NASA/TM-20210022137

Planting a Flag in the Tropics
The Essential Tropical Geometric Background for Networking Applications

Jacob Cleveland
Glenn Research Center, Cleveland, Ohio

Brendan Mallery
University at Albany – State University of New York, Albany, New York

Alan Hylton and Robert Short
Glenn Research Center, Cleveland, Ohio

January 2022

NASA STI Program . . . in Profile

NASA/TM-20210022137

Planting a Flag in the Tropics

The Essential Tropical Geometric Background for Networking Applications

Jacob Cleveland
Glenn Research Center, Cleveland, Ohio

Brendan Mallery
University at Albany – State University of New York, Albany, New York

Alan Hylton and Robert Short
Glenn Research Center, Cleveland, Ohio

National Aeronautics and
Space Administration

Glenn Research Center
Cleveland, Ohio 44135

January 2022

Acknowledgments

This work was supported by the NASA Nebraska Space Grant and the SUNY Albany NASA Grant during Fall 2020 and Spring 2021 as well as the NASA Internship Project and the Pathways program during Summer 2021. We would like to thank Robert Green, Robert Cardona, Justin Curry, and John Nowakowski for many helpful discussions and for providing advice and guidance while writing this paper.

Available from

NASA STI Program
Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
703-605-6000

This report is available in electronic form at http://www.sti.nasa.gov/ and http://ntrs.nasa.gov/

Planting a Flag in the Tropics
The Essential Tropical Geometric Background for Networking Applications

Jacob Cleveland
National Aeronautics and Space Administration
Cleveland, Ohio 44135

Brendan Mallery
University at Albany – State University of New York
Albany, New York 12222

Alan Hylton and Robert Short
National Aeronautics and Space Administration
Cleveland, Ohio 44135

This paper serves to guide future NASA researchers and the public at large in terms of how tropical geometry can be applied to optimization problems, namely in networking. We give a basic overview of tropical geometry through a collection of resources that we have summarized. We also describe the connections between optimal transport, power diagrams, and tropical varieties. Lastly, we include a discussion of power diagrams viewed as stratified spaces which may be a useful tool to study tropical varieties.

I. Introduction

This paper is intended to serve as a collection of ideas from tropical geometry with potential applications for NASA's mission of communicating in space. Tropical geometry is a relatively young field that has picked up a lot of interest lately - mostly thanks to the many applications it has found in scheduling, job assignment, and shortest path searches, among many others. In addition, it has been used for optimizing train schedules, which is of great interest since a delay tolerant network is in some ways analogous to a train network. We believe tropical geometry is an important avenue worthy of investigation for many of the problems we intend to solve.

II. Tropical Algebra Introduction

A. Tropical geometry

Note that we will only briefly describe the main definitions and ideas of tropical geometry in this section. The interested reader should reference [1] for a proper treatment of the subject.

We define the tropical algebra over the extended real numbers $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\}$ to be $\mathbb{T}$, with tropical addition given by $x \oplus y = \max\{x, y\}$ and tropical multiplication given by $x \otimes y = x + y$. Note that we may define an isomorphic algebra by setting $x \oplus y$ to be $\min(x, y)$ and mapping $-\infty \mapsto \infty$, and we will interchange between the min-plus and max-plus convention for different applications.

Exercise II.1. *Prove that* $\overline{\mathbb{R}}$ *with* $\oplus$ *and* $\otimes$ *forms a semiring.*

Exercise II.2. *What does it mean to take a tropical exponent?*

A tropical monomial in k-variables $x_1, x_2, ..., x_k$ can be written $x_1^{a_1} x_2^{a_2}, ..., x_k^{a_k}$, where $(a_1, a_2, ..., a_k) \in \mathbb{R}^k$. We may concisely write this as x^a, where $x = (x_1, ..., x_k)$ and $a = (a_1, a_2, ..., a_k)$. Notice that $x^a = \langle x, a\rangle$, where $\langle ., .\rangle$ denotes the dot product of (real) vectors.

A tropical polynomial in k-variables is defined for $p(\mathbf{x}) \in \mathbb{T}[x_1, x_2, ..., x_k] = \mathbb{T}[x]$ as

$$p(x) = \bigoplus_{i=1}^{n} x^{a_i} \otimes h_i,$$

where $a_i = (a_{i1}, a_{i2}, ..., a_{ik}) \in \mathbb{R}^k$ and $h_i \in \mathbb{R}$ for all i. Note that k-variate Tropical polynomial are piecewise linear functions in $\mathbb{R}^k$. Recall that a function $f : X \to \overline{\mathbb{R}}$ is convex if and only if its epigraph $\{(x, y) \in X \times \overline{\mathbb{R}} | y \geq f(x)\}$ is convex, for any nonempty set X. An example of a tropical polynomial with plots for max-plus and min-plus semirings is shown in Figure 1. In three dimensions, the surface defined by a tropical polynomial can be seen to be the upper (lower) envelope of the planes defined by each monomial of the polynomial in the min-plus (max-plus resp.) semiring. This can be seen Figures 1(a) and 1(b).

Exercise II.3. *Show that $p(x) \in \mathbb{T}(x)$ is convex as a function from $\mathbb{R}^k$ to $\overline{\mathbb{R}}$.*

B. Tropical Hypersurface

A k-variate tropical polynomial p defines a tropical hypersurface $\mathcal{T}(p)$ in $\mathbb{R}^k$ as follows: Write $p(x) = \bigoplus_{i=1}^{n} f_i(x)$, where $f_i(x)$ is a tropical monomial $x^{a_i} \otimes h_i$. Then

$$\mathcal{T}(p) = \{x \in \mathbb{R}^k | \exists i \neq j, p(x) = f_i(x) = f_j(x)\}.$$

In other words, a tropical hypersurface is (the following are equivalent):

- $x \in \mathbb{R}^k$ such that the value of $p(x)$ is realized by (at least) two different monomials;
- $x \in \mathbb{R}^k$ such that

$$\max_{1 \leq l \leq k} \{f_l(x)\} = f_i(x) = f_j(x) \text{ for } i \neq j;$$

- The subset of $\mathbb{R}^k$ where $p(x)$ fails to be linear.

The corresponding hypersurfaces for the example $p(x, y)$ in both min-plus and max-plus semirings is also shown in Figure 1. One can see that projecting down the 'corners' of the plots for $p(x)$ onto the $x - y$ plane gives the respective tropical hypersurface. These 'corners' are points on the plot where two or more planes intersect, meaning two or more monomials agreed and gave the minimum (or maximum) of all monomials. Note that $(x, y) = (3, 2)$ yields both the maximum and minimum value for $p(x, y)$ and is the only point where the tropical hypersurfaces intersect.

The complement of $\mathcal{T}(p)$ is a decomposition of $\mathbb{R}^k$ into convex polyhedral cells U_i, where (the following are equivalent):

- $p(x) = f_i(x)$ on U_i, i.e. the region where the max is realized by a single monomial.
- $\nabla p(x) = a_i$ on W_i, where ∇ is the gradient operator.

As we proceed we will sometimes switch between tropical addition and multiplication and the usual addition and multiplication on $\mathbb{R}$. We will always refer to the tropical operations as $\oplus, \otimes$ and the usual operations as $+, \times$.

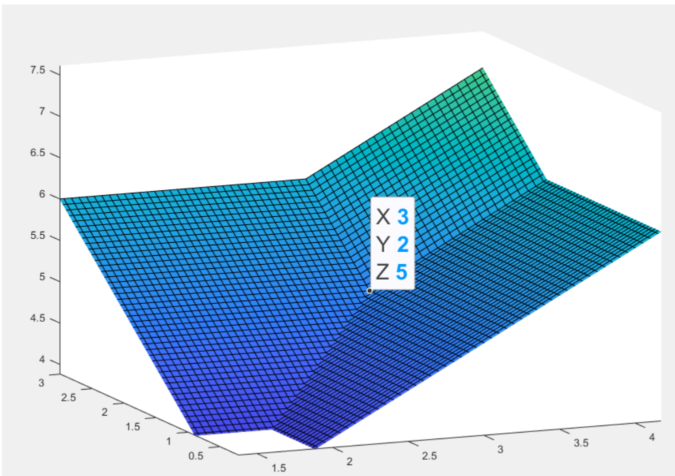
III. Tropical Matrices

In this section we will introduce and relate tropical matrix multiplication to a graph optimization problem. This is a condensed version of the highly recommended lecture notes [2].

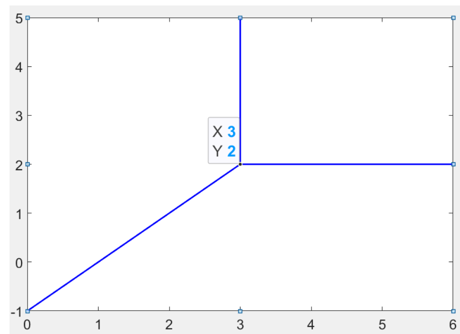
First, a few words about graphs. Given a graph $G = (V, E)$ consisting of a set of *vertices*, V, also called *nodes*, as well as a set of *edges*, E, a *walk* is a sequence of edges (equiv. vertices), while a *path* is a walk without repeated vertices (hence no repeated edges). From a graph G with $n = |V|$ vertices and no *loops* (edge from a vertex to itself), one can create the n by n adjacency matrix A_G which has a '1' in the (i, j) entry whenever there is an edge from vertex i to vertex j and a '0' otherwise.

Exercise III.1. *Given that there are no loops, what value goes in the (i, i) entry of the adjacency matrix?*

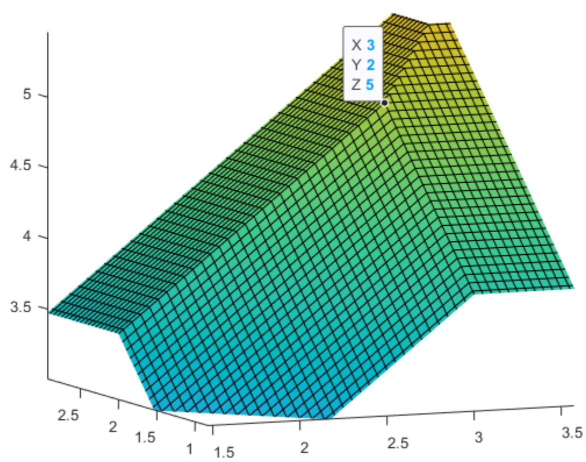
For an undirected graph, the adjacency matrix is symmetric, while for a directed graph, it is generally not symmetric.



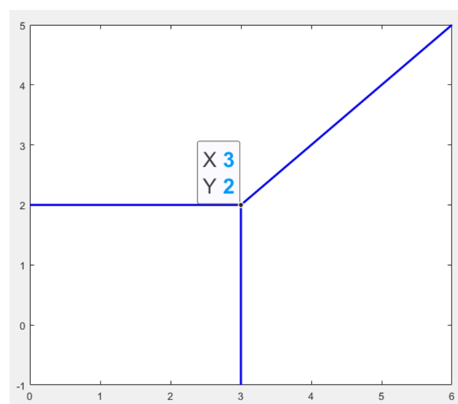
(a) Plot of $p(x, y)$ in max-plus semiring.



(b) Tropical hypersurface $\mathcal{T}(p)$ in max-plus semiring.



(c) Plot of $p(x, y)$ in min-plus semiring.



(d) Tropical hypersurface $\mathcal{T}(p)$ in min-plus semiring.

Figure 1 Plots for $p(x, y) = (x \otimes 2) \oplus (y \otimes 3) \oplus (x \otimes y)$.

Exercise III.2. *Prove the above statements regarding symmetry in the adjacency matrix.*

The adjacency matrix is a useful tool for studying graphs, as taking powers of it yield the number of walks from node i to node j with length the power the matrix was taken to. However, a more useful tool would perhaps be some operation leveraging matrices to give the total weight of the optimal walk (and hence, path, at least under the assumption that cycles have positive weight, since any walk with a cycle would induce a shorter path without the cycle). With minimal overhead, this can further give us the actual hops taken on each optimal path. Tropical matrix multiplication provides such a tool.

Now given matrices X and Y with entries in $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ we can define the tropical matrix sum $X \oplus Y$ as

$$(X \oplus Y)_{ij} = X_{ij} \oplus Y_{ij} = \min(X_{ij}, Y_{ij}).$$

That is, tropical matrix addition is just entry-wise minimum.

If X is an m by n matrix and Y is an n by l matrix, one can also define the tropical matrix product $X \otimes Y$ as

$$(X \otimes Y)_{ij} = \bigoplus_{k=1}^{n} X_{ik} \otimes Y_{kj} = \min_{1 \le k \le n} (X_{ik} + Y_{kj}).$$

That is, tropical matrix multiplication is minimizing over all pairwise sums in the ith row of X and jth column of Y.

Moreover, if X is an n by n matrix then one can define tropical matrix powers by $X^{\otimes k} = X^{\otimes k-1} \otimes X$ with $X^{\otimes 0} = \hat{I}_n$.

Define the tropical identity matrix as follows:

$$\hat{I}_n = \begin{bmatrix} 0 & \infty & \dots & \infty \\ \infty & 0 & \ddots & \\ \vdots & \ddots & \ddots & \infty \\ \infty & & \infty & 0 \end{bmatrix}$$

such that there are 0's on the diagonal and ∞ everywhere else. It is not difficult to show that for X any m by n matrix,

$$\hat{I}_m \otimes X = X = X \otimes \hat{I}_n$$

We can leverage the above definitions by fixing a weighted graph G and its weighted adjacency matrix M_G to compute shortest paths through G. Essentially, tropical matrix powers of M_G, namely the lth power, takes a path of length $l-2$ from node v_i to some node v_e and builds a path of length l from i to j by checking each intermediary node v_k and picking the optimal one in terms of total path cost; this process is illustrated by Figure 2. If there is no edge between two vertices, we consider the weight to be ∞.

Now, for the main result we wish to apply, which is that given a graph G with vertices V, edges E, and weight function $w : E \to \mathbb{R}$ such that there are no negative weight cycles, one can define the weight matrix W with entries

$$W_{k,j} = \begin{cases} 0 & \text{if j=k} \\ w(v_k, v_j) & \text{if } (v_k, v_j) \in E \\ \infty & \text{otherwise} \end{cases}.$$

Theorem III.3. *If $G(V, E, w)$ is a weighted graph (directed or undirected) on n vertices in which all cycles have positive weight, then the weight $d(v_i, v_j)$ of a minimal-weight path from v_i to v_j is given by*

$$d(v_i, v_j) = W^{\otimes(n-1)}_{i,j}$$

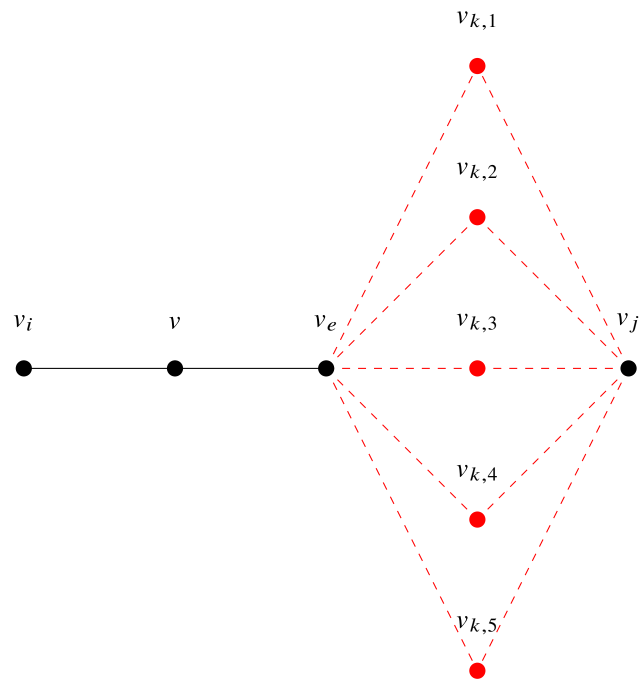


Figure 2 Graph showing the 'selection process' of tropical matrix multiplication. The black solid path shows edges that have already been selected by previous matrix powers. The red dotted lines show potential edges for selection by the l-th power of M_G. The edge weights have been omitted for simplicity.

This theorem is stated without proof since it can be found in the previously mentioned lecture notes [2].

A. Graph routing with parameterized weights

As detailed above, for a weighted graph G with n vertices with (weighted) adjacency matrix M_G, we may obtain solutions to the shortest path problem between any pair of vertices by computing $M_G^{\otimes(n-1)}$, the $(n-1)$-th tropical power of M_G. This follows from the observation that tropical matrix products provide solutions to Bellman equations of dynamical programming.

In [3], the authors consider a more general problem: suppose the weights on the edges of G are not constant, but vary according to a set of variables. What is the appropriate structure to study this parameterized optimization problem?

The main observation in [3] is that, when the edge weights are linear in the set of variables $\{x_1, x_2, ..., x_k\}$, we may replace the weights with tropical polynomials in $\mathbb{T}[x_1, ..., x_k]$. For concreteness, suppose each edge $e_{ij} = (v_i, v_j) \in G$ has weight given by a tropical indeterminate x_{ij}. Then the parameterized adjacency matrix M_G is a symmetric matrix with entry (i, j) given by x_{ij}. We may take tropical multiples of this matrix to obtain $M_G^{\otimes(n-1)} = D_G \in \mathbb{T}[\{x_{ij}\}_{i,j}]^{n\times n}$. The (i, j) entry is now a polynomial $p(x) \in \mathbb{T}[\{x_{ij}]_{i,j}$, with a monomial $x_{il_1}x_{l_1l_2}...x_{l_mj}$ for each path $x_i \to x_{j_1} \to x_{j_2} \to ... \to x_j$ in G.

Remark III.4. *For our exposition, we chose the weight on edge e_{ij} to be the (tropical) indeterminate x_{ij}. In theory, we may choose our weights to be any tropical polynomial in our set of variables. In [3], the authors place certain restrictions on their edge weights so that the matrix M_G has what they term "separated variables" (see section 2 of [3] for details).*

We can view the collection of edge weights (which for simplicity we label $x_1, x_2, ..., x_k$) as a vector $x \in \mathbb{R}^k$. The matrix D is thus a function of x, where $D(x)$ is the matrix in $\mathbb{R}^{n\times n}$ that solves the $n \times n$ path optimization problems in G with weights given by x.

Taking the product over all entries $\prod_{i,j} D_{ij}$ returns another k-variate tropical polynomial. The tropical hypersurface $\mathcal{T}(\prod_{i,j} D_{ij})$ is referred to as the *hypersurface induced by D*. This hypersurface gives a partition of parameter space $\mathbb{R}^k$ into cells, where associated to each cell is the shortest path tree that is optimal for the points within that cell. In other words, for all x

in the same cell defined by $\mathcal{T}(\prod_{i,j} D_{ij})$, the (i, j) entry of $M_G(x)$ is given by the same monomial.

This decomposition of $\mathbb{R}^k$ gives a complete solution to the parameterized shortest path problem on G. The authors of [3] then describes an algorithm that produces this decomposition, as well as the solution matrix associated to each cell. This algorithm is slower than methods for solving a specific instance of the parameterized shortest path problem, but comes with the obvious benefit of completely describing the solution space, which for instance reduces the cost of repeated computation.

IV. The equivalence of semi-discrete optimal transport, power diagrams, and tropical geometry in $\mathbb{R}^d$

In this section, we outline a connection between semi-discrete optimal transport, tropical geometry and power diagrams. This connection was observed by Na Lei in [4], though different parts of the story can be found in e.g. [5], [6], and [7]. We attempt to extend this connection in several directions. We begin by reviewing the definitions of semi-discrete optimal transport, power diagrams, and tropical geometry. This will be a fairly incomplete treatment, and we will tend to reference results rather than reproduce them.

A. Semi-discrete Optimal Transport

The optimal transport problem was first formalized by Monge[8], and can be stated as follows:

Let X and Y be metric spaces, and let $\mathcal{P}(X)$ and $\mathcal{P}(Y)$ to be the space of probability distributions on X and Y. Let $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$. We call $T : X \to Y$ a *transport map* if $T_*(\mu) = \nu$, where $T_*(\mu)$ is the pushforward distribution of μ along T. Given a cost function $c : X \times Y \to \mathbb{R}$, we define the *cost* of a transport map to be $\int_X c(x, T(x))d\mu$ The *optimal transport map* between μ and ν is T^* realizing the minimal cost of transport, i.e. the solution to

$$\min_T \{ \int_X c(x, T(x))d\mu \}.$$

When both distributions μ and ν are discrete, we refer to this as a *discrete optimal transport problem*, in which case T is a bijection. If our domain distribution is continuous and our target is discrete, we refer to this as a *semi-discrete optimal transport problem*.

Remark IV.1. *In modern optimal transport literature, the optimal transport problem is usually replaced by the Kantorovich relaxation[9]: Given $(\mu, \nu) \in \mathcal{P}(X) \times \mathcal{P}(Y)$, we let $\Pi(\mu, \nu)$ be the set of couplings of μ and ν (i.e. those distributions γ on $X \times Y$ with marginals $\pi_X(\gamma)$ and $\pi_Y(\gamma)$ equal to μ and ν, respectively. Then the optimal transport between μ and ν is the coupling satisfying*

$$W_1(\mu, \nu) = \inf_{\gamma \in \Pi(\mu, \nu)} \{ \int_{X \times Y} c(x, y)d\gamma(x, y) \}.$$

W_1 is referred to as the Wasserstein distance between μ and ν.

This has a number of benefits over the Monge reformulation. For instance, we can now "split" mass at a point to multiple targets, e.g. if μ is a Dirac distribution at x, T can only assign mass to a single point in the codomain, $T(x)$, whereas a map defined by a coupling can pair the mass at x with the mass of a distribution over e.g. a closed subset $U \subset Y$. However, in our setting the two notions will be equivalent and we will spend no more time on the distinction.

We will initially consider the optimal transport problem in $\mathbb{R}^d$, with $c(x, y)$ given by $||x - y||^2$, the L^2 norm in $\mathbb{R}^d$.

B. Power diagrams in Euclidean space

Let $\{p_i\}_{i=1}^n$ be a set of n points in $\mathbb{R}^d$, and let $W = \{w_i\}_{i=1}^n$ be a set of non-negative real numbers. We define the *power diagram* $P(\{p_i, w_i\})$ to be the decomposition of $\mathbb{R}^d$ into cells

$$U_i = \{x \in \mathbb{R}^d | ||x - p_i||^2 - w_i < ||x - p_j||^2 - w_j, \forall j \neq i\}.$$

The intuitive picture to have in mind is that the points p_i (which we will call sites) transmit signals with power proportional to w_i, and the power diagram is the decomposition of $\mathbb{R}^d$ into n cells, where the i'th cell U_i are those points which are best covered by site p_i. The cells U_i are convex polyhedra. When the weight vector is constant, i.e. $w_i = w_j \forall i, j$ we recover the *Voronoi diagram* of the points $\{p_i\}_{i=1}^n$. The cells of the power diagram arrange into a polyhedral decomposition of $\mathbb{R}^d$, meaning that if two cells U_i and U_j intersect, then they do so along a proper common face of the two, which will once again be polyhedral. Points y in these intersections are those where some set of the inequalities defining the power diagram are in fact equalities, e.g. if x is in $U_i \cap U_j$ then $||x - p_i||^2 - w_i = ||x - p_j||^2 - w_j$. In fact, the codimension of the cell which x lives gives the number of such equalities which x satisfies.

C. From semi-discrete Optimal Transport to power diagrams

The connection between semi-discrete Optimal Transport and power diagrams is essentially outlined in a theorem of Aurenhammer et al [6] (also see [7]). Before stating the theorem, we offer a slight change in perspective on power diagrams: Given a set of sites $\{(p_i\}_{i=1}^n$ and a vector of weights $W = (w_1, w_2, ..., w_n)$, we can define the power diagram $P(\{p_i, w_i)\})$ by defining an *assignment* $A_W : \mathbb{R}^d \to \{p_i\}$ that sends $y \in \mathbb{R}^d$ to p_i iff $||p_i - y||^2 - w_i \leq ||p_j - y||^2 - w_j$ for all $j \neq i$. Our use of the term assignment instead of function is intentional, as there exist points y such that $||p_i - y||^2 - w_i = ||p_j - y||^2 - w_j$ for some $i \neq j$, which is consistent with the discussion in the previous section. We now define $A_W^{-1}(p_i)$ to be the set of all y that are associated to p_i, i.e. $A_W^{-1}(p_i) = U_i$. $A_W^{-1}(p_i)$ may be viewed as a multivalued map, though we may also view this as a function from the set of sites to the set of polyhedral domains in $\mathbb{R}^d$. With this terminology in hand, we now present the theorem. We remark that, although the theorem was proved in [6], the statement above is more similar to a reproduced version in [7].

Theorem IV.2. *([6]) Let $X = \{p_i\}_{i=1}^n$ be a set of n points in $\mathbb{R}^d$. Let μ be a measure supported on a compact set $\Omega \subset \mathbb{R}^k$. Let ν_i be a set of nonnegative real numbers such that $\sum_{i=1}^n \nu_i = \mu(\Omega)$. Then there exists a weight vector W such that $\mu(A_W^{-1}(p_i)) = \nu_i$ for all $p_i \in X$. Furthermore, A_W defines an optimal transport map from the measure μ to the discrete distribution on the set $\{p_i\}_{i=1}^n$, assigning mass ν_i to p_i. The converse also holds, i.e. any assignment A_W defining a power diagram is also an optimal transport map from a measure supported on $\cup_{i=1}^n A_W^{-1}(p_i)$ to a discrete measure on $\{p_i\}_{i=1}^n$*

The above theorem shows that power diagrams give solutions to semi-discrete optimal transport between compactly supported μ to discrete $\nu = (p_i, \nu_i)$, and that conversely any power diagram with sites $\{p_i\}$ induces an optimal transport map from some compactly supported μ to some discrete distribution on $\{p_i\}$. This repeated emphasis on compactly supported distributions μ is warranted: Notice that a power diagram induces a decomposition of the entire $\mathbb{R}^d$, and hence the optimal map is only defined once a domain distribution μ on a subset of $\mathbb{R}^d$ is specified. In particular, the optimal map is the restriction of A_W^{-1} to $supp(\mu)$.

Clearly, the two important parameters (given a fixed $\mu, supp(\nu)$) that go into this map are the weight vector W and the mass vector $V = \{\nu_i\}_{i=1}^n$, which are dependent on one another. In [6] (also [5],[7]) it was shown that given a fixed V one can compute the W that induces the appropriate transportation map A_W by optimizing a certain convex function on $\mathbb{R}^d$. To define this function, we first define

$$f_{A_W}(W) = \sum_{i=1}^n \int_{A_W^{-1}(p_i)} ||x - A_W(x)||^2 - w_i d\mu$$

which is essentially integrating over the power distance $||x - y||^2 - w_i$ for the assignment induced by W. This function can be rewritten

$$f_{A_W}(W) = \sum_{i=1}^{n} \int_{A_W^{-1}(p_i)} ||x - A_W(x)||^2 d\mu - w_i \mu(A_W^{-1}(p_i)).$$

It is not hard to show that $f_{A_W}(W)$ is concave: observe that A_W is a fixed assignment coming from a power diagram and hence is the pointwise minimum of the set of assignment functions $\{f_A\}$, where A ranges over all assignments from Ω to $\{p_i\}$ with $\mu(A^{-1}(p_i)) = \mu(A_W^{-1}(p_i))$. These functions are all linear in the variables $\{w_i\}_{i=1}^n = W$, and hence f_{A_W} is the pointwise minimum of linear functions. Therefore, $f_{A_W}(W)$ is concave. We then define $E(W) = f_{A_W} + \sum_{i=1}^n \nu_i w_i$. $E(W)$ is an "energy" function, and is concave (as the sum of a concave and linear function). By inspection, it is clear that $\nabla E(w) = 0$ iff $\mu(A_W^{-1}(p_i)) = \nu_i$, and hence we can determine the W vector using gradient descent on $E(W)$. In other words, given a fixed μ, ν and an arbitrary weight vector W', we can define the *measured power diagram* $P(supp(\nu), V_{W'}, \mu, W')$, where $V_{W'}$ is the vector storing how much mass is assigned to each $p_i \in supp(\nu)$ by the map A'_W. More explicitly, let $\nu_i^{W'}$ be the i'th coordinate of $V_{W'}$. Then $\nu_i^{W'} = \mu(A_W^{-1}(p_i) \cap \Omega) = \mu(A_W^{-1}(p_i))$ (since μ is only supported on Ω). We can then use gradient descent to obtain a sequence of power diagrams $P(supp(\nu), V_{W_i}, \mu, W_i)$ that converges to $P(supp(\nu), V_W, \mu, W)$ such that $V_W = V = (\nu_1, ...\nu_n)$, the masses of the target distribution ν. It follows that $P(supp(\nu), V_W, \mu, W)$ is also the data of the optimal transport map from μ to ν.

D. From power diagrams to tropical geometry

To obtain the equivalence of power diagrams and tropical geometry, one must first observe that the condition $||p_i - x||^2 - w_i < ||p_j - x||^2 - w_j$ can be rewritten into an inequality involving linear functions (linear in y). The derivation can be found in [5].

Proposition IV.3. *The following are equivalent:*

- $||p_i - x||^2 - w_i < ||p_j - x||^2 - w_j$
- $\langle x, p_i \rangle + \frac{1}{2}(w_i - |p_i|^2) > \langle x, p_j \rangle + \frac{1}{2}(w_j - |p_j|^2)$

Notice that the second inequality can be phrased in terms of a tropical polynomial: Let $\frac{1}{2}(w_i - |p_i|^2) = h_i$. Then the second inequality can be written $x^{p_i} + h_i = x^{p_i} + h_i \oplus x^{p_j} + h_j$. Hence the system of inequalities that define a power diagram with sites $\{p_i\}_{i=1}^n$ and weights $W = (w_1, w_2, ..., w_n)$ can in fact be given by the tropical polynomial $p_h(x) = \bigoplus_{i=1}^n x^{p_i} \otimes h_i$, where we apply the transformation from W to the vector $h = (h_1, ..., h_n)$. The i'th power cell $A_W^{-1}(p_i)$ is the cell where $x^{p_i} \otimes h_i = p(x)$, and the boundaries between cells (where the power distance between two sites is equal) is nothing but the tropical hypersurface defined by $p_h(x)$. We also see that, under this correspondence $A_W(x) = \nabla p_h(x)$. The fact that the optimal transport map is given by the gradient of a convex function is an important result in optimal transport theory, and in fact holds in the continuous setting [10].

In [5] a version of Theorem 4.2 is proved. Their approach slightly differed from Aurenhammer et al. in that they started from the function $p_h(x)$ (though they did not take the tropical geometry perspective) and defined a *convex* energy function $E(h)$ that was the pointwise *maximum* of a family of linear functions, given by the monomials in $p_h(x)$. As such, the approach in [5] lets one directly optimize the masses assigned to the target distribution by the map ∇p_h in terms of h, as opposed to W (from which we can invert $W \to h$ transformation to obtain W). Apart from this, their constructions are essentially the same.

E. An Optimal Transport - Power Diagram - Tropical Geometry dictionary

We now have a fairly complete description of the equivalence between optimal transport and power diagrams, and also between power diagrams and tropical geometry. However, drawing the connection from optimal transport to tropical geometry is slightly more awkward. In particular, there is (to our knowledge) no precedent of associating a measure to the decomposition of $\mathbb{R}^k$ induced by the tropical polynomial p_h. Note that this issue is also present in the optimal transport and power diagram correspondence, but (in the authors' view) this feels more unusual in the optimal transport and tropical geometry case. Furthermore, while there is a clear transformation from the coefficients of p_h and the weights of the associated power diagram,

Table 1 Optimal Transport - Power Diagrams - Tropical Geometry Dictionary

Semi-discrete optimal transport	Power diagrams	Tropical geometry
Optimal transport from (Ω, μ) to $(supp(\nu), \nu)$	"Measured"power diagram $P(\{p_i\}_{i=1}^n, V_W, \mu, W)$	"Measured" tropical polynomial $(p_h = \bigoplus x^{p_i} \otimes h_i, \mu)$
$supp(\nu) = \{p_1, ..., p_n\}$	Sites $\{p_1, ..., p_n\}$	Tropical exponents $(p_1, p_2, ..., p_n)$
$\max\{dim(\Omega), dim(supp(\nu))\}$	dimension of the (top-dim cells) power diagram	k-variable polynomial
$\|supp(\nu)\| = n$	n sites	n-monomials
Transport map $T : \Omega \to supp(\nu)$	("Measured") assignment map $A_W(x)$	("Measured") gradient $\nabla p(x)$
"Implicit constraints" on the optimal transport problem	$W = (w_1, ..., w_n)$ $(w_i = 2h_i + \|p_i\|^2)$	$h = (h_1, ..., h_n)$ $(h_i = \frac{1}{2}(w_i - \|p_i\|^2))$
Mass at target distribution $V = (\nu_1, ...\nu_n)$	$V_W = (\mu(A_W^{-1}(p_i))_{i=1}^n$	$\mu((\nabla^{-1} p_h(p_i)))$

these quantities are only implicitly present in the optimal transport problem, and can only be determined by computation, (e.g. by variational methods as in [5]).

With this in mind, we offer a tentative dictionary summarizing the below connections (Table 1). For a similar figure, see [4]. We emphasize this dictionary is at this point merely suggestive: for instance, we refer to a "measured" tropical polynomial as a pair (p_h, μ) for a tropical polynomial p_h and a measure μ on $\mathbb{R}^k$. We take this informal perspective in order to stimulate future discussion on whether these connections can (and should) be made rigorous.

We now explicitly reiterate the aforementioned issues present in the dictionary: first, there is no natural analog of the mass of target points $V = (\nu_1, ..., \nu_n)$ in the tropical setting, apart from the (in the authors' opinion, unnatural) quantity $\mu(\nabla^{-1} p_h(p_i))$, for a measure μ on $\mathbb{R}^k$. It would be desirable to find some natural analogue of this quantity in the tropical setting (if it exists). A first step would be to determine a natural interpretation of the volume in terms of the e.g. Lebesgue measure on $\mathbb{R}^k$ of a cell in the decomposition induced by p_h. Similarly, there is no direct analogue of the tropical coefficients h or the weight vector W in the optimal transport setting. For instance, the method in 4.3 gives a method (via gradient descent of $E(W)$) for determining the weight/coefficient vector for a given target distribution ν, but it is not obvious what target distribution ν is associated to a given weight/coefficient vector.

V. Stratified Spaces

Stratified spaces have come up several times through several investigations into singular points of manifolds [11–13], which can be thought of heuristically as points of your topological space at which it fails to be a manifold. In other words, singularities occur in the localities not homeomorphic to $\mathbb{R}^n$. Whenever a space fails to be a manifold, one method of studying it is to possibly view it as a stratified space, and study the singularities within. Tropical varieties and Voronoi diagrams are such spaces.

We will briefly give two definitions for stratified spaces from [14] and [15], as well as describe their connection to Voronoi diagrams and hence tropical varieties.

A. Definition 1

First, from Gunnells [14], who describes a stratified space essentially as a topological space built up from manifolds in a 'nice way.' More precisely, a closed subset $X \subset \mathbb{R}^n$ is a stratified space when there is a poset I and a locally finite collection of disjoint locally closed subsets S_i for $i \in I$ such that:

1) $X = \bigcup S_i$
2) $S_i \cap \overline{S_j} \neq \emptyset$ if and only if $S_i \subset \overline{S_j}$, and this happens if and only if $i \leq j$. This is known as the axiom of the frontier.
3) Each S_i is a locally closed smooth submanifold of $\mathbb{R}^n$

These S_i are known as strata, hence the term stratified space. The first condition essentially states that the strata have to combine to create the whole space, the second condition stating that strata intersect when contained in the closure, and the last condition stating what type of space the strata must be.

B. Example 1

Consider space X in Figure 3 composed of a sphere S^2 and an open line segment T intersecting at point P. Colloquially one can think of this as a cup with a straw in it. Upon inspection, it does not appear that this space is a manifold. Unfortunately, there is a point at which it fails to be locally homeomorphic to euclidean space, namely P.

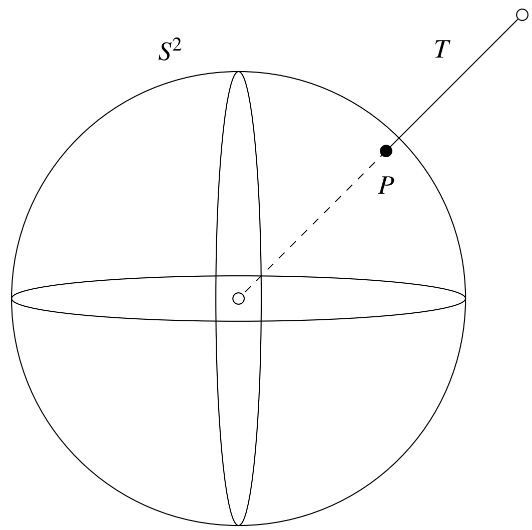


Figure 3 Example stratified space.

This is where stratified spaces come in. According to the definition above, we must produce both the partition of the space into strata as well as determine the induced poset. One possible partition is to have the following (Figure 4):

$$S_1 := S^2 - \{P\} = \{\text{punctured 2-sphere}\}$$
$$S_2 := \{\text{lower portion of the segment T without boundary}\}$$
$$S_3 := \{\text{upper portion of the segment T without boundary}\}$$
$$S_4 := \{P\}$$

We would like to see that this collection of strata constitutes a stratified space. To this end, we must first confirm that the collection of S_i is in fact a locally finite collection. According to [16], for a topological space X and a collection A of subsets of X, A is *locally finite* in X if every point of X has a neighborhood intersecting only finitely many elements of A. In our case, this is true because $A = \{S_1, S_2, S_3, S_4\}$ is only a finite collection of four subsets of X, so any neighborhood at any point is going to intersect with a finite number of elements of A. Moreover, condition 1 is seen to be satisfied since A is a partition of the space.

Now, for the axiom of the frontier, condition 2, which will determine our poset, note that $S_1 \subset \overline{S_4}$, $S_1 \subset \overline{S_2}$, and $S_1 \subset \overline{S_3}$, and that this is the only containment that occurs. This yields the poset $I = \{1, 2, 3, 4\}$ with $1 \leq 2, 1 \leq 3$, and $1 \leq 4$. Lastly, condition 3, which is left as an exercise to the reader. Note that the meaning of locally closed according to [14] is that each subset is the intersection of a closed and open subset in the ambient space.

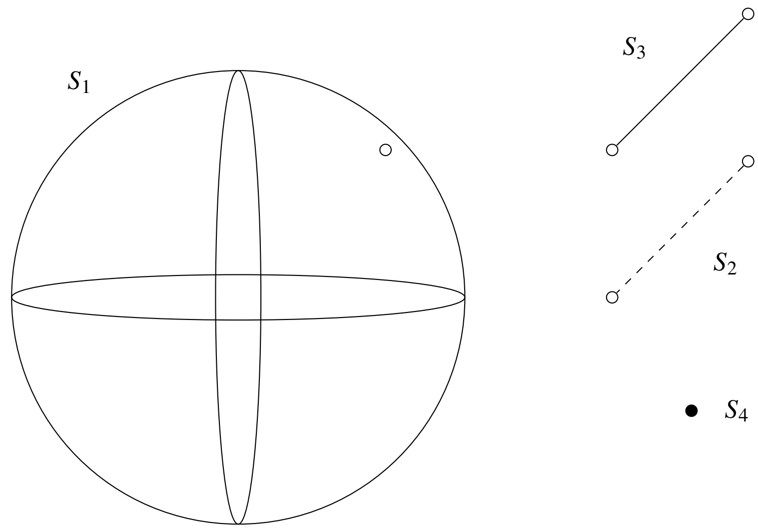


Figure 4 Example stratified space with noted strata

C. Definition 2

An alternative, perhaps not equivalent, definition of a (conically) stratified space is found in [15].

Definition: [Conically Stratified Spaces] An **n-dimensional conically stratified space** is a Hausdorff space equipped with an n-step filtration

$$\emptyset = X^{-1} \subseteq \cdots \subseteq X^n = X$$

where for each integer $d \geq 0$ and each point $x \in X^d - X^{d-1}$, there exists a distinguished open set $U_x \subseteq X$ containing x, a filtered space L_x of formal dimension $(n - d - 1)$, and a filtration-preserving homeomorphism

$$h_x : \mathbb{R}^d \times C(L_x) \rightarrow U_x$$

taking $(0, *)$ to x. We call L_x the **link** about x and the open set U_x a **basic open**. A connected component of $X^d - X^d - 1$ for any $d \geq 0$ is called a **d-dimensional stratum.** We can denote the set of all strata by S and will refer to X along with its stratification using the pair (X, S).

D. Example 1, again

To revisit the straw and cup example from figure 3, but now under this alternative definition, we must first supply the n-step filtration. One such filtration is

$$\emptyset \subseteq \{P\} \subseteq T \subseteq S^2 \cup T = X$$

To see that this is a proper stratification, let us work through one example of d value, leaving the rest for the reader. Namely, $d = 0$. In this case, our space in consideration is $X^0 - X^{-1} = \{P\} - \emptyset = \{P\}$. We can take as our basic open any neighborhood containing some of the sphere and some of both sides of the straw, leaving us with something homeomorphic to a closed disk with a closed line segment intersecting it on the interior of said segment, as seen in figure 5. It will become apparent in a moment why it is useful to think of these as closed.

With this top-like structure in mind, all that remains is to show that it is homeomorphic to

$$\mathbb{R}^0 \times C(L_x) \cong C(L_x)$$

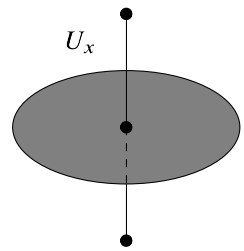


Figure 5 Basic open U_x of point P.

Exercise V.1. *We claim that for the U_x above, the link is $L_x \cong S^1 \sqcup S^0$, i.e. a circle with two points. Verify this by finding the cone of the link and showing it is homeomorphic to our basic open.*

VI. Suggestions for Future Work

A. Tropical approach to satellite routing:

It would be interesting to apply tropical approach to parameterized graph routing described in section 3.1 to routing in temporal networks, e.g. networks of orbiting satellites. We now offer one possible approach:

We identify a collection of n satellites with vertices $v_1, v_2, ..., v_n$, and we define a graph Γ by adding an edge between any pair of vertices that are able to communicate. For now we assume that this set of k edges E is fixed. We will define weight functions on the edges in our graph, where the weight gives a "score" assessing communications that use this edge. These weights are linear functions of certain state variables at each vertex. Explicitly, to each vertex v_i, we associate a set of real variables $x_{i,1}, x_{i,2}, ..., x_{i,n_i}$. These variables could represent quantities such as (Euclidean) distance to a satellite v_j, computing power, signal power, bit-rate constraints, "reliability" of the connection, current load, or any other quantity that could affect an edge's score. By repeating the construction in [3] we obtain a tropical hypersurface giving a complete description of routing over our network as a function of the state variables at each satellite.

There are a numerous directions one might wish to take this basic set-up:

- The above model assumes a fixed topology, i.e. a fixed set of edges for our network. This is not representative of satellite networks. For instance, if two satellites are initially close together but end up on opposite sides of a planet, the edge between them should disappear, as they are no longer able to directly communicate. Hence perhaps a more faithful model would be one in which the graph Γ is in fact a sequence of graphs $\{\Gamma^t\}$ indexed by time $t \in \mathbb{Z}$, with a different edge set E^t for each Γ^t. A straightforward way to produce this evolving topology is to impose a maximum communication radius on the satellites, i.e. a constant K such that if there is an edge between v_i and v_j iff $d(v_i, v_j) < K$, where d is Euclidean distance. However, to do this with the framework defined in [3], one would have to modify the functions on the edges by adding singularities, i.e. by requiring the weight become ∞ once the distance variable surpassed K. This is impossible to express as a continuous tropical polynomial. In short, the parameterized graphs considered in [3] are good for networks with fixed topology and varying weights, but not networks with varying topology and varying weights. We outline some potential modifications:
  - Perhaps the most straightforward solution would be to simply repeat the construction for each different A_{Γ^t}, i.e. construct a tropical hypersurface for each Γ^t in our sequence. For implementation purposes, this would only be feasible if the number of distinct topologies we need to consider are small, which could happen for instance if our orbits were periodic and we only consider an edge present if it exists for a long enough time period. From this we obtain a sequence of tropical curves, parameterized by time. In full generality, this is perhaps all there is to be said about the curves in this sequence. However, suppose we impose additional assumptions on the sequence Γ^t, for instance, that the topology of Γ^t and Γ^{t+1} differs by only a single edge (under mild assumptions on the trajectory of

our satellites this should always be possible). Then is there something to be said about the relationship between tropical hypersurfaces $\mathcal{T}(\Gamma^t)$ and $\mathcal{T}(\Gamma^{t+1})$? A challenge here is that the dimension of the parameter space is not constant with time, as e.g. introducing new variables associated to an edge will increase dimension.

  - Although introducing singularities to the weight equations would leave the realm of classical tropical geometry, the objects we obtain still retain many of the nice properties of tropical functions. For instance, such a "polynomial" $p(x)$ would still give a convex decomposition of our space of variables $\mathbb{R}^m$. This decomposition would have cell boundaries given by "traditional" nonlinearities obtained when $p(x)$ was equal to at least two of its "monomial terms", finitely many "bounded singularities" where the function value rapidly jumps from the value of the monomial $p_i(x)$ to the value of a different monomial $p_j(x)$ (this happens when the monomial $p_i(x) > K$ and becomes singular, i.e. $p_i(x) = \infty$, in which case the polynomial p assumes the value of finite monomial $p_j(x)$, which is optimal in a neighboring cell) and "unbounded singularities", where the value of $p(x) = \infty$, which for instance happens when $x \notin [0, K]^m$. At the bounded singularities, $p(x)$ would not be differentiable, but have left and right derivatives. It would be interesting to further explore the resulting singular tropical geometry arising from these equations, which are well suited to describe networks with changing topology.

- Each cell in the decomposition induced by $\mathcal{T}(\Gamma)$ corresponds to a region in parameter space where a fixed set of paths is optimal. As each cell $U_i \subset \mathbb{R}^k$ is open, any parameter vector $x \in U_i$ should remain in U_i under a small enough perturbation. Intuitively, the volume of the cell U_i should give some measure of how difficult it is to perturb a vector $x \in U_i$ into another cell, and hence should also give a notion of stability for the "optimality" of the set of paths associated to U_i. Can this notion of stability be used as an invariant to study weighted networks? A proper understanding of this question will no doubt also aid in answering the questions raised by the suggestive notion of "measured tropical polynomials" put forward in section 4.
- In the above set-up we assign each vertex a set of real variables that determine weight variables on the edges. Suppose we consider the variables at each vertex and edge as bases of real vector spaces of the appropriate dimension (so for instance, each edge is assigned a copy of $\mathbb{R}$). With these identifications, for any vertex v and incident edge e, we have vector spaces V_v and $V_e \cong \mathbb{R}$ and a linear map $f_{v\to e} : V_v \to V_e$ which details how the edge weight is computed from the vertex variables. This is all the data needed to define a cellular sheaf [17] over the graph Γ. Cellular sheaves have recently gained much attention in the applied topology community. With this perspective can we leverage any of the tools developed from cellular sheaves to study this problem? For instance, the most immediate thing we can do is replace the copies of $\mathbb{R}$ on edges with higher dimensional vector spaces. What object do we get if we naively repeat the above steps to produce a "tropical hypersurface" from this data?

B. Stratified spaces of power diagrams

We now outline some possible applications of stratified space theory to power diagrams and tropical geometry. First, note that tropical hypersurfaces are locally smooth (linear) manifolds apart from finitely many separated singularities, and hence may be described as stratified spaces. While this may already be enough reason to introduce stratified spaces to this document, there is a larger, more conjectural picture we would like to explore:

Let $P = \{p_1, p_2, ..., p_n\}$ be a set of points in $\mathbb{R}^n$. Then we may define a power diagram by specifying a weight vector $w \in \mathbb{R}^n$, where w_i is the power of the site p_i. We can view the power diagram as a map $h_w : \mathbb{R}^n \to P$ assigning each point in $\mathbb{R}^n$ to its associated site, i.e. $h_w(x) = p_i$ iff x is in the cell associated to the site p_i. If we vary w, the function h_w may change.

Suppose we wish to determine which weight vectors $w \in \mathbb{R}^n$ determine "equivalent" power diagram h_w. Even in the most rigid case, where $w \sim w' \in \mathbb{R}^n$ are equivalent if h_w and $h_{w'}$ are equal as functions from $\mathbb{R}^n \to J$ (i.e. induce the same power diagram), we see that $[w]$ is infinitely large:

However, suppose we consider a looser topological notion of equivalence. Clearly, if we perturb w by a small amount to

w', we get a power diagram that topologically similar to the original, in that no cells have been collapsed, and if two cells were adjacent in the diagram h_w then they are still adjacent in $h_{w'}$. A first question is: how can we formalize this notion of topological equivalence for diagrams h_w and $h_{w'}$? A promising candidate is to take the one-point compactification of $\mathbb{R}^n$ and then deleting the top dimensional cells, giving a compact, $(n-1)$-dimensional CW complex, and calling two power diagrams $h_w, h_{w'}$ equivalent if they have isomorphic complexes, and use this to define equivalence classes $[w] \subset \mathbb{R}^n$. In the case of $n = 2$, this means $w \sim w'$ iff h_w and $h_{w'}$ have boundaries given by isomorphic graphs (on the surface of S^2).

Once a good definition is determined, one may ask how the different equivalence classes $[w]$ are organized in $\mathbb{R}^n$. For instance, suppose that $h_{w'}$ is obtained from h_w by collapsing a cell and $h_{w''}$ is obtained from h_w by letting a cell ”grow” into existence (notice that there are at most J cells in the power diagram h_w, but there could be fewer). Intuitively, it would seem that any path from $[w']$ to $[w'']$ in $\mathbb{R}^n$ should in some sense "factor through" $[w]$, since the diagram h_w is a kind of intermediate between the diagrams $h_{w'}$ and $h_{w''}$.

The above discussion shows that the set of weight vectors for power diagrams on a given set of sites J with the appropriately defined topological equivalence relation bears numerous similarities to a stratified space with $[w]$ as strata: The set of equivalence classes $[w]$ forms a partition $\mathbb{R}^n$, around each $w \in \mathbb{R}^n$ there is an open neighborhood homeomorphic to $\mathbb{R}^n$ contained in the class $[w]$ and intuitively the classes $[w]$ can be organized into some kind of ”skeleton” determined by the arrangement of the sites J (witnessed by topological changes in the associated diagrams h_w). This last feature should reflect the axiom of the frontier, or some equivalent condition in one of the many other definitions of stratified spaces.

Seemingly similar ideas appear in the study of *tropical moduli spaces*, see e.g. [18]. Here, tropical curves are represented via objects called *marked, vertex-weighted, metric graphs*. These provide a coarse notion of equivalence for tropical curves based on their combinatorial structure, in much the same way we hope the (tentatively defined) compactified boundary complexes provide a coarse equivalence for power diagrams. Given the connections outlined in Section 4, this is perhaps unsurprising, and it would be interesting to see how these notions relate.

C. Tropical Arithmetic/Logic Unit

Applying the path optimization algorithm described in § 3 is theoretically pleasing due to its aesthetic simplicity, however much work is to be done to make it feasible and more applicable than existent path optimization algorithms. Work has been done investigating path optimization where the weights can change in [3] and this may be very useful for space communication applications, however this solution is very computationally intensive, both in time and space. A solution method that scales better is more desirable.

One such solution is to implement an Arithmetic/Logic Unit (ALU) specifically optimized to perform tropical operations on the max plus semiring. This allows the tropical operations to be performed much quicker than on a standard arithmetic ALU. The hope is that building these on an Field-Programmable Gate Array (FPGA) will give the scaling required by arbitrary matrix calculations as encountered above.

This work would be approached from two separate directions: implementing a tropical ALU in hardware, such as on an FPGA, as well as a way to distribute these ALU's in such a way to perform matrix multiplication in parallel. The other direction, which is mostly software, would be developing a package linking Satellite Orbit Analysis Program (SOAP) data to the hardware by parsing the data into matrices, then transferring these matrices accross to the hardware and combining the results to be used.

References

[1] Maclagan, D., and Sturmfels, B., *Introduction to Tropical Geometry*, 1st ed., Graduate Studies in Mathematics, Vol. 161, American Mathematical Society, 2015.

[2] Muldoon, M., “Tropical Arithmetic and Shortest Paths,” , 2020. URL `https://personalpages.manchester.ac.uk/staff/mark.muldoon/Teaching/DiscreteMaths/LectureNotes/TropicalShortestPaths.pdf`.

[3] Joswig, M., and Schröter, B., “Parametric Shortest-Path Algorithms Via Tropical Geometry,” , Nov 2020.

[4] Lei, N., “Tropical Geometry, OMT and Neural Network,” *Geometric Analysis Approach to AI*, Vol. CMSA, Harvard, 2019, 2019.

[5] Gu, X., Luo, F., Sun, J., and Yau, S.-T., “Variational principles for Minkowski type problems, discrete optimal transport, and discrete Monge–Ampère equations,” *Asian Journal of Mathematics*, 2016. https://doi.org/10.4310/ajm.2016.v20.n2.a7.

[6] Aurenhammer, F., Hoffmann, F., and Aronov, B., “Minkowski-Type Theorems and Least-Squares Clustering,” *Algorithmica*, Vol. 20, No. 1, 1998. https://doi.org/10.1007/pl00009187.

[7] Lévy, B., “A Numerical Algorithm for L_2 Semi-Discrete Optimal Transport in 3D,” *ESAIM: Mathematical Modelling and Numerical Analysis*, Vol. 49, No. 6, 2015. https://doi.org/10.1051/m2an/2015055.

[8] G.Monge, “Mémoire sur la théorie des déblais et des remblais,” *Histoire de l’Académie Royale des Sciences de Paris*, Vol. 1, No. 1, 1781, p. 666–704.

[9] Villani, C., *Topics in Optimal Transportation*, AMS, 2003.

[10] Brenier, Y., “Polar factorization and monotone rearrangement of vector-valued functions,” *Communications on Pure and Applied Mathematics*, Vol. 44, No. 4, 1991, p. 375–417. https://doi.org/10.1002/cpa.3160440402.

[11] Bendich, P., Cohen-Steiner, D., Edelsbrunner, H., Harer, J., and Morozov, D., “Inferring local homology from sampled stratified spaces,” 2007.

[12] Pflaum, M., *Analytic and geometric study of stratified spaces: contributions to analytic and geometric aspects*, Springer Science & Business Media, 2001.

[13] Weinberger, S., *The Topological Classification of Stratified Spaces*, The University of Chicago Press, 1994.

[14] Gunnells, P. E., “Stratified Spaces Twigs,” , 2003.

[15] Curry, J., and Patel, A., “Classification of Constructible Cosheaves,” *Theory and Applications of Categories*, Vol. 35, No. 27, 2020, pp. 1012–1047.

[16] Munkres, J., *Topology*, 2^{nd} ed., Prentice Hall, INC, 2000.

[17] Curry, J., “Sheaves, Cosheaves and Applications,” , Dec 2014. URL https://arxiv.org/abs/1303.3255.

[18] Chan, M., “Lectures on tropical curves and their moduli spaces,” , Feb 2017. URL https://arxiv.org/abs/1606.02778.