The Journal of Fourier Analysis and Applications

On Determinants of Constrained Random Fourier Minors

Maurice J. Jansen and Kenneth W. Regan

ABSTRACT. *For parameters n, l and r, we consider the problem of maximizing the determinant of an $l \times l$ Vandermonde matrix V, with nodes selected from the set Ω_n of nth roots of unity, but avoiding a forbidden subset R of size r. An asymptotically tight lower bound is given for the expected value of $\ln |\det V|$ in case the nodes are selected uniformly at random from Ω_n / R. We apply our result to give a discrete uncertainty relation for so-called ϵ, l-index-limited vectors.*

1. Introduction

The minors of the $n \times n$ discrete Fourier matrix $DFT_n = (e^{2\pi ist/n})_{0 \leq s,t \leq n-1}$ have been studied more traditionally in the literature under *existential* and *universal* modes of quantification. To give an example of the latter, there is the classic result that for prime p, any minor of DFT_p is non-singular. The first proof of this fact is attributed to Chebotarëv, who proved it in 1926 (see [8], and also see [9] for an elementary proof). Recently added to this, Candès, Romberg and Tao [1] engaged in a study of *randomly* quantified Fourier minors. They prove that for any set of rows R of size $O(\frac{n}{\log n})$, if one selects a set S of columns by independently choosing for each column to be in S with some fixed probability τ, then with high probability for the minor M of DFT_n with rows R and columns S , denoted by $M = DFT^n_{R,S}$, the determinant $\det(MM^*)$ is "not too small."

We are interested in $\forall\exists$-quantified Fourier minors of the following kind. Suppose an adversary specified l rows R and r columns C of the DFT_n matrix. Does

Math Subject Classifications. 4202, 15A15, 15A52.
Keywords and Phrases. Discrete Fourier Transform, Random Vandermonde Matrix, Determinant, Uncertainty Principle.
Note. Part of this work by both authors was supported by NSF Grant CCR-9821040.

ISSN 1069-5869

there exist a minor of DFT_n, that has exactly as its rows the set R and has all its columns *disjoint* from C, with determinant of magnitude larger than some given value B? That is, for given parameters l, r, n, B, when does it hold that *for all* sets of l many required rows R and r many excluded columns C, *there exists* a set of l many columns D disjoint from C, such that the minor $M = DFT^n_{R,D}$ has $|\det(M)| > B$? Second, what can be said about $|\det(DFT^n_{R,D})|$ if D is selected uniformly at random among all such allowed sets ? Let us remark that the the requirement to avoid C seems to upset attempts to extend the main proof idea in [1], which relies on the *cancellation property* of the roots of unity. We are especially interested in the case where the set of rows R selected by the adversary is contiguous.

Toward these ends, we give some motivating open problems. For a set S of complex numbers, $|S| = l < \infty$, define its *chordal product* $\mathcal{CP}(S)$ by

$$\mathcal{CP}(S) = \prod_{p \neq q \in S} |p - q|^{1/2} = |\det(V_S)|,$$

where V_S is an $l \times l$ Vandermonde matrix whose second row comprises S. We consider S on the unit circle, and further restrict S to be a subset of $\Omega_n = \{e^{2\pi ik/n} : 0 \leq k \leq n-1\}$. Let T stand for subsets of on the unit circle S^1, respectively subsets of Ω_n, that are "off-limits", in the sense that we require $S \cap T = \emptyset$. Given $l \geq 1$ and T, define

$$f(T, l) = \sup\{\mathcal{CP}(S) : |S| = l,\ S \cap T = \emptyset\}.$$

Alternately given integers $l, n \geq 1$ and $T \subseteq \Omega_n$, define

$$f(T, n, l) = \max\{\mathcal{CP}(S) : S \subseteq \Omega_n,\ |S| = l,\ S \cap T = \emptyset\}.$$

Finally given $\alpha > 0$ and an integer r, define

$$\begin{aligned} g(\alpha, l) &= \inf\{f(T, l) : \mu(T) = \alpha\}, \\ g(r, n, l) &= \min\{f(T, n, l) : T \subseteq \Omega_n,\ |T| = r\}. \end{aligned}$$

Here μ is Lebesgue measure on S^1 except that we take $\mu(S^1) = 1$ instead of 2π.

Proposition 1.

- *(a) If T is closed, then for any l, $\lim_{n\to\infty} f(T, n, l)$ is well-defined and equals $f(T, l)$.*
- *(b) For every fixed l, the function $g(\alpha, l)$ is continuous in α.*
- *(c) For any $\alpha > 0$ and ℓ, taking $r = \lfloor n\alpha \rfloor$, $\lim_{n\to\infty} g(r, n, l)$ exists and equals $g(\alpha, l)$.*

The proof, given in Appendix A, exploits the uniform continuity of $\mathcal{CP}(S)$ over the compact set S^1.

Problem 1. *Given l, n, r and $T \subseteq \Omega_n$ of size r, is there an easily described strategy to compute S of size l achieving the maximum in $f(T, n, l)$? Same question for $f(T, l)$ when T is open.*

Related to this question, what sets T in the above provide the worst-case scenario? That is:

Problem 2. *Given l, n, r, which subsets T of Ω_n, $|T| = r$, minimize $f(T, n, l)$, and what is the minimax value? Same question for (closed) sets T of measure α achieving or approaching the infimum that defines $g(\alpha, l)$, and what is its value?*

We believe that in both cases a minimizing T is an interval of size r in Ω_n, respectively of measure α in S^1. Intuitively, an adversary making T one contiguous block bunches the allowable domain of S most closely together around the circle, preventing one from profiting by choosing points "inside/between" T that make long chords to other regions. There are two caveats in particular against believing this is obvious, however.

First, Donoho and Stark [2] considered the related question of which sets R and C maximize $||\boldsymbol{DFT}^n_{R,C}||_2$, and stated:

Conjecture 1 [2]. *For interval R and set C with $|C| \cdot |R| = n$, $||\boldsymbol{DFT}^n_{R,C}||_2$ is maximized when C is also an interval.*

This conjecture is still open, indicating that strategies that similarly appeal strongly to intuition can be surprisingly hard to verify.

Second, Vandermonde matrices can be numerically volatile. A Vandermonde matrix with nodes selected to be real numbers can be highly ill-conditioned [4]. For Vandermonde matrices with nodes on the unit circle the situation can be nicer, provided the nodes are spread out relatively evenly [3]. However, in our situation the set T can prevent this—intuitively most strongly when T is an interval.

The relationship between the continuous and discrete cases shown in Proposition 1 further motivates our attention to these problems. In this paper, we evaluate the efficacy of approaching Problem 1 with a randomized strategy. Our main result is to give an explicit lower bound for the expected value of $\mathcal{CP}(S)$, if we select S uniformly at random. Considering l and r to be functions of n, this result is shown to be asymptotically tight for a wide range of parameter settings. From this we conclude that essentially the worst the adversary can do to frustrate the uniform random selection strategy, is to pick T to be an interval in Ω_n.

The effects we observe are drastic. We have $\ln g(0, l) = \ln |\det(\boldsymbol{DFT}_l)| = \frac{l}{2} \ln l$, as witnessed by $S =$ the l-th roots of unity. For a randomly selected subset S of size l when no points are disallowed, we will observe that $E[\ln \mathcal{CP}(S)] = \Omega(l^2/n)$. However, for a player selecting randomly from among allowed nth-roots, we observe expected value $E[\ln \mathcal{CP}(S)] = -\Theta(l^2)$, in case the adversary disallows an interval of constant measure ϵ, for any small $\epsilon > 0$.

The rest of this paper is organized as follows. Section 2 contains some mathematical prerequisites. In Section 3, we take a preliminary look at the $\forall\exists$-quantified

optimization problem about minors of the *DFT*$_n$ matrix we mentioned above. In Section 4, we present and evaluate our main result about random Vandermonde matrices. Finally in Section 5, we give an application of the main theorem. We introduce so-called ϵ, l*-index-limited vectors*, which combines Donoho and Stark's [2] notion of ϵ-concentration of a vector with some measure of a vector being limited to an interval or band. For our notion we give an uncertainty relation that is applicable in cases where the Donoho-Stark discrete uncertainty principle trivializes.

2. Prerequisites

Theorem 1 (Weyl Perturbation). *Let A and E be Hermitian matrices. Then*

$$\max_j |\lambda_j(A) - \lambda_j(A+E)| \leq ||E||_2.$$

Theorem 2 (Courant-Fisher). *Let A be an $m \times n$ with $m \geq n$, matrix then for any $i = 1, 2, \ldots, n$, the ith singular value σ_i of A is given by*

$$\sigma_i(A) = \max_{\substack{S \subseteq \mathbf{C}^n \\ dim(S)=i}} \min_{x \in S/\{0\}} \frac{||Ax||_2}{||x||_2},$$

where S ranges over all linear subspaces of dimension i.

Theorem 3 (Binet-Cauchy). *Let A be an $m \times n$ matrix and let B be an $n \times m$ matrix with $n \geq m$. Then*

$$\det(AB) = \sum_{\substack{I \subseteq \{1,2,\ldots,n\} \\ |I|=m}} \det(A^I) \det(B_I),$$

where A^I is the $m \times m$ minor of A consisting of all columns in I, and B_I is the $m \times m$ minor of B consisting of all rows in I.

3. Fourier Matrix Games

Definition 1. We define the Fourier matrix game *DFT-Game*(n, l, k, B) to be the following single-round game against an adversary agent:

Adversary:	selects l distinct rows $r_1, r_2, \ldots, r_l$ and k distinct columns $c_1, c_2, \ldots, c_k$ in $\{0, 1, \ldots, (n-1)\}$.
Player:	selects an $l \times l$ minor M of the $n \times n$ Fourier matrix *DFT*$_n$ with rows $r_1, r_2, \ldots, r_l$ and columns disjoint from $c_1, c_2, \ldots, c_k$.
Result:	The player wins if and only if $\lvert\det(M)\rvert > B$.

We define *DFT-Game*$^*(n, l, k, B)$ the same game as above, but with the modification that the adversary can only choose sets of rows R that are contiguous in the cyclic sense: $R = \{b + i \bmod n : 0 \leq i \leq l - 1\}$ for some *base point* b. For this game, it is not hard to see, we can assume without loss of generality that the adversary's set of rows is fixed to be the first l rows of the DFT_n matrix.

Proposition 2. *If $n = l \cdot k$, then the adversary has a winning strategy for DFT-Game$(n, l, k, 0)$.*

Proof. A winning strategy for the adversary is to take rows $r_i = ki$, for $i = 0, 1, \ldots, (l - 1)$, and columns $c_i = li$, for $i = 0, 1, \ldots, (k - 1)$. Let A be the $l \times n$ minor of DFT_n with rows $r_0, r_1, \ldots r_{l-1}$. The rth column A_r of A equals $(1, \alpha^r, \alpha^{2r}, \ldots, \alpha^{(l-1)r})^T$, where $\alpha = e^{\frac{2\pi i}{n}k} = e^{\frac{2\pi i}{l}}$. Hence for any r, $A_r = A_{r+l \bmod n}$. With columns $0, l, 2l, \ldots, (k - 1)l$ disallowed, there are therefore only $l - 1$ distinct columns in the remaining set, so any $l \times l$ minor of A that avoids the disallowed columns will be singular. □

So if $n = l \cdot k$, there is not much honor to achieve for the player. For $k \cdot l$ below n, Theorem 1 and 3 guarantee the existence of a minor with an at least "fairly reasonable" lower bound on the magnitude of its determinant.

Proposition 3. *The player has a winning strategy for DFT-Game(n, l, k, B), provided $k \cdot l < n$ and $B < (n - kl)^{l/2}\binom{n-k}{l}^{-1/2}$.*

Proof. Suppose the adversary chooses l rows R and k columns C. Let $N = \{0, 1, \ldots, n-1\}$. Let $A = DFT_{R,N/C}$ and $B = DFT_{R,C}$. Then $AA^* = nI - BB^*$. Both AA^* and BB^* are Hermitian, so by Theorem 1, provided $||BB^*||_2 \leq n$, for each i, the ith eigenvalue $\lambda_i(AA^*) \geq n - ||BB^*||_2$. We can write $BB^* = \sum_{i \in C} c_i c_i^*$, where c_i is the ith column of $DFT_{R,N}$. Since $||c_i c_i^*||_2 = ||c_i||_2^2 = l$, then by subadditivity of the ℓ_2-norm, $||BB^*||_2 \leq kl$. Hence $\det(AA^*) \geq (n - kl)^l$. By Theorem 3, $\det(AA^*) = \sum_{|S|=l} |\det(A_{R,S})|^2$. Hence we conclude there exists S of size l such that $|\det DFT_{R,S}| \geq (n - kl)^{l/2}\binom{n-k}{l}^{-1/2}$. □

Our main interest in this paper however, will be the contiguous variant of the Fourier matrix game, in particular for cases where $k \cdot l \geq n$. In the following section we evaluate a randomized player strategy for this game.

4. Random Vandermonde Matrices

For complex numbers $z_0, z_1, \ldots, z_{l-1}$, denote by $V = V(z_0, z_1, \ldots, z_{l-1})$ the $l \times l$ Vandermonde matrix defined by $V_{ij} = z_i^j$ for $0 \leq i, j \leq l - 1$. We have the following theorem:

Theorem 4 (Main Result). *For any $n \geq 7$ and l, r with $0 < r < \frac{n}{\pi}$ and $l + r \leq n$, Let R be a subset of Ω_n of size r. Consider the process of picking $\{z_0, \ldots, z_{l-1}\} \subset \Omega_n \backslash R$ uniformly at random among all subsets of $\Omega_n \backslash R$ of size l. Then for the Vandermonde matrix $V = V(z_0, z_1, \ldots, z_{l-1})$ we have that $E[\ln |\det V|]$ is at least*

$$\frac{\binom{l}{2}}{(n-r)(n-r-1)} \left((n-2r)(\frac{2n}{\pi} \sin \frac{\pi}{n} - \ln 2) - r^2 \ln \frac{n}{r\pi} - r^2 - \frac{r^4 \pi^2}{36 n^2} \right).$$

Let us stress that in the above random selection of a subset of $\Omega_n \backslash R$ we do not assume any particular order among $z_0, z_1, \ldots, z_{l-1}$. In other words, a subset is selected uniformly at random, and uniformly at random we associate variable names $z_0, z_1, \ldots, z_{l-1}$ to its elements. This allows us to treat the z_i variables symmetrically in the proof. For the proof of Theorem 4 we need an estimate involving the "ln-of-chord-length" function $f(t) = \ln |1 - e^{it}|$, for $t \in \mathbf{R} \setminus \{k2\pi : k \in \mathbf{Z}\}$. Straightforward geometry gives us:

$$f(t) = \frac{1}{2} \ln(2 - 2\cos t),$$

which can be rewritten using the relation $\sin^2 \frac{\alpha}{2} = \frac{1-\cos\alpha}{2}$ as

$$f(t) = \ln 2 + \ln |\sin \frac{t}{2}|.$$

We will also consider a discretized version of this function, which per abuse of notation will also be denoted by f; which version "f" refers to will be clear from the context.

Lemma 1. *Let $n \geq 7$, and let $\omega = e^{2\pi i/n}$. Define the discrete function $f(d) = \ln |1 - \omega^d|$, for $d = 1, 2, \ldots, n-1$. Then*

- **(a)** $\sum_{d=1}^{n-1} f(d) \geq \frac{2n}{\pi} \sin \frac{\pi}{n} - \ln 2$, *and*
- **(b)** $\sum_{d=1}^{n-1} f(d) \leq 2 \ln \frac{n}{2\pi} + 2 + \ln 2 + \frac{2\pi^2}{9n^2}$.

A proof of the above lemma can be found in Appendix B. To a subset $R \subset N = \{0, 1, \ldots, n-1\}$ we associate its characteristic function $\chi : N \to \{0, 1\}$, which is defined by $\chi(i) = 1$ iff $i \in R$. Given a characteristic function χ, we define the function $c_\chi : N \to \mathbf{N}$ by

$$c_\chi(d) = \sum_{i=0}^{n-1} \chi(i)\chi(i + d \bmod n).$$

We identify subsets of $\Omega_n = \{e^{2\pi ik/n} : 0 \leq k \leq n-1\}$ and N in the obvious manner.

Lemma 2. *For any n, l, and r with $l + r \leq n$, Let R be an arbitrary subset of Ω_n of size r. Consider the process of picking $\{z_0, \ldots, z_{l-1}\} \subset \Omega_n \backslash R$ uniformly at random among all subsets of $\Omega_n \backslash R$ of size l. Then for the Vandermonde matrix $V = V(z_0, z_1, \ldots, z_{l-1})$ we have*

$$E[\ln |\det V|] = \frac{\binom{l}{2}}{(n-r)(n-r-1)} \left((n-2r) \sum_{d=1}^{n-1} f(d) + \sum_{d=1}^{n-1} f(d) c_\chi(d) \right) \quad ,$$

where χ is the characteristic function of R.

Proof.

$$\begin{aligned}
E[\ln |\det V|] &= E[\ln \prod_{i<j} |z_i - z_j|] \\
&= E[\sum_{i<j} \ln |z_i - z_j|] \\
&= \sum_{i<j} E[\ln |z_i - z_j|] && \text{(by linearity of E)} \\
&= \binom{l}{2} E[\ln |z_0 - z_1|]. && \text{(by symmetry)}
\end{aligned}$$

Let $\eta = E[\ln |z_0 - z_1|]$. We can write the following expression for η:

$$\eta = \sum_{p \in \Omega_n \backslash R} \sum_{q \in \Omega_n \backslash R, q \neq p} \Pr[z_0 = p \text{ and } z_1 = q] \ln |p - q|,$$

where $\Pr[(z_0 = p \text{ and } z_1 = q] = \frac{1}{(n-r)(n-r-1)}$. Let $\overline{\chi}$ correspond to the characteristic function of $\Omega_n \backslash R$. We have that

$$\begin{aligned}
(n-r)(n-r-1) \cdot \eta &= \sum_{p \in \Omega_n \backslash R} \sum_{q \in \Omega_n \backslash R, q \neq p} \ln |p - q| \\
&= \sum_{i=0}^{n-1} \sum_{j=0, j \neq i}^{n-1} \overline{\chi}(i) \overline{\chi}(j) \ln |\omega^i - \omega^j| \\
&= \sum_{i=0}^{n-1} \sum_{d=1}^{n-1} \overline{\chi}(i) \overline{\chi}(i + d \bmod n) \ln |\omega^i - \omega^{i+d}| \\
&= \sum_{i=0}^{n-1} \sum_{d=1}^{n-1} \overline{\chi}(i) \overline{\chi}(i + d \bmod n) \ln |1 - \omega^d| \\
&= \sum_{i=0}^{n-1} \sum_{d=1}^{n-1} \overline{\chi}(i) \overline{\chi}(i + d \bmod n) f(d) \\
&= \sum_{d=1}^{n-1} f(d) \sum_{i=0}^{n-1} \overline{\chi}(i) \overline{\chi}(i + d \bmod n).
\end{aligned}$$

The lemma now follows by rewriting $\overline{\chi}(i) = 1 - \chi(i)$, where χ is the characteristic function of R, and observing that

$$\begin{aligned}\sum_{i=0}^{n-1} \overline{\chi}(i)\overline{\chi}(i+d \bmod n) &= \sum_{i=0}^{n-1}(1-\chi(i))(1-\chi(i+d \bmod n)) \\ &= (n-2r) + \sum_{i=0}^{n-1}\chi(i)\chi)(i+d \bmod n).\end{aligned}$$

□

We now prove Theorem 4.

Proof. From Lemma 2 we have that

$$E[\ln|\det V|] = \frac{\binom{l}{2}}{(n-r)(n-r-1)}\left((n-2r)\sum_{d=1}^{n-1} f(d) + \sum_{d=1}^{n-1} f(d)c_\chi(d)\right) \quad ,$$

where χ is the characteristic function of R. We have that

$$\begin{aligned}\sum_{d=1}^{n-1} c_\chi(d) &= \sum_{d=1}^{n-1}\sum_{i=0}^{n-1}\chi(i)\chi(i+d \bmod n) \\ &= \sum_{i=0}^{n-1}\chi(i)\sum_{d=1}^{n-1}\chi(i+d \bmod n) \\ &= r^2 - r.\end{aligned}$$

We know that $\sum_{d=1}^{n-1} f(d)c_\chi(d)$ is smallest if the total mass r^2-r is placed at much as possible at places where $f(d)$ is the smallest. Note that for any d, $0 \leq c_\chi(d) \leq r$. Define $\epsilon(t) = \ln|t| - f(t)$. By the concavity of f, in case r is odd,

$$\begin{aligned}\sum_{d=1}^{n-1} f(d)c_\chi(d) &\geq \sum_{d=1}^{(r-1)/2} f(d)r + \sum_{d=n-\frac{r-1}{2}}^{n-1} f(d)r \\ &= 2r\sum_{d=1}^{(r-1)/2} f(d) \\ &\geq 2r\frac{n}{2\pi}\int_0^{\frac{r-1}{2}\frac{2\pi}{n}} f(t)dt \\ &= \frac{rn}{\pi}\int_0^{\frac{(r-1)\pi}{n}} \ln t - \epsilon(t)dt \\ &\geq \frac{rn}{\pi}[t\ln t - t - \frac{t^3}{36}]_0^{\frac{(r-1)\pi}{n}} \qquad \{\text{by Lemma B.1}\} \\ &= r(r-1)\ln\frac{(r-1)\pi}{n} - (r-1)r - \frac{1}{36}\frac{r(r-1)^3\pi^2}{n^2} \\ &\geq r^2\ln\frac{r\pi}{n} - r^2 - \frac{r^4\pi^2}{36n^2},\end{aligned}$$

In case r is even, the same lower bound can be obtain similarly. Together with Lemma 1 this yields the theorem. □

Corollary 1. *For any $n \geq 7$ and any l, r with $0 < r < \frac{n}{\pi}$ and $l + r \leq n$, the player has a winning strategy for DFT-Game$^*(n, l, r, e^C)$, provided C is less than*

$$\frac{\binom{l}{2}}{(n-r)(n-r-1)} \left((n-2r)(\frac{2n}{\pi} \sin \frac{\pi}{n} - \ln 2) - r^2 \ln \frac{n}{r\pi} - r^2 - \frac{r^4\pi^2}{36n^2} \right).$$

Proof. Recalling our remark after Definition 1, we can assume with loss of generality that the adversary chooses rows $R = \{0, 1, \ldots, l-1\}$. Any $l \times l$ minor of DFT_n with rows R is a Vandermonde matrix. Let C be the set of columns the adversary chooses. Theorem 4 gives a lower bound on $E[\ln |\det(M)|]$ for randomly selected $l \times l$ minor M of DFT_n with rows R avoiding columns C. There must exist at least one minor M' that has $\ln |\det(M')| \geq E[\ln |\det(M)|]$. So the player chooses such a minor, for which we then have the lower bound on the absolute value of its determinant as stated in the corollary. □

4.1 Sharpness of the Result

We are interested in the asymptotic growth as a function of n of the lower bound given in Theorem 4, where we consider the variables r and l to be certain functions of n, which if we want to be explicit about this will be denoted by $l(n)$ and $r(n)$. For certain growth rates of $l(n)$ and $r(n)$, Problem 1 is trivial. For example, assuming $l(n)$ divides n for simplicity, for $r(n) < \frac{n}{l(n)}$, there always exist some equally spaced selection of $l(n)$ points that is not blocked. This then yields an optimal magnitude for the determinant of $l^{l/2}$. At the end of this section we will give an example that compares our random strategy with this optimal value. First however, we study the optimality of our analysis for cases where this triviality is avoided. For this, it is good to keep in mind that we typically are interested in both $l(n)$ and $r(n)$ growing much faster than $\sqrt{n}$. In that case, straightforward perturbation techniques to yield strategies for the Fourier matrix game like Proposition 3 appear to stop working.

4.1.1 Small Value Example

Note that for $r(n) = \Omega(\sqrt{n})$ and $r = o(n)$, the growth of the expression for the (ln of the) determinant in Theorem 4 is dominated by the term $-\frac{\binom{l}{2}}{(n-r)(n-r-1)} r^2 \ln \frac{n}{r\pi}$. To give an idea how bad things can get, for $r(n) > \frac{n}{\sqrt{l(n)}}$ this starts comparing unfavorably even to the natural log of the *reciprocal* of the optimal value of $l^{l/2}$ for l equally spaced points. This may raise doubts about the tightness of Theorem 4, but we will show that the adversary can indeed frustrate the random player strategy to such extents. We will prove our result to be asymptotically tight, at least for a wide range of functions r. We will show that essentially the worst-case scenario arises

when the adversary chooses the set of disallowed roots of unity as a contiguous block, and that in this case we get an upper bound matching the lower bound of Theorem 4. This sheds some light on Open Problem 2 mentioned in the introduction. We believe that picking contiguous blocks is optimal for the adversary, not only against the "uniform random player", but against *any* player.

For fixed l, r and n, minimization of the expression given in Lemma 2 is done by minimizing the $\sum_{d=1}^{n-1} f(d)c_\chi(d)$ term. Define $\rho_{n,r} = \sum_{d=1}^{n-1} f(d)c_\chi(d)$, where χ is the characteristic function of an arbitrary contiguous subset of N (in the cyclic sense) of size r. Note that $\rho_{n,r}$ is well-defined. One easily observes that for any $0 \leq r < n$, $\rho_{n,r} = 2\sum_{d=1}^{r-1}(r-d)f(d)$. Note this does not differ by too much from our conservative estimate $\sum_{d=1}^{n-1} f(d)c_\chi(d) \geq 2r\sum_{d=1}^{(r-1)/2} f(d)$, we used to prove Theorem 4. We have the following upper and lower bounds (See Appendix C for a proof):

Proposition 4. *For $2 \leq r \leq \frac{n}{2\pi}$,*

(a) $\rho_{n,r} \leq -r^2 \ln \frac{n}{2\pi r} - \frac{3r^2}{2} + (1-2r)\ln\frac{n}{2\pi} + 2r - \frac{1}{2}$.

(b) $\rho_{n,r} \geq -(r^2-1)\ln\frac{n}{2\pi(r-1)} - \frac{3r^2}{2} + r + \frac{1}{2} - \frac{2r^4\pi^2}{9n^2}$.

Applying Lemma's 1 and 2 and Proposition 4, the following theorem follows straightforwardly:

Theorem 5. *For some large enough constants $c_0 > 0$ and small enough constant $c_1 > 0$, for all large enough n, provided $c_0\sqrt{n\ln n} < r < c_1 n$ and $l + r \leq n$, then the following holds: Consider the process of picking l distinct roots $z_0, z_1, \ldots, z_{l-1}$ uniformly at random from the nth roots of unity, where a contiguous block of r many roots is disallowed. Then we have for the Vandermonde matrix $V = V(z_0, z_1, \ldots, z_{l-1})$ that*

$$E[\ln|\det V|] = -\Theta\left(\frac{l^2r^2}{(n-r)^2}\ln\frac{n}{r}\right).$$

To give a striking example, say $r(n) = \lfloor \alpha n \rfloor$, for some small enough constants α. Then $E[\ln|\det V|] = -\Theta(l^2)$. Optimally for l points, $\ln|\det V| = \frac{l}{2}\ln l$. In other words, cut out any small constant size sector of the unit circle, and randomly selecting a Vandermonde supported on the remainder of the circle, is expected to do even worse than reciprocal of the optimum value. Note also Lemma's 1 and 2 imply that if no points are disallowed, for a randomly selected Vandermonde V, we have $E[\ln|\det V|] = \Omega(l^2/n)$. For r as considered in Theorem 5, the lower bound given in Theorem 4 is

$$E[\ln|\det V|] = -\Omega\left(\frac{l^2r^2}{(n-r)^2}\ln\frac{n}{r}\right).$$

We conclude that picking a contiguous block of disallowed points, is the worst the adversary can do to the player that selects uniformly at random, at least for r as

considered in Theorem 5. We leave it as an open problem to prove or disprove that this is true against any optimal player strategy.

4.1.2 Large Value Example

Assume both l and r divide n, and let us consider the scenario where the forbidden set R consists of r equally spaced points, and that it is possible to select l equally spaced points disjoint from R. For the resulting Vandermonde matrix V we then have $\ln|\det V| = \frac{l}{2}\ln l$. Quantitatively, how does randomly selecting the Vandermonde compare to this ?

Letting χ denote the characteristic function of R, one can verify that the function $c_\chi(d)$ mentioned in Lemma 2 equals r for d being multiples of $\frac{n}{r}$ and is zero otherwise. Hence we can use Lemma 1 to bound the term $\sum_{d=1}^{n-1} f(d)c_\chi(d)$ of Lemma 2, with the only difference being that we are now summing over rth roots instead of nth roots of unity. We hence conclude that picking l points at random in the current scenario gives $E[\ln|\det V|]$ to be at least

$$\frac{\binom{l}{2}}{(n-r)(n-r-1)}\left((n-2r+1)(\frac{2n}{\pi}\sin\frac{\pi}{n}-\ln 2)-\frac{2\pi^2}{3r^2}\right).$$

For $r \leq \alpha n$, for some small enough constant $\alpha > 0$, we thus have

$$E[\ln|\det V|] = \Omega(\frac{l^2}{n}).$$

We conclude in this case, that the random strategy features positive growth for $l = \omega(\sqrt{n})$.

5. Application

For an n-vector f, define the *support of* f to be the set $\text{supp}(f) = \{i : f_i \neq 0\}$. Following Donoho and Stark [2], we say an n-vector f is ϵ*-concentrated* on a set T of indices if

$$\sqrt{\sum_{i\notin T}|f_i|^2} \leq \epsilon.$$

Theorem 6 [2]. *For any n-vector f with $||f||_2 = 1$ that is ϵ_T-concentrated on a set T and $\hat{f} = F_n f$ being ϵ_Ω-concentrated on a set Ω, we have that*

$$|T|\cdot|\Omega| \geq n(1-(\epsilon_T+\epsilon_\Omega))^2. \tag{5.1}$$

As a side note, Theorem 6 yields a "fairly" good strategy for playing the Fourier matrix game, comparable in strength to Proposition 3, which is slightly stronger. See Appendix D for a proof.

Proposition 5. *For any l, r with $lr \leq n$ and $l + r \leq n$, the player has a winning strategy for DFT-Game(n, l, r, B). for any $B < (\sqrt{n} - \sqrt{lr})^l \binom{n-r}{l}^{-1/2}$.*

Definition 2. An n-vector f is called *l-index-limited* if supp$(f) \subseteq \{b + i \bmod n : 0 \leq i \leq l-1\}$, for some number b. An n-vector f is called *ϵ, l-index-limited* if there exists g with $||g||_2 \leq \epsilon$ such that $f - g$ is l-index-limited.

Theorem 6 is trivialized in case $|T| \cdot |\Omega| > n$. We use Corollary 1, to give an uncertainty type relation that does manage to express non-trivial lower-bounds on concentration in case $|T| \cdot |\Omega| > n$, when dealing with ϵ, l-index-limited vectors.

Lemma 3. *Suppose the player has a winning strategy for DFT-Game$^*(n, l, k, B)$. Then for any n-vector f with $||f||_2 = 1$ that is ϵ, l-index-limited, and any set Ω of size r with $r \leq k$, $\hat{f} = F_n f$ is ϵ_Ω-concentrated on Ω with*

$$\epsilon_\Omega > (1 - \epsilon)\frac{B}{n^{l/2}} - \epsilon.$$

Proof. Consider an arbitrary Fourier transform pair $(f, \hat{f})$ and let $T = \{b + i \bmod n : 0 \leq i \leq l-1\}$ be a contiguous set of indices containing supp$(f - g)$ with g some vector with $||g||_2 \leq \epsilon$, and $||f||_2 = 1$. Consider an arbitrary set of indices Ω of size r with $r \leq k$. By definition of the relaxed Fourier game and the fact that the Fourier matrix is symmetric, there exists $l \times l$ minor V of DFT_n with columns T and rows avoiding Ω such that

$$|\det(V)|^2 \geq B^2.$$

We get for the smallest singular value σ_l of $\frac{1}{\sqrt{n}}V$

$$\sigma_l(\frac{1}{\sqrt{n}}V) > \frac{B}{n^{l/2}}.$$

Let Ω' be the rows of V. Write

$$(F_n f)_{\Omega'} = (F_n(f - g) + F_n g)_{\Omega'} = \frac{1}{\sqrt{n}}V(f - g)_T + (F_n g)_{\Omega'}.$$

By the max-min characterization of singular values given by Theorem 2, we have that

$$||\frac{1}{\sqrt{n}}V(f - g)_T)||_2 \geq \sigma_l(\frac{1}{\sqrt{n}}V)||f - g||_2 > (1 - \epsilon)\frac{B}{n^{l/2}}.$$

Since $||(F_n g)_{\Omega'}|| \leq \epsilon$, we get by the triangle inequality that

$$||\hat{f}_{\Omega'}||_2 > (1 - \epsilon)\frac{B}{n^{l/2}} - \epsilon.$$

Since Ω' is disjoint from Ω we conclude $\hat{f}$ is ϵ_Ω concentrated on Ω with $\epsilon_\Omega > (1 - \epsilon)\frac{B}{n^{l/2}} - \epsilon$. □

Theorem 7. *Let $n \geq 7$. Suppose f is an n-vector with $||f||_2 = 1$ that is ϵ, l-index-limited with Fourier transform $\hat{f} = F_n f$. Then for any set Ω of size r with $0 < r < \frac{n}{\pi}$ and $l + r \leq n$, $\hat{f}$ is ϵ_Ω-concentrated on Ω with*

$$\epsilon_\Omega \geq (1-\epsilon)\frac{e^B}{n^{l/2}} - \epsilon,$$

where B is given by

$$\frac{\binom{l}{2}}{(n-r)(n-r-1)}\left((n-2r)(\frac{2n}{\pi}\sin\frac{\pi}{n} - \ln 2) - r^2 \ln\frac{n}{r\pi} - r^2 - \frac{r^4\pi^2}{36n^2}\right).$$

Proof. This follows immediately from the player strategy shown to exist in Corollary 1 and applying Lemma 3. □

The lower-bound on concentration on Ω is fairly weak, but we should stress this bound is given for any conceivable set Ω, not just contiguous ones. The most notable fact is that our theorem still yields non-trivial lower bounds on concentration in case both $l, r \gg \sqrt{n}$, which is a breaking point for typical straightforward calculations. For example, Theorem 6 yields a trivial lower-bound of $\epsilon_\Omega \geq 0$ in case $|T| \cdot |\Omega| \geq n$. To give an extreme example, for $r = \lfloor \alpha n \rfloor$, for some small enough absolute constant $\alpha > 0$, we get concentration $\epsilon_\Omega \geq (1-\epsilon)e^{-\Theta(l^2)} - \epsilon$, for any set Ω of size r.

6. Future Directions

Remaining are the Open Problems 1 and 2 mentioned in the introduction. Will one be able to observe the extreme squashing of the determinant value as seen in Section 4.1, when the player's strategy is optimal instead of just random ? We believe this will occur, because of eigenvalue clustering phenomena similar to those studied in the landmark work of Slepian [7]. We ask whether a discrete analogue of the results by Slepian can be be developed to properly study this. A first step in this direction has been taken by Grünbaum [5].

Acknowledgments

We thank Prof. Dr. Karlheinz Gröchenig and Prof. Terence Tao for helpful communications.

References

[1] E.J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.

[2] D.L. Donoho and P.B. Stark. Uncertainty principles and signal recovery. *SIAM J. App. Math.*, 49:906–931, 1989.

[3] P. J. S. G. Ferreira. Superresolution, the recovery of missing samples, and Vandermonde matrices on the unit circle. In *Proc. Workshop on Sampling Theory and App.*, pages 216–220, 1999.

[4] W. Gautschi. Norm estimates for inverses of Vandermonde matrices. *Numer. Math.*, 23:337–347, 1975.

[5] F. A. Grünbaum. Eigenvectors of a Toeplitz matrix: discrete version of the prolate spheroidal wave functions. *SIAM J. Alg. Disc. Meth.*, 2:136–141, 1981.

[6] L. Rade and B. Westergren. *Mathematics Handbook for Science and Engineering, 5th ed.* Springer Verlag, 2004.

[7] D. Slepian. Prolate spheroidal wave functions, Fourier analysis, and uncertainty - v: The discrete case. *Bell System Technical Journal*, 57(5):1371–1430, 1978.

[8] P. Stevenhagen and H.W. Lenstra Jr. Chebotarëv and his density theorem. *Mathematical Intelligencer*, 18(2):26–37, 1996.

[9] T. Tao. An uncertainty principle for cyclic groups of prime order. *Mathematical Research Letters*, 12:121–127, 2005.

Department of Computer Science, University of Aarhus
IT-Parken, Aabogade 34, DK-8200 Aarhus N, Denmark
email: mjjansen@daimi.au.dk

Department of Computer Science and Engineering, University at Buffalo
201 Bell Hall, Buffalo, NY 14260-2000, USA
email: regan@cse.buffalo.edu

A. Proof of Proposition 1

Let μ stand for Lebesgue measure on the unit circle S_1, with $\mu(S_1) = 1$ (not 2π). Let d_1 stand for circular distance, so for $x, y \in S_1$, $|x-y| \leq d_1(x,y) \leq |x-y|\pi/2$. We extend the definition of chordal product to be a symmetric function on S_1^l, and given an l-vector $S \in S_1^l$, let $\hat{S}$ denote the multi-set of its components. Note that if S has a duplicate entry then $\mathcal{CP}(S) = 0$. By continuity of $\det$, and hence uniform continuity of $\mathcal{CP}$ on the compact set S_1^l, for all $\epsilon > 0$ there exists $\delta > 0$ such that whenever $||S - S'||_\infty \leq \delta$, $|\mathcal{CP}(S) - \mathcal{CP}(S')| \leq \epsilon$. When $||S - S'||_\infty$ is much less than the minimum distance between distinct points of $\hat{S}$ or of $\hat{S'}$, then it does no harm to ignore the distinction between S and $\hat{S}$. For $T \subset S_1$, let $\sim T$ stand for $S_1 \setminus T$, and note that the definitions of f and g also extend naturally:

$$\begin{aligned} f(T,l) &= \sup\{\, \mathcal{CP}(S) : S \in S_1^l,\ \hat{S} \cap T = \emptyset \,\} \\ f(T,n,l) &= \max\{\, \mathcal{CP}(S) : S \in \Omega_n^l,\ \hat{S} \cap T = \emptyset \,\}. \\ g(\alpha,l) &= \inf\{\, f(T,l) : \mu(T) = \alpha \,\}, \\ g(r,n,l) &= \min\{\, f(T,n,l) : T \subset \Omega_n, |T| = r \,\}. \end{aligned}$$

Lemma A.1. *Let $\delta > 0$, let T be a proper subset of S^1 with $\mu(T) > 2\delta$, and let $U = \{\, u \in T : (\exists y \in \sim T)\ d_1(u,y) \leq \delta \,\}$. Then $\mu(U) \geq 2\delta$, with equality iff T equals the union of an interval and a set of measure zero.*

Proof. Let $p_0 = q_0 =$ some point in $\sim T$. Let $\eta > 0$, $\eta < \delta$. At any stage $i \geq 0$ we will have a point p_i clockwise ("negative") from p_0 and a point q_i counterclockwise ("positive") from q_0, with $d_1(p_i, q_i) \geq i\eta$, and the invariant that every point in T from p_i up to q_i belongs to U. If $[p_i - \delta, p_i - \eta] \cup [q_i + \eta, q_i + \delta] \subseteq T$, then these intervals are in U, and so $\mu(U) \geq 2\delta - 2\eta$. Else, it is possible to pick $p_{i+1} \notin T$ in the former interval, and/or $q_{i+1} \notin T$ in the latter interval (if only one is possible, the other stays unchanged), thus meeting the conditions for the next stage. Since the minimum "step" η is constant, after finitely many stages either p_i and q_i meet at the other end of the circle, whence $T = U$, or we've proved $\mu(U) \geq 2\delta - 2\eta$. Since η is arbitrary and $\mu(T) > 2\delta$, either way gives $\mu(U) \geq 2\delta$.

Clearly equality holds when T is an interval plus a nullset. For inequality otherwise, let $V \subseteq T$ be a maximum interval that is closed in T. If $\mu(V) \leq 2\delta$ then $U = T$ and so $\mu(U) > 2\delta$. Wlog. assume $T \setminus V$ is not a nullset. Hence it is possible to find points $p_0, q_0 \notin T$ such that the interval from p_0 up to q_0 contains V, and the interval from p_0 down to q_0 the other way contains a positive-measure subset T' of T. Then V contributes 2δ to U, while on repeating the argument of the first part of the proof, T' contributes nonzero measure to U. Hence we conclude that $\mu(U) > 2\delta$. □

To re-state Proposition 1, we need to prove that

(a) $f(T, l) = \lim_{n\to\infty} f(T, n, l)$,
(b) $g(\cdot, l)$ is continuous, and
(c) $g(\alpha, l) = \lim_{n\to\infty} g(\lfloor \alpha n \rfloor, n, l)$.

Proof. (of Proposition 1): (a) For all T and n, clearly $f(T, n, l) \leq f(T, l)$. Given $\epsilon > 0$, find S_0 giving $f(T, l) - \mathcal{CP}(S_0) \leq \epsilon/2$. Since $S_0 \cap T = \emptyset$, S_0 is finite, and T is closed, the minimum distance $\delta_0 = \min\{\, d_1(x, y) : x \in S_0, y \in T \,\}$ from S_0 to T is well-defined and positive. By uniform continuity we may find $\delta > 0$, with also $\delta < \delta_0$, such that whenever $|S'| = l$ and $||S_0 - S'||_\infty \leq \delta$, $|\mathcal{CP}(S') - \mathcal{CP}(S_0)| \leq \epsilon/2$. Thus whenever $n > 1/\delta$, there is $S' \in \Omega_n^l$ with $\hat{S}' \cap T = \emptyset$ such that $||S_0 - S'||_\infty \leq \delta$, and so

$$f(T, n, l) \geq \mathcal{CP}(S') \geq \mathcal{CP}(S_0) - \epsilon/2 \geq f(T, l) - \epsilon.$$

Hence for all sufficiently large n, $|f(T, n, l) - f(T, l)| \leq \epsilon$.

(b) Let $\epsilon > 0$. Similarly to (a) we can choose $\delta_0 > 0$ such that for all $S, S' \in S_1^l$ such that $||S - S'||_\infty \leq \delta_0$, $|\mathcal{CP}(S) - \mathcal{CP}(S')| \leq \epsilon/3$. Since $g(\cdot, l)$ is nonincreasing, it suffices to show that for any $\delta \leq \delta_0$, and all $\alpha > \delta$, $g(\alpha - \delta, l) \leq g(\alpha, l) + \epsilon$. We can choose a closed set T with $\mu(T) = \alpha$ such that $f(T, l) \leq g(\alpha, l) + \epsilon/3$, since $g(\alpha, l)$ is an infimum.

By Lemma A.1 (for "$\delta/2$") we can find $U \subseteq T$ such that $\mu(U) = \delta$ and every point in U is within $\delta/2$ of the boundary of T. Define $T' = T \setminus U$, so that $\mu(T') = \alpha - \delta$. Then since $f(T', l)$ is a supremum, we can take $S' \in S_1^l$ such that

$\hat{S'} \cap T' = \emptyset$ and $f(T', l) \leq \mathcal{CP}(S') + \epsilon/3$. Finally, by construction of U there exists $S \in S_1^l$ such that $\hat{S} \cap T = \emptyset$ and $||S - S'||_\infty \leq \delta$. This yields

$$\begin{aligned} g(\alpha - \delta, l) &\leq f(T', l) \qquad \text{(since } g(\cdots) \text{ is an infimum)} \\ &\leq \mathcal{CP}(S') + \epsilon/3 \\ &\leq \mathcal{CP}(S) + 2\epsilon/3 \\ &\leq f(T, l) + 2\epsilon/3 \\ &\leq g(\alpha, l) + \epsilon. \end{aligned}$$

(c) Let $\epsilon > 0$, and choose $\delta > 0$ such that for all $S, S' \in S_1^l$, $||S - S'||_\infty \leq \delta$ implies $|\mathcal{CP}(S) - \mathcal{CP}(S')| \leq \epsilon/2$. First, we show that for $n > 1/\delta$, $g(\lfloor \alpha n \rfloor, n, l) \geq g(\alpha, l) - \epsilon/2$. Setting $r = \lfloor \alpha n \rfloor$, take $T_n \subset \Omega_n$ with $|T_n| = r$ and $S_n \in \Omega_n^l$ giving $g(r, n, l) = f(T_n, n, l) = \mathcal{CP}(S_n)$ and $\hat{S_n} \cap T_n = \emptyset$. Define $T \subset S_1$ of measure r/n by unioning closed intervals of width $1/n$ of the circle centered on the points of T_n. Then $\hat{S_n} \cap T = \emptyset$, so $\mathcal{CP}(S_n) \leq f(T, l)$. Moreover, for every $S \in S_1^l$ such that $S \cap T = \emptyset$, there exists $S' \in \Omega_n^l$ such that $\hat{S'} \cap T = \hat{S'} \cap T_n = \emptyset$ and $||S - S'||_\infty \leq \delta$ (actually, $\leq \delta/2$). Hence $f(T, l) \leq f(T_n, n, l) + \epsilon/2$. Since $r/n \leq \alpha$, this gives

$$g(\alpha, l) \leq g(r/n, l) \leq f(T, l) \leq f(T_n, n, l) + \epsilon/2 = g(r, n, l) + \epsilon/2$$

as needed. It remains to rule out the possibility that $g(r, n, l) > g(\alpha, l) + \epsilon$.

Now since $f(\bar{T}, l) \leq f(T, l)$, we may take a closed T_0 of measure α giving $f(T_0, l) - g(\alpha, l) \leq \epsilon/2$. Take a covering C of T_0 by open intervals of total measure (at most) $\alpha + \delta$. By compactness, C has a finite sub-covering C', and define T_1 to be the closure of C'. Then T_1 consists of a finite number of closed intervals, which can be further regarded as a (possibly smaller) finite number m of disjoint closed intervals. Take $n > 1/\delta$ and set $T_n = T_1 \cap \Omega_n$. (It is not necessary to arrange also that every pair of consecutive intervals in T_1 going around the circle is separated by a point in Ω_n.) Then T_n consists of (at most) m-many disjoint intervals in Ω_n, and $|T_n| \leq r + \delta n$.

In case $|T_n| < r$, we define T' by adding $r - |T_n|$ arbitrary points to T_n.

$$\begin{aligned} g(r, n, l) \leq f(T', n, l) &\leq f(T_n, n, l) \\ &= f(T_1, n, l) \\ &\leq f(T_1, l) \\ &\leq f(T_0, l) \qquad \text{(since } T_0 \subseteq T_1) \\ &\leq g(\alpha, l) + \epsilon/2. \end{aligned}$$

In case $|T_n| \geq r$, define T' by removing $q = |T_n| - r \leq \delta n$ points from the end of one of the intervals that comprise T_n, so that $|T'| = r$ exactly. Then for every $S' \in \Omega_n^l$ such that $\hat{S'} \cap T' = \emptyset$ there exists $S_n \in \Omega_n^l$ such that $\hat{S_n} \cap T_n = \emptyset$

and $||S' - S_n||_\infty \leq q/n \leq \delta$. Thus $f(T', n, l) \leq f(T_n, n, l) + \epsilon/2$. In full this gives:

$$\begin{aligned} g(r, n, l) \leq f(T', n, l) &\leq f(T_n, n, l) + \epsilon/2 \\ &\leq f(T_1, l) + \epsilon/2 \\ &\leq f(T_0, l) + \epsilon/2 \qquad \text{(since } T_0 \subseteq T_1\text{)} \\ &\leq g(\alpha, l) + \epsilon. \end{aligned}$$

This gives that for all $n > 1/\delta$, $|g(\lfloor \alpha n \rfloor, n, l) - g(\alpha, l)| \leq \epsilon$. □

Lemma A.1 and its use in the proof of Proposition 1.1 suggest an attack on the open problems in our Introduction. Notice that the proof used only the conclusion $\mu(U) \geq 2\delta$, not the strict inequality when T is not an interval (plus a nullset). The strict inequality suggests "slack" that might be used to derive a contradiction from the infimum of $\{ f(T, l) : \mu(T) = \alpha \}$ being achieved by some T that is not an interval (plus a nullset), or not being approachable by sets T_ϵ that converge pointwise to an interval (give or take a nullset).

Given any $\delta > 0$, say that a set T' *δ-guards* T if for every l-set S' disjoint from T', there is an l-set S disjoint from T that is pointwise within δ of S'. Then Lemma A.1 can be read as saying that intervals maximize the infimum of $\alpha' = \mu(T')$ such that T' is a subset of T that *δ-guards* T, namely at measure $\alpha - 2\delta$ with $T' = T \setminus U$. Thus if T is not an interval (plus a nullset), then for some $\delta > 0$, T' *δ-guards* T but T' has measure α' smaller than $\alpha - 2\delta$.

The objective then becomes to argue, using particularities of the chordal-product function, that all sets T' of measure α' must allow an S' whose chordal product is greater than $f(T, l)+$ the bound for $|\mathcal{CP}(S) - \mathcal{CP}(S')|$ when S and S' pointwise differ by at most δ. Thus we need to examine further the gradient of $g(\alpha, l)$ with α, compared to the continuity bound on $\mathcal{CP}(S)$. The argument might have the character of an induction on α, using the basis that when $\alpha = 1/l$, the infimum is trivially achieved by an interval because sets T of that measure have no effect in keeping $\sup\{ \mathcal{CP}(S) : S \cap T = \emptyset \}$ below its maximum $l^{l/2}$. Steps that may help further it are:

- Showing that the infimum is always achieved by some set T, and limiting the Borel complexity of T.
- Showing that if T achieves the infimum for measure α, and $\alpha' < \alpha$, then the infimum for measure α' is achievable by a subset of T.

In any event, we suspect that progress would require a finer perturbative analysis of the chordal-product function than we have employed in this paper.

B. Proof of Lemma 1

Lemma B.1. *Let $\epsilon(t) = \ln|t| - f(t)$. Then for any t with $|t| < 1$, $0 < \epsilon(t) < \frac{t^2}{12}$.*

Proof. First of all for any t, $f(t) = \ln|1 - e^{it}| < \ln|t|$. We thus see that $\epsilon(t)$ is non-negative. For $t \in (0, 2\pi)$, we have for the error function $\epsilon(t) = \ln|t| - f(t) = \ln\frac{|t|}{2\sin\frac{t}{2}}$. For $t > 0$, $\sin t \geq t - \frac{t^3}{6}$. So on this interval, $\epsilon(t) \leq \ln\frac{t}{t-\frac{t^3}{24}} = -\ln(1 - \frac{t^2}{24})$. For $\frac{-1}{24} < x < \frac{1}{24}$, $\ln(1+x) \geq x - \frac{x^2}{2}$. So for $0 < t < 1$, $0 < \epsilon(t) < \frac{t^2}{24} + \frac{t^4}{1152} < \frac{t^2}{12}$. The lemma follows by symmetry of $f(t)$ and $\ln|t|$. □

Proof. *(of Lemma 1)* We have

$$\int_0^{\pi\backslash 2} \ln(\sin x)dx = -\frac{\pi}{2}\ln 2,$$

See e.g. [6], p. 182, equation 55. Hence

$$\begin{aligned}\int_0^{2\pi} f(t)dt &= 2\pi\ln 2 + \int_0^{2\pi} \ln\sin\frac{t}{2}dt \\ &= 2\pi\ln 2 + 2\int_0^{\pi} \ln\sin\frac{t}{2}dt \\ &= 2\pi\ln 2 + 4\int_0^{\pi\backslash 2} \ln\sin x dx = 0\end{aligned}$$

For $j = 0, 1, , \ldots, n-1$, define interval $I_j = [j\frac{2\pi}{n}, (j+1)\frac{2\pi}{n}]$. By the above,

$$\begin{aligned}&\frac{2\pi}{n}\sum_{d=1}^{n-1} f(d) \\ =\ &\frac{2\pi}{n}\sum_{d=1}^{n-1} f(d) - \int_0^{2\pi} f(t)dt \\ =\ &\frac{2\pi}{n}f(1) - 2\int_0^{2\pi/n} f(t)dt + \frac{2\pi}{n}\sum_{d=2}^{n-1} f(d) - \int_{2\pi/n}^{(n-1)\frac{2\pi}{n}} f(t)dt. \quad \text{(B.1)}\end{aligned}$$

We will now bound the last two terms in the above expression. We assume n

is even. The case when n is odd follows similarly.

$$\begin{aligned}
& \frac{2\pi}{n}\sum_{d=2}^{n-1} f(d) - \int_{2\pi/n}^{(n-1)\frac{2\pi}{n}} f(t)dt \\
= \; & \frac{2\pi}{n}\sum_{d=2}^{n-1} f(d) - \sum_{d=1}^{n-2}\int_{I_d} f(t)dt \\
= \; & \frac{2\pi}{n}\sum_{d=2}^{n/2}[f(d)+f(n+1-d)] - 2\sum_{d=1}^{\frac{n}{2}-1}\int_{I_d} f(t)dt \\
= \; & \frac{2\pi}{n}\sum_{d=2}^{n/2}[f(d)+f(d-1)] - 2\sum_{d=1}^{\frac{n}{2}-1}\int_{I_d} f(t)dt \\
= \; & \frac{2\pi}{n}\sum_{d=1}^{n/2-1}[f(d)+f(d+1)] - 2\sum_{d=1}^{\frac{n}{2}-1}\int_{I_d} f(t)dt \\
= \; & \sum_{d=1}^{n/2-1}\left(\frac{2\pi}{n}[f(d)+f(d+1)] - 2\int_{I_d} f(t)dt\right). \qquad \text{(B.2)}
\end{aligned}$$

We now give a lower bound to prove Item (a) of Lemma 1. Since for $1 \le d \le n/2-1$, $f(t)$ is strict monotone increasing, we know that

$$\frac{2\pi}{n}[f(d)+f(d+1)] - 2\int_{I_d} f(t)dt \ge -\frac{2\pi}{n}[f(d+1)-f(d)].$$

Hence (B.2) is at least[1]

$$\frac{2\pi}{n}\sum_{d=1}^{n/2-1}[f(d)-f(d+1)] \;=\; \frac{2\pi}{n}[f(1)-f(\frac{n}{2})]$$

Hence (B.1) is at least

$$\begin{aligned}
& 2\frac{2\pi}{n}f(1) - 2\int_0^{2\pi/n} f(t)dt - \frac{2\pi}{n}f(\frac{n}{2}) \\
\ge \; & 2\int_{-f(1)}^{\infty} e^{-y}dy - \frac{2\pi}{n}f(\frac{n}{2}) \\
= \; & 4\sin\frac{\pi}{n} - \frac{2\pi}{n}\ln 2.
\end{aligned}$$

[1]Note that applying the Trapezoidal Rule to (B.2) only would bound the magnitude of this term by $O(1)$ bound instead of $o(1)$, as $f''(t) = -1/(2-2\cos t)$ equals approximately $-n^2/4\pi^2$ for $t = 2\pi/n$.

Hence we conclude that

$$\sum_{d=1}^{n-1} f(d) \geq \frac{2n}{\pi} \sin \frac{\pi}{n} - \ln 2.$$

We now prove Item (b) of Lemma 1. Similar to the reasoning that bounded (B.2) from below, we get that (B.2) is at most

$$\frac{2\pi}{n}[f(\frac{n}{2}) - f(1)].$$

Hence (B.1) is at most

$$-2\int_0^{2\pi/n} f(t)dt + \frac{2\pi}{n} f(\frac{n}{2})$$

Provided $n \geq 7$, we have by Lemma B.1 that

$$\begin{aligned}
\int_0^{2\pi/n} f(t)dt &\geq \int_0^{2\pi/n} \ln t - \epsilon(t)dt \\
&\geq [t \ln t - t]_0^{2\pi/n} - [\frac{t^3}{36}]_0^{2\pi/n} \\
&= \frac{2\pi}{n} \ln \frac{2\pi}{n} - \frac{2\pi}{n} - \frac{2\pi^3}{9n^3}.
\end{aligned}$$

Hence (B.1) is at most $\frac{4\pi}{n} \ln \frac{n}{2\pi} + \frac{4\pi}{n} + \frac{4\pi^3}{9n^3} + \frac{2\pi}{n} \ln 2$. Hence

$$\sum_{d=1}^{n-1} f(d) \leq 2 \ln \frac{n}{2\pi} + 2 + \ln 2 + \frac{2\pi^2}{9n^2}.$$

□

C. Proof of Proposition 4

We first prove Item (a). Define the function $g(t) = (r - t\frac{n}{2\pi})f(t)$. Then

$$\begin{aligned}
\sum_{d=1}^{r-1}(r-d)f(d) &\leq \frac{n}{2\pi} \int_{\frac{2\pi}{n}}^{r\frac{2\pi}{n}} g(t)dt \\
&\leq \frac{n}{2\pi} \int_{\frac{2\pi}{n}}^{r\frac{2\pi}{n}} (r - t\frac{n}{2\pi}) \ln t dt \\
&= \frac{rn}{2\pi} \int_{\frac{2\pi}{n}}^{r\frac{2\pi}{n}} \ln t dt - \frac{n^2}{4\pi^2} \int_{\frac{2\pi}{n}}^{r\frac{2\pi}{n}} t \ln t dt \\
&= \frac{rn}{2\pi}[t \ln t - t]_{\frac{2\pi}{n}}^{r\frac{2\pi}{n}} - \frac{n^2}{4\pi^2}[\frac{t^2 \ln t}{2} - \frac{t^2}{4}]_{\frac{2\pi}{n}}^{r\frac{2\pi}{n}} \\
&= \frac{r^2}{2} \ln \frac{2\pi r}{n} + (\frac{1}{2} - r) \ln \frac{2\pi}{n} - \frac{3r^2}{4} + r - \frac{1}{4}
\end{aligned}$$

We now prove Item (b). Let $\alpha = (r-1)\frac{2\pi}{n}$. Then

$$\begin{aligned}
& \sum_{d=1}^{r-1}(r-d)f(d) \\
\geq\ & \frac{n}{2\pi}\int_0^{\alpha} g(t)dt \\
\geq\ & \frac{n}{2\pi}\int_0^{\alpha}(r-t\frac{n}{2\pi})(\ln t - \frac{t^2}{12})dt \quad \{\text{By Lemma B.1}\} \\
=\ & \frac{rn}{2\pi}\int_0^{\alpha}\ln t dt - \frac{n^2}{4\pi^2}\int_0^{\alpha} t\ln t dt - \frac{rn}{24\pi}\int_0^{\alpha} t^2 dt + \frac{n^2}{48\pi^2}\int_0^{\alpha} t^3 dt \\
\geq\ & \frac{rn}{2\pi}[t\ln t - t]_0^{(r-1)\frac{2\pi}{n}} - \frac{n^2}{4\pi^2}[\frac{t^2\ln t}{2} - \frac{t^2}{4}]_0^{(r-1)\frac{2\pi}{n}} - \frac{r^4\pi^2}{9n^2} \\
\geq\ & -\frac{r^2-1}{2}\ln\frac{n}{2\pi(r-1)} - \frac{3r^2}{4} + \frac{r}{2} + \frac{1}{4} - \frac{r^4\pi^2}{9n^2}.
\end{aligned}$$

□

D. Proof of Proposition 5

Suppose the adversary chooses a set of rows R of size l and set of columns T of size r. Let M be the minor of F_n with rows R and columns T. By Theorem 6, for any unit vector f that is 0-concentrated on T, $\hat{f} = F_n f$ is ϵ_R concentrated on R, where $\epsilon_R \geq 1 - \sqrt{\frac{lr}{n}}$. Hence $||M||_2 = \max_{||a||_2=1}||Ma||_2 \leq \sqrt{\frac{lr}{n}}$. Let N be the $l \times (n-r)$ minor of F_n corresponding to rows R and columns not in T. Since $NN^* + MM^* = I$, λ is an eigenvalue of MM^* if-and-only if $(1-\lambda)$ is an eigenvalue of NN^*. The singular values of M are the square roots of the eigenvalues of MM^*. Hence we conclude the smallest singular value of N is at least $\sigma_l^2(N) \geq 1 - \sqrt{\frac{lr}{n}}$, and hence that $\sigma_l^2(\sqrt{n}N) \geq \sqrt{n} - \sqrt{lr}$. Therefore

$$\det(\frac{1}{n}NN^*) \geq (\sqrt{n} - \sqrt{lr})^{2l}.$$

By Theorem 3,

$$\det(\frac{1}{n}NN^*) = \sum_{|S|=l, S\cap T=\emptyset} |\det(\boldsymbol{DFT}_{R,S})|^2.$$

Hence we conclude there exists a minor M_1 with rows R and columns avoiding T that has determinant at least

$$|\det(M_1)| \geq (\sqrt{n} - \sqrt{lr})^l \binom{n-r}{l}^{-1/2}.$$

□