INDEPENDENT VERSUS REPEATED MEASUREMENTS: A PERFORMANCE QUANTIFICATION VIA STATE EVOLUTION

Yang Lu and Wei Dai

Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom

ABSTRACT

The paper quantifies and compares the exact asymptotic performance of multiple measurement vector (MMV) and distributed sensing (DS) models. Both models assume multiple measurement instances $\boldsymbol{y}_k = \boldsymbol{A}_k\boldsymbol{x}_k + \boldsymbol{w}_k,\ k = 1, 2, ..., K$. The difference is that MMV involves identical measurement matrices whereas DS allows different matrices for different measurement instances. It has been recognized that DS works better than MMV empirically. However, the quantification of the performance difference is not available in the literature. Our contribution is to quantify the asymptotic performance of MMV and DS in the asymptotic regime that the dimensions of the measurement matrices approach infinity proportionally but the number of measurement instances K remains a constant. The case study and numerical results justify the accuracy of the performance quantification. The analysis technique is based on the state evolution for approximate message passing.

***Index Terms*—** Distributed sensing, message passing, multiple measurement vector (MMV), sparse recovery.

1. INTRODUCTION

Compressed sensing (CS) has achieved tremendous success in the recent decade due to its wide applications in sensing and learning. Its core is to solve a linear inverse problem where the solution is supposed to be sparse, i.e., the number of nonzeros in the solution is much smaller than the signal dimension. In the standard setting only one measurement instance is involved, that is, one linear system is concerned. Nevertheless in many application scenarios, multiple linear systems are involved and the solutions of these linear systems are supposed to be somewhat dependent [1, 2]. Refer to this setting as multiple measurement instances. The applications can be found in distributed sensor networks [3], direction of arrival (DOA) estimation [4] and parallel magnetic resonance imaging (pMRI) [5], to name a few.

The focus of this paper is to quantify and compare the performance of multiple measurement vector (MMV) and distributed sensing (DS). Consider the following model for multiple measurement instances: $\boldsymbol{y}_k = \boldsymbol{A}_k\boldsymbol{x}_k + \boldsymbol{w}_k,\ k = 1, 2, ..., K$. MMV assumes that all the measurement matrices $\boldsymbol{A}_k$'s are identical whereas DS allows different matrices for different measurement instances. When the signals $\boldsymbol{x}_k$'s are sparse, we assume that they share the same supports, i.e., the locations of the nonzero components. This signal model is often referred to as common sparse supports (or JSM-2 model in [1]). In this paper, both sparse and non-sparse signals are analyzed.

Various methods and analysis are reported in the literature. Examples include MMV focal under-determined system solution (M-FOCUSS) [6], DCS-tailored Simultaneous Orthogonal Matching Pursuit (DCS-SOMP) [1], subspace augmented multiple signal classifier (SA-MUSIC) [7], side-information based OMP (SiOMP) [8], distributed and collaborative OMP (DC-OMP) [9, 10], approximate message passing MMV (AMP-MMV) [11], and distributed subspace pursuit (DiSP) [2]. Empirical results have shown that generally speaking DS outperforms MMV. However, exact understanding of the performance improvement is not available in the literature.

The main contribution of this paper is to exactly quantify the asymptotic performance of MMV and DS in the asymptotic regime that the dimensions of the measurement matrices approach infinity proportionally but the number of measurement instances remains a constant. It is assumed that the measurement matrices are drawn from the standard Gaussian random matrix ensemble and the signals obey a Bernoulli-Gaussian model. In DS, we further assume that the measurement matrices are independent. Our analysis reveals that the benefits of DS come from the independent views of correlated signals (the more correlated the signals, the larger performance gain DS achieves). A case study of non-sparse Gaussian signals is given where performance quantification can be obtained via random matrix theory and matches our analysis. Both case study and numerical results verify the accuracy of the asymptotic quantification. The analysis tool used is the state evolution technique for the approximate message passing (AMP) [12, 13]. It is noteworthy that similar analysis appeared in [14, 15] where [14] focuses on MMV while [15] concentrates on DS. However, these two works don't characterize the exact evolution of the estimation error for the MMV model, nor the performance gap between MMV and DS.

This work was supported in part by Defence Science and Technology Laboratory (Dstl) under Grant No: DSTLX-1000081291

2. PRELIMINARIES

2.1. System model

Consider K measurement instances, at each of which a linear measurement operator is acting on an unknown signal:

$$\boldsymbol{y}_k \;=\; \boldsymbol{A}_k\boldsymbol{x}_k+\boldsymbol{w}_k,\quad k\in[K]\,, \tag{1}$$

where $\boldsymbol{y}_k\in\mathbb{R}^m$ is the measurement vector, $\boldsymbol{A}_k\in\mathbb{R}^{m\times n}$ stands for the measurement matrix, $\boldsymbol{x}_k\in\mathbb{R}^n$ represents the unknown signal vector, $\boldsymbol{w}_k\in\mathbb{R}^m$ denotes the measurement noise and $[K]\coloneqq\{1,2,...,K\}$. The MMV model assumes that all the measurement matrices are identical, i.e. $\boldsymbol{A}_k=\boldsymbol{A}_\ell$ for all $k,\ell\in[K]$ [11, 16], whereas in DS model the measurement matrices are allowed to be different [1]. In this paper for the convenience of analysis, we assume that the measurement matrices are generated from the Gaussian random matrix ensemble for both MMV and DS cases. In the MMV model, let all the entries of $\boldsymbol{A}_1$ be independently drawn from the Gaussian distribution $\mathcal{N}\left(0,\frac{1}{m}\right)$ and then $\boldsymbol{A}_k=\boldsymbol{A}_1$ for all $1<k\leq K$. In the DS case, all the entries of $\boldsymbol{A}_k$'s are independently drawn from the same distribution $\mathcal{N}\left(0,\frac{1}{m}\right)$. For compositional convenience, sometimes we write the overall system in the equivalent form

$$\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{w}, \tag{2}$$

where $\boldsymbol{y}=\left[\boldsymbol{y}_1^T,\boldsymbol{y}_2^T,...,\boldsymbol{y}_K^T\right]^T$, $\boldsymbol{w}=\left[\boldsymbol{w}_1^T,\boldsymbol{w}_2^T,...,\boldsymbol{w}_K^T\right]^T$, $\boldsymbol{x}=\left[\boldsymbol{x}_1^T,\boldsymbol{x}_2^T,...,\boldsymbol{x}_K^T\right]^T$ and $\boldsymbol{A}=\mathrm{diag}\left(\boldsymbol{A}_1,\boldsymbol{A}_2,...,\boldsymbol{A}_K\right)$.

For MMV and distributed compressed sensing cases, the common sparse supports model has been widely adopted to model the unknown signals $\boldsymbol{x}_k$'s [1, 11]. Let $\mathrm{supp}\left(\boldsymbol{x}\right)=\{i:\;x_i\neq0\}$ be the support set of the signal vector $\boldsymbol{x}$. Common sparse supports means that $\mathrm{supp}\left(\boldsymbol{x}_k\right)=\mathrm{supp}\left(\boldsymbol{x}_\ell\right)$ for all $k\neq\ell\in[K]$. It is typically assumed that the nonzero signal components are independent. A more delicate model is that though the nonzero components in $\boldsymbol{x}_k$ are independent, those across different measurement instances can be correlated. In particular, let $x_{k,j}$ be the i-th component of the signal $\boldsymbol{x}_k$ at the k-th measurement instance. Define

$$\boldsymbol{x}_{:,i}\coloneqq\left[x_{1,i},x_{2,i},\cdots,x_{K,i}\right]^{\mathrm{T}},\,for\;i\in[n]\,, \tag{3}$$

which contains the i-th components of all K signals. We assume a commonly used Bernoulli-Gaussian model: $\boldsymbol{x}_{:,i}\in\mathbb{R}^K$, $i\in[n]$, are independently drawn from the distribution

$$p_{\boldsymbol{X}_{:,i}}\left(\boldsymbol{x}\right) \;=\; \left(1-\epsilon\right)\delta_{\boldsymbol{x}=\boldsymbol{0}}+\epsilon p_G\left(\boldsymbol{x};\,\boldsymbol{0},\,\boldsymbol{\Sigma}_x\right), \tag{4}$$

where $\epsilon\in(0,1]$ defines the nonzero probability (sparsity level), $\delta_{\boldsymbol{x}=\boldsymbol{0}}$ is the Dirac delta function, $p_G\left(\boldsymbol{x};\,\boldsymbol{0},\,\boldsymbol{\Sigma}_x\right)$ is the multivariate Gaussian probability density, $\boldsymbol{\Sigma}_x\in\mathbb{R}^{K\times K}$ is the covariance matrix, and we assume the diagonal elements of $\boldsymbol{\Sigma}_x$ are all identical, denoted by σ_x^2. It is clear that if all nonzero components across different measurement instances are independent then $\boldsymbol{\Sigma}_x$ is diagonal, otherwise $\boldsymbol{\Sigma}_x$ is not diagonal. It is also noteworthy that the analysis presented in this paper can be generalized to the non-sparse case where $\epsilon=1$.

2.2. AMP and state evolution

The AMP algorithm was originally developed in [12] and then rigorously analyzed in [13]. It is designed to solve the linear inverse problem (2) when $\boldsymbol{A}$ is generated from the Gaussian random matrix ensemble. It is an iterative algorithm and at the t-th iteration it executes the following two operations:

$$\boldsymbol{x}^{t+1}=\boldsymbol{\eta}\left(\boldsymbol{A}^T\boldsymbol{r}^t+\boldsymbol{x}^t\right), \tag{5}$$

$$\boldsymbol{r}^t=\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}^t+\frac{1}{\delta}\boldsymbol{r}^{t-1}\left\langle\boldsymbol{\eta}'\left(\boldsymbol{A}^T\boldsymbol{r}^{t-1}+\boldsymbol{x}^{t-1}\right)\right\rangle, \tag{6}$$

where $\boldsymbol{x}^{t+1}$ is the new estimated signal, $\boldsymbol{\eta}\left(\cdot\right)$ is a component-wise estimator, $\boldsymbol{A}^T$ is the transpose of $\boldsymbol{A}$, $\boldsymbol{r}^t$ is the residual vector, $\delta=\frac{m}{n}$ is a constant, $\langle\boldsymbol{v}\rangle=\frac{1}{n}\sum_{i=1}^n v_i$ computes the average of the vector $\boldsymbol{v}$, and $\boldsymbol{\eta}'\left(\cdot\right)$ denotes the derivative of the function $\boldsymbol{\eta}\left(\cdot\right)$.

Two advantages of AMP include its low complexity and the exact quantification of its asymptotic performance when $(m,n)\to\infty$ simultaneously with $\frac{m}{n}\to\delta\in\mathbb{R}^+$. The performance quantification is based on the following observation. Let $\boldsymbol{x}_0$ be the ground truth signal that generates $\boldsymbol{y}$ in (2). Let $\tilde{\boldsymbol{x}}^t$ denotes the input of $\boldsymbol{\eta}\left(\cdot\right)$ in (5). Based on the heuristic analysis for iterative approaches in [12], write $\tilde{\boldsymbol{x}}^t$ in the form

$$\begin{aligned}\tilde{\boldsymbol{x}}^t &= \boldsymbol{A}^T\boldsymbol{r}^t+\boldsymbol{x}^t \;=\; \boldsymbol{A}^T\left(\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}^t\right)+\boldsymbol{x}^t\\ &= \boldsymbol{x}_0+\left(\boldsymbol{A}^T\boldsymbol{A}-\boldsymbol{I}\right)\left(\boldsymbol{x}_0-\boldsymbol{x}^t\right)+\boldsymbol{A}^T\boldsymbol{w}\\ &= \boldsymbol{x}_0+\boldsymbol{w}_x^t+\boldsymbol{A}^T\boldsymbol{w} \;=\; \boldsymbol{x}_0+\boldsymbol{w}_e^t,\end{aligned} \tag{7}$$

where $\boldsymbol{w}_e^t$ denotes the equivalent noise at the t-th iteration. Then the rigorous analysis in [13] shows that $\boldsymbol{w}_e^t$ can be modeled as white Gaussian noise and will be independent of the equivalent noise at previous iterations $\boldsymbol{w}_e^{t'}$ where $1\leq t'<t$. The mean of this equivalent noise is zero and the variance can be computed via the so called state evolution. By tracking the variance of the equivalent noise across iterations, the asymptotic performance of AMP can be exactly quantified.

3. PERFORMANCE QUANTIFICATION VIA STATE EVOLUTION

This section derives the state evolution for both MMV and DS cases. In the analysis, it is assumed that $(m,n)\to\infty$ with $\frac{m}{n}\to\delta\in\mathbb{R}^+$ but K is fixed.

For the performance analysis, the covariance matrix of $\boldsymbol{x}_{:,i}$, $i\in[n]$, needs to be tracked across iterations. Towards this end, consider the following model

$$\tilde{\boldsymbol{x}}=\boldsymbol{x}+\boldsymbol{w}_e, \tag{8}$$

where $\boldsymbol{x}\in\mathbb{R}^K$ obeys the distribution in (4) and $\boldsymbol{w}_e\in\mathbb{R}^K$ is the additive Gaussian noise with zero mean and variance $\boldsymbol{\Sigma}_e$. The minimum mean squared error (MMSE) estimator is given by

$$\hat{\boldsymbol{x}}=\mathrm{E}\left[\boldsymbol{x}|\tilde{\boldsymbol{x}}\right]=:\boldsymbol{\eta}\left(\tilde{\boldsymbol{x}}\right). \tag{9}$$

Algorithm 1 AMP based Joint Reconstruction

Imput: $\boldsymbol{y}$, $\boldsymbol{A}$, $\boldsymbol{\Sigma}_x$ and σ_w^2.
Initialization:$\boldsymbol{r}^0 = \boldsymbol{y}$, $\boldsymbol{x}^0 = \boldsymbol{0}$, $\boldsymbol{\Sigma}^0_{\eta,:,i} = \boldsymbol{\Sigma}_x$
Iteration: In each iteration t, do following until meet the stop criteria.
1.For the system model, $\tilde{\boldsymbol{x}}^t = \boldsymbol{x}^t + \boldsymbol{A}^T\boldsymbol{r}^t$.
2.For $i \in [n]$, calculate $\boldsymbol{\Sigma}^t_{e,:,i} = \boldsymbol{\Sigma}^t_{w_x,:,i} + \sigma_w^2\boldsymbol{I}$ where

$$\boldsymbol{\Sigma}^t_{w_x,:,i} = \begin{cases} \frac{1}{\delta}\boldsymbol{\Sigma}^t_{\eta,:,i} & \text{for MMV case,} \\ \frac{1}{\delta}\boldsymbol{D}\left(\boldsymbol{\Sigma}^t_{\eta,:,i}\right) & \text{for DS case,} \end{cases}$$

3.For $i \in [n]$, update $\boldsymbol{x}^{t+1}_{:,i} = \boldsymbol{\eta}\left(\tilde{\boldsymbol{x}}^t_{:,i}\right)$.
4.For $i \in [n]$, update $\boldsymbol{\Sigma}^{t+1}_{\eta,:,i} = \Psi\left(\boldsymbol{\Sigma}^t_{e,:,i}\right)$.
5.For the system model. $\boldsymbol{r}^t = \boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}^t + \frac{1}{\delta}\boldsymbol{r}^{t-1}\left\langle \boldsymbol{\eta}'\left(\boldsymbol{A}^T\boldsymbol{r}^{t-1} + \boldsymbol{x}^{t-1}\right)\right\rangle$

For later use, we also define $\boldsymbol{\eta}'\left(\tilde{\boldsymbol{x}}\right) \in \mathbb{R}^K$ of which the i-th component is given by

$$\eta_i'\left(\tilde{\boldsymbol{x}}\right) = \frac{d\hat{x}_i}{d\tilde{x}_i}. \tag{10}$$

The covariance matrix of the MMSE estimate error is given by

$$\boldsymbol{\Sigma}_\eta = \boldsymbol{\Psi}\left(\boldsymbol{\Sigma}_e\right) := \mathrm{E}\left[\boldsymbol{x}\boldsymbol{x}^T|\tilde{\boldsymbol{x}}\right] - \boldsymbol{\eta}\left(\tilde{\boldsymbol{x}}\right)\boldsymbol{\eta}^T\left(\tilde{\boldsymbol{x}}\right). \tag{11}$$

The specific formulas to compute $\boldsymbol{\eta}$, $\boldsymbol{\eta}'$, and $\boldsymbol{\Psi}$ can be derived using Bayes' rule. See, e.g., [15]. With these definitions, the AMP algorithms for MMV and DS have the same form in (5) and (6) except that $\boldsymbol{\eta}$ and $\boldsymbol{\eta}'$ are specifically defined in (9) and (10) and they are acting on the subvector of the input vector, e.g., $\left(\boldsymbol{A}^T\boldsymbol{r}^t + \boldsymbol{x}^t\right)_{:,i}$, $i \in [n]$.

We now focus on the input of $\boldsymbol{\eta}\left(\cdot\right)$. Similar to (7), we obtain $\tilde{\boldsymbol{x}}^t = \boldsymbol{x}_0 + \boldsymbol{w}_x^t + \boldsymbol{A}^T\boldsymbol{w}$ where $\boldsymbol{w}_x^t = \left(\boldsymbol{A}^T\boldsymbol{A} - \boldsymbol{I}\right)\left(\boldsymbol{x}_0 - \boldsymbol{x}^t\right)$. Following the same analysis as in [13], $\boldsymbol{w}_x^t$ is approximately Gaussian and $\boldsymbol{A}^T\boldsymbol{w}$ are Gaussian conditional on $\boldsymbol{w}$ both with zero mean. The covariance matrix of $\boldsymbol{A}^T\boldsymbol{w}$ can be easily computed which is $\sigma_w^2\boldsymbol{I}$. The nontrivial part is to compute the covariance matrix of $\boldsymbol{w}_x^t$. Note that different from (7) where the matrix $\boldsymbol{A}^T\boldsymbol{A} - \boldsymbol{I}$ is dense, the matrix $\boldsymbol{A}^T\boldsymbol{A} - \boldsymbol{I}$ for MMV and DS is block diagonal and its k-th block is given by $\boldsymbol{A}_k^T\boldsymbol{A}_k - \boldsymbol{I}$, $k \in [K]$. The entries of a given block $\boldsymbol{A}_k^T\boldsymbol{A}_k - \boldsymbol{I}$ can be viewed as independent Gaussian random variables with distribution $\mathcal{N}\left(0, \frac{1}{m}\right)$. However, it is clear that in MMV the entries of $\boldsymbol{A}_k^T\boldsymbol{A}_k - \boldsymbol{I}$ for different k's are identical whereas in DS they are independent across k. For convenience, define $\tilde{\boldsymbol{A}} := \boldsymbol{A}^T\boldsymbol{A} - \boldsymbol{I}$, let $\tilde{\boldsymbol{A}}_k$ be the k-th block of $\tilde{\boldsymbol{A}}$, and $\tilde{A}_k\left(i,j\right)$ be the (i,j)-th component of $\tilde{\boldsymbol{A}}_k$, $i,j \in [n]$. Then $w^t_{x,k,i} = \sum_j \tilde{A}_k\left(i,j\right)\left(x_{0,k,i} - x^t_{k,i}\right)$. Let $\boldsymbol{w}^t_{x,:,i} = \left[\boldsymbol{w}^t_{x,1,i}, \cdots, \boldsymbol{w}^t_{x,K,i}\right]^T$ and $\boldsymbol{\Sigma}^t_{\eta,:,i}$ be the covariance matrix of $\left(\boldsymbol{x}_0 - \boldsymbol{x}^t\right)_{:,i}$. Then the covariance matrix of $\boldsymbol{w}^t_{x,:,i}$ is given by

$$\boldsymbol{\Sigma}^t_{w_x,:,i} = \begin{cases} \frac{1}{\delta}\boldsymbol{\Sigma}^t_{\eta,:,i} & \text{for MMV case,} \\ \frac{1}{\delta}\boldsymbol{D}\left(\boldsymbol{\Sigma}^t_{\eta,:,i}\right) & \text{for DS case,} \end{cases} \tag{12}$$

Algorithm 2 The State Evolution of Joint Reconstruction

Imput: $\boldsymbol{\Sigma}_x$ and σ_w^2.
Initialization: $\boldsymbol{\Sigma}^0_{\eta,:,i} = \boldsymbol{\Sigma}_x$, $\forall i \in [n]$.
Iteration: In each iteration t, do following until meet the stop criteria.
1. Calculate $\boldsymbol{\Sigma}^t_{e,:,i} = \boldsymbol{\Sigma}^t_{w_x,:,i} + \sigma_w^2\boldsymbol{I}$ where

$$\boldsymbol{\Sigma}^t_{w_x,:,i} = \begin{cases} \frac{1}{\delta}\boldsymbol{\Sigma}^t_{\eta,:,i} & \text{for MMV case,} \\ \frac{1}{\delta}\boldsymbol{D}\left(\boldsymbol{\Sigma}^t_{\eta,:,i}\right) & \text{for DS case,} \end{cases}$$

3.Compute $\boldsymbol{\Sigma}^{t+1}_{\eta,:,i}$ based on $\Psi\left(\Sigma^t_{e,:,i}\right)$.

where $\boldsymbol{D}\left(\cdot\right)$ maps an $K \times K$ matrix to a diagonal $K \times K$ matrix by retaining diagonal entries and setting off-diagonal entries to zero. Furthermore, $\boldsymbol{w}^t_{x,:,i}$'s are independent across $i \in [n]$. With this, it can be concluded that $\tilde{\boldsymbol{x}}^t = \boldsymbol{x}_0 + \boldsymbol{w}_e^t$ where $\boldsymbol{w}^t_{e,:,i} \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{\Sigma}^t_{w_x,:,i} + \sigma_w^2\boldsymbol{I}\right)$ and $\boldsymbol{w}^t_{e,:,i}$'s are independent.

By now, we can apply the model in (8) and track the covariance matrix $\boldsymbol{\Sigma}^t_\eta$ across iterations by using (11). The performance of MMV and DS can be obtained by finding the covariance matrix at steady state, i.e., $\boldsymbol{\Sigma}^\infty_\eta$.

It is noteworthy that the performance difference between MMV and DS comes from (12). It is straightforward to verify that the performance of MMV and DS will be the same when $\boldsymbol{\Sigma}_x$ in (4) is diagonal; performance difference only appears when $\boldsymbol{\Sigma}_x$ is not diagonal. See the whole algorithm in Algorithm 1 and state evolution in Algorithm 2.

4. CASE STUDY

This section quantifies the asymptotic performance when $\boldsymbol{x}$ is drawn from the Bernoulli-Gaussian model (4). Simulation results are provided to demonstrate that the asymptotic performance quantification is accurate for finite dimensional systems. Furthermore, when $\epsilon = 1$, simple closed-form formulas are derived for performance quantification for some special cases. These formulas coincide with those obtained from random matrix theory, which once again verifies the correctness of the analysis in this paper.

With given prior distribution (4), the MMSE estimator $\boldsymbol{\eta}\left(\cdot\right)$ in (9) is given by

$$\begin{aligned} \mathrm{E}\left[\boldsymbol{x}|\tilde{\boldsymbol{x}}\right] &= \int \boldsymbol{x}p\left(\boldsymbol{x}|\tilde{\boldsymbol{x}}\right)d\boldsymbol{x} \\ &= \frac{\boldsymbol{R}\tilde{\boldsymbol{x}}}{Cexp\left(-\frac{1}{2}\tilde{\boldsymbol{x}}^T\boldsymbol{\Sigma}_e^{-1}\boldsymbol{R}\tilde{\boldsymbol{x}}\right)+1}, \qquad (13) \\ \boldsymbol{R} &= \left(\boldsymbol{\Sigma}_x^{-1} + \boldsymbol{\Sigma}_e^{-1}\right)^{-1}\boldsymbol{\Sigma}_e^{-1}, \qquad (14) \\ C &= \frac{(1-\epsilon)}{\epsilon}\left|\boldsymbol{\Sigma}_x\left(\boldsymbol{\Sigma}_x^{-1} + \boldsymbol{\Sigma}_e^{-1}\right)\right|^{\frac{1}{2}}. \qquad (15) \end{aligned}$$

The covariance matrix defined in (12) can be computed via

$$\boldsymbol{\Sigma}_\eta\left(p,q\right) = \epsilon\boldsymbol{\Sigma}_x\left(p,q\right) - \boldsymbol{\Sigma}_z\left(p,q\right), \tag{16}$$

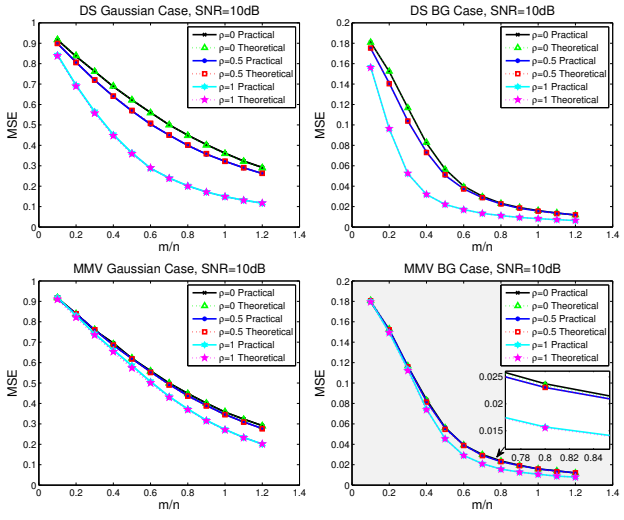

Fig. 1. Compare DS and MMV performances

where $\mathbf{\Sigma}_z(p,q) = \mathrm{E}_{\tilde{\boldsymbol{x}}}\left[\mathrm{E}_x\left[x_p|\tilde{\boldsymbol{x}}\right]\mathrm{E}_x\left[x_q|\tilde{\boldsymbol{x}}\right]\right]$, and

$$\mathbf{\Sigma}_z(p,q) = \int \frac{[\boldsymbol{R}\tilde{\boldsymbol{x}}]_p [\boldsymbol{R}\tilde{\boldsymbol{x}}]_q \epsilon p_G(\tilde{\boldsymbol{x}};\mathbf{0},\mathbf{\Sigma}_x+\mathbf{\Sigma}_e)}{Cexp\left(-\frac{1}{2}\tilde{\boldsymbol{x}}^T\mathbf{\Sigma}_e^{-1}\boldsymbol{R}\tilde{\boldsymbol{x}}\right)+1}\mathrm{d}\tilde{\boldsymbol{x}}, \quad (17)$$

where $[\boldsymbol{R}\tilde{\boldsymbol{x}}]_p$ represents the single element of vector $\boldsymbol{R}\tilde{\boldsymbol{x}}$ at position p. The $\Psi(\mathbf{\Sigma}_e)$ function calculate the matrix element value separately.

Fig. 1 compares the theoretical performance quantification and the empirical performance. We assume $K=2$ measurement instances. We set $\epsilon = 0.2$ for the Bernoulli-Gaussian case and $\epsilon = 1$ for the Gaussian case. The covariance matrix $\mathbf{\Sigma}_x \in \mathbb{R}^{2\times 2}$ has unit diagonal entries and its off-diagonal entries have the same value $\rho \in [0,1]$. For empirical study, the dimension of the signal $\boldsymbol{x}_k$ is set by $n = 1000$. The numerical results are obtained from the average of 100 trials. The results in Fig. 1 clearly demonstrate the accuracy of the asymptotic analysis.

4.1. Gaussian prior: $\epsilon = 1$

When $\epsilon = 1$, the MMSE estimator and the associated covariance matrix of estimation error have nice closed forms:

$$\eta(\tilde{\boldsymbol{x}}) = \mathrm{E}\left[\boldsymbol{x}|\tilde{\boldsymbol{x}}\right] = \boldsymbol{R}\tilde{\boldsymbol{x}}, \quad (18)$$

$$\Psi(\mathbf{\Sigma}_e) = \mathbf{\Sigma}_\eta = \left(\mathbf{\Sigma}_x^{-1}+\mathbf{\Sigma}_e^{-1}\right)^{-1}. \quad (19)$$

When the signal covariance matrix $\mathbf{\Sigma}_x$ are of special forms, i.e., either $\mathbf{\Sigma}_x = \sigma_x^2\boldsymbol{I}$ or $\mathbf{\Sigma}_x = \sigma_x^2\mathbf{1}$ where all the entries of the matrix $\mathbf{1}\in\mathbb{R}^{K\times K}$ are one, the state evolution admits simple closed forms. Define the average recovery distortion at the steady state by $d^\infty = \frac{1}{Kn}\left\|\boldsymbol{x}-\boldsymbol{x}^\infty\right\|_2^2$. It can be verified, via steady state analysis, that when $\mathbf{\Sigma}_x = \sigma_x^2\boldsymbol{I}$ (independent signals),

$$d^\infty_{MMV} = d^\infty_{DS} = \frac{\delta}{2}\left(\left(\frac{1-\delta}{\delta}\sigma_x^2-\sigma_w^2\right)+\sqrt{\left(\frac{1-\delta}{\delta}\sigma_x^2+\sigma_w^2\right)^2+4\sigma_x^2\sigma_w^2}\right), \quad (20)$$

and when $\mathbf{\Sigma}_x = \sigma_x^2\mathbf{1}$ (repeated signals: $\boldsymbol{x}_1 = \cdots = \boldsymbol{x}_K$),

$$d^\infty_{MMV} = \frac{\delta}{2}\left(\left(\frac{1-\delta}{\delta}\sigma_x^2-\frac{\sigma_w^2}{K}\right)+\sqrt{\left(\frac{1-\delta}{\delta}\sigma_x^2+\frac{\sigma_w^2}{K}\right)^2+\frac{4\sigma_w^2\sigma_x^2}{K}}\right), \quad (21)$$

$$d^\infty_{DS} = \frac{\delta}{2}\left(\left(\frac{1-K\delta}{\delta}\sigma_x^2-\sigma_w^2\right)+\sqrt{\left(\frac{1-K\delta}{\delta}\sigma_x^2+\sigma_w^2\right)^2+4K\sigma_x^2\sigma_w^2}\right). \quad (22)$$

It is interesting to observe that the same results can be obtained by solely applying random matrix theory. Specifically, consider a linear system $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}+\boldsymbol{w}$ where $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ is Gaussian random matrix with i.i.d. entries drawn from $\mathcal{N}\left(0,\frac{1}{m}\right)$, $\boldsymbol{x}\in\mathbb{R}^n$ is the signal drawn from $\mathcal{N}\left(\mathbf{0},\sigma_x^2\boldsymbol{I}\right)$, and $\boldsymbol{w}\in\mathbb{R}^m$ is the noise drawn from $\mathcal{N}\left(\mathbf{0},\sigma_w^2\boldsymbol{I}\right)$. Let $(m,n)\to\infty$ simultaneously with $\frac{m}{n}\to\delta$. The empirical distribution of the eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$ converges the Marchenko-Pastur distribution weakly. Based on this fact, the average distortion of MMSE estimate, i.e., $\frac{1}{n}\left\|\boldsymbol{x}-\hat{\boldsymbol{x}}_{MMSE}\right\|_2^2$, can be computed as [13, 17, 18]

$$f(\delta,\sigma_x^2,\sigma_w^2) = \frac{1}{n}\mathrm{tr}\left(\mathbf{\Sigma}_{x|y}\right) = \frac{1}{n}\mathrm{tr}\left(\left(\sigma_x^{-2}\boldsymbol{I}+\sigma_w^{-2}\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\right)$$
$$= \frac{\delta}{2}\left(\left(\frac{1-\delta}{\delta}\sigma_x^2-\sigma_w^2\right)+\sqrt{\left(\frac{1-\delta}{\delta}\sigma_x^2+\sigma_w^2\right)^2+4\sigma_x^2\sigma_w^2}\right), \quad (23)$$

where $\mathrm{tr}(\boldsymbol{M})$ calculates the trace of matrix $\boldsymbol{M}$. In the following, this result is used to analyze MMV and DS models.

MMV and DS models with independent signals ($\mathbf{\Sigma}_x = \sigma_x^2\boldsymbol{I}$). The covariance matrix of recovery error turns out to be block diagonal $\mathbf{\Sigma}_{x|y} = \mathrm{diag}\left(...,\sigma_x^{-2}\boldsymbol{I}+\sigma_w^{-2}\boldsymbol{A}_k^T\boldsymbol{A}_k,...\right)^{-1}$ $= \mathrm{diag}\left(...,\left(\sigma_x^{-2}\boldsymbol{I}+\sigma_w^{-2}\boldsymbol{A}_k^T\boldsymbol{A}_k\right)^{-1},...\right) =: \mathrm{diag}\left(...,\mathbf{\Sigma}_{x_k|y_k},...\right)$. To compute the average distortion, apply (23) to each block $\mathbf{\Sigma}_{x_k|y_k}$. The final result is the same as (20).

MMV and DS models with repeated signals ($\mathbf{\Sigma}_x = \sigma_x^2\mathbf{1}$, i.e., $\boldsymbol{x}_1 = \cdots = \boldsymbol{x}_K$). With repeated signals, the system model can be written as $\boldsymbol{y} = \hat{\boldsymbol{A}}\boldsymbol{x}_1+\boldsymbol{w}$ where $\hat{\boldsymbol{A}} = \left[\boldsymbol{A}_1^T,\boldsymbol{A}_2^T,...,\boldsymbol{A}_K^T\right]^T$. In the MMV model, $\boldsymbol{A}_1 = \boldsymbol{A}_2 = \cdots = \boldsymbol{A}_K$. The covariance matrix of recovery error becomes $\mathbf{\Sigma}_{x_1|y} = \left(\sigma_x^{-2}\boldsymbol{I}+\sigma_w^{-2}\hat{\boldsymbol{A}}^T\hat{\boldsymbol{A}}\right)^{-1} = \left(\sigma_x^{-2}\boldsymbol{I}+K\sigma_w^{-2}\boldsymbol{A}_1^T\boldsymbol{A}_1\right)^{-1}$. Apply (23) and replace the σ_w^2 term in (23) with σ_w^2/K. One obtains (21). By contrast, in the DS model, $\frac{1}{\sqrt{K}}\hat{\boldsymbol{A}}$ is a standard Gaussian random matrix with dimension $Km\times n$. Hence, when Equation (23) is applied to $\mathbf{\Sigma}_{x_1|y} = \left(\sigma_x^{-2}\boldsymbol{I}+K\sigma_w^{-2}\frac{1}{K}\hat{\boldsymbol{A}}^T\hat{\boldsymbol{A}}\right)^{-1}$, σ_w^2 and δ in (23) are replaced with σ_w^2/K and $K\delta$ respectively. The average distortion is the same as (22).

5. REFERENCES

[1] M.F. Duarte, S. Sarvotham, D. Baron, M.B. Wakin, and R.G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," *Asilomar Conference on Signals, Systems and Computers*, pp. 1537–1541, Oct, 2005.

[2] D. Sundman, S. Chatterjee, and M. Skoglund, "Distributed greedy pursuit algorithms," *Signal Processing*, vol. 105, pp. 298–315, 2014.

[3] A.Y. Yang, M. Gastpar, R. Bajcsy, and S.S. Sastry, "Distributed sensor perception via sparse representation," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1077–1088, Jun, 2010.

[4] G. Tzagkarakis, D. Milioris, and P. Tsakalides, "Multiple-measurement Bayesian compressed sensing using GSM priors for DOA estimation," *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp. 2610–2613, Mar, 2010.

[5] S.S. Vasanawala, M.J. Murphy, M.T. Alley, P. Lai, K. Keutzer, J.M. Pauly, and M. Lustig, "Practical parallel imaging compressed sensing MRI: Summary of two years of experience in accelerating body MRI of pediatric patients," *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pp. 1039–1043, Mar, 2011.

[6] B.D. Rao, K. Engan, and S. Cotter, "Diversity measure minimization based method for computing sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, pp. ii–369–72 vol.2, May, 2004.

[7] K. Lee and Y. Bresler, "Subspace-augmented MUSIC for joint sparse recovery with any rank," *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pp. 205–208, Oct, 2010.

[8] W. Zhang, C. Ma, W. Wang, Y. Liu, and L. Zhang, "Side information based orthogonal matching pursuit in distributed compressed sensing," *IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC)*, pp. 80–84, Sept, 2010.

[9] T. Wimalajeewa and P.K. Varshney, "Cooperative sparsity pattern recovery in distributed networks via distributed-omp," pp. 5288–5292, May 2013.

[10] T. Wimalajeewa and P.K. Varshney, "Omp based joint sparsity pattern recovery under communication constraints," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 5059–5072, Oct 2014.

[11] J. Ziniel and P. Schniter, "Efficient high-dimensional inference in the multiple measurement vector problem," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 340–354, Jan, 2013.

[12] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 106, no. 45, pp. 18914–18919, 2009.

[13] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 764–785, Feb, 2011.

[14] J. Kim, W. Chang, B. Jung, D. Baron, and J.C. Ye, "Belief propagation for joint sparse recovery," *arXiv preprint arXiv:1102.3289*, 2011.

[15] X. Zhao and W. Dai, "On joint recovery of sparse signals with common supports," *IEEE International Symposium on Information Theory (ISIT)*, Jun, 2015.

[16] S.M.S. Zobly and Y.M. Kadah, "Multiple measurements vectors compressed sensing for Doppler ultrasound signal reconstruction," *International Conference on Computing, Electrical and Electronics Engineering (ICCEEE)*, pp. 319–322, Aug, 2013.

[17] S. Verdu and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 622–640, Mar, 1999.

[18] D.N.C. Tse and S.V. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 641–657, Mar, 1999.