J. Nonlinear Var. Anal. 5 (2021), No. 6, pp. 865-880
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.5.2021.6.02

AN ITERATIVELY REGULARIZED STOCHASTIC GRADIENT METHOD FOR ESTIMATING A RANDOM PARAMETER IN A STOCHASTIC PDE. A VARIATIONAL INEQUALITY APPROACH

BAASANSUREN JADAMBA[1,*], AKHTAR A. KHAN[1], MIGUEL SAMA[2], YIDAN YANG[1]

[1]*School of Mathematical Sciences, Rochester Institute of Technology, 85 Lomb Memorial Drive, Rochester, New York, 14623, USA*
[2]*Departamento de Matemática Aplicada, Universidad Nacional de Educación a Distancia, Calle Juan del Rosal, 12, 28040 Madrid, Spain*

Abstract. We develop a variational inequality approach for the inverse problem of identifying a stochastic parameter in a stochastic partial differential equation. An iteratively regularized projected stochastic gradient scheme for variational inequalities posed in a Hilbert space is proposed. By employing the martingale theory, we give a complete convergence analysis for the iterative scheme under weaker conditions on the random noise than those commonly imposed in the available literature. Preliminary numerical results on the inverse problem demonstrate the efficacy of the developed framework.
Keywords. Energy least-squares; Inverse problem; Iterative regularization; Stochastic parameter identification; Stochastic PDEs.

1. Introduction

Let $(\Omega,\mathbb{F},\mathbb{P})$ be a probability space, and let $D\subset\mathbb{R}^n$ be a bounded domain with sufficiently smooth boundary ∂D. Given two random fields $a:\Omega\times D\to\mathbb{R}$ and $f:\Omega\times D\to\mathbb{R}$, we consider the SPDE of finding a random field $u:\Omega\times D\to\mathbb{R}$ that almost surely satisfies

$$-\nabla\cdot(a(\omega,x)\nabla u(\omega,x)) = f(\omega,x), \text{ in } D, \tag{1.1a}$$

$$u(\omega,x) = 0, \text{ on } \partial D. \tag{1.1b}$$

SPDE (1.1) models critical real-world phenomena and has been studied extensively, see [1, 2].

This work focuses on the inverse problem of identifying the parameter a from a measurement of the solution u of (1.1). We note that another inverse problem related to (1.1) is the linear inverse problem of identifying the source term f from a measurement of the solution u. This linear inverse problem becomes the optimal control problem when f is viewed as the control variable. We will study the inverse problem as a stochastic optimization problem of the form:

$$\min_{a\in\mathbb{K}} \mathbb{J}(a) := \mathbb{E}\left[J(a,\omega)\right]. \tag{1.2}$$

*Corresponding author.
E-mail addresses: bxjsma@rit.edu (B. Jadamba), aaksma@rit.edu (A.A. Khan), msama@ind.uned.es (M. Sama), yy2513@rit.edu (Y. Yang).
Received June 11, 2021; Accepted August 5, 2021.

Here $\mathbb{K}$ is a constraint set, which is a subset of a real Hilbert space H, $J(a,\omega)$ is a suitable misfit function, and $\mathbb{E}$ is the expectation with respect to the probability space $(\Omega, \mathbb{F}, \mathbb{P})$.

For the aforementioned inverse problem, we will employ a convex misfit function, and hence a necessary and sufficient optimality condition for the stochastic optimization problem is a stochastic variational inequality. Therefore, we will work in a general variational inequality framework and develop an iteratively regularized projected stochastic gradient approach. We note that the dynamic field of stochastic approximation that began by Robbins and Monro [3] has been applied to a wide variety of research domains, see [4, 5, 6, 7, 8] and the cited references. Recent developments in machine learning and stochastic variational inequalities have rekindled interest in stochastic approximation, see [9, 10, 11, 12, 13, 14, 15, 16, 17, 18]. On the other hand, stochastic PDE-constrained optimization problems attracted a great deal of attention in recent years. Such problems emerge from two sources, the inverse problems and optimal control problems, see [19, 20, 21, 22, 23, 24]. For an overview of inverse problems, see [25, 26, 27, 28, 29, 30, 31, 32].

This work is a continuation of our recent research [33], where a variant of the above ELS functional was recently used for estimating a deterministic coefficient in SPDEs. The recent works by Geiersbach and Pflug [34] and Martin, Krumschield, and Nobile [35] should also be mentioned where the stochastic approximation framework was used to study optimal control problems. We note that [34, 35] focused on a deterministic control in stochastic PDEs. So far, no research has addressed estimating stochastic parameters or control variables by a stochastic approximation approach. In this work, we develop a new stochastic approximation approach for nonlinear inverse problems of identifying stochastic parameters.

We organize the contents of this paper into five sections. Section 2 presents a new iteratively regularized stochastic gradient method and provides its convergence. Section 3 focuses on the inverse problem and develops an ELS-based stochastic optimization framework. Besides providing technical details on the two functionals in a continuous setting, we also provide a discretization scheme, including discrete formulas for the objective functionals and their gradient. Numerical experiments, given in Section 4, demonstrates the feasibility and the efficacy of the developed framework. The paper concludes with some remarks.

2. An Iteratively Regularized Projected Stochastic Gradient Method

Let H be a real Hilbert space, K be a closed, and convex subset of H, and $F : H \to H$ be a given map. We consider the variational inequality of finding $u \in K$ such that

$$\langle F(u), v-u\rangle \geq 0, \quad \text{for every } v \in K. \tag{2.1}$$

Let $\mathscr{S}(F,K)$ be the set of all solutions of variational inequality (2.1).

Let $\{\varepsilon_n\}$ be sequence of positive regularization parameters such that $\varepsilon_n \to 0$ as $n \to \infty$. Along with (2.1) , we consider the regularized variational inequality of finding $u_{\varepsilon_n} \in K$ such that

$$\langle F(u_{\varepsilon_n}) + \varepsilon_n u_{\varepsilon_n}, v - u_{\varepsilon_n}\rangle \geq 0, \quad \text{for every } v \in K. \tag{2.2}$$

The following well-known result for the regularized solutions will be used shortly [36].

Theorem 2.1. *Let H be a real Hilbert space, $K \subset H$ be closed and convex, and $F : H \to H$ be monotone and hemicontinuous. Let $\{\varepsilon_n\}$ be a decreasing sequence of positive reals with $\varepsilon_n \to 0$*

as $n \to \infty$, and let $\{u_{\varepsilon_n}\}$ be the sequence of the regularized solutions of (2.2). *Then,*

$$\|u_{\varepsilon_n} - u_{\varepsilon_{n-1}}\| \leq \frac{|\varepsilon_n - \varepsilon_{n-1}|}{\varepsilon_n}\|u_{\varepsilon_n}\|. \tag{2.3}$$

If variational inequality (2.1) *is solvable, then the following estimate holds*

$$\|u_{\varepsilon_n} - u_{\varepsilon_{n-1}}\| \leq \frac{|\varepsilon_n - \varepsilon_{n-1}|}{\varepsilon_n}\|\bar{u}\|, \tag{2.4}$$

where $\bar{u}$ is the minimal norm solution of (2.1).

Let $(\Omega, \mathbb{F}, \mathbb{P})$ be a probability space, and let $\{\omega_n\}$ be a sequence of H-valued random variables on $(\Omega, \mathbb{F}, \mathbb{P})$. We consider the following algorithm:

1: Initialization. A random $u_1 \in K$ with $\mathbb{E}\left[\|u_1\|^2\right] < \infty$.
2: At step n, compute $u_{n+1} \in K$ by

$$u_{n+1} = P_K\left[u_n - \alpha_n\left(F(u_n) + \varepsilon_n u_n + \omega_n\right)\right], \tag{2.5}$$

where α_n is the step-size, ε_n is the regularization parameter, and P_K is the projection.

We recall that, given the probability space $(\Omega, \mathbb{F}, \mathbb{P})$, a filtration $\{\mathbb{F}_n\} \subset \mathbb{F}$ is an increasing sequence of σ-algebras. A sequence of random variable $\{\omega_n\}$ is said to be adapted to a filtration $\mathbb{F}_n$, if and only if, $\omega_n \in \mathbb{F}_n$ for all $n \in \mathbb{N}$, that is, ω_n is $\mathbb{F}_n$-measurable. Moreover, the natural filtration is the one generated by the sequence $\{\omega_n\}$ and is given by $\mathbb{F}_n = \sigma(\omega_n : m \leq n)$.

The following result by Robbins and Siegmund [37] will also be used.

Theorem 2.2. *Let $\mathbb{F}_n$ be an increasing sequence of σ-algebras, and V_n, a_n, b_n, and c_n be non-negative random variables adapted to $\mathbb{F}_n$. Assume that $\sum_{n=1}^{\infty} a_n < \infty$ and $\sum_{n=1}^{\infty} b_n < \infty$, almost surely, and*

$$\mathbb{E}[V_{n+1}|\mathbb{F}_n] \leq (1 + a_n)V_n - c_n + b_n.$$

Then $\{V_n\}$ is almost surely convergent and $\sum_{n=1}^{\infty} c_n < \infty$, almost surely.

The following result gives the convergence analysis for the scheme (2.5):

Theorem 2.3. *Let H be a real Hilbert space, $K \subset H$ be closed and convex, and $F : H \to H$ be monotone. Let $\{u_n\}$ be the sequence generated by* (2.5)*, and let $\mathbb{F}_n$ be a filtration on the probability space $(\Omega, \mathbb{F}, \mathbb{P})$ such that $\{u_n\}$ is $\mathbb{F}_n$-measurable. Let the solution set $\mathscr{S}(F, K)$ be nonempty. Assume that the following conditions hold:*

(A_1): *There is a constant $c > 0$ such that $\|F(u)\| \leq c(1 + \|u\|)$, for every $u \in K$.*
(A_2): *There are constants $c_1 \geq 0$ and $c_2 > 0$ such that*

$$\|\mathbb{E}\left[\omega_n|\mathbb{F}_n\right]\| \leq c_1\beta_n\left(1 + \|F(u_n)\|\right), \quad \beta_n > 0, \tag{2.6}$$

$$\mathbb{E}\left[\|\omega_n\|^2|\mathbb{F}_n\right] \leq c_2\left(1 + \frac{1}{\delta_n}\|F(u_n)\|^2\right), \quad \delta_n > 0. \tag{2.7}$$

(A_3): *The bounded sequences $\{\varepsilon_n\}$, $\{\alpha_n\}$, $\{\beta_n\}$, and $\{\delta_n\}$ satisfy the following:*

$$\sum_{n\in\mathbb{N}} \varepsilon_n\alpha_n = \infty, \quad \sum_{n\in\mathbb{N}} \alpha_n^2 < \infty, \quad \sum_{n\in\mathbb{N}} \frac{\alpha_n^2}{\delta_n} < \infty, \quad \sum_{n\in\mathbb{N}} \alpha_n\beta_n < \infty, \quad \sum_{n\in\mathbb{N}} \left(\frac{1 + \alpha_n\varepsilon_n}{\alpha_n\varepsilon_n}\right)\left|\frac{\varepsilon_{n-1} - \varepsilon_n}{\varepsilon_n}\right|^2 < \infty.$$

Then, $\|u_{n+1} - u_{\varepsilon_n}\| \to 0$, almost surely.

Proof. For $n \in \mathbb{N}$, let $u_{\varepsilon_n} \in K$ be the solution of the regularized variational inequality with the regularization parameter ε_n. That is, the element $u_{\varepsilon_n} \in K$ satisfies the variational inequality

$$\langle F(u_{\varepsilon_n}) + \varepsilon_n u_{\varepsilon_n}, z - u_{\varepsilon_n} \rangle \geq 0, \quad \text{for every } z \in K,$$

which, by the variational characterization of the projection map P_K from H onto K, implies that

$$u_{\varepsilon_n} = P_K[u_{\varepsilon_n} - \alpha_n(F(u_{\varepsilon_n}) + \varepsilon_n u_{\varepsilon_n})].$$

By using the iterative scheme (2.5) defining u_{n+1} and the above identity, we obtain

$$\begin{aligned}
&\|u_{n+1} - u_{\varepsilon_n}\|^2 \\
&= \|P_K[u_n - \alpha_n(F(u_n) + \varepsilon_k u_n + \omega_n)] - P_K[u_{\varepsilon_n} - \alpha_n(F(u_{\varepsilon_n}) + \varepsilon_n u_{\varepsilon_n})]\|^2 \\
&\leq \|\,[u_n - \alpha_n(F(u_n) + \varepsilon_n u_n + \omega_n)] - [u_{\varepsilon_n} - \alpha_n(F(u_{\varepsilon_n}) + \varepsilon_n u_{\varepsilon_n})]\,\|^2 \\
&= \|u_n - u_{\varepsilon_n} - \alpha_n(F(u_n) - F(u_{\varepsilon_n})) - \alpha_n\varepsilon_n(u_n - u_{\varepsilon_n}) - \alpha_n\omega_n\|^2 \\
&= \|u_n - u_{\varepsilon_n}\|^2 + \alpha_n^2\|F(u_n) - F(u_{\varepsilon_n})\|^2 + \alpha_n^2\varepsilon_n^2\|u_n - u_{\varepsilon_n}\|^2 + \alpha_n^2\|\omega_n\|^2 \\
&- 2\alpha_n\varepsilon_n\|u_n - u_{\varepsilon_n}\|^2 - 2\alpha_n\langle F(u_n) - F(u_{\varepsilon_n}), u_n - u_{\varepsilon_n}\rangle \\
&- 2\alpha_n\langle u_n - u_{\varepsilon_n}, \omega_n\rangle + 2\alpha_n^2\langle F(u_n) - F(u_{\varepsilon_n}), \omega_n\rangle \\
&+ 2\alpha_n^2\varepsilon_n\langle u_n - u_{\varepsilon_n}, \omega_n\rangle + 2\alpha_n^2\varepsilon_n\langle F(u_n) - F(u_{\varepsilon_n}), u_n - u_{\varepsilon_n}\rangle,
\end{aligned}$$

and by taking the expectation past $\mathbb{F}_n$, we obtain

$$\begin{aligned}
&\mathbb{E}\,[\|\,u_{n+1} - u_{\varepsilon_n}\|^2|\mathbb{F}_n] \\
&\leq (1 - 2\alpha_n\varepsilon_n + \alpha_n^2\varepsilon_n^2)\|u_n - u_{\varepsilon_n}\|^2 + \alpha_n^2\|F(u_n) - F(u_{\varepsilon_n})\|^2 + \alpha_n^2\mathbb{E}\left[\|\omega_n\|^2|\mathbb{F}_n\right] \\
&+ 2\alpha_n^2\varepsilon_n\langle u_n - u_{\varepsilon_n}, F(u_n) - F(u_{\varepsilon_n})\rangle + 2\alpha_n\|u_n - u_{\varepsilon_n}\|\|\mathbb{E}\,[\omega_n|\mathbb{F}_n]\,\| \\
&+ 2\alpha_n^2\|F(u_n) - F(u_{\varepsilon_n})\|\|\mathbb{E}\,[\omega_n|\mathbb{F}_n]\,\| + 2\alpha_n^2\varepsilon_n\|u_n - u_{\varepsilon_n}\|\|\mathbb{E}\,[\omega_n|\mathbb{F}_n]\,\| \\
&\leq (1 - 2\alpha_n\varepsilon_n + 2\alpha_n^2)\|u_n - u_{\varepsilon_n}\|^2 + 2\alpha_n^2\|F(u_n) - F(u_{\varepsilon_n})\|^2 + \alpha_n^2\mathbb{E}\left[\|\omega_n\|^2|\mathbb{F}_n\right] \\
&+ 2\alpha_n\|u_n - u_{\varepsilon_n}\|\|\mathbb{E}\,[\omega_n|\mathbb{F}_n]\,\| + 2\alpha_n^2\|F(u_n) - F(u_{\varepsilon_n})\|\|\mathbb{E}\,[\omega_n|\mathbb{F}_n]\,\| \\
&+ 2\alpha_n^2\|u_n - u_{\varepsilon_n}\|\|\mathbb{E}\,[\omega_n|\mathbb{F}_n]\,\|.
\end{aligned} \tag{2.8}$$

To find bounds on the terms in (2.8), we begin by noticing that the sequence $\{u_{\varepsilon_n}\}$ is bounded, and hence there exists a constant $c_0 > 0$ such that $\|u_{\varepsilon_n}\| \leq c_0$, for every $n \in \mathbb{N}$. Therefore,

$$\|F(u_n) - F(u_{\varepsilon_n})\| \leq k_1(1 + \|u_n - u_{\varepsilon_n}\|), \tag{2.9}$$

where $k_1 := 2c(1 + c_0)$, and hence with $k_2 := 8c^2(1 + c_0)^2$, we obtain

$$\|F(u_n) - F(u_{\varepsilon_n})\|^2 \leq k_2(1 + \|u_n - u_{\varepsilon_n}\|^2). \tag{2.10}$$

Moreover, using the inequality $a \leq 1 + a^2$, which holds for every $a \in \mathbb{R}$, we can show that

$$2\alpha_n\|u_n - u_{\varepsilon_n}\|\|\mathbb{E}\,[\omega_n|\mathbb{F}_n]\,\| \leq c_1\alpha_n\beta_n(2 + k_1) + (2 + 2c + k_1)c_1\alpha_n\beta_n\|u_n - u_{\varepsilon_n}\|^2,$$

and hence by setting $k_3 := c_1(2 + 2c + k_1)$, we obtain

$$2\alpha_n\|u_n - u_{\varepsilon_n}\|\|\mathbb{E}\,[\omega_n|\mathbb{F}_n]\,\| \leq k_3\alpha_n\beta_n(1 + \|u_n - u_{\varepsilon_n}\|^2). \tag{2.11}$$

Analogously, we can show that

$$2\alpha_n^2\|u_n - u_{\varepsilon_n}\|\|\mathbb{E}\,[\omega_n|\mathbb{F}_n]\,\| \leq k_3\alpha_n^2\beta_n(1 + \|u_n - u_{\varepsilon_n}\|^2). \tag{2.12}$$

Furthermore, using (2.6) and (2.9), we obtain

$$2\alpha_n^2\|F(u_n)-F(u_{\varepsilon_n})\|\|\mathbb{E}[\omega_n|\mathbb{F}_n]\| \le 2k_1(2+k_1)\alpha_n^2\beta_n(1+\|u_n-u_{\varepsilon_n}\|^2),$$

and by taking $k_4 := 2k_1(2+k_1)$, we obtain

$$2\alpha_n^2\|F(u_n)-F(u_{\varepsilon_n})\|\|\mathbb{E}[\|\omega_n\||\mathbb{F}_n]\| \le k_4\alpha_n^2\beta_n(1+\|u_n-u_{\varepsilon_n}\|^2). \tag{2.13}$$

Finally, as in the proof of (2.9), we have

$$\alpha_n^2\mathbb{E}\left[\|\omega_n\|^2|\mathbb{F}_n\right] \le \alpha_n^2 c_2\left(1+\frac{\|F(u_n)^2\|}{\delta_n}\right) \le c_2\alpha_n^2+\frac{c_2k_1^2\alpha_n^2}{\delta_n}(1+\|u_n-u_{\varepsilon_n}\|^2),$$

and hence for a constant $k_5 := c_2\max\{1,k_1^2\}$, we obtain

$$\alpha_n^2\mathbb{E}\left[\|\omega_n\|^2|\mathbb{F}_n\right] \le k_5\alpha_n^2+\frac{k_5\alpha_n^2}{\delta_n}+\frac{k_5\alpha_n^2}{\delta_n}\|u_n-u_{\varepsilon_n}\|^2. \tag{2.14}$$

Consequently, due to (2.8), and the above inequalities, there is a constant $k>0$ such that

$$\begin{aligned}&\mathbb{E}[\|u_{n+1}-u_{\varepsilon_n}\|^2|\mathbb{F}_n]\\ &\le \left(1-2\alpha_n\varepsilon_n+k\alpha_n^2+k\alpha_n\beta_n+k\frac{\alpha_n^2}{\delta_n}\right)\|u_n-u_{\varepsilon_n}\|^2+k\alpha_n^2+k\alpha_n\beta_n+k\frac{\alpha_n^2}{\delta_n}.\end{aligned}$$

The above estimate, with the aid of the following inequality, which holds for all $a,b\in\mathbb{R}$,

$$(a+b)^2 \le (1+\alpha_n\varepsilon_n)a^2+\left(1+\frac{1}{\alpha_n\varepsilon_n}\right)b^2,$$

for the choices $a := \|u_n-u_{\varepsilon_{n-1}}\|$ and $b := \|u_{\varepsilon_n}-u_{\varepsilon_{n-1}}\|$, yields

$$\begin{aligned}&\mathbb{E}[\|u_{n+1}-u_{\varepsilon_n}\|^2|\mathbb{F}_n]\\ &\le \left(1-2\alpha_n\varepsilon_n+k\alpha_n^2+k\alpha_n\beta_n+k\frac{\alpha_n^2}{\delta_n}\right)(1+\alpha_n\varepsilon_n)\|u_n-u_{\varepsilon_{n-1}}\|^2\\ &+k\alpha_n^2+k\alpha_n\beta_n+k\frac{\alpha_n^2}{\delta_n}\\ &+\left(1-2\alpha_n\varepsilon_n+k\alpha_n^2+k\alpha_n\beta_n+k\frac{\alpha_n^2}{\delta_n}\right)\left(1+\frac{1}{\alpha_n\varepsilon_n}\right)\|u_{\varepsilon_n}-u_{\varepsilon_{n-1}}\|^2\|\bar{u}\|^2\\ &\le \left(1-\alpha_n\varepsilon_n+s\alpha_n^2+s\alpha_n\beta_n+s\frac{\alpha_n^2}{\delta_n}\right)\|u_n-u_{\varepsilon_{n-1}}\|^2\\ &+k\alpha_n^2+k\alpha_n\beta_n+k\frac{\alpha_n^2}{\delta_n}+s\left(\frac{1+\alpha_n\varepsilon_n}{\alpha_n\varepsilon_n}\right)\left|\frac{\varepsilon_{n-1}-\varepsilon_n}{\varepsilon_n}\right|^2,\end{aligned}$$

where we used Theorem 2.1, and s is a positive constant such that

$$\begin{aligned}s_1 &:= \sup_{n\in\mathbb{N}}\left(1+k\alpha_n^2+k\alpha_n\beta_n+k\frac{\alpha_n^2}{\delta_n}\right)\|\bar{u}\|^2,\\ s_2 &:= k\sup_n(1+\alpha_n\varepsilon_n),\\ s &:= \max\{s_1,s_2\}.\end{aligned}$$

Due to the summability condition on all the sequences involved, all the terms in sequences in the above term must converge to zero, and hence they remain bounded.

Therefore, we have

$$\mathbb{E}\left[\|u_{n+1}-u_{\varepsilon_n}\|^2|\mathbb{F}\right] \leq (1+t_n)\|u_n-u_{\varepsilon_{n-1}}\|^2-\gamma_n+\kappa_n,$$

where

$$t_n := s\alpha_n^2+s\alpha_n\beta_n+s\frac{\alpha_n^2}{\delta_n},$$

$$\kappa_n := k\alpha_n^2+k\alpha_n\beta_n+k\frac{\alpha_n^2}{\delta_n}+s\left(\frac{1+\alpha_n\varepsilon_n}{\alpha_n\varepsilon_n}\right)\left|\frac{\varepsilon_{n-1}-\varepsilon_n}{\varepsilon_n}\right|^2,$$

$$\gamma_n := \alpha_n\varepsilon_n\|u_n-u_{\varepsilon_{n-1}}\|^2.$$

Since $\{t_n\}$ and $\{\kappa_n\}$ generate summable series, Robbins and Siegmund Theorem 2.2, ensures that $\|u_n-u_{\varepsilon_{n-1}}\|$ converges almost surely, and $\sum_{n\in\mathbb{N}}\alpha_n\varepsilon_n\|u_n-u_{\varepsilon_{n-1}}\|^2<\infty$, which due to the divergence of the series $\sum_{n\in\mathbb{N}}\alpha_n\varepsilon_n$, implies that $\liminf_{n\to\infty}\|u_n-u_{\varepsilon_{n-1}}\|=0$, almost surely. □

We now give a variant of Theorem 2.3 under a different set of conditions on the random noise.

Theorem 2.4. *Let H be a real Hilbert space, $K\subset H$ be closed and convex, and $F:H\to H$ be monotone and hemicontinuous. Let $\{u_n\}$ be the sequence generated by (2.5), and let $\mathbb{F}_n$ be a filtration on the probability space $(\Omega,\mathbb{F},\mathbb{P})$ such that $\{u_n\}$ is $\mathbb{F}_n$-measurable. Let the solution set $\mathscr{S}(F,K)$ be nonempty. Assume that the following conditions hold:*

- **(A_1):** *There is a constant $c>0$ such that $\|F(u)\|\leq c(1+\|u\|)$, for every $u\in K$.*
- **(A_2):** *The random noise $\{\omega_n\}$ and the sequences $\{\varepsilon_n\}$ and $\{\alpha_n\}$ satisfy:*

$$\mathbb{E}\left[\omega_n|\mathbb{F}_n\right]=0.$$

$$\mathbb{E}\left[\|\omega_n\|^2|\mathbb{F}_n\right]\leq\beta(1+\|u_n\|^2),\ \beta>0.$$

$$\sum_{n\in\mathbb{N}}\varepsilon_n\alpha_n=\infty,\quad \sum_{n\in\mathbb{N}}\alpha_n^2<\infty,\quad \sum_{n\in\mathbb{N}}\left(\frac{1+\alpha_n\varepsilon_n}{\alpha_n\varepsilon_n}\right)\left|\frac{\varepsilon_{n-1}-\varepsilon_n}{\varepsilon_n}\right|^2<\infty.$$

Then, $\|u_{n+1}-u_{\varepsilon_n}\|\to 0$, almost sure.

Proof. As in the proof of Theorem 2.3, we have

$$\begin{aligned}\|u_{n+1}-u_{\varepsilon_n}\|^2 \leq{}& \|u_n-u_{\varepsilon_n}\|^2+\alpha_n^2\|F(u_n)-F(u_{\varepsilon_n})\|^2+\alpha_n^2\varepsilon_n^2\|u_n-u_{\varepsilon_n}\|^2\\ &+\alpha_n^2\|\omega_n\|^2-2\alpha_n\varepsilon_n\|u_n-u_{\varepsilon_n}\|^2-2\alpha_n\langle F(u_n)-F(u_{\varepsilon_n}),u_n-u_{\varepsilon_n}\rangle\\ &-2\alpha_n\langle u_n-u_{\varepsilon_n},\omega_n\rangle+2\alpha_n^2\langle F(u_n)-F(u_{\varepsilon_n}),\omega_n\rangle\\ &+2\alpha_n^2\varepsilon_n\langle u_n-u_{\varepsilon_n},\omega_n\rangle+2\alpha_n^2\varepsilon_n\langle F(u_n)-F(u_{\varepsilon_n}),u_n-u_{\varepsilon_n}\rangle.\end{aligned}$$

We take the expectation past $\mathbb{F}_n$ in the above estimate, use $\mathbb{E}\left[\omega_n|\mathbb{F}_n\right]=0$ and the monotonicity of F, and rearrange the terms to obtain

$$\begin{aligned}&\mathbb{E}\left[\|u_{n+1}-u_{\varepsilon_n}\|^2|\mathbb{F}_n\right]\\ &\leq(1-2\alpha_n\varepsilon_n+2\alpha_n^2)\|u_n-u_{\varepsilon_n}\|^2+2\alpha_n^2\|F(u_n)-F(u_{\varepsilon_n})\|^2+\alpha_n^2\mathbb{E}\left[\|\omega_n\|^2|\mathbb{F}_n\right],\end{aligned}$$

and by (A_2), (2.10), and a rearrangement of terms, we deduce that there is a constant $k>0$ with

$$\mathbb{E}\left[\|u_{n+1}-u_{\varepsilon_n}\|^2|\mathbb{F}_n\right]\leq\left(1-2\alpha_n\varepsilon_n+k\alpha_n^2\right)\|u_n-u_{\varepsilon_n}\|^2+k\alpha_n^2.$$

As in the proof of Theorem 2.3, for a constant $s > 0$, we have

$$\begin{aligned}&\mathbb{E}[\| u_{n+1} - u_{\varepsilon_n}\|^2 | \mathbb{F}_n] \\ &\leq \left(1 - \alpha_n \varepsilon_n + s\alpha_n^2\right) \|u_n - u_{\varepsilon_{n-1}}\|^2 + k\alpha_n^2 + s\left(\frac{1+\alpha_n\varepsilon_n}{\alpha_n\varepsilon_n}\right)\left|\frac{\varepsilon_{n-1}-\varepsilon_n}{\varepsilon_n}\right|^2,\end{aligned}$$

which can be written as

$$\mathbb{E}\left[\|u_{n+1} - u_{\varepsilon_n}\|^2 | \mathbb{F}\right] \leq (1+t_n)\|u_n - u_{\varepsilon_{n-1}}\|^2 - \gamma_n + \kappa_n,$$

where

$$\begin{aligned} t_n &:= s\alpha_n^2, \\ \kappa_n &:= k\alpha_n^2 + s\left(\frac{1+\alpha_n\varepsilon_n}{\alpha_n\varepsilon_n}\right)\left|\frac{\varepsilon_{n-1}-\varepsilon_n}{\varepsilon_n}\right|^2, \\ \gamma_n &:= \alpha_n\varepsilon_n\|u_n - u_{\varepsilon_{n-1}}\|^2, \end{aligned}$$

and then the conclusion that $\liminf_{n\to\infty} \|u_n - u_{\varepsilon_{n-1}}\| = 0$, almost surely, ensues, completing the proof. □

Remark 2.1. Iterative regularization has been extensively used for deterministic variational inequalities. Koshal, Nedić, and Shanbhag [38] used iterative regularization for finite-dimensional stochastic variational inequalities with a Lipschitz continuous map. However, our results cannot be derived from their results even in the finite-dimensional setting due to more general assumptions on the noise used here.

3. Regularized Stochastic Gradient for Estimating Random Parameters

3.1. **Function spaces.** Given the domain D, for $1 \leq p < \infty$, by $L^p(D)$, we represent the space of pth Lebesgue integrable functions, that is,

$$L^p(D) = \left\{ y : D \mapsto \mathbb{R} \text{ is measurable, and } \int_D |y|^p\, dx < +\infty \right\}.$$

The space $L^\infty(D)$ consists of measurable functions that are bounded almost everywhere (a.e.) on D. We also recall that the Sobolev spaces are given by

$$\begin{aligned} H^1(D) &= \left\{ y \in L^2(D),\ \ \partial_{x_i} y \in L^2(D),\ i = 1, \ldots, n \right\}, \\ H_0^1(D) &= \left\{ y \in H^1(D),\ \ y|_{\partial D} = 0 \right\}, \end{aligned}$$

and $H^{-1}(D) = (H_0^1(D))^*$ is the topological dual of $H_0^1(D)$. For $m \in \mathbb{N}$, higher-order Sobolev spaces $H^m(D)$ consist of $L^2(D)$ functions with all partial derivatives up to order m in $L^2(D)$.

Bochner spaces of random variables provide a convenient functional framework to study variational problems emerging from stochastic PDEs, see [39]. Given a real Banach space X, a probability space $(\Omega, \mathbb{F}, \mathbb{P})$, and an integer $p \in [1, \infty)$, the Bochner space $L^p(\Omega, X)$ consists of Bochner integrable functions $u : \Omega \to X$ with finite p-th moment, that is,

$$\|u\|_{L^p(\Omega,X)} := \left(\int_\Omega \|u(\omega)\|_X^p d\mathbb{P}(\omega)\right)^{1/p} = \mathbb{E}\left[\|u(\omega)\|_X^p\right]^{1/p} < \infty.$$

If $p = \infty$, then $L^\infty(\Omega, X)$ is the space of Bochner measurable functions $u : \Omega \to X$ such that

$$\operatorname{ess\,sup}_{\omega\in\Omega} \|u(\omega)\|_X < \infty.$$

3.2. **Inverse problem as a stochastic optimization problem.** There are two natural variational formulations for (1.1). The so-called path-wise formulation seeks, for a given realization $\omega \in \Omega$, $u_a(\cdot,\omega) \in H_0^1(D)$ such that

$$\int_D a(\omega,x)\nabla u_a(x,\omega)\cdot\nabla v(x)dx = \int_D f(\omega,x)v(x)\,dx, \text{ for all } v \in H_0^1(D). \tag{3.1}$$

The second variational formulation of (1.1), which is commonly referred to as the so-called integral formulation, seeks $u_a \in L^2(\Omega, H_0^1(D))$ such that for every $v \in L^2(\Omega, H_0^1(D))$, we have

$$\int_\Omega\int_D a(\omega,x)\nabla u_a(\omega,x)\cdot\nabla v(\omega,x)dx\,d\mathbb{P} = \int_\Omega\int_D f(\omega,x)v(\omega,x)\,dx\,d\mathbb{P}. \tag{3.2}$$

The two formulations are equivalent in a certain sense. From a numerical viewpoint, variational form (3.2) permits discretization over $\Omega\times D$, and the two most commonly used methods are the Stochastic Galerkin method and the Stochastic collocation method, see [1, 2, 40]. On the other hand, variational form (3.1) has been the basis of extensively studied Monte-Carlo type methods, which act on the principle of generating realizations of the solution $u(\omega,\cdot)$ from the realizations of the data, see [41]. The use of the stochastic approximation is also based on the realizations and revolves round (3.1). In the following, we state some results in the integral formulation, but they can be equivalently written in the path-wise formulation.

In the following, we will assume that there are constants k_0 and k_1 such that

$$0 < k_0 \le a(\omega,x) \le k_1 < \infty, \text{ almost everywhere in } \Omega\times D. \tag{3.3}$$

In particular, $a \in L^\infty(\Omega\times D)$.

The analytical properties of the parameter-to-solution map $a \mapsto u_a(\omega,x)$, that assigns to a, the unique solution $u_a(\omega,x)$ of (3.2) are of significant importance in the study of inverse problem. For this, let $K \subset L^\infty(\Omega; L^\infty(D))$ be the set of feasible parameters with a nonempty interior.

We will focus on the energy least-squares (ELS) functional recently proposed in [42]:

$$J(a) = \frac{1}{2}\mathbb{E}\left[\int_D a(\omega,x)\nabla(u_a(\omega,x)-z(\omega,x))\cdot\nabla(u_a(\omega,x)-z(\omega,x))dx\right], \tag{3.4}$$

where $u_a(\omega,x)$ is the solution of (3.2) for $a(\omega,x)$ and $z(\omega,x) \in L^2(\Omega, H_0^1(D))$ is the data.

The following result summarizes a derivative characterization of the parameter-to-solution map and useful properties of the above ELS objective:

Theorem 3.1. [42] *For each $a(\omega,x)$ in the interior of K, the map $a(\omega,x) \to u_a(\omega,x)$ is differentiable at $a(\omega,x)$. The derivative $\delta u_a := Du_a(\delta a)$ of $u_a(\omega,x)$ at $a(\omega,x)$ in the direction $\delta a(\omega,x)$ is the unique solution of the variational problem: Find $\delta u_a(\omega,x) \in V$ such that for every $v(\omega,x) \in V$, we have*

$$\mathbb{E}\left[\int_D a(\omega,x)\nabla\delta u_a(\omega,x)\cdot\nabla v(\omega,x)dx\right] = -\mathbb{E}\left[\int_D \delta a(\omega,x)\nabla u_a(\omega,x)\cdot\nabla v(\omega,x)\,dx\right]. \tag{3.5}$$

Moreover, the ELS functional given in (3.4) is convex in the interior of the set K. Finally, the first derivative of J at a is given by

$$DJ(a)(\delta a) = -\frac{1}{2}\mathbb{E}\left[\int_D \delta a(\omega,x)\nabla(u_a(\omega,x)+z(\omega,x))\nabla(u_a(\omega,x)-z(\omega,x))\,dx\right].$$

Defining

$$\mathscr{J}(a,\omega) = \int_D a(\omega,x)\nabla(u_a(\omega,x)-z(\omega,x))\cdot\nabla(u_a(\omega,x)-z(\omega,x))\,dx,$$

we obtain

$$\nabla\mathscr{J}(a)(\delta a) = -\frac{1}{2}\int_D \delta a(\omega,x)\nabla(u_a(\omega,x)+z(\omega,x))\nabla(u_a(\omega,x)-z(\omega,x))\,dx,$$

and consequently

$$\nabla J(a) = \nabla\mathbb{E}[\mathscr{J}(a,\omega)] = \mathbb{E}[\nabla\mathscr{J}(a,\omega)].$$

The inverse problem of identifying stochastic parameters in partial differential equations is ill-posed, and some form of regularization is essential for achieving a stable identification process. For this, we tailor a general setting by defining the following admissible set:

$$K := \left\{a \in \mathscr{H} := L^2(\Omega,H):\ 0 < k_0 \le a(\omega,x) \le k_1 \text{ a.s. } \Omega\times D\right\},$$

where $\mathscr{H}$ is a separable Hilbert space compactly embedded into $B := L^\infty(\Omega, L^\infty(\Omega))$, and H is continuously embedded in $L^\infty(\Omega)$.

We consider the following regularized energy least-squares

$$\min_{a\in K} J_{\varepsilon_n}(a) := \frac{1}{2}\mathbb{E}\left[\int_D a(\omega,x)\nabla(u(\omega,x)-z)\cdot\nabla(u(\omega,x)-z)dx + \frac{\varepsilon_n}{2}\|a(\omega,x)\|_H^2\right], \tag{3.6}$$

where $u_a(\omega,x)$ is the solution of (3.2) for $a(\omega,x)$, $z(\omega,x)\in L^2(\Omega,L^2(D))$ is the data, $\varepsilon_n > 0$ is a fixed regularization parameter, and $\|\cdot\|_{\mathscr{H}}^2$ is the regularizer. Optimization problem (3.6) is uniquely solvable, see [42]. The norm $\|\cdot\|_{\mathscr{H}}^2$ includes the expectation of the norm $\|\cdot\|_H^2$.

A fruitful technique in the study of stochastic PDEs and stochastic optimization problems is the finite-dimensional noise representation of random fields appearing as the parameter, source term, or control variable, see [39]. A random field $v \in L^2(\Omega,L^2(D))$ is called a finite-dimensional noise if it has the form $v(x,\xi(\omega))$, $x\in D$ and $\omega\in\Omega$, where $\xi = (\xi_1,\xi_2,\dots,\xi_M)$: $\Omega\to\Gamma := \Gamma_1\times\Gamma_2\cdots\times\Gamma_M \subset \mathbb{R}^M$ and for $k = 1,\dots,m$, $\xi_k:\Omega\to\Gamma_k$, are real-valued random variables. At the core of the finite-dimensional noise property is the fact that it represents a random field expressed into a finite number of real-valued random variables. This process permits a change of variables, recasting a stochastic PDE, posed in the space $D\times\Omega$, into a deterministic albeit high-dimensional parametric PDE, posed in the space $D\times\Gamma$.

In the stochastic approximation approach developed here, we assume that the random fields are finite-dimensional noise; however, performing a change of variable is not required. We assume that the unknown random field is expressed as a finite linear expansion of the form:

$$a(\omega,x) = a_0(x) + \sum_{i=1}^{M} a_t(x)Y_t(\omega) = \sum_{t=0}^{M} a_t(x)Y_t(\omega), \tag{3.7}$$

where $Y_t(\omega)$ are random variables for $t = 0,\dots,M$, and by convention, $Y_0(\omega) = 1$, see [39].

We note that under (3.7), the variational problem reads: Find $u_a \in L^2(\Omega,H_0^1(D))$ such that for every $v\in L^2(\Omega,H_0^1(D))$, we have

$$\mathbb{E}\left[\int_D\left(\sum_{t=0}^{M} a_t(x)Y_t(\omega)\right)\nabla u(\omega,x)\cdot\nabla v(\omega,x)dx\right] = \mathbb{E}\left[\int_D f(\omega,x)v(\omega,x)\,dx\right]. \tag{3.8}$$

Moreover,

$$J(a)=\mathbb{E}[J(\omega,a)],$$

where

$$J(\omega,a)=\frac{1}{2}\sum_{t=0}^{M}Y_t(\omega)\int_D a_t(x)\nabla(u_a(\omega,x)-z(\omega,x))\cdot\nabla(u_a(\omega,x)-z(\omega,x))dx.$$

Furthermore, it can be shown that

$$DJ(a)(\delta a)=\mathbb{E}\left[T(\omega,a)(\delta a)\right],$$

where

$$T(\omega,a)(\delta a):=\sum_{t=0}^{M}\left[-\frac{1}{2}\int_D \delta a_t(x)\nabla(u_a(\omega,x)+z(\omega,x))\nabla(u_a(\omega,x)-z(\omega,x))\,dx\right]Y_t(\omega).$$

Consequently,

$$\nabla_a J(a)=\nabla_a\mathbb{E}\left[J(\omega,a)\right]=\mathbb{E}\left[\nabla_a J(\omega,a)\right],$$
$$\nabla_a J(\omega,a)=\nabla_a T(\omega,a)=\left(\frac{\partial T}{\partial a_t}(\omega,a)\right)_{t=0,\ldots,M},$$

where, for $t=0,...,M$, we have

$$\frac{\partial T}{\partial a_t}(\omega,a)(\cdot)=\left[-\frac{1}{2}\int_D(\cdot)\nabla(u_a(\omega,x)+z(\omega,x))\nabla(u_a(\omega,x)-z(\omega,x))\,dx\right]Y_t(\omega) \qquad (3.9)$$

3.3. **Finite element discretization.** We will use a finite element discretization of the spaces $V=H_0^1(D)$ and H. For this, we define a triangulation (or finite element mesh) $\mathscr{T}_h$ on D. Let V_h and H_h be the spaces of piecewise linear continuous polynomials relative to $\mathscr{T}_h$, and let $\{\phi_1,\phi_2,\ldots,\phi_m\}$ and $\{\varphi_1,\varphi_2,\ldots,\varphi_l\}$ be the corresponding bases, respectively.

Given a realization of $\omega\in\Omega$, the discrete pathwise formulation (3.1) seeks $U=U(\omega,A)\in\mathbb{R}^m$ by solving

$$K(A)U(\omega,A)=F(\omega),$$

where A shows the dependence of the random parameter $a(\omega,x)$, and $K(A)\in\mathbb{R}^{m\times m}$ and $F(\omega_n)\in\mathbb{R}^m$ are the stiffness matrix and the load vector defined by

$$K(A)_{i,j}=\int_D a(\omega,x)\nabla\phi_j\cdot\nabla\phi_i dx,\quad \text{for } i,j=1,\ldots,m,$$
$$F(\omega)_i=\int_D f(\omega,x)\phi_i dx,\quad \text{for } i=1,\ldots,m.$$

Due to the expansion

$$a(\omega,x)=a_0(x)+\sum_{i=1}^{M}a_t(x)Y_t(\omega)=\sum_{t=0}^{M}a_t(x)Y_t(\omega),$$

the stiffness matrix can be decomposed as follows:

$$\begin{aligned} K(A(\omega))_{i,j} &= \int_D a(\omega,x)\nabla\phi_j\cdot\nabla\phi_i dx, \quad \text{for } i,j=1,\dots,m, \\ &= \int_D \left(\sum_{t=0}^M a_t(x)Y_t(\omega)\right)\nabla\phi_j\cdot\nabla\phi_i dx, \quad \text{for } i,j=1,\dots,m, \\ &= \sum_{t=0}^M Y_t(\omega)\int_D a_t(x)\nabla\phi_j\cdot\nabla\phi_i dx, \quad \text{for } i,j=1,\dots,m, \\ &= \sum_{t=0}^M Y_t(\omega)K(A_t), \end{aligned} \tag{3.10}$$

where $Y_0(\omega)=1$ by convention, and

$$K(A_t)_{i,j} := \int_D a_t(x)\nabla\phi_j\cdot\nabla\phi_i dx, \quad \text{for } i,j=1,\dots,m.$$

In the following, denoting by A_t the discrete version of $a_t(x)$, we run the stochastic approximation scheme in terms of the spatial components:

$$A^{(n)} = (A_0^{(n)},\dots,A_M^{(n)}).$$

By using standard finite element discretization arguments, from (3.9), we get

$$\frac{\partial T}{\partial A_t}(\omega,A) = -\frac{1}{2}\mathbb{L}(U(\omega,A)+Z(\omega))^T\left(U(\omega,A)-Z(\omega)\right)Y_t(\omega), \text{ for } t=0,\dots,M,$$

where $\mathbb{L}\in\mathbb{R}^{m\times l}$ is the so-called adjoint stiffness matrix defined by the condition that

$$\mathbb{L}(V)A = K(A)V, \text{ for every } V\in\mathbb{R}^m,\ A\in\mathbb{R}^l.$$

Given the random vector $(Y_0(\bar\omega),\dots,Y_M(\bar\omega))$ by $A^{(n)}(\bar\omega)$, we denote the corresponding realization

$$A^{(n)}(\bar\omega) = \sum_{t=0}^M A_t^{(n)}Y_t(\bar\omega).$$

In view of the above preparation, we propose the following for the inverse problem:

1: Choose an initial guess $A^{(0)}$, step-lengths $\{\alpha_n\}$, regularization parameters $\{\varepsilon_n\}$, the sample rate $\{s_n\}\subset\mathbb{N}$, and initial samples $\{\omega_j^0\}_{j=1}^{s_0}$ of the random variable ω.
2: Generate random $(1,y_1(\omega_n),\dots,y_m(\omega_n))\in\Upsilon$.
3: Compute $U^{(n)}=U(\omega_n,A^{(n)})$ by solving the following system:

$$K(A^{(n)}(\omega_n))U^{(n)} = F(\omega_n).$$

4: Given $A_n\in K$, generate samples $\{\omega_j^n\}_{j=1}^{s_n}$ of ω and compute $A^{(n+1)}\in K$ by the following

$$A_{n+1} = P_K\left[A_n - \frac{\alpha_n}{s_n}\sum_{j=1}^{s_n} G_{\varepsilon_n}\left(\omega_j^n, A_n\right)\right], \tag{3.11}$$

where G_{ε_n} is the discrete variant of gradient of the regularized ELS objective given by

$$G_{\varepsilon_n}(\omega_n, A^{(n)}) = \begin{pmatrix} -\frac{1}{2}\mathbb{L}(U^{(n)} + Z(\omega_n))^T \left(U^{(n)} - Z(\omega_n)\right) \\ -\frac{1}{2}\mathbb{L}(U^{(n)} + Z(\omega_n))^T \left(U^{(n)} - Z(\omega_n)\right) y_1(\omega_n) \\ \vdots \\ -\frac{1}{2}\mathbb{L}(U^{(n)} + Z(\omega_n))^T \left(U^{(n)} - Z(\omega_n)\right) y_m(\omega_n) \end{pmatrix}$$
$$+ \varepsilon_n \begin{pmatrix} K_A & 0 & \cdots & 0 \\ 0 & K_A & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & K_A \end{pmatrix}.$$

Remark 3.1. In the above algorithm, we modify the classical stochastic gradient scheme by sampling the regularized gradient at a given sampling rate. Moreover, the regularizer corresponds to the $H_0^1(\Omega)$ semi-norm, but can easily be replaced by any other regularizer.

4. A Computational Example

In this section, we present results of a numerical experiment where $M = 2$ and the coefficient a to be identified and u (solution of the direct problem) are given by:

$$a(\omega, x) = 1 + 3x^3 + Y_1(\omega)\cos\left(\frac{\pi}{4}x\right) + Y_2(\omega)\cos\left(\frac{\pi}{2}x\right),$$
$$u(\omega, x) = x(1 - x^2) + Y_1(\omega)\sin(2\pi x) + Y_2(\omega)\sin(4\pi x).$$

Spatial domain Ω is $[0,1]$ and both $Y_1(\omega)$ and $Y_2(\omega)$ are uniformly distributed on the interval $[0,1]$. Coefficients $a_t(x)$ are components of a vector function $\tilde{a}$ that has 3 components, that is, we have $\tilde{a} = [a_0(x), a_1(x), a_2(x)]^\top$ and $a(\omega, x) = a_0(x) + a_1(x)Y_1(\omega) + a_2(x)Y_2(\omega)$. Once the problem is discretized, we look for a solution vector $A = [A_0, A_1, A_2]^\top \in \mathbb{R}^{3l}$ which is a finite-dimensional approximation $\tilde{a}$. We use subscript h to indicate the approximation $a_h(\omega, x) = a_{0,h}(x) + a_{1,h}(x)Y_1(\omega) + a_{1,h}(x)Y_1(\omega)$ on uniform mesh with size h. Each component $a_t(x)$ of the vector function $\tilde{a}$ is represented by an l-vector A_t where l is the number of nodes in the mesh. We also choose a constraint set

$$\{A \in \mathbb{R}^{3l} | A_{\text{lower}} \le A \le A_{\text{upper}}\},$$

where A_{lower} and A_{upper} are (constant) vectors containing lower and upper bounds for the components $a_t(x)$, $t = 0, 1, 2$. Piecewise linear basis functions were used to represent both $u(\omega, x)$ and the components $a_t(x)$, and the nodal interpolant of the the direct problem solution u is taken as the data. Results of a typical simulation are shown in Figures 1 and 2. The top row of the Figure 1 shows some realizations of the exact coefficients $a(\omega, x)$ and the exact solution $u(\omega, x)$. Plots in the bottom row show realizations of the identified coefficient $a_h(\omega, x)$ and the corresponding simulated solution $u_h(\omega, x)$. In Figure 2, we show the comparison of the mean of the exact coefficient $a(\omega, x)$ which is

$$\mathbb{E}[a(\omega, x)] = 1 + 3x^3 + \frac{1}{2}\cos\left(\frac{\pi}{4}x\right) + \frac{1}{2}\cos\left(\frac{\pi}{2}x\right),$$

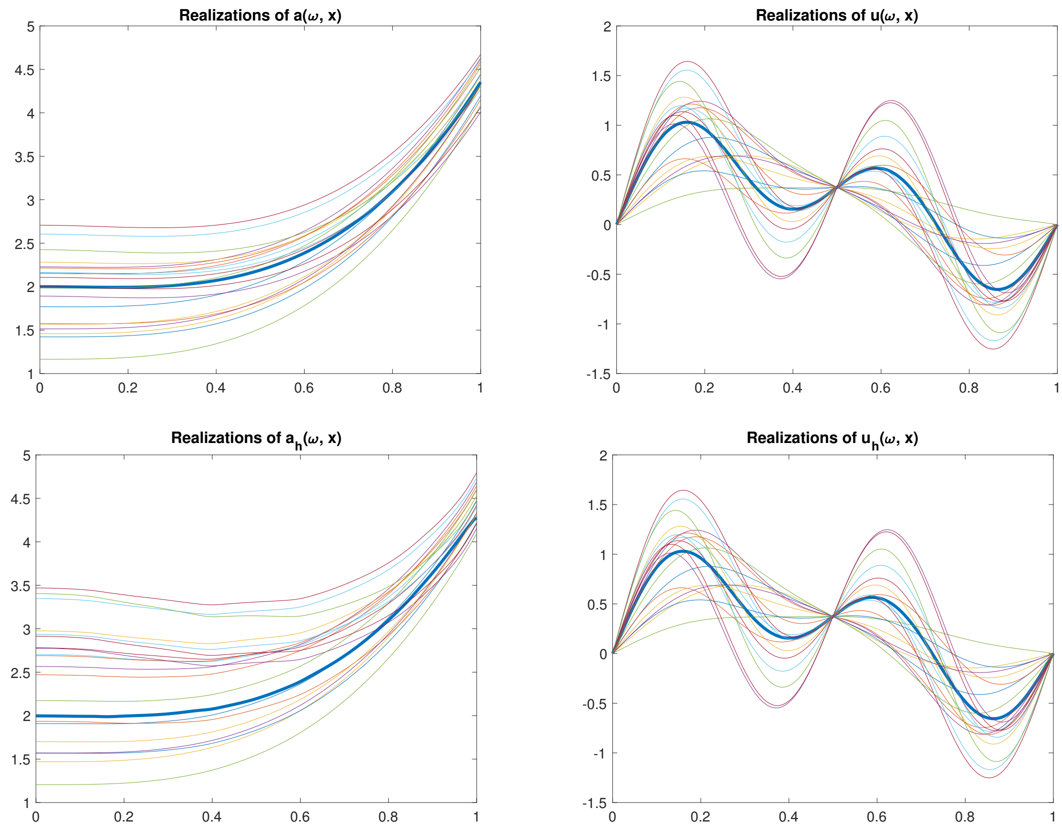

FIGURE 1. A comparison of 20 realizations of the exact coefficient a, the identified coefficient a_h, exact solution u, and the simulated solution $u_h = u_h(a_h)$. Mesh size used in the computations is $h = 1/160$. The thick blue lines in each plot represent the means.

and the estimated coefficient a_h which is given by

$$\mathbb{E}[a_h(\omega,x)] = A_0 + \frac{1}{2}A_1 + \frac{1}{2}A_2.$$

As we see from the figure, the identification is very stable in one dimensional computations and the estimated mean matches the exact mean very closely. This is observed for all computations with reasonably small mesh sizes such as $h = 1/100$ and $h = 1/160$ (this one is used in producing the plots shown in Figures 1 and 2). We plan to conduct more numerical studies with problems in higher space dimensions as well as a study of effects of a noise in the data in a future work.

5. CONCLUDING REMARKS

We studied the nonlinear inverse problem of estimating the stochasticity of a random parameter in stochastic partial differential equations. One of the main contributions is a new iteratively regularized projected stochastic gradient scheme for a general variational inequality. We reformulated the nonlinear inverse problem of parameter identification as a stochastic convex optimization problem. Therefore, the necessary and sufficient optimality condition is a variational

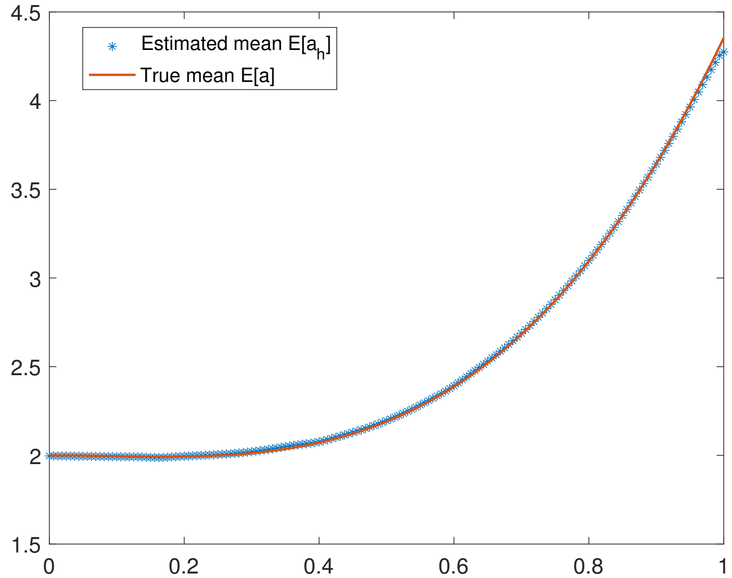

FIGURE 2. Comparison of the means $\mathbb{E}[a(\omega,x)]$ and $\mathbb{E}[a_h(\omega,x)]$.

inequality, and the developed framework is readily applicable. In the unknown parameter, we separate the stochastic and the deterministic components by employing a Karhunen-Loeve type expansion. Since the stochastic approximation framework samples the uncertainty, we identify the deterministic components in the expansion of the parameter. The preliminary numerical results are quite encouraging, and they show the feasibility of a stochastic approximation approach for the estimation of random parameters. It would be of genuine interest to conduct a more detailed numerical comparison of the available techniques for solving stochastic inverse problems, such as the stochastic Galerkin method and the stochastic collocation method.

Acknowledgments

The research of Baasansuren Jadamba, Akhtar Khan, and Yidan Yang was supported by the National Science Foundation under Award No. 1720067. Miguel Sama was supported by Ministry for Science and Innovation by the Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación (Spain) under project PID2020-112491GB-I00 / AEI / 10.13039/501100011033 and grant 2021-MAT11 (ETSI Industriales, UNED).

REFERENCES

[1] J. Bäck, F. Nobile, L. Tamellini, R. Tempone, Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: a numerical comparison, In: Spectral and High Order Methods for Partial Differential Equations, vol. 76 of Lect. Notes Comput. Sci. Eng., pp. 43-62, Springer, Heidelberg, 2011.

[2] I. Babuška, R. Tempone, G. E. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, SIAM J. Numer. Anal. 42 (2004), 800-825.

[3] H. Robbins, S. Monro, A stochastic approximation method, Ann. Math. Statistics 22 (1951), 400-407.

[4] K. Barty, J.-S. Roy, C. Strugarek, Hilbert-valued perturbed subgradient algorithms, Math. Oper. Res. 32 (2007), 551-562.

[5] D.P. Bertsekas, J.N. Tsitsiklis, Gradient convergence in gradient methods with errors, SIAM J. Optim. 10 (2000), 627-642.
[6] J.-C. Culioli, G. Cohen, Decomposition/coordination algorithms in stochastic optimization, SIAM J. Control Optim. 28 (1990), 1372-1403.
[7] H. J. Kushner, A. Shwartz, Stochastic approximation in Hilbert space: identification and optimization of linear continuous parameter systems, SIAM J. Control Optim. 23 (5) (1985), 774-793.
[8] G. Yin, Y. M. Zhu, On H-valued Robbins-Monro processes, J. Multivariate Anal. 34 (1990), 116–140.
[9] R.I. Boţ, A. Bö, Variable smoothing for convex optimization problems using stochastic gradients, J. Sci. Comput. 85 (2020), 33.
[10] A. Dieuleveut, A. Durmus, F. Bach, Bridging the gap between constant step size stochastic gradient descent and Markov chains, Ann. Statist. 48 (2020), 1348-1382.
[11] A.N. Iusem, A. Jofré, R.I. Oliveira, P. Thompson, Extragradient method with variance reduction for stochastic variational inequalities, SIAM J. Optim. 27 (2017), 686-724.
[12] A.N. Iusem, A. Jofré, R.I. Oliveira, P. Thompson, Variance-based extragradient methods with line search for stochastic variational inequalities, SIAM J. Optim. 29 (2019), 175-206.
[13] P. Mertikopoulos, M. Staudigl, Stochastic mirror descent dynamics and their convergence in monotone variational inequalities, J. Optim. Theory Appl. 179 (3) (2018) 838–867.
[14] K. Scaman, F. Bach, S. Bubeck, Y.T. Lee, L. Massoulié, Optimal convergence rates for convex distributed optimization in networks, J. Mach. Learn. Res. 20 (2019), 159.
[15] D. Scieur, A. d'Aspremont, F. Bach, Regularized nonlinear acceleration, Math. Program. Ser. A 179 (2020), 47-83.
[16] F.Q. Xia, Q. H. Ansari, J.C. Yao, A new incremental constraint projection method for solving monotone variational inequalities, Optim. Methods Softw. 32 (2017), 470–502.
[17] X.-J. Zhang, X.-W. Du, Z.-P. Yang, G.-H. Lin, An infeasible stochastic approximation and projection algorithm for stochastic variational inequalities, J. Optim. Theory Appl. 183 (2019), 1053–1076.
[18] J. Zhu, J.C. Spall, Stochastic approximation with nondecaying gain: error bound and data-driven gain-tuning, Internat. J. Robust Nonlinear Control 30 (2020), 5820–5870.
[19] R. Aboulaich, N. Fikal, E. El Guarmah, N. Zemzemi, Stochastic finite element method for torso conductivity uncertainties quantification in electrocardiography inverse problem, Math. Model. Nat. Phenom. 11 (2016), 1-19.
[20] V. A. Badri Narayanan, N. Zabaras, Stochastic inverse heat conduction using a spectral approach, Internat. J. Numer. Methods Engrg. 60 (2004), 1569–1593.
[21] M. Heinkenschloss, B. Kramer, T. Takhtaganov, Adaptive reduced-order model construction for conditional value-at-risk estimation, SIAM/ASA J. Uncertain. Quantif. 8 (2020), 668–692.
[22] M. Morzfeld, X. Tu, J. Wilkening, A.J. Chorin, Parameter estimation by implicit sampling, Commun. Appl. Math. Comput. Sci. 10 (2015), 205–225.
[23] B.V. Rosić, H.G. Matthies, Identification of properties of stochastic elastoplastic systems, In: Computational Methods in Stochastic Dynamics, vol. 2, pp. 237–253, Springer, Dordrecht, 2013.
[24] N. Zabaras, B. Ganapathysubramanian, A scalable framework for the solution of stochastic inverse problems using a sparse grid collocation approach, J. Comput. Phys. 227 (2008), 4697-4735.
[25] N. Cahill, B. Jadamba, A.A. Khan, M. Sama, B. Winkler, A first-order adjoint and a second-order hybrid method for an energy output least squares elastography inverse problem of identifying tumor location, Boundary Value Probl. 263 (2013), 1–14.
[26] E. Crossen, M.S. Gockenbach, B. Jadamba, A.A. Khan, B. Winkler, An equation error approach for the elasticity imaging inverse problem for predicting tumor location, Comput. Math. Appl. 67 (2014), 122–135.
[27] M.M. Doyley, B. Jadamba, A.A. Khan, M. Sama, B. Winkler, A new energy inversion for parameter identification in saddle point problems with an application to the elasticity imaging inverse problem of predicting tumor location, Numer. Funct. Anal. Optim. 35 (2014), 984–1017.
[28] A. Gibali, B. Jadamba, A.A. Khan, F. Raciti, B. Winkler, Gradient and extragradient methods for the elasticity imaging inverse problem using an equation error formulation: a comparative numerical study, In: Nonlinear Analysis and Optimization, vol. 659 of Contemp. Math., Amer. Math. Soc., pp. 65–89, Providence, RI, 2016.

[29] M.S. Gockenbach, A.A. Khan, An abstract framework for elliptic inverse problems: Part I. An output least-squares approach, Math. Mech. Solids 12 (2007), 259–276.
[30] M.S. Gockenbach, A.A. Khan, An abstract framework for elliptic inverse problems. II. An augmented Lagrangian approach, Math. Mech. Solids 14 (2009), 517–539.
[31] B. Jadamba, A.A. Khan, A. Oberai, M. Sama, First-order and second-order adjoint methods for parameter identification problems with an application to the elasticity imaging inverse problem, Inverse Probl. Sci. Eng. 25 (2017), 1768-1787.
[32] B. Jadamba, A.A. Khan, G. Rus, M. Sama, B. Winkler, A new convex inversion framework for parameter identification in saddle point problems with an application to the elasticity imaging inverse problem of predicting tumor location, SIAM J. Appl. Math. 74 (2014), 1486–1510.
[33] R. Hawks, B. Jadamba, A.A. Khan, M. Sama, Y. Yang, A variational inequality based Stochastic approximation for inverse problems in stochastic partial differential equations, In: Nonlinear Analysis and Global Optimization, vol. 167 of Springer Optim. Appl., pp. 207–226, Springer, Cham, 2021.
[34] C. Geiersbach, G.C. Pflug, Projected stochastic gradients for convex constrained problems in Hilbert spaces, SIAM J. Optim. 29 (2019), 2079–2099.
[35] M. Martin, S. Krumschield, F. Nobile, Analysis of stochastic gradient methods for pde-constrained optimal control problems with uncertain parameters, MATHICSE Technical Report, (2018), Nr. 04.2018.
[36] Y. Alber, I. Ryazantseva, Nonlinear ill-posed problems of monotone type, Springer, Dordrecht, 2006.
[37] H. Robbins, D. Siegmund, A convergence theorem for non negative almost supermartingales and some applications, (1971), 233–257. https://doi.org/10.1016/B978-0-12-604550-5.50015-8.
[38] J. Koshal, A. Nedić, U.V. Shanbhag, Regularized iterative stochastic approximation methods for stochastic variational inequality problems, IEEE Trans. Automat. Control 58 (2013), 594–609.
[39] G.J. Lord, C. E. Powell, T. Shardlow, An introduction to computational stochastic PDEs, Cambridge Texts in Applied Mathematics, Cambridge University Press, New York, 2014.
[40] I. Babuška, F. Nobile, R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, SIAM J. Numer. Anal. 45 (2007), 1005–1034.
[41] A. Barth, C. Schwab, N. Zollinger, Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients, Numer. Math. 119 (2011), 123–161.
[42] B. Jadamba, A.A. Khan, M. Sama, C. Tammer, H.-J. Starkloff, A convex optimization framework for the inverse problem of identifying a random parameter in a stochastic partial differential equation, SIAM J. Uncertain. Quantif. 9 (2021), 922–952.