Stochastic Filtering via Reweighted-ℓ_1

Adam S. Charles, and Christopher J. Rozell
School of ECE, Georgia Institute of Technology, Atlanta, GA, 30332-0250 USA (e-mail: acharles6,crozell@gatech.edu)

Many modern signal processing applications can be stated in terms of performing inference on sparse signals that have significant dependencies between those signals. Consider the task of recovering signals $\boldsymbol{x}_n \in \mathbb{R}^N$ from incomplete and noisy linear measurements $\boldsymbol{y}_n = \boldsymbol{\Phi}_n \boldsymbol{x}_n + \boldsymbol{\epsilon}_n$, where $\boldsymbol{\Phi}_n \in \mathbb{R}^{M \times N}$ is the potentially time-varying measurement matrix, $\boldsymbol{\epsilon}_n$ is the measurement noise, and $n \in K$ is an index (e.g., temporal or spatial). Many methods estimate each $\boldsymbol{x}_n$ independently by assuming that $\boldsymbol{x}_n$ is sparse in some dictionary ($\boldsymbol{x}_n = \boldsymbol{D}\boldsymbol{a}_n$ with $\boldsymbol{a}$ sparse). More recent methods have sought to leverage inter-signal dependencies through joint inference. For example, in some applications (e.g. recovery of video sequences or dynamic MRI) $n \in K \subset \mathbb{Z}$ represents a time index and we wish to recover ordered sparse signals with a known dynamics model

$$\boldsymbol{x}_n = f\left(\boldsymbol{x}_{n-1}\right) + \boldsymbol{\nu}_n, \tag{1}$$

where $f\left(\cdot\right): \mathbb{R}^N \to \mathbb{R}^N$ describes the dynamics and $\boldsymbol{\nu}$ is the model error (called an innovation). In applications such as hyperspectral imagery (HSI), $n \in K \subset \mathbb{Z} \times \mathbb{Z}$ represents pixel location and we wish to recover sparse spectral signatures with spatial regularity [1].

We desire a stochastic filtering approach that leverages sparse structure of single signals, exploits dependencies between signals, and is computationally efficient enough to apply on high-dimensional datasets. Classic approaches to stochastic filtering either have restrictive assumptions (i.e., Gaussian assumptions in Kalman filtering) or are so general that it is difficult to incorporate specific strong sparsity assumptions (i.e., particle filtering). More recent approaches have addressed aspects of jointly estimating sparse signals. For example a number of algorithms have been developed for recovering temporally evolving signals when the support stays constant [2], [3], or when a known function relates the signals through time [4], [5]. Other algorithms have been proposed for when there is no temporal ordering but the signals have common support or correlated signal values (e.g. multiple measurement vector [6]).

We propose an approach to stochastic filtering for sparse signals based on reweighted ℓ_1 (RWL1) optimization that can incorporate a number of different signal dependency structures. Standard RWL1 iteratively solves a weighted basis pursuit de-noising (BPDN) program

$$\widehat{\boldsymbol{a}}_n^t = \arg\min_{\boldsymbol{a}} \|\boldsymbol{y}_n - \boldsymbol{\Phi}_n \boldsymbol{D}\boldsymbol{a}\|_2^2 + \lambda_0 \sum_i \widehat{\boldsymbol{\lambda}}_n^{t-1}[i] |\boldsymbol{a}[i]|$$

while updating the weights as

$$\widehat{\boldsymbol{\lambda}}_n^t[i] = \frac{\tau}{|\widehat{\boldsymbol{a}}_n^t[i]| + \eta},$$

where λ_0, τ and η are algorithmic constants and the variables $\boldsymbol{\lambda}_n$ are the weights. This optimization corresponds to using the EM algorithm to perform inference in a hierarchical sparsity model (called a Laplacian Scale Mixture model [7]) where the variances of the sparse coefficients are themselves random variables. In our general stochastic filtering approach, we modify the RWL1 program to update the weights based on information from correlated signals

$$\widehat{\boldsymbol{\lambda}}_n^t[i] = \frac{\tau}{g(\widehat{\boldsymbol{a}}_{k\in K}^t)[i] + \eta}$$

where $g(\cdot)$ captures inter-signal dependencies. In the case of dynamic filtering, an example function is $g(\widehat{\boldsymbol{a}}_{k\in K})[i] = |\widehat{\boldsymbol{a}}_n[i]| + |\boldsymbol{D}^{-1} f\left(\boldsymbol{D}\boldsymbol{z}\right)[i]|$, which essentially predicts the sparse coefficients via the known dynamics and uses it to bias the weights in the next iteration of the RWL1 algorithm. In the case of spatial filtering, this function captures unordered dependencies through operations such as a linear summation of coefficients in a neighborhood of pixels (i.e. using spatial regularity to bias the weights for a coefficient).

The proposed RWL1 approach to stochastic filtering has many benefits. The approach is very general (i.e., treating temporal and spatial dependencies in the same framework), and can leverage the recent advances in specialized ℓ_1 solvers for high-dimensional data. Furthermore, by leveraging a hierarchical probabilistic model, this approach provides explicit ways to incorporate new signal dependencies into the conditional distributions of the coefficient variances. Finally, by incorporating signal dependencies in the second order statistics rather than via direct support set estimates, the proposed approach appears to be much more robust to model errors than existing approaches.

Previous work has shown state-of-the-art performance for the proposed dynamic filtering algorithm (RWL1-DF) on compressed sensing video recovery. In this abstract, we focus on more recent results illustrating the proposed spatial filtering algorithm (RWL1-SF) on spectral superresolution of HSI from multispectral measurements. In particular, we use HSI data to simulate multispectral measurements and demonstrate that RWL1-SF can be used to recover the ground-truth HSI with a mean and median relative mean-squared error (rMSE) of 2.45% and 1.89% respectively. This represents a significant improvement over using BPDN or RWL1, and is particularly interesting because the biggest performance gains come from areas of the scene where the sparsity model is *not* performing well at capturing the data (i.e., highlighting the robustness benefits of the approach). As part of this work we will also show more detailed comparisons to other work for our previous dynamic filtering algorithm RWL1-DF.

References

[1] M. Iordache, J. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 50, no. 11, 2012.
[2] J. Ziniel, L. C. Potter, and P. Schniter, "Tracking and smoothing of time-varying sparse signals via approximate belief propagation," *Proceesings of the Asilomar Conference on Signals, Systems and Computers*, 2010.
[3] N. Vaswani, "Ls-cs-residual (ls-cs): Compressive sensing on the lease squares residual," *IEEE Trans on Sig Proc*, vol. 58, no. 8.
[4] A. Carmi, P. Gurfil, and D. Kanevsky, "Methods for sparse signal recovery using Kalman filtering pseudo-measuremennt norms and quasi-norms," *IEEE Trans on Sig Proc*, vol. 58, no. 4, pp. 2405–2409, Apr 2010.
[5] A. Charles, M. S. Asif, J. Romberg, and C. Rozell, "Sparsity penalties in dynamical system estimation," *Proc of the CISS*, March 2011.
[6] S. Cotter, B. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. on Sig. Proc.*, vol. 53, no. 7, pp. 2477–2488, 2005.
[7] P. Garrigues and B. Olshausen, "Group sparse coding with a laplacian scale mixture prior," *Adv in Neural Info Proc Sys*, vol. 24, 2010.