Asymptotic properties of Monte Carlo methods for PDE constrained optimization under uncertainty

W. Römisch and T. M. Surowiec[+]

Humboldt-University Berlin
Institute of Mathematics
www.math.hu-berlin.de/~romisch

[+]Simula Research Laboratory
Oslo, Norway

Congratulations René

Workshop *Frontiers of Stochastic Optimization and its Applications in Industry*
WIAS Berlin, May 10–12, 2023

Introduction

PDE constrained optimization problems under uncertainty attracted more interest since the Oberwolfach meeting in 2013. Such problems are motivated by economic and engineering applications.

Their solution requires a combination of discretization and sampling techniques, and of specific iterative solution methods.

More specific PDEs with random coefficients are within the reach of efficient computational methods.

Here, we study the impact of Monte Carlo methods in the presence of infinite dimensions and present some numerical results.

PDE constrained optimization under uncertainty

Let $D \subset \mathbb{R}^m$ be an open bounded domain with Lipschitz boundary, $V = H_0^1(D)$ the classical Sobolev space with inner product $(\cdot,\cdot)_V$, $V^\star = H^{-1}(D)$ its dual with norm $\|\cdot\|_\star$ und dual pairing $\langle\cdot,\cdot\rangle$ and $H = L^2(D)$ with inner product $(\cdot,\cdot)_H$. Let Ξ be a metric space and $\mathbb{P}$ be a Borel probability measure on Ξ.
We consider the bilinear form $a(\cdot,\cdot;\xi) : V \times V \to \mathbb{R}$ defined by

$$a(u,v;\xi) = \int_D \sum_{i,j=1}^n b_{ij}(x,\xi) \frac{\partial u(x)}{\partial x_i} \frac{\partial v(x)}{\partial x_j} \mathrm{d}x \quad (\xi \in \Xi).$$

We impose the condition that the functions $b_{ij} : D \times \Xi \to \mathbb{R}$ are measurable on $D \times \Xi$ and there exist $L > \gamma > 0$ such that

$$\gamma \sum_{i=1}^n y_i^2 \le \sum_{i,j=1}^n b_{ij}(x,\xi) y_i y_j \le L \sum_{i=1}^n y_i^2 \quad (\forall y \in \mathbb{R}^n)$$

for a.e. $x \in D$ and $\mathbb{P}$-a.e. $\xi \in \Xi$. This implies that each b_{ij} is essentially bounded on $D \times \Xi$ from both sides with respect to the associated product measure.

For $\mathbb{P}$-a.e. $\xi \in \Xi$ we define the mapping $A(\xi) : V \to V^\star$ by means of the Riesz representation theorem

$$\langle A(\xi)u, v\rangle = a(u, v; \xi) \quad (u, v \in V).$$

Then $A(\xi)$ is a linear, uniformly positive definite (with $\gamma > 0$) and uniformly bounded (with $L > 0$) random elliptic operator.

PDE constrained stochastic optimization problem:
Minimize the functional

$$\begin{aligned} F(z) \;&=\; \frac{1}{2}\int_\Xi \int_D |u(x,\xi) - \widetilde{u}(x)|^2 \,\mathrm{d}x\mathrm{d}\mathbb{P}(\xi) + \frac{\alpha}{2}\int_D |z(x)|^2 \,\mathrm{d}x \\ &=\; \frac{1}{2}\mathbb{E}_\mathbb{P}[\|u - \widetilde{u}\|_H^2] + \frac{\alpha}{2}\|z\|_H^2 \end{aligned}$$

subject to $A(\xi)u = z + g(\xi)$ and $z \in Z_{\mathrm{ad}} \quad (\mathbb{P}\text{-a.e. } \xi \in \Xi)$,

where $Z_{\mathrm{ad}} \subset H$ denotes a closed convex bounded subset, $\alpha > 0$, $\widetilde{u} \in H$ and $g : D \times \Xi \to \mathbb{R}$ is measurable on $D \times \Xi$ and at least square integrable on D.

Existence and quadratic growth

Infinite-dimensional stochastic optimization problem:
We consider the integrand

$$f(z,\xi) = \frac{1}{2}\|A(\xi)^{-1}(z+g(\xi)) - \widetilde{u}\|_H^2 + \frac{\alpha}{2}\|z\|_H^2$$

for any $z \in Z_{\mathrm{ad}}$ and $\mathbb{P}$-a.e. $\xi \in \Xi$, and the optimization problem

$$\min\Big\{F(z) = \int_\Xi f(z,\xi)\mathrm{d}\mathbb{P}(\xi) : z \in Z_{\mathrm{ad}}\Big\},$$

where $g \in L_2(\Xi,\mathbb{P};V^\star)$ and $A(\xi)$ as defined earlier.

Proposition:
The functional F is finite, strongly convex and lower semicontinuous, hence, weakly lower semicontinuous on the weakly compact set Z_{ad}. Hence, there exists a unique minimizer $z(\mathbb{P}) \in Z_{\mathrm{ad}}$ of F and it holds

$$\|z - z(\mathbb{P})\|^2 \leq \frac{8}{\alpha}(F(z) - F(z(\mathbb{P}))) \quad (\forall z \in Z_{\mathrm{ad}}).$$

Weak convergence and metric distances

Weak convergence in $\mathcal{P}(\Xi)$: $(\mathbb{P}_N)$ converges weakly to $\mathbb{P}$ iff

$$\lim_{N\to\infty}\int_\Xi f(\xi)\mathrm{d}\mathbb{P}_N(\xi)=\int_\Xi f(\xi)\mathrm{d}\mathbb{P}(\xi)\qquad (\forall f\in C_b(\Xi,\mathbb{R})).$$

The topology of weak convergence is metrizable if Ξ is separable.
A sequence of random variables converges in distribution if their probability distributions converge weakly.

Distance on $\mathcal{P}(\Xi)$: (Zolotarev 83)

$$d_{\mathfrak{F}}(\mathbb{P},\mathbb{Q})=\sup_{f\in\mathfrak{F}}\left|\int_\Xi f(\xi)\mathrm{d}\mathbb{P}(\xi)-\int_\Xi f(\xi)\mathrm{d}\mathbb{Q}(\xi)\right|,$$

where $\mathfrak{F}$ is a family of real-valued Borel measurable functions on Ξ.

A number of important probability metrics are of the form $d_{\mathfrak{F}}$, for example, the bounded Lipschitz metric (metrizing the topology of weak convergence) and the Fortet-Mourier type metrics.

Whether convergence with respect to $d_{\mathfrak{F}}$ implies or is implied by weak convergence depends on the richness and on analytical properties of $\mathfrak{F}$.

Quantitative stability

Compared with classical probability metrics we consider much smaller families $\mathfrak{F}$:

$$\begin{aligned}\mathfrak{F}_{\mathrm{mi}} &= \{f(z,\cdot) : z \in Z_{\mathrm{ad}}\}\\ \mathfrak{F}_{\mathrm{di}} &= \{f'_z(z,\cdot)(h) : z \in Z_{\mathrm{ad}}, \|h\|_H \le 1\}.\end{aligned}$$

In this case we arrive at semi-metrics which we call minimal information (m.i.) and minimal (Fréchet) derivative information (d.i.) distances, respectively.

Theorem:

Under the standing assumptions we obtain the following estimates for the infimum $v(\mathbb{P})$ and the minimizer $z(\mathbb{P})$ of F with respect to Z_{ad}:

$$\begin{aligned}|v(\mathbb{Q}) - v(\mathbb{P})| &\le d_{\mathfrak{F}_{\mathrm{mi}}}(\mathbb{Q},\mathbb{P})\\ \|z(\mathbb{Q}) - z(\mathbb{P})\|_H &\le 2\sqrt{\frac{2}{\alpha}} d_{\mathfrak{F}_{\mathrm{mi}}}(\mathbb{Q},\mathbb{P})^{\frac{1}{2}}\\ \|z(\mathbb{Q}) - z(\mathbb{P})\|_H &\le \frac{8}{\alpha} d_{\mathfrak{F}_{\mathrm{di}}}(\mathbb{Q},\mathbb{P})\end{aligned}$$

for any $\mathbb{Q} \in \mathcal{P}(\Xi)$.

Properties of the integrands

Lemma: (Topsøe 67)
Let $\mathfrak{F}$ be uniformly bounded and $\mathbb{P}(\{\xi \in \Xi : \mathfrak{F}$ is not equicontinuous at $\xi\}) = 0$ holds. Then $\mathfrak{F}$ is a $\mathbb{P}$-uniformity class, i.e., weak convergence of $(\mathbb{P}_N)$ to $\mathbb{P}$ implies

$$\lim_{N\to\infty} d_{\mathfrak{F}}(\mathbb{P}_N, \mathbb{P}) = 0.$$

Theorem:
Assume that all functions $b_{ij}(x,\cdot)$, $i,j = 1,\ldots,n$, and $g(x,\cdot)$ are Lipschitz continuous on Ξ uniformly with respect to $x \in D$. Furthermore, let $g \in L_\infty(\Xi, \mathbb{P}; V^\star)$. Then the families $\mathfrak{F}_{\mathrm{mi}}$ and $\mathfrak{F}_{\mathrm{di}}$ are uniformly bounded and equi-Lipschitz continuous on Ξ and, hence, $\mathbb{P}$-uniformity classes.

Let $\Xi \subset \mathbb{R}^d$ be a bounded, convex set having the property $\Xi \subseteq \mathrm{cl}\,\mathrm{int}\,\Xi$, let $k \in \mathbb{N}$. Assume that all functions $b_{ij}(x,\cdot)$, $i,j = 1,\ldots,n$, $g(x,\cdot)$, $x \in D$, have continuous mixed partial derivatives up to order k which are in addition all measurable and essentially bounded on $D \times \Xi$.
Then both classes $\mathfrak{F}_{\mathrm{mi}}$ and $\mathfrak{F}_{\mathrm{di}}$ are bounded subsets of $C^k(\Xi)$.

Monte Carlo approximations

Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be independent identically distributed Ξ-valued random variables on some probability space $(\Omega, \mathcal{F}, P)$ having the common distribution $\mathbb{P}$, i.e., $\mathbb{P} = P\xi_1^{-1}$. We consider the empirical measures

$$\mathbb{P}_n(\cdot) = \frac{1}{n} \sum_{i=1}^{n} \delta_{\xi_i(\cdot)} \qquad (n \in \mathbb{N})$$

and the empirical or Monte Carlo approximation of the original stochastic program with sample size n, namely, the program

$$\min \left\{ \frac{1}{n} \sum_{i=1}^{n} f(z, \xi_i(\cdot)) : z \in Z_{\mathrm{ad}} \right\}.$$

Its optimal value $v(\mathbb{P}_n(\cdot))$ is a real random variable and its solution $z(\mathbb{P}_n(\cdot))$ an H-valued random element.

Qualitative and quantitative results on the asymptotic behavior of optimal values and solutions are known in finite-dimensional settings so far.
(Dupačová-Wets 88, and the surveys by Shapiro 03 and Pflug 03)

It is well known that $(\mathbb{P}_n(\cdot))$ converges weakly to $\mathbb{P}$ P-almost surely.

Corollary: The sequences $(v(\mathbb{P}_n(\cdot)))$ and $(z(\mathbb{P}_n(\cdot)))$ of empirical optimal values and solutions converge P-almost surely to the true optimal values and solutions $v(\mathbb{P})$ and $z(\mathbb{P})$, respectively.

Quantitative information on the asymptotic behavior of $v(P_n(\cdot))$ and $z(P_n(\cdot))$ is closely related to uniform convergence properties of the empirical process

$$\left\{\mathbb{G}_n f = \sqrt{n}(\mathbb{P}_n(\cdot) - \mathbb{P})f = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(f(\xi_i(\cdot)) - \mathbb{P}f)\right\}_{f\in\mathfrak{F}}$$

indexed by $\mathfrak{F} = \{f(z,\cdot) : z \in Z_{\mathrm{ad}}\}$, where we set $\mathbb{Q}f = \int_\Xi f(\xi)\mathrm{d}\mathbb{Q}(\xi)$ for any probability distribution $\mathbb{Q}$ and any $f \in \mathfrak{F}$.

We are interested in obtaining quantitative information on the mean or probability estimates of

$$\sqrt{n}\, d_{\mathfrak{F}}(\mathbb{P}_n(\cdot), \mathbb{P}) = \sqrt{n} \sup_{f\in\mathfrak{F}} |\mathbb{P}_n(\cdot)f - \mathbb{P}f|$$

and in a (functional) central limit theorem for $\{\mathbb{G}_n f\}_{f\in\mathfrak{F}}$.

The supremum is non-measurable in general, but since Z is a subset of a separable Hilbert space and all functions in $\mathfrak{F}$ are continuous, the supremum may be restricted to a countable set of functions and is, hence, measurable.

A collection $\mathfrak{F}$ of measurable functions on Ξ is called $\mathbb{P}$-Donsker if the empirical process $\{\mathbb{G}_n f\}_{f\in\mathfrak{F}}$ converges in distribution to a tight random variable $\mathbb{G}$

$$\mathbb{G}_n = \sqrt{n}(\mathbb{P}_n - \mathbb{P}) \xrightarrow{d} \mathbb{G}$$

in the space $\ell^\infty(\mathfrak{F})$ of bounded functions on $\mathfrak{F}$, where the limit $\mathbb{G} = \{\mathbb{G}f : f \in \mathfrak{F}\}$ is a Gaussian process with zero mean and covariance function

$$E_P[\mathbb{G}f_1\mathbb{G}f_2] = \mathbb{P}[(f_1 - \mathbb{P}f_1)(f_2 - \mathbb{P}f_2)] \quad (f_1, f_2 \in \mathfrak{F}).$$

The limit $\mathbb{G}$ is sometimes called a $\mathbb{P}$-Brownian bridge process in $\ell^\infty(\mathfrak{F})$.

Whether such a functional central limit theorem or rates for convergence in mean or probability are valid, depends on the size of the class $\mathfrak{F}$ measured in terms of bracketing and entropy numbers. To introduce the latter concepts, let $\mathfrak{F}$ be a subset of the linear normed space $L_p(\Xi, \mathbb{P})$ (for some $p \geq 1$) equipped with the usual norm

$$\|f\|_{\mathbb{P},p} = (\mathbb{P}|f|^p)^{\frac{1}{p}} = \left(\int_\Xi |f(\xi)|^p \mathrm{d}\mathbb{P}(\xi)\right)^{\frac{1}{p}}.$$

The bracketing number $N_{[\,]}(\varepsilon, \mathfrak{F}, \|\cdot\|_{\mathbb{P},p})$ is the minimal number of brackets

$$[l, u] = \{f \in L_p(\Xi, \mathbb{P}) : l \le f \le u\}$$

needed to cover $\mathfrak{F}$, where $l, u \in L_p(\Xi, \mathbb{P})$ and $\|l - u\|_{\mathbb{P},p} < \varepsilon$.
The metric entropy number $H_{[\,]}(\varepsilon, \mathfrak{F}, \|\cdot\|_{\mathbb{P},p})$ with bracketing for $\mathfrak{F}$ is defined by

$$H_{[\,]}(\varepsilon, \mathfrak{F}, \|\cdot\|_{\mathbb{P},p}) = \log N_{[\,]}(\varepsilon, \mathfrak{F}, \|\cdot\|_{\mathbb{P},p}).$$

Both numbers are finite if $\mathfrak{F}$ is a totally bounded subset of $L_p(\Xi, \mathbb{P})$.

Proposition: (van der Vaart 96)
There exists a universal constant $C > 0$ such that for any class $\mathfrak{F}$ of measurable functions with envelope function $\hat{F}$ (i.e., $|f| \le \hat{F}$ for every $f \in \mathfrak{F}$) belonging to $L_2(\Xi, \mathbb{P})$ the estimate

$$\sqrt{n}\, \mathbb{E}[d_{\mathfrak{F}}(\mathbb{P}_n(\cdot), \mathbb{P})] \le C \int_0^1 \sqrt{1 + H_{[\,]}(\varepsilon \|\hat{F}\|_{\mathbb{P},2}, \mathfrak{F}, \|\cdot\|_{\mathbb{P},2})}\, d\varepsilon\, \|\hat{F}\|_{\mathbb{P},2}$$

holds. If the integral is finite, then the class $\mathfrak{F}$ is $\mathbb{P}$-Donsker.

Note that the integral can only be finite if $H_{[\,]}(\varepsilon, \mathfrak{F}, \|\cdot\|_{\mathbb{P},2})$ grows at most like $\varepsilon^{-\beta}$ with $0 < \beta < 2$ for $\varepsilon \to +0$.

Below it turns out that the integral is finite if $\mathfrak{F}$ is a bounded subset of classical linear normed spaces of smooth functions.

Let $\Xi \subset \mathbb{R}^d$ be convex, bounded with the property $\Xi \subseteq \mathrm{cl}\ \mathrm{int}\,\Xi$, $k \in \mathbb{N}_0$. We consider the classical Banach space $C^k(\Xi)$ of real functions on Ξ having continuous partial derivatives up to order k.

Proposition: (Kolmogorov/Tikhomirov 61)
Let $\mathbb{B}_k(\rho)$ denote the ball around the origin with radius ρ in $C^k(\Xi)$. Then there exists a constant $K > 0$ depending only on d, k, ρ and the diameter of Ξ such that we have for its metric entropy with bracketing

$$H_{[\,]}(\varepsilon\rho, \mathbb{B}_k(\rho), \|\cdot\|_{\mathbb{P},2}) \leq K\varepsilon^{-\frac{d}{k}}$$

for every $\varepsilon > 0$.

Corollary:
Bounded subsets $\mathfrak{F}$ of $C^k(\Xi)$ are $\mathbb{P}$-Donsker if $d < 2k$.

Delta theorem

Proposition: (Shapiro 91)
Let B_1 and B_2 be linear normed spaces equipped with their Borel σ-fields. Let (X_n) be random elements of B_1, and (τ_n) be positive and such that $\tau_n \to \infty$ as $n \to \infty$. Let

$$\tau_n(X_n - \bar{x}) \xrightarrow{d} X$$

for some $\bar{x} \in B_1$ and some random element X of B_1, and let $\Phi : B_1 \to B_2$ be Hadamard directionally differentiable at $\bar{x}$, i.e.,

$$\lim_{\substack{t \to 0+ \\ y \to \bar{y}}} \frac{\Phi(\bar{x} + ty) - \Phi(\bar{x})}{t} = \Phi'(\bar{x}; \bar{y})$$

for all $\bar{y} \in B_1$. Then it holds

$$\tau_n(\Phi(X_n) - \Phi(\bar{x})) \xrightarrow{d} \Phi'(\bar{x}; X),$$

where $\stackrel{d}{\to}$ means convergence in distribution.

(King 89, Dümbgen 93)

Application:

$B_1 = \ell^\infty(Z_{\mathrm{ad}})$, $B_2 = \mathbb{R}$ and $\Phi(h) = \inf_{z \in Z_{\mathrm{ad}}} h(z)$.

Proposition:

The infimum mapping Φ is Hadamard directionally differentiable at each $h \in \ell^\infty(Z_{\mathrm{ad}})$ and it holds

$$\Phi'(h;\theta) = \lim_{\varepsilon \to 0+} \inf \left\{ \theta(z) : z \in Z_{\mathrm{ad}}, h(z) \le \inf_{z \in Z_{\mathrm{ad}}} h(z) + \varepsilon \right\}.$$

If h and θ are weakly lower semicontinuous on Z_{ad}, then

$$\Phi'(h;\theta) = \min \left\{ \theta(z) : z \in \arg \min_{z \in Z_{\mathrm{ad}}} h(z) \right\}.$$

(Lachout 06, Eichhorn-Römisch 07)

Application:

$h(z) := F(z) = \mathbb{P} f(z, \cdot)$, $\theta(z) := \mathbb{G} f(z, \cdot)$ and $\Phi'(h, \theta) = \mathbb{G} f(z(\mathbb{P}), \cdot)$.

Main result

Theorem:
Let $\Xi \subset \mathbb{R}^d$ be a bounded, convex set having the property $\Xi \subseteq \mathrm{cl}\ \mathrm{int}\,\Xi$ and let $k \in \mathbb{N}$ be such that $d < 2k$. Assume that all functions $b_{ij}(x,\cdot) : \Xi \to \mathbb{R}$, $i,j = 1,\dots,m$, and $g(x,\cdot) : \Xi \to \mathbb{R}$, $x \in D$, have continuous partial derivatives up to order k which are all measurable and essentially bounded on $D \times \Xi$.
Then the classes $\mathfrak{F}_{\mathrm{mi}}$ and $\mathfrak{F}_{\mathrm{di}}$ are $\mathbb{P}$-Donsker and it holds that

$$\begin{aligned} \mathbb{E}[|v(\mathbb{P}_n(\cdot)) - v(\mathbb{P})|] &= O(n^{-\frac{1}{2}}) \\ \mathbb{E}[\|z(\mathbb{P}_n(\cdot)) - z(\mathbb{P})\|_H] &= O(n^{-\frac{1}{2}}) \end{aligned}$$

and the sequence $(\sqrt{n}(v(\mathbb{P}_n(\cdot)) - v(\mathbb{P})))$ converges in distribution to a normal random variable with mean zero and variance $\mathbb{E}[(\mathbb{G}f(z(\mathbb{P}),\cdot))^2]$.

Remark:
An extension to unbounded Ξ is possible if an additional condition is satisfied which connects the tail behaviour of $\mathbb{P}$ and bounds for the integrand on bounded subsets.

Illustrating example

We construct examples that are flexible in the number of random variables and smoothness in the coefficients.
Set $D = (0,1)^2$, $\Xi = [0,1]^d$ and $\mathbb{P}$ the uniform distribution, where $d = 2^q$, $q \in \mathbb{N}_0$.
Next, we partition the interval $[0,1]$ into d closed intervals

$$D_i = \left[\frac{i-1}{d}, \frac{i}{d}\right] \qquad (i = 1, \dots, d)$$

with $\mathbb{1}_{D_i}$ be the associated characteristic function.
Next we define the mapping $\hat{b} : D \times \mathbb{R}^d \to \mathbb{R}$ by

$$\hat{b}(x, \eta) = \sum_{i=1}^{d} (\eta_i + 10^{-2} x_2 + 10^{-3}) \mathbb{1}_{D_i}(x_1).$$

where $x \in D$ and $\eta \in \mathbb{R}^d$. We will use $\hat{b}$ to define the random coefficients inside the differential operator. For each $k = 0, 1, 2, \dots$ we define $w_k : \Xi \to \mathbb{R}^d$ by

$$(w_k(\xi))_j = (\xi_j - 10^{-1})^k \max\{0, \xi_j - 10^{-1}\} \qquad (j = 1, \dots, d),$$

and for every $i, j \in \{1, 2\}$ we set

$$b_{ij}(x, \xi) = \hat{b}(x, w_k(\xi)).$$

Note that if $k = 0$, then $b_{ij}(x, \xi)$ is only Lipschitz in ξ. However, for $k = 1, 2, \ldots$ and fixed $x \in D$, $b_{ij}(x, \cdot)$ is in $C^k(\Xi)$.

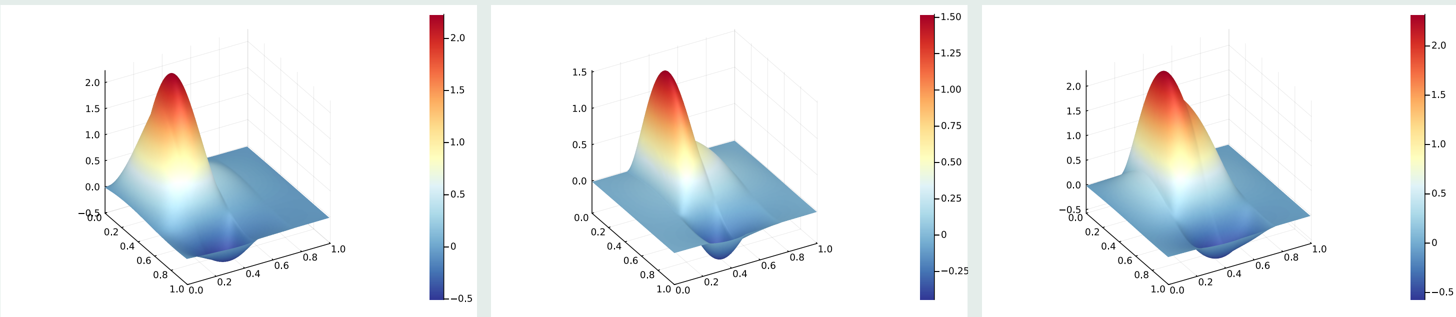
For the right hand side of the differential equation, we consider

$$g(x, \nu) = \sin(2x_1)\sin(2x_2) + 10^{-2}\nu$$

where $(x_1, x_2) \in D$ and $\nu \in \mathbb{R}$. The parameter ν is understood to be part of the vector ξ and Ξ will be a subset of $\mathbb{R}^{d+1}$. Given $z \in H$, we then consider the random elliptic PDE

$$A(\xi)u = z + g(\cdot, \xi).$$

For three randomly chosen $\xi_1, \xi_2, \xi_3 \in \Xi$ and $z \equiv 0$ we plot the resulting solutions:



Uncontrolled, random states: Three realizations of $u(\xi)$ computed by setting $z \equiv 0$.

The PDE constrained stochastic optimization problem is

$$\min\ F(z)\ =\ \frac{1}{2}\mathbb{E}_{\mathbb{P}}[\|u-\widetilde{u}\|_H^2]+\frac{\alpha}{2}\|z\|_H^2$$

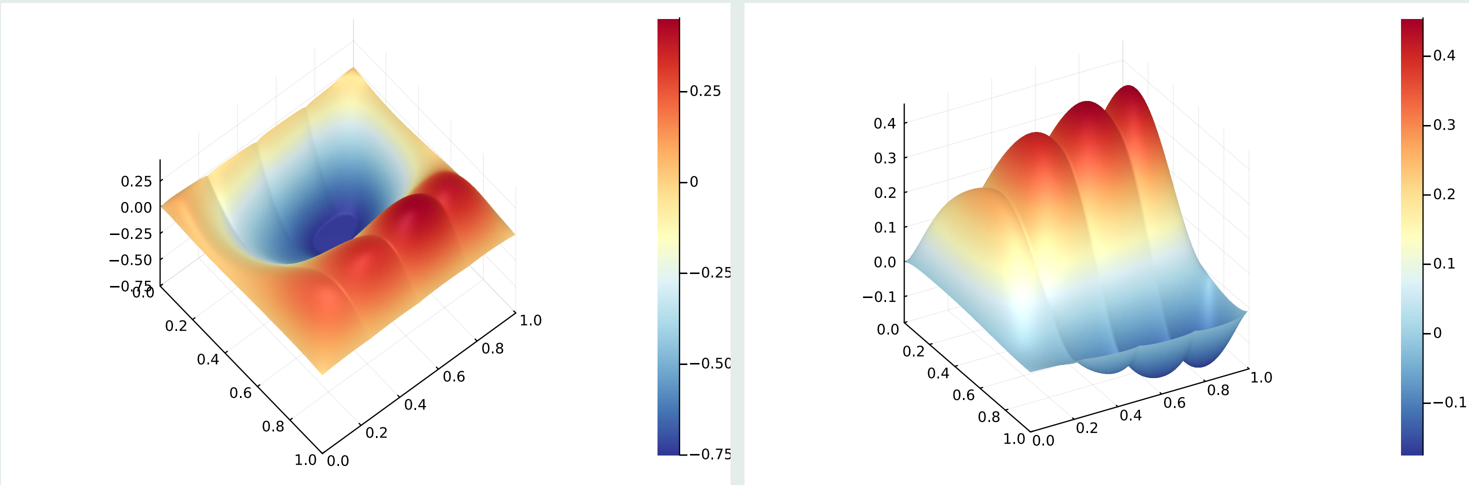
subject to $A(\xi)u = z + g(\xi)$ and $z \in Z_{\text{ad}} \quad (\mathbb{P}\text{-a.e. } \xi \in \Xi)$,

where we set $\alpha = 1.0$, $\widetilde{u} \equiv 1/2$, $Z_{\text{ad}} = \left\{ z \in H : -\frac{3}{4} \leq z(x) \leq \frac{3}{4}, \text{ a.e. } x \in D \right\}$.

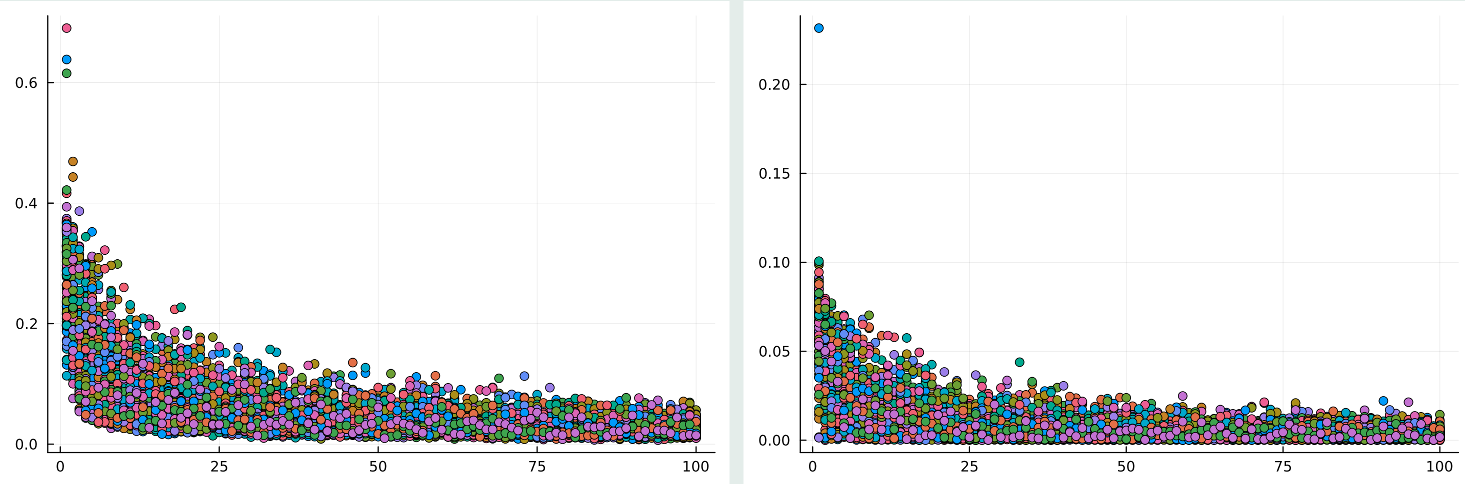
Solution method:

- For fixed $n \in \mathbb{N}$, we let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be independent identically distributed (iid) Ξ-valued random variables on a probability space $(\Omega, \mathcal{F}, P)$ with common distribution $\mathbb{P} = P \circ \xi_1^{-1}$.
- The domain D is triangulated by a uniform mesh rule and we use a standard H^1 finite element discretization for solving the PDE. The theoretical optimality conditions indicate that the optimal solutions $z(\mathbb{P})$, $z(\mathbb{P}_n)$ share the regularity in $V = H_0^1(D)$. Hence, we use the same discretization for the controls.
- The unique solutions satisfy nonsmooth equations which are solved by a semi-smooth Newton method.

We have plotted the result applying this scheme to our problem with $n = 500$ and a mesh defined by 128 x 128 grid. This corresponds to 16129 degrees of freedom for the control variables z and approximately 8 million degrees of freedom for the state variables associated with the 500 elliptic PDEs. The average controlled state is much closer to the desired state of $\widetilde{u} \equiv 0.5$ than observed in the uncontrolled states.



Optimal solution and average states: The left picture shows the optimal control computed for a random sample of size $n = 500$ on a uniform mesh with 16129 degrees of freedom. The right picture shows the effect of the optimal control $z(\mathbb{P}_n)$ on the state variables $u(\xi)$ for a new sample of size 500 by computing $\frac{1}{n}\sum_{i=1}^{n} A(\xi_i)^{-1}(z(\mathbb{P}_n) + g(\xi_i))$.



Stability statistics: The figures show the experimental convergence rates of the optimal solutions and optimal values. A coarser uniform mesh was chosen that corresponded to 900 degrees of freedom was used.

The left figure exhibits an experimental rate of $O(m^{-0.53656})$ for $\|z(\mathbb{P}_m) - z(\mathbb{P}_n)\|_H$.

The right figure exhibits an experimental rate of $O(m^{-0.66035})$ for $|v(\mathbb{P}_m) - v(\mathbb{P}_n)|$.

Conclusions and future work

- The empirical central limit theorem enables the application of resampling techniques like bootstrapping or subsampling to determine asymptotic confidence intervals.
- Monte Carlo methods have several asymptotic properties, but the slow convergence rate $O(n^{-\frac{1}{2}})$ requires a large sample size and, thus, a high number of PDE solves.
- Randomized Quasi-Monte Carlo methods could be a viable alternative due to the better convergence rate of about $O(n^{-1})$. However, their justification requires a completely different methodology, they produce dependent samples and the computation of confidence intervals is unsolved.

References

A. Eichhorn and W. Römisch: Stochastic integer programming: Limit theorems and confidence intervals, *Mathematics of Operations Research* 32 (2007), 118–135.

P. A. Guth, V. Kaarnioja, F. Y. Kuo, C. Schillings, I. H. Sloan: A Quasi-Monte Carlo method for optimal control under uncertainty, *SIAM/ASA Journal Uncertainty Quantification* 9 (2021), 354–383.

M. Hoffhues, W. Römisch and T. M. Surowiec: On quantitative stability in infinite-dimensional optimization under uncertainty, *Optimization Letters* 15 (2021), 2733–2756.

A. N. Kolmogorov and V. M. Tikhomirov: Epsilon-entropy and epsilon-capacity of sets in function spaces, *American Mathematical Society Translation Series* 2 17 (1961), 277–364.

D. N. Politis, J. P. Romano, and M. Wolf: *Subsampling*, Springer Series in Statistics, Springer, New York, 1999.

W. Römisch: Delta method, infinite dimensional, in: *Encyclopedia of Statistical Sciences* (S. Kotz, C. B. Read, N. Balakrishnan, B. Vidakovic eds.), Wiley, 2006.

A. Ruszczyński and A. Shapiro (eds.): *Stochastic Programming*, Handbooks in Operations Research and Management Science, Volume 10, Elsevier, Amsterdam 2003.

A. Shapiro: Asymptotic analysis of stochastic programs, *Annals of Operations Research* 30 (1991), 169–186.

F. Topsøe: On the connection of P-continuity and P-uniformity in weak convergence, *Probability Theory and its Applications* 12 (1967), 281–290.

A. W. van der Vaart: New Donsker classes, *The Annals of Probability* 24 (1996), 2128–2140.

V. M. Zolotarev: Probability metrics, *Theory of Probability and its Applications* 28 (1967), 278–302.