DOI: 10.1111/j.1467-8659.2011.02026.x

Eurographics Symposium on Geometry Processing 2011
Mario Botsch and Scott Schaefer
(Guest Editors)

*Volume 30 (2011), Number 5*

# Multiscale Biharmonic Kernels

Raif M. Rustamov

Drew University, Madison NJ, USA

**Abstract**

*This paper introduces a general principle for constructing multiscale kernels on surface meshes, and presents a construction of the multiscale pre-biharmonic and multiscale biharmonic kernels. Our construction is based on an optimization problem that seeks to minimize a smoothness criterion, the Laplacian energy, subject to a sparsity inducing constraint. Namely, we use the lasso constraint, which sets an upper bound on the $l_1$-norm of the solution, to obtain a family of solutions parametrized by this upper-bound parameter. The interplay between sparsity and smoothness results in smooth kernels that vanish away from the diagonal. We prove that the resulting kernels have gradually changing supports, consistent behavior over partial and complete meshes, and interesting limiting behaviors (e.g. in the limit of large scales, the multiscale biharmonic kernel converges to the Green's function of the biharmonic equation); in addition, these kernels are based on intrinsic quantities and so are insensitive to isometric deformations. We show empirically that our kernels are shape-aware, are robust to noise, tessellation, and partial object, and are fast to compute. Finally, we demonstrate that the new kernels can be useful for function interpolation and shape correspondence.*

## 1. Introduction

Recently, there has been a renewed interest in multiscale methods as it was realized that a multiscale kernel, namely the heat kernel, allows extracting information about shapes at multiple levels. This realization was immediately found useful in applications such as distance measurement, segmentation, shape matching and retrieval. To increase the scope of applications and to provide a greater assortment of choices, there is a need for new multiscale kernels. Desirable properties of such kernels include gradually changing local support, consistent behavior over partial and complete meshes, non-trivial limiting behavior, being intrinsic, and being fast to compute.

Devising multiscale constructions on non-trivial geometries is challenging. Wavelet theory – the natural context for such constructions – provides only a few approaches that can be adapted to the surface setting: 1) based on the Laplace-Beltrami eigenfunctions (also used in the heat kernel), but eigenfunctions are global so the resulting functions have full support; 2) lifting scheme, which is hard to make intrinsic in general due to the non-trivial use of the underlying mesh; 3) diffusion wavelets, which are critically sampled so one cannot obtain kernels centered at every point; 4) projection methods, which in general do not provide constructions on the entire surface.

In this work, we introduce a general principle for constructing multiscale kernels together with the construction of two such kernels. We design the multiscale pre-biharmonic kernel $P_t(x,y)$ and the multiscale biharmonic kernel $B_t(x,y)$ that have supports varying from local (for small values of the scale parameter $t$) to global (for large $t$), see Figure 1. Our construction is based on a convex optimization problem seeking to minimize a smoothness criterion, the Laplacian energy, subject to a sparsity inducing constraint. Namely, we use the lasso constraint, which sets an upper bound on the $l_1$-norm of the solution, to obtain a family of solutions parametrized by this upper-bound parameter $t$. The interplay between the smoothness objective and the sparsity constraint results in smooth solutions with local supports that gradually increase with the parameter $t$.

This approach was chosen for several reasons. First, since the lasso constraint induces true sparsity, our kernels exactly vanish away from the diagonal; this is in contrast to the heat kernel which only decays away from the diagonal but has full support on the surface. Second, the underlying optimization problem makes it possible to argue that our kernels computed on the partial object are exactly equal to the kernels computed on the complete object (assuming the kernel is fully supported within the partial object). Third, within our construction we are able to prove that the multiscale bi-

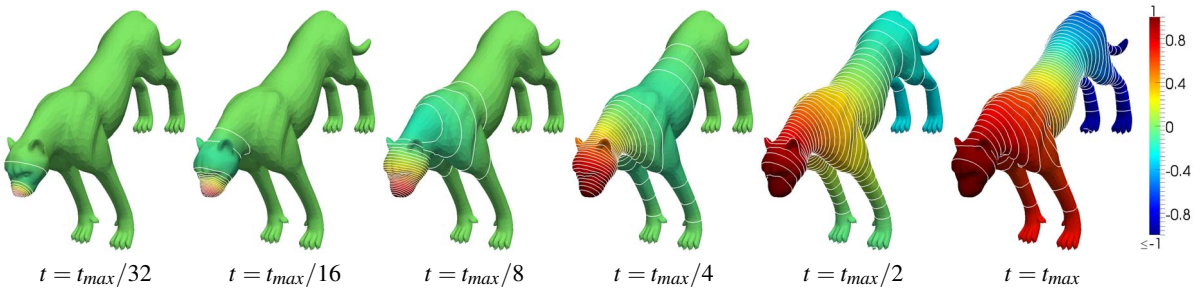



**Figure 1:** *The multiscale biharmonic kernel $B_t(x,y)$ is shown for different values of the scale parameter $t$. Point $y$ is located on the tip of cheetah's nose. When $t = t_{max}$, our kernel coincides with the (scaled) global biharmonic kernel – Green's function of the biharmonic equation.*

harmonic kernel $B_t(x,y)$ converges to the global biharmonic kernel (Green's function of the biharmonic equation) which has recently been used to construct a global distance measure on meshes [LRF10] and seems to strike a balance in the normalization of eigenfunctions. Fourth, our construction uses only intrinsic quantities of the surface including the Laplace-Beltrami operator, which renders the kernels shape-aware and insensitive to near isometric deformations. Finally, our approach leads to a convex quadratic optimization problem with linear constraints, which can be solved using off-the-shelf optimization software very efficiently.

The main contribution of this paper is the proposed principle for constructing multiscale kernels together with the construction of two such kernels (Section 3). In addition, we investigate the theoretical properties of these kernels (Section 4). After providing more detail about their computation (Section 5), we present their empirical evaluation (Section 6) and applications (Section 7).

## 2. Previous work

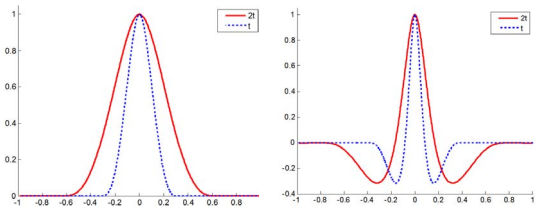
Multiscale constructions, usually discussed in the context of wavelets, are of great importance in many applied areas and in theoretical mathematics. Among such constructions we will analyze ones that can be adapted to the setting of surfaces or surface meshes; for a recent review with an emphasis on projection methods see [ARV10].

The first class of multiscale constructions is based on the eigenvalues and eigenfunctions of the Laplace-Beltrami operator. Let $\{\lambda_i, \phi_i\}_{i=0}^{\infty}$ be the Laplace-Beltrami eigenvalues and orthonormal eigenfunctions; note that $\lambda_0 = 0$ with the corresponding eigenfunction $\phi_0 = const$. Consider the general expression $K_t(x,y) = \sum_i g(t\lambda_i)\phi_i(x)\phi_i(y)$, where $g$ is some real function. To be multiscale, $K_t(x,y)$ should localize around the diagonal and the degree of localization should depend on the scale parameter $t$.

Choosing $g(s) = e^{-s}$ results in the heat kernel $H_t(x,y)$, which is well-known to be multiscale and has already found important applications in geometry processing, see for example De Goes et al. [dGGV08], Sun et al. [SOG09], and Ovsjanikov et al. [OMMG10]. Geller and Mayeli [GM09] make the choice of $g(s) = se^{-s}$ to obtain the counterpart of the Mexican hat wavelet on manifolds. By assuming that $g(s)$ is a band-pass filter with special behavior around $s = 0$, Hammond et al. [HVG11] prove that the obtained kernels (wavelets) $K_t(x,y)$ are multiscale. While having provable decay properties away from the diagonal, note that all of these kernels/wavelets have full supports and depend on eigenfunctions which are global by their very nature. Thus, at every scale these kernels are aware of the entire manifold, and, for example, one cannot expect that a kernel computed on a partial manifold (sub-manifold) will be exactly equal to the kernel computed on the complete manifold.

The lifting scheme [SS95, Swe98] and diffusion wavelet construction [CM06] revolutionized wavelet theory. Both of these approaches are universal and can be applied to construct wavelets on complicated geometries. However, the lifting scheme makes a non-trivial use of the underlying mesh and its multiresolution decomposition, which makes it hard in general to obtain wavelets/kernels that are intrinsic. Diffusion wavelet construction, on the other hand, results in critically sampled wavelets; while this is a very desirable property in wavelet theory, for our purposes it means that we cannot obtain kernels centered at every point on the mesh. Finally, another approach is the projection method [ARV10] which transfers a wavelet construction from the plane to the surface by using an appropriate projection map. In general, this results in a wavelet construction within a single manifold chart; gluing these charts together and obtaining an intrinsic construction is difficult.

Our construction is based on the well established principle in compressed sensing and statistics that the $l_1$-norm regularizers/constraints induce sparsity. Specifically we use the Least Absolute Shrinkage and Selection Operator, lasso, constraint introduced by Tibshirani [Tib96] for statistical variable selection; it also has close relationship to the basis pursuit method [CDS98]. In computer graphics sparse




a) Pre-biharmonic b) Biharmonic

**Figure 2:** *Multiscale pre-biharmonic and multiscale biharmonic kernels for two different values of t in 1D case.*

methods have been used in the construction of $L_1$ splines [Lav02] and for surface reconstruction [ASGCO10]. Note that in both of these applications the $l_1$-norm is applied to the error term, which forces the error to concentrate at sparse locations; we, on the other hand, apply the $l_1$-norm to the solution vector in order to make the solution locally supported.

## 3. Formulation

In this section we construct the multiscale pre-biharmonic kernel $P_t(x,y)$ and the multiscale biharmonic kernel $B_t(x,y)$, both of which depend on the scale parameter $t$. Our general construction principle is to build these kernels as solutions to optimization problems which require the kernels to be smooth and have local supports. We first discuss the continuous version of our constructions and then discretize.

**Continuous setting:** Given a compact surface $S$, let $\Delta$ be its Laplace-Beltrami operator. Consider the following constrained (variational) least squares problem

$$\min_f \int_S (\Delta f)^2 \quad \text{subject to} \int_S |f| \le t \quad \text{and} \quad f(y) = 1,$$

where $f : S \to R$ belongs to the set of real-valued functions on the surface, and $y$ is a fixed point on the surface. This optimization problem has a unique solution (see Section 4) which is a function $g_{t,y} : S \to R$ that depends on the point $y$ and parameter $t$. Thus, we can define the function $P_t(x,y) = g_{t,y}(x)$. Since our optimization problem is a pre-cursor to the multiscale biharmonic kernel optimization problem, we will refer to $P_t(x,y)$ as the multiscale pre-biharmonic kernel.

Now we modify the optimization problem by adding another constraint: we require that the integral of the solution over the surface must vanish, $\int_S f = 0$. This requirement mimics the property of global poly-harmonic kernels which do not contain the constant eigenfunction term and, therefore, integrate to zero. Similarly to above, we obtain another function $B_t(x,y)$ which we will call the multiscale biharmonic kernel. The name is justified by the fact that for values of $t$ large enough, $B_t$ is the scaled Green's function of the biharmonic equation (see Section 4).

Figure 2 shows the one-dimensional versions of the multiscale pre-biharmonic and biharmonic kernels centered at $y = 0$ for two different values of $t$. The kernels can be seen to exhibit the following desirable properties. First, both of the kernels are smooth because the objective functional, the Laplacian energy, is well known to provide a measure of smoothness for functions on manifolds. Second, both of the kernels have local supports around the center; as we will see, this is due to the constraint on the $L^1$ norm of the solution which forces the solution to be sparse. Finally, both kernels are multiscale – their supports gradually grow with $t$. In this one-dimensional case this growth is particularly simple: the kernel for $2t$ is the twice horizontally scaled version of the kernel for $t$, and so it has a twice larger support. We will fully discuss these and further properties of our kernels in the next section.

**Discrete setting:** Consider a discretization of a surface $S$ as a triangle mesh with vertices $v_i$, $i = 1,\ldots,n$. Let $W$ be the $n \times n$ diagonal matrix with entry $W_{ii}$ equal to the point/Voronoi area at vertex $v_i$. We denote the conformal (with no area weights) cotangent Laplacian matrix [PP93] by $L$. The Laplace-Beltrami operator $\Delta$ can be discretized [MDSB03] as $W^{-1}L$. We discretize the surface function $f : S \to R$ as the column-vector $\vec{f} \in R^n$ with $i$-th entry $f_i = f(v_i)$. Surface integrals are discretized as follows $\int_S f \approx \sum_i W_{ii} f_i$.

Now we will discretize all the terms appearing in our optimization problems. Assuming that $y = v_k$, after some algebraic manipulation the pre-biharmonic kernel optimization problem becomes

$$\min_{\vec{f}} \ \|W^{-1/2} L \vec{f}\|_2^2 \quad \text{subject to} \ \|W\vec{f}\|_1 \le t \quad \text{and} \quad f_k = 1,$$

where the one-half power arises as the integration weights are pushed inside the squared 2-norm. The multiscale biharmonic kernel optimization problem is discretized similarly and only has the additional constraint $\sum_i W_{ii} f_i = 0$ which captures vanishing of the integral $\int_S f = 0$.

## 4. Properties

**Smoothness:** The Laplacian energy $\int_S (\Delta f)^2$ is well-known to provide a measure of smoothness for functions on manifolds [BNS06, BS08]. As a result, our optimization problem is finding the smoothest function subject to given constraints. This results in smooth kernels whose support have smooth boundaries and include the center point. In addition, we have empirically observed that for both of the kernels the (closure of) support regions are connected. For the pre-biharmonic kernel this property is easy to prove: any island not containing the center point can be eliminated by setting the function equal to zero on that island, this will result in a smaller objective functional value without violating the constraints.

Note that a variety of choices for the smoothness functional exist, such as the Dirichlet energy or the energies based on iterated Laplacians. Our specific choice is made in order to force $B_t(x,y)$ to converge to Green's function of the biharmonic equation, which has recently been used

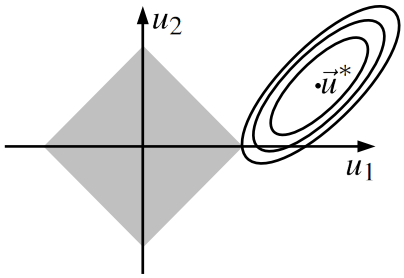



**Figure 3:** *The sparsifying effect of the lasso constraint. Ellipses depict the isocontours of $\|A\vec{u}-\vec{b}\|_2^2$ and shaded in gray is the feasible region for lasso constraint $\|\vec{u}\|_1 \leq t$. Solutions at corners have a vanishing coordinate.*

to construct a global distance measure on meshes [LRF10] and seems to strike a balance in the normalization of eigenfunctions. We avoid using the Dirichlet energy because this would have resulted in convergence to Green's function of the harmonic equation, which has a logarithmic singularity along the diagonal (cf. [LRF10]).

**Local support:** Both of the kernels have local supports around the center point $y$. This is a consequence of the interplay between the smoothness inducing property of the objective functional and the sparsifying effect of the constraint on the $l_1$ norm of the solution. We quickly review how this constraint forces the solution to have vanishing entries.

As shown in Appendix A, the discrete pre-biharmonic optimization problem can be written as a standard least squares problem with *the lasso constraint* [Tib96]:

$$\min_{\vec{u}} \|A\vec{u}-\vec{b}\|_2^2 \qquad \text{subject to} \quad \|\vec{u}\|_1 \leq t. \tag{1}$$

The multiscale biharmonic kernel problem can be written similarly with an additional equality constraint corresponding to the vanishing integral requirement.

The sparsifying effect of the lasso constraint $\|\vec{u}\|_1 \leq t$ when $\vec{u}$ is two-dimensional is demonstrated in Figure 3 (redrawn from [Tib96]). The diamond shaded in gray shows the feasible region for the lasso constraint; the ellipses depict the isocontours of the least squares error function being minimized. When the solution is at a corner of the feasible region, one of its coordinates will be zero. In high dimensions, sparse solutions become much more likely not only because the number of feasible region corners grows exponentially, but also because additional sharp structures which force vanishing entries, such as edges and higher dimensional cells, are present.

Note that the additional equality constraint in the multiscale biharmonic kernel will constrain the solution to lay on a hyperplane, so the feasible region (the intersection of this hyperplane with the lasso region) will retain sharp structures enforcing vanishing entries – the sparsity inducing effect of the lasso constraint will survive.

**Gradual support:** The number of vanishing entries for a lasso constrained problem can be controlled by varying the parameter $t$. For our kernels this results in supports that gradually grow as the scale parameter $t$ is increased. The case of the pre-biharmonic kernel lends itself to a more detailed analysis. In the discrete case, we claim that with increasing $t$ the support of the pre-biharmonic kernel gradually expands by passing through the two-ring neighbors. To be more precise, assuming non-degeneracy, for all $t$ and small enough $\varepsilon$, we have $supp(P_{t+\varepsilon}(\cdot,y)) \subseteq N_2(supp(P_t(\cdot,y)))$; here for any subset of mesh vertices $V$, we denote by $N_2(V)$ its two-ring neighborhood ($V \subseteq N_2(V)$).

To obtain this result, we need the following fact: the solutions to the standard least squares problem with the lasso constraint, Eq. (1), can be obtained via a procedure called LARS. Since we do not use LARS in our implementation to compute the kernel, we will only quickly sketch the procedure; for full details see [EHJT04]. Start with $\vec{u}=0$ and let the active set be empty. Find the column of $A$, say $i$-th column $A_i$, most correlated with the initial residual $\vec{r}=\vec{b}-A\vec{u}=\vec{b}$. The corresponding entry $u_i$ of the solution vector enters the active set and we start changing $u_i$ so as to decrease the correlation of $A_i$ with the current residual. This continues until another column of $A$, say $A_j$, has as much correlation with the current residual as $A_i$. Now the corresponding solution entry $u_j$ enters the active set, and both $u_i$ and $u_j$ are allowed to change. The process is repeated until the active set includes all the entries of the solution vector and eventually yields the ordinary least squares solution. With an appropriate provision for dropping variables from the active set, one can prove that during this process $\vec{u}$ traverses all the lasso constrained solutions for increasing values of $t$.

The pre-biharmonic optimization problem can be written in the form of Eq. (1), and so this kernel can be computed using the sketched LARS procedure. The LARS active set corresponds to the support of our kernel, and it can only grow by adding vertices that have non-zero correlation with the current residual. One can see that the residual $\vec{r}=\vec{b}-A\vec{u}$ is supported within the two-ring neighborhood of the active set. Indeed, in our case, matrix $A$ has the same sparsity pattern as the Laplacian, and vector $\vec{b}$ is supported within the one ring neighborhood of the vertex $y=v_k$, see Appendix A. This proves our assertion about the gradual growth of the pre-biharmonic kernel support. While due to the vanishing integral condition, the LARS procedure cannot be applied to the multiscale biharmonic kernel problem directly, we believe that qualitatively a similar behavior can be expected.

**Large $t$ behavior:** For large enough values of the parameter $t$, the lasso constraint becomes inactive and the kernels become independent of $t$. Indeed, let $f^*$ be the solution of either of our optimization problems with the lasso constraint removed. When the scale parameter $t=\|f^*\|_1$, this solution $f^*$ is within the lasso feasible region, and so it will be the solution of the lasso constrained problem as well. Clearly, increasing $t$ beyond this value will not change the solution

– the lasso constraint will remain inactive — so one can define $t_{max} = \|f^*\|_1$. Therefore, to investigate the behavior of our kernels for large values of $t$ and to find $t_{max}$ we will drop the lasso constraint and find the solution of the remaining problem.

For the pre-biharmonic kernel, it is easy to see that the choice $f^* \equiv 1$ satisfies the constraint $f(y) = 1$ and makes the objective functional equal to zero, which is its smallest possible value. We conclude that $t_{max} = \|f^*\|_1 = \int_S |f^*| = Area(S)$, and that the pre-biharmonic kernel $P_t(x,y) = 1$ for all $t \geq Area(S)$ and all $x, y \in S$.

Dropping the lasso constraint from the multiscale biharmonic kernel problem leads to a more difficult problem. As shown in Appendix B, the solution to this problem is given by $f^*(x) = G(x,y)/G(y,y)$; here $G(x,y)$ is the global biharmonic kernel. As a result, we find that $t_{max}$ depends on the point $y$, and is given by $t_{max} = \int_S |f^*| = \int_S |G(x,y)/G(y,y)|$. For $t \geq t_{max}$,we have $B_t(x,y) = G(x,y)/G(y,y)$ for all $x \in S$.

The relevance of the biharmonic kernel in this context can be better seen after rewriting our smoothness objective using integration by parts: $\int_S (\Delta f)^2 = \int_S f \Delta^2 f$, assuming the appropriate boundary conditions. It is true that if we wanted to construct a kernel that converges to Green's function of $\Delta^p$, we could have chosen the smoothness objective as $\int_S f \Delta^p f$. The discussion of these more general kernels is beyond the scope of this paper.

**Partial model behavior:** A kernel computed on a partial model coincides with the kernel on the complete model, as long as the partial model is big enough to include the (complete model) kernel's support. The precise statement, say in the case of the multiscale biharmonic kernel, is as follows. Consider the kernel $B_t(\cdot, y)$ for some value of $t$ and a point $y \in S$. Let $R$ be a subset (sub-manifold with boundary) of surface $S$ such that the support of $B_t(\cdot, y)$ is contained within $R$, and let $B_t^R(\cdot, y)$ be the multiscale biharmonic kernel computed on $R$. Now we can easily see that $B_t(\cdot, y)$ and $B_t^R(\cdot, y)$ coincide. Indeed, the solution of the multiscale biharmonic problem on $S$ is supported within $R$ and is feasible for the optimization problem on $R$, and vice-versa, the solution of problem for $R$ is feasible for the problem for $S$. Noting that the objective functional does not change when switching between $R$ and $S$, we see that the solutions of these optimization problems must coincide.

**Isometry invariance:** If the surface is isometrically deformed, then the kernels will stay the same. Indeed, all of the ingredients used in our construction involve only intrinsic quantities on the surface, namely the Laplace-Beltrami operator and surface integrals. As a result, the optimal function will stay the same, and the computed kernels will be the same as before the deformation.

**Unique solution:** The optimization problems for both of the kernels have unique solutions. This can be most clearly seen in the discrete case. Indeed, both of the problems reduce to convex problems in the standard form (see Section 5), and the objective function $\|W^{-1/2} L \vec{f}\|_2^2$ is strictly convex over the feasible set, which implies the uniqueness of minimizer (see [BV04] pp.136-137). Strict convexity of the objective can be seen as follows. The Laplacian matrix $L$ is positive definite except on the subspace spanned by the constant eigenvector; shifts by constant vectors leave the objective the same. However, since the constraints enforce a specific value at one of the mesh vertices ($f_k = 1$), such shifts are not possible within the feasible region.

## 5. Computation

Due to widespread applications in compressed sensing and statistics, there are numerous algorithms for solving the least squares problems with $l_1$ norm constraints/regularizers. Following practical needs, most of such approaches assume that the involved matrices are dense and so they trade accuracy for efficiency. In our case the Laplacian matrix is sparse and this allows obtaining solutions efficiently without employing any approximations. Namely, we rewrite our optimization problems as convex quadratic problems with linear constraints (see e.g. [KKL*07]), and solve them using the interior point method [BV04].

To reformulate the multiscale biharmonic optimization problem as a convex quadratic problem we introduce two auxiliary vector variables $\vec{h}$ and $\vec{r}$, and we let $\vec{w} = diag(W)$ be the column vector containing the diagonal entries of the area weight matrix $W$. The equivalent quadratic problem is as follows:

$$
\begin{array}{lllll}
\min\limits_{\vec{f},\vec{h},\vec{r}} & \vec{r}^T W^{-1} \vec{r} & & & \\
\text{subject to} & \vec{w}^T \vec{f} = 0 & (a) & \vec{f} + \vec{h} \geq 0 & (e) \\
& \vec{w}^T \vec{h} \leq t & (b) & -\vec{f} + \vec{h} \geq 0 & (f) \\
& f_k = 1 & (c) & \vec{h} \geq 0 & (g) \\
& L\vec{f} - \vec{r} = \vec{0} & (d) & &
\end{array}
$$

Here the inequalities involving vectors apply element-wise, and the $T$ superscript refers to the vector transpose. First, from the constraints $(e, f, g)$ notice that the auxiliary variable $\vec{h}$ captures the absolute values of entries of $\vec{f}$ needed in the lasso constraint; more precisely these constraints enforce the inequalities $|f_i| \leq h_i$. Second, constraints $(a)$, $(b)$, and $(c)$ capture the integral equals to zero, the lasso, and $f(y) = 1$ constraints respectively. Finally, constraint $(d)$ allows rewriting the objective function in terms of the residual; this presumably leads to a sparser Hessian [MOS11].

To obtain the corresponding formulation for the pre-biharmonic kernel, one only needs to drop the constraint $(a)$ from the problem. After experimenting extensively with this formulation, we found that the pre-biharmonic kernel never attains noticeable negative values. As a result, for all practical purposes the following simple formulation can be used

to compute the pre-biharmonic kernel:

$$\min_{\vec{f},\vec{r}} \quad \vec{r}^T W^{-1} \vec{r}$$
$$\text{subject to} \quad \vec{w}^T \vec{f} \leq t \quad f_k = 1 \quad L\vec{f} - \vec{r} = \vec{0} \quad \vec{f} \geq 0$$

This is the formulation we used for the pre-biharmonic kernel examples presented in this paper.

Both of the above formulations are convex quadratic problems with linear constraints; we solve them using the log-barrier interior point method. Our implementation uses the MOSEK [MOS11] interior point solver which takes advantage of sparsity in our problems, can be conveniently called from MATLAB, and is freely available for academic purposes. Without any parameter tweaking, this approach produces high quality solutions efficiently as will be shown in Section 6.

## 6. Results

In order to investigate the empirical properties of the multiscale kernels introduced in this paper, we ran a set of experiments using a variety of 3D meshes. In these experiments, our goals are to understand how the kernels behave on different surfaces, to test their robustness to noise and tessellation, to confirm their theoretical properties empirically, and to make qualitative comparisons to the heat kernel.

For the experiments, we compute the multiscale kernels $P_t(\cdot, y)$ and $B_t(\cdot, y)$ for a surface point $y$ at different time scales $t$. The time scales are shown in terms of $t_{max}$ which is defined for each of the kernels in Section 4; in case of the multiscale biharmonic kernel we set $t_{max}$ somewhat larger than the theoretical value so as to tolerate numerical errors. For ease of visualization, unless specified otherwise, all kernels are scaled to have the maximum value of 1 and the following color coding is used: dark red represents large positive values, neutral green corresponds to zero, and dark blue represents large negative values, see the color legend in Figure 1; this same color legend applies to all of the figures. To enhance the visual information content, we also include the isocontours, shown as white curves, of kernel function equally spaced within the function's range; note that the number of isocontours may vary from figure to figure in order to avoid the aliasing when the support of the kernel is small.

Figure 1 depicts the multiscale biharmonic kernel on the cheetah model for six values of the parameter $t$. The kernel displays multiscale behavior: its support gradually increases from very local to global. In fact, at $t \geq t_{max}$ the kernel converges to the scaled global biharmonic kernel. For this model, the maximum absolute difference between our kernel at $t = t_{max}$ and the scaled global biharmonic kernel [LRF10] is $4 \times 10^{-7}$.

Figure 4 shows both of multiscale pre-biharmonic and biharmonic kernels on different models. Note that the kernels vary smoothly over the surface, are isotropic around the point $y$ for small $t$; they become increasingly "shape-aware" for larger values of $t$, and they follow the natural cross-sections of the shape away from $y$. We also point out Figure 5 g-h which shows that the kernels are well-defined on surfaces with boundaries.

Figure 5 demonstrates the various insensitivity properties of the multiscale biharmonic kernel; due to limited space similar figures for the pre-biharmonic kernel are not included. The top two images (Figure 5 a-b) demonstrate the insensitivity to noise. Here the addition of Gaussian noise to the mesh ($\sigma = 400\%$ of the average edge length) leads to little change in the kernel as evidenced by the coloring and the isolines. The second row of two images (Figure 5 c-d) shows insensitivity to tessellation – as the mesh is simplified from 50K vertices to 5K, the kernel remains almost unchanged.

The third row (Figure 5 e-f) confirms the theoretical isometry invariance property of the kernel – the isolines and coloring are in correspondence between the David model and its nearly isometric deformation. The last row (Figure 5 g-h) demonstrates the behavior of the kernel on partial models. Here, function values along the isolines match; we can see that the kernel computed on the hand alone matches the kernel computed on the human model for the same point $y$ and value of $t$; similarly for the Cat model. Here the partial models are obtained by cutting the complete model by a plane and then remeshing.

Figure 6 provides a comparison between our kernels and the heat kernel. Here we scale the model to have unit area and then compute the heat kernel at values of $t = 1/256,\ 1/64,\ 1/16$. Next, in each case we found multiscale pre-biharmonic and biharmonic kernels that have supports similar to the heat kernel's support. Upon visual inspection we find that heat kernel and pre-biharmonic kernel behave similarly. Interestingly, they also have the same limiting behavior: both of them start with support at a single vertex, and as $t \to \infty$ (in fact, $t \geq t_{max}$ for pre-biharmonic kernel) both of them become global and converge to a constant function. One consequence of this limiting behavior is that even for relatively small $t$ the kernels have a smeared look near the center-point $y$ and "forget" its location – the isocontours do not make circles around $y$. The multiscale biharmonic kernel, on the other hand, has circular isocontours around the point $y$ even when the kernel's support is large; such preservation of center point location information can be useful in shape matching.

**Timing:** Table 1 reports time in seconds to compute the multiscale kernels, including the heat kernel, at different values of $t$ for a variety of surface mesh sizes on a 3.06Ghz Intel T9900 processor with 4G RAM, single threaded. The entire implementation is done in MATLAB, except for optimization we call MOSEK through its convenient MATLAB interface. For heat kernel we scale the models to have unit area

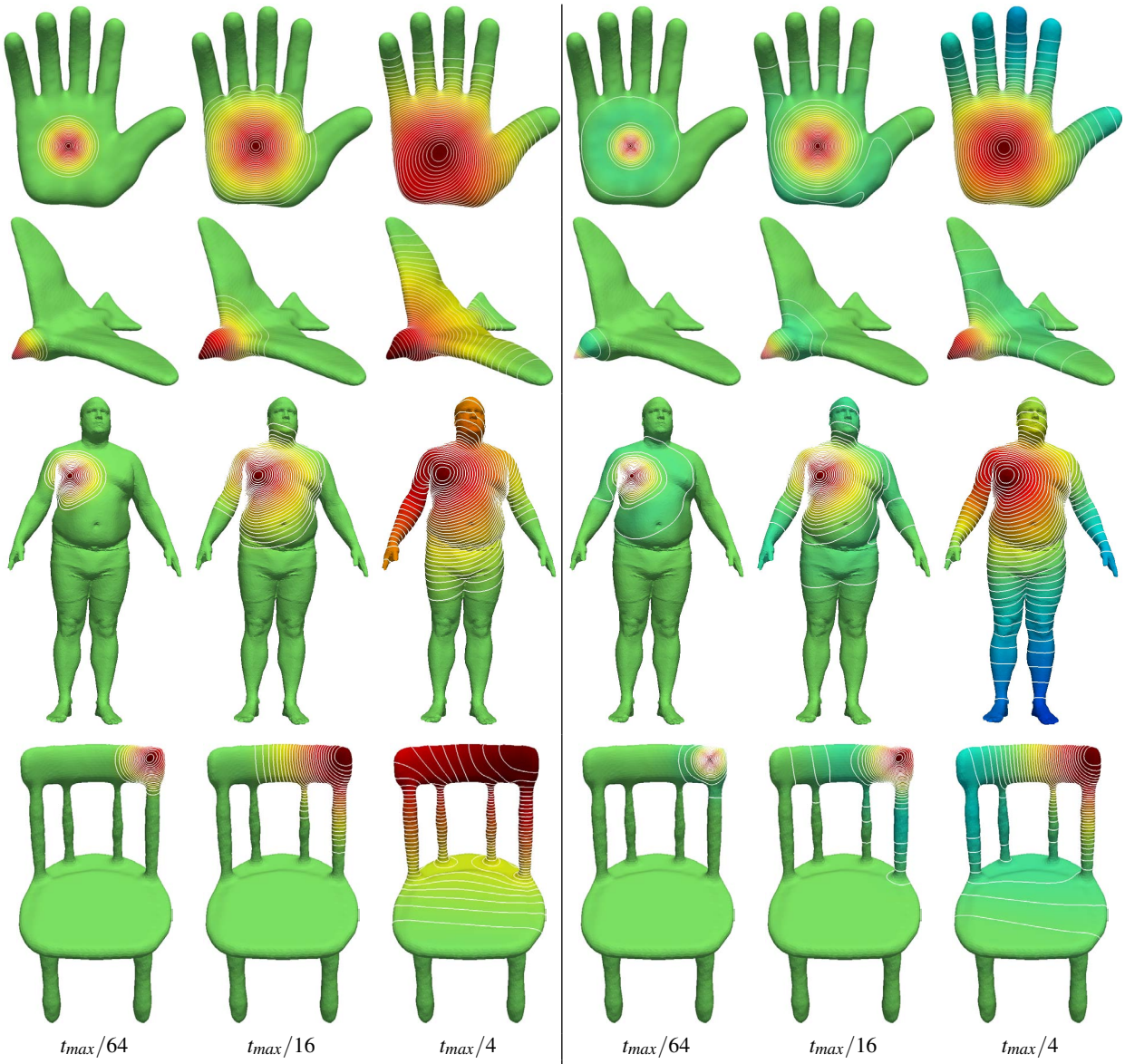



**Figure 4:** *Examples of multiscale pre-biharmonic and multiscale biharmonic kernels at different scales on a variety of models.*

| Size | Multiscale biharmonic | | | | Multiscale pre-biharmonic | | | | Heat – Matrix exponential | | | | Heat – Spectral | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\|V\|$ | $t = 1/4$ | $1/16$ | $1/64$ | $1/256$ | $1/4$ | $1/16$ | $1/64$ | $1/256$ | $1/4$ | $1/16$ | $1/64$ | $1/256$ | $1/4$ | $1/16$ | $1/64$ | $1/256$ |
| 2K | 0.47 | 0.42 | 0.41 | 0.36 | 0.31 | 0.31 | 0.31 | 0.27 | 3.01 | 0.91 | 0.33 | 0.12 | 0.07 | 0.11 | 0.40 | 9.29 |
| 4.3K | 0.94 | 0.97 | 1.00 | 0.92 | 0.56 | 0.61 | 0.59 | 0.51 | 29.94 | 7.56 | 2.52 | 0.83 | 0.16 | 0.18 | 0.98 | 17.25 |
| 8K | 2.08 | 2.01 | 1.65 | 1.92 | 1.09 | 1.33 | 1.23 | 1.06 | 151.81 | 49.42 | 15.66 | 4.83 | 0.33 | 0.41 | 3.32 | 38.67 |
| 16K | 5.16 | 5.66 | 7.35 | 6.30 | 3.20 | 3.60 | 4.21 | 3.28 | 206.63 | 65.08 | 21.34 | 6.40 | 0.70 | 0.71 | 5.11 | 46.93 |
| 27.9K | 8.74 | 14.52 | 9.75 | 8.86 | 5.12 | 5.72 | 6.33 | 6.45 | >600 | >600 | >600 | 599.08 | 1.60 | 1.60 | 11.30 | 107.03 |
| 52.6K | 28.91 | 26.41 | 25.10 | 22.78 | 12.84 | 14.10 | 17.58 | 15.83 | >1200 | >1200 | >1200 | >1200 | 0.17 | 0.17 | 0.74 | 14.78 |

**Table 1:** *Time (in seconds) to compute multiscale kernels for different values of the scale parameter $t$.*

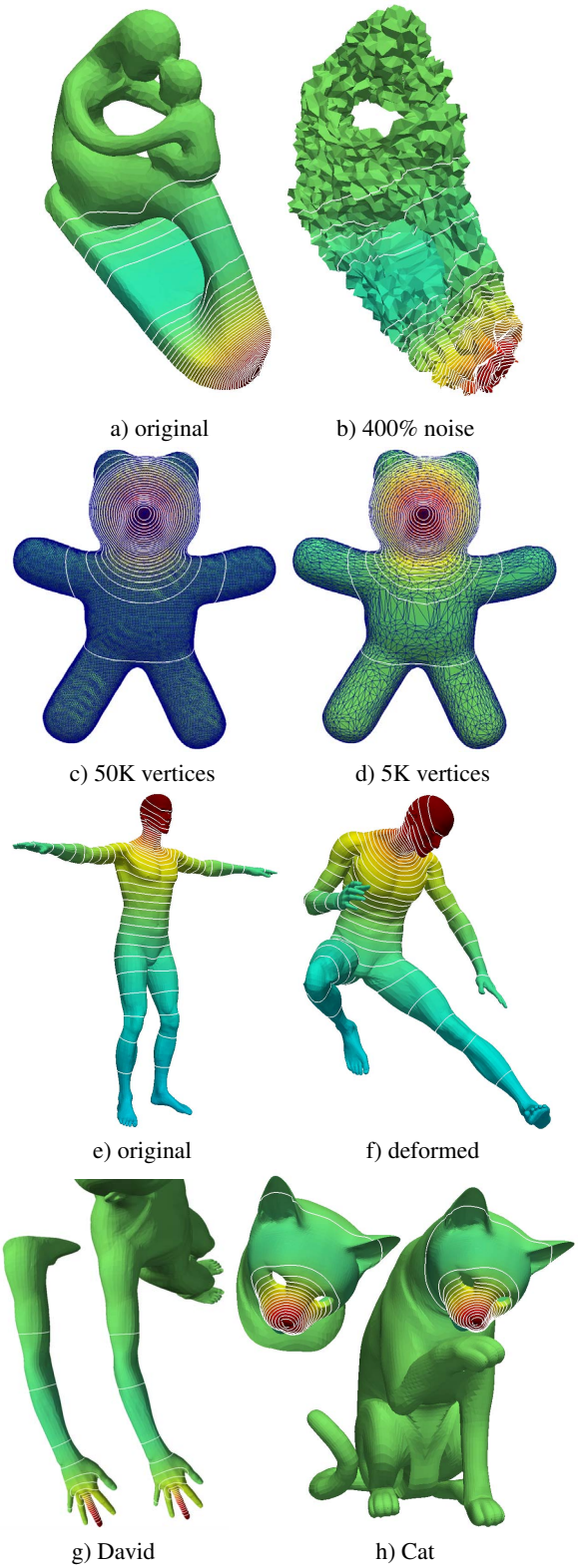



**Figure 5:** *Insensitivity of the multiscale biharmonic kernel to noise, tessellation, deformation, and partial object.*

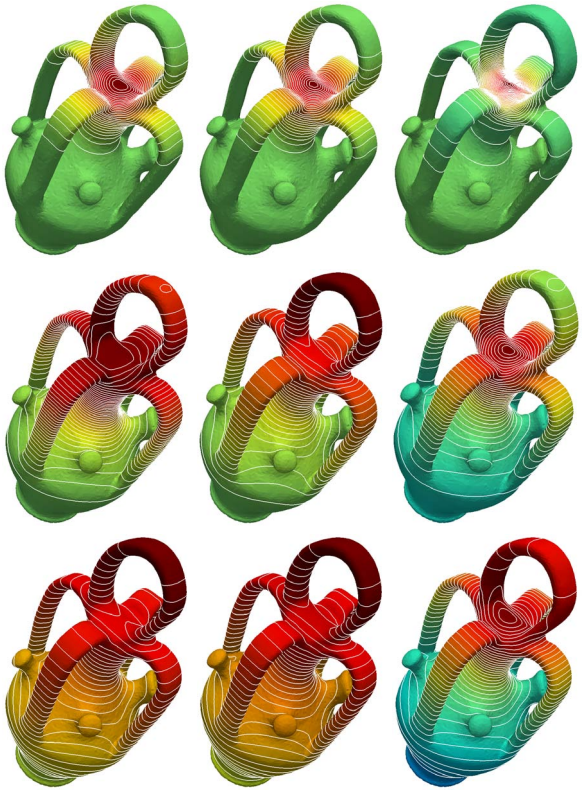



**Figure 6:** *Comparison to heat kernel for different t.*

and report timings for two different computation approaches. First is the approach of Ovsjanikov et al. [OMMG10] who compute the kernel as a matrix exponential using the Expokit [Sid98]. Second is the approach based on the spectral expansion of the heat kernel and requires computing eigenvalues and eigenvectors of the Laplacian; for each $t$ we compute enough eigenvalues to attain $e^{-t\lambda_{max}} < 10^{-6}$. All of the kernels require the Laplace-Beltrami matrix so its assembly time is not reported.

We empirically observe that the maximum computation time (over different scales $t$) for both multiscale pre-biharmonic and biharmonic kernels is approximately linear in the mesh size and is generally better than or is as good as the matrix exponential heat kernel computation time at the smallest $t$. The spectral approach to the heat kernel is relatively fast, but it slows down for smaller values of $t$ because more and more eigenvalues are required.

For the last two models (Cat0 and Michael0 from TOSCA [BBK08]) the heat kernel computation using matrix exponential requires drastically more time than previously (this effect is not due to swapping); on the contrary, the spectral approach leads to smaller compute times for the largest mesh. This means that the heat kernel computation time can depend critically on subtle mesh properties, whereas our ker-

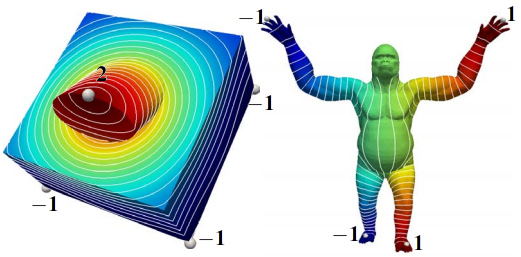



**Figure 7:** *Interpolation using multiscale pre-biharmonic kernel as RBF. Grey spheres depict the anchor points, the annotation shows the prescribed function values at anchors. In the left model there is an invisible anchor point at the fourth vertex in the base with the value of* $-1$.

nel compute times are insensitive to such subtleties. Finally, while a multiresolution approach to computing heat kernel has recently been proposed [VBCG10], we observe that such an approach can also be devised for our kernels due to their insensitivity to tessellation (Figure 5 c-d).

## 7. Applications

**Function interpolation:** An application that can benefit from kernels defined on surfaces is interpolation. For a set of anchor points $y_i$, $i = 1,..,m$, where function values $g_i$ are given, we would like to construct a function on the entire surface which attains these values as prescribed. To this end, we can use either of our kernels within the Radial Basis Function (RBF) interpolation framework [Wen05, Fas07]. The RBF interpolant based on the multiscale pre-biharmonic kernel has the form $g(x) = \sum_j c_j P_t(x, y_j)$, where $x$ is any point on the surface, $t$ is a choice of the scale parameter, and $c_j$ are some coefficients. These coefficients are found by solving a linear system of equations that forces the interpolant to attain the prescribed values at the anchors, $g(y_i) = g_i$. Namely, define the $m \times m$ matrix $M$ by $M_{ij} = P_t(y_i, y_j)$ and let $\vec{c}$ and $\vec{g}$ be column-vectors of coefficients and the prescribed function values; the system we solve is $M\vec{c} = \vec{g}$. The pre-biharmonic kernel is more suitable for interpolation both because it is "non-oscillatory" in the RBF terminology, and because it allows a uniform and intuitive choice for the scale parameter, $t = t_{max}/m = Area/m$. Examples of interpolants obtained via this procedure are shown in Figure 7.

**Correspondence propagation:** In the pioneering work [OMMG10], Ovsjanikov et al. observe that a multiscale kernel, namely the heat kernel, can be used for extending a sparse correspondence between two isometric surfaces into a dense correspondence. They prove that theoretically a single landmark correspondence is enough to reconstruct the isometry map; to obtain good maps in practice, they use two landmark correspondences together with a local shape descriptor. In this context we will show that the multiscale biharmonic kernel provides a practical alternative to the heat kernel.

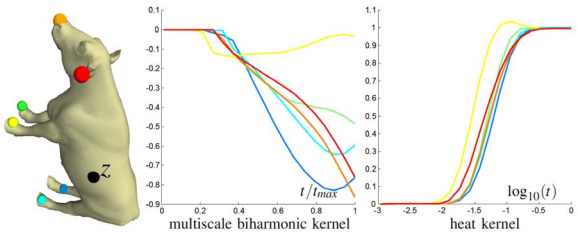



**Figure 8:** *Distinguishing surface points by looking at the t dependence of their kernels (multiscale biharmonic vs. heat) with respect to a fixed point z.*

Given a fixed point $z$ on a surface, for every point $x$ consider the function $h_x(t) = H_t(x, z)$ of variable $t$; here $H_t(\cdot,\cdot)$ is the heat kernel. Under mild genericity conditions, Ovsjanikov et al. [OMMG10] show that $h_x(\cdot) \equiv h_y(\cdot)$ if and only if $x = y$. In other words, to distinguish any two surface points it is enough to look at the $t$ dependence of their heat kernels with respect to a fixed point. This result crucially depends on the multiscale nature of the heat kernel, which naturally leads us to defining a similar function $b_x(t) = B_t(x, z)$ in terms of the multiscale biharmonic kernel $B_t(\cdot,\cdot)$. While proving a theoretical result as above for our kernel can be difficult and is beyond the scope of this paper, we were able to observe practically that $b_x(t)$ can provide a better degree of distinction than $h_x(t)$.

In Figure 8, the fixed point $z$ is shown as a black sphere; for each of the remaining points, shown as a sphere of some color, we plot the corresponding graphs of functions $b_x(t)$ (left) and $h_x(t)$ (right) using the same color. Following [OMMG10] we scale the surface to have unit area, and compute the heat kernel based function $h_x(t)$ for twenty values of $t$ *logarithmically* sampled from $[0.001, 1]$. The multiscale biharmonic kernel based function $b_x(t)$ is computed for twenty values of $t/t_{max}$ *linearly* sampled from the same interval. While both of the functions distinguish all the points, note that the heat kernel has difficulty in distinguishing the orange (nose) and the cyan (back foot) points. One of the reasons for this seems to be the limiting behavior of the kernels: since the heat kernel converges to a constant function with increasing $t$, in a sense it has a smaller "window of opportunity" for detecting the differences between points. On the other hand, our kernel converges to the global biharmonic kernel as $t/t_{max} \to 1$ which readily results in a better distinction. Note also that our multiscale kernel captures more information than the global biharmonic kernel: the darkest blue (back knee) and the red (ear) points are not distinguished by the global kernel as seen by the coinciding function values at $t/t_{max} = 1$, whereas for $t/t_{max} < 1$ the graphs are distinct.

To obtain a dense correspondence map between two surfaces $S$ and $\tilde{S}$ we use the following setup. Given pairs $(z_i, \tilde{z}_i) \in S \times \tilde{S}$, $i = 1,...,m$, of corresponding landmark points, consider the functions $b^i_x(t) = B_t(x, z_i)$ for $x \in S$ and similarly the functions $\tilde{b}^i_y(t) = \tilde{B}_t(y, \tilde{z}_i)$ for $y \in \tilde{S}$. These func-

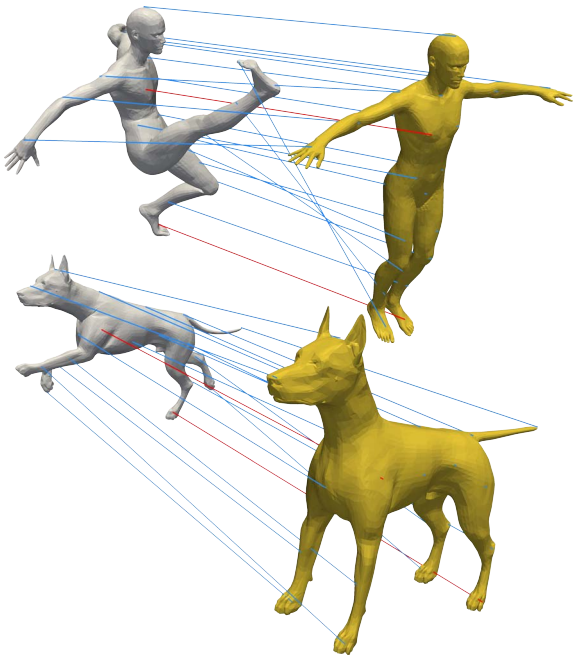


**Figure 9:** *Given two landmark correspondences shown in red, we obtain a dense correspondence shown in blue.*

tions are computed at values $\{t_j\}$ as in the above experiment. For each point $x \in S$ we define the corresponding point $\tilde{x} \in \tilde{S}$ via $\tilde{x} = \arg\min_y \sum_{i,j} (\tilde{b}^i_y(t_j) - b^i_x(t_j))^2$. Figure 9 depicts dense correspondences obtained from two landmark correspondences (shown in red) using this approach. Note that all of the points are placed in the correct relative locations. We believe that the precision of placements can be further improved by incorporating local shape descriptors as in [OMMG10].

## 8. Discussion and future work

This work presents a novel construction of two kernels that share common properties of being multiscale, insensitive to isometric deformations, stable on partial models, robust to noise and tessellation, and fast to compute. The multiscale pre-biharmonic kernel is more amenable to theoretical investigation, has smaller supports that are easier to control via a uniform setting of the scale parameter, and is faster to compute. The multiscale biharmonic kernel has a non-trivial limiting behavior, has a better memory of the center-point, and exhibits wavelet properties. These differences may make one kernel more suitable than the other in a given context, which provides a greater assortment of choices and may eventually lead to an increase in the scope of applications that can benefit from multiscale kernels.

This work provides a small, first step and therefore has limitations that suggest topics for future work. A first topic suitable for further investigation is to more formally characterize the theoretical properties of these multiscale kernels. The vanishing integral property of the multiscale biharmonic kernel hints to its wavelet nature; thus, studying its wavelet properties is an attractive topic. In addition, investigating the locally supported RBFs obtained from our kernels both on the plane (more generally in $\mathbb{R}^n$) and the sphere can be interesting. A second topic for future research is developing multiresolution approaches to kernel computation. Indeed, since tessellation has little effect on the kernels, one can compute an initial approximation on a coarse mesh, and then use this approximate solution to "warm" start the optimizer on the finer mesh – allowing to obtain a fast and high quality solution that incorporates the second order features of the mesh. Finally, an interesting avenue for future work would be to investigate applications of multiscale kernels in computer graphics and geometric modeling. Furthermore, graph Laplacians can be used instead of the Laplace-Beltrami operator in our construction in order to generalize the kernels to the graph setting. As a result, we expect these kernels to have applications in areas such as high dimensional data analysis.

## Acknowledgments

We thank the reviewers for the constructive comments and suggestions that helped to streamline and clarify the exposition of the paper significantly. For the 3D models we thank Daniela Giorgi, AIM@SHAPE, and Project TOSCA [BBK08].

## References

[ARV10] ANTOINE J.-P., ROŞCA D., VANDERGHEYNST P.: Wavelet transform on manifolds: old and new approaches. *Appl. Comput. Harmon. Anal. 28*, 2 (2010), 189–202. 2

[ASGCO10] AVRON H., SHARF A., GREIF C., COHEN-OR D.: $l_1$-sparse reconstruction of sharp point set surfaces. *ACM Trans. Graph. 29* (November 2010), 135:1–135:12. 3

[BBK08] BRONSTEIN A., BRONSTEIN M., KIMMEL R.: *Numerical Geometry of Non-Rigid Shapes*. Springer, 2008. 8, 10

[BNS06] BELKIN M., NIYOGI P., SINDHWANI V.: Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *Journal of Machine Learning Research 7* (2006), 2399–2434. 3

[BS08] BOTSCH M., SORKINE O.: On linear variational surface deformation methods. *IEEE Trans. Vis. Comput. Graph. 14*, 1 (2008), 213–230. 3

[BV04] BOYD S., VANDENBERGHE L.: *Convex optimization*. Cambridge University Press, Cambridge, 2004. 5

[CDS98] CHEN S. S., DONOHO D. L., SAUNDERS M. A.: Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput. 20*, 1 (1998), 33–61. 2

[CM06] COIFMAN R. R., MAGGIONI M.: Diffusion wavelets. *Appl. Comput. Harmon. Anal. 21*, 1 (2006), 53–94. 2

[dGGV08] DE GOES F., GOLDENSTEIN S., VELHO L.: A hierarchical segmentation of articulated bodies. *Computer Graphics Forum 27*, 5 (2008), 1349–1356. 2

[EHJT04] EFRON B., HASTIE T., JOHNSTONE I., TIBSHIRANI R.: Least angle regression. *Ann. Statist. 32*, 2 (2004), 407–499. With discussion, and a rejoinder by the authors. 4

[Fas07] FASSHAUER G. E.: *Meshfree approximation methods with MATLAB*, vol. 6 of *Interdisciplinary Mathematical Sciences*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2007. 9

[GM09] GELLER D., MAYELI A.: Continuous wavelets on compact manifolds. *Math. Z. 262*, 4 (2009), 895–927. 2

[HVG11] HAMMOND D. K., VANDERGHEYNST P., GRIBONVAL R.: Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis 30*, 2 (2011), 129 – 150. 2

[KKL*07] KIM S.-J., KOH K., LUSTIG M., BOYD S., GORINEVSKY D.: An Interior-Point Method for Large-Scale L1-Regularized Least Squares. *IEEE Journal of Selected Topics in Signal Processing 1*, 4 (2007), 606–617. 5

[Lav02] LAVERY J. E.: Shape-preserving, multiscale interpolation by univariate curvature-based cubic l1 splines in cartesian and polar coordinates. *Computer Aided Geometric Design 19*, 4 (2002), 257–273. 3

[LRF10] LIPMAN Y., RUSTAMOV R. M., FUNKHOUSER T. A.: Biharmonic distance. *ACM Trans. Graph. 29*, 3 (2010). 2, 4, 6

[MDSB03] MEYER M., DESBRUN M., SCHRÖDER P., BARR A. H.: Discrete differential-geometry operators for triangulated 2-manifolds. In *Visualization and Mathematics III*, Hege H.-C., Polthier K., (Eds.). Springer-Verlag, Heidelberg, 2003, pp. 35–57. 3

[MOS11] *The MOSEK optimization toolbox for MATLAB manual. Version 6.0 (Revision 106)*. MOSEK ApS, Denmark, 2011. 5, 6

[OMMG10] OVSJANIKOV M., MÉRIGOT Q., MÉMOLI F., GUIBAS L. J.: One point isometric matching with the heat kernel. *Comput. Graph. Forum 29*, 5 (2010), 1555–1564. 2, 8, 9, 10

[PP93] PINKALL U., POLTHIER K.: Computing discrete minimal surfaces and their conjugates. *Experimental Mathematics 2* (1993), 15–36. 3

[Sid98] SIDJE R. B.: Expokit: a software package for computing matrix exponentials. *ACM Trans. Math. Softw. 24* (March 1998), 130–156. 8

[SOG09] SUN J., OVSJANIKOV M., GUIBAS L.: A concise and provably informative multi-scale signature based on heat diffusion. *Computer Graphics Forum 28*, 5 (2009), 1383–1392. 2

[SS95] SCHRÖDER P., SWELDENS W.: Spherical wavelets: efficiently representing functions on the sphere. In *SIGGRAPH* (1995), pp. 161–172. 2

[Swe98] SWELDENS W.: The lifting scheme: a construction of second generation wavelets. *SIAM J. Math. Anal. 29*, 2 (1998), 511–546. 2

[Tib96] TIBSHIRANI R.: Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B 58*, 1 (1996), 267–288. 2, 4

[VBCG10] VAXMAN A., BEN-CHEN M., GOTSMAN C.: A multi-resolution approach to heat kernels on discrete surfaces. *ACM Transactions on Graphics 29*, 4 (2010), 121:1–121:10. 9

[Wen05] WENDLAND H.: *Scattered data approximation*, vol. 17 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2005. 9

## Appendices

**Appendix A:** The pre-biharmonic kernel optimization problem can be transformed into the lasso formulation of Eq. (1) as follows. To get rid of the weighting within the $l_1$ norm, we make the substitution $\vec{g} = W\vec{f}$. Let $M = W^{-1/2}LW^{-1}$, and note that $M\vec{g} = W^{-1/2}L\vec{f}$. The constraint $f_k = 1$ becomes $g_k = W_{kk}$. To drop this constraint, let $\vec{b} = W_{kk}M_k$, where $M_k$ is the $k$-th column of matrix $M$, and define the $n \times (n-1)$ matrix $A$ as equal to $M$ with the $k$-th column deleted. It is easy to check that the problem

$$\min_{\vec{u}} \|A\vec{u} - \vec{b}\|_2^2 \qquad \text{subject to} \quad \|\vec{u}\|_1 \leq t',$$

is equivalent to our pre-biharmonic kernel problem. Note that $\vec{u} \in R^{n-1}$ contains all the entries of $\vec{g} \in R^n$ except the entry $g_k$ which is constrained to be equal to $W_{kk}$. Thus, the new parameter $t' = t - W_{kk}$. Note that $\vec{b}$ is essentially a scaled $k$-th column of the cotangent Laplacian, and so $\vec{b}$ has vanishing entries except for the entries corresponding to mesh vertex $v_k$ and its one-ring neighbors.

**Appendix B:** Here we show that the solution to the multiscale biharmonic optimization problem when the lasso constraint becomes inactive is given by $f^*(x) = G(x,y)/G(y,y)$ for all $x \in S$, where $G(x,y)$ is the global biharmonic kernel. Let $\{\lambda_i, \phi_i\}_{i=0}^{\infty}$ the Laplace-Beltrami eigenvalues and eigenfunctions; the eigenfunctions are orthonormal, and $\lambda_0 = 0$ with corresponding eigenfunction $\phi_0 = const$. The global biharmonic kernel can be written as $G(x,y) = \sum_{i>0} \phi_i(x)\phi_i(y)/\lambda_i^2$.

Writing $f^*$ in term of the eigenfunction basis we obtain

$$f^* = c_1\phi_1 + c_2\phi_2 + c_3\phi_3 + \cdots, \tag{2}$$

where we do not have the term $c_0\phi_0$ because $c_0 = \int_S f^*\phi_0 = const \times \int_S f^* = 0$ by the vanishing integral constraint in the optimization problem. Since all eigenfunctions $\phi_i$ for $i \neq 0$ are orthogonal to $\phi_0$, we obtain that $\int_S \phi_k = 0$. This discussion implies that requiring the expansion of Eq. (2) is equivalent to the vanishing integral condition $\int_S f^* = 0$.

Now we rewrite the objective functional and the remaining constraint $f^*(y) = 1$ in terms of $\{c_i\}_{i=1}^{\infty}$ to obtain the following problem:

$$\min_{\{c_k\}} \sum_{i>0} c_i^2\lambda_i^2 \qquad \text{subject to} \quad \sum_{i>0} c_i\phi_i(y) = 1.$$

Here we used the facts that $\Delta\phi_i = \lambda_i\phi_i$ and that the eigenfunctions are orthonormal, $\int_S \phi_i\phi_j = \delta_{ij}$. To solve the problem, make a substitution $z_i = c_i\lambda_i$, which gives the problem,

$$\min_{\{z_i\}} \sum_{i>0} z_i^2 \qquad \text{subject to} \quad \sum_{i>0} z_i\phi_i(y)/\lambda_i = 1$$

which is equivalent to finding the point $\{z_i\}_{i=1}^{\infty}$ on the hyperplane $\sum z_i\phi_i(y)/\lambda_i = 1$ closest to the origin. It is easy to see that this point must have the coordinates $z_i = \alpha\phi_i(y)/\lambda_i$, where $\alpha$ is some constant. We find that $c_i = \alpha\phi_i(y)/\lambda_i^2$ and using the expansion Eq. (2) we obtain

$$f^*(x) = \alpha \sum_{i>0} \frac{\phi_i(x)\phi_i(y)}{\lambda_i^2} = \alpha G(x,y).$$

From the constraint $f^*(y) = 1$, we find the value of $\alpha = 1/G(y,y)$ , which establishes that $f^*(x) = G(x,y)/G(y,y)$.