A Variational Inequality Based Stochastic Approximation for Inverse Problems in Stochastic Partial Differential Equations

Rachel Hawks, Baasansuren Jadamba, Akhtar A. Khan, Miguel Sama, and Yidan Yang

Abstract The primary objective of this work is to study the inverse problem of identifying a parameter in partial differential equations with random data. We explore the nonlinear inverse problem in a variational inequality framework. We propose a projected-gradient-type stochastic approximation scheme for general variational inequalities and give a complete convergence analysis under weaker conditions on the random noise than those commonly imposed in the available literature. The proposed iterative scheme is tested on the inverse problem of parameter identification. We provide a derivative characterization of the solution map, which is used in computing the derivative of the objective map. By employing a finite element based discretization scheme, we derive the discrete formulas necessary to test the developed stochastic approximation scheme. Preliminary numerical results show the efficacy of the developed framework.

2010 Mathematics Subject Classification 35R30, 49N45, 65J20, 65J22, 65M30

1 Introduction

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, that is, Ω is a nonempty set whose elements are termed as elementary events, $\mathscr{F}$ is a σ-algebra of subsets of Ω, and $\mathbb{P}$ a probability measure. Let $D \subset \mathbb{R}^n$ be a sufficiently smooth bounded domain and ∂D be the boundary of Ω. Given two random fields $a : \Omega \times D \mapsto R$ and $f : \Omega \times D \to \mathbb{R}$, we consider the stochastic partial differential equation (SPDE) of

R. Hawks · B. Jadamba · A. A. Khan (✉) · Y. Yang
School of Mathematical Sciences, Rochester Institute of Technology, Rochester, NY, USA
e-mail: rh9784@rit.edu; bxjsma@rit.edu; aaksma@rit.edu; yy2513@rit.edu

M. Sama
Departamento de Matemática Aplicada, Universidad Nacional de Educación a Distancia, Madrid, Spain
e-mail: msama@ind.uned.es

T. M. Rassias, P. M. Pardalos (eds.), *Nonlinear Analysis and Global Optimization*, Springer Optimization and Its Applications 167,
https://doi.org/10.1007/978-3-030-61732-5_9

finding a random field $u : \Omega \times D \to \mathbb{R}$ that almost surely satisfies the following:

$$-\nabla \cdot (a(\omega, x)\nabla u(\omega, x)) = f(\omega, x), \text{ in } D, \tag{1a}$$

$$u(\omega, x) = 0, \text{ on } \partial D. \tag{1b}$$

The above SPDE models interesting real-world phenomena and has been studied in great detail in the deterministic setting. For example, in (1), u may represent the steady-state temperature at a given point of a body; then a would be a variable thermal conductivity coefficient, and f an external heat source. The system (1) also models underground steady-state aquifers in which the parameter a is the aquifer transmissivity coefficient, u is the hydraulic head, and f is the recharge. Some details on inverse problems can be found in [6–8, 10, 12, 16, 18, 20–22].

A natural interpretation of (1) is that realizations of the data lead to deterministic PDEs. That is, for a fixed $\omega \in \Omega$, SPDE (1), under appropriate conditions, admits a weak solution $u(\omega, \cdot) \in H_0^1(D)$.

The objective of this work is to study the nonlinear inverse problem of identifying the parameter a from a measurement of the solution u of (1) by solving a stochastic optimization problem of the following form:

$$\min_{a \in \mathbb{A}} \mathbb{J}(a) := \mathbb{E}\left[J(a, \omega)\right]. \tag{2}$$

Here $\mathbb{A}$ is the set of feasible parameters, which is a subset of a real Hilbert space H, $J(a, \omega)$ is a misfit function, which we will define shortly, and $\mathbb{E}$ is the expectation with respect to the probability measure.

If the expected value $\mathbb{E}[J(a, \omega)]$ is readily assessable, either analytically or numerically, then (2) is practically a deterministic optimization problem that can be solved by a wide variety of available numerical methods. However, the evaluation of $\mathbb{E}[J(a, \omega)]$ is a challenging task. For instance, even when the random vector ω has a known probability distribution, the computation of the expected value $\mathbb{E}[J(a, \omega)]$ could involve computationally expensive multi-dimensional integral evaluations. A likely scenario is when the function $J(a, \omega)$ is known, but the distribution of ω is unknown, and any information on ω can only be gathered using sampling. Another challenging situation occurs when the expected value $\mathbb{E}[J(a, \omega)]$ is not observable, and it must be evaluated through a simulation process. In all such situations, the existing methods for deterministic optimization problems are not applicable and ought to be appropriately modified.

Our objective is to employ the stochastic approximation approach (SAA) in a Hilbert space setting for solving the nonlinear inverse problem of parameter identification in stochastic PDEs. The SAA has a long history and has been used for a wide variety of problems. On the other hand, SPDEs have also received a great deal of attention in recent years. To the best of our knowledge, this is the first time that the SAA approach is being used for nonlinear inverse problems. Before describing the main contributions and our strategy, let us briefly discuss the key ideas that have been used in these two disciplines.

During the past several decades, the dynamic field of stochastic approximation, initiated by the seminal paper by Robbins and Monro [32], witnessed an explosive growth in numerous directions. To give a brief review of some work relevant to this research, we note that Kiefer and Wolfowitz [25] extended the stochastic approximation approach to finding a unique minimum/maximum of the one-dimensional unknown regression function. An informative survey of the earlier developments in the stochastic approximation is by Lai [27]. Many authors have studied stochastic approximation in general space inspired by applications in control theory and related areas. A small sample of such contributions includes the research by Barty, Roy, and Strugarek [3], Goldstein [17], Kushner and Shwartz [26], Salov [35], Yin and Zhu [37], where more references can be found. Interesting related results have been given by Bertsekas and Tsitsiklis [5], Culioli and Cohen [9], and others.

On the other hand, the stochastic PDE-constrained optimization also attracted a great deal of attention in recent years. Such problems emerged from two sources, the inverse problems of parameter or source identification and optimal control problems. For example, Narayanan and Zabaras [2] studied the inverse problem in the presence of uncertainties in the material data and developed an adjoint approach based identification process using the spectral stochastic finite element method. In [38], the authors developed a scalable methodology for the stochastic inverse problem using a sparse grid collocation approach. In [36], the authors developed a robust and efficient method by employing generalized polynomial chaos expansion to identifying uncertain elastic parameters from experimental modal data. In [30], the authors presented an implicit sampling for parameter identification. In [34], the authors studied the parameter identification in a Bayesian setting for the elastoplastic problem. In [31], the authors explored the optimal control problem for the stochastic diffusion equation. In [24], the authors focused on determining the optimal thickness subjected to stochastic force. In [1], the authors investigated the impact of errors and uncertainties of the conductivity on the electrocardiography imaging solution.

Since the stochastic approximation approach is designed for problems with either noisy experimental values or noisy samples of the function, it seems ideal for solving inverse and ill-posed problems. However, the use of the stochastic approximation approach is mostly non-existent. Note that Bertran [4], who studied a stochastic projected gradient algorithm in a Hilbert space, gave an application to optimal control problems where the data was uncertain. A formal study of the stochastic approximation approach for optimal control in stochastic PDEs was initiated independently by Martin, Krumschield, and Nobile [29] and Geiersbach and Pflug [11]. Since the control-to-state map is linear, these problems involve a convex objective function. On the other hand, the inverse problem we consider in the present work is nonlinear, and the commonly used output least-squares (OLS) objective functional is nonconvex, in general. Therefore, the classical results of convex optimization cannot be combined with the SAA approach. We circumvent this difficulty by employing a modified output least-squares (MOLS) objective functional that uses the energy norm and is always convex. The use of the MOLS functional permits us to use the stochastic approximation to the inverse problem of identifying a parameter in stochastic PDEs.

The contents of this paper are organized into five sections. Section 2 presents a projected gradient scheme for variational inequalities in the general stochastic approximation framework. We provide complete convergence analysis for the proposed iterative scheme under weaker conditions on the random noise. In Section 3, we focus on the inverse problem and present three optimization formulations, namely, the OLS functional, the MOLS, functional, and the equation error (EE) approach. The primary focus, however, remains on the MOLS approach. Besides providing technical details on the three functionals in a continuous setting, we also provide a discretization scheme, including discrete formulas for the objective functionals and the gradient for the MOLS functional. Two numerical examples, given in Section 4, demonstrate the feasibility and the efficacy of the developed framework. The paper concludes with some remarks and a discussion of future objectives.

2 Stochastic Approximation for Variational Inequalities

Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$, $K \subset H$ be nonempty, closed, and convex, and $F : H \mapsto H$ be a given map.

We consider the variational inequality of finding $u \in K$ such that

$$\langle F(u), v - u \rangle \geq 0, \quad \text{for all } v \in K. \tag{3}$$

We aim to develop iterative methods for (3) in the general framework of stochastic approximation, that is, when the map F can only be accessed with some random noise. As a particular case, we will explore the variational inequality of finding $u \in K$ such that

$$\langle \mathbb{E}[G(u, \omega)], v - u \rangle \geq 0, \quad \text{for every } v \in K, \tag{4}$$

where $G(\cdot, \cdot) : \Omega \times H \mapsto H$, and $\mathbb{E}[G(u, \omega)]$ is the expected value of $G(u, \omega)$.

Our focus is on the following projected stochastic approximation scheme for (3):

$$u_{n+1} = P_K[u_n - \alpha_n(F(u_n) + \omega_n)], \ \alpha_n > 0. \tag{5}$$

Here $F(u_n)$ is the true value of F at u_n, $F(u_n) + \omega_n$ is an approximation of F at u_n, and ω_n is a stochastic error. In the context of (4), $F(u_n) + \omega_n = G(u_n, \omega_n)$, where ω_n is a sample of the random variable ω. To be specific, here at iteration n, we use a sample ω_n of ω to calculate $G(u_n, \omega)$ and treat it as an approximation of $\mathbb{E}[G(u_n, \omega)] = F(u_n)$. Evidently, $F(u_n)$ can be approximated without any information on the probability distribution of ω.

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

The following result by Robbins and Siegmund [33] will be used shortly:

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

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

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

We also need the following notions of monotonicity and continuity:

Definition 1 Given the Hilbert space H, let $F : X \mapsto X^*$ be a nonlinear map. The map F is called:

1. **monotone**, if

$$\langle Fu - Fv, u - v\rangle \geq 0, \quad \textit{for all } u, v \in X. \tag{6}$$

2. m-**strongly monotone**, if there exists a constant $m > 0$ such that

$$\langle Fu - Fv, u - v\rangle \geq m\|u - v\|^2, \quad \textit{for all } u, v \in X. \tag{7}$$

3. L-**Lipschitz continuous**, if there exists a constant $L > 0$ such that

$$\|Fu - Fv\| \leq L\|u - v\|, \quad \textit{for all } v, v \in X. \tag{8}$$

4. **hemicontinuous**, if the real function $t \mapsto \langle F(u+tv), w\rangle$ is continuous on $[0, 1]$, for all $u, v, w \in X$.

The following result provides the convergence analysis for the scheme (5):

Theorem 1 *Let H be a real Hilbert space, $K \subset H$ be nonempty, closed, and convex, and $F : H \mapsto H$ be given. Let $\{\omega_n\}$ be an H-valued sequence of random variables on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\{u_n\}$ be the sequence generated by (5) and $\mathscr{F}_n := \sigma(u_0, \ldots, u_n)$ be a filtration on $(\Omega, \mathscr{F}, \mathbb{P})$ such that $\{u_n\}$ is $\mathscr{F}_n$-measurable. 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) *F is m-strongly monotone and hemicontinuous.*
(A_3) *There are constants $c_1 \geq 0$ and $c_2 > 0$ such that*

$$\|\mathbb{E}\left[\omega_n|\mathscr{F}_n\right]\| \leq c_1\beta_n\left(1+\|F(u_n)\|\right), \quad \beta_n > 0, \tag{9}$$

$$\mathbb{E}\left[\|\omega_n\|^2|\mathscr{F}_n\right] \leq c_2\left(1+\frac{1}{\delta_n}\|F(u_n)\|^2\right), \quad \delta_n > 0. \tag{10}$$

(A_4) *The sequences $\{\alpha_n\}$, $\{\beta_n\}$, and $\{\delta_n\}$ of positive reals satisfy:*

$$\sum_{n\in\mathbb{N}}\alpha_n = \infty, \quad \sum_{n\in\mathbb{N}}\alpha_n^2 < \infty, \quad \sum_{n\in\mathbb{N}}\frac{\alpha_n^2}{\delta_n} < \infty, \quad \sum_{n\in\mathbb{N}}\alpha_n\beta_n < \infty. \tag{11}$$

Then, $\{u_n\}$ converges, almost surely, to the unique solution $\bar{u}$ of (3).

Proof Due to the strong monotonicity of F, variational inequality (3) has a unique solution $\bar{u} \in K$. Then, we have $\bar{u} = P_K(\bar{u})$, and by (5) and the m-strong monotonicity of F, we get

$$\begin{aligned}
\|u_{n+1}-\bar{u}\|^2 &= \|P_K(u_n-\alpha_n(F(u_n)+\omega_n)) - P_K(\bar{u})\|^2 \\
&\leq \|u_n-\bar{u}-\alpha_n(F(u_n)+\omega_n)\|^2 \\
&= \|u_n-\bar{u}\|^2+\alpha_n^2\|F(u_n)+\omega_n\|^2 - 2\alpha_n\langle F(u_n)+\omega_n, u_n-\bar{u}\rangle \\
&\leq (1-2m\alpha_n)\|u_n-\bar{u}\|^2+2\alpha_n^2\|F(u_n)\|^2+2\alpha_n^2\|\omega_n\|^2 \\
&\quad - 2\alpha_n\langle\omega_n, u_n-\bar{u}\rangle,
\end{aligned}$$

where we used the m-strong monotonicity of F to deduce that

$$\langle F(u_n), u_n-\bar{u}\rangle \geq m\|u_n-\bar{u}\|^2 + \langle F(\bar{u}), u_n-\bar{u}\rangle \geq m\|u_n-\bar{u}\|^2.$$

Next, by taking conditional expectation with respect to $\mathscr{F}_n$, we deduce

$$\begin{aligned}
\mathbb{E}\left[\| u_{n+1}-\bar{u}\|^2|\mathscr{F}_n\right] &\leq (1-2m\alpha_n)\|u_n-\bar{u}\|^2+2\alpha_n^2\|F(u_n)\|^2 \\
&\quad + 2\alpha_n^2\mathbb{E}\left[\|\omega_n\|^2|\mathscr{F}_n\right] + 2\alpha_n\|u_n-\bar{u}\|\|\mathbb{E}\left[\omega_n|\mathscr{F}_n\right]\|.
\end{aligned} \tag{12}$$

To find bounds on the terms in (12), we begin by noticing that

$$\begin{aligned}
\|F(u_n)\| &\leq c(1+\|u_n\|) \\
&\leq c(1+\|\bar{u}\|)+c\|u_n-\bar{u}\| \\
&\leq k_1(1+\|u_n-\bar{u}\|),
\end{aligned} \tag{13}$$

where $k_1 := c(1+\|\bar{u}\|)$, and hence by setting $k_2 := 4k_1^2$, we obtain

$$2\alpha_n^2 \|F(u_n)\|^2 \leq k_2 \alpha_n^2 (1 + \|u_n - \bar{u}\|^2). \tag{14}$$

Moreover, by the inequality $a \leq 1 + a^2$, which holds for all $a \in \mathbb{R}$, and (13), we get

$$\begin{aligned}
\|u_n - \bar{u}\| \|\mathbb{E}[\omega_n | \mathscr{F}_n]\| &\leq \beta_n \|u_n - \bar{u}\| (1 + \|F(u_n)\|) \\
&\leq \beta_n \|u_n - \bar{u}\| (1 + k_1 + k_1 \|u_n - \bar{u}\|)) \\
&\leq \beta_n (1 + k_1) \|u_n - \bar{u}\| + k_1 \beta_n \|u_n - \bar{u}\|^2 \\
&\leq \beta_n (1 + k_1)(1 + \|u_n - \bar{u}\|^2) + k_1 \beta_n \|u_n - \bar{u}\|^2 \\
&\leq \beta_n (1 + k_1) + (1 + 2k_1) \beta_n \|u_n - \bar{u}\|^2,
\end{aligned}$$

and hence setting $k_3 := 2(1 + 2k_1)$, we obtain

$$2\alpha_n \|u_n - \bar{u}\| \|\mathbb{E}[\omega_n | \mathscr{F}_n]\| \leq k_3 \alpha_n \beta_n (1 + \|u_n - \bar{u}\|^2). \tag{15}$$

Finally, using (13) again, we obtain

$$\begin{aligned}
\mathbb{E}\left[\|\omega_n\|^2 | \mathscr{F}_n\right] &\leq c_2 \left(1 + \frac{\|F(u_n)\|^2}{\delta_n}\right) \\
&\leq c_2 \left(1 + \frac{2k_1^2 (1 + \|u_n - \bar{u}\|^2)}{\delta_n}\right) \\
&\leq c_2 + \frac{2c_2 k_1^2}{\delta_n} (1 + \|u_n - \bar{u}\|^2),
\end{aligned}$$

and hence, setting $k_4 := 4c_2 k_1^2$, we obtain

$$2\alpha_n^2 \mathbb{E}\left[\|\omega_n\|^2 | \mathscr{F}_n\right] \leq 2c_2 \alpha_n^2 + \frac{k_4 \alpha_n^2}{\delta_n} + \frac{k_4 \alpha_n^2}{\delta_n} \|u_n - \bar{u}\|^2. \tag{16}$$

Summarizing, due to (12), (14), (15), and (16), there is a constant $k > 0$ with

$$\begin{aligned}
\mathbb{E}[\| u_{n+1} - \bar{u}\|^2 | \mathscr{F}_n] \leq &\left(1 - 2m\alpha_n + k\alpha_n^2 + k\alpha_n \beta_n + \frac{k\alpha_n^2}{\delta_n}\right) \|u_n - \bar{u}\|^2 \\
&+ k\alpha_n^2 + k\alpha_n \beta_n + \frac{k\alpha_n^2}{\delta_n},
\end{aligned} \tag{17}$$

which can be written as

$$\mathbb{E}\left[\|u_{n+1} - \bar{u}\|^2 | \mathscr{F}_n\right] \leq (1 + a_n) \|u_n - \bar{u}\|^2 - c_n + b_n,$$

where

$$
\begin{aligned}
a_n &:= k\alpha_n^2 + k\alpha_n\beta_n + \frac{k\alpha_n^2}{\delta_n},\\
b_n &:= k\alpha_n^2 + k\alpha_n\beta_n + \frac{k\alpha_n^2}{\delta_n},\\
c_n &:= 2m\alpha_n\|u_n - \bar{u}\|^2.
\end{aligned}
$$

Since $\sum_{n\in\mathbb{N}} a_n < \infty$ and $\sum_{n\in\mathbb{N}} b_n < \infty$, as a consequence of Theorem 1, it follows that $\|u_n - \bar{u}\|^2$ converges, almost surely, and

$$
\sum_{n\in\mathbb{N}} 2m\alpha_n\|u_n - \bar{u}\|^2 < +\infty,
$$

which, due to $\sum_{n\in\mathbb{N}} \alpha_n = \infty$, confirms that $\|u_n - \bar{u}\| \to 0$, almost surely. The proof is complete. □

We shall now discuss two special cases of the above result:

Corollary 1 *Let H be a real Hilbert space, $K \subset H$ be nonempty, closed, and convex, and $F : H \mapsto H$ be given. Let $\{\omega_n\}$ be an H-valued sequence of random variables on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\{u_n\}$ be the sequence generated by (5) and $\mathscr{F}_n := \sigma(u_0, \ldots, u_n)$ be a filtration on $(\Omega, \mathscr{F}, \mathbb{P})$ such that $\{u_n\}$ is $\mathscr{F}_n$-measurable. Assume that the following conditions hold:*

(C_1) *F is m-strongly monotone and L-Lipschitz continuous.*
(C_2) $\mathbb{E}[\omega_n|\mathscr{F}_n] = 0$, *and* $\sum_n \alpha_n^2\mathbb{E}\left[\|\omega_n\|^2|\mathscr{F}_n\right] < \infty.$
(C_3) $\alpha_n \in (0, 2m/L^2).$

Then, $\{u_n\}$ converges, almost surely, to the unique solution $\bar{u}$ of (3).

Proof Note that $\bar{u} = P_K(\bar{u} - \alpha_n F(\bar{u}))$, and hence

$$
\begin{aligned}
\|u_{n+1} - \bar{u}\|^2 &= \|P_K(u_n - \alpha_n(F(u_n) + \omega_n)) - P_K(\bar{u} - \alpha_n F(\bar{u}))\|^2\\
&\le \|u_n - \bar{u} - \alpha_n(F(u_n) - F(\bar{u}) + \omega_n)\|^2\\
&\le \|u_n - \bar{u}\|^2 + \alpha_n^2\|F(u_n) - F(\bar{u}) + \omega_n\|^2\\
&- 2\alpha_n\langle F(u_n) - F(\bar{u}) + \omega_n, u_n - \bar{u}\rangle\\
&\le (1 - 2m\alpha_n + 2\alpha_n^2L^2)\|u_n - \bar{u}\|^2 + 2\alpha_n\|\omega_n\|^2 - 2\alpha_n\langle\omega_n, u_n - \bar{u}\rangle,
\end{aligned}
$$

and by taking the expectation past $\mathscr{F}_n$, we deduce

$$\mathbb{E}\,[\|\,u_{n+1}-\bar{u}\|^2|\mathscr{F}_n] \le (1-2m\alpha_n+2\alpha_n^2L^2)\|u_n-\bar{u}\|^2+\alpha_n^2\mathbb{E}\left[\|\omega_n\|^2|\mathscr{F}_n\right],$$

which can be written as

$$\mathbb{E}\left[\|u_{n+1}-\bar{u}\|^2|\mathscr{F}_n\right] \le (1+a_n)\|u_n-\bar{u}\|^2-c_n+b_n,$$

where for a positive constant $k>0,$ we have

$$\begin{aligned} a_n &:= 0,\\ b_n &:= \alpha_n^2\mathbb{E}\left[\|\omega_n\|^2|\mathscr{F}_n\right],\\ s_n &:= 2(\alpha_n m-\alpha_n^2L^2),\\ c_n &:= s_n\|u_n-\bar{u}\|^2. \end{aligned}$$

Due to imposed conditions, we have $\sum_{n=\in\mathbb{N}} a_n<\infty$, and $\sum_{n\in\mathbb{N}} b_n<\infty$, almost surely. As a consequence, $\|u_n-\bar{u}\|^2$ converges almost surely, and $\sum_{n=1}^{\infty} c_n<\infty$, almost surely. Furthermore, since s_n is bounded away from zero, we infer that the sequence $\{u_n\}$ converges strongly to $\bar{u}$, almost surely. The proof is complete. □

Corollary 2 *Let H be a real Hilbert space, $K\subset H$ be nonempty, closed, and convex, and $F:H\mapsto H$ be given. Let $\{\omega_n\}$ be an H-valued sequence of random variables on a probability space $(\Omega,\mathscr{F},\mathbb{P})$. Let $\{u_n\}$ be the sequence generated by (5) and $\mathscr{F}_n:=\sigma(u_0,\ldots,u_n)$ be a filtration on $(\Omega,\mathscr{F},\mathbb{P})$ such that $\{u_n\}$ is $\mathscr{F}_n$-measurable.*

(H_1) *There is a constant $c>0$ such that $\|F(u)\|\le c(1+\|u\|)$, for every $u\in K$.*
(H_2) *F is m-strongly monotone and hemicontinuous.*
(H_3) $\mathbb{E}\,[\omega_n|\mathscr{F}_n]=0$, *and* $\sum_n \alpha_n^2\mathbb{E}\left[\|\omega_n\|^2|\mathscr{F}_n\right]<\infty$.
(H_4) *The sequence $\{\alpha_n\}$ of positive reals satisfies:*

$$\sum_{n\in\mathbb{N}}\alpha_n=\infty,\quad \sum_{n\in\mathbb{N}}\alpha_n^2<\infty. \tag{18}$$

Then, $\{u_n\}$ converges, almost surely, to the unique solution $\bar{u}$ of (3).

Proof The proof is based on the arguments used above. □

Remark 1 Hiriart-Urruty [19] extended the stochastic approximation approach to nonlinear variational inequalities when some random noise contaminated the data. He proposed a variety of projection-type iterative methods in Hilbert spaces, even considered variational inequalities with multi-valued maps, and provided several convergence theorems in quadratic mean and almost certain sense. Theorem 1 is given under the same condition F as in [19]; however, we have more general conditions on random noise, which were inspired by Barty, Roy, and Strugarek [3]. Jiang and Xu [23] initiated a detailed study of the stochastic approximation framework for the expected value formulation of variational inequalities. Corollary 1 is similar to the results [23], given for the particular case of an expected value formulation of a variational inequality.

3 Stochastic Approximation for Inverse Problems

In this section, we will study the inverse problem of identifying a deterministic parameter in a stochastic partial differential equation. In the final section, we will discuss the extension of the present framework to the case of a stochastic parameter.

3.1 *Optimization Formulation of the Inverse Problem*

We recall that given a real Banach space X, a measure space $(\Omega, \mathscr{F}, \mu)$, 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 pth moment, that is,

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

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

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

For $\omega \in \Omega$, variational formulation of (1) seeks $u_\omega \in V := H_0^1(D)$ such that

$$\int_D a(\omega, x) \nabla u_\omega(a) \cdot \nabla v dx = \int_D f(\omega, x) v \, dx, \text{ for all } v \in V. \tag{19}$$

Assume that there are constants k_0 and k_1 such that

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

The following is a well-known result for (19):

Lemma 2 *Let $f \in L^2(\Omega, H^{-1}(D))$. Then, there is a positive constant c such that*

$$\|u_\omega(a)\|_{H_0^1(D)} \leq c\|f\|_{H^{-1}(D)} \quad \text{for a.e. } \omega \in \Omega,$$

$$\|u(a)\|_{L^2(\Omega, H_0^1(D))} \leq c\|f\|_{L^2(\Omega, H_0^{-1}(D))}.$$

In the following, we shall assume that a is deterministic. Moreover, for positive k_0 and k_1, we define the set of admissible parameters:

$$\mathbb{A} := \left\{a \in L^\infty(D) : \ 0 < k_0 \leq a(x) \leq k_1 < \infty, \ x \in D\right\}. \tag{20}$$

We now state some technical results. Since these results are stated for realizations, their proofs are natural generalizations of the results given in [15] for the corresponding deterministic case.

Theorem 2 *For $\omega \in \Omega$, the map $\mathbb{A} \ni a \mapsto u_\omega(a)$ is Lipschitz continuous.*

Theorem 3 *For $\omega \in \Omega$, and a in the interior of $\mathbb{A}$, the map $a \mapsto u_\omega(a)$ is differentiable at a. The derivative $\delta u_\omega := Du_\omega(a)(\delta a)$ of $u_\omega(a)$ at a in the direction δa is the unique solution of the stochastic variational problem: Find $\delta u_\omega \in V$ such that*

$$\int_D a(x)\nabla\delta u_\omega \cdot \nabla v dx = -\int_D \delta a \nabla u_\omega(a) \cdot \nabla v \, dx, \ \text{for all } v \in V. \tag{21}$$

One of the most commonly used optimization formulations is the following output least-squares (OLS) objective functional:

$$\widehat{\mathbb{J}}(a) = \frac{1}{2}\mathbb{E}\left[\|u_\omega(a) - z_\omega\|^2_{L^2(D)}\right], \tag{22}$$

where $u_\omega(a)$ is the solution of (19) for a and z_ω is the measured data.

One of the shortcomings of the above functional is that it is nonconvex, in general. Although it is known that the gradient of the OLS functional, with the aid of a regularization, can be made strongly monotone, it runs into the risk of over-regularizing the identification process, see [14].

We now define the modified output least-squares (MOLS) objective functional:

$$\mathbb{J}(a) = \frac{1}{2}\mathbb{E}\left[\int_D a(x)\nabla(u_\omega(a) - z_\omega) \cdot \nabla(u_\omega(a) - z_\omega) dx\right], \tag{23}$$

where $u_\omega(a)$ is the solution of (19) for a and z_ω is the measured data.

The following result summarizes some properties of the MOLS objective:

Theorem 4 *Let a be in the interior of $\mathbb{A}$. Then:*

1. *The first derivative of $\mathbb{J}$ at a is given by*

$$D\mathbb{J}(a)(\delta a) = -\frac{1}{2}\mathbb{E}\left[\int_D \delta a \nabla(u_\omega(a) + z_\omega)\nabla(u_\omega(a) - z_\omega)\,dx\right].$$

2. *The second derivative of* $\mathbb{J}$ *at* a *is given by*

$$D^2\mathbb{J}(a)(\delta a, \delta a) = \mathbb{E}\left[\int_D a(x)\nabla u_\omega(a)\nabla u_\omega(a)\,dx\right].$$

Consequently, the MOLS functional is convex in the interior of the set $\mathbb{A}$.

For the sake of a comparison, we would also describe another commonly used method, the so-called equation error approach (see [13]), which consists of minimizing, for $\omega \in \Omega$, and for the data $z_\omega \in H_0^1(D)$, the quadratic objective functional:

$$\min_{a\in\mathbb{A}} \widetilde{\mathbb{J}}(a) = \frac{1}{2}\mathbb{E}\left[\|e_\omega(a, z_\omega)\|^2_{H_0^1}\right], \tag{24}$$

where $e_\omega(a, u_\omega) \in H_0^1(D)$ satisfies the following variational problem:

$$\langle e_\omega(a, u_\omega), v\rangle_{H_0^1(D)} = \int_D a\nabla u_\omega \cdot \nabla v - \int_D f(\omega, x)v, \quad \text{for all } v \in H_0^1(D).$$

Since the inverse problem of identifying parameters in partial differential equations is ill-posed, and for a stable identification process, some regularization is needed. For this, we assume that the set of admissible parameters A belongs to a Hilbert space that is compactly embedded into $L^\infty(D)$.

Therefore, we consider the following regularized analogues of the three functionals described above:

$$\min_{a\in\mathbb{A}} \widehat{\mathbb{J}}_\kappa(a) := \frac{1}{2}\mathbb{E}\left[\|u_\omega(a) - z_\omega\|^2_{L^2(D)}\right] + \frac{\kappa}{2}\|a\|^2_H, \tag{25}$$

$$\min_{a\in\mathbb{A}} \mathbb{J}_\kappa(a) := \frac{1}{2}\mathbb{E}\left[\int_D a(x)\nabla(u_\omega(a) - z_\omega)\cdot\nabla(u_\omega(a) - z_\omega)dx\right] + \frac{\kappa}{2}\|a\|^2_H, \tag{26}$$

$$\min_{a\in\mathbb{A}} \widetilde{\mathbb{J}}_\kappa(a) := \frac{1}{2}\mathbb{E}\left[\|e_\omega(a, z_\omega)\|^2_{H_0^1}\right] + \frac{\kappa}{2}\|a\|^2_H. \tag{27}$$

Here $u_\omega(a)$ is the solution of (19) for $a(x)$, z_ω is the measured data, $\kappa > 0$ is a fixed regularization parameter, and $\|\cdot\|^2_H$ is the regularizer.

Since $\mathbb{J}$ is convex and $\mathbb{A}$ is closed and convex, the following variational inequality is a necessary and sufficient optimality condition for (26): Find $a \in \mathbb{A}$ such that

$$\langle\nabla\mathbb{J}(a), b - a\rangle + \kappa\langle a, b - a\rangle \geq 0, \quad \text{for every } b \in \mathbb{A}. \tag{28}$$

Note that by defining

$$J(a,\omega) = \int_D a(x)\nabla(u_\omega(a) - z_\omega) \cdot \nabla(u_\omega(a) - z_\omega)\, dx,$$

we can show that

$$\nabla\mathbb{J}(a,\omega)(\delta a) = -\frac{1}{2}\int_D \delta a(x)\nabla(u_\omega(a) + z_\omega)\nabla(u_\omega(a) - z_\omega)\, dx,$$

and, consequently,

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

Therefore, it follows that

$$\nabla\mathbb{J}_\kappa(a) = \nabla\mathbb{E}\left[J(a,\omega) + \kappa a\right] = \mathbb{E}\left[\nabla J(a,\omega) + \kappa a\right] = \mathbb{E}\left[G(a,\omega)\right], \tag{29}$$

where we set $G(a,\omega) = \nabla J(a,\omega) + \kappa u$.

Analogous statements can be made for the OLS objective and the EE objective.

3.2 Discrete Formulas

We will use a standard finite element discretization of the spaces V and H. We begin, therefore, with a triangulation $\mathcal{T}_h$ on D. Let V_h and H_h be the spaces of piecewise linear continuous polynomials relative to $\mathcal{T}_h$. Let $\{\phi_1, \phi_2, \dots, \phi_m\}$ and $\{\varphi_1, \varphi_2, \dots, \varphi_l\}$ be the corresponding bases for V_h and H_h, respectively. The space H_h is then isomorphic to $\mathbb{R}^l$, and for any $a \in H_h$, we define $A \in \mathbb{R}^l$ by $A_i = a(x_i)$, $i = 1, 2, \dots, l$, where the nodal basis $\{\varphi_1, \varphi_2, \dots, \varphi_l\}$ corresponds to the nodes $\{x_1, x_2, \dots, x_l\}$. Conversely, each $A \in \mathbb{R}^l$ corresponds to $a \in H_h$ defined by $a(x) = \sum_{i=1}^{l} A_i\varphi_i$. Analogously, $u \in V_h$ will correspond to $U \in \mathbb{R}^m$, where $U_i = u(y_i)$, $i = 1, 2, \dots, m$, and $u = \sum_{i=1}^{m} U_i\phi_i$, where $y_1, y_2, \dots, y_m$ are the interior nodes of the finite element mesh (triangulation) $\mathcal{T}_h$.

Given a realization/sample $\omega \in \Omega$, the discrete version of variational problem (19) seeks $U = U(\omega, A) \in \mathbb{R}^m$ by solving

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

where $K(A) \in \mathbb{R}^{m\times m}$ and $F(\omega_n) \in \mathbb{R}^m$ are the stiffness matrix and the load vector defined by

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

$$F(\omega)_i = \int_D f_h(\omega, x)\phi_i dx, \quad \text{for } i = 1, \dots, m.$$

To compute the gradient of the MOLS objective, it is convenient to define the so-called adjoint stiffness matrix $L(\cdot) \in \mathbb{R}^{m\times l}$ by the condition

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

Then,

$$\begin{aligned}\nabla J(A, \omega)(\delta A) &= -\frac{1}{2}(U(\omega, A) + Z(\omega))^\top K(\delta A)(U(\omega, A) - Z(\omega))\\ &= -\frac{1}{2}\delta A^\top L(U(\omega, A) + Z(\omega))^\top (U(\omega, A) - Z(\omega)),\end{aligned}$$

which yields

$$\nabla J_\kappa(A, \omega) = -\frac{1}{2}L(U(\omega, A) + Z(\omega))^T (U(\omega, A) - Z(\omega)) + \kappa(\mathbb{M} + \mathbb{K})A,$$

where $\mathbb{M}, \mathbb{K} \in \mathbb{R}^{m\times m}$ are the corresponding mass and stiffness matrices in H_h:

$$\mathbb{M}_{i,j} = \int_D \varphi_j\varphi_i dx, \quad \text{for } i, j = 1, \dots, l,$$

$$\mathbb{K}_{i,j} = \int_D \nabla\varphi_j \cdot \nabla\varphi_i dx, \quad \text{for } i, j = 1, \dots, l.$$

The above preparation permits to define the following stochastic approximation scheme for computing a solution of the discrete variant of (28):

In the classical stochastic gradient, a single sampling is done at each iterative step. However, in the above algorithm, instead of sampling the random variable at each step once, at step n, we sample a predetermined number s_n times, called the sample rate, and use the empirical average to approximate the expected value.

4 Computational Experiments

In this section, we present results from our numerical computations. We consider the domain $D = (0, 1)$ and choose functions $a(x)$ and $u(\omega, x) = u(Y_1(\omega), Y_2(\omega), x)$ and compute the corresponding $f(\omega, x)$ by $f(\omega, x) = -(a(x)u_x(\omega, x))_x$. We choose a uniform mesh on $(0, 1)$ with mesh size $h = 1/(N + 1)$, where N stands

Algorithm 1 Stochastic approximation for parameter identification

1: Choose an initial guess $A_0 \in \mathbb{R}^m$, positive step-lengths $\{\alpha_n\}$ satisfying (18), the sample rate $\{s_n\} \subset \mathbb{N}$, and initial samples $\{\omega_j^0\}_{j=1}^{s_0}$ of the random variable ω.

2: Given $A_n \in A$, generate samples $\{\omega_j^n\}_{j=1}^{s_n}$ of ω and define

$$A_{n+1} = P_{\mathbb{A}}\left[A_n - \frac{\alpha_n}{s_n}\sum_{j=1}^{s_n} G\left(\omega_j^n, A_n\right)\right], \tag{30}$$

where G is the discrete variant of gradient of the regularized MOLS objective (see (29)).

3: Stop if some stopping criteria are met.

for the number of interior nodes. The same set of piecewise linear finite element basis functions is used for the representations of $a(x)$ and $u(\omega, x)$; therefore, $U(\omega, A) \in \mathbb{R}^N$ (for a fixed ω) and $A \in \mathbb{R}^{N+2}$ (i.e., $m = N + 2$). The constraint set K is defined by

$$\mathbb{A} = \{a \in H^1(\Omega) : \ a_0 \le a(x) \le a_1\}.$$

Example 1 For this example, we choose

$$a(x) = 1 + x^2,$$
$$u(\omega, x) = Y_1(\omega)x(1-x) + Y_2(\omega)\sin(3\pi x),$$

where $Y_1(\omega), Y_2(\omega) \sim U[0, 1]$, i.e., random variables Y_1 and Y_2 are uniformly distributed over interval $[0, 1]$. We choose $a_0 = 0.5$ and $a_1 = 3$ and use $N = 99$, $s_n = 5$, $\alpha_n = 0.5\alpha_0/n$ with $\alpha_0 = 10^4$ in Algorithm 1 for this example. Iterations are terminated once the L^2 norm of the expected value of the gradient drops below $\gamma = 10^{-7}$. Results of this computation using the MOLS method are shown in Figure 1. Regularization parameter $\kappa = 10^{-6}$ is used to produce these figures.

Example 2 In this example, we choose the same $u(\omega, x)$ as in Example 1, but with a slightly more interesting function $a(x)$ defined by

$$a(x) = 2\sin(\pi(x - 0.2)) - 2\tanh(20x - 8) + 4.$$

Figure 2 shows results of a run using parameters $N = 159$, $s_n = 10$, $\alpha_0 = 10^5$, and $\kappa = 10^{-7}$ using the MOLS method. For the constraints, we use $a_0 = 1$ and $a_1 = 8$. Figure 3 shows some realizations of the random fields $u(\omega, x)$ and $f(\omega, x)$. Note that Figures 1 and 2 represent results of a typical simulation. Regularization parameter κ is chosen after we do several test runs for a particular set of parameter values. The method gives us a very stable reconstruction of the coefficient $a(x)$ in each case regardless of the choice of the initial approximate $A^{(0)}$.

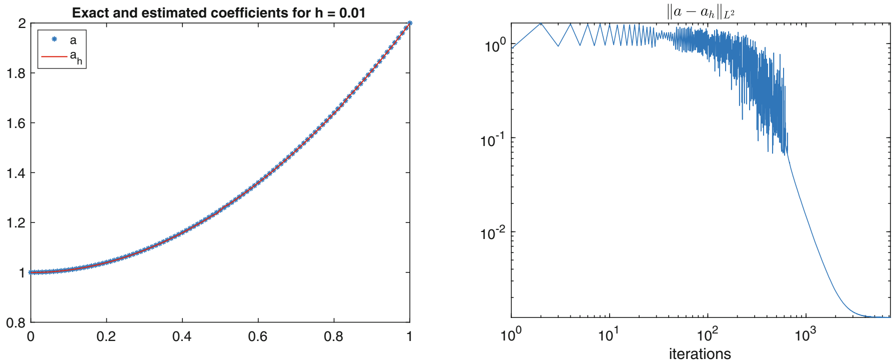

Fig. 1 Example 1: Comparison of exact coefficient a and the approximated coefficient a_h using MOLS method (left) and loglog plot of the error $\|a - a_h\|_{L^2}$ versus iterations (right)

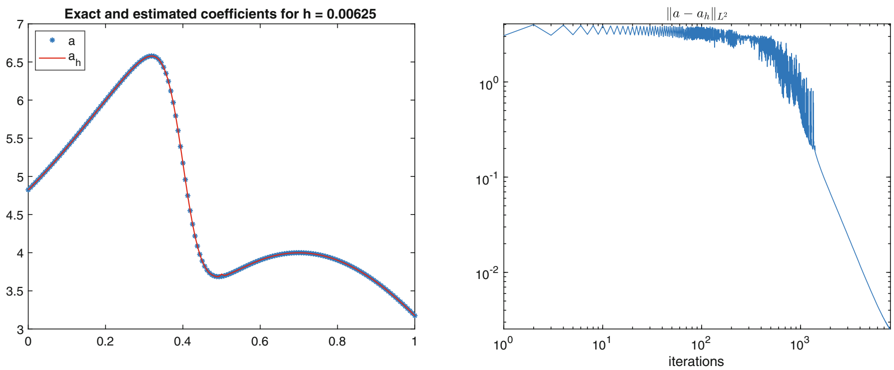

Fig. 2 Example 2: Comparison of exact coefficient a and the approximated coefficient a_h using MOLS method (left) and loglog plot of the error $\|a - a_h\|_{L^2}$ versus iterations (right)

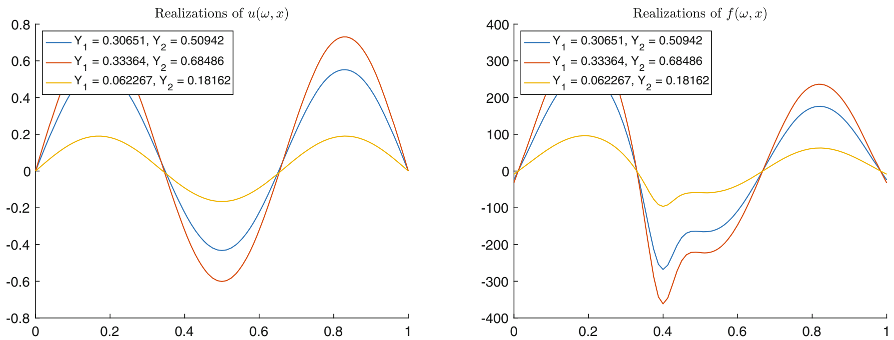

Fig. 3 Typical realizations of the random fields u and f from Example 2

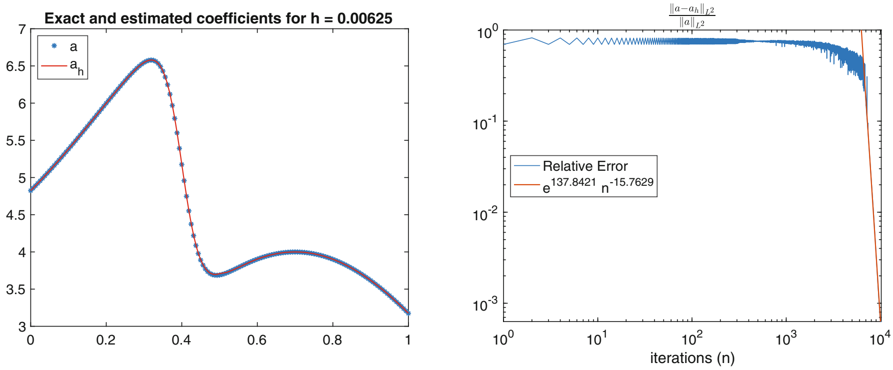

Fig. 4 Example 2: Comparison of exact coefficient a and the approximated coefficient a_h using EE method (left) and loglog plot of the relative error $\|a - a_h\|_{L^2}/\|a\|_{L^2}$ versus iterations (right)

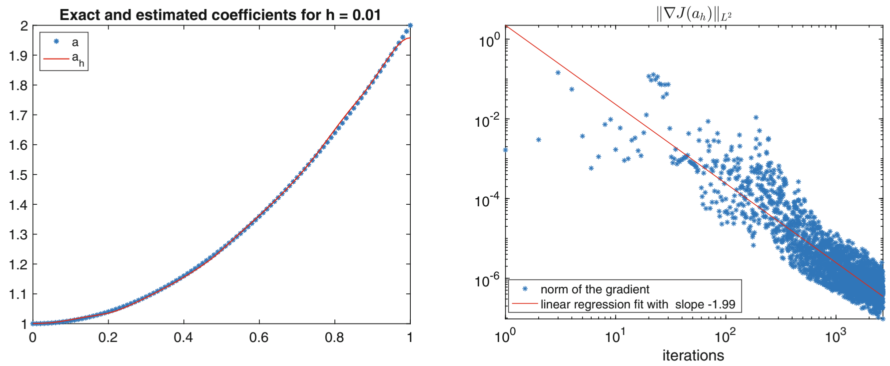

Fig. 5 Example 1: Comparison of exact coefficient a and the approximated coefficient a_h using the OLS method (left) and loglog plot of the norm of the gradient $\|\nabla J(a_h)\|_{L^2}$ versus iterations (right)

Results of Computations by EE and OLS Methods We compare the performance of the MOLS method with those of the OLS and EE methods (see equations (27) and (25) for regularized objective functional definitions). Figure 4 shows the results of a run with parameters $N = 159$, $s_n = 20$, $\alpha_0 = 10^6$, and $\kappa = 5 \cdot 10^{-7}$ for Example 2 using EE method. The quality of the estimation is excellent and the results are comparable with those of the MOLS method. No significant gain in the computational cost for the EE method is observed as our examples are in 1D (these computations take only a minute or two in MATLAB). However, the EE method is expected to have considerable computational cost advantage for problems in two or three space dimensions compared to MOLS and OLS methods. Figure 5 shows results of a simulation using OLS method for Example 1. Parameter values used in the computation are $N = 99$, $s_n = 1$, $\alpha_0 = 10^5$, and $\kappa = 5 \cdot 10^{-6}$. Tolerance for the L^2 norm of the gradient is set to $\gamma = 10^{-7}$ (see the right plot in the figure referenced above which shows the decrease of this norm as iterations progress). The quality of the estimation seems to be not as good as the ones we obtained from the MOLS and EE methods, and there is a mismatch close to the right boundary of the domain. While applying the OLS method to both examples, we observed that the method requires a more careful tuning of the parameters compared to the other two methods we used in our experiments.

5 Concluding Remarks

We developed a stochastic approximation approach for identifying a deterministic parameter in a stochastic partial differential equation. Besides considering more general stochastic PDEs such as linear elasticity or fourth-order plate models, a desirable extension of this work is to identify a stochastic parameter $a(\omega, x)$. A natural approach would be to separate the deterministic and stochastic components by using the so-called finite-dimensional noise assumption (see [28]). The deterministic components can then be identified by extending the stochastic approximation framework. We aim to pursue this approach in future work.

Acknowledgments Contributions of R. Hawks, B. Jadamba, A. A. Khan, and Y. Yang are supported by the National Science Foundation under Award No. 1720067. M. Sama's work is partially supported by the Ministerio de Ciencia, Innovacion y Universidades (MCIU), Agencia Estatal de Investigacion (AEI) (Spain), and Fondo Europeo de Desarrollo Regional (FEDER) under project PGC2018-096899-B-I00 (MCIU/AEI/FEDER, UE).

References

1. 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**(2), 1–19 (2016)
2. V.A. Badri Narayanan, N. Zabaras, Stochastic inverse heat conduction using a spectral approach. Internat. J. Numer. Methods Eng. **60**(9), 1569–1593 (2004)
3. K. Barty, J.-S. Roy, C. Strugarek, Hilbert-valued perturbed subgradient algorithms. Math. Oper. Res. **32**(3), 551–562 (2007)
4. J.-P. Bertran, Optimisation stochastique dans un espace de Hilbert. Méthode de gradient. C. R. Acad. Sci. Paris Sér. A-B **276**, A613–A616 (1973)
5. D.P. Bertsekas, J.N. Tsitsiklis, Gradient convergence in gradient methods with errors. SIAM J. Optim. **10**(3), 627–642 (2000)
6. 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 Prob. **263**, 1–14 (2013)
7. M. Cho, B. Jadamba, R. Kahler, A.A. Khan, M. Sama, First-order and second-order adjoint methods for the inverse problem of identifying nonlinear parameters in PDEs, in *Industrial Mathematics and Complex Systems* (Springer, Berlin, 2017), pp. 1–16.

8. 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**(1), 122–135 (2014)
9. J.-C. Culioli, G. Cohen, Decomposition/coordination algorithms in stochastic optimization. SIAM J. Control Optim. **28**(6), 1372–1403 (1990)
10. 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**(7–9), 984–1017 (2014)
11. C. Geiersbach, G.C. Pflug, Projected stochastic gradients for convex constrained problems in Hilbert spaces. SIAM J. Optim. **29**(3), 2079–2099 (2019)
12. 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*. Contemporary Mathematics, vol. 659 (American Mathematical Society, Providence, 2016), pp. 65–89
13. M.S. Gockenbach, B. Jadamba, A.A. Khan, Numerical estimation of discontinuous coefficients by the method of equation error. Int. J. Math. Comput. Sci. **1**(3), 343–359 (2006)
14. M.S. Gockenbach, A.A. Khan, Identification of Lamé parameters in linear elasticity: a fixed point approach. J. Ind. Manag. Optim. **1**(4), 487–497 (2005)
15. M.S. Gockenbach, A.A. Khan, An abstract framework for elliptic inverse problems: part 1. an output least-squares approach. Math. Mech. Solids **12**(3), 259–276 (2007)
16. M.S. Gockenbach, A.A. Khan, An abstract framework for elliptic inverse problems. II. An augmented Lagrangian approach. Math. Mech. Solids **14**(6), 517–539 (2009)
17. L. Goldstein, Minimizing noisy functionals in Hilbert space: an extension of the Kiefer–Wolfowitz procedure, J. Theoret. Probab. **1**(2), 189–204 (1988)
18. J. Gwinner, B. Jadamba, A.A. Khan, M. Sama, Identification in variational and quasi-variational inequalities. J. Convex Anal. **25**(2), 545–569 (2018)
19. J.-B. Hiriart-Urruty, Algorithmes stochastiques de résolution d'équations et d'inéquations variationnelles. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **33**(3), 167–186 (1975/1976)
20. B. Jadamba, R. Kahler, A.A. Khan, F. Raciti, B. Winkler, Identification of flexural rigidity in a Kirchhoff plates model using a convex objective and continuous Newton method. Math. Probl. Eng. **2015**, 290301 (2015)
21. 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 Prob. Sci. Eng. **25**(12), 1768–1787 (2017)
22. 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**(5), 1486–1510 (2014)
23. H. Jiang, H. Xu, Stochastic approximation approaches to the stochastic variational inequality problem. IEEE Trans. Automat. Control **53**(6), 1462–1475 (2008)
24. M. Keyanpour, A.M. Nehrani, Optimal thickness of a cylindrical shell subject to stochastic forces. J. Optim. Theory Appl. **167**(3), 1032–1050 (2015)
25. J. Kiefer, J. Wolfowitz, Stochastic estimation of the maximum of a regression function. Ann. Math. Stat. **23**, 462–466 (1952)
26. H.J. Kushner, A. Shwartz, Stochastic approximation in Hilbert space: identification and optimization of linear continuous parameter systems. SIAM J. Control Optim. **23**(5), 774–793 (1985)
27. T.L. Lai, Stochastic approximation. Ann. Stat. **31**(2), 391–406 (2003)
28. G.J. Lord, C.E. Powell, T. Shardlow, in *An Introduction to Computational Stochastic PDEs*. Cambridge Texts in Applied Mathematics (Cambridge University Press, New York, 2014)
29. M. Martin, S. Krumschield, F. Nobile, Analysis of stochastic gradient methods for PDE-constrained optimal control problems with uncertain parameters. Preprint 1–39 (2018)
30. M. Morzfeld, X. Tu, J. Wilkening, A.J. Chorin, Parameter estimation by implicit sampling. Commun. Appl. Math. Comput. Sci. **10**(2), 205–225 (2015)

31. R. Naseri, A. Malek, Numerical optimal control for problems with random forced SPDE constraints. ISRN Appl. Math. 2014, 974305 (2014)
32. H. Robbins, S. Monro, A stochastic approximation method. Ann. Math. Stat. **22**, 400–407 (1951)
33. H. Robbins, D. Siegmund, A convergence theorem for non negative almost supermartingales and some applications, in *Optimizing Methods in Statistics* (Academic Press, Cambridge, 1971) pp. 233–257
34. B. V. Rosić, H.G. Matthies, Identification of properties of stochastic elastoplastic systems, in *Computational Methods in Stochastic Dynamics*, vol. 2 (Springer, Dordrecht, 2013), pp. 237–253
35. G.I. Salov, Stochastic approximation in a Hilbert space in the problem of the detection of the appearance of an object in a sequence of noisy images. Sib. Zh. Ind. Mat. **12**(1), 127–135 (2009)
36. K. Sepahvand, S. Marburg, On construction of uncertain material parameter using generalized polynomial chaos expansion from experimental data. Proc. IUTAM **6**, 4–17 (2013)
37. G. Yin, Y.M. Zhu, On H-valued Robbins–Monro processes. J. Multivar. Anal. **34**(1), 116–140 (1990)
38. N. Zabaras, B. Ganapathysubramanian, A scalable framework for the solution of stochastic inverse problems using a sparse grid collocation approach. J. Comput. Phys. **227**(9), 4697–4735 (2008)