Parameter estimation for linear parabolic SPDEs in two space dimensions based on high frequency data

Yozo Tonaki1 | Yusuke Kaino2 | Masayuki Uchida1,3

1Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka, Japan
2Graduate School of Maritime Sciences, Kobe University, Higashinada-ku, Kobe, Japan
3Center for Mathematical Modeling and Data Science (MMDS), Osaka University, Toyonaka, Osaka, Japan

Correspondence
Yozo Tonaki, Graduate School of Engineering Science, Osaka University, 1-3, Machikaneyama, Toyonaka, Osaka 560-8531, Japan.
Email: tonaki@sigmath.es.osaka-u.ac.jp

Abstract
We consider parameter estimation for a linear parabolic second-order stochastic partial differential equation (SPDE) in two space dimensions driven by two types of Q-Wiener processes based on high frequency data in time and space. We first estimate the parameters which appear in the eigenfunctions of the differential operator of the SPDE using the minimum contrast estimator based on the thinned data with respect to space, and then construct an approximate coordinate process of the SPDE. Furthermore, we propose estimators of the coefficient parameters of the SPDE utilizing the approximate coordinate process based on the thinned data with respect to time. We also give some simulation results.

KEYWORDS
adaptive estimation, high frequency data, Q-Wiener process, stochastic partial differential equations in two space dimensions

1 | INTRODUCTION

We consider the following linear parabolic stochastic partial differential equation (SPDE) in two space dimensions

\[
\begin{align*}
\mathrm{d}X_t^Q(y, z) &= \left\{ \theta_2 \left( \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + \theta_1 \frac{\partial}{\partial y} + \eta_1 \frac{\partial}{\partial z} + \theta_0 \right\} X_t^Q(y, z) \mathrm{d}t \\
&\quad + \sigma \mathrm{d}W_t^Q(y, z), \quad (t, y, z) \in [0, 1] \times D, \\
X_0^Q(y, z) &= \xi(y, z), \quad (y, z) \in D, \\
X_t^Q(y, z) &= 0, \quad (t, y, z) \in [0, 1] \times \partial D,
\end{align*}
\]

(1)

\[\text{© 2023 The Board of the Foundation of the Scandinavian Journal of Statistics.}\]
where \( D = (0, 1)^2 \), \( W_t^Q \) is a \( Q \)-Wiener process in a Sobolev space on \( D \), the initial value \( \xi \) is independent of \( W_t^Q \), \( \theta = (\theta_0, \theta_1, \eta_1, \theta_2) \) and \( \sigma \) are unknown parameters and \((\theta_0, \theta_1, \eta_1, \theta_2, \sigma) \in \mathbb{R}^3 \times (0, \infty)^2 \). Moreover, the parameter space \( \Theta \) is a compact convex subset of \( \mathbb{R}^3 \times (0, \infty)^2 \), \( \Theta^* = (\theta_0^*, \theta_1^*, \eta_1^*, \theta_2^*) \) and \( \sigma^* \) are true values of the parameters and we assume that \((\theta^*, \sigma^*) \in \text{Int} \ \Theta \). The original data are discrete observations \( X_{N,M} = \{ X_{t_i}^Q(y_j, z_j); i = 0, \ldots, N, j = 0, \ldots, M_1, j_2 = 0, \ldots, M_2 \} \) with \( t_i = i\Delta_N \), \( y_{j_1} = j_1/M_1 \) and \( z_{j_2} = j_2/M_2 \), where \( \Delta_N = 1/N \) and \( M := M_1M_2 \).

SPDEs appear in various fields such as physics, engineering, biology, and economics. In particular, the linear parabolic SPDE is a fundamental equation, and its typical example is the stochastic heat equation. The heat equation is known as the equation which describes the heat conduction in an object and the diffusion phenomenon of particles, and it is an essential equation which appears in various situations. For example, in the heat equation with one space dimension, we can consider the temperature variability of a thin object such as a wire, or a sea surface on a straight line. However, it is insufficient to consider practical problems with only a heat equation in one space dimension. Actually, in our daily problems, we often deal with two- or three-dimensional heat phenomena, such as temperature variability of a thin steel plate, sea surface, solid, or seawater. Thus, it is important to analyze the SPDEs in two space dimensions because it can address more general problems than SPDEs in one space dimension. For applications of linear parabolic SPDEs to biology and physics, see Altmeyer et al. (2022) and Piterbarg and Ostrovskii (1997).

Statistical inference for SPDE models based on discrete observations has been studied by many researchers, see, for example, Markussen (2003), Bibinger and Trabs (2020), Chong (2019, 2020), Cialenco et al. (2020), Cialenco and Huang (2020), Kaino and Uchida (2020), and references therein. Recently, Kaino and Uchida (2021) proposed the adaptive maximum likelihood type estimation for the coefficient parameters of linear parabolic second-order SPDEs in one space dimension with a small noise. Hildebrandt and Trabs (2021a, 2021b) studied the estimation for the coefficient parameters of linear parabolic and semilinear SPDEs in one space dimension, respectively, using a contrast function with double increments in time and space.

In particular, Bibinger and Trabs (2020) treated the following linear parabolic SPDE in one space dimension

\[
\begin{align*}
\frac{dX_t(y)}{dt} &= \left( \theta_2 \frac{\partial^2}{\partial y^2} + \theta_1 \frac{\partial}{\partial y} + \theta_0 \right) X_t(y) + \sigma dB_t(y), \quad (t, y) \in [0, T] \times (0, 1), \\
X_0(y) &= \xi(y), \quad y \in [0, 1], \quad X_t(1) = 0, \quad t \in [0, T],
\end{align*}
\]  

(2)

where \( T > 0 \), \( B_t \) is a cylindrical Brownian motion in a Sobolev space on \((0, 1)\), \( \xi \) is the initial value, \( \theta_0, \theta_1, \theta_2 \) and \( \sigma \) are unknown parameters and \((\theta_0, \theta_1, \theta_2, \sigma) \in \mathbb{R}^2 \times (0, \infty)^2 \). They proposed minimum contrast estimators for \( \kappa = \theta_1/\theta_2 \), \( \sigma_0^2 = \sigma^2/\theta_2^{1/2} \) in the case where \( T \) is fixed. Since the coordinate process \( x_k(t), k \geq 1 \) of the SPDE (2) is expressed as

\[
x_k(t) = \sqrt{2} \int_0^1 X_t(y) \sin(\pi ky)e^{\pi y/2} dy,
\]

and is a diffusion process satisfying the stochastic differential equation

\[
\begin{align*}
dx_k(t) &= -\lambda_k x_k(t) dt + \sigma dw_k(t), \quad x_k(0) = \sqrt{2} \int_0^1 \xi(y) \sin(\pi ky)e^{\pi y/2} dy,
\end{align*}
\]  

(3)
where \( \{w_k\}_{k \geq 1} \) are independent real-valued standard Brownian motions and \( \lambda_k = -\theta_0 + \frac{\theta_1^2}{4\theta_2} + \pi^2k^2\theta_2 \). Kaino and Uchida (2020) constructed an approximate coordinate process by a Riemann sum

\[
\hat{x}_k(t) = \frac{1}{M} \sum_{j=1}^{M} \sqrt{2}X_t(y_j) \sin(\pi ky_j)e^{\hat{\kappa}y_j/2},
\]

with observations in space \( \{y_j\}_{j=1}^{M} \) and an estimator \( \hat{\kappa} \) of \( \kappa \), and proposed estimators of \( \sigma^2, \theta_2, \) and \( \theta_1 \) using the approximate coordinate process based on the thinned data with respect to time and statistical inference for diffusion processes. Furthermore, they extended the results of Bibinger and Trabs (2020) to the case where \( T \) is large and proposed estimators of \( \sigma^2, \theta_2, \theta_1, \) and \( \theta_0 \). For details of statistical inference for diffusion processes based on discrete observations, see Genon-Catalot and Jacod (1993), Kessler (1997), and Uchida and Yoshida (2012, 2013).

In this paper, we apply the estimation method based on the approximate coordinate process to SPDEs in two space dimensions. Specifically, we provide minimum contrast estimators based on quadratic variations, which are used to construct an approximate coordinate process and estimate the coefficient parameters of the SPDE (1). In this case, we need to be careful in setting a noise \( W_t^Q \) because \( X_t^Q \) may not be square integrable for any \( t > 0 \) depending on the choice of the Q-Wiener process. Moreover, since the method of Kaino and Uchida (2020) is based on the property that the random field \( X_t \) admits a spectral decomposition

\[
X_t(y) = \sum_{k=1}^{\infty} x_k(t)e_k(y),
\]

with \( e_k(y) = \sqrt{2}\sin(\pi ky)e^{-\kappa y/2} \) and \( x_k(t) \) in (3), it is also required to set a noise \( W_t^Q \) such that the random field \( X_t^Q \) in the SPDE (1) can be decomposed. See Chong and Dalang (2020) for power variations of SPDEs on general domains that include the stochastic heat equation and Hübner et al. (1993) and Cialenco and Glatt-Holtz (2011) for parameter estimation for SPDEs driven by a Q-Wiener process. See also the surveys by Lototsky (2009) and Cialenco (2018) for statistical inference based on the spectral approach.

The main contribution of this paper is to provide asymptotically normal estimators for the coefficient parameters \( (\theta_0, \theta_1, \eta_1, \theta_2, \sigma^2) \) with the convergence rate \( \sqrt{n} \) for some \( n \leq N \), where \( \theta_0 \) can be estimated if the SPDE (1) is driven by the \( Q_1 \)-Wiener process given in (4) below.

This paper is organized as follows. In Section 2, we present the setting of our model. We also discuss how to choose a \( Q \)-Wiener process and introduce two types of \( Q \)-Wiener processes. In Section 3, we first propose minimum contrast estimators of the parameters appearing in the eigenfunctions of the differential operator of the SPDE by using the thinned data with respect to space. Next, we construct an approximate coordinate process by using these minimum contrast estimators, and provide estimators of the coefficient parameters of the SPDE based on the thinned data with respect to time. We then show that the estimators of the coefficient parameters are asymptotically normal. In Section 4, we give some simulation studies. The proofs of our results in Section 3 are provided in the supporting information.
2 PRELIMINARIES

Let \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)\) be a stochastic basis with usual conditions, and let \(\{w_{k,\ell}\}_{k,\ell \in \mathbb{N}}\) be independent real valued standard Brownian motions on this basis.

Let \(A_\theta\) be the differential operator defined by

\[-A_\theta = \theta_2 \left( \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + \theta_1 \frac{\partial}{\partial y} + \eta_1 \frac{\partial}{\partial z} + \theta_0.\]

Note that the SPDE (1) is represented as

\[dX_t^Q(y, z) = -A_\theta X_t^Q(y, z)dt + \sigma dW_t^Q(y, z).\]

The domain of \(A_\theta\) is given by \(\mathcal{D}(A_\theta) = H^2(D) \cap H^1_0(D)\), where \(H^p(D)\) denotes the \(L^2\)-Sobolev space of order \(p \in \mathbb{N}\) and \(H^1_0(D)\) is the closure of \(C_c^\infty(D)\) in \(H^1(D)\). For \(k, \ell \in \mathbb{N}\), the eigenfunctions \(e_{k,\ell}\) of \(A_\theta\) and the corresponding eigenvalues \(\lambda_{k,\ell}\) are given by

\[e_{k,\ell}(y, z) = 2 \sin(\pi k y) \sin(\pi \ell z) e^{-\frac{\theta_1}{4\theta_2^2} y^2 - \frac{\eta_1}{4\theta_2^2} z^2}, \quad y, z \in [0, 1],\]

\[\lambda_{k,\ell} = -\theta_0 + \frac{\theta_1^2}{4 \theta_2^2} + \frac{\pi^2 (k^2 + \ell^2)}{2}.\]

We then obtain that \(A_\theta e_{k,\ell} = \lambda_{k,\ell} e_{k,\ell}\) for \(k, \ell \in \mathbb{N}\). We set \(H_\theta = \{f : D \to \mathbb{R} |||f||_\theta < \infty\}\) with

\[\langle f, g \rangle_\theta = \int_0^1 \int_0^1 f(y, z)g(y, z)e^{\theta_1 y} e^{\eta_1 z} dy dz, \quad ||f||_\theta = \langle f, f \rangle_\theta^{1/2}.\]

We consider two types of \(Q\)-Wiener processes, \(\{W_t^{Q_1}\}_{t \geq 0}\) and \(\{W_t^{Q_2}\}_{t \geq 0}\), defined as follows.

\[\langle W_t^{Q_1}, f \rangle_\theta = \sum_{k, \ell \geq 1} \lambda_{k,\ell}^{-a/2} \langle f, e_{k,\ell} \rangle_\theta w_{k,\ell}(t), \quad (4)\]

\[\langle W_t^{Q_2}, f \rangle_\theta = \sum_{k, \ell \geq 1} \mu_{k,\ell}^{-a/2} \langle f, e_{k,\ell} \rangle_\theta w_{k,\ell}(t), \quad (5)\]

for \(f \in H_\theta\) and \(t \geq 0\), where \(a \in (0, 1)\), \(\mu_{k,\ell} = \pi^2 (k^2 + \ell^2) + \mu_0\) and \(\mu_0 \in (-2\pi^2, \infty)\). \(\mu_0\) is a known or unknown parameter, the parameter space of \(\mu_0\) is a compact convex subset of \((-2\pi^2, \infty)\) and the true value \(\mu_0^\ast\) belongs to its interior. Since we do not have any consistent estimator of \(\alpha\), we assume that \(\alpha\) is known. It is future work to study the estimability of \(\alpha\). The restriction of \(\alpha\) guarantees the existence of solutions and the estimability of the parameters, see Remarks 2 and 4 for details.

The SPDE driven by the \(Q_1\)-Wiener process is a model where \(\lambda_{k,\ell} = -\theta_0 + \frac{\theta_1^2 + \eta_1^2}{4 \theta_2^2} + \frac{\pi^2 (k^2 + \ell^2)}{2}\) is included in the noise and \(\theta_2, \theta_1, \eta_1, \sigma^2\), and especially \(\theta_0\) can be estimated, see Theorem 2. In contrast, the SPDE driven by the \(Q_2\)-Wiener process is a model whose noise does not depend on the coefficient parameters, and we can construct consistent estimators for the parameters \(\theta_2, \theta_1, \eta_1, \sigma^2\), and \(\mu_0\) except for \(\theta_0\), see Theorem 4. Although it seems that the choice of the \(Q\)-Wiener process depends on statistical modeling of SPDE, note that there is no difference in the fact that the coefficient parameters \(\theta_2, \theta_1, \eta_1, \sigma^2\) can be estimated for both models, and even if one chooses a \(Q\)-Wiener process with a damping factor such that the random field \(X^Q_t(y, z)\) is spectrally...
decomposable as in (6) besides these two, the coefficient parameters \( \theta_2, \theta_1, \eta_1 \), and \( \sigma^2 \) can still be estimated by following our estimation technique. In this paper, we consider only two types of statistical models (1) whose driving processes are the \( Q \)-Wiener processes given by (4) or (5).

We assume that \( \xi \in H_\theta \) and \( \lambda_{2,1} = -\theta_1^* + \frac{q_2^*(\eta_1^*)^2}{4\theta_2^*} + 2\pi^2 \theta_2^* > 0 \). \( X_t^Q \) is called a mild solution of (1) on \( D \) if it satisfies that for any \( t \in [0,1] \),

\[
X_t^Q = e^{-t\lambda_0} \xi + \sigma \int_0^t e^{-(t-s)\lambda_0} dW_s^Q \quad \text{a.s.,}
\]

where \( e^{-t\lambda_0} u = \sum_{k,\ell \geq 1} e^{-\lambda_{k,\ell} t} (u, e_{k,\ell})_\theta e_{k,\ell} \) for \( u \in H_\theta \). The random field \( X_t^Q(y, z) \) is expressed by using the coordinate process \( x_{k,\ell}^Q(t) = (X_t^Q, e_{k,\ell})_\theta \) as follows.

\[
x_t^Q(y, z) = \sum_{k,\ell \geq 1} x_{k,\ell}^Q(t) e_{k,\ell}(y, z),
\]

where \( x_{k,\ell}^Q \) and \( x_{k,\ell}^Q \) satisfy the Ornstein–Uhlenbeck processes

\[
\begin{align*}
\text{d}x_{k,\ell}^Q(t) &= -\lambda_{k,\ell} x_{k,\ell}^Q(t) \text{d}t + \sigma \lambda_{k,\ell}^{-1/2} \text{d}w_{k,\ell}(t), \quad x_{k,\ell}^Q(0) = (\xi, e_{k,\ell})_\theta, \\
\text{d}x_{k,\ell}^Q(t) &= -\lambda_{k,\ell} x_{k,\ell}^Q(t) \text{d}t + \sigma \mu_{k,\ell}^{-1/2} \text{d}w_{k,\ell}(t), \quad x_{k,\ell}^Q(0) = (\xi, e_{k,\ell})_\theta,
\end{align*}
\]

respectively. We assume the following conditions on the initial value \( \xi \in H_\theta \) of the SPDE (1).

**Assumption 1.** The initial value \( \xi \) satisfies either (i) or (ii), and both (iii) and (iv).

(i) \( \mathbb{E}[\langle \xi, e_{k,\ell} \rangle_\theta] = 0 \) for all \( k, \ell \geq 1 \) and \( \sup_{k,\ell \geq 1} \lambda_{k,\ell}^{1+a} \mathbb{E}[\langle \xi, e_k \rangle^2_\theta] < \infty \).

(ii) \( \mathbb{E}[\|A_\theta^{(1+a)/2} \xi\|^2_\theta] < \infty \).

(iii) \( \sup_{k,\ell \geq 1} \lambda_{k,\ell}^{1+a} \mathbb{E}[\langle \xi, e_{k,\ell} \rangle^2_\theta] < \infty \).

(iv) \( \{\langle \xi, e_{k,\ell} \rangle_\theta \}_{k,\ell \geq 1} \) are independent.

**Remark 1.** (1) From \( \|A_\theta^{(1+a)/2} \xi\|^2_\theta = \sum_{k,\ell \geq 1} \lambda_{k,\ell}^{1+a} \mathbb{E}[\langle \xi, e_{k,\ell} \rangle^2_\theta] \), (ii) in Assumption 1 can be replaced by

\[
\sum_{k,\ell \geq 1} \lambda_{k,\ell}^{1+a} \mathbb{E}[\langle \xi, e_{k,\ell} \rangle^2_\theta] < \infty.
\]

Moreover, for a nonrandom function \( \xi(y, z) = \xi_1(y) \xi_2(z) \in H_\theta \), if \( \xi_1, \xi_2 \in C^2([0,1]) \) for \( 0 < \alpha < 1/2 \), or if \( \xi_1, \xi_2 \in C^3([0,1]) \) for \( 1/2 \leq \alpha < 1 \), then \( \xi \) satisfies (ii) and (iii).

Indeed, noting that

\[
\langle \xi, e_{k,\ell} \rangle_\theta = 2 \int_0^1 \xi_1(y) \sin(\pi k y) e^{\frac{\pi k y}{2}} dy \int_0^1 \xi_2(z) \sin(\pi \ell z) e^{\frac{\pi \ell z}{2}} dz,
\]

and that for \( \xi_1 \in C^p([0,1]) \) \((p = 2, 3)\) such that \( \xi_1(0) = \xi_1(1) = 0 \),

\[
\left| \int_0^1 \xi_1(y) \sin(\pi k y) e^{\frac{\pi k y}{2}} dy \right| \leq \frac{C_1}{k^p}.
\]
Specifically, (4) can be obtained by setting \( \lambda_{k,\ell} \leq C_3(k^2 + \ell^2) \) and

\[
\sum_{k,\ell \geq 1} \lambda_{k,\ell}^{1+a} \langle \xi, e_{k,\ell} \rangle_\theta^2 \leq C \sum_{k,\ell \geq 1} \frac{(k^2 + \ell^2)^{1+a}}{k^{2p}\ell^{2p}} \leq C' \sum_{k \geq 1} \frac{1}{k^{2(p-1)-2a}} \sum_{\ell \geq 1} \frac{1}{\ell^{2p}},
\]

\( \sum_{k,\ell \geq 1} \lambda_{k,\ell}^{1+a} \langle \xi, e_{k,\ell} \rangle_\theta^2 \) converges if \( p > \alpha + 3/2 \), and therefore (ii) is satisfied by setting \( p = 2 \) for \( 0 < \alpha < 1/2 \) and \( p = 3 \) for \( 1/2 \leq \alpha < 1 \). Similarly, \( \xi \) also satisfies (iii).

(2) If \( \xi \) is distributed according to the stationary distribution of (1) driven by the \( Q_1 \)-Wiener process, where \( \langle \xi, e_{k,\ell} \rangle_\theta \) are independently \( N(0, \frac{\sigma^2}{2k_{\ell}^\alpha}) \)-distributed, then (i), (iii), and (iv) in Assumption 1 are satisfied.

**Remark 2.** Consider the case where the SPDE (1) is driven by a cylindrical Brownian motion \( \{B_t\}_{t \geq 0} \) defined as

\[
\langle B_t, f \rangle_\theta = \sum_{k,\ell \geq 1} \langle f, e_{k,\ell} \rangle_\theta w_{k,\ell}(t), \quad f \in H_\theta, \quad t \geq 0,
\]

as in the setting of previous studies on parameter estimation for SPDEs in one space dimension, in other words, consider the SPDE represented by

\[
dX_t^I(y, z) = -A_0X_t^I(y, z)dt + \sigma dW_t^0(y, z),
\]

where \( I \) is the identity operator and \( B_t = W_t^0 \). In this case, it follows that \( \sup_{0 \leq t \leq 1} \mathbb{E}[\|X_t^I\|_\theta^2] = \infty \), see Walsh (1986). For this reason, we need to introduce a noise \( W_t^Q \) with a damping factor such as (4) and satisfying \( \sup_{0 \leq t \leq 1} \mathbb{E}[\|X_t^Q\|_\theta^2] < \infty \). Specifically, (4) can be obtained by setting \( Q_1 \) as follows. The domain \( \mathcal{D}(A_\theta^{-1/2}) \supset H_\theta \) of \( A_\theta^{-1/2} \) is a Hilbert space with inner product

\[
\langle u, v \rangle_{\theta, -1/2} = \langle A_\theta^{-1/2}u, A_\theta^{-1/2}v \rangle_\theta,
\]

and its corresponding induced norm \( \|u\|_{\theta, -1/2} = \|A_\theta^{-1/2}u\|_\theta \), and \( \{v_{k,\ell}\}_{k,\ell \geq 1}, v_{k,\ell} = e_{k,\ell}/\|e_{k,\ell}\|_{\theta, -1/2} \) is the complete orthonormal system on \( \mathcal{D}(A_\theta^{-1/2}) \). By defining the covariance operator \( Q_1 \) on \( \mathcal{D}(A_\theta^{-1/2}) \) by

\[
Q_1 u = \sum_{k,\ell \geq 1} \lambda_{k,\ell}^{-a} \langle u, v_{k,\ell} \rangle_{\theta, -1/2} v_{k,\ell},
\]

for \( u = \sum_{k,\ell \geq 1} \langle u, v_{k,\ell} \rangle_{\theta, -1/2} v_{k,\ell} \in \mathcal{D}(A_\theta^{-1/2}) \) and \( \alpha > 0 \), \( Q_1 \) is of trace class for \( \alpha > 0 \): \( \text{Tr} \ Q_1 = \sum_{k,\ell \geq 1} \lambda_{k,\ell}^{-a} < \infty \), the \( Q_1 \)-Wiener process \( \{W_t^{Q_1}\}_{t \geq 0} \) is well-defined in \( \mathcal{D}(A_\theta^{-1/2}) \) and it follows that for \( f \in H_\theta \),

\[
\langle W_t^{Q_1}, f \rangle_\theta = \langle A_\theta^{1/2}W_t^{Q_1}, A_\theta^{1/2}f \rangle_{\theta, -1/2}
= \sum_{k,\ell \geq 1} \langle W_t^{Q_1}, A_\theta^{1/2}v_{k,\ell} \rangle_{\theta, -1/2} \langle f, A_\theta^{1/2}v_{k,\ell} \rangle_{\theta, -1/2}
= \sum_{k,\ell \geq 1} \lambda_{k,\ell}^{-a/2} \langle f, e_{k,\ell} \rangle_\theta w_{k,\ell}(t).
\]
If we set $W_t^Q = W_t^{Q_1}$, then there exists a unique mild solution $X_t^{Q_1}$ of the SPDE (1) which satisfies $\sup_{0 \leq t \leq 1} \mathbb{E}[||X_t^{Q_1}||_\theta^2] < \infty$ under $\lambda_{1,1}^* > 0$ and Assumption 1. Indeed, since $\sup_{k,\ell \geq 1} \lambda_{k,\ell}^{1+\alpha} \mathbb{E}[\langle \xi, e_{k,\ell} \rangle^2] < \infty$ under (i) or (ii) in Assumption 1, it holds that

$$
\mathbb{E}[||X_t^{Q_1}||_\theta^2] = \sum_{k,\ell \geq 1} \mathbb{E}[x_{k,\ell}(t)^2] = \sum_{k,\ell \geq 1} e^{-2\lambda_{k,\ell} t} \mathbb{E}[\langle \xi, e_{k,\ell} \rangle^2] + \sigma^2 \sum_{k,\ell \geq 1} \frac{1 - e^{-2\lambda_{k,\ell} t}}{2\lambda_{k,\ell}^{1+\alpha}}$

$$
\leq \sum_{k,\ell \geq 1} \frac{1}{\lambda_{k,\ell}^{1+\alpha}} \lambda_{k,\ell}^{1+\alpha} \mathbb{E}[\langle \xi, e_{k,\ell} \rangle^2] + \sigma^2 \sum_{k,\ell \geq 1} \frac{1}{2\lambda_{k,\ell}^{1+\alpha}}$

$$
\leq C \sum_{k,\ell \geq 1} \frac{1}{\lambda_{k,\ell}^{1+\alpha}} < \infty,$n

for $\alpha > 0$, and the mild solution $X_t^{Q_1}$ satisfies $\sup_{0 \leq t \leq 1} \mathbb{E}[||X_t^{Q_1}||_\theta^2] < \infty$. Refer to Lord et al. (2014), Da Prato and Zabczyk (2014), and Lototsky and Rozovsky (2017) for details on the Q-Wiener process.

**Remark 3.** Unlike the damping factor $\{\lambda_{k,\ell}^{-a/2}\}_{k,\ell \geq 1}$ of the $Q_1$-Wiener process, the damping factor $\{\mu_{k,\ell}^{-a/2}\}_{k,\ell \geq 1}$ of the $Q_2$-Wiener process does not include the parameter $\theta$ of the differential operator $A_\theta$. Moreover, by setting $\kappa = \theta_1/\theta_2$, $\eta = \eta_1/\theta_2$ and $\zeta = \left(\frac{k^2 + \eta^2}{4} - \mu_0, \kappa, \eta, 1\right)$, it can be regarded as

$$
\mu_{k,\ell} = \pi^2(k^2 + \eta^2) + \mu_0 = -\left(\frac{k^2 + \eta^2}{4} - \mu_0\right) + \frac{k^2 + \eta^2}{4} + \pi^2(k^2 + \eta^2) \cdot 1,$n

and it holds that $\langle \cdot, \cdot \rangle_\zeta = \langle \cdot, \cdot \rangle_\theta$, $H_\zeta = H_\theta$ and $A_\zeta e_{k,\ell} = \mu_{k,\ell} e_{k,\ell}$, where

$$
-A_\zeta = \left(\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + \kappa \frac{\partial}{\partial y} + \eta \frac{\partial}{\partial z} + \left(\frac{k^2 + \eta^2}{4} - \mu_0\right).$

Hence, the $Q_2$-Wiener process can be constructed in the same way. Indeed, by choosing $Q_2$ as the covariance operator on $\mathcal{D}(A_\zeta^{-1/2}) \supset H_\zeta(= H_\theta)$ that satisfies

$$
Q_2u = \sum_{k,\ell \geq 1} \mu_{k,\ell}^{-1+\alpha} \langle u, v_{k,\ell} \rangle_{\zeta,-1/2} v_{k,\ell},$

for $v_{k,\ell} = e_{k,\ell}/\|e_{k,\ell}\|_{\zeta,-1/2}$, $u \in \mathcal{D}(A_\zeta^{-1/2})$ and $\alpha > 0$, the $Q_2$-Wiener process $\{W_t^{Q_2}\}_{t \geq 0}$ is well-defined in $\mathcal{D}(A_\zeta^{-1/2})$ and

$$
\langle W_t^{Q_2}, f \rangle_\theta = \langle W_t^{Q_1}, f \rangle_\zeta = \sum_{k,\ell \geq 1} \mu_{k,\ell}^{-a/2} \langle f, e_{k,\ell} \rangle_{\zeta} w_{k,\ell}(t) = \sum_{k,\ell \geq 1} \mu_{k,\ell}^{-a/2} \langle f, e_{k,\ell} \rangle_\theta w_{k,\ell}(t),$

for $f \in H_\theta$ and $t \geq 0.$
3 | MAIN RESULTS

3.1 | SPDE driven by \( Q_1 \)-Wiener process

In this subsection, we deal with the SPDE \( (1) \) driven by the \( Q_1 \)-Wiener process defined as \( (4) \). We first consider the estimation for the parameters which appear in the eigenfunctions \( e_k^{(1)} \) based on the thinned data with respect to space. Specifically, we set the thinned data \( X_{N,m}^{\delta,1} = \{ X_i^{\delta,1}(\bar{y}_j, \bar{z}_j); i = 0, \ldots, N, j_1 = 0, \ldots, m_1, j_2 = 0, \ldots, m_2 \} \) as follows. Set \( \bar{m}_1 \leq M_1 \) and \( \bar{m}_2 \leq M_2 \), and let

\[
\bar{y}_{j_1} = \left[ \frac{M_1}{\bar{m}_1} \right] \frac{J_{1} + j_1}{M_1}, \quad j_1 = 0, \ldots, \bar{m}_1,
\]

\[
\bar{z}_{j_2} = \left[ \frac{M_2}{\bar{m}_2} \right] \frac{J_{2} + j_2}{M_2}, \quad j_2 = 0, \ldots, \bar{m}_2.
\]

Let \( \delta \in (0, 1/2) \). Since there exist \( J_1, J_2 \geq 1, m_1, m_2 \geq 1 \) such that

\[
\bar{y}_{j_1} < \delta \leq \bar{y}_{j_1 + 1} < \cdots < \bar{y}_{j_1 + m_1} \leq 1 - \delta < \bar{y}_{j_1 + m_1 + 1},
\]

\[
\bar{z}_{j_2} < \delta \leq \bar{z}_{j_2 + 1} < \cdots < \bar{z}_{j_2 + m_2} \leq 1 - \delta < \bar{z}_{j_2 + m_2 + 1},
\]

we set

\[
\bar{y}_{j_1} = \bar{y}_{j_1 + j_1} = \left[ \frac{M_1}{\bar{m}_1} \right] \frac{J_{1} + j_1}{M_1}, \quad j_1 = 1, \ldots, m_1,
\]

\[
\bar{z}_{j_2} = \bar{z}_{j_2 + j_2} = \left[ \frac{M_2}{\bar{m}_2} \right] \frac{J_{2} + j_2}{M_2}, \quad j_2 = 1, \ldots, m_2,
\]

and \( D_\delta = [\delta, 1 - \delta]^2 \subset D \). Let \( m := m_1 m_2 \) and note that \( (\bar{y}_{j_1}, \bar{z}_{j_2}) \in D_\delta \) for any \( j_1 = 1, \ldots, m_1 \) and \( j_2 = 1, \ldots, m_2 \). Throughout, we will write \( m \to \infty \) when \( m_1 \to \infty \) and \( m_2 \to \infty \).

We write \( \Delta_i X^Q(y, z) = X_i^Q(y, z) - X_{i-1}^Q(y, z) \) for \( X^Q \). Let \( \Gamma(s) = \int_0^s x^{-1} e^{-s} dx, s > 0 \).

**Proposition 1.** Under Assumption 1, it holds that uniformly in \((y, z) \in D_\delta\),

\[
\mathbb{E}[(\Delta_i X^Q)^2(y, z)] = \Delta_N^a \frac{\sigma^2 \Gamma(1 - \alpha)}{4 \pi a \theta_2} e^{-\frac{\sigma}{a} y} e^{-\frac{\sigma}{a} z} + r_{N,i} + O(\Delta_N), \tag{10}
\]

where \( \sum_{i=1}^N r_{N,i} = O(\Delta_N^a) \), and thus

\[
\mathbb{E} \left[ \frac{1}{N \Delta_N^a} \sum_{i=1}^N (\Delta_i X^Q)^2(y, z) \right] = \frac{\sigma^2 \Gamma(1 - \alpha)}{4 \pi a \theta_2} e^{-\frac{\sigma}{a} y} e^{-\frac{\sigma}{a} z} + O(\Delta_N^a).
\]

**Remark 4.** It follows from the proofs of Lemmas C.3 and C.4 in the Data S1 and Proposition 1 that

\[
\mathbb{E}[(\Delta_i X^Q)^2(y, z)] = \sigma^2 \sum_{k,c \geq 1} \frac{1 - e^{-\lambda_c \Delta_N}}{\lambda_{k,c}^{1+a}} \left( 1 - \frac{1 - e^{-\lambda_k \Delta_N}}{2} e^{-2 \lambda_k (1 - \Delta_N)} \right) e_{k,c}^2(y, z) + r_{N,i}, \tag{11}
\]
where \( r_{N,i} \) is the sequence in (10). According to Lemmas C.2 and C.4 in the Data S1, the restriction \( \alpha \in (0, 1) \) allows us to approximate the summation in (11) containing \( \lambda_{k,\ell} \) with unknown parameters by an explicit expression such as the main part in (10). This makes it possible to estimate \( \sigma^2 / \theta_2, \theta_1 / \theta_2 \) and \( \eta_1 / \theta_2 \). In the case where \( \alpha \geq 1 \), we cannot use the same parameter estimation technique as in the case where \( \alpha \in (0, 1) \) because Lemma C.1 in the Data S1 is not applicable to \( f(s) = \frac{1-e^{-s}}{s^{\alpha+1}} (\alpha \geq 1) \) and the first term of the right-hand side of (11) is not approximated by the first term of the right-hand side of (10). We leave the construction of the estimators in the case where \( \alpha \geq 1 \) for future work.

For \( X^Q \), set

\[
Z^Q_N(y, z) = \frac{1}{N \Delta_N^\alpha} \sum_{i=1}^N (\Delta_i X^Q(y, z), z).
\]

Let \( s = \frac{\sigma^2}{\theta_2}, \kappa = \frac{\theta_1}{\theta_2} \) and \( \eta = \frac{\eta_1}{\theta_2} \). By setting the contrast function

\[
U^{(1)}_{N,m}(s, \kappa, \eta) = \sum_{j_1=1}^{m_1} \sum_{j_2=1}^{m_2} \left( Z^Q_N(\tilde{y}_{j_1}, \tilde{z}_{j_2}) - \frac{\Gamma(1 - \alpha)}{4\pi s^{\alpha+1}} e^{-(\kappa \tilde{y}_{j_1} + \eta \tilde{z}_{j_2})} \right)^2,
\]

the minimum contrast estimators of \( s, \kappa \) and \( \eta \) are defined as

\[
(\hat{s}, \hat{\kappa}, \hat{\eta}) = \operatorname{arg inf}_{(s, \kappa, \eta) \in \Xi_1} U^{(1)}_{N,m}(s, \kappa, \eta),
\]

where \( \Xi_1 \) is a compact convex subset of \( (0, \infty) \times \mathbb{R}^2 \). We suppose that the true value \((s^*, \kappa^*, \eta^*)\) belongs to \( \text{Int} \ \Xi_1 \).

**Theorem 1.** Under Assumption 1, it holds that as \( N \to \infty \) and \( m \to \infty \),

\[
N^{-1/\alpha (1-\alpha)} \begin{pmatrix}
\hat{s} - s^* \\
\hat{\kappa} - \kappa^* \\
\hat{\eta} - \eta^*
\end{pmatrix} = O_p(1).
\]

**Remark 5.** (1) Bibinger and Trabs (2020), who dealt with the parameter estimation for linear parabolic SPDEs in one space dimension, showed that the estimators have asymptotic normality. On the other hand, in two space dimensions, the evaluation of the remainder in (10) is worse than that in one space dimension because of the increase in dimension, and the asymptotic normality cannot be derived. However, it is possible to estimate each coefficient parameter even if the estimators do not have asymptotic normality.

(2) In contrast to the one-dimensional case (proposition 3.1 in Bibinger & Trabs, 2020), there is no profit from the number of spatial observations \( m \) in Theorem 1 because of the fact \( \text{Var}[Z^Q_N(y, z)] = O(\Delta_N^{\alpha/2(1-\alpha)}) \) uniformly in \((y, z) \in D_\delta\), see Proof of Theorem 1 in the Data S1. Moreover, the rate \( N^{-1/\alpha (1-\alpha)} \) is attributed to this fact.

Once we estimate \((s, \kappa, \eta)\), we can construct an approximation for the coordinate process \( X^Q_{k,\ell}(t) \) by a Riemann sum. Furthermore, noting that the coordinate process \( X^Q_{k,\ell}(t) \) satisfies (7),
we can estimate the volatility parameter \( \sigma_{k, \ell} := \sigma \lambda_{k, \ell}^{-a/2} \) by using statistical inference for diffusion processes. Since

\[
\lambda_{1,2} = \left( \frac{\sigma_{1,1}^2}{\sigma_{1,2}^2} \right)^{1/a} \lambda_{1,1}, \quad \lambda_{1,1} = \left( \frac{\sigma_2^2}{\sigma_{1,1}^2} \right)^{1/a} \left( \frac{s \theta_2}{\sigma_{1,1}^2} \right)^{1/a},
\]

and \( \lambda_{1,2} - \lambda_{1,1} = 3\pi^2 \theta_2, \theta_2 \) can be expressed by using \( s, \sigma_{1,1} \), and \( \sigma_{1,2} \) as follows.

\[
\theta_2 = \left\{ \frac{3\pi^2}{s^{1/a}} \left( \frac{1}{\sigma_{1,2}^2} - \frac{1}{\sigma_{1,1}^2} \right)^{-1} \right\}^{\sum_{i=0}^{n}}^{\sum_{i=0}^{n}}.
\]

\( \sigma^2, \theta_1, \eta_1, \) and \( \theta_0 \) can also be expressed by using \( \theta_2, s, \kappa, \eta, \) and \( \lambda_{1,1} \) as follows.

\[
\sigma^2 = s \theta_2, \quad \theta_1 = 2^\kappa \theta_2, \quad \eta_1 = \eta \theta_2,
\]

\[
\theta_0 = -\lambda_{1,1} + \left( \frac{\kappa^2 + \eta^2}{4} + 2\pi^2 \right) \theta_2.
\]

With the above in mind, we construct an approximate coordinate process by using the thinned data with respect to time, and consider the estimation for each coefficient parameter. Specifically, we consider the thinned data \( \mathbb{X}^{(2)}_{n,M} = \{ X_{Q}^{(1)}(y_j, z_j); i = 0, \ldots, n, j_1 = 0, \ldots, j_2 = 0, \ldots, M \} \) with \( n \leq N \) and \( \bar{t}_i = \lfloor \frac{N}{n} \rfloor \bar{t}_i \). Since the coordinate process \( X_{k, \ell}^{(1)}(t) \) can be expressed as

\[
X_{k, \ell}^{(1)}(t) = 2 \int_0^1 \int_0^1 X_{Q}^{(1)}(y, z) \sin(\pi \ell y) \sin(\pi \ell z) e^{2y} e^{2z} dxdz,
\]

we consider

\[
\tilde{X}_{k, \ell}^{(1)}(\bar{t}_i) = 2 \sum_{j_1=1}^{M_1} \sum_{j_2=1}^{M_2} X_{Q}^{(1)}(y_j, z_j) \sin(\pi \ell y_j) \sin(\pi \ell z_j) e^{2y_j} e^{2z_j}, \quad i = 1, \ldots, n,
\]

as an approximation of \( X_{k, \ell}^{(1)}(t) \). By using the thinned data based on the approximate coordinate process \( \{ \tilde{X}_{k, \ell}^{(1)}(\bar{t}_i) \}_{i=1}^n \), the estimator of \( \sigma^2_{k, \ell} \) is defined as

\[
\hat{\sigma}^2_{k, \ell} = \sum_{i=1}^{n} \left( \tilde{X}_{k, \ell}^{(1)}(\bar{t}_i) - \tilde{X}_{k, \ell}^{(1)}(\bar{t}_{i-1}) \right)^2.
\]

Moreover, the estimators of \( \theta_2, \sigma^2, \theta_1, \eta_1, \) and \( \theta_0 \) are defined as

\[
\hat{\theta}_2 = \left\{ \frac{3\pi^2}{s^{1/a}} \left( \frac{1}{(\hat{\sigma}^2_{1,2})^{1/a}} - \frac{1}{(\hat{\sigma}^2_{1,1})^{1/a}} \right)^{-1} \right\}^{\sum_{i=0}^{n}}^{\sum_{i=0}^{n}}.
\]
\[ \hat{\sigma}^2 = \frac{\hat{s}^2}{\hat{\theta}_2}, \quad \hat{\theta}_1 = \hat{k} \hat{\sigma}^2, \quad \hat{\eta}_1 = \hat{\eta} \hat{\sigma}^2, \]

\[ \hat{\theta}_0 = -\hat{\lambda}_{1,1} + \left( \frac{\kappa^2 + \hat{\eta}^2}{4} + 2\pi^2 \right) \hat{\theta}_2, \quad \hat{\lambda}_{1,1} = \left( \frac{\hat{s}^2}{\hat{\sigma}_{1,1}^2} \right)^{1/\alpha}. \]

Let \( \vartheta_{-1} = (\vartheta_1^*, \eta_1^*, \theta_2^*, (\sigma^*)^2)^T, \)

\[
c_1 = (\lambda_{1,1}^*)^2 \left( \frac{\lambda_{1,2}^*}{\alpha} - \frac{\theta_0^*}{1 - \alpha} \right)^2 + (\lambda_{1,2}^*)^2 \left( \frac{\lambda_{1,1}^*}{\alpha} - \frac{\theta_0^*}{1 - \alpha} \right)^2,
\]

\[
c_2 = -\frac{1}{1 - \alpha} \left\{ (\lambda_{1,1}^*)^2 \left( \frac{\lambda_{1,2}^*}{\alpha} - \frac{\theta_0^*}{1 - \alpha} \right) + (\lambda_{1,2}^*)^2 \left( \frac{\lambda_{1,1}^*}{\alpha} - \frac{\theta_0^*}{1 - \alpha} \right) \right\},
\]

\[
c_3 = \frac{1}{(1 - \alpha)^2} \left\{ (\lambda_{1,1}^*)^2 + (\lambda_{1,2}^*)^2 \right\}
\]

\[
J = \frac{2}{9\pi^4(\theta_2^*)^2} \begin{bmatrix} c_1 & c_2(\vartheta_{-1}^*)^T \\ c_2 \vartheta_{-1} & c_3 \vartheta_{-1}(\vartheta_{-1}^*)^T \end{bmatrix},
\]

where T denotes the transpose.

**Theorem 2.** Suppose Assumption 1 holds.

1. Under \( \frac{n^{1-\epsilon}}{M_1^\alpha M_2^\alpha} \to 0 \) for some \( \epsilon < \alpha \), it holds that as \( n \to \infty, M_1 \to \infty \) and \( M_2 \to \infty \),

\[
(\hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \hat{\sigma}^2) \xrightarrow{p} (\theta_0^*, \theta_1^*, \eta_1^*, \theta_2^*, (\sigma^*)^2). \tag{12}
\]

2. Under \( \frac{n^{1-\epsilon}}{N^{1/(\alpha_1 + \alpha_2)}} \to 0 \) and \( \frac{n^{2-\epsilon}}{M_1^\alpha M_2^\alpha} \to 0 \) for some \( \epsilon < \alpha \), it holds that as \( n \to \infty, M_1 \to \infty \) and \( M_2 \to \infty \),

\[
\sqrt{n} \begin{pmatrix} \hat{\theta}_0 - \theta_0^* \\ \hat{\theta}_1 - \theta_1^* \\ \hat{\eta}_1 - \eta_1^* \\ \hat{\theta}_2 - \theta_2^* \\ \hat{\sigma}^2 - (\sigma^*)^2 \end{pmatrix} \xrightarrow{d} N(0, J). \tag{13}
\]

**Corollary 1.** Suppose Assumption 1 holds. Let \( M_1 = M_2, M = O(N^\beta), \beta > 0 \) and \( n = O(N^\gamma), 0 < \gamma \leq 1 \).

1. If \( \zeta < \alpha \beta \), then (12) holds.
2. If \( \zeta < \frac{\alpha \beta (1 - \alpha)}{2 - \alpha} \), then (13) holds.

**Remark 6.** (1) We can estimate \( \theta_0 \) because this parameter is contained in the noise. In general, \( \theta_0 \) can only be estimated through noise.

(2) We use two coordinate processes \( x_{1,1}^{Q_1}(t) \) and \( x_{1,2}^{Q_1}(t) \) when we estimate the coefficient parameters, but the choice of the coordinate processes is arbitrary.
Therefore, we can construct $K$ estimators $\hat{\theta}_{2,1}, \ldots, \hat{\theta}_{2,K}$ of $\theta_2$ using the coordinate processes $\{x_{k,1}^Q(t), x_{k,2}^Q(t)\}_{k=1}^K$. Since the coordinate processes $\{x_{k,\ell}^Q(t)\}_{k,\ell \geq 1}$ are independent, it holds that
\[
\sqrt{n}(\hat{\theta}_{2,1} - \theta_2^*, \ldots, \hat{\theta}_{2,K} - \theta_2^*) \xrightarrow{d} N(0, \Sigma_K).
\]
where \( \Sigma_K = \frac{2}{\eta_1} \text{diag}((\lambda_{k,1}^*)^2, (\lambda_{k,2}^*)^2) \). If there were a sequence \( \{b_{k,K}\} \) such that \( b_{k,K} \geq 0, \lim_{K \to \infty} \sum_{k=1}^K b_{k,K} = 1 \) and \( \lim_{K \to \infty} \sum_{k=1}^K b_{k,K}^2 (\lambda_{k,1}^*)^2 + (\lambda_{k,2}^*)^2 = 0 \), then the convergence rate could be improved by setting \( \hat{\theta}_2 = \sum_{k=1}^K b_{k,K} \hat{\theta}_{2,k} \) as the estimator of $\theta_2$. However, such a sequence \( \{b_{k,K}\} \) does not exist.

Indeed, by assuming its existence, it follows from the Schwarz inequality that
\[
\left( \sum_{k=1}^K b_{k,K} \right)^2 \leq \left( \sum_{k=1}^K b_{k,K}^2 (\lambda_{k,1}^*)^2 + (\lambda_{k,2}^*)^2 \right) \left( \sum_{k=1}^K \frac{1}{(\lambda_{k,1}^*)^2 + (\lambda_{k,2}^*)^2} \right),
\]
which yields together with \( \lim_{K \to \infty} (\lambda_{k,1}^*)^2 + (\lambda_{k,2}^*)^2/k^d \in (0, \infty) \) that \( 1 \leq 0 \) as \( K \to \infty \).

For this reason, we do not expect to improve the convergence rate of the estimators when the driving noise is the $Q_1$-Wiener process even if a large number of coordinate processes are used.

(3) For some $\zeta < \frac{1+2(1-a)}{2-a}$ in Corollary 1, in order to obtain asymptotic normality of the estimators, it is necessary to make $\beta$ larger than that required for consistency to hold. Moreover, if $\alpha$ is small, then $\beta$ concerned with the number of spatial observations $M$ has to be large, which implies that one needs a large $\beta$ to approximate the coordinate processes.

### 3.2 SPDE driven by $Q_2$-Wiener process

In this subsection, we consider the SPDE (1) driven by the $Q_2$-Wiener process defined as (5). In this case, as in Section 3.1, we first estimate the parameters $\theta_1/\theta_2$ and $\eta_1/\theta_2$ which appear in the eigenfunctions $e_{k,\ell}$, and then estimate the coefficient parameters using the approximate coordinate process.

In a similar way to Proposition 1, the following proposition holds.

**Proposition 2.** Under Assumption 1, it holds that uniformly in $(y, z) \in D_\delta$,
\[
\mathbb{E}[(\Delta_y X^Q_{k,1})^2(y, z)] = \Delta^a N \frac{\sigma^2 \Gamma(1-a)}{4\pi a} e^{-\frac{\eta_1}{\eta_2} y} e^{-\frac{\eta_2}{\eta_1} z} + r_{N,i} + O(\Delta^a N), \tag{14}
\]
where $\sum_{i=1}^N |r_{N,i}| = O(\Delta^a N)$, and thus
\[
\mathbb{E} \left[ \frac{1}{N^\Delta N} \sum_{i=1}^N (\Delta_y X^Q_{k,1})^2(y, z) \right] = \frac{\sigma^2 \Gamma(1-a)}{4\pi a} e^{-\frac{\eta_1}{\eta_2} y} e^{-\frac{\eta_2}{\eta_1} z} + O(\Delta^a N) + r_{N,i}.
\]

**Remark 7.** The only difference between (10) and (14) is the exponent of $\theta_1$ in the denominator. This is caused by the fact that the coefficients of $k^2 + \ell^2$ in $\lambda_{k,\ell}$ and $\mu_{k,\ell}$
are different, which are $\theta_2 \pi^2$ and $\pi^2$ respectively. See Lemmas C.1 and C.5 (in the Data S1) for details. Proposition 2 allows us to estimate $\sigma^2/\theta_2^{1-\alpha}$, $\theta_1/\theta_2$ and $\eta_1/\theta_2$ when the SPDE (1) is driven by the $Q_2$-Wiener process.

Let $S = \sigma^2/\theta_2^{1-\alpha}$, $\kappa = \theta_1/\theta_2$, and $\eta = \eta_1/\theta_2$, and let

$$U_{N,m}^{(2)}(S, \kappa, \eta) = \sum_{j_1=1}^{m_1} \sum_{j_2=1}^{m_2} \left( Z_{N}^{Q_2}(\tilde{q}_{j_1}, \tilde{z}_{j_2}) - \frac{\Gamma(1 - \alpha)}{4\pi \alpha} S e^{(i\tilde{q}_{j_1} + \tilde{q}_{j_2})} \right)^2,$$

which is the contrast function of $S$, $\kappa$ and $\eta$. Let $\tilde{S}$, $\tilde{\kappa}$ and $\tilde{\eta}$ be minimum contrast estimators defined as

$$(\tilde{S}, \tilde{\kappa}, \tilde{\eta}) = \arg\inf_{(S, \kappa, \eta) \in \Xi_2} U_{N,m}^{(2)}(S, \kappa, \eta),$$

where $\Xi_2$ is a compact convex subset of $(0, \infty) \times \mathbb{R}^2$, and we assume that the true value $(S^*, \kappa^*, \eta^*)$ belongs to Int $\Xi_2$.

**Theorem 3.** Under Assumption 1, it holds that as $N \to \infty$ and $m \to \infty$,

$$N^{1/2(1-\alpha)} \begin{pmatrix} \tilde{S} - S^* \\ \tilde{\kappa} - \kappa^* \\ \tilde{\eta} - \eta^* \end{pmatrix} = O_p(1).$$

We construct the following approximate coordinate process by using the estimators $\tilde{\kappa}$ and $\tilde{\eta}$.

$$\tilde{x}_{k,\ell}^{Q_2}(\tilde{t}_i) = \frac{2}{M} \sum_{j_1=1}^{M_1} \sum_{j_2=1}^{M_2} X_{Q_2}^{j_1, j_2}(y_{j_1}, z_{j_2}) \sin(\pi ky_{j_1}) \sin(\pi \ell z_{j_2}) e^{\hat{\kappa} y_{j_1}} e^{\hat{\eta} z_{j_2}}, \quad i = 1, \ldots, n.$$

Noting that the coordinate process $x_{k,\ell}^{Q_2}(t)$ is a diffusion process given by (8), we can estimate the volatility parameter $\tau_{k,\ell} := \sigma \mu_{k,\ell}^{-a/2}$.

If $\mu_0$ is known, then the estimator of $\sigma^2$ is defined as

$$\hat{\sigma}^2 = \mu_{1,1}^{-1} \sum_{i=1}^{n} \left( \tilde{x}_{Q_2}^{1,1}(\tilde{t}_i) - \tilde{x}_{Q_2}^{1,1}(\tilde{t}_{i-1}) \right)^2,$$

and the estimators of $\theta_2$, $\theta_1$ and $\eta_1$ are defined as

$$\hat{\theta}_2 = \left( \frac{\hat{\sigma}^2}{S} \right)^{\frac{1}{1-a}}, \quad \hat{\theta}_1 = \hat{\kappa} \hat{\theta}_2, \quad \hat{\eta}_1 = \hat{\eta} \hat{\theta}_2.$$

On the other hand, since

$$\mu_{1,1} = \left( \frac{\tau_{1,2}^2}{\tau_{1,1}^2} \right)^{1/a} \mu_{1,2}, \quad \mu_{1,2} = \mu_{1,1} + 3\pi^2,$$
\(\mu_{1,1}\) and \(\sigma^2\) can be expressed as
\[
\begin{align*}
\mu_{1,1} &= \frac{3\pi^2 (r_{1,2}^2 / r_{1,1}^2)^{1/\alpha}}{1 - (r_{1,2}^2 / r_{1,1}^2)^{1/\alpha}} = \frac{3\pi^2}{\tau_{1,1}^2 \tau_{1,2}} \left( \frac{1}{\tau_{1,2}^{2/\alpha}} - \frac{1}{\tau_{1,1}^{2/\alpha}} \right)^{-1}, \\
\sigma^2 &= \tau_{1,1}^2 \mu_{1,1}^2 = \left\{ 3\pi^2 \left( \frac{1}{\tau_{1,2}^{2/\alpha}} - \frac{1}{\tau_{1,1}^{2/\alpha}} \right)^{-1} \right\}^\alpha.
\end{align*}
\]

Therefore, if \(\mu_0\) is unknown, then the estimators of \(\tau_{k,\ell}^2\) and \(\sigma^2\) are defined as
\[
\tau_{k,\ell}^2 = \sum_{i=1}^{n} \left( x_{k,\ell}^Q_i(t_i) - x_{k,\ell}^Q_i(t_{i-1}) \right)^2,
\]
\[
\bar{\sigma}^2 = \left\{ 3\pi^2 \left( \frac{1}{(\tau_{1,2}^2)^{1/\alpha}} - \frac{1}{(\tau_{1,1}^2)^{1/\alpha}} \right)^{-1} \right\}^\alpha,
\]
and the estimators of \(\theta_2, \theta_1, \eta_1,\) and \(\mu_0\) are defined as
\[
\begin{align*}
\bar{\theta}_2 &= \left( \bar{\sigma}^2 \right)^{1/2}, \\
\bar{\theta}_1 &= \bar{k}\bar{\theta}_2, \\
\bar{\eta}_1 &= \bar{\eta}\bar{\theta}_2, \\
\bar{\mu}_0 &= \bar{\mu}_{1,1} - 2\pi^2 = \frac{3\pi^2}{(\tau_{1,1}^2)^{1/\alpha}} \left( \frac{1}{(\tau_{1,2}^2)^{1/\alpha}} - \frac{1}{(\tau_{1,1}^2)^{1/\alpha}} \right)^{-1} - 2\pi^2.
\end{align*}
\]

Let \(v_{-1}^* = (\theta_1^*, \eta_1^*, \theta_2^*, (1 - \alpha)(\sigma^*)^2)^T,\)
\[
d_1 = \frac{2(\mu_{1,1}^*)^2(\mu_{1,2}^*)^2}{\alpha^2}, \\
d_2 = \frac{\mu_{1,1}^* \mu_{1,2}^* (\mu_{1,1}^* + \mu_{1,2}^*)}{\alpha(1 - \alpha)}, \\
d_3 = \frac{(\mu_{1,1}^*)^2 + (\mu_{1,2}^*)^2}{(1 - \alpha)^2},
\]
\[
\mathcal{K} = \frac{2}{(1 - \alpha)^2} v_{-1}^*(v_{-1}^*)^T, \\
\mathcal{L} = \frac{2}{9\pi^4} \left( \begin{array}{ccc} d_1 & d_2 (v_{-1}^*)^T \\
 \end{array} \right),
\]

**Theorem 4.** Suppose Assumption 1 holds.

(a) Let \(\mu_0\) be known.

\(1\) Under \(\frac{n^{1+\epsilon}}{M_1^\alpha M_2^\alpha} \to 0\) for some \(\epsilon < \alpha\), it holds that as \(n \to \infty, M_1 \to \infty\) and \(M_2 \to \infty,\)
\[
(\bar{\theta}_1, \bar{\eta}_1, \bar{\theta}_2, \bar{\sigma}^2) \xrightarrow{p} (\theta_1^*, \eta_1^*, \theta_2^*, (\sigma^*)^2).
\]

\(2\) Under \(\frac{n^{1+\epsilon}}{M_1^\alpha M_2^\alpha} \to 0\) and \(\frac{n^{2+\epsilon}}{M_1^\alpha M_2^\alpha} \to 0\) for some \(\epsilon < \alpha\), it holds that as \(n \to \infty, M_1 \to \infty\) and \(M_2 \to \infty,\)
\[
\sqrt{n} \begin{pmatrix}
\bar{\theta}_1 - \theta_1^* \\
\bar{\eta}_1 - \eta_1^* \\
\bar{\theta}_2 - \theta_2^* \\
\bar{\sigma}^2 - (\sigma^*)^2
\end{pmatrix} \xrightarrow{d} N(0, \mathcal{K}).
\]
(b) Let \( \mu_0 \) be unknown.

1. Under \( \frac{n^{1-\epsilon}}{M_1^2 \wedge M_2^2} \to 0 \) for some \( \epsilon < \alpha \), it holds that as \( n \to \infty \), \( M_1 \to \infty \) and \( M_2 \to \infty \),

\[
(\bar{\mu}_0, \bar{\theta}_1, \bar{\eta}_1, \bar{\theta}_2, \bar{\sigma}^2) \xrightarrow{D} (\mu_0^*, \theta_1^*, \eta_1^*, \theta_2^*, (\sigma^*)^2).
\]

2. Under \( \frac{n^{1-\epsilon}}{N^{1/(3-\alpha)}} \to 0 \) and \( \frac{n^{1-\epsilon}}{M_1^2 \wedge M_2^2} \to 0 \) for some \( \epsilon < \alpha \), it holds that as \( n \to \infty \), \( M_1 \to \infty \) and \( M_2 \to \infty \),

\[
\sqrt{n} \begin{pmatrix}
\bar{\mu}_0 - \mu_0^* \\
\bar{\theta}_1 - \theta_1^* \\
\bar{\eta}_1 - \eta_1^* \\
\bar{\theta}_2 - \theta_2^* \\
\bar{\sigma}^2 - (\sigma^*)^2
\end{pmatrix} \xrightarrow{D} N(0, \mathcal{L}).
\]

Remark 8. (1) Since the cylindrical Brownian motion \( \{B_t\}_{t \geq 0} \) given by (9) is the noise without unknown parameters, the \( Q_2 \)-Wiener process with known \( \mu_0 \) can be regarded as equivalent to \( \{B_t\}_{t \geq 0} \) in the sense that the noise is known. Indeed, (a)-(2) in Theorem 4, especially for \( \alpha = 0.5 \), corresponds to the result of Kaino and Uchida (2020). On the other hand, when \( \mu_0 \) is unknown, the \( Q_2 \)-Wiener process can be regarded as equivalent to the \( Q_1 \)-Wiener process in the sense that the damping factor is unknown. Theorem 2 and (b) in Theorem 4 show that \( \theta_0 \) can be estimated for the SPDE (1) driven by the \( Q_1 \)-Wiener process, while \( \mu_0 \) can be estimated instead of \( \theta_0 \) for the SPDE (1) driven by the \( Q_2 \)-Wiener process.

2. Note that \( 1 < \frac{(\mu_0^*)^2 + (\mu_0^*)^2}{9\sigma^2} \). The variances of the estimators of \( \theta_1 \), \( \eta_1 \), \( \theta_2 \), and \( \sigma^2 \) when \( \mu_0 \) is unknown are \( \frac{2}{9} \left( \frac{\mu_0^*}{\sigma^2} + 2 \left( \frac{\mu_0^*}{\sigma^2} + 5 \right) \right) \) times larger than those of \( \theta_1 \), \( \eta_1 \), \( \theta_2 \) and \( \sigma^2 \) when \( \mu_0 \) is known, respectively. For example, if \( \mu_0 = 0 \), then the variances of the estimators when \( \mu_0 \) is unknown are \( \frac{20}{9} \approx 2.2 \) times larger than those when \( \mu_0 \) is known.

3. Unlike the discussion in Remark 6-(2), when the driving noise is the \( Q_2 \)-Wiener process with a known \( \mu_0 \), it is expected to improve the convergence rate of the estimators in (a)-(2) of Theorem 4 by using a larger number of coordinate processes \( \{x_{k,1}^Q(t), x_{k,2}^Q(t)\}_{k=1}^K \) and setting \( \tilde{\theta}^{(K)} = \frac{1}{K} \sum_{i=1}^K \tilde{\theta}_i \) with \( \tilde{\theta}_i = (\tilde{\theta}_{1,i}, \tilde{\eta}_{1,i}, \tilde{\theta}_{2,i}, (\tilde{\sigma}_i^2)^T) \). This is future work.

4 | SIMULATIONS

The numerical solution of the SPDE (1) is generated by

\[
\hat{X}_{t_i}^Q \left( y_{j_1}, z_{j_2} \right) = \sum_{k=1}^K \sum_{\ell \in \mathcal{L}} x_{k,\ell}^Q(t_i) e_{k,\ell}(y_{j_1}, z_{j_2}), \quad i = 1, \ldots, N, j_1 = 1, \ldots, M_1, j_2 = 1, \ldots, M_2.
\]

In this simulation, the true values of parameters \( (\theta^*_0, \theta^*_1, \eta^*_1, \theta^*_2, \sigma^*) = (0, 0.2, 0.2, 0.2, 1) \). We set that \( N = 10^3, M_1 = M_2 = 200, K = L = 10^4, \xi = 0, \alpha = 0.5, \lambda_{1,1}^* \approx 4.05 \). When \( N = 10^3, M_1 = M_2 = \)
TABLE 1 Simulation results of $\hat{s}$, $\hat{k}$, and $\hat{\eta}$.

|       | $\hat{s}$ | $\hat{k}$ | $\hat{\eta}$ |
|-------|-----------|-----------|--------------|
| True value | 5         | 1         | 1            |
| Mean   | 4.805     | 0.986     | 1.003        |
| SD     | (0.140)   | (0.039)   | (0.032)      |

TABLE 2 Simulation results of $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$.

|       | $\hat{\theta}_0$ | $\hat{\theta}_1$ | $\hat{\eta}_1$ | $\hat{\theta}_2$ | $\hat{\sigma}^2$ |
|-------|-------------------|-------------------|-----------------|-------------------|------------------|
| True value | 0                 | 0.2               | 0.2             | 0.2               | 1                |
| Mean   | −0.317            | 0.152             | 0.154           | 0.154             | 0.739            |
| SD     | (1.950)           | (0.077)           | (0.078)         | (0.079)           | (0.376)          |

TABLE 3 Simulation results of $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$.

|       | $\hat{\theta}_0$ | $\hat{\theta}_1$ | $\hat{\eta}_1$ | $\hat{\theta}_2$ | $\hat{\sigma}^2$ |
|-------|-------------------|-------------------|-----------------|-------------------|------------------|
| True value | 0                 | 0.2               | 0.2             | 0.2               | 1                |
| Mean   | −0.643            | 0.190             | 0.192           | 0.192             | 0.923            |
| SD     | (1.985)           | (0.074)           | (0.075)         | (0.075)           | (0.360)          |

200, the size of data is about 10 GB. We used R language to compute the estimators of Theorems 1 and 2. The computation time of (15) is directly proportional to $K \times L$. Therefore, the computation time for the numerical solution of the SPDE (1) is directly proportional to $N \times M_1 \times M_2 \times K \times L$. In the setting of this simulation, $N \times M_1 \times M_2 \times K \times L = 4 \times 10^{15}$. Three personal computers were used for this simulation, and it takes about 100 h to generate one sample path of the SPDE (1). The number of iteration is 200.

4.1 $m_1 = m_2 = 5$

First, we estimated $s = \sigma^2 / \theta_2$, $\kappa = \theta_1 / \theta_2$ and $\eta = \eta_1 / \theta_2$. Table 1 is the simulation results of the means and the SDs of $\hat{s}$, $\hat{k}$, and $\hat{\eta}$ with $(N, m_1, m_2, \alpha) = (10^3, 5, 5, 0.5)$. It seems from Table 1 that the biases of $\hat{k}$ and $\hat{\eta}$ are both very small and the result of $\hat{s}$ is not bad. The estimators of $s$, $\kappa$, and $\eta$ have good performances.

Next, we estimated $(\hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \hat{\sigma}^2)$. Table 2 is the simulation results of the means and the SDs of $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$ with $(N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 5, 5, 0.5, 0.499)$. In this case, $\frac{n^{1-\alpha \epsilon}}{M_1 \wedge M_2} \approx 0.25$. It seems from Table 2 that the biases of $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$ are not too large.

Table 3 is the simulation results of the means and the SDs of $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$ with $(N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 5, 5, 0.5, 0.499)$. In this case, $\frac{n^{1-\alpha \epsilon}}{M_1 \wedge M_2} \approx 0.50$. Table 3 shows that $\hat{\theta}_0$ has a bias, but $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$ have good performances.

Table 4 is the simulation results of the means and the SDs of $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$ with $(N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 5, 5, 0.5, 0.499)$. In this case, $\frac{n^{1-\alpha \epsilon}}{M_1 \wedge M_2} \approx 0.75$. By Table 4, it seems that the biases of $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$ are not too large and that $\hat{\theta}_0$ has a small bias.
Table 4 Simulation results of $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$.

|       | $\hat{\theta}_0$ | $\hat{\theta}_1$ | $\hat{\eta}_1$ | $\hat{\theta}_2$ | $\hat{\sigma}^2$ |
|-------|-------------------|-------------------|-----------------|-------------------|------------------|
| True value | 0                 | 0.2               | 0.2             | 0.2               | 1                |
| Mean   | $-0.581$          | 0.165             | 0.168           | 0.168             | 0.805            |
| SD     | (1.861)           | (0.058)           | (0.060)         | (0.060)           | (0.284)          |

Table 5 Simulation results of $\hat{s}$, $\hat{k}$, and $\hat{\eta}$.

|       | $\hat{s}$ | $\hat{k}$ | $\hat{\eta}$ |
|-------|-----------|-----------|--------------|
| True value | 5         | 1         | 1            |
| Mean   | 4.633     | 0.964     | 0.968        |
| SD     | (0.074)   | (0.024)   | (0.018)      |

Table 6 Simulation results of $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$.

|       | $\hat{\theta}_0$ | $\hat{\theta}_1$ | $\hat{\eta}_1$ | $\hat{\theta}_2$ | $\hat{\sigma}^2$ |
|-------|-------------------|-------------------|-----------------|-------------------|------------------|
| True value | 0                 | 0.2               | 0.2             | 0.2               | 1                |
| Mean   | $-0.873$          | 0.161             | 0.162           | 0.167             | 0.774            |
| SD     | (2.412)           | (0.077)           | (0.077)         | (0.081)           | (0.370)          |

Table 7 Simulation results of $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$.

|       | $\hat{\theta}_0$ | $\hat{\theta}_1$ | $\hat{\eta}_1$ | $\hat{\theta}_2$ | $\hat{\sigma}^2$ |
|-------|-------------------|-------------------|-----------------|-------------------|------------------|
| True value | 0                 | 0.2               | 0.2             | 0.2               | 1                |
| Mean   | $-0.212$          | 0.169             | 0.176           | 0.171             | 0.815            |
| SD     | (1.888)           | (0.065)           | (0.064)         | (0.067)           | (0.310)          |

From Tables 2–4, we can see that $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$ have good performances when $n = 100$ and $\hat{\theta}_0$ has a smallest bias when $n = 50$.

4.2 $m_1 = m_2 = 10$

Table 5 is the simulation results of the means and the SDs of $\hat{s}$, $\hat{k}$, and $\hat{\eta}$ with $(N, m_1, m_2, \alpha) = (10^3, 10, 10, 0.5)$. It seems from Table 5 that the biases of $\hat{k}$ and $\hat{\eta}$ are both very small and the result of $\hat{s}$ is not bad. The estimators of $s$, $\kappa$, and $\eta$ have good performances.

Table 6 is the simulation results of the means and the SDs of $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$ with $(N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 10, 10, 50, 0.5, 0.499)$. In this case, $\frac{\eta^{1-\alpha\epsilon}}{M_{1}^{\alpha}M_{2}^{\epsilon}} \approx 0.25$. It seems from Table 6 that the biases of $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$ are not too large.

Table 7 is the simulation results of the means and the SDs of $\hat{\theta}_0$, $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$ with $(N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 10, 10, 100, 0.5, 0.499)$. In this case, $\frac{\eta^{1-\alpha\epsilon}}{M_{1}^{\alpha}M_{2}^{\epsilon}} \approx 0.50$. Table 7 shows that $\hat{\theta}_0$ has a bias, but $\hat{\theta}_1$, $\hat{\eta}_1$, $\hat{\theta}_2$, and $\hat{\sigma}^2$ have good performances.
Table 8 is the simulation results of \( \hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \) and \( \hat{\sigma}^2 \) with \((N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 10, 10, 150, 0.5, 0.499)\). In this case, \( \frac{n^{1+\epsilon}}{M_1^7 \& M_2^5} \approx 0.75 \). By Table 8, it seems that the biases of \( \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \) and \( \hat{\sigma}^2 \) are not too large and that \( \hat{\theta}_0 \) has a small bias.

From Tables 6–8, we can see that \( \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \) and \( \hat{\sigma}^2 \) have good performances when \( n = 100 \) and \( \hat{\theta}_0 \) has a smallest bias when \( n = 100 \).

### 4.3 | \( m_1 = m_2 = 50 \)

Table 9 is the simulation results of the means and the SDs of \( \hat{s}, \hat{k}, \) and \( \hat{\eta} \) with \((N, m_1, m_2, \alpha) = (10^3, 50, 50, 0.5)\). It seems from Table 9 that the biases of \( \hat{k}, \hat{\eta} \) and \( \hat{s} \) are not too large.

Table 10 is the simulation results of the means and the SDs of \( \hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2 \) and \( \hat{\sigma}^2 \) with \((N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 50, 50, 50, 0.5, 0.499)\). In this case, \( \frac{n^{1+\epsilon}}{M_1^7 \& M_2^5} \approx 0.25 \). It seems from Table 10 that \( \hat{\theta}_0 \) has a bias, but the biases of \( \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \) and \( \hat{\sigma}^2 \) are not too large.

Table 11 is the simulation results of the means and the SDs of \( \hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2 \) and \( \hat{\sigma}^2 \) with \((N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 50, 50, 100, 0.5, 0.499)\). In this case, \( \frac{n^{1+\epsilon}}{M_1^7 \& M_2^5} \approx 0.50 \). Table 11 shows that \( \hat{\theta}_0 \) has a bias, but \( \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2 \) and \( \hat{\sigma}^2 \) have good performances.

Table 12 is the simulation results of the means and the SDs of \( \hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2 \) and \( \hat{\sigma}^2 \) with \((N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 50, 50, 150, 0.5, 0.499)\). In this case, \( \frac{n^{1+\epsilon}}{M_1^7 \& M_2^5} \approx 0.75 \). By Table 12, it seems that the biases of \( \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2 \) and \( \hat{\sigma}^2 \) are not too large and that \( \hat{\theta}_0 \) has a small bias.

From Tables 10–12, we can see that \( \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \) and \( \hat{\sigma}^2 \) have good performances when \( n = 100 \) and \( \hat{\theta}_0 \) has a smallest bias when \( n = 150 \).
TABLE 11 Simulation results of \( \hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \text{ and } \hat{\sigma}^2 \).

|       | \( \hat{\theta}_0 \) | \( \hat{\theta}_1 \) | \( \hat{\eta}_1 \) | \( \hat{\theta}_2 \) | \( \hat{\sigma}^2 \) |
|-------|------------------------|------------------------|------------------------|------------------------|------------------------|
| True value | 0                      | 0.2                    | 0.2                    | 0.2                    | 1                      |
| Mean   | -1.075                 | 0.194                  | 0.195                  | 0.214                  | 0.915                  |
| SD     | (2.278)                | (0.065)                | (0.065)                | (0.072)                | (0.306)                |

TABLE 12 Simulation results of \( \hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \text{ and } \hat{\sigma}^2 \).

|       | \( \hat{\theta}_0 \) | \( \hat{\theta}_1 \) | \( \hat{\eta}_1 \) | \( \hat{\theta}_2 \) | \( \hat{\sigma}^2 \) |
|-------|------------------------|------------------------|------------------------|------------------------|------------------------|
| True value | 0                      | 0.2                    | 0.2                    | 0.2                    | 1                      |
| Mean   | -0.761                 | 0.161                  | 0.161                  | 0.177                  | 0.756                  |
| SD     | (2.178)                | (0.058)                | (0.057)                | (0.063)                | (0.271)                |

TABLE 13 Simulation results of \( \hat{s}, \hat{\kappa}, \text{ and } \hat{\eta} \).

|       | \( \hat{s} \) | \( \hat{\kappa} \) | \( \hat{\eta} \) |
|-------|----------------|-----------------|-----------------|
| True value | 5                | 1               | 1               |
| Mean   | 4.881           | 0.995           | 0.993           |
| SD     | (0.074)         | (0.022)         | (0.021)         |

TABLE 14 Simulation results of \( \hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \text{ and } \hat{\sigma}^2 \).

|       | \( \hat{\theta}_0 \) | \( \hat{\theta}_1 \) | \( \hat{\eta}_1 \) | \( \hat{\theta}_2 \) | \( \hat{\sigma}^2 \) |
|-------|------------------------|------------------------|------------------------|------------------------|------------------------|
| True value | 0                      | 0.2                    | 0.2                    | 0.2                    | 1                      |
| Mean   | -0.744                 | 0.155                  | 0.155                  | 0.156                  | 0.760                  |
| SD     | (2.320)                | (0.077)                | (0.077)                | (0.078)                | (0.378)                |

4.4 \( m_1 = m_2 = 100 \)

Table 13 is the simulation results of the means and the SDs of \( \hat{s}, \hat{\kappa}, \text{ and } \hat{\eta} \) with \( (N, m_1, m_2, \alpha) = (10^3, 100, 100, 0.5) \). It seems from Table 13 that the biases of \( \hat{\kappa} \) and \( \hat{\eta} \) are both very small and the result of \( \hat{s} \) is not bad. The estimators of \( s, \kappa \) and \( \eta \) have good performances.

Table 14 is the simulation results of the means and the SDs of \( \hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \text{ and } \hat{\sigma}^2 \) with \( (N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 100, 100, 50, 0.5, 0.499) \). In this case, \( \frac{n^{1.499}}{M_1^\alpha M_2} \approx 0.25 \). It seems from Table 14 that the biases of \( \hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \text{ and } \hat{\sigma}^2 \) are not too large.

Table 15 is the simulation results of the means and the SDs of \( \hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \text{ and } \hat{\sigma}^2 \) with \( (N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 100, 100, 100, 0.5, 0.499) \). In this case, \( \frac{n^{1.499}}{M_1^\alpha M_2} \approx 0.50 \). Table 15 shows that the bias of \( \hat{\theta}_0 \) is not too large, and \( \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \text{ and } \hat{\sigma}^2 \) have good performances.

Table 16 is the simulation results of the means and the SDs of \( \hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2, \text{ and } \hat{\sigma}^2 \) with \( (N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 100, 100, 150, 0.5, 0.499) \). In this case, \( \frac{n^{1.499}}{M_1^\alpha M_2} \approx 0.75 \). By Table 16, it seems that the biases of \( \hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2 \) and \( \hat{\sigma}^2 \) are not too large.
TABLE 15  Simulation results of $\hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2$, and $\hat{\sigma}^2$.

|           | $\hat{\theta}_0$ | $\hat{\theta}_1$ | $\hat{\eta}_1$ | $\hat{\theta}_2$ | $\hat{\sigma}^2$ |
|-----------|------------------|------------------|----------------|------------------|------------------|
| True value| 0                | 0.2              | 0.2            | 0.2              | 1                |
| Mean      | $-0.774$         | 0.186            | 0.185          | 0.187            | 0.910            |
| SD        | (2.027)          | (0.066)          | (0.067)        | (0.067)          | (0.327)          |

TABLE 16  Simulation results of $\hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2$, and $\hat{\sigma}^2$.

|           | $\hat{\theta}_0$ | $\hat{\theta}_1$ | $\hat{\eta}_1$ | $\hat{\theta}_2$ | $\hat{\sigma}^2$ |
|-----------|------------------|------------------|----------------|------------------|------------------|
| True value| 0                | 0.2              | 0.2            | 0.2              | 1                |
| Mean      | $-0.606$         | 0.158            | 0.158          | 0.159            | 0.774            |
| SD        | (1.945)          | (0.057)          | (0.057)        | (0.057)          | (0.279)          |

TABLE 17  Simulation results of $\hat{s}, \hat{\kappa}$, and $\hat{\eta}$.

|       | $\hat{s}$ | $\hat{\kappa}$ | $\hat{\eta}$ |
|-------|-----------|----------------|--------------|
| True value | 5        | 1              | 1            |
| Mean   | 4.557     | 0.957          | 0.957        |
| SD     | (0.018)   | (0.006)        | (0.006)      |

TABLE 18  Simulation results of $\hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2$, and $\hat{\sigma}^2$.

|           | $\hat{\theta}_0$ | $\hat{\theta}_1$ | $\hat{\eta}_1$ | $\hat{\theta}_2$ | $\hat{\sigma}^2$ |
|-----------|------------------|------------------|----------------|------------------|------------------|
| True value| 0                | 0.2              | 0.2            | 0.2              | 1                |
| Mean      | $-0.917$         | 0.164            | 0.164          | 0.171            | 0.781            |
| SD        | (2.453)          | (0.079)          | (0.082)        | (0.079)          | (0.376)          |

From Tables 14–16, we can see that $\hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2$ and $\hat{\sigma}^2$ have good performances when $n = 100$ and $\hat{\theta}_0$ has a smallest bias when $n = 150$.

4.5  $m_1 = m_2 = 196$

Table 17 is the simulation results of the means and the SDs of $\hat{s}, \hat{\kappa}$, and $\hat{\eta}$ with $(N, m_1, m_2, \alpha) = (10^3, 196, 196, 0.5)$. It seems from Table 17 that the biases of $\hat{\kappa}$ and $\hat{\eta}$ are both very small and the result of $\hat{s}$ is not bad. The estimators of $s, \kappa$, and $\eta$ have good performances.

Table 18 is the simulation results of the means and the SDs of $\hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2$, and $\hat{\sigma}^2$ with $(N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 196, 196, 100, 0.5, 0.499)$. In this case, $\frac{\mu_{1-\alpha\epsilon}}{M_{1}^{\mu_{1}}M_{\eta}^{\mu_{2}}} \approx 0.25$. It seems from Table 18 that $\hat{\theta}_0$ has a bias, but the biases of $\hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2$, and $\hat{\sigma}^2$ are not too large.

Table 19 is the simulation results of the means and the SDs of $\hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2$, and $\hat{\sigma}^2$ with $(N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 196, 196, 100, 0.5, 0.499)$. In this case, $\frac{\mu_{1-\alpha\epsilon}}{M_{1}^{\mu_{1}}M_{\eta}^{\mu_{2}}} \approx 0.50$. Table 19 shows that $\hat{\theta}_0$ has a bias, but $\hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2$ and $\hat{\sigma}^2$ have very good performances.
Table 19 is the simulation results of the means and the SDs of $\hat{\theta}_0, \hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2,$ and $\hat{\sigma}^2$ with $(N, m_1, m_2, n, \alpha, \epsilon) = (10^3, 196, 196, 150, 0.5, 0.499)$. In this case, $\frac{m_1^{1-\epsilon}}{M_2^{\alpha}M_2^{2\epsilon}} \approx 0.75$. By Table 20, it seems that the biases of $\hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2,$ and $\hat{\sigma}^2$ are not too large and that $\hat{\theta}_0$ has a small bias.

From Tables 18–20, we can see that $\hat{\theta}_1, \hat{\eta}_1, \hat{\theta}_2,$ and $\hat{\sigma}^2$ have good performances when $n = 100$ and $\hat{\theta}_0$ has a smallest bias when $n = 150$.

ACKNOWLEDGMENTS
The authors thank the associate editor and the two referees for careful reading of the manuscript and constructive comments which led to significant improvements of the paper. This work was partially supported by JST CREST Grant Number JPMJCR2115, JSPS KAKENHI Grant Number JP17H01100, MEXT Grant Number JPJ010217, and ISM Cooperative Research Program (2022-ISMCRP-1013).

ORCID
Yozo Tonaki https://orcid.org/0000-0002-4554-4701
Yusuke Kaino https://orcid.org/0000-0002-0219-5636

REFERENCES
Altmeyer, R., Bretschneider, T., Janák, J., & Reiß, M. (2022). Parameter estimation in an SPDE model for cell repolarization. SIAM/ASA Journal on Uncertainty Quantification, 10(1), 179–199.
Bibinger, M., & Trabs, M. (2020). Volatility estimation for stochastic PDEs using high-frequency observations. Stochastic Processes and Their Applications, 130(5), 3005–3052.
Chong, C. (2019). High-frequency analysis of parabolic stochastic PDEs with multiplicative noise: Part I. arXiv:1908.04145.
Chong, C. (2020). High-frequency analysis of parabolic stochastic PDEs. The Annals of Statistics, 48(2), 1143–1167.
Chong, C., & Dalang, R. C. (2020). Power variations in fractional Sobolev spaces for a class of parabolic stochastic PDEs. arXiv:2006.15817.
Cialenco, I. (2018). Statistical inference for SPDEs: An overview. Statistical Inference for Stochastic Processes, 21(2), 309–329.
Cialenco, I., Delgado-Vences, F., & Kim, H. J. (2020). Drift estimation for discretely sampled SPDEs. Stochastics and Partial Differential Equations: Analysis and Computations, 8, 895–920.
Cialenco, I., & Glatt-Holtz, N. (2011). Parameter estimation for the stochastically perturbed Navier-Stokes equations. Stochastic Processes and Their Applications, 121(4), 701–724.
Cialenco, I., & Huang, Y. (2020). A note on parameter estimation for discretely sampled SPDEs. *Stochastics and Dynamics*, 20(3), 2050016.

Da Prato, G., & Zabczyk, J. (2014). *Stochastic equations in infinite dimensions*. In *Encyclopedia of mathematics and its applications* (2nd ed.). Cambridge University Press.

Genon-Catalot, V., & Jacod, J. (1993). On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Annales de l’Institut Henri Poincaré Probabilités et Statistiques*, 29(1), 119–151.

Hildebrandt, F., & Trabs, M. (2021a). Nonparametric calibration for stochastic reaction-diffusion equations based on discrete observations. *arXiv:2102.13415*.

Hildebrandt, F., & Trabs, M. (2021b). Parameter estimation for SPDEs based on discrete observations in time and space. *Electronic Journal of Statistics*, 15(1), 2716–2776.

Hübner, M., Khasminskii, R., & Rozovskii, B. L. (1993). *Two examples of parameter estimation for stochastic partial differential equations*, 149–160. Springer.

Kaino, Y., & Uchida, M. (2020). Parametric estimation for a parabolic linear SPDE model based on discrete observations. *Journal of Statistical Planning and Inference*, 211, 190–220.

Kaino, Y., & Uchida, M. (2021). Adaptive estimator for a parabolic linear SPDE with a small noise. *Japanese Journal of Statistics and Data Science*, 4, 513–541.

Kessler, M. (1997). Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics*, 24(2), 211–229.

Lord, G. J., Powell, C. E., & Shardlow, T. (2014). *An introduction to computational stochastic PDEs* Cambridge texts in applied mathematics. Cambridge University Press.

Lototsky, S. V. (2009). Statistical inference for stochastic parabolic equations: A spectral approach. *Publicaciones Matemàtiques*, 53(1), 3–45.

Lototsky, S. V., & Rozovsky, B. L. (2017). *Stochastic partial differential equations*. Springer.

Markussen, B. (2003). Likelihood inference for a discretely observed stochastic partial differential equation. *Bernoulli*, 9, 745–762.

Piterbarg, L., & Ostrovskii, A. (1997). *Advection and diffusion in random media: Implications for sea surface temperature anomalies*. Springer Science & Business Media.

Uchida, M., & Yoshida, N. (2012). Adaptive estimation of an ergodic diffusion process based on sampled data. *Stochastic Processes and their Applications*, 122(8), 2885–2924.

Uchida, M., & Yoshida, N. (2013). Quasi likelihood analysis of volatility and nondegeneracy of statistical random field. *Stochastic Processes and their Applications*, 123(7), 2851–2876.

Walsh, J. B. (1986). *An introduction to stochastic partial differential equations*. In P. L. Hennequin (Ed.), *École d’Été de probabilités de saint flour XIV-1984* (pp. 265–439). Springer.

**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Tonaki, Y., Kaino, Y., & Uchida, M. (2023). Parameter estimation for linear parabolic SPDEs in two space dimensions based on high frequency data. *Scandinavian Journal of Statistics*, 50(4), 1568–1589. [https://doi.org/10.1111/sjos.12663](https://doi.org/10.1111/sjos.12663)