ON THE CHARACTERIZATION AND UNIQUENESS OF CENTROIDAL VORONOI TESSELLATIONS

JOHN C. URSCHEL *

Abstract. Vector quantization is a classical signal-processing technique with significant applications in data compression, pattern recognition, clustering, and data stream mining. It is well known that for critical points of the quantization energy, the tessellation of the domain is a centroidal Voronoi tessellation. However, for dimensions greater than one, rigorously verifying a given centroidal Voronoi tessellation is a local minimum can prove difficult. Using variational techniques, we give a full characterization of the second variation of a centroidal Voronoi tessellation and give sufficient conditions for a centroidal Voronoi tessellation to be a local minimum. In addition, the conditions under which a centroidal Voronoi tessellation for a given density and domain is unique have been elusive for dimensions greater than one. We prove that there does not exist a unique two generator centroidal Voronoi tessellation for dimensions greater than one.

Key words. centroidal Voronoi tessellation, vector quantization, data compression, clustering

AMS subject classifications. 52B55, 62H30, 65C50, 68U05

1. Preliminaries. Consider the vector space $\mathbb{R}^N$ over the field $\mathbb{R}$, with inner product $\langle x, y\rangle := x^T y$ and induced norm $\|x\| := \langle x, x\rangle^{1/2}$. Let $p(x)$ be a positive real-valued C^2 function, $p : \mathbb{R}^N \to \mathbb{R}$, with compact and convex support Ω. We call $p(x)$ the *density function* associated with Ω. Let μ be the measure associated with $p(x)$, namely

$$\mu(A) = \int_A p(x)dx.$$

We consider the problem of approximating $p(x)$ by a piecewise-constant function taking only K many values, namely $\hat{p} : \Omega \to \{p(u_1), ..., p(u_K)\}$, where $\{u_1, ..., u_K\} \subset \Omega$ and

$$\hat{p}(x) = p(u_i) \quad \text{if } x \in \Omega_i,$$

$$\Omega_i = \{x \in \Omega | \|x - u_i\| \leq \|x - u_j\| \text{ for all } j\}.$$

The process of approximating a function by its value at finitely many points is known as *quantization*. Quantization is a classical signal-processing technique with applications in data compression, pattern recognition, clustering, and data stream mining [9]. For $N = 1$ and $N > 1$, it is referred to as *scalar quantization* and *vector quantization*, respectively.

We have $\Omega = \cup_{i=1}^K \Omega_i$. Let us define $\lambda_N(A)$ as the N-dimensional Lebesgue measure of the set $A \subset \mathbb{R}^N$. We have $\lambda_N(\Omega_i \cap \Omega_j) = 0$ for all $i \neq j$. If $\Omega_i \cap \Omega_j \neq \emptyset$, we define $\sigma_{i,j}$ as the unique element of $\Omega_i \cap \Omega_j$ that is collinear with $\{u_i, u_j\}$, namely

$$\sigma_{i,j} = \frac{u_i + u_j}{2}.$$

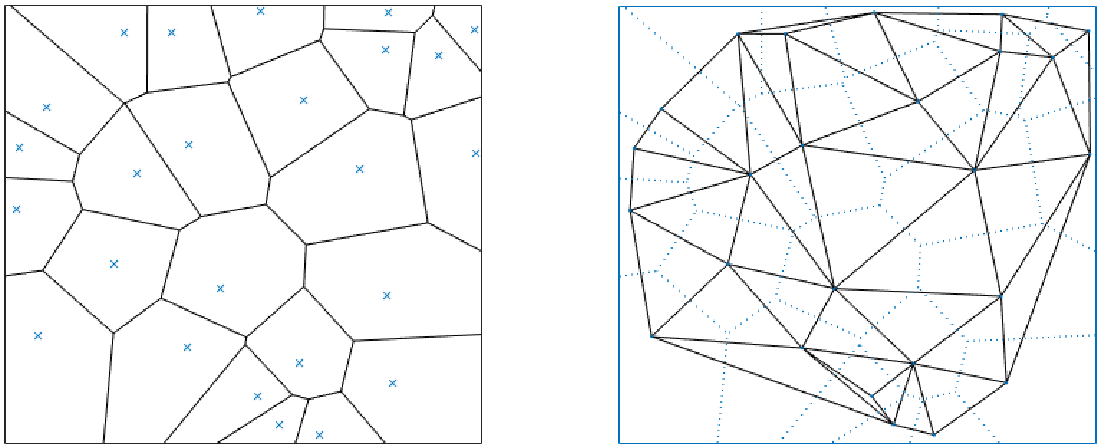
We note that the set $\{\Omega_i\}_{i=1}^K$ defines a tessellation of Ω and is referred to as the *Voronoi tessellation* associated with the generating points $\{u_i\}_{i=1}^K$. The region Ω_i is referred to as the *Voronoi region* associated with the generator u_i. Each of these regions Ω_i is convex. The dual to a Voronoi tessellation defines a Delaunay triangulation [1]. See Figure 1.

The connectivity of the Voronoi regions $\{\Omega_i\}_{i=1}^K$ gives rise to a graphical structure. We define the *Voronoi graph* $G = (V, E)$ as the graph with vertices $i \in V$, $i = 1, ..., K$ and edges $e_{i,j} \in E$ if $\Omega_i \cap \Omega_j \neq \emptyset$.

If the Voronoi generators satisfy the property

$$u_i = \mu(\Omega_i)^{-1} \int_{\Omega_i} xp(x)dx,$$

*Baltimore Ravens, NFL, Owings Mills, MD 21117; Department of Mathematics, MIT, Cambridge, MA 02138; and Department of Mathematics, Penn State, University Park, PA 16801; Email: jcurschel@gmail.com



(a) Voronoi tessellation

(b) Delaunay triangulation

FIG. 1. *A Voronoi tessellation of the square and the dual Delaunay triangulation.*

then the tessellation is a *centroidal Voronoi tessellation*.

We define the error of the approximating function $\hat{p}$, uniquely defined by the generators $\{u_i\}_{i=1}^K$, by

$$J[u] = \sum_{i \in V} \int_{\Omega_i} \|x - u_i\|^2 p(x) dx.$$

This is also often referred to as the quantization energy or energy of the tessellation. Any set of generators that minimizes the functional $J[u]$ must necessarily generate a centroidal Voronoi tessellation. We define the quantization error corresponding to a particular generator u_i by

$$J_i[u] = \int_{\Omega_i} \|x - u_i\|^2 p(x) dx.$$

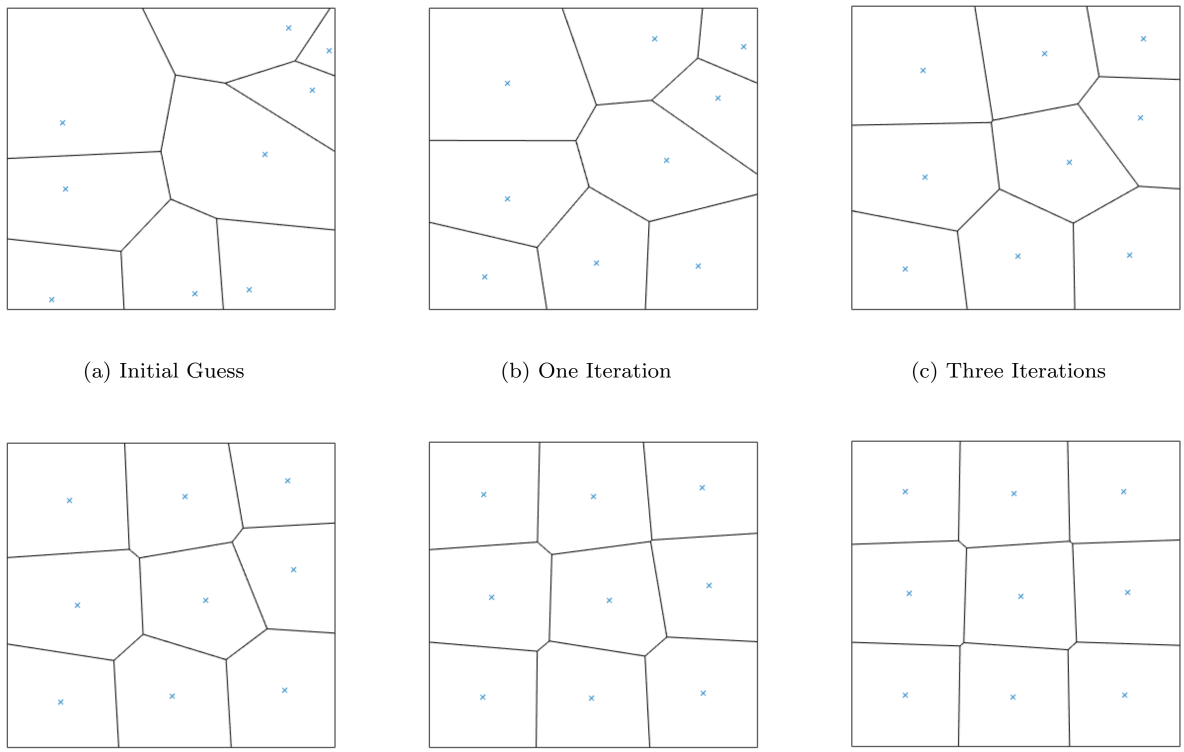
While the support Ω is compact, the domain of $\{u_i\}_{i=1}^K$ is not. The set of valid choices for $\{u_i\}_{i=1}^K$, denoted by $\mathcal{U} \subset \Omega^K$, does not contain points for which $u_i = u_j$ for some $i \neq j$. Such points are called degenerate and do not define a K-quantizer. Analysis of such points is not needed; the energy of a degenerate point is greater than some local K-quantizer in an ϵ-neighborhood for any ϵ greater than zero [2, 5]. For this reason, we restrict ourselves to

$$\mathcal{U} = \{u | u_i \in \Omega, i = 1, .., k, u_i \neq u_j, \text{ for } i \neq j\}.$$

The study of quantizers and centroidal Voronoi tessellations is closely related to Lloyd's algorithm [12], described as follows:

Lloyd's Algorithm:
1. Choose initial set of generators $\{u_i\}_{i=1}^K$ on Ω.
2. Construct associated Voronoi regions $\{\Omega_i\}_{i=1}^K$ of Ω.
3. Compute centroids of each Voronoi region, $\{\tilde{u}_i\}_{i=1}^K$, $\tilde{u}_i = \mu(\Omega_i)^{-1} \int_{\Omega_i} xp(x)dx$.
4. If $\{u_i\}_{i=1}^K$ and $\{\tilde{u}_i\}_{i=1}^K$ meet the given convergence criterion, return $\{\tilde{u}_i\}_{i=1}^K$ and terminate; otherwise let $\{u_i\}_{i=1}^K := \{\tilde{u}_i\}_{i=1}^K$ and repeat steps 2 and 3.

The algorithm is intended to converge to a centroidal Voronoi tessellation minimizing $J[u]$ over all acceptable choices of u. See Figure 2. Centralized Voronoi tessellations and their relation to the convergence



(a) Initial Guess (b) One Iteration (c) Three Iterations

(d) Five Iterations (e) Ten Iterations (f) Twenty-Five Iterations

Fig. 2. *Iterations of Lloyds algorithm on the square. The tessellations are approaching the globally minimizing centroidal Voronoi tessellation.*

of Lloyd's algorithm have been studied extensively. For excellent and thorough reviews of the literature, we refer the reader to [3, 4].

Initially in [7], and later in [11], it was shown that Lloyd's algorithm is a local contraction for $N = 1$, given that the density function is logarithmically concave, namely

$$\frac{d^2}{dx^2} \ln p(x) < 0 \text{ for all } x \in \Omega.$$

In [15], this result was extended to continuous and positive densities. Similar convergence results are tougher for $N > 1$, but not intractable. Convergence was shown in [10, 14] by defining the Lloyd mapping for degenerate points. Most notably, in [2] global convergence of Lloyd's algorithm was shown under a number of different conditions, one of the largest recent results in the field:

Theorem 1.1. *If any one of the following conditions occur*

(i) there is a unique fixed point,

(ii) the set of fixed points with any particular distortion value is finite,

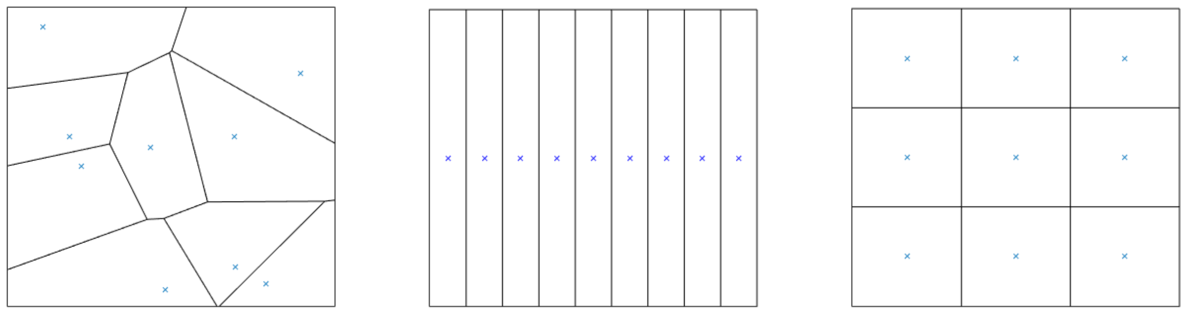
(iii) the Lloyd iterations stay in a compact set for which the Lloyd map is continuous,

then the Lloyd's algorithm converges globally.

While this result proves convergence to a centroidal Voronoi tessellation, it does not give conditions for uniqueness. In fact, there may be multiple centroidal Voronoi tessellations for a given density and domain, each with a different locally minimizing energy. See Figure 3.

In [7], Fleischer showed that for $N = 1$ the logarithmic concavity condition implies that any centroidal Voronoi tessellation is a local minimum, and, further, that there is a unique centroidal Voronoi tessellation that is both a local and global minimum of $J[u]$. Namely, he proved the following theorem:

Theorem 1.2 (Fleischer, 1964). *Let a given continuous probability density of finite second moment obey*



(a) Voronoi tessellation (b) Saddle point CVT (c) Globally minimizing CVT

FIG. 3. *Both figure (B) and (C) are centroidal Voronoi tessellations of the square with uniform density. Tessellation (B) is only a saddle point with respect to energy, while tessellation (C) globally minimizes energy.*

the inequality $[\ln p(x)]'' < 0$. Then the energy of an K-level quantizer has a unique stationary point. This point is a relative and absolute minimum.

As noted in numerous papers, including [2, 4], there has been no extension of Fleischer's theorem to $N > 1$. The need for vector quantizers in higher dimensions, rather than a superposition of N scalar quantizers, was shown in [13]. Convergence of Lloyd's algorithm and conditions for a unique centroidal Voronoi tessellation remain a relevant and open area of research.

However, rigorously verifying that a given centroidal Voronoi tessellation is a minimum can be difficult. Using a variational formulation, we determine when critical points are guaranteed to be minima and give sufficient conditions that can be used in practice to verify a given centroidal Voronoi tessellation is indeed a minimum. Furthermore, we address the open question of the existence of conditions for a unique centroidal Voronoi tessellation in higher dimensions. Namely, we prove that there exists multiple two generator centroidal Voronoi tessellations for any density and multidimensional domain. Such results have both clear practical and theoretical applications to Lloyd's algorithm and the many fields that use quantizers.

The remainder of the paper is as follows. In section two, we introduce the problem through variational calculus. We recall classical results and find explicit representations for the first and second variation of $J[u]$. In section three, we examine sufficient conditions for minima and illustrate how the conditions can be used in practice. Finally, in section four, we prove that there does not exist a unique two generator centroidal Voronoi tessellation for dimensions greater than one.

2. Variational Formulation. First, we recall the following differential notation. The gradient and Hessian of a function $f(x)$, denoted $\nabla f(x)$ and $H[f(x)]$, respectively, are defined by

$$\langle \nabla f(x), \varphi \rangle := \lim_{\epsilon \to 0^+} \frac{f(x+\epsilon\varphi) - f(x)}{\epsilon} \quad \text{and} \quad H[f(x)] := \nabla[\nabla f(x)]^T.$$

When referring to a set A, we will denote the interior, boundary, and closure by A^o, ∂A, and $\bar{A}$, respectively. Here, and in what follows, we will make use of the theory of variational calculus to assist in the study of the critical points of the functional $J[u]$. Fleischer's proof for $N = 1$ uses no such machinery but instead uses basic differential calculus and optimization concepts, namely, first and second order conditions of the form

$$\nabla J[u] = 0, H[\, J[u]\,] > 0,$$

where now and in what follows, "$>$" and "$\geq$" refer to positive definiteness and semi-definiteness, respectively. Differential calculus works well for $N = 1$, but it is more difficult for $N > 1$, due to the complex nature of the subregions Ω_i in higher dimensions.

We recall certain basic concepts from variational calculus that will prove useful in our analysis. For additional reading, the author suggests [8], a classic text by Gelfand on the subject.

Let us fix a point $u \in \mathcal{U}$. Define $u^* = u + \epsilon\varphi$, for some $\epsilon \in \mathbb{R}$, $\varphi \in \mathbb{R}^{N\times K}$, $\varphi = (\varphi_1^T, ..., \varphi_K^T)^T$, $\|\varphi_i\| \leq 1$ for all $i = 1, .., K$. The *increment* of $J[u]$ is given by

$$\Delta J[\epsilon, \varphi; u] = J[u^*] - J[u].$$

We say the functional $J[u]$ is twice-differentiable if its increment can be written in the form

$$\Delta J[\epsilon, \varphi; u] = \delta J[\varphi; u]\epsilon + \delta^2 J[\varphi; u]\epsilon^2 + M[\epsilon, \varphi; u]\epsilon^2,$$

where $\delta J[\varphi; u]$ and $\delta^2 J[\varphi; u]$ are linear and quadratic functionals in φ, respectively, and $\lim_{\epsilon\to 0} M[\epsilon, \varphi; u] = 0$. We say the functionals $\delta J[\varphi; u]$ and $\delta^2 J[\varphi; u]$ are the *first* and *second variation* of $J[u]$, respectively.

We recall the following results regarding extremal points of a functional, taken from [8].

THEOREM 2.1. *A necessary condition for the differentiable functional $J[u]$ to have an extremum at $u = \hat{u}$ is that its first variation vanishes, namely*

$$\delta J[\varphi; \hat{u}] = 0$$

for all admissible φ.

THEOREM 2.2. *A sufficient condition for a twice differentiable functional $J[u]$ to have a minimum at $u = \hat{u}$, given that the first variation $\delta J[\varphi; u]$ vanishes at $u = \hat{u}$, is that its second variation $\delta^2 J[\varphi; \hat{u}]$ is strongly positive, namely, there exists a constant $C > 0$ such that*

$$\delta^2 J[\varphi; \hat{u}] \geq C\|\varphi\|^2$$

for all admissible φ.

Theorems 2.1 and 2.2 will not be used until Section 3, but they motivate the work in this section. With the necessary machinery in hand, we examine the increment $\Delta J[\epsilon, \varphi; u]$.

Just as we have

$$\Omega_i = \{x \in \Omega | \|x - u_i\| \leq \|x - u_j\| \text{ for all } j\},$$

let us define

$$\Omega_i^* = \{x \in \Omega | \|x - u_i^*\| \leq \|x - u_j^*\| \text{ for all } j\}.$$

In this way, just as

$$J[u] = \sum_{i\in V} \int_{\Omega_i} \|x - u_i\|^2 p(x) dx,$$

we have

$$J[u^*] = \sum_{i\in V} \int_{\Omega_i^*} \|x - u_i^*\|^2 p(x) dx.$$

In addition, if $\Omega_i^* \cap \Omega_j^* \neq \emptyset$, let us define $\sigma_{i,j}^*$ as the unique element of $\Omega_i^* \cap \Omega_j^*$ that is collinear with $\{u_i^*, u_j^*\}$.

Because of the complex nature of the regions of integration Ω_i and Ω_i^*, the uncertain variation between the regions Ω_i and Ω_i^*, and the uncertainty in the change of boundaries or creation of new boundaries between $\{\Omega_i\}_{i=1}^K$ and $\{\Omega_i^*\}_{i=1}^K$, we divide the increment into two portions.

We introduce another functional, $\hat{J}[u^*; u]$, defined by

$$\hat{J}[u^*; u] = \sum_{i\in V} \int_{\Omega_i} \|x - u_i^*\|^2 p(x) dx.$$

We can now decompose the increment $\Delta J[\epsilon,\varphi;u]$ into two parts and write

$$\Delta J[\epsilon,\varphi;u] = \Delta J_u[\epsilon,\varphi;u] + \Delta J_\Omega[\epsilon,\varphi;u],$$

where

$$\Delta J_u[\epsilon,\varphi;u] = \hat{J}[u^*;u] - J[u],$$

$$\Delta J_\Omega[\epsilon,\varphi;u] = J[u^*] - \hat{J}[u^*;u].$$

The subscripts u and Ω have meaning, in that $\Delta J_u[\epsilon,\varphi;u]$ obtains its variation from the difference in functions of u and u^* integrated over identical subregions, and $\Delta J_\Omega[\epsilon,\varphi;u]$ obtains its variation from the difference in identical functions of u^* integrated over differing subregions. However, before we compute the representations of $\Delta J_u[\epsilon,\varphi;u]$ and $\Delta J_\Omega[\epsilon,\varphi;u]$, we give the following lemma, which will be useful in analyzing both components of the increment.

Lemma 2.3.

$$\|x-y\|^2 - \|x-z\|^2 = 2\langle x - \frac{y+z}{2}, z-y\rangle$$

With the assistance of Lemma 2.3, we can efficiently compute the representation of $\Delta J_u[\epsilon,\varphi;u]$.

Lemma 2.4.

$$\Delta J_u[\epsilon,\varphi;u] = -2\epsilon\sum_{i\in V}\int_{\Omega_i}\langle x-u_i,\varphi_i\rangle p(x)dx + \epsilon^2\sum_{i\in V}\|\varphi_i\|^2\mu(\Omega_i)$$

Proof.

$$\begin{aligned}\Delta J_u[\epsilon,\varphi;u] = \hat{J}[u^*;u] - J[u] &= \sum_{i\in V}\int_{\Omega_i}\|x-u_i^*\|^2 p(x)dx - \sum_{i\in V}\int_{\Omega_i}\|x-u_i\|^2 p(x)dx \\ &= \sum_{i\in V}\int_{\Omega_i}\|x-u_i^*\|^2 - \|x-u_i\|^2 p(x)dx\end{aligned}$$

By Lemma 2.3,

$$\begin{aligned}\|x-u_i^*\|^2 - \|x-u_i\|^2 &= 2\langle x - \frac{u_i+u_i^*}{2}, u_i-u_i^*\rangle = 2\langle x-u_i-\frac{\epsilon}{2}\varphi_i, -\epsilon\varphi_i\rangle \\ &= -2\epsilon\langle x-u_i,\varphi_i\rangle + \epsilon^2\|\varphi_i\|^2.\end{aligned}$$

□

The representation of the component $\Delta J_u[\epsilon,\varphi;u]$ has been straightforward to obtain. However, the component $\Delta J_\Omega[\epsilon,\varphi;u]$ requires precision and care, due to the complex nature of the subregions and the boundaries between them. We give the following lemma.

Lemma 2.5. *Let $\{\Omega_i^*\}_{i=1}^K$ and $\{\Omega_j\}_{j=1}^K$ be two Voronoi tessellations of Ω. Then*

$$\Omega = \cup_{1\le i\le K}\cup_{1\le j\le K}(\Omega_i^*\cap\Omega_j)$$

and

$$\partial(\Omega_i^*\cap\Omega_j) = \big(\cup_{\ell\ne i}\,\Omega_i^*\cap\Omega_\ell^*\cap\Omega_j\big)\cup\big(\cup_{\ell\ne j}\,\Omega_i^*\cap\Omega_\ell\cap\Omega_j\big)\cup\big(\Omega_i^*\cap\Omega_j\cap(\partial\Omega)\big).$$

Proof. We have

$$\cup_{1\le i\le K}\cup_{1\le j\le K}(\Omega_i^*\cap\Omega_j) = \cup_{1\le i\le K}(\Omega_i^*\cap(\cup_{1\le j\le K}\Omega_j)) = \cup_{1\le i\le K}(\Omega_i^*\cap\Omega) = \cup_{1\le i\le K}\Omega_i^* = \Omega.$$

Now consider $x \in \partial(\Omega_i^* \cap \Omega_j)$. We have either $x \in \partial\Omega_i^*$, $x \in \partial\Omega_j$, or both. Without loss of generality, suppose $x \in \partial\Omega_i^*$. Then $x \in \Omega_i^* \cap \Omega_\ell^*$ for some $\ell \neq i$, or $x \in \partial\Omega$. Therefore,

$$\partial(\Omega_i^* \cap \Omega_j) \subset \big(\cup_{\ell\neq i}\, \Omega_i^* \cap \Omega_\ell^* \cap \Omega_j\big) \cup \big(\cup_{\ell\neq j}\, \Omega_i^* \cap \Omega_\ell \cap \Omega_j\big) \cup \big(\Omega_i^* \cap \Omega_j \cap (\partial\Omega)\big).$$

Because $\Omega_j \cap \Omega_\ell \subset \partial\Omega_j$ for all $\ell \neq j$ and $\Omega_j \cap \partial\Omega \subset \partial\Omega_j$, we also have

$$\partial(\Omega_i^* \cap \Omega_j) \supset \big(\cup_{\ell\neq i}\, \Omega_i^* \cap \Omega_\ell^* \cap \Omega_j\big) \cup \big(\cup_{\ell\neq j}\, \Omega_i^* \cap \Omega_\ell \cap \Omega_j\big) \cup \big(\Omega_i^* \cap \Omega_j \cap (\partial\Omega)\big).$$

□

We have the following representation for $\Delta J_\Omega[\epsilon, \varphi; u]$.

Lemma 2.6.

$$\Delta J_\Omega[\epsilon, \varphi; u] = -\epsilon^2 \sum_{e_{i,j}\in E} \int_{\Omega_i\cap\Omega_j} \frac{[\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle]^2}{\|u_j-u_i\|} p(x)dx + M_\Omega[\epsilon,\varphi;u]\epsilon^2,$$

where $\lim_{\epsilon\to 0} M_\Omega[\epsilon, \varphi; u] = 0$.

Proof. For the sake of space, we introduce the notation

$$u_{i\to j} = u_j - u_i, \quad u_{i\to j}^* = u_j^* - u_i^*,$$

and

$$\Omega_{i^*,j} = \Omega_i^* \cap \Omega_j, \quad \Omega_{i^*,j,k} = \Omega_i^* \cap \Omega_j \cap \Omega_k, \quad \Omega_{i^*,j^*,k} = \Omega_i^* \cap \Omega_j^* \cap \Omega_k.$$

By Lemmas 2.3 and 2.5, we have

$$\begin{aligned}
\Delta J_\Omega[\epsilon, \varphi; u] &= J[u^*] - \hat{J}[u^*, u] = \sum_{i\in V}\int_{\Omega_i^*} \|x-u_i^*\|^2 p(x)dx - \sum_{j\in V}\int_{\Omega_j} \|x-u_j^*\|^2 p(x)dx \\
&= \sum_{i\in V}\sum_{j\in V}\int_{\Omega_{i^*,j}} \big(\|x-u_i^*\|^2 - \|x-u_j^*\|^2\big)p(x)dx = \sum_{i\neq j}\int_{\Omega_{i^*,j}} 2\langle x-\sigma_{i,j}^*, u_{i\to j}^*\rangle p(x)dx \\
&= \sum_{i\neq j}\int_{\Omega_{i^*,j}} \frac{2}{\|u_{i\to j}^*\|^2}\langle x-\sigma_{i,j}^*, u_{i\to j}^*\rangle\langle u_{i\to j}^*, u_{i\to j}^*\rangle p(x)dx \\
&= \sum_{i\neq j}\int_{\Omega_{i^*,j}} \frac{2}{\|u_{i\to j}^*\|^2}(x-\sigma_{i,j}^*)^T[u_{i\to j}^*(u_{i\to j}^*)^T]u_{i\to j}^* p(x)dx.
\end{aligned}$$

Applying integration by parts, we obtain

$$\Delta J_\Omega[\epsilon, \varphi; u] = \sum_{i\neq j}\Bigg[\int_{\partial(\Omega_{i^*,j})} \frac{\langle x-\sigma_{i,j}^*, u_{i\to j}^*\rangle^2\langle u_{i\to j}^*, n\rangle}{\|u_{i\to j}^*\|^2} p(x)dx - \int_{\Omega_{i^*,j}} \frac{\langle x-\sigma_{i,j}^*, u_{i\to j}^*\rangle^2\langle u_{i\to j}^*, \nabla p(x)\rangle}{\|u_{i\to j}^*\|^2} dx\Bigg],$$

where n is the unit normal vector, oriented outward with respect to $\Omega_{i^*,j}$.

By Lemma 2.5,

$$\partial(\Omega_{i^*,j}) = \big(\cup_{\ell\neq i}\, \Omega_i^* \cap \Omega_\ell^* \cap \Omega_j\big) \cup \big(\cup_{\ell\neq j}\, \Omega_i^* \cap \Omega_\ell \cap \Omega_j\big) \cup \big(\Omega_i^* \cap \Omega_j \cap (\partial\Omega)\big).$$

For $x \in \Omega_{i^*,j^*,j}$, $\langle u_{i\to j}^*, x-\sigma_{i,j}^*\rangle = 0$. Therefore the integral over $\Omega_{i^*,j^*,j}$ is zero. For $x \in \Omega_{i^*,i,j}$, $\langle u_{i\to j}, x - \sigma_{i,j}\rangle = 0$, giving

$$\begin{aligned}
\langle u_{i\to j}^*, x-\sigma_{i,j}^*\rangle &= \langle u_{i\to j}, x-\sigma_{i,j}\rangle + \langle u_{i\to j}^* - u_{i\to j}, x-\sigma_{i,j}\rangle + \langle u_{i\to j}^*, \sigma_{i,j}-\sigma_{i,j}^*\rangle \\
&= \epsilon\langle\varphi_j - \varphi_i, x-\sigma_{i,j}\rangle - \frac{\epsilon}{2}\langle u_{i\to j}^*, \varphi_i+\varphi_j\rangle \\
&= \epsilon\langle\varphi_j - \varphi_i, x-\sigma_{i,j}\rangle - \frac{\epsilon}{2}\langle u_{i\to j}, \varphi_i+\varphi_j\rangle - \frac{\epsilon^2}{2}\langle\varphi_j-\varphi_i, \varphi_i+\varphi_j\rangle \\
&= \epsilon\langle\varphi_j, x-\sigma_{i,j} - \frac{u_{i\to j}}{2}\rangle - \epsilon\langle\varphi_i, x-\sigma_{i,j} + \frac{u_{i\to j}}{2}\rangle - \frac{\epsilon^2}{2}\langle\varphi_j-\varphi_i, \varphi_i+\varphi_j\rangle \\
&= \epsilon(\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle) - \frac{\epsilon^2}{2}\langle\varphi_j-\varphi_i, \varphi_i+\varphi_j\rangle.
\end{aligned}$$

Therefore,

$$\langle u^*_{i\to j}, x-\sigma^*_{i,j}\rangle^2 = \epsilon^2(\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle)^2 - \frac{\epsilon^3}{2}\langle\varphi_j-\varphi_i,\varphi_i+\varphi_j\rangle(\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle - \frac{\epsilon}{2}\langle\varphi_j-\varphi_i,\varphi_i+\varphi_j\rangle).$$

For $\Omega_{i^*,i,j}$ oriented with respect to $\Omega_{i^*,j}$,

$$\langle u^*_{i\to j}, n\rangle = -\frac{\langle u^*_{i\to j}, u_{i\to j}\rangle}{\|u_{i\to j}\|} = -\frac{\langle u^*_{i\to j}, u^*_{i\to j}\rangle + \langle u^*_{i\to j}, u_{i\to j}-u^*_{i\to j}\rangle}{\|u_{i\to j}\|} = -\frac{\|u^*_{i\to j}\|^2}{\|u_{i\to j}\|} + \epsilon\frac{\langle u^*_{i\to j}, \varphi_j-\varphi_i\rangle}{\|u_{i\to j}\|},$$

giving

$$\frac{\langle u^*_{i\to j}, x-\sigma^*_{i,j}\rangle^2\langle u^*_{i\to j}, n\rangle}{\|u^*_{i\to j}\|^2} = \epsilon\frac{\langle u^*_{i\to j},\varphi_j-\varphi_i\rangle\langle u^*_{i\to j}, x-\sigma^*_{i,j}\rangle^2}{\|u_{i\to j}\|\|u^*_{i\to j}\|^2} - \epsilon^2\frac{[\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle]^2}{\|u_{i\to j}\|} + \epsilon^3\frac{\langle\varphi_j-\varphi_i,\varphi_i+\varphi_j\rangle}{2\|u_{i\to j}\|}(\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle - \frac{\epsilon}{2}\langle\varphi_j-\varphi_i,\varphi_i+\varphi_j\rangle)$$

for $x\in\Omega_{i^*,i,j}$ oriented with respect to $\Omega_{i^*,j}$.

Noting that $p(x)=0$ on $\partial\Omega$ and that $\Omega_{i,j}=\cup_{\ell\in V}\Omega_{i,j,\ell^*}$, we can write

$$\Delta J_\Omega[\epsilon,\varphi;u] = -\epsilon^2\sum_{e_{i,j}\in E}\int_{\Omega_i\cap\Omega_j}\frac{[\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle]^2}{\|u_{i\to j}\|}p(x)dx + M_\Omega[\epsilon,\varphi;u]\epsilon^2,$$

where

$$\begin{aligned}M_\Omega[\epsilon,\varphi;u] = \frac{1}{\epsilon^2}\sum_{i\neq j}\Bigg[&\int_{(\cup_{\ell\neq i,j}\Omega_{i^*,\ell,j})\cup(\cup_{\ell\neq i,j}\Omega_{i^*,\ell^*,j})}\frac{\langle x-\sigma^*_{i,j}, u^*_{i\to j}\rangle^2\langle u^*_{i\to j}, n\rangle}{\|u^*_{i\to j}\|^2}p(x)dx \\ &-\int_{\Omega_{i^*,j}}\frac{\langle x-\sigma^*_{i,j}, u^*_{i\to j}\rangle^2\langle u^*_{i\to j}, \nabla p(x)\rangle}{\|u^*_{i\to j}\|^2}dx + \epsilon\int_{\Omega_{i^*,i,j}}\frac{\langle u^*_{i\to j},\varphi_j-\varphi_i\rangle\langle u^*_{i\to j}, x-\sigma^*_{i,j}\rangle^2}{\|u_{i\to j}\|\|u^*_{i\to j}\|^2}p(x)dx \\ &+\epsilon^3\int_{\Omega_{i^*,i,j}}\frac{\langle\varphi_j-\varphi_i,\varphi_i+\varphi_j\rangle}{2\|u_{i\to j}\|}(\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle - \frac{\epsilon}{2}\langle\varphi_j-\varphi_i,\varphi_i+\varphi_j\rangle)p(x)dx\Bigg] \\ &+\sum_{e_{i,j}\in E}\int_{\cup_{\ell\neq i,j}\Omega_{i,j,\ell^*}}\frac{[\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle]^2}{\|u_{i\to j}\|}p(x)dx.\end{aligned}$$

The proof that $\lim_{\epsilon\to 0} M_\Omega[\epsilon,\varphi;u]=0$ is tedious and adds little insight to the result. Therefore, we leave it to Appendix A. □

With an explicit representation of both $\Delta J_u[\epsilon,\varphi;u]$ and $\Delta J_\Omega[\epsilon,\varphi;u]$, we can write the increment $\Delta J[\epsilon,\varphi;u]$ explicitly.

Theorem 2.7.

$$\Delta J[\epsilon,\varphi;u] = -2\epsilon\sum_{i\in V}\int_{\Omega_i}\langle x-u_i,\varphi_i\rangle p(x)dx + \epsilon^2\sum_{i\in V}\|\varphi_i\|^2\mu(\Omega_i) - \epsilon^2\sum_{e_{i,j}\in E}\int_{\Omega_i\cap\Omega_j}\frac{[\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle]^2}{\|u_j-u_i\|}p(x)dx + M[\epsilon,\varphi;u]\epsilon^2,$$

where $\lim_{\epsilon\to 0} M[\epsilon,\varphi;u]=0$.

Explicit representations of the first and second variation $\delta J[\varphi;u]$ and $\delta^2 J[\varphi;u]$ follow naturally.

Corollary 2.8.

$$\delta J[\varphi;u] = -2\sum_{i\in V}\int_{\Omega_i}\langle x-u_i,\varphi_i\rangle p(x)dx$$

Corollary 2.9.

$$\delta^2 J[\varphi;u] = \sum_{i\in V}\|\varphi_i\|^2\mu(\Omega_i) - \sum_{e_{i,j}\in E}\int_{\Omega_i\cap\Omega_j}\frac{[\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle]^2}{\|u_j-u_i\|}p(x)dx$$

3. Conditions for Extrema. With explicit representations of both the first and second variations in hand, we may now investigate necessary conditions for extremal points and give sufficient conditions for which those points are minima. We have the following necessary condition with respect to the first variation.

THEOREM 3.1. *u is an extremal point of J only if*

$$u_i = \mu(\Omega_i)^{-1} \int_{\Omega_i} xp(x)dx \text{ for all } i = 1, ..., K.$$

Proof. By Theorem 2.1, we require

$$\sum_{i \in V} \int_{\Omega_i} \langle x - u_i, \varphi_i \rangle p(x)dx = 0$$

for all admissible φ. This implies that

$$\int_{\Omega_i} (x - u_i)p(x)dx = 0$$

for all $i = 1, ..., K$. Rearranging, the desired result follows. □

Before we make use of the second variation for sufficient conditions, we give the following lemmas:

LEMMA 3.2. *Let f be a C^2 function on a convex region Ω. Then we have*

$$\min_{x \in \Omega} \lambda_{min}(H[f(x)]) \leq \frac{\langle x - y, \nabla_x f(x) - \nabla_y f(y) \rangle}{\langle x - y, x - y \rangle} \text{ for all } x \neq y \in \Omega.$$

Proof. The result follows from the definition of the Hessian, namely

$$H[f(x)]\varphi = \lim_{\epsilon \to 0} \frac{1}{\epsilon}[\nabla f(x + \epsilon\varphi) - \nabla f(x)].$$

□

LEMMA 3.3.

$$arg \min_{y \in \Omega} \int_{\Omega} \|x - y\|^2 p(x)dx = \mu(\Omega)^{-1} \int_{\Omega} xp(x)dx$$

Proof. Observing the first order condition

$$\nabla_y \int_{\Omega} \|x - y\|^2 p(x)dx = \int_{\Omega} 2(y - x)p(x)dx = 0$$

gives us

$$y = \mu(\Omega)^{-1} \int_{\Omega} xp(x)dx$$

as the unique critical point. We note that

$$H[\int_{\Omega} \|x - y\|^2 p(x)dx] = \nabla_y \int_{\Omega} 2(y - x)p(x)dx = 2\mu(\Omega)I.$$

□

We also recall the following result with respect to block diagonally dominant matrices, taken from [6].

LEMMA 3.4. *Let $f(x) = \sum_{i=1}^{K} \sum_{j=1}^{K} x_i^T a_{i,j} x_j$, with $a_{i,j} = a_{j,i}^T$ for all $i, j = 1, ..., K$. The function $f(x)$ is strongly positive if*

(i) $a_{i,i}$ *is SPD,* $i = 1, ..., K$,

(ii) $(\|a_{i,i}^{-1}\|)^{-1} \geq \sum_{j \neq i} \|a_{i,j}\|$, $i = 1, .., K$, *with strict inequality holding for at least one* i,

(iii) $A = (a_{i,j})_{i,j=1}^{K}$ *is block irreducible.*

We prove the following sufficient condition with respect to the second variation.

THEOREM 3.5. *Suppose*

$$u_i = \mu(\Omega_i)^{-1} \int_{\Omega_i} xp(x)dx$$

and

$$\frac{J_i[u]}{2} \min_{x\in\Omega} \lambda_{min}(H[-\ln p(x)]) \geq (N-1)\mu(\Omega_i) + \sum_{j\sim i} \int_{\Omega_i\cap\Omega_j} \frac{\|x_i - \sigma_{i,j}\|^2}{\|u_i - \sigma_{i,j}\|} p(x)dx$$

for all $i = 1, ..., K$, with strict inequality holding for at least one i. Then u is a minimum.

Proof. The condition

$$u_i = \mu(\Omega_i)^{-1} \int_{\Omega_i} xp(x)dx, \quad i = 1, ..., K$$

implies that $\delta J[\varphi; u] = 0$. Therefore, by Theorem 2.2 all we require is that $\delta^2 J[\varphi; u]$ is strongly positive. However, before we can apply the result of Lemma 3.4, we must verify that $\delta^2 J[\varphi; u]$ can be put in the necessary form. We have

$$\begin{aligned}[\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle]^2 = \varphi_j^T(x-u_j)(x-u_j)^T\varphi_j - \varphi_i^T(x-u_i)(x-u_j)^T\varphi_j \\ - \varphi_j^T(x-u_j)(x-u_i)^T\varphi_i + \varphi_i^T(x-u_i)(x-u_i)^T\varphi_i.\end{aligned}$$

From this we obtain

$$\delta^2 J[\varphi; u] = \sum_{i\in V}\sum_{j\in V} \varphi_i^T a_{i,j}\varphi_j,$$

where

$$a_{i,i} = \mu(\Omega_i)I - \sum_{j\sim i} \int_{\Omega_i\cap\Omega_j} \frac{(x-u_i)(x-u_i)^T}{\|u_j - u_i\|} p(x)dx,$$

$$a_{i,j} = \int_{\Omega_i\cap\Omega_j} \frac{(x-u_i)(x-u_j)^T}{\|u_j - u_i\|} p(x)dx, \quad i \neq j.$$

We note that

$$(\|a_{i,i}^{-1}\|)^{-1} = \inf \frac{\|A_{i,i}x\|}{\|x\|} = \inf \frac{\|sx - Bx\|}{\|x\|} \geq \inf \frac{s\|x\| - \|Bx\|}{\|x\|} = s - \sup \frac{\|Bx\|}{\|x\|} = s - \|B\|.$$

We require

$$\mu(\Omega_i) \geq \left\| \sum_{j\sim i} \int_{\Omega_i\cap\Omega_j} \frac{(x-u_i)(x-u_i)^T}{\|u_j - u_i\|} p(x)dx \right\| + \sum_{j\sim i} \left\| \int_{\Omega_i\cap\Omega_j} \frac{(x-u_i)(x-u_j)^T}{\|u_j - u_i\|} p(x)dx \right\|.$$

From this condition, it quickly follows that $a_{i,i}$ is SPD. In addition, $A = (a_{i,j})_{i,j=1}^K$ is block irreducible, from the connectedness of the Voronoi graph G. Examining the right hand side, we have

$$\begin{aligned}
&\left\| \sum_{j\sim i} \int_{\Omega_i\cap\Omega_j} \frac{(x-u_i)(x-u_i)^T}{\|u_j-u_i\|} p(x)dx \right\| + \sum_{j\sim i} \left\| \int_{\Omega_i\cap\Omega_j} \frac{(x-u_i)(x-u_j)^T}{\|u_j-u_i\|} p(x)dx \right\| \\
&\qquad \leq \sum_{j\sim i} \int_{\Omega_i\cap\Omega_j} \frac{\|(x-u_i)(x-u_i)^T\| + \|(x-u_i)(x-u_j)^T\|}{\|u_j-u_i\|} p(x)dx \\
&\qquad \leq \sum_{j\sim i} \int_{\Omega_i\cap\Omega_j} \frac{\|x-u_i\|^2 + \|x-u_i\|\|x-u_j\|}{\|u_j-u_i\|} p(x)dx = \sum_{j\sim i} \int_{\Omega_i\cap\Omega_j} \frac{2\|x-u_i\|^2}{\|u_j-u_i\|} p(x)dx \\
&\qquad \leq \sum_{j\sim i} \int_{\Omega_i\cap\Omega_j} \frac{2\|x-u_i\|^2 - \langle x-u_i, x-u_j\rangle + \langle x-u_i, x-u_j\rangle}{\|u_j-u_i\|} p(x)dx \\
&\qquad \leq \sum_{j\sim i} \int_{\Omega_i\cap\Omega_j} \frac{\langle x-u_i, u_j-u_i\rangle}{\|u_j-u_i\|} p(x)dx + \sum_{j\sim i} \int_{\Omega_i\cap\Omega_j} \frac{2\langle x-u_i, x-\sigma_{i,j}\rangle}{\|u_j-u_i\|} p(x)dx \\
&\qquad \leq \int_{\partial\Omega_i} \langle x-u_i, n\rangle p(x)dx + \sum_{j\sim i} \int_{\Omega_i\cap\Omega_j} \frac{\|x-\sigma_{i,j}\|^2}{\|u_i-\sigma_{i,j}\|} p(x)dx.
\end{aligned}$$

Using Stokes' Theorem, we have

$$\int_{\partial\Omega_i} \langle x-u_i, n\rangle p(x)dx = \int_{\Omega_i} \langle \nabla p(x), x-u_i\rangle dx + N\mu(\Omega_i).$$

Therefore, to prove the theorem it suffices to have

$$\int_{\Omega_i} \langle \nabla p(x), u_i - x\rangle dx \geq (N-1)\mu(\Omega_i) + \sum_{j\sim i} \int_{\Omega_i\cap\Omega_j} \frac{\|x-\sigma_{i,j}\|^2}{\|u_i-\sigma_{i,j}\|} p(x)dx.$$

We also have

$$\begin{aligned}
\int_{\Omega_i} \langle \nabla_x p(x), u_i - x\rangle dx &= \mu(\Omega_i)^{-1}\Big(\langle \int_{\Omega_i} xp(x)dx, \int_{\Omega_i} \nabla_x p(x)dx\rangle - \int_{\Omega_i} p(x)dx \int_{\Omega_i} \langle x, \nabla_x p(x)\rangle dx \Big) \\
&= \mu(\Omega_i)^{-1}\Big(\langle \int_{\Omega_i} xp(x)dx, \int_{\Omega_i} \nabla_y p(y)dy\rangle - \int_{\Omega_i} p(x)dx \int_{\Omega_i} \langle y, \nabla_y p(y)\rangle dy \Big) \\
&= \mu(\Omega_i)^{-1} \int_{\Omega_i}\int_{\Omega_i} p(x)\langle x, \nabla_y p(y)\rangle - p(x)\langle y, \nabla_y p(y)\rangle dxdy \\
&= \mu(\Omega_i)^{-1} \int_{\Omega_i}\int_{\Omega_i} p(x)\langle x-y, \nabla_y p(y)\rangle dxdy \\
&= \mu(\Omega_i)^{-1}\frac{1}{2} \int_{\Omega_i}\int_{\Omega_i} \langle x-y, p(x)\nabla_y p(y) - p(y)\nabla_x p(x)\rangle dxdy \\
&= \mu(\Omega_i)^{-1}\frac{1}{2} \int_{\Omega_i}\int_{\Omega_i} \langle x-y, \nabla_y \ln p(y) - \nabla_x \ln p(x)\rangle p(x)p(y)dxdy.
\end{aligned}$$

Applying Lemmas 3.2 and 3.3,

$$\begin{aligned}
\int_{\Omega_i} \langle \nabla_x p(x), u_i - x\rangle dx &\geq \mu(\Omega_i)^{-1}\frac{1}{2}\big[\min_{x\in\Omega} \lambda_{min}(H[-\ln p(x)])\big] \int_{\Omega_i} \Big[\int_{\Omega_i} \|x-y\|^2 p(x)dx \Big] p(y)dy \\
&\geq \frac{J_i[u]}{2} \min_{x\in\Omega} \lambda_{min}(H[-\ln p(x)]).
\end{aligned}$$

□

With this, we can easily prove Fleischer's sufficient conditions for local minima for $N = 1$.

Corollary 3.6. *Let $\Omega \subset \mathbb{R}$. A sufficient condition for u to be a minima of J is*

$$\tilde{u}_i = \mu(\Omega_i)^{-1} \int_{\Omega_i} xp(x)dx$$

and

$$\frac{d^2}{dx^2}\ln p(x) < 0$$

for all $x \in \Omega$.

Proof. When $N = 1$, $\Omega_i \cap \Omega_j = \{\sigma_{i,j}\}$. By Theorem 3.5, a sufficient condition is

$$\min_{x\in\Omega} \lambda_{min}(H[-\ln p(x)]) = \min_{x\in\Omega} \frac{d^2}{dx^2}[-\ln p(x)] > 0.$$

□

When the dimension increases, the additional terms play a role, and the multidimensional version of the log-concavity condition

$$\min_{x\in\Omega} \lambda_{min}(H[-\ln p(x)]) > 0$$

is not sufficient to guarantee a critical point to be a minimum. We have the following example for $N = 2$.

EXAMPLE 3.7. *Consider the function* $f : \Omega := [0, M]^2 \to [0, 1]$, $p(x) = \exp\{-\|x\|^2\}$, *and the case* $K = 2$. *We introduce the subregions*

$$\Omega_1 = \{x \in \Omega | x_1 \geq x_2\}, \quad \Omega_2 = \{x \in \Omega | x_2 \geq x_1\}.$$

The centroidal locations are given by

$$u_1 = \frac{\sqrt{2\pi}}{8}\begin{pmatrix} erf(\sqrt{2}M) - \sqrt{2}e^{-M^2} erf(M) \\ \sqrt{2}erf(M) - erf(\sqrt{2}M) \end{pmatrix}, \quad u_2 = \frac{\sqrt{2\pi}}{8}\begin{pmatrix} \sqrt{2}erf(M) - erf(\sqrt{2}M) \\ erf(\sqrt{2}M) - \sqrt{2}e^{-M^2} erf(M) \end{pmatrix}.$$

The centroidal locations are reflections about the line $x_1 = x_2$, *implying that they are generators for* Ω_1, Ω_2 *and, therefore, produce a centroidal Voronoi tessellation. We note that we have the log-concavity condition*

$$\lambda_{min}(H[-\ln p(x)]) = \lambda_{min}(H[\|x\|^2]) > 0.$$

We investigate $\delta^2 J[\varphi^*; u]$ *for* $\varphi_1^* = (0, 1)^T$, $\varphi_2^* = (0, 0)^T$. *For illustration, we consider the result as* $M \to \infty$, *though the result holds for choices for which* M *is finite. We have*

$$\begin{aligned}
\delta^2 J[\varphi^*; u] &= \int_{\Omega_1} p(x)dx - \int_{\Omega_1\cap\Omega_2} \frac{\langle \varphi_1, x - u_1\rangle^2}{\|u_2 - u_1\|} p(x)dx \\
&= \int_0^\infty \int_0^{x_1} e^{-(x_1^2+x_2^2)} dx_2 dx_1 - \int_0^\infty \frac{(\frac{x}{\sqrt{2}} - u_1(2))^2}{\|u_2 - u_1\|} e^{-x^2} dx \\
&= \frac{\pi}{8} - \frac{\frac{\sqrt{\pi}}{64}((3 - 2\sqrt{2})\pi + 8(2 - \sqrt{2}))}{\frac{\sqrt{2\pi}}{4}(\sqrt{2} - 1)} = \frac{\pi}{16\sqrt{2}(\sqrt{2} - 1)} - \frac{1}{2} < 0,
\end{aligned}$$

so therefore the centroidal Voronoi diagram of u_1, u_2 *cannot be a minimum.*

Finally, we give a simple example of Theorem 3.5 used in practice to verify a given centroidal Voronoi tessellation is indeed a minimum.

EXAMPLE 3.8. *Let* $\Omega = [-10, 10]^2$ *and* $p(x) = e^{-\|x\|^2/2}$. *The generators*

$$u_1 = \begin{pmatrix} (1 - e^{-50})\sqrt{2\pi} erf(5\sqrt{2}) \\ 0 \end{pmatrix}, \quad u_2 = -u_1$$

produce a centroidal Voronoi tessellation with regions $\Omega_1 = [0, 10] \times [-10, 10]$, $\Omega_2 = [-10, 0] \times [-10, 10]$. *We have*

$$\lambda_{min}(H[-\ln p(x)]) = \lambda_{min}(H[\|x\|^2/2]) = 1,$$

$$\mu(\Omega_1) = \mu(\Omega_2) = \frac{1}{2}\int_{\Omega} e^{-\|x\|^2/2}dx = \pi\, erf(5\sqrt{2})^2 \approx \pi,$$

$$\begin{aligned} J[u_1] = J[u_2] &= \int_{\Omega_1} \|x-u_1\|^2 e^{-\|x\|^2/2}dx \\ &= \sqrt{\pi/2}\,erf(5\sqrt{2})\Big((1-e^{-50})\sqrt{2\pi}\,erf(5\sqrt{2})((1-e^{-50})2\pi\,erf(5\sqrt{2})-4) \\ &\quad + 2\sqrt{2\pi}\,erf(5\sqrt{2}) + 4e^{-50}((1-e^{-50})\sqrt{2\pi}\,erf(5\sqrt{2})-10)\Big) \\ &\approx 2\pi(\pi-1), \end{aligned}$$

and

$$\int_{\Omega_{1,2}} \frac{\|x\|^2}{\|u_1\|} e^{-\|x\|^2/2}dx = \frac{1}{1-e^{-50}}\left(1 - \frac{20e^{-50}}{\sqrt{2\pi}\,erf(5\sqrt{2})}\right) \approx 1.$$

We see that

$$\pi(\pi-1) - \pi - 1 = (\pi-1)^2 > 0,$$

implying that centroids u_1 and u_2 satisfy the conditions of Theorem 3.5, and, therefore, the centroidal Voronoi tessellation is a strict minimum.

4. Non-Uniqueness for Two Generators. Following and expanding upon the techniques of Fleischer, the sufficient conditions in Theorem 3.5 can be used to give sufficient conditions for a unique centroidal Voronoi tessellation. This procedure is done in Appendix B, however, the resulting conditions in dimensions greater than one prove unwieldy, and may describe an empty set in general. In particular, we show that for any density and multidimensional domain, there exists multiple two generator centroidal Voronoi tessellations.

First, we must first show that the energy $J[u]$ is stable with respect to small changes in the density $p(x)$. For the remainder of the section, we assume all densities are positive real-valued C^2 functions with compact and convex support contained in Ω, but not necessarily given by Ω. Let us denote the energy of the generator u over a density $p(x)$ by $J[u;p]$. Let

$$\Delta(\Omega) := \max_{x,y\in\Omega} \|x-y\|$$

be the diameter of Ω. Using the diameter and the measure of the domain Ω, we can bound the change in energy resulting from a change in density.

Lemma 4.1. *If $\|p(x)-\hat{p}(x)\| < \epsilon$, then $|J[u;p] - J[u;\hat{p}]| < \epsilon\Delta(\Omega)^2\mu(\Omega)^{1/2}$.*

Proof. We may write $\hat{p}(x)$ as

$$\hat{p}(x) = p(x) + \epsilon\gamma(x), \quad \|\gamma(x)\| < 1.$$

By linearity of J with respect to the density, we have

$$|J[u;\hat{p}] - J[u;p]| = \epsilon\Big|\sum_{i\in V}\int_{\Omega_i} \|x-u_i\|^2\gamma(x)dx\Big| \le \epsilon\sum_{i\in V}\int_{\Omega_i}\|x-u_i\|^2|\gamma(x)|dx.$$

Applying the Cauchy-Schwartz inequality, we obtain

$$\begin{aligned} |J[u;\hat{p}] - J[u;p]| &\le \epsilon\int_{\Omega}\big(\sum_{i\in V}\|x-u_i\|^2\mathbf{1}\{x\in\Omega_i\}\big)|\gamma(x)|dx \\ &\le \epsilon\bigg(\int_{\Omega}\big(\sum_{i\in V}\|x-u_i\|^2\mathbf{1}\{x\in\Omega_i\}\big)^2dx\bigg)^{1/2}\bigg(\int_{\Omega}\gamma(x)^2dx\bigg)^{1/2} \\ &< \epsilon[\Delta(\Omega)^4\mu(\Omega)]^{1/2} = \epsilon\Delta(\Omega)^2\mu(\Omega)^{1/2}. \end{aligned}$$

□

If we have a given centroidal Voronoi tessellation u for a density $p(x)$, for sufficiently small perturbations of the density, we can bound the distance between u and the closest centroidal Voronoi tessellation for the new, perturbed density. However, first we must define notions of distance on Ω^K. We define a natural inner product, given by

$$\langle u, v\rangle_{\Omega^K} := \sum_{i=1}^{K} \langle u_i, v_i\rangle,$$

with induced norm

$$\|u - v\|_{\Omega^K}^2 = \sum_{i=1}^{K} \|u_i - v_i\|^2.$$

We have the following theorem.

THEOREM 4.2. *Let u be a centroidal Voronoi tessellation of $p(x)$ and $J[u;p]$ be σ-strongly convex in $B(u,R) := \{v \in \Omega^K | \|v-u\|_{\Omega^K} \leq R\}$. Then for any $\hat{p}(x)$ satisfying*

$$\|p(x) - \hat{p}(x)\| < \epsilon := \frac{\sigma R^2}{4\Delta(\Omega)^2 \mu(\Omega)^{1/2}}$$

there exists some centroidal Voronoi tessellation $\hat{u}$ of $\hat{p}(x)$ such that

$$\|u - \hat{u}\|_{\Omega^K} < \sqrt{\epsilon}.$$

Proof. Because $J[u;p]$ is σ-strongly convex in $B(u,R)$ and u is a centroidal Voronoi tessellation of $p(x)$, we have that for any $v \in B(u,R)$,

$$J[v;p] \geq J[u;p] + \langle \nabla J[u;p], v-u\rangle + \frac{\sigma}{2}\|u-v\|_{\Omega^K}^2 = J[u;p] + \frac{\sigma}{2}\|u-v\|_{\Omega^K}^2.$$

This gives

$$\begin{aligned} J[v;\hat{p}] - J[u;\hat{p}] &= (J[v;\hat{p}] - J[v;p]) + (J[u;p] - J[u;\hat{p}]) + (J[v;p] - J[u,p]) \\ &\geq \frac{\sigma}{2}\|u-v\|_{\Omega^K}^2 - |J[v;\hat{p}] - J[v;p]| - |J[u;\hat{p}] - J[u;p]|. \end{aligned}$$

Now, suppose $\|u-v\|_{\Omega^K} = R$. Using Lemma 4.1, we have

$$J[v;\hat{p}] - J[u;\hat{p}] > \frac{\sigma R^2}{2} - 2\epsilon\Delta(\Omega)^2\mu(\Omega)^{1/2}.$$

Setting

$$\epsilon = \frac{\sigma R^2}{4\Delta(\Omega)^2\mu(\Omega)^{1/2}}$$

implies that

$$J[v, t^* + \tau] - J[u, t^* + \tau] > 0$$

for all v such that $\|u-v\|_{\Omega^K} = R$. Therefore $B(u,R)$ must contain a local minimum in its interior for $\hat{p}(x)$. This local minimum is by definition a centroidal Voronoi tessellation. □

Given Theorem 4.2, we are now prepared to prove that for any density on a multidimensional domain, there exists multiple two generator centroidal Voronoi tessellations.

THEOREM 4.3. *Let $\Omega \subset \mathbb{R}^N$, $N > 1$. Then there exists multiple two generator centroidal Voronoi tessellations that are distinct not counting permutations.*

Proof. Suppose, to the contrary, that there is a unique centroidal Voronoi tessellation, given by $u = (u_1, u_2)$. Let the mass center of Ω with respect to $p(x)$ be given by z. We necessarily have $z = \alpha u_1 + (1-\alpha)u_2$ for some $0 < \alpha < 1$.

For each line L in $\mathbb{R}^N$ that passes through z, we can consider the projection of $p(x)$ onto L, given by

$$p_L(x) = \int_{\{y|(y-x)\perp L\}} p(x)dy, \quad x \in L.$$

Each p_L has convex and compact support on a one dimensional interval of length at most $\Delta(\Omega)$, where we will assume z to correspond to the origin. Let

$$\Theta = \{u | u \text{ is a local minimum two generator centroidal Voronoi tessellation of } p_L \text{ for some } L\} \subset \Omega^2.$$

We note that any centroidal Voronoi tessellation in one dimension that is a local minimum must be a strict local minimum. This is easily seen from considering $\delta^2 J[\varphi; v]$ in the one dimensional case and noting that p_L has convex support. This implies that each minimum is σ-strongly convex, for some $\sigma > 0$.

Let $v = (v_1, v_2)$ be an element of Θ that maximizes the energy $J[v]$ over Θ. Let the line that contains v_1, v_2 be denoted L_v. Because of the assumption of uniqueness of u, u uniquely minimizes energy over all choices in Ω^2. Therefore v and u are distinct, even with respect to permutation. Let us denote the Voronoi regions of v_1 and v_2 by Ω_1 and Ω_2, respectively. Let the centroids of Ω_1 and Ω_2 be given by w_1 and w_2, respectively. We have $z = \alpha w_1 + (1-\alpha)w_2$ for some $0 < \alpha < 1$. Because v is a centroidal Voronoi tessellation of p_{L_v}, $(v_1 - w_1) \perp L_v$ and $(v_2 - w_2) \perp L_v$. Therefore, $\delta J[\varphi; u] = 0$ for all perturbations φ along the line L_v.

It suffices to show that $v_1 = w_1$ and $v_2 = w_2$. Because $\Delta(\Omega) < \infty$ and $p(x)$ is C^2, we have that for a sufficiently small angular perturbation of L_v about z, p_L is also perturbed a sufficiently small amount. Therefore, by Theorem 4.2, for any sufficiently small angular perturbation, there is an element of Θ on the perturbed line that is sufficiently close to v. This creates a local, convex $(N-1)$ dimensional subset M of Θ, that contains v and a unique element of Θ for every directional angular perturbation of L_v. However, v maximizes energy over this local convex set M, and therefore $\delta J[\varphi; u] = 0$ for all φ pointing along M.

Because $\delta J[\varphi; u] = 0$ for all perturbations φ along the line L_v, and the linearity of $\delta J[\varphi; v]$ with respect to φ, we have $\delta J[\varphi; v]$ equals zero for $\varphi_1 = w_1 - v_1$, $\varphi_2 = w_2 - v_2$. However, because w_1 and w_2 are the centroids of Ω_1 and Ω_2, respectively, we have

$$\int_{\Omega_1} \langle x - v_1, w_1 - v_1 \rangle p(x)dx \geq 0 \quad \text{and} \quad \int_{\Omega_2} \langle x - v_2, w_2 - v_2 \rangle p(x)dx \geq 0.$$

Therefore

$$\int_{\Omega_1} \langle x - v_1, w_1 - v_1 \rangle p(x)dx = \int_{\Omega_2} \langle x - v_2, w_2 - v_2 \rangle p(x)dx = 0,$$

which implies that $v_1 = w_1$ and $v_2 = w_2$. ☐

While the machinery in the proof of Theorem 4.3 does not extend easily to $K > 2$, it is likely that this result holds for general K. Finally, we make the following conjecture.

CONJECTURE 4.4. *There does not exist a unique centroidal Voronoi tessellation for generators and dimensions both greater than one.*

Acknowledgements. The author would like to thank Vadim Kaloshin and Ludmil Zikatanov for invaluable comments and discussion and Louisa Thomas for improving the style of the presentation. The author would also like to thank Alexander Rand for pointing out errors in an earlier draft.

Appendix A. $\lim_{\epsilon \to 0} M_\Omega[\epsilon, \varphi; u] = 0$ **.**

We note that for ϵ sufficiently small, the change in tessellation from $\{\Omega_i\}_{i=1}^K$ to $\{\Omega_i^*\}_{i=1}^K$ is of order ϵ, by the definition of Ω_i. Stated more precisely, we have the following lemma:

LEMMA A.1. *For sufficiently small* ϵ,

$$\sup_{x,y \in \Omega_i^* \cap \Omega_j} |\langle x - y, u_j - u_i \rangle| < M_1 \epsilon, \quad \lambda_N(\Omega_i^* \cap \Omega_j) < M_2 \epsilon, \quad \textit{and} \quad \lambda_{N-1}(\Omega_i \cap \Omega_j \cap \Omega_k^*) < M_3 \epsilon$$

for all $i \neq j \neq k$ *and some fixed* $M_1, M_2, M_3 < \infty$.

Proof. We first aim to bound $\sup_{x,y\in\Omega_i^*\cap\Omega_j}|\langle x-y,u_j-u_i\rangle|$. It suffices to show that this holds for all u^* such that $\varphi_j=0$ for all j. We can decompose φ_i into two components, the projection onto u_j-u_i and the component orthogonal to $u_{i\to j}$. The result of the projection component is that the hyperplanes containing $\Omega_{i,j}$ and Ω_{i^*,j^*}, respectively are parallel, and at most an ϵ distance apart. The orthogonal component has norm at most ϵ. By the boundedness of Ω, namely

$$\Delta(\Omega):=\max_{x,y\in\Omega}\|x-y\|<\infty,$$

the maximum distance between $\Omega_{i,j}$ and Ω_{i^*,j^*} along u_j-u_i is bounded above by $\frac{\epsilon\Delta(\Omega)}{\|u_j-u_i\|}$. Therefore, we have $\sup_{x,y\in\Omega_i^*\cap\Omega_j}|\langle x-y,u_j-u_i\rangle|<M\epsilon$, $M<\infty$. From here, the remaining two results come from noting that the N-dimensional volume of $\Omega_i^*\cap\Omega_j$ and the $(N-1)$-dimensional volume $\Omega_i\cap\Omega_j\cap\Omega_k^*$ are bounded by $\partial(\Omega)^{N-1}M\epsilon$ and $\partial(\Omega)^{N-2}M\epsilon$, respectively. □

The assumption on ϵ is non-restrictive, in that we are concerned with behavior in an ϵ-neighborhood of u as $\epsilon\to 0$. We are now prepared to prove the following:

LEMMA A.2. $\lim_{\epsilon\to 0}M_\Omega[\epsilon,\varphi;u]=0$

Proof. It suffices to show $|M_\Omega[\epsilon,\varphi;u]|<C\epsilon$, for some $C<\infty$. By Lemma A.1,

$$\begin{aligned}
&\int_{(\cup_{\ell\neq i,j}\Omega_{i^*,\ell,j})\cup(\cup_{\ell\neq i,j}\Omega_{i^*,\ell^*,j})}\frac{\langle x-\sigma_{i,j}^*,u_{i\to j}^*\rangle^2\langle u_{i\to j}^*,n\rangle}{\|u_{i\to j}^*\|^2}p(x)dx\\
&\quad\leq\int_{(\cup_{\ell\neq i,j}\Omega_{i^*,\ell,j})\cup(\cup_{\ell\neq i,j}\Omega_{i^*,\ell^*,j})}\frac{M_1^2\epsilon^2}{\|u_{i\to j}^*\|}p(x)dx\\
&\quad\leq\frac{M_1^2\epsilon^2\max_{x\in\Omega}p(x)}{\min_{i\neq j}\|u_{i\to j}^*\|}\Big[\sum_{\ell\neq i,j}\lambda_{N-1}(\Omega_{i^*,j,\ell})+\sum_{\ell\neq i,j}\lambda_{N-1}(\Omega_{i^*,j,\ell^*})\Big]\\
&\quad\leq\frac{M_1^2\epsilon^2\max_{x\in\Omega}p(x)}{\min_{i\neq j}\|u_{i\to j}^*\|}2(K-2)M_3\epsilon,
\end{aligned}$$

$$\begin{aligned}
\int_{\Omega_{i^*,j}}\frac{\langle x-\sigma_{i,j}^*,u_{i\to j}^*\rangle^2\langle u_{i\to j}^*,\nabla p(x)\rangle}{\|u_{i\to j}^*\|^2}dx&\leq\int_{\Omega_{i^*,j}}\frac{M_1^2\epsilon^2\|\nabla p(x)\|}{\|u_{i\to j}^*\|}dx\\
&\leq\frac{M_1^2\epsilon^2\max_{x\in\Omega}\|\nabla p(x)\|}{\min_{i\neq j}\|u_{i\to j}^*\|}\lambda_N(\Omega_{i^*,j})\\
&\leq\frac{M_1^2\epsilon^2\max_{x\in\Omega}\|\nabla p(x)\|}{\min_{i\neq j}\|u_{i\to j}^*\|}M_2\epsilon,
\end{aligned}$$

$$\begin{aligned}
\int_{\Omega_{i^*,i,j}}\frac{\langle u_{i\to j}^*,\varphi_j-\varphi_i\rangle\langle u_{i\to j}^*,x-\sigma_{i,j}^*\rangle^2}{\|u_{i\to j}\|\|u_{i\to j}^*\|^2}p(x)dx&\leq\int_{\Omega_{i^*,i,j}}\frac{\|\varphi_j-\varphi_i\|M_1^2\epsilon^2}{\|u_{i\to j}\|\|u_{i\to j}^*\|}p(x)dx\\
&\leq\frac{(\|\varphi_j\|+\|\varphi_i\|)M_1^2\epsilon^2\max_{x\in\Omega}p(x)}{\min_{i\neq j}\|u_{i\to j}\|\min_{i\neq j}\|u_{i\to j}^*\|}\lambda_{N-1}(\Omega_{i^*,i,j})\\
&\leq\frac{2M_1^2\epsilon^2\max_{x\in\Omega}p(x)}{\min_{i\neq j}\|u_{i\to j}\|\min_{i\neq j}\|u_{i\to j}^*\|}\Delta(\Omega)^{N-1},
\end{aligned}$$

$$\begin{aligned}
&\int_{\Omega_{i^*,i,j}}\frac{\langle\varphi_j-\varphi_i,\varphi_i+\varphi_j\rangle}{2\|u_{i\to j}\|}(\langle\varphi_j,x-u_j\rangle-\langle\varphi_i,x-u_i\rangle-\frac{\epsilon}{2}\langle\varphi_j-\varphi_i,\varphi_i+\varphi_j\rangle)p(x)dx\Big]\\
&\quad\leq\int_{\Omega_{i^*,i,j}}\frac{\|\varphi_j-\varphi_i\|\|\varphi_i+\varphi_j\|}{2\|u_{i\to j}\|}(\|\varphi_j\|\Delta(\Omega)+\|\varphi_i\|\Delta(\Omega)+\frac{\epsilon}{2}\|\varphi_j-\varphi_i\|\|\varphi_i+\varphi_j\|)p(x)dx\Big]\\
&\quad\leq\frac{(\|\varphi_j\|+\|\varphi_i\|)^2}{2\min_{i\neq j}\|u_{i\to j}\|}(2\Delta(\Omega)+\frac{\epsilon}{2}(\|\varphi_j\|+\|\varphi_i\|)^2)\lambda_{N-1}(\Omega_{i^*,i,j})max_{x\in\Omega}p(x)\\
&\quad\leq\frac{2(2\Delta(\Omega)+2\epsilon)\Delta(\Omega)^{N-1}max_{x\in\Omega}p(x)}{\min_{i\neq j}\|u_{i\to j}\|},
\end{aligned}$$

$$\int_{\cup_{\ell\neq i,j}\Omega_{i,j,\ell^*}} \frac{[\langle\varphi_j, x-u_j\rangle - \langle\varphi_i, x-u_i\rangle]^2}{\|u_{i\to j}\|} p(x)dx \leq \int_{\cup_{\ell\neq i,j}\Omega_{i,j,\ell^*}} \frac{[\|\varphi_j\|\Delta(\Omega) + \|\varphi_i\|\Delta(\Omega)]^2}{\|u_{i\to j}\|} p(x)dx$$
$$\leq \frac{[2\Delta(\Omega)]^2 \max_{x\in\Omega} p(x)}{\min_{i\neq j}\|u_{i\to j}\|} \sum_{\ell\neq i,j} \lambda_{N-1}(\Omega_{i,j,\ell^*})$$
$$\leq \frac{4\Delta(\Omega)^2 \max_{x\in\Omega} p(x)}{\min_{i\neq j}\|u_{i\to j}\|}(K-2)M_3\epsilon.$$

This gives us

$$|M_\Omega[\epsilon,\varphi;u]| \leq \sum_{i\neq j}\Bigg[\frac{2M_1^2M_3(K-2)\max_{x\in\Omega} p(x)}{\min_{i\neq j}\|u^*_{i\to j}\|} + \frac{M_1^2M_2\max_{x\in\Omega}\|\nabla p(x)\|}{\min_{i\neq j}\|u^*_{i\to j}\|}$$
$$+ \frac{2M_1^2\Delta(\Omega)^{N-1}\max_{x\in\Omega} p(x)}{\min_{i\neq j}\|u_{i\to j}\|\min_{i\neq j}\|u^*_{i\to j}\|} + \frac{4\Delta(\Omega)^{N-1}(\Delta(\Omega)+\epsilon) max_{x\in\Omega} p(x)}{\min_{i\neq j}\|u_{i\to j}\|}\Bigg]\epsilon$$
$$+ \sum_{e_{i,j}\in E} \frac{4M_3(K-2)\Delta(\Omega)^2\max_{x\in\Omega} p(x)}{\min_{i\neq j}\|u_{i\to j}\|}\epsilon.$$

We have $|M_\Omega[\epsilon,\varphi;u]| \leq C\epsilon$, where $C<\infty$ follows from $p(x), \|\nabla p(x)\| < \infty$ for all $x\in\Omega$, $\|u_j - u_i\|$ bounded away from zero, and $N, K, \Delta(\Omega) < \infty$. □

Appendix B. Resulting Sufficient Conditions for Uniqueness.

First we must restrict the domain $\mathcal{U}$. Consider the set of permutations acting on the elements $u_1, ..., u_K$, denoted $\mathcal{P}_K$. If $u\in\mathcal{U}$ is a minimum, then $\sigma(u)$ must also be a minimum for any $\sigma\in\mathcal{P}_K$. Therefore, we must redefine our concept of unique. We must either search for a centroidal Voronoi tessellation that is unique up to permutation, or consider an alternate domain which does not contain permutations of itself. Namely, we consider closed convex regions $\mathcal{W}\subset\Omega^K$ satisfying the following two conditions:

(i) $\sigma(\mathcal{W}^o)\cap\mathcal{W}=\emptyset$ for all non-trivial permutations $\sigma\in\mathcal{P}_K$,
(ii) $\cup_{\sigma\in\mathcal{P}_K}\sigma(\mathcal{W})=\Omega^K$.

The region $\mathcal{W}$ contains every element of $\mathcal{U}$ up to permutation, but does not contain any non-trivial permutations of itself. Therefore, every unlabeled set of K-quantizers appears only once.

In the scalar setting, this is easily done by setting

$$\mathcal{W} = \{u | u_i \leq u_j, \forall i<j\}.$$

In the multi-dimensional setting, there is no clear "best" choice of $\mathcal{W}$. To show uniqueness, the boundary of $\mathcal{W}$ must not contain any points that locally minimize energy. In one dimension, this is guaranteed, for the boundary consists exclusively of degenerate points.

We also note that while $\mathcal{U}$ is not convex, $\mathcal{W}$ is, allowing existing optimization theory to be applied. In addition, because $\mathcal{W}$ is closed, it necessarily contains the set of degenerate points on its boundary. We recall the following optimization lemma, from [7].

Lemma B.1. *Let C_1 be a connected open region in N-dimensional Euclidean space and let C be a convex closed region in C_1. Let $f(x)$ be a function defined in C_1 which has the following properties:*

(i) $\nabla f(x)$ exists and is continuous in C_1.
(ii) At every point where $\nabla f(x)=0$, the function attains a strict local minimum.
(iii) At every point on the boundary of C there exists a vector pointing into C along which the directional derivative of $f(x)$ is negative.

Then, in the region C, $f(x)$ possesses a unique stationary point. The point is interior to C; it is a relative and absolute minimum of $f(x)$ in C.

We are now prepared to give conditions for uniqueness.

Theorem B.2. *Suppose $\mathcal{W}$ is such that no element of $\partial\mathcal{W}$ locally minimizes energy with respect to $\mathcal{W}^o$, and for every $u\in\mathcal{W}$ such that*

$$u_i = \mu(\Omega_i)^{-1}\int_{\Omega_i} xp(x)dx$$

for all $i = 1, ..., K$, we have

$$\frac{J_i[u]}{2} \min_{x \in \Omega} \lambda_{min}(H[-\ln p(x)]) \geq (N-1)\mu(\Omega_i) + \sum_{j \sim i} \int_{\Omega_i \cap \Omega_j} \frac{\|x_i - \sigma_{i,j}\|^2}{\|u_i - \sigma_{i,j}\|} p(x)dx$$

for all $i = 1, ..., K$, with strict inequality holding for at least one i. Then $J[u]$ achieves a unique centroidal Voronoi tessellation on $\mathcal{W}$, which is also unique on $\mathcal{U}$ up to permutation. That centroidal Voronoi tessellation is the unique local and global minimizer of the quantization energy on $\mathcal{W}$, and also on $\mathcal{U}$ up to permutation.

Proof. By the assumption on $\partial\mathcal{W}$, we immediately have condition (iii) of Lemma B.1. By the assumptions of the theorem and Theorem 3.5, condition (ii) holds. We have continuity of $\delta J[\varphi; u]$ for all K-quantizers. We will show continuity of $\delta J[\varphi; u]$ for degenerate points in Appendix C. □

Fleischer's result for $N = 1$ (Theorem 1.2) is a direct corollary of Theorem B.2. However, as noted in Section 4, when $N > 1$ the given conditions are significantly more cumbersome. Further, by Theorem 4.3, the conditions of Theorem B.2 are not satisfiable for $K = 2$, and, as suggested by Conjecture 4.4, may indeed describe an empty set in general.

Appendix C. Continuity of $\delta J[\varphi; u]$ for Degenerate Points.

The continuity of $\delta J[\varphi; u]$ for degenerate points is immediate when φ points along a curve of degenerate points. Therefore, it suffices to consider $\delta J[\varphi; u]$ when $u^* = u + \epsilon\varphi$ is a K-quantizer, for ϵ sufficiently small. We will treat the case when u is a $(K-1)$-quantizer. All other cases of degeneracy follow immediately from successive application of the following analysis.

Without loss of generality, suppose the degenerate point of u is given by u_{K-1}, and the corresponding two points in u^* are given by u^*_{K-1} and u^*_K. Let us introduce the point $\hat{u} = u + \epsilon\hat{\varphi}$, where

$$\hat{\varphi}_i = \begin{cases} 0 & i = 1, ..., K-2 \\ \varphi_i & i = K-1, K \end{cases},$$

and the functional $\hat{J}[\hat{u}, u]$, given by

$$\hat{J}[\hat{u}, u] = \sum_{i=1}^{K-2} \int_{\Omega_i} \|x - \hat{u}_i\|^2 p(x)dx + \int_{\tilde{\Omega}_{K-1}} \|x - \hat{u}_{K-1}\|^2 p(x)dx + \int_{\tilde{\Omega}_K} \|x - \hat{u}_K\|^2 p(x)dx,$$

where $\tilde{\Omega}_{K-1}$ and $\tilde{\Omega}_K$ are the resulting tessellations from the generators $\hat{u}_{K-1}$ and $\hat{u}_K$ acting on the domain Ω_{K-1}.

We will decompose $\Delta J[\epsilon, \varphi; u]$ into three components, namely

$$\Delta J[\epsilon, \varphi; u] = [J[u^*] - J[\hat{u}]] + [J[\hat{u}] - \hat{J}[\hat{u}, u]] + [\hat{J}[\hat{u}, u] - J[u].$$

We aim to find the terms of order epsilon in each component. The component $J[u^*] - J[\hat{u}]$ is the difference in energy between two K-quantizers, and therefore

$$\begin{aligned} J[u^*] - J[\hat{u}] &= -2\epsilon \sum_{i=1}^{K-2} \int_{\Omega_i} \langle x - \hat{u}_i, \varphi_i \rangle p(x)dx + M_1[u^*, \hat{u}]\epsilon \\ &= -2\epsilon \sum_{i=1}^{K-2} \int_{\Omega_i} \langle x - u_i, \varphi_i \rangle p(x)dx + M_1[u^*, \hat{u}]\epsilon, \end{aligned}$$

where $M_1[u^*, \hat{u}] \to 0$ as $\epsilon \to 0$. The component $J[\hat{u}] - \hat{J}[\hat{u}, u]$ consists of the same generators, integrated over ϵ-varying subdomains. By Lemma 2.6 and the results of the previous Appendix,

$$J[\hat{u}] - \hat{J}[\hat{u}, u] = M_2[\hat{u}, u]\epsilon,$$

where $M_2[\hat{u}, u] \to 0$ as $\epsilon \to 0$. All that remains is to consider $\hat{J}[\hat{u}, u] - J[u]$. We have

$$\begin{aligned}\hat{J}[\hat{u}, u] - J[u] &= \int_{\tilde{\Omega}_{K-1}} \|x - \hat{u}_{K-1}\|^2 p(x)dx + \int_{\tilde{\Omega}_K} \|x - \hat{u}_K\|^2 p(x)dx - \int_{\Omega_{K-1}} \|x - u_{K-1}\|^2 p(x)dx \\ &= \int_{\tilde{\Omega}_{K-1}} [\|x - \hat{u}_{K-1}\|^2 - \|x - u_{K-1}\|^2] p(x)dx + \int_{\tilde{\Omega}_K} [\|x - \hat{u}_K\|^2 - \|x - u_{K-1}\|^2] p(x)dx \\ &= -2\epsilon \left[\int_{\tilde{\Omega}_{K-1}} \langle x - u_{K-1}, \varphi_{K-1} \rangle p(x)dx + \int_{\tilde{\Omega}_K} \langle x - u_{K-1}, \varphi_K \rangle p(x)dx \right] + M_3[\hat{u}, u]\epsilon,\end{aligned}$$

where $M_3[\hat{u}, u] \to 0$ as $\epsilon \to 0$. We are now prepared to give an explicit representation of $\delta J[\varphi; u]$ for degenerate points.

Theorem C.1. *Let $u \in \Omega^K$ be degenerate, namely*

$$u_1 = u_2 = ...u_{\ell(1)}, u_{\ell(1)+1} = u_{\ell(1)+2} = ... = u_{\ell(2)}, ..., u_{\ell(k-1)+1} = u_{\ell(k-1)+2} = ... = u_{\ell(k)},$$

with $\ell(k) = K$. Suppose $u^* = u + \epsilon\varphi$ is non-degenerate for all ϵ sufficiently small. Then

$$\delta J[\varphi; u] = -2 \sum_{i \in V} \int_{\tilde{\Omega}_i} \langle x - u_i, \varphi_i \rangle p(x)dx,$$

where $\{\tilde{\Omega}_i\}_{i=\ell(j-1)+1}^{\ell(j)}$ is the tessellation of the degenerate subregion corresponding to the j^{th} element degenerate k-quantizer by the elements $\{u^*\}_{i=\ell(j-1)+1}^{\ell(j)}$ as $\epsilon \to 0$.

REFERENCES

[1] Franz Aurenhammer, Rolf Klein, and Der-Tsai Lee. *Voronoi diagrams and Delaunay triangulations.* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2013.

[2] Qiang Du, Maria Emelianenko, and Lili Ju. Convergence of the Lloyd algorithm for computing centroidal Voronoi tessellations. *SIAM J. Numer. Anal.*, 44(1):102–119, 2006.

[3] Qiang Du, Vance Faber, and Max Gunzburger. Centroidal Voronoi tessellations: applications and algorithms. *SIAM Rev.*, 41(4):637–676, 1999.

[4] Qiang Du, Max Gunzburger, and Lili Ju. Advances in studies and applications of centroidal Voronoi tessellations. *Numer. Math. Theory Methods Appl.*, 3(2):119–142, 2010.

[5] Maria Emelianenko, Lili Ju, and Alexander Rand. Nondegeneracy and weak global convergence of the Lloyd algorithm in $\mathbb{R}^D$. *SIAM J. Numer. Anal.*, 46(3):1423–1441, 2008.

[6] David G. Feingold and Richard S. Varga. Block diagonally dominant matrices and generalizations of the Gerschgorin circle theorem. *Pacific J. Math.*, 12:1241–1250, 1962.

[7] P Fleischer. Sufficient conditions for achieving minimum distortion in a quantizer. *IEEE Int. Conv. Rec*, 12:104–111, 1964.

[8] I. M. Gelfand and S. V. Fomin. *Calculus of variations.* Revised English edition translated and edited by Richard A. Silverman. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

[9] Allen Gersho and Robert M. Gray. *Vector Quantization and Signal Compression.* Kluwer Academic Publishers, Norwell, MA, USA, 1991.

[10] R. M. Gray, J. C. Kieffer, and Y. Linde. Locally optimal block quantizer design. *Inform. and Control*, 45(2):178–198, 1980.

[11] John C. Kieffer. Uniqueness of locally optimal quantizer for log-concave density and convex error weighting function. *IEEE Trans. Inform. Theory*, 29(1):42–47, 1983.

[12] Stuart P. Lloyd. Least squares quantization in pcm. *IEEE Transactions on Information Theory*, 28:129–137, 1982.

[13] Alon Orlitsky. Scalar versus vector quantization: worst case analysis. *IEEE Trans. Inform. Theory*, 48(6):1393–1409, 2002. Special issue on Shannon theory: perspective, trends, and applications.

[14] Michael J. Sabin and Robert M. Gray. Global convergence and empirical consistency of the generalized Lloyd algorithm. *IEEE Trans. Inform. Theory*, 32(2):148–155, 1986.

[15] Xiaolin Wu. On convergence of Lloyd's method I. *IEEE Trans. Inform. Theory*, 38(1):171–174, 1992.