Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

IZASKUN GARRIDO AND MARC C. STEINBACH

A Multistage Stochastic Programming Approach in Real-Time Process Control

ZIB-Report 01-05 (March 2001)

A Multistage Stochastic Programming Approach in Real-Time Process Control

Izaskun Garrido and Marc C. Steinbach

Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB),
Takustr. 7, 14195 Berlin

Abstract. Standard model predictive control for real-time operation of industrial production processes may be inefficient in the presence of substantial uncertainties. To avoid overly conservative disturbance corrections while ensuring safe operation, random influences should be taken into account explicitly. We propose a multistage stochastic programming approach within the model predictive control framework and apply it to a distillation process with a feed tank buffering external sources. A preliminary comparison to a probabilistic constraints approach is given and first computational results for the distillation process are presented.

1 Introduction

The work reported here is part of a joint research effort aiming at real-time control of chemical processes under uncertainty. Two different stochastic optimization approaches are studied, with the intention to explore and compare their respective general properties and their usefulness in certain practical situations. A specific distillation process serves as a prototypical application example which is investigated under various aspects.

Distillation processes are used to separate liquid or vapor mixtures of several substances into products with different compositions of a desired purity, by the application and removal of heat. Distillation is the most widely used separation process in chemical industry; it consumes large amounts of energy.

The specific process under investigation is the separation of a binary mixture of methanol and water in a continuously running system of two energetically coupled distillation columns. The process is fed from a buffer tank that collects several external sources. In practice, uncertainty occurs when the inflow into the tank may vary at random due to disturbances in the upstream processes. A robust extraction strategy is then required to prevent the tank from running dry or spilling over while keeping the process in favorable operating conditions. Specifically, we assume that the total energy consumption is to be minimized over a given planning horizon; cf. [12, §1].

A pilot system of the process just described is installed at the Institute of Process Dynamics of the Technical University of Berlin. Optimization results have been obtained in the partner project for a simplified one-stage column model called a *flash unit* [14]. In that work, the composition of the inflow mixture and its temperature are assumed to be deterministic whereas the

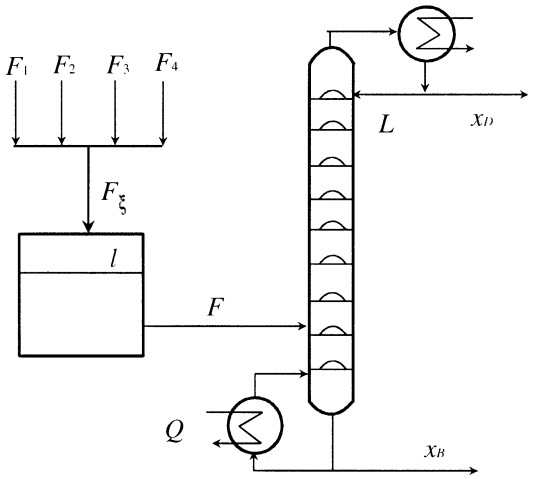

Fig. 1. Distillation column with feed tank. F_ξ: random inflow rate; F: feed rate; L: reflux flow rate; Q: reboiler duty; x_D: distillate; x_B: bottom.

inflow *rate* may vary at random. The rate is modeled as an autocorrelated Gaussian process representing the superposition of many independent inflows, as indicated in Figure 1. A rectangular inflow profile modeling a single event with known rate and duration but random starting time is also considered in [14]. Each of these stochastic models represents disturbances that occur in usual process operation, as opposed to exceptional events like failures.

The stochastic optimization approach presented here is based on a scenario tree model; cf. [12, §§2, 3]. Possible deviations from the expected inflow profile over the entire planning horizon are thus represented as a discrete-time stochastic process with finitely many realizations. An optimal solution in this framework minimizes the mathematical expectation of the cost (energy consumption) over all scenarios. The optimal control strategy is itself a *nonanticipative* stochastic process: it determines a different extraction profile for each scenario, thus specifiying *a priori* how to *react* to future measurements of actual inflows. This approach requires that (within a linearized model) the constraints can be satisfied for any possible sequence of random disturbances (the tank filling level can be kept feasible for any sequence of inflows). More generally, the approach is applicable if *hard* constraints can always be satisfied and costs for potential violations of *soft* constraints can be quantified. In the latter case, soft constraints will be satisfied only if this is possible and economic.

The probabilistic constraints approach pursued in the partner project also models the inflow history as a discrete-time process but allows continuous probability distributions given by a density function; cf. [12, §§2, 4]. The

optimal control strategy in this case is *deterministic*: it does not react to actual disturbances. Instead, it minimizes the cost under the restriction that constraints will be satisfied with high probability (for instance, in at least 95% of all cases). This approach is appropriate if it is not possible to satisfy all constraints with certainty (in certain extreme cases one cannot prevent the tank from running dry or spilling over), or if such events cause substantial costs for which no precise model is available.

Both stochastic optimization approaches are naturally applicable within a moving horizon framework. In the present context, *real-time* process control means response times in the order of 10 to 15 minutes. This is appropriate for the distillation process with a planning horizon of about a day and reoptimizations every 2 or 3 hours.

As indicated in [12, §1], uncertainty may influence the planning and operation of chemical production processes in various ways. For instance, [23,24] study the *design* of chemical plants under uncertainty, with the aim of guaranteeing the existence of feasible control strategies after observing the random event(s). Process *operation* under uncertainty is investigated, e.g., in [16,25] and, for the case of random feed streams, in [7] and in the partner project [13,14]. A stochastic integer programming approach to online scheduling of batch processes is given in [8]. The rigorous treatment of exceptional events by scenario-based DAE models is described in [1]. A coarse classification of relevant types of uncertainty and a general discussion of the topic can be found in [17]. The specific area of distillation processes is particularly well-studied under various aspects. See, e.g., [20] for a recent general survey, [5] for optimal control in the presence of random feed, or [2] for a large-scale industrial application. For the background in stochastic optimization required in this paper we refer to [12, §§2, 3] and standard textbooks [3,15].

The current investigation treats the same general situation as [14], using a tracking approach presented in §2. The discretization of the Gaussian inflow process in a scenario tree framework is described in §3, and a new, straightforward technique of evaluating integrals of the multivariate normal density is proposed in §4. Finally we present first computational results in §5 and give some conclusions in §6.

2 Optimization Model

A schematic diagram of a distillation column with a buffer tank is shown in Figure 1. Uncertainty occurs in the tank inflow F_ξ; control variables are the reboiler duty Q, the feed extraction F directed from the tank to the column, and the reflux flow L. For more detailed descriptions ot the system we refer to [12, §1] and [9,13].

Our first investigations, as reported here, are aimed at answering the following question: given a desired extraction profile $\hat{F}$, under which inflow conditions is it possible to satisfy the level constraints in the tank, and how

difficult is it? Difficulty is measured as the expectation of the accumulated quadratic deviation between actual extraction and target profile. That is, we solve on-line a stochastic *tracking problem* where the target profile typically results from an *off-line* process optimization. Uncertainty is thus effectively decoupled from the process dynamics. Of course, such a simplified approach will be practically useful only in situations where the distillation column is always capable of processing the extracted amount of liquid without violating the purity constraints, and the total energy consumption $\int Q\,dt$ is not too sensitive to the deviations.

2.1 Continuous Time

Since only the basic model structure is of interest here, we formulate a deterministic tracking problem for simplicity (without uncertainty in the inflow rate F_ξ). Given a target feed rate $\hat{F}$ and denoting by v the liquid volume in the tank, the model reads

$$\min_F \int_0^T \frac{1}{2}[F(t) - \hat{F}(t)]^2\,dt \tag{1}$$

$$\text{s.t. } \dot{v}(t) = F_\xi(t) - F(t), \tag{2}$$

$$v(0) = \hat{v}_0, \tag{3}$$

$$v(T) = \hat{v}_T, \tag{4}$$

$$v(t) \in [v^{\min}, v^{\max}], \tag{5}$$

$$F(t) \in [F^{\min}, F^{\max}]. \tag{6}$$

Here $\hat{v}_0$ is the known initial volume, and $\hat{v}_T$ is a prespecified final level.

Note that some terminal condition on the liquid volume is always required in the original problem since minimizing the total energy consumption of the reboiler would otherwise result in processing as little liquid as possible, and hence yield a full tank at the end of the planning horizon. Here we simply fix $v(T)$; the nature of the condition (called a *cycling constraint*) will be discussed in more detail in the following section.

2.2 Discrete Time

Considering T periods (not necessarily of equal physical length) in discrete time $t = 0, 1, \ldots, T$, we denote by v_t the liquid volume in the tank at time t, by f_t the feed volume extracted during $(t, t+1)$, by $\hat{f}_t$ the associated target extraction volume, and by ξ_t the random inflow volume during $(t-1, t)$. In our approach, the latter is assumed to vary only within some known finite interval $[\xi_t^{\min}, \xi_t^{\max}]$. We also assume that the volume f_t is extracted at a constant rate, which is consistent with standard practice in process operation. The tank filling volume now evolves according to

$$v_t = v_{t-1} - f_{t-1} + \xi_t, \qquad t = 1, \ldots, T.$$

Note that the time index on f_t and $\hat{f}_t$ refers to the following period whereas on ξ_t it refers to the previous period. This reflects the fact that decisions on extraction volumes must be made at the beginning of each period whereas inflow volumes are only measured at the end. An obvious consequence is that the control f_t will always lag behind one period in compensating for undesired inflows. It can never steer the state v_t to a precise value but only into some range determined by the inflow variation: $v_t \in (v_{t-1} - f_{t-1}) + [\xi_t^{\min}, \xi_t^{\max}]$.

In particular, the cycling constraint $v_T = \hat{v}_T$ cannot be satisfied with certainty but only in an average sense. It is thus replaced by

$$\mathbf{E}(v_T) = \hat{v}_T \tag{7}$$

in [14]. This is the best one can do in a probabilistic constraints framework, but the condition is actually quite weak in our approach: final values of v_T may vary over the entire feasible range rather than being clustered around $\hat{v}_T$ (as intended). In fact we can do better: it is possible to satisfy a similar condition *independently* for every realization of the preceding state v_{T-1}, which amounts to prescribing the *conditional expectation*

$$\mathbf{E}(v_T|v_{T-1}) = \hat{v}_T. \tag{8}$$

(This condition can also be interpreted as a limiting case of (7) when the feasible range $[v_T^{\min}, v_T^{\max}]$ is continuously reduced in an appropiate manner.) We will present optimization results for both alternatives.

The discrete-time stochastic tracking problem minimizes the *expected* tracking error. In variables $v = (v_0, \ldots, v_T)$ and $f = (f_0, \ldots, f_{T-1})$ it reads

$$\min_{(v,f)} \sum_{t=0}^{T-1} \frac{1}{2}\mathbf{E}[(f_t - \hat{f}_t)^2] \tag{9}$$

$$\text{s.t. } v_t = v_{t-1} - f_{t-1} + \xi_t, \quad t = 0, \ldots, T, \tag{10}$$

$$\mathbf{E}(v_T) = \hat{v}_T \quad \text{or} \quad \mathbf{E}(v_T|v_{T-1}) = \hat{v}_T, \tag{11}$$

$$v_t \in [v^{\min}, v^{\max}], \quad t = 1, \ldots, T, \tag{12}$$

$$f_t \in [f^{\min}, f^{\max}], \quad t = 0, \ldots, T-1. \tag{13}$$

Uncertainty occurs only in the right-hand side of the dynamic equations (10). The special case $t = 0$ is formally handled by setting $v_{-1} := f_{-1} := 0$ and $\xi_0 := \hat{v}_0$, which is equivalent to using the physical quantities v_{-1}, f_{-1}, and ξ_0 from the actual, continously running process.

Given a scenario tree with vertex set V, we denote by $L_t \subseteq V$ the level set of nodes at time t and by $L \equiv L_T$ the set of leaves; further by $0 \in L_0$ the root, by $j \in L_t$ the "current" node, by $i \equiv \pi(j) \in L_{t-1}$ its unique predecessor (if $t > 0$), and by $S(j) \subseteq L_{t+1}$ its set of successors. The node probabilities are $p_j > 0$, $j \in V$. For further details see [12, §2], or [18, §2.4] where alternative scenario models (with explicit nonanticipativity constraints) are discussed.

In the numerical formulation, the state variable is $x_t := v_t$ and the control is defined as the tracking error, $u_t := f_t - \hat{f}_t$ (with limits $u_t^{\min} := f^{\min} - \hat{f}_t$ and $u_t^{\max} := f^{\max} - \hat{f}_t$). Correspondingly, we define $h_t := \xi_t - \hat{u}_{t-1}$. Objective terms then simplify to $\mathbf{E}(u_t^2)$, and the remaining equations and constraints retain their original form with proper variable replacements—except that the control bounds are now time-dependent. On the scenario tree, stochastic quantities x_t, u_t, h_t are represented by their realizations x_j, u_j, h_j, $j \in L_t$. Letting $V^* := V \setminus \{0\}$ and recalling $i \equiv \pi(j)$, the first optimization problem (with cycling constraint $\mathbf{E}(v_T) = \hat{v}_T$) then reads

$$\min_{(x,u)} \sum_{j \in V \setminus L} \frac{1}{2} p_j u_j^2 \tag{14}$$

$$\text{s.t. } x_j = x_i - u_i + h_j \quad \forall j \in V, \tag{15}$$

$$x_j \in [x^{\min}, x^{\max}] \quad \forall j \in V^*, \tag{16}$$

$$u_j \in [u_t^{\min}, u_t^{\max}] \quad \forall j \in V \setminus L, \tag{17}$$

$$\sum_{j \in L} p_j x_j = \hat{x}_T. \tag{18}$$

The second problem can be written in the same form where the cycling constraint $\mathbf{E}(v_T | v_{T-1}) = \hat{v}_T$ replacing (18) translates to

$$\sum_{k \in S(j)} \frac{p_k}{p_j} v_k = \hat{v}_T \quad \forall j \in L_{T-1}.$$

However, this (set of) condition(s) is not explicitly specified in the problem formulation. Instead, we use it to pre-eliminate the final period entirely as follows. Substituting the dynamic equation (10) into (8) yields

$$\hat{v}_T = \mathbf{E}(v_{T-1} - f_{T-1} + \xi_T | v_{T-1}) = v_{T-1} - f_{T-1} + \mathbf{E}(\xi_T | v_{T-1}).$$

Hence, the final-period feed extraction is uniquely determined as

$$f_{T-1} = v_{T-1} - \hat{v}_T + \hat{\xi}_{T-1} \tag{19}$$

where $\hat{\xi}_{T-1} := \mathbf{E}(\xi_T | v_{T-1})$ is the conditional expectation of the final-period inflow. The uniqueness of f_{T-1} shows that (8) is actually the strongest possible cycling constraint in our framework.

It remains to clarify the roles of conditions (12) at $t = T$ and (13) at $t = T - 1$. The bounds on v_T translate to a restriction of the problem *data*,

$$\xi_T - \hat{\xi}_{T-1} \in [v^{\min} - \hat{v}_T, v^{\max} - \hat{v}_T].$$

This yields an *a priori* feasibility check: no feasible solution can exist if the final inflow ξ_T (conditioned on v_{T-1}) varies too much. On the other hand, since

$$\xi_T - \hat{\xi}_{T-1} \in [\xi_T^{\min} - \xi_T^{\max}, \xi_T^{\max} - \xi_T^{\min}],$$

the restriction is always satisfied if the latter range is contained in the former.

The bounds on f_{T-1} simply imply a further restriction of v_{T-1},

$$v_{T-1} \in (\hat{v}_T - \hat{\xi}_{T-1}) + [f^{\min}, f^{\max}].$$

In the numerical formulation, the final state variable is now defined as the final-period tracking error, $x_{T-1} := (v_{T-1} - \hat{v}_T + \hat{\xi}_{T-1}) - \hat{f}_{T-1}$ by (19). Accordingly, we have $h_{t-1} := \xi_{t-1} - \hat{v}_T + \hat{\xi}_{T-1} - \hat{f}_{T-1}$ and limits

$$\begin{aligned} x_{T-1}^{\min} &:= \max(f^{\min}, v^{\min} - \hat{v}_T + \hat{\xi}_{T-1}) - \hat{f}_{T-1}, \\ x_{T-1}^{\max} &:= \min(f^{\max}, v^{\max} - \hat{v}_T + \hat{\xi}_{T-1}) - \hat{f}_{T-1}, \end{aligned}$$

yielding a possibly empty interval (which is also checked a priori). At times $t = 0, \ldots, T-2$ we use the same variables as before. That is, $x_t := v_t$, $u_t := f_t - \hat{f}_t$, (with limits $u_t^{\min}, u_t^{\max}$), and $h_t := \xi_t - \hat{u}_{t-1}$. The state bounds for $0 < t < T-1$ are $x_t^{\min} := v^{\min}$ and $x_t^{\max} := v^{\max}$.

Due to the eliminations, the leaves of the original scenario tree are now obsolete (typically a drastical reduction in size!), and the optimization problem is defined on the subtree with vertex set $V_c := V \setminus L$ and leaf set $L_c := L_{T-1}$. It reads

$$\min_{(x,u)} \sum_{j \in V_c \setminus L_c} \frac{1}{2} p_j u_j^2 + \sum_{j \in L_c} \frac{1}{2} p_j x_j^2 \tag{20}$$

$$\text{s.t. } x_j = x_i - u_i + h_j \quad \forall j \in V_c, \tag{21}$$

$$x_j \in [x_t^{\min}, x_t^{\max}] \quad \forall j \in V_c^*, \tag{22}$$

$$u_j \in [u_t^{\min}, u_t^{\max}] \quad \forall j \in V_c \setminus L_c. \tag{23}$$

This is still very similar to problem (14–18), but instead of a terminal condition we now have objective terms in the final period, and the state bounds in $T-1$ are now defined by the (possibly empty) intersection of two intervals.

3 Discretizing the Gaussian Process

The autocorrelated Gaussian process model for the inflow rate F_ξ leads immediately to an autocorrelated discrete-time Gaussian process of inflow volumes ξ_t, see Figure 2. The latter is given by a general multivariate normal distribution $\mathcal{N}(\bar{\xi}, \Sigma)$ whose dimension is the number of time periods, T. The density function φ of the normal distribution is positive on the entire space, that is, its support is $\mathbb{R}^T$. Thus, although with small probability, it allows arbitrarily large inflows and even negative ones (which are physically impossible unless the tank leaks). On the other hand, a scenario tree representation necessarily corresponds to a *discrete* probability distribution, with *compact* support. Since we intend to compare the two stochastic optimization approaches, this raises several nontrivial questions: How should the support be chosen, how should the scenario tree be constructed, and how should the node

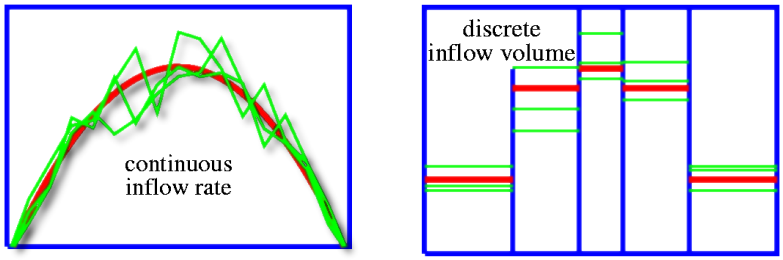

Fig. 2. Expected inflow profiles (bold) and realizations. Left: continuous time. Right: discrete time.

probabilities be assigned to obtain a "good discretization" of the continuous normal distribution? Such questions would be less relevant in practice: scenario trees would be constructed directly from measurements, and a Gaussian process would be seen as just one possible approximation of the real data, with the property of being particularly well tractable in the probabilistic constraints approach.

Compact Support. Due to the absence of real data we make the following assumptions that seem reasonable for a basic investigation: the actual inflow rate may only vary within a (possibly variable) symmetric bandwith around the expected rate, and "negative inflows" are impossible,

$$F_\xi(t) \in [F_\xi^{\min}(t), F_\xi^{\max}(t)] \equiv [\bar{F}_\xi(t) - \Delta F_\xi(t), \bar{F}_\xi(t) + \Delta F_\xi(t)] \subset \mathbb{R}_+. \quad (24)$$

Integration over the subintervals then yields a similar relation for the inflow volumes,

$$\xi_t \in [\xi_t^{\min}, \xi_t^{\max}] \equiv [\bar{\xi}_t - \Delta\xi_t, \bar{\xi}_t + \Delta\xi_t] \subset \mathbb{R}_+. \quad (25)$$

This means that the discrete distribution will be supported by a T-dimensional compact box centered at the mean and lying entirely in the positive orthant,

$$\Xi := [\bar{\xi} - \Delta\xi, \bar{\xi} + \Delta\xi] \equiv \prod_{t=1}^{T} [\bar{\xi}_t - \Delta\xi_t, \bar{\xi}_t + \Delta\xi_t] \subset \mathbb{R}_+^T. \quad (26)$$

We now define a new density function (whose support is precisely this box) by restricting the given normal density to Ξ and renormalizing the weight,

$$\varphi_\Xi(\xi) := \frac{1}{P(\Xi)} \chi_\Xi(\xi)\varphi(\xi), \qquad P(\Xi) = \int_\Xi \varphi(\xi)\, d\xi. \quad (27)$$

Probabilities are thus replaced by *conditional* probabilities with respect to Ξ. Obviously the construction leaves expected inflows invariant by symmetry. The correlations of inflows, however, will deviate from their original values,

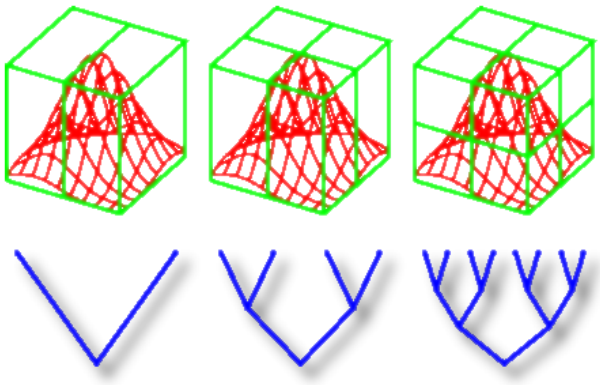

Fig. 3. Construction of a scenario tree by recursive partitioning of the compact support Ξ. Node probabilities correspond to the weights of (unions of) sub-boxes.

giving increasingly inaccurate approximations with decreasing weight $P(\Xi)$. Moreover, assumption (24) on the inflow rate guarantees that all constraints are satisfied in continuous time if this is true in the discrete-time model.

Scenario Tree. Currently we construct the scenario tree from a uniform recursive partitioning of Ξ. Each scenario corresponds to an elementary box of full dimension. The stage-t scenarios correspond to unions of elementary boxes having the same geometry in the first t dimensions (i.e., identical projections into the associated subspace). This means that the same nodes are traversed up to level t in the scenario tree or, equivalently, that inflow volumes are identical during the first t periods, see Figure 3. The number k_t of partitions in dimension t is the number $|S(j)|$ of successors of each node j on level $t-1$. The resulting total number of boxes (scenarios) is $N = \prod_{t=1}^{T} k_t$.

Scenario Probabilites. As scenario probability we define the weight of the associated elementary box (with respect to the renormalized density φ_Ξ). This weight is assumed to be concentrated in the geometric center of the box, whose coordinates represent the sequence of associated inflow volumes. Although the center of gravity would yield a more accurate approximation, we prefer the geometric center since this allows to choose an exact range of inflow variations *a priori*: realizations of the discrete distribution will be evenly spaced at a distance of $2\Delta\xi_t/k_t$ between limits

$$\xi^{\min} + \frac{\Delta\xi_t}{k_t}, \quad \xi^{\max} - \frac{\Delta\xi_t}{k_t}. \tag{28}$$

(With the center of gravity, minimal and maximal realizations would depend on the density.)

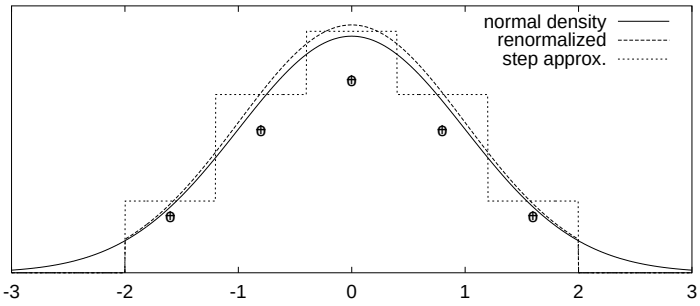

Fig. 4. Discretization of a univariate normal distribution on $[-2, 2]$.

For a standard univariate normal distribution, Figure 4 shows the probability density function, the renormalized density for the interval $\Xi = [-2, 2]$ (having weight 0.9545), an approximation by a piecewise constant density for $k_1 = 5$ scenarios, and the weights of the five subintervals concentrated in their respective midpoints.

4 Calculating Scenario Probabilities

Consider without loss of generality a centralized normal distribution $\mathcal{N}(0, \Sigma)$ (with mean $\bar{\xi} = 0$) in $\mathbb{R}^T$. The density function reads

$$\varphi(\xi) = \frac{1}{\sqrt{(2\pi)^T \det(\Sigma)}} \exp\left(-\frac{1}{2}\xi^* \Sigma^{-1} \xi\right). \tag{29}$$

In our optimization model with equidistant time discretization we make the same assumptions as our colleagues [13,14]: random inflow volumes ξ_t have the same variance σ^2 in all periods, and their correlations c_{st} decrease linearly with the distance $|s - t|$ such that the elements of the covariance matrix are

$$\Sigma_{st} = \sigma^2 c_{st}, \quad c_{st} = 1 - \frac{1}{T}|s - t|, \qquad s, t = 1, \ldots, T. \tag{30}$$

We have to calculate the scenario probabilities which are defined as multivariate integrals of the density function over rectangular domains. Multidimensional numerical integration is generally hard since the required effort in direct generalizations of univariate integration methods grows exponentially with the dimension. Thus, Monte Carlo techniques are often applied. In the special case of normal distributions, some alternative approaches are reported in the literature. Schervish [19] employs an adaptive quadrature routine using an error estimate based on the Newton-Cotes approximation with non-local modifications. Deák [4] combines a transformation to spherical coordinates with Monte-Carlo techniques. The method of Genz [10] transforms the integration domain to the unit cube and applies either Monte-Carlo, adaptive subregions, or lattice rules to the transformed integral.

Here we propose a straightforward, easily implementable approach based on direct integration of the second order Taylor series expansion in combination with adaptive refinement. This appears to be a new method; it seems appropriate in our situation since we have to evaluate integrals over comparatively small domains: the elementary boxes representing scenarios.

Taylor Approximation. In scalar product notation $\langle \xi, \eta \rangle := \xi^* \Sigma^{-1} \eta$, the first three derivatives of the density function (29) are

$$\begin{aligned} D\varphi(\xi)[h] &= -\varphi(\xi)\langle \xi, h\rangle, \\ D^2\varphi(\xi)[h,h] &= +\varphi(\xi)\left[\langle \xi, h\rangle^2 - \langle h, h\rangle\right], \\ D^3\varphi(\xi)[h,h,h] &= -\varphi(\xi)\left[\langle \xi, h\rangle^3 - 3\langle \xi, h\rangle\langle h, h\rangle\right]. \end{aligned}$$

Hence, with some $\theta \in [0,1]$, the third-order Taylor series expansion reads

$$\begin{aligned} \varphi(\xi + h) = \varphi(\xi) - \varphi(\xi)\langle \xi, h\rangle + \frac{1}{2}\varphi(\xi)\left[\langle \xi, h\rangle^2 - \langle h, h\rangle\right] \\ - \frac{1}{6}\varphi(\xi)\left[\langle \xi, h\rangle^3 - 3\langle \xi, h\rangle\langle h, h\rangle\right] + \frac{1}{24} D^4\varphi(\xi + \theta h)[h,h,h,h]. \end{aligned}$$

We need to integrate φ over rectangular boxes $\xi + A = \{\xi + h : h \in A\}$ where

$$A = [-a, a] := \prod_{t=1}^{T} [-a_t, a_t], \qquad \mathrm{Vol}(A) = \prod_{t=1}^{T} 2a_t. \tag{31}$$

For the expansion above, this integration is easily evaluated in closed form. In terms of the Hessian

$$H(\xi) := \varphi(\xi)(\Sigma^{-1}\xi\xi^*\Sigma^{-1} - \Sigma^{-1})$$

one obtains

$$\int_{\xi + A} \varphi(h)\, dh = \varphi(\xi)\,\mathrm{Vol}(A)\left[1 + \frac{1}{24}\sum_{t=1}^{T} H_{tt}(\xi) a_t^2 + O(\|a\|^4)\right]. \tag{32}$$

Observe that, due to symmetry, all odd-order terms vanish in the integration. Thus we get an asymptotic error of order four by adding just one correction (of second order) to the trivial approximation $P(\xi + A) \approx \varphi(\xi)\,\mathrm{Vol}(A)$.

Asymptotic Error Control. To ensure sufficient accuracy, a simple adaptive strategy is employed. After evaluating (32), A is partitioned into a "left" half A_l and a "right" half A_r, and the same weight approximation is applied to A_l and A_r, yielding values w, w_l, w_r. The bisection procedure is recursively repeated with each box until the relative difference falls below a given tolerance,

$$\frac{|w_l + w_r - w|}{w_l + w_r} < \epsilon.$$

Two heuristic strategies have been tested for determining the side s in which to cut the box:

- the largest side, $a_s = \max_t a_t$;
- the side giving the largest second order term, $H_{ss}(\xi)a_s^2 = \max_t H_{tt}(\xi)a_t^2$.

The second strategy was found to perform better and was used in numerical calculations, with the tolerance set to $\epsilon = 10^{-3}$.

Matrix Determinant and Inverse. It turns out that both the determinant and the inverse of the specific correlation matrix (30) have closed-form representations,

$$\det(\Sigma) = \frac{T+1}{4}\left(\frac{2\sigma^2}{T}\right)^T$$

and (for $T \geq 3$)

$$\Sigma^{-1} = \frac{T}{2\sigma^2}\begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix} + \frac{T}{2\sigma^2(T+1)}\begin{bmatrix} 1 & 0 & \dots & 0 & 1 \\ 0 & 0 & & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & 0 & 0 \\ 1 & 0 & \dots & 0 & 1 \end{bmatrix}.$$

Using these formulae in (32) yields significant savings in the numerical computation, especially for higher dimensions. The computational effort is still immense (several hours for a complete discretization), but it should be kept in mind that the weight calculation is an *off-line* task which is required just once for a given inflow distribution.

5 Computational Results

In this section we report on computational experiments with the stochastic tracking problems stated in §2.2, where either the expectation or the conditional expectation of the final liquid volume are prescribed, $\mathbf{E}(v_T) = \hat{v}_T$ or $\mathbf{E}(v_T|v_{T-1}) = \hat{v}_T$. All problems are solved by a primal-dual interior point method combined with a tree-sparse KKT solver [21,22].

5.1 Prescribed Expectation of v_T

Problem data. To allow a comparison of optimization approaches, our computations are based on the following problem data for which optimization runs have also been performed in the partner project. They are slight modifications of the computations described in [14].

The planning horizon has a length of 16 hours and is equidistantly partitioned into eight discretization intervals of two hours each, $\Delta t = 2$ h. Only

Table 1. Target extraction volumes for each two-hour period.

Target type	1	2	3	4	5	6	7	8
Expected inflow	75.44	162.9	221.2	250.3	250.3	221.2	162.9	75.44
Deterministic	139.8	186.3	186.3	186.3	186.3	186.3	186.3	162.0
Probabilistic	126.2	193.0	196.9	198.0	198.0	198.0	198.0	111.7

the flow rate into the buffer tank is assumed to vary at random, whereas its temperature and the respective concentrations of methanol and water remain constant. A parabolic profile of the expected inflow rate is assumed, starting and ending with 11.0 ml/h and reaching a maximum of 127.6 ml/h after eight hours, at $t = 4$. The associated variance of the inflow volume in each two-hour period is $\sigma^2 = 20.0\,\mathrm{ml}^2$, yielding by (30) the covariance matrix

$$\Sigma_{st} = 20c_{st}, \quad c_{st} = 1 - \frac{1}{8}|s-t|, \qquad s,t = 1,\ldots,8.$$

The liquid volume in the tank is to be kept between 440 ml and 1320 ml, with an initial filling level of 1210 ml. This value is also specified as the final level so that the distillation process can be repeated periodically if there are no disturbances. In the presence of disturbances, preventing violations of the upper limit $v^{\max}$ is a major concern since the initial level and desired final level are quite close to that limit.

Problem types. We consider three problem types corresponding to the following target extraction profiles:

1. the expected inflow;
2. the optimal extraction strategy of a deterministic optimization based on the expected inflow ("deterministic solution");
3. the optimal extraction strategy under probabilistic constraints as obtained in the partner project ("probabilistic solution").

Tracking the expected inflow is a very simplistic approach; we include this case only for comparison purposes. The deterministic and probabilistic cases are discussed in [14], where both are solved for a DAE model of a flash unit, using a finer discretization of 32 time periods. The extraction limit is $f^{\max} = 186.34$ (average inflow plus 5%) in the deterministic case, and $f^{\max} = 198$ in the probabilistic case. For the problem data given above (with 8 periods), the expected inflow and optimal profiles are displayed in Table 1 and Figure 5. Here the probabilistic solution satisfies the lower (less critical) level constraint with certainty, and the upper level constraint with a probability of 0.95.

Inflow and extraction bounds. For a given inflow limit $\xi^{\max}$, we choose as support Ξ (on which the normal distribution is discretized) the largest

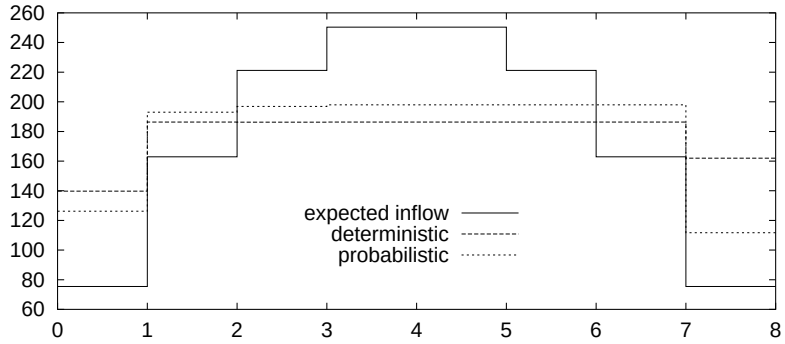

Fig. 5. Target extraction profiles over eight two-hour periods.

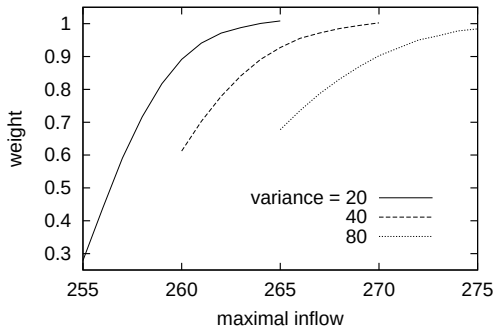

Fig. 6. Weight of support versus inflow limit.

cube centered at the mean $\bar{\xi}$ and lying entirely in $[0, \xi^{\max}]^T$,

$$\Xi := \bar{\xi} + [-\Delta, \Delta]^T, \qquad \Delta := \min_t \min(\bar{\xi}_t, \xi^{\max} - \bar{\xi}_t). \tag{33}$$

Recalling that $\xi^{\max} = F_\xi^{\max} \Delta t$, this obviously models a fixed bandwidth of the inflow rate F_ξ. To study the feasibility question stated in §2, we vary the inflow limit $\xi^{\max}$ in a suitable range of values slightly larger than the largest expected inflow $\max_t \bar{\xi}_t = 250.3\,\mathrm{ml}$, namely $\xi^{\max} \in \{255, 256, \ldots, 265\}$. The discrete bandwidth is then determined by the largest expected inflow as $\Delta = \xi^{\max} - \max_t \bar{\xi}_t$, covering the range $[4.7, 14.7]$ and yielding cubes Ξ of different sizes with weights roughly between 0.3 and 1; see the solid line in Figure 6. (For our standard deviation $\sigma \approx 4.5$, this allows to compare "good" and "poor" discretizations of the normal distribution.)

Obviously, feasibility is harder to achieve when $\xi^{\max}$ is increased (giving larger inflow variations) or when $f^{\max}$ is decreased (giving a smaller control range). For each $\xi^{\max}$ value we therefore set $f^{\max} = \xi^{\max}$ first and then decrease $f^{\max}$ until the problem becomes infeasible. For all combinations of $\xi^{\max}$ and $f^{\max}$ we solve problem (14–18). Data for variances $\sigma^2 = 40$ and

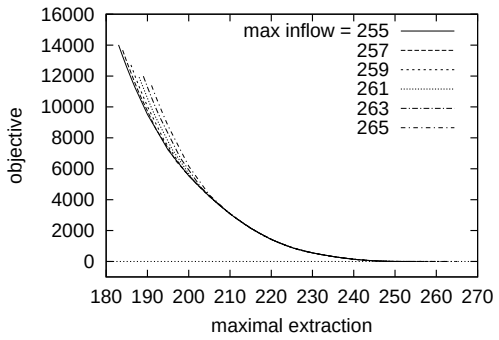

Fig. 7. Tracking error versus extraction limit for expected inflow.

$\sigma^2 = 80$ appearing in the plots are associated with problems in § 5.2 and will be discussed below.

Expected Inflow. First we consider optimal solutions for the expected inflow as target profile. Figure 7 plots the optimal objective value, i.e., the (expected) "tracking error", versus extraction limit $f^{\max}$ for selected inflow limits $\xi^{\max}$ covering the entire range $[255, 265]$. One observes that the tracking error increases with decreasing extraction limit; closer inspection of the data reveals that its value is actually zero in all problems with $f^{\max} \geq 251$. The first observation confirms precisely the expected behavior: since the target profile has a peak inflow of 250.3 in the middle, large extractions are required in earlier and later periods when $f^{\max}$ is reduced below that peak value. The second observation says that (accumulated) inflow deviations can never violate a limit if precisely the expected inflow is extracted. This is easily verified; it indicates that inflow variations are moderate even in the case $\xi^{\max} = 265$ which covers about 99.9% of the distribution.

It is also observed that the tracking error is almost independent of the inflow limit $\xi^{\max}$ for large values of $f^{\max}$ whereas there are significant differences for small values. This is also easily explained: A large extraction limit is not a severe restriction in any case, whereas a small limit will be binding in most periods until the problem becomes eventually infeasible.

Deterministic Solution. Next we consider the deterministic solution as target profile. Figure 8 plots the tracking error versus extraction limit $f^{\max}$ for the same $\xi^{\max}$ values as before. Here it is observed that in all cases the tracking error remains constant (on a low level) over a wide range of $f^{\max}$ values but increases rapidly when infeasibility is approached. Smaller inflow limits yield smaller tracking errors for all extraction limits. This might come as a surprise but can be explained when the target profile is inspected: the deterministically optimal extraction is constantly at its upper limit 186.34,

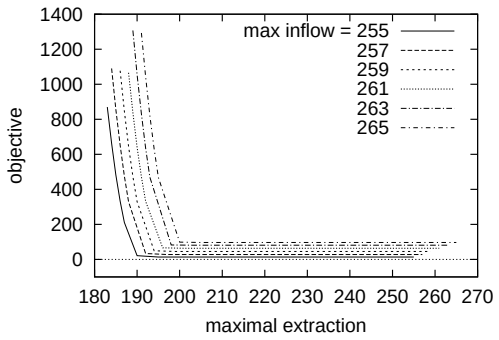

Fig. 8. Tracking error versus extraction limit for deterministic solution.

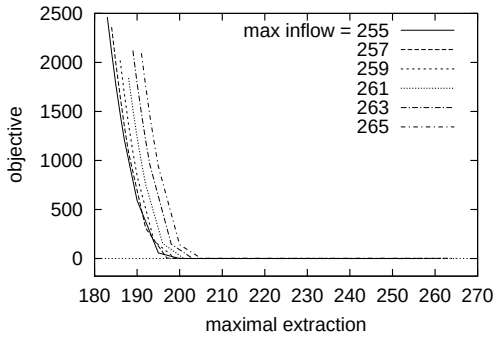

Fig. 9. Tracking error versus extraction limit for probabilistic solution.

except for the first and last interval. Again, only small $f^{\max}$ values are a severe restriction, and a violation of the upper level constraint by accumulated inflow deviations can be avoided with comparatively small corrections. This is consistent with the results in [14]: although many trajectories violate the upper level constraint for the given profile, the limit is only slightly exceeded.

Probabilistic Solution. The probabilistic solution as target profile yields almost identical results, as shown in Figure 9. However, the constant level for large $f^{\max}$ values is consistently smaller than in the deterministic case: at most 1.86, but often exactly zero. This confirms the robustness of the probabilistic solution (certain feasibility is achieved with negligible extra effort) and is again a consequence of the precise shape of the target profile. On the other hand, for small $f^{\max}$ values the tracking error is *larger* than in the deterministic case, which can be seen in the direct comparison of all three target profiles displayed in Figure 10 (where $\xi^{\max} = 265$). The latter fact

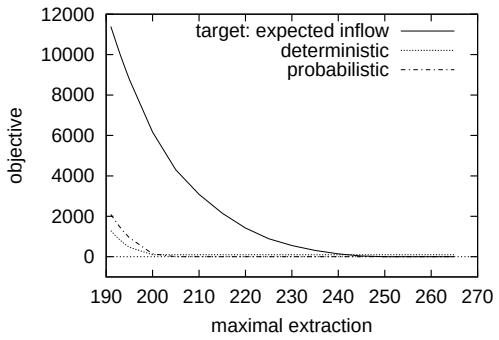

Fig. 10. Tracking error versus extraction limit for all target profiles, with inflow limit $\xi^{\max} = 265$.

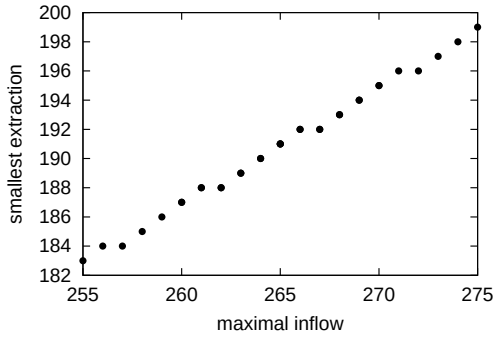

Fig. 11. Smallest feasible extraction limit versus inflow limit.

indicates that reducing the *probability* of constraint violations may go along with an increase of their *size*.

Feasibility. Figure 11 shows how the smallest feasible extraction limit $f^{\max}$ depends on the inflow limit $\xi^{\max}$. (There is obviously no dependence on the target profile.) At $\xi^{\max} = 265$, a feasible solution can still be obtained with $f^{\max} = 191$, and at $\xi^{\max} = 274$ with $f^{\max} = 198$. This demonstrates the flexibility of the stochastic programming approach: with reasonable extraction limits, constraint violations can be avoided with certainty by suitable predetermined reactions to inflow measurements.

5.2 Prescribed Conditional Expectation of v_T

Feasibility was comparatively easy to achieve in the previous problem due to the moderate inflow variance and weak cycling constraint. Therefore we also

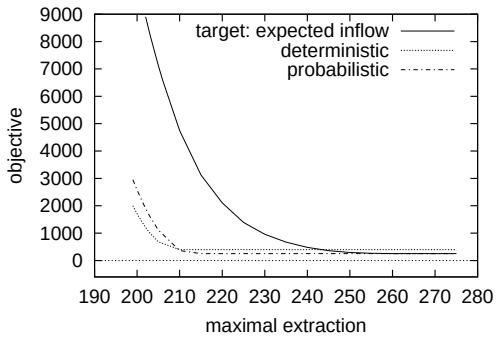

Fig. 12. Tracking error versus extraction limit for all target profiles, with inflow limit $\xi^{\max} = 275$ and conditional expectation in cycling constraint.

investigate problem (20–23), using the variance values $\sigma^2 \in \{20, 40, 80\}$ and respective $\xi^{\max}$ ranges 255–265, 260–270, and 265–275. Associated weights of the support Ξ and smallest feasible extraction limits are included in Figures 6 and 11. A comparison of all three target profiles for $\sigma^2 = 80$ and $\xi^{\max} = 275$ is given in Figure 10. As expected, increasing the variance (and accordingly the inflow limit, to cover sufficient weight of the distribution) requires larger extraction limits to achieve feasibility. Moreover, the optimal tracking error is never zero in all these cases. This shows that, even if no level constraints are violated by the nominal extraction, corrective action is required to meet the stronger cycling constraint. In some of the higher variance problems, part of the tracking error is also due to corrections preventing violation of the upper level constraint. However, the stronger cycling constraint has no significant impact on feasibility. This is because the prescribed final level is close to the critical upper level constraint: if no violations occur as long as inflows are large, the control will be able to extract a sufficient amount of liquid during the last period so that the final level is met on average. (If accumulated inflow deviations are negative, the control just extracts less liquid than expected.) In contrast, test runs show that certain feasibility *is* harder to achieve if the final level is close to the middle of the feasible range. To sum up, results for the stronger cycling constraint are similar to the previous case, and they scale in some sense for the larger variance values. Differences for the three target profiles are slightly accentuated; cf. Figures 10 and 12.

6 Conclusions

Multistage stochastic programming has been proposed as a new approach in real-time control of chemical processes. This can be seen as a generalization of standard model predictive control in the sense that reactions to measured disturbances are combined with the prevention of unfavorable future events

by means of a stochastic model. The basic concept of the approach has been demonstrated for the problem of controlling the buffer tank of a distillation column with random inflows, and a preliminary comparison with a probabilistic constraints approach has been given. Future research should extend this work in different directions. For instance, a more realistic, integrated treatment of stochasticity and process dynamics in the application example is intended. Further, scenario reduction techniques as in [6] appear promising in obtaining better discrete distributions and at the same time allowing finer time discretizations. Finally, warm start techniques and other algorithmic improvements may increase the efficiency of the approach so that faster processes can be controlled in real time.

Acknowledgments

We would like to express our gratitude to G. Wozny, P. Li, and M. Wendt (Technische Universität Berlin) and to R. Henrion and A. Möller (Weierstraß-Institut Berlin) for numerous vital discussions and for providing the process model and optimization data used in the comparison. We also wish to thank L. T. Biegler and G. Rodriguez (Carnegie Mellon University) who suggested this research and introduced the second author to basic robust control issues during two enjoyable research visits. Finally, we gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft.

References

1. O. Abel and W. Marquardt. Scenario-integrated optimization of dynamic systems. *AIChE J.*, 46(4):803–823, 2000.
2. O. E. Agamennoni, J. L. Figueroa, G. W. Barton, and J. A. Romagnoli. Advanced controller design for a distillation column. *Int. J. Control*, 59:817–839, 1994.
3. J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming.* Springer-Verlag, New York, 1997.
4. I. Deák. Three digit accurate multiple normal probabilities. *Numer. Math.*, 35:369–380, 1980.
5. U. M. Diwekar and J. R. Kalagnanam. Efficient sampling technique for optimization under uncertainty. *AIChE J.*, 43:440–447, 1997.
6. J. Dupačová, N. Gröwe-Kuska, and W. Römisch. Scenario reduction in stochastic programming: An approach using probability metrics. Preprint 00-09, Institut für Mathematik, Humboldt-Universität Berlin, 2000.
7. A. M. Eliceche, M. Sanchez, and L. Fernandez. Feasible operating region of natural gas plants under feed perturbations. *Comput. Chem. Eng.*, 22:S879–S882, 1998.
8. S. Engell, A. Märkert, G. Sand, R. Schultz, and C. Schulz. Online scheduling of multiproduct batch plants under uncertainty. Preprint SM-DU-494, Universität Duisburg, 2001. Submitted to [11].

9. H. A. Garcia, R. Henrion, P. Li, A. Möller, W. Römisch, M. Wendt, and G. Wozny. A model for the online optimization of integrated distillation columns under stochastic constraints. Preprint 98-32, DFG Research Center "Echtzeit-Optimierung großer Systeme", Nov. 1998.
10. A. Genz. Numerical computation of the multivariate normal probabilities. *J. Comput. Graph. Statist.*, 1:141–150, 1992.
11. M. Grötschel, S. O. Krumke, and J. Rambau, editors. *Online Optimization of Large Systems.* Lecture Notes in Computational Science and Engineering. Springer-Verlag, 2001. In preparation.
12. R. Henrion, P. Li, A. Möller, M. C. Steinbach, M. Wendt, and G. Wozny. Stochastic optimization for operating chemical processes under uncertainty. Technical Report ZR-01-04, ZIB, 2001. Submitted to [11].
13. R. Henrion, P. Li, A. Möller, M. Wendt, and G. Wozny. Optimization of a continuous distillation process under probabilistic constraints. Submitted to [11].
14. R. Henrion and A. Möller. Optimization of a continuous distillation process under random inflow rate. Preprint 00-4, DFG Research Center "Echtzeit-Optimierung großer Systeme", Mar. 2000. Submitted to Comput. Math. Appl.
15. P. Kall and S. W. Wallace. *Stochastic Programming.* Wiley, New York, 1994.
16. S. Orçun, I. K. Altinel, and Ö. Hortaçsu. Scheduling of batch processes with operational uncertainty. *Comput. Chem. Eng.*, 20:S1191–S1196, 1996.
17. E. N. Pistikopoulos. Uncertainty in process design and operation. *Comput. Chem. Eng.*, 19:S553–S563, 1995.
18. W. Römisch and R. Schultz. Multistage stochastic integer programs: An introduction. Preprint SM-DU-496, Universität Duisburg, 2001. Submitted to [11].
19. M. Schervish. Multivariate normal probabilities with error bound. *J. Appl. Statist.*, 33:81–87, 1984.
20. S. Skogestad. Dynamics and control of distillation columns—a critical survey. *Model. Identif. Control*, 18:177–217, 1997.
21. M. C. Steinbach. Recursive direct algorithms for multistage stochastic programs in financial engineering. In P. Kall and H.-J. Lüthi, editors, *Operations Research Proceedings 1998*, pages 241–250, New York, 1999. Springer-Verlag.
22. M. C. Steinbach. Hierarchical sparsity in multistage convex stochastic programs. In S. Uryasev and P. M. Pardalos, editors, *Stochastic Optimization: Algorithms and Applications*, Kluwer Academic Publishers, 2001. Dordrecht, The Netherlands.
23. D. A. Straub and I. E. Grossmann. Design optimization of stochastic flexibility. *Comput. Chem. Eng.*, 17:S339–S354, 1993.
24. R. Swaney and I. E. Grossmann. An index for operational flexibility in chemical process design. *AIChE J.*, 31:621–630, 1985.
25. P. Terwiesch, D. Ravemark, B. Schenker, and D. W. T. Rippin. Semi-batch process optimization under uncertainty: Theory and experiments. *Comput. Chem. Eng.*, 22:201–213, 1998.