

---

# On Bounds for Diffusion, Discrepancy and Fill Distance Metrics

Steven B Damelin

Department of Mathematical Sciences, Georgia Southern University, PO Box 8093, Statesboro, GA, 30460  
damelin@georgiasouthern.edu

Criteria for optimally *discretizing* measurable sets in Euclidean space is a difficult and old problem which relates directly to the problem of good numerical integration rules or finding points of low discrepancy. On the other hand, *learning* meaningful descriptions of a finite number of given points in a measure space is an exploding area of research with applications as diverse as dimension reduction, data analysis, computer vision, critical infrastructure, complex networks, clustering, imaging neural and sensor networks, wireless communications, financial marketing and dynamic programming. The purpose of this paper is to show that a general notion of extremal energy as defined and studied recently by Damelin, Hickernell and Zeng on measurable sets  $\mathcal{X}$  in Euclidean space, defines a *diffusion metric* on  $\mathcal{X}$  which is equivalent to a *discrepancy* on  $\mathcal{X}$  and at the same time bounds the *fill distance* on  $\mathcal{X}$  for suitable measures with discrete support. The diffusion metric is used to *learn* via normalized graph Laplacian dimension reduction and the discrepancy is used to *discretize*.

Dedicated to Alexander Gorban, Andrei Zinovyev and their co-organisers for the wonderful international workshop on large data sets held at the University of Leicester in August 2006.

**AMS(MOS) Classification:** Primary 41A05; Secondary 41A15, 41A63

**Keywords:** Affinity Matrix, Diffusion Distance, Discrepancy, Distribution, Discretization, Dimension Reduction, Eigenfunction Expansion, Energy, Equilibrium measure, Euclidean space, Fill Distance, Graph, Laplacian, Kernels, Mesh Norm, Numerical Integration, Orthonormal Basis, Positive Definite.

## 1 Table of Contents

The outline of this paper is as follows:

- Introduction.
- Energy, discrepancy, distance and integration on measurable sets in Euclidean space.
- Set learning via normalized Laplacian dimension reduction and diffusion distance.
- Main result: Bounds for discrepancy, diffusion and fill distance metrics.

## 2 Introduction

Criteria for optimally *discretizing* measurable sets in Euclidean space is a difficult and old problem which relates directly to the problem of good numerical integration rules or finding points of low discrepancy. Indeed, let us consider the problem of uniformly distributing points on spheres, more generally, on compact sets in  $d \geq 1$  dimensional Euclidean space. It is folklore, that such problems were discussed already by Carl Friedrich Gauss in his famous *Disquisitiones arithmeticae*, although it is most likely that similar problems appeared in mathematical writings even before that time. For  $d \geq 2$ , let  $S^{d-1}$  denote the  $d$ -dimensional unit sphere in  $\mathbb{R}^d$ , given by

$$\mathbf{x}_1^2 + \cdots + \mathbf{x}_d^2 = 1.$$

For  $d = 2$ , the problem is reduced to uniformly distributing  $n \geq 1$  points on a circle, and equidistant points provide an obvious answer. For  $d \geq 3$ , the problem becomes much more difficult; in fact, there are numerous criteria for uniformity, resulting in different optimal configurations on the sphere. See (6; 3; 8; 9; 10; 12; 11) and the references cited therein.

On the other hand, *learning* meaningful descriptions of a finite number of given points in a measure space (set learning) is an exploding area of research with applications as diverse as dimension reduction, data analysis, computer vision, critical infrastructure, complex networks, clustering, imaging neural and sensor networks, wireless communications, financial marketing and dynamic programming. See (1)-(25) and the references cited therein.

The purpose of this paper is to show that a general notion of extremal energy as defined and studied recently by Damelin, Hickernell and Zeng on measurable sets  $\mathcal{X}$  in Euclidean space, defines a *diffusion metric* on  $\mathcal{X}$  which is equivalent to a *discrepancy* on  $\mathcal{X}$  and at the same time bounds the *fill distance* on  $\mathcal{X}$  for suitable measures with discrete support. The diffusion metric is used to *learn* via normalized graph Laplacian dimension reduction and the discrepancy is used to *discretize*.

The remainder of this paper is structured as follows. In Section 2, we introduce needed ideas of energy, discrepancy, integration and distance on measurable sets in Euclidean space. In Section 3, we discuss set learning via normalized Laplacian dimension reduction and diffusion distance. Finally, Section 4 is devoted to our main result on bounds for discrepancy, diffusion and fill distance metrics.

## 3 Energy, discrepancy, distance and integration on measurable sets in Euclidean space

Here and throughout, let  $\mathcal{X}$  be a measurable subset of  $d \geq 1$  Euclidean space  $\mathbb{R}^d$  and let  $\mathcal{M}(\mathcal{X})$  denote the space of all (non zero), finite signed measures (distributions)  $\mu$  on  $\mathcal{X}$  so that  $Q(\mu) := \int_{\mathcal{X}} d\mu$  exists and is finite. We will henceforth call  $Q(\mu)$  the total charge (mass) of  $\mu$ . If the space  $\mathcal{M}(\mathcal{X})$  is endowed with a norm  $\|\cdot\|_{\mathcal{M}(\mathcal{X})}$ , then the discrepancy problem measures the difference between any two measures in  $\mathcal{M}(\mathcal{X})$  in the norm  $\|\cdot\|_{\mathcal{M}(\mathcal{X})}$ . Following, Damelin, Hickernell and Zeng, (6), let  $K(\mathbf{x}, \mathbf{y})$  be a positive definite function on  $\mathcal{X} \times \mathcal{X}$ , where  $\mathcal{X} \subseteq \mathbb{R}^d$ . This means that  $\int_{\mathcal{X}^2} K(\mathbf{x}, \mathbf{y}) d\mu(\mathbf{x}) d\mu(\mathbf{y})$  exists, is finite and is positive for  $\mu \in \mathcal{M}(\mathcal{X})$ . Also we assume that  $K$  is symmetric, ie,  $K(\mathbf{x}, \mathbf{y}) = K(\mathbf{y}, \mathbf{x})$  for all  $\mathbf{x}, \mathbf{y} \in \mathcal{X}$  and  $\int_{\mathcal{X}} K(\mathbf{x}, \mathbf{y}) d\mu(\mathbf{y})$  exists. We call  $K$  an *energy kernel*, which means that the *potential field*  $\phi_{K,\mu}$  induced by the *charge distribution*  $\mu$  on  $\mathcal{X}$  is

$$\phi_{K,\mu}(\mathbf{x}) = \int_{\mathcal{X}} K(\mathbf{x}, \mathbf{y}) d\mu(\mathbf{y}), \mathbf{x} \in \mathcal{X}.$$

The *energy* of a charge distribution  $\mu \in \mathcal{M}(\mathcal{X})$  is

$$E_K(\mu) = \int_{\mathcal{X}^2} K(\mathbf{x}, \mathbf{y}) d\mu(\mathbf{x}) d\mu(\mathbf{y}),$$

and the *energy* of the charge distribution  $\mu$  in the field  $f_{K,\mu}(\mathbf{x}) = \int_{\mathcal{X}} K(\mathbf{x}, \mathbf{y}) d\mu_f(\mathbf{y})$  induced by the charge distribution  $\mu_f$  is

$$E_K(\mu, \mu_f) = \int_{\mathcal{X}} f(\mathbf{x}) d\mu(\mathbf{x}) = \int_{\mathcal{X}^2} K(\mathbf{x}, \mathbf{y}) d\mu(\mathbf{x}) d\mu_f(\mathbf{y}) = \langle \mu, \mu_f \rangle_{\mathcal{M}}.$$

Here we see that  $E_K(\mu, \mu_f)$  defines an inner product on the space  $\mathcal{M}$  of signed measures (charge distributions) for which the energy is finite. Sometimes we need only assume that  $K$  is *conditionally positive definite*, meaning

$$\int_{\mathcal{X}^2} K(\mathbf{x}, \mathbf{y}) d\mu(\mathbf{x}) d\mu(\mathbf{y}) > 0 \quad \forall \mu \neq 0 \text{ with } Q(\mu) = 0.$$

For conditionally positive definite kernels the energy  $E_K(\mu)$  may be negative. The energy  $E_K(\mu, \mu_f)$  may be of either sign in general, even for positive definite kernels. In what follows,  $|\cdot|$  and  $|\cdot|_2^2$  denotes the Euclidean and squared  $L_2$  norms in  $\mathbb{R}^d$  respectively.

We also call  $K$  the *reproducing kernel* of a Hilbert space,  $H(K)$  which is a Hilbert space of functions  $f : \mathcal{X} \rightarrow \mathbb{R}$ . This means that  $K(\cdot, \mathbf{y})$  is the representer of the linear functional that evaluates  $f \in H(K)$  at  $\mathbf{y}$ :

$$f(\mathbf{y}) = \langle K(\cdot, \mathbf{y}), f \rangle_{H(K)} \quad \forall f \in H(K), \mathbf{y} \in \mathcal{X}.$$

For any  $f, g \in H(K)$  with  $f(\mathbf{x}) = \int_{\mathcal{X}} K(\mathbf{x}, \mathbf{y}) d\mu_f(\mathbf{y})$  and  $g(\mathbf{x}) = \int_{\mathcal{X}} K(\mathbf{x}, \mathbf{y}) d\mu_g(\mathbf{y})$  it follows that their inner product is the energy of the two corresponding charge distributions:

$$\langle f, g \rangle_{H(K)} = E_K(\mu_f, \mu_g) = \int_{\mathcal{X}^2} K(\mathbf{x}, \mathbf{y}) d\mu_f(\mathbf{x}) d\mu_g(\mathbf{y}) = \langle \mu_f, \mu_g \rangle_{\mathcal{M}}$$

Note that a crucial feature of the function space  $H(K)$  is that it depends directly on the kernel  $K$ .

In **electrostatics**, a positive charge  $\mu$  placed upon a conductor (compact set) will distribute itself so as to minimize its energy. Equilibrium will be reached when the total energy is minimal amongst all possible charge distributions on  $\mathcal{X}$ . One would thus expect that the potential field  $\int_{\mathcal{X}} K(x, y) d\mu(\mathcal{Y})$  should be **constant** on compact  $\mathcal{X}$  most of the time for otherwise charge would flow from one point of  $\mathcal{X}$  to the next disturbing the equilibrium. The kernel  $K$  describes the interaction of the electrons (positive charges) on the conductor  $\mathcal{X}$  and a charge distribution which gives a constant potential field is an example of an equilibrium measure. More precisely, we have, see (6):

**Theorem 1.** *Let  $K$  be an energy kernel. Then*

$$E_K(\mu) = \int_{\mathcal{X}^2} K(\mathbf{x}, \mathbf{y}) d\mu(\mathbf{x}) d\mu(\mathbf{y}) \geq \frac{[Q(\mu)]^2}{C_K(\mathcal{X})}, \quad \mu \in \mathcal{M}(\mathcal{X})$$

for a constant  $C_K(\mathcal{X})$  depending only on  $\mathcal{X}$  and  $K$  (called the *capacity* of  $\mathcal{X}$ ) and with equality holding for any equilibrium charge distribution or *equilibrium measure*,  $\mu_{e,K}$ , defined as one that induces a constant field,

$$\phi_{K, \mu_{e,K}}(\mathbf{x}) = \int_{\mathcal{X}} K(\mathbf{x}, \mathbf{y}) d\mu_{e,K}(\mathbf{y}) = \frac{Q(\mu_{e,K})}{C_K(\mathcal{X})} \quad \forall \mathbf{x} \in \mathcal{X}.$$

1

Examples of energy kernels are:

- (a) The weighted *Riesz kernel* on compact subsets of  $d \geq 1$  Euclidean space  $\mathbb{R}^d$

$$K_{s,w}(\mathbf{x}, \mathbf{y}) = \begin{cases} w(\mathbf{x}, \mathbf{y})|\mathbf{x} - \mathbf{y}|^{-s}, & 0 < s < d, \quad \mathbf{x}, \mathbf{y} \in \mathcal{X}, \\ -w(\mathbf{x}, \mathbf{y}) \log |\mathbf{x} - \mathbf{y}|, & s = 0, \quad \mathbf{x}, \mathbf{y} \in \mathcal{X}, \\ w(\mathbf{x}, \mathbf{y})(c - |\mathbf{x} - \mathbf{y}|^{-s}), & -1 \leq s < 0, \quad \mathbf{x}, \mathbf{y} \in \mathcal{X} \end{cases}$$

where  $w : \mathcal{X} \times \mathcal{X} \rightarrow (0, \infty)$  is chosen to be lower semi-continuous such that  $K$  is an energy kernel. Such kernels (in the case  $w \equiv 1$ ,  $s > 0$ ), arise naturally in describing the distributions of electrons on rectifiable manifolds such as the sphere  $S^d$ . The case when  $w$  is active, comes about for example in problems in computer modeling in which points are not necessarily uniformly distributed on  $\mathcal{X}$ . See (3; 8; 9; 10; 12). The case when  $-1 \leq s < 0$  appears more frequently in discrepancy theory. Here  $c$  is chosen so that the kernel is positive definite.

- (b) If  $\rho(\text{dist}_K(\mathbf{x}, \mathbf{y}))$  is conditionally negative semi-definite and  $\rho(0) = 0$ , then  $K(\mathbf{x}, \mathbf{y}) = \Psi(\rho(\text{dist}_K(\mathbf{x}, \mathbf{y})))$  is an energy kernel for any non constant, completely monotonic function  $\Psi$  on  $\mathcal{X}$  where  $\text{dist}_K$  is a metric on  $\mathcal{X}^2$  which is defined by way of (2) below. For example, typical examples of such kernels are the heat kernel  $\exp(-c|\mathbf{x} - \mathbf{y}|_2^2)$ ,  $c > 0$  on  $\mathcal{X}$  and Hamming distance kernel used in the construction of linear codes when well defined.

In what follows, we will assume that our kernels are finite. Good numerical integration rules or good **discretizations** of  $\mathcal{X}$  are obtained generally by optimizing the placement of points and weights assigned to function values on  $\mathcal{X}$ . In the classical setting, the domain of integration is the interval  $[-1, 1]$  and as is known, the nodes of the celebrated Gaussian quadrature formula, may be uniquely determined by the following characteristic property of the nodes of an  $n \geq 1$  point Gauss quadrature: The  $n$  nodes are the zeros of the unique monic polynomial of minimal mean-square deviation on a real interval. In other words, the nodes are the zeros of the unique solution of an extremal problem. In particular, recent work of Damelin and his collaborators, see (3; 6; 8; 9) has shown that minimal energy points are good for integrating functions with respect to equilibrium distributions. More precisely, in numerical integration or cubature we approximate the integral of a function  $f \in H(K)$ ,

$$I(f; \tilde{\mu}) = \int_{\mathcal{X}} f(\mathbf{x}) d\tilde{\mu}(\mathbf{x}) = E_K(\tilde{\mu}, \mu_f) = \langle \tilde{\mu}, \mu_f \rangle_{\mathcal{M}} = \langle \phi_{\tilde{\mu}}, f \rangle_{H(K)}$$

by the cubature rule

$$I(f; \hat{\mu}) = \int_{\mathcal{X}} f(\mathbf{x}) d\hat{\mu}(\mathbf{x}) = \sum_{i=1}^n c_i f(\mathbf{x}_i) = E_K(\hat{\mu}, \mu_f) = \langle \hat{\mu}, \mu_f \rangle_{\mathcal{M}} = \langle \phi_{\hat{\mu}}, f \rangle_{H(K)}$$

<sup>1</sup> Note that the capacity depends of course on the dimension  $d$  but since  $d$  is fixed when  $\mathcal{X}$  is chosen, this dependence is suppressed for notational convenience. Also the equilibrium measure depends on  $\mathcal{X}$  and  $d$  but again it is supported on  $\mathcal{X}$  by definition so we suppress this notation again. We adopt the same convention for other qualities as well throughout this paper which will be clear to the reader.

where  $\hat{\mu}$  is the charge distribution (signed measure) with support on the set of  $n$  points,  $\mathbf{x}_1, \dots, \mathbf{x}_n$  and charge  $c_i$  at each point  $f_i$ . Moreover,

$$\phi_{\tilde{\mu}}(\mathbf{x}) = \int_{\mathcal{X}} K(\mathbf{x}, \mathbf{y}) d\tilde{\mu}(\mathbf{y})$$

is the representer of the integration functional, and

$$\phi_{\hat{\mu}}(\mathbf{x}) = \int_{\mathcal{X}} K(\mathbf{x}, \mathbf{y}) d\hat{\mu}(\mathbf{y}) = \sum_{i=1}^n c_i K(\mathbf{x}, \mathbf{x}_i)$$

is the representer of the cubature rule functional. The error of this numerical approximation is

$$\begin{aligned} \int_{\mathcal{X}} f(\mathbf{x}) d\tilde{\mu}(\mathbf{x}) - \sum_{i=1}^n c_i f(\mathbf{x}_i) &= I(f; \tilde{\mu}) - I(f; \hat{\mu}) = \int_{\mathcal{X}} f(\mathbf{x}) d[\tilde{\mu} - \hat{\mu}](\mathbf{x}) \\ &= E_K(\tilde{\mu} - \hat{\mu}, \mu_f) = \langle \tilde{\mu} - \hat{\mu}, \mu_f \rangle_{\mathcal{M}} = \langle \phi_{\tilde{\mu}} - \phi_{\hat{\mu}}, f \rangle_{H(K)}. \end{aligned}$$

The *worst-case integration error* is defined as the largest absolute value of this error for integrands,  $f$ , with unit norm. By the Cauchy-Schwartz inequality we see that this occurs when  $f$  is parallel to  $\phi_{\tilde{\mu}} - \phi_{\hat{\mu}}$ , or equivalently,  $\mu_f$  is parallel to  $\tilde{\mu} - \hat{\mu}$ . Thus, we have, see (6):

**Theorem 2.**

$$\begin{aligned} D_K(\tilde{\mu}, \hat{\mu}) &:= \min_{\|f\|_{H(K)} \leq 1} \left| \int_{\mathcal{X}} f(\mathbf{x}) d\tilde{\mu}(\mathbf{x}) - \sum_{i=1}^n c_i f(\mathbf{x}_i) \right| \quad (1) \\ &= \sqrt{E_K(\tilde{\mu} - \hat{\mu})} = \|\tilde{\mu} - \hat{\mu}\|_{\mathcal{M}} = \|\phi_{\tilde{\mu}} - \phi_{\hat{\mu}}\|_{H(K)} \\ &= \left\{ \int_{\mathcal{X}^2} K(\mathbf{x}, \mathbf{y}) d\tilde{\mu}(\mathbf{x}) d\tilde{\mu}(\mathbf{y}) - 2 \sum_{i=1}^n c_i \int_{\mathcal{X}} K(\mathbf{x}_i, \mathbf{y}) d\tilde{\mu}(\mathbf{y}) \right. \\ &\quad \left. + \sum_{i,k=1}^n c_i c_k K(\mathbf{x}_i, \mathbf{x}_k) \right\}^{1/2} \end{aligned}$$

The quantity  $D_K(\tilde{\mu}, \hat{\mu})$ , defined by (1) which depends both on the placement and magnitude of the point charges defining  $\hat{\mu}$ , is called the *discrepancy*. We see that it is equivalent to the square root of an energy.

For a fixed choice of points  $\mathcal{Y} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ , the best cubature rule, i.e., the choice of  $c_i$  that minimizes the discrepancy, is obtained by choosing the potential induced by  $\hat{\mu}$  to match the potential induced by  $\tilde{\mu}$  on  $\mathcal{Y}$ , i.e.,

$$\phi_{\hat{\mu}}(\mathbf{x}_i) = \phi_{\tilde{\mu}}(\mathbf{x}_i), \quad i = 1, \dots, n.$$

In this case

$$D_K(\tilde{\mu}, \hat{\mu}) = \{E_K(\tilde{\mu}) - E_K(\hat{\mu})\}^{1/2}.$$

The best choice of locations and magnitude of the charges is to find the set  $\mathcal{Y}$  consisting of  $n$  points that has *maximum energy* under the given constraint.

It is now possible to define a distance on  $\mathcal{X}$  by way of:

$$\text{dist}_K(\mathbf{x}, \mathbf{y}) := \sqrt{K(\mathbf{x}, \mathbf{x}) - 2K(\mathbf{x}, \mathbf{y}) + K(\mathbf{y}, \mathbf{y})}, \mathbf{x}, \mathbf{y} \in \mathcal{X}. \quad (2)$$

In the next section, we will show that the distance defined in (2), for certain energy kernels, arises as a diffusion metric which may be used to *learn*  $\mathcal{X}$  via normalized graph Laplacian dimension reduction. Our main result in Section 4 shows that this later distance is essentially equivalent to the discrepancy defined by way of (2).

## 4 Set learning via normalized Laplacian dimension reduction and diffusion distance

In this section, we show that the distance defined by way (2) arises in a natural way in set learning. See (2). In what follows, we will need to define a class of energy kernels which we will call *admissible* in the sense below. Suppose first that  $K$  is non-negative. Next, suppose that  $K$  is zonal in the sense that it depends only on the Euclidean distance between two points  $\mathbf{x}$  and  $\mathbf{y}$  in  $\mathcal{X}$ . For example, Damelin, Leversley and Sun have shown in (8) that if  $\mathcal{X}$  is a compact  $d$  dimensional homogenous reflexive manifold with invariant group  $G$ , then if  $K$  is invariant under actions of  $G$ , it is zonal. Suppose that  $\mathcal{X}$  is a finite collection of points which are the vertices of an oriented graph and if  $b(\mathbf{x}, \mathbf{y})$  denotes the associated adjacency matrix, then we assume that  $b(\mathbf{x}, \mathbf{y}) = 1$  if there is an edge going from  $\mathbf{x}$  to  $\mathbf{y}$  and  $b(\mathbf{x}, \mathbf{y}) = 0$  otherwise. Let  $\hat{\mu}$  be a positive counting measure with support  $\mathcal{X}$ .  $K$  is now scaled in the following way. First, normalize  $K$  to be stochastic (to have sum 1 along its rows) by setting

$$v^2(x) = \int_{\mathcal{X}} K(\mathbf{x}, \mathbf{y}) d\hat{\mu}(\mathbf{y}), \mathbf{x} \in \mathcal{X}$$

and defining

$$a(\mathbf{x}, \mathbf{y}) = \frac{K(\mathbf{x}, \mathbf{y})}{v(\mathbf{x})v(\mathbf{y})}, \mathbf{x}, \mathbf{y} \in \mathcal{X}.$$

Also define  $A : L_2(\hat{\mu}) \rightarrow L_2(\hat{\mu})$  by

$$Af(\mathbf{x}) := \int_{\mathcal{X}} a(\mathbf{x}, \mathbf{y})f(\mathbf{y})d\hat{\mu}(\mathbf{y}), \mathbf{x} \in \mathcal{X}.$$

Then  $A$  is bounded, has norm 1, and positive definite. As is well known, we may also assume that  $A$  is symmetric. Assuming that  $A$  is compact, we may also write

$$a(\mathbf{x}, \mathbf{y}) = \sum_{j \geq 0} \lambda_j \phi_j(\mathbf{x})\phi_j(\mathbf{y}), \mathbf{x}, \mathbf{y} \in \mathcal{X}$$

where

$$A\phi_j(\mathbf{x}) = \lambda_j \phi_j(\mathbf{x}), \mathbf{x} \in \mathcal{X}$$

where the  $\lambda_j$  are discrete, non-increasing and non-negative. Given  $m \geq 1$ , following (23), we define  $A^{(m)}(x, y)$  to be the kernel for  $A^m$  so that

$$a^{(m)}(\mathbf{x}, \mathbf{y}) = \sum_{j \geq 0} \lambda_j^m \phi_j(\mathbf{x})\phi_j(\mathbf{y}), \mathbf{x}, \mathbf{y} \in \mathcal{X}.$$

At the level of the points in  $\mathcal{X}$ ,  $a^{(m)}$  denotes the probability for a Markov chain with transition matrix  $a$  to reach  $\mathbf{y}$  from  $\mathbf{x}$  in  $m$  steps. The mapping

$$\Psi(\mathbf{x}) := (\psi_0(\mathbf{x}), \psi_1(\mathbf{x}), \psi_2(\mathbf{x}), \dots)^T$$

maps  $\mathcal{X}$  into an Euclidean space and each eigenfunction is a coordinate on this set. Thus the mapping provides a dimension controlled representation of the points in  $\mathcal{X}$  in Euclidean space.

Define for each  $m$  on  $\mathcal{X}$ ,  $\text{dist}_{a^{(m)}}(\cdot)$  as given by (2). Then  $\text{dist}_{a^{(m)}}(\cdot)$  is a diffusion distance which measures the rate of connectivity between points on the data set  $\mathcal{X}$ .

## 5 Main result: Bounds for discrepancy, diffusion and fill distance metrics

Following is our main result.

**Theorem 3.** *Let  $\mathcal{X}$  be a finite measure space in  $\mathbb{R}^d$  for some  $d \geq 1$  consisting of a finite collection of  $n \geq 1$  points which we label as  $\mathbf{x}_i$ ,  $i = 1, \dots, n$ . Let  $\hat{\mu}$  be a positive counting measure with support  $\mathcal{X}$  and let  $\tilde{\mu}$  be any measure in  $\mathcal{M}(\mathcal{X})$ . Let  $m \geq 1$ , let  $K : \mathcal{X}^2 \rightarrow \mathbb{R}$  be admissible and define  $\text{dist}_{a^{(m)}}(\cdot)$  as given by (2) and Section 3. Then the following hold:*

(a)

$$\begin{aligned} & D_{a^{(m)}}^2(\tilde{\mu}, \hat{\mu}) \\ &= -\frac{1}{2} \int_{\mathcal{X}^2} \text{dist}_{a^{(m)}}^2(\mathbf{x}, \mathbf{y}) d\tilde{\mu}(\mathbf{x}) d\tilde{\mu}(\mathbf{y}) \\ &+ \frac{1}{n} \sum_{i=1}^n \int_{\mathcal{X}} \text{dist}_{a^{(m)}}^2(\mathbf{x}_i, \mathbf{x}) d\tilde{\mu}(\mathbf{x}) - \frac{1}{2n^2} \sum_{ij=1}^n \text{dist}_{a^{(m)}}^2(\mathbf{x}_i, \mathbf{x}_j) \\ &\leq \text{fill}_{a^{(m)}}^2(\hat{\mu}) \end{aligned} \tag{1}$$

where

$$\text{fill}_{a^{(m)}}^2(\hat{\mu}) := \sup_{\mathbf{x} \in \mathcal{X}} \min_{\mathbf{y} \in \mathcal{X}} (\text{dist}_{a^{(m)}}(\mathbf{y}, \mathbf{x}))$$

is the fill distance or mesh norm of  $\mathcal{X}$ .

(b)

$$\text{dist}_{a^{(m)}}^2(\mathbf{x}, \mathbf{y}) = \sum_{j \geq 0} \lambda_j^m (\phi_j(\mathbf{x}) - \phi_j(\mathbf{y}))^2, \mathbf{x}, \mathbf{y} \in \mathcal{X}.$$

Part (a) of the theorem gives an equivalence between discrepancy on  $\mathcal{X}$  and diffusion distance. Notice that  $\tilde{\mu}$  can be arbitrary in  $\mathcal{M}(\mathcal{X})$ . Part (b) says that discrepancy or diffusion metrics can be computed as weighted Euclidean distance in the embedded space of  $\Psi$ , the weights being precisely the eigenvalues of the operator  $A$ .

**Proof** Part (a) of the theorem follows from Theorem(2), the definition of  $A$  and (2). Part (b) follows from (2).  $\square$

**Acknowledgement** S. B. Damelin was supported, in part, by EP/C000285, NSF-DMS-0439734 and NSF-DMS-0555839. This work began while S. B. Damelin was on sabbatical at the Institute for Mathematics and Applications, University of Minnesota during the 2005-2006 academic year.

## References

- [1] B. Bajnok, S.B. Damelin, J. Li, and G. Mullen, *A constructive finite field method for scattering points on the surface of a  $d$ -dimensional sphere*, *Computing*, **68** (2002), pp 97–109.
- [2] F. Chung, *Spectral Graph Theory*, in CBNS-AMS, **92**, Amer. Math.Soc.Providence, RI.
- [3] S. B. Damelin and P. Grabner, *Numerical integration, energy and asymptotic equidistribution on the sphere*, *Journal of Complexity*, **19** (2003), pp 231-246.
- [4] S.B. Damelin, Marcinkiewicz-Zygmund inequalities and the Numerical approximation of singular integrals for exponential weights: Methods, Results and Open Problems, some new, some old; *Journal of Complexity*, 19(2003), pp 406-415.
- [5] S. B. Damelin, A. Hero and S. J. Wang, *Kernels, Hitting time metrics and diffusion in clustering and data analysis*, in preparation.
- [6] S. B. Damelin, F. Hickernell and X. Zeng, *On the equivalence of discrepancy and energy on closed subsets of Euclidean space*, in preparation.
- [7] S.B. Damelin and A. Kuijlaars, The support of the extremal measure for monomial external fields on  $[-1, 1]$ , *Trans.Amer.Math. Soc.* 351 (1999), 4561-4584.
- [8] S. B. Damelin, J. Levesley and X. Sun, *Energies, Group Invariant Kernels and Numerical Integration on Compact Manifolds*, submitted to *Journal of Complexity*.
- [9] S. B. Damelin, J. Levesley and X. Sun, *Energy estimates and the Weyl criterion on compact homogeneous manifolds*, *Algorithms for Approximation V*, Springer, to appear.
- [10] S. B. Damelin and V. Maymeskul, *On Point Energies, Separation Radius and Mesh Norm for  $s$ -Extremal Configurations on Compact Sets in  $\mathbb{R}^n$* , *Journal of Complexity*, Volume 21(6)(2005), pp 845-863.
- [11] S. B. Damelin and V. Maymeskul, *Minimal discrete energy problems and numerical integration on compact sets in Euclidean spaces*, *Algorithms for Approximation V*, Springer, to appear.
- [12] S. B. Damelin and V. Maymeskul, *On regularity and dislocation properties of minimal energy configurations on compact sets*, submitted.
- [13] Q. Du, V. Faber and M. Gunzburger, *Centroidal Voronoi tessellations: applications and algorithms*, *SIAM Rev*, **41**(4), pp 637-676.
- [14] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces (Graduate Studies in Mathematics) (Hardcover)*, American Mathematical Society (July 2001).
- [15] F. J. Hickernell, *Goodness of fit statistics, discrepancies and robust designs*, *Statist. Probab. Lett.* **44**, pp 73-78.
- [16] F. J. Hickernell, *What affects the accuracy of quasi-Monte Carlo quadrature?*, in *Monte Carlo and Quasi-Monte Carlo Methods 1998* (H. Niederreiter and J. Spanier, eds.), Springer-Verlag, Berlin, 2000, pp. 16-55.
- [17] F. J. Hickernell, *A generalized discrepancy and quadrature error bound*, *Mathematics of Computation*, **67**(221), pp 299-322.
- [18] F. J. Hickernell, *An algorithm driven approach to error analysis for multidimensional integration*, submitted.
- [19] F. J. Hickernell, *My dream quadrature rule*, *J. Complexity*, 19 (2003), pp. 420-427.
- [20] W. C. Huffman and V. Pless, *Fundamentals of error-correcting codes*, Cambridge University Press, Cambridge.
- [21] M. E. Johnson, L. M. Moore and D. Ylvisacker, *Minimax and maxmin distance designs*, *J. Statist. Plann. Inference* **26**, pp 131-148.
- [22] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley-Interscience, New York, 1974.



- [23] S. Lafon and R. R. Coifman, *Diffusion maps*, Applied and Computational Harmonic Analysis, **21**(2006), pp 5-30.
- [24] H. Niederreiter, *Random Number Generation and Quasi- Monte Carlo Methods*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia.
- [25] A. Lubotzky, R. Phillips, and P. Sarnak, *Hecke operators and distributing points on the sphere (I-II)*, Comm. Pure App. Math. **39-40** (1986,1987), 148-186, 401-420.