PURE and APPLIED
ANALYSIS

PAA

GABRIEL KHAN AND JUN ZHANG

THE KÄHLER GEOMETRY OF CERTAIN OPTIMAL TRANSPORT PROBLEMS

msp

vol. 2 no. 2 2020

PURE and APPLIED
ANALYSIS
Vol. 2, No. 2, 2020
https://doi.org/10.2140/paa.2020.2.397

THE KÄHLER GEOMETRY OF CERTAIN OPTIMAL TRANSPORT PROBLEMS

GABRIEL KHAN AND JUN ZHANG

Let X and Y be domains of $\mathbb{R}^n$ equipped with probability measures μ and ν, respectively. We consider the problem of optimal transport from μ to ν with respect to a cost function $c : X \times Y \to \mathbb{R}$. To ensure that the solution to this problem is smooth, it is necessary to make several assumptions about the structure of the domains and the cost function. In particular, Ma, Trudinger, and Wang established regularity estimates when the domains are strongly *relatively c-convex* with respect to each other and the cost function has nonnegative *MTW tensor*. For cost functions of the form $c(x, y) = \Psi(x - y)$ for some convex function $\Psi : \mathcal{M} \to \mathbb{R}$, we find an associated Kähler manifold on $T\mathcal{M}$ whose orthogonal antibisectional curvature is proportional to the MTW tensor. We also show that relative c-convexity geometrically corresponds to geodesic convexity with respect to a dual affine connection on $\mathcal{M}$. Taken together, these results provide a geometric framework for optimal transport which is complementary to the pseudo-Riemannian theory of Kim and McCann (*J. Eur. Math. Soc.* **12**:4 (2010), 1009–1040).

We provide several applications of this work. In particular, we find a complete Kähler surface with nonnegative orthogonal antibisectional curvature that is not a Hermitian symmetric space or biholomorphic to $\mathbb{C}^2$. We also address a question in mathematical finance raised by Pal and Wong (2018, arXiv:1807.05649) on the regularity of *pseudoarbitrages*, or investment strategies which outperform the market.

1. Introduction

Optimal transport is a classic field of mathematics combining ideas from geometry, probability, and analysis. The problem was first formalized by Gaspard Monge [1781]. In his work, he considered a worker who is tasked with moving a large pile of sand into a prescribed configuration and wants to minimize the total effort required to complete the job. Trying to determine the optimal way of transporting the sand leads into deep and subtle mathematical phenomena and is a thriving field of research to this day. Furthermore, optimal transport has many practical applications. Monge's work was originally inspired by a problem in engineering, but these same ideas can be applied to logistics, economics, computer imaging processing, and many other fields [Peyré and Cuturi 2019].

The modern framework for optimal transport, due to Kantorovich [1958], considers arbitrary couplings between two probability measures. In this formulation, we consider X and Y as Borel subsets of two metric spaces equipped with probability measures μ and ν, respectively. Intuitively, $d\mu$ is the shape of the original sand pile and $d\nu$ is the target configuration. To transport the sand from μ to ν, we consider a *coupling* of μ and ν, which is a nonnegative measure on $X \times Y$ whose marginal distributions are μ and ν,

MSC2010: primary 49Q20, 53C55; secondary 46N10, 46N30.
Keywords: optimal transportation, Kähler metrics, regularity of optimal maps, complex geometry, MTW condition, curvature, Monge–Kantorovich, tube domains, tangent bundle.

respectively. To measure the efficiency of a plan for transport μ to ν, we consider a lower-semicontinuous cost function $c : X \times Y \to \mathbb{R}$. The solution to the Kantorovich optimal transport problem is the coupling γ which achieves the smallest total cost

$$\min_{\gamma \in \Gamma(\mu,\nu)} \int_{X \times Y} c(x, y)\, d\gamma(x, y).$$

Here $\Gamma(\mu, \nu)$ is the set of all couplings of μ and ν. In this case, a minimizing measure γ is referred to as the *optimal coupling*. An optimal coupling exists for very general measures and cost functions, so the Kantorovich approach is a flexible and powerful framework to study optimal transport.

In Monge's work, it is assumed that the mass at a given point will not be subdivided and sent to multiple locations. This is known as *deterministic* optimal transport, which seeks to find a measurable map $\mathbb{T} : X \to Y$ so that the optimal coupling is entirely contained within the graph of $\mathbb{T}$. When this occurs, the map $\mathbb{T}$ is known as the *optimal map*. A priori, there is no guarantee that optimal transport is deterministic, so a Monge solution may not exist for a given optimal transport problem. We will discuss certain sufficient conditions for the optimal transport to be deterministic in Section 2.

For deterministic optimal transport, it is natural to ask whether the optimal map is continuous or even smooth. This is known as the *regularity problem for optimal transport*. Historically, most of the work on this problem was done in Euclidean space for the cost $c(x, y) = \|x - y\|^2$, better known as the quadratic cost.

For more general cost functions (such as quadratic costs on Riemannian manifolds), the groundbreaking work was done by Ma, Trudinger and Wang [Ma et al. 2005], who proved that the transport map is smooth under the assumptions that

(1) a certain nonlinear fourth-order quantity, known as the MTW tensor (denoted by $\mathfrak{S}$), is nonnegative, and that

(2) the sets X and Y are *relatively c-convex* with respect to each other.[1]

These results were refined by Loeper [2009], who showed that the nonnegativity of $\mathfrak{S}$ is necessary to establish continuity for the optimal transport between smooth measures. Furthermore, he gave some insight into the geometric significance of the MTW tensor. Later work of Kim and McCann [2010] furthered this understanding by presenting a pseudo-Riemannian framework for optimal transport in which the MTW tensor is the curvature of certain light-like planes.

1.1. *Our results.*

In this paper, we primarily consider Ψ-costs, which we define as follows.

Definition (Ψ-cost). Let $\Psi : \mathcal{M} \to \mathbb{R}$ be a locally strongly convex C^4 function[2] on an open domain $\mathcal{M}$ in Euclidean space.

For open domains X and Y in $\mathbb{R}^n$, a Ψ-cost is a cost function of the form

$$c : X \times Y \to \mathbb{R}, \quad c(x, y) = \Psi(x - y).$$

[1] More precisely, the assumption is that the supports of μ and ν are relatively c-convex.

[2] Here, a function being locally strongly convex means that the Hessian is positive definite. Furthermore, it is possible to work with less regular convex functions, but we will not do so in this paper.

These costs were previously studied by Gangbo and McCann [1995] and by Ma, Trudinger and Wang [2005]. For such a cost to be well-defined, $\mathcal{M}$ must contain the difference set $X - Y$, defined as

$$X - Y := \{z \in \mathbb{R}^n \mid \text{there exists } x \in X,\ y \in Y \text{ such that } z = x - y\}.$$

We can now summarize the main results of our work, which associate a complex manifold to a given Ψ-cost. To do so, we consider $\mathcal{M}$ as a Hessian manifold, using Ψ as its potential function (i.e., setting $g_{ij} = \partial^2\Psi/(\partial u^i \partial u^j)$). Such manifolds naturally admit a dual pair of flat connections, which we denote by D and D^* [Shima 2007].

Using the primal flat connection D and the metric g, there is a canonical Kähler metric on the tangent bundle, known as the Sasaki metric and denoted by $(T\mathcal{M}, g^D, J^D)$. Our main result shows the following correspondence between the curvature of this metric and the MTW tensor.

Theorem. *Let X and Y be open sets in $\mathbb{R}^n$ and c be a Ψ-cost. Then the MTW tensor $\mathfrak{S}$ satisfies the identity*

$$\tfrac{1}{2}\mathfrak{S}(\eta, \xi) = \mathfrak{R}_{g^D}(\xi, J^D\eta^\sharp, \xi, J^D\eta^\sharp) - \mathfrak{R}_{g^D}(\eta^\sharp, \xi, \eta^\sharp, \xi),$$

where ξ and η are an orthogonal real vector-covector pair (which we extend[3] to $T\mathcal{M}$) and $\mathfrak{R}_{g^D}$ is the curvature of $(T\mathcal{M}, g^D, J^D)$ (where the metric is induced by the potential Ψ).

For reasons that we will explain later, we call the right-hand expression the *orthogonal antibisectional curvature*. We furthermore show that relative c-convexity of sets is geodesic convexity with respect to the dual affine connection on $\mathcal{M}$.

Proposition. *For a Ψ-cost, a set Y is c-convex relative to X if and only if, for all $x \in X$, the set $x - Y$ is geodesically convex with respect to the dual connection D^*. Here, D^* is the connection on $\mathcal{M}$ satisfying*

$$\mathcal{X}(g(\mathcal{Y}, \mathcal{Z})) = g(D_{\mathcal{X}}\mathcal{Y}, \mathcal{Z}) + g(\mathcal{Y}, D^*_{\mathcal{X}}\mathcal{Z})$$

for all vector fields $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$.

Apart from providing a new geometric framework for the regularity problem, we can use these results to address several questions of independent interest.

1.1.1. *Applications to complex geometry.* This approach can be used to construct several examples of interesting metrics with subtle nonnegativity properties. In particular, we find a complete complex surface which is neither biholomorphic to $\mathbb{C}^2$ nor Hermitian symmetric but whose orthogonal antibisectional curvature is nonnegative. Many of the complex manifolds constructed using this approach are of independent interest, and we will provide a few examples which we will study in depth in future work.

1.1.2. *Applications to mathematical finance.* Our second main application is to establish regularity for a certain problem in portfolio design theory. The recent work [Pal and Wong 2016] studies the problem of finding *pseudoarbitrages*, which are investment strategies that outperform the market almost surely in the long run under mild and realistic assumptions on the stock market. Their work shows that this is

[3] To be more precise, we consider certain lifts of these vectors to $T\mathcal{M}$. For a more formal statement, see Theorem 6.

equivalent to solving an optimal transport problem where the cost function is a divergence function (in information-geometric language) that is closely related to the free energy in statistical physics.

For this problem, our approach relates the MTW tensor of this cost to a Kähler manifold with constant positive holomorphic sectional curvature. As such, this cost function satisfies the MTW(0) condition (and also satisfies a stronger condition known as *nonnegative cost-curvature* [Figalli et al. 2011]). We further show that relative c-convexity corresponds precisely to the standard notion of convexity on the probability simplex. Combining these calculations, we can apply the results of [Trudinger and Wang 2009] to obtain a regularity theory of portfolio maps and their associated displacement interpolations. This addresses a question asked in [Pal and Wong 2018b], and intuitively shows that when the market conditions change slightly, the investment strategy similarly does not change by much.

A preliminary announcement of some of these results (stated in terms of the so-called $\mathcal{D}_{\Psi}^{(\alpha)}$-divergences) appeared in [Khan and Zhang 2019].

1.2. ***Layout of the paper.*** In Section 2 we discuss some background information on optimal transport. In Section 3 we review some complex and Kähler geometry. Section 4 discusses some background information on Hessian manifolds and the curvature of the Sasaki metric. In Section 5, we state our main results, which show the precise interaction between complex/information geometry and the regularity theory of optimal transport. In Section 6, we explore various applications of this result. In Section 7, we conclude with a section of open questions, which we hope to explore in future work.

1.3. ***Notation.*** We have attempted to preserve the notation from [De Philippis and Figalli 2014; Satoh 2007] as much as possible, while minimizing abuse of notation or overlap. For clarity, we introduce some notational conventions now.

Throughout the paper, X and Y will denote open domains in $\mathbb{R}^n$. Invariably, these will be smooth and bounded. We will use $\{x^i\}_{i=1}^n$ as coordinates on X and $\{y^i\}_{i=1}^n$ as coordinates on Y. To study optimal transport, we will use $c(x, y)$ to denote a cost function $c : X \times Y \to \mathbb{R}$, which will generally be C^4 in this paper. Often times, the domain of c will be larger than $X \times Y$, but we will ignore this. To avoid confusion with coordinate functions and the notation for tangent spaces, we denote the solutions to equations of Monge–Ampère type by U, and the associated optimal map by $\mathbb{T}_{\mathrm{U}}$.

For the most part, $\mathcal{M}$ will be an open domain in Euclidean space which contains $X - Y$, and Ψ will denote a convex function $\Psi : \mathcal{M} \to \mathbb{R}$. It is instructive to also consider $\mathcal{M}$ as an affine manifold, and we will use $\{u^i\}_{i=1}^n$ as its coordinates. When considering the tangent bundle of $\mathcal{M}$ (denoted by $T\mathcal{M}$), we will use bundle coordinates $\{(u^i, v^i)\}_{i=1}^n$. This notation is a change from [Satoh 2007] and is done to avoid overusing x and y.

In order to prescribe $T\mathcal{M}$ with a Hermitian structure, it is necessary to consider a flat affine connection on $\mathcal{M}$, which we denote by D. More precisely, this will be the affine connection induced by differentiation within the u-coordinates. Furthermore, we use $\mathcal{W}$, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$ and ξ to denote tangent vectors on $\mathcal{M}$ (i.e., elements of $T\mathcal{M}$). This is the convention of [Satoh 2007], except with calligraphic font to avoid confusion with our notation for domains. When computing the MTW tensor, we will denote the vectors in the MTW tensor by ξ and the covectors by η.

To simplify the derivative notation, for a two-variable function $c(x, y)$, we use $c_{I,J}$ to denote $\partial_{x^I}\partial_{y^J}c$ for multi-indices I and J. Furthermore, $c^{i,j}$ denotes the matrix inverse of the mixed derivative $c_{i,j}$. For a convex function Ψ, we use the notation Ψ_J to denote $\partial_{u^J}\Psi$ for a multi-index J and the notation Ψ^{ij} to denote the matrix inverse of Ψ_{ij}. Finally, we will use Einstein summation notation throughout the paper.

2. Background on the regularity theory of optimal transport

The main focus of our paper is to study the assumptions needed to ensure optimal transport is regular. In order to understand these, we first review several preliminary results on the regularity theory of optimal transport.

As our primary interest is the geometric structure of the regularity problem, we will not make use of the sharpest possible regularity estimates. The material in this section is based on the survey [De Philippis and Figalli 2014], which provides a more complete background for the regularity theory. For a more thorough overview on optimal transport, see [Villani 2009].

The regularity problem arises when the optimal coupling in the Kantorovich optimal transport problem is induced by a deterministic transport map. As such, we first discuss some conditions which ensure the Kantorovich optimal transport problem has a deterministic solution. The following theorem was originally proven in [Brenier 1987] for the quadratic cost and in more generality in [Gangbo and McCann 1996]. It gives sufficient conditions for deterministic transport and shows that the optimal maps can be found by solving an equation of Monge–Ampère type.

Theorem 1. *Let X and Y be two open domains of $\mathbb{R}^n$ and consider a cost function $c : X \times Y \to \mathbb{R}$. Suppose that $d\mu$ is a smooth probability density supported on X and that $d\nu$ is a smooth probability density supported on Y. Suppose that the following conditions hold*:

(1) *The cost function c is of class C^4 with $\|c\|_{C^4(X\times Y)} < \infty$.*

(2) *For any $x \in X$, the map $Y \ni y \to c_x(x, y) \in \mathbb{R}^n$ is injective.*

(3) *For any $y \in Y$, the map $X \ni x \to c_y(x, y) \in \mathbb{R}^n$ is injective.*

(4) $\det(c_{x,y})(x, y) \neq 0$ *for all $(x, y) \in X \times Y$.*

Then there exists a c-convex function $\mathrm{U} : X \to \mathbb{R}$ such that the map $\mathbb{T}_{\mathrm{U}} : X \to Y$ defined by $\mathbb{T}_{\mathrm{U}}(x) := c\text{-}\exp_x(\nabla \mathrm{U}(x))$ is the unique optimal transport map sending μ onto ν. Furthermore, $\mathbb{T}_{\mathrm{U}}$ is injective $d\mu$-a.e.,

$$|\det(\nabla \mathbb{T}_{\mathrm{U}}(x))| = \frac{d\mu(x)}{d\nu(\mathbb{T}_{\mathrm{U}}(x))} \quad d\mu\text{-a.e.,} \tag{1}$$

and its inverse is given by the optimal transport map sending ν onto μ.

In order to express (1) more concretely, we recall the notion of the c-exponential map (denoted by $c\text{-}\exp_x$).

Definition (c-exponential map). For any $x \in X$, $y \in Y$, $p \in \mathbb{R}^n$, the c-exponential map satisfies the identity

$$c\text{-}\exp_x(p) = y \quad \iff \quad p = -c_x(x, y).$$

For the squared-distance cost on a Riemannian manifold, the c-exponential is exactly the standard exponential map, which motivates its name. For this cost in Euclidean space, (1) becomes the standard Monge–Ampère equation

$$\det(\nabla^2 \mathrm{U}(x)) = \frac{f(x)}{g(\nabla \mathrm{U}(x))}. \tag{2}$$

Due to the comparatively simple form for (2), much of the initial work on the regularity problem was done for the quadratic cost in Euclidean space. In this setting, Caffarelli [1992] and others proved a priori estimates under certain convexity and smoothness assumptions on the measures (for a more complete history, see [De Philippis and Figalli 2014]). Caffarelli also observed there is no hope of proving interior regularity for $\mathbb{T}_{\mathrm{U}}$ without assuming that the support of the target measure is convex.

For more general cost functions, Ma, Trudinger and Wang's breakthrough work [Ma et al. 2005] gave three conditions that ensure C^2 regularity for the solutions of the Monge–Ampère equation (1). In this paper, we will use a stronger version of this result, originally proven in [Trudinger and Wang 2009].

Theorem 2. *Suppose that $c : X \times Y \to \mathbb{R}$, μ, and ν satisfy the hypothesis of Theorem 1, and the densities $d\mu$ and $d\nu$ are bounded away from zero and infinity on their respective supports X and Y. Suppose further that the following hold*:

(1) *X and Y are smooth.*

(2) *The domain X is strictly c-convex relative to the domain Y.*

(3) *The domain Y is strictly c^*-convex relative to the domain X.*

(4) *For all vectors $\xi, \eta \in \mathbb{R}^n$ with $\xi \perp \eta$, the following inequality holds*:

$$\mathfrak{S}(\xi, \eta) := \sum_{i,j,k,l,p,q,r,s} (c_{ij,p} c^{p,q} c_{q,rs} - c_{ij,rs}) c^{r,k} c^{s,l} \xi^i \xi^j \eta^k \eta^l \geq 0. \tag{3}$$

Then $\mathrm{U} \in C^\infty(\bar{X})$ and $\mathbb{T}_{\mathrm{U}} : \bar{X} \to \bar{Y}$ is a smooth diffeomorphism, where $\mathbb{T}_{\mathrm{U}}(x) = c\text{-}\exp_x(\nabla \mathrm{U}(x))$.

While we will not discuss the proof in detail, we note that the main challenge is obtaining an a priori C^2 estimate on U. Once such an estimate is established, the Monge–Ampère equation can be linearized at U, at which point standard elliptic bootstrapping yields estimates of all orders and implies that $\mathbb{T}_{\mathrm{U}}$ is smooth.

The main results of this paper concern the assumptions of Theorem 2, so we discuss these in more detail. The first condition is self-explanatory, while the second and third provide the proper notions of convexity for the supports. To explain this in detail, we recall the notion of c-convexity for sets.

Definition (c-segment). A c-segment in X with respect to a point y is a solution set $\{x\}$ to $c_{,y}(x, y) \in \ell$ for ℓ a line segment in $\mathbb{R}^n$. A c^*-segment in Y with respect to a point x is a solution set $\{y\}$ to $c_{x,}(x, y) \in \ell$, where ℓ is a line segment in $\mathbb{R}^n$.

Definition (c-convexity). A set E is c-convex relative to a set E^* if for any two points $x_0, x_1 \in E$ and any $y \in E^*$, the c-segment relative to y connecting x_0 and x_1 lies in E. Similarly we say E^* is c^*-convex relative to E if for any two points $y_0, y_1 \in E^*$ and any $x \in E$, the c^*-segment relative to x connecting y_0 and y_1 lies in E^*.

Finally, we discuss inequality (3), which is known as the MTW(0) condition and is a weakened version of the MTW(κ) condition.

Definition (MTW(κ), $\kappa > 0$). A cost function c satisfies the MTW(κ) condition if for any orthogonal vector-covector pair ξ and η we have $\mathfrak{S}(\xi, \eta) \geq \kappa |\xi|^2 |\eta|^2$ for $\kappa > 0$.

Ma, Trudinger and Wang's original work relied on the MTW(κ) assumption, and this stronger condition is used in many applications. Although it is not immediately apparent, $\mathfrak{S}(\xi, \eta)$ is tensorial (coordinate-invariant) so long as one considers η as a covector [Kim and McCann 2010], which we will do throughout the rest of the paper. Furthermore, it transforms quadratically in η and ξ, but is highly nonlinear and nonlocal in the cost function.

The geometric significance of the MTW tensor is an active topic of research. On a Riemannian manifold, Loeper [2009] gave some insight into its behavior. His work showed that for the quadratic cost, the MTW tensor is proportional to the sectional curvature on the diagonal $x = y$. In this paper, he also showed that c-convexity and nonnegativity of the MTW tensor are essentially necessary conditions to prove regularity of optimal transport.

Building on Loeper's results, Kim and McCann [2010] gave a geometric framework for optimal transport. In their formulation, optimal transport is expressed in terms of a pseudo-Riemannian metric on the manifold $X \times Y$ and the MTW tensor becomes the curvature of light-like planes. This interpretation holds for arbitrary cost functions, and gives intrinsic geometric structure to the regularity problem. Our geometric interpretation is different, but many of the formulas appear similar, in part due to the fact that Kim and McCann chose notation reminiscent of complex geometry.

Before concluding our background discussion on optimal transport, we will introduce one more strengthening of the MTW(0) condition, known as nonnegative "cross-curvature" [Figalli et al. 2011].

Definition (nonnegative cross-curvature). A cost function c has nonnegative (resp. strictly positive $\kappa > 0$) cross-curvature if, for any vector-covector pair η and ξ,

$$\mathfrak{S}(\xi, \eta) \geq 0 \quad (\text{resp. } \kappa |\xi|^2 |\eta|^2 \text{ for some } \kappa > 0).$$

Note that nonnegative cross-curvature is stronger than MTW(0), as the nonnegativity must hold for all pairs η and ξ, not simply orthogonal ones. Cross-curvature was introduced by Figalli, Kim, and McCann [Figalli et al. 2011] to study a problem in microeconomics. In later work, they also showed that stronger regularity for optimal maps can be proven with this assumption [Figalli et al. 2013]. Cross-curvature was also studied in [Sei 2013] for an application in statistics.

3. Background on Kähler geometry

In order to connect optimal transport with Kähler geometry, we will review some background on Kähler manifolds. We will only discuss what is needed for this work, and refer the reader to [Zheng 2000] for a more complete reference on complex geometry.

Given a smooth manifold[4] $\mathbb{X}$, an *almost complex structure* J is a smoothly varying endomorphism of $T\mathbb{X}$ which satisfies $J^2 = -\mathrm{Id}$. In this case, we say that the pair $(\mathbb{X}, J)$ is an *almost-complex manifold*. From the definition, it immediately follows that any almost-complex manifold must be even-dimensional and orientable.

We say that an almost complex manifold $\mathbb{X}$ is *complex* if it admits an atlas of holomorphic coordinate charts satisfying $z^i = u^i + \sqrt{-1}y^i$ such that $J\partial_{x^i} = \partial_{y^i}$ and $J\partial_{y^i} = -\partial_{x^i}$. In other words, around each point in $\mathbb{X}$, a complex manifold admits a local biholomorphism to a subset of $\mathbb{C}^n$ (in which J acts on tangent vectors as multiplication by $\sqrt{-1}$). If this can be done, we say that the almost complex structure is *integrable*. Due to a deep theorem of [Newlander and Nirenberg 1957], integrability of an almost complex structure J is equivalent to the vanishing of the so-called Nijenhuis tensor, which is defined as

$$N_J(\mathcal{X}, \mathcal{Y}) = -J^2[\mathcal{X}, \mathcal{Y}] + J[\mathcal{X}, J\mathcal{Y}] + J[J\mathcal{X}, \mathcal{Y}] - [J\mathcal{X}, J\mathcal{Y}].$$

Showing that this condition is necessary is relatively straightforward,[5] but it highly nontrivial to show that it is also sufficient.

We say that an almost complex structure is *compatible* with a Riemannian metric g if it satisfies $g(\mathcal{X}, \mathcal{Y}) = g(J\mathcal{X}, J\mathcal{Y})$ for all tangent vectors $\mathcal{X}$ and $\mathcal{Y}$. In this case, the triple $(\mathbb{X}, g, J)$ is said to be a *Hermitian* manifold. Furthermore, we say a Hermitian manifold is *Kähler* if J is integrable and the Kähler form $\omega = g(J\,\cdot\,, \cdot\,)$ is closed (i.e., $d\omega = 0$). This closedness has many important consequences for the geometry of Kähler metrics. Most importantly, in any set of holomorphic coordinates $\{z^i\}_{i=1}^n$, we can express the Kähler form as

$$\omega = \sqrt{-1}\frac{\partial^2\Phi}{\partial z^i\,\bar{\partial} z^j} dz^i \wedge d\bar{z}^j$$

for some strictly plurisubharmonic potential Φ. This leads to many important geometric properties, only a few of which we will explore here.

3.1. ***The curvature of Kähler manifolds.*** In this paper, we will study the curvature for a certain class of Kähler manifolds. As such, it is necessary to review some background on the curvature of Kähler metrics. We will specialize our focus to the curvature of the Levi-Civita connection on $\mathbb{X}$, which we denote as ∇. For non-Kähler Hermitian manifolds, there are several canonical connections which have distinct curvature tensors. Fortunately, all these connections coincide for Kähler manifolds. As such, we can unambiguously denote the curvature tensor as $\mathfrak{R}$, which is defined[6] as

$$\mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) = g(\nabla_{\mathcal{X}}\nabla_{\mathcal{Y}}\mathcal{Z} - \nabla_{\mathcal{Y}}\nabla_{\mathcal{X}}\mathcal{Z} - \nabla_{[\mathcal{X},\mathcal{Y}]}\mathcal{Z}, \mathcal{W}). \tag{4}$$

[4]Not to be confused with our notation for domains.

[5]The Nijenhuis tensor vanishes if the Lie bracket of any two holomorphic vectors is holomorphic, which is automatically true for complex structures.

[6]For this definition to be meaningful, we must extend each of the vectors to vector fields, but since the expression is tensorial, the choice of extension does not matter.

Apart from the usual symmetries of the Riemannian curvature, the curvature of a Kähler metric satisfies the following identity (when $\mathbb{X}$ is regarded as a (real) manifold):

$$\mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) = \mathfrak{R}(J\mathcal{X}, J\mathcal{Y}, \mathcal{Z}, \mathcal{W}). \tag{5}$$

After repeatedly applying this identity and using the other symmetries of the curvature tensor, it is possible to show that we can determine the entire curvature tensor from the values $\mathfrak{R}(\mathcal{X}, \overline{\mathcal{Y}}, \mathcal{Z}, \overline{\mathcal{W}})$, where $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}$ are holomorphic[7] vector fields and the overline represents conjugation.[8]

3.1.1. *Sectional and bisectional curvature.* Aside from the full curvature tensor, there are various notions of sectional and bisectional curvature on Kähler manifolds, which are important in the study of complex differential geometry.

As a preliminary, we first recall the definition of sectional curvature for Riemannian manifolds. Let $\mathcal{X}, \mathcal{Y}$ be nonparallel tangent vectors. The sectional curvature is defined as

$$K(\mathcal{X}, \mathcal{Y}) = \frac{\mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Y}, \mathcal{X})}{g(\mathcal{X}, \mathcal{X})g(\mathcal{Y}, \mathcal{Y}) - g(\mathcal{X}, \mathcal{Y})^2}.$$

It is a classic theorem in Riemannian geometry that the sectional curvature completely determines the entire curvature tensor, which can be proven using the polarization formula and a careful application of the Bianchi identity. The sectional curvature is a fundamental concept in Riemannian geometry, and many theorems depend on either upper or lower bounds for it. Furthermore, the assumption that the sectional curvature is nonnegative greatly restricts the topology and geometry of a given Riemannian manifold.

For a Kähler manifold, there are several notions of curvature closely related to the sectional curvature. One natural type of sectional curvature on a Kähler manifold is the *holomorphic sectional curvature*. For a tangent vector $\mathcal{X} \in T\mathbb{X}$, this is defined as

$$\mathfrak{H}(\mathcal{X}) = \frac{\mathfrak{R}(\mathcal{X}, J\mathcal{X}, J\mathcal{X}, \mathcal{X})}{\|\mathcal{X}\|^4}.$$

Similarly to the sectional curvature, the holomorphic sectional curvature determines the entire curvature tensor of a Kähler manifold; see [Ballmann 2006, Proposition 4.51].

Of particular interest are Kähler manifolds whose holomorphic sectional curvature is constant c. In this case, the polarization formula can be used to show that such a Kähler manifold $\mathbb{X}$ is a Hermitian symmetric space whose curvature satisfies

$$\mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) = \frac{c}{4}\begin{pmatrix} g(\mathcal{X}, \mathcal{Z})g(\mathcal{Y}, \mathcal{W}) - g(\mathcal{X}, \mathcal{W})g(\mathcal{Y}, \mathcal{Z}) + g(\mathcal{X}, J\mathcal{Z})g(\mathcal{Y}, J\mathcal{W}) \\ -g(\mathcal{X}, J\mathcal{W})g(\mathcal{Y}, J\mathcal{Z}) + 2g(\mathcal{X}, J\mathcal{Y})g(\mathcal{Z}, J\mathcal{W}) \end{pmatrix}.$$

Spaces with constant holomorphic sectional curvature serve as the Kähler analogues of manifolds of constant sectional curvature. It is worth noting that when the complex dimension is greater than 1, a complex manifold cannot have constant sectional curvature (unless it is flat), so the definition of a complex space form is not a Kähler manifold with constant sectional curvature, but rather a Kähler manifold with

[7]In other words, vectors which satisfy $J\mathcal{X} + \sqrt{-1}\mathcal{X} = 0$ in a holomorphic coordinate chart.

[8]In fact, this shows that $\mathfrak{R}(\overline{\mathcal{X}}, \overline{\mathcal{Y}}, \mathcal{Z}, \mathcal{W})$, $\mathfrak{R}(\mathcal{X}, \overline{\mathcal{Y}}, \mathcal{Z}, \mathcal{W})$, and $\mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W})$ all vanish on a Kähler manifold.

constant holomorphic sectional curvature. The sectional curvature of such a metric (with $n > 1$) ranges between c and $\frac{1}{4}c$, so the metric is quarter-pinched from the view of Riemannian geometry.

Despite the similarity between the holomorphic sectional curvature and the sectional curvature, the former is a much more subtle invariant than the latter. For instance, nonnegative holomorphic sectional curvature does not imply nonnegative Ricci curvature, but does imply nonnegative scalar curvature [Ni and Zheng 2018].

As such, it is worthwhile to consider other curvature quantities which more directly control the geometry of a Kähler manifold. One such example is the *bisectional curvature*,[9] which was introduced in [Goldberg and Kobayashi 1967]. For two unit vectors $\mathcal{X}$ and $\mathcal{Y}$, it is defined as

$$\mathfrak{B}(\mathcal{X}, \mathcal{Y}) = \mathfrak{R}(\mathcal{X}, J\mathcal{X}, J\mathcal{Y}, \mathcal{Y}).$$

The reason that this is known as the bisectional curvature is that it satisfies the identity

$$\mathfrak{B}(\mathcal{X}, \mathcal{Y}) = \mathfrak{R}(\mathcal{X}, J\mathcal{Y}, J\mathcal{Y}, \mathcal{X}) + \mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Y}, \mathcal{X}),$$

which can be proven using the Bianchi identity.

A metric is said to have nonnegative bisectional curvature if $\mathfrak{B}(\mathcal{X}, \mathcal{Y}) \geq 0$ for all vectors $\mathcal{X}$ and $\mathcal{Y}$. Nonnegative bisectional curvature is a weaker condition than nonnegative sectional curvature (as the bisectional curvature is the sum of two sectional curvatures) but still provides very strong control over the geometry of a Kähler manifold.

There are several further curvatures of interest. For instance, it is possible to consider the bisectional curvature when it is restricted to unit tangent vectors $\mathcal{X}$ and $\mathcal{Y}$ satisfying $g(\mathcal{X}, \mathcal{Y}) = g(\mathcal{X}, J\mathcal{Y}) = 0$. This is known as the orthogonal bisectional curvature. We say that a Kähler manifold has (*NOB*) (nonnegative orthogonal bisectional curvature) if for all unit tangent vectors $\mathcal{X}$ and $\mathcal{Y}$ satisfying $g(\mathcal{X}, \mathcal{Y}) = g(\mathcal{X}, J\mathcal{Y}) = 0$ we have $\mathfrak{B}(\mathcal{X}, \mathcal{Y}) \geq 0$.

In this paper, we will need to consider a curvature tensor we call the *antibisectional curvature*. For totally real vectors[10] $\mathcal{X}$ and $\mathcal{Y}$, we define this to be

$$\mathfrak{A}(\mathcal{X}, \mathcal{Y}) = \mathfrak{R}(\mathcal{X}, J\mathcal{Y}, J\mathcal{Y}, \mathcal{X}) - \mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Y}, \mathcal{X}).$$

Similarly, we define the orthogonal antibisectional curvature (denoted by $\mathfrak{O}\mathfrak{A}$) to be the restriction of the antibisectional curvature to vectors $\mathcal{X}, \mathcal{Y}$ satisfying $g(\mathcal{X}, \mathcal{Y}) = g(\mathcal{X}, J\mathcal{Y}) = 0$. More precisely,

$$\mathfrak{O}\mathfrak{A}(\mathcal{X}, \mathcal{Y}) := \mathfrak{A}(\mathcal{X}, \mathcal{Y})|_{\{\mathcal{X}, \mathcal{Y} \,|\, g(\mathcal{X}, \mathcal{Y}) = g(\mathcal{X}, J\mathcal{Y}) = 0\}}.$$

The reason for the term "antibisectional" curvature is that $\mathfrak{A}$ and $\mathfrak{B}$ differ only in that we subtract rather than add the sectional curvatures. However, these curvatures are very different. For instance, the bisectional curvature is J-invariant, in that we can multiply either $\mathcal{X}$ or $\mathcal{Y}$ by J and get the same result.

---

[9]This is more commonly called the *holomorphic bisectional curvature*, but we will omit the "holomorphic" for the sake of exposition.

[10]In other words, vectors whose imaginary component is zero within a particular holomorphic chart.

On the other hand, $\mathfrak{A}$ changes sign if we multiply one of the vectors by J.[11] As such, it takes some care to define nonnegative orthogonal antibisectional curvature.

We say that a Kähler metric on a domain in $\mathbb{C}^n$ has (*NAB*) if, for all totally real vectors $\mathcal{X}, \mathcal{Y}$,

$$\mathfrak{A}(\mathcal{X}, \mathcal{Y}) \geq 0.$$

Similarly, we say that a Kähler metric has (*NOAB*) if, for all orthogonal totally real vectors $\mathcal{X}, \mathcal{Y}$,

$$\mathfrak{O}\mathfrak{A}(\mathcal{X}, \mathcal{Y}) \geq 0.$$

Due the previous discussion, we define nonnegativity for antibisectional curvature when restricted to totally real vectors. This definition inherently relies on a canonical decomposition of $T\mathbb{X}$ into real and imaginary vectors (i.e., an embedding into $\mathbb{C}^n$). For the spaces of interest in this paper, this can be done in a natural way. However, for more general Kähler manifolds, formulating nonnegative antibisectional curvature is less clear.

It is worth observing that if a Kähler manifold $\mathbb{X}$ has constant holomorphic sectional curvature, then the orthogonal antibisectional curvature identically vanishes. In fact, if we require that $\mathfrak{O}\mathfrak{A}(\mathcal{X}, \mathcal{Y}) \geq 0$ for all orthogonal vectors $\mathcal{X}$ and $\mathcal{Y}$, then the polarization formula shows that Hermitian symmetric spaces are the only spaces satisfying this property.[12]

3.2. ***Positively curved Kähler metrics.*** One question of considerable interest in complex geometry is to understand complete Kähler metrics with various nonnegativity properties. Most famously, Frankel conjectured that if a compact Kähler manifold has positive holomorphic bisectional curvature, it is biholomorphic to the complex projective space $\mathbb{CP}^n$. This conjecture was independently proven in [Mori 1979; Siu and Yau 1980].

For compact Kähler manifolds, it is possible to obtain this result under weaker curvature assumptions. For instance, all compact Kähler manifolds with positive orthogonal bisectional curvature are biholomorphic to $\mathbb{CP}^n$; see [Chen 2007; Feng et al. 2017; Gu and Zhang 2010]. Furthermore, all compact irreducible Kähler manifolds with nonnegative isotropic curvature are either Hermitian symmetric or else biholomorphic to $\mathbb{CP}^n$ [Seshadri 2009]. For complex surfaces, nonnegative orthogonal bisectional curvature is equivalent to nonnegative isotropic curvature[13] [Li and Ni 2019] so the previous result gives a classification of such surfaces.

For noncompact manifolds, it is natural to ask whether similar results hold. The most famous conjecture in this direction is Yau's uniformization conjecture [1994], which states that any complete irreducible noncompact Kähler metric with nonnegative bisectional curvature is biholomorphic to $\mathbb{C}^n$. Although the full conjecture is still open, Liu [2019] proved it under certain volume growth assumptions.

Although these results are not directly related to the work in this paper, they provide much of the intuition for how positive curvature of a Kähler metric imposes strong restrictions on the geometry of

[11] This was pointed out to us by Fangyang Zheng.

[12] Thanks to Fangyang Zheng for this observation.

[13] In higher dimensions, nonnegative isotropic curvature is a stronger assumption than nonnegative orthogonal bisectional curvature.

that manifold. In this spirit, by drawing a connection between the curvature of Kähler metrics and the MTW tensor in optimal transport, we hope to provide strong restrictions on cost functions which have nonnegative MTW tensor. We plan to develop this idea further in future work.

4. Hessian manifolds and the Sasaki metric

In order to interpret the MTW tensor as a complex-geometric curvature, we study the Sasaki metric, which is an almost-Hermitian metric on the tangent bundle of a Riemannian manifold. We discuss some background on this metric, focusing on the case of Hessian manifolds, in which case the Sasaki metric is Kähler.

4.1. ***The Sasaki metric on the tangent bundle.*** On a Riemannian manifold $(\mathcal{M}, g)$ with a flat[14] affine connection D, the tangent bundle naturally inherits a Hermitian structure $(T\mathcal{M}, g^D, J^D)$ [Dombrowski 1962]. The metric g^D is known as the *Sasaki metric* and the complex structure J^D is called the *canonical complex structure*. For completeness, we present a brief overview of this construction. For a more complete reference, we refer to [Satoh 2007].

Since D is flat, we can find local coordinates $\{u^i\}_{i=1}^n$ on $\mathcal{M}$ in which the Christoffel symbols of D vanish. Using these coordinates, define smooth functions $v^1, \ldots, v^n$ on the tangent bundle $T\mathcal{M}$ by $v^j(\mathcal{X}) = \mathcal{X}^j$ for a vector $\mathcal{X} = \mathcal{X}^i \partial_{u^i}$. The collection of functions $\{(u^i, v^i)\}_{i=1}^n$ then forms local coordinates for $T\mathcal{M}$. Then, for a tangent vector $\xi \in T_u\mathcal{M}$ (which we consider as a point in the tangent bundle $T\mathcal{M}$) and a tangent vector $\mathcal{X} = \mathcal{X}^i \partial_{u^i} \in T_u\mathcal{M}$, we can define vertical and horizontal lifts of $\mathcal{X}$ at ξ, denoted by $\mathcal{X}_\xi^V$ and $\mathcal{X}_\xi^H$, respectively. These are elements of $T_\xi(T\mathcal{M})$, which are defined as

$$\mathcal{X}_\xi^V = \mathcal{X}^i \partial_{v^i}, \quad \mathcal{X}_\xi^H = \mathcal{X}^i \partial_{u^i}. \tag{6}$$

This yields a decomposition of $T_\xi(T\mathcal{M})$ into horizontal and vertical subspaces, which depends on the choice of connection D:

$$T_\xi(T\mathcal{M}) = H_\xi(T\mathcal{M}) \oplus V_\xi(T\mathcal{M}).$$

As such, there is a natural identification $H_\xi(T\mathcal{M}) \cong V_\xi(T\mathcal{M}) \cong T_u\mathcal{M}$, which we use to construct the Sasaki metric [Satoh 2007, Definition 2.1].

Definition (the Sasaki metric and canonical complex structure). Let $(\mathcal{M}^n, g)$ be a Riemannian manifold with a flat affine connection D. For $\mathcal{X}, \mathcal{Y} \in T_u\mathcal{M}$ and $\xi \in T\mathcal{M}$ with $\xi = (u, v)$ in bundle coordinates, the canonical complex structure J^D is defined as

$$J^D \mathcal{X}_\xi^H = \mathcal{X}_\xi^V, \quad J^D \mathcal{X}_\xi^V = -\mathcal{X}_\xi^H.$$

Furthermore, the Sasaki metric g^D is defined as

$$\tilde{g}^D(\mathcal{X}_\xi^H, \mathcal{Y}_\xi^H) = \tilde{g}^D(\mathcal{X}_\xi^V, \mathcal{Y}_\xi^V) = g(\mathcal{X}, \mathcal{Y}), \quad \tilde{g}^D(\mathcal{X}_\xi^H, \mathcal{Y}_\xi^V) = 0.$$

[14] It is possible to define the Sasaki metric for arbitrary connections, but that will not be necessary for this paper.

This induces a Hermitian structure on $T\mathcal{M}$, which depends on both the choice of metric and flat connection on $\mathcal{M}$. To see that this is indeed a Hermitian manifold (and not merely almost Hermitian), we rely on the following result.

Theorem 3 [Dombrowski 1962]. *Let $(\mathcal{M}, g)$ be a Riemannian manifold with an affine connection D. The almost-Hermitian manifold $(T\mathcal{M}, g^D, J^D)$ satisfies the following*:

(1) *The almost-complex structure J^D is integrable whenever the connection D is flat.*

(2) *$(T\mathcal{M}, g^D, J^D)$ is Kähler if and only if D and D^* are both flat connections, which further implies that g is a Hessian metric.*

4.2. *Hessian manifolds.* We are primarily interested in the case where $T\mathcal{M}$ is Kähler, for which we must study Hessian manifolds (also known as *affine-Kähler* manifolds, due to the parallel with Kähler geometry). There are two equivalent definitions for such manifolds; with the former definition primarily used in differential geometry and the latter primarily used in information geometry.

Definition (Hessian manifold: differential-geometric). A Riemannian manifold $(\mathcal{M}, g)$ is Hessian if there is an atlas of local coordinates $\{u^i\}_{i=1}^n$ so that for each coordinate chart there is a convex potential Ψ such that

$$g_{ij} = \frac{\partial^2 \Psi}{\partial u^i\, \partial u^j}.$$

Furthermore, the transition maps between these coordinate charts are affine (i.e., $\mathcal{M}$ is an affine manifold).

Definition (Hessian manifold: information-geometric). A Riemannian manifold $(\mathcal{M}, g)$ is said to be *Hessian* if it admits dually flat connections. That is to say, it admits two flat (torsion- and curvature-free) connections D and D^* satisfying

$$\mathcal{X}(g(\mathcal{Y}, \mathcal{Z})) = g(D_{\mathcal{X}}\mathcal{Y}, \mathcal{Z}) + g(\mathcal{Y}, D^*_{\mathcal{X}}\mathcal{Z}) \tag{7}$$

for all vector fields $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$. Because of these dual flat connections, a Hessian manifold is often said to be *dually flat*.

Although these definitions initially appear different, they are actually equivalent. If we choose an atlas of coordinate charts (we will abuse notation and refer to this chart as u) in which the metric g is of Hessian form, we can induce a flat connection D by differentiation with respect to the u-coordinates. The requirement that the transition maps be affine is exactly what is necessary for this connection to be well-defined when we switch coordinates.

Before moving on, we take a moment to discuss the properties of the connection D in more detail. Firstly, by definition we have that the Christoffel symbols of D vanish in the u-coordinates. As a result, the D-geodesic equations

$$\frac{d^2u^i}{ds^2} + \Gamma^i_{jk}\frac{du^j}{ds}\frac{du^k}{ds} = 0$$

simplify to the equations

$$\frac{d^2u^i}{ds^2} = 0.$$

As a result, D-geodesics correspond to straight lines in the u-coordinates. It is worth noting that these geodesics are distinct from the geodesics with respect to the Levi-Civita connection.

We now turn our attention to the dual connection D^*. We can induce the dual connection by differentiation with respect to the so-called dual coordinates θ, which are defined as

$$\theta^i := \frac{\partial \Psi}{\partial u^i}. \tag{8}$$

It is a straightforward calculation to show that the connection induced by the u-coordinates and the connection induced by the θ-coordinates are indeed dual, in the sense of (7). For more information, see Chapter 2 of [Shima 2007]. Similarly to the situation for the primal connection, the Christoffel symbols of D^* vanish in the θ-coordinates and D^*-geodesics correspond to straight lines in the θ-coordinates. Furthermore, we can define D^*-convexity for subsets of $\mathcal{M}$ in terms of whether a subset entirely contains the D^*-geodesics between its points. This can be extended to define strict convexity as well. Strict convexity will be important in Proposition 8, so we draw attention to it here.

In the dual coordinates, the metric is also of Hessian form, where the potential is the Legendre transform Ψ^*, defined as

$$\Psi^*(\theta) = \sup_{u \in \mathcal{M}} \langle \theta, u \rangle - \Psi(u).$$

When Ψ is a convex function, Ψ^* is as well. In this case, we say that Ψ and Ψ^* are Legendre duals.[15] For further details on this correspondence, we refer the reader to [Shima 2007, Chapter 2].

There are topological and geometric obstructions for a given Riemannian manifold to admit a Hessian structure. In dimensions 4 and higher, there are local curvature obstructions as well; see [Amari and Armstrong 2014]. As all of the manifolds of interest in this paper are open domains in $\mathbb{R}^n$ (which admit a global coordinate chart), we can construct Hessian metrics simply by choosing a convex potential.

4.3. *The curvature of the Sasaki metric.* We now calculate the curvature of a Kähler Sasaki metric. To do so, we use the curvature formulas for a general Sasaki metric [Satoh 2007, Proposition 2.3] and then simplify them using the dually flat structure. Applying Satoh's Proposition 2.3 in the case where D is a flat connection, we have the following.

Proposition 4. *Let $(\mathcal{M}, g, D)$ be an affine manifold with flat connection D and Levi-Civita connection ∇. Let $\mathfrak{R}_{g^D}$ be the Riemannian curvature tensor of the Sasaki metric g^D on $T\mathcal{M}$. For vectors $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}, \xi \in T_u\mathcal{M}$,*

$$\mathfrak{R}_{g^D}(\mathcal{Z}^H_\xi, \mathcal{W}^H_\xi, X^H_\xi, \mathcal{Y}^H_\xi) = R^\nabla_g(\mathcal{Z}, \mathcal{W}, \mathcal{X}, \mathcal{Y}),$$

$$\mathfrak{R}_{g^D}(\mathcal{Z}^V_\xi, \mathcal{W}^V_\xi, \mathcal{X}^V_\xi, \mathcal{Y}^V_\xi) = -\frac{1}{4}\sum_i [(D_{e_i} g)(\mathcal{X}, \mathcal{Z})(D_{e_i} g)(\mathcal{Y}, \mathcal{W}) - (D_{e_i} g)(\mathcal{Y}, \mathcal{Z})(D_{e_i} g)(\mathcal{X}, \mathcal{W})],$$

$$\mathfrak{R}_{g^D}(\mathcal{Z}^H_\xi, \mathcal{W}^V_\xi, \mathcal{X}^V_\xi, \mathcal{Y}^V_\xi) = \mathfrak{R}_{g^D}(\mathcal{Z}^H_\xi, \mathcal{W}^V_\xi, \mathcal{X}^H_\xi, \mathcal{Y}^H_\xi) = 0,$$

$$\mathfrak{R}_{g^D}(\mathcal{Z}^H_\xi, \mathcal{W}^V_\xi, \mathcal{X}^H_\xi, \mathcal{Y}^V_\xi) = -\frac{1}{2}(D^2_{\mathcal{X}\mathcal{Z}} g)(\mathcal{Y}, \mathcal{W}) - \frac{1}{2}(D_{\gamma(\mathcal{X},\mathcal{Z})} g)(\mathcal{Y}, \mathcal{W}) + \frac{1}{4}\sum_i (D_{\mathcal{X}} g)(\mathcal{W}, e_i) \cdot (D_{\mathcal{Z}} g)(\mathcal{Y}, e_i).$$

[15] This is indeed a duality: for a convex function Ψ, $\Psi = \Psi^{**}$.

Here, $\{e_i\}$ is an orthonormal basis of $T_u\mathcal{M}$ and γ^D is the difference between D and the Levi-Civita connection on $\mathcal{M}$:

$$\gamma^D(\mathcal{X}, \mathcal{Y}) = D_{\mathcal{X}}\mathcal{Y} - \nabla_{\mathcal{X}}\mathcal{Y}.$$

When $(\mathcal{M}, g, D)$ is dually flat (i.e., Hessian), the situation simplifies further. To ease the computations, recall that we are working in coordinates $\{u^i\}_{i=1}^n$ where the Christoffel symbols of D vanish. Doing so, we find the following identities:

(1) The Riemannian metric g is given by the Hessian of a convex potential Ψ.

(2) In the induced coordinates $\{(u^i, v^i)\}_{i=1}^n$ on the tangent bundle $(T\mathcal{M}, g^D, J^D)$, the complex structure can be written as $J^D\partial_{u^i} = \partial_{v^i}$ and $J^D\partial_{v^i} = -\partial_{u^i}$. As such, the coordinate chart $(u^1, v^1, \dots, u^n, v^n)$ is biholomorphic to an open set in $\mathbb{C}^n$ under the natural identification.

(3) There are simple expressions for the Riemannian curvature and the Christoffel symbols of the Levi-Civita connection.

(a) The Riemannian curvature of the $(\mathcal{M}, g)$ is (see [Shima 2007, Proposition 3.2])

$$R_g^{\nabla}(\partial_{u^i}, \partial_{u^j}, \partial_{u^k}, \partial_{u^l}) = -\tfrac{1}{4}\Psi^{pq}(\Psi_{jlp}\Psi_{ikq} - \Psi_{ilp}\Psi_{jkq}).$$

(b) The Christoffel symbols of the Levi-Civita connection satisfy the identity

$$\Gamma_{ijk} = \tfrac{1}{2}\Psi_{ijk}, \quad \Gamma_{ji}^k = \tfrac{1}{2}\Psi_{ijm}\Psi^{km}.$$

(4) Using these formulas for the Christoffel symbols, we obtain a simple formula for $D_{\gamma^D(\mathcal{X},\mathcal{Z})}$ for two vector fields $\mathcal{X} = \mathcal{X}^i\partial_{u^i}$ and $\mathcal{Z} = \mathcal{Z}^k\partial_{u^k}$:

$$D_{\gamma^D(\mathcal{X},\mathcal{Z})} = -\mathcal{X}^i\mathcal{Z}^k\Gamma_{ik}^r D_{\partial_{u^r}} = -\mathcal{X}^i\mathcal{Z}^k\Psi_{iks}\Psi^{sr}D_{\partial_{u^r}}$$

Combining these identities with the curvature formulas for the Sasaki metric, we find the following proposition.

Proposition 5 (curvature of a Kähler Sasaki metric). *Let $(\mathcal{M}, g, D)$ be a Hessian manifold. The Riemannian curvature of the Sasaki metric on $(T\mathcal{M}, g^D, J^D)$ (in the (u, v)-coordinates defined on page 16) is*

$$\mathfrak{R}_{g^D}(\partial_{u^i}, \partial_{u^j}, \partial_{u^k}, \partial_{u^l}) = \mathfrak{R}_{g^D}(\partial_{v^i}, \partial_{v^j}, \partial_{v^k}, \partial_{v^l}) = -\tfrac{1}{4}\Psi^{rs}(\Psi_{jlr}\Psi_{iks} - \Psi_{ilr}\Psi_{jks}), \tag{9}$$

$$\mathfrak{R}_{g^D}(\partial_{u^i}, \partial_{v^j}, \partial_{u^k}, \partial_{v^l}) = -\tfrac{1}{2}\Psi_{ijkl} + \tfrac{1}{4}(\Psi_{iks}\Psi^{sr}\Psi_{jlr}) + \tfrac{1}{4}(\Psi^{sr}\Psi_{jks}\Psi_{ilr}). \tag{10}$$

Furthermore, when stated in terms of holomorphic vectors, the curvature of $T\mathcal{M}$ satisfies the identity

$$(\mathfrak{R}_{g^D})_{i\bar{j}k\bar{l}} = \mathfrak{R}_{g^D}(\partial_{u^i}, \partial_{v^j}, \partial_{u^k}, \partial_{v^l}) - \mathfrak{R}_{g^D}(\partial_{u^i}, \partial_{u^j}, \partial_{u^k}, \partial_{u^l}) = -\tfrac{1}{2}\Psi_{ijkl} + \tfrac{1}{2}\Psi^{rs}\Psi_{iks}\Psi_{jlr}. \tag{11}$$

We remark that for a Hessian manifold, Shima [2007] defined the Hessian curvature to be the negative of formula (11). We will not use this convention and instead work in terms of complex geometry.

4.4. ***The geometry when $\mathcal{M}$ is a domain.*** The previous calculation provides the curvature of $T\mathcal{M}$ for an arbitrary Hessian manifold. However, in the special case where $\mathcal{M}$ is a domain in Euclidean space and g is induced by a global potential Ψ, it is possible to construct this Kähler manifold explicitly as a subset of $\mathbb{C}^n$.

In this case, we consider the associated global coordinates $\{u\}_{i=1}^n$ as being defined on domain in $\mathbb{R}^n$. By a slight abuse of notation, we will denote the domain of the u-coordinates as $\mathcal{M}$. We can then write $z^i = u^i + \sqrt{-1}v^i$ as the standard holomorphic coordinate on $\mathbb{C}^n$, and $T\mathcal{M}$ can be identified with the domain

$$T\mathcal{M} = \{(u, v) \mid u \in \mathcal{M}, v \in \mathbb{R}^n\} \subset \mathbb{C}^n.$$

This class of domains are known as *tube domains* and have been studied in various contexts. For an introduction on these spaces, we refer the reader to page 41 of [Hörmander 1973] and for a more detailed study of their geometry, we refer the reader to [Yang 1982].

As for the associated Kähler metric, we can write the Kähler form as

$$\omega = \sqrt{-1}\Psi_{ij} dz^i \wedge d\bar{z}^j$$

and the Sasaki metric as

$$g = \begin{pmatrix} \Psi_{ij}(u) & 0 \\ 0 & \Psi_{ij}(u) \end{pmatrix}.$$

As can be seen from these formulas, the Kähler Sasaki metric is translation symmetric in its fibers (since Ψ does not depend on the fiber coordinates v). However, from (9) we can see that the fiber directions are not "flat" unless the underlying Hessian manifold $\mathcal{M}$ is Riemannian curvature free.

5. Optimal transport and complex geometry

With the background concluded, we can now state the central results of this paper, which relate the regularity theory of optimal transport to the complex geometry of the Sasaki metric.

5.1. ***The MTW tensor and the curvature of $T\mathcal{M}$.*** When $c : X \times Y \to \mathbb{R}$ is a Ψ-cost (as in the definition on page 398), Ma, Trudinger and Wang observed that the MTW tensor takes the form

$$\mathfrak{S}_{(x,y)}(\xi, \eta) = (\Psi_{ijp}\Psi_{rsq}\Psi^{pq} - \Psi_{ijrs})\Psi^{rk}\Psi^{sl}\xi^i\xi^j\eta^k\eta^l. \tag{12}$$

In this formula, k and l are summed over, despite the double superscript resulting from the vector-covector ambiguity.

To make the connection between (11) and (12) precise, we do the following. Firstly, we induce $\mathcal{M}$ with the structure of a Hessian manifold. To do so, we use Ψ as a potential for a Riemannian metric and let D be the flat connection induced by differentiation with respect to the u-coordinates. This then induces $(T\mathcal{M}, g^D, J^D)$ with a Kähler metric. Secondly, given a tangent vector $\xi \in T_x X$ and a point $y \in Y$, we induce a tangent vector (also denoted by ξ) in $T_{x-y}\mathcal{M}$ by shifting the base by y and leaving the components unchanged. We also use this same construction to induce cotangent vectors in $T^*_{x-y}\mathcal{M}$, given a point $x \in X$ and a cotangent vector $\eta \in T^*_y Y$.

After doing so, by comparing (9), (10) and (12), we obtain the following result.

Theorem 6. *Let X and Y be open sets in $\mathbb{R}^n$ and c be a Ψ-cost. Furthermore, let (ξ, η) be a vector-covector pair in $T_xX \times T_y^*Y$ such that the associated vector-covector pair on $T_{x-y}\mathcal{M} \times T_{x-y}^*\mathcal{M}$ is orthogonal. Finally, let ζ be an arbitrary vector in $T_{x-y}\mathcal{M}$.*

Then the MTW tensor satisfies the identity

$$\mathfrak{S}(\xi, \eta) = 2\mathfrak{R}_{g^D}(\xi_\zeta^H, (\eta^\sharp)_\zeta^V, \xi_\zeta^H, (\eta^\sharp)_\zeta^V) - 2\mathfrak{R}_{g^D}((\eta^\sharp)_\zeta^H, \xi_\zeta^H, (\eta^\sharp)_\zeta^H, \xi_\zeta^H). \tag{13}$$

Here, $\mathfrak{R}_{g^D}$ is the curvature of the Sasaki metric on $(T\mathcal{M}, g^D, J^D)$ after sharping η (recall that the η in the MTW tensor is a covector with $\eta(\xi) = 0$). Furthermore, the cross curvature satisfies the same identity when we allow ξ and η to be an arbitrary vector-covector pair.

Here, due to the symmetries of the Kähler Sasaki metric, the choice of ζ is arbitrary. We will discuss this fact in Section 7. Note that it is important to be careful with the indices in the previous result.[16]

Recalling our previous discussion on the curvature of Kähler metrics, if we consider $T\mathcal{M}$ as a tube domain, then the right-hand side of (13) is twice the orthogonal antibisectional curvature, which implies the following corollary.

Remark 7. The MTW tensor for a Ψ-cost is nonnegative if and only if $T\mathcal{M}$ has (NOAB) on the set $T(X - Y) \subset T\mathcal{M}$.

5.2. ***Relative c-convexity of sets and dual geodesic convexity.*** In order to establish regularity for optimal transport (as done in Theorem 2), not only is it necessary to assume that the MTW tensor is nonnegative, there are also assumptions about the relative c-convexity of the supports of μ and ν. For Ψ-costs, there is a natural geometric interpretation for this notion, which we establish here.

Proposition 8. *For a Ψ-cost, a set Y is c-convex relative to X if and only if, for all $x \in X$, the set $x - Y \subset \mathcal{M}$ is geodesically convex with respect to the dual connection D^*.*

Proof. Recall that for $x \in X$, a c-segment in Y is the curve $c\text{-}\exp_x(\ell)$ for some line segment ℓ and a set Y is c-convex relative to a set X if, for all $x \in X$, Y contains all c-segments between points in Y. For a Ψ-cost, relative c-convexity corresponds with geodesic convexity with respect to the dual connection[17] D^*.

We now apply (8) to see

$$-c_{i,} = -\Psi_i(x - y) = -\theta^i(x - y),$$

where $\theta^i(x - y)$ is the point $x - y \in \mathcal{M}$ in terms of the dual coordinates θ^i. By the definition on page 402, c-segments correspond to straight lines in the θ-coordinates. From the discussion after (8), this shows that c-segments are geodesics with respect to the dual connection D^*.

As such, a set Y contains all its c-segments if and only if, for all $x \in X$, $x - Y$ contains all of D^* geodesics, which is another way of saying that $x - Y$ is geodesically convex with respect to D^*. □

[16] A previous version of this paper [Khan and Zhang 2019] mistakenly switched the roles of j and k, leading to an incorrect claim of a correspondence of the MTW tensor to the orthogonal bisectional curvature.

[17] Recall that the dual connection D^* satisfies (7), where D is the flat connection induced by differentiation with respect to the u-coordinates.

An analogous result holds for relative c-convexity of X relative to Y. Combining the previous two results, we can restate Theorem 2 in this new language.

Theorem. *Suppose X and Y are smooth bounded domains in $\mathbb{R}^n$ and that $d\mu$ and $d\nu$ are smooth probability densities supported on X and Y, respectively, bounded away from zero and infinity on their supports. Consider a Ψ-cost for some convex function $\Psi : \mathcal{M} \to \mathbb{R}$ and suppose the following conditions hold*:

(1) *Ψ is C^4 and locally strongly convex (i.e., its Hessian is positive definite).*

(2) *For all $x \in X$, $x - Y \subset \mathcal{M}$ is strictly geodesically convex with respect to the dual connection D^*.*

(3) *For all $y \in Y$, $X - y \subset \mathcal{M}$ is strictly geodesically convex with respect to the dual connection D^*.*

(4) *The Kähler manifold $(T\mathcal{M}, g^D, J^D)$ has (NOAB) on the subset $T(X - Y)$.*

Let $\mathbb{T}_{\mathrm{U}}$ be the c-optimal transport map carrying μ to ν as in Theorem 1. Then $\mathrm{U} \in C^\infty(\overline{X})$ and $\mathbb{T}_{\mathrm{U}} : \overline{X} \to \overline{Y}$ is a smooth diffeomorphism.

We should note that for many Ψ-costs of interest, Ψ will not be uniformly strongly convex over its entire domain. This is no issue for the regularity theory, as we will restrict our attention to bounded sets X and Y, so that $X - Y$ is precompact. As such, Ψ will be strongly convex on $X - Y$.

5.3. ***Information-geometric interpretation.*** The previous results provide new interpretations of the regularity theory, but we can also use this approach to find new examples of cost functions which satisfy MTW(0). Before doing so, we will briefly discuss information geometry, which is where many of these examples originate.

Information geometry studies the geometry of parametrized statistical models. For a more complete background, we refer readers to [Amari 2016]. All of the examples in this paper are constructed from exponential families, so we will focus only on the information geometry of exponential families.

Given a sample space S (about which we make no assumptions), an exponential family is a parametrized family of probability distributions whose probability density functions are of the form

$$f_S(s \mid u) = h(s) \exp(\boldsymbol{\eta}(u) \cdot \boldsymbol{T}(s) - A(u)) \tag{14}$$

for some known functions $h : S \to \mathbb{R}$, $\boldsymbol{\eta} : U \to \mathbb{R}^n$, $\boldsymbol{T} : S \to \mathbb{R}^n$, and $A : U \to \mathbb{R}$. Here, $u \in U$ serve as parameters of the distributions, and they are generally defined on an open domain U in $\mathbb{R}^n$. Note that $\boldsymbol{\eta}$ should not be confused with the covector notation; it is instead a function of the parameters. This class of statistical models includes many commonly used parametrized families, such as the univariate normal and multinomial distributions (both of which we consider in the application section).

An exponential family is said to be in *canonical form* if $\boldsymbol{\eta}(u) = u$, in which case u is said to be the *natural parameters*. In this case, we can rewrite (14) as

$$f_S(s \mid u) = h(s) \exp(u \cdot \boldsymbol{T}(s) - \Psi(u)). \tag{15}$$

Here, $\Psi(u)$ is known as the *log-partition function*, and serves to preserve the probability normalization (i.e., $\int_S f_S(s \mid u)\, ds = 1$). When an exponential family is written in terms of its natural parameters, $\Psi(u)$ is a convex function whose domain is also convex.

5.3.1. *The Fisher metric.* Given any parametrized statistical model (not just an exponential family), there is a canonical Riemannian metric that can be induced on the parameter space. This metric, known as the *Fisher metric*, takes the form

$$g\left(\frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^j}\right) = \int_S \frac{\partial \log f(s \mid u)}{\partial u^i} \frac{\partial \log f(s \mid u)}{\partial u^j} f(s \mid u)\, ds. \tag{16}$$

There are several reasons to consider the Fisher metric as a canonical metric, and this is just one of several equivalent definitions for it. However, a more complete discussion of this topic would take us too far from the central aim of the this project. For more information, we refer the reader to the paper by Ay, Jost, Vân Lê, and Schwachhöfer [Ay et al. 2015].

For an exponential family in canonical form, the Fisher metric takes a special form. More precisely, it can be written as a Hessian metric

$$g\left(\frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^j}\right) = \frac{\partial^2}{\partial u^i \partial u^j} \Psi(u), \tag{17}$$

where $\Psi(u)$ is the log-partition function (which is guaranteed to be convex). From this, there is a natural statistical reason to consider Hessian manifolds, which we can further use to construct cost functions for optimal transport.

This paper is not the first work to consider using the log-partition function to find interesting Ψ-costs. This construction was first introduced in [Pal 2017], who developed some of the optimal transport theory for costs of this form. We thank the reviewer for bringing this paper to our attention.

5.3.2. *The case of $\mathcal{D}_\Psi^{(\alpha)}$-divergences.* Although our main results are stated in terms of Ψ-costs, they also hold (with minor modifications) for cost functions that are $\mathcal{D}_\Psi^{(\alpha)}$-divergences, which were previously studied by the second author [Zhang 2004] in the context of information geometry.

Definition ($\mathcal{D}_\Psi^{(\alpha)}$-divergence). Let $\Psi : \mathcal{M} \to \mathbb{R}$ be a convex function on a convex domain $\mathcal{M}$ in Euclidean space. For two points $x, y \in \mathcal{M}$ and $\alpha \in \mathbb{R}$, a $\mathcal{D}_\Psi^{(\alpha)}$-divergence is a function of the form

$$\mathcal{D}_\Psi^{(\alpha)}(x, y) = \frac{4}{1-\alpha^2}\left[\frac{1-\alpha}{2}\Psi(x) + \frac{1+\alpha}{2}\Psi(y) - \Psi\left(\frac{1-\alpha}{2}x + \frac{1+\alpha}{2}y\right)\right].$$

For cost functions of this form, we use Ψ to construct a Hessian metric on $\mathcal{M}$ and consider X and Y as subsets of $\mathcal{M}$. Theorem 6 relates[18] the MTW tensor of a $\mathcal{D}_\Psi^{(\alpha)}$-divergence on $\mathcal{M}$ to the orthogonal antibisectional curvature of $T\mathcal{M}$ and Proposition 8.

There are several reasons to extend our results to the case of a $\mathcal{D}_\Psi^{(\alpha)}$-divergence. Firstly, they are a natural class of divergences which interpolate between dual Bregman divergences [1967] (as α approaches ± 1).

[18]There is a scaling factor of $\frac{1}{2}(1-\alpha^2)$ between the curvature and the MTW tensor for a $\mathcal{D}_\Psi^{(\alpha)}$-divergence.

Secondly, such divergences satisfy a natural "biduality", which is important to the study of information geometry [Zhang 2004]. Thirdly, $\mathcal{D}_\Psi^{(\alpha)}$-divergences are often more natural than Ψ-costs for optimal transport on statistical manifolds for the following reason.

Both Ψ-costs and $\mathcal{D}_\Psi^{(\alpha)}$-divergences involve a convex function defined on an open domain of $\mathbb{R}^n$. When choosing to use one versus the other as a cost function, the primary difference is whether to assume that $X - Y \subset \mathcal{M}$ (as for the former), or that $X,\ Y,\ \frac{1}{2}(1-\alpha)X + \frac{1}{2}(1+\alpha)Y \subset \mathcal{M}$ (as for the latter).

For a $\mathcal{D}_\Psi^{(\alpha)}$-divergence induced by a log-partition function Ψ, the domain $\mathcal{M}$ is convex. As such, if we consider X and Y to be subsets of the natural parameters of an exponential family and let $\alpha \in (-1, 1)$, we are assured that $\frac{1}{2}(1-\alpha)X + \frac{1}{2}(1+\alpha)Y \subset \mathcal{M}$. Because of this, the $\mathcal{D}_\Psi^{(\alpha)}$-divergence is guaranteed to be well-defined. We will give an example of such a divergence function and prove the regularity for an associated optimal transport problem in Section 6.1.3.

More broadly, divergences are a generalization of distance functions, where the assumptions of symmetry and the triangle inequality are dropped. Such functions are widely used in statistics and information geometry because they can be seen as generalizations of the relative entropy. Using divergences as cost functions in order to connect information geometry with optimal transport is an active field of research [Wong and Yang 2019], and we expect that there are interesting connections yet to be found.

6. Applications

As the results in the previous section give a new interpretation for prior work, it is natural to ask for original results that can be found using this approach. In this section, we give several such applications. We will not provide the derivations of the identities in this section, as they are very involved but otherwise routine. In order to compute the associated curvature tensors, we have written a Mathematica notebook, which is available online [Khan 2018].

6.1. ***A complete, complex surface with (NOAB).*** Since the antibisectional curvature appears similar to the bisectional curvature, we can also ask how much control nonnegative antibisectional curvature or (NOAB) provides over the geometry of a Kähler manifold. Using the polarization formula [Hawley 1953], it can be shown that any metric of constant holomorphic sectional curvature has vanishing orthogonal antibisectional curvature, so any Hermitian symmetric space satisfies (NOAB). One can then ask whether there are other examples.

The following example gives a very interesting metric which satisfies (NOAB) and completeness but is neither Hermitian symmetric nor biholomorphic to $\mathbb{C}^n$.

Example 9 (a complete surface with (NOAB)). Consider the negative half-plane

$$\mathcal{M} = \mathbb{H} := \{(u^1, u^2) \mid u^2 < 0\}.$$

Prescribe a Hessian metric associated with the potential function $\Psi : \mathbb{H} \to \mathbb{R}$ given by

$$\Psi(u) = -\frac{(u^1)^2}{4u^2} - \frac{1}{2}\log(-2u^2).$$

For a vector $\xi = \partial_{u^1} + a\partial_{u^2}$ and a covector $\eta = a\,du^1 - du^2$, the associated orthogonal antibisectional curvature[19] on $T\mathbb{H}$ is given by

$$\mathfrak{OA}(\eta^\sharp, \xi) = \frac{6a^2(-a(u^1)^2 + u^2)^2}{(u^2)^2}.$$

As such, the metric has (NOAB). This metric is of independent interest, and for a more complete discussion, we refer the reader to [Molitor 2014]. We will note in passing a few of its curvature properties. For a vector $\xi = \partial_{u^1} + a\partial_{u^2}$ and a covector $\eta = du^1 + a\,du^2$, the antibisectional curvature is given by

$$\mathfrak{A}(\eta, \xi) = 2 - 12a^2 - 12\frac{au^1}{u^2} + 6\left(\frac{au^1}{u^2}\right)^2.$$

As such, the antibisectional curvature does not have a definite sign. It can similarly be shown that the orthogonal bisectional curvature also does not have a definite sign. However, the metric does have constant negative scalar curvature. This manifold is complete and Stein (it is biholomorphic to an open set in $\mathbb{C}^2$). However, it has the standard complex structure on a half-space in $\mathbb{R}^4$, so is *not* biholomorphic to $\mathbb{C}^2$.

6.1.1. *The Fisher metric of the normal distribution $\mathcal{N}(\mu, \sigma)$.* Although this example has interesting theoretical properties, it may appear to be a somewhat ad hoc construction without context. In fact, it is a natural example from information geometry. If we consider u^1 and u^2 as the natural parameters of the normal distribution with mean μ and standard deviation σ (i.e., $u^1 = \mu/\sigma^2$ and $u^2 = -1/(2\sigma^2)$), then the Riemannian metric $g_{ij} = \Psi_{ij}$ is the Fisher metric on the statistical manifold of univariate normal distributions (with unknown mean and variance). As a Riemannian manifold, $(\mathbb{H}, g)$ is a complete hyperbolic surface (which motivated our choice of notation). Note, however, that the (u^1, u^2)-coordinates do *not* induce the standard half-plane model of hyperbolic space.[20]

6.1.2. *A closely related example.* Using the normal statistical manifold, it is possible to construct another Kähler metric which satisfies (MTW). This space is actually Hermitian symmetric and was first constructed by Shima [2007, Example 6.7].

Consider the domain

$$\widetilde{\mathcal{M}} := \{(\theta^1, \theta^2) \mid \theta^2 - (\theta^1)^2 > 0\}$$

and prescribe it with a Hessian metric with potential $\Psi^*(\theta) = -\frac{1}{2} - \log(\theta^2 - (\theta^1)^2)$.

This potential arises from the parametrization of the univariate normal distribution in terms of its dual parameters $\theta^1 = \mu$ and $\theta^2 = \mu^2 + \sigma^2$ and the potential $\Psi^*(\theta)$ is the Legendre dual of the above potential in Example 9. Computing the antibisectional curvature for a vector ξ and covector η, we find that it satisfies

$$\mathfrak{A}(\xi, \eta^\sharp) = -\eta(\xi)^2.$$

As such, the orthogonal antibisectional curvature vanishes and the holomorphic sectional curvature is a negative constant. From this, we can see that the geometry of $T\widetilde{\mathcal{M}}$ is of independent interest, as it is a

[19] By computing antibisectional curvature solely on real vectors and covectors, we are slightly abusing notation. To formalize this, extend ξ and η to their real counterparts on $T\mathcal{M}$.

[20] In (μ, σ)-coordinates, the Fisher metric is $ds^2 = (1/\sigma^2)(d\mu^2 + 2d\sigma^2)$, which is much closer to the standard model.

complete Hermitian symmetric space with constant negative holomorphic sectional curvature. This is an example of a Siegel upper half-space.

We note that it is possible to construct other Kähler metrics with (NOAB) that are very similar to $T\widetilde{\mathcal{M}}$. Using a similar construction for round multivariate Gaussian distributions, it is possible to construct such a Hermitian symmetric space in arbitrary dimensions. For another example, we can consider the potential $\Psi(\theta^1, \theta^2) = -\frac{1}{2} - \log(\theta^2 - (\theta^1)^4)$, which also has (NOAB).

6.1.3. *Regularity for an associated cost function.* We can also use the potential

$$\Psi(u) = -\frac{(u^1)^2}{4u^2} - \frac{1}{2}\log(-2u^2)$$

to construct a cost function with a natural regularity theory. Due to the fact that the domain of Ψ is convex, it is more natural to consider a $\mathcal{D}_\Psi^{(\alpha)}$-divergence rather than a Ψ-cost. As such, we will consider the cost function

$$\begin{aligned} c(x, y) &= \mathcal{D}_\Psi^{(0)}(x, y) \\ &= 2\left(-\frac{(x^1)^2}{4x^2} - \frac{1}{2}\log(-2x^2)\right) + 2\left(\frac{(y^1)^2}{4y^2} + \frac{1}{2}\log(-2y^2)\right) + 4\left(\frac{(x^1+y^1)^2}{8(x^2+y^2)} + \frac{1}{2}\log(-x^2-y^2)\right). \end{aligned}$$

For this cost function, we can apply our previous calculations to obtain the following result.

Corollary. *Suppose μ and ν are probability measures supported on bounded subsets X and Y of the normal statistical manifold $\mathcal{M}$. Suppose further that the following regularity assumptions hold.*

(1) *μ and ν are absolutely continuous with respect to the Lebesgue measure. Furthermore, $d\mu$ and $d\nu$ are smooth and bounded away from zero and infinity on their respective supports.*

(2) *For all $x \in X$, $\frac{1}{2}(x + Y)$ is strictly convex with respect to the coordinates $\theta^1 = \mu$ and $\theta^2 = \mu^2 + \sigma^2$. Furthermore, the same property holds for $\frac{1}{2}(X + y)$ for all $y \in Y$.*

Let $c(x, y)$ be the cost function given by

$$c(x, y) = \mathcal{D}_\Psi^{(0)}(x, y) = 2\Psi(x) + 2\Psi(y) - 4\Psi\left(\frac{x+y}{2}\right),$$

where Ψ is the convex function given in Example 9. Then the c-optimal map $\mathbb{T}_{\mathrm{U}}$ taking μ to ν is smooth.

6.2. ***The regularity of pseudoarbitrages.*** Recently, a series of papers [Pal and Wong 2016; 2018a; 2018b; Wong 2018; 2019] has studied the problem of finding *pseudoarbitrages*, which are investment strategies which outperform the market portfolio under "mild and realistic assumptions". Their work combines information geometry with optimal transport and mathematical finance to reduce the problem to solving optimal transport problems where the cost function is given by a so-called log-divergence.

A central result in [Pal and Wong 2018a] shows that a portfolio map π outperforms the market portfolio almost surely in the long run if and only if it is a solution to the Monge problem for the cost function $c : \mathbb{R}^{n-1} \times \mathbb{R}^{n-1} \to \mathbb{R}$ given by

$$c(x, y) := \log\left(1 + \sum_{i=1}^{n-1} e^{x^i - y^i}\right) - \log(n) - \frac{1}{n}\sum_{i=1}^{n-1} x^i - y^i. \tag{18}$$

To give some context for this cost function, it is instructive to consider x and y as the natural parameters of the multinomial distribution. For natural parameters $\{x^i\}_{i=1}^{n-1}$, we can compute the probability p_i of the i-th event (in this context, the i-th market weight) using the formulas

$$p_i = \frac{e^{x^i}}{1+\sum_{j=1}^{n-1} e^{x^j}} \quad \text{for } 1 \le i < n, \tag{19}$$

$$p_n = \frac{1}{1+\sum_{j=1}^{n-1} e^{x^j}}. \tag{20}$$

To write this cost function in a more familiar form, we similarly find probabilities q_i associated to the y-parameters and fix $\boldsymbol{\pi} = (1/n, \dots, 1/n) \in \triangle^n$. Rewriting our cost in these terms,[21] we have

$$\hat{c}(p, q) := \log\left(\sum_{i=1}^{n} \boldsymbol{\pi}_i \frac{p_i}{q_i}\right) - \sum_{i=1}^{n} \boldsymbol{\pi}_i \log\left(\frac{p_i}{q_i}\right).$$

This quantity is known as the free energy in statistical physics [Pal and Wong 2018a] and by various different names in finance (such as the "diversification return", the "excess growth rate", the "rebalancing premium" and the "volatility return"). Since Pal and Wong refer to this as a *logarithmic divergence*, we refer to this cost as the logarithmic cost. This cost function is not symmetric, so is not induced by any distance function. However, Jensen's inequality shows that it is a divergence.

The main focus of Pal and Wong's work is to study the information-geometric properties of divergence functions induced by exponentially concave functions, of which $\hat{c}$ is only a single example. For any exponentially concave function, one can define a corresponding divergence which has a self-dual representation in terms of the logarithmic cost; see [Pal and Wong 2018a, Proposition 3.7]. In order to study optimal transport, we do not specify the exponentially concave function a priori. In fact, such a function induces the *solution* to an optimal transport problem.

For the logarithmic cost, only the first term affects optimal transport. As such, we instead consider the cost function

$$\tilde{c}(x, y) := \log\left(1 + \sum_{i=1}^{n-1} e^{x^i - y^i}\right).$$

This is now a Ψ-cost for the convex function

$$\Psi(u) = \log\left(1 + \sum_{i=1}^{n-1} e^{u^i}\right).$$

As such, we can apply Theorem 6 to compute the MTW tensor for the cost $\tilde{c}$. For a vector ξ and a covector η, the antibisectional curvature of $T\mathbb{R}^{n-1}$ (denoted by $\mathfrak{A}$) with Hessian metric induced by Ψ is

$$\mathfrak{A}(\xi, \eta^\sharp) = 2(g(\eta^\sharp, \xi))^2.$$

As such, the MTW tensor identically vanishes and the cost has nonnegative cross-curvature. A proof for this identity can be found in [Shima 2007, Proposition 3.9]. From the curvature formulas, we see that

[21] This is what Pal and Wong denote by $T(p \mid q)$.

this potential induces a Kähler metric on $\mathbb{C}^n$ with constant positive holomorphic sectional curvature. As such, it is a Hermitian symmetric space (although it is not complete).

In order to apply the result of [Trudinger and Wang 2009], we must also determine what relative c-convexity means in this context. To do so, we solve for the dual coordinates to the natural parameters u^i by calculating $\partial_{u^i}\Psi$ for $i = 1, \dots, n-1$. Doing so, we find that the dual coordinates are

$$\theta^i = \frac{e^{u^i}}{1+\sum_{j=1}^{n-1} e^{u^j}} \quad \text{for } 1 \le i < n,$$

which are exactly the formulas for the market weights p_i (i.e., $\theta^i = p_i$). This is initially surprising, but has a natural interpretation in terms of information geometry.

6.2.1. *The information geometry of the multinomial distribution.* It is worth discussing the geometry of this example in more detail. It turns out that if we consider the $\{x^i\}$ as natural parameters, the potential Ψ induces the Fisher metric of the multinomial distribution, which is an important exponential family in statistics. Geometrically, this is the round metric on the positive orthant of a sphere, which immediately shows that neither the underlying Hessian metric nor the Sasaki metric is complete. It is worth mentioning that this metric cannot be extended to a Kähler Sasaki metric on the tangent bundle of the entire sphere, due to the fact that the sphere is *not* an affine manifold.

For an exponential family of probability distributions, the dual coordinates are the expected values of the natural sufficient statistics. More specifically, for the multinomial distribution the dual coordinates are precisely the original market weights, which explains the relationship between the market weights and the partial derivatives of the potential function. As such, if we let $\mathcal{P}$ be the coordinate transformation from the natural parameters x to the market weights p (i.e., $\mathcal{P}(x)$ is as given in (19)), a subset $X \subset \mathbb{R}^{n-1}$ is relatively c-convex if and only if the set $\mathcal{P}(X)$ is convex as a subset of the probability simplex in the usual sense. Using this transformation, we say that a subset $\mathcal{P}(X)$ of the probability simplex has *uniform probability* if X is a precompact set. More concretely, a subset $\mathcal{P}(X)$ has uniform probability if and only if there exists $\delta > 0$ so that for all $p \in \mathcal{P}(X)$ and $1 \le i \le n,\ p_i > \delta$.

6.2.2. *Regularity of optimal transport.* From these observations and the previous identity for the MTW tensor (13), we can derive the following regularity result.

Corollary 10. *Suppose μ and ν are smooth probability measures supported respectively on subsets X and Y of the probability simplex $\triangle^n$. Suppose further that the following regularity assumptions hold*:

(1) *X and Y are smooth and strictly convex. Furthermore, both have uniform probability* (*as defined above*).

(2) *μ and ν are absolutely continuous with respect to the Lebesgue measure and $d\mu$ and $d\nu$ are bounded away from zero and infinity on their supports.*

Let $\hat{c}(p, q)$ be the cost function given by

$$\hat{c}(p,q) = \log\left(\frac{1}{n}\sum_{i=1}^{n}\frac{q_i}{p_i}\right) - \frac{1}{n}\sum_{i=1}^{n}\log\frac{q_i}{p_i}.$$

Then the $\hat{c}$-optimal map $\mathbb{T}_{\mathrm{U}}$ taking μ to ν is smooth.

Pal and Wong [2018b] study the cost function $\hat{c}$ and use it to define a displacement interpolation between two probability measures. In their paper, they inquire about the regularity problem for this interpolation. We can now answer this question using the previous result.

Corollary 11. *Suppose that μ and ν are smooth probability measures satisfying the assumptions of Corollary 10 and that $\mathbb{T}_{\mathrm{U}}$ is the $\hat{c}$-optimal map transporting μ to ν. Suppose further that $\mathbb{T}(t)\mu$ is the displacement interpolation from μ to ν induced by the one-parameter family of exponentially concave potentials $\varphi(t)$, with*

$$\varphi(t) = t\mathrm{U} + (1-t)\varphi_0, \quad \textit{where } \varphi_0 = \frac{1}{n}\sum_{i=1}^{n} \log p_i.$$

Then $\mathbb{T}_{\varphi(t)}$ is smooth, both as a map on the probability simplex for fixed t and also in terms of the t-parameter.

For $t = 1$, the solution to the interpolation problem is simply $\mathbb{T}_{\mathrm{U}}$ and so Corollary 10 shows that the potential U is smooth. Since the displacement interpolation linearly interpolates between smooth potential functions, the associated displacement interpolation is also smooth for $0 \le t \le 1$. The displacement induced by linearly interpolating the potential functions is exactly the transport considered in [Pal and Wong 2018b] (see Definition 6), and so this establishes regularity for this transport.

In closing, we note that the cost function considered here is very similar, but not identical, to the radial antennae cost, which was studied in [Wang 2004]. It is of interest to determine whether there is some deeper connection between these two costs which explains their apparent similarity.

6.3. ***Other examples in complex geometry and optimal transport.*** While writing this paper, we were able to find several more examples of Hessian manifolds whose tangent bundles have nonnegative bisectional curvature or (NOAB).

Relatively few examples of positively curved metrics are known (for some examples, see [Wu and Zheng 2011]), so this method may be helpful for finding new ones. One limitation of this approach is that many of the manifolds are not complete as metric spaces. It would be of interest to determine which convex functions induce complete Kähler metrics with nonnegative or positive orthogonal bisectional curvature, and we plan to study this problem in future work.

Each of these examples further induces a cost with nonnegative MTW tensor. Furthermore, since many of these examples are obtained from statistical manifolds, it may be possible to use them to induce meaningful statistical divergences.

(1) $\Psi(u) = -\log\big(1 - \sum_{i=1}^{n} e^{u^i}\big)$ defined on the set $\mathcal{M} = \big\{u \mid \sum_{i=1}^{n} e^{u^i} < 1\big\}$. This potential induces $T\mathcal{M}$ with a Sasaki metric of constant negative holomorphic sectional curvature:

$$\mathfrak{A}(\xi, \eta^\sharp) = -\eta(\xi)^2.$$

From an information-geometric point of view, this is the Fisher metric of the negative multinomial distribution. As a Hessian manifold, $\mathcal{M}$ is a noncompact metric of constant negative holomorphic sectional curvature, but is not complete.

(2) $\Psi(u) = (e^{u^1} + e^{u^2})^p$ for $0 < p < 1$. For a vector $\xi = \partial_{u^1} + a\partial_{u^2}$ and covector $\eta = a du^1 - du^2$, the associated orthogonal antibisectional curvature of the Sasaki metric is given by

$$\mathfrak{O}\mathfrak{A}(\xi, \eta^\sharp) = \frac{2(1/p-1)(a-1)^2(e^{u^1} + ae^{u^2})^2}{(e^{u^1} + e^{u^2})^{2+p}}.$$

This is nonnegative, and so the Kähler metric has (NOAB). As Hessian manifolds, this family of metrics is neither compact nor complete.

(3) $\Psi(u) = \log(\cosh(u^1) + \cosh(u^2))$. This potential induces a Sasaki metric on $T\mathbb{R}^2$ whose bisectional curvature is nonnegative. For a vector $\xi = \xi_1 \partial_{u^1} + \xi_2 \partial_{u^2}$ and covector $\eta = \eta_1\, du^1 + \eta_2\, du^2$, the associated bisectional curvature is given by

$$\mathfrak{B}(\xi, \eta^\sharp) = |\xi|^2 |\eta|^2 + 4\xi_1\xi_2\eta_1\eta_2,$$

where $|\xi|^2 = \xi_1^2 + \xi_2^2$ and $|\eta|^2 = \eta_1^2 + \eta_2^2$.

Furthermore, the antibisectional curvature also satisfies the same formula:

$$\mathfrak{A}(\xi, \eta^\sharp) = |\xi|^2 |\eta|^2 + 4\xi_1\xi_2\eta_1\eta_2.$$

As such, this metric has (NOAB) and nonnegative bisectional curvature. As a Hessian manifold, this metric is bounded, and so is not complete. Note that the curvature of this metric is in fact parallel with respect to D, which makes it an interesting example. We will explore this metric further in future work.

7. Open questions

7.1. ***The complex geometry of optimal maps.*** It is of interest to understand the geometry of optimal maps from the perspective of the complex geometry. Although we were able to give a complex-/information-geometric interpretation of the Ma–Trudinger–Wang conditions, we do not have a complex-geometric interpretation for Theorem 1.

It is worth comparing the situation to the pseudo-Riemannian theory of optimal transport of [Kim and McCann 2010]. One striking feature in this theory is the natural geometric interpretation for optimal maps. More precisely, if one deforms the pseudometric by a particular conformal factor (which is determined by the respective densities), the optimal map is induced by a maximal codimension-n surface with respect to the conformal pseudometric [Kim et al. 2010].

For a Ψ-cost, we hope that [Gangbo and McCann 1995] will allow us to encode optimal transport problems within $T\mathcal{M}$ so that the solution corresponds to some submanifold (or a less regular subset when the transport is discontinuous). Intuitively, this should indicate the "direction" in which the mass is transported from (X, μ) to (Y, ν). However, at present we cannot make this intuition rigorous, so we leave it for future work.

If we are able to complete the previous step, a natural follow-up question would be to try to establish the regularity theory for optimal transport with Ψ-costs in terms of complex Monge–Ampère equations. For an overview on complex Monge–Ampère equations, we refer the reader to the paper by Phong, Song and Sturm [Phong et al. 2012].

7.2. ***A potential non-Kähler generalization.*** While Ψ-costs yield many interesting examples, there are many relevant cost functions which are not of this form. As such, one natural generalization of the construction considered here is to instead consider a Lie group $\mathcal{G}$ and cost functions of the form $\Psi(x \cdot y^{-1})$ for $x, y \in \mathcal{G}$. Our work thus far can be interpreted as doing this calculation in the special case where $\mathcal{G}$ is Abelian. For non-Abelian groups, we hope it is possible to recover the MTW tensor as a curvature tensor in almost-complex geometry. In this case, there would be correction terms due to the non-Abelian nature of the group. Furthermore, the natural connections on a non-Abelian Lie group are not flat, so the associated almost-complex structure on the tangent bundle $T\mathcal{G}$ would fail to be integrable.[22]

There are several key difficulties in making this program rigorous. Firstly the curvature of an almost-Hermitian manifold is much more complicated than that of a Kähler manifold, in that it does not satisfy (5). Furthermore, there is not one, but several canonical connections to choose from, and it's not immediately clear which is the right one to use. Finally, the cut locus of a Lie group can be nontrivial, which plays an important role in the regularity theory of optimal transport on manifolds. It seems that before any of these issues can be addressed, it will be necessary to understand the optimal map in terms of complex geometry. Hopefully this will provide insight into the correct generalization in the almost-complex setting.

However, there is reason to be hopeful about this approach, as there are several examples of $\mathrm{MTW}(\kappa)$ costs induced from this construction. Most strikingly, it is known that the squared distance cost on $\mathbb{RP}^3$ with its round metric satisfies a stronger version of the MTW condition which implies regularity [Loeper and Villani 2010]. However, $\mathbb{RP}^3$ is diffeomorphic to $\mathrm{SO}(3)$ and the round metric is one example of a left-invariant Berger metric (see [Brown et al. 2007] for a more complete discussion of left-invariant metrics). As such, we hope that this approach can be used to find other examples of cost functions satisfying $\mathrm{MTW}(\kappa)$, either by considering other Lie groups or by considering other left-invariant metrics on $\mathrm{SO}(3)$.

7.3. ***Implications for optimal transport.*** The primary focus of this work is to use optimal transport theory to study complex geometry and information geometry. However, it remains an open question what can be proven about optimal transport using this approach. For instance, it is hard to find cost functions which satisfy $\mathrm{MTW}(0)$. Similarly, relatively few Kähler metrics of positive curvature are known and there are various stability and gap theorems about them (see, e.g., [Liu 2019; Ni and Niu 2019]). In optimal transport, we expect that similar results can be proven using complex geometry. We plan to explore this topic in future work.

7.4. ***The complex geometry of the antibisectional curvature.*** The orthogonal antibisectional curvature is a very subtle invariant, and its geometry is quite mysterious. It does not determine the full curvature tensor of the manifold, since all spaces with constant holomorphic sectional curvature have vanishing orthogonal antibisectional curvature. However, it is nonnegative for some important metrics which do not have any other obvious nonnegativity properties. In future work, we hope to understand this curvature more fully and to understand what sort of control it exerts over the geometry of a complex manifold.

[22] If we restrict our attention to torsion-free connections, we may be able to recover an almost-Kähler theory (see [Satoh 2007, Theorem 1.1]), but it's not clear that this is the best option.

Acknowledgments

Khan would like to thank Fangyang Zheng and Bo Guan for their helpful discussions with this project. He would also like to thank the referee for helpful suggestions for the paper. The project is supported by DARPA/ARO Grant W911NF-16-1-0383 ("Information Geometry: Geometrization of Science of Information", PI: Zhang). A preliminary announcement of these results appeared [Khan and Zhang 2019].

References

[Amari 2016] S. Amari, *Information geometry and its applications*, Applied Mathematical Sciences **194**, Springer, 2016. MR Zbl

[Amari and Armstrong 2014] S. Amari and J. Armstrong, "Curvature of Hessian manifolds", *Differential Geom. Appl.* **33** (2014), 1–12. MR Zbl

[Ay et al. 2015] N. Ay, J. Jost, H. Vân Lê, and L. Schwachhöfer, "Information geometry and sufficient statistics", *Probab. Theory Related Fields* **162**:1-2 (2015), 327–364. MR Zbl

[Ballmann 2006] W. Ballmann, *Lectures on Kähler manifolds*, European Mathematical Society (EMS), Zürich, 2006. MR Zbl

[Bregman 1967] L. M. Bregman, "A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming", *Ž. Vyčisl. Mat i Mat. Fiz.* **7** (1967), 620–631. In Russian; translated in *USSR Comput. Math. and Math. Phys.* **7**:3 (1967), 200–217. MR Zbl

[Brenier 1987] Y. Brenier, "Décomposition polaire et réarrangement monotone des champs de vecteurs", *C. R. Acad. Sci. Paris Sér. I Math.* **305**:19 (1987), 805–808. MR Zbl

[Brown et al. 2007] N. Brown, R. Finck, M. Spencer, K. Tapp, and Z. Wu, "Invariant metrics with nonnegative curvature on compact Lie groups", *Canad. Math. Bull.* **50**:1 (2007), 24–34. MR Zbl

[Caffarelli 1992] L. A. Caffarelli, "The regularity of mappings with a convex potential", *J. Amer. Math. Soc.* **5**:1 (1992), 99–104. MR Zbl

[Chen 2007] X. X. Chen, "On Kähler manifolds with positive orthogonal bisectional curvature", *Adv. Math.* **215**:2 (2007), 427–445. MR Zbl

[De Philippis and Figalli 2014] G. De Philippis and A. Figalli, "The Monge–Ampère equation and its link to optimal transportation", *Bull. Amer. Math. Soc.* (*N.S.*) **51**:4 (2014), 527–580. MR Zbl

[Dombrowski 1962] P. Dombrowski, "On the geometry of the tangent bundle", *J. Reine Angew. Math.* **210** (1962), 73–88. MR Zbl

[Feng et al. 2017] H. Feng, K. Liu, and X. Wan, "Compact Kähler manifolds with positive orthogonal bisectional curvature", *Math. Res. Lett.* **24**:3 (2017), 767–780. MR Zbl

[Figalli et al. 2011] A. Figalli, Y.-H. Kim, and R. J. McCann, "When is multidimensional screening a convex program?", *J. Econom. Theory* **146**:2 (2011), 454–478. MR Zbl

[Figalli et al. 2013] A. Figalli, Y.-H. Kim, and R. J. McCann, "Hölder continuity and injectivity of optimal maps", *Arch. Ration. Mech. Anal.* **209**:3 (2013), 747–795. MR Zbl

[Gangbo and McCann 1995] W. Gangbo and R. J. McCann, "Optimal maps in Monge's mass transport problem", *C. R. Acad. Sci. Paris Sér. I Math.* **321**:12 (1995), 1653–1658. MR Zbl

[Gangbo and McCann 1996] W. Gangbo and R. J. McCann, "The geometry of optimal transportation", *Acta Math.* **177**:2 (1996), 113–161. MR Zbl

[Goldberg and Kobayashi 1967] S. I. Goldberg and S. Kobayashi, "Holomorphic bisectional curvature", *J. Differential Geometry* **1** (1967), 225–233. MR Zbl

[Gu and Zhang 2010] H. Gu and Z. Zhang, "An extension of Mok's theorem on the generalized Frankel conjecture", *Sci. China Math.* **53**:5 (2010), 1253–1264. MR Zbl

[Hawley 1953] N. S. Hawley, "Constant holomorphic curvature", *Canad. J. Math.* **5** (1953), 53–56. MR Zbl

[Hörmander 1973] L. Hörmander, *An introduction to complex analysis in several variables*, 2nd ed., North-Holland Mathematical Library **7**, North-Holland, Amsterdam, 1973. MR Zbl

[Kantorovitch 1958] L. Kantorovitch, "On the translocation of masses", *Management Sci.* **5** (1958), 1–4. MR Zbl

[Khan 2018] G. Khan, "MTW notebook", Mathematica notebook, 2018, available at https://sites.google.com/a/umich.edu/gabekhan/code/optimaltransport.

[Khan and Zhang 2019] G. Khan and J. Zhang, "Hessian curvature and optimal transport", pp. 423–430 in *Geometric science of information*, edited by F. Nielsen and F. Barbaresco, Lecture Notes in Comput. Sci. **11712**, Springer, 2019. MR

[Kim and McCann 2010] Y.-H. Kim and R. J. McCann, "Continuity, curvature, and the general covariance of optimal transportation", *J. Eur. Math. Soc.* (*JEMS*) **12**:4 (2010), 1009–1040. MR Zbl

[Kim et al. 2010] Y.-H. Kim, R. J. McCann, and M. Warren, "Pseudo-Riemannian geometry calibrates optimal transportation", *Math. Res. Lett.* **17**:6 (2010), 1183–1197. MR Zbl

[Li and Ni 2019] X. Li and L. Ni, "Kähler–Ricci shrinkers and ancient solutions with nonnegative orthogonal bisectional curvature", preprint, 2019. arXiv

[Liu 2019] G. Liu, "On Yau's uniformization conjecture", *Camb. J. Math.* **7**:1-2 (2019), 33–70. MR Zbl

[Loeper 2009] G. Loeper, "On the regularity of solutions of optimal transportation problems", *Acta Math.* **202**:2 (2009), 241–283. MR Zbl

[Loeper and Villani 2010] G. Loeper and C. Villani, "Regularity of optimal transport in curved geometry: the nonfocal case", *Duke Math. J.* **151**:3 (2010), 431–485. MR Zbl

[Ma et al. 2005] X.-N. Ma, N. S. Trudinger, and X.-J. Wang, "Regularity of potential functions of the optimal transportation problem", *Arch. Ration. Mech. Anal.* **177**:2 (2005), 151–183. MR Zbl

[Molitor 2014] M. Molitor, "Gaussian distributions, Jacobi group, and Siegel–Jacobi space", *J. Math. Phys.* **55**:12 (2014), art. id. 122102. MR Zbl

[Monge 1781] G. Monge, "Mémoire sur la théorie des déblais et des remblais", pp. 666–704 in *Histoire de l'Académie Royale des Sciences de Paris*, De l'Imprimerie Royale, Paris, 1781.

[Mori 1979] S. Mori, "Projective manifolds with ample tangent bundles", *Ann. of Math.* (2) **110**:3 (1979), 593–606. MR Zbl

[Newlander and Nirenberg 1957] A. Newlander and L. Nirenberg, "Complex analytic coordinates in almost complex manifolds", *Ann. of Math.* (2) **65** (1957), 391–404. MR Zbl

[Ni and Niu 2019] L. Ni and Y. Niu, "Gap theorem on Kähler manifolds with nonnegative orthogonal bisectional curvature", *J. Reine Angew. Math.* (online publication April 2019).

[Ni and Zheng 2018] L. Ni and F. Zheng, "Comparison and vanishing theorems for Kähler manifolds", *Calc. Var. Partial Differential Equations* **57**:6 (2018), art. id. 151. MR Zbl

[Pal 2017] S. Pal, "Embedding optimal transports in statistical manifolds", *Indian J. Pure Appl. Math.* **48**:4 (2017), 541–550. MR Zbl

[Pal and Wong 2016] S. Pal and T.-K. L. Wong, "The geometry of relative arbitrage", *Math. Financ. Econ.* **10**:3 (2016), 263–293. MR Zbl

[Pal and Wong 2018a] S. Pal and T.-K. L. Wong, "Exponentially concave functions and a new information geometry", *Ann. Probab.* **46**:2 (2018), 1070–1113. MR Zbl

[Pal and Wong 2018b] S. Pal and T.-K. L. Wong, "Multiplicative Schrödinger problem and the Dirichlet transport", preprint, 2018. arXiv

[Peyré and Cuturi 2019] G. Peyré and M. Cuturi, "Computational optimal transport", *Found. Trends Mach. Learn.* **11**:5–6 (2019), 355–607.

[Phong et al. 2012] D. H. Phong, J. Song, and J. Sturm, "Complex Monge–Ampère equations", pp. 327–410 in *Algebra and geometry: in memory of C. C. Hsiung*, edited by H.-D. Cao and S.-T. Yau, Surv. Differ. Geom. **17**, Int. Press, Boston, 2012. MR Zbl

[Satoh 2007] H. Satoh, "Almost Hermitian structures on tangent bundles", pp. 105–118 in *Proceedings of the Eleventh International Workshop on Differential Geometry*, edited by Y. J. Suh and B. H. Kim, Kyungpook Nat. Univ., Taegu, 2007. MR Zbl

[Sei 2013] T. Sei, "A Jacobian inequality for gradient maps on the sphere and its application to directional statistics", *Comm. Statist. Theory Methods* **42**:14 (2013), 2525–2542. MR Zbl

[Seshadri 2009] H. Seshadri, "Manifolds with nonnegative isotropic curvature", *Comm. Anal. Geom.* **17**:4 (2009), 621–635. MR Zbl

[Shima 2007] H. Shima, *The geometry of Hessian structures*, World Scientific, Hackensack, NJ, 2007. MR Zbl

[Siu and Yau 1980] Y. T. Siu and S. T. Yau, "Compact Kähler manifolds of positive bisectional curvature", *Invent. Math.* **59**:2 (1980), 189–204. MR Zbl

[Trudinger and Wang 2009] N. S. Trudinger and X.-J. Wang, "On the second boundary value problem for Monge–Ampère type equations and optimal transportation", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **8**:1 (2009), 143–174. MR Zbl

[Villani 2009] C. Villani, *Optimal transport: old and new*, Grundlehren der Mathematischen Wissenschaften **338**, Springer, 2009. MR Zbl

[Wang 2004] X.-J. Wang, "On the design of a reflector antenna, II", *Calc. Var. Partial Differential Equations* **20**:3 (2004), 329–341. MR Zbl

[Wong 2018] T.-K. L. Wong, "Logarithmic divergences from optimal transport and Rényi geometry", *Inf. Geom.* **1**:1 (2018), 39–78. MR Zbl

[Wong 2019] T.-K. L. Wong, "Information geometry in portfolio theory", pp. 105–136 in *Geometric structures of information*, edited by F. Nielsen, Springer, 2019. MR Zbl

[Wong and Yang 2019] T.-K. L. Wong and J. Yang, "Optimal transport and information geometry", preprint, 2019. arXiv

[Wu and Zheng 2011] H.-H. Wu and F. Zheng, "Examples of positively curved complete Kähler manifolds", pp. 517–542 in *Geometry and analysis, Vol. 1*, edited by L. Ji, Adv. Lect. Math. (ALM) **17**, International Press, Somerville, MA, 2011. MR Zbl

[Yang 1982] P. C. Yang, "Automorphism of tube domains", *Amer. J. Math.* **104**:5 (1982), 1005–1024. MR Zbl

[Yau 1994] S.-T. Yau, "Open problems in geometry", pp. v+235 in *Lectures on differential geometry*, edited by R. Schoen and S.-T. Yau, Conference Proceedings and Lecture Notes in Geometry and Topology **1**, International Press, Cambridge, MA, 1994.

[Zhang 2004] J. Zhang, "Divergence function, duality, and convex analysis", *Neural Comput.* **16**:1 (2004), 159–195.

[Zheng 2000] F. Zheng, *Complex differential geometry*, AMS/IP Studies in Advanced Mathematics **18**, American Mathematical Society, Providence, RI, 2000. MR Zbl

Received 3 Jun 2019. Revised 17 Oct 2019. Accepted 19 Nov 2019.

GABRIEL KHAN: gabekhan@umich.edu
University of Michigan, Ann Arbor, MI, United States

JUN ZHANG: junz@umich.edu
University of Michigan, Ann Arbor, MI, United States

PURE and APPLIED ANALYSIS

msp.org/paa

Cover image: The figure shows the outgoing scattered field produced by scattering a plane wave, coming from the northwest, off of the (stylized) letters P A A. The total field satisfies the homogeneous Dirichlet condition on the boundary of the letters. It is based on a numerical computation by Mike O'Neil of the Courant Institute.

See inside back cover or msp.org/paa for submission instructions.

The subscription price for 2020 is US $505/year for the electronic version, and $565/year (+$25, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Pure and Applied Analysis (ISSN 2578-5885 electronic, 2578-5893 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

PAA peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

PURE and APPLIED ANALYSIS

vol. 2 no. 2 2020