Fast Polynomial Approximation of Heat Kernel Convolution on Manifolds and Its Application to Brain Sulcal and Gyral Graph Pattern Analysis

Shih-Gu Huang, Ilwoo Lyu, Anqi Qiu, and Moo K. Chung

Abstract—Heat diffusion has been widely used in brain imaging for surface fairing, mesh regularization and cortical data smoothing. Motivated by diffusion wavelets and convolutional neural networks on graphs, we present a new fast and accurate numerical scheme to solve heat diffusion on surface meshes. This is achieved by approximating the heat kernel convolution using high degree orthogonal polynomials in the spectral domain. We also derive the closed-form expression of the spectral decomposition of the Laplace-Beltrami operator and use it to solve heat diffusion on a manifold for the first time. The proposed fast polynomial approximation scheme avoids solving for the eigenfunctions of the Laplace-Beltrami operator, which is computationally costly for large mesh size, and the numerical instability associated with the finite element method based diffusion solvers. The proposed method is applied in localizing the male and female differences in cortical sulcal and gyral graph patterns obtained from MRI in an innovative way. The MATLAB code is available at http://www.stat.wisc.edu/~mchung/chebyshev.

Index Terms—Heat diffusion, Laplace-Beltrami operator, brain cortical sulcal curves, diffusion wavelets, Chebyshev polynomials.

I. INTRODUCTION

Heat diffusion has been widely used in brain image processing as a form of smoothing and noise reduction starting with Perona and Malik’s ground-breaking study [1]. Many techniques have been developed for surface mesh fairing, regularization [2], [3] and surface data smoothing [4]–[7]. The diffusion equation has been solved by various numerical techniques [4]–[6], [8], [9]. In [6], [10], [11], the heat diffusion was solved iteratively by the discrete estimate of the LB-operator using the finite element method (FEM) and the FDM. However, the FDM are known to suffer numerical instability if the sufficiently small step size is not chosen in the forward Euler scheme. In [8], [9], [12], [13], diffusion was solved by expanding the heat kernel as a series expansion of the LB-eigenfunctions. Although the LB-eigenfunction approach avoids the numerical instability associated with the FEM based diffusion solver [10], the computational complexity of computing eigenfunctions is very high for large-scale surface meshes.

In this paper, motivated by the diffusion wavelet transform [14]–[18] and convolutional neural networks [19] on graphs that all use Chebyshev polynomials, we propose a new spectral method to solve the heat diffusion by approximating the heat kernel by orthogonal polynomials. The previous works did the spectral decomposition on mostly graph Laplacian exclusively using Chebyshev polynomials. The LB-operator with other polynomials were not considered before. We present a new general theory for the LB-operator on an arbitrary manifold that works with an arbitrary orthogonal polynomial. Besides the Chebyshev polynomials, we provide three other polynomials to show the generality of the proposed method. We further derive the closed-form expression of the spectral decomposition of the LB-operator and use it to solve heat diffusion on a manifold for the first time. Taking the advantage of the recurrence relations of orthogonal polynomials [20]–[22], the computational run time of the proposed method is significantly reduced. The proposed method is faster than the LB-eigenfunction approach and FEM based diffusion solvers [9]. We further applied the fast polynomial approximation method to iterative convolution to obtain multiscale features, which is shown to be as good as the diffusion wavelet in detecting localized surface signals [14]–[18].

The proposed method is applied in quantifying brain sulcal and gyral patterns. The sulcal and gyral features such as gyrification index, sulcal depth, curvature, sulcal length and sulcal area were widely used in revealing significant differences between populations [23]. In [24], the difference of the superior temporal sulcus length was analyzed. Shi et al. [25] computed the lengths of sulcal curves in characterizing the Alzheimer’s disease (AD). Seong et al. [26] measured the sulcal depth and average mean curvature along the sulcal lines in the AD study. In [27], various metrics were proposed to
measure the difference between sulcal graph features including sulcal pits, sulcal basins and ridge points. Lyu et al. [28] computed local gyrification index using shape-adaptive kernels by performing wavefront propagation with sulcal and gyral curves as the source. Im et al. [29] measured the similarity between two sulcal graphs.

The main contributions of the paper are as follows. 1) The development of a general polynomial approximation theory for LB-operator and heat kernel and its application to solving diffusion equations fast. The derivation of the closed-form solutions of the expansion that enables faster computation of the heat diffusion than before. 2) New multiscale shape analysis framework on manifolds that utilizes the iterative heat kernel convolution property that is as powerful as diffusion wavelets.

II. METHODS

We present a new general spectral theory for diffusion equations and the heat kernel using four different types of polynomials (Jacobi, Chebyshev, Hermite, Laguerre) to show the generality of the method. The analytic closed-form solutions to the expansion coefficients are derived and used to solve the heat diffusion fast. The new theory works for an arbitrary orthogonal polynomial.

A. Diffusion on Manifolds

Let functional data \( f \in L^2(M) \), the space of square integrable functions on manifold \( M \) with inner product

\[
(f, h) = \int_M f(p)h(p)d\mu(p),
\]

where \( \mu(p) \) is the Lebesgue measure such that \( \mu(M) \) is the total area or volume of \( M \). Let \( \Delta \) denote the LB-operator on \( M \). Let \( \psi_j \) be the eigenfunctions of the LB-operator with eigenvalues \( \lambda_j \), i.e., \( \Delta \psi_j = \lambda_j \psi_j \). Let us order the eigenvalues as \( 0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \).

The isotropic heat diffusion on \( M \) with \( f \) as the initial observed data is given by

\[
\frac{\partial g(p, \sigma)}{\partial \sigma} + \Delta g = 0, \quad g(p, \sigma = 0) = f(p), \quad (1)
\]

where \( \sigma \) is the diffusion time. It has been shown that the convolution of \( f \) with heat kernel \( K_\sigma \) is the unique solution of (1) [8], [9], [30], [31]:

\[
g(p, \sigma) = K_\sigma * f(p) = \int_M K_\sigma(p, q)f(q)d\mu(q),
\]

with the heat kernel given by

\[
K_\sigma(p, q) = \sum_{j=0}^{\infty} e^{-\lambda_j \sigma} \psi_j(p)\psi_j(q). \quad (2)
\]

The heat kernel convolution can be written as

\[
g(p, \sigma) = K_\sigma * f(p) = \sum_{j=0}^{\infty} e^{-\lambda_j \sigma} f_j \psi_j(p) \quad (3)
\]

with coefficients \( f_j \) computed as

\[
f_j = \int_M f(p)\psi_j(p)d\mu(p).
\]

B. Basic Idea in 1D Diffusion

We explain the core idea with 1D example. Consider time series data \( f \) on \([0, 1]\). The solution of heat diffusion (1) is given by the weighted cosine series representation [32],

\[
g(p, \sigma) = K_\sigma * f = \sum_{j=0}^{\infty} e^{-j^2 \pi^2 \sigma} f_j \psi_j, \quad (4)
\]

where \( \psi_j(p) = \sqrt{\frac{2}{\pi}} \cos(j\pi p) \) are the eigenfunctions of \( \Delta = -\frac{d^2}{dp^2} \). From Taylor expansion \( e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} \),

\[
K_\sigma * f = \sum_{n=0}^{\infty} \frac{(-\sigma)^n}{n!} \sum_{j=0}^{\infty} f_j (j^2 \pi^2)^n \psi_j.
\]

Since \( \Delta \psi_j = j^2 \pi^2 \psi_j \) and \( \Delta^n \psi_j = (j^2 \pi^2)^n \psi_j \), we have

\[
K_\sigma * f = \sum_{n=0}^{\infty} \frac{(-\sigma)^n}{n!} \Delta^n f.
\]

Thus, heat diffusion can be computed simply by using the power of Laplacian. If we can further compute the power of Laplacian quickly using some recursion, the computation can be done more quickly.

C. Fast Polynomial Approximation

Here we present a general new theory for an arbitrary manifold that works in any type of image domain including surface and volumetric meshes. Consider an orthogonal polynomial \( P_n \) over interval \([a, b]\) with inner product \( \int_a^b P_n(\lambda)P_m(\lambda)w(\lambda)d\lambda = \delta_{nk} \), the Dirac delta. The weight \( w(\lambda) \) differs for polynomials. \( P_n \) is often defined using the second order recurrence [22],

\[
P_{n+1}(\lambda) = (A_n \lambda + B_n)P_n(\lambda) + C_n P_{n-1}(\lambda) \quad (5)
\]

with initial conditions \( P_{-1}(\lambda) = 0 \) and \( P_0(\lambda) = 1 \). We expand the exponential weight of the heat kernel by polynomials \( P_n \):

\[
e^{-\lambda \sigma} = \sum_{n=0}^{\infty} c_{\sigma,n} P_n(\lambda), c_{\sigma,n} = \int_a^b e^{-\lambda \sigma} P_n(\lambda)w(\lambda)d\lambda. \quad (6)
\]

Substituting (6) into (3), the solution of heat diffusion can be expressed in terms of the polynomials:

\[
K_\sigma * f = \sum_{n=0}^{\infty} c_{\sigma,n} \sum_{j=0}^{\infty} P_n(\lambda_j) f_j \psi_j. \quad (7)
\]
Since $\Delta \psi_j = \lambda_j \psi_j$, we have $\Delta^l \psi_j = \lambda^l_j \psi_j$. Assuming the form $P_n(\lambda) = \sum_{l=0}^n d_l \lambda^l$, we have

$$P_n(\lambda_j) \psi_j = \sum_{l=0}^n d_l \lambda_j^l \psi_j = \sum_{l=0}^n d_l \Delta^l \psi_j = P_n(\Delta) \psi_j. \quad (8)$$

By substituting (8) into (7), the heat diffusion equation is solved by polynomial expansion involving the LB-operator but without the LB-eigenfunctions,

$$K_\sigma \ast f = \sum_{n=0}^\infty c_{\sigma,n} P_n(\Delta) f.$$

Since $P_n$ is a polynomial of degree $n$, the direct computation of $P_n(\Delta) f$ requires the costly computation of $\Delta f, \Delta^2 f, \cdots, \Delta^n f$. Instead, we compute $P_n(\Delta) f$ by the following recurrence

$$P_{n+1}(\Delta) f = (A_n \Delta + B_n) P_n(\Delta) f + C_n P_{n-1}(\Delta) f$$

with initial conditions $P_{-1}(\Delta) f = 0$ and $P_0(\Delta) f = f$.

In practice, the expansion is truncated at degree $m$, which is empirically determined. The expansion coefficients $c_{\sigma,n}$ can be computed from the closed-form solution to (6). In the following, we present three examples of the fast polynomial approximation methods based on the Jacobi, Hermite and Laguerre polynomials.

**Jacobi Polynomials:** The Jacobi polynomials $P_n^{(\alpha,\beta)}(\lambda)$, which are orthogonal in $[-1, 1]$ for $\alpha, \beta > -1$, are defined by the recurrence (5) with parameters given by [22],

$$A_n = \frac{2n + \alpha + \beta + 2}{2(n+1)(n+\alpha + \beta + 1)}, \quad B_n = \frac{2(n+1)(n+\alpha + \beta + 1)(2n+\alpha + \beta)}{(n+\alpha)(n+\beta)(2n+\alpha + \beta + 2)}, \quad C_n = \frac{2(n+1)(n+\alpha + \beta + 1)(2n+\alpha + \beta)}{(n+1)(n+\alpha + \beta + 1)(2n+\alpha + \beta)}.$$

The Jacobi polynomials are orthogonal over interval $[-1, 1]$ with inner product [22],

$$\int_{-1}^1 P_n^{(\alpha,\beta)}(\lambda) P_k^{(\alpha,\beta)}(\lambda)(1 - \lambda)^\alpha (1 + \lambda)^\beta d\lambda = \frac{2^{\alpha+\beta+1} \Gamma(n + \alpha + 1) \Gamma(n + \beta + 1)}{(2n + \alpha + \beta + 1) \Gamma(n + \alpha + \beta + 1) n!} \delta_{nk}.$$

Many polynomials such as Chebyshev, Legendre and Gegenbauer polynomials defined in $[-1, 1]$ are the special cases of the Jacobi polynomials [22].

The eigenvalue $\lambda$ of the LB-operator ranges over $[0, \infty)$. Expanding the exponential weight $e^{-\lambda \sigma}$ by the Jacobi polynomials may not be able to provide a good fit outside the interval $[-1, 1]$. Hence, we shift and scale Jacobi polynomials with parameter $b > 0$

$$P_n^{(\alpha,\beta)}(\lambda) = P_n^{(\alpha,\beta)} \left( \frac{2\lambda}{b} - 1 \right), \quad (9)$$

which are orthogonal over $[0, b]$. Then, $e^{-\lambda \sigma}$ is expanded in terms of $P_n^{(\alpha,\beta)}$. \hfill

**Theorem 1:** The Jacobi polynomial expansion of the solution to heat diffusion (1) is given by

$$K_\sigma \ast f = \sum_{n=0}^\infty c_{\sigma,n} \overline{T}_n^{(\alpha,\beta)}(\Delta) f, \quad (10)$$

where the coefficients $c_{\sigma,n}$ have the closed-form solution

$$c_{\sigma,n} = \frac{\Gamma(\alpha + \beta + n + 1)}{\Gamma(\alpha + \beta + 2n + 1)} (-b \sigma)^n \times F_1 \left( \frac{\beta + n + 1}{\alpha + \beta + 2n + 2}; -b \sigma \right),$$

and $p F_q$ is the generalized hypergeometric function [22].

**Proof:** We first derive the expansion of $e^{-\lambda \sigma}$ using the Jacobi polynomials $P_n^{(\alpha,\beta)}$. The algebraic derivation will show that the expansion of $e^{-\lambda \sigma}$ is given by

$$e^{-\lambda \sigma} = \sum_{n=0}^\infty \gamma_n (-2 \sigma)^n e^{\sigma - \frac{\lambda}{2}} F_1 \left( \frac{\beta + n + 1}{\alpha + \beta + 2n + 2}; -b \sigma \right) P_n^{(\alpha,\beta)}(\lambda).$$

where $\gamma_n = \frac{\Gamma(\alpha + \beta + n + 1)}{\Gamma(\alpha + \beta + 2n + 1)}$, and $p F_q$ is the generalized hypergeometric function [22], [33]. The expansion is only valid in interval $[-1, 1]$. To obtain the expansion of $e^{-\lambda \sigma}$ in terms of the shifted and scaled Jacobi polynomial (9), we replace $\lambda$ by $\frac{2\lambda}{b} - 1$ and $\sigma$ by $\frac{b}{2}$ and expand $e^{-\lambda \sigma + \frac{b}{2}}$ as

$$\sum_{n=0}^\infty \gamma_n (-b \sigma)^n e^{\frac{b \sigma}{2}} F_1 \left( \frac{\beta + n + 1}{\alpha + \beta + 2n + 2}; -b \sigma \right) \overline{P}_n^{(\alpha,\beta)}(\lambda).$$

We divide the both sides of the equation by $e^{\frac{b \sigma}{2}}$, and the expansion of $e^{-\lambda \sigma}$ follows. ■

**Chebyshev Polynomials:** The Chebyshev polynomials $T_n(\lambda) = \cos(n \cos^{-1} \lambda)$ defined in interval $[-1, 1]$ are the special cases of the Jacobi polynomials [22],

$$T_n(\lambda) = \frac{4^n (n!)^2}{(2n)!} P_n(-\frac{1}{2}, -\frac{1}{2})(\lambda). \quad (11)$$

The Chebyshev polynomials satisfy the recurrence relation (5) with parameters $A_n = 2 - \delta_{0n}$, $B_n = 0$ and $C_n = -1$. Similar to using the shifted and scaled Jacobi polynomials in Theorem 1, we shift and scale the Chebyshev polynomials to

$$\overline{T}_n(\lambda) = T_n \left( \frac{2\lambda}{b} - 1 \right)$$

for the expansion of exponential weight over interval $[0, b]$.

**Theorem 2:** The Chebyshev polynomial expansion of the solution to heat diffusion (1) is given by

$$K_\sigma \ast f = \sum_{n=0}^\infty c_{\sigma,n} \overline{T}_n(\Delta) f,$$

where the coefficients $c_{\sigma,n}$ have the closed-form solution

$$c_{\sigma,n} = (2 - \delta_{0n})(-1)^n e^{-\frac{b \sigma}{2}} I_n \left( \frac{b \sigma}{2} \right),$$

and $I_n$ is the modified Bessel function of the first kind [22].

**Proof:** We provide two different proofs. The first proof is based on Theorem 1. The Chebyshev polynomial is a special case of the Jacobi polynomial (11), and thus their shifted and
scaled versions have the relation $T_n(\lambda) = \sum_{l=0}^{\infty} \frac{(n+l)!}{l!(2n)!} T_{n+l}(\lambda)$. Identifying $\alpha = \beta = -\frac{1}{2}$ in Theorem 1 and noting $\frac{1}{\Gamma(a+\beta+1)} = 1$ when $n = 0$, we have

$$c_{\lambda,n} = \frac{2 - \delta_{00}(-b\sigma)^n}{2^{n+1}} \frac{n+1/2}{2n+1} \cdot \frac{-b\sigma}{2n+1}.$$

(12)

The modified Bessel function is closely related to the generalized hypergeometric function [22],

$$I_n(z) = \frac{z^n e^{\pm z}}{2^n n!} F_1 \left( n + 1/2 ; 2n + 1 ; \mp 2z \right).$$

(13)

Substitute $I_n(z)$ in (13) with $z = \frac{b\sigma}{2}$ for the term 1 $F_1$ in (12), and the result follows.

The second proof is based on the generating function of the modified Bessel functions [22]:

$$e^{z \cos \theta} = I_0(z) + 2 \sum_{n=1}^{\infty} I_n(z) \cos(n\theta).$$

(14)

We use the generating function to develop the relation between exponential function and the Chebyshev polynomials. Let $\theta = \cos^{-1} \lambda$, and then (14) can be rewritten in terms of the Chebyshev polynomials $T_n(\lambda) = \cos(n \cos^{-1} \lambda)$,

$$e^{z \lambda} = I_0(z) T_0(\lambda) + 2 \sum_{n=1}^{\infty} I_n(z) T_n(\lambda),$$

(15)

where $T_0(\lambda) = 1$. Replacing $\lambda$ by $\frac{z}{2n+1}$ and identifying $z = -\frac{b\sigma}{2}$ in (15) give the expansion of $e^{-i\sigma + \frac{b\sigma}{2}}$ in terms of the shifted and scaled Chebyshev polynomials $T_n$: $$e^{-i\sigma + \frac{b\sigma}{2}} = I_0 \left( \frac{-b\sigma}{2} \right) T_0(\lambda) + 2 \sum_{n=1}^{\infty} I_n \left( \frac{-b\sigma}{2} \right) T_n(\lambda).$$

(16)

We divide the both sides of the equation by $e^{\frac{b\sigma}{2}}$, and the expansion of $e^{-i\sigma}$ follows. Note that $I_n \left( \frac{-b\sigma}{2} \right) = (-1)^n I_n \left( \frac{b\sigma}{2} \right)$. ■

In numerical implementation, given the maximum eigenvalue $\lambda_{\text{max}}$ of the discrete LB-operator, we set $b = \lambda_{\text{max}}$ such that the Chebyshev polynomials provide good approximation of the exponential weight over $[0, \lambda_{\text{max}}]$. [14]

**Hermite Polynomials:** The Hermite polynomials

$$H_n(\lambda) = (-1)^n e^{\lambda^2} \frac{d^n}{d\lambda^n} e^{-\lambda^2}$$

with $H_0(\lambda) = 0$ and $H_1(\lambda) = 1$ in $(-\infty, \infty)$ satisfy the recurrence relation (5) with parameters $A_n = 2, B_n = 0, C_n = -2n$.

The orthogonal condition of the Hermite polynomials [22] is given by

$$\int_{-\infty}^{\infty} H_n(\lambda) H_m(\lambda) e^{-\lambda^2} d\lambda = \delta_{nm} \sqrt{\pi} 2^n n! \delta_{nm}.$$

**Theorem 3:** The Hermite polynomial expansion of the solution to heat diffusion (1) is given by

$$K_\sigma \ast f = \sum_{n=0}^{\infty} c_{\sigma,n} H_n(\Delta) f,$$

where the coefficients $c_{\sigma,n}$ have the closed-form solution

$$c_{\sigma,n} = \frac{1}{n!} \left( \frac{-\sigma}{2} \right)^n e^{\frac{\sigma^2}{4}}.$$

**Proof:** It follows that the expansion of $e^{-i\sigma}$ in terms of the Hermite polynomials has coefficients

$$c_{\sigma,n} = \frac{1}{\sqrt{\pi} 2^n n!} \int_{-\infty}^{\infty} e^{-i\sigma} H_n(\lambda) e^{-\lambda^2} d\lambda.$$

The closed-form solution of the expansion coefficients can be derived through the integral property involving the Hermite polynomials $\int_{-\infty}^{\infty} e^{-(\lambda - y)^2} H_n(\lambda) d\lambda = \sqrt{\pi} 2^n y^n$ [34] with $y = -\frac{\sigma}{2}$.

The statement can be also proved using the exponential generating function [22],

$$e^{2z\lambda - z^2} = \sum_{n=0}^{\infty} \frac{z^n}{n!} H_n(\lambda).$$

Here, it is used to derive the expansion coefficients by dividing the both sides of the equation by $e^{-z^2}$ and then identifying $z = -\frac{\sigma}{2}$. ■

**Laguerre Polynomials:** The Laguerre polynomials $L_n$ satisfy the recurrence relation (5) with parameters

$$A_n = -\frac{1}{n+1}, B_n = \frac{n+1}{n+1}, C_n = -\frac{n}{n+1}$$

and $L_1(\lambda) = 0$ and $L(\lambda) = 1$ in $[0, \infty)$ [22].

**Theorem 4:** The Laguerre polynomial expansion of the solution to heat diffusion (1) is given by

$$K_\sigma \ast f = \sum_{n=0}^{\infty} c_{\sigma,n} L_n(\Delta) f,$$

where the coefficients $c_{\sigma,n}$ have the closed-form solution

$$c_{\sigma,n} = \frac{\sigma^n}{(\sigma + 1)^{n+1}}.$$

**Proof:** From the orthogonal condition of the Laguerre polynomials [22],

$$\int_{0}^{\infty} L_n(\lambda) L_k(\lambda) e^{-\lambda^2} d\lambda = \delta_{nk},$$

the expansion of $e^{-i\sigma}$ in terms of the Laguerre polynomials has coefficients given as the inner product of $e^{-i\sigma}$ and $L_n$:

$$c_{\sigma,n} = \int_{0}^{\infty} e^{-i\sigma} L_n(\lambda) e^{-\lambda^2} d\lambda.$$

The closed-form solution of the expansion coefficients can be derived through the integral property $\int_{0}^{\infty} e^{-\lambda^2} L_n(\lambda) d\lambda = \sqrt{\pi} n! y^{-n-1}$ [34] with $y = \sigma + 1$.

Alternately, we can prove the theorem using the exponential generating function of the Laguerre polynomials [22],

$$\frac{1}{1-z} e^{-\frac{\sigma}{1-z}} = \sum_{n=0}^{\infty} \frac{z^n}{n!} L_n(\lambda).$$

Multiply the both sides of the equation by $1-z$. Let $\frac{z}{1-z} = \sigma$, i.e., $z = \frac{\sigma}{1+\sigma}$, and then the expansion of $e^{-i\sigma}$ follows. ■
D. Numerical Implementation

The MATLAB code is available at http://www.stat.wisc.edu/~mchung/chebyshev.

Expansion degree. The expansion degree $m$ is empirically determined to the sufficiently small MSE. Fig. 1 displays the heat diffusion on the left hippocampus surface mesh with 2338 vertices and 4672 triangles, with diffusion time $\sigma = 1.5$ and expansion degree $m = 100$. The reconstruction error is measured by the mean squared error (MSE) between the heat diffusion solver [10], [11] and the LB-eigenfunction approach [8], [9] on the unit sphere $S^2$, where the ground truth can be analytically obtained by the spherical harmonics (SPHARM) $Y_{lm}$, which are the eigenfunctions of the LB-operator with $V(T)$ is the Voronoi region (gray area) computed following [35]. Let $p_i$ and $p_k$ denote the other two vertices of $T$ with angles $\angle p_j p_i$ and $\angle p_k p_i$ and edge lengths $|p_i p_j|$ and $|p_i p_k|$. Then, the Voronoi region area $V(T)$ at $p_i$ is given by $\frac{1}{2}((p_i p_j)^2 \cot \angle p_k + (p_i p_k)^2 \cot \angle p_j)$ (gray area of Fig. 2-right). The computation of $A(T)$ is done using the Heron’s formula involving the three edge lengths of $T$. A simpler cotan discretization in [9], [11], [12] can be also used.

Iterative kernel smoothing. We can obtain diffusion related multiscale features at different time points by iteratively performing heat kernel smoothing. Instead of applying the polynomial approximation separately for each $\sigma$, the computation can also be realized in an iterative fashion. The solution to heat diffusion with larger diffusion time can be broken into iterative heat kernel convolutions with smaller diffusion time [9].

$$K_{\sigma_1+\sigma_2+\ldots+\sigma_m} \ast f = K_{\sigma_1} \ast K_{\sigma_2} \ast \ldots \ast K_{\sigma_m} \ast f.$$  

Thus, if we compute $K_{0.25} \ast f$, then $K_{0.5} \ast f$ can be simply computed as two repeated kernel convolutions, $K_{0.25} \ast (K_{0.25} \ast f)$. Heat diffusion with much larger diffusion time can be done similarly. Fig. 3 displays heat diffusion with $\sigma = 0.25$, 0.5, 0.75 and 1 realized by iteratively applying the Chebyshev approximation method with $\sigma = 0.25$ sequentially four times. As $\sigma$ increases, we are smoothing the surface more smoothly and MSE increases.

E. Validation

We compared the Chebyshev method against the FEM based diffusion solver [10], [11] and the LB-eigenfunction approach [8], [9] on the unit sphere $S^2$, where the ground truth can be analytically obtained by the spherical harmonics (SPHARM) $Y_{lm}$, which are the eigenfunctions of the LB-operator with degree $m$. The expansion degree $m$ is empirically determined to the sufficiently small MSE. The expansion degree $m$ is empirically determined to the sufficiently small MSE. The expansion degree $m$ is empirically determined to the sufficiently small MSE. The expansion degree $m$ is empirically determined to the sufficiently small MSE.

Fig. 1. Left: left hippocampus surface, heat diffusion with $\sigma = 1.5$ using the Chebyshev, Hermite and Laguerre approximation methods with degree $m = 100$. Right: MSE between the original surface and the polynomial approximation methods for different expansion degree $m$. The Hermite approximation method has the slowest convergence, while the Chebyshev method converges slightly faster than the Laguerre method. But all the methods converge quickly with degree $m = 100$.

Fig. 2. Left: $\phi_i$ and $\theta_i$, angles opposite to the edge connecting $p_i$ and $p_j$ corresponding to $T_{ij}^+$ and $T_{ij}^-$, used for computing the global coefficient matrix $C = (C_{ij})$. Right: Computation of area $A_i$ at vertex $p_i$. The neighboring triangles are decomposed into three sets: $\mathcal{O}_i$ is the set of nonobtuse triangles, $\tilde{\mathcal{O}}_i$ is the set of obtuse triangles with obtuse angle at $p_i$, and $\hat{\mathcal{O}}_i$ is the set of obtuse triangles with nonobtuse angle at $p_i$.

Fig. 3. We iteratively applied Chebyshev approximation method with $\sigma = 0.25$ four times to the left hippocampus surface mesh coordinates to obtain heat diffusion with $\sigma = 0.25, 0.5, 0.75$ and 1. As $\sigma$ increases, we are smoothing the surface more smoothly and MSE increases.

$$A_i = \sum_{T \in \mathcal{O}_i} V(T) + \frac{1}{2} \sum_{T \in \tilde{\mathcal{O}}_i} A(T) + \frac{1}{2} \sum_{T \in \hat{\mathcal{O}}_i} A(T),$$
evolution $l(l+1)$. Given surface data $f$ on the sphere
\[ f(p) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_{lm} Y_{lm}(p), \quad p \in S^2. \tag{16} \]

The heat kernel convolution at time $\sigma$ is given as [32]
\[ g(p, \sigma) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sigma^{l(l+1)} f_{lm} Y_{lm}(p), \tag{17} \]
where $f_{lm} = \int_{S^2} f(p)Y_{lm}(p)d\mu(p)$.

**Ground Truth:** Assign value 1 within one circular region, $-1$ within the other circular region, and all other regions were assigned value 0 on the spherical meshes with 2562, 10242, 40962, 163842, 655362 and 2621442 vertices (Fig. 4). We fitted the above signal using SPHARM with degree $l = 100$, which is high enough degree to provide numerical accuracy up to 4 decimal places in terms of MSE. The above signals were smoothed with $\sigma = 0.01$ using (17) and taken as the ground truth.

We applied the three methods with different $\sigma$ values (0.005, 0.01, 0.02 and 0.05). Fig. 4 displays the result of the LB-eigenfunction approach with 210 eigenfunctions, FEM based diffusion solver with 405 iterations, and Chebyshev approximation method with degree 45 that achieved the similar reconstruction error of about $10^{-5}$ MSE.

**Computational runtime over mesh sizes.** To achieve the similar reconstruction error, the FEM based diffusion solver and Chebyshev approximation method need more iterations and higher degree for larger meshes, while the LB-eigenfunction approach is nearly unaffected by the mesh size (Fig. 5-left). Fig. 5-right displays the computational time of the three methods at the similar accuracy (MSE about $10^{-5}$).

**Computational Run Time Over Diffusion Times:** The computational run time for different $\sigma$ (0.005, 0.01, 0.02, 0.05) with fixed spherical mesh resolution (40962 vertices) was also investigated. To achieve the similar reconstruction error, the FEM based diffusion solver and Chebyshev expansion method need more iterations and higher degree for larger $\sigma$, while the LB-eigenfunction approach requires less number of eigenfunctions (Fig. 6-left). Fig. 6-right displays the computational run time over $\sigma$ at the same MSE of about $10^{-7}$.

From Figs. 5 and 6, the LB-eigenfunction method is the slowest. The polynomial approximation method is up to 40 times faster than the FEM based diffusion solver and took 5.7 seconds for $\sigma = 0.01$ on the sphere with 2621442 vertices.

**III. APPLICATION**

**A. HCP Dataset**

We used the T1-weighted MRI of 268 females and 176 males in the Human Connectome Project (HCP) database [36]. MRI were obtained using a Siemens 3T Connectome Skyra scanner with a 32-channel head coil. The details on image acquisition parameters and image processing can be found in [37], [38].

A bias field correction was performed, and the T1-weighted image was registered to the MNI space with a FLIRT affine and then a FNIRT nonlinear registration [39]. The distortion-and bias-corrected T1-weighted image was then undergone the FreeSurfer’s recon-all pipeline [40]–[42] that includes the segmentation of volume into predefined structures, reconstruction of white and pial cortical surfaces, and FreeSurfer’s standard folding-based surface registration to their surface atlas. Then, the white, pial and spherical surfaces of the left and right hemispheres were produced.

**B. Sulcal and Gyral Curve Extraction**

The automatic sulcal curve extraction method (TRACE) [43], [44] was used to detect concave regions (sulcal fundi) along which sulcal curves are traced. The method consists of two main steps: (1) sulcal point detection and (2) curve delineation by tracing the detected sulcal points. For sulcal
The paired robustness to noise in TRACE than the existing method [47].

The paired robustness to noise in TRACE than the existing method [47]. The experimental results from [44] (Table I) show higher reproducibility and robustness to noise in TRACE than the existing method [47]. The paired $T$-test showed significant differences between these two methods in both the average and Hausdorff distances, with $p$-values $0.0045$ and $0.003$ respectively in reproducibility and $p$-values $< 10^{-16}$ in robustness. For the comparison with manually labeled primary curves [50], [51], the MRIs Surfaces Curves dataset (http://sipi.usc.edu/~ajoshi/MSC) consisting of 12 subjects was used. The mean values of the average and Hausdorff distances of the 26 primary curves are $1.32$ and $3.77$ mm in the TRACE method, which are smaller than $1.38$ and $4.20$ mm in [47]. Even though the paired $T$-test found no significant difference between the two methods in the average distance ($p$-value $= 0.0713$), we found significant difference in the Hausdorff distance ($p$-value $= 7.3 \times 10^{-6}$).

### C. Diffusion Maps on Sulcal and Gyral Curves

The junctions between sulci are highly variable [52]. A sulcus corresponding to a long elementary fold in one subject may be made up of several small elementary folds in another subject [5]. Each subject has different number of vertices and edges in sulcal and gyral graphs, and they don’t exactly match across subjects even after registration. It is difficult to directly compare such graphs at the vertex level across subjects. Thus, the proposed polynomial approximation was used to smooth out the sulcal and gyral curves and obtain the smooth representation of curves that enables the vertex-level comparisons.

The extracted gyral curves were assigned heat value $1$, and sulcal curves were assigned heat value $-1$. All other parts of surface mesh vertices were assigned value $0$. Then heat diffusion was performed on these values. The diffusion map values range from $-1$ to $1$. The close to the value of $1$ indicates the likelihood of the gyral curves while the close to the value of $-1$ indicates the likelihood of the sulcal curves. The proposed method is motivated by the voxel-based morphometry point detection, concave points are initially obtained from the vertices of the input surface mesh by thresholding mean curvatures. The concave points are further filtered by employing the line simplification method [45] that simplifies the sulcal regions without significant loss of their morphological details. For curve delineation, the selected sulcal points are connected to form a graph, and the curves are delineated by tracing shortest paths on the graph. Finally, the sulcal curves are traced over the graph by the Dijkstra’s algorithm [46]. We use similar idea to gyral curve extraction by finding convex regions.

The TRACE method only identified the major gyral and sulcal curves. Minor curves in almost flat regions like plateau or with very low curvature, shallow depth or short length were not extracted. Fig. 7 displays the sulcal and gyral curves and the smoothed mean curvature of four subjects. In the enlarged regions, the first three subjects have no sulcal curves between the two gyral curves due to very low mean curvature, while the fourth subject has sulcal curve in the same region because of higher mean curvature.

The TRACE method was validated using the Kirby reproducibility dataset with 21 T1-weighted scans [48]. The reproducibility was measured by the distance between two corresponding surfaces (scan and re-scan sessions). The robustness to noise compared to [47] was done using synthetic noisy surfaces, which were generated by adding vertex-wise random displacements to the original surfaces. The displacement at each vertex follows an independent and identically distributed uniform distribution between 0 and 1.0 mm. We used the average and Hausdorff distances [49]. The experimental results from [44] (Table I) show higher reproducibility and robustness to noise in TRACE than the existing method [47]. The paired $T$-test showed significant differences between these two methods in both the average and Hausdorff distances, with $p$-values $0.0045$ and $0.003$ respectively in reproducibility and $p$-values $< 10^{-16}$ in robustness. For the comparison with manually labeled primary curves [50], [51], the MRIs Surfaces Curves dataset (http://sipi.usc.edu/~ajoshi/MSC) consisting of 12 subjects was used. The mean values of the average and Hausdorff distances of the 26 primary curves are $1.32$ and $3.77$ mm in the TRACE method, which are smaller than $1.38$ and $4.20$ mm in [47]. Even though the paired $T$-test found no significant difference between the two methods in the average distance ($p$-value $= 0.0713$), we found significant difference in the Hausdorff distance ($p$-value $= 7.3 \times 10^{-6}$).
(VBM) [53], [54], where the segmented white or gray matter regions are compared in 3D volume. Due to the difficulty of exactly aligning the white or gray matter regions separately, Gaussian kernel smoothing with large bandwidth was used to mask the shape variations across subjects and approximately align the segmented regions. Also a similar approach was used in the tract-based spatial statistics (TBSS) [55], [56] in analyzing white matter regions in diffusion tensor imaging that does not exactly align across subjects.

In the numerical implementation, a sufficient large expansion degree \( m = 1000 \) were used. In a desktop with 4.2 GHz Intel Core i7 processor, the construction of the discrete LB-operator took 5.76 seconds, the computation of Chebyshev coefficients took \( 2.6 \times 10^{-4} \) seconds, and heat diffusion by the Chebyshev polynomials took 3.19 seconds for the both hemispheres in average. The total computation took 8.95 seconds per subject in average. The diffusion maps were then subsequently used in localizing the male and female differences. One example of diffusion map is displayed in Fig. 8-left.

### D. Univariate Two-Sample T-Test

The diffusion maps with \( \sigma = 0.001 \) were constructed for 268 females 176 males. The average diffusion maps in Fig. 9-left displays the major differences in the temporal lobe, which is responsible for processing sensory input into derived meanings for the appropriate retention of visual memory, language comprehension, and emotion association [57].

The two-sample \( T \)-statistics maps are in the range of \([-6.5, 7.02]\). Any \( T \)-statistic with absolute value above 2.75 (red and blue regions) is considered as statistically significant using the false discovery rate (FDR) at 0.05. If the \( T \)-statistic map shows high \( T \)-statistic value at a particular vertex, it indicates that one group has consistently more gyral curves than sulcal curves at the vertex. If we use slightly different diffusion time \( \sigma \), we still obtain similar results.

We did an additional analysis using the mean curvature and shape index (SI). We estimated the curvatures and SI, which are the functions of curvature, by fitting the local quadratic surface in the first neighboring vertices [53], [58]

\[
f(x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 y_2 + \frac{1}{2} \beta_3 x_1^2 + \beta_4 x_1 y_2 + \frac{1}{2} \beta_5 y_2^2.
\]

The curvature and SI are expected to be noisy and require smoothing to increase statistical sensitivity and the signal-to-noise ratio [6], [7], [59] (Fig. 8). Smoothing surface data before statistical analysis is often done in various cortical surface features. Even the FreeSurfer package output the smoothed mean and Gaussian curvatures [60], [61]. Fig. 8 displays the results of smoothed mean curvature and SI maps.

We performed the two-sample \( T \)-test on the smoothed mean curvature and SI maps (Fig. 9-middle and right). The results show significant gender difference mainly in the temporal lobe, consistent to the findings in the proposed sulcal/gyral curve analysis (Fig. 9-left).

### E. Multivariate Two-Sample T-Test

The iterative kernel convolution was used to compute the diffusion at different time points quickly. The values of diffusion at different time points were then used in constructing the multiscale features. In this study, we adopted 10 time points \( \sigma = 0.0005, 0.001, \cdots, 0.0045, 0.005 \). Fig. 10 shows the diffusion maps of one representative subject. At each vertex, the multiscale diffusion features are used to determine the significant difference between the females and males. We used the two-sample Hotelling’s \( T^2 \)-statistic, which is the multivariate generalization of the two-sample \( T \)-statistic [16], [62]. Fig. 13 shows the Hotelling’s \( T^2 \)-statistics and the corresponding \( p \)-values in the log-scale. The heat diffusion has \( T^2 \)-statistics in the range of \([0.13, 8.2] \) with minimum \( p \)-value \( 3.4 \times 10^{-12} \). Any \( T^2 \)-statistic above 2.28 (yellow and red regions) is considered as significant at FDR 0.05.

In comparison, we used the diffusion wavelet features [14]–[17] at ten different scales and showed that the proposed method can achieve similar performance in localizing signal

---

**Fig. 9.** The average diffusion maps of the sulcal/gyral curves (left), mean curvature (middle) and SI (right) of 268 females and 176 males displayed on the average surface template. The \( T \)-statistic maps (bottom) show the localized sulcal and gyral pattern, mean curvature and SI differences (female - male) thresholded at \( \pm 4.96 \) (uncorrected \( p \)-value of \( 10^{-6} \)), having significant difference mainly in the temporal lobe.
values at different time points are used in constructing the Hotelling’s $T^2$-statistic to contrast males and females.

The iterative convolution of subject 130114. At each vertex, 10 diffusion wavelet transform values at different scales are used in constructing the Hotelling’s $T^2$-statistic to contrast males and females.

regions as the wavelet features. The diffusion wavelet [14], [15], [17] has the similar algebraic form as the heat kernel:

$$ W_t(p, q) = \sum_{j=0}^{\infty} g(\lambda_j t) \psi_j(p) \psi_j(q). $$

The difference between the heat kernel and diffusion wavelet transform is the weight function $g$, which determines the spectral distribution.

Compared to the heat kernel, the weight function $g$ attenuating all low and high frequencies outside the passband makes the diffusion wavelet work as a band-pass filter. The wavelet transform of $f$ is then given by

$$ W_t \ast f(p) = \sum_{j=0}^{\infty} g(\lambda_j t) f_j(p), \quad f_j = \int_{\mathcal{M}} f(p) \psi_j(p) d\mu(p). $$

The proposed polynomial approximation scheme can be applied to the diffusion wavelet transform through expanding $g(\lambda t)$ by orthogonal polynomials. In this paper, we used the following cubic spine as $g(\lambda t)$ [14]

$$ g(x) = \begin{cases} x^{1-\alpha} x^\alpha, & x < x_1 \\ -5 + 11x - 6x^2 + x^3, & x_1 \leq x \leq x_2 \\ 2^{1-\beta} x^\beta, & x > x_2 \end{cases}, $$

where $\alpha = \beta = 2$, $x_1 = 1$ and $x_2 = 2$. The scaling parameter $t$ controls the passband of the diffusion wavelet (Fig. 11).

Diffusion at different diffusion time $\sigma$ and diffusion wavelets at different scaling parameter $t$ contain different spectral information of input data $f$ (Fig. 11). Thus, the heat diffusion with a varying $\sigma$ and diffusion wavelet with varying $t$ provide multiscale features of $f$. All the heat diffusion features contain low-frequency components. If the initial surface data suffer from significant low-frequency noise, the diffusion wavelet transform would be more suitable. On the other hand, if most noises are in high frequencies, performance of the both methods would be similar and we do not really need diffusion wavelet features [63].

In this study, we adopted 10 different values of $t = 0.002, 0.003, \ldots, 0.011$. Fig. 12 shows the flattened diffusion wavelet maps of one representative subject. The values of $t$ were chosen empirically to match the amount of smoothing (FWHM) in the wavelet to the amount of smoothing in heat diffusion. Using the two-sample Hotelling’s $T^2$-statistic on the multiscale diffusion wavelet features, we also contrasted 268 females and 176 males. Fig. 13 shows the Hotelling’s $T^2$-statistics and the corresponding $p$-values in the log-scale. The diffusion wavelet transform has $T^2$-statistics in the range of $[0.09, 7.6]$ with minimum $p$-value $3.4 \times 10^{-11}$. For multiple comparisons, any $T^2$-statistic above 2.37 (yellow and red regions) is considered as significant using the FDR 0.05. Although there are slight differences, the both methods show the similar localization of sulcal and gyral graph patterns, mainly in the temporal lobe.

The exponential weight in the heat diffusion has only one parameter, i.e., the diffusion time $\sigma$, and leads to the analytic closed-form solutions to the expansion coefficients. The weight function in the diffusion wavelet transform is more complicated, and it may not possible to derive the closed-form expression for the expansion coefficients. The simpler weight function in heat kernel and the iterative convolution scheme lead to faster computational run time compared to the diffusion wavelets. In heat diffusion, we only needed to compute the
expansion coefficients for $\sigma = 0.0005$ and reused these coefficients in the iterative convolution to obtain the other nine features. The computation of the 1000 degree expansion coefficients by the proposed closed-form solution costed only $2.6 \times 10^{-4}$ seconds. In the diffusion wavelet transform, due to the more complicated weight function, the 1000 degree expansion coefficients were computed numerically, which took 1.26 seconds [14], [16].

**F. Comparing to Other Cortical Folding Features**

We computed the correlations between the diffusion maps of the sulcal/gyral curves and the mean curvature, SI and cortical thickness across all subjects with diffusion time $\sigma = 0.001$ [64], [65] (Fig. 14). We can observe strongly negative correlation to mean curvature $(-0.77 \pm 0.12)$ and strongly positive correlation to SI $(0.76 \pm 0.13)$. Although the correlation to the cortical thickness $(0.23 \pm 0.20)$ is not as high as the correlation to the mean curvature and SI, many regions have correlation value larger than 0.4, especially in the temporal lobe. Most of regions are all statistically significant after FDR correction at 0.05.

**IV. DISCUSSION AND CONCLUSION**

The numerical computation of solving diffusion equations has been thoroughly investigated and mostly solved using the finite element method (FEM) and finite difference method (FDM) on triangulated surface meshes for many decades [4]–[7], [9]–[11]. It is extremely difficult if not impossible to speed up the computation over existing methods any further. Utilizing the proposed spectral decomposition of heat kernel, we were able to come up with a new numerical scheme that speeds up the computation 8-40 times (depending on the mesh size) over these existing methods, which is a significant contribution itself.

Due to the advancement of imaging techniques, we are beginning to see larger meshes. For instance, [7] smoothed the mean curvature of triangulated cortical surfaces with 1.4 million vertices. Fan et al. [66] computed the heat flux signature in a cortical tetrahedral mesh with more than 1.5 million vertices. Kay et al. [67] used six parallel surfaces between the pial and white surfaces with 5 million vertices in modeling fMRI BOLD activity patterns at sub-millimeter resolution. Warner et al. [68] generated 3D tetrahedral head and cortical surface meshes with 2.7 million vertices to build high-resolution head and brain computer models for fMRI and EEG. Thus, the increase of computational run time would be of great interest.

In the sulcal and gyral graph pattern analysis, the sulcal and gyral curves were assigned value -1 and 1 respectfully. In the regions of higher diffusion value close to 1, there are more gyral curves than sulcal curves. In the regions of lower diffusion value close to -1, there are likely more sulcal curves than gyral curves. If a group consistently higher diffusion value in a particular region, it indicates there are likely to be more gyral curves in that region. The regions of interwinding complex sulcal and gyral curves will result in diffusion maps close to 0 (Fig. 14). Thus, the statistically significant group differences are not due to the complexity of interwinding sulcal/gyral patterns but the consistent concentration of more gyral or sulcal curves.

We found that the differences are mainly in the temporal lobe, especially in the superior temporal gyrus and sulcus, which is consistent with the literature. Harasty et al. [69] reported that females have proportionally larger language areas compared to males, such as the superior temporal cortex and Broca’s area. Ochiai et al. [24] reported statistically differences between males and females in the right superior temporal sulcus and the most posterior point and center of the left superior temporal sulcus. The significant gender differences in sulcal width and depth were reported in the superior temporal, collateral, and cingulate sulci in [70]. Also, there were significant gender differences in the cortical area of the left frontal lobe and in the gyrification index of the right temporal lobe [71]. Luders et al. [72] detected higher gray matter concentrations in females in the left posterior superior temporal gyrus and left inferior frontal gyrus. Crespo-Facorro et al. [73] found significant differences between males and females in the sulcal curvature index of the temporal and occipital lobes. Lyu et al. [28] reported that females showed higher gyration in the superior temporal, right inferior frontal, and parieto-occipital sulcal regions. In [74], the mean curvature of the left superior temporal sulcus was identified as a highly discriminative feature of sex classification. The consistent result with previous studies shows that the sulcal/gyral curves are reliable cortical surface features.

Sulcal and gyral curves can be interpreted with respect to cortical folding. The cortical folding is usually measured by the mean curvature and SI [60], [75]. Our result show that the sulcal and gyral curves are almost linearly related to the existing mean curvature and SI. This is the reason that we got the similar statistical results in all three methods. The cortical folding is known to correlate to cortical thickness. The gyri are thicker than the sulci [76]–[78]. Observing higher diffusion value at a vertex implies that the vertex is closer to gyri than sulci, and thus larger thickness is expected at the vertex.
In the cortical growth and folding development in human fetal brains, many studies have reported changes in surface and shape features such as the curvatures, sulcal depth, gyri- fication index, sulcal pit based graphs and sulcal skeletons [23], [79]–[83]. At 25 weeks, the cortical surface is still very smooth [81]. There are few major sulcal and gyral curves, and most surface vertices will have heat diffusion values close to zero. With increasing gestational age, the cortical folds become more complex with more sulcal and gyral curves and branches, which will likely result in higher variability in diffusion values across vertices.

The proposed general polynomial approximation of the Laplace-Beltrami (LB) operator works for an arbitrary orthogonal polynomial. The proposed polynomial expansion method speeds up the computation compared to existing numerical schemes for diffusion equations. Our method avoids various numerical issues associated with the LB-eigenfunction method and FEM based diffusion solvers. The proposed fast and accurate scheme can be further extended to any arbitrary domain without much computational bottlenecks. Thus, the method rate scheme can be further extended to any arbitrary domain numerical issues associated with the LB-eigenfunction method and non-orthogonal polynomial. The proposed polynomial expansion method which will likely result in higher variability in diffusion values existing methods may not be applicable without additional computational resources. Beyond the sulcal and gyral graph analysis on 2D surface meshes, the proposed method can be applied to 3D volumetric meshes [84]. This is left as a future study.

REFERENCES


