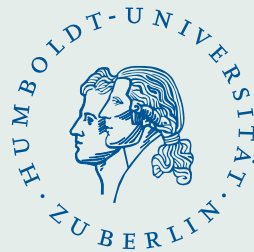


Stochastic Optimization: Approximation and Scenarios

W. Römisch

Humboldt-University Berlin
Institute of Mathematics

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



Colloquium "Continuous Optimization and Optimal Control"
TU Berlin, 23.06.17, 2017

Introduction

Many **stochastic programming models** may be traced back to minimizing an expectation functional on some closed subset of a Euclidean space or, eventually in addition, relative to some expectation constraint. Their general form is

$$(SP) \quad \min \left\{ \int_{\Xi} f_0(x, \xi) P(d\xi) : x \in X, \int_{\Xi} f_1(x, \xi) P(d\xi) \leq 0 \right\}$$

where X is a closed subset of \mathbb{R}^m , Ξ a closed subset of \mathbb{R}^s , P is a Borel probability measure on Ξ abbreviated by $P \in \mathcal{P}(\Xi)$. The functions f_0 and f_1 from $\mathbb{R}^m \times \Xi$ to the extended reals $\overline{\mathbb{R}} = (-\infty, \infty]$ are normal integrands.

For general continuous multivariate probability distributions P such stochastic optimization models (SP) are **not solvable** in general.

Many approaches for solving such optimization models computationally are based on finding a **discrete** probability measure P_n in

$$\mathcal{P}_n(\Xi) := \left\{ \sum_{i=1}^n p_i \delta_{\xi^i} : \xi^i \in \Xi, p_i \geq 0, i = 1, \dots, n, \sum_{i=1}^n p_i = 1 \right\}$$

for some $n \in \mathbb{N}$, which approximates P at least such that the corresponding optimal values of (SP) are close. The atoms ξ^i , $i = 1, \dots, n$, of P_n are often called **scenarios** in this context.

For example, typical integrands in **linear two-stage stochastic programming models** are

$$f_0(x, \xi) = \begin{cases} g(x) + \Phi(q(\xi), h(x, \xi)) & , q(\xi) \in D \\ +\infty & , \text{else} \end{cases} \quad \text{and } f_1(x, \xi) \equiv 0,$$

where X and Ξ are convex polyhedral, $g(\cdot)$ is a linear function, $q(\cdot)$ is affine, $D = \{q \in \mathbb{R}^{\bar{m}} : \{z \in \mathbb{R}^r : W^\top z - q \in Y^*\} \neq \emptyset\}$ denotes the convex polyhedral dual feasibility set, $h(\cdot, \xi)$ is affine for fixed ξ and $h(x, \cdot)$ is affine for fixed x , and Φ denotes the infimal function of the linear (second-stage) optimization problem

$$\Phi(q, t) := \inf\{\langle q, y \rangle : Wy = t, y \in Y\}$$

with (r, \bar{m}) matrix W and convex polyhedral cone $Y \subset \mathbb{R}^{\bar{m}}$.

Typical integrands f_1 appearing in **chance constrained programming** are of the form

$$f_1(x, \xi) = p - \mathbf{1}_{\mathcal{P}(x)}(\xi),$$

where $p \in (0, 1)$ is a probability level and $\mathbf{1}_{\mathcal{P}(x)}$ is the characteristic function of the polyhedron $\mathcal{P}(x) = \{\xi \in \Xi : h(x, \xi) \leq 0\}$ depending on x , where Ξ and h have the same properties as above.

Stability-based scenario generation

Let $v(P)$ and $S(P)$ denote the infimum and solution set of (SP). We are interested in their dependence on the underlying probability distribution P .

To state a stability result we introduce the following sets of functions and of probability distributions (both defined on Ξ)

$$\mathcal{F} = \{f_j(x, \cdot) : j = 0, 1, x \in X\},$$

$$\mathcal{P}_{\mathcal{F}} = \left\{ Q \in \mathcal{P}(\Xi) : -\infty < \int_{\Xi} \inf_{x \in X} f_j(x, \xi) Q(d\xi), \sup_{x \in X} \int_{\Xi} f_j(x, \xi) Q(d\xi) < +\infty, \forall j \right\}$$

and the (pseudo-) distance on $\mathcal{P}_{\mathcal{F}}$

$$d_{\mathcal{F}}(P, Q) = \sup_{f \in \mathcal{F}} \left| \int_{\Xi} f(\xi)(P - Q)(d\xi) \right| \quad (P, Q \in \mathcal{P}_{\mathcal{F}}).$$

At first sight the set $\mathcal{P}_{\mathcal{F}}$ seems to have a complicated structure. For typical applications, however, like for linear two-stage and chance constrained models, the sets $\mathcal{P}_{\mathcal{F}}$ or appropriate subsets allow a simple characterization, for example, as subsets of $\mathcal{P}(\Xi)$ satisfying certain moment conditions.

Proposition: We consider (SP) for $P \in \mathcal{P}_{\mathcal{F}}$, assume that X is compact and

- (i) the function $x \rightarrow \int_{\Xi} f_0(x, \xi)P(d\xi)$ is Lipschitz continuous on X ,
- (ii) the set-valued mapping $y \rightrightarrows \{x \in X : \int_{\Xi} f_1(x, \xi)P(d\xi) \leq y\}$ satisfies the Aubin property at $(0, \bar{x})$ for each $\bar{x} \in S(P)$.

Then there exist constants $L > 0$ and $\delta > 0$ such that the estimates

$$\begin{aligned} |v(P) - v(Q)| &\leq L d_{\mathcal{F}}(P, Q) \\ \sup_{x \in S(Q)} d(x, S(P)) &\leq \Psi_P(L d_{\mathcal{F}}(P, Q)) \end{aligned}$$

hold whenever $Q \in \mathcal{P}_{\mathcal{F}}$ and $d_{\mathcal{F}}(P, Q) < \delta$. The real-valued function Ψ_P is given by $\Psi_P(r) = r + \psi_P^{-1}(2r)$ for all $r \in \mathbb{R}_+$, where ψ_P is the growth function

$$\psi_P(\tau) = \inf_{x \in X} \left\{ \int_{\Xi} f_0(x, \xi)P(d\xi) - v(P) : d(x, S(P)) \geq \tau, x \in X, \int_{\Xi} f_1(x, \xi)P(d\xi) \leq 0 \right\}.$$

In case $f_1 \equiv 0$ only lower semicontinuity is needed in (i) and the estimates hold with $L = 1$ and for any $\delta > 0$. Furthermore, Ψ_P is lower semicontinuous and increasing on \mathbb{R}_+ with $\Psi_P(0) = 0$. (Rachev-Römisch 02)

The stability result suggests to choose discrete approximations from $\mathcal{P}_n(\Xi)$ for solving (SP) such that they solve the **best approximation problem**

$$(OSG) \quad \min_{P_n \in \mathcal{P}_n(\Xi)} d_{\mathcal{F}}(P, P_n).$$

at least approximately. Determining the scenarios of some solution to (OSG) may be called **optimal scenario generation**. This optimal choice of discrete approximations is **challenging** and not possible in general.

It was suggested in (Rachev-Römisch 02) to eventually enlarge the function class \mathcal{F} such that $d_{\mathcal{F}}$ becomes a metric distance and has further nice properties. This may lead, however, to **nonconvex nondifferentiable minimization problems (OSG)** for determining the optimal scenarios and to **unfavorable convergence rates** of

$$\left(\min_{P_n \in \mathcal{P}_n(\Xi)} d_{\mathcal{F}}(P, P_n) \right)_{n \in \mathbb{N}}.$$

Typical examples are to choose \mathcal{F} as bounded subset of some Banach space $C^{r,\alpha}(\Xi)$ with $r \in \mathbb{N}_0$, $\alpha \in (0, 1]$, and **convergence rate** $O(n^{-\frac{r+\alpha}{s}})$.

The road of probability metrics

Motivated by linear two-stage models one may consider

Fortet-Mourier metrics:

$$\zeta_r(P, Q) := d_{\mathcal{F}_r(\Xi)}(P, Q) := \sup \left| \int_{\Xi} f(\xi)(P - Q)(d\xi) : f \in \mathcal{F}_r(\Xi) \right|,$$

where the function class \mathcal{F}_r for $r \geq 1$ is given by

$$\begin{aligned} \mathcal{F}_r(\Xi) &:= \{f : \Xi \mapsto \mathbb{R} : f(\xi) - f(\tilde{\xi}) \leq c_r(\xi, \tilde{\xi}), \forall \xi, \tilde{\xi} \in \Xi\}, \\ c_r(\xi, \tilde{\xi}) &:= \max\{1, \|\xi\|^{r-1}, \|\tilde{\xi}\|^{r-1}\} \|\xi - \tilde{\xi}\| \quad (\xi, \tilde{\xi} \in \Xi). \end{aligned}$$

Proposition: (Rachev-Rüschendorf 98)

If Ξ is bounded, ζ_r may be reformulated as **dual transportation problem**

$$\zeta_r(P, Q) = \inf \left\{ \int_{\Xi \times \Xi} \hat{c}_r(\xi, \tilde{\xi}) \eta(d\xi, d\tilde{\xi}) : \pi_1 \eta = P, \pi_2 \eta = Q \right\},$$

where the **reduced cost** \hat{c}_r is a metric with $\hat{c}_r \leq c_r$ and given by the minimal cost flow problem

$$\hat{c}_r(\xi, \tilde{\xi}) := \inf \left\{ \sum_{i=1}^{n-1} c_r(\xi_{l_i}, \xi_{l_{i+1}}) : n \in \mathbb{N}, \xi_{l_i} \in \Xi, \xi_{l_1} = \xi, \xi_{l_n} = \tilde{\xi} \right\}.$$

The problem of optimal scenario generation (OSG) then reads

$$\min_{P_n \in \mathcal{P}_n(\Xi)} \zeta_r(P, P_n)$$

or

$$\min_{(\xi^1, \dots, \xi^n) \in \Xi^n} \int_{\Xi} \min_{j=1, \dots, n} \hat{c}_r(\xi, \xi^j) P(d\xi).$$

The function $(\xi^1, \dots, \xi^n) \mapsto \int_{\Xi} \min_{j=1, \dots, n} \hat{c}_r(\xi, \xi^j) P(d\xi)$ is continuous on Ξ^n and has compact level sets, but is **nonconvex and nondifferentiable** in general. Hence, optimal scenarios exist, but their computation is difficult.

If P itself is discrete with possibly many (say $N \gg n$) scenarios and the minimization is restricted to $\Xi = \text{supp}(P)$ one arrives at the **optimal scenario reduction** problem. This problem can be shown to **decompose** into finding the optimal scenario set J to remain and into determining the optimal new probabilities given J . The background is that the Fortet-Mourier metric is of a special form if both probability measures are discrete.

Let P and Q be two discrete distributions, where ξ^i are the scenarios with probabilities p_i , $i = 1, \dots, N$, of P and $\tilde{\xi}^j$ the scenarios and q_j , $j = 1, \dots, n$, the probabilities of Q . Let Ξ denote the union of both scenario sets. Then

$$\begin{aligned}
 \zeta_r(P, Q) &= \inf \left\{ \int_{\Xi \times \Xi} \hat{c}_r(\xi, \tilde{\xi}) \eta(d\xi, d\tilde{\xi}) : \pi_1 \eta = P, \pi_2 \eta = Q \right\} \\
 &= \inf \left\{ \sum_{i=1}^N \sum_{j=1}^n \eta_{ij} \hat{c}_r(\xi_i, \tilde{\xi}_j) : \sum_{j=1}^n \eta_{ij} = p_i, \sum_{i=1}^N \eta_{ij} = q_j, \eta_{ij} \geq 0, \right. \\
 &\quad \left. i = 1, \dots, N, j = 1, \dots, n \right\} \\
 &= \sup \left\{ \sum_{i=1}^N p_i u_i - \sum_{j=1}^n q_j v_j : p_i - q_j \leq \hat{c}_r(\xi_i, \tilde{\xi}_j), i = 1, \dots, N, \right. \\
 &\quad \left. j = 1, \dots, n \right\}
 \end{aligned}$$

These two formulas represent **primal and dual representations of $\zeta_r(P, Q)$** and **primal and dual linear programs**.

Now, let P and Q be two discrete distributions, where ξ^i are the scenarios with probabilities p_i , $i = 1, \dots, N$, of P and ξ^j , $j \in J$, the scenarios and q_j , $j \in J$, the probabilities of Q . Let Ξ denote the support of P .

The **best approximation of P with respect to ζ_r** by such a distribution Q exists and is denoted by Q^* . It has the distance

$$D_J := \zeta_r(P, Q^*) = \min_{Q \in \mathcal{P}_n(\Xi)} \zeta_r(P, Q) = \sum_{i \notin J} p_i \min_{j \in J} \hat{c}_r(\xi^i, \xi^j)$$

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

and $j(i) \in \arg \min_{j \in J} \hat{c}_r(\xi^i, \xi^j)$, $\forall i \notin J$ (**optimal redistribution**).

(Dupačová–Gröwe-Kuska–Römisch 03)

Determining the **optimal scenario set J** with prescribed cardinality n is, however, a **combinatorial optimization problem: (metric n -median problem)**

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

The problem of finding the optimal set J of remaining scenarios is known to be **\mathcal{NP} -hard** (Kariv-Hakimi 79) and **polynomial time algorithms are not available**.

Reformulation of the (metric) n -median problem as combinatorial program:

$$\begin{aligned} \min \quad & \sum_{i,j=1}^N p_i x_{ij} \hat{c}_r(\xi^i, \xi^j) \quad \text{subject to} \\ \sum_{i=1}^N x_{ij} &= 1 \quad (j = 1, \dots, N), \quad \sum_{i=1}^N y_i \leq n, \\ x_{ij} &\leq y_i, \quad x_{ij} \in \{0, 1\} \quad (i, j = 1, \dots, N), \\ y_i &\in \{0, 1\} \quad (i = 1, \dots, N). \end{aligned}$$

The variable y_i decides whether scenario ξ^i remains and x_{ij} indicates whether scenario ξ^j minimizes the \hat{c}_r -distance to ξ^i .

The combinatorial program can, of course, be solved by standard software. However, meanwhile there is a well developed theory of polynomial-time **approximation algorithms** for solving it.. The current best approximation algorithm provides always an approximation guarantee of $1 + \sqrt{3} + \varepsilon$ (Li-Svensson 16).

The simplest algorithms are **greedy heuristics**, namely, backward (or reverse) and forward heuristics.

Starting point ($n = N - 1$): $\min_{l \in \{1, \dots, N\}} p_l \min_{j \neq l} \hat{c}_r(\xi_l, \xi_j)$

Algorithm: (Backward reduction)

Step [0]: $J^{[0]} := \emptyset$.

Step [i]: $l_i \in \arg \min_{l \notin J^{[i-1]}} \sum_{k \in J^{[i-1]} \cup \{l\}} p_k \min_{j \notin J^{[i-1]} \cup \{l\}} \hat{c}_r(\xi_k, \xi_j)$.
 $J^{[i]} := J^{[i-1]} \cup \{l_i\}$.

Step [N-n+1]: Optimal redistribution.

Starting point ($n = 1$): $\min_{u \in \{1, \dots, N\}} \sum_{k=1}^N p_k \hat{c}_r(\xi_k, \xi_u)$

Algorithm: (Forward selection)

Step [0]: $J^{[0]} := \{1, \dots, N\}$.

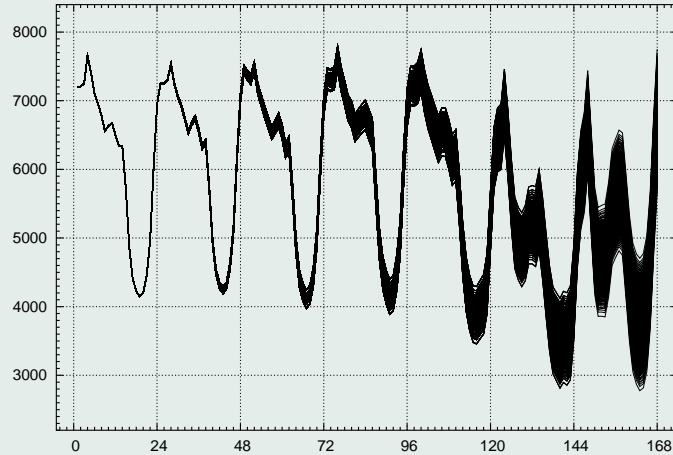
Step [i]: $u_i \in \arg \min_{u \in J^{[i-1]}} \sum_{k \in J^{[i-1]} \setminus \{u\}} p_k \min_{j \in J^{[i-1]} \setminus \{u\}} \hat{c}_r(\xi_k, \xi_j)$,
 $J^{[i]} := J^{[i-1]} \setminus \{u_i\}$.

Step [n+1]: Optimal redistribution.

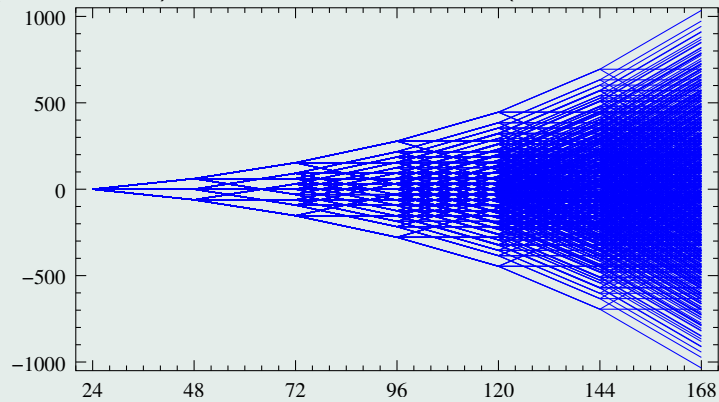
Although the approximation ratio of forward selection is known to be unbounded (Rujeerapaiboon-Schindler-Kuhn-Wiesemann 17), it worked well in many practical instances.

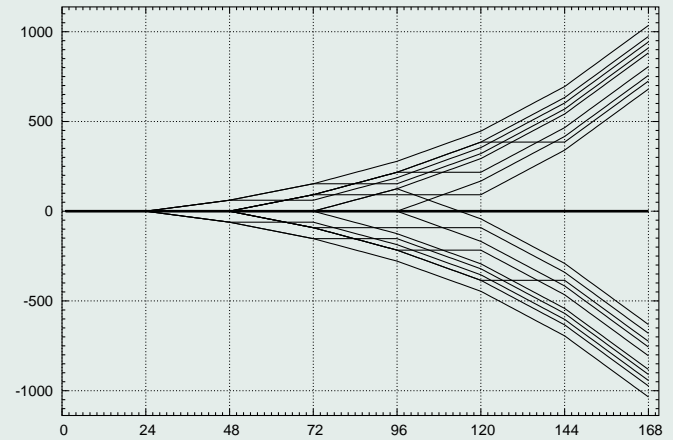
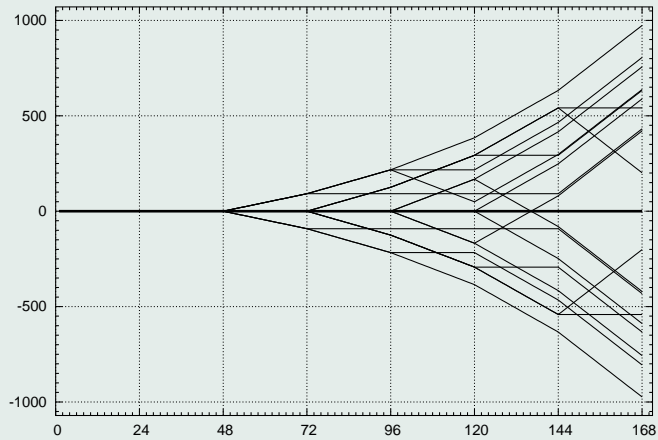
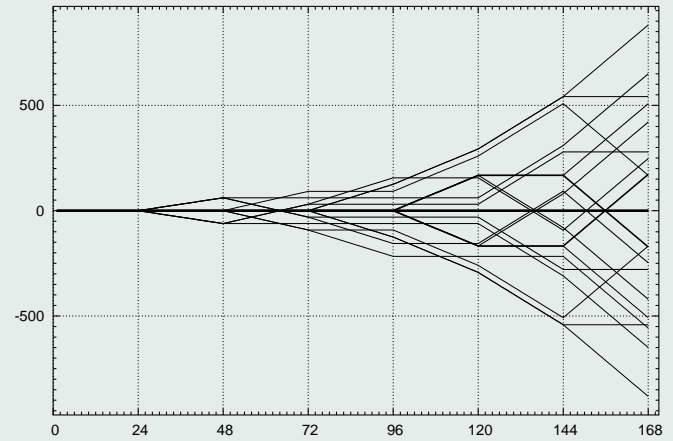
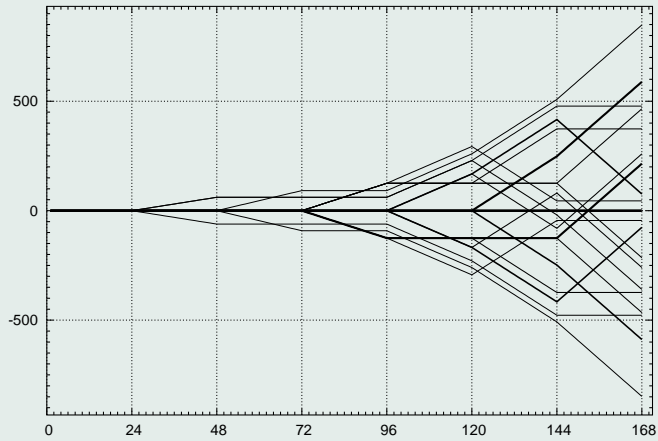
Example: (Weekly electrical load scenario tree)

Ternary load scenario tree (N=729 scenarios)



(Mean shifted) Ternary load scenario tree (N=729 scenarios)





Reduced load scenario trees obtained by forward selection with respect to the Fortet-Mourier distances ζ_r , $r = 1, 2, 4, 7$ and $n = 20$ (starting above left) (Heitsch-Römisch 07)

Optimal scenario generation for linear two-stage models

We consider linear two-stage stochastic programs as introduced earlier and impose the following conditions:

(A0) X is a bounded polyhedron and Ξ is convex polyhedral.

(A1) $h(x, \xi) \in W(Y)$ and $q(\xi) \in D$ are satisfied for every pair $(x, \xi) \in X \times \Xi$,

(A2) P has a second order absolute moment.

Then the infima $v(P)$ and $v(P_n)$ are attained and the estimate

$$\begin{aligned} |v(P) - v(P_n)| &\leq \sup_{x \in X} \left| \int_{\Xi} f_0(x, \xi) P(d\xi) - \int_{\Xi} f_0(x, \xi) P_n(d\xi) \right| \\ &= \sup_{x \in X} \left| \int_{\Xi} \Phi(q(\xi), h(x, \xi)) P(d\xi) - \int_{\Xi} \Phi(q(\xi), h(x, \xi)) P_n(d\xi) \right| \end{aligned}$$

holds due to the stability result for every $P_n \in \mathcal{P}_n(\Xi)$.

Hence, the **optimal scenario generation problem (OSG)** with uniform weights may be reformulated as: Determine $P_n^* \in \mathcal{P}_n(\Xi)$ such that it solves the **best uniform approximation problem**

$$\min_{(\xi^1, \dots, \xi^n) \in \Xi^n} \sup_{x \in X} \left| \int_{\Xi} \Phi(q(\xi), h(x, \xi)) P(d\xi) - \frac{1}{n} \sum_{i=1}^n \Phi(q(\xi^i), h(x, \xi^i)) \right|.$$

The class of functions $\{\Phi(q(\cdot), h(x, \cdot)) : x \in X\}$ from Ξ to $\overline{\mathbb{R}}$ enjoys specific properties. All functions are finite, continuous and piecewise linear-quadratic on Ξ . They are linear-quadratic on each convex polyhedral set

$$\Xi_j(x) = \{\xi \in \Xi : (q(\xi), h(x, \xi)) \in \mathcal{K}_j\} \quad (j = 1, \dots, \ell),$$

where the convex polyhedral cones \mathcal{K}_j , $j = 1, \dots, \ell$, represent a decomposition of the domain of Φ , which is itself a convex polyhedral cone in $\mathbb{R}^{\bar{m}+r}$.

Theorem: (Henrion-Römisch 17)

Assume (A0)–(A2). Then (OSG) is equivalent to the generalized semi-infinite program

$$\min_{t \geq 0, (\xi^1, \dots, \xi^n) \in \Xi^n} \left\{ t \left| \begin{array}{l} \frac{1}{n} \sum_{i=1}^n \langle h(x, \xi^i), z_i \rangle \leq t + F_P(x) \\ F_P(x) \leq t + \frac{1}{n} \sum_{i=1}^n \langle q(\xi^i), y_i \rangle \\ \forall (x, y, z) \in \mathcal{M}(\xi^1, \dots, \xi^n) \end{array} \right. \right\},$$

where the set $\mathcal{M} = \mathcal{M}(\xi^1, \dots, \xi^n)$ and the function $F_P : X \rightarrow \mathbb{R}$ are given by

$$\mathcal{M} = \{(x, y, z) \in X \times Y^n \times \mathbb{R}^{rn} : W y_i = h(x, \xi^i), W^\top z_i - q(\xi^i) \in Y^*, \forall i\},$$

$$F_P(x) := \int_{\Xi} \Phi(q(\xi), h(x, \xi)) P(d\xi).$$

The latter is the convex expected recourse function of the two-stage model.

Theorem: Let the function h be affine, assume (A0)–(A2) and that either h or q be random. Then the set-valued mapping $\mathcal{M} : \Xi^n \rightrightarrows \mathbb{R}^m \times \mathbb{R}^{\bar{m}n} \times \mathbb{R}^{rn}$ has convex polyhedral graph and is Hausdorff Lipschitz continuous on Ξ^n .

The feasible set M is closed and convex.

We note that $F_P(x)$ can only be calculated approximately even if the probability measure P is completely known. For example, this could be done by Monte Carlo or Quasi-Monte Carlo methods with a large sample size $N > n$. Let

$$F_P(x) \approx \frac{1}{N} \sum_{j=1}^N \Phi(q(\hat{\xi}^j), h(x, \hat{\xi}^j))$$

be such an approximate representation of $F_P(x)$ based on a sample $\hat{\xi}^j$, $j = 1, \dots, N$. The corresponding generalized semi-infinite program is of the form

$$\min_{t \geq 0, (\xi^1, \dots, \xi^n) \in \Xi^n} \left\{ t \left| \begin{array}{l} \frac{1}{n} \sum_{i=1}^n \langle h(x, \xi^i), z_i \rangle \leq t + \frac{1}{N} \sum_{j=1}^N \langle q(\hat{\xi}^j), \hat{y}_j \rangle \\ \frac{1}{N} \sum_{j=1}^N \langle h(x, \hat{\xi}^j), \hat{z}_j \rangle \leq t + \frac{1}{n} \sum_{i=1}^n \langle q(\xi^i), y_i \rangle \\ \forall (x, \hat{y}, \hat{z}) \in \mathcal{M}(\hat{\xi}^1, \dots, \hat{\xi}^N) \\ \forall (x, y, z) \in \mathcal{M}(\xi^1, \dots, \xi^n) \end{array} \right. \right\}.$$

Monte Carlo and Quasi-Monte Carlo methods

Monte Carlo: Let $\xi^i(\cdot)$, $i \in \mathbb{N}$, denote independent and identically distributed random vectors with common distribution P and P_n be the empirical measure

$$P_n(\cdot) = \frac{1}{n} \sum_{i=1}^n \delta_{\xi^i(\cdot)} \quad (n \in \mathbb{N})$$

defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The law of large numbers implies that the sequence $(P_n(\cdot))_{n \in \mathbb{N}}$ converges \mathbb{P} -almost surely weakly to P .

To study the convergence rate one considers the **empirical process**

$$\{\beta_n(P_n(\cdot) - P)f\}_{f \in \mathcal{F}} \quad (n \in \mathbb{N})$$

indexed by a function class \mathcal{F} with sequence (β_n) , where $Qf = \int_{\Xi} f(\xi)Q(d\xi)$ for any Borel probability measure Q on Ξ . The empirical process is called **bounded in probability with tail function** $\tau_{\mathcal{F}}$ if for all $\varepsilon > 0$ and $n \in \mathbb{N}$ the estimate

$$\mathbb{P}(\{\beta_n d_{\mathcal{F}}(P_n(\cdot), P) \geq \varepsilon\}) \leq \tau_{\mathcal{F}}(\varepsilon)$$

holds. Whether the empirical process is bounded in probability, depends on the size of the class \mathcal{F} measured in terms of covering numbers in $L_2(\Xi, P)$. Typically, one has an **exponential tail** $\tau_{\mathcal{F}}(\varepsilon) = C(\varepsilon) \exp(-\varepsilon^2)$ and $\beta_n = \sqrt{n}$.

Quasi-Monte Carlo: The basic idea of Quasi-Monte Carlo (QMC) methods is to use **deterministic points that are (in some way) uniformly distributed in $[0, 1]^s$** and to consider first the approximate computation of

$$I_s(f) = \int_{[0,1]^s} f(\xi) d\xi \quad \text{by} \quad Q_{n,s}(f) = \frac{1}{n} \sum_{i=1}^n f(\xi^i)$$

with (non-random) points ξ^i , $i = 1, \dots, n$, from $[0, 1]^s$.

The uniform distribution property of point sets may be defined in terms of the so-called **L_p -discrepancy of ξ^1, \dots, ξ^n** for $1 \leq p \leq \infty$

$$d_{p,n}(\xi^1, \dots, \xi^n) = \left(\int_{[0,1]^s} |\text{disc}(\xi)|^p d\xi \right)^{\frac{1}{p}}, \quad \text{disc}(\xi) := \prod_{j=1}^d \xi_j - \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{[0,\xi]}(\xi^i).$$

A sequence $(\xi^i)_{i \in \mathbb{N}}$ is called **uniformly distributed in $[0, 1]^s$** if

$$d_{p,n}(\xi^1, \dots, \xi^n) \rightarrow 0 \quad \text{for} \quad n \rightarrow \infty$$

There exist sequences (ξ^i) in $[0, 1]^s$ such that for all $\delta \in (0, \frac{1}{2}]$

$$d_{\infty,n}(\xi^1, \dots, \xi^n) = O(n^{-1}(\log n)^s) \quad \text{or} \quad d_{\infty,n}(\xi^1, \dots, \xi^n) \leq C(s, \delta)n^{-1+\delta}.$$

Using a suitable randomization of such sequences may lead to a root mean square convergence rate $\sqrt{\mathbb{E}[d_{2,n}^2(\xi^1, \dots, \xi^n)]} \leq C(\delta)n^{-1+\delta}$ with a constant $C(\delta)$ not depending on the dimension s and $\delta \in (0, \frac{1}{2}]$.

Example: Randomly shifted lattice rule (Sloan-Kuo-Joe 02).

With a random vector Δ which is uniformly distributed on $[0, 1]^s$, we consider the randomly shifted lattice rule

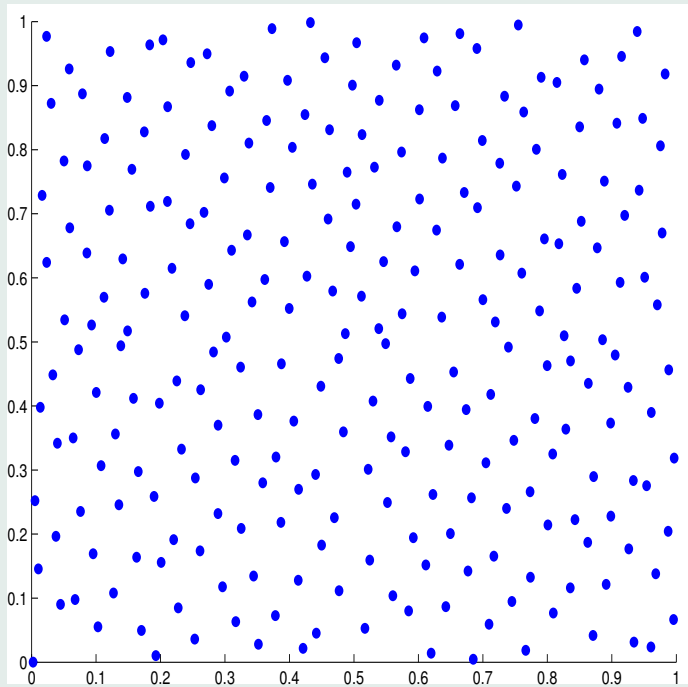
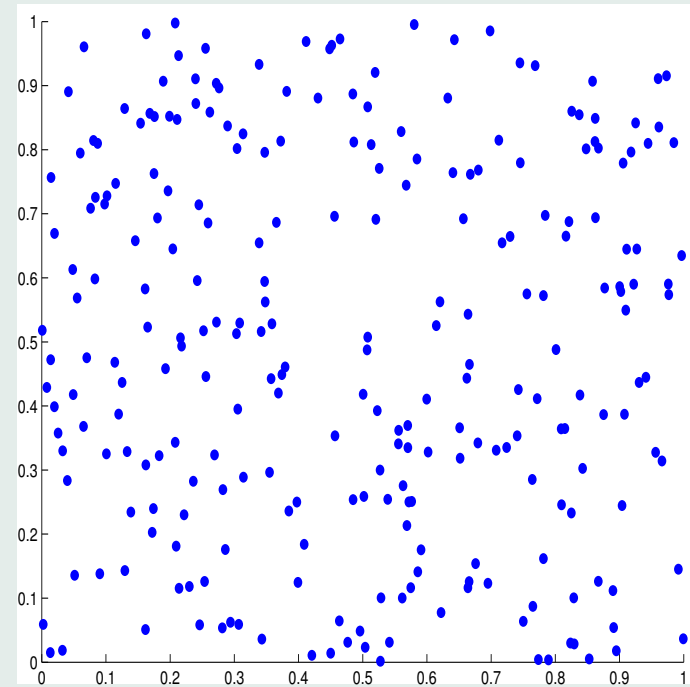
$$Q_{n,s}(\omega)(f) = \frac{1}{n} \sum_{j=1}^n f\left(\left\{\frac{(j-1)}{n}\mathbf{g} + \Delta(\omega)\right\}\right).$$

Theorem: Let $n \in \mathbb{N}$ be prime and f belong to the weighted tensor product Sobolev space $\mathcal{W}_{2,\gamma,\text{mix}}^{(1,\dots,1)}([0, 1]^s)$. Then $\mathbf{g} \in \mathbb{Z}_+^s$ can be constructed componentwise such that for each $\delta \in (0, \frac{1}{2}]$ there exists a constant $C(\delta) > 0$ with

$$\sup_{\|f\|_\gamma \leq 1} \sqrt{\mathbb{E}|Q_{n,s}(\omega)(f) - I_s(f)|^2} \leq C(\delta) n^{-1+\delta},$$

where $C(\delta)$ increases if δ decreases, but does not depend on s if the sequence (γ_j) of coordinate weights satisfies $\sum_{j=1}^{\infty} \gamma_j^{\frac{2}{2(1-\delta)}} < \infty$ (e.g. $\gamma_j = \frac{1}{j^3}$).

Note that piecewise polynomial functions f do almost belong to $\mathcal{W}_{2,\gamma,\text{mix}}^{(1,\dots,1)}([0, 1]^s)$ if its effective dimension is small (Heitsch-Leövey-Römisch 16).



Comparison of $n = 2^7$ Monte Carlo Mersenne Twister points and randomly binary shifted Sobol' points in dimension $s = 500$, projection (8,9)

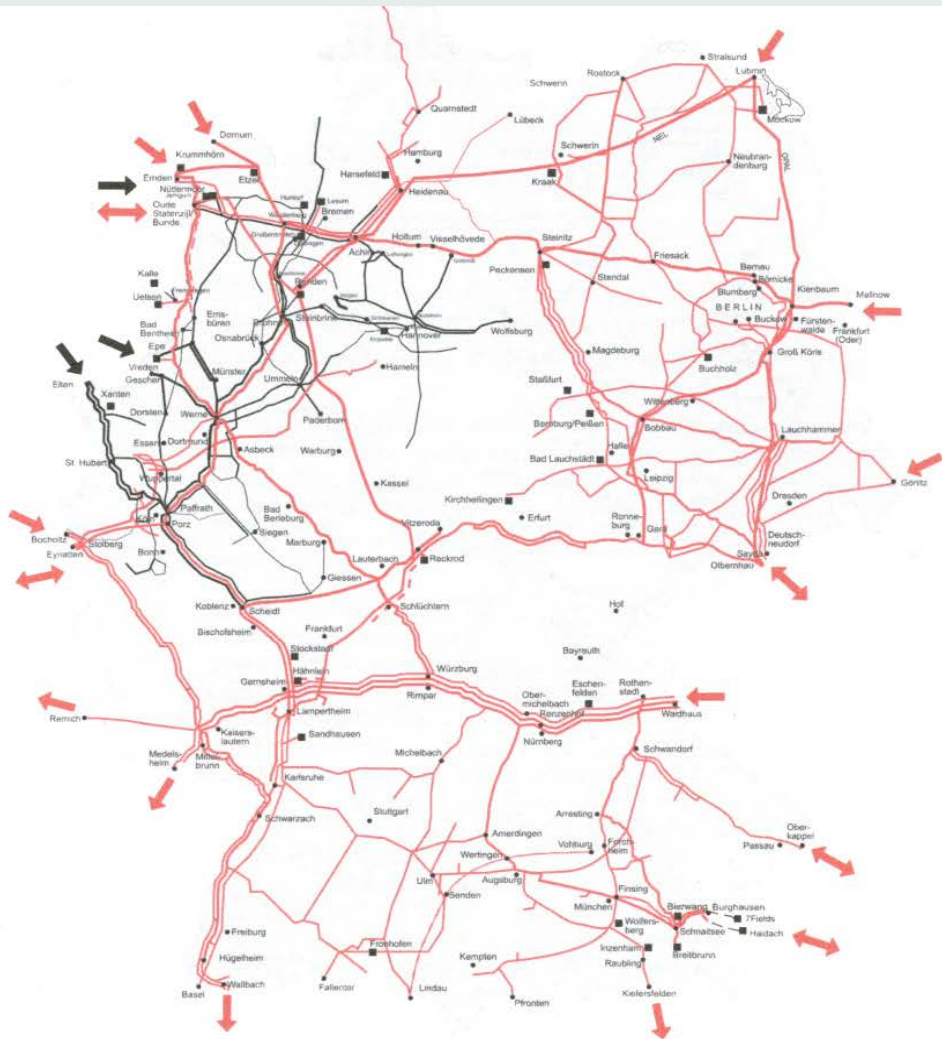
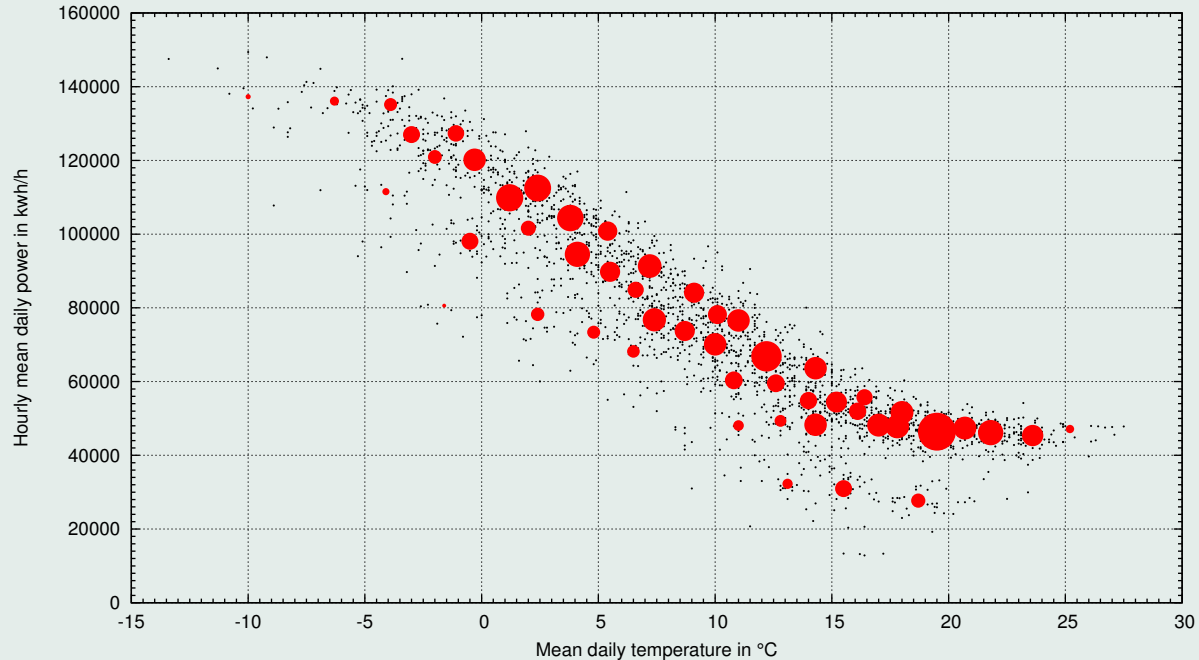


Figure 1.4. German H-gas (red) and L-gas (black) network systems. The arrows indicate entry and exit nodes. Gas storages are represented by black squares. (Source: OGE.)

Illustration:

$N = 2340$ samples based on randomized Sobol' points are generated for several hundred exits and later reduced by scenario reduction to $n = 50$ scenarios. The result is shown below for a specific exit where the diameters of the red balls are proportional to the new probabilities.



(Chapters 13 and 14 in Koch-Hiller-Pfetsch-Schewe 15)

References

- J. Dick, F. Y. Kuo, I. H. Sloan: High-dimensional integration – the Quasi-Monte Carlo way, *Acta Numerica* 22 (2013), 133–288.
- J. Dupačová, N. Gröwe-Kuska, W. Römisch: Scenario reduction in stochastic programming: An approach using probability metrics, *Mathematical Programming* 95 (2003), 493–511.
- H. Heitsch, H. Leövey, W. Römisch: Are Quasi-Monte Carlo algorithms efficient for two-stage stochastic programs?, *Computational Optimization and Applications* 65 (2016), 567–603.
- H. Heitsch, W. Römisch: A note on scenario reduction for two-stage stochastic programs, *Operations Research Letters* 35 (2007), 731–738.
- R. Henrion, W. Römisch: Optimal scenario generation and reduction in stochastic programming, *Stochastic Programming E-Print Series* 2 (2017) and submitted.
- O. Kariv, S. L. Hakimi: An algorithmic approach to network location problems, II: The p -medians, *SIAM Journal Applied Mathematics* 37 (1979), 539–560.
- T. Koch, B. Hiller, M. E. Pfetsch, L. Schewe (Eds.): Evaluating Gas Network Capacities, SIAM-MOS Series on Optimization, Philadelphia, 2015.
- S. Li, O. Svensson: Approximating k -median via pseudo-approximation, *SIAM Journal on Computing* 45 (2016), 530–547.
- S. T. Rachev, W. Römisch: Quantitative stability in stochastic programming: The method of probability metrics, *Mathematics of Operations Research* 27 (2002), 792–818.
- S. T. Rachev, L. Rüschendorf: *Mass Transportation Problems*, Vol. I, Springer, Berlin 1998.
- N. Rujeerapaiboon, K. Schindler, D. Kuhn, W. Wiesemann: Scenario reduction revisited: Fundamental limits and guarantees, *Stochastic Programming E-Print Series* 1 (2017) and submitted.