Compressive Harmonic Retrieval via Matrix Completion

Yuxin Chen
Stanford University

Yuejie Chi
Ohio State University

Abstract—We study the problem of recovering a mixture of $r$ complex sinusoids at continuous-valued frequencies from $[0, 1)^2$, when a random subset of its $n_1 \times n_2$ regularly spaced time-domain samples are observed. Addressing this problem is important in many applications such as imaging system, control, radar, and can help overcome the basis mismatch issue in compressed sensing. We develop a new algorithm based on an enhanced Hankel structure and nuclear norm minimization. Under certain incoherence condition, this method admits accurate recovery from the order of $r \log^2(n_1, n_2)$. Additionally, our results extend to the more general problem of low-rank Hankel matrix completion.

I. PROBLEM FORMULATION

A large class of signals of interest entails a superposition of spikes in the continuous frequency domain, which arises in numerous applications including radar, imaging systems, wireless networks, etc. The resolution of signal acquisition devices is often limited by physics and hardwares, precluding sampling with desired resolution. It is of great interest to identify the underlying multi-dimensional frequencies from $\mathbb{R}^m$ to $\mathbb{C}^m$ regular samples as introduced in [1]

Recent addressed in [1].

It is of great interest to identify the underlying frequencies from $\mathbb{R}^m$ to $\mathbb{C}^m$ regular samples as introduced in [1].

A large class of signals of interest entails a superposition of spikes in the continuous frequency domain, which arises in numerous applications including radar, imaging systems, wireless networks, etc. The resolution of signal acquisition devices is often limited by physics and hardwares, precluding sampling with desired resolution. It is of great interest to identify the underlying multi-dimensional frequencies from $\mathbb{R}^m$ to $\mathbb{C}^m$ regular samples as introduced in [1].

Recent addressed in [1].

For concreteness, we restrict our attention to 2-D frequency models. Consider an $n_1 \times n_2$ data matrix $X = (x_{i,j})_{0 \leq i < n_1, 0 \leq j < n_2}$ that can be expressed as $x_{i,j} = \sum_{l=1}^{r} d_l y_{i,l} z_{i,l}$, where $y_{i,l} := \exp(j2\pi f_{i,l})$ and $z_{i,l} := \exp(j2\pi f_{j,l})$ for some set of frequency pairs $\{(f_{i,l}, f_{j,l}) \mid 1 \leq i \leq r\}$. Suppose that there exists a location set $\Omega$ of size $m$ such that $x_{k,l}$ is observed iff $(k,l) \in \Omega$. We are interested in perfectly recovering $X$ from a small set of observations. Once the data matrix $X$ is recovered, the underlying frequencies can be retrieved using conventional methods such as MEMP [2].

We first convert $X$ to an enhanced form $X_e$ as introduced in [2]

$$X_e := \begin{bmatrix} X_0 & X_1 & \cdots & X_{n_1-1,k} \\ X_1 & X_2 & \cdots & X_{n_1-2,k+1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{k-1} & X_{k} & \cdots & X_{n_1-1} \end{bmatrix}^{t}$$

(1)

where each block $X_t$ is a Hankel matrix defined as

$$X_t := \begin{bmatrix} x_{t,0} & x_{t,1} & \cdots & x_{t,n_2-k-2} \\ x_{t,1} & x_{t,2} & \cdots & x_{t,n_2-k+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{t,k-2} & x_{t,k-1} & \cdots & x_{t,n_2-1} \end{bmatrix}^t$$

This enhanced form admits a low-rank decomposition

$$X_e = E_L D E_R,$$

for some matrices $E_L$ and $E_R$, where $D := \text{diag}(d_l)$. This reveals that rank($X_e$) $\leq r$. We also let the SVD of the enhanced form $X_e$ be $X_e = UV^*$.

The enhanced matrix completion (EMaC) algorithm can be presented as

$$\text{minimize } \|M_e\|_* \quad \text{subject to } P_\Omega(M) = P_\Omega(X)$$

(2)

where $M_e$ denotes the enhanced form of $M$.

II. PERFORMANCE GUARANTEE

We define a measure of incoherence as follows. Let $G_L$ and $G_R$ be two $r \times r$ matrices such that

$$(G_L)_{k,l} := \frac{1}{k_1 k_2} \begin{bmatrix} 1 & -(y^*_l y^*_k) & 1 & -(z^*_l z^*_k) \\ 1 & -y^*_l y^*_k & 1 & -z^*_l z^*_k \end{bmatrix},$$

$$(G_R)_{k,l} := \frac{1}{(n_1 - k_1 + 1) (n_2 - k_2 + 1)} \begin{bmatrix} 1 & -(y^*_l y^*_k) & 1 & -(z^*_l z^*_k) \\ 1 & -y^*_l y^*_k & 1 & -z^*_l z^*_k \end{bmatrix},$$

with the convention that $(G_L)_{k,l} = (G_R)_{k,l} = 1$. We let $\Omega_e(i,l)$ be the set of locations in the enhanced matrix $X_e$ containing copies of $x_{i,l}$, and denote $k(i,l) = \Omega_e(i,l)$. Besides, we define $A_{i,l}$ such that $A_{i,l} \equiv 1/\sqrt{\Omega_e(i,l)}$ if $(\alpha, \beta) \in \Omega_e(i,l)$ and 0 otherwise. The incoherence can then be defined as follows.

**Definition 1 (Incoherence).** $X$ is said to have incoherence $(\mu_1, \mu_2, \mu_3)$ if they are respectively the smallest values obeying

$$\sigma_{\min}(G_L) \geq \frac{1}{\mu_1}, \quad \sigma_{\min}(G_R) \geq \frac{1}{\mu_2};$$

$$\max_{(i, l) \in [n_1] \times [n_2]} \frac{1}{\text{dist}(i, l)} \left| (UV^*, A_{i,l}) \right|^2 \leq \frac{\mu_3 r}{n_1^2 n_2};$$

$$\forall b \in [n_1] \times [n_2]: \sum_{a} \left| (UU^* A_b VV^*, \sqrt{\mu_a} A_a) \right|^2 \leq \frac{\mu_3 n_1 n_2}{n_1 n_2}.$$  

(3)

**Theorem 1.** Define $c_l := \max_n \left( \min(n_{1-l}, n_{1-k+l}) \right)$.

Then there exist constants $c_0, c_1 > 0$ such that under either of the following conditions:

1) Condition (3), (4), and (5) hold and $m > c_0 \max \{n_1 n_2, 1/\mu_1, \mu_2, \mu_3 r \log^2(n_1 n_2);$

2) Condition (3) holds and $m > c_0 \mu_1 c_2 r^2 \log^2(n_1 n_2);$  

$X$ is the unique solution of EMaC with high probability.

III. NUMERICAL RESULTS

Set $n := n_1 = n_2 = 11$. We run 200 Monte Carlo trials for each triple $(n, m, r)$ when the frequencies are randomly generated. A trial is declared successful if the relative error $\|X^* - X\|_F/\|X\|_F \leq 10^{-5}$. The phase transition diagram is plotted in Fig. 1, which illustrates the practicability of EMaC.

Figure 1. Phase transition of EMaC when $m = n_1 = n_2 = 11$.

REFERENCES
