A model for dynamic chance constraints in hydro power reservoir management *

L. Andrieu†, R. Henrion‡, W. Römisch§

January 18, 2009

2000 Mathematics Subject Classification. 90C15, 90B05.

Keywords: dynamic chance constraints, probabilistic constraints, multistage, water reservoir management

Abstract

In this paper, a model for (joint) dynamic chance constraints is proposed and applied to an optimization problem in water reservoir management. The model relies on discretization of the decision variables but keeps the probability distribution continuous. Our approach relies on calculating probabilities of rectangles which is particularly useful in the presence of independent random variables but works for a moderate number of stages equally well in case of correlated variables. Numerical results are provided for two and three stages.

1 Introduction

A conventional optimization problem under chance constraints is given by

$$\min \{f(x) \,|\mathbb{P}(g(x, \xi(\omega)) \geq 0) \geq p\}. \tag{1}$$

Here, $f : \mathbb{R}^n \to \mathbb{R}$ is some objective function, $g : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^m$ is a constraint mapping, ξ is some s-dimensional random vector on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and $p \in [0, 1]$ is some specified safety level. The meaning of the chance constraint is as follows: a decision

*This work was supported by the OSIRIS Department of *Electricité de France R&D* and by the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin

†Electricité de France, Recherche & Développement, 1 Avenue du Général de Gaulle, 92141 Clamart, France, E-Mail: laetitia.andrieu@edf.fr

‡Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, 10117 Berlin, Germany, E-Mail: henrion@wias-berlin.de

§Humboldt University Berlin, Department of Mathematics, Rudower Chaussee 25, 12489 Berlin, Germany, E-Mail: romisch@mathematik.hu-berlin.de

x is declared to be feasible, whenever the probability of satisfying the random inequality system

$$g_i(x, \xi(\omega)) \geq 0 \quad (i = 1, \ldots, m)$$

is at least p. Such constraints have importance in many engineering problems affected by random parameters the realization of which can be observed only after a (an optimal) decision has been taken. As basic references to theory, algorithms, stability and applications of chance constrained optimization problems we refer to [13], [14] and [5].

In the setting above, no possible temporal character of randomness is taken into account. The decision is determined once and for ever, thus making (1) a *static* model. Often, however, decision making is *dynamic*, i.e., it is a process in the course of which an increasing amount of information is available due to successive realization of random variables. Then, intuitively, one could expect better solutions to problem (1) when decisions also rely on previously observed realizations of the random parameter rather than ignoring them. This is a standard viewpoint in multistage stochastic optimization problems. These, however, after discretizing the probability distribution, have the privilege of not being burdened by all the difficulties inherent to chance constraints. It is not clear, whether recent advantages in integer programming methods for chance constraints (e.g., [8]) would allow to follow a similar way by discretizing the distribution and working with some appropriate scenario tree formulation of (1).

In this paper, we propose a semi-discrete (or semicontinuous) approach in that we suggest to discretize the dynamic decisions but not the distribution. This builds a bridge to existing methods for (continuous) static chance-constrained problems. The method will be applied to a single water reservoir management model. We keep the model as simple as possible in order to focus on aspects related to the dynamic character of chance constraints. The importance of chance constrained programming in the context of water reservoir management has been recognized a long time ago (see, e.g., the basic monograph [7] or [2], [3],[6],[9]). We emphasize that, in contrast to most related papers, we consider the more appropriate and more difficult case of joint chance constraints rather than individual ones which would allow for simple quantile-based reformulations of the chance constraints. The difference between the two models is that under joint constraints the probability is evaluated for the entire inequality system as in (1), whereas under individual constraints probabilities are computed for each inequality of the system separately. Though appealing by their easy application, individual chance constraints have the drawback, for instance when dealing with time-dependent constraints, of guaranteeing a certain probability level pointwise only, but usually one is interested in safety over a whole time horizon.

Dealing with joint probability distributions in the context of chance constraints has been pioneered by Prékopa and corresponding models for the control of water reservoirs are found, for instance, in the early papers [11] and [12]. Here already, a dynamic approach is used by taking into account previously observed data via conditional distributions. On the other hand, in this model, no reaction of future decisions on future observations is anticipated, so it is quasi-static. In our proposal, we adopt the viewpoint of multistage stochastic programs, where decisions at any future time stage are understood as functions of all previously observed random data (see, e.g., [10]). In contrast to the typical

alternating sequence of decisions and observations in multistage programming, we let this sequence always terminate by an observation (the last one). In this way the last decision made does not have the meaning of just reacting on the last observation made but rather reacts on the observation before last and hedges against the last observation. This point of view (see also [4]) is coherent with the original philosophy of chance constraints.

2 Dynamic Chance Constraints

Let $\xi := (\xi_1, \ldots, \xi_T)$ be an s-dimensional discrete time stochastic process on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$.Denote by

$$x = (x_1, x_2(\xi_1), x_3(\xi_1, \xi_2), \ldots, x_T(\xi_1, \ldots, \xi_{T-1}))$$

some adapted decision process, where the x_t are $\mathbb{R}^n$-valued. Furthermore, let

$$g : \mathbb{R}^{nT} \times \mathbb{R}^{sT} \to \mathbb{R}^m$$

be some constraint mapping. Then, we call the inequality

$$\mathbb{P}(g(x(\omega), \xi(\omega)) \geq 0) \geq p \quad (p \in [0,1])\,.$$

a *dynamic chance constraint*. Denoting by $\mu := \mathbb{P} \circ \xi^{-1}$ the law of ξ, we may rewrite the chance constraint as

$$\int\limits_{A(x)} \mu(dz) \geq p, \tag{2}$$

where

$$A(x) = \bigcap_{i=1,\ldots,m} \{z \mid g_i(x_1, \ldots, x_T(z_1, \ldots, z_{T-1}), z_1, \ldots, z_T) \geq 0\}.$$

Here, the unknowns are the functions $x_1, x_2(z_1), \ldots, x_T(z_1, \ldots, z_{T-1})$. Choosing all these functions as (unknown) constants $x_1, x_2, \ldots, x_T$, leads us back to the conventional static model of a chance constraint. Dealing with constraints like (2), where the variable x is of functional type seems to be very hard. Therefore, it is natural to discretize or parameterize x. More precisely, we put

$$x_t(z_1, \ldots, z_{t-1}) := \alpha_t\,(\pi_t, z_1, \ldots, z_{t-1}) \quad (t = 1, \ldots, T)\,, \tag{3}$$

where $\pi_t \in \mathbb{R}^{P_t}$ is some finite-dimensional parameter and $\alpha_t : \mathbb{R}^{P_t} \times \mathbb{R}^{s(t-1)} \to \mathbb{R}^n$ is a given mapping. Then, (2) turns into

$$\int\limits_{\tilde{A}(\pi)} \mu(dz) \geq p, \tag{4}$$

where

$$\begin{aligned}\tilde{A}(\pi) &= \bigcap_{i=1,\ldots,m} \{z \mid g_i(\pi_1, \alpha_2(\pi_2, z_1), \ldots, \alpha_T(\pi_T, z_1, \ldots, z_{T-1}), z_1, \ldots, z_T) \geq 0\} \\ &= \bigcap_{i=1,\ldots,m} \{z \mid \tilde{g}_i(\pi, z) \geq 0\}.\end{aligned}$$

In other words, (4) has the form of a conventional (static) chance constraints with a finite dimensional decision vector $\pi = (\pi_1, \ldots, \pi_T)$ representing the parameters of the decision functions in the dynamic chance constraint. The choice of an appropriate parameterization is crucial for an efficient numerical treatment of constraints like (4).

3 A simple single water reservoir model

In the following we consider a simple model of a single water reservoir designed for hydro power generation. By $\xi := (\xi_1, \ldots, \xi_T)$, we denote a discrete scalar random process indicating stochastic inflows (precipitation) to the reservoir at time periods $1, \ldots, T$. Associated with ξ, we consider a scalar, non-anticipative decision process

$$x = (x_1, x_2(\xi_1), x_3(\xi_1, \xi_2), \ldots, x_T(\xi_1, \ldots, \xi_{T-1}))$$

reflecting the policy of water release from the reservoir. Given an initial water level l_0 in the reservoir, we require that over the whole period of time the water level remains between some upper value l^* (flood reserve) and some lower value l_* (dead storage). Since the current water level is the sum of the initial level and the cumulative inflows until the given time minus the cumulative release at that time, one arrives at the following system of stochastic inequalities:

$$l_* \leq l_0 + \sum\nolimits_{i=1}^{t} \xi_i - \sum\nolimits_{i=1}^{t} x_i(\xi_1, \ldots, \xi_{i-1}) \leq l^* \quad (t = 1, \ldots, T). \tag{5}$$

As the realizations of future observations $\xi_t(\omega), \ldots, \xi_T(\omega)$ are not known when taking a decision at time t, the constraint system (5) is required to hold at least with some probability $p \in [0, 1]$. The released water is used to produce energy which is sold on the power market. The objective is to maximize the expected profit subject to satisfying the level constraints with a given probability. The energy produced E by releasing a quantity r of water is a function of the current water level l which may be approximated by the expression $r(cl + d)$ (see, e.g., [3]). Thus, the (stochastic) amount of energy produced in period t is given by

$$E_t(x, \xi) = x_t(\xi_1, \ldots, \xi_{t-1}) \left(c \left(l_0 + \sum\nolimits_{i=1}^{t-1} \xi_i - \sum\nolimits_{i=1}^{t-1} x_i(\xi_1, \ldots, \xi_{i-1}) \right) + d \right).$$

Now, the expected total amount of energy produced over the whole time period becomes

$$f(x) := \sum\nolimits_{t=1}^{T} \mathbb{E} E_t(x, \xi), \tag{6}$$

where '$\mathbb{E}$' denotes expectation. In addition to the level constraints (to be formulated as a chance constraint), we do not allow negative releases and we require some condition on the final filling level of the reservoir, namely that it coincides in expectation with the initial level (cycling constraint). Using the expression in (5) for the current filling level at $t = T$, this last constraint can be written as

$$\sum\nolimits_{i=1}^{T} \mathbb{E}\xi_i = \sum\nolimits_{i=1}^{T} \mathbb{E}x_i\left(\xi_1, \ldots, \xi_{i-1}\right)$$

Summarizing, we arrive at the following optimization problem under dynamic chance constraints:

$$\begin{aligned} &\max f(x) \quad \text{subject to} \\ &\mathbb{P}\left(l_* \leq l_0 + \sum\nolimits_{i=1}^{t} \xi_i - \sum\nolimits_{i=1}^{t} x_i\left(\xi_1, \ldots, \xi_{i-1}\right) \leq l^* \quad (t = 1, \ldots, T)\right) \geq p \\ &x_t(\xi_1, \ldots, \xi_{t-1}) \geq 0 \quad \mathbb{P}\text{-almost surely} \quad (t = 1, \ldots, T) \\ &\sum\nolimits_{i=1}^{T} \mathbb{E}\xi_i = \sum\nolimits_{i=1}^{T} \mathbb{E}x_i\left(\xi_1, \ldots, \xi_{i-1}\right) \end{aligned} \tag{7}$$

Evidently, the chance constraint in (7) is of type (2) with $n = s = 1$ (scalar processes) and $m = T$ (time constraints).

4 Discretization of the decision policies

In principle one could proceed as suggested in Section 2 and parameterize the decision policies according to (3), in order to arrive at a chance constraint of type (4) in a finite dimensional setting. It turns out, however, that for the convenience of numerical treatment, a simple transformation $x \to \psi$ is useful, which can easily be inverted after having calculated optimal transformed policies ψ. More precisely, we introduce the functions

$$\psi_t\left(z_1, \ldots, z_{t-1}\right) := l^* - l_0 + \sum\nolimits_{i=1}^{t} x_i\left(z_1, \ldots, z_{i-1}\right) - \sum\nolimits_{i=1}^{t-1} z_i \quad (t = 1, \ldots, T)\,. \tag{8}$$

Note, that for $t = 1$, one gets the deterministic (transformed) decision variable $\psi_1 = l^* - l_0 + x_1$. With this transformation , one can rewrite the chance constraint (7) as

$$\mathbb{P}\left(\psi_t\left(\xi_1, \ldots, \xi_{t-1}\right) + l_* - l^* \leq \xi_t \leq \psi_t\left(\xi_1, \ldots, \xi_{t-1}\right) \quad (t = 1, \ldots, T)\right) \geq p. \tag{9}$$

Now, we approximate the policies ψ_t by piecewise constant functions $\tilde{\psi}_t$. Accordingly we make the step function ansatz

$$\tilde{\psi}_t\left(z_1, \ldots, z_{t-1}\right) := \sum_{i_1, \ldots, i_{t-1}=1}^{N} a\left(i_1, \ldots, i_{t-1}\right) \chi_{Q(i_1, \ldots, i_{t-1})}\left(z_1, \ldots, z_{t-1}\right) \quad (t = 1, \ldots, T)\,, \tag{10}$$

where χ refers to the characteristic function and $Q\left(i_1, \ldots, i_{t-1}\right) \subseteq \mathbb{R}^{t-1}$ are disjoint subsets covering the whole space. For the case $t = 1$, we have $\tilde{\psi}_1 = a$ for some scalar a. Our aim

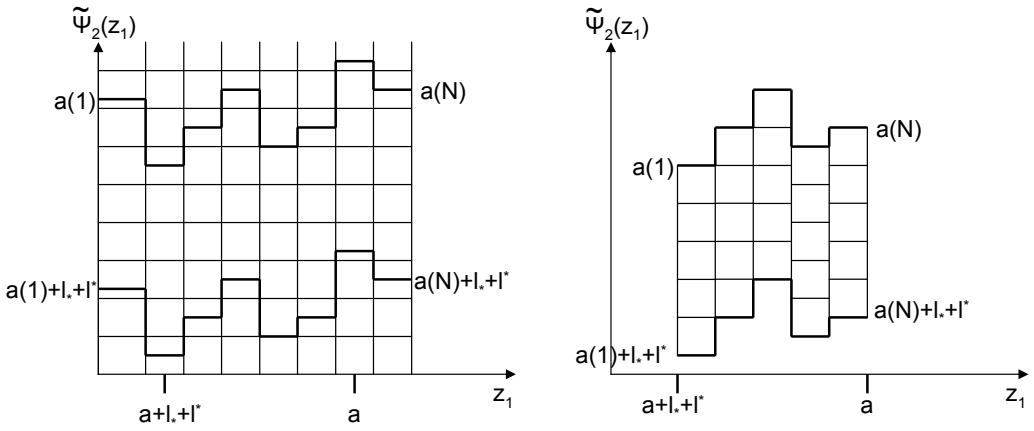

Figure 1: Illustration of the step functions $\tilde{\psi}_2$ defined on a fixed (left part) and a variable grid (right part).

is to substitute the infinite dimensional optimization problem (7) by a finite dimensional one which just depends on the parameter vector

$$\vec{a} = (a, a(1), \ldots, a(N), a(1,1), \ldots, a(N,N), \ldots, a(1,\ldots,1), \ldots, a(N,\ldots,N)). \quad (11)$$

defining the functions $\tilde{\psi}_t$ via (10). To start with, we replace the chance constraint (9) by the approximate chance constraint

$$\mathbb{P}\left(\tilde{\psi}_t(\xi_1,\ldots,\xi_{t-1}) + l_* - l^* \leq \xi_t \leq \tilde{\psi}_t(\xi_1,\ldots,\xi_{t-1}) \quad (t=1,\ldots,T)\right) \geq p. \quad (12)$$

Figure 1 (left part) illustrates the second ansatz function $\tilde{\psi}_2$ ($\tilde{\psi}_1$ reducing to the scalar a as mentioned before) defined by the step coefficients $a(1), \ldots, a(n)$ on some fixed grid. If the problem was just a two-stage one (i.e., $T = 2$), then (12) would mean that the probability of the set under the step function and between the limits $a + l_* - l^*$ and a be not smaller than p. Considering a fixed grid leads to certain difficulties due to the fact that the components of $\vec{a}$ will not lie on a grid point in the course of iterations which complicates the calculus of the desired probability. Moreover, for checking (12), it is not necessary that the step functions $\tilde{\psi}_t$ be determined on their whole domain spaces. It is sufficient to know them on the domain defined by the previous relations

$$\tilde{\psi}_\tau(\xi_1,\ldots,\xi_{\tau-1}) + l_* - l^* \leq \xi_\tau \leq \tilde{\psi}_\tau(\xi_1,\ldots,\xi_{\tau-1})$$

for $\tau < t$ (see right part of Figure 1). This leads us to propose a kind of dynamic grid adapted to the current values of the variable $\vec{a}$. To be more precise, given a discretization parameter $N \in \mathbb{N}$, let for $t = 1, \ldots, T$

$$\begin{aligned} Q(i_1,\ldots,i_t) &: = [r_{i_1}, r_{i_1+1}) \times [r_{i_1,i_2}, r_{i_1,i_2+1}) \times \cdots \times [r_{i_1,i_2,\ldots,i_t}, r_{i_1,i_2,\ldots,i_t+1}) \quad (13)\\ r_{i_1,i_2,\ldots,i_t} &: = a(i_1,\ldots,i_{t-1}) + l_* - l^* + \Delta_{i_t} \\ \Delta_i &: = (i-1)(l^* - l_*)/N \quad (i = 1,\ldots,N+1). \end{aligned}$$

Observe that $\Delta_1 = 0$ and $\Delta_{N+1} = l^* - l_*$, whence $r_{i_1,i_2,\ldots,i_{t-1},1} = a(i_1,\ldots,i_{t-2}) + l_* - l^*$ and $r_{i_1,i_2,\ldots,i_{t-1},N+1} = a(i_1,\ldots,i_{t-2})$. The two-stage grid defined by the half open rectangles

$$Q(i_1,i_2) := [r_{i_1}, r_{i_1+1}) \times [r_{i_1,i_2}, r_{i_1,i_2+1})$$

is illustrated in the right part of Figure 1. Of course, one might object that choosing a grid size depending on the (arbitrary) variable size of the intervals

$$[a(i_1,\ldots,i_{t-1}) + l_* - l^*, a(i_1,\ldots,i_{t-1})]$$

may not be the best idea from the numerical point of view, but we are not concerned with this detail in the present paper yet. Rather, we benefit from the conceptual advantage of the chosen variable grid in order to facilitate the calculus of (12). As a preparatory result we have:

Proposition 1 *A vector $z \in \mathbb{R}^T$ satisfies the inequality system*

$$\tilde{\psi}_t(z_1,\ldots,z_{t-1}) + l_* - l^* \le z_t < \tilde{\psi}_t(z_1,\ldots,z_{t-1}) \quad (t = 1,\ldots,T) \tag{14}$$

with $\tilde{\psi}_t$ as in (10) and $Q(i_1,\ldots,i_t)$ as in (13) if and only if

$$z \in \bigcup_{i_1=1}^{N} \cdots \bigcup_{i_T=1}^{N} Q(i_1,\ldots,i_T). \tag{15}$$

Proof. Let z satisfy (14). Recalling that $\tilde{\psi}_1 = a$, the first relation in (14) reads as

$$r_1 = a + l_* - l^* \le z_1 < a = r_{N+1}.$$

Consequently, there is some unique index $j_1 \in \{1,\ldots,N\}$ such that $z_1 \in [r_{j_1}, r_{j_1+1})$. Hence,

$$\tilde{\psi}_2(z_1) = \sum_{i_1=1}^{N} a(i_1)\,\chi_{Q(i_1)}(z_1) = \sum_{i_1=1}^{N} a(i_1)\,\chi_{[r_{i_1}, r_{i_1+1})}(z_1) = a(j_1).$$

Now, the second relation in (14) can be written as

$$r_{j_1,1} = a(j_1) + l_* - l^* \le z_2 < a(j_1) = r_{j_1,N+1}.$$

Again, there is some unique index $j_2 \in \{1,\ldots,N\}$ such that $z_2 \in [r_{j_1,j_2}, r_{j_1,j_2+1})$. Proceeding by induction, one derives the relations

$$z_t \in [r_{j_1,j_2,\ldots,j_t}, r_{j_1,j_2,\ldots,j_t+1}) \quad (t = 1,\ldots,T)$$

for certain indices $j_1, j_2, \ldots, j_T \in \{1,\ldots,N\}$. This shows, that z satisfies (15). The proof of the reverse implication runs the same way but the opposite direction. ■

Now, we are in a position to make the probability (12) explicit in terms of the coefficients contained in $\vec{a}$:

Theorem 2 *Let the random vector ξ have an absolutely continuous distribution. Then,*

$$\mathbb{P}\left(\tilde{\psi}_t\left(\xi_1,\ldots,\xi_{t-1}\right)+l_*-l^*\leq\xi_t\leq\tilde{\psi}_t\left(\xi_1,\ldots,\xi_{t-1}\right)\quad(t=1,\ldots,T)\right)=$$
$$\sum\nolimits_{i_1=1}^{N}\cdots\sum\nolimits_{i_{T-1}=1}^{N}\mathbb{P}\left(\xi\in A_{i_1}\left(\vec{a}\right)\times A_{i_1,i_2}\left(\vec{a}\right)\cdots\times A_{i_1,i_2,\ldots,i_{T-1}}\left(\vec{a}\right)\times B_{i_1,i_2,\ldots,i_{T-1}}\left(\vec{a}\right)\right),$$

where $\vec{a}$ refers to (11) and for $t=1,\ldots,T$, the one-dimensional intervals $A_{i_1,i_2,\ldots,i_t}\left(\vec{a}\right)$ and $B_{i_1,i_2,\ldots,i_t}$ are defined as

$$\begin{aligned}A_{i_1,i_2,\ldots,i_t}\left(\vec{a}\right)&:=\left[a\left(i_1,\ldots,i_{t-1}\right)+l_*-l^*+\Delta_{i_t},a\left(i_1,\ldots,i_{t-1}\right)+l_*-l^*+\Delta_{i_t+1}\right]\\B_{i_1,i_2,\ldots,i_t}\left(\vec{a}\right)&:=\left[a\left(i_1,\ldots,i_t\right)+l_*-l^*,a\left(i_1,\ldots,i_t\right)\right].\end{aligned}$$

Proof. We first observe that

$$\begin{aligned}\bigcup\nolimits_{i_1=1}^{N}\cdots\bigcup\nolimits_{i_T=1}^{N}Q\left(i_1,\ldots,i_T\right)&=\\\bigcup\nolimits_{i_1=1}^{N}\cdots\bigcup\nolimits_{i_{T-1}=1}^{N}\bigcup\nolimits_{i_T=1}^{N}&\\\left[r_{i_1},r_{i_1+1}\right)\times\cdots\times\left[r_{i_1,i_2,\ldots,i_{T-1}},r_{i_1,i_2,\ldots,i_{T-1}+1}\right)\times\left[r_{i_1,i_2,\ldots,i_T},r_{i_1,i_2,\ldots,i_T+1}\right)&=\\\bigcup\nolimits_{i_1=1}^{N}\cdots\bigcup\nolimits_{i_{T-1}=1}^{N}&\\\left[r_{i_1},r_{i_1+1}\right)\times\cdots\times\left[r_{i_1,\ldots,i_{T-1}},r_{i_1,i_2,\ldots,i_{T-1}+1}\right)\times\left\{\bigcup\nolimits_{i_T=1}^{N}\left[r_{i_1,i_2,\ldots,i_T},r_{i_1,i_2,\ldots,i_T+1}\right)\right\}&=\\\bigcup\nolimits_{i_1=1}^{N}\cdots\bigcup\nolimits_{i_{T-1}=1}^{N}&\\\left[r_{i_1},r_{i_1+1}\right)\times\cdots\times\left[r_{i_1,i_2,\ldots,i_{T-1}},r_{i_1,i_2,\ldots,i_{T-1}+1}\right)\times\left[r_{i_1,i_2,\ldots,i_{T-1},1},r_{i_1,i_2,\ldots,i_{T-1},N+1}\right).&\end{aligned}$$

Then, the absolute continuity of the distribution of ξ along with Proposition 1 yield:

$$\begin{aligned}\mathbb{P}\left(\tilde{\psi}_t\left(\xi_1,\ldots,\xi_{t-1}\right)+l_*-l^*\leq\xi_t\leq\tilde{\psi}_t\left(\xi_1,\ldots,\xi_{t-1}\right)\quad(t=1,\ldots,T)\right)&=\\\mathbb{P}\left(\tilde{\psi}_t\left(\xi_1,\ldots,\xi_{t-1}\right)+l_*-l^*\leq\xi_t<\tilde{\psi}_t\left(\xi_1,\ldots,\xi_{t-1}\right)\quad(t=1,\ldots,T)\right)&=\\\mathbb{P}\left(\xi\in\bigcup\nolimits_{i_1=1}^{N}\cdots\bigcup\nolimits_{i_T=1}^{N}Q\left(i_1,\ldots,i_T\right)\right)&=\\\mathbb{P}\left(\xi\in\bigcup\nolimits_{i_1=1}^{N}\cdots\bigcup\nolimits_{i_{T-1}=1}^{N}A_{i_1}\left(\vec{a}\right)\times A_{i_1,i_2}\left(\vec{a}\right)\cdots\times A_{i_1,i_2,\ldots,i_{T-1}}\left(\vec{a}\right)\times B_{i_1,i_2,\ldots,i_{T-1}}\left(\vec{a}\right)\right)&=\\\sum\nolimits_{i_1=1}^{N}\cdots\sum\nolimits_{i_{T-1}=1}^{N}\mathbb{P}\left(\xi\in A_{i_1}\left(\vec{a}\right)\times A_{i_1,i_2}\left(\vec{a}\right)\cdots\times A_{i_1,i_2,\ldots,i_{T-1}}\left(\vec{a}\right)\times B_{i_1,i_2,\ldots,i_{T-1}}\left(\vec{a}\right)\right).&\end{aligned}$$

Here, the last equality relies on the fact that for different $(T-1)$- tuples $(i_1,i_2,\ldots,i_{T-1})$, the sets

$$A_{i_1}\left(\vec{a}\right)\times A_{i_1,i_2}\left(\vec{a}\right)\cdots\times A_{i_1,i_2,\ldots,i_{T-1}}\left(\vec{a}\right)\times B_{i_1,i_2,\ldots,i_{T-1}}\left(\vec{a}\right)$$

are almost disjoint in the sense that they overlap only at sets of Lebesgue measure zero. ■

The chosen approach turns out to be particularly useful under the assumption of ξ having independent components because this allows to reduce the calculus of probabilities to one-dimensional distribution functions. As an immediate consequence of Theorem 2 we have:

Corollary 3 *Let the random vector ξ have an absolutely continuous distribution with independent components. Then,*

$$\mathbb{P}\left(\tilde{\psi}_t\left(\xi_1,\ldots,\xi_{t-1}\right)+l_*-l^*\leq\xi_t\leq\tilde{\psi}_t\left(\xi_1,\ldots,\xi_{t-1}\right)\quad(t=1,\ldots,T)\right)=$$
$$\sum\nolimits_{i_1=1}^{N}\cdots\sum\nolimits_{i_{T-1}=1}^{N}\left(\Phi_T\left(a\left(i_1,\ldots,i_{T-1}\right)\right)-\Phi_T\left(a\left(i_1,\ldots,i_{T-1}\right)+l_*-l^*\right)\right)\cdot$$
$$\prod\nolimits_{t=1}^{T-1}\left(\Phi_t\left(a\left(i_1,\ldots,i_{t-1}\right)+l_*-l^*+\Delta_{i_t+1}\right)-\Phi_t\left(a\left(i_1,\ldots,i_{t-1}\right)+l_*-l^*+\Delta_{i_t}\right)\right),$$

where Φ_t refers to the one-dimensional distribution function of the component ξ_t. In the two-stage situation $(T=2)$ one gets

$$\mathbb{P}\left(\tilde{\psi}_t\left(\xi_1,\ldots,\xi_{t-1}\right)+l_*-l^*\leq\xi_t\leq\tilde{\psi}_t\left(\xi_1,\ldots,\xi_{t-1}\right)\quad(t=1,2)\right)\;=$$
$$\sum\nolimits_{i=1}^{N}\left(\Phi_2\left(a\left(i\right)\right)-\Phi_2\left(a\left(i\right)+l_*-l^*\right)\right)\cdot$$
$$\left(\Phi_1\left(a+l_*-l^*+\Delta_{i+1}\right)-\Phi_1\left(a+l_*-l^*+\Delta_i\right)\right).$$

But also in the more realistic case of correlated components of ξ one may benefit from the obtained structure by the possibility to reduce rectangle probabilities to (now: multidimensional) distribution functions. At least for a small number T of stages this seems to be a viable procedure. We present a corresponding result for the special case of two stages:

Corollary 4 *Let the random vector ξ have an absolutely continuous distribution. In the two-stage situation $(T=2)$, one gets*

$$\mathbb{P}\left(\tilde{\psi}_t\left(\xi_1,\ldots,\xi_{t-1}\right)+l_*-l^*\leq\xi_t\leq\tilde{\psi}_t\left(\xi_1,\ldots,\xi_{t-1}\right)\quad(t=1,2)\right)\;=$$
$$\sum\nolimits_{i=1}^{N}\Phi\left(a+l_*-l^*+\Delta_{i+1},a\left(i\right)\right)-\Phi\left(a+l_*-l^*+\Delta_i,a\left(i\right)\right)+$$
$$\Phi\left(a+l_*-l^*+\Delta_i,a\left(i\right)+l_*-l^*\right)-\Phi\left(a+l_*-l^*+\Delta_{i+1},a\left(i\right)+l_*-l^*\right),$$

where Φ denotes the two-dimensional distribution function of $\xi=(\xi_1,\xi_2)$.

Proof. The proof follows immediately from Theorem 2 upon observing that, for any rectangle $[\alpha_1,\alpha_2]\times[\beta_1,\beta_2]$ one has

$$\mathbb{P}\left(\xi\in[\alpha_1,\alpha_2]\times[\beta_1,\beta_2]\right)=\Phi\left(\alpha_2,\beta_2\right)-\Phi\left(\alpha_1,\beta_2\right)+\Phi\left(\alpha_1,\beta_1\right)-\Phi\left(\alpha_2,\beta_1\right).$$

■

5 The finite dimensional optimization problem

In the last section, we were able via Theorem 2 to represent a finite dimensional approximation (in terms of the finite dimensional variable $\vec{a}$) of the infinite dimensional chance constraint in problem (7). To do so with the whole optimization problem (7), we have to express first the original unknown x in terms of $\vec{a}$ and then to substitute the corresponding expressions in the remaining constraints as well as in the objective of (7). At the same time, this will show us how to recover x from a solution $\vec{a}^*$ of the finite dimensional problem. Note that (8) is easily inverted in terms of x:

$$\begin{aligned} x_1 &= \psi_1 + l_0 - l^*, \\ x_t\left(z_1, \ldots, z_{t-1}\right) &= \psi_t\left(z_1, \ldots, z_{t-1}\right) - \psi_{t-1}\left(z_1, \ldots, z_{t-2}\right) + z_{t-1} \quad (t = 2, \ldots, T). \end{aligned}$$

Now, using the approximations $\tilde{\psi}_t$ of ψ_t, we get corresponding approximations $\hat{x}_t$ of x_t. Using (10), the explicit dependence of $\hat{x}_t$ on the unknown $\vec{a}$ is revealed:

$$\begin{aligned} \hat{x}_1 &= \tilde{\psi}_1 + l_0 - l^* = a + l_0 - l^*, \\ \hat{x}_t\left(z_1, \ldots, z_{t-1}\right) &= \tilde{\psi}_t\left(z_1, \ldots, z_{t-1}\right) - \tilde{\psi}_{t-1}\left(z_1, \ldots, z_{t-2}\right) + z_{t-1} \\ &= \sum_{i_1, \ldots, i_{t-1}=1}^{N} a\left(i_1, \ldots, i_{t-1}\right) \chi_{Q(i_1, \ldots, i_{t-1})}\left(z_1, \ldots, z_{t-1}\right) - \\ &\quad \sum_{i_1, \ldots, i_{t-2}=1}^{N} a\left(i_1, \ldots, i_{t-2}\right) \chi_{Q(i_1, \ldots, i_{t-2})}\left(z_1, \ldots, z_{t-2}\right) + z_{t-1} \\ &\quad (t = 2, \ldots, T). \end{aligned}$$

Observe, that $\hat{x}_t$ above is not globally defined as a consequence of the same fact holding true for the functions $\tilde{\psi}_t$. The latter were defined only on a subdomain corresponding to the region of interest of the chance constraint. This raises the question of how to extend them globally (which is important for instance, in order to calculate their expectations as needed in (7)). One could agree that the $\hat{x}_t$ remain free outside the mentioned subdomain thus serving the aim of maximizing the objective in (7). However, one could also adopt the viewpoint that arguments outside the subdomain represent violations of the filling level constraint, and in such cases no benefit in terms of profit from energy production should be allowed. This idea, which shall follow from now on, reflects some additional stimulus for not violating the filling level constraints which, in addition to the imposed chance constraint, acts in the same direction by not rewarding profit under constraint violation. Therefore, we put $\hat{x}_t \equiv 0$ outside the domain

$$\bigcup_{i_1=1}^{N} \cdots \bigcup_{i_{t-1}=1}^{N} Q\left(i_1, \ldots, i_{t-1}\right).$$

The computation of the expectation $\mathbb{E}\hat{x}_t$ is simple as far as the contributions of step functions to the definition of $\hat{x}_t$ are concerned. However, the remaining contribution coming from z_{t-1} is less obvious because it amounts to the expectation of ξ_t restricted to

the mentioned subdomain. Calculating this expectation exactly may be complicated or not even possible. To overcome this difficulty, we propose to approximate the functions z_{t-1} as step functions defined on the same grid as the $\hat{x}_t$:

$$\tilde{z}_t(z_1,\ldots,z_t) := \sum_{i_1,\ldots,i_t=1}^{N} \bar{a}(i_1,\ldots,i_t)\,\chi_{Q(i_1,\ldots,i_t)}(z_1,\ldots,z_t) \quad (t=1,\ldots,T).$$

Here, the coefficients $\bar{a}(i_1,\ldots,i_t)$ are chosen as the midpoints of the intervals

$$[r_{i_1,i_2,\ldots,i_t}, r_{i_1,i_2,\ldots,i_t+1}),$$

or stated explicitly (see (13)):

$$\begin{aligned}\bar{a}(i_1,\ldots,i_t) :=& \quad (r_{i_1,i_2,\ldots,i_t}, r_{i_1,i_2,\ldots,i_t+1})/2 \\ =& \quad a(i_1,\ldots,i_{t-1}) + l_* - l^* + \left(i_t - \frac{1}{2}\right)\frac{l^*-l_*}{N}. \end{aligned} \tag{16}$$

With this approximation $\tilde{z}_t$ of z_t, we arrive, starting from the previously obtained expressions for $\hat{x}_t$, at the following second approximation $\tilde{x}_t$ of x_t which finally is fully amenable to the desired formulation of our finite dimensional optimization problem:

$$\begin{aligned}\tilde{x}_1 &= a + l_0 - l^*, \\ \tilde{x}_t(z_1,\ldots,z_{t-1}) &= \sum_{i_1,\ldots,i_{t-1}=1}^{N} \left(a(i_1,\ldots,i_{t-1}) + \bar{a}(i_1,\ldots,i_{t-1})\right)\chi_{Q(i_1,\ldots,i_{t-1})}(z_1,\ldots,z_{t-1}) - \\ &\qquad \sum_{i_1,\ldots,i_{t-2}=1}^{N} a(i_1,\ldots,i_{t-2})\,\chi_{Q(i_1,\ldots,i_{t-2})}(z_1,\ldots,z_{t-2}) \quad (t=2,\ldots,T).\end{aligned}$$

Note that the characteristic functions $\chi_{Q(i_1,\ldots,i_{t-2})}$ themselves can be made explicit in terms of the unknown $\vec{a}$ via (13). Now, for instance, the nonnegativity constraint in (7) can be written as

$$\begin{aligned} a &\geq l^* - l_0 \\ \sum_{i_1,\ldots,i_{t-1}=1}^{N} \left(a(i_1,\ldots,i_{t-1}) + \bar{a}(i_1,\ldots,i_{t-1})\right)\chi_{Q(i_1,\ldots,i_{t-1})}(z_1,\ldots,z_{t-1}) &\geq \\ \sum_{i_1,\ldots,i_{t-2}=1}^{N} a(i_1,\ldots,i_{t-2})\,\chi_{Q(i_1,\ldots,i_{t-2})}(z_1,\ldots,z_{t-2}) \quad (t=2,\ldots,T).& \end{aligned}$$

Evaluating pointwise the last system of inequalities, we may invoke the definition of the rectangles $Q(i_1,\ldots,i_t)$ to derive the identity

$$\chi_{Q(i_1,\ldots,i_{t-1})}(z_1,\ldots,z_{t-1}) = \chi_{Q(i_1,\ldots,i_{t-2})}(z_1,\ldots,z_{t-2}) = 1 \quad \forall\,(z_1,\ldots,z_{t-1}) \in Q(i_1,\ldots,i_{t-1}).$$

Since the $Q(i_1,\ldots,i_{t-1})$ form a partition of the domain of $\tilde{x}_t$, we may further exploit (16) in order to reduce the relations above to the following explicit inequality system in the components of the unknown $\vec{a}$:

$$\begin{aligned} a &\geq l^* - l_0 \\ a(i_1,\ldots,i_{t-1}) &\geq l^* - l_* + \left(i_{t-1} - \frac{1}{2}\right)\frac{l_* - l^*}{N} \\ &\quad \forall (i_1,\ldots,i_{t-1}) \in \{1,\ldots,N\}^{t-1}\ \forall t \in \{2,\ldots,T\}. \end{aligned}$$

Next, we obtain the expectations

$$\begin{aligned} \mathbb{E}\tilde{x}_1 &= a + l_0 - l^* \\ \mathbb{E}\tilde{x}_t &= \sum_{i_1,\ldots,i_{t-1}=1}^{N} \left(a(i_1,\ldots,i_{t-1}) + \bar{a}(i_1,\ldots,i_{t-1})\right) \mathbb{P}\left((\xi_1,\ldots,\xi_{t-1}) \in Q(i_1,\ldots,i_{t-1})\right) - \\ &\quad \sum_{i_1,\ldots,i_{t-2}=1}^{N} a(i_1,\ldots,i_{t-2})\, \mathbb{P}\left((\xi_1,\ldots,\xi_{t-2}) \in Q(i_1,\ldots,i_{t-2})\right) \quad (t = 2,\ldots,T). \end{aligned}$$

Similarly to the calculus of the chance constraints, we are again faced with the need of computing probabilities of rectangles, this time, however, in one dimension less. If, for instance, $T = 2$, then the computation of $\mathbb{E}\tilde{x}_2$ requires just the probability of intervals etc. Again, as for the chance constraints, one might benefit from the assumption of independent components for larger values of T.

Now, that we are able to calculate expectations $\mathbb{E}\tilde{x}_t$, we can immediately rephrase the last constraint (cycling constraint) in (7) in terms of the unknown $\vec{a}$.

As far as the objective (6) is concerned, we approximate the terms $E_t(x)$ in the vein of the previous derivations as

$$E_t(x,z) \approx E_t(\tilde{x},z) = \tilde{x}_t(z_1,\ldots,z_{t-1})\left(c\left(l_0 + \sum\nolimits_{i=1}^{t-1} \tilde{z}_i(z_1,\ldots,z_i) - \sum\nolimits_{i=1}^{t-1} \tilde{x}_i(z_1,\ldots,z_{i-1})\right) + d\right).$$

Using the already developed step function representations for $\tilde{x}_t$ and $\tilde{z}_i$ one may again explicitly calculate the expectations $\mathbb{E}E_t(\tilde{x},\xi)$. Rather than deriving the involved general formula here, we present its expression for the two-stage case $T = 2$. Trivially,

$$\mathbb{E}E_1(\tilde{x},\xi) = \tilde{x}_1(cl_0 + d) = (a + l_0 - l^*)(cl_0 + d).$$

Next, using the step function representations for $\tilde{x}_2$ and $\tilde{z}_1$, we obtain

$$\begin{aligned} E_2(\tilde{x},z) &= \tilde{x}_2(z)\left(c\left(l_0 + \tilde{z}_1(z) - \tilde{x}_1\right) + d\right) \\ &= \left(\sum_{i=1}^{N} \left(a(i) + \bar{a}(i)\right)\chi_{Q(i)}(z) - a\right)\left(c\left(l_0 + \sum_{j=1}^{N} \bar{a}(j)\,\chi_{Q(j)}(z) - \tilde{x}_1\right) + d\right). \end{aligned}$$

The further development of this expression relies on the fact that the product of characteristic functions of two sets equals the characteristic function of the intersection of these sets and on the sets $Q(i)$ and $Q(j)$ being disjoint for $i \neq j$ (see (13)). One then finally arrives at

$$E_2(\tilde{x}, z) = \sum_{i=1}^{N} (a(i) + \bar{a}(i) - a)((c(l^* - a + \bar{a}(i)) + d))\chi_{Q(i)}(z).$$

Recalling from (13) that

$$Q(i) = [a + l_* - l^* + \Delta_i, a + l_* - l^* + \Delta_{i+1}),$$

we get

$$\begin{aligned} \mathbb{E}E_2(\tilde{x}, \xi) &= \sum_{i=1}^{N} (a(i) + \bar{a}(i) - a)((c(l^* - a + \bar{a}(i)) + d)) \cdot \\ &\quad (\Phi_1(a + l_* - l^* + \Delta_{i+1}) - \Phi_1(a + l_* - l^* + \Delta_i)), \end{aligned}$$

where Φ_1 refers to the one-dimensional distribution function of ξ_1.

6 Numerical Results

In this section, we illustrate our approach to dynamic chance constraints by numerical examples in a two-stage and three-stage setting. We consider the water reservoir problem presented in section 3 with the following data:

$$T = 2, l_* = 1, l^* = 3, p = 0.9, c = 2, d = 1, \xi \sim \mathcal{N}\left((1,1), \begin{pmatrix} 0.09 & 0 \\ 0 & 0.09 \end{pmatrix}\right),$$

the latter expression meaning that the random vector $\xi = (\xi_1, \xi_2)$ has a bivariate normal distribution with independent components each of them having mean 1 and variance 0.09 (or standard deviation 0.3). We take the initial (and expected final) filling level l_0 as a parameter moving between l_* and l^*. We solve the finite dimensional optimization problem formulated in section 5 with a discretization parameter $N = 160$. For dealing with the dynamic chance constraint, we make use of the second statement in Corollary 3. Similarly, we make use of the explicit representation of the objective developed for the two stage case at the end of section 5.

Figure 2 illustrates the obtained results. The left upper part provides a plot of optimal first stage release decisions x_1 as a function of the initial filling level l_0. The right upper part plots the optimal second stage release policies x_2 as functions of the first inflow realization ξ_1 for inititial filling levels l_0 reaching from 1.5 to 2.6 (curves from left to right). As can be seen, the optimal releas policies become piecewise linear for extreme initial levels whereas they are strongly nonlinear in between. The lower left diagram

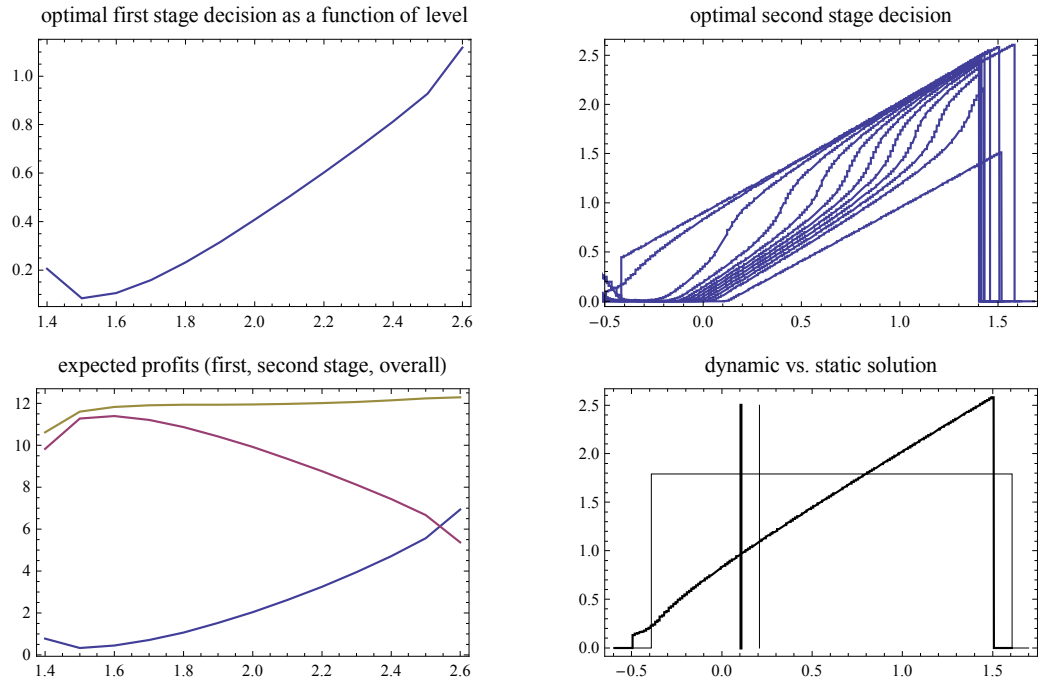

Figure 2: Illustration of numerical results for a two stage instance of the water reservoir problem assuming independent components of the random inflow. For details, see text.

plots the expected profits (from bottom to top) of the optimal first and second stage decisions as well as the overall expected profit as functions of l_0. Roughly speaking, the same overall expected profit can be realized over the whole interval of initial filling levels but with strongly varying contributions by the first and second order release decisions. Finally, the lower right diagram opposes the optimal static (thin lines) and dynamic (thick lines) solutions of the problem for a fixed $l_0 = 1.6$. The (scalar) first stage decisions are represented by vertical lines identifying the corresponding value on the first axis. In contrast to the dynamic release policy, the optimal second stage decisions in the static case reduces to a constant (on the domain of feasible operation).

It is obvious that a static solution cannot realize the same value of the objective function as the dynamic one, but has to be worse due to a lack of flexibility in its reaction on observations of first stage inflow. Indeed, the expected profit of the static solution amounts to 9.89 as opposed to an expected profit of 11.83 for the dynamic solution. The difference 11.83-9.89=1.94 can be interpreted as the *Value of the Dynamic Solution* (VDS). Taken as a relative figure, the improvement of the objective reaches approximately 16% in this example. Of course, for other data, even larger improvements can be expected.

If correlations between the random inflow components shall be taken into account, then only the computation of the chance constraint is affected. One can no longer rely on Corollary 3. Nevertheless, one may employ Corollary 4 which comes at the price of

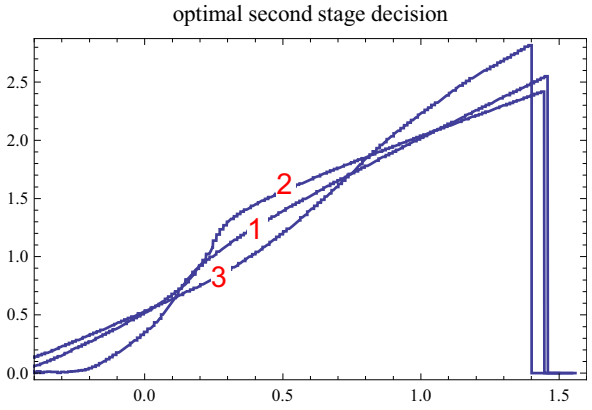

Figure 3: Second stage release policies for the previous example (at $l_0 = 1.7$) but with an assumed correlation between inflow components of $\rho = -0.3$ (1), $\rho = 0$ (2) and $\rho = 0.9$ (3), respectively.

calculating bivariate distribution functions. This, however, is the only change as compared to the case of independent components. Figure 3 illustrates the optimal second stage decision policies in the previous example for different correlation coefficients between the components of the inflow variables.

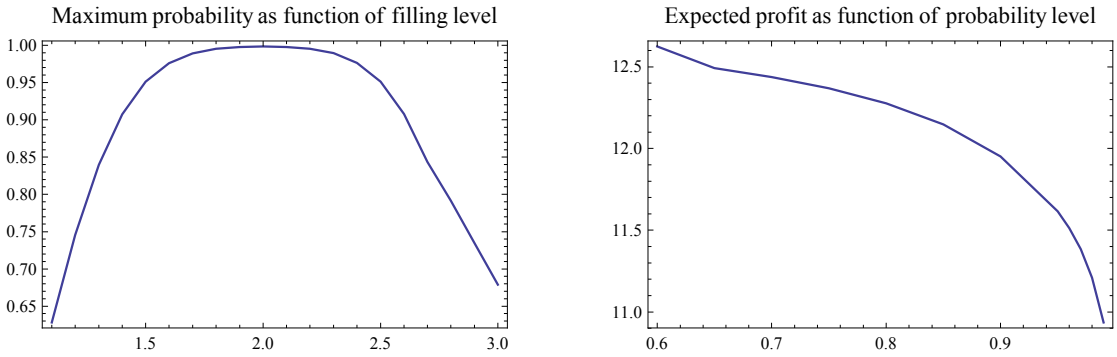

Figure 4: Maximum probability level p as a function of the initial filling level (left part) and expected profit as a function of the chosen probability level (right part)

Another issue of interest in optimization problems with chance constraints is the right choice of the probability level p. Of course, first of all, this is a matter of experience, insight into the problem and of the right interpretation of results. It has to be taken into account however, that increasing p corresponds to shrinking the feasible set of the problem. Accordingly, there may exist some maximum probability level p^* which can be realized at most under the remaining constraints of the problem. In other words, further increasing p leads to empty feasible sets then, which should be avoided in the

numerical treatment of the problem. Moreover, increasing p even below this maximum value will decrease the optimal value of the objective function, i.e., the expected profit, in the optimization problem. Both aspects are illustrated in Figure 4 for the two-stage case with uncorrelated components. It can be seen that, over a reasonably large range of probability levels, the expected profit decreases only moderately whereas it quickly collapses for excessively high probabilities. Such diagrams can help the decision maker to find a good compromise between safe operation and profit making.

Finally, we want to illustrate the three-stage case. To do so we consider again independent components, now of the three-dimensional random vector $\xi = (\xi_1, \xi_2, \xi_3)$. These components are supposed to have the same distribution parameters as before (normal distribution with expectation = 1 and standard deviation = 0.3). To simplify the representation and numerical computations, we restrict the objective of the problem to the first two stages which means, that the expected profit is calculated until the second stage only, wheras all constraints have to be satisfied for the whole of three stages. Due to the increasing complexity of computations, we restrict the discretization parameter to the value $N = 20$. The resulting plots of the optimal second stage (as a function of ξ_1) and third stage (as a function of ξ_1, ξ_2) release policies are provided in Figure 5.

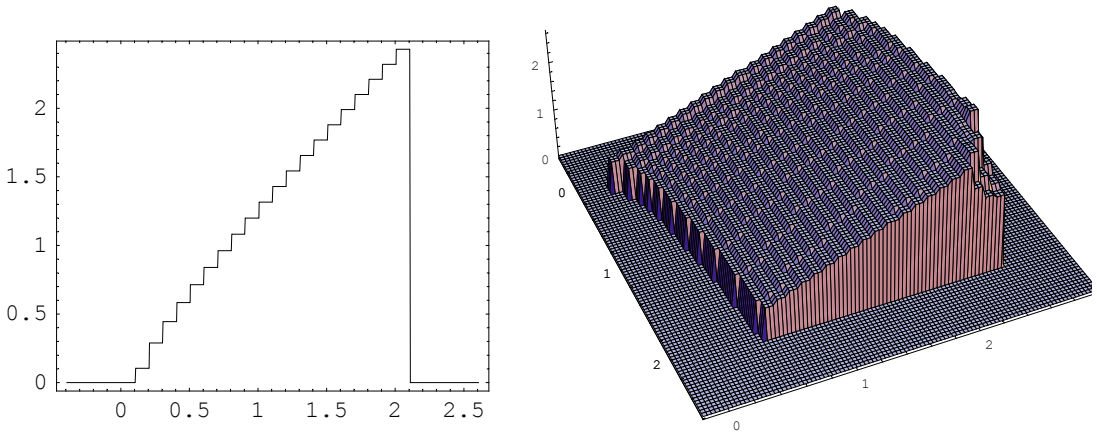

Figure 5: Optimal second stage and third stage release policies for a three-stage example with independent components

7 Conclusion

The numerical results have shown that the presented dynamic model can be efficiently treated for a small number (≤ 3) of stages. Obviously, the elementary discretization grid considered here becomes an obstacle for larger number of stages for complexity reasons. Therefore, a future adaptation of sparse grid ideas (see, e.g., [1]) to our setting shall open

a new perspective to this issue. It is also emphasized that our simple reservoir model is easily extended to more complicated ones, for instance to a network of reservoirs.

References

[1] Bungartz, H.-J. and Griebel, M., Sparse grids, Acta Numerica (2004), 1-123.

[2] Duranyildiz, I., Önöz, B. and Bayazit, M., A chance-constrained LP model for short term reservoir operation optimization, Turkish Journal of Engineering 23 (1999), 181-186.

[3] Edirisinghe, N.C.P., Patterson, E.I. and Saadouli, N., Capacity planning model for a multipurpose water reservoir with target-priority operation, Annals of Operations Research 100 (2000), 273-303.

[4] Gassmann, H.I. and Prékopa, A., On stages and consistency checks in stochastic programming, Operations Research Letters 33 (2005), 171-175.

[5] Henrion, R. and Römisch, W., Hlder and Lipschitz Stability of Solution Sets in Programs with Probabilistic Constraints, Mathematical Programming 100 (2004) 589-611.

[6] Loiaciga, H.A., On the use of chance constraints in reservoir design and operation modeling, Water Resources Research 24 (1988), 1969-1975

[7] Loucks, D.P., Stedinger, J.R. and Haith, D.A., Water Resource Systems Planning and Analysis, Prentice Hall, New Jersey, 1981.

[8] Luedtke, J., Ahmed, S. and Nemhauser, G.L., An integer programming approach for linear programs with probabilistic constraints, to appear in Mathematical Programming (appeared online at: DOI 10.1007/s10107-008-0247-4)

[9] Morgan, D.R., Eheart, J.W. and Valocchi, A.J., Aquifer remediation design under uncertainty using a new chance constraint programming technique, Water Resources Research 29 (1993), 551-561.

[10] Pereira, M.V.F. and Pinto, L.M.V.G., Multi-stage stochastic optimization applied to energy planning, Mathematical Programming 52 (1991), 359-375.

[11] Prékopa, A. and Szàntai, T., Flood control reservoir system design using stochastic programming, Mathematical Programming Study 9 (1978), 138-151.

[12] Prékopa, A. and Szàntai, T., On optimal regulation of a storage level with application to the water level regulation of a lake, European Journal of Operations Research 3 (1979), 175-189.

[13] Prékopa, A., Stochastic Programming, Kluwer, Dordrecht, 1995.

[14] Prékopa, A., Probabilistic Programming, Chapter V in: A. Ruszczynski and A. Shapiro (eds.): Stochastic Programming, Hamdbooks in Operations Research and Management Science, Vol. 10, Elsevier, Amsterdam, 2003.