Scenario Generation in Stochastic Programming with Application to Optimizing Electricity Portfolios under Uncertainty

W. Römisch

Humboldt-University Berlin
Department of Mathematics

www.math.hu-berlin.de/~romisch

IMA Hot Topics Workshop *Uncertainty Quantification in Industrial and Energy Applications: Experiences and Challenges*, Minneapolis, June 2-4, 2011

DFG Research Center MATHEON
Mathematics for key technologies

Contents

Stochastic programming and approximation issues

We consider a stochastic program of the form

$$\min\left\{\int_{\Xi}\Phi(x,\xi)P(d\xi) : x\in X\right\},$$

where $X\subseteq\mathbb{R}^m$ is a constraint set, P a probability distribution on $\Xi\subseteq\mathbb{R}^d$, and $f:=\Phi(x,\cdot)$ is a decision-dependent integrand.

Any approach to solving such models computationally requires to replace the integral by a quadrature rule

$$Q_{n,d}(f)=\sum_{i=1}^{n}w_i f(\xi^i),$$

with weights $w_i\in\mathbb{R}$ and scenarios $\xi^i\in\Xi$, $i=1,\dots,n$.

If the natural condition $w_i\geq 0$ and $\sum_{i=1}^{n}w_i=1$ is satisfied, $Q_{n,d}(f)$ allows the interpretation as integral with respect to the discrete probability measure Q_n having scenarios ξ^i with probabilities w_i, $i=1,\dots,n$.

Example 1: Linear two-stage stochastic programs

We consider two-stage linear stochastic programs:

$$\min \left\{ \langle c, x \rangle + \int_{\Xi} \varphi(q(\xi), h(\xi) - Tx) P(d\xi) : x \in X \right\}$$

where $c \in \mathbb{R}^m$, X is a convex polyhedral subset of $\mathbb{R}^m$, Ξ a closed subset of $\mathbb{R}^d$, T a (r, m)-matrix, $h(\cdot)$ and $q(\cdot)$ are affine mappings on $\mathbb{R}^d$, P a Borel probability measure on Ξ and

$$\begin{aligned} \varphi(q, t) &= \inf\{\langle q, y \rangle : Wy = t, y \geq 0\} \\ &= \sup\{\langle t, z \rangle : W^\top z \leq q\} \end{aligned}$$

where $q \in \mathbb{R}^{\bar{m}}$, W a $(r, \bar{m})$-matrix (having rank r) and t varies in the polyhedral cone $W(\mathbb{R}^{\bar{m}}_+)$. There exist matrices C_j and polyhedral cones $\mathcal{K}_j$, $j = 1, \ldots, \ell$, decomposing $\operatorname{dom} \varphi$ such that $\varphi(q, t) = \langle C_j q, t \rangle$, $\forall (q, t) \in \mathcal{K}_j$. Hence, the integrand is

$$\Phi(x, \xi) = \langle c, x \rangle + \max_{j=1,\ldots,\ell} \langle C_j q(\xi), h(\xi) - Tx \rangle$$

Example 2: Linear multi-stage stochastic programs

Let $\{\xi_t\}_{t=1}^T$ be a discrete-time stochastic data process defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and with ξ_1 deterministic. The stochastic decision x_t at period t is assumed to be measurable with respect to $\mathcal{F}_t := \sigma(\xi_1, \ldots, \xi_t)$ (nonanticipativity).

$$\min \left\{ \mathbb{E}\left(\sum_{t=1}^{T} \langle b_t(\xi_t), x_t \rangle \right) \middle| \begin{array}{l} x_t \in X_t,\ x_t \in L_p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^{m_t}) \\ \sum_{\tau=0}^{t-1} A_{t,\tau} x_{t-\tau} = h_t(\xi_t) \\ (t = 1, ..., T) \end{array} \right\}$$

where the sets X_t are convex polyhedral in $\mathbb{R}^{m_t}$, $A_{t,\tau}$, $\tau = 0, \ldots, t-1$, are matrices and the vectors $b_t(\cdot)$ and $h_t(\cdot)$ depend affine linearly on ξ_t, $t = 1, \ldots, T$.

The integrand $\Phi = \Phi_1$ is given by dynamic programming

$$\Phi_{t-1}(x^{t-1}, \xi^t) = \inf_{x_t \in X_t} \left\{ \langle b_t(\xi_t), x_t \rangle + \mathbb{E}(\Phi_t(x^t, \xi^{t+1}) | \mathcal{F}_t) \middle| \sum_{\tau=0}^{t-1} A_{t,\tau} x_{t-\tau} = h_t(\xi_t) \right\},$$

where $t = 2, \ldots, T$, $\Phi_T \equiv 0$, $x^t = (x_1, \ldots, x_t)$, $\xi^t = (\xi_t, \ldots, \xi_T)$.

Assumption: P has a density ρ w.r.t. λ^d.

Now, we set $\mathcal{F} = \{\Phi(\cdot, x)\rho(\cdot) : x \in X\}$ and assume that the set $\mathcal{F}$ is a bounded subset of some linear normed space F_d with norm $\|\cdot\|_d$ and unit ball $\mathbb{B}_d = \{f \in F_d : \|f\|_d \leq 1\}$.

The absolute error of the quadrature rule $Q_{n,d}$ is

$$e(Q_{n,d}) = \sup_{f \in \mathbb{B}_d} \left| \int_\Xi f(\xi) d\xi - \sum_{i=1}^{n} w_i f(\xi^i) \right|$$

and the approximation criterion is based on the relative error and a given tolerance $\varepsilon > 0$, namely, it consists in finding the smallest number $n_{\min}(\varepsilon, Q_{n,d}) \in \mathbb{N}$ such that

$$e(Q_{n,d}) \leq \varepsilon e(Q_{0,d}),$$

holds, where $Q_{0,d}(f) = 0$ and, hence, $e(Q_{0,d}) = \|I_d\|$ with

$$I_d(f) = \int_\Xi f(\xi) d\xi.$$

The behavior of both quantities depends heavily on the normed space F_d and the set $\mathcal{F}$, respectively.
It is desirable that an estimate of the form

$$n_{\min}(\varepsilon, Q_{n,d}) \leq C\, d^q \varepsilon^{-p} \qquad (\text{'tractability'})$$

is valid for some constants $q \geq 0$, $C, p > 0$ and for every $\varepsilon \in (0,1)$.
Of course, $q = 0$ is highly desirable for high-dimensional problems.

Proposition: (Stability)
Let the set X be compact. Then there exists $L > 0$ such that

$$\left| \inf_{x \in X} \int_\Xi \Phi(\xi, x)\rho(\xi)d\xi - \inf_{x \in X} \sum_{i=1}^n w_i \Phi(\xi^i, x)\rho(\xi^i) \right| \leq L\, e(Q_{n,d}).$$

The solution set mapping is outer semicontinuous at P.

Alternatively, we look for a suitable set $\mathcal{F}$ of functions such that $\{C\Phi(\cdot, x) : x \in X\} \subseteq \mathcal{F}$ for some constant $C > 0$ and, hence,

$$e(Q_{n,d}) \leq \frac{1}{C} \sup_{f \in \mathcal{F}} \left| \int_\Xi f(\xi) P(d\xi) - \int_\Xi f(\xi) Q_n(d\xi) \right| = D(P, Q_n),$$

and that D is a metric distance between probability distributions.

Example: L_p-minimal metric ℓ_p (or Wasserstein metric) of order $p \geq 1$

$$\ell_p(P, Q) := \left(\inf \left\{ \int_{\Xi \times \Xi} \|\xi - \tilde{\xi}\|^p \eta(d\xi, d\tilde{\xi}) \middle| \pi_1 \eta = P, \pi_2 \eta = Q \right\} \right)^{\frac{1}{p}}$$

It holds

$$\ell_p(P, Q) = \inf\{\|\xi - \tilde{\xi}\|_p \,|\, \mathcal{L}(\xi) = P, \mathcal{L}(\tilde{\xi}) = Q\}$$

$$\ell_1(P, Q) = \sup \left\{ \left| \int_\Xi f(\xi)(P - Q)(d\xi) \right| : |f(\xi) - f(\tilde{\xi})| \leq \|\xi - \tilde{\xi}\| \right\}$$

by definition and duality, respectively.

Examples of normed spaces F_d:

(a) The Banach space $F_d = \mathrm{Lip}(\mathbb{R}^d)$ of Lipschitz continuous functions equipped with the norm

$$\|f\|_d = |f(0)| + \sup_{\xi \neq \tilde{\xi}} \frac{|f(\xi) - f(\tilde{\xi})|}{\|\xi - \tilde{\xi}\|}.$$

The best possible convergence rate is $e(Q_{n,d}) = O(n^{-\frac{1}{d}})$. It is attained for $w_i = \frac{1}{n}$ and certain ξ^i, $i = 1, \ldots, n$, if P has finite moments of order $1 + \delta$ for some $\delta > 0$. (Graf-Luschgy 00)

(b) The tensor product Sobolev space

$$F_{d,\gamma} = \mathcal{W}_{2,\mathrm{mix}}^{(1,\ldots,1)}([0,1]^d) = \bigotimes_{j=1}^{d} W_2^1([0,1])$$

of real functions on $[0,1]^d$ having first order mixed weak derivatives with the (weighted) norm

$$\|f\|_{d,\gamma} = \left(\sum_{u \subset D} \gamma_u^{-1} \int_{[0,1]^{|u|}} \left| \frac{\partial^{|u|}}{\partial \xi^u} f(\xi^u, 1^{-u}) \right|^2 d\xi^u \right)^{\frac{1}{2}},$$

where $D = \{1, \ldots, d\}$, $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_d > 0$, $\gamma_\emptyset = 1$ and

$$\gamma_u = \prod_{j \in u} \gamma_j \quad (u \subseteq D).$$

For n prime, $w_i = \frac{1}{n}$, and a suitable choice of (γ_j), points $\xi^i \in [0,1]^d$, $i = 1, \ldots, n$, can be constructed such that

$$e(Q_{n,d}) \leq C(\delta) n^{-1+\delta} \|I_d\|$$

for some $C(\delta) > 0$ (not depending on d) and all $0 < \delta \leq \frac{1}{2}$.

(Sloan, Woźniakowski 98, Kuo 03)

Scenario generation methods

We will discuss the following three scenario generation methods for stochastic programs *without nonanticipativity constraints*:

(a) Monte Carlo sampling from the underlying probability distribution P on $\mathbb{R}^d$ (Shapiro 03).

(b) Optimal quantization of probability distributions (Pflug-Pichler 10).

(c) Quasi-Monte Carlo methods (Koivu-Pennanen 05, Homem-de-Mello 06).

Monte Carlo sampling

Monte Carlo methods are based on drawing independent identically distributed (iid) Ξ-valued random samples $\xi^1(\cdot), \ldots, \xi^n(\cdot), \ldots$ (defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$) from an underlying probability distribution P (on Ξ) such that

$$Q_{n,d}(\omega)(f) = \frac{1}{n} \sum_{i=1}^{n} f(\xi^i(\omega)),$$

i.e., $Q_{n,d}(\cdot)$ is a random functional, and it holds

$$\lim_{n \to \infty} Q_{n,d}(\omega)(f) = \int_{\Xi} f(\xi) P(d\xi) = \mathbb{E}(f) \quad \mathbb{P}\text{-almost surely}$$

for every real continuous and bounded function f on Ξ.

If P has finite moment of order $r \geq 1$, the error estimate

$$\mathbb{E}\left(\left| \frac{1}{n} \sum_{i=1}^{n} f(\xi^i(\omega)) - \mathbb{E}(f) \right|^r \right) \leq \frac{\mathbb{E}\left((f - \mathbb{E}(f))^r\right)}{n^{r-1}}$$

is valid.

Hence, the mean square convergence rate is

$$\|Q_{n,d}(\omega)(f) - \mathbb{E}(f)\|_{L_2} = \sigma(f)n^{-\frac{1}{2}},$$

where $\sigma^2(f) = \mathbb{E}\left((f - \mathbb{E}(f))^2\right)$.
The latter holds without any assumption on f except $\sigma(f) < \infty$.

Advantages:
(i) MC sampling works *for (almost) all integrands*.
(ii) The machinery of probability theory is available.
(iii) The convergence *rate does not depend on* d.

Deficiencies: (Niederreiter 92)
(i) There exist 'only' *probabilistic error bounds*.
(ii) Possible regularity of the integrand *does not improve* the rate.
(iii) Generating (independent) random samples is *difficult*.

Practically, iid samples are approximately obtained by pseudo random number generators as uniform samples in $[0, 1]^d$ and later transformed to more general sets Ξ and distributions P.

Excellent pseudo random number generator: Mersenne Twister (Matsumoto-Nishimura 98).

Survey: L'Ecuyer 94.

Optimal quantization of probability measures

Let D be a distance of probability measures on $\mathbb{R}^d$ such that the underlying stochastic program behaves stable w.r.t. D (Römisch 03).

Example:

L_p-minimal metric ℓ_p for $p \geq 1$, i.e.

$$\ell_p(P,Q) = \inf\{(\mathbb{E}(\|\xi - \eta\|^p))^{\frac{1}{p}} : \mathcal{L}(\xi) = P,\, \mathcal{L}(\eta) = Q\}$$

Let P be a given probability distribution on $\mathbb{R}^d$. We are looking for a discrete probability measure Q_n with support

$$\mathrm{supp}(Q_n) = \{\xi^1, \ldots, \xi^n\} \quad \text{and} \quad Q_n(\{\xi^i\}) = \frac{1}{n},\, i = 1, \ldots, n,$$

that is the best approximation to P with respect to D, i.e.,

$$D(P, Q_n) = \min\{D(P,Q) : |\mathrm{supp}(Q)| = n, Q \text{ is uniform}\}.$$

Existence of best approximations, called optimal quantizers, and their best possible convergence rate $O(n^{-\frac{1}{d}})$ is well known for ℓ_p (Graf-Luschgy 00).

However, in general, the function

$$\Psi_D(\xi^1, \ldots, \xi^n) := D\Big(P, \frac{1}{n}\sum_{i=1}^{n} \delta_{\xi^i}\Big)$$

and, in particular,

$$\Psi_{\ell_p}(\xi^1, \ldots, \xi^n) = \Big(\int_{\mathbb{R}^d} \min_{i=1,\ldots,n} \|\xi - \xi^i\|^p P(d\xi)\Big)^{\frac{1}{p}}$$

is nonconvex and nondifferentiable on $\mathbb{R}^{dn}$.
Hence, the global minimization of Ψ_D is not an easy task.

Algorithmic procedures for minimizing Ψ_{ℓ_r} globally may be based on stochastic gradient (type) algorithms, stochastic approximation methods and stochastic branch-and-bound techniques (e.g. Pflug 01, Hochreiter-Pflug 07, Pagés 97, Pagés et al 04).

However, asymptotically optimal quantizers can be determined explicitly in a number of cases (Pflug-Pichler 10).

Quasi-Monte Carlo methods

The idea of Quasi-Monte Carlo (QMC) methods is to replace random samples in Monte Carlo methods by deterministic points ξ^i, $i \in \mathbb{N}$, that are uniformly distributed in $[0,1]^d$. QMC is of the form

$$Q_{n,d}(f) = \frac{1}{n} \sum_{i=1}^{n} f(\xi^i)$$

The uniform distribution property may be defined in terms of the so-called star-discrepancy of $\xi^1, \ldots, \xi^n$

$$D_n^*(\xi^1, \ldots, \xi^n) := \sup_{\xi \in [0,1]^d} \left| \lambda^d([0,\xi)) - \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{[0,\xi)}(\xi^i) \right|,$$

by calling a sequence $(\xi^i)_{i \in \mathbb{N}}$ uniformly distributed in $[0,1]^d$ if

$$D_n^*(\xi^1, \ldots, \xi^n) \to 0 \quad \text{for} \quad n \to \infty\,.$$

A **classical result** due to Roth 54 states

$$D_n^*(\xi^1, \ldots, \xi^n) \geq B_d \frac{(\log n)^{\frac{d-1}{2}}}{n}$$

for some constant B_d and all sequences (ξ^i) in $[0,1]^d$.

Classical convergence results:

Theorem: (Proinov 88)
If the real function f is continuous on $[0,1]^d$, then there exists $C > 0$ such that

$$|Q_{n,d}(f) - I_d(f)| \leq C\omega_f\left(D_n^*(\xi^1, \ldots, \xi^n)^{\frac{1}{d}}\right),$$

where $\omega_f(\delta) = \sup\{|f(\xi) - f(\tilde{\xi})| : \|\xi - \tilde{\xi}\| \leq \delta,\ \xi, \tilde{\xi} \in [0,1]^d\}$ is the modulus of continuity of f.

Theorem: (Koksma-Hlawka 61)
If f is of bounded variation in the sense of Hardy and Krause, it holds

$$|I_d(f) - Q_{n,d}(f)| \leq V_{\mathrm{HK}}(f) D_n^*(\xi^1, \ldots, \xi^n)\,.$$

for any $n \in \mathbb{N}$ and any $\xi^1, \ldots, \xi^n \in [0,1]^d$.

There exist sequences (ξ^i) in $[0,1]^d$ such that

$$D_n^*(\xi^1, \ldots, \xi^n) = O(n^{-1}(\log n)^{d-1}).$$

First general construction: Nets (Sobol 69, Niederreiter 87)

Elementary subintervals E of $[0,1)^d$ in base b:

$$E = \prod_{j=1}^{d} \left[\frac{a_j}{b^{d_j}}, \frac{a_j+1}{b^{d_j}}\right),$$

with $a_i, d_i \in \mathbb{Z}_+, 0 \le a_i < d_i, i = 1, \ldots, d$.

A set of b^m points in $[0,1]^d$ is a (t,m,d)-net in base b if every elementary subinterval E in base b with $\lambda^d(E) = b^{t-m}$ contains b^t points ($m, t \in \mathbb{Z}_+$, $m > t$).

A sequence (ξ^i) in $[0,1]^d$ is a (t,d)-sequence in base b if, for all integers $k \in \mathbb{Z}_+$ and $m > t$, the set

$$\{\xi^i : kb^m \le i < (k+1)b^m\}$$

is a (t,m,d)-net in base b.

Proposition: $(0,d)$-sequences exist if $d \le b$.

Theorem: A $(0,m,d)$-net $\{\xi^i\}$ in base b satisfies

$$D_n^*(\xi^i) \leq A_d(b)\frac{(\log n)^{d-1}}{n} + O\left(\frac{(\log n)^{d-2}}{n}\right).$$

with reasonably small constants $A_d(b)$.

Special cases: Sobol, Faure and Niederreiter sequences.

Second general construction: Lattices (Korobov 59, Sloan-Joe 94)

Let $g \in \mathbb{Z}^d$ and consider the lattice points

$$\left\{\xi^i = \left\{\frac{i}{n}g\right\} : i = 1,\ldots,n\right\},$$

where $\{z\}$ is defined componentwise and for $z \in \mathbb{R}_+$ it is the *fractional part* of z, i.e., $\{z\} = z - \lfloor z \rfloor \in [0,1)$.

Randomly shifted lattice points with a uniform random vector $\triangle$:

$$\left\{\xi^i = \left\{\frac{i}{n}g + \triangle\right\} : i = 1,\ldots,n\right\},$$

There is a component-by-component construction algorithm for g such that for some constant $C(\delta)$ and all $0 < \delta \leq \frac{1}{2}$

$$e(Q_{n,d}) \leq C(\delta) n^{-1+\delta} \|I_d\| \quad \text{(Sloan-Kuo 05, Kuo 03).}$$

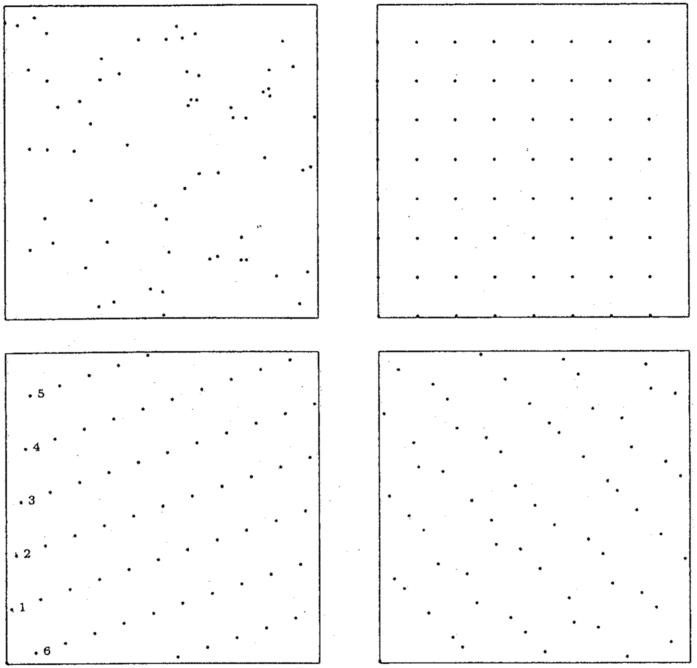


Fig. 5.3 Four different point sets with $n = 64$: random (top left), rectangular grid (top right), Korobov lattice (bottom left), and Sobol' (bottom right).

Convergence rates for unbounded integrands ?

(Kuo-Sloan-Wasilkowski-Waterhouse 10)

Let us consider

$$I_{d,\rho}(f) = \int_{\mathbb{R}^d} f(\xi)\rho(\xi)d\xi \quad \text{with} \quad \rho(\xi) = \prod_{j=1}^{d} \phi(\xi_j)$$

and strictly positive ϕ (w.l.o.g.).

Transformation:

$$I_{d,\rho}(f) = I_d(g) = \int_{(0,1)^d} g(u)du, \quad \text{where}$$

$$g(u) = f(\Phi^{-1}(u)) := f(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)) \text{ and } \Phi(\xi) = \int_{-\infty}^{\xi} \phi(t)dt$$

Absolute error:

$$e(Q_{n,d}) = \sup_{f \in \mathbb{B}_d} \left| \int_{(0,1)^d} f(\Phi^{-1}(u))du - \frac{1}{n}\sum_{i=1}^{n} f(\Phi^{-1}(u^i)) \right|$$

where $u^i \in (0,1)^d$, $i = 1, \ldots, n$.

Rates of convergence for unbounded integrands are known for several densities ϕ and close to those for $[0,1]^d$.

Is QMC efficient in stochastic programming ?

Problem: Typical integrands in linear stochastic programming are not smooth and, hence, do not belong to the relevant function spaces in general.
Idea: Study of the efficient dimension of typical integrands.

ANOVA-decomposition of f:

$$f = \sum_{u \subseteq D} f_u \,,$$

where $f_\emptyset = I_d(f) = I_D(f)$ and recursively

$$f_u = I_{-u}(f) + \sum_{v \subseteq u} (-1)^{|u|-|v|} I_{u-v}(I_{-u}(f)) \,,$$

where I_{-u} means integration with respect to ξ_j in $[0,1]$, $j \in D \setminus u$ and $D = \{1, \ldots, d\}$. Hence, f_u is essentially as smooth as $I_{-u}(f)$ and does not depend on ξ^{-u}.

Proposition: The functions $\{f_u\}_{u \subseteq D}$ are orthogonal in $L_2([0,1]^d)$.

We set $\sigma^2(f) = \|f - I_d(f)\|_{L_2}^2$ and have

$$\sigma^2(f) = \|f\|_{L_2}^2 - (I_d(f))^2 = \sum_{\emptyset \neq u \subseteq D} \|f_u\|_{L_2}^2 .$$

The truncation dimension d_t of f is the smallest $d_t \in \mathbb{N}$ such that

$$\sum_{u \subseteq \{1,\ldots,d_t\}} \|f_u\|_{L_2}^2 \geq \alpha \sigma^2(f) \quad \text{(where } \alpha \in (0,1) \text{ is close to } 1\text{)}.$$

Then

$$\|f - \sum_{u \subseteq \{1,\ldots,d_t\}} f_u\|_{L_2}^2 \leq (1-\alpha)\sigma^2(f).$$

Most of the ANOVA terms f_u may be smoother than f under certain conditions.

(Griebel-Kuo-Sloan 10).

A note on scenario reduction

Assume that the stochastic program behaves stable with respect to ℓ_p for some $p \geq 1$.

Let us consider discrete probability distributions P with scenarios ξ^i and probabilities p_i, $i = 1, \ldots, N$, and Q being supported by a given subset of scenarios ξ^j, $j \notin J \subset \{1, \ldots, N\}$, of P.

The best approximation of P with respect to ℓ_p given an index set J exists and is denoted by Q^*. It has the distance

$$D_J := \ell_p(P, Q^*) = \min_Q \ell_p(P, Q) = \Big(\sum_{i \in J} p_i \min_{j \notin J} \|\xi^i - \xi^j\|^p \Big)^{\frac{1}{p}}$$

and the probabilities $q_j^* = p_j + \sum_{i \in J_j} p_i$, $\forall j \notin J$, where

$J_j := \{i \in J : j = j(i)\}$ and $j(i) \in \arg\min_{j \notin J} \|\xi^i - \xi^j\|$, $\forall i \in J$

(optimal redistribution) (Dupačová-Gröwe-Römisch 03).

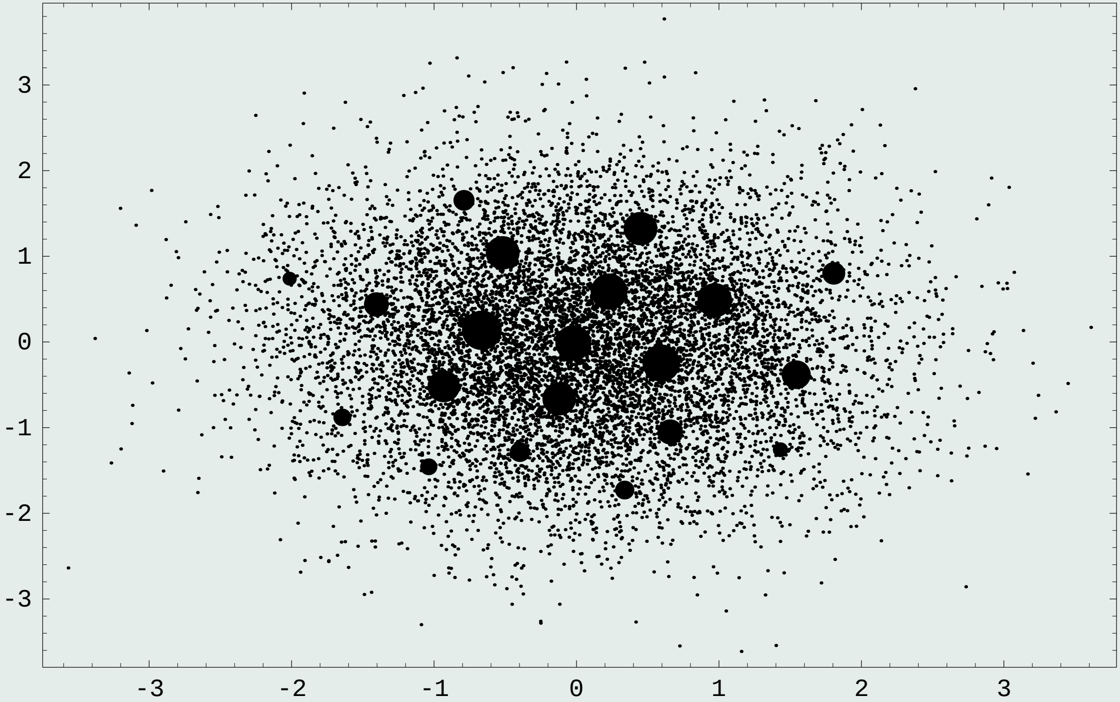
For mixed-integer two-stage stochastic programs the relevant distance is a polyhedral discrepancy. In that case, the new weights have to be determined by linear programming (Henrion-Küchler-Römisch 08, 09).

Determining the optimal index set J with prescribed cardinality $N-n$ is a combinatorial optimization problem:

$$\min \{D_J : J \subset \{1, ..., N\}, |J| = N-n\}$$

Hence, the problem of finding the optimal index set J of scenarios to delete is $\mathcal{NP}$-hard and polynomial time algorithms are not available in general.

$\Longrightarrow$ Heuristics are used to determine J.



Scenario reduction w.r.t. ℓ_1 from N=10 000 MC samples of $N(0, I)$ in $\mathbb{R}^2$ to $n = 20$. The diameters of the circles are proportional to their probabilities

Generation of scenario trees

In multistage stochastic programs the decisions x have to satisfy the additional information constraint that x_t is measurable with respect to $\mathcal{F}_t = \sigma(\xi_\tau, \tau = 1, \ldots, t)$, $t = 1, \ldots, T$. The increase of the σ-fields $\mathcal{F}_t$ w.r.t. t is reflected by approximating the underlying stochastic process $\xi = (\xi_t)_{t=1}^T$ by scenarios forming a scenario tree.

Some recent approaches:

(1) Bound-based approximation methods: Kuhn 05, Casey-Sen 05.

(2) Monte Carlo-based schemes: Shapiro 03, 06.

(3) Quasi-Monte Carlo methods: Pennanen 06, 09 .

(4) Moment-matching principle: Høyland-Kaut-Wallace 03.

(5) Optimal quantization: Pagés et al. 03.

(6) Stability-based approximations: Hochreiter-Pflug 07, Mirkov-Pflug 07, Pflug-Pichler 10, Heitsch-Römisch 05, 09.

Survey: Dupačová-Consigli-Wallace 00

Theoretical basis of (6):

Quantitative stability results for multi-stage stochastic programs.

(Heitsch-Römisch-Strugarek 06; Mirkov-Pflug 07, Pflug 09)

Scenario tree generation: (Heitsch-Römisch 09)

(i) Generate a number of scenarios by one of the methods discussed earlier.

(ii) Construction of a scenario tree out of these scenarios by recursive scenario reduction and bundling over time such that the optimal value stays within a prescribed tolerance.

Implementation: GAMS-SCENRED 2.0 (developed by H. Heitsch)

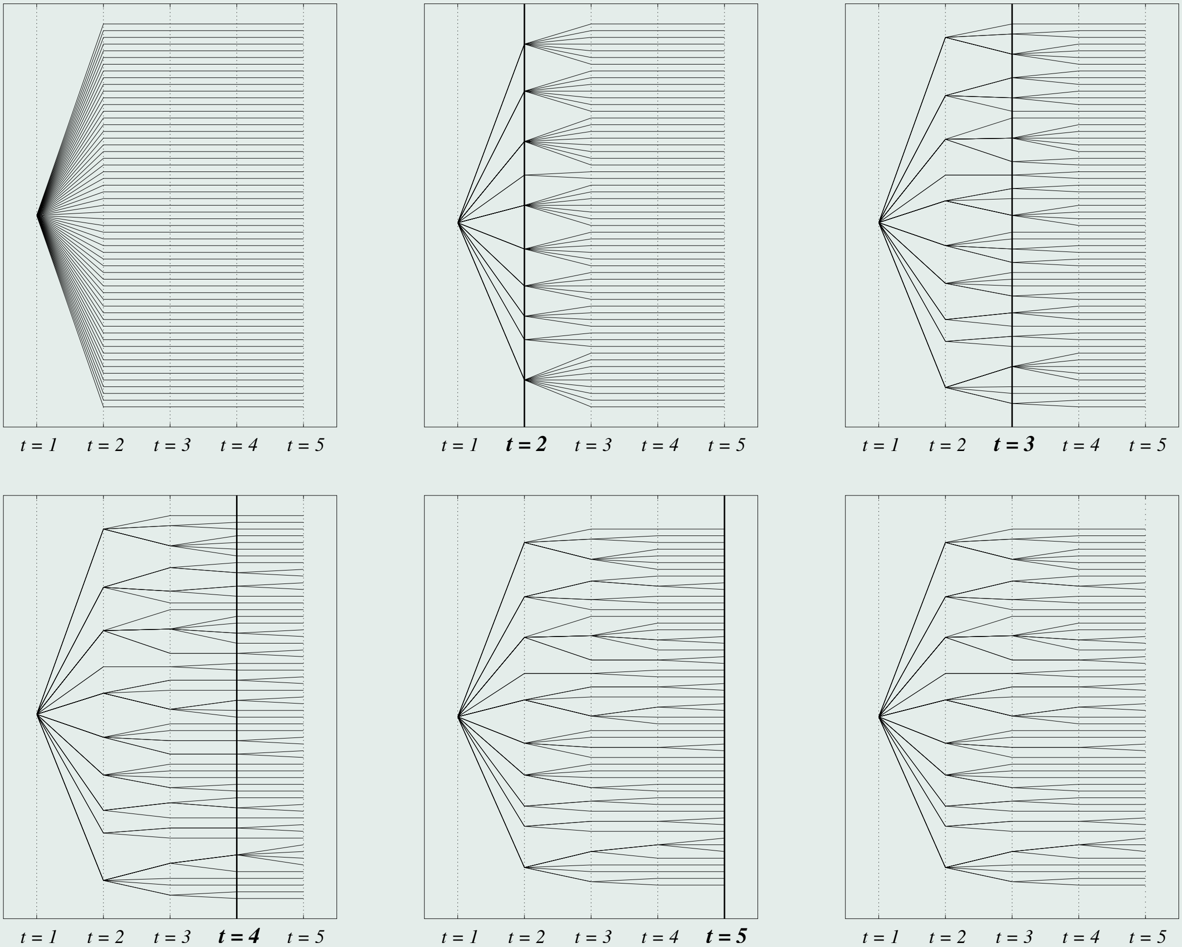
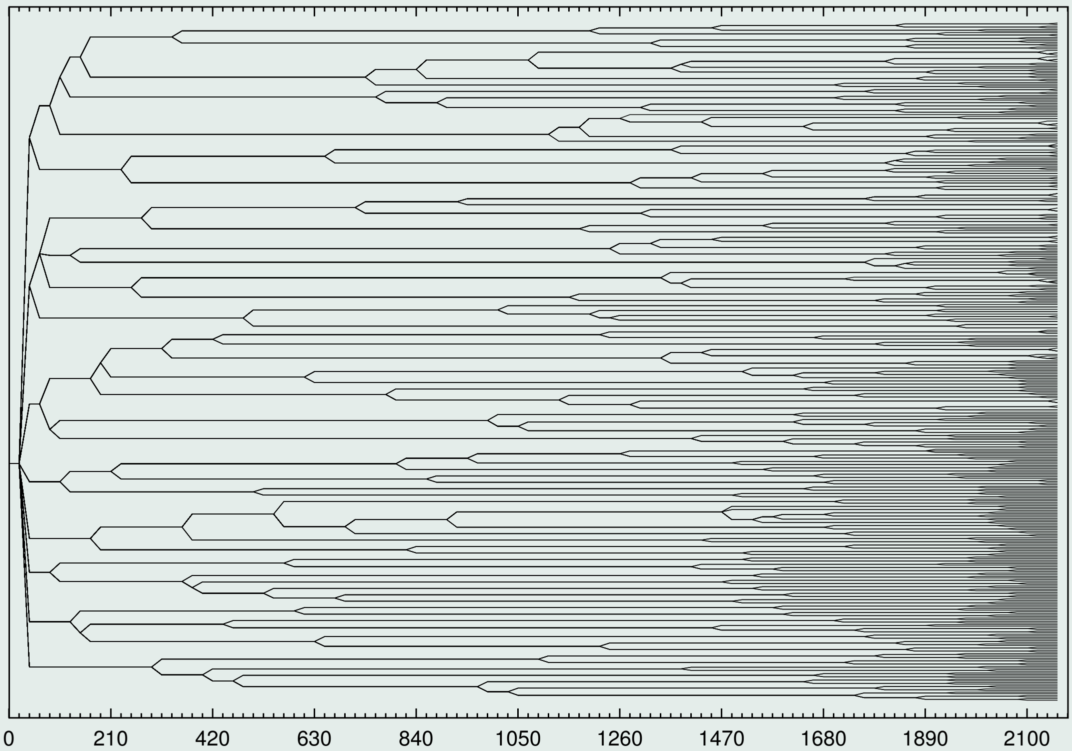


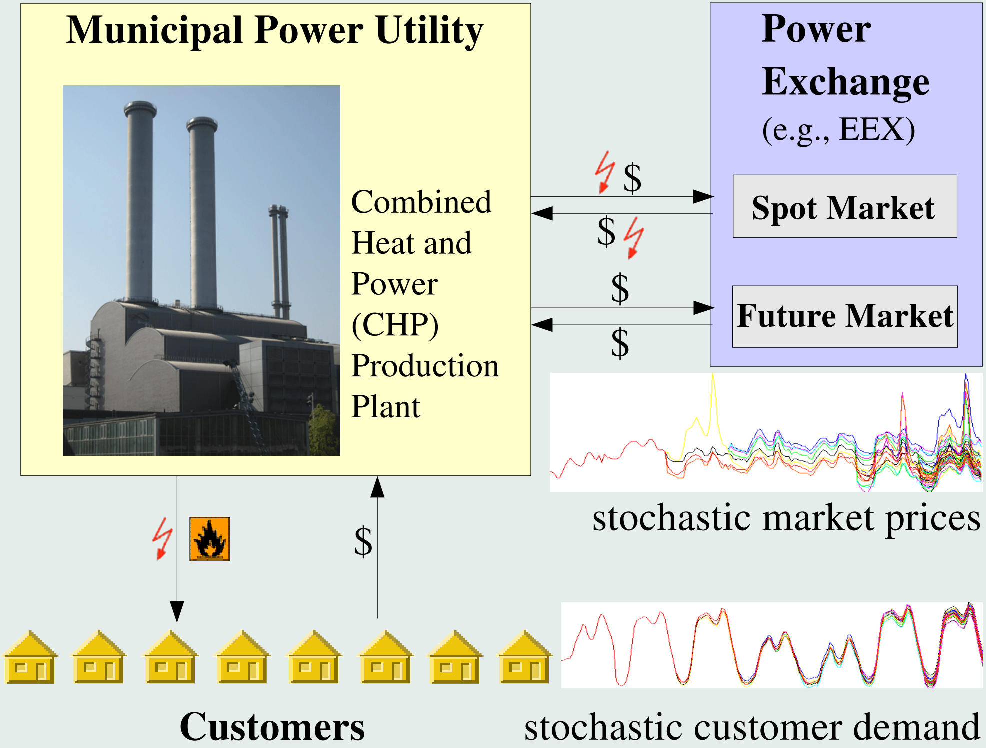
Illustration of the forward tree generation for an example including T=5 time periods starting with a scenario fan containing N=58 scenarios

<Start Animation>



Two-yearly demand-inflow scenario tree with weekly branchings for French EDF

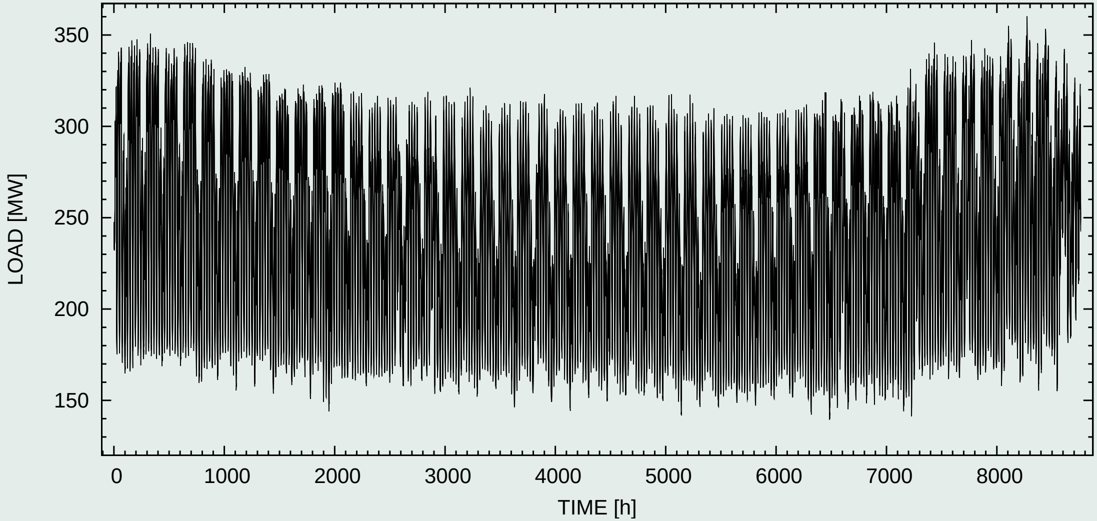
Mean-Risk Electricity Portfolio Management



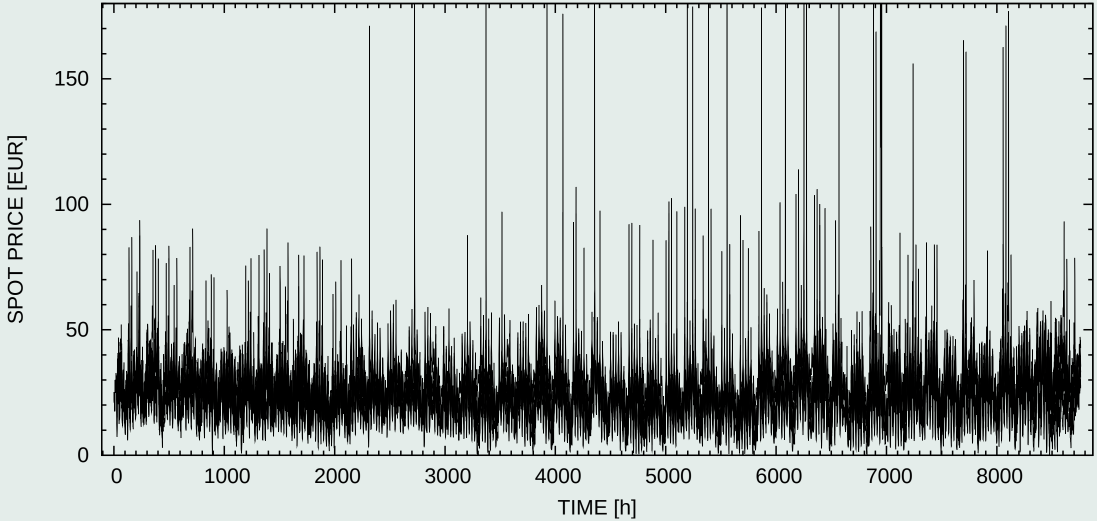
We consider the electricity portfolio management of a German municipal electric power company. Its portfolio consists of the following positions:

- power production (based on company-owned thermal units),
- bilateral contracts,
- (physical) (day-ahead) spot market trading (e.g., European Energy Exchange (EEX)) and
- (financial) trading of futures.

The time horizon is discretized into hourly intervals. The underlying stochasticity consists in a multivariate stochastic load and price process that is approximately represented by a finite number of scenarios. The objective is to maximize the total expected revenue and to minimize the risk. The portfolio management model is a large scale (mixed-integer) multi-stage stochastic program.



Time plot of yearly load profile



Time plot of yearly spot price profile

Statistical models and scenario trees

For the stochastic input data of the optimization model (here yearly electricity and heat demand, and electricity spot prices), a statistical model is employed.

- cluster classification for the intra-day (demand and price) profiles,

- Three-dimensional time series model for the daily average values (deterministic trend functions, a trivariate ARMA model for the (stationary) residual time series),

- Generation of scenarios by computing Monte Carlo samples from the multivariate normal distribution that corresponds to the ARMA process, and adding on trend functions as well as matched intra-day profiles from the clusters afterwards,

Intended modification: QMC samples instead of MC.

- generation of scenario trees (Heitsch-Römisch 09).

Numerical results

Test runs were performed on real-life data of a German municipal power company leading to a linear program containing $T = 365 \cdot 24 = 8760$ time steps, a scenario tree with 40 demand-price scenarios (see below) with about 150.000 nodes. The objective function is of the form

$$\text{Minimize} \quad \gamma\rho(z) - (1-\gamma)\mathbb{E}(z_T)$$

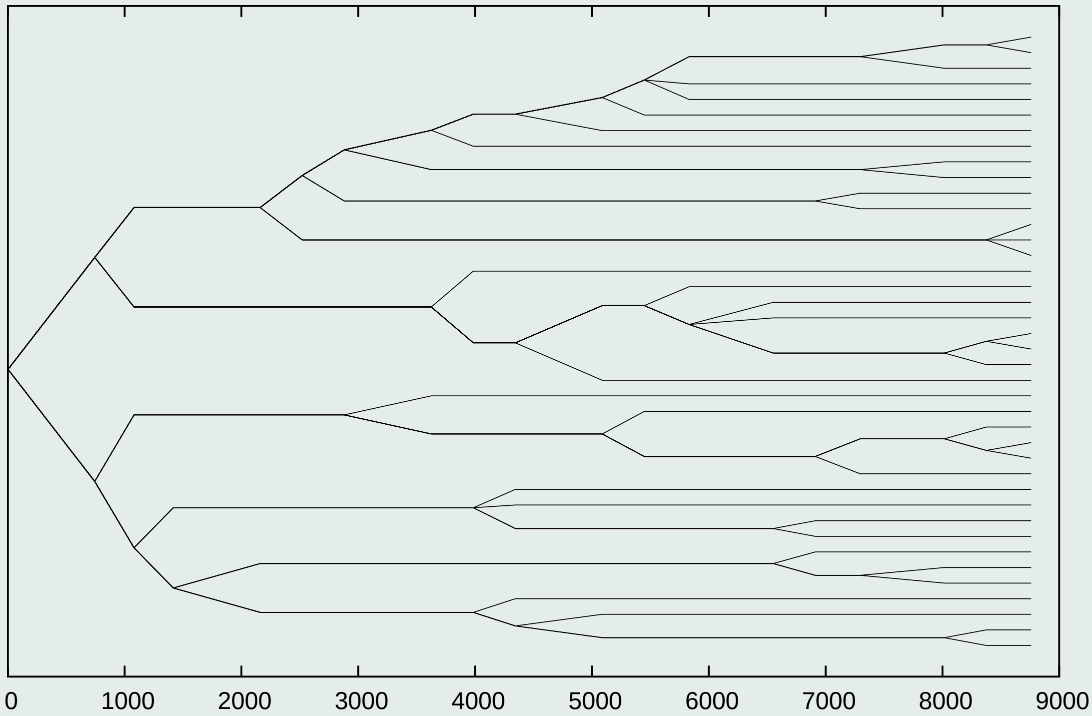
with a (multiperiod) risk measures ρ with risk aversion parameter $\gamma \in [0,1]$ ($\gamma = 0$ corresponds to the risk-neutral case).

Two risk measures:

(1) $\rho(z) = \mathbb{AV@R}_{0.05}(z_T)$ (Average or Conditional Value-at-risk)

(2) $\rho(z) = \rho_m(z) = \mathbb{AV@R}_{0.05}(\min_{j=1,\dots,J} z_{t_j})$

(t_j, $j = 1, \dots, J = 52$, are the risk measuring time steps; they correspond to 11 pm at the last trading day of each week).



Yearly scenario tree for the trivariate load-price process

Home Page
Title Page
Contents
◀◀ ▶▶
◀ ▶
Page 38 of 100
Go Back
Full Screen
Close
Quit

It turns out that the numerical results for the expected maximal revenue and minimal risk

$$\mathbb{E}(z_T^{*\gamma}) \qquad \text{and} \qquad \rho(z_{t_1}^{*\gamma}, \ldots, z_{t_J}^{*\gamma})$$

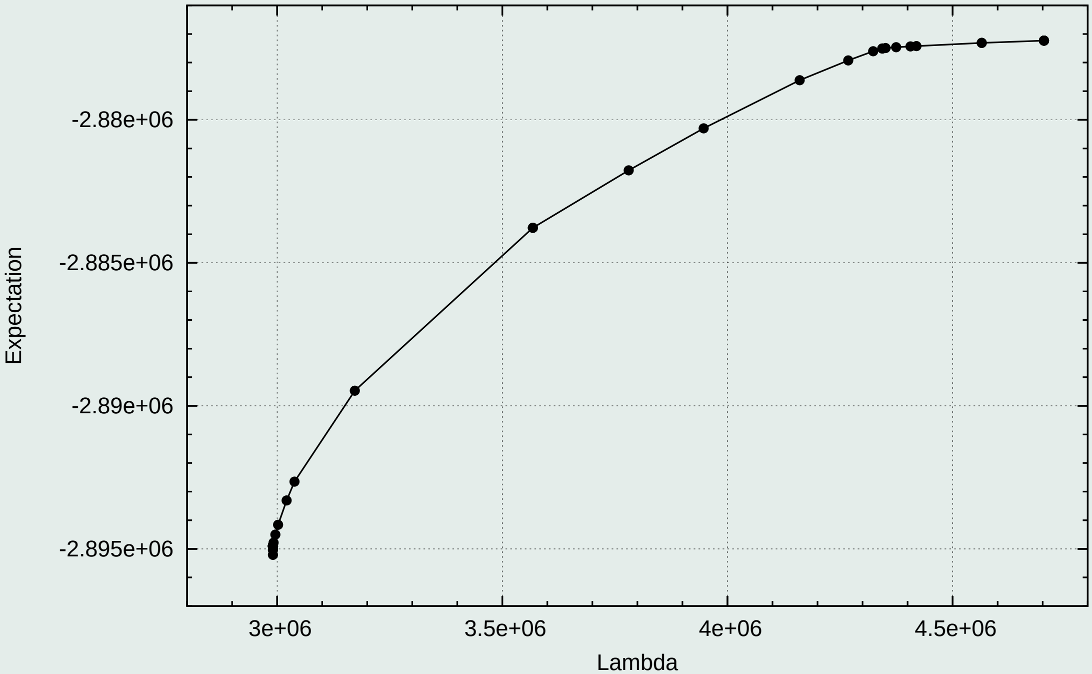
with the optimal revenue process $z^{*\gamma}$ are (almost) **identical** for $\gamma \in [0.15, 0.95]$ and the risk measures used in the test runs.

The efficient frontier

$$\gamma \mapsto \left(\rho(z_{t_1}^{*\gamma}, \ldots, z_{t_J}^{*\gamma}), \mathbb{E}(z_T^{*\gamma})\right)$$

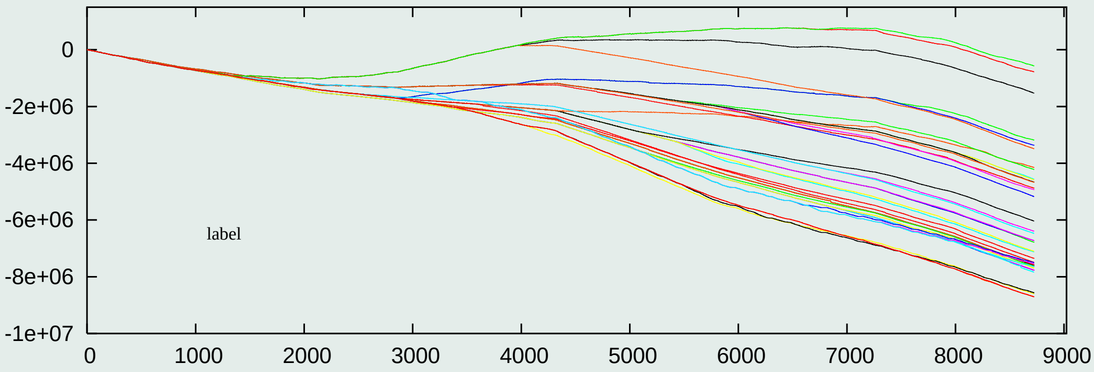
is concave for $\gamma \in [0, 1]$.

Risk aversion costs less than 1% of the expected overall revenue.

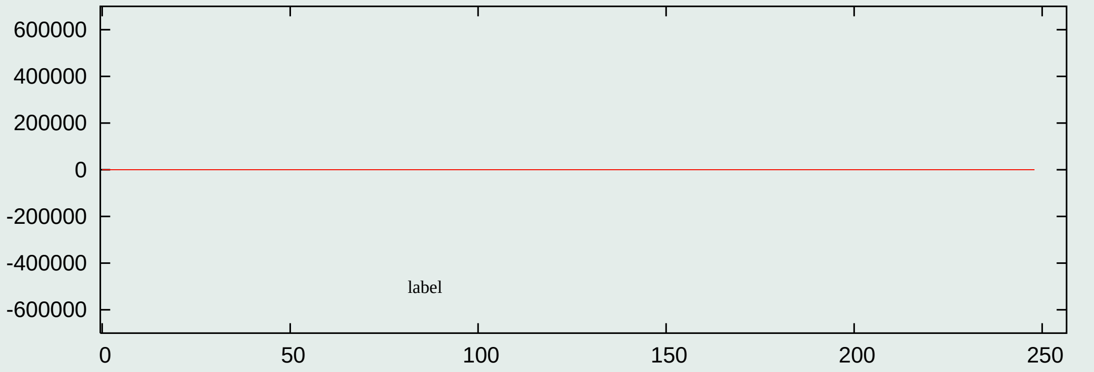


Efficient frontier

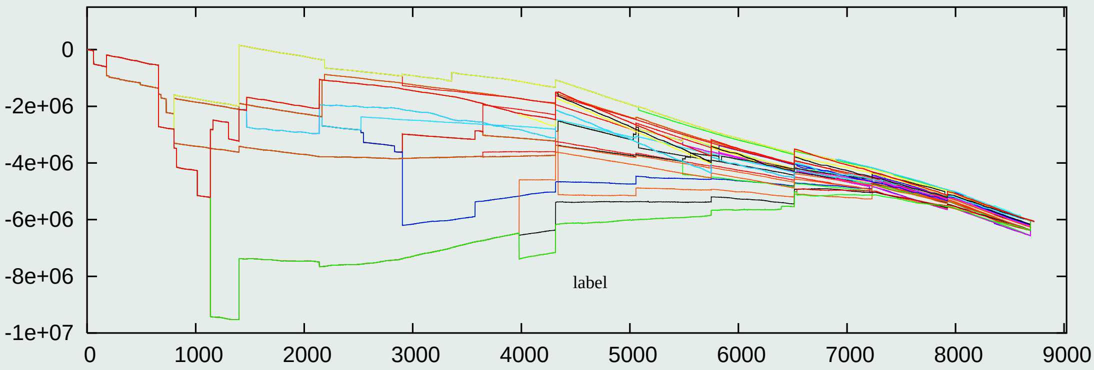
The LP is solved by CPLEX 9.1 in about 1 h running time on a 2 GHz Linux PC with 1 GB RAM.



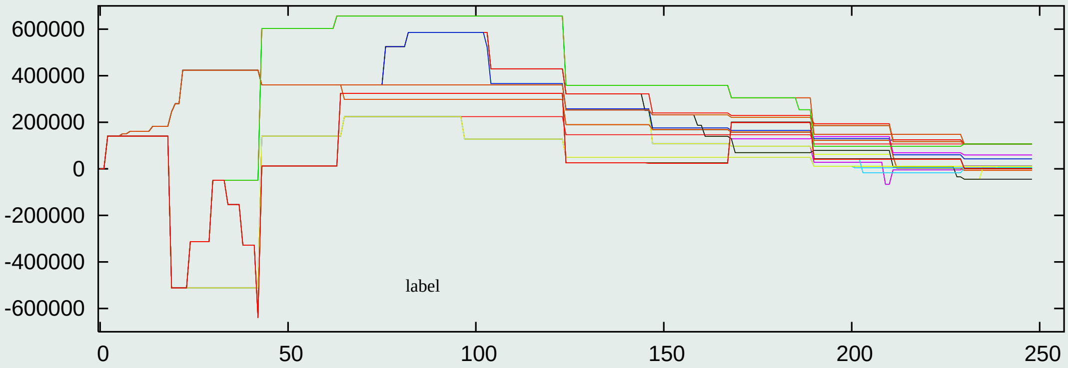
Overall revenue scenarios for $\gamma = 0$



Future trading for $\gamma = 0$



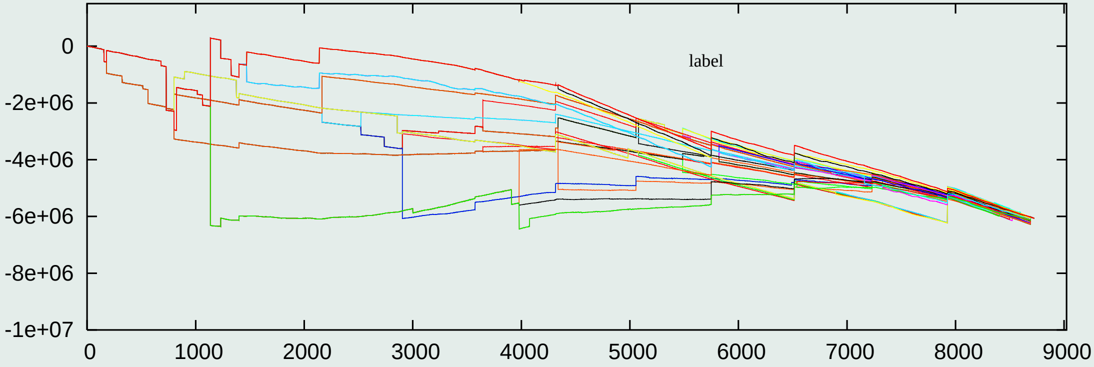
Overall revenue scenarios with $\mathbb{A}\mathrm{V@R}_{0.05}$ and $\gamma = 0.9$



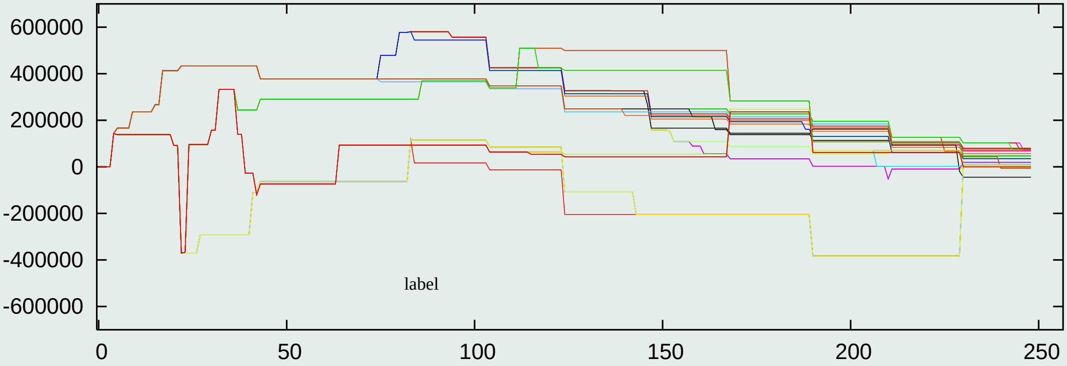
Future trading with $\mathbb{A}\mathrm{V@R}_{0.05}$ and $\gamma = 0.9$

Home Page
Title Page
Contents
◀◀ ▶▶
◀ ▶

Go Back
Full Screen
Close
Quit



Overall revenue scenarios with ρ_m and $\gamma = 0.9$



Future trading with ρ_m and $\gamma = 0.9$

Home Page
Title Page
Contents
Page 43 of 100
Go Back
Full Screen
Close
Quit

Conclusions

- A survey of approaches for scenario generation in stochastic optimization was presented.
- We outlined that a theoretical basis for applying Quasi-Monte Carlo in stochastic programming is still open.
- Strategies for scenario reduction and scenario tree generation were briefly discussed.
- Numerical results for a risk-neutral and risk-averse yearly electricity portfolio management model were presented.

References

Monographs:

H. Niederreiter: *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992.

C. Lemieux: *Monte Carlo and Quasi-Monte Carlo Sampling*, Springer, 2009.

J. Dick, F. Pillichshammer: *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*, Cambridge University Press, 2010.

I. H. Sloan, S. Joe: *Lattice Methods for Multiple Integration*, Clarendon Press, Oxford 1994.

S. Graf, H. Luschgy: *Foundations of Quantization for Probability Distributions*, Lecture Notes in Mathematics, Vol. 1730, Springer, Berlin, 2000.

Monte Carlo:

P. L'Ecuyer: Uniform random number generation, *Annals of Operations Research* 53 (1994), 77–120.

M. Matsumoto, T. Nishimura: Mersenne Twister: A 623-dimensionally equidistributed uniform pseudo-random number generator, *ACM Transactions on Modeling and Computer Simulation* 8 (1998), 3–30.

A. Genz: Numerical computation of multivariate normal probabilities, J. Comp. Graph. Stat. 1 (1992), 141–149.

A. Shapiro: Monte Carlo sampling methods, in: *Stochastic Programming* (A. Ruszczyński, A. Shapiro eds.), Handbooks in Operations Research and Management Science, Volume 10, Elsevier, Amsterdam 2003, 353–425.

A. Shapiro: Inference of statistical bounds for multistage stochastic programming problems, *Mathematical Methods of Operations Research* 58 (2003), 57–68.

Quantization:

W. Römisch: Stability of stochastic programming problems, in: *Stochastic Programming* (A. Ruszczyński, A. Shapiro eds.), Handbooks in Operations Research and Management Science, Volume 10, Elsevier, Amsterdam 2003, 483–554.

G. Pagès: A space vector quantization method for numerical integration, *J. Computat. Appl. Math.* 89 (1997), 1–38.

V. Bally, G. Pagès, J. Printems: First order schemes in the numerical quantization method, *Math. Finance* 13 (2003), 1–16.

V. Bally, G. Pagès: A quantization algorithm for solving multidimensional discrete-time optimal stopping problems, *Bernoulli* 9 (2003), 1003–1049.

G. Pagès, H. Pham, J. Printems: Optimal quantization methods and applications to numerical problems in finance, in *Handbook of Computational and Numerical Methods in Finance* (S.T. Rachev Ed.), Birkhäuser, Boston 2004, 253–297.

G. Ch. Pflug, A. Pichler: Scenario generation for stochastic optimization problems, in:

Stochastic Optimization Methods in Finance and Energy (M.I. Bertocchi, G. Consigli, M.A.H. Dempster eds.) (to appear).

G. Ch. Pflug: Scenario tree generation for multiperiod financial optimization by optimal discretization, *Mathematical Programming* 89 (2001), 251–271.

Quasi-Monte Carlo:

T. Pennanen, M. Koivu: Epi-convergent discretizations of stochastic programs via integration quadratures, *Numerische Mathematik* 100 (2005), 141–163.

T. Homem-de-Mello: On rates of convergence for stochastic optimization problems under non-i.i.d. sampling, *SIAM Journal on Optimization* 19 (2008), 524-551.

S. S. Drew, T. Homem-de-Mello: Quasi-Monte Carlo strategies for stochastic optimization, *Proceedings of the 2006 Winter Simulation Conference*, IEEE, 2006, 774–782.

I. M. Sobol': *Multidimensional Quadrature Formulas and Haar Functions*, Nauka, Moscow, 1969 (in Russian).

I. H. Sloan, H. Woźniakowski: When are Quasi Monte Carlo algorithms efficient for high-dimensional integration, *Journal of Complexity* 14 (1998), 1–33.

F. Y. Kuo, I. H. Sloan: Lifting the curse of dimensionality, *Notices of the AMS* 52 (2005), 1320–1328.

F. Y. Kuo: Component-by-component constructions achieve the optimal rate of convergence in weighted Korobov and Sobolev spaces, *Journal of Complexity* 19 (2003), 301-320.

F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, B. J. Waterhouse: Randomly shifted lattice rules with the optimal rate of convergence for unbounded integrands, *Journal of Complexity* 26 (2010), 135–160.

A. B. Owen: Multidimensional variation for Quasi-Monte Carlo, in J. Fan, G. Li (Eds.), International Conference on Statistics, World Scientific Publ., 2005, 49–74.

X. Wang, K.-T. Fang: The effective dimension and Quasi-Monte Carlo integration, *Journal of Complexity* 19 (2003), 101–124.

F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, H. Woźniakowski: On decomposition of multivariate functions, *Mathematics of Computation* 79 (2010), 953–966.

M. Griebel, F. Y. Kuo, I. H. Sloan: The smoothing effect of the ANOVA decomposition, *Journal of Complexity* 26 (2010), 523–551.

Scenario reduction:

J. Dupačová, N. Gröwe-Kuska, W. Römisch: Scenario reduction in stochastic programming: An approach using probability metrics, *Mathematical Programming* 95 (2003), 493–511.

W. Römisch and R. J-B Wets: Stability of epsilon-approximate solutions to convex stochastic programs, *SIAM Journal on Optimization* 18 (2007), 961–979.

H. Heitsch, W. Römisch: A note on scenario reduction for two-stage stochastic programs, *Operations Research Letters* 35 (2007), 731–738.

R. Henrion, C. Küchler, W. Römisch: Discrepancy distances and scenario reduction in two-stage stochastic mixed-integer programming, *Journal of Industrial and Management Optimization* 4 (2008), 363–384.

R. Henrion, C. Küchler, W. Römisch: Scenario reduction in stochastic programming with respect to discrepancy distances, *Computational Optimization and Applications* 43 (2009), 67–93.

S, T. Rachev, L. Rüschendorf: *Mass Transportation Problems*, Vol. I, Springer, Berlin, 1998.

Scenario trees:

J. Dupačová, G. Consigli, S. W. Wallace: Scenarios for multistage stochastic programs, *Annals of Operations Research* 100 (2000), 25–53.

M. Casey, S. Sen: The scenario generation algorithm for multistage stochastic linear programming, *Mathematics of Operations Research* 30 (2005), 615–631.

K. Frauendorfer: Barycentric scenario trees in convex multistage stochastic programming, *Mathematical Programming* Ser. B, 75 (1996), 277–293.

H. Heitsch, W. Römisch: Scenario tree modeling for multistage stochastic programs, *Mathematical Programming* 118 (2009), 371–406.

H. Heitsch, W. Römisch: Scenario tree reduction for multistage stochastic programs, *Computational Management Science* 6 (2009), 117–133.

R. Hochreiter, G. Ch. Pflug: Financial scenario generation for stochastic multi-stage decision processes as facility location problem, *Annals of Operations Research* 152 (2007), 257–272.

K. Høyland, M. Kaut, S. W. Wallace: A heuristic for moment-matching scenario generation, *Computational Optimization and Applications* 24 (2003), 169–185.

D. Kuhn: *Generalized bounds for convex multistage stochastic programs*, Lecture Notes in Economics and Mathematical Systems, Vol. 548, Springer, Berlin, 2005.

T. Pennanen: Epi-convergent discretizations of multistage stochastic programs via integration quadratures, *Mathematical Programming* 116 (2009), 461–479.

R. Mirkov, G. Ch. Pflug: Tree approximations of dynamic stochastic programs, *SIAM Journal on Optimization* 18 (2007), 1082–1105.

G. Ch. Pflug: Version independence and nested distributions in multistage stochastic optimization, *SIAM Journal on Optimization* 20 (2009), 1406–1420.

Electricity portfolio management:

A. Eichhorn, H. Heitsch, W. Römisch: Stochastic optimization of electricity portfolios: Scenario tree modeling and risk management, in: Handbook of Power Systems, Vol. II (S. Rebennack, P. Pardalos, M. Pereira, N. Iliadis eds.), Springer, 2010.

A. Eichhorn, W. Römisch and I. Wegner: Mean-risk optimization of electricity portfolios using multiperiod polyhedral risk measures, IEEE St. Petersburg Power Tech 2005.