Efficient Stochastic Galerkin Methods for Maxwell's Equations with Random Inputs

Zhiwei Fang, Jichun Li, Tao Tang & Tao Zhou

Journal of Scientific Computing

ISSN 0885-7474

J Sci Comput
DOI 10.1007/s10915-019-00936-z

Springer

Journal of Scientific Computing
https://doi.org/10.1007/s10915-019-00936-z

Efficient Stochastic Galerkin Methods for Maxwell's Equations with Random Inputs

Zhiwei Fang[1] · Jichun Li[1] · Tao Tang[2] · Tao Zhou[3]

Received: 19 September 2018 / Revised: 20 February 2019 / Accepted: 28 February 2019

Abstract
In this paper, we are concerned with the stochastic Galerkin methods for time-dependent Maxwell's equations with random input. The generalized polynomial chaos approach is first adopted to convert the original random Maxwell's equation into a system of deterministic equations for the expansion coefficients (the Galerkin system). It is shown that the stochastic Galerkin approach preserves the energy conservation law. Then, we propose a finite element approach in the physical space to solve the Galerkin system, and error estimates is presented. For the time domain approach, we propose two discrete schemes, namely, the Crank–Nicolson scheme and the leap-frog type scheme. For the Crank–Nicolson scheme, we show the energy preserving property for the fully discrete scheme. While for the classic leap-frog scheme, we present a conditional energy stability property. It is well known that for the stochastic Galerkin approach, the main challenge is how to efficiently solve the coupled Galerkin system. To this end, we design a modified leap-frog type scheme in which one can solve the coupled system in a decouple way—yielding a very efficient numerical approach. Numerical examples are presented to support the theoretical finding.

Keywords Maxwell's equations · Finite element method · Random inputs · Polynomial chaos methods · Stochastic Galerkin

Mathematics Subject Classification 65N30 · 35L15 · 78-08

1 Introduction

With the increase of computer power, simulation of complex physical systems governed by various partial differential equations (PDEs) with random inputs becomes necessary, since uncertainty is ubiquitous in many complex physical systems. For example, in simulating

Jichun Li: Work partially supported by NSF Grant DMS-1416742 and NSFC Project 11671340. Tao Tang: Work supported by the NSF of China (under the Grant No. 11731006) and the Science Challenge Project (No. TZ2018001). Tao Zhou: Work supported by the NSF of China (under Grant Nos. 11822111, 11688101, 91630203, 11571351, 11731006), the Science Challenge Project (No. TZ2018001), the National Key Basic Research Program (No. 2018YFB0704304), NCMIS, and the Youth Innovation Promotion Association (CAS).

Extended author information available on the last page of the article

Published online: 07 March 2019

electric current flow in coaxial cables [21], the cables may be corrugated, which leads to a physical domain with some randomness. In the producing process of electromagnetic materials, random fluctuations often allow researchers to treat the permittivity and permeability as uncertain parameters (e.g., [4,6]). In the past two decades, the study of uncertainty quantification (UQ) got great attentions across different disciplines of sciences and engineering. For stochastic PDEs (SPDEs), uncertain inputs can appear in coefficients, forcing terms, boundary and initial conditions, and physical domains, etc (cf. [7,9,11,28]).

Due to the high dimensionality of random variables, it is very challenging to efficiently solve PDEs with random inputs. Through the great efforts of many researchers, two major class of numerical methods have become very popular in solving random PDEs. One is the so-called non-intrusive stochastic collocation method (cf. [1,30,36]), which is simple in implementation and the system of resulting equations is decoupled and hence is efficient to solve. The stochastic collocation method can achieve fast convergence when the solutions are sufficiently smooth in the random space. Another popular approach is the intrusive stochastic Galerkin method [2,13–15], which shows fast convergence rates with increasing order of expansions, provided that the solution of the underlying differential equation is sufficiently smooth in the random space. However, the system of equations resulting from the stochastic Galerkin methods is in general coupled and quite expensive to solve especially for problems requiring high-dimensional random spaces. The stochastic Galerkin method is based on the polynomial chaos (PC) approximation, originally developed by Ghanem and Spanos [13] using Wiener–Hermite expansion and finite element discretization for a wide range of problems. It was later extended by Xiu and Karniadakis [35] to generalized polynomial chaos (gPC) expansion by using general orthogonal polynomials. Based on gPC expansion and stochastic Galerkin projection, a given random PDE can be transformed into a system of deterministic PDEs (the Galerkin system) which can be solved by any existing popular numerical methods. So far, both the stochastic Galerkin method and the stochastic collocation method have been widely used to solve various problems, such as elliptic problems (e.g., [8,10,27,32,38]), parabolic equations (e.g., [25]), hyperbolic equations (e.g., [16,17,24,31,33,39], to name a few. More details can be found in recent review articles [15,26,29]) and monographs [22,34].

Compared to many papers published for those PDEs mentioned above, there are much less existing works for solving stochastic Maxwell's equations. In 2006, Chauviere et al. [6] developed both the stochastic Galerkin method and stochastic collocation method to solve the time-dependent Maxwell's equations. Detailed comparisons of both methods are made for uncertainties caused by physical materials, by the source wave and by the physical domain. In 2015, Benner and Schneider [4] described several techniques for uncertainty quantification for the time-harmonic Maxwell's equations by using stochastic collocation method. However, how uncertainty propagates through the stochastic Maxwell's equations and the relevant regularity analysis are not investigated yet as Chauviere et al. pointed out [6, p. 774]. Recently, these issues were investigated through the stochastic collocation method in [18,19].

We are well aware that the collocation methods as investigated in [19] are non-intrusive and can be easily applied. However, the stochastic Galerkin methods, which is mathematically elegant, can also be a good candidate. In particular, to get the same polynomial accuracy, the Galerkin methods use much less degree of freedom compared to the stochastic collocation methods. However, it is well known that the main drawback in using stochastic Galerkin methods is that one has to deal with a coupled Galerkin system. In this work, we shall explore stochastic Galerkin methods by designing efficient solvers to the Galerkin system. More precisely, we propose a finite element approach in the physical space and error estimates will be presented. For the time domain approach, we propose two discrete schemes, namely, the Crank–Nicolson scheme and the leap-frog type scheme. For the Crank–Nicolson scheme, we show the energy preserving property for the fully discrete scheme. While for the classic leap-frog scheme, we present a conditional energy stability property. To further reduce the computational cost, we design a revised leap-frog type scheme in which one can solve the coupled system in a decouple way—yielding a very efficient numerical approach. Numerical examples are presented to support the theoretical finding.

The rest of the paper is organized as follows. In Sect. 2, we first carry out some analysis of the gPC method for Maxwell's equations. In Sect. 3, we develop and analyze both the semi-discrete and fully-discrete finite element schemes for solving the system resulting from the gPC method. Numerical results are presented in Sect. 4 to support our theoretical analysis. We conclude the paper in Sect. 5.

2 The gPC Approach for Maxwell's Equations

Consider the three-dimensional Maxwell's equations with random coefficients

$$\epsilon(\boldsymbol{x}, \boldsymbol{y})\partial_t \boldsymbol{E}(t, \boldsymbol{x}, \boldsymbol{y}) = \nabla \times \boldsymbol{H}(t, \boldsymbol{x}, \boldsymbol{y}), \tag{1}$$

$$\mu(\boldsymbol{x}, \boldsymbol{y})\partial_t \boldsymbol{H}(t, \boldsymbol{x}, \boldsymbol{y}) = -\nabla \times \boldsymbol{E}(t, \boldsymbol{x}, \boldsymbol{y}), \tag{2}$$

where $\boldsymbol{x}$ denotes the spatial variable in the three-dimensional domain D and $\boldsymbol{y} = (y_1, y_2, \ldots, y_N)^T \in R^N, N \geq 1$, is a random vector with independent and identically distributed components. Furthermore, we assume that the Eqs. (1)–(2) are subject to the initial conditions

$$\boldsymbol{E}(0, \boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{E}_0(\boldsymbol{x}, \boldsymbol{y}), \quad \boldsymbol{H}(0, \boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{H}_0(\boldsymbol{x}, \boldsymbol{y}), \tag{3}$$

and the perfectly conducting (PEC) boundary condition

$$\boldsymbol{n} \times \boldsymbol{E} = \boldsymbol{0}, \quad \text{on } \partial\Omega, \tag{4}$$

where $\boldsymbol{n}$ denotes the unit outward normal vector on $\partial\Omega$, and $\boldsymbol{E}_0$ and $\boldsymbol{H}_0$ are some given functions.

Following the standard gPC notation [37, p. 268], we let $\{\Phi_m(\boldsymbol{y})\}_{m=1}^M$ be the N-variate orthonormal polynomials of degree up to p, where $M = \begin{pmatrix} N+p \\ N \end{pmatrix}$. Note that $\{\Phi_m(\boldsymbol{y})\}_{m=1}^M$ are constructed as products of univariate polynomials in each direction $y_i, i = 1, \ldots, N$, i.e.,

$$\Phi_m(\boldsymbol{y}) = \phi_{m_1}(y_1)\phi_{m_2}(y_2)\ldots\phi_{m_N}(y_N), \quad m_1 + \cdots + m_N \leq p, \tag{5}$$

where m_i is the degree of the univariate polynomial $\phi_{m_i}(y_i)$ for $1 \leq i \leq N$. These univariate polynomials are orthonormal, i.e.,

$$\int \phi_j(y_i)\phi_k(y_i)\rho_i(y_i)dy_i = \delta_{jk}, \quad 1 \leq i, j, k \leq N, \tag{6}$$

where δ_{jk} is the Kronecker delta function and $\rho_i(y_i)$ is the probability distribution function for the random variable y_i. Note that the choice of polynomials $\phi_{m_i}(y_i)$ depends on the underlying probability density functions $\rho_i(y_i)$. For example, Hermite polynomials are associated with the Gaussian distribution, and Legendre polynomials are adopted for uniformly distributed random variables. More details can be found in [35]. Let $\rho(\boldsymbol{y}) = \Pi_{i=1}^N \rho_i(y_i)$, the N-variate basis polynomials $\{\Phi_m(\boldsymbol{y})\}_{m=1}^M$ are also orthonormal

$$\mathbb{E}[\Phi_m(\boldsymbol{y})\Phi_n(\boldsymbol{y})] := \int \Phi_m(\boldsymbol{y})\Phi_n(\boldsymbol{y})\rho(\boldsymbol{y})d\boldsymbol{y} = \delta_{mn}, \quad 1 \le m, n \le M, \tag{7}$$

We expand the solution of (1)–(2) using polynomial chaos expansions

$$\boldsymbol{E}(t, \boldsymbol{x}, \boldsymbol{y}) = \sum_{m=1}^{\infty} \boldsymbol{E}_m(t, \boldsymbol{x})\Phi_m(\boldsymbol{y}), \quad \boldsymbol{H}(t, \boldsymbol{x}, \boldsymbol{y}) = \sum_{m=1}^{\infty} \boldsymbol{H}_m(t, \boldsymbol{x})\Phi_m(\boldsymbol{y}). \tag{8}$$

Substituting (8) into (1)–(2), we obtain

$$\sum_{m=1}^{\infty} \left(\epsilon\partial_t \boldsymbol{E}_m(t, \boldsymbol{x}) - \nabla \times \boldsymbol{H}_m(t, \boldsymbol{x})\right) \Phi_m(\boldsymbol{y}) = 0, \tag{9}$$

$$\sum_{m=1}^{\infty} \left(\mu\partial_t \boldsymbol{H}_m(t, \boldsymbol{x}) + \nabla \times \boldsymbol{E}_m(t, \boldsymbol{x})\right) \Phi_m(\boldsymbol{y}) = 0. \tag{10}$$

Multiplying (9)–(10) by $\Phi_k(\boldsymbol{y})\rho(\boldsymbol{y})$ for any $k \ge 1$, and integrating the resultant, and using the orthonormality (7), we obtain

$$\sum_{m=1}^{\infty} A^{\epsilon}_{k,m}(\boldsymbol{x})\partial_t \boldsymbol{E}_m(t, \boldsymbol{x}) - \nabla \times \boldsymbol{H}_k(t, \boldsymbol{x}) = 0, \tag{11}$$

$$\sum_{m=1}^{\infty} A^{\mu}_{k,m}(\boldsymbol{x})\partial_t \boldsymbol{H}_m(t, \boldsymbol{x}) + \nabla \times \boldsymbol{E}_k(t, \boldsymbol{x}) = 0, \tag{12}$$

where

$$A^{\epsilon}_{k,m}(\boldsymbol{x}) = \int \epsilon(\boldsymbol{x}, \boldsymbol{y})\Phi_m(\boldsymbol{y})\Phi_k(\boldsymbol{y})\rho(\boldsymbol{y})dy, \quad A^{\mu}_{k,m}(\boldsymbol{x}) = \int \mu(\boldsymbol{x}, \boldsymbol{y})\Phi_m(\boldsymbol{y})\Phi_k(\boldsymbol{y})\rho(\boldsymbol{y})dy.$$

If we consider the pth-order gPC approximations of $\boldsymbol{E}$ and $\boldsymbol{H}$, i.e.,

$$\boldsymbol{E}_M(t, \boldsymbol{x}, \boldsymbol{y}) := \sum_{m=1}^{M} \widehat{\boldsymbol{E}}_m(t, \boldsymbol{x})\Phi_m(\boldsymbol{y}), \quad \boldsymbol{H}_M(t, \boldsymbol{x}, \boldsymbol{y}) := \sum_{m=1}^{M} \widehat{\boldsymbol{H}}_m(t, \boldsymbol{x})\Phi_m(\boldsymbol{y}). \tag{13}$$

The the coefficients $\widehat{\boldsymbol{E}}_m$ and $\widehat{\boldsymbol{H}}_m$ satisfy the following Galerkin system:

$$\sum_{m=1}^{M} A^{\epsilon}_{k,m}(\boldsymbol{x})\partial_t \widehat{\boldsymbol{E}}_m(t, \boldsymbol{x}) - \nabla \times \widehat{\boldsymbol{H}}_k(t, \boldsymbol{x}) = 0, \tag{14}$$

$$\sum_{m=1}^{M} A^{\mu}_{k,m}(\boldsymbol{x})\partial_t \widehat{\boldsymbol{H}}_m(t, \boldsymbol{x}) + \nabla \times \widehat{\boldsymbol{E}}_k(t, \boldsymbol{x}) = 0. \tag{15}$$

Let us denote $\widehat{\boldsymbol{E}} = \left(\widehat{\boldsymbol{E}}_1, \ldots, \widehat{\boldsymbol{E}}_M\right)'$, $\widehat{\boldsymbol{H}} = \left(\widehat{\boldsymbol{H}}_1, \ldots, \widehat{\boldsymbol{H}}_M\right)'$, and matrices $A^{\epsilon}(\boldsymbol{x}) = (A^{\epsilon}_{k,m})_{1\le k,m\le M}$ and $A^{\mu}(\boldsymbol{x}) = (A^{\mu}_{k,m})_{1\le k,m\le M}$. Then the above gPC Galerkin system can be written as

$$A^{\epsilon}(\boldsymbol{x})\partial_t \widehat{\boldsymbol{E}}(t, \boldsymbol{x}) - \nabla \times \widehat{\boldsymbol{H}}(t, \boldsymbol{x}) = 0, \tag{16}$$

$$A^{\mu}(\boldsymbol{x})\partial_t \widehat{\boldsymbol{H}}(t, \boldsymbol{x}) + \nabla \times \widehat{\boldsymbol{E}}(t, \boldsymbol{x}) = 0, \tag{17}$$

which are subject to the PEC boundary condition

$$\boldsymbol{n} \times \widehat{\boldsymbol{E}} = \boldsymbol{0}, \quad \text{on } \partial\Omega, \tag{18}$$

and the initial conditions

$$\widehat{\boldsymbol{E}}(0, \boldsymbol{x}) = \widehat{\boldsymbol{E}}_0(\boldsymbol{x}), \quad \widehat{\boldsymbol{H}}(0, \boldsymbol{x}) = \widehat{\boldsymbol{H}}_0(\boldsymbol{x}). \tag{19}$$

Here $\widehat{\boldsymbol{E}}_0(\boldsymbol{x})$ and $\widehat{\boldsymbol{H}}_0(\boldsymbol{x})$ are the gPC expansion coefficient vectors obtained by expressing the initial condition of (3) in the form of (13).

For practical applications, we shall assume that the random permittivity and permeability are bounded below and above, i.e., there exist constants ϵ_{min}, ϵ_{max}, μ_{min} and μ_{max} such that

$$0 < \epsilon_{min} \le \epsilon(\boldsymbol{x}, \boldsymbol{y}) \le \epsilon_{max}, \quad 0 < \mu_{min} \le \mu(\boldsymbol{x}, \boldsymbol{y}) \le \mu_{max}, \quad \forall \quad \boldsymbol{x} \in D, \boldsymbol{y} \in R^N. \tag{20}$$

Theorem 1 *Under the assumption (20), the matrices $A^\epsilon(\boldsymbol{x})$ and $A^\mu(\boldsymbol{x})$ are positive definite for any $\boldsymbol{x} \in D$, and satisfy the following*

$$0 < \epsilon_{min}\|\boldsymbol{u}\|^2_{L_2(D)} \le (A^\epsilon(\boldsymbol{x})\boldsymbol{u}, \boldsymbol{u}) \le \epsilon_{max}\|\boldsymbol{u}\|^2_{L_2(D)}, \tag{21}$$

$$0 < \mu_{min}\|\boldsymbol{u}\|^2_{L_2(D)} \le (A^\mu(\boldsymbol{x})\boldsymbol{u}, \boldsymbol{u}) \le \mu_{max}\|\boldsymbol{u}\|^2_{L_2(D)}, \tag{22}$$

hold true for any M dimensional non-zero vector $\boldsymbol{u}$.

Proof Let $\boldsymbol{u} = (\widehat{u}_1, \ldots, \widehat{u}_M)^T$ be an arbitrary non-zero vector, and $u(\boldsymbol{y}) = \sum_{k=1}^M \widehat{u}_k \Phi_k(\boldsymbol{y})$ be a random variable constructed by the vector $\boldsymbol{u}$. Using the definition of $A^\epsilon(\boldsymbol{x})$, we easily have: for any $\boldsymbol{x} \in D$,

$$\begin{aligned}\boldsymbol{u}^T A^\epsilon(\boldsymbol{x})\boldsymbol{u} &= \sum_{k=1}^M \sum_{m=1}^M \widehat{u}_k \int \epsilon(\boldsymbol{x}, \boldsymbol{y})\Phi_k(\boldsymbol{y})\Phi_m(\boldsymbol{y})\rho(\boldsymbol{y})d\boldsymbol{y}\widehat{u}_m \\ &= \int \epsilon(\boldsymbol{x}, \boldsymbol{y})u^2(\boldsymbol{y})\rho(\boldsymbol{y})d\boldsymbol{y} > 0,\end{aligned} \tag{23}$$

which shows the positive definiteness of $A^\epsilon(\boldsymbol{x})$. The boundness (21) is straightforward from (23) and (20). The conclusion for $A^\mu(\boldsymbol{x})$ follows the same argument. □

Furthermore, we can show that the Galerkin system of (16)–(17) satisfies the energy conservation property.

Theorem 2 *For the solution $(\widehat{\boldsymbol{E}}(t, \boldsymbol{x}), \widehat{\boldsymbol{H}}(t, \boldsymbol{x}))$ of (16)–(17) subject to the PEC boundary condition (18), the following energy identity holds true for any $t \in [0, T]$ and $k \ge 0$:*

$$\begin{aligned}&\left(\left\|A^{\epsilon/2}\partial_{t^k}\widehat{\boldsymbol{E}}\right\|^2_{L^2(D)} + \left\|A^{\mu/2}\partial_{t^k}\widehat{\boldsymbol{H}}\right\|^2_{L^2(D)}\right)\Big|_t \\ &= \left(\left\|A^{\epsilon/2}\partial_{t^k}\widehat{\boldsymbol{E}}\right\|^2_{L^2(D)} + \left\|A^{\mu/2}\partial_{t^k}\widehat{\boldsymbol{H}}\right\|^2_{L^2(D)}\right)\Big|_{t=0}.\end{aligned} \tag{24}$$

Proof Multiplying (16) and (17) by $\widehat{\boldsymbol{E}}$ and $\widehat{\boldsymbol{H}}$ and integrating over D, respectively, then summing up the resultants and using the PEC boundary condition (18), we easily see that (24) holds true for $k = 0$.

To prove (24) for any $k \ge 1$, we take the kth time derivative of (16)–(18), and follow the same step as for the $k = 0$ case. □

3 The Finite Element Time-Domain Schemes

To solve the problem (16)–(17) by a finite element method, we partition the physical domain Ω by a family of regular cubic or tetrahedral mesh T^h with maximum mesh size h, and adopt the rth ($r \geq 1$ order Raviart–Thomas–Nédélec (RTN) mixed finite element spaces U_h and $\boldsymbol{V}_h$ [20,23]: For any $r \geq 1$,

$$\begin{aligned} \boldsymbol{U}_h &= \left\{ \boldsymbol{u}_h \in H(div;\Omega) \mid \boldsymbol{u}_h|_K \in (p_{r-1})^3 \oplus \tilde{p}_{r-1}\boldsymbol{x}, \ \ \forall\, K \in T^h \right\}, \\ \boldsymbol{V}_h &= \left\{ \boldsymbol{v}_h \in H(\mathrm{curl};\Omega) \mid \boldsymbol{v}_h|_K \in (p_{r-1})^3 \oplus S_r, \ \ \forall\, K \in T^h \right\}, \\ S_r &= \left\{ \vec{p} \in (\tilde{p}_r)^3, \boldsymbol{x}\cdot\vec{p} = 0 \right\}, \end{aligned}$$

or RTN cubic elements:

$$\begin{aligned} \boldsymbol{U}_h &= \left\{ \boldsymbol{u}_h \in H(div;\Omega) \mid \boldsymbol{u}_h|_K \in Q_{r,r-1,r-1} \times Q_{r-1,r,r-1} \times Q_{r-1,r-1,r}, \ \ \forall\, K \in T^h \right\}, \\ \boldsymbol{V}_h &= \left\{ \boldsymbol{v}_h \in H(\mathrm{curl};\Omega) \mid \boldsymbol{v}_h|_K \in Q_{r-1,r,r} \times Q_{r,r-1,r} \times Q_{r,r,r-1}, \ \ \forall\, K \in T^h \right\}. \end{aligned}$$

Here $\tilde{p}_r$ denotes the space of homogeneous polynomials of degree r, and $Q_{i,j,k}$ denotes the space of polynomials whose degrees are less than or equal to i, j, k in variables x, y, z, respectively. To impose the PEC boundary condition, we denote $\boldsymbol{V}_h^0 = \{\boldsymbol{v} \in \boldsymbol{V}_h : \boldsymbol{v} \times \boldsymbol{n} = 0 \ \text{ on } \partial\Omega\}$.

3.1 The Semi-discrete Scheme and Its Analysis

Let us first consider a semi-discrete scheme for the Galerkin system: find $\widehat{\boldsymbol{E}}_h \in (\boldsymbol{V}_h^0)^M$, $\widehat{\boldsymbol{H}}_h \in (U_h)^M$ such that

$$(A^{\epsilon}\partial_t \widehat{\boldsymbol{E}}_h, \phi_{E,h})_D - (\widehat{\boldsymbol{H}}_h, \nabla \times \phi_{E,h})_D = 0, \ \ \forall\, \phi_{E,h} \in (\boldsymbol{V}_h^0)^M, \tag{25}$$

$$(A^{\mu}\partial_t \widehat{\boldsymbol{H}}_h, \phi_{H,h})_D + (\nabla \times \widehat{\boldsymbol{E}}_h, \phi_{H,h})_D = 0, \ \ \forall\, \phi_{H,h} \in (\boldsymbol{U}_h)^M, \tag{26}$$

subject to the initial conditions

$$\widehat{\boldsymbol{E}}_h(0,\boldsymbol{x}) = \Pi_h^c \widehat{\boldsymbol{E}}_0(\boldsymbol{x}), \quad \widehat{\boldsymbol{H}}_h(0,\boldsymbol{x}) = \Pi_h^d \widehat{\boldsymbol{H}}_0(\boldsymbol{x}), \tag{27}$$

where we denote Π_h^c for the Nédélec interpolation operator and Π_h^d for the L_2 projection into the space $\boldsymbol{U}_h$. It is known that [20,23]:

$$\begin{aligned} &\|\boldsymbol{u} - \Pi_h^c \boldsymbol{u}\|_{L_2(D)} + \|\nabla \times (\boldsymbol{u} - \Pi_h^c \boldsymbol{u})\|_{L_2(D)} \leq Ch^r \|\boldsymbol{u}\|_{H^r(curl;D)}, \\ &\quad \forall\, \boldsymbol{u} \in H^r(curl;D), \end{aligned} \tag{28}$$

$$\|\boldsymbol{v} - \Pi_h^d \boldsymbol{v}\|_{L_2(D)} \leq Ch^r \|\boldsymbol{u}\|_{H^r(D)}, \ \ \forall\, \boldsymbol{v} \in H^r(D). \tag{29}$$

Below we provide the error estimate for the semi-discrete scheme (25)–(27). Let $\boldsymbol{E}(t,\boldsymbol{x},\boldsymbol{y})$ and $\boldsymbol{H}(t,\boldsymbol{x},\boldsymbol{y})$ be the analytical solution of (1)–(2) subject to the initial conditions (3) and the PEC boundary condition (4), and $\boldsymbol{E}_h(t,\boldsymbol{x},\boldsymbol{y})$ and $\boldsymbol{H}_h(t,\boldsymbol{x},\boldsymbol{y})$ be the numerical solution

$$\boldsymbol{E}_h(t,\boldsymbol{x},\boldsymbol{y}) = \sum_{m=1}^{M} \widehat{\boldsymbol{E}}_{m,h}^t \Phi_m(\boldsymbol{y}), \ \ \boldsymbol{H}_h(t,\boldsymbol{x},\boldsymbol{y}) = \sum_{m=1}^{M} \widehat{\boldsymbol{H}}_{m,h}^t \Phi_m(\boldsymbol{y}), \tag{30}$$

where $\widehat{\boldsymbol{E}}_{m,h}$ and $\widehat{\boldsymbol{H}}_{m,h}$ are the mth component of $\widehat{\boldsymbol{E}}_h(t,\boldsymbol{x})$ and $\widehat{\boldsymbol{H}}_h(t,\boldsymbol{x})$ of (25)–(27). We consider the errors

$$\boldsymbol{E}-\boldsymbol{E}_h=(\boldsymbol{E}-\boldsymbol{E}_M)+(\boldsymbol{E}_M-\boldsymbol{E}_h),\quad \boldsymbol{H}-\boldsymbol{H}_h=(\boldsymbol{H}-\boldsymbol{H}_M)+(\boldsymbol{H}_M-\boldsymbol{H}_h),\tag{31}$$

where $\boldsymbol{E}_M$ and $\boldsymbol{H}_M$ are the gPC approximations given in (13). We shall will show that the error bound is optimal and the error grows only linearly in time.

Theorem 3 *Denote the M-dimensional vectors R^E and R^H with kth components given by*

$$R_k^E=\sum_{m=M+1}^{\infty}A_{k,m}^{\epsilon}(\boldsymbol{x})\partial_t\boldsymbol{E}_m(t,\boldsymbol{x}),\ \ R_k^H=\sum_{m=M+1}^{\infty}A_{k,m}^{\mu}(\boldsymbol{x})\partial_t\boldsymbol{H}_m(t,\boldsymbol{x}),\ 1\le k\le M.\tag{32}$$

Then we have the optimal error estimate: for any $t\in[0,T]$,

$$\begin{aligned}
&\left(\mathbb{E}\left[\|\boldsymbol{E}-\boldsymbol{E}_h\|_{L_2(D)}^2\right]\right)^{1/2}+\left(\mathbb{E}\left[\|\boldsymbol{H}-\boldsymbol{H}_h\|_{L_2(D)}^2\right]\right)^{1/2}\\
&\le C\max_{0\le t\le T}\left(\sum_{m=M+1}^{\infty}\|\boldsymbol{E}_m(t,\boldsymbol{x})\|_{L_2(D)}^2+\sum_{m=M+1}^{\infty}\|\boldsymbol{H}_m(t,\boldsymbol{x})\|_{L_2(D)}^2\right)^{1/2}\\
&\quad+CTh^r\max_{0\le t\le T}\left(\left\|\partial_t\widehat{\boldsymbol{E}}\right\|_{H^r(curl;D)}^2+\left\|\widehat{\boldsymbol{E}}\right\|_{H^r(curl;D)}^2+\left\|\partial_t\widehat{\boldsymbol{H}}\right\|_{H^r(D)}^2+\left\|\widehat{\boldsymbol{H}}\right\|_{H^r(D)}^2\right)^{1/2}\\
&\quad+CT\max_{0\le t\le T}\left(\left\|A^{-\epsilon/2}R^E\right\|_{L_2(D)}^2+\left\|A^{-\mu/2}R^H\right\|_{L_2(D)}^2\right)^{1/2},
\end{aligned}\tag{33}$$

where the constant $C>0$ is independent of T and h.

Proof The proof is divided into two major parts, corresponding to bounds for the gPC approximation error and the semi-discretization error, respectively.

(I) By the expansions (8) and (13), and the orthonormality condition of $\Phi_n(\boldsymbol{y})$, we easily see that the mean of the gPC approximation error

$$\begin{aligned}
\mathbb{E}\left[\|\boldsymbol{E}-\boldsymbol{E}_M\|_{L_2(D)}^2\right]&:=\int\|\boldsymbol{E}-\boldsymbol{E}_M\|_{L_2(D)}^2\,\rho(\boldsymbol{y})d\boldsymbol{y}\\
&=\int\int_D\left|\sum_{m=1}^{M}(\boldsymbol{E}_m-\widehat{\boldsymbol{E}}_m)\Phi_m(\boldsymbol{y})+\sum_{m=M+1}^{\infty}\boldsymbol{E}_m\Phi_m(\boldsymbol{y})\right|^2d\boldsymbol{x}\,\rho(\boldsymbol{y})d\boldsymbol{y}\\
&=\int_D\int\left[\left|\sum_{m=1}^{M}(\boldsymbol{E}_m-\widehat{\boldsymbol{E}}_m)\Phi_m(\boldsymbol{y})\right|^2+\left|\sum_{m=M+1}^{\infty}\boldsymbol{E}_m\Phi_m(\boldsymbol{y})\right|^2\right]\rho(\boldsymbol{y})d\boldsymbol{y}d\boldsymbol{x}\\
&=\sum_{m=1}^{M}\left\|\boldsymbol{E}_m(t,\boldsymbol{x})-\widehat{\boldsymbol{E}}_m(t,\boldsymbol{x})\right\|_{L_2(D)}^2+\sum_{m=M+1}^{\infty}\|\boldsymbol{E}_m(t,\boldsymbol{x})\|_{L_2(D)}^2.
\end{aligned}\tag{34}$$

By the same argument, we have

$$\mathbb{E}\left[\|\boldsymbol{H}-\boldsymbol{H}_M\|_{L_2(D)}^2\right]=\sum_{m=1}^{M}\left\|\boldsymbol{H}_m(t,\boldsymbol{x})-\widehat{\boldsymbol{H}}_m(t,\boldsymbol{x})\right\|_{L_2(D)}^2+\sum_{m=M+1}^{\infty}\|\boldsymbol{H}_m(t,\boldsymbol{x})\|_{L_2(D)}^2.\tag{35}$$

To investigate the error $\boldsymbol{E}-\boldsymbol{E}_M$ and $\boldsymbol{H}-\boldsymbol{H}_M$, let us introduce $\widetilde{\boldsymbol{E}}=(\boldsymbol{E}_1,\ldots,\boldsymbol{E}_M)'$ and $\widetilde{\boldsymbol{H}}=(\boldsymbol{H}_1,\ldots,\boldsymbol{H}_M)'$, where $\boldsymbol{E}_i$ and $\boldsymbol{H}_i$ are the coefficients in the expansion (8). From (11)–(12), we see that $\widetilde{\boldsymbol{E}}$ and $\widetilde{\boldsymbol{H}}$ satisfy the following equations

$$A^{\epsilon}(\boldsymbol{x})\partial_t \widetilde{\boldsymbol{E}}(t,\boldsymbol{x}) - \nabla \times \widetilde{\boldsymbol{H}}(t,\boldsymbol{x}) = R^E, \tag{36}$$
$$A^{\mu}(\boldsymbol{x})\partial_t \widetilde{\boldsymbol{H}}(t,\boldsymbol{x}) + \nabla \times \widetilde{\boldsymbol{E}}(t,\boldsymbol{x}) = R^H, \tag{37}$$

subject to the PEC boundary condition

$$\boldsymbol{n} \times \widetilde{\boldsymbol{E}} = \boldsymbol{0}, \quad \text{on } \partial\Omega, \tag{38}$$

and the initial conditions

$$\widetilde{\boldsymbol{E}}(0,\boldsymbol{x}) = \widetilde{\boldsymbol{E}}_0(\boldsymbol{x}), \quad \widetilde{\boldsymbol{H}}(0,\boldsymbol{x}) = \widetilde{\boldsymbol{H}}_0(\boldsymbol{x}), \tag{39}$$

where $\widetilde{\boldsymbol{E}}_0(\boldsymbol{x})$ and $\widetilde{\boldsymbol{H}}_0(\boldsymbol{x})$ are the gPC expansion coefficient vectors obtained by expressing the initial condition of (3) in the form of (8). Moreover, the kth components of R^E and R^H are given by (32).

By (16)–(17) and (36)–(37), we can obtain the following error equations in the weak form:

$$\begin{aligned}
&\left(A^{\epsilon}\partial_t(\widetilde{\boldsymbol{E}} - \widehat{\boldsymbol{E}}), \phi_E\right)_D - \left(\widetilde{\boldsymbol{H}} - \widehat{\boldsymbol{H}}, \nabla \times \phi_E\right)_D = \left(R^E, \phi_E\right)_D, \\
&\quad \forall \phi_E \in \left(H_0(curl; D)\right)^M,
\end{aligned} \tag{40}$$

$$\begin{aligned}
&\left(A^{\mu}\partial_t(\widetilde{\boldsymbol{H}} - \widehat{\boldsymbol{H}}), \phi_H\right)_D + \left(\nabla \times (\widetilde{\boldsymbol{E}} - \widehat{\boldsymbol{E}}), \phi_H\right)_D = \left(R^H, \phi_H\right)_D, \\
&\quad \forall \phi_H \in \left(H(div; D)\right)^M,
\end{aligned} \tag{41}$$

subject to the PEC boundary condition

$$\boldsymbol{n} \times (\widetilde{\boldsymbol{E}} - \widehat{\boldsymbol{E}}) = \boldsymbol{0}, \quad \text{on } \partial\Omega, \tag{42}$$

and the initial conditions

$$(\widetilde{\boldsymbol{E}} - \widehat{\boldsymbol{E}})(0,\boldsymbol{x}) = \boldsymbol{0}, \quad (\widetilde{\boldsymbol{H}} - \widehat{\boldsymbol{H}})(0,\boldsymbol{x}) = \boldsymbol{0}. \tag{43}$$

Choosing $\phi_E = 2(\widetilde{\boldsymbol{E}} - \widehat{\boldsymbol{E}})(t,\boldsymbol{x})$ and $\phi_H = 2(\widetilde{\boldsymbol{H}} - \widehat{\boldsymbol{H}})(t,\boldsymbol{x})$ in (40) and (41), respectively, adding the resultants together, and using the Cauchy–Schwarz inequality, we have

$$\begin{aligned}
&\frac{d}{dt}\left(\|A^{\epsilon/2}(\widetilde{\boldsymbol{E}} - \widehat{\boldsymbol{E}})\|^2_{L_2(D)} + \|A^{\mu/2}(\widetilde{\boldsymbol{H}} - \widehat{\boldsymbol{H}})\|^2_{L_2(D)}\right) \\
&\quad = 2\left(R^E, \widetilde{\boldsymbol{E}} - \widehat{\boldsymbol{E}}\right)_D + 2\left(R^H, \widetilde{\boldsymbol{H}} - \widehat{\boldsymbol{H}}\right)_D \\
&\quad \le \delta\left(\|A^{\epsilon/2}(\widetilde{\boldsymbol{E}} - \widehat{\boldsymbol{E}})\|^2_{L_2(D)} + \|A^{\mu/2}(\widetilde{\boldsymbol{H}} - \widehat{\boldsymbol{H}})\|^2_{L_2(D)}\right) \\
&\qquad + \frac{1}{\delta}\left(\left\|A^{-\epsilon/2}R^E\right\|^2_{L_2(D)} + \left\|A^{-\mu/2}R^H\right\|^2_{L_2(D)}\right).
\end{aligned}$$

Integrating the above inequality from $t = 0$ and any $t \le T$ and taking the maximum of right hand side with respect to $t \in [0,T]$, we obtain

$$\begin{aligned}
&\left(\|A^{\epsilon/2}(\widetilde{\boldsymbol{E}} - \widehat{\boldsymbol{E}})\|^2_{L_2(D)} + \|A^{\mu/2}(\widetilde{\boldsymbol{H}} - \widehat{\boldsymbol{H}})\|^2_{L_2(D)}\right)(t) \\
&\quad \le \delta T \max_{0\le t\le T}\left(\|A^{\epsilon/2}(\widetilde{\boldsymbol{E}} - \widehat{\boldsymbol{E}})\|^2_{L_2(D)} + \|A^{\mu/2}(\widetilde{\boldsymbol{H}} - \widehat{\boldsymbol{H}})\|^2_{L_2(D)}\right) \\
&\qquad + \frac{T}{\delta}\max_{0\le t\le T}\left(\left\|A^{-\epsilon/2}R^E\right\|^2_{L_2(D)} + \left\|A^{-\mu/2}R^H\right\|^2_{L_2(D)}\right).
\end{aligned} \tag{44}$$

Taking the maximum of left hand side with respect to $t \in [0, T]$, and choosing δ such that $\delta = \frac{1}{2T}$, we obtain

$$\max_{0\le t\le T}\left(\|A^{\epsilon/2}(\widetilde{\boldsymbol{E}}-\widehat{\boldsymbol{E}})\|^2_{L_2(D)}+\|A^{\mu/2}(\widetilde{\boldsymbol{H}}-\widehat{\boldsymbol{H}})\|^2_{L_2(D)}\right)$$
$$\le 4T^2\max_{0\le t\le T}\left(\left\|A^{-\epsilon/2}R^E\right\|^2_{L_2(D)}+\left\|A^{-\mu/2}R^H\right\|^2_{L_2(D)}\right), \tag{45}$$

which leads to

$$\max_{0\le t\le T}\left(\|A^{\epsilon/2}(\widetilde{\boldsymbol{E}}-\widehat{\boldsymbol{E}})\|^2_{L_2(D)}+\|A^{\mu/2}(\widetilde{\boldsymbol{H}}-\widehat{\boldsymbol{H}})\|^2_{L_2(D)}\right)^{1/2}$$
$$\le CT\max_{0\le t\le T}\left(\left\|A^{-\epsilon/2}R^E\right\|^2_{L_2(D)}+\left\|A^{-\mu/2}R^H\right\|^2_{L_2(D)}\right)^{1/2}. \tag{46}$$

(II) Multiplying (16)–(17) by $\phi_{E,h}$ and $\phi_{H,h}$ and integrating over D, we obtain

$$\left(A^{\epsilon}\partial_t\widehat{\boldsymbol{E}},\phi_{E,h}\right)_D-\left(\widehat{\boldsymbol{H}},\nabla\times\phi_{E,h}\right)_D=0,\quad\forall\,\phi_{E,h}\in(\boldsymbol{V}_h^0)^M, \tag{47}$$
$$\left(A^{\mu}\partial_t\widehat{\boldsymbol{H}},\phi_{H,h}\right)_D+\left(\nabla\times\widehat{\boldsymbol{E}},\phi_{H,h}\right)_D=0,\quad\forall\,\phi_{H,h}\in(\boldsymbol{U}_h)^M. \tag{48}$$

Subtracting (25)–(26) from (47)–(48), we obtain the error equations:

$$\left(A^{\epsilon}\partial_t(\widehat{\boldsymbol{E}}-\widehat{\boldsymbol{E}}_h),\phi_{E,h}\right)_D-\left(\widehat{\boldsymbol{H}}-\widehat{\boldsymbol{H}}_h,\nabla\times\phi_{E,h}\right)_D=0,\quad\forall\,\phi_{E,h}\in(\boldsymbol{V}_h^0)^M, \tag{49}$$
$$\left(A^{\mu}\partial_t(\widehat{\boldsymbol{H}}-\widehat{\boldsymbol{H}}_h),\phi_{H,h}\right)_D+\left(\nabla\times(\widehat{\boldsymbol{E}}-\widehat{\boldsymbol{E}}_h),\phi_{H,h}\right)_D=0,\quad\forall\,\phi_{H,h}\in(\boldsymbol{U}_h)^M. \tag{50}$$

Let us introduce the short notations

$$\widehat{\boldsymbol{E}}_I:=\Pi_h^c\widehat{\boldsymbol{E}},\quad\widehat{\boldsymbol{H}}_I:=\Pi_h^d\widehat{\boldsymbol{H}}.$$

Choosing $\phi_{E,h}=2(\widehat{\boldsymbol{E}}_I-\widehat{\boldsymbol{E}}_h)$ and $\phi_{H,h}=2(\widehat{\boldsymbol{H}}_I-\widehat{\boldsymbol{H}}_h)$ in (49)–(50), respectively, we have

$$\begin{aligned}
&\frac{d}{dt}\left(\|A^{\epsilon/2}(\widehat{\boldsymbol{E}}_I-\widehat{\boldsymbol{E}}_h)\|^2_{L_2(D)}+\|A^{\mu/2}(\widehat{\boldsymbol{H}}_I-\widehat{\boldsymbol{H}}_h)\|^2_{L_2(D)}\right)\\
&=2\left(A^{\epsilon}\partial_t(\widehat{\boldsymbol{E}}_I-\widehat{\boldsymbol{E}}),\widehat{\boldsymbol{E}}_I-\widehat{\boldsymbol{E}}_h\right)_D+2\left(\widehat{\boldsymbol{H}}-\widehat{\boldsymbol{H}}_I,\nabla\times(\widehat{\boldsymbol{E}}_I-\widehat{\boldsymbol{E}}_h)\right)_D\\
&\quad+2\left(A^{\mu}\partial_t(\widehat{\boldsymbol{H}}_I-\widehat{\boldsymbol{H}}),\widehat{\boldsymbol{H}}_I-\widehat{\boldsymbol{H}}_h\right)_D-2\left(\nabla\times(\widehat{\boldsymbol{E}}-\widehat{\boldsymbol{E}}_I),\widehat{\boldsymbol{H}}_I-\widehat{\boldsymbol{H}}_h\right)_D\\
&\le A^{\epsilon}_{max}\left(\frac{1}{\delta}\|\partial_t(\widehat{\boldsymbol{E}}_I-\widehat{\boldsymbol{E}})\|^2_{L_2(D)}+\delta\|\widehat{\boldsymbol{E}}_I-\widehat{\boldsymbol{E}}_h\|^2_{L_2(D)}\right)\\
&\quad+A^{\mu}_{max}\left(\frac{1}{\delta}\|\partial_t(\widehat{\boldsymbol{H}}_I-\widehat{\boldsymbol{H}})\|^2_{L_2(D)}+\delta\|\widehat{\boldsymbol{H}}_I-\widehat{\boldsymbol{H}}_h\|^2_{L_2(D)}\right)\\
&\quad+\frac{1}{\delta}\|\nabla\times(\widehat{\boldsymbol{E}}-\widehat{\boldsymbol{E}}_I)\|^2_{L_2(D)}+\delta\|\widehat{\boldsymbol{H}}_I-\widehat{\boldsymbol{H}}_h\|^2_{L_2(D)}\\
&\le A^{\epsilon}_{max}\left(\frac{Ch^{2r}}{\delta}\|\partial_t\widehat{\boldsymbol{E}}\|^2_{H^r(curl;D)}+\delta\|\widehat{\boldsymbol{E}}_I-\widehat{\boldsymbol{E}}_h\|^2_{L_2(D)}\right)\\
&\quad+A^{\mu}_{max}\left(\frac{Ch^{2r}}{\delta}\|\partial_t\widehat{\boldsymbol{H}}\|^2_{H^r(D)}+\delta\|\widehat{\boldsymbol{H}}_I-\widehat{\boldsymbol{H}}_h\|^2_{L_2(D)}\right)\\
&\quad+\frac{Ch^{2r}}{\delta}\|\widehat{\boldsymbol{E}}\|^2_{H^r(curl;D)}+\delta\|\widehat{\boldsymbol{H}}_I-\widehat{\boldsymbol{H}}_h\|^2_{L_2(D)},
\end{aligned} \tag{51}$$

where we used the fact that $\nabla\times(\widehat{\boldsymbol{E}}_I-\widehat{\boldsymbol{E}}_h)\in\boldsymbol{U}_h$, the Cauchy–Schwarz inequality, and the interpolation and projection error estimates (28)–(29).

Integrating (51) from $t = 0$ to any $t \leq T$, and taking the maximum of the right hand side with respect to t, we obtain

$$\begin{aligned}
&\left(\|A^{\epsilon/2}(\widehat{\boldsymbol{E}}_I - \widehat{\boldsymbol{E}}_h)\|^2_{L_2(D)} + \|A^{\mu/2}(\widehat{\boldsymbol{H}}_I - \widehat{\boldsymbol{H}}_h)\|^2_{L_2(D)}\right)(t) \\
&\quad \leq \left(\|A^{\epsilon/2}(\widehat{\boldsymbol{E}}_I - \widehat{\boldsymbol{E}}_h)\|^2_{L_2(D)} + \|A^{\mu/2}(\widehat{\boldsymbol{H}}_I - \widehat{\boldsymbol{H}}_h)\|^2_{L_2(D)}\right)(0) \\
&\qquad + \frac{TCh^{2r}}{\delta} \max_{0\leq t\leq T} \left(\|\partial_t \widehat{\boldsymbol{E}}\|^2_{H^r(curl;D)} + \|\partial_t \widehat{\boldsymbol{H}}\|^2_{H^r(D)} + \|\widehat{\boldsymbol{E}}\|^2_{H^r(curl;D)}\right) \\
&\qquad + C\delta T \max_{0\leq t\leq T} \left(\|\widehat{\boldsymbol{E}}_I - \widehat{\boldsymbol{E}}_h\|^2_{L_2(D)} + \|\widehat{\boldsymbol{H}}_I - \widehat{\boldsymbol{H}}_h\|^2_{L_2(D)}\right).
\end{aligned} \tag{52}$$

Noting that the first term on the right hand side of (52) is zero due to (27), then taking the maximum of the left hand side with respect to t, and choosing δ such that $\delta = \frac{1}{2CT}$, we have

$$\begin{aligned}
&\max_{0\leq t\leq T} \left(\|A^{\epsilon/2}(\widehat{\boldsymbol{E}}_I - \widehat{\boldsymbol{E}}_h)\|^2_{L_2(D)} + \|A^{\mu/2}(\widehat{\boldsymbol{H}}_I - \widehat{\boldsymbol{H}}_h)\|^2_{L_2(D)}\right)^{1/2} \\
&\quad \leq CTh^r \max_{0\leq t\leq T} \left(\|\partial_t \widehat{\boldsymbol{E}}\|^2_{H^r(curl;D)} + \|\partial_t \widehat{\boldsymbol{H}}\|^2_{H^r(D)} + \|\widehat{\boldsymbol{E}}\|^2_{H^r(curl;D)}\right)^{1/2}.
\end{aligned} \tag{53}$$

Using the interpolation and projection error estimates (28)–(29) and the triangle inequality, from (53) we have

$$\begin{aligned}
&\max_{0\leq t\leq T} \left(\|A^{\epsilon/2}(\widehat{\boldsymbol{E}} - \widehat{\boldsymbol{E}}_h)\|^2_{L_2(D)} + \|A^{\mu/2}(\widehat{\boldsymbol{H}} - \widehat{\boldsymbol{H}}_h)\|^2_{L_2(D)}\right)^{1/2} \\
&\quad \leq CTh^r \max_{0\leq t\leq T} \left(\|\partial_t \widehat{\boldsymbol{E}}\|^2_{H^r(curl;D)} + \|\widehat{\boldsymbol{E}}\|^2_{H^r(curl;D)} + \|\partial_t \widehat{\boldsymbol{H}}\|^2_{H^r(D)} + \|\widehat{\boldsymbol{H}}\|^2_{H^r(D)}\right)^{1/2}.
\end{aligned} \tag{54}$$

(III) By the error definition (31) and the obtained error estimates (46) and (54), we conclude the proof of (33). □

Remark 1 For any given small number $\epsilon_M > 0$, under the assumption that there exists a sufficiently large M in (13) so that

$$\begin{aligned}
&\max_{0\leq t\leq T} \Bigg[\left(\sum_{m=M+1}^{\infty} \|\boldsymbol{E}_m(t,\boldsymbol{x})\|^2_{L_2(D)} + \sum_{m=M+1}^{\infty} \|\boldsymbol{H}_m(t,\boldsymbol{x})\|^2_{L_2(D)}\right)^{1/2} \\
&\qquad + \left(\left\|A^{-\epsilon/2}R^E\right\|^2_{L_2(D)} + \left\|A^{-\mu/2}R^H\right\|^2_{L_2(D)}\right)^{1/2}\Bigg] \leq \epsilon_M,
\end{aligned} \tag{55}$$

and the solutions $(\widehat{\boldsymbol{E}}, \widehat{\boldsymbol{H}})$ of (16)–(17) are smooth enough and bounded above:

$$\max_{0\leq t\leq T} \left(\|\partial_t \widehat{\boldsymbol{E}}\|^2_{H^r(curl;D)} + \|\widehat{\boldsymbol{E}}\|^2_{H^r(curl;D)} + \|\partial_t \widehat{\boldsymbol{H}}\|^2_{H^r(D)} + \|\widehat{\boldsymbol{H}}\|^2_{H^r(D)}\right)^{1/2} \leq C, \tag{56}$$

then the optimal error estimate (33) becomes

$$\left(\mathbb{E}\left[\|\boldsymbol{E} - \boldsymbol{E}_h\|^2_{L_2(D)}\right]\right)^{1/2} + \left(\mathbb{E}\left[\|\boldsymbol{H} - \boldsymbol{H}_h\|^2_{L_2(D)}\right]\right)^{1/2} \leq C(T+1)\epsilon_M + CTh^r. \tag{57}$$

Of course, rigorous conditions for our assumption (55) to hold true are unclear since they involve dedicate regularity assumptions for the solution and the underlying polynomial basis. Such issues have been explored for stochastic Helmholtz equation [5] and stochastic Darcy's equation [12].

3.2 The Fully-Discrete Schemes

To construct a fully discrete finite element scheme, we assume that the time interval $[0, T]$ is partitioned uniformly into $0 = t_0 < t_1 < \cdots < t_{N_t} = T$, where $t_i = i\tau, i = 0, \ldots, N_t$, and $\tau = \frac{T}{N_t}$ denotes the time step size. Furthermore, we introduce the following backward difference operators: For any discrete time solution u^n,

$$\delta_\tau u^{n+1} := \frac{u^{n+1} - u^n}{\tau}, \quad \delta_\tau^2 u^{n+1} := \delta_\tau(\delta_\tau u^{n+1}) = \frac{\delta_\tau u^{n+1} - \delta_\tau u^n}{\tau} = \frac{u^{n+1} - 2u^n + u^{n-1}}{\tau^2},$$
$$\delta_\tau^{k+1} u^{n+1} := \delta_\tau(\delta_\tau^k u^{n+1}), \quad k \geq 1.$$

Crank–Nicolson Scheme We first consider a Crank–Nicolson scheme (Scheme 1) for solving (16)–(17): given proper initial approximations $\widehat{\boldsymbol{H}}_h^0 \in \boldsymbol{U}_h$ and $\widehat{\boldsymbol{E}}_h^0 \in \boldsymbol{V}_h^0$, for $k \geq 0$ find $\widehat{\boldsymbol{E}}_h^{k+1} \in \boldsymbol{V}_h^0$ and $\widehat{\boldsymbol{H}}_h^{k+1} \in \boldsymbol{U}_h$ such that

$$\left(A^\epsilon \frac{\widehat{\boldsymbol{E}}_h^{k+1} - \widehat{\boldsymbol{E}}_h^k}{\tau}, \phi_{E,h} \right)_D - \left(\frac{\widehat{\boldsymbol{H}}_h^{k+1} + \widehat{\boldsymbol{H}}_h^k}{2}, \nabla \times \phi_{E,h} \right)_D = 0, \quad \forall\, \phi_{E,h} \in \boldsymbol{V}_h^0, \tag{58}$$

$$\left(A^\mu \frac{\widehat{\boldsymbol{H}}_h^{k+1} - \widehat{\boldsymbol{H}}_h^k}{\tau}, \phi_{H,h} \right)_D + \left(\nabla \times \frac{\widehat{\boldsymbol{E}}_h^{k+1} + \widehat{\boldsymbol{E}}_h^k}{2}, \phi_{H,h} \right)_D = 0, \quad \forall\, \phi_{H,h} \in \boldsymbol{U}_h. \tag{59}$$

For the fully discrete Crank–Nicolson scheme (58)–(59), we can prove the following energy conservation property as that given in Theorem 2 for the continuous case.

Theorem 4 *For any $k \geq 0$, the solution $\left(\widehat{\boldsymbol{E}}_h^{k+1}, \widehat{\boldsymbol{H}}_h^{k+1}\right)$ of (58)–(59) satisfies the energy identity:*

$$\left\| A^{\epsilon/2} \widehat{\boldsymbol{E}}_h^{k+1} \right\|_{L^2(D)}^2 + \left\| A^{\mu/2} \widehat{\boldsymbol{H}}_h^{k+1} \right\|_{L^2(D)}^2 = \left\| A^{\epsilon/2} \widehat{\boldsymbol{E}}_h^{0} \right\|_{L^2(D)}^2 + \left\| A^{\mu/2} \widehat{\boldsymbol{H}}_h^{0} \right\|_{L^2(D)}^2. \tag{60}$$

Proof Choosing $\phi_{E,h} = \tau\left(\widehat{\boldsymbol{E}}_h^{k+1} + \widehat{\boldsymbol{E}}_h^k\right)$ and $\phi_{H,h} = \tau\left(\widehat{\boldsymbol{H}}_h^{k+1} + \widehat{\boldsymbol{H}}_h^k\right)$ in (58) and (59), respectively, then adding the resultants together, we obtain

$$\begin{aligned} &\left(\left\| A^{\epsilon/2} \widehat{\boldsymbol{E}}_h^{k+1} \right\|_{L^2(D)}^2 - \left\| A^{\epsilon/2} \widehat{\boldsymbol{E}}_h^{k} \right\|_{L^2(D)}^2 \right) \\ &+ \left(\left\| A^{\mu/2} \widehat{\boldsymbol{H}}_h^{k+1} \right\|_{L^2(D)}^2 - \left\| A^{\mu/2} \widehat{\boldsymbol{H}}_h^{k} \right\|_{L^2(D)}^2 \right) = 0, \end{aligned} \tag{61}$$

from which we conclude the proof. □

Leap-Frog Scheme Due to the high computational cost of the Crank–Nicolson scheme, we now construct the following leap-frog type scheme (Scheme 2) for solving the Galerkin system (16)–(17): given proper initial approximations $\widehat{\boldsymbol{H}}_h^0 \in \boldsymbol{U}_h$ and $\widehat{\boldsymbol{E}}_h^{-\frac{1}{2}} \in \boldsymbol{V}_h^0$, for $k \geq 0$ find $\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}} \in \boldsymbol{V}_h^0$ and $\widehat{\boldsymbol{H}}_h^{k+1} \in \boldsymbol{U}_h$ such that

$$\left(A^\epsilon \frac{\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}} - \widehat{\boldsymbol{E}}_h^{k-\frac{1}{2}}}{\tau}, \phi_{E,h} \right)_D - \left(\widehat{\boldsymbol{H}}_h^k, \nabla \times \phi_{E,h} \right)_D = 0, \quad \forall\, \phi_{E,h} \in \boldsymbol{V}_h^0, \tag{62}$$

$$\left(A^\mu \frac{\widehat{\boldsymbol{H}}_h^{k+1} - \widehat{\boldsymbol{H}}_h^k}{\tau}, \phi_{H,h} \right)_D + \left(\nabla \times \widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}, \phi_{H,h} \right)_D = 0, \quad \forall\, \phi_{H,h} \in \boldsymbol{U}_h. \tag{63}$$

Notice that the above leap-frog scheme decouples the original coupled system by first solving for $\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}$ through (62), and then solving $\widehat{\boldsymbol{H}}_h^{k+1}$ from (63). We like to remark that the leap-frog scheme does not conserve the energy anymore due to time staggering, but the scheme is conditionally stable as shown below.

Theorem 5 *Let $C_{inv} > 0$ be the constant appearing in the standard inverse estimate*

$$\|\nabla \times \boldsymbol{u}_h\|_{L^2(D)} \leq C_{inv} h^{-1} \|\boldsymbol{u}_h\|_{L^2(D)}. \tag{64}$$

Then under the time step constraint

$$\tau \leq (\epsilon_{min}\mu_{min})^{1/2} h / C_{inv}, \tag{65}$$

for any $k \geq 0$, the solution $\left(\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}, \widehat{\boldsymbol{H}}_h^{k+1}\right)$ of (62)–(63) satisfies the energy stability:

$$\left\|A^{\epsilon/2}\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}\right\|_{L^2(D)}^2 + \left\|A^{\mu/2}\widehat{\boldsymbol{H}}_h^{k+1}\right\|_{L^2(D)}^2 \leq 3\left[\left\|A^{\epsilon/2}\widehat{\boldsymbol{E}}_h^{-\frac{1}{2}}\right\|_{L^2(D)}^2 + \left\|A^{\mu/2}\widehat{\boldsymbol{H}}_h^{0}\right\|_{L^2(D)}^2\right]. \tag{66}$$

Proof In (62) and (63), we choose

$$\phi_{E,h} = \tau\left(\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}} + \widehat{\boldsymbol{E}}_h^{k-\frac{1}{2}}\right), \quad \phi_{H,h} = \tau\left(\widehat{\boldsymbol{H}}_h^{k+1} + \widehat{\boldsymbol{H}}_h^{k}\right).$$

Then summing up the resultants, we have

$$\begin{aligned}
&\left(\left\|A^{\epsilon/2}\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}\right\|_{L^2(D)}^2 - \left\|A^{\epsilon/2}\widehat{\boldsymbol{E}}_h^{k-\frac{1}{2}}\right\|_{L^2(D)}^2\right) + \left(\left\|A^{\mu/2}\widehat{\boldsymbol{H}}_h^{k+1}\right\|_{L^2(D)}^2 - \left\|A^{\mu/2}\widehat{\boldsymbol{H}}_h^{k}\right\|_{L^2(D)}^2\right) \\
&= \tau\left(\widehat{\boldsymbol{H}}_h^k, \nabla \times \left(\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}} + \widehat{\boldsymbol{E}}_h^{k-\frac{1}{2}}\right)\right)_D - \tau\left(\nabla \times \widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}, \widehat{\boldsymbol{H}}_h^{k+1} + \widehat{\boldsymbol{H}}_h^{k}\right)_D \\
&= \tau\left[\left(\widehat{\boldsymbol{H}}_h^k, \nabla \times \widehat{\boldsymbol{E}}_h^{k-\frac{1}{2}}\right)_D - \left(\widehat{\boldsymbol{H}}_h^{k+1}, \nabla \times \widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}\right)_D\right].
\end{aligned} \tag{67}$$

Summing up (67) from $k = 0$ to any $k \leq N_t$, we obtain

$$\begin{aligned}
&\left(\left\|A^{\epsilon/2}\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}\right\|_{L^2(D)}^2 + \left\|A^{\mu/2}\widehat{\boldsymbol{H}}_h^{k+1}\right\|_{L^2(D)}^2\right) - \left(\left\|A^{\epsilon/2}\widehat{\boldsymbol{E}}_h^{-\frac{1}{2}}\right\|_{L^2(D)}^2 + \left\|A^{\mu/2}\widehat{\boldsymbol{H}}_h^{0}\right\|_{L^2(D)}^2\right) \\
&= \tau\left[\left(\widehat{\boldsymbol{H}}_h^0, \nabla \times \widehat{\boldsymbol{E}}_h^{-\frac{1}{2}}\right)_D - \left(\widehat{\boldsymbol{H}}_h^{k+1}, \nabla \times \widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}\right)_D\right].
\end{aligned} \tag{68}$$

By using the Cauchy–Schwarz inequality, the inverse estimate (64), and Theorem 2, we have

$$\begin{aligned}
\tau\left(\widehat{\boldsymbol{H}}_h^{k+1}, \nabla \times \widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}\right)_D &\leq \tau\left\|\widehat{\boldsymbol{H}}_h^{k+1}\right\|_{L^2(D)} C_{inv} h^{-1}\left\|\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}\right\|_{L^2(D)} \\
&\leq \tau\mu_{min}^{-1/2}\left\|A^{\mu/2}\widehat{\boldsymbol{H}}_h^{k+1}\right\|_{L^2(D)} C_{inv} h^{-1}\epsilon_{min}^{-1/2}\left\|A^{\epsilon/2}\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}\right\|_{L^2(D)} \\
&\leq \frac{\tau h^{-1} C_{inv}\epsilon_{min}^{-1/2}\mu_{min}^{-1/2}}{2}\left(\left\|A^{\mu/2}\widehat{\boldsymbol{H}}_h^{k+1}\right\|_{L^2(D)}^2 + \left\|A^{\epsilon/2}\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}\right\|_{L^2(D)}^2\right).
\end{aligned} \tag{69}$$

Similarly, we have

$$\tau\left(\widehat{\boldsymbol{H}}_h^0, \nabla\times\widehat{\boldsymbol{E}}_h^{-\frac{1}{2}}\right)_D \le \frac{\tau h^{-1}C_{inv}\epsilon_{min}^{-1/2}\mu_{min}^{-1/2}}{2}\left(\left\|A^{\mu/2}\widehat{\boldsymbol{H}}_h^0\right\|_{L^2(D)}^2 + \left\|A^{\epsilon/2}\widehat{\boldsymbol{E}}_h^{-\frac{1}{2}}\right\|_{L^2(D)}^2\right). \tag{70}$$

The proof is completed by substituting the estimates (69) and (70) into (68), and using the time step constraint (65). □

We like to remark that $C_v := \frac{1}{\sqrt{\epsilon_{min}\mu_{min}}}$ denotes the wave propagation speed in a medium with permittivity ϵ_{min} and permeability μ_{min}. Hence the time constraint (65) actually becomes $\tau \le \frac{h}{C_{inv}C_v}$, which is similar to the stability constraint obtained for the leap-frog scheme developed for solving the metamaterial Drude model (cf. [20, Theorem 3.11]).

Modified Leap-Frog Scheme To further reduce the computational cost, wee consider a more efficient scheme than both the Crank–Nicolson and leap-frog schemes. Following the same idea as [37, Lemma 3.3], it can be proved that the matrices $A^\epsilon(\boldsymbol{x})$ and $A^\mu(\boldsymbol{x})$ are strictly diagonal dominant, and we can rewrite them as

$$A^\epsilon(\boldsymbol{x}) = D^\epsilon(\boldsymbol{x}) + S^\epsilon(\boldsymbol{x}), \quad A^\mu(\boldsymbol{x}) = D^\mu(\boldsymbol{x}) + S^\mu(\boldsymbol{x}), \tag{71}$$

where $D^\epsilon(\boldsymbol{x})$, $D^\mu(\boldsymbol{x})$ and $S^\epsilon(\boldsymbol{x})$, $S^\mu(\boldsymbol{x})$ are the diagonal and off-diagonal parts.

By using the Taylor expansion, we can see that

$$\frac{2u^{k-\frac{1}{2}} - 3u^{k-\frac{3}{2}} + u^{k-\frac{5}{2}}}{\tau} = \partial_t u^k + O(\tau^2). \tag{72}$$

Using (71) and (74), we propose the following modified leap-frog type scheme (Scheme 3) for solving (16)–(17): given proper initial approximations $\widehat{\boldsymbol{H}}_h^0 \in \boldsymbol{U}_h$ and $\widehat{\boldsymbol{E}}_h^{-\frac{1}{2}} \in \boldsymbol{V}_h^0$, for $k \ge 0$ find $\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}} \in \boldsymbol{V}_h^0$ and $\widehat{\boldsymbol{H}}_h^{k+1} \in \boldsymbol{U}_h$ such that

$$\begin{aligned}
&\left(D^\epsilon\frac{\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}} - \widehat{\boldsymbol{E}}_h^{k-\frac{1}{2}}}{\tau}, \phi_{E,h}\right)_D - \left(\widehat{\boldsymbol{H}}_h^k, \nabla\times\phi_{E,h}\right)_D \\
&\quad + \left(S^\epsilon\frac{2\widehat{\boldsymbol{E}}_h^{k-\frac{1}{2}} - 3\widehat{\boldsymbol{E}}_h^{k-\frac{3}{2}} + \widehat{\boldsymbol{E}}_h^{k-\frac{5}{2}}}{\tau}, \phi_{E,h}\right)_D = 0, \\
&\left(D^\mu\frac{\widehat{\boldsymbol{H}}_h^{k+1} - \widehat{\boldsymbol{H}}_h^{k}}{\tau}, \phi_{H,h}\right)_D + \left(\nabla\times\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}, \phi_{H,h}\right)_D \\
&\quad + \left(S^\mu\frac{2\widehat{\boldsymbol{H}}_h^{k} - 3\widehat{\boldsymbol{H}}_h^{k-1} + \widehat{\boldsymbol{H}}_h^{k-2}}{\tau}, \phi_{H,h}\right)_D = 0
\end{aligned}$$

hold true for any $\phi_{E,h} \in \boldsymbol{V}_h^0$ and $\phi_{H,h} \in \boldsymbol{U}_h$.

Though we could not rigorously prove the stability of this revised leap-frog scheme at this moment, it is a second order in time scheme and much more efficient than the leap-frog scheme (62)–(63) as shown by our numerical results given below.

4 Numerical Results

In this section, we will perform some numerical tests to verify our theoretical analysis and present some applications of random Maxwell's equations. All our numerical experiments are carried out on a 2017 MacBook Pro laptop with processor of 2.8 GHz Intel Core i7, and memory of 16 GB 2133 MHz LPDDR3.

To implement the leap-frog scheme (62)–(63) for solving (16)–(17), we assume some given proper initial approximations $\widehat{\boldsymbol{H}}_h^0 \in \boldsymbol{U}_h$ and $\widehat{\boldsymbol{E}}_h^{-\frac{1}{2}} \in \boldsymbol{V}_h^0$, then for $k \geq 0$ find $\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}} \in \boldsymbol{V}_h^0$ and $\widehat{\boldsymbol{H}}_h^{k+1} \in \boldsymbol{U}_h$ such that

$$\left(A^{\epsilon} \frac{\widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}} - \widehat{\boldsymbol{E}}_h^{k-\frac{1}{2}}}{\tau}, \phi_{E,h}\right)_D - \left(\widehat{\boldsymbol{H}}_h^k, \nabla \times \phi_{E,h}\right)_D = \left(\mathbf{f}, \phi_{E,h}\right), \quad \forall\, \phi_{E,h} \in \boldsymbol{V}_h^0, \tag{73}$$

$$\left(A^{\mu} \frac{\widehat{\boldsymbol{H}}_h^{k+1} - \widehat{\boldsymbol{H}}_h^{k}}{\tau}, \phi_{H,h}\right)_D + \left(\nabla \times \widehat{\boldsymbol{E}}_h^{k+\frac{1}{2}}, \phi_{H,h}\right)_D = \left(\mathbf{g}, \phi_{H,h}\right), \quad \forall\, \phi_{H,h} \in \boldsymbol{U}_h. \tag{74}$$

where $\mathbf{f}$ and $\mathbf{g}$ are added source terms to allow us to give an exact solution for testing the convergence rate of the numerical scheme.

We partition the physical domain D into $N_e = N_t^2$ rectangular elements with N_d edges. By applying the finite element discretization, we assume that the coefficients in (30) have the following form:

$$\widehat{\boldsymbol{E}}_{m,h}^t = \sum_{j=1}^{N_E} \widehat{\boldsymbol{E}}_{m,j,h}^t \phi_{E,j,h}(\boldsymbol{x}), \quad \widehat{\boldsymbol{H}}_{m,h}^t = \sum_{j=1}^{N_H} \widehat{\boldsymbol{H}}_{m,j,h}^t \phi_{H,j,h}(\boldsymbol{x}),$$

where N_E and N_H are the number of basis functions of $\boldsymbol{E}$ and $\boldsymbol{H}$, respectively. Hence we obtain the algebraic equations for $\boldsymbol{E}_h$ and $\boldsymbol{H}_h$ as following:

$$\mathcal{A}_{m,n,i,j}^{\epsilon} \otimes E_{n,j,h}^{k+1/2} = \mathcal{A}_{m,n,i,j}^{\epsilon} \otimes E_{n,j,h}^{k-1/2} + \tau H_{n,j,h}^k M_{i,j} + \tau F_{n,j}^k, \tag{75}$$

$$\mathcal{A}_{m,n,i,j}^{\mu} \otimes H_{n,j,h}^{k+1} = \mathcal{A}_{m,n,i,j}^{\mu} \otimes H_{n,j,h}^{k} + \tau E_{n,j,h}^{k+1/2} M_{i,j}^{\top} + \tau G_{n,j}^k, \tag{76}$$

where $\mathcal{A}_{m,n,i,j}^{\epsilon}$ and $\mathcal{A}_{m,n,i,j}^{\mu}$ are fourth order tensors whose components are given by

$$\mathcal{A}_{m,n,i,j}^{\epsilon} = \int_{D\times\Omega} \epsilon(\boldsymbol{x}, \boldsymbol{y}) \Phi_m(\boldsymbol{y}) \Phi_n(\boldsymbol{y}) \phi_{E,i,h}(\boldsymbol{x}) \cdot \phi_{E,j,h}(\boldsymbol{x}) \rho(\boldsymbol{y}) d\boldsymbol{x} d\boldsymbol{y},$$

$$\mathcal{A}_{m,n,i,j}^{\mu} = \int_{D\times\Omega} \mu(\boldsymbol{x}, \boldsymbol{y}) \Phi_m(\boldsymbol{y}) \Phi_n(\boldsymbol{y}) \phi_{H,i,h}(\boldsymbol{x}) \cdot \phi_{H,j,h}(\boldsymbol{x}) \rho(\boldsymbol{y}) d\boldsymbol{x} d\boldsymbol{y},$$

and

$$F_{n,j}^k = \left[\int_{D\times\Omega} \rho(\boldsymbol{y}) \Phi_n(\boldsymbol{y}) \mathbf{f}(t_k, \boldsymbol{x}, \boldsymbol{y}) \cdot \phi_{E,j,h}(\boldsymbol{x}) d\boldsymbol{x} d\boldsymbol{y}\right]_{M\times N_E}, \quad 1 \leq n \leq M, \;\; 1 \leq j \leq N_E,$$

$$G_{n,j}^k = \left[\int_{D\times\Omega} \rho(\boldsymbol{y}) \Phi_n(\boldsymbol{y}) \mathbf{g}(t_k, \boldsymbol{x}, \boldsymbol{y}) \cdot \phi_{H,j,h}(\boldsymbol{x}) d\boldsymbol{x} d\boldsymbol{y}\right]_{M\times N_H}, \quad 1 \leq n \leq M, \;\; 1 \leq j \leq N_H.$$

Denote $E^t_{n,j,h}$ and $H^t_{n,j,h}$ for the two second order tensors:

$$E^t_{n,j,h} = \left[\widehat{\boldsymbol{E}}^t_{n,j,h}\right]_{M\times N_E}, \quad 1 \le n \le M, \ \ 1 \le j \le N_E,$$
$$H^t_{n,j,h} = \left[\widehat{\boldsymbol{H}}^t_{n,j,h}\right]_{M\times N_H}, \quad 1 \le n \le M, \ \ 1 \le j \le N_H,$$

and $\otimes$ for a tensor product like operator:

$$\begin{aligned}\mathcal{A}^{\epsilon}_{m,n,i,j} \otimes E^t_{n,j,h} &= \sum_{m=1}^{M}\sum_{i=1}^{N_E} \widehat{\boldsymbol{E}}^t_{n,j,h} \int_{D\times\Omega} \epsilon(\boldsymbol{x},\boldsymbol{y})\Phi_m(\boldsymbol{y})\Phi_n(\boldsymbol{y})\phi_{E,i,h}(\boldsymbol{x})\\ &\quad\cdot \phi_{E,j,h}(\boldsymbol{x})\rho(\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y},\\ \mathcal{A}^{\mu}_{m,n,i,j} \otimes H^t_{n,j,h} &= \sum_{m=1}^{M}\sum_{i=1}^{N_H} \widehat{\boldsymbol{H}}^t_{n,j,h} \int_{D\times\Omega} \mu(\boldsymbol{x},\boldsymbol{y})\Phi_m(\boldsymbol{y})\Phi_n(\boldsymbol{y})\phi_{H,i,h}(\boldsymbol{x})\\ &\quad\cdot \phi_{H,j,h}(\boldsymbol{x})\rho(\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y},\end{aligned}$$

and $M_{i,j}$ for the stiffness matrix:

$$M_{i,j} = \left[\int_{\Omega} \phi_{H,i,h} \cdot \nabla\times\phi_{E,j,h}(\boldsymbol{x})d\boldsymbol{x}\right]_{N_H\times N_E}, \quad 1 \le i \le N_H, \ \ 1 \le j \le N_E.$$

Therefore, once we have $\boldsymbol{E}^t_{n,j,h}$ and $\boldsymbol{H}^t_{n,j,h}$, we can compute the numerical solutions $\boldsymbol{E}_h$ and $\boldsymbol{H}_h$ by the following quadratic forms:

$$\boldsymbol{E}_h(t,\boldsymbol{x},\boldsymbol{y}) = (\Phi_1(\boldsymbol{y}),\ldots,\Phi_M(\boldsymbol{y}))\cdot E^t_{n,j,h}\cdot(\phi_{E,1,h}(\boldsymbol{x}),\ldots,\phi_{E,N_E,h}(\boldsymbol{x}))^{\top}, \ \text{at } t = t_{k+1/2},$$
$$\boldsymbol{H}_h(t,\boldsymbol{x},\boldsymbol{y}) = (\Phi_1(\boldsymbol{y}),\ldots,\Phi_M(\boldsymbol{y}))\cdot H^t_{n,j,h}\cdot(\phi_{H,1,h}(\boldsymbol{x}),\ldots,\phi_{H,N_H,h}(\boldsymbol{x}))^{\top}, \ \text{at } t = t_k.$$

4.1 Example 1: Test of Convergence and CPU Time

For simplicity, we solve the two-dimensional TE_z mode equation:

$$\begin{aligned}\epsilon(\boldsymbol{x},\boldsymbol{y})\partial_t E_x(t,\boldsymbol{x},\boldsymbol{y}) &= \partial_y H(t,\boldsymbol{x},\boldsymbol{y}) + f_1(t,\boldsymbol{x},\boldsymbol{y}),\\ \epsilon(\boldsymbol{x},\boldsymbol{y})\partial_t E_y(t,\boldsymbol{x},\boldsymbol{y}) &= -\partial_x H(t,\boldsymbol{x},\boldsymbol{y}) + f_2(t,\boldsymbol{x},\boldsymbol{y}),\\ \mu(\boldsymbol{x},\boldsymbol{y})\partial_t H(t,\boldsymbol{x},\boldsymbol{y}) &= -(\partial_x E_y(t,\boldsymbol{x},\boldsymbol{y}) - \partial_y E_x(t,\boldsymbol{x},\boldsymbol{y})) + g(t,\boldsymbol{x},\boldsymbol{y}),\end{aligned}$$

subject to the PEC boundary condition (42). We solve this system on $D\times\Omega\times[0,T]$, where $\Omega = D = [0,1]^2$ and $T = 10^{-5}$. The domain D is partitioned uniformly into $N_e = N_t^2$ rectangular elements with a total edge number N_d, where N_t is the total time steps. We solve the problem by using the lowest order edge element on D, hence $N_E = N_d$ and $N_H = N_e$. We choose the permittivity and permeability as follows:

$$\begin{aligned}\epsilon(\boldsymbol{x},\boldsymbol{y}) &= 1 + 0.1(\sin(x_1)\cos(y_1) + \cos(x_2)\sin(y_2)),\\ \mu(\boldsymbol{x},\boldsymbol{y}) &= 1 + 0.1(\cos(x_1)\sin(y_1) + \sin(x_2)\cos(y_2)),\end{aligned}$$

for any $\boldsymbol{x} = (x_1, x_2) \in D$ and $\boldsymbol{y} = (y_1, y_2) \in \Omega$. The following exact solution is used to test the accuracy of our numerical scheme:

$$E_x(t, \boldsymbol{x}, \boldsymbol{y}) = \sin(\pi x_2)e^{-t}(1 + 0.1(\sin(\pi y_1))\cos(\pi y_2)),$$
$$E_y(t, \boldsymbol{x}, \boldsymbol{y}) = \sin(\pi x_1)e^{-t}(1 + 0.1(\cos(\pi y_1))\sin(\pi y_2)),$$
$$H(t, \boldsymbol{x}, \boldsymbol{y}) = \pi(\cos(\pi x_1) - \cos(\pi x_2))e^{-t}(1 + 0.1(\sin(\pi y_1))\sin(\pi y_2))$$

with appropriate source terms $\mathbf{f} = (f_1, f_2)$ and $\boldsymbol{g} = g$. We use orthogonal polynomials of degrees up to 5 and assume that $\boldsymbol{y}$ has a 2-dimensional uniform distribution on $[0, 1]^2$. Hence, $\rho(\boldsymbol{y}) = 1$ for $\boldsymbol{y} \in \Omega$ and our $M = 21$. We calculate the errors of $\boldsymbol{E} := (E_x, E_y)$ and H at the final time T by the following norm:

$$\|u - u_h\|^2_{l^2(L^2)} := \int_\Omega \rho(\boldsymbol{y}) \sum_{i=1}^{N_e} |u(\boldsymbol{x}_i, \boldsymbol{y}) - u_h(\boldsymbol{x}_i, \boldsymbol{y})|^2 |K_i| d\boldsymbol{y},$$

where $\boldsymbol{x}_i$ is the middle point of element K_i, $|K_i|$ is the area of element K_i, and u represents $\boldsymbol{E}$ or H.

In our test, we let $N_e = N_t^2$ with varying N_t. The solution errors are presented in Table 1, which clearly shows a second order convergence for both $\boldsymbol{E}$ and H. This is consistent with the theoretical result of leap-frog scheme [20].

To test the convergence of the gPC expansion on Ω, we solved the problem by using different orders of orthogonal polynomials for $\boldsymbol{E}$ and H with a fixed $N_t = 20$. Observing the error of H in Fig. 1, we can find that the error is decreasing when the degree p of the gPC orthonormal polynomials is increasing. Note that the error stops going down further when $p \geq 4$. This is because the gPC error is so small for $p \geq 4$ that the spatial and temporal error of the scheme will dominate the total error. Therefore, in the above numerical example, we choose $p = 5$ so that the gPC error will not affect the total error.

Considering that the standard leap-frog scheme (62)–(63) involves the full matrices $A^{\epsilon(\boldsymbol{x})}$ and $A^{\mu(\boldsymbol{x})}$, we expect that the revised leap-frog type scheme (Scheme 3) would be more

Table 1 Errors of $\boldsymbol{E}$ and H obtained by Scheme 2

Mesh	$\|\boldsymbol{E} - \boldsymbol{E}_h\|_{l^2(L^2)}$	Rate	$\|H - H_h\|_{l^2(L^2)}$	Rate	CPU time (s)
2×2	8.517593e−02	–	3.259975e−01	–	13.243185
4×4	1.760587e−02	2.2744	8.343959e−02	1.9925	104.981369
8×8	4.620225e−03	2.1022	2.110498e−02	2.0072	727.997221
16×16	1.323839e−03	1.9953	5.162830e−03	2.0313	5630.592848

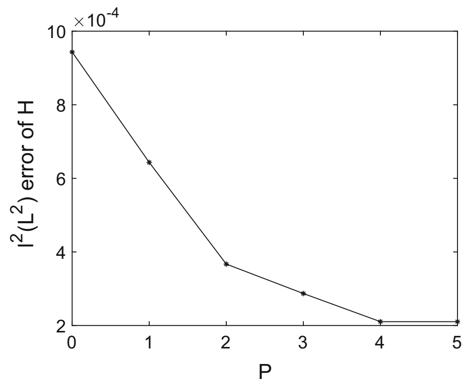

Fig. 1 The error of H vs the degree p of gPC orthonormal polynomials

Table 2 Errors of $\boldsymbol{E}$ and H obtained by Scheme 3

Mesh	$\|\boldsymbol{E}-\boldsymbol{E}_h\|_{l^2(L^2)}$	Rate	$\|\boldsymbol{H}-\boldsymbol{H}_h\|_{l^2(L^2)}$	Rate	CPU time (s)
2×2	8.517593e−02	–	3.259975e−01	–	0.893071
4×4	1.760587e−02	2.2744	8.343959e−02	1.9847	6.487596
8×8	4.520225e−03	2.1180	2.110498e−02	1.9942	44.077161
16×16	1.284596e−03	2.0115	5.257084e−03	2.0052	358.162832

efficient. By using this scheme, we just need to handle the diagonal matrices $D^{\epsilon(\boldsymbol{x})}$ and $D^{\mu(\boldsymbol{x})}$. The CPU time and errors calculated by this revised scheme are shown in Table 2, which shows that the new scheme produces almost the same accuracy with a significant saving of CPU time.

4.2 Example 2: Application with Random permittivity

Here we will display one numerical experiment for wave scattering problem solved by using our method. The numerical test is done by using 1600 rectangular edge elements in the physical domain $D=[-0.5,0.5]^2$. The time domain for the test is $[0,0.9]$ with $N_t=100$. We still choose $p=5$ for the orthonormal polynomial in Ω since it is accurate enough as discussed above. A similar setup as the experiment 5.2.2 of [6] is used. Namely, we solve the scattered fields $\boldsymbol{E}^s$ and $\boldsymbol{H}^s$ governed by the following equations:

$$\epsilon\frac{\partial \boldsymbol{E}^s}{\partial t}=\nabla\times\boldsymbol{H}^s+\sigma\boldsymbol{E}^s+\mathbf{S}^E,$$
$$\mu\frac{\partial \boldsymbol{H}^s}{\partial t}=-\nabla\times\boldsymbol{E}^s+\mathbf{S}^H,$$

and the source terms $\mathbf{S}^E$ and $\mathbf{S}^H$ are given by

$$\mathbf{S}^E=-(\epsilon-\epsilon^i)\frac{\partial \boldsymbol{E}^i}{\partial t}+(\sigma-\sigma^i)\boldsymbol{E}^i,$$
$$\mathbf{S}^H=-(\mu-\mu^i)\frac{\partial \boldsymbol{H}^i}{\partial t}.$$

Here the incident field $(\boldsymbol{E}^i,\boldsymbol{H}^i)$ is a solution of Maxwell's equation with permittivity ϵ^i, permeability μ^i, and conductivity σ^i. More specifically, we choose

$$E_x^i=\sin(\pi y)\sin(\pi t),$$
$$E_y^i=\sin(\pi x)\sin(\pi t),$$
$$H^i=(\cos(\pi x)-\cos(\pi y))\cos(\pi t),$$

$\sigma=\sigma^i=0$, and

$$\epsilon^i(\boldsymbol{x},\boldsymbol{y})=\begin{cases}2.25e^{0.1y} & \text{if } \boldsymbol{x}\in B(0.1),\\ 1 & \text{otherwise,}\end{cases}$$

where $B(r)$ denotes a disc located at the center of the physical domain with radius r. In other words, $\epsilon^i(\boldsymbol{x},\boldsymbol{y})$ is a univariate function on $\Omega=[0,1]$ and $\rho(y)=1$.

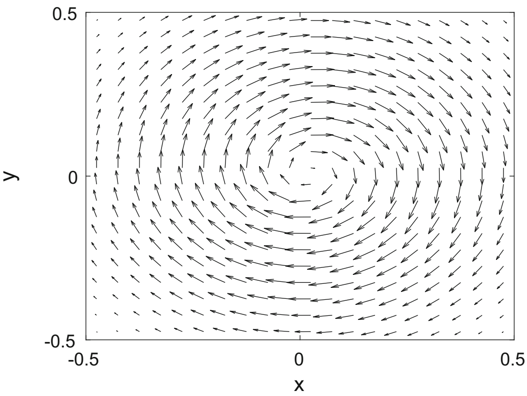

Fig. 2 The plot of vector field $\boldsymbol{E}$

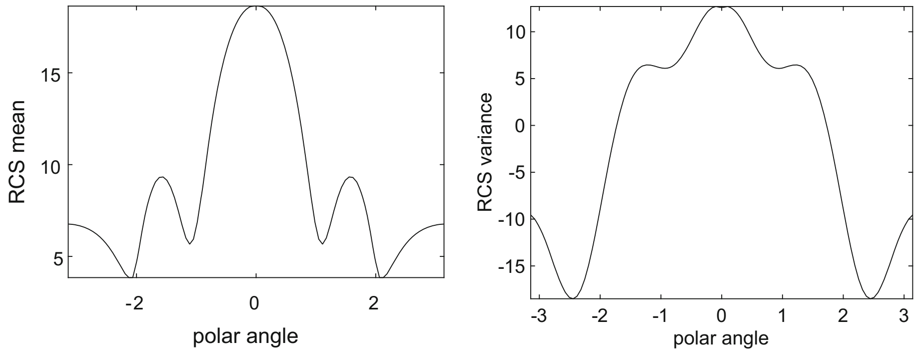

Fig. 3 The mean and variance of the RCS

A good measurement of scattering problem is the so-called radar cross section (RCS) [3, Sec.11.3]:

$$RCS(\phi) = \lim_{r\to\infty} 10 \ln \left(2\pi\rho \frac{|\boldsymbol{E}^s(\phi)|^2}{|\boldsymbol{E}^i|^2} \right) \tag{77}$$

where $\phi \in [-\pi, \pi]$ is the polar angle. In Fig. 2 we plot the electronic field $\boldsymbol{E} = (E_x, E_y)$ on D, and in Fig. 3 we present the mean and variance of the RCS given by (77).

5 Conclusion

In this work, we proposed the stochastic Galerkin methods for time-dependent Maxwell's equations with random input. It is shown that the stochastic Galerkin approach preserves the energy conservation law. Moreover, we propose a finite element approach in the physical space and propose three schemes to deal with the time discretizition. It is shown that the fully discrete Crank–Nicolson scheme preserves the energy. While the classic leap-frog scheme admits a conditional energy stability property. A modified leap-frog type scheme is designed to further reduce the computational cost. Numerical examples are presented to support our theoretical analysis and demonstrate the effectiveness of our algorithm in solving the wave scattering problem.

References

1. Babuska, I.M., Nobile, F., Tempone, R.: A stochastic collocation method for elliptic partial differential equations with random input data. SIAM J. Numer. Anal. **435**, 1005–1034 (2007)
2. Babuska, I., Tempone, R., Zouraris, G.: Galerkin finite element approximations of stochastic elliptic differential equations. SIAM J. Numer. Anal. **42**, 800–825 (2004)
3. Balanis, C.A.: Advanced Engineering Electromagnetics, 2nd edn. Wiley, Hoboken, NJ (2012)
4. Benner, P., Schneider, J.: Uncertainty quantification for Maxwell's equations using stochastic collocation and model order reduction. Int. J. Uncertain. Quantif. **5**(3), 195–208 (2015)
5. Cao, Y.: On convergence rate of Wiener–Ito expansion for generalized random variables. Stoch. Int. J. Probab. Stoch. Process. **78**(3), 179–187 (2006)
6. Chauviére, C., Hesthaven, J.S., Lurati, L.: Computational modeling of uncertainty in time-domain electromagnetics. SIAM J. Sci. Comput. **28**(2), 751–775 (2006)
7. Deang, J., Du, Q., Gunzburger, M.D.: Modeling and computation of random thermal fluctuations and material defects in the Ginzburg–Landau model for superconductivity. J. Comput. Phys. **181**, 45–67 (2002)
8. Deb, M.K., Babuska, I.M., Oden, J.T.: Solution of stochastic partial differential equations using Galerkin finite element techniques. Comput. Methods Appl. Mech. Eng. **190**, 6359–6372 (2001)
9. Dostert, P., Efendiev, Y., Hou, T.Y.: Multiscale finite element methods for stochastic porous media flow equations and application to uncertainty quantification. Comput. Methods Appl. Mech. Eng. **197**, 3445–3455 (2008)
10. Elman, H.C., Furnival, D.G., Powell, C.E.: $H(div)$ preconditioning for a mixed finite element formulation of the diffusion problem with random data. Math. Comput. **79**, 733–760 (2010)
11. Fouque, J., Garnier, J., Papanicolaou, G., Solna, K.: Wave Propogation and Time Reversal in Randomly Layered Media. Springer, Berlin (2007)
12. Galvis, J., Sarkis, M.: Approximating infinity-dimensional stochastic Darcy's equations without uniform ellipticity. SIAM J. Numer. Anal. **47**(5), 3624–3651 (2009)
13. Ghanem, R., Spanos, P.: Stochastic Finite Elements: A Spectral Approach. Springer, New York (1991)
14. Graham, I.G., Kuo, F.Y., Nichols, J.A., Scheichl, R., Schwab, Ch., Sloan, I.H.: Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients. Numer. Math. **131**, 329–368 (2015)
15. Gunzburger, M.D., Webster, C.G., Zhang, G.: Stochastic finite element methods for partial differential equations with random input data. Acta Numer. **23**, 521–650 (2014)
16. Jin, S., Xiu, D., Zhu, X.: A well-balanced stochastic Galerkin method for scalar hyperbolic balance laws with random inputs. J. Sci. Comput. **67**(3), 1198–1218 (2016)
17. Kovacs, M., Larsson, S., Saedpanah, F.: Finite element approximation of the linear stochastic wave equation with additive noise. SIAM J. Numer. Anal. **48**, 408–427 (2010)
18. Li, J., Fang, Z.: Analysis and application of stochastic collocation methods for Maxwell's equations with random inputs. Adv. Appl. Math. Mech. **10**, 1305–1326 (2018)
19. Li, J., Fang, Z., Lin, G.: Regularity analysis of metamaterial Maxwells equations with random coefficients and initial conditions. Comput. Methods Appl. Mech. Eng. **335**, 24–51 (2018)
20. Li, J., Huang, Y.: Time-Domain Finite Element Methods for Maxwell's Equations in Metamaterials. Springer Series in Computational Mathematics, vol. 43. Springer, Berlin (2013)
21. Li, J., Machorro, E.A., Shields, S.: Numerical study of signal propagation in corrugated coaxial cables. J. Comput. Appl. Math. **309**, 230–243 (2017)
22. Lord, G., Powell, C.E., Shardlow, T.: An Introduction to Computational Stochastic PDEs. Cambridge University Press, Cambridge (2014)
23. Monk, P.: Finite Element Methods for Maxwell's Equations. Oxford University Press, Oxford (2003)
24. Motamed, M., Nobile, F., Tempone, R.: A stochastic collocation method for the second order wave equation with a discontinuous random speed. Numer. Math. **123**, 493–536 (2013)
25. Musharbash, E., Nobile, F., Zhou, T.: Error analysis of the dynamically orthogonal approximation of time dependent random PDEs. SIAM J. Sci. Comput. **37**(2), A776–A810 (2015)
26. Narayan, A., Zhou, T.: Stochastic collocation methods on unstructured meshes. Commun. Comput. Phys. **18**, 1–36 (2015)
27. Nobile, F., Tempone, R., Webster, C.G.: A sparse grid stochastic collocation method for partial differential equations with random input data. SIAM J. Numer. Anal. **46**(5), 2309–2345 (2008)
28. Oden, J.T., Belytschko, T., Babuska, I., Hughes, T.J.R.: Research directions in computational mechanics. Comput. Methods Appl. Mech. Eng. **192**, 913–922 (2003)
29. Schwab, C., Gittelson, C.J.: Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs. Acta Numer. **20**, 291–467 (2011)

30. Tang, T., Zhou, T.: Convergence analysis for stochastic collocation methods to scalar hyperbolic equations with random wave speed. Commun. Comput. Phys. **8**(1), 226–248 (2010)
31. Tryoen, J., LeMaitre, O., Ndjinga, M., Ern, A.: Intrusive Galerkin methods with upwinding for uncertain nonlinear hyperbolic systems. J. Comput. Phys. **229**, 6485–6511 (2010)
32. Wan, X., Karniadakis, G.E.: Long-term behavior of polynomial chaos in stochastic flow simulations. Comput. Methods Appl. Mech. Eng. **195**, 5582–5596 (2006)
33. Wu, K., Tang, H., Xiu, D.: A stochastic Galerkin method for first-order quasilinear hyperbolic systems with uncertainty. J. Comput. Phys. **345**, 224–244 (2017)
34. Xiu, D.: Numerical Methods for Stochastic Computations: A Spectral Method Approach. Princeton University Press, Princeton (2010)
35. Xiu, D., Karniadakis, G.E.: The Wiener–Askey polynomial chaos for stochastic differential equations. SIAM J. Sci. Comput. **24**(2), 619–644 (2002)
36. Xiu, D., Hesthaven, J.S.: High-order collocation methods for differential equations with random inputs. SIAM J. Sci. Comput. **27**(3), 1118–1139 (2005)
37. Xiu, D., Shen, J.: Efficient stochastic Galerkin methods for random diffusion equations. J. Comput. Phys. **228**(2), 266–281 (2009)
38. Zhou, T.: Stochastic Galerkin methods for elliptic interface problems with random input. J. Comput. Appl. Math. **236**, 782–792 (2011)
39. Zhou, T., Tang, T.: Galerkin methods for stochastic hyperbolic problems using bi-orthogonal polynomials. J. Sci. Comput. **51**, 274–292 (2012)

Affiliations

Zhiwei Fang[1] · Jichun Li[1] · Tao Tang[2] · Tao Zhou[3]

✉ Jichun Li
jichun.li@unlv.edu

Zhiwei Fang
fangz1@unlv.nevada.edu

Tao Tang
tangt@sustc.edu.cn

Tao Zhou
tzhou@lsec.cc.ac.cn

1 Department of Mathematical Sciences, University of Nevada Las Vegas, Las Vegas, NV 89154-4020, USA

2 Department of Mathematics, Southern University of Science and Technology, Shenzhen 518055, Guangdong, China

3 LSEC, Institute of Computational Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China