Blind Demixing and Deconvolution at Near-Optimal Rate

Dominik Stöger

Technical Unversity of Munich,
Department of Mathematics

dominik.stoeger@ma.tum.de

Joint work with Peter Jung (TU Berlin), Felix Krahmer (TU München)

supported by DFG priority program "Compressed Sensing in Information Processing" (CoSIP)

Overview

General Framework

Recovery and guarantees

Proof sketch

Overview

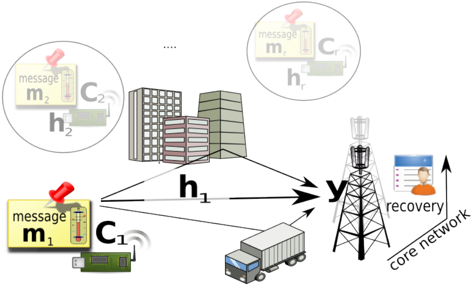
General Framework

Recovery and guarantees

Proof sketch

A problem in Wireless Communication

- r different devices
- each device wants to deliver a message $m_i \in \mathbb{C}^N$
- **Channel model:**
  Only few active paths $w_i = Bh_i$, where $B \in \mathbb{C}^{L \times K}$
- **Linear encoding:**
  $x_i = C_i m_i$ with $C_i \in \mathbb{C}^{L \times N}$
  Device i transmits x_i
- **Received signal:**



$$y = \sum_{i=1}^{r} w_i * x_i \in \mathbb{C}^L$$

Goal: recover **all** m_i from y

Assumptions on B_i and C_i

- Assume w_i is concentrated on the first few entries (most direct paths)
- B: First K columns of the $L \times L$ identity $\Rightarrow$ extends h_i by zeros
- (More general models for B are possible.)
- Choice of C_i arbitrary $\Rightarrow$ randomize
- Choose C_i to have i.i.d. standard complex normal entries, i.e., $(C_i)_{jk} \sim \mathcal{CN}(0,1)$

Lifting

- $w_i * x_i = Bh_i * C_i m_i$ bilinear in h_i and m_i
- $\Rightarrow$ There is a unique linear map $\mathcal{A}_i : \mathbb{C}^{K\times N} \to \mathbb{C}^L$ such that $Bh_i * C_i m_i = \mathcal{A}_i \left(h_i m_i^* \right)$ for arbitrary h_i and m_i
- $$y = \sum_{i=1}^{r} \mathcal{A}_i \left(h_i m_i^* \right) = \mathcal{A}\left(X^0 \right),$$
  where
  $$X^0 = \left(h_1 m_1^*, \ldots, h_r m_r^* \right)$$
- **Low rank matrix recovery problem**
- Corresponding combinatorial problem NP-hard in general
  $\rightarrow$ convex relaxation

Overview

A convex approach for recovery

[Ling, Strohmer 2017]

- Solve the semi-definite program

$$\text{minimize} \sum_{i=1}^{r} \|Y_i\|_* \qquad \text{subject to} \sum_{i=1}^{r} \mathcal{A}_i(Y_i) = y. \qquad \text{(SDP)}$$

- $\|\cdot\|_*$: nuclear norm, i.e., the sum of the singular values
- Recovery is guaranteed with high probability, if

$$L \geq C r^2 \left(K + \mu_h^2 N\right) \log^3 L \log r$$

- μ_h coherence parameter, ranges between $1 \leq \mu_h^2 \leq K$
- (Near-)optimal dependence on K, N, suboptimal dependence on r.
- Previously established for $r = 1$ in [Ahmed, Recht, Romberg 2015]

Main result

Theorem (Jung, Krahmer, S., 2017)

Let $\omega \geq 1$. Assume that

$$L \geq C_\omega r \left(K \log K + N \mu_h^2 \right) \log^3 L, \tag{1}$$

where C_ω is a universal constant only depending on ω. Then with probability $1 - \mathcal{O}\left(L^{-\omega}\right)$ the recovery program (SDP) is successful, i.e., X^0 is its unique minimizer.

- (Near) optimal dependence on K, N, and r

Overview

Proof overview

Our proof follows the same strategy as [Ling, Strohmer 2016].
It consists of the following two main steps:

- Establishing sufficient conditions for recovery
  $\Rightarrow$ approximate dual certificate
- Constructing the dual certificate ("Golfing Scheme")

Proof overview

Our proof follows the same strategy as [Ling, Strohmer 2016].
It consists of the following two main steps:

- Establishing sufficient conditions for recovery
  $\Rightarrow$ approximate dual certificate
- Constructing the dual certificate ("Golfing Scheme")

The following subspace is important for both steps of the proof:

$$\mathcal{T} = \Big\{ (u_1 m_1^* + h_1 v_1^*, \ldots, u_r m_r^* + h_r v_r^*) : \\ u_1, \ldots, u_r \in \mathbb{C}^K, v_1, \ldots, v_r \in \mathbb{C}^N \Big\}$$

$\mathcal{T}_i$ is defined by

$$\mathcal{T}_i = \left\{ u m_i^* + h_i v^* : \; u \in \mathbb{C}^K, \; v \in \mathbb{C}^N \right\}.$$

Local Isometry Property

- Crucial ingredient for the proof:

Definition

We say that $\mathcal{A}$ fulfills the δ-local isometry property, if

$$(1-\delta)\sum_{i=1}^{r}\|X_i\|_F^2 \le \left\| \sum_{i=1}^{r} \mathcal{A}_i\left(X_i\right) \right\|_{\ell_2}^2 \le (1+\delta)\sum_{i=1}^{r}\|X_i\|_F^2$$

for all $X = (X_1, \ldots, X_r) \in \mathcal{T}$.

Local Isometry Property

- Crucial ingredient for the proof:

Definition

We say that $\mathcal{A}$ fulfills the δ-local isometry property, if

$$(1-\delta)\sum_{i=1}^{r}\|X_i\|_F^2 \leq \left\| \sum_{i=1}^{r} \mathcal{A}_i\left(X_i\right) \right\|_{\ell_2}^2 \leq (1+\delta)\sum_{i=1}^{r}\|X_i\|_F^2$$

for all $X = (X_1, \dots, X_r) \in \mathcal{T}$.

- Our goal: Show that $\mathcal{A}$ fulfills the local isometry property, if L scales linearly with r

Local isometry property

- Define $\hat{\mathcal{T}} = \left\{X = (X_1, \dots, X_r) \in \mathcal{T} : \sum_{i=1}^r \|X_i\|_F^2 = 1\right\}$
- δ-local isometry property is equivalent to

$$\begin{aligned}
\delta &\geq \sup_{X \in \hat{\mathcal{T}}} \left| \|\sum_{i=1}^r \mathcal{A}_i(X_i)\|_{\ell_2}^2 - \sum_{i=1}^r \|X_i\|_F^2 \right| \\
&= \sup_{X \in \hat{\mathcal{T}}} \left| \|\sum_{i=1}^r \mathcal{A}_i(X_i)\|_{\ell_2}^2 - \mathbb{E}\left[\|\sum_{i=1}^r \mathcal{A}_i(X_i)\|_{\ell_2}^2 \right] \right| \\
&= \sup_{X \in \hat{\mathcal{T}}} \left| \|V_X \operatorname{vec}([C_1, \dots, C_r])\|_{\ell_2}^2 - \mathbb{E}\left[\|V_X \operatorname{vec}([C_1, \dots, C_r])\|_{\ell_2}^2 \right] \right|,
\end{aligned}$$

where for $X = (u_1 m_1^* + h_1 v_1^*, \dots, u_r m_r^* + h_r v_r^*) \in \mathcal{T}$

$$V_X\left(\operatorname{vec}([C_1, \dots, C_r])\right) = \sum_{i=1}^r \mathcal{A}_i(X_i)$$

Suprema of Chaos Processes

Theorem (Krahmer, Mendelson, Rauhut 2014)

Let $\mathcal{X}$ be a symmetric set of matrices, i.e., $\mathcal{X} = -\mathcal{X}$, and let ξ be a random vector whose entries ξ_i are independent and have distribution $\mathcal{CN}(0,1)$. Then, for $t > 0$,

$$P\left(\sup_{A\in\mathcal{X}}\left|\|A\xi\|_{\ell_2}^2 - \mathbb{E}\|A\xi\|_{\ell_2}^2\right| \geq c_1 E + t\right) \leq 2\exp\left(-c_2 \min\left(\frac{t^2}{V^2}, \frac{t}{U}\right)\right)$$

where, setting $\mathcal{D}(\mathcal{X}) = \int_0^{+\infty}\sqrt{\log \mathcal{N}(\mathcal{X}, \|\cdot\|_{2\to2}, t)}dt$ the quantities E, V, and U are defined as

$$\begin{aligned}
E &= \mathcal{D}(\mathcal{X})\left(\mathcal{D}(\mathcal{X}) + d_F(\mathcal{X})\right)\\
V &= d_{2\to2}(\mathcal{X})\left(\mathcal{D}(\mathcal{X}) + d_F(\mathcal{X})\right)\\
U &= d_{2\to2}^2(\mathcal{X}).
\end{aligned}$$

The next steps

- Apply Theorem for $\mathcal{X} = \{V_X : X \in T;\ \sum_{i=1}^{r} \|X_i\|_F^2 = 1\}$.
- Bound covering numbers

$\Longrightarrow$ δ-local isometry property holds with high probability