Stochastic Optimal Control Problems of Residential Heating Systems With a Geothermal Energy Storage

Von der Fakultät 1- MINT – Mathematik, Informatik, Physik,
Elektro- und Informationstechnik
der Brandenburgischen Technischen Universität Cottbus-Senftenberg
genehmigte Dissertation
zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von

Paul Honore Takam
geboren am 27.05.1992 in Bansoa, Kamerun

Vorsitzender:	Prof. Dr. Gennadiy Averkov
Gutachter:	Prof. Dr. Ralf Wunderlich
Gutachter:	Prof. Dr. Olivier Menoukeu Pamen
Gutachter:	Prof. Dr. Martin Redman

Tag der mündlichen Prüfung: 4. Juli 2023

https://doi.org/10.26127/BTUOpen-6450

Abstract

In this thesis we consider a residential heating system equipped with several heat production and consumption units and investigate the stochastic optimal control problem for its cost-optimal management. As a special feature the manager has access to a geothermal storage which allows for inter-temporal transfer of heat energy by storing leftover solar thermal energy generated in summer for satisfying demand later. It is charged and discharged via heat exchanger pipes filled with a moving fluid. Further, the manager of that system faces uncertainties about the future fuel price and heat demand.

The main goal is to minimize the expected aggregated cost for generating heat and running the system. This leads to a challenging mathematical optimization problem. The optimization problem is formulated first as a non-standard continuous-time stochastic optimal control problem for a controlled state process whose dynamics is described by a system of ordinary differential equations (ODEs), stochastic differential equations (SDEs) and a partial differential equation (PDE).

In order to convert the problem into an optimization problem for a controlled diffusion process in standard form, the PDE (heat equation), which describes the temperature distribution in the geothermal energy storage, is first converted into a high-dimensional system of ODEs by semi-discretizing the space variables and its stability is investigated. The subsequent time-discretization and the associated stability analysis for the finite difference scheme used makes it possible to numerically simulate the spatio-temporal temperature distribution in the storage. This also makes it possible to compute some aggregated characteristics which are useful for the operation of the geothermal energy storage embedded in the residential heating system. They describe the input-output behavior of the geothermal storage and the associated energy flows as well as its response to charging and discharging processes.

Since knowledge of these aggregated characteristics is sufficient for the cost-optimal management of the heating system, model order reduction (MOR) method is used in a second step to the solution of the associated optimal control problem. This replaces the high-dimensional system of ODEs obtained by the semi-discretization of the heat equation by a suitable low-dimensional system that approximates the input-output behaviour of the dynamics of the geothermal energy storage sufficiently accurate. First, the linear time-varying system of ODEs is approximated by a suitable linear time-invariant system. This allows the Lyapunov balanced truncation MOR method to be applied.

Finally, we investigate the solution of the resulting standard optimal control problem for a controlled multi-dimensional diffusion process using dynamic programming methods and derive the corresponding Hamilton-Jacobi-Bellman (HJB) equation. However, no analytical solution of the HJB equation can be expected for the control problem under investigation. Therefore, we transform the continuous-time optimal control problem into a discrete-time control problem for a controlled Markov chain with finitely many states by discretizing both the time and the states. After determining the transition probabilities, the problem is solved using methods from the theory of Markovian decision processes.

The thesis presents results of extensive numerical experiments carried out with the developed methods. Results of a first group of experiments show the dependence of the input-output behaviour of the geothermal energy storage on the topology and arrangement of the heat exchangers pipes and on the timing sequence of the charging and discharging processes. A second group of experiments illustrates the efficiency of the applied MOR methods. Finally, numerical results are presented which reveal typical properties of the value function and the optimal

strategy of the optimization problem.

We end this thesis by describing some alternative methods to overcome the curse of dimensionality. These methods include Least Square Monte Carlo and Approximate Dynamic Programming.

Kurzfassung

In dieser Arbeit betrachten wir ein mit mehreren Wärmeerzeugungs- und -verbrauchseinheiten ausgestattetes Gebäudeheizungssystem und untersuchen ein stochastisches Optimalsteuerungsproblem für dessen kostenoptimale Bewirtschaftung. Als Besonderheit steht dem Manager des Heizungssystems ein Erdwärmespeicher zur Verfügung, der eine zeitliche Übertragung von Wärmeenergie ermöglicht, indem die Überproduktion von Solarwärme z.B. im Sommer für die spätere Deckung des Wärmebedarfs gespeichert wird. Die Be- und Entladung erfolgt über Wärmetauscherrohre, in denen eine Flüssigkeit zirkuliert. Der Manager dieses Systems muss seine Entscheidungen unter Unsicherheiten über die Brennstoffpreise und den zu deckenden Wärmebedarf in der Zukunft treffen.

Das Hauptziel ist es, die zu erwartenden Gesamtkosten für die Wärmeerzeugung und den Betrieb des Systems zu minimieren. Dies führt zu einem anspruchsvollen mathematischen Optimierungsproblem. Das Optimierungsproblem wird zunächst als ein zeitstetiges stochastisches Optimalsteuerungsproblem in Nichtstandardform formuliert, bei dem die Dynamik des gesteuerten Zustandsprozesses durch ein System gewöhnlicher Differentialgleichungen (ODEs), stochastischer Differentialgleichungen (SDEs) und einer partiellen Differentialgleichung (PDE) gegeben ist.

Um das Problem in ein Optimierungsproblem für einen gesteuerten Diffusionsprozess in Standardform zu überführen, wird zunächst die PDE (Wärmeleitgleichung), welche die Temperaturverteilung im Erdwärmespeicher beschreibt, durch Semidiskretisierung der Raumvariablen in ein hochdimensionales System von ODEs überführt und es wird dessen Stabilität gezeigt. Die anschließende Zeitdiskretisierung und die zugehörige Stabilitätsanalyse für das verwendete Differenzenverfahrens ermöglicht die numerische Simulation der räumlichen und zeitlichen Temperaturverteilung im Speicher. Damit gelingt es auch, wichtige Kenngrößen für den Betrieb des Erdwärmespeichers innerhalb eines Heizungssystems zu berechnen. Diese beschreiben das Input-Output-Verhalten des Speichers und die damit verbundenen Energieflüsse sowie dessen Antwort auf Lade- und Entladevorgänge.

Da die Kenntnis dieser Kenngrößen ausreichend für die kostenoptimale Bewirtschaftung des Heizungssystems ist, werden zur Lösung des zugehörigen Optimalsteuerungsproblem in einem zweiten Schritt Modellreduktionsverfahren eingesetzt. Diese ersetzen das durch die Semidiskretiseierung der Wärmeleitgleichung entstandene hochdimensionale System von ODEs durch ein geeignetes niedrigdimensionales System, welches das Input-Output-Verhalten des Erdwärmespeichers hinreichend genau approximiert. Dabei wird zunächst das lineare zeitvariable System von ODEs durch ein geeignetes lineares zeitinvariantes System approximiert. Dies erlaubt es dann die Lyapunov-Balanced-Truncation-Methode für die Modellreduktion einzusetzen.

Schließlich untersuchen wir die Lösung des resultierenden Standard-Optimalsteuerungsproblems für einen gesteuerten mehrdimensionalen Diffusionsprozess unter Verwendung von Methoden des Dynamic Programming und leiten die zugehörige Hamilton-Jacobi-Bellman-Gleichung (HJB) her. Für das untersuchte Steuerungsproblem kann jedoch keine analytische Lösung der HJB-Gleichung erwartet werden. Daher überführen wir das zeitstetige Optimalsteuerungsproblem durch Diskretisierung sowohl der Zeit als auch der Zustände in ein zeitdiskretes Kontrollproblem für eine gesteuerte Markovkette mit endlich vielen Zuständen. Nach der Bestimmung der Übergangswahrscheinlichkeiten wird das Problem mit Methoden der Theorie der Markovschen Entscheidungsprozesse gelöst.

Die Arbeit präsentiert Ergebnisse umfangreicher numerischer Experimente, die mit den

entwickelten Methoden durchgeführt wurden. Ergebnisse einer ersten Gruppe von Experimenten zeigen die Abhängigkeit des Input-Output-Verhaltens des Erdwärmespeichers von der Topologie und Anordnung der Wärmetauscher und von der zeitlichen Abfolge der Lade- und Entladevorgänge. Eine zweite Gruppe von Experimenten illustriert die Effizienz der angewandten Modellreduktionsmethoden. Schließlich werden numerische Ergebnisse präsentiert, welche typische Eigenschaften der Wertfunktion und der optimalen Steuerung des Optimalsteuerungsproblems sichtbar machen.

Die Arbeit schließt mit der Beschreibung einiger alternativer Methoden zur Lösung der Kontrollprobleme, mit denen der „Fluch der Dimension“ überwunden werden kann. Zu diesen Methoden gehören die Least-Squares-Monte-Carlo-Methode und Approximative Dynamic Programming.

Acknowledgements

I would like to take this opportunity to thank all those who motivated me, supported me, and contributed to the success of my doctoral thesis.

A special thanks goes to Professor Ralf Wunderlich for his guidance and his support since my Master's degree and throughout the research work on my doctoral thesis. First of all, without him I would never have come to BTU Cottbus-Senftenberg to do research in a passionate and prominent domain of mathematics. He has shown a great patience and understanding of the difficulties I encountered during the work period. I would like to thank him for supervising me in the course of my research work on my dissertation project and for his extensive suggestions for improvement not only in my research but also in private conversations. Further, his suggestions are always very constructive and his response time to emails very fast, even when I write very late he replies as quickly as possible.

I would like to express my gratitude to Professor Oliver Menoukeu Pamen (University of Liverpool, AIMS-Ghana) for his support, his contribution to enriching discussions and his suggestions to improve the work during my research.

A special thanks goes to Prof. Carsten Hartmann (BTU Cottbus–Senftenberg) for his suggestions for improvement of the work during the weekly stochastic research seminars, where I had the opportunity to present my work progress several times. I am grateful to all participants of the stochastic research seminar for the many fruitful discussions and suggestions.

I would like to thank Thomas Apel (Universität der Bundeswehr München), Martin Bähr (DLR), Michael Breuss Gerd Wachsmuth (BTU Cottbus–Senftenberg), Andreas Witzig (ZHAW Winterhur), Karsten Hartig (Energie-Concept Chemnitz), Dietmar Deunert (eZeit Ingenieure Berlin), Martin Redmann (Martin-Luther-Universität Halle-Wittenberg) for valuable discussions that improved this doctoral thesis. Furthermore, I would also like to thank all professors and colleagues of the institute of Mathematics who provided a pleasant atmosphere and good cooperation during this research period at the BTU Cottbus-Senftenberg. I gratefully acknowledge the support by the German Academic Exchange Service (DAAD) within the project No. 57417894.

I would like to thank Prof. Tadmon Calvin (University of Dschang, Cameroon), Prof. Takam Soh Patrice (University of Yaounde I, Cameroon) and Prof. Mama Foupouagnigni (AIMS-Cameroon) for their support and advice which inspired me and helped me to overcome all the difficulties encountered during this research period.

I am particularly grateful to my parents who instilled in me great human and moral values such as perseverance, honesty, integrity and love of a job well done. All this allowed me to overcome all the difficulties encountered during the research period of my doctoral thesis. A special thank you to my darling, my lovely wife Wati Bolang Doriane, my guardian angel who permanently watches over the sentimental and emotional logistics, my health, necessary for the successful completion of the thesis. I am deeply grateful to her for her incredibly support and understanding during this difficult period of our life. I am deeply grateful to my family and friends for their incredibly support and understanding in the preparation of this doctoral thesis.

Contents

CHAPTER 1

Introduction

1.1 Motivation

Climate change and energy dependency require urgent measures for the improvement of energy efficiency in all areas. District heating and cooling systems play an important role for increasing energy efficiency in buildings and for including renewable energy sources. Besides of numerous technical issues also economic issues such as the cost-optimal control and management of such heating systems play a central role. The latter leads to challenging mathematical optimization problems which require advanced and sophisticated solution techniques. One of these problems arising in the optimal management of a residential heating system with access to an additional geothermal energy storage will be addressed in this thesis.
We consider a residential heating system equipped with a local renewable production unit such as a solar thermal collector, a non-renewable production unit using fossil fuels such as oil,gas,coal or power grid, an internal storage (IS) which is a water tank with small capacity, a geothermal storage (GS) with large capacity, and several consumption units in the house. The local thermal energy produced by the solar thermal collector is used to satisfy the demand of the building and the unsatisfied demand must always be satisfied by taking heat from the IS. In case of overproduction the left over thermal energy is transferred into the IS. A heating system consisting of the IS and the solar collector is not sufficient because the IS is a small water tank and cannot store heat for several weeks or months and in winter the demand of thermal energy is high and the production of the solar collector is low. Hence, one needs to add another heat production unit into the model. The natural choice will be to fire fuel or use electricity to generate heat in times of high demand. This option may not be efficient because it does not favor a manager who wants to minimize the expected aggregated costs for producing heat and running the system. Therefore, this motivates us to include the GS into our model. As a special feature, the GS has large capacity which allows for inter-temporal transfer of heat energy by storing leftover solar thermal energy generated in summer for satisfying demand in autumn and winter. Further, such thermal storages may help to mitigate peaks in the electricity grid by converting electrical energy into heat energy.

The manager wants to choose at any time the decision called policy, that will optimize the performance of the system over some finite time horizon. The decision consists of charging the

IS by discharging the GS or discharging the IS to charge the GS, charging the IS by firing fuel, and do nothing, i.e. waiting. In our model we assume that we can always satisfy the demand.

On the one hand, the production of renewable energies and the heat consumption is affected by the uncertainties resulting from weather and environmental conditions. On the other hand, time fluctuations of the market prices for fuel or electricity cannot be predicted or can only be predicted imperfectly. Therefore, the controller of the heating system has to make decisions always under uncertainties about the future dynamics of several factors such as the electricity or fuel prices, and the uncertainty about the future residual demand resulting from the superposition of the demand for thermal energy in the building and supply of thermal energy of solar collector.

We have to impose some constraints on the state, meaning that the average temperature in the IS and GS must always be within certain comfort zone. This leads to a challenging mathematical optimization problem. That optimization problem is treated first as a continuous-time stochastic optimal control problem for a multidimensional controlled state process whose dynamics is described by a system of random ordinary differential equations (ODEs), stochastic differential equations (SDEs) and partial differential equation (PDE). Second, the control problem is transformed into a Markov decision process (MDP) for a controlled finite state Markov chain characterized by the associated transition probabilities.

Main objectives. The main goals of this thesis are to:

1) investigate the mathematical modeling and the analysis of the GS and to perform numerical simulations of its short-term behavior. These simulations support the design and the choice of the topological structure of the heat exchanger pipes (PHXs) in the GS,

2) embed the GS into a residential heating system and formulate the stochastic optimal control problem for the cost-optimal management of such heating systems,

3) apply suitable model reduction techniques to reduce the dimension of the state of the GS. This dimension reduction helps to reduce the complexity of the associated optimal control problem and to make its numerical solution tractable and more efficient,

4) solve numerically the stochastic optimal control problem to determine the optimal charging and discharging decisions of the controller of the heating system over a finite decision making horizon.

Now we give more detailed explanation of our main goals.

Geothermal energy storage. In this thesis we are going to pay special attention to the thermal storage facilities which help to mitigate and to manage temporal fluctuations of heat supply and demand for heating and cooling systems of single buildings as well as for district heating systems. They allow heat to be stored in form of thermal energy and be used hours, days, weeks or months later. This is attractive for space heating, domestic or process hot water production, or generating electricity. Note that thermal energy may also be stored in the way of cold.

The first goal of this thesis is the modeling and the analysis of the GSs as depicted in Fig. 1.1. Such storages gain more and more importance and are quite attractive for residential heating systems since construction and maintenance are relatively inexpensive. Furthermore, they can be integrated both in new buildings and in renovations. We consider a 2D-model of a geothermal thermal energy storage, see Fig. 1.2, where a defined volume under or aside of a

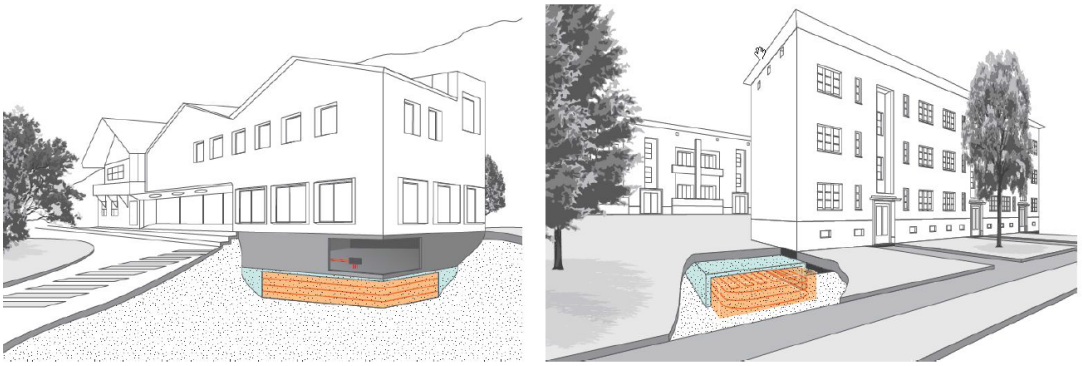


Figure 1.1: Geothermal storage: in the new building, under a building (left) and in the renovation, aside of the building (right), see `www.ezeit-ingenieure.eu`, `www.geo-ec.de`.

building is filled with soil and insulated to the surrounding ground. Thermal energy is stored by raising the temperature of the soil inside the storage. It is charged and discharged via pipe heat exchangers (PHX) filled with some fluid (e.g. water). These PHXs can be connected to a short-term storage such as a water tank or directly to a solar collector and (heat) pumps move the fluid carrying the thermal energy. A special feature of the storage in this work is that it is not insulated at the bottom such that thermal energy can also flow into deeper layers as it can be seen in Fig. 1.2. This can be considered as a natural extension of the storage capacity since this heat can to some extent be retrieved if the storage is sufficiently discharged (cooled) and a heat flux back to storage is induced. Of course, there are unavoidable diffusive losses to the environment but due to the open architecture, the GS can benefit from higher temperatures in deeper layers of the ground and serve as a production unit similar to a downhole heat exchanger. Note that in many regions in Europe the temperature in a depth of only 10 meter is almost constant around 10 $^{\circ}C$ over the year.

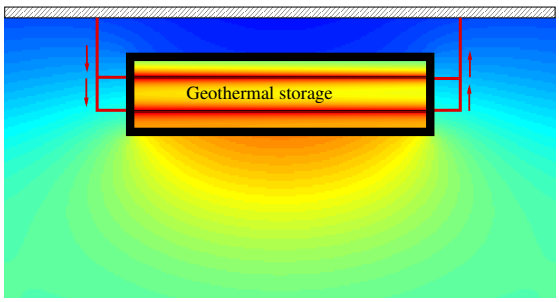


Figure 1.2: 2D-model of a GS insulated at the top and the sides while open at the bottom and spatial temperature distribution.

Geothermal energy storages enable an extremely efficient operation of heating and cooling systems in buildings. Further, they can be used to mitigate peaks in the electricity grid by converting electrical energy into heat energy (power to heat). Pooling several GSs within the framework of a virtual power plant gives the necessary capacity which allows to participate in the balancing energy market. This type of GS was first studied by Bähr et al. [9, 8]. In this thesis, we extends and complements the results in Bähr et al. [9, 8] where the authors focus on the numerical simulation of the long-term behavior over weeks and months of the spatial temperature distribution in a GS and the interaction between a GS and its surrounding domain. For

simplicity charging and discharging was described by a simple source term but not by PHXs. Here, we focus on the computation of the short-term behavior of the spatial temperature distribution. This is needed for storages embedded into residential heating systems and the study of the storage's response to charging and discharging operations on time scales from a few minutes to a few days. We extend the setting in [8, 9] and include PHXs for a more realistic model of the storage's charging and discharging process. However, for the sake of simplicity we do not consider the surrounding medium but reduce the computational domain to the storage depicted in Fig. 2.3 by a solid black rectangle. Instead, we set appropriate boundary conditions to mimic the interaction between storage and environment. For the management and control of a storage which is embedded into a residential system one needs to know the amount of available thermal energy that can be stored in or extracted from the storage in a given short period of time. Such questions can only be answered if one knows the spatial temperature distribution, in particular around the PHXs. Charging and discharging is not efficient or even impossible if there are only small differences between the temperatures inside and in the vicinity of the PHXs. Long periods of (dis)charging may lead to saturation in the vicinity of the PHXs. As a consequence charging or discharging is no longer efficient and should be stopped since propagation of heat to regions away from the PHXs takes time.

Aggregated characteristics. For the operation of a GS within a residential system, the controller or manager of that system needs to know certain aggregated characteristics of the spatio-temporal temperature distribution in the storage, their dynamics and response to charging and discharging decisions. Note that the latter means to decide whether the fluid is pumped through the PHXs to the storage or whether it is at rest and pumps are off. Further, if pumps are on, one has to decide on an appropriate temperature of the fluid. An example of such an aggregated characteristic is the average temperature in the storage medium from which one can derive the amount of available thermal energy that can be stored in or extracted from the storage. Another example is the average temperature at the outlet which allows to determine the amount of energy injected to or withdrawn from the storage. Further, the average temperature at the bottom of the storage allows to quantify the heat transfer to and from the ground via the open bottom boundary. The above aggregated characteristics can be computed by post-processing the spatio-temporal temperature distribution in the storage. On the other hand, to derive the charging and discharging decisions, the manager of the heating system does not need to know the complete spatio-temporal temperature distribution. It is enough to know a rough approximation of the dynamics described by a few aggregate quantities such as the average temperature in the storage and the temperature at the outlet of the heat exchanger pipes. The latter allows by comparison with the (controllable) inlet temperature to determine the amount of heat energy that can be charged or discharged in next period of time. Further, during charging that temperature difference indicates whether the storage is saturated and charging is not no longer efficient while during discharging the manager can learn if the storage is still ready to deliver energy.

Stochastic optimal control. The second focus of this thesis is the incorporation of a GS into a residential heating system and its implication to the optimal control problem of the cost-optimal management of the system. From a mathematical point of view this leads to an interesting and challenging stochastic optimization problem for a controlled stochastic process whose dynamics is governed by SDEs, ODEs and a parabolic PDE (heat equation with convection and appropriate boundary and interface conditions). The fact that the dynamics of the GS is described by a PDE is a non-standard feature of the optimal control problem and does not fit to the stan-

dard framework for stochastic optimal control problems where the state is a multi-dimensional stochastic process described by a system of SDEs (and ODEs). Further, this non-standard feature makes the optimal control problem much more difficult and challenging. However, we replace the PDE describing the dynamics of the temperature in the GS by a system of ODEs resulting from the semi-discretization w.r.t. spatial variables. Following this approach the complete space-time dynamics of the temperature in the GS is available with any given precision. Therefore, the state of the optimal control problem becomes a high-dimensional vector. Even though the optimal control problem is now in *standard form*, we do not expect to have an explicit or analytical solution. Also numerically solving will not be tractable and efficient. Since, we are facing the so-called curse of dimensionality. This leads us to the problem of MOR which we pay a special attention in this thesis.

Model order reduction. The aim of the MOR is to find an approximation of the dynamics of the aggregated characteristics describing the storage response to the manager's charging and discharging decisions by an appropriate low-dimensional system of ODEs. This will allow for a dimension reduction of the state process of the optimal control problem. The time-dependent velocity considered in the dynamics of the GS makes the MOR more demanding and more challenging. Therefore, we restrict to the case of a piece-wise constant velocity of the fluid in the PHXs. This is often observed in real-world systems which operate with constant velocity during charging and discharging if pumps are on, while the velocity is zero if pumps are off. Then the high-dimensional system of ODEs constitutes a system of linear non-autonomous ODEs since the system and the input matrices depend on time through the fluid velocity. The latter varies over time and is only piece-wise constant. Thus, the obtained linear system is not linear time-invariant (LTI). The latter is a crucial assumption for many of MOR methods. We circumvent this problem by approximating the model for the GS by a so-called *analogous model* which is LTI. The key idea for the construction of such an analogue is to mimic the original model by a LTI system, where pumps are always on such that the fluid velocity is constant all the time. During the waiting periods we use at the inlet and outlet boundary the same type of boundary conditions as during charging and discharging. However, we choose the inlet temperature to be equal to the average temperature in the PHX. Numerical examples presented in Chapter 3 show that the analogous system approximates the original system quite well. To the derived linear LTI system, we apply the Lyapunov balanced truncation model order reduction method which is well suited for our purposes.

Solving the stochastic optimal control problem using MDP theory. Finally, the resulting continuous-time stochastic optimal control problem obtained by replacing the dynamics of the GS by the reduced-order system is now in a standard form and can be solved using well-known tools in stochastic optimal control. We first explore the solution of the continuous-time stochastic optimal control problem using dynamic programming methods. Applying dynamic programming to the continuous-time stochastic optimal control problem enables us to derive the associated Hamilton-Jacobi-Bellman (HJB) equation. The latter is a nonlinear degenerated PDE for which an analytical solution is not expected and due the curse of dimensionality, a numerical solution cannot be expected. Therefore, we discretized the optimal control problem in time to obtain a continuous-state Markov decision processes (MDPs) with finite time horizon and finite action space. The optimal policies are found using discrete-time dynamic programming.
MDPs are a class of stochastic sequential decision processes in which the set of feasible actions, the rewards and the transition kernel depend only on the current state and action and not on past

states and actions. In the MDP setting a decision rule specifies the action to be chosen at a particular time. In the problem considered in this thesis the decision rule depends not only on time but also on the current state. A policy or strategy is a sequence of decision rules which provides the storage manager a prescription for choosing his actions in any possible future state.

The discrete-time settings require the manager of the heating system to choose his actions taking into account future consequences because the action chosen at present time affects the future evolution of the system. Another requirement for the discrete-time setting is that the decisions can only be made at every fix time point, i.e, once an action is taken at current time, the manager cannot change it before the next time point in contrast to the continuous-time model where the action can be changed at any time. Further, the dimension of the system must be such that the computation of the optimal policies is feasible with dynamics programming. Therefore, special attention is paid to the state and control constraint sets.

Numerical solution of the MDP. For small size systems, the dynamic programming appears to be tractable approach to find the optimal policies. This fits to our problem when we consider a low-dimensional reduced-order system of the GS. The application of the dynamic programming leads the so called dynamic programming equation, from which the optimal policies is computed by solving the associated pointwise optimization problem. In a stochastic setting finding the optimal policies are challenging due to the complicated nature of the computation of the conditional expectation appearing in the dynamic programming equation. Further, in many practical problems no closed-form expressions of this expectation are available and when the dimension of the system is high it becomes computationally non-tractable. One focus of this thesis is to find a method of computing this conditional expectation in a very fast and efficient manner.

Since the actions are state-dependent and the state constraints require the average temperature in the IS and GS to be in some intervals called comfort zones. First, we transform the continuous state MDP into a MDP for a controlled finite-state Markov chain and express the conditional expectation in terms of the transition probabilities of the states of the Markov chain. Second, we transform the reduced order system into a basis where average temperature is up to a scaling constant the last reduced order state of the GS. This transformation helps in state discretization since we can allow to use a fine grid only for the last coordinate and use coarse grid for other coordinates of the reduced order state.

Finally, we construct the transition probabilities such that the numerical computations will be efficient and very fast. However, when the dimension of the state space is very high, solving the control problem using the MDP approach may not be efficient. In this case, we have to resort to an approximate solution using some numerical methods such as least-square Monte Carlo and approximate dynamic programming. The focus of the last chapter is to describe some of these methods necessary to overcome the cure of dimensionality.

1.2 Literature Review

Alternative energy systems. District heating and cooling systems comprise a network of pipes connecting the buildings in a neighborhood, town district or whole city, so that they can be served from centralized plants or a number of distributed heat producing units. This approach allows any available source of heat to be used. The inclusion of district heating in future sustainable cities allows for the wide use of combined heat and power (CHP) together with the

utilization of heat from waste-to-energy and various industrial surplus heat sources as well as the inclusion of geothermal and solar thermal heat.

The development of future district heating systems and technologies involves energy efficiency and conservation measures as an important part of the technology, see Chen et al. [29]. The design and perspective of low-energy buildings have been intensively analyzed and described in Tommerup et al. [114, 115], including concepts like energy efficient buildings, see Heiselberg [58], zero emission buildings, and plus energy houses, see Abel [1], Neilsen and Möller [79]. However, such papers mostly deal with future buildings and not with the existing building stock which, due to the long lifetime of buildings, is expected to constitute the major part of the heat demand for many decades to come. Furthermore, the "near zero energy buildings" requirement cannot be achieved by high energy performance buildings alone. The new buildings have to be fitted with new technologies to produce energy from local renewable energy sources. These buildings will be both heat consumers and heat producers at the same time. According to Saeb-Gilani et al. [94], if in some periods more heat is produced than it is consumed, then energy efficiency requires to transfer the excess heat to storages or to other buildings, reducing the share of fossil fuels in the energy mix of the network.

Renewable energy sources such as solar thermal collectors and geothermal energy as well as heat pumps can be integrated in the district heating systems. This will lead to multiple distributed energy sources of different temperature levels being present in the network at the same time. Also, instead of having energy sources on the one hand, and heat consumers, on the other hand, some of the consumers will feed energy into the network at the same time.

Thermal energy storages. Thermal energy storages can significantly increase both the flexibility and the performance of district energy systems and enhancing the integration of intermittent renewable energy sources into thermal networks (see Guelpa and Verda [50], Kitapbayev et al. [65]). Since heat production is still mainly based on burning fossil fuels (gas, oil, coal) these are important contributions for the reduction of carbon emissions and an increasing energy independence of societies. Thermal energy storages have attracted the interest of several authors over the last decades. In Zalba et al. [128] a review has been carried out for the history of thermal energy storages with solid–liquid phase change and focused in three aspects: materials, heat transfer and applications. An overview of the European ~~and in particular the Spanish~~ thermal energy storage potential is presented in Arce et al. [7]. The authors show that thermal energy storages make an important contribution to the reduction of CO_2-emissions. In Soltani et al. [103] the authors provide a comprehensive review on the evolution of geothermal energy production from its beginnings to the present time by reporting production data from individual countries and collective data of worldwide production.

The efficient operation of thermal storages requires a thorough design and planning because of the considerable investment cost. For that purpose mathematical models and numerical simulations are widely used. We refer to Dahash et al. [32] and the references therein. In that paper the authors investigate large-scale seasonal thermal energy storages allowing for buffering intermittent renewable heat production in district heating systems. Numerical simulations are based on a multi-physics model of the thermal energy storage which was calibrated to measured data for a pit thermal energy storage in Dronninglund (Denmark). Another contribution is Major et al. [73] which considers heat storage capabilities of deep sedimentary reservoirs. The governing heat and flow equations are solved using finite element methods. Further, Regnier et al. [92] study the numerical simulation of aquifer thermal energy storages and focus on dynamic mesh optimisation for the finite element solution of the heat and flow equations. For an overview on

thermal energy storages we refer to Dincer and Rosen [38] and for further contributions on the numerical simulation of such storages to [13, 38, 56, 70, 103, 127].

Lyapunov balanced truncation MOR. The balanced truncation method was first introduced by Mullis and Roberts [78] and later in the linear systems and control literature by Moore [77]. The idea of this method is first to transform the system into an appropriate coordinate system for the state-space in which the states that are difficult to reach, i.e., require a large input energy to be reached. They are simultaneously difficult to observe, i.e., produce a small observation output energy. Then, the reduced model is obtained by truncating the states which are simultaneously difficult to reach and to observe. Among the various model order reduction methods balanced truncation is characterized by the preservation of several system properties like stability and passivity, see Pernebo and Silverman [84]. Further, it provides error bounds that permit an appropriate choice of the dimension of the reduced-order model depending on the desired accuracy of the approximation, see Enns [41].

Besides the Lyapunov balancing method, there exist other types of balancing techniques such as stochastic balancing, bounded real balancing, positive real balancing and frequency weighted balancing, see Gugercin and Antoulas [51]. The Lyapunov balanced truncation model reduction method applied in this paper was first introduced by Mullis and Roberts [78] and later in the systems and control literature by Moore [77]. The idea of the balanced truncation model reduction is first to transform the system into an appropriate coordinate system for the state-space in which the states that are difficult to reach, that is, require a large input energy to be reached are simultaneously difficult to observe, i.e., produce a small observation output energy. This is achieved by simultaneously diagonalizing the controllability and the observability Gramians, which are solutions to the controllability and the observability Lyapunov equations. Then, the reduced model is obtained by truncating the states which are simultaneously difficult to reach and to observe [77]. In the book of Benner et al. [22], an efficient implementation of MOR methods such as modal truncation, balanced truncation, and other balancing-related truncation techniques is presented. In this book, the authors discussed various aspects of balancing-related techniques for large-scale systems, structured systems, and descriptor systems. The results presented in [22] also cover the MOR techniques for time-varying as well as the model reduction for second- and higher-order systems, which can be considered as one of the major research directions in dimension reduction for linear systems. In addition, surveys on system approximation and MOR can be found in [3, 5, 21, 26, 46, 51, 62, 67, 76, 91, 123] and the references therein.

Optimal control of energy storages. Energy storage is one of the key underpinnings of the vision of the smart grid which aims to support sustainable energy provisioning across the world. As shown in Urgaonkar et al.[117], and Vytelingum et al.[124], incorporating storage technology into the electricity grid design can significantly improve energy management and result in huge cost saving in electricity delivery. In fact, the role of storage technology is even more pronounced when a portion of injected electricity to the grid is obtained from renewable source (e.g., wind power or solar energy). This is due to uncertain and hardly predictable fluctuations of renewable energy production due to weather conditions. Energy storage units allow to save the current excess energy and use it whenever there is energy shortage in the grid, they increase flexibility in the energy management. Likewise, consumers seeking reduced electricity costs by shifting electricity purchases away from times of peak tariffs, together with a desire for increased energy self-sufficiency, are beginning to consider an energy storage as a viable option.

With economically feasible residential storage on the horizon, in recent years researchers have moved from the analysis of relatively rudimentary and largely uncoordinated battery energy storage systems to systems of increasing scale and sophistication as well, see Huq et al. [63], and Levron et al. [69].

It is well-known that energy storages can be used to create profit by trading in the energy market and taking advantage of the fluctuating energy price by applying an active storage management, see Bäuerle and Riess [12], Chen and Forsyth [30, 31], Ware [125], Shardin and Wunderlich[99]. The basic principle is minimize to the intermediate storage and operating costs, and unavoidable dissipative losses. Note that the energy prices as well as the residual demand typically do not only vary in time, but are also unpredictable. Therefore, the control decisions must be made in the face of uncertainty about future energy prices and the future residual demand. In district energy systems powered by CHP plants, thermal storage can significantly increase CHP flexibility to respond to real time market signals and therefore improve the business case of such demand response schemes in a smart grid environment, see Kitapbayev [65]. However, one of the main challenges is to determine the optimal control of the inter-temporal storage operation in the presence of uncertain market prices. In view of the current development of the CHP including exchange of energy between consumers and producers also the control of smart thermal grids has to be investigated. Here, households or consumers in the grid are equipped with a solar collector and thermal storage facilities.

Dynamic programming techniques. The dynamic programming can be considered as collection of mathematical tools used to analyze sequential decision processes. The concept of dynamic programming was first introduced by Bellman in the early 1950s and was extensively developed and applied in his numerous papers [15, 16, 18, 19] and in his book [17]. These methods are applied by several authors in different domains of sciences to approach various problems where inter-temporal relationships are important or the problems where the impact of the current decisions on future decisions are considered. Here we pay attention to finance and in particular to energy management sector which fit to the problem that we consider in this thesis. Edwin et al. [40] used these tools in finance to solve the problem of bond refinancing decision. There, the use of the dynamic programming methods was motivated by the fact the bonds can be refunded over time which create a recursive relationship. In Henning and Gregor [112], the dynamic programming methods are applied to determine the optimal management strategy of a smart home, which is equipped with a fuel cell, photovoltaic system, an electric car, a battery and a storage unit for thermal energy. In this paper the fuel cell is used to co-generate heat and power which are stored in the thermal energy storage unit and battery, respectively. Further, many other recent applications of the dynamic programming methods in the computation of the optimal management decision of the energy system are given in [42, 59, 72, 104, 126, 130], whereas for other applications in finance we refer to [28, 105] and the references therein.

MDP methods. Several authors used the MDP approach to solve problems in different domains of science. For example Puterman [90] applied the MDP approach to solve many problems in different domains, among which maintenance and engine replacement problems, and inventory management problem. Another important contribution in this domain is the book by Bäuerle and Rieder [11]. In this book the authors described the MDP methods for finite, random, and infinite time horizon and applied them to solve many problems in finance and insurance. For further contributions and applications of MDP, see [4, 35, 43, 89] and references therein. However, in many practical problems, the dimension of the state space is very high and solving the

control problem using the MDP approach may not be efficient. In this case, we have to resort to an approximate solution using some numerical methods such as Least-Squares Monte Carlo [10, 52, 71, 81] and Approximate Dynamic Programming [2, 88, 101] and references therein.

1.3 Main Contributions

As mentioned above, the main goal is to find a tractable mathematical model of residential heating systems, in particular, the dynamics of the geothermal energy storage and to determine the manager's optimal charging and discharging decisions that minimize the cost for generating heat and running the system equipped with several storages and heat production units. The heat can be generated by the local renewable heat source such as solar thermal collector or by firing fuel. We also mentioned above that the manager of this system is exposed to uncertainties about future fuel price and future residual demand. The thesis provides the following contributions.

- We build up the mathematical model of the GS and we incorporate it into the stochastic optimal control problem of the cost-optimal management of a residential heating system. To the best of our knowledge a GS has only been considered in Bähr et al [8], where the authors investigated the long-term simulation of the heat equation. Further, the optimal control problems considered in this thesis have not been considered yet in the literature. The optimization problem is not in standard form because one variable of the state process is described by a PDE.

- We introduce the concept of analogous system. This helps to transform the linear time-varying system into a linear time-invariant system, suitable for balanced truncation model order reduction that we apply to reduce the dimension of the state of the GS.

- We solve that PDE describing the dynamics of the GS using finite difference schemes. In a first step we study the semi-discretization with respect to spatial variables leading to a system of linear ODEs. In a second step, we consider full space-time discretization and derive implicit finite-difference schemes.

- We prove that the chosen semi-discretization ensures a system of linear ODEs with a stable system matrix.

- We provide a detailed stability analysis for the implicit finite-difference schemes of the fully discretized PDE and establish a stability condition.

- We perform extensive numerical experiments for the GS. In a first group of experiments, to study its short-term behaviour, where simulations results for the temporal behavior of the spatial temperature distribution are used to determine how much energy can be stored in or taken from the storage within a given period of time. Special focus is laid on the dependence of these quantities on the arrangement of the PHXs within the storage.

- In the second group of experiments, we apply model reduction techniques known from control theory such as balanced truncation to derive low-dimensional approximations of aggregated characteristics of the temporal behavior of the spatial temperature distribution. There, extensive numerical experiments are carried out which show the efficiency and the accuracy of the method. Simulations show that only a few suitable chosen ODEs

are sufficient to produce good approximations of the input-output behaviour of the storage. The latter is crucial if the GS is embedded into a residential heating system and the cost-optimal management of such systems is studied mathematically in terms of optimal control problems.

- Non-standard optimal control problem is transformed first into a continuous-time optimal control problem for a controlled degenerated diffusion process. Second, we apply time discretization to transform the resulting standard continuous-time control problem into a continuous-state MDP with no discretization error. Finally, we transform the continuous-state MDP into a MDP for a controlled finite-state Markov chain, where we derive associated transition probabilities.
- We construct the transition probabilities of the controlled Markov chain in such away that the numerical computations of the conditional expectation will be efficient and very fast.
- We solve the MDP for a low-dimensional state numerically and compute the optimal policies and optimal value function. By means of numerical experiments we investigate the properties of the value function and the optimal control. This will be helpful for future solution approaches for MDPs with a higher dimensional state using approximate solution techniques which are introduced in Chapter 7.

1.4 Outline

We conclude this chapter with an overview of the different chapters of this dissertation. The thesis consists of 8 chapters and an appendix. The first chapter is devoted to the introduction.

In Chapter 2, a mathematical model for the residential heating system is developed and based on it a stochastic optimal control problem for the cost-optimal management of this heating system is formulated.

Chapter 3 describes the semi-discretization with respect to spatial variables of the initial boundary value problem describing the dynamics of the GS. For the resulting system of linear ODEs we show that the system matrix is stable and the full space-time discretization is studied where we derive implicit finite-difference schemes and provide the associated stability analysis. We also explain numerical approximation of the aggregated characteristics and derive an LTI analogous model of the GS that mimics the most important features of the original non-LTI model of the GS. We end the chapter with results of extensive numerical experiments where we determine how much energy can be stored in or taken from the storage within a given short period of time.

The formulation of the general model reduction problem is discussed in Chapter 4, where we present the Lyapunov balanced truncation method which is based on the computation of the observability and controllability Gramians as solutions of two algebraic Lyapunov equations. In this chapter, we demonstrate the efficiency of Lyapunov balanced truncation by numerical experiments for various settings of the output variables describing the aggregated characteristics of the temperature distribution in the GS.

In Chapter 5, we first reconsider the stochastic optimal control problem that we formulated in Chapter 2, where we replace the PDE describing the dynamics of the GS by a low-dimensional system of ODEs obtained by applying Lyapunov balanced truncation model order reduction to a high-dimensional semi-discretized system. Then, we explore its solution using dynamic programming and derive the associated Hamilton-Jacobi-Bellman (HJB) equation.

It turns out that the solution of the HJB equation suffers from the curse of dimensionality. Therefore, we discretize in Chapter 6 the problem in time to obtain a MDP and solve it numerically using backward recursion method. We first discretize the time interval $[0,T]$ into finite number of time points and solve the ODEs and SDEs describing the dynamics of the state variables within a short time interval defined by two arbitrary consecutive time points. Then we investigate the marginal distribution of individual state variables and the joint distribution of the correlated variables and formulate the discrete-time stochastic optimal control problem. Further, using state-discretization, we approximate the control problem by a MDP for a controlled finite state Markov chain and investigate the associated transition probabilities. This chapter ends with intensive numerical experiments where we study the behaviour of the value function and the optimal strategy with respect to time and different states.

In Chapter 7 we briefly introduce some alternative approximation methods to overcome the curse of dimensionality in the computation of the value function and the optimal strategy. These methods include Least Square Monte Carlo and Approximate Dynamic Programming for solving the discrete-time optimal control problem.

The dissertation ends with a summary of the results and an outlook on some possible future works in Chapter 8.

CHAPTER 2

Problem Formulation

Introduction

In this chapter, we consider a mathematical model of a residential heating system equipped with a GS, see Figure 2.1 and formulate the stochastic optimal problem for the cost-optimal management that heating system. In this model, we do not describe all the technical details of heat transfer to and from the consumption units of the building and the contribution of the solar collector. Instead, we only consider the aggregated residual demand resulting from the superposition of intermittent demand for thermal energy in the building and supply of thermal energy of solar collector. The main goal of this chapter is to embed the GS into a residential heating system equipped with a buffer storage and the renewable heat production unit such as the solar thermal collector and to formulate the associated cost-optimal management problem.

This chapter is organized as follows. We begin with a brief description of the residential heating system in Sec. 2.1 and in Sec. 2.2 we describe the control variables. Sec. 2.3 is devoted to the problem setup in which we describe various state processes of the residential heating system in continuous-time. In Sec. 2.4 we formulate the stochastic optimal control problem

2.1 Residential Heating System

In this section we briefly sketch the technical functionality of the residential heating system considered in this work.

This system is designed to provide thermal energy for heating and hot water supply of a building. Here the notion *building* is used for single family homes, office buildings, small companies or even small districts with a couple of buildings sharing a common heat and water supply. The building is equipped with some local production units for thermal energy such as solar collectors or other units using renewable energies. The supply of theses units usually does not meet exactly the demand of thermal energy due to the immanent temporal fluctuations and seasonality effects in both supply and demand. That imbalance is modeled by a stochastic process $R = (R(t)_{t\in[0,T]}$ where $R(t)$ is the residual demand at time t. Details will be given below in Subsec. 2.3.1

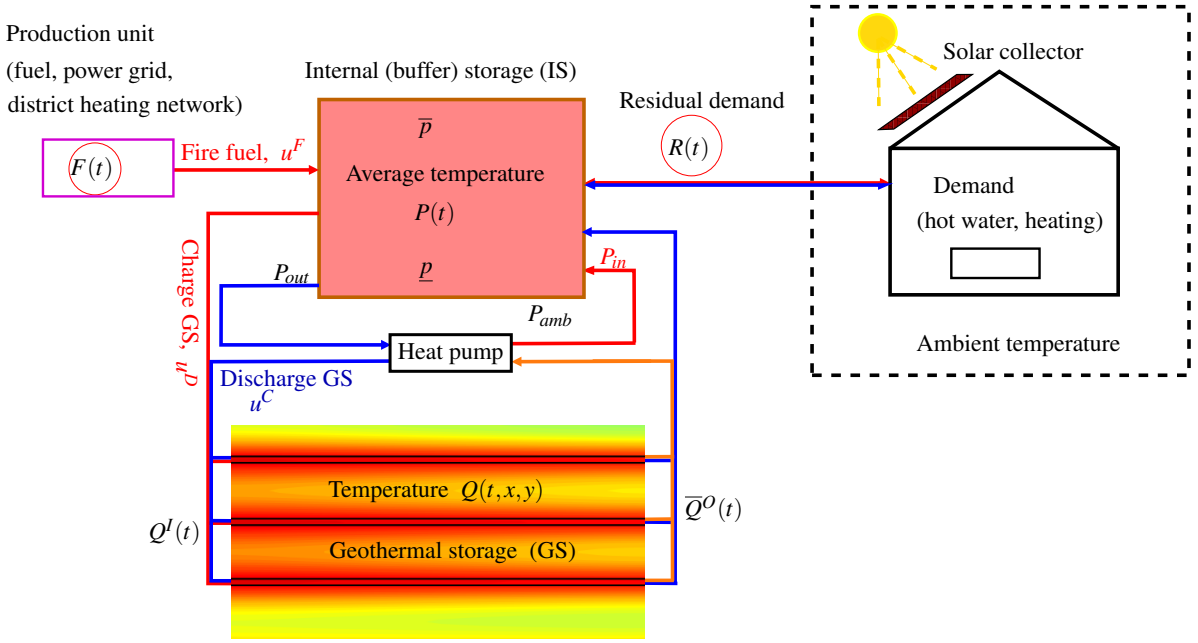


Figure 2.1: Simplified model of a residential heating system

If the demand exceeds the supply from local production then the excess demand $R(t) > 0$ is satisfied by an internal buffer storage. It also stores overproduction of thermal energy which is not needed to satisfy the demand. In that case $R(t)$ is negative. That internal storage is designed to balance supply and demand on shorter time scales of some hours or only few days. However, the capacity is not sufficient to serve as buffer for seasonal fluctuations on time scales of weeks and months.

As often observed in reality, we assume the internal storage to be a stratified hot water storage tank. The warmest storage layer is at the top and below there are colder layers through natural layering. For simplicity we assume that the storage can keep a constant temperature $\overline{p}$ at the top and also a constant temperature $\underline{p} < \overline{p}$ at the bottom. We do not model the vertical temperature profile in the storage but consider only the storage's average temperature which is denoted by $P(t)$ for time $t \in [0,T]$. Note that $P(t) - \underline{p}$ is proportional to stored thermal energy at time t available for extracting whereas $\overline{p} - \underline{p}$ and $\overline{p} - P(t)$ are proportional to the total storage capacity and the remaining storage capacity for future charging processes, respectively. The factor of proportionality is the mass times the specific heat capacity of the water in the tank.

Due to its limited capacity the internal storage cannot provide the necessary heat supply for a permanent or very strong unsatisfied demand of the building. Therefore the system is equipped with another production unit which is able to generate enough heat also on short time scales and to prevent the internal storage to become completely empty. That unit may fire fossil fuels (gas, oil, coal), convert electricity to heat using an immersion heater or obtain additional heat from a district heating system. In our work we address that unit as a fuel fired boiler but have the other options in mind. In all cases that heat production comes with additional costs arising from the consumption of fuel (or electricity or other respective heat sources). We denote the fuel price at time t by $F(t)$. Uncertainties about the future prices will be captured by modeling F as a stochastic process. More details will be given below in Subsec. 2.3.1

In periods of permanent or strong overproduction the internal storage may reach its capacity limits and can no longer accommodate more leftover heat from the local production. In order

to enable a later usage of that leftover heat, the heating system is equipped with an additional external thermal storage which in this work is a geothermal storage. Compared to the internal storage its capacity is much larger but it is also characterized by a lower temperature level. Therefore, heat pumps are required for transferring heat from the geothermal to the internal storage. Further, the transfer of thermal energy to and from the external storage depends on the often slow operation of heat exchangers. The geothermal storage is characterized by a non-homogeneous spatial and temporal temperature distribution describes the temperature $Q(t,x,y)$ of the storage medium at time t and position (x,y). The storage medium is assumed to be dry soil.

If the internal storage is already (almost) fully charged and there is still overproduction of thermal energy in the building then heat can be transferred from the internal to the external storage. This is obtained by sending a fluid of high temperature from the internal storage through the heat exchanger pipes of the geothermal storage. The fluid arrives at the (possibly multiple) inlets of the heat exchangers with the inlet temperature denoted by $Q^I(t)$. After passing the geothermal storage the fluid will leave the heat exchangers with a lower temperature. The average temperature at the (possibly multiple) outlets is denoted by $\overline{Q}^O(t)$ (for details see Subsection 2.3.3). This is also the temperature at which the fluid returns to the internal storage. Since the efficiency of charging the geothermal storage is improved by increasing the inlet temperature $Q^I(t)$, we assume in this work that during charging that temperature is equal to the maximum available temperature provided by the system which we denote by the constant Q^I_C.

On the other hand, if the internal storage is (almost) empty and there is still unsatisfied demand in the building then instead of producing heat from firing fuel, thermal energy can be also be transferred from the geothermal storage to the internal storage. For that process the system uses a heat pump for raising the temperature of the fluid arriving from the outlet of the geothermal storage to a higher level $P_{in} > \underline{p}$. Here P_{in} is a pre-specified temperature at which the fluid coming from the heat pump arrives at the internal storage. For simplicity we assume that P_{in} is constant. Without such a heat pump the heat transfer would be impossible due to the lower temperature level in the geothermal storage. The outlet temperature $\overline{Q}^O$ usually does not exceed $\underline{p}$, i.e., it will be not high enough to inject heat into the internal storage.

The heat pump connects two cycles in which moving fluids carry heat. A first cycle is connected to the geothermal storage. The fluid arrives from the storage's outlet at the inlet of the heat pump with temperature $\overline{Q}^O$. Due to the operation of the heat pump it leaves the pump with the temperature $Q^I_D < \overline{Q}^O$ and returns to the inlet of the geothermal storage. The thermal energy extracted from the fluid of the first cycle is transferred to the fluid in the second cycle. The latter connects the heat pump with the internal storage. At the pump's inlet arrives cold water of temperature P_{out} which is raised using the extracted heat in the first cycle and additional electrical energy to the temperature $P_{in} > P_{out}$ at which the fluid returns to the internal storage. In this work we assume for simplicity that the heat pump's operation is such that in the first cycle the outlet temperature Q^I_D is a constant. A modification to the case $Q^I_D(t) = \overline{Q}^O(t) - \Delta$ with some fixed temperature spread Δ is straightforward.

We recall that the geothermal storage is in some sense a hybrid storage since it may also serve as a production unit due to its open bottom boundary. This allows to use the thermal energy available in layers of the ground below the storage if the temperature in the storage falls below the temperature under the storage.

2.2 Control Process

In this section we describe the control variables. We use the notation introduced in the Sec. 2.1.

The GS is charged and discharged via the PHXs connected to the IS and filled with some fluid. Charging the GS is always by discharging the IS and vice-versa. When we charge the GS the fluid arrives at the inlet of the GS with a constant temperature of Q_C^I. This fluid goes though the GS and the heat propagates to the surrounding medium, leaves the GS and returns into the IS with a temperature $\overline{Q}^O < Q_C^I$. When we discharge the GS the fluid arrives at the inlet of the PHXs with a constant temperature $Q_D^I < Q_C^I$ and leaves the GS with some output temperature $\overline{Q}^O > Q_D^I$ which constitutes the inlet temperature of the heat pump. The heat pump uses additional electrical to raises the temperature to some constant $P_{in} > \underline{p}$.

We may also charge the IS by firing fuel or using electricity. When the IS is full, we switch off the pumps and the fuel fired-boiler (we wait or do noting). We assume that charging/discharging the GS is always at the maximum rate. Further, we assume that discharging or charging the GS and firing fuel simultaneously is never optimal (does not generate minimum cost). Therefore, the control $u(t)$ at time $t \in [0,T]$ is determine by the set of labels

$$\{u^C, u^D, u^F, u^W\},$$

with the following meaning.

- The control u^C is for charging the IS at the maximum rate by discharging the GS with fuel fired-boiler off.
- The control u^D means discharging the IS at the maximum rate to charge the GS with fuel fired-boiler off.
- The control u^F for heat-production using fuel/electricity at a maximum rate to charge the IS with pumps off.
- The control u^W for pumps off and no heat-production using fuel/electricity.

Now we emphasize on the control u^W (pumps and fuel fired-boiler off) for the case of strong negative residual demand. When the control u^W is applied, the change in the temperature of the IS is only due to the residual demand and the ambient temperature through the Newton's law of cooling. This control means do nothing (fuel boiler off, heat pump and conventional pump off).

However, the control u^W cannot be applied if we have a strong negative exogenous residual demand (overproduction of heat by the solar collector that must be stored in the IS) and the IS and the GS are simultaneously full (charging and discharging not possible). To handle such cases we introduce another control called *over-spilling* denoted by u^O which is similar to *do nothing* or *waiting* (with zero cost). We assume that during over-spilling the change in the temperature of the IS is only due to loss to the environment at ambient temperature P_{amb} by the Newton's law of cooling. Hence, the control $u(t)$ at any time t takes values in the set of labels

$$\overline{\mathcal{U}} = \{u^O, u^D, u^W, u^C, u^F\}.$$

2.3 Dynamics of the State Variables

In this section we formulate the mathematical model for a residential heating system which consists of production and consumption units, an internal and an external storage. The various

components are connected by pumps among them are heat pumps as well as ordinary pumps. The pumps control the flow rate between the IS and external storage. The model of the heating system is depicted in Fig. 2.1. We focus on the interplay of the IS and the GS. The IS allows for fast operation but has only a small capacity. The capacity of the GS is much higher, but operation is slow. In order to keep the exposition readable and understandable we restrict our study to a simple model which captures some physical aspects, but does not take all the engineering details into account. We setup the model for continuous-time $t \in [0,T]$ where $T > 0$ is a finite time horizon.

The state of the control system at $t \in [0,T]$ is given by the following quantities

$P(t)$	average temperature in the IS	$[^\circ C]$
$Q(t) = Q(t,x,y)$	temperature in the geothermal storage	$[^\circ C]$
$R(t)$	residual demand	[kW]
$F(t)$	energy price	[EUR/kWh]

We define the state process by $X = (R,F,P,Q)^\top$ taking values in some space $\mathcal{X}$.

Uncertainties. The uncertainties are modeled by a 2-dimensional standard Wiener process $W = (W_R, W_F)^\top$ on $[0,T]$ defined on a filtered probability space $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$. The filtration $\mathbb{G}$ is assumed to be generated by $(W(t))_{t\in[0,T]}$, i.e., $\mathbb{G} = \mathbb{G}^W = (\mathcal{G}^W(t))_{t\in[0,T]}$ and $\mathcal{G} = \mathcal{G}^W(T)$. In the following, we introduce each component of the state process X.

2.3.1 Residual Demand and Fuel Price

We denote by $R(t)$ the residual demand measured in kW which is the difference of the demand for thermal energy in the building and supply of thermal energy of solar collector at time $t \in [0,T]$. Further, we assume that $R(t) \in \mathcal{R} \subseteq \mathbb{R}$ and can be positive or negative. The residual demand is positive ($R(t) > 0$) when the demand exceeds supply and negative ($R(t) < 0$) when supply exceeds demand (overproduction). The dynamics of the residual demand $R(t)$ is given by

$$dR(t) = \beta_R(\mu_R(t) - R(t))dt + \sigma_R(t)dW_R(t), \quad R(0) = r_0 \in \mathcal{R} \subseteq \mathbb{R}, \tag{2.1}$$

with the mean-reversion level is given by

$$\mu_R(t) = \widetilde{\mu}_R(t) + \frac{1}{\beta_R}\dot{\widetilde{\mu}}_R(t). \tag{2.2}$$

Here, $\widetilde{\mu}_R(t) : [0,T] \to \mathbb{R}$ is a bounded deterministic differentiable function describing the residual demand's seasonality which we describe more detailed below in (2.3). The main idea of the shift is to have $\mathbb{E}[R(t)] - \widetilde{\mu}_R(t) \to 0$, as $t \to \infty$ and $\mathbb{E}[R(t) \mid R(0) = \widetilde{\mu}_R(0)] = \widetilde{\mu}_R(t)$. For constant $\widetilde{\mu}_R$ the latter holds true but for time dependent function $\widetilde{\mu}_R(t)$ the latter only holds true when we apply the shift (2.2).

Lemma 2.3.1 Under the transformation (2.2), the residual remand given by (2.1) satisfies the following properties:
$\mathbb{E}[R(t)] - \widetilde{\mu}_R(t) \to 0$, as $t \to \infty$ and $\mathbb{E}[R(t) \mid R(0) = \widetilde{\mu}_R(0)] = \widetilde{\mu}_R(t)$.

Proof. Using the Itô formula and the integration by part, the closed form solution of the SDE (2.1) is given by

$$\begin{aligned}
R(t) &= r_0 e^{-\beta_R t} + \int_0^t e^{-\beta_R(t-s)}\beta_R \mu_R(s)ds + \int_0^t e^{-\beta_R(t-s)}\sigma_R(s)dW_R(s) \\
&= r_0 e^{-\beta_R t} + \int_0^t e^{-\beta_R(t-s)}\beta_R\left(\widetilde{\mu}_R(s) + \frac{1}{\beta_R}\dot{\widetilde{\mu}}_R(s)\right)ds + \int_0^t e^{-\beta_R(t-s)}\sigma_R(s)dW_R(s) \\
&= r_0 e^{-\beta_R t} + \int_0^t e^{-\beta_R(t-s)}\beta_R\widetilde{\mu}_R(s)ds + \int_0^t e^{-\beta_R(t-s)}\dot{\widetilde{\mu}}_R(s)ds + \int_0^t e^{-\beta_R(t-s)}\sigma_R(s)dW_R(s) \\
&= r_0 e^{-\beta_R t} + \int_0^t e^{-\beta_R(t-s)}\beta_R\widetilde{\mu}_R(s)ds + \left[e^{-\beta_R(t-s)}\widetilde{\mu}_R(s)\right]_0^t - \int_0^t \beta_R e^{-\beta_R(t-s)}\widetilde{\mu}_R(s)ds \\
&\quad + \int_0^t e^{-\beta_R(t-s)}\sigma_R(s)dW_R(s) \\
&= r_0 e^{-\beta_R t} + \widetilde{\mu}_R(t) - \widetilde{\mu}_R(0)e^{-\beta_R t} + \int_0^t e^{-\beta_R(t-s)}\sigma_R(s)dW_R(s).
\end{aligned}$$

Therefore, $\mathbb{E}[R(t)] - \widetilde{\mu}_R(t) \to 0$, as $t \to \infty$. Further, $\mathbb{E}[R(t) \mid R(0) = \widetilde{\mu}_R(0)] = \widetilde{\mu}_R(t)$.
□

The parameter β_R is a constant mean-reversion speed and $\sigma_R : [0,T] \to \mathbb{R}$ is the deterministic and positive bounded volatility. For the seasonality functions mentioned above we work with functions of the form

$$\widetilde{\mu}_R(t) = k_0^R + \sum_{i=1}^{M} k_i^R \cos\frac{2\pi(t - t_i^R)}{\delta_i^R}, \tag{2.3}$$

where $k_0^R > 0$ and $k_i^R \geq 0, i = 1,\ldots,M$, are constants with k_0^R the long-term mean, k_i^R the amplitude of the seasonality, δ_i^R the length of the seasonal period and t_i^R some time shift parameter for the i-th seasonality component (representing the time of the seasonal peak of the residual demand) and M is the number of components. The quantity

$$\frac{1}{\beta_R}\dot{\widetilde{\mu}}_R(t) = -\sum_{i=1}^{M} k_i^R \frac{2\pi}{\beta_R \delta_i^R}\sin\frac{2\pi(t - t_i^R)}{\delta_i^R}$$

represents the shift in the seasonality function. A typical choice is $M = 2$ and $\delta_1^R = 1$ year, $\delta_2^R = 1$ day, and the reference time $t_1^R = t_2^R = 0$.

Fuel/Electricity price. The fuel or electricity price F is a stochastic mean reverting process (with the same structure like (2.1)) given by

$$dF(t) = \beta_F(\mu_F(t) - F(t))dt + \sigma_F(t)dW_F(t), \quad F(0) = f_0 \in \mathcal{F} = \mathbb{R}, \tag{2.4}$$

with seasonality $\mu_F(t)$ having the same structure like (2.2). For further simplification, we can assume that fuel or electricity price is constant or a deterministic function of time. In this case the dimension of the state process is reduced by one variable.

Assumption 2.3.2 We assume that $\sigma_R(t) \geq \underline{\sigma}_R > 0$ and $\sigma_F(t) \geq \underline{\sigma}_F > 0$, for some constants $\underline{\sigma}_R$ and $\underline{\sigma}_R$.

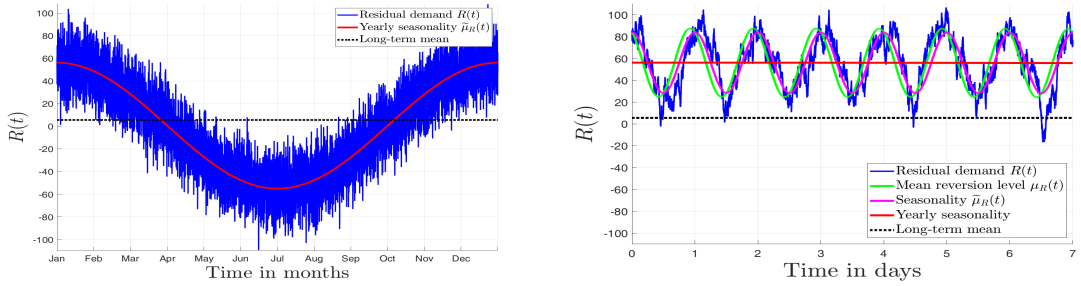


Figure 2.2: Residual demand over a period of one year (left) and a one week zoom in at the beginning of February (right) with parameters $\beta = 0.5$, $\mu_R(0) = 47.26$, $\sigma^R = 13.95$, $k_0^R = 0.56$ $k_1^R = 55.6$, $k_2^R = 13.9$, $\delta_1^R = 365$, and $\delta_2^R = 365 \times 24$.
Red solid line for the yearly seasonality component, blue solid line for the residual demand, green solid line for the mean reversion level, magenta solid line for the seasonality function, and black dotted line for the long-term mean level.

2.3.2 Spatial Temperature Distribution in the Geothermal Storage

In this section we describe the dynamics of the spatial temperature distribution in a GS. In contrast to the other state variables the temperature in the GS, denoted by $Q = Q(t,x,y)$ depends not only on time but also on the location in space. Compared to the IS which is a water tank (buffer storage) the capacity of the GS is much larger but it is also characterized by a lower temperature level. Therefore, heat pumps are required for transferring heat from the geothermal to the internal storage. Charging and discharging is not efficient or even impossible if there are only small differences between the temperatures inside and in the vicinity of the pipes. Long periods of (dis)charging may lead to saturation in the vicinity of the pipes such that (dis)charging is no longer efficient and should be stopped since propagation of heat to regions away from the PHXs takes time. Mathematically, the dynamics of the spatial temperature distribution in a GS is described by a linear heat equation with convection term and appropriate boundary and interface conditions. In this model we focus only on the storage without the surrounding region (see black solid rectangle in Fig. 2.3). These physical effects imply that the cost-optimal management of a heating system equipped with a GS depends on the space-time dynamics of the temperature of the storage. The control has to be *forward-looking* and account for the slow response of the storage. Therefore, we now study a mathematical model of the evolution of the spatial temperature distribution inside the storage.

2D-Model

We assume that the domain of the GS is a cuboid and consider a two-dimensional rectangular cross-section. We denote by $Q = Q(t,x,y)$ the temperature at time $t \in [0,T]$ at the point $(x,y) \in \mathcal{D} = (0,l_x) \times (0,l_y)$ with l_x, l_y denoting the width and height of the storage. The domain $\mathcal{D}$ and its boundary $\partial\mathcal{D}$ are depicted in Fig. 2.4. $\mathcal{D}$ is divided into three parts. The first is $\mathcal{D}^M$ and is filled with a homogeneous medium (soil) characterized by constant material parameters ρ^M, κ^M and c_p^M denoting mass density, thermal conductivity and specific heat capacity, respectively. The second is $\mathcal{D}^F$, it represents the PHXs filled with a fluid (water) with constant material parameters ρ^F, κ^F and c_p^F. The fluid moves with time-dependent velocity $v_0(t)$ along the pipe. For the sake of simplicity, we restrict to the case, often observed in applications, where the

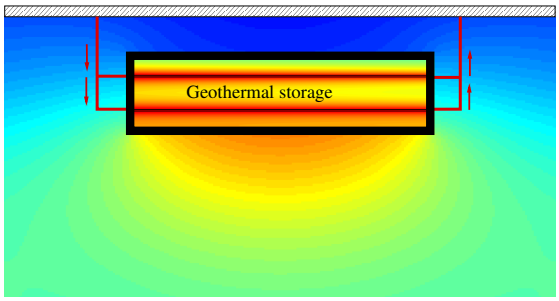


Figure 2.3: 2D-model of a GS insulated at the top and the sides while open at the bottom and spatial temperature distribution

pumps moving the fluid are either on or off. Thus the velocity $v_0(t)$ is piece-wise constant taking values $\bar{v}_0 > 0$ and zero, only. Finally, the third part is the interface $\mathcal{D}^J$ between $\mathcal{D}^M$ and $\mathcal{D}^F$. That interface is split into upper and lower interfaces $\overline{\mathcal{D}}^J$ and $\underline{\mathcal{D}}^J$, respectively. Observe that we neglect modeling the wall of the pipe and suppose perfect contact between the pipe and the soil. Details are given below in (2.11) and (2.12). Summarizing we make the following

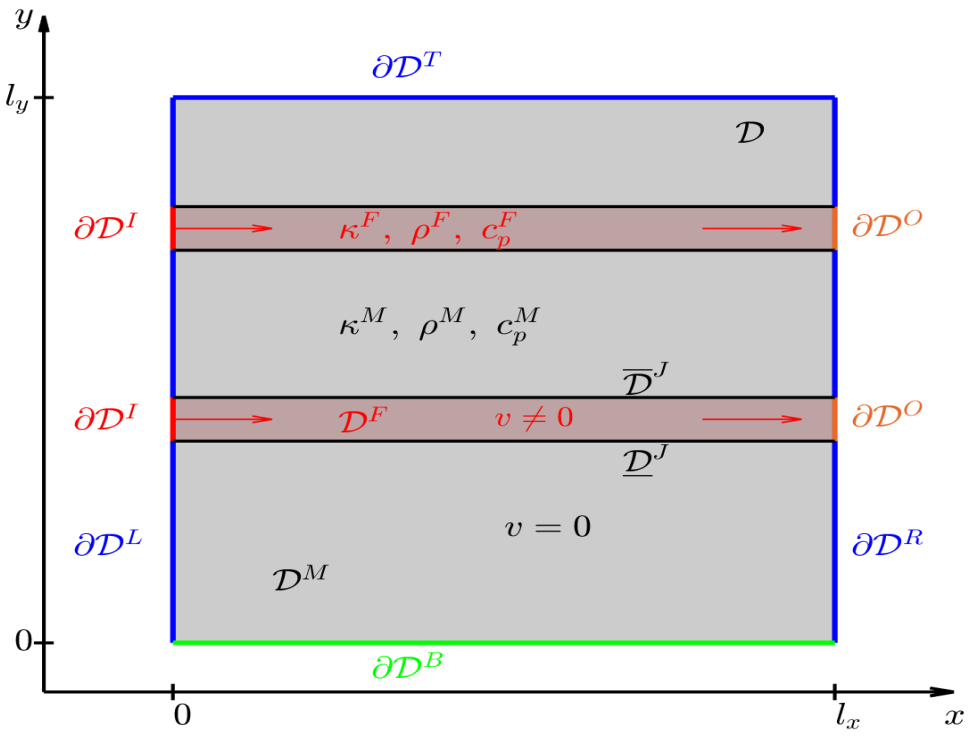


Figure 2.4: 2D-model of the GS: decomposition of the domain $\mathcal{D}$ and the boundary $\partial\mathcal{D}$.

Assumption 2.3.3

1. Material parameters of the medium ρ^M, κ^M, c_p^M in the domain $\mathcal{D}^M$ and of the fluid ρ^F, κ^F, c_p^F in the domain $\mathcal{D}^F$ are constants.

2. Fluid velocity is piecewise constant, i.e. $v_0(t) = \begin{cases} \bar{v}_0 > 0, & \text{pump on,} \\ 0, & \text{pump off.} \end{cases}$

3. Perfect contact at the interface between fluid and medium.

4. There are $n_P \in \mathbb{N}$ straight horizontal pipes, the fluid moves in positive x-direction.

Heat equation. The temperature $Q = Q(t,x,y)$ in the external storage is governed by the linear heat equation with convection term

$$\rho c_p \frac{\partial Q}{\partial t} = \nabla \cdot (\kappa \nabla Q) - \rho v \cdot \nabla (c_p Q), \quad (t,x,y) \in (0,T] \times \mathcal{D} \backslash \mathcal{D}^J, \tag{2.5}$$

where the first term on the right hand side describes diffusion, while the second represents convection of the moving fluid in the pipes. Further, $v = v(t,x,y) = v_0(t)(v^x(x,y), v^y(x,y))^\top$ denotes the velocity vector with $(v^x, v^y)^\top$ being the normalized directional vector of the flow. According to Assumption 2.3.3 the material parameters ρ, κ, c_p depend on the position (x,y) and take the values ρ^M, κ^M, c_p^M for points in $\mathcal{D}^M$ (medium) and ρ^F, κ^F, c_p^F in $\mathcal{D}^F$ (fluid).

Note that there are no sources or sinks inside the storage and therefore the above heat equation appears without forcing term. Based on this assumption, the heat equation (2.5) can be written as

$$\frac{\partial Q}{\partial t} = a \Delta Q - v \cdot \nabla Q, \quad (t,x,y) \in (0,T] \times \mathcal{D} \backslash \mathcal{D}^J, \tag{2.6}$$

where $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the Laplace operator, $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)$ the gradient operator, and $a = a(x,y)$ is the thermal diffusivity which is piecewise constant with values $a^M = \frac{\kappa^M}{\rho^M c_p^M}$ for $(x,y) \in \mathcal{D}^M$ and $a^F = \frac{\kappa^F}{\rho^F c_p^F}$ for $(x,y) \in \mathcal{D}^F$. The initial condition $Q(0,x,y) = Q_0(x,y)$ is given by the initial temperature distribution Q_0 of the storage.

Remark 2.3.4 In real-world GSs heat exchanger pipes are often designed in a snake form located in the storage domain at multiple horizontal layers as it is sketched in Fig. 1.2. Typically there is only a single inlet and outlet. We will mimic that design by a computationally more tractable design characterized by multiple horizontal straight pipes. On the one hand, this allows to control the different pipes separately. On the other hand, an approximation of the widely used snake-shaped pipe design can be derived by connecting the outlet of one pipe with the inlet of the next.

Boundary and interface conditions

For the description of the boundary conditions we decompose the boundary $\partial\mathcal{D}$ into several subsets as depicted in Fig. 2.4 representing the insulation at the top and the side, the open bottom, the inlet and outlet of the pipes. Further, we have to specify conditions at the interface between pipes and soil. The inlet, outlet and the interface conditions model the heating and cooling of the storage via PHXs. We distinguish between the two regimes 'pump on' and 'pump off' where for simplicity we assume perfect insulation at inlet and outlet if the pump is off. This leads to the following boundary conditions.

- *Homogeneous Neumann condition* describing perfect insulation at the top and the side

$$\frac{\partial Q}{\partial \mathbf{n}} = 0, \qquad (x,y) \in \partial\mathcal{D}^T \cup \partial\mathcal{D}^L \cup \partial\mathcal{D}^R, \tag{2.7}$$

where $\partial\mathcal{D}^L = \{0\} \times [0, l_y] \backslash \partial\mathcal{D}^I$, $\partial\mathcal{D}^R = \{l_x\} \times [0, l_y] \backslash \partial\mathcal{D}^O$ and $\partial\mathcal{D}^T = [0, l_x] \times \{l_y\}$.

- *Robin condition* describing heat transfer at the bottom

$$-\kappa^M \frac{\partial Q}{\partial \mathbf{n}} = \lambda^G (Q - Q^G(t)), \qquad (x,y) \in \partial\mathcal{D}^B, \tag{2.8}$$

with $\partial\mathcal{D}^B = [0, l_x] \times \{0\}$, where $\lambda^G > 0$ denotes the heat transfer coefficient and $Q^G(t)$ the underground temperature. For more interpretation we refer to Remark 2.3.6.

- *Dirichlet condition* at the inlet if the pump is on ($v_0(t) > 0$), i.e. the fluid arrives at the storage with a given temperature $Q^I(t)$. If pump is off ($v_0(t) = 0$), we set a homogeneous Neumann condition describing perfect insulation.

$$\begin{cases} Q = Q^I(t), & \text{pump on,} \\ \frac{\partial Q}{\partial \mathbf{n}} = 0, & \text{pump off,} \end{cases} \qquad (x,y) \in \partial\mathcal{D}^I. \tag{2.9}$$

- *Do Nothing condition* at the outlet in the following sense. If the pump is on ($v_0(t) > 0$) then the total heat flux directed outwards can be decomposed into a diffusive heat flux given by $k^F \frac{\partial Q}{\partial \mathbf{n}}$ and a convective heat flux given by $v_0(t) \rho^F c_p^F Q$. Since in real-world applications the latter is much larger than the first we neglect the diffusive heat flux. This leads to a homogeneous Neumann condition

$$\frac{\partial Q}{\partial \mathbf{n}} = 0, \qquad (x,y) \in \partial\mathcal{D}^O. \tag{2.10}$$

If the pump is off then we assume (as already for the inlet) perfect insulation which is also described by the above condition.

- *Smooth heat flux* at interface $\mathcal{D}^J$ between fluid and soil leading to a coupling condition

$$\kappa^F \frac{\partial Q^F}{\partial \mathbf{n}} = \kappa^M \frac{\partial Q^M}{\partial \mathbf{n}}, \qquad (x,y) \in \mathcal{D}^J. \tag{2.11}$$

Here, Q^F, Q^M denote the temperature of the fluid inside the pipe and of the soil outside the pipe, respectively. Moreover, we assume that the contact between the pipe and the medium is perfect which leads to a smooth transition of a temperature, i.e., we have

$$Q^F = Q^M, \qquad (x,y) \in \mathcal{D}^J. \tag{2.12}$$

Remark 2.3.5 If the contact between the pipe and the medium is not perfect (e.g., in case of contact resistance) then the transition of the temperature at the interface $\mathcal{D}^J$ will not be smooth, that is, $Q^F \neq Q^M$. This leads to a temperature jump between the pipe and the medium. That phenomenon occurs in the heat transfer between the medium and an insulation as shown in [9].

Remark 2.3.6 Imposing the Robin condition (2.8) at the bottom boundary is indeed only an attempt to mimic the thermal behavior at the bottom boundary. A more realistic description requires embedding the storage domain $\mathcal{D}$ into a larger computational domain including the surrounding regions as in Fig. 1.2. This allows for warming and cooling in the vicinity of the storage resulting from the outflow and inflow of the storage heat. Contrary to that, condition

(2.8) assumes an exogenously given underground temperature Q_G independent of the temperature in the storage.

The heat transfer coefficient λ^G describes the resistance to the heat flux at the boundary. For the limiting case $\lambda^G \to 0$ we get a homogeneous Neumann condition, i.e., perfect insulation, while in the limit for $\lambda^G \to \infty$ condition (2.8) is the Dirichlet condition $Q = Q^G(t)$. The underground temperature in general shows seasonal fluctuations which can be described by $Q^G(t) = k_1^G \cos\left(\frac{2\pi t}{T_a}\right) + k_2^G$, where k_1^G is the intensity of the fluctuation, k_2^G is the average ground temperature and T_a the number of time units per year. Since our focus is on the short-term behavior, we assume in the sequel that the underground temperature is constant over time, i.e. $k_1^G = 0$.

2.3.3 Aggregated Characteristics

The PDE (2.5) allows to describe the spatio-temporal temperature distribution in the GS. In many applications it is not necessary to know the complete information about that distribution. An example is the optimal management of a residential heating system equipped with such storage that we will consider in Chapter 5. Here it is sufficient to know only a few aggregated characteristics of the temperature distribution which can be computed via post-processing after solving the PDE. In this section we introduce some of these aggregated characteristics.

Aggregated characteristics related to the amount of stored energy. We start with aggregated characteristics given by the average temperature in some given subdomain of the storage which are related to the amount of stored energy in that domain.

Let $\mathcal{B} \subset \mathcal{D}$ be a generic subset of the 2D computational domain. We denote by $|\mathcal{B}| = \iint_{\mathcal{B}} dxdy$ the area of $\mathcal{B}$. Then $W_{\mathcal{B}}(t) = l_z \iint_{\mathcal{B}} \rho c_p Q(t,x,y)dxdy$ represents the thermal energy contained in the 3D spatial domain $\mathcal{B} \times [0, l_z]$ at time $t \in [0,T]$ for $l_z > 0$. Then for $0 \le t_0 < t_1 \le T$ the difference $G_{\mathcal{B}}(t_0,t_1) = W_{\mathcal{B}}(t_1) - W_{\mathcal{B}}(t_0)$ is the gain of thermal energy during the period $[t_0,t_1]$. While positive values correspond to warming of $\mathcal{B}$, negative values indicate cooling and $-G_{\mathcal{B}}(t_0,t_1)$ represents the magnitude of the loss of thermal energy.

For $\mathcal{B} = \mathcal{D}^\dagger$, where $\dagger = M,F$, we can use that the material parameters on $\mathcal{D}^\dagger$ equal to the constants $\rho = \rho^\dagger, c_p = c_p^\dagger$. Thus, for the corresponding gain of thermal energy we obtain

$$\begin{aligned} G^\dagger = G^\dagger(t_0,t_1) &:= G_{\mathcal{D}^\dagger}(t_0,t_1) = \rho^\dagger c_p^\dagger |\mathcal{D}^\dagger| l_z \left(\overline{Q}^\dagger(t_1) - \overline{Q}^\dagger(t_0)\right), \\ \text{where} \quad \overline{Q}^\dagger(t) &= \frac{1}{|\mathcal{D}^\dagger|} \iint_{\mathcal{D}^\dagger} Q(t,x,y)dxdy, \quad \dagger = M,F, \end{aligned} \tag{2.13}$$

denotes the average temperature in the medium ($\dagger = M$) and the fluid ($\dagger = F$), respectively. We denote by $\overline{Q}^S$ the average temperature in the whole storage. It can be obtained from $\overline{Q}^M$ and $\overline{Q}^F$ by

$$\overline{Q}^S(t) = \frac{1}{|\mathcal{D}|}\left(\overline{Q}^M(t)\,|\mathcal{D}^M| + \overline{Q}^F(t)\,|\mathcal{D}^F|\right).$$

Further, the total gain in the storage denoted by G^S is obtained by

$$G^S = G^S(t_0,t_1) = G^M(t_0,t_1) + G^F(t_0,t_1).$$

Aggregated characteristics related to the heat flux at the boundary. Now we consider the

convective heat flux at the inlet and outlet boundary and the diffusive heat flux at the bottom boundary. Let $\mathcal{C} \subset \partial\mathcal{D}$ be a generic curve on the boundary, then we denote by $|\mathcal{C}| = \int_{\mathcal{C}} ds$ the curve length.

The rate at which the energy is injected or withdrawn via the pipe is given by

$$\begin{aligned} R^P(t) &= \rho^\dagger c_p^\dagger v_0(t) \Big[\int_{\mathcal{D}^I} Q(t,x,y)\, ds - \int_{\mathcal{D}^O} Q(t,x,y)\, ds \Big] \\ &= \rho^\dagger c_p^\dagger v_0(t) |\partial\mathcal{D}^O| [Q^I(t) - \overline{Q}^O(t)], \qquad (2.14) \\ \text{where} \quad \overline{Q}^O(t) &= \frac{1}{|\partial\mathcal{D}^O|} \int_{\partial\mathcal{D}^O} Q(t,x,y) ds \end{aligned}$$

is the average temperature at the outlet boundary. Here, we have used that in our model we have horizontal pipes such that $|\partial\mathcal{D}^I| = |\partial\mathcal{D}^O|$ and a uniformly distributed inlet temperature at the inlet boundary $\partial\mathcal{D}^I$. Note that the fluid moves at time t with velocity $v_0(t)$ and arrives at the inlet with temperature $Q^I(t)$ while it leaves at the outlet with the average temperature $\overline{Q}^O(t)$. For a given interval of time $[t_0, t_1]$ the quantity

$$G^P = G^P(t_0,t_1) = l_z \int_{t_0}^{t_1} R^P(t)\, dt$$

describes the amount of heat injected ($G^P > 0$) to or withdrawn ($G^P < 0$) from the storage due to convection of the fluid.

Next we look at the diffusive heat transfer via the bottom boundary and define the rate

$$\begin{aligned} R^B(t) &= \int_{\mathcal{D}^B} \kappa^M \frac{\partial Q}{\partial \mathbf{n}}\, ds = \int_{\mathcal{D}^B} \lambda^G (Q^G(t) - Q(t,x,y))\, ds \\ &= \lambda^G |\partial\mathcal{D}^B| (Q^G(t) - \overline{Q}^B(t)), \qquad (2.15) \\ \text{where} \quad \overline{Q}^B(t) &= \frac{1}{|\partial\mathcal{D}^B|} \int_{\partial\mathcal{D}^B} Q(t,x,y) ds \end{aligned}$$

is the average temperature at the bottom boundary. Note that the second equation in the first line follows from the Robin boundary condition. The quantity

$$G^B = G^B(t_0,t_1) = l_z \int_{t_0}^{t_1} R^B(t)\, dt$$

describes the amount of heat transferred via the bottom boundary of the storage.

Energy balance. In our model we assume perfect thermal insulation at all boundaries except the inlet, outlet and the bottom boundary. At the outlet we impose a homogeneous Neumann condition describing zero diffusive heat transfer. At the inlet we also have a zero diffusive heat transfer under the reasonable assumption that the temperature in the supply pipe is constant and equals $Q^I(t)$, thus the normal derivative $\frac{\partial Q}{\partial \mathbf{n}}$ is zero. This implies that gains and losses of thermal energy in the storage are caused either by injections or withdrawals via the heat exchanger pipes or by heat transfer via the open bottom boundary. Thus, we can decompose the total gain G^S to obtain the following energy balance

$$G^S = G^M + G^F = G^P + G^B.$$

2.3.4 Analogous Model

Note that in the dynamics of the GS described by equation (2.5), the velocity $v_0 = v_0(t)$ is time-dependent. Further, it is assumed that the fluid velocity is constant $\overline{v}_0$ during (dis)charging when the pump is on, and zero during waiting when the pump is off. Thus, the dynamics of the GS has two regimes, say, pump on and pump off. In order to apply MOR that we consider later in Chapter 4, the PDE (2.5) together with the boundary condition at the inlet of the pipe (2.9) need to be slightly modified to obtain a model with only one regime, pump on. The latter is a crucial assumption for most of model reduction methods such as the Lyapunov balanced truncation technique. We circumvent this problem by replacing the model for the GS by a so-called *analogous model*

The key idea for the construction of such an analogue is based on the observation that under the assumption 2.3.3, the "original model" has a differential operator with piece-wise time-invariant coefficients. This is due to our assumption that the fluid velocity is constant $\overline{v}_0$ during (dis)charging when the pump is on, and zero during waiting when the pump is off. This leads to the following approximation of the original by an analogous model.

For the analogous model we assume that contrary to the original model the fluid is also moving with constant velocity $\overline{v}_0$ during pump-off periods. During these waiting periods in the original model the fluid is at rest and only subject to the diffusive propagation of heat. In order to mimic that behavior of the resting fluid by a moving fluid we assume that the temperature Q^I at the pipe's inlet is equal to the average temperature of the fluid in the pipe $\overline{Q}^F$. From a physical point of view we will preserve the average temperature of the fluid but a potential temperature gradient along the pipe is not preserved and replaced by an almost flat temperature distribution. It can be expected that the error induced by this "mixing" of the fluid temperature in the pipe is small after sufficiently long (dis)charging periods leading to saturation with an almost constant temperature along the pipe.

In the mathematical description by an initial boundary value problem for the heat equation (2.6), the above approximation leads to a modified boundary condition at the inlet. During waiting the homogeneous Neumann boundary condition in (2.9) is replaced by a non-local coupling condition such that the inlet boundary condition reads as

$$Q(t) = \begin{cases} Q^I(t), & \text{pump on,} \\ \overline{Q}^F(t), & \text{pump off,} \end{cases} \qquad (x,y) \in \partial\mathcal{D}^I. \tag{2.16}$$

That condition is termed 'non-local' since the inlet temperature is not only specified by a condition to the local temperature distribution at the inlet boundary $\partial\mathcal{D}^I$ but it depends on the whole spatial temperature distribution in the fluid domain $\mathcal{D}^F$.

2.3.5 Internal Storage

The IS is assumed to be a non-stratified water tank. For the ease of exposition we assume that the technical implementation is such that there is a constant and known bottom temperature $\underline{p}$ and top temperature $\overline{p} > \underline{p}$. Further, we assume that the average temperature $P(t)$, considered as state variable satisfies the state constraint $P(t) \in [\underline{p}, \overline{p}]$.
We assume that charging the GS by discharging the IS is such that a (conventional) pump sends fluid with an inlet temperature Q_C^I from the IS to the GS and the fluid returns to the IS with a temperature $\overline{Q}^O(t)$. Charging the IS by discharging the GS is such that heat pump raises the

temperature of the fluid from P_{out}, to a given constant and known temperature P_{in} with $P_{in} > \underline{p}$. More details about charging and discharging cycles are given in Sec. 2.1

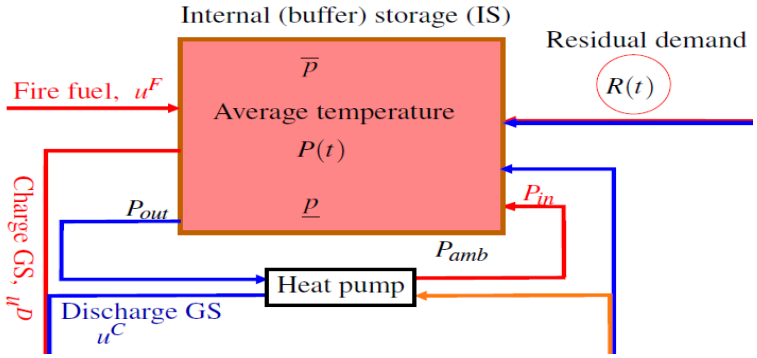


Figure 2.5: Changes of thermal energy in the IS

Further, we assume that changes of thermal energy in the IS are due to inflow of energy from overproduction, from the GS, by firing fuel. Further, there may be outflow of energy to satisfy the positive residual demand, to the GS, due to the loss to environment as depicted in Fig. 2.5. The environment is assumed to be at a constant temperature $P_{amb} < \underline{p}$. The dynamics of the IS is then given by

$$dP(t) = (\psi_p(R(t), F(t), \overline{Q}^O(t), u(t)) - \gamma(P(t) - P_{amb}))dt, \;\; P(0) = p_0 \in \mathcal{P} \subset \mathbb{R} \quad (2.17)$$

where R is the residual demand given by equation (2.1), F the fuel price, and $\overline{Q}^O$ is the average outlet temperature of the PHX . The quantity $-\gamma(P(t) - P_{amb}))$ is the heat loss to the environment at given time t, where $\gamma = \frac{\kappa_h A_h}{m^P c_P^F}$ is a constant with m^P the mass of the water in the IS, c_P^F the specific heat capacity of the water, κ_h the overall heat transfer coefficient, A_h the total surface of the IS and P_{amb} the ambient temperature. The function ψ_p is given by

$$\psi_P(r, f, \overline{Q}^O, \nu) = \begin{cases} -k_P r + \kappa_F & \nu = u^F, \\ -k_P r + \kappa_C(P_{in} - P_{out}) & \nu = u^C, \\ -k_P r & \nu = u^W, \\ -k_P r - \kappa_D(Q_C^I - \overline{Q}^O) & \nu = u^D, \\ 0 & \nu = u^O. \end{cases} \quad (2.18)$$

where $k_P = \frac{1}{m^P c_P^F}$, $\kappa_D = \frac{k_D}{m^P c_P^F}$, $\kappa_C = \frac{k_C}{m^P c_P^F}$ and $\kappa_F = \frac{k_F}{m^P c_P^F}$ are positive constants. The increment of the total thermal energy in the IS at time t is given by $m^P c_P^F dP(t)$. In the dynamics of the IS we assume that $P_{in} \geq P_{out}$ and $Q_C^I \geq \overline{Q}^O$ for all $t \in [0, T]$. The residual demand appears in the dynamics of P with negative sign because a positive residual demand decreases the temperature in the IS and a negative residual demand increases temperature in the IS. In the next section we will setup a continuous-time optimal control problem for the cost-optimal management of the residential heating system equipped with a buffer storage (water tank), a GS and a local renewable heat production unit.

2.4 Continuous-Time Stochastic Optimal Control Problem

2.4.1 Controlled State

In this section we formulate the optimal control problem for the residential heating system with GS in continuous-time. We focus on the embedding of the GS into a residential heating system and derive the associated optimal control problem. Let $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ be the filtered probability space introduced in Sec. 2.3 and carrying two Brownian motions W_R and W_F. Then, the state process $X = (R, F, P, Q)^\top$ is adapted to the filtration $\mathbb{G}$.

State dynamics. Let $\mathcal{X} = \mathcal{R} \times \mathcal{F} \times \mathcal{P} \times \mathcal{Q}$ be a suitable chosen state space, with $\mathcal{R} \times \mathcal{F} \times \mathcal{P} \subset \mathbb{R}^3$ and $\mathcal{Q}$ a function space defined by $\mathcal{Q} = \mathcal{C}^{1,2}((0,T] \times \mathcal{D} \backslash \mathcal{D}^J) \cap \mathcal{C}((0,T] \times \mathcal{D} \cup \partial\mathcal{D})$. The infinite dimensional state process $X = (R, F, P, Q)^\top$, $X \in \mathcal{X}$ is governed by the following equations

$$
\begin{aligned}
dR(t) &= \beta_R(\mu_R(t) - R(t))dt + \sigma_R(t)dW_R(t), && R(0) = r_0,\\
dF(t) &= \beta_F(\mu_F(t) - F(t))dt + \sigma_F(t)dW_F(t), && F(0) = f_0,\\
dP(t) &= (\psi_p(R(t), F(t), \overline{Q}^O(t), u(t)) - \gamma(P(t) - P_{amb}))dt, && P(0) = p_0,\\
\frac{\partial Q(t,x,y)}{\partial t} &= a(x,y)\Delta Q(t,x,y) - v(t,x,y) \cdot \nabla Q(t,x,y),\ (t,x,y) \in (0,T] \times \mathcal{D} \backslash \mathcal{D}^J,\ Q(0,x,y) = Q_0,
\end{aligned}
$$

+boundary and interface conditions given in (2.7) through (2.12).

The state can be decomposed into two parts: $X^u = (\widehat{X}(t), \overline{X}^u(t))^\top$ where $\widehat{X} \in \mathcal{R} \times \mathcal{F} \subset \mathbb{R}^2$ is an exogenous state variable and $\overline{X}^u \in \mathcal{P} \times \mathcal{Q}$ is an endogenous state variable. The uncontrolled state $\widehat{X} = (R, F)^\top \in \mathcal{R} \times \mathcal{F}$satisfies the SDE

$$
d\widehat{X}(t) = \widehat{\mu}(t, \widehat{X}(t))\, dt + \widehat{\sigma}(t) d\widehat{W}(t), \quad \widehat{X}(0) = \widehat{x}_0 = (r_0, f_0)^\top \in \mathcal{R} \times \mathcal{F},
$$

where the uncertainty $\widehat{W} = (W_R, W_F)^\top$, the drift coefficient $\widehat{\mu} : [0,T] \times \mathcal{R} \times \mathcal{F} \to R^2$ and the volatility matrix $\widehat{\sigma} : [0,T] \times \mathcal{R} \times \mathcal{F} \to \mathbb{R}^{2\times 2}$ are defined for $\widehat{x} = (r, f)$ as

$$
\widehat{\mu}(t, \widehat{x}) = \begin{pmatrix} \beta_R(\mu_R(t) - r) \\ \beta_F(\mu_F(t) - f) \end{pmatrix} \in \mathbb{R}^2, \qquad \widehat{\sigma}(t) = \begin{pmatrix} \sigma_R(t) & 0 \\ 0 & \sigma_F(t) \end{pmatrix} \in \mathbb{R}^{2\times 2}. \tag{2.19}
$$

For deterministic known fuel price F, the exogenous state variable $\widehat{X} = R \in \mathcal{R}$ given by equation (2.1) and $\widehat{\mu}(t,r) = \beta_R(\mu_R(t) - r)$, $\widehat{\sigma}(t) = \sigma_R(t) \in \mathbb{R}$.
The controlled state $\overline{X}^u = (P, Q)^\top$ satisfies the a system of ODE and a PDE given by

$$
\begin{aligned}
dP(t) &= (\psi_p(R(t), F(t), \overline{Q}^O(t), u(t)) - \gamma(P(t) - P_{amb}))dt, && P(0) = p_0,\\
\frac{\partial Q(t,x,y)}{\partial t} &= a(x,y)\Delta Q(t,x,y) - v(t,x,y) \cdot \nabla Q(t,x,y),\ (t,x,y) \in (0,T] \times \mathcal{D} \backslash \mathcal{D}^J,\ Q(0,x,y) = Q_0,
\end{aligned}
$$

+boundary and interface conditions given in (2.7) through (2.12).

2.4.2 Control and State Constraints

Due to some environmental regulations fixed by the authorities, we require that the state process X always satisfied the state constraint

$$X(t) \in \mathcal{K} = \Big\{ x \in \mathcal{X},\ \ p \in [\underline{p}, \overline{p}],\ \overline{Q}^M(t) \in [\underline{q}, \overline{q}] \Big\},$$

where $\overline{Q}^M$ is the average temperature in the GS (pipes not included), given by (2.13). For the storage rate process $u = (u(t))_{t\in[0,T]}$, we require that at any time t

$$u(t) \in \overline{\mathcal{U}} = \{u^O, u^D, u^W, u^C, u^F\}.$$

In addition to that constraints the controller will face further operational constraints leading to time- and state-dependent control constraints and a restriction of the set $\overline{\mathcal{U}}$ to the set of (actually) feasible controls of the form $\mathcal{U}(t,x)$ where $\mathcal{U}$ is a set-valued function mapping $(t,x) \in [0,T] \times \mathcal{X}$ to subsets of $\overline{\mathcal{U}}$. At each time $t \in [0,T]$ we require that the Markovian control $\widetilde{u}(t,x)$ is such that the control $\widetilde{u}(t,x)$ lies in that set, i.e, $\widetilde{u}(t,x) \in \mathcal{U}(t,x)$, for some measurable function $\widetilde{u}$. Examples of state-dependent constraints are

- no charging for a full storage but discharging is allowed
- no discharging for an empty storage but charging is allowed.

We say that the GS is empty if $\overline{Q}^M \leq \underline{q}$ and the GS is full if $\overline{Q}^M \geq \overline{q}$. The IS is empty if $p \leq \underline{p}$ and the IS is full if $p \geq \overline{p}$. Then, the set-valued mapping $\mathcal{U} : [0,T] \times \mathcal{X} \to \overline{\mathcal{U}}$ can formally be described for all $t \in [0,T]$ as

$$\mathcal{U}(t, X(t)) = \begin{cases} \{u^O\} & P(t) \geq \overline{p} \quad \text{and} \quad \overline{Q}^M \geq \overline{q} \\ \{u^D, u^W\} & P(t) \geq \overline{p} \quad \text{and} \quad \overline{Q}^M \in [\underline{q}, \overline{q}) \\ \{u^D, u^W, u^F\} & P(t) \in (\underline{p}, \overline{p}) \quad \text{and} \quad \overline{Q}^M \leq \underline{q} \\ \{u^C, u^W, u^F\} & P(t) \in (\underline{p}, \overline{p}) \quad \text{and} \quad \overline{Q}^M \geq \overline{q} \\ \{u^C, u^F\} & P(t) \leq \underline{p} \quad \text{and} \quad \overline{Q}^M \in (\underline{q}, \overline{q}] \\ \{u^F\} & P(t) \leq \underline{p} \quad \text{and} \quad \overline{Q}^M \leq \underline{q} \\ \{u^D, u^W, u^C, u^F\} & \text{else.} \end{cases} \tag{2.20}$$

Interpretation.

- If the IS is full ($P(t) \geq \overline{p}$) we have many cases depending on the state of the GS and the sign of the residual demand.
  - If the current residual demand is negative ($R(t) < 0$) and the GS not full, then the over-production will drive the temperature in the IS above the maximum level $\overline{p}$ if we do nothing. In this case we have to discharge the IS to charge the GS as long as the latter is not full ($\overline{Q}^M(t) < \overline{q}$). However, if GS full ($\overline{Q}^M(t) \geq \overline{q}$), then discharging the IS is not possible and the overproduction will drive the temperature in the IS above the maximum level $\overline{p}$ if we do nothing. In this critical case, we have to apply the so-called over-spilling and the set of feasible controls taking into account the sign of the residual is restricted to $\mathcal{U} = \{u^O\}$.

  - If the residual demand is positive ($R(t) > 0$) and the GS not full ($\overline{Q}^M(t) < \overline{q}$), we can just wait so that the unsatisfied demand drives the temperature in the IS below the maximum level with no cost. During this operation, no additional heat production using fuel is needed, e.g the set of feasible control is restricted to $\mathcal{U} = \{u^D, u^W\}$.

- If $P(t) \in (\underline{p}, \overline{p})$ we have again various cases depending on the state of the GS and the sign of the residual demand.

  - If the residual demand is negative ($R(t) < 0$), then the overproduction will increase the temperature in the internal with no cost. However, if the residual demand is positive ($R(t) > 0$), then the unsatisfied demand will decreases the temperature in the IS. In this case, we can just have to wait as long as the average temperature in the IS remains in the comfort zone.
  - If the GS is empty ($\overline{Q}^M(t) \leq \underline{q}$), we can no longer charge the IS by discharging the geothermal storage. In this case, the set of feasible control is restricted to $\mathcal{U} = \{u^D, u^W, u^F\}$.
  - If the GS is full ($\overline{Q}^M(t) \geq \overline{q}$), we can no longer discharge the IS to charge the GS. In this case, discharging the IS is not possible and the set of feasible control is restricted to $\mathcal{U} = \{u^C, u^W, u^F\}$.

- If the IS is empty, $P(t) \leq \underline{p}$, then we have again various cases depending on the state of the GS and the sign of the residual demand. The most critical cases occur when there is unsatisfied demand.

  - If the GS is not empty ($\overline{Q}^M(t) \in (\underline{q}, \overline{q}]$) and there is unsatisfied demand ($R(t) > 0$), then the unsatisfied demand will drive the temperature below the minimum level if we do nothing. In this case, we have to charge the IS by discharging the GS as long as the latter is not empty or we fire fuel. Discharging the IS or waiting are not possible and the set of feasible controls is restricted to $\mathcal{U} = \{u^C, u^F\}$.
  - If the GS is empty ($\overline{Q}^M(t) \leq \underline{q}$) and there is unsatisfied demand, we have to charge the IS by firing fuel. In this case, charging the IS by discharging the GS or waiting are not possible and the set of feasible control is restricted to $\mathcal{U} = \{u^F\}$.

2.4.3 Performance Criterion

To define the running cost, we introduce notation of heat pump which is used during discharging of the GS to raise the temperature to a pre-specified temperature P_{in}.

Heat pump. A heat pump is a device that transfers heat from a colder area to a warmer area using external energy, such as electricity. Heat energy naturally transfers from warmer to colder spaces. However, a heat pump can reverse this process, by absorbing heat from a cold space and releasing it to a warmer one. This process requires some amount of external energy, such as electricity. It uses external power to accomplish the work of transferring energy from the heat source to the heat sink. Heat pumps are also increasingly used to heat domestic hot water, the hot water used for kitchens, bathrooms, clothes washers, etc. The coefficient of performance (COP) is a measure of a heat pump's efficiency. It is determined by dividing the energy output of the heat pump by the electrical energy needed to run the heat pump, at a specific temperature.

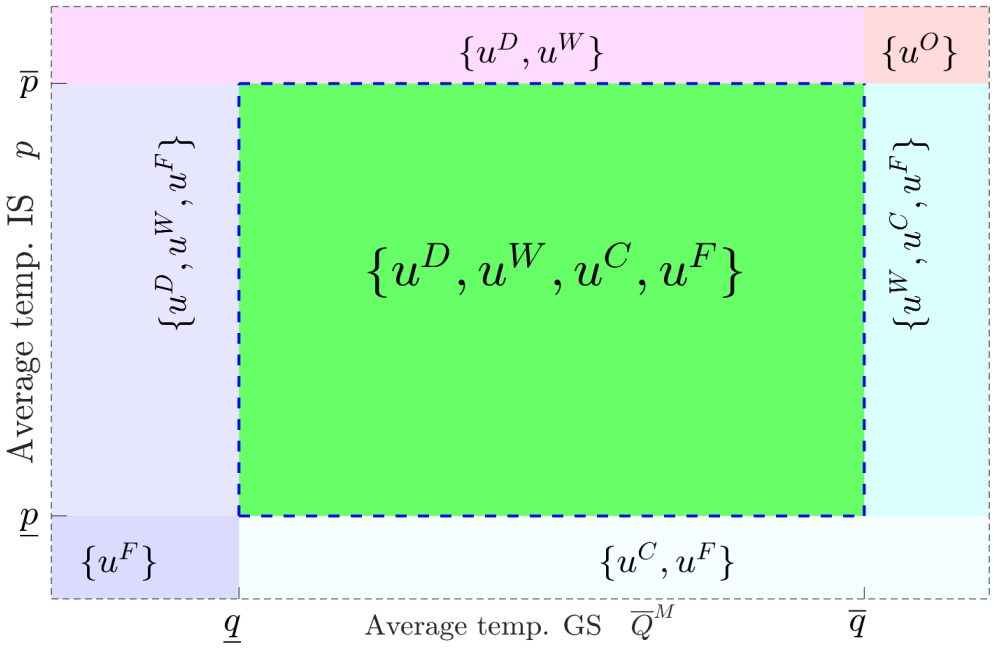


Figure 2.6: Set of feasible controls $\mathcal{U}(t, X(t))$

The higher the COP, the more efficient the heat pump. In electrically-powered heat pumps, the heat transferred can be three or four times larger than the electrical power consumed, giving the system a COP of 3 or 4, as opposed to a COP of 1 for a conventional electrical resistance heater, in which all heat is produced from input electrical energy. Reversible heat pumps work in either direction to provide heating or cooling to the internal space. In our model we use heat pump only to charge the IS by discharging the GS.

Running cost. Let $x = (r, f, p, q)$, where r, f, p and q are the residual demand, the fuel price, the average temperature in the IS and GS, respectively. The running cost contains the

- cost for charging the IS by firing fuel

$$\psi_F(x, \nu) = \begin{cases} k_F f & \nu = u^F, \\ 0 & \text{else;} \end{cases}$$

- cost ψ_D for charging the GS by discharging the IS

$$\psi_D(x, \nu) = \begin{cases} \zeta_D & \nu = u^D, \\ 0 & \text{else;} \end{cases}$$

- cost ψ_C for charging the IS by the discharging the GS

$$\psi_C(x, \nu) = \begin{cases} k_C(P_{in} - \overline{Q}^O) + \zeta_D & \nu = u^C, \\ 0 & \text{else,} \end{cases}$$

where ζ_D $[EUR/h]$ is the cost rate for the power consumption of the ordinary pump, k_F $[EUR/kWh]$ is the cost rate for fuel consumption, and k_C $[EUR/hK]$ is the cost rate for the power consumption of the heat pumps to raise the temperature from $\overline{Q}^O$ to P_{in}. They are all positive constants.

The running reward describing the cost for generating heat and managing such a storage is given for $\nu \in \mathcal{U}$ by

$$\Psi(x,\nu) = \psi_F(x,\nu) + \psi_D(x,\nu) + \psi_C(x,\nu). \tag{2.21}$$

Assumption 2.4.1 We assume that charging the IS by firing fuel is more expensive that discharging the GS (using heat pump) which is more expensive than discharging the IS(using ordinary pump), i.e.,

$$\psi_F > \psi_C > \psi_D.$$

Remark 2.4.2 To relax the strict constraint to P and $\overline{Q}^M$ which is the starting point for simplification, we may also assume for $\nu \in \mathcal{U}$

- a penalty $\psi_P^+(x,\nu)$ if the average temperature in the IS at time t exceeds the maximum level $p > \overline{p}$, with $\psi_P^+(x,\nu) = 0$ for $p < \overline{p}$.
- A penalty $\psi_P^-(x,\nu)$ if the average temperature in the IS at time t exceeds the minimum level $p < \underline{p}$, with $\psi_P^-(x,\nu) = 0$ for $p > \underline{p}$.
- A penalty $\psi_Q^+(x,\nu)$, if the average temperature in the GS at time t exceeds the maximum level $\overline{Q}^M > \overline{q}$, with $\psi_Q^+(x,\nu) = 0$ for $\overline{Q}^M < \overline{q}$.
- A penalty $\psi_Q^-(x,\nu)$ if the average temperature in the GS at time t exceeds the minimum level $\overline{Q}^M < \underline{q}$, with $\psi_Q^-(x,\nu) = 0$ for $\overline{Q}^M > \underline{q}$.

where $\psi_P^+(x,\nu)$, $\psi_P^-(x,\nu)$, $\psi_Q^+(x,\nu)$, $\psi_Q^-(x,\nu)$ are some increasing convex functions.
The running reward describing the cost for generating heat and managing such a storage is then given by

$$\widetilde{\Psi}(x,\nu) = \Psi(x,\nu) + \Psi_{pen}(x,\nu),$$

where Ψ is described in equation (2.21) and Ψ_{pen} by

$$\Psi_{pen}(x,\nu) = \psi_P^-(x,\nu) + \psi_P^+(x,\nu) + \psi_Q^-(x,\nu) + \psi_Q^+(x,\nu).$$

Terminal cost. We also consider a terminal cost depending on the state $X(T)$ at time T given by the function $\phi(X(T))$. This function depends on the storage contract and may include penalties for failing to leave the storage with a pre-specified temperature. Typical examples include:

- Zero cost $\phi(X(T)) = 0$, no penalty and no reward at the terminal time.
- Penalty if at the terminal time the GS or the IS is not filled appropriately, i.e. if the average temperature of the GS $\overline{Q}^M(T)$ is smaller than some pre-specified temperature q_{pen} (usually the initial average temperature in the GS $q_{pen} = \overline{Q}^M(0)$) or if the average temperature of the IS $P(T)$ is smaller than some pre-specified temperature p_{pen} (usually the initial average temperature in the IS storage $p_{pen} = P(0)$). Then a penalty price ζ_{pen}^Q is charged

for every quantity $m^Q c_p^M (q_{pen} - \overline{Q}^M(T))$ of energy needed to raise the temperature from $\overline{Q}^M(T)$ to q_{pen} and there is no reward for the surplus. A penalty price ζ_{pen}^P is charged for every quantity of energy $m^P c_P^F (p_{pen} - P(T))$ and there is no reward for the surplus. Then the terminal cost Φ is of the form

$$\phi(X(T)) = \zeta_{pen}^Q m^Q c_p^M (q_{pen} - \overline{Q}^M(T))^+ + \zeta_{pen}^P m^P c_P^F (p_{pen} - P(T))^+ \qquad (2.22)$$

where $x^+ = \max(x, 0)$.

- Non-negative pay-off (*liquidation of the storage*) obtained by selling all the leftover heat in the GS at some fixed price ζ_{liq}^Q. Then the sales profit is of the form

$$\phi(X(T)) = \zeta_{liq}^Q m^Q c_p^M (\overline{Q}^M(T) - q_{ref}),$$

where $m^Q c_p^M (Q_{av}(T) - q_{ref})$ is the quantity of the leftover heat in the GS. Here, $\overline{Q}^M$ is the average temperature in the GS at the terminal time T and q_{ref} is a reference or pre-specified temperature of the GS. For example, $q_{ref} = \underline{q}$.

Let the control process $u = (u(t))_{t\in[0,T]}$, $u(t) \in \mathcal{U}(t,x)$ be given. The function $J : [0,T] \times \mathcal{K} \times \mathcal{U} \to \mathbb{R}$ defined by

$$J(t,x;u) = \mathbb{E}_{t,x}\left[\int_t^T \Psi(X(t), u(t))dt + \phi(X(T))\right]$$

is the expected aggregated costs over the time interval $[t,T]$, where $\mathbb{E}_{t,x}[\cdot] = \mathbb{E}[\cdot \mid X(t) = x]$ is the conditional expectation given that at time t the state $X(t) = x = (r,f,p,q)^\top \in \mathcal{K}$, Ψ is the the running cost and ϕ is the terminal cost.

2.4.4 Optimal Control Problem

Admissible control. We denote by $\mathcal{A}(x)$ the class of admissible controls, consisting of Markovian control processes u being progressively measurable w.r.t. the filtration $\mathbb{G}$, satisfying certain integrability conditions and control constraints (described above) such that the controlled state X^u takes at any time t values in the prescribed state space $\mathcal{K}$, i.e.,

$$\mathcal{A}(x) = \Big\{ (u(t))_{t\in[0,T]} \mid \begin{array}{l} u \text{ is } \mathbb{G}\text{-progressively measurable}, u(t) = \widetilde{u}(t, X(t)) \text{ for all } t \in [0,T], \\ \widetilde{u}(t,x) \in \mathcal{U}(t,x) \text{ for all } (t,x) \in [0,T] \times \mathcal{X},\ X(t) \in \mathcal{K},\ t \in [0,T], \\ \text{and } \mathbb{E}_{t,x}\left[\int_t^T |\Psi(t, X(t), u(t))|\, dt + |\phi(X(T))|\right] < \infty \Big\}. \end{array}$$

The objective is to minimize the performance criterion (2.4.3) over all admissible controls (2.4.4). We define the value function for all $x \in \mathcal{X}$ by

$$V(x) = \inf_{u\in\mathcal{A}(x)} J(t,x;u).$$

A control $u^* \in \mathcal{A}(x)$ is called optimal control if $V(x) = J(t,x;u^*)$.
Note the dynamics of all states of the control system except the temperature $Q = Q(t,x,y)$ are described by ODEs or SDEs, while Q satisfies the heat equation (2.5) which is a PDE. This is

a non-standard feature and does not fit to the standard framework for stochastic optimal control problems where the state is a multi-dimensional stochastic process described by a system of SDEs (and ODEs). The fact that the state Q follows a PDE makes the optimal control problem much more difficult and challenging. The main idea is to replace PDE (2.5) by a system of ODEs resulting from the semi-discretization w.r.t. spatial variables. However, the discretization approach for solving the heat equation described in Chapter 3 leads to a high-dimensional system of ODEs describing the dynamics of the temperature in the GS. After the discretization we have to further reduce the dimension of the system to make the problem tractable.

In the next chapter we are going to discuss the semi-discretization of the PDE and present some numerical results.

CHAPTER 3

Numerical Analysis of the Dynamics of a Geothermal Storage

Introduction

This chapter aims to investigate the numerical analysis and the simulation of the GS. Such simulations are needed for the optimal control and management of residential heating systems equipped with an underground thermal storage. We work with a 2D-model of a geothermal thermal energy storage, see Fig. 1.2, where a defined volume under or aside of a building is filled with soil and insulated to the surrounding ground. Thermal energy is stored by raising the temperature of the soil inside the storage. It is charged and discharged via heat exchanger pipes filled with some fluid (e.g. water).

In this chapter we focus on the computation of the short-term behavior of the spatial temperature distribution and choose the computational domain to be the storage depicted in Fig. 2.3 by a solid black rectangle. For the sake of simplicity we do not consider the surrounding medium but set appropriate boundary conditions to mimic the interaction between storage and environment. However, we extend the setting in [8, 9] and include heat exchanger pipes for a more realistic model of the storage's charging and discharging process.

We discretize that PDE using finite difference schemes, see Duffy [39]. In a first step we study the semi-discretization with respect to spatial variables leading to a system of linear ODEs. In a second step, we consider full space-time discretization and derive implicit finite-difference schemes.

This chapter provides the following theoretical contributions. First, we prove that the chosen semi-discretization ensures a system of linear ODEs with a stable system matrix. Second, we provide a detailed stability analysis for the implicit finite-difference schemes of the fully discretized PDE and establish a stability condition.

Afterward, we perform extensive numerical experiments, where simulation results for the temporal behavior of the spatial temperature distribution are used to determine how much energy can be stored in or taken from the storage within a given short period of time. Special focus is laid on the dependence of these quantities on the arrangement of the heat exchanger pipes within the storage.

This chapter is organized as follows. In Sec. 3.1 we present the semi-discretization with respect to spatial variables of the initial boundary value problem for that heat equation. For the resulting

system of linear ODEs we show that the system matrix is stable. The full space-time discretization is studied in Sec. 3.2 where we derive implicit finite-difference schemes and provide the associated stability analysis. We end this section by explaining numerical approximation of the aggregated characteristics and by deriving an LTI analogous model of the GS that mimics the most important features of the original non-LTI model of the GS. In Sec. 3.5 we present results of extensive numerical experiments where we use simulations results for the temporal behavior of the spatial temperature distribution to determine how much energy can be stored in or taken from the storage within a given short period of time for the case of horizontal straight pipes. Some technical details of the finite difference scheme which were removed from the main text are provided in Appendix A.

3.1 Semi-Discretization of the Dynamics of a Geothermal Storage

We focus in this section on the spatial discretization of the dynamics of the GS. We recall that the spatial-temporal distribution of the temperature in the GS is described by equation (2.5) together with the boundary and initial conditions and given by

$$\rho c_p \frac{\partial Q}{\partial t} = \nabla \cdot (\kappa \nabla Q) - \rho v \cdot \nabla (c_p Q), \quad (t,x,y) \in (0,T] \times \mathcal{D} \backslash \mathcal{D}^J.$$

For the sake of simplification and tractability of our analysis we restrict ourselves to the following assumption on the arrangement of pipes and impose conditions on the location of grid points along the pipes.

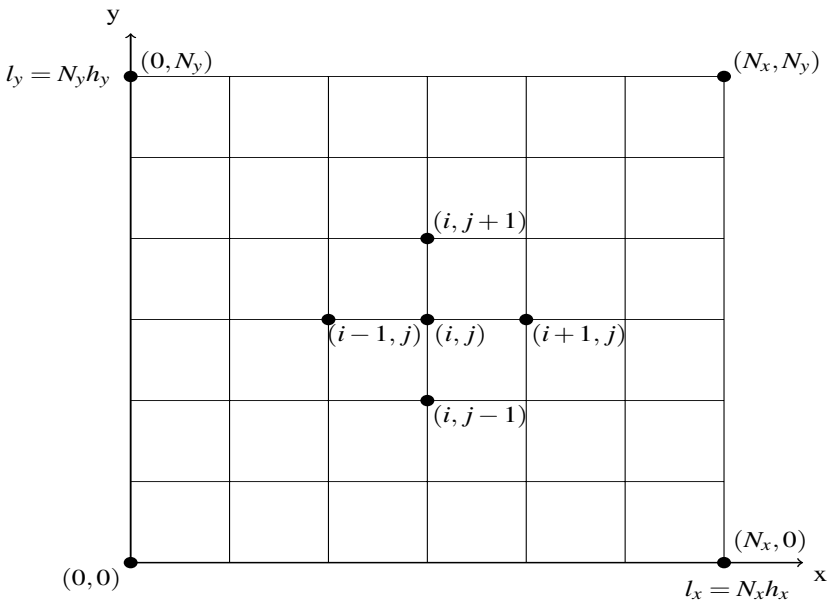


Figure 3.1: Computational grid.

Assumption 3.1.1

1. The interior of pipes contains grid points.

2. Each interface between medium and fluid contains grid points.

3.1.1 Semi-Discretization of the Heat Equation

Let N_x and N_y be the grid size in x-direction and y-direction, respectively, and $h_x = l_x/N_x$ and $h_y = l_y/N_y$ the mesh sizes in x-direction and y-direction, respectively. The spatial domain is discretized by means of a mesh with grid points (x_i, y_j) as in Fig. 3.1 where

$$x_i = ih_x, \quad y_j = jh_y, \quad i = 0,...,N_x, \quad j = 0,...,N_y.$$

We denote by $Q_{ij}(t) \simeq Q(t,x_i,y_j)$ the semi-discrete approximation of the temperature Q and by $v_0(t)(v^x_{ij}, v^y_{ij})^\top = v_0(t)(v^x(x_i,y_j), v^y(x_i,y_j))^\top = v(t,x_i,y_j)$ the velocity vector at time t at the grid point (x_i,y_j). Further, we introduce the following set on indices

$$\begin{aligned}
\mathcal{N}_x &= \{1,...,N_x-1\}, \quad \mathcal{N}_y = \{1,...,N_y-1\},\\
\mathcal{N}^M &= \{(i,j):\ (i,j) \in \mathcal{N}_x \times \mathcal{N}_y \ \text{ with } \ (x_i,y_j) \in \mathcal{D}^M\},\\
\mathcal{N}^F &= \{(i,j):\ (i,j) \in \mathcal{N}_x \times \mathcal{N}_y \ \text{ with } \ (x_i,y_j) \in \mathcal{D}^F\},\\
\mathcal{N}^J &= \{(i,j):\ (i,j) \in \mathcal{N}_x \times \mathcal{N}_y \ \text{ with } \ (x_i,y_j) \in \mathcal{D}^J\},\\
\mathcal{N}^{\mathcal{C}} &= \{(i,j):\ (i,j) \in \{0,...,N_x\} \times \{0,...,N_y\} \ \text{ with } \ (x_i,y_j) \in \partial\mathcal{D}\},
\end{aligned}$$

which we identify with the corresponding sets of grid points. We denote by $\mathcal{N}^S = \mathcal{N}^F \cup \mathcal{N}^M$ the set of grid points in the inner domain $\mathcal{D}^S = \mathcal{D}^F \cup \mathcal{D}^M$, $\mathcal{N}^J = \mathcal{N}^J_L \cup \mathcal{N}^J_U$ the set of grid points on the interface $\mathcal{D}^J = \underline{\mathcal{D}}^J \cup \overline{\mathcal{D}}^J$ between the fluid and medium. Here, $\underline{\mathcal{D}}^J$ and $\overline{\mathcal{D}}^J$ denote the lower and upper interface, respectively, see Fig. 2.4. Further, we decompose the set $\mathcal{N}^{\mathcal{C}}$ of grid points on the boundary domain $\partial\mathcal{D}$ according to the decomposition of $\partial\mathcal{D}$ given in Fig. 2.4 into $\mathcal{N}^{\mathcal{C}} = \mathcal{N}^{\mathcal{C}}_I \cup \mathcal{N}^{\mathcal{C}}_O \cup \mathcal{N}^{\mathcal{C}}_L \cup \mathcal{N}^{\mathcal{C}}_R \cup \mathcal{N}^{\mathcal{C}}_T \cup \mathcal{N}^{\mathcal{C}}_B$. The derivatives in the PDE (2.5) are approximated by linear combinations of values of Q at the grid points (x_i,y_j) in $\mathcal{D}^{fm}$ at time t. We use central second-order finite difference for the diffusion term:

$$\begin{aligned}
\frac{\partial^2 Q(t,x_i,y_j)}{\partial x^2} &= \frac{Q_{i+1,j}(t) - 2Q_{ij}(t) + Q_{i-1,j}(t)}{h_x^2} + \mathcal{O}(h_x^2),\\
\frac{\partial^2 Q(t,x_i,y_j)}{\partial y^2} &= \frac{Q_{i,j+1}(t) - 2Q_{ij}(t) + Q_{i,j-1}(t)}{h_y^2} + \mathcal{O}(h_y^2).
\end{aligned}$$

For the convection term we use the upwind discretization to get

$$\begin{aligned}
v^x(x_i,y_j)\frac{\partial Q(t,x_i,y_j)}{\partial x} &= v^x_{ij}\mathbb{1}_{\{v^x_{ij}>0\}}\frac{Q_{ij}(t) - Q_{i-1,j}(t)}{h_x}\\
&\quad + v^x_{ij}\mathbb{1}_{\{v^x_{ij}<0\}}\frac{Q_{i+1,j}(t) - Q_{ij}(t)}{h_x} + \mathcal{O}(h_x),\\
v^y(x_i,y_j)\frac{\partial Q(t,x_i,y_j)}{\partial y} &= v^y_{ij}\mathbb{1}_{\{v^y_{ij}>0\}}\frac{Q_{ij}(t) - Q_{i,j-1}(t)}{h_y}\\
&\quad + v^y_{ij}\mathbb{1}_{\{v^y_{ij}<0\}}\frac{Q_{i,j+1}(t) - Q_{ij}(t)}{h_y} + \mathcal{O}(h_y).
\end{aligned}$$

We have to point out that above upwind approximations of the convection terms need to be

applied only to the set of grid points $\mathcal{N}^F$ in the fluid domain $\mathcal{D}^F$, since there is no convection outside the fluid and we can set $v_{ij}^x = v_{ij}^y = 0$.

Then for grid points in the domain $\mathcal{D}^S$ the semi-discrete scheme is given by

$$\begin{aligned}\frac{dQ_{ij}(t)}{dt} &= \alpha_{ij}^+(t)Q_{i+1,j}(t) + \alpha_{ij}^-(t)Q_{i-1,j}(t) + \beta_{ij}^+(t)Q_{i,j+1}(t) + \beta_{ij}^-(t)Q_{i,j-1}(t) \\ &\quad + \gamma_{ij}(t)Q_{ij}(t).\end{aligned} \tag{3.1}$$

For grid points $(i,j) \in \mathcal{N}^F$ in the *fluid* domain $\mathcal{D}^f$ the coefficients are given by

$$\begin{aligned}\alpha_{ij}^+(t) &= \alpha^{F+} = \tfrac{a^F}{h_x^2}, \quad \alpha_{ij}^-(t) = \alpha^{F-}(t) = \tfrac{a^F}{h_x^2} + \tfrac{v_0(t)}{h_x}, \quad \beta_{ij}^\pm(t) = \beta^F = \tfrac{a^F}{h_y^2}, \\ \gamma_{ij}(t) &= \gamma^F(t) = -2a^F\left(\tfrac{1}{h_x^2} + \tfrac{1}{h_y^2}\right) - \tfrac{v_0(t)}{h_x}, \qquad a^F = \tfrac{\kappa^F}{\rho^F c_p^F}.\end{aligned} \tag{3.2}$$

In the *medium* domain $\mathcal{D}^m$ the convection terms disappear and the coefficients of the scheme (3.1) become time-independent and are given for $(i,j) \in \mathcal{N}^M$, by

$$\alpha_{ij}^\pm(t) = \alpha^M = \frac{a^M}{h_x^2}, \;\; \beta_{ij}^\pm(t) = \beta^M = \frac{a^M}{h_y^2}, \;\; \gamma_{ij}(t) = \gamma^M = -2a^M\left(\frac{1}{h_x^2} + \frac{1}{h_y^2}\right), \tag{3.3}$$

and $a^M = \frac{\kappa^M}{\rho^M c_p^M}$. Note that for the neighboring grid points to the interfaces we have to slightly modify the above scheme (3.1) due to the extra contribution from the interfaces, see equations (3.8) and (3.9) below.

3.1.2 Semi-Discretization of the Boundary and Interface Conditions

Semi-discretization of the boundary conditions

In this paragraph we consider the discretization of boundary conditions. We start with the homogeneous Neumann conditions (2.7) and (2.10) for the top, left, right and the outlet boundary, where the normal vector $\mathbf{n}$ is equal to $(0,1)^\top, (-1,0)^\top, (1,0)^\top$ and $(1,0)^\top$, respectively. Using first-order differences for the normal derivative we obtain for all $t \in [0,T]$

$$\begin{cases} Q_{iN_y}(t) = Q_{iN_y-1}(t) & \text{for } (i,N_y) \in \mathcal{N}_T^{\mathcal{C}}, \\ Q_{0j}(t) = Q_{1j}(t) & \text{for } l(0,j) \in \mathcal{N}_L^{\mathcal{C}}, \\ Q_{N_x j}(t) = Q_{N_x-1\,j}(t) & \text{for } (N_x,j) \in \mathcal{N}_R^{\mathcal{C}} \cup \mathcal{N}_O^{\mathcal{C}}. \end{cases} \tag{3.4}$$

Next we discretize the Robin condition (2.8) at the bottom boundary $\partial\mathcal{D}^B$. We have $\mathbf{n} = (0,-1)^\top$ such that for all grid points $(i,0) \in \mathcal{N}_B^{\mathcal{C}}$, we have for all $t \in [0,T]$

$$Q_{i0}(t) = \frac{\kappa^M}{\kappa^M + \lambda^G h_y} Q_{i1}(t) + \frac{\lambda^G h_y}{\kappa^M + \lambda^G h_y} Q^G(t). \tag{3.5}$$

On the inlet boundary $\partial\mathcal{D}^I$ we have according to (2.9) a Dirichlet boundary condition during pumping and a Neumann condition if the pump is off. Then for all grid points $(0,j) \in \mathcal{N}_I^{\mathcal{C}}$, we

have $\mathbf{n} = (-1,0)^\top$ which implies for all $t \in [0,T]$

$$\begin{cases} Q_{0j}(t) = Q^I(t) & \text{if pump on,} \\ Q_{0j}(t) = Q_{1j}(t) & \text{if pump off.} \end{cases} \tag{3.6}$$

The relations (3.4) through (3.6) represent linear algebraic equations which allow to express the grid values $Q_{ij}(t)$ in the boundary grid points $(i,j) \in \mathcal{N}^{\mathcal{B}}$ in terms of the corresponding values in the neighbouring points in the interior of the domain and the input data to the boundary conditions. Thus, in the finite difference scheme these values $Q_{ij}(t)$ can be removed from the set of unknowns.

Semi-discretization of interface condition

Now we consider grid points on the interface $\mathcal{D}^J$ between fluid and medium which are by Assumption 2.3.3 straight lines in x-direction. That interface can be decomposed as $\mathcal{D}^J = \underline{\mathcal{D}}^J \cup \overline{\mathcal{D}}^J$, with $\underline{\mathcal{D}}^J$ and $\overline{\mathcal{D}}^J$ representing the lower and upper interface, respectively, see Fig. 3.2. We define the outer normal by $\mathbf{n} = (0,1)^\top$ on the upper interface and by $\mathbf{n} = (0,-1)^\top$ for lower

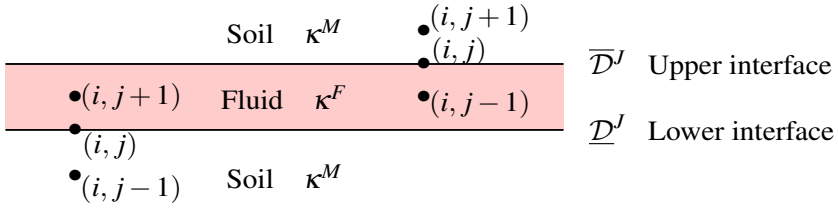


Figure 3.2: Interface between the fluid and soil.

interface. Note that we have n_P pipes and each pipe has two interfaces. Then, we have in total $2n_P$ interface subdomains.

For a grid point (x_i, y_j) on the interface $\mathcal{D}^J$ the perfect contact condition (2.12) implies that at a given time t the temperature of the fluid $Q^f(t,x_i,y_j)$ is equal to the temperature $Q^M(t,x_i,y_j)$ of the medium at that point. As usual, $Q_{ij}(t)$ denotes the semi-discrete approximation of that temperature. Then discretization of the interface condition (2.11) leads to

$$\begin{aligned} \kappa^F \frac{Q_{ij}(t) - Q_{i,j+1}(t)}{h_y} &= \kappa^M \frac{Q_{i,j-1}(t) - Q_{ij}(t)}{h_y} && \text{for lower interface} \\ \kappa^M \frac{Q_{i,j+1}(t) - Q_{ij}(t)}{h_y} &= \kappa^F \frac{Q_{ij}(t) - Q_{i,j-1}(t)}{h_y} && \text{for upper interface.} \end{aligned}$$

We obtain the following coupling between the grid values in an interface grid point $(i,j) \in \mathcal{N}^J$ and its neighbours in vertical direction a time $t \in [0,T]$,

$$\begin{aligned} Q_{ij}(t) &= \psi^F Q_{i,j+1}(t) + \psi^M Q_{i,j-1}(t), \qquad (i,j) \in \mathcal{N}_L^J, \\ Q_{ij}(t) &= \psi^F Q_{i,j-1}(t) + \psi^M Q_{i,j+1}(t), \qquad (i,j) \in \mathcal{N}_U^J, \end{aligned} \tag{3.7}$$

$$\text{where} \quad \psi^F = \frac{\kappa^F}{\kappa^F + \kappa^M} \quad \textit{and} \quad \psi^M = 1 - \psi^F.$$

The above relations show that the grid values $Q_{ij}(t)$ in the interface grid points $(i,j) \in \mathcal{N}^J$ can be expressed as linear combinations of the grid values in the two vertical neighbouring points in the fluid and medium. Thus, in the finite difference scheme these values $Q_{ij}(t)$ can be removed from the set of unknowns. Now, let $(i,j) \in \mathcal{N}_L^J$ be an interface point on the lower interface. Then substituting the above expressions for $Q_{ij}(t)$ into the finite differences scheme (3.1) applied to the lower neighbour $(i,j-1) \in \mathcal{N}^M$ in the medium leads to

$$\frac{d}{dt}Q_{i,j-1}(t) = \alpha^M Q_{i+1,j-1}(t) + \alpha^M Q_{i-1,j-1}(t) + \beta^M Q_{i,j-2}(t) + \beta_I^M Q_{i,j+1}(t) + \gamma_I^M Q_{i,j-1}(t)$$
$$\text{with} \quad \beta_I^M = \psi^F \beta^M \quad \text{and} \quad \gamma_I^M = \gamma + \psi^M \beta^M, \tag{3.8}$$

whereas for the upper neighbour $(i,j+1) \in \mathcal{N}^F$ in the fluid it holds

$$\frac{d}{dt}Q_{i,j+1}(t) = \alpha^{F+} Q_{i+1,j+1}(t) + \alpha^{F-} Q_{i-1,j+1}(t) + \beta^F Q_{i,j+2}(t) + \beta_I^F Q_{i,j-1}(t) + \gamma_I^F Q_{i,j+1}(t)$$
$$\text{with} \quad \beta_I^F = \psi^M \beta^F \quad \text{and} \quad \gamma_I^F = \gamma + \psi^F \beta^F. \tag{3.9}$$

Similar expressions can be derived for points $(i,j) \in \mathcal{N}_U^J$ on the upper interface. We obtain for the neighbour $(i,j-1) \in \mathcal{N}^F$ in the fluid

$$\frac{d}{dt}Q_{i,j-1}(t) = \alpha^{F+} Q_{i+1,j-1}(t) + \alpha^{F-} Q_{i-1,j-1}(t) + \beta^F Q_{i,j-2}(t) + \beta_I^F Q_{i,j+1}(t) + \gamma_I^F Q_{i,j-1}(t)$$
$$\text{with} \quad \beta_I^F = \psi^M \beta^F \quad \text{and} \quad \gamma_I^F = \gamma + \psi^F \beta^F,$$

whereas for the upper neighbour $(i,j+1) \in \mathcal{N}^M$ in the medium it holds

$$\frac{d}{dt}Q_{i,j+1}(t) = \alpha^M Q_{i+1,j+1}(t) + \alpha^M Q_{i-1,j+1}(t) + \beta^M Q_{i,j+2}(t) + \beta_I^M Q_{i,j-1}(t) + \gamma_I^M Q_{i,j+1}(t)$$
$$\text{with} \quad \beta_I^M = \psi^F \beta^M \quad \text{and} \quad \gamma_I^M = \gamma + \psi^M \beta^M,$$

3.1.3 Matrix Form of the Semi-Discrete Scheme

We are now in a position to establish a semi-discretized version of the heat equation (2.6) in terms of a system of ODEs by summarizing relations (3.1), (3.8) and (3.9). To this end we recall that the temperature at the boundary grid points can be obtained by the linear algebraic equations (3.4) through (3.6) derived from the boundary conditions. Further, the values at the interface points are obtained by the interpolation formulas in (3.7) derived from the perfect contact condition. Thus, we can exclude these grid points from the subsequent considerations where we collect the semi-discrete approximations of the temperature $Q(t,x_i,y_j)$ at the remaining points of the grid in the vector function $Y(t) = (Y_1(t), Y_2(t), \ldots, Y_n(t))^T$. The enumeration of the entries of Y is such that we start with the first inner grid point $(1,1)$ next to the lower left corner of the domain. Then we number grid points consecutively in vertical direction where we exclude the $2n_P$ points of the interfaces of the n_P pipes such that we have

$$q = N_y - 2n_P - 1$$

points in each "column" of the grid. Thus, $Y_{(i-1)q+1}$ corresponds to grid point $(i,1)$ for $i = 1,\ldots,N_x-1$, and the last entry Y_n to the inner grid point (N_x-1, N_y-1) next to the domain's

upper right corner. The dimension of Y is

$$n = (N_x - 1)q = (N_x - 1)(N_y - 2n_P - 1).$$

The enumeration described above can be expressed formally by a mapping $\mathcal{K} : \mathcal{N}^S \to \{1, \ldots, n\}$ with $(i, j) \mapsto l = \mathcal{K}(i, j)$ which maps pairs of indices (i, j) of grid point $(x_i, y_j) \in \mathcal{D}$ to the single index l of the corresponding entry in the vector Y.

Using the above notations we can rewrite relations (3.1), (3.8) and (3.9) as the following system of ODEs for the vector function Y representing the semi-discretized heat equation (2.6) together with the given boundary and interface conditions.

$$\frac{dY(t)}{dt} = A(t)Y(t) + B(t)g(t), \;\; t \in (0, T], \tag{3.10}$$

with the initial condition $Y(0) = y_0$. Here, the vector $y_0 \in \mathbb{R}^n$ contains the initial temperatures $Q(0, ., .)$ at the corresponding grid points. The system matrix A results from the spatial discretization of the convection and diffusion term in the heat equation (2.6) together with the Robin and linear heat flux boundary conditions. It has tridiagonal structure consisting of $(N_x - 1) \times (N_x - 1)$ block matrices of dimension q given by

$$A = \begin{pmatrix} A_L & D^+ & & & & 0 \\ D^- & A_M & D^+ & & & \\ & D^- & A_M & D^+ & & \\ & & \ddots & \ddots & \ddots & \\ & & & D^- & A_M & D^+ \\ 0 & & & & D^- & A_R \end{pmatrix}. \tag{3.11}$$

The inner block matrices $A_M, i = 2, \ldots N_x - 2$ of dimension q have tridiagonal structure and are sketched for the case of one pipe in Table 3.1. The matrix entries β^F, γ^F are given in (3.2), β^M, γ^M in (3.3), β_I^M, γ_I^M in (3.8) and β_I^F, γ_I^F in (3.9). The first and last diagonal entries read as $\gamma_D^M = \gamma^M + \frac{\kappa^M}{\kappa^M + \lambda^G h_y}\beta^M$, $\gamma_U^M = \gamma^M + \beta^M$, respectively. They are obtained if the discretized top and bottom boundary conditions (3.4) and (3.5) are substituted into (3.1).

For the matrices A_L and A_R containing entries resulting from the discretization of boundary conditions at the left and right boundary we refer to [110] which is also given in Appendix A.1. The lower and upper block matrices $D^\pm \in \mathbb{R}^{q \times q}$, $i = 1, \ldots, N_x - 1$, are diagonal matrices of the form

$$D^\pm = D^\pm(t) = \mathrm{diag}\,(\alpha^M, \ldots, \alpha^M | \alpha^{F\pm}(t), \ldots, \alpha^{F\pm}(t) | \alpha^M, \ldots, \alpha^M), \tag{3.12}$$

where α^M is given in (3.3) and $\alpha^{F\pm}$ in (3.2). Here, we denote by $|$ the location of the interfaces where we only sketched the case of one pipe. For the convenience of the reader we provide a comprehensive list of all entries of matrix A showing the dependence on model and discretization parameters in Appendix 3.1.4.

The $n \times 2$ input matrix B is a result from the discretization of the inlet and Robin boundary

$$
A_M = \left(\begin{array}{ccccccccccc}
\gamma_D^M & \beta^M & & & & & & & & & \\
\beta^M & \gamma^M & \beta^M & & & & & & & & \\
 & \ddots & \ddots & \ddots & & & & & & & \\
 & & \beta^M & \gamma^M & \beta^M & & & & & & \\
 & & & \beta^M & \gamma_I^M & \beta_I^M & & & & & \\
\hline
 & & & & \beta_I^F & \gamma_I^F & \beta^F & & & & \\
 & & & & & \beta^F & \gamma^F & \beta^F & & & \\
 & & & & & & \ddots & \ddots & \ddots & & \\
 & & & & & & & \beta^F & \gamma^F & \beta^F & \\
 & & & & & & & & \beta^F & \gamma_I^F & \beta_I^F \\
\hline
 & & & & & & & & & \beta_I^m & \gamma_I^M \\
 & & & & & & & & & \beta^M & \gamma^M \\
 & & & & & & & & & & \ddots \\
 & & & & & & & & & & \\
 & & & & & & & & & &
\end{array}\right)
$$

Lower Boundary
Medium
Lower interface
Fluid
Upper interface
Medium
Uppper Boundary

Table 3.1: Sketch of inner block matrices $A_M, i = 2,\dots,N_x-2$ for the case of one pipe

conditions, its entries $B_{lr},\ l = 1,\dots,n,\ \ r = 1,2$, are derived in Appendix A.1 and are given by

$$
\begin{aligned}
B_{l1} &= B_{l1}(t) = \begin{cases} \frac{a^F}{h_x^2} + \frac{\overline{v}_0}{h_x}, & \text{pump on,} \\ 0, & \text{pump off,} \end{cases} && l = \mathcal{K}(1,j),(0,j) \in \mathcal{N}_I^{\mathcal{C}}, \\
B_{l2} &= \frac{\lambda^G h_y}{\kappa^M + \lambda^G h_y}\beta^M, && l = \mathcal{K}(i,1),(i,0) \in \mathcal{N}_B^{\mathcal{C}}.
\end{aligned} \tag{3.13}
$$

The entries for other l are zero. The input function $g:\ [0,T] \to \mathbb{R}^2$ is defined by

$$
g(t) = \begin{cases} (Q^I(t),\ Q^G(t))^\top, & \text{pump on,} \\ (0,\ Q^G(t))^\top, & \text{pump off.} \end{cases} \tag{3.14}
$$

Recall that Q^I is the inlet temperature of the pipe during pumping and Q^G is the underground temperature.

Remark 3.1.2 Eq. (3.10) represents a system of n linear non-autonomous ODEs. Since some of the coefficients in the matrices A,B resulting from the discretization of convection terms in the heat equation (2.6) depend on the velocity $v_0(t)$, it follows that A,B are time-dependent. With some abuse of notation we will write $A = A(t) = A(v_0(t))$ and $B = B(t) = B(v_0(t))$. In Ass. 2.3.3 we assume that $v_0(t)$ is piecewise constant with $v_0(t) = \overline{v}_0$ during charging and discharging when the pump is on whereas $v_0(t) = 0$ during waiting periods when the pump is off. Therefore, the matrices A,B share this property. They take only the two values $A^P = A(\overline{v}_0), B^P = B(\overline{v}_0)$ during pumping and $A^N = A(0), B^N = B(0)$ if the pump is off.

3.1.4 Matrix Stability Analysis

Auxiliary results from matrix analysis

In this subsection we collect some results from matrix analysis taken from the literature. They will be used in the proofs of Theorem 3.1.13 and Lemma 3.2.2. Let $M \in \mathbb{C}^{n\times n}$ be some generic

matrix. For $i = 1, \ldots, n$ we introduce the notations

$$R_i(M) = \sum_{j \neq i} |M_{ij}|, \quad J_i(M) = |M_{ii}| - R_i(M), \quad S_i(M) = |M_{ii}| + R_i(M) = \sum_j |M_{ij}|. \quad (3.15)$$

Note that the maximum norm of M is given by $\|M\|_\infty = \max_i S_i(M)$. The quantities $R_i(M)$ appear as radii of Gershgorin's circles of M and the $J_i(M)$ are used to describe diagonal dominance of M.

Lemma 3.1.3 (Gershgorin's Circle Theorem, Varga [120]) Let $M \in \mathbb{C}^{n\times n}$ and for $i = 1, \ldots, n$ let $D_i = \{z \in \mathbb{C} : |z - M_{ii}| \leq R_i\}$ be the closed discs in the complex plane centred at M_{ii} with radius $R_i = R_i(M)$ given in (3.15). Then all the eigenvalues of M lie in the union of the discs $D_1, \ldots, D_n$.

Definition 3.1.4 (Diagonal Dominance) Row $i \in \{1, \ldots, n\}$ of a matrix $M \in \mathbb{C}^{n\times n}$ is called
strictly diagonal dominant if $J_i(M) > 0$,
weakly diagonal dominant if $J_i(M) \geq 0$,
The matrix M is called strictly (weakly) diagonal dominant if all of its rows are strictly (weakly) diagonal dominant.

The following result says that strictly diagonal dominant matrices are invertible and provides a upper bound for the maximum norm of the inverse.

Lemma 3.1.5 (Varah [118], Theorem 1) Let $M \in \mathbb{C}^{n\times n}$ strictly diagonal dominant matrix. Then M is invertible and

$$\left\|M^{-1}\right\|_\infty \leq \frac{1}{J(M)}, \quad \text{where} \quad J(M) = \min_{1\leq i\leq n} J_i(M).$$

Matrices which are weakly but not strictly diagonal dominant can be singular. A criterion for non-singularity is based on the following property of a matrix and the subsequent lemma. This property was used in Horn and Johnson [61, Definition 6.2.7] and termed *property SC*. In the literature it is also known as *strongly connected*.

Definition 3.1.6 (Strongly Connected Matrix) A matrix $M \in \mathbb{C}^{n\times n}$ is called strongly connected (or of property SC) if for each pair of distinct integers $p, q \in \{1, \ldots, n\}$ there is a sequence of distinct integers $k_1 = p, k_2, ..., k_m = q$ such that each entry $M_{k_1k_2}, M_{k_2k_3}, \ldots, M_{k_{m-1}k_m}$ is non-zero.

For strongly connected matrices Horn and Johnson [61, Corollary 6.2.9] give the following criterion for non-singularity.

Lemma 3.1.7 (Better's Corollary) Suppose that the matrix $M \in \mathbb{C}^{n\times n}$ is strongly connected, weakly diagonally dominant and there exists one strictly diagonal dominant row. Then M is non-singular.

Properties of the system matrix A

We recall that the time-dependence of $A(t)$ is a result of the discretization of convection terms in the heat equation (2.6). The latter depend on the time-dependent velocity $v_0(t)$ for which we assume in Ass. 2.3.3 that $v_0(t)$ is piecewise constant with $v_0(t) = \bar{v}_0$ during charging and

discharging when the pump is on whereas $v_0(t) = 0$ if the pump is off. Therefore, $A(t)$ is also piecewise constant taking only the two values A^P and A^N introduced in Remark 3.1.2. Thus, for studying properties of $A(t)$ on $[0,T]$ or of $A^k = A(k\tau)$ for $k = 0,\dots,N_\tau$ it is sufficient to look at the properties of A^P and A^N.

We want to have a closer look to the entries of the block matrices A_M, A_L, A_R given in Tables 3.1, A.1 and of $D^{\pm}$ given in (3.12), forming the system matrix A. It turns out that for the diagonal entries and the row characteristics R_i, J_i, S_i given in (3.15) one has to distinguish 14 different cases. Instead of n rows it is sufficient to consider only 14 representative rows whose indices we denote by $i_l, l = 1,\dots,14$. Table A.2 provides a list of diagonal entries $A_{i_l i_l}$ and the row characteristics $R_{i_l}(A), J_{i_l}(A), S_{i_l}(A)$ in terms of the model and discretization parameters. For the convenience of the reader we give below that information also for the individual non-diagonal entries of A.

$$\beta^M = \frac{a^M}{h_y^2},\ \beta^F = \frac{a^F}{h_y^2},\ \beta_I^F = \frac{\kappa^F}{\kappa^F+\kappa^M}\beta^M,\quad \beta_I^F = \frac{\kappa^M}{\kappa^F+\kappa^M}\beta^F,$$
$$\alpha^M = \frac{a^M}{h_x^2},\ \alpha^{F+} = \frac{a^F}{h_x^2},\ \alpha^{F-} = \frac{a^F}{h_x^2} + \frac{\overline{v}_0}{h_x}.$$

Lemma 3.1.8 The matrix $A = A(t)$ is weakly diagonal dominant for all $t \in [0,T]$.

Proof. Inspecting the quantities $J_{i_l}(A)$ and Table A.2 it can be seen that it holds $J_{i_l}(A) \geq 0$, hence by Definition 3.1.4 the matrix is diagonal dominant. □
Note that A is weakly but not strictly diagonal dominant since not all of its rows are strictly diagonal dominant.

Lemma 3.1.9 The Gershgorin circles of the matrix $A = A(t)$ are subsets of $\mathbb{C}_- \cup \{0\}$ for all $t \in [0,T]$. Here, $\mathbb{C}_-$ denotes the set of complex numbers with negative real part.

Proof. Let us examine the Gershgorin's circles of A for the 14 different representative rows denoted by $D_{i_l} = D_{i_l}(C_{i_l}, R_{i_l})$ with centres $C_{i_l} = A_{i_l i_l}$ and the radii $R_{i_l}(A)$, $l = 1,\dots,14$, given in Table A.2. Since all entries of A are real, the centres $C_{i_l} = A_{i_l i_l} < 0$ of the discs are on the negative real axis. Lemma 3.1.8 shows that A is diagonal dominant, i.e., $J_{i_l}(A) = |C_{i_l}| - R_{i_l}(A) \geq 0$. Hence, the radii $R_{i_l}(A)$ of the Gershgorin circles never exceed $|C_{i_l}|$ and it holds $D_{i_l} \subset \mathbb{C}_- \cup \{0\}$. □

Lemma 3.1.10 The matrix $A = A(t)$ is strongly connected for all $t \in [0,T]$.

Proof. Let (p,q) be a pair of distinct integers with $p,q \in \{1,\dots,n\}$. Then we can choose the sequence of distinct integers $k_1, k_2, \dots, k_m$, such that $m = |p-q|+1$ and $k_j = p+j-1$ (for $p < q$) and $k_j = p-j+1$ (for $p > q$). It holds $A_{k_j k_{j+1}} \neq 0$ since these entries are located on the upper and lower subdiagonal of A for which we have

$$A_{k_j k_{j+1}} = \begin{cases} \beta^{F/M}, & \text{for} \quad (k_j, k_{j+1}) \in \mathcal{N}^{FM} \setminus \mathcal{N}_n^J, \\ \beta_I^{F/M}, & \text{for} \quad (k_j, k_{j+1}) \in \mathcal{N}_n^J, \end{cases}$$

where $\mathcal{N}^{FM}$ is the set of grid points in the fluid and medium $\mathcal{D}^F \cup \mathcal{D}^M$ and $\mathcal{N}_n^J$ the set of neighbouring grid points to the interface. Since $\beta^{F/M}$ given in (3.2), (3.3) and $\beta_I^{F/M}$ given in (3.8),

(3.9) are positive, we have $A_{k_j k_{j+1}} \neq 0,\ j = 1,2,\ldots,m$. Thus, the matrix A is strongly connected. □

Lemma 3.1.11 The matrix $A = A(t)$ is non-singular for all $t \in [0,T]$.

Proof. From Lemma 3.1.8 and 3.1.10 it is known that $A(t)$ is weakly diagonal dominant and strongly connected for all $t \in [0,T]$. Table A.2 shows that there exist strictly diagonal dominant rows. Hence, Better's Corollary (see Lemma 3.1.7) implies that $A(t)$ is nonsingular. □

Lemma 3.1.12 For maximum norm of the matrix $A = A(t)$ it holds

$$\max_{t\in[0,T]} \|A(t)\|_\infty = \max\left\{\left\|A^P\right\|_\infty,\ \left\|A^N\right\|_\infty\right\} \leq 4\max\{a^F, a^M\}\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right)+\frac{2\bar{v}_0}{h_x}.$$

Proof. $A(t)$ is piecewise constant taking only the two values A^P and A^N. From the last column of Table A.2 showing the 14 different row sums $S_i(A^{P/N})$ of the two matrices it can be easily seen that $\|A^{P/N}\|_\infty \leq \max\{S_{i_6}(A^{P/N}), S_{i_7}(A^{P/N})\}$ yielding the estimate in the lemma. □

Stability of the system matrix

The finite difference semi-discretization of the heat equation (2.6) given by the system of ODEs (3.10) is expected to preserve the dissipativity of the PDE. This property is related to the stability of the system matrix $A = A(t)$ in the sense that all eigenvalues of A lie in left open complex half plane. This property will play a crucial role in Chapter 4, where we study model reduction techniques for (3.10) based on balanced truncation. The next theorem confirms the expectations on the stability of A.

Theorem 3.1.13 (Stability of Matrix A)
Under Assumption 2.3.3 on the model and Assumption 3.1.1 on the discretization, the matrix $A = A(t)$ given in (3.11) is stable for all $t \in [0,T]$, i.e., all eigenvalues $\lambda(A)$ of A lie in the open left half plane.

Proof. Lemma 3.1.9 in subsection 3.1.4 shows using Gershgorin's circle theorem, that the eigenvalues are either located in left open complex half plane or zero. Further, Lemma 3.1.11 (also in subsection 3.1.4) shows that $A(t)$ is non-singular for all $t \in [0,T]$ and thus excludes the case $\lambda(A) = 0$. Thus, for all eigenvalues it holds $\lambda(A) \in \mathbb{C}_-$ and A is stable. □

3.2 Full Discretization of the Model

After discretizing the heat equation (2.6) w.r.t. spatial variables, we are going to discretize the temporal derivative and derive a family of implicit finite difference schemes for which we perform a stability analysis.

3.2.1 Implicit Finite Difference Scheme

We introduce the notation N_τ for the mesh size in the t-direction, $\tau = T/N_\tau$ the time step and $t_k = k\tau,\ k \in \mathcal{N}_\tau = \{0,...,N_\tau\}$. Let A^k, B^k, g^k, v_0^k be the values of $A(t), B(t), g(t), v_0(t)$ at time $t = t_k$. Further, we denote by $Y^k = (Y_1^k, \ldots, Y_n^k)^\top$ the discrete-time approximation of the vector function $Y(t)$ at time $t = t_k$. Recall that Y contains the temperatures $Q = Q(t,x,y)$ at the points of the grid excluding points on the boundary and interface. Discretizing the temporal derivative in (3.10) with the forward difference gives

$$\frac{dY(t_k)}{dt} = \frac{Y^{k+1} - Y^k}{\tau} + \mathcal{O}(\tau). \tag{3.16}$$

Substituting (3.16) into (3.10) and replacing the r.h.s. of (3.10) by a linear combination of the values at time t_k and t_{k+1} with the weight $\theta \in [0,1]$ gives the following general θ-implicit finite difference scheme

$$\frac{Y^{k+1} - Y^k}{\tau} = \theta[A(t_{k+1})Y^{k+1} + B(t_{k+1})g^{k+1}] + (1-\theta)[A(t_k)Y^k + B(t_k)g^k]$$

from which we derive for $k = 0, \ldots, N_\tau - 1$ the recursion

$$\begin{aligned} &G^{k+1}Y^{k+1} = H^k Y^k + \tau F^k \quad \text{where} \\ &G^k = \mathbb{I}_n - \tau\theta A^k,\ H^k = \mathbb{I}_n + \tau(1-\theta)A^k, \text{ and } F^k = \theta B^{k+1}g^{k+1} + (1-\theta)B^k g^k, \end{aligned} \tag{3.17}$$

with the initial value $Y^0 = Y(0)$ and the notation $\mathbb{I}_n$ for the $n \times n$ identity matrix.

The above general θ-implicit scheme leads for $\theta = 0, 1/2$ and 1 to special cases which are known in the literature as forward Euler or fully explicit scheme for $\theta = 0$, Crank-Nicolson scheme for $\theta = 1/2$ and backward Euler or fully implicit scheme for $\theta = 1$. For our numerical experiments in Sec. 3.5 we use an explicit scheme which is obtained for $\theta = 0$ and given by the recursion as

$$Y^{k+1} = (\mathbb{I}_n + \tau A^k)Y^k + \tau B^k g^k, \quad k = 0, \ldots, N_\tau - 1, \tag{3.18}$$

The advantage of an explicit scheme is that it avoids the time-consuming solution of systems of linear equations but on has to take care of the appropriate choice of the time step to ensure stability of the scheme.

3.2.2 Stability Analysis of the Finite Difference Scheme

In this subsection we investigate the stability of the finite difference scheme (3.17) in the maximum norm and present in Theorem 3.2.3 below a stability condition to the time discretization. The use of such stability results is twofold. First it ensures "robustness" w.r.t. round-off errors of the problems's input data, which are the initial condition and the inlet and underground temperature, in the sense that we can run the recursion for an arbitrarily long time without a total loss of accuracy. Second, stability of the scheme is a key ingredient in every analysis of convergence of the solution of the finite difference scheme to the solution of the given initial boundary value problem for the PDE for an infinite refinement of space and time discretization.

Note that a complete convergence analysis is beyond the scope of this paper. In particular, we do not investigate consistency issues. Consistency roughly says that the finite differences

scheme approximates correctly the PDE. The proof of consistency is straightforward and based on Taylor series expansions. We refer to the Lax-Richtmyer Equivalence Theorem, see Sanz-Serna and Palencia [96], Thomas [111, Theorem 2.5.3], stating that a consistent finite difference scheme for a well-posed linear initial boundary value problem, is convergent if and only if it is stable. Hence, for a consistent scheme, convergence is synonymous with stability.

Our stability result is given in terms of maximum norms which are defined for a vector $X \in \mathbb{R}^n$ by $\|X\|_\infty = \max\limits_{1\le i\le n} |X_i|$ and for a square matrix $M \in \mathbb{C}^{n\times n}$ by $\|M\|_\infty = \max\limits_{1\le i\le n} \sum\limits_{j=1}^{n} |M_{ij}|$.

Definition 3.2.1 (Stability of difference scheme in the maximum norm) The finite difference scheme (3.17) is stable in the maximum norm if there exist constants $C_0,\ C_g > 0$ such that

$$\|Y^k\|_\infty \le C_0\|Y^0\|_\infty + C_g \max_{0\le j\le k} \|g^j\|_\infty \ \text{ for } \ k = 1,2,\ldots,\mathcal{N}_\tau. \tag{3.19}$$

We state the following lemma, proved in Appendix A.2, which is useful for the proof of stability theorem.

Lemma 3.2.2 Under Assumption 2.3.3 on the model and Assumption 3.1.1 on the discretization it holds for all $k = 0,\ldots,N_\tau - 1$ and $\theta \in [0,1]$ that

1. the matrices G^{k+1} given in (3.17) are invertible and $\|(G^{k+1})^{-1}\|_\infty \le 1$ with equality for $\theta = 0$;

2. the matrices H^k given in (3.17) satisfy $\|H^k\|_\infty \le 1$ for all $\tau > 0$ if $\theta = 1$;
   and for $\tau \le \frac{1}{(1-\theta)\eta}$ if $\theta \in [0,1)$, where η is given in (3.20);

3. the vectors F^k given in (3.17) satisfy $\|F^k\|_\infty \le C_B \max\limits_{0\le j\le k+1} \left\|g^j\right\|_\infty$ where C_B given in (3.21).

Theorem 3.2.3 (Stability of θ-implicit scheme) Under Assumption 2.3.3 on the model and Assumption 3.1.1 on the discretization it holds

1. For $\theta \in [0,1)$, the semi-implicit finite difference scheme (3.17) is stable if the time step τ satisfies the condition

$$\tau \le \frac{1}{(1-\theta)\eta}, \quad \text{where} \quad \eta = 2\max\{a^F, a^M\}\Big(\frac{1}{h_x^2} + \frac{1}{h_y^2}\Big) + \frac{\overline{v}_0}{h_x}. \tag{3.20}$$

2. For $\theta = 1$, the fully implicit finite difference scheme (3.17) is unconditionally stable, i.e, stable for all $\tau > 0$.

The constants C_0, C_g in (3.19) can be chosen as

$$C_0 = 1 \ \text{ and } \ C_g = C_B T \quad \text{where} \quad C_B = \max\left\{\left\|B^P\right\|_\infty,\ \left\|B^N\right\|_\infty\right\}. \tag{3.21}$$

Proof. of Theorem 3.2.3. From the invertibility of G^k (see Lemma 3.2.2,1.) and the iteration of the recursion (3.17) we obtain for $k = 1,\ldots,\mathcal{N}_\tau$ the explicit representation

$$\begin{aligned} Y^k &= (G^k)^{-1}H^{k-1}Y^{k-1} + \tau(G^k)^{-1}F^{k-1} \\ &= (G^k)^{-1}H^{k-1}(G^{k-1})^{-1}H^{k-2}Y^{k-2} + \tau(G^k)^{-1}H^{k-1}(G^{k-1})^{-1}F^{k-2} + \tau(G^k)^{-1}F^{k-1} \end{aligned}$$

$$= \ldots = \Big(\prod_{j=1}^{k}(G^{k-j+1})^{-1}H^{k-j}\Big)Y^0 + \tau\sum_{j=0}^{k-1}\Big(\prod_{i=1}^{j}(G^{k-i+1})^{-1}H^{k-i}\Big)(G^{k-j})^{-1}F^{k-j-1},$$

where we define $\prod_{j=1}^{0}(\cdot) = \mathbb{I}_n$. Taking the maximum norm on both sides and applying the triangular and Cauchy-Schwarz inequality gives

$$\begin{aligned}\|Y^k\|_\infty \le &\Big(\prod_{j=1}^{k}\|(G^{k-j+1})^{-1}\|_\infty\|H^{k-j}\|_\infty\Big)\|Y^0\|_\infty \\ &+\tau\sum_{j=0}^{k-1}\Big(\prod_{i=1}^{j}\|(G^{k-i+1})^{-1}\|_\infty\|H^{k-i}\|_\infty\Big)\|(G^{k-j})^{-1}\|_\infty\|F^{k-j-1}\|_\infty.\end{aligned}$$

Substituting the estimates for $\|(G^k)^{-1}\|_\infty, \|H^k\|_\infty$ and $\|F^k\|_\infty$ given in Lemma 3.2.2 into the above inequality yields

$$\|Y^k\|_\infty \le \|Y^0\|_\infty + \tau k C_B \max_{0\le j\le k}\|g^j\|_\infty \le \|Y^0\|_\infty + C_B T \max_{0\le j\le k}\|g^j\|_\infty,$$

where we used $\tau k \le \tau N_\tau = T$. According to the second assertion of Lemma 3.2.2 the above estimate holds for all $\tau > 0$ if $\theta = 1$ and for $\tau \le \frac{1}{(1-\theta)\eta}$ if $\theta \in [0,1)$. □

This theorem tells us that the global scheme is only conditionally stable, i.e. the stability is conditioned by the choice of the time step $\tau = \tau(h_x, h_y, v_0)$. Note that as the velocity $v_0(t)$ increases the time step decreases.

Example 3.2.4 Consider the following parameters given in Table 3.2. Mesh sizes $h_x = 10^{-1}\ m$ and $h_y = 10^{-2}\ m$, the thermal diffusivity $a^M = 9.9375\times 10^{-7}\ m^2/s$, $a^F = 1.4558\times 10^{-7}\ m^2/s$ and the velocity $v_0 = 10^{-2}\ m/s$ ($v = 0$ for pure diffusion case). Then

$$\eta^{-1} = \left[2\times 9.94643\times 10^{-7}\left(\frac{1}{10^{-2}}+\frac{1}{10^{-4}}+\right)+\frac{10^{-2}}{10^{-1}}\right]^{-1} = 8.395\ (\eta^{-1} = 52.307,\ v_0 = 0)$$

and the time step can be chosen for $\theta \in [0,1)$ as follows

$$0 \le \tau \le \frac{8.395}{(1-\theta)},\ \ v_0 \ne 0 \qquad \text{or} \quad 0 \le \tau \le \frac{52.307}{(1-\theta)},\ \ v_0 = 0, \quad .$$

This result can be interpreted as follows. Given the above data, for $\theta = 0$ (fully explicit scheme), the choice $\tau > 8.395\ s$ (rep. $\tau > 52.307\ s$, for $v_0 = 0$) of the time step will make the explicit difference scheme unstable. Note that for $\theta \to 1$, $\tau \ge 0$. Therefore, there is no restriction for the time step, hence, the fully implicit scheme is unconditionally stable.

3.3 Numerical Computation of the Aggregated Characteristics

In this subsection we consider the approximate computation of aggregated characteristics introduced in the previous subsections by using finite difference approximations of the temperature

$Q = Q(t,x,y)$. The approximations are given in terms of the entries of the vector function $Y(t)$ satisfying the system of ODEs (3.10) and containing the semi-discrete finite difference approximations of the temperature in the inner grid points of the computational domain $\mathcal{D}$. Recall that the temperatures on boundary and interface grid points can be determined by linear combinations from the entries of $Y(t)$. The extension to approximations based of the solution of the fully discretized PDE (3.18) is straightforward using the relation $Y(t_k) = Y(k\tau) \approx Y^k, k = 0,\dots,N_\tau$.

Let us start with the average temperatures $\overline{Q}^M$ and $\overline{Q}^F$, where the temperature $Q(t,x,y)$ is averaged over unions of disjoint rectangular subsets of the computational domain $\mathcal{D}$. Assume that $\mathcal{B} \subset \mathcal{D}$ is a generic rectangular subset with corners defined by the grid points (x_i, y_j) with indices $(\underline{i},\underline{j}),(\bar{i},\underline{j}),(\bar{i},\bar{j}),(\underline{i},\bar{j})$, where $0 \le \underline{i} < \bar{i} \le N_x$ and $0 \le \underline{j} < \bar{j} \le N_y$. We assume further that the domain $\mathcal{B}$ contains at least one layer of horizontal and vertical inner grid points, respectively. Thus we require $\underline{i}+2 \le \bar{i}$ and $\underline{j}+2 \le \bar{j}$. We denote by $\overline{Q}^{\mathcal{B}} = \overline{Q}^{\mathcal{B}}(t) = \frac{1}{|\mathcal{B}|}\iint_{\mathcal{B}} Q(t,x,y)dxdy$ the average temperature in $\mathcal{B}$. Rewriting the double integral as two iterated single integrals and applying trapezoidal rule to the single integrals the average temperature $\overline{Q}^{\mathcal{B}}$ can be approximated by (for details see Appendix A.3)

$$\overline{Q}^{\mathcal{B}} = \frac{1}{|\mathcal{B}|}\iint_{\mathcal{B}} Q(t,x,y)dxdy \approx \sum_{(i,j)\in\mathcal{N}_{\mathcal{B}}} \mu_{ij} Q_{ij}, \tag{3.22}$$

where $\mathcal{N}_{\mathcal{B}} = \{(i,j) : i = \underline{i},\dots,\bar{i}, j = \underline{j},\dots,\bar{j}\}$ and the coefficients d_{ij} of the above quadrature formula are given by

$$\mu_{ij} = \frac{1}{(\bar{i}-\underline{i})(\bar{j}-\underline{j})} \begin{cases} 1, & \text{for } \underline{i} < i < \bar{i},\ \underline{j} < j < \bar{j}, & \text{(inner points)} \\ \frac{1}{2}, & \text{for } i = \underline{i},\bar{i},\ \underline{j} < j < \bar{j}, & \text{(boundary points, except corners)} \\ & \quad j = \underline{j},\bar{j},\ \underline{i} < i < \bar{i}, & \\ \frac{1}{4}, & \text{for } i = \underline{i},\bar{i},\ j = \underline{j},\bar{j} & \text{(corner points).} \end{cases} \tag{3.23}$$

Next we want to rewrite approximation (3.22) in terms of the vector $Y = Y(t)$. Recall that Y contains the finite difference approximations of the temperature in the inner grid points of the computational domain $\mathcal{D}$. Let us introduce the vector $\overline{Y}$ of dimension $\overline{n} = (N_x+1)(N_y+1) - n$ containing the temperature approximations at the remaining grid points located on the boundary $\partial\mathcal{D}$ and the interface $\mathcal{D}^J$. These values can be determined by the discretized boundary and interface conditions and expressed as linear combinations of the entries of Y. This allows for a representation $\overline{Y} = \overline{C}Y$ with some $\overline{n}\times n-$matrix $\overline{C}$.

Now, let $\mathcal{N}_{\mathcal{B}}^0 \subset \mathcal{N}_{\mathcal{B}}$ and $\overline{\mathcal{N}_{\mathcal{B}}^0} = \mathcal{N}_{\mathcal{B}} \setminus \mathcal{N}_{\mathcal{B}}^0$ be the subsets (of index pairs $(i,j) \in \mathcal{N}_{\mathcal{B}}$ of grid points) for which the finite difference approximation Q_{ij} is contained in the vector Y and the vector $\overline{Y}$, respectively. Further, let $\mathcal{K} : \mathcal{N}_{\mathcal{B}}^0 \to \{1,\dots,n\}$ and $\overline{\mathcal{K}} : \overline{\mathcal{N}_{\mathcal{B}}^0} \to \{1,\dots,\overline{n}\}$ denote the mappings $(i,j) \mapsto l = \mathcal{K}(i,j)$ and $(i,j) \mapsto \bar{l} = \overline{\mathcal{K}}(i,j)$ of pairs of indices (i,j) to the single indices l and $\bar{l}$ of the corresponding entries in the vectors Y and $\overline{Y}$, respectively. Then it holds

$$Q_{ij} = \begin{cases} Y_{\mathcal{K}(i,j)}, & (i,j) \in \mathcal{N}_{\mathcal{B}}^0, \\ \overline{Y}_{\overline{\mathcal{K}}(i,j)}, & (i,j) \in \overline{\mathcal{N}_{\mathcal{B}}^0}, \end{cases}$$

and we can rewrite approximation (3.22) as

$$\begin{aligned} \overline{Q}^{\mathcal{B}} &\approx \sum_{(i,j)\in\mathcal{N}_{\mathcal{B}}^0} \mu_{ij} Q_{ij} + \sum_{(i,j)\in\overline{\mathcal{N}_{\mathcal{B}}^0}} \mu_{ij} Q_{ij} \\ &= \sum_{l=\mathcal{K}(i,j):(i,j)\in\mathcal{N}_{\mathcal{B}}^0} d_l Y_l + \sum_{\bar{l}=\overline{\mathcal{K}}(i,j):(i,j)\in\overline{\mathcal{N}_{\mathcal{B}}^0}} \overline{d}_{\bar{l}}\, \overline{Y}_{\bar{l}} \\ &= DY + \overline{D}\overline{Y}, \end{aligned} \tag{3.24}$$

with an $1\times n-$matrix D and an $1\times\overline{n}-$matrix $\overline{D}$, whose entries are given for $l=1,\dots,n$, $\bar{l}=1,\dots,\overline{n}$ by

$$d_l = \begin{cases} \mu_{ij}, & l=\mathcal{K}(i,j),\ (i,j)\in\mathcal{N}_{\mathcal{B}}^0, \\ 0 & \text{else,} \end{cases} \quad \text{and} \quad \overline{d}_{\bar{l}} = \begin{cases} \mu_{ij}, & \bar{l}=\overline{\mathcal{K}}(i,j),\ (i,j)\in\overline{\mathcal{N}_{\mathcal{B}}^0}, \\ 0 & \text{else,} \end{cases} \tag{3.25}$$

respectively. Finally, substituting $\overline{Y}=\overline{C}Y$ into (3.24) yields a representation of the average temperature $\overline{Q}^{\mathcal{B}}$ as a linear combination of entries of the vector Y which reads as

$$\overline{Q}^{\mathcal{B}} \approx C^{\mathcal{B}} Y \quad \text{with} \quad C^{\mathcal{B}} = D + \overline{D}\,\overline{C}. \tag{3.26}$$

Based on the above representation we can derive similar approximations for the average temperatures $\overline{Q}^{M}$ and $\overline{Q}^{F}$ in the medium and the fluid, respectively. Our model assumptions imply that for a storage with n_P pipes the domain $\mathcal{D}^F$ splits into n_P disjoint rectangular subsets $\mathcal{D}_j^F$, $j=1,\dots,n_P$ (pipes), whereas $\mathcal{D}^M$ consists of n_P+1 of such subsets between the pipes and the top and bottom boundary of $\mathcal{D}$ which we denote by $\mathcal{D}_j^M$, $j=0,\dots,n_P$. Then we can apply (3.22) to derive the approximation

$$\overline{Q}^{F} = \frac{1}{|\mathcal{D}^F|}\sum_{j=1}^{n_P} |\mathcal{D}_j^F|\,\overline{Q}^{\mathcal{D}_j^F} \approx C^F Y \quad \text{where} \quad C^F = \frac{1}{|\mathcal{D}^F|}\sum_{j=1}^{n_P} |\mathcal{D}_j^F|\, C^{\mathcal{D}_j^F}. \tag{3.27}$$

An approximation of the form $\overline{Q}^{M}\approx C^M Y$ can be obtained analogously. In the next subsection we derive approximations $\overline{Q}^{O}\approx C^O Y$ and $\overline{Q}^{B}\approx C^B Y$ for the average temperatures at the outlet and the bottom boundary, respectively. Here, the line integrals in the definitions (2.14) and (2.15) of these two characteristics are approximated by trapezoidal rule.

Numerical approximation of $\overline{Q}^{O}$ and $\overline{Q}^{B}$

Now we consider the average temperatures $\overline{Q}^{O}$ and $\overline{Q}^{B}$ where the temperature $Q(t,x,y)$ is averaged over one-dimensional curves at the boundary $\partial\mathcal{D}$. Assume that $\mathcal{C}\subset\partial\mathcal{D}$ is a generic curve on one of the four outer boundaries. For the ease of exposition we restrict $\mathcal{C}$ to be a line between the grid points $(x_{\underline{i}},y_0)$ and $(x_{\bar{i}},y_0)$ at the bottom boundary, where $0\le\underline{i},\,\underline{i}+2\le\bar{i}\le N_x$. We denote by $\overline{Q}^{\mathcal{C}}=\overline{Q}^{\mathcal{C}}(t)=\frac{1}{|\mathcal{C}|}\int_{\mathcal{C}} Q(t,x,y)\,ds$ the average temperature in $\mathcal{C}$. Applying trapezoidal rule to the line integral we obtain (suppressing the time variable t)

$$\int_{\mathcal{C}} Q(x,y)\,ds = \int_{x_{\underline{i}}}^{x_{\bar{i}}} Q(x,y_0)\,dx \approx h_x\left(\frac{1}{2}Q(x_{\underline{i}},y_0) + \sum_{i=\underline{i}+1}^{\bar{i}-1} Q(x_i,y_0) + \frac{1}{2}Q(x_{\bar{i}},y_0)\right).$$

Since the length of the curve $\mathcal{C}$ is given by $(\bar{i}-\underline{i})h_x$ the average temperature $\overline{Q}^{\mathcal{C}}$ can be approximated by

$$\overline{Q}^{\mathcal{C}} = \frac{1}{|\mathcal{C}|}\int_{\mathcal{C}} Q(t,x,y)\,ds \approx \sum_{(i,j)\in\mathcal{N}_{\mathcal{C}}} \mu_{ij}\,Q_{ij}, \tag{3.28}$$

where $\mathcal{N}_{\mathcal{C}} = \{(i,j) : i = \underline{i},\ldots,\bar{i}, j = 0\}$ and the coefficients μ_{ij} of the above quadrature formula are given by

$$\mu_{ij} = \frac{1}{(\bar{i}-\underline{i})}\begin{cases} 1, & \text{for } \ \underline{i} < i < \bar{i}, \ \ j = 0, \quad \text{(inner points)} \\ \frac{1}{2}, & \text{for } \ i = \underline{i}, \bar{i}, \qquad\quad \text{(end points).} \end{cases}$$

Using the same notation and approach as above we can rewrite approximation (3.28) as

$$\overline{Q}^{\mathcal{C}} \approx \sum_{(i,j)\in\mathcal{N}_{\mathcal{C}}^0} \mu_{ij}\,Q_{ij} + \sum_{(i,j)\in\overline{\mathcal{N}_{\mathcal{C}}}^0} \mu_{ij}\,Q_{ij} = DY + \overline{D}\,\overline{Y}, \tag{3.29}$$

where the matrices D and $\overline{D}$ are defined as in (3.25) with $\mathcal{N}_{\mathcal{B}}^0$ and $\overline{\mathcal{N}_{\mathcal{B}}^0}$ replaced by $\mathcal{N}_{\mathcal{C}}^0$ and $\overline{\mathcal{N}_{\mathcal{C}}}^0$, respectively. Note that in our finite difference scheme the grid values of boundary points are not contained in Y. Thus, we have $\mathcal{N}_{\mathcal{C}}^0 = \varnothing$ and $D = 0_{1\times n}$. Finally, substituting $\overline{Y} = \overline{C}Y$ into (3.29) yields a representation of the average temperature $\overline{Q}^{\mathcal{C}}$ as a linear combination of entries of the vector Y which reads as

$$\overline{Q}^{\mathcal{C}} \approx C^{\mathcal{C}}\,Y \quad \text{with} \quad C^{\mathcal{C}} = D + \overline{D}\,\overline{C}. \tag{3.30}$$

For $\mathcal{C} = \partial\mathcal{D}^B$, i.e., $\underline{i} = 0, \bar{i} = N_x$ the above representation directly gives the approximation of $\overline{Q}^B = C^{\partial\mathcal{D}^B}Y$. For the average temperature $\overline{Q}^O$ at the outlet of a storage with n_P pipes the outlet boundary $\mathcal{D}^O$ splits into n_P disjoint curves $\mathcal{D}_j^O, j = 1,\ldots,n_P$. Then we can apply (3.30) to derive the approximation

$$\overline{Q}^O = \frac{1}{|\partial\mathcal{D}^O|}\sum_{j=1}^{n_P} |\partial\mathcal{D}_j^O|\,\overline{Q}^{\partial\mathcal{D}_j^O} \approx C^O Y \quad \text{where} \quad C^O = \frac{1}{|\partial\mathcal{D}^O|}\sum_{j=1}^{n_P} |\partial\mathcal{D}_j^O|\,C^{\mathcal{D}_j^O}.$$

3.4 Analogous Linear Time-Invariant System

The topic considered this subsection is the starting point of the next chapter in which we aim to approximate the dynamics of certain aggregated characteristics for the infinite dimensional spatial distribution of the temperature $Q = Q(t,x,y)$ by a low-dimensional system of ODEs. The key idea of the analogous model has already been presented in Subsec.2.3.2. Here we provide further details and specify the average temperature in the PHX in terms of the vector function Y containing the temperatures in the grid points. Recall that the dynamics of the spatial distribution of Q is governed by the heat equation (2.6). We semi-discretize the PDE to obtained a finite-dimensional approximation (3.10) which reads as $\frac{dY(t)}{dt} = A(t)Y(t) + B(t)g(t)$ and constitutes a high-dimensional system of ODEs for the vector function Y containing the temperatures in the grid points. In Chapter 4 that system of ODEs is the starting point for the application of model reduction techniques to find a suitable low-dimensional system of ODEs from which the aggregated characteristics can be obtained with a reasonable degree of accuracy.

Eq.(3.10) represents a system of n linear non-autonomous ODEs. Since some of the coefficients in the matrices A, B resulting from the discretization of convection terms in the heat equation (2.6) depend on the velocity $v_0(t)$, it follows that A, B are time-dependent. Thus, (3.10) does not constitute a linear time-invariant (LTI) system. The latter is a crucial assumption for most of model reduction methods such as the Lyapunov balanced truncation technique that is considered in Chapter 4. We circumvent this problem by replacing the model for the GS by a so-called *analogous model* which is LTI. Following key idea presented in Subsc. 2.3.2, we obtain the following approximation of the original by an analogous model which is performed in two steps.

Approximation Step 1. We assume that contrary to the original model the fluid is also moving with constant velocity $\bar{v}_0$ during pump-off periods. During these waiting periods we assume that the temperature Q^I at the pipe's inlet is equal to the average temperature of the fluid in the pipe $\overline{Q}^F$. This approximation leads to a modified boundary condition at the inlet given by (2.16). Semi-discretization of the boundary condition (2.16) using approximation (3.27) of the average fluid temperature $\overline{Q}^F(t) = C^F Y(t)$ leads to a modification of the input term $g(t)$ of the system of ODEs (3.10) given in (3.14). That input term now reads as

$$g(t) = \begin{cases} (Q^I(t),\ Q^G(t))^\top, & \text{pump on,} \\ (C^F Y(t),\ \ Q^G(t))^\top, & \text{pump off.} \end{cases} \tag{3.31}$$

Further, the non-zero entries B_{l1} of the input matrix B given in (3.13) are modified. They are now no longer time-dependent but given by the constant $B_{l1} = \frac{a^F}{h_x^2} + \frac{\bar{v}_0}{h_x}$ which was already used during pump-on periods.

Approximation Step 2. From (3.31) it can be seen that the input term g during pumping depends on the state vector Y via $C^F Y$ and can no longer considered as exogenous. This has to be corrected and leads to an additional contribution to the system matrix A given by $B_{\bullet 1} C^F$ where $B_{\bullet 1}$ denotes the first column of B. Thus, the system matrix again would be time-dependent and the system not LTI. In order to obtain a LTI system we therefore perform a second approximation step and treat $\overline{Q}^F$ as an exogenously given quantity (such as Q_C^I, Q_D^I, Q_G). This leads to a tractable approach for model reduction by the Lyapunov balanced truncation technique applied in the next chapter to generates low-dimensional systems depending only on the system matrices A, B but not on the input term g. Further, from an algorithmic or implementation point of view this is not a problem since given the solution Y of (3.10) at time t, the average fluid temperature $\overline{Q}^F(t)$ can be computed as a linear combination of the entries of $Y(t)$.

3.5 Numerical Results

In this section we present results of numerical experiments based on the finite difference discretization (3.18) of the heat equation (2.6) and determine the spatio-temporal temperature distribution in the storage. Further, we study the impact of the heat exchanger pipe topology and vary the number and arrangement of the pipes. In Subsecs. 3.5.1, 3.5.2 and 3.5.3 we present results for a storage with one, two, and three pipes, respectively. For these experiments we also compute and compare certain aggregated characteristics which are introduced in Sec. 2.3.3 and computed via post processing of the temperature distribution.

Note that we provide additional video material showing animations of the temporal evolution of the spatial temperature distribution for which in the following we can present snapshots

only.
The videos are available at `www.b-tu.de/owncloud/s/D68fmqXRcgbesKj`.

Parameters		Values	Units
Geometry			
width	l_x	10	m
height	l_y	1	m
depth	l_z	10	m
diameter of pipe	d_P	0.02	m
number of pipes	n_P	$1,2,3$	
Material			
medium (dry soil)			
mass density	ρ^M	2000	kg/m^3
specific heat capacity	c_p^m	800	$J/kg\,K$
thermal conductivity	κ^M	1.59	$W/m\,K$
thermal diffusivity $\kappa^M(\rho^M c_p^m)^{-1}$	a^M	9.9375×10^{-7}	m^2/s
fluid (water)			
mass density	ρ^F	997	kg/m^3
specific heat capacity	c_p^F	4182	$J/kg\,K$
thermal conductivity	κ^F	0.607	$W/m\,K$
thermal diffusivity $\kappa^F(\rho^F c_p^F)^{-1}$	a^F	1.4558×10^{-7}	m^2/s
velocity during pumping	$\bar{v}_0$	10^{-2}	m/s
heat transfer coeff. to underground	λ^G	10	$W/(m^2\,K)$
initial temperature	Q_0	10 and 35	°C
inlet temperature: charging	Q_C^I	40	°C
discharging	Q_D^I	5	°C
underground temperature	Q^G	15	°C
Discretization			
mesh size	h_x	0.1	m
mesh size	h_y	0.01	m
time step	τ	1	s
time horizon	T	36 and 72	h

Table 3.2: Model and discretization parameters

Experimental settings

The model and discretization parameters are given in Table 3.2. We run the experiments for $\theta = 0$ (fully explicit scheme). The storage is charged and discharged via heat exchanger pipes filled with a moving fluid and thermal energy is stored by raising the temperature of the storage medium. We recall the open architecture of the storage which is only insulated at the top and the side but not at the bottom. This leads to an additional heat transfer to the underground for which we assume a constant temperature of $Q^G(t) = 15$ °C. In the simulations the fluid is assumed to be water while the storage medium is dry soil. During charging a pump moves the fluid with constant velocity $\bar{v}_0$ arriving with constant temperature $Q^I(t) = Q_C^I = 40$ °C at the inlet. If this temperature is higher than in the vicinity of the pipes, then a heat flux into the storage

medium is induced. During discharging the inlet temperature is $Q^I(t) = Q_D^I = 5$ °C leading to a cooling of the storage. At the outlet we impose a vanishing diffusive heat flux, i.e. during pumping there is only a convective heat flux. In some experiments we also consider waiting periods where the pump is off. This helps to mitigate saturation effects in the vicinity of the pipes reducing the injection and extraction efficiency. During that waiting periods the injected heat (cold) can propagate to other regions of the storage. Since pumps are off we have only diffusive propagation of heat in the storage and the transfer over the bottom boundary.

3.5.1 Storage With one Horizontal Straight PHX

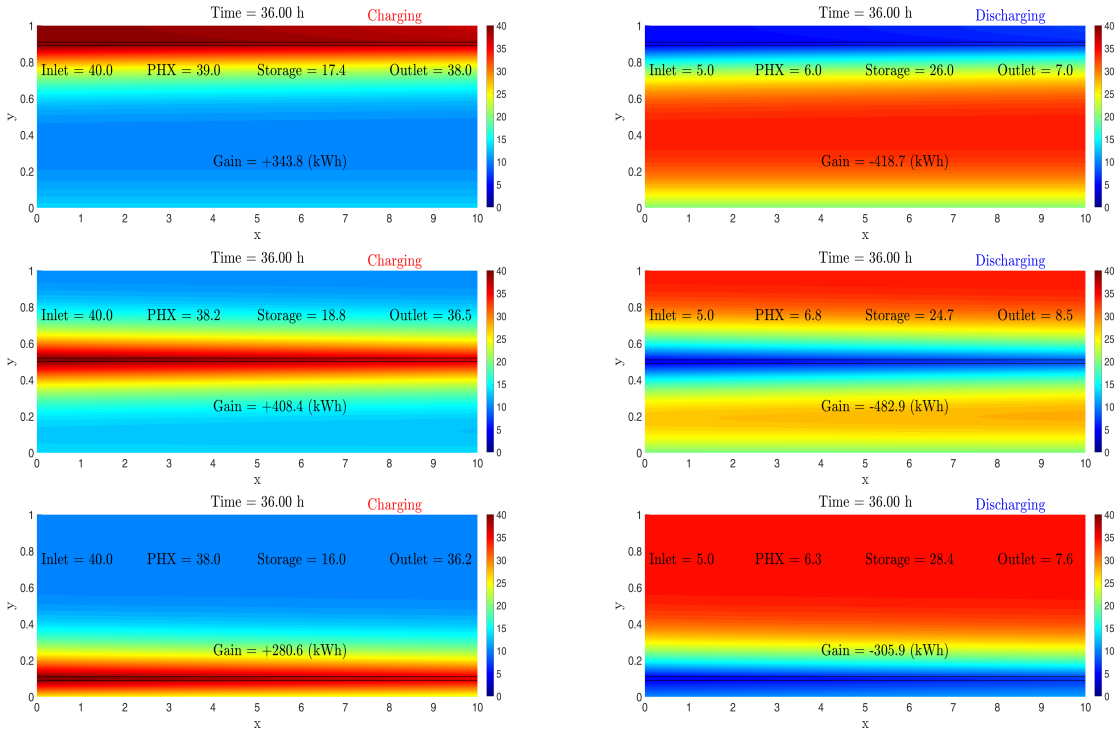


Figure 3.3: Spatial distribution of the temperature in the storage with one horizontal PHX at vertical position p after of 36 hours of charging (left) and discharging (right).
Top: $p = 90$ *cm*. Middle: $p = 50$ *cm*. Bottom: $p = 10$ *cm*.

In this experiment we run simulations with one horizontal pipe located at different vertical positions p between the bottom ($p = 0$ *cm*) and the top ($p = l_y = 100$ *cm*) of the storage. We compare the spatial temperature distributions as well as aggregated characteristics such as the average temperature $\overline{Q}(t)$, the average outlet temperature $\overline{Q}^O(t)$, and the gain or loss of energy $G(0,T)$ in the storage during a period of $T = 36$ hours. Charging is realized by sending fluid through the pipe for 36 hours. It arrives at the inlet with constant temperature $Q_C^I(t) = 40$ °C. We start with a homogeneous initial temperature distribution with $Q(0,x,y) = 10$ °C, uniformly distributed in the storage. In the experiment with discharging we start with a homogeneous initial temperature distribution with 35 °C. For 36 hours the storage is cooled by the moving fluid arriving at the storage inlet with constant temperature $Q_D^I(t) = 5$ °C.

Fig. 3.3 shows the spatial distribution of the temperature in the storage after 36 hours of charging (left) and discharging (right) where we used three different vertical positions p of the pipe. In the top panels the pipe is located close to the insulated top boundary ($p = 90$ *cm*). The

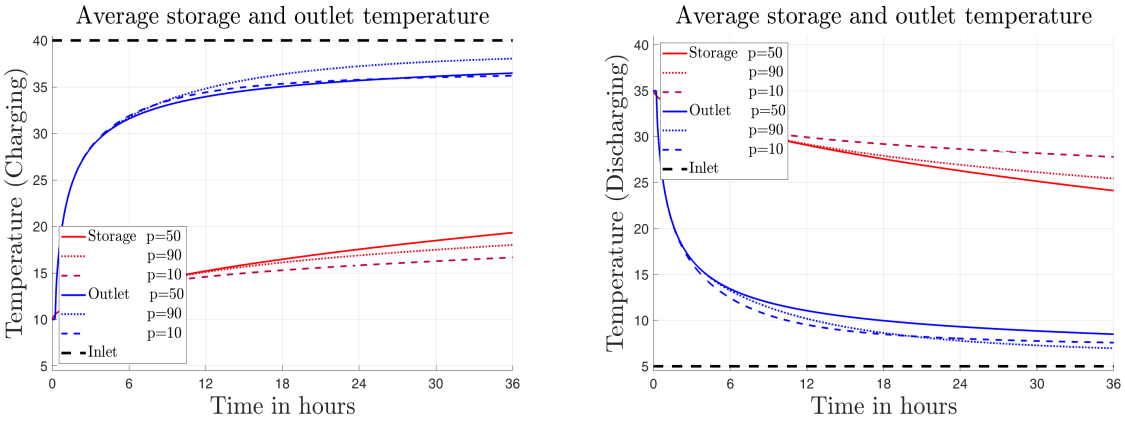


Figure 3.4: Average temperature in the storage $\overline{Q}^S$ and average outlet temperature $\overline{Q}^O$ after 36 hours for a storage with one horizontal PHX at different vertical positions. Left: Charging. Right: Discharging.

panels in the middle show the results for a pipe in the center ($p = 50\ cm$) while in the bottom panels the pipe is close to the bottom boundary ($p = 10\ cm$). Recall that the bottom is open and allows for heat transfer to the underground with constant temperature $Q_G(t) = 15\ ^\circ$C. Fig. 3.4 plots the corresponding average temperatures in the storage and at the outlet against time. In Fig. 3.3 it can be seen that warming and cooling mainly takes places in a vicinity of the pipe and after 36 hours the temperature in more distant storage domains is only slightly changed. Due to the direction of the moving fluid from left to right, warming and cooling in the left part of the storage is slightly stronger than in the right part. A closer inspection of the results shows that except in the experiment with the pipe close to the bottom boundary ($p = 10\ cm$), after 36 hours of charging the temperatures in the vicinity of that boundary are below the underground temperature $Q^G = 15\ ^\circ$C. Thus in addition to the injection of heat via the pipe we also have an inflow of thermal energy from the warmer underground into the storage. This results in a "boundary layer" which is slightly warmer than in the inner storage region. The reverse effect can be observed during discharging where close to the bottom boundary the temperature is always above $Q^G = 15\ ^\circ$C. This induces a heat flux from the storage to the colder underground which contributes together with the extraction of heat via the pipe to the total loss of thermal energy in the storage.

In Fig. 3.5 we plot in the upper panels the gain G^S (respectively loss $-G^S$) of thermal energy during 36 hours of charging (respectively discharging) against time for vertical positions $p = 10,\ 20, \ldots, 90\ cm$. The lower panel shows the gain and the loss at the end of the 36 hours charging and discharging period, respectively, depending on the vertical pipe position p. In the first 4 hours of charging there are almost no visible deviations in the gains and losses, but after 36 hours we can see a clear dependence of the pipe's vertical position p. Further, for all p we observe a decaying slope of the curves in the upper plots. This can be explained by the "thermal saturation" in the vicinity of the pipe and the slow diffusive propagation of the heat to the more distant regions of the storage. It shows that (dis)charging the storage becomes less efficient after longer periods of operation. Injecting (extracting) a certain amount of energy takes longer and needs more electricity consumed by the pumps. This effect suggests to interrupt (dis)charging and include waiting periods in which the heat (cold) in the vicinity of the pipes can propagate to other regions of the storage. The impact of such waiting periods will be studied in more detail in Subsec. 3.5.2.

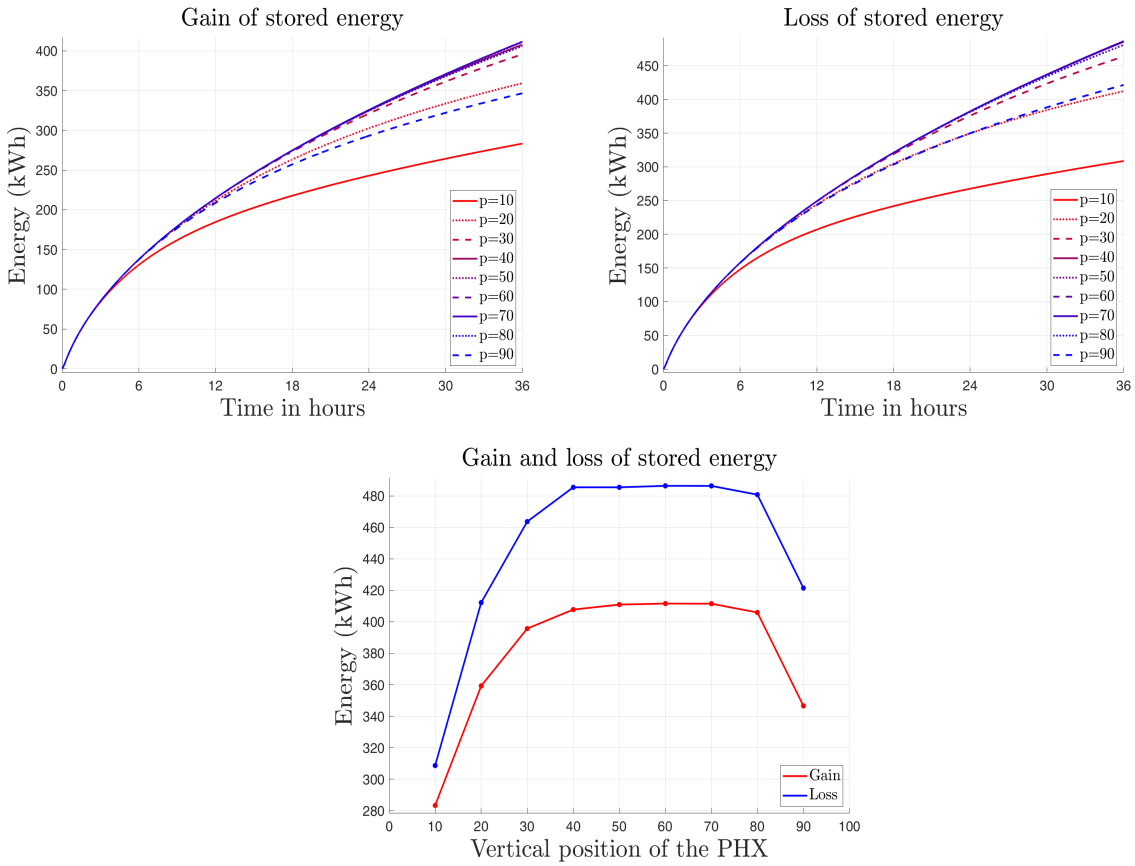


Figure 3.5: Gain and loss of stored energy for a storage with one horizontal PHX at different vertical positions.
Top left: Gain of stored energy G^S during charging. Top right: Loss of stored energy $-G^S$ during discharging.
Bottom: Gain G^S and loss $-G^S$ of stored energy after 36 hours of charging and discharging, respectively, depending on vertical PHX position p.

The results for $p = 40, \ldots, 70$ *cm* are quite similar. However, for pipe locations close to the open bottom boundary ($p = 10,\ 20$ *cm*) and the insulated top boundary ($p = 90$ *cm*) we observe remarkable deviations. Here charging and discharging is considerably slower and gains and losses of thermal energy are smaller. For a pipe close to the top this can be explained by the saturation of the storage domain in the vicinity of the pipe. During charging (discharging) the boundary and its insulation prevent the propagation of heat into (from) the inner storage regions. On the bottom boundary that effect is combined with heat transfer to the underground. During charging a part of the injected heat is lost to the underground while during discharging the vicinity of the pipe is also cooled by the colder underground. Thus as expected, for an efficient operation of the storage the pipe should be located in the central region of the storage.

3.5.2 Storage With Two Horizontal Straight PHX*s*

In this experiment we run the simulations with two horizontal pipes located symmetrically to the vertical mid level of $p = 50$ *cm* and separated by a distance d varying between 10 *cm* and 90 *cm*. Recall that placing a single pipe at $p = 50$ *cm* showed quite good performance in the last subsection. First we study the spatial temperature distribution and some aggregated char-

acteristics during (dis)charging for $T = 36$ hours. Then we introduce waiting periods allowing the injected heat (cold) to spread within the storage.

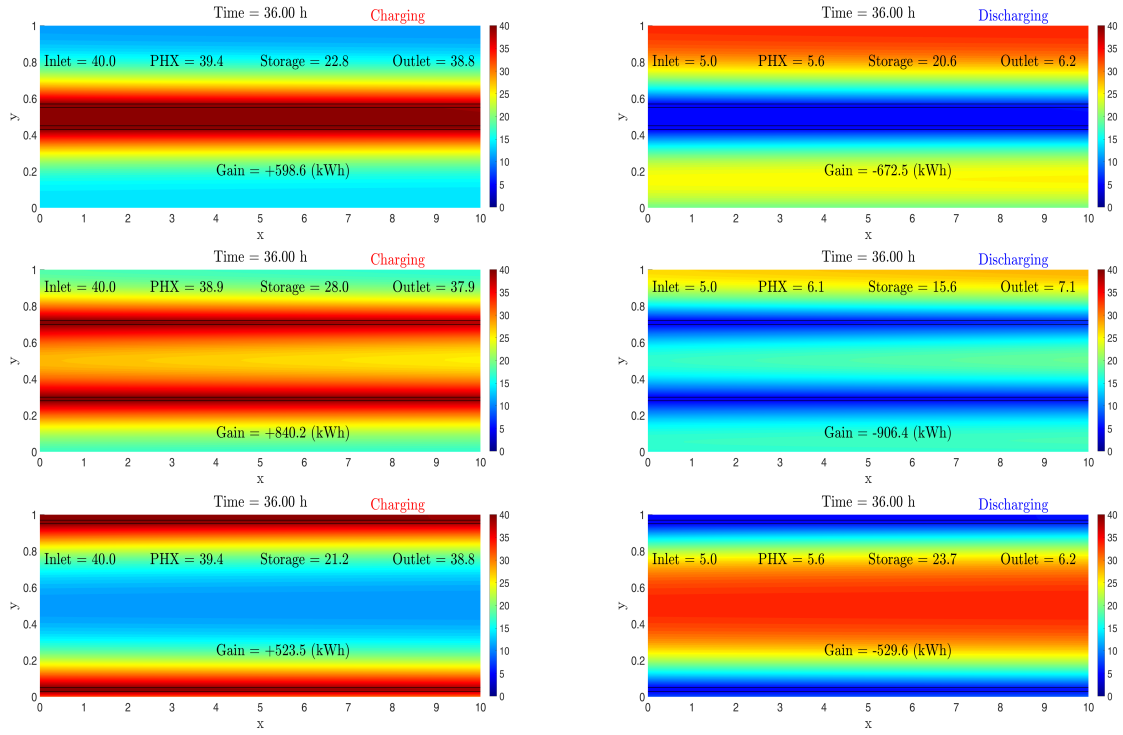


Figure 3.6: Spatial distribution of the temperature in the storage with two horizontal PHX*s* of vertical distance d after 36 hours of charging (left) and discharging (right).
Top: $d = 10$ *cm*. Middle: $d = 40$ *cm*. Bottom: $d = 90$ *cm*.

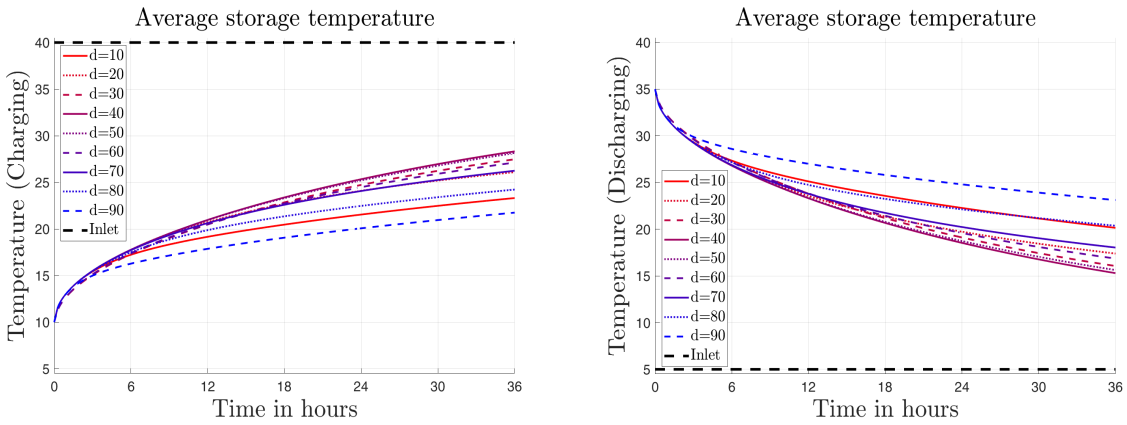


Figure 3.7: Average temperature in the storage $\overline{Q}^S$ during 36 hours for a storage with two horizontal PHX*s* of different vertical distances. Left: Charging. Right: Discharging.

Charging and discharging without waiting periods

Fig. 3.6 shows for three different distances d of the two pipes the spatial distribution of the temperature in the storage after 36 h of charging (left) and discharging (right). In the top panels

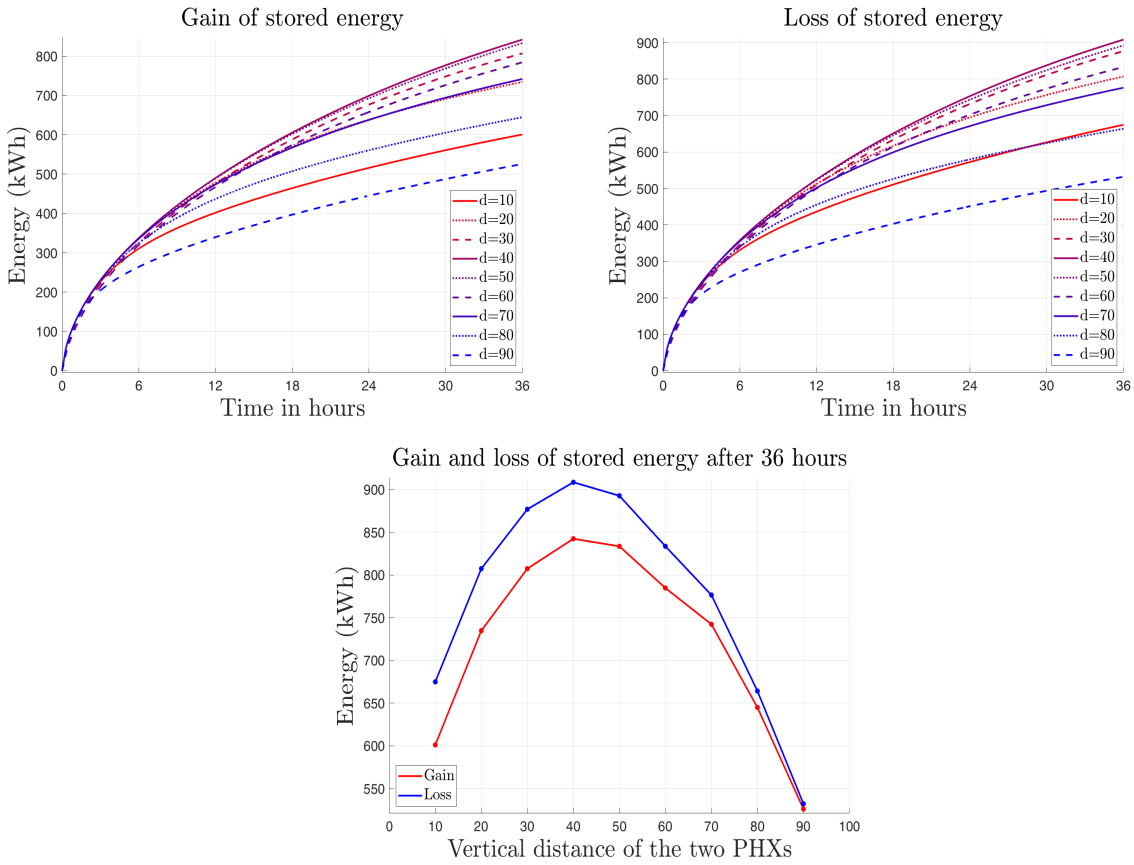


Figure 3.8: Gain and loss of stored energy for a storage with two horizontal PHX*s* of different distance d. to 90 cm.
Top left: Gain G^S of stored energy during charging. Top right: Loss $-G^S$ of stored energy during discharging.
Bottom: Gain G^S and loss $-G^S$ of stored energy after 36 hours of charging and discharging, respectively, depending on distance d.

the pipes are very close ($d = 10$ cm). The panels in the middle show the results for two pipes at a distance $d = 40$ cm and in the bottom panels one pipe is located close to the top and the other close to the bottom boundary ($d = 90$ cm). As in the experiment with only one pipe it can be seen that warming and cooling in the left part of the storage is slightly stronger than in the right part. It mainly takes places in a vicinity of the pipe whereas after 36 h temperatures in more distant regions are only slightly changed. Thus, the spatial temperature distributions differ considerably for the three arrangements of two pipes. For a small distance ($d = 10$ cm), we observe a strong saturation at a level close to the inlet temperature in the small region between the pipes while the region at the top is almost at the initial temperature and the region at the bottom is only slightly warmed (cooled) by the underground. For the pipes at distance $d = 90$ cm, we observe an extreme saturation in the small layer between the upper pipe and the top boundary while the lower pipe is also warming (cooling) the underground.

Next we will have a look at aggregated characteristics. In Fig. 3.7 the average temperatures are plotted against time for distances of the pipes $d = 10, 20, \ldots, 90$ cm. Fig. 3.8 presents the gain and loss of thermal energy in the storage at the end of the charging and discharging period, respectively. The figures reveal that apart from the first 4 hours there is a strong impact

of the pipe distance. The most efficient mode of operation is obtained for the pipes distance of $d = 40\ cm$. Here, the gain (loss) of thermal energy during charging (discharging) is at maximum. These quantities strongly decay for smaller and larger distances because of the saturation effect which becomes stronger if pipes are arranged closer to each other or closer to the top and bottom boundary of the storage.

Charging and discharging with waiting periods

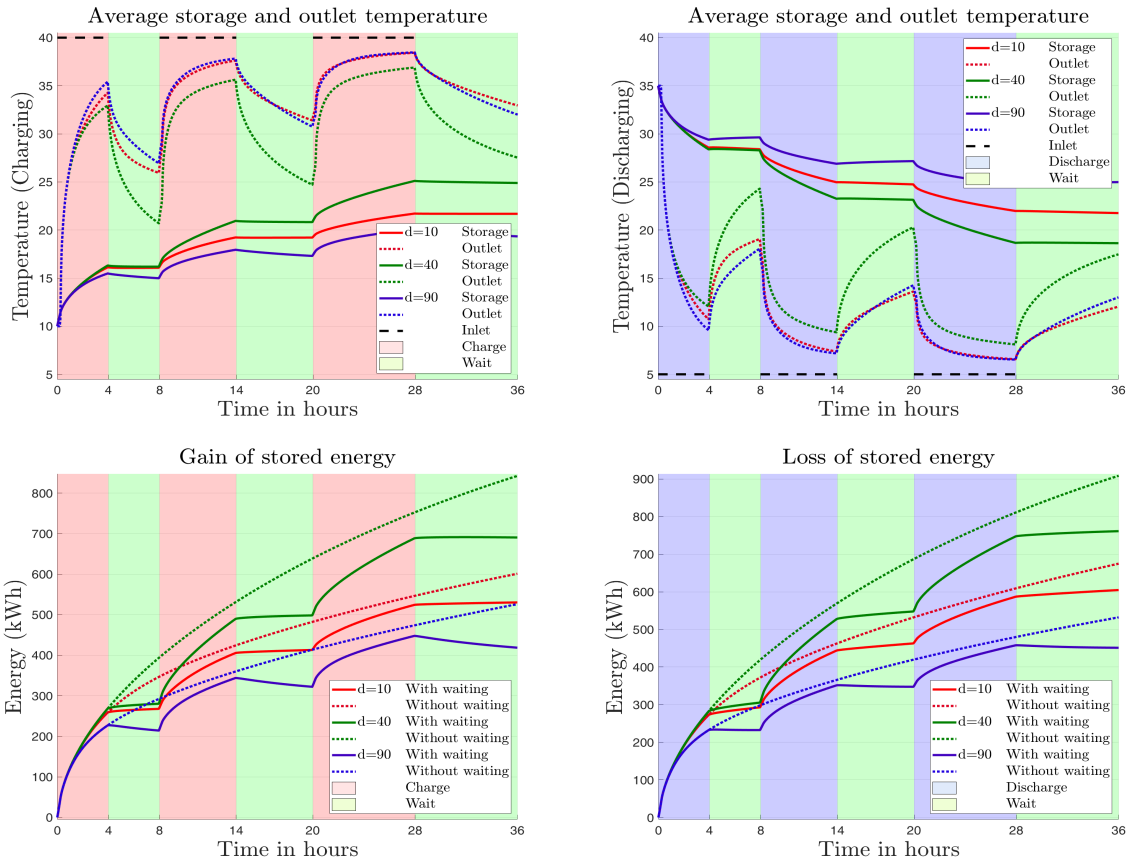


Figure 3.9: Charging and discharging during 36 hours with several waiting periods for a storage with two horizontal PHX*s* at distance $d = 10\ cm$, $d = 40$, and $d = 90\ cm$.
Top: Aggregated characteristics $\overline{Q}^S$ and $\overline{Q}^O$. Bottom: Gain G^S / loss $-G^S$ of stored energy.
Left: Charging. Right: Discharging.

The above experiments have shown how saturation effects can be mitigated by choosing an appropriate vertical distance of the two pipes. This option is only available in the design of the storage architecture and not during the operation of an already existing storage. Therefore, we now want to examine another option, which is the interruption of (dis)charging cycles allowing the heat injected to (extracted from) the vicinity of the pipes to propagate to the other storage regions. The idea is that after a sufficiently long waiting period the saturation in the vicinity of the pipe is considerably reduced such that (dis)charging can resumed with higher efficiency. Although, the introduction of such waiting period will increase the time needed to inject (extract) a given amount of thermal energy it reduces the saturation effect and helps to save operational costs for electricity used for running the pumps.

In our experiments we divide the time interval $[0,T]$ into three sub-intervals of length $8, 12, 16$ hours. In each sub-interval (dis)charging is followed by a waiting period of the same length as it can be seen in Fig. 3.9 where charging, waiting and discharging periods are represented by red, green and blue background color. The top panels show the average temperatures in the storage $\overline{Q}$ and at the outlet $\overline{Q}^O$, respectively, during charging and discharging. We compare a storage with two pipes of distance $d = 40$ *cm* and a storage with more close-by pipes $d = 10$ *cm* and two pipes at distance $d = 90$ *cm*. Recall that in the previous subsection we have seen that $d = 40$ *cm* allows for much more efficient operation than for $d = 10,\ 90$ *cm*. As expected, during the waiting periods the average temperatures at the outlet and in the pipe decay after charging and rise after discharging. This is due to the diffusion of heat in the storage, in particular the heat flux induced by the different temperatures inside and outside the pipe. During waiting the average temperature in the storage $\overline{Q}$ is almost constant since injection or extraction of heat is stopped. However, the heat transfer to and from the underground at the bottom boundary continues also during waiting but it does not produce a visible change of $\overline{Q}$. In the two lower panels of Fig. 3.9 we compare the storage operation with and without waiting periods. We plot the gain (loss) of thermal energy in the storage during charging (discharging) over time. Note that for operation with waiting (dis)charging takes place only 50% of the time. However, for the "optimal" pipe distance $d = 40$ *cm* the resulting gain (loss) reaches more than 80% of the values for uninterrupted operation. For the less efficient cases of pipes at distance $d = 10$ *cm* and pipes at distance $d = 90$ *cm* that cause strong saturation effects the differences are smaller and the gaps are quickly reduced to almost zero after resuming (dis)charging.

3.5.3 Storage With Three Horizontal Straight PHX*s*

In this example we add a third pipe to the storage architecture and study two different pipe arrangements. We proceed with the experimental design including the same waiting periods considered in the previous subsection but now we "glue" together the two periods of charging and discharging each of length 36 *h*. The result is a total period of length $T = 72$ *h* starting with a storage at temperature $Q(0,x,y) = 10$ °C. Within the the first 36 hours the storage is charged by the moving fluid arriving at the pipe inlet with temperature $Q_C^I(t) = 40$ °C and then discharged using the inlet temperature $Q_D^I(t) = 5$ °C. The charging, waiting and discharging periods can be seen in Fig. 3.11. Contrary to the above experiments, discharging now starts not with a temperature 35 °C but with a non-uniformly temperature distribution which is obtained after $36h$ of charging (and waiting). In this more realistic setting, temperatures typically are higher in the vicinity of the pipes and lower in other regions.

Fig. 3.10 shows snapshots of the spatial temperature distribution during the last charging period (at $t = 27h$), during the subsequent waiting period (at $t = 35h$) and during of the last discharging period (at $t = 63h$), respectively. We compare two storage architectures with three pipes. In the first, the pipes are located symmetrically w.r.t. the vertical mid level. For the second, the central pipe was moved upwards such that we get a non-symmetric arrangement with two quite close-by pipes in the upper region. The snapshots show a strong saturation between the two upper pipes of the non-symmetric pipe arrangement while for symmetric pipes the temperature distribution is much more uniform, in particular during the waiting period as it can be seen in the middle panel for time $t = 35$ *h*.

In Fig. 3.11 we present aggregated characteristics which are plotted over time and observe similar patterns as in the experiment with a two-pipe storage considered in the previous subsection. During the waiting periods after charging the average outlet and pipe temperatures decay

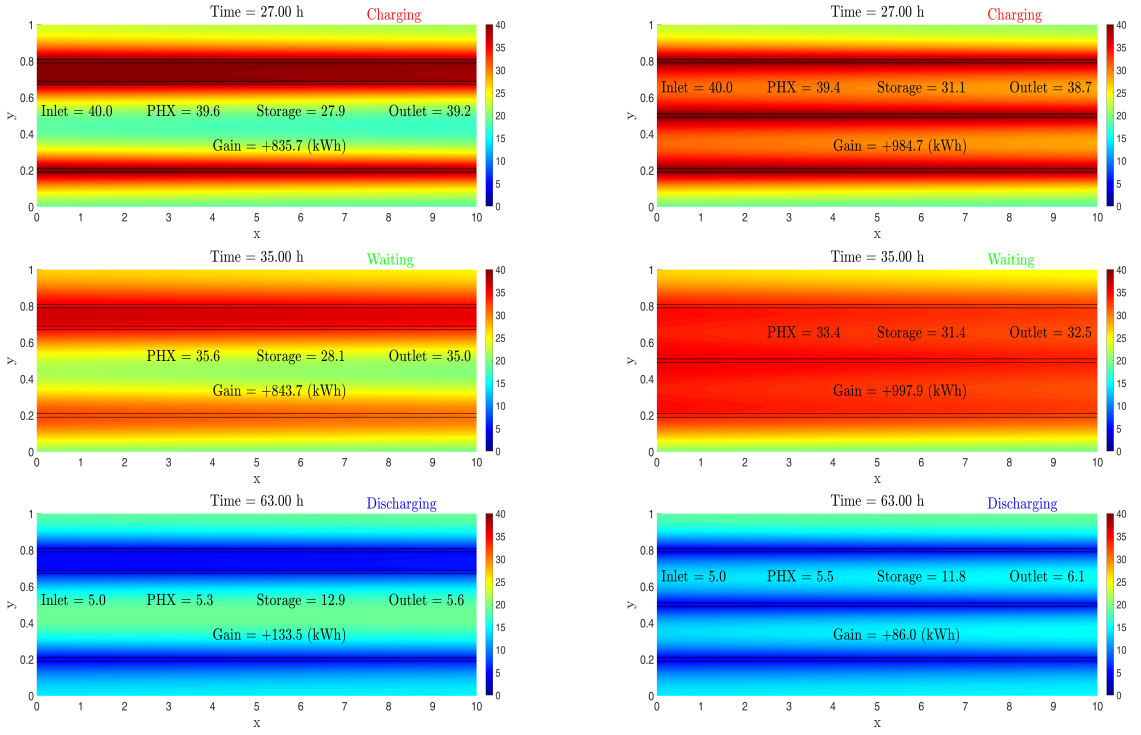


Figure 3.10: Spatial distribution of the temperature in the storage with three horizontal PHXs during charging (top), waiting (middle) and discharging (bottom) period.
Left: Non-symmetric PHXs, Right: Symmetric PHXs

at a faster rate for symmetric pipes than for non-symmetric pipes and vice versa for waiting periods after discharging. This is a consequence of the stronger saturation for non-symmetric pipes which prohibits a faster cooling (warming) of the pipe during waiting. For symmetric pipes the average storage temperature during charging increases faster and during discharging decreases faster than for non-symmetric pipes. This explains the similar patterns for the gain of stored energy which are plotted in the right panel. It shows that the storage with symmetric pipes (dis)charges faster than the storage with non-symmetric pipes.

3.5.4 Numerical Results for Analogous LTI System

In Figs. 3.12 and 3.13 we present some numerical results where we compare the spatio-temporal temperature distribution and its aggregated characteristics of the original and the associated analogous model. These results are based on the experimental design in the Subsec. 3.5.3 for a storage architecture with three symmetric pipes and waiting periods. Fig. 3.12 compares snapshots of the spatial temperature distribution in the storage for the original and analogous model. One snapshot is taken during charging and the other at the end of the last waiting period after preceding discharging periods. At first glance there are no visible differences. A look at the aggregated characteristics in Fig. 3.13 shows negligible approximation errors for the average temperature in the storage and the fluid. However, the approximation of the average outlet temperature suffers slightly from the replacement of a resting fluid by a moving fluid during the waiting period. The resulting “mixing of the temperature profile” inside the pipe adjusts the outlet to the average temperature in the pipe. This can be seen in the right panel where the relative error for the outlet temperature dominates the errors for the two average storage

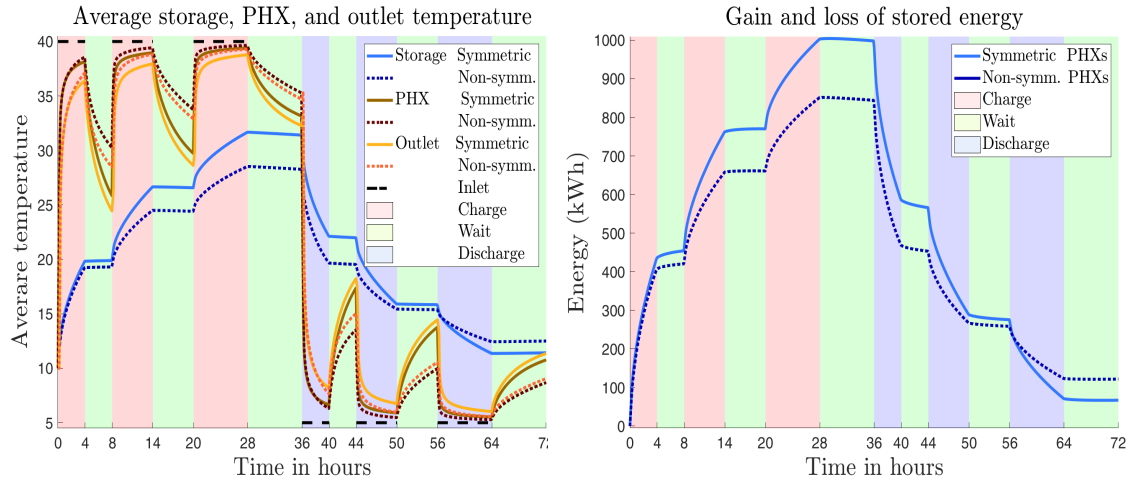


Figure 3.11: Storage with three horizontal PHX*s* during 72 hours with charging, waiting and discharging periods. Left: Aggregated characteristics $\overline{Q}^S, \overline{Q}^F, \overline{Q}^O$. Right: Gain of stored energy G^S.

and fluid temperature. The experiment indicates that apart from some noticeable approximation errors in the pipe during waiting periods, in particular at the outlet, the other deviations are negligible. Finally, it can be nicely seen that during the (dis)charging periods the errors decrease and vanish almost completely, i.e., in the long run there is no accumulation of errors.

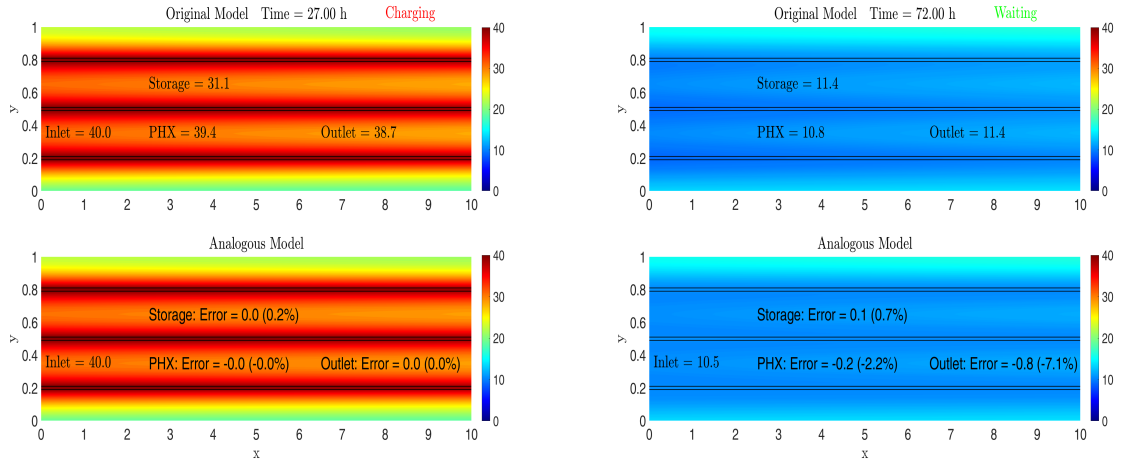


Figure 3.12: Spatial distribution of the temperature in the storage with three horizontal symmetric PHX*s* during charging (left) and waiting (right).
Top: Original system. Bottom: Analogous system.

Remark 3.5.1 The poor precision of the outlet temperature approximation by the analogous model during waiting periods is of no relevance for the management and operation of the GS within a residential heating system. Here, the outlet temperature is required only during charging and discharging but not during the waiting periods. The interesting quantity for which a good approximation precision is required is the average temperature in the storage and this is provided by the analogous model.

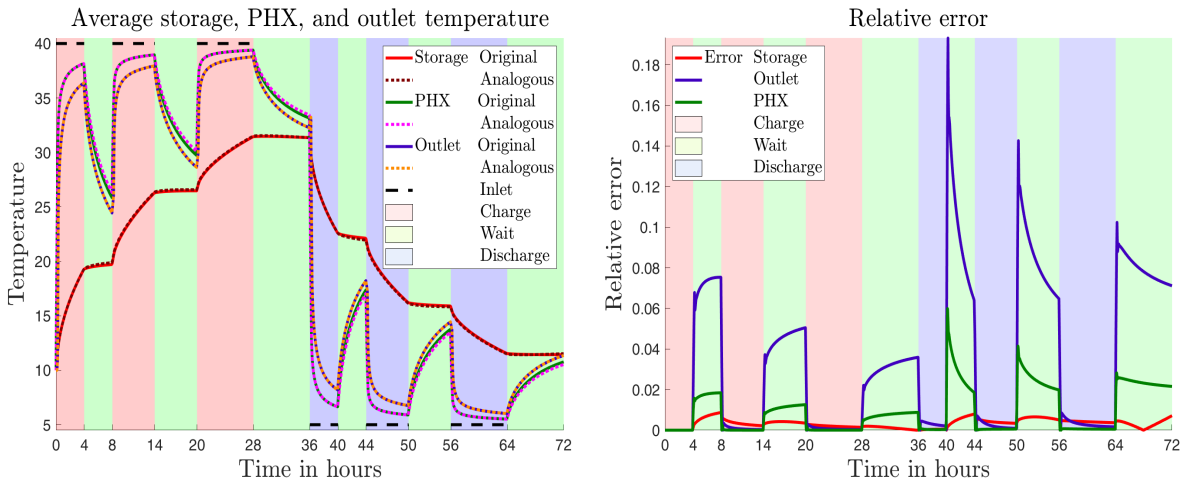


Figure 3.13: Original and analogous system of a storage with three horizontal non-symmetric PHX*s* during 72 *h* of charging, waiting and discharging. Left: Comparison of aggregated characteristics $\overline{Q}^S, \overline{Q}^F, \overline{Q}^O$. Right: Relative error of approximation by analogous system

CHAPTER 4

Model Order Reduction of the Dynamics of a Geothermal Storage

Introduction

The aim of this chapter is to reduce the dimension of system (3.10) to facilitate the computation of the charging and discharging decisions of the storage manager. This technique is known as MOR. Balanced truncation model reduction method adopted for this purpose is one of the most common model reduction techniques for standard state space systems. System (3.10) that we aim to reduce the dimension is linear time-varying with time-dependent system matrix caused by the fluid velocity $v_0(t)$. Indeed, the velocity $v_0(t) = v_0$, v_0 is constant during charging/discharging periods (when the pump is on) and $v_0(t) = 0$ during waiting period (when the pump is off). This system could be considered as linear switched system since the velocity of the fluid changes from one mode to another and leads to different systems matrices when switching from one mode to another. Balanced truncation for such a linear switched system exists, but it requires too much computational effort. In fact, one has to compute pairs of controllability and observability Gramians corresponding to each active mode by solving systems of coupled Lyapunov equations, see Gosea et al. [49]. To remedy this, we consider the analogous model presented in in Sec. 3.4 which mimics the original model by a linear time-invariant system where the pump is always on and the fluid velocity $v(t) = v_0$ is constant on the entire interval $\in [0, T]$. The idea of the analogous model is that we use at the inlet and outlet boundary also during the waiting period the same type of boundary conditions as during charging and discharging. However, we choose the inlet temperature to be equal to the average temperature in the pipe. Numerical examples presented in Subsec. 3.5.4 show that the analogous system approximates the original system quite well.

In this chapter the goal is to use Lyapunov balanced truncation MOR to reduce the dimension of the linear time-invariant system (analogous system) resulting from the semi-discretization of the heat equation describing the dynamics of the geothermal energy storage. The latter has many advantages such as preservation of several system properties like stability and passivity, see Pernabo and Silverman [83] and guarantees the existence of a priori error bound, see Enns [41] for the difference between the outputs of the full and the reduced model. This a priori error bound permits an appropriate choice of the order of the reduced-order model depending on how accurate the approximation is needed.

The rest of the chapter is organized as follows. Sec. 4.1 we start with the formulation of the general MOR problem. Then we present the Lyapunov balanced truncation method which is based on the computation of the observability and controllability Gramians as solutions of two algebraic Lyapunov equations in Sec. 4.2. In Sec. 4.3 we demonstrate the efficiency of Lyapunov balanced truncation by numerical experiments for various settings of the output variables describing the aggregated characteristics of the temperature distribution in the GS.

4.1 Model Order Reduction

4.1.1 Problem Setup

In the previous sections we have seen that the spatio-temporal temperature distribution describing the input-output behavior of the GS can be approximately computed by solving the system of ODEs (3.10) for the n-dimensional function Y resulting from semi-discretization of the heat equation (2.6). Aggregated characteristics can be obtained by linear combinations of the entries of Y in a post-processing step, see Sec. 2.3.3. In the following we work with the approximation using an analogous system introduced in Sec. 3.4. Then the input-output behavior of the GS can be described by a LTI system, i.e., its response to any arbitrary input signal does not depend on absolute time. We consider in the sequel a pair of a linear autonomous differential and a linear algebraic equation which is well-known from linear system and control theory and of the form

$$\begin{aligned} \dot{Y}(t) &= AY(t) + Bg(t), \\ Z(t) &= CY(t). \end{aligned} \tag{4.1}$$

Here, $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{n_o\times n}$ for $n, m, n_o \in \mathbb{N}$ are called *system, input, output matrix*, respectively. Further, $g : [0,T] \to \mathbb{R}^m$ is the *input* (or control), $Y : [0,T] \to \mathbb{R}^n$ the *state* and $Z : [0,T] \to \mathbb{R}^{n_o}$ is the *output*. Given some initial value $Y(0) = y_0$ the input-output behavior, i.e., the mapping of the input g to the output Z is fully described by the triple of matrices (A,B,C) which is called *realization* of the above system.

For the analogous system, A and B are constant matrices which are given in (3.11) and (3.13) for the case of constant velocity $v_0(t) = \overline{v}_0$, i.e., the pump is on. From Theorem 3.1.13 we know that A is stable. The input dimension is $m = 2$ while the dimension n of the state depends on the discretization of the spatial domain $\mathcal{D}$. The two entries of the input g are the temperatures at the inlet and of the underground at the bottom boundary. The output Z contains the desired aggregated characteristics such as the average temperatures $\overline{Q}^\dagger$, $\dagger = M, F, O, B$, of the medium, the fluid, at the outlet or the bottom boundary. Thus, g is piecewise continuous and bounded. The associated row matrices $C^\dagger$ of the approximation $\overline{Q}^\dagger = C^\dagger Y(t)$ given in (3.26) form the n_o rows of the output matrix C. The output dimension n_o is the number of characteristics included in the problem and typically small while the state dimension n will be very large in order to obtain a reasonable accuracy for the semi-discretized solution of the heat equation.

For the above systems with high-dimensional state the simulation of the input-output behavior and the solution of optimal control problems suffer from the curse of dimensionality because of the computational complexity and memory requirements. This motivates us to apply model order reduction (MOR).

The general goal of MOR is to approximate the high-dimensional linear time-invariant sys-

tem (4.1) given by the realization (A,B,C) by a low-dimensional reduced-order system

$$\begin{aligned} \dot{\widetilde{Y}}(t) &= \widetilde{A}\widetilde{Y}(t)+\widetilde{B}g(t), \quad \widetilde{Y}(0)=\widetilde{y}_0 \\ \widetilde{Z}(t) &= \widetilde{C}\widetilde{Y}(t), \end{aligned} \tag{4.2}$$

i.e., a realization $(\widetilde{A},\widetilde{B},\widetilde{C})$ where $\widetilde{A}\in\mathbb{R}^{\ell\times\ell}$, $\widetilde{B}\in\mathbb{R}^{\ell\times m}$, $\widetilde{C}\in\mathbb{R}^{n_o\times\ell}$, $\widetilde{Y},\widetilde{Y}(0)\in\mathbb{R}^{\ell}$, $\widetilde{Z}\in\mathbb{R}^{n_o}$ and $\ell\ll n$ denotes the dimension of the reduced-order state. The reduced order initial condition $\widetilde{Y}(0)$ is obtained by projection of y_0 onto a low ℓ-dimensional subspace. We notice that the input variable g is the same for systems (4.1) and (4.2). Next we introduce the concept of transfer function of a LTI system.

Transfer function. Consider the Laplace transform of a function $f(t), t\in\mathbb{R}$, defined by

$$\mathbf{f}(s)=\int_0^\infty e^{-st}f(t)dt,$$

where s a complex variable called frequency. Then, taking the Laplace transform in (4.1) gives

$$\mathbf{Y}(s)=(s\mathbb{I}_n-A)^{-1}B\mathbf{g}(s)+(s\mathbb{I}_n-A)^{-1}Y(0), \tag{4.3}$$

$$\mathbf{Z}(s)=C(s\mathbb{I}_n-A)^{-1}B\mathbf{g}(s)+C(s\mathbb{I}_n-A)^{-1}Y(0), \tag{4.4}$$

where $\mathbf{Y}(s)$, $\mathbf{Z}(s)$ and $\mathbf{g}(s)$ are the Laplace transforms of $Y(t)$, $Z(t)$ and $g(t)$ respectively. Indeed, taking the Laplace transform in (4.1) and using the integration by part, we obtain

$$\begin{aligned} \int_0^\infty e^{-st}\dot{Y}(t)dt &= \int_0^\infty e^{-st}AY(t)dt+\int_0^\infty e^{-st}Bg(t)dt \\ e^{-st}Y(t)\Big|_0^\infty+s\int_0^\infty e^{-st}Y(t)dt &= A\int_0^\infty e^{-st}Y(t)dt+B\int_0^\infty e^{-st}g(t)dt \\ s\mathbf{Y}(s)-Y(0) &= A\mathbf{Y}(s)+B\mathbf{g}(s), \end{aligned}$$

from which we derive relation (4.3). Since $\mathbf{Z}(s)=C\mathbf{Y}(s)$, we immediately have relation (4.4). Define the rational matrix-valued function called transfer function of the continuous-time LTI system (4.1) by

$$G(s)=C(s\mathbb{I}_n-A)^{-1}B$$

We can observe that if initial state $Y(0)=0$, then (4.4) implies that $\mathbf{Z}(s)=G(s)\mathbf{g}(s)$. Hence, the transfer function $G(s)$ gives the relation between the Laplace transforms of the input $g(t)$ and the output $Z(t)$. Therefore, the transfer function of $G(s)$ describes the input-output behaviour of the LTI system (4.1) in the frequency domain under the assumption that the initial state is zero. If for any rational matrix-valued function $G(s)$ there exist matrices A, B and C such that $G(s)=C(s\mathbb{I}_n-A)^{-1}B$, Then the $(A,\ B,\ C)$ is called a realization of $G(s)$.

Remark 4.1.1 Note that the realization of $G(s)$ is not unique, see Stykel [107]. If the assumption on the initial state does not hold, i.e., $Y(0)\neq 0$, the above description of the input-output behaviour is not applicable. In addition, the transfer function representation does not reveal the behaviour inside the system such as unobservable unstable modes if the initial state is nonzero. This is due to the fact that observable modes can be excited due to a nonzero initial state but may not appear in the transfer function due to pole-zero cancellation, see De Almeida et al. [36].

Therefore, the transfer function matrix cannot always be used to study the stability properties of an LTI system.

Requirements of the MOR. The reduced-order system should capture the most dominant dynamics of the original system, in particular preserve the main physical system properties, e.g., stability. Further, it should provide a reasonable approximation of the original output Z by $\widetilde{Z}$ to given input g where the output error $\|Z-\widetilde{Z}\|_{\mathcal{L}^2([0,T))}$ satisfies the desired error tolerance

$$\|\widetilde{Z}-Z\|_{\mathcal{L}^2([0,T))} \leq \varepsilon_z \cdot \|g\|_{\mathcal{L}^2([0,T))}$$

for every input g, where ε_z is some fixed tolerance and the function space $\mathcal{L}^2$ is the space of all square integrable functions and the $\mathcal{L}^2$-norm of a function $f \in \mathcal{L}^2([0,T))$ is defined by

$$\|f\|_{\mathcal{L}^2([0,T])} = \left(\int_0^T \|f(t)\|_2^2 dt\right)^{1/2},$$

where $\|.\|_2$ is a Euclidean norm on the n-dimensional Euclidean space. For the limiting case $T \to \infty$, we simply write $\mathcal{L}^2$.
Equivalently, the transfer function $\widetilde{G}(s) = \widetilde{C}(s\mathbb{I}_\ell - \widetilde{A})^{-1}\widetilde{B}$ of the reduced order system should approximate the transfer function $G(s) = C(s\mathbb{I}_n - A)^{-1}B$ of the original system with a small error satisfying the desired error tolerance

$$\|G(.) - \widetilde{G}(.)\|_{\mathcal{H}^\infty} \leq \varepsilon_z,$$

where $\mathcal{H}^\infty$ is the Hardy space (a vector space of bounded holomorphic functions on the disk) with the $\|.\|_{\mathcal{H}^\infty}$ a norm associated to a system characterized by a transfer function $G(s)$ defined as in Stykel and Reis [108] by

$$\|G\|_{\mathcal{H}^\infty} = \sup_{\omega \in \mathbb{R}} \|G(i\omega)\|_2.$$

In addition, the computation of the reduced order system should be numerically stable and efficient. For $g \neq 0$ we have from the definition of $\mathcal{H}^\infty$ norm the following equivalent relation for the error measure,

$$\|\widetilde{Z}-Z\|_{\mathcal{L}^2} \leq \|G-\widetilde{G}\|_{\mathcal{H}^\infty}\|g\|_{\mathcal{L}^2}.$$

MOR techniques can be broadly classified in truncation/projection based methods and moment matching methods. These methods are singular value decomposition (SVD) based methods and Krylov based methods, respectively. Further, SVD can also be subdivided into two classes depending on the structure of the problem. The first class of methods among which the POD and the Gramian based approximations are suitable for nonlinear systems. The second class among which Balanced truncation and Hankel approximation methods are suitable for linear system. For a general overview see Antoulas et al. [5, 6] and Schilders et al. [97].
Below we will describe the general truncation/projection based model reduction procedure following by the detailed description of a specific truncation based method known as Truncated Balanced Realization method or balanced truncation. This method is well studied and known to produce reduced-order models that preserve properties of the original model and guarantee stable error bound. The basic idea of the truncation/projection based method is to truncate the

dynamical system studied at some point or in an appropriate basis. The latter is illustrated in Subsec.4.1.2. Although the balanced truncation method is not the fastest MOR method, our choice is motivated by the fact that it produces a very low-dimensional reduced-order system that suits our optimal control problem.

4.1.2 Projection-Based Methods

The underlying idea of projection-based methods is that the state dynamics can be well approximated by the dynamics of a projection of the n-dimensional state Y onto a suitable low-dimensional subspace of $\mathbb{R}^n$ of dimension $\ell < n$. Then the aim is to describe the dynamics of the projection by a ℓ-dimensional system of ODEs. Prominent examples are the modal truncation and balanced truncation method, see Antoulas [5, Secs. 7, 9.2]. Here, the projection is found by applying a suitable linear state *transformation* $\overline{Y} = \mathcal{T}Y$ with some non-singular $n \times n$-matrix $\mathcal{T}$ which allows to define the projection by *truncation* of $\overline{Y}$.

Transformation. The above mentioned transformation allows to derive the following alternative equivalent realization of the system (4.1). which is proven in [109, Appendix B.1].

Lemma 4.1.2 Let $\mathcal{T}$ be a $n \times n$ constant non-singular transformation matrix. If we define the transformation $\overline{Y} = \mathcal{T}Y$, then the state and output equation in (4.1) become

$$\begin{aligned} \dot{\overline{Y}}(t) &= \overline{A}\,\overline{Y}(t) + \overline{B}g(t), \\ Z(t) &= \overline{C}\,\overline{Y}(t). \end{aligned} \tag{4.5}$$

The realization of the system is given by $(\overline{A}, \overline{B}, \overline{C})$ with

$$\overline{A} = \mathcal{T}A\mathcal{T}^{-1}, \quad \overline{B} = \mathcal{T}B \quad \text{and} \quad \overline{C} = C\mathcal{T}^{-1}.$$

The following shows that the transfer is invariant under the above mentioned transformation.

Lemma 4.1.3 Let (A,B,C) be the realization of the transfer function $G(s)$ of the LTI system (4.1) and $(\overline{A}, \overline{B}, \overline{C})$ the realization of the transfer function $\overline{G}(s)$ of the LTI system (4.5).
Then $G(s) = \overline{G}(s)$.

Proof.

$$\overline{G}(s) = \overline{C}(s\mathbb{I}_n - \overline{A})^{-1}\overline{B}. \tag{4.6}$$

Substituting the the transformed matrices in equation (4.6) yields

$$\begin{aligned} \overline{G}(s) &= (C\mathcal{T}^{-1})(s\mathbb{I}_n - \mathcal{T}A\mathcal{T}^{-1})^{-1}(\mathcal{T}B) = C[(s\mathbb{I}_n - \mathcal{T}A\mathcal{T}^{-1})\mathcal{T}]^{-1}(\mathcal{T}B) \\ &= C[\mathcal{T}^{-1}(s\mathbb{I}_n - \mathcal{T}A\mathcal{T}^{-1})\mathcal{T}]^{-1}B = C(s\mathbb{I}_n - A)^{-1}B = G(s). \end{aligned}$$

□

Truncation. After transforming the system we proceed to the truncation step. It is assumed that the transformation $\mathcal{T}$ is such that the first ℓ entries of the transformed state $\overline{Y}$ forming the ℓ-dimensional vector $\overline{Y}_1$ represent the *most dominant* states while the remaining $n-\ell$ entries which are collected in the $n-\ell$-dimensional vector $\overline{Y}_2$ comprise the *less dominant* and *negligible* states. Based on this decomposition of $\overline{Y}$ into $\overline{Y}_1$ and $\overline{Y}_2$ the following partition of system

(4.5) into blocks is obtained

$$\begin{pmatrix}\dot{\overline{Y}}_1\\ \dot{\overline{Y}}_2\end{pmatrix} = \begin{pmatrix}\overline{A}_{11} & \overline{A}_{12}\\ \overline{A}_{21} & \overline{A}_{22}\end{pmatrix}\begin{pmatrix}\overline{Y}_1\\ \overline{Y}_2\end{pmatrix} + \begin{pmatrix}\overline{B}_1\\ \overline{B}_2\end{pmatrix} g(t), \qquad Z = \begin{pmatrix}\overline{C}_1 & \overline{C}_2\end{pmatrix}\begin{pmatrix}\overline{Y}_1\\ \overline{Y}_2\end{pmatrix}.$$

The above block partition of the equivalent realization (4.5) is used to define the reduced-order system by assuming that the truncation of $\overline{Y}$ to the first ℓ entries defines the desired projection. Then truncating the *less dominant* states $\overline{Y}_2$ and keeping only the first ℓ dominant states $\overline{Y}_1$ leads to the reduced system

$$\dot{\overline{Y}}_1 = \overline{A}_{11}\overline{Y}_1 + \overline{B}_1 g(t), \qquad \overline{Z} = \overline{C}_1\overline{Y}_1. \tag{4.7}$$

Thus is the desired reduced-order system (4.2) is given by the realization $(\widetilde{A},\widetilde{B},\widetilde{C}) = (\overline{A}_{11},\overline{B}_1,\overline{C}_1)$ and approximates the output Z of the original system (4.1) by $\widetilde{Z} = \overline{Z}$.

4.2 Lyapunov Balanced Truncation

Setting. The crucial question for the above introduced projection-based MOR methods is the choice of a *suitable* transformation matrix $\mathcal{T}$ which was left out and will be addressed in this subsection. The transformation should be such that the first entries of the transformed state $\overline{Y}$ provide the largest contribution to the input-output behaviour of the system. They carry the essential information for approximating the system output Z to a given input g with sufficiently high accuracy. On the other hand the last entries should deliver the smallest contribution to the input-output behaviour and thus can be neglected.

Lyapunov balanced truncation uses ideas from control theory, in particular the notion of controllability and observability which we sketch below. The basic idea is to define a transformation $\mathcal{T}$ that "balances" the state in a way that the first ℓ entries of $\overline{Y}$ are the states which are the easiest to observe and to reach. Then states which are simultaneously difficult to reach and to observe can be neglected and are truncated.

This method appears to be well-suited for the present problem of approximation input-output behavior of the GS. Balanced truncation MOR was first presented by Moore [77] who exploited results of Mullis and Roberts [78]. The preservation of stability was addressed by Pernebo and Silverman [84], error bounds derived by Enns [41] and Glover [48]

In the following we always assume that the linear system (4.1) is stable, i.e., the system matrix A is stable. From Theorem 3.1.13 it is known that this is the case in the problem under consideration. This allows that the system dynamics is not only considered on finite time intervals $[0,T]$ but also on $[0,\infty)$.

Given the initial state $Y(0) = y_0$ and the control g, there exists a unique solution to the continuous-time dynamical system (4.1) given by

$$\begin{aligned} Y(t) = \psi(t,y_0,g) &:= e^{At}y_0 + \int_0^t e^{A(t-s)}Bg(s)ds, \\ Z(t) = CY(t) &= Ce^{At}y_0 + \int_0^t Ce^{A(t-s)}Bg(s)ds. \end{aligned} \tag{4.8}$$

Note that the first term in the above state equation $\psi(t,y_0,0)$ representing the response to an initial condition y_0 and a zero input, while the second term is $\psi(t,0,g)$ and represents the

response to a zero initial state and input g.

4.2.1 Controllability and Observability

We now introduce some concepts from linear system theory which play a crucial role in the derivation of the balanced truncation method. Let us denote by $\mathcal{L}^2([t_1,t_2])$ the space of square integrable functions on $[t_1,t_2]$. We write $\mathcal{L}^2(0,\infty)$ for functions on $[0,\infty)$.

Definition 4.2.1 (Controllability)

1. Let the linear system (4.1) be given. A state $y \in \mathbb{R}^n$ is said to be **controllable** or **reachable** from zero initial state $y_0 = 0$ if there exist a finite time t^* and an input $g \in \mathcal{L}^2([0,t^*])$ such that the solution given in (4.8) satisfies $Y(t^*) = \psi(t^*,0,g) = y$.

2. The **controllable** or reachable **subspace** $\mathcal{Y}_C$ is the set of states that can be obtained from zero initial state and a given input $g \in \mathcal{L}^2([0,t^*])$.

3. The linear system (4.1) is said to be (completely) controllable if $\mathcal{Y}_C = \mathbb{R}^n$.

We note that for LTI systems the controllable subspace $\mathcal{Y}_C$ is invariant w.r.t. the choice of the initial state y_0. This allows the above restriction to controllability from zero initial state.

Remark 4.2.2 For linear systems the notions of controllability and reachability are equivalent. The latter is not true for nonlinear systems.

Definition 4.2.3 (Observability)

1. Let the linear system (4.1) be given and let $g = 0$. A state $y \in \mathbb{R}^n$ is said to be **unobservable** if the output $Z(t) = C\psi(t,y,0) = 0$ for all $t \geq 0$, i.e., for all $t \geq 0$ the output $Z(t)$ of the system to initial state $Y(0) = y$ is indistinguishable from the output to zero initial state. Otherwise the state y is called **observable**.

2. The **observable subspace** $\mathcal{Y}_O$ is the set all observable states.

3. System (4.1) is said to be (completely) observable if $y = 0$ is the only unobservable state.

Equivalently, the system is said controllable if the following controllability matrix has full rank

$$\mathcal{C}_n(A,B) = [B \quad AB \quad A^2B \quad \dots \quad A^{n-1}B],$$

and it is said to be observable if the following observability matrix has full rank

$$\mathcal{O}_n(C,A) = [C \quad CA \quad CA^2 \quad \dots \quad CA^{n-1}]^\top.$$

A realization (A,B,C) is said to be minimal if it is controllable and observable [36]. Minimal state-space realizations play a crucial role in balanced truncation in the sense that since it guarantees to obtain a reduced order state-space model which is minimal given a high dimensional system. Since minimal realization is both controllable and observable, it is a good basis for designing an observer to estimate the states of the system from measurements of the outputs.

The reduced order realization obtain from balanced truncation may not be minimal if the original state-space realization is not minimal. Nevertheless, different techniques exist to reduce non-minimal realizations to minimal realization. In [36] the authors presented a procedure to reduce non-minimal realizations to minimal realization in just two steps. In fact, this procedure is similar to the standard algorithm for reducing matrices to echelon from [75] but with small modification.

Remark 4.2.4 Let Y_1, Y_2 and Z_1, Z_2 denote state and output of system (4.1) to initial states y_1, y_2 and the same input g. If system (4.1) is observable then the equality of outputs $Z_1(t) = Z_2(t)$ for all $t \geq 0$ implies that $y_1 = y_2$. Otherwise it holds $y_1 \neq y_2$. This can easily be seen from the consideration for $Y = Y_1 - Y_2$ and $Z = Z_1 - Z_2$ which are the state and the output to initial state $y = 0$ and input $g = 0$.

The input-output behavior of system (4.1) can be quantified by the following measures of the "degree of controllability and observability". They are based on the (squared) $\mathcal{L}^2$-norms of the input and output functions which in the literature are often called "input and output energy".

Definition 4.2.5 (Controllability and Observability Function) The controllability function $\mathcal{E}_C : \mathcal{Y}_C \to [0, t^*)$, $t^* > 0$ is given for $y \in \mathcal{Y}_C$ by

$$\mathcal{E}_C(y) = \min_{\substack{g \in \mathcal{L}^2(0,\infty) \\ Y(0)=0, Y(\infty)=y}} \|g\|^2_{\mathcal{L}^2(0,\infty)} = \inf_{t^*>0} \min_{\substack{g \in \mathcal{L}^2([0,t^*)) \\ Y(0)=0,\ Y(t^*)=y}} \int_0^{t^*} \|g(t)\|_2^2 dt.$$

and the observability function $\mathcal{E}_O : \mathcal{Y}_O \to [0, \infty)$ is given for $y \in \mathcal{Y}_O$ by

$$\mathcal{E}_O(y) = \|Z\|^2_{\mathcal{L}^2(0,\infty)} = \int_0^{\infty} \|Z(t)\|_2^2 dt, \quad Y(0) = y,\ g(t) = 0 \quad \text{for all } t \geq 0. \quad (4.9)$$

The controllability function $\mathcal{E}_C(y)$ is the smallest input energy required to reach the state $y \in \mathcal{Y}_C$ from zero initial state in infinite time ($t^* \to \infty$). In view of the definition of a controllable state y given in Def. 4.2.1 the minimization is w.r.t. both the input function and the terminal time t^*. For more details see Antoulas [5, Lemma 4.29]. The observability function $\mathcal{E}_O(y)$ is obtained as the limit of the output energy $\int_0^{t^*} \|Z(t)\|_2^2 dt$ on a finite time interval $[0, t^*)$ for $t^* \to \infty$. Since the output energy increases in t^* we can consider $\mathcal{E}_O(y)$ as the maximum output energy produced by the system when it is released from initial state y for zero input.

The above quantities allow the following interpretation. States which are difficult to reach are characterized by large values of the controllability function $\mathcal{E}_C(y)$. They require large input energy to reach them. On the other hand, for large values of the observability function $\mathcal{E}_O(y)$ the state y is easy to observe whereas small values of $\mathcal{E}_O(y)$ indicate that state y is difficult to observe since it produces only small output energy. Note, that unobservable states do not produce output energy at all, and it holds $\mathcal{E}_O(y) = 0$ for $y \notin \mathcal{Y}_O$.

Gramians. Below we see that the controllability and observability function can be expressed in terms of the matrices called Gramians.

Definition 4.2.6 (Controllability and Observability Gramian)
Consider a stable LTI system (4.1). The matrices defined by

$$\mathcal{G}_C = \int_0^\infty e^{At} B B^\top e^{A^\top t} dt,$$
$$\mathcal{G}_O = \int_0^\infty e^{A^\top t} C^\top C e^{At} dt$$

are called (infinite) controllability and observability Gramians, respectively.

Note that the above Gramians are well-defined if and only if the LTI system (4.1) is stable. They have the following properties.

Lemma 4.2.7 The Gramians $\mathcal{G}_C$ and $\mathcal{G}_O$ are symmetric, positive semi-definite matrices.
If the linear system (4.1) is stable, controllable and observable then the Gramians are strictly positive definite.

Proof. The first two properties follow directly from the definition of the Gramians. The third property is proven in Theorems 2.2 and 3.2 of Davis et al. [34]. □
The following establishes the relation between the controllability and observability functions and the Gramians.

Theorem 4.2.8 Let the linear system (4.1) be stable, controllable and observable. Then the maximum output energy $\mathcal{E}_O(y)$ that can be obtained by observing the system with initial state $Y(0)$ and zero input can be written in terms of the observability Gramian $\mathcal{G}_\mathbf{O}$ as follows

$$\mathcal{E}_O(y) = y^\top \mathcal{G}_O y \quad \text{for } y \in \mathbb{R}^n, \tag{4.10}$$

and the minimal input energy $\mathcal{E}_C(y)$ required to steer the system to a fixed state $Y(t^*)$ from the zero state can be written in terms of the controllability Gramian $\mathcal{G}_C$ as follows.

$$\mathcal{E}_C(y) = y^\top \mathcal{G}_C^{-1} y \quad \text{for } y \in \mathcal{Y}_C, \tag{4.11}$$

Proof. For zero input $g = 0$ and initial state $Y(0) = y$ the state equation of system (4.1) has a unique solution $Y(t) = e^{At} y$ and the output is given by $Z(t) = CY = Ce^{At} y$ for $t \geq 0$. Hence

$$y^\top \mathcal{G}_O y = \int_0^\infty y^\top e^{A^\top t} C^\top C e^{At} y dt = \int_0^\infty (y e^{A^t} C)^\top C e^{At} y dt = \int_0^\infty \left| C e^{At} y \right|_2^2 dt = \int_0^\infty |Z(t)|_2^2 dt.$$

this prove relation (4.10). The proof of (4.11) can be derived from the results established in [107, Sec. 7.4]. □
The above relations show that states y in the span of the eigenvectors corresponding to small (large) eigenvalues of $\mathcal{G}_C$ lead to large (small) values of $\mathcal{E}_C(y)$. Thus such states require a high (small) input energy and are difficult (easy) to reach. On the other hand, states y in the span of the eigenvectors corresponding to small (large) eigenvalues of $\mathcal{G}_O$ produce only small (large) output energy $\mathcal{E}_O(y)$ and are difficult (easy) to observe.
The Gramians are therefore efficient tool to quantify the degree of controllability and observability of a given state. Further the eigenvectors of $\mathcal{G}_C$ and $\mathcal{G}_O$ span the controllable and observable subspace, respectively.

Balanced truncation model order reduction strategy consists of eliminating the states of the system that are simultaneously difficult to reach (require a large input energy to be reached) and

difficult to observe (produce a small observation output energy). Next we are going to show that the Gramians can be computed by solving some linear matrix equations.

Lyapunov equations. The major task of the balanced truncation is based on the computation of the Gramians in order to transform the system in its balanced form. The computation of the Gramians according to the relations (4.2.6) is time consuming but for the control problem consider in this thesis, the Grampians are computed once ans saved for future uses. This reduces the computational time for solving the control problem. A computationally feasible method is based on the fact that the Gramians satisfy some linear matrix equations.

Theorem 4.2.9 Let the LTI system (4.1) be stable. Then the controllability Gramian $\mathcal{G}_C$ and observability Gramian $\mathcal{G}_O$ satisfy the algebraic Lyapunov equations

$$A\mathcal{G}_C + \mathcal{G}_C A^\top = -BB^\top,$$
$$\mathcal{G}_O A + A^\top \mathcal{G}_O = -C^\top C.$$

The proof can be found in [109, Appendix B.3] but for the convenience of the reader, it is also provided in Appendix B.1 .

Remark 4.2.10 For solving to the above Lyapunov matrix equations usually numerical methods have to be applied. Such methods have been addressed in a large range of literature. For a low-dimensional and dense matrix A, the Lyapunov equations can be solved using Hammarling method [55, 60], for medium- to large-scale Lyapunov equations, a sign function method [20, 27] can be used, in the case of a large and sparse matrix A, projection-type methods such as $\mathcal{H}$-matrices based methods [53], Krylov subspace methods [64, 93, 102] and alternating direction implicit method [23, 24, 82] are more appropriate techniques for solving Lyapunov equations.

4.2.2 Balancing

We now come back to the construction of a *suitable* transformation matrix $\mathcal{T}$ which should be such that the first entries of the transformed state $\overline{Y} = \mathcal{T}Y$ provide the largest contribution to the input-output behaviour of the system. The idea of Lyapunov balanced truncation is to use a transformation $\mathcal{T}$ that *balances* the state such that the first entries of $\overline{Y}$ are simultaneously easy to reach and easy to observe. This allows to truncate the remaining states which are difficult to reach and difficult to observe.

We recall the interpretation of the Gramians according to Theorem 4.2.8. States that are easy to reach, i.e., those that require a small amount of input energy to reach, are found in the span of the eigenvectors of the controllability Gramian $\mathcal{G}_C$ corresponding to large eigenvalues. Further, states that are easy to observe, i.e., those that produce large amounts of output energy, lie in the span of eigenvectors of the observability Gramian $\mathcal{G}_O$ corresponding to large eigenvalues as well. However, in an arbitrary coordinate system, a state y that is easy to reach might be difficult to observe. On the other hand, there might exist a different state y' that is difficult to reach but easy to observe. consequently, it is hard to decide which of the two states y and y' is more important for the input-output behavior of the system.

This observation suggests to transform the coordinate system of the state space using a suitable transformation matrix $\mathcal{T}$ in which easy reachable states are simultaneously easy to

observe and vice versa. This is the case if the transformed Gramians coincide. Below we give a transformation for which the two Gramians are even diagonal.

We recall Lemma 4.1.2 which describes the transformation of the realization (A,B,C) of system (4.1) to the realization $(\overline{A},\overline{B},\overline{C})$ of the transformed system (4.5). The following lemma describes the transformation of the corresponding Gramians.

Lemma 4.2.11 Let $\mathcal{T}$ be a $n \times n$ non-singular transformation matrix defining the transformed system (4.5) for the transformed state $\overline{Y} = \mathcal{T}Y$.
For the transformed Gramians $\overline{\mathcal{G}_C}$ and $\overline{\mathcal{G}_O}$ of that system it holds

$$\overline{\mathcal{G}_C} = \mathcal{T}\mathcal{G}_C\mathcal{T}^\top, \quad \overline{\mathcal{G}_O} = \mathcal{T}^{-\top}\mathcal{G}_O\mathcal{T}^{-1} \quad \text{and} \quad \overline{\mathcal{G}_C}\overline{\mathcal{G}_O} = \mathcal{T}\mathcal{G}_C\mathcal{G}_O\mathcal{T}^{-1}.$$

The proof of Lemma 4.2.11 can be found in [109, Appendix B.4] and is also provided in Appendix B.2.
The last relation of the above lemma shows that the product $\overline{\mathcal{G}_C}\overline{\mathcal{G}_O}$ results from a similarity transformation of the product $\mathcal{G}_C\mathcal{G}_O$. Thus the eigenvalues of the product of the two Gramians are preserved under transformations of the coordinate system and can be considered as system invariants. They can be expressed as squares of the Hankel singular values of the system.

Definition 4.2.12 (Hankel Singular Values) The Hankel singular values are the square roots of the eigenvalues of the product of the controllability and observability Gramian

$$\sigma_i = \sqrt{\lambda_i(\mathcal{G}_C\mathcal{G}_O)}, \quad i = 1,\ldots,n.$$

Here $\lambda_i(G)$ denotes the i-th eigenvalue of the matrix G, ordered as $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n \geq 0$.

For linear systems (4.1) which are controllable, observable and stable it is known that the Hankel singular values are strictly positive, see Antoulas [5, Lemma 5.8].

The next proposition presents the main result of Lyapunov balanced truncation and gives the transformation matrix $\mathcal{T}$ that balances the system such that the two Gramians are equal and diagonal.

Theorem 4.2.13 (Balancing transformation)
Let the linear system (4.1) be stable, controllable and observable. Further, let

$\mathcal{G}_C = UU^\top$	be the Cholesky decomposition of the controllability Gramian,
$U^\top\mathcal{G}_OU = K\Sigma^2K^\top$	be the eigenvalue decomposition of $U^\top\mathcal{G}_OU$ such that
$\Sigma = \operatorname{diag}(\sigma_1,\ldots,\sigma_n)$	is the diagonal matrix of Hankel singular values from (4.2.12).

Then for the transformation matrix

$$\mathcal{T} = \Sigma^{\frac{1}{2}}K^\top U^{-1} \tag{4.12}$$

the transformed system (4.5) is balanced and the controllability $\overline{\mathcal{G}_C}$ and observability $\overline{\mathcal{G}_O}$ Gramians are given by

$$\overline{\mathcal{G}_C} = \overline{\mathcal{G}_O} = \Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{bmatrix}$$

i.e., they are diagonal and equal to a diagonal matrix containing the Hankel singular values as diagonal entries.

The proof of Theorem 4.2.13 can be found in [109, Appendix B.5] and is also provided in Appendix B.3.

Remark 4.2.14 The above approach can be numerically inefficient and ill-conditioned. The reason is that for large-scale systems the Gramians $\mathcal{G}_C, \mathcal{G}_O$ often have a numerically low rank compared to the dimension n. This is due to the fast decay of the eigenvalues of the Gramians and also of the Hankel singular values. Then the computation of inverse matrix such as U^{-1} should be avoided from a numerical point of view. In the literature there are several suggestions of alternative approaches which are identical in theory but yield algorithms with different numerical properties, see [5, 21, 119]. One of them is given below.

Theorem 4.2.15 (Square Root Algorithm, Antoulas [5], Sec. 7.4) Let the assumptions of Theorem 4.2.13 be fulfilled. Further, let

$\mathcal{G}_C = UU^\top$ be the Cholesky decomposition of the controllability Gramian,
$\mathcal{G}_O = LL^\top$ be the Cholesky decomposition of the observability Gramian,
where U and L are lower triangular matrices,
$U^\top L = W\Sigma V^\top$ be the singular value decomposition of $U^\top L$ with the
orthogonal matrices W and V.

Then the transformation matrix $\mathcal{T}$ given in (4.12) and its inverse can be represented as

$$\mathcal{T} = \Sigma^{-1/2} V^\top L^\top \quad \text{and} \quad \mathcal{T}^{-1} = UW\Sigma^{-1/2}. \tag{4.13}$$

Note that the computation of $\mathcal{T}$ according to (4.13) does not require the inversion of the full matrix U as in (4.12).

Truncation. The final step of the Lyapunov balanced truncation MOR is the truncation of the balanced system as it is explained in Subsec. 4.1.2. Then the truncated balanced system as in (4.7) is the desired reduced-order system (4.2). Antoulas [5] shows in Theorem 7.9 that for $\sigma_\ell > \sigma_{\ell+1}$ the reduced-order system is again stable, controllable and observable.

Lyapunov balanced truncation algorithm. Given a minimal realization (A, B, C) of the stable high-dimensional LTI system (4.1). Then, the balanced truncation procedure can be summarized in the following algorithm.

Algorithm 1: Balanced truncation algorithm

Result: Given a minimal realization (A,B,C) find the reduced realization $(\widetilde{A},\widetilde{B},\widetilde{C})$

1. Compute the Gramians $\mathcal{G}_C$ and $\mathcal{G}_O$ by solving the Lyapunov equations $A\mathcal{G}_C + \mathcal{G}_C A^\top = -BB^\top$ and $\mathcal{G}_O A + A^\top \mathcal{G}_O = -C^\top C$.
2. Compute the Cholesky decomposition of the Gramians $\mathcal{G}_C = UU^\top$ and $\mathcal{G}_O = LL^\top$.
3. Compute the singular value decomposition $U^\top L = W\Sigma V^\top$ with $\Sigma = \operatorname{diag}(\sigma_1,\ldots,\sigma_n)$ and $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_\ell > \sigma_{\ell+1} \geq \ldots \geq \sigma_n \geq 0$.
4. Construct the transformation matrices $\mathcal{T} = \Sigma^{-\frac{1}{2}} V^\top L^\top$ and $\mathcal{T}^{-1} = UW\Sigma^{-\frac{1}{2}}$.
5. Transforming the state to $\overline{Y} = \mathcal{T}Y$ leads to the balanced system defined by the matrices
$$(\overline{A},\overline{B},\overline{C}) = (\mathcal{T}A\mathcal{T}^{-1}, \mathcal{T}B, C\mathcal{T}^{-1}) = \left(\begin{pmatrix} \overline{A}_{11} & \overline{A}_{12} \\ \overline{A}_{21} & \overline{A}_{22} \end{pmatrix}, \begin{pmatrix} \overline{B}_1 \\ \overline{B}_2 \end{pmatrix}, \begin{pmatrix} \overline{C}_1 & \overline{C}_2 \end{pmatrix} \right).$$
6. Truncate the transformed state $\overline{Y} = \mathcal{T}Y$ keeping the first ℓ entries and choose
$$(\widetilde{A},\widetilde{B},\widetilde{C}) = (\overline{A}_{11},\overline{A}_1,\overline{C}_1), \quad \widetilde{A} \in \mathbb{R}^{\ell\times\ell}, \ \widetilde{B} \in \mathbb{R}^{\ell\times m}, \ \widetilde{C} \in \mathbb{R}^{n_o\times\ell}.$$

Remark 4.2.16

1. From an implementation point of view in the above algorithm the truncation in step 6 can be moved to step 4 which allows the construction of the reduced order system without forming the high-dimensional balanced system in step 5 and no full SVD is necessary. This leads to the following modification of Algorithm 1.

   4*) Construct new transformation matrices with the $\ell \times n$ matrix $\mathcal{T}^+ = \Sigma_\ell^{-1/2} V_\ell^\top L^\top$ and the $n \times \ell$ matrix $\mathcal{T}^- = UW_\ell \Sigma_\ell^{-1/2}$.
   Note that $\mathcal{T}^+\mathcal{T}^- = I_\ell$ is the identity matrix of dimension ℓ.

   5*) Omitted.

   6*) The reduced-order system is obtained directly by
   $$(\widetilde{A},\widetilde{B},\widetilde{C}) = (\mathcal{T}^+ A\mathcal{T}^-, \mathcal{T}^+ B, C\mathcal{T}^-).$$

   This modification is also interesting from an computational point of view since it requires not the full but only a partial singular value decomposition of $U^\top L$ in step 3.

2. The above modification is also useful for linear systems which are not (fully) observable or controllable as it is often the case if system (4.1) results from the semi-discretization of PDEs. We also observe this for the system derived above in (3.10). Then the Gramians $\mathcal{G}_C, \mathcal{G}_O$ and also the product $\mathcal{G}_C\mathcal{G}_O$ might be singular and there are zero Hankel singular values such that we have $\sigma_1 \geq \ldots \geq \sigma_{n_0} > 0 = \sigma_{n_0+1} = \ldots = \sigma_n$ for some $n_0 < n$. In this case the transformation matrix $\mathcal{T}$ and its inverse $\mathcal{T}^{-1}$ are not defined and the above algorithm breaks down. However, $\mathcal{T}^+$ and $\mathcal{T}^-$ are well-defined for any $\ell \leq n_0$ and formally the above described modification works well.

A theoretical justification of that approach is given in Tombs and Postlethwaite [113, Theorem. 2.2]. The authors show that for $\ell = n_0$ the reduced-order system is stable, fully controllable and observable and has the same transfer function as the original system, i.e. there is no change of the input-output behavior. This allows to fit a stable but not necessarily fully observable and controllable system into the above framework from the beginning of this section. In a pre-processing step balanced truncation is formally applied to obtain a reduced-order system of dimension $\ell = n_0$. The latter system is stable, controllable and observable and can be further reduced applying the standard methods to obtain an approximation of the input-output behavior.

4.2.3 Error Bounds

One of the advantages of Lyapunov balanced truncation is that there exist error estimates which can be given in terms of the discarded Hankel singular values. They allow to select the dimension ℓ of the reduced-order system such that a prescribed accuracy of the output approximation is guaranteed. In the literature these error estimates are given for the transfer functions of the original and the reduced-order system in the $\mathcal{H}^\infty$-norm in terms of the Hankel singular values that are truncated as

$$\|G - \widetilde{G}\|_{\mathcal{H}^\infty} = \sup_{\omega \in \mathbb{R}} \|G(i\omega) - \widetilde{G}(i\omega)\|_2 \leq 2 \sum_{i=\ell+1}^{n} \sigma_i,$$

where n is the total number of singular values. From which one can derive the estimates given below for the error measured in the $\mathcal{L}^2(0,t)$-norm. The following theorem is proven in Enns [41] and Glover [48].

Theorem 4.2.17 Let the linear system (4.1) be a stable, controllable and observable with zero initial value, i.e. $Y(0) = 0$. Further, let Hankel singular values be pairwise distinct with $\sigma_1 > \ldots > \sigma_n > 0$. Then it holds for all $t \geq 0$

$$\|Z - \widetilde{Z}\|_{\mathcal{L}^2(0,t)} \leq 2 \sum_{i=\ell+1}^{n} \sigma_i \, \|g\|_{\mathcal{L}^2(0,t)}. \tag{4.14}$$

The error bound given in (4.14) depends on the reduced order ℓ only via the sum of discarded Hankel singular values for which we have

$$\sum_{i=\ell+1}^{n} \sigma_i = \operatorname{tr}(\Sigma_2) = \operatorname{tr}(\Sigma) - \operatorname{tr}(\Sigma_1),$$

where $\Sigma_1 = \operatorname{diag}(\sigma_1, \ldots, \sigma_\ell)$ and $\Sigma_2 = \operatorname{diag}(\sigma_{\ell+1}, \ldots, \sigma_n)$. According to proposition 4.2.13 we have that $\overline{\mathcal{G}}_C = \overline{\mathcal{G}}_O = \Sigma$, i.e., the controllability and observability Gramians of the balanced system are equal to the diagonal matrix Σ. Further, the two Gramians of the reduced-order system are also equal and given by the diagonal matrix Σ_1. Relations (4.9) and (4.10) state that the output energy contained in the output Z when the system is released from initial state y for zero input is given by $\mathcal{E}_O(y) = \|Z\|^2_{\mathcal{L}^2(0,\infty)} = y^\top \mathcal{G}_O y$. Given the balanced realization of the original system (4.1) the output energy related to the initial state $Y(0) = \mathbf{1}_n$, i.e., the vector with all entries equal one, is $\mathcal{E}_O(\mathbf{1}_n) = \mathbf{1}_n^\top \Sigma \mathbf{1}_n = \operatorname{tr}(\Sigma)$. Analogously, for the reduced-order system the output energy related to the initial state $\widetilde{Y}(0) = \mathbf{1}_\ell$ obtained by truncation of $\mathbf{1}_n$ to the first ℓ

entries is given by $\mathcal{E}_O(\mathbf{1}_\ell) = \|\widetilde{Z}\|^2_{\mathcal{L}^2(0,\infty)} = \operatorname{tr}(\Sigma_1)$. Thus the sum of discarded Hankel singular values which is $\operatorname{tr}(\Sigma) - \operatorname{tr}(\Sigma_1)$ can be interpreted as the loss of output energy due to truncation if the balanced system starts with an initial *excitation* where all (balanced) states are equal one and no forcing is applied, i.e., the input g is zero.

Obviously, the above sum is decreasing in ℓ and becomes zero for $\ell = n$. This motivates the introduction of the following relative selection criterion

$$\mathcal{S}(\ell) = \frac{\operatorname{tr}(\Sigma_1)}{\operatorname{tr}(\Sigma)}, \quad \text{for } \ell = 1, \ldots, n,$$

with values in $(0,1]$. It is increasing in ℓ with $\mathcal{S}(\ell) = 1$ for $\ell = n$. $\mathcal{S}(\ell)$ can be used as a measure of the proportion of the output energy which can be captured by a reduced-order system of dimension ℓ. The selection of an appropriate dimension ℓ can be based on a prescribed threshold level $\alpha \in (0,1]$ for that proportion for which the minimal reduced order reaching that level is defined by

$$\ell_\alpha = \min\{\ell : \mathcal{S}(\ell) \geq \alpha\}. \tag{4.15}$$

Remark 4.2.18

1. Since the Hankel singular values only depend on the original model (4.1) the error bound in (4.14) can be computed a priori. Given the input g this allows to control the approximation error by the selection of the reduced order ℓ.

2. The error bound can be generalized to systems with Hankel singular values with multiplicity larger than one. In this case they only need to be counted once, leading to tighter bounds, see Glover [48].

3. The assumption of a zero initial state is quite restrictive and usually not fulfilled in applications. We refer to Beattie et al. [14], Daraghmeh at al. [33], Heinkenschloss et al. [57] and Schröder and Voigt [98] where the authors study the general case of linear systems with non-zero initial conditions and derive error bounds with extra terms accounting for the initial condition.

4. The linear systems considered in the present paper are obtained by semi-discretization of the heat equation (2.6) with the associated boundary and initial conditions. The initial value $Y(0) = y_0 \in \mathbb{R}^n$ represents the initial temperatures $Q(0,\cdot,\cdot)$ at the corresponding grid points. In general we have $y_0 \neq 0_n$. However, for the case of a homogeneous initial temperature distribution with $Q(0,x,y) = Q_0$ for all $(x,y) \in \mathcal{D}$ and some constant Q_0 one can derive an equivalent linear system with zero initial value. That case is considered in our numerical experiments in Sec. 4.3.

   The idea is to shift the temperature scale by Q_0 and describe the temperature distribution by $\widehat{Q}(t,x,y) = Q(t,x,y) - Q_0$. Then the initial condition reads as $\widehat{Q}(0,x,y) = 0$. Thanks to linearity of the heat equation (2.6) and the boundary conditions, $\widehat{Q}$ also satisfies the heat equation together with the boundary conditions if the inlet and ground temperature appearing in the Dirichlet and Robin condition are shifted accordingly, i.e., Q^I and Q^G are replaced by $Q^I - Q_0$ and $Q^G - Q_0$. Semi-discretization of the modified heat equation generates a linear system (4.1) with zero initial condition $Y(0) = 0_n$ for which we can apply the error estimate given in (4.14). The aggregated characteristics of the temperature

distribution corresponding to the model with constant but non-zero initial temperature Q_0 and forming the system output Z can easily be derived from the output of the modified system. In a post-processing step the inverse shift of the temperature scale is applied where Q_0 is added to all temperatures.

4.3 Numerical Results

In this section we present results of numerical experiments on model order reduction for the system of ODEs (3.10) resulting from semi-discretization of the heat equation (2.6) which models the spatio-temporal temperature distribution of a GS. For describing the input-output behavior of that storage we use the aggregated characteristics of the spatial temperature distribution introduced in Sec. 2.3.3. Further we work with the approximation of (3.10) by an analogous system as explained in Sec. 3.4.

The experiments are based on Algorithm 1 and are performed for the cases of one and three PHXs . While a model with three PHXs is more realistic and shows more structure of the spatial temperature distribution in the storage the case of only one PHX is an interesting benchmark model allowing for reduced-order systems of very low dimension since the spatial temperature distribution is less heterogeneous.

After explaining the experimental settings in Subsec. 4.3 we start in Subsec. 4.3.1 with an experiment where the system output consists of only one variable which is the average temperature $\overline{Q}^M$ in the storage medium. For that case we restrict ourselves to charging and discharging the storage without intermediate waiting periods. Note that the analogous system requires during the waiting periods the knowledge of the average PHX temperature which is not included in the output variables.

In Subsec. 4.3.2 we add a second output variable which is the average PHX temperature $\overline{Q}^F$. This now allows to work with waiting periods. Then we add a third output variable which is in Subsec. 4.3.3 the average temperature at the outlet $\overline{Q}^O$ whereas in Subsec. 4.3.4 the average temperature at the bottom $\overline{Q}^B$ of the storage is used as third output variable. The outlet temperature $\overline{Q}^O$ is interesting for the management of heating systems equipped with a GS while $\overline{Q}^B$ allows to quantify the transfer of thermal energy to the environment at the bottom boundary of the GS. Finally, Subsec. 4.3.5 presents results for a model where the output contains all of the four above mentioned quantities.

Experimental settings

For our numerical examples we use the model and discretization parameters given in Sec. 3.5, Table 3.2 but we restrict to the case of 1 PHX and 3 PHXs. The storage is charged and discharged either by a single PHX or by three PHXs filled with a moving fluid, see Fig. 4.1. Thermal energy is stored by raising the temperature of the storage medium. We recall the open architecture of the storage which is only insulated at the top and the side but not at the bottom. This leads to an additional heat transfer to the underground for which we assume a constant temperature of $Q^G(t) = 15$ °C. In the simulations the fluid is assumed to be water while the storage medium is dry soil. During charging a pump moves the fluid with constant velocity $\overline{v}_0$ arriving with constant temperature $Q^I(t) = Q^I_C = 40$ °C at the inlet. This temperature is higher than in the vicinity of the PHXs, thus induces a heat flux into the storage medium. During discharging the inlet temperature is $Q^I(t) = Q^I_D = 5$ °C leading to a cooling of the storage. At the outlet we impose a vanishing diffusive heat flux, i.e. during pumping there is only a convective heat

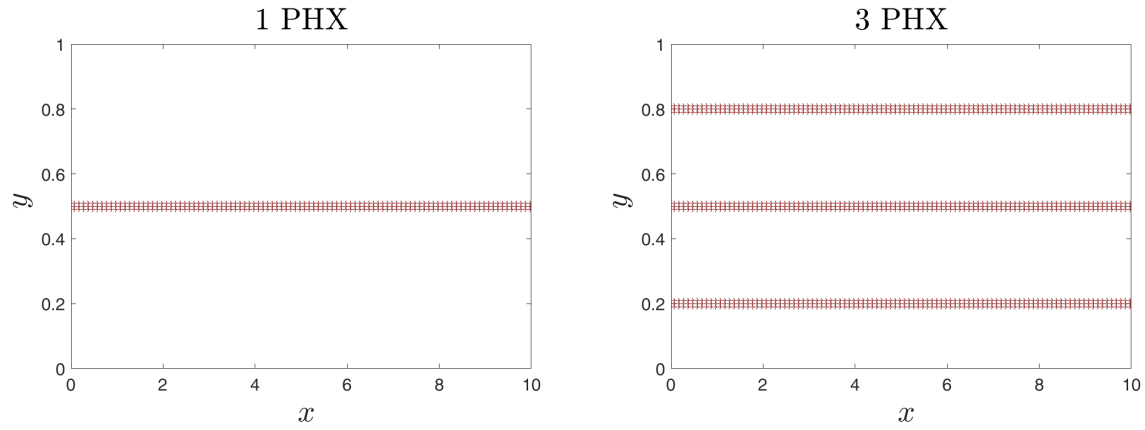


Figure 4.1: Computational domain with horizontal straight PHXs. Left: one PHX . Right: three PHXs.

flux. We also consider waiting periods where the pump is off. This helps to mitigate saturation effects in the vicinity of the PHXs which reduce the injection and extraction efficiency. During that waiting periods the injected heat (cold) can propagate to other regions of the storage. Since pumps are off we have only diffusive propagation of heat in the storage and the transfer over the bottom boundary.

For the chosen discretization parameters the dimension of the state equation (4.1) resulting from the space-discretization of the heat equation is $n = 10201$. Instead of the non-autonomous linear system of ODEs (4.1) we work with the LTI system obtained by approximating (4.1) by an analogous system as explained in Sec. 3.4. Recall, that here it is assumed that the pump is always on and during the waiting periods the inlet temperature Q^I is set to be the average temperature $\overline{Q}^F$ in the PHX (in the fluid). The output matrix C depends on the number of output variables and changes in the various experiments.

4.3.1 One Aggregated Characteristic: $\overline{Q}^M$

In this example we consider a model with only a single output variable $Z_1 = \overline{Q}^M$, the average temperature of the medium. Then the system output does not contain the average fluid temperature $\overline{Q}^F$ which serves in the analogous system as inlet temperature during waiting periods. Therefore, we consider only charging and discharging periods without intermediate waiting periods. For time horizon $T = 72$ hours we divide the interval $[0,T]$ into a charging period $I_C = [0,T/2]$ followed by a discharging period is $I_D = (T/2,T]$. The input function $g:\ [0,T] \to \mathbb{R}^2$ is defined as

$$g(t) = (Q^I(t), Q^G(t))^\top \quad \text{with} \quad Q^I(t) = \begin{cases} Q^I_C = 40\ ^\circ\text{C} & \text{for } t \in I_C \ \text{(charging)}, \\ Q^I_D = 5\ ^\circ\text{C} & \text{for } t \in I_D \ \text{(discharging)}, \end{cases}$$

with the piece-wise constant inlet temperature $Q^I(t)$ and the constant underground temperature $Q^G(t) = 15\ ^\circ$C. The output matrix in this case is given by $C = C^M$ which is given in Subsec. 3.3

In Fig. 4.2, the left panel shows first 50 largest Hankel singular values associated to the most observable and most reachable states, whereas in the right panel we plot the selection criterion against the reduced order ℓ (red for 1 PHX and blue for 3 PHXs). For the first 50 singular

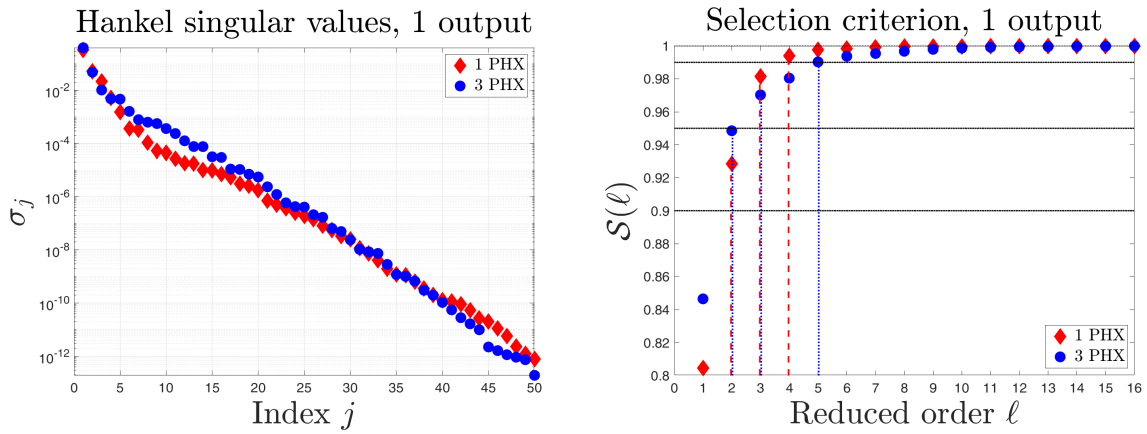


Figure 4.2: Model with one output $Z = \overline{Q}^M$:
Left: first 50 largest Hankel singular values, Right: selection criterion

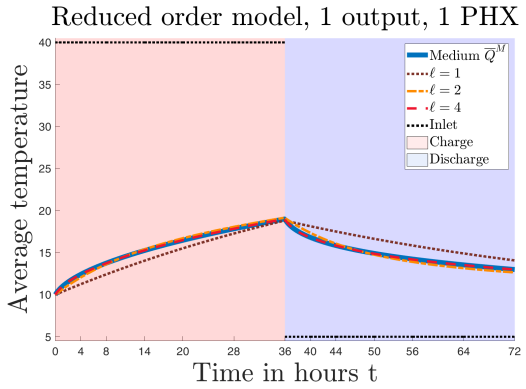
values, we observe for both models a fast decrease by 12 orders of magnitude. The first 20 singular values decrease faster for the model with 1 PHX than for the 3 PHX model.

Output	α	90%	95%	99%
$Z = \overline{Q}^M$		2/2	3/3	4/5
$Z = (\overline{Q}^M, \overline{Q}^F)^\top$		4/4	5/6	11/11
$Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^B)^\top$		5/6	7/8	12/13
$Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^O)^\top$		8/8	10/9	15/14
$Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^O, \overline{Q}^B)^\top$		9/9	11/11	17/16

Table 4.1: Minimal reduced orders $\ell_\alpha = \min\{\ell : \mathcal{S}(\ell) \geq \alpha\}$, 1 PHX / 3 PHXs

We recall that the selection criterion $\mathcal{S}(l)$ provides an estimate of the proportion of output energy of the original system that can be captured by the reduced-order system of dimension ℓ. In the right panel of Fig. 4.2 we draw vertical red dashed (one PHX) and blue dotted lines (three PHX) to indicate the reduced orders ℓ for which the selection criterion $\mathcal{S}(l)$ exceeds the threshold values $\alpha = \{90\%, 95\%, 99\%\}$ for the first time, respectively. This allows to determine graphically the associated minimal reduced orders $\ell_\alpha = \min\{\ell : \mathcal{S}(\ell) \geq \alpha\}$ which already have been introduced in (4.15). The resulting values are also given in Table 4.1. It can be seen that already with $\ell_{0.9} = 2$ states the reduced-order system can capture more than 90% of the output energy while with $\ell_{0.99} = 4$ states the level 99% is exceeded for the one PHX model whereas the three PHX model is only slightly below that level and requires $\ell_{0.99} = 5$ states.

Fig. 4.3 allows to evaluate the actual quality of the output approximation. It plots the average temperature in the medium $Z(t) = \overline{Q}^M(t)$ against time and compares that system output of the original system (solid blue line) with the approximation $\widetilde{Z}(t)$ from the reduced-order model (brown, orange, red lines) for $\ell = \{1, 2, 4\}$. The figures also show the inlet temperature $Q^I(t)$ (black dotted line) which are constant and equal to $Q^I_C = 40$ °C during charging (light red region) and $Q^I_D = 5$ °C during discharging (light blue region), respectively. The figure shows that both for one and three PHXs , the reduced-order system captures well the input-output behaviour



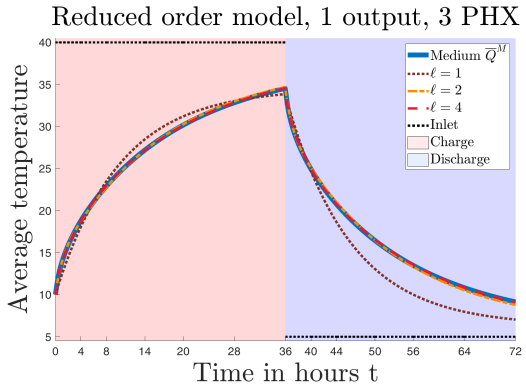
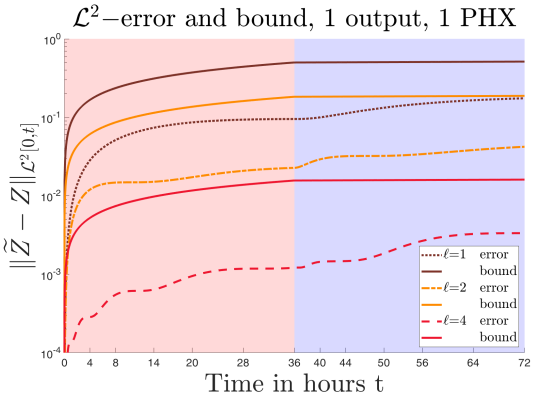


Figure 4.3: Model with one output $Z = \overline{Q}^M$: Left: one PHX , Right: three PHXs. Approximation of the output for $\ell = \{1,2,4\}$.



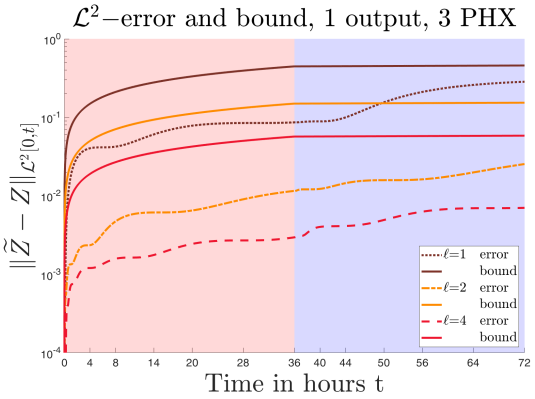


Figure 4.4: Model with one output $Z = \overline{Q}^M$: Left: one PHX , Right three PHXs. $\mathcal{L}^2$-error and error bound for $\ell = \{1,2,4\}$.

of the original high-dimensional system already for $\ell \geq 2$. For $\ell = 1$ the approximation is less good as expected in view of the low value of the selection criterion (see Fig. 4.2).

Finally, Fig. 4.4 plots the $\mathcal{L}^2$-error $\|Z - \widetilde{Z}\|_{\mathcal{L}^2[0,t]}$ against time t and compares with the associated error bound given in (4.14) for $\ell = 1,2,4$. Note that both quantities are non-decreasing in t. The error bounds grow less during discharging since here the growth of the $\mathcal{L}^2$-norm of the input g is smaller due to the smaller inlet temperature Q^I, see (4.14). As expected from this selection criterion, the error bounds decrease with ℓ and this is also the case for the actual error.

The next examples show that the number of states needed to capture well the input-output behavior of the system may increase considerably if we add more aggregated characteristics to the system output.

4.3.2 Two Aggregated Characteristics: $\overline{Q}^M, \overline{Q}^F$

In this example we add to the system output the average temperature of the PHX fluid leading to the two-dimensional output $Z = (\overline{Q}^M, \overline{Q}^F)^\top$. Since in the analogous system $\overline{Q}^F$ is used as inlet temperature, we now can include also waiting periods between periods of charging and discharging allowing the storage to mitigate saturation effects. For time horizon $T = 72$ hours we divide the simulation time interval $[0,T]$ into charging, discharging and waiting periods with

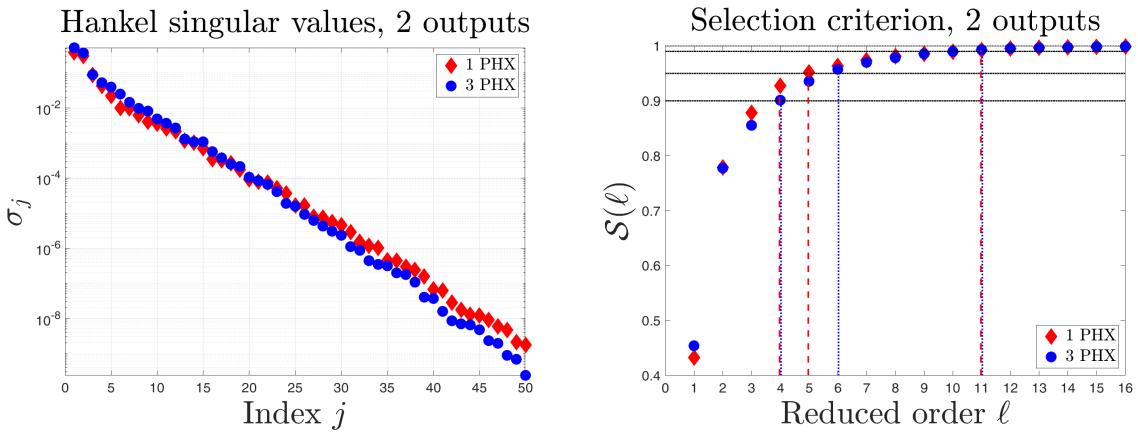


Figure 4.5: Model with two outputs $Z = (\overline{Q}^M, \overline{Q}^F)^\top$:
Left: first 50 largest Hankel singular values, Right: selection criterion

$$\begin{aligned} I_C &= [0,4] \cup [8,14] \cup [20,28], && \text{charging,} \\ I_D &= [36,40] \cup [44,50] \cup [56,64], && \text{discharging,} \\ I_W &= [0,72] \setminus (I_C \cup I_D), && \text{waiting.} \end{aligned}$$

which are also depicted in Fig. 4.6. The two-dimensional input function g is defined as

$$g(t) = (Q^I(t), Q^G(t))^\top \text{ with } Q^I(t) = \begin{cases} Q^I_C = 40\ ^\circ\text{C} & \text{for } t \in I_C \text{ (charging),} \\ Q^I_D = 5\ ^\circ\text{C} & \text{for } t \in I_D \text{ (discharging),} \\ \overline{Q}^F(t) & \text{for } t \in I_W \text{ (waiting).} \end{cases} \quad (4.16)$$

Here, the inlet temperature $Q^I(t)$ is piece-wise constant during charging and discharging but time-dependent and equal to $\overline{Q}^F(t)$ during waiting periods. The two rows of the $2 \times n$ output matrix C are C^M and C^F which are given in Subsec. 3.3.

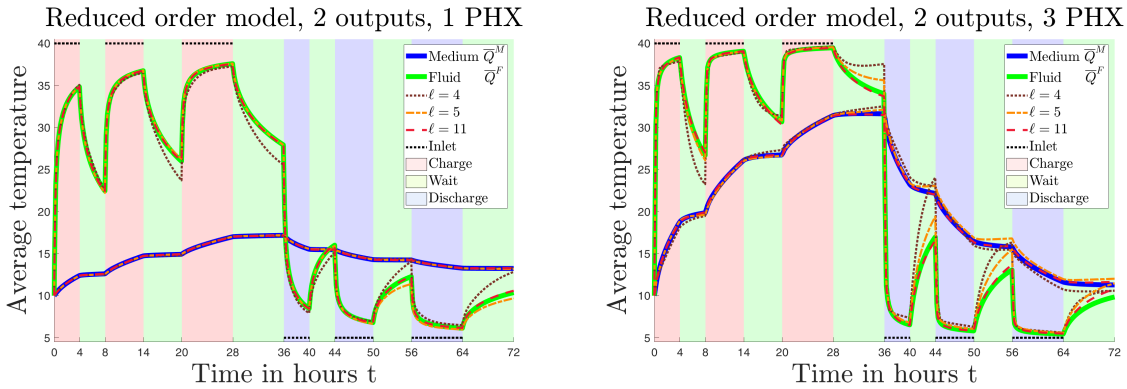


Figure 4.6: Model with two outputs $Z = (\overline{Q}^M, \overline{Q}^F)^\top$: Approximation of the output for $\ell = \{4,5,11\}$.
Left: one PHX , Right: three PHXs.

Fig. 4.5 depicts in the left panel the first 50 largest Hankel singular values, whereas the right panel shows the selection criteria (red for 1 PHX and blue for 3 PHXs). For the first 50 singular values we observe for both models a decrease by 9 orders of magnitude. As in the example with

a single output the first 20 singular values decrease faster for the model with one PHX than for the 3 PHX model. The selection criterion for the model with 1 PHX is for all $\ell \geq 2$ larger than for 3 PHXs. From Fig. 4.5 and also from the minimal reduced orders reported in Table 4.1 it can be seen that a reduced-order system with $\ell_{0.9} = 4$ states can capture more than 90% of the output energy of the original system. For the level threshold 95% the 1 PHX model requires $\ell_{0.95} = 5$ states, while for the 3 PHX model $\ell_{0.95} = 6$ states are needed. In both cases the level of 99% is exceeded for the first time for $\ell_{0.99} = 11$. Hence, for dimension $\ell \geq 11$ an almost perfect approximation of the input-output behavior can be expected.

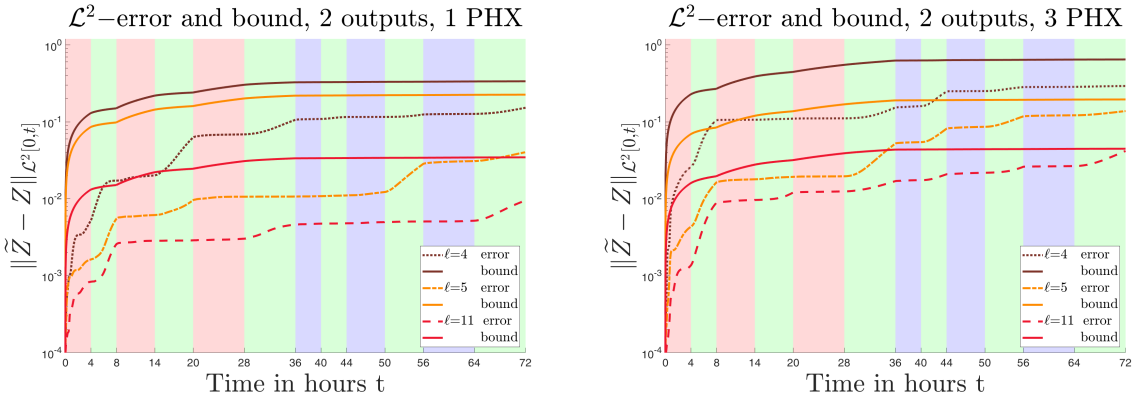


Figure 4.7: Model with two outputs $Z = (\overline{Q}^M, \overline{Q}^F)^\top$: $\mathcal{L}^2$-error and error bound for $\ell = 4, 5, 11$. Left: one PHX , Right: three PHXs.

For the evaluation of the actual quality of the output approximation we plot in Fig. 4.6 the output variables of the original and reduced-order system against time. The average temperatures $Z_1(t) = \overline{Q}^M(t)$ and $Z_2(t) = \overline{Q}^F(t)$ in the medium and fluid are drawn as solid blue and green lines, respectively, and its approximations as brown, orange and red lines for $\ell = \{4, 5, 11\}$. Further, the inlet temperature $Q^I(t)$ during the charging and discharging periods is shown as black dotted line. The figures show that the approximation of $\overline{Q}^M$ is better than for $\overline{Q}^F$. A possible explanation is that $\overline{Q}^M$ is an average of the spatial temperature distribution over the quite large subdomain $\mathcal{D}^M$ (medium) while for $\overline{Q}^F$ the temperature is averaged only over the much smaller subdomain $\mathcal{D}^F$ of the PHX fluid. Further, the temporal variations of $\overline{Q}^F$ are much larger than those of $\overline{Q}^M$ due to the impact of the changing inlet temperature during charging, discharging and waiting. Errors are more pronounced during waiting periods than during charging and discharging. For the three PHXs model the pointwise errors are slightly larger. As noted above, for $\ell = 11$ the selection criterion exceeds 99% and the approximation errors are almost negligible. This was also observed for $\ell > 11$.

Fig. 4.7 plots for the reduced orders ℓ considered above the $\mathcal{L}^2$-error $\|Z - \widetilde{Z}\|_{\mathcal{L}^2[0,t]}$ against time t together with the error bounds from (4.14). This allows an alternative evaluation of the approximation quality. As expected, the error bounds and also the actual errors decrease with ℓ. While the error bounds increase more during the charging periods due to the larger norm of the input g caused by the higher inlet temperature, the actual error increase more during the waiting periods. This corresponds to the above observed larger errors in the output approximation during these periods.

4.3.3 Three Aggregated Characteristics I: $\overline{Q}^M, \overline{Q}^F, \overline{Q}^O$

This example extends the example considered in Sec. 4.3.2 by adding a third variable to the system output which is the average temperature at the PHX outlet, i.e., we consider the three-dimensional output $Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^O)^\top$. The outlet temperature is needed if the GS is embedded into a residential system. Then the management of the heating system and the interaction between the geothermal and the internal buffer storage rely on the knowledge of $\overline{Q}^O$. Further, the difference $Q^I(t) - \overline{Q}^O(t)$ between inlet and outlet temperature is the key quantity for the quantification of the amount of heat injected to or withdrawn from the storage due to convection of the fluid in the PHX, we refer to Eq. (2.14) and the explanations in Sec. 2.3.3.

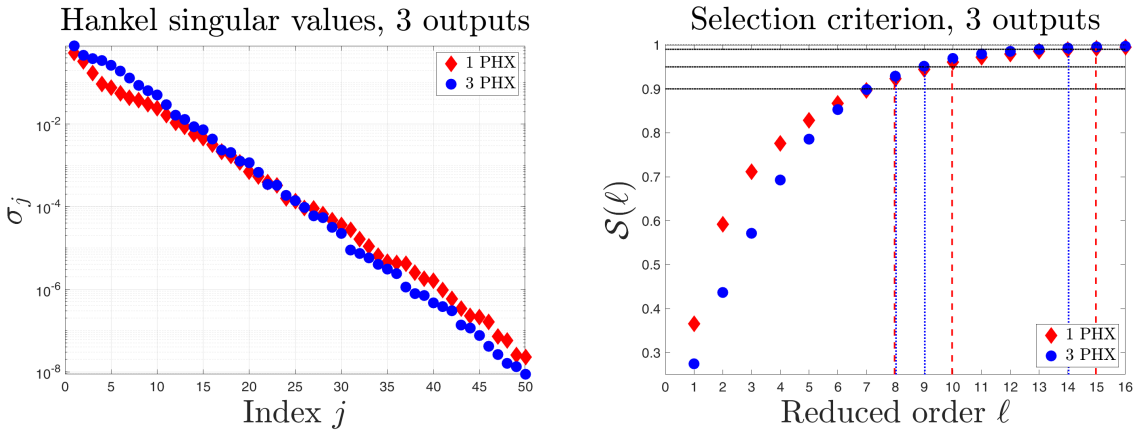


Figure 4.8: Model with three outputs $Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^O)^\top$:
Left: first 50 largest Hankel singular values, Right: selection criterion

The setting is analogous to Subsec. 4.3.2. The input function g is given in (4.16) and the $3 \times n$ output matrix C is formed by the three rows C^M, C^F, C^O which are given in Subsec. 3.3.

Fig. 4.8 shows, as in the previous experiments, in the left panel the first 50 largest Hankel singular values, whereas the right panel shows the selection criteria. For the first 50 singular values we observe for both models a decrease by 8 orders of magnitude which is slightly less than for the case of only two outputs. As in the examples with one and two outputs the first 20 singular values decrease faster for the model with 1 PHX than for the 3 PHX model. The selection criterion for the model with 1 PHX is for $\ell \leq 6$ larger than for 3 PHXs and for $\ell \geq 7$ slightly smaller. Table 4.1 shows that for reaching threshold levels of $\alpha = \{90\%, 95\%, 99\%\}$ in the one PHX case $\ell_\alpha = \{8, 10, 15\}$ states are required while for three PHXs one needs $\ell_\alpha = \{8, 9, 14\}$ states, respectively. Thus, for dimension $\ell \geq 15$ an almost perfect approximation of the input-output behavior can be expected. Note that in the previous experiment with two outputs (without outlet temperature) reaching the above thresholds requires about 4 to 5 states less.

In Fig. 4.9 we plot the output variables of the original and reduced-order system against time. The top panels show the average temperatures $Z_1(t) = \overline{Q}^M(t)$ and $Z_2(t) = \overline{Q}^F(t)$ in the medium and fluid which are drawn as solid blue and green lines, respectively. The bottom panels depict the average temperature at the outlet $Z_3(t) = \overline{Q}^O(t)$ by a solid green line. The reduced-order approximations as drawn for $\ell = \{8, 10, 15\}$. As in the previous experiments with two outputs it can be observed that the approximation of $\overline{Q}^M$ is better than for $\overline{Q}^F$. The approximation errors for the outlet temperature $\overline{Q}^O$ are quite similar to the errors for the fluid temperature. Note that the outlet temperature represents an average of the spatial temperature

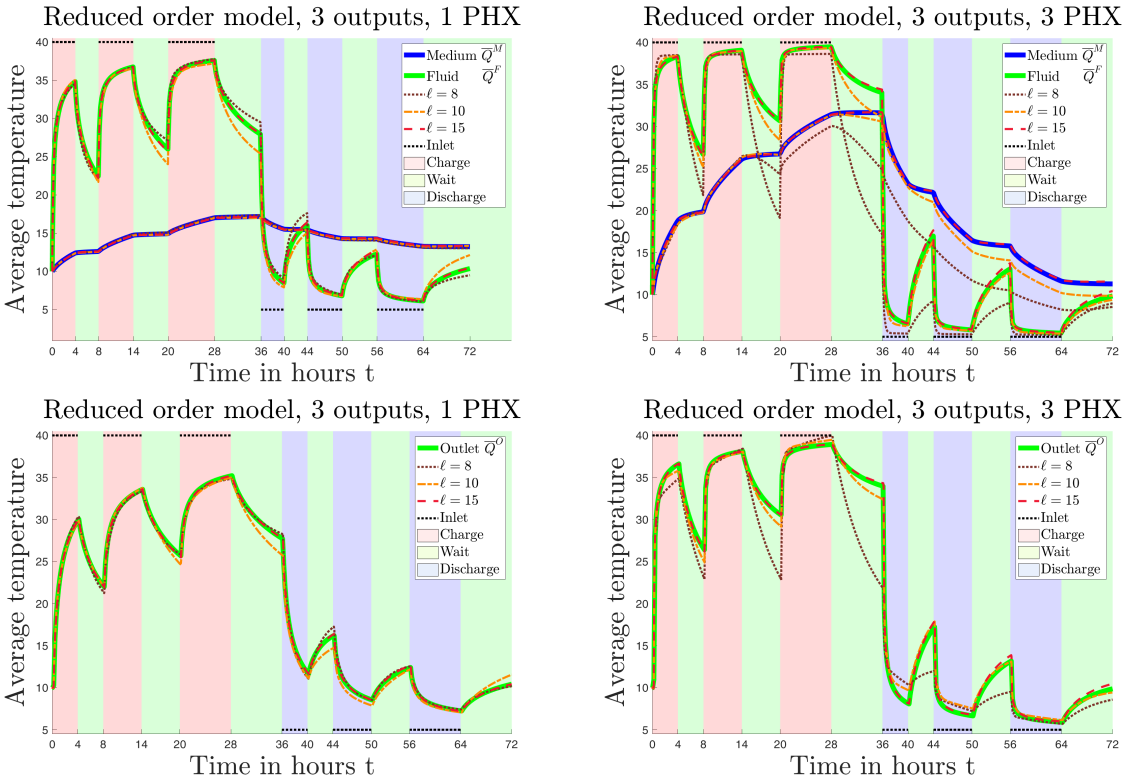


Figure 4.9: Model with three outputs $Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^O)^\top$: Approximation of the output for $\ell = 8, 10, 15$. Top: Average temperatures in the medium $\overline{Q}^M$ and the fluid $\overline{Q}^F$, Bottom: Average temperature at the outlet $\overline{Q}^O$, Left: one PHX , Right: three PHXs .

distribution over the quite small subdomain $\mathcal{D}^O$ on the boundary which is still smaller than the subdomain $\mathcal{D}^F$ over which the average is taken for the fluid temperature $\overline{Q}^F$. Both, the fluid and the outlet temperature show much larger temporal variations than the temperature in medium $\overline{Q}^M$. Again, errors are more pronounced during waiting periods than during charging and discharging, and for the 3 PHX model the pointwise errors are larger than for the 1 PHX model. For $\ell \geq 15$ states the selection criterion is above 99% and the approximation errors are almost negligible.

Fig. 4.9 also shows that the average temperatures of the fluid and at the outlet pipe, $\overline{Q}^M$ and $\overline{Q}^O$, exhibit almost the same pattern during the charging, discharging and waiting periods. Hence, knowing the average fluid temperature one can simply predict the outlet temperature and remove $\overline{Q}^O$ from the output variables. Then we are back in the setting of the two output experiment in Subsec. 4.3.2 and need 4 to 5 states less to capture the input-output behaviour with the same approximation quality. Below in Subsec. 4.3.4 we consider a model where instead of removing $\overline{Q}^O$ from the output this quantity is replaced by the average bottom temperature $\overline{Q}^B$ leading again to a model with three outputs.

An alternative evaluation of the approximation quality can be derived from Fig. 4.10 which plots for the reduced orders ℓ considered above the $\mathcal{L}^2$-error $\|Z - \widetilde{Z}\|_{\mathcal{L}^2[0,t]}$ against time t together with the error bounds from (4.14). The results are similar to Fig. 4.7 and we refer for the interpretation to the end of the previous subsection.

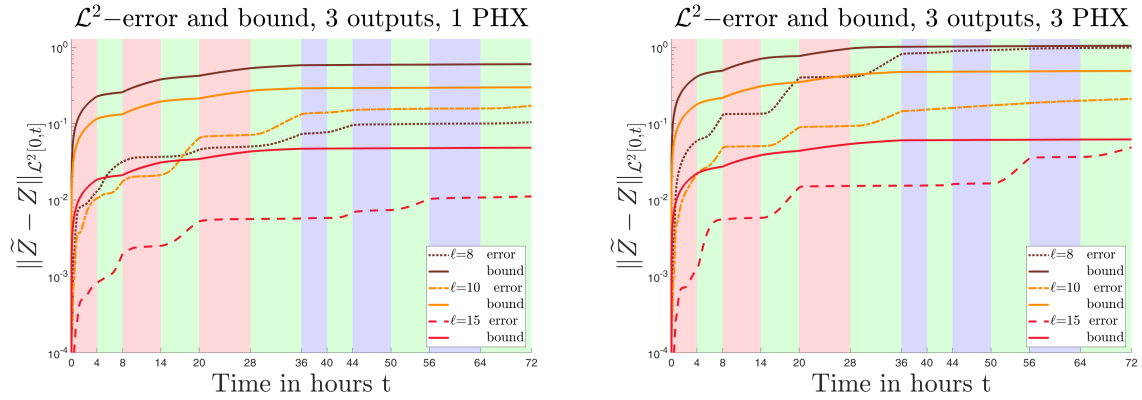


Figure 4.10: Model with three outputs $Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^O)^\top$: $\mathcal{L}^2$-error and error bound for $\ell = \{8, 10, 15\}$. Left: one PHX , Right: three PHXs.

4.3.4 Three Aggregated Characteristics II: $\overline{Q}^M, \overline{Q}^F, \overline{Q}^B$

As already announced above in this experiment, we again consider a model with three outputs but instead of the outlet temperature $\overline{Q}^O$ now the third output is the average temperature at the bottom boundary $\overline{Q}^B$. Hence, the output is $Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^B)^\top$. We recall that the bottom boundary is open and not insulated and the temperature $\overline{Q}^B$ is of crucial importance for the quantification of gains and losses of thermal energy resulting from the heat transfer to the underground from the GS. We refer to Eq. (2.15) and the explanations in Subsec. 2.3.3 and the Robin boundary condition (2.8) modeling that heat transfer from the storage to the underground.

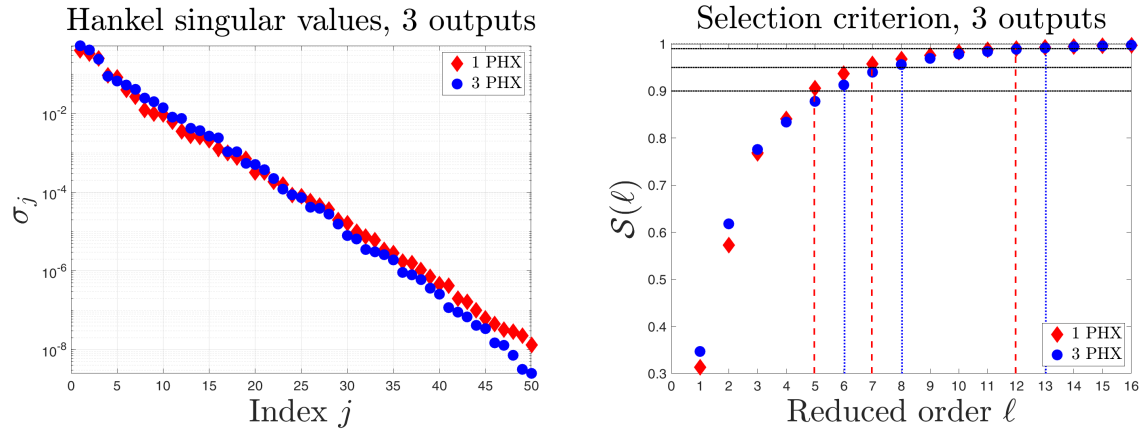


Figure 4.11: Model with three outputs $Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^B)^\top$:
Left: first 50 largest Hankel singular values, Right: selection criterion.

The setting is analogous to Subsec. 4.3.2. The input function g is given in (4.16) and the $3 \times n$ output matrix C is formed by the four rows C^M, C^F, C^B which are given in Subsec. 3.3

Fig. 4.11 shows in the left panel the first 50 largest Hankel singular values, whereas the right panel shows the selection criteria. For both PHX models the first 50 singular values decrease by more than 8 orders of magnitude which is only slightly less than for the case of two outputs. As in the previous experiments the first 20 singular values decrease faster for the model with 1 PHX than for the 3 PHX model. The selection criterion for the model with one PHX is for $\ell \leq 3$ smaller than for 3 PHXs and for $\ell \geq 4$ slightly larger. From the figure and also from

Table 4.1 it can be seen that for reaching threshold levels of $\alpha = \{90\%, 95\%, 99\%\}$ in the 1 PHX case $\ell_\alpha = \{5, 7, 12\}$ states are required while for 3 PHXs one needs $\ell_\alpha = \{6, 8, 13\}$ states, respectively. Thus, for dimension $\ell \geq 13$ an almost perfect approximation of the input-output behavior can be expected.

A comparison with the two-output model in Subsec. 4.3.2 with output $Z = (\overline{Q}^M, \overline{Q}^F)^\top$ shows that the additional third output variable $\overline{Q}^B$ requires only one or two more state variables to ensure the same approximation quality. However, the three-output model considered above in Subsec. 4.3.3 where the third output is the average outlet temperature $\overline{Q}^O$ requires two or three states more in the reduced-order system to ensure the same approximation quality. This shows that $\overline{Q}^B$ is much easier to reconstruct by a reduced order model than $\overline{Q}^O$. This can be explained by the strong dependence of the average outlet temperature $\overline{Q}^O$ on the average temperature $\overline{Q}^F$ in the PHX. Further, for $\overline{Q}^O$ the spatial temperature distribution is averaged over the subdomain $\mathcal{D}^O$ which is much smaller than the corresponding domain $\mathcal{D}^B$ over which the average is taken for $\overline{Q}^B$.

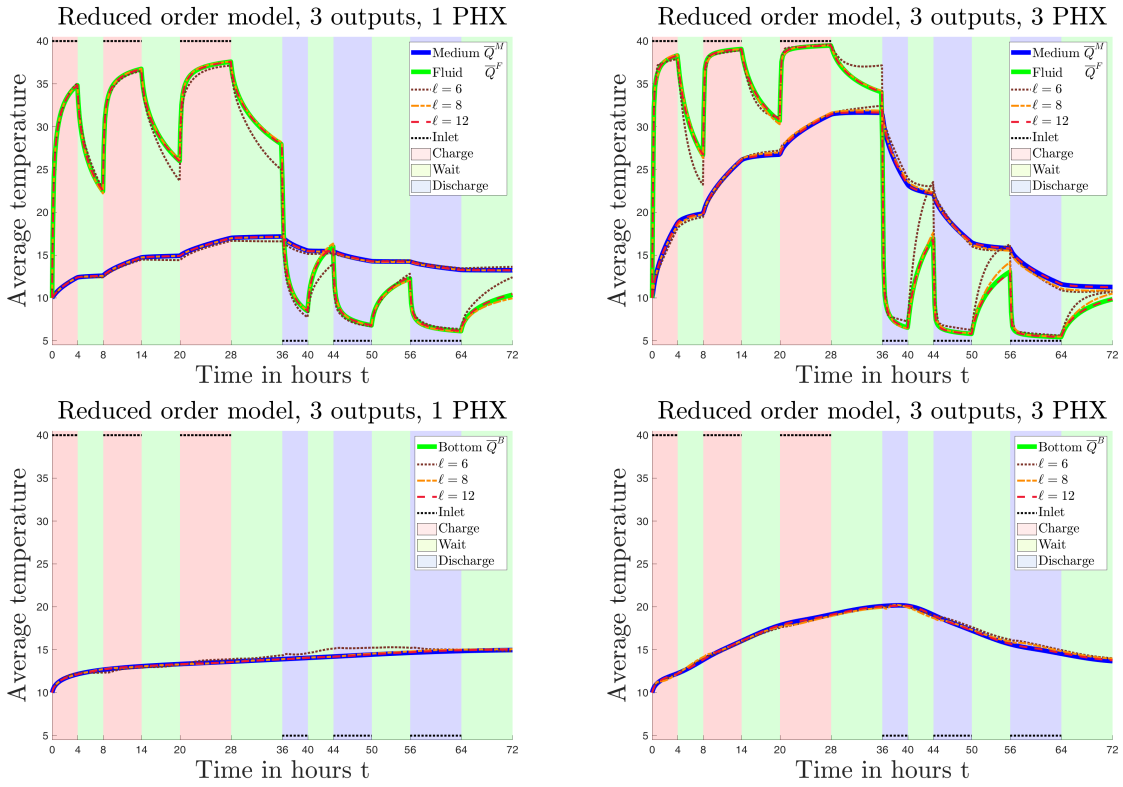


Figure 4.12: Model with three outputs $Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^B)^\top$: Approximation of the output for $\ell = \{6, 8, 12\}$. Top: Average temperatures in the medium $\overline{Q}^M$ and the fluid $\overline{Q}^F$,
Bottom: Average bottom temperature $\overline{Q}^B$,
Left: one PHX , Right: three PHXs .

Fig. 4.12 shows the output variables of the original and reduced-order system which are plotted against time. In the top panels the average temperatures $Z_1(t) = \overline{Q}^M(t)$ and $Z_2(t) = \overline{Q}^F(t)$ in the medium and fluid are drawn as solid blue and green lines, respectively. The bottom panel depicts the average temperature at the bottom boundary $\overline{Q}^B$ by a blue solid line. The reduced-order approximations as drawn for $\ell = \{6, 8, 12\}$. As in the previous experiments the approximation of $\overline{Q}^M$ is much better than for $\overline{Q}^F$. The approximation of the third output variable

$\overline{Q}^B$ is quite good although it represents an average of the spatial temperature distribution over the rather small subdomain $\mathcal{D}^B$ at the bottom boundary. Possible explanations are the relatively small temporal fluctuations of that quantity and the large distance of the bottom boundary to the PHXs, where the charging and discharging generates large temporal and spatial fluctuations.

In Fig. 4.13 we show for the reduced orders ℓ considered above the $\mathcal{L}^2$-error $\|Z-\widetilde{Z}\|_{\mathcal{L}^2[0,t]}$ which plotted against time t together with the error bounds from (4.14). The results are similar to Fig. 4.7 and we refer for the interpretation to the end of Subsec. 4.3.2, Fig. 4.13.

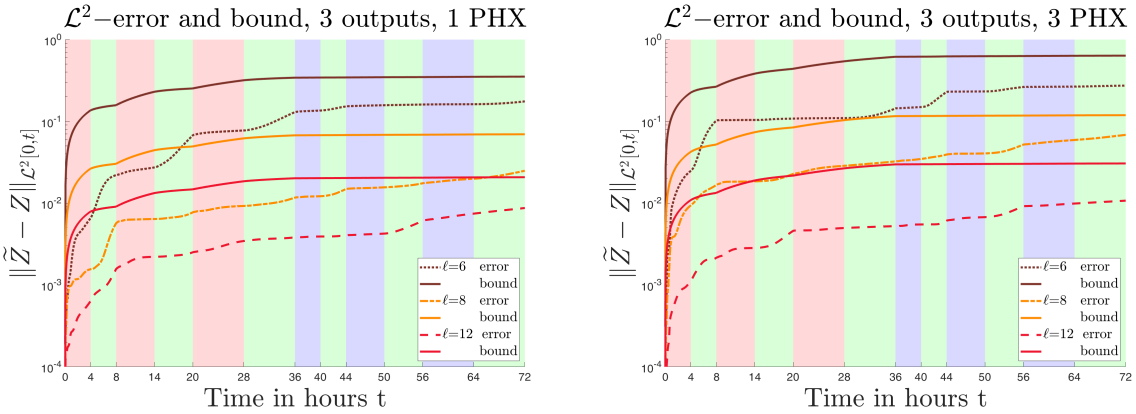


Figure 4.13: Model with three outputs $Z=(\overline{Q}^M,\overline{Q}^F,\overline{Q}^B)^\top$: $\mathcal{L}^2$-error and error bound for $\ell=\{6,8,12\}$. Left: one PHX , Right three PHXs.

4.3.5 Four Aggregated Characteristics: $\overline{Q}^M,\overline{Q}^F,\overline{Q}^O,\overline{Q}^B$

In this last experiment we consider a model with the four-dimensional output $Z=(\overline{Q}^M,\overline{Q}^F,\overline{Q}^O,\overline{Q}^B)^\top$ which contains all of aggregated characteristics appearing in the above experiments.

The setting is analogous to Subsec. 4.3.2. The input function g is given in (4.16) and the $4\times n$ output matrix C is formed by the four rows C^M,C^F,C^O,C^B which are given in Subsec. 3.3.

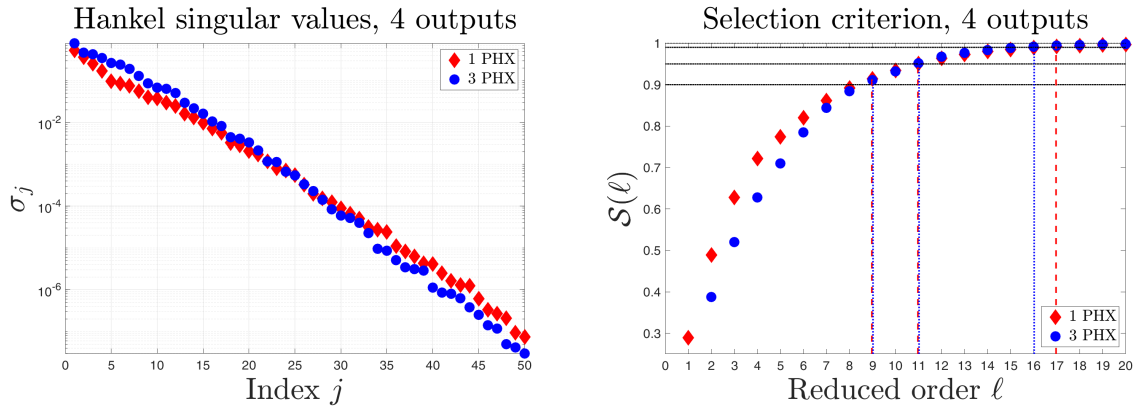


Figure 4.14: Model with four outputs $Z=(\overline{Q}^M,\overline{Q}^F,\overline{Q}^O,\overline{Q}^B)^\top$:
Left: first 50 largest Hankel singular values, Right: selection criterion

Fig. 4.14 shows in the left panel the first 50 largest Hankel singular values, whereas the right panel shows the selection criteria. For both PHX models the first 50 singular values decrease by almost 8 orders of magnitude which is only slightly less than for the case of three outputs. As in

the previous experiments the first 20 singular values decrease faster for the model with one PHX than for the 3 PHX model. The selection criterion for the model with one PHX is for $\ell \leq 10$ larger than for 3 PHXs and for $\ell \geq 11$ slightly smaller. From the figure and also from Table 4.1 it can be seen that for reaching threshold levels of $\alpha = \{90\%, 95\%, 99\%\}$ in the one PHX case $\ell_\alpha = \{9, 11, 17\}$ states are required while for three PHXs one needs $\ell_\alpha = \{9, 11, 16\}$ states, respectively. Thus, for dimension $\ell \geq 17$ an almost perfect approximation of the input-output behavior can be expected. A comparison with the three output model in Subsec. 4.3.3 with output $Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^O)^\top$ shows that the additional fourth output variable $\overline{Q}^B$ requires only one or two more state variables to ensure the same approximation quality. This corresponds to our previous observations for the augmentation of the output $Z = (\overline{Q}^M, \overline{Q}^F)^\top$ of the model considered in Subsec. 4.3.2 by adding as third output the average bottom temperature $\overline{Q}^B$, see Subsec. 4.3.4. There the minimal reduced orders ℓ_α also increase only by 1 ore 2.

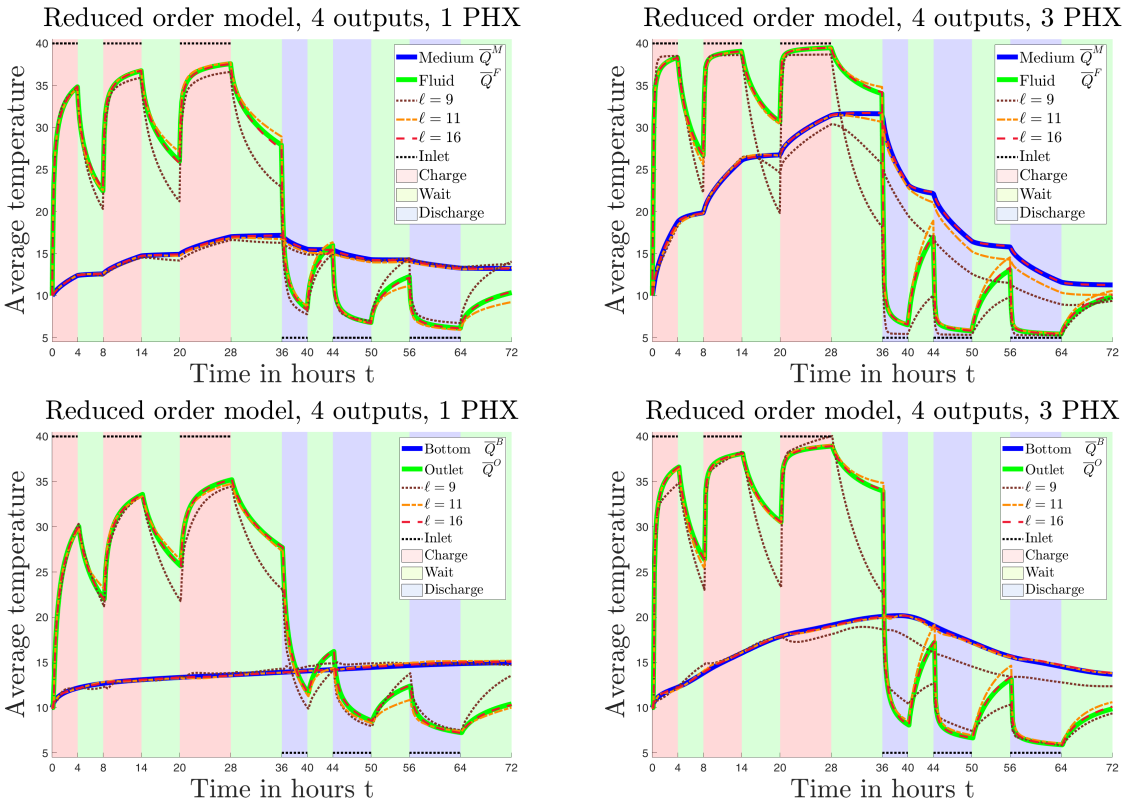


Figure 4.15: Model with four outputs $Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^O, \overline{Q}^B)^\top$: Approximation of the output for $\ell = \{9, 11, 16\}$. Top: Average temperatures in the medium $\overline{Q}^M$ and the fluid $\overline{Q}^F$, Bottom: Average temperatures at the outlet $\overline{Q}^O$ and the bottom boundary $\overline{Q}^B$, Left: one PHX , Right: three PHXs .

Fig. 4.15 depicts the output variables of the original and reduced-order system which are plotted against time. In the top panels the average temperatures $Z_1(t) = \overline{Q}^M(t)$ and $Z_2(t) = \overline{Q}^F(t)$ in the medium and fluid are drawn as solid blue and green lines, respectively. The bottom panel shows the average temperatures at the outlet $Z_3(t) = \overline{Q}^O(t)$ and at the bottom boundary $\overline{Q}^B$ by a blue and green line, respectively. The reduced-order approximations are drawn for $\ell = \{9, 11, 16\}$. The results for the first three outputs $\overline{Q}^M, \overline{Q}^F, \overline{Q}^O$ are similar to the experiment with those three outputs considered in Subsec. 4.3.3. The approximation of the fourth output variable $\overline{Q}^B$ is quite good and comparable to the results in Subsec. 4.3.4. For the model with 3

PHXs we notice some visible errors for the smallest order $\ell = 9$.

Finally, Fig. 4.10 shows for the reduced orders ℓ considered above the $\mathcal{L}^2$-error $\|Z - \widetilde{Z}\|_{\mathcal{L}^2[0,t]}$ which plotted against time t together with the error bounds from (4.14). The results are similar to Fig. 4.7 and we refer for the interpretation to the end of Subsec. 4.3.2.

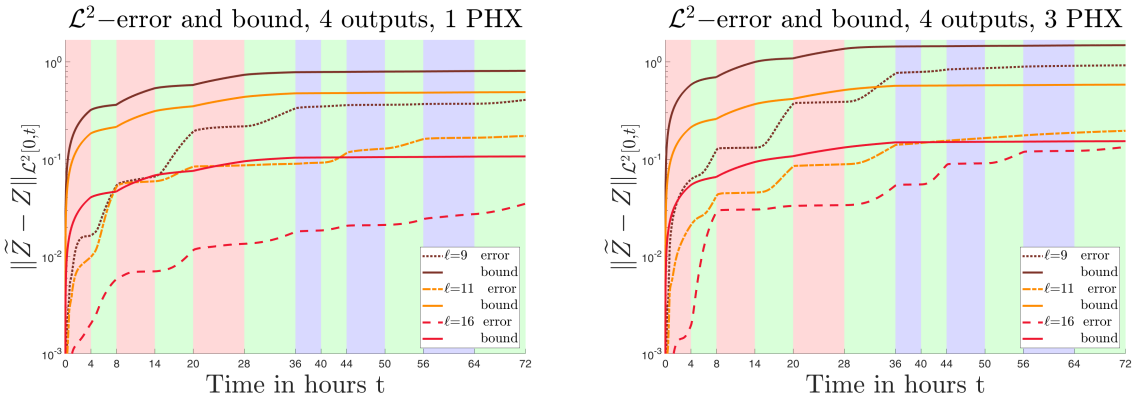


Figure 4.16: Model with four outputs $Z = (\overline{Q}^M, \overline{Q}^F, \overline{Q}^O, \overline{Q}^B)^\top$:
$\mathcal{L}^2$-error and error bound for $\ell = \{9, 11, 16\}$. Left: one PHX , Right: three PHXs.

CHAPTER 5

Continuous-Time Stochastic Optimal Control Problem

Introduction

In this chapter we want to reconsider the optimization problem formulated in Subsec.2.4 arising in the cost-optimal management of the residential heating system equipped with a geothermal energy storage. Since one of the state components, the temperature $Q = Q(t,x,y)$ in the GS, depends not only on time t but also on spatial variables and its dynamics is governed by a PDE, the optimal control problem formulated in Subsec.2.4 appears as a *non-standard* (some components of the state process satisfy PDEs) stochastic optimal control problem with a state process $X = X(t)$ taking values in an infinite-dimensional space. However, using the results of Chapter 4 on model reduction in a first step we can replace Q by its finite-dimensional approximation given by the temperatures $Y_1,\ldots,Y_n$ in the grid points of the mesh used in the finite difference discretization. Then, the state X of the control problem becomes a high-dimensional vector and all components satisfy SDEs and ODEs. This is now a control problem in *standard form*, while we are facing the curse of dimensionality.
A detailed inspection of the mathematical description of the control system and the associated performance criterion shows that we do not need the temperatures Q in every individual grid point but only some aggregated quantities such as the average temperature in the storage, in the PHX and at the outlet boundary of the PHX . Chapter 4 on model reduction has shown that we can find quite accurate approximations of these aggregated quantities from the output of a suitable chosen reduced-order model with a state variable $\widetilde{Y} = (\widetilde{Y}_1,\ldots,\widetilde{Y}_\ell)^\top$ of dimension $\ell \ll n$. Therefore, in a second step we setup a stochastic optimal control problem where instead of the infinite-dimensional original state component $Q = Q(t,x,y)$, we use the ℓ components of $\widetilde{Y}$ as state variables. This chapter is organized as follows. In Sec. 5.1 we reconsider the state process in which we replace the PDE describing the dynamics of GS by a low-dimensional reduced-order system of ODEs. In Sec. 5.2 we reformulate the control problem for a controlled diffusion process describing in Sec.5.1 and derive the associated Hamilton-Jacobi-Bellman equation in Sec. 5.3.

5.1 Dynamics of the Controlled Diffusion Process

In this section we reconsider the stochastic optimal control problem considered in Chapter 2 in which we replace the state Q by ℓ components of $\widetilde{Y}$ whose dynamics is described by a linear system of ODE of dimension ℓ. Note that this modification does not affect the uncontrolled state variables R and F but it modifies the dynamics of the IS which is connected to the GS.

Approximate dynamics of the geothermal storage. The temperature Q in the GS is described by a heat equation. After applying semi-discretization to that PDE we obtained a large n-dimensional system of ODEs. Then, balanced truncation model order reduction applied to the resulting large n-dimensional system of ODEs enables us to reduce its dimension to $\ell \ll n$. For $u(t) \in \overline{U}$, the approximate dynamics of the GS becomes

$$\begin{aligned} &d\widetilde{Y}(t) = \psi_Y(t,\widetilde{Y}(t),u(t))dt, \quad \widetilde{Y}(0) = y_0 \in \mathcal{Y} \subset \mathbb{R}^\ell, \\ &Z(t) = C\widetilde{Y}(t) \end{aligned} \tag{5.1}$$

with

$$\psi_Y(t,y,\nu) = \widetilde{A}(\nu)y + \widetilde{B}(\nu)g(t,\nu), \tag{5.2}$$

where $y = (y^1, y^2, \ldots, y^\ell)^\top \in \mathcal{Y} \subset \mathbb{R}^\ell$ is the reduced-order state and y_0 is a given reduced-order state of the GS at time $t = 0$. Equation (5.1) is an algebraic equation called output equation in which C is some $n_o \times \ell$-matrix described in Subsec. 3.3, whose entries depend on the type of information the manager wishes to get from the system and $Z \in \mathbb{R}^{n_o}$ is a vector of aggregated characteristics. We recall that in the analogous model described in Sec. 3.4 the input function $g(t,\nu) = (C^F y, Q^G(t))^\top$ during the waiting period ($\nu = u^W$). Then the control-dependent reduced-order system matrix $\widetilde{A}(\nu)$ and input matrix $\widetilde{B}(\nu)$ are such that

$$\widetilde{A}(\nu) = \begin{cases} \widetilde{A}_\ell & \nu \in \{u^C, u^D\}, \\ \widetilde{A}_\ell + \widetilde{B}^1_\ell C^F & \text{otherwise}, \end{cases} \quad \widetilde{B}(\nu) = \begin{cases} \widetilde{B}_\ell & \nu \in \{u^C, u^D\}, \\ \widetilde{B}^2_\ell & \text{otherwise}, \end{cases} \tag{5.3}$$

with $\widetilde{A}(\nu) \in \mathbb{R}^{\ell\times\ell}$, $\widetilde{B}_\ell = (\widetilde{B}^1_\ell, \widetilde{B}^2_\ell) \in \mathbb{R}^{\ell\times 2}$ the constant system and input matrices, respectively. Indeed, the matrices $\widetilde{A}(\nu)$ and $\widetilde{B}(\nu)$ depend on time through the control ν which changes with time. The input function g is given by

$$g(t,\nu) = \begin{cases} g_1(t,\nu) & \nu = u^C, u^D, \\ Q_G(t) & \text{otherwise}, \end{cases} \tag{5.4}$$

where $g_1(t,\nu) = (Q^I_\nu(t), Q_G(t))^\top \in \mathbb{R}^2$. Here $Q_G(t)$ is the underground temperature and $Q^I_\nu(t)$ the inlet temperature of the GS given by

$$Q^I_\nu(t) = \begin{cases} Q^I_C & \nu = u^D \quad \text{(charge GS)}, \\ Q^I_D & \nu = u^C \quad \text{(discharge GS)}. \end{cases}$$

Typically, $Q^I_C > Q^I_D$. The manager or controller of such a storage has to make sure that the temperature in the GS is always in some interval called *comfort zone*. Let the vector of the aggregated characteristics be given by $Z = (Z_1, Z_2, \ldots, Z_{n_o})^\top \in \mathbb{R}^{n_o}$, where for all $i =$

$1,\dots,n_o$, $Z_i(t) = C^\dagger \widetilde{Y}(t)$ is the average temperature in given domain in the GS. For example $Z(t) = (Z_1(t), Z_2(t), Z_3(t), Z_4(t))^\top = C\widetilde{Y}$, with $C = (C^M, C^F, C^O, C^B)^\top \in \mathbb{R}^{4\times l}$ if we are interested in 4 aggregated characteristics such as the average temperature in the medium (soil without the pipe), in the pipe, at the outlet of the pipe and at the bottom of the storage. Here $Z_1(t) = \overline{Q}^M(t) = C^M\widetilde{Y}(t)$ is the average temperature in the medium, $Z_2(t) = \overline{Q}^F(t) = C^F\widetilde{Y}(t)$ the average temperature in the pipe, $Z_3(t) = \overline{Q}^O(t) = C^O\widetilde{Y}(t)$ the average temperature at the outlet boundary of the pipe, and $Z_4(t) = \overline{Q}^B(t) = C^B\widetilde{Y}(t)$ is the average temperature at the bottom boundary of the storage. The quantity Z_3 is used to calculate how much energy we can extract from the storage through the pipe and Z_4 is used to calculate the amount of energy we gain or loose through the open bottom boundary. In order to keep the temperature in the comfort zone, we have to impose constraints on the reduced state variable $\widetilde{Y}(t)$ or on the aggregated characteristics. It turns out that it is enough to impose the constraints on the average temperature in the GS since a non-homogeneous spatial temperature will be averaged after a while due to the diffusion. The storage manager has to ensure that the average temperature in the GS is always in the *comfort zone*, say $\overline{Q}^M = C^M\widetilde{Y}(t) \in \mathcal{Q} = [\underline{q}, \overline{q}]$.

State dynamics. With some abuse of notation we denote again by $\mathcal{X}$ the finite-dimensional state space and by X the state process of the new (approximate reduced-order) control problem, where we want to emphasize the dependence on the control u by writing $X = X^u$. The state X^u can be decomposed into two groups of state variables. The first group consists of controlled processes with deterministic dynamics described by ODEs. These are the temperature in the IS, together with the state variables of the reduced-order system. We denote this $(1+\ell)$-dimensional state component by

$$\overline{X}^u = (P, \widetilde{Y}^1, \widetilde{Y}^2, \dots, \widetilde{Y}^\ell)^\top.$$

The second group contains variables which do not depend on the control and also not on the variables of the first group. Their dynamics is subject to exogenously given seasonality and uncertainties and described by SDEs. In our model these are the fuel/electricity price F and the residual demand R. We denote this 2-dimensional second state component of uncontrolled variables by

$$\widehat{X} = (R, F)^\top.$$

The state process is now $X = \begin{pmatrix} \widehat{X} \\ \overline{X}^u \end{pmatrix} = (R, F, P, \widetilde{Y}^1, \widetilde{Y}^2, \dots, \widetilde{Y}^\ell)^\top \in \mathcal{X} \subset \mathbb{R}^{l+3}$, where

$$\mathcal{X} = \mathcal{R} \times \mathcal{F} \times \mathcal{P} \times \mathcal{Y} = \{(r, f, p, y) \mid r \in \mathcal{R},\ f \in \mathcal{F},\ p \in \mathcal{P},\ y \in \mathcal{Y}\},$$

with the dynamics given by a stochastic differential equation

$$dX(t) = \mu(t, X(t), u(t))dt + \sigma(t)dW(t),\ \ X(0) = x_0 = (r_0, f_0, p_0, y_0^\top)^\top \in \mathcal{X}. \tag{5.5}$$

where the drift coefficient $\mu : [0,T] \times \mathcal{X} \times \mathcal{U} \to \mathbb{R}^{3+l}$ and the diffusion coefficient $\sigma : [0,T] \to \mathbb{R}^{(3+l)\times 2}$ are giving by

$$\mu(t, x, \nu) = \begin{pmatrix} \beta_R(\mu_R(t) - r) \\ \beta_F(\mu_F(t) - f) \\ \widetilde{\psi}_P(r, f, p, y, \nu) \\ \psi_Y(t, y, \nu) \end{pmatrix},\ \ \sigma(t) = \begin{pmatrix} \sigma_R(t) & 0 \\ 0 & \sigma_F(t) \\ \mathbf{0}_{(l+1)\times 1} & \mathbf{0}_{(l+1)\times 1} \end{pmatrix} \in \mathbb{R}^{(3+l)\times 2},$$

$W(t) = (W_R(t), W_F(t))^\top$, $t \in [0,T]$ is a two-dimensional standard Brownian motion, and the initial state x_0 at time $t = 0$ is given. Here, $\widetilde{\psi}_P(r,f,p,y,\nu) = \psi_p(r,f,y,\nu) - \gamma(p - P_{amb})$, where a function ψ_p is given by (2.18), and ψ_Y is given by equation (5.2). We require that for all $t \in [0,T]$, $\sigma_R(t) > \underline{\sigma}_R > 0$ and $\sigma_F(t) > \underline{\sigma}_F > 0$. The state dynamics (5.5) can be written componentwise as follows

$$\begin{aligned}
dR(t) &= \beta_R(\mu_R(t) - R(t))dt + \sigma_R(t)dW_R(t), && R(0) = r_0,\\
dF(t) &= \beta_F(\mu_F(t) - F(t))dt + \sigma_F(t)dW_F(t), && F(0) = f_0,\\
dP(t) &= (\psi_p(R(t),F(t),\widetilde{Y}(t),u(t)) - \gamma(P(t) - P_{amb}))dt, && P(0) = p_0,\\
d\widetilde{Y}(t) &= \psi_Y(t,\widetilde{Y}(t),u(t))dt, && \widetilde{Y}(0) = y_0.
\end{aligned}$$

As defined in Subsec. 2.4.1 the state can be decomposed into two parts: $X^u = (\widehat{X}(t), \overline{X}^u(t))^\top$ where $\widehat{X} \in \mathcal{R} \times \mathcal{F} \subset \mathbb{R}^2$ is an exogenous state variable and $\overline{X}^u \in \mathcal{P} \times \mathcal{Y} \subset \mathbb{R}^{1+l}$ is an endogenous state variable. The uncontrolled state $\widehat{X} = (R,F)^\top \in \mathcal{R} \times \mathcal{F}$ satisfies the SDE

$$d\widehat{X}(t) = \widehat{\mu}(t,\widehat{X}(t))\,dt + \widehat{\sigma}(t)d\widehat{W}(t), \quad \widehat{X}(0) = \widehat{x}_0 = (r_0,f_0)^\top \in \mathcal{R} \times \mathcal{F},$$

where the uncertainty $\widehat{W} = (W_R, W_F)^\top$, the drift coefficient $\widehat{\mu} : [0,T] \times \mathcal{R} \times \mathcal{F} \to R^2$ and the volatility matrix $\widehat{\sigma} : [0,T] \times \mathcal{R} \times \mathcal{F} \to \mathbb{R}^{2\times 2}$ are defined for $\widehat{x} = (r,f)$ as

$$\widehat{\mu}(t,\widehat{x}) = \begin{pmatrix} \beta_R(\mu_R(t) - r) \\ \beta_F(\mu_F(t) - f) \end{pmatrix} \in \mathbb{R}^2, \qquad \widehat{\sigma}(t) = \begin{pmatrix} \sigma_R(t) & 0 \\ 0 & \sigma_F(t) \end{pmatrix} \in \mathbb{R}^{2\times 2}. \tag{5.6}$$

For deterministic known fuel price F, the exogenous state variable is $\widehat{X} = R \in \mathcal{R}$ and given by Equation (2.1) and $\widehat{\mu}(t,r) = \beta_R(\mu_R(t) - r)$, $\widehat{\sigma}(t) = \sigma_R(t) \in \mathbb{R}$.
The controlled state $\overline{X}^u = (P^u, \widetilde{Y}^1, \widetilde{Y}^2, \ldots, \widetilde{Y}^\ell)^\top$ satisfies the system of ODEs

$$d\overline{X}^u(t) = \overline{\mu}(t,\widehat{X}(t),\overline{X}^u(t),u(t))\,dt, \qquad \overline{X}^u(0) = \overline{x}_0 = (p_0, y_0^\top)^\top \in \mathcal{P} \times \mathcal{Y}$$

where

$$\overline{\mu}(t,\widehat{x},\overline{x},\nu) = \begin{pmatrix} \psi_P(\widehat{x},\overline{x},\nu) \\ \psi_Y(t,y^1,y^2,\ldots,y^\ell,\nu) \end{pmatrix} \in \mathbb{R}^{1+\ell}.$$

Note that the drift coefficient $\overline{\mu}$ of the controlled variables depends on the control as well as of the second state component $\widehat{X}$. The latter represents the coupling of the controlled with the uncontrolled states. The drift coefficient $\widehat{\mu}$ and diffusion coefficient $\widehat{\sigma}$ in the SDE for $\widehat{X}$ depend neither on the controlled variables nor on the control u.

Assumption 5.1.1 We assume that the control is of Markov type, i.e, it can be written in the form $u(t) = \widetilde{u}(t,X^u(t)), t \in [0,T]$, where $\widetilde{u} : [0,T] \times \mathcal{X} \to \mathbb{R}^q$ is a (Borel) measurable function.

This property states that the control at any time t depends only on the current time and also on the current state $X^u(t)$ but does not depend on past history. In addition, it is assumed that the control takes values in some subset $\mathcal{U} \subset \mathbb{R}^q$, called the set of feasible controls defined by (2.20). The subset $\mathcal{U}$ is often assumed to be compact and convex, but in most practical cases the set $\mathcal{U}$ is non-convex. The latter fits to the control problem we consider in this thesis. Indeed, the set $\mathcal{U} \subset \mathbb{N}$ only takes finite number of elements. Next we state under which condition the SDE (5.5)

is well-posed and if it has a unique solution. A solution of (5.5) with initial data $X^u(0) = x_0$ can be interpreted as a solution of the stochastic integral equation

$$X^u(t) = x_0 + \int_0^t \mu(s, X^u(s), u_s)ds + \int_0^t \sigma(s)dW_s, \quad 0 \leq s \leq t \leq T. \tag{5.7}$$

Definition 5.1.2 (Strong solution) A stochastic process $(X^u(t))_{t\in[0,T]}$ is called a strong solution to the SDE (5.5) if it is continuous, $\mathbb{G}-$adapted and satisfies equation (5.7) almost surely (i.e., with probability 1). A solution is said to be local if it exists up to a stopping time.
A solution $(X^u(t))_{t\in[0,T]}$ is said to be unique if any other solution $(\widetilde{X}^u(t))_{t\in[0,T]}$ is indistinguishable from $(X^u(t))_{t\in[0,T]}$, i.e., $\mathbb{P}\{X^u(t) = \widetilde{X}^u(t),\ \forall t \in [0,T]\} = 1$.

Remark 5.1.3 In the dynamics of the diffusion process $(X^u(t))_{t\in[0,T]}$, given by equation (5.5), the drift coefficient depends explicitly on time and on the control. However, the diffusion coefficient does not depend on the control but depends only on time. In general the diffusion coefficient may also depend on state and on the control ($\sigma = \sigma(t, X^u, u)$) but such cases are not considered in this thesis, we assume that we can only control the drift and the diffusion is uncontrolled. The case involving time-dependent coefficients is only suitable for finite time horizon problems since the typical requirement for infinite time horizon problems is that the coefficients should not be explicitly time-dependent.

In order to quote the well known result about the existence and the uniqueness of the solution of the SDE, we have to make some assumptions on the coefficients μ and σ where we consider the general case with $\sigma = \sigma(t,x)$.

Assumption 5.1.4 Let $\mu : [0,T] \times \mathbb{R}^d \times \mathcal{U} \to \mathbb{R}^d$ and $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be to measurable functions for which the following apply:

- For all $t \in [0,T]$, and $x \in \mathbb{R}^d$ the function $\mu(t,x,\nu)$ is continuous in $\mathcal{U}$.
- For all $t \in [0,T]$ and $\nu \in \mathcal{U}$ there exists a constant $M_1 > 0$ such that
$$|\mu(t,0,\nu)| + \|\sigma(t,0)\| \leq M_1.$$
- Growth condition: for all $x \in \mathbb{R}^d, \nu \in \mathcal{U}$ and $t \in [0,T]$, there exists a constant $M_2 > 0$ such that
$$|\mu(t,x,\nu)| + \|\sigma(t,x)\| \leq M_2(1+|x|).$$
- Lipschitz condition: for all $x, y \in \mathbb{R}^d$, $t \in [0,T]$ and $\nu \in \mathcal{U}$, there exists a constant $M_3 > 0$, such that
$$|\mu(t,x,\nu) - \mu(t,y,\nu)| + \|\sigma(t,x) - \sigma(t,y)\| \leq M_3|x-y|.$$

where $|.|$ is the Euclidean norm in $\mathbb{R}^d$ and $\|.\|$ is the Frobenius norm, i.e.,

$$\|\widetilde{\sigma}\| = \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{m} |\widetilde{\sigma}_{ij}|^2}.$$

The following lemma gives the well-known existence and uniqueness result for SDEs, which is proven in Mao [74, Sec.2.3].

Lemma 5.1.5 (Classical solution) Assume that the drift μ and the diffusion coefficient σ satisfy the assumption (5.1.4) with σ is non-degenerated. Then, there exist a unique strong solution to the Itô stochastic differential equation (5.7).

In most practical problems the conditions in assumption 5.1.4 are not fulfilled, i.e., the drift may be discontinuous, non-Lipschitz and may also be unbounded; the diffusion coefficient may also be non-Lipschitz, degenerated and unbounded. In such extreme cases ensuring the existence of the solution is more demanding and not straightforward. If Assumption 5.1.4 fails, we can still get a unique solution. Such cases are given in the following remark.

Remark 5.1.6

1) If the drift μ is measurable and bounded but not Lipschitz, then there exists a unique strong solution to (5.5) if the diffusion coefficient σ is bounded and Lipschitz and the infinitesimal generator of the SDE (5.5) is uniformly elliptic, i.e., $\exists c > 0$ such that $\forall x \in \mathbb{R}^d$ and $\forall \zeta \in \mathbb{R}^d$, we have $\zeta^\top \sigma\sigma^\top \zeta \geq c|\zeta|^2$, see the work by Zvonkin [132] for one dimensional case and Veretenikov [122] for multidimensional case. Furthermore, if the drift coefficient is partially Lipschitz, i.e., Lipschitz in some components and non-Lipschitz in others, then there exists a strong solution when the diffusion coefficient is uniformly elliptic only in the components in which the drift is non-Lipschitz, we refer to the result by Veretenikov [121].

2) If the drift is only locally integrable and the diffusion σ non-degenerated, then there exists a strong solution to (5.5), see Zhang [131].

3) If the drift coefficient is discontinuous but increasing in each variable and the diffusion coefficient σ a Lipschitz continuous, then the existence of continuous strong solution is ensure by the work or Halidias and Kloeden [54].

4) If the drift coefficient μ is discontinuous, non-Lipschitz (may also be unbounded and decreasing in each variable), the diffusion coefficient σ singular and the infinitesimal generator is not uniformly elliptic, then under some conditions the existence of the unique maximal local solution to equation (5.5) can be ensured by the work by Leobacher et al. [68].

5) Let the drift coefficient μ be non-Lipschitz but measurable and bounded, the diffusion coefficient σ degenerated but Lipschitz with respect to the non-degenerated components. Further, assume that μ and σ are twice continuously differentiable functions with bounded derivatives with respect to the degenerated components. Then, without considering any regularization in the degenerated components, there exists a unique strong solution to equation (5.5), see Zvonkin [132].

Note that for our problem, the state variable P is a piece-wise linear continuous function and the set of feasible controls $\mathcal{U} \subset \mathbb{N}$. Then, the drift μ maybe discontinuous in P and $\widetilde{Y}$, and even non-Lipschitz with respect to some components. In addition, the diffusion coefficient is degenerated but locally Lipschitz with respect to the non-degenerated components. In this case the existence of the solution to SDE (5.5) can be guaranteed by items 4) and 5) in the above remark 5.1.6.

5.2 Stochastic Optimal Control Problem

5.2.1 Performance Criterion

Let the control process $u = (u(t))_{t\in[0,T]}$, $u(t) \in \mathcal{U}(t,x)$ be given, the initial residual demand $R(t) = r$, the initial fuel price $F(t) = f$, the initial temperature in the IS $P(t) = p$, and the initial temperature in the GS generated by the reduced-order state $\widetilde{Y}(t) = y$ a time $t \in [0,T]$. The cost function at time t, $J : [0,T] \times \mathcal{X} \times \mathcal{U} \to \mathbb{R}$ is the expected aggregated cost over the time interval $[0,T]$ given by

$$J(t,x;u) = \mathbb{E}_{t,x}\left[\int_t^T \Psi(s,X^u(s),u(s))ds + \phi(X^u(T))\right] \tag{5.8}$$

for $X^u = (R,P,\widetilde{Y}^1,\widetilde{Y}^2,\ldots,\widetilde{Y}^\ell)^\top \in \mathcal{X} \subset R^{l+3}$. Here, $\mathbb{E}_{t,x}[\cdot]$ is the conditional expectation given that at time t the state $X^u(t) = x = (r,f,p,y^\top)^\top \in \mathcal{X}$, Ψ is the running cost and ϕ is the terminal cost described in Subsec. 2.4. Note that the terminal cost only applies to finite time horizon problems. For infinite time horizon problems, the terminal cost does not apply but we have to take into account some discounting factor for the running costs. We make the following assumptions on the functions Ψ and ϕ.

Assumption 5.2.1 Let $\Psi : [0,T] \times \mathcal{X} \times \overline{\mathcal{U}} \to \mathbb{R}$ and $\phi : \mathcal{X} \to \mathbb{R}$ be two measurable functions for which the following apply

- Growth condition: for all $x \in \mathcal{X}$, $t \in [0,T]$ and $\nu \in \overline{\mathcal{U}}$, there exists a constant $M_4, M_5 > 0$ such that

$$\sup_{\nu\in\overline{\mathcal{U}}} |\Psi(t,x,\nu)| \le M_4(1+|x|), \quad |\phi(x)| \le M_5(1+|x|)$$

- Lipschitz condition: for all $x,y \in \mathcal{X}$, $t \in [0,T]$ and $\nu \in \overline{\mathcal{U}}$, there exists a constant $M_6, M_7 > 0$, such that

$$\sup_{\nu\in\overline{\mathcal{U}}} |\Psi(t,x,\nu) - \Psi(t,y,\nu)| \le M_6|x-y|, \quad |\phi(x) - \phi(y)| \le M_7|x-y|.$$

The running cost and ϕ is the terminal cost described in Subsec. 2.4.3 satisfy Assumption. 5.2.1.

Admissible control. We denote by $\mathcal{A}(x)$ the class of admissible controls, consisting of Markovian control processes u being progressively measurable w.r.t. the filtration $\mathbb{G}$, satisfying certain integrability conditions and control constraints (described above) such that the controlled state X^u takes at any time t values in the prescribed state space $\mathcal{X}$, i.e.,

$$\begin{aligned}\mathcal{A}(x) = \Big\{(u(t))_{t\in[0,T]} \mid\ & u \text{ is } \mathbb{G}\text{-progressively measurable}, u_t = \widetilde{u}(t,X^u(t)) \text{ for all} \\ & t \in [0,T], u_t \in \mathcal{U}(t,x) \text{ for all } (t,x) \in [0,T]\times\mathcal{X},\ X^u(t) \in \mathcal{X},\ t \in [0,T], \\ & \text{and } \mathbb{E}_{t,x}\left[\int_t^T |\Psi(s,X^u(s),u(s))|\,ds + |\phi(X^u(T))|\right] < \infty\Big\}.\end{aligned} \tag{5.9}$$

5.2.2 Optimal Control Problem

The objective is to minimize the performance criterion (5.8) over all admissible controls (5.9). We define the value function for all $(t,x) \in [0,T] \times \mathcal{X}$ by

$$V(t,x) = \inf_{u \in \mathcal{A}(x)} J(t,x;u). \tag{5.10}$$

A control $u^* \in \mathcal{A}(x)$ is called optimal control if $V(t,x) = J(t,x;u^*)$.

Note that very often one is interested in the value function at the initial time $t = 0$. However, solving the optimization problem (5.9) with dynamic programming techniques requires embedding it into a family of optimization problems and solve for each pair $[0,T] \times \mathcal{X}$.
Next we consider the solution of the control problem using the dynamic programming and derive the associated Hamilton-Jacobi-Bellman equation.

5.3 Hamilton-Jacobi-Bellman Equation

In this subsection we use the dynamic programming principle (DPP) initiated by Bellman in the 1950s to solve the control problem. It is a fundamental principle in the theory of stochastic control which states that an optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision, see Bellman [15].
We recall that our goal is to minimize the performance criterion over all admissible controls. To obtain an optimal control on the whole time interval $[t,T]$ the basic idea is to split the optimization problem into two parts. First, search for an optimal control from time $\tau_h \in [t,T]$ given the state value $X^u(\tau_h)$, i.e., compute the value function $V(\tau_h, X^u(\tau_h))$. Second, minimize over all admissible controls on $[t,\tau_h]$ the quantity

$$\mathbb{E}_{t,x}\left[\int_t^{\tau_h} \Psi(s,X^u(s),u(s))ds + V(\tau_h,X^u(\tau_h))\right].$$

This principle is called the dynamic programming principle and is formulated as follows:

Theorem 5.3.1 (Dynamic programming principle) Let $(t,x) \in [0,T] \times \mathcal{X}$ as $\mathcal{T}_{t,T}$ the set of stopping times valued in $[t,T]$. Then, it holds

$$\begin{aligned} V(t,x) &= \inf_{u \in \mathcal{A}(x)} \sup_{\tau_h \in \mathcal{T}_{t,T}} \mathbb{E}_{t,x}\left[\int_t^{\tau_h} \Psi(s,X^u(s),u(s))ds + V(\tau_h,X^u(\tau_h))\right] \\ &= \inf_{u \in \mathcal{A}(x)} \inf_{\tau_h \in \mathcal{T}_{t,T}} \mathbb{E}_{t,x}\left[\int_t^{\tau_h} \Psi(s,X^u(s),u(s))ds + V(\tau_h,X^u(\tau_h))\right]. \end{aligned}$$

The proof of this theorem can be found in Pham [87, Chapter 3, Theorem 3.3.1], Fleming and Soner [45, Chapter 4], Fleming and Rishel [44, Chapter 4], Oksendal [80, Chapter 11], Touzi [116, Chapter 1].
A stronger version of this DPP is given as

$$V(t,x) = \inf_{u \in \mathcal{A}(x)} \mathbb{E}_{t,x}\left[\int_t^{\tau_h} \Psi(s,X^u(s),u(s))ds + V(\tau_h,X^u(\tau_h))\right], \tag{5.11}$$

for any stopping time valued in $[t,T]$. From the DPP, we formally derive the dynamic programming equation, which is also called the Hamilton-Jacobi-Bellman (HJB) equation, by sending the stopping time τ_h to t. Therefore, the HJB equation is the infinitesimal version of the DPP describing the local behavior of the value function.

Hamilton-Jacobi-Bellman equation. After reformulating the optimal control problem obtained by replacing the PDE describing the spatial distribution of the temperature in the GS with a low-dimensional system of ODE resulting from model reduction as described in Chapter 4, we now consider its solution using dynamic programming techniques. The basic idea of this approach is to embed the control problem into a family of control problems by varying the initial state values of the controlled diffusion process which leads to a nonlinear partial differential equation (PDE) of second order called the Hamilton-Jacobi-Bellman (HJB) equation that we present below. We begin with the generator of the state process X^{ν} for a constant strategy $u(t) = \nu$. The generator of the state process reads as

$$\mathcal{L}^{\nu} = \overline{\mathcal{L}}^{\nu} + \widehat{\mathcal{L}}, \tag{5.12}$$

where for $G \in C^2(\mathcal{X})$, $x = (\widehat{x}, \overline{x})$

$$\begin{aligned}
\overline{\mathcal{L}}^{\nu} G &= \overline{\mu}^{\top}(t,x,\nu) D_{\overline{x}} G = \psi_P(t,x,\nu)\frac{\partial G}{\partial p} + \sum_{i=1}^{\ell} \psi_Y^i(t,x,\nu)\frac{\partial G}{\partial y_i} \\
\widehat{\mathcal{L}}\, G &= \widehat{\mu}^{\top}(t,\widehat{x}) D_{\widehat{x}} G + \frac{1}{2}\mathrm{tr}[(\widehat{\sigma}\widehat{\sigma}^{\top})(t) D^2_{\widehat{x}} G] \\
&= \widehat{\mu}_R(t,\widehat{x})\frac{\partial G}{\partial r} + \frac{1}{2}(\widehat{\sigma}_R(t))^2\frac{\partial^2 G}{\partial r^2} + \widehat{\mu}_F(t,\widehat{x})\frac{\partial G}{\partial f} + \frac{1}{2}(\widehat{\sigma}_F(t))^2\frac{\partial^2 G}{\partial f^2} \\
&= \widehat{\mathcal{L}}_R\, G + \widehat{\mathcal{L}}_F\, G.
\end{aligned}$$

Applying the dynamic programming principle given in Theorem 5.3.1 yields the following.

Theorem 5.3.2 (HJB equation) The value function (2.4.4) satisfies the Hamilton-Jacobi-Bellman equation given by

$$\frac{\partial}{\partial t} V + \widehat{\mathcal{L}}\, V + \inf_{\nu \in \mathcal{U}(t,x)} \left\{ \overline{\mathcal{L}}^{\nu} V + \Psi(t,x,\nu) \right\} = 0, \quad (t,x) \in [0,T) \times \mathcal{X} \tag{5.13}$$

$$V(t,x) = \phi(x).$$

For a proof see Appendix C.1.
A candidate for the optimal decision rule $\widetilde{u}^{\star}$ control is obtained from the solution of the pointwise optimization problem in the above HJB equation and reads as

$$\widetilde{u}^{\star}(t,x) = \operatorname*{argmin}_{\nu \in \mathcal{U}(t,x)} \left\{ \overline{\mathcal{L}}^{\nu} V(t,x) + \Psi(t,x,\nu) \right\}.$$

Now we are going to discuss under which condition equation (5.13) has a unique solution and in which case the solution coincides with the solution of the optimal control problem. With enough regularity on V, the so-called verification theorem guarantees that the solution of the HJB equation coincides with the solution of the optimal control problem.

Theorem 5.3.3 (Verification theorem) Let w be a function in $C^{1,2}([0,T]\times\mathcal{X})\cap C([0,T]\times\mathcal{X})$, satisfying the quadratic growth condition

$$|w(t,x)| \leq M_7(1+|x|^2), \quad \forall\,(t,x)\in[0,T]\times\mathcal{X},$$

for some constant M_7. Assume that w is a solution to the HJB equation

$$\frac{\partial}{\partial t}w + \widehat{\mathcal{L}}\,w + \inf_{v\in\mathcal{U}(t,x)}\left\{\overline{\mathcal{L}}^{v}w + \Psi(t,x,v)\right\} = 0, \quad (t,x)\in[0,T)\times\mathcal{X}$$

$$w(T,x) = \phi(x), \quad x\in\mathcal{X}.$$

Further, assume that there exists a measurable function $\widetilde{u}:[0,T]\times\mathcal{X}\to\mathcal{U}$ such that the SDE

$$dX(s) = \mu(s,X(s),\widetilde{u}(s,X(s)))ds + \sigma(s)dW(s),\;\; X(t) = x\in\mathcal{X},\;\; s\in[t,T],$$

admits a unique solution $\widetilde{X}$, the process $(\widetilde{u}(t))_{s\in[t,T]}$ lies in $\widetilde{A}(x)$ and the following holds

$$\widetilde{u}(t,x) = \operatorname*{argmin}_{v\in\mathcal{U}(t,x)}\left\{\overline{\mathcal{L}}^{v}w(t,x) + \Psi(t,x,v)\right\}.$$

Then, w coincides with the value function, i.e., $w(t,x) = V(t,x)$ on $[0,T]\times\mathcal{X}$ and $\widetilde{u}$ is the optimal control.

For a proof see Pham [87, Chapter 3, Theorem 3.5.2].

Remarks on the solution to the HJB equation

To derive the manager's charging and discharging decision, the optimal control problem (5.10) had to be solved. Due to the complexity of the problem, the corresponding HJB equation (5.13) must be solved for the value function and the corresponding optimal strategy. To ensure that a smooth solution to the HJB equation coincides with the value function and the candidate for the optimal decision rule (5.3) is indeed the optimal control, the classical method is to use the so-called verification theorem. When there is a sufficiently regular (classical) solution to the HJB equation with appropriate terminal condition, then a verification theorems states that the solution of the HJB equation is the value function of the stochastic optimal control problem and the optimal control as Markov control in the feedback form can be chosen. For the application of verification theorems from the literature (see for example Fleming and Soner [45, Chapter 4], Fleming and Rishel [44, Chapter 6, Sec. 4], Pham [87, Chapter 3] or Touzi [116, Chapter 1]) it is required that the solution to the HJB equation is regular, i.e., twice continuously differentiable with respect to the space component and once continuous and differentiable with respect to time, with bounded derivatives.

In general, the HJB equation does not have a sufficiently smooth solution, in particular, when the generator of the SDE is not uniformly elliptic, i.e., when there is no constant $\lambda > 0$ such that $\forall x\in\mathcal{X}$ and $\forall\zeta\in\mathcal{X}$, we have $\zeta^{\top}\sigma\sigma^{\top}\zeta \geq \lambda|\zeta|^2$. In such cases one must resort to solutions which hold in some weaker sense, in particular, viscosity solutions. Possible references for more comprehensive details on the topic of viscosity solution are Pham [87, Chapter 4], Fleming and Soner [45, Chapter 5] or Touzi [116, Chapter 2].

Note that the generator $a = \sigma\sigma^{\top}$ of the SDE (5.5) is singular (not of full rank) since the

noise does not enter certain components of the system, specially the controlled components. Then, uniform ellipticity of the differential operator cannot be guaranteed and the dynamic programming equation (5.13) is degenerated or non parabolic. As a consequence, we no longer know that a regular solution exists as required in the verification theorem. Therefore, the existence of the classical solution of the HJB equation (5.13) cannot be guaranteed. Instead, the concept of the viscosity solutions can be used here to investigate the existence of the weak solution to the HJB equation (5.13). However, the concept of the viscosity solutions does not provide an explicit form of the optimal strategy and should therefore not be pursued further here. In addition, the feasible control set is not convex and the optimal strategy $u^\star$ is taking only discrete values in a finite set. Thus, the state $X^{u^\star}$ associated with optimal control $u^\star$ may be continuous but not differentiable as the drift coefficient μ in the SDE (5.5) is discontinuous. Furthermore, the diffusion coefficient σ is degenerate. This leads to a delicate mathematical problem which consists in checking the admissibility of the optimal control process $u^\star$ and its solution.

We have already mentioned above that the differential operator of the HJB equation is not elliptic and that the existence of the classical solution cannot be guaranteed. In addition, we have also mentioned that one way to overcome this problem is to use the viscosity solution concept to ensure the existence of the weak solution to the HJB equation (5.13). Another way to overcome this problem is use the regularization techniques. Several authors have resorted to this technique to guarantee the solution of the HJB equation (see, i.e., Frey et al. [47], Krylov [66, Chapter 4, Sec. 6] and Fleming and Soner [45, Sec. IV.6], Fleming and Rishel [44, Chapter 6, Sec. 8]). The main idea of this technique is to modify the diffusion coefficient by adding some small $\varepsilon > 0$ so that the resulting SDE is non-singular. For example, with small perturbation ε , the SDE (5.5) can be written as

$$\begin{aligned} &dX^\varepsilon(t) = \mu(t,X(t),u(t))dt + \sigma^\varepsilon(t,X(t))dW^\varepsilon(t),\ \ X^\varepsilon(0) = x_0^\varepsilon \in \mathcal{X}^\varepsilon \\ &\mathcal{X}^\varepsilon = \mathcal{R}\times\mathcal{F}\times\mathcal{P}^\varepsilon\times\mathcal{Y}^\varepsilon = \{(r,f,p^\varepsilon,y^\varepsilon) \mid r\in\mathcal{R},\ f\in\mathcal{F},\ p^\varepsilon\in\mathcal{P}^\varepsilon,\ y^\varepsilon\in\mathcal{Y}^\varepsilon\}, \quad (5.14) \end{aligned}$$

where $W^\varepsilon = (\widehat{W},\widetilde{W}) \in \mathbb{R}^{\ell+3}$, with $\widetilde{W} = (W_P, W_{\widetilde{Y}_1},\ldots,W_{\widetilde{Y}_\ell})$ an $(\ell+1)$-dimensional wiener process independent of $\widehat{W} = (W_R, W_F)$ and

$$\sigma^\varepsilon(t,x^\varepsilon) = \begin{pmatrix} \widehat{\sigma}(t,\widehat{x}(t)) & \mathbf{0}_{2\times(l+1)} \\ \mathbf{0}_{(l+1)\times 2} & \sqrt{2\varepsilon}\mathbb{I}_{(\ell+1)} \end{pmatrix} \in \mathbb{R}^{(l+3)\times(l+3)}$$

with $\widehat{\sigma}\in\mathbb{R}^{2\times 2}$ given in (5.6) and $\mathbb{I}_{(\ell+1)}$ is and $(\ell+1)\times(\ell+1)-$ identity matrix. The generator associated with these perturbed dynamics has an additional term

$$\varepsilon\left(\frac{\partial^2 V}{\partial p^2} + \frac{\partial^2 V}{\partial y_1^2} + \ldots + \frac{\partial^2 V}{\partial y_\ell^2}\right).$$

Therefore, it is uniformly elliptic. Hence, the existence of a classical solution V^ε to the perturbed HJB equation can be obtained by applying the above mentioned results to the perturbed SDE ((5.14)). Moreover, it can be shown (see, [47], [99]) that for sufficiently small ε the value function V^ε of the perturbed problem converges to the value function V of the original (unperturbed) problem. In addition, for ε sufficiently small, the optimal strategy for the perturbed problem is approximately the optimal strategy of the original problem.

Remark 5.3.4 Note that in many cases the HJB equation (5.13) has no closed-form solution

due to its non-linear structure. Therefore, the HJB equation has to be solved numerically but when the dimension of the state process is high many numerical methods based on finite difference schemes such as Semi-Lagrangian techniques and splitting method become intractable due to the curse of dimensionality. However, an alternative approach is to approximate the above continuous-time optimal control problem to a discrete-time problem based on Markov decision processes (MDP) and solve by using the backward recursion method or the approximate dynamic programming (ADP) which is considered below.

CHAPTER 6

Discrete-Time Stochastic Optimal Control Problem

The aim of this chapter is to find an alternative method to overcome the curse of dimensionality. We recall that the dimension of the state space $d = \ell + 3$, where ℓ is the dimension of the reduced-order states which can be high. Due to this, the HJB equation (5.13) associated to the continuous-time optimal control problem cannot be handled using well known numerical techniques. Instead, we are going to consider the time discretization of the control problem which leads to a Markov decision process with a finite time horizon and action spaces. Further, we derive the associated dynamic programming equation or Bellman equation and solve it using the backward recursion method. A reader who is interested in the theory of Markov Decision Process can consult the book by Bäuerle and Rieder [11] and some references given in the introduction. This chapter is organized as follows. In Sec. 6.1 we present the discrete-time MDP resulting from the time discretization of the continuous-time model and study the transition kernel. The novelty here is that, we manage to solve the SDEs and system of ODEs in closed form in Subsec. 6.1.1 in order to obtain a discrete-time model with no discretization error. Contrary to the continuous-time model, where the control can be changed an anytime, in discrete-time the action can only be changed at the discrete time points. This motivates us to reformulate the state-dependent control constraints for the discrete-time model in Subsec. 6.1.3. In Sec. 6.2 we apply state-discretization to transform the MDP into a finite actions and finite states Markov chain and construct associated the transition probabilities. Finally, we discuss the numerical solution of the discrete-time optimal control problem in Sec. 6.3 and carry out intensive numerical experiments to determine the behaviour of the value function and the optimal strategy. In addition, we investigate the sensitivity analysis of the behaviour of the value function and the optimal control with respect some selected parameters.

6.1 Discrete-time Markov Decision Processes

We recall that the dimension of the state space $d = \ell + 3$, where ℓ is the dimension of the reduced-order states. Let $t_n = n\Delta_N,\ n = 0, 1, \ldots, N$, be the discrete time points with N the number of time steps and $\Delta_N = T/N = t_{n+1} - t_n$ the step size. We sample the state process $X(t), t \in [0, T]$ at discrete time points $t_0, \ldots, t_N$. Let $X_n = X(t_n)$ be the short-hand notation of the

sampling of the state process at time n, $u_n = u(t_n)$ the sampling of the decision rule at time t_n, where we assume $u(t) = \sum_{n=1}^{N} u_n \mathbb{1}_{[t_n,t_{n+1})}(t)$.

Control. The control process or strategy is defined by

$$\mathbf{u} = (u_0, u_1, u_2, \ldots, u_{N-1}),$$

where for $n = 0, 1, \ldots, N-1, u_n = \widetilde{u}(n, X_n^{\mathbf{u}})$, with the mapping

$$\widetilde{u} : \{0, 1, \ldots, N-1\} \times \mathcal{X} \to \overline{\mathcal{U}}$$

is the decision rule at time n and for all $x \in \mathcal{X}$ and $n \in \{0, 1, \ldots, N-1\}$,

$$u_n(x) \in \overline{\mathcal{U}} = \{u^O, u^C, u^W, u^D, u^F\}$$

is the action taken in state x at time n.
In the following, we are going to first write down closed-form solutions to the state equations (SDEs and ODEs) on the time interval $[t_n, t_{n+1})$. Second, based on these expressions we show that the conditional distribution of $X(t_{n+1})$ given $X(t_n) = x$ is a (degenerated) multivariate Gaussian. Finally, we derive the recursion defining the transition operator

$$X_{n+1} = \mathcal{T}_n(X_n, u_n, \mathcal{E}_{n+1}), \quad X_0 = X(0) = x_0 \;\; n = 0, 1, \ldots N-1,$$

where $(\mathcal{E}_1, \ldots, \mathcal{E}_N)$ is a sequence of standard normally distributed random variables that we will specify later.

6.1.1 Time-Discretization of the State Variables

This subsection is devoted to the time-discretization of the dynamics of the state process. The state process $X = (X_n)_{n=0,1,\ldots,N}$ is taking values in $\mathcal{X} \subset \mathbb{R}^d$, i.e, $X_n = (R_n, F_n, P_n, \widetilde{Y}_n) = (R_n, F_n, P_n, \widetilde{Y}_n^1, \widetilde{Y}_n^2, \ldots, \widetilde{Y}_n^\ell) \in \mathcal{X}$. We begin with the following assumption on the parameters which it crucial for this chapter.

Assumption 6.1.1 We assume that $\mu_R(t)$, $\mu_F(t)$, $\sigma_R(t)$ and $\sigma_F(t)$ are piece-wise constant, i.e, for $\dagger = R, F$,

$$\mu_\dagger(t) = \sum_{n=0}^{N-1} \mu_{\dagger,n} \mathbb{1}_{[t_n,t_{n+1})}(t), \;\; \mu_\dagger(T) = \mu_{\dagger,N}, \text{ and } \sigma_\dagger(t) \;\; = \sum_{n=0}^{N-1} \sigma_{\dagger,n} \mathbb{1}_{[t_n,t_{n+1})}(t), \;\; \sigma_\dagger(T) = \sigma_{\dagger,N}.$$

with some constants $\mu_{\dagger,n}, \sigma_{\dagger,n}, \quad n = 0, \ldots, N$. An example can be $\mu_{\dagger,n} = \mu_\dagger(t_n)$ and $\sigma_{\dagger,n} = \sigma_\dagger(t_n)$.

To avoid time-discretization error, instead of apply Euler scheme, we rather solve the state equations in closed-form on the time interval $[t_n, t_{n+1})$, $n = 0, 1, \ldots, N-1$.

Discrete-time dynamics of the geothermal storage. We recall that the continuous-time dy-

namics of the GS given in equation (5.2) can be written as

$$d\widetilde{Y}(t) = (\widetilde{A}\,\widetilde{Y}(t) + \widetilde{B}g(t,u(t))dt, \quad \widetilde{Y}(0) = y_0 \in \mathcal{Y} \subset \mathbb{R}^\ell, \tag{6.1}$$

where $\widetilde{A}$ and $\widetilde{B}$ are constant system and input matrices given in (5.3) and the input function $g(t,\nu)$ is given by relation (5.4). We make the following assumption on the input function g.

Assumption 6.1.2 We assume that for a fixed control $u_n = \nu$ the input function $g(t,\nu) = g_n^\nu$ is constant on the time interval $[t_n, t_{n+1})$.

The following lemma gives the the closed-form solution of the ODE (6.1).

Lemma 6.1.3 Let $t \in [t_n, t_{n+1})$. Under Assumption 6.1.2, the closed-form solution of the ODE (6.1) on the time interval $[t_n, t_{n+1})$, with initial value $\widetilde{Y}(t_n) = \widetilde{Y}_n$, is given as

$$\widetilde{Y}(t) = e^{\widetilde{A}(t-t_n)}\widetilde{Y}_n + (e^{\widetilde{A}(t-t_n)} - \mathbb{I}_\ell)\widetilde{A}^{-1}\widetilde{B}g_n^\nu, \tag{6.2}$$

In particular, when $t \to t_{n+1}$, the a above closed-form solution yields the recursion

$$\widetilde{Y}_{n+1} = e^{\widetilde{A}\Delta_N}\widetilde{Y}_n + (e^{\widetilde{A}\Delta_N} - \mathbb{I}_\ell)\widetilde{A}^{-1}\widetilde{B}g_n^\nu. \tag{6.3}$$

where $\mathbb{I}_\ell$ is an $\ell \times \ell$ identity matrix.

The proof of this lemma can be found in Appendix C.2.1.

For sufficiently small Δ_N, the first Taylor expansion of the solution $\widetilde{Y}_{n+1}$ given in (6.3), is given by

$$\begin{aligned}\widetilde{Y}_{n+1} = \widetilde{Y}(t_{n+1}) &= (\mathbb{I}_\ell + \widetilde{A}\Delta_N)\widetilde{Y}_n + \widetilde{A}\widetilde{A}^{-1}\widetilde{B}g_n^\nu\Delta_N + \mathcal{O}(\Delta_N) \\ &= (\mathbb{I}_\ell + \widetilde{A}\Delta_N)\widetilde{Y}_n + \widetilde{B}g_n^\nu\Delta_N + \mathcal{O}(\Delta_N).\end{aligned}$$

This correspond to the Euler discretization of the ODE (6.1) which is a good approximation only for sufficiently small Δ_N.

Discrete-time dynamics of the residual demand. Recall that the continuous-time residual demand is defined by

$$dR(t) = \beta_R(\mu_R(t) - R(t))dt + \sigma_R(t)dW_R(t). \tag{6.4}$$

Let $R_n = R(t_n)$ be the sampling of $R(t)$ a time t_n. The closed-form solution of the SDE (6.4) is given my the following:

Lemma 6.1.4 Let $t \in [t_n, t_{n+1})$. Then, under Assumption 6.1.1 the closed-form solution of the SDE (2.1) on the time interval $[t_n, t_{n+1})$, with initial value $R(t_n) = R_n$ is given by

$$R(t) = R_n e^{-\beta_R(t-t_n)} + \mu_{R,n}(1 - e^{-\beta_R(t-t_n)}) + \sigma_{R,n}\int_{t_n}^{t} e^{-\beta_R(t-s)}dW_R(s). \tag{6.5}$$

The proof of this lemma can be found in Appendix C.2.2.

Discrete-time dynamics of the fuel price. We assume that the fuel price is stochastic and follows the OU-process with piece-wise $\mu_F(t)$ and piece-wise $\sigma_F(t)$, and constant mean reversion speed β_F. Using de same procedure presented above for the case of residual demand we obtain the closed-form solution of the SDE (2.4) given in the following lemma:

Lemma 6.1.5 Under Assumption 6.1.1, the closed-form solution of the SDE (2.4) on the time interval $[t_n, t_{n+1})$, with initial value $F(t_n) = F_n$ is given by

$$F(t) = F_n e^{-\beta_F(t-t_n)} + \mu_{F,n}(1 - e^{-\beta_F(t-t_n)}) + \sigma_{F,n} \int_{t_n}^{t} e^{-\beta_F(t-s)} dW_F(s). \tag{6.6}$$

The proof of this lemma is similar to the proof of Lemma 6.1.4.

Discrete-time dynamics of the internal storage . We recall that the continuous-time dynamics of the temperature in the IS given by equation (2.17) can be written for $u(t) \in \overline{\mathcal{U}}$ as

$$dP(t) = (\psi_p(R(t), \widetilde{Y}(t), u(t))) - \gamma(P(t) - P_{amb})dt, \quad P(0) = p_0, \tag{6.7}$$

where $\gamma = \frac{\kappa_h A_h}{m^P c_P^F}$ is a constant, $R(t)$ is the residual demand given by equation (6.5), and $\widetilde{Y}(t)$ is the reduced order state of the GS given by (6.2). The function ψ_p is given by (2.18). The following lemma provides the closed-form solution to equation (2.17).

Lemma 6.1.6 Assume that at time t_n the average temperature in the IS $P(t_n) = P_n$. Under Assumptions 6.1.1 and 6.1.2, for $n \in \{0, 1, \ldots, N-1\}$, the closed-form solution to equation (6.7) on the time interval $[t_n, t_{n+1})$ is given by

1) For $u_n = u^O$

$$P(t) = e^{-\gamma(t-t_n)} P_n + P_{amb}(1 - e^{-\gamma(t-t_n)}). \tag{6.8}$$

2) For $u_n \neq u^O$

$$P(t) = e^{-\gamma(t-t_n)} P_n + \Upsilon_n(u_n, \widetilde{Y}_n) - k_P \sigma_{R,n} \int_{t_n}^{t} e^{-\gamma(t-s)} \left(\int_{t_n}^{s} e^{-\beta_R(s-u)} dW_R(u) \right) ds, \tag{6.9}$$

where the function Υ_n is given by

$$\Upsilon_n(\nu, y) = \eta_n + \begin{cases} \left(P_{amb} + \frac{\kappa^F}{\gamma}\right)(1 - e^{-\gamma(t-t_n)}), & \nu = u^F \\ \left(P_{amb} + \frac{\kappa_C(p_{in} - P_{out})}{\gamma}\right)(1 - e^{-\gamma(t-t_n)}), & \nu = u^D \\ P_{amb}(1 - e^{-\gamma(t-t_n)}), & \nu = u^W \\ \left(P_{amb} - \frac{\kappa_D Q_C^I}{\gamma}\right)(1 - e^{-\gamma(t-t_n)}) + e^{-\gamma(t-t_n)} \psi_n(y), & \nu = u^D \end{cases} \tag{6.10}$$

with ψ_n given for $g_n^D = g(t_n, u^D)$ by

$$\psi_n(y) = \kappa_D C^O \Big\{ (e^{(\gamma \mathbb{I}_\ell + \widetilde{A})(t-t_n)} - \mathbb{I}_\ell)(\gamma \mathbb{I}_\ell + \widetilde{A})^{-1} y +$$

$$\left[(e^{(\gamma\mathbb{I}_\ell+\widetilde{A})(t-t_n)}-\mathbb{I}_\ell)(\gamma\mathbb{I}_\ell+\widetilde{A})^{-1}-\frac{1}{\gamma}(e^{\gamma(t-t_n)}-1)\mathbb{I}_\ell\right]\widetilde{A}^{-1}\widetilde{B}g_n^D\Bigg\},$$

$\mathbb{I}_\ell$ is an $\ell\times\ell$ identity matrix, and

$$\eta_n=\frac{k_P}{\beta_R-\gamma}(\mu_{R,n}-R_n)\left(e^{-\gamma(t-t_n)}-e^{-\beta_R(t-t_n)}\right)-\frac{k_P\mu_{R,n}}{\gamma}(1-e^{-\gamma(t-t_n)}).$$

The proof of this lemma is given in Appendix C.2.3.

Now, let $R_{n+1}=R(t_{n+1})$, $F_{n+1}=F(t_{n+1})$, and $P_{n+1}=P(t_{n+1})$ be the sampling of the residual demand, the fuel or electricity price, and the average temperature in the IS at time t_{n+1}, respectively. Next, we are going to determine based on the above closed-form solutions, the (marginal and) joint conditional distribution of $(R_{n+1},F_{n+1},P_{n+1})$ given $(R_n,F_n,P_n)=(r,f,p)$ and the action $u_n=\nu$.

6.1.2 Marginal and Joint Conditional Distributions of the State Variables

We first note that, the last term in equations (6.5) and (6.6) are integral functionals of Wiener processes with deterministic integrand. Further, the last term of equations (6.9) is an integrated process which is also integral functional of the Wiener process with deterministic integrand. Therefore, for $t\in[t_n,t_{n+1})$, $R(t)$, $F(t)$, and $P(t)$ given by (6.5), (6.6), and (6.9), respectively, are normally distributed. In particular, when $t\to t_{n+1}$, the processes R_{n+1}, F_{n+1}, and P_{n+1} are Gaussian random variables. In addition, the last term in equations (6.5) and (6.9) are integral functional of the same Wiener process. Therefore, when $t\to t_{n+1}$, the Gaussian processes R_{n+1} and P_{n+1} are dependent random variables. We denote by $m_{R,n}=m_{R,n}(r)=\mathbb{E}[R_{n+1}\mid R_n=r]$, $m_{F,n}=m_{F,n}(f)=\mathbb{E}[F_{n+1}\mid F_n=f]$, $m^\nu_{P,n}=m^\nu_{P,n}(p)=\mathbb{E}[P_{n+1}\mid(R_n,F_n,P_n,\widetilde{Y}_n)=(r,f,p,y),u_n=\nu]$, and $m^\nu_{Y,n}=\mathbb{E}[\widetilde{Y}_{n+1}\mid\widetilde{Y}_n=y,u_n=\nu]$ the conditional means of R_{n+1},F_{n+1},P_{n+1}, and $\widetilde{Y}_{n+1}$ given $X_n=x=(r,f,p,y)$ and $u_n=\nu$, respectively. We also denote by $\Sigma^2_{R,n}=Var(R_{n+1}\mid R_n=r)$, $\Sigma^2_{F,n}=Var(F_{n+1}\mid F_n=f)$, and $\Sigma^2_{P,n}=Var(P_{n+1}\mid X_n=x,u_n=\nu)$ the conditional variances of R_{n+1},F_{n+1}, and P_{n+1}, respectively.
Since R_{n+1} and P_{n+1} are dependent random variables, we denote by $\Sigma^{2,\nu}_{RP,n}=Cov(R_{n+1},P_{n+n}\mid X_n=x,u_n=\nu)$ their conditional covariance and by $\rho^\nu_{RP,n}$ their correlation coefficient. In the following we will derive explicit formulas for each parameter stated below in equation (6.17). Through out this chapter we consider that Assumption 6.1.1 is fulfilled. We use for simplicity the short-hand notations $m_{P,n}=m^\nu_{P,n}$ and $m_{Y,n}=m^\nu_{Y,n}$ for the conditional means of P_{n+n} and $\widetilde{Y}_{n+1}$, respectively. We use the short-hand notations $\Sigma^2_{RP,n}=\Sigma^{2,\nu}_{RP,n}$ and $\rho_{RP,n}=\rho^\nu_{RP,n}$ for the conditional covariance and the conditional correlation coefficient of (R_{n+1},P_{n+1}). We begin with the marginal distribution of R_{n+1}.

Conditional distribution of $\mathbf{R_{n+1}}$ given $\mathbf{R_n=r}$. We recall that the discrete-time process R_{n+1} is a Gaussian process. Then, the following proposition provides its mean and its variance.

Proposition 6.1.7 (Conditional mean and variance of $\mathbf{R_{n+1}}$) Under Assumption 6.1.1, the residual demand R_{n+1} is a normally distributed random variable with a conditional mean given by

$$m_{R,n}=re^{-\beta_R\Delta_N}+\mu_{R,n}(1-e^{-\beta_R\Delta_N}),$$

and a conditional variance given by

$$\Sigma_{R,n}^2 = \frac{\sigma_{R,n}^2}{2\beta_R}(1 - e^{-2\beta_R \Delta_N}). \tag{6.11}$$

The proof is given in Appendix C.2.4.

Conditional distribution of $\mathbf{F_{n+1}}$ given $\mathbf{F_n} = \mathbf{f}$. Similarly, the discrete-time process F_{n+1} given by the recursion (6.6) is a Gaussian process and the following proposition provides its mean and its variance.

Proposition 6.1.8 (Conditional mean and variance of $\mathbf{F_{n+1}}$) Under Assumption 6.1.1, the Fuel price F_{n+1} is a standard normally distributed random variable with a conditional mean given by

$$m_{F,n} = f e^{-\beta_F \Delta_N} + \mu_{F,n}(1 - e^{-\beta_F \Delta_N}),$$

and a conditional variance given by

$$\Sigma_{F,n}^2 = \frac{\sigma_{F,n}^2}{2\beta_F}(1 - e^{-2\beta_F \Delta_N}).$$

The proof this proposition is similar the proof of proposition 6.1.7 given in Appendix C.2.4.

Conditional distribution of $\mathbf{P_{n+1}}$ given $\mathbf{X_n} = \mathbf{x} = (\mathbf{r}, \mathbf{f}, \mathbf{p}, \mathbf{y})$ and $\mathbf{u_n} = \nu$. We recall that for all feasible control $u_n \in \overline{\mathcal{U}} \backslash \{u^O\}$ the discrete-time process P_{n+1} given in (6.9) is a normally distributed random variable. The following theorem provides its conditional mean and variance given $X_n = x$ and $u_n = \nu$.

Theorem 6.1.9 (Condition distribution of $\mathbf{P_{n+1}}$) Let Assumption 6.1.1 be fulfilled and the rate of heat loss to the environment $\gamma \neq 0$. Then, for all feasible control $\nu \in \overline{\mathcal{U}} \backslash \{u^O\}$ the conditional distribution of the process P_{n+1} given $X_n = x = (r, f, p, y)$ and $u_n = \nu$ is Gaussian with the following parameters:

1. The conditional mean is given by

$$m_{P,n} = e^{-\gamma \Delta_N} p + \Upsilon_n(\nu, y),$$

   where Υ_n is given by (6.10).

2. The conditional variance is given by

$$\Sigma_{P,n}^2 = \frac{k_P^2 \sigma_{R,n}^2}{2\beta_R(\beta_R - \gamma)^2(\beta_R + \gamma)} \Big\{ \gamma + 4\beta_R e^{-(\beta_R + \gamma)\Delta_N} - (\beta_R + \gamma) e^{-2\beta_R \Delta_N} - \beta_R \left(2 + e^{-2\gamma \Delta_N}\right) + \frac{\beta_R^2}{\gamma} \left(1 - e^{-2\gamma \Delta_N}\right) \Big\}. \tag{6.12}$$

The proof of this theorem is given in Appendix C.2.5.

Next, we want to derive $\rho_{RP,n}$ based on the joint conditional distribution of R_{n+1} and P_{n+1} given $X_n = x$ and $u_n = \nu$.

Joint conditional distribution of $\mathbf{R_{n+1}}$ and $\mathbf{P_{n+1}}$. Recall that when $t \to t_{n+1}$ the residual demand R_{n+1} and the temperature in the IS P_{n+1} given by the recursions (6.1.4) and (6.9), respectively, are correlated normally distributed random variables. Therefore, the conditional distribution of the pair (R_{n+1}, P_{n+1}) is a bivariate normal distribution. We denote by $\Sigma^2_{RP,n} = cov(R_{n+1}, P_{n+1} \mid X_n = x, u_n = \nu)$ the conditional covariance of R_{n+1} and P_{n+1} given $X_n = x$ and $u_n = \nu$. In the following theorem we derive their covariance and their correlation coefficient.

Theorem 6.1.10 (Conditional correlation between $\mathbf{R_{n+1}}$ and $\mathbf{P_{n+1}}$) For $t \to t_{n+1}$, R_{n+1} and P_{n+1} given by (6.5) and (6.9), respectively, are negatively correlated with the conditional correlation coefficient $\rho_{RP,n} \in [-1, 0)$ given by

$$\rho_{RP,n} = \frac{\Sigma^2_{RP,n}}{\Sigma_{R,n}\Sigma_{P,n}}. \tag{6.13}$$

Here, $\Sigma_{P,n}$ and $\Sigma_{R,n}$ are given by (6.12) and (6.11), respectively, and $\Sigma^2_{RP,n}$ is the conditional covariance of R_{n+1} and P_{n+1} given by

$$\Sigma^2_{RP,n} = -\frac{k_P \sigma^2_{R,n}}{2\beta_R(\beta_R^2 - \gamma^2)} \left(\beta_R - \gamma - 2\beta_R e^{-(\beta_R+\gamma)\Delta_N} + (\beta_R + \gamma)e^{-2\beta_R \Delta_N}\right) \tag{6.14}$$

For the proof see Appendix C.3.1.

Joint conditional density function of $\mathbf{P_{n+1}}$ and $\mathbf{R_{n+1}}$. We recall that giving the current state $X_n = x$ and the current control $u_n = \nu \in \overline{\mathcal{U}} \backslash \{u^O\}$, the random variables R_{n+1} and P_{n+1} are normally distributed and correlated, i.e., $P_{n+1} \sim \mathcal{N}(m_{P,n}, \Sigma^2_{P,n})$ and $R_{n+1} \sim \mathcal{N}(m_{R,n}, \Sigma^2_{P,n})$. In addition, the pair (R_{n+1}, P_{n+1}) is jointly normal with parameters $m_{P,n}$, $\Sigma^2_{P,n}$, $m_{R,n}$, $\Sigma^2_{R,n}$ and $\rho_{RP,n}$. Now we want to investigate their joint distribution. We first introduce some useful notations for the remaining section. Let $Z_P = \frac{P_{n+1} - m_{P,n}}{\Sigma_{P,n}} \sim \mathcal{N}(0,1)$ and $Z_R = \frac{R_{n+1} - m_{R,n}}{\Sigma_{R,n}} \sim \mathcal{N}(0,1)$ be two standard normal variables.

Proposition 6.1.11 The random variable

$$Z_{RP} = -\frac{\rho_{RP,n}}{\sqrt{1-\rho^2_{RP,n}}} \left(\frac{R_{n+1} - m_{R,n}}{\Sigma_{R,n}}\right) + \frac{1}{\sqrt{1-\rho^2_{RP,n}}} \left(\frac{P_{n+1} - m_{P,n}}{\Sigma_{P,n}}\right) = \frac{Z_P - \rho_{RP,n} Z_R}{\sqrt{1-\rho^2_{RP,n}}}$$

is standard normally distributed, i.e., $Z_{RP} \sim \mathcal{N}(0,1)$.

The proof is given in Appendix C.3.2.

Let $Z \sim \mathcal{N}(\mu_Z, \sigma^2_Z)$. Then the probability density function of a normal random variable Z is given by

$$\varphi_Z(z) = \frac{1}{\sigma_Z \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{z-\mu_Z}{\sigma_Z}\right)^2},$$

and the cumulative distribution of a standard normal random variable Z is define by

$$\Phi_Z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{s^2}{2}} ds.$$

Definition 6.1.12 The joint conditional density function of the random variables R_{n+1} and P_{n+1} given $X_n = x$ and $u_n = \nu \in \overline{\mathcal{U}} \backslash \{u^O\}$ is defined by

$$\varphi_{RP}(r,p) = \frac{1}{2\pi\Sigma_{R,n}\Sigma_{P,n}\sqrt{1-\rho_{RP,n}^2}} \times \exp\left\{ -\frac{1}{2(1-\rho_{RP,n}^2)} \left[\left(\frac{p-m_{P,n}}{\Sigma_{P,n}} \right)^2 - 2\rho_{RP,n} \frac{(r-m_{R,n})(p-m_{P,n})}{\Sigma_{R,n}\Sigma_{P,n}} + \left(\frac{r-m_{R,n}}{\Sigma_{R,n}} \right)^2 \right] \right\}.$$

Now, let the function ζ_z be defined by $\zeta_z(R_{n+1}, P_{n+1}) = Z_{RP}$ with

$$\zeta_z(r,p) = -\frac{\rho_{RP,n}}{\sqrt{1-\rho_{RP,n}^2}} z_r + \frac{1}{\sqrt{1-\rho_{RP,n}^2}} z_p, \tag{6.15}$$

where $z_r = \frac{r-m_{R,n}}{\Sigma_{R,n}}$ and $z_p = \frac{p-m_{P,n}}{\Sigma_{P,n}}$.
In the following proposition we express the joint conditional density function ϕ_{RP} in terms of the function ζ_z.

Proposition 6.1.13 The joint conditional density function of the random variables R_{n+1} and P_{n+1} given $X_n = x$ and $u_n = \nu \in \overline{\mathcal{U}} \backslash \{u^O\}$ can be written as

$$\varphi_{RP}(r,p) = \varphi_R(r) \frac{1}{\Sigma_{P,n}\sqrt{2\pi}\sqrt{1-\rho_{RP,n}^2}} e^{-\frac{1}{2}\zeta_z^2(r,p)},$$

where $\varphi_R(r)$ is the probability density function of a normal random variable R_{n+1} and $\zeta_z(r,p)$ is a function given by (6.15).

The proof can be found in Appendix C.3.3.

Now, we are in the position to define the joint cumulative distribution function of R_{n+1} and P_{n+1} in terms of φ_R and $\Phi(\zeta_z)$.
Let us first recall the definition of the joint cumulative function of two random variables. Let X and Y be two random variables. Then, the joint cumulative distribution function of X and Y is defined as

$$F_{XY}(x_1, y_1) = \mathbb{P}(X \leq x_1,\ Y \leq y_1) = \int_{-\infty}^{x_1} \int_{-\infty}^{y_1} f_{XY}(x,y) dx dy,$$

where $f_{XY}(x,y)$ is a joint density function of X and Y.

Proposition 6.1.14 Let $p_2, r_2 \in \mathbb{R}$. The joint cumulative distribution function of the random variables R_{n+1} and P_{n+1} can be expressed in terms of φ_R and $\Phi(\zeta_z)$ as follows:

$$F_{RP}(r_2, p_2) = \mathbb{P}(P_{n+1} \leq p_2,\ R_{n+1} \leq r_2) = \int_{-\infty}^{r_2} \varphi_R(r) \Phi(\zeta_z(r, p_2)) dr.$$

For proof see Appendix C.3.4.

This transformation helps to reduce the computational time for the computation of the transition probabilities for the Markov chain that we will study later in this chapter. The following corollary gives some properties of the joint cumulative distribution function defined above.

Corollary 6.1.15 Let $r_1, r_2, p_1, p_2 \in \mathbb{R}$. Then, Proposition 6.1.14 can be extended as follows:

1. $\mathbb{P}(P_{n+1} \leq p_2,\ r_1 \leq R_{n+1} \leq r_2) = \int_{r_1}^{r_2} \varphi_R(r)\Phi(\zeta_z(r,p_2))dr.$

2. $\mathbb{P}(P_{n+1} \geq p_2,\ r_1 \leq R_{n+1} \leq r_2) = \int_{r_1}^{r_2} \varphi_R(r)(1-\Phi(\zeta_z(r,p_2)dr.$

3. $\mathbb{P}(p_1 \leq P_{n+1} \leq p_2,\ r_1 \leq R_{n+1} \leq\ r_2) = \int_{r_1}^{r_2} \varphi_R(r)(\Phi(\zeta_z(r,p_2)) - \Phi(\zeta_z(r,p_1))dr.$

Statements 1,2, and 3 of Corollary 6.1.15 also hold true for $r_1 \to -\infty$ and $r_2 \to +\infty$. The proof of this corollary is analogous to the proof of Proposition 6.1.14.

Transition operator. A time-discretization leads to a Markov Decision Process (MDP) with a finite time horizon N, with state space $\mathcal{X} \subset \mathbb{R}^{\ell+3}$ and finite action space $\overline{\mathcal{U}}$. Based on the above closed-form expressions and the (marginal and) joint distribution of state variables we can derive the recursions defining the linear transition operator associated to the MDP, $\mathcal{T}_n$: $\mathcal{X} \times \overline{\mathcal{U}} \times \mathcal{Z} \to \mathcal{X}$ given by

$$X_{n+1} = \mathcal{T}_n(X_n, u_n, \mathcal{E}_{n+1}),\ \ n = 0,1,\ldots N-1, \tag{6.16}$$

where $X_{n+1} = (R_{n+1}, F_{n+1}, P_{n+1}, \widetilde{Y}_{n+1})$ with individual states given for $X_n = x$, by

$$\begin{aligned}
R_{n+1} &= m_{R,n} + \Sigma_{R,n}\mathcal{E}^R_{n+1},\\
F_{n+1} &= m_{F,n} + \Sigma_{F,n}\mathcal{E}^F_{n+1},\\
P_{n+1} &= m_{P,n} + \Sigma_{P,n}\left(\sqrt{1-\rho^2_{RP,n}}\mathcal{E}^P_{n+1} + \rho_{RP,n}\mathcal{E}^R_{n+1}\right),\\
\widetilde{Y}_{n+1} &= m_{Y,n},
\end{aligned}$$

where $m_{\dagger,n}$ and $\Sigma^2_{\dagger,n}$, $\dagger = R,F,P$ are the mean and the variance of the processes R_{n+1}, F_{n+1}, and P_{n+1}, respectively and $m_{Y,n}$ is the mean of the process $\widetilde{Y}_{n+1}$ given by

$$m_{Y,n} = ye^{\widetilde{A}\Delta_N} \qquad + (e^{\widetilde{A}\Delta_N} - \mathbb{I}_\ell)\widetilde{A}^{-1}\widetilde{B}g^\nu_n.$$

In the above transition operator $(\mathcal{E}_1, \ldots, \mathcal{E}_N)$ with $\mathcal{E}_n = (\mathcal{E}^R_n, \mathcal{E}^F_n, \mathcal{E}^P_n)^\top \in \mathcal{Z} \subset \mathbb{R}^3$, for $n = 1,2,\ldots N$, is a sequence of multivariate standard normally distributed random variables.
Next, we will investigate the conditional distribution of $X(t_{n+1})$ given $X(t_n) = x$ and $u_n = \nu$ based on the marginal distribution of the independents variables and the joint distribution of the dependent variables.

Transition Kernel. In this paragraph we use the results of the marginal and joint distributions of the discrete-time state variables R_{n+1}, F_{n+1}, P_{n+1}, and $\widetilde{Y}_{n+1}$ to derive the conditional distribution of $X(t_{n+1})$ given $X(t_n) = x$ and $u_n = \nu$. We recall that under Assumption 6.1.1 there is no discretization error in the discrete-time approximation of the continuous-time process. Therefore, the law of the discrete-time process $(X_0, X_1, \ldots, X_N)$ and the law of the continuous-time process $(X(t_0), X(t_1), \ldots, X(t_N))$ sampled at $t_0, t_1, \ldots, t_N$ coincide. We have shown based on the closed-form expressions (6.5), (6.6), (6.9), and (6.3) that, for $u_n \in \overline{\mathcal{U}} \backslash \{u^O\}$, the discrete-time state variables R_{n+1}, F_{n+1} and P_{n+1} are normally distributed random variables and the discrete-time state variable $\widetilde{Y}_{n+1}$ is degenerated (Dirac). However, for $u_n = u^O$ the closed-form expression (6.8) shows that P_{n+1} is degenerated (Dirac). Then, given the state $X_n = x$ and action $u_n = \nu \in \overline{\mathcal{U}}$ at time t_n the conditional distribution of the process $X_{n+1} = X(t_{n+1})$ is a (degenerate) multivariate Gaussian with the mean $m_{X,n}$ and the covariance matrix $\Sigma_{X,n}$ represented as

$$m_{X,n} = \begin{pmatrix} m_{R,n} \\ m_{F,n} \\ m_{P,n} \\ m_{Y,n} \end{pmatrix}, \quad \Sigma_{X,n} = \begin{pmatrix} \Sigma^2_{R,n} & 0 & \Sigma^2_{RP,n} & 0 \\ 0 & \Sigma^2_{F,n} & 0 & 0 \\ \Sigma^2_{RP,n} & 0 & \Sigma^2_{P,n} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Sigma^2_{RP,n} = \rho_{RP,n}\Sigma_{R,n}\Sigma_{P,n}. \tag{6.17}$$

Uncertainty. We model the uncertainty by a sequence $\mathcal{E} = (\mathcal{E}_n)_{n=1,\ldots,N}$ of independent identically distributed random variables with values in $\mathcal{Z}$, i.e, for $n = 1, \ldots, N$, $\mathcal{E}_n = (\mathcal{E}^R_n, \mathcal{E}^F_n, \mathcal{E}^P_n)^\top \in \mathcal{Z}$, where $\mathcal{E}^\dagger_n \sim \mathcal{N}(0,1)$, $\dagger = R, F, P$.

Filtration. We consider the filtration $\mathbb{F} = (\mathcal{F}_n)_{n=0,1,\ldots,N}$ with $\mathcal{F}_n = \sigma(\{\mathcal{E}_1, \ldots, \mathcal{E}_n\})$ is the sigma-algebra generated by the first n independent random variables $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ and $\mathcal{F}_0 = \{\emptyset, \Omega\}$ is the trivial sigma-algebra.

6.1.3 State-Dependent Control Constraints for MDP

For Markov decision processes in each time step $n \in \{0, 1, \ldots, N-1\}$ a set of feasible actions depending on the current state is required. The latter should be defined such that within the next period the internal and GS are not full or empty. We denote by

$$\mathcal{U}(n,x) = \{\nu \in \overline{\mathcal{U}} \mid X^\nu_{n+1} \text{ satisfies the state constraint, given } X_n = x\},$$

where $\overline{\mathcal{U}}$ is the set of feasible actions. We allow that $\mathcal{U} = \mathcal{U}(n,x)$ depends on the current time n and the current state.

The state-dependent control constraints result from the constraints to the state which are box constraints to the average temperature in the IS, $P_n \in [\underline{p}, \overline{p}]$ and to the average temperature in the medium of GS, $\overline{Q}^M_n = C^M \widetilde{Y}_n \in [\underline{q}, \overline{q}]$, for all $n = 0, 1, \ldots, N$. We note that the Gaussian nature of the process P_{n+1} does not allow it to satisfy the box constraint to P with certainty. Instead we have to allow over- and undershooting, i.e. $P_{n+1} > \overline{p}$ and $P_n < \overline{p}$, but constrain the probabilities for the events by very small tolerance value $\varepsilon \ll 1$.
In the view of the state discretization such an approach appears to be acceptable since the grid points on the boundaries of the truncated state space, in particular the points on $P_{n+1} = \underline{p}$ and $P_{n+1} = \overline{p}$ will represent all points in the state space with $P_{n+1} \leq \underline{p}$ and $P_{n+1} \geq \overline{p}$, respectively. Since we have to satisfy two state constraints by an appropriate set of feasible actions, we first

define the sets

$$\begin{aligned}\mathcal{U}_P(n,x) = \big\{ \nu \in \overline{\mathcal{U}} \mid\ & P^{\nu}_{n+1} \text{ satisfies the state constraint to } P,\\ & i.e.\ \mathbb{P}\big(P^{\nu}_{n+1} \in [\underline{p},\overline{p}]\big) \geq 1-2\varepsilon,\ \text{given } X_n = x\big\}\\ \mathcal{U}_Y(n,x) = \big\{ \nu \in \overline{\mathcal{U}} \mid\ & \widetilde{Y}^{\nu}_{n+1} \text{ satisfies the state constraint to } \widetilde{Y},\\ & i.e.\ C^M \widetilde{Y}^{\nu}_{n+1} \in [\underline{q},\overline{q}],\ \text{given } X_n = x\big\}. \end{aligned} \tag{6.18}$$

Then the desired set of feasible actions is given by

$$\mathcal{U}(n,x) = \mathcal{U}_P(n,x) \cap \mathcal{U}_Y(n,x).$$

Next, we describe the $\mathcal{U}_P(n,x)$ and $\mathcal{U}_P(n,x)$ separately.

Set of feasible actions $\mathcal{U}_Y(n,x)$ related to the state constraint to $\widetilde{Y}$. In our continuous-time control problem we impose constraints to the state $\widetilde{Y}$ of the form

$$\overline{Q}^M(t) = C^M \widetilde{Y}(t) \in [\underline{q},\overline{q}],\ t \in [0,T],$$

i.e. box constraints to the average temperature in the storage medium which is a linear combination of the entries of $\widetilde{Y}$. This constraint implies state-dependent control constraints saying that an empty GS ($\overline{Q}^M(t) = \underline{q}$) can no longer be discharged ($u(t) \neq u^C$) and a full GS ($\overline{Q}^M(t) = \overline{q}$) can no longer be charged ($u(t) \neq u^D$). However, in the discrete-time setting the control $u(t)$ can no longer be changed at any time t but only at the discrete-time points $t_0,\ t_1, \ldots, t_{N-1}$. Hence, charging or discharging the GS must be stopped already for average temperatures $\overline{Q}^M(t_n)$ slightly below $\overline{q}$ or slightly above $\underline{q}$. This leads to a change of the feasible set of control depending on the current state if one changes from continuous to discrete-time setting.
The set $\mathcal{U}_Y(n,x)$ can be subdivided into 3 subsets, see Fig 6.1, corresponding to $\mathcal{Y}_C, \mathcal{Y}_D, \mathcal{Y}_W$, given by

$$\begin{aligned}\mathcal{Y}_C &= \{y \in \mathcal{Y} :\ C^M \mathcal{T}^y_n(y,u^D) > \overline{q}\},\\ \mathcal{Y}_D &= \{y \in \mathcal{Y} :\ C^M \mathcal{T}^y_n(y,u^C) < \underline{q}\},\\ \mathcal{Y}_W &= \mathcal{Y} \backslash (\mathcal{Y}_C \cup \mathcal{Y}_D),\end{aligned}$$

where for all $n = 0,1,\ldots,N-1$, $\mathcal{T}^y_n$ is the transition operator for $(\widetilde{Y}_n)_n$ given by

$$\mathcal{T}^y_n(y,\nu) = y e^{\widetilde{A}\Delta_N} \quad + (e^{\widetilde{A}\Delta_N} - \mathbb{I}_\ell)\widetilde{A}^{-1}\widetilde{B} g^{\nu}_n.$$

This partitioning of the set $\mathcal{U}_Y(n,x)$ can also be seen in first row of Fig 6.4, where we compute all elements of the intersection $\mathcal{U}_P(n,x) \cap \mathcal{U}_Y(n,x)$.

Note that for extreme charging and discharging rates, there might be an empty $\mathcal{Y}_W$ region and an overlap of the two other regions $\mathcal{Y}_C$ and $\mathcal{Y}_D$.
The most critical situation occurs when there is a possibility that the average temperature in the GS exceeds the maximum at the next time step, $\overline{Q}^M_{n+1} \geq \overline{q}$, given the current average temperature $\overline{Q}^M_n$. In this case, charging the GS is no longer possible and the set of feasible actions becomes $\mathcal{U}_Y(n,x) = \overline{\mathcal{U}} \backslash u^D$. Moreover, when it is possible that, given the current average temperature, the average temperature in the GS at the next time step is lower than the minimum, $\overline{Q}^M_{n+1} \leq \underline{q}$. Then discharging GS is no longer possible and in this case, the set of feasible actions is the given by

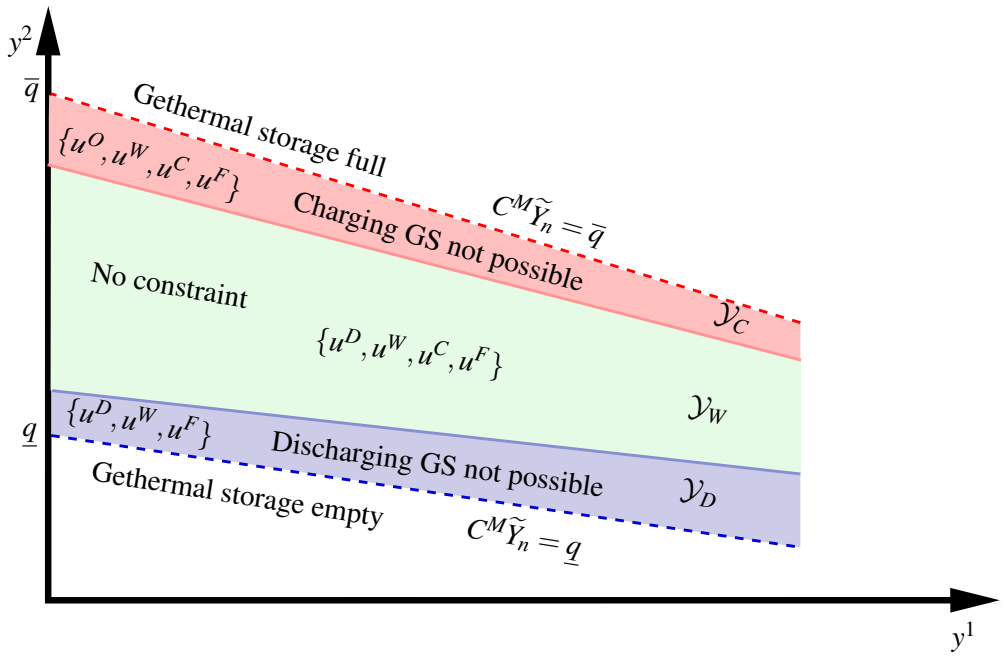


Figure 6.1: Characterization of the set of feasible control $\mathcal{U}_Y(n,x)$ for $\ell = 2$

$\mathcal{U}_Y(n,x) = \{u^D, u^W, u^F\}$.

Remark 6.1.16 It is just the matter of taste if for the box constraint (6.18) related to the state $\widetilde{Y}$ we use the average temperature in the complete GS, $\overline{Q}^S = C^S\widetilde{Y}^{\nu}_{n+1}$ or only in the medium (soil), $C^M\widetilde{Y}^{\nu}_{n+1}$. We motivate the restriction of just a single constraint to some average temperature and allow for local temperatures outside $[\underline{q}, \overline{q}]$ by the fact that the diffusion will average out fluctuations in the local temperature. Fast varying local temperature fluctuations can be found in the fluid of the pipe and starting charging or discharging changes the temperature of the fluid withing a very short period of time. Therefore, imposing constraints to the average temperature $\overline{Q}^M$ in the medium only, seems to be more appropriate.

Set of feasible actions $\mathcal{U}_P(n,x)$ related to the state constraint to P. Let $0 < \varepsilon \ll 1$ be a small probability (tolerance level) which we accept for the violation of the strict constraint $P_n \in [\underline{p}, \overline{p}]$, $n \in \{0, \dots N\}$. We define the conditional probabilities

$$\overline{\pi}^{\nu} = \overline{\pi}^{\nu}(n,x) = \mathbb{P}(P^{\nu}_{n+1} > \overline{p} \,|\, X_n = x, u_n = \nu),$$
$$\underline{\pi}^{\nu} = \underline{\pi}^{\nu}(n,x) = \mathbb{P}(P^{\nu}_{n+1} < \underline{p} \,|\, X_n = x, u_n = \nu).$$

Remark 6.1.17 The above conditional probabilities depend only on $P_n = p$ and $R_n = r$ but not on $\widetilde{Y}_n = y$ except for $\nu = u^D$ (discharging IS to charge the geothermal). For $\nu = u^D$ there is a dependence of P on the outlet temperature of the PHX which is a function of y.

Assumption 6.1.18 We assume that the following hold

1) For $X_n = x$

$$P^{\nu=u^F}_{n+1} > P^{\nu=u^C}_{n+1} > P^{\nu=u^W}_{n+1} > P^{\nu=u^D}_{n+1} \quad \text{a.s.}$$

2) A full internal or GS cannot be driven empty (or almost empty) within one period of length Δ_N. Similarly, charging the internal or GS must be such that an empty internal or GS is not full (or almost full) at the end of one period of length Δ_N.

Under Assumption 6.1.18, the following monotonicity properties of $\underline{\pi}^\nu$ and $\overline{\pi}^\nu$ hold

$$\overline{\pi}^F > \overline{\pi}^C > \overline{\pi}^W > \overline{\pi}^D \text{ and } \underline{\pi}^F < \underline{\pi}^C < \underline{\pi}^W < \underline{\pi}^D.$$

Further, we define the subsets of the state space $\mathcal{X} \subset \mathbb{R}^{\ell+3}$ by

$$\overline{\mathcal{X}}_P^\nu(n) = \{x \in \mathcal{X} \mid \overline{\pi}^\nu(n,x) \le \varepsilon\} \text{ and } \underline{\mathcal{X}}_P^\nu(n) = \{x \in \mathcal{X} \mid \underline{\pi}^\nu(n,x) \le \varepsilon\}.$$

For a given state $X_n = x$ at time $t = t_n$, an action is feasible with respect to the constraint to P if $x \in \overline{\mathcal{X}}_P^\nu(n) \cap \underline{\mathcal{X}}_P^\nu(n)$. Thus the set of feasible controls with respect to P can be written as

$$\mathcal{U}_P(n,x) = \bigcup_{\nu:\, x \in \overline{\mathcal{X}}_P^\nu(n) \cap \underline{\mathcal{X}}_P^\nu(n)} \{\nu\},$$

and contains those actions ν for which x is in the above mentioned intersection. In view of our model setting and the dynamics of the state process it can be deduced that the projection of the above subsets onto the sub-state spaces

$$\mathcal{X}_{RP}^\nu = \{(r,p) \mid x = (r,f,p,y) \in \mathcal{X}\},$$

have the following form

$$\overline{\mathcal{X}}_{RP}^\nu(n) = \{(r,p) \mid p \le \overline{h}_n^\nu(r,y)\}, \qquad \underline{\mathcal{X}}_{RP}^\nu(n) = \{(r,p) \mid p \ge \underline{h}_n^\nu(r,y)\},$$

for some functions $\overline{h}_n^\nu, \underline{h}_n^\nu : \mathbb{R} \times \mathcal{Y} \to \mathbb{R}$, see Fig. 6.2.
Note that for $\nu = u^O,\ u^W,\ u^C,\ u^F$ there is no dependence of the functions $\overline{h}_n^\nu, \underline{h}_n^\nu$ on the variable y since in that case the dynamics of the process P does not depend on the state of the GS. However, for $\nu = u^D$ the state process P depends on the outlet temperature in the GS, i.e., there is a dependence of the functions $\overline{h}_n^\nu,\ \underline{h}_n^\nu$ on the variable y.

The characterization of the set of feasible control related to P considered below requires to truncate the space for the residual demand into some bounded interval $\overline{\mathcal{R}} = [\underline{r}, \overline{r}]$, where $\underline{r}$ and $\overline{r}$ are the minimum and the maximum residual demand, respectively, which are determine using the so called 3-sigma rule. Indeed, the residual demand is modeled using Ornstein–Uhlenbeck (OU) process. Its range $\mathcal{R} = \mathbb{R}$ can be replaced by a closed interval $\overline{\mathcal{R}} = [\underline{r}, \overline{r}]$, in which the values of the random process R lie with high probability. With regard to the stationary distribution of the OU-process with parameters $\mu_R(t) \in [\underline{\mu}_R, \overline{\mu}_R]$ and $\beta_R \in \mathbb{R}$, the 3-σ rule motivates the following choice of the limits for $\underline{r}$ and $\overline{r}$

$$\overline{\mathcal{R}} = [\underline{r}, \overline{r}] = \left[\underline{\mu}_R - 3\frac{\sigma_R}{\sqrt{2\beta_R}}, \overline{\mu}_R + 3\frac{\sigma_R}{\sqrt{2\beta_R}}\right].$$

Characterization of the set $\mathcal{U}_P(n,x)$. The set $\mathcal{U}_P(n,x)$ can be subdivided into 7 or 8 subsets, see Fig 6.3. This partitioning of the set $\mathcal{U}_P(n,x)$ can also be seen in the first column of Fig 6.4,

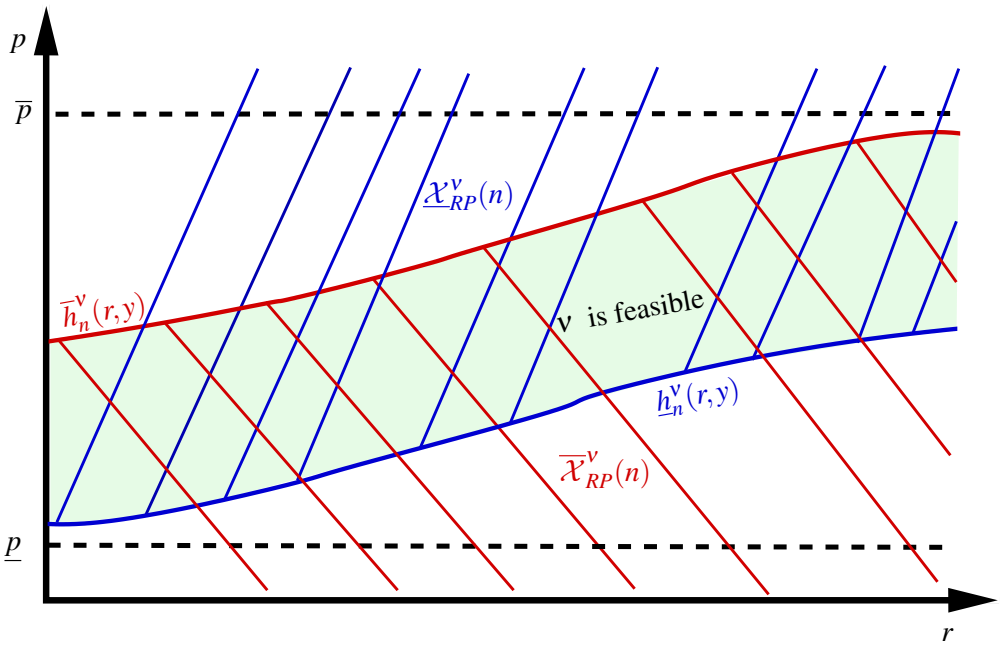


Figure 6.2: Projection of $\underline{\mathcal{X}}^{\nu}_{P}(n)$ and $\overline{\mathcal{X}}^{\nu}_{P}(n)$ onto $\mathcal{X}_{RP}$

where we compute all elements of the intersection $\mathcal{U}_P(n,x) \cap \mathcal{U}_Y(n,x)$. The critical situation occurs when the IS is full, $P_n \geq \overline{p}$ and there is maximal overproduction, $R_n = \underline{r} < 0$. In this case, we can only discharge the IS if the GS is not full, $\overline{Q}^M_n = C^M \widetilde{Y}_n < \overline{q}$ or apply over-spilling if the latter is full and the set of feasible actions is given by $\mathcal{U}_P(n,x) = \{u^O, u^D\}$. Now when the IS is empty, $P_n \leq \underline{p}$ and there is unsatisfied demand, $R_n > 0$ the right action to be taken is to charge the IS by firing fuel if the GS is empty or by discharging the GS if it is not empty. In this case, the set of feasible actions is given by $\mathcal{U}_P(n,x) = \{u^C, u^F\}$. However the truncated parameters $\underline{r}$ and $\overline{r}$ for the residual demand R and the rate of charging/discharging should be chosen such that

- when the IS is empty and there is maximal unsatisfied demand, $(R_n, P_n) = (\overline{r}, \underline{p})$ firing fuel is enough to keep the temperature P_{n+1} above $\underline{p}$ (with probability $1-\varepsilon$, i.e. $\underline{\pi}^F < \varepsilon$),

- when the IS is full and there is maximal overproduction, $(R_n, P_n) = (\underline{r}, \overline{p})$ discharging the IS with not completely full GS is enough to keep the temperature P_{n+1} below $\overline{p}$ (with probability $1-\varepsilon$, i.e. $\overline{\pi}^D < \varepsilon$).

Remark 6.1.19 The above requirements for $(R_n, P_n) = (\underline{r}, \overline{p})$ can be modified such that we include only over-spilling for $R_n = \underline{r}$ and $P_n \in [\overline{p} - \delta, \overline{p}]$, for some small $\delta > 0$. This corresponds to a down-shift of $\overline{h}^D_n$ in Fig. 6.3.

Finally the state dependent control constraints given by the intersection of the sets $\mathcal{U}_P(n,x)$ and $\mathcal{U}_Y(n,x)$ is then given in Fig. 6.4 which contain only 12 potential different subsets. Further, this can be reduced to 10 subsets by changing and merging subsets for $\mathcal{U}_P(n,x)$. For example $\{u^F\}$ and $\{u^C, u^F\}$ can be merged to $\{u^F\}$, and $\{u^D, u^W\}$ and $\{u^O, u^D\}$ can be merged to $\{u^O, u^D\}$. This leads to a slight modification of the set of feasible controls related to P from 7 subsets to 5 and consequently reduce the set of feasible control $\mathcal{U}_P(n,x)$ from 12 subset to 10.

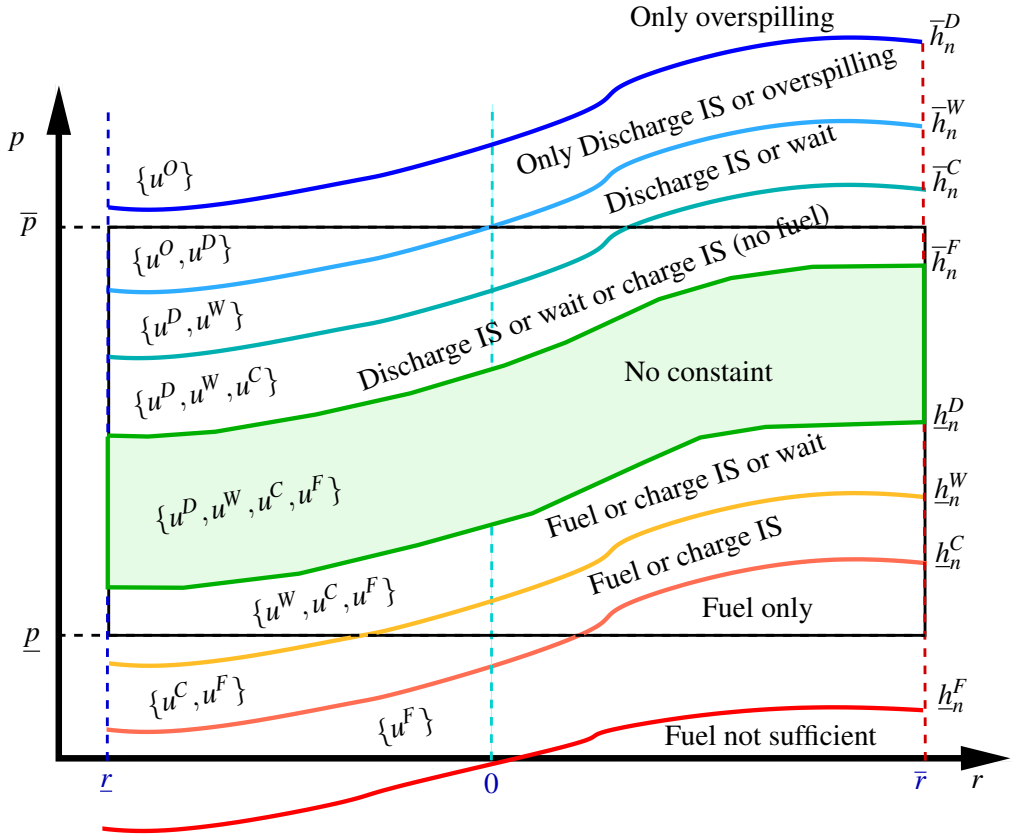


Figure 6.3: Characterization of the set of feasible control $\mathcal{U}_P(n,x)$

Geothermal storage

Internal storage

$\mathcal{U}_P(n,x)$ \ $\mathcal{U}_Y(n,x)$	GS empty $\{u^D, u^W, u^F\}$	No constraint $\{u^D, u^W, u^C, u^F\}$	GS full $\{u^D, u^W, u^C, u^F\}$
IS empty $\{u^F\}$	$\{u^F\}$	$\{u^F\}$	$\{u^F\}$
$\{u^C, u^F\}$	$\{u^F\}$	$\{u^C, u^F\}$	$\{u^C, u^F\}$
$\{u^W, u^C, u^F\}$	$\{u^W, u^F\}$	$\{u^W, u^C, u^F\}$	$\{u^W, u^C, u^F\}$
No constraint $\{u^D, u^W, u^C, u^F\}$	$\{u^D, u^W, u^F\}$	$\{u^D, u^W, u^C, u^F\}$	$\{u^W, u^C, u^F\}$
$\{u^D, u^W, u^C\}$	$\{u^D, u^W\}$	$\{u^D, u^W, u^C\}$	$\{u^W, u^C\}$
$\{u^D, u^W\}$	$\{u^D, u^W\}$	$\{u^D, u^W\}$	$\{u^W\}$
IS full $\{u^O, u^D\}$	$\{u^D\}$	$\{u^D\}$	$\{u^O\}$

Figure 6.4: Set of feasible controls $\mathcal{U}(n,x) = \mathcal{U}_P(n,x) \cap \mathcal{U}_Y(n,x)$

6.1.4 Discrete-Time Optimal Control Problem

In this subsection we consider a Markov Decision Processes with a finite time horizon T and finite action spaces $\mathcal{U}(n,x)$ and the state $\mathcal{X}$ described above. We show that the associated optimization problem can be solved by a backward recursion algorithm. We refer to [11] and references therein for more details in the theory of Markov decision processes with general state and action spaces.

Admissible control. We denote by $\mathcal{A}$ the class of admissible controls, consisting of Markovian control processes $\mathbf{u}$ being adapted w.r.t. the filtration $\mathbb{G}$, satisfying the control constraints (described above) such that the controlled state X^u takes at any time t_n values in the state space $\mathcal{X}$, i.e.,

$$\mathcal{A} = \Big\{ \mathbf{u} = (u_0, \dots, u_{N-1}) \big| \quad \mathbf{u} \text{ is } \mathcal{U}(n,x)\text{-adapted}, u_n = \widetilde{u}(n, X_n^u) \text{ for all } n = 0,1,\dots,N-1, \\ \widetilde{u}(n,x) \in \mathcal{U}(n,x) \text{ for all } (n,x) \in \{0,1,\dots,N-1\} \times \mathcal{X} \Big\}.$$

Performance criterion. Given a control process
$\mathbf{u} = (u_0, u_1, u_2, \dots, u_{N-1})$ the performance criterion $J : \{0,1,\dots,N\} \times \mathcal{X} \times \overline{\mathcal{U}} \to \mathbb{R}$ is the expected aggregated cost over the time $n = 0,1,\dots,N$ given by

$$J(n,x;\mathbf{u}) = \mathbb{E}_{n,x}\left[\sum_{k=n}^{N-1} \Psi(k, X_k^{\mathbf{u}}, u_k) + \phi(X_N^{\mathbf{u}}) \right]$$

for $X_n^{\mathbf{u}} \in \mathcal{X} \subset R^{l+3}$. Here, Ψ is the the running cost given in (2.21), ϕ is the terminal cost given in (2.22) and $\mathbb{E}_{n,x}[.]$ denotes the conditional expectation given that at time n the state $X_n^{\mathbf{u}} = x$.

Optimal control problem. The objective is to minimize the performance criterion J given above over all admissible controls $u \in \mathcal{A}$. We define the value function for all $x \in \mathcal{X}$ and $n = 0,1,\dots,N$ by

$$V(n,x) = \inf_{\mathbf{u} \in \mathcal{A}} J(n,x;\mathbf{u}).$$

A control $\mathbf{u}^* = (u_0^*, u_1^*, \dots, u_{N-1}^*) \in \mathcal{A}$ is called optimal control if $V(n,x) = J(n,x;\mathbf{u}^*)$.

Dynamic programming equation. The Bellman principle presented in Bäuerle and Rieder [11] leads to the following necessary optimality condition called Bellman equation or dynamic programming equation (DPE).

Theorem 6.1.20 (Bellman equation) The value function satisfies the Bellman equation

$$V(N,x) = \phi(x), \quad x \in \mathcal{X}, \\ V(n,x) = \inf_{a \in \mathcal{U}(n,x)} \Big\{ \Psi(n,x,a) + \mathbb{E}_{n,x}\big[V(n+1, X_{n+1}^{\mathbf{u}})\big] \Big\}, \; x \in \mathcal{X}, \; n = 0,1,\dots,N-1. \tag{6.19}$$

For all $n = 0, 1, \ldots, N-1$, the candidate for the optimal control is $u_n^* = \widetilde{u}^*(n, X_n^{\mathbf{u}^*})$, with the optimal decision rule given by

$$\widetilde{u}^*(n,x) = \underset{a \in \mathcal{U}(n,x)}{\operatorname{argmin}} \left\{ \Psi(n,x,a) + \mathbb{E}_{n,x}\left[V(n+1, X_{n+1}^{\mathbf{u}})\right] \right\}.$$

The dynamic programming equation (6.19) can be solved using the following backward recursion algorithm starting at the terminal time N.

Algorithm 2: Backward recursion algorithm

Result: Find the value function V and the optimal strategy $\mathbf{u}^*$

Step 1 Compute for all $x \in \mathcal{X}$

$$V(N,x) = \phi(x)$$

Step 2 For $n := N-1, \ldots, 1, 0$ compute for all $x \in \mathcal{X}$

$$V(n,x) = \inf_{a \in \mathcal{U}(n,x)} \left\{ \Psi(n,x,a) + \mathbb{E}_{n,x}\left[V(n+1, X_{n+1}^{\mathbf{u}})\right] \right\}.$$

Compute the minimizer u_n^* of V_{n+1} given by

$$\widetilde{u}^*(n,x) = \underset{a \in \mathcal{U}(n,x)}{\operatorname{argmin}} \left\{ \Psi(n,x,a) + \mathbb{E}_{n,x}\left[V(n+1, X_{n+1}^{\mathbf{u}})\right] \right\}.$$

The challenge of the implementation of the backward recursion algorithm is that it becomes computationally intractable if the the dimension of the state space is high or if no closed-form expressions of the expectation $\mathbb{E}_{n,x}\left[V(n+1, X_{n+1}^{\mathbf{u}})\right]$ are available. In the next section we discretize the state space to form the MDP for controlled finite-state Markov chain. Then we can express the conditional expectation $\mathbb{E}_{n,x}\left[V(n+1, X_{n+1}^{\mathbf{u}})\right]$ in terms of the transition probabilities of the state of the Markov chain.

6.2 Numerical Solution of the Markov Decision Process

In this section we approximate the continuous-state MDP by a MDP for controlled finite-state Markov chain and compute the associated transition probabilities. For the sake of simplicity we restrict to a model with deterministic fuel price $F(t) = F, F > 0$ and remove F from the state process. The case of random fuel price F can be treated analogously to R. The state space of the MDP is given by

$$\mathcal{X} = \mathcal{R} \times \mathcal{P} \times \mathcal{Y}^1 \times \mathcal{Y}^2 \times \ldots \times \mathcal{Y}^\ell \subset \mathbb{R}^{\ell+2}.$$

Since the state constraint for the GS involve only the average temperature, we first want to construct a suitable basis for the state space $\mathcal{Y}$ in which the average can be given by only one component of the reduced order states.

Suitable basis vectors for the state space $\mathcal{Y}$. Recall that balanced truncation model order reduction yields a ℓ-dimensional subspace $\mathcal{Y} \subset \mathbb{R}^\ell$ for the state component $\widetilde{Y}$. The optimal control problem imposes state constraints to $\widetilde{Y}$ of the form $\underline{q} \leq C^M \widetilde{Y} \leq \overline{q}$. For the approximation of the continuous-state MDP into a MDP for finite-states Markov chain one has to keep in mind that the hyperplane defined by the above state constraint to $\widetilde{Y}$ may not always be parallel to the axis and such that these hyperplanes contain grid points. The state discretization requires the truncation of $\mathbb{R}^\ell$ to a bounded subset (see below in Subsec. 6.2.1) in which a finite subset of grid points is chosen. Typically the bounded subset is an ℓ-dimensional rectangle and the grid points are placed on lines parallel to the axis, see Fig. 6.5. To keep the approximate solution of

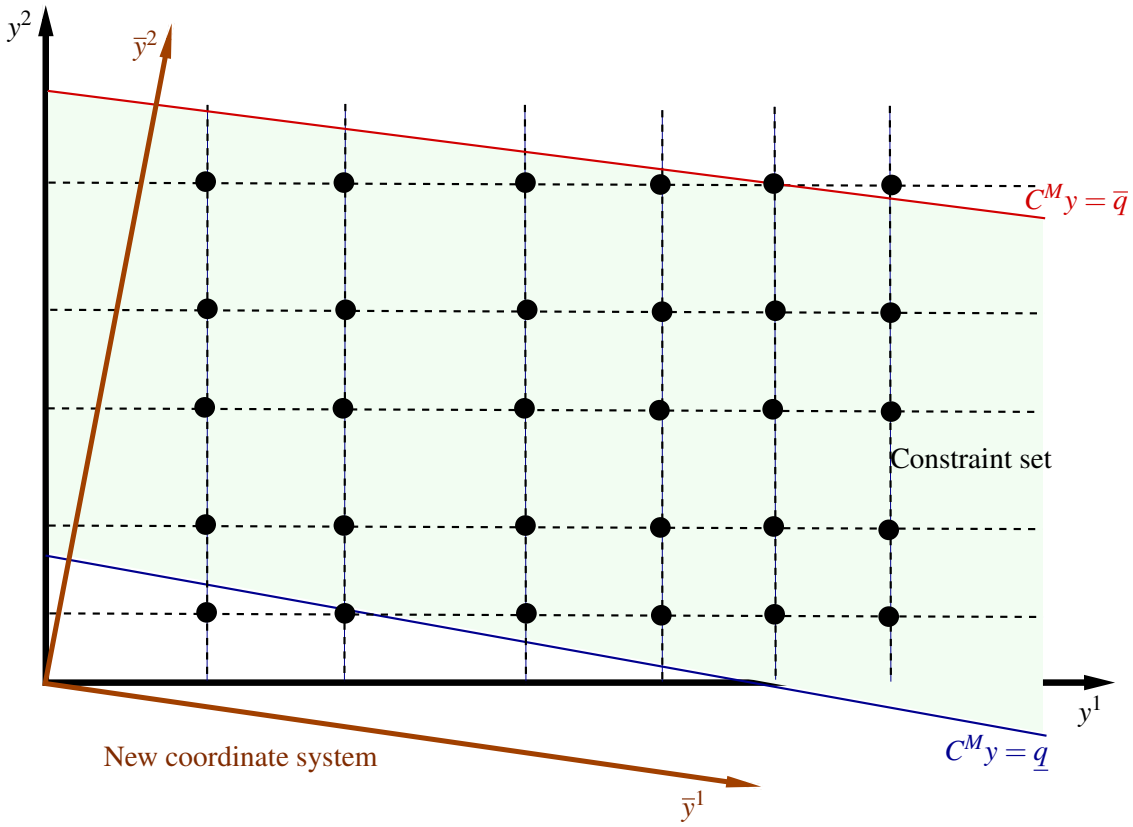


Figure 6.5: Change of coordinate system for the reduced order system

the control problem computation tractable, the number of grid points in the domain $\mathcal{Y}$ should be as small as possible. Thus it is important to find a appropriate

- truncation to a bounded subset of $\mathbb{R}^\ell$
- location of the grid points within this bounded subset

In the view of the geometry of the constraint set with respect to the reduced order state $\widetilde{Y}$ which is a subset of $\mathcal{Y}$ between the two hyperplanes defined by

$$C^M y = \underline{q} \quad \text{and} \quad C^M y = \overline{q},$$

it is advisable to place the grid points on the hyperplanes parallel to the above two hyperplanes. This can be simplified by choosing new coordinates in $\mathcal{Y}$ such that the new basis vectors are

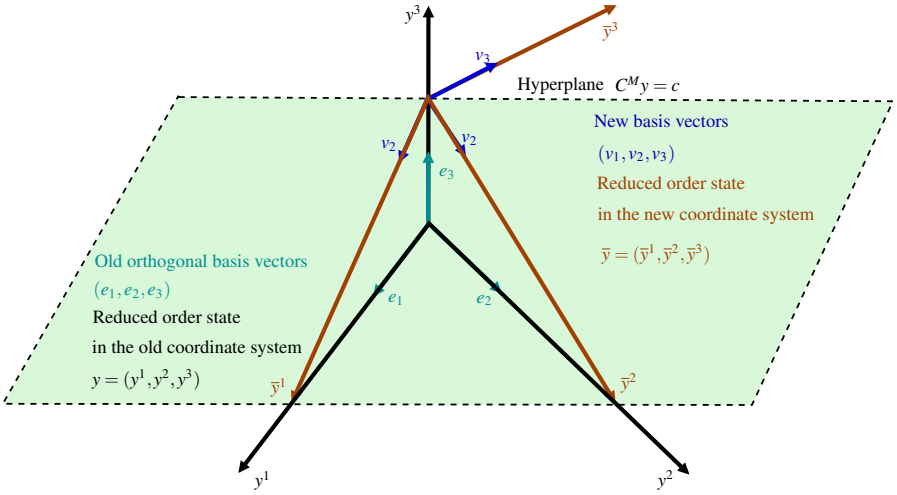


Figure 6.6: Basis vectors

either contained in the hyperplanes or orthogonal to the hyperplanes.

Construction of the new basis. To find the new basis we can proceed as follows

1. choose $\ell - 1$ linear independent vectors $v_1, v_2, \ldots, v_{\ell-1}$ on the hyperplane

2. find the vector v_ℓ such that $\langle v_\ell, v_k \rangle = 0$ for $k = 1, \ldots, \ell - 1$, i.e. v_ℓ is orthogonal to the first $\ell - 1$ vectors, see Fig. 6.6. The orthogonal vector v_ℓ obtained using Gram-Schmidt orthonalization method.

Details on how new basis vectors can be practically chosen are given in Appendix C.4.1. For simplicity we denote again by $v_1, v_2, \ldots, v_\ell$ the orthogonal new basis vectors. The following lemma gives the average temperature and the output matrix in the transformed coordinates system.

Lemma 6.2.1 Let $v_1, v_2, \ldots, v_\ell$ be the orthogonal new basis vectors. Then, in the new basis the

1. hyperplanes $C^M \overline{Y} = c$ for a constant c are parallel to $v_1, v_2, \ldots, v_{\ell-1}$ and orthogonal to v_ℓ,

2. last coordinate of $\overline{Y}$ denoted by $\overline{Y}^\ell$ is up to a scaling constant equal to the average temperature in the GS,i.e. $\overline{Q}^M = \overline{C}^{M,\ell} \overline{Y}^\ell$,

3. Row matrix $\overline{C}^M$ is given by

$$\overline{C}^M = (0, \ldots, 0, \overline{C}^{M,\ell}) \quad \text{with} \quad \overline{C}^{M,\ell} = C^M v_\ell.$$

The proof of this lemma can be found in Appendix C.4.2.
It is helpful to work with a non-equidistant discretization in the $\overline{Y}^\ell$-direction and to place more grid points close to $\overline{C}^{M,\ell} \overline{Y}^\ell = \underline{q}$, $\overline{q}$. This allows a more sensitive response to the state constraint.

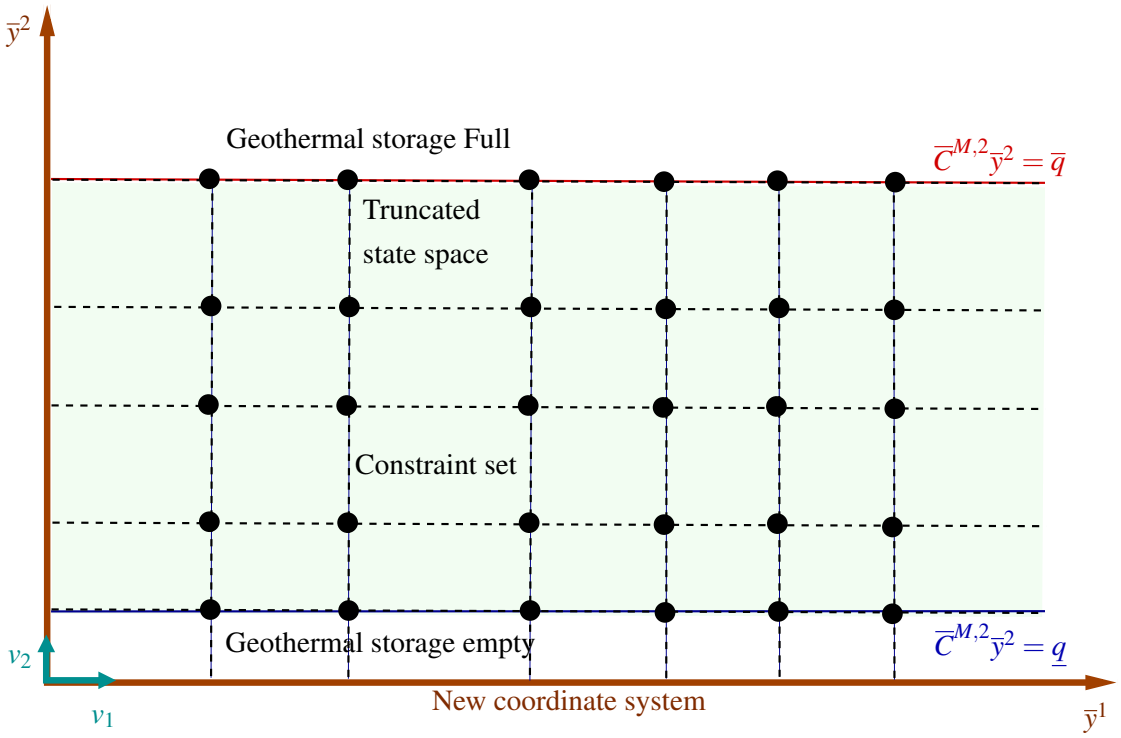


Figure 6.7: New basis vectors for the truncated state space ($\ell = 2$)

6.2.1 State Discretization

In this subsection we approximate the above continuous-state MDP by a MDP for finite-states Markov chain.
Let N_r and N_p be the number of grid points in r and p-directions, respectively and let N_{y^k}, $k = 1,\dots,\ell$ be the number of grid points in y^k-direction. Let $r_0 < r_1 < \dots < r_{N_r}$ and $p_0 < p_1 < \dots < p_{N_p}$ be finitely many grid points in r and p-directions, respectively. Let $y_0^1 < y_2^1 < \dots < y_{N_{y^1}}^1$, $y_0^2 < y_2^2 < \dots < y_{N_{y^2}}^2, \dots$, $y_0^\ell < y_2^\ell < \dots < y_{N_{y^\ell}}^\ell$ be finitely many grid points in y^1, $y^2,\dots,y^\ell$-directions, respectively. Let h_{r_i}, $i = 0,1,\dots,N_r-1$, h_{p_j}, $j = 0,1,\dots,N_p-1$, and $h_{y^i_{k_i}}$, $k_i = 0,\dots,N_{y^i}-1$, $i = 1,\dots,\ell$ be non equidistant step sizes in r, p and y^i-directions, respectively, with

$$\begin{aligned} h_{r_i} &= r_{i+1} - r_i, \quad h_{p_j} = p_{j+1} - p_j, \\ h_{y^1_{k_1}} &= y^1_{k_1+1} - y^1_{k_1}, \quad h_{y^2_{k_2}} = y^2_{k_2+1} - y^2_{k_2}, \dots, \quad h_{y^\ell_k} = y^\ell_{k_2+1} - y^\ell_{k_2}. \end{aligned} \tag{6.20}$$

Then, the $(\ell+2)$-dimensional discretized state space is given by

$$\widetilde{\mathcal{X}} = \widetilde{\mathcal{R}} \times \widetilde{\mathcal{P}} \times \widetilde{\mathcal{Y}} = \{r_0,\dots,r_{N_r}\} \times \{p_0,\dots,p_{N_p}\} \times \{y_0^1,\dots,y_{N_{y^1}}^1\} \times \dots \times \{y_0^\ell,\dots,y_{N_{y^\ell}}^\ell\}$$

Let now denote

$$\mathcal{N}_r = \{0,1,...,N_r\}, \quad \mathcal{N}_p = \{0,1,...,N_p\}, \text{ and } \mathcal{N}_{y^k} = \{0,1,...,N_{y^k}\}, \ k = 1,2\dots,\ell,$$

the set of indices for R, P and $\widetilde{Y}^k$, $k = 1,\dots,\ell$, respectively and let $\widetilde{\mathcal{N}}$ be the set of $(\ell+2)$-tuples

of multi-indices defined by

$$\begin{aligned}\widetilde{\mathcal{N}} &= \mathcal{N}_r \times \mathcal{N}_p \times \mathcal{N}_{y^1} \times \ldots \times \mathcal{N}_{y^\ell} \\ &= \{(i,j,k_1,k_2,\ldots,k_\ell),\ i \in \mathcal{N}_r,\ j \in \mathcal{N}_p,\ k_1 \in \mathcal{N}_{y^1},\ldots,\ k_\ell \in \mathcal{N}_{y^\ell}\}.\end{aligned}$$

A point $x_m \in \widetilde{\mathcal{X}}$, with $m = (i,j,k_1,k_2,\ldots,k_\ell) \in \widetilde{\mathcal{N}}$ is defined by

$$x_m = (r_i, p_j, y^1_{k_1}, y^2_{k_2}, \ldots, y^\ell_{k_\ell}), \ \text{ with } \ r_i \in \widetilde{\mathcal{R}},\ p_j \in \widetilde{\mathcal{P}},\ y^1_{k_1} \in \widetilde{\mathcal{Y}}^1, \ldots, y^\ell_{k_\ell} \in \widetilde{\mathcal{Y}}^\ell.$$

This discretization converts the given MDP with state space $\mathcal{X}$ into a MDP for controlled finite-state Markov chain with state space $\widetilde{\mathcal{X}}$. Note that the discrete-state approximation is a Markov process. This property is inherited from the continuous-state process X.
Let us denote by $\mathcal{N}_n = \{0,...,N\}$ the set of time indices for discrete time points $t_0,\ldots,t_N$. For $(n,x_m) \in \mathcal{N}_n \times \widetilde{\mathcal{X}}$, we define the approximate value function and decision rule on the grid point $x_m = (r_i, p_j, y^1_{k_1}, y^2_{k_2}, \ldots, y^\ell_{k_\ell}) \in \widetilde{\mathcal{X}}$ at time $n \in \mathcal{N}_n$ by

$$V^D(n,x_m) = V(t_n, r_i, p_j, y^1_{k_1}, y^2_{k_2}, \ldots, y^\ell_{k_\ell}) \text{ and } u_n = u_n(x_m) = u(t_n, r_i, p_j, y^1_{k_1}, y^2_{k_2}, \ldots, y^\ell_{k_\ell}),$$

respectively.
We recall that $X^{\mathbf{u}}_n = (R_n, P_n, \widetilde{Y}^1_n, \ldots, \widetilde{Y}^\ell_n) \in \overline{\mathcal{X}}$ is the continuous-state process at discrete time n and $X^{\mathbf{u}}_{n+1} = (R_{n+1}, P_{n+1}, \widetilde{Y}^1_{n+1}, \ldots, \widetilde{Y}^\ell_{n+1}) \in \overline{\mathcal{X}}$ is the continuous-state process at discrete time $n+1$.
We denote by $X^{\mathbf{u},D}_n = (R^D_n, P^D_n, \widetilde{Y}^{1,D}_n, \ldots, \widetilde{Y}^{\ell,D}_n) \in \widetilde{\mathcal{X}}$ the discrete-state process at discrete time n and by $X^{\mathbf{u},D}_{n+1} = (R^D_{n+1}, P^D_{n+1}, \widetilde{Y}^{1,D}_{n+1}, \ldots, \widetilde{Y}^{\ell,D}_{n+1}) \in \widetilde{\mathcal{X}}$ the discrete-state process at discrete time $n+1$.
Assume that at time $n \in \mathcal{N}_n$ the state is on the grid point $x_{m_1} \in \widetilde{\mathcal{X}}$ and the action $\nu \in \mathcal{U}$ is taken. Then, the state moves to a grid point $x_{m_2} \in \widetilde{\mathcal{X}}$ at time $n+1$ with some probability $\mathbf{P}^\nu_{x_{m_1},x_{m_2}}$. This probability is the so-called transition probability which is the probability that the state moves from x_{m_1} at time n to x_{m_2} at time $n+1$ under the action ν and it is defined by

$$\mathbf{P}^\nu_{x_{m_1},x_{m_2}} = \mathbb{P}^\nu(X^{\mathbf{u},D}_{n+1} = x_{m_2} \ \mid \ X^{\mathbf{u},D}_n = x_{m_1},\ u_n = \nu).$$

These probabilities are required for the above algorithm for the computation of the conditional expectation of the value function at time $n+1$ given that at time n the state $X^{\mathbf{u}}_n = x_{m_1} \in \widetilde{\mathcal{X}}$. It holds

$$\begin{aligned}\mathbb{E}\big[V^D(n+1, X^{\mathbf{u},D}_{n+1}) \mid \ X^{\mathbf{u},D}_n = x_{m_1}\big] &= \sum_{x_{m_2} \in \widetilde{\mathcal{X}}} \mathbb{P}^\nu(X^{\mathbf{u},D}_{n+1} = x_{m_2} \ \mid \ X^{\mathbf{u},D}_n = x_{m_1},\ u_n = \nu) V^D(n+1, x_{m_2}) \\ &= \sum_{x_{m_2} \in \widetilde{\mathcal{X}}} \mathbf{P}^\nu_{x_{m_1},x_{m_2}} V^D(n+1, x_{m_2}).\end{aligned}$$

The following relations between the discrete-state process and the continuous-state process hold true

$$R^D_n = r_i \Longleftrightarrow R_n \in \left(r_i - \frac{1}{2}h_{r_{i-1}}, r_i + \frac{1}{2}h_{r_i}\right], \quad i = 1,\ldots,N_r - 1$$

$$
\begin{aligned}
P_n^D = p_j \Longleftrightarrow P_n \in \left(p_j - \frac{1}{2}h_{p_{j-1}}, p_j + \frac{1}{2}h_{p_j} \right], \quad j = 1, \ldots, N_p - 1 \\
\widetilde{Y}_n^{1,D} = y_{k_1}^1 \Longleftrightarrow \widetilde{Y}_n^{1,D} \in \left(y_{k_1}^1 - \frac{1}{2}h_{y_{k_1-1}^1}, y_{k_1}^1 + \frac{1}{2}h_{y_{k_1}^1} \right], \quad k_1 = 1, \ldots, N_{y^1} \\
\vdots \\
\widetilde{Y}_n^{\ell,D} = y_{k_\ell}^\ell \Longleftrightarrow \widetilde{Y}_n^{\ell,D} \in \left(y_{k_\ell}^\ell - \frac{1}{2}h_{y_{k_\ell-1}^\ell}, y_{k_\ell}^\ell + \frac{1}{2}h_{y_{k_\ell}^\ell} \right], \quad k_\ell = 1, \ldots, N_{y^\ell},
\end{aligned}
$$

where, $h_{r_i},\ i = 0, 1, \ldots, N_r - 1$, $h_{p_j},\ j = 0, 1, \ldots, N_p - 1$, and $h_{y_{k_i}^i}$, $k_i = 0, \ldots, N_{y^i} - 1$, $i = 1, \ldots, \ell$ are non equidistant step sizes given by (6.20).
The next subsection will be devoted to the computation of the transition probabilities.

6.2.2 Computation of the Transition Probabilities

In this subsection we are going describe how the transition probabilities introduced above can be computed practically. We recall that given the state X_n and the decision rule u_n a time n, the state of the residual demand R_{n+1} and the state of the temperature in the storage P_{n+1} (for $u_n \neq u^O$) at time $n+1$ are Gaussian random variables and the pair (R_{n+1}, P_{n+1}) is bivariate Gaussian. Further, the reduced order states $\widetilde{Y}_{n+1}^1, \widetilde{Y}_{n+1}^2, \ldots, \widetilde{Y}_{n+1}^\ell$ of the GS at time $n+1$ are degenerated (Dirac). Then, $\widetilde{Y}_{n+1}^1, \widetilde{Y}_{n+1}^2, \ldots, \widetilde{Y}_{n+1}^\ell$ are mutually independent and independent of R_{n+1} and P_{n+1}. Therefore, $\widetilde{Y}_{n+1}^1, \widetilde{Y}_{n+1}^2, \ldots, \widetilde{Y}_{n+1}^\ell$ are independent of the pair (R_{n+1}, P_{n+1}).
For fixed $y^1, y^2, \ldots, y^\ell$ the computational grid in (r,p)-plane is sketched in Fig. 6.8. For the computation of the transition probabilities one has to distinguish the inner and boundary points (including the corners). Let $r_i \in \widetilde{\mathcal{R}}, i \in \mathcal{N}_r$ and $p_j \in \widetilde{\mathcal{P}}, j \in \mathcal{N}_p$. Then, in (r,p)-plane the inner points are $(r_i, p_j) \in \widetilde{\mathcal{R}} \times \widetilde{\mathcal{P}}$ with $i = 1, 2, \ldots, N_r - 1$, and $j = 1, 2, \ldots, N_p - 1$, and the boundary points are $(r_i, p_j) \in \widetilde{\mathcal{R}} \times \widetilde{\mathcal{P}}$ with $i \in \mathcal{N}_r$ and $j = 0,\ N_p$ or $j \in \mathcal{N}_p$ and $i = 0,\ N_r$.

- For inner grid points we denote by

$$
\begin{aligned}
\mathcal{B}_{r_i} &= \left(r_i - \frac{1}{2}h_{r_{i-1}}, r_i + \frac{1}{2}h_{r_i} \right] = \left(\frac{1}{2}(r_i + r_{i-1}), \frac{1}{2}(r_i + r_{i+1}) \right], \ \ i = 1, 2, \ldots, N_r - 1, \\
\mathcal{B}_{p_j} &= \left(p_j - \frac{1}{2}h_{p_{j-1}}, p_j + \frac{1}{2}h_{p_j} \right] = \left(\frac{1}{2}(p_j + p_{j-1}), \frac{1}{2}(p_j + p_{j+1}) \right], \ \ j = 1, 2, \ldots, N_p - 1,
\end{aligned}
$$

  the neighborhoods of r_i andp_j, respectively.

- For the boundary grid points we denote by

$$
\begin{aligned}
\mathcal{B}_{r_0} &= \left(-\infty, r_0 + \frac{1}{2}h_{r_0} \right] = \left(-\infty, \frac{1}{2}(r_0 + r_1) \right], \\
\mathcal{B}_{r_{N_r}} &= \left(r_{N_r} - \frac{1}{2}h_{r_{N_r-1}},\ +\infty \right) = \left(\frac{1}{2}(r_{N_r} + r_{N_r-1}), +\infty \right), \\
\mathcal{B}_{p_0} &= \left(-\infty, p_0 + \frac{1}{2}h_{p_0} \right] = \left(-\infty, \frac{1}{2}(p_0 + p_1) \right],
\end{aligned}
$$

$$\mathcal{B}_{p_{N_p}} = \left(p_{N_p} - \frac{1}{2} h_{p_{N_p-1}},\ +\infty \right) = \left(\frac{1}{2}(p_{N_p} + p_{N_p-1}), +\infty \right),$$

the neighborhoods of r_0, r_{N_r}, p_0, and p_{N_p}, respectively.

In this setting the joint probability that at time $n+1$ the pair (R^D_{n+1}, P^D_{n+1}) is at the grid point $(r_i, p_j), (i,j) \in \mathcal{N}_r \times \mathcal{N}_p$ is set to be the probability that (R_{n+1}, P_{n+1}) is located in the neighborhood $\mathcal{B}_{ij} = \mathcal{B}_{r_i} \times \mathcal{B}_{p_j}$ of (r_i, p_j), $i \in \mathcal{N}_r$ and $j \in \mathcal{N}_p$ (including the boundary grid points).

Similarly, for $i = 1, \ldots, \ell$, we define the neighborhoods of $y^i_{k_i} \in \widetilde{\mathcal{Y}}^i$, $k_i \in \mathcal{N}_{y^i}$ by

$$\begin{aligned}
\mathcal{B}y^i_{k_i} &= \left(y^i_{k_i} - \frac{1}{2} h_{y^i_{k_i-1}}, y^i_{k_i} + \frac{1}{2} h_{y^i_{k_i}} \right] = \left(\frac{1}{2}(y^i_{k_i} + y^i_{k_i-1}), \frac{1}{2}(y^i_{k_i} + y^i_{k_i+1}) \right], \quad k_i = 1, \ldots, N_{y^i} - 1 \\
\mathcal{B}y^i_0 &= \left(-\infty, y^i_0 + \frac{1}{2} h_{y^i_0} \right] = \left(-\infty, \frac{1}{2}(y^i_0 + y^i_1) \right], \\
\mathcal{B}y_{N_{y^i}} &= \left(y_{N_{y^i}} - \frac{1}{2} h_{N_{y^i}},\ +\infty \right) = \left(\frac{1}{2}(y_{N_{y^i}} + y_{N_{y^i}-1}), +\infty \right).
\end{aligned}$$

In this setting the probability that at time $n+1$ the discrete reduced-order state $\widetilde{Y}^{i,D}_{n+1}$ is at the grid point $y^i_{k_i} \in \widetilde{\mathcal{Y}}^i$, $k_i \in \mathcal{N}_{y_i}$, $i = 1, \ldots, \ell$, is set to be the probability that the continuous reduced-order state $\widetilde{Y}^i_{n+1}$ is located in the neighborhood $\mathcal{B}y^i_{k_i}$ of $y^i_{k_i}$, $k_i \in \mathcal{N}_{y_i}$, $i = 1, \ldots, \ell$.

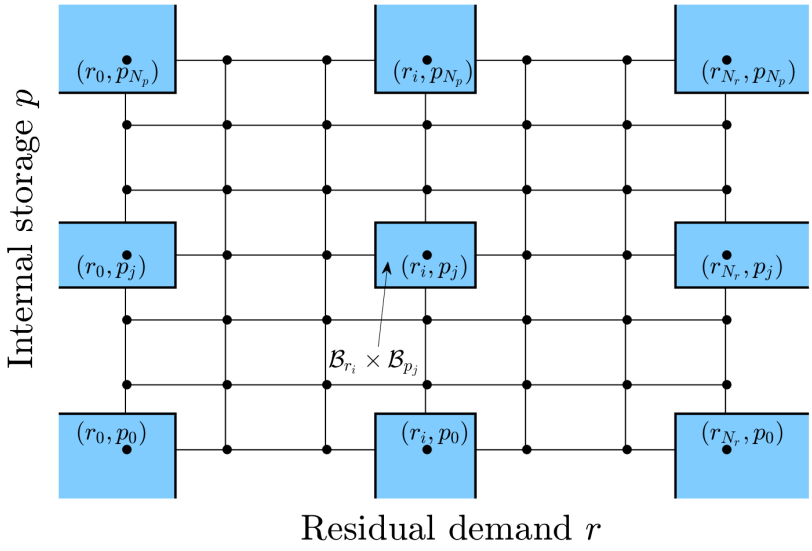


Figure 6.8: Computational grid in (r,p)-plane for fixed $y^1, y^2, \ldots, y^\ell$

Given two points $x_{m_1} = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^\ell_{k^1_\ell}) \in \widetilde{\mathcal{X}}$ and $x_{m_2} = (r_{i_2}, p_{j_2}, y^1_{k^2_1}, \ldots, y^\ell_{k^2_\ell}) \in \widetilde{\mathcal{X}}$, the transition probability that the state moves from x_{m_1} at time n to x_{m_2} at time $n+1$ under the action $u_n = \nu \neq u^O$ is computed as follows

$$\begin{aligned}
\mathbf{P}^\nu_{x_{m_1}, x_{m_2}} &= \mathbb{P}^\nu(X^{\mathbf{u},D}_{n+1} = x_{m_2} \mid X^{\mathbf{u},D}_n = x_{m_1},\ u_n = \nu) \\
&= \mathbb{P}^\nu(X^{\mathbf{u},D}_{n+1} = (r_{i_2}, p_{j_2}, y^1_{k^2_1}, \ldots, y^\ell_{k^2_\ell}) \mid X^{\mathbf{u},D}_n = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^\ell_{k^1_\ell}),\ u_n = \nu)
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{P}^{\nu}((R^D_{n+1}, P^D_{n+1}) = (r_{i_2}, p_{j_2}) \mid X^{\mathbf{u},D}_n = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^{\ell}_{k^1_\ell}),\ u_n = \nu) \\
&\quad \times \Pi^{\ell}_{i=1} \mathbb{P}^{\nu}(\widetilde{Y}^{i,D}_{n+1} = y^i_{k^2_i} \mid \widetilde{Y}^{i,D}_n = y^i_{k^1_i}, u_n = \nu) \\
&= \mathbb{P}^{\nu}((R_{n+1}, P_{n+1}) \in \mathcal{B}_{r_{i_2}} \times \mathcal{B}_{p_{j_2}} \mid X^{\mathbf{u}}_n = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^{\ell}_{k^1_\ell}),\ u_n = \nu) \\
&\quad \times \Pi^{\ell}_{i=1} \mathbb{P}^{\nu}(\widetilde{Y}^{i}_{n+1} \in \mathcal{B}y^i_{k^2_i} \mid \widetilde{Y}^{i}_n = y^i_{k^1_i}, u_n = \nu) \\
&= \mathbb{P}^{\nu}((R_{n+1}, P_{n+1}) \in \mathcal{B}_{i_2 j_2} \mid X^{\mathbf{u}}_n = x_{m_1},\ u_n = \nu) \\
&\quad \times \Pi^{\ell}_{i=1} \mathbb{P}^{\nu}(\widetilde{Y}^{i}_{n+1} \in \mathcal{B}y^i_{k^2_i} \mid \widetilde{Y}^{i}_n = y^i_{k^1_i}, u_n = \nu).
\end{aligned}
$$

For $i = 1, \ldots, \ell$, the probability that at time $n+1$ the state $\widetilde{Y}^i_{n+1}$ is located in the neighborhood $\mathcal{B}y^i_{k^2_i}$ of $y^i_{k^2_i}$ given that at time n, $Y^i_n = y^i_{k^1_i}$ and the action $u_n = \nu$ is taken, is given by

$$
\mathbb{P}^{\nu}(\widetilde{Y}^i_{n+1} \in \mathcal{B}y^i_{k^2_i} \mid \widetilde{Y}^i_n = y^i_{k^1_i}, u_n = \nu) = \begin{cases} 1 & \text{if } \widetilde{Y}^i_{n+1} \in \mathcal{B}y^i_{k^2_i} \\ 0 & \text{else} \end{cases}
$$

Therefore, the product of the probabilities of the independent state variables is given by

$$
\Pi^{\ell}_{i=1} \mathbb{P}^{\nu}(\widetilde{Y}^i_{n+1} \in \mathcal{B}y^i_{k^2_i} \mid \widetilde{Y}^i_n = y^i_{k^1_i}, u_n = \nu) = \begin{cases} 1 & \text{if } \widetilde{Y}^1_{n+1} \in \mathcal{B}y^1_{k^2_1}, \ldots, \widetilde{Y}^{\ell}_{n+1} \in \mathcal{B}y^{\ell}_{k^2_\ell} \\ 0 & \text{otherwise.} \end{cases}
$$

Next, we want to compute the conditional probability that at time t_{n+1} the pair $(R_{n+1}, P_{n+1}) \in \mathcal{B}_{r_{i_2}} \times \mathcal{B}_{p_{j_2}}$ given that the state process $X^{\mathbf{u}}_n = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^{\ell}_{k^1_\ell})$, i.e.

$$
\mathbb{P}^{\nu}((R_{n+1}, P_{n+1}) \in \mathcal{B}_{r_{i_2}} \times \mathcal{B}_{p_{j_2}} \mid X^{\mathbf{u}}_n = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^{\ell}_{k^1_\ell}),\ u_n = \nu).
$$

For the inner grid points (r_{i_2}, p_{j_2}) with $i_2 = 1, \ldots, N_r - 1$ and $j_2 = 1, \ldots, N_p - 1$, we have

$$
\begin{aligned}
&\mathbb{P}^{\nu}((R_{n+1}, P_{n+1}) \in \mathcal{B}_{r_{i_2}} \times \mathcal{B}_{p_{j_2}} \mid X^{\mathbf{u}}_n = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^{\ell}_{k^1_\ell}),\ u_n = \nu) \\
&\quad = \mathbb{P}^{\nu}\left((R_{n+1}, P_{n+1}) \in \left(\frac{1}{2}(r_{i_2} + r_{i_2 - 1}), \frac{1}{2}(r_{i_2} + r_{i_2+1}) \right] \right. \\
&\qquad \left. \times \left(\frac{1}{2}(p_{j_2} + p_{j_2 - 1}), \frac{1}{2}(p_{j_2} + p_{j_2+1}) \right] \mid X^{\mathbf{u}}_n = x_{m_1},\ u_n = \nu \right) \\
&\quad = \int_{\frac{1}{2}(r_{i_2} + r_{i_2 - 1})}^{\frac{1}{2}(r_{i_2} + r_{i_2+1})} \varphi_R(r) \big(\Phi(\zeta_z(r, \delta_{p_{j_2+1}})) - \Phi(\zeta_z(r, \delta_{p_{j_2-1}})) \big) dr,
\end{aligned}
$$

where $\delta_{p_{j_2-1}} = \frac{1}{2}(p_{j_2} + p_{j_2+1})$ and $\delta_{p_{j_2+1}} = \frac{1}{2}(p_{j_2} + p_{j_2+1})$, φ_R is the probability density function of a normal random variable R_{n+1}, $\zeta_z(r,p)$ is a function given by (6.15), and Φ is the cumulative distribution function of a normal random variable ζ_z.
For the boundary grid points (r_{i_2}, p_0) with $i_2 = 1, 2, \ldots, N_r - 1$, we have

$$
\begin{aligned}
&\mathbb{P}^{\nu}((R_{n+1}, P_{n+1}) \in \mathcal{B}_{r_{i_2}} \times \mathcal{B}_{p_0} \mid X^{\mathbf{u}}_n = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^{\ell}_{k^1_\ell}),\ u_n = \nu) \\
&\quad = \mathbb{P}^{\nu}\left((R_{n+1}, P_{n+1}) \in \left(\frac{1}{2}(r_{i_2} + r_{i_2 - 1}), \frac{1}{2}(r_{i_2} + r_{i_2+1}) \right] \right.
\end{aligned}
$$

$$\times \left(-\infty, \frac{1}{2}(p_0+p_1)\right] \mid X_n^{\mathbf{u}} = x_{m_1},\ u_n = \nu \Big)$$
$$= \int_{\frac{1}{2}(r_{i_2}+r_{i_2-1})}^{\frac{1}{2}(r_{i_2}+r_{i_2+1})} \varphi_R(r)\Phi(\zeta_z(r,\delta_{p_1}))dr,$$

where $\delta_{p_1} = \frac{1}{2}(p_0+p_1)$.
For the boundary grid points (r_{i_2}, p_{N_p}) with $i_2 = 1,2,\ldots,N_r-1$, we have

$$\mathbb{P}^\nu((R_{n+1},P_{n+1}) \in \mathcal{B}_{r_{i_2}} \times \mathcal{B}_{p_{N_p}} \mid X_n^{\mathbf{u}} = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^\ell_{k^1_\ell}),\ u_n = \nu)$$
$$= \mathbb{P}^\nu\Big((R_{n+1},P_{n+1}) \in \Big(\frac{1}{2}(r_{i_2}+r_{i_2-1}), \frac{1}{2}(r_{i_2}+r_{i_2+1})\Big]$$
$$\times \Big(\frac{1}{2}(p_{N_p}+p_{N_p-1}), +\infty\Big) \mid X_n^{\mathbf{u}} = x_{m_1},\ u_n = \nu\Big)$$
$$= \int_{\frac{1}{2}(r_{i_2}+r_{i_2-1})}^{\frac{1}{2}(r_{i_2}+r_{i_2+1})} \varphi_R(r)(1-\Phi(\zeta_z(r,\delta_{p_{N_p}})))dr,$$

where $\delta_{p_{N_p}} = \frac{1}{2}(p_{N_p}+p_{N_p-1})$.

Similar reasoning can be adopted for the computations of the transition probabilities for the corner points $(r_0,p_0),(r_{N_r},p_0),(r_0,p_{N_p})$, and (r_{N_r},p_{N_p}).

Remark 6.2.2 Note that the above computations hold only for $u_n \neq u^O$. However, for $u_n = u^O$ the state of the GS P_{n+1} at time $n+1$ is degenerated (Dirac), i.e., the probability that P_{n+1} is located in the neighborhood of p_{j_2} given $X_n^{\mathbf{u}} = x_{m_1} = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^\ell_{k^1_\ell})$ is defined by

$$\mathbb{P}^\nu(P_{n+1} \in \mathcal{B}_{p_{j_2}} \mid X_n^{\mathbf{u}} = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^\ell_{k^1_\ell}),\ u_n = \nu) = \begin{cases} 1 & \text{if } P_{n+1} \in \mathcal{B}_{p_{j_2}}, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, for $u_n = u^O$, P_{n+1} and R_{n+1} are independent.Therefore, the above transition probability can be written as follows

$$\mathbf{P}^\nu_{x_{m_1},x_{m_2}} = \mathbb{P}^\nu\big(X^{\mathbf{u},D}_{n+1} = x_{m_2} \mid X^{\mathbf{u},D}_n = x_{m_1},\ u_n = \nu\big)$$
$$= \mathbb{P}^\nu((R^D_{n+1}, P^D_{n+1}, \widetilde{Y}^{1,D}_{n+1}, \ldots, \widetilde{Y}^{\ell,D}_{n+1}) = (r_{i_2}, p_{j_2}, y^1_{k^2_1}, \ldots, y^\ell_{k^2_\ell}) \mid X^{\mathbf{u},D}_n = x_{m_1},\ u_n = \nu)$$
$$= \mathbb{P}^\nu(R^R_{n+1} = r_{i_2} \mid R^D_n = r_{i_1}) \times \mathbb{P}^\nu(P^D_{n+1} = p_{j_2} \mid X^{\mathbf{u},D}_n = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^\ell_{k^1_\ell}),\ u_n = \nu)$$
$$\times \Pi_{i=1}^\ell \mathbb{P}^\nu(\widetilde{Y}^{i,D}_{n+1} = y^i_{k^2_i} \mid \widetilde{Y}^{i,D}_n = y^i_{k^1_i}, u_n = \nu)$$
$$= \mathbb{P}^\nu(R_{n+1} \in \mathcal{B}_{r_{i_2}} \mid R_n = r_{i_1}) \times \mathbb{P}^\nu(P_{n+1} \in \mathcal{B}_{p_{j_2}} \mid X_n^{\mathbf{u}} = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^\ell_{k^1_\ell}),\ u_n = \nu)$$
$$\times \Pi_{i=1}^\ell \mathbb{P}^\nu(\widetilde{Y}^i_{n+1} \in \mathcal{B}y^i_{k^2_i} \mid \widetilde{Y}^i_n = y^i_{k^1_i}, u_n = \nu)$$
$$= \begin{cases} \mathbb{P}^\nu(R_{n+1} \in \mathcal{B}_{r_{i_2}} \mid R_n = r_{i_1}) & \text{if } P_{n+1} \in \mathcal{B}_{p_{j_2}},\ \widetilde{Y}^1_{n+1} \in \mathcal{B}y^1_{k^2_1}, \ldots, \widetilde{Y}^\ell_{n+1} \in \mathcal{B}y^\ell_{k^2_\ell} \\ 0 & \text{otherwise.} \end{cases}$$
$$= \begin{cases} \Phi\big(\frac{1}{2}(r_{i_2}+r_{i_2-1})\big) - \Phi\big(\frac{1}{2}(r_{i_2}+r_{i_2+1})\big) & \text{if } P_{n+1} \in \mathcal{B}_{p_{j_2}},\ \widetilde{Y}^1_{n+1} \in \mathcal{B}y^1_{k^2_1}, \ldots, \widetilde{Y}^\ell_{n+1} \in \mathcal{B}y^\ell_{k^2_\ell} \\ 0 & \text{otherwise.} \end{cases}$$

For boundary grid point $x_{m_2} = (r_0, p_{j_2}, y^1_{k^2_1}, \ldots, y^\ell_{k^2_\ell})$, $j_2 \in \mathcal{N}_p$, $k^2_i \in \mathcal{N}_{y^i}$, $i = 1, \ldots, \ell$, the above transition probability becomes

$$\begin{aligned}
\mathbf{P}^{\nu}_{x_{m_1}, x_{m_2}} &= \mathbb{P}^{\nu}(R_{n+1} \in \mathcal{B}_{r_0} \mid R_n = r_{i_1}) \times \mathbb{P}^{\nu}(P_{n+1} \in \mathcal{B}_{p_{j_2}} \mid X^{\mathbf{u}}_n = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^\ell_{k^1_\ell}),\ u_n = \nu) \\
&\quad \times \Pi^\ell_{i=1} \mathbb{P}^{\nu}(\widetilde{Y}^i_{n+1} \in \mathcal{B} y^i_{k^2_i} \mid \widetilde{Y}^i_n = y^i_{k^1_i}, u_n = \nu) \\
&= \begin{cases} \Phi\big(\frac{1}{2}(r_0 + r_1)\big) & \text{if} \quad P_{n+1} \in \mathcal{B}_{p_{j_2}},\ \widetilde{Y}^1_{n+1} \in \mathcal{B} y^1_{k^2_1}, \ldots, \widetilde{Y}^\ell_{n+1} \in \mathcal{B} y^\ell_{k^2_\ell} \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

For boundary grid point $x_{m_2} = (r_{N_r}, p_{j_2}, y^1_{k^2_1}, \ldots, y^\ell_{k^2_\ell})$, the above transition probability becomes

$$\begin{aligned}
\mathbf{P}^{\nu}_{x_{m_1}, x_{m_2}} &= \mathbb{P}^{\nu}(R_{n+1} \in \mathcal{B}_{r_{N_r}} \mid R_n = r_{i_1}) \times \mathbb{P}^{\nu}(P_{n+1} \in \mathcal{B}_{p_{j_2}} \mid X^{\mathbf{u}}_n = (r_{i_1}, p_{j_1}, y^1_{k^1_1}, \ldots, y^\ell_{k^1_\ell}),\ u_n = \nu) \\
&\quad \times \Pi^\ell_{i=1} \mathbb{P}^{\nu}(\widetilde{Y}^i_{n+1} \in \mathcal{B} y^i_{k^2_i} \mid \widetilde{Y}^i_n = y^i_{k^1_i}, u_n = \nu) \\
&= \begin{cases} 1 - \Phi\big(\frac{1}{2}(r_{N_r} + r_{N_r - 1})\big) & \text{if} \quad P_{n+1} \in \mathcal{B}_{p_{j_2}},\ \widetilde{Y}^1_{n+1} \in \mathcal{B} y^1_{k^2_1}, \ldots, \widetilde{Y}^\ell_{n+1} \in \mathcal{B} y^\ell_{k^2_\ell} \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

In the next section we present the numerical results of the stochastic optimal control problem.

6.3 Numerical Results

This section is devoted to the numerical experiments of the optimal control problem. The intentions of the numerical experiments are to first explore the capabilities but also the limits of the solution of MDP described above by backward recursion. We consider the state space of small dimension but already meaningful for practical purposes and the curse of dimensionality is still under control. Second, to gain inside to properties of the value function and the optimal decision rules. Such properties are helpful for approximate solutions based on ansatz functions using least squares Monte Carlo or approximate dynamic programming.

The control problem is obtained by replacing the state Q of the GS by a low-dimensional state $\widetilde{Y}$ of the reduced order system. This reduced order state results from applying model order reduction (presented in Chapter 4) to a system of ODEs (3.10) resulting from semi-discretization of the heat equation (2.6) which models the spatio-temporal temperature distribution of a GS presented in Chapter 7.
For describing the input-output behavior of that storage we use the aggregated characteristics of the spatial temperature distribution introduced in Sec. 2.3.3. Further we work with the approximation of (3.10) by an analogous system as explained in Sec. 3.4 and the output matrix is $C = (C^M, C^F)^\top$. This choice is motivated by the fact that the control constraint related to the reduced-order state $\widetilde{Y}$ is restricted to the average temperature in the storage medium of the GS but the average temperature in the fluid is needed for the analogous model. Recall, in the analogous model presented in Sec 3.4 it is assumed that the pump is always on and during the waiting periods the inlet temperature Q^I is set to be the average temperature $\overline{Q}^F = C^F \widetilde{Y}$ in the PHX fluid. With this choice of the output matrix, the numerical results presented in Subsec. 4.3.2 show that with 3 states, the reduced-order system can capture almost 90% of the output energy of the analogous high-dimensional system. This motivates us to choose for all numerical experiments the dimension of the reduced-order state $\ell = 3$. Thus, the discretized state space is then given by $\widetilde{\mathcal{X}} = \widetilde{\mathcal{R}} \times \widetilde{\mathcal{P}} \times \widetilde{\mathcal{Y}}^1 \times \widetilde{\mathcal{Y}}^2 \times \widetilde{\mathcal{Y}}^3$.

Numerical experiments are based on the backward recursion in Algorithm 2 and are performed for the cases of one PHX with diameter 2 *cm* and take into account several waiting periods. Numerical results presented in Chapter 4 showed that the model with one PHX can capture well the output energy of the high-dimensional system with only 3 states. The numerical experiments are based on the time and state discretization of the resulting optimal control problem presented in Sec. 6. The simulations aim to determine the value function and the optimal charging and discharging decisions of the storage's manager at any discrete time points and at any grid point in the discretized state-space $\widetilde{\mathcal{X}}$. For these experiments we also compute and plot some optimally controlled paths of the temperature in the internal and GS.

After explaining the experimental settings in Subsec. 6.3.1 we start in Subsec. 6.3.2 with the experiments where we compute and plot the value function and the optimal strategy a time $t = 0, T-2, T-1, T$, and study the sensitivity analysis with respect to the individual state variable. We end this chapter by showing some optimally controlled paths of the temperatures in both storages, presented in Subsec. 6.3.3.

6.3.1 Experimental Setting

For the numerical results presented below we use for the GS the discretization parameters given in Table 3.2 and for the solution of the discrete-time optimal control problem via backward recursion we use the parameters given in Table 6.1. Thermal energy is stored by raising the

temperature of the storage medium. The fluid is assumed to be water while the storage medium is dry soil. During charging a pump moves the fluid with constant velocity $\overline{v}_0$ arriving with constant temperature $Q^I(t) = Q^I_C = 40$ °C at the inlet. This temperature is higher than in the vicinity of the PHXs, thus induces a heat flux into the storage medium. During discharging the inlet temperature is $Q^I(t) = Q^I_D = 5$ °C which corresponds to the temperature of the fluid returning from the heat pump and leading to a cooling of the GS. At the outlet we impose a vanishing diffusive heat flux, i.e. during pumping there is only a convective heat flux. We also consider waiting periods where the pump should be off but in that analogous model we consider here the pump is always on and we choose the inlet temperature of the PHX to be the average temperature of the pipe ($\overline{Q}^F = C^F\widetilde{Y}$). We assume a fixed and constant fuel price or $F = 1.6$ EUR/kWh and we assume that when the fuel-fired boiler is on, it produces $k_F = 92.97$ kW of energy in one hour. This corresponds to a fuel cost of $\psi_F = k_F F = 150$ EUR/h. This raises the temperature in the IS by 40 °C in one hour. When we discharge the IS to charge the GS we use the classical pump which costs $\psi_D = 5$ EUR/h. However, to charge the IS by discharging the GS we use the heat pump with the inlet temperature $\overline{Q}^O(t) = C^O\widetilde{Y}(t)$ to raise the temperature to $P_{in} = 35$ °C. This operation costs $\psi_C(t,x) = 3.157(P_{in} - C^O y) + 5$ EUR/h. We also consider the terminal cost including the penalty if the internal and the GSs are not properly filled at the terminal time $T = 72$ h given for $X(T) = x = (r,p,y)$, $y = (y_1,y_2,y_3)$ by

$$\Phi(x) = 20(q_{pen} - C^M y)^+ + 30(p_{pen} - p)^+ \quad EUR.$$

We choose the finite action set $\overline{\mathcal{U}} = \{u^O, u^D, u^W, u^C, u^F\} = \{-2,-1,0,+1,+2\}$, where $u^F = +2$ stands for charging the IS by firing fuel at maximum rate (classical pumps and heat pump off), $u^C = +1$ for charging the IS by discharging the GS at maximum rate (classical pump and fuel-fired boiler off), $u^W = 0$ for wait or do nothing (pumps and fuel-fired boiler off), $u^D = -1$ for charging the GS by discharging the IS at maximum rate (heat pump and fuel-fired boiler off) and $u^O = -2$ when none of the above control can keep the temperature in the IS below the maximum. We assume that charging the IS by discharging the GS is such that within one hour the temperature in the IS increases by 20 °C and charging the GS by discharging the IS is such that in one hour the temperature in the IS decreases by 20 °C. We choose the minimum and maximum temperature in the IS $\underline{p} = 30$ °C and $\overline{p} = 90$ °C, respectively. We choose the minimum and maximum temperature in the GS $\underline{q} = 10$ °C and $\overline{q} = 30$ °C, respectively. We choose using the 3-sigma rule the minimum and the maximum residual demand $\underline{r} = -16.7 \times 10^7$ $J/h = -46.38$ kW and $\overline{r} = 13.4 \times 10^7$ $J/h = 37.22$ kW, respectively. For the reduced-order system we choose $\ell = 3$ and choose the bounds $[\underline{y}_1, \overline{y}_1] = [4500, 13750]$, $[\underline{y}_2, \overline{y}_2] = [200, 800]$, and $[\underline{y}_3, \overline{y}_3] = [4084.5, 12255]$ for the reduced-order states $\widetilde{Y}^1$, $\widetilde{Y}^2$ and $\widetilde{Y}^3$, respectively. We assume that the mass of the IS is $m^P = 2000$ kg. Then we have $\overline{p} - \underline{p} = 60$ K and the maximum amount of energy that can be stored in the IS is $m^P c^F_P(\overline{p} - \underline{p}) = 2000 \times 4182 \times 60 = 501.84$ $MJ \simeq 139.4$ kWh. We also assume the volume of the GS without the pipes is $V^Q = 100$ m^3 and its mass is $m^Q = \rho^M V^Q = 2000 \times 100 = 2 \times 10^5$ kg. Then we have $\overline{q} - \underline{q} = 20$ K and the maximum amount of energy that can be stored in the GS is $m^Q C^M(\overline{q} - \underline{q}) = 200000 \times 800 \times 20 = 3200$ $MJ \simeq 888.88$ kWh$\simeq 0.89$ MWh. For simplicity, in all figures write the value of the residual demand given in joule per hour (J/h) without the factor 10^7 and we label the axis of the residual demand without this factor but we keep in mind that this factor multiplies all values of the residual demand in this section. We choose $P_{out} = \underline{p}$.

Parameters		Values	Units
Discretization			
dimension of the reduced order state $\widetilde{Y}$	ℓ	3	
time horizon	T	72	h
time step	Δ_N	1	h
number of grid points in r- direction	N_r	8	
number of grid points in p- direction	N_p	11	
number of grid points in y^1- direction	N_{y_1}	5	
number of grid points in y^2- direction	N_{y_2}	5	
number of grid points in y^3- direction	N_{y_3}	11	
Material			
diffusion coefficient of R	σ_R	13.95	$kW/\sqrt{h}$
drift coefficient of R	μ_R	-4.64	kW
mean reversion speed of R	β_R	0.5	$1/h$
total surface area of the IS	A_h	9.096	m^2
overall heat transfer coefficient	κ_h	12	W/m^2K
specific heat capacity of the IS medium	c_P^F	4184	J/kgK
specific heat capacity of the GS medium	c_P^M	800	J/kgK
mass of the IS	m^P	2000	kg
mass of the GS	m^Q	200000	kg
rate of heat loss to the environment	$\gamma = \frac{\kappa_h A_h}{m^P c_P^F}$	0.0118	$1/h$
penalty threshold temperature of the IS at time T	P_{pen}	60	°C
penalty threshold temperature of the GS at time T	q_{pen}	20	°C
ambient temperature around the IS	P_{amb}	20	°C
maximum temperature in the GS	$\overline{q}$	30	°C
minimum temperature in the GS	$\underline{q}$	10	°C
maximum temperature in the IS	$\overline{p}$	90	°C
minimum temperature in the IS	$\underline{p}$	30	°C
rate of energy produced by firing fuel	k_F	92.97	kW
increase in the temperature in the IS by firing fuel	κ_F	40	K/h
Constants			
energy conversion rate	$\kappa_P = \frac{1}{m^P c_P^F}$	0.4302	K/kWh
rate of energy produced by discharging the GS	k_C	1.67	kW/K
rate of energy loss by discharging the IS	k_D	1.38	kW/K
related to the inflow of energy to the IS	$\kappa_C = k_C \kappa_P$	0.7184	$1/h$
outflow of energy to the GS	$\kappa_D = k_D \kappa_P$	0.5937	$1/h$
conversion factor	$\zeta_{pen}^P m^P c_P^F$	30	EUR/K
conversion factor	$\zeta_{pen}^Q m^Q c_p^M$	20	EUR/K
efficiency of the heat pump	η^Q	95%	
tolerance	ε	5%	

Table 6.1: Constants and Material parameters for MDP

6.3.2 Optimal Strategy and Value Function

Terminal value function

The left panel of Fig. 6.9 gives the value function at terminal time $T = 72$ for an empty and full GS, and the right panel of shows the terminal value function for a GS at the penalty threshold ($\overline{Q}^M \geq q_{pen}$). This figure shows that when the average temperature in the IS is above

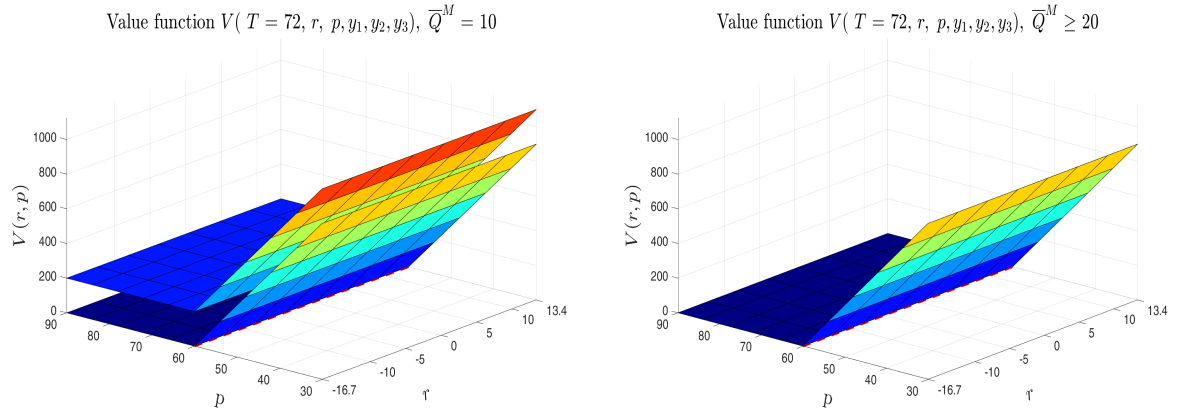


Figure 6.9: Terminal value function. Left: empty (upper graph) and full (lower graph) GS. Right: GS above the penalty threshold level ($\overline{Q}^M \geq q_{pen} = 20$ °C).

the threshold ($P(T) \geq p_{pen} = 60$ °C), the terminal value function is constant (positive for $\overline{Q}^M = C^M y < q_{pen}$ and zero for $\overline{Q}^M \geq q_{pen}$) and when the average temperature in the IS is below the threshold ($P(T) <= p_{pen}$) the terminal value function increases as the average temperature in the IS moves away from the threshold p_{pen}.

Value function and optimal strategy at time $t = T - 1 = 71$ h

In this this paragraph we study the behaviour of the value function and the optimal strategy with respect to individual state variables. We begin with the case of an empty GS.

Geothermal storage is empty. Fig. 6.10 gives the value function and optimal strategy at time $t = 71$ h for an empty GS ($C^M y = \underline{q} = 10$ °C) as a function of (r, p). Fig. 6.11 shows a comparison of value functions (left) and optimal strategies (right) at time $t = 71$ h as a function of p for three fixed values of the residual demand. Fig. 6.12 plots a comparison of value functions (left) and optimal strategies (right) as a function of the residual demand r for three fixed values of the average temperature in the IS. The left panel of Fig. 6.11 shows that the value function decreases as the average temperature in the IS increases whereas the left panel of Fig. 6.12 shows that the value function increases as the residual demand increases. The right panels of Fig. 6.11 shows that when the IS and the GS are empty one hour before the terminal time, we must fire fuel (even if there is overproduction) to increase the temperature in order to avoid high penalty at the terminal time. However, if the temperature in the IS is above certain level ($p > 40$ °C), we are required to fire fuel only in case of unsatisfied demand and when $p > 70$ °C it is sufficient to wait and only discharge the IS in case of overproduction. Similarly, the right panel of Fig. 6.12 shows that we must fire fuel as long the IS remain empty (including the case of overproduction) and when the IS is full we must wait and only discharge it when there is overproduction. This justifies the large gap in the value functions (see the left panel of Fig. 6.12) when the state of

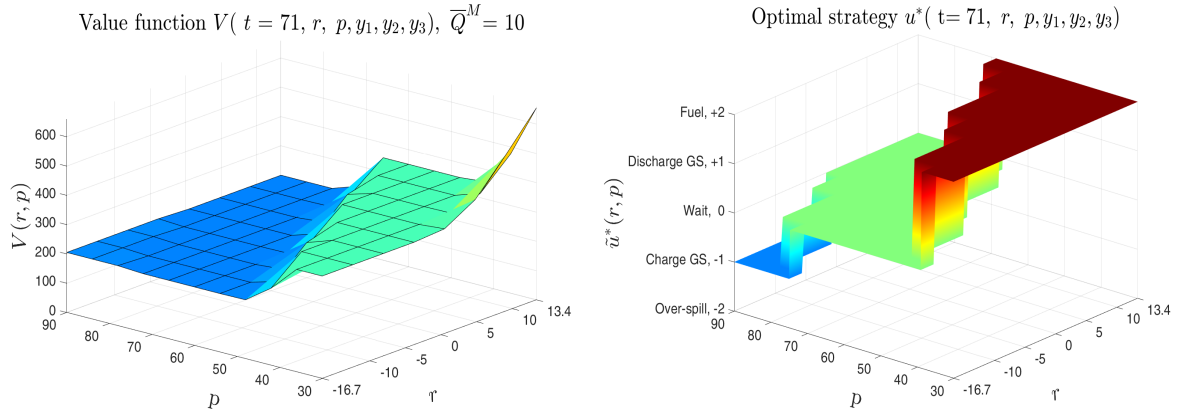


Figure 6.10: Value function (left) and optimal strategy (right) at time $t = 71\ h$ as a function of (r, p) for an empty GS ($\overline{Q}^M = \underline{q} = 10\ °\text{C}$).

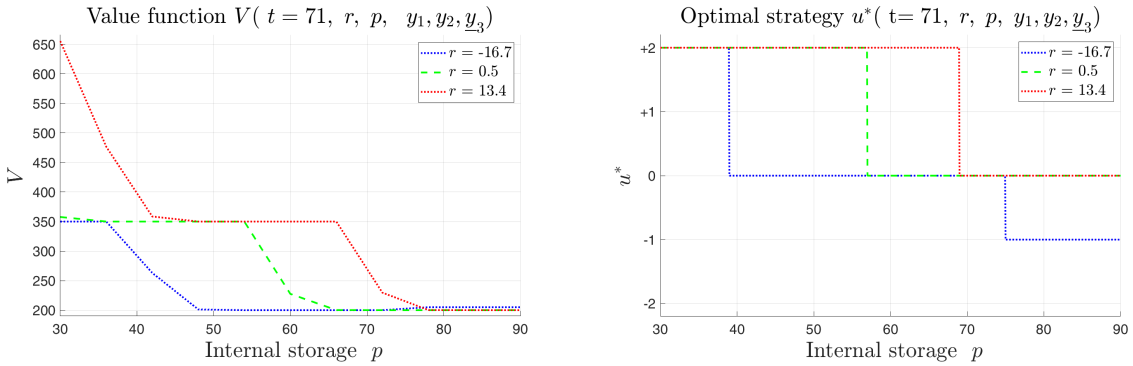


Figure 6.11: Comparison of value functions (left) and optimal strategies (right) at time $t = 71\ h$ as a function of the average temperature in the IS p for different values of the residual demand and an empty GS.
Blue dotted line for strong overproduction ($r = \underline{r} = -16.7$), green dashed line for very small unsatisfied demand ($r = 0.5$) and red dotted line for strong unsatisfied demand ($r = \overline{r} = 13.4$).

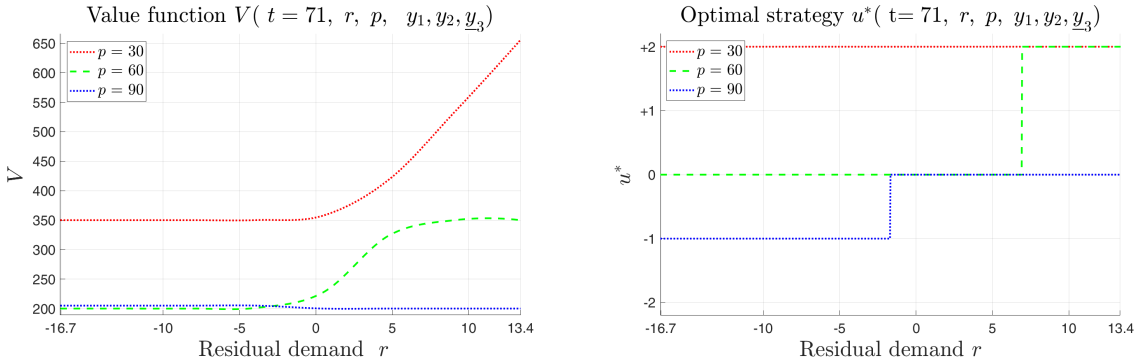


Figure 6.12: Comparison of value functions (left) and optimal strategies (right) at time $t = 71\ h$ as a function of the residual demand r for different values of the temperature in the IS and an empty GS.
Blue dotted line for IS full ($p = \overline{p} = 90\ °\text{C}$), green dashed line for IS at the penalty threshold ($p = p_{pen} = 60\ °\text{C}$), and red dotted line for IS empty ($p = \underline{p} = 30\ °\text{C}$).

the IS changes from $p = 30\ °\text{C}$ to $p = 90\ °\text{C}$. The right panel shows that when the temperature in the IS is at the penalty threshold level ($p = p_{pen} = 60\ °\text{C}$), it is optimal to wait and only fire fuel in case of strong unsatisfied demand. This justifies the increase in the value function as r

approaches the maximum value $\overline{r}$.

Geothermal storage at intermediate filling level. Fig. 6.13 shows the value function and optimal strategy at time $t = 71\ h$ when the GS is at the penalty threshold level ($\overline{Q}^M = p_{pen} = 20\ ^\circ$C) as a function of (r, p). The right panel of this figure shows that when the average temperature in the IS p is above the penalty threshold level ($p > p_{pen} = 60\ ^\circ$C) it is optimal to wait and only discharge the IS to charge the GS if the latter is full and there is strong overproduction. However, when the average temperature in the IS p is below the penalty threshold level ($p < p_{pen} = 60\ ^\circ$C) we have to discharge the GS to charge the IS and only fire fuel when the IS is empty and there is strong unsatisfied demand. This justifies the fact that, when there is overproduction, the value function increases as the average temperature in the IS P approaches the minimum level ($p = \underline{p} = 30\ ^\circ$C) and increases slowly as p approaches the maximum level ($p = \overline{p} = 90\ ^\circ$C). Further, when there is unsatisfied demand the value function strongly increases as the p decreases. Similarly, the value function increases strongly as the residual demand increases when the IS is empty. It increases slowly as the residual demand decreases when the IS is full.

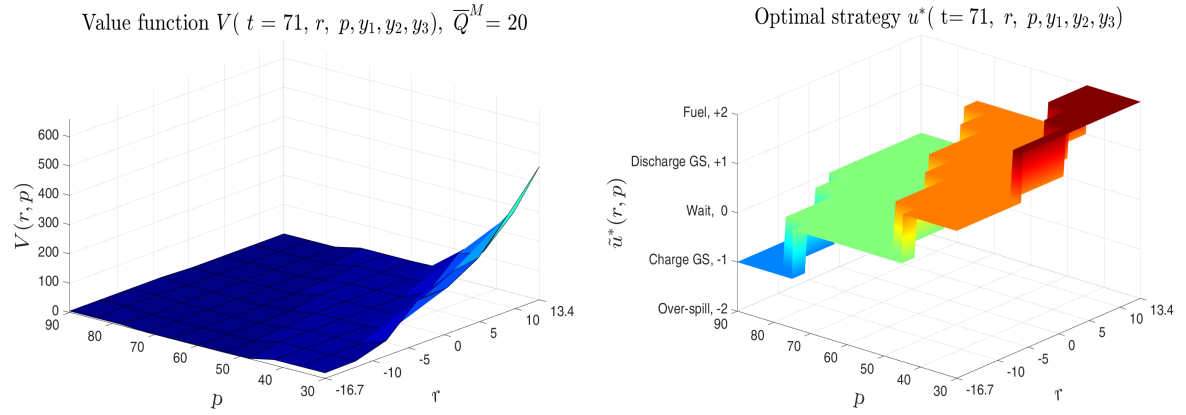


Figure 6.13: Value function (left) and optimal strategy (right) at time $t = 71\ h$ as a function of (r, p) for GS at the penalty threshold level ($\overline{Q}^M = q_{pen} = 20\ ^\circ$C).

Fig. 6.14 shows comparison of the value function and optimal strategy at time $t = 71\ h$ as a function the average temperature of the GS $\overline{Q}^M$ for different values of the temperature in the IS. Further, in Fig. 6.14, we consider two cases: strong overproduction in the top panels and strong unsatisfied demand in the bottom panels. Note that in the transformed coordinates system the average temperature in the GS is proportional to the last reduced order state of the GS y_3 and the value function is varies slowly with respect to the first two coordinates. Therefore, we are going to study the value function with respect to the last reduced order states y_3 or with respect to the average temperature in the GS.
The left panels of Fig. 6.14 show that at time $t = 71\ h$, the value function decreases as the average temperature in the GS increases and it becomes constant when the average temperature in the GS exceeds the penalty threshold level ($\overline{Q}^M = q_{pen} = 20\ ^\circ$C). The right panels of Fig. 6.14 show that when the IS is full it is optimal to wait when there is strong unsatisfied demand (bottom panel). In case of strong overproduction we charge the GS by discharging the IS (as long as the geothermal is not full). Contrary, when the IS is empty we must stop firing fuel as soon as the GS in no longer empty and there is strong overproduction but in case of strong unsatisfied demand we must fire fuel as long as the IS is empty, no matter the state of the GS. This justifies the large gap in the value function when the state of the IS changes from $p = 30\ ^\circ$C

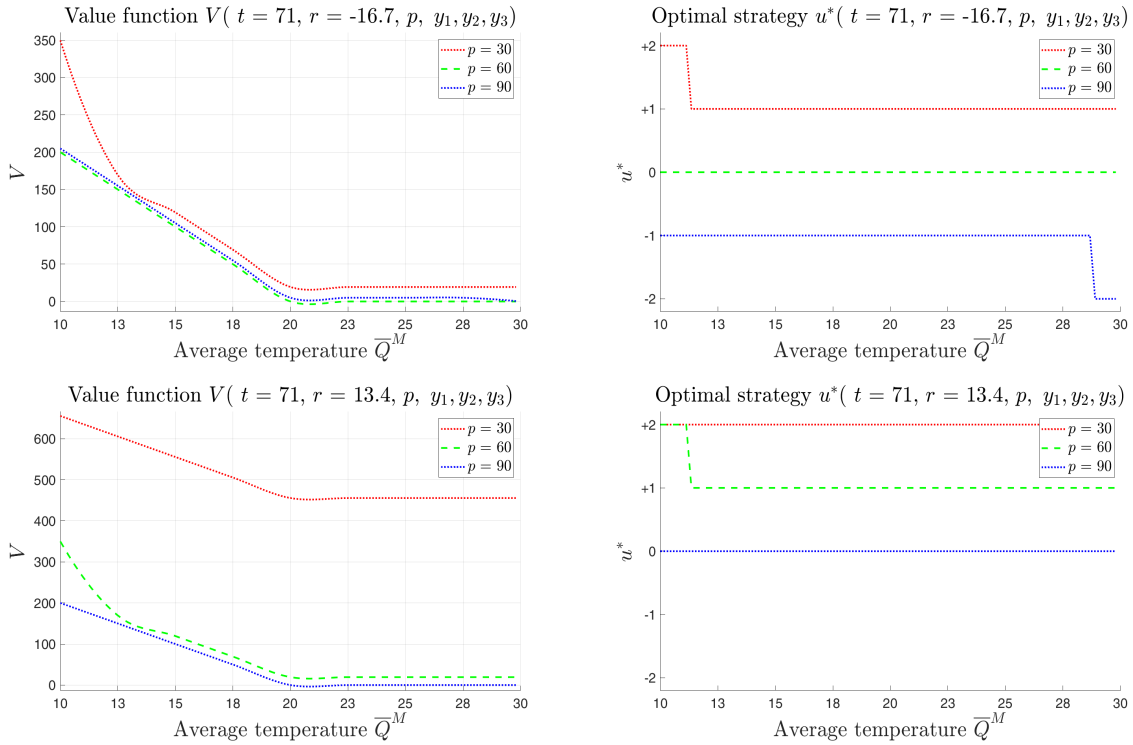


Figure 6.14: Comparison of value function (left) and optimal strategy (right) at time $t = 71$ h as a function of the average temperature in the GS $\overline{Q}^M$ for a strong overproduction (top panels) and strong unsatisfied demand (bottom panels) and different values of the temperature in the IS.
Blue dotted line for IS full ($p = \overline{p} = 90$ °C), green dashed line for IS at the penalty threshold ($p = p_{pen} = 60$ °C), and red dotted line for IS empty ($p = \underline{p} = 30$ °C).

to $p = 90$ °C, see the dotted red and blue lines in the left panel of Fig. 6.14. The right panel shows that when the temperature in the IS is at the penalty threshold level ($p = p_{pen} = 60$ °C), it is optimal to wait and only fire fuel if there is strong unsatisfied demand and the GS is empty. This justifies the strong decay in the value function as $\overline{Q}^M$ approaches the minimum value $\underline{q} = 10$ °C.

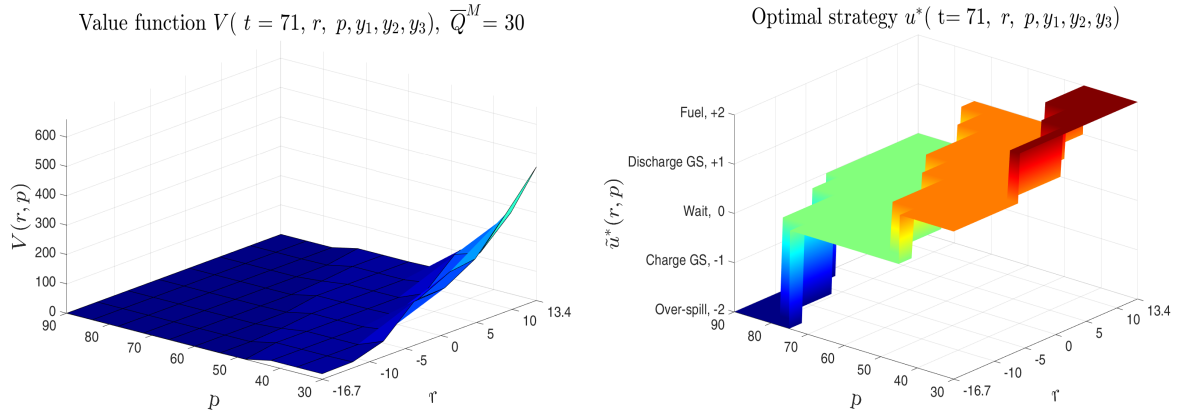


Figure 6.15: Value function (left) and optimal strategy (right) at time $t = 71$ h as a function of (r, p) for a full GS ($\overline{Q}^M = \overline{q} = 30$ °C).

Geothermal storage is full. Fig. 6.15 depicts the value function and optimal strategy at time

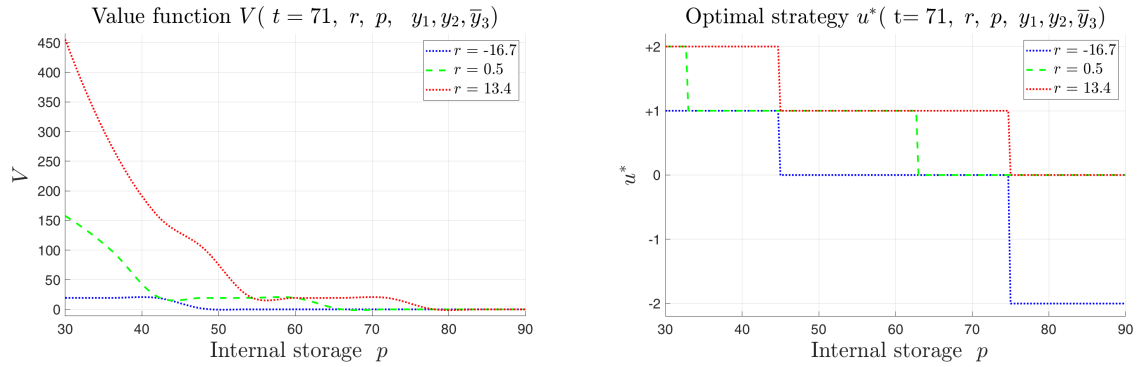


Figure 6.16: Comparison of value function (left) and optimal strategy (right) at time $t = 71\ h$ as a function of p for different values of the residual demand and a full GS ($\overline{Q}^M = \overline{q} = 30\ °\text{C}$).
Blue dotted line for strong overproduction ($r = \underline{r} = -16.7$), green dashed line for very small unsatisfied demand ($r = 0.5$) and red dotted line for strong unsatisfied demand ($r = \overline{r} = 13.4$).

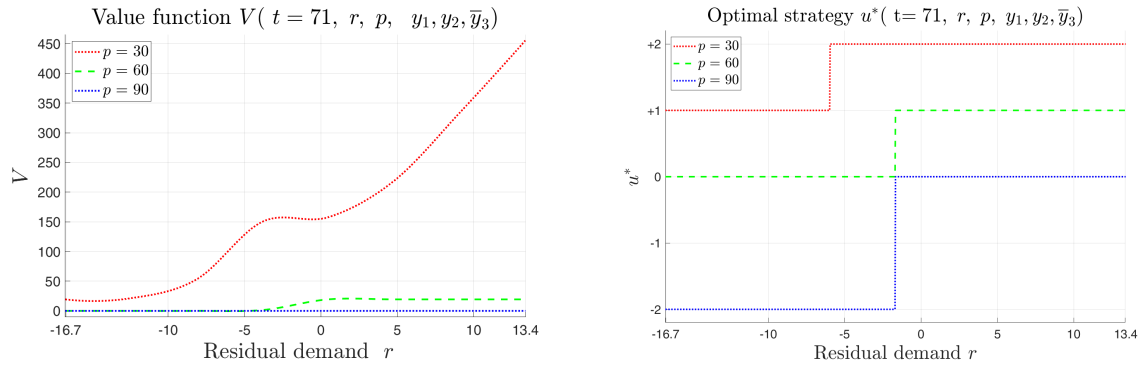


Figure 6.17: Comparison of value functions (left) and optimal strategies (right) at time $t = 71\ h$ as a function of the residual demand r for different values of the average temperature in the IS and a full GS ($\overline{Q}^M = \overline{q} = 30\ °\text{C}$).
Blue dotted line for IS full ($p = \overline{p} = 90\ °\text{C}$), green dashed line for IS at the penalty threshold ($p = p_{pen} = 60\ °\text{C}$), and red dotted line for IS empty ($p = \underline{p} = 30\ °\text{C}$).

$t = 71\ h$ for a full GS ($\overline{Q}^M = \underline{q} = 30\ °\text{C}$) as a function of (r, p). Fig. 6.16 plots a comparison of value functions (left) and optimal strategies (right) at time $t = 71\ h$ as a function of p for three fixed values of the residual demand and Fig. 6.17 shows a comparison of value functions (left) and optimal strategies (right) at time $t = 71\ h$ as a function of the residual demand r for three fixed values of the average temperature in the IS. The left panel of Fig. 6.16 shows that the value function decreases as the average temperature in the IS increases and converges to zero as p approaches $\overline{p}$. The left panels of Fig. 6.15 and Fig. 6.17 show that the value function is constant zero when the temperature in the IS exceeds the penalty threshold level ($p > 60\ °\text{C}$) and increases as the residual demand increases and the temperature in the IS is below the penalty threshold level ($p \leq 60\ °\text{C}$). The right panel of Fig. 6.16 shows that when the GS is full one hour before the terminal time, we only fire fuel or discharge the GS if the temperature in the IS is below the penalty threshold level ($p \leq 60\ °\text{C}$) and if $p > 60\ °\text{C}$ we wait when there is a small unsatisfied demand or apply spill-over when there is overproduction. Similarly, the right panels of Fig. 6.15 and Fig. 6.17 show that when the temperature in the IS is above the penalty threshold level ($p > 60\ °\text{C}$) we have to wait or apply spill-over (this justifies the fact that the value function is zero) and only discharge the GS when $p = 60\ °\text{C}$ and we have unsatisfied demand (this justifies the small increase in the value function when $p = 60\ °\text{C}$). The latter shows that we must fire fuel as long the IS is empty and there is unsatisfied demand, this

justifies the strong increase in the value function as when the the IS is empty.

Value function and optimal strategy at time $t = T - 2 = 70\ h$

In this paragraph we study the behaviour of the value function and the optimal strategy 2 hours before the terminal time as the average temperature in the GS increases from $\underline{q} = 10\ \ ^\circ\text{C}$ to $\overline{q} = 30\ \ ^\circ\text{C}$.

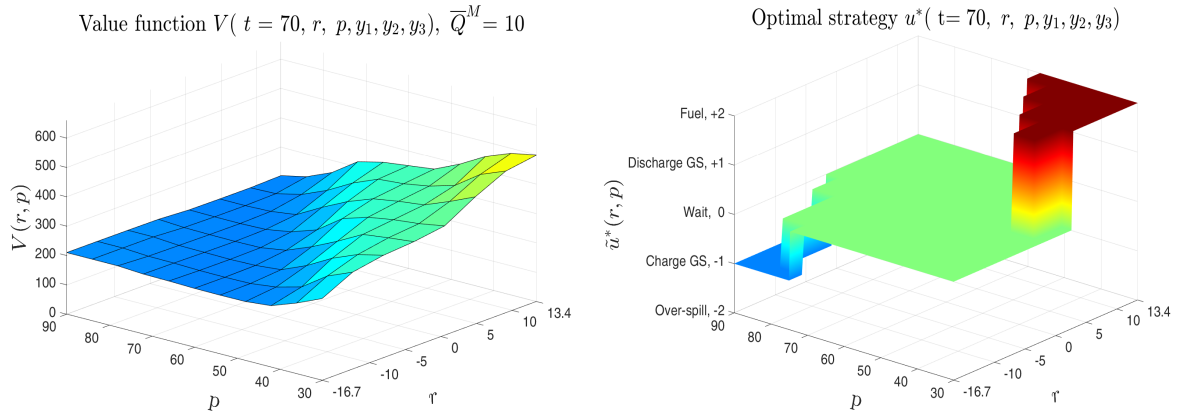


Figure 6.18: Value function (left) and optimal strategy (right) at time $t = 70\ h$ as a function of (r, p) for an empty GS ($\overline{Q}^M = \underline{q} = 10\ \ ^\circ\text{C}$).

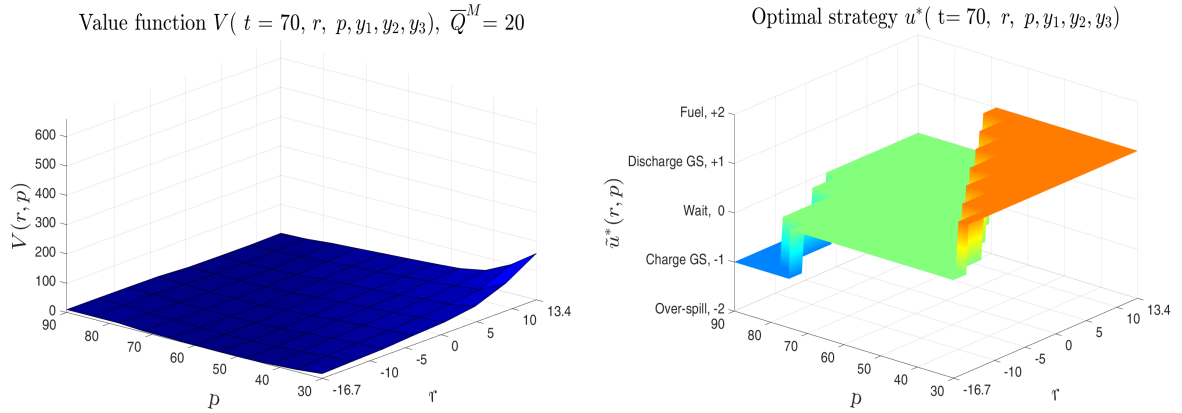


Figure 6.19: Value function (left) and optimal strategy (right) at time $t = 70\ h$ as a function of (r, p) for a GS at the penalty threshold level ($\overline{Q}^M = q_{pen} = 20\ \ ^\circ\text{C}$).

In particular, we consider three fixed values of the average temperature in the GS. Fig. 6.18,6.19, and 6.20 show the value function (left) and optimal strategy (right) at time $t = 70\ h$ for an empty GS ($\overline{Q}^M = \underline{q} = 10\ \ ^\circ\text{C}$), the storage at the penalty threshold level ($\overline{Q}^M = q_{pen} = 20\ \ ^\circ\text{C}$) and a full storage ($\overline{Q}^M = \overline{q} = 30\ \ ^\circ\text{C}$) as a function of (r, p), respectively. The left panels of these figures show that the value function decreases as the average temperature in the internal or GS increases and it increases as the residual demand increases. The right panels of these figures show that we only fire fuel if the GS is empty and there is strong unsatisfied demand but we stop firing fuel as soon as the GS is no longer empty.

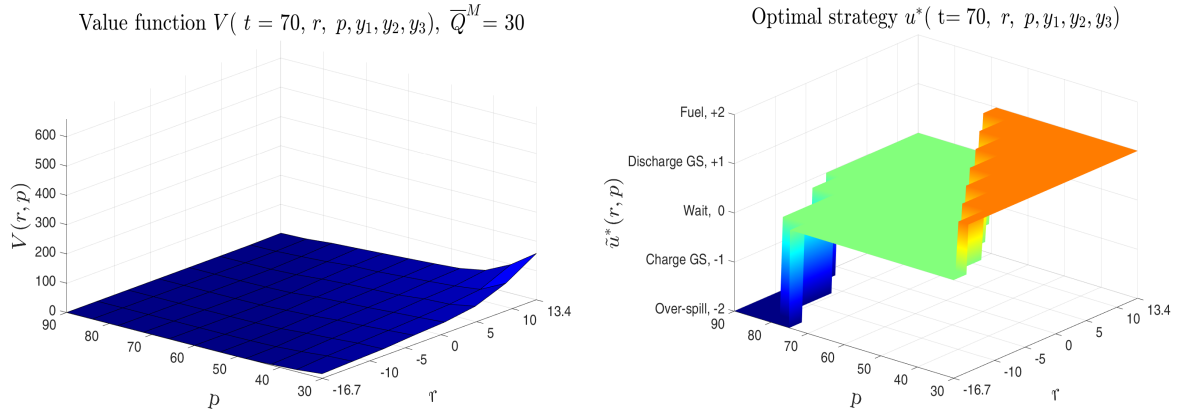


Figure 6.20: Value function (left) and optimal strategy (right) at time $t = 70\ h$ as a function of (r, p) for a full GS ($Qm = \overline{q} = 30\ \ ^\circ$C).

Value function and optimal strategy at initial time $t = 0$

In this paragraph we study the behaviour of the value function and the optimal charging or discharging decision of the storage at the initial time with respect to individual state variables. Further, we compare the results with $t = 70\ h$ and $t = 71\ h$. As in the case of times $t = 71\ h$ and $t = 70\ h$ we begin with case of an empty GS ($\overline{Q}^M = \underline{q} = 10\ \ ^\circ$C).

Geothermal storage is empty. Fig. 6.21 shows the value function and optimal strategy at the initial time $t = 0\ h$ for an empty GS as a function of (r, p). Fig. 6.22 shows a comparison of value functions and optimal strategies at time $t = 0\ h$ as a function of p for three fixed values of the residual demand. Fig. 6.23 shows a comparison of value functions and optimal strategies at time $t = 0\ h$ as a function of the residual demand r for three fixed values of the temperature in the IS. The left panel of Fig. 6.22 shows that the value functions strongly decrease as the average temperature in the IS increases and there is unsatisfied demand (green dashed and red dotted lines). Further, we observe a slow increase in the value function when there is overproduction and the average temperature approaches the maximum (blue dotted line). This observation is

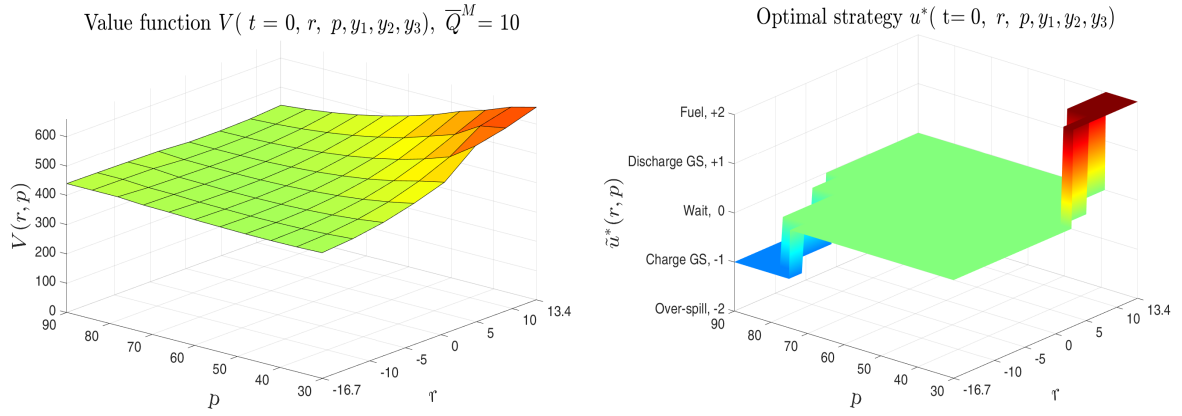


Figure 6.21: Value function (left) and optimal strategy (right) at the initial time $t = 0\ h$ as a function of (r, p) for an empty GS ($\overline{Q}^M = \underline{q} = 10\ \ ^\circ$C).

due to the fact that we fire fuel when there is a strong unsatisfied demand and the IS is empty,

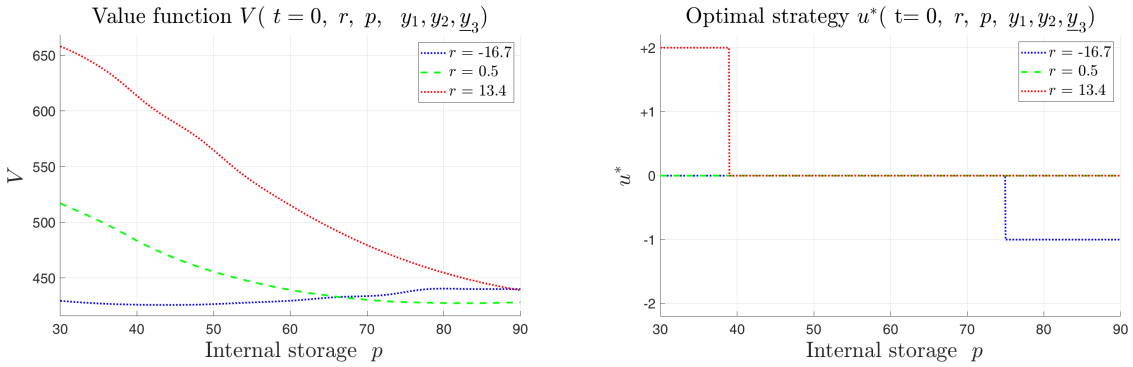


Figure 6.22: Comparison of value functions (left) and optimal strategies (right) at time $t = 0$ h as a function of p for three fixed values of the residual demand and an empty GS ($\overline{Q}^M = \underline{q} = 10$ °C).
Blue dotted line for strong overproduction ($r = \underline{r} = -16.7$), green dashed line for very small unsatisfied demand ($r = 0.5$) and red dotted line for strong unsatisfied demand ($r = \overline{r} = 13.4$).

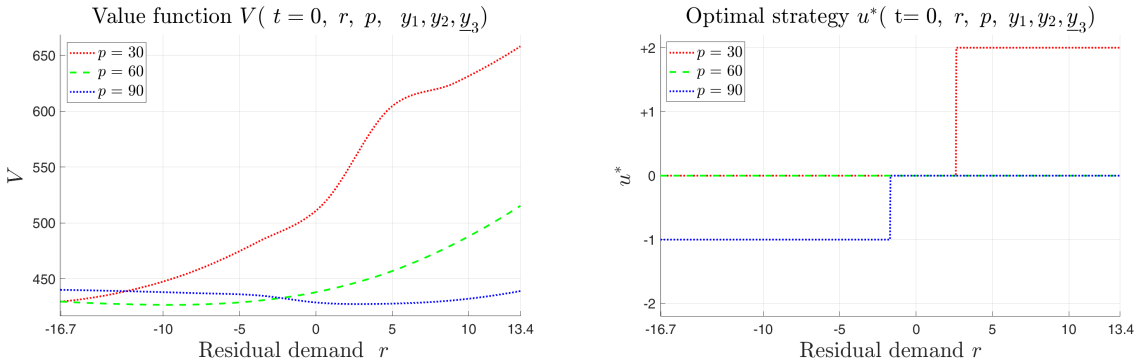


Figure 6.23: Comparison of value functions (left) and optimal strategies (right) at time $t = 0$ h as a function of the residual demand r for three fixed values of the average temperature in the IS and an empty GS.
Blue dotted line for IS full ($p = \overline{p} = 90$ °C), green dashed line for IS at the penalty threshold ($p = p_{pen} = 60$ °C), and red dotted line for IS empty ($p = \underline{p} = 30$ °C).

and we discharge the internal to charge the GS when there is overproduction and the IS is full, see the right panel of Fig. 6.22.

The left panel of Fig. 6.23 shows that when the residual demand increases, the value function strongly increases if the IS is empty and slowly increases when the average temperature in the IS is at the penalty threshold level ($p = p_{pen} = 60$ °C). This is due the fact that we have to fire fuel when the IS is empty and there is unsatisfied demand, and when the average temperature in the IS is at the penalty threshold level ($p = p_{pen} = 60$ °C) we only have to wait, see the green dashed and dotted red lines in right panel of Fig. 6.23. However, the right panel of Fig. 6.23 shows that when the IS is full, we discharge the IS to charge the GS if there is overproduction and we wait when there is unsatisfied demand. This justifies the slow decrease of the value function when there is overproduction and slow increase in the value function when there is unsatisfied demand, see the blue dotted line in the left panel of Fig. 6.23. We observe that the value function is much larger for $t = 0$ h than $t = 70$ h or $t = 71$ h, even for $p > 60$ °C. This can be justified by the fact that from $t = 70$ h there are two periods and from $t = 71$ h there is only one period ahead to the terminal time but from $t = 0$ h there are many periods ahead until the terminal time T.

Geothermal storage at intermediate filling level. Fig. 6.24 shows the value function and

optimal strategy at time $t = 0\ h$ when GS is at the penalty threshold level ($\overline{Q}^M = p_{pen} = 20\ ^\circ$C) as a function of (r, p). The result depicted in the right panel of this figure shows that when the IS is full and there is strong overproduction, we have to discharge the IS to charge the GS and when there is unsatisfied demand we only have to wait. However, when the IS is empty we have to wait and only charge it by discharging the GS when there is strong unsatisfied demand. This justifies the fact that, when there is overproduction, value function remains constant as long as the average temperature in the IS p is below the penalty threshold level, $p < p_{pen} = 60$ and only increases slowly as p approaches the maximum level, $p = \overline{p} = 90\ ^\circ$C. Further, when there is unsatisfied demand it remains constant as long as p is above the penalty threshold level, $p > p_{pen} = 60\ ^\circ$C. It only increases slowly as p approaches the minimum level, $p = \underline{p} = 30\ ^\circ$C. Similarly, the value function increases as the residual demand increases, when the IS is empty and decreases as the residual demand increases, when the IS is full, and remains constant otherwise.

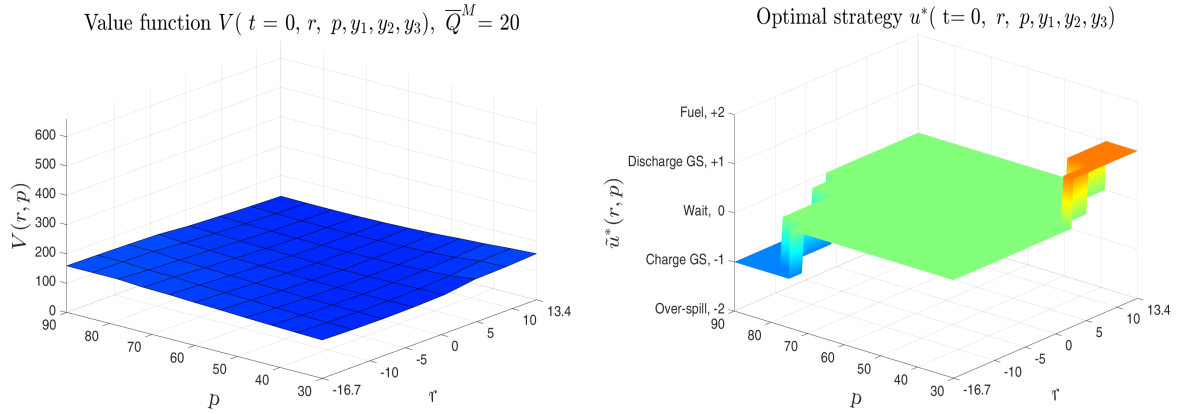


Figure 6.24: Value function (left) and optimal strategy (right) at the initial time $t = 0\ h$ as a function of (r, p) for an average temperature in the GS at the threshold ($\overline{Q}^M = q_{pen} = 20\ ^\circ$C).

Fig. 6.25 shows a comparison of the value functions (left) and the optimal strategies (right) at time $t = 0\ h$ as a function the average temperature in the GS $\overline{Q}^M$ for three fixed values of the temperature in the IS. As in Figure 6.14, here we consider the case of strong overproduction in the top panels and strong unsatisfied demand in the bottom panels. We recall that the value function is almost constant with respect to the first two reduced-order states. Therefore, we want to study the value function with respect to the average temperature in the GS which is proportional to the last reduced-order state y_3. The left panels of Fig. 6.14 shows that, the value function decreases as the average temperature in the GS increases. The right panels of Fig. 6.25 shows that when the IS is full we have to wait if there is strong unsatisfied demand (blue dotted line in the bottom right panel). In case of strong overproduction we discharge the IS as long as the GS is not full, see the blue dotted line in the top right panel. However, this figure shows that when the IS is empty and there strong overproduction it is optimal to wait but when there is strong unsatisfied demand we discharge the GS as long as the latter is not empty and we only fire fuel when it is empty. The right panel of Fig. 6.25 shows that when the temperature in the IS is at the penalty threshold level ($p = p_{pen} = 60\ ^\circ$C), it is optimal to wait independent of the state of the residual demand.

Geothermal storage is full. Fig. 6.26 shows the value function and optimal strategy at time

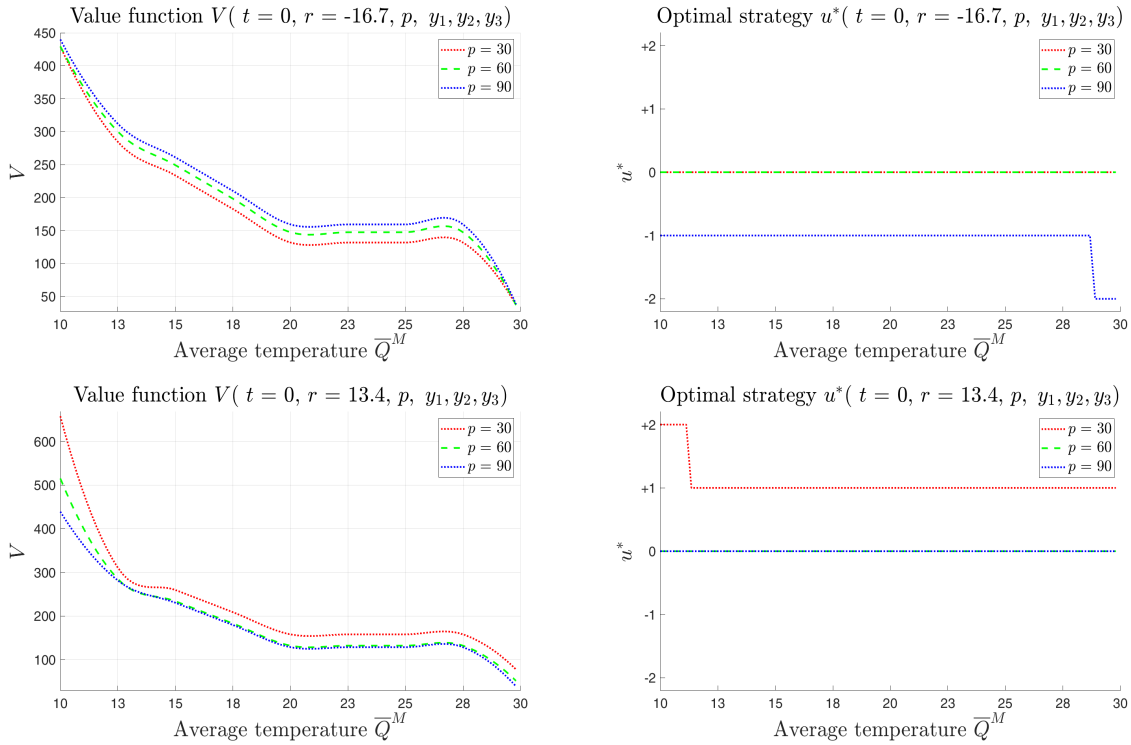


Figure 6.25: Comparison of value function (left) and optimal strategy (right) at time $t=0\ h$ as a function average temperature of the GS $\overline{Q}^M$ for a strong overproduction (top panel) and a strong unsatisfied demand (bottom panel) and for three fixed values of the average temperature in the IS.
Blue dotted line for IS full ($p=\overline{p}=90\ ^\circ$C), green dashed line for IS at the penalty threshold ($p=p_{pen}=60\ ^\circ$C), and red dotted line for IS empty ($p=\underline{p}=30\ ^\circ$C).

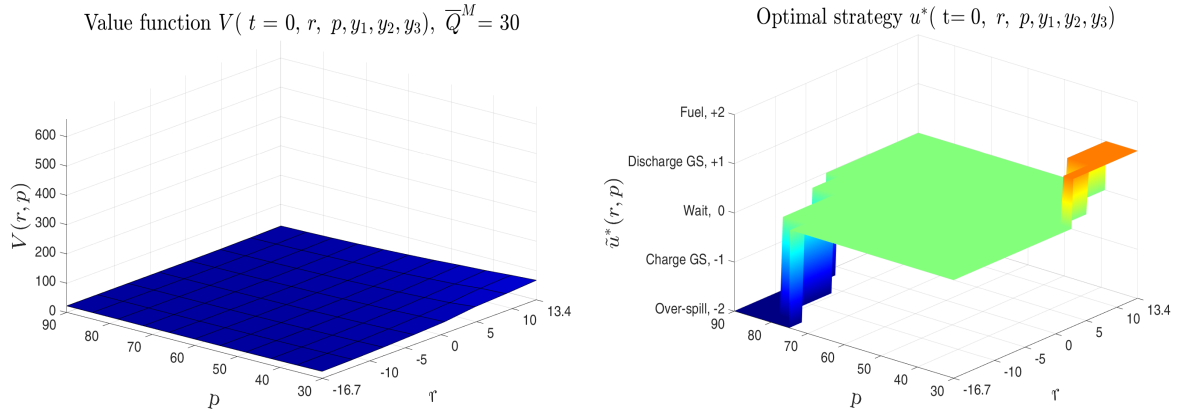


Figure 6.26: Value function (left) and optimal strategy (right) at the initial time $t=0\ h$ as a function of (r,p) for a full GS ($\overline{Q}^M=\overline{q}=30\ ^\circ$C).

$t=0\ h$ for a full GS ($\overline{Q}^M=\underline{q}=30\ ^\circ$C) as a function of (r,p). Fig. 6.27 shows a comparison of value functions and optimal strategies at time $t=0\ h$ as a function of p for three different values of the residual demand. Fig. 6.28 shows a comparison of value functions and optimal strategies at time $t=0\ h$ as a function of the residual demand r for three fixed values of the average temperature in the IS. The right panel of Fig. 6.27 shows that when there is a small unsatisfied demand (green dashed line) or strong overproduction (blue dotted line), it is optimal to only wait or apply spill-over if the IS is full whereas when there strong unsatisfied demand,

we have to discharge the GS to charge the IS if the latter is empty and only wait as soon as it is no longer empty. This results to a strong decay (red dotted line in the left panel) and a slow decay (blue dotted line in the left panel) in the value function as the average temperature in the IS increases.

Similarly, the right panels of Fig. 6.26 and Fig. 6.28 show that when the average temperature in the IS is above the penalty threshold level ($p > 60$ °C) it is optimal to wait or apply spill-over whereas when the average temperature in the IS is below the penalty threshold level ($p \leq 60$ °C), it is optimal to wait and only discharge the GS to charge the IS in case of strong unsatisfied demand. This results to a strong increase in (red dotted line in the left panel) and a slow increase (blue dotted line in the left panel) in the value function as the residual demand increases.

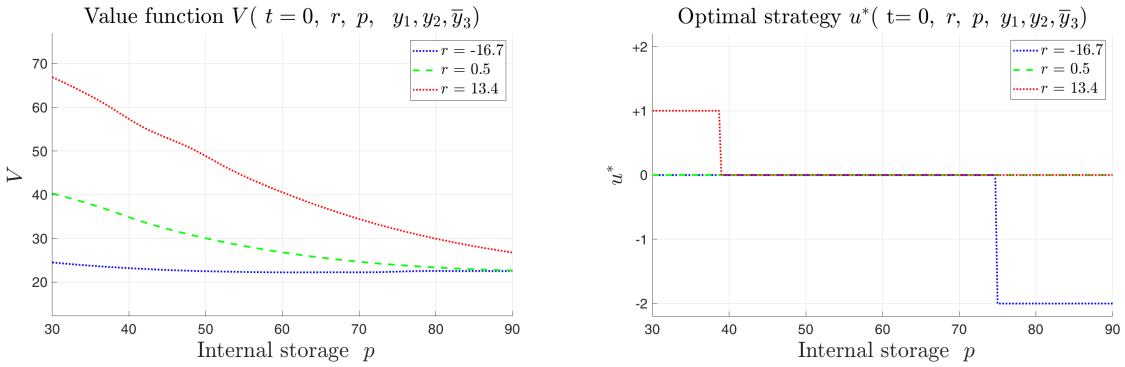


Figure 6.27: Comparison of value functions (left) and optimal strategies (right) at time $t = 0$ h as a function of p for three different values of the residual demand and a full GS ($\overline{Q}^M = \overline{q} = 30$ °C).
Blue dotted line for strong overproduction ($r = \underline{r} = -16.7$), green dashed line for very small unsatisfied demand ($r = 0.5$) and red dotted line for strong unsatisfied demand ($r = \overline{r} = 13.4$).

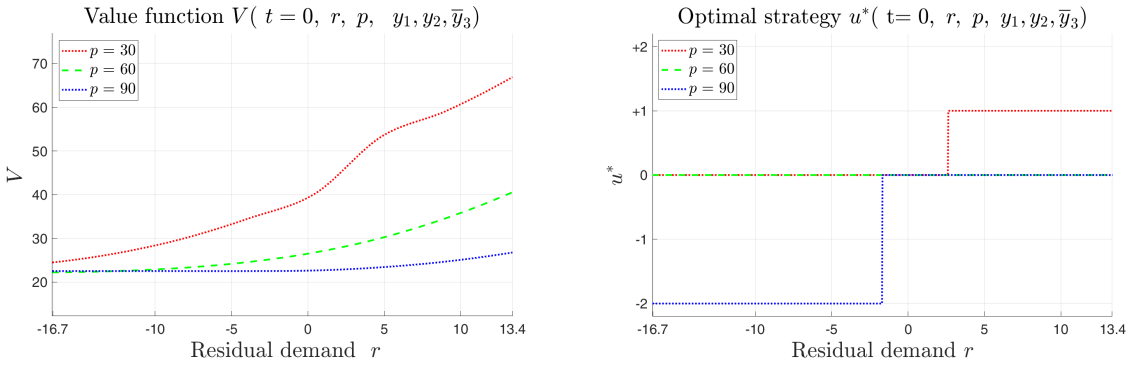


Figure 6.28: Comparison of value functions (left) and optimal strategies (right) at time $t = 0$ h as a function of the residual demand r for three fixed values of the average temperature in the IS and a full GS ($\overline{Q}^M = \overline{q} = 30$ °C).
Blue dotted line for IS full ($p = \overline{p} = 90$ °C), green dashed line for IS at the penalty threshold ($p = p_{pen} = 60$ °C), and red dotted line for IS empty ($p = \underline{p} = 30$ °C).

6.3.3 Optimal Paths of the State Process

In this subsection we present optimal paths of individual state variables. We aggregate the states of the reduced-order system to form the average temperature of the GS. In all figures the red,

blue, and black solid lines represent the average temperature in the IS and GS, and the residual demand, respectively. The values of the average temperature in the IS and GS are depicted on the left red y-axis and right blue y-axis, respectively. In all figures the black dotted lines (at 30 and 90) represent the minimum and the maximum values of the average temperature of the internal and GSs. The red marker at 60 (left axis) and the blue marker at 20 (right axis) are penalty level indicators for internal and GS, respectively. We also add the black solid line to show the residual demand relative to the zero level shown as a black dashed line. When the residual demand is above zero there is unsatisfied demand and when it is below there is over-production. In the background of these figures the red color represents the action of charging the IS by firing fuel at the maximum rate, orange represents charging the IS by discharging the GS at the maximum rate, green represents waiting periods (pumps and fuel fired-boiler off), light blue represents charging the GS by discharging the IS at the maximum rate and dark blue indicates over-spilling. We consider 5 cases in which we vary the initial temperatures in the IS and GS at time $t = 0$.

Start with full IS and empty GS. Fig. 6.29 shows optimal paths of the average temperatures in the IS and GS together with the residual demand when we start with a full IS and an empty GS. We observe that when the IS is full and the GS empty, it is more likely that we discharge the IS to charge the GS when there is overproduction. When the IS is empty and there is unsatisfied demand we charge it by discharging the GS and do not fire fuel as long as the GS is not empty. When t approaches the terminal time we have to do everything to avoid high penalty a the terminal time.

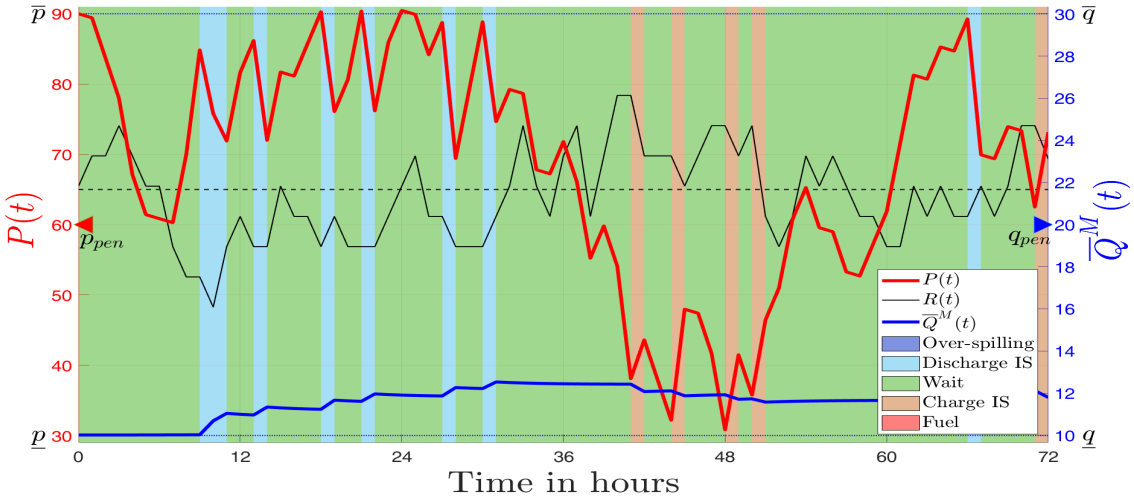


Figure 6.29: Optimal paths of the average temperature in the IS (red solid line), GS (blue solid line) for a full initial IS ($P(0) = \overline{p} = 90$ °C) and an empty initial GS ($\overline{Q}^M(0) = \underline{q} = 10$ °C), path of the residual demand (black solid line).

Start with empty IS and full GS. Fig. 6.30 shows the results when we start with an empty IS and a full GS. We observe that the states are controlled such that the state constraints are not violated and to avoid penalty at the terminal time. In this case it is more likely that we discharge the GS to charge the IS and never fire fuel to save cost. This also shows that the residual demand ans the average temperature in the IS are negatively correlated.

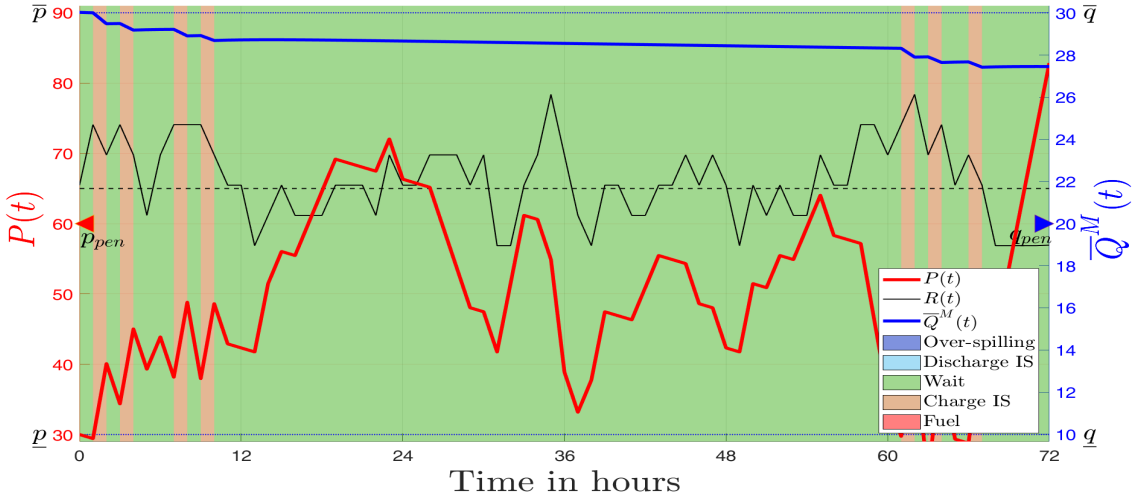


Figure 6.30: Optimal paths of the average temperature in the IS (red solid line), GS (blue solid line) for an empty initial IS ($P(0) = \underline{p} = 30$ °C) and a full initial GS ($\overline{Q}^M(0) = \overline{q} = 30$ °C), path of the residual demand (black solid line).

Start with empty IS and empty GS. Fig. 6.31 shows the results when we start with both IS and GS empty. As in the Figures 6.30 and 6.29, the states are controlled such that the state constraints are not violated and to avoid high penalty at the terminal time. When both storages are empty and there is unsatisfied demand we have to fire fuel. When the IS is full and and there is overproduction we discharge the IS to charge the GS but we make sure that the average temperature in the IS is above the penalty threshold 60 °C at the terminal time.

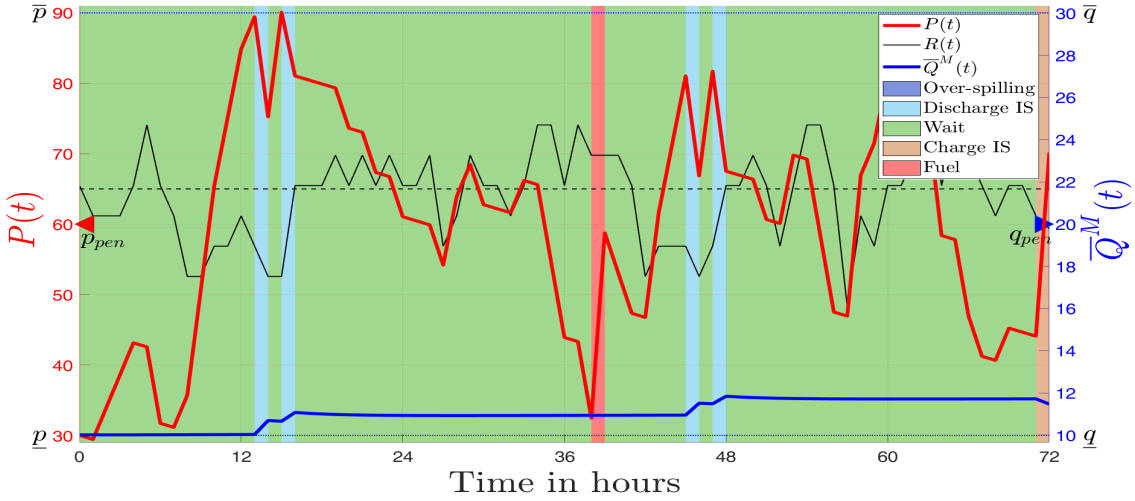


Figure 6.31: Optimal paths of the average temperature in the IS (red solid line), GS (blue solid line) for an empty initial IS ($P(0) = \underline{p} = 30$ °C) and an empty initial GS ($\overline{Q}^M(0) = \underline{q} = 10$ °C), path of the residual demand (black solid line).

Start with full IS and full GS. Fig. 6.32 shows the results when we start with both IS and GS full, and a small residual demand. As in the Figures 6.29, 6.30, and 6.31 the states are controlled such that the state constraints are not violated and to avoid penalty at the terminal time. When

both storages are full and there is overproduction $(R < 0)$ we have to apply over-spilling and when the IS is empty we charge it by discharging the GS an never fire fuel to save cost.

we have to

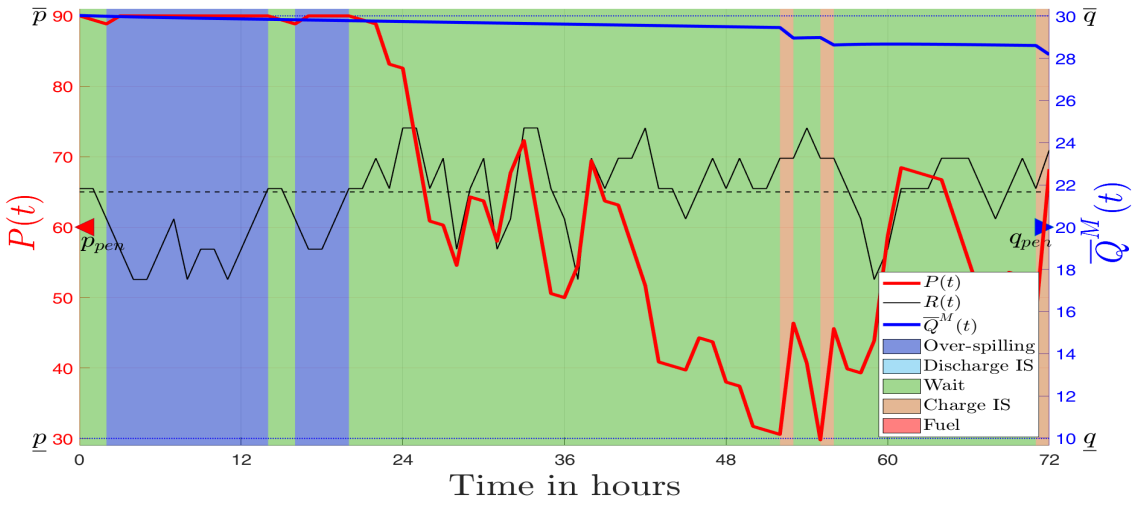


Figure 6.32: Optimal paths of the average temperature in the IS (red solid line), GS (blue solid line) for a full initial IS $(P(0) = \overline{p} = 90\ ^\circ\text{C})$ and a full initial GS $(\overline{Q}^M(0) = \overline{q} = 30\ ^\circ\text{C})$, path of the residual demand (black solid line).

Start with IS and GS at penalty thresholds. Fig. 6.33 shows the results when we start with

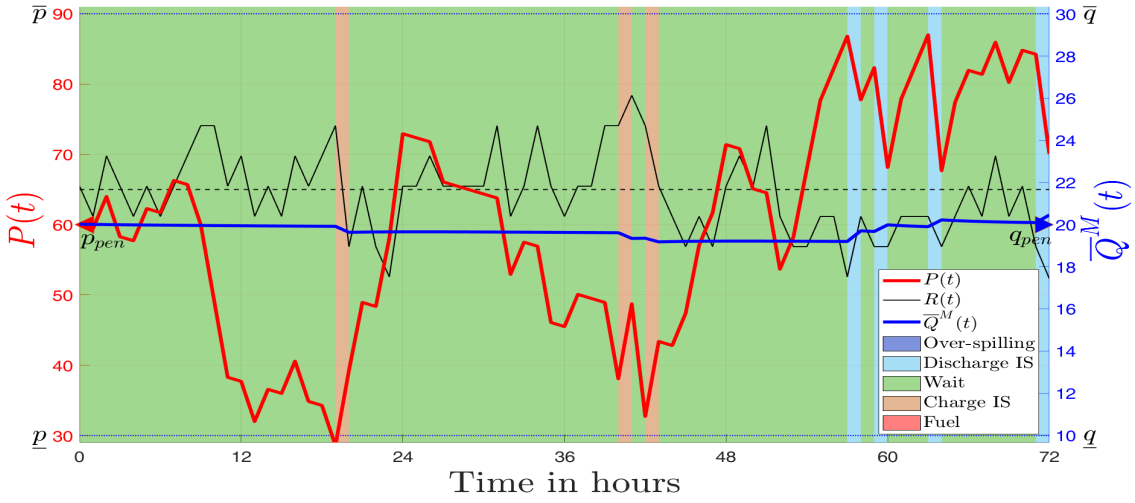


Figure 6.33: Optimal paths of the average temperature in the IS (red solid line), GS (blue solid line) for initial IS at the penalty threshold $(P(0) = p_{pen} = 60\ ^\circ\text{C})$ and initial GS at the penalty threshold $(\overline{Q}^M(0) = q_{pen} = 20\ ^\circ\text{C})$, path of the residual demand (black solid line).

initial IS and initial GS at the penalty thresholds, and a small residual demand. The states must be controlled such that the state constraints are not violated and to avoid high penalty at the terminal time. We observe in this case that, it is more likely to wait or do nothing but when the IS is full we discharge it to charge the GS. When the IS is empty we charge it only by discharging the GS and never fire fuel to save cost.

CHAPTER 7

Approximate Solution of the Markov Decision Processes

Introduction

This chapter aims to approximate the value function of the discrete-time optimal control problem and to derive the approximate optimal control using numerical methods. We recall the value function of the MPD satisfies the Bellman (6.19) which contains the conditional expectation $\mathbb{E}_{n,x}\left[V(n+1,X^{\mathbf{u}}_{n+1})\right]$. We have mentioned in Chapter 5 that the challenge of many numerical methods such as backward recursion method is that it becomes computationally intractable if the dimension of the state space is high or if no closed-form expressions of the conditional expectation are available. This leads to the so-called curse of dimensionality and the goal of this chapter is to introduce some methods to overcome this curse of dimensionality. Note that in the optimal control problem for a residential heating system considered in this thesis the dimension of the state process increases when the dimension of the reduced-order system of the geothermal ℓ increases. When the dimension of the reduced-order system ℓ is high, the control problem cannot be solved using direct methods such as backward recursion. In this case, we have to resort to an approximate solution using some numerical methods such as least-squares Monte Carlo [10, 52, 71, 81], approximate dynamic programming [2, 88, 101], Q-learning, neural network, and deep learning methods. In this chapter we are going to briefly describe the first two methods and show that they can be used to efficiently approximate the value function. This is based on a large number of simulated paths, where conditional expectations are replaced with least-squares regression approximations. The least-squares Monte Carlo techniques help to generate samples from the distribution of the random perturbations $\mathcal{E}_1,\ldots,\mathcal{E}_N$, and to approximate the expectation $\mathbb{E}_{n,x}\left[V(n+1,X^{\mathbf{u}}_{n+1})\right]$ under this distribution. On the other hand, the approximate dynamic programming uses parametric and non-parametric approaches to approximate the so-called post decision value function which solves an equivalent dynamic programming equation. The idea of the approximate dynamics programming is to first split the transition operator (6.16) into two parts called post- and pre-decision operators, then derive the equivalent Bellman equation called post-decision dynamic programming equation and solve it using iterative methods. The details in the theory and the application of approximate dynamics programming can be found in Powell [88].

The rest of this chapter is organized as follows. In Sec. 7.1we briefly introduce the Least

Square Monte Carlo methods for solving the Bellman equation and in Sec. 7.2 we introduce the concepts of post-decision state, the post-decision value function, and present some methods for solving the associated post-decision dynamic programming equation.

7.1 Least-Squares Monte Carlo Methods

Solving the dynamic programming equation (6.19) using backward recursion requires the computation of the value function and the optimal decision rule for all points (n,x) in the time-state space $\{0,1,\ldots,N\} \times \mathcal{X}$. This makes the computational time very large and the computation becomes impossible when the dimension gets higher. Now, assume that for a given optimal control problem, we already know some properties of the value function and of the optimal decision rule. For example, for the stochastic optimal control problem considered in this thesis we know based on the numerical results presented in Sec. 6.3 some properties of the value function and the decision rule. We know that the value function decreases as the last reduced-order state of the GS increases and it is almost constant with respect to the first $\ell - 1$ reduced-order states of the GS. Further, it increases as the residual demand increases and decreases as the average temperature in the internal storage increases. In addition, the properties and the form of the value function at the terminal time N is known. Therefore, it is possible to guess a reasonable regression ansatz for the value function of the form

$$V(n,x) \simeq \overline{V}(n,x) = \sum_{i=0}^{L} \Theta_i(n)\Gamma_i(x), \tag{7.1}$$

where $\Theta_i(n)$ are unknown coefficients and $\Gamma_i(x)$ are known ansatz functions which can be some suitable polynomials. The regression ansatz function (7.1) can be considered as an abstract Fourier series or Galerkin ansatz for the value function $V(n,x), x \in \mathcal{X}$. We may also consider the optimal decision to be given in such a form, i.e.,

$$\widetilde{u}^{\star}(n,x) = \sum_{i=0}^{L} \overline{\Theta}_i(n)\overline{\Gamma}_i(x),$$

where $\overline{\Theta}_i(n)$ are unknown coefficients and $\overline{\Gamma}_i(x)$ are known ansatz functions.
For example, for $x \in \mathcal{X} \subset \mathbb{R}$, the regression ansatz can be chosen as quadratic polynomial

$$\overline{V}(n,x) = \widetilde{\Theta}_0(n) + \widetilde{\Theta}_1(n)x + \widetilde{\Theta}_2(n)x^2,$$

where the unknown coefficients $\widetilde{\Theta}_0(n), \widetilde{\Theta}_1(n), \widetilde{\Theta}_2(n) \in \mathbb{R}$, or as transcendental functions of the form

$$\overline{V}(n,x) = \overline{\Theta}_0(n) + \overline{\Theta}_1(n)\cosh(x), \quad \text{or} \quad \overline{V}(n,x) = \overline{\Theta}_2(n) + \overline{\Theta}_3(n)\arctan(x),$$

with unknown coefficients $\overline{\Theta}_0(n), \overline{\Theta}_1(n), \overline{\Theta}_2(n), \overline{\Theta}_3(n) \in \mathbb{R}$.
For the multidimensional case, $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d, d \in \mathbb{N}$, the ansatz function can be chosen as a multivariate polynomial of the form

$$\overline{V}(n,x) = \Theta_0(n) + \Theta_1(n)^{\top}x + x^{\top}\Theta_2(n)x, \tag{7.2}$$

where, $\Theta_0(n) \in \mathbb{R}$, $\Theta_1(n) \in \mathbb{R}^d$, and $\Theta_2(n) \in \mathbb{R}^{d\times d}$. Based on the dimension of the state process $x = (r, f, p, \widetilde{Y}^1, \widetilde{Y}^2, \ldots, \widetilde{Y}^\ell) \in \mathcal{X} \subset \mathbb{R}^{3+\ell}$ and the above mentioned properties of the value function, a suitable regression ansatz for the value function can be a multivariate polynomial of degree 2 of the form (7.2).

The main idea of least-squares Monte Carlo (LSM) is to consider (7.1) as a regression ansatz for the value function with known ansatz functions $\Gamma_i(x)$ and unknown regression coefficients $\Theta_i(n)$. Instead of computing the value function $V(n,x)$ for all $(n,x) \in \{0,1,\ldots,N-1\} \times \mathcal{X}$ we only compute the $L+1$ unknown coefficients $\Theta_0(n), \Theta_1(n), \ldots, \Theta_L(n)$, in each time step by backward recursion.

The starting point of the LSM algorithm is the terminal time N where we already know the value function

$$V(N,x) = \psi(x) \simeq \sum_{i=0}^{L} \Theta_i(N)\Gamma_i(x).$$

We choose $M \in \mathbb{N}$ representative points $z_1, z_2, \ldots, z_M$, in the state space $\mathcal{X}$ which serve as samples of the independent variable x of the regression problem. Then we compute the responses $y_j = \psi(x_j),\ j = 1,\ldots,M$, to get the associated sample of the dependent variable y. Then, we compute the regression coefficients $\Theta(N) = (\Theta_0(N), \ldots, \Theta_L(N))^\top$ as least-square estimators of the regression problem

$$\widehat{\Theta}(N) = \underset{\Theta(N)\in\mathbb{R}^{L+1}}{\operatorname{argmin}} \left\{ \sum_{j=1}^{M} \left(\psi(z_j) - \sum_{i=0}^{L} \Theta_i(N)\Gamma_i(z_j) \right)^2 \right\}.$$

We are now in the position to start the backward recursion. For $n = N-1, N-2, \ldots, 1, 0$, given the already known coefficients $\Theta_0(n+1), \ldots, \Theta_L(n+1)$, we derive the unknown coefficients $\Theta_1(n), \ldots, \Theta_L(n)$ using the dynamic programming equation (6.19) given as follow:

$$V(n,x) = \inf_{a\in\mathcal{U}(n,x)} \Big\{ \Psi(n,x,a) + \mathbb{E}_{n,x}\big[V(n+1, X_{n+1}^a)\big] \Big\}.$$

Substituting $V(n+1, X_{n+1}^a)$ by the known regression ansatz $\overline{V}(n+1, X_{n+1}^a)$ and using the transition operator $X_{n+1}^a = \mathcal{T}_n(x, a, \mathcal{E}_{n+1})$ given in (6.16), the above equation becomes

$$\begin{aligned}
V(n,x) &= \inf_{a\in\mathcal{U}(n,x)} \Big\{ \Psi(n,x,a) + \mathbb{E}_{n,x}\big[V(n+1, X_{n+1}^a)\big] \Big\} \\
&\simeq \inf_{a\in\mathcal{U}(n,x)} \Big\{ \Psi(n,x,a) + \mathbb{E}_{n,x}\big[\overline{V}(n+1, X_{n+1}^a)\big] \Big\} \\
&= \inf_{a\in\mathcal{U}(n,x)} \Big\{ \Psi(n,x,a) + \sum_{i=0}^{L} \Theta_i(n+1)\mathbb{E}_{n,x}\big[\Gamma_i(X_{n+1}^a)\big] \Big\} \\
&= \inf_{a\in\mathcal{U}(n,x)} \Big\{ \Psi(n,x,a) + \sum_{i=0}^{L} \Theta_i(n+1)\mathbb{E}_{n,x}\big[\Gamma_i(\mathcal{T}_n(x,a,\mathcal{E}_{n+1}))\big] \Big\},
\end{aligned}$$

where $\mathcal{T}_n$ is the transition operator given in (6.16). Note that given x and a the conditional expectation $\mathbb{E}_{n,x}\big[\Gamma_i(\mathcal{T}_n(x,a,\mathcal{E}_{n+1}))\big]$ is the unconditional expectation with respect to the random variable $\mathcal{E}_{n+1}$. Then, using Monte Carlo method, we can generate $K \geq 1$ realizations

$\mathcal{E}_{n+1}(\omega_1),\ldots,\mathcal{E}_{n+1}(\omega_K)$ from the distribution of the random perturbations $\mathcal{E}_{n+1}$ to approximate the expectation, we get

$$\mathbb{E}_{n,x}\big[\Gamma_i(\mathcal{T}_n(x,a,\mathcal{E}_{n+1}))\big] \simeq \frac{1}{K}\sum_{k=1}^{K}\Gamma_i(\mathcal{T}_n(x,a,\mathcal{E}_{n+1}(\omega_k))) =: \widehat{G}_i(n,x,a).$$

Hence, we get the approximation

$$\begin{aligned} V(n,x) &\simeq \inf_{a\in\mathcal{U}(n,x)}\left\{\Psi(n,x,a)+\sum_{i=0}^{L}\Theta_i(n+1)\widehat{G}_i(n,x,a)\right\} \\ &= \inf_{a\in\mathcal{U}(n,x)}\left\{\Psi(n,x,a)+g(n,x,a)\right\} \\ &=: \widehat{V}(n,x), \end{aligned}$$

where $g(n,x,a)=\sum_{i=0}^{L}\Theta_i(n+1)\widehat{G}_i(n,x,a)$. Then, we are left with above deterministic pointwise optimization problem with respect to $a\in\mathcal{U}(n,x)$. Now, assume that this problem can be solved by standard methods. We choose again $M\in\mathbb{N}$ representative points $z_1,z_2,\ldots,z_M$, in the state space $\mathcal{X}$ which serve as samples of the independent variable x, repeat the above steps, and compute the Monte Carlo approximations

$$y_j=\widehat{V}(n,z_j)\simeq V(n,z_j),\quad j=1,\ldots,M.$$

Then, we compute the unknown coefficients $\Theta(n)=(\Theta_0(n),\ldots,\Theta_L(n))^\top$ as least-squares estimators of the regression problem

$$\widehat{\Theta}(n)=\underset{\Theta\in\mathbb{R}^{L+1}}{\operatorname{argmin}}\left\{\sum_{j=1}^{M}\left(\widehat{V}(n,z_j)-\sum_{i=0}^{L}\Theta_i(n)\Gamma_i(z_j)\right)^2\right\}.$$

We summarize the LSM method in the following algorithm. We refer to [10, 37, 106, 129] for the convergence of the LSM method.

Algorithm 3: Least-squares Monte Carlo method

Result: Approximation of the value function V and the decision rule $\widetilde{u}^\star$

Step 1 At time N:

- Choose the ansatz functions $\Gamma_0(x),\ldots,\Gamma_L(x)$;
- Choose M representative points $z_1,z_2,\ldots,z_M$, in $\mathcal{X}$ as samples for the variable x;
- Compute the responses $y_j=\psi(z_j),\ j=1,\ldots,M$, for the associated samples of the dependent variable y;
- Compute the regression coefficients $\Theta(N)=(\Theta_0(N),\ldots,\Theta_L(N))^\top$ as least-squares estimators of the regression problem

$$\widehat{\Theta}(N)=\underset{\Theta(N)\in\mathbb{R}^{L+1}}{\operatorname{argmin}}\left\{\sum_{j=1}^{M}\left(\psi(z_j)-\sum_{i=0}^{L}\Theta_i(N)\Gamma_i(z_j)\right)^2\right\}.$$

Step 2 **for** `n=N-1 to 0` **do**

a. Generate K realizations $\mathcal{E}_{n+1}(\omega_1),\ldots,\mathcal{E}_{n+1}(\omega_K)$, from the distribution of the random perturbation $\mathcal{E}_{n+1}$ and compute

$$\mathbb{E}_{n,x}\big[\Gamma_i(\mathcal{T}_n(x,a,\mathcal{E}_{n+1}))\big]\simeq\widehat{G}_i(n,x,a)=\tfrac{1}{K}\sum_{k=1}^{K}\Gamma_i(\mathcal{T}_n(x,a,\mathcal{E}_{n+1}(\omega_k)));$$

b. Find $\mathbb{E}_{n,x}\big[V(n+1,X_{n+1})\big]\simeq g(n,x,a)=\sum_{i=0}^{L}\Theta_i(n+1)\widehat{G}_i(n,x,a)$, the Monte Carlo approximation of the condtional expection;

c. Find $\widehat{V}(n,x)\simeq V(n,x)$, the approximation of $V(n,x)$ and $\widehat{u}^\star(n,x)\simeq\widetilde{u}^\star(n,x)$, the approximation of $\widetilde{u}^\star(n,x)$ at time n by solving

$$\widehat{V}(n,x)=\inf_{a\in\mathcal{U}(n,x)}\Big\{\Psi(n,x,a)+g(n,x,a)\Big\},$$
$$\widehat{u}^\star(n,x)=\operatorname*{argmin}_{a\in\mathcal{U}(n,x)}\Big\{\Psi(n,x,a)+g(n,x,a)\Big\};$$

d. Use the same representative points $z_1,z_2,\ldots,z_M$, as samples of the independent variable x generated above for $n=N$ to compute the Monte Carlo approximations

$$y_j=\widehat{V}(n,z_j)\simeq V(n,z_j),\quad j=1,\ldots,M;$$

c. Compute the coefficients $\Theta(n)=(\Theta_0(n),\ldots,\Theta_L(n))^\top$ as least-squares estimators of the regression problem

$$\widehat{\Theta}(n)=\underset{\Theta\in\mathbb{R}^{L+1}}{\operatorname{argmin}}\left\{\sum_{j=1}^{M}\left(\widehat{V}(n,z_j)-\sum_{i=0}^{L}\Theta_i(n)\Gamma_i(z_j)\right)^2\right\}.$$

end

7.2 Approximate Dynamic Programming Methods

In this section we briefly describe the approximate dynamic programming method for solving the optimal control problem as prospective method to overcome the curse of dimensionality. This method is based on the transition operator associated to the MDP, $\mathcal{T}_n : \mathcal{X} \times \overline{\mathcal{U}} \times \mathcal{Z} \to \mathcal{X}$ given in (6.16), which for the convenience of the reader is also given below

$$X_{n+1} = \mathcal{T}_n(X_n, u_n, \mathcal{E}_{n+1}), \;\; n = 0, 1, \ldots N-1,$$

where $X_{n+1} = (R_{n+1}, F_{n+1}, P_{n+1}, \widetilde{Y}_{n+1})$ with individual states given by the recursion

$$\begin{aligned}
R_{n+1} &= R_n e^{-\beta_R \Delta_N} + \mu_{R,n}(1 - e^{-\beta_R \Delta_N}) + \Sigma_{R,n} \zeta^R_{n+1}, \\
F_{n+1} &= F_n e^{-\beta_F \Delta_N} + \mu_{F,n}(1 - e^{-\beta_F \Delta_N}) + \Sigma_{F,n} \zeta^F_{n+1}, \\
P_{n+1} &= P_n e^{-\gamma \Delta_N} + \Upsilon_n(u_n, \widetilde{Y}_n) + \Sigma_{P,n} \left(\sqrt{1 - \rho^2_{RP,n}} \mathcal{E}^P_{n+1} + \rho_{RP,n} \mathcal{E}^R_{n+1} \right), \\
\widetilde{Y}_{n+1} &= \widetilde{Y}_n e^{\widetilde{A} \Delta_N} + (e^{\widetilde{A} \Delta_N} - \mathbb{I}_\ell) \widetilde{A}^{-1} \widetilde{B} g^{\nu}_n,
\end{aligned}$$

where $\mathcal{X} \in \mathbb{R}^{3+\ell}$ and Υ_n is given by (6.10). In the transition operator $(\mathcal{E}_1, \ldots, \mathcal{E}_N)$ is a sequence of i.i.d multivariate standard normally distributed random variables.
The main idea is to decompose the above transition operator into post- and pre-decision operator and to derive the dynamic programming equation (DPE) associated to the post-decision state called post-decision dynamic programming equation (PDPE). Then, find the approximate solution the PDPE using the parametric or non-parametric method.

Post-Decision state. The above transition operator $\mathcal{T}_n : \mathcal{X} \times \overline{\mathcal{U}} \times \mathcal{Z} \to \mathcal{X}$ is decomposed into two parts:

$$\mathcal{T}_n = \mathcal{T}^{(2)}_n \circ \mathcal{T}^{(1)}_n$$

where $\mathcal{T}^{(1)}_n : \mathcal{X} \times \overline{\mathcal{U}} \to \mathcal{X}$ is a post-decision operator defined by

$$\mathcal{T}^{(1)}_n(x, a) = \mathcal{T}_n(x, a, 0),$$

and $\mathcal{T}^{(2)}_n : \mathcal{X} \times \mathcal{Z} \to \mathcal{X}$ is a pre-decision operator defined by

$$\mathcal{T}^{(2)}_n(x, \varepsilon) = \mathcal{T}_n(x, 0, \varepsilon).$$

Then

$$\mathcal{T}_n(x, a, \varepsilon) = \mathcal{T}^{(2)}_n(\mathcal{T}^{(1)}_n(x, a), \varepsilon).$$

The $X^P_n = X^{\mathbf{u},P}_n = \mathcal{T}^{(1)}_n(X^{\mathbf{u}}_n, u_n)$ associated with the post-decision operator $\mathcal{T}^{(1)}$ is called post-decision state of $X^{\mathbf{u}}_n$ and it holds that $X^{\mathbf{u}}_{n+1} = \mathcal{T}^{(2)}_n(X^{\mathbf{u},P}_n, \mathcal{E}_{n+1})$. The post decision state is the state immediately after the action is taken but before the uncertainty given by the exogenous perturbation is realized. The post decision state only captures the deterministic part of the action u_n. However, This decomposition is helpful since the transition operator $\mathcal{T}_n$ consists of a

deterministic and a stochastic part. In this setting, we separate the effect of the decision and the incoming stochastic perturbation. The state associated with the pre-decision operator is called pre-decision state. It is the state of the system after the uncertainty is realized. Next, we show that this formulation can be used to simplify the complexity of the optimal control problem. Throughout this section, we will use z as a placeholder for post-decision state.

Post-decision Value Function. The post-decision value function is defined by

$$V^P(n,z) = \mathbb{E}\big[V(n+1, X_{n+1}^{\mathbf{u}}) \big| X_n^{\mathbf{u},P} = z\big], \;\; n = 0,1,\ldots,N-1. \tag{7.3}$$

This can be interpreted as the minimum expected cost that the controller can achieve immediately after the action u_n has been taken at time n and the post-decision state is z. Since the post-decision state $X_n^{\mathbf{u},P}$ is $\mathcal{F}_n$-measurable and the stochastic perturbation $\mathcal{E}_{n+1}$ is independent of $\mathcal{F}_n$, it holds

$$V^P(n,z) = \mathbb{E}\big[V(n+1, \mathcal{T}_n^{(2)}(X_n^{\mathbf{u},P}, \mathcal{E}_{n+1})) \big| X_n^{\mathbf{u},P} = z\big] = \mathbb{E}\big[V(n+1, \mathcal{T}_n^{(2)}(z, \mathcal{E}_{n+1}))\big].$$

Now, since the state X_n and the control $u_n = a$ at time n are $\mathcal{F}_n$-measurable we have

$$\begin{aligned}
\mathbb{E}\big[V(n+1, X_{n+1}^{\mathbf{u}}) \big| X_n^{\mathbf{u}} = x\big] &= \mathbb{E}\big[V(n+1, X_{n+1}^{\mathbf{u}}) \big| X_n^{\mathbf{u},P} = \mathcal{T}_n^{(1)}(x,a)\big] \\
&= \mathbb{E}\big[V(n+1, \mathcal{T}_n^{(2)}(\mathcal{T}_n^{(1)}(x,a), \mathcal{E}_{n+1}))\big] \\
&= V^P(n, \mathcal{T}_n^{(1)}(x,a)).
\end{aligned}$$

Substituting into the dynamic programming equation (6.19), we get

$$V(n,x) = \inf_{a \in \overline{\mathcal{U}}(x)} \Big\{ \Psi(n,x,a) + V^P(n, \mathcal{T}_n^{(1)}(x,a)) \Big\}, \; x \in \mathcal{X}, \; n = 0,1,\ldots,N-1. \tag{7.4}$$

Therefore, once the exact or approximate post-decision value function $V^P(n,z)$ is known for all $(n,z) \in \{0,1,\ldots,N-1\} \times \mathcal{X}$, the value function $V(n,x)$ and the optimal decision rule $\widetilde{u}^*(n,x)$ can be found by solving the pointwise deterministic optimization problem in (7.4), instead of the pointwise deterministic optimization problem in (6.19). Hence, solving the optimal control problem is reduced to computing the post-decision value function V^P. In Equation (7.4), V is the minimal value of the function $f(a) = \Psi(n,x,a) + V^P(n, \mathcal{T}_n^{(1)}(x,a))$ and $\widetilde{u}^*$ is the minimizer.

Remark 7.2.1 If we only need the value function $V(n,x)$ and the optimal decision rule $\widetilde{u}^*(n,x)$ for the points $(n, X_n^{\mathbf{u}^*})$ along the path of the optimal state process $X^{\mathbf{u}^*}$, then we only need to know the post-decision value function V^P for all (n,x) whereas the pre-decision value function V has to be determined only along the path of $X^{\mathbf{u}^*}$.

7.2.1 Post-Decision Dynamic Programming Equation

In this part we derive the dynamic programming equation for the post-decision value function V^P. The following theorem gives the recursion for the post-decision value function called post-decision dynamic programming equation (PDPE).

Theorem 7.2.2 (Post-decision DPE) The value function satisfies the so-called post-decision dynamic programming equation or post-decision Bellman equation given for $n = 0, 2, \ldots, N$, by

$$V^P(n,z) = \mathbb{E}\left[\inf_{a\in\overline{\mathcal{U}}(\mathcal{T}_n^{(2)}(z,\mathcal{E}_{n+1}))} \left\{\Psi(n+1,\mathcal{T}_n^{(2)}(z,\mathcal{E}_{n+1}),a) + V^P(n+1,\mathcal{T}_{n+1}^{(1)}(\mathcal{T}_n^{(2)}(z,\mathcal{E}_{n+1}),a))\right\}\right]. \tag{7.5}$$

Proof. From the definition of the post-decision value function V^P given in (7.3), we have

$$V^P(n,z) = \mathbb{E}\big[V(n+1,X_{n+1}^{\mathbf{u}})\big|X_n^{\mathbf{u},P} = z\big] = \mathbb{E}\big[V(n+1,\mathcal{T}_n^{(2)}(z,\mathcal{E}_{n+1}))\big]. \tag{7.6}$$

Then, relation (7.4) for the dynamic programming equation implies for $n = 0, 1, \ldots, N-1$

$$V(n+1,X_{n+1}^{\mathbf{u}}) = \inf_{a\in\overline{\mathcal{U}}(X_{n+1}^{\mathbf{u}})} \left\{\Psi(n+1,X_{n+1}^{\mathbf{u}},a) + V^P(n+1,\mathcal{T}_{n+1}^{(1)}(X_{n+1}^{\mathbf{u}},a))\right\}. \tag{7.7}$$

Substituting (7.7) into (7.6) yields

$$\begin{aligned} V^P(n,z) &= \mathbb{E}\left[\inf_{a\in\overline{\mathcal{U}}(X_{n+1}^{\mathbf{u}})} \left\{\Psi(n+1,X_{n+1}^{\mathbf{u}},a) + V^P(n+1,\mathcal{T}_{n+1}^{(1)}(X_{n+1}^{\mathbf{u}},a))\right\}\middle| X_n^{\mathbf{u},P} = z\right] \\ &= \mathbb{E}\left[\inf_{a\in\overline{\mathcal{U}}(\mathcal{T}_n^{(2)}(z,\mathcal{E}_{n+1}))} \left\{\Psi(n+1,\mathcal{T}_n^{(2)}(z,\mathcal{E}_{n+1}),a) + V^P(n+1,\mathcal{T}_{n+1}^{(1)}(\mathcal{T}_n^{(2)}(z,\mathcal{E}_{n+1}),a))\right\}\right]. \end{aligned}$$

□

Note that the unconditional expectation in equation (7.5) is taken with respect to the random variable $\mathcal{E}_n$. Now, we recall the dynamic programming equation also called pre-decision DPE given by

$$V(n,x) = \inf_{a\in\overline{\mathcal{U}}(x)} \left\{\Psi(n,x,a) + \mathbb{E}\left[V(n+1,\mathcal{T}_n^{(2)}(\mathcal{T}_n^{(1)}(x,a),\mathcal{E}_{n+1}))\right]\right\}. \tag{7.8}$$

We observe that in the post-decision DPE the minimization (inf) is inside the expectation, see equation (7.5). However, in the pre-decision DPE the minimization is over the expectation, see equation (7.8). The latter is a more involved optimization problem than the one in the PDPE. In the PDPE, we have a deterministic minimization problem for a fixed realization of the stochastic perturbation $\mathcal{E}_n$.

The main advantage of the post-decision DPE over the pre-decision or classical DPE is that the repeated computation of the expectation within the optimization can be avoided. Instead the PDPE requires the computation of a single expectation after solving the pointwise optimization problems. This is advantageous because in most cases the computation of the expectation is very difficult or impossible (for example, when the transition probabilities are unknown or complicated).

7.2.2 Non-Parametric Approximation of the Post-Decision Value Function

In this section we are going to describe the iterative procedure for the approximation of the post decision value function $V^P = V^P(n,z),\ n = 0,1,\ldots,N-1,\ z \in \mathcal{X}$. The starting point of the method is initial guess $\overline{V}^0(n,z)$ of the post-decision value function V^P for all (n,z). Such a guess can be derived by exploiting the structural properties of the control problem and its post-decision value function such as monotonicity, convexity, non-negativity, asymptotic behavior, etc. If no prior information is available, one can choose $\overline{V}^0(n,z) = 0$.

Remark 7.2.3 Note that as in every iterative method, the required number of iterations for obtaining a prescribed level of approximation accuracy depends critically on the initial guess. Further, we note that the pre-decision value function $V(n,x)$ is already known at the terminal time ($n = N$). Therefore, the post-decision value function has to be determined only for $n = 0,1,\ldots,N-1$.

Before describing the method we note that the updates or the improvements of the post-decision value function $\overline{V}^k$ are typically generated only at finite visited states $z_1,\ldots,z_k$. For a state space $\mathcal{X}$, with $|\mathcal{X}| = d$ (d very large or $d = \infty$), this will need some interpolation or smoothing between the visited states. This can be done using a non-parametric smoothing with kernel functions or with a parametric smoothing using regression ansatz functions. We will describe these two methods below. Now we describe the iterative procedure, adopted from the one established in Powell [88] and Dimitri [25], based on a value iteration approach and the post-decision DPE for the post-decision value function V^P.
In Algorithm 4, we consider a non-parametric procedure in which we include smoothing of of update in the neighborhood of the sample points z_n^k. For each time n we assume that we already know the previous approximation $\overline{V}^{k-1}(n,z_n^k)$ of the post-decision value function from the first $k-1$ iterations. Then, we generate in the iteration k a new observation of V called $\widehat{V}_n^k$ and update the approximation by $\overline{V}^k(n,z_n^k) = (1-\gamma^k)\overline{V}^{k-1}(n,z_n^k) + \gamma^k\widehat{V}_n^k$, with the learning rate $\gamma \in (0,1)$. The updating given in step 2b represents a weighted mean of the previous approximation $\overline{V}^{k-1}$ and the new approximation $\widehat{V}^k$.

Algorithm 4: Non-parametric method for the approximation of the post-decision value function

Result: Approximation of the post-decision value function V^P

Step 0 initialization:

- Choose $\overline{V}^0(n,z),\ \ (n,z) \in \{0,1,\ldots,N-1\} \times \mathcal{X}$;
- Choose the maximal number of iteration $K \in \mathbb{N}$;
- Choose the tolerance $\varepsilon_z > 0$;
- Set the counter $k = 1$;

Step 1 Initial value for the state process X

- Generate N realizations $\varepsilon_1^k, \varepsilon_2^k, \ldots, \varepsilon_N^k$, of the stochastic perturbation;
- Generate the initial value z_0^k of the post-decision state;

Step 2 **for** `n=0 to N-1` **do**

a. Find the next pre-decision state $x = \mathcal{T}_n^{(2)}(z_n^k, \varepsilon_{n+1}^k)$;

b. Compute the new approximation $\widehat{V}_n^k$ of the post-decision value function $V^P(n,z)$ at $z = z_n^k$ by solving

$$\widehat{V}_n^k = \inf_{a \in \overline{\mathcal{U}}(x)} \left\{ \Psi(n+1,x,a) + \overline{V}^{k-1}(n+1, \mathcal{T}_{n+1}^{(1)}(x,a)) \right\};$$

$$a_n^k = \operatorname*{argmin}_{a \in \overline{\mathcal{U}}(x)} \left\{ \Psi(n+1,x,a) + \overline{V}^{k-1}(n+1, \mathcal{T}_{n+1}^{(1)}(x,a)) \right\};$$

Update the previous approximation $\overline{V}^{k-1}$ using $\widehat{V}_n^k$ and the weighting factor/learning rate $\gamma^k \in (0,1)$, $\overline{V}^k(n, z_n^k) = (1-\gamma^k)\overline{V}^{k-1} + \gamma^k \widehat{V}_n^k$;

c. Compute the next post-decision state $z_{n+1}^k = \mathcal{T}_{n+1}^{(1)}(x, a_n^k)$;

end

Step 3 Check convergence criterion;

if $\left|\overline{V}^k - \overline{V}^{k-1}\right| < \varepsilon_z$ *or* $k = K$ **then**

Stop and return $V^P = \overline{V}^k$;

else

Increment k. Go to step 1.

end

Ideas for improving the update

We recall the update formula from step 2b of the algorithm 4: $\overline{V}^k(n, z_n^k) = (1-\gamma^k)\overline{V}^{k-1}(n, z_n^k) + \gamma^k \widehat{V}_n^k$, with the learning rate $\gamma^k \in (0,1)$. It turns out that in each iteration k and for each time n the new approximation $\overline{V}^k$ differs from the previous approximation $\overline{V}^{k-1}$ only in a single point

($z = z_n^k$) in state space.
This seems to be not efficient if the state state space $\mathcal{X}$ is continuous or of high cardinality and if we expect a slowly varying value function. In this case it is reasonable to share the new information about the post-decision value function learned by $\widehat{V}_n^k$ to some vicinity of the point $z = z_n^k$. This can be achieved by choosing a learning rate that depends on the state z. We replace γ^k by

$$\overline{\gamma}^k(z) = \gamma^k \zeta(z - z_n^k),$$

where ζ is some kernel function controlling the information sharing with $\zeta : \mathbb{R}^d \to [0,1]$ having the following properties:

- $\zeta(0) = 1$,
- $\zeta(|x|)$ is decreasing,
- $\zeta(x) \to 0$ for $|x| \to \infty$,

One example could be the Gaussian kernel $\zeta(x) = e^{-(\frac{x}{\delta_b})^2}$, where δ_b is the kernel bandwidth. The kernel bandwidth $\delta_b > 0$ controls how fast the learning rate $\overline{\gamma}^k(z)$ decays to zero if the distance from z to z_n^k becomes larger. It is reasonable to assume that $\overline{\gamma}^k$ decreases in $|z - z_n^k|$ and tends to zero for $|z - z_n^k| \to \infty$. The update formula in Algorithm 4 step 2b then reads for all z as

$$\overline{V}^k(n,z) = (1 - \overline{\gamma}^k(z))\overline{V}^{k-1}(n,z) + \overline{\gamma}^k(z)\widehat{V}_n^k.$$

Figure 7.1 shows the example of update with some information sharing by a Gaussian kernel in the vicinity of the point $z_n^k = 1$ with the learning rate $\gamma^k = 0.4$ and kernel bandwidths $\delta_b = 1/2$ (left) and $\delta_b = 1/10$ (right).

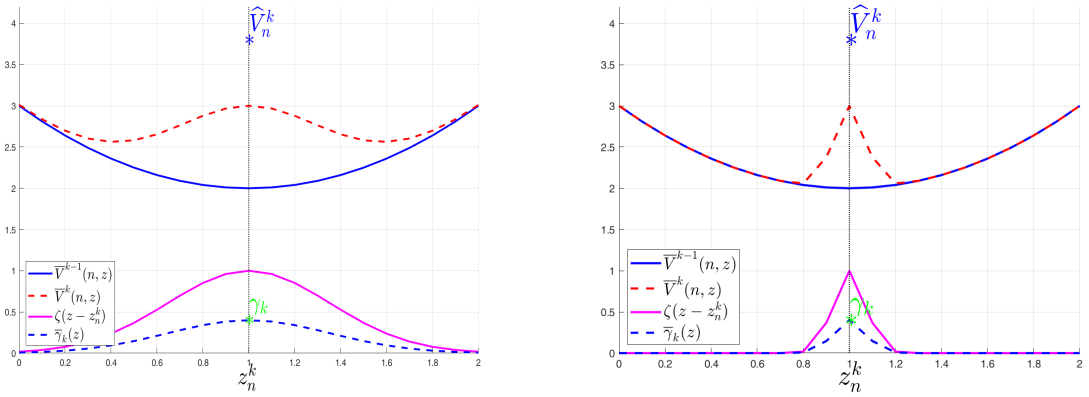


Figure 7.1: Update with sharing of information in the vicinity of $z_n^k = 1$, with $\gamma^k = 0.4$.
Left: bandwidth $\delta_b = 0.5$. Right: bandwidth $\delta_b = 0.1$.

Remark 7.2.4 In computer implementation the new update $\overline{V}^k$ in the Algorithm 4 is only stored for finitely many grid points $z^{(1)}, \ldots, z^{(M)}$ of the state space. A simple approximate solution of

$$\widehat{V}_n^k = \inf_{a \in \overline{\mathcal{U}}(x)} \left\{ \Psi(n+1, x, a) + \overline{V}^{k-1}(n+1, \mathcal{T}_{n+1}^{(1)}(x,a)) \right\}$$

can be done as follows:

- sample the set $\overline{\mathcal{U}}_{n+1}(x)$ of feasible actions by sufficiently many trial points $a^{(1)},\dots,a^{(K)}$,
- compute for all trial points $a^{(i)}, i=1,\dots,K$, the objective function

$$f(a^{(i)}) = \Psi(n+1,x,a^{(i)}) + \overline{V}^{k-1}(n+1,\mathcal{T}_{n+1}^{(1)}(x,a^{(i)}))$$

  and find the minimum of these K values.

This requires the values of $\overline{V}^{k-1}(n+1,\mathcal{T}_{n+1}^{(1)}(x,a^{(i)})),\ i=1,\dots,K$, at the points $\mathcal{T}_{n+1}^{(1)}(x,a^{(i)})$ which are in general no grid points $z^{(1)},\dots,z^{(M)}$. Hence, an interpolation procedure must be performed to efficiently compute the values $\overline{V}^{k-1}$ at the points $\mathcal{T}_{n+1}^{(1)}(x,a^{(i)})$ given the grid values $\overline{V}^{k-1}(n+1,z^{(j)}),\ j=1,\dots,M$, at the grid points $z^{(j)}),\ j=1,\dots,M$.

7.2.3 Parametric Approximation of the Post-Decision Value Function

In this subsection, we extend and combine the non-parametric approach with a regression approach. This will lead us to an approximation which does not suffer from the curse of dimensionality. We will not need any grid for the state space as in the previous approaches.
The basic idea is to work with a regression ansatz for the post-decision value function of the form

$$V^P(n,z) = \sum_{i=0}^{L} \Theta_i(n)\Gamma_i(z),$$

with known ansatz functions $\Gamma_i(z)$ and unknown regression coefficients $\Theta_i(n)$ to be approximated. Starting with an initial guess $\overline{\Theta}_i^0(n), i=0,\dots,L, n=0,\dots,N-1$, for the coefficients, i.e., the initial guess for the post-decision value function is

$$V^P(n,z) \simeq \overline{V}^0(n,z) = \sum_{i=0}^{L} \overline{\Theta}_i^0(n)\Gamma_i(z),$$

we will iteratively find new approximations for the regression coefficients $\overline{\Theta}_i^k(n)$ and the post-decision value function

$$\overline{V}^k(n,z) = \sum_{i=0}^{L} \overline{\Theta}_i^k(n)\Gamma_i(z),$$

for $k=1,2,\dots,K$. This is achieved by generating $n_j \in \mathbb{N}$ sample points $(z_n^j, \overline{V}_n^j),\ j=1,\dots,n_j$, for each $n=0,\dots,N-1$, which we fit to the regression ansatz to determine the new approximations of $\overline{\Theta}_i^k$.
Note that generating the above sample points follows the non-parametric approach considered in Subsec. 7.2.2 but without information sharing to neighboring points (no sharing during sampling) and using the same previous approximation of $\overline{V}^{k-1}$ for all sample points.

Algorithm 5: Parametric method to approximate the post-decision value function

Result: Approximation of the post-decision value function V^P

Step 0 initialization:

- Choose the ansatz functions $\Gamma_0(z),\ldots,\Gamma_L(z)$;
- Choose an initial guess for the regression coefficients $\overline{\Theta}_i^0(n),\ i=0,\ldots,L$, $n=0,\ldots,N-1$;
- Set $\overline{V}^0(n,z)=\sum_{i=0}^{L}\overline{\Theta}_i^0(n)\Gamma_i(z)$, $n=0,\ldots,N-1$;
- Choose the maximum number of iterations K, set the counter $k=1$;
- Choose the sample size n_j for the learning step;
- Choose the tolerance threshold ε_z and set the iteration counter $k=1$;

Step 1 Generate sample for the regression:

for $j=1$ `to` n_j **do**

- Generate realizations $\varepsilon_1^j,\varepsilon_2^j,\ldots,\varepsilon_N^j$, of random perturbations;
- Generate the initial value z_0^j of the post-decision state;

for `n=0 to N-1` **do**

a. Find the next pre-decision state $x=\mathcal{T}_n^{(2)}(z_n^j,\varepsilon_{n+1}^j)$;

b. Compute the new approximation $\widehat{V}_n^j$ of the post-decision value function $V^P(n,z)$ at $z=z_n^j$ by solving

$$\widehat{V}_n^j=\inf_{a\in\overline{\mathcal{U}}(x)}\left\{\Psi(n+1,x,a)+\overline{V}^{k-1}(n+1,\mathcal{T}_{n+1}^{(1)}(x,a))\right\};$$

$$a_n^j=\operatorname*{argmin}_{a\in\overline{\mathcal{U}}(x)}\left\{\Psi(n+1,x,a)+\overline{V}^{k-1}(n+1,\mathcal{T}_{n+1}^{(1)}(x,a))\right\};$$

Update the previous approximation $\overline{V}^{k-1}$ using $\widehat{V}_n^j$ and the weighting factor or learning rate $\gamma^k\in(0,1)$, $\overline{V}_n^j=(1-\gamma^k)\overline{V}^{k-1}(n,z_n^j)+\gamma^k\widehat{V}_n^j$;

c. Compute the next post-decision state $z_{n+1}^j=\mathcal{T}_{n+1}^{(1)}(x,a_n^j)$;

end

end

Step 2 Regression:
for $n = 0 \;\; to \;\; N-1$ **do**

- Compute new approximations of the regression coefficients $\overline{\Theta}_i^k(n),\ i = 0,\ldots,L$, from the sample poits $(z_n^j, \overline{V}_n^j)$, $j = 1,\ldots,n_j$;
- Set $\overline{V}^k(n,z) = \sum_{i=0}^{L} \overline{\Theta}_i^k(n)\Gamma_i(z)$

Step 3 Check convergence criterion;
if $\left|\overline{V}^k - \overline{V}^{k-1}\right| < \varepsilon_z$ *or* $k = K$ **then**
Stop and return $V^P = \overline{V}^k;\ \Theta_i = \overline{\Theta}_i^k,\ i = 0,\ldots,L$;
else
Increment k. Go to step 1.

To ensure the convergence of the Algorithms 4 and 5 the learning rate must be nonnegative and should satisfy the following properties.

$$\sum_{k=0}^{\infty} \gamma^k = \infty, \quad \sum_{k=0}^{\infty} (\gamma^k)^2 < \infty.$$

The proof of the convergence of the Algorithms 4 and 5 can be adapted from the proof of convergence of the iterative methods given in Dimitri [25, Chapter 4]

CHAPTER 8

Summary and Outlook

8.1 Summary

In this thesis, we have investigated the stochastic optimal control problem for the cost-optimal management of the residential heating equipped with several production and consumption units. As a special feature, the manager of the heating system has access to an external GS allowing for storing local overproduction over longer periods and providing heat in times of high demand to save cost of heat production using fuel or electricity. The main focus was to minimize the expected aggregated costs of generating thermal energy and to running the system. This led to a challenging optimization problem whose one state variable was described by a PDE. This is a non-standard feature and does not fit to the standard framework for stochastic optimal control problems where the state is a multi-dimensional stochastic process described by a system of SDEs (and ODEs). The main idea is to transform the problem into standard form by replacing the PDE describing the dynamics of the GS by a system of ODEs resulting from the semi-discretization w.r.t. spatial variables. However, the semi-discretization approach described in Chapter 3 led to a high-dimensional system of ODEs which makes the control problem non tractable. Therefore, one focus of this thesis was to reduce this high-dimensional system to low-dimensional system of ODEs in order make the control problem tractable. This led to a problem of model order reduction which we discussed in Chapter 4.

In order to formulate the model reduction problem, the underlying initial boundary value problem for the heat equation with a convection term has been discretized using finite difference schemes. In a first step we studied the semi-discretization with respect to spatial variables. For the resulting system of linear ODEs we proved that the system matrix is stable. In a second step the full space-time discretization has been considered. Here we derived explicit and implicit finite-difference schemes and investigated associated stability problems. In a large number of numerical experiments we have shown how these simulations can support the design and operation of a GS. Examples are the dependence of the charging and discharging efficiency on the topology and arrangement of heat exchanger PHX*s* and on the length of charging, discharging and waiting periods. In the third step we have considered the approximate description of the input-output behavior of a GS by a low-dimensional system of linear ODEs. Starting point was the semi-discretization of that PDE w.r.t. spatial variables which led to a high-dimensional

system of non-autonomous ODEs. The latter was approximated by an analogous LTI system. Reduced-order models in which the state dynamics is described by a low-dimensional system of linear ODEs were derived by the Lyapunov balanced truncation method. In our numerical experiments we considered aggregated characteristics describing the input-output behavior of the storage which are required for the operation of the GS within a residential heating system. The results showed that it is possible to obtain quite accurate approximations from reduced-order systems with only a few state variables. This allowed to treat the cost-optimal management of residential heating systems as a decision making problem under uncertainty which mathematically can be formulated as a tractable stochastic optimal control problem.

Another focus of this thesis was to investigate numerical solutions of this stochastic optimal control problem. We have considered the solution using dynamic programming and derived the associated HJB equation. However, the analytical solution of the HJB equation was not expected and due to the curse of dimensionality a numerical solution was not tractable. Therefore, the continuous-state MDP resulting from the discrete-time approximation of the continuous-time problem was considered and its solution has been investigated. Further, the MDP was transformed into a MDP for a controlled finite-states Markov chains and the associated transition probabilities have been investigated. The value function and the optimal strategy are determined by numerically solving the MDP using backward recursion. Based on Matlab implementation, extensive numerical experiments are carried out. The results showed some properties of the value function necessary to select the ansatz functions required in the alternative approximation methods to overcome the curse of dimensionality presented in Chapter 7.

8.2 Outlook

Based on the investigations in this thesis, the following extensions appear to be promising for future work.

Geothermal storage. One of the main goal of this thesis was to incorporate the geothermal into a residential heating system and study the interplay between the IS and GS. For the analysis and simulation purposes some restrictions have been made which should be relaxed to be closer to reality.
-*Model*. In this work we assume that the dynamics of the GS follows a 2D heat equation. This restriction provides more inside in the model and reduces the complexity of the stochastic optimal control problem when the interplay between the internal and GS is considered. The 2D model can be considered as cross section of a 3D model which represents the real world GS. A 3D model of a geothermal has already been considered by Bähr et al. [8] where the storage charging and discharging process was modeled by a simple source term and long-term simulation has been investigated. The results in [8] showed that the 3D model does not provide more significant information than 2D, specially for pure diffusion problem when charging is modeled by a source term. In our model we assume charging and discharging the geothermal by using heat exchanger pipes. We believe that a 3D model with heat exchanger pipes will provide more information in comparison to 2D model. However, this extension will be mathematically more involved and will require more work.
-*Topology of heat exchanger pipes*. In this thesis we considered a storage with several straight pipes. However, this restriction to straight pipes is only for simulation purpose since in the real world storage pipes are usually U-shaped or snake shaped. For more realistic setting, one can consider 2D or 3D models with snake pipes and study its short- or long-term behaviour. In this setting the finite difference scheme can no longer be used as a discretization method because of the bending of the pipes which cannot be captured appropriately. Therefore, to capture the behaviour at bending of the snake pipes a semi-discretization of the PDE has to done by other numerical schemes such as finite element methods.
-*Interfaces between the pipes and storage*. In the 2D model considered in this thesis we assume that the interfaces were straight lines. This consideration is already complicated enough and helps to gain more inside in this 2D model. However, in a more realistic 3D setting the pipes have cylindrical form and the interfaces which are the contact surfaces between the pipes and the storage medium are no longer straight lines. Therefore, the 2D interface condition does not capture the true 3D reality.

Residential heating system. We considered the heating system consisting of several heat production and consumption units in which we use one heat pump to connect the internal and GS for charging the IS. As we mentioned in the introduction one possible extension of the model will be to connect the IS to a district heating network. In this case, instead of firing fuel or using electricity to generate heat in times of high demand one takes thermal energy from the district heating system using the heat pump. One can also connect the GS to the solar collector directly to store overproduction directly in the GS without going through the IS. This will lead to another challenging mathematical control problem.

Model reduction. Balanced truncation model reduction method has been considered in this thesis to reduce the dimension of the dynamics of the GS. We have transformed the linear time-

varying system into an analogous LTI system since balanced truncation methods are known to perform well for LTI systems. However, some balanced truncation approaches for linear time-varying systems are available in the literature, see Sandberg and Rantzer [95] or Shokoohi et al. [100]. Since the system is linear time-varying through the time-dependent velocity which takes values in a finite set, we could consider a system as a linear switching system and apply balanced truncation for switching systems, see Gosea et al. [49], and Petreczky et al. [86, 85].
We noticed that the HJB equation associated to the control problem suffers from the curse of dimensionality. Assume that we are interested only on "good" (which may not be the best) controls u and the associated values of the performance criterion $J(t,x;u)$. Then to given and fixed control u we have to solve a high-dimensional linear PDE which is the HJB but without the pointwise optimization problem (*inf*) creating the non-linearity. We believe that it will be possible to apply the model order reduction techniques to reduce the dimension of the JHB without any further discretization.

Stochastic optimal control problem. The continuous-time optimal control problem has been transformed into a discrete-time problem with no discretization error by solving explicitly the SDEs and ODEs describing the state variables and the numerical solution has been investigated using backward recursion. However, this was possible only under Assumption 6.1.1 on the drift and diffusion coefficients of the exogenous state variables. This assumption can be relaxed and the the state equation can still be solved explicitly but this will be more involved and will require more effort.
The backward recursion used to numerically solving the discrete-time optimal control problem is suitable for problems whose state process has small dimension since the computation of the conditional expectation requires state discretization. We have been able to solve it in our computer implementation with state process of dimension 5, i.e. with reduced-order system of dimension 3. However, for more accuracy or if we are interested in more output quantities the dimension of the reduced-order system will be higher and the problem can no longer be handled using backward recursion. Therefore, one should resort to the approximation methods presented in Chapter 7.

APPENDIX A

Semi-Discretization Details

A.1 Block Matrices A_L and A_R

$$A_{L/R} = \left(\begin{array}{cccccccccccc}
\gamma^M_{DB} & \beta^M & & & & & & & & & & \\
\beta^M & \gamma^M_B & \beta^M & & & & & & & & & \\
 & \ddots & \ddots & \ddots & & & & & & & & \\
 & & \beta^M & \gamma^M_B & \beta^M & & & & & & & \\
 & & & \beta^M & \gamma^M_{IB} & \beta^M_I & & & & & & \\
\hline
 & & & & \beta^F_I & \gamma^F_{IL/R} & \beta^F & & & & & \\
 & & & & & \beta^F & \gamma^F_{L/R} & \beta^F & & & & \\
 & & & & & & \ddots & \ddots & \ddots & & & \\
 & & & & & & & \beta^F & \gamma^F_{L/R} & \beta^F & & \\
 & & & & & & & & \beta^F & \gamma^F_{IL/R} & \beta^F_I & \\
\hline
 & & & & & & & & & \beta^M_I & \gamma^M_{IB} & \beta^M \\
 & & & & & & & & & & \beta^M & \gamma^M_B \\
 & & & & & & & & & & & \ddots \\
 & & & & & & & & & & & \beta^M \\
 & & & & & & & & & & & \gamma^M_{UB}
\end{array}\right).$$

Lower Boundary
Medium
Lower interface
Fluid
Upper interface
Medium
Uppper Boundary

Table A.1: Sketch of the matrices A_L and A_R for the case of one pipe

This appendix gives the first and the last diagonal block matrices A_L and $A_R \in \mathbb{R}^{q\times q}$ of the matrix A given in (3.11). Its entries result from the discretization of boundary conditions at the left and right boundary. Both block matrices are tridiagonal and sketched for the case of only one pipe in Table A.1. The entries in the first and last row are related to the inner grid points next to the four corners of the domain and obtained by substituting homogeneous Neumann

condition (3.4) and Robin condition (3.5) into (3.1). For the grid points next to the lower left we obtain

$$\begin{aligned}\frac{d}{dt}Q_{11}(t) &= \alpha^M Q_{21}(t) + \alpha^M Q_{01}(t) + \beta^M Q_{12}(t) + \beta^M Q_{10}(t) + \gamma^M Q_{11}(t)\\ &= \alpha^M Q_{21}(t) + \beta^M Q_{12}(t) + \left(\alpha^M + \frac{\kappa^M}{\kappa^M + \lambda^G h_y}\beta^M + \gamma^M\right) Q_{11}(t) + \frac{\lambda^G h_y}{\kappa^M + \lambda^G h_y}\beta^M Q^G(t)\\ &= \alpha^M Q_{21}(t) + \beta^M Q_{12}(t) + \gamma^M_{DB} Q_{11}(t) + \frac{\lambda^G h_y}{\kappa^M + \lambda^G h_y}\beta^M Q^G(t),\end{aligned}$$

where $\gamma^M_{DB} = \alpha^M + \frac{\kappa^M}{\kappa^M + \lambda^G h_y}\beta^M + \gamma^M$. Recall, that $\alpha^M, \beta^M, \gamma^M$ are given (3.3).

Analogously, we derive for the lower right corner

$$\frac{d}{dt}Q_{N_x-1,1}(t) = \alpha^M Q_{N_x-2,1}(t) + \beta^M Q_{N_x,2}(t) + \gamma^M_{DB} Q_{N_x,1}(t) + \frac{\lambda^G h_y}{\kappa^M + \lambda^G h_y}\beta^M Q^G(t).$$

Note that the last terms on the r.h.s. of the above equations are contributions to the input term $B(t)g(t)$ given in (3.13) and (3.14). For the grid points next to the upper left and right corner we have to apply the homogeneous Neumann conditions (3.4) and obtain from (3.1)

$$\begin{aligned}\frac{d}{dt}Q_{1,N_y-1}(t) &= \alpha^M Q_{2,N_y-1}(t) + \alpha^M Q_{0,N_y-1}(t) + \beta^M Q_{1,N_y}(t) + \beta^M Q_{1,N_y-2}(t) + \gamma^M Q_{1,N_y-1}(t)\\ &= \alpha^M Q_{2,N_y-1}(t) + \beta^M Q_{1,N_y-2}(t) + (\alpha^M + \beta^M + \gamma^M) Q_{1,N_y-1}(t)\\ &= \alpha^M Q_{2,N_y-1}(t) + \beta^M Q_{1,N_y-2}(t) + \gamma^M_{UB} Q_{1,N_y-1}(t),\end{aligned}$$

where $\gamma^M_{UB} = \alpha^M + \beta^M + \gamma^M$. Analogously, we derive for the upper right corner

$$\frac{d}{dt}Q_{N_x-1,N_y-1}(t) = \alpha^M Q_{N_x-2,N_y-1}(t) + \beta^M Q_{N_x-1,N_y-2}(t) + \gamma^M_{UB} Q_{N_x-1,N_y-1}(t).$$

For "inner" grid points located next to insulated left and right boundary but not next to the upper and lower boundary or the interface we have to combine (3.1) with the homogeneous Neumann condition (3.4). This leads to the coefficient $\gamma^M_B = \gamma^M + \alpha^M$ on the main diagonal.

For the grid points next to the inlet boundary we apply Dirichlet condition during pumping and homogeneous Neumann condition if the pump is off, see (3.6). For j with $(0, j) \in \mathcal{N}_I^{\mathcal{C}}$ it holds

$$\begin{aligned}\frac{d}{dt}Q_{1j}(t) &= \alpha^{F+} Q_{2j}(t) + \alpha^{F-} Q_{0j}(t) + \beta^F Q_{1,j+1}(t) + \beta^F Q_{1,j-1}(t) + \gamma^F Q_{1j}(t)\\ &= \alpha^{F+} Q_{2j}(t) + \beta^F Q_{1j+1}(t) + \beta^F Q_{1,j-1}(t) + \begin{cases}\gamma^F Q_{1j}(t) + \alpha^{F-} Q^I(t) & \text{pump on}\\ (\gamma^F + \alpha^{F-}) Q_{1j}(t) & \text{pump off}\end{cases}\\ &= \alpha^{F+} Q_{2j}(t) + \beta^F Q_{1,j+1}(t) + \beta^F Q_{1,j-1}(t) + \gamma^F_L Q_{1j}(t) + b_{k1} Q^I(t),\end{aligned}$$

where $\gamma^F_L = \gamma^F_L(t) = \begin{cases}\gamma^F & \text{pump on,}\\ \gamma^F + \frac{a^F}{h_x^2} & \text{pump off,}\end{cases}$ and $B_{l1} = B_{l1}(t) = \begin{cases}\frac{a^F}{h_x^2} + \frac{\bar{v}_0}{h_x} & \text{pump on,}\\ 0 & \text{pump off,}\end{cases}$

with $l = \mathcal{K}(1, j)$. We note that $\alpha^{F\pm}, \beta^F, \gamma^F$ are given in (3.2) and point out that the term $b_{k1} Q^I(t)$

contributes to the input term $B(t)g(t)$.

At the outlet boundary we have homogeneous Neumann condition and for the grid points next to the outlet we obtain from the discretized boundary condition (3.4) for j with $(N_x, j) \in \mathcal{N}_O^{\mathcal{C}}$ it holds

$$\begin{aligned}\frac{d}{dt}Q_{N_x-1,j}(t) &= \alpha^{F+}Q_{N_x,j}(t) + \alpha^{F-}Q_{N_x-2,j}(t) + \beta^F Q_{N_x-1,j+1}(t) + \beta^F Q_{N_x-1,j-1}(t) + \gamma^F Q_{N_x-1,j}(t)\\ &= \alpha^{F-}Q_{N_x-2,j}(t) + \beta^F Q_{N_x-1,j+1}(t) + \beta^F Q_{N_x-1,j-1}(t) + \gamma_R^F Q_{N_x-1,j}(t),\end{aligned}$$

where $\gamma_R^F = \gamma^F + \alpha^{F+}$ and $\alpha^{F\pm}, \beta^F, \gamma^F$ are given in (3.2).

Finally, for the grid points next to the interface we obtain by an analogous procedure as described in Subsec. 3.1.2 the coefficients

$$\gamma_{IB}^M = \gamma_B^M + \psi^M \beta^M, \quad \gamma_{IL}^F = \gamma_{IL}^F(t) = \gamma_L^F(t) + \psi^F \beta^F, \quad \gamma_{IR}^F = \gamma_R^F + \psi^F \beta^F,$$

where ψ^M and ψ^F are given (3.7). Recall that the off-diagonal coefficients β_I^M, β_I^F are given in (3.8), (3.9).

A.2 Proof of Lemma 3.2.2

Proof. **First assertion.** Table A.2 shows that the diagonal entries of the matrices A^k, $k = 1, \ldots, N_\tau$ are all negative. Thus, we have for all $i = 1, \ldots, n$

$$J_i(G^k) = |G_{ii}^k| - \sum_{j=1, j\neq i} |G_{ij}^k| = 1 + \tau\theta(|A_{ii}^k| - \sum_{j=1, j\neq i} |A_{ij}^k|) = 1 + \tau\theta J_i(A^k) \geq 1,$$

since by Lemma 3.1.8 the matrices A^k are diagonal dominant and it holds $J_i(A^k) \geq 0$. Therefore, the matrices $G^k = \mathbb{I}_n - \tau\theta A^k$, $k = 1, \ldots, N_\tau$ are strictly diagonal dominant. Lemma 3.1.5 implies that G^k is invertible and $\|(G^k)^{-1}\|_\infty \leq 1/J(G^k) \leq 1$. For $\theta = 0$ it holds $G^k = \mathbb{I}_n$, hence $\|(G^k)^{-1}\|_\infty = \|\mathbb{I}_n\|_\infty = 1$ and the above inequality holds with equality.

Second assertion. We recall the definition of H^k given in (3.17) which reads as $H^k = \mathbb{I}_n + \tau(1-\theta)A^k$. For $\theta = 1$, we have $H^k = \mathbb{I}_n$, thus for all $\tau > 0$ it holds $\left\|H^k\right\|_\infty = 1$ which proves the claim for $\theta = 1$.

Now, let $\theta \in [0, 1)]$. We recall that $A^k = A(k\tau)$ takes only the values A^P and A^N. Thus it is sufficient to show that the claim holds for H^P and H^N where $H^{P/N} = \mathbb{I}_n + \tau(1-\theta)A^{P/N}$. It holds

$$\left\|H^P\right\|_\infty = \max_{1\leq i\leq n} \left\{S_i(H^P)\right\}, \text{ with } S_i(H^P) = |1 + \tau(1-\theta)A_{ii}^P| + \tau(1-\theta) \sum_{j=1, j\neq i}^{n} |A_{ij}^P|.$$

Using the fact that all diagonal entries of the matrix A^P are negative, we have for $\tau_i^P = \frac{1}{(1-\theta)|A_{ii}^P|}$, $i = 1, \ldots, n$,

$$|1 + \tau(1-\theta)A_{ii}^P| = \begin{cases} 1 - \tau(1-\theta)|A_{ii}^P|, & \text{for } \tau \leq \tau_i^P \\ \tau(1-\theta)|A_{ii}^P| - 1, & \text{for } \tau > \tau_i^P. \end{cases}$$

This implies that for $i = 1, \dots, n$, we have

$$S_i(H^P) = \begin{cases} 1 - \tau(1-\theta)\Big[|A^P_{ii}| - \sum\limits_{j=1, j\neq i}^{n} |A^P_{ij}|\Big] = 1 - \tau(1-\theta)R_i(A^P), & \text{for } \tau \leq \tau^P_i, \\ \tau(1-\theta)\Big[|A^P_{ii}| + \sum\limits_{j=1, j\neq i}^{n} |A^P_{ij}|\Big] - 1 = -1 + \tau(1-\theta)S_i(A^P), & \text{for } \tau > \tau^p_i. \end{cases}$$

Since A^P is weakly diagonal dominant, we distinguish the two cases $J_i(A^P) > 0$ and $J_i(A^P) = 0$. For $J_i(A^P) > 0$, the sum $S_i(H^P)$ is strictly decreasing in τ on $[0, \tau^P_i]$ and strictly increasing in τ on $(\tau^P_i, +\infty)$ and it holds

$$S_i(H^P) \leq 1 \;\text{ for }\; \tau \leq \overline{\tau}^P_i := \frac{2}{(1-\theta)S_i(A^P)} \;\text{ and }\; S_i(H^P) > 1 \;\text{ for }\; \tau > \overline{\tau}^P_i.$$

For $J_i(A^P) = 0$, we have $S_i(A^P) = 2|A^P_{ii}|$. It holds $S_i(H^P) = 1$ for $\tau \in [0, \overline{\tau}^P_i]$ while $S_i(H^P)$ is strictly increasing in τ on $(\overline{\tau}^P_i, +\infty)$, hence $S_i(H^P) > 1$ for $\tau > \overline{\tau}^P_i$.
Summarizing we obtain

$$\left\|H^P\right\|_\infty = \max_{1\leq i\leq n} S_i(H^P) = 1 \;\text{ for }\; \tau \leq \overline{\tau}^P = \min_{1\leq i\leq n} \overline{\tau}^P_i = \frac{2}{(1-\theta)\max\limits_{1\leq i\leq n} S_i(A^P)} = \frac{2}{(1-\theta)\|A^P\|_\infty},$$

and $\left\|H^P\right\|_\infty > 1$ for $\tau > \overline{\tau}^P$. For $A = A^N$ the proof is analogous. Thus, we have

$$\left\|H^k\right\|_\infty \leq 1 \quad \text{for} \quad \tau \leq \min\{\overline{\tau}^P, \overline{\tau}^N\} = \frac{2}{(1-\theta)\max\left\{\left\|A^P\right\|_\infty, \left\|A^N\right\|_\infty\right\}}.$$

Finally, Lemma 3.1.12 shows that $\max\left\{\left\|A^P\right\|_\infty, \left\|A^N\right\|_\infty\right\} = 4\max\{a^F, a^M\}\left(\frac{1}{h_x^2} + \frac{1}{h_y^2}\right) + \frac{2\overline{v}_0}{h_x} = 2\eta$ which proves the claim.

Third assertion. From the definition of F^k given in (3.17) it follows that for $k = 0, \dots, N_\tau - 1$

$$\begin{aligned} \left\|F^k\right\|_\infty &= \left\|\theta B^{k+1}g^{k+1} + (1-\theta)B^k g^k\right\|_\infty \leq \theta\left\|B^{k+1}\right\|_\infty\left\|g^{k+1}\right\|_\infty + (1-\theta)\left\|B^k\right\|_\infty\left\|g^k\right\|_\infty \\ &\leq (\theta + 1 - \theta)C_B \max_{j=k,k+1}\left\|g^j\right\|_\infty \leq C_B \max_{0\leq j\leq k+1}\left\|g^j\right\|_\infty. \end{aligned}$$

where we have used that $B^k = B(k\tau)$ takes only the values B^P and B^N. □

l	Gershgorin circles: centres $C_{i_l} = A_{i_l i_l}$ (diagonal entries)		Gershgorin circles: radii $R_{i_l}(A) = \sum_{j=1, j\neq i_l}^{n} \lvert A_{i_l,j}\rvert$	Differences $J_{i_l}(A) = \lvert A_{i_l i_l}\rvert - \sum_{j=1, j\neq i_l}^{n} \lvert A_{i_l,j}\rvert$	Row sums $S_{i_l}(A) = \sum_{j=1}^{n} \lvert A_{i_l,j}\rvert$
1	γ_{DB}^m	$-a^M\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right) - \left(\frac{\lambda^G h_y}{\kappa^M+\lambda^G h_y}\right)\frac{a^M}{h_y^2}$	$a^M\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right)$	$\left(\frac{\lambda^G h_y}{\kappa^M+\lambda^G h_y}\right)\frac{a^M}{h_y^2}$	$2a^M\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right) + \left(\frac{\lambda^G h_y}{\kappa^M+\lambda^G h_y}\right)\frac{a^M}{h_y^2}$
2	γ_{UB}^m	$-a^M\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right)$	$a^M\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right)$	0	$2a^M\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right)$
3	γ_{D}^m	$-a^M\left(\frac{2}{h_x^2}+\frac{1}{h_y^2}\right) - \left(\frac{\lambda^G h_y}{\kappa^M+\lambda^G h_y}\right)\frac{a^M}{h_y^2}$	$a^M\left(\frac{2}{h_x^2}+\frac{1}{h_y^2}\right)$	$\left(\frac{\lambda^G h_y}{\kappa^M+\lambda^G h_y}\right)\frac{a^M}{h_y^2}$	$2a^M\left(\frac{2}{h_x^2}+\frac{1}{h_y^2}\right) + \left(\frac{\lambda^G h_y}{\kappa^M+\lambda^G h_y}\right)\frac{a^M}{h_y^2}$
4	γ_{U}^m	$-a^M\left(\frac{2}{h_x^2}+\frac{1}{h_y^2}\right)$	$a^M\left(\frac{2}{h_x^2}+\frac{1}{h_y^2}\right)$	0	$2a^M\left(\frac{2}{h_x^2}+\frac{1}{h_y^2}\right)$
5	γ_{B}^m	$-a^M\left(\frac{1}{h_x^2}+\frac{2}{h_y^2}\right)$	$a^M\left(\frac{1}{h_x^2}+\frac{2}{h_y^2}\right)$	0	$2a^M\left(\frac{1}{h_x^2}+\frac{2}{h_y^2}\right)$
6	γ^M	$-2a^M\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right)$	$2a^M\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right)$	0	$4a^M\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right)$
7	γ^F	$-2a^F\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right) - \frac{v_0(t)}{h_x}$	$2a^F\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right) + \frac{v_0(t)}{h_x}$	0	$4a^F\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right) + \frac{2v_0(t)}{h_x}$
8	γ_L^f	$\begin{cases} -2a^F\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right) - \frac{\overline{v}_0}{h_x}, & A = A^P \\ -a^F\left(\frac{1}{h_x^2}+\frac{2}{h_y^2}\right), & A = A^N \end{cases}$	$a^F\left(\frac{1}{h_x^2}+\frac{2}{h_y^2}\right)$	$\begin{cases} \frac{a^F}{h_x^2} + \frac{\overline{v}_0}{h_x}, & A = A^P \\ 0, & A = A^N \end{cases}$	$\begin{cases} a^F\left(\frac{3}{h_x^2}+\frac{4}{h_y^2}\right) + \frac{\overline{v}_0}{h_x}, & A = A^P \\ 2a^F\left(\frac{1}{h_x^2}+\frac{2}{h_y^2}\right), & A = A^N \end{cases}$
9	γ_R^f	$-a^F\left(\frac{1}{h_x^2}+\frac{2}{h_y^2}\right) - \frac{v_0(t)}{h_x}$	$a^F\left(\frac{1}{h_x^2}+\frac{2}{h_y^2}\right) + \frac{v_0(t)}{h_x}$	0	$2a^F\left(\frac{1}{h_x^2}+\frac{2}{h_y^2}\right) + \frac{2v_0(t)}{h_x}$
10	γ_I^M	$-a^M\left(\frac{2}{h_x^2}+\frac{1}{h_y^2}\right) - \left(\frac{\kappa^F}{\kappa^M+\kappa^F}\right)\frac{a^M}{h_y^2}$	$\frac{2a^M}{h_x^2} + \left(1+\frac{\kappa^F}{\kappa^M+\kappa^F}\right)\frac{a^M}{h_y^2}$	0	$2a^M\left(\frac{2}{h_x^2}+\frac{1}{h_y^2}\right) + \left(\frac{2\kappa^F}{\kappa^M+\kappa^F}\right)\frac{a^M}{h_y^2}$
11	γ_{IB}^M	$-a^M\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right) - \left(\frac{\kappa^F}{\kappa^M+\kappa^F}\right)\frac{a^M}{h_y^2}$	$\frac{a^M}{h_x^2} + \left(1+\frac{\kappa^F}{\kappa^M+\kappa^F}\right)\frac{a^M}{h_y^2}$	0	$2a^M\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right) + \left(\frac{2\kappa^F}{\kappa^M+\kappa^F}\right)\frac{a^M}{h_y^2}$
12	γ_I^F	$-a^F\left(\frac{2}{h_x^2}+\frac{1}{h_y^2}\right) - \left(\frac{\kappa^M}{\kappa^M+\kappa^F}\right)\frac{a^F}{h_y^2} - \frac{v_0(t)}{h_x}$	$\frac{2a^F}{h_x^2} + \left(1+\frac{\kappa^M}{\kappa^M+\kappa^F}\right)\frac{a^F}{h_y^2} + \frac{v_0(t)}{h_x}$	0	$2a^F\left(\frac{2}{h_x^2}+\frac{1}{h_y^2}\right) + \left(\frac{2\kappa^M}{\kappa^M+\kappa^F}\right)\frac{a^F}{h_y^2} + \frac{2v_0(t)}{h_x}$
13	γ_{IL}^F	$\begin{cases} -a^F\left(\frac{2}{h_x^2}+\frac{1}{h_y^2}\right) - \left(\frac{\kappa^M}{\kappa^M+\kappa^F}\right)\frac{a^F}{h_y^2} - \frac{\overline{v}_0}{h_x}, & A = A^P \\ -a^F\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right) - \left(\frac{\kappa^M}{\kappa^M+\kappa^F}\right)\frac{a^F}{h_y^2}, & A = A^N \end{cases}$	$\frac{a^F}{h_x^2} + \left(1+\frac{\kappa^M}{\kappa^M+\kappa^F}\right)\frac{a^F}{h_y^2}$	$\begin{cases} \frac{a^F}{h_x^2} + \frac{\overline{v}_0}{h_x}, & A = A^P \\ 0, & A = A^N \end{cases}$	$\begin{cases} a^F\left(\frac{3}{h_x^2}+\frac{2}{h_y^2}\right) + \left(\frac{2\kappa^M}{\kappa^M+\kappa^F}\right)\frac{a^F}{h_y^2} + \frac{\overline{v}_0}{h_x}, & A = A^P \\ 2a^F\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right) + \left(\frac{2\kappa^M}{\kappa^M+\kappa^F}\right)\frac{a^F}{h_y^2}, & A = A^N \end{cases}$
14	γ_{IR}^F	$-a^F\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right) - \left(\frac{\kappa^M}{\kappa^M+\kappa^F}\right)\frac{a^F}{h_y^2} - \frac{v_0(t)}{h_x}$	$\frac{a^F}{h_x^2} + \left(1+\frac{\kappa^M}{\kappa^M+\kappa^F}\right)\frac{a^F}{h_y^2} + \frac{v_0(t)}{h_x}$	0	$2a^F\left(\frac{1}{h_x^2}+\frac{1}{h_y^2}\right) + \left(\frac{2\kappa^M}{\kappa^M+\kappa^F}\right)\frac{a^F}{h_y^2} + \frac{2v_0(t)}{h_x}$

Table A.2: Diagonal entries $A_{i_l i_l}$ (centres of Gershgorin circles), radii of Gershgorin circles R_{i_l}, differences J_{i_l} and row sums S_{i_l} of matrices A^P and A^N, $l = 1, \ldots, 14$

A.3 Derivation of Quadrature Formula

Rewriting the double integral as two iterated single integrals and applying trapezoidal rule to the outer integral we obtain (suppressing the time variable t)

$$J = \iint_{\mathcal{B}} Q(x,y)dxdy = \int_{x_{\underline{i}}}^{x_{\overline{i}}} \left(\int_{y_{\underline{j}}}^{y_{\overline{j}}} Q(x,y)\,dy \right) dx$$
$$\approx \int_{x_{\underline{i}}}^{x_{\overline{i}}} h_y \left(\frac{1}{2}Q(x,y_{\underline{j}}) + \sum_{j=\underline{j}+1}^{\overline{j}-1} Q(x,y_j) + \frac{1}{2}Q(x,y_{\overline{j}}) \right) dx.$$

Approximating the inner integrals again by trapezoidal rule we get

$$\int_{x_{\underline{i}}}^{x_{\overline{i}}} Q(x,y_j)\,dx \approx h_x \left(\frac{1}{2}Q(x_{\underline{i}},y_j) + \sum_{i=\underline{i}+1}^{\overline{i}-1} Q(x_i,y_j) + \frac{1}{2}Q(x_{\overline{i}},y_j) \right), \quad j = \underline{j},\ldots,\overline{j}.$$

Substituting into the above expression for J yields

$$J \approx h_x h_y \left(\frac{1}{4}\left[Q(x_{\underline{i}},y_{\underline{j}}) + Q(x_{\overline{i}},y_{\underline{j}}) + Q(x_{\underline{i}},y_{\overline{j}}) + Q(x_{\overline{i}},y_{\overline{j}}) \right] \right.$$
$$\left. + \frac{1}{2}\left[\sum_{i=\underline{i}+1}^{\overline{i}-1} \left[Q(x_i,y_{\underline{j}}) + Q(x_i,y_{\overline{j}}) \right] + \sum_{j=\underline{j}+1}^{\overline{j}-1} \left[Q(x_{\underline{i}},y_j) + Q(x_{\overline{i}},y_j) \right] \right] + \sum_{i=\underline{i}+1}^{\overline{i}-1} \sum_{j=\underline{j}+1}^{\overline{j}-1} Q(x_i,y_j) \right)$$
$$= h_x h_y \left(\frac{1}{4}\left[Q_{\underline{i}\underline{j}}(t) + Q_{\overline{i}\underline{j}}(t) + Q_{\underline{i}\overline{j}}(t) + Q_{\overline{i}\overline{j}}(t) \right] \right.$$
$$\left. + \frac{1}{2}\left[\sum_{i=\underline{i}+1}^{\overline{i}-1} \left[Q_{i\underline{j}} + Q_{i\overline{j}} \right] + \sum_{j=\underline{j}+1}^{\overline{j}-1} \left[Q_{\underline{i}j} + Q_{\overline{i}j} \right] \right] + \sum_{i=\underline{i}+1}^{\overline{i}-1} \sum_{j=\underline{j}+1}^{\overline{j}-1} Q_{ij} \right).$$

Since the area of the rectangle $\mathcal{B}$ is given by $(\overline{i}-\underline{i})(\overline{j}-\underline{j})h_x h_y$ the average temperature $\overline{Q}^{\mathcal{B}}$ can be approximated by

$$\overline{Q}^{\mathcal{B}} = \frac{1}{|\mathcal{B}|} \iint_{\mathcal{B}} Q(t,x,y)dxdy \approx \sum_{(i,j)\in\mathcal{N}_{\mathcal{B}}} \mu_{ij} Q_{ij}$$

with the coefficients μ_{ij} given in (3.23).

APPENDIX B

Model Reduction Details

B.1 Proof of Theorem 4.2.9

Proof. Following the results in Antouslas [5, Proposition 4.27], we can derive the following. Assume that a system matrix A is stable, then we have:

$$\begin{aligned} A\mathcal{G}_C + \mathcal{G}_C A^\top &= \int_0^\infty \left(A e^{At} BB^\top e^{A^\top t} + e^{At} BB^\top e^{A^\top t} A^\top \right) dt = \int_0^\infty \frac{d}{dt} \left(e^{At} BB^\top e^{A^\top t} \right) dt \\ &= \lim_{T\to\infty} e^{At} BB^\top e^{A^\top t} \Big|_0^T = \lim_{T\to\infty} e^{AT} BB^\top e^{A^\top T} - BB^\top = -BB^\top. \\ \mathcal{G}_O A + A^\top \mathcal{G}_O &= \int_0^\infty \left(e^{A^\top t} C^\top C e^{At} A + A^\top e^{A^\top t} C^\top C e^{At} \right) dt = \int_0^\infty \frac{d}{dt} \left(e^{A^\top t} C^\top C e^{At} \right) dt \\ &= \lim_{T\to\infty} e^{A^\top t} C^\top C e^{At} \Big|_0^T = \lim_{T\to\infty} e^{A^\top T} C^\top C e^{AT} - C^\top C = -C^\top C. \end{aligned}$$

□

B.2 Proof of Lemma 4.2.11

Proof. let $\overline{\mathcal{G}}_C$ and $\overline{\mathcal{G}}_O$ be the controllability and observability Gramians of the transformed system, respectively. Then $\overline{\mathcal{G}}_C$ satisfies the following Lyapunov equation:

$$0 = \overline{A}\,\overline{\mathcal{G}}_C + \overline{\mathcal{G}}_C \overline{A}^\top + \overline{B}\,\overline{B}^\top = \mathcal{T}A\mathcal{T}^{-1}\overline{\mathcal{G}}_C + \overline{\mathcal{G}}_C(\mathcal{T}A\mathcal{T}^{-1})^\top + \mathcal{T}B(\mathcal{T}B)^\top.$$

Multiplying by $\mathcal{T}^{-1}$ from left and by $\mathcal{T}^{-\top}$ from right gives

$$0 = A(\mathcal{T}^{-1}\overline{\mathcal{G}}_C\mathcal{T}^{-\top}) + (\mathcal{T}^{-1}\overline{\mathcal{G}}_C\mathcal{T}^{-\top})A^\top + BB^\top.$$

Comparing with the Lyapunov equation for the Gramian $\mathcal{G}_C$ of the original system which reads as $0 = A\mathcal{G}_C + \mathcal{G}_C A^\top + BB^\top$ gives $\mathcal{G}_C = \mathcal{T}^{-1}\overline{\mathcal{G}}_C\mathcal{T}^{-\top}$ and finally $\overline{\mathcal{G}}_C = \mathcal{T}\mathcal{G}_C\mathcal{T}^\top$. Similar reasoning gives $\overline{\mathcal{G}}_O = \mathcal{T}^{-\top}\mathcal{G}_O\mathcal{T}^{-1}$.
Substituting into the product of the transformed Gramians yields $\overline{\mathcal{G}}_C\overline{\mathcal{G}}_O = \mathcal{T}\mathcal{G}_C\mathcal{G}_O\mathcal{T}^{-1}$. □

B.3 Proof of Theorem 4.2.13

Proof. We have to prove that the system is balanced under the transformation $\mathcal{T} = \Sigma^{\frac{1}{2}}K^\top U^{-1}$. For the Gramians of the transformed system, we obtain

$$\begin{aligned}
\overline{\mathcal{G}}_C = \mathcal{T}\mathcal{G}_C\mathcal{T}^\top &= \Sigma^{\frac{1}{2}}K^\top U^{-1}\mathcal{G}_C U^{-\top}K\Sigma^{\frac{1}{2}} = \Sigma^{\frac{1}{2}}K^\top U^{-1}UU^\top U^{-\top}K\Sigma^{\frac{1}{2}} \\
&= \Sigma^{\frac{1}{2}}K^\top K\Sigma^{\frac{1}{2}} = \Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}} = \Sigma. \\
\overline{\mathcal{G}}_O = \mathcal{T}^{-\top}\mathcal{G}_O\mathcal{T}^{-1} &= \Sigma^{-\frac{1}{2}}K^\top U^\top\mathcal{G}_O UK\Sigma^{-\frac{1}{2}} = \Sigma^{-\frac{1}{2}}K^\top K\Sigma^2K^\top K\Sigma^{-\frac{1}{2}} \\
&= \Sigma^{-\frac{1}{2}}\Sigma^2\Sigma^{-\frac{1}{2}} = \Sigma.
\end{aligned}$$

We used $\mathcal{G}_C = UU^\top$, $U^\top U^{-\top} = I_n = U^{-1}U$, $U^\top\mathcal{G}_O U = K\Sigma^2K^\top$ and $K^\top K = I_n$. □

APPENDIX C

Optimal Control Details

C.1 Proof of Theorem 5.3.2

Proof. Let $\tau_h = t+h$ and a constant control $u(s) = \nu$, for some arbitrary ν in $\mathcal{U}(t,x)$. The dynamic programming principle yields

$$V(t,x) \leq \mathbb{E}_{t,x}\left[\int_t^{t+h} \Psi(s,X^u(s),\nu)ds + V(t+h,X^u(t+h))\right]. \tag{C.1}$$

Assuming that V is smooth enough, applying the Itô's formula on $V(t+h,X^u(t+h))$ between t and $t+h$ and substituting into (C.1) yields

$$0 \leq \mathbb{E}_{t,x}\left[\int_t^{t+h} \Psi(s,X^u(s),\nu) + \left(\frac{\partial}{\partial t}V + \mathcal{L}^\nu V\right)(s,X^u(s))ds\right],$$

where $\mathcal{L}^\nu$ is the generator associated to the diffusion process for the constant control ν and given by equation (5.12).
Dividing by h and sending h to 0, this yields by the mean-value theorem

$$0 \leq \Psi(t,x,\nu) + \frac{\partial}{\partial t}V(t,x) + \overline{\mathcal{L}}^\nu V(t,x) + \widehat{\mathcal{L}}V(s,x).$$

Since the control ν in $\mathcal{U}(t,x)$ was chosen arbitrary, we obtain the inequality

$$\frac{\partial}{\partial t}V(t,x) + \widehat{\mathcal{L}}\,V(t,x) + \inf_{\nu\in\mathcal{U}(t,x)}\left\{\overline{\mathcal{L}}^\nu V(t,x) + \Psi(t,x,\nu)\right\} \geq 0. \tag{C.2}$$

On the other hand, assume that $u^\star$ is an optimal control and $X^\star = X^{u^\star}$ the associated solution to the SDE (5.5) starting from x a time t. Then, the stronger version of the DPP (5.11) yields

$$V(t,x) = \mathbb{E}_{t,x}\left[\int_t^{\tau_h} \Psi(s,X^\star(s),u^\star(s))ds + V(\tau_h,X^\star(\tau_h))\right],$$

Using similar argument as above, we obtain

$$\frac{\partial}{\partial t}V(t,x)+\widehat{\mathcal{L}}\,V(t,x)+\overline{\mathcal{L}}^{u^\star}V(t,x)+\Psi(t,x,u^\star(s))=0. \tag{C.3}$$

Combining (C.2) and (C.3), gives

$$\frac{\partial}{\partial t}V(t,x)+\widehat{\mathcal{L}}\,V(t,x)+\inf_{v\in\mathcal{U}(t,x)}\left\{\overline{\mathcal{L}}^{v}V(t,x)+\Psi(t,x,v)\right\}=0,\quad (t,x)\in[0,T)\times\mathcal{X},$$

if the above infimum in v is finite. From the definition of the reward function (5.8) considered at the time horizon T, we immediately obtain the terminal condition associated to this PDE, i.e.

$$V(t,x)=\psi(x).$$

□

C.2 Time-Discretization Details

C.2.1 Proof of Lemma 6.1.3

The flowing remark plays a crucial role for the poof of the above Lemma 6.1.3 given below.

Remark C.2.1 Let A be a diagonalizable matrix, then there exists a diagonal matrix $D=diag(\lambda_1,\dots,\lambda_l)$ and a regular matrix V such that $A=VDV^{-1}$ is the eigenvalue decomposition of A. Then it holds

$$A^m=(VDV^{-1})(VDV^{-1})\dots(VDV^{-1})=VD^mV^{-1}.$$

and

$$\begin{aligned}e^{At}&=\sum_{m=0}^{\infty}\frac{1}{m!}A^mt^m=V\left(\sum_{m=0}^{\infty}\frac{1}{m!}D^mt^m\right)V^{-1}=Ve^{Dt}V^{-1}\\ \int e^{As}ds&=V\int e^{Ds}dsV^{-1}=VD^{-1}e^{Ds}V^{-1}+c=(VD^{-1}V^{-1})(Ve^{Ds}V^{-1})+c\\ &=A^{-1}e^{As}+c=e^{As}A^{-1}+c.\end{aligned}$$

Note that the matrices A^{-1} and e^{As} commute in the product sense since $D^{-1}e^{Ds}=e^{Ds}D^{-1}$.

Proof. Let $t\in[t_n,t_{n+1})$. Under Assumption 6.1.2, the closed-form solution of the ODE (6.1) on the time interval $[t_n,t_{n+1})$ given $\widetilde{Y}(t_n)=\widetilde{Y}_n$ is given as

$$\widetilde{Y}(t)=e^{\widetilde{A}(t-t_n)}\left(\widetilde{Y}_n+\int_{t_n}^{t}e^{-\widetilde{A}(s-t_n)}\widetilde{B}g_n^v ds\right)=e^{\widetilde{A}(t-t_n)}\widetilde{Y}_n+f_n,$$

where

$$f_n=e^{\widetilde{A}(t-t_n)}\int_{t_n}^{t}e^{-\widetilde{A}(s-t_n)}ds\widetilde{B}g_n^v=e^{\widetilde{A}(t-t_n)}\left[-e^{-\widetilde{A}(s-t_n)}\right]_{t_n}^{t}\widetilde{A}^{-1}\widetilde{B}g_n^v,$$

$$= e^{\widetilde{A}(t-t_n)}(-e^{-\widetilde{A}(t-t_n)} + \mathbb{I}_\ell)\widetilde{A}^{-1}\widetilde{B}g_n^\nu = (e^{\widetilde{A}(t-t_n)} - \mathbb{I}_\ell)\widetilde{A}^{-1}\widetilde{B}g_n^\nu.$$

For $\widetilde{Y}_{n+1} = \widetilde{Y}(t_{n+1})$, we immediately have the discrete-time dynamics of the GS (6.3). □

C.2.2 Proof of Lemma 6.1.4

Proof. For all $t \in [t_n, t_{n+1})$ we apply the Itô formula to the function $f(r,t) = re^{\beta_R t}$. We have $f_t = \beta_R r e^{\beta_R t}$, $f_r = e^{\beta_R t}$, $f_{rr} = 0$. Then,

$$\begin{aligned}
df(R(t),t) &= f_t(R(t),t)dt + f_r(R(t),t)dR(t) + \frac{1}{2}f_{rr}(R(t),t)(dR(t))^2,\\
\int_{t_n}^t df(R(s),s) &= \int_{t_n}^t \beta_R R(s)e^{\beta_R s}ds + \int_{t_n}^t e^{\beta_R s}dR(s),\\
R(t)e^{\beta_R t} - R_n e^{\beta_R t_n} &= \int_{t_n}^t \beta_R R(s)e^{\beta_R s}ds + \int_{t_n}^t e^{\beta_R s}\beta_R(\mu_{R,n} - R(s))ds + \int_{t_n}^t e^{\beta_R s}\sigma_{R,n}dW_R(s).
\end{aligned}$$

This implies that

$$\begin{aligned}
R(t) &= e^{-\beta_R t}\left(R_n e^{\beta_R t_n} + \mu_{R,n}(e^{\beta_R t} - e^{\beta_R t_n})\right) + e^{-\beta_R t}\int_{t_n}^t e^{\beta_R s}\sigma_{R,n}dW_R(s)\\
&= R_n e^{-\beta_R(t-t_n)} + \mu_{R,n}(1 - e^{-\beta_R(t-t_n)}) + \sigma_{R,n}\int_{t_n}^t e^{-\beta(t-s)}dW_R(s). \qquad \text{(C.4)}
\end{aligned}$$

□

C.2.3 Proof of Lemma 6.1.6

We recall that the continuous time dynamics of the temperature in the IS can be written for $u(t) \in \overline{\mathcal{U}}$ as

$$dP(t) = (\psi_p(R(t), \widetilde{Y}(t), u(t))) - \gamma(P(t) - P_{amb})dt, \quad P(0) = p_0, \qquad \text{(C.5)}$$

where $\gamma = \frac{\kappa_h A_h}{m^P c_P^F}$ is a constant and $R(t)$ is the residual demand given by equation (6.5). The function ψ_p if given by

$$\psi_P(r,y,\nu) = \begin{cases} -k_P r + \kappa_F & \nu = u^F \\ -k_P r + \kappa_C(P_{in} - P_{out}) & \nu = u^C \\ -k_P r & \nu = u^W \\ -k_P r - \kappa_D(Q_C^I - C^O y) & \nu = u^D \\ 0 & \nu = u^O \end{cases}$$

where, $k_P = \frac{1}{m^P c_P^F}$, $\kappa_D = \frac{k_D}{m^P c_P^F}$, $\kappa_C = \frac{k_C}{m^P c_P^F}$ and $\kappa_F = \frac{k_F}{m^P c_P^F}$.

Proof. Let $t \in [t_n, t_{n+1})$. For $u_n = u^O$, and $P(t_n) = P_n$, a closed form solution of the ODE (C.5) on the interval $[t_n, t_{n+1})$, is given by

$$P(t) = e^{-\gamma(t-t_n)}(p + P_{amb}(e^{\gamma(t-t_n)} - 1)).$$

For $u_n \neq u^O$ and $P(t_n) = P_n$, a closed form solution of the equation (C.5) is given by

$$\begin{aligned} P(t) &= e^{-\gamma(t-t_n)}\left\{P(t_n) + \int_{t_n}^{t} e^{\gamma(s-t_n)}\zeta(\widetilde{Y}(s), u_n)ds - k_P\int_{t_n}^{t} e^{\gamma(s-t_n)}R(s)ds\right\} \\ &= e^{-\gamma(t-t_n)}\left\{P_n + \int_{t_n}^{t} e^{\gamma(s-t_n)}\zeta(\widetilde{Y}(s), u_n)ds\right\} - k_P\int_{t_n}^{t} e^{-\gamma(t-s)}R(s)ds \\ &= e^{-\gamma(t-t_n)}P_n + f(u_n, \widetilde{Y}(t)) - k_P\int_{t_n}^{t} e^{-\gamma(t-s)}R(s)ds \end{aligned} \tag{C.6}$$

where

$$f(u_n, \widetilde{Y}(t)) = \int_{t_n}^{t} e^{-\gamma(t-s)}\zeta(u_n, \widetilde{Y}(s))ds$$

and

$$\zeta(\nu, y) = \begin{cases} \gamma P_{amb} + \kappa_F & \nu = u^F, \\ \gamma P_{amb} + \kappa_C(P_{in} - P_{out}) & \nu = u^C, \\ \gamma P_{amb} & \nu = u^W, \\ \gamma P_{amb} - \kappa_D(Q_C^I - C^O y) & \nu = u^D. \end{cases}$$

The process $P(t)$ given in (C.6) is an integrated process since it is the integral functional of a Gaussian process $R(t)$ given by (C.4). Under Assumption 6.1.1, substituting $R(s)$ given by (C.4) in (C.6) gives

$$\begin{aligned} P(t) &= e^{-\gamma(t-t_n)}P_n + f(u_n) \\ &\quad - k_P\int_{t_n}^{t} e^{-\gamma(t-s)}\left[re^{-\beta(s-t_n)} + \mu_{R,n}(1 - e^{-\beta(s-t_n)}) + \int_{t_n}^{s}\sigma_{R,n}e^{-\beta(s-u)}dW_R(u)\right]ds\Bigg\} \\ &= e^{-\gamma(t-t_n)}P_n - \frac{k_P\mu_{R,n}}{\gamma}(1 - e^{-\gamma(t-t_n)}) + \frac{k_P}{\beta_R - \gamma}(\mu_{R,n} - R_n)\left(e^{-\gamma(t-t_n)} - e^{-\beta_R(t-t_n)}\right) \\ &\quad + f(u_n) - k_p\sigma_{R,n}\int_{t_n}^{t} e^{-\gamma(t-s)}\left(\int_{t_n}^{s} e^{-\beta(s-u)}dW_R(u)\right)ds \\ &= e^{-\gamma(t-t_n)}P_n + \Upsilon_n(u_n, \widetilde{Y}_n) + M(t), \end{aligned} \tag{C.7}$$

where $M(t), t \in [t_n, t_{n+1})$ is a martingale given by

$$M(t) = -k_p\sigma_{R,n}\int_{t_n}^{t} e^{-\gamma(t-s)}\left(\int_{t_n}^{s} e^{-\beta(s-u)}dW_R(u)\right)ds \tag{C.8}$$

and the function Υ_n is given by

$$\Upsilon_n(\nu, y) = f(\nu, y) + \eta_n$$

with

$$\eta_n = \frac{k_P}{\beta_R - \gamma}(\mu_{R,n} - R_n)\left(e^{-\gamma(t-t_n)} - e^{-\beta_R(t-t_n)}\right) - \frac{k_P \mu_{R,n}}{\gamma}(1 - e^{-\gamma(t-t_n)}).$$

The function f is given by

$$f(\nu, y) = \begin{cases} \int_{t_n}^{t} (\gamma P_{amb} + \kappa^F) e^{-\gamma(t-s)} ds, & \nu = u^F \\ \int_{t_n}^{t} \left[\gamma P_{amb} + \kappa_C (P_{in} - P_{out})\right] e^{-\gamma(t-s)} ds, & \nu = u^C \\ \int_{t_n}^{t} \gamma P_{amb} e^{-\gamma(t-s)} ds, & \nu = u^W \\ \int_{t_n}^{t} e^{-\gamma(t-s)} (\gamma P_{amb} - \kappa_D Q_C^I) ds + e^{-\gamma(t-t_n)} \int_{t_n}^{t} e^{\gamma(s-t_n)} \kappa_D C^O \widetilde{Y}(s)) ds, & \nu = u^D \end{cases}$$

$$= \begin{cases} \left(P_{amb} + \frac{\kappa^F}{\gamma}\right)(1 - e^{-\gamma(t-t_n)}), & \nu = u^F \\ \left(P_{amb} + \frac{\kappa_C (P_{in} - P_{out})}{\gamma}\right)(1 - e^{-\gamma(t-t_n)}), & \nu = u^C \\ P_{amb}(1 - e^{-\gamma(t-t_n)}), & \nu = u^W \\ \left(P_{amb} - \frac{\kappa_D Q_C^I}{\gamma}\right)(1 - e^{-\gamma(t-t_n)}) + e^{-\gamma(t-t_n)} \psi_n(y), & \nu = u^D \end{cases}$$

with ψ_n given by

$$\psi_n(y) = \int_{t_n}^{t} e^{\gamma(s-t_n)} \kappa_D C^O \widetilde{Y}(s)) ds$$

Plugging the closed-form solution of the ODE (5.1) given in equation 6.2 in this integral and using the Assumption 6.1.2, for $g_n^D = g_n^{u^D}$ and $\widetilde{Y}_n(t) = y$ gives

$$\begin{aligned} \psi_n(y) &= \kappa_D C^O \int_{t_n}^{t} e^{\gamma(s-t_n)} \left(e^{\widetilde{A}(s-t_n)} \widetilde{Y}(t_n) + (e^{\widetilde{A}(s-t_n)} - \mathbb{I}_\ell) \widetilde{A}^{-1} \widetilde{B} g_n^D\right) ds \\ &= \kappa_D C^O \left\{ \int_{t_n}^{t} e^{(\gamma \mathbb{I}_\ell + \widetilde{A})(s-t_n)} \widetilde{Y}(t_n) ds + \int_{t_n}^{t} \left(e^{(\gamma \mathbb{I}_\ell + \widetilde{A})(s-t_n)} - e^{\gamma(s-t_n)} \mathbb{I}_\ell\right) ds \widetilde{A}^{-1} \widetilde{B} g_n^D \right\}. \\ &= \kappa_D C^O \Bigg\{ (e^{(\gamma \mathbb{I}_\ell + \widetilde{A})(t-t_n)} - \mathbb{I}_\ell)(\gamma \mathbb{I}_\ell + \widetilde{A})^{-1} y + \\ &\qquad \left[(e^{(\gamma \mathbb{I}_\ell + \widetilde{A})(t-t_n)} - \mathbb{I}_\ell)(\gamma \mathbb{I}_\ell + \widetilde{A})^{-1} - \frac{1}{\gamma}(e^{\gamma(t-t_n)} - 1)\mathbb{I}_\ell \right] \widetilde{A}^{-1} \widetilde{B} g_n^D \Bigg\}. \end{aligned}$$

□

C.2.4 Proof of Proposition 6.1.7

The following intermediate result is required for the proof of second statement of Proposition 6.1.7 presented below and for the proof of Theorem 6.1.9 presented in Appendix C.2.5.

Intermediate result.

Lemma C.2.2 Let $s,t \in [t_n, t_{n+1})$. Under Assumption 6.1.1, the conditional covariance of $R(t)$ and $R(s)$ given $R(t_n) = r$, is given by

$$cov(R(t), R(s) \mid R(t_n) = r) = \frac{\sigma_{R,n}^2}{2\beta_R}(e^{-\beta_R|t-s|} - e^{-\beta_R(t+s)+2\beta_R t_n}). \tag{C.9}$$

Proof. Under Assumption 6.1.1, applying Itô isometry property, the closed-form expression (C.4) yields

$$\begin{aligned}
cov(R(t),R(s)\,|R(t_n) = r) &= \mathbb{E}[(R(t) - \mathbb{E}[R(t) \mid R(t_n) = r])(R(s) - \mathbb{E}[R(s) \mid R(t_n) = r]) \mid R(t_n) = r] \\
&= \mathbb{E}\left[\left(\int_{t_n}^{t} \sigma_{R,n} e^{-\beta_R(t-u)} dW_R(u)\right)\left(\int_{t_n}^{s} \sigma_{R,n} e^{-\beta_R(s-v)} dW_R(v)\right) \,\middle|\, R(t_n) = r\right] \\
&= \sigma_{R,n}^2 e^{-\beta_R(t+s)} \mathbb{E}\left[\left(\int_{t_n}^{t} e^{\beta_R u} dW_R(u)\right)\left(\int_{t_n}^{s} e^{\beta_R v} dW_R(v)\right) \,\middle|\, R(t_n) = r\right] \\
&= \sigma_{R,n}^2 e^{-\beta_R(t+s)} \mathbb{E}\left[\left(\int_{t_n}^{\min(t,s)} e^{\beta_R u} dW_R(u)\right)^2 \,\middle|\, R(t_n) = r\right] \\
&= \sigma_{R,n}^2 e^{-\beta_R(t+s)} \mathbb{E}\left[\int_{t_n}^{\min(t,s)} e^{2\beta_R u} du \,\middle|\, R(t_n) = r\right] \\
&= \frac{\sigma_{R,n}^2}{2\beta_R} e^{-\beta_R(t+s)}(e^{2\beta_R \min(t,s)} - e^{2\beta_R t_n}) \\
&= \frac{\sigma_{R,n}^2}{2\beta_R}(e^{-\beta_R|t-s|} - e^{-\beta_R(t+s)+2\beta_R t_n}).
\end{aligned}$$

□

Proof of Proposition 6.1.7.

Proof. Let $R_{n+1} = R(t_{n+1})$ be the sampling of the residual demand at time t_{n+1}. Taking the conditional expectation on both sides of the expression of the closed-form solution of the SDE given by equation (C.4), and using the fact that, for $t \in [t_n, t_{n+1})$, $\int_{t_n}^{t} e^{-\beta(t-s)} dW_R(s)$ is a martingale, yields

$$\begin{aligned}
m_{R,n} &= \mathbb{E}[R_{n+1} \mid R_n = r] \\
&= \mathbb{E}[R_n e^{-\beta_R(t_{n+1}-t_n)} + \mu_{R,n}(1 - e^{-\beta_R(t_{n+1}-t_n)}) + \sigma_{R,n} \int_{t_n}^{t_{n+1}} e^{-\beta(t_{n+1}-s)} dW_R(s) \mid R_n = r] \\
&= r e^{-\beta_R \Delta_N} + \mu_{R,n}(1 - e^{-\beta_R \Delta_N}) + \sigma_{R,n} \mathbb{E}\left[\int_{t_n}^{t_{n+1}} e^{-\beta(t_{n+1}-s)} dW_R(s) \mid R_n = r\right] \\
&= r e^{-\beta_R \Delta_N} + \mu_{R,n}(1 - e^{-\beta_R \Delta_N}).
\end{aligned}$$

For $t = t_{n+1}$, using the definition of the conditional variance, equation (C.9) yields

$$\Sigma_{R,n}^2 = Var[R_{n+1} \mid R_n = r] = cov(R(t_{n+1}), R(t_{n+1}) \mid R_n = r) = \frac{\sigma_{R,n}^2}{2\beta_R}(1 - e^{-\beta_R \Delta_N}).$$

□

C.2.5 Proof of Theorem 6.1.9

To prove the proof of second statement of Theorem 6.1.9 the following intermediate result is required.

Conditional Covariance of $P(t)$ and $P(s)$ Given $X_n = x$ and $u_n = \nu$

Let $\nu \in \overline{\mathcal{U}} \backslash \{u^O\}$. We want to find the covariance function and derive the variance. We use the short-hand notations $cov_r(R(t),R(s)) = cov(R(t),R(s) \mid R_n = r)$ for the conditional covariance of $R(t)$ and $R(s)$given $R(t_n) = R_n$ and $cov_p(P(t),P(s)) = cov(P(t),P(s) \mid X_n = x, u_n = \nu)$ for the conditional covariance of $P(t)$ and $P(s)$ given $X_n = x$ and $u_n = \nu$.

Lemma C.2.3 Let $s,t \in [t_n,t_{n+1})$. Under Assumption 6.1.1, the conditional covariance of $P(t)$ and $P(s)$ given $X(t_n) = x$ and $u_n = \nu$ is given by

$$\begin{aligned} cov_p(P(t),P(s)) = \frac{k_P^2\sigma_{R,n}^2}{2\beta_R(\beta_R-\gamma)^2(\beta_R+\gamma)}\Big\{ & 2\beta_R e^{-\beta_R(s-t_n)-\gamma(t-t_n)} + 2\beta_R e^{-\beta_R(t-t_n)-\gamma(s-t_n)} \\ & - (\beta_R-\gamma)e^{-\beta_R|t-s|} + \left(\frac{\beta_R^2}{\gamma}-\beta_R\right)e^{-\gamma|t-s|} - (\beta_R+\gamma)e^{-\beta_R(t+s)+2\beta_R t_n} \\ & - \left(\frac{\beta_R^2}{\gamma}+\beta_R\right)e^{-\gamma(t+s)+2\gamma t_n}\Big\}. \end{aligned} \tag{C.10}$$

Proof. We first recall the following properties of the conditional covariance of a random variables X and Y given a filtration $\mathcal{G}$. For $a,b,c,d \in \mathbb{R}$, we have

$$cov(a+bX,c+dY \mid \mathcal{G}) = cov(bX,dY \mid \mathcal{G}) = bdcov(X,Y \mid \mathcal{G}).$$

Under assumption 6.1.1, using the the expression of $P(t)$ given in (C.6) and the property of the covariance given above, we obtain

$$\begin{aligned} cov_p(P(t),P(s)) &= cov\Bigg(e^{-\gamma(t-t_n)}P_n + f(u_n) - k_P\int_{t_n}^t e^{-\gamma(t-u)}R(u)du, e^{-\gamma(s-t_n)}P_n \\ &\qquad + f(u_n) - k_P\int_{t_n}^s e^{-\gamma(s-v)}R(v)dv \;\Bigg|\; X_n = x, u_n = \nu\Bigg) \\ &= cov\Bigg(-k_P\int_{t_n}^t e^{-\gamma(t-u)}R(u)du, k_P\int_{t_n}^s e^{-\gamma(s-v)}R(v)dv \;\Bigg|\; X_n = x, u_n = \nu\Bigg) \\ &= k_p^2 e^{-\gamma(t+s)}\int_{t_n}^t\int_{t_n}^s e^{\gamma(u+v)}cov(R(u),R(v) \mid R_n = r)dudv \end{aligned}$$

Plugging the $cov(R(u),R(v) \mid R_n = r)$ given in (C.9) in the above expression yields

$$\begin{aligned} cov_p(P(t),P(s)) &= \frac{k_P^2\sigma_{R,n}^2}{2\beta_R}e^{-\gamma(t+s)}\int_{t_n}^t\int_{t_n}^s e^{\gamma(u+v)}\left(e^{-\beta|u-v|} - e^{-\beta(u+v)+2\beta_R t_n}\right)dudv \\ &= \frac{k_P^2\sigma_{R,n}^2}{2\beta_R}\left(e^{-\gamma(t+s)}J_1(s,t) - e^{-\gamma(t+s)+2\beta_R t_n}J_2(s,t)\right), \end{aligned} \tag{C.11}$$

where

$$J_1(s,t) = \int\int_{\Delta} e^{\gamma(u+v)-\beta|u-v|}dudv \quad \text{and} \quad J_2(s,t) = \int\int_{\Delta} e^{-(\beta_R-\gamma)(u+v)}dudv.$$

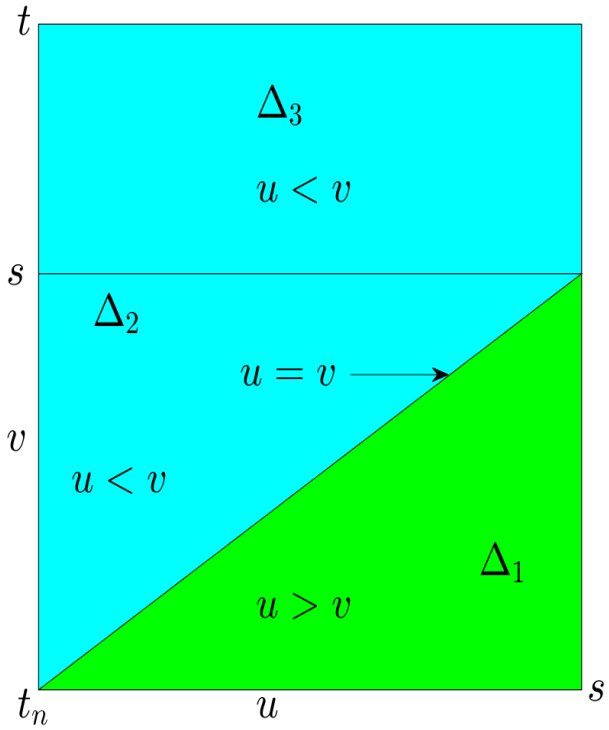


Figure C.1: Domain of integration for $t > s$

Without loss of generality, we assume that $s \leq t$. Then the domain of integration reads as $\Delta = \{(u,v) \in [t_n,s] \times [t_n,t], t_n \leq s \leq t \leq t_{n+1}\}$. For the integrand with absolute value the domain of integration Δ can be divided into 3 subdomains Δ_1, Δ_2 and Δ_3, see Figure C.1 below where

$$\begin{aligned}
\Delta_1 &= \{(u,v) \in \Delta, t_n \leq u \leq s, t_n \leq v \leq u\} \\
\Delta_2 &= \{(u,v) \in \Delta, t_n \leq u \leq s, u \leq v \leq s\} \\
\Delta_3 &= \{(u,v) \in \Delta, t_n \leq u \leq s, s \leq v \leq t\}.
\end{aligned}$$

Then, the integral $J_1(s,t)$ can be split as follows:

$$\begin{aligned}
J_1(s,t) &= \int\int_{\Delta} e^{\gamma(u+v)-\beta|u-v|}dudv \\
&= \int\int_{\Delta_1} e^{-(\beta_R-\gamma)u+(\beta_R+\gamma)v}dvdu + \int\int_{\Delta_2} e^{-(\beta_R-\gamma)v+(\beta_R+\gamma)u}dvdu \\
&\quad + \int\int_{\Delta_3} e^{-(\beta_R-\gamma)v+(\beta_R+\gamma)u}dudv \\
&= J_{\Delta_1}(s,t) + J_{\Delta_2}(s,t) + J_{\Delta_3}(s,t)
\end{aligned}$$

where

$$J_{\Delta_1}(s,t) = \int\int_{\Delta_1} e^{-(\beta_R-\gamma)u+(\beta_R+\gamma)v}dvdu$$

$$= \frac{1}{\beta_R+\gamma}\int_{t_n}^{s}\left(e^{2\gamma u} - e^{-(\beta_R-\gamma)u+(\beta_R+\gamma)t_n}\right)du$$
$$= \frac{1}{2(\beta_R+\gamma)\gamma}\left(e^{2\gamma s} - e^{2\gamma t_n}\right) + \frac{1}{\beta_R^2-\gamma^2}\left(e^{-\beta_R(s-t_n)+\gamma(s+t_n)} - e^{2\gamma t_n}\right),$$

$$J_{\Delta_2}(s,t) = \int\int_{\Delta_2} e^{-(\beta_R-\gamma)v+(\beta_R+\gamma)u}dvdu$$
$$= \frac{1}{\beta_R-\gamma}\int_{t_n}^{s}\left(e^{2\gamma u} - e^{-(\beta_R-\gamma)s+(\beta_R+\gamma)u}\right)du$$
$$= \frac{1}{2(\beta_R-\gamma)\gamma}\left(e^{2\gamma s} - e^{2\gamma t_n}\right) + \frac{1}{(\beta_R^2-\gamma^2)}\left(e^{-\beta_R(s-t_n)+\gamma(s+t_n)} - e^{2\gamma s}\right),$$

and

$$J_{\Delta_3}(s,t) = \int\int_{\Delta_3} e^{-(\beta_R-\gamma)v+(\beta_R+\gamma)u}dudv$$
$$= \frac{1}{\beta_R+\gamma}\int_{s}^{t}\left(e^{-(\beta_R-\gamma)v+(\beta_R+\gamma)s} - e^{-(\beta_R-\gamma)v+(\beta_R+\gamma)t_n}\right)dv$$
$$= \frac{1}{\beta_R^2-\gamma^2}\left\{e^{-\beta_R(t-t_n)+\gamma(t+t_n)} - e^{-\beta_R(t-s)+\gamma(t+s)} - e^{-\beta_R(s-t_n)+\gamma(s+t_n)} + e^{2\gamma s}\right\}.$$

Combining $J_{\Delta_1}, J_{\Delta_2}$ and J_{Δ_3}, and using the identity $\frac{1}{2(\beta_R+\gamma)\gamma} + \frac{1}{2(\beta_R-\gamma)\gamma} = \frac{\beta_R}{\gamma(\beta_R^2-\gamma^2)}$, we obtain

$$J_1(s,t) = \frac{1}{\beta_R^2-\gamma^2}\left\{e^{-\beta_R(s-t_n)+\gamma(s+t_n)} + e^{-\beta_R(t-t_n)+\gamma(t+t_n)} - e^{-\beta_R(t-s)+\gamma(t+s)}\right.$$
$$\left.\frac{\beta_R}{\gamma}\left(e^{2\gamma s} - e^{2\gamma t_n}\right) - e^{2\gamma t_n}\right\}.$$

This implies that

$$J^1(s,t) = e^{-\gamma(t+s)}J_1(s,t)$$
$$= \frac{1}{\beta_R^2-\gamma^2}\left\{e^{-\beta_R(s-t_n)-\gamma(t-t_n)} + e^{-\beta_R(t-t_n)-\gamma(s-t_n)} - e^{-\beta_R(t-s)} + \frac{\beta_R}{\gamma}e^{-\gamma(t-s)}\right.$$
$$\left. - \left(\frac{\beta_R}{\gamma}+1\right)e^{-\gamma(t+s)+2\gamma t_n}\right\}.$$

The integral $J_2(s,t)$ is given by

$$J_2(s,t) = \int\int_{\Delta} e^{-(\beta_R-\gamma)(u+v)}dudv$$
$$= \frac{1}{(\beta_R-\gamma)^2}\left\{e^{-(\beta_R-\gamma)(t+s)} - e^{-(\beta_R-\gamma)(t+t_n)} - e^{-(\beta_R-\gamma)(s+t_n)} + e^{-2(\beta_R-\gamma)t_n}\right\}.$$

This implies that

$$J^2(s,t) = e^{-\gamma(t+s)+2\beta_R t_n}J_2(s,t)$$

$$= \frac{1}{(\beta_R-\gamma)^2}\left\{e^{-\beta_R(t+s)+2\beta_R t_n} - e^{-\beta_R(t-t_n)-\gamma(s-t_n)} - e^{-\beta_R(s-t_n)-\gamma(t-t_n)} + e^{-\gamma(t+s)+2\gamma t_n}\right\}.$$

Substituting $J^1(s,t)$ and $J^2(s,t)$ in (C.11) and using the identities $\frac{1}{(\beta_R-\gamma)^2} + \frac{1}{\beta_R^2-\gamma^2} = \frac{2\beta}{(\beta_R-\gamma)^2(\beta_R+\gamma)}$ and $\frac{1}{(\beta_R-\gamma)^2} - \frac{1}{\beta_R^2-\gamma^2} = \frac{2\gamma}{(\beta_R-\gamma)^2(\beta_R+\gamma)}$ we obtain for $s \le t$:

$$\begin{aligned} cov(P(t),P(s)) = \frac{k_P^2\sigma_{R,n}^2}{2\beta_R(\beta_R-\gamma)^2(\beta_R+\gamma)}\Bigg\{ & 2\beta_R e^{-\beta_R(s-t_n)-\gamma(t-t_n)} + 2\beta_R e^{-\beta_R(t-t_n)-\gamma(s-t_n)} \\ & - (\beta_R-\gamma)e^{-\beta_R(t-s)} + \left(\frac{\beta_R^2}{\gamma} - \beta_R\right)e^{-\gamma(t-s)} - (\beta_R+\gamma)e^{-\beta_R(t+s)+2\beta_R t_n} \\ & - \left(\frac{\beta_R^2}{\gamma} + \beta_R\right)e^{-\gamma(t+s)+2\gamma t_n}\Bigg\}. \end{aligned}$$

Interchanging t and s, we obtain (C.10) for all t and s. □

Proof of Theorem 6.1.9

Proof. Let $\nu \in \overline{\mathcal{U}}\backslash\{u^O\}$. To prove the first statement we use the fact that for $s \in [t_n, t_{n+1})$ the process $\int_{t_n}^{s} e^{-\beta(s-u)}dW^R(u)$ is a martingale, i.e., $\mathbb{E}\left[\int_{t_n}^{s} e^{-\beta(s-u)}dW^R(u)\right] = 0$.
Taking the expectation in (C.8) and applying Fubini's theorem yields

$$\begin{aligned} \mathbb{E}\big[M(t) \mid X_n = x, u_n = \nu\big] &= \mathbb{E}\left[-k_p\sigma_R\int_{t_n}^{t} e^{-\gamma(t-s)}\left(\int_{t_n}^{s} e^{-\beta(s-u)}dW^R(u)\right)ds\right] \\ &= -k_p\sigma_R\int_{t_n}^{t} e^{-\gamma(t-s)}\mathbb{E}\left[\int_{t_n}^{s} e^{-\beta(s-u)}dW^R(u)\right] = 0 \end{aligned}$$

Then, for $x = (r,f,p,y)$ and $t = t_{n+1}$ the closed-form solution (C.7) yields

$$\begin{aligned} m_{P,n} &= \mathbb{E}[P_{n+1} \mid X_n = x, u_n = \nu] = \mathbb{E}[e^{-\gamma\Delta_N}p + \Upsilon_n(\nu,y) + M(t_{n+1}) \mid X_n = x, u_n = \nu] \\ &= e^{-\gamma\Delta_N}p + \Upsilon_n(\nu,y) + \mathbb{E}[M(t_{n+1}) \mid X_n = x, u_n = \nu] = e^{-\gamma\Delta_N}p + \Upsilon_n(\nu,y). \end{aligned}$$

Given that the conditional variance is given by

$$\Sigma_{P,n}^2 = cov(P(t),P(t) \mid X_n = x,\ u_n = \nu),$$

for $t = t_{n+1}$, we immediately have the conditional variance (6.12). □

Limiting case, $\gamma = 0$. Assume that the IS is perfectly insulated (no heat loss to the environment). Then, the following corollary can be derived from Theorem 6.1.9 as a limiting case for $\gamma = 0$.

Corollary C.2.4 (Perfectly insulated IS) Assume that Assumption 6.1.1 is fulfilled and the rate of heat loss to the environment $\gamma = 0$. Then, the parameters of the Gaussian process P_{n+1} given in Theorem 6.1.9 become:

1. The conditional mean is given by

$$m_{P_0,n} = p - k_P \mu_{R,n} \Delta_N + \frac{k_P}{\beta_R}(\mu_{R,n} - r)\left(1 - e^{-\beta_R \Delta_N}\right) + \Upsilon_n^0(\nu, y),$$

where Υ_n^0 is given by

$$\Upsilon_n^0(\nu, y) = \begin{cases} \kappa^F \Delta_N, & \nu = u^F \\ \kappa_C (P_{in} - P_{out}) \Delta_N, & \nu = u^C \\ 0, & \nu = u^W \\ \kappa_D Q_C^I \Delta_N + \psi_n^0(y), & \nu = u^D \end{cases} \tag{C.12}$$

with ψ_n^0 given for $g_n^D = g(t_n, u^D)$ by

$$\psi_n^0(y) = \kappa_D C^O \left\{ (e^{A\Delta_N} - \mathbb{I}_\ell) A^{-1} (y + A^{-1} B g_n^D) - \Delta_N A^{-1} B g_n^D \right\},$$

where $\mathbb{I}_\ell$ is an $\ell \times \ell$ identity matrix.

2. The conditional variance is given by

$$\Sigma_{P_0,n}^2 = \frac{k_P^2 \sigma_{R,n}^2}{2\beta_R^3} \left\{ 2\beta_R \Delta_N + 4e^{-\beta_R \Delta_N} - e^{-2\beta_R \Delta_N} - 3 \right\}. \tag{C.13}$$

The proof this corollary is given below

Proof. For $\gamma = 0$ and $t, s \in [t_n, t_{n+1})$, we have the following limits

$$\lim_{\gamma \to 0} \frac{\beta_R^2}{\gamma} (e^{-\gamma|t-s|} - e^{-\gamma(t+s)+2\gamma t_n}) = 2\beta_R^2 (\min(s,t) - t_n),$$
$$\lim_{\gamma \to 0} \frac{\beta_R^2}{\gamma} \left(1 - e^{-2\gamma \Delta_N}\right) = 2\beta_R^2 \Delta_N, \qquad \lim_{\gamma \to 0} \frac{1}{\gamma} \left(1 - e^{\gamma \Delta_N}\right) = -\Delta_N.$$

Then, plugging the above approximations into relation (C.7) and yields equation (C.12). Plugging the above approximations into relation (C.10) reduces the conditional covariance to

$$cov_p(P(t), P(s)) = \frac{k_P^2 \sigma_{R,n}^2}{2\beta_R^3} \Big\{ 2\beta_R (\min(s,t) - t_n) + 2e^{-\beta_R (s - t_n)} + 2e^{-\beta_R (t - t_n)} - e^{-\beta_R |t-s|} - e^{-\beta_R (t+s) + 2\beta_R t_n} - 2 \Big\},$$

and for $t = s = t_{n+1}$ the conditional variance (C.13) follows. □

C.3 Details on Joint Conditional Distribution

C.3.1 Proof of Theorem 6.1.10

Proof. We want to show that Equation (6.14) holds true. We denote by $cov_{rp}(P(t),R(t)) = cov(P(t),R(t) \mid X_n = x,\ u_n = \nu)$, $t \in [t_n, t_{n+1})$ the conditional covariance of $(R(t),P(t))$ given that at time t_n, $X_n = x$ and $u_n = \nu$. Using the expression of $P(t)$, $t \in [t_n, t_{n+1})$ given in equation (C.6) and the property of the conditional covariance given in Appendix C.2.5, we have

$$\begin{aligned}
cov_{rp}(P(t),R(t)) &= cov\left(e^{-\gamma(t-t_n)}P_n + f(u_n) - k_P \int_{t_n}^{t} e^{-\gamma(t-s)}R(s)ds, R(t) \;\middle|\; X_n = x,\ u_n = \nu \right) \\
&= cov\left(-k_P \int_{t_n}^{t} e^{-\gamma(t-u)}R(s)ds, R(t) \;\middle|\; X_n = x,\ u_n = \nu \right) \\
&= -k_p \int_{t_n}^{t} e^{\gamma(t-s)} cov(R(s),R(t) \mid R_n = r)ds \\
&= -\frac{k_P \sigma_{R,n}^2}{2\beta_R} \int_{t_n}^{t} e^{-\gamma(t-u)} \left(e^{-\beta_R|t-s|} - e^{-\beta_R(t+s)+2\beta_R t_n} \right) ds \\
&= -\frac{k_P \sigma_{R,n}^2}{2\beta_R} \int_{t_n}^{t} \left(e^{-(\beta_R+\gamma)(t-s)} - e^{-(\beta_R-\gamma)s-(\beta_R+\gamma)t+2\beta_R t_n} \right) ds \\
&= -\frac{k_P \sigma_{R,n}^2}{2\beta_R} \left\{ \frac{1}{\beta_R+\gamma} \left(1 - e^{-(\beta_R+\gamma)(t-t_n)} \right) + \frac{1}{\beta_R-\gamma} \left(e^{-2\beta_R(t-t_n)} - e^{-(\beta_R+\gamma)(t-t_n)} \right) \right\}
\end{aligned}$$

Using the identity $\frac{1}{\beta_R+\gamma} + \frac{1}{\beta_R-\gamma} = \frac{2\beta_R}{\beta_R^2-\gamma^2}$, for $t = t_{n+1}$, we immediately have (6.14). Now, Let $\mathcal{E}_{n+1}^R$ and $\mathcal{E}_{n+1}^P$ be two independent normally distributed random variables. Since P_{n+1} and R_{n+1} are correlated Gaussian processes, from the closed-form expression (6.9) and (6.1.4), we can derive the following recursions.

$$\begin{aligned}
R_{n+1} &= m_{R,n} + \Sigma_{R,n}\mathcal{E}_{n+1}^R, \\
P_{n+1} &= m_{P,n} + \Sigma_{P,n}\left(\sqrt{1-\rho_{RP,n}^2}\mathcal{E}_{n+1}^P + \rho_{RP,n}\mathcal{E}_{n+1}^R \right).
\end{aligned}$$

$$\begin{aligned}
\Sigma_{RP,n}^2 &= cov(R_{n+1}, P_{n+1} \mid X_n = x,\ u_n = \nu) \\
&= cov\left(m_{R,n} + \Sigma_{R,n}\mathcal{E}_{n+1}^R, m_{P,n} + \Sigma_{P,n}\left(\sqrt{1-\rho_{RP,n}^2}\mathcal{E}_{n+1}^P + \rho_{RP,n}\mathcal{E}_{n+1}^R \right) \;\middle|\; X_n = x,\ u_n = \nu \right) \\
&= cov\left(\Sigma_{R,n}\zeta_{n+1}^R, \Sigma_{P,n}\left(\sqrt{1-\rho_{RP,n}^2}\zeta_{n+1}^P + \rho_{RP,n}\zeta_{n+1}^R \right) \;\middle|\; X_n = x,\ u_n = \nu \right) \\
&= \Sigma_{R,n}\Sigma_{P,n}\left[\sqrt{1-\rho_{RP,n}^2} cov(\zeta_{n+1}^R, \zeta_{n+1}^P \mid X_n = x,\ u_n = \nu) + \rho_{RP,n} cov(\zeta_{n+1}^R, \zeta_{n+1}^R \mid R_n = r) \right] \\
&= \rho_{RP,n}\Sigma_{R,n}\Sigma_{P,n} cov(\zeta_{n+1}^R, \zeta_{n+1}^R \mid R_n = r) \\
&= \rho_{RP,n}\Sigma_{R,n}\Sigma_{P,n},
\end{aligned}$$

from which we derive relation (6.13). □

Corollary C.3.1 (Perfectly insulated IS, $\gamma = 0$) For perfectly insulated IS, the conditional covariance of R_{n+1} and P_{n+1} given by (6.14) reads as

$$\Sigma_{RP,n} = -\frac{k_P \sigma_{R,n}^2}{2\beta_R^2}\left(1 - 2e^{-\beta_R \Delta_N} + e^{-2\beta_R \Delta_N}\right).$$

This corollary shows that P_{n+1} and R_{n+1} are negatively correlated also when $\gamma = 0$ since

$$1 - 2e^{-\beta_R \Delta_N} + e^{-2\beta_R \Delta_N} = \left(1 - e^{-\beta_R \Delta_N}\right)^2 > 0.$$

C.3.2 Proof of Proposition 6.1.11

Proof. Z_{RP} is a normally distributed as a linear combination of two jointly normally distributed random variables, we have

$$\begin{aligned}
\mathbb{E}[Z_{RP} \mid X_n = x, u_n = v] &= \mathbb{E}\left[\frac{Z_P - \rho_{RP,n} Z_R}{\sqrt{1-\rho_{RP,n}^2}} \;\middle|\; X_n = x, u_n = v\right] \\
&= \frac{1}{\sqrt{1-\rho_{RP,n}^2}} \mathbb{E}[Z_P - \rho_{RP,n} Z_R \mid X_n = x, u_n = v] \\
&= \frac{1}{\sqrt{1-\rho_{RP,n}^2}} (\mathbb{E}[Z_P \mid X_n = x, u_n = v] - \rho_{RP,n} \mathbb{E}[Z_R \mid R_n = r]) = 0. \\
Var[Z_{RP} \mid X_n = x, u_n = v] &= Var\left[\frac{Z_P - \rho_{RP,n} Z_R}{\sqrt{1-\rho_{RP,n}^2}} \;\middle|\; X_n = x, u_n = v\right] \\
&= \frac{1}{1-\rho_{RP,n}^2} Var[Z_P - \rho_{RP,n} Z_R \mid X_n = x, u_n = v] \\
&= \frac{1}{1-\rho_{RP,n}^2}\Big(Var[Z_P \mid X_n = x, u_n = v] + \rho_{RP,n}^2 Var[Z_R \mid R_n = r] \\
&\qquad - 2\rho_{RP,n} cov(Z_P, Z_R \mid X_n = x, u_n = v)\Big) \\
&= \frac{1}{1-\rho_{RP,n}^2}(1 + \rho_{RP,n}^2 - 2\rho_{RP,n}^2) = \frac{1-\rho_{RP,n}^2}{1-\rho_{RP,n}^2} = 1.
\end{aligned}$$

□

C.3.3 Proof of Proposition 6.1.13

Proof. Using the variables z_r and z_p defined above, we obtain

$$\varphi_{RP}(r,p) = \frac{1}{2\pi\Sigma_{R,n}\Sigma_{P,n}\sqrt{1-\rho_{RP,n}^2}} \exp\left\{-\frac{1}{2(1-\rho_{RP,n}^2)}\left(z_p^2 - 2\rho_{RP,n} z_P z_r + z_r^2\right)\right\}$$

$$= \frac{1}{2\pi\Sigma_{R,n}\Sigma_{P,n}\sqrt{1-\rho_{RP,n}^2}} \exp\left\{ -\frac{1}{2(1-\rho_{RP,n}^2)} \left((z_p - \rho_{RP,n} z_r)^2 + (1-\rho_{RP,n}^2) z_r^2 \right) \right\}$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{2\pi}\Sigma_{P,n}\Sigma_{P,n}\sqrt{1-\rho_{RP,n}^2}} \exp\left\{ -\frac{1}{2} \left(\frac{z_p - \rho_{RP,n} z_r}{\sqrt{1-\rho_{RP,n}^2}} \right)^2 - \frac{z_r^2}{2} \right\}$$

$$= \frac{1}{\Sigma_{P,n}\sqrt{2\pi}\sqrt{1-\rho_{RP,n}^2}} \exp\left\{ -\frac{z_{RP}^2}{2} \right\} \frac{1}{\Sigma_{R,n}\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left(\frac{r - m_{R,n}}{\Sigma_{R,n}} \right)^2 \right\}$$

$$= \varphi_R(r) \frac{1}{\Sigma_{P,n}\sqrt{2\pi}\sqrt{1-\rho_{RP,n}^2}} e^{-\frac{1}{2}\zeta_z^2(r,p)}.$$

□

C.3.4 Proof of Proposition 6.1.14

Proof. Let $p_2, r_2 \in \mathbb{R}$. Relation (6.1.2) implies that

$$\begin{aligned}
\mathbb{P}(P_{n+1} \le p_2,\ R_{n+1} \le r_2) &= \int_{-\infty}^{r_2} \int_{-\infty}^{p_2} \varphi_{PR}(r,p)\, dp\, dr \\
&= \int_{-\infty}^{r_2} \int_{-\infty}^{p_2} \varphi_R(r) \frac{1}{\Sigma_{P,n}\sqrt{2\pi}\sqrt{1-\rho_{RP,n}^2}} e^{-\frac{1}{2}\zeta_z^2(r,p)}\, dp\, dr \\
&= \int_{-\infty}^{r_2} \int_{-\infty}^{\zeta_z(r,p_2)} \varphi_R(r) \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, dz\, dr \\
&= \int_{-\infty}^{r_2} \varphi_R(r) \int_{-\infty}^{\zeta_z(r,p_2)} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, dz\, dr \\
&= \int_{-\infty}^{r_2} \varphi_R(r) \Phi(\zeta_z(r,p_2))\, dr.
\end{aligned}$$

□

C.4 Construction of New Basis for Reduced Order System

C.4.1 Practical Construction of New Basis Vectors

In this subsection we give details on how we can practically choose the new basis vectors. Choose $c \neq 0$. Then the hyperplane $C^M \widetilde{Y} = c$ does not contain the origin 0, there are intersection points $A_1, A_2, \ldots, A_\ell$, with the axes with coordinates $a_1, a_2, \ldots, a_\ell$. Let us assume that there is an intersection with the axis ℓ with coordinate a_ℓ. For $k = 1, 2, \ldots, \ell - 1$, define the vectors $v_1, v_2, \ldots, v_{\ell-1}$, by

$$v_k = \begin{cases} \overrightarrow{A_\ell A_k} = (0, \ldots, 0, a_k, 0, \ldots, 0, -a_\ell)^\top & \text{if there is intersection with axis k} \\ e_k \quad = (0, \ldots, 0, 1, 0, \ldots, 0)^\top & \text{else} \end{cases}$$

Then, the vectors $v_1, v_2, \ldots, v_{\ell-1}$, are linearly independent.

Now we choose the vector v_ℓ such that $\langle v_\ell, v_k \rangle = 0$, for $k = 1, \ldots, \ell-1$. Therefore, the vectors $v_1, v_2, \ldots, v_\ell$, form a sequence of linear independent vectors in $\mathbb{R}^\ell$. Let $v_\ell = (b^1, b^2, \ldots, b^\ell)$ be the unknown coordinates of v_ℓ. Then, $\langle v_\ell, v_k \rangle = 0$, is a homogeneous system of equations

$$\begin{pmatrix} v_1^1 & v_1^2 & \ldots & v_1^\ell \\ v_2^1 & v_2^2 & \ldots & v_2^\ell \\ \vdots & & & \vdots \\ v_{\ell-1}^1 & v_{\ell-1}^2 & \ldots & v_{\ell-1}^\ell \end{pmatrix} \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^\ell \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The coefficient matrix of the above system has rank $\ell - 1$ and there are infinitely many solutions of the system containing one free parameter. Finally, the linear independent vectors $v_1, v_2, \ldots, v_\ell$ can be transformed by Gram-Schmidt orthogonalization into ℓ orthogonal vectors.

C.4.2 Proof of Lemma 6.2.1

Proof.

1. Assume that the output matrix $C^M \neq 0$. Let

   $$\mathcal{Y}_c = \{Y \in \mathbb{R}^\ell : C^M Y = c\} \subset \mathbb{R}^\ell$$

   be a subset of $\mathbb{R}^\ell$ containing all points with constant average temperature $\overline{Q}^M = c$. Since $\overline{Q}^M = C^M \widetilde{Y}$, the subset $\mathcal{Y}_c$ defines an hyperplane in $\mathbb{R}^\ell$ and for $c = 0$ we have

   $$\mathcal{Y}_0 = \{Y \in \mathbb{R}^\ell : C^M Y = 0\}$$

   is an $(\ell - 1)$-dimensional subspace of $\mathbb{R}^\ell$ containing the origin 0 and all points with average temperature $\overline{Q}^M = 0$. By construction, the new basis vectors $v_1, v_2, \ldots, v_\ell$, are such that $v_1, v_2, \ldots, v_{\ell-1}$, span the subspace $\mathcal{Y}_0$ while the last vector v_ℓ is orthogonal to $\mathcal{Y}_0$. Thus all linear combination of $v_1, v_2, \ldots, v_{\ell-1}$, have the same average temperature $\overline{Q}^M = 0$. Therefore, the hyperplanes $C^M \overline{Y} = c$ are parallel to $v_1, v_2, \ldots, v_{\ell-1}$, and orthogonal to v_ℓ.

2. The average temperature $\overline{Q}^M = C^M \widetilde{Y}$ can be written in terms of the new coordinates as $\overline{Q}^M = C^M \overline{V}\ \overline{Y} = \overline{C}^M \overline{Y}$ where $\overline{v} = (v_1, v_2, \ldots, v_\ell)$ and $\overline{C}^M = C^M \overline{v} \in \mathbb{R}^{1 \times \ell}$. For $\overline{Y}_p = (\overline{Y}^1, \ldots, \overline{Y}^{\ell-1}, 0)\top$ and $\overline{Y}_q = (0, \ldots, 0, \overline{Y}^l)$, the decomposition of $\overline{Y} = \overline{Y}_p + \overline{Y}_q$ into the projection of $\overline{Y}$ onto span$\{v_1, v_2, \ldots, v_{\ell-1}\}$ and v_ℓ gives

   $$\overline{Q}^M = \overline{C}^M \overline{Y} = \overline{C}^M \overline{Y}_p + \overline{C}^M \overline{Y}_q = \overline{C}^{M,\ell} \overline{Y}^\ell.$$

   Since $\overline{Y}_p$ is a linear combination of the first $\ell - 1$ basis vectors, it belongs to the subspace $\mathcal{Y}_0$ with $\overline{C}^M \overline{Y}_p = 0$. Hence, in the new coordinate it holds $\overline{Q}^M = \overline{C}^{M,\ell} \overline{Y}^\ell$, i.e. the last coordinate $\overline{Y}^\ell$ is up to a scaling constant the average temperature.

3. Finally, the relation $0 = \overline{C}^M \overline{Y}_p = \sum_{k=1}^{\ell-1} \overline{C}^{M,k} \overline{Y}^k = 0$ holds for all $\overline{Y}_p \in \mathbb{R}^{\ell-1}$. This implies that all the first $\ell - 1$ entries of the row matrix $\overline{C^M}$ must vanish, i.e. $\overline{C}^{M,1} = \ldots = \overline{C}^{M,\ell-1} = 0$.

Using $\overline{Q}^M = \overline{C}^{M,\ell}\overline{Y}^\ell$ the row matrix $\overline{C}^M$ reads as

$$\overline{C}^M = (0,\dots,0,\overline{C}^{M,\ell}).$$

Since $\overline{C}^M = C^M\overline{V}$ the last entry $\overline{C}^{M,\ell}$ is given by

$$\overline{C}^{M,\ell} = C^M V_\ell.$$

□

Nomenclature

List of Abbreviations

PHX	pipe heat exchanger
ADP	Approximate dynamic programming
BT	Balanced truncation
MOR	Model order reduction
MDP	Markov decision process
IS	Internal storage
GS	Geothermal storage
ODE	Ordinary differential equation
PDE	Partial differential equation
HJB	Hamilton-Jacobi-Bellman
SDE	Stochastic differential equation
COP	Coefficient of performance of the heat pump
LTI	Linear time-invariant
DPE	Dynamic programming equation
PDPE	Post-decision dynamic programming equation
LMS	Least square Monte Carlo
CHP	Combined heat and power

List of Symbols

$Q = Q(t,x,y)$	temperature in the GS
T	finite time horizon
l_x, l_y, l_z	width, height and depth of the storage
$\mathcal{D} = (0,l_x) \times (0,l_y)$	domain of the GS
$\mathcal{D}^F$, $\mathcal{D}^M$	domain inside and outside the pipes
$\mathcal{D}^J = \underline{\mathcal{D}}^J \cup \overline{\mathcal{D}}^J$	interface between the pipes and the medium
$\partial\mathcal{D}$	boundary of the domain
$\partial\mathcal{D}^I$, $\partial\mathcal{D}^O$	inlet and outlet boundaries of the pipe
$\partial\mathcal{D}^L, \partial\mathcal{D}^R, \partial\mathcal{D}^T$, $\partial\mathcal{D}^B$	left, right, top and bottom boundaries of the domain
$\mathcal{N}_*^*$	subsets of index pairs for grid points
$\mathcal{K}, \overline{\mathcal{K}}$	mappings $(i,j) \mapsto l$ of index pairs to single indices
$v = v_0(t)(v^x, v^y)^\top$	time-dependent velocity vector
$\overline{v}_0$	constant velocity during pumping
c_p^F, c_p^M	specific heat capacity of the fluid and medium
ρ^F, ρ^M	mass density of the fluid and medium
κ^F, κ^M	thermal conductivity of the fluid and medium
a^F, a^M	thermal diffusivity of the fluid and medium
λ^G	heat transfer coefficient between storage and underground
Q_0	initial temperature distribution of the GS
Q^G	underground temperature
Q^I, Q_C^I, Q_D^I	inlet temperature of the pipe, during charging and discharging
$\overline{Q}^M, \overline{Q}^F$	average temperature in the storage medium and fluid
$\overline{Q}^O, \overline{Q}^B$	average temperature at the outlet and bottom boundary
G^*	gain of thermal energy in a certain subdomain
I_C, I_W, I_D	time interval for charging, waiting, discharging periods
∇, $\Delta = \nabla \cdot \nabla$	gradient, Laplace operator
N_x, N_y,	number of grid points in x,y-direction
h_x, h_y	mesh size in x and y-direction
n_P	number of pipes
$\mathfrak{n}$	outward normal to the boundary $\partial\mathcal{D}$
n	dimension of state vector Y
ℓ	dimension of the reduced-order system
$\mathbb{I}_n$	$n \times n$ identity matrix
A, B, C	$n \times n$ system matrix, $n \times m$ input matrix, $n_0 \times n$ output matrix of original system
$\overline{A}$, $\overline{B}$, $\overline{C}$	$n \times n$ system matrix, $n \times m$ input matrix, $n_0 \times n$ output matrix of transformed original system
$\widetilde{A}$, $\widetilde{B}$, $\widetilde{C}$	$\ell \times \ell$ system matrix, $\ell \times m$ input matrix, $n_0 \times \ell$ output matrix of the reduced-order system
$D^{\pm}$, A_L, A_M, A_R	block matrices of matrix A
Y, $\overline{Y}$	n-dimensional state of original and transformed original system
$\widetilde{Y}$	ℓ-dimensional state of reduced-order system
Z	n_o-dimensional output of original system
$\widetilde{Z}$	n_o-dimensional output of reduced-order system

g	input variable of the system
$\mathcal{G}_C$, $\mathcal{G}_O$	controllability and observability Gramians
$\overline{\mathcal{G}_C}$, $\overline{\mathcal{G}_O}$	transformed controllability and observability Gramians
$\mathcal{T}$	transformation matrix
$\sigma_i > 0$	Hankel singular values
Σ	diagonal matrix of Hankel singular values
U, L	upper/ lower triangular matrix from Cholesky decomp. of $\mathcal{G}_C/\mathcal{G}_O$
K	orthogonal matrix from the eigenvalue decomposition of $U^\top \mathcal{G}_O U$
W, V	unitary matrices from the singular value decomposition
$\mathcal{S}(\ell)$	selection criterion
$\mathcal{L}^2(0,t)$	set of square integrable functions on $[0,t]$
$\mathbb{1}_X(.)$	indicator function of X
I_C, I_W and I_D	time interval for charging, waiting and discharging periods
$P(t) = P(x,t)$	average temperature in the IS
γ	rate of heat loss to the environment
$\underline{p}, \overline{p}$	minimum and maximum temperature in the IS
$\underline{q}, \overline{q}$	minimum and maximum temperature in the GS
R, F	residual demand and fuel price
$u^D(t)$	maximum discharging rate of the IS
$u^C(t)$	maximum charging rate of the IS
u^F	maximum rate of firing fuel
u^W, u^O	waiting and over-spilling control
$u = (u(t))_{t\in[0;T]}$	control process
$\widetilde{u}(t,x)$	Markov decision rule
u^*	optimal control
$\overline{\mathcal{U}} = \{u^O, u^D, 0, u^C, u^F\}$	set of feasible controls
$\mathcal{U}(t,x) \subset \overline{\mathcal{U}}$	continuous time control constraint
$\mathcal{A}$	set of admissible controls
$\mathcal{K}(X)$	state constraint set
Ω	sample space
W_R, W_F	Wiener processes
$\mathcal{G}_t$	sigma-algebra generated by $\{(W_R(s), W_F(s)),\ s \in [0,t]\}$
$\mathbb{G}$	filtration generated by the Wiener process $\{(W_R(t), W_F(t)),\ t \in [0,T]\}$
$\mathbb{P}$	probability measure on a measurable space $(\Omega, \mathcal{G}_T)$
$\widehat{X} = (R,F)$	generic variable for uncontrolled states
$\overline{X} = (P,Q) = (P,\widetilde{Y})$	generic variable for controlled states
$X = (\widehat{X}, \overline{X}) = (R,F,P,\widetilde{Y})$	state variable
$\mathbb{E}_{t,x}[\cdot] = \mathbb{E}[.\|X(t) = x]$	conditional expectation given that at time t the state $X(t) = x$
$\mu_{R/F}$	mean reversion level for residual demand/ fuel
$\beta_{R/F}$	mean-reversion speed for residual demand/ fuel
$\widehat{\mu}$, $\widehat{\sigma}$	drift and volatility of the uncontrolled state process $\widehat{X}$
$\delta^i_{R/F}$	length of the seasonal period for the i-th seasonality component for $R/\ F$
$t^i_{R/F}$	reference time for the i-th seasonality component for $R/\ F$

$K^i_{R/F}$	amplitude of the i-th seasonality component for $R/\ F$
$J(t,x;u),\ V(t,x)$	performance criterion and value function at time t
$\Psi,\ \phi$	running and terminal cost
$V(x)$	terminal value function
A_h	area of the IS
P_{amb}	constant ambient temperature around the IS
$\mathcal{F},\ \mathcal{R}$	continuous state space of the fuel price and the residual demand
$\mathcal{P}$	continuous state space of the temperature in the IS
$\mathcal{Y}$	continuous reduced-order state space of the GS
$\mathcal{X} = \mathcal{R} \times \mathcal{F} \times \mathcal{P} \times \mathcal{Y}$	continuous state space of the generic state variable
$\overline{\mathcal{F}} = [\underline{f}, \overline{f}],\ \overline{\mathcal{R}} = [\underline{r}, \overline{r}]$	truncated continuous state space of the fuel price and the residual demand
$\overline{\mathcal{P}} = [\underline{p}, \overline{p}]$	truncated continuous state space of the average temperature in the IS
$\overline{\mathcal{Y}} = [\underline{y}_1, \overline{y}_1] \times \ldots \times [\underline{y}_\ell, \overline{y}_\ell]$	truncated reduced order state space of the GS
$\overline{\mathcal{X}} = \overline{\mathcal{R}} \times \overline{\mathcal{F}} \times \overline{\mathcal{P}} \times \overline{\mathcal{Y}}$	truncated state space of the generic state variable
$\widetilde{\mathcal{F}},\ \widetilde{\mathcal{R}}$	discrete state space of the fuel price and the residual demand
$\widetilde{\mathcal{P}}$	discrete state space of the average temperature in the IS
$\widetilde{\mathcal{Y}}$	discrete state space of the GS
$\widetilde{\mathcal{X}} = \widetilde{\mathcal{R}} \times \widetilde{\mathcal{F}} \times \widetilde{\mathcal{P}} \times \widetilde{\mathcal{Y}}$	discrete state space of the generic state variable
$\mathcal{T}_{t,T}$	set of stopping times valued in $[t,T]$
$\mathcal{L}^\nu = \overline{\mathcal{L}}^\nu + \widehat{\mathcal{L}}$	generator of the state process
X^ε	perturbed generic state variable
$\mathcal{X}^\varepsilon = \mathcal{R} \times \mathcal{F} \times \mathcal{P}^\varepsilon \times \mathcal{Y}^\varepsilon$	perturbed state space of the generic state variable
$\sigma^\varepsilon \in \mathbb{R}^{(l+3)\times(l+3)}$	perturbed non-singular volatility matrix
$t_n = n\Delta_N, n = 0,1,\ldots,N$	discrete time points with N the number of time steps
$\Delta_N = T/N = t_{n+1} - t_n$	step size
$u_n = u(t_n)$	approximation of the decision rule at discrete time n
$\mathcal{E} = (\mathcal{E}_n)_{n=1,\ldots,N}$	sequence of i.i.d random variables with values in $\mathbb{R}^3$
$\mathcal{F}_n = \sigma(\{\mathcal{E}_1, \ldots, \mathcal{E}_n\})$	sigma-algebra generated by $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$
$\mathbb{F} = (\mathcal{F}_n)_{n=0,1,\ldots,N}$	discrete-time filtration with $\mathcal{F}_0 = \{\emptyset, \Omega\}$ the trivial sigma-algebra
$R_n = R(t_n) \in \mathcal{R}$	residual demand at discrete time n
$F_n = F(t_n) \in \mathcal{F}$	fuel price at discrete time n
$P_n = P(t_n) \in \mathcal{P}$	average temperature in the IS at discrete time n
$\widetilde{Y}_n = \widetilde{Y}(t_n) \in \mathcal{Y}$	reduced-order state of the GS discrete time n
$X_n = (R_n, F_n, P_n, \widetilde{Y}_n) \in \mathcal{X}$	generic state process at discrete time n
$R^D_n \in \widetilde{\mathcal{R}}$	approximation of the residual demand at discrete time n
$F^D_n \in \widetilde{\mathcal{F}}$	approximation of the fuel price at discrete time n
$P^D_n \in \widetilde{\mathcal{P}}$	approximation average temperature in IS at discrete time n
$\widetilde{Y}^D_n \in \widetilde{\mathcal{Y}}$	approximation of the reduced-order state of the GS at discrete time n
$X^D_n = (R^D_n, F^D_n, P^D_n, \widetilde{Y}^D_n)$	approximation of the generic state process at discrete time n
$\mathcal{N}(0,1)$	standard normal random variable
$(\zeta^\dagger_1, \ldots, \zeta^\dagger_N)$,	i.i.d sequence of standard normal random variables with $\dagger = R, F, P$

$m_{\dagger,n},\ \Sigma^2_{\dagger,n}$	conditional mean and variance of the random variable $\dagger_{n+1}$, with $\dagger = R, F, P, \widetilde{Y}$
$X_{n+1} = \mathcal{T}_n(X_n, u_n, \mathcal{E}_{n+1})$	transition operator with $\mathcal{T}_n : \mathcal{X} \times \overline{\mathcal{U}} \times \mathbb{R}^3 \to \mathcal{K}$
$\rho_{RP,n}$	conditional correlation between R_{n+1} and P_{n+1}
$\Sigma^2_{RP,n}$	conditional covariance of R_{n+1} and P_{n+1}
$\varphi_R,\ \Phi$	density and cumulative distribution function
φ_{RP}	joint conditional density function of R_{n+1} and P_{n+1}
$\mathcal{U}_P(n,x)$	set of feasible actions related to the state constraint to P
$\mathcal{U}_Y(n,x)$	set of feasible actions related to the state constraint to $\widetilde{Y}$
$\mathcal{U}_P(n,x) \cap \mathcal{U}_Y(n,x)$	state-dependent control constraints for MDP
$\langle , \rangle$	scalar product
$v_1, v_2, \ldots, v_{\ell-1}$	basis vectors for the state space $\mathcal{Y}$
$\mathcal{N}_\dagger = \{0, 1, ..., N_\dagger\}$	set of indices for R, P and $\widetilde{Y}^k$, with $\dagger = r, p, y_1, \ldots, y_\ell$
$\mathcal{N}_r \times \mathcal{N}_p \times \mathcal{N}_{y_1} \times \ldots \times \mathcal{N}_{y_\ell}$	generic index set the in $(\ell+2)$-dimensional discretized apace
$x_m = (r_i, p_j, y^1_{k_1}, y^2_{k_2}, \ldots, y^\ell_{k_\ell})$	a point in the $(\ell+2)$-dimensional discretized apace $\widetilde{\mathcal{X}}$
$V(n,x_m),\ u_n = u_n(x_m)$	approximate value function and decision rule in the grid point $x_m = (r_i, p_j, y^1_{k_1}, y^2_{k_2}, \ldots, y^\ell_{k_\ell}) \in \widetilde{\mathcal{X}}$ at time n
$\mathbf{P}^{\nu}_{x_{m_1}, x_{m_2}}$	probability that the state moves from x_{m_1} at time n to x_{m_2} at time $n+1$ under the action $u_n = \nu$
$\mathcal{B}_{\dagger_i}$	$\delta_\dagger$ neighborhood of $\dagger_i,\ i = 0, 2, \ldots, N_\dagger$ with $\dagger = r, p$
$\mathcal{B}_{ij} = \mathcal{B}_{r_i} \times \mathcal{B}_{p_j}$	neighborhood of (r_i, p_j), $i \in \mathcal{N}_r$ and $j \in \mathcal{N}_p$
$\mathcal{B}y^i_{k_i}$	neighborhood of $y^i_{k_i},\ k_i \in \mathcal{N}_{y_i}$

List of Figures

List of Tables

Bibliography

[1] E. Abel. Low-energy buildings. *Energy and Buildings*, 21(3):169 – 174, 1994.

[2] Asma Al-Tamimi, Frank L Lewis, and Murad Abu-Khalaf. Discrete-time nonlinear hjb solution using approximate dynamic programming: Convergence proof. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 38(4):943–949, 2008.

[3] David Amsallem and Charbel Farhat. Stabilization of projection-based reduced-order models. *International Journal for Numerical Methods in Engineering*, 91(4):358–377, 2012.

[4] Jesper Fink Andersen, Anders Reenberg Andersen, Murat Kulahci, and Bo Friis Nielsen. A numerical study of markov decision process algorithms for multi-component replacement problems. *European Journal of Operational Research*, 299(3):898–909, 2022.

[5] Athanasios C Antoulas. *Approximation of large-scale dynamical systems*. SIAM, 2005.

[6] Athanasios C Antoulas, Roxana Ionutiu, Nelson Martins, E Jan W ter Maten, Kasra Mohaghegh, Roland Pulch, Joost Rommes, Maryam Saadvandi, and Michael Striebel. Model order reduction: methods, concepts and properties. *Coupled multiscale simulation and optimization in nanoelectronics*, pages 159–265, 2015.

[7] Pablo Arce, Marc Medrano, Antoni Gil, Eduard Oró, and Luisa F Cabeza. Overview of thermal energy storage potential energy savings and climate change mitigation in spain and europe. *Applied Energy*, 88(8):2764–2774, 2011.

[8] Martin Bähr and Michael Breuß. Efficient long-term simulation of the heat equation with application in geothermal energy storage. *Mathematics*, 10(13):2309, 2022.

[9] Martin Bähr, Michael Breuß, and Ralf Wunderlich. Fast explicit diffiusion for long-time integration of parabolic problems. In *AIP Conference Proceedings*, volume 1863, page 410002. AIP Publishing, 2017.

[10] Daniel Bauer and Hongjun Ha. A Least-Squares Monte Carlo approach to the calculation of capital requirements. In *World Risk and Insurance Economics Congress, Munich, Germany, August*, volume 6, pages 2–6, 2015.

[11] Nicole Bäuerle and Ulrich Rieder. *Markov decision processes with applications to finance*. Springer Science & Business Media, 2011.

[12] Nicole Bäuerle and Viola Riess. Gas storage valuation with regime switching. *Energy Systems*, 7(3):499–528, 2016.

[13] Shahab Bazri, Irfan Anjum Badruddin, Abdullah Yousuf Usmani, Saleem Anwar Khan, Sarfaraz Kamangar, Mohammad Sajad Naghavi, Abdul Rahman Mallah, and Ali H Abdelrazek. Thermal hysteresis analysis of finned-heat-pipe-assisted latent heat thermal energy storage application for solar water heater system. *Case Studies in Thermal Engineering*, 40:102490, 2022.

[14] Christopher Beattie, Serkan Gugercin, and Volker Mehrmann. Model reduction for systems with inhomogeneous initial conditions. *Systems & Control Letters*, 99:99–106, 2017.

[15] Richard Bellman. An introduction to the theory of dynamic programming. Technical report, RAND CORP SANTA MONICA CA, 1953.

[16] Richard Bellman. The theory of dynamic programming. *Bulletin of the American Mathematical Society*, 60(6):503–515, 1954.

[17] Richard Bellman. Dynamic programming. *Princeton, USA: Princeton University Press*, 1(2):3, 1957.

[18] Richard Bellman and Robert E Kalaba. *Dynamic programming and modern control theory*, volume 81. Citeseer, 1965.

[19] Richard Bellman and E Stanley Lee. Functional equations in dynamic programming. *Aequationes mathematicae*, 17(1):1–18, 1978.

[20] Peter Benner and Enrique S Quintana-Ortí. Solving stable generalized Lyapunov equations with the matrix sign function. *Numerical Algorithms*, 20(1):75–100, 1999.

[21] Peter Benner, Enrique S Quintana-Ortí, and Gregorio Quintana-Ortí. Balanced truncation model reduction of large-scale dense systems on parallel computers. *Mathematical and Computer Modelling of Dynamical Systems*, 6(4):383–405, 2000.

[22] Peter Benner, Volker Mehrmann, and Danny C Sorensen. *Dimension reduction of large-scale systems*, volume 45. Springer, 2005.

[23] Peter Benner, Jing-Rebecca Li, and Thilo Penzl. Numerical solution of large-scale Lyapunov equations, Riccati equations, and linear-quadratic optimal control problems. *Numerical Linear Algebra with Applications*, 15(9):755–777, 2008.

[24] Peter Benner, Patrick Kürschner, and Jens Saak. Self-generating and efficient shift parameters in adi methods for large Lyapunov and Sylvester equations. *Electronic Transactions on Numerical Analysis (ETNA)*, 43:142–162, 2014.

[25] Dimitri Bertsekas and John N Tsitsiklis. *Neuro-dynamic programming*. Athena Scientific, 1996.

[26] Bart Besselink, Umut Tabak, Agnieszka Lutowska, Nathan Van de Wouw, H Nijmeijer, Daniel J Rixen, ME Hochstenbach, and WHA Schilders. A comparison of model reduction techniques from structural dynamics, numerical mathematics and systems and control. *Journal of Sound and Vibration*, 332(19):4403–4422, 2013.

[27] Ralph Byers, Chunyang He, and Volker Mehrmann. The matrix sign function method and the computation of invariant subspaces. *SIAM Journal on Matrix Analysis and Applications*, 18(3):615–632, 1997.

[28] Yongyang Cai and Kenneth L Judd. Advances in numerical dynamic programming and new applications. In *Handbook of computational economics*, volume 3, pages 479–516. Elsevier, 2014.

[29] X. Chen, L. Wang, L. Tong, S. Sun, X. Yue, S. Yin, and L. Zheng. Energy saving and emission reduction of China's urban district heating. *Energy Policy*, 55:677 – 682, 2013.

[30] Zhuliang Chen and Peter A Forsyth. A semi-Lagrangian approach for natural gas storage valuation and optimal operation. *SIAM Journal on Scientific Computing*, 30(1):339–368, 2007.

[31] Zhuliang Chen and Peter A Forsyth. Implications of a regime-switching model on natural gas storage valuation and optimal operation. *Quantitative Finance*, 10(2):159–176, 2010.

[32] Abdulrahman Dahash, Fabian Ochs, Alice Tosatto, and Wolfgang Streicher. Toward efficient numerical modeling and analysis of large-scale thermal energy storage for renewable district heating. *Applied Energy*, 279:115840, 2020.

[33] Adnan Daraghmeh, Carsten Hartmann, and Naji Qatanani. Balanced model reduction of linear systems with nonzero initial conditions: Singular perturbation approximation. *Applied Mathematics and Computation*, 353:295–307, 2019.

[34] John M Davis, Ian A Gravagne, Billy J Jackson, and Robert J Marks II. Controllability, observability, realizability, and stability of dynamic linear systems. *Electronic Journal of Differential Equations*, 2009(37):1–32, 2009.

[35] Mariana de Almeida Costa, Joaquim Pedro de Azevedo Peixoto Braga, and Antonio Ramos Andrade. A data-driven maintenance policy for railway wheelset based on survival analysis and markov decision process. *Quality and Reliability Engineering International*, 37(1):176–198, 2021.

[36] Bart De Schutter. Minimal state-space realization in linear system theory: an overview. *Journal of computational and applied mathematics*, 121(1-2):331–354, 2000.

[37] Aymeric Dieuleveut, Nicolas Flammarion, and Francis Bach. Harder, better, faster, stronger convergence rates for least-squares regression. *The Journal of Machine Learning Research*, 18(1):3520–3570, 2017.

[38] Ibrahim Dincer and Marc A Rosen. *Thermal energy storage: systems and applications*. John Wiley & Sons, 2021.

[39] Daniel J Duffy. *Finite Difference Methods in Financial Engineering: a Partial Differential Equation Approach*. John Wiley & Sons, 2013.

[40] Edwin J Elton and Martin J Gruber. Dynamic programming applications in finance. *The journal of finance*, 26(2):473–506, 1971.

[41] Dale F Enns. Model reduction with balanced realizations: An error bound and a frequency weighted generalization. In *The 23rd IEEE conference on decision and control*, pages 127–132. IEEE, 1984.

[42] Bérenger Favre and Bruno Peuportier. Application of dynamic programming to study load shifting in buildings. *Energy and Buildings*, 82:57–64, 2014.

[43] Eugene A Feinberg and Adam Shwartz. *Handbook of Markov decision processes: methods and applications*, volume 40. Springer Science & Business Media, 2012.

[44] Wendell H Fleming and Raymond W Rishel. *Deterministic and stochastic optimal control*, volume 1. Springer Science & Business Media, 2012.

[45] Wendell H Fleming and Halil Mete Soner. *Controlled Markov Processes and Viscosity Solutions*, volume 25. Springer Science and Business Media, 2006.

[46] Roland W Freund. Krylov-subspace methods for reduced-order modeling in circuit simulation. *Journal of Computational and Applied Mathematics*, 123(1-2):395–421, 2000.

[47] Rüdiger Frey, Abdelali Gabih, and Ralf Wunderlich. Portfolio optimization under partial information with expert opinions: a dynamic programming approach. *Communications on Stochastic Analysis*, 8(1):5, 2014.

[48] Keith Glover. All optimal hankel-norm approximations of linear multivariable systems and their L^∞-error bounds. *International journal of control*, 39(6):1115–1193, 1984.

[49] Ion Victor Gosea, Mihaly Petreczky, Athanasios C Antoulas, and Christophe Fiter. Balanced truncation for linear switched systems. *Advances in Computational Mathematics*, 44(6):1845–1886, 2018.

[50] Elisa Guelpa and Vittorio Verda. Thermal energy storage in district heating and cooling systems: A review. *Applied Energy*, 252:113474, 2019.

[51] Serkan Gugercin and Athanasios C Antoulas. A survey of model reduction by balanced truncation and some new results. *International Journal of Control*, 77(8):748–766, 2004.

[52] Lajos Gergely Gyurkó, Ben M Hambly, and Jan Hendrik Witte. Monte Carlo methods via a dual approach for some discrete time stochastic control problems. *Mathematical Methods of Operations Research*, 81(1):109–135, 2015.

[53] Wolfgang Hackbusch. A sparse matrix arithmetic based on H-matrices. part I: Introduction to h-matrices. *Computing*, 62(2):89–108, 1999.

[54] Nikolaos Halidias and PE Kloeden. A note on strong solutions of stochastic differential equations with a discontinuous drift coefficient. *Journal of Applied Mathematics and Stochastic Analysis*, 2006, 2006.

[55] Sven J Hammarling. Numerical solution of the stable, non-negative definite lyapunov equation lyapunov equation. *IMA Journal of Numerical Analysis*, 2(3):303–323, 1982.

[56] Hafiz MKU Haq, Birgitta Martinkauppi, Erkki Hiltunen, and Timo Sivula. Simulated thermal response test for ground heat storage: Numerical and analytical modeling of borehole. In *2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA)*, pages 291–296. IEEE, 2016.

[57] Matthias Heinkenschloss, Timo Reis, and Athanasios C Antoulas. Balanced truncation model reduction for systems with inhomogeneous initial conditions. *Automatica*, 47(3): 559–564, 2011.

[58] P. Heiselberg, H. Brohus, A. Hesselholt, H. Rasmussen, E. Seinre, and S. Thomas. Application of sensitivity analysis in design of sustainable buildings. *Renewable Energy*, 34 (9):2030 – 2036, 2009.

[59] Gregor P Henze, Robert H Dodier, and Moncef Krarti. Development of a predictive optimal controller for thermal energy storage systems. *HVAC&R Research*, 3(3):233–264, 1997.

[60] A Scottedward Hodel and Kameshwar Poolla. Parallel solution of large Lyapunov equations. *SIAM Journal on Matrix Analysis and Applications*, 13(4):1189–1203, 1992.

[61] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge University Press, 2012.

[62] Sumon Hossain et al. Efficient solution of Lyapunov equation for descriptor system and application to model order reduction. 2017.

[63] K. M. M. Huq, M. E. Baran, S. Lukic, and O. E. Nare. An energy management system for a community energy storage system. In *2012 IEEE Energy Conversion Congress and Exposition (ECCE)*, pages 2759–2763, Sept 2012. doi: 10.1109/ECCE.2012.6342532.

[64] Imad M Jaimoukha and Ebrahim M Kasenally. Krylov subspace methods for solving large lyapunov equations. *SIAM Journal on Numerical Analysis*, 31(1):227–251, 1994.

[65] Y. Kitapbayev, J. Moriarty, and P. Mancarella. Stochastic control and real options valuation of thermal storage-enabled demand response from flexible district energy systems. *Applied Energy*, 137:823 – 831, 2015.

[66] Nikolai V Krylov. Some new results in the theory of controlled diffusion processes. *Mathematics of the USSR-Sbornik*, 37(1):133, 1980.

[67] Patrick Kürschner. Balanced truncation model order reduction in limited time intervals for large systems. *Advances in Computational Mathematics*, 44(6):1821–1844, 2018.

[68] Gunther Leobacher, Michaela Szölgyenyi, and Stefan Thonhauser. On the existence of solutions of a class of sdes with discontinuous drift and singular diffusion. *Electronic Communications in Probability*, 20:1–14, 2015.

[69] L. Levron, J. Guerrero, and Y. Beck. Optimal power flow in microgrids with energy storage. *IEEE trans on Power Systems*, 28:3226–3234, 2013.

[70] Hong Li, Kun Ji, Ye Tao, and Chun'an Tang. Modelling a novel scheme of mining geothermal energy from hot dry rocks. *Applied Sciences*, 12(21):11257, 2022.

[71] Francis A Longstaff and Eduardo S Schwartz. Valuing american options by simulation: a simple least-squares approach. *The review of financial studies*, 14(1):113–147, 2001.

[72] Javad Mahmoudimehr and Leila Loghmani. Optimal management of a solar power plant equipped with a thermal energy storage system by using dynamic programming method. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, 230(2):219–233, 2016.

[73] Márton Major, Søren Erbs Poulsen, and Niels Balling. A numerical investigation of combined heat storage and extraction in deep geothermal reservoirs. *Geothermal Energy*, 6(1):1–16, 2018.

[74] Xuerong Mao. *Stochastic differential equations and applications*. Elsevier, 2007.

[75] D Mayne. An elementary derivation of rosenbrock's minimal realization algorithm. *IEEE Transactions on Automatic Control*, 18(3):306–307, 1973.

[76] Volker Mehrmann and Tatjana Stykel. Balanced truncation model reduction for large-scale systems in descriptor form. In *Dimension Reduction of Large-Scale Systems*, pages 83–115. Springer, 2005.

[77] Bruce Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE transactions on automatic control*, 26(1):17–32, 1981.

[78] C Mullis and RA Roberts. Synthesis of minimum roundoff noise fixed point digital filters. *IEEE Transactions on Circuits and Systems*, 23(9):551–562, 1976.

[79] S. Nielsen and B. Möller. Excess heat production of future net zero energy buildings within district heating areas in Denmark. *Energy*, 48(1):23 – 31, 2012.

[80] Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

[81] Hacer Öz Bakan, Fikriye Yilmaz, and Gerhard-Wilhelm Weber. An efficient algorithm for stochastic optimal control problems by means of a least-squares monte-carlo method. *Optimization*, 71(11):3133–3146, 2022.

[82] Thilo Penzl. A cyclic low-rank smith method for large sparse Lyapunov equations. *SIAM Journal on Scientific Computing*, 21(4):1401–1418, 1999.

[83] Lars Pernebo and Leonard Silverman. Model reduction via balanced state space representations. *IEEE Transactions on Automatic Control*, 27(2):382–387, 1982.

[84] Lars Pernebo and Leonard Silverman. Model reduction via balanced state space representations. *IEEE Transactions on Automatic Control*, 27(2):382–387, 1982.

[85] Mihály Petreczky, Rafael Wisniewsk, and John Leth. Theoretical analysis of balanced truncation for linear switched systems. *IFAC Proceedings Volumes*, 45(9):240–247, 2012.

[86] Mihály Petreczky, Rafael Wisniewski, and John Leth. Balanced truncation for linear switched systems. *Nonlinear Analysis: Hybrid Systems*, 10:4–20, 2013.

[87] Huyen Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science and Business Media, 2009.

[88] Warren B Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality*, volume 703. John Wiley & Sons, 2007.

[89] Martin L Puterman. Markov decision processes. *Handbooks in operations research and management science*, 2:331–434, 1990.

[90] Martin L Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[91] Martin Redmann. *Balancing related model order reduction applied to linear controlled evolution equations with Lévy Noise*. PhD thesis, Otto-von-Guericke Universität Magdeburg, 2016.

[92] G Regnier, P Salinas, C Jacquemyn, and MD Jackson. Numerical simulation of aquifer thermal energy storage using surface-based geologic modelling and dynamic mesh optimisation. *Hydrogeology Journal*, 30(4):1179–1198, 2022.

[93] Youcef Saad. Numerical solution of large lyapunov equations. 1989.

[94] B. Saeb-Gilani, B. Giorgi, M. Bachmann, and M. Kriegel, editors. *Potential analysis of heat sharing at different temperature levels in a district*, volume 3, CLIMA 2016 - proceedings of the 12th REHVA World Congress, 9 2016. Department of Civil Engineering, Aalborg University.

[95] Henrik Sandberg and Anders Rantzer. Balanced truncation of linear time-varying systems. *IEEE Transactions on automatic control*, 49(2):217–229, 2004.

[96] JM Sanz-Serna and C Palencia. A general equivalence theorem in the theory of discretization methods. *Mathematics of computation*, 45(171):143–152, 1985.

[97] Wilhelmus HA Schilders, Henk A Van der Vorst, and Joost Rommes. *Model order reduction: theory, research aspects and applications*, volume 13. Springer, 2008.

[98] Christian Schröder and Matthias Voigt. Balanced truncation model reduction with a priori error bounds for lti systems with nonzero initial value. *Journal of Computational and Applied Mathematics*, page 114708, 2022.

[99] Anton A Shardin and Ralf Wunderlich. Partially observable stochastic optimal control problems for an energy storage. *Stochastics*, 89(1):280–310, 2017.

[100] Shahriar Shokoohi, L Silverman, and Paul Van Dooren. Linear time-variable systems: Balancing and model reduction. *IEEE Transactions on Automatic Control*, 28(8):810–822, 1983.

[101] Jennie Si, Andrew G Barto, Warren B Powell, and Don Wunsch. *Handbook of learning and approximate dynamic programming*, volume 2. John Wiley & Sons, 2004.

[102] Valeria Simoncini. A new iterative method for solving large-scale lyapunov matrix equations. *SIAM Journal on Scientific Computing*, 29(3):1268–1288, 2007.

[103] Majid Soltani, F Moradi Kashkooli, AR Dehghani-Sanij, A Nokhosteen, A Ahmadi-Joughi, K Gharali, SB Mahbaz, and MB Dusseault. A comprehensive review of geothermal energy evolution and development. *International Journal of Green Energy*, 16(13): 971–1009, 2019.

[104] Marco Sorrentino, Kréhi Serge Agbli, Daniel Hissel, Frédéric Chauvet, and Tony Letrouve. Application of dynamic programming to optimal energy management of grid-independent hybrid railcars. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 235(2):236–247, 2021.

[105] Jerome L Stein. Applications of stochastic optimal control/dynamic programming to international finance and debt crises. *Nonlinear Analysis: Theory, Methods & Applications*, 63(5-7):e2033–e2041, 2005.

[106] Lars Stentoft. Convergence of the Least-Squares Monte Carlo approach to American option valuation. *Management Science*, 50(9):1193–1203, 2004.

[107] Tatjana Stykel. *Analysis and numerical solution of generalized Lyapunov equations*. PhD thesis, Institut für Mathematik, Technische Universität, Berlin, 2002.

[108] Tatjana Stykel and Timo Reis. Balanced truncation model reduction of second-order systems. 2007.

[109] Paul Honore Takam and Ralf Wunderlich. On the input-output behavior of a geothermal energy storage: Approximations by model order reduction. *arXiv preprint arXiv:2209.14761*, 2022.

[110] Paul Honore Takam, Ralf Wunderlich, and Olivier Menoukeu Pamen. Short-term behavior of a geothermal energy storage: Modeling and theoretical results. *arXiv preprint arXiv:2104.05005*, 2021.

[111] James William Thomas. *Numerical partial differential equations: finite difference methods*, volume 22. Springer Science Business Media, 2013.

[112] Henning Tischer and Gregor Verbic. Towards a smart home energy management system-a dynamic programming approach. In *2011 IEEE PES Innovative Smart Grid Technologies*, pages 1–7. IEEE, 2011.

[113] Michael S Tombs and Ian Postlethwaite. Truncated balanced realization of a stable non-minimal state-space system. *International Journal of Control*, 46(4):1319–1330, 1987.

[114] H. Tommerup and S. Svendsen. Energy savings in Danish residential building stock. *Energy and Buildings*, 38(6):618 – 626, 2006.

[115] H. Tommerup, J. Rose, and S. Svendsen. Energy-efficient houses built according to the energy performance requirements introduced in Denmark in 2006. *Energy and Buildings*, 39(10):1123 – 1130, 2007.

[116] Nizar Touzi. *Stochastic control problems, viscosity solutions and application to finance*. Scuola normale superiore, 2004.

[117] R. Urgaonkar, B. Urgaonkar, M. J. Neely, and A. Sivasubramaniam. Optimal power cost management using stored energy in data centers. In *Proceedings of the ACM SIGMETRICS Joint International Conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '11, pages 221–232. ACM, 2011.

[118] James M Varah. A lower bound for the smallest singular value of a matrix. *Linear Algebra and its applications*, 11(1):3–5, 1975.

[119] Andras Varga. Balancing free square-root algorithm for computing singular perturbation approximations. In *[1991] Proceedings of the 30th IEEE Conference on Decision and Control*, pages 1062–1065. IEEE, 1991.

[120] Richard S Varga. Geršgorin-type eigenvalue inclusion theorems. In *Geršgorin and his circles*, pages 35–72. Springer, 2004.

[121] A Yu Veretennikov. On stochastic equations with degenerate diffusion with respect to some of the variables. *Mathematics of the USSR-Izvestiya*, 22(1):173, 1984.

[122] Alexander Ju Veretennikov. On strong solutions and explicit formulas for solutions of stochastic integral equations. *Mathematics of the USSR-Sbornik*, 39(3):387, 1981.

[123] Stefan Volkwein. Proper orthogonal decomposition: Theory and reduced-order modelling. *Lecture Notes, University of Konstanz*, 4(4):1–29, 2013.

[124] P. Vytelingum, T. D. Voice, S. D. Ramchurn, A. Rogers, and N. R. Jennings. Agent-based micro-storage management for the smart grid. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1*, AAMAS '10, pages 39–46. International Foundation for Autonomous Agents and Multiagent Systems, 2010.

[125] Antony Ware. Accurate semi-Lagrangian time stepping for stochastic optimal control problems with application to the valuation of natural gas storage. *SIAM Journal on Financial Mathematics*, 4(1):427–451, 2013.

[126] Xiaohua Wu, Xiaosong Hu, Xiaofeng Yin, and Scott J Moura. Stochastic optimal energy management of smart home with pev energy storage. *IEEE Transactions on Smart Grid*, 9(3):2065–2075, 2016.

[127] Yangyang Wu, Dong Li, Ruitong Yang, Arıcı Müslüm, and Changyu Liu. Enhancing heat transfer and energy storage performance of shell-and-tube latent heat thermal energy storage unit with unequal-length fins. *Journal of Thermal Science*, pages 1–14, 2022.

[128] Belen Zalba, Jose Ma Marın, Luisa F Cabeza, and Harald Mehling. Review on thermal energy storage with phase change: materials, heat transfer analysis and applications. *Applied thermal engineering*, 23(3):251–283, 2003.

[129] Daniel Z Zanger. Convergence of a least-squares monte carlo algorithm for american option pricing with dependent sample data. *Mathematical Finance*, 28(1):447–479, 2018.

[130] Lizhi Zhang, Jiyuan Kuang, Bo Sun, Fan Li, and Chenghui Zhang. A two-stage operation optimization method of integrated energy systems with demand response and energy storage. *Energy*, 208:118423, 2020.

[131] Xicheng Zhang. Strong solutions of sdes with singular drift and sobolev diffusion coefficients. *Stochastic Processes and their Applications*, 115(11):1805–1818, 2005.

[132] Alexander K Zvonkin. A transformation of the phase space of a diffusion process that removes the drift. *Mathematics of the USSR-Sbornik*, 22(1):129, 1974.