Manifold parametrizations by eigenfunctions of the Laplacian and heat kernels

Peter W. Jones*, Mauro Maggioni†‡, and Raanan Schul§

*Department of Mathematics, Yale University, 10 Hillhouse Avenue, New Haven, CT 06510; †Department of Mathematics, Duke University, Box 90320, Durham, NC 27708; and §Department of Mathematics, University of California, Box 951555, Los Angeles, CA 90095-1555

Communicated by Ronald R. Coifman, Yale University, New Haven, CT, October 29, 2007 (received for review September 3, 2007)

We use heat kernels or eigenfunctions of the Laplacian to construct local coordinates on large classes of Euclidean domains and Riemannian manifolds (not necessarily smooth, e.g., with $\mathcal{C}^\alpha$ metric). These coordinates are bi-Lipschitz on large neighborhoods of the domain or manifold, with constants controlling the distortion and the size of the neighborhoods that depend only on natural geometric properties of the domain or manifold. The proof of these results relies on novel estimates, from above and below, for the heat kernel and its gradient, as well as for the eigenfunctions of the Laplacian and their gradient, that hold in the non-smooth category, and are stable with respect to perturbations within this category. Finally, these coordinate systems are intrinsic and efficiently computable, and are of value in applications.

spectral geometry | nonlinear dimensionality reduction

In many recent applications, one attempts to find local parametrizations of data sets. A recurrent idea is to approximate a high dimensional data set, or portions of it, by a manifold of low dimension. A variety of algorithms for this task have been proposed (1–8). Unfortunately, such techniques seldomly come with guarantees on their capabilities of indeed finding local parametrization (but see, for example, refs. 8 and 9) or on quantitative statements on the quality of such parametrizations. Examples of such disparate applications include document analysis, face recognition, clustering, machine learning (10–13), nonlinear image denoising and segmentation (11), processing of articulated images (8), and mapping of protein energy landscapes (14). It has been observed in many cases that the eigenfunctions of a suitable graph Laplacian on a data set provide robust local coordinate systems and are efficient in dimensional reduction (1, 4, 5). The purpose of this paper is to provide a partial explanation for this phenomenon by proving an analogous statement for manifolds as well as introducing other coordinate systems via heat kernels, with even stronger guarantees. Here, we should point out the 1994 paper of Bérard *et al.* (15) where a weighted infinite sequence of eigenfunctions is shown to provide a global coordinate system. (Points in the manifold are mapped to ℓ_2.) To our knowledge, this was the first result of this type in Riemannian geometry. If a given data set has a piece that is statistically well approximated by a low dimensional manifold, it is then plausible that the graph eigenfunctions are well approximated by the Laplace eigenfunctions of the manifold. One of our results is that, with the normalization that the volume of a d-dimensional manifold $\mathcal{M}$ equals one, any suitably embedded ball $B_r(z)$ in $\mathcal{M}$ has the property that one can find (exactly) d eigenfunctions that are a "robust"coordinate system on $B_{cr}(z)$ (for a constant c depending on elementary properties of $\mathcal{M}$). In addition, these eigenfunctions, which depend on z and r, "blow up" the ball $B_{cr}(z)$ to diameter at least one. In other words, one can find d eigenfunctions that act as a "microscope" on $B_{cr}(z)$ and "magnify" it up to size ~ 1. Another of our results is as follows. We introduce simple "heat coordinate" systems on manifolds. Roughly speaking (and in the language of the previous paragraph), these are d choices of manifold heat kernels that form a robust coordinate system on $B_{cr}(z)$. We call this method "heat triangulation" in analogy with triangulation as practiced in surveying, cartography, navigation, and modern GPS. Indeed, our method is a simple translation of these classical triangulation methods.

The embeddings we propose can be computed efficiently and therefore, together with the strong guarantees we prove, are expected to be useful in a variety of applications, from dimensionality reduction to data set compression and navigation.

Given these results, it is plausible to guess that analogous results should hold for a local piece of a data set if that piece has in some sense a "local dimension" approximately d. There are certain difficulties with this philosophy. The first is that graph eigenfunctions are global objects and any definition of "local dimension" must change from point to point in the data set. A second difficulty is that our manifold results depend on classical estimates for eigenfunctions. This smoothness is often lacking in graph eigenfunctions.

For data sets, *heat triangulation* is a much more stable object than eigenfunction coordinates because

- heat kernels are local objects;
- if a manifold $\mathcal{M}$ is approximated by discrete sets X, the corresponding graph heat kernels converge rather nicely to the manifold heat kernel (4, 5);
- one has good statistical control on smoothness of the heat kernel, simply because one can easily examine it and because one can use the Hilbert space $\{f \in L^2 : \nabla f \in L^2\}$;
- our results that use eigenfunctions rely in a crucial manner on Weyl's lemma, whereas heat kernel estimates do not.

The philosophy used in this paper is as follows.

Step 1. Find suitable points y_j, $1 \leq j \leq d$ and a time t so that the mapping given by heat kernels

$$(x \to K_t(x, y_1), \ldots, K_t(x, y_d))$$

is a good local coordinate system on $B(z, cr)$. (This is heat triangulation.)

Step 2. Use Weyl's lemma to find suitable eigenfunctions φ_{i_j} so that (with $K_j(x) = K_t(x, y_j)$) one has $\nabla\varphi_{i_j}(x) \approx c_j \nabla K_j(x)$, $x \in B(z, cr)$ for an appropriate constant c.

Results

Euclidean Domains. We first present the case of Euclidean domains. Although our results in this setting follow from the more general results for manifolds discussed in the next section, the case of Euclidean domains is of independent interest, and the exposition of the theorem is simpler.

We consider the heat equation in Ω, a finite volume domain in $\mathbb{R}^d$, with either Dirichlet or Neumann boundary conditions:

Author contributions: P.W.J., M.M., and R.S. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

‡To whom correspondence should be addressed. E-mail: mauro.maggioni@duke.edu.

$$\begin{cases}\left(\Delta - \frac{\partial}{\partial t}\right)u(x,t) = 0 \\ u|_{\partial\Omega} = 0\end{cases} \quad \text{or} \quad \begin{cases}\left(\Delta - \frac{\partial}{\partial t}\right)u(x,t) = 0 \\ \partial_\nu u|_{\partial\Omega} = 0\end{cases}$$

Here, ν is the outer normal on $\partial\Omega$. Independently of the boundary conditions, we will denote by Δ the Laplacian on Ω. For the purpose of this paper, in both the Dirichlet and Neumann case, we restrict our study to domains where the spectrum is discrete and the corresponding heat kernel can be written as

$$K_t^{\Omega}(z,w) = \sum_{j=0}^{+\infty} \varphi_j(z)\varphi_j(w)e^{-\lambda_j t}, \qquad \textbf{[0.1]}$$

where the $\{\varphi_j\}$ form an orthonormal basis of eigenfunctions of Δ, with eigenvalues $0 \le \lambda_0 \le \cdots \le \lambda_j \le \ldots$. We also require that the following Weyl's estimate holds, i.e., there is a constant $C_{Weyl,\Omega}$ such that for any $T > 0$

$$\#\{j : \lambda_j \le T\} \le C_{Weyl,\Omega}\, T^{d/2}|\Omega|. \qquad \textbf{[0.2]}$$

(This condition is always satisfied in the Dirichlet case, where in fact $C_{Weyl,\Omega}$ can be chosen independent of Ω.) It should however be noted that these conditions are not always true and the Neumann case is especially problematic (16, 17).

Theorem 1. Embedding via Eigenfunctions, for Euclidean Domains. *Let Ω be a domain in $\mathbb{R}^d$ satisfying all the conditions above, rescaled so that $|\Omega| = 1$. There are constants $c_1, \ldots, c_6 > 0$ that depend only on d and $C_{Weyl,\Omega}$, such that the following hold. For any $z \in \Omega$, let $R_z \le \operatorname{dist}(z, \partial\Omega)$. Then there exist $i_1, \ldots, i_d$ and constants $c_6 R_z^{d/2} \le \gamma_1 = \gamma_1(z), \ldots, \gamma_d = \gamma_d(z) \le 1$ such that*

(a) the map

$$\Phi : B_{c_1 R_z}(z) \to \mathbb{R}^d \qquad \textbf{[0.3]}$$

$$x \to (\gamma_1\varphi_{i_1}(x), \ldots, \gamma_d\varphi_{i_d}(x)) \qquad \textbf{[0.4]}$$

satisfies, for any $x_1, x_2 \in B(z, c_1 R_z)$,

$$\frac{c_2}{R_z}\|x_1 - x_2\| \le \|\Phi(x_1) - \Phi(x_2)\| \le \frac{c_3}{R_z}\|x_1 - x_2\|; \qquad \textbf{[0.5]}$$

(b) the associated eigenvalues satisfy

$$c_4 R_z^{-2} \le \lambda_{i_1}, \ldots, \lambda_{i_d} \le c_5 R_z^{-2}.$$

Remark 1. The dependence on the constant $C_{Weyl,\Omega}$ is only needed in the Neumann case.

Manifolds with C^α Metric. The results above can be extended to certain classes of manifolds. To formulate a result corresponding to *Theorem 1*, we must first carefully define the manifold analogue of $\operatorname{dist}(z, \partial\Omega)$. Let $\mathcal{M}$ be a smooth, d-dimensional compact manifold, possibly with boundary. Suppose we are given a metric tensor g on $\mathcal{M}$ is C^α for some $\alpha > 0$. For any $z_0 \in \mathcal{M}$, let (U, u) be a coordinate chart such that $z_0 \in U$ and

(*i*) $g^{il}(u(z_0)) = \delta^{il}$;
(*ii*) for any $x \in U$, and any $\xi, \nu \in \mathbb{R}^d$,

$$c_{\min}(g)\|\xi\|^2_{\mathbb{R}^d} \le \sum_{i,j=1}^{d} g^{ij}(u(x))\xi_i\xi_j \quad \sum_{i,j=1}^{d} g^{ij}(u(x))\xi_i\nu_j \le c_{\max}(g)\|\xi\|_{\mathbb{R}^d}\|\nu\|_{\mathbb{R}^d}. \qquad \textbf{[0.6]}$$

We let $r_{\mathcal{M}}(z_0) = \sup\{r > 0 : B_r(u(z_0)) \subseteq u(U)\}$. Observe that, when g is at least C^2, $r_{\mathcal{M}}$ can be taken to be the inradius, with local coordinate chart given by the exponential map at z. We denote by $\|g\|_{\alpha\wedge 1}$ the maximum over all i, j of the $\alpha \wedge 1$-Hölder norm of g^{ij} in the chart (U, u). The natural volume measure $d\mu$ on the manifold is given, in any local chart, by $\sqrt{\det g}$; conditions **0.6** guarantee that $\det g$ is uniformly bounded below from 0. Let $\Delta_{\mathcal{M}}$ be the Laplace Beltrami operator on $\mathcal{M}$ In a local chart, we have

$$\Delta_{\mathcal{M}} f(x) = -\frac{1}{\sqrt{\det g}} \sum_{i,j=1} \partial_j(\sqrt{\det g}\, g^{ij}(u(x))\partial_i f)(u(x)), \qquad \textbf{[0.7]}$$

where (g^{ij}) is the inverse of g_{ij}. Conditions **0.6** are the usual uniform ellipticity conditions for the operator **0.7**. With Dirichlet or Neumann boundary conditions, $\Delta_{\mathcal{M}}$ is self-adjoint on $L^2(\mathcal{M}, \mu)$. We will assume that the spectrum is discrete, denote by $0 \le \lambda_0 \le \cdots \le \lambda_j \le$ its eigenvalues and by $\{\varphi_j\}$ the corresponding orthonormal basis of eigenfunctions, and write Eqs. **0.1** and **0.2** with Ω replaced by $\mathcal{M}$.

Theorem 2. *Let $(\mathcal{M}, g)$, $z \in \mathcal{M}$ and (U, u) be as above. Also, assume $|\mathcal{M}| = 1$. There are constants $c_1, \ldots, c_6 > 0$, depending on d, $c_{\min}$, $c_{\max}$, $\|g\|_{\alpha\wedge 1}$, $\alpha \wedge 1$, and $C_{Weyl,\mathcal{M}}$, such that the following hold. Let $R_z = r_{\mathcal{M}}(z)$. Then there exist $i_1, \ldots, i_d$ and constants $c_6 R_z^{d/2} \le \gamma_1 = \gamma_1(z), \ldots, \gamma_d = \gamma_d(z) \le 1$ such that*

(a) the map

$$\Phi : B_{c_1 R_z}(z) \to \mathbb{R}^d \qquad \textbf{[0.8]}$$

$$x \to (\gamma_1\varphi_{i_1}(x), \ldots, \gamma_d\varphi_{i_d}(x)) \qquad \textbf{[0.9]}$$

such that for any $x_1, x_2 \in B(z, c_1 R_z)$

$$\frac{c_2}{R_z} d_{\mathcal{M}}(x_1, x_2) \le \|\Phi(x_1) - \Phi(x_2)\| \le \frac{c_3}{R_z} d_{\mathcal{M}}(x_1, x_2). \qquad \textbf{[0.10]}$$

(b) the associated eigenvalues satisfy

$$c_4 R_z^{-2} \le \lambda_{i_1}, \ldots, \lambda_{i_d} \le c_5 R_z^{-2}.$$

Remark. In both *Theorem 1* and *Theorem 2*, the constants γ_j are given by

$$\gamma_j = \left(\fint_{B_{cR_z}(z)} \varphi_{i_j}(z')^2 dz' \right)^{-1/2},$$

where c is some fixed constant, depending on d, $c_{\min}$, $c_{\max}$, $\|g\|_{\alpha\wedge 1}$, $\alpha \wedge 1$, and C_{Weyl}.

Remark 2. The constant $C_{Weyl,\mathcal{M}}$ is only needed in the Neumann case.

Remark 3. Most of the proof is done on one local chart containing z which we choose (one which contains a large enough ball around z). An inspection of the proof shows that we use only the norm $\|g\|_{\alpha\wedge 1}$ of the g restricted to this chart. In particular, the theorem holds also for $R_z \le r_{\mathcal{M}}(z)$.

Remark 4. When rescaling *Theorem 2,* it is important to note that if f is a Hölder function with $\|f\|_{C^{\alpha\wedge 1}} = A$ and $f_r(z) = f(r^{-1}z)$, then $\|f_r\|_{C^{\alpha\wedge 1}} = Ar^{\alpha\wedge 1}$. Since we will have $r < 1$, f_r satisfies a better Hölder estimate then f, i.e., $\|f_r\|_{C^{\alpha\wedge 1}} = Ar^{\alpha\wedge 1} \le A = \|f\|_{C^{\alpha\wedge 1}}$.

Remark 5. We do not know, in both *Theorem 1* and *Theorem 2,* whether it is possible to choose eigenfunctions such that $\gamma_1 = \ldots = \gamma_d$.

Another result is true. One may replace the d chosen eigenfunctions above by d chosen heat kernels, i.e., $\{K_t(z, y_i)\}_{i=1,\ldots,d}$. In fact, such heat kernels arise naturally in the main steps of the proofs of *Theorem 1* and *Theorem 2.* This leads to an embedding map with even stronger guarantees:

Theorem 3. Heat Triangulation Theorem. *Let* $(\mathcal{M}, g)$, $z \in \mathcal{M}$ *and* (U, u) *be as above, where we now allow* $|\mathcal{M}| = +\infty$. *Let* $R_z \leq \min\{1, r_{\mathcal{M}}(z)\}$. *Let* $p_1, \ldots, p_d$ *be d linearly independent directions. There are constants* $c_1, \ldots, c_6 > 0$, *depending on d,* $c_{\min}$, $c_{\max}$, $\|g\|_{\alpha\wedge 1}$, $\alpha \wedge 1$, *and the smallest and largest eigenvalues of the Gramian matrix* $(p_i, p_j)_{i,j=1,\ldots,d}$, *such that the following holds. Let* y_i *be so that* $y_i - z$ *is in the direction* p_i, *with* $c_4R_z \leq d_{\mathcal{M}}(y_i, z) \leq c_5R_z$ *for each* $i = 1, \ldots, d$ *and let* $t_z = c_6R_z^2$. *The map* $\Phi : B_{c_1R_z}(z) \to \mathbb{R}^d$, *defined by*

$$x \to (R_z^d K_{t_z}(x, y_1)), \ldots, R_z^d K_{t_z}(x, y_d)) \qquad \textbf{[0.11]}$$

satisfies, for any $x_1, x_2 \in B_{c_1R_z}(z)$,

$$\frac{c_2}{R_z} d_{\mathcal{M}}(x_1, x_2) \leq \|\Phi(x_1) - \Phi(x_2)\| \leq \frac{c_3}{R_z} d_{\mathcal{M}}(x_1, x_2).$$

This holds for the manifold and Euclidean case alike and depends only on estimates for the heat kernel and its gradient.

Remark 6. One may replace the (global) heat kernel above with a local heat kernel, i.e., the heat kernel for the ball $B(z, R_z)$ with the metric induced by the manifold and Dirichlet boundary conditions. In fact, this is a key idea in the proof of all of the above theorems.

Remark 7. All theorems hold for more general boundary condi-tions. This is especially true for the *Heat Triangulation Theorem,* which does not even depend on the existence of a spectral expansion for the heat kernel.

Example 1. It is a simple matter to verify this theorem for the case where the manifold in $\mathbb{R}^d$. For example, if $d = 2$, $R_z = 1$, and $z = 0$, $y_1 = (-1, 0)$ and $y_2 = (0, -1)$. Then if $K_t(x, y)$ is the Euclidean heat kernel,

$$x \to (K_1(x, y_1), K_1(x, y_2))$$

is a (nice) biLipschitz map on $B_{1/2}((0, 0))$. (The result for arbitrary radii then follows from a scaling argument.) This is because on can simply evaluate the heat kernel $K_t(x, y) = [1/(4\pi t)]e^{-(|x-y|^2)/4t}$. In $B_{1/2}((0, 0))$,

$$\nabla K_1(x, y_1) \sim \frac{1}{2\pi} e^{-1/4} (1, 0) \textit{ and } \nabla K_1(x, y_2) \sim \frac{1}{2\pi} e^{-1/4}(0, 1).$$

Notation. In what follows, we will write $f(x) \lesssim_{c_1,\ldots,c_n} g(x)$ if there exists a constant C depending only on $c_1, \ldots, c_n$, and not on f, g, or x, such that $f(x) \leq Cg(x)$ for all x (in a specified domain). We will write $f(x) \sim_{c_1,\ldots,c_n} g(x)$ if both $f(x) \lesssim_{c_1,\ldots,c_n} g(x)$ and $g(x) \lesssim_{c_1,\ldots,c_n} f(x)$. We will write $a \sim_{C_1}^{C_2} b$ for a, b vectors, if $a_i \sim_{C_1}^{C_2} b_i$ for all i.

The Proofs

The proofs in the Euclidean and manifold case are similar. In this section, we present the main steps of the proof. A full presentation is given in ref. 2. Some remarks about the manifold case:

(a) As mentioned in *Remark 3,* we will often restrict to working on a single (fixed) chart in local coordinates. When we discuss moving in a direction *p,* we mean in the local coordinates.

(b) Let us say a few words about how the dependence on $\|g\|_{\alpha\wedge 1}$ comes into play. Generally speaking, in all places except one (which we will mention momentarily), the $\alpha \wedge 1$-Hölder condition is used to get local bi-Lipschitz bounds on the perturbation of the metric (resp. the ellipticity constants) from the Euclidean metric (resp. the Laplacian). The place where one really uses the Hölder condition is an estimate on how much the gradient of a (global) eigenfunction changes in a ball.

(c) We will use Brownian motion arguments (on the manifold). To have existence and uniqueness, one needs smoothness assumptions on the metric (say, C^2). Therefore, we will first prove the theorem in the manifold case in the C^2 metric category, and then use perturbation estimates to obtain the result for $g^{ij} \in C^\alpha$. To this end, we will often have dependence on the C^α norm of the coordinates of the g^{ij} even though we will be (for a specific lemma or proposition) assuming the g has C^2 entries.

(d) We will use estimates from ref. 18. The theorems in ref. 18 are stated only for the case of $d \geq 3$. Our theorems are true also for the case $d = 2$ (and trivially, $d = 1$). This can beseen indirectly by considering $\tilde{\mathcal{M}} := \mathcal{M} \times \mathbb{T}$ and noting that the eigenfunctions of $\tilde{\mathcal{M}}$ and the heat kernel of $\tilde{\mathcal{M}}$ both factor.

The idea of the proof of *Theorem 1* and *2* is as follows. We start by fixing a direction p_1 at z. We would like to find an eigenfunction φ_{i_1} such that $|\partial_{p_1}\varphi_{i_1}| \gtrsim R_z^{-1}$ on $B_{c_1R_z}(z)$. To achieve this, we start by showing that the heat kernel has large gradient in an annulus of inner and outer radius $\sim R_z^{-1}$ around y_1 (y_1 chosen such that z is in this annulus, in direction p_1). We then show that the heat kernel and its gradient can be approximated on this annulus by as the partial sum of **0.1** over eigenfunctions φ_λ which satisfy both $\lambda \sim R_z^{-2}$ and $R_z^{-d/2} \|\varphi_\lambda\|_{L^2(B_{c_1R_z}(z))} \gtrsim 1$. By the pigeon-hole principle, at least one such eigenfunction, let it be φ_{i_1}, has a large partial derivative in the direction p_1. We then consider $\nabla\varphi_{i_1}$ and pick $p_2 \perp \nabla\varphi_{i_1}$ and by induction we select $\varphi_{i_1}, \ldots, \varphi_{i_d}$, making sure that at each stage we can find φ_{i_k}, not previously chosen, satisfying $|\partial_{p_k}\varphi_{i_k}| \sim R_z^{-1}$ on $B_{c_1R_z}(z)$. We finally show that the $\Phi := (\varphi_{i_1}, \ldots, \varphi_{i_d})$ satisfies the desired properties.

Step 1. Estimates on the heat kernel and its gradient. Let K be the Dirichlet or Neumann heat kernel on Ω or $\mathcal{M}$, corresponding to one of the Laplacian operators considered above associated with g. We have the spectral expansion

$$K_t(x, y) = \sum_{j=0}^{+\infty} e^{-\lambda_j t}\varphi_j(x)\varphi_j(y).$$

When working on a manifold, we can assume in what follows that we fix a local chart containing $B_{R_z}(z)$.

Assumption A.1. *Let the constants* $\delta_0, \delta_1 > 0$ *depend on d,* $c_{\min}$, $c_{\max}$, $\|g\|_{\alpha\wedge 1}$, $\alpha \wedge 1$. *We consider* $z, w \in \Omega$ *satisfying* $(\delta_1/2) R_z < t^{1/2} < \delta_1 R_z$ *and* $|z - w| < \delta_0 R_z$.

Proposition 4. *Under Assumption* A.1, *let* $g \in C^\alpha$, δ_0 *sufficiently small, and* δ_1 *is sufficiently small depending on* δ_0. *Then there are constants* $C_1, C_2, C_1', C_2', C_9 > 0$, *that depend on d,* δ_0, δ_1, $c_{\min}$, $c_{\max}$, $\|g\|_{\alpha\wedge 1}$, $\alpha \wedge 1\}$ *and* C_1', C_2', C_9 *dependent also on* C_{Weyl}, *such that the following hold:*

(i) the heat kernel satisfies

$$K_t(z, w) \sim_{C_1}^{C_2} t^{-d/2}; \qquad \textbf{[0.12]}$$

(ii) if $(1/2)\delta_0 R_z < |z - w|$, *p is a unit vector in the direction of* $z - w$, *and q is arbitrary unit vector, then*

$$|\nabla K_t(z, w) \sim_{C_1'}^{C_2'} t^{-d/2} \frac{R_z}{t} \quad \text{and} \quad |\partial_p K_t(z, w)| \sim_{C_1'}^{C_2'} t^{-d/2} \frac{R_z}{t} \qquad \textbf{[0.13]}$$

MATHEMATICS

$$\left|\partial_q K_t(z,w) - C_2'\left\langle q, \frac{z-w}{\|z-w\|}\right\rangle t^{-d/2}\frac{R_z}{t}\right| \le C_9 t^{-d/2}\frac{R_z}{t}, \quad \mathbf{[0.14]}$$

where $C_9 \to 0$ as $\delta_1 \to 0$ (with δ_0 fixed);

(iii) if in addition $g^{ij} \in C^2$, $(1/2)\delta_0 R_z < |z-w|$, and q is as above, then for $s \le t$,

$$K_s(z,w) \lesssim_{C_2} t^{-d/2}, \quad |\nabla K_s(z,w)| \lesssim_{C_2'} t^{-d/2}\frac{R_z}{t}$$

$$\text{and} \quad |\partial_q K_s(z,w)| \lesssim_{C_2'} t^{-d/2}\frac{R_z}{t}; \quad \mathbf{[0.15]}$$

(iv) C_1, C_2 both tend to a single function of $\{d, c_{\min}, c_{\max}, \delta_0, C_{Weyl}\}$, as δ_1 tends to 0 with δ_0 fixed;

(v) if $g \in C^2$, then also C_1', C_2', C_9 can be chosen independently of C_{Weyl}. Furthermore, the above estimates also hold for $|\mathcal{M}| \le +\infty$.

At this point we can side track and choose heat kernels $\{K_t(\cdot, y_i)\}_{i=1,\ldots,d}$, with $t \sim R_z^2$, that provide a local coordinate chart with the properties claimed in *Theorem 3.*

Proof of Theorem 3. We start with the case $g \in C^2$. Let us consider the Jacobian $\tilde{J}(x)$, for $x \in B_{c_1 R_z}(z)$, of the map

$$\tilde{\Phi} := R_z^{-d} t^{+d/2}(t/R_z^2)\Phi.$$

By **0.14** we have $|\tilde{J}_{ij}(x) - C_2'\langle p_i, (x-y_j)/\|x-y_j\|\rangle R_z^{-1}| \le C_9 R_z^{-1}$, with C_2' independent of C_{Weyl}. As dictated by *Proposition 4,* by choosing δ_0, δ_1 appropriately (and, correspondingly, c_1 and c_6), we can make the constant C_9 smaller than any chosen ϵ, for all entries, and for all x at distance no greater than $c_1 R_z$ from z, where we use $t = t_z = c_6 R_z^2$ for $\tilde{\Phi}$. Therefore, for c_1 small enough compared to c_4 we can write $R_z \tilde{J}(x) = G_d + E(x)$, where G_d is the Gramian matrix $\langle p_i, p_j\rangle$ (indepedent of x), and $|E_{ij}(x)| < \epsilon$, for $x \in B_{c_1 R_z}(z)$. This implies that $R_z^{-1}(\sigma_{\min} - C_d\epsilon)\|v\| \le \|\tilde{J}(x)v\| \le R_z^{-1}(\sigma_{\max} + C_d\epsilon)\|v\|$, with C_d depending linearly on d, where $\sigma_{\max}$ and $\sigma_{\min}$ are the largest and, respectively, smallest eigenvalues of G_d. At this point, we choose ϵ small enough, so that the above bounds imply that the Jacobian is essentially constant in $B_{c_1 R_z}(z)$, and by integrating along a path from x_2 to x_2 in $B_{c_1 R_z}(z)$, we obtain the Theorem (Φ and $\tilde{\Phi}$ differ only by scalar multiplication). We note that $\epsilon \sim 1/d$ suffices. To get the result when g is only C^α we use perturbation techniques for the heat kernel (2).

We proceed towards the proof of *Theorem 1* and *2.* The following steps aim at replacing appropriately chosen heat kernels by a set of eigenfunctions, by extracting the "leading terms"in their spectral expansion.

Step 2. Heat kernel and eigenfunctions. Let $\mathrm{Ave}_R^z(f) = (⨍_{B_R(z)} |f|^2)^{1/2}$. We record the following (2):

Proposition 5. *Assume $g^{ij} \in C^\alpha$. There exists $b_1 < 1$, that depends on d, $c_{\min}$, $c_{\max}$, $\|g\|_{\alpha\wedge 1}$, $\alpha \wedge 1$ such that the following holds. For an eigenfunction φ_j of $\Delta_{\mathcal{M}}$, corresponding to the eigenvalue λ_j, and $R \le R_z$, the following estimates hold. For $w \in B_{b_1 R}(z)$ and $x, y \in B_{b_1 R}(z)$,*

$$|\varphi_j(w)| \lesssim P_1(\lambda_j R^2)\mathrm{Ave}_R^z(\varphi_j)$$

$$\|\nabla\varphi_j(w)\| \lesssim R^{-1}P_2(\lambda_j R^2)\mathrm{Ave}_R^z(\varphi_j)$$

$$\frac{\|\nabla\varphi_j(x) - \nabla\varphi_j(y)\|}{\|x-y\|^{\alpha\wedge 1}} \lesssim R^{-1-\alpha\wedge 1}P_3(\lambda_j R^2)\mathrm{Ave}_R^z(\varphi_j)$$

with constants depending only on d, $c_{\max}$, $c_{\min}$, $\|g^{ij}\|_{\alpha\wedge 1}$, and $P_1(x) = (1+x)^{(1/2)+\beta}$, $P_2(x) = (1+x)^{(3/2)+\beta}$, $P_3(x) = (1+x)^{(5/2)+\beta}$, with β the smallest integer larger than or equal to $(d-2)/4$.

We start by restricting our attention to eigenfunctions do not have too high frequency. Let $\Lambda_L(A) = \{\lambda_j : \lambda_j \le At^{-1}\}$ and $\Lambda_H(A') = \{\lambda_j : \lambda_j > A't^{-1}\} = \Lambda_L(A')^c$.

A first connection between the heat kernel and eigenfunctions is given by the following truncation Lemma.

Lemma 1. *Assume $g \in C^2$. Under Assumption A.1, for $A > 1$ large enough and $A' < 1$ small enough, depending on δ_0, δ_1, C_1, C_2, C_1', C_2' (as in Proposition 4), there exist constants C_3, C_4 (depending on A, A' as well as $\{d, c_{\min}, c_{\max}, \|g\|_{\alpha\wedge 1}, \alpha\wedge 1\}$) such that*

(i) The heat kernel is approximated by the truncated expansion

$$K_t(z,w) \sim_{C_3}^{C_4} \sum_{j\in\Lambda_L(A)} \varphi_j(z)\varphi_j(w)e^{-\lambda_j t}.$$

(ii) If $(1/2)\delta_0 R_z < |z-w|$ and p is a unit vector parallel to $z-w$, then

$$\nabla_w K_t(z,\cdot) \sim_{C_3}^{C_4} \sum_{j\in\Lambda_L(A)\cap\Lambda_H(A')} \varphi_j(z)\nabla_w\varphi_j(\cdot)e^{-\lambda_j t}$$

$$\partial_p K_t(z,\cdot) \sim_{C_3}^{C_4} \sum_{j\in\Lambda_L(A)\cap\Lambda_H(A')} \varphi_j(z)\partial_p\varphi_j(\cdot)e^{-\lambda_j t}.$$

(iii) C_3, C_4 both tend to 1 as $A \to \infty$ and $A' \to 0$.

This lemma implies that in the heat kernel expansion, we do not need to consider eigenfunctions corresponding to eigenvalues larger than At^{-1}. However, in our search for eigenfunctions with the desired properties, we need to restrict our attention further, by discarding eigenfunctions that have too small a gradient around z. As a proxy for gradient, we use local energy. Recall $\mathrm{Ave}_R^z(f) = (⨍_{B_R(z)} |f|^2)^{1/2}$, and let

$$\Lambda_E(z, R_z, \delta_0, c_0) := \{\lambda_j \in \sigma(\Delta) : \mathrm{Ave}^z_{\frac{1}{2}\delta_0 R_z}(\varphi_j) \ge c_0\}.$$

The truncation *Lemma 1* can be strengthened into

Lemma 2. *Assume $g \in C^2$. Under Assumption A.1, for C_3, C_4 close enough to 1 (as in Lemma 1), and c_0 small enough (depending on d, $c_{\min}$, $c_{\max}$, $\|g\|_{\alpha\wedge 1}$, $\alpha\wedge 1$, and $C_{Weyl,\mathcal{M}}$), there exist constants C_5, C_6 (depending only on C_3, C_4, c_0, and $C_{Weyl,\mathcal{M}}$) such that the heat kernel satisfies*

$$K_t(z,w) \sim_{C_5}^{C_6} \sum_{\lambda_j\in\Lambda_L(A)\cap\Lambda_E(z,R_z,\delta_0,c_0)} \varphi_j(z)\varphi_j(w)e^{-\lambda_j t}$$

and if $(1/2)\delta_0 R_z < |z-w|$, then, if $\Lambda := \Lambda_L(A) \cap \Lambda_H(A') \cap \Lambda_E(z, R_z, \delta_0, c_0)$,

$$\partial_p K_t(z,w) \sim_{C_5}^{C_6} \sum_{\lambda_j\in\Lambda} \varphi_j(z)\partial_p\varphi_j(w)e^{-\lambda_j t}.$$

C_5, C_6 tend to 1 as C_3, C_4 tend to 1 and c_0 tends to 0.

Step 3. Choosing appropriate eigenfunctions.

The set of eigenfunctions with eigenvalues in Λ (as in *Lemma 2*) is well suited for our purposes, in view of:

Lemma 3. *Assume $g \in C^2$. Under Assumption A.1, for δ_0 small enough, there exists a constant C_7 depending on $\{C_1, C_2, C_1', C_2', C_5, \delta_1\}$ and C_8 depending on $\{\delta_0, c_{\min}, c_{\max}, \|g\|_{\alpha\wedge 1}, \alpha\wedge 1\}$ such that the following holds. For any direction p there exist $j \in \Lambda := \Lambda_L(A) \cap\Lambda_H(A') \cap\Lambda_E(z, R_z, \delta_0, c_0)$ such that*

$$|\partial_p\varphi_j(z)| \sim_{C_7}^{C_8} R_z^{-1}\,\mathrm{Ave}^z_{\frac{1}{2}\delta_u R_z}\,\varphi_j,$$

and moreover, if $\|z - z'\| \le b_1 R_z$, where b_1 is a constant that depends on C_7, C_8, d, $c_{\min}$, $c_{\max}$, $\|g\|_{\alpha\wedge 1}$, $\alpha\wedge 1$, then

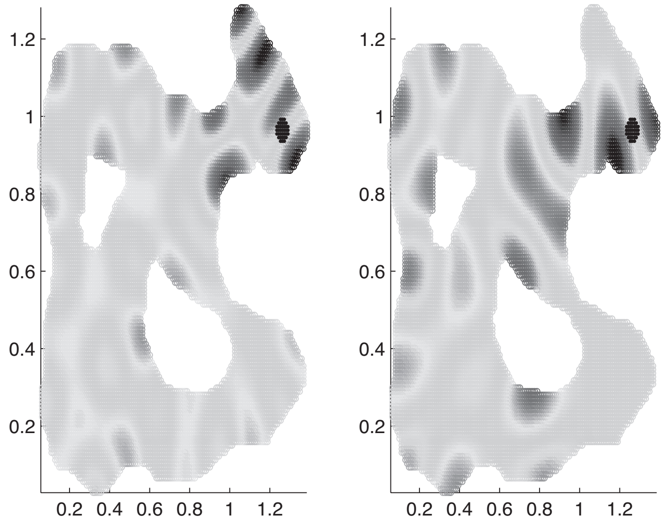

Fig. 1. A non-simply connected domain in $\mathbb{R}^2$. The dark circle is the neighborhood to be mapped, and grayscale intensity represents two (left and right) eigenfunctions for the embedding. Each of them has about half an oscillation in the neighborhood, and these two half-oscillations are in roughly orthogonal directions.

$$|\partial_p \varphi_j(z')| \sim^{C_8}_{C_7} R_z^{-1} \, \mathrm{Ave}^z_{\frac{1}{2}\delta_0 R_z} \, \varphi_j.$$

This lemma yields an eigenfunction that serves our purpose in a given direction. From this point onwards we let, with abuse of notation, $\varphi_j(z') = (\mathrm{Ave}^z_{\frac{1}{2}\delta_0 R_z} \varphi_j)^{-1} \cdot \varphi_j(z')$. To complete the proof of the theorems, we need to cover d linearly independent directions. Pick an arbitrary direction p_1. By *Lemma 3* we can find $j_1 \in \Lambda$, (in particular $j_1 \sim t^{-1}$) such that $|\partial_{p_1} \varphi_{j_1}(z)| \geq c_0 R_z^{-1}$. Let p_2 be a direction orthogonal to $\nabla \varphi_{j_1}(z)$. We apply again *Lemma 3,* and find $j_2 < At^{-1}$ so that $|\partial_{p_2}\varphi_{j_2}(z)| \geq c_0 R_z^{-1}$. Note that necessarily $j_2 \neq j_1$ and p_2 is linearly independent of p_1. In fact, by choice of p_2, $\partial_{p_2}\varphi_{j_1} = 0$. We proceed in this fashion. By induction, once we have chosen $j_1, \ldots, j_k$ $(k < d)$, and the corresponding $p_1, \ldots, p_k$, such that $|\partial_{p_l} \varphi_{j_l}(z)| \geq c_0 R_z^{-1}$, for $l = 1, \ldots, k$, we pick p_{k+1} orthogonal to $\langle \{\nabla \varphi_{j_n}\}_{n=1,\ldots,k} \rangle$ and apply *Lemma 3,* that yields j_{k+1} such that $|\partial_{p_{k+1}} \varphi_{j_{k+1}}(z)| \geq c_0 R_z^{-1}$.

We claim that the matrix $A_{k+1} := (\partial_{p_n} \varphi_{j_m})_{m,n=1,\ldots,k+1}$ is lower triangular and $\{p_1, \ldots, p_{k+1}\}$ is linearly independent. Lower-triangularity of the matrix follows by induction and the choice of p_{k+1}. Assume $a \in \mathbb{R}^{k+1}$ and $\sum_{n=1}^{k+1} a_n p_n = 0$, then $\langle \sum_{n=1}^{k+1} a_n p_n, \nabla \varphi_{j_l} \rangle = 0$ for all $l = 1, \ldots, k+1$, i.e., a solves the linear system $A_{k+1} a = 0$. But A_{k+1} is lower triangular with all diagonal entries non-zero, hence $a = 0$.

For $l \leq k$, we have $\langle \nabla \varphi_{j_l}, p_{k+1} \rangle = 0$ and, by *Lemma 3,* $|\langle \nabla \varphi_{j_l}, p_l \rangle| \gtrsim R_z^{-1}$. Now let $\Phi_k = (\varphi_{j_1}, \ldots, \varphi_{j_k})$ and $\Phi = \Phi_d$. We start by showing that $\|\nabla\Phi|_z (w - z)\| \gtrsim_d 1/R_z \|w - z\|$. Indeed, suppose that $\|\nabla\Phi_k|_z (w - z)\| \leq c/R_z \|w - z\|$, for all $k = 1, \ldots, d$. For c small enough, this will lead to a contradiction. Let $w - z = \sum_l a_l p_l$. We have (using say *Lemma 3*)

$$\|\nabla\Phi_k|_z(w-z)\| = \left\| \sum_{l\leq k} a_l \partial_{p_l} \Phi_k|_z \right\| \gtrsim \left(|a_k| - c \sum_{l<k} |a_l| \right) \frac{1}{R_z}.$$

By induction, $|a_k| \leq \sum_{l=1}^{k} c^l \|w - z\|$. For c small enough, $|a_i| \leq \|w - z\|/d$. This is a contradiction since $\|\sum_i a_i p_i\| = \|w - z\|$ and $\|p_i\| = 1$. We also have, by *Proposition 5,*

$$\|\nabla\Phi|_w - \nabla\Phi|_z\| \lesssim \left(\frac{\|z - w\|}{R_z} \right)^{\alpha \wedge 1} \frac{1}{R_z}. \qquad \textbf{[0.16]}$$

Finally, by ensuring $\|z - w_i\|/R_z$ is smaller than a universal constant for $i = 1, 2$, we get from Eq. **0.16**

$$\|\Phi(w_1) - \Phi(w_2)\| = \left| \int_0^1 \nabla\Phi|_{tw_1+(1-t)w_2}(w_1 - w_2)dt \right|$$

$$= \left| \int_0^1 (\nabla\Phi|_{w_1} + (\nabla\Phi|_{tw_1+(1-t)w_2} - \nabla\Phi|_{w_1})) \right.$$

$$\left. \cdot (w_1 - w_2)dt \right| \gtrsim \int_0^1 \frac{1}{R_z} \|w_1 - w_2\| dt$$

$$\gtrsim \frac{1}{R_z} c_0 \|w_1 - w_2\|,$$

which proves the lower bound **0.5.** To prove the upper bound of **0.5,** we observe that from *Proposition 5* we have the upper bound $|\partial_{p_l} \varphi_{i_l}(z)| \lesssim \mathrm{Ave}^z_{R_z}(\varphi_{i_l})$, and since φ_{i_l} is L^2-normalized, the right hand side can be as big as $\sim R_z^{-d/2}$. Therefore, we can choose γ_l as in the statement of the theorem so as to satisfy the upper bound **0.5.** This completes the proof for the Euclidean case.

In the manifold case, we consider first the case $g \in C^2$, and thus let $\alpha \wedge 1 = 1$ in what follows. Let R_z be as in the theorem. We take $c_1 \leq (1/2)\delta_0$ chosen so that

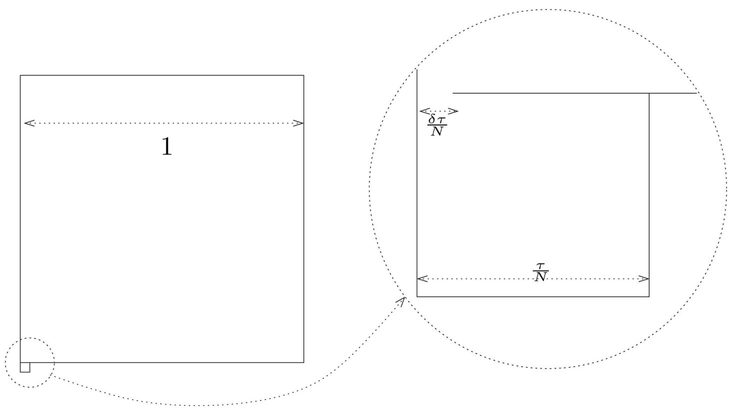

Fig. 2. Example of localization. When the entrance in the room is small enough, all of the eigenfunctions that *Theorem 1* selects for mapping a neighborhood of a point z in the small room need to be rescaled by a factor $R_z^{-d/2}$ to map a neighborhood of z to a ball of size ~ 1.

$$|g^{il}(x) - \delta^{il}| = |g^{il}(x) - g^{il}(z)| < \|g\|_{\alpha\wedge 1}\|x - z\|^{\alpha\wedge 1} < \epsilon_0$$

for all $x \in B_{2c_1R_z}(z)$. For this g, the above is carried on in local coordinates. It is then left to prove that the Euclidean distance in the range of the coordinate map is equivalent to the geodesic distance on the manifold. For all $x, y \in B_{c_1R_z}(z)$

$$d_{\mathcal{M}}(x, y) \leq \int_0^1 \left\|\frac{x - y}{\|x - y\|}\right\|_{\mathbb{R}^d} (1 + \|g\|_{\alpha\wedge 1} t^{\alpha\wedge 1})dt$$

$$\lesssim_{\alpha\wedge 1}(1 + \|g\|_{\alpha\wedge 1})\|x - y\|.$$

For the converse, let γ: $[0, 1] \to \mathcal{M}$ be the geodesic from x to y. γ is contained in $B_{2d_{\mathcal{M}}(x,y)}(x)$ on the manifold, whose image in the local chart is contained in $B_{2(1+\|g\|\alpha\wedge 1)d_{\mathcal{M}}(x,y)}(x)$. We have

$$d_{\mathcal{M}}(x, y) \gtrsim (1 - \|g\|_{\alpha\wedge 1})\int_\gamma \|\dot{\gamma}(t)\|_{\mathbb{R}^d} \gtrsim (1 - \|g\|_{\alpha\wedge 1})\|x - y\|.$$

Finally, when $g \in C^\alpha$ we will need:

Lemma 4. *Let $J > 0$ be given. If $\|\tilde{g}_n^{il} - g^{il}\|_{L^\infty(B_R(z))} \to_n 0$ with $\|\tilde{g}_n^{il}\|_{C^\alpha}$ uniformly bounded, then for $j <$ J*

$$\|\varphi_j - \tilde{\varphi}_{j,n}\|_{L^\infty(B_R(z))} \to_n 0,$$

$$\|\nabla(\varphi_j - \tilde{\varphi}_{j,n})\|_{L^\infty(B_R(z))} \to_n 0,$$

$$|\lambda_j - \tilde{\lambda}_{j,n}| \to_n 0.$$

To conclude the proof of the theorem, let $J = c_5R_z^{-2}$, depending on d, $(1/2)c_{\min}$, $2c_{\max}$, $\|g\|_{\alpha\wedge 1}$, $\alpha \wedge 1$. We may approximate g in C^α norm arbitrarily well by a C^2 $(\mathcal{M})$ metric. By the above lemma, and our main theorem for the case of C^2 metric, we obtain the theorem for the C^α case.

Examples

Example 2. Mapping with eigenfunctions, non simply-connected domain. We consider the planar, non-simply connected domain Ω in Fig. 1. We fix a point $z \in \Omega$, as in the figure, and display two eigenfunctions whose existence is implied by *Theorem 1.*

Example 3. Localized eigenfunctions. In this example, we show that the factors $\gamma_1, \ldots, \gamma_d$ in *Theorem 1* and *2* may in fact be required to be as small as $R_z^{d/2}$. We consider the "two-drums" domain in Fig. 2, consisting of a unit-size square drum, connected by a small aperture to a small square drum, with size τ/N, where τ is the golden ratio. The with of the connecting aperture is $\delta\tau/N$, for small δ. For this domain, for small enough δ, and for z in the smaller square, it can be shown that all possible eigenfunctions that may get chosen in the theorem are localized in the smaller square This is essentially a consequence of the fact that the proper frequencies of the two drums, for $\delta = 0$, are all irrational with respect to each other, and therefore eigenfunctions are perfectly localized on each drum. For δ small enough, a perturbation argument shows that the eigenfunctions will be essentially localized on each drum. But the eigenfunctions localized on the small drum, being normalized to L^2 norm, will have L^∞ norm as large as $R_z^{-d/2}$, and therefore the lower bound for the γ_i's is sharp.

ACKNOWLEDGMENTS. We thank K. Burdzy, R. R. Coifman, P. Gressman, H. Smith, and A. D. Szlam for useful discussions during the preparation of the manuscript, as well as IPAM for hosting these discussions and more. P.W.J., M.M., and R.S. are grateful for partial support from, respectively, NSF DMS 0501300, NSF DMS 0650413 and ONR N0001407-1-0625, NSF DMS 0502747.

1. Coifman RR, Lafon S, Lee AB, Maggioni M, Nadler B, Warner F, Zucker SW (2005) Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proc Natl Acad Sci USA* 102:7426–7431.
2. Jones PW, Maggioni M, Schul R (2007) Universal local manifold parametrizations via heat kernels and eigenfunctions of the Laplacian, http://arxiv.org/abs/0709.1975.
3. Roweis ST, Saul LK (2000) Nonlinear dimensionality reduction by locally linear embedding. *Science* 290:2323–2326.
4. Belkin M, Niyogi P (2003) Using manifold structure for partially labelled classification. *Adv NIPS* 15.
5. Coifman RR, Lafon S (2006) Diffusion maps. *Appl Comput Harm Anal* 21:5–30.
6. Tenenbaum JB, de Silva V, Langford JC (2000) A global geometric framework for nonlinear dimensionality reduction. *Science* 290:2319–2323.
7. Zhang Z, Zha H (2002) *Principal Manifolds and Nonlinear Dimension Reduction via Local Tangent Space Alignment* (Dept of Computer Science, Pennsylvania State Univ, University Park, PA), Tech Rep CSE-02-019.
8. Donoho DL, Grimes C (2003) Hessian eigenmaps: New locally linear embedding techniques for high-dimensional data. *Proc Natl Acad Sci USA* 100:5591– 5596.
9. Donoho DL, Grimes C (2002) *When Does Isomap Recover Natural Parameterization of Families of Articulated Images?* (Dept of Statistics, Stanford Univ, Stanford, CA), Tech Rep 2002-27.
10. Niyogi P, Matveeva I, Belkin M (2003) *Regression and Regularization on Large Graphs* (Univ of Chicago, Chicago), Tech Rep.
11. Szlam AD, Maggioni M, Coifman RR (2006) *A General Framework for Adaptive Regularization Based on Diffusion Processes on Graphs* (Dept of Computer Science, Yale Univ, New Haven, CT), Tech Rep 1365.
12. Mahadevan S, Maggioni M (2007) Proto-value functions: A spectral framework for solving Markov decision processes. *J Mach Learn Res* 8:2169–2231.
13. Szlam AD, Maggioni M, Coifman RR, Bremer JC, Jr (2005) Diffusion-driven multiscale analysis on manifolds and graphs: Top-down and bottom-up constructions. *Proc SPIE* 5914:59141D.
14. Das P, Moll M, Stamati H, Kavraki LE, Clementi C (2006) Low-dimensional, free-energy landscapes of protein-folding reactions by nonlinear dimensionality reduction. *Proc Natl Acad Sci USA* 103:9885–9890.
15. Bérard P, Besson G, Gallot S (1994) Embedding Riemannian manifolds by their heat kernel. *Geom Funct Anal* 4:374–398.
16. Netrusov Y, Safarov Y (2005) Weyl asymptotic formula for the Laplacian on domains with rough boundaries. *Commun Math Phys* 253:481–509.
17. Hempel R, Seco L, Simon B (1991) The essential spectrum of Neumann laplacians on some bounded singular domains. *J Funct Anal* 102:448–483.
18. Grüter M, Widman K-O (1982) The Green function for uniformly elliptic equations. *Man Math* 37:303–342.