Infinite-dimensional spherical-radial decomposition for probabilistic functions

Kewei Wang, Georg Stadler

Courant Institute School of Mathematics, Computing and Data Science, New York University

https://arxiv.org/abs/2603.19907

Estimating Probabilistic Functions for Optimization

Consider stochastic optimization with chance constraint:

$$\min_{u \in U_{\text{ad}}} \quad \mathcal{J}(u) \quad \text{s.t.} \quad \varphi(u) \geq p,$$

- $\varphi(u) := \mathbb{P}(g(u, \xi) \leq 0)$ is a probabilistic function, with ξ an elliptical random variable taking values in a (inf-dim) Hilbert space $\mathcal{H}$.
- $g : U_{\text{ad}} \times \mathcal{H} \to \mathbb{R}$ is convex in ξ.

We want to find an estimator of $\varphi(u)$ that is **differentiable**, **unbiased**, and of **low variance**.

Existing Methods

Simple Monte Carlo (MC): $\varphi(u) \approx \bar{\varphi}_N(u) := \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\{g(u,z_i)\leq 0\}}$.
Unbiased, but high variance and not differentiable.

Spherical-radial decomposition (SRD): By KL expansion

$$\xi = \bar{\xi} + \textstyle\sum_{k=1}^{\infty} \xi_k \phi_k, \quad \bar{\xi} = \mathbb{E}[\xi],\ \xi_k \text{ scalar random variables,}$$

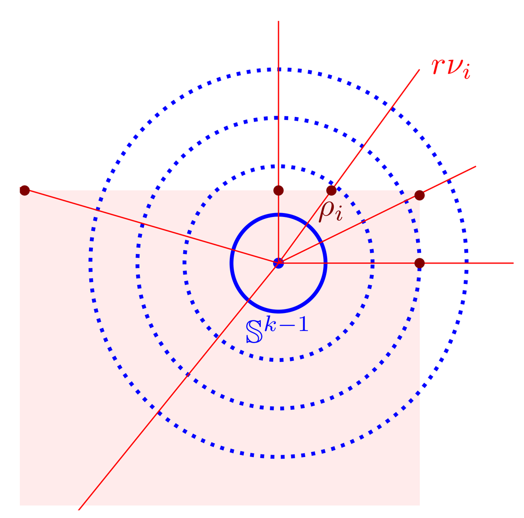
we truncate ξ to the first K KL modes: $\xi \approx \xi_K = \bar{\xi} + \sum_{k=1}^K \xi_k \phi_k$.

SRD for Gaussian: $\xi_K = \bar{\xi} + \tau L_K \boldsymbol{\nu}_K$.

- $\tau \sim \chi_K$, $\boldsymbol{\nu}_K \sim \mathcal{U}(\mathbb{S}^{K-1})$.
- $L_K : \boldsymbol{x} \mapsto \sum_{k=1}^K x_k \phi_k$, $\boldsymbol{x} \in \mathbb{R}^K$.

Estimator ($\rho_i := \sup_{r \geq 0} \{g(u, \bar{\xi} + r L_K \boldsymbol{v}_i) \leq 0\}$):

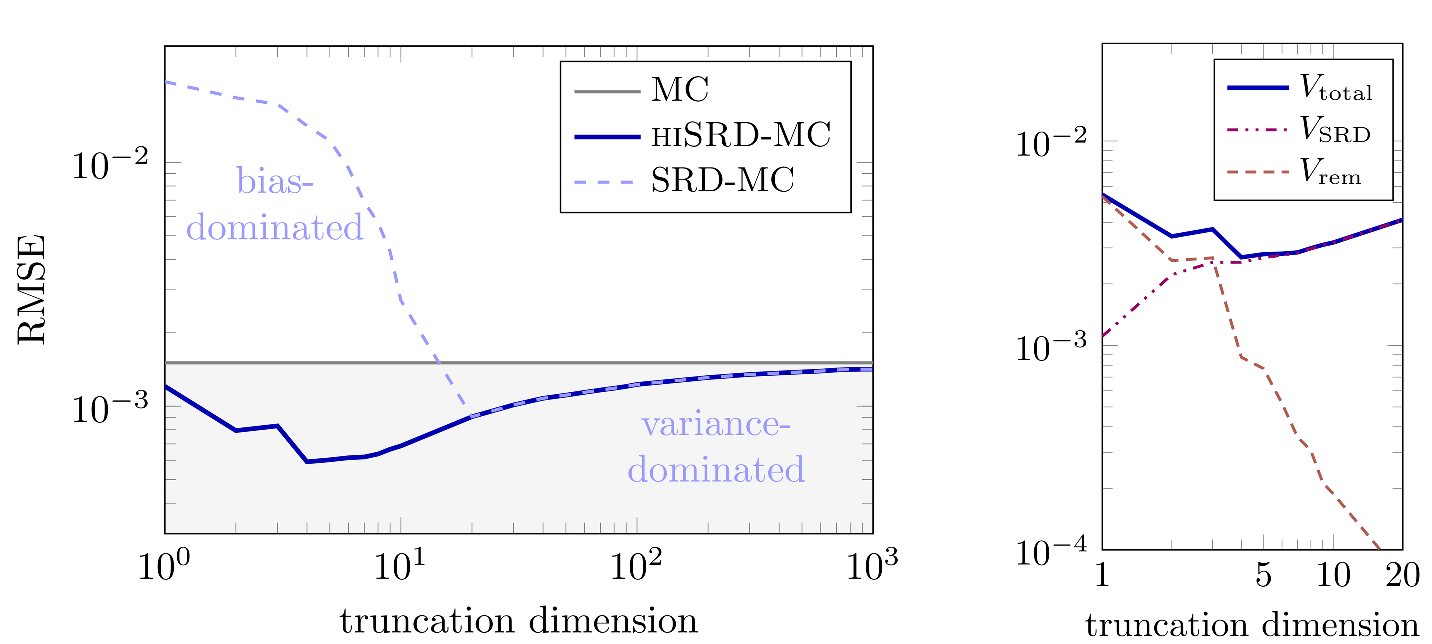
$$\varphi(u) \approx \tilde{\varphi}_N^{(K)}(u) := \frac{1}{N}\sum_{i=1}^N F_{\chi_K}(\rho_i).$$



Derivative:

$$\nabla \tilde{\varphi}_N^{(K)}(u) = -\frac{1}{N}\sum_{i=1}^N \frac{f_{\chi_K}(\rho_i)}{\left\langle \nabla_z g(u, \bar{\xi} + \rho_i L_K \boldsymbol{v}_i), L_K \boldsymbol{v}_i \right\rangle} \nabla_u g(u, \bar{\xi} + \rho_i L_K \boldsymbol{v}_i).$$

There is a trade-off between truncation bias (large for small K) and variance (converges to the variance MC at $O(1/\sqrt{K})$ for large K).



Left: root mean squared error (RMSE) versus truncation dimension K, for standard Monte Carlo (gray horizontal line), finite-dimensional SRD (dashed) and hiSRD (solid, **our method**). Right: the individual components V_{SRD} and V_{rem} contributing to the variance of hiSRD.

Hybrid Infinite-Dimensional SRD (hiSRD)

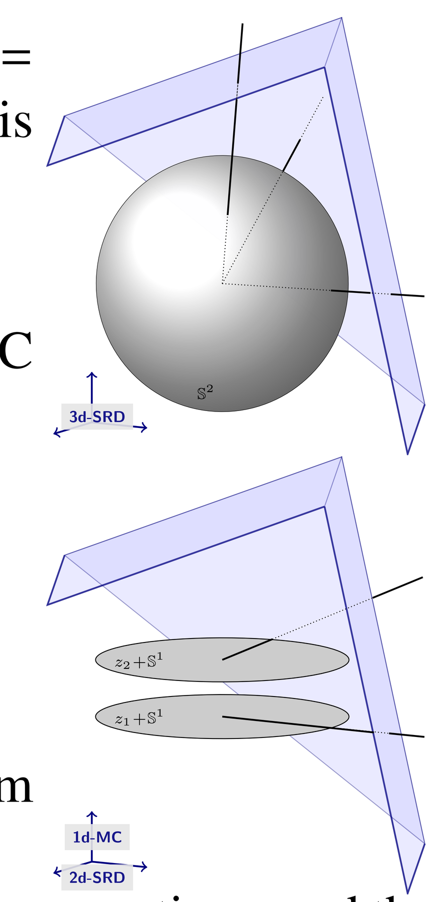
Space decomposition: $\mathcal{H} = \mathcal{H}_K \oplus \mathcal{H}_R$, $\mathcal{H}_K := \text{span}\{\phi_k : k \leq K\}$, $\mathcal{H}_R := \text{span}\{\phi_k : k \geq K+1\}$. ξ is correspondingly decomposed as

$$\xi = \bar{\xi} + \xi_K + \xi_R = \bar{\xi} + \tau L_K \boldsymbol{\nu}_K + \xi_R.$$

hiSRD: perform SRD on ξ_K, and keep standard MC integration for the remainder ξ_R:

$$\varphi(u) = \int_{\mathcal{H}_R}\int_{\boldsymbol{v}\in\mathbb{S}^{K-1}} e(u, \boldsymbol{v}, z) d\mu_{\mathcal{U}}(\boldsymbol{v}) d\mu_{\xi_R}(z)$$
$$\approx \frac{1}{N}\sum_{i=1}^N e(u, \boldsymbol{v}_i, z_i) =: \tilde{\varphi}_N(u),$$
$$e(u, \boldsymbol{v}, z) := \mu_{\chi_K}\left(\left\{r \geq 0 : g(u, \bar{\xi} + z + r L_K \boldsymbol{v}) \leq 0\right\}\right);$$

$\{\boldsymbol{v}_i\}_{i=1}^N$ and $\{z_i\}_{i=1}^N$ are independently sampled from $\mathcal{U}(\mathbb{S}^{K-1})$ and ξ_R.



Differentiability: $\tilde{\varphi}_N(u)$ is differentiable under mild assumptions, and the form of $\nabla \tilde{\varphi}_N(u)$ is generalized from $\nabla \tilde{\varphi}_N^{(K)}(u)$.

Variance decomposition

$$\text{Var}(\tilde{\varphi}_N(u)) = \frac{1}{N}\text{Var}_{\boldsymbol{\nu},\xi_R}(e(u, \boldsymbol{\nu}, \xi_R)) = \frac{1}{N}(V_{\text{SRD}} + V_{\text{rem}}).$$

- $V_{\text{SRD}} := \mathbb{E}_{\xi_R}[\text{Var}_{\nu}(e(u, \boldsymbol{\nu}, \xi_R)|\xi_R)]$ is the the expected conditional variance of the SRD estimator.
- $V_{\text{rem}} := \text{Var}_{\xi_R}(\mathbb{E}_{\nu}[e(u, \boldsymbol{\nu}, \xi_R)|\xi_R])$ captures the variance of sampling the remainder.

Entry/Exit Ray Evaluation

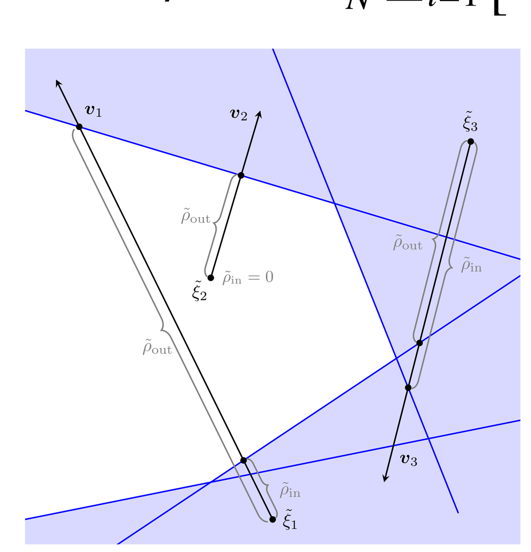
Assumption: $g(u, \xi) = \max_{j=1,\ldots,M} g_j(u, \xi)$, each g_j affine in ξ.
Unlike standard SRD, the shifted means $\bar{\xi} + z_i$ may be outside of the feasible set. We define a variant of entry and exit ray distances for this geometry:

$$\tilde{\rho}_{\text{in}}^{(i)} := \max_{j \in J_o} \inf\{r \geq 0 : g_j(u, \bar{\xi} + z_i + r L_K \boldsymbol{v}_i) \leq 0\},$$
$$\tilde{\rho}_{\text{out}}^{(i)} := \min_{j \in J_i} \sup\{r \geq 0 : g_j(u, \bar{\xi} + z_i + r L_K \boldsymbol{v}_i) \leq 0\}.$$

Partition of the index set:

$$J_{\text{i}} := \{j \in \{1, \ldots, M\} : g_j(u, \bar{\xi} + z_i) < 0\},$$
$$J_{\text{o}} := \{j \in \{1, \ldots, M\} : g_j(u, \bar{\xi} + z_i) > 0\}.$$

Evaluation: $\tilde{\varphi}_N(u) = \frac{1}{N}\sum_{i=1}^N \left[\max\{0, F_{\chi_K}(\tilde{\rho}_{\text{out}}^{(i)}) - F_{\chi_K}(\tilde{\rho}_{\text{in}}^{(i)})\}\right]$.



Three different intersection scenarios: a ray entering and exiting the feasible set from the outside, with $0 < \tilde{\rho}_{\text{in}} < \tilde{\rho}_{\text{out}}$ $(i = 1)$, a ray originating inside the feasible set, with $0 = \tilde{\rho}_{\text{in}} < \tilde{\rho}_{\text{out}}$ $(i = 2)$, and a ray that completely misses the feasible set, illustrating the case in which $\tilde{\rho}_{\text{out}} < \tilde{\rho}_{\text{in}}$ $(i = 3)$.

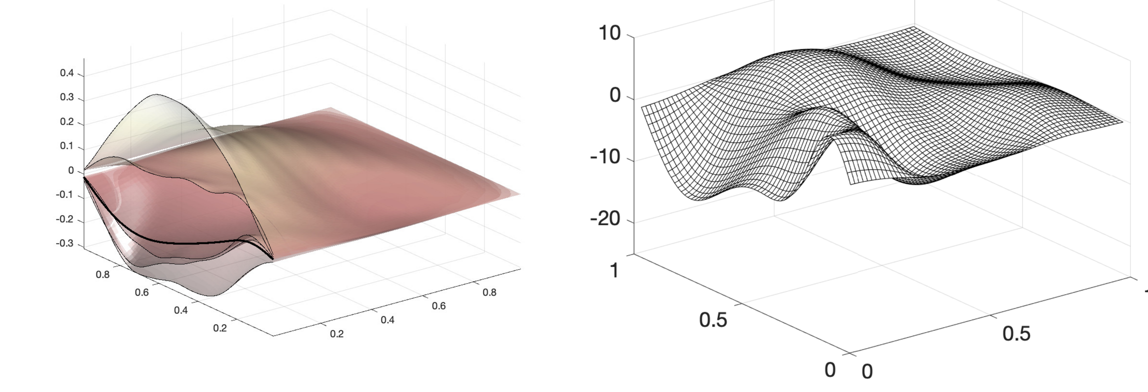
Optimal Control under Uncertainty

Setup: u deterministic control, $y(\xi)$ random state variable, ξ a Gaussian random field as Neumann boundary data on $\partial\mathcal{D}_2 = \{0\} \times [0, 1] \subseteq \partial\mathcal{D}$.
Objective: Minimize a tracking objective (with target state y_d and regularization α) subject to the elliptic PDE $-\Delta y = f + u$, $x \in \mathcal{D} = (0, 1)^2$:

$$\mathcal{J}(u) := \frac{1}{2}\mathbb{E}_{\xi}\left[\int_{\mathcal{D}} (y(u) - y_d)^2\, dx\right] + \frac{\alpha}{2}\int_{\mathcal{D}} u^2\, dx.$$

Constraint: A joint chance constraint ensuring the state stays below a threshold $\bar{y}$ almost everywhere, with 90% probability:

$$\mathbb{P}(y(x; u) \leq \bar{y} \ \text{ for a.a. } x \in \mathcal{D}) \geq 0.9.$$



Optimal control solution for $p = 0.9$ and chance constraints with $\bar{y} = 0.3$. Left: samples from the state variable y. Right: optimal control u. Source: [2].

Gaussian Process Regression (GPR)

Setup: Observations $(\boldsymbol{X}, \boldsymbol{y})$, optimize hyperparameters $u = (l, \sigma, \sigma_{\text{n}})$:

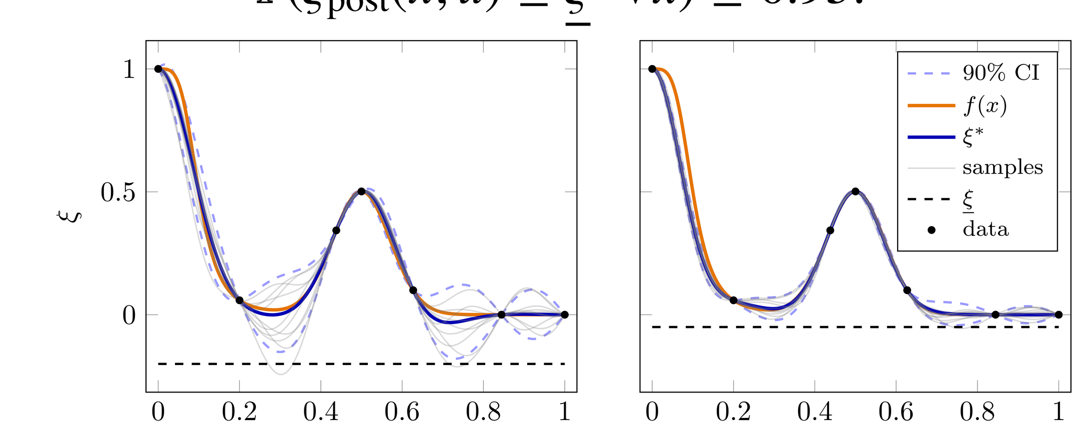
$$\mathcal{K}_u(x, x') = \sigma^2 \exp\left(\frac{-\|x - x'\|_2^2}{2l^2}\right) + \sigma_n^2 \delta_{x,x'}.$$

Objective: Minimize the negative marginal log-likelihood

$$\mathcal{J}(u) := -\log p(\boldsymbol{y}|u) = \frac{1}{2}((\boldsymbol{y} - \boldsymbol{\xi}_0)^T K_u^{-1}(\boldsymbol{y} - \boldsymbol{\xi}_0) + \log|K_u| + N\log(2\pi)).$$

Constraint: A joint chance constraint ensuring the posterior predictive process $\xi_{\text{post}} \sim \mathcal{N}(m_{\text{post}}, \mathcal{K}_{\text{post}})$ stays above $\underline{\xi}$ with 95% probability:

$$\mathbb{P}(\xi_{\text{post}}(x; u) \geq \underline{\xi} \quad \forall x) \geq 0.95.$$



Left: $\underline{\xi} \equiv -0.2$. Right: $\underline{\xi} \equiv -0.05$.

References

[1] K. Wang and G. Stadler, Infinite-dimensional spherical-radial decomposition for probabilistic functions, with application to constrained optimal control and Gaussian process regression, arXiv:2603.19907 (2026).

[2] R. Henrion, G. Stadler, and F. Wechsung, Optimal Control under Uncertainty with Joint Chance State Constraints: Almost-Everywhere Bounds, Variance Reduction, and Application to (Bi) linear Elliptic PDEs, SIAM/ASA Journal on Uncertainty Quantification, 13 (2025).

Acknowledgements

Partially supported by the U.S. DOE under #DE-SC0024721 and NSF under #2411229.