A Non-Convex Approach to Blind Calibration for Linear Random Sensing Models

Valerio Cambareri*, Laurent Jacques*

ICTEAM/ELEN, ISPGroup, Université catholique de Louvain, Belgium.

***Abstract*— Performing blind calibration is highly important in modern sensing strategies, particularly when calibration aided by multiple, accurately designed training signals is infeasible or resource-consuming. We here address it as a naturally-formulated non-convex problem for a linear model with sub-Gaussian random sensing vectors in which both the sensor gains and the signal are unknown. A sample complexity bound is derived to assess that solving the problem by projected gradient descent with a suitable initialisation provably converges to the global optimum. These findings are supported by numerical evidence on the phase transition of blind calibration and by an imaging example.**

1 Introduction

The problem of acquiring an unknown signal under sensing model errors is crucial for modern sensing strategies such as Compressed Sensing (CS), in which such errors are inevitable in physical implementations, have direct impact on the attainable quality [1], and may be to some extent exploited for security purposes [2]. However, if the signals and model errors are stationary, the use of random sensing operators in CS also suggests that repeating the acquisition (*i.e.*, taking more *snapshots*) under new draws of the sensing model could suffice to diversify the measurements and learn both unknown quantities. We address the case of sensing a single unstructured vector $\boldsymbol{x} \in \mathbb{R}^n$ by collecting p snapshots of m random projections, *i.e.*,

$$\boldsymbol{y}_l = \bar{\boldsymbol{d}}\,\boldsymbol{A}_l\boldsymbol{x},\ \bar{\boldsymbol{d}} \coloneqq \mathrm{diag}(\boldsymbol{d}) \in \mathbb{R}^{m\times m},\ l \in [p] \coloneqq \{1,\dots,p\}, \tag{1}$$

where $\boldsymbol{y}_l = (y_{1,l},\cdots,y_{m,l})^\top \in \mathbb{R}^m$ is the l-th snapshot; $\boldsymbol{d} = (d_1,\cdots,d_m)^\top \in \mathbb{R}^m_+$ is an unknown, positive and bounded gain vector that is identical throughout the p snapshots; the random matrices $\boldsymbol{A}_l \in \mathbb{R}^{m\times n}$ are independent and identically distributed (i.i.d.) and each $\boldsymbol{A}_l$ has i.i.d. rows, the i-th row $\boldsymbol{a}_{i,l}^\top \in \mathbb{R}^n$ being a centred isotropic sub-Gaussian random vector (*i.e.*, $\mathbb{E}[\boldsymbol{a}_{i,l}] = \boldsymbol{0}_n, \mathbb{E}[\boldsymbol{a}_{i,l}\boldsymbol{a}_{i,l}^\top] = \boldsymbol{I}_n$).

This essentially bilinear model, referred to as *blind calibration* [3, 4] is motivated by compressive imaging applications, where unknown $\boldsymbol{d}$ are associated to physical gains and attenuations in a sensor array, while the random projections of an image $\boldsymbol{x}$ are obtained by light modulation (*e.g.*, by random convolution [5–8]). Similarly, model errors concern radar signal processing applications (*e.g.*, as discussed in [9, 10]). In such contexts, the knowledge of $\boldsymbol{d}$ could critically improve the accuracy of the sensing model. This is typically achieved by solving convex or alternating minimisation problems [3, 4, 11] that use multiple training signals $\boldsymbol{x}_l$ instead of randomising the sensing operator. More recently, *lifting* approaches [9, 10] have been proposed to jointly recover $(\boldsymbol{x}, \boldsymbol{d})$ in (1) (as well as more general *blind deconvolution* models, [12–14]). Their main limitation is in that a semidefinite program is solved to recover a very large rank-one matrix $\boldsymbol{x}\boldsymbol{d}^\top$, an approach that rapidly becomes computationally inefficient as m and n exceed a few hundreds.

*LJ and VC are funded by the Belgian FNRS. Part of this study is funded by the project ALTERSENSE (MIS-FNRS).

Quantity	Finite-sample ($p < \infty$)	Expectation ($\mathbb{E}_{\boldsymbol{a}_{i,l}}, p \to \infty$)
Objective: $f(\boldsymbol{\xi},\boldsymbol{\gamma})$	$\frac{1}{2mp}\sum_{l=1}^p \lVert\boldsymbol{\gamma}\boldsymbol{A}_l\boldsymbol{\xi} - \bar{\boldsymbol{d}}\boldsymbol{A}_l\boldsymbol{x}\rVert_2^2$	$\frac{1}{2m}\left[\lVert\boldsymbol{\xi}\rVert_2^2\lVert\boldsymbol{\gamma}\rVert_2^2 + \lVert\boldsymbol{x}\rVert_2^2\lVert\boldsymbol{d}\rVert_2^2 - 2(\boldsymbol{\gamma}^\top\boldsymbol{d})(\boldsymbol{\xi}^\top\boldsymbol{x})\right]$
Signal gradient: $\nabla_{\boldsymbol{\xi}} f(\boldsymbol{\xi},\boldsymbol{\gamma})$	$\frac{1}{mp}\sum_{l=1}^p \boldsymbol{A}_l^\top\bar{\boldsymbol{\gamma}}\left(\boldsymbol{\gamma}\boldsymbol{A}_l\boldsymbol{\xi} - \bar{\boldsymbol{d}}\boldsymbol{A}_l\boldsymbol{x}\right)$	$\frac{1}{m}\left[\lVert\boldsymbol{\gamma}\rVert_2^2\boldsymbol{\xi} - (\boldsymbol{\gamma}^\top\boldsymbol{d})\boldsymbol{x}\right]$
Gain gradient: $\nabla_{\boldsymbol{\gamma}} f(\boldsymbol{\xi},\boldsymbol{\gamma})$	$\frac{1}{mp}\sum_{l=1}^p \overline{\boldsymbol{A}_l\boldsymbol{\xi}}\left(\boldsymbol{\gamma}\boldsymbol{A}_l\boldsymbol{\xi} - \bar{\boldsymbol{d}}\boldsymbol{A}_l\boldsymbol{x}\right)$	$\frac{1}{m}\left[\lVert\boldsymbol{\xi}\rVert_2^2\boldsymbol{\gamma} - (\boldsymbol{\xi}^\top\boldsymbol{x})\boldsymbol{d}\right]$
Hessian: $\mathcal{H}f(\boldsymbol{\xi},\boldsymbol{\gamma})$	$\frac{1}{mp}\sum_{l=1}^p \begin{bmatrix} \boldsymbol{A}_l^\top\boldsymbol{\gamma}^2\boldsymbol{A}_l & \boldsymbol{A}_l^\top\overline{2\boldsymbol{\gamma}\boldsymbol{A}_l\boldsymbol{\xi} - \bar{\boldsymbol{d}}\boldsymbol{A}_l\boldsymbol{x}} \\ \overline{2\boldsymbol{\gamma}\boldsymbol{A}_l\boldsymbol{\xi} - \bar{\boldsymbol{d}}\boldsymbol{A}_l\boldsymbol{x}}\,\boldsymbol{A}_l & \overline{\boldsymbol{A}_l\boldsymbol{\xi}}^2 \end{bmatrix}$	$\frac{1}{m}\begin{bmatrix} \lVert\boldsymbol{\gamma}\rVert_2^2\boldsymbol{I}_n & 2\boldsymbol{\xi}\boldsymbol{\gamma}^\top - \boldsymbol{x}\boldsymbol{d}^\top \\ 2\boldsymbol{\gamma}\boldsymbol{\xi}^\top - \boldsymbol{d}\boldsymbol{x}^\top & \lVert\boldsymbol{\xi}\rVert_2^2\boldsymbol{I}_m \end{bmatrix}$
Initialisation: $(\boldsymbol{\xi}_0,\boldsymbol{\gamma}_0)$	$\left(\frac{1}{mp}\sum_{l=1}^p (\boldsymbol{A}_l)^\top\bar{\boldsymbol{d}}\boldsymbol{A}_l\boldsymbol{x},\,\boldsymbol{1}_m\right)$	$\left(\frac{\lVert\boldsymbol{d}\rVert_1}{m}\boldsymbol{x},\,\boldsymbol{1}_m\right)$

Table 1: Finite-sample and expected values of the objective function; its gradient components and Hessian matrix; the initialisation point $(\boldsymbol{\xi}_0,\boldsymbol{\gamma}_0)$. $\bar{\cdot}$ abbreviates the $\mathrm{diag}(\cdot)$ operator.

2 A Non-Convex Approach

Inspired by recent results on fast, provably convergent non-convex approaches to phase retrieval [15–18] we argue that the blind calibration problem of recovering the two unstructured vectors $(\boldsymbol{x}, \boldsymbol{d})$ in (1) can be solved exactly by a non-convex problem described hereafter. Since no *a priori* structure is assumed on $(\boldsymbol{x}, \boldsymbol{d})$ we operate in the case $mp \geq n + m$ and solve

$$(\hat{\boldsymbol{x}}, \hat{\boldsymbol{d}}) = \underset{\boldsymbol{\xi}\in\mathbb{R}^n, \boldsymbol{\gamma}\in\Pi_+^m}{\mathrm{argmin}} \frac{1}{2mp}\sum_{l=1}^p \lVert\bar{\boldsymbol{\gamma}}\boldsymbol{A}_l\boldsymbol{\xi} - \boldsymbol{y}_l\rVert_2^2, \tag{2}$$

given $\{\boldsymbol{y}_l\}_{l=1}^p, \{\boldsymbol{A}_l\}_{l=1}^p$ with $\Pi_+^m = \{\boldsymbol{\gamma} \in \mathbb{R}_+^m, \boldsymbol{1}_m^\top\boldsymbol{\gamma} = m\}$ being the scaled probability simplex. The geometry of this problem is well understood by observing the objective $f(\boldsymbol{\xi},\boldsymbol{\gamma})$ with its gradient and Hessian matrix, computed in Table 1 for finite and asymptotic p. All finite-sample expressions therein are unbiased estimates of their expectation w.r.t. $\boldsymbol{a}_{i,l}$. There, we evince that all points in $\left\{(\boldsymbol{\xi},\boldsymbol{\gamma}) \in \mathbb{R}^n \times \mathbb{R}^m : \boldsymbol{\xi} = \frac{1}{\alpha}\boldsymbol{x}, \boldsymbol{\gamma} = \alpha\boldsymbol{d}, \alpha \in \mathbb{R}\setminus\{0\}\right\}$ are global minimisers of $f(\boldsymbol{\xi},\boldsymbol{\gamma})$. Moreover, $f(\boldsymbol{\xi},\boldsymbol{\gamma})$ is easily shown to be generally non-convex[1] as there exist plenty of counterexamples $(\boldsymbol{\xi}', \boldsymbol{\gamma}')$ for which $\mathcal{H}f(\boldsymbol{\xi}', \boldsymbol{\gamma}') \not\succeq 0$.

By applying the constraint $\boldsymbol{\gamma} \in \Pi_+^m$, one minimiser $(\boldsymbol{x}^\star, \boldsymbol{d}^\star) = \left(\frac{\lVert\boldsymbol{d}\rVert_1}{m}\boldsymbol{x}, \frac{m}{\lVert\boldsymbol{d}\rVert_1}\boldsymbol{d}\right)$ remains for $mp \geq n + m$, which is exact up to $\alpha = m/\lVert\boldsymbol{d}\rVert_1$. Assuming that $\boldsymbol{d} \in \mathbb{R}_+^m$ in (1) is bounded amounts to letting[2] $\boldsymbol{d}^\star \coloneqq \boldsymbol{1}_m + \boldsymbol{\omega} \in \mathcal{C}_\rho \subset \Pi_+^m$, $\boldsymbol{\omega} \in \boldsymbol{1}_m^\perp \cap \rho\,\mathbb{B}_\infty^m$ for a maximum deviation $\rho = \lVert\boldsymbol{d}^\star - \boldsymbol{1}_m\rVert_\infty < 1$ where $\mathcal{C}_\rho \coloneqq \boldsymbol{1}_m + \boldsymbol{1}_m^\perp \cap \rho\,\mathbb{B}_\infty^m$. The test vector is also cast as $\boldsymbol{\gamma} \coloneqq \boldsymbol{1}_m + \boldsymbol{\varepsilon} \in \mathcal{C}_\rho$. While the problem clearly remains non-convex, by defining a *distance* $\Delta(\boldsymbol{\xi},\boldsymbol{\gamma}) \coloneqq \lVert\boldsymbol{\xi} - \boldsymbol{x}^\star\rVert_2^2 + \frac{\lVert\boldsymbol{x}^\star\rVert_2^2}{m}\lVert\boldsymbol{\gamma} - \boldsymbol{d}^\star\rVert_2^2$, and a *neighbourhood* of the global minimiser $(\boldsymbol{x}^\star, \boldsymbol{d}^\star)$ as $\mathcal{D}_{\kappa,\rho} \coloneqq \{(\boldsymbol{\xi},\boldsymbol{\gamma}) \in \mathbb{R}^n \times \mathcal{C}_\rho : \Delta(\boldsymbol{\xi},\boldsymbol{\gamma}) \leq \kappa^2\lVert\boldsymbol{x}^\star\rVert_2^2\}$ for $\kappa, \rho \in [0, 1)$, it is easily shown that (2) is locally convex in expectation on $\mathcal{D}_{\kappa,\rho}$ for sufficiently small κ. Thus, we require an *initialisation* point $(\boldsymbol{\xi}_0,\boldsymbol{\gamma}_0)$ to lie in a small $\mathcal{D}_{\kappa,\rho}$ for which a notion of local convexity holds for finite p. Similarly to [15] we see that initialising $\boldsymbol{\xi}_0$ as in Table 1 grants $\mathbb{E}[\boldsymbol{\xi}_0] \equiv \boldsymbol{x}^\star$, *i.e.*, asymptotically in p this initialisation yields the exact solution. We also let $\boldsymbol{\gamma}_0 = \boldsymbol{1}_m$ ($\boldsymbol{\varepsilon}_0 = \boldsymbol{0}_m$) and run projected gradient descent to solve (2), as summarised in Algorithm 1. Note that, since we assume ρ to be small (*i.e.*, $\boldsymbol{d}^\star$ is far from the vertices of Π_+^m) and the optimisation starts on

[1] It is *biconvex* [10], *i.e.*, convex once either $\boldsymbol{\xi}$ or $\boldsymbol{\gamma}$ are fixed.

[2] The orthogonal complement of $\boldsymbol{1}_m$ is $\boldsymbol{1}_m^\perp = \{\boldsymbol{v} \in \mathbb{R}^m : \boldsymbol{1}_m^\top\boldsymbol{v} = 0\}$.

1: Initialise $\boldsymbol{\xi}_0 \coloneqq \frac{1}{mp}\sum_{l=1}^{p}(\boldsymbol{A}_l)^\top \boldsymbol{y}_l$, $\boldsymbol{\gamma}_0 \coloneqq \mathbf{1}_m$, $k \coloneqq 0$.
2: **while** stop criteria not met **do**
3: $\quad \mu_{\boldsymbol{\xi}} \coloneqq \operatorname{argmin}_{\upsilon\in\mathbb{R}} f(\boldsymbol{\xi}_k - \upsilon\nabla_{\boldsymbol{\xi}} f(\boldsymbol{\xi}_k,\boldsymbol{\gamma}_k),\boldsymbol{\gamma}_k)$ {Line search in $\boldsymbol{\xi}$}
4: $\quad \mu_{\boldsymbol{\gamma}} \coloneqq \operatorname{argmin}_{\upsilon\in\mathbb{R}} f(\boldsymbol{\xi}_k, \boldsymbol{\gamma}_k - \upsilon\nabla^{\perp}_{\boldsymbol{\gamma}} f(\boldsymbol{\xi}_k,\boldsymbol{\gamma}_k))$ {Line search in $\boldsymbol{\gamma}$}
5: $\quad \boldsymbol{\xi}_{k+1} \coloneqq \boldsymbol{\xi}_k - \mu_{\boldsymbol{\xi}}\nabla_{\boldsymbol{\xi}} f(\boldsymbol{\xi}_k,\boldsymbol{\gamma}_k)$
6: $\quad \underline{\boldsymbol{\gamma}}_{k+1} \coloneqq \boldsymbol{\gamma}_k - \mu_{\boldsymbol{\gamma}}\nabla^{\perp}_{\boldsymbol{\gamma}} f(\boldsymbol{\xi}_k,\boldsymbol{\gamma}_k)$
7: $\quad \boldsymbol{\gamma}_{k+1} \coloneqq P_{\mathcal{C}_\rho}\underline{\boldsymbol{\gamma}}_{k+1}$ {Projection on $\mathcal{C}_\rho$}
8: $\quad k \coloneqq k+1$
9: **end while**

Algorithm 1: Non-Convex Blind Calibration by Projected Gradient Descent.

$\boldsymbol{\gamma}_0 = \mathbf{1}_m$, we are allowed to first project the gradient step as $\nabla^{\perp}_{\boldsymbol{\gamma}} f(\boldsymbol{\xi},\boldsymbol{\gamma}) \coloneqq \boldsymbol{P}_{\mathbf{1}^{\perp}_m}\nabla_{\boldsymbol{\gamma}} f(\boldsymbol{\xi},\boldsymbol{\gamma})$ with $\boldsymbol{P}_{\mathbf{1}^{\perp}_m} = \boldsymbol{I}_m - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m^\top$, then project it on $\mathcal{C}_\rho$ (*i.e.*, $P_{\mathcal{C}_\rho}$; this step is useful for the proofs and not required in all the experiments). The line searches in 3:-4: of Algorithm 1 are solved in closed-form and introduced to improve the convergence rate. Below, we obtain the conditions that ensure convergence of this descent algorithm to the exact solution $(\boldsymbol{x}^\star, \boldsymbol{d}^\star)$ for fixed step sizes $\mu_{\boldsymbol{\xi}}, \mu_{\boldsymbol{\gamma}}$.

3 Recovery Guarantees

Having assumed that all mp sensing vectors $\boldsymbol{a}_{i,l}$ are i.i.d. sub-Gaussian, we establish the sample complexity required for the initialisation to lie in $\mathcal{D}_{\kappa,\rho}$ with very high probability. This and all other statements rely on Bernstein-type concentration inequalities [19] and some simple geometry of (2) in $\mathcal{D}_{\kappa,\rho}$.

Proposition 1 (Initialisation Proximity)**.** *Let $(\boldsymbol{\xi}_0,\boldsymbol{\gamma}_0)$ be as in Algorithm 1. For any $\epsilon \in (0,1), t > 1$, provided $n \gtrsim t\log(mp)$ and $mp \gtrsim \epsilon^{-2}(n+m)\log\left(n\epsilon^{-1}\right)$ we have that, with probability exceeding $1 - Ce^{-c\epsilon^2 mp} - (mp)^{-t}$ for some $C, c > 0$, $\|\boldsymbol{\xi}_0 - \boldsymbol{x}^\star\|_2 \le \epsilon\|\boldsymbol{x}^\star\|_2$. Since $\boldsymbol{\gamma}_0 = \mathbf{1}_m$ we also have $\|\boldsymbol{\gamma}_0 - \boldsymbol{d}^\star\|_\infty \le \rho < 1$. Thus $(\boldsymbol{\xi}_0,\boldsymbol{\gamma}_0) \in \mathcal{D}_{\kappa,\rho}$ with the same probability and $\kappa \coloneqq \sqrt{\epsilon^2+\rho^2}$.*

With analogue tools we are able to state when a single gradient descent step from any $(\boldsymbol{\xi},\boldsymbol{\gamma})$ reduces, with very high probability, the distance w.r.t. the global minimum. This requires establishing a *regularity condition* on $\nabla f(\boldsymbol{\xi},\boldsymbol{\gamma})$ holding uniformly for all $(\boldsymbol{\xi},\boldsymbol{\gamma}) \in \mathcal{D}_{\kappa,\rho}$ (similarly to [15, Condition 7.9]).

Proposition 2 (Regularity condition in $\mathcal{D}_{\kappa,\rho}$)**.** *For any $\delta \in (0,1), \rho \in [0,1), t > 1$, provided $\rho < \frac{1-2\delta}{9}$, $n \gtrsim t\log(mp)$, $p \gtrsim \delta^{-2}\log m$ and $\sqrt{mp} \gtrsim \delta^{-2}(n+m)\log(\frac{n}{\delta})$, with probability exceeding $1 - C\left[me^{-c\delta^2 p} + e^{-c\delta^2\sqrt{mp}} + (mp)^{-t}\right]$ for some $C, c > 0$, we have that for all $(\boldsymbol{\xi},\boldsymbol{\gamma}) \in \mathcal{D}_{\kappa,\rho}$,*

$$\left\langle \nabla^{\perp} f(\boldsymbol{\xi},\boldsymbol{\gamma}), \begin{bmatrix}\boldsymbol{\xi}-\boldsymbol{x}^\star \\ \boldsymbol{\gamma}-\boldsymbol{d}^\star\end{bmatrix}\right\rangle \ge \tfrac{1}{2}\eta\,\Delta(\boldsymbol{\xi},\boldsymbol{\gamma}) \quad \text{(Bounded curvature)}$$

$$\|\nabla^{\perp} f(\boldsymbol{\xi},\boldsymbol{\gamma})\|_2^2 \le L^2\,\Delta(\boldsymbol{\xi},\boldsymbol{\gamma}) \quad \text{(Lipschitz gradient)}$$

for $\eta \coloneqq 2(1-9\rho-2\delta) > 0$, $L \coloneqq 4\sqrt{2}[1+\rho+(1+\kappa)\|\boldsymbol{x}^\star\|_2]$.

From Proposition 2 it follows that the neighbourhood $\mathcal{D}_{\kappa,\rho}$ is a basin of attraction to the solution $(\boldsymbol{x}^\star,\boldsymbol{d}^\star)$, *i.e.*, we can show that the update from any $(\boldsymbol{\xi}_k,\boldsymbol{\gamma}_k) \in \mathcal{D}_{\kappa,\rho}$ to $(\boldsymbol{\xi}_{k+1},\boldsymbol{\gamma}_{k+1})$ is so that the distance $\Delta(\boldsymbol{\xi},\boldsymbol{\gamma})$ decreases under an upper bound on $\mu_{\boldsymbol{\xi}}, \mu_{\boldsymbol{\gamma}}$ specified in the following result.

Theorem 1 (Provable Convergence to the Exact Solution)**.** *Under the conditions of Proposition 1, 2 we have that, with probability exceeding $1 - C\left[me^{-c\delta^2 p} + e^{-c\delta^2\sqrt{mp}} + e^{-c\epsilon^2 mp} + (mp)^{-t}\right]$ for some $C, c > 0$, Algorithm 1 with $\mu_{\boldsymbol{\xi}} \coloneqq \mu, \mu_{\boldsymbol{\gamma}} \coloneqq \mu\frac{m}{\|\boldsymbol{x}^\star\|_2^2}$ has error decay*

$$\Delta(\boldsymbol{\xi}_k,\boldsymbol{\gamma}_k) \le \left(1-\eta\mu+\tfrac{L^2}{\tau}\mu^2\right)^k(\epsilon^2+\rho^2)\|\boldsymbol{x}^\star\|_2^2, (\boldsymbol{\xi}_k,\boldsymbol{\gamma}_k) \in \mathcal{D}_{\kappa,\rho}$$

at any iteration $k > 0$ provided $\mu \in \left(0, \tau\eta/L^2\right)$, $\tau \coloneqq \min\{1, \|\boldsymbol{x}^\star\|_2^2/m\}$. Hence, $\Delta(\boldsymbol{\xi}_k,\boldsymbol{\gamma}_k) \underset{k\to\infty}{\longrightarrow} 0$.

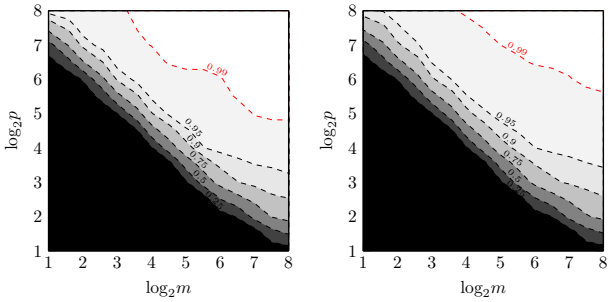

Figure 1: Empirical phase transition of (2) for $n = 2^8$, $\rho = 10^{-3}$ (left) and $\rho = 10^{-0.5}$ (right). The probability value P_ζ is reported on each level set.

(a) $\boldsymbol{x}$ (b) $\boldsymbol{d}$, $\rho = 1/2$ (c) $\hat{\boldsymbol{x}}$ by LS (d) $\hat{\boldsymbol{x}}$ by (2) (e) $\hat{\boldsymbol{d}}$ by (2)

Figure 2: A practical, high-dimensional example of blind calibration against unknown, unstructured gains $\boldsymbol{d}$ with $m = n, p = 4$ snapshots.

Let us mention finally that if additive measurement noise $\boldsymbol{N} \coloneqq (\boldsymbol{n}_1,\cdots,\boldsymbol{n}_p) \in \mathbb{R}^{m\times p}$ corrupts the p snapshots of (1), then it can be shown that $\Delta(\boldsymbol{\xi}_k,\boldsymbol{\gamma}_k) \to \sigma^2$ as $k \to \infty$, with $\sigma^2 \gtrsim \frac{1}{mp}\|\boldsymbol{N}\|_F^2$. Thus, Algorithm 1 is robust to noise, its solution degrading gracefully with σ^2.

4 Numerical Experiments

Empirical Phase Transition: To characterise the phase transition of (2) solved by Algorithm 1 we ran some extensive simulations by generating 256 random instances of (2) for each configuration $\log_2 n \times \log_2 m \times \log_2 p \times \log_{10}\rho \in \{1,\dots,8\}^3 \times \{-3,\dots,0\}$, drawing $\boldsymbol{d}^\star = \mathbf{1}_m + \boldsymbol{\omega}$ with $\boldsymbol{\omega} \in \mathbf{1}_m^{\perp}$ drawn uniformly at random on the sphere $\rho\,\mathbb{S}_\infty^{m-1}$. Then we evaluated $\mathrm{P}_\zeta \coloneqq \mathbb{P}\left[\max\left\{\frac{\|\hat{\boldsymbol{d}}-\boldsymbol{d}^\star\|_2}{\|\boldsymbol{d}^\star\|_2}, \frac{\|\hat{\boldsymbol{x}}-\boldsymbol{x}^\star\|_2}{\|\boldsymbol{x}^\star\|_2}\right\} < \zeta\right]$ on the trials with $\zeta = 10^{-3}$ chosen according to the stop criterion $f(\boldsymbol{\xi},\boldsymbol{\gamma}) < 10^{-7}$. Of this large dataset we only report the case $n = 2^8$ in Fig. 1, highlighting the contour levels of P_ζ for $\rho = \{10^{-3}, 10^{-0.5}\}$ as a function of $\log_2 m$ and $\log_2 p$.

Blind Calibration of an Imaging System: To test (2) in a realistic context, we assume that $\boldsymbol{x}$ is a $n = 64\times 64$ pixel image acquired by an imaging system following (1), in which its $m = 64\times 64$ pixel sensor array suffers from large *pixel response non-uniformity* [20]. This is simulated by generating $\boldsymbol{d}$ as before with $\rho = 1/2$. We capture $p = 4$ snapshots with $\boldsymbol{a}_{i,l} \sim \mathcal{N}(\mathbf{0}_n, \boldsymbol{I}_n)$. By running Algorithm 1, the recovered $(\hat{\boldsymbol{x}},\hat{\boldsymbol{d}}) \equiv (\boldsymbol{x},\boldsymbol{d})$ attains $\max\left\{\frac{\|\hat{\boldsymbol{d}}-\boldsymbol{d}^\star\|_2}{\|\boldsymbol{d}^\star\|_2}, \frac{\|\hat{\boldsymbol{x}}-\boldsymbol{x}^\star\|_2}{\|\boldsymbol{x}^\star\|_2}\right\} \approx -147.38\,\mathrm{dB}$. Instead, by fixing $\boldsymbol{\gamma} = \mathbf{1}_m$ and solving (2) only in $\boldsymbol{\xi}$, *i.e.*, finding a least squares (LS) $\hat{\boldsymbol{x}}$, we obtain $\frac{\|\hat{\boldsymbol{x}}-\boldsymbol{x}^\star\|_2}{\|\boldsymbol{x}^\star\|_2} \approx -5.50\,\mathrm{dB}$.

5 Conclusion

We presented and solved a non-convex formulation of the blind calibration problem for a random linear model. The obtained sample complexity could be further improved by exploiting, *e.g.*, a sparse model for $\boldsymbol{x},\boldsymbol{d}$; in absence of such priors, the exact solution is achieved at a rate $\sqrt{mp} \gtrsim (n+m)\log(n)$.

References

[1] M. A. Herman and T. Strohmer, "General deviants: An analysis of perturbations in compressed sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 342–349, 2010.

[2] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "Low-complexity multiclass encryption by compressed sensing," *IEEE Transactions on Signal Processing*, vol. 63, no. 9, pp. 2183–2195, 2015.

[3] L. Balzano and R. Nowak, "Blind calibration of networks of sensors: Theory and algorithms," in *Networked Sensing Information and Control*. Springer, 2008, pp. 9–37.

[4] C. Bilen, G. Puy, R. Gribonval, and L. Daudet, "Convex Optimization Approaches for Blind Sensor Calibration Using Sparsity," *IEEE Transactions on Signal Processing*, vol. 62, no. 18, pp. 4847–4856, Sep. 2014.

[5] J. Romberg, "Compressive sensing by random convolution," *SIAM Journal on Imaging Sciences*, vol. 2, no. 4, pp. 1098–1128, 2009.

[6] S. Bahmani and J. Romberg, "Compressive deconvolution in random mask imaging," *IEEE Transactions on Computational Imaging*, vol. 1, no. 4, pp. 236–246, 2015.

[7] T. Bjorklund and E. Magli, "A parallel compressive imaging architecture for one-shot acquisition," in *2013 IEEE Picture Coding Symposium (PCS)*. IEEE, 2013, pp. 65–68.

[8] K. Degraux, V. Cambareri, B. Geelen, L. Jacques, G. Lafruit, and G. Setti, "Compressive Hyperspectral Imaging by Out-of-Focus Modulations and Fabry-Pérot Spectral Filters," in *International Traveling Workshop on Interactions between Sparse models and Technology (iTWIST)*, 2014.

[9] B. Friedlander and T. Strohmer, "Bilinear compressed sensing for array self-calibration," in *2014 48th Asilomar Conference on Signals, Systems and Computers*, Nov. 2014, pp. 363–367.

[10] S. Ling and T. Strohmer, "Self-calibration and biconvex compressive sensing," *Inverse Problems*, vol. 31, no. 11, p. 115002, 2015.

[11] J. Lipor and L. Balzano, "Robust blind calibration via total least squares," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 4244–4248.

[12] A. Ahmed, B. Recht, and J. Romberg, "Blind deconvolution using convex programming," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1711–1732, 2014.

[13] S. Ling and T. Strohmer, "Blind deconvolution meets blind demixing: Algorithms and performance bounds," *arXiv preprint arXiv:1512.07730*, 2015.

[14] A. Ahmed, A. Cosse, and L. Demanet, "A convex approach to blind deconvolution with diverse inputs," in *2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec 2015, pp. 5–8.

[15] E. Candès, X. Li, and M. Soltanolkotabi, "Phase Retrieval via Wirtinger Flow: Theory and Algorithms," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1985–2007, Apr. 2015.

[16] C. D. White, S. Sanghavi, and R. Ward, "The local convexity of solving systems of quadratic equations," *arXiv:1506.07868 [math, stat]*, Jun. 2015, arXiv: 1506.07868.

[17] Y. Chen and E. Candès, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," in *Advances in Neural Information Processing Systems*, 2015, pp. 739–747.

[18] J. Sun, Q. Qu, and J. Wright, "A geometric analysis of phase retrieval," in *2016 IEEE International Symposium on Information Theory (ISIT)*, July 2016, pp. 2379–2383.

[19] R. Vershynin, "Introduction to the non-asymptotic analysis of random matrices," in *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012, pp. 210–268.

[20] M. M. Hayat, S. N. Torres, E. Armstrong, S. C. Cain, and B. Yasuda, "Statistical algorithm for nonuniformity correction in focal-plane arrays," *Applied Optics*, vol. 38, no. 5, pp. 772–780, 1999.