Schrödinger principal-component analysis: On the duality between principal-component analysis and the Schrödinger equation

Ziming Liu,[1,*] Sitian Qian [ORCID],[2] Yixuan Wang,[3,†] Yuxuan Yan [ORCID],[2] and Tianyi Yang[2]

[1]*Department of Physics, Massachusetts Institute of Technology, Cambridge Massachusetts 02139, USA*
[2]*School of Physics, Peking University, Beijing 100871, China*
[3]*Applied and Computational Mathematics, California Institute of Technology, Pasadena, California 91125, USA*

Principal component analysis (PCA) has been applied to analyze random fields in various scientific disciplines. However, the explainability of PCA remains elusive unless strong domain-specific knowledge is available. This paper provides a theoretical framework that builds a duality between the PCA eigenmodes of a random field and eigenstates of a Schrödinger equation. Based on the duality we propose the *Schrödinger PCA* algorithm to replace the expensive PCA solver with a more sample-efficient Schrödinger equation solver. We verify the validity of the theory and the effectiveness of the algorithm with numerical experiments.

I. INTRODUCTION

Random fields are prevalent and important in many scientific disciplines, e.g., cosmology [1], high energy physics [2], fluid dynamics [3], and material science [4,5]. Since fields values are usually correlated with neighboring points, a low-dimensional description of random fields is not only possible, but also more computationally efficient. One standard method for dimensionality reduction [6–13] is principal component analysis (PCA), which attempts to diagonalize the correlation matrix via an orthogonal transformation, whose column vectors are known as eigenmodes.

A random field can be characterized by a two-point covariance function $C(\mathbf{x}, \mathbf{y})$, so the question is: How is $C(\mathbf{x}, \mathbf{y})$ related to the eigenmodes $\phi_i(\mathbf{x})$? The answer is the Karhunen–Loève expansion [14], an established theory in the statistics community. The aim of this paper is to answer this question in the taste of physics by revealing the fact that PCA eigenmodes of random field configurations are (approximately) equivalent to eigenstates of a steady-state Schrödinger equation. The aspects of such duality is summarized in Table I.

For example, we study a toy 2D random field shown in Fig. 1. Many realizations of the random field are drawn from the probabilistic distribution determined by $C(\mathbf{x}, \mathbf{y})$. PCA eigenmodes of these realizations have interesting structures: the eigenmodes look like quantum wave functions of a 2D harmonic oscillator. In this paper, we argue that this is not a mere coincidence but rather a direct consequence of the duality between the PCA problem and the quantum steady-state problem. Definitions of these quantities in the table will be made precise in the next section.

Besides offering a phyiscal perspective of the well-known PCA algorithm, we can also achieve a significant computational speedup benefiting from the duality. We can solve a PCA problem with a Schrödinger solver (i.e., elliptic equation solver), which is the main idea of our proposed Schrödinger PCA algorithm. Numerical experiments show that Schrödinger PCA requires $100\times$ fewer anchor points to accurately recover eigenvalues and eigenvectors compared to PCA. In the literature, there have been various methods proposed to accelerate the PCA algorithm, including Hilbert space methods exploiting the structure of RKHS [15] and stochastic PDE methods that use solutions of SPDE to approximate Gaussian rathom fields [16]. Our Schrödinger PCA method instead studies the case where the variance function could be as general as position-dependent, and the correlation length is relatively small, rendering previous methods inapplicable or inefficient. We notice that Refs. [17,18] are the PCA variants closest to our goals, but they do not point out the duality nor take advantage of the elliptic partial differential equation solvers. Although connections between convolutions and differential operators have been well studied in the mathematics literature [19–21], we provide a particular case where an intuitive physics picture is available.

II. METHOD

A. Notations and backgrounds

1. Random field

A random field is a stochastic process where $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ is a stochastic function, which we refer to as *field* below. Two-point covariance $C(\mathbf{x}, \mathbf{y})$, one-point variance $A(\mathbf{x})$, and two-point correlation $K(\mathbf{x}, \mathbf{y})$ of the random field are defined as

$$C(\mathbf{x}, \mathbf{y}) \equiv \langle \phi(\mathbf{x})\phi(\mathbf{y}) \rangle,$$
$$A(\mathbf{x}) \equiv C(\mathbf{x}, \mathbf{x}) = \langle \phi(\mathbf{x})\phi(\mathbf{x}) \rangle, \tag{1}$$
$$K(\mathbf{x}, \mathbf{y}) \equiv C(\mathbf{x}, \mathbf{y})/\sqrt{A(\mathbf{x})A(\mathbf{y})},$$

*zmliu@mit.edu
†roywang@caltech.edu

TABLE I. The duality between the PCA problem and the quantum steady-state problem.

PCA	Schrödinger
Variance $A(\mathbf{x})$	Potential $V(\mathbf{x})$
Correlation Kernel $\mathbf{\Sigma}(\mathbf{x})$	Inverse mass $\mathbf{\Sigma}_m(\mathbf{x})$
Eigenvalue λ_i	Eigen-energy E_i
Eigenvector ϕ_i	Eigenstate ψ_i

where $\langle \cdots \rangle$ means averaging over an ensemble of realizations of the random field or integral over the probability distribution:

$$P[\phi(\mathbf{x}_1) = v_1, \ldots, \phi(\mathbf{x}_n) = v_n]$$

$$\propto \exp\left[- \sum_{ij} C(\mathbf{x}_i, \mathbf{x}_j) v_i v_j \right]. \quad (2)$$

Usually $K(\mathbf{x}, \mathbf{y}) = f(|\mathbf{x} - \mathbf{y}|)$, where $|\cdot|$ can be an arbitrary norm, and f is a monotonically decreasing function that approaches to 0 as $|\mathbf{x} - \mathbf{y}| \to \infty$ due to locality. To make things concrete, we focus on Gaussian random fields,[1] where

$$K(\mathbf{x}, \mathbf{y}) = \exp\left[- \tfrac{1}{2} (\mathbf{y} - \mathbf{x})^T \mathbf{\Sigma}^{-1} (\mathbf{y} - \mathbf{x}) \right] \quad (3)$$

Without loss of generality, we have assumed $\langle \phi(x) \rangle = 0$.[2] In the special case where $\mathbf{\Sigma} = \sigma^2 \mathbf{I}$, σ characterizes the correlation length of the Gaussian random field. The correlation length can originate from the smoothness of functions or from physical entities, e.g., the width of a particle [22]. A large class of spatial data (including but not limited to physics) can be modeled with Gaussian random fields [1,2,23–25], in that Gaussian random fields capture two key features: fluctuations and (local) correlations. In numerical experiments, Gaussian random fields are efficiently generated by smearing out completely uncorrelated random fields with a Gaussian filter (see Appendix D).

2. PCA

PCA plays an important role in dimension reduction, pattern recognition, and partial differential equations (PDE) [26–30]. The key idea of PCA is to diagonalize the correlation matrix via an orthogonal transformation. For the random fields described by Eq. (1), eigenmodes $\phi_i(\mathbf{x})(i = 1, 2, \cdots)$ should satisfy the eigen-equations and orthogonality constraints:

$$\int_{\mathbf{y}} d\mathbf{y} C(\mathbf{x}, \mathbf{y}) \phi_i(\mathbf{y}) = \lambda_i \phi_i(\mathbf{x}), \quad \int_{\mathbf{x}} d\mathbf{x} \phi_i(\mathbf{x}) \phi_j(\mathbf{x}) = \delta_{ij}, \quad (4)$$

where eigenvalues λ_i are ordered as $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant 0$. The nonnegative eigenvalues and existence of orthogonal eigenmodes are due to the positive definiteness of $C(\mathbf{x}, \mathbf{y})$. Eigen-decomposition of random field is also known is Karhunen–Loève expansion in the statistics literature [14].

---

[1]Our analysis can also extend to laplace-type correlation and possibly many other tailed correlations (please see Appendix B).

[2]Principal component analysis will remove the mean value before solving the eigen-problem.

3. Schrödinger equation

The Schrödinger equation (SE) [31] uses a wave function $\psi(\mathbf{x})$ ($\mathbf{x} \in \mathbb{R}^d$) to characterize a quantum particle. The steady-state SE describes a particle in a potential field $V(\mathbf{x})$ with constant energy E:

$$\hat{H}\psi(\mathbf{x}) = -\frac{\mathbf{\Sigma}_m(\mathbf{x})}{2} \Delta\psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}) = E\psi(\mathbf{x}), \quad (5)$$

where the positive definite matrix $\mathbf{\Sigma}_m \in \mathbb{R}^{d \times d}$ is the inverse mass matrix of the particle, Δ is the Laplacian operator: $\Delta = \nabla^2 = \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2}$. The first term $-\frac{\mathbf{\Sigma}_m(\mathbf{x})}{2} \Delta\psi(\mathbf{x})$ and the second term $V(\mathbf{x})\psi(\mathbf{x})$ correspond to the kinetic and potential energy of a particle, respectively. Because $\hat{H}$ is an Hermitian operator, we have a complete set of real eigenstates $\psi_i(\mathbf{x})$ that satisfy

$$\hat{H}\psi_i(\mathbf{x}) = E_i\psi_i(\mathbf{x}), \quad \int_{\mathbf{x}} d\mathbf{x}\psi_i(\mathbf{x})\psi_j(\mathbf{x}) = \delta_{ij}, \quad (6)$$

where $\psi_i(\mathbf{x})$ and E_i are the ith *eigenstate* and *eigen-energy*, respectively.

B. The duality between PCA and Schrödinger equation

We shall build the correspondence between PCA and Schrödinger equation under some mild assumptions. Comparing Eqs. (4) and (6), similar structures can be observed, and intuitively the eigenmodes with the largest eigenvalues of the PCA problem could be mapped one-by-one to the first few lowest energy states of the Schrödinger equation. We aim to establish the duality in Theorem 1.

Theorem 1. (Informal) The PCA problem of Eq. (4) can be approximated by the Schrödinger problem of Eq. (6), i.e., two systems have the same equation hence same eigenvalues and eigenmodes up to a (possible) minus sign,[3] up to a second-order approximation provided that these quantities are equal:

$$\phi(\mathbf{x}) \iff \psi(\mathbf{x}),$$
$$-A(\mathbf{x}) \iff V(\mathbf{x}), \quad (7)$$
$$A(\mathbf{x})\mathbf{\Sigma}(\mathbf{x}) \iff \mathbf{\Sigma}_m(\mathbf{x}).$$

This approximation is valid for the first few eigenmodes, but not for higher eigenmodes.

We summarize the key idea of the proof here. Formal formulation of Theorem 1 and proof details can be found in Appendix A. We characterize a general operator $\mathbf{M}$ by its Rayleigh quotient (Ω is the integral domain):

$$R(\mathbf{M}, \phi) = \frac{\int_{\mathbf{x} \in \Omega} d\mathbf{x} \phi(\mathbf{x}) \mathbf{M} \phi(\mathbf{x})}{\int_{\mathbf{x} \in \Omega} d\mathbf{x} \phi^2(\mathbf{x})}$$

$$= \int_{\mathbf{x} \in \Omega} d\mathbf{x} \phi(\mathbf{x}) \mathbf{M} \phi(\mathbf{x}), \quad (8)$$

$$\left[||\phi||^2 = \int_{\mathbf{x} \in \Omega} \phi^2(\mathbf{x}) d\mathbf{x} = 1 \right].$$

---

[3]The eigenvalues of PCA and SE match each other by a minus sign. In terms of eigenmodes, there is a possible minus sign as in most eigen-problems: ψ_i is an eigenstate if and only if $-\psi_i$ is an eigenstate.
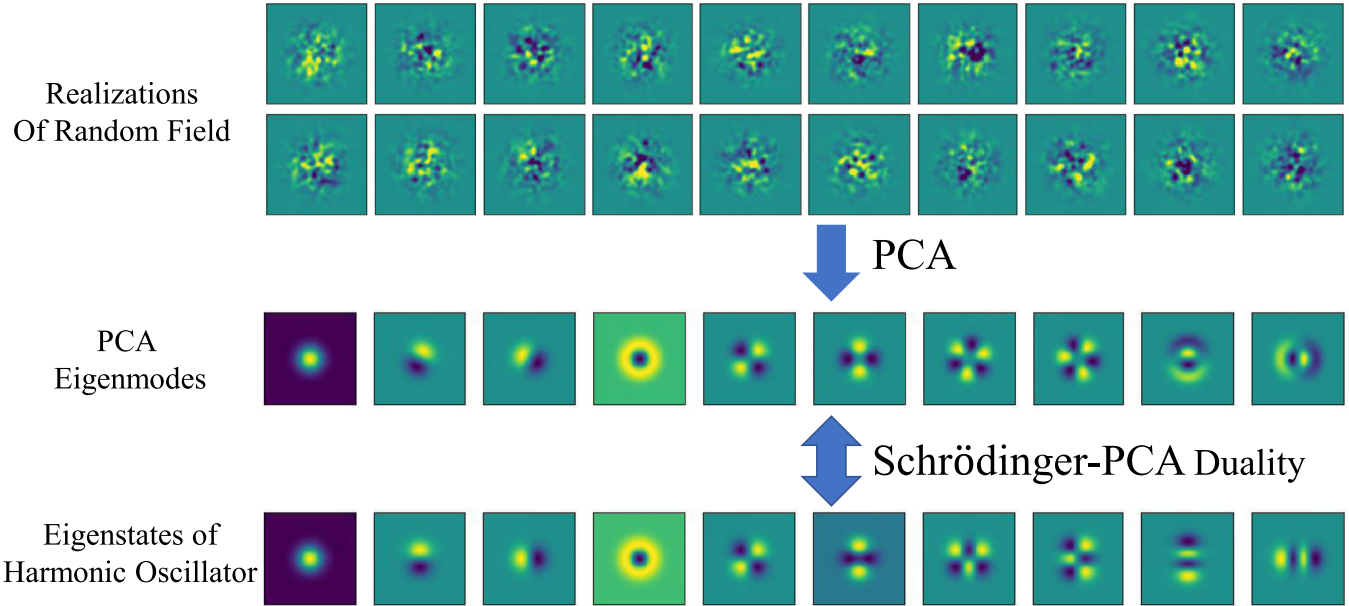
FIG. 1. PCA eigenmodes of random field configurations share much similarity with eigenstates of a 2D quantum harmonic oscillator.

The Rayleigh quotient of PCA and SE are defined as $R_{PCA}(C, \phi)$ and $R_{SE}(\hat{H}, \psi)$, respectively:

$$R_{PCA}(C, \phi) = \int_{\mathbf{x}\in\Omega, \mathbf{y}\in\Omega} d\mathbf{x} d\mathbf{y} \phi(\mathbf{x}) C(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y})$$

$$= \int_{\mathbf{x}\in\Omega} d\mathbf{x} \phi(\mathbf{x}) \left\{ \int_{\mathbf{y}\in\Omega} d\mathbf{y} C(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) \right\}, \quad (9)$$

$$R_{SE}(\hat{H}, \psi) = \int_{\mathbf{x}\in\Omega} d\mathbf{x} \psi(\mathbf{x}) \hat{H} \psi(\mathbf{x})$$

$$= \int_{\mathbf{x}\in\Omega} d\mathbf{x} \psi(\mathbf{x}) \left\{ \left[-\frac{\mathbf{\Sigma}_m}{2} \Delta + V(\mathbf{x}) \right] \psi(\mathbf{x}) \right\}. \quad (10)$$

The only thing left is to prove the (approximate) equivalence of $\{\cdots\}$ in Eqs. (9) and (10) by invoking Guassian integrals and integration by parts (details in Appendix A). To sum up, the approximation is valid for lower eigenmodes in practice, but we do not know *a priori* how many eigenmodes can be well approximated. Rather, the formal version of Theorem 1 in Appendix A sheds some light on how to post-process and verify *a posteriori* that the approximation is valid.

Although highly speculative, this duality between PCA and the quantum steady-state problem might shed light on deep theoretical links between quantum mechanics (Schrödinger) and information compression (PCA), which will be explored in future works.

C. Discussions

1. Implications of the Duality

The duality also explains scientists' conventions when they analyze a signal/field by decomposing it into a combination of hand-designed modes, e.g., Fourier bases, spherical harmonics, Bessel functions, etc. We argue that scientists are implicitly doing some optimal compression (PCA) when choosing the hand-designed modes. For simplicity, we con-

sider a uniform random field such that $\mathbf{\Sigma}(\mathbf{x})$ and $A(\mathbf{x})$ are independent of $\mathbf{x}$. These conditions are then translated to uniform potential energy $V(\mathbf{x}) = V_0$ [WLOG $V(\mathbf{x}) = 0$] and constant mass in the Schrödinger equation. Consequently, the Schrödinger equation describes a free particle on the domain Ω, where both the random field and Schrödinger equation are defined on.

Suppose Ω is a sphere (2D). On the one hand, eigenstates of a free particle on a sphere are spherical harmonics, i.e., eigenstates of the angular momentum operator. On the other hand, in cosmology, spherical harmonics are leveraged to analyze the (approximately) isotropic cosmic microwave background (CMB) [32]. We speculate this is not coincident but a direct consequence of the duality. If we humans are living in a universe where CMB is largely anisotropic, then the spherical harmonics are no longer "optimal" in the sense of PCA (see the anisotropic example in Sec. IV B). Similarly, a particle trapped in a deep circular well (2D) has Bessel functions as eigenstates, a free particle in a periodic box has plane wave solutions (i.e., Fourier). These basis functions are used in waveguide design and fluid dynamics, respectively. We summarize the relations in Table II.

2. No quantum tunneling for PCA

Quantum mechanics is distinguished from classical mechanics by its uncertainty principle, leading to the intriguing phenomenon of quantum tunneling. When a particle (with state-independent mass) is trapped in a finite deep poten-

TABLE II. "Optimal" bases for different problems.

Domain Ω	Eigenstates/Bases	Physics Example
Sphere (2D)	Spherical harmonics	CMB in cosmology
Circular well (2D)	Bessel function	Circular waveguide
Periodic box (nD)	Fourier bases	Fluid dynamics

Algorithm 1. Principal Component Analysis for Mode Decomposition.

Input: Data points $\mathbf{X} = \{\mathbf{x}_i \in \mathbb{R}^d\}$ and corresponding field
value $\mathbf{\Phi} = \{\phi(\mathbf{x}_i, t) \in \mathbb{R}\}$
$(i = 1, \dots, m; t = 1, \dots, n)$ where i labels different
anchor points in space, and t refers to different time of
measurements or realizations. We have removed the
mean value of fields such that $\frac{1}{n} \sum_{t=1}^{n} \phi(\mathbf{x}_i, t) = 0$ for
all i.
(1) Compute *covariance* matrix
$\mathbf{C}_{ij} = \text{Cov}(\phi(\mathbf{x}_i), \phi(\mathbf{x}_j)) = \frac{1}{n} \sum_{t=1}^{n} \phi(\mathbf{x}_i, t) \phi(\mathbf{x}_j, t)$;
(2) Diagonalization of $\mathbf{C}$ such that $\mathbf{C} = \sum_k \lambda_k \phi_k \phi_k^T$ where λ_k
and ϕ_k are the kth eigenvalue and eigenmode;
(3) Interpolation of finite-dimensional eigenmode
$\phi_k \in \mathbb{R}^m (k = 1, 2, \cdots)$ to a function $\phi_k(\mathbf{x})$ that satisfies
$\phi_k(\mathbf{x}_l) = \phi_{k,l}$ where $\phi_{k,l}$ is the lth entry of ϕ_k which
corresponds to the field value at $\mathbf{x}_l$ in mode ϕ_k;
Output: $\{\lambda_k, \phi_k(\mathbf{x})\}(k = 1, 2, \cdots)$

Algorithm 2. Schrödinger Principal Component Analysis for Mode Decomposition.

Input: Data points $\mathbf{X} = \{\mathbf{x}_i \in \mathbb{R}^d\}$, field value
$\mathbf{\Phi} = \{\phi(\mathbf{x}_i, t) \in \mathbb{R}\}$ and estimated correlation kernel
$\mathbf{\Sigma} = \{\mathbf{\Sigma}(\mathbf{x}_i) \in \mathbb{R}^{d \times d}\}$
$(i = 1, \dots, m; t = 1, \dots, N)$
(1) Compute *variance* vector: $A(\mathbf{x}_i) = \frac{1}{N} \sum_{t=1}^{N} |\phi(\mathbf{x}_i, t)|^2$;
(2) Define corresponding potential vector as $V_i = -A(\mathbf{x}_i)$ and
interpolation of $V \in \mathbb{R}^m$ to a function $V(\mathbf{x})$ satisfying
$V(\mathbf{x}_i) = V_i$;
(3) Define corresponding inverse mass matrix as
$\mathbf{\Sigma}_{m,i} = A(\mathbf{x}_i)\mathbf{\Sigma}(\mathbf{x}_i)$ and interpolation of $\mathbf{\Sigma}_m \in \mathbb{R}^{d \times d}$ to a
matrix function $\mathbf{\Sigma}_m(\mathbf{x})$ satisfying $\mathbf{\Sigma}_m(\mathbf{x}_i) = \mathbf{\Sigma}_{m,i}$;
(4) Solve the Schrödinger equation
$-\frac{\mathbf{\Sigma}_m(\mathbf{x})}{2}\Delta \psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}) = E\psi(\mathbf{x})$ and obtain
eigenstates $\psi_k(\mathbf{x})$ and corresponding energies E_k;
Output: $\{E_k, \psi_k(\mathbf{x})\}(k = 1, 2, \cdots)$

---

tial well, i.e., $V(\mathbf{x}) = -V_0 < 0 (\mathbf{x} \in \Omega)$ and $V(\mathbf{x}) = 0 (\mathbf{x} \notin \Omega)$, the wave function of the particle is nonzero outside Ω even when the particle has negative energy $-V_0 < E < 0$. However, there does not exist a random field model dual to the quantum system. According to Theorem 1, one-point variance is determined $A(\mathbf{x}) = V_0 > 0 (\mathbf{x} \in \Omega)$ and $A(\mathbf{x}) = 0 (\mathbf{x} \notin \Omega)$. For $\mathbf{x} \notin \Omega$, inverse mass matrix $\mathbf{\Sigma}_m(\mathbf{x}) = A(\mathbf{x})\mathbf{\Sigma}(\mathbf{x}) = 0$, i.e., an infinite mass. The consequence is as follows: Outside Ω the particle is infinitely heavy thus cannot move at all, i.e., quantum tunneling is impossible. This actually makes perfect sense for PCA: If a field value never fluctuates, then it should contribute exactly zero to all eigenmodes.

3. Extension to Laplace-type random field

The duality can be similarly extended to the Laplace-type kernel, with only coefficients different from the Gaussian case. Please refer to Appendices B and C for details.

III. SCHRÖDINGER PCA ALGORITHM

We aim to design an algorithm, dubbed as Schrödigner PCA, to convert a PCA problem to the corresponding Schrödinger problem. The main idea is already clear in the aforementioned duality. We start this section by revealing the limitation of PCA, motivating Schrödinger PCA.

A. Motivation and algorithm

1. Why PCA degrades

In practice, only a finite set of *anchor points* $S = \{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ are available for field measurements i.e., only $[\phi(\mathbf{x}_1), \cdots, \phi(\mathbf{x}_n)]$ are accessible rather than $\phi(\mathbf{x})$ for any $\mathbf{x} \in \Omega \subset \mathbb{R}^d$. The procedure of PCA on discrete anchor points is shown in Algorithm 1. In brief, one applies PCA to many realizations of the field values at anchor points, obtain eigenmodes defined on these anchor points and interpolate eigenvectors to continuous functions.

In the case $\mathbf{\Sigma} = \sigma^2 \mathbf{I}_{d \times d}$ where σ is the correlation length, two neighboring anchor points should be close enough (e.g.,

distance $\lesssim \sigma$) to make sure the covariance is significant enough to reveal local correlation. As a rough estimate, we suppose Ω has size L on each dimension, then the number of anchor points required scales as $(\frac{L}{\sigma})^d$ which increases as $\frac{L}{\sigma}$ increases (multiscale) or d increases (curse of dimensionality). We will attack the multiscale issue with Schrödinger PCA algorithm.

2. Schrödinger PCA

Instead of trying to detect local correlation with densely distributed anchor points, Schrödinger PCA bypasses this multiscale issue by intergrating out local Gaussian correlation analytically thus transforming the PCA problem to a Schrödinger equation. The key steps are summarized here (also in Algorithm 2): (1) compute only *variance* of fields at each anchor point; (2) leveraging dualities in Eq. (7) and interpolate to obtain the continuous potential function $V(\mathbf{x})$ and inverse mass matrix $\mathbf{\Sigma}_m(\mathbf{x})$; (3) solving the Schrödinger equation, Eq. (5), with elliptic equation solvers.

B. Discussions

1. When Schrödinger PCA outperforms PCA

It is worth emphasizing the key factor that contributes to the success of Schrödinger PCA lies in scale separation, i.e., $\phi(\mathbf{x})$ have smaller scales than $A(\mathbf{x})$ such that $\nabla\nabla^T A(\mathbf{x})/A(\mathbf{x}) \ll \nabla\nabla^T \phi(\mathbf{x})/\phi(\mathbf{x})$. While PCA requires anchor points to characterize $\phi(\mathbf{x})$ (small scale), Schrödinger PCA only requires anchor points to characterize $A(\mathbf{x})$ (large scale). For this scale separation to work, Schrödinger PCA requires extra estimation of the small-scale correlation kernel $K(\mathbf{x}, \mathbf{y})$ to determine the coefficients in the Schrödinger equation, but this is cheap because either (1) the correlation kernel is known a prior (e.g., can be derived from physical principles); (2) a few *detector points* can be added locally around each anchor point to estimate the correlation kernel. In the isotropic case $\mathbf{\Sigma} = \sigma^2 \mathbf{I}$, only one detector point is required to determine σ so the extra computational costs can be ignored. For simplicity, we choose $\mathbf{\Sigma} = \sigma^2 \mathbf{I}$ and assume

σ is known *a priori* (but can depend on **x**) in the following numerical experiments.

2. Error analysis of the second-order approximation of Schrödinger PCA

Since the duality between PCA and Schrödinger equation is only approximated up to second-order approximations, higher-order terms should be small enough for this duality to be valid. The only two approximations used in proof of Theorem 1 (Appendix A) are: (1) The derivative of $A(\mathbf{x})$ is truncated to the first order $\nabla A(\mathbf{x})$, thus higher-order terms $\nabla^n A(\mathbf{x})(n \geqslant 2)$ can be neglected; (2) The derivative of $\phi(\mathbf{x})$ is truncated to the second order $\nabla\nabla\phi(\mathbf{x})$, thus higher-order terms $\nabla^n \phi(\mathbf{x})(n \geqslant 3)$ can be neglected. In Sec. IV we show in practice, the errors can be almost neglected for leading eigenmodes/eigenvalues. We leave the full analysis of how perturbation of high-order terms influence eigenmodes/eigenvalues in future works.

IV. NUMERICAL EXPERIMENTS

First, we revisit the example mentioned in Fig. 1 to test the correctness of the correspondence between PCA and Schrödinger equation stated in Theorem 1. We find PCA and Schrödinger PCA have similar eigenvalues and eigenmodes for the fine grid (i.e., a large number of anchor points), but PCA fails for the coarse grid (i.e., a small number of anchor points) while Schrödinger PCA can still obtain accurate eigenvalues and eigenmodes. Second, we apply our method to the global climate example which is modeled as a random field on a 2D sphere S^2. The climate modes discovered by Schrödinger PCA admit a nice physical interpretation and demonstrate the potential of our method to attack problems on graphs and manifolds beyond the Euclidean space.

A. Two-dimensional Gaussian Process

The two-dimensional Gaussian random field lives on the grid $[-50, 50] \times [-50, 50]$ (size $L \times L \equiv 100 \times 100$). The field generation process can be found in Appendix D. The generated profiles are denoted as $\phi_t(\mathbf{x})$ where $\mathbf{x} = (x_1, x_2)$ and $t = 1, \cdots, N$ index to distinguish among different realizations. The random field satisfies the following statistical properties:

$$A(\mathbf{x}) = \begin{cases} 1 - (x_1^2 + x_2^2)/40^2 & (x_1^2 + x_2^2 \leqslant 40^2), \\ 0 & (x_1^2 + x_2^2 > 40^2), \end{cases} \quad (11)$$

$$C(\mathbf{x}, \mathbf{y}) = \sqrt{A(\mathbf{x})A(\mathbf{y})}\exp\left(-\frac{(\mathbf{x} - \mathbf{y})^T(\mathbf{x} - \mathbf{y})}{2\sigma^2}\right) \quad (\sigma = 3). \quad (12)$$

Note that $\sigma = 3 \ll L = 100$; this corresponds to a multiscale scenario where Schrödinger PCA obtains more accurate results than PCA in the undersampling regime.

1. Oversampling regime with fine grids

In this part, we use numerical results to verify the equivalence between PCA and Schrödinger equation, as pointed out in Theorem 1. PCA (fine grid) is treated as the gold standard method and its results are treated as ground truths.
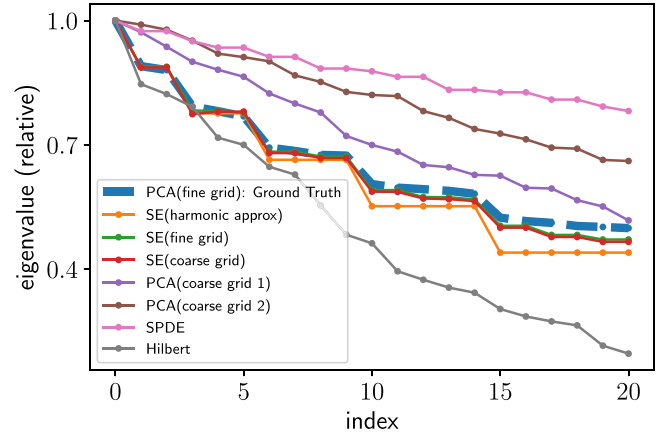

FIG. 2. The (relative) eigenvalues of first 21 modes with different algorithms. Schrödinger PCA is able to recover eigenvalues on both the fine grid and the corase grid, while other methods either fail (SPDE and Hilbert), or succeed only for the fine grid (PCA).

2. PCA (fine grid)

We use all $101 \times 101 = 10\,201$ points as anchor points and generate 40 000 realizations of Gaussian random field satisfying Eq. (11). We apply PCA to obtain the first $k = 21$ eigenmodes and corresponding eigenvalues. The distance of two neighboring anchor points is 1 (smaller than $\sigma = 3$). In this sense, the anchor points oversample the random field, so covariance information among different spatial points can be captured. We implement PCA by using `sklearn.decomposition.pca`.

3. SE (fine grid)

We use all 10 201 anchor points to evaluate $A(\mathbf{x})$ to obtain the potential function $V(\mathbf{x})$ and inverse mass matrix $\boldsymbol{\Sigma}_m(\mathbf{x})$ as indicated in Theorem 1, then the Schrödinger equation, Eq. (5), is discretized on the same grid with the finite difference method. The eigenvalue problem is solved by the python eigen-solver named `numpy.linalg.eigh`. We collect first $k = 21$ eigenstates and eigenenergies. These eigenenergies are negative and are negated below to compare with PCA results.

4. SE (harmonic approximation)

The potential function $V(\mathbf{x}) = -A(\mathbf{x})$ has a quadratic form that is reminiscent of the harmonic oscillator in quantum physics. Although $\boldsymbol{\Sigma}_m(\mathbf{x})$ depends explicitly on **x** it can be nicely approximated as $\boldsymbol{\Sigma}_m(\mathbf{0})$ for ground states and low excited states (eigenstates with low energies) because the wavefunction $\psi(\mathbf{x})$ centered around $\mathbf{x} = 0$ where the potential energy is global minima. As soon as the system is identified as a harmonic oscillator, the eigenenergies and eigenstates are immediately accessible via our knowledge of quantum mechanics, detailed in Appendix E.

As shown in Fig. 2, three methods listed above agree well in terms of eigenvalues for first $k = 10$ modes. For $k > 10$ modes, the staircase degeneracy structure still remains the same for all methods above, while the values have deviations ~ 0.06, which is because the approximation
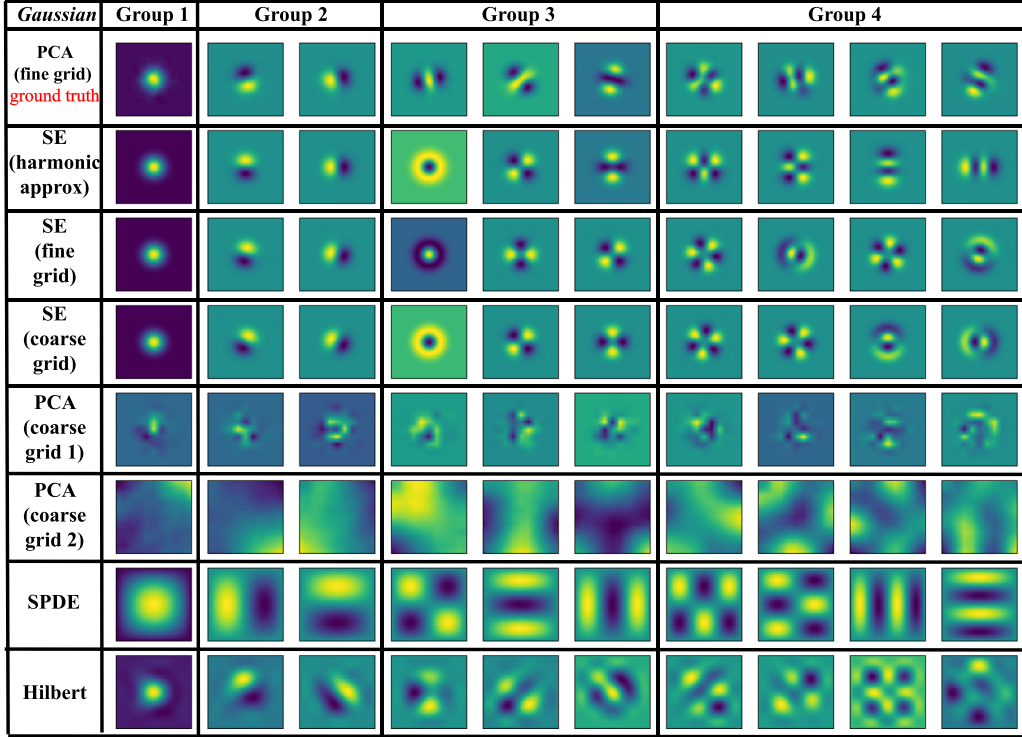
FIG. 3. The eigenmodes of first 10 modes from different algorithms. Each square subfigure is a 2D heat diagram where bright (yellow) and dark (dark blue) colors represent high and low values, respectively. Schrödinger PCA is able to recover eigenmodes on both the fine grid and the corase grid, while other methods either fail (SPDE and Hilbert), or succeed only for the fine grid (PCA).

$\mathbf{\Sigma}_m(\mathbf{x}) \approx \mathbf{\Sigma}_m(\mathbf{0})$ becomes poorer for states with higher energies. The staircase degeneracy structure can be elegantly understood by the energy spectrum of 2D quantum harmonic oscillator: an energy level with quantum number $n \geqslant 0$ admits $n + 1$ ways to choose two integers $n_1, n_2 \geqslant 0$ such that $n_1 + n_2 = n$. Likewise, eigenmodes for all methods are illustrated and compared in Fig. 3 and they show similar behavior for the three aforementioned methods.

As mentioned in the Introduction, the Stochastic PDE (SPDE) method and the reproducing kernel Hilbert space (RKHS) method can also be used to solve the eigendecomposition of random fields. We show below that neither methods are proper to attack our particular case with position-dependent variance and small correlation length.

5. SPDE

The Matern-type random field has eigenmodes which are solutions of the following equality [16]:

$$(\kappa^2 - \Delta)^{\alpha/2}\phi_i(x) = \lambda_i\phi_i(x). \tag{13}$$

Although Eq. (13) looks similar to Eq. (5), our toy random field is not the Matern-type random field, so Eq. (13) is inapplicable to our example. The Matern-type variance function should be at least position-independent, but our variance function is indeed position-dependent. Consequently, we do not expect SPDE to work, but for illustration we choose $\alpha = 2$ and $\kappa = 1/\sigma = 1/3$. Equation (13) therefore becomes a wave equation whose eigenmodes are 2D Fourier bases with wave-number $(k_1, k_2) = \frac{2\pi}{L}(n_1, n_2)$. And according to the wave-number spectrum obtained in Ref. [16], eigenvalues

are

$$\lambda_i = C\big(\kappa^2 + k_{i,1}^2 + k_{i,2}^2\big)^{-\alpha}. \tag{14}$$

Eigenvalues and eigenmodes obtained from SPDE are plotted in Figs. 2 and 3.

6. Reproducing Kernel Hilbert Space (RKHS)

The space of bandlimited continuous functions (truncated Fourier space) is a RKHS [15]. We first project random fields onto the leading 100 Fourier bases (i.e., truncated Fourier transformation), obtaining Fourier coefficients, followed by applying PCA to those Fourier coefficients. Eigenmodes of Fourier coefficients are then transformed back via an inverse Fourier transformation. The RKHS method is insufficient to reproduce eigenvalues and eigenmodes in Figs. 2 and 3 because our toy problem has small correlation length, and the leading 100 Fourier bases cannot capture such small-scale information.

7. Undersampling regime with coarse grids

When the number of anchor points decreases, PCA will fail in terms of both eigenvalues and eigenmodes; by contrast, Schrödinger PCA remains quite robust and behaves nearly the same as the Schrödinger PCA in the oversampling regime, as illustrated in Figs. 2 and 3.

8. PCA (coarse grid 1)

We set anchor points on the uniform coarse grid $[-50, 40, \ldots, 50]^2$ (121 anchor points in total). The same 40 000 random field realizations are used as in the
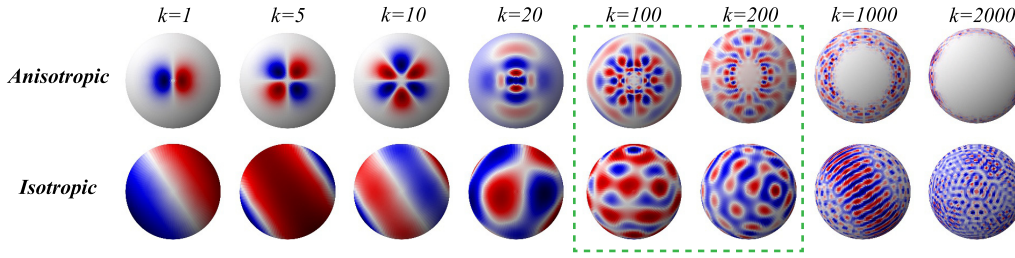
FIG. 4. Eigenmodes of the global climate model obtained from Schrödinger PCA (top view). First row: Anisotropic globe; second row: Isotropic globe. Each globe is a heat diagram where bright (red) and dark (blue) colors represent high and low values, respectively.

oversampling test, but only the field values on the coarse grid are available. We apply PCA to obtain eigenmodes and eigenvalues defined on the coarse grid, and interpolate the eigenmodes back to the fine grid.

9. PCA (coarse grid 2)

Same with *PCA (coarse grid 1)* except that switching the order of PCA and interpolation.

10. SE (coarse grid)

We evaluated $A(\mathbf{x})$ at anchor points (coarse grid) and interpolate $A(\mathbf{x})$ back to the fine grid. Correspondingly the values of $V(\mathbf{x})$ and $\mathbf{\Sigma}_m(\mathbf{x})$ can be determined on the fine grid based on Eq. (7). Finally the SE is discretized with the finite difference method and solved with the `eigh` eigen-solver on the fine grid.

B. Global climate modes

In this example, we demonstrate the effectiveness of the proposed Schrödinger PCA to solve a mode decomposition problem on a unit sphere S^2.

Our earth has various kinds of climates, depending on longitude and latitude of the location. In the following, we use spherical coordinate $\mathbf{x} = (\theta, \varphi)$ and $\phi(\mathbf{x})$ refers to any climate-related scalar field (e.g., temperature). Both fluctuation magnitude $A(\mathbf{x})$ and (local) correlation length $\sigma(\mathbf{x}) = \sqrt{\mathbf{\Sigma(x)}}$ depends explicitly on $\mathbf{x}$: (1) $A(\mathbf{x})$ is small for a place where it is like spring all the year around, while large for another place with scorching summers and cold winters; (2) $\sigma(\mathbf{x})$ is large in oceans due to ocean circulations, while small on a land without any winds.

We consider two "earths": isotropic earth (below $a = 0$) and anisotropic earth ($a = 3$):

$$A(\theta, \varphi) = \frac{1}{20}(3 + a\cos\theta),$$
$$\sigma^2(\theta, \varphi) = \frac{1}{50}(3 - a\cos\theta)d^2. \tag{15}$$

To discretize the Laplacian oeprator on a graph, we utilze icosahedron mesh which contains 2562 vertices, 7680 edges, and 5120 faces. Here d is the (averaged) length of all edges. The details of generating the mesh can be found in Appendix F. We replace the Laplacian operator in Eq. (5) with the graph Laplacian matrix and obtain the eigenmodes on the graph. Finally we interpolate eigenmodes back to the sphere and show them (top view) in Fig. 4. Here $k = 0, 1, 2 \cdots$ refers to the index of eigenstates, ordered from lowest energy to highest energy. The isotropic case provides a baseline for spherical function decomposition, and in fact they correspond

to spherical harmonic functions. We observe that eigenmodes of the anisotropic earth can shed light on the the fluctuation pattern: (1) Fluctuations are large around the poles, so the first few patterns ($k = 1, 5, 10, 20$) only concentrate around poles; (2) intermediate patterns ($k = 100, 200$) are particularly interesting because such patterns have similar magnitude around the pole and the equator, but finer structure is observed around the pole revealing that the correlation length is smaller around the pole than the equator; (3) the last few patterns ($k = 1000, 2000$) capture local fluctuations around the equator, not revealing collective behavior of global temperature.[4] These observations are particularly interesting because it allows one to infer the random field model from eigenmode structures. We will investigate quantitatively the question of "Can one hear the shape of random fields?"[5] in future works.

V. CONCLUSIONS AND DISCUSSIONS

In this paper, we first build the duality between PCA eigenmodes of a random field and eigenstates of a Schrödinger equation. Based on this observation, we propose the Schrödinger PCA algorithm which is more sample-efficient than PCA when the random fields have multiscale structures. Besides the numerical examples in the previous section, the proposed algorithm has the potential to speed up many scientific computations: (1) model reduction of turbulent fluid dynamics [3]; (2) analysis of the cosmic microwave background [1]; (3) collective behavior of relativistic heavy-ion collisions [2]. All of these scenarios have a very narrow correlation kernel, which can be nicely handled by our algorithm, while vanilla PCA will fail on these cases in the undersampling regime.

In the future, it would be interesting to investigate deeper physical implications of the duality between PCA and the Schrödinger equation. Moreover, we would like to extend Schrödinger PCA algorithm to solve higher-dimensional problems, or problems on graphs and Riemannian manifolds, with the help of state-of-the-art elliptic partial differential equation solvers.

---

[4]Although we do not provide PCA results here, the averaged edge length d is much greater than the (averaged) correlation length σ, which in principle prevents PCA from working.

[5]Similar to the spirit of "Can one hear the shape of a drum?" [33].

APPENDIX A: PROOF OF THEOREM 1

In this Appendix, we shall present a formal version of Theorem 1, with rigorous mathematical formulations along with assumptions on the kernel and the function space. Then we make relevant comments on the implications of the assumptions and how it is related to practical scenarios.

Theorem 1. (Formal) Assume that the coefficients satisfy $-A(\mathbf{x}) = V(\mathbf{x})$ and $A(\mathbf{x})\Sigma(\mathbf{x}) = \Sigma_m(\mathbf{x})$, and that $\|\sqrt{A(\mathbf{x})}\|_{L^\infty}\|(-\frac{A'^{\otimes 2}(\mathbf{x})}{8A(\mathbf{x})^{3/2}} + \frac{A''(\mathbf{x})}{4\sqrt{A(\mathbf{x})}})\|_{L^\infty} \leqslant c_1$ for some $c_1 > 0$. In the function space $\mathcal{H}_{c_2}$, the PCA problem of Eq. (4) can be approximated by the Schrödinger problem of Eq. (6), in the sense that $|\lambda_i - E_i| \leqslant O(c_1 + c_2)$ for all i. Here the function space $\mathcal{H}_{c_2}$ is defined via $\mathcal{H}_{c_2} = \{\phi \in L^2 \cap C^3, \|\phi'''\|_{L^\infty} \leqslant c_2\|\phi\|_{L^2}\}$.

Proof. Since the ith eigenvalue λ_i of a positive operator M is related to the min-max formulation of the Rayleigh quotient:

$$\lambda_i = \min_{S:\dim(S)=i} \max_{x \in S, \|x\|=1} (Mx, x),$$

we only need to bound the differences between Eqs. (9) and (10) for $\|\phi\|_{L^2} = 1$. In the following we aim to build a correspondence between two expressions in $\{\cdots\}$ of Eqs. (9) and (10). First, we shall introduce an approximation via Taylor's expansion:

$$\int_{\mathbf{y}} d\mathbf{y} C(\mathbf{x}, \mathbf{y})\phi(\mathbf{y}) \tag{A1}$$

$$= \int_{\mathbf{y}} d\mathbf{y}\sqrt{A(\mathbf{x})A(\mathbf{y})}\exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{x})^T\Sigma^{-1}(\mathbf{y} - \mathbf{x})\right]\phi(\mathbf{y}) \tag{A2}$$

$$= \int_{\mathbf{r}} d\mathbf{r}\sqrt{A(\mathbf{x})A(\mathbf{x} + \mathbf{r})}\exp\left(-\frac{1}{2}\mathbf{r}^T\Sigma^{-1}\mathbf{r}\right)\phi(\mathbf{x} + \mathbf{r}) \tag{A3}$$

$$\approx \int_{\mathbf{r}} d\mathbf{r}\exp\left(-\frac{1}{2}\mathbf{r}^T\Sigma^{-1}\mathbf{r}\right)\left[A(\mathbf{x}) + \frac{1}{2}\nabla A(\mathbf{x}) \cdot \mathbf{r}\right]$$

$$\times \left[\phi(\mathbf{x}) + \nabla\phi(\mathbf{x}) \cdot \mathbf{r} + \frac{1}{2}\mathbf{r}^T\nabla\nabla^T\phi(\mathbf{x})\mathbf{r}\right] \tag{A4}$$

$$= \int_{\mathbf{r}} d\mathbf{r}\exp\left(-\frac{1}{2}\mathbf{r}^T\Sigma^{-1}\mathbf{r}\right)\left\{\frac{1}{2}\mathbf{r}^T[\nabla A(\mathbf{x})\nabla^T\phi(\mathbf{x})\right.$$

$$\left. + A(\mathbf{x})\nabla\nabla^T\phi(\mathbf{x})]\mathbf{r} + \phi(\mathbf{x})A(\mathbf{x})\right\} \tag{A5}$$

$$= \int_{\mathbf{r}} d\mathbf{r}\exp\left(-\frac{1}{2}\mathbf{r}^T\Sigma^{-1}\mathbf{r}\right)\left(\frac{1}{2}\mathbf{r}^T\{\nabla^T \cdot [A(\mathbf{x})\nabla\phi(x)]\}\mathbf{r}\right.$$

$$\left. + \phi(\mathbf{x})A(\mathbf{x})\right) \tag{A6}$$

$$= (2\pi)^{\frac{d}{2}}(\det\Sigma)^{\frac{1}{2}}\left(\phi(x)A(\mathbf{x}) + \frac{1}{2}\mathrm{Tr}\{\Sigma\nabla^T \cdot [A(\mathbf{x})\nabla\phi(\mathbf{x})]\}\right). \tag{A7}$$

From Eqs. (A3) to (A4), $A(\mathbf{x})$ and $\phi(\mathbf{x})$ are Taylor expanded to first and second order, respectively. From Eqs. (A4) to (A5), odd terms of $\mathbf{r}$ vanish due to symmetry. From Eqs. (A6) to (A7), we leverage the Gaussian integral:

$$\int_{\mathbf{r}} d\mathbf{r}\exp\left(-\frac{1}{2}\mathbf{r}^T\Sigma^{-1}\mathbf{r}\right)(\mathbf{r}^T\mathbf{B}\mathbf{r}) = (2\pi)^{\frac{d}{2}}(\det\Sigma)^{\frac{1}{2}}\mathrm{Tr}(\Sigma\mathbf{B}). \tag{A8}$$

We insert the trace term back to Eq. (9) and invoke integration by parts:

$$\int_{\mathbf{x}} d\mathbf{x}\phi(\mathbf{x}) \cdot \mathrm{Tr}\{\Sigma\nabla^T \cdot [A(\mathbf{x})\nabla\phi(\mathbf{x})]\} \tag{A9}$$

$$= \mathrm{Tr}\left\{\Sigma\int_{\mathbf{x}} d\mathbf{x}\phi(\mathbf{x})\nabla^T \cdot [A(\mathbf{x})\nabla\phi(\mathbf{x})]\right\} \tag{A10}$$

$$= \mathrm{Tr}\left(\Sigma\int_{\mathbf{x}} d\mathbf{x}\{\nabla^T \cdot [\phi(\mathbf{x})A(\mathbf{x})\nabla\phi(\mathbf{x})] - \nabla\phi(\mathbf{x})\right.$$

$$\left.\times [A(\mathbf{x})\nabla\phi(\mathbf{x})]\}^T\right) \tag{A11}$$

$$= -\mathrm{Tr}\left[\Sigma\int_{\mathbf{x}} d\mathbf{x}A(\mathbf{x})\nabla\phi(\mathbf{x})\nabla^T\phi(\mathbf{x})\right] \tag{A12}$$

$$= -\int_{\mathbf{x}} d\mathbf{x}A(\mathbf{x})\nabla^T\phi(\mathbf{x})\Sigma\nabla\phi(\mathbf{x}). \tag{A13}$$

By inserting Eqs. (A7) and (A13) back to Eq. (9) we have

$$R_{\mathrm{PCA}}(C, \phi) = -(2\pi)^{\frac{d}{2}}(\det\Sigma)^{\frac{1}{2}}\int_{\mathbf{x}} d\mathbf{x}\left\{-A(\mathbf{x})\phi^2(\mathbf{x})\right.$$

$$\left. + \nabla^T\phi(\mathbf{x})\left[\frac{1}{2}A(\mathbf{x})\Sigma\right]\nabla\phi(\mathbf{x})\right\}. \tag{A14}$$

After integration by parts, Eq. (6) now becomes

$$R_{\mathrm{SE}}(\hat{H}, \psi) = \int_{\mathbf{x}} d\mathbf{x}\left[V(\mathbf{x})\psi^2(\mathbf{x}) + \nabla^T\psi(\mathbf{x})\frac{1}{2}\Sigma_m\nabla\psi(\mathbf{x})\right]. \tag{A15}$$

To equate Eqs. (A14) and (A15) [ignoring the constant factor in Eq. (A14)], we only need Eq. (7).

Therefore, we only need to give an estimate of the approximation to obtain a qualitative bound on the error.

In fact, the Taylor's residue in Eq. (A4) in approximation of $\sqrt{A(\mathbf{x})A(\mathbf{x} + \mathbf{r})}$ is $\sqrt{A(\mathbf{x})}(-\frac{A'^{\otimes 2}(\mathbf{x}+\epsilon\mathbf{r})}{8A(\mathbf{x}+\epsilon\mathbf{r})^{3/2}} + \frac{A''(\mathbf{x}+\epsilon\mathbf{r})}{4\sqrt{A(\mathbf{x}+\epsilon\mathbf{r})}}) \cdot \mathbf{r}^{\otimes 2}$. Similarly, the Taylor's residue in line Eq. (A4) in approximation of $\phi(\mathbf{x} + \mathbf{r})$ is $\frac{1}{6}\phi'''(\mathbf{x} + \epsilon\mathbf{r}) \cdot \mathbf{r}^{\otimes 3}$. Thus, we can bound the approximation error as

$$\int_{\mathbf{r}} d\mathbf{r}\exp\left(-\frac{1}{2}\mathbf{r}^T\Sigma^{-1}\mathbf{r}\right)\left\{|c_1\mathbf{r}^{\otimes 2}|\left[\phi(\mathbf{x}) + \nabla\phi(\mathbf{x}) \cdot \mathbf{r}\right.\right.$$

$$\left.+ \frac{1}{2}\mathbf{r}^T\nabla\nabla^T\phi(\mathbf{x})\mathbf{r}\right] + \left|\left[A(\mathbf{x}) + \frac{1}{2}\nabla A(\mathbf{x}) \cdot \mathbf{r}\right]\right|$$

$$\left.\times \frac{1}{6}\|\phi'''\|_{L^\infty}|\mathbf{r}^{\otimes 3}| + |c_1\mathbf{r}^{\otimes 2}|\frac{1}{6}\|\phi'''\|_{L^\infty}|\mathbf{r}^{\otimes 3}|\right\}.$$

Evaluating the Gaussian integrals, we obtain the desired error bound of $O(c_1 + c_2)$. ∎

To avoid confusion, we provide a few remarks here:

(1) The only approximation made in the proof of Theorem 1 lies in Eq. (A4) where $A(\mathbf{x})$ and $\phi(\mathbf{x})$ are taylor expanded to first-derivative and second-derivative, respectively. The underlying assumption is that (a) the correlation lengths (eigenvalues of $\boldsymbol{\Sigma}$) should be much smaller than the size of whole integral domain, so that the Gaussian factor decays fast enough before Taylor expansion fails; (b) the variation of $A(\mathbf{x})$ is milder than variation of $\phi(\mathbf{x})$, or $\nabla\nabla^T A(\mathbf{x})/A(\mathbf{x}) \ll \nabla\nabla^T \phi(\mathbf{x})/\phi(\mathbf{x})$.

(2) The assumptions on the coefficients and the function space might appear strange. The purpose is to impose a control of higher-order derivatives, so as to make the approximation via Tayler's expansion valid. In practical implementations, we shall relax the constraint and search for eigenfunctions in L^2. For the numerical experiments and its lower eigenmodes, the functions are with low oscillation with associated constants c_1, c_2 small, so the approximation is indeed accurate. There shall be more oscillations for high eigenmode functions, in the regime beyond the applicability of our theorem, which explains why our approximation is more valid and mostly applicable to the approximation of the first few eigenmodes.

Unfortunately, we do not know *a priori* how many lower eigenmodes are applicable for this theorem. The statement of the theorem can only serve as *a posteriori* criteria in the numerical experiments since the desired *a priori* estimates on the first few eigenmodes are not available.

(3) Although we have assumed $C(\mathbf{x}, \mathbf{y})$ as Gaussian, our proof also works for any function that can be decomposed into a linear combination of Gaussian radial basis function. The equivalence between PCA and SE remains valid when the covariance function $C(\mathbf{x}, \mathbf{y})$ is of the form of an exponential function, which is reminiscent of many condensed matter systems in physics, e.g., Ising model. Please refer to Appendices B and C for details.

(4) The corresponding relations Eq. (7) are particularly interesting in terms of physics. Normally one locates the system at disposal on a compact support $\Omega_c \in \Omega$, i.e., $A(\mathbf{x}) > 0$ for $x \in \Omega_s$ while $A(\mathbf{x}) = 0$ for $x \in \Omega/\Omega_s$, $V(\mathbf{x}) = -A(\mathbf{x})$ implies that the quantum particle is trapped in a potential well on Ω_s. Further, if $A(\mathbf{x})$ can be approximated as a quadratic form around the global maxima, then the Schrödinger equation is readily identified as a well-studied system in physics: the quantum harmonic oscillator, the properties of which are summarized in Appendix E.

APPENDIX B: GENERALIZATION OF THEOREM 1 TO EXPONENTIAL CORRELATIONS

In the physics community, we are often concerned with covariance fields of the following type:

$$C(x, y) = \sqrt{A(\mathbf{x})A(\mathbf{y})}\exp\left(-\left|\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{y}-\mathbf{x})\right|\right)/Z_{\boldsymbol{\Sigma}}, \quad (B1)$$

where $\boldsymbol{\Sigma}$ is a positive definite function, c_{d-1} is the area surface of the unit sphere in R^d, $Z_{\boldsymbol{\Sigma}} := (d-1)!c_{d-1}(\det\boldsymbol{\Sigma})^{\frac{1}{2}}$ is the normalization factor of the covariance function, and $|\cdot|$ refers to vector 2-norm.

Similar to the Gaussian case, we shall use Schödinger equation to perform PCA approximations. Here we identify when the PCA problem described by Eq. (4) can be well ap-

proximated by the Schrödinger problem described by Eq. (6) as in Theorem 1.

Theorem 2. (Informal) The PCA problem described by Eq. (4) can be approximated by the Schrödinger problem described by Eq. (6); i.e., two systems have the same equation hence same eigenvalues and eigenmodes, up to a second-order approximation provided that these quantities are equal:

$$\phi(\mathbf{x}) \Longleftrightarrow \psi(\mathbf{x}),$$
$$-A(\mathbf{x}) \Longleftrightarrow V(\mathbf{x}), \quad (B2)$$
$$(d+1)A(\mathbf{x})\boldsymbol{\Sigma} \Longleftrightarrow \boldsymbol{\Sigma}_m.$$

Proof. In the following we aim to build a correspondence between two expressions in $\{\cdots\}$ of Eqs. (9) and (10):

$$\int_{\mathbf{y}} d\mathbf{y}C(\mathbf{x}, \mathbf{y})\phi(\mathbf{y}) \quad (B3)$$

$$= \int_{\mathbf{y}} d\mathbf{y}\sqrt{A(\mathbf{x})A(\mathbf{y})}\exp\left(-\left|\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{y}-\mathbf{x})\right|\right)\phi(\mathbf{y})/Z_{\boldsymbol{\Sigma}} \quad (B4)$$

$$= \int_{\mathbf{r}} d\mathbf{r}\sqrt{A(\mathbf{x})A(\mathbf{x}+\mathbf{r})}\exp\left(-\left|\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{r})\right|\right)\phi(\mathbf{x}+\mathbf{r})/Z_{\boldsymbol{\Sigma}} \quad (B5)$$

$$\approx \int_{\mathbf{r}} d\mathbf{r}\exp\left(-\left|\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{r})\right|\right)\left[A(\mathbf{x}) + \frac{1}{2}\nabla A(\mathbf{x})\cdot\mathbf{r}\right]$$
$$\times \left[\phi(\mathbf{x}) + \nabla\phi(\mathbf{x})\cdot\mathbf{r} + \frac{1}{2}\mathbf{r}^T\nabla\nabla^T\phi(\mathbf{x})\mathbf{r}\right]\Big/Z_{\boldsymbol{\Sigma}} \quad (B6)$$

$$= \int_{\mathbf{r}} d\mathbf{r}\exp\left[-\left|\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{r})\right|\right]\left\{\frac{1}{2}\mathbf{r}^T[\nabla A(\mathbf{x})\nabla^T\phi(\mathbf{x})\right.$$
$$\left. + A(\mathbf{x})\nabla\nabla^T\phi(\mathbf{x})]\mathbf{r} + \phi(\mathbf{x})A(\mathbf{x})\right\}\Big/Z_{\boldsymbol{\Sigma}} \quad (B7)$$

$$= \int_{\mathbf{r}} d\mathbf{r}\exp\left[-\left|\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{r})\right|\right]\left(\frac{1}{2}\mathbf{r}^T\{\nabla^T\cdot[A(\mathbf{x})\nabla\phi(x)]\}\mathbf{r}\right.$$
$$\left. + \phi(\mathbf{x})A(\mathbf{x})\right)\Big/Z_{\boldsymbol{\Sigma}} \quad (B8)$$

$$= \phi(x)A(\mathbf{x}) + \frac{1}{2}(d+1)\mathrm{Tr}\left\{\boldsymbol{\Sigma}^{\frac{1}{2}}\nabla^T\cdot[A(\mathbf{x})\nabla\phi(\mathbf{x})]\boldsymbol{\Sigma}^{\frac{1}{2}}\right\}. \quad (B9)$$

From Eqs. (B5) to (B6), $A(\mathbf{x})$ and $\phi(\mathbf{x})$ are Taylor expanded to first and second order, respectively. From Eqs. (B6) to (B7), odd terms of $\mathbf{r}$ vanish due to symmetry. From Eqs. (B8) to (B9), we leverage the Exponential integral:

$$\int_{\mathbf{r}} d\mathbf{r}\exp\left[-\left|\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{r})\right|\right](\mathbf{r}^T\mathbf{B}\mathbf{r})/Z_{\boldsymbol{\Sigma}} = (d+1)\mathrm{Tr}\left(\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{B}\boldsymbol{\Sigma}^{\frac{1}{2}}\right). \quad (B10)$$

We insert the trace term back to Eq. (9) and invoke integration by parts:

$$\int_{\mathbf{x}} d\mathbf{x}\phi(\mathbf{x})\cdot\mathrm{Tr}\left\{\boldsymbol{\Sigma}^{\frac{1}{2}}\nabla^T\cdot[A(\mathbf{x})\nabla\phi(\mathbf{x})]\boldsymbol{\Sigma}^{\frac{1}{2}}\right\} \quad (B11)$$

$$= \mathrm{Tr}\left\{\boldsymbol{\Sigma}^{\frac{1}{2}}\int_{\mathbf{x}} d\mathbf{x}\phi(\mathbf{x})\nabla^T\cdot[A(\mathbf{x})\nabla\phi(\mathbf{x})]\boldsymbol{\Sigma}^{\frac{1}{2}}\right\} \quad (B12)$$

$$= \text{Tr}\left(\mathbf{\Sigma}^{\frac{1}{2}} \int_{\mathbf{x}} d\mathbf{x} \{ \nabla^T \cdot [\phi(\mathbf{x}) A(\mathbf{x}) \nabla \phi(\mathbf{x})] \right.$$
$$\left. - \nabla \phi(\mathbf{x}) [A(\mathbf{x}) \nabla \phi(\mathbf{x})]^T \} \mathbf{\Sigma}^{\frac{1}{2}} \right) \qquad \text{(B13)}$$

$$= -\text{Tr}\left[\mathbf{\Sigma}^{\frac{1}{2}} \int_{\mathbf{x}} d\mathbf{x} A(\mathbf{x}) \nabla \phi(\mathbf{x}) \nabla^T \phi(\mathbf{x}) \mathbf{\Sigma}^{\frac{1}{2}} \right] \quad \text{(B14)}$$

$$= -\int_{\mathbf{x}} d\mathbf{x} A(\mathbf{x}) \nabla^T \phi(\mathbf{x}) \mathbf{\Sigma} \nabla \phi(\mathbf{x}). \qquad \text{(B15)}$$

By inserting Eqs. (B9) and (B15) back to Eq. (9) we have

$$R_{\text{PCA}}(C, \phi) = -\int_{\mathbf{x}} d\mathbf{x} \left\{ -A(\mathbf{x}) \phi^2(\mathbf{x}) \right.$$
$$\left. + \frac{1}{2}(d+1) \nabla^T \phi(\mathbf{x}) [A(\mathbf{x}) \mathbf{\Sigma} \nabla \phi(\mathbf{x})] \right\}. \quad \text{(B16)}$$

After integration by parts, Eq. (6) now becomes

$$R_{\text{SE}}(\hat{H}, \psi) = \int_{\mathbf{x}} d\mathbf{x} \left[V(\mathbf{x}) \psi^2(\mathbf{x}) + \frac{1}{2} \nabla^T \psi(\mathbf{x}) \mathbf{\Sigma}_m \nabla \psi(\mathbf{x}) \right].$$
$$\text{(B17)}$$

To equate Eqs. (B16) and (B17), we only need Eq. (B2). ∎

APPENDIX C: EXPERIMENTAL RESULTS FOR LAPLACE KERNEL

For the Laplace kernel case, we did same experiments as the Gaussian kernel case. Eigenvalues and eigenmodes are shown in Figs. 5 and 6, respectively.
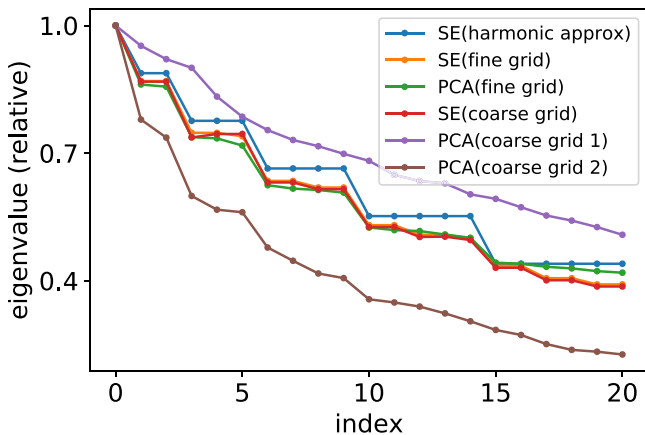

FIG. 5. Exponential correlation example: the (relative) eigenvalues of first 21 modes in five different cases. Note: (1) On the fine grid, Schrödinger PCA and PCA obtain similar eigenvalues; (2) On the coarse grid, Schrödinger PCA still works while PCA fails; (3) Harmonic approximation is nearly exact for first 15 eigenvalues;

APPENDIX D: GENERATION OF GAUSSIAN PROCESS

In this section, we will discuss the generation of the rectangular Gaussian process (GP) in our numerical experiments. We simply set $\mathbf{\Sigma} = \sigma^2 \mathbf{I}$ in Eq. (1).

The generation process contains two stages: (1) we add Gaussian filtering to match the correlation kernel, i.e., $\exp(-\frac{(\mathbf{y}-\mathbf{x})^T(\mathbf{y}-\mathbf{x})}{2\sigma^2})$ in Eq. (1). (2) We rescale the field dependent on $\mathbf{x}$ to match the variance $A(\mathbf{x})$ and magnitude of covariance $\sqrt{A(\mathbf{x})A(\mathbf{y})}$ in Eq. (1).

1. Covariance Generation

Since the two-dimensional GP lies on grid$[-50, 50] \times [-50, 50]$(size 101×101), we start from generating a 101×101 matrix with each component independently sampled from a standard normal distribution.

As we mentioned, an isotropic correlation length σ is introduced in our experiment. The width of the Gaussian filter, denoted as σ_0, is set to match the correlation length σ. It is easy to see that $\sigma = \sqrt{2}\sigma_0$.

For the standard Gaussian filtering algorithm, a truncation along each direction is introduced for efficient calculation. In our numerical experiment, we set this truncation level at 5σ.[6]

$$C(\mathbf{x}, \mathbf{y}) = C_{\text{ideal}}(\mathbf{x}, \mathbf{y}) \eta(5\sigma - |x_1 - y_1|) \eta(5\sigma - |x_2 - y_2|).$$
$$\text{(D1)}$$

Here the $\eta(\cdot)$ is the step function.

$$\eta(x) = \begin{cases} 0, & x < 0; \\ 1, & x > 0; \\ \frac{1}{2}, & x = 0. \end{cases} \qquad \text{(D2)}$$

With this truncation, the Gaussian filtering can be achieved by convolution with a Gaussian kernel G_σ.

2. Modifying magnitudes

Then, due to the scaling property of Gaussian distribution, we design a window function $A'(\mathbf{x})$ to modulate the original GP. Field values at point $\mathbf{x}$ are multiplied by $\sqrt{A(\mathbf{x})}$.

3. The overall algorithm

Based on the above discussion, the overall algorithm to generate n number of samples GP with predefined variance $A(\mathbf{x})$ and the Gaussian-like covariance $C(\mathbf{x}, \mathbf{y})$ with predefined isotropic correlation length σ is as Algorithm 3.

APPENDIX E: PROPERTIES OF QUANTUM HARMONIC OSCILLATOR

The standard equation (in physics) of a one-dimensional harmonic oscillator is written as follows:[7]

$$-\frac{1}{2m} \nabla^2 \psi(x) + \frac{1}{2} m \omega^2 x^2 \psi(x) = E \psi(x), \qquad \text{(E1)}$$

---

[6]This truncation results in a $\mathcal{O}(10^{-7})$ relative error level.

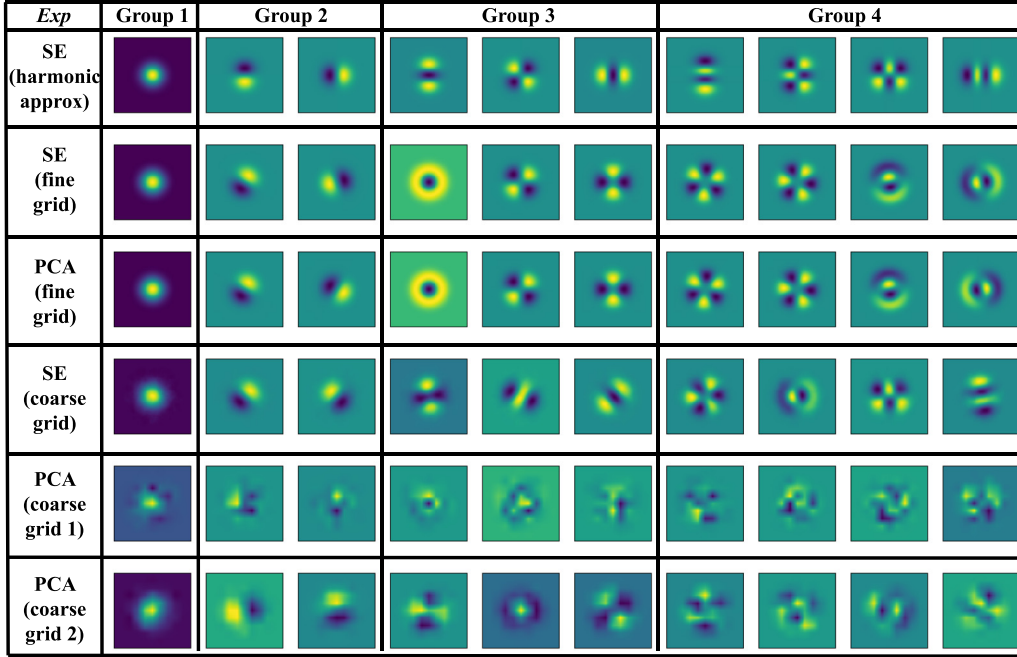[7]We have set the Planck constant $\hbar = 1$.

Exp	Group 1	Group 2		Group 3			Group 4			
SE (harmonic approx)										
SE (fine grid)										
PCA (fine grid)										
SE (coarse grid)										
PCA (coarse grid 1)										
PCA (coarse grid 2)										

FIG. 6. The exponential correlation example: eigenmodes of first 10 modes in 5 different cases. (1) On the fine grid, Schrödinger PCA and PCA obtain similar eigenmodes; (2) Harmonic approximation is nearly exact for first 10 eigenmodes; (3) On the coarse grid, Schrödinger PCA still works while PCA fails.

whose eigenstates $\psi_n(\mathbf{x})$ and eigenenergies E_n are

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}} \left(\frac{m\omega}{\pi}\right)^{\frac{1}{4}} e^{-\frac{m\omega x^2}{2}} H_n(\sqrt{m\omega}x),$$

$$E_n = \left(n + \frac{1}{2}\right)\omega \quad (n = 0, 1, 2, \cdots), \tag{E2}$$

where $H_n(z) = (-1)^n e^{z^2} \frac{d^n}{dz^n}(e^{-z^2})$ are Hermite polynomials. For the two-dimensional harmonic oscillator in our case, it can be decoupled into two independent oscillators along x_1 and x_2, respectively. The total energy is equal to the sum of energy in both directions $E = E_1 + E_2$, and the total wavefunction is equal to the product of wavefunction in both directions $\psi(x_1, x_2) = \psi_1(x_1)\psi_2(x_2)$.

APPENDIX F: GENERATION OF ICOSAHEDRON MESH

There are many ways to build meshes on a sphere: UV sphere, normalized cube, spherified cube, and icosahedron.

UV sphere has singularities around two poles, and normalized cube and spherified cube do not sustain enough rational symmetry. Given the above considerations, we choose icosahedron mesh to discretize the sphere.

Our mesh starts from an icosahedron (20 faces, 30 edges, and 12 vertices). Each face of icosahedron is an equidistant triangle. *Subdivision* is to partition one triangle into four smaller equidistant triangles, as shown in Fig. 7. Then middle points are rescaled to project to the surface of the sphere. A larger number of subdivisions generates finer meshes, shown in Fig. 8. Both Figs. 7 and 8 courtesy of Ref. [34].

After the icosahedron mesh is generated, corresponding degree matrix $\mathbf{D}$, adjacency matrix $\mathbf{A}$ and Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$ can be computed. In the global climate example (Sec. IV B), we replace the Laplacian operator ∇^2 with the Laplacian matrix $\mathbf{L}$ in the Schrödinger equation, i.e.,

$$\nabla^2 \to -\mathbf{L}d^2, \tag{F1}$$

where d is the averaged length of edges.

Algorithm 3. Generation of Gaussian Process.

Input: sample size n, grid infomation $[-50, 50] \times [-50, 50] \sim 101 \times 101$;
variance definition $A(\mathbf{x})$, isotropic correlation length σ
Determine $A'(\mathbf{x})$ with given σ and $A(\mathbf{x})$;
for $i = 1, 2, \ldots, n$ **do**
    (1) Get the initial GP sample $\phi_i^0(\mathbf{x}), \phi_i^0(\mathbf{x}_{j,k}) \sim$ i.i.d.$N(0, 1^2)$ for $j, k \in [-50, 50]$, here $N(0, 1^2)$ stands for the standard normal
      distribution;
    (2) Add correlations via Gaussian filtering at 5σ truncation level. $\phi_i(\mathbf{x}) = \phi_i^0(\mathbf{x}) \otimes G_\sigma$. ;
    (3) Obtain $\psi_i(\mathbf{x})$ via modulating $\phi_i(\mathbf{x})$ through $\psi_i(\mathbf{x}_{j,k}) = A'(\mathbf{x}_{j,k}) \cdot \phi_i(\mathbf{x}_{j,k})$.
end
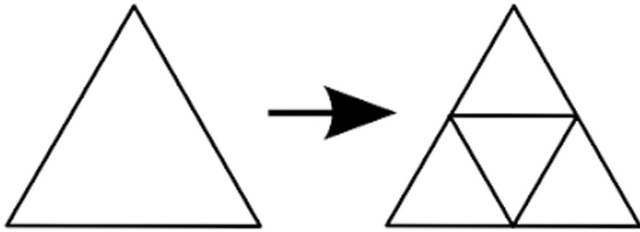**Output:** GP samples $\{\psi_i(\mathbf{x})\}_{i=1}^n$

FIG. 7. Divide one equidistant triangle to four smaller equidistant triangles

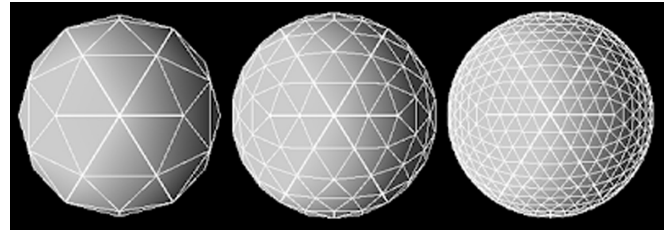
FIG. 8. More subdivisions generate finer meshes.

[1] U.-L. Pen, Generating cosmological Gaussian random fields, Astrophys. J. Lett. **490**, L127 (1997).

[2] J. Novak, K. Novak, S. Pratt, J. Vredevoogd, C. E. Coleman-Smith, and R. L. Wolpert, Determining fundamental properties of matter created in ultrarelativistic heavy-ion collisions, Phys. Rev. C **89**, 034917 (2014).

[3] R. M. Pereira, C. Garban, and L. Chevillard, A dissipative random velocity field for fully developed fluid turbulence, J. Fluid Mech. **794**, 369 (2016).

[4] G. A. Pavliotis, *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*, Vol. 60 (Springer, Berlin, 2014).

[5] A. Garbuno-Inigo, F. Hoffmann, W. Li, and A. M. Stuart, Interacting langevin diffusions: Gradient structure and ensemble Kalman sampler, SIAM J. Appl. Dynam. Syst. **19**, 412 (2020).

[6] R. S. Strichartz, Harmonic analysis as spectral theory of Laplacians, J. Funct. Anal. **87**, 51 (1989).

[7] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, Tensor decomposition for signal processing and machine learning, IEEE Trans. Signal Process. **65**, 3551 (2017).

[8] T. G. Kolda and B. W. Bader, Tensor decompositions and applications, SIAM Rev. **51**, 455 (2009).

[9] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. Nathan Kutz, On dynamic mode decomposition: Theory and applications, J. Comput. Dyn. **1**, 391 (2014).

[10] S. Wold, K. Esbensen, and P. Geladi, Principal component analysis, Chemom. Intell. Lab. Syst. **2**, 37 (1987).

[11] A. Hyvärinen and E. Oja, Independent component analysis: algorithms and applications, Neural netw. **13**, 411 (2000).

[12] I. Tošić and P. Frossard, Dictionary learning, IEEE Signal Process. Mag. **28**, 27 (2011).

[13] D. D. Lee and H. S. Seung, Algorithms for nonnegative matrix factorization, in T. K. Leen, T. G. Dietterich, and V. Tresp, editors, *Advances in Neural Information Processing Systems 13* (MIT Press, Cambridge, MA, 2001), pp. 556–562.

[14] C. Schwab and R. A. Todor, Karhunen–Loève approximation of random fields by generalized fast multipole methods, J. Comput. Phys. **217**, 100 (2006).

[15] A. A Amini, M. J. Wainwright *et al.*, Sampled forms of functional PCA in reproducing kernel Hilbert spaces, Ann. Stat. **40**, 2483 (2012).

[16] F. Lindgren, H. Rue, and J. Lindström, An explicit link between Gaussian fields and Gaussian-Markov random fields: The stochastic partial differential equation approach, J. Roy. Stat. Soc.: Series B (Stat. Methodol.) **73**, 423 (2011).

[17] A. A. Akinduko and A. N. Gorban, Multiscale principal component analysis, J. Phys.: Conf. Ser. **490**, 012081 (2014).

[18] U. Demšar, P. Harris, C. Brunsdon, A. S. Fotheringham, and S. McLoone, Principal component analysis on spatial data: An overview, Ann. Assoc. Amer. Geogr. **103**, 106 (2013).

[19] R. Bellman, R. Kalaba, and B. Kotkin, Differential approximation applied to the solution of convolution equations, Math. Comput. **18**, 487 (1964).

[20] N. Distéfano, On alternative representations of time varying viscoelastic materials, Int. J. Solids Struct. **6**, 1021 (1970).

[21] B. R. Draganov, Exact estimates of the rate of approximation of convolution operators, J. Approx. Theory **162**, 952 (2010).

[22] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg, Glauber modeling in high-energy nuclear collisions, Annu. Rev. Nucl. Part. Sci. **57**, 205 (2007).

[23] R. M. Neal, Monte carlo implementation of Gaussian process models for Bayesian regression and classification, arXiv:physics/9701026.

[24] B. Bócsi, D. Nguyen-Tuong, L. Csató, B. Schölkopf, and J. Peters, Learning inverse kinematics with structured prediction, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, USA* (IEEE, CA, 2011), pp. 698–703.

[25] R. B. Gramacy and H. K. H. Lee, Bayesian treed Gaussian process models with an application to computer modeling, J. Am. Stat. Assoc. **103**, 1119 (2008).

[26] T. Zhang and B. Yang, Big data dimension reduction using PCA, in *Proceedings of the IEEE International Conference on Smart Cloud (SmartCloud'16)* (IEEE, Piscataway, NJ, 2016), pp. 152–157.

[27] H. Hoffmann, Kernel PCA for novelty detection, Pattern Recogn. **40**, 863 (2007).

[28] V. Perlibakas, Distance measures for PCA-based face recognition, Pattern Recogn. Lett. **25**, 711 (2004).

[29] C. Ding and X. He, K-means clustering via principal component analysis, in *Proceedings of the 21st International Conference on Machine Learning* (Association for Computing Machinery, New York, NY, 2004), p. 29.

[30] H. Owhadi, Bayesian numerical homogenization, Multiscale Model. Simul. **13**, 812 (2015).

[31] D. J. Griffiths and D. F. Schroeter, *Introduction to Quantum Mechanics* (Cambridge University Press, Cambridge, UK, 2018).

[32] R. Durrer, *The Cosmic Microwave Background* (Cambridge University Press, Cambridge, UK, 2020).

[33] M. Kac, Can one hear the shape of a drum? Amer. Math. Monthly **73**, 1 (1966).

[34] http://blog.andreaskahler.com/2009/06/creating-icosphere-mesh-in-code.html