

Recovering Spikes from Noisy Neuronal Calcium Signals via Structured Sparse Approximation

Eva L. Dyer^r, Marco F. Duarte^p, Don H. Johnson^r, and Richard G. Baraniuk^r

^rDept. of Electrical and Computer Engineering, Rice University, Houston, TX 77005, USA

^pProgram in Applied and Comp. Math., Princeton University, Princeton, NJ 08544, USA

Abstract. Two-photon calcium imaging is an emerging experimental technique that enables the study of information processing within neural circuits *in vivo*. While the spatial resolution of this technique permits the calcium activity of individual cells within the field of view to be monitored, inferring the precise times at which a neuron emits a spike is challenging because spikes are hidden within noisy observations of the neuron’s calcium activity. To tackle this problem, we introduce the use of sparse approximation methods for recovering spikes from the time-varying calcium activity of neurons. We derive sufficient conditions for exact recovery of spikes with respect to (i) the decay rate of the spike-evoked calcium event and (ii) the maximum firing rate of the cell under test. We find—both in theory and in practice—that standard sparse recovery methods are not sufficient to recover spikes from noisy calcium signals when the firing rate of the cell is high, suggesting that in order to guarantee exact recovery of spike times, additional constraints must be incorporated into the recovery procedure. Hence, we introduce an iterative framework for *structured sparse approximation* that is capable of achieving superior performance over standard sparse recovery methods by taking into account knowledge that spikes are non-negative and also separated in time. We demonstrate the utility of our approach on simulated calcium signals in various amounts of additive Gaussian noise and under different degrees of model mismatch.

Key words: Two-photon calcium imaging, structured sparse approximation, spike recovery, exact support recovery, coherent dictionaries

1 Introduction

Despite the tight link between chemical and electrical signaling in neural systems, experimental studies of neural information processing have mainly been limited to observations of the electrical activity of single neurons and small neural networks. Recently however, a new experimental technique known as two-photon calcium imaging has enabled the study of the concentration of calcium within the dendritic trees of single neurons as well as the time-varying calcium activity of neural populations *in vivo* [1]. Studies of neuronal calcium activity have already confirmed the existence of functional maps in the visual cortex [2]; however, in order to use this technique to study temporal correlations between distinct neurons, calcium imaging must also be able to uncover the precise times at which each cell emits a spike or action potential (AP). Fortunately,

the calcium activity of the cell may be used to infer the times at which the neuron emits a spike from its cell body due to the fact that immediately after a spike is emitted from the cell, calcium rushes into the cell body and this results in a spike-evoked calcium transient that has roughly exponential decay with a decay time of approximately 0.5–1 second. Due to the long time course of spike-evoked calcium transients and high levels of photon shot noise incurred during the sensing process, the problem of inferring spikes hidden within the calcium activity of a cell is extremely challenging. Furthermore, the question of whether precise spike timing information can be extracted from noisy calcium signals is still open to debate, especially for fast-spiking interneurons.

In this paper, we introduce the use of sparse approximation methods for recovering spike times from noisy observations of the time-varying calcium activity of an individual neuron. By using sparse approximation techniques for spike inference, we are afforded a great deal of theory that enables a principled study of the limits of spike recovery from calcium signals. We use these results to derive sufficient conditions that ensure exact recovery of spike times with respect to (i) the decay rate of the spike-evoked calcium event and (ii) the maximum firing rate of the cell under test. This theoretical study, coupled with our experimental findings, suggests that in order to guarantee exact recovery of spikes from noisy calcium signals, additional constraints must be incorporated into the recovery procedure. Hence, we introduce the use of *structured sparse approximation* for spike recovery and show that, when spikes are separated in time and this knowledge is leveraged in the recovery procedure, spike times can be extracted faithfully in higher levels of noise than when no constraint is imposed. We detail a structured sparse recovery algorithm capable of exploiting this sort of knowledge and study the utility of our approach to structured sparse recovery on simulated calcium signals in various levels of noise.

Related Work. The problem of spike recovery from calcium imaging data was previously studied by Vogelstein et al., who develop sequential Monte Carlo methods [3] and fast methods for non-negative deconvolution [4]. Both of these methods assume that the spikes are distributed according to a Poisson distribution and that the spike-evoked calcium transients have exponential decay. In contrast to both of these approaches, the methods we develop here can be used to perform spike inference for arbitrary calcium waveforms and do not leverage any model of the dynamical behavior of the spikes nor do they assume a particular noise distribution. For this reason, our methods require far fewer parameters than methods that rely on estimates of the distributions of the noise and spiking dynamics. In addition, our work is the first to address the question of whether spike timing information can be reliably extracted from calcium signals with respect to parameters that govern the generation of spike-evoked calcium signals.

Paper Organization. We introduce sparse representations, nonlinear approximation methods for sparse recovery, and metrics required to derive dictionary dependent conditions for exact spike recovery in Section 2. Following this, we show how the problem of spike recovery from neuronal calcium signals can be posed as a sparse approximation problem and derive conditions for exact spike recovery in coherent dictionaries in Section 3. In Section 4, we introduce methods for structured sparse approximation, detail our algorithmic approach to structured sparse signal recovery, and then present our experimental results. We end with conclusions in Section 5.

2 Background

Sparse Representation. In many applications, it is advantageous to transform a signal into a new domain where it admits a compact or *sparse* representation. This idea lies at the heart of transform coding-based compression algorithms such as JPEG, where orthonormal bases are typically employed. When the collection of elements used to represent the signal are redundant, the representation is no longer unique; however, redundant systems often produce even sparser representations than orthonormal systems. To make this precise, we will refer to a finite collection of M unit-norm atoms as a dictionary, $\mathcal{D} = \{\varphi_m\}_{m=1}^M$, where $\|\varphi_m\|_2 = 1$, for all m . Upon stacking the atoms in \mathcal{D} into the columns of a dictionary matrix $\Phi \in \mathbb{R}^{N \times M}$, an exact reconstruction of the input signal $\mathbf{x} \in \mathbb{R}^N$ can be obtained by finding a linear combination of the atoms in the dictionary weighted by the coefficient vector $\mathbf{a} \in \mathbb{R}^M$, where $\mathbf{x} = \sum_{m=1}^M a(m)\varphi_m = \Phi\mathbf{a}$. The dictionary is considered complete if it spans \mathbb{R}^N with $M = N$ and overcomplete if it spans \mathbb{R}^N with $M > N$.

When using an overcomplete dictionary, the simplest explanation of the data or the sparsest representation may be found by finding an approximate solution to the following non-convex problem,

$$\min_{\mathbf{a}} \|\mathbf{a}\|_0 \quad \text{subject to} \quad \mathbf{x} = \Phi\mathbf{a}, \quad (1)$$

where $\|\mathbf{a}\|_0$, denoted as the “ ℓ_0 -norm” of \mathbf{a} , counts the number of non-zeros in \mathbf{a} . We call a signal k -sparse if $\|\mathbf{a}\|_0 \leq k$ and refer to the index set containing the indices of atoms with non-zero coefficients as the *support* of the signal or $\text{supp}(\mathbf{a}) = \Lambda$. We denote the corresponding sub-dictionary as Φ_Λ , the complement set of atoms as Φ_{Λ^c} , and can express the entire dictionary matrix as, $\Phi = [\Phi_\Lambda \ \Phi_{\Lambda^c}]$. The ℓ_0 -norm penalty in (1) is non-convex; however, we can relax (1) by replacing the ℓ_0 term with the ℓ_1 -norm $\|\mathbf{a}\|_1 = \sum_{m=1}^D |a_m|$, resulting in a method known as basis pursuit (BP) [5]. Alternatively, greedy methods such as orthogonal matching pursuit (OMP) [6] can be employed which select atoms iteratively, subtracting the contribution of each selected atom from the current signal residual.

Exact Recovery Conditions. In order to derive conditions that guarantee that sparse approximation methods like BP and OMP will recover a unique representation of a signal in a particular dictionary, two measures related to the similarity of atoms in the dictionary can be used. For unit-norm atoms, the *maximum coherence* is defined as $\mu \equiv \max_{i \neq j} |\langle \varphi_i, \varphi_j \rangle|$ and the *cumulative coherence* is

$$\mu(k) \equiv \max_{|\Lambda|=k} \max_{m \in \Lambda^c} \sum_{n \in \Lambda} |\langle \varphi_n, \varphi_m \rangle| \leq k\mu. \quad (2)$$

When atoms are sufficiently incoherent, a signal synthesized from a collection of these atoms is guaranteed to be unique and hence recoverable by standard sparse approximation methods like BP and OMP. In [7], Gribonval and Nielsen derive a sufficient condition for exact recovery of a signal drawn from the sub-dictionary Φ_Λ that depends on what we call the intra-support coherence, μ_Λ , and the inter-support coherence, μ_{Λ^c} of the dictionary.

Proposition 1 (Neumann Exact Recovery Coefficient (ERC) [7]) *Let a signal be supported over the index set Λ . If the Neumann ERC of a dictionary $\mathcal{D} = \{\varphi_m\}$,*

$$\mu_\Lambda + \mu_{\Lambda^c} = \max_{n \in \Lambda} \sum_{m \in \Lambda, m \neq n} |\langle \varphi_m, \varphi_n \rangle| + \max_{n \in \Lambda^c} \sum_{m \in \Lambda} |\langle \varphi_m, \varphi_n \rangle| < 1, \quad (3)$$

holds then OMP will recover the support set Λ exactly.

This result implies a slightly weaker condition stated below in (4) that is also sufficient to guarantee exact recovery of k -sparse signals. In [8], Tropp demonstrates that this condition guarantees that a quantity related to the Neumann ERC is positive, which is sufficient for exact recovery of all k -sparse signals via BP as well.

$$\mu(k) + \mu(k-1) < 1. \quad (4)$$

3 Exact Recovery of Spikes from Neuronal Calcium Signals

Before delving into conditions that ensure exact recovery of spikes from calcium signals, we briefly introduce a model for spike-evoked calcium signals that is a generalization of the linear model introduced in [3].

Signal Model. When a spike is emitted from a neuron, voltage-gated calcium channels across the cell's membrane cause an influx of calcium into the cell. The calcium transients that emerge are marked by a quick rise time and roughly exponential decay on the order of 0.5–1 second. Thus, the detected fluorescence signal $\mathbf{y} \in \mathbb{R}^N$ can be modeled as

$$\mathbf{y} = \sum_{m \in \Lambda} a_m \varphi_m + \eta = \Phi \mathbf{a} + \eta = \mathbf{x} + \eta, \quad (5)$$

where φ_m is a N -dimensional vector of zeros with a spike-evoked waveform placed at the m^{th} sample, $\mathbf{x} \in \mathbb{R}^N$ is a linear combination of calcium waveforms drawn from the dictionary Φ weighted by the coefficient vector $\mathbf{a} \in \mathbb{R}^M$, and η is a vector containing additive noise. With this linear model, sparse approximation techniques such as those described in section 2 can be used to recover the index set Λ and corresponding non-zero coefficients $\mathbf{a}_\Lambda \in \mathbb{R}^{|\Lambda|}$ for arbitrary calcium waveforms. We note that because the coefficients in \mathbf{a} correspond to the amplitude of each spike, the coefficients are all non-negative but are not necessarily equal due to slight variations in spike generation process. If we wish to enforce knowledge that coefficients used in the representation must be non-negative, a non-negative version of OMP and BP can easily be obtained. In the case of greedy algorithms like OMP and CoSaMP [9], non-negativity can be imposed by simply selecting atoms with large non-negative coefficients (instead of coefficients that are large in absolute magnitude) and replacing the least-squares step at each iteration with a non-negative least-squares step.

Sufficient Conditions for Exact Spike Recovery. We will now study conditions that allow for exact recovery of spike times from neuronal calcium signals. We say that exact recovery is obtained when the index set Λ used to synthesize a signal can be recovered without any errors. Although one of the advantages of using sparse approximation

methods is that they can be used to recover spikes from dictionaries consisting of shifted versions of arbitrary spike-evoked calcium signals, to simplify our subsequent analysis, we will assume that each spike-evoked calcium signal has exponential decay and a finite window of length L over which it is non-vanishing. Hence, each atom $\varphi_m \in \mathbb{R}^N$ can be written as $\varphi_m(n) = \beta^{n-m}$, for all $n = \{m, \dots, m+(L-1)\}$ and 0 otherwise, where L is chosen so that β^L negligible. Thus, the maximum coherence between any two atoms is simply the normalized inner product between atoms at neighboring points in time, $\mu \approx \beta$, where the approximation is due to the fact that the waveforms are truncated to L samples. Likewise, the cumulative coherence is defined over k calcium waveforms placed at neighboring samples: $\mu(k) \approx \sum_{n=1}^{\min(k,L)} \beta^n = \frac{\beta(\beta^k-1)}{\beta-1} = \frac{\mu(\mu^k-1)}{\mu-1} \leq \mu(L)$. By simply plugging this definition into the condition (4), we obtain the following sufficient condition for exact support recovery from shift-invariant dictionaries of spike-evoked calcium transients with exponential decay.

Proposition 2 (Exact Spike Recovery from Calcium Signals) *For any k -sparse signal synthesized from a dictionary with cumulative coherence $\mu(k) = \mu(\mu^k - 1)/\mu - 1$, where μ is the maximum coherence, if $3\mu - \mu^k(\mu - 1) > 1$, then this condition is sufficient to guarantee exact recovery via OMP and BP.*

This implies that to guarantee that two spikes can be recovered faithfully, we need $\mu^2 + 2\mu - 1 < 0$, and for $\mu \in (0, 1)$ this implies that $\mu < \sqrt{2} - 1$ or that $\mu < 0.414$. We show the relationship between this condition for exact recovery for different values of k in Figure 1(b).

4 Structured Sparse Approximation for Spike Recovery

In the previous section, we found that in order to guarantee that two spikes that occur at subsequent samples can be reliably separated, $\mu < 0.414$. In practice, we find that this condition is overly pessimistic and that sparse approximation are quite robust to high levels of coherence. However, when noise is introduced into observations, standard sparse approximation methods begin to fail when the noise exceeds 20% of the spike amplitude. Hence, to ensure exact recovery of spike times in the levels of noise present in real calcium imaging data, we must go “beyond sparsity” and incorporate additional constraints into our recovery procedure.

Drawing upon inspiration from recent results in structured sparse signal recovery [10], we introduce the use of a flexible iterative framework for structured sparse approximation that can be used to incorporate a number of different model-driven constraints into our recovery procedure. We find that if we leverage the fact that spikes are non-negative as we described in the previous section, this in turn leads to improved recovery; however, if we also assume that spikes are separated in time, we can obtain even better recovery performance in noisy settings. Although this type of separation constraint may not be applicable in all settings, e.g, when the cell under test is a fast-spiking cell, there are many settings where there exist physical constraints on the spike rate, either due to the refractory period or the maximum firing rate of the cell. Hence, when these assumptions are satisfied, it is advantageous to leverage knowledge about the maximum spiking rate into the recovery procedure.

Algorithm 1 Non-Negative Separated Support Model CoSaMP (SSM-CoSaMP)Inputs: Dictionary Φ , observations \mathbf{y} , K , Δ , model approximation algorithm $\mathbb{M}(\cdot, K, \Delta)$ Output: K -sparse approximation $\hat{\mathbf{x}}$ to true signal \mathbf{x} $\hat{\mathbf{x}}_0 = 0$, $\mathbf{r} = \mathbf{y}$; $\mathbf{b} = 0$; $i = 0$ {initialize}**while** halting criterion false **do**1. $i \leftarrow i + 1$ 2. $\mathbf{e} \leftarrow \Phi^\dagger \mathbf{r}$ {form signal residual estimate}3. $\Omega \leftarrow \text{supp}(\mathbb{M}(\mathbf{e}, 2K, \Delta))$ {prune signal residual estimate according to model}4. $\Lambda \leftarrow \Omega \cup \text{supp}(\hat{\mathbf{x}}_{i-1})$ {merge supports}5. $\mathbf{b}|_\Lambda \leftarrow \mathbb{N}(\mathbf{y}, \Phi_\Lambda)$ {compute non-negative least-squares signal estimate}6. $\hat{\mathbf{x}}_i \leftarrow \mathbb{M}(\mathbf{b}, K, \Delta)$ {prune signal estimate according to model}7. $\mathbf{r} \leftarrow \mathbf{y} - \Phi \hat{\mathbf{x}}_i$ {update measurement residual}**end while**return $\hat{\mathbf{x}} \leftarrow \hat{\mathbf{x}}_i$

To provide insight into why the introduction of a constraint on the firing rate of the cell leads to improved recovery, we examine the two terms in the Neumann ERC, $\mu_\Lambda + \mu_{\Lambda^c}$. For a signal drawn from a collection of atoms in Λ that are all separated by at least Δ samples, the intra-support coherence is bounded by $\mu_\Lambda \leq \sum_{n=1}^{\min(k,L)} \beta^{\Delta n} = \frac{\mu_\Delta((\mu_\Delta)^k - 1)}{\mu_\Delta - 1}$, where the maximum coherence $\mu_\Delta = \beta^\Delta$ (see Fig. 1(a) for a demonstration of how Δ and β contribute to the maximum coherence between atoms in the representation). Although signals generated under this model exhibit reduced intra-support coherence for all $\Delta > 1$, when we compute the inter-support coherence $\mu_{\Lambda^c} = \max_{n \in \Lambda^c} \sum_{m \in \Lambda} |\langle \varphi_m, \varphi_n \rangle|$, an atom that lies in between any of the spikes in the support will be selected to maximize this term. This means that even when spikes are separated in time, when our dictionary is not modified to reflect this constraint, the inter-support coherence term remains large. Hence, when our recovery algorithm is agnostic to the model that generated our observations, we do not observe improved recovery performance. In contrast, when we employ a structured sparse recovery method that leverages knowledge that there is some separation between atoms in the support, we effectively reduce the total coherence of our representation and this leads to better estimates of spike times in noisy settings. To enforce this knowledge within an algorithmic framework, we introduce an iterative method for structured sparse approximation that is detailed in Algorithm 1. We denote the minimizer of the non-negative least-squares problem $\mathbb{N}(\mathbf{s}, \Phi_\Lambda) = \arg \min_{\mathbf{a} \geq 0} \|\Phi_\Lambda \mathbf{a} - \mathbf{s}\|_2^2$ and the model-based pruning algorithm described in [10] as $\mathbb{M}(\mathbf{s}, K, \Delta)$ which finds the closest ℓ_2 approximation to a signal \mathbf{s} that has exactly K coefficients that are at least Δ samples apart from one another.

Results. To generate the data used for our simulations, we generate a sequence of spikes according to a Poisson distribution with rate λ and then pass the spikes through the model pruning algorithm $\mathbb{M}(\cdot, K, \Delta)$ to ensure that all spikes are at least Δ samples apart. We note that we never assume knowledge of the dynamical behavior of the spikes and only use a Poisson distribution to simulate realistic spiking behavior. Afterwards, the spikes are passed through a first-order IIR filter with difference equation, $x(n) = \beta x(n-1)$, and i.i.d. Gaussian noise with zero mean and variance σ^2 is added.

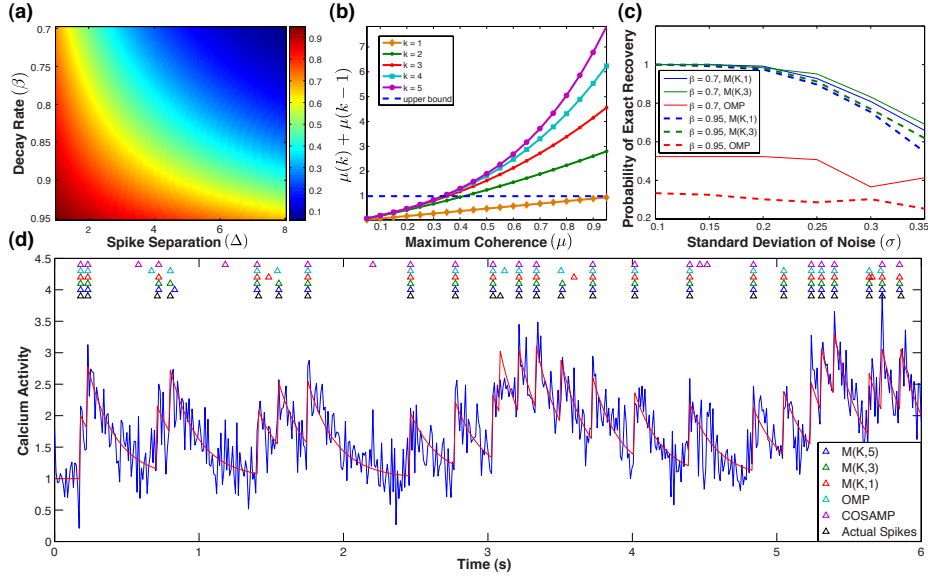


Fig. 1. (a) Maximum coherence μ_Δ for different values of Δ and β ; (b) comparison of μ_Δ with $\mu_\Delta(k) + \mu_\Delta(k-1)$ and different sparsity levels (exact recovery is guaranteed when curve lies below one); (c) Probability of exact recovery of a spike train with $\Delta^* = 3$ and $\beta = \{0.7, 0.95\}$ for different values of σ ; (d) Synthetic calcium signals generated with $K = 25$, $\Delta^* = 5$, $\beta = 0.95$, $\sigma = 0.3$. On top, we show the actual spike train (black) and recovered spikes using Algorithm (1) with $\Delta = \{5, 3, 1\}$ (blue, green, and red) and for OMP and COSAMP (cyan and magenta).

In Fig. 1(d) we show an example of a calcium signal in red, the noisy fluorescence signal in blue, and the corresponding spike trains recovered via Algorithm (1) and compare these results with those obtained via CoSaMP [9] and OMP. For this example, we set $\Delta^* = 5$ samples (for a sampling rate of 100 Hz, $\Delta = 5 = 50$ ms) and employ Algorithm (1) for $\Delta = \{1, 3, 5\}$, where $\Delta = 1$ corresponds to the case where no separation constraint is imposed but the non-negativity constraint is in effect. We show improved recovery over conventional methods such as CoSaMP and OMP when a non-negativity constraint is imposed but even further improvement when we assume $\Delta = \{3, 5\}$. In Fig. 1(c) we show the probability of exact recovery for various amounts of additive noise after averaging over 10 trials. We find that near-perfect recovery is obtained until the standard deviation of the noise is equal to 20% of the amplitude of the spikes. In contrast, OMP fails to recover the correct support even in small amounts of noise due to the high levels of coherence in the dictionary.

5 Discussion

In this paper, we have introduced sparse approximation for spike recovery and studied the limits of spike recovery from noisy observations of neuronal calcium activity. This study marks the first effort of its kind to explore the limits of using calcium imaging data for extracting precise spike timing information from single neurons. After studying the performance of these methods in noise, we introduced the use of structured sparse

approximation for improved spike recovery and show that by enforcing knowledge that spikes are separated in time in conjunction with the fact that the spikes are non-negative, we obtain significantly better estimates of the spiking activity than when standard sparse approximation methods are used. Our study suggests that although spikes may not be reliably extracted from calcium imaging data collected from fast-spiking cells, for cells that exhibit slower spiking rates, structured sparse approximation methods may be used to reliably recover spike timing information, even from very noisy data.

Furthermore, there are many settings where exact spike times are not essential for the study of a particular phenomenon of interest. In this case, structured sparse approximation methods like those we introduce here can also be used to obtain reliable estimates of spike counts over windows where bursts of spikes occur in addition to being able to reliably estimate individual spike times when sparse firing occurs. In this way, structured sparse approximation can be used to compute histograms of spike counts with more accuracy than when non-overlapping bins of a particular width are used. Due to the fact that our framework for structured sparse approximation can be used to enforce a number of model-driven constraints, this framework can also be extended to incorporate a number of different constraints into the recovery of spikes and in the analysis and/or denoising of calcium signals collected from diverse cell types, e.g., astrocytes and interneurons. Our future efforts include a large-scale study of spike recovery methods for *in vivo* data and extending our analysis to study conditions for exact recovery when structured sparse approximation methods are employed.

Acknowledgments. We thank Dr. Andreas Tolias, James Cotton, Dimitri Yatsenko and Chinmay Hegde for insightful discussions. This work was supported by AFOSR-FA9550-07-1-0301 and ONR-N00014-08-1-1112. ED was supported by a training fellowship from NLM-5T15LM007093. MFD was supported by NSF DMS-0439872.

References

1. Göbel, W., Helmchen, F.: In Vivo Calcium Imaging of Neural Network Function. *Physiology* 22, 358–365 (2007)
2. Ohki K., Chung S., Ch’ng Y.H., Kara P., Reid C.: Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature* 433, 597–603 (2005)
3. Vogelstein, J., Watson, B., et al.: Spike inference from calcium imaging using sequential Monte Carlo methods. *Biophysical Journal*, Vol. 97, Issue 2, 636–655 (2009)
4. Vogelstein, J., Packer, A., et al.: Fast non-negative deconvolution for spike train inference from population calcium imaging. *J Neurophysiology*, in press, (2010)
5. Chen, S., Donoho, D., Saunders, M.: Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.*, Vol. 20, No. 1, 33–61 (1999)
6. Davis, G., Mallat, S., Avellaneda, M.: Greedy adaptive approximation. *J. Constr. Approx.*, Vol. 13, 57–98 (1997)
7. Gribonval, R., Nielsen, M.: Beyond sparsity: Recovering structured representations by ℓ_1 minimization and greedy algorithms. *Adv. in Comp. Math*, No. 28, 23–41 (2008)
8. Tropp, J.A.: Greed is good: Algorithmic results for sparse approximation. *IEEE Trans. Info. Theory*, Vol. 50, No. 10, 2231–2242 (2004)
9. Needell, D., Tropp, J.: CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Appl. Comp. Harmonic Anal.*, Vol. 26, 301–321 (2008)
10. Hegde, C., Duarte, M.F., Cevher V.: Compressive Sensing Recovery of Spike Trains Using a Structured Sparsity Model. *SPARS* (2009)