University of Bath

The scaling and skewness of optimally transported meshes on the sphere

Chris J. Budd[a,*], Andrew T. T. McRae[a,b], Colin J. Cotter[c]

[a]*Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, UK*
[b]*Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, OX1 3PU, UK*
[c]*Department of Mathematics, Imperial College London, London, SW7 2AZ, UK*

Abstract

In the context of numerical solution of PDEs, dynamic mesh redistribution methods (r-adaptive methods) are an important procedure for increasing the resolution in regions of interest, without modifying the connectivity of the mesh. Key to the success of these methods is that the mesh should be sufficiently refined (locally) and flexible in order to resolve evolving solution features, but at the same time not introduce errors through skewness and lack of regularity. Some state-of-the-art methods are bottom-up in that they attempt to prescribe both the local cell size and the alignment to features of the solution. However, the resulting problem is overdetermined, necessitating a compromise between these conflicting requirements. An alternative approach, described in this paper, is to prescribe only the local cell size and augment this an optimal transport condition to provide global regularity. This leads to a robust and flexible algorithm for generating meshes fitted to an evolving solution, with minimal need for tuning parameters. Of particular interest for geophysical modelling are meshes constructed on the surface of the sphere. The purpose of this paper is to demonstrate that meshes generated on the sphere using this optimal transport approach have good a-priori regularity and that the meshes produced are naturally aligned to various simple features. It is further shown that the sphere's intrinsic curvature leads to more regular meshes than the plane. In addition to these general results, we provide a wide range of examples relevant to practical applications, to showcase the behaviour of optimally transported meshes on the sphere. These range from axisymmetric cases that can be solved analytically to more general examples that are tackled numerically. Evaluation of the singular values and singular vectors of the mesh transformation provides a quantitative measure of the mesh anisotropy, and this is shown to match analytic predictions.

Keywords: Mesh adaptation, mesh regularity, optimal transport

*Corresponding author
 Email address: `c.j.budd@bath.ac.uk` (Chris J. Budd)

1. Introduction

1.1. Overview

Many partial differential equations are naturally formulated on the sphere, S^2, or on a thin spherical shell. A notable example is those equations describing atmospheric or oceanic flows on the Earth, which are essential in weather forecasting and climate predictions. To find approximate solutions to these, it is common to first define a mesh on the sphere (perhaps extended in vertical columns in the case of a thin shell). The equations are then discretized with respect to this mesh, using, for example, a finite difference or finite volume method.

There are many considerations for constructing a suitable mesh. Firstly, the mesh must be a reasonable approximation of the analytic domain. It is important that the solution can be faithfully represented on the mesh; this may be non-trivial if the solution develops small-scale features that evolve and move around over time. In [1], it is highlighted that a major limitation of the development of climate models is the lack of mesh resolution when faced with climatic phenomena on many scales. The resulting need for using some form of mesh refinement to resolve important atmospheric features is emphasised. Issues can arise from interactions between the mesh and the numerical method being used, as described in [2], where the desirability of particular degree-of-freedom ratios between different fields (such as air pressure and wind speed) leads to constraints on the topology of the mesh. Furthermore, if computational efficiency is important, a structured mesh is desirable. These allow direct addressing to be used, and so generally lead to faster calculations than unstructured meshes – although the difference can be minimised in the spherical shell case since the radial direction always provides exploitable structure [3, 4].

Crucial to the success of such methods is the development of algorithms which (with minimal operator intervention) can generate a mesh rapidly. This mesh must be able to resolve the small scale features of the underlying solution, align itself to anisotropic solution features, and yet be sufficiently regular to avoid errors due to excessive mesh skewness. One method for doing this is r-adaptive mesh relocation, in which a fixed number N of mesh points, with prescribed connectivity, is moved around the sphere so that points are concentrated in regions requiring higher mesh resolution. In earlier papers [5, 6], we have demonstrated that r-adaptive mesh relocation using optimal transport regularisation can be implemented, flexibly and rapidly, both in the plane and on the surface of the sphere. These meshes have been shown to avoid tangling and to be capable of following time-evolving features with excellent resolution of small scales. The unchanging mesh topology in the r-adaptive approach also means that all data structures remain constant throughout the simulation; these data structures can be based on well-established static mesh topologies such as cubed-sphere or icosahedral meshes.

The purpose of this paper is to provide novel analytical estimates for the regularity of such meshes on the sphere by exploiting the structure of the Monge–Ampère equation. We will show that, in general, meshes on the sphere have

better regularity than those on the plane. The general theory will be illustrated by firstly looking at some specific examples relevant to practical applications, where we have exact analytic descriptions of optimally transported meshes, and then looking at some numerically-generated meshes for more challenging examples. Comparisons will be made with other methods for mesh generation on the sphere. We aim to convince the reader that the ease and flexibility of use of the optimal transport methods, combined with their good regularity properties, make them very suitable for calculating rapidly evolving PDEs defined on the surface of the sphere, such as those used in geophysical modelling.

1.2. Some existing mesh generation methods and their properties

We can broadly define three approaches to mesh generation for transient simulations. In the first approach, the mesh is (reasonably) uniform and is *static* – the mesh does not evolve with the simulation. In the second approach, the mesh is again static, but now non-uniform, with user-prescribed local resolution. These meshes are often unstructured. A typical example is an ocean simulation in which the coastline is resolved. In the third approach, the mesh is *dynamic*, changing as the solution evolves. This case further separates into mesh refinement methods and mesh relocation methods (r-adaptive methods), which we discuss in this paper.

The meteorological community have traditionally used a latitude–longitude mesh of the sphere, in which the sphere is divided into cells by lines of constant latitude and constant longitude. This mesh has the advantages of being fully structured and of having quadrilateral cells and orthogonal gridlines, which are beneficial for certain numerical schemes. However, the severe resolution clustering at the poles is problematic for numerical methods and for parallel efficiency. As a response to this problem, more uniform meshes have been considered, particularly in recent years [7, 2]. Numerous 'next-generation' models use varieties of cubed-sphere or icosahedral meshes [8] (the spectral community has a longer history of using 'reduced' grids [9]). Both cubed-sphere and icosahedral meshes can be represented as a collection of structured patches. The standard varieties of these meshes are fairly uniform, with cell areas that vary by at most a factor of two. These are, therefore, not immediately suitable for resolving evolving small-scale features in the solution. However, we can generalise to using meshes that are similar to these, but which also make use of r-adaptive mesh redistribution or a local refinement strategy to increase the resolution in particular regions. For atmospheric flows, this may be appropriate when the solutions of the governing PDEs develop small-scale features in particular locations, such as atmospheric fronts or equatorial waves.

Dynamic mesh adaptivity on the sphere is used in [10] and [11] for a shallow-water model, with impressive results. These use local mesh refinement, rather than r-adaptivity, on an underlying cubed-sphere grid. By exploiting the structure of the cubed-sphere grid – each of the six panels is locally cartesian – the resulting adaptivity problem is similar to adaptive mesh refinement on the plane, with extra care needed where the panels join. Existing mesh relocation methods on the sphere appear to be aimed at static, rather than dynamic,

3

adaptivity. An approach based on spring dynamics is used to generate meshes for the NICAM dynamical core [12], although this is motivated by producing maximally-uniform meshes – arguably the opposite of adaptivity! A prescribed mesh is projected onto the sphere. The mesh is then smoothed by considering the vertices to be connected by springs and allowing the mesh to relax to a minimal-energy configuration. The paper [13] attempts to iron out the small non-uniformities in refined icosahedral grids by varying the natural spring length that is used. The later paper [14] shows that a natural spring length of zero is a natural choice from energy considerations. However, this leads to mesh distortion around the 12 five-neighbour points in the mesh, and so a custom transformation is introduced in the neighbourhood of these points. The more recent papers [15, 16] deliberately generate grids with higher resolution near the equator. This is done by introducing meshes with more cells near the equator (see Figure 1 in both papers), rather than basic icosahedral grids. The paper [17] uses spherical centroidal Voronoi tessellation, with a spatially-varying mass distribution, to produce a static mesh with increased resolution in one part of the world. The method is not strictly r-adaptivity: the number of mesh points remains the same, but the connectivity is allowed to vary in order to reduce cell stretching.

These mesh relocation methods are not particularly well-suited to dynamic adaptivity. The NICAM grids are very pleasing, but their generation required significant human intervention and the tuning of free parameters. This is reasonable for static adaptivity, but the methods do not easily generalise to producing meshes that can resolve arbitrary time-varying features. A distinct approach to introducing mesh adaptivity using spring dynamics would be to vary the spring constants and/or the natural lengths. It is feasible that this process could be performed each timestep. However, the relationship between the spring parameters and the resulting mesh resolution is complicated – it is far from clear what spring parameters are needed in order to produce the desired resolution distribution. The centroidal Voronoi approach described in [17] relies on Lloyd's algorithm, which is extremely expensive. As a result, this approach is likely infeasible for dynamic adaptivity. Furthermore, the changing unstructured connectivity would be hard to couple to a PDE solver.

An alternative approach for constructing an r-adaptive mesh is to explicitly prescribe the local scale of the mesh, via a (solution-dependent) mesh density monitor function, while imposing global regularity in the form of an optimal transport constraint. This has been implemented and analysed on the plane [18, 19, 20, 21, 22], and in recent papers, we have extended this to produce solution-adapted meshes on the sphere [5, 6]. These methods used an optimal transport approach linked to the solution of a (version of the) Monge–Ampère equation, posed on the tangent bundle to the sphere, to produce regular-looking meshes with the desired (spatially-varying) density of mesh points. The r-adaptive approach, in which a fixed number of mesh points with constant connectivity is moved around, is well-suited to PDE computations as it allows the use of fixed and simple data structures, and requires no modification on parallel architectures. The papers [5, 6] were primarily computational. In this

4

companion paper, we look at the geometry of the resulting meshes generated using the optimal transport methods, particularly local mesh scaling, local mesh regularity, and mesh alignment. We demonstrate that these methods have the merit of being a systematic approach to mesh generation that delivers meshes of prescribed local scale and of provable regularity, through the use of robust algorithms which are relatively simple to implement. They also have significant flexibility in the control of the mesh points. Optimal transport not only gives a cheap, reliable, robust and flexible means of generating a (solution-dependent) mesh, but also has provable regularity bounds. Furthermore, we show that there are a number of exact solutions of the Monge–Ampère equation on the sphere which generate meshes that are appropriate for the solution of certain meteorological problems described in, for example, [1].

1.3. Summary of this paper

The remainder of this paper is structured as follows. In section 2, we introduce some basic theory of r-adaptive mesh relocation strategies, based on controlling the density of mesh points. In particular, we consider meshes which equidistribute a monitor function, m, of the mesh density. We then describe some measures of mesh quality for meshes obtained by relocation strategies. In section 3, we consider the construction of meshes on the sphere, and possibly more general two-dimensional Riemannian manifolds. This is achieved through the use of optimal transport maps from the sphere to itself, which act on a base mesh of the sphere. The maps are obtained by solving scalar partial differential equations of Monge–Ampère type. We then make significant use of the structure provided by the optimal transport formulation to derive some a priori estimates of mesh regularity. These results follow directly from the analytic theory of optimally transported maps. We draw the surprising conclusion that, due to the positive intrinsic curvature of the sphere, meshes on the sphere tend to have better formal regularity than analogous meshes on the plane. In section 4, we consider the specific case of monitor functions, m, which are axisymmetric about a particular axis. By solving the equidistribution equation exactly in this case, we derive analytic expressions for the related optimally transported meshes, and the associated mesh quality measures. In section 5, we then apply the optimal transport maps, with particular monitor functions, to some common meshes of the sphere, including latitude–longitude, cubed-sphere and icosahedral meshes, to construct examples of meshes of proven regularity. In section 6, we briefly compare a spherical example to a 'matching' planar example, and show that the geometry of the sphere leads to a mesh of higher quality. In section 7, we provide more general computational examples on the sphere, looking at some challenging examples. In each case, we compute the regularity of the mesh and also show that the resulting meshes are naturally aligned to the prescribed features. Some comparisons are made with similar meshes obtained using different mesh construction algorithms such as the spring dynamics method. Finally, in section 8, we draw some conclusions and consider future work and applications of these methods.

5

2. Mesh construction through mesh point relocation, and measures of mesh regularity.

2.1. Mesh construction

In this paper, we consider an r-adaptive approach to mesh construction. As described earlier, this involves a fixed number of mesh points being *relocated*, while the topology of the mesh remains unchanged. Suppose that we wish to construct a mesh τ_P for simulating a physical problem in a domain Ω_P, where Ω_P lies in a manifold M. We assume that τ_P will be specially adapted for the problem and may be highly non-uniform. In this work, we are particularly interested in the case where M is the surface of the sphere, S^2, with $\Omega_P = M$. However, we will frequently draw comparisons to the planar case $M = \mathbb{R}^2$. Previous work on this planar case can be found in [22].

We also define a 'computational' (or 'logical') domain $\Omega_C \subseteq M$, with a computational mesh τ_C of prescribed connectivity. We assume the cells of τ_C are reasonably uniform—perhaps fully uniform in the planar case—having shapes and sizes that do not vary too much. This is the case in an icosahedral mesh on the sphere refined through repeated bisection, and in a gnomonic cubed-sphere mesh. In the r-adaptive approach, we assume the existence of a bijective map $F : \Omega_C \to \Omega_P$, with τ_P the image of τ_C under the action of this map. It follows that τ_P will have the same topology (connectivity) as τ_C.

We use $\vec{\xi}$ to denote a position vector in Ω_C, and $\vec{x}$ to denote the corresponding position in Ω_P: $F(\vec{\xi}) = \vec{x}$. Let U be a small open set containing $\vec{\xi}$, and let V be the image of this set under the action of F (hence containing $\vec{x}$). We may compute the ratio of the volumes (areas) of these two sets, $|V|/|U|$. In the limit $|U| \to 0$, we define

$$r(\vec{\xi}) = \lim_{|U| \to 0} \frac{|V|}{|U|} \tag{1}$$

to be the limiting area ratio. If $M = \mathbb{R}^2$, we have

$$r(\vec{\xi}) = |\det J|, \tag{2}$$

where J is the Jacobian of the map F. If M is a general Riemannian manifold of dimension d embedded in some $\mathbb{R}^n$, r is now the product of the first d singular values of J. This coincides with the *pseudodeterminant* of J—the product of the non-zero singular values—as long as F is not degenerate.

In r-adaptive methods, controlling this area ratio is always a primary concern. An *equidistribution* principle is widely used: let $m(\vec{x})$ be a suitable *monitor function*, traditionally related to the error in representing the solution on the physical mesh. We then seek a mesh where the area ratio is inversely proportional to m:

$$m(\vec{x})r(\vec{\xi}) = \alpha, \tag{3}$$

where α is a normalisation constant that ensures the domains Ω_C and Ω_P have the correct size (alternatively, we could impose a condition on m to ensure the

correct scaling). The error in representing the solution would then be *equidistributed* between cells of the physical mesh. The monitor function m does not have to be a proxy for interpolation error; a far more general monitor function can be used (for example, [5] shows a mesh of the Earth adapted to the amount of precipitation that fell on a particular day). By eq. (3), if $m(\vec{x})$ is large in a region, the cells of the physical mesh τ_P are small there (as $r(\vec{\xi})$ is forced to be small in the preimage of this region). This is desirable if higher resolution is sought in that area.

We refer to eq. (3) as the *equidistribution condition*. It is clear that it does not, on its own, lead to a well-posed mesh generation problem (other than in one-dimension), since the resulting map is far from unique. It is necessary to augment the equidistribution condition with further conditions. Traditionally, these have been constraints on mesh regularity such as orthogonality [23] or alignment to a prescribed tensor field [24]. In the latter approach, the resulting mesh is then chosen to minimise some weighted sum of terms, each attempting to enforce a separate condition. In particular, the mesh is not designed to satisfy eq. (3) exactly.

An alternative, and powerful, technique is to use the concept of *optimal transport* [25, 26]; previous work using this approach includes [18, 19, 20, 21]. We now seek a map F satisfying eq. (3) exactly (up to discretisation error, at least) so that the resulting mesh τ_P is "as close as possible" to τ_C. This "distance" between the meshes is defined as

$$\int_{\Omega_C} \|\vec{x}(\vec{\xi}) - \vec{\xi}\|^2 \, \mathrm{d}\vec{\xi}, \tag{4}$$

the integral of the squared Riemannian distance. In optimal transport terminology, this is the cost of a candidate map $F : \Omega_C \to \Omega_P$.

It is well-known that a unique solution exists for this problem; this was established in [27] for Euclidean space and in [28] for the sphere. In Euclidean space, the appropriate map can be written in the form

$$\vec{x} = F(\vec{\xi}) = \vec{\xi} + \nabla u, \tag{5}$$

for a suitable scalar 'potential' u. The corresponding area ratio is

$$r(\vec{\xi}) = \det(I + H(u)), \tag{6}$$

where $H(u)$ denotes the *Hessian* of u. In section 3, we extend this to the sphere. Combining eq. (3) with eq. (6) then leads to a nonlinear partial differential equation for u, a *Monge–Ampère* equation. Such equations, defined over general manifolds M, have been well-studied. Many results are available on the formal regularity of the solutions, including estimates for the various derivatives of the function u in terms of m [29, 30, 31, 32, 33, 34, 35, 36]. The formal regularity properties of the resulting mesh τ_P can then be determined from these estimates.

In a transient simulation, the monitor function will vary with time: $m = m(\vec{x}, t)$. In contrast to some other r-adaptive techniques (particularly *velocity-based* moving mesh methods), our approach to mesh generation is effectively

7

quasi-static. The mapping $\vec{x}(\vec{\xi}, t)$ is uniquely defined by the monitor function $m(\vec{x}, t)$ at that moment; our meshes *do not know their history*. For this reason, the effect of time does not enter our analysis.

2.2. Geometrical measures of mesh regularity

A general mesh τ defined on a two-dimensional Riemannian manifold M is comprised of a set of nodes on M connected together by edges, defining a set of cells. In a well-behaved mesh, the nodes are regularly spaced and the edges meet at carefully-controlled angles. The resulting cells are therefore not too skew. The spherical meshes mentioned previously have these desirable properties, as does the uniform mesh on the plane. In an r-adaptive context, such meshes are appropriate for a computational mesh τ_C. Under the action of the map F, the physical mesh τ_P is typically less regular than τ_C. In particular, the equidistribution condition eq. (3) controls the size of the mesh cells; if this varies, it leads to a degree of skewness in the mesh. A general criticism of r-adaptive meshes is that they can lead to excessively skew meshes. It can, however, be shown that the use of the optimal transport regularisation condition (which forces τ_P to be "as close as possible" to τ_C) is beneficial for controlling the degree of skewness [37].

Assume that we are constructing meshes on a two-dimensional manifold. We can then define the *local scaling* and *local skewness* in terms of the linearisation of the map F.

Definition 1. Let $F : \Omega_C \to \Omega_P$ have Jacobian J, and suppose that this linear operator has leading singular values σ_1, σ_2.

The **local scaling** s is defined as

$$s = \sigma_1 \sigma_2. \tag{7}$$

The **local skewness** Q is defined as

$$Q = \frac{1}{2}\left(\frac{\sigma_1}{\sigma_2} + \frac{\sigma_2}{\sigma_1}\right). \tag{8}$$

In lemma 1, we will obtain estimates for both of these quantities for a certain class of meshes induced by axisymmetric monitor functions.

This two-dimensional local skewness measure, Q, is equivalent to the mesh quality measure Q_{geo} defined in [24, p. 205], where more analysis is given. In Section 5.1 of that book, it is shown how the interpolation error for a mesh, using different types of interpolant, can be calculated directly in terms of the scaling and skewness. Both need to be controlled to get a low overall error, but it may well be that one can be large provided that the other is small.

If F is the identity, so that the physical mesh is equal to the computational mesh, then s is constant and $Q = 1$. This is optimal for Q, but may be suboptimal for s if the underlying solution we are trying to represent on the mesh has very small scales. For an adapted mesh it is expected that both will vary. The local scaling s is controlled directly via the equidistribution condition

eq. (3). In contrast, the local skewness Q follows indirectly from properties of the Monge–Ampère equation. We will consider using the local skewness as a general measure of the quality of the mesh.

These general considerations lead to the following challenges for adaptive mesh generation:

Challenge 1 Derive a mesh which has minimal skewness Q for a given scaling distribution s.

Challenge 2 Derive a mesh which leads to minimal solution error, expressed as a combination of scale, skewness and other factors such as alignment properties [24].

Both questions are very hard to answer in general. However we will show that the use of the optimal transport regularisation gives some partial answers in terms of mesh generation, and that we can give a much more complete analysis in the case of axisymmetric monitor functions.

We remark that on the plane $\mathbb{R}^2$, the use of optimal transport techniques leads to J being symmetric, and so the expressions above in eqs. (7) and (8) can be written with eigenvalues replacing singular values. However, this is not true for a general manifold embedded in Euclidean space.

Our measure of skewness above is based only on the local linearisation of the map F. There are many other measures of mesh quality that are calculated directly from the vertices and edges of the mesh, such as those used in [5]. On the other hand, our skewness measure doesn't formally provide this geometric information; rather it is a function of the map $F : \Omega_C \to \Omega_P$. The geometric information of τ_P is only recovered when the computational mesh τ_C is specified.

For example, if τ_C is made up of small square elements, these will be mapped to small quadrilaterals, whose precise shape depends on how the original square is aligned to the (orthogonal) eigenvectors/singular vectors of F'. If we consider the ratio, r, of the lengths of the two sides of the quadrilateral to be a measure of its skewness, Q is roughly the mean of r taken over all possible alignments.

Alternative regularity measures, derived directly from the mesh, consider quantities such as non-orthogonality of certain angles, and mismatches between primal and dual grid components. Loosely speaking, these measures are based on larger scale properties of the mesh, often quantifying the extent to which neighbouring cells differ from each other. This is equivalent to analysing higher-order spatial derivatives in the expansion of F, which we do not perform in this paper.

3. Optimally transported meshes on the sphere

3.1. The definition of an optimally transported mesh.

There are various approaches for solving the Monge–Kantorovich problem of constructing the optimal transport map that minimises the appropriate cost function. Some approaches attack the mass transportation problem directly, solving the problem "from first principles". An alternative approach, which we describe here, reduces the problem to solving a partial differential equation of

9

Monge–Ampère type. This can be done with fast algorithms which are amenable to analysis. It is shown in [28] that an optimal transport map on a manifold can again be expressed in terms of the gradient of a scalar function. Let M be a general Riemannian manifold and $u(\vec{\xi}) : M \to \mathbb{R}$ be the scalar 'potential' generated by the optimal transport procedure. We then define the *McCann map* as an *exponential map* $F : M \to M$:

$$\vec{x} = F(\vec{\xi}) = e^{\nabla u}\vec{\xi}. \tag{9}$$

The quantity $\nabla u(\vec{\xi})$ lies in the cotangent space at $\vec{\xi}$, which can be trivially associated with the tangent space $T_\xi M$. The exponential map maps this onto M itself. Intuitively, one selects the geodesic that passes through $\vec{\xi}$ and coincides with ∇u there, then travels a distance $|\nabla u|$ along this geodesic. The expression eq. (9) reduces to the earlier expression eq. (5) if M is some $\mathbb{R}^n$.

From optimal transport theory, the function u automatically inherits a convexity property: it is c-convex, as defined in [28], where c denotes the cost function used for optimal transport. According to [28] (see also [36]), if the monitor function m is sufficiently smooth then so is the potential u, and so the map eq. (9) is well-defined and locally bijective. The McCann map thus associates a well-defined map $F : M \to M$ with the scalar-valued function $u(\xi)$. The map F induces a well-defined area map $r(\vec{\xi})$, which can be constructed in terms of u. For problems posed in Euclidean space, as stated in eq. (6), $r(\vec{\xi})$ is given by

$$r(\vec{\xi}) = \det(I + H(u)). \tag{10}$$

Requiring that $m(\vec{x})r(\vec{\xi})$ is constant over the mesh then leads to a variant of the celebrated Monge–Ampère equation, a fully-nonlinear second-order partial differential equation. If M is a more general manifold, the resulting expression for r in terms of u involves some Monge–Ampère-like operator, conceptually

$$r = \text{MA}(u). \tag{11}$$

Setting $m(\vec{x})r(\vec{\xi})$ to be constant then leads to an equation of *Monge–Ampère-type*. In the case of the sphere S^2, an explicit form of this equation is derived in [6] and is given in eq. (57).

Conversely, suppose we use eq. (1) to associate an area map $r(\xi)$ with a general map $F : M \to M$. [28] showed that if $r(\xi)$ is absolutely continuous with respect to the Lebesgue measure on M—which is certainly true if the monitor function m is continuous and bounded away from zero—then there exists a unique optimal transport map of the form eq. (9) with this associated area map.

Significantly, if $M = S^2$, the unit sphere, the geodesics are segments of great circles, and the resulting exponential map in eq. (9) can be calculated easily. Indeed, it can be written in closed form as

$$\vec{x} = \cos(\delta)\,\vec{\xi} + \sin(\delta)\frac{\nabla u}{|\nabla u|}, \quad \delta = |\nabla u|, \tag{12}$$

10

or, equivalently,

$$\vec{x} = \cos(\delta)\,\vec{\xi} + \frac{\sin(\delta)}{\delta}\nabla u, \quad \delta = |\nabla u|. \tag{13}$$

This is a simple case of Rodrigues' rotation formula. The form eq. (12) makes it clear that the destination $\vec{x}$ is an appropriate combination of orthogonal unit vectors, while the alternative form eq. (13) highlights the reduction to the planar expression eq. (5) in the small-δ limit. We use these closed-form expressions for all our subsequent calculations on S^2.

3.2. A-priori estmates of the local and global regularity of optimally transported meshes

In a series of papers (extending earlier work of, among others, Pogorelov, Lions, Gilbarg, Trudinger, and Urbas in $\mathbb{R}^n$), Loeper [36] and McCann [28] have derived estimates for the derivatives of the McCann map acting on S^2. These can be used to help determine the mesh regularity through eq. (8). The principal results from this work are summarised as follows.

Theorem 1. (Trudinger et al.) *Suppose that u satisfies a Monge–Ampère equation in Euclidean space of the form*

$$\det(I + H(u)) = g(\xi, \nabla u) \equiv 1/m(\xi, \nabla u). \tag{14}$$

There then exists a constant C that depends only on g, the domain Ω_C, and any boundary conditions, such that

$$\sup_{\Omega_C} |H(u)| \leq C. \tag{15}$$

This result is then extended by Loeper:

Theorem 2. (Loeper) *Suppose, analogously to eq. (14), that u is the solution of a problem of Monge–Ampère-type on the sphere, arising from an optimal transport problem as defined in eq. (11). As long as $g > 0$ (i.e., $m > 0$), then*

1. *if $g \in C^{1,1}$ then $u \in C^{3,\alpha}$, and*
2. *if $g \in C^\infty$ then $u \in C^\infty$.*

These results are significant for mesh generation: the map from the computational domain Ω_C to the physical domain Ω_P is given by the exponential map of the gradient of u. The smoothness of u therefore implies smoothness of the map. We can deduce that if $g \in C^\infty$, the map is a C^∞-function of ξ. Consequently, the singular values σ_j of the Jacobian of the map are bounded, differentiable functions of ξ over S^2. It follows further from the convexity properties that both σ_1 and σ_2 are uniformly bounded away from zero. We deduce from this that the skewness Q of the mesh is a bounded, differentiable function of ξ over S^2. Intuitively, we expect that the mesh we generate cannot become too skew.

It is shown further by Loeper that the formal regularity of the function u is slightly better on the sphere than for the plane. This is because the *cost-sectional curvature*—defined in [36] and closely related to the usual curvature—is uniformly positive on the sphere, but is zero on the plane. In fact, it is possible to get certain formal regularity results on the sphere even if the source measure vanishes. On the sphere we also avoid problems on the plane, seen in [22], where the mesh loses regularity as one approaches the boundary. In section 6, we make a direct comparison between the sphere and the plane.

3.3. A local coordinate-based approach

It is helpful to see how the exponential map can be expressed in terms of a *local* two-dimensional coordinate basis mapping from $\mathbb{R}^2$ to S^2. A natural basis to use is spherical angles (θ, ϕ) with respect to some (unit) axis $\vec{\omega}$. This maps a coordinate patch $(\theta, \phi) \in \mathbb{R}^2$ directly onto S^2. We can then consider—locally, at least—$u \equiv u(\theta, \phi)$. To avoid singularities in the coordinate mapping, we will assume that we are working in a region well-separated from the poles $\pm\vec{\omega}$.

In this local basis, we have

$$\vec{\xi} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)^T, \tag{16}$$

with local unit vectors

$$\vec{e}_\theta = (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta)^T, \quad \vec{e}_\phi = (-\sin\phi, \cos\phi, 0)^T; \tag{17}$$

these are orthogonal and also orthogonal to $\vec{\xi}$. We then have

$$\nabla u = u_\theta \vec{e}_\theta + \frac{u_\phi}{\sin\theta} \vec{e}_\phi, \tag{18}$$

where, by assumption, we are working on a patch where $\sin\theta$ is bounded away from zero. Note that, using the relation $\cos\theta = \vec{\xi}\cdot\vec{\omega}$, we can express these vectors as

$$\vec{e}_\theta = \frac{\cos(\theta)\,\vec{\xi} - \vec{\omega}}{\sin\theta}, \quad \vec{e}_\phi = \vec{\xi}\times\vec{e}_\theta = -\frac{\vec{\xi}\times\vec{\omega}}{\sin\theta}. \tag{19}$$

We can then substitute eq. (18) into eq. (13) to find the McCann map explicitly:

$$\vec{x} = \cos(\delta)\,\vec{\xi} + \frac{\sin(\delta)}{\delta}\left(u_\theta\vec{e}_\theta + \frac{u_\phi}{\sin\theta}\vec{e}_\phi\right), \quad \delta = \sqrt{u_\theta^2 + \left(\frac{u_\phi}{\sin\theta}\right)^2}. \tag{20}$$

The map eq. (20) induces a map on the local coordinate space from $(\theta, \phi) \to (\theta', \phi')$. It follows from standard geometry that the area ratio is

$$r(\vec{\xi}) = \frac{\sin\theta'}{\sin\theta}\,|K|, \quad \text{where } K = \frac{\partial(\theta', \phi')}{\partial(\theta, \phi)}, \tag{21}$$

and $|K|$ is the determinant of this. After some manipulation, it follows from eq. (20) that

$$\cos\theta' = \cos\delta\cos\theta - \frac{\sin(\delta)}{\delta}\sin(\theta)\,u_\theta, \quad \sin(\phi' - \phi) = \frac{\sin(\delta)\,u_\phi}{\delta\sin\theta'\sin\theta} \tag{22}$$

12

The scaling factor r can then be derived by differentiating eq. (22) with respect to θ and ϕ and applying eq. (21). Specifying the value of r leads to a form of the Monge–Ampère equation.

3.4. Summary

The formulation described in section 3.1 has several advantages for mesh generation. Firstly, we need only work with scalar quantities—monitor functions m and mesh potentials u—in order to compute the map $\vec{x}(\vec{\xi})$. This leads directly to flexible and robust methods for time-dependent mesh generation which are relatively easy to implement. Secondly, there is a substantial body of theory giving formal regularity estimates for solutions of the optimal transportation problem, as mentioned in section 3.2. This can be exploited to give formal regularity properties for the map and hence the mesh. Thirdly, provided the exponential map can be calculated easily (as on the sphere), it gives a systematic and straightforward way of finding a map $M \to M$ which can be used to calculate the mesh in a natural manner. This avoids the many ad-hoc approaches to mesh construction that can be found in the literature, which often involve fine-tuning at a local level.

4. Exact maps generated by axisymmetric monitor functions and their regularity

For general monitor functions, the optimal transport map cannot be expressed analytically. However, it is possible to create exact solutions for certain classes of meshes on the sphere by considering the maps arising from axisymmetric monitor functions on the sphere. The action of these maps on certain computational meshes τ_C can be studied to generate physical meshes τ_P (which need not themselves be axisymmetric). The purpose of doing this is two-fold. Firstly, the regularity of the resulting physical mesh τ_P can be deduced directly from this calculation. We can then obtain exact expressions for the scaling and skewness of the resulting meshes. The second reason for this study is that a number of the meshes so generated are appropriate to be used with PDE problems on the sphere, such as some of those described in [1]. For example, it is easy to generate smooth meshes which can resolve specific regions of the sphere, which may be appropriate for fronts and cyclones. We return to the case of calculating more general meshes in section 7.

4.1. The basic geometry of the maps

An axisymmetric function u satisfies $u(\theta, \phi) \equiv u(\theta)$, where the coordinates are defined with respect to some axis $\vec{\omega}$. We then have $u_\phi = 0$, so

$$\nabla u = u_\theta \vec{e}_\theta, \tag{23}$$

It then follows from eqs. (12) and (20) that

$$\vec{x} = \cos(\delta)\, \vec{\xi} + \frac{\sin(\delta)}{\delta} u_\theta \vec{e}_\theta, \quad \text{where } \delta = |u_\theta|. \tag{24}$$

13

It follows immediately from eq. (22) that if the local coordinates for $\vec{\xi}$ are (θ, ϕ), the new local coordinates (θ', ϕ') for $\vec{x}$ are given by

$$\theta' = \theta + \mathrm{d}u/\mathrm{d}\theta, \quad \phi' = \phi. \tag{25}$$

Thus, from eq. (21), the area scaling is given by

$$r(\theta, \phi) = \frac{\sin \theta'}{\sin \theta} \frac{\mathrm{d}\theta'}{\mathrm{d}\theta} = \frac{\sin \theta'}{\sin \theta}(1 + u_{\theta\theta}). \tag{26}$$

It is also useful to also consider the axisymmetric map in a coordinate-free form. As before, let $\vec{\omega}$ be the (unit) axis of symmetry. It follows from eq. (19) that

$$\nabla u = \frac{\mathrm{d}u}{\mathrm{d}\theta} \frac{\cos(\theta)\,\vec{\xi} - \vec{\omega}}{\sin \theta} = (\theta' - \theta) \frac{\cos(\theta)\,\vec{\xi} - \vec{\omega}}{\sin \theta}. \tag{27}$$

Combining the previous results, we have

$$\vec{x} = \cos(\theta' - \theta)\,\vec{\xi} + \sin(\theta' - \theta) \frac{\cos(\theta)\,\vec{\xi} - \vec{\omega}}{\sin \theta}, \tag{28}$$

Using eqs. (25) and (28), we can generate a map $\vec{x}(\vec{\xi})$ from the sphere to itself for any suitable $u(\theta)$ and axis of symmetry $\vec{\omega}$.

4.2. Calculation of θ' from a monitor function

Consider an axisymmetric monitor function $m(\vec{x}) \equiv m(\theta')$, where $\cos \theta' = \vec{x} \cdot \vec{\omega}$. We can use the results of the previous subsection to calculate the map that equidistributes this monitor function. Using eq. (26), the equidistribution condition eq. (3) gives

$$m(\theta') \sin \theta' \frac{\mathrm{d}\theta'}{\mathrm{d}\theta} = \alpha \sin \theta \tag{29}$$

where α is a normalisation constant. The axis of symmetry must map to itself, implying

$$\theta'(0) = 0, \quad \theta'(\pi) = \pi. \tag{30}$$

Integrating eq. (29), we have

$$F(\theta') \equiv \int_0^{\theta'} m(t) \sin t \, \mathrm{d}t = \alpha(1 - \cos \theta), \tag{31}$$

where the normalisation constant α satisfies

$$2\alpha = \int_0^{\pi} m(t) \sin t \, \mathrm{d}t. \tag{32}$$

For particular monitor functions $m(\theta')$, we can use eqs. (29) to (32) to calculate θ' directly from θ by inverting F. Applying eq. (28) then lets us calculate $\vec{x}(\vec{\xi})$ directly. We can apply this transformation to the vertices of some reasonably uniform mesh τ_C to generate new meshes adapted to the given monitor function.

14

4.3. The local regularity of the map

Two important mesh properties are its scale and regularity. Here, scale refers to the sizes of its cells, per eq. (7), and regularity to the skewness Q of cells, per eq. (8). Both of these are local measures of mesh quality, and relate to the resulting errors which can be expected when solving PDEs on the mesh. For these axisymmetric maps, we can obtain analytic expressions for the scale and skewness quantities in terms of the monitor function.

Consider a local quadrilateral coordinate patch in $\mathbb{R}^2$ centred on (θ, ϕ) and of sides $\delta\theta \times \delta\phi$. This corresponds to a patch in S^2 of sides $\delta x \times \delta y = \sin\theta \; \delta\phi \times \delta\theta$. The patch in $\mathbb{R}^2$ is mapped to a patch centred on (θ', ϕ') of sides $\delta\theta' \times \delta\phi'$, which corresponds to a patch on S^2 of sides $\delta x' \times \delta y' = \sin\theta' \; \delta\phi' \times \delta\theta'$. In the limit of the patch going to zero,

$$\frac{\mathrm{d}x'}{\mathrm{d}x} = \frac{\sin\theta'}{\sin\theta}\frac{\mathrm{d}\phi'}{\mathrm{d}\phi}, \quad \frac{\mathrm{d}y'}{\mathrm{d}y} = \frac{\mathrm{d}\theta'}{\mathrm{d}\theta}. \tag{33}$$

For an axisymmetric equidistribution map generated by the monitor function $m(\theta')$, eq. (29) implies

$$\frac{\mathrm{d}\phi'}{\mathrm{d}\phi} = 1, \quad \frac{\mathrm{d}\theta'}{\mathrm{d}\theta} = \frac{\alpha}{m(\theta')}\frac{\sin\theta}{\sin\theta'}. \tag{34}$$

Combining these, we have

$$\frac{\mathrm{d}x'}{\mathrm{d}x} = \frac{\sin\theta'}{\sin\theta}, \quad \frac{\mathrm{d}y'}{\mathrm{d}y} = \frac{\alpha}{m(\theta')}\frac{\sin\theta}{\sin\theta'}. \tag{35}$$

It follows that the Jacobian matrix, J, of the map from the local patches of S^2 is diagonal and has eigenvalues or singular values

$$\sigma_1 = \frac{\sin\theta'}{\sin\theta}, \quad \sigma_2 = \frac{\alpha}{m(\theta')}\frac{\sin\theta}{\sin\theta'}. \tag{36}$$

From this, we can deduce expressions for the scaling and skewness of the resulting map.

Lemma 1. (i) *The local scaling of the map is given by*

$$s = \sigma_1\sigma_2 = \alpha/m. \tag{37}$$

(ii) *The local skewness of the map is given by*

$$Q = \frac{1}{2}\left(\frac{\alpha}{m(\theta')}\frac{\sin^2\theta}{\sin^2\theta'} + \frac{m(\theta')}{\alpha}\frac{\sin^2\theta'}{\sin^2\theta}\right). \tag{38}$$

PROOF. Result *(i)* follows directly from eq. (36), and is consistent with the equidistribution condition eq. (3). Similarly, *(ii)* follows from eq. (36) and eq. (8).

We now show that $Q \to 1$ at the poles, as long as m is continuous. This implies that Q is close to unity for open regions around each pole. Therefore, the resulting mesh is extremely regular in these regions even if there is significant mesh contraction.

Lemma 2. *If m is continuous at the poles, $Q \to 1$ as $\theta' \to 0$ or $\theta' \to \pi$.*

PROOF. It follows from the regularity of the map that $d\theta'/d\theta$ exists. Furthermore, since $\theta' = 0$ when $\theta = 0$, it follows that $\sin \theta' / \sin \theta \to d\theta'/d\theta$ as $\theta \to 0$. Now, as in eq. (29), the equidistribution condition leads to

$$m \sin \theta' \frac{d\theta'}{d\theta} = \alpha \sin \theta. \tag{39}$$

It follows that, as $\theta \to 0$, we have

$$\frac{m}{\alpha} \frac{\sin^2 \theta'}{\sin^2 \theta} \to \frac{m}{\alpha} \frac{\sin \theta'}{\sin \theta} \frac{d\theta'}{d\theta} = 1. \tag{40}$$

Thus $Q \to 1$. A similar calculation can be performed for $\theta \to \pi$.

We immediately see, stemming from the fact that S^2 is compact and the result of lemma 2, that controlling the variation of Q is easier than on the plane. The challenge of designing a mesh adaptation strategy in the axisymmetric case can thus (at least locally) be summarised as an ODE-constrained optimisation problem.

Challenge 3 Determine a suitable monitor function $m(\theta')$, satisfying eq. (29), so that with scaling s and mesh skewness Q given by eq. (37) and eq. (38), we can obtain bounds for the solution error.

A general answer to Challenge 3 is difficult and we leave it as a subject for future research.

5. Examples of meshes generated by using axisymmetric monitor functions

5.1. Overview of the meshes generated

In this section, we consider two examples of monitor functions which are axisymmetric about an axis ω, and the maps, and hence meshes, that these induce. These monitor functions are chosen to give meshes which can be analysed and are of potential practical importance for meteorological applications. The examples will be (i) meshes which concentrate points in localised regions, and (ii) meshes which concentrate points in rings. (We note that very similar methods, with a monitor function such as $m(\theta') = \gamma e^{-\gamma(1-\cos \theta')} + 1/2$, where γ is assumed to be large, can be used to concentrate mesh points close to a single point on S^2.)

For any axisymmetric monitor function $m(\theta')$, the formula eq. (31) can be used to evaluate the map $\theta'(\theta)$. The desired meshes τ_P are then produced

by applying the resulting axisymmetric maps to various computational meshes τ_C, such as cubed-sphere and icosahedral meshes, with known topology and connectivity. The definition of the skewness, as in eq. (8), is a property of the *map*, not of the meshes themselves. However, as long as τ_C is reasonably uniform and has regular angles, such as the examples considered above, large skewness values will coincide with highly-skew cells in the adapted mesh τ_P, and low skewness with regular cells. We can hence estimate the regularity of the resulting mesh.

5.2. Meshes concentrating points into regions

5.2.1. Analytical construction of the meshes

We firstly consider monitor functions which induce meshes that concentrate more points uniformly into specified regions, such as a disc centred on the axis $\vec{\omega}$. In a meteorological context, this mesh could be used to represent a localised feature such as a moving hurricane or a vortex patch, or a static feature such as a country [17]. Ideally the algorithm for constructing such a mesh will resolve the region in fine detail without making the mesh in other regions too coarse or irregular, or introducing too much skewness in the transition between the two regions. Using optimal transport methods, we can produce and analyse such meshes. To achieve this, we consider both a discontinuous "top-hat" monitor function, for which we can express the map analytically, and a smoothed version of this.

The *top-hat* monitor function is given by the expression

$$m(\theta') = \begin{cases} \rho_1, & \theta' < \Theta' \\ \rho_2, & \theta' > \Theta', \end{cases} \tag{41}$$

where Θ' marks the boundary between high- and low-resolution regions. The ratio $\gamma = \rho_2/\rho_1$ sets the ratio of mesh density between the two regions. We will assume that $\gamma < 1$, indeed for meshes with high compression, we expect that γ will be small. The top-hat function is discontinuous, so formally we have less regularity than discussed in section 3.2, but the resulting optimal transport map is still continuous. Integrating eq. (29) from the 'boundaries' 0 and π, we have

$$\rho_1(1 - \cos\theta') = \alpha(1 - \cos\theta), \quad \theta' < \Theta', \tag{42}$$
$$\rho_2(1 + \cos\theta') = \alpha(1 + \cos\theta), \quad \theta' > \Theta'. \tag{43}$$

It follows immediately from eq. (32) that the normalisation constant α satisfies

$$2\alpha = \rho_1(1 - \cos\Theta') + \rho_2(1 + \cos\Theta'). \tag{44}$$

Finally, define Θ to be the preimage of Θ' under the map. By invoking continuity of the mesh, it follows that

$$\frac{\rho_1(1 - \cos\Theta')}{1 - \cos\Theta} = \frac{\rho_2(1 + \cos\Theta')}{1 + \cos\Theta}. \tag{45}$$
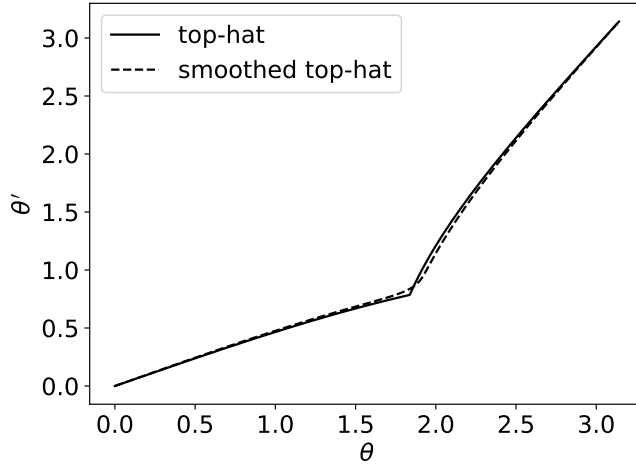
17

Figure 1: The map $\theta'(\theta)$ produced by the top-hat monitor function eq. (41), with $\rho_1/\rho_2 = 10$ and $\Theta' = \pi/4$, and by a smoothed approximation eq. (48), with $\epsilon = \pi/50$. The top-hat monitor function has a discontinuity in m, which leads to a visible discontinuity in $\frac{d\theta'}{d\theta}$, per eq. (29). In the smoothed top-hat, the transition occurs over a distance $\mathcal{O}(\epsilon)$.

Using standard trigonometrical identities, this can be written

$$\rho_1 \tan^2(\Theta'/2) = \rho_2 \tan^2(\Theta/2). \tag{46}$$

Given a monitor function of the form eq. (41), we can find Θ from Θ' via eq. (46). We then apply eqs. (42) and (43) to find θ' as a function of θ in the separate ranges $\theta < \Theta$ and $\theta > \Theta$. Having determined θ', we can calculate the image point $\vec{x}$ for a given $\vec{\xi}$ by using eq. (28). For example, if $\rho_1 = 10, \rho_2 = 1$ and $\Theta' = \pi/4$ then

$$\alpha = 2.318\ldots, \quad \Theta = 1.837\ldots \tag{47}$$

The resulting map $\theta \to \theta'$ is given in fig. 1.

A smoothed form of the top-hat monitor function was proposed by [17] in the context of a global weather forecasting model with increased resolution over the United States. In that paper, a mesh generation algorithm using Lloyd's algorithm was considered. The same monitor function was also used as a test problem in the recent paper [5] [1]. This smoother monitor function takes the form

$$m(\theta') = \sqrt{\frac{1-\gamma^2}{2}\left(\tanh\left(\frac{\Theta'-\theta'}{\epsilon}\right) + 1\right) + \gamma^2}, \tag{48}$$

---

[1] The expression in the original paper is incorrect; a correct version is given in our recent paper [6], and is used here.

where we assume ϵ and γ are small. For the region $\theta' < \Theta'$ we have $m(\theta') \approx 1 = \rho_1$, while for $\theta' > \Theta'$, $m(\theta') \approx \gamma = \rho_2$. This monitor function therefore has a similar profile to the top-hat monitor function eq. (41), but with a smooth transition over a lengthscale ϵ. There is no closed-form integral of $m(\theta')\sin(\theta')$, so we cannot use eq. (31) to get a closed-form expression for $\theta'(\theta)$. However, we can use numerical quadrature to obtain an arbitrarily good approximation. We expect that it will have similar behaviour to the top-hat monitor function if the parameters are chosen carefully. For comparison, we present a calculation for the smoothed monitor function, taking $\gamma = 1/10$, $\Theta' = \pi/4$, and $\epsilon = \pi/50$, and the resulting map $\theta \to \theta'$ is given in fig. 1.

We observe that the general piecewise-constant monitor function

$$m(\theta') = \rho_i, \quad \theta_i < \theta' < \theta_{i+1}, \quad i = 1, \ldots, N, \tag{49}$$

or smoother versions of this, can be used to concentrate points in annular regions, such as close to the equator or in the tropical zones. (Examples of meshes which concentrate points in equatorial regions are given in [16]). The calculations (and indeed the resulting mesh regularity) for the monitor function eq. (49) are very similar to those for the top-hat monitor function eq. (41).

5.2.2. Analytical estimates of the regularity of the regional meshes

We now study the regularity of the resulting maps by calculating the skewness function Q. We first consider the map induced by the top-hat monitor function eq. (41). It follows from eq. (38) that the local skewness of the map is given by

$$Q = \frac{1}{2}\left(\frac{\alpha}{\rho_i}\frac{\sin^2\theta}{\sin^2\theta'} + \frac{\rho_i}{\alpha}\frac{\sin^2\theta'}{\sin^2\theta}\right) \tag{50}$$

in each region $i = 1, 2$. In fig. 2, we plot Q as a function of θ' for the case studied previously, where we have large mesh compression with $\rho_1/\rho_2 = 10$, and $\Theta' = \pi/4$. We see that Q takes its largest value just outside the "top-hat region", *i.e.*, at $\theta' = \Theta'_+$. We therefore expect to see the most significant mesh distortion in this region, as was also observed in [5]. The value of $Q_{\max} \approx 2.273$, implies the resulting mesh will have some *moderately* skew cells. However, this is in the context of significant mesh compression and thus greatly enhanced resolution of the underlying solution in the inner region.

When we consider the smoother top-hat monitor function eq. (48), with the same parameters as before, the resulting skewness factor is smaller and is plotted in fig. 2. The skewness Q now takes its maximum value $Q_{max} \approx 1.6$ in the outer part of the transition region. We observe that $Q = 1$ in the middle of the transition region, implying that the mesh is very regular there. This can be explained by continuity: approaching the transition from inside, the cells are stretched in one direction (zonally), but are stretched in the other direction (meridionally) when approaching from outside. By continuity, there must be some intermediate value of θ' where the cells are stretched equally in both directions, so that $Q = 1$.
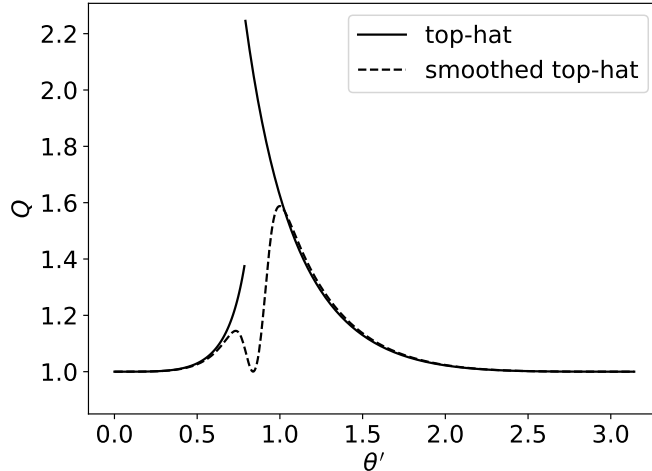
Figure 2: The skewness functions for the maps generated by the top-hat and smoothed top-hat monitor functions. For the top-hat monitor function, Q is maximised when approaching the transition from the outside; it can be shown analytically that the resulting $Q_{\max} \approx 2.273$. For the smoothed approximation, $Q_{\max}$ is notably smaller, despite the maps in fig. 1 being very similar. Perhaps surprisingly, Q returns to 1 in the interior of the transition region; this can be attributed to continuity. For both monitor functions, as expected, $Q \to 1$ at the poles.

5.2.3. The resulting regional meshes

In this and in the next subsection, the adapted meshes are generated as follows. For some vertex of τ_C, located at $\vec{\xi} \in S^2$, we calculate the corresponding value of θ from the expression $\cos\theta = \vec{\xi} \cdot \vec{\omega}$. Applying the axisymmetric map gives the value of θ' for the image point $\vec{x}$. We can then use the expression eq. (28) to determine $\vec{x}$ explicitly – it is on the same great circle as $\vec{\omega}$ and $\vec{\xi}$, and is at an angle θ' from $\vec{\omega}$ (on the same "side" as $\vec{\xi}$). This is performed for each vertex of τ_C; the image vertices form the adapted mesh τ_P, which has the same connectivity as τ_C.

We use this method to look at the action of the resulting maps on three types of meshes: a cubed-sphere mesh, an icosahedral mesh, and a latitude–longitude mesh. We also calculate the resulting skewness for the top-hat and smoothed top-hat monitor functions. For ease of visualisation, the axis of rotational symmetry, ω, is always taken to be proportional to $(0.7, -1.0, 2.0)^T$, and the meshes are viewed from the negative-y direction. This allows the mesh behaviour around the pole ω, and in any "inner region", to be seen clearly. The mesh behaviour around the opposite pole, $-\omega$, cannot be seen directly, but (in all cases that we use) it is visible that the mesh is becoming increasingly regular towards the pole. This is consistent with lemma 2. The computational meshes each have some rotational symmetry about the z-axis $(0.0, 0.0, 1.0)^T$. However, they have no symmetry about ω, the symmetry axis of the monitor function. This is done deliberately, in order to illustrate the more general behaviour.
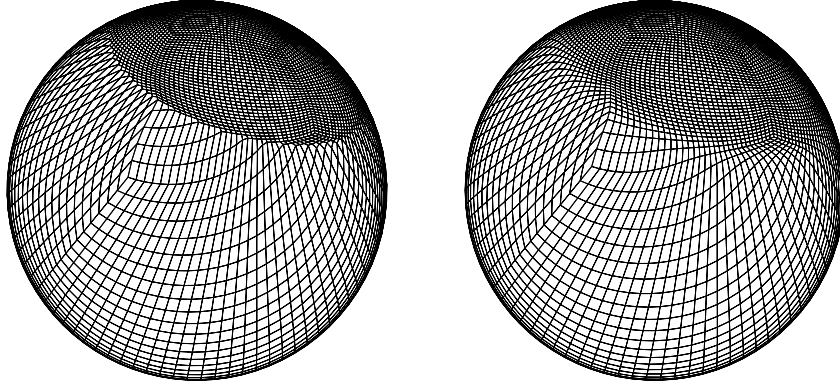
20

Figure 3: A cubed-sphere mesh, adapted to the top-hat and smoothed top-hat monitor functions.

Figure 3 shows the effect of the resulting maps on a cubed-sphere mesh, made up of six patches of 'squares', while fig. 4 shows the same for an icosahedral mesh, formed of triangles. In each case, the high concentration of mesh points, and resulting mesh compression, in the inner region is clear. We also see, as predicted from lemma 2, that there is good regularity of the meshes near both poles, which follows from Q approaching unity there. The meshes produced by the top-hat monitor function have a visible sharp transition at $\theta' = \Theta'$; however this transition is rather smoother for the tanh-based monitor function. As predicted, there is a narrow band of 'regular' cells in the transition region, where $Q \approx 1$, and the skewness of the mesh is modest elsewhere. In general, these meshes are suitable for a computation; if desired, the transition could be further smoothed by increasing ϵ. We also consider the map applied to a latitude–longitude mesh, which already has a large variation in cell-size. The resulting mesh is shown in fig. 5. This no longer has orthogonal grid lines, one of the main advantages of the latitude–longitude mesh, while the disadvantage of large variations in cell-size is still present. This mesh is unlikely to be useful for numerical weather prediction calculations, and the cubed-sphere or icosahedral meshes are probably far more suitable.

5.3. Meshes concentrating points in rings

5.3.1. Analytical construction of the meshes

Ring-like structures, or, more generally, solutions of PDEs with features strongly aligned in a certain direction, arise naturally in many applications. These include certain types of laser-driven optical phenomena [38], cross sections of bubbles, and also in low-amplitude Rossby waves of the form seen in [1]. A ring can also be regarded as an extended form of a locally anisotropic and strongly aligned feature such as a weather front. Understanding the skewness
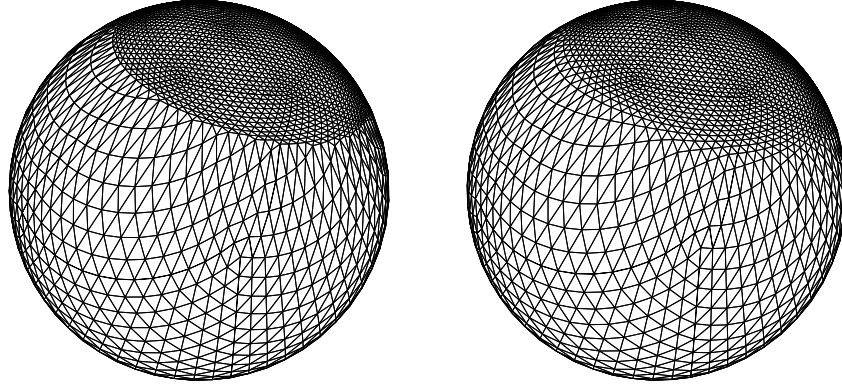
Figure 4: An icosahedral mesh, adapted to the top-hat and smoothed top-hat monitor functions.
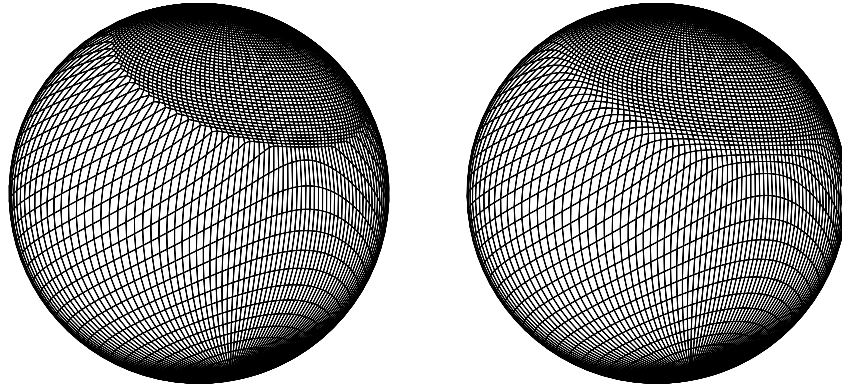

Figure 5: A latitude–longitude mesh, adapted to the top-hat and smoothed top-hat monitor functions.

and alignment properties of an analytically-generated mesh for a ring gives us insight into the corresponding features of a numerically generated mesh used to follow such a front, see for example [39]. It is known that, in such cases, it is vital for representing the solution accurately that the mesh is aligned with the anisotropic feature, even if this comes at the expense of a certain amount of mesh skewness [24].

We can consider such a feature to be one in which information is concentrated in a neighbourhood of width $\mathcal{O}(\epsilon) \ll 1$ of a circle at an angle Θ' from the axis $\vec{\omega}$. A monitor function leading to meshes which concentrate points in such a

22

ring is given by

$$m(\theta') = 1 + \frac{\beta}{\epsilon} \operatorname{sech}^2 \left(\frac{\theta'^2 - \Theta'^2}{\epsilon} \right). \tag{51}$$

The parameter β in this expression controls the density of mesh points in the ring. In the limit of $\epsilon \to 0$, m approximates an expression involving a delta function given by

$$m(\theta') = 1 + \lambda \, \delta(\theta' - \Theta'). \tag{52}$$

Here, λ controls the density of mesh points in the ring, and comparison with eq. (51), $\lambda = \beta/\Theta'$ to leading order. The analysis of the monitor function eq. (52) is more straightforward than that of eq. (51) and gives a leading-order expression for the mesh generated by eq. (51), although the resulting mesh transformation for eq. (52) (or, rather, its inverse) is discontinuous. By eq. (32), we have

$$\alpha = 1 + \frac{\lambda}{2} \sin \Theta'. \tag{53}$$

Integrating from 0 and from π, we then have

$$1 - \cos \theta' = \alpha(1 - \cos \theta), \quad \theta' < \Theta', \tag{54}$$

$$1 + \cos \theta' = \alpha(1 + \cos \theta), \quad \theta' > \Theta'. \tag{55}$$

The monitor function eq. (52) leads to a jump in the value of the computational coordinate θ when the physical coordinate satisfies $\theta' = \Theta'$. If θ_1 and θ_2 are mapped to the inner and outer edge of the ring, so that $\theta_1 \to \Theta'_-$ and $\theta_2 \to \Theta'_+$, then eqs. (54) and (55) give

$$\cos \theta_1 = 1 - \frac{1 - \cos(\Theta')}{\alpha}, \quad \cos \theta_2 = \frac{1 + \cos(\Theta')}{\alpha} - 1 \tag{56}$$

For example, if we take $\Theta' = \pi/4$ and $\gamma = 5$, then $\theta_1 = 0.464\ldots$, $\theta_2 = 1.964\ldots$, $\alpha = 2.768\ldots$, and the resulting map is presented in fig. 6.

Returning to the smoother monitor function eq. (51), we take $\Theta' = \pi/4$, $\beta = 5\pi/4$ – compatible with the value of λ used in the delta function example – and $\epsilon = \pi/50$. The relevant integrals are evaluated with numerical quadrature, and the resulting map is shown in fig. 6.

5.3.2. Analytical estimates of the regularity of the ring meshes

We firstly consider the smoother monitor function eq. (51) with the same parameters as before. The skewness factor Q is plotted in fig. 7. Again, Q is continuous, taking the value of 1 at the poles. Unsurprisingly it takes its maximum value inside the ring region. This value, close to 6.4 (representing anisotropic stretching by a factor of between 12 and 13), is fairly moderate, given that the monitor function varies by a factor of 63.5. We note though, that due to continuity, Q briefly touches 1 close to the ring. The resulting mesh cells are, as we will see later, well-aligned with the ring itself, and such a mesh is well-suited for representing a function that is aligned with the ring [24, 40].
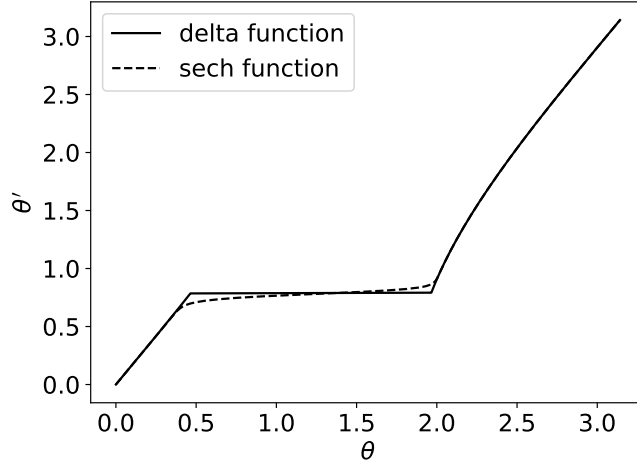
Figure 6: The map $\theta'(\theta)$ produced by a ring monitor function based on a delta function eq. (52), with $\lambda = 5$ and $\Theta' = \pi/4$, and one based on a smoother sech function eq. (51), with $\epsilon = \pi/50$. For the delta-function-based monitor function, all $\theta \in [\theta_1, \theta_2]$ are mapped to Θ'. For the smoother approximation, a similar transition occurs, but now over a distance (in θ') of $\mathcal{O}(\epsilon)$.
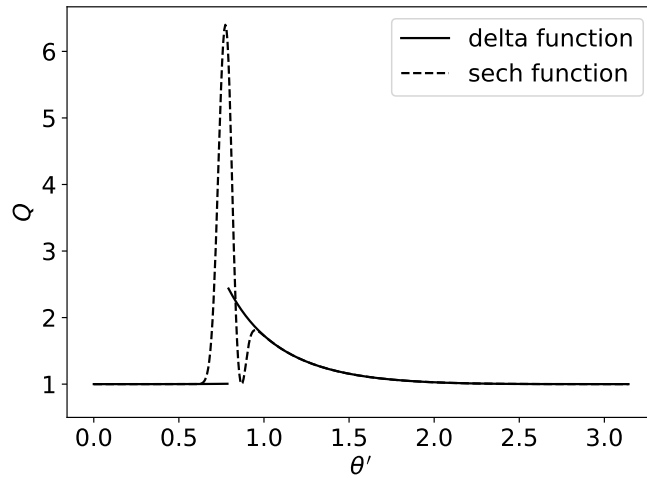

Figure 7: The skewness functions for the maps generated by the delta-function-based and sech-function-based monitor functions for ring-like features. For the delta-function example, we have not tried to represent the infinite skewness in the ring itself. The maximum skewness outside the ring is approximately 2.467. Intuitively, the ring 'swallows' cells due to the delta function. The remaining cells must then be stretched out to cover the entire area, leading to skewness. For the smooth example, the maximum skewness in the ring region is around 6.4. There is a secondary peak in Q outside the ring, as in the delta-function case, and Q briefly touches 1 due to continuity.
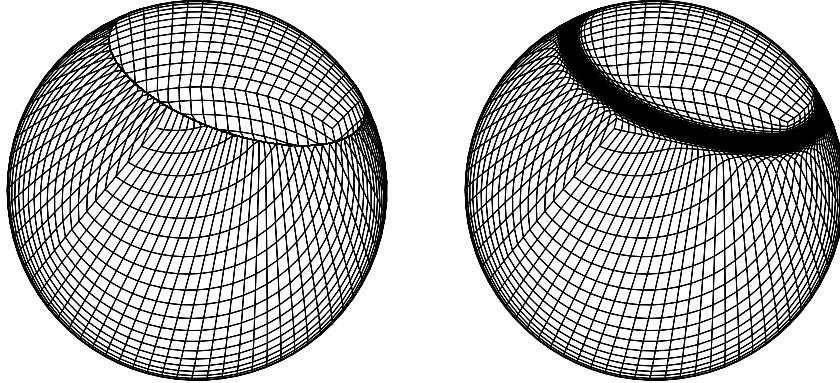
Figure 8: A cubed-sphere mesh, adapted to the delta-function-based and sech-function-based ring monitor functions.

This is important: when calculating such anisotropic functions, the alignment of the mesh with the features of the function can lead to interpolation error estimates which are much lower than those arising from a uniform mesh [24]. In a sense, the good features of the alignment outweigh the bad features of the consequent mesh skewness.

We also consider the delta function eq. (52), with Q plotted in fig. 7. This function is theoretically infinite in the ring, as cells have length zero in the meridional direction. However, the maximum skewness outside of this is just 2.467..., obtained by substituting $m(\theta') = 1$, $\theta' = \Theta'$, $\theta = \theta_2$, and α in eq. (38). The values of θ_2 and α are the same as those used previously.

5.3.3. The resulting ring meshes

We next show ring meshes induced by these monitor functions. Cubed-sphere meshes are shown in fig. 8, while icosahedral meshes are shown in fig. 9. While the meshes adapted to delta functions are obviously unsuitable for calculations, outside the singular ring they are remarkably similar to those adapted to the corresponding smooth sech function. The smoother meshes are, as predicted, quite skew in the ring region. However, they are well-aligned with the ring itself, this will lead to lower interpolation estimates of the underlying solution [24].

5.4. Direct numerical calculations of the skewness of the resulting meshes

Later, in section 7, we will apply the general numerical methods developed in [6] to generate examples of meshes produced by non-axisymmetric monitor functions. We firstly apply this approach to the axisymmetric examples above so that we can compare the output of the numerical method to the analytic results of mesh regularity we have obtained so far.
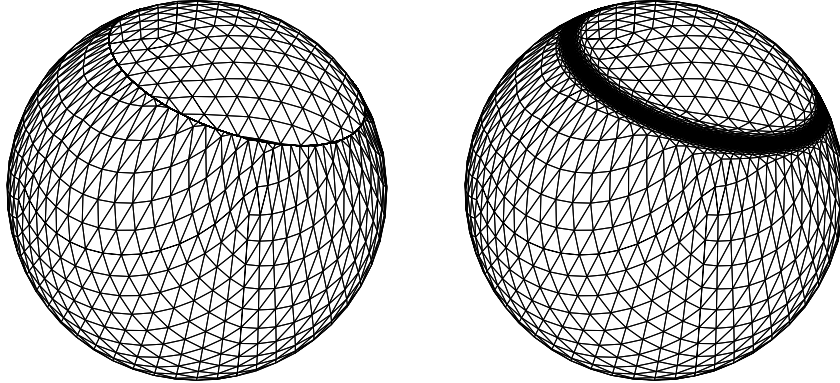
25

Figure 9: An icosahedral mesh, adapted to the delta-function-based and sech-function-based ring monitor functions.

Given an analytically prescribed monitor function, the numerical method in [6] solves a Monge–Ampère equation of the form

$$m(\vec{x}) \det \left((\nabla \exp(\nabla u)\vec{\xi}) \cdot P_\xi + \exp(\nabla u)\vec{\xi} \otimes \vec{\xi} \right) = \alpha \qquad (57)$$

for the scalar mesh potential $u(\vec{\xi})$, where P_ξ is a projection matrix and α a normalisation constant. The coordinate mapping $\vec{x}(\vec{\xi})$ can then be derived from u via the exponential map eq. (9), and lives in the finite element space $(P_2)^3$. The derivation of eq. (57) is given in [6]; briefly, the determinant term is a way to express the area ratio $r(\vec{\xi})$ in eq. (3), under the assumption that all calculations are done with the sphere immersed in $\mathbb{R}^3$.

Given the coordinate map $\vec{x}(\vec{\xi})$ obtained by this method, we can do post-processing to obtain an approximation to the skewness quantity Q. We define the raw Jacobian J as the L^2-projection of $\nabla_\xi \vec{x}$ into the finite element space $(P_2)^{3\times3}$. Analytically, the column space of J would have no component normal to the sphere, but this is not true in the presence of discretisation errors. We therefore form $J' = (I - \vec{x}\vec{x}) \cdot J$ to eliminate this component completely. At each node, we then perform an SVD: $J' = U\Sigma V$. The matrix Σ is a diagonal matrix containing the singular values of the map σ_1, σ_2, and a third entry $\sigma_3 \approx 0$. The first column of U, $\vec{u}_1$, is a vector in the direction of maximum local stretching, while $\vec{u}_2$ is at right angles to this. The vectors $\sigma_1\vec{u}_1$ and $\sigma_2\vec{u}_2$ therefore represent the local stretching of the mesh.

In the following figures, we show the skewness factor Q, together with the vectors $\sigma_1\vec{u}_1$ and $\sigma_2\vec{u}_2$, for several of the examples considered previously. For brevity, we use an icosahedral mesh in each case. Figure 10 shows the smoother top-hat regional mesh example, and fig. 11 shows the sech-based ring example.

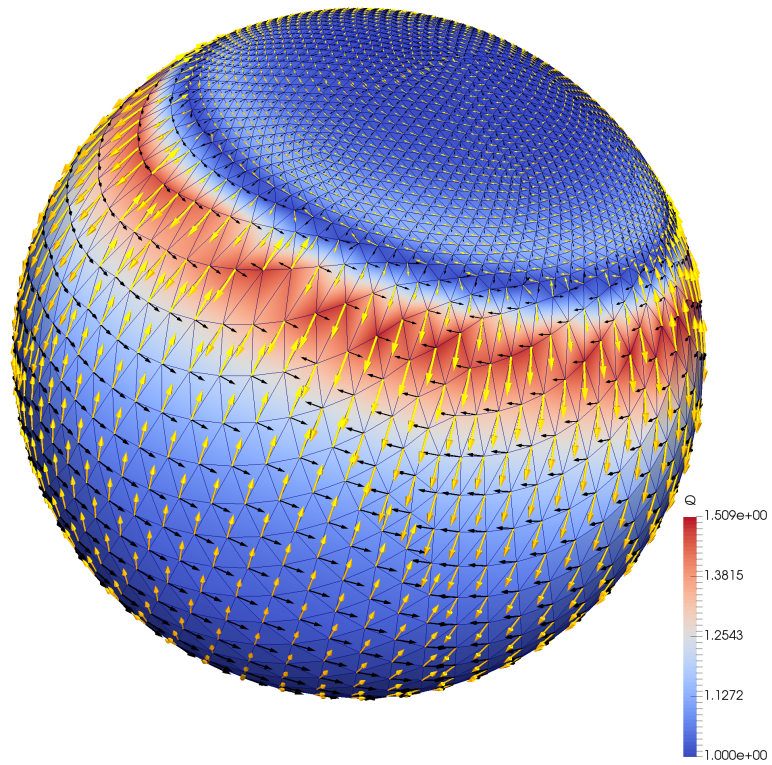The results of these numerical calculations agree with the earlier analytical

26

Figure 10: Numerical skewness calculation for the smoother top-hat regional monitor function. The background colour shows the skewness, Q, while the arrows represent the scaled singular vectors $\sigma_1 \vec{u}_1$ (orange) and $\sigma_2 \vec{u}_2$ (black). The 'flipping' of arrows between neighbouring cells is an artifact of the sign-ambiguity of the SVD, and is not supposed to be physically significant. It is clear that Q approaches 1 at the poles $\pm\vec{\omega}$. The maximum skewness is found just outside the high-resolution region, as predicted by fig. 2, and the intermediate band of $Q = 1$ is visible. In the outer region, there is meridional stretching, so the leading singular vector is in the meridional direction. In the inner region, there is meridional compression, so the leading singular vector is in the zonal direction.
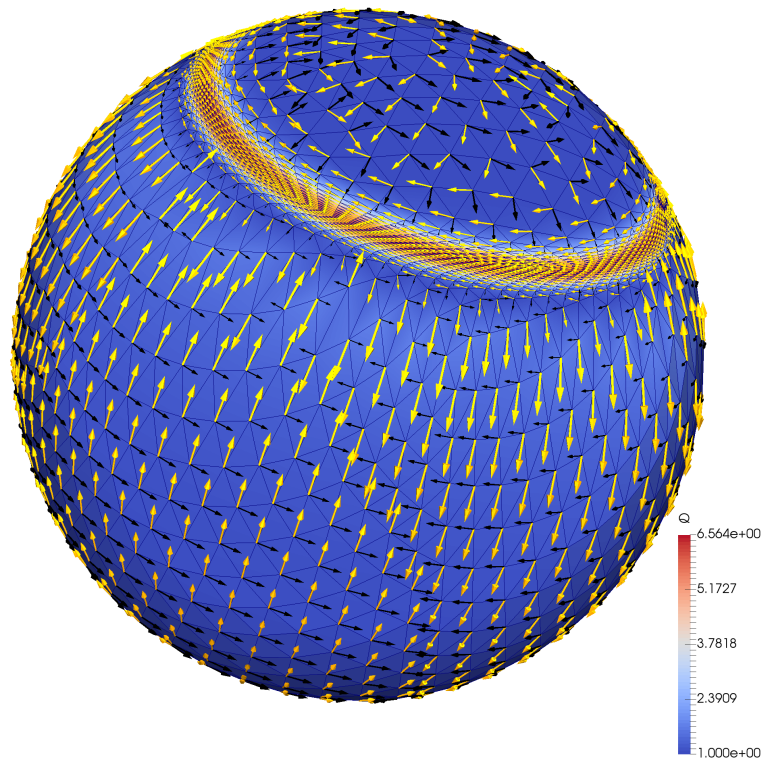
Figure 11: As fig. 10, for the sech-based ring monitor function. The maximum skewness is found in the high-resolution ring, and is close to the earlier prediction of approximately 6.4. The return to $Q = 1$ is visible as a darker blue band just outside the ring. Within the ring, there is extreme meridional compression, so the leading singular vector is in the zonal direction. Outside the ring, there is mild meridional stretching, so the leading singular vector is in the meridional direction. Observe as a consequence, the excellent alignment of the mesh with the ring feature.

estimates. It can be seen that the vectors $\vec{u}_1$ and $\vec{u}_2$ generally point towards the pole and perpendicular to it, a consequence of the axisymmetry of the examples. This is occasionally violated in the regions where $Q \approx 1$, in which there is no dominant direction of local stretching. The maximum skewness is somewhat underestimated in the point singularity example. This is perhaps not surprising: the large skewness is caused by stretching of cells in the meridional direction. The meridional resolution is therefore poor precisely where the skew is large, and fairly large discretisation errors can be expected. Using a once-refined mesh (not shown) gives the closer estimate of 5.59.

6. A comparison with meshes on the plane

We now briefly compare the regularity of a mesh on the sphere with an 'equivalent' mesh on a subset of the plane, extending the analysis in [22]. The conclusion of this section is that meshes on the sphere have, as expected, a higher regularity than those on the plane.

On the plane, we consider a radially-symmetric monitor function, which induces a map $R(r)$ from computational space to physical space. The equidistribution condition then leads to

$$m(R)\frac{\mathrm{d}R}{\mathrm{d}r}R = \alpha r, \tag{58}$$

with α a normalisation constant. This can be compared with eq. (29) for the sphere. For a disc of radius π, α satisfies

$$\frac{1}{2}\alpha\pi^2 = \int_0^\pi m(R)R\,\mathrm{d}R, \tag{59}$$

c.f. eq. (32). In [22], it is shown that the eigenvalues of the map are $\mathrm{d}R/\mathrm{d}r$ and R/r. By following similar steps to section 4.3, the skewness Q is given by

$$Q = \frac{1}{2}\left(\frac{\alpha}{m(R)}\left(\frac{r}{R}\right)^2 + \frac{m(R)}{\alpha}\left(\frac{R}{r}\right)^2\right). \tag{60}$$

For r and R small, we have $R/r \to dR/dr$ as $r \to 0$. As on the sphere, it follows that

$$Q \to 1 \quad \text{as} \quad r, R \to 0, \tag{61}$$

and hence the mesh is very regular in this limit. However, as we approach the boundary (which we assume is mapped to itself, so that $R(a) = a$), we have

$$Q = \frac{1}{2}\left(\frac{\alpha}{m(a)} + \frac{m(a)}{\alpha}\right), \tag{62}$$

This value is completely arbitrary; there is no control over Q as we approach the boundary, and the mesh could be very skew there, as was observed in [22]. This behaviour could not occur on the sphere due to the absence of a boundary. To
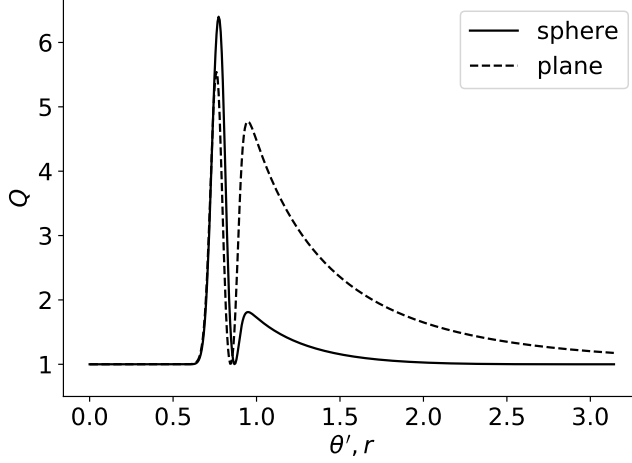
Figure 12: The skewness functions for the maps generated by the sech-based ring monitor function on the sphere and on the plane. Both cases lead to large skewness in the ring itself, which is expected, and is likely to be desirable. However, the differing geometries result in the plane having a far larger secondary peak than the sphere. This secondary peak corresponds to cells outside the ring being radially stretched in order to cover the remaining area.

give a specific numerical example, consider the sech-based ring monitor function eq. (51),

$$m(\theta') = 1 + \frac{\beta}{\epsilon} \operatorname{sech}^2\left(\frac{\theta'^2 - \Theta'^2}{\epsilon}\right), \tag{63}$$

together with an 'equivalent' radially-symmetric monitor function for a disk of radius π,

$$m(R) = 1 + \frac{\beta}{\epsilon} \operatorname{sech}^2\left(\frac{R^2 - \Theta'^2}{\epsilon}\right). \tag{64}$$

For comparison we take $\Theta' = \pi/4$, $\beta = 5\pi/4$, and $\epsilon = \pi/50$. The resulting skewness estimate is shown in fig. 12. Both cases naturally lead to large skewness in the ring itself. However, the secondary peak is much larger in the planar case than for the sphere, corresponding to (unwanted) radial stretching of cells. This is consistent with the mesh shown in Fig 4.3 of [22].

7. More general meshes and comparison to some other methods

In the previous sections, we considered several examples that made use of axisymmetric monitor functions to generate meshes. Whilst these meshes are, in certain cases, computationally useful, and can be analysed exactly, they do not, of course, have the generality of the meshes required for calculating the solution of most PDEs. Without the axisymmetric restriction, the equidistribution requirement and the optimal transport condition leads to a generalised
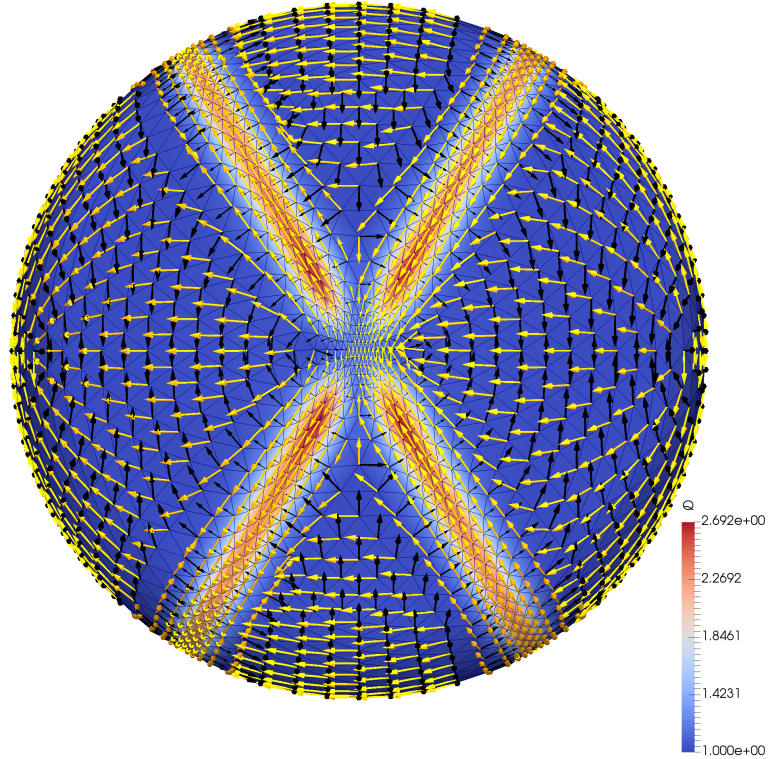
30

Figure 13: As fig. 10, for the 'cross' monitor function eq. (65). The mesh is well-aligned to the intersecting ring features. Within the rings, the dominant singular vector is in the direction of the ring. Away from the cross, the mesh undergoes very mild stretching in order to 'provide' resolution to the ring features.

Monge–Ampère equation which can be expressed with respect to spherical angles, as in section 3.3, or with respect to the background Cartesian space $\mathbb{R}^3$, as was done in [6].

The purpose of this section is to now consider some interesting non-axisymmetric examples of meshes obtained from more general monitor functions. This is done by solving the Monge–Ampère-like equation numerically using the methods described in [6]. We will look at the compression, skewness and alignment of the resulting meshes as well as demonstrating the robustness and flexibility of the method. We also present some examples that can be compared with meshes generated by other methods.

7.1. Two intersecting rings

Our first example is shown in fig. 13 and comprises two intersecting ring features. This is based on (though different to) the 'cross' example in [6]. The

monitor function that generates this is given by

$$m(\vec{x}) = \prod_i^N \left(1 + \alpha_i \operatorname{sech}^2(\beta_i(\|\vec{x} - \omega_i\|^2 - (\pi/2)^2)))\right). \tag{65}$$

We take $N = 2$, $\alpha_1 = \alpha_2 = 5$, $\beta_1 = \beta_2 = 5$, and the axes $\omega_1, \omega_2 = (\pm\sqrt{3}/2, 0, 1/2)$ are such that the rings cross at an angle of 60 degrees. The mesh cells in the ring are reasonably skew, as expected, and locally the mesh compression and regularity are very similar to those for the single ring considered in section 5. The cells outside the rings are almost unaffected and show great regularity. A key observation from this figure is that at the point where the rings intersect, and where we might expect to see a very distorted mesh, we see instead that the skewness is small, and that the mesh cells are very nearly uniform.

7.2. Sinusoidally-varying ring

Our second example is shown in fig. 14. This mesh is induced by a monitor function which concentrates cells within a sinusoidal pattern in the northern hemisphere. It is inspired by the related planar example in [22], and is not dissimilar to a Rossby–Hauritz wave (see, for example, test case 6 from [41]). The monitor function is given by

$$m(\vec{x}) = 1 + \alpha \operatorname{sech}(\beta\theta'), \qquad \theta' = \theta - (\theta_c + \frac{1}{2}\theta_a \sin(k\phi)), \tag{66}$$

where θ is the latitude (now measured from the equator), ϕ is the longitude, $\alpha = 15$, $\beta = 25$, $\theta_c = \frac{\pi}{6}$, $\theta_a = \frac{\pi}{6}$, and we use a wavenumber $k = 3$. The mesh cells are well-aligned with the sinusoidal pattern. There is also only slight stretching of cells outside this high-resolution region and the grid is very regular there. Hence this mesh would be very suitable for computing a Rossby-Hauritz wave with good resolution. The local properties of this mesh close to the wave are very similar to that of the ring examples considered earlier.

7.3. Equatorially-enhanced mesh

Inspired by [16], our next example is a mesh that concentrates points in an equatorial region. That paper uses a specially-constructed triangular mesh with an elaborate topology that places far more cells around the equator than a normal icosahedral mesh. A spring dynamics approach is then used to smooth the mesh, followed by analytical transformations around problematic points. Here, we use the optimal transport procedure to generate a non-uniform mesh from a nearly-uniform icosahedral mesh, requesting a similar distribution of resolution.

We define a target grid spacing d^*, which is a function of latitude θ' only (expressed in degrees). This takes the values

$$d^*(\theta') = \begin{cases} 0.064, & |\theta'| < 13 \\ 0.064 + \frac{|\theta'|-13}{31-13}(0.23 - 0.064), & 13 \leq |\theta'| \leq 31 \\ 0.23, & |\theta'| > 31. \end{cases} \tag{67}$$
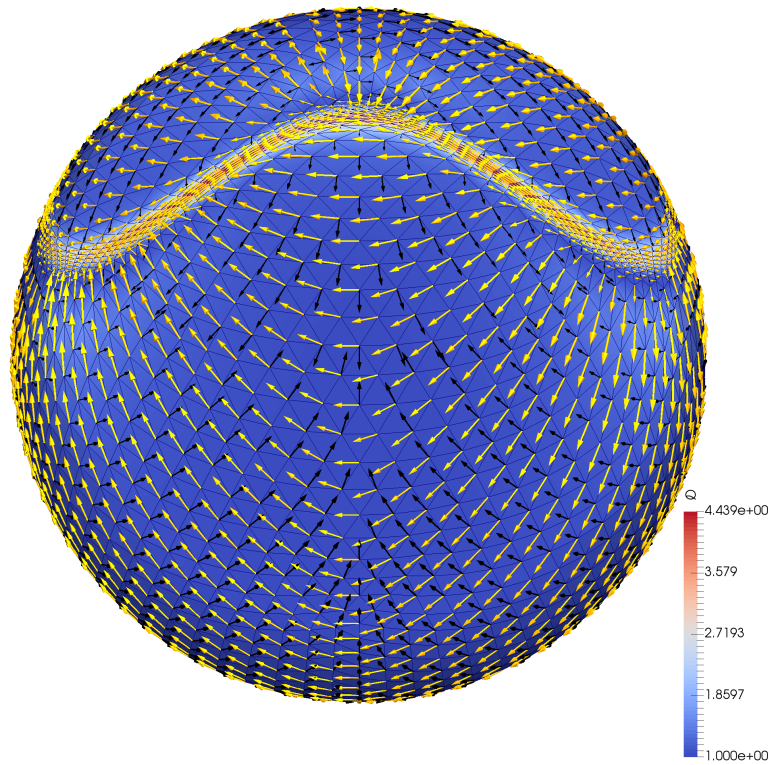
32

Figure 14: As fig. 10, for the sinusoidal monitor function eq. (66). The dominant singular vectors show that the mesh is very well aligned to the sinusoidal feature. There is some mild stretching of the mesh immediately outside the feature in order to provide the enhanced resolution. Away from the feature, the mesh is incredibly regular.
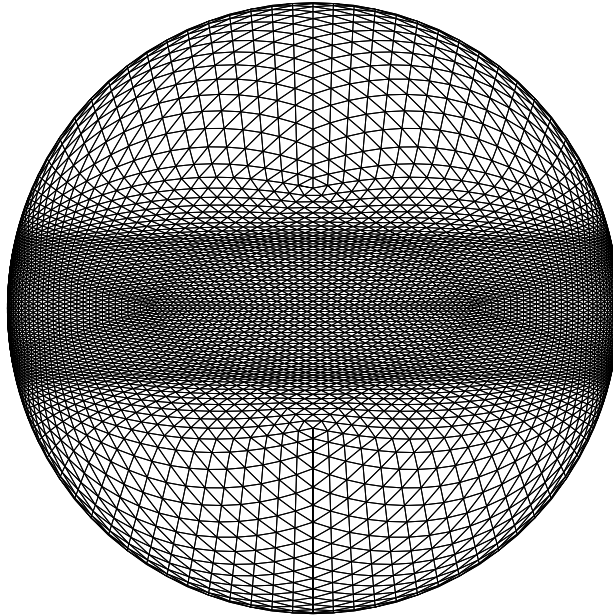
Figure 15: An equatorially-enhanced mesh produced using the optimal transport approach.

This is a similar target grid spacing to the 'analytic resolution distribution' in Figure 8a of [16] (note that this is *derived from their mesh topology*, rather than being specified in advance). We then use a monitor function $m \propto 1/d^{*2}$ to control the cell area in the solution of the Monge–Ampère equation. The mesh generation technique automatically obtains the correct constant of proportionality.

The resulting mesh is shown in fig. 15, and a graph of mesh spacing against latitude is given in fig. 16. It is clear from both figures that the desired mesh compression has been achieved, and that the resulting mesh is smooth and has good regularity. In comparison to the meshes in [16], our technique leads to slightly more stretching of cells around the equator. This is unavoidable, as our topology is fixed, and the only way to reduce cell area is to relocate cells towards the equator from the mid-latitudes.

7.4. A more-uniform icosahedral mesh

Our final example uses optimal transport to tackle the minor nonuniformities in a refined icosahedral mesh. The standard approach to generating a refined icosahedral mesh is to refine the faces of an icosahedron, then to project mesh vertices radially outwards onto the surface of the sphere. The resulting mesh is reasonably uniform, but the ratio of maximum to minimum cell area is approximately 2. Using our optimal transport approach, we can equalise the cell areas.
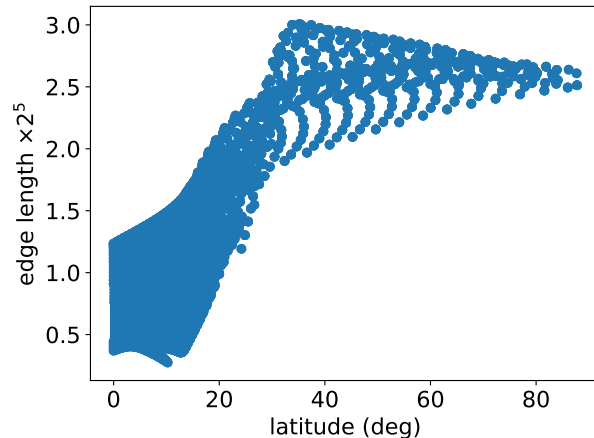
34

Figure 16: A graph of normalized edge length against latitude for the equatorially-enhanced mesh (for symmetry reasons, only the northern hemisphere is shown).

This problem is slightly different to what we have done previously: compared to eq. (3), we instead have $m(\vec{\xi})r(\vec{\xi}) = \alpha$. The monitor function now depends on the computational coordinate $\vec{\xi}$ rather than the physical coordinate $\vec{x}$. The computation of the map is slightly easier, as a result, since one source of nonlinearity is removed from the resulting PDE. We represent m as a piecewise-constant function defined by the areas of each cell on the unadjusted mesh.

We give some graphs analysing the effect of this procedure. A graph of cell area against angle from a specific 5-neighbour point is shown in fig. 17; this can be compared with Figures 1 and 13 in [14]. The resulting cell areas still vary by about 1%, due to discretisation error, but this compares very favourably with other methods. The optimal transport approach controls the area scaling, so it is natural that we have strong control over the area of the resulting mesh cells. In fig. 18, we plot the normalised edge lengths against angle. Globally, there is less variation in the edge lengths after the optimal transport procedure, which is consistent with the equalisation of cell areas. In fig. 19, we plot a discrete measure of mesh regularity. This is defined (for ease of comparison with other methods) as the minimum-to-maximum edge ratio for each cell. The adjustment procedure leads to cells which are generally slightly less regular (equilateral) than the original. Overall, we see that the optimal transport approach has acted as an effective 'mesh smoother', giving a more uniform mesh than the original. The adjusted mesh is thus very suitable for computing solutions of certain PDEs.
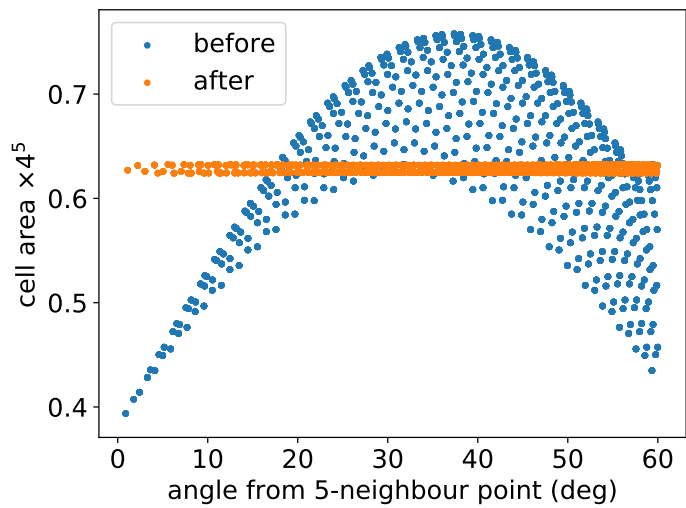
35

Figure 17: A graph of cell area against angle from the special 5-neighbour points, before and after the optimal transport procedure. Before adjustment, the smallest cells are gathered around these special points, and the max-min area ratio is approximately 1.925. After adjustment, the max-min area ratio is just 1.013.
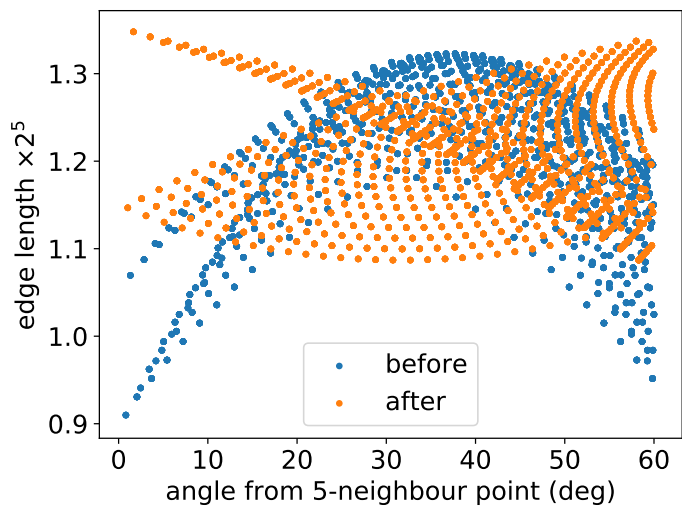

Figure 18: A graph of normalised edge length against angle from the special 5-neighbour points, before and after the optimal transport procedure. Globally, the edge lengths become more uniform, with a resulting variation of approximately 20%.
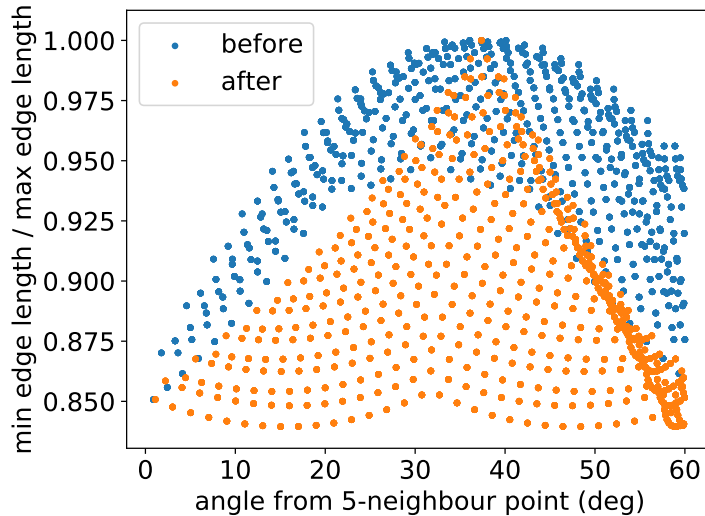
Figure 19: A graph of cell regularity, defined as the ratio of minimum to maximum edge length, against angle from the special 5-neighbour points, before and after the optimal transport procedure. While the total range has hardly changed, the equal-area mesh has slightly less regular cells than the unadjusted mesh.

8. Conclusions

In this paper, we have described a flexible method that produces dynamic adapted meshes on the sphere which are topologically identical to a given logical mesh. This involves calculating a map from the sphere to itself, by solving a PDE of Monge–Ampère type, and applying this map to the input mesh. By using an optimal transport strategy to find this map, we have a robust method of constructing such a dynamic mesh. Significantly, a-priori mesh regularity properties are inherited from the regularity of this optimal transport map. In particular, we show, both theoretically and by example, that such meshes on the sphere are more regular than similar meshes on the plane.

By looking at a particular class of solutions, we analytically derived some specific maps and also their associated skewness properties. Hence we could generate various meshes, with provable regularity estimates, which could be used for practical computations. We also considered more general examples, calculating the maps and meshes numerically, and showed that the resulting meshes still have good regularity properties and compare favourably to those given by other methods. We observe that the meshes generated through the optimal transport approach are capable of producing anisotropic meshes aligned with the solution features, even though the monitor function is only scalar-valued. In another example we have demonstrated that optimal transport is effective as a 'mesh smoother' for increasing the uniformity of a mesh.

In this paper, we only considered the case where the domain is the entire

37

sphere S^2, which has no boundary. Some applications may require a mesh for only a subset of the sphere, particularly a geophysical simulation such as an ocean on the earth. In Euclidean space, optimal transport theory requires convex domains, which would seem to rule out this sort of application. Further-more, [28] only considers boundary-free manifolds. However, it is sometimes possible to bypass the convexity requirement by working on an extended, convex, domain, and setting the monitor to zero outside the true domain. This is done in, for example, [42]. We have not attempted to replicate this on the sphere.

A realistic geophysical application on the sphere would likely require a three-dimensional mesh. The earlier paper [21] already considered fully three-dimensional optimal transport mesh adaptivity in a cuboid domain. Unfortunately, the analogous spherical shell is not a convex domain, so our method would not work without further modifications. However, it is unlikely that full three-dimensional adaptivity is desirable in a geophysical application, since accurate representation of pressure gradient terms necessitates that cells should be in vertically-aligned columns. It is more likely that some kind of 2+1D adaptivity would be used. The base mesh can be adapted following the methods described in this paper, then the nodes in each column can be relocated up or down separately. This also reduces the computational complexity of the problem significantly.

Of course, for useful problems, finding a mesh is just one part of solving a physical problem represented by the time-evolving solution of a PDE. The equation must then be discretised on the mesh using, for example, a finite volume or finite element approach. There are numerous issues that must still be investigated. In the context of advection-dominated flows, this includes the accuracy and stability of the solution, the dispersion relations of any waves, the ability of the mesh to support balanced flows, and the construction of suitable monitor functions to achieve these. This ongoing research will be the subject of future papers.

Acknowledgements

Appendix A. Code availability

The numerical calculations made use of *SciPy* [43], particularly the integration and optimisation routines. When running the numerical mesh-generation

methods developed in [6], we used the *Firedrake* finite element software [44], in-

₇₉₅ cluding specialist functionality developed in [45, 46, 47, 48, 49, 50, 51]. Firedrake itself relies on *PETSc* [52, 53] and *petsc4py* [54].

The code for the numerical experiments can be found in the supplementary material to this paper. While the code should be compatible with Firedrake for the foreseeable future, the precise versions of Firedrake components that were

₈₀₀ used in this paper are archived at [55].

References

[1] J. Slingo, K. Bates, N. Nikiforakis, M. Piggott, M. Roberts, L. Shaffrey, I. Stevens, P. L. Vidale, H. Weller, Developing the next-generation climate system models: challenges and achievements, Philosophical Transactions
₈₀₅ of the Royal Society A: Mathematical, Physical and Engineering Sciences 367 (1890) (2009) 815–831. `doi:10.1098/rsta.2008.0207`.

[2] A. Staniforth, J. Thuburn, Horizontal grids for global weather and climate prediction models: a review, Quarterly Journal of the Royal Meteorological Society 138 (662) (2012) 1–26. `doi:10.1002/qj.958`.

₈₁₀ [3] A. E. MacDonald, J. Middlecoff, T. Henderson, J.-L. Lee, A general method for modeling on irregular grids, International Journal of High Performance Computing Applications 25 (4) (2010) 392–403. `doi:10.1177/1094342010385019`.

[4] G.-T. Bercea, A. T. T. McRae, D. A. Ham, L. Mitchell, F. Rathgeber,
₈₁₅ L. Nardi, F. Luporini, P. H. J. Kelly, A structure-exploiting numbering algorithm for finite elements on extruded meshes, and its performance evaluation in Firedrake, Geoscientific Model Development 9 (10) (2016) 3803–3815. `doi:10.5194/gmd-9-3803-2016`.

[5] H. Weller, P. Browne, C. Budd, M. Cullen, Mesh adaptation on the sphere
₈₂₀ using optimal transport and the numerical solution of a Monge–Ampère type equation, Journal of Computational Physics 308 (2016) 102–123. `doi:10.1016/j.jcp.2015.12.018`.

[6] A. T. T. McRae, C. J. Cotter, C. J. Budd, Optimal-Transport–Based Mesh Adaptivity on the Plane and Sphere Using Finite Elements, SIAM Jour-
₈₂₅ nal on Scientific Computing 40 (2) (2018) A1121–A1148. `doi:10.1137/16M1109515`.

[7] D. L. Williamson, The Evolution of Dynamical Cores for Global Atmospheric Models, Journal of the Meteorological Society of Japan 85B (2007) 241–269. `doi:10.2151/jmsj.85B.241`.

₈₃₀ [8] P. A. Ullrich, C. Jablonowski, J. Kent, P. H. Lauritzen, R. Nair, K. A. Reed, C. M. Zarzycki, D. M. Hall, D. Dazlich, R. Heikes, C. Konor, D. Randall, T. Dubos, Y. Meurdesoif, X. Chen, L. Harris, C. Kühnlein, V. Lee,

A. Qaddouri, C. Girard, M. Giorgetta, D. Reinert, J. Klemp, S.-H. Park, W. Skamarock, H. Miura, T. Ohno, R. Yoshida, R. Walko, A. Reinecke, K. Viner, DCMIP2016: a review of non-hydrostatic dynamical core design and intercomparison of participating models, Geoscientific Model Development 10 (12) (2017) 4477–4509. `doi:10.5194/gmd-10-4477-2017`.

[9] M. Hortal, A. J. Simmons, Use of Reduced Gaussian Grids in Spectral Models, Monthly Weather Review 119 (4) (1991) 1057–1074. `doi:10.1175/1520-0493(1991)119<1057:UORGGI>2.0.CO;2`.

[10] P. McCorquodale, P. A. Ullrich, H. Johansen, P. Colella, An adaptive multi-block high-order finite-volume method for solving the shallow-water equations on the sphere, Communications in Applied Mathematics and Computational Science 10 (2) (2015) 121–162. `doi:10.2140/camcos.2015.10.121`.

[11] J. O. Ferguson, C. Jablonowski, H. Johansen, P. McCorquodale, P. Colella, P. A. Ullrich, Analyzing the Adaptive Mesh Refinement (AMR) Characteristics of a High-Order 2D Cubed-Sphere Shallow-Water Model, Monthly Weather Review 144 (12) (2016) 4641–4666. `doi:10.1175/MWR-D-16-0197.1`.

[12] M. Satoh, H. Tomita, H. Yashiro, H. Miura, C. Kodama, T. Seiki, A. T. Noda, Y. Yamada, D. Goto, M. Sawada, T. Miyoshi, Y. Niwa, M. Hara, T. Ohno, S.-i. Iga, T. Arakawa, T. Inoue, H. Kubokawa, The Non-hydrostatic Icosahedral Atmospheric Model: description and development, Progress in Earth and Planetary Science 1 (2014) 18. `doi:10.1186/s40645-014-0018-1`.

[13] H. Tomita, M. Satoh, K. Goto, An Optimization of the Icosahedral Grid Modified by Spring Dynamics, Journal of Computational Physics 183 (1) (2002) 307–331. `doi:10.1006/jcph.2002.7193`.

[14] S.-i. Iga, H. Tomita, Improved smoothness and homogeneity of icosahedral grids using the spring dynamics method, Journal of Computational Physics 258 (2014) 208–226. `doi:10.1016/j.jcp.2013.10.013`.

[15] S.-i. Iga, Smooth, seamless, and structured grid generation with flexibility in resolution distribution on a sphere based on conformal mapping and the spring dynamics method, Journal of Computational Physics 297 (2015) 381–406. `doi:10.1016/j.jcp.2015.05.026`.

[16] S.-i. Iga, An equatorially enhanced grid with smooth resolution distribution generated by a spring dynamics method, Journal of Computational Physics 330 (2017) 794–809. `doi:10.1016/j.jcp.2016.10.017`.

[17] T. D. Ringler, D. Jacobsen, M. Gunzburger, L. Ju, M. Duda, W. Skamarock, Exploring a Multiresolution Modeling Approach within the Shallow-Water Equations, Monthly Weather Review 139 (11) (2011) 3348–3368. `doi:10.1175/MWR-D-10-05049.1`.

40

[18] C. J. Budd, J. F. Williams, Parabolic Monge–Ampère methods for blow-up problems in several spatial dimensions, Journal of Physics A: Mathematical and General 39 (19) (2006) 5425–5444. `doi:10.1088/0305-4470/39/19/S06`.

[19] C. J. Budd, J. F. Williams, Moving Mesh Generation Using the Parabolic Monge–Ampère Equation, SIAM Journal on Scientific Computing 31 (5) (2009) 3438–3465. `doi:10.1137/080716773`.

[20] L. Chacón, G. Delzanno, J. Finn, Robust, multidimensional mesh-motion based on Monge–Kantorovich equidistribution, Journal of Computational Physics 230 (1) (2011) 87–103. `doi:10.1016/j.jcp.2010.09.013`.

[21] P. Browne, C. Budd, C. Piccolo, M. Cullen, Fast three dimensional r-adaptive mesh redistribution, Journal of Computational Physics 275 (2014) 174–196. `doi:10.1016/j.jcp.2014.06.009`.

[22] C. Budd, R. Russell, E. Walsh, The geometry of r-adaptive meshes generated using optimal transport methods, Journal of Computational Physics 282 (2015) 113–137. `doi:10.1016/j.jcp.2014.11.007`.

[23] J. F. Thompson, B. K. Soni, N. P. Weatherill (Eds.), Handbook of grid generation, CRC Press, 1998.

[24] W. Huang, R. D. Russell, Adaptive Moving Mesh Methods, Applied Mathematical Sciences, Springer-Verlag New York, 2011. `doi:10.1007/978-1-4419-7916-2`.

[25] C. Villani, Topics in Optimal Transportation, Graduate Studies in Mathematics, American Mathematical Society, 2003.

[26] C. Villani, Optimal Transport: Old and New, Grundlehren der mathematischen Wissenschaften, Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, 2009. `doi:10.1007/978-3-540-71050-9`.

[27] Y. Brenier, Polar Factorization and Monotone Rearrangement of Vector-Valued Functions, Communications on Pure and Applied Mathematics 44 (4) (1991) 375–417. `doi:10.1002/cpa.3160440402`.

[28] R. J. McCann, Polar factorization of maps on Riemannian manifolds, Geometric And Functional Analysis 11 (3) (2001) 589–608. `doi:10.1007/PL00001679`.

[29] N. S. Trudinger, J. I. Urbas, On second derivative estimates for equations of Monge-Ampère type, Bulletin of the Australian Mathematical Society 30 (3) (1984) 321–334. `doi:10.1017/S0004972700002069`.

[30] L. A. Caffarelli, Interior $W^{2,p}$ estimates for solutions of the Monge-Ampère equation, Annals of Mathematics 131 (1) (1990) 135–150. `doi:10.2307/1971510`.

41

[31] X.-J. Wang, Some counterexamples to the regularity of Monge-Ampère equations, Proceedings of the American Mathematical Society 123 (3) (1995) 841–845. `doi:10.1090/S0002-9939-1995-1223269-0`.

[32] C. E. Gutiérrez, The Monge-Ampère Equation, Birkhäuser Boston, Boston, MA, 2001. `doi:10.1007/978-1-4612-0195-3`.

[33] P. Delanoë, G. Loeper, Gradient estimates for potentials of invertible gradient–mappings on the sphere, Calculus of Variations and Partial Differential Equations 26 (3) (2006) 297–311. `doi:10.1007/s00526-006-0006-4`.

[34] L. A. Caffarelli, C. E. Gutiérrez, Q. Huang, On the regularity of reflector antennas, Annals of Mathematics 167 (1) (2008) 299–323. `doi:10.4007/annals.2008.167.299`.

[35] G. Loeper, On the regularity of solutions of optimal transportation problems, Acta Mathematica 202 (2) (2009) 241–283. `doi:10.1007/s11511-009-0037-8`.

[36] G. Loeper, Regularity of Optimal Maps on the Sphere: the Quadratic Cost and the Reflector Antenna, Archive for Rational Mechanics and Analysis 199 (1) (2011) 269–289. `doi:10.1007/s00205-010-0330-x`.

[37] G. Delzanno, L. Chacón, J. Finn, Y. Chung, G. Lapenta, An optimal robust equidistribution method for two-dimensional grid adaptation based on Monge–Kantorovich optimization, Journal of Computational Physics 227 (23) (2008) 9841–9864. `doi:10.1016/j.jcp.2008.07.020`.

[38] G. Fibich, The Nonlinear Schrödinger Equation: Singular Solutions and Optical Collapse, Applied Mathematical Sciences, Springer International Publishing, 2015. `doi:10.1007/978-3-319-12748-4`.

[39] C. Budd, M. Cullen, E. Walsh, Monge–Ampére based moving mesh methods for numerical weather prediction, with applications to the Eady problem, Journal of Computational Physics 236 (2013) 247–270. `doi:10.1016/j.jcp.2012.11.014`.

[40] C. Budd, R. Russell, E. Walsh, Near optimal mesh generation using optimal transport based methods, In preparation.

[41] D. L. Williamson, J. B. Drake, J. J. Hack, R. Jakob, P. N. Swarztrauber, A Standard Test Set for Numerical Approximations to the Shallow Water Equations in Spherical Geometry, Journal of Computational Physics 102 (1) (1992) 211–224. `doi:10.1016/S0021-9991(05)80016-6`.

[42] J.-D. Benamou, B. D. Froese, A. M. Oberman, Numerical solution of the Optimal Transportation problem using the Monge–Ampère equation, Journal of Computational Physics 260 (2014) 107–126. `doi:10.1016/j.jcp.2013.12.015`.

[43] E. Jones, T. Oliphant, P. Peterson, et al., SciPy: Open Source Scientific Tools for Python, [Online; accessed 2017-06-26] (2001–).
URL http://www.scipy.org/

[44] F. Rathgeber, D. A. Ham, L. Mitchell, M. Lange, F. Luporini, A. T. T. McRae, G.-T. Bercea, G. R. Markall, P. H. J. Kelly, Firedrake: Automating the Finite Element Method by Composing Abstractions, ACM Transactions on Mathematical Software 43 (3) (2016) 24:1–24:27. doi:10.1145/2998441.

[45] M. E. Rognes, D. A. Ham, C. J. Cotter, A. T. T. McRae, Automating the solution of PDEs on the sphere and other manifolds in FEniCS 1.2, Geoscientific Model Development 6 (6) (2013) 2099–2119. doi:10.5194/gmd-6-2099-2013.

[46] M. S. Alnæs, A. Logg, K. B. Ølgaard, M. E. Rognes, G. N. Wells, Unified Form Language: A Domain-Specific Language for Weak Formulations of Partial Differential Equations, ACM Transactions on Mathematical Software 40 (2) (2014) 9:1–9:37. doi:10.1145/2566630.

[47] A. T. T. McRae, G.-T. Bercea, L. Mitchell, D. A. Ham, C. J. Cotter, Automated generation and symbolic manipulation of tensor product finite elements, SIAM Journal on Scientific Computing 38 (5) (2016) S25–S47. doi:10.1137/15M1021167.

[48] M. Homolya, D. A. Ham, A parallel edge orientation algorithm for quadrilateral meshes, SIAM Journal on Scientific Computing 38 (5) (2016) S48–S61. doi:10.1137/15M1021325.

[49] F. Luporini, D. A. Ham, P. H. J. Kelly, An algorithm for the optimization of finite element integration loops, ACM Transactions on Mathematical Software 44 (1) (2017) 3:1–3:26. doi:10.1145/3054944.

[50] M. Homolya, L. Mitchell, F. Luporini, D. A. Ham, TSFC: A Structure-Preserving Form Compiler, SIAM Journal on Scientific Computing 40 (3) (2018) C401–C428. doi:10.1137/17M1130642.

[51] M. Homolya, R. C. Kirby, D. A. Ham, Exposing and exploiting structure: optimal code generation for high-order finite element methods, Submitted to ACM Transactions on Mathematical Software.
URL https://arxiv.org/abs/1711.02473

[52] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, K. Rupp, B. F. Smith, S. Zampini, H. Zhang, H. Zhang, PETSc Users Manual, Tech. Rep. ANL-95/11 - Revision 3.7, Argonne National Laboratory (2016).
URL http://www.mcs.anl.gov/petsc

[53] S. Balay, W. D. Gropp, L. C. McInnes, B. F. Smith, Efficient Management of Parallelism in Object Oriented Numerical Software Libraries, in: E. Arge, A. M. Bruaset, H. P. Langtangen (Eds.), Modern Software Tools in Scientific Computing, Birkhäuser Press, 1997, pp. 163–202.

[54] L. D. Dalcin, R. R. Paz, P. A. Kler, A. Cosimo, Parallel distributed computing using Python, Advances in Water Resources 34 (9) (2011) 1124–1139. `doi:10.1016/j.advwatres.2011.04.013`.

[55] Software used in 'The scaling and skewness of optimally transported meshes on the sphere' (Aug. 2018). `doi:10.5281/zenodo.1345278`. URL `https://doi.org/10.5281/zenodo.1345278`