Coherence-Based Probabilistic Recovery Guarantees for Sparsely Corrupted Signals

Annina Bracher*, Graeme Pope*, and Christoph Studer†

*Dept. of IT & EE, ETH Zurich, 8092 Zurich, Switzerland; e-mail: brachera@student.ethz.ch, gpope@nari.ee.ethz.ch
†Dept. of Electrical and Computer Engineering, Rice University, Houston, TX 77005, USA; e-mail: studer@rice.edu

Abstract—In this paper, we present novel probabilistic recovery guarantees for sparse signals subject to sparse interference, covering varying degrees of knowledge of the signal and interference support. Our results assume that the sparsifying dictionaries are characterized by coherence parameters and we require randomness only in the signal and/or interference. The obtained recovery guarantees show that one can recover sparsely corrupted signals with overwhelming probability, even if the sparsity of both the signal and interference scale (near) linearly with the number of measurements.

I. INTRODUCTION

We consider the problem of recovering the sparse signal vector $\mathbf{x} \in \mathbb{C}^{n_a}$ with support set $\mathcal{X}$ (containing the locations of the non-zero entries of $\mathbf{x}$) from m linear measurements [2]

$$\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{e}. \tag{1}$$

Here, $\mathbf{A} \in \mathbb{C}^{m \times n_a}$ and $\mathbf{B} \in \mathbb{C}^{m \times n_b}$ are given (and known) dictionaries, i.e., matrices that are possibly over-complete and whose columns have unit Euclidean norm. The vector $\mathbf{e} \in \mathbb{C}^{n_b}$ with support set $\mathcal{E}$ represents the sparse interference. We investigate the following models for $\mathbf{x}$ and $\mathbf{e}$, and their support sets $\mathcal{X}$ and $\mathcal{E}$:

- The interference support set $\mathcal{E}$ is *arbitrary*, i.e., $\mathcal{E} \subseteq \{1, \ldots, n_b\}$ can be any subset of cardinality n_e. In particular, $\mathcal{E}$ may depend upon the sparse signal vector $\mathbf{x}$ and/or the dictionary $\mathbf{A}$, and hence, may also be chosen adversarially. The support set $\mathcal{X}$ of $\mathbf{x}$ is chosen uniformly at random from all subsets of $\{1, \ldots, n_a\}$ with cardinality n_x.
- The support set $\mathcal{E}$ of the sparse interference vector $\mathbf{e}$ is chosen uniformly at random from all subsets of $\{1, \ldots, n_b\}$ with cardinality n_e. The support set $\mathcal{X}$ is assumed to be arbitrary and of cardinality n_x.
- Both $\mathcal{X}$ and $\mathcal{E}$, the support sets of the signal and of the interference with cardinality n_x and n_e, respectively, are chosen uniformly at random from all subsets of $\{1, \ldots, n_a\}$ and $\{1, \ldots, n_b\}$ with sizes n_x and n_e, respectively.

In addition, for each model on $\mathcal{X}$ and $\mathcal{E}$ we may or may not know either of the support sets prior to recovery.

As discussed in [2], recovery of $\mathbf{x}$ from the corrupted observation $\mathbf{z}$ in (1) is relevant in a large number of practical applications. In particular, the restoration of saturated signals or signals impaired by impulse noise is captured by the input-output relation (1). Furthermore, (1) enables one to investigate sparsity-based super-resolution and in-painting, as well as signal separation. Hence, identifying the fundamental limits on the recovery of $\mathbf{x}$ (and $\mathbf{e}$ if appropriate) from the sparsely corrupted observation $\mathbf{z}$ is of significant practical interest.

Recovery guarantees for sparsely corrupted signals have been partially studied in [2]–[9]. In particular, [2], [3] investigated coherence-based recovery guarantees for both support sets $\mathcal{X}$ and $\mathcal{E}$ being arbitrary and for varying levels of support-set knowledge; [4] analyzed the special case where both support sets are unknown, but one is chosen arbitrarily and the other at random. The recovery guarantees in [5] require a random matrix $\mathbf{A}$ and that $\mathbf{B}$ is unitary, whereas the results in [6]–[9] characterize $\mathbf{A}$ by the restricted isometry property (RIP), which is difficult to verify in practice. The specific models and assumptions underlying the results in [5]–[9] reduce their utility for the applications outlined above.

A. Contributions

In this paper, we focus exclusively on results where the randomness is in the signal and/or the interference but *not* in the dictionaries $\mathbf{A}$ or $\mathbf{B}$. Furthermore, the dictionaries will be characterized only by their coherence parameters, their dimensions, and their spectral norms. Our results refine or improve upon the recovery guarantees in [2]–[4], [10] and cover novel cases for varying degrees of knowledge of the signal and interference support sets. In particular, we present novel recovery guarantees for the situations where the support sets $\mathcal{X}$ and/or $\mathcal{E}$ are chosen at random, and for the cases where knowledge of neither, one, or both support sets is available prior to recovery. We furthermore show that ℓ_1-norm minimization is able to recover the vectors $\mathbf{x}$ and $\mathbf{e}$ with overwhelming probability, even if the number of non-zero entries in both scales (near) linearly with the number of measurements.

A summary of all the cases studied in this paper is given in Table I; the theorems highlighted in dark gray indicate

TABLE I
SUMMARY OF THE RECOVERY GUARANTEES FOR SPARSELY CORRUPTED
SIGNALS; THE SETTING WITH $\mathcal{X}$ CHOSEN ARBITRARILY AND $\mathcal{E}$ CHOSEN
UNIFORMLY AT RANDOM FOLLOWS ANALOGOUSLY

	$\mathcal{X}$, $\mathcal{E}$ arbitrary	$\mathcal{X}$ rand., $\mathcal{E}$ arb.	$\mathcal{X}$, $\mathcal{E}$ rand.
$\mathcal{X}$, $\mathcal{E}$ known	Case 1a [2, Thm. 3]	Case 1b **Thm. 1**	Case 1c **Thm. 1**
$\mathcal{E}$ known	Case 2a [2, Thm. 4]	Case 2b **Thm. 2**	Case 2d **Thm. 3**
$\mathcal{X}$ known	*Case 2a* [2, Cor. 6]	Case 2c **Thm. 4**	*Case 2d* Thm. 3
neither known	Case 3a [4, Thms. 2, 3]	Case 3b **Thm. 5, [4, Thm. 6]**	Case 3c **Thm. 6**

novel results, light gray indicates refined results. The cases shown in italics follow by reversing the roles of **x** and **e** in the appropriate symmetric case.

B. Notation

Lowercase and uppercase boldface letters stand for column vectors and matrices, respectively. For the matrix $\mathbf{M}$, we denote its transpose, adjoint, and (Moore–Penrose) pseudo-inverse by $\mathbf{M}^T$, $\mathbf{M}^H$, and $\mathbf{M}^\dagger$, respectively; the spectral norm of $\mathbf{M}$ is $\|\mathbf{M}\|_{2,2}$, and the ith column of $\mathbf{M}$ is $\mathbf{m}_i$. Sets are designated by upper-case calligraphic letters; the cardinality of the set $\mathcal{S}$ is $|\mathcal{S}|$. The support set of $\mathbf{v}$ is given by $\mathrm{supp}(\mathbf{v})$. The matrix $\mathbf{M}_\mathcal{S}$ is obtained from $\mathbf{M}$ by retaining the columns of $\mathbf{M}$ with indices in $\mathcal{S}$; the vector $\mathbf{v}_\mathcal{S}$ is obtained analogously from $\mathbf{v}$. We define $\mathbb{1}[\mu \neq 0]$ to be equal to 1 if $\mu \neq 0$ and 0 otherwise. Throughout the paper, $\mathcal{X} = \mathrm{supp}(\mathbf{x})$ and $\mathcal{E} = \mathrm{supp}(\mathbf{e})$ with cardinality n_x and n_e, respectively. We furthermore define $\mathbf{D} = [\,\mathbf{A}\ \mathbf{B}\,]$, $\mathbf{D}_{\mathcal{X},\mathcal{E}} = [\,\mathbf{A}_\mathcal{X}\ \mathbf{B}_\mathcal{E}\,]$, $\mathbf{s} = [\,\mathbf{x}^T\ \mathbf{e}^T\,]^T$, and $\mathbf{s}_{\mathcal{X},\mathcal{E}} = [\,\mathbf{x}_\mathcal{X}^T\ \mathbf{e}_\mathcal{E}^T\,]^T$. For two functions f and g we write $f \sim g$ to indicate that $f(n)/g(n) \to 1$ as $n \to \infty$, and we say that "f scales with g."

II. RELEVANT PRIOR WORK

We next summarize relevant prior work on sparse signal recovery and sparsely corrupted signals.

A. Coherence-Based Recovery Guarantees

During the last decade, numerous deterministic and probabilistic guarantees for the recovery of sparse signals from linear (and non-adaptive) measurements have been developed, e.g., [10]–[14]. These results give sufficient conditions as to when one can reconstruct the sparse signal vector $\mathbf{x}$ from the (interference-less) observation $\mathbf{y} = \mathbf{A}\mathbf{x}$ by solving

$$(\text{P0}) \quad \underset{\hat{\mathbf{x}}}{\text{minimize}}\ \|\hat{\mathbf{x}}\|_0 \quad \text{subject to } \mathbf{y} = \mathbf{A}\hat{\mathbf{x}},$$

or its convex relaxation, basis pursuit, defined as

$$(\text{BP}) \quad \underset{\hat{\mathbf{x}}}{\text{minimize}}\ \|\hat{\mathbf{x}}\|_1 \quad \text{subject to } \mathbf{y} = \mathbf{A}\hat{\mathbf{x}}.$$

In particular, in [11], [12] it is shown that if $\|\mathbf{x}\|_0 \leqslant n_x$ for some $n_x < (1 + 1/\mu_a)/2$ with the coherence parameter

$$\mu_a = \max_{i,j,i\neq j} |\langle \mathbf{a}_i, \mathbf{a}_j \rangle|, \tag{2}$$

then (P0) and (BP) are guaranteed to perfectly recover the sparse signal vector $\mathbf{x}$. Such coherence-based recovery guarantees are, however, subject to the "square-root bottleneck," so that recovery is only guaranteed for sparsity levels scaling as $n_x \sim \sqrt{m}$ [10]. In order to overcome the square-root bottleneck, one must either resort to a RIP-based analysis, e.g., [13], [14], which typically requires randomness in the dictionary $\mathbf{A}$ to show that a certain RIP holds, or a *probabilistic* analysis that considers randomness only in the vector $\mathbf{x}$. Probabilistic and coherence-based recovery guarantees that overcome the square-root bottleneck have been derived in [10]. The corresponding results, however, do not exploit the structure of the problem (1), i.e., the fact that we are dealing with two dictionaries and that knowledge of $\mathcal{X}$ and/or $\mathcal{E}$ may be available prior to recovery.

B. Recovery Guarantees for Sparsely Corrupted Signals

Guarantees for the recovery of sparsely corrupted signals as modeled by (1) have been developed recently in [2]–[4]. The references [2], [3] consider deterministic and coherence-based results for several cases which arise in different applications: 1) $\mathcal{X} = \mathrm{supp}(\mathbf{x})$ and $\mathcal{E} = \mathrm{supp}(\mathbf{e})$ are known prior to recovery, 2) only one of $\mathcal{X}$ and $\mathcal{E}$ is known, and 3) neither $\mathcal{X}$ nor $\mathcal{E}$ are known. For case 1), the non-zero entries of both the signal and interference vectors can be recovered by [2]

$$\mathbf{s}_{\mathcal{X},\mathcal{E}} = \mathbf{D}_{\mathcal{X},\mathcal{E}}^\dagger \mathbf{z}, \tag{3}$$

if the recovery guarantee in [2, Thm. 3] is satisfied. For case 2), recovery is performed by using modified versions of (P0) and (BP); the associated recovery guarantees can be found in [2, Thm. 4 and Cor. 6]. For case 3), recovery guarantees for the standard (P0) or (BP) algorithms are given in [4, Thms. 2 and 3]. However, all these recovery guarantees suffer from the square-root bottleneck, as they guarantee recovery for *all* signal and *all* interference vectors satisfying the given sparsity constraints. A notable exception for case 3) was discussed in [4, Thm. 6]. There, $\mathbf{e}$ is assumed to be random, but $\mathbf{x}$ is assumed to be arbitrary, which significantly improves upon the corresponding deterministic (and coherence-based) recovery guarantees in [4, Thms. 2 and 3]. To overcome the square-root bottleneck for various degrees of support-set knowledge, we next propose a generalization of the probabilistic signal models developed in [4], [10] for the cases 1), 2), and 3) outlined above.

III. MAIN RESULTS

The recovery guarantees developed next rely upon the models $\mathcal{M}(\text{P0})$ and $\mathcal{M}(\text{BP})$ summarized in Model 1 and Model 2, respectively. In addition to these models, our results require the coherence parameters of the dictionaries $\mathbf{A} \in \mathbb{C}^{m \times n_a}$ and $\mathbf{B} \in \mathbb{C}^{m \times n_b}$, i.e., the coherence μ_a of $\mathbf{A}$ in (2), the coherence μ_b of $\mathbf{B}$ given by $\mu_b = \max_{i,j,i\neq j} |\langle \mathbf{b}_i, \mathbf{b}_j \rangle|$, and the mutual coherence μ_m between $\mathbf{A}$ and $\mathbf{B}$, defined as $\mu_m = \max_{i,j} |\langle \mathbf{a}_i, \mathbf{b}_j \rangle|$. Our main results for the cases highlighted in Table I are detailed next.

Model 1 $\mathcal{M}(\text{P0})$

- Let $\mathbf{x} \in \mathbb{C}^{n_a}$ and $\mathbf{e} \in \mathbb{C}^{n_b}$ have support set $\mathcal{X}$ and $\mathcal{E}$, respectively, of which at least one is chosen uniformly at random from all subsets of $\{1, \ldots, n_a\}$ and $\{1, \ldots, n_b\}$ with sizes n_x and n_e, as appropriate. If a support set is chosen at random, then assume that the corresponding non-zero entries of the associated vector are drawn from a continuous distribution.
- The observation $\mathbf{z}$ is given by $\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{e}$.

Model 2 $\mathcal{M}(\text{BP})$

- The conditions of $\mathcal{M}(\text{P0})$ hold.
- If $\mathcal{X}$ or $\mathcal{E}$ is chosen at random, then assume that the corresponding non-zero entries of the associated vector(s) are drawn from a continuous distribution, where the phases of the individual components are independent and uniformly distributed on $[0, 2\pi)$.

A. Cases 1b and 1c: $\mathcal{X}$ and $\mathcal{E}$ Known

We start with the case where both support sets $\mathcal{X}$ and $\mathcal{E}$ are known prior to recovery. The following theorem guarantees recovery of $\mathbf{x}$ and $\mathbf{e}$ from $\mathbf{z}$ using (3), with high probability.

Theorem 1 (Cases 1b and 1c): Let $\mathbf{x}$ and $\mathbf{e}$ be signals satisfying the conditions of $\mathcal{M}(\text{P0})$, assume that both $\mathcal{X}$ and $\mathcal{E}$ are known, and choose $\beta \geqslant \log(n_x)$. If $\mathcal{X}$ is chosen uniformly at random, $\mathcal{E}$ is arbitrary, and if

$$\delta e^{-\frac{1}{4}} \geqslant \|\mathbf{A}\|_{2,2} \|\mathbf{B}\|_{2,2} \sqrt{\frac{n_x}{n_a}} + 12\mu_a \sqrt{\beta n_x} + (n_e - 1)\mu_b$$
$$+ \mathbb{1}[\mu_a \neq 0]\frac{2n_x}{n_a} \|\mathbf{A}\|_{2,2}^2 + 3\mu_m \sqrt{2\beta n_e} \qquad (4)$$

holds with[1] $\delta = 1$, then we can recover $\mathbf{x}$ and $\mathbf{e}$ using (3) with probability at least $1 - e^{-\beta}$.

If both $\mathcal{X}$ and $\mathcal{E}$ are chosen at random and if

$$\delta e^{-\frac{1}{4}} \geqslant 12\sqrt{\beta}\left(\mu_a\sqrt{n_x} + \mu_b\sqrt{n_e}\right) + \mathbb{1}[\mu_a \neq 0]\frac{2n_x}{n_a}\|\mathbf{A}\|_{2,2}^2$$
$$+ \mathbb{1}[\mu_b \neq 0]\frac{2n_e}{n_b}\|\mathbf{B}\|_{2,2}^2 + \min\left\{3\mu_m\sqrt{2\beta n_x}\right.$$
$$\left.+ \sqrt{\frac{n_e}{n_b}}\left\|\mathbf{A}^H\mathbf{B}\right\|_{2,2}, 3\mu_m\sqrt{2\beta n_e} + \sqrt{\frac{n_x}{n_a}}\left\|\mathbf{A}^H\mathbf{B}\right\|_{2,2}\right\} \quad (5)$$

holds with $\delta = 1$ and $\beta \geqslant \max\{\log(n_x), \log(n_e)\}$, then we can recover $\mathbf{x}$ and $\mathbf{e}$ using (3) with probability at least $1 - e^{-\beta}$.

Proof: The proof can be found in [1]. ∎

B. Cases 2b and 2d: $\mathcal{E}$ Known

Consider the case where only the support set $\mathcal{E}$ of $\mathbf{e}$ is known prior to recovery. In this case, recovery of $\mathbf{x}$ (and the non-zero entries of $\mathbf{e}$) from $\mathbf{z}$ can be achieved by solving[2]

$$(\text{P0}, \mathcal{E}) \quad \begin{cases} \underset{\hat{\mathbf{x}}, \hat{\mathbf{e}}_\mathcal{E}}{\text{minimize}} & \|\hat{\mathbf{x}}\|_0 + \|\hat{\mathbf{e}}_\mathcal{E}\|_0 \\ \text{subject to} & \mathbf{z} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}_\mathcal{E}\hat{\mathbf{e}}_\mathcal{E}, \end{cases}$$

---

[1]Later, we will require (4) to hold for different values of δ.
[2]Note that since $\mathcal{E}$ is known, the norm $\|\hat{\mathbf{e}}_\mathcal{E}\|_0$ in $(\text{P0}, \mathcal{E})$ can be omitted.

---

or its convex relaxation[3]

$$(\text{BP}, \mathcal{E}) \quad \begin{cases} \underset{\hat{\mathbf{x}}, \hat{\mathbf{e}}_\mathcal{E}}{\text{minimize}} & \|\hat{\mathbf{x}}\|_1 + \|\hat{\mathbf{e}}_\mathcal{E}\|_1 \\ \text{subject to} & \mathbf{z} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}_\mathcal{E}\hat{\mathbf{e}}_\mathcal{E}. \end{cases}$$

The following theorems guarantee the recovery of $\mathbf{x}$ and $\mathbf{e}$ from $\mathbf{z}$, using $(\text{P0}, \mathcal{E})$ or $(\text{BP}, \mathcal{E})$, with high probability.

Theorem 2 (Case 2b): Let $\mathbf{x}$ and $\mathbf{e}$ be signals satisfying the conditions of $\mathcal{M}(\text{P0})$, assume that $\mathcal{E}$ is known prior to recovery and chosen arbitrarily, and assume that $\mathcal{X}$ is unknown and drawn uniformly at random. Choose $\beta \geqslant \log(n_x)$. If (4) holds for some $0 < \delta < 1$ and if

$$n_x\mu_a^2 + n_e\mu_m^2 < 1 - \delta, \qquad (6)$$

then we can recover $\mathbf{x}$ and $\mathbf{e}$ using $(\text{P0}, \mathcal{E})$ with probability at least $1 - e^{-\beta}$.

Moreover, if $\mathbf{x}$ and $\mathbf{e}$ are signals satisfying the conditions of $\mathcal{M}(\text{BP})$, and, in addition to (4) if

$$n_x\mu_a^2 + n_e\mu_m^2 < \frac{(1-\delta)^2}{2(\log(n_a) + \beta)} \qquad (7)$$

holds then we can recover $\mathbf{x}$ and $\mathbf{e}$ using $(\text{BP}, \mathcal{E})$ with probability at least $1 - 3e^{-\beta}$.

Proof: The proof can be found in [1]. ∎

Theorem 3 (Case 2d): Let $\mathbf{x}$ and $\mathbf{e}$ be signals satisfying the conditions of $\mathcal{M}(\text{P0})$, assume that $\mathcal{E}$ is known but $\mathcal{X}$ is unknown prior to recovery, and assume that both $\mathcal{X}$ and $\mathcal{E}$ are drawn uniformly at random. If (5) and (6) hold for some $0 < \delta < 1$ and $\beta \geqslant \max\{\log(n_x), \log(n_e)\}$, then we can recover $\mathbf{x}$ and $\mathbf{e}$ using $(\text{P0}, \mathcal{E})$ with probability at least $1 - e^{-\beta}$.

Moreover, if $\mathbf{x}$ and $\mathbf{e}$ satisfy the conditions of $\mathcal{M}(\text{BP})$ and if (7) holds in addition to (5) and (6), then we can recover $\mathbf{x}$ and $\mathbf{e}$ using $(\text{BP}, \mathcal{E})$ with probability at least $1 - 3e^{-\beta}$.

Proof: The proof can be found in [1]. ∎

C. Case 2c: $\mathcal{X}$ Known

The case where $\mathcal{X}$ is random and known, and $\mathcal{E}$ is unknown and arbitrary, differs slightly to the case where $\mathcal{X}$ is random and unknown, and $\mathcal{E}$ is arbitrary and known (covered by Thm. 2). Hence, we need to consider the two cases separately. The recovery problems $(\text{P0}, \mathcal{X})$ and $(\text{BP}, \mathcal{X})$ required here are defined analogously to $(\text{P0}, \mathcal{E})$ and $(\text{BP}, \mathcal{E})$.

Theorem 4 (Case 2c): Let $\mathbf{x}$ and $\mathbf{e}$ be signals satisfying the conditions of $\mathcal{M}(\text{P0})$, assume that $\mathcal{X}$ is known and chosen uniformly at random, and assume that $\mathcal{E}$ is unknown and arbitrary. If

$$\delta e^{-\frac{1}{4}} \geqslant \|\mathbf{A}\|_{2,2}\|\mathbf{B}\|_{2,2}\sqrt{\frac{n_e}{n_b}} + 12\mu_b\sqrt{\beta n_e} + (n_x - 1)\mu_a$$
$$+ \mathbb{1}[\mu_b \neq 0]\frac{2n_e}{n_b}\|\mathbf{B}\|_{2,2}^2 + 3\mu_m\sqrt{2\beta n_x} \qquad (8)$$

holds for some $0 < \delta < 1$ and $\beta \geqslant \log(n_e)$, and if

$$n_x\mu_m^2 + n_e\mu_b^2 < 1 - \delta, \qquad (9)$$

---

[3]Note that we consider a slightly different convex optimization problem to that proposed in [2] for the case where $\mathcal{E}$ is known prior to recovery.

then we can recover $\mathbf{x}$ and $\mathbf{e}$ using $(\mathrm{P0}, \mathcal{X})$ with probability at least $1 - e^{-\beta}$.

Moreover, if $\mathbf{x}$ and $\mathbf{e}$ are signals satisfying the conditions of $\mathcal{M}(\mathrm{BP})$, and, in addition to (8), if

$$n_x \mu_m^2 + n_e \mu_b^2 < \frac{(1-\delta)^2}{2(\log(n_b) + \beta)} \tag{10}$$

holds then we can recover $\mathbf{x}$ and $\mathbf{e}$ using $(\mathrm{BP}, \mathcal{X})$ with probability at least $1 - 3e^{-\beta}$.

Proof: The proof can be found in [1]. ∎

D. Cases 3b and 3c: No Support-Set Knowledge

Recovery guarantees for the case of no support-set knowledge, but where one support set is chosen at random and the other arbitrarily can be found in [4]. The theorem shown next refines the result in [4, Thm. 6]. The refinements are due to the following facts: i) We allow for arbitrary $0 < \delta < 1$, whereas $\delta = 1/2$ in [4, Thm. 6], ii) we do not use a global coherence parameter $\mu = \max\{\mu_a, \mu_b, \mu_m\}$, but rather exploit the individual coherence parameters μ_a, μ_b, and μ_m of $\mathbf{A}$ and $\mathbf{B}$, and iii) the indicator functions in our recovery conditions improve our results in the cases where $\mathbf{A}$ and/or $\mathbf{B}$ are unitary.

Theorem 5 (Case 3b): Let $\mathbf{x}$ and $\mathbf{e}$ be signals satisfying conditions of $\mathcal{M}(\mathrm{P0})$, assume that $\mathcal{X}$ is chosen uniformly at random, and assume that $\mathcal{E}$ is arbitrary. If (4), (6), and (9) hold for some $0 < \delta < 1$ and $\beta \geqslant \log(n_x)$, then

$$(\mathrm{P0}^*) \quad \underset{\hat{\mathbf{x}}, \hat{\mathbf{e}}}{\text{minimize}} \ \|\hat{\mathbf{x}}\|_0 + \|\hat{\mathbf{e}}\|_0 \quad \text{subject to } \mathbf{z} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{e}},$$

recovers $\mathbf{x}$ and $\mathbf{e}$ with probability at least $1 - e^{-\beta}$.

Moreover, if $\mathbf{x}$ and $\mathbf{e}$ satisfy the conditions of $\mathcal{M}(\mathrm{BP})$ and if (7) and (10) hold in addition to (4), (6), and (9), then

$$(\mathrm{BP}^*) \quad \underset{\hat{\mathbf{x}}, \hat{\mathbf{e}}}{\text{minimize}} \ \|\hat{\mathbf{x}}\|_1 + \|\hat{\mathbf{e}}\|_1 \quad \text{subject to } \mathbf{z} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{e}},$$

recovers $\mathbf{x}$ and $\mathbf{e}$ with probability at least $1 - 3e^{-\beta}$.

Proof: The proof can be found in [1]. ∎

The last theorem considers the case where neither support-set is known and both are chosen uniformly at random.

Theorem 6 (Case 3c): Let $\mathbf{x}$ and $\mathbf{e}$ be signals satisfying the conditions of $\mathcal{M}(\mathrm{P0})$ and assume that $\mathcal{X}$ and $\mathcal{E}$ are both unknown and chosen uniformly at random. If (5), (6), and (9) hold for some $0 < \delta < 1$ and $\beta \geqslant \max\{\log(n_x), \log(n_e)\}$, then $(\mathrm{P0}^*)$ recovers $\mathbf{x}$ and $\mathbf{e}$ with probability at least $1 - e^{-\beta}$.

Moreover, if $\mathbf{x}$ and $\mathbf{e}$ are signals from $\mathcal{M}(\mathrm{BP})$ and if (7) and (10) hold in addition to (5), (6), and (9), then (BP^*) recovers $\mathbf{x}$ and $\mathbf{e}$ with probability at least $1 - 3e^{-\beta}$.

Proof: The proof can be found in [1]. ∎

IV. DISCUSSION OF THE RECOVERY GUARANTEES

We now briefly discuss a subset of the theorems presented in Sec. III. A detailed discussion of all recovery guarantees can be found in the journal version of the paper [1].

In what follows, we assume $\mathbf{A}$ and $\mathbf{B}$ are unitary, i.e., $n_a = n_b = m$ and $\mu_a = \mu_b = 0$, and maximally incoherent, i.e., $\mu_m = 1/\sqrt{m}$. For example, $\mathbf{A}$ could be the discrete Fourier transform matrix and $\mathbf{B}$ the identity matrix. We
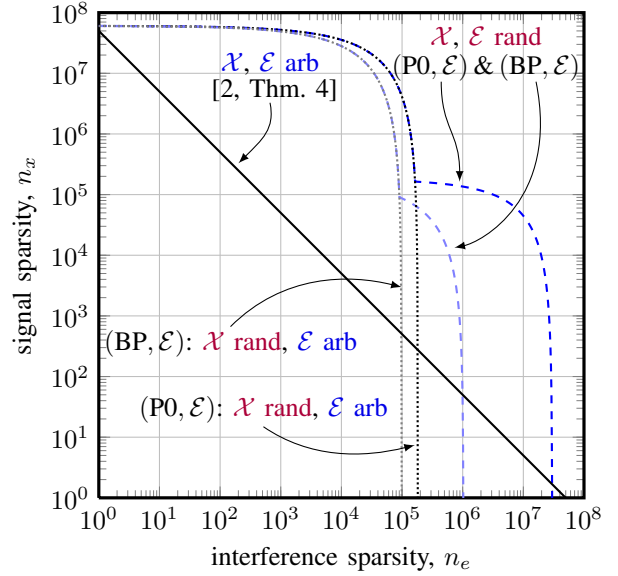

Fig. 1. Comparison of the recovery conditions for the case where only $\mathcal{E}$ is known prior to recovery. $\mathbf{A}$ and $\mathbf{B}$ are assumed to be unitary with $m = n_a = n_b = 10^8$ and $\mu_m = 1/\sqrt{m}$.

furthermore set $\beta = \log(m)$, so that recovery is guaranteed with probability at least $1 - 1/m$ and $1 - 3/m$ when solving the ℓ_0-norm and ℓ_1-norm-based recovery problems, respectively.

A. Recovery Guarantees

1) Only $\mathcal{E}$ known: Fig. 1 shows the recovery conditions from Theorems 2 and 3 for the cases where only $\mathcal{E}$ is known prior to recovery (the case of only $\mathcal{X}$ known behaves analogously). We see that for a random $\mathcal{X}$ and random $\mathcal{E}$ successful recovery at high probability is guaranteed for significantly larger n_x and n_e compared to the case where one or both support sets are arbitrary. Hence, allowing for randomness in the support sets leads to less restrictive recovery conditions.

2) No support-set knowledge: Fig. 2 shows the recovery conditions for (BP^*) for the case of no support-set knowledge. We see that for random $\mathcal{X}$ and $\mathcal{E}$, successful recovery is guaranteed for significantly larger n_x and n_e compared to the case where one or both support sets are assumed to be arbitrary. As a comparison, we also show the recovery conditions derived in [4, Thm. 6] and the condition from [10], which does not take into account the structure of the problem (1). We see that the recovery conditions derived in Theorems 5 and 6 guarantee the successful recovery for a larger number of nonzero coefficients in both the sparse signal vector $\mathbf{x}$ and the sparse interference $\mathbf{e}$.

B. Asymptotic Behavior of the Recovery Conditions

We now compare the asymptotic behavior of probabilistic and deterministic recovery conditions, i.e., we study the scaling behavior of n_x and n_e. To this end, we are interested in the largest n_x for which recovery of $\mathbf{x}$ (and $\mathbf{e}$) from $\mathbf{z}$ can be guaranteed with high probability. In particular, we consider the following models for the sparse interference vector $\mathbf{e}$: i) constant sparsity, i.e., $n_e = 10^3$, ii) sparsity proportional
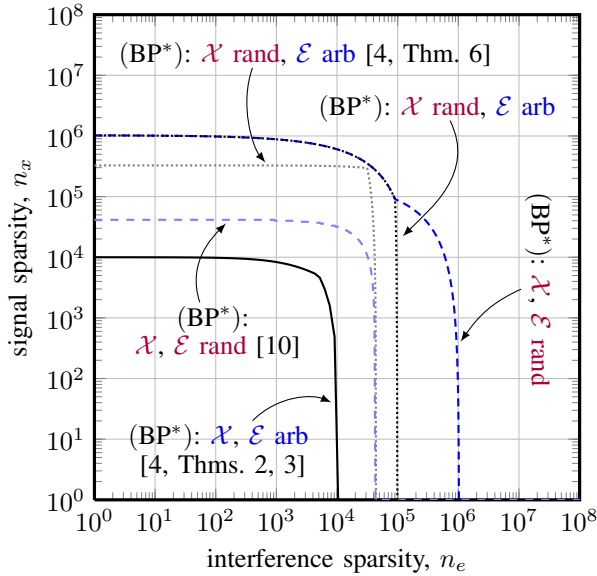
Fig. 2. Comparison of the recovery conditions for the case of no support-set knowledge. $\mathbf{A}$ and $\mathbf{B}$ are assumed to be unitary with $m = n_a = n_b = 10^8$ and $\mu_m = 1/\sqrt{m}$.
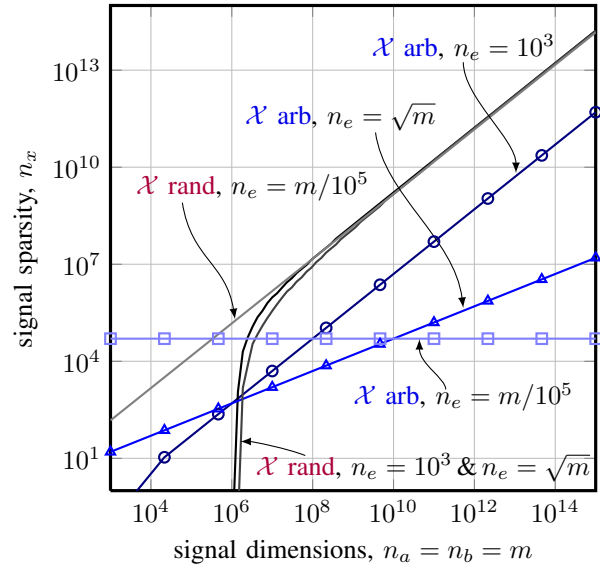

Fig. 3. Maximum signal sparsity n_x that ensures recovery of $\mathbf{x}$ for $\mathcal{E}$ known and arbitrary. We assume $n_e = 10^3$, $n_e = \sqrt{m}$, and $n_e = m/10^5$. The probability of successful recovery is set to be at least $1 - 10^{-15}$.

to the square root of the problem size, i.e., $n_e = \sqrt{m}$, and iii) sparsity proportional to the problem size, i.e., $n_e = m/10^5$.

Fig. 3 shows the largest n_x (for a given n_e) for which recovery can be guaranteed using $(\mathrm{BP}, \mathcal{E})$. Here, $\mathcal{E}$ is assumed to be known and arbitrary and $\mathcal{X}$ is unknown and chosen either at random or arbitrarily. Note that the other cases of support-set knowledge and arbitrary/random exhibit the same scaling behavior. We see from Fig. 3 that for a constant interference sparsity (i.e., $n_e = 10^3$), the probabilistic and deterministic results show the same scaling behavior. For the cases where n_e scales with $\sqrt{m}$ or m, however, the deterministic thresholds developed in [2] result in worse scaling, while the behavior of the probabilistic guarantees derived here remains unaffected.

We now investigate the scaling behavior observed in Fig. 3 analytically. Again, we only consider the case where $\mathcal{X}$ is unknown and chosen at random and $\mathcal{E}$ is known and chosen arbitrarily; an analysis of the other cases yields similar results. From Thm. 2, the recovery of $\mathbf{x}$ from $\mathbf{z}$ using $(\mathrm{BP}, \mathcal{E})$ is guaranteed with probability at least $1 - 3/n_a$ (i.e., for $\beta = \log(n_a)$) for maximally incoherent and unitary $\mathbf{A}$ and $\mathbf{B}$ (i.e., $\mu_a = \mu_b = 0$, $n_a = n_b = m$, and $\mu_m = 1/\sqrt{m}$) if for some $\delta \in (0, 1)$, both $\delta e^{-1/4} \geqslant \sqrt{n_x/n_a} + 3\mu_m\sqrt{2\beta n_e}$ and $2n_e\mu_m^2(\log(n_a) + \beta) < (1 - \delta)^2$ hold, that is, if

$$e^{-\frac{1}{4}}\sqrt{m} > \sqrt{n_x} + (3\sqrt{2} + 2e^{-\frac{1}{4}})\sqrt{n_e\log(m)}. \qquad (11)$$

Hence, if $n_x \sim m$ and $n_e \sim m/\log(m)$, the condition (11) can be satisfied. Consequently, recovery of $\mathbf{x}$ (and of $\mathbf{e}$) is guaranteed with probability at least $1 - 3/m$ even if n_x scales *linearly* in the number of (corrupted) measurements m and n_e scales near-linearly (i.e., with $m/\log(m)$) in m.

We finally note that the recovery guarantees in [5] also allow for the sparsity of the interference vector to scale near-linearly in the number of measurements. The results in [5], however, require the matrix $\mathbf{A}$ to be random and $\mathbf{B}$ to be orthogonal,

whereas the recovery guarantees shown here are for *arbitrary* pairs of dictionaries $\mathbf{A}$ and $\mathbf{B}$ (characterized by the coherence parameters) and for varying degrees of support-set knowledge.

REFERENCES

[1] G. Pope, A. Bracher, and C. Studer, "Probabilistic recovery guarantees for sparsely corrupted signals," *submitted to IEEE Trans. Inf. Theory, arXiv:1203.6001v1*, Mar. 2012.
[2] C. Studer, P. Kuppinger, G. Pope, and H. Bölcskei, "Recovery of sparsely corrupted signals," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3115–3130, May 2012.
[3] C. Studer and R. G. Baraniuk, "Stable restoration and separation of approximately sparse signals," *submitted, arXiv:1107.0420v1*, July 2011.
[4] P. Kuppinger, G. Durisi, and H. Bölcskei, "Uncertainty relations and sparse signal recovery for pairs of general signal sets," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 263–277, Jan. 2012.
[5] X. Li, "Compressed sensing and matrix completion with constant proportion of corruptions," *arXiv:1104.1041v2*, Jan. 2012.
[6] J. N. Laska, P. T. Boufounos, M. A. Davenport, and R. G. Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing," *App. Comp. Harm. Anal.*, vol. 31, no. 3, pp. 429–443, Nov. 2011.
[7] J. N. Laska, M. A. Davenport, and R. G. Baraniuk, "Exact signal recovery from sparsely corrupted measurements through the pursuit of justice," *Proc. of 43rd Asilomar Conf. on Signals, Systems, and Comput.*, pp. 1556–1560, Nov. 2009.
[8] N. Vaswani and W. Lu, "Modified-CS: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Sig. Proc.*, vol. 58, no. 9, pp. 4595–4607, Sep. 2010.
[9] L. Jacques, "A short note on compressed sensing with partially known signal support," *Sig. Proc.*, vol. 90, no. 12, pp. 3308–3312, Dec. 2010.
[10] J. A. Tropp, "On the conditioning of random subdictionaries," *App. Comp. Harm. Anal.*, vol. 25, pp. 1–24, July 2008.
[11] ——, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
[12] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.
[13] E. J. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure and Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.
[14] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *C. R. Acad. Sci. Paris*, vol. 346, no. 9-10, pp. 589–592, 2008.