The Constrained Earth Mover Distance Model, with Applications to Compressive Sensing

Ludwig Schmidt, Chinmay Hegde, Piotr Indyk
Computer Science and Artificial Intelligence Lab
Massachusetts Institute of Technology

***Abstract*—Sparse signal representations have emerged as powerful tools in signal processing theory and applications, and serve as the basis of the now-popular field of compressive sensing (CS). However, several practical signal ensembles exhibit additional, richer structure beyond mere sparsity. Our particular focus in this paper is on signals and images where, owing to physical constraints, the positions of the nonzero coefficients do not change significantly as a function of spatial (or temporal) location. Such signal and image classes are often encountered in seismic exploration, astronomical sensing, and biological imaging. Our contributions are threefold: (i) We propose a simple, deterministic model based on the *Earth Mover Distance* that effectively captures the structure of the sparse nonzeros of signals belonging to such classes. (ii) We formulate an approach for approximating any arbitrary signal by a signal belonging to our model. The key idea in our approach is a min-cost max-flow graph optimization problem that can be solved efficiently in polynomial time. (iii) We develop a CS algorithm for efficiently reconstructing signals belonging to our model, and numerically demonstrate its benefits over state-of-the-art CS approaches.**

I. INTRODUCTION

A signal (or image) is said to be k-sparse if only k of its coefficients in a given basis expansion are nonzero; in other words, the intrinsic information content in the signal is miniscule relative to its apparent size. This simple notion enables a wide variety of conceptual and algorithmic techniques to compress, reconstruct, denoise, and process practical high-dimensional signals and images. Notably, sparsity serves as the cornerstone of the field of compressive sensing (CS), an interesting alternative to the classical Shannon/Nyquist theory for signal sampling and reconstruction [1, 2]. A canonical result in CS states that for a k-sparse signal of length n, merely $O(k \log n/k)$ non-adaptive, *linear* measurements (samples) suffice to ensure robust, efficient reconstruction. When $k \ll n$, this can lead to significant practical benefits.

In several practical applications, the nonzero coefficients of signal ensembles exhibit additional, richer relationships that cannot be captured by mere sparsity. Consider, for example, a 2D "image" constructed by column-wise stacking of seismic time traces (or shot records) measured by geophones positioned on a uniform linear array. Assuming the presence of only a few subsurface reflectors, the physics of wave propagation dictates that such a 2D image would essentially consist of a number of curved lines, possibly contaminated with noise (see Figure 1). A convenient model for such an image is to simply assume that each column is sparse; indeed, such a sparsity assumption has been proven to be beneficial for efficient shot record sampling and reconstruction [3]. However, while this assumption may suffice for some situations, such a model cannot capture the the fact that the indices of the nonzeros change smoothly across adjacent columns. Such settings are commonplace; for example, similar "line" singularities are encountered in applications such as biological imaging and radio-astronomy.

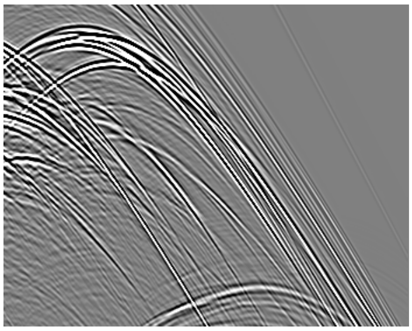

Fig. 1. *Example of a seismic shot record (Sigsbee2A data set). The horizontal axis corresponds with space (receiver) and the vertical axis with time. Note that the large coefficients of neighboring columns are at similar locations.*

In this paper, we propose a deterministic model for sparse signal ensembles where the locations of the nonzeros, or the support, of a signal transforms continuously as a function of spatial (or temporal) location. A key ingredient in our model is the classical Earth Mover Distance (EMD) [4], and we will call it the *Constrained EMD* model. Informally, our proposed model assumes that: (i) each signal in our ensemble is k-sparse, and (ii) the cumulative EMD between pairs of adjacent signal supports is constrained to be no greater than a nonnegative parameter B. The parameter B controls how dramatically the support can vary across different signals; a value of $B = 0$ indicates that the support remains invariant across all signals in our ensemble, while a large value of B admits potentially drastic changes across adjacent supports.

Next, given an arbitrary input signal (ensemble) x, we develop an efficient algorithm to find a near-optimal ℓ_2-approximation of x in the Constrained EMD model. We show that the support of the optimal approximation can be discovered by solving a small number of *min-cost max-flow* [5] problems over a specially defined graph. Each intermediate problem can be solved using existing, highly efficient network optimization methods, and therefore the overall signal approximation can be obtained in polynomial time.

Additionally, we demonstrate the advantages of the Constrained EMD model, and the associated approximation algorithm, in the context of compressive sensing. Geometrically, the model is equivalent to a particular *union of subspaces* of the ambient signal space. Therefore, we can leverage the framework of *model-based compressive sensing* [6] to build a new CS reconstruction algorithm that is specially tailored to signal ensembles well-described by the Constrained EMD model. We illustrate the numerical benefits of the new algorithm in comparison with existing state-of-the-art CS recovery approaches.

The rest of this paper is organized as follows. Section II provides a brief introduction to structured sparsity and compressive sensing. Section III introduces the constrained EMD model and describes our main algorithm. Section IV illustrates the advantages of our method with example reconstructions of images and quantitative results of algorithm performance. Section V concludes with a discussion of further directions.

II. Background

A. Preliminaries

A signal $x \in \mathbb{R}^n$ is said to be k-sparse in the orthobasis Ψ if at most $k < n$ coefficients of the basis expansion $\alpha = \Psi^T x$ are nonzero. In this paper, we assume that the basis Ψ is the identity matrix, while noting that all our 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 be represented by a binary vector $s(x) \in \{0,1\}^n$ with at most k ones. Denote the set of *all* k-sparse signals by Σ_k. Geometrically, this set is equivalent to the union of the $\binom{n}{k}$ canonical k-dimensional subspaces of $\mathbb{R}^n$.

B. Structured sparsity

Often, 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* $\mathcal{A}$ [7], comprising 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, define:

$$\mathcal{A} = \bigcup_{m=1}^{m_k} \mathcal{X}_m, \ \mathcal{X}_m := \{x : x|_{\Omega_m} \in \mathbb{R}^k, x|_{\Omega_m^C} = 0\}, \qquad (1)$$

where each subspace $\mathcal{X}_m$ contains all signals x with $\mathrm{supp}(x) \in \Omega_m$. 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 $x^* \in \mathcal{A}$, and $\|x - x^*\|_2$ is minimized. We define a *model-projection* algorithm as a procedure $\mathbb{M}(x, k)$ which returns the best k-term approximation of a given signal under the model $\mathcal{A}$, i.e., $x^* = \mathbb{M}(x, k)$.

C. Compressive Sensing

Suppose instead of collecting all the coefficients of a vector $x \in \mathbb{R}^n$, we merely record $m = O(k \log n/k)$ inner products (measurements) of x with $m < n$ pre-selected vectors, i.e., we observe an m-dimensional $y = \Phi x$, where $\Phi \in \mathbb{R}^{m \times n}$. The central tenet of compressive sensing (CS) is that x can be *exactly* recovered from y, even though Φ is rank-deficient (and therefore has a nontrivial nullspace). Numerous algorithms for signal recovery have been developed; particularly, iterative support selection algorithms (such as CoSaMP [8] and IHT [9]) have emerged that are both numerically stable and computationally efficient. Also, an added advantage is that such iterative algorithms can be easily *tailored to any arbitrary structured sparsity model*; this forms the central premise of *model-based* compressive sensing framework, initially proposed in [6]. In Section III below, we describe this further.

D. Related Work

There has been prior research on reconstructing time sequences of spatially sparse signals (e.g., [10]). Such approaches assume that the support of the signal (or even the signal itself) does not change much between two consecutive time steps. However, the variation between two columns a and b was defined according to the ℓ_0 distance between the supports $\|s(a) - s(b)\|_0$. In contrast, in this paper we measure this difference according to the classical Earth Mover Distance (EMD) (also variously known as the *Mallows* or the *Wasserstein* distance) between the supports. As a result, our model easily handles signals such as those in Figure 3, where the supports of any two consecutive columns can potentially be *even disjoint*, yet differ very little according to the EMD.

Another related work is that of [11], who proposed the use of the EMD in a compressive sensing context in order to measure the *approximation error* of the recovered signal. In contrast, in this paper we are using the EMD to constrain the *support set* of the signals.

III. The Constrained EMD Model

Below, we interpret the signal $x \in \mathbb{R}^n$ as a matrix $X \in \mathbb{R}^{h \times w}$ with $n = h\,w$. Furthermore, we denote the individual columns of X with $x_i \in \mathbb{R}^h$ for $i \in [w]$.

A. Definitions

Definition 1: The EMD of two index sets A and B with $|A| = |B|$ is defined as:

$$\mathrm{EMD}(A, B) = \min_{\pi : A \to B} \sum_{a \in A} |a - \pi(a)|, \qquad (2)$$

where π ranges over all one-to-one mappings from A to B.

Definition 2: The support-EMD of two k-sparse vectors $a, b \in \mathbb{R}^h$ is defined as:

$$\mathrm{sEMD}(a, b) = \mathrm{EMD}(\mathrm{supp}(a), \mathrm{supp}(b)). \qquad (3)$$

Definition 3: The *Constrained EMD model* is the set:

$$\mathcal{A}_{k,B} = \{X \in \mathbb{R}^{h \times w} : |\mathrm{supp}(x_i)| = k \text{ for } i \in [w], \sum_{i=1}^{w-1} \mathrm{sEMD}(x_i, x_{i+1}) \le B\}. \qquad (4)$$

$$X = \begin{bmatrix} 1 & 3 \\ 0 & -1 \\ 2 & 1 \end{bmatrix} \qquad G_{X,k,\lambda} =$$

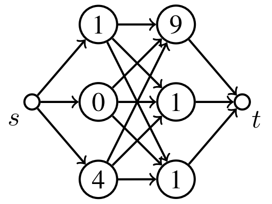

Fig. 2. *A signal X with the corresponding flow network $G_{X,k,\lambda}$. The node costs are the squared amplitudes of the corresponding signal components (negation omitted here). The capacities and edge costs are omitted for clarity. All capacities in the flow network are 1. The edge costs are the vertical distances between the start and end nodes.*

The set $\mathcal{A}_{k,B}$ in (4) is a subset of the set of all $k-$sparse signals Σ_k, and therefore the Constrained EMD model constitutes a specific instance of a structured sparsity model (1). For given dimensions of X, the Constrained EMD model has two parameters: (i) k, the sparsity of each column x_i and (ii) B, the cumulative support-EMD of adjacent columns x_i and x_{i+1}. Importantly, we note that we only constrain the EMD between adjacent signal *supports* and not the actual signal coefficients.

B. Graph-Based Model-Projection

In order to use our Constrained EMD signal model within a model-based compressive sensing framework, we need an algorithm that approximates arbitrary signals with signals in our model. Formally, we need a model-projection algorithm $\mathbb{M}(x, k, B)$ that returns a $\widehat{x} \in \mathcal{A}_{k,B}$ minimizing $\|x - x'\|_2$ for all $x' \in \mathcal{A}_{k,B}$.

To achieve this, we use the following *graph-based* approach. Observe that the support-EMD (3) of a pair of signals is the minimal cost of a maximum bipartite matching of the two support sets, where the edge costs are given by the absolute difference between the indices. We extend this intuition to ensembles of signals, via the notion of a *flow network*.

Definition 4: For a given signal X, sparsity k and parameter λ, the *flow network* $G_{X,k,\lambda}$ consists of the following elements:

- The *nodes* comprise a source s, a sink t and a node $v_{i,j}$ for $i \in [h]$, $j \in [w]$, i.e. one node per signal coefficient.
- G has an *edge* from every $v_{i,j}$ to every $v_{k,j+1}$ for $i, k \in [h]$, $j \in [w-1]$. Moreover, there is an edge from s to every $v_{i,1}$ and from every $v_{i,w}$ to t for $i \in [h]$.
- The *capacity* of every edge and node is 1.
- The *cost of a node* $v_{i,j}$ is $-x_{i,j}^2$. The *cost of an edge* from $v_{i,j}$ to $v_{k,j+1}$ is $\lambda|i-k|$. The cost of the source, the sink and all edges incident to the source or sink is 0.
- The *supply* at the source, and the *demand* at the sink, both equal k.

Figure 2 illustrates the construction of an example $G_{X,k,\lambda}$. Observe that for any $G_{X,k,\lambda}$, a standard min-cost max-flow optimization [5] through this network reveals a subset of nodes S that corresponds to exactly k indices per column. Moreover, this optimal flow minimizes the cost $-\|X|_S\|^2 + \lambda \sum_{i=1}^{w-1} \mathrm{EMD}(s_i, s_{i+1})$ over all choices of S. This cost includes both the fidelity of the signal projection as well as the cumulative support-EMD across columns. The trade-off between these two quantities is determined by the parameter λ; for small values of λ, the resulting flow has a large support-EMD and vice versa. Setting $\lambda = 0$ removes the EMD-constraint while $\lambda = +\infty$ is equivalent to selecting the k rows with the largest amplitude sums. By systematically varying the parameter λ, we can find a support S that belongs to the Constrained EMD model $\mathcal{A}_{k,B}$ for a target B and simultaneously maximizes the quality of the projection under this constraint.

Algorithm 1 Model projection $\mathbb{M}(x, k, B)$

$\lambda_l \leftarrow 0, \lambda_h \leftarrow 1$
do
  $\lambda_h \leftarrow 2\lambda_h$
  Run min-cost max-flow on G_{X,k,λ_h}
while resulting support has total support-EMD $> B$.
do
  $\lambda_m \leftarrow (\lambda_h + \lambda_l)/2$
  Run min-cost max-flow on G_{X,k,λ_m}
  if resulting support has total support-EMD $> B$
    $\lambda_l \leftarrow \lambda_m$
  else
    $\lambda_h \leftarrow \lambda_m$
while $\lambda_h - \lambda_l > \epsilon_\lambda$
return $\widehat{x}$ corresponding to min-cost max flow on G_{X,k,λ_h}

Algorithm 1 describes the entire model projection algorithm. In order to solve the min-cost max-flow instances, it is possible to exploit the special structure of the graph. Since all edges and nodes have unit capacity, it is sufficient to find k cheapest augmenting paths in the flow network. Using Dijkstra's algorithm and assuming a square X, i.e. $h = w = \sqrt{n}$, each min-cost max-flow can be found in $O(kn^{3/2})$ time.

C. Compressive Sensing

The model projection method (Alg. 1) is useful in a number of contexts. Here, we use Alg. 1 in order to develop a new compressive sensing (CS) reconstruction algorithm specially tailored to signals and images with line singularities. Since the constrained EMD model essentially is a special structured sparsity model $\mathcal{A}_k$, as in (1), Alg. 1 provides an projection algorithm for this model. Given such a projection algorithm, the framework of model-based compressive sensing [6] suggests that iterative support selection algorithms, such as CoSaMP and IHT, can easily be modified in order to be tailored for signals belonging to the constrained EMD model. Further, the modified algorithms are provably stable, as well as provably achieve successful recovery using fewer measurements than the conventional (unmodified) algorithms.

We summarize our proposed CS recovery method as Alg. 2; we call it EMD-CoSaMP. The modification is simple: simply replace the signal thresholding steps (3 and 6) by an appropriate model projection step. A similar modification of IHT can also be developed (the description of which we omit); we will call it EMD-IHT. Below, we empirically illustrate the benefits of our proposed model-based CS recovery algorithms.

Algorithm 2 EMD-CoSaMP(Φ, y)

$\widehat{x}_0 \leftarrow 0, r \leftarrow y, i \leftarrow 0$
while not converged **do**
1. $i \leftarrow i + 1$
2. $e \leftarrow \Phi^T r$
3. $\Omega \leftarrow \mathrm{supp}(\mathbb{M}(e, 2k, 2B))$
4. $T \leftarrow \Omega \cup \mathrm{supp}(\widehat{x}_{i-1})$
5. $z|_T \leftarrow \Phi_T^{\dagger} y,\ z|_{T^C} = 0$
6. $\widehat{x}_i \leftarrow \mathbb{M}(z, k, B)$
7. $r \leftarrow y - \Phi \widehat{x}_i$

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

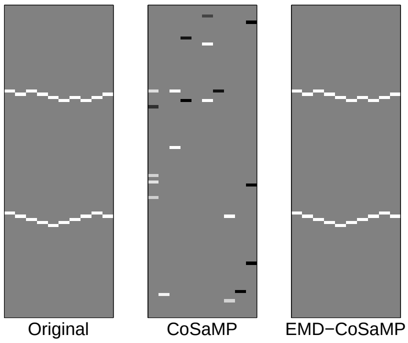

Fig. 3. *Benefits of CS reconstruction using EMD-CoSaMP. (left) Original image with parameters $h = 100$, $w = 10$, $k = 2$, $B = 20$, $m = 80$. (center) CS reconstruction using CoSaMP [8]. (right) CS reconstruction using EMD-CoSaMP. CoSaMP fails, while our proposed algorithm is able to perfectly recover the image.*

IV. Numerical Experiments

In all our experiments, we use the LEMON library [12] in order to solve the min-cost max-flow subroutine in Alg. 1. Figure 3 displays a test grayscale image of size 100×10 with edge discontinuities such that the total sparsity is $2 \times 10 = 20$ and the cumulative EMD across pairs of adjacent columns is equal to $B = 20$. We measure linear samples of this image using merely $m = 80$ random Gaussian measurements, and reconstruct using CoSaMP as well our proposed approach (EMD-CoSaMP). Each iteration of EMD-CoSaMP takes less than three seconds to execute. As visually evident from Fig. 3, CoSaMP fails to reconstruct the image, while our proposed algorithm provides an accurate reconstruction.

Figure 4 displays the results of a Monte Carlo experiment to quantify the effect of the number of random measurements M required by different CS reconstruction algorithms to enable accurate reconstruction. Each data point in Fig. 4 was generated using 100 sample trials over randomly generated measurement matrices. Successful recovery is declared when the converged solution is within an ℓ_2 distance of 5% relative to the Euclidean norm of the original image. We observe that our proposed EMD-CoSaMP and EMD-IHT algorithms achieve successful recovery with far fewer measurements than their conventional (unmodified) counterparts.

V. Conclusions

We have proposed a deterministic structured sparsity model, and associated model projection algorithm, based on the Earth Mover Distance (EMD) for signals and images with line singularities. We leverage this algorithm to develop a new compressive sensing (CS) recovery algorithm with significant numerical benefits. We defer a full theoretical characterization of our proposed CS recovery algorithm, as well as a thorough study of practical applications such as seismic shot record acquisition, to a future expanded version of this work.

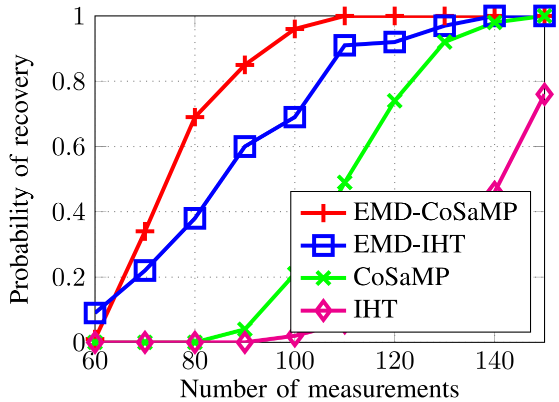

Fig. 4. *Comparison of several reconstruction algorithms. The signal is the same as in Figure 3. The probability of recovery is with respect to the measurment matrix and generated using 100 trial runs. The recovery algorithms using our constrained EMD model have a higher probability of recovery than standard algorithms.*

Acknowledgements

The authors would like to thank Lei Hamilton, Chris Yu, and Detlef Hohl for helpful discussions. This work was supported in part by a grant from the MITEI-Shell program, by the MADALGO center, and by the Packard Foundation.

References

[1] D. Donoho, “Compressed sensing,” *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, September 2006.

[2] E. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[3] F. Herrmann, “Randomized sampling and sparsity: Getting more information from fewer samples,” *Geophysics*, vol. 75, no. 6, pp. WB173–WB187, 2010.

[4] Y. Rubner, C. Tomasi, and L. J. Guibas, “The Earth Mover’s Distance as a metric for image retrieval,” *Intl. J. Comp. Vision*, vol. 40, no. 2, pp. 99–121, 2000.

[5] R. Ahuja, T. Magnanti, and J. Orlin, *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, 1993.

[6] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, “Model-based compressive sensing,” *IEEE Trans. Inform. Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[7] Y. Eldar and M. Mishali, “Robust recovery of signals from a structured union of subspaces,” *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 5302–5316, 2009.

[8] D. Needell and J. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, 2009.

[9] T. Blumensath and M. Davies, “Iterative hard thresholding for compressive sensing,” *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, 2009.

[10] N. Vaswani and W. Lu, “Modified-CS: Modifying compressive sensing for problems with partially known support,” in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Seoul, Korea, Jun. 2009.

[11] P. Indyk and E. Price, “K-median clustering, model-based compressive sensing, and sparse recovery for Earth Mover Distance,” in *Proc. ACM Symp. Theory of Comput.*, San Jose, CA, June 2011.

[12] B. Dezs, A. Jüttner, and P. Kovács, “LEMON: an open source C++ graph template library,” *Electron. Notes Theor. Comput. Sci.*, vol. 264, no. 5, pp. 23–45, 2011.