Perturbed Orthogonal Matching Pursuit

Oguzhan Teke, *Student Member, IEEE*, Ali Cafer Gurbuz, *Member, IEEE*, and Orhan Arikan, *Member, IEEE*

Abstract—Compressive Sensing theory details how a sparsely represented signal in a known basis can be reconstructed with an underdetermined linear measurement model. However, in reality there is a mismatch between the assumed and the actual bases due to factors such as discretization of the parameter space defining basis components, sampling jitter in A/D conversion, and model errors. Due to this mismatch, a signal may not be sparse in the assumed basis, which causes significant performance degradation in sparse reconstruction algorithms. To eliminate the mismatch problem, this paper presents a novel perturbed orthogonal matching pursuit (POMP) algorithm that performs controlled perturbation of selected support vectors to decrease the orthogonal residual at each iteration. Based on detailed mathematical analysis, conditions for successful reconstruction are derived. Simulations show that robust results with much smaller reconstruction errors in the case of perturbed bases can be obtained as compared to standard sparse reconstruction techniques.

Index Terms—Compressive sensing, basis perturbation, basis mismatch, perturbed OMP.

I. INTRODUCTION

SPARSE signal representations and the compressive sensing (CS) theory [1], [2] has received considerable attention in recent years in many research communities. In particular, CS changed the way data is acquired by significantly reducing the data acquisition number or cost, which has been applied to a wide range of important applications, such as computational photography [3], medical imaging [4], radar [5], [6] and sensor networks [7].

Compressive sensing states that a sparse signal in some known basis can be efficiently acquired using a small set of nonadaptive and linear measurements. Suppose N dimensional signal $\boldsymbol{x}$ has a K-sparse representation in a transform domain $\boldsymbol{\Psi}$, as $\boldsymbol{x} = \boldsymbol{\Psi s}$ and $\|\boldsymbol{s}\|_0 = K$. Given linear measurements in the form $\boldsymbol{y} = \boldsymbol{\Phi x}$, by using compressive sensing techniques, the sparse signal $\boldsymbol{s}$, hence $\boldsymbol{x}$, can be recovered exactly with very

high probability from $O(K \log N)$ measurements by solving a convex ℓ_1 optimization problem of the following form:

$$\min \|\boldsymbol{s}\|_1, \qquad \text{subject to} \quad \boldsymbol{y} = \boldsymbol{\Phi\Psi s}, \qquad (1)$$

which can be solved efficiently using linear programming. Stable reconstruction methods for noisy measurements or compressible signals based on ℓ_1 minimization have been developed [8]–[10] for the known basis case. Suboptimal greedy algorithms have also been used in many applications. Matching pursuit (MP) [11], orthogonal matching pursuit (OMP) [12], compressive sampling matching pursuit (CoSaMP) [13], iterative hard/soft thresholding (IHT) [14] are among the most commonly used greedy algorithms. Apart from greedy algorithms, approximate message passing (AMP) uses the idea of belief propagation to achieve high reconstruction performance with low complexity [15]. If the sparse signal $\boldsymbol{s}$ has a structure, such as a wavelet tree, techniques proposed in [16] can exploit those models for better reconstruction. The study in [17] assumes a Markov-tree structure in the sparse coefficients and adapts the AMP algorithm in a Bayesian framework.

Commonly used sparse reconstruction techniques assume that the basis $\boldsymbol{\Psi}$ is exactly known and the signal is sparse in that basis. However, in some applications there is a mismatch between the assumed basis and the actual but unknown one. For example in applications like target localization [18], radar [19], [20], time delay and doppler estimation, beamforming [21], [22] or shape detection [23], the sparsity of the signal is in a continuous parameter space and the sparsity basis $\boldsymbol{\Psi}$ is constructed through discretization or gridding of these parameter spaces. In general, a signal will not be sparse in such a dictionary created through discretization, since no matter how fine the grid dimensions are, the signal parameters may not, and generally do not, lie in the center of the grid cells. As a simple example; consider a general signal which is sparse in the continuous frequency domain. This signal may not be sparse in the DFT basis defined by the frequency grid. A continuous frequency parameter lying between two successive DFT grid cells will affect not the only the closest two cells, but the whole grid with amplitude decaying with $1/T$, where T is the sampling time interval. This off-grid phenomena violates the sparsity assumption, resulting in a decrease in reconstruction performance. In addition to these structured perturbations, random time jitter in A/D conversion, modeling errors in construction of the dictionary $\boldsymbol{\Psi}$ create perturbations on the dictionary columns. Hence, in general, the signal $\boldsymbol{x}$ will be sparse in an unknown basis $\hat{\boldsymbol{\Psi}} = \boldsymbol{\Psi} + \boldsymbol{P}$ where $\boldsymbol{\Psi}$ is the adopted basis and $\boldsymbol{P}$ is the unknown perturbation matrix.

In the literature, the effect of this basis mismatch has been observed and analyzed in some applications such as radar [19], [24] and beamforming [25]. In problems due to parameter space discretization, a simplistic approach is to use multi-resolution refinement and decrease the grid size. Decreasing the grid size is

not a direct solution to the basis mismatch problem, because it increases the coherence between dictionary columns, violating the restricted isometry property (RIP) [26] and increasing the computational complexity of the reconstruction. In [27]–[29] the effect of the basis mismatch problem on the reconstruction performance of CS has been analyzed and the resultant performance degradation levels and analytical ℓ_2 norm error bounds due to the basis mismatch have been investigated. However, these works do not offer a systematic approach for sparse reconstruction under random perturbation models. In [30], the dictionary is extended to several dictionaries and solution is pursued not in a single orthogonal basis, but in a set of bases using a tree structure, assuming that the given signal is sparse in at least one of the basis. However, this strategy does not provide solutions if the signal is not-sparse in the extended dictionary. In the Continuous Basis Pursuit approach [31], perturbations are assumed to be continuously shifted features of the functions on which the sparse solution is searched for, and ℓ_1 based minimization is proposed. Also in this method, perturbations are assumed to have structure, and are modeled with a first order Taylor approximation or polar interpolators. In [32], ℓ_1 minimization based algorithms are proposed for linear structured perturbations on the sensing matrix. In [33] a total least square (TLS) solution is proposed for the problem, in which an optimization over all signals $\boldsymbol{x}$, perturbation matrix $\boldsymbol{P}$ and error vector spaces should be solved. To reduce complexity, suboptimal optimization techniques have been pursued in [33].

In this paper, a novel perturbed orthogonal matching pursuit (POMP) algorithm is presented. In the standard OMP algorithm [12] the column vector that has the largest correlation with the current residual is selected and the new residual is calculated by projecting the measurements onto the subspace defined by the span of all selected columns. This procedure is repeated until the termination criteria is met. In the proposed POMP algorithm, controlled perturbation mechanism is applied on the selected columns. The selected column vectors are perturbed in directions that decrease the orthogonal residual at each iteration. Proven limits on perturbations are obtained. The proposed method is fast, simple to implement and successful in recovering sparse signals under random basis perturbations. A preliminary form of this approach has been presented in [34].

The organization of the paper is as follows. Section II outlines the formulation of the problem and details the development of the proposed POMP algorithm with a pseudo code and supporting theorems. Results covering performance comparisons of the proposed method and applications on a set of examples are given in Section III. In Section IV, concluding remarks are presented and detailed proof of a theorem is given in the Appendix.

II. PERTURBED ORTHOGONAL MATCHING PURSUIT

In compressive sensing, a sparse signal $\boldsymbol{x}$ is reconstructed from its given underdetermined linear measurements:

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n}, \qquad (2)$$

where $\boldsymbol{A} = [\boldsymbol{a}_1 \boldsymbol{a}_2 \dots \boldsymbol{a}_N] \in \Re^{M \times N}$ with $M < N$ and $\boldsymbol{n}$ is additive noise, typically assumed to be independent and identically distributed (i.i.d.) Gaussian noise with known variance σ^2. The

matrix $\boldsymbol{A}$ may be known, as in the case of row decimated versions of Fourier or wavelet transformation matrices, or it might be a randomly selected one. The ideal reconstruction problem can be formulated as an ℓ_0 minimization problem,

$$\min \|\boldsymbol{x}\|_0 \qquad \text{s.t.} \qquad \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2 < \epsilon, \qquad (3)$$

where $\|\boldsymbol{x}\|_0$ is the cardinality of $\boldsymbol{x}$, and the fit error constraint has a bound ϵ that is related to the measurement noise statistics. However, solution of the ℓ_0 minimization problem requires a combinatorial search, thus it is not feasible for practical applications.

Another challenge is reconstruction in the presence of errors in $\boldsymbol{A}$. When $\boldsymbol{A}$ is not known precisely, $\boldsymbol{b}$ may not be sparse in the assumed $\boldsymbol{A}$. However, sparsity of the signal can be revealed under a certain perturbation, $\delta \boldsymbol{A}$, on the given $\boldsymbol{A}$. In this case, the ℓ_0 problem can be recast as,

$$\min \|\boldsymbol{x}\|_0 \qquad \text{s.t.} \qquad \min_{\delta \boldsymbol{A} \in \Delta} \|\boldsymbol{b} - (\boldsymbol{A} + \delta \boldsymbol{A})\,\boldsymbol{x}\|_2 < \epsilon, \quad (4)$$

where Δ is some bounded perturbation space. This problem can be viewed as a generalized version of the problem given in [31], in which $\delta \boldsymbol{A}$ is considered as a Taylor series or polar approximation of $\boldsymbol{A}$ and Δ is the sufficient limits, also relaxing ℓ_0 with ℓ_1. The solution to this general ℓ_0 problem is also combinatoric in nature, and infeasible in practice.

To reduce the complexity of the ℓ_0 problem of (4), sub-optimal greedy techniques can be developed as well. In these greedy approaches, the support set of the reconstruction is iteratively increased until the constraints are satisfied. Assuming that at iteration k, the support set contains k columns of $\boldsymbol{A}$, which we will call as $\boldsymbol{S}_k = \{\boldsymbol{s}_1, \boldsymbol{s}_2, \dots, \boldsymbol{s}_k\}$ for simplicity. At the $(k+1)$th iteration, new vector $\boldsymbol{s}_{k+1}$ is obtained from the solution of the following optimization problem:

$$\boldsymbol{s}_{k+1} = \arg \max_{\boldsymbol{u} \in \boldsymbol{S}_k'} \left\{ \max_{\delta \boldsymbol{S}_k, \delta \boldsymbol{u}} \left\| \boldsymbol{P}_{([\boldsymbol{S}_k \boldsymbol{u}] + [\delta \boldsymbol{S}_k \delta \boldsymbol{u}])}\, \boldsymbol{b} \right\|_2 \right.$$
$$\left. \text{s.t.} \quad [\delta \boldsymbol{S}_k \delta \boldsymbol{u}] \in \Delta \right\}, \quad (5)$$

where matrix $\boldsymbol{P}$ is the projection operator to the column space of perturbed $[\boldsymbol{s}_1, \dots, \boldsymbol{s}_k, \boldsymbol{u}]$ and $\boldsymbol{S}_k' = \boldsymbol{A} \setminus \boldsymbol{S}_k$ is the set of all basis vectors that are not contained in $\boldsymbol{S}_k$. For each $\boldsymbol{u} \in \boldsymbol{S}_k'$, this perturbation problem can be solved by using the technique given in [35]. However, due to its associated gradient descent based iterations, the complexity of solution is still a practical limitation for large N. In this work, we propose a simpler non-iterative perturbation for each $\boldsymbol{s}_i \in \boldsymbol{S}_k$, to maximize the projection under bounded perturbations. At any iteration k, the measurement $\boldsymbol{b}$ can be decomposed as:

$$\boldsymbol{b} = \boldsymbol{b}_\perp + \boldsymbol{b}_{//}, \qquad (6)$$

where $\boldsymbol{b}_{//}$ is the projection of $\boldsymbol{b}$ onto the span of vectors in $\boldsymbol{S}_k$ and $\boldsymbol{b}_\perp$ is the orthogonal residual. Since vectors in $\boldsymbol{S}_k$ are linearly independent, this projection can be uniquely expressed as:

$$\boldsymbol{b}_{//} = \sum_{i=1}^{k} \alpha_i\, \boldsymbol{s}_i, \qquad (7)$$

where α_i is the weight of the corresponding ith column vector. In the proposed approach, as shown in Fig. 1, the $\boldsymbol{s}_i$'s will be
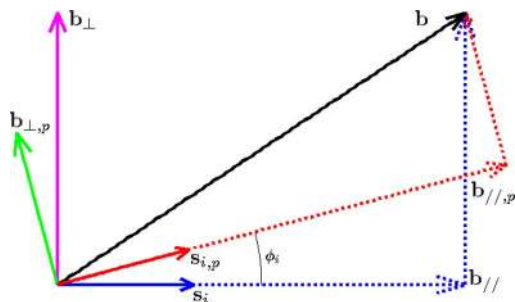
Fig. 1. One dimensional example for perturbed and unperturbed column vectors. As the basis vector rotates, residual decreases.

perturbed by rotating them towards to $\widehat{\boldsymbol{b}_\perp}$ by an angle ϕ_i, where $\widehat{\boldsymbol{b}_\perp} = \boldsymbol{b}_\perp / \|\boldsymbol{b}_\perp\|_2$ is the normalized residual:

$$\boldsymbol{s}_{i,p}(\phi_i) = \boldsymbol{s}_i \cos(\phi_i) + \widehat{\boldsymbol{b}_\perp} \operatorname{sgn}(\alpha_i) \sin(\phi_i). \quad (8)$$

Since both $\boldsymbol{s}_i$ and $\widehat{\boldsymbol{b}_\perp}$ have unit norms and they are orthogonal to each other, $\boldsymbol{s}_{i,p}$, which is the perturbed version of $\boldsymbol{s}_i$, has also unit norm. If the angle of rotation, or equivalently, the allowed amount of perturbation on $\boldsymbol{s}_i$ is large enough, then $\boldsymbol{s}_{i,p}$ can be aligned with $\boldsymbol{b}$ and there will be no residual left in the perturbed basis. If the $\boldsymbol{s}_i$'s are rotated more than adequate, $\boldsymbol{s}_{i,p}$'s can overlap with each other. When more than one vector span the overlapping region, uniqueness of the projection is lost. In order to avoid such overlaps, rotation of $\boldsymbol{s}_i$ should be limited to the half of the minimum angle between $\boldsymbol{s}_i$ and other $\boldsymbol{s}_j$ for $j \neq i$. More precisely, the maximum perturbation angle, $\phi_{\max}$, for a vector $\boldsymbol{a}_i$ should satisfy:

$$\cos(2\phi_{\max}) \leq \mu(\boldsymbol{A}) = \max_{j \neq i} |\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle| \text{ for } i, j \in 1, \ldots, N, \quad (9)$$

where $\mu(\boldsymbol{A})$ is the mutual coherence of $\boldsymbol{A}$ and $\phi_{\max} = \cos^{-1}(\mu(\boldsymbol{A}))/2$. This case is illustrated in Fig. 2, where the maximum allowed perturbation of a vector is such that the cones around the columns of $\boldsymbol{A}$ do not overlap with each other. Perturbations beyond this limit generate switch-over between the chosen set of vectors and cause non-unique projection. Therefore, at each step of the proposed approach, only perturbations satisfying this limit will be considered.

The approach of perturbations embedded in the iterations of Orthogonal Matching Pursuit(OMP) provides a practical technique for sparse reconstruction when basis mismatch is present. In the following discussion, we will provide a detailed theoretical investigation of this approach that will be referred to as Perturbed-OMP, or POMP in short.

Theorem 1: For the largest perturbation angle satisfying $\phi_i < \phi_{\max} < \frac{\pi}{2}$ and $\sum_{i=1}^{k} |\alpha_i| \tan(\phi_i) \leq \|\boldsymbol{b}_\perp\|_2$, the perturbed support vectors $\boldsymbol{s}_{i,p}$ defined in (8), has an orthogonal residual $\boldsymbol{b}_{\perp,p}$ whose ℓ_2 norm is upper bounded as:

$$\|\boldsymbol{b}_{\perp,p}\|_2 \leq \|\boldsymbol{b}_\perp\|_2 \left(1 - \frac{1}{\|\boldsymbol{b}_\perp\|_2} \sum_{i=1}^{k} |\alpha_i| \tan(\phi_i)\right)$$

Proof: From (8), $\boldsymbol{s}_i$ can be written as:

$$\boldsymbol{s}_i = \boldsymbol{s}_{i,p}(\phi_i)\sec(\phi_i) - \widehat{\boldsymbol{b}_\perp} \operatorname{sgn}(\alpha_i) \tan(\phi_i) \quad 1 \leq i \leq k. \quad (10)$$
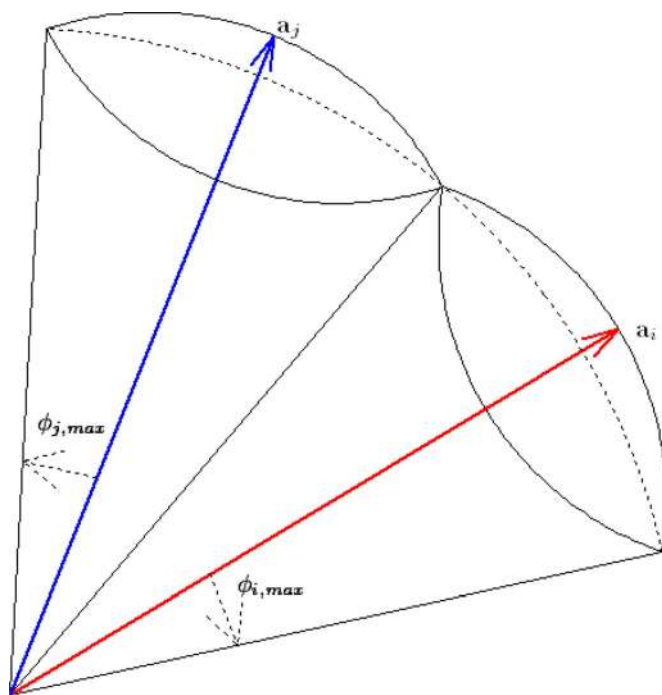

Fig. 2. Each unit column of $\boldsymbol{A}$ has a maximum perturbation angle so that the perturbed vectors do not overlap with each other.

Replacing this decomposition of $\boldsymbol{s}_i$ in terms of $\boldsymbol{s}_{i,p}$ and $\boldsymbol{b}_\perp$ in (7), measurements can be written as:

$$\boldsymbol{b} = \boldsymbol{b}_\perp + \sum_{i=1}^{k} \alpha_i \left(\boldsymbol{s}_{i,p}(\phi_i) \sec(\phi_i) - \widehat{\boldsymbol{b}_\perp} \operatorname{sgn}(\alpha_i) \tan(\phi_i)\right), \quad (11)$$

which can be regrouped to obtain:

$$\boldsymbol{b} = \boldsymbol{b}_\perp \left(1 - \frac{1}{\|\boldsymbol{b}_\perp\|_2} \sum_{i=1}^{k} |\alpha_i| \tan(\phi_i)\right) + \sum_{i=1}^{k} \alpha_i \boldsymbol{s}_{i,p}(\phi_i) \sec(\phi_i). \quad (12)$$

In this decomposition of $\boldsymbol{b}$, the second term is in the span of the perturbed vectors $\boldsymbol{s}_{i,p}$. Therefore, the orthogonal decomposition of $\boldsymbol{b}$ in the perturbed basis is:

$$\boldsymbol{b} = \boldsymbol{b}_{\perp,p} + \boldsymbol{b}_{//,p}, \quad (13)$$

where $\boldsymbol{b}_{//,p}$ is the projection of $\boldsymbol{b}$ onto the span of the perturbed basis vectors, and $\boldsymbol{b}_{\perp,p}$ is the corresponding residual. Thus,

$$\boldsymbol{b}_{\perp,p} = \overset{\perp}{\boldsymbol{P}}_{[\boldsymbol{s}_{i,p}, \ldots, \boldsymbol{s}_{k,p}]} \boldsymbol{b}_\perp \left(1 - \frac{1}{\|\boldsymbol{b}_\perp\|_2} \sum_{i=1}^{k} |\alpha_i| \tan(\phi_i)\right). \quad (14)$$

Since, the ℓ_2 norm of the projection operation is less than one, we have:

$$\|\boldsymbol{b}_{\perp,p}\|_2 \leq \left\|\boldsymbol{b}_\perp \left(1 - \frac{1}{\|\boldsymbol{b}_\perp\|_2} \sum_{i=1}^{k} |\alpha_i| \tan(\phi_i)\right)\right\|_2$$

$$= \|\boldsymbol{b}_\perp\|_2 \left|1 - \frac{1}{\|\boldsymbol{b}_\perp\|_2} \sum_{i=1}^{k} |\alpha_i| \tan(\phi_i)\right|. \quad (15)$$

From the hypothesis, $\sum_{i=1}^{k} |\alpha_i| \tan(\phi_i) \leq \|\boldsymbol{b}_\perp\|_2$, thus we obtain the desired upper bound:

$$\|\boldsymbol{b}_{\perp,p}\|_2 \leq \|\boldsymbol{b}_\perp\|_2 \left(1 - \frac{1}{\|\boldsymbol{b}_\perp\|_2} \sum_{i=1}^{k} |\alpha_i| \tan(\phi_i) \right), \quad \text{q.e.d.} \tag{16}$$

∎

The proof of Theorem 1 also reveals the following fact:

Corollary 1: The upper bound derived in Theorem 1 is monotonically decreasing as a function of ϕ_i as long as the condition $\sum_{i=1}^{k} |\alpha_i| \tan(\phi_i) \leq \|\boldsymbol{b}_\perp\|_2$ is satisfied.

Proof: In Theorem 1, it is shown that:

$$\|\boldsymbol{b}_{\perp,p}\|_2 \leq \|\boldsymbol{b}_\perp\|_2 \left(1 - \frac{1}{\|\boldsymbol{b}_\perp\|_2} \sum_{i=1}^{k} |\alpha_i| \tan(\phi_i) \right).$$

As long as $\sum_{i=1}^{k} |\alpha_i| \tan(\phi_i) \leq \|\boldsymbol{b}_\perp\|_2$, the right hand side is a valid non-negative upper bound for $\|\boldsymbol{b}_{\perp,p}\|_2$, and is a continuous function of all $\phi_i < \phi_{i,\max}, 1 \leq i \leq k$, with partial derivatives:

$$\frac{-|\alpha_i|}{1 + \tan^2(\phi_i)} < 0, \quad 1 \leq i \leq k, \tag{17}$$

which are all strictly negative for $|\alpha_i| \neq 0$. Thus, the upper bound decreases monotonically as a function of $\phi_i, 1 \leq i \leq k$, q.e.d. ∎

The first implication of Corollary 1 is that once a candidate support set is available and $\sum_{i=1}^{k} |\alpha_i| \tan(\phi_i) \leq \|\boldsymbol{b}_\perp\|_2$, rather than searching for a perturbation that minimizes $\|\boldsymbol{b}_{\perp,p}\|_2$, one can minimize its upper bound simply by perturbing each vector in the support up to its allowed limit.

The second implication of Corollary 1 is that if the perturbation angles are increased, i.e., more freedom for the adjustment of the basis exists, the derived upper bound decreases. Therefore, for a certain amount of perturbation, this raises the possibility of driving the upper bound to zero.

Corollary 2: Let $\phi_i = \phi^*, 1 \leq i \leq k$. Then if $\phi^* = \tan^{-1}\left(\frac{\|\boldsymbol{b}_\perp\|_2}{\|\boldsymbol{\alpha}\|_1} \right)$, perturbation of the support vectors up to ϕ^*, results in $\|\boldsymbol{b}_{\perp,p}\|_2 = 0$, yielding a k-sparse reconstruction with no residual error.

Proof: Simply replace $\tan(\phi_i) = \tan(\phi^*) = \frac{\|\boldsymbol{b}_\perp\|_2}{\sum_{i=1}^{k} |\alpha_i|}$ in the upper bound given in the (16) to get:

$$0 \leq \|\boldsymbol{b}_{\perp,p}\|_2 \leq \|\boldsymbol{b}_\perp\|_2$$
$$\times \left(1 - \frac{1}{\|\boldsymbol{b}_\perp\|_2} \sum_{i=1}^{k} |\alpha_i| \tan(\phi^*) \right) = 0, \quad \text{q.e.d.} \tag{18}$$

∎

It is important to note that Theorem 1 and Corollaries 1 and 2 are valid for any set of linearly independent k vectors from the columns of $\boldsymbol{A}$. Even though chosen subset $\boldsymbol{S}_k$ does not include any component from the correct support, this theorem guarantees that by using perturbation, the residual can be decreased. If $\boldsymbol{S}_k$ is the correct support, or some subset of the correct support, then $\boldsymbol{b}^* = \boldsymbol{S}_k \boldsymbol{S}_k^\dagger \boldsymbol{b}$ becomes a good approximation of $\boldsymbol{b}$. In this case, $\|\boldsymbol{b} - \boldsymbol{b}^*\|_2$ becomes smaller and hence ϕ^*, the angle of perturbation at which the upper bound becomes zero, is a smaller angle.

Proposition 1: A perturbed set of support vectors $\{\boldsymbol{s}_{i,p}(\phi), 1 \leq i \leq k\}$ will have $\|\boldsymbol{b}_{\perp,p}\|_2 = 0$ if and only if the upper bound in (16) is zero.

Proof: Using perturbations up to ϕ^*, the residual will be zero; hence we can expand $\boldsymbol{b}$ as:

$$\boldsymbol{b} = \sum_{i=1}^{k} \beta_i \boldsymbol{s}_{i,p}(\phi^*). \tag{19}$$

Assume there is another angle $\widetilde{\phi} < \phi^*$ resulting also a zero residual. Hence, we can expand $\boldsymbol{b}$ as:

$$\boldsymbol{b} = \sum_{i=1}^{k} \gamma_i \boldsymbol{s}_{i,p}(\widetilde{\phi}). \tag{20}$$

If we subtract (19) from (20) term by term, we get the following:

$$0 = \sum_{i=1}^{k} (\beta_i \, \boldsymbol{s}_{i,p}(\phi^*) - \gamma_i \, \boldsymbol{s}_{i,p}(\widetilde{\phi})),$$
$$= \sum_{i=1}^{k} ((\beta_i \cos(\phi^*) - \gamma_i \cos(\widetilde{\phi}))\boldsymbol{s}_i + \cdots$$
$$(\beta_i \sin(\phi^*) - \gamma_i \sin(\widetilde{\phi}))\widehat{\boldsymbol{b}_\perp} \, \text{sgn}\,(\alpha_i)). \tag{21}$$

Since $\boldsymbol{s}_i$ and $\widehat{\boldsymbol{b}_\perp}$ are orthogonal to each other, $\beta_i \sin(\phi^*) = \gamma_i \sin(\widetilde{\phi})$ and $\beta_i \cos(\phi^*) = \gamma_i \cos(\widetilde{\phi})$. If we divide both equations term by term, we get $\tan(\phi^*) = \tan(\widetilde{\phi})$. Since we only consider acute angles, $\phi^* = \widetilde{\phi}$, which contradicts the assumption. Hence, the perturbation angle which results upper bound in (16) to be zero is unique, q.e.d. ∎

Although in the proposed perturbation approach norm of the residual can be made zero for large enough perturbation angles, a relaxed constrained optimization problem, where $\|\boldsymbol{b}_{\perp,p}\|_2 < \epsilon$, is more appropriate when there is noise in $\boldsymbol{b}$. Under this relaxation, more sparse reconstructions can also be obtained. In Fig. 3 as an illustrative case, the error residual is shown as a function of perturbation angle. It is seen that the constraint line $\|\boldsymbol{b}_{\perp,p}\| = \epsilon$ intersects with the error norm curve at ϕ_ϵ^*, which is the first intersecting angle smaller than ϕ^*. The following theorem provides an estimate for ϕ^* by using the intersection of $\|\boldsymbol{b}_{\perp,p}\|_2 = \epsilon$ with the upper bound in (16).

Theorem 2: If the perturbation angles of support vectors in $\boldsymbol{S}_k$ are all chosen as $\phi_\epsilon^* = \tan^{-1}\left(\frac{\|\boldsymbol{b}_\perp\|_2 - \epsilon}{\|\boldsymbol{\alpha}\|_1} \right)$, where $\boldsymbol{b} = \boldsymbol{b}_\perp + \sum_{i=1}^{k} \alpha_i \, \boldsymbol{s}_i$, then $\|\boldsymbol{b}_{\perp,p}\|_2 \leq \epsilon$.
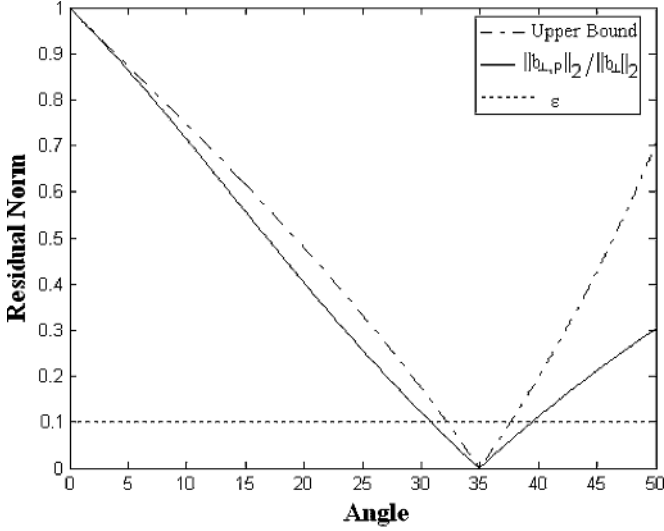
Fig. 3.　Upper bound and $\|\boldsymbol{b}_{\perp,p}\|_2/\|\boldsymbol{b}_\perp\|_2$ as a function of ϕ.

Proof: If ϕ_i in (16) is replaced with ϕ_ϵ^*, then the upper bound becomes:

$$
\begin{aligned}
\|\boldsymbol{b}_{\perp,p}\|_2 &\leq \|\boldsymbol{b}_\perp\|_2 \left(1 - \frac{1}{\|\boldsymbol{b}_\perp\|_2} \sum_{i=1}^k |\alpha_i| \tan(\phi_\epsilon^*) \right) \\
&= \|\boldsymbol{b}_\perp\|_2 \left(1 - \frac{\|\boldsymbol{\alpha}\|_1}{\|\boldsymbol{b}_\perp\|_2} \frac{(1 - \epsilon/\|\boldsymbol{b}_\perp\|_2)\|\boldsymbol{b}_\perp\|_2}{\|\boldsymbol{\alpha}\|_1} \right) \\
&= \|\boldsymbol{b}_\perp\|_2 \frac{\epsilon}{\|\boldsymbol{b}_\perp\|_2} = \epsilon, \quad \text{q.e.d.} \qquad (22)
\end{aligned}
$$

∎

Note that for small enough ϵ, the residual error of projection onto the perturbed support vectors, i.e., $\|\boldsymbol{b}_{\perp,p}\|_2$, the above bound will also be tight and $\|\boldsymbol{b}_{\perp,p}\|$ will be close to ϵ. Therefore, for the kth iteration of the proposed POMP approach, there are three limits on the perturbation angles: first, $\phi_{\max}$, which is the limit predefined by the coherence of $\boldsymbol{A}$ as defined in (9); second ϕ_ϵ^*, beyond which the residual error norm is lower than ϵ; and third ϕ', the user-defined limit. A user may want a solution with larger or smaller perturbations and, therefore may want $\phi_i \leq \phi_i'$. Hence, the allowable perturbation angle ϕ_i for the ith support vector should be

$$\phi_i = \min \{\phi_{\max}, \phi_\epsilon^*, \phi_i'\}. \qquad (23)$$

In general, $\phi_{\max}$ is a larger angle compared to other two. In the early phases of POMP, where the number of chosen support vectors are lower than actual sparsity level, ϕ_ϵ^* is expected to be larger than the allowed perturbation, ϕ'. Therefore, in the early phases of the iterations, the residual error norm of the perturbed support vectors is typically larger than the termination criterion ϵ. Yet, it is expected that as new columns are added to the current support during the iterations, ϕ_ϵ^* will decrease and will be less than ϕ' eventually. Once $\phi_i = \min\{\phi_{\max}, \phi_\epsilon^*, \phi_i'\} = \phi_\epsilon^*$, iterations will stop, since after the perturbation the residual will have a norm less than ϵ. The following theorem, whose proof is provided in the Appendix, states this expectation formally.

Theorem 3: Let $\boldsymbol{S}_k$ be the support estimation in the kth iteration and ϕ_k^* be the required perturbation angle as derived in Corollary 2 and let $\boldsymbol{s}_{k+1}$ be the basis vector chosen in the

TABLE I
PERTURBED-OMP (POMP) ALGORITHM

Inputs:
$(\boldsymbol{A}, \boldsymbol{b}, \epsilon, \phi')$
Initialization:
$\boldsymbol{b}_{\perp,0} = \boldsymbol{b}$, $S_0 = \{\}$, $e = \|\boldsymbol{b}_{\perp,0}\|_2$, $k = 1$
Keep iterating until $e < \epsilon$
Orthogonal Pursuit:
$\quad \boldsymbol{U}_k = \boldsymbol{A}/\boldsymbol{S}_{k-1}$
$\quad j^* = \arg \max_j |\boldsymbol{u}_j^T \boldsymbol{b}_{\perp,k-1}|$
$\quad \boldsymbol{S}_k = \boldsymbol{S}_{k-1} \bigcup \{\boldsymbol{u}_{j^*}\}$
$\quad \boldsymbol{x}_k = \boldsymbol{S}_k^\dagger \boldsymbol{b}$
$\quad \boldsymbol{b}_{\perp,k} = \boldsymbol{b} - \boldsymbol{S}_k \boldsymbol{x}_k$
Perturbation Procedure:
If $\|\boldsymbol{b}_{\perp,k}\|_2 > \epsilon$
$\quad \widehat{\boldsymbol{b}}_{\perp,k} = \boldsymbol{b}_{\perp,k}/\|\boldsymbol{b}_{\perp,k}\|_2$
$\quad \phi_k^* = \tan^{-1}\left(\frac{\|\boldsymbol{b}_{\perp,k}\|_2 - \epsilon}{\|\boldsymbol{x}_k\|_1} \right)$
$\quad \phi^k = \min(\phi', \phi_k^*)$,
$\quad$ For $i = 1$ to k;
$\quad\quad \boldsymbol{s}_{i,p} = \boldsymbol{s}_i \cos(\phi_i^k) + \widehat{\boldsymbol{b}}_{\perp,k}\, \text{sgn}(x_{k,i}) \sin(\phi_i^k)$
$\quad\quad \boldsymbol{S}_{k,p} = \{\boldsymbol{s}_{i,p}, 1 \leq i \leq k\}$
$\quad\quad \boldsymbol{b}_{\perp,k} = (\boldsymbol{I} - \boldsymbol{S}_{k,p}\, \boldsymbol{S}_{k,p}^\dagger) \boldsymbol{b}$
$\quad\quad e = \|\boldsymbol{b}_{\perp,k}\|_2$
Return with $k = k + 1$.
Output: $\boldsymbol{x}^* = \boldsymbol{S}_{k,p}^\dagger \boldsymbol{b}$

$(k+1)$th iteration. If $\|\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}\|_1 \leq 1 + \gamma$, then, $\phi_{k+1}^* < \phi_k^*$, where $\gamma > 0$.

Since $\boldsymbol{A} \in \Re^{M \times N}$ is full rank, iterations stop when k reaches M without any perturbation in the worst case. Therefore, ϕ_M^* is zero. According to Theorem 3, when the required condition is satisfied, ϕ_k^* produces a monotonically decreasing sequence throughout the iterations. Since we know that $\phi_M^* = 0$, there exists a $K \leq M$ such that $\phi_K^* \leq \phi'$ due to monotonicity of the angles. Therefore, Theorem 3 indirectly guarantees the termination of iterations.

It is important to note that Tropp's Exact Recovery Condition for OMP [36], which simply states that if $\|\boldsymbol{S}_k^\dagger \boldsymbol{a}_i\|_1 \leq 1$ for all un-selected basis vectors, OMP will select one of the correct support vectors in the next iteration, is a stricter version of Theorem 3. Hence, we can safely guarantee that maximum required angle, ϕ_ϵ^*, will always decrease throughout the iterations as long as OMP is guaranteed to provide exact reconstruction. This point will be revisited with numerical simulations in Section III.

The steps of the proposed Perturbed Orthogonal Matching Pursuit (POMP) algorithm is detailed in Table I. Starting with a set of unit norm vectors, in the kth iteration, POMP searches over the dictionary to find the vector providing the largest absolute inner product with the residual. After the selection of the new vector, POMP computes the projection of the measurement, $\boldsymbol{b}$, onto the new larger support and finds the residual. OMP continues with the next iteration here. However, POMP proceeds with the perturbation. Given ϵ and ϕ', ϕ is computed according to (23). It is assumed that user provides an angle, ϕ', less than the mutual coherence of the basis, $\phi' \leq \cos^{-1}(\mu(\boldsymbol{A}))/2$, otherwise perturbations are limited by the basis itself. After that POMP starts to perturb each vector in the current support as given in (8). Then, the measurement vector, $\boldsymbol{b}$, is projected onto the perturbed support and the new residual is found. If the norm

of the residual is less than ϵ, iterations are terminated, otherwise POMP continues with the next iteration.

One important characteristic of POMP is the promise of the sparser solutions. This property can be revealed as follows. Assume that POMP has produced a K-sparse solution. Iterations can be terminated due to two reasons. If the observation $\boldsymbol{b}$ is already sparse in $\boldsymbol{A}$, then at the Kth iteration we get $\|\boldsymbol{b}_\perp\|_2 \le \epsilon$. Since OMP and POMP have the same selection criteria, OMP also chooses the very same $\boldsymbol{S}_K$ and obtains the same K sparse solution. However, if $\boldsymbol{b}$ is not sparse in $\boldsymbol{A}$, POMP obtains the K sparse solution using the perturbation. Since $\|\boldsymbol{b}_{\perp,p}\|_2 \le \epsilon$ but $\|\boldsymbol{b}_\perp\|_2 > \epsilon$, OMP iterates at least one more time resulting a denser solution. Therefore, for any observation $\boldsymbol{b}$, OMP produces denser, or equally sparse at best, solutions than POMP.

OMP is preferred in many applications due to its computational efficiency. In the perturbation stage of the proposed algorithm, the inverse tangent operation can be well approximated by using tables and low order power series expansion, therefore the computational order is determined by the least-squares solution on the perturbed basis, which has a complexity of $O(k^2 M)$ in the kth iteration, which is the same as the standard OMP algorithm. If the algorithm terminates in the Kth iteration, overall complexity becomes $O(K^3 M)$ for both POMP and OMP. In the worst case, algorithm will terminate eventually in the Mth iteration, which results in complexity of $O(M^4)$.

III. SIMULATION RESULTS

In this section, performance of POMP algorithm will be investigated and compared with alternative techniques under random basis mismatch. For this purpose, sparse reconstruction of sinusoids from their time samples will be considered. In this example, following dictionary vectors are used:

$$\boldsymbol{a}_i = \cos(2\pi f_i \boldsymbol{t}) \quad 1 \le i \le N, \qquad (24)$$

where f_i is the frequency of the ith dictionary vector and $\boldsymbol{t}$ is the vector of time samples at which the signal is sampled. However, in practice, due to time jitter in the sampling, the observed signal $\boldsymbol{b}$ is not sampled at the nominal sampling times, resulting in a dictionary mismatch:

$$\boldsymbol{b} = \sum_{i=1}^{K} x_i \cos(2\pi f_i(\boldsymbol{t} + \boldsymbol{t}^j)) = \sum_{i=1}^{K} x_i \widetilde{\boldsymbol{a}}_i, \qquad (25)$$

where $\boldsymbol{t}^j$ can be modeled as a vector of independent and uniformly distributed random variables in the range of $[-\delta T/2, \delta T/2]$ where δT represents the level of jitter. The overall effect of time-jitter can be considered as a random perturbation on the dictionary $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N]$. In this more realistic scenario, $\boldsymbol{b}$ is typically non-sparse in the assumed dictionary $\boldsymbol{A}$.

In the first simulation, the dictionary is constructed by using frequencies $f_i = i \cdot \delta_f + 100$ Hz for $1 \le i \le 200$ and frequency separation of $\delta_f = 1$ Hz. The sampling time vector $\boldsymbol{t}$ is created over the $[0, 1]$ second interval with randomly chosen $M = 180$ time samples. A 25-sparse signal $\boldsymbol{x}$ is randomly generated with non-zero coefficients selected uniformly random from the range $[-10, -1] \cup [1, 10]$. Observations, $\boldsymbol{b}$, are generated with time jitter of $\delta T = 8.3$ μsec. The same termination criteria of $\epsilon = 10^{-3}$ is used for all compared techniques.

To select a proper maximum perturbation angle for POMP, the expected value of the perturbation angles are calculated. Based on the following, the normalized inner products of $\boldsymbol{a}_i$ and $\widetilde{\boldsymbol{a}}_i$:

$$\cos(\phi_i) = \frac{\boldsymbol{a}_i^T \widetilde{\boldsymbol{a}}_i}{\|\boldsymbol{a}_i\|_2 \|\widetilde{\boldsymbol{a}}_i\|_2}$$

$$= \frac{\sum_{k=1}^{M} 0.5 \cos\left(2\pi f_i \left(2t_k + t_k^j\right)\right) + 0.5 \cos\left(2\pi f_i t_k^j\right)}{M \boldsymbol{a}_{i,\mathrm{RMS}} \widetilde{\boldsymbol{a}}_{i,\mathrm{RMS}}}.$$

Since the sampling jitter is modeled as an i.i.d. sequence, $E\left\{\frac{1}{M} \sum_{k=1}^{M} 0.5 \cos\left(2\pi f_i \left(2t_k + t_k^j\right)\right)\right\} \simeq 0$ and $E\{\boldsymbol{a}_{i,\mathrm{RMS}}\} \simeq E\{\widetilde{\boldsymbol{a}}_{i,\mathrm{RMS}}\} \simeq \sqrt{0.5}$ for large M. Therefore, the expected value of the cosine of the perturbation angles can be approximated as:

$$E\{\cos(\phi_i)\} \simeq \frac{E\left\{0.5 \cos\left(2\pi f_i t_k^j\right)\right\}}{\sqrt{0.5}\sqrt{0.5}} = E\left\{\cos\left(2\pi f_i t_k^j\right)\right\}. \qquad (26)$$

Using the small angle approximation of cosines, (26) can be further simplified as:

$$E\{\cos(\phi_i)\} \simeq E\left\{1 - 2\pi^2 f_i^2 t_k^{j^2}\right\} = 1 - 2\pi^2 f_i^2 \sigma_j^2,$$

where σ_j^2 is the jitter variance. Since jitter is assumed to be uniform in the interval $[-\delta T/2, \delta T/2]$, $\sigma_j^2 = \delta T^2/12$. Finally, using a small angle approximation for ϕ_i, we obtain,

$$E\{\phi_i\} \approx 60° \sqrt{3} f_i \delta T. \qquad (27)$$

Therefore, in the implementation of POMP, the allowed perturbation angles are selected according to (27) for each column, respectively. For the ith column of $\boldsymbol{A}$ that corresponds to $f_i = i + 100$ Hz, the maximum perturbation angle is $\phi_i = (i + 100) \cdot 8.63° \cdot 10^{-4}$.

Fig. 4 shows actual and reconstructed signal coefficients for OMP and POMP techniques for a random realization of jitter. In this specific case, while POMP correctly reconstructs the sparse signal coefficients, OMP generates a highly non-sparse signal. For this realization, Fig. 5 shows $\|\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}\|_1$, the Tropp's Exact Recovery Condition (ERC) [36] and the proposed bound of Theorem 3. Even though ERC is not satisfied for $k > 16$, the proposed conditions of Theorem 3 are satisfied for this case. Therefore, this example shows that, the guarantees given by Theorem 3 provide a larger regime in which POMP is successful.

In this section of the simulations, we will investigate performance of POMP when the condition of the Theorem 3 is not satisfied. For this purpose, we conduct a large set of Monte-Carlo simulations in excess of 10^6 trials in which the sparsity level K and number of measurement M are swept from 2 to 85 and K to 200, respectively. For each (K, M) pair, 100 cases are simulated. In each of these simulations, cases where the perturbation angle decrease, i.e., $\phi_{k+1}^* \le \phi_k^*$, are identified as event E_1, and the sufficiency condition of Theorem 3 is satisfied, i.e., $\|\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}\|_1 \le 1 + \gamma$, are identified as E_2. Out of all the runs at each sparsity level, the cases where only E_1 is valid and both
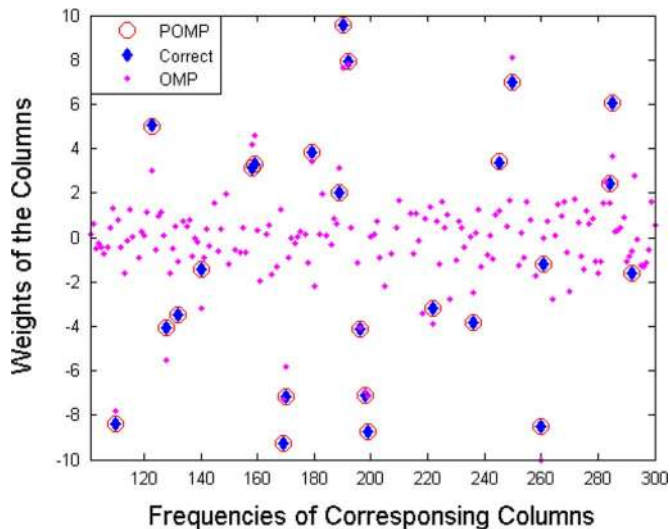
Fig. 4. A realization of the reconstruction problem under time-jitter where OMP drastically fails and produces 167-sparse solution.
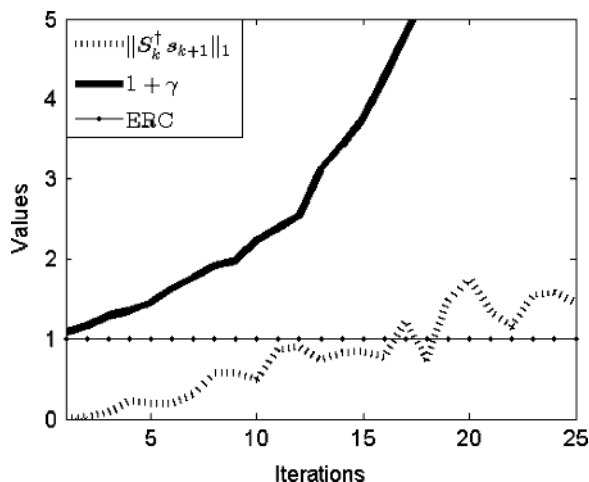

Fig. 5. $\|S_k^\dagger s_{k+1}\|_1$ and its bound $(1 + \gamma)$ as a function of iterations. Even though ERC is not satisfied for $k > 16$, condition in Theorem 3 is satisfied due to large value of γ, ensuring the decrease of maximum perturbation angle during the iterations.

E_1 and E_2 are valid are counted separately. In Fig. 6, the statistics of the cases are shown for all the experiments conducted.

It can be observed that for $K \leq 10$, the case E_2 is satisfied with empirical probability of 1, and as sparsity level increases, the probability of the case E_2 being satisfied decreases. The maximum perturbation angle decreases consistently when the case E_2 is satisfied. Note that, Theorem 3 is only a sufficiency condition and in the simulated scenario for $K \geq 10$ its requirements are not guaranteed to met. Nevertheless, in all the simulations conducted, it is observed that the perturbation angle decreases.

To compare the overall performance in the above described set of simulations, the following metrics are used: the normalized signal reconstruction error, $\|x - x^*\|_2 / \|x\|_2$; the level of sparsity, $\|x^*\|_0$; the distance between signal supports, $1 - \dfrac{|S^* \cap S|}{\max\{|S^*|, |S|\}}$; and the normalized residual norm, $\|b - S^* x^*\|_2 / \|b\|_2$, where (x, S) is the correct signal and its support, and (x^*, S^*) is the obtained solution and its corresponding support [37].
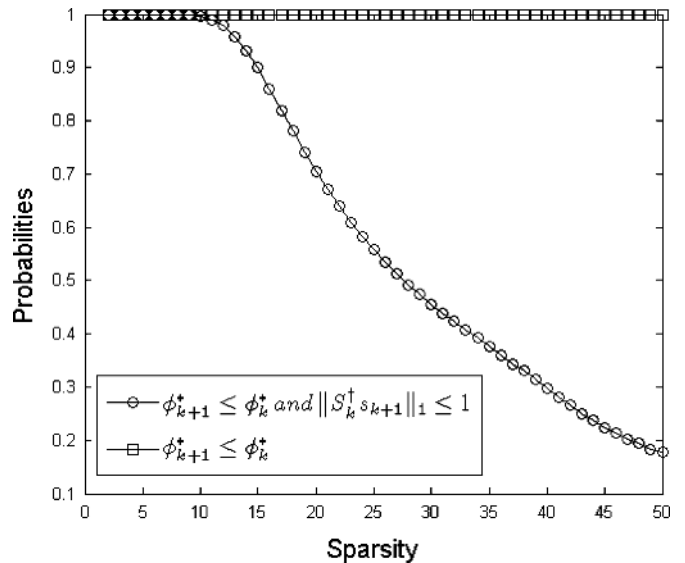

Fig. 6. Empirical probabilities of E_1 is valid and (E_1, E_2) is jointly valid.

In Figs. 7 and 8, the average performance results obtained for OMP and POMP are shown. Fig. 7 shows the normalized reconstruction error for both OMP and POMP in dB, while the corresponding support distances can be seen in Fig. 8. For the same number of measurements and range of sparsity levels, it can be observed from Fig. 7 that both OMP and POMP have similar phase transition curves; however POMP produces significantly lower reconstruction errors. In Fig. 8, it can be seen that POMP provides more reliable supports as well.

Instead of using perturbation procedure, one can try to use OMP in a finely discretized frequency domain. This way, the required perturbations on the dictionary vectors can be reduced. However, increasing the density of the basis also increases the coherence of the dictionary which adversely affects performance of the CS techniques. In the simulated scenario, the effect of increasing frequency density in the dictionary is shown in Fig. 9. In this case number of measurements is kept constant at $M = 180$, however the size of the dictionary is increased as $N = 199/\delta_f + 1$. The frequency range of $[101, \quad 300]$ Hz is considered in the simulations. Random time samples are chosen in $[0, 1/\delta_f]$ time interval with a jitter of $\delta T = 8.3\ \mu\text{sec}$ time jitter. Compared to $\delta_f = 1$, when $\log_{10}(\delta_f) = -3$ is used, this corresponds to approximately 1000 times denser sampled frequency dictionary with $N = 199001$. Sparsity of the signal is kept constant at $K = 10$, and the SNR of the observed signal is kept at 60 dB to better observe the effect of using a denser dictionary in the reconstructions. As seen in the Fig. 9(a), using a denser basis can decrease the reconstruction error up to approximately -50 dB for OMP. However, beyond this point, OMP reaches an error floor. Also, as seen from Fig. 9(b), obtained sparsity is reduced significantly. On the other hand, reconstruction performance of POMP is almost independent from the density of the dictionary. For all tested cases of N, POMP has a reconstruction error below -60 dB and yields a 10 sparse solution, that matches the actual sparsity. Note that using a denser basis significantly increases the mutual correlation of the dictionary vectors as well as the required number of inner-products. Therefore, rather than using OMP
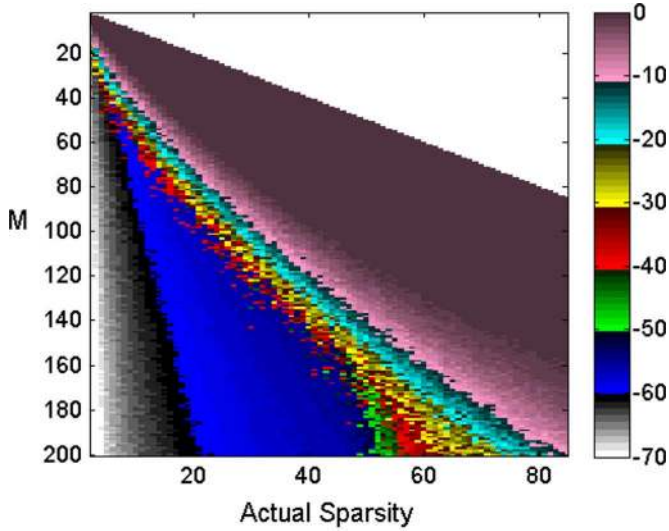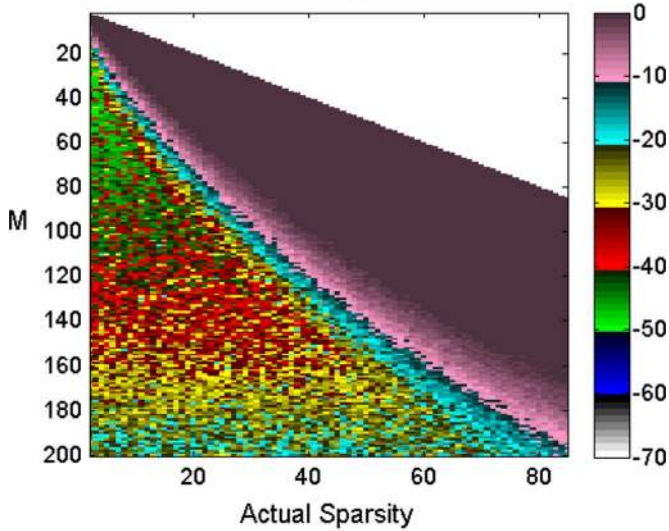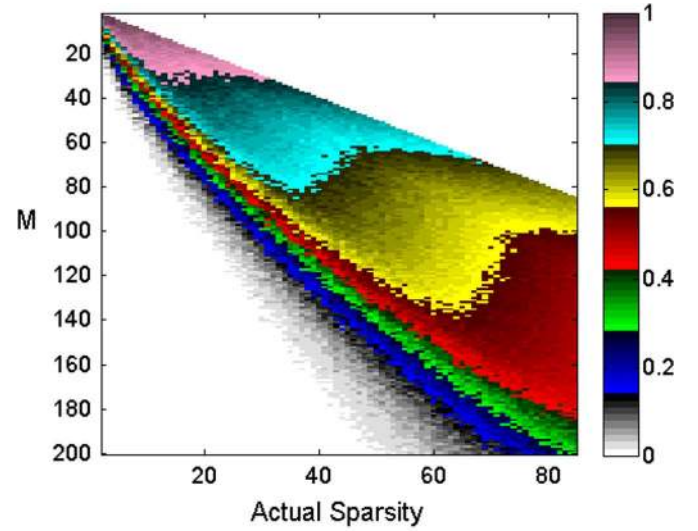
Fig. 7. Sparse signal reconstruction average error is in dB, $20 \log_{10}\left(\frac{\|\boldsymbol{x}^* - \boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}\right)$ as a function of measurement number and sparsity. (a) POMP. (b) OMP.

Fig. 8. Average distances between actual and obtained supports, $1 - \frac{|\boldsymbol{S}^* \cap \boldsymbol{S}|}{\max\{|\boldsymbol{S}^*|, |\boldsymbol{S}|\}}$ as a function of measurement number and sparsity. (a) POMP. (b) OMP.

over a larger and denser dictionary, it is advisable to use POMP over a moderate size dictionary. Note that if the frequencies of the signal components were known precisely, the linearization based techniques proposed in [32], [33], [38] could have been used for the estimation of the jitter in the sampling times.
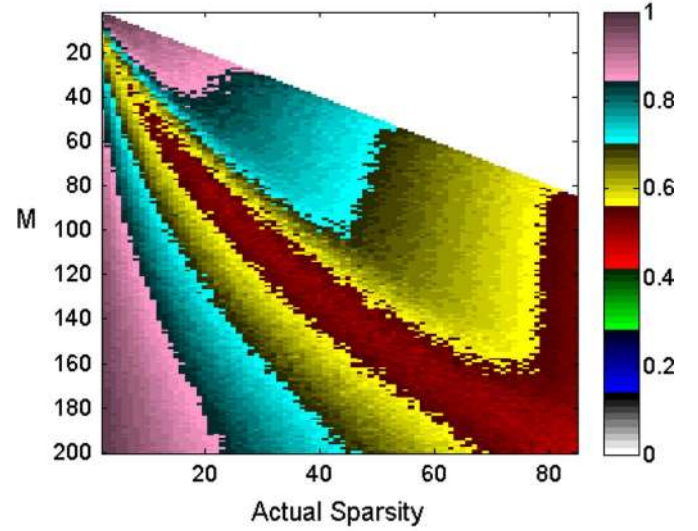
In this part of the simulations, we compare POMP with OMP, CoSaMP, Basis Pursuit, and Sparse Total Least Squares (S-TLS) [33] algorithms. Since POMP is based on OMP iterations, they have similar phase transition characteristics as given in Fig. 7. Hence, in the following simulations, we will stay in the K/M regime in which OMP works successfully.

Unlike OMP and POMP, CoSaMP requires the correct sparsity level of the signal. Since, such information is not available in general, CoSaMP reconstructions obtained at all sparsity levels starting from $K = 1$ until the residual error is below the specified ϵ level.

Basis Pursuit (BP) which is also known as ℓ_1 reconstruction is implemented using the convex optimization toolbox CVX. To

induce sparsity, hard thresholding is applied to BP reconstruction results. Let $\hat{\boldsymbol{x}} = [\hat{\boldsymbol{x}}_l^T \ \hat{\boldsymbol{x}}_{l'}^T]^T$ be the BP reconstruction sorted in the absolute sense and $\hat{\boldsymbol{x}}_l^T$ be the vector containing l largest coefficients. Then, in the reported results here, the threshold, τ, is selected as $\tau = \hat{\boldsymbol{x}}(l)$ such that $\|\hat{\boldsymbol{x}}_l\|_2 \geq (1 - 10^{-6})\|\hat{\boldsymbol{x}}\|_2$.

In [33], two algorithms have been proposed for S-TLS; one finds the global optimum with highly demanding computational cost, and the other one is computationally more efficient but only guaranteed to converge to a local minima. Here, this more efficient technique which is called as coordinate descend (CD) based S-TLS is used for comparison. Since S-TLS has an iterative structure, the stopping criteria is important. To obtain a consistent comparison of algorithms, S-TLS is terminated when the residual error is below the specified ϵ at the end of an iteration. Rather than an all-zero vector, iterations are started with the solution of the obtained BP reconstruction to achieve faster convergence for S-TLS. In the following results, STLS-1 corresponds to $\lambda = 10^{-1}$ and STLS-2 corresponds to $\lambda = 10^{-2}$,
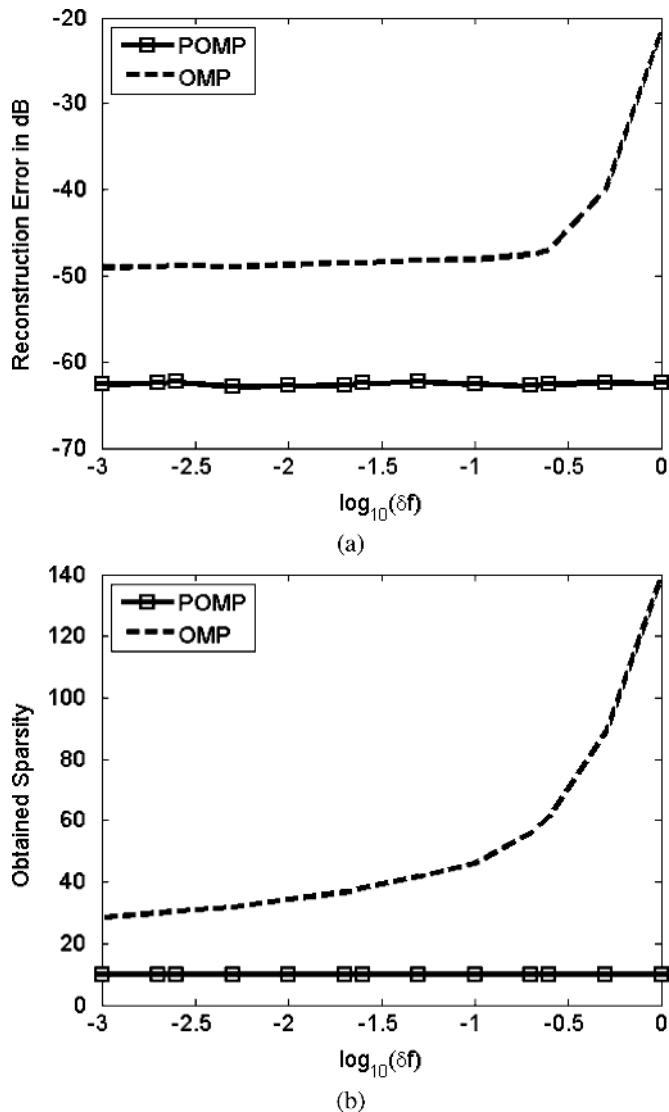
Fig. 9. Reconstruction performance with respect to density of the complex exponentials. (a) Sparse signal reconstruction error in dB, (b) Obtained sparsity level for POMP and OMP algorithms.
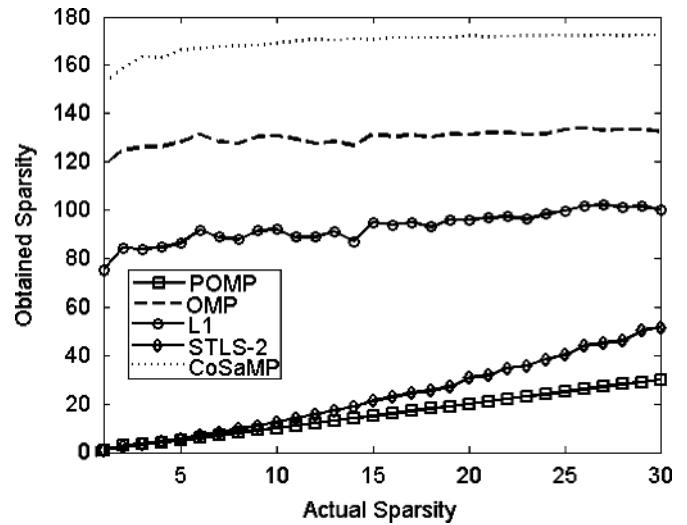

Fig. 10. Obtained sparsity of the reconstructed signal as a function of actual sparsity.
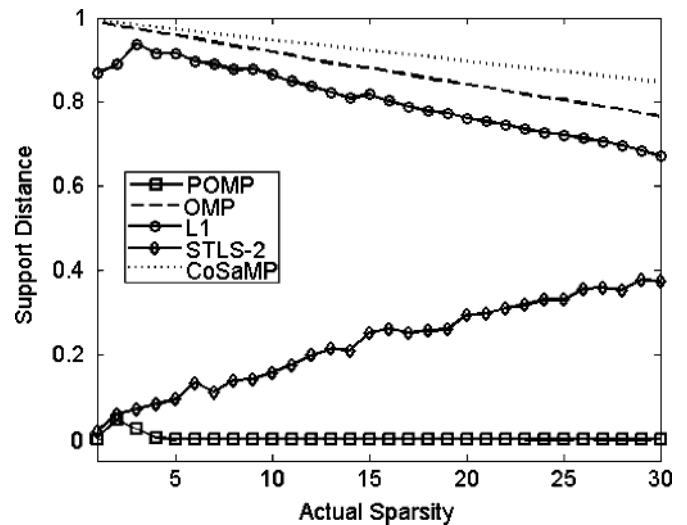
Fig. 11. Distances between actual and obtained supports as a function of actual sparsity.

where λ is the sparsity-tuning parameter of the CD based S-TLS algorithm.

For all the techniques compared, Figs. 10 and 11 shows the obtained sparsity levels and support distances as a function of sparsity, respectively. POMP and S-TLS, being perturbation based techniques, can achieve the correct sparsity level for smaller K. As the actual sparsity increases, S-TLS fails to obtain the correct sparsity, whereas POMP successfully finds the support up to $K = 30$. Techniques that do not employ dictionary perturbation provided results that are significantly inferior than POMP and S-TLS at all sparsity levels. The performance of these techniques saturates and produce solutions at about the same sparsity, irrespective of the actual sparsity level. Since the observed signal is not sparse in the assumed dictionary $\boldsymbol{A}$ with $\mathrm{rank}(\boldsymbol{A}) = M = 180$, OMP goes up to two-third of the rank whereas CoSaMP uses approximately 90% of the rank. On the other hand, POMP obtains the correct sparsity level and the support as shown in Fig. 11. S-TLS gradually produces higher support distances due to denser solutions obtained in less sparse

signals. Since techniques that do not employ perturbation has a saturated sparsity estimate, $\max\{|\boldsymbol{S}^*|, |\boldsymbol{S}|\} = |\boldsymbol{S}^*|$, and $|\boldsymbol{S}^* \cap \boldsymbol{S}|$ increases gradually as $|\boldsymbol{S}|$ increases. Therefore, their distance metric produces smaller values. However, this situation should not be considered as performance improvement for less sparse signals. It is because they simply fail to recover the correct support at all tested sparsity levels.

Fig. 12 shows the average normalized reconstruction error in dB as a function of true sparsity for each of the compared reconstruction techniques. As expected, BP performs better than OMP. However, even though S-TLS employs perturbation, it has very similar performance to BP. POMP, on the other hand, has a significantly lower, -20 dB, signal reconstruction error as compared to S-TLS and BP. CoSaMP has the worst performance among all the compared techniques with significantly higher reconstruction errors.

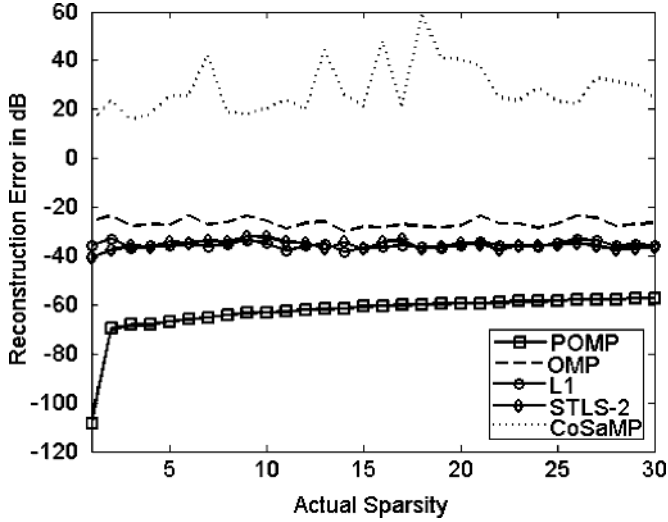In the above simulations, termination criteria of all algorithms are determined by the residual norm level ϵ. We assume that

Fig. 12. Signal reconstruction error in dB as a function of actual sparsity.
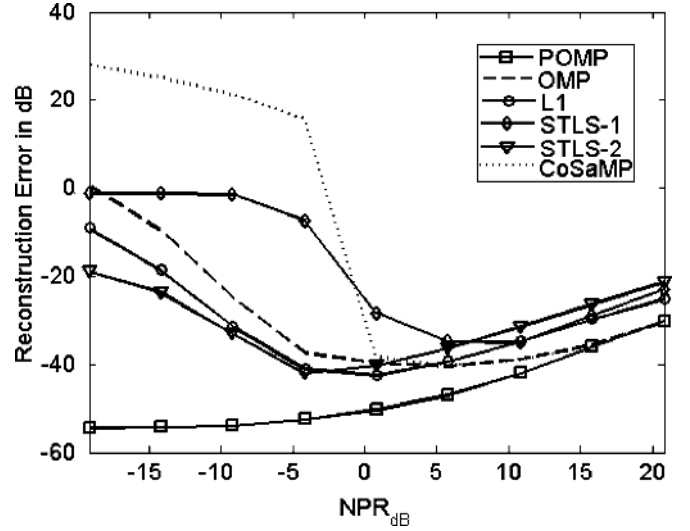

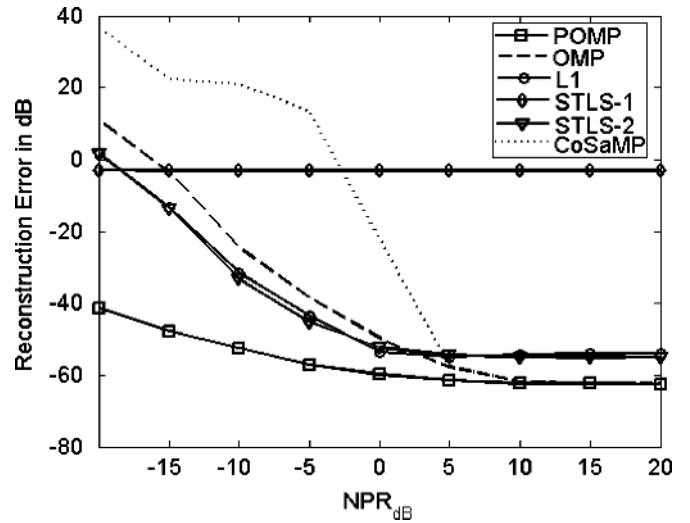Fig. 13. Signal reconstruction error in dB for fixed $\mathrm{SPR}_{\mathrm{dB}} = 40$ dB and varying SNR from 20 dB to 60 dB.

Fig. 14. Signal reconstruction error in dB for fixed $\mathrm{SNR}_{\mathrm{dB}} = 50$dB and varying SPR from 30 dB to 70 dB.

observation $\boldsymbol{b}$ has the form of $\boldsymbol{b} = \boldsymbol{Ax} + \boldsymbol{n} = \boldsymbol{b}_0 + \boldsymbol{n}$. If we write the residual as,

$$\|\boldsymbol{Ax} - \boldsymbol{b}\|_2 = \|\boldsymbol{Ax} - \boldsymbol{b}_0 - \boldsymbol{n}\|_2 \simeq E\{\|\boldsymbol{n}\|_2\}, \qquad (28)$$

hence $\|\boldsymbol{Ax} - \boldsymbol{b}\|_2 / \|\boldsymbol{Ax}\|_2 \simeq 1/\mathrm{SNR}$ if noise variance is small. Therefore, $\epsilon = 1/\mathrm{SNR}$ can be chosen. More specifically, if the noise has a variance σ^2, then $\epsilon = \sigma\sqrt{M}/\|\boldsymbol{b}_0\|_2$. Since the signals are error free in previous simulations, $\epsilon = 10^{-3}$ is chosen. However, if the signal comes from a perturbed dictionary, then $\boldsymbol{b} = (\boldsymbol{A} + \delta\boldsymbol{A})\boldsymbol{x} + \boldsymbol{n}$, hence,

$$\|\boldsymbol{Ax} - \boldsymbol{b}\|_2 = \|\boldsymbol{Ax} - \boldsymbol{Ax} - \delta\boldsymbol{Ax} - \boldsymbol{n}\|_2$$
$$\simeq \sqrt{E\left\{\|\boldsymbol{n}\|_2^2\right\} + E\left\{\|\delta\boldsymbol{Ax}\|_2^2\right\}}, \qquad (29)$$

where noise and perturbations are assumed to be independent. It is appropriate to define

$$\mathrm{SPR} = \|\boldsymbol{Ax}\|_2 / \|\delta\boldsymbol{Ax}\|_2, \qquad (30)$$

as the Signal-to-Perturbation Ratio (SPR). Therefore according to (29), $\|\boldsymbol{Ax} - \boldsymbol{b}\|_2 / \|\boldsymbol{b}\|_2 \simeq \sqrt{1/\mathrm{SNR} + 1/\mathrm{SPR}}$. In our framework, $\|\boldsymbol{a}_{i,p}^T \boldsymbol{a}_i\|_2 = \cos(\phi_i)$, since $\|\boldsymbol{a}_{i,p}\|_2 = \|\boldsymbol{a}_i\|_2 = 1$. Using simple geometry, the perturbation amount can be written as $\|\boldsymbol{a}_{i,p} - \boldsymbol{a}_i\|_2 = 2\sin(\phi_i/2)$. Hence, the perturbation amount is $E\{\|\delta\boldsymbol{Ax}\|_2\} \simeq \|\boldsymbol{b}\|_2 \times E\{\phi\}$, where $E\{\phi\} = 2\sin(60° \sqrt{3} E\{f\} \delta T/2)$. Therefore, $\mathrm{SPR}_{\mathrm{dB}} = 20\log(\|\boldsymbol{b}\|_2 / (E\{\phi\} \|\boldsymbol{b}\|_2)) = 20\log(1/E\{\phi\})$. If the perturbations are small, it can be approximated as:

$$\mathrm{SPR}_{\mathrm{dB}} = -20\log(\pi E\{f\}/\sqrt{3}) - 20\log(\delta T). \qquad (31)$$

In the following simulations, we aim to analyze the SNR and SPR regions in which POMP and compared techniques can successfully work. For this purpose sparsity is fixed at $K = 10$ and average reconstruction performance is found for varying SNR and SPR levels.

In Fig. 13, the perturbation level is kept constant at $\mathrm{SPR}_{\mathrm{dB}} = 40$ dB, which corresponds to $\delta T = 25$ μsec according to (31), and the SNR is varied from 20 dB to 60 dB. In Fig. 14, we keep the noise level constant at $\mathrm{SNR}_{\mathrm{dB}} = 50$ dB and the $\mathrm{SPR}_{\mathrm{dB}}$ is varied from 30 dB to 70 dB. In both figures, the reconstruction

performances are shown as a function of the Noise-to-Perturbation Ratio (NPR) defined as $\mathrm{NPR}_{\mathrm{dB}} = \mathrm{SPR}_{\mathrm{dB}} - \mathrm{SNR}_{\mathrm{dB}}$. It is clear that two distinct performance regions are determined by the sign of $\mathrm{NPR}_{\mathrm{dB}}$.

As shown in Fig. 13, although POMP has the lowest reconstruction error, all compared techniques have acceptable performances for positive $\mathrm{NPR}_{\mathrm{dB}}$. In that regime, noise power dominates the perturbation on the dictionary. Since the perturbations are relatively insignificant, OMP and CoSaMP can achieve similar results to POMP. However, when $\mathrm{NPR}_{\mathrm{dB}}$ is negative, the perturbation on the dictionary is more dominant than the noise level. OMP, BP and CoSaMP have no mechanism to handle these perturbations, hence they treat the perturbations as additional noise. However, in this case it is highly improbable to decrease the residual error to the actual noise level. Thus, they saturate and produce reconstructions with larger errors. On the other hand, even if perturbations dominate the noise, POMP can handle perturbations and produces reconstruction with significantly lower errors. Fig. 14 displays similar results, where SNR is fixed

and SPR is varied. For high NPR regime, all algorithms result in similar errors, whereas for low NPR, only POMP produces acceptable reconstructions. Higher perturbation causes larger reconstruction errors in POMP. However, it is much more robust to dictionary mismatch than the compared techniques.

IV. CONCLUSION

Compressive sensing reconstruction techniques suffer from significant performance degradation when there is a mismatch between the actual and the chosen signal dictionaries. Such mismatch generally occurs in practice due to modeling errors, parameter space discritization or simply by sampling jitter. In general, this mismatch problem cannot be solved using denser basis. Also, density of the basis adversely affects the algorithms due to high correlation. In this paper, a novel perturbed greedy reconstruction technique is proposed for the case of signal reconstruction in the presence of perturbations in the signal dictionary. The proposed Perturbed OMP (POMP) technique performs controlled perturbations of selected support vectors. Limits on the required perturbation for exact fit to observation signal at any sparsity is found. Also, for a given acceptable residual level, limits on the required perturbation are also provided. A sufficiency condition that assures monotonic decrease on the required perturbations as a function of sparsity is given. The performance of the proposed algorithm is investigated on a time jitter example that causes a random dictionary mismatch. As an extension to SNR, Signal-to-Perturbation (SPR) and Noise-to-Perturbation (NPR) are used for better characterization of the performance limits of POMP. Results show that, in comparison with well known CS reconstruction techniques, POMP provides efficient reconstructions with significantly lower reconstruction error in a wide range of sparsity levels.

APPENDIX
PROOF OF THEOREM 3

Let $\boldsymbol{S}_k = [\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_k]$ be the matrix whose columns correspond to the current estimate for the support of $\boldsymbol{x}$. Since $\boldsymbol{S}_k$ is full rank, we can define $\boldsymbol{M}_k = (\boldsymbol{S}_k^T \boldsymbol{S}_k)^{-1}$ and $\boldsymbol{S}_k^\dagger = \boldsymbol{M}_k \boldsymbol{S}_k^T$. The following equations provide recursive relations for $\boldsymbol{M}_{k+1}$ and $\boldsymbol{S}_{k+1}^\dagger$:

$$\boldsymbol{M}_{k+1} = \begin{pmatrix} \boldsymbol{M}_k + d(\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1})(\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1})^T & -d\left(\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}\right) \\ -d\left(\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}\right)^T & d \end{pmatrix},$$

(32)

$$\boldsymbol{S}_{k+1}^\dagger = \boldsymbol{M}_{k+1} \boldsymbol{S}_{k+1}^T = \begin{pmatrix} \boldsymbol{S}_k^\dagger - d\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1} \boldsymbol{s}_{k+1,r}^T \\ d\,\boldsymbol{s}_{k+1,r}^T \end{pmatrix}, \quad (33)$$

where $d = 1/\|\boldsymbol{s}_{k+1,r}\|_2^2$ and $\boldsymbol{s}_{k+1,r} = \boldsymbol{s}_{k+1} - \boldsymbol{S}_k \boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}$. Since the ℓ_1 norm of a partitioned vector is the sum of the ℓ_1 norms of each partition:

$$\|\boldsymbol{S}_{k+1}^\dagger \boldsymbol{b}\|_1 = \left\| \boldsymbol{S}_k^\dagger \boldsymbol{b} - d\,\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}\,\boldsymbol{s}_{k+1,r}^T \boldsymbol{b} \right\|_1 + \left\| d\,\boldsymbol{s}_{k+1,r}^T \boldsymbol{b} \right\|_1,$$

(34)

which can be written as:

$$\|\boldsymbol{S}_{k+1}^\dagger \boldsymbol{b}\|_1 = \|\boldsymbol{S}_k^\dagger(\boldsymbol{b} - \boldsymbol{u})\|_1 + |\lambda|, \quad (35)$$

where $\lambda = d\,\boldsymbol{s}_{k+1,r}^T \boldsymbol{b}$ and $\boldsymbol{u} = \lambda\,\boldsymbol{s}_{k+1}$.

To have $\phi_{k+1}^* \leq \phi_k^*$, we need the following inequality to hold true:

$$\frac{\left\|\boldsymbol{\alpha}^{(k)}\right\|_1}{\left\|\boldsymbol{\alpha}^{(k+1)}\right\|_1} \leq \frac{\left\|\boldsymbol{b}_\perp^{(k)}\right\|_2}{\left\|\boldsymbol{b}_\perp^{(k+1)}\right\|_2} = \beta. \quad (36)$$

Using the triangle inequality on (35), the following bound can be obtained,

$$|\lambda| - \|\boldsymbol{S}_k^\dagger \boldsymbol{u}\|_1 \leq \|\boldsymbol{S}_{k+1}^\dagger \boldsymbol{b}\|_1 - \|\boldsymbol{S}_k^\dagger \boldsymbol{b}\|_1. \quad (37)$$

Since $\|\boldsymbol{S}_k^\dagger \boldsymbol{u}\|_1 = \|\boldsymbol{S}_k^\dagger \lambda\,\boldsymbol{s}_{k+1}\|_1 = |\lambda|\,\|\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}\|_1$, (37) becomes:

$$|\lambda|\,(1 - \|\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}\|_1) \leq \|\boldsymbol{S}_{k+1}^\dagger \boldsymbol{b}\|_1 - \|\boldsymbol{S}_k^\dagger \boldsymbol{b}\|_1. \quad (38)$$

By adding $(\beta - 1)\|\boldsymbol{S}_{k+1}^\dagger \boldsymbol{b}\|_1$ to both sides of (38), we obtain:

$$(\beta - 1)\|\boldsymbol{S}_{k+1}^\dagger \boldsymbol{b}\|_1 + |\lambda|\,(1 - \|\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}\|_1)$$
$$\leq \beta\|\boldsymbol{S}_{k+1}^\dagger \boldsymbol{b}\|_1 - \|\boldsymbol{S}_k^\dagger \boldsymbol{b}\|_1. \quad (39)$$

The desired condition on angle decrease requires $\beta\|\boldsymbol{S}_{k+1}^\dagger \boldsymbol{b}\|_1 \geq \|\boldsymbol{S}_k^\dagger \boldsymbol{b}\|_1$, which is always achieved if:

$$0 \leq (\beta - 1)\|\boldsymbol{S}_{k+1}^\dagger \boldsymbol{b}\|_1 + |\lambda|\,(1 - \|\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}\|_1), \quad (40)$$

or equivalently:

$$\|\boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}\|_1 \leq 1 + \frac{(\beta - 1)\left\|\boldsymbol{\alpha}^{(k+1)}\right\|_1}{|\lambda|}. \quad (41)$$

Therefore, by using (36), (41) and definition of λ, the non-negative constant γ in the statement of Theorem 3 can be obtained as:

$$\gamma$$
$$= \frac{\left(\dfrac{\left\|\boldsymbol{b}_\perp^{(k)}\right\|_2}{\left\|\boldsymbol{b}_\perp^{(k+1)}\right\|_2} - 1\right)\left\|\boldsymbol{\alpha}^{(k+1)}\right\|_1 \|\boldsymbol{s}_{k+1} - \boldsymbol{S}_k \boldsymbol{S}_k^\dagger \boldsymbol{s}_{k+1}\|_2^2}{\left|\boldsymbol{s}_{k+1}^T \boldsymbol{b}_\perp^{(k)}\right|},$$
$$\text{q.e.d.} \quad (42)$$

REFERENCES

[1] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
[2] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, pp. 489–509, 2006.
[3] M. D. ad, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, 2008.
[4] M. Lustig, D. Donoho, and J. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, Dec. 2007.

[5] R. Baraniuk and P. Steeghs, "Compressive radar imaging," in *Proc. IEEE Radar Conf.*, 2007, pp. 128–133.

[6] V. Patel, G. Easley, D. H. Jr., and R. Chellappa, "Compressed synthetic aperture radar," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 244–254, 2010.

[7] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressive wireless sensing," in *Proc. Int. Conf. Inf. Process. Sens. Netw. (IPSN)*, Apr. 2006, pp. 134–142.

[8] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.

[9] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4036–4048, 2006.

[10] D. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[11] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, 1993.

[12] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[13] D. Needell and J. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Computat. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, 2009.

[14] T. Blumensath, M. Yaghoobi, and M. Davies, "Iterative hard thresholding and ℓ_0 regularisation," in *Proc. ICASSP*, 2007, vol. 3, pp. 877–880.

[15] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18 914–18 919, Nov. 2009.

[16] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hedge, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.

[17] S. Som and P. Schniter, "Compressive imaging using approximate message passing and a Markov-tree prior," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3439–3448, Jul. 2010.

[18] V. Cevher, M. Duarte, and R. Baraniuk, "Distributed target localization via spatial sparsity," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2008, pp. 134–142.

[19] J. H. Ender, "On compressive sensing applied to radar," *Signal Process.*, vol. 90, no. 5, pp. 1402–1414, 2010.

[20] A. C. Gurbuz, J. H. McClellan, and W. R. Scott Jr., "A compressive sensing data acquisition and imaging method for stepped frequency GPRS," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2640–2650, 2009.

[21] A. C. Gurbuz, V. Cevher, and J. H. McClellan, "A compressive beamformer," in *Proc. ICASSP 2008*, Las Vegas, NV, Mar. 30–Apr. 4 2008.

[22] Y. Yu, A. P. Petropulu, and H. V. Poor, "MIMO radar using compressive sampling," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 1, pp. 146–163, 2010.

[23] N. Aggarwal and W. C. Karl, "Line detection in images through regularized Hough transform," *IEEE Trans. Image Process.*, vol. 15, pp. 582–590, 2006.

[24] M. A. Tuncer and A. C. Gurbuz, "Analysis of unknown velocity and target off the grid problems in compressive sensing based subsurface imaging," in *Proc. ICASSP 2011*, Prague, Czech Republic, pp. 2880–2883.

[25] Y. Chi, L. Scharf, A. Pezeshki, and R. Calderbank, "The sensitivity to basis mismatch of compressed sensing in spectrum analysis and beamforming," in *Proc. 6th Workshop on Defense Appl. Signal Process. (DASP)*, Lihue, HI, Oct. 2009.

[26] R. Baraniuk, M. Davenport, R. Devore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constr. Approx.*, vol. 2008, 2007.

[27] M. Herman and T. Strohmer, "General deviants: An analysis of perturbations in compressed sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 342–349, 2010.

[28] Y. Chi, L. Scharf, A. Pezeshki, and R. Calderbank, "Sensitivity of basis mismatch to compressed sensing," *IEEE Trans. Signal Process.*, vol. 59, pp. 2182–2195, 2011.

[29] D. H. Chae, P. Sadeghi, and R. A. Kennedy, "Effects of basis-mismatch in compressive sampling of continuous sinusoidal signals," in *Int. Conf. Future Comput. Commun. (ICFCC)*, 2010, pp. 739–743.

[30] G. Peyre, "Best basis compressed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2613–2622, May 2010.

[31] C. Ekanadham, D. Tranchina, and E. Simoncelli, "Recovery of sparse translation-invariant signals with continuous basis pursuit," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4735–4744, Oct. 2011.

[32] Z. Yang, C. Zhang, and L. Xie, "Robustly stable signal recovery in compressed sensing with structured matrix perturbation," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4658–4671, Sep. 2012.

[33] H. Zhu, G. Leus, and G. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2002–2016, 2011.

[34] O. Teke, A. C. Gurbuz, and O. Arikan, "A new omp technique for sparse recovery," in *Proc. 20th Signal Process. Commun. Appl. Conf. (SIU)*, Apr. 2012, pp. 1–4.

[35] M. Pilanci, O. Arikan, and M. Pinar, "Structured least squares problems and robust estimators," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2453–2465, May 2010.

[36] J. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[37] M. Elad, *Sparse and Redundant Representations*. New York, NY, USA: Springer, 2010.

[38] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, 2013.

Oguzhan Teke (S'12) was born in 1990 in Tokat, Turkey. In 2012, he received the B.Sc. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey.

He is currently working toward the M.S. degree under the supervision of Prof. Orhan Arikan in the Department of Electrical and Electronics Engineering, Bilkent University. His research interests are in applied linear algebra and numerical analysis, inverse problems, and compressed sensing.

Ali Cafer Gurbuz (M'08) received the B.Sc. degree from Bilkent University, Ankara, Turkey, in 2003 in electrical and electronics engineering, and the M.S. and Ph.D. degrees from the Georgia Institute of Technology (Georgia Tech), Atlanta, in 2005 and 2008, both in electrical and computer engineering.

From 2003 to 2008, he participated in multimodal landmine detection system research as a Graduate Research Assistant and from 2008 to 2009, as Postdoctoral Fellow, all with Georgia Tech. He is currently an Associate Professor with TOBB University of Economics and Technology, Ankara, with the Department of Electrical and Electronics Engineering. His research interests include compressive sensing applications, ground penetrating radar, array signal processing, remote sensing, and imaging.

Orhan Arikan (M'91) was born in 1964 in Manisa, Turkey. In 1986, he received the B.Sc. degree in electrical and electronics engineering from the Middle East Technical University, Ankara, Turkey. He received the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois, Urbana-Champaign, in 1988 and 1990, respectively.

Following his graduate studies, he was a Research Scientist with Schlumberger-Doll Research Center, Ridgefield, CT. In 1993, he joined the Electrical and Electronics Engineering Department, Bilkent University, Ankara, and since 2011, has been Department chairman. His current research interests include statistical signal processing, time frequency analysis and remote sensing.

Dr. Arikan has served as Chairman of IEEE Signal Processing Society Turkey Chapter and President of IEEE Turkey Section.