Compressive Sensing Recovery of Spike Trains Using A Structured Sparsity Model

Chinmay Hegde, Marco F. Duarte, Volkan Cevher

Compressive Sensing Recovery of Spike Trains Using A Structured Sparsity Model

Chinmay Hegde, Marco F. Duarte, and Volkan Cevher
Deparment of ECE
Rice University
Houston, Texas 77098
Email: {chinmay, duarte, volkan}@rice.edu

***Abstract*—The theory of Compressive Sensing (CS) exploits a well-known concept used in signal compression – sparsity – to design new, efficient techniques for signal *acquisition*. CS theory states that for a length-N signal x with sparsity level K, $M = O(K \log(N/K))$ random linear projections of x are sufficient to robustly recover x in polynomial time. However, richer models are often applicable in real-world settings that impose additional *structure* on the sparse nonzero coefficients of x. Many such models can be succinctly described as a union of K-dimensional subspaces. In recent work, we have developed a general approach for the design and analysis of robust, efficient CS recovery algorithms that exploit such signal models with structured sparsity.**

We apply our framework to a new signal model which is motivated by *neuronal spike trains*. We model the firing process of a single Poisson neuron with absolute refractoriness using a union of subspaces. We then derive a bound on the number of random projections M needed for stable embedding of this signal model, and develop a algorithm that provably recovers any neuronal spike train from M measurements. Numerical experimental results demonstrate the benefits of our model-based approach compared to conventional CS recovery techniques.

I. Introduction

Many methods for signal compression are commonly based on the *transform coding* approach. In such methods, the assumption is that a signal $x \in \mathbb{R}^N$ can be represented as a *sparse* linear combination of elements from a fixed, known basis $\Psi \in \mathbb{R}^{N \times N}$. In other words, $x = \Psi\alpha$, where the number of nonzero elements K of the transform coefficient vector α is much smaller than N. Sparsity is fundamental to our understanding and processing of several real-world signals; for instance, piecewise smooth signals and images are compressible in the wavelet basis [1]. Thus a transform coder operates on a signals of size N and obtains a condensed representation at its "true" information rate K.

Compressive Sensing (CS) [2, 3] offers an intriguing alternative to this classical process of acquiring and compressing signals. As opposed to uniform signal samples, a signal may be sampled by measuring its inner product with $M \ll N$ vectors. Therefore, $y = \Phi x = \Phi\Psi\alpha$ where $\Phi \in \mathbb{R}^{M \times N}$ is a non-invertible matrix. Interestingly, if α is K-sparse and if the entries of the matrix Φ are chosen *randomly* from certain types of probability distributions, CS theory dictates that if $M = O(K \log(N/K))$, the coefficient vector α, and consequently, the signal x, can be exactly reconstructed from the measurements y using efficient recovery algorithms, such as convex programming techniques and greedy methods [4–7].

Sparsity is a popular model for compression and tractable processing of several interesting classes of signals. Nonetheless, the sparse signal model can be termed as simplistic in the sense that it assumes *no additional structure* about the inter-relations between the transform coefficients α. On the other hand, consider the class of one-dimensional piecewise smooth signals; we not only know that such signals are sparse in the wavelet basis, but also that the wavelet coefficients lie approximately on a connected tree [1]. This has given rise to the design and development of sophisticated compression algorithms that operate on a given signal x according to *structured sparsity models*.

In recent work [8], we have developed a comprehensive framework that leverages additional structure in sparse models to develop novel methods for CS recovery. Our framework offers a systematic method for designing provably efficient CS reconstruction algorithms for signals belonging to these models. We have examined a number of instances of our framework; in particular, we have shown that for the connected wavelet tree model, our recovery algorithm requires merely $M = O(K)$ measurements for robust reconstruction.

In this paper, as a specific instance of our approach, we study the compressive acquisition of neuronal spike trains [9]; particularly, we are interested in temporal point processes in which consecutive spikes occur with a time delay no smaller than a known quantity Δ. We introduce a empirically motivated union-of-subspaces model for this class of signals. We compute a bound on the minimum number of measurements required to stably project this set into a lower dimensional subspace. Additionally, we develop a new algorithm for recovery of neuronal signals from these measurements. An interesting consequence of our analysis is that the number of measurements required for robust recovery scales as $M = O(K \log(N/K - \Delta))$, i.e., M *decreases* with increasing minimum inter-arrival time Δ.

The rest of the paper is organized as follows. Section II provides a brief review of the mathematical theory of compressive sensing, and describes our new model-based framework for CS recovery. In Section III, we provide an overview of neuronal spike trains, model this class of signals as a union of subspaces and use our framework to formulate a novel algorithm for CS recovery with provable guarantees. Experimental results that demonstrate the utility of our method are presented in Section IV. Section V lists our conclusions.

II. Background

A. Sparsity as a union of subspaces

Given a signal $x \in \mathbb{R}^N$ and a basis $\Psi \in \mathbb{R}^{N\times N}$, we may represent x in terms of its basis coefficients α, so that $x = \Psi\alpha$. We say that x is *K-sparse* in Ψ if no more than $K \ll N$ coefficients of α are nonzero. In the rest of the paper, we assume that the sparsity basis Ψ is the identity matrix (i.e., $x = \alpha$), while noting that the results are conceptually valid for general Ψ. The *support* of x is defined as the set of indices corresponding to nonzero entries of x; this can alternately be mapped to a binary vector $s(x)$ of length N with no greater than K entries being equal to 1.

Consider the set Σ_K of *all* K-sparse signals. It is easy to identify this set as the union of $\binom{N}{K}$ K-dimensional subspaces of $\mathbb{R}^N$, with each subspace being equivalent to the linear span of exactly K canonical unit vectors in $\mathbb{R}^N$. The notion of *compressibility* is based on this geometric intuition; we say that a signal x is compressible if it lies *close* to this union of subspaces Σ_K. Given a compressible signal x, compression techniques are interested in a K-sparse signal x_K so that $\|x - x_K\|_2$ is minimized, with $\|\cdot\|_2$ being the ℓ_2-norm. This is achieved by a simple approximation procedure $\mathbb{T}(x, K)$, that basically selects the K largest coefficients of x.

In many situations, we possess some additional information about the support of a sparse signal x. For example, suppose we are interested in K-sparse signals with only a few permitted configurations of $s(x)$. This defines a *union of subspaces model* $\mathcal{M}_K$ consisting of only m_K canonical K-dimensional subspaces of $\mathbb{R}^N$, with $m_K < \binom{N}{K}$. Let $x|_\Omega$ represent the entries of x corresponding to the set of indices $\Omega \subseteq \{1, \ldots, N\}$, and let Ω^C denote the complement of the set Ω. Then, $\mathcal{M}_K$ is defined as:

$$\mathcal{M}_K = \bigcup_{m=1}^{m_K} \mathcal{X}_m,\ \mathcal{X}_m := \{x : x|_{\Omega_m} \in \mathbb{R}^K, x|_{\Omega_m^C} = 0\}, \quad (1)$$

where each subspace $\mathcal{X}_m$ contains all signals x with $\mathrm{supp}(x) \in \Omega_m$. Thus, $\mathcal{M}_K$ is characterized by the set of permitted supports $\{\Omega_1, \ldots, \Omega_{m_K}\}$.

In light of this definition, we view any such union of subspaces as a *structured sparsity model*. As in the general K-sparse case, given a signal x, we seek a signal x^* such that $x^* \in \mathcal{M}_k$, and $\|x - x^*\|_2$ is minimized. We define a *model-approximation* algorithm as a procedure $\mathbb{M}(x, K)$ which returns the best K-term approximation of a given signal under the model $\mathcal{M}_K$, i.e., $x^* = \mathbb{M}(x, K)$. In addition, we note the set $\mathcal{M}_{2K} := \{x - y : x, y \in \mathcal{M}_K\}$ forms a union of $2K$-dimensional subspaces. For reasons described in the sequel, we also define an approximation procedure $\mathbb{M}_2(x, 2K)$ as a function that computes x_2^*, the best approximation of x in this difference set $\mathcal{M}_{2K}$, i.e., $x_2^* = \mathbb{M}_2(x, 2K)$.

B. A brief review of Compressive Sensing

Compressive Sensing (CS) arguably represents a paradigm shift in the way we sample and process signals [2, 3]. In essence, CS exploits prior knowledge about the sparsity of the signal of interest x to greatly reduce sampling rates, while guaranteeing stable reconstruction of x from its samples. In CS, we do not observe a K-sparse signal x directly; instead we record $M < N$ nonadaptive linear measurements $y = \Phi x$, where $\Phi \in \mathbb{R}^{M\times N}$ is a measurement matrix. The central premise in CS is that in specific circumstances, x can be efficiently and accurately reconstructed from y even though Φ possesses a nontrivial nullspace. In particular, this is possible if Φ satisfies the *restricted isometry property* (RIP):

Definition 1: [2] An $M \times N$ matrix Φ has the K-RIP with constant δ_K if, for all $x \in \Sigma_K$,

$$(1 - \delta_K)\|x\|_2^2 \le \|\Phi x\|_2^2 \le (1 + \delta_K)\|x\|_2^2. \quad (2)$$

In other words, we only desire that all submatrices of size $M \times K$ of Φ are approximate isometric transformations, so that the norms of K-sparse signals are (approximately) preserved. Practical CS recovery algorithms require Φ to satisfy the slightly stronger bK-RIP, so that the norms of b-wise differences of K-sparse signals are preserved as well; here, b is a small integer (typically 2 or 3.)

CS is characterized by two hallmarks. The first hallmark involves the *design* of desirable sampling operators Φ. The design of a matrix Φ having the K-RIP with a prescribed constant δ_K is NP-complete [2]; nevertheless, random matrices whose entries are i.i.d. subgaussian random variables[1] work with high probability provided $M = O(K \log(N/K))$. Thus, the number of required samples scales linearly with signal sparsity, and is only logarithmic in signal length.

The second hallmark addresses the issue of *feasible, stable recovery* of a signal from its measurements. The development of efficient algorithms for CS recovery has received considerable attention in the literature [4–6]. More recently, algorithms based on the tenet of iterative sparse approximation [7, 10] have been demonstrated to yield uniform, stable guarantees for signal recovery while expending minimum computational resources. In particular, given noisy measurements of any signal $x \in \mathbb{R}^N$ so that $y = \Phi x + n$, if Φ is known to possess RIP, then the signal estimate $\widehat{x}$ obtained by these algorithms [7, 10, 11] is given by:

$$\|x - \widehat{x}\|_2 \le C_1\|x - x_K\|_2 + \frac{C_2}{\sqrt{K}}\|x - x_K\|_1 + C_3\|n\|_2,$$

where x_K is the best K-sparse approximation to x and C_1, C_2 are constants. An important implication of this result is that given noiseless measurements, a K-sparse signal can be recovered perfectly using these algorithms.

C. Compressive sensing recovery using structured sparsity

CS principles exploit signal sparsity in order to develop efficient signal sampling and reconstruction methods. A very natural question to ask is this: is it possible to develop analogous sampling methods *as well as* recovery algorithms for structured sparsity models? Several approaches have been

[1] A random variable X is called subgaussian if there exists $c > 0$ such that $\mathbb{E}\left(e^{Xt}\right) \le e^{c^2t^2/2}$ for all $t \in \mathbb{R}$. Examples include the Gaussian and Bernoulli random variables, as well as any bounded random variable.

adopted in the literature to address this issue [12–15]. However, these approaches exhibit one or more of the following symptoms: (i) they are based on heuristics and lack mathematical rigor; (ii) they do not provide uniform guarantees; (iii) they are tailored to specific signal models.

For target signals x that belonging to a structured sparsity model, we may impose a less stringent constraint on the CS measurement matrix Φ and still achieve approximate isometry in the compressive measurements y. This gives rise to the notion of a *model-based* RIP which requires that (2) holds only for signals $x \in \mathcal{M}_K$ [14, 16]; we denote this new property as $\mathcal{M}_K$-RIP to specify the dependence on the chosen signal model. A recent result [14] quantifies the number of measurements M necessary for a subgaussian CS matrix to have the $\mathcal{M}_K$-RIP with constant $\delta_{\mathcal{M}_K}$ and with probability $1 - e^{-t}$ to be

$$M \geq \frac{2}{c\delta^2_{\mathcal{M}_K}} \left(\ln(2m_K) + K \ln \frac{12}{\delta_{\mathcal{M}_K}} + t \right). \qquad (3)$$

This bound can be used to recover the conventional CS result by substituting $m_K = \binom{N}{K} \approx (Ne/K)^K$.

In previous work [8], we have shown how this result could be used to design a provably robust algorithm for CS recovery of signals in $\mathcal{M}_K$. A quick description of our approach is as follows. Consider CoSaMP [7], an efficient greedy algorithm that provides state-of-the-art guarantees. CoSaMP possesses an iterative structure that hinges on computing the best K-term approximation to an intermediate signal estimate x_j; in other words, the algorithm requires computing $\mathbb{T}(x, K)$. We simply replace this sparse approximation step by a best *model-approximation* step, i.e., we compute $\mathbb{M}(x_j, K)$. Our recovery method is detailed in pseudocode form in Algorithm 1.

In this way, given an efficient (polynomial time) algorithm $\mathbb{M}(x, K)$ that performs pruning of a signal x according to a given signal model $\mathcal{M}_K$, Algorithm 1 can perform efficient *reconstruction* of a signal belonging to the model, given M measurements as specified in Equation 3. In cases where $m_K \ll \binom{N}{K}$, our algorithm provides provable reconstruction from measurements sampled proportional to the *information rate* K of the signal, i.e., $M = O(K)$. The sampling bound (3) and the associated recovery algorithm are for signals exactly lying on the union of subspaces $\mathcal{M}_K$; in [8], we develop parallel theorems that prove robustness to noise in the measurements y as well as model mismatch.

Summarizing, we possess the recipe to develop a stable, efficient algorithm for CS recovery of signals belonging to any structured sparsity model, provided we have the following ingredients:

Formulation of model: we need to establish the signal model as a well-defined union of subspaces $\mathcal{M}_K$.

A sampling bound: we need to calculate the number of subspaces in the model m_K in order to obtain a requisite number of random measurements M for stable embedding.

An approximation algorithm: given an arbitrary $x \in \mathbb{R}^N$, we need feasible methods to compute the model approximations $\mathbb{M}(x, K)$ and $\mathbb{M}_2(x, 2K)$.

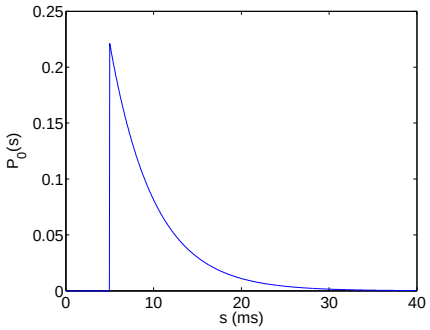

Fig. 1. *Interval distribution $P_0(s)$ for a Poisson process with refractoriness $\Delta = 5$ ms and $v = 0.2$.*

III. A MODEL FOR NEURONAL SPIKE TRAINS

As a practical example of our approach, we study a representative structured sparsity model for spike trains emitted by a single neuron; see [9] for a detailed statistical analysis of such signals. We consider one-dimensional spike trains generated by a stochastic neuronal firing process that can be modeled as an input-dependent renewal system. A key concept in the statistical description of such a process is the *inter-spike interval distribution* $P_0(s)$, which is defined as the probability density function of the inter-arrival time s between consecutive spikes; for stationary processes, $P_0(s)$ contains all the information required to describe the system. A related quantity is the *hazard function* $\rho_0(s)$, which is related to $P_0(s)$ as follows: for $s > 0$

$$\rho_0(s) = \frac{P_0(s)}{1 - \int_0^s P_0(t)dt}.$$

Suppose we observe a homogenous Poisson process with rate ν. It is well known that the interval distribution of this process is given by the exponential $P_0(s) = \nu e^{-\nu s}$, and the hazard function is given by the constant function $\rho(s) = \nu$. Notice that the interval distribution is maximum at $\nu = 0$, i.e., this model dictates that spikes are very likely to be generated within a very short interval of each other. In contrast, empirical studies [9] have shown that real neurons indicate refractoriness, i.e., the interval distribution vanishes as $s \to 0$, which implies that there exists a minimum nonzero time delay between consecutive spikes. Thus, a Poisson process with *absolute refractoriness* Δ is defined as a process with interval distribution as follows:

$$P_0(s) = \begin{cases} 0 & \text{for } s \leq \Delta, \\ \nu e^{-\nu(s-\Delta)} & \text{for } s > \Delta, \end{cases}$$

The associated hazard function remains zero for $0 \leq s \leq \Delta$ and then jumps to a constant ν for $s > \Delta$. A plot of the interval distribution for such a process is displayed in Figure 1.

A. Ingredient 1: the model

Consider the discrete time analogue of the above process, so that we observe a length-N time signal with K spikes (of unknown magnitude). If there were no restrictions on the

Algorithm 1 Model-based CoSaMP

Inputs: Projection matrix Φ, measurements y, model approximation algorithm $\mathbb{M}_K$
Output: K-sparse approximations $\widehat{x}$ to true signal x
$\widehat{x}_0 = 0$, $r = y$; $i = 0$ {initialize}
while halting criterion false **do**
1. $i \leftarrow i + 1$
2. $e \leftarrow \Phi^T r$ {form signal residual estimate}
3. $\Omega \leftarrow \text{supp}(\mathbb{M}_2(e, 2K))$ {prune merged signal residual estimate according to model}
4. $T \leftarrow \Omega \cup \text{supp}(\widehat{x}_{i-1})$ {merge supports}
5. $b|_T \leftarrow \Phi_T^{\dagger} y$, $b|_{T^C}$ {form signal estimates}
8. $\widehat{x}_i \leftarrow \mathbb{M}(b, K)$ {prune signal estimate according to model}
10. $r \leftarrow y - \Phi \widehat{x}_i$ {update measurement residual}

end while
return $\widehat{x} \leftarrow \widehat{x}_i$

locations of the spikes, it would correspond to the classical K-sparse model. However, if we require that no two consecutive spikes occur within a time interval Δ, we clearly obtain a structured sparsity model as defined in (1). Mathematically, we define the model as follows:

Definition 2: Suppose $r/2 < \Delta < N/K$. The (K, Δ, r)-*model* is defined as the set of all K-sparse one-dimensional signals $x \in \mathbb{R}^N$ such that no contiguous set of Δ consecutive locations contain *greater* than r spikes.

It is clear that the Poisson process with nonzero absolute refractoriness corresponds to a $(K, \Delta, 1)$-model.[2] We also state the following simple lemma without proof:

Lemma 1: If $\mathcal{M}_K$ is a (K, Δ, r)-model, then the set of pairwise differences $\mathcal{M}_{2K}$ is a $(2K, \Delta, 2r)$-model.

B. Ingredient 2: the sampling bound

The following theorem prescribes a number of random linear measurements that suffice for the stable embedding of all possible signals in a $(K, \Delta, 1)$-model.

Theorem 1: Let $\mathcal{M}_K$ be the $(K, \Delta, 1)$-model as defined above. Then, for any $t > 0$ and any

$$M \geq O\left(\frac{1}{\delta^2_{\mathcal{M}_K}}\left(K \log(N/K - \Delta) + K \ln \frac{1}{\delta_{\mathcal{M}_K}} + t\right)\right),$$

an $M \times N$ i.i.d. subgaussian random matrix has the $\mathcal{M}_K$-RIP with constant $\delta_{\mathcal{M}_K}$ with probability at least $1 - e^{-t}$.

Proof: Given x belonging to the model, we first observe that there are $K + 1$ contiguous blocks of zeros; blocks at either tail of the one-dimensional signal may be of zero size, while blocks in the interior must be at least of size $\Delta - 1$. We rewrite the binary support vector $s(x)$ in a run-length coding fashion; in other words, $s(x)$ maps to a length-N vector X such that $X = (X_1, \mathbf{1}, X_2, \ldots, \mathbf{1}, X_{K+1})$, where X_k denotes the *length* of the corresponding block of zeros and $\mathbf{1}$ denotes a spike. The following relations hold:

$$\begin{aligned} \sum_{k=1}^{K+1} X_k &= N - K, \\ X_1, X_{K+1} &\geq 0, \\ X_j &\geq \Delta - 1, j \neq \{1, K+1\}. \end{aligned} \tag{4}$$

Let $Y_1 = X_1, Y_{K+1} = X_{K+1}$ and $Y_j = X_j - (\Delta - 1)$ for $j \neq \{1, K+1\}$. Rewriting (4) in terms of Y_i, we observe that our answer m_K is nothing but the number of nonnegative integer solutions to the equation:

$$\sum_{k=1}^{K+1} Y_k = N - K\Delta + \Delta - 1.$$

From elementary combinatorics, this number is easily calculated as:

$$m_K = \binom{N - (K-1)(\Delta - 1)}{K}.$$

Substituting this value in (3) while ignoring constant factors, we derive the desired on M. □

The sampling bound for the $(2K, \Delta, 2)$ model can be similarly derived, and can be shown to be no more than twice the number of measurements prescribed in Theorem 1. We observe that as $\Delta \to 0$, we approach the bound proffered by conventional CS (i.e., $M = O(K \log(N/K))$). Significant advantages are achieved when Δ is large; hence, M decreases with increasing minimum interval time. Note that $\Delta > N/K$ is not possible, since we cannot have K spikes packed within a space of N locations under this condition.

C. Ingredient 3: the approximation algorithm

We now turn to the final step in the development of our CS recovery algorithm. Given an *arbitrary* signal $x \in \mathbb{R}^N$, we need to solve for the best $(K, \Delta, 1)$-approximation to x. Let $x = (x_1, \ldots, x_N)^{\top}$. If $s = (s_1, \ldots, s_N)$ is any binary support vector of length N, let:

$$x_{|s} := (s_1 x_1, s_2 x_2, \ldots, s_N x_N),$$

[2]We ignore further probabilistic dependencies among the locations of the spikes (such as expressed by the hazard parameter ν) and assume that every subspace belonging to this model is equally likely.

so that $x_{|s}$ is the portion of the signal x lying within the support. Our aim is to solve for that choice of support s so that s belongs to the $(K, \Delta, 1)$-model, and $\|x - x_{|s}\|_2$ is minimized. The following constraints on s follow from the definition of $(K, \Delta, 1)$-signals:

$$\begin{aligned} s_1 + s_2 + \ldots + s_N &\leq K, \\ s_1 + \ldots + s_\Delta &\leq 1, \\ s_2 + \ldots + s_{\Delta+1} &\leq 1, \\ &\ldots \\ s_{N-\Delta+1} + \ldots + s_N &\leq 1. \end{aligned}$$

Further, it is easy to see that minimizing $\|x - x_{|s}\|_2$ is equivalent to maximizing $c^\top s$ where $c = (x_1^2, x_2^2, \ldots, x_N^2)$, i.e., we maximize that portion of the signal energy that lies within s. The above optimization can be posed as an integer program as follows: let $W \in (N - \Delta + 2) \times N$ such that $W_{1j} = 1, j = \{1, \ldots, N\}$, $W_{2j} = 1, j = \{1, \ldots, \Delta\}$, $W_{3j} = 1, j = \{2, \ldots, \Delta + 1\}$ and so on; this represents the matrix inequality constraints. Next, define $u \in \mathbb{R}^{N-\Delta+2} = (K, 1, 1, \ldots, 1)$; this represents the RHS of the constraints. If we denote c to be the vector formed by the absolute values of the entries in x, we obtain the following integer program:

$$\begin{aligned} s^* &= \arg\min c^\top s, \\ Ws &\leq u. \end{aligned} \tag{5}$$

and s is binary. To solve the integer program, we make use of the following theorem.

Theorem 2: The solution to the binary integer program (5) is identical to the solution of the linear program obtained by relaxing the integer constraints.

Proof: First, we note that the binary matrix W is *totally unimodular* (TU), i.e., the determinant of every square submatrix of W is equal to 0, 1 or -1. This follows from the fact that W is a binary matrix with the 1's in every row occurring in consecutive blocks, i.e., W is a so-called "interval matrix"; interval matrices are well-known to be totally unimodular [17].

We next show that the polytope formed by the system of inequalities in the relaxed linear program has integer basic feasible solutions. Since W is TU, every basic feasible solution is determined by a nonsingular $r \times r$ submatrix of W, $r = N - \Delta + 2$. Since the determinant of this submatrix is ± 1 and the RHS of the linear system is integer, by Cramer's rule we get that all components of every basic feasible solution are integers. The desired result follows from this observation, since the optimum of the linear program has to occur in a basic feasible solution. Indeed, the solution has to be binary, since every s_i is positive and constrained to be lesser than 1. □

Thus, the approximation step $\mathbb{M}(x, K)$ in the neuronal spike model can efficiently be performed by solving a linear program (LP). The approximation $\mathbb{M}_2(x, 2K)$ can similarly be solved by a linear system similar to (5), with u being replaced by $2u$ on the right hand side. The number of variables in the linear program is equal to the dimension of the signal N; thus the model-approximation step can be performed in $O(N^{3.5})$ operations using state-of-the-art interior-point LP methods [17]. For very large N, this could be computationally expensive; however, it is possible that faster approximation algorithms can be developed to exploit the special Toeplitz-like structure of W.

D. Model-based recovery

Having developed all the ingredients of the algorithm, we insert the appropriate model approximation steps into Algorithm 1 to obtain a robust recovery algorithm for CS recovery of signals belonging to the $(K, \Delta, 1)$-model. Thus, we obtain the following theorem characterizing our derived algorithm.

Theorem 3: [8] Let x be a signal from the $(K, \Delta, 1)$-model and $\Phi \in \mathbb{R}^{M \times N}$. Let $y = \Phi x + n$ be a set of noisy CS measurements. If M is at least as great as specified by Theorem (1) with $\delta_{\mathcal{M}_{4K}} \leq 0.1$, the estimate $\widehat{x}$ obtained from iteration i of Algorithm 1 satisfies:

$$\|x - \widehat{x}_i\|_2 \leq 2^{-i}\|x\|_2 + 15\|n\|_2.$$

While we have presented results only for signals exactly lying on the given union-of-subspaces model, an analogous theorem can be developed for model-compressible signals using the theory described in [8]; we defer a full study of such signals to future work.

IV. Experiments

In this section, we present numerical results demonstrating the utility of our new algorithm. We generate spike trains of length N, and perform a model-approximation step to obtain a test signal which belongs to the $(K, \Delta, 1)$-model. The signal is measured via M random projections and reconstructed using Algorithm 1. For a baseline comparison, we use the reconstruction obtained by standard CoSaMP; the parameters K and Δ are assumed to be known.

Figure 2 indicates the potential performance gains of our algorithm. We measure a length-1024 neuronal signal with $K = 50$ spikes and minimum inter-arrival time $\Delta = 15$ using 150 random Gaussian measurements, and plot the error signal obtained by either recovery algorithm. Clearly, our method requires fewer measurements for accurate recovery compared to conventional sparse approximation.

Figure 3 illustrates the results of a Monte Carlo study on the effect of the number of measurements M on the conventional and model-based approaches. Each data point was generated using 300 sample trials. Successful recovery is defined as an instance when the solution is within an ℓ_2-distance of 1% relative to the original signal. We observe that our approach We observe that our method achieves successful recovery with probability over 95% with only $3.5K$ measurements, while CoSaMP can only achieve this with $M = 5K$.

V. Discussion

In this paper, we have described a general tool for developing algorithms for CS recovery when the signal of interest lies on a union of low-dimensional subspaces. The flexibility of our approach potentially enables to extend this general outline for

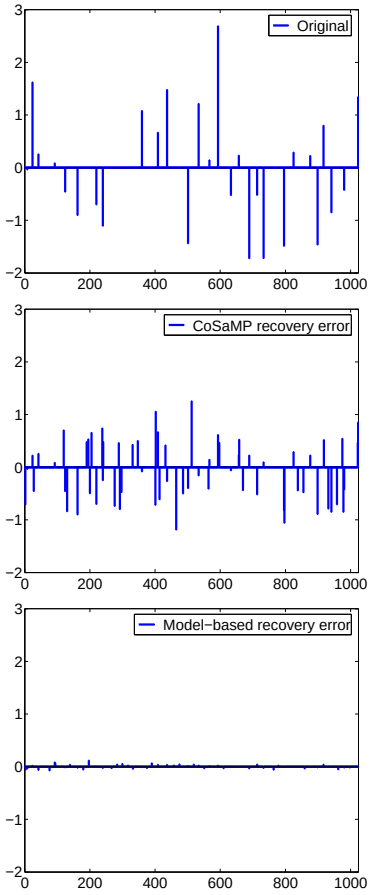

Fig. 2. *Performance of model-based neuronal spike train recovery. (top) Example spike train with parameters* $N = 1024, K = 50, \Delta = 10$*; (middle) CoSAMP recovery from* $M = 150$ *measurements. Distortion = 1.76 dB. (bottom) Model-based recovery from the same* $M = 150$ *measurements. Distortion = 25.53 dB.*

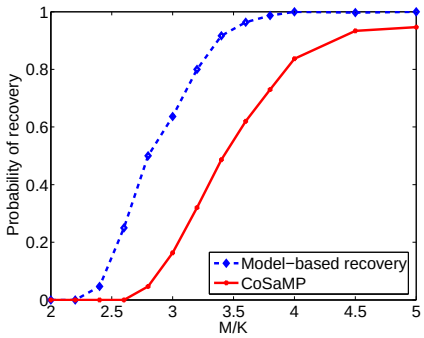

Fig. 3. *Probability of signal recovery for varying measurement ratios.* $N = 1024, K = 40, \Delta = 20$.

algorithm design to several diverse applications such as image processing, sensor networks and computer vision.

As an instantiation of our model-based framework, we study the class of neuronal spike trains with nonzero refractory period Δ. We systematically construct a suitable model, a sampling bound and a provably accurate algorithm which are specifically tailored to the model describing this class of signals. The improvement in the number of measurements M is captured by the parameter Δ; as Δ grows larger (i.e., consecutive spikes occur further apart), we may expect to see considerable gains in CS recovery performance.

Acknowledgments

The authors would like to thank Richard Baraniuk and Don Johnson for valuable suggestions. This work was supported by the grants NSF CCF-0431150, CCF-0728867, CNS-0435425, and CNS-0520280, DARPA/ONR N66001-08-1-2065, ONR N00014-07-1-0936, N00014-08-1-1067, N00014-08-1-1112, and N00014-08-1-1066, AFOSR FA9550-07-1-0301, ARO MURI W311NF-07-1-0185, and the Texas Instruments Leadership University Program.

References

[1] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, 1999.

[2] E. J. Candès. Compressive sampling. In *Proc. International Congress of Mathematicians*, volume 3, pages 1433–1452, Madrid, Spain, 2006.

[3] D. L. Donoho. Compressed sensing. *IEEE Trans. Info. Theory*, 52(4):1289–1306, Sept. 2006.

[4] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, 20:33, 1998.

[5] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright. Gradient projections for sparse reconstruction: Application to compressed sensing and other inverse problems. 2007. Preprint.

[6] J. Tropp and A. C. Gilbert. Signal recovery from partial information via orthogonal matching pursuit. *IEEE Trans. Info. Theory*, 53(12):4655–4666, Dec. 2007.

[7] D. Needell and J. Tropp. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, June 2008. To be published.

[8] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde. Model based compressive sensing. August 2008. Preprint.

[9] W. Gerstner and W.M. Kistler. *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge University Press, New York, NY, USA, 2002.

[10] T. Blumensath and M. E. Davies. Iterative hard thresholding for compressed sensing. July 2008. Preprint.

[11] W. Dai and O. Milenkovic. Subspace pursuit for compressive sensing: Closing the gap between performance and complexity. Mar. 2008. Preprint.

[12] M. F. Duarte, M. B. Wakin, and R. G. Baraniuk. Fast reconstruction of piecewise smooth signals from random projections. In *Proc. SPARS05*, Rennes, France, Nov. 2005.

[13] C. La and M. N. Do. Tree-based orthogonal matching pursuit algorithm for signal reconstruction. In *IEEE International Conference on Image Processing (ICIP)*, pages 1277–1280, Atlanta, GA, Oct. 2006.

[14] T. Blumensath and M. E. Davies. Sampling theorems for signals from the union of linear subspaces. *IEEE Trans. Info. Theory*, July 2007. Submitted.

[15] Y. Eldar and M. Mishali. Robust recovery of signals from a union of subspaces. 2008. Preprint.

[16] Y. M. Lu and M. N. Do. Sampling signals from a union of subspaces. *IEEE Signal Processing Mag.*, 25(2):41–47, Mar. 2008.

[17] G. L. Nemhauser and L. A. Wolsey. *Integer and Combinatorial Optimization*. Wiley-Interscience, 1999.