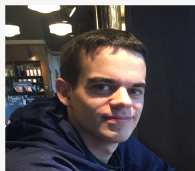


Approximating 1-Qubit Gates: Energy and Discrepancy

Steven Damelin²
Joint work with: Alec Greene¹, QingZhong Liang¹

¹University of Michigan, ²The American Math Society: Homepage:
<http://www.umich.edu/~damelin>

Computation and Optimization of Energy, Packing, and Covering
ICERM, April 10, 2018



Thanks to the organizers for this lucky invitation, in particular:
Ed Saff, Rob Womersley and Peter Grabner. Thanks to my
collaborators and students who did 99 percent of the work.
Alec and QingZhong are my collaborators on the Quantum
gate part of this talk.

Some common themes

- Interacting particle problem:

Some common themes

- ▶ **Interacting particle problem:**
- ▶ **Energy and discrepancy:**

Some common themes

- ▶ Interacting particle problem:
- ▶ Energy and discrepancy:
- ▶ Finite fields, designs, codes:

Some common themes

- ▶ Interacting particle problem:
- ▶ Energy and discrepancy:
- ▶ Finite fields, designs, codes:
- ▶ Quantum gate problem:

Thompson

The "Thompson" problem aims to understand a well defined periodic table by packing repulsively interacting particles-electrons on a set with some topology. When formulated for Euclidean space (some metric), the problem is that of finding a collection (or packing) of non-overlapping equal balls with the largest density in space for example a d -dimensional sphere S^d or a d -dimensional ball/torus.

The problem can be formulated generally and has deep applications in many areas for example Crystal structure, Special Functions, Orthogonal Polynomials, Random Matrix Theory, Integer Lattices, Cryptography and distribution of primes, Discrepancy, Computer Vision, Learning theory, Network design (for example on classes of Riemannian manifolds), Designs-codes, approximation theory and many others.

Interacting particle-1

In Euclidean space: Choose a repulsive potential and then minimize the energy of $N \geq 2$ particles-electrons that are constrained to S^d for some $d \geq 1$ and interact pairwise with this potential with Euclidean metric. What can be said about the minimizing configurations? In particular, we may ask about **regularity and dislocation** properties of these later configurations. A related problem is the Wigner crystal where one considers positively charged particles in a uniform background of negative charge so that the whole system is neutral. One expects that the energy minimizing configuration forms a crystal.

In electrostatics, a positive charge μ placed upon a conductor will distribute itself so as to minimize its energy. Equilibrium will be reached when the total energy is minimal amongst all possible charge distributions.

Discrepancy-Energy

- S. B. Damelin, F. Hickernell, D. Ragozin and X. Zeng, On energy, discrepancy and G invariant measures on measurable subsets of Euclidean space, Journal of Fourier Analysis and its Applications (2010) (16), pp 813-839.

Discrepancy-Energy

- ▶ S. B. Damelin, F. Hickernell, D. Ragozin and X. Zeng, On energy, discrepancy and G invariant measures on measurable subsets of Euclidean space, Journal of Fourier Analysis and its Applications (2010) (16), pp 813-839.
- ▶ S. B. Damelin, J. Levesley, D. L. Ragozin and X. Sun, Energies, Group Invariant Kernels and Numerical Integration on Compact Manifolds, Journal of Complexity, 25(2009), pp 152-162.

Discrepancy-Potential

Henceforth, X is a measurable subset of Euclidean space and let $\mathcal{M}(X)$ denote the space of all (non zero) signed Borel measures (distributions) μ on X with finite charge (mass) $Q(\mu) := \int_X d\mu$. If the space $\mathcal{M}(X)$ is endowed with a norm $\|\cdot\|_{\mathcal{M}(X)}$, then the discrepancy problem measures the difference between any two measures in $\mathcal{M}(X)$ in the norm $\|\cdot\|_{\mathcal{M}(X)}$.

Let $K : X^2 \rightarrow \mathbb{R}$ be a positive definite function. This means that $\int_{X^2} K(x, y) d\mu(x) d\mu(y)$ exists, is finite and is positive for $\mu \in \mathcal{M}(X)$. Also, we assume that K is symmetric, ie, $K(x, y) = K(y, x)$ for all $x, y \in X$. We call K an *energy kernel*, which means that the *potential field* $\phi_{K, \mu}$ induced by the *charge distribution* μ on X exists and is given by

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

Energy-2

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

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

Sometimes we need to assume that K is *conditionally positive definite*, meaning

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

For conditionally positive definite kernels the energy $E_K(\mu)$ may be negative.

Example in Learning

- Sung Jin Hwang, Steven B. Damelin, Alfred O. Hero III, Shortest Path through Random Points, The Annals of Applied Probability, 2016, Vol. 26, No. 5, pp 2791-2823.

Example in Learning

- ▶ Sung Jin Hwang, Steven B. Damelin, Alfred O. Hero III, Shortest Path through Random Points, The Annals of Applied Probability, 2016, Vol. 26, No. 5, pp 2791-2823.
- ▶ Raviv Raich, Jose A. Costa, Steven B. Damelin, Alfred O. Hero, Classification Constrained Dimensionality Reduction, arxiv: 0802.2906.

Nowdays, we are constantly flooded with information of all sorts and forms and a common denominator of data analysis in many emerging fields of current interest are large amounts of measurable observations X that may sit or lie near or on a manifold embedded in some high dimensional Euclidean space. Think of X as a discrete metric space. We call this the "Manifold hypothesis problem".

For example the data could be the frames of your favorite movie produced by a digital camera or the pixels of a hyperspectral image in a computer vision problem or unlabelled face recognition labels.

Kernel correlation

The crux of the matter is the following essential observation. Given a discrete set X of data, there is often a (local or global) correlation between the members of X which is defined by way of an energy kernel. Examples of energy kernels which arise in this way:

The weighted *Riesz/Newtonian kernel* on d dimensional compact subsets of $\mathbb{R}^{d'}$, $d' \geq d \geq 1$

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

where $w : (R^{d'})^2 \rightarrow (0, \infty)$ is chosen such that K is an energy kernel.

Active Newtonian

The case when w is active, comes about for example in problems in computer modeling in which points are not necessarily uniformly distributed. The case when $-1 \leq s < 0$ appears more frequently in discrepancy theory. Here c is chosen so that the kernel is positive definite.

For a suitable action ρ , if $\rho(\text{dist}_K(x, y))$ is conditionally negative semi-definite and $\rho(0) = 0$, then $\Psi(\rho(\text{dist}_K(x, y)))$ is an energy kernel for any non constant, completely monotonic function Ψ on $\mathbb{R}^{d'}$ where dist_K is a suitable metric on $(R^{d'})^2$.

Examples-Brachial Plexus-Hamming-Codes-A sample of some papers: See my homepage.

- For example, typical examples of such kernels are the heat kernel $\exp(-c|x - y|^2)$, $c > 0$ on X and certain Hamming distance kernels used in the construction of linear codes when well defined.

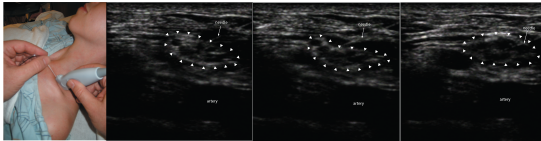
- J. H. Ann, S. B. Damelin and P. Bigeleisen, Medical Image segmentation using modified Mumford segmentation methods, Ultrasound-Guided Regional Anesthesia and Pain Medicine, eds P. Bigeleisen, Chapter 40, Birkhauser.

- J. H. Ann, S. B. Damelin and P. Bigeleisen, Medical Image segmentation using modified Mumford segmentation methods, Ultrasound-Guided Regional Anesthesia and Pain Medicine, eds P. Bigeleisen, Chapter 40, Birkhauser.

- ▶ J. H. Ann, S. B. Damelin and P. Bigeleisen, Medical Image segmentation using modified Mumford segmentation methods, Ultrasound-Guided Regional Anesthesia and Pain Medicine, eds P. Bigeleisen, Chapter 40, Birkhauser.
- ▶ Kerry Cawse, Steven B. Damelin, Amandine Robin, Michael Sears, A parameter free approach for determining the intrinsic dimension of a hyperspectral image using Random Matrix Theory, IEEE Transaction on Image Processing, 22(4), 1301-1310,

- ▶ J. H. Ann, S. B. Damelin and P. Bigeleisen, Medical Image segmentation using modified Mumford segmentation methods, Ultrasound-Guided Regional Anesthesia and Pain Medicine, eds P. Bigeleisen, Chapter 40, Birkhauser.
- ▶ Kerry Cawse, Steven B. Damelin, Amandine Robin, Michael Sears, A parameter free approach for determining the intrinsic dimension of a hyperspectral image using Random Matrix Theory, IEEE Transaction on Image Processing, 22(4), 1301-1310,
- ▶ S. B. Damelin, On Bounds for Diffusion, Discrepancy and Fill Distance Metrics, Springer Lecture Notes in Computational Science and Engineering, Vol. 58, pp 32-42.

Brachial Plexus



Point energies

- Let us consider the problem of regularity for arrangements of $N \geq 2$ points on a class of d -dimensional compact sets A embedded in $\mathbb{R}^{d'}$ (ie sphere S^d , ball and torus). We assume that these $N \geq 2$ -arrangements interact through the Riesz kernel:

$$K_s(x, y) = \begin{cases} |x - y|^{-s}, & 0 < s < d, \quad x, y \in \mathbb{R}^{d'}, \\ -\log |x - y|, & s = 0, \quad x, y \in \mathbb{R}^{d'} \\ (c - |x - y|^{-s}), & -1 \leq s < 0, \quad x, y \in \mathbb{R}^{d'} \end{cases}$$

Point Energies-2

- Given a compact set $A \subset \mathbb{R}^{d'}$ and a collection $\omega_N = \{x_1, \dots, x_N\}$ of $N \geq 2$ distinct points on A , the discrete Riesz s -energy associated with ω_N is given by

$$E_s(A, \omega_N) := \sum_{1 \leq i < j \leq N} |x_i - x_j|^{-s}.$$

Let $\omega_s^*(A, N) := \{x_1^*, \dots, x_N^*\} \subset A$ be a configuration for which $E_s(A, \omega_N)$ attains its minimal value, that is,

$$\mathcal{E}_s(A, N) := \min_{\omega_N \subset A} E_s(A, \omega_N) = E_s(A, \omega_s^*(A, N)).$$

- In accordance with convention, we shall call such minimal configurations *s-extremal configurations*. It is well-known that, in general, *s*-extremal configurations are not always unique. For example, in the case of the unit sphere S^d , they are invariant under rotations.

$[-1, 1]$ [L,MF-M-R-S]

- The interval $[-1, 1]$, $\text{meas}([-1, 1]) = 1$:

$[-1, 1]$ [L,MF-M-R-S]

- ▶ The interval $[-1, 1]$, $\text{meas}([-1, 1]) = 1$:
- ▶ In the limiting cases, i.e., $s = 0$ (logarithmic interactions) and $s = \infty$ (best-packing problem), the s -extremal configurations are Fekete points and equally spaced points, respectively.

$[-1, 1]$ [L,MF-M-R-S]

- ▶ The interval $[-1, 1]$, $\text{meas}([-1, 1]) = 1$:
- ▶ In the limiting cases, i.e., $s = 0$ (logarithmic interactions) and $s = \infty$ (best-packing problem), the s -extremal configurations are Fekete points and equally spaced points, respectively.
- ▶ Fekete points are distributed on $[-1, 1]$ according to the arcsine measure, which has the density $\mu'_0(x) := (1/\pi)(1 - x^2)^{-1/2}$.

$[-1, 1]$ [L,MF-M-R-S]

- ▶ The interval $[-1, 1]$, $\text{meas}([-1, 1]) = 1$:
- ▶ In the limiting cases, i.e., $s = 0$ (logarithmic interactions) and $s = \infty$ (best-packing problem), the s -extremal configurations are Fekete points and equally spaced points, respectively.
- ▶ Fekete points are distributed on $[-1, 1]$ according to the arcsine measure, which has the density $\mu'_0(x) := (1/\pi)(1 - x^2)^{-1/2}$.
- ▶ Equally spaced points, $-1 + 2(k - 1)/(N - 1)$, $k = 1, \dots, N$, have the arclength distribution, as $N \rightarrow \infty$.

Critical transition: Movement of Mass

- $s = 1$ is the critical value in the sense that s -extremal configurations are distributed on $[-1, 1]$ differently for $s < 1$ and $s \geq 1$.

Critical transition: Movement of Mass

- ▶ $s = 1$ is the critical value in the sense that s -extremal configurations are distributed on $[-1, 1]$ differently for $s < 1$ and $s \geq 1$.
- ▶ For $s < 1$, the limiting distribution of s -extremal configurations has an arcsine-type density

$$\mu'_s(x) := \frac{\Gamma(1 + s/2)}{\sqrt{\pi} \Gamma((1 + s)/2)} (1 - x^2)^{(s-1)/2}.$$

Critical transition: Movement of Mass

- ▶ $s = 1$ is the critical value in the sense that s -extremal configurations are distributed on $[-1, 1]$ differently for $s < 1$ and $s \geq 1$.
- ▶ For $s < 1$, the limiting distribution of s -extremal configurations has an arcsine-type density

$$\mu'_s(x) := \frac{\Gamma(1 + s/2)}{\sqrt{\pi} \Gamma((1 + s)/2)} (1 - x^2)^{(s-1)/2}.$$

- ▶ For $s \geq 1$, the limiting distribution is the arclength distribution.

- This dependence of the distribution of s -extremal configurations over $[-1, 1]$ and the asymptotics for minimal discrete s -energy on s can be easily explained from potential theory point of view.

- This dependence of the distribution of s -extremal configurations over $[-1, 1]$ and the asymptotics for minimal discrete s -energy on s can be easily explained from potential theory point of view.
- For a probability Borel measure ν on $[-1, 1]$, its s -energy integral is defined to be

$$I_s([-1, 1], \nu) := \iint_{[-1, 1]^2} |x - y|^{-s} d\nu(x) d\nu(y)$$

(which can be finite or infinite).

- Let, for a set of points $\omega_N = \{x_1, \dots, x_N\}$ on $[-1, 1]$,

$$\nu^{\omega_N} := \frac{1}{N} \sum_{i=1}^N \delta_{x_i}$$

denote the normalized counting measure of ω_N (so that $\nu^{\omega_N}([-1, 1]) = 1$). Then the discrete Riesz s -energy, associated with ω_N can be written as

$$E_s([-1, 1], \omega_N) = (1/2)N^2 \iint_{x \neq y} |x - y|^{-s} d\nu^{\omega_N}(x) d\nu^{\omega_N}(y)$$

where the integral represents a discrete analog of the s -energy integral for the point-mass measure ν^{ω_N} .

- If $s < 1$, then the energy integral is minimized uniquely by an arcsine-type measure ν_s^* , whose density $\mu'_s(x)$ with respect to the Lebesgue measure.

- ▶ If $s < 1$, then the energy integral is minimized uniquely by an arcsine-type measure ν_s^* , whose density $\mu'_s(x)$ with respect to the Lebesgue measure.
- ▶ On the other hand, the normalized counting measure $\nu_{s,N}^*$ of an s -extreme configuration minimizes the discrete energy integral over all configurations ω_N on $[-1, 1]$.

- ▶ If $s < 1$, then the energy integral is minimized uniquely by an arcsine-type measure ν_s^* , whose density $\mu'_s(x)$ with respect to the Lebesgue measure.
- ▶ On the other hand, the normalized counting measure $\nu_{s,N}^*$ of an s -extreme configuration minimizes the discrete energy integral over all configurations ω_N on $[-1, 1]$.
- ▶ Thus one can reasonably expect that, for N large Thus one can reasonably expect that, for N large, $\nu_{s,N}^*$ is “close” to ν_s^* .

- ▶ If $s \geq 1$, then the energy integral diverges for every measure ν .
- ▶ Of course, depending on s , “far” neighbors still incorporate some energy in $\mathcal{E}_s([-1, 1], N)$, but the closest neighbors are dominating. So, s -extremal points distribute themselves over $[-1, 1]$ in an equally spaced manner.

- ▶ If $s \geq 1$, then the energy integral diverges for every measure ν .
- ▶ Concerning the distribution of s -extremal points over $[-1, 1]$, the interactions are now strong enough to force them to stay away from each other as far as possible.
- ▶ Of course, depending on s , “far” neighbors still incorporate some energy in $\mathcal{E}_s([-1, 1], N)$, but the closest neighbors are dominating. So, s -extremal points distribute themselves over $[-1, 1]$ in an equally spaced manner.

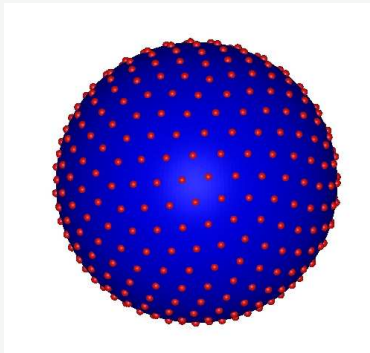
The Sphere [HS]

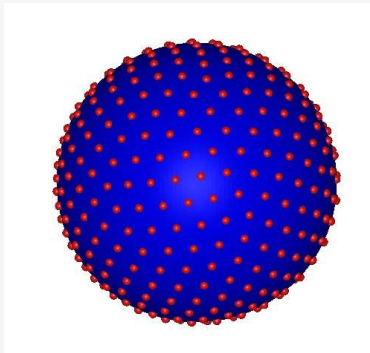
- The unit sphere S^d , $d_H(S^d) = d$: Here we again see three distinct cases: $s < d$, $s = d$, and $s > d$. Although it turns out that, for any s , the limiting distribution of s -extremal configurations is given by the normalized area measure on S^d .

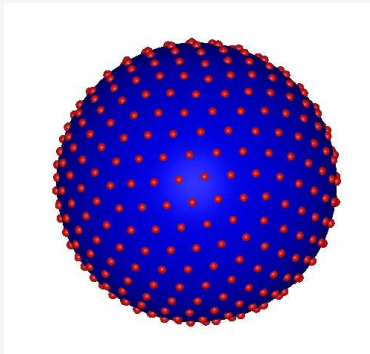
- Consider the sphere S^2 embedded in \mathbb{R}^3 . The minimum Riesz s -energy points presented are close to global minimum. In the table below, ρ denotes fill distance(mesh norm); 2δ denotes separation angle which is twice the separation (packing) radius and a denotes mesh ratio which is ρ/δ . Plots 1-4 illustrate $s = 1, 2, 3, 4$ extremal configurations for 400 points respectively. Because area measure is equilibrium measure in all cases due to symmetry, the points are similar for all values of s considered.

$s,$	$\rho,$	$2\delta,$	a
1,	0.113607,	0.175721,	1.2930
2,	0.127095,	0.173361,	1.4662
3,	0.128631,	0.173474,	1.4830
4,	0.134631,	0.172859,	1.5577

Some common themes
Interacting particle-Energy-Discrepancy
The Quantum Problem
Conjecture on the Covering Radius







Energy Equilibrium and Discrepancy Equivalence

We recall that the *energy* of a charge distribution $\mu \in \mathcal{M}(X)$ is

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

and the *energy* of the charge distribution μ in the field $f_{K,\mu}(x) = \int_X K(x, y) d\mu_f(y)$ induced by the charge distribution μ_f is

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

Here we see that $E_K(\mu, \mu_f)$ defines an inner product on the space of signed measures (charge distributions) for which the energy is finite.

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

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

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

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

Note that a crucial feature of the function space $H(K)$ is that it depends directly on the kernel K . More precisely, we have:

- S. B. Damelin, F. Hickernell, D. Ragozin and X. Zeng, On energy, discrepancy and G invariant measures on measurable subsets of Euclidean space, Journal of Fourier Analysis and its Applications (2010) (16), pp 813-839.

- S. B. Damelin, F. Hickernell, D. Ragozin and X. Zeng, On energy, discrepancy and G invariant measures on measurable subsets of Euclidean space, Journal of Fourier Analysis and its Applications (2010) (16), pp 813-839.
- S. B. Damelin, J. Levesley, D. L. Ragozin and X. Sun, Energies, Group Invariant Kernels and Numerical Integration on Compact Manifolds, Journal of Complexity, 25(2009), pp 152-162.

Let K be a conditionally positive definite energy kernel. Then

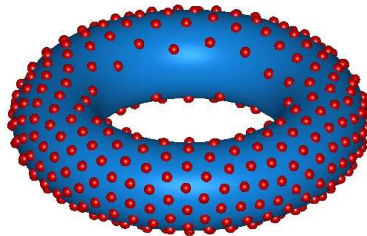
$$E_K(\mu) = \int_{X^2} K(x, y) d\mu(x) d\mu(y) \geq \frac{[Q(\mu)]^2}{C_K(X)}, \mu \in \mathcal{M}(X)$$

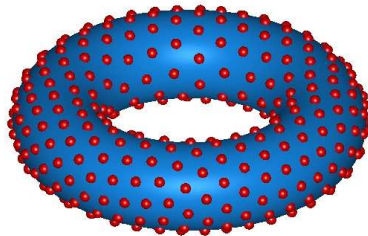
for the capacity constant $C_K(X)$ depending only on X and K with equality holding for any *equilibrium charge distribution* $\mu_{e,K}$, defined as one that induces a constant field,

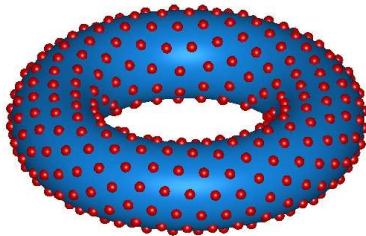
$$\phi_{K, \mu_{e,K}}(x) = \int_X K(x, y) d\mu_{e,K}(y) = \frac{Q(\mu_{e,K})}{C_K(X)} \quad \forall x \in X.$$

Torus

- Consider a torus embedded in \mathbb{R}^3 with inner radius 1 and outer radius 3. In this case, no longer have symmetry and so the three cases presented below for the minimum Riesz s -energy points $s = 1, 2, 3$ are not similar. Again we have 400 points.







- S. B. Damelin, 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)(2006), pp 845-863.

- S. B. Damelin, 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)(2006), pp 845-863.
- We define the *point energies* associated with $\omega_s^*(A, N)$ by

$$\mathcal{E}_{j,s}(A, N) := \sum_{\substack{i=1 \\ i \neq j}}^N |x_j^* - x_i^*|^{-s}, \quad j = 1, \dots, N.$$

Let $A \in \mathcal{A}^d$ and $s > d$. Then, for all $1 \leq j \leq N$,

$$\mathcal{E}_{j,s}(A, N) \leq CN^{s/d}.$$

Regularity: Separation $s > d$

- For $j = 1, \dots, N$ and a set $\omega_N = \{x_1, \dots, x_N\}$ of distinct points on $A \in \mathcal{A}^d$, we let

$$\delta_j(\omega_N) := \min_{i \neq j} \{|x_i - x_j|\}$$

and define

$$\delta(\omega_N) := \min_{1 \leq j \leq N} \delta_j(\omega_N).$$

The quantity $\delta(\omega_N)$ is called the *separation or packing radius* and gives the minimal distance between points in ω_N .

fill-distance and covering radius

We also define the *fill distance*(*mesh norm*) $\rho(A, \omega_N)$ of ω_N by

$$\rho(A, \omega_N) := \max_{y \in A} \min_{x \in \omega_N} |y - x|.$$

Geometrically, $\rho(A, \omega_N)$ means the maximal radius of a cap on A , which does not contain points from ω_N .

These two quantities, $\delta(\omega_N)$ and $\rho(A, \omega_N)$, give a good enough description of the distribution of ω_N over the set A . It is worth mentioning that, even for a sequence $\{\omega_N\}$ of *asymptotically* s -extremal configurations, i.e., configurations satisfying

$$\lim_{N \rightarrow \infty} \frac{E_s(A, \omega_N)}{\mathcal{E}_s(A, N)} = 1,$$

one can get only trivial estimates for the separation radius. Namely,

$$\delta(\omega_N) \geq cN^{-(1/d+1/s)}, \quad s > d.$$

However, for s -extremal configurations on A much better (best possible) estimate for the separation radius holds.

- S. B. Damelin, 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)(2006), pp 845-863.

- S. B. Damelin, 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)(2006), pp 845-863.
- For A , $s > d$, and any s -extremal configuration $\omega_s^*(A, N)$ on A ,

$$\delta_s^*(A, N) := \delta(\omega_s^*(A, N)) \geq cN^{-1/d}.$$

Regularity: Separation, $s < d - 1$

- S. B. Damelin, 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)(2006), pp 845-863.

Regularity: Separation, $s < d - 1$

- ▶ S. B. Damelin, 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)(2006), pp 845-863.
- ▶ Separation results for $s < d$ are far more difficult to find in the literature for the sets A . A reason for such a lack of results for weak interactions ($s < d$) is that this case require more delicate considerations based on the minimizing property of $\omega_s^*(A, N)$ while strong interactions ($s > d$) prevent points to be very close to each other without affecting the total energy.

A separation estimate in the case $s < d - 1$ for the unit sphere S^d .

- S. B. Damelin, 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)(2006), pp 845-863.

A separation estimate in the case $s < d - 1$ for the unit sphere S^d .

- S. B. Damelin, 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)(2006), pp 845-863.
- For $d \geq 2$ and $s < d - 1$,

$$\delta_s^*(S^d, N) \geq cN^{-1/(s+1)}.$$

For any $0 < s < d - 1$,

$$\lim_{N \rightarrow \infty} \frac{\max_{1 \leq j \leq N} \mathcal{E}_{j,s}(S^d, N)}{\min_{1 \leq j \leq N} \mathcal{E}_{j,s}(S^d, N)} = 1.$$

Numerical computations for a sphere and a torus suggest that, for any $s > 0$, the point energies are nearly equal for almost all points (which are of so called “hexagonal” type). However, some points (“pentagonal”) have elevated energies and some (“heptagonal”) have low energies. The transition from points that are “hexagonal” to those that are “pentagonal” and “heptagonal” induces **dislocation** (scar) defects, which are conjectured to vanish for N large enough. Thus, the corollary confirms this conjecture for $0 < s < d - 1$.

The estimate above can be improved for $d \geq 3$
and $s \leq d - 2$.

- S. B. Damelin, 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)(2006), pp 845-863.

The estimate above can be improved for $d \geq 3$
and $s \leq d - 2$.

- S. B. Damelin, 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)(2006), pp 845-863.
- Let $d \geq 3$ and $s \leq d - 2$. Then

$$\delta_s^*(S^d, N) \geq cN^{-1/(s+2)}.$$

Finite Field Algorithm

- ▶ S. B. Damelin, G. Mullen and G. Michalski, The cardinality of sets of k independent vectors over finite fields, *Monatsh.Math*, 150(2008), pp 289-295.
- ▶ S.B. Damelin, G. Mullen, G. Michalski and D. Stone, On the number of linearly independent binary vectors of fixed length with applications to the existence of completely orthogonal structures, *Monatsh Math*, (1)(2003), pp 1-12.
- ▶ B.Bajnok, S.B. Damelin, J. Li and G. Mullen, A constructive method of scattering points on d dimensional spheres using finite fields, *Computing (Springer)*, 68 (2002), pp 97-109, arxiv:1512.02984.

For an odd prime p , let F_p denote the finite field of integers modulo p . Consider the quadratic form given above over F_p . Let $N = N(d, p)$ denote the number of solutions of this form.

Step 1 ,We have:

$$N(d, p) = \begin{cases} p^d - p^{(d-1)/2} \eta((-1)^{(d+1)/2}) & \text{if } d \text{ is odd} \\ p^d + p^{d/2} \eta((-1)^{d/2}) & \text{if } d \text{ is even} \end{cases}$$

Here η is the quadratic character defined on F_p by $\eta(0) = 0$, $\eta(a) = 1$ if a is a square in F_p , and $\eta(a) = -1$ if a is a non-square in F_p .

Step 2

We now scale and centre around the origin.

Given a solution vector

$$X = (x_1, \dots, x_{d+1}), x_i \in F_p, 1 \leq i \leq d+1,$$

we may assume without loss of generality that the points x_i are scaled so that they are centered around the origin and are contained in the set

$$\{-(p-1)/2, \dots, (p-1)/2\}.$$

More precisely, if $x_i \in X$, define

$$w_i = \begin{cases} x_i & \text{if } x_i \in \{0, \dots, (p-1)/2\} \\ x_i - p & \text{if } x_i \in \{(p+1)/2, \dots, p-1\}. \end{cases}$$

Then $w_i \in \{-(p-1)/2, \dots, (p-1)/2\}$ and the scaled vector

$$W = (w_1, \dots, w_{d+1}), \quad 1 \leq i \leq d+1$$

solves the above if and only if X solves the above.

Step 3 Denoting by $\|\cdot\|$ the usual Euclidean metric, we multiply each solution vector W by $\frac{1}{\|W\|}$. Clearly each of these normalized points is now on the surface of the unit sphere S^d .

Use of the finite field F_p for larger primes p provides a method to increase the number N of points that are placed on the surface of S^d for any fixed $d \geq 1$. For increasing values of p , we obtain an increasing number $N = O(p^d)$ of points scattered on the surface of the unit sphere S^d ; in particular, as $p \rightarrow \infty$ through all odd primes, it is clear that $N \rightarrow \infty$.

For each prime p and integer $d \geq 1$, we will henceforth denote the set of points arising from our finite field construction by $X = X(d, p)$.

Examples

Let us now describe the point set X produced by the finite field construction and provide some explicit examples for small values of p and q . In each case, we may start with a well chosen set $V = V(d, p)$ of vectors.

In order to construct the full set of points $X(d, p)$, we need to consider all points obtained from V by taking ± 1 times the entry in each coordinate, and by permuting the coordinates of each vector, in all possible ways. For small values of d and p , this construction is summarized in the following table.

d	p	$N(d, p)$	$V(d, p)$
1	3	4	$\{(1, 0)\}$
1	5	4	$\{(1, 0)\}$
1	7	8	$\{(1, 0), \frac{1}{\sqrt{2}}(1, 1)\}$
2	3	6	$\{(1, 0, 0)\}$
2	5	30	$\{(1, 0, 0), \frac{1}{\sqrt{2}}(2, 1, 1)\}$
2	7	42	$\{(1, 0, 0), \frac{1}{\sqrt{2}}(1, 1, 0), \frac{1}{\sqrt{22}}(3, 3, 2)\}$

Observe that for $p = 3, 5, 7$ and $d = 1$, our construction gives the optimal solution, namely the vertices of the regular N -gon. This, however, is not the case for $p > 7$.

Spherical t -designs

Definition A finite set X of points on the d -sphere S^d is a *spherical t -design* or a *spherical design of strength t* , if for every polynomial f of total degree t or less, the average value of f over the whole sphere is equal to the arithmetic average of its values on X . If this only holds for homogeneous polynomials of degree t , then X is called a *spherical design of index t* .

For every odd positive integer k , odd prime p , and dimension $d \geq 1$, $X(d, p)$ is a spherical design of index k . Furthermore, $X(d, p)$ is a spherical 3-design.

Extension to finite fields of odd prime powers

Solve the same quadratic form over a general finite field F_q , where $q = p^e$ is an odd prime power and in this way distribute points on S^d as well. One way to do this is as follows.

Assume that $q = p^e$, with $e \geq 1$. Then the field F_q is an e -dimensional vector space over the field F_p . Let $\alpha_1, \dots, \alpha_e$ be a basis of F_q over F_p . Thus if $\alpha \in F_q$, then α can be uniquely written as $\alpha = a_1\alpha_1 + \dots + a_e\alpha_e$, where each $a_i \in F_p$. Moreover, we may assume that each a_i satisfies $-(p-1)/2 \leq a_i \leq (p-1)/2$.

If (x_1, \dots, x_{d+1}) is a solution to the quadratic form (1.1) over F_q , then each x_i is of the form $x_i = \alpha \in F_q$. Corresponding to the finite field element $x_i = \alpha$, we may now naturally associate the integer $M_i = a_1 + a_2 p + \dots + a_e p^{e-1}$. It is an easy exercise to check that indeed $-(p^e - 1)/2 \leq M_i \leq (p^e - 1)/2$. We then map the vector $V = (M_1, \dots, M_{d+1})$ to the surface of the unit sphere S^d by normalizing the vector V . We note that when $e = 1$, this reduces to our original construction. In particular, for increasing values of e , we obtain an increasing number N_e of points scattered on the surface of the unit sphere S^d , so that as $e \rightarrow \infty$, it is clear that $N_e \rightarrow \infty$.

Who Wins: Covering radius

- ρ : Points X_N randomly and independently distribution by area measure on S^d : $E\rho(X_N)$ has limit $((\log N)/N)^{1/d}$.
- Not extremal on A needed:

$$\delta(\omega_N) \geq cN^{-(1/d+1/s)}, \quad s > d.$$

Who Wins: Covering radius

- ρ : Points X_N randomly and independently distribution by area measure on S^d : $E\rho(X_N)$ has limit $((\log N)/N)^{1/d}$.
- Not extremal on A needed:

$$\delta(\omega_N) \geq cN^{-(1/d+1/s)}, \quad s > d.$$

- Extremal on A : For A , $s > d$, and any s -extremal configuration $\omega_s^*(A, N)$ on A ,

$$\delta_s^*(A, N) := \delta(\omega_s^*(A, N)) \geq cN^{-1/d}.$$

Who Wins: Covering radius

- ρ : Points X_N randomly and independently distribution by area measure on S^d : $E\rho(X_N)$ has limit $((\log N)/N)^{1/d}$.
- Not extremal on A needed:

$$\delta(\omega_N) \geq cN^{-(1/d+1/s)}, \quad s > d.$$

- Extremal on A : For A , $s > d$, and any s -extremal configuration $\omega_s^*(A, N)$ on A ,

$$\delta_s^*(A, N) := \delta(\omega_s^*(A, N)) \geq cN^{-1/d}.$$

- For $d \geq 2$ and $s < d - 1$,

$$\delta_s^*(S^d, N) \geq cN^{-1/(s+1)}.$$

Who Wins: Covering radius

- ρ : Points X_N randomly and independently distribution by area measure on S^d : $E\rho(X_N)$ has limit $((\log N)/N)^{1/d}$.
- Not extremal on A needed:

$$\delta(\omega_N) \geq cN^{-(1/d+1/s)}, \quad s > d.$$

- Extremal on A : For A , $s > d$, and any s -extremal configuration $\omega_s^*(A, N)$ on A ,

$$\delta_s^*(A, N) := \delta(\omega_s^*(A, N)) \geq cN^{-1/d}.$$

- For $d \geq 2$ and $s < d - 1$,

$$\delta_s^*(S^d, N) \geq cN^{-1/(s+1)}.$$

Who Wins: Covering radius-2

- Integer lattices (as I will use in the Quantum Section):
Sarnak Conjecture $N^{-1/4}$.

Who Wins: Covering radius-2

- ▶ Integer lattices (as I will use in the Quantum Section):
Sarnak Conjecture $N^{-1/4}$.
- ▶ FF Field and spherical design.

Who Wins: Covering radius-2

- ▶ Integer lattices (as I will use in the Quantum Section): Sarnak Conjecture $N^{-1/4}$.
- ▶ FF Field and spherical design.
- ▶ The finiteness and complexity.

Who Wins: Covering radius-2

- ▶ Integer lattices (as I will use in the Quantum Section): Sarnak Conjecture $N^{-1/4}$.
- ▶ FF Field and spherical design.
- ▶ The finiteness and complexity.
- ▶ I will assume for my integer lattices a lower bound of $(\log N)^b N^{-1/4}$ any b which will suffice for my approximation.

Who Wins: Covering radius-2

- ▶ Integer lattices (as I will use in the Quantum Section): Sarnak Conjecture $N^{-1/4}$.
- ▶ FF Field and spherical design.
- ▶ The finiteness and complexity.
- ▶ I will assume for my integer lattices a lower bound of $(\log N)^b N^{-1/4}$ any b which will suffice for my approximation.
- ▶ Hyperuniform points, Salvatore was talking about and Peter Grabner?

Invariant kernels on compact, reflexive homogenous spaces

Let $X \subset \mathbb{R}^{d+k}$ be a $d \geq 1, k \geq 0$ dimensional embedded reflexive, compact homogeneous C^∞ manifold; i.e. there is a compact group G of isometries of \mathbb{R}^{d+k} such that for some $\eta \in X$ (often referred to as the pole) $X = \{g\eta : g \in G\}$. The *reflexive* condition means that for each pair $x, y \in X$ there is a $g \in G$ with $gx = y$ and $gy = x$. A natural example to keep in mind is S^d , the d dimensional sphere realized as a subset of \mathbb{R}^{d+1} which is the orbit of any unit vector under the action of $SO(d+1)$, the group of $d+1$ dimensional orthogonal matrices of determinant 1.

A kernel $K : X \times X \rightarrow \mathbb{R}$ is termed *zonal* (or G -invariant) if $K(x, y) = K(gx, gy)$ for all $g \in G$ and $x, y \in X$. Since the maps in G are isometries of Euclidean space, they preserve both Euclidean distance and the (arc-length) metric $d(\cdot, \cdot)$ induced on the components of X by the Euclidean metric. Thus the distance kernel $d(x, y)$ on S^d is zonal. The manifold X carries a normalized *surface* (G -invariant) measure which we call μ_e .

In what follows, we will assume henceforth that a zonal kernel is continuous off the diagonal, lower semi-continuous everywhere and satisfies the following two conditions below:



$$\int_X K(x, y) d\mu(y)$$

exists for every $x \in X$.

- For each non-trivial continuous function ϕ on X , we have

$$\int_X \int_X K(x, y) \phi(x) \phi(y) d\mu(x) d\mu(y) > 0.$$

The archetype for such kernels is the weighted *Riesz kernel*

$$\kappa(x, y) = w(x, y) \|x - y\|^{-s}, \quad s > 0, \quad x, y \in X.$$

where $w : X \times X \rightarrow (0, \infty)$ is G invariant, positive definite, continuous off the diagonal and lower semi continuous everywhere.

Such kernels (in the case $w \equiv 1$) 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 modelling where points are do not have a uniform density. Note that when $s > -d$, K is absolutely integrable on X .

The error in integration is defined by

$$R(f, Z) := \int_X f(y) d\mu(y) - \frac{1}{N} \sum_{z \in Z} f(z).$$

For example, our result below applies to the space Pi_n of polynomials of degree at most $n \geq 1$ on X . Here, the space Π_n on X is realized as the space

$$\left\{ p \in C(X) : p = p_n|_X, \text{ for some polynomial, } p_n \in C(\mathbb{R}^{d+k}) \right\}$$

of degree at most n .

Harmonic analysis on X , in our case, requires the construction of harmonic polynomials on X . In this regard, if Π_j is the space of all polynomials of total degree $j \geq 1$ in the ambient space \mathbb{R}^{d+k} then $P_j := \Pi_j|_X$ is the space of degree j polynomials on X . We can also construct sets of *harmonic polynomials* $H_j := P_j \cap P_{j-1}^\perp$, where the orthogonality is with respect to the inner product on X .

Harmonic as in Laplace annihilation in the container space.

Remarkably in the case of a sphere—sum= of 2 spaces:
Harmonic, homogenous of degree n and harmonic
homogenous of degree $n - 1$ which gives the whole
polynomial space.

Let us consider integration over a finite interval $[a, b]$, $a < b$. In this case, as is well known, the nodes of the celebrated Gaussian quadrature formula can 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 $[a, b]$. In other words, the nodes are the zeros of the unique solution of an extremal problem. In the work of Damelin and Grabner, this idea was extended to the sphere whereby the authors related numerical integration via an extremal problem using Riesz energy and a class of G invariant kernels defined on the sphere.

Theorem Let K be admissible on X and $Z \subset X$ be a point subset of cardinality $N \geq 1$. If $q \in \Pi_n$ is a polynomial of degree at most $n \geq 1$ on X then,

$$|R(q, Z, \mu)| \leq \max_{j \leq n, l \leq \nu_n} \frac{1}{a_{j,l}} (K)^{1/2} \|q\|_2 (E_K(Z) - a_{0,0}(K))^{1/2}.$$

- **Theorem** $|R(f, Z, \mu)| = O\left(\frac{1}{\sqrt{\log N}}\right)$., far away interactions. O depends on smooth $f : S^d \rightarrow \mathbb{R}$. Not sharp for sure. Ergodic actions I think will reduce it.
- S.B. Damelin and P. Grabner, Energy functionals, Numerical integration and Asymptotic equidistribution on the sphere, Journal of Complexity, 19(2003), pp 231-246. (Postscript) Corrigendum, Journal of Complexity, (20)(2004), pp 883-884.

In numerical integration or cubature we approximate the integral of a function $f \in H(K)$,

$$\begin{aligned} I(f; \tilde{\mu}) &= \int_X f(x) d\tilde{\mu}(x) = E_K(\tilde{\mu}, \mu_f) = \\ &= \langle \tilde{\mu}, \mu_f \rangle_M = \langle \phi \tilde{\mu}, f \rangle_{H(K)} \end{aligned}$$

by the cubature rule

$$\begin{aligned} I(f; \hat{\mu}) &= \int_X f(x) d\hat{\mu}(x) = \sum_{i=1}^n c_i f(x_i) = E_K(\hat{\mu}, \mu_f) = \\ &= \langle \hat{\mu}, \mu_f \rangle_M = \langle \phi \hat{\mu}, f \rangle_{H(K)} \end{aligned}$$

where $\hat{\mu}$ is the charge distribution (signed measure) with support N points, x_1, \dots, x_N and charge c_i at each point f_i .

Moreover,

$$\phi_{\tilde{\mu}}(x) = \int_X K(x, y) d\tilde{\mu}(y)$$

is the representer of the integration functional, and

$$\phi_{\hat{\mu}}(X) = \int_X K(x, y) d\hat{\mu}(y) = \sum_{i=1}^N c_i K(x, x_i)$$

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

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

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, μ_f is parallel to $\tilde{\mu} - \hat{\mu}$. Thus, we have:

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

The quantity $D_K(\tilde{\mu}, \hat{\mu})$, defined by above 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 provided the right hand side is well defined.

For a fixed choice of points $Y = \{x_1, \dots, x_N\}$, 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 Y , i.e.,

$$\phi_{\hat{\mu}}(x_i) = \phi_{\tilde{\mu}}(x_i), \quad i = 1, \dots, N.$$

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 Y consisting of n points that has *maximum energy* under the given constraint.

It is now possible to define a distance on X by way of:

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

The Quantum Problem

A key problem in Quantum Computation is how to approximate quantum gates. The 2 main notions in classical computing are then translated into Quantum equivalents.

The Quantum Problem

A key problem in Quantum Computation is how to approximate quantum gates. The 2 main notions in classical computing are then translated into Quantum equivalents.

- A 1-qubit, is a quantum bit of information represented by $\alpha|0\rangle + \beta|1\rangle$ where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. n -qubits are defined as n tensor products of some 1-qubits.

The Quantum Problem

A key problem in Quantum Computation is how to approximate quantum gates. The 2 main notions in classical computing are then translated into Quantum equivalents.

- ▶ A 1-qubit, is a quantum bit of information represented by $\alpha|0\rangle + \beta|1\rangle$ where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. n -qubits are defined as n tensor products of some 1-qubits.
- ▶ A n -bit quantum gate is viewed as a linear function on n qubits, which is represented by an element in $U(2^n, \mathbb{C})$. A 1-qubit gate is taken to be an element of $SU(2, \mathbb{C})$ to preserve the norm.

The Quantum Problem

The Quantum Problem is how to find small sets of quantum gates which generate $SU(2)$. However, since $SU(2)$ is not finitely generated, no finite set will completely generate $SU(2)$. We consider a well defined notion of approximation. Good coverings S^3 . Henceforth, G refers to either the group $SU(2)$ or $PSU(2)$.

The Quantum Problem

The Quantum Problem is how to find small sets of quantum gates which generate $SU(2)$. However, since $SU(2)$ is not finitely generated, no finite set will completely generate $SU(2)$. We consider a well defined notion of approximation. Good coverings S^3 . Henceforth, G refers to either the group $SU(2)$ or $PSU(2)$.

- $d_G(M, N) = \sqrt{1 - \frac{|Tr(M^\dagger N)|}{2}}$, which is invariant under the group action. \dagger is complex conjugation.

The Quantum Problem

The Quantum Problem is how to find small sets of quantum gates which generate $SU(2)$. However, since $SU(2)$ is not finitely generated, no finite set will completely generate $SU(2)$. We consider a well defined notion of approximation. Good coverings S^3 . Henceforth, G refers to either the group $SU(2)$ or $PSU(2)$.

- ▶ $d_G(M, N) = \sqrt{1 - \frac{|Tr(M^\dagger N)|}{2}}$, which is invariant under the group action. \dagger is complex conjugation.
- ▶ Denote the Haar measure on G by μ . Then $\mu(B_G(M, \varepsilon)) = \mu(B_G(I, \varepsilon))$, $M \in G$, $\varepsilon > 0$.

The Quantum Problem

The Quantum Problem is how to find small sets of quantum gates which generate $SU(2)$. However, since $SU(2)$ is not finitely generated, no finite set will completely generate $SU(2)$. We consider a well defined notion of approximation. Good coverings S^3 . Henceforth, G refers to either the group $SU(2)$ or $PSU(2)$.

- ▶ $d_G(M, N) = \sqrt{1 - \frac{|\text{Tr}(M^\dagger N)|}{2}}$, which is invariant under the group action. \dagger is complex conjugation.
- ▶ Denote the Haar measure on G by μ . Then $\mu(B_G(M, \varepsilon)) = \mu(B_G(I, \varepsilon))$, $M \in G$, $\varepsilon > 0$.
- ▶ A **universal set** Γ is a subset of G that generates a dense subgroup with respect to d_G

Some Quick Picture

Since $SU(2)$ is diffeomorphic to S^3 , the Quantum problem can be envisioned as figuring out how to best cover S^3 with balls of radius ε . To help envision this, some basic instances were simulated in the 2-dimensional case.

Some Quick Picture

Since $SU(2)$ is diffeomorphic to S^3 , the Quantum problem can be envisioned as figuring out how to best cover S^3 with balls of radius ε . To help envision this, some basic instances were simulated in the 2-dimensional case.

- ▶ Left picture is of all products of length 5 over a universal set with positive coordinates.
- ▶ Right picture is of all products of length 7 over the same universal set with positive coordinates.

Some Quick Picture

Since $SU(2)$ is diffeomorphic to S^3 , the Quantum problem can be envisioned as figuring out how to best cover S^3 with balls of radius ε . To help envision this, some basic instances were simulated in the 2-dimensional case.

- ▶ Left picture is of all products of length 5 over a universal set with positive coordinates.
- ▶ Right picture is of all products of length 7 over the same universal set with positive coordinates.
- ▶ Two arcs of the same size are pictured, centered at generated points. Endpoints of the arcs are included for clarity, but are not generated points.

Some Quick Picture

Since $SU(2)$ is diffeomorphic to S^3 , the Quantum problem can be envisioned as figuring out how to best cover S^3 with balls of radius ε . To help envision this, some basic instances were simulated in the 2-dimensional case.

- ▶ Left picture is of all products of length 5 over a universal set with positive coordinates.
- ▶ Right picture is of all products of length 7 over the same universal set with positive coordinates.
- ▶ Two arcs of the same size are pictured, centered at generated points. Endpoints of the arcs are included for clarity, but are not generated points.
- ▶ Notice that the arcs can cover the circle in the right picture, but cannot cover the circle in the left picture.

Some Quick Pictures

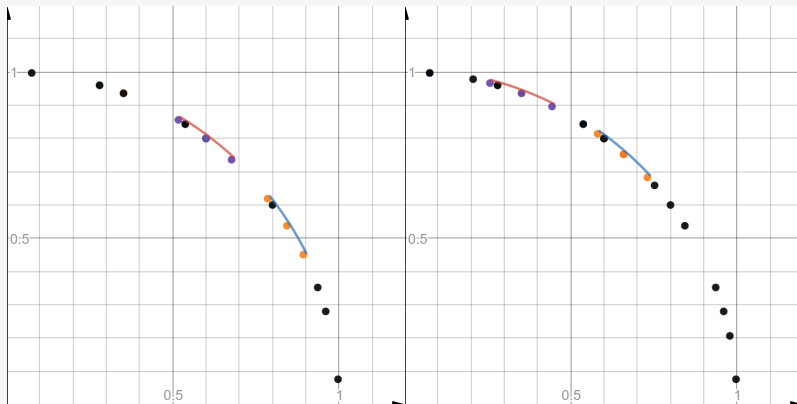


Figure: Rough covering the unit circle by arcs, the exact representation is in 4 dimensions

The Quantum Problem

Basis for the approximation theory.

Theorem (Solovay-Kitaev)

Let Γ be a finite universal set in $SU(2)$ and $\varepsilon > 0$. Then there exists $c > 0$ such that for any $X \in SU(2)$, there is a finite product S of gates in Γ of length $O(\log^c(\frac{1}{\varepsilon}))$ such that $d_G(S, X) < \varepsilon$.

This theorem suggests that instead of measuring how efficiently a universal set can approximate one gate, that it is practical to measure how efficient a universal set can approximate all of $SU(2)$.

The Covering Exponent

To expand on this notion on the efficiency of universal sets, we look to measure how well universal sets can approximate any quantum gate. Let Γ be a universal set in G equipped with a positive weight w .

The Covering Exponent

To expand on this notion on the efficiency of universal sets, we look to measure how well universal sets can approximate any quantum gate. Let Γ be a universal set in G equipped with a positive weight w .

- For any $\gamma \in \langle \Gamma \rangle$, define the **height** of γ

$$h(\gamma) = \min \left\{ \sum_i w(c_i) : c_i \in \Gamma, \prod_i c_i = \gamma \right\}$$

The Covering Exponent

To expand on this notion on the efficiency of universal sets, we look to measure how well universal sets can approximate any quantum gate. Let Γ be a universal set in G equipped with a positive weight w .

- For any $\gamma \in \langle \Gamma \rangle$, define the **height** of γ

$$h(\gamma) = \min \left\{ \sum_i w(c_i) : c_i \in \Gamma, \prod_i c_i = \gamma \right\}$$

- The height represents the cost of the gate γ , and very often the weight is chosen so that most of Γ has weight 1.

The Covering Exponent

The goal is to effectively measure how much it costs to approximate any gate within a tolerance of ε by a gate generated over Γ . For convenience, define for each $t > 0$, the sets:

The Covering Exponent

The goal is to effectively measure how much it costs to approximate any gate within a tolerance of ε by a gate generated over Γ . For convenience, define for each $t > 0$, the sets:

- ▶ $U_{\Gamma}(t) = \{\gamma \in \langle \Gamma \rangle : h(\gamma) = t\}$
- ▶ $V_{\Gamma}(t) = \{\gamma \in \langle \Gamma \rangle : h(\gamma) \leq t\}$

The Covering Exponent

The goal is to effectively measure how much it costs to approximate any gate within a tolerance of ε by a gate generated over Γ . For convenience, define for each $t > 0$, the sets:

- ▶ $U_\Gamma(t) = \{\gamma \in \langle \Gamma \rangle : h(\gamma) = t\}$
- ▶ $V_\Gamma(t) = \{\gamma \in \langle \Gamma \rangle : h(\gamma) \leq t\}$
- ▶ Then the **Covering Height** for a tolerance $\varepsilon > 0$ is given by

$$t_\varepsilon = \min \left\{ t > 0 : G \subset \bigcup_{\gamma \in V_\Gamma(t)} B_G(\gamma, \varepsilon) \right\}$$

The Covering Exponent

The goal is to effectively measure how much it costs to approximate any gate within a tolerance of ε by a gate generated over Γ . For convenience, define for each $t > 0$, the sets:

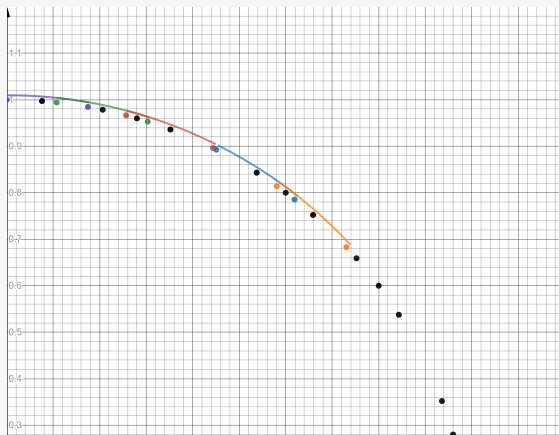
- ▶ $U_\Gamma(t) = \{\gamma \in \langle \Gamma \rangle : h(\gamma) = t\}$
- ▶ $V_\Gamma(t) = \{\gamma \in \langle \Gamma \rangle : h(\gamma) \leq t\}$
- ▶ Then the **Covering Height** for a tolerance $\varepsilon > 0$ is given by

$$t_\varepsilon = \min \left\{ t > 0 : G \subset \bigcup_{\gamma \in V_\Gamma(t)} B_G(\gamma, \varepsilon) \right\}$$

- ▶ t_ε represents the minimal height required to cover S^3 with balls of radius ε centered at points generated by Γ . However, t_ε is extraordinarily hard to compute.

The Covering Exponent

Here is an example of an almost covering with tolerance $\varepsilon = 0.1$. Then $t_{0.1} = 7 + 1 = 8$.



The Covering Exponent [Sar]

The Covering Exponent [Sar]

- Notice that $G \subset \bigcup_{\gamma \in V_\Gamma(t_\varepsilon)} B_G(\gamma, \varepsilon)$ for any ε , and thus

$$\mu(B_G(\varepsilon)) |V_\Gamma(t)| \geq 1$$

The Covering Exponent [Sar]

- Notice that $G \subset \bigcup_{\gamma \in V_\Gamma(t_\varepsilon)} B_G(\gamma, \varepsilon)$ for any ε , and thus

$$\mu(B_G(\varepsilon)) |V_\Gamma(t)| \geq 1$$

- Define the **Covering Exponent** of Γ over G with respect to w as

$$K(\Gamma) = \limsup_{\varepsilon \rightarrow 0} \frac{\log |V_\Gamma(t_\varepsilon)|}{\log(1/\mu(B_G(\varepsilon)))}$$

The Covering Exponent [Sar]

- Notice that $G \subset \bigcup_{\gamma \in V_\Gamma(t_\varepsilon)} B_G(\gamma, \varepsilon)$ for any ε , and thus

$$\mu(B_G(\varepsilon)) |V_\Gamma(t)| \geq 1$$

- Define the **Covering Exponent** of Γ over G with respect to w as

$$K(\Gamma) = \limsup_{\varepsilon \rightarrow 0} \frac{\log |V_\Gamma(t_\varepsilon)|}{\log(1/\mu(B_G(\varepsilon)))}$$

- The covering exponent compares how quickly the number of points Γ generates grows against how quickly the measure of the balls being used grows. Thus is a good way of seeing the efficiency of a universal set.

Constructing an Efficient Universal Set

For $S \subset \mathbb{R}$, define $H(S)$ as the set $\{a + bi + cj + dk : a, b, c, d \in S\}$. Furthermore, denote the unit sphere over the quaternions as $H_1(S)$.

Constructing an Efficient Universal Set

For $S \subset \mathbb{R}$, define $H(S)$ as the set $\{a + bi + cj + dk : a, b, c, d \in S\}$. Furthermore, denote the unit sphere over the quaternions as $H_1(S)$.

- Every Matrix $M \in SU(2)$ can be written as $\begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix}$ and thus can be related to the pair of complex numbers α, β

Constructing an Efficient Universal Set

For $S \subset \mathbb{R}$, define $H(S)$ as the set $\{a + bi + cj + dk : a, b, c, d \in S\}$. Furthermore, denote the unit sphere over the quaternions as $H_1(S)$.

- ▶ Every Matrix $M \in SU(2)$ can be written as $\begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix}$ and thus can be related to the pair of complex numbers α, β
- ▶ Note that $SU(2)$ is isomorphic to $H_1(\mathbb{R})$ via the map

$$\begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix} \mapsto \alpha + \beta j$$

Constructing an Efficient Universal Set

For $S \subset \mathbb{R}$, define $H(S)$ as the set $\{a + bi + cj + dk : a, b, c, d \in S\}$. Furthermore, denote the unit sphere over the quaternions as $H_1(S)$.

- ▶ Every Matrix $M \in SU(2)$ can be written as $\begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix}$ and thus can be related to the pair of complex numbers α, β
- ▶ Note that $SU(2)$ is isomorphic to $H_1(\mathbb{R})$ via the map

$$\begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix} \mapsto \alpha + \beta j$$

- ▶ Define $\Phi : U(2) \rightarrow H(\mathbb{R})$ so that $\Phi(M) = \alpha + \beta j$

Constructing an Efficient Universal Set

For $S \subset \mathbb{R}$, define $H(S)$ as the set $\{a + bi + cj + dk : a, b, c, d \in S\}$. Furthermore, denote the unit sphere over the quaternions as $H_1(S)$.

- ▶ Every Matrix $M \in SU(2)$ can be written as $\begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix}$ and thus can be related to the pair of complex numbers α, β
- ▶ Note that $SU(2)$ is isomorphic to $H_1(\mathbb{R})$ via the map

$$\begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix} \mapsto \alpha + \beta j$$

- ▶ Define $\Phi : U(2) \rightarrow H(\mathbb{R})$ so that $\Phi(M) = \alpha + \beta j$
- ▶ Conveniently, $\Phi(M^\dagger) = \overline{\Phi(M)}$. Thus, we have the identity $Tr(M^\dagger N) = \frac{2\langle \Phi(M), \Phi(N) \rangle}{|\Phi(M)| |\Phi(N)|}$

Constructing an Efficient Universal Set

The aim is to construct integer lattices in $H(\mathbb{Z})$, and use the identity from the last slide to get a good bound for the covering exponent. Let $q \equiv 1 \pmod{4}$ be prime.

Constructing an Efficient Universal Set

The aim is to construct integer lattices in $H(\mathbb{Z})$, and use the identity from the last slide to get a good bound for the covering exponent. Let $q \equiv 1 \pmod{4}$ be prime.

- Consider $L_k = \{a \in H(\mathbb{Z}) : |a| = q^k\}$. Then $L_k \subset \langle L_1 \rangle$ and thus define $L = \{a \in H(\mathbb{Z}) : |a| = q^k, k \in \mathbb{N}\}$.

Constructing an Efficient Universal Set

The aim is to construct integer lattices in $H(\mathbb{Z})$, and use the identity from the last slide to get a good bound for the covering exponent. Let $q \equiv 1 \pmod{4}$ be prime.

- ▶ Consider $L_k = \{a \in H(\mathbb{Z}) : |a| = q^k\}$. Then $L_k \subset \langle L_1 \rangle$ and thus define $L = \{a \in H(\mathbb{Z}) : |a| = q^k, k \in \mathbb{N}\}$.
- ▶ The problem can be simplified by considering representations of q^{2k} as a sum of four squares. Note that when $q \equiv 5$ that L_1 is generated by

$$A = \{1 + 2i, 1 + 2j, 1 + 2k, 1 - 2i, 1 - 2j, 1 - 2k, i, j, k\}$$

Constructing an Efficient Universal Set

The aim is to construct integer lattices in $H(\mathbb{Z})$, and use the identity from the last slide to get a good bound for the covering exponent. Let $q \equiv 1 \pmod{4}$ be prime.

- ▶ Consider $L_k = \{a \in H(\mathbb{Z}) : |a| = q^k\}$. Then $L_k \subset \langle L_1 \rangle$ and thus define $L = \{a \in H(\mathbb{Z}) : |a| = q^k, k \in \mathbb{N}\}$.
- ▶ The problem can be simplified by considering representations of q^{2k} as a sum of four squares. Note that when $q \equiv 5$ that L_1 is generated by

$$A = \{1 + 2i, 1 + 2j, 1 + 2k, 1 - 2i, 1 - 2j, 1 - 2k, i, j, k\}$$

- ▶ We will take $T = \frac{1}{\sqrt{5}} \Phi^{-1}(A)$ as our universal set.

Constructing an Efficient Universal Set

Consider the Pauli matrices denoted here as

$$X = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, Z = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Notice that for any of the matrices above say M , $iM \in SU(2)$ and has an inverse of $-iM$. Thus, we take $G = PSU(2)$. Let

$$s_x = \frac{1}{\sqrt{5}}(1 + 2iX), s_y = \frac{1}{\sqrt{5}}(1 + 2iY), s_z = \frac{1}{\sqrt{5}}(1 + 2iZ)$$
$$s_x^{-1} = \frac{1}{\sqrt{5}}(1 - 2iX), s_y^{-1} = \frac{1}{\sqrt{5}}(1 - 2iY), s_z^{-1} = \frac{1}{\sqrt{5}}(1 - 2iZ)$$

Define T as

$$T = \{X, Y, Z, s_x, s_y, s_z, s_x^{-1}, s_y^{-1}, s_z^{-1}\}$$

Constructing an Efficient Universal Set

As claimed,

$$\Phi(iX) = i, \Phi(iY) = j, \Phi(iZ) = k$$

$$\Phi(s_x) = \frac{1}{\sqrt{5}}(1 + 2i), \Phi(s_y) = \frac{1}{\sqrt{5}}(1 + 2j), \Phi(s_z) = \frac{1}{\sqrt{5}}(1 + 2k)$$

$$\Phi(s_x^{-1}) = \frac{1}{\sqrt{5}}(1 - 2i), \Phi(s_y^{-1}) = \frac{1}{\sqrt{5}}(1 - 2j), \Phi(s_z^{-1}) = \frac{1}{\sqrt{5}}(1 - 2k)$$

$$\text{Thus } T = \frac{1}{\sqrt{5}}\Phi^{-1}(A).$$

Constructing an Efficient Universal Set

The first obstacle in computing $K(\Gamma)$ comes from reliably expressing $|V_\Gamma(t_\varepsilon)|$. However, since $T = \frac{1}{\sqrt{5}}\Phi^{-1}(A)$ it is by construction that $U_\Gamma(t)$ is in bijection with L_t for all t .

Note that if $\gamma \in U_\Gamma(t)$, then $5^t\Phi(\gamma) \in L_t$ implies that $5^{t+1}\Phi(\gamma) \in L_{t+1}$. Thus, $V_\Gamma(t)$ is in bijection with L_t since for all $k \leq t$ each matrix in $U_\Gamma(k)$ has a representative in L_t . The number of ways to express an odd integer n as a sum of 4 squares is

$$r_4(n) = 8 \sum_{m|n} m$$

Thus,

$$|V_\Gamma(t)| = r_4(5^t) = 8 \sum_{k=0}^t 5^k = 2 \cdot 5^{t+1} - 2$$

Constructing an Efficient Universal Set

Then we have all the tools to bound $K(T)$.

Constructing an Efficient Universal Set

Then we have all the tools to bound $K(T)$.

$$\blacktriangleright |V_T(t_\varepsilon)| = 2 \cdot 5^{t_\varepsilon+1} - 2$$

Constructing an Efficient Universal Set

Then we have all the tools to bound $K(T)$.

- ▶ $|V_T(t_\varepsilon)| = 2 \cdot 5^{t_\varepsilon+1} - 2$
- ▶ As shown in [Sar], Hecke operators can be used to show that there is exists $c > 0$ so that $|V_T(t_\varepsilon)| \leq \frac{4\pi c t_\varepsilon}{\varepsilon^4}$.
Standard computations also give $\mu(B_G(\varepsilon)) \sim \varepsilon^2$.

Constructing an Efficient Universal Set

Then we have all the tools to bound $K(T)$.

- ▶ $|V_T(t_\varepsilon)| = 2 \cdot 5^{t_\varepsilon+1} - 2$
- ▶ As shown in [Sar], Hecke operators can be used to show that there exists $c > 0$ so that $|V_T(t_\varepsilon)| \leq \frac{4\pi c t_\varepsilon}{\varepsilon^4}$.
 Standard computations also give $\mu(B_G(\varepsilon)) \sim \varepsilon^2$.
- ▶ Recall: $K(T) = \limsup_{\varepsilon \rightarrow 0} \frac{\log(|V_T(t_\varepsilon)|)}{\log(\mu(1/\mu(B_G(\varepsilon))))}$
- ▶ It follows from the results of [DGLM] that $|V_T(t_\varepsilon)|$ is bounded by some function of ε which gives that $K(T) \leq 2$.

Constructing an Efficient Universal Set

Then we have all the tools to bound $K(T)$.

► $|V_T(t_\varepsilon)| = 2 \cdot 5^{t_\varepsilon+1} - 2$

► Recall: $K(T) = \limsup_{\varepsilon \rightarrow 0} \frac{\log(|V_T(t_\varepsilon)|)}{\log(\mu(1/\mu(B_G(\varepsilon))))}$

- Since the logarithms share the same base in the numerator and the denominator, the choice of q is irrelevant as long as $q \equiv 1 \pmod{4}$. Thus, T is the most efficient construction of this type.

Conjecture on the Covering Radius

Ultimately, we would like to compute the covering radius of G by $V_T(t_\varepsilon)$ with respect to d_G . Recall that G can be either $PSU(2)$ or $SU(2)$. Consider the following optimistic conjecture from [DGLM]:

Conjecture

There is $0 < \delta < 1$ so that for all $a \in H_1(\mathbb{R})$ and $\varepsilon > 0$ there is an $m \in \mathbb{N}$ and $b \in L_m$ with $\frac{\langle a, b \rangle}{|b|} > 1 - 5^{\frac{-2m}{2-\delta}}$

Note that as t_ε grows without bound as ε goes to zero, $1 - 5^{\frac{-2t_\varepsilon}{2-\delta}}$ approaches 1 monotonically from the left.

The Conjecture Visualized

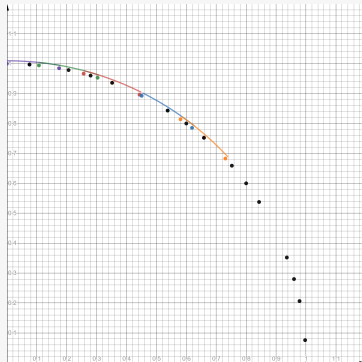


Figure: In terms of the earlier scenario, the conjecture gives that size of the arcs required to cover the circle is slightly larger than 0.1 and that any ε it does work for satisfies $\varepsilon \leq 5^{\frac{-14}{2-\delta}}$.

Conjecture on the Covering Radius

If the conjecture holds true, then for any $M \in SU(2)$ and $\varepsilon > 0$ there is an $N \in V_T(t_\varepsilon)$ so that

$$\text{Tr}(M^\dagger N) = 2\langle \Phi(M), \Phi(N) \rangle = 2 \left\langle \frac{\Phi(5^{t_\varepsilon} M)}{5^{t_\varepsilon}}, \Phi(N) \right\rangle > 2(1 - 5^{\frac{2t_\varepsilon}{2-\delta}})$$

which implies

$$d_G(M, N) = \sqrt{1 - \frac{|\text{Tr}(M^\dagger N)|}{2}} \leq 5^{\frac{t_\varepsilon}{2-\delta}}$$

and since this bound is constructed off t_ε , if multiple ε generate the same t_ε then $\varepsilon \leq 5^{\frac{t_\varepsilon}{2-\delta}}$. Thus, $\varepsilon \leq 5^{\frac{t_\varepsilon}{2-\delta}}$ will always be true.

Conjecture on the Covering Radius

Why does this matter?

- ▶ When the conjecture holds, it follows from [DGLM] that $\varepsilon \leq 5^{\frac{t\varepsilon}{2-\delta}}$ and in turn $K(T) \leq 2 - \delta$. Practically speaking, the conjecture allows the covering exponent to be estimated using simpler calculations over quaternions.

- The conjecture can be reworded in several ways. For example [DGLM]:

Conjecture

Let ρ be the euclidean distance on $\Phi(PSU(2))$. Then for all $a \in H_1(\mathbb{R})$, there is a $b \in L_{t_\varepsilon}$ such that

$$\frac{1}{b}\rho(a, b) \leq \varepsilon\sqrt{2}$$

- $(\log(N))^b N^{-1/4}$ any b works.

- ▶ S.B. Damelin and B. Mode, On a qualitative Solovay-Kitaev theorem for Quantum gates, arxiv: 1709.03007.
- ▶ S.B. Damelin, Q. Liang and B. Mode, Golden Gates and Discrepancy, arxiv: 1506.05785.

- Shortest paths through random points and K-Nearest and Random Graphs.

- ▶ Shortest paths through random points and K-Nearest and Random Graphs.
- ▶ Whitney extensions and interpolation in high dimensions.

- ▶ Shortest paths through random points and K-Nearest and Random Graphs.
- ▶ Whitney extensions and interpolation in high dimensions.
- ▶ Codes-designs.

- ▶ M. Hua, S. Damelin, J. Sun and M. Yu, The Truncated and Supplemented Pascal Matrix and Applications, Involve, Vol. 11, No. 2, 2018.
- ▶ S. B. Damelin, J. Sun, M. Yu and D. Kaiser, An Algebraic-Combinatorial proof for a Ring formulation of the MDS conjecture, arxiv: 1611.02354.
- ▶ S. B. Damelin, J. Sun and M. Yu, An Analytic and Probabilistic Approach to the Problem of Matroid Representibility, arxiv: 1506.06146.
- ▶ Sung Jin Hwang, Steven B. Damelin, Alfred O. Hero III, Shortest Path through Random Points, The Annals of Applied Probability, 2016, Vol. 26, No. 5, pp 2791-2823.
- ▶ S.B. Damelin and C. Fefferman, On the Whitney extension-interpolation-alignment problem for almost isometries with small distortion in \mathbb{R}^D , arxiv: 1411.2468.

Thank you very much again for this invitation.