Multilevel Monte Carlo for Bayesian Inverse Problems

R. Gantner
ETH Zürich

Affiliation: *Seminar for Applied Mathematics*, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland

Email: robert.gantner@sam.math.ethz.ch – **URL**: http://www.sam.math.ethz.ch

Master: MSc. CSE ETH Zürich

Ph.D. (2013-2016): ETH Zürich

Supervisor(s): Prof. Christoph Schwab (ETH Zürich) and Dr. Claudia Schillings (ETH Zürich)

Introduction In recent years, various methods have been developed for solving parametric operator equations, focusing on the estimation of parameters given measurements of the parametric solution, subject to a stochastic observation error model. The *Bayesian approach* [7] to such inverse problems for PDEs will be considered here and solved using adaptive, deterministic sparse tensor Smolyak quadrature schemes from [4, 5]. Multiple solutions of the Bayesian inverse problem based on different measurements are often averaged using a standard Monte Carlo approach. We develop a multilevel Monte Carlo method achieving an error of the same order while requiring less work [1, 2, 3].

Bayesian Inversion of Parametric Operator Equations We assume an operator equation depending on a distributed, uncertain parameter u with values in a separable Banach space X of the form

$$\text{Given } u \in \tilde{X} \subseteq X, \text{ find } q \in \mathcal{X}: \quad A(u;q) = F(u) \quad \text{in} \quad \mathcal{Y}' \tag{1}$$

where we denote by $\mathcal{X}$ and $\mathcal{Y}$ two reflexive Banach spaces over $\mathbb{R}$ with (topological) duals $\mathcal{X}'$ and $\mathcal{Y}'$, respectively and $A(u;\cdot) \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$. Assuming that the forcing function $F : \tilde{X} \mapsto \mathcal{Y}'$ is known, and the uncertain operator $A(u;\cdot) : \mathcal{X} \mapsto \mathcal{Y}'$ is locally boundedly invertible for uncertain input u in a sufficiently small neighborhood $\tilde{X}$, we define the *uncertainty-to-observation map* $\mathcal{G} : \tilde{X} \mapsto \mathbb{R}^K$ with the structure

$$X \supseteq \tilde{X} \ni u \mapsto \mathcal{G}(u) := \mathcal{O}(G(u;F)) \in Y\ . \tag{2}$$

Here, $\tilde{X} \ni u \mapsto q(u) = G(u;F) \in \mathcal{X}$ denotes the response of the forward problem for a given instance of $u \in \tilde{X}$ and $\mathcal{O}$ an *observation operator* $\mathcal{O} \in \mathcal{L}(\mathcal{X}, \mathbb{R}^K),\ K < \infty$. The goal of computation is the low-order statistics of a *quantity of interest* (QoI) ϕ given noisy observational data δ of the form $\delta = \mathcal{G}(u) + \eta$, where δ represents the observation $\mathcal{G}(u)$ perturbed by the normally distributed noise η. We assume u to be parametrized by $u = u(\mathbf{y}) := \langle u \rangle + \sum_{j\in\mathbb{J}} y_j \psi_j \in X$ for some "nominal" value $\langle u \rangle$ and coefficient sequence $\mathbf{y} = (y_j)_{j\in\mathbb{J}}$, $\mathbb{J} = \{1, \ldots, J\}$ where the y_j are uniformly distributed on $[-1, 1]$.

Bayes' theorem characterizes moments of the QoI as mathematical expectations with respect to the prior measure μ_0 on U, which here is given as the countable product of uniform measures. In particular, we are interested in $\phi = G$, the response of the system. To this end, we use Bayes' Theorem to obtain an expression for $\mathbf{y}|\delta$, as in [6, 7]. This yields our desired expectation as an integral over the prior measure μ_0. Defining $Z_\delta := \int_U \exp(-\Phi(\mathbf{y};\delta))\, \mu_0(\mathrm{d}\mathbf{y}) > 0$, we obtain

$$\mathbb{E}^{\mu^\delta}[\phi] = \int_U \phi(\mathbf{y})\, \mu^\delta(\mathrm{d}\mathbf{y}) = \frac{1}{Z_\delta} \int_U \phi(\mathbf{y}) \exp\left(-\frac{1}{2}\|\delta - \mathcal{G}(\mathbf{y})\|_\Gamma^2\right) \mu_0(\mathrm{d}\mathbf{y}) =: \frac{Z'_\delta}{Z_\delta}. \tag{3}$$

This formulation of the expectation $\mathbb{E}^{\mu^\delta}[\cdot]$ is based on just one measurement δ. For a given model for the measurement errors η, we would like to additionally compute the expectation over the assumed error distribution, in this case $\gamma_\Gamma^K(\eta)$, the K-variate Gaussian measure with s.p.d. covariance matrix Γ.

Here, we assume the observation noise η to be statistically independent from the uncertain parameter u in (1). Thus, the total expectation of the QoI ϕ in terms of Z'_δ and Z_δ is $\mathbb{E}^{\gamma_\Gamma^K}\left[\mathbb{E}^{\mu^\delta}[\phi]\right] = \int_{\mathbb{R}^K} \frac{Z'_\delta}{Z_\delta}\Big|_{\delta=\mathcal{G}(\mathbf{y}_0)+\eta} \gamma_\Gamma^K(\mathrm{d}\eta)$, where $\mathcal{G}(\mathbf{y}_0)$ denotes the observation at the unknown, exact parameter $\mathbf{y}_0$.

In practice, we are given a set of measurements $\Delta := \{\delta_i,\ i = 1, \dots, M\}$ with which this outer expectation should be approximated. The measurements can be taken at different positions, i.e. with respect to different observation maps $\mathcal{O}_i$ in (2). We consider the notationally more convenient case where the measurements are all obtained using the same observation map. We do, however, impose the restriction that the measurements are homoscedastic, i.e. δ_i is Gaussian with the same covariance Γ for all $\delta_i \in \Delta$.

Approximation of Posterior Expectation The inner expectation over the posterior distribution μ^δ is replaced by an approximation $E^{\mu^\delta}_{\tau_L}[\phi]$ with tolerance parameter $\tau_L > 0$. We assume the work required to compute this approximation to be bounded by $C(\Gamma)\tau_L^{-s}$, with $C(\Gamma) > 0$ independent of τ_L and $s > 0$. Our method of choice for approximating $\mathbb{E}^{\mu^\delta}[\phi]$ is the adaptive Smolyak quadrature algorithm developed in [4, 5], which adaptively constructs a sparse tensor quadrature rule that approximates Z_δ and Z'_δ. For forward problems belonging to a certain sparsity class, analytic regularity of the Bayesian posterior suggests dimension-independent convergence rates for the adaptive, deterministic Smolyak quadrature fulfilling the work bound $C(\Gamma)\tau_L^{-s}$, where s depends on the sparsity class.

Binned Multilevel Monte Carlo The approach proposed here is based on the multilevel Monte Carlo method originally applied by [3] and formulated in the current form for PDEs by [1]. Our approximation to $\mathbb{E}^{\gamma_\Gamma^K}\left[\mathbb{E}^{\mu^\delta}[\phi]\right]$ is given by

$$E^{\gamma_\Gamma^K}_{\mathrm{ML},L}[E^{\mu^\delta}_{\tau_L}[\phi]] := \sum_{\ell=0}^{L} E^{\gamma_\Gamma^K}_{M_\ell}\left[E^{\mu^\delta}_{\tau_\ell}[\phi] - E^{\mu^\delta}_{\tau_{\ell-1}}[\phi]\right], \tag{4}$$

where $E^{\gamma_\Gamma^K}_{M_\ell}[\cdot]$ denotes the sample mean over M_ℓ samples and $E^{\mu^\delta}_{\tau_\ell}[\cdot]$ denotes the posterior expectation approximation introduced above. We show that, assuming a certain distribution of samples per level, one can find a tolerance for each level such that the rate of convergence of the error e_{tot} vs. the work W^L_{tot} fulfills the optimal relationship $e_{\mathrm{tot}} = \mathcal{O}\left(\left(W^L_{\mathrm{tot}}\right)^{-\frac{1}{2}}\right)$, which is superior to Monte Carlo depending on the sparsity class of the underlying problem.

Applications The proposed approach is applicable for instance for definite and indefinite elliptic and parabolic evolution problems with scalar and tensoral unknowns. Furthermore, uncertainty in domains and high-dimensional initial value problems can be treated. Numerical experiments yielding the optimal rate of convergence when using the binned multilevel Monte Carlo algorithm will be presented and compared to standard Monte Carlo simulations.

References

[1] A. Barth, Ch. Schwab, and N. Zollinger. Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients. *Numer. Math.*, 119(1):123–161, September 2011.

[2] R. Gantner, Cl. Schillings, and Ch. Schwab. Multilevel Monte Carlo for Bayesian inverse problems. (submitted), 2014.

[3] M. B. Giles. Multilevel Monte Carlo path simulation. *Operations Research*, pages 1–25, 2008.

[4] Cl. Schillings and Ch. Schwab. Sparse, adaptive Smolyak quadratures for Bayesian inverse problems. *Inverse Problems*, 29(6):065011, 2013.

[5] Cl. Schillings and Ch. Schwab. Sparsity in Bayesian inversion of parametric operator equations. Technical Report 2013-17, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2013.

[6] Ch. Schwab and A. M. Stuart. Sparse deterministic approximation of Bayesian inverse problems. *Inverse Problems*, 28(4):045003, 2012.

[7] A. M. Stuart. Inverse problems: a Bayesian perspective. *Acta Numer.*, 19:451–559, 2010.

Short biography – After completing his master's degree in Computational Science and Engineering at ETH Zurich, Robert Gantner is now pursuing a PhD in the group of Prof. Schwab at the Seminar for Applied Mathematics. The main focus is the development and high-performance implementation of novel algorithms for Bayesian inverse problems, including Multilevel Monte Carlo and Smolyak quadratures.