Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 10(8): 19816–19844.
DOI: 10.3934/math.2025884
Received: 02 June 2025
Revised: 18 August 2025
Accepted: 22 August 2025
Published: 28 August 2025

Research article

Numerical analysis of the MLMC ensemble scheme for transient heat equations with uncertain inputs

Tingfu Yao[1,*], Changlun Ye[2] and Xianbing Luo[3,*]

[1] College of Science, Guiyang University, Guiyang 550005, China

[2] School of Mathematical Sciences, Guizhou Normal University, Guiyang 550001, China

[3] School of Mathematics and Statistics, Guizhou University, Guiyang 550025, China

* **Correspondence:** Email: ytfwdm520@163.com, xbluo1@gzu.edu.cn.

Abstract: A multilevel Monte Carlo ensemble (MLMCE) coupled with the finite element method is applied to address numerically transient heat equations characterized by random diffusion and Robin coefficients. By incorporating two ensemble averages for the Robin boundary and diffusion coefficients, we present an extended Monte Carlo ensemble scheme tailored for the uncertain transient heat equation. The suggested MLMCE approach resolves a single linear system that entails multiple right-hand-side vectors for a group during each time step, thereby decreasing both the storage requirements and the computational expenses associated with the solution process. Stability analysis and error estimates of the method are derived under some conditions involving two ratios between fluctuations in the thermal conductivity and a random Robin coefficient corresponding to their mean. Numerical experiments are presented to confirm the theoretical results and verify the feasibility and effectiveness of the proposed approach.

Keywords: multilevel Monte Carlo method; ensemble; finite element; transient heat equation; Robin coefficient; diffusion coefficient; uncertain inputs
Mathematics Subject Classification: 65C05, 65M60, 65N12

1. Introduction

We consider the numerical solutions of a transient heat model characterized by random diffusion coefficients and Robin coefficients, aiming to find a random function y such that it holds almost surely

(a.s.)

$$\begin{cases} y_t - \nabla \cdot [a(\mathbf{x},\omega)\nabla y] = f(t,\mathbf{x},\omega), & \text{in } [0,T]\times D\times\Omega, \\ a\nabla y(t,\mathbf{x},\omega)\cdot\mathbf{n} = 0, & \text{on } [0,T]\times\partial D_0\times\Omega, \\ a\nabla y(t,\mathbf{x},\omega)\cdot\mathbf{n} = \alpha(\mathbf{x},\omega)(u(t,\mathbf{x},\omega)-y(t,\mathbf{x},\omega)), & \text{on } [0,T]\times\partial D_1\times\Omega, \\ y(0,\mathbf{x}) = y^0(\mathbf{x},\omega), & \text{in } D\times\Omega, \end{cases} \tag{1.1}$$

where $D \subseteq \mathbb{R}^d$ ($d = 2,3$) denotes a Lipschitz domain. The boundary ∂D is split into two distinct sections ∂D_0 and ∂D_1. The vector $\mathbf{n}$ represents the outward unit normal vector to the boundary ∂D. In this context, $(\Omega, \mathcal{F}, P)$ denotes a complete probability space, with Ω as the sample space, $\mathcal{F} \subset 2^\Omega$ being the σ-algebra of the events, and $P : \mathcal{F} \to [0,1]$ representing a probability measure.

Uncertainty is prevalent in various challenges related to physics and engineering. Numerous uncertain factors influence the processes of heat and mass transfer, including random initial temperatures, ambient temperatures, material properties, thermal conductivities (diffusion coefficients), convective heat transfer coefficients (also known as Robin coefficients), and geometric variations. The problem stated in Eq (1.1) involves random diffusion coefficients and Robin coefficients that may vary with the time variable $t \in [0,T]$. However, for simplicity, it is assumed to be stationary, as referenced in [3,4]. The model in (1.1) is also relevant to boundary control issues, as discussed in [27,28].

There are numerous numerical algorithms aimed at resolving partial differential equations (PDEs) characterized by random coefficients (see, for example, [1,9,12,24,29,31,38]). In addition to the polynomial chaos method, the stochastic collocation method, and the stochastic finite element method, the Monte Carlo (MC) method stands out as a significant technique (see, for instance, [8,10,13,26]). The MC method is non-intrusive and its convergence is independent of the dimensionality of the random model's parameters. This method is straightforward to implement: It involves initial independent sampling followed by separate numerical simulations for each sample. These simulations account for variations in the initial and boundary conditions, physical influences, diffusion coefficients, and Robin coefficients. Let us assume that the simulations consist of J independent samples, where the j-th sample corresponds to

$$\begin{cases} y_{j,t} - \nabla \cdot \left[a_j(\mathbf{x})\nabla y_j\right] = f_j(t,\mathbf{x}), & \text{in } (0,T)\times D, \\ a_j\nabla y_j(t,\mathbf{x})\cdot\mathbf{n} = 0, & \text{on } (0,T)\times\partial D_0, \\ a_j\nabla y_j(t,\mathbf{x})\cdot\mathbf{n} = \alpha_j(\mathbf{x})(u_j(t,\mathbf{x})-y_j(t,\mathbf{x})), & \text{on } (0,T)\times\partial D_1, \\ y_j(0,\mathbf{x}) = y_j^0(\mathbf{x}), & \text{in } D, \end{cases} \tag{1.2}$$

$j \in \{1,2,\cdots,J\}$. Here, we can assume $a_j(\mathbf{x}) = a(\mathbf{x},\omega_j)$, $\alpha_j(\mathbf{x}) = \alpha(\mathbf{x},\omega_j)$.

In the numerical simulations, many linear equations of the form $\mathbf{A}_j\mathbf{z}_j = \mathbf{b}_j$, $j \in \{1,2,\cdots,J\}$ must be resolved. Given that the number J is relatively substantial, the computational expense becomes significant. Therefore, ensemble methods are frequently employed to enhance computational efficiency (see, for instance, [14,18,25,26,34–37]). The ensemble approach transforms the problem of solving $\mathbf{A}_j\mathbf{z}_j = \mathbf{b}_j$, $j \in \{1,2,\cdots,J\}$ into solving $\tilde{\mathbf{A}}\tilde{\mathbf{z}}_j = \tilde{\mathbf{b}}_j$, $j \in \{1,2,\cdots,J\}$. By [25], the authors examined a parabolic problem characterized by random coefficients utilizing the ensemble method, resulting in an error estimate. However, this error estimate is not optimal concerning the spatial considerations. To address this issue, the researchers in [18–21] merged the ensemble method with the hybridizable discontinuous Galerkin (HDG) approach to derive an optimal error estimate in terms of space. Concerning the heat conduction problem with random Robin coefficients, we find that there are relatively few research articles about the ensemble method, other than our team's recent contributions [34–37]. However, we have not discovered any relevant findings regarding the multilevel Monte Carlo ensemble (MLMCE) method for this model. The main objective of this article is to analyze the MLMCE algorithm numerically for transient heat equations with uncertain inputs. The content primarily focuses on the numerical stability conditions of the MLMCE algorithm, along with stability analysis, error estimates, and computational cost analysis.

The coefficient matrix of the resulting linear system, obtained via spatial discretization, is independent of j. This key feature of the ensemble algorithm should be emphasized. For a given ensemble of simulations, the corresponding discrete systems share the same coefficient matrix, while the right-hand-side (RHS) vectors vary from one ensemble member to another. Therefore, for small-scale problems, the solution for the group can be obtained by performing Lower Upper (LU) decomposition of the coefficient matrix just once (see, for example, [25, 26]); for large-scale problems, an efficient computation will employ the block Krylov subspace iteration method (refer to, for instance, [7, 17, 22, 30]).

The structure of this article is outlined as follows. Section 2 presents some notations and preliminary concepts. The full discretization ensemble framework is detailed in Section 3. The discussion of the random transient heat equation, including its stability and error analysis, is provided in Section 4. Finally, Section 5 showcases several numerical tests, while Section 6 offers some concluding remarks.

2. Basic preliminaries

We introduce several notations in this section, which are primarily sourced from books [6, 28]. For the sake of simplicity, we will omit the terms $d\mathbf{x}, d\mathbf{s}$, and dt in certain expressions where it does not lead to confusion. The boundaries ∂D_0 and ∂D_1 refer to the parts that are experimentally accessible and inaccessible, respectively.

We use $\|\cdot\|$ and $(\cdot,\cdot)$ to represent the $L^2(D)$ norm and inner product, respectively. Additionally, $\|\cdot\|_{\partial D}$ and $(\cdot,\cdot)_{\partial D}$ correspond to the $L^2(\partial D)$ norm and inner product. The Sobolev space $W^{s,q}(D)$ is equipped with the norm $\|v\|_{W^{s,q}(D)}$, where $s \in \mathbb{N}$ (the set of natural numbers) and $q \geq 1$. We denote $H^s(D) = W^{s,2}(D)$. Notably, the norm $H^1(D)$ is defined as $\|\cdot\|_1 = \|\cdot\|_{1,D}$

$$\|y\|_{1,D} = \left(\|y\|^2 + \|\nabla y\|^2\right)^{1/2}.$$

Let $H^{-s}(D)$ be the dual space of bounded linear functions on $H^s(D)$, with the norm

$$\|f\|_{-s} = \sup_{0\neq v\in H^s(D)} (f, v)/\|v\|_s.$$

The norm $\|\cdot\|_{1,\partial D_1}$ defined by

$$\|y\|_{1,\partial D_1} = \left(\int_D |\nabla y|^2 + \int_{\partial D_1} |y|^2\right)^{1/2},$$

is equivalent to the standard norm $\|\cdot\|_1$ (see also [15, 16]) and will be used in this work. In subsequent statements, we will treat the norms $\|\cdot\|_{1,\partial D_1}$ and $\|\cdot\|_1 = \|\cdot\|_{H^1(D)}$ as equivalent. We define

$$L^\infty(D) = \{v : v \text{ is a measurable function and } |v|_\infty < +\infty\},$$

where $|v|_\infty = \text{ess } \sup_{\mathbf{x}\in D}|v|$.

Let $Z \in L^1_P(\Omega)$ be a random variable and denote $\mathbb{E}[Z] = \int_\Omega Z(\omega)dP(\omega)$. Let $\delta = (\delta_1, \cdots, \delta_d)$ be a d-tuple with the length $|\delta| = \sum_{i=1}^{d} \delta_i, \delta_i \in N^+$. The stochastic Sobolev spaces $\widetilde{W}^{s,q}(D) = L^q_P(\Omega, W^{s,q}(D))$ consists of stochastic functions, represented as $v : \Omega \times D \to R$, which are measurable with respect to (w.r.t.) the product σ-algebra $\mathcal{F} \otimes B(D)$, where $B(D)$ is a Borel σ-algebra. The norm of $\widetilde{W}^{s,q}(D)$ is defined by

$$\|v\|_{\widetilde{W}^{s,q}(D)} = \left(\mathbb{E}\left[\|v\|^q_{W^{s,q}(D)}\right]\right)^{1/q} = \left(\mathbb{E}\left[\sum_{|\delta|\le s}\int_D \left|\partial^\delta v\right|^q\right]\right)^{1/q}, 1 \le q < +\infty.$$

Let $\widetilde{H}^s(D) = \widetilde{W}^{s,2}(D) \simeq L^2_P(\Omega) \otimes H^s(D)$. We will use the tensor product Hilbert space

$$X = \widetilde{L}^2\left(0, T; H^1(D)\right) \simeq L^2_P\left(0, T; H^1(D); \Omega\right)$$

with its inner product

$$(v, u)_X \equiv \mathbb{E}\left[\int_0^T \int_D (\nabla v \cdot \nabla u + vu)\right].$$

The induced norm is given by

$$\|v\|_X = \left(\mathbb{E}\left[\int_0^T \int_D \left(|\nabla v|^2 + v^2\right)\right]\right)^{1/2}.$$

A weak solution to the problem (1.1) is defined in the following manner: A function $y \in X$ qualifies as a weak solution of (1.1) if it satisfies the initial condition $y(0, \mathbf{x}, \omega) = y^0(\mathbf{x}, \omega) \in L^2_P\left(H^1(D); \Omega\right)$, and for $T > 0$

$$\begin{aligned} &\int_0^T (y_t, v)\, dt + \int_0^T \int_\Omega \int_D a\nabla y \cdot \nabla v d\mathbf{x} dP(\omega)dt + \int_0^T \int_\Omega \int_{\partial D_1} \alpha y v d\mathbf{s} dP(\omega)dt \\ =& \int_0^T \int_\Omega \int_D f v d\mathbf{x} dP(\omega)dt + \int_0^T \int_\Omega \int_{\partial D_1} \alpha u v d\mathbf{s} dP(\omega)dt \end{aligned} \tag{2.1}$$

for all $v \in X$, where $d\mathbf{s}$ is the Lebesgue measure on ∂D_1.

In this paper, we use C to denote a positive constant that takes various values in different contexts, independent of the time step Δt_l, the sample size J_l, and the mesh size h_l.

3. Multilevel MC ensemble scheme for the transient heat equation

When statistical data regarding the inputs of random transient heat equations are provided, uncertainty quantification serves to ascertain statistical details about the outputs of interest that are influenced by the solutions of these transient heat equations. After employing MC sampling, to tackle the model (1.1), it becomes essential to approximate the solutions associated with a set of independent realizations, which involves solving deterministic PDEs at random sample values. Typically, numerical simulations are conducted in isolation, leading to an incremental increase in the total computational costs as the sampling set enlarges. In order to enhance efficiency, this study introduces a method based on multilevel MC (MLMC) ensembles, which extends the MC ensemble technique previously discussed in [25, 34]. The newly proposed method demonstrates superior performance compared with its predecessor in terms of both accuracy and efficiency. This improvement can be attributed to the combination of a second-order ensemble-based time-stepping strategy with the MLMC approach.

Next, we introduce the algorithm within the framework of numerical approximations for the random PDE (1.1). In terms of spatial discretization, conforming finite elements are utilized. For a polygonal (or polyhedral) region D, to align with the hierarchical structure of MLMC methods, we consider a series of quasi-uniform grids consisting of a collection of shape-regular triangles (or tetrahedra), $\{\mathcal{T}_l\}_{l=0}^{L}$. Denote the mesh size of $\mathcal{T}_l$ as $h_l = \max_{K\in\mathcal{T}_l} \operatorname{diam} K$. Assume that the mesh sequence is generated through uniform refinements, meeting the criteria of h_l satisfying

$$h_l = h_0 2^{-l}. \tag{3.1}$$

For a non-negative integer k, let $\mathbf{P}_k$ denote the set of polynomials of degree k, and define the finite element (FE) space as follows

$$X_l := \{v \in C(\bar{D}) : v|_K \in \mathbf{P}_k, \forall K \in \mathcal{T}_l\},$$

and the series of FE spaces satisfies

$$X_0 \subset X_1 \subset \cdots \subset X_l \subset \cdots \subset X_L.$$

Let $y_l(t_n, \mathbf{x}, \omega)$ denote the FE solution in X_l at $t = t_n$. The MLMC FE solution at the L-th level mesh can be expressed as

$$y_L(t_n, \mathbf{x}, \omega) = \sum_{l=1}^{L} \left(y_l(t_n, \mathbf{x}, \omega) - y_{l-1}(t_n, \mathbf{x}, \omega)\right) + y_0(t_n, \mathbf{x}, \omega).$$

By the linearity of $\mathbb{E}[\cdot]$, we get

$$\begin{aligned}
\mathbb{E}\left[y_L(t_n, \mathbf{x}, \omega)\right] &= \mathbb{E}\left[\sum_{l=1}^{L} \left(y_l(t_n, \mathbf{x}, \omega) - y_{l-1}(t_n, \mathbf{x}, \omega)\right) + y_0(t_n, \mathbf{x}, \omega)\right] \\
&= \sum_{l=1}^{L} \mathbb{E}\left[y_l(t_n, \mathbf{x}, \omega) - y_{l-1}(t_n, \mathbf{x}, \omega)\right] + \mathbb{E}\left[y_0(t_n, \mathbf{x}, \omega)\right].
\end{aligned}$$

From a numerical perspective, the expected value of the FE solution at the the l-th level, $\mathbb{E}\left[y_l(t_n, \mathbf{x}, \omega)\right]$ is estimated by the sampling average $\Phi_{J_l}^n = \Phi_{J_l}\left[y_l(t_n, \mathbf{x}, \omega)\right] = \frac{1}{J_l}\sum_{j=1}^{J_l} y_l\left(t_n, \mathbf{x}, \omega_j\right)$, where J_l is the sample size. Similarly, $\mathbb{E}\left[y_L(t_n, \mathbf{x}, \omega)\right]$ is approximated by an unbiased estimator

$$\Phi\left[y_L(t_n, \mathbf{x}, \omega)\right] := \sum_{l=1}^{L} \left(\Phi_{J_l}\left[y_l(t_n, \mathbf{x}, \omega) - y_{l-1}(t_n, \mathbf{x}, \omega)\right]\right) + \Phi_{J_0}\left[y_0(t_n, \mathbf{x}, \omega)\right]. \tag{3.2}$$

At each mesh level, it becomes apparent that multiple simulations must be implemented. Hence, it makes sense to apply ensemble-based time-stepping to decrease computational expenses. Subsequently, we introduce the MLMCE technique to fulfill this goal.

At the l-th level, we let N_l be a positive integer and set both $\Delta t_l = \frac{T}{N_l}$ and $t_n = n\Delta t_l$ for $0 \le n \le N_l$. Then $[0,T] = \bigcup_{n=0}^{N_l-1} [t_n, t_{n+1}]$ is a uniform partition of the time interval. J_l independent identically distributed (i.i.d.) samples are selected, and the associated random functions are denoted by $a_j \equiv a\left(\cdot, \omega_j\right), \alpha_j \equiv \alpha\left(\cdot, \omega_j\right), f_j \equiv f\left(\cdot, \cdot, \omega_j\right), u_j \equiv u\left(\cdot, \cdot, \omega_j\right)$, and $y_j^0 \equiv y^0\left(\cdot, \omega_j\right)$ for $j = 1, \cdots, J_l$. Denote

$$\bar{a}_l := \frac{1}{J_l} \sum_{j=1}^{J_l} a\left(\mathbf{x}, \omega_j\right) \quad \text{and} \quad \bar{\alpha}_l := \frac{1}{J_l} \sum_{j=1}^{J_l} \alpha\left(\mathbf{x}, \omega_j\right)$$

are the ensemble average for the diffusion coefficient and the Robin coefficient, respectively,

$$a'_j := a_j - \bar{a}_l, \alpha'_j := \alpha_j - \bar{\alpha}_l.$$

It is important to observe that the respective exact solutions $\left\{y\left(t, \mathbf{x}, \omega_j\right)\right\}_{j=1}^{J_l}$ are i.i.d. Let us take $y_{j,l}^n = y_l\left(t_n, \mathbf{x}, \omega_j\right)$ as the FE approximation of $y\left(t_n, \mathbf{x}, \omega_j\right)$ at the l-th level. Let f_j^n, u_j^n be the values of the functions f_j, u_j at $t = t_n$.

The MLMCE method is applied to (1.1) solves the following group of simulations at the l-th level: for $j = 1, \cdots, J_l$ and $n = 1, \cdots, N_l - 1$, given $y_{j,l}^0$ and $y_{j,l}^1$, to find $y_{j,l}^{n+1} \in X_l$ such that,

$$\begin{aligned}
&\left(\frac{3y_{j,l}^{n+1} - 4y_{j,l}^n + y_{j,l}^{n-1}}{2\Delta t_l}, v_l\right) + \left(\bar{a}_l \nabla y_{j,l}^{n+1}, \nabla v_l\right) + \left(\bar{\alpha}_l y_{j,l}^{n+1}, v_l\right)_{\partial D_1} \\
&= \left(f_j^{n+1}, v_l\right) - \left(a'_j \nabla \left(2y_{j,l}^n - y_{j,l}^{n-1}\right), \nabla v_l\right) + \left(\alpha_j u_j^{n+1}, v_l\right)_{\partial D_1} - \left(\alpha'_j \left(2y_{j,l}^n - y_{j,l}^{n-1}\right), v_l\right)_{\partial D_1}, \quad \forall v_l \in X_l. \quad (3.3)
\end{aligned}$$

After determining the numerical solutions at all L levels, the MLMCE uses (3.2) to estimate $\mathbb{E}\left[y\left(t_n\right)\right]$, the random PDE solution at $t = t_n$. Simultaneously, given an interest quantity $P(y)$, one can evaluate the results from the ensemble simulations to derive the underlying stochastic characteristics of the system. The MLMCE successfully combines the ensemble-based sampling method and the ensemble-based time-stepping scheme, inheriting the advantages of both. Like the MLMC, it can also reduce the computational expenses by adaptively choosing the time step size, mesh size, and the number of samples at each level. Moreover, the ensemble-based time-stepping scheme leads to a discrete linear system (3.3) with the coefficient matrix being independent of j.

In the subsequent sections, we will critically analyze the stability and asymptotic error estimation inherent in the MLMCE methodology.

4. Stability and error estimates

This section presents the results of our stability and error estimation analyses. The MLMCE approximation is contingent on the MC solutions derived from various levels. Therefore, our exploration commences with a thorough investigation of the ensemble-based single-level MC in Subsection 4.1. Following this, we subsequently deduce the error estimation for the MLMCE method in Subsection 4.2.

Assume that the exact solution of Eq (1.1) possesses sufficient smoothness, specifically

$$\begin{aligned}
y_j \in\ & \widetilde{L}^2\left(0, T; H^1(\bar{D}) \cap H^{k+1}(\bar{D})\right) \cap \widetilde{H}^1\left(0, T; H^{k+1}(\bar{D})\right) \\
&\cap \widetilde{H}^2\left(0, T; H^1(\bar{D})\right) \cap \widetilde{H}^3\left(0, T; L^2(\bar{D})\right),
\end{aligned}$$

and assume

$$f_j \in \tilde{L}^2\left(0,T;H^{-1}(D)\right), u_j \in \tilde{L}^2\left(0,T;L^2(\partial D_1)\right).$$

We highlight that the required regularity for the random fields only requires that they are square integrable. We assume finite-dimensional noise, that is, all the involved random input data depend on a finite number of real-valued random variables. Assumptions of input data for two coefficients are considered, as follows:

(i) $a = a(\mathbf{x},\omega) \in L_P^2\left(L^2(D);\Omega\right)$ is uniformly bounded and coercive, i.e., $\exists\, a_{\min}$, $a_{\max} > 0$ such that $P\{\omega \in \Omega : a(\mathbf{x},\omega) \in [a_{\min}, a_{\max}], \forall\, \mathbf{x} \in D\} = 1$.

(ii) $\alpha = \alpha(\mathbf{x},\omega) \in L_P^2\left(L^2\left(\partial D_1\right);\Omega\right)$ is also uniformly bounded and coercive, i.e., $\exists\, \alpha_{\min}$, $\alpha_{\max} > 0$ such that $P\{\omega \in \Omega : \alpha(\mathbf{x},\omega) \in [\alpha_{\min}, \alpha_{\max}], \forall\, \mathbf{x} \in \partial D_1\} = 1$, and in addition,

$$\int_\Omega \int_{\partial D_1} \alpha(\mathbf{x},\omega) d\mathbf{s} dP(\omega) > 0.$$

Denote $a^* = \max\limits_{1\leqslant j\leqslant J_l} \left\|\frac{a'_j}{\bar{a}}\right\|_\infty$ and $\alpha^* = \max\limits_{1\leqslant j\leqslant J_l} \left\|\frac{\alpha'_j}{\bar{\alpha}}\right\|_\infty$, we suppose the following condition (4.1) is valid for the second-order ensemble time-stepping method (3.3):

$$1 - 3a^* > 0 \text{ and } 1 - 3\alpha^* > 0. \tag{4.1}$$

4.1. Single-level MC ensemble FE method

When the expression $\mathbb{E}\left[y\left(t_n\right)\right]$ is approximated numerically by $\Phi_{J_l}^n$, the resulting approximation error can be partitioned into two components

$$\mathbb{E}\left[y\left(t_n\right)\right] - \Phi_{J_l}^n = \left(\mathbb{E}\left[y_j\left(t_n\right)\right] - \mathbb{E}\left[y_{j,l}^n\right]\right) + \left(\mathbb{E}\left[y_{j,l}^n\right] - \Phi_{J_l}^n\right) := \mathcal{E}_l^n + \mathcal{E}_S^n,$$

where we acknowledge that $\mathbb{E}\left[y\left(t_n\right)\right] = \mathbb{E}\left[y_j\left(t_n\right)\right]$. The error arising from FE discretization, denoted $\mathcal{E}_l^n = \mathbb{E}\left[y_j\left(t_n\right) - y_{j,l}^n\right]$, is influenced by the dimensions of the spatial triangulations $\mathcal{T}_l$ and the time step, while the statistical sampling error, represented as $\mathcal{E}_S^n = \mathbb{E}\left[y_{j,l}^n\right] - \Phi_{J_l}^n$, is chiefly affected by the number of realizations and the variance involved. We will first address the stability of the ensemble method presented in Eq (3.3) at the l-th level (Theorem 4.1), establish bounds for $\mathcal{E}_S^n$ (Theorem 4.2) and $\mathcal{E}_l^n$ (Theorem 4.3), followed by deriving the asymptotic error estimation (Theorem 4.4). The stability of the ensemble scheme in Eq (3.3) yields the following result.

Theorem 4.1. *Under the statements (i) and (ii), the scheme* (3.3) *is stable according to the premise of condition* (4.1). *Moreover, the numerical solution to* (3.3) *satisfies*

$$\begin{aligned}
&\mathbb{E}\left[\left\|y_{j,l}^{N_l}\right\|^2\right] + \mathbb{E}\left[\left\|2y_{j,l}^{N_l} - y_{j,l}^{N_l-1}\right\|^2\right] + 2\Delta t_l\left(1 - 3a^*\right)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^{n+1}\right\|^2\right] \\
&+ 2\Delta t_l\left(1 - 3\alpha^*\right)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right] + 6\Delta t_l a^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^{N_l}\right\|^2\right] + 2\Delta t_l a^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^{N_l-1}\right\|^2\right] \\
&+ 6\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{N_l}\right\|_{\partial D_1}^2\right] + 2\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{N_l-1}\right\|_{\partial D_1}^2\right]
\end{aligned}$$

$$\leq C\Delta t_l \sum_{n=1}^{N_l-1} \mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1}^2\right] + C\Delta t_l \sum_{n=1}^{N_l-1} \mathbb{E}\left[\left\|u_j^{n+1}\right\|_{\partial D_1}^2\right] + \mathbb{E}\left[\left\|y_{j,l}^1\right\|^2\right] + \mathbb{E}\left[\left\|2y_{j,l}^1 - y_{j,l}^0\right\|^2\right]$$
$$+ C\Delta t_l \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^1\right\|^2\right] + C\Delta t_l \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^0\right\|^2\right] + C\Delta t_l \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^1\right\|_{\partial D_1}^2\right] + C\Delta t_l \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^0\right\|_{\partial D_1}^2\right]. \quad (4.2)$$

Proof. Taking $v_l = y_{j,l}^{n+1}$ in (3.3), we have

$$\begin{aligned}
&\left(\frac{3y_{j,l}^{n+1} - 4y_{j,l}^n + y_{j,l}^{n-1}}{2\Delta t_l}, y_{j,l}^{n+1}\right) + \left(\bar{a}_l \nabla y_{j,l}^{n+1}, \nabla y_{j,l}^{n+1}\right) + \left(\bar{\alpha}_l y_{j,l}^{n+1}, y_{j,l}^{n+1}\right)_{\partial D_1} \\
=&\left(f_j^{n+1}, y_{j,l}^{n+1}\right) - \left(a_j' \nabla\left(2y_{j,l}^n - y_{j,l}^{n-1}\right), \nabla y_{j,l}^{n+1}\right) \\
&+ \left(\alpha_j u_j^{n+1}, y_{j,l}^{n+1}\right)_{\partial D_1} - \left(\alpha_j'\left(2y_{j,l}^n - y_{j,l}^{n-1}\right), y_{j,l}^{n+1}\right)_{\partial D_1}. \quad (4.3)
\end{aligned}$$

Applying the polarization identity, multiplying both sides by Δt_l, and integrating over the probability space, we obtain

$$\begin{aligned}
&\frac{1}{4}\mathbb{E}\left[\left\|y_{j,l}^{n+1}\right\|^2 + \left\|2y_{j,l}^{n+1} - y_{j,l}^n\right\|^2\right] - \frac{1}{4}\mathbb{E}\left[\left\|y_{j,l}^n\right\|^2 + \left\|2y_{j,l}^n - y_{j,l}^{n-1}\right\|^2\right] \\
&+ \frac{1}{4}\mathbb{E}\left[\left\|y_{j,l}^{n+1} - 2y_{j,l}^n + y_{j,l}^{n-1}\right\|^2\right] + \Delta t_l \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{n+1}\right\|^2\right] + \Delta t_l \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right] \\
=&\Delta t_l \mathbb{E}\left[\left(f_j^{n+1}, y_{j,l}^{n+1}\right)\right] - \Delta t_l \mathbb{E}\left[\left(a_j' \nabla\left(2y_{j,l}^n - y_{j,l}^{n-1}\right), \nabla y_{j,l}^{n+1}\right)\right] \\
&+ \Delta t_l \mathbb{E}\left[\left(\alpha_j u_j^{n+1}, y_{j,l}^{n+1}\right)_{\partial D_1}\right] - \Delta t_l \mathbb{E}\left[\left(\alpha_j'\left(2y_{j,l}^n - y_{j,l}^{n-1}\right), y_{j,l}^{n+1}\right)_{\partial D_1}\right]. \quad (4.4)
\end{aligned}$$

The Cauchy-Schwarz and the Young's inequalities applied to the RHS terms of (4.4) give the following estimation, for $\forall \epsilon_i > 0, i = 1, \cdots, 6$:

$$\begin{aligned}
&\mathbb{E}\left[\left(f_j^{n+1}, y_{j,l}^{n+1}\right)\right] \\
\leq&\mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1} \left\|y_{j,l}^{n+1}\right\|_1\right] \\
\leq&\mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1} C_1\left(\left\|\nabla y_{j,l}^{n+1}\right\|^2 + \left\|y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right)^{\frac{1}{2}}\right] \\
\leq&\mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1} \frac{C_1}{\min\{a_{\min}^{\frac{1}{2}}, \alpha_{\min}^{\frac{1}{2}}\}}\left(\bar{a}_l\left\|\nabla y_{j,l}^{n+1}\right\|^2 + \bar{\alpha}_l\left\|y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right)^{\frac{1}{2}}\right] \\
\leq&\frac{C_1^2}{2\epsilon_1 \min\{a_{\min}, \alpha_{\min}\}}\mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1}^2\right] + \frac{\epsilon_1}{2}\left(\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{n+1}\right\|^2\right] + \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right]\right), \quad (4.5)
\end{aligned}$$

$$\begin{aligned}
&- \mathbb{E}\left[\left(a_j' \nabla\left(2y_{j,l}^n - y_{j,l}^{n-1}\right), \nabla y_{j,l}^{n+1}\right)\right] \\
\leq&\mathbb{E}\left[\left|\left(2a_j' \nabla y_{j,l}^n, \nabla y_{j,l}^{n+1}\right) - \left(a_j' \nabla y_{j,l}^{n-1}, \nabla y_{j,l}^{n+1}\right)\right|\right] \\
\leq&\frac{\epsilon_2 + \epsilon_3}{2} a^* \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{n+1}\right\|^2\right] + \frac{2}{\epsilon_2} a^* \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^n\right\|^2\right] + \frac{1}{2\epsilon_3} a^* \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{n-1}\right\|^2\right], \quad (4.6)
\end{aligned}$$

$$
\begin{aligned}
&\mathbb{E}\left[\left(\alpha_j u_j^{n+1}, y_{j,l}^{n+1}\right)_{\partial D_1}\right] \\
\leq &\mathbb{E}\left[\left|\left(\alpha_j u_j^{n+1}, y_{j,l}^{n+1}\right)\right|_{\partial D_1}\right] \\
\leq &\mathbb{E}\left[\alpha_{\max}\left\|u_j^{n+1}\right\|_{\partial D_1}\left\|y_{j,l}^{n+1}\right\|_{\partial D_1}\right] \\
\leq &\frac{\alpha_{\max}^2}{2\epsilon_4\alpha_{\min}}\mathbb{E}\left[\left\|u_j^{n+1}\right\|_{\partial D_1}^2\right] + \frac{\epsilon_4}{2}\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right],
\end{aligned} \tag{4.7}
$$

and

$$
\begin{aligned}
&-\mathbb{E}\left[\left(\alpha_j'\left(2y_{j,l}^n - y_{j,l}^{n-1}\right), y_{j,l}^{n+1}\right)_{\partial D_1}\right] \\
\leq &\mathbb{E}\left[\left|\left(2\alpha_j' y_{j,l}^n, y_{j,l}^{n+1}\right)_{\partial D_1} - \left(\alpha_j' y_{j,l}^{n-1}, y_{j,l}^{n+1}\right)_{\partial D_1}\right|\right] \\
\leq &\frac{\epsilon_5+\epsilon_6}{2}\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right] + \frac{2}{\epsilon_5}\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n}\right\|_{\partial D_1}^2\right] + \frac{1}{2\epsilon_6}\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n-1}\right\|_{\partial D_1}^2\right].
\end{aligned} \tag{4.8}
$$

Substituting (4.5)–(4.8) into (4.4), choosing $\epsilon_2 = \epsilon_5 = 2, \epsilon_3 = \epsilon_6 = 1$ and dropping the non-negative term $\mathbb{E}\left[\frac{1}{4}\|y_{j,l}^{n+1} - 2y_{j,l}^n + y_{j,l}^{n-1}\|^2\right]$, we obtain

$$
\begin{aligned}
&\frac{1}{4}\mathbb{E}\left[\left\|y_{j,l}^{n+1}\right\|^2 + \left\|2y_{j,l}^{n+1} - y_{j,l}^n\right\|^2\right] - \frac{1}{4}\mathbb{E}\left[\left\|y_{j,l}^n\right\|^2 + \left\|2y_{j,l}^n - y_{j,l}^{n-1}\right\|^2\right] \\
&+\Delta t_l\left(1 - \frac{3a^*}{2} - \frac{\epsilon_1}{2}\right)\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^{n+1}\right\|^2\right] + \Delta t_l\left(1 - \frac{3\alpha^*}{2} - \frac{\epsilon_1}{2} - \frac{\epsilon_4}{2}\right)\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right] \\
&-\Delta t_l a^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^n\right\|^2\right] - \frac{\Delta t_l}{2}a^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^{n-1}\right\|^2\right] \\
&-\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^n\right\|_{\partial D_1}^2\right] - \frac{\Delta t_l}{2}\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n-1}\right\|_{\partial D_1}^2\right] \\
\leq &\Delta t_l\frac{C_1^2}{2\epsilon_1\min\{a_{\min}, \alpha_{\min}\}}\mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1}^2\right] + \Delta t_l\frac{\alpha_{\max}^2}{2\epsilon_4\alpha_{\min}}\mathbb{E}\left[\left\|u_j^{n+1}\right\|_{\partial D_1}^2\right].
\end{aligned} \tag{4.9}
$$

Multiplying this by 4 and rearranging, we obain

$$
\begin{aligned}
&\mathbb{E}\left[\left\|y_{j,l}^{n+1}\right\|^2 + \left\|2y_{j,l}^{n+1} - y_{j,l}^n\right\|^2\right] - \mathbb{E}\left[\left\|y_{j,l}^n\right\|^2 + \left\|2y_{j,l}^n - y_{j,l}^{n-1}\right\|^2\right] \\
&+4\Delta t_l\left(1 - 3a^* - \frac{\epsilon_1}{2}\right)\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^{n+1}\right\|^2\right] + 4\Delta t_l\left(1 - 3\alpha^* - \frac{\epsilon_1}{2} - \frac{\epsilon_4}{2}\right)\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right] \\
&+6\Delta t_l a^*\left(\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^{n+1}\right\|^2\right] - \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^n\right\|^2\right]\right) \\
&+2\Delta t_l a^*\left(\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^n\right\|^2\right] - \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^{n-1}\right\|^2\right]\right) \\
&+6\Delta t_l\alpha^*\left(\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right] - \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^n\right\|_{\partial D_1}^2\right]\right) \\
&+2\Delta t_l\alpha^*\left(\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^n\right\|_{\partial D_1}^2\right] - \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n-1}\right\|_{\partial D_1}^2\right]\right)
\end{aligned}
$$

$$\leq \Delta t_l \frac{2C_1^2}{\epsilon_1 \min\{a_{\min}, \alpha_{\min}\}} \mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1}^2\right] + \Delta t_l \frac{2\alpha_{\max}^2}{\epsilon_4 \alpha_{\min}} \mathbb{E}\left[\left\|u_j^{n+1}\right\|_{\partial D_1}^2\right]. \tag{4.10}$$

Choosing $\epsilon_4 = 1 - 3\alpha^*$ and $\epsilon_1 = \min\left\{1 - 3a^*, \frac{1-3\alpha^*}{2}\right\}$ and using the conditions in (4.1), we can obtain

$$\begin{aligned}
&\mathbb{E}\left[\left\|y_{j,l}^{n+1}\right\|^2 + \left\|2y_{j,l}^{n+1} - y_{j,l}^{n}\right\|^2\right] - \mathbb{E}\left[\left\|y_{j,l}^{n}\right\|^2 + \left\|2y_{j,l}^{n} - y_{j,l}^{n-1}\right\|^2\right] \\
&+ 2\Delta t_l (1 - 3a^*) \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{n+1}\right\|^2\right] + 2\Delta t_l (1 - 3\alpha^*) \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right] \\
&+ 6\Delta t_l a^* \left(\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{n+1}\right\|^2\right] - \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{n}\right\|^2\right]\right) \\
&+ 2\Delta t_l a^* \left(\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{n}\right\|^2\right] - \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{n-1}\right\|^2\right]\right) \\
&+ 6\Delta t_l \alpha^* \left(\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right] - \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n}\right\|_{\partial D_1}^2\right]\right) \\
&+ 2\Delta t_l \alpha^* \left(\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n}\right\|_{\partial D_1}^2\right] - \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n-1}\right\|_{\partial D_1}^2\right]\right) \\
\leq &\frac{2\Delta t_l C_1^2}{\min\left\{1 - 3a^*, \frac{1-3\alpha^*}{2}\right\} \min\{a_{\min}, \alpha_{\min}\}} \mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1}^2\right] + \frac{2\Delta t_l \alpha_{\max}^2}{(1 - 3\alpha^*)\, \alpha_{\min}} \mathbb{E}\left[\left\|u_j^{n+1}\right\|_{\partial D_1}^2\right].
\end{aligned} \tag{4.11}$$

Summing over n from 1 to $N_l - 1$ yields

$$\begin{aligned}
&\mathbb{E}\left[\left\|y_{j,l}^{N_l}\right\|^2\right] + \mathbb{E}\left[\left\|2y_{j,l}^{N_l} - y_{j,l}^{N_l-1}\right\|^2\right] \\
&+ 2\Delta t_l (1 - 3a^*) \sum_{n=1}^{N_l-1} \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{n+1}\right\|^2\right] + 2\Delta t_l (1 - 3\alpha^*) \sum_{n=1}^{N_l-1} \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{n+1}\right\|_{\partial D_1}^2\right] \\
&+ 6\Delta t_l a^* \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{N_l}\right\|^2\right] + 2\Delta t_l a^* \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{N_l-1}\right\|^2\right] \\
&+ 6\Delta t_l \alpha^* \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{N_l}\right\|_{\partial D_1}^2\right] + 2\Delta t_l \alpha^* \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{N_l-1}\right\|_{\partial D_1}^2\right] \\
\leq &\frac{2\Delta t_l C_1^2}{\min\left\{1 - 3a^*, \frac{1-3\alpha^*}{2}\right\} \min\{a_{\min}, \alpha_{\min}\}} \sum_{n=1}^{N_l-1} \mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1}^2\right] \\
&+ \frac{2\Delta t_l \alpha_{\max}^2}{(1 - 3\alpha^*)\, \alpha_{\min}} \sum_{n=1}^{N_l-1} \mathbb{E}\left[\left\|u_j^{n+1}\right\|_{\partial D_1}^2\right] + \mathbb{E}\left[\left\|y_{j,l}^{1}\right\|^2\right] + \mathbb{E}\left[\left\|2y_{j,l}^{1} - y_{j,l}^{0}\right\|^2\right] \\
&+ 6\Delta t_l a^* \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{1}\right\|^2\right] + 2\Delta t_l a^* \mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}} \nabla y_{j,l}^{0}\right\|^2\right] \\
&+ 6\Delta t_l \alpha^* \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{1}\right\|_{\partial D_1}^2\right] + 2\Delta t_l \alpha^* \mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}} y_{j,l}^{0}\right\|_{\partial D_1}^2\right].
\end{aligned} \tag{4.12}$$

This completes the proof. □

Remark 4.1. *C_1 is a constant of two equivalent norms, i.e., $\|\cdot\|_1 \leq C_1\|\cdot\|_{1,\partial D_1}$. The time-stepping scheme based on ensemble methods, denoted as* (3.3)*, will be stable if Condition* (4.1) *is fulfilled. Additionally, it reaches unconditional stability when the ensemble size is reduced to one, resulting in a^* and α^* diminishing to zero. Consequently, when addressing a set of problems, one may utilize Condition* (4.1) *as a criterion to categorize these problems into subgroups, ensuring that Condition* (4.1) *is valid within each subgroup. The smallest subgroup might comprise just a single element, as no stability conditions would be necessary in that case.*

Furthermore, by applying the conventional error estimate associated with the MC method (refer to [23]), it is feasible to establish a bound on the statistical error $\mathcal{E}_S^n$ in the following manner.

Theorem 4.2. *Assume $\mathcal{E}_S^n = \mathbb{E}\left[y_{j,l}^n\right] - \Phi_{J_1}^n$, where $y_{j,l}^n$ represents the output of the scheme* (3.3) *and $\Phi_{J_l}^n = \frac{1}{J_l}\sum_{j=1}^{J_l} y_{j,l}^n$. Assuming Conditions (i) and (ii) as well as the stability condition* (4.1) *are satisfied, a generic positive constant C exists such that*

$$
\begin{aligned}
&\mathbb{E}\left[\left\|\mathcal{E}_S^{N_l}\right\|^2\right] + \mathbb{E}\left[\left\|2\mathcal{E}_S^{N_l} - \mathcal{E}_S^{N_l-1}\right\|^2\right] + 2\Delta t_l\,(1-3a^*)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\mathcal{E}_S^{n+1}\right\|^2\right] \\
&+ 2\Delta t_l\,(1-3\alpha^*)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\mathcal{E}_S^{n+1}\right\|_{\partial D_1}^2\right] + 6\Delta t_l a^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\mathcal{E}_S^{N_l}\right\|^2\right] \\
&+ 2\Delta t_l a^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\mathcal{E}_S^{N_l-1}\right\|^2\right] + 6\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\mathcal{E}_S^{N_l}\right\|_{\partial D_1}^2\right] + 2\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\mathcal{E}_S^{N_l-1}\right\|_{\partial D_1}^2\right] \\
\leq&\frac{C}{J_l}\left(\Delta t_l\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1}^2\right] + \Delta t_l\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|u_j^{n+1}\right\|_{\partial D_1}^2\right] + \mathbb{E}\left[\left\|y_{j,l}^1\right\|^2\right] + \mathbb{E}\left[\left\|2y_{j,l}^1 - y_{j,l}^0\right\|^2\right]\right. \\
&\left.+\Delta t_l\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^1\right\|^2\right] + \Delta t_l\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^0\right\|^2\right] + \Delta t_l\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}y_{j,l}^1\right\|_{\partial D_1}^2\right] + \Delta t_l\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}y_{j,l}^0\right\|_{\partial D_1}^2\right]\right).
\end{aligned}
\tag{4.13}
$$

Proof. First, for $n = 1, \cdots, N_l - 1$, we assess $\mathbb{E}\left[\left\|\nabla\mathcal{E}_S^n\right\|^2\right]$.

$$
\begin{aligned}
\mathbb{E}\left[\left\|\nabla\mathcal{E}_S^n\right\|^2\right] &= \mathbb{E}\left[\left(\frac{1}{J_l}\sum_{i=1}^{J_l}\left(\nabla\mathbb{E}\left[y_{i,l}^n\right] - \nabla y_{i,l}^n\right), \frac{1}{J_l}\sum_{j=1}^{J_l}\left(\nabla\mathbb{E}\left[y_{j,l}^n\right] - \nabla y_{j,l}^n\right)\right)\right] \\
&= \frac{1}{J_l^2}\sum_{i,j=1}^{J_l}\mathbb{E}\left[\left(\nabla\mathbb{E}\left[y_l^n\right] - \nabla y_{i,l}^n, \nabla\mathbb{E}\left[y_l^n\right] - \nabla y_{j,l}^n\right)\right] \\
&= \frac{1}{J_l^2}\sum_{j=1}^{J_l}\mathbb{E}\left[\left(\nabla\mathbb{E}\left[y_l^n\right] - \nabla y_{j,l}^n, \nabla\mathbb{E}\left[y_l^n\right] - \nabla y_{j,l}^n\right)\right]
\end{aligned}
$$

The last equality is due to the fact that $y_{1,l}^n, \cdots, y_{J_l,l}^n$ are i.i.d., and $\mathbb{E}\left[\left(\nabla\mathbb{E}\left[y_l^n\right] - \nabla y_{i,l}^n, \nabla\mathbb{E}\left[y_l^n\right] - \nabla y_{j,l}^n\right)\right]$ is a zero for $i \neq j$. We now calculate $\mathbb{E}\left[\left(\nabla\mathbb{E}\left[y_l^n\right] - \nabla y_{j,l}^n, \nabla\mathbb{E}\left[y_l^n\right] - \nabla y_{j,l}^n\right)\right]$ and use the fact that $\mathbb{E}\left[\nabla y_{j,l}^n\right] = \nabla\mathbb{E}\left[y_{j,l}^n\right]$ and $\mathbb{E}\left[y_l^n\right] = \mathbb{E}\left[y_{j,l}^n\right]$ to obtain

$$
\mathbb{E}\left[\left\|\nabla\mathcal{E}_S^n\right\|^2\right] = -\frac{1}{J_l}\left\|\nabla\mathbb{E}\left[y_l^n\right]\right\|^2 + \frac{1}{J_l}\mathbb{E}\left[\left\|\nabla y_{j,l}^n\right\|^2\right]
$$

which yields

$$\mathbb{E}\left[\left\|\nabla\mathcal{E}_S^n\right\|^2\right] \le \frac{1}{J_l}\mathbb{E}\left[\left\|\nabla y_{j,l}^n\right\|^2\right],$$

then

$$\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\mathcal{E}_S^{n+1}\right\|^2\right] \le \frac{1}{J_l}\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^{n+1}\right\|^2\right].$$

With the help of Theorem 4.1, we have

$$\begin{aligned}
&2\Delta t_l(1-3a^*)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\mathcal{E}_S^{n+1}\right\|^2\right]\\
\le&\frac{C}{J_l}\Bigg(\Delta t_l\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1}^2\right]+\Delta t_l\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|u_j^{n+1}\right\|_{\partial D_1}^2\right]+\mathbb{E}\left[\left\|y_{j,l}^1\right\|^2\right]+\mathbb{E}\left[\left\|2y_{j,l}^1-y_{j,l}^0\right\|^2\right]\\
&+\Delta t_l\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^1\right\|^2\right]+\Delta t_l\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla y_{j,l}^0\right\|^2\right]+\Delta t_l\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}y_{j,l}^1\right\|_{\partial D_1}^2\right]+\Delta t_l\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}y_{j,l}^0\right\|_{\partial D_1}^2\right]\Bigg).
\end{aligned}\tag{4.14}$$

The remaining terms on the left-hand side of (4.13) can be addressed in an analogous fashion, which completes the proof. □

We will now assess the finite element discretization error $\mathcal{E}_l^n$.

Theorem 4.3. *Assume* $\mathcal{E}_l^n=\mathbb{E}\left[y_j(t_n)-y_{j,l}^n\right]$*, where* $y_j(t_n)$ *refers to the solution of Eq* (1.1) *at* $\omega=\omega_j$ *and* $t=t_n$*, while* $y_{j,l}^n$ *represents the outcome derived from the scheme* (3.3)*. We assume that the initial errors* $\left\|y_j(t_0)-y_{j,l}^0\right\|$, $\left\|y_j(t_1)-y_{j,l}^1\right\|$, $\left\|\nabla\left(y_j(t_0)-y_{j,l}^0\right)\right\|$*, and* $\left\|\nabla\left(y_j(t_1)-y_{j,l}^1\right)\right\|$ *are all at least* $O\left(h_l^k\right)$*. Under the assumptions of Conditions (i) and (ii), along with the stability condition* (4.1)*, then a constant C exists and satisfies*

$$\begin{aligned}
&\mathbb{E}\left[\left\|\mathcal{E}_l^{N_l}\right\|^2\right]+\mathbb{E}\left[\left\|2\mathcal{E}_l^{N_l}-\mathcal{E}_l^{N_l-1}\right\|^2\right]+2\Delta t_l(1-3a^*)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\mathcal{E}_l^{n+1}\right\|^2\right]\\
&+2\Delta t_l(1-3\alpha^*)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\mathcal{E}_l^{n+1}\right\|_{\partial D_1}^2\right]+6\Delta t_l a^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\mathcal{E}_l^{N_l}\right\|^2\right]\\
&+2\Delta t_l a^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\mathcal{E}_l^{N_l-1}\right\|^2\right]+6\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\mathcal{E}_l^{N_l}\right\|_{\partial D_1}^2\right]+2\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\mathcal{E}_l^{N_l-1}\right\|_{\partial D_1}^2\right]\\
\le&C\left(\Delta t_l^4+h_l^{2k}\right).
\end{aligned}\tag{4.15}$$

Proof. Firstly, the error equation for (3.3) is derived. The true solution $y_j(t_{n+1})$ of Eq (1.2) at t_{n+1} $(n=1,\cdots,N_l-1)$ and tested by v_l produces

$$\begin{aligned}
&\left(\frac{3y_j(t_{n+1})-4y_j(t_n)+y_j(t_{n-1})}{2\Delta t_l},v_l\right)+\left(a_j\nabla y_j(t_{n+1}),\nabla v_l\right)+\left(\alpha_j y_j(t_{n+1}),v_l\right)_{\partial D_1}\\
=&\left(f_j^{n+1},v_l\right)+\left(\alpha_j u_j^{n+1},v_l\right)_{\partial D_1}-\left(R_j^{n+1},v_l\right),\quad\forall v_l\in X_l,
\end{aligned}\tag{4.16}$$

where $f_j^{n+1} = f_j(t_{n+1})$, $u_j^{n+1} = u_j(t_{n+1})$, and

$$R_j^{n+1} = y_{j,t}(t_{n+1}) - \frac{3y_j(t_{n+1}) - 4y_j(t_n) + y_j(t_{n-1})}{2\Delta t_l}.$$

Let $e_j^n := y_j(t_n) - y_{j,l}^n$. Subtracting (3.3) from (4.16) obtains the error equation

$$\begin{aligned}
&\left(\frac{3e_j^{n+1} - 4e_j^n + e_j^{n-1}}{2\Delta t_l}, v_l\right) + \left(R_j^{n+1}, v_l\right) \\
&+ \left(\bar{a}_l \nabla e_j^{n+1}, \nabla v_l\right) + \left(a_j' \nabla\left(2e_j^n - e_j^{n-1}\right), \nabla v_l\right) + \left(a_j' \nabla\left(y_j^{n+1} - 2y_j^n + y_j^{n-1}\right), \nabla v_l\right) \\
&+ \left(\bar{\alpha}_l e_j^{n+1}, v_l\right)_{\partial D_1} + \left(\alpha_j'\left(2e_j^n - e_j^{n-1}\right), v_l\right)_{\partial D_1} + \left(\alpha_j'\left(y_j^{n+1} - 2y_j^n + y_j^{n-1}\right), v_l\right)_{\partial D_1} \\
=&0.
\end{aligned} \tag{4.17}$$

Secondly, the error can be decomposed. Let $Q_l\left(y_j(t_n)\right)$ be the L^2 projection of $y_j(t_n)$ onto $X_l \cap \partial D_1$ satisfying

$$\left(y_j(t_n) - Q_l\left(y_j(t_n)\right), v_l\right) = 0, \quad \forall v_l \in X_l,$$

and

$$\left(y_j(t_n) - Q_l\left(y_j(t_n)\right), v_l\right)_{\partial D_1} = 0, \quad \forall v_l \in X_l \cap \partial D_1.$$

We now decompose the error

$$e_j^n = \rho_{j,l}^n - \phi_{j,l}^n, \text{ where } \rho_{j,l}^n = y_j(t_n) - Q_l\left(y_j(t_n)\right) \text{ and } \phi_{j,l}^n = y_{j,l}^n - Q_l\left(y_j(t_n)\right).$$

Upon substituting this decomposition into (4.17) and choosing $v_l = \phi_{j,l}^{n+1}$, we obtain

$$\begin{aligned}
&\left(\frac{3\phi_{j,l}^{n+1} - 4\phi_{j,l}^n + \phi_{j,l}^{n-1}}{2\Delta t_l}, \phi_{j,l}^{n+1}\right) + \left(\bar{a}_l \nabla \phi_{j,l}^{n+1}, \nabla \phi_{j,l}^{n+1}\right) + \left(\bar{\alpha}_l \phi_{j,l}^{n+1}, \phi_{j,l}^{n+1}\right)_{\partial D_1} \\
=& -\left(a_j' \nabla\left(2\phi_{j,l}^n - \phi_{j,l}^{n-1}\right), \nabla \phi_{j,l}^{n+1}\right) + \left(\frac{3\rho_{j,l}^{n+1} - 4\rho_{j,l}^n + \rho_{j,l}^{n-1}}{2\Delta t_l}, \phi_{j,l}^{n+1}\right) \\
&+ \left(\bar{a}_l \nabla \rho_{j,l}^{n+1}, \nabla \phi_{j,l}^{n+1}\right) + \left(a_j' \nabla\left(2\rho_{j,l}^n - \rho_{j,l}^{n-1}\right), \nabla \phi_{j,l}^{n+1}\right) + \left(R_j^{n+1}, \phi_{j,l}^{n+1}\right) \\
&+ \left(a_j' \nabla\left(y_j^{n+1} - 2y_j^n + y_j^{n-1}\right), \nabla \phi_{j,l}^{n+1}\right) + \left(\alpha_j'\left(y_j^{n+1} - 2y_j^n + y_j^{n-1}\right), \phi_{j,l}^{n+1}\right)_{\partial D_1} \\
&+ \left(\bar{\alpha}_l \rho_{j,l}^{n+1}, \phi_{j,l}^{n+1}\right)_{\partial D_1} - \left(\alpha_j'\left(2\phi_{j,l}^n - \phi_{j,l}^{n-1}\right), \phi_{j,l}^{n+1}\right)_{\partial D_1} + \left(\alpha_j'\left(2\rho_{j,l}^n - \rho_{j,l}^{n-1}\right), \phi_{j,l}^{n+1}\right)_{\partial D_1}.
\end{aligned} \tag{4.18}$$

For the left-hand-side (LHS) of (4.18), using the polarization identity and integrating over probability space, we get

$$\begin{aligned}
&\left(\frac{3\phi_{j,l}^{n+1} - 4\phi_{j,l}^n + \phi_{j,l}^{n-1}}{2\Delta t_l}, \phi_{j,l}^{n+1}\right) + \left(\bar{a}_l \nabla \phi_{j,l}^{n+1}, \nabla \phi_{j,l}^{n+1}\right) + \left(\bar{\alpha}_l \phi_{j,l}^{n+1}, \phi_{j,l}^{n+1}\right)_{\partial D_1} \\
=&\frac{1}{4\Delta t_l}\mathbb{E}\left[\left\|\phi_{j,l}^{n+1}\right\|^2 + \left\|2\phi_{j,l}^{n+1} - \phi_{j,l}^n\right\|^2\right] - \frac{1}{4\Delta t_l}\mathbb{E}\left[\left\|\phi_{j,l}^n\right\|^2 + \left\|2\phi_{j,l}^n - \phi_{j,l}^{n-1}\right\|^2\right]
\end{aligned}$$

$$+\frac{1}{4\Delta t_l}\mathbb{E}\left[\left\|\phi_{j,l}^{n+1}-2\phi_{j,l}^{n}+\phi_{j,l}^{n-1}\right\|^2\right]+\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n+1}\right\|^2\right]+\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{n+1}\right\|_{\partial D_1}^2\right]. \tag{4.19}$$

The Cauchy-Schwarz and Young's inequalities are applied to the RHS of (4.18) one by one give the following estimations:

$$\begin{aligned}
&\mathbb{E}\left[\left|\left(a_j'\nabla\left(2\phi_{j,l}^{n}-\phi_{j,l}^{n-1}\right),\nabla\phi_{j,l}^{n+1}\right)\right|\right]\\
\le&2a^*\mathbb{E}\left[\left|\left(\bar{a}_l\nabla\phi_{j,l}^{n},\nabla\phi_{j,l}^{n+1}\right)\right|\right]+a^*\mathbb{E}\left[\left|\left(\bar{a}_l\nabla\phi_{j,l}^{n-1},\nabla\phi_{j,l}^{n+1}\right)\right|\right]\\
\le&a^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n}\right\|^2\right]+\frac{a^*}{2}\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n-1}\right\|^2\right]+\frac{3a^*}{2}\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n+1}\right\|^2\right],
\end{aligned} \tag{4.20}$$

$$\mathbb{E}\left[\left|\left(\frac{3\rho_{j,l}^{n+1}-4\rho_j^{n}+\rho_{j,l}^{n-1}}{2\Delta t_l},\phi_{j,l}^{n+1}\right)\right|\right]=0, \tag{4.21}$$

$$\begin{aligned}
&\mathbb{E}\left[\left|\left(\bar{a}_l\nabla\rho_{j,l}^{n+1},\nabla\phi_{j,l}^{n+1}\right)\right|\right]\\
\le&\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\rho_{j,l}^{n+1}\right\|\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n+1}\right\|\right]\\
\le&\frac{a_{\max}}{2\epsilon_1}\mathbb{E}\left[\left\|\nabla\rho_{j,l}^{n+1}\right\|^2\right]+\frac{\epsilon_1}{2}\mathbb{E}\left[\left\|\bar{a}^{\frac{1}{2}}\nabla\phi_{j,l}^{n+1}\right\|^2\right],
\end{aligned} \tag{4.22}$$

$$\begin{aligned}
&\mathbb{E}\left[\left|\left(a_j'\nabla\left(2\rho_{j,l}^{n}-\rho_{j,l}^{n-1}\right),\nabla\phi_{j,l}^{n+1}\right)\right|\right]\\
\le&a^*\left(\frac{a_{\max}}{2\epsilon_2}\mathbb{E}\left[\left\|\nabla\left(2\rho_{j,l}^{n}-\rho_{j,l}^{n-1}\right)\right\|^2\right]+\frac{\epsilon_2}{2}\mathbb{E}\left[\left\|\bar{a}^{\frac{1}{2}}\nabla\phi_{j,l}^{n+1}\right\|^2\right]\right),
\end{aligned} \tag{4.23}$$

$$\begin{aligned}
&\mathbb{E}\left[\left|\left(R_j^{n+1},\phi_{j,l}^{n+1}\right)\right|\right]\\
\le&\mathbb{E}\left[\left\|R_j^{n+1}\right\|_{-1}\left\|\phi_{j,l}^{n+1}\right\|_1\right]\\
\le&\mathbb{E}\left[\left\|R_j^{n+1}\right\|_{-1}C_1\left(\left\|\nabla\phi_{j,l}^{n+1}\right\|^2+\left\|\phi_{j,l}^{n+1}\right\|_{\partial D_1}^2\right)^{\frac{1}{2}}\right]\\
\le&\mathbb{E}\left[\left\|R_j^{n+1}\right\|_{-1}\frac{C_1}{\min\{a_{\min}^{\frac{1}{2}},\alpha_{\min}^{\frac{1}{2}}\}}\left(\bar{a}_l\left\|\nabla\phi_{j,l}^{n+1}\right\|^2+\bar{\alpha}_l\left\|\phi_{j,l}^{n+1}\right\|_{\partial D_1}^2\right)^{\frac{1}{2}}\right]\\
\le&\frac{C_1^2}{2\epsilon_3\min\{a_{\min},\alpha_{\min}\}}\mathbb{E}\left[\left\|R_j^{n+1}\right\|_{-1}^2\right]+\frac{\epsilon_3}{2}\left(\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n+1}\right\|^2\right]+\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{n+1}\right\|_{\partial D_1}^2\right]\right)\\
\le&\frac{C_1^2C_2^2}{2\epsilon_3\min\{a_{\min},\alpha_{\min}\}}\mathbb{E}\left[\left\|R_j^{n+1}\right\|^2\right]+\frac{\epsilon_3}{2}\left(\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n+1}\right\|^2\right]+\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{n+1}\right\|_{\partial D_1}^2\right]\right)\\
\le&\frac{C_1^2C_2^2\Delta t_l^3}{2\epsilon_3\min\{a_{\min},\alpha_{\min}\}}\mathbb{E}\left[\int_{t_{n-1}}^{t_{n+1}}\left\|y_{j,ttt}\right\|^2dt\right]+\frac{\epsilon_3}{2}\left(\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n+1}\right\|^2\right]+\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{n+1}\right\|_{\partial D_1}^2\right]\right),
\end{aligned} \tag{4.24}$$

where C_1 is a constant of two equivalent norms, C_2 is a constant w.r.t. the Sobolev space embedding theorem, and the last inequality of (4.24) holds by the integral form of the Taylor's theorem and the Cauchy-Schwarz inequality. The remaining terms can be examined in the manner below.

$$
\begin{aligned}
&\mathbb{E}\left[\left|\left(a'_j\nabla\left(y_j^{n+1}-2y_j^n+y_j^{n-1}\right),\nabla\phi_{j,l}^{n+1}\right)\right|\right]\\
\leq& a^*\left(\frac{a_{\max}}{2\epsilon_4}\mathbb{E}\left[\left\|\nabla\left(y_j^{n+1}-2y_j^n+y_j^{n-1}\right)\right\|^2\right]+\frac{\epsilon_4}{2}\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n+1}\right\|^2\right]\right)\\
\leq& a^*\left(\frac{a_{\max}}{2\epsilon_4}C_3\Delta t_l^3\mathbb{E}\left[\int_{t_{n-1}}^{t_{n+1}}\left\|\nabla y_{j,tt}\right\|^2dt\right]+\frac{\epsilon_4}{2}\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n+1}\right\|^2\right]\right),
\end{aligned}
\tag{4.25}
$$

$$
\begin{aligned}
&\mathbb{E}\left[\left|\left(\alpha'_j\left(y_j^{n+1}-2y_j^n+y_j^{n-1}\right),\phi_{j,l}^{n+1}\right)_{\partial D_1}\right|\right]\\
\leq& \alpha^*\left(\frac{\alpha_{\max}}{2\epsilon_5}\mathbb{E}\left[\left\|\left(y_j^{n+1}-2y_j^n+y_j^{n-1}\right)\right\|_{\partial D_1}^2\right]+\frac{\epsilon_5}{2}\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{n+1}\right\|_{\partial D_1}^2\right]\right)\\
\leq& \alpha^*\left(\frac{\alpha_{\max}}{2\epsilon_5}C_4\Delta t_l^3\mathbb{E}\left[\int_{t_{n-1}}^{t_{n+1}}\left\|y_{j,tt}\right\|_{\partial D_1}^2dt\right]+\frac{\epsilon_5}{2}\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{n+1}\right\|_{\partial D_1}^2\right]\right),
\end{aligned}
\tag{4.26}
$$

$$
\begin{aligned}
&\mathbb{E}\left[\left|-\left(\alpha'_j\left(2\phi_{j,l}^n-\phi_{j,l}^{n-1}\right),\phi_{j,l}^{n+1}\right)_{\partial D_1}\right|\right]\\
\leq& 2\alpha^*\mathbb{E}\left[\left|\left(\bar{\alpha}_l\phi_{j,l}^n,\phi_{j,l}^{n+1}\right)_{\partial D_1}\right|\right]+\alpha^*\mathbb{E}\left[\left|\left(\bar{\alpha}_l\phi_{j,l}^{n-1},\phi_{j,l}^{n+1}\right)_{\partial D_1}\right|\right]\\
\leq& \alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^n\right\|_{\partial D_1}^2\right]+\frac{\alpha^*}{2}\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{n-1}\right\|_{\partial D_1}^2\right]+\frac{3\alpha^*}{2}\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{n+1}\right\|_{\partial D_1}^2\right],
\end{aligned}
\tag{4.27}
$$

$$
\mathbb{E}\left[\left|\left(\bar{\alpha}_l\rho_{j,l}^{n+1},\phi_{j,l}^{n+1}\right)_{\partial D_1}\right|\right]=0,\quad \mathbb{E}\left[\left|\left(\alpha'_j\left(2\rho_{j,l}^n-\rho_{j,l}^{n-1}\right),\phi_{j,l}^{n+1}\right)_{\partial D_1}\right|\right]=0.
\tag{4.28}
$$

Substituting (4.19) to (4.28) into (4.18), multiplying by 4, and dropping the non-negative term

$$
\frac{1}{4\Delta t_l}\mathbb{E}\left[\left\|\phi_{j,l}^{n+1}-2\phi_{j,l}^n+\phi_{j,l}^{n-1}\right\|^2\right]
$$

on the LHS, we have

$$
\begin{aligned}
&\frac{1}{\Delta t_l}\mathbb{E}\left[\left\|\phi_{j,l}^{n+1}\right\|^2+\left\|2\phi_{j,l}^{n+1}-\phi_{j,l}^n\right\|^2\right]-\frac{1}{\Delta t_l}\mathbb{E}\left[\left\|\phi_{j,l}^n\right\|^2+\left\|2\phi_{j,l}^n-\phi_{j,l}^{n-1}\right\|^2\right]\\
&+4\left(1-3a^*-\frac{\epsilon_2+\epsilon_4}{2}a^*-\frac{\epsilon_1+\epsilon_3}{2}\right)\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n+1}\right\|^2\right]\\
&+6a^*\left(\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n+1}\right\|^2\right]-\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^n\right\|^2\right]\right)+2a^*\left(\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^n\right\|^2\right]-\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{n-1}\right\|^2\right]\right)\\
&+4\left(1-3\alpha^*-\frac{\epsilon_3+\epsilon_5}{2}\right)\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{n+1}\right\|_{\partial D_1}^2\right]+6\alpha^*\left(\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{n+1}\right\|_{\partial D_1}^2\right]-\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^n\right\|_{\partial D_1}^2\right]\right)\\
&+2\alpha^*\left(\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^n\right\|_{\partial D_1}^2\right]-\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{n-1}\right\|_{\partial D_1}^2\right]\right)
\end{aligned}
$$

$$
\begin{aligned}
&\le\frac{2a_{\max}}{\epsilon_1}\mathbb{E}\left[\left\|\nabla\rho_{j,l}^{n+1}\right\|^2\right]+\frac{2a^*a_{\max}}{\epsilon_2}\mathbb{E}\left[\left\|\nabla\left(2\rho_{j,l}^{n}-\rho_{j,l}^{n-1}\right)\right\|^2\right]\\
&+\frac{2C_1^2C_2^2\Delta t_l^3}{\epsilon_3\min\{a_{\min},\alpha_{\min}\}}\mathbb{E}\left[\int_{t_{n-1}}^{t_{n+1}}\left\|y_{j,ttt}\right\|^2dt\right]\\
&+\frac{2a^*a_{\max}}{\epsilon_4}C_3\Delta t_l^3\mathbb{E}\left[\int_{t_{n-1}}^{t_{n+1}}\left\|\nabla y_{j,tt}\right\|^2dt\right]+\frac{2\alpha^*\alpha_{\max}}{\epsilon_5}C_4\Delta t_l^3\mathbb{E}\left[\int_{t_{n-1}}^{t_{n+1}}\left\|y_{j,tt}\right\|^2_{\partial D_1}dt\right].
\end{aligned}
\tag{4.29}
$$

Now, choosing

$$
\epsilon_1=(1-3a^*)/4,\epsilon_2=\epsilon_4=(1-3a^*)/(4a^*),
$$

$$
\epsilon_3=\min\{(1-3a^*)/4,(1-3\alpha^*)/2\},\epsilon_5=(1-3\alpha^*)/2,
$$

multiplying by Δt_l, and summing from $n=1$ up to $n=N_l-1$, we have

$$
\begin{aligned}
&\mathbb{E}\left[\left\|\phi_{j,l}^{N_l}\right\|^2\right]+\mathbb{E}\left[\left\|2\phi_{j,l}^{N_l}-\phi_{j,l}^{N_l-1}\right\|^2\right]+2\Delta t_l(1-3a^*)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{a}_l^{\frac12}\nabla\phi_{j,l}^{n+1}\right\|^2\right]\\
&+2\Delta t_l(1-3\alpha^*)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac12}\phi_{j,l}^{n+1}\right\|^2_{\partial D_1}\right]+6\Delta t_la^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac12}\nabla\phi_{j,l}^{N_l}\right\|^2\right]\\
&+2\Delta t_la^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac12}\nabla\phi_{j,l}^{N_l-1}\right\|^2\right]+6\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac12}\phi_{j,l}^{N_l}\right\|^2_{\partial D_1}\right]+2\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac12}\phi_{j,l}^{N_l-1}\right\|^2_{\partial D_1}\right]\\
&\le\Delta t_l\sum_{n=1}^{N_l-1}\left(\frac{8a_{\max}}{1-3a^*}\mathbb{E}\left[\left\|\nabla\rho_{j,l}^{n+1}\right\|^2\right]+\frac{8(a^*)^2a_{\max}}{1-3a^*}\mathbb{E}\left[\left\|\nabla\left(2\rho_{j,l}^{n}-\rho_{j,l}^{n-1}\right)\right\|^2\right]\right.\\
&+\frac{8C_1^2C_2^2\Delta t_l^3}{\min\{1-3a^*,2(1-3\alpha^*)\}\min\{a_{\min},\alpha_{\min}\}}\mathbb{E}\left[\int_{t_{n-1}}^{t_{n+1}}\left\|y_{j,ttt}\right\|^2dt\right]\\
&\left.+\frac{8(a^*)^2a_{\max}}{1-3a^*}C_3\Delta t_l^3\mathbb{E}\left[\int_{t_{n-1}}^{t_{n+1}}\left\|\nabla y_{j,tt}\right\|^2dt\right]+\frac{4\alpha^*\alpha_{\max}}{1-3\alpha^*}C_4\Delta t_l^3\mathbb{E}\left[\int_{t_{n-1}}^{t_{n+1}}\left\|y_{j,tt}\right\|^2_{\partial D_1}dt\right]\right)\\
&+\mathbb{E}\left[\left\|\phi_{j,l}^{1}\right\|^2\right]+\mathbb{E}\left[\left\|2\phi_{j,l}^{1}-\phi_{j,l}^{0}\right\|^2\right]+6\Delta t_la^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac12}\nabla\phi_{j,l}^{1}\right\|^2\right]\\
&+2\Delta t_la^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac12}\nabla\phi_{j,l}^{0}\right\|^2\right]+6\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac12}\phi_{j,l}^{1}\right\|^2_{\partial D_1}\right]+2\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac12}\phi_{j,l}^{0}\right\|^2_{\partial D_1}\right].
\end{aligned}
\tag{4.30}
$$

By using the initial error assumption that $\left\|e_{j,l}^0\right\|,\left\|e_{j,l}^1\right\|,\left\|\nabla e_{j,l}^0\right\|$, and $\left\|\nabla e_{j,l}^1\right\|$ are at least $O\left(h_l^k\right)$; the regularity assumption; and the standard finite element estimates of L^2 projection error in the H^1 norm (see, e.g., [2]), for any $y_j^n\in H^{k+1}(\bar{D})\cap H^1(\bar{D})$

$$
\left\|\rho_{j,l}^n\right\|^2\le Ch_l^{2k+2}\left\|y_j(t_n)\right\|^2_{k+1}\quad\text{and}\quad\left\|\nabla\rho_{j,l}^n\right\|^2\le Ch_l^{2k}\left\|y_j(t_n)\right\|^2_{k+1},
\tag{4.31}
$$

we have

$$
\mathbb{E}\left[\left\|\phi_{j,l}^{N_l}\right\|^2\right]+\mathbb{E}\left[\left\|2\phi_{j,l}^{N_l}-\phi_{j,l}^{N_l-1}\right\|^2\right]+2\Delta t_l(1-3a^*)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{a}_l^{\frac12}\nabla\phi_{j,l}^{n+1}\right\|^2\right]
$$

$$
\begin{aligned}
&+2\Delta t_l\left(1-3\alpha^*\right)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{n+1}\right\|_{\partial D_1}^2\right]+6\Delta t_l a^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{N_l}\right\|^2\right]\\
&+2\Delta t_l a^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\phi_{j,l}^{N_l-1}\right\|^2\right]+6\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{N_l}\right\|_{\partial D_1}^2\right]+2\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\phi_{j,l}^{N_l-1}\right\|_{\partial D_1}^2\right]\\
\leq&\frac{C\Delta t_l^4}{\min\{1-3a^*,2\left(1-3\alpha^*\right)\}\min\{a_{\min},\alpha_{\min}\}}\mathbb{E}\left[\int_0^T\left\|y_{j,ttt}\right\|^2dt\right]\\
&+\frac{\left(1+\left(a^*\right)^2\right)a_{\max}C}{1-3a^*}h_l^{2m}+\frac{\left(a^*\right)^2a_{\max}C\Delta t_l^4}{1-3a^*}\mathbb{E}\left[\int_0^T\left\|\nabla y_{j,tt}\right\|^2dt\right]\\
&+\frac{\alpha^*\alpha_{\max}C\Delta t_l^4}{1-3\alpha^*}\mathbb{E}\left[\int_0^T\left\|y_{j,tt}\right\|_{\partial D_1}^2dt\right]+h_l^{2k}+\left(a^*+\alpha^*\right)C\Delta t_lh_l^{2k}.
\end{aligned}
\tag{4.32}
$$

By the triangle inequality, we have

$$
\begin{aligned}
&\mathbb{E}\left[\left\|y_j\left(t_{N_l}\right)-y_{j,l}^{N_l}\right\|^2\right]+\mathbb{E}\left[\left\|2\left(y_j\left(t_{N_l}\right)-y_{j,l}^{N_l}\right)-\left(y_j\left(t_{N_l-1}\right)-y_{j,l}^{N_l-1}\right)\right\|^2\right]\\
&+2\Delta t_l\left(1-3a^*\right)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\left(y_j\left(t_{n+1}\right)-y_{j,l}^{n+1}\right)\right\|^2\right]\\
&+2\Delta t_l\left(1-3\alpha^*\right)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\left(y_j\left(t_{n+1}\right)-y_{j,l}^{n+1}\right)\right\|_{\partial D_1}^2\right]\\
&+6\Delta t_la^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\left(y_j\left(t_{N_l}\right)-y_{j,l}^{N_l}\right)\right\|^2\right]+2\Delta t_la^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\left(y_j\left(t_{N_l-1}\right)-y_{j,l}^{N_l-1}\right)\right\|^2\right]\\
&+6\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\left(y_j\left(t_{N_l}\right)-y_{j,l}^{N_l}\right)\right\|_{\partial D_1}^2\right]+2\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\left(y_j\left(t_{N_l-1}\right)-y_{j,l}^{N_l-1}\right)\right\|_{\partial D_1}^2\right]\\
\leq&C\left(\Delta t_l^4+h_l^{2k}\right).
\end{aligned}
\tag{4.33}
$$

This finishes the proof. □

The aggregated contributions to the error from MC sampling and FE approximation culminate in the following estimation for the l-th level MC ensemble approximation.

Theorem 4.4. *Denote $y\left(t_n\right)$ as the solution to Eq* (1.1)*, and let $\Phi_{J_l}^n=\frac{1}{J_l}\sum_{j=1}^{J_l}y_{j,l}^n$. Provided that Conditions (i) and (ii) are satisfied, alongside the stability condition given in* (4.1)*, we have*

$$
\begin{aligned}
&\mathbb{E}\left[\left\|\mathbb{E}\left[y\left(t_{N_l}\right)\right]-\Phi_{J_l}^{N_l}\right\|^2\right]+\mathbb{E}\left[\left\|2\left(\mathbb{E}\left[y\left(t_{N_l}\right)\right]-\Phi_{J_l}^{N_l}\right)-\left(\mathbb{E}\left[y\left(t_{N_l-1}\right)\right]-\Phi_{J_l}^{N_l-1}\right)\right\|^2\right]\\
&+2\Delta t_l\left(1-3a^*\right)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\left(\mathbb{E}\left[y\left(t_{n+1}\right)\right]-\Phi_{J_l}^{n+1}\right)\right\|^2\right]\\
&+2\Delta t_l\left(1-3\alpha^*\right)\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\left(\mathbb{E}\left[y\left(t_{n+1}\right)\right]-\Phi_{J_l}^{n+1}\right)\right\|_{\partial D_1}^2\right]\\
&+6\Delta t_la^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\left(\mathbb{E}\left[y\left(t_{N_l}\right)\right]-\Phi_{J_l}^{N_l}\right)\right\|^2\right]+2\Delta t_la^*\mathbb{E}\left[\left\|\bar{a}_l^{\frac{1}{2}}\nabla\left(\mathbb{E}\left[y\left(t_{N_l-1}\right)\right]-\Phi_{J_l}^{N_l-1}\right)\right\|^2\right]
\end{aligned}
$$

$$+ 6\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\left(\mathbb{E}\left[y\left(t_{N_l}\right)\right]-\Phi_{J_l}^{N_l}\right)\right\|_{\partial D_1}^2\right]+2\Delta t_l\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\left(\mathbb{E}\left[y\left(t_{N_l-1}\right)\right]-\Phi_{J_l}^{N_l-1}\right)\right\|_{\partial D_1}^2\right]$$
$$\leq\frac{C}{J_l}\left(\Delta t_l\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1}^2\right]+\Delta t_l\sum_{n=1}^{N_l-1}\mathbb{E}\left[\left\|u_j^{n+1}\right\|_{\partial D_1}^2\right]+\mathbb{E}\left[\left\|y_{j,l}^1\right\|^2\right]+\mathbb{E}\left[\left\|2y_{j,l}^1-y_{j,l}^0\right\|^2\right]\right.$$
$$\left.+\Delta t_l\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\nabla y_{j,l}^1\right\|^2\right]+\Delta t_l\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}\nabla y_{j,l}^0\right\|^2\right]+\Delta t_l\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}y_{j,l}^1\right\|_{\partial D_1}^2\right]+\Delta t_l\mathbb{E}\left[\left\|\bar{\alpha}_l^{\frac{1}{2}}y_{j,l}^0\right\|_{\partial D_1}^2\right]\right)$$
$$+ C\left(\Delta t_l^4+h_l^{2k}\right). \tag{4.34}$$

Proof. We consider the first term on the LHS of (4.34). By the triangle and Young's inequalities, we get

$$\mathbb{E}\left[\left\|\mathbb{E}\left[y\left(t_{N_l}\right)\right]-\Phi_{J_l}^{N_l}\right\|^2\right]\leq 2\left(\mathbb{E}\left[\left\|\mathbb{E}\left[y_j\left(t_{N_l}\right)\right]-\mathbb{E}\left[y_{j,l}^{N_l}\right]\right\|^2\right]+\mathbb{E}\left[\left\|\mathbb{E}\left[y_{j,l}^{N_l}\right]-\Phi_{J_l}^{N_l}\right\|^2\right]\right).$$

The conclusion is derived from Theorems 4.2 and 4.3. The remaining terms on the LHS of (4.34) can be evaluated using a similar approach, which finishes the proof. □

4.2. Multilevel MC ensemble FE method

Next, we will formulate the error estimation for the MLMCE method.

Theorem 4.5. *Given the assumption of the validity of Conditions (i) and (ii), and that the stability condition* (4.1) *is satisfied, then the error of the MLMCE approximation can be described as follows:*

$$\mathbb{E}\left[\left\|\mathbb{E}\left[y\left(t_{N_L}\right)\right]-\Phi\left[y_L\left(t_{N_L}\right)\right]\right\|^2\right]+\mathbb{E}\left[\|\mathbb{E}\left[y^{N_L}\right]-\Phi\left[y_L\left(t_{N_L}\right)\right]-\left(\mathbb{E}\left[y^{N_L-1}\right]\right.\right.$$
$$\left.\left.-\Phi\left[y_L\left(t_{N_L-1}\right)\right]\right)\|^2\right]+2\Delta t_L\left(1-3a^*\right)\sum_{n=1}^{N_L-1}\mathbb{E}\left[\left\|\bar{a}_L^{\frac{1}{2}}\left(\nabla\mathbb{E}\left[y\left(t_{n+1}\right)\right]-\nabla\Phi\left[y_L\left(t_{n+1}\right)\right]\right)\right\|^2\right]$$
$$+ 2\Delta t_L\left(1-3\alpha^*\right)\sum_{n=1}^{N_L-1}\mathbb{E}\left[\left\|\bar{\alpha}_L^{\frac{1}{2}}\left(\mathbb{E}\left[y\left(t_{n+1}\right)\right]-\Phi\left[y_L\left(t_{n+1}\right)\right]\right)\right\|_{\partial D_1}^2\right]$$
$$+ 6\Delta t_La^*\mathbb{E}\left[\left\|\bar{a}_L^{\frac{1}{2}}\nabla\left(\mathbb{E}\left[y\left(t_{N_L}\right)\right]-\Phi\left[y_L\left(t_{N_L}\right)\right]\right)\right\|^2\right]$$
$$+ 2\Delta t_La^*\mathbb{E}\left[\left\|\bar{a}_L^{\frac{1}{2}}\nabla\left(\mathbb{E}\left[y\left(t_{N_L-1}\right)\right]-\Phi\left[y_L\left(t_{N_L-1}\right)\right]\right)\right\|^2\right]$$
$$+ 6\Delta t_L\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_L^{\frac{1}{2}}\left(\mathbb{E}\left[y\left(t_{N_L}\right)\right]-\Phi\left[y_L\left(t_{N_L}\right)\right]\right)\right\|_{\partial D_1}^2\right]$$
$$+ 2\Delta t_L\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_L^{\frac{1}{2}}\left(\mathbb{E}\left[y\left(t_{N_L-1}\right)\right]-\Phi\left[y_L\left(t_{N_L-1}\right)\right]\right)\right\|_{\partial D_1}^2\right]$$
$$\leq C\left(h_L^{2k}+\Delta t_L^4+\sum_{l=1}^{L}\frac{1}{J_l}\left(h_l^{2k}+\Delta t_l^4\right)\right)$$
$$+\frac{C}{J_0}\left(\Delta t_0\sum_{n=1}^{N_0-1}\mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1}^2\right]+\Delta t_0\sum_{n=1}^{N_0-1}\mathbb{E}\left[\left\|u_j^{n+1}\right\|_{\partial D_1}^2\right]+\mathbb{E}\left[\left\|y_{j,0}^1\right\|^2\right]+\mathbb{E}\left[\left\|2y_{j,0}^1-y_{j,0}^0\right\|^2\right]\right.$$

$$+\Delta t_0\mathbb{E}\left[\left\|\bar{a}_0^{\frac{1}{2}}\nabla y_{j,0}^1\right\|^2\right]+\Delta t_0\mathbb{E}\left[\left\|\bar{a}_0^{\frac{1}{2}}\nabla y_{j,0}^0\right\|^2\right]+\Delta t_0\mathbb{E}\left[\left\|\bar{\alpha}_0^{\frac{1}{2}}y_{j,0}^1\right\|_{\partial D_1}^2\right]+\Delta t_0\mathbb{E}\left[\left\|\bar{\alpha}_0^{\frac{1}{2}}y_{j,0}^0\right\|_{\partial D_1}^2\right]\Bigg). \tag{4.35}$$

Proof. Let $y_{-1}(t) = 0$. Our analysis solely focuses on the initial term on the LHS, as the remaining terms can be addressed in a similar way.

$$\begin{aligned}
&\mathbb{E}\left[\left\|\mathbb{E}\left[y\left(t_{N_L}\right)\right]-\Phi\left[y_L\left(t_{N_L}\right)\right]\right\|^2\right]\\
=&\mathbb{E}\left[\left\|\mathbb{E}\left[y\left(t_{N_L}\right)\right]-\mathbb{E}\left[y_L\left(t_{N_L}\right)\right]+\mathbb{E}\left[y_L\left(t_{N_L}\right)\right]-\sum_{l=0}^{L}\Phi_{J_l}\left[y_l\left(t_{N_L}\right)-y_{l-1}\left(t_{N_L}\right)\right]\right\|^2\right]\\
\leq&C\Bigg(\mathbb{E}\left[\left\|\mathbb{E}\left[y\left(t_{N_L}\right)\right]-\mathbb{E}\left[y_L\left(t_{N_L}\right)\right]\right\|^2\right]+\sum_{l=0}^{L}\mathbb{E}\left[\|\left(\mathbb{E}\left[y_l\left(t_{N_L}\right)-y_{l-1}\left(t_{N_L}\right)\right]\right.\right.\\
&\left.\left.-\Phi_{J_l}\left[y_l\left(t_{N_L}\right)-y_{l-1}\left(t_{N_L}\right)\right]\right)\|^2\right]\Bigg).
\end{aligned} \tag{4.36}$$

By Jensen's inequality and Theorem 4.3, we get

$$\mathbb{E}\left[\left\|\mathbb{E}\left[y\left(t_{N_L}\right)\right]-\mathbb{E}\left[u_L\left(t_{N_L}\right)\right]\right\|^2\right]\leq\mathbb{E}\left[\left\|y\left(t_{N_L}\right)-y_L\left(t_{N_L}\right)\right\|^2\right]\leq C\left(\Delta t_L^4+h_L^{2k}\right). \tag{4.37}$$

By Theorems 4.2 and 4.3, and applying the triangle inequality, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\left\|\mathbb{E}\left[y_l\left(t_{N_L}\right)-y_{l-1}\left(t_{N_L}\right)\right]-\Phi_{J_l}\left[y_l\left(t_{N_L}\right)-y_{l-1}\left(t_{N_L}\right)\right]\right\|^2\right]\\
=&\mathbb{E}\left[\left\|\left(\mathbb{E}-\Phi_{J_l}\right)\left[y_l\left(t_{N_L}\right)-y_{l-1}\left(t_{N_L}\right)\right]\right\|^2\right]\\
\leq&\frac{1}{J_l}\mathbb{E}\left[\left\|y_l\left(t_{N_L}\right)-y_{l-1}\left(t_{N_L}\right)\right\|^2\right]\\
\leq&\frac{2}{J_l}\left(\mathbb{E}\left[\left\|y\left(t_{N_L}\right)-y_l\left(t_{N_L}\right)\right\|^2\right]+\mathbb{E}\left[\left\|y\left(t_{N_L}\right)-y_{l-1}\left(t_{N_L}\right)\right\|^2\right]\right)\\
\leq&\frac{C}{J_l}\left(\Delta t_l^4+h_l^{2k}+\Delta t_{l-1}^4+h_{l-1}^{2k}\right)\\
\leq&\frac{C}{J_l}\left(\Delta t_l^4+h_l^{2k}\right).
\end{aligned} \tag{4.38}$$

Meanwhile, by Theorem 4.4, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\left\|\mathbb{E}\left[y_0\left(t_{N_L}\right)\right]-\Phi_{J_0}\left[y_0\left(t_{N_L}\right)\right]\right\|^2\right]\\
\leq&\frac{C}{J_0}\Bigg(\Delta t_0\sum_{n=1}^{N_0-1}\mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1}^2\right]+\Delta t_0\sum_{n=1}^{N_0-1}\mathbb{E}\left[\left\|u_j^{n+1}\right\|_{\partial D_1}^2\right]+\mathbb{E}\left[\left\|y_{j,0}^1\right\|^2\right]+\mathbb{E}\left[\left\|2y_{j,0}^1-y_{j,0}^0\right\|^2\right]\\
&+\Delta t_0\mathbb{E}\left[\left\|\bar{a}_0^{\frac{1}{2}}\nabla y_{j,0}^1\right\|^2\right]+\Delta t_0\mathbb{E}\left[\left\|\bar{a}_0^{\frac{1}{2}}\nabla y_{j,0}^0\right\|^2\right]+\Delta t_0\mathbb{E}\left[\left\|\bar{\alpha}_0^{\frac{1}{2}}y_{j,0}^1\right\|_{\partial D_1}^2\right]+\Delta t_0\mathbb{E}\left[\left\|\bar{\alpha}_0^{\frac{1}{2}}y_{j,0}^0\right\|_{\partial D_1}^2\right]\Bigg).
\end{aligned} \tag{4.39}$$

Plugging (4.37)–(4.39) into (4.36), we have

$$\mathbb{E}\left[\left\|\mathbb{E}\left[y\left(t_{N_L}\right)\right]-\Phi\left[y_L\left(t_{N_L}\right)\right]\right\|^2\right]$$

$$
\begin{aligned}
&\leq C\left(\Delta t_L^4 + h_L^{2k} + \sum_{l=1}^{L}\frac{1}{J_l}\left(\Delta t_l^4 + h_l^{2k}\right)\right)\\
&+\frac{C}{J_0}\left(\Delta t_0\sum_{n=1}^{N_0-1}\mathbb{E}\left[\left\|f_j^{n+1}\right\|_{-1}^2\right] + \Delta t_0\sum_{n=1}^{N_0-1}\mathbb{E}\left[\left\|u_j^{n+1}\right\|_{\partial D_1}^2\right] + \mathbb{E}\left[\left\|y_{j,0}^1\right\|^2\right] + \mathbb{E}\left[\left\|2y_{j,0}^1 - y_{j,0}^0\right\|^2\right]\right.\\
&\left.+\ \Delta t_0\mathbb{E}\left[\left\|\bar{a}_0^{\frac{1}{2}}\nabla y_{j,0}^1\right\|^2\right] + \Delta t_0\mathbb{E}\left[\left\|\bar{a}_0^{\frac{1}{2}}\nabla y_{j,0}^0\right\|^2\right] + \Delta t_0\mathbb{E}\left[\left\|\bar{\alpha}_0^{\frac{1}{2}} y_{j,0}^1\right\|_{\partial D_1}^2\right] + \Delta t_0\mathbb{E}\left[\left\|\bar{\alpha}_0^{\frac{1}{2}} y_{j,0}^0\right\|_{\partial D_1}^2\right]\right). \qquad (4.40)
\end{aligned}
$$

The other terms on the LHS in (4.35) can be treated in the same way. This finishes the proof. □

In general, as the mesh is refined, the cost associated with FE simulations tends to rise. To achieve an optimal convergence rate, we can adjust the time step size Δt_l, mesh size h_l, and sampling size J_l in accordance with the preceding error estimation. The subsequent corollary can be derived directly from Theorem 4.5.

Corollary 4.1. *By choosing*

$$
\Delta t_l = O\left(h_l^{k/2}\right) \quad \text{and} \quad J_l = O\left(l^{1+\epsilon}2^{2k(L-l)}\right)
$$

for $\forall\ \epsilon > 0$ *and* $l = 0, 1, \cdots, L$*, the MLMCE approximation is such that*

$$
\begin{aligned}
&\mathbb{E}\left[\left\|\mathbb{E}\left[y\left(t_{N_L}\right)\right] - \Phi\left[y_L\left(t_{N_L}\right)\right]\right\|^2\right] + \mathbb{E}\left[\|\mathbb{E}\left[y^{N_L}\right] - \Phi\left[y_L\left(t_{N_L}\right)\right] - \left(\mathbb{E}\left[y^{N_L-1}\right]\right.\right.\\
&\left.\left.-\Phi\left[y_L\left(t_{N_L-1}\right)\right]\right)\|^2\right] + 2\Delta t_L\left(1-3a^*\right)\sum_{n=1}^{N_L-1}\mathbb{E}\left[\left\|\bar{a}_L^{\frac{1}{2}}\left(\nabla\mathbb{E}\left[y\left(t_{n+1}\right)\right] - \nabla\Phi\left[y_L\left(t_{n+1}\right)\right]\right)\right\|^2\right]\\
&+ 2\Delta t_L\left(1-3\alpha^*\right)\sum_{n=1}^{N_L-1}\mathbb{E}\left[\left\|\bar{\alpha}_L^{\frac{1}{2}}\left(\mathbb{E}\left[y\left(t_{n+1}\right)\right] - \Phi\left[y_L\left(t_{n+1}\right)\right]\right)\right\|_{\partial D_1}^2\right]\\
&+ 6\Delta t_L a^*\mathbb{E}\left[\left\|\bar{a}_L^{\frac{1}{2}}\nabla\left(\mathbb{E}\left[y\left(t_{N_L}\right)\right] - \Phi\left[y_L\left(t_{N_L}\right)\right]\right)\right\|^2\right]\\
&+ 2\Delta t_L a^*\mathbb{E}\left[\left\|\bar{a}_L^{\frac{1}{2}}\nabla\left(\mathbb{E}\left[y\left(t_{N_L-1}\right)\right] - \Phi\left[y_L\left(t_{N_L-1}\right)\right]\right)\right\|^2\right]\\
&+ 6\Delta t_L\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_L^{\frac{1}{2}}\left(\mathbb{E}\left[y\left(t_{N_L}\right)\right] - \Phi\left[y_L\left(t_{N_L}\right)\right]\right)\right\|_{\partial D_1}^2\right]\\
&+ 2\Delta t_L\alpha^*\mathbb{E}\left[\left\|\bar{\alpha}_L^{\frac{1}{2}}\left(\mathbb{E}\left[y\left(t_{N_L-1}\right)\right] - \Phi\left[y_L\left(t_{N_L-1}\right)\right]\right)\right\|_{\partial D_1}^2\right]\\
&\leq Ch_L^{2k}. \qquad (4.41)
\end{aligned}
$$

4.3. The cost of the MLMCE FE method

Similar to the MLMC method [5, 10, 33], the sample size in MLMCE can be determined by minimizing the total computational cost, all while maintaining a predetermined level of error. By setting $\Delta t_l = O\left(h_l^{k/2}\right)$, one can align the errors in both the spatial and temporal domains. It is posited that as the mesh size decreases, the average cost associated with solving the PDE at level l will increase. Conversely, the average variance will decrease, following these relationships accordingly:

$$
C_l = C_\gamma h_l^{-\gamma_1} \text{ and } \sigma_l = C_\mu h_l^{\mu}, \qquad (4.42)
$$

where $C_\gamma, C_\mu, \gamma_1$, and μ are positive constants. To determine the optimal number of samples at the l-th level, denoted as J_l, it is imperative to minimize the overall sampling cost, ensuring that the statistical error does not exceed a predefined tolerance ϵ. This objective can be effectively formulated as an unconstrained optimization problem using the Lagrangian methodology

$$\min_{J_l} \sum_{l=0}^{L} J_l C_l + \lambda \left[(L+1) \sum_{l=0}^{L} \frac{\sigma_l}{J_l} - \frac{\epsilon^2}{4} \right].$$

Upon utilizing the Euler-Lagrange condition, we derive the optimal number of samples,

$$J_l = \frac{4(L+1)}{\epsilon^2} \left(\sum_{l=0}^{L} \sqrt{\sigma_l C_l} \right) \sqrt{\frac{\sigma_l}{C_l}}$$

and the corresponding total cost is denoted as C_{total}

$$C_{total} = \sum_{l=0}^{L} J_l C_l = \frac{4(L+1)}{\epsilon^2} \left(\sum_{l=0}^{L} \sqrt{\sigma_l C_l} \right)^2.$$

It is noteworthy that within this framework, MLMCE utilizes the same expressions for J_l and C_{total} as those in MLMC. However, utilizing the scheme (3.3) in MLMCE results in a lower average cost for solving the PDE compared with MLMC. Let the average cost of MLMC at level l be denoted as $Ch_l^{-\gamma_2}$; it follows that $\gamma_1 < \gamma_2$ when either direct or block iterative methods are employed in the linear solver. Take C_{MLMCE} and C_{MLMC} as the total costs of the MLMCE and MLMC methods, respectively, we have

$$\frac{C_{MLMCE}}{C_{MLMC}} = \left(\frac{\sum_{l=0}^{L} \sqrt{\sigma_l h_l^{-\gamma_1}}}{\sum_{l=0}^{L} \sqrt{\sigma_l h_l^{-\gamma_2}}} \right)^2 = \left(\frac{\sum_{l=0}^{L} \sqrt{h_l^{\mu-\gamma_1}}}{\sum_{l=0}^{L} \sqrt{h_l^{\mu-\gamma_2}}} \right)^2.$$

Then

$$\frac{C_{MLMCE}}{C_{MLMC}} = \begin{cases} h_0^{\mu-\gamma_1}/h_0^{\mu-\gamma_2} = h_0^{\gamma_2-\gamma_1} & \text{if } \gamma_2 < \mu, \\ h_0^{\mu-\gamma_1}/h_L^{\mu-\gamma_2} = 2^{L(\mu-\gamma_2)} h_0^{\gamma_2-\gamma_1} & \text{if } \gamma_1 < \mu < \gamma_2, \\ h_L^{\mu-\gamma_1}/h_L^{\mu-\gamma_2} = h_L^{\gamma_2-\gamma_1} & \text{if } \mu < \gamma_1. \end{cases}$$

The overall computational complexity of MLMCE is found to be less than that of standard MLMC in all scenarios.

If the standard LU factorization is used in the linear solver, then a sharper computational complexity can be established. Denote d as the dimension of the domain. The complexity of LU factorization is Ch^{-3d}, and the complexity for solving triangular systems is Ch^{-2d} (see, e.g., [11]). Therefore, the total computational cost for sampling is $\sum_{l=0}^{L} \left(J_l h_l^{-2d} + h_l^{-3d} \right)$ since only one LU factorization per level is needed. By minimizing the total cost while maintaining an error margin of ϵ, the optimal sample size J_l is obtained as follows:

$$J_l = \frac{4(L+1)}{\epsilon^2} \left(\sum_{l=0}^{L} \sqrt{\sigma_l h_l^{-2d}} \right) \sqrt{\sigma_l h_l^{2d}}. \tag{4.43}$$

The corresponding computational complexity is denoted as C_{MLMCE}

$$C_{MLMCE} = \frac{4(L+1)}{\epsilon^2}\left(\sum_{l=0}^{L}\sqrt{\sigma_l h_l^{-2d}}\right)^2 + \sum_{l=0}^{L} h_l^{-3d}. \tag{4.44}$$

The optimized MLMC complexity is

$$C_{MLMC} = \frac{4(L+1)}{\epsilon^2}\left(\sum_{l=0}^{L}\sqrt{\sigma_l\left(h_l^{-2d}+h_l^{-3d}\right)}\right)^2. \tag{4.45}$$

5. Numerical experiments

This section present an example to prove the validity and efficiency of the proposed MLMCE algorithm. All tests are implemented by using Matlab code and run on a Hewlett-Packard laptop PC with an Intel Core i5-6200U CPU and 12 GB memory.

We consider the model (1.1) by choosing $D = [-1/4, 1/4] \times [0,1]$ and $T = 1$. Let $\partial D_1 = \Gamma_l \cup \Gamma_r, \partial D_0 = \Gamma_u \cup \Gamma_d$, where Γ_l and Γ_r are the left and right edges of D, respectively; Γ_u and Γ_d are the upside and downside edges. Inspired by numerical experiments in [32] , we construct an exact solution for the model is $y = (1+\omega)(T-t)\cos(\pi x_1)\cos(\pi x_2)$, $\forall\, \mathbf{x} = (x_1, x_2) \in D, t \in [0,T]$, and $\omega \sim U(0,1)$.

We take $a(\mathbf{x}, \omega_a) = \frac{1+\omega_a}{2}\cos(\pi x_1 x_2)$, $\alpha(\mathbf{x}, \omega_\alpha) = \frac{1+\omega_\alpha}{2}\cos(\pi x_1 x_2)$, where $\omega, \omega_a, \omega_\alpha \sim U(0,1)$ are i.i.d. While the initial condition y^0, the boundary condition u, and the source term f are chosen to match the prescribed exact solution. We can see that the expectation of the solution, $\mathbb{E}[y(T)]$, is zero. We adopt the following a non-ensemble one-step stable second-order time-stepping method (NE/CN) in which we have omitted the subscript l to calculate y_j^1. NE/CN [36]: Find y_j^{n+1}, such that

$$\left(\frac{y_j^{n+1}-y_j^n}{\Delta t}, v\right) + \left(a_j \nabla y_j^{n+1/2}, \nabla v\right) + \left(\alpha_j y_j^{n+1/2}, v\right)_{\partial D_1} = \left(f_j^{n+1/2}, v\right) + \left(\alpha_j u_j^{n+1/2}, v\right)_{\partial D_1},$$

where $y_j^{n+1/2} = \frac{y_j^{n+1}+y_j^n}{2}$, $f_j^{n+1/2} = f_j(n+1/2, \mathbf{x})$, and $u_j^{n+1/2} = u_j(n+1/2, \mathbf{x})$. We consider the quantity of interest, $P(y) = \|\mathbb{E}[y(T)]\|$. One can analyze the expectation of the solution at the final time from the ensemble simulations and extract the underlying stochastic information of the system. For the level $l \in \{0, 1, \cdots, L\}$, let $P_l = \|\Phi[y_l(t_{N_l})]\|$. In terms of spatial discretization, we employ linear finite elements on uniform triangulations, i.e., $k = 1$. Some tests are as follows.

(1) Basic hypothesis condition tests of MLMC and MLMCE: In this testing session, we set $L = 5$, $N_0 = 10$ (initial samples on the coarsest levels), and $N = 1000$ (samples for convergence tests). From Figures 1 and 2, both the multilevel correction average variance $\mathbb{E}[Var[P_l - P_{l-1}]]$ and the mean $\mathbb{E}[P_l - P_{l-1}]$ (so $\mathbb{E}[P_L - P]$) decrease exponentially with level, while Figure 3 displays the average costs increasing exponentially. These results align with the basic hypothesis conditions (4.42), and also conform to the geometric MLMC theorem (see, e.g., Theorem 1 in [10]). In addition, the MLMCE cost shows below the MLMC cost with the level from Figure 3.

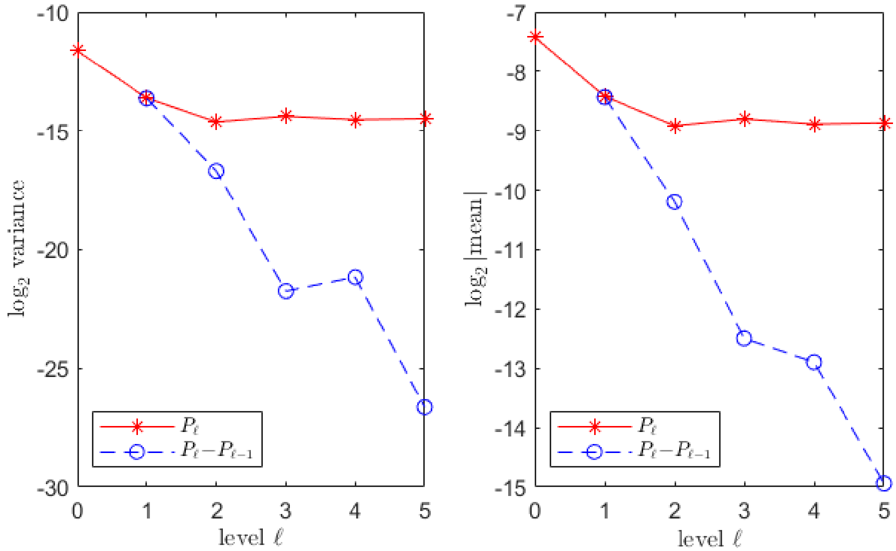

Figure 1. The variance and mean of the quantity of interest for MLMC.

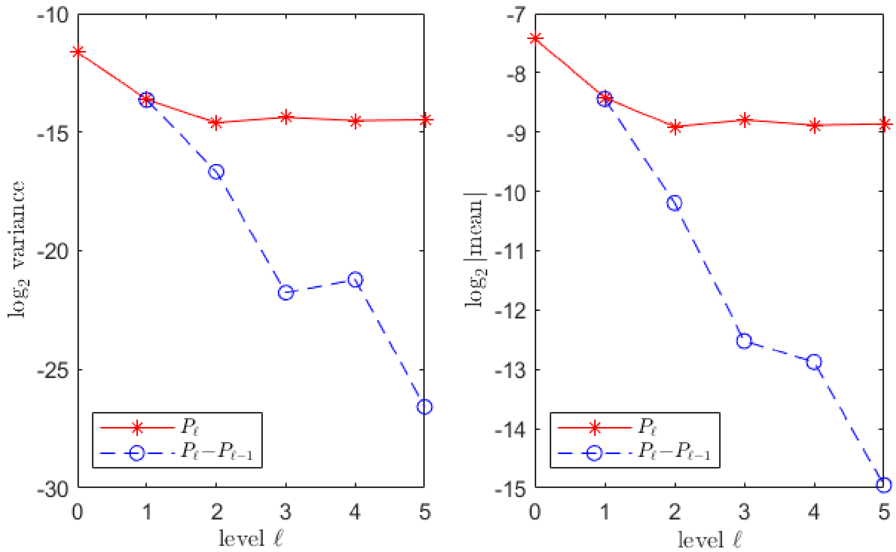

Figure 2. The variance and mean of the quantity of interest for MLMCE.

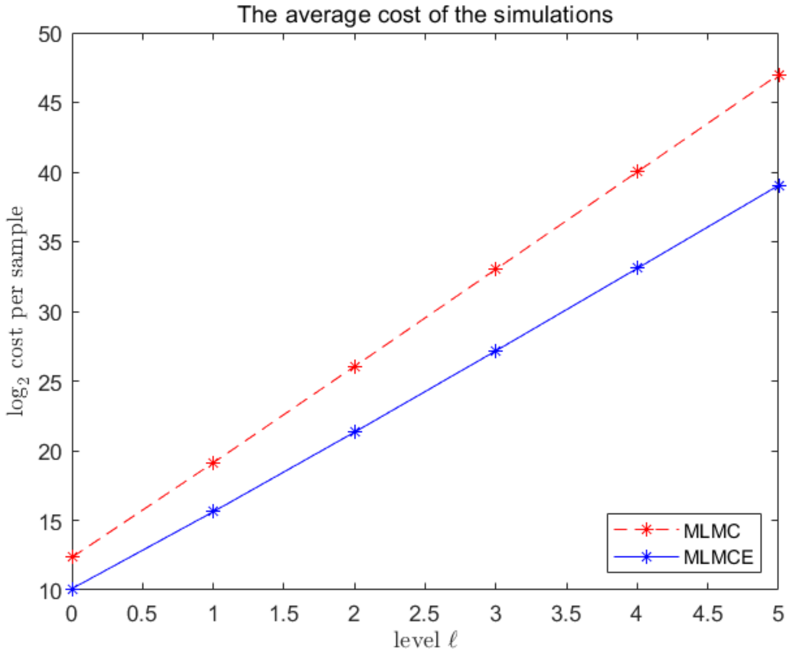

Figure 3. The average cost of the simulations.

(2) Convergence rate of the simulations: To validate the analysis presented in (4.41), we keep L constant and select $h_l = \sqrt{2} \cdot 2^{-2-l}$, $\Delta t_l = 2^{-3-l}$, and $J_l = 2^{2(L-l)+1}$ at the l-th level of simulations for $l = 0, \cdots, L$. The test is conducted $R = 12$ times. Take the numerical error in L^2 norm

$$\mathcal{E}_{L^2(D)} = \sqrt{\frac{1}{R}\sum_{r=1}^{R} \|\mathbb{E}\left[y(T)\right] - \Phi\left[y_L^{(r)}(t_{N_L})\right]\|^2},$$

where y is the exact solution and $y_L^{(r)}$ is the numerical solution of the r-th replica. The numerical errors as L varies from 1 to 3 are listed in Table 1. It is observed that both the MLMC and MLMCE numerical errors converge at the order of nearly 1 w.r.t. h_L, the convergence order aligns with the theoretical findings of Corollary 4.1.

Table 1. Numerical errors and convergence rates of the simulations.

L	$\mathcal{E}_{L^2(D)}^{MLMC}$	Rate	$\mathcal{E}_{L^2(D)}^{MLMCE}$	Rate
1	7.59e-02		5.71e-02	
2	1.89e-02	1.00	1.42e-02	1.00
3	2.76e-03	1.38	2.70e-03	1.19

(3) The numerical results of MLMC and MLMCE: We utilize the MLMCE method with the maximum level $L = 2$, $h_l = \sqrt{2} \cdot 2^{-2-l}$, and $\Delta t_l = 2^{-3-l}$. Due to the limited scope of the problem, we employ LU factorization to solve the linear systems. Given a numerical error of $\epsilon = 10^{-3}$, we select the number of samples $J_l = 2^{2(L-l)+1}$ at the l-th level of simulations, for $l = 0, \cdots, L$, according to (4.43) with $d = 2$ and $\mu = 4$. It is important to note that if the sample set does not comply with the stability condition (6), we will partition the sample set into smaller subsets to ensure that the

condition (6) is satisfied within each subset. The MLMCE solution at the final time T is $\Phi^E(\mathbf{x}) = \Phi\left[y_L^E(t_{N_L})\right]$, as illustrated in Figure 4 (left). We assess the outcomes against those yielded by the standard MLMC FE simulations, maintaining an individual computational framework. Let us denote the approximated expected value obtained from the latter method as $\Phi^I(\mathbf{x}) = \Phi\left[y_L^I(t_{N_L})\right]$, represented in Figure 4 (middle). For a just comparison, it is important to mention that we also implement LU factorization for solving the linear systems across all individual simulations. The observed difference, represented as $\Phi^E - \Phi^I$, is illustrated in Figure 4 (right). The analysis indicates that this difference is on the order of 10^{-3}, demonstrating that the MLMCE method achieves a similar level of accuracy to the individual simulations.

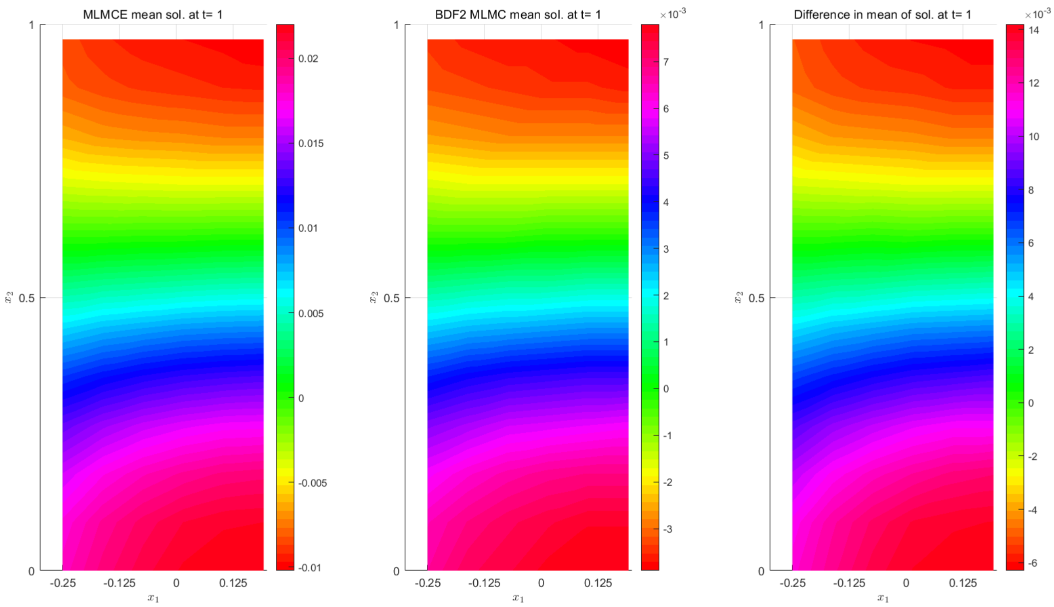

Figure 4. The comparison of simulation means is depicted as follows: Simulations using MLMCE (left) and MLMC simulations (middle), and the subsequent difference (right).

(4) The total costs of MLMC and MLMCE: Following the previous section, we find that the computational complexity linked to the MLMCE simulation is notably lower than that of the separate MLMC simulations. In the theory, by (49) and (50), we have

$$C_{MLMC} = \frac{4(L+1)}{\epsilon^2}\left(\sum_{l=0}^{L}\sqrt{1+h_l^{-2}}\right)^2 \approx 4.49\times 10^9,$$

$$C_{MLMCE} = \frac{4(L+1)^3}{\epsilon^2} + \sum_{l=0}^{L} h_l^{-6} \approx 1.10\times 10^8.$$

In the simulations, the total costs are calculated by

$$C_{MLMC}^{sim} = \sum_{l=0}^{L} J_l \cdot N_l\left(h_l^{-4}+h_l^{-6}\right) \approx 6.75\times 10^9,$$

$$C^{sim}_{MLMCE} = \sum_{l=0}^{L} J_l \cdot N_l h_l^{-4} + \sum_{l=0}^{L} N_l h_l^{-6} \approx 1.66 \times 10^8.$$

Thus

$$C_{MLMC}/C_{MLMCE} \approx 40.69, \quad C^{sim}_{MLMC}/C^{sim}_{MLMCE} \approx 40.66.$$

Our complexity estimations almost match the theoretical calculations, and the MLMCE method is more cost-effective than the MLMC method.

From Test (1), it is evident that the feasibility of the MLMC and MLMCE methods is demonstrated. The accuracy of the MLMCE numerical results aligns well with the theoretical findings in Test (2). In Tests (3) and (4), it is observed that not only is the numerical accuracy of MLMCE nearly identical to that of MLMC, but the MLMCE method also significantly reduces the computational costs.

6. Conclusions

In this study, the MLMCE algorithm has been employed to enhance the computational efficiency for numerical solutions of a transient heat equation characterized by random Robin boundary conditions and diffusion coefficients. The coefficients of the linear systems are computed and stored in a potentially resource-intensive offline phase, allowing for an extremely fast and J-independent assembly of the linear equations in the online phase. Thereby, the MLMCE method is implemented to greatly reduce the computational costs. Stability analysis and error estimates of the algorithms are derived under some conditions involving two ratios between the fluctuations of the thermal conductivity and random Robin boundary coefficient corresponding their mean. Numerical tests are presented to confirm the theoretical results and verify the efficiency of the proposed algorithm.

However, one disadvantage is that the error estimate is not optimal concerning the spatial considerations in this article. This also outlines the direction for our next steps in improvement, and our team's new proof method in [37] could be applied to drive these enhancements. Meanwhile, the MLMCE algorithm should be applied to nonlinear parabolic equations. Additionally, stochastic Robin boundary control problems, which are worth investigating, can be simulated using the MLMCE algorithm.

Author contributions

Tingfu Yao: Methodology, writing-original draft, and software; Changlun Ye: Software and checking; Xianbing Luo: Review, supervision, and validation.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant Nos. 12361033, 12461076), the Natural Science Research Foundation of Education Department of

Guizhou Province (Grant No. QJJ[2024]190), and the Guiyang Univeristy Multidisciplinary Team Construction Projects in 2025 [Gyxk202503].

Conflict of interest

The authors declare there is no conflict of interest.

References

1. I. Babuška, F. Nobile, R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, *SIAM J. Numer. Anal.*, **52** (2010), 317–355. https://doi.org/10.1137/10078635
2. S. C. Brenner, L. R. Scott, *The mathematical theory of finite element methods*, New York: Springer, 2008. https://doi.org/10.1007/978-0-387-75934-0
3. R. Chiba, *Stochastic analysis of heat conduction and thermal stresses in solids: a review*, IntechOpen, 2012. https://doi.org/10.5772/50994
4. R. Chiba, Stochastic heat conduction analysis of a functionally graded annular disc with spatially random heat transfer coefficients, *Appl. Math. Model.*, **33** (2009), 507–523. https://doi.org/10.1016/j.apm.2007.11.014
5. K. A. Cliffe, M. B. Giles, R. Scheichl, A. L. Teckentrup, Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients, *Comput. Visual Sci.*, **14** (2011), 3. https://doi.org/10.1007/s00791-011-0160-x
6. L. C. Evans, *Partial differential equations*, American Mathematical Society, 2010.
7. X. Feng, Y. Luo, L. Vo, Z. Wang, An efficient iterative method for solving parameter-dependent and random convection-diffusion problems, *J. Sci. Comput.*, **90** (2022), 72. https://doi.org/10.1007/s10915-021-01737-z
8. G. S. Fishman, *Monte Carlo: Concepts, algorithms, and applications*, New York: Springer, 1996. https://doi.org/10.1007/978-1-4757-2553-7
9. B. Ganapathysubramanian, N. Zabaras, Sparse grid collocation schemes for stochastic natural convection problems, *J. Comput. Phys.*, **225** (2007), 652–685. https://doi.org/10.1016/j.jcp.2006.12.014
10. M. B. Giles, Multilevel Monte Carlo methods, *Acta Numer.*, **24** (2015), 259–328. https://doi.org/10.1017/S096249291500001X
11. G. H. Golup, C. F. Van Loan. *Matrix computation*, Baltimore: The Johns Hopkins University Press, 2013.
12. M. Gunzburger, C. Webster, G. Zhang, Stochastic finite element methods for partial differential equations with random input data, *Acta Numer.*, **23** (2014), 521–650. https://doi.org/10.1017/S0962492914000075
13. J. C. Helton, F. J. Davis, Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems, *Reliab. Eng. Syst. Safe.*, **81** (2003), 23–69. https://doi.org/10.1016/S0951-8320(03)00058-9

14. N. Jiang, W. Layton, An algorithm for fast calculation of flow ensembles, *Int. J. Uncertain. Quan.*, **4** (2014), 273–301. https://doi.org/10.1615/Int.J.UncertaintyQuantification.2014007691

15. B. Jin, J. Zou, Numerical estimation of the Robin coefficient in a stationary diffusion equation, *IMA J. Numer. Anal.*, **30** (2010), 677–701. https://doi.org/10.1093/imanum/drn066

16. B. Jin, X. Lu, Numerical identification of a Robin coefficient in parabolic problems, *Math. Comp.*, **81** (2012), 1369–1398. https://doi.org/10.1090/S0025-5718-2012-02559-2

17. L. Ju, W. Leng, Z. Wang, S. Yuan, Numerical investigation of ensemble methods with block iterative solvers for evolution problems, *Discrete Contin. Dyn. B*, **25** (2020), 4905–4923. https://doi.org/10.3934/dcdsb.2020132

18. M. Li, X. Luo, An EMC-HDG scheme for the convection-diffusion equation with random diffusivity, *Numer. Algor.*, **90** (2022), 1755–1776. https://doi.org/10.1007/s11075-021-01250-2

19. M. Li, X. Luo, An MLMCE-HDG method for the convection diffusion equation with random diffusivity, *Comput. Math. Appl.*, **127** (2022), 127–143. https://doi.org/10.1016/j.camwa.2022.10.002

20. M. Li, X. Luo, An ensemble Monte Carlo HDG method for parabolic PDEs with random coefficients, *Int. J. Comput. Math.*, **100** (2023), 405–421. https://doi.org/10.1080/00207160.2022.2119082

21. M. Li, X. Luo, A multilevel Monte Carlo ensemble and hybridizable discontinuous Galerkin method for a stochastic parabolic problem, *Numer. Meth. Part. D. E.*, **39** (2023), 2840–2864. https://doi.org/10.1002/num.22990

22. X. Lin, M. Ng, An all-at-once preconditioner for evolutionary partial differential equations, *SIAM J. Sci. Comput.*, **43** (2021), A2766–A2784. https://doi.org/10.1137/20M1316354

23. K. Liu, B. M. Riviere, Discontinuous galerkin methods for elliptic partial differential equations with random coefficients, *Int. J. Comput. Math.*, **90** (2013), 2477–2490. https://doi.org/10.1080/00207160.2013.784280

24. G. J. Lord, C. E. Powell, T. Shardlow, *An introduction to computational stochastic PDEs*, New York: Cambridge University Press, 2014. https://doi.org/10.1017/CBO9781139017329

25. Y. Luo, Z. Wang, An ensemble algorithm for numerical solutions to deterministic and random parabolic PDEs, *SIAM J. Numer. Anal.*, **56** (2018), 859–876. https://doi.org/10.1137/17M1131489

26. Y. Luo, Z. Wang, A multilevel Monte Carlo ensemble scheme for solving random parabolic PDEs, *SIAM J. Sci. Comput.*, **41** (2019), A622–A642. https://doi.org/10.1137/18M1174635

27. J. Martínez-Frutos, M. Kessler, A. Münch, F. Periago, Robust optimal Robin boundary control for the transient heat equation with random input data, *Int. J. Numer. Meth. Eng.*, **108** (2016), 116–135. https://doi.org/10.1002/nme.5210

28. J. Martínez-Frutos, F. P. Esparza, *Optimal control of PDEs under uncertainty: An introduction with application to optimal shape design of structures*, Cham: Springer, 2018. https://doi.org/10.1007/978-3-319-98210-6

29. L. Mathelin, M. Y. Hussaini, T. A. Zang, Stochastic approaches to uncertainty quantification in CFD simulations, *Numer. Algor.*, **38** (2005), 209–236. https://doi.org/10.1007/BF02810624

30. J. Meng, P. Y. Zhu, H. B. Li, A block GCROT (m,k) method for linear systems with multiple right-hand sides, *J. Comput. Appl. Math.*, **255** (2014), 544–554. https://doi.org/10.1016/j.cam.2013.06.014

31. D. Xiu, J. S. Hesthaven, High-order collocation methods for differential equations with random inputs, *SIAM J. Sci. Comput.*, **27** (2005), 1118–1139. https://doi.org/10.1137/040615201

32. K. Ma, T. Sun, A non-overlapping DDM for optimal boundary control problems governed by parabolic equations, *Appl. Math. Optim.*, **79** (2019), 769–795. https://doi.org/10.1007/s00245-017-9456-7

33. A. L. Teckentrup, P. Jantsch, C. G. Webster, M. Gunzburger, A multilevel stochastic collocation method for partial differential equations with random input data, *SIAM/ASA J. Uncertain.*, **3** (2015), 1046–1074. https://doi.org/10.1137/140969002

34. T. Yao, C. Ye, X. Luo, S. Xiang, An ensemble scheme for the numerical solution of a random transient heat equation with uncertain inputs, *Numer. Algor.*, **94** (2023), 643–668. https://doi.org/10.1007/s11075-023-01514-z

35. T. Yao, C. Ye, X. Luo, S. Xiang, A variational MAX ensemble numerical algorism for a transient heat model with random inputs, *Netw. Heterog. Media*, **19** (2024), 1013–1037. https://doi.org/10.3934/nhm.2024045

36. C. Ye, T. Yao, H. Bi, X. Luo, A variational Crank-Nicolson ensemble Monte Carlo algorithm for a heat equation under uncertainty, *J. Comput. Appl. Math.*, **451** (2024), 116068. https://doi.org/10.1016/j.cam.2024.116068

37. J. Yong, C. Ye, X. Luo, S. Sun, Improved error estimates of ensemble Monte Carlo methods for random transient heat equations with uncertain inputs, *Comput. Appl. Math.*, **44** (2025), 58.https://doi.org/10.1007/s40314-024-03022-9

38. X. Zhu, E. M. Linebarger, D. Xiu, Multi-fidelity stochastic collocation method for computation of statistical moments, *J. Comput. Phys.*, **341** (2017), 386–396. https://doi.org/10.1016/j.jcp.2017.04.022

AIMS Press