Non-Convex Recovery from Phaseless Low-Resolution Blind Deconvolution Measurements using Noisy Masked Patterns

Samuel Pinilla†, Kumar Vijay Mishra‡, and Brian Sadler‡ †Faculty of Information Technology and Communication Sciences, Tampere University
‡United States CCDC Army Research Laboratory, Adelphi, MD 20783 USA

Here, we provide some mathematical proofs that support the theoretical analysis of the paper.

Appendix A: Proof of Theorem 1

Due to homogeneity in (10), it suffices to work with the case where $\|\mathbf{x}\| = 1$. Instrumental in proving Theorem 2 is the following result. Additionally, in order to model the effect of the low-pass filter $\mathcal{G}$ in Algorithm 1 consider the diagonal matrix $\mathbf{L}_{\mathcal{G}} \in \mathbb{R}_{+}^{n\times n}$ whose entries satisfy $0 < (\mathbf{L}_{\mathcal{G}})_{i,i} < 1$. This inequality constraint is assumed without loss of generality to model the contribution of the frequencies in the Fourier domain.

Lemma 1. Consider the noisy data $|\mathbf{a}_i^H\mathbf{x}|^2 + (\boldsymbol{\omega})_i = |\mathbf{a}_i^H\mathbf{F}^H\boldsymbol{\alpha}|^2 + (\boldsymbol{\omega})_i$, with $\boldsymbol{\alpha} = \mathbf{F}\mathbf{x}$ such that $\|\boldsymbol{\omega}\|_\infty \leq c\|\mathbf{x}\|_\infty$ for some $c > 0$. For any unit vector $\boldsymbol{\alpha} \in \mathbb{C}^n$, there exists a vector $\mathbf{u} \in \mathbb{C}^n$ with $\mathbf{u}^H\boldsymbol{\alpha} = 0$ and $\|\mathbf{u}\| = 1$, such that

$$\frac{1}{2}\|\boldsymbol{\alpha}\boldsymbol{\alpha}^H - \hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}^H\|_F^2 \leq \frac{\|\mathbf{S}\mathbf{L}_{\mathcal{G}}\mathbf{u}\|_2^2}{\|\mathbf{S}\mathbf{L}_{\mathcal{G}}\boldsymbol{\alpha}\|_2^2} + \mathcal{O}(\|\boldsymbol{\omega}\|_\infty)^1, \tag{19}$$

where $\mathbf{S} = \left[\frac{\mathbf{b}_1}{\|\mathbf{b}_1\|_2}, \cdots, \frac{\mathbf{b}_J}{\|\mathbf{b}_J\|_2}\right]^H$ for $i \in \mathcal{I}_0$, with $J = card(\mathcal{I}_0)$ is the cardinality of $\mathcal{I}_0$. Recall that the vectors $\mathbf{b}_i$ are the rows of the matrix $\mathbf{B} = \mathbf{A}\mathbf{F}$, and $\hat{\boldsymbol{\theta}}$ is the solution of (9).

Proof. Notice that

$$\begin{aligned}\frac{1}{2}\|\boldsymbol{\alpha}\boldsymbol{\alpha}^H - \hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}^H\|_F^2 &= \frac{1}{2}\|\boldsymbol{\alpha}\|_2^4 + \frac{1}{2}\|\hat{\boldsymbol{\theta}}\|_2^4 - |\boldsymbol{\alpha}^H\hat{\boldsymbol{\theta}}|^2 \\ &= 1 - |\boldsymbol{\alpha}^H\hat{\boldsymbol{\theta}}|^2 = 1 - \cos^2(\beta) \\ &= \sin^2(\beta),\end{aligned} \tag{20}$$

where $\beta \in [0, \pi/2]$ is the angle between the spaces spanned by $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\alpha}$. Then one can write

$$\boldsymbol{\alpha} = \cos(\beta)\hat{\boldsymbol{\theta}} + \sin(\beta)\hat{\boldsymbol{\theta}}^{\perp}, \tag{21}$$

where $\hat{\boldsymbol{\theta}}^{\perp} \in \mathbb{C}^n$ is a unit vector orthogonal to $\hat{\boldsymbol{\theta}}$ and the real part of its inner product with $\boldsymbol{\alpha}$ is non-negative. Then from (21) we have that

$$\boldsymbol{\alpha}^{\perp} = -\sin(\beta)\hat{\boldsymbol{\theta}} + \cos(\beta)\hat{\boldsymbol{\theta}}^{\perp}, \tag{22}$$

in which $\boldsymbol{\alpha}^{\perp} \in \mathbb{C}^n$ is a unit vector orthogonal to $\boldsymbol{\alpha}$. Thus, considering (21), (22) and appealing to Lemma 1 in [1], we obtain that

$$\frac{1}{2}\|\boldsymbol{\alpha}\boldsymbol{\alpha}^H - \hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}^H\|_F^2 \leq \frac{\|\mathbf{S}\mathbf{L}_{\mathcal{G}}\boldsymbol{\alpha}^{\perp}\|_2^2}{\|\mathbf{S}\mathbf{L}_{\mathcal{G}}\boldsymbol{\alpha}\|_2^2} + \mathcal{O}(\|\boldsymbol{\omega}\|_\infty). \tag{23}$$

Then, taking $\mathbf{u} = \boldsymbol{\alpha}^{\perp}$ the result holds. □

We now turn to prove Theorem 2. The first step consists in upper-bounding the term on the right hand-side of (23). Specifically, its numerator term will be upper bounded, and the denominator term lower bounded, which are summarized in the following lemmas.

Lemma 2. Assume that $\sum_{\ell=1}^{L}\mathbf{D}_\ell^H\mathbf{D}_\ell = r\mathbf{I}$ for some $r > 0$ with $L \geq c_0 n$ for some sufficiently large constant $c_0 > 0$. In the setup of Lemma 1, if $card(\mathcal{I}_0) \geq C_I n$, then the next

$$\|\mathbf{S}\mathbf{L}_{\mathcal{G}}\mathbf{u}\|_2^2 \leq \gamma_1 r(1 + \delta - \zeta)card(\mathcal{I}_0) \tag{24}$$

holds for $\delta, \zeta, \gamma_1 \in (0, 1)$ with probability at least $1 - 2e^{-Cn}$ provided that L is sufficiently large.

Proof. In order to proof this lemma observe that

$$\mathbf{E} = \sum_{i=1}^{nL}\frac{\mathbf{b}_i\mathbf{b}_i^H}{\|\mathbf{b}_i\|_2^2} = \sum_{\ell=1}^{L}\frac{\mathbf{D}_\ell^H\mathbf{D}_\ell}{\|\mathbf{D}_\ell\|_F^2} \approx r\mathbf{I},$$

since the Fourier transform matrix $\mathbf{F}$ is an orthogonal matrix i.e $\mathbf{F}^H\mathbf{F} = \mathbf{F}\mathbf{F}^H = \mathbf{I}$. Then, from standard concentration inequality on the sum of random positive semi-definite matrices with sub-Gaussian rows [2], it can be obtained that

$$(1-\delta) \leq \sigma_{min}\left(\frac{1}{r}\mathbf{E}\right) \leq \sigma_{max}\left(\frac{1}{r}\mathbf{E}\right) \leq (1+\delta), \tag{25}$$

with probability at least $1 - 2e^{-Cn}$ as long as L is sufficiently large, for some constant $\delta \in (0, 1)$ and $C > 0$, where $\sigma_{max}(\cdot)$ and $\sigma_{min}(\cdot)$ denote the largest and smallest singular value, respectively. Given the fact that $\mathbf{S}$ is a sub-matrix of $\mathbf{E}$, from (25) we get

$$\begin{aligned}\sigma_{max}\left(\frac{1}{rcard(\mathcal{I}_0)}\mathbf{S}\mathbf{L}_{\mathcal{G}}\right) &\leq \left(\sigma_{max}\left(\frac{1}{rcard(\mathcal{I}_0)}\mathbf{E}\right) - \zeta\right)\gamma_1 \\ &\leq (1 + \delta - \zeta)\gamma_1,\end{aligned} \tag{26}$$

[1]The notation $\varphi(w) = \mathcal{O}(g(w))$ means there exists a numerical constant $c > 0$ such that $\varphi(w) \leq cg(w)$.

for some constants $\zeta, \delta \in (0,1)$, where $\gamma_1 = \sigma_{max}(\mathbf{L}_{\mathcal{G}}) < 1$. Thus, from (26) we have that

$$\|\mathbf{Su}\|_2^2 = \left|\mathbf{u}^H\mathbf{S}^H\mathbf{Su}\right| \leq \gamma_1 r(1+\delta-\zeta)card(\mathcal{I}_0), \quad (27)$$

holds with probability at least $1-2e^{-Cn}$, provided that L is sufficiently large. Considering the fact that $\boldsymbol{\alpha}^{\perp} = \mathbf{u}$, then we have that

$$\|\mathbf{SL}_{\mathcal{G}}\boldsymbol{\alpha}^{\perp}\|_2^2 \leq \gamma_1 r(1+\delta-\zeta)card(\mathcal{I}_0), \quad (28)$$

with high probability. Thus, the result holds. □

Lemma 3. In the setup of Lemma 1, the following holds with probability at least $1-2e^{-Cn}$

$$\|\mathbf{SL}_{\mathcal{G}}\boldsymbol{\alpha}\|_2^2 \geq \gamma_2 r(1-\delta)card(\mathcal{I}_0), \quad (29)$$

with $\delta, \gamma_2 \in (0,1)$ provided that L is sufficiently large.

Proof. Notice that the left side term in (29) can be seen as

$$\|\mathbf{SL}_{\mathcal{G}}\boldsymbol{\alpha}\|_2^2 = \sum_{i\in\mathcal{I}_0} \frac{|\mathbf{b}_i^H\mathbf{L}_{\mathcal{G}}\boldsymbol{\alpha}|^2}{\|\mathbf{L}_{\mathcal{G}}\mathbf{b}_i\|_2^2}. \quad (30)$$

Given the fact that $\mathbf{S}$ is a sub-matrix of $\mathbf{E}$, from (25) we get

$$\begin{aligned}\sigma_{min}\left(\frac{1}{r\,card(\mathcal{I}_0)}\mathbf{SL}_{\mathcal{G}}\right) &\geq \sigma_{min}\left(\frac{1}{r\,card(\mathcal{I}_0)}\mathbf{E}\right)\gamma_2 \\ &\geq (1-\delta)\,\gamma_2, \quad (31)\end{aligned}$$

for some constant $\delta \in (0,1)$, and $\gamma_2 = \sigma_{min}(\mathbf{L}_{\mathcal{G}}) < 1$ with probability at least $1-2e^{-Cn}$ as long as L is sufficiently large, for some constant $C > 0$. Thus, from (31) it can be concluded that

$$\sum_{i\in\mathcal{I}_0} \frac{|\mathbf{b}_i^H\mathbf{L}_{\mathcal{G}}\boldsymbol{\alpha}|^2}{\|\mathbf{L}_{\mathcal{G}}\mathbf{b}_i\|_2^2} \geq \gamma_2 r(1-\delta)card(\mathcal{I}_0), \quad (32)$$

with probability at least $1-2e^{-Cn}$ as long as L is sufficiently large. Thus, from (32) we have that

$$\|\mathbf{SL}_{\mathcal{G}}\boldsymbol{\alpha}\|_2^2 \geq \gamma_2 r(1-\delta)card(\mathcal{I}_0), \quad (33)$$

holds with probability at least $1-2e^{-Cn}$, provided that L is sufficiently large. Therefore, from (33) the result holds. □

Hence, putting together (20) and (33) it can be concluded that

$$\frac{\|\mathbf{SL}_{\mathcal{G}}\mathbf{u}\|_2^2}{\|\mathbf{SL}_{\mathcal{G}}\boldsymbol{\alpha}\|_2^2} \leq \left(\frac{1+\delta-\zeta}{1-\delta}\right)\frac{\gamma_1}{\gamma_2} \triangleq \kappa_{\mathcal{G}} < 1, \quad (34)$$

by taking $\delta < \zeta\frac{\gamma_2}{2\gamma_1}$. Thus, putting together (19) and (34) it can be obtained that

$$\sin^2(\beta) = 1-\cos^2(\beta) \leq \kappa_{\mathcal{G}}. \quad (35)$$

On the other hand, notice that

$$\begin{aligned}dist^2(\hat{\mathbf{z}},\mathbf{x}) &\leq \|\mathbf{x}\|_2^2 + \|\hat{\mathbf{z}}\|_2^2 - 2|\mathbf{x}^H\hat{\mathbf{z}}| \\ &= \|\boldsymbol{\alpha}\|_2^2 + \|\hat{\boldsymbol{\theta}}\|_2^2 - |\boldsymbol{\alpha}^H\hat{\boldsymbol{\theta}}| \\ &= \|\boldsymbol{\alpha}\|_2^2 + \|\hat{\boldsymbol{\theta}}\|_2^2 - 2\cos(\beta) \\ &\leq \underbrace{2(1-\sqrt{1-\kappa_{\mathcal{G}}})}_{\delta_{\mathcal{G}}} + \mathcal{O}(\|\boldsymbol{\omega}\|_\infty), \quad (36)\end{aligned}$$

where the first equality comes from the fact that $\mathbf{Fx} = \boldsymbol{\alpha}$, and $\mathbf{F}$ is an orthogonal matrix. Then, combining (35) and (36) we finally conclude that

$$dist^2(\hat{\mathbf{z}},\mathbf{x}) < \delta_{\mathcal{G}} + \mathcal{O}(\|\boldsymbol{\omega}\|_\infty). \quad (37)$$

Thus, in (37) the result holds.

Analysis of the proof

Observe that (36) establishes the effect of the low-pass filter $\mathcal{G}$ to approximate $\mathbf{x}$. Additionally, notice that $\frac{\gamma_1}{\gamma_2}$ is the condition number of $\mathbf{L}_{\mathcal{G}}$. From this observation we can compare the performance between two different low-pass filters to solver (9). In fact, a filter $\mathcal{G}_1$ is able to better approximate the complex signal $\mathbf{x}$ compared to $\mathcal{G}_2$ if the condition number of $\mathbf{L}_{\mathcal{G}_1}$ is smaller than $\mathbf{L}_{\mathcal{G}_2}$.

On the other hand, performing an analog procedure to proof Lemmas 1,2 and 3 assuming that the matrix $\mathbf{L}_{\mathcal{G}}$ contains zeros in its diagonal, that is, $(\mathbf{L}_{\mathcal{G}})_{i,i} = 0$ for several $i \in \{1,\cdots,n\}$, we obtain the accuracy of solving (7) to approximate $\mathbf{x}$. Therefore, from the previous analysis we have that a matrix $\mathbf{L}_{\mathcal{G}_1}$ containing zeros in its diagonal will have a poorer performance compared to a matrix $\mathbf{L}_{\mathcal{G}_2}$ that does not contain zeros its diagonal, to approximate $\mathbf{x}$, because the rank of $\mathbf{L}_{\mathcal{G}_1}$ will be smaller than $\mathbf{L}_{\mathcal{G}_2}$. Thus, it is theoretically establishes that solving (9) using Algorithm 1 will return a closer approximation of $\mathbf{x}$ compared to a solution obtained by solving (7).

REFERENCES

[1] G. Wang, G. B. Giannakis, and Y. C. Eldar, "Solving systems of random quadratic equations via truncated amplitude flow," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 773–794, 2018.

[2] R. Vershynin, "Introduction to the non-asymptotic analysis of random matrices," *arXiv preprint arXiv:1011.3027*, 2010.

[3] G. Prémont and Y. Sheng, "Fast design of circular-harmonic filters using simulated annealing," *App. Opt.*, vol. 32, no. 17, pp. 3116–3121, 1993.

[4] P. Zi-Liang and E. Dalsgaard, "Synthetic circular-harmonic phase-only filter for shift, rotation, and scaling-invariant correlation," *App. Opt.*, vol. 34, no. 32, pp. 7527–7531, 1995.