On the Sample Complexity of “Super-Resolution Radar”
Mohammad Mahdi Kamjoo, Saeed Razavi, Sajad Daei

Abstract—In this work, we consider linear time-varying systems whose output signal is described by scaled time-frequency shifted versions of a known input with finite bandwidth $B$. The aim is to identify the time and frequency shifts by observing the received signal over a time interval with duration $T$. This problem is considered in [1] and is shown that the recovery of time-frequency pairs is possible by using an atomic norm minimization provided that the time-frequency shifts are separated by at least $2.37/B$ and $2.37/T$, respectively. For $S$ targets, we theoretically show that the number of required samples for perfect recovery is $O(S)$ up to a logarithmic factor in $BT$ instead of $O(S^2)$ which was obtained in [1].

Index Terms—Off-the-grid compressed sensing, radar system, super resolution, single-input-single-output systems.

I. INTRODUCTION

Super resolution is the problem of obtaining fine-scale information from low resolution data. This problem appears in many practical applications such as optical imaging systems [2], [3], channel estimation [4], image/video processing [5], microscopy [6], localization problems [7] and detection/classification of radar targets [8].

In a typical pulse-Doppler radar system, the relative distances and velocities of the targets are determined by estimating the induced delays $\tau_1, \ldots, \tau_S$ and Doppler shifts $\nu_1, \ldots, \nu_S$ corresponding to the reflections of $S$ targets arising from a transmitted probing signal $x$. Due to physical limitations, the resolution of radar systems is restricted by the bandwidth $B$ of the probing signal and the time interval $T$ over which the responses are observed. As a matter of fact, the estimation accuracy of $\tau$ and $\nu$ is proportional to $1/B$ and $1/T$, respectively. In the super-resolution radar problem, it is supposed to recover the exact timefrequency shifts beyond the natural resolution limitation i.e. $(1/B, 1/T)$ which is achieved by standard pulse-Doppler radar via performing digital matched filtering.

The pulse-Doppler radar problem can be considered as a time-varying linear system. Consequently, traditional spectral estimation techniques can not be directly applied [9]. However, for a single input antenna and with either known, constant delays [10] or known, constant Doppler shifts, the target locations can be recovered with standard spectral estimation techniques such as Pronys method, MUSIC, and ESPRIT [11]. Yet in general, the super-resolution radar problem cannot be reduced to the classical line spectral estimation problem.

The emergence of compressed sensing (CS) methods in the past decade has led to a significant improvement in reducing the required sampling rate for recovery of delay and Doppler pairs compared to the Nyquist sampling rate [12]–[14]. However, CS methods are applicable for the scenarios in which the unknown parameters are assumed to lie on a sufficiently coarse grid. Subsequently, Candès et. al in [15], [16] introduced a grid-less convex formulation for a line spectral estimation problem. Besides, they demonstrated that $S$ continuous frequencies can be recovered from $O(S)$ samples provided a minimum separation condition between the frequencies is satisfied. Inspired by this approach, in [17], the authors developed an atomic norm minimization algorithm and showed that taking $O(S \log(S) \log(N))$ random time-domain samples of an $N$ set index are sufficient for line spectral estimation. The atomic norm minimization is extended to the stationary [18] and non-stationary [19] blind super-resolution problem by utilizing a subspace model for the unknown waveforms and a lifting trick. Moreover, it is shown in [19] that for unknown waveforms with dimension $K$, the blind super-resolution problem is solvable with high probability provided the number of measurements is $O(SK)$ up to a poly-logarithmic factor.

In [1], the authors apply atomic norm minimization to recover delays and Doppler shifts in a radar system. In fact, they adapted super-resolution techniques of [15], [12] to both single-input single-output (SISO) and multi-input multi-output (MIMO) [20] radar system with a random band-limited probing signal. Afterwards, along with this manner, they extended atomic norm minimization to the case of generalized line spectral estimation problem in [21].

In this paper, we focus on the estimation of continuous delay-Doppler pairs of the targets via a tailored atomic norm minimization. The locations and velocities of the targets are determined provided that they are sufficiently separated. Furthermore, we theoretically demonstrate that the delay-Doppler pairs provably can be recovered under a mild condition. In fact, we obtain a upper-bound on the required number of measurements for prefect recovery by using sharped concentration inequalities. More precisely, the delay-Doppler pairs of $S$ targets are recovered from $O(S)$ (up to a poly-logarithmic factor) measurements with high probability. This result reduces the total required number of measurements from quadratic in [1] to linear with a log-factor in $BT$. It is worth noting that the relation between $L$ and $S$ in [1, Theorem 3.1] is wrongly interpreted as linear while it seems quadratic with close scrutiny.

A. Notations

Throughout the paper, scalars are denoted by lowercase letters, and vectors and matrices by lower and upper-case
boldface letters, respectively. The transpose and Hermitian of a matrix $X$ are represented by $X^T$ and $X^H$, respectively. Following the conventional notations, we denote $j$th entry of vector $x$ by $|x|_j$ or $x_j$. Also, we define $|X|$ as the spectral norm of the matrix $X$. Also $|x|_1$, $|x|_2$, and $|x|_\infty$ are defined as the element-wise $\ell_1$, $\ell_2$, and $\ell_\infty$ norms of $x$, respectively. $e_j$ represents the $j$-th canonical basis vector. For an integer $N$, $[N]$ stands for $\{1, 2, \ldots, N\}$. Identity matrix of size $N \times N$ is represented by $I_N$. We adopt notations from the previous work \cite{1} in which they used two-dimensional index for vectors and matrices, e.g., we write $|x|_{(k, \ell)}$, $k, \ell = -N, \ldots, N$ for $x = [x(-N,-N), x(-N,-N+1), \ldots, x(-N,N), x(-N+1,-N), \ldots, x(N,N)]$. Also, $\text{sign}(b) := e^{i2\pi b}$ is sign of a complex number $b$ with polar decomposition $|b|e^{i\theta}$. The $(m, n)$-th partial derivative of a function of two variables $G(x, y)$ is denoted by $G^{(m,n)}(x, y) := \frac{\partial^m G(x, y)}{\partial x^m \partial y^n}$. Finally, we reserved the calligrapher alphabet $S$ for denoting the $\ell$-th canonical basis vector. For an integer $\ell$, we denote $S_\ell$ as the $\ell$-th canonical basis vector. By sampling (1), the samples $x_{\ell} := x(p/B)$ are given by (see \cite{1} Appendix A) for more details):

$$y(t) = \sum_{j=1}^{S} b_j x(t - \tau_j) e^{i2\pi \nu_j t},$$

where $b_j$ denotes attenuation factors corresponding to the delayDoppler pairs $(\tau_j, \nu_j), j = 1, \ldots, S$. According to the 2WT-Theorem \cite{22, 23}, at least $L = B^T$ samples are required for identifying all delayDoppler pairs. By sampling (1) at a rate of $\frac{1}{B^T}$, then applying the discrete Fourier transform (DFT) and the inverse DFT (IDFT) to (1), the samples $y_p := y(p/B)$ are given by (see \cite{1} Appendix A) for more details):

$$y_p = \sum_{k=-N}^{N} \sum_{\ell=-N}^{N} [a(r_j)]_{(k, \ell)} x_p e^{i2\pi (kp/L)};$$

$$p = -N, \ldots, N,$$

where $N = (L - 1)/2$, and

$$[a(r_j)]_{(k, \ell)} := \frac{1}{L^2} \sum_{m, n=-N}^{N} e^{i2\pi (m(\frac{\ell}{L}) - \tau_j + n(\frac{k}{L}) - \nu_j)}.$$  (3)

Here, $r_j$ denotes the $j$-th time-frequency shift pair i.e. $r_j = [\tau_j, \nu_j]_T$ where $\tau_j = \frac{\tau_j}{T}$ and $\nu_j = \frac{\nu_j}{B}$. The aim is to recover all the $S$ parameters $(\tau_j, \nu_j)$ from the $p$ linear measurements $\{y_p\}_{p=-N}^{N}$. We can rewrite these measurements in matrix notation as below

$$y = Gz \in \mathbb{C}^L,$$  (4)

where $z \in \mathbb{C}^{L^2}$ is a sparse linear combination of time and frequency shifted versions of the atom set $A := \{e^{i2\pi \phi}(r), r \in [0, 1]^2, \phi \in [0, 1]\}$ i.e.

$$z = \sum_{j} b_j a(r_j).$$  (5)

and $G \in \mathbb{C}^{L \times L^2}$ is the Gabor matrix defined as

$$[G]_{p, (k, \ell)} := x_p e^{-i2\pi (kp/L)} \quad \forall \ k, \ell, p = -N, \ldots, N. \quad (6)$$

Thus, the estimation of the triplets $(\tau_j, \nu_j)$ amounts to recovering the $S$-sparse vector $z \in \mathbb{C}^{L^2}$ from the measurement vector $y \in \mathbb{C}^{L}$. By utilizing the fact that $z$ is a sparse linear combination of atoms from the set $A$, the atomic norm of such signals is defined as \cite{17}:

$$\|z\|_A := \inf_{b_j \in \mathbb{C}, r_j \in [0, 1]^2} \left\{ \sum_{j} |b_j| z = \sum_{j} b_j a(r_j) \right\}.$$  

To estimate $z$, the following basis pursuit type atomic norm minimization is used

$$\min_{\tilde{z} \in \mathbb{C}^{L^2}} \|z\|_A \quad \text{s.t.} \quad y_{L \times 1} = G\tilde{z}. \quad (7)$$

This problem is a convex optimization and can be solved by reformulating its dual as an semi-definite programming \cite{1}. In what follows, we investigate and discuss the minimum $L$ required for prefect recovery via (7) which is the main emphasis of this paper.

### III. MAIN RESULTS

The main result of this paper is that if we assume the probing signal has normal distribution under a mild separation condition for the delays and velocities with the targets, Problem (7) provably recovers all $S$ triples $(\tau_j, \nu_j) \in [S]$, with high probability as long as $L$ is large enough. More precisely, the difficulty of recovery depends on the wrap-around distance separations between the spikes in $S$ which is formally defined as

$$\Delta(S) := \inf_{r_j, r_j' \in \mathbb{R} \neq 0} \min_{n} \|r_j - r_j' + n\|_\infty.$$  (8)

**Theorem 1.** Assume that the samples of the probing signal \{xt\}_{t=-N} are i.i.d. \mathcal{N}(0, \Sigma_L) random variables and $L = 2N + 1 \geq 1024$. Let $y = Gz$ with $G \in \mathbb{C}^{L \times L^2}$ defined in (6) and $z = \sum_{j=1}^{S} b_j a(r_j)$. Also, assume that the signs of the coefficients $b_j \in \mathbb{C}$ are chosen independently from symmetric distributions on the complex unit circle, and the vectors $r_j = (\tau_j, \nu_j)$ satisfying the minimum separation $\Delta(S) \geq \frac{2\pi}{S}$. Then as long as

$$L \geq cS \log^3(L/\delta),$$  (9)

where $c$ is a numerical constant, with probability at least $1-\delta$, $z$ is the unique minimizer of (7).
Remark 1. Based on (9), the spikes are perfectly recovered when the number of measurements $L$ is linearly greater than the number of spikes up to a logarithmic factor i.e. $L = O(S \log^3(L))$. The previous result was obtained in [1 Theorem 3.1], where the required number of samples $L = O(S^2 \log^3(L))$ is quadratically related to $S$. Noteworthy that the relation between $L$ and $S$ is wrongly interpreted as linear in [1 Theorem 3.1]. The improvement of our bound is due to using a strong version of the matrix Bernstein inequality in the way of the proof (see Lemma A.7) instead of the Hanson-Wright inequality [24] used in [1].

Theorem [1] suggests that under a separation condition and $L$ on the order of $O(S)$ up to logarithmic factors, we provably recover the spike signal with high probability.

IV. PROOF SKETCH

To enhance the readability, we provide a proof sketch of Theorem [1]. First in Proposition [1] we describe the desired form of a valid dual polynomial and its properties that will guarantee the uniqueness of the solution to (7). We next design the dual polynomial by using the squared Fejer kernel [15]. The reminder of the proof is then to cautiously evaluate the dual polynomial in order to fulfill all the required properties.

Proposition 1. If there exists a trigonometric polynomial $Q(r)$. $r = [r, \nu]$, of the form

$$Q(r) = \langle q, G \alpha(r) \rangle$$

(10)

with complex coefficients $q \in \mathbb{C}^L$ such that

$$Q(r_j) = \text{sign}(b_j), \text{ for all } r_j \in S,$$

(11a)

$$|Q(r)| < 1 \text{ for all } r \in [0, 1]^2 \setminus S,$$

(11b)

where $S = \{r_1, \ldots, r_S\}$, and sign(·) is the complex sign function. Then, the solution to (7) is unique.

The proof of Proposition [1] is rather standard (see e.g., [1] Proof of Proposition 1.3), we do not repeat technical details here again. The proof of Theorem [1] is concluded by constructing an appropriate dual certificate satisfying the conditions of Proposition [1] see Appendix A. Based on Proposition [1] we only need to find a dual polynomial $Q(r)$ with the valid form of (10) that satisfies (11a) and (11b). The main strategy of constructing a dual certificate $Q(r)$ is inspired by the method introduced in [15], [17]. In fact, from [15] Proposition C.1, we know that a deterministic 2D trigonometric polynomial such as

$$Q(r) = \langle q, f(r) \rangle, \quad [f(r)]_{k, \nu} = e^{-i2\pi(k\nu + r)}$$

for $k, \nu = -N, \ldots, N$ and with deterministic coefficients $q \in \mathbb{C}^L$ satisfies the interpolation (11a) and boundedness (11b) conditions under a minimum separation constraint defined in [15, Definition 7.9]. To build the certificate, we follow a similar program. Precisely, the atom $a(r)$ can be expressed as inverse Fourier transform of $f(r)$ i.e.

$$a(r) = F^Hf(r),$$

(12)

where $F \in \mathbb{C}^{L^2 \times L^2}$ is discrete 2D Fourier transform matrix whose entry of the $(k, \nu)$-th row and $(k', \nu')$-th column equals to $(1/L^2)e^{-i2\pi((k'k + \nu')/L)}$. Moreover, based on (10), $Q(r)$ can be considered as a 2D random trigonometric polynomial in the variables $r$ and $\nu$ with random coefficient vector $FGHq$ where the randomness comes from the random nature of the probing signal. Therefore, it remains to show that $Q(r)$ tends to $Q(r)$ with high probability, which in turn implies that $Q(r)$ satisfies both interpolation and boundedness conditions with high probability. To do so, we first prove that $Q(r)$ concentrates around $\bar{Q}(r)$ on a finite support of $r$. Unlike Hanson Wright inequality used in [1], we use sharp matrix Bernstein inequality which culminates in fewer number of required sample for recovery. In other words, invoking this inequality instead of Hanson Wright inequality, reducing order of sample complexity from quadratic term $O(S^2)$ to linear term $O(S)$. Then, it is extend to be hold uniformly for all points with high probability, using Bernstein’s polynomial inequality. The technical details are provided in the Appendix A.

V. CONCLUSION

In this paper, we studied a single transmit and single receive antenna radar system which was modeled as a linear time-varying system. We focused on the sample complexity of this problem. In [1], it is theoretically shown that the required number of samples at the receiver is related quadratically to the number of delay-Doppler pairs. However, in this paper, we showed that the relation to the number of delay-Doppler pairs is linear rather than quadratic. Our approach was to use non-commutative matrix Bernstein inequality instead of Hanson-Wright inequality [24] used in [1].

APPENDIX

A. Construction of the dual polynomial

In this section, we describe how to construct the polynomial $Q(r)$ of the form (10) to satisfy conditions in Proposition 1. We first take a brief detour to take into account a basic technique for the construction of a deterministic dual polynomial named $\bar{Q}(r)$. In order to construct $\bar{Q}(r)$, Candès et al. in [25] used a fast-decaying kernel $\bar{G}$ and its partial derivatives according to

$$\bar{Q}(r) = \sum_{j=1}^{S} \bar{\alpha}_j \bar{G}(r - r_j) + \bar{\beta}_{j1} \bar{G}^{(1,0)}(r - r_j) + \bar{\beta}_{j2} \bar{G}^{(0,1)}(r - r_j),$$

(13)

where $\bar{G}(r) := F(\tau)F(\nu)$, and $F(t)$ is the squared Fejer kernel defined by $F(t) := \left(\frac{\sin(M \pi t)}{M \sin(\pi t)}\right)^4$. It follows that $F(t)$ is a trigonometric polynomial of degree $N$ i.e.

$$F(t) = \frac{1}{M} \sum_{k=-N}^{N} g_k e^{i2\pi tk}.$$ 

(14)

Here, $M := N/2 + 1$ for even $N$. In [15], it is suggested to obtain coefficients $\bar{\alpha}_j, \bar{\beta}_{j1},$ and $\bar{\beta}_{j2}$ such that

$$\bar{Q}(r_j) = u_j \quad \text{and} \quad \nabla \bar{Q}(r_j) = 0 \quad \forall \ r_j \in S,$$ 

(15)
where \( u_j = \text{sign}(b_j) \), \( S = \{r_1, r_2, \ldots, r_S\} \), and \( \nabla(\cdot) \) represents derivation respect to both \( \tau \) and \( \nu \) which leads to
\[
\nabla \tilde{Q}(r_j) = 0 \rightarrow \tilde{Q}(1,0)(r_j) = 0, \quad \tilde{Q}(0,1)(r_j) = 0.
\]
To alleviate notation, let us rewrite this equation i.e. (15) in matrix form
\[
\begin{bmatrix}
\tilde{D}^{(0,0)}(1,0) & \frac{1}{\kappa} \tilde{D}^{(0,1)}(1,0) & \frac{1}{\kappa^2} \tilde{D}^{(0,2)}(1,0)
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\
\kappa \beta_1 \\
\kappa \beta_2
\end{bmatrix}
= \begin{bmatrix} u \end{bmatrix},
\]
(16)
where \( [\tilde{D}^{(m,n)}]_{j,j'} := \tilde{G}^{(m,n)}(r_{j'} - r_j), [\alpha]_j := \tilde{\alpha}_j, [\beta]_j := \beta_{j'}, r_j := [r_j, u_j]^T, \) where \( g_k \) is the coefficient corresponding to the squared Fejer kernel in (14).

With this notation, we define
\[
G_{(m,n)}(r_j, r_j) := \frac{L^2}{M^2} (G G^H g_{(m,n)}(r_j), G G^H f(r_j))
\]
where \( G \) is the Gabor matrix. Recall that \( G^H \) is the inverse two-dimensional discrete Fourier Transform matrix with the entry in the \( (k, \ell) \)th row and \( (k', \ell') \)th column given by \( [G^H]_{(k, \ell), (k', \ell')} := (1/L^2) e^{-i2\pi (k + \ell') / L} \) and \( [f(r_j)]_{(k, \ell)} = e^{-i2\pi (k + \ell') / L} \) with \( k, \ell, k', \ell' = -N, \ldots, N \). Then, the polynomial \( Q(r) \) is constructed by interpolating the points \( \{r_j, u_j\} \) with functions \( G_{(m,n)}(r_j, r_j), m, n = 0, 1 \) as:
\[
Q(r) = \sum_{j=1}^{S} \alpha_j G(r, r_j) + \beta_{1j} G(1,0)(r, r_j) + \beta_{2j} G(0,1)(r, r_j)
\]
(19)
This ensures that \( Q(r) \) in (19) is of the form (10). Hence, the proof is outlined as follows.

- **Step 1:** For the interpolation condition (11a) and boundedness condition in (11b), we will show that, for a randomly chosen \( x \), there exists a specific choice of coefficients \( \alpha_j, \beta_{1j}, \beta_{2j} \) such that
\[
Q(r_j) = u_j \quad \text{and} \quad \nabla Q(r_j) = 0 \quad \forall \ r_j \in \mathcal{S},
\]
(20)
with probability at least \( 1 - \delta/48 \). This is necessary to guarantee that \( Q(r) \) reaches local maxima at the \( r_j \) which satisfies both conditions in (20), simultaneously.

- **Step 1a:** In Lemma [A.1] we show with probability at least \( 1 - \delta/48 \) that \( ||\tilde{D} - \bar{D}|| \leq 1/4 \) by using non-commutative Bernstein inequality of Lemma [A.10].

- **Step 1b:** From [1, Proposition 8.2], we know that \( \bar{D} \) is invertible and \( ||\bar{D} - I_{3S}|| \leq 0.1908 \). With triangle inequality and the result of Step 1a, we prove that \( \tilde{D} \) is invertible with probability at least \( 1 - \delta/48 \). Therefore, condition (20) for \( \forall r_j \in \mathcal{S} \) is satisfied.

- **Step 2:** By using an \( \epsilon \)-net argument, we show that with the coefficients chosen as in Step 1, \( |Q(r)| < 1 \), with probability at least \( 1 - \delta \), uniformly for all \( r \not\in \mathcal{S} \). This completes the proof.

- **Step 2a:** For every \( r \in \Omega \) which \( \Omega \in [0,1]^2 \) is a (finite) set of grid points, we prove that \( Q(r) \) is close to \( \tilde{Q}(r) \) with probability at least \( 1 - \delta \).

- **Step 2b:** By using Bernstein's polynomial inequality used in [1, Lemma 8.9], we extend this result to hold with high probability uniformly for all \( r \in [0,1]^2 \).

- **Step 2c:** Finally, in the Lemma [A.7] we apply [1, Lemma 8.12] to conclude \( |Q(r)| < 1 \) holds with probability at least \( 1 - \delta \) uniformly for all \( r \not\in \mathcal{S} \). Indeed, this Lemma is inspired by a result of (15) that shows \( |Q(r)| < 1 \) for all \( r \not\in \mathcal{S} \).

**B. Selecting the coefficients of \( Q(r) \)**

In this subsection, we demonstrate that there exists a set of the coefficients \( \alpha_j, \beta_{1j}, \beta_{2j} \) such that \( Q(r) \) satisfies (20). First, the corresponding system of (20) can be recast in matrix form as follows:
\[
\begin{bmatrix}
D^{(0,0)}(0,0) & \frac{1}{\kappa} D^{(0,1)}(0,0) & \frac{1}{\kappa^2} D^{(0,2)}(0,0)
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\
\kappa \beta_1 \\
\kappa \beta_2
\end{bmatrix}
= \begin{bmatrix} u \end{bmatrix},
\]
(21)
where \( D^{(m,n)}(r_{j'}, r_{j'}) := G^{(m,n)}(r_{j'}, r_{j'}), [\alpha]_j := \alpha_j, [\beta]_j := \beta_{j'}, r_j := [r_j, u_j]^T \).
Solving this system of equations yields the coefficients. To show that the system of equations (20) has a solution, and prove existence of the coefficients, we require matrix \( D \) to be invertible. To this end, from [1, Proposition 8.2], we know that \( \bar{D} \) is invertible, and we have the following result on that
\[
||\bar{D} - I_{3S}|| \leq 0.1908.
\]
Therefore, we just need to show that w.h.p the following event occurs i.e.
\[
\Upsilon_\xi = \{ ||\tilde{D} - \bar{D}|| \leq \xi \}.
\]
(22)
So for all \( \xi \in (0, 1/4) \) by using triangle inequality, we can arrive to our desire i.e.
\[
||I_{3S} - \bar{D}|| \leq ||\tilde{D} - \bar{D}|| + ||\bar{D} - I_{3S}||
\]
\[
\leq \xi + 0.1908 \leq 0.4408.
\]
(23)
Hence $D$ is invertible on $\mathbb{Y}_\xi$, and the coefficients $\alpha_j, \beta_{1j}, \beta_{2j}$ can be obtained as follows

$$
\begin{bmatrix}
\alpha \\
\kappa \beta_1 \\
\kappa \beta_2
\end{bmatrix} = D^{-1} \begin{bmatrix}
u \\
0
\end{bmatrix} = L u,
$$
(24)

where $L$ is the $3S \times S$ submatrix of $D^{-1}$ corresponding to its first $S$ columns. It only reminds to prove that the event $\mathbb{Y}_\xi$ can be appointed with high probability, which stated in the following Lemma.

**Lemma A.1.** For all $\xi > 0$

$$
P[\mathbb{Y}_\xi] \geq 1 - \frac{\delta}{48},
$$
(25)

provided that

$$
L \geq \frac{c_1 S \log(288S\delta)}{\xi^2},
$$
(26)

where $c_1$ is a numerical constant.

**Proof.** See Appendix [E] for technical details. \qed

Also, on the event $\mathbb{Y}_\xi$, we state two useful inequalities which will be used in the next steps, about $L$ and its deviation from $\mathbb{Y}_\xi$ in the lemma below.

**Lemma A.2.** [Lemma 8.4]. On the event $\mathbb{Y}_\xi$ with $\xi \in (0, 1/4]$, the following identities hold:

$$
\|L\| \leq 2.5,
$$
(27)

$$
\|L - L\| \leq 2.5\xi.
$$
(28)

**C. $Q(r)$ and $\bar{Q}(r)$ are close on a grid**

The goal of this step is to show that $Q(r)$ and $\bar{Q}(r)$ and their partial derivatives are close on a set of (grid) points as stated in the following Lemma.

**Lemma A.3.** Let $\Omega \subset \{0, 1\}^2$ be a finite set of points. Fix $0 \leq \epsilon \leq 1$ and $\delta > 0$. Suppose that

$$
L \geq \frac{S}{\epsilon^2} \max \left( c_2 \log^2 \left( \frac{576S|\Omega|}{\delta}\right) \log \left( \frac{384|\Omega|}{\delta}\right),
\right.
$$

$$
\left. c_4 \log \left( \frac{192|\Omega|}{\delta}\right) \log \left( \frac{288S}{\delta}\right) \right).
$$
(29)

Then, for $r \in \Omega$, we have

$$
\frac{1}{\kappa^{n+m}} |Q^{(m,n)}(r) - \bar{Q}^{(m,n)}(r)| \leq \frac{\epsilon}{6},
$$
(30)

with probability at least $1 - \frac{\delta}{12}$.

**Proof.** The $(m, n)$-th partial derivative of $Q(r)$ after normalization with $\frac{1}{\kappa^{n+m}}$ reads to

$$
\frac{1}{\kappa^{n+m}} Q^{(m,n)}(r) = \frac{1}{\kappa^{n+m}} \sum_{k=1}^{S} \left( \alpha k G^{(m,n)}(0,0)(r, r_j) + \beta_{1j} k^{-1} G^{(m,n)}(1,0)(r, r_j) + \beta_{2j} k^{-1} G^{(m,n)}(0,1)(r, r_j) \right)
$$
$$
= (u, L^H v^{(m,n)}(r)).
$$
(31)

Here, the last equality comes from (24) and the shorthand $v^{(m,n)}(r)$ is defined by

$$
v^{(m,n)}(r) := \frac{1}{\kappa^{n+m}} G^{(m,n)}(0,0)(r, r_1)
$$

$$
G^{(m,n)}(1,0)(r, r_1)
$$

$$
G^{(m,n)}(0,1)(r, r_1),
$$
(32)

with $\mathbb{E}[v^{(m,n)}(r)] = \bar{v}^{(m,n)}(r)$. With simple manipulations, the derivatives of $Q(r)$ are divided into three terms as bellow

$$
\frac{1}{\kappa^{n+m}} Q^{(m,n)}(r) = \frac{1}{\kappa^{n+m}} \bar{Q}^{(m,n)}(r) + I_1^{(m,n)}(r) + I_2^{(m,n)}(r),
$$
(33)

where

$$
I_1^{(m,n)}(r) = \langle u, L^H (v^{(m,n)}(r) - \bar{v}^{(m,n)}(r)) \rangle,
$$

$$
I_2^{(m,n)}(r) = \langle u, (L - \bar{L})^H v^{(m,n)}(r) \rangle,
$$

and,

$$
\frac{1}{\kappa^{n+m}} \bar{Q}^{(m,n)}(r) = \mathbb{E}[Q^{(m,n)}(r)] = \langle u, L^H \bar{v}^{(m,n)}(r) \rangle.
$$
(34)

Now, in order to prove the lemma, we need to show that the perturbations $I_1^{(m,n)}(r)$ and $I_2^{(m,n)}(r)$ are small on a set of finite points $\Omega$ with high probability. From (33), we obtain, for all $\xi > 0$ satisfying (39),

$$
P[\max_{r \in \Omega} |I_1^{(m,n)}(r) + I_2^{(m,n)}(r)| \geq \frac{\epsilon}{6}] \\ \leq P[\max_{r \in \Omega} I_1^{(m,n)}(r) \geq \frac{\epsilon}{12}] + P[\max_{r \in \Omega} I_2^{(m,n)}(r) \geq \frac{\epsilon}{12}],
$$
(35)

where the last inequality follows from the union bound. Then, by using the inequality $P[A] \leq P[B] + P[A|B]$ for events $B = \mathbb{Y}_\xi$ and $A = \{\max_{r \in \Omega} I_2^{(m,n)}(r) \geq \frac{\epsilon}{12}\}$, we have

$$
P[\max_{r \in \Omega} I_1^{(m,n)}(r) \geq \frac{\epsilon}{12}] \\
\leq P[\max_{r \in \Omega} I_1^{(m,n)}(r) \geq \frac{\epsilon}{12}] \\
+ P[\max_{r \in \Omega} I_2^{(m,n)}(r) \geq \frac{\epsilon}{12}] + P[\mathbb{Y}_\xi].
$$
(36)

In order to bound these two terms, we use two useful lemmas from (1).

**Lemma A.4 (Lemma 8.8).** Let $\Omega \subset [0, 1]^2$ be a finite set of points and suppose that $m + n \leq 2$. Then, for all $0 < \epsilon \leq 1$ and for all $\delta > 0$,

$$
P[\max_{r \in \Omega} I_2^{(m,n)}(r) \geq \frac{\epsilon}{12}] \leq \frac{\delta}{48} + P[\mathbb{Y}_{1/4}],
$$
(37)
provided that
\[ L \geq \frac{S}{e^2c_2} \log \left( \frac{576S\Omega}{\delta} \right) \log \left( \frac{384\Omega}{\delta} \right). \] (38)

**Lemma A.5 (Lemma 8.9).** Let \( \Omega \subset [0, 1]^2 \) be a finite set of points and suppose that \( m + n \leq 2 \). Then, for all \( 0 < \epsilon \leq 1 \) and \( \delta > 0 \), and for all \( \xi > 0 \) with
\[ \xi \leq \frac{cc_3}{\sqrt{\log(192\Omega/\delta)}} \leq 1/4, \] (39)
where \( c_3 \leq 1/4 \) is a numerical constant, it follows
\[ \mathbb{P} \left[ \max_{r \in \Omega} |I_{2}^{(m,n)}(r)| \geq \frac{\epsilon}{12} \right] \leq \frac{\delta}{48}. \] (40)

With using Lemmas A.4 and A.5, then setting \( \xi \) to be \( \epsilon c_3 \log^{-1/2}(192\Omega/\delta) \), which leads to
\[ L \geq S(c_4/\epsilon^2) \log(192\Omega/\delta) \log(288S/\delta), \] (41)
by choosing \( c_4 = c_1/c_2^2 \) in Lemma A.1 and it yields \( \mathbb{P} \left[ \bar{Y}_{\xi} \right] \leq \frac{\delta}{48} \). Thus, we observe that
\[ \mathbb{P} \left[ \max_{r \in \Omega} \frac{|Q^{(m,n)}(r) - \bar{Q}^{(m,n)}(r)|}{\kappa^{n+m}} \geq \frac{\epsilon}{6} \right] \leq \frac{\delta}{12}. \] (42)

This concludes the proof.

\[ \square \]

**D. \( Q(r) \) and \( \bar{Q}(r) \) are close for all \( r \)**

We next use an \( \epsilon \)-net argument together with Lemma A.3 to establish that \( Q^{(m,n)}(r) \) is close to \( \bar{Q}^{(m,n)}(r) \) with high probability uniformly for all \( r \in [0,1]^2 \). To this end, we barrow two useful lemmas from [1] to avoid repeating the technical details here.

**Lemma A.6 (Lemma 8.11).** Let \( \epsilon, \delta > 0 \). If
\[ L \geq S \frac{c_6}{\epsilon^2} \log^3 \left( \frac{c_5 L^6}{\delta \epsilon^2} \right), \] (43)
then, with probability at least \( 1 - \delta \),
\[ \max_{r \in [0,1]^2, (m,n):m+n \leq 2} \frac{1}{\kappa^{n+m}} |Q^{(m,n)}(r) - \bar{Q}^{(m,n)}(r)| \leq \epsilon. \] (44)

Finally, we used a result in [1] to satisfy the bound Lemma 8.11]ness condition for all \( r \). Since that result is valid, we do not repeat technical details here.

**Lemma A.7 (Lemma 8.12).** Suppose that
\[ L \geq S c_6 \log^3 \left( \frac{c_5 L^6}{\delta} \right). \] (45)

Then with probability at least \( 1 - \delta \), the following statements hold:
1. For all \( r \) that satisfy \( \min_{r_j \in \mathcal{S}} |r - r_j| \geq 0.2447/N \), we have \( |Q(r)| < 0.9963 \).
2. For all \( r \notin \mathcal{S} \) that satisfy \( 0 < |r - r_j| \geq 0.2447/N \), for some \( r_j \in \mathcal{S} \), we have \( |Q(r)| < 1 \).

**E. Proof Of Lemma A.7**

Let \( M := N/2 + 1 \), then for matrix \( D \), one can write
\[ D = \frac{1}{M^2} \sum_{k,\ell = -N}^{N} g_k g_\ell \bar{w}_{k,\ell} w_{k,\ell}^H, \] (46)
where \( \bar{w}_{k,\ell} = \mathbb{E}[w_{k,\ell}] \), and
\[ \mathbb{E}[D] = \frac{1}{M^2} \sum_{k,\ell = -N}^{N} g_k g_\ell \bar{w}_{k,\ell} w_{k,\ell}^H. \] (47)

where \( p_{k,\ell} \in \mathbb{C}^{L^2} \) is the \( (k, \ell) \)-th row of matrix \( L^2 \mathbf{F} \mathbf{G}^H \mathbf{G}^H \mathbf{F} \) and \( e_{k,\ell} \in \mathbb{R}^{L^2} \) is canonical basis vector whose \( (k, \ell) \)-th element is one and the other ones are zero. Also, by using the fact that \( \mathbb{E}[p_{k,\ell}] = e_{k,\ell} \), it is convenient to verify \( \mathbb{E}[D] = \frac{1}{M} \sum_{k,\ell = -N}^{N} g_k g_\ell \bar{w}_{k,\ell} w_{k,\ell}^H. \) Then, we have
\[ D - \bar{D} = \frac{1}{M^2} \sum_{k,\ell = -N}^{N} g_k g_\ell \bar{w}_{k,\ell} (w_{k,\ell}^H - \bar{w}_{k,\ell}^H) = \sum_{k,\ell = -N}^{N} X_{k,\ell}, \] (48)

where \( X_{k,\ell} = \frac{1}{M^2} g_k g_\ell \bar{w}_{k,\ell} (w_{k,\ell}^H - \bar{w}_{k,\ell}^H) \). Now, in order to use the Bernstein’s inequality Lemma A.10, we require to obtain the upper bound of \( \|X_{k,\ell}\| \) and its variance \( \sigma^2 = \mathbb{E} \left[ \sum_{k,\ell = -N}^{N} X_{k,\ell} X_{k,\ell}^H \right] \). Therefore, we have
\[ \|X_{k,\ell}\| = \left\| \frac{1}{M^2} \bar{w}_{k,\ell} (w_{k,\ell}^H - \bar{w}_{k,\ell}^H) \right\| \leq \frac{2}{M^2} \|\bar{w}_{k,\ell}\| \|w_{k,\ell}\|, \] (49)

where inequality (49) comes from the fact that for any random vector \( \mathbf{v} \in \mathcal{V} \), we have \( \|\mathbf{v} - \mathbb{E}[\mathbf{v}]\| \leq 2 \sup_{\mathbf{v} \in \mathcal{V}} \|\mathbf{v}\| \|\bar{\mathbf{v}}\| \). By using the fact that \( \|\bar{w}_{k,\ell}\| \leq \|w_{k,\ell}\| \max_{j \in [S]} \|p_{k,\ell}, f(r_j)\| \), one can attain to
\[ \|X_{k,\ell}\| \leq \frac{2}{M^2} \|\bar{w}_{k,\ell}\| \max_{j \in [S]} \|p_{k,\ell}, f(r_j)\| \leq \frac{2}{M^2} 27S \max_{j \in [S]} \|p_{k,\ell}, f(r_j)\| \] (50)

\[ = \frac{54S}{M^2} \max_{j \in [S]} \|p_{k,\ell}, f(r_j)\|. \] (51)
In (50), we have used the obtained upper bound in \[ \text{Lemma 4.4} \] where \( \| \tilde{w}_{k,\ell} \|_2^2 \leq 27S \). Then, using Lemma \[ \text{A.9} \] for (51) gives us \( \| X_{k,\ell} \| \leq \frac{108SL}{M} \). Next, for the variance, we have

\[
\left\| \mathbb{E} \left[ \sum_{k,\ell=-N}^{N} X_{k,\ell}X_{k,\ell}^H \right] \right\| 
\leq \frac{1}{M^2} \left\| \sum_{k,\ell=-N}^{N} g_k^2g_{\ell}^2 \tilde{w}_{k,\ell}^2 \right\|.
\]

(52)

Now, we require to bound \( \mathbb{E} \left[ \| w_{k,\ell} \|^2 \right] \). To do so, we have

\[
\mathbb{E} \left[ \| w_{k,\ell} \|^2 \right] \leq \mathbb{E} \left[ \| w_{k,\ell} \|^2 \left\| (p_{k,\ell}f(r_j))^2 \right\| \right] 
\leq 27SF^2f^2(r_j)f(p_{k,\ell})f(r_j) 
= 27SL^2f^2(r_j)f^2(r_j) 
\leq 28SL.
\]

(54)

Substituting (53) into (52), gives

\[
\left\| \mathbb{E} \left[ \sum_{k,\ell=-N}^{N} X_{k,\ell}X_{k,\ell}^H \right] \right\| 
\leq \frac{56SL}{M^2} \left\| \sum_{k,\ell=-N}^{N} g_k^2g_{\ell}^2 \tilde{w}_{k,\ell}^2 \right\| 
\leq \frac{56SL}{M^2} \max_{k,\ell} |g_kg_{\ell}| \|D\|.
\]

(55)

where the last inequality comes from \[ \text{Lemma 4.3} \]. Now, by using \[ \text{I} \] Proposition 8.2, one can arrive to

\[
\left\| \mathbb{E} \left[ \sum_{k,\ell=-N}^{N} X_{k,\ell}X_{k,\ell}^H \right] \right\| \leq \frac{1}{M^2} 68SL.
\]

(56)

By following similar steps, we have that

\[
\left\| \mathbb{E} \left[ \sum_{k,\ell=-N}^{N} X_{k,\ell}X_{k,\ell}^H \right] \right\| \leq \frac{1}{M^2} 68SL.
\]

Therefore, with invoking the Lemma \[ \text{A.10} \] inequality, we write

\[
\mathbb{P} \left( \| D - \overline{D} \| \geq \xi \right) \leq 6S \exp \left( - \frac{\xi^2}{48S^2 + 68SL} \right) 
\leq \frac{\delta}{48},
\]

(57)

that is satisfied if \( M^2 \geq \frac{8S^2}{\xi^2} \log \left( \frac{48S}{\delta} \right) \) or equivalently

\[
L \geq \frac{8S^2}{\xi^2} \log \left( \frac{48S}{\delta} \right).
\]

(58)

**Lemma A.8.** Let \( p_{k,\ell} \in \mathbb{C}^{L^2} \) be the \( (k, \ell) \)-th row of matrix \( L^2 A \), then we have

\[
\mathbb{E} \left[ p_{k,\ell}p_{k,\ell}^H \right] = \frac{1}{L} I_L + \frac{1 + L}{L} e_{k,\ell}e_{k,\ell}^H
\]

for all \( k, \ell = -N, \ldots, N \).

**Proof.** See Appendix \[ \text{C} \].

**Lemma A.9.** If \( L \geq 4 \log (4/\delta_1) \) and \( 1 > \delta_1 > 0 \) be a positive constant, we have

\[
\mathbb{P} \left( \| p_{k,\ell}f(r_j) \| \geq 2L \right) \leq \delta_1,
\]

for all \( k, \ell = -N, \ldots, N \) and \( r_j \in S \).

**Proof.** See Appendix \[ \text{C} \] for the proof.

**F. Proof Of Lemma A.8**

Each element of matrix \( p_{k,\ell}p_{k,\ell}^H \in \mathbb{C}^{L^2 \times L^2} \) equals to

\[
[p_{k,\ell}p_{k,\ell}^H]_{(k_1,k_2),(\ell_1,\ell_2)} = \varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2},
\]

(61)

where \( \varphi_{k_1,k_2}, \varphi_{k_1,k_2} \in \mathbb{C}^L \), and \( \varphi_{k_1,k_2} \) are random vectors with distribution \( N(0, I_L/2) \). Now, for its expectation, we first consider non-diagonal elements that are independent which leads to zero expectation i.e.,

\[
\mathbb{E}[\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}] = \mathbb{E}[\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}] = 0.
\]

Moreover, for diagonal elements in the case \( (k_1,k_1) = (\ell_1,\ell_2) \neq (k,\ell) \), we have

\[
\mathbb{E}[\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}] = \text{Tr} \left( \mathbb{E}[\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}] \right)
\leq \frac{1}{L} \text{Tr} \left( I_L I_L \right)
= \frac{L}{L^2} = \frac{1}{L}.
\]

(62)

Finally, in the case of \( (k_1,k_1) = (\ell_1,\ell_2) = (k,\ell) \), the expectation equals 4-th cumulant of Gaussian vector which is

\[
\mathbb{E}[\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}^{H}\varphi_{k_1,k_2}] = \mathbb{E}[\| \varphi_{k_1,k_2} \|_2^4] = \frac{2}{L} + 1.
\]

(63)

Therefore, we can conclude that

\[
\mathbb{E}[p_{k,\ell}p_{k,\ell}^H]_{(k_1,k_2),(\ell_1,\ell_2)} = \begin{cases} 
1/L & \text{if } (k_1, k_2) = (\ell_1, \ell_2) \\
2/L + 1 & \text{if } (k_1, k_2) = (k, \ell) \\
0 & \text{otherwise.}
\end{cases}
\]

(64)

**G. Proof Of Lemma A.9**

Let \( p_{k,\ell} \in \mathbb{C}^{L^2} \) be \( \langle \varphi_{k,\ell}, \varphi_{k',\ell'} \rangle \) for \( k',\ell' = -N, \ldots, N \), where \( \varphi_{k',\ell'} \) is the \( (k',\ell') \)-th column of matrix \( GF^H \), then we have that

\[
\langle p_{k,\ell}, f(r_j) \rangle = \sum_{k',\ell'=-N}^{N} [f(r_j)]_{k',\ell'} \langle \varphi_{k,\ell}, \varphi_{k',\ell'} \rangle
\]

(65)

\[
= \langle \varphi_{k,\ell} \sum_{k',\ell'=-N}^{N} [f(r_j)]_{k',\ell'} \varphi_{k',\ell'} \rangle + \| f(r_j) \|_2 \| \varphi_{k,\ell} \|_2^2
\]

(66)

where \( \varphi \) is a random vector of dimension \( L \) with distribution \( N(0, I_L/(L^2-1)) \). Moreover, we know that \( \varphi_{k,\ell} \) has \( N(0, I_L/L) \). With triangular inequality, it gives us

\[
\| p_{k,\ell}f(r_j) \| \leq \| \langle \varphi_{k,\ell}, \varphi \rangle \| + \| \varphi_{k,\ell} \|_2^2.
\]

(67)
We used the fact that $|\mathbf{f}(r_j)|_{k,\ell} = 1$. Afterwards, we obtained upper bound of each term in (67), separately. At first step, by using Chernoff inequality [26 Proposition 12], the inner product in (67) can be bounded as below

$$
P(|\varphi_{k,\ell}, \varphi| \geq L) \leq \min_{t > 0} \mathbb{E}[e^{\lambda (\varphi_{k,\ell}, \varphi) t} e^{-Lt}].$$

(68)

From [27 Equation 28], we know that moment of two independent normal variable $x$ and $y$ with distribution $N(0, \sigma_x^2)$ and $N(0, \sigma_y^2)$, respectively, is equal to $\frac{1}{\sqrt{1-\sigma_x^2 \sigma_y^2}}$. Hence, we have that

$$
P(|\varphi_{k,\ell}, \varphi| \geq L) \leq \mathbb{E}[e^{\lambda (\varphi_{k,\ell}, \varphi) t} e^{-Lt}]
= \mathbb{E}\left[\prod_{i} e^{\lambda (\varphi_{i,\ell}, \varphi_{i}) t} e^{-Lt}\right]
= \mathbb{E}\left[\prod_{i} e^{\lambda (\varphi_{i,\ell}, \varphi_{i}) t} e^{-Lt}\right]
= (1 - \lambda^2 t^2)^{-L/2} e^{-Lt}
= \frac{L}{2} (\log(1 - \lambda^2 t^2) + 2t),$$

(69)

where $\lambda$ is the positive constant that equals to $\sqrt{\frac{L^2}{\sigma_x^2 \sigma_y^2}}$. Also, equality (69) comes from independency of each element of inner product $\langle \varphi_{k,\ell}, \varphi \rangle$. In order to minimize the upper bound in (70), we derive it with respect to $t$, and it could be upper bounded with $\exp\left(-\frac{L}{2} \left(\log \left(\frac{\sqrt{5} - 1}{2}\right) + \sqrt{5} - 1\right)\right)$ which gives us to

$$
P(|\varphi_{k,\ell}, \varphi| \geq L) \leq e^{-\frac{3L}{8}},$$

(71)

where $-\log\left(\frac{\sqrt{5} - 1}{2}\right) - \sqrt{5} + 1 > \frac{3}{4}$. Thus, for $L \geq \frac{3}{4} \log(4/\delta_1)$, the probability is less than $\delta_1/2$, i.e.

$$
P(|\varphi_{k,\ell}, \varphi| \geq L) \leq P(|\varphi_{k,\ell}, \varphi| \geq L) + P(|\varphi_{k,\ell}, \varphi| \leq -L)
\leq 2e^{-\frac{3L}{8}} \leq \delta_1/2.$$  

(72)

Utilizing a standard concentration inequality for the norm of a Gaussian random vector, e.g. [28 Equation 1.6], we obtain upper bound of the second term in (66), for $L \geq 4 \log(2/\delta_1)$ as

$$
P\left[\|\varphi_{k,\ell}\|_2^2 \geq 3\right] \leq P\left[\|\varphi_{k,\ell}\|_2^2 \geq 2\left(1 + \frac{2 \log(2/\delta_1)}{L}\right)\right]
\leq P\left[\|\varphi_{k,\ell}\|_2^2 \geq \left(1 + \frac{2 \log(2/\delta_1)}{\sqrt{L}}\right)\right]
\leq e^{-2 \log(2/\delta_1)/2} \leq \delta_1/2.$$  

(73)

(74)

Here, we used $\sqrt{2(1 + \gamma^2)} \geq (1 + \gamma)$ in (75). Therefore, in order to show that (66) hold with high probability at least $1 - \delta_1$, it is required that $L \geq \max\{\frac{3}{4} \log(4/\delta_1), 4 \log(2/\delta_1)\}$, e.g.

$$
|\langle p_{k,\ell}, f(r_j) \rangle| \leq |\langle \varphi_{k,\ell}, \varphi \rangle| + \|\varphi_{k,\ell}\|_2^2
\leq L + 3 \leq 2L.$$  

(75)

This concludes the proof.

$H$. Matrix Bernstein inequality

Lemma A.10 ([29 Theorem 1.4]). Let $X_i$ be a finite sequence of independent zero-mean self-adjoint random matrices of dimension $d \times d$ such that $\|X_i\| \leq B$ almost surely for a certain constant $B$. For all $t \geq 0$ and a positive constant $\sigma^2$

$$
P\left[\|\sum_i X_i\| \geq t\right] \leq 2e \exp\left(-\frac{t^2/2}{\sigma^2 + Bt/3}\right),$$

where $\sigma^2 := \max\{\|\mathbb{E}[\sum_i X_i X_i^H]\|, \|\mathbb{E}[\sum_i X_i^H X_i]\|\}$.

REFERENCES

[1] R. Heckel, V. I. Morgenshtern, and M. Soltanolkotabi, “Super-resolution radar,” Information and Inference: A Journal of the IMA, vol. 5, no. 1, pp. 22–75, 2016.

[2] S. Razavikia, A. Amini, and S. Daei, “Reconstruction of binary shapes from blurred images via Hankel-structured low-rank matrix recovery,” IEEE Transactions on Image Processing, vol. 29, pp. 2452–2462, 2020.

[3] S. Razavikia, H. Zamani, and A. Amini, “Sampling and recovery of binary shapes via low-rank structures,” in 2019 13th International conference on Sampling Theory and Applications (SampTA), 2019, pp. 1–4.

[4] C. Shin, R. W. Heath, and E. J. Powers, “Blind channel estimation for mimo-ofdm systems,” IEEE Transactions on Vehicular Technology, vol. 56, no. 2, pp. 670–685, 2007.

[5] M. Aghzani, A. Esmaeili, K. Behdin, and F. Marvasti, “Missing low-rank and sparse decomposition based on smoothed nuclear norm,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 6, pp. 1550–1558, 2020.

[6] A. Alberti, C. Robens, W. Alt, S. Brakhane, M. Karski, R. Reimann, A. Widera, and D. Meschede, “Super-resolution microscopy of single atoms in optical lattices,” New Journal of Physics, vol. 18, no. 5, p. 053010, May 2016.

[7] J. Helland, M. B. Wakin, and G. Tang, “A super-resolution algorithm for extended target localization,” in 2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2019, pp. 386–390.

[8] F. Xi, S. Chen, and Z. Liu, “Super-resolution delay-doppler estimation for sub-nuquist radar via atomic norm minimization,” in 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2017, pp. 4326–4330.

[9] R. Heckel, “7 super-resolution radar imaging via convex optimization,” Compressed Sensing in Radar Signal Processing, p. 193, 2019.

[10] T. Strohmer and H. Wang, Adventures in Compressive Sensing Based MIMO Radar. Springer International Publishing, address:Cham, 2015, vol. 3, pp. 285–326.

[11] K. D. Rao and M. Swamy, Spectral Analysis of Signals. Singapore: Singapore Individual Publishing, 2018, pp. 721–751.

[12] A. Herman and T. Strohmer, “High-resolution radar via compressive sensing,” IEEE Transactions on Signal Processing, vol. 57, no. 6, pp. 2275–2284, 2009.

[13] F. Baraniuk and P. Steeghs, “Compressive radar imaging,” in 2007 IEEE Radar Conference, 2007, pp. 128–133.

[14] R. Heckel and H. Bolcskei, “Identification of sparse linear operators,” IEEE Transactions on Information Theory, vol. 59, no. 12, pp. 7985–8000, 2013.

[15] E. J. Candés and C. Fernandez-Granda, “Towards a mathematical theory of super-resolution,” Communications on pure and applied Mathematics, vol. 67, no. 6, pp. 906–956, 2014.

[16] ———, “Super-resolution from noisy data,” Journal of Fourier Analysis and Applications, vol. 19, no. 6, pp. 1229–1254, 2013.

[17] G. Tang, B. N. Bluskar, P. Shah, and B. Recht, “Compressed sensing off the grid,” IEEE transactions on information theory, vol. 59, no. 11, pp. 7465–7490, 2013.

[18] Y. Chi, “Guaranteed blind sparse peaks deconvolution via lifting and convex optimization,” IEEE Journal of Selected Topics in Signal Processing, vol. 10, no. 4, pp. 782–794, 2016.

[19] D. Yang, G. Tang, and M. B. Wakin, “Super-resolution of complex exponentials from modulations with unknown waveforms,” IEEE Transactions on Information Theory, vol. 62, no. 10, pp. 5809–5830, 2016.

[20] R. Heckel, “Super-resolution mimo radar,” in 2016 IEEE International Symposium on Information Theory (ISIT), 2016, pp. 1416–1420.
[21] M. Hong and Z.-Q. Luo, “On the linear convergence of the alternating direction method of multipliers,” Mathematical Programming, vol. 162, no. 1-2, pp. 165–199, 2017.

[22] G. Durisi, V. I. Morgenshtern, and H. Bolcskei, “On the sensitivity of continuous-time noncoherent fading channel capacity,” IEEE Transactions on Information Theory, vol. 58, no. 10, pp. 6372–6391, 2012.

[23] D. Slepian, “On bandwidth,” Proceedings of the IEEE, vol. 64, no. 3, pp. 292–300, 1976.

[24] M. Rudelson, R. Vershynin et al., “Hanson-wright inequality and sub-gaussian concentration,” Electronic Communications in Probability, vol. 18, 2013.

[25] E. J. Candès and B. Recht, “Exact matrix completion via convex optimization,” Foundations of Computational mathematics, vol. 9, no. 6, p. 717, 2009.

[26] Ž. Stojanac, D. Suess, and M. Kliesch, “On products of gaussian random variables,” arXiv preprint arXiv:1711.10516, 2017.

[27] A. Seijas-Macias and A. Oliveira, “An approach to distribution of the product of two normal variables,” Discussiones Mathematicae Probability and Statistics, vol. 32, no. 1-2, pp. 87–99, 2012.

[28] L. Michel and T. Michel, Probability in Banach Spaces: Isoperimetry and Processes. Springer, 1991.

[29] J. A. Tropp, “User-friendly tail bounds for sums of random matrices,” Foundations of computational mathematics, vol. 12, no. 4, pp. 389–434, 2012.