

Nonconvex optimization on data manifold by accelerated curvature transport*

E. Jonckheere^{1,*}, H. Krovi², and A. A. Rompokos¹

Abstract—In this paper, we explore a new approach to optimization of cost or utility functions defined over a surface, a manifold, or its simplicial decomposition. In the era of Big Data, heterogeneous signal samples sometimes embed with less distortion in a lower dimensional space if the embedding space is a manifold rather than the traditional Euclidean space. If a utility function is defined over the data and if there is a need to identify significant events defined by extreme values of the utility function, we are faced with the problem of identifying the extreme minima/maxima points of the cost/utility function defined over the manifold or its triangulation. The fundamental idea developed here is to observe that at the extreme points the graph of the utility function has extreme curvature. Accordingly, the celebrated Ricci/Yamabe flow for uniformization of the curvature of the graph will show significant “curvature transport” in the vicinity of the extreme values, hence allowing their rapid identification, obviating the classical sorting. The novel theoretical contribution is to accelerate the process by compounding the Laplace operator.

I. INTRODUCTION

In convex optimization, much effort has been devoted to the problem of *accelerating* the process of reaching the minimum (maximum) of some cost (utility) function f . If f is defined over \mathbb{E}^n , the best example is provided by Nesterov’s *accelerated* gradient descent, allowing the convergence rate to increase from $O(1/t)$ to $O(1/t^2)$. Already over \mathbb{E}^n , the landscape—the graph of f —can be very challenging, with sharp peaks and narrow valleys [21, Fig. 3]. If f is defined over a (greedy hyperbolic embedding) manifold, a data analytics simplicial complex, or a data stream graph, the problem becomes even more challenging.

In this paper, we precisely exploit the challenging nature of the landscape, by observing that f has its minima (maxima) encoded in the curvature of the graph as $\text{Hess}_f(x_*) \gg 0$ ($\text{Hess}_f(x_*) \ll 0$) and that an algorithm sensitive to the curvature should efficiently single out the minima (maxima). This algorithm is the Yamabe curvature uniformization running on the graph of f . This algorithm treats curvature as mass, to be moved in the Monge-Kantorovic sense, to level off the curvature landscape, a process that singles out the maxima (minima) as areas of maximum mass transport.

In the case of a simplicial complex, or surface triangulation, with vertex set $\{v_i\}_{i=1}^N$ (or a graph with N vertices) this process must be confronted with the “brutal” $\max_i f(v_i)$, which classically requires $\Omega(N)$ queries and $O(\sqrt{N})$ quantum queries [1]. (The `qsort` requires $\Omega(N \log_2(N))$ steps

¹ Dept. of Elec. Eng., Univ. of Southern California, Los Angeles, CA 90089 {jonckhee, rompokos}@usc.edu

² Quantum Information Processing Group, Raytheon BBN Technologies, Cambridge, MA hari.krovi@raytheon.com

* Supported by NSF CCF-1423624

TABLE I
FUNDAMENTAL COMBINATORIAL TOPOLOGY STRUCTURES

combinatorial notation	
$\mathcal{V} = \{v_i\}_{i=1}^N$	vertex set
$\mathcal{E} = \{v_i v_j\} \subseteq \mathcal{V} \times \mathcal{V}$	edge set
$\mathcal{G} = ((\mathcal{V}, \mathcal{E}), w)$	graph with $w : \mathcal{E} \rightarrow R_{>0}$
\mathcal{L}	Laplacian of graph \mathcal{G}
$\Sigma = \{\sigma_{jik} = v_j v_i v_k\} \subset \mathcal{V} \times \mathcal{V} \times \mathcal{V}$	2-simplex set subject to rules of simplicial topology
$T(\mathcal{V}, \mathcal{E}, \Sigma)$	triangulation of surface
$K(v_i)$	curvature of T at v_i
$\bar{K}(v_i)$	curvature of graph of f at v_i

for both classical and quantum sorting [14].) Here, the sorting is obviated by having each vertex of the combinatorial manifold monitor its curvature rate of change and returning the vertex that first reached a preset rate threshold (**Algorithm 1**). To reach the **return** even faster, it is essential to *accelerate* the curvature uniformization—to make the curvature flow faster—by compounding the Laplace operator of a heat equation formulation of the Yamabe flow. This has the effect that a very short amount of integration time of the Yamabe flow is enough for one of the vertices of the manifold to reach the threshold, singling out a minimum (maximum).

A comparison can be drawn with the way Adiabatic Quantum Computations (AQC) avoid sorting by identifying a (minimum energy) curve that terminates at the minimum. Here, the situation is somehow time-reversed. As soon as the flow algorithm starts, we identify a curve by its maximal rate of change, a curve that itself identifies the max (min).

A. Rough paper outline

The paper is divided into two parts: The first part (Secs. III, IV) deals with optimization over continuous geometry domains, mostly dealing with the Riemannian geometry definition of curvature; the second part (Secs. V, VI) deals with optimization over discrete geometry domains, more specifically triangulations of compact surfaces, where the notion of curvature is considered in terms of Alexandrov angles.

The commonality between the two parts is secured by the Yamabe flow, where the crucial transition from continuous to combinatorial geometry was developed in a series of papers by Feng Luo [17], [18], [7].

The fundamental combinatorial structures are outlined in Table I-A.

B. Some motivation

Contrary to traditional wisdom, several interesting optimization problems involve non-convex functions. One of

our motivations is quantum tomography and quantum control [16], [21], where optimization of non-convex functions is at the heart of several problems. One example is to determine the optimal strategy to estimate an unknown phase shift in a Hamiltonian. Here optimization over adaptive strategies where successive measurements depend on the outcomes of the previous ones, can lead to a non-convex optimization problem [12].

II. THE “EARTH MOVING” AND HEAT DIFFUSION METAPHORS

The *Earth Moving* concept traces back to Monge and was later revived by Kantorovich. Assume $f(x, y)$ is the height of the sand on a landscape and let us visualize the process of moving sand from the hill tops to the valleys. Assume we are able to register the rate of sand removal from the hill tops to the valleys. The guiding idea is that the point that has received (given up) sand at the highest rate to level off the landscape is the minimum (maximum). *However, here, we are going one step further than this classical idea and argue that the process can be accelerated by moving curvature instead of earth of the landscape.*

Probably at this juncture the heat diffusion provides a better metaphor. Clearly, the point of maximum temperature could be identified at the one shedding “calories” at the highest rate, having the largest heat flux, something that can be identified very early on in the course of the heat diffusion process. As we show in Secs. III-C, III-D, taking the spatial derivative of the temperature yields the curvature of the temperature profile. This explains at least partially why the (Ollivier-Ricci) curvature appears to be the correct network measure to anticipate how efficiently the Heat Diffusion protocol would run on such network [3], [22], [23].

III. SIMPLEST EXAMPLES OF 1-DIMENSIONAL OPTIMIZATION BY CURVATURE TRANSPORT

A. Function on unit circle, graph plotted on cylinder

Consider a twice-differentiable function $f : S^1 \rightarrow \mathbb{R}$, $\theta \mapsto f(\theta)$, defined on the unit circle together with the Haar measure $\mu(d\theta) = d\theta/2\pi$. The graph of f could be thought of as plotted on the cylinder $S^1 \times \mathbb{R}$. Nevertheless, it could also be plotted on the “flat cylinder” $[0, 2\pi] \times \mathbb{R}$, with the convention that $f^{(0,1,2)}(0) = f^{(0,1,2)}(2\pi)$. On the flat cylinder, the arc length is $ds = \sqrt{1 + f'(\theta)^2}d\theta$ and the curvature of the graph, defined as the inverse radius of the osculating circle at $(\theta, f(\theta))$, is $\kappa(\theta) = f''(\theta)/(1 + f'(\theta)^2)^{3/2}$. Our objective is to locate the minima/maxima by curvature uniformization. A first question is to what constant value should the curvature be uniformized? This constant value is precisely a topological invariant and is revealed by the total curvature:

$$\oint \kappa(\theta)ds = \int_0^{2\pi} \frac{df'(\theta)}{1 + f'(\theta)^2} = [\arctan f'(\theta)]_0^{2\pi} = 0.$$

Consider now a function with its curvature $\kappa(\theta)$ varying along the unit circle, together with the problem of finding its minimum. One can easily visualize a deformation of the

graph of f on the cylinder $S^1 \times \mathbb{R}$ to a minimal length curve on the cylinder [8], but this deformation has to be smooth. This is formalized in the concept of isotopy from f to the constant function 0 on the unit circle. To be precise, this isotopy is a smooth map $F : S^1 \times [0, 1] \rightarrow \mathbb{R} \times [0, 1]$, $(\theta, t) \rightarrow (F_t(\theta), t)$ such that F_t is an embedding and $F_0 = f$ and $F_1 = 0$. Under such process, the minima of f are shedding their curvatures and the maxima are receiving them, in a process called curvature transport. Focusing on the transport, rather than the deformation, the former could take the form of several popular processes. Here in Sec. III-C we propose a heat diffusion transport process, although other transport processes, e.g., the gradient flow of the Wasserstein distance between the current curvature measure and the vanishing one, could be considered.

B. Function on unit circle, graph plotted on flat cylinder

Instead of plotting the graph of $f : S^1 \rightarrow \mathbb{R}$ in the cylinder we could plot it radially outside the circle. If κ is now the curvature of the graph of the function f , the invariant becomes $\oint \kappa(s)ds = 2\pi$. By the Whitney-Graustein theorem, it is an isotopy invariant.

C. Heat equation formulation of curvature transport

To begin in a simple context, consider the function $f : \mathbb{R} \rightarrow \mathbb{R}$. Clearly, $f''(x) > 0$ indicates a local minimum while $f''(x) < 0$ indicates a local maximum. At this stage, we could envision the Heat Equation algorithm

$$\frac{\partial F(x, t)}{\partial t} = \frac{\partial^2 F(x, t)}{\partial x^2}, \quad F(x, 0) = f(x),$$

with asymptotic solution $F(x, \infty) = \text{constant}$. This equation *indirectly* smooths over the curvature by leveling off the values of f . Clearly, at $t = 0$, the narrow valleys have large $F'' := \frac{\partial^2 F}{\partial x^2} \gg 0$ and hence the rate of change will be very high and easily registrable. The preceding can easily be generalized to the classical Laplacian of \mathbb{E}^3 in the RHS, even the Laplace Beltrami operator on a Riemannian manifold.

D. Accelerating the heat equation

As a prelude to the Hamilton-Ricci flow on a surface (Eq. (2)) or the Yamabe flow (Eq. (4)) on a surface triangulation [17], we go one step further: instead of uniformizing the level of the function f we uniformize its curvature, and we run the heat equation on the curvature:

$$\frac{\partial}{\partial t} \frac{\partial^2 F(x, t)}{\partial x^2} = \frac{\partial^2}{\partial x^2} \frac{\partial^2 F(x, t)}{\partial x^2}, \quad (1)$$

under the assumption that F'' is an approximation of the curvature $F''/(1 + F'^2)^{3/2}$ if F' is small. Observe that in the RHS of Eq. (1) we have compounded the Laplace operator. Of course, one might question why this latter equation involving the biharmonic operator $\frac{\partial^4}{\partial x^4}$ should do any better than the simplified equation involving the harmonic operator $\frac{\partial^2}{\partial x^2}$. The issue is that, if a function has a minimum in a sharp valley, as we already know its second derivative $F'' := \frac{\partial^2 F}{\partial x^2}$ is large, but its 4th, 6th, 8th etc. derivatives

could be even larger, hence could yield a faster and faster rise of the function at its original minimum point.

As an illustration, consider the usual Gauss distribution

$$g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}.$$

Clearly, $g''(0) = -1/(\sqrt{2\pi}\sigma^3)$; however, we could make the maximum even sharper by taking the even order n th derivatives at the maximum point $x = 0$:

$$\frac{d^n g(0)}{dx^n} = (-1)^{n/2} \frac{1}{\sqrt{2\pi}} \frac{(n-1)!!}{\sigma^{n+1}}.$$

The drawback is that if we look at the function $\frac{d^n g(x)}{dx^n}$ it shows some additional oscillations in the neighborhood of $x = 0$; however, the ratio of the heights of those oscillations over the height of the maximum oscillation at $x = 0$ goes to zero as $\sigma \downarrow 0$.

Clearly, the algorithm reveals its full power in those problems where the landscape has very narrow “valleys”/very sharp “peaks.” These are precisely the local minima/maxima that are difficult to capture, for the step of a gradient algorithm might have to be taken unreasonably small, resulting in an unreasonably large amount of time to reach the minimum/maximum. The proposed variant of the Earth Moving algorithm is based on the obvious fact that, if the valley is really narrow, it would rise very quickly if sand is moved in.

E. Jet local structure

It remains to define those functions that benefit from taking higher even-order derivatives in the RHS of the heat related equations. Obviously, this is an issue of the local structure of the function f around the singularity. Recall that the k -jet of a C^k function f at a point x_* , $j^k f(x_*)$, is the collections of all derivatives of f up to order k at x_* . The infinite jet $j^\infty f(x_*)$ of a C^∞ function is the set of all derivatives of all orders [9]. The nontrivial fact is that, even if the sequence in $j^\infty f(x_*)$ grows without bound, there still exists a smooth map germ whose derivatives match the coefficients in $j^\infty f(x_*)$ (see [15, Th. D.9].)

Definition 1: A function $f \in C^\infty$ has a ∞ -sharp minimum at x_* if the sequence of coefficients in its jet $j^\infty f(x_*)$ starts at $0 = f'(x_*)$ and then grows without bound on the even components. $f \in C^k$ has a k -sharp minimum if the restricted sequence in $j^k f(x_*)$ starts at 0 and is then monotone increasing on the even components.

F. Optimization under time-varying landscape

Under some conditions, this algorithm can be made adaptive under deformation of the cost/utility function f , provided the algorithm is *accelerated* consistently with the speed of the varying landscape. The problem is that, since the algorithm relies on the singularity of the cost/utility function, this adaptation would work only if the singularity structure of the function survives the deformation. This is the concept of stability of maps [10]:

Definition 2: Consider a smooth function $f : M \rightarrow \mathbb{R}$ defined over a (compact) smooth manifold M . The function $f \in C^\infty(M, \mathbb{R})$ is said to be stable if for any sufficiently small deformation in the Whitney (weak or compact-open) topology $\tilde{f} \in C^\infty(M, \mathbb{R})$ of f , there are maps d_x, d_y , smooth together with their inverses such that $\tilde{f} = d_y \circ f \circ d_x$.

Lemma 1 ([6]): The set of smooth functions defined over a compact smooth manifold M has a stratification $C^\infty(M, \mathbb{R}) = F^0 \sqcup F^1 \sqcup \dots$ such that F^i has codimension i and F^0 is the subset of Morse functions with distinct critical values. The latter are stable; hence stable functions are everywhere dense. ■

As a corollary, generically, the singularity structure is preserved.

The problem is that we accelerate the algorithm by subjecting f to the Laplace operator and it turns out that under this process the Morse property of f might be lost:

Counter-Example 1: Consider the Morse function $f(x) = x^2 + \sum_{k=1}^{\infty} x^{4k}$ defined for $x \in (-1, +1)$ and observe that $(\partial/\partial^2)^2 f(x) = 24 + (8!/4!)x^4 + \text{higher order terms}$ ¹. The latter is not Morse as the $x = 0$ singularity has codimension 2. Hence, using the notation of Lemma 1, $f \in F^2$.

However, this negative result is by far compensated by the remarkable property that the heat equation “discovers all by itself minimal $[F^0]$ Morse functions”²:

Lemma 2 ([19], [20]): Let M be a torus or a sphere. Then for some generic initial condition $f(\cdot, 0)$, there exists a T such that, for $t > T$, the solution of the heat equation $f(\cdot, t) \in F^0$.

IV. HAMILTON’S RICCI FLOW ON GRAPH OF FUNCTION DEFINED ON SURFACE

Here we redo on a surface what we did in Sec. III-A on the circle.

A. Graph of a function on surface

Algorithm (1) is nothing other than the Ricci flow to uniformize the curvature of a manifold (or graph, or simplicial complex) subject to the global Euler characteristic invariant. Consider for the time-being a (compact, oriented) surface S with uniform curvature K together with a distance $d(\cdot, \cdot)$ and a real-valued function $f : S \rightarrow \mathbb{R}$. Let (x, y) be local coordinates of S in a coordinate patch and let ds^2 be the intrinsic geometric Gauss metric. Consider the vector bundle (E, p, S) over $S \subset \mathbb{E}^3$ where the bundle projection $p : E \rightarrow S$ is the orthogonal projection onto S . Consider the tubular neighborhood [13, p. 109] of S defined by (E, p, S) . In this tubular neighborhood, define the surface \tilde{S} with local coordinates $(x, y, \varepsilon f(x, y))$ in the same patch as that of S . Note the scaling factor ε to avoid the emergence of swallow tail singularities [2, Fig. 9] in the surface \tilde{S} .

Lemma 3: Let $R_1(x, y) \leq R_2(x, y)$ be the principal curvature radii of the surface S at (x, y) . Then \tilde{S} has no swallow tail singularities in the local patch of (x, y) if and only if $\varepsilon f(x, y) < R_1(x, y)$. ■

¹Many thanks to Prof. F. Bonahon for constructing this counterexample.

²Quoted verbatim from [19].

From here on, it is assumed the condition of Lemma 3 holds. The surface \tilde{S} receives from S a Riemannian (or Gauss) metric

$$d\tilde{s}^2 = ds^2 + \varepsilon^2 \left((f'_x)^2 dx^2 + (f'_y)^2 dy^2 \right) = \tilde{g}_{xx} dx^2 + \tilde{g}_{yy} dy^2.$$

The integration of $d\tilde{s}$ along a geodesic yields the distance \tilde{d} . The nonuniformity of the curvature of the surface \tilde{S} indicates the presence of minima and maxima. The action of uniformizing the curvature would reveal the dips, the valleys, the tops of the graph \tilde{S} of the function f defined on the surface S .

B. Hamilton's Ricci flow

We consider Hamilton's formulation [11] of the Ricci flow since it provides the generalization the closest in spirit to Eq. (1). Let \tilde{K} be the scalar curvature of \tilde{S} . Close to the original Ricci flow, the Riemannian metric \tilde{g} is updated as follows:

$$\frac{\partial \tilde{g}}{\partial t} = -(\tilde{K} - \tilde{K}_{\text{av}}) \tilde{g},$$

where $\tilde{K}_{\text{av}} = \frac{\int_{\tilde{S}} \tilde{K} dA}{\int_{\tilde{S}} dA}$. For aesthetic reason, the flow is reformulated in terms of the evolution of the curvature

$$\frac{\partial \tilde{K}}{\partial t} = \Delta \tilde{K} + \tilde{K}(\tilde{K} - \tilde{K}_{\text{av}}), \quad (2)$$

where Δ is the Laplace-Beltrami operator. This continuous geometry equation does not quite compare with the heat equation (1) on the curvature, unless $\tilde{K} \approx \tilde{K}_{\text{av}}$. However, as we shall soon see, the connection becomes much clearer on a triangulated surface.

C. Jet space

We follow up on Sec. III-E with the idea that the function f might be imprecisely known. The jet space [5] is defined as $J^k(S, \mathbb{R}) = \{j^k f(x) : f \in C^k(S, \mathbb{R}), x \in S\}$. The infinite jet space $J^\infty(S, \mathbb{R})$ is defined as the inverse limit of the $J^k(S, \mathbb{R})$'s. The jet space is a manifold. The acceleration of the process applies to the subset of J^k or J^∞ of jets with their coefficients growing.

V. YAMABE FLOW ON GRAPH OF FUNCTION DEFINED ON TRIANGULATION OF SURFACE

A. Curvature of surface triangulation

Let $T_w(\mathcal{V}, \mathcal{E}, \Sigma)$ be a triangulation of the compact, oriented surface S , endowed with a weight $w : \mathcal{E} \rightarrow \mathbb{R}_{>0}$, which itself yields a metric $d : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}_{\geq 0}$. The local curvature at a vertex v_i of the triangulation is defined via angles at v_i in all triangles sharing v_i as apex. The angle $\theta(j\hat{i}k)$ at vertex v_i in the triangle $v_j v_i v_k$ is defined via the cosine law as

$$\theta(j\hat{i}k) = \arccos \left(\frac{d(v_i, v_j)^2 + d(v_i, v_k)^2 - d(v_j, v_k)^2}{2d(v_i, v_j)d(v_i, v_k)} \right).$$

The curvature at vertex v_i is defined as

$$K(v_i) = 2\pi - \sum_{\text{triangles } v_j v_i v_k} \theta(j\hat{i}k),$$

where the sum is extended to all triangles having v_i as apex.

The combinatorial Gauss-Bonnet theorem for the PL metric [17] asserts that $\sum_{i=1}^N K(v_i) = 2\pi\chi(S)$, where $\chi(S) = V - E + F$, where V, E, F are the number of vertices, edges, and faces, resp., regardless of the triangulation, which can be taken arbitrarily.

Clearly the preceding definitions apply to the triangulation \tilde{T} of the surface \tilde{S} . The following lemma connects the triangulation T of S with the triangulation \tilde{T} of \tilde{S} .

Lemma 4: Let S be a compact, oriented surface. Under the condition of Lemma 3, $\chi(S) = \chi(\tilde{S})$.

Proof: Since χ is an invariant independent of the triangulation, take an arbitrary triangulation T of S . Under the condition of Lemma 3, the orthogonal projection $p_\perp : S \rightarrow \tilde{S}$ is one to one. Hence project T orthogonally to \tilde{S} to obtain the triangulation \tilde{T} of \tilde{S} . Combinatorially, \tilde{T} and T are the same triangulation; hence, $\chi(S) = \chi(T) = V - E + F = \tilde{V} - \tilde{E} + \tilde{F} = \chi(\tilde{T}) = \chi(\tilde{S})$. ■

B. Curvature transport

The uniformization of the curvature of \tilde{S} is done by the Yamabe flow [17, Eq. 1.3] on the triangulation \tilde{T} of the surface \tilde{S} :

$$\frac{du(v_i, t)}{dt} = - \left(\tilde{K}_{u(t)}(v_i) - \tilde{K}_{\text{av}} \right) u(v_i, t), \quad u(v_i, 0) = 1, \quad (3)$$

where $\tilde{K}_{u(t)}(v_i)$ is the curvature of the triangulation \tilde{T} of \tilde{S} at vertex v_i , after changing the metric in a conformal manner via

$$\tilde{d}_{u(t)}(v_i, v_j) = u(v_i, t) \tilde{d}(v_i, v_j) u(v_j, t)$$

and $\tilde{K}_{\text{av}} = \sum_{i=1}^N \tilde{K}(v_i)/N = 2\pi\chi(T)/N$ is the average of the curvature of the triangulation of \tilde{T} with its invariance derived from the combinatorial Gauss-Bonnet theorem. The evolution of the curvature $\tilde{K}_{u(t)}$ under the conformal transformation u follows the heat equation:

$$\frac{d\tilde{K}_{u(t)}}{dt} = \mathcal{L} \left(\tilde{K}_{u(t)} \right), \quad \tilde{K}_{u(0)}(v_i) = \tilde{K}(v_i), \quad (4)$$

where \mathcal{L} denotes the "graph Laplacian," defined as

$$\mathcal{L}_{mn} = \sum_{j, k \neq m} \frac{\partial \theta_u(j\hat{m}k)}{\partial u_n} u_n, \quad m, n \in \{i, j, k\},$$

where $\theta_u(j\hat{m}k)$ denote the angle θ at vertex v_m in the triangle $v_j v_m v_k$ subject to the conformally changed metric. The terminology of *graph Laplacian* is justified by the following result:

Lemma 5: The matrix \mathcal{L} is symmetric negatively semi-definite; moreover, $\mathcal{L}_{m \neq n} \geq 0$ and $\mathcal{L}_{nn} = -\sum_{m \neq n} \mathcal{L}_{mn} \leq 0$. Moreover, the upper triangular part of the 3×3 submatrix of \mathcal{L} corresponding to the triangle $v_i v_j v_k$ reads

$$\mathcal{L} = \begin{pmatrix} * & \cot \theta_u(i\hat{k}j) & \cot \theta_u(i\hat{j}k) \\ * & * & \cot \theta_u(k\hat{i}j) \\ * & * & * \end{pmatrix}$$

Proof: The first part is available in [17, Th. 2.1]. The second part is direct from the same theorem after some trigonometric manipulations in the triangle $v_i v_j v_k$. ■

Putting the lemma in other words, \mathcal{L} is the graph Laplacian of $(\mathcal{V}, \mathcal{E})$, but subject to the new metric $D(v_i, v_j) = \tan \theta_u \left(\hat{i} \hat{k} \hat{j} \right)$, $D(v_i, v_k) = \tan \theta_u \left(\hat{i} \hat{j} \hat{k} \right)$, $D(v_j, v_k) = \tan \theta_u \left(\hat{j} \hat{i} \hat{k} \right)$.

Since \mathcal{L} is a genuine graph Laplacian, Eq. (4) is the heat equation on a graph [3]. Hence, $\tilde{K}_{u(t)}(v_i)$ can be thought of as the temperature at the vertices v_i at time t and the corresponding heat flux is the “Earth Moving.”

Equation (3) for the conformal factors *apparently* requires constant update on the curvature $\tilde{K}_{u(t)}(v_i)$, which would be prohibitively time-consuming. However, computation can be cut down by running the heat equation (4) parallel to (3) to get the curvature. Another possibility is to observe that (3) is decentralized and hence the equations for the various v_i 's can be run in parallel.

C. Accelerated curvature transport

The Yamabe flow (4) is to be compared with the heat equation (1) on the curvature F'' introduced in Sec. III-C in the continuous geometry setup. But in order to follow the line of thoughts of this section—especially, the idea of using a biharmonic operator instead of a harmonic operator to speed up the rate of change of the curvature at the extremal points—one should attempt to define a “biharmonic graph Laplacian,” if such concept exists at all. Attempting to define the “biharmonic Laplacian” as \mathcal{L}^2 fails, as it is easily seen that \mathcal{L}^2 does not qualify as a graph Laplacian in the sense of Lemma 5 ($(\mathcal{L}^2)_{m \neq n}$ need not be nonnegative). Here we follow Bauer et al [4]: We define the p -neighborhood graph $\tilde{T}^{[p]}$ of the triangulation \tilde{T} , as having the same vertices as \tilde{T} but with edge set defined as those pairs of vertices exactly within p hops of each other. The p -neighborhood graph Laplacian $\mathcal{L}^{[p]}$ is defined on $\tilde{T}^{[p]}$ for edge weights defined in terms of the p -hop transition probabilities. Therefore, instead of the traditional heat equation for the curvature transport, we propose the “accelerated” heat equation:

$$\frac{d\tilde{K}_u}{dt} = \mathcal{L}^{[p]} \left(\tilde{K}_u \right).$$

The speed up afforded by the $\mathcal{L}^{[p]}$ operator is related to its spectrum. As shown by Bauer, Jost and Liu [4], the spectrum of the normalized operator can be related to the Ollivier-Ricci curvature.

The problem is that with an “efficient” triangulation (least amount of vertices & edges), the p -neighborhood graph might become trivial. One approach is to keep the same triangulation but attempt to approximate the p -neighborhood graph as follows: Let $D = \text{diag}\{|\mathcal{L}_{mm}|\} \geq 0$ be the diagonal matrix of the absolute values of the diagonal entries of the Laplacian. Define the normalized graph Laplacian $\mathcal{L}_{\text{norm}} = D^{-1/2} \mathcal{L} D^{-1/2}$. The pseudo- p -neighborhood graph Laplacian is defined as

$$\mathcal{L}^{(p)} = D^{p/2} (-I + (I + \mathcal{L}_{\text{norm}})^p) D^{p/2} \leq 0$$

Lemma 6: $\mathcal{L}^{(p)} \leq 0$, the largest eigenvalue of $\mathcal{L}^{(p)}$ vanishes, and $(\mathcal{L}^{(p)})_{m \neq n} \geq 0$ (but the same Laplacian falls short of achieving $\sum_{m \neq n} (\mathcal{L}^{(p)})_{mn} = (\mathcal{L}^{(p)})_{nn}$.)

Proof: Since, as shown in [4], $-2 \leq \lambda_i(\mathcal{L}_{\text{norm}}) \leq 0$, it is easily seen by arguing on the eigenvalues that $-I + (I + \mathcal{L}_{\text{norm}})^p \leq 0$, and the rest follows from the Sylvester law of inertia. ■

Despite the above shortcoming, the pseudo-accelerated heat equation

$$\frac{d\tilde{K}_u}{dt} = \mathcal{L}^{(p)} \left(\tilde{K}_u \right) \quad (5)$$

allows for a substantial speed up of the procedure. More precisely, at the very beginning of the algorithm, the rate of change of the curvature at the extrema is much stronger, as we show in Sec. VI-B.

VI. CODE AND NUMERICAL EXAMPLES

The overall algorithm is summarized in Algorithm 1.

Algorithm 1 Compute $\arg \max_i f(v_i)$

procedure CURVTRANS($f, T(\mathcal{V}, \mathcal{E}, \Sigma), w$)

- ▷ w -weighted triangulation $T(\mathcal{V}, \mathcal{E}, \Sigma)$ of surface S
- ▷ triangulation \tilde{T} of graph \tilde{S} of f
- ▷ threshold $h_K < 0$ for curvature

while $\exists i \ni dK_i/dt < h_K$ **do**

Run Eqs. (3)-(4)

if $dK_{i^*}/dt < h_K$ **then**

return $i^* = \arg \max_i f(v_i)$

end if

end while

end procedure

A. Example: icosahedron

To illustrate the utilization of the Yamabe flow in the hunt for maxima/minima of functions defined over triangulations of surfaces, we consider a regular icosahedron (one of the Platonic solids) together with a function f defined over its vertices. The icosahedron has 12 vertices, 30 edges and 20 faces, that is, an Euler characteristics $\chi = F - E + V = 2$, which endows it, via the discrete Gauss-Bonnet theorem, with an average curvature $\bar{K}_{\text{av}} = \pi/3$ (in radians). The triangulation is shown in Figure 1. To fix the geometry, the edge length is set to 1, which gives a radius of the circumscribed sphere $r = \sin(2\pi/5) \approx 0.9511$.

To illustrate the idea of the optimization algorithm, we define a somewhat trivial function

$$f = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 10).$$

We plot the function f along axes orthonormal to, and pointing outside of, the circumscribed sphere at the selected vertices. (There is no need for a scaling, since the swallow tail singularities appear *inside* the convex disconnected component of the space \mathbb{R}^2 broken by the surface.) We run the Yamabe flow (3) and the resulting evolution of the conformal factor is shown in Figure 2.

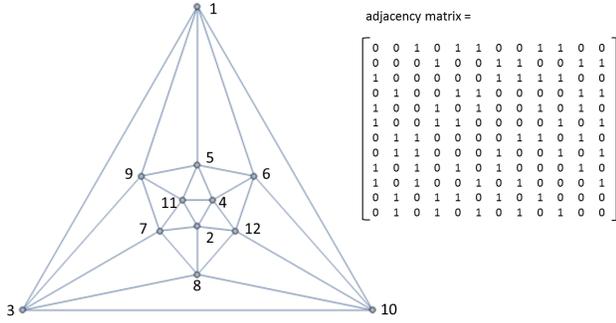


Fig. 1. Triangulation of the icosahedron along with its adjacency matrix. Note that 1-3-10 is a face of the triangulation.

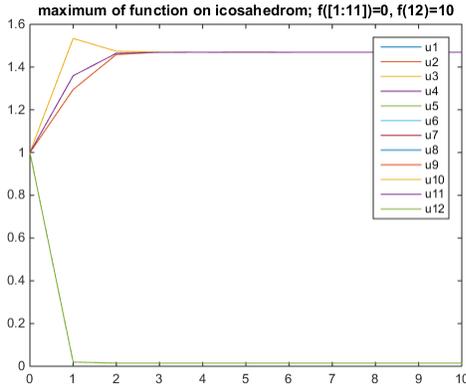


Fig. 2. Evolution of the conformal factors $u(v_i)$, $i=1, \dots, 12$, during Yamabe flow. Observe that most significant activity at the 12^{th} vertex in terms of “curvature transport.” The evolution is the fastest at that vertex; it is negative, meaning that there is a significant amount of curvature transport away from that vertex, indicating a maximum at that point.

Consider another function defined on the same icosahedron:

$$f = (-1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -0.5).$$

The results are shown in Figure 3.

Finally, we consider a less trivial function

$$f = (2, 5, 2, 5, 3, 5, 2, 5, 1, 4, 2, 10).$$

The results are shown in Figs. 4, 5, and 6.

For the same icosahedron and the same function f , the results of the pseudo-accelerated case (Eq. 5) are shown in Fig. 7. Clearly, early in the run of the algorithm, the rate of change of the gain at the extremal vertex is significantly larger than in the non-accelerated case.

B. Example: icosahedron: effect of acceleration

We evaluated the function $f(x) = -(x^2 - 8)^2 - (x - 7)^2 + 500$ in 12 points evenly spaced between -5 and 5 with local maximum at $x = 4$, $f(4) = 405.9$ and global at $x = 10$, $f(10) = 480.9$ on the same icosahedron. The significance of the pseudo-accelerated heat equation on accelerating the curvature flow and reducing the simulation running time in the same function $f(x)$ are shown in Figs. 8, 9 and 10.

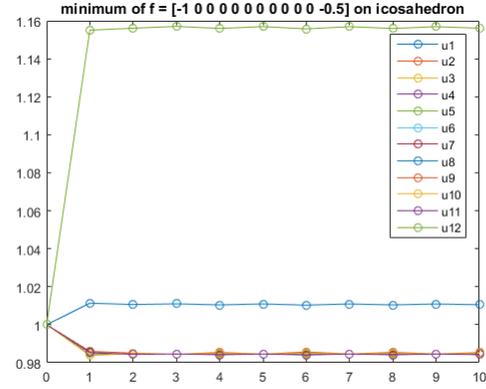


Fig. 3. Evolution of conformal factors showing significant curvature transport to the local minima.

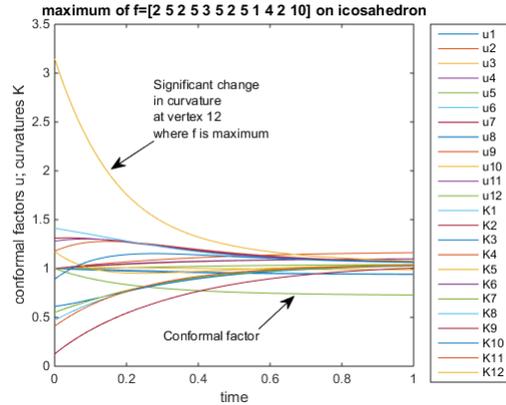


Fig. 4. Evolution of conformal factors and curvatures showing significant curvature transport away from the minimum at 12.

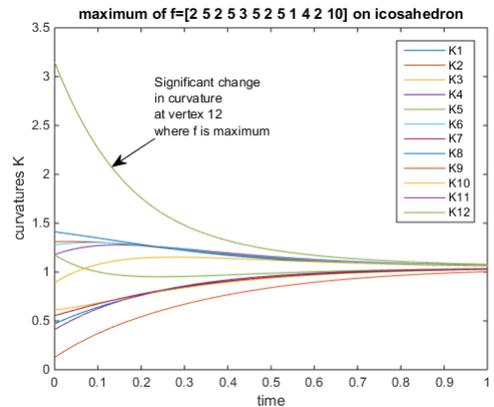


Fig. 5. Evolution of curvatures showing significant curvature transport away from the maximum at 12.

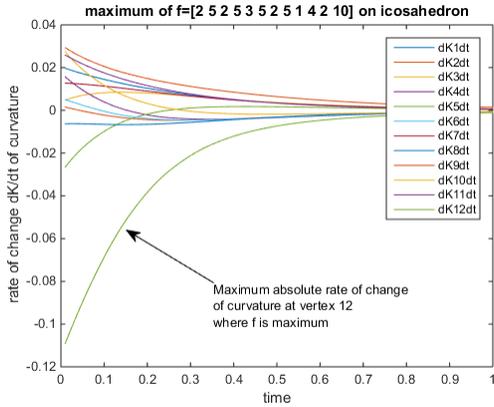


Fig. 6. Rate of change of the curvatures showing maximum absolute rate of change at vertex 12.

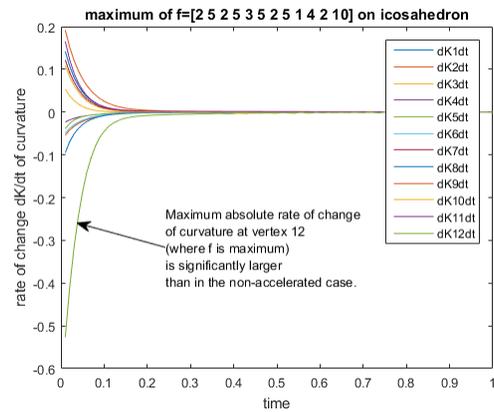


Fig. 7. Rate of change of the curvatures showing maximum absolute rate of change at vertex 12 using the pseudo-accelerated heat equation (5)

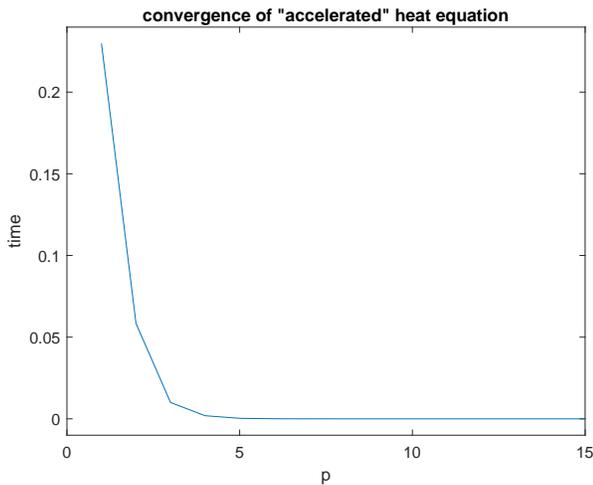


Fig. 8. Run time of the accelerated heat equation on the icosahedron of $f(x) = -(x^2 - 8)^2 - (x - 7)^2 + 500$. The convergence time of the equation is significantly decreased for $p \geq 4$.

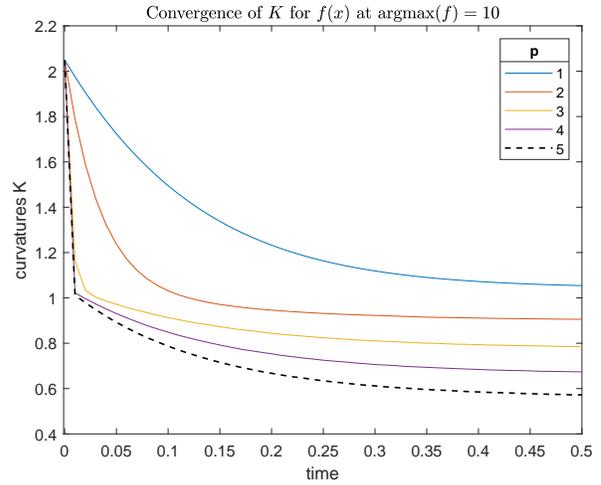


Fig. 9. Curvature evolution at vertex 10 where $f(x) = -(x^2 - 8)^2 - (x - 7)^2 + 500$ is maximum for $p = 1 \dots 5$.

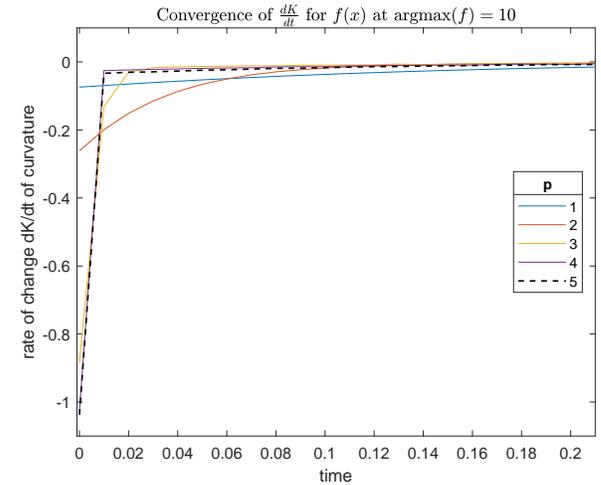


Fig. 10. Rate of change of curvature at vertex 10 where $f(x) = -(x^2 - 8)^2 - (x - 7)^2 + 500$ is maximum for $p = 1 \dots 5$. The rate of change of curvature converges faster to 0 for increasing values of p and the fastest convergence is observed for $p = 5$.

C. Example: flat torus

We now consider a surface of zero curvature: the “flat torus.” Its triangulation is shown in Figure 11.

The triangulation has 7 vertices, 14 faces, and 22 edges, leading to $\chi = F - E + V = 0$, hence vanishing curvature and $\bar{K}_{av} = 0$.

The following function is defined over the flat torus:

$$f = (1.2 \quad 0.9 \quad 1.9 \quad 1.3 \quad 2.0 \quad 2.2 \quad -5.0).$$

The behavior of the conformal factors for uniformizing the curvature of the graph of the function f defined over the two-torus is shown in Figure 12.

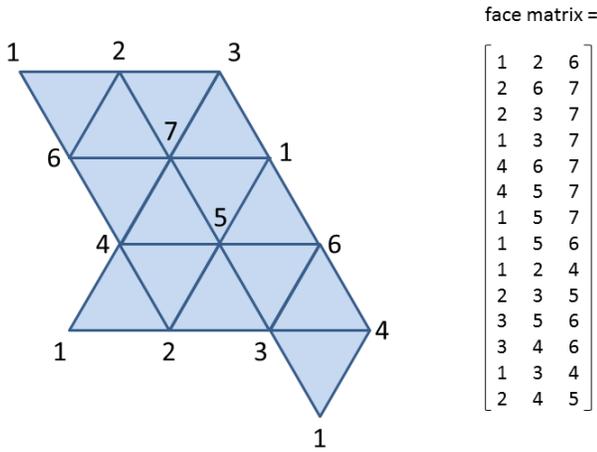


Fig. 11. “Economic” triangulation of a 2-torus, along with its “face matrix,” where faces are displayed row-wise.

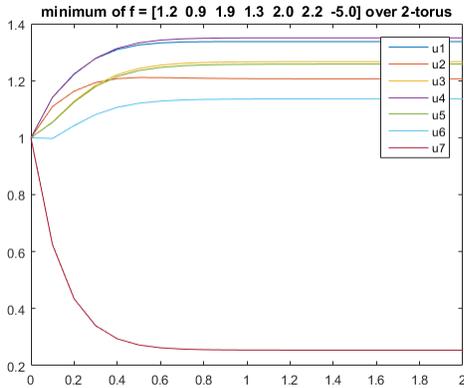


Fig. 12. Evolution of conformal weights to uniformize curvature of function f defined on 2-torus, showing maximum curvature transport at minimum point 7.

D. Connection with random walks

Let us define a random walker over the vertices of the icosahedron. At every time instant, the walker has some probability to leave its current vertex and proceed along some edge. Many probability laws can be defined and displayed in a Markov matrix. The *unique feature* that is suggested here is to define the transition probabilities proportional to the difference of curvature at the end vertices of the edge.

VII. CONCLUSION

We have introduced an optimization procedure that capitalizes on the sharp peaks & deep valleys of a challenging landscape. The fundamental observation is that the minima & maxima have a curvature signature, which the Ricci-Yamabe flow is able to detect by curvature transport—nominally on a triangulated surface—with the differential-geometric novelty of *acceleration* of the curvature flow by compounding the Laplace operator. The next challenge is to dispense of the surface structure and run the flow on the graph \mathcal{G} . A *faster* scalar curvature $K(v_i)$ could be defined as the average of

the Ollivier-Ricci or Forman-Ricci curvature along the edges abutting v_i , but whether a Yamabe flow can be defined on such restricted structure is open.

REFERENCES

- [1] A. Ambainis. Quantum search algorithms (survey). *SIGACT News*, 35:22–35, 2004. quant-ph/0504012.
- [2] V. I. Arnold. *The Theory of Singularities and its Applications*. Accademia Nazionale Dei Lincei; Scuola Normale Superiore Lezioni Fermiane. Press Syndicate of the University of Cambridge, Pisa, Italy, 1993.
- [3] R. Banirazi, E. Jonckheere, and B. Krishnamachari. Heat-Diffusion: Pareto optimal dynamic routing for time-varying wireless networks. *IEEE/ACM Transactions on Networking*, March 2020. in press.
- [4] F. Bauer, J. Jost, and S. Liu. Ollivier-Ricci curvature and the spectrum of the normalized graph Laplace operator. arXiv:1105.3803v1[math.CO] 19 May 2011, 2011.
- [5] J. M. Boardman. Singularities of differentiable maps. *Publications Mathématiques, Institut des Hautes Etudes Scientifiques (I.H.E.S.)*, (33):21–57, 1967.
- [6] J. Cerf. La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie. *Publications Mathématiques, Institut des Hautes Etudes Scientifiques (I.H.E.S.)*, (39):5–173, 1970.
- [7] Bennett Chow and Feng Luo. Combinatorial Ricci flows on surfaces. *J. Differential Geom.*, 63(1):97–129, 2003.
- [8] L. E. Dubins. On curves of minimal length with a constraint on average curvature, and with prescribed initial and terminal positions and tangents. *American Journal of Mathematics*, 79(3):497–516, Jul. 1957.
- [9] C. Ehresmann. Les prolongements d’une variété différentiable. *C. R. Acad. Sci. Paris*, 233:598–600, 1951.
- [10] M. Golubitsky and V. Guillemin. *Stable Mappings and Their Singularities*, volume 14 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1973.
- [11] R. S. Hamilton. The Ricci flow on surfaces. *Mathematics and General Relativity, Contemporary Mathematics*, 71:237262, 1988.
- [12] Alexander Hentschel and Barry C. Sanders. Machine learning for precise quantum measurement. *Phys. Rev. Lett.*, 104:063603, 2010.
- [13] M. W. Hirsch. *Differential Topology*, volume 33 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1976.
- [14] P. Hoyer, J. Neerbeck, and Y. Shi. Quantum complexities of ordered searching, sorting, and element distinctness. *Algorithmica*, 34(4):429–448, November 2002.
- [15] E. A. Jonckheere. *Algebraic and Differential Topology of Robust Stability*. Oxford, New York, 1997.
- [16] F. Langbein, S. Schirmer, and E. Jonckheere. Time optimal information transfer in spintronics networks. In *IEEE Conference on Decision and Control*, pages 6454–6459, Osaka, Japan, December 2015.
- [17] Feng Luo. Combinatorial Yamabe flow on surfaces. *Communications in Contemporary Mathematics*, 6(5):765–780, 2004.
- [18] Feng Luo. A combinatorial curvature flow for compact 3-manifolds with boundary. *Electronic research Announcement of the American Mathematical Society*, 11:12–20, 2005. Available at <http://www.citebase.org/abstract?id=oai:arXiv.org:math/0405295>.
- [19] C. C. Moreno and J. D. Vélez Caicedo. A remark on the heat equation and minimal Morse functions on tori and spheres. *Ingeniería y Ciencia*, 9(17):11–20, 2013.
- [20] S. Munoz Munoz and A. Quintero Velez. Heat equation and stable minimal Morse functions on real and complex projective spaces. *Revista Colombiana de Matemáticas*, 51:71 – 82, 06 2017. arXiv:1703.01105v1.
- [21] S. Schirmer, E. Jonckheere, and F. Langbein. Design of feedback control laws for spintronics networks. *IEEE Transactions on Automatic Control*, 63(8):2523–2536, August 2018. Available at arXiv:1607.05294.
- [22] Chi Wang, E. Jonckheere, and R. Banirazi. Wireless network capacity versus Ollivier-Ricci curvature under Heat Diffusion (HD) protocol. In *American Control Conference (ACC)*, pages 3536–3541, Portland, OR, June 04-06 2014. Available at <http://eudoxus2.usc.edu>.
- [23] Chi Wang, E. Jonckheere, and R. Banirazi. Interference constrained network performance control based on curvature control. In *2016 American Control Conference (ACC)*, pages 6036–6041, Boston, MA, July 6-8 2016.