Recovery of Compressible Signals in Unions of Subspaces

Marco F. Duarte, Chinmay Hegde, Volkan Cevher, and Richard G. Baraniuk

Department of Electrical and Computer Engineering

Rice University

Abstract—**Compressive sensing (CS) is an alternative to Shannon/Nyquist sampling for acquisition of sparse or compressible signals; instead of taking periodic samples, we measure inner products with $M < N$ random vectors and then recover the signal via a sparsity-seeking optimization or greedy algorithm. Initial research has shown that by leveraging stronger signal models than standard sparsity, the number of measurements required for recovery of an structured sparse signal can be much lower than that of standard recovery. In this paper, we introduce a new framework for structured compressible signals based on the unions of subspaces signal model, along with a new sufficient condition for their recovery that we dub the restricted amplification property (RAmP). The RAmP is the natural counterpart to the restricted isometry property (RIP) of conventional CS. Numerical simulations demonstrate the validity and applicability of our new framework using wavelet-tree compressible signals as an example.**

Index Terms—**Compressive sensing, compressible signals, unions of subspaces**

I. INTRODUCTION

Compressive sensing (CS) is a new approach to simultaneous sensing and compression that enables a potentially large reduction in the sampling and computation costs at a sensor for signals having a sparse representation in some basis. CS builds on the work of Candès, Romberg, and Tao [1] and Donoho [2], who showed that a signal having a sparse representation in one basis can be recovered from a small set of projections onto a second, measurement basis that is *incoherent* with the first.[1] *Random projections* play a central role as a *universal* measurement basis in the sense that they are incoherent with any fixed basis with high probability. The CS measurement process is nonadaptive, and the recovery process is nonlinear, for which a variety of algorithms have been proposed [1–6].

While this represents significant progress from Nyquist-rate sampling, it is possible to do even better by more fully leveraging concepts from state-of-the-art signal compression and processing algorithms. In many such algorithms, the key ingredient is a more realistic *signal model* that goes beyond simple sparsity by codifying the inter-dependency *structure* among the signal coefficients. Coding the coefficients according to a model for this structure enables these algorithms to compress signals close to the maximum amount possible — significantly better than a naive coder that just processes each large coefficient independently.

We have previously developed a new model-based CS algorithmic framework [7] that parallels the conventional theory by establishing an underlying *signal model*, favoring certain configurations for the magnitudes and indices of the significant coefficients of the signal, and designing recovery algorithms that exploit the knowledge of such a model. Building on this work, we developed CS recovery algorithms in [8] for the recently introduced *union-of-subspaces* models for strictly sparse signals [9–11]. We have shown that these algorithms achieve the same performance as standard CS recovery [8]. In both cases, by reducing the degrees of freedom of a sparse or compressible signal, signal models provide two immediate benefits to CS. First, they enable a reduction in the number of measurements M required to stably recover a signal. Second, during signal recovery, they enable us to better differentiate true signal information from recovery artifacts, which leads to a more robust recovery.

In this paper, we build on previous work by expanding the union-of-subspaces signal models beyond strictly sparse signals. To precisely quantify the benefits of model-based CS, we introduce and study several new theoretical concepts that could be of more general interest. We define a class of *model-compressible signals* that can be approximated by increasingly complex sig-

[1]Roughly speaking, *incoherence* means that no element of one basis has a sparse representation in terms of the other basis.

nals under a established union-of-subspaces model with exponentially decaying error. We also formulate a new *restricted amplification property* (RAmP) that guarantees stable recovery of model-compressible signals. For some compressible signal models, the number of random measurements M required by the RAmP is independent of N.

This paper is organized as follows. A review of the CS theory is given in Section II. Section III develops the concept of model-sparse signals and introduces the concept of model-compressible signals. We also quantify how signal models improve the measurement and recovery process by exploiting the model-based RIP for model-sparse signals and by introducing the RAmP for model-compressible signals in Section IV. Section V then specializes our theory to the case of wavelet tree-sparse signal models and Section VI reports on a series of numerical experiments that validate our theoretical claims. We conclude with a discussion in Section VII.

II. BACKGROUND ON COMPRESSIVE SENSING

Given a basis Ψ, we can represent every signal $x \in \mathbb{R}^N$ in terms of the coefficient vector α as $x = \Psi\alpha$. In this section we will assume without loss of generality that the signal x is sparse or compressible in the canonical domain so that the sparsity basis Ψ is the identity and $\alpha = x$. A signal x is *K-sparse* if only $K \ll N$ entries of x are nonzero. We call the set of indices corresponding to the nonzero entries the *support* of x and denote it by supp(x). The set Σ_K of all K-sparse signals is the union of the $\binom{N}{K}$, K-dimensional subspaces aligned with the coordinate axes in $\mathbb{R}^N$.

While many natural and manmade signals are not strictly sparse, they can be approximated as such; we say that such signals are *compressible*. An example is a signal x whose coefficients, when sorted in order of decreasing magnitude, decay according to the power law

$$\left| x_{\mathcal{I}(i)} \right| \le S\, i^{-1/r}, \quad i = 1, \ldots, N, \tag{1}$$

where $\mathcal{I}$ indexes the sorted coefficients. Thanks to the rapid decay of their coefficients, such signals are well-approximated by K-sparse signals. Let $x_K \in \Sigma_K$ represent the best K-term approximation of x, which is obtained by keeping just the first K terms in $x_{\mathcal{I}(i)}$ from (1). Denote the error of this approximation as

$$\sigma_K(x) := \arg\min_{\bar{x} \in \Sigma_K} \|x - \bar{x}\|_2 = \|x - x_K\|_2.$$

We then have that

$$\sigma_K(x) \le (rs)^{-1/2} S K^{-s}, \tag{2}$$

with $s = \frac{1}{r} - \frac{1}{2}$. That is, the signal's best approximation error has a power-law decay with exponent s as K increases. We dub such a signal *s-compressible*.

Compressive sensing (CS) integrates the signal acquisition and compression steps into a single process [1–3]. In CS we do not acquire x directly but rather acquire $M < N$ linear measurements $y = \Phi x$ using an $M \times N$ measurement matrix Φ. We then recover x by exploiting its sparsity or compressibility. Our goal is to push M as close as possible to K in order to perform as much signal "compression" during acquisition as possible. In order to recover a good estimate of x from the M compressive measurements, the measurement matrix Φ should satisfy the *restricted isometry property* (RIP).

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

$$(1 - \delta_K)\|x\|_2^2 \le \|\Phi x\|_2^2 \le (1 + \delta_K)\|x\|_2^2. \tag{3}$$

In words, the K-RIP ensures that all submatrices of Φ of size $M \times K$ are close to an isometry, and therefore distance (and information) preserving. Practical recovery algorithms typically require that Φ have a slightly stronger $2K$-RIP or higher-order RIP in order to preserve differences of K-sparse vectors (which are $2K$-sparse in general) and other higher-order structures [1, 6, 12].

While the design of a measurement matrix Φ satisfying the K-RIP is an NP-Complete problem in general [1], random matrices whose entries are i.i.d. Gaussian, Bernoulli (± 1), or more generally subgaussian[2] work with high probability, provided $M = \mathcal{O}\left(K \log(N/K)\right)$. These random matrices also have a so-called *universality* property in that, for any choice of orthonormal basis matrix Ψ, $\Phi\Psi$ has the K-RIP with high probability. This is useful when the signal is sparse in a basis other than the identity. A random Φ corresponds to an intriguing data acquisition protocol in which each measurement y_j is a randomly weighted linear combination of the entries of x.

A number of different CS signal recovery algorithms, both from optimization and greedy approaches [1–3, 5, 6], offer provably stable signal recovery with performance close to optimal K-term approximation. For a matrix Φ that has the $2K$-RIP and noisy measurements $y = \Phi x + n$, the recovered signal $\hat{x}$ holds the guarantee

$$\|x - \hat{x}\|_2 \le C_1 \|x - x_K\|_2 + \frac{C_2}{\sqrt{K}} \|x - x_K\|_1 + C_3 \|n\|_2, \tag{4}$$

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

with C_1 and C_2 denoting constants. This result has many implications. Under noiseless measurements, K-sparse signals are recovered perfectly; under noisy measurements, an s-compressible signal has recovery error

$$\|x - \widehat{x}\|_2 \leq \frac{C_1 S K^{-s}}{\sqrt{2s}} + \frac{C_2 S K^{-s}}{s - 1/2} + C_3 \|n\|_2. \quad (5)$$

III. MODEL-SPARSE AND COMPRESSIBLE SIGNALS

While many natural and manmade signals and images can be described to first-order as sparse or compressible, the support of their large coefficients often has an underlying inter-dependency structure. This phenomenon has received only limited attention by the CS community to date [9–11, 13–16]. In this section, we present a theory of CS that captures such structure using a union-of-subspaces model. Such a model reduces the degrees of freedom of a sparse/compressible signal by permitting only certain configurations of supports for the large coefficients. As we will show, this allows us to reduce, in some cases significantly, the number of compressive measurements M required to stably recover a signal.

A. Model-sparse signals

Recall from Section II that a K-sparse signal vector x lives in $\Sigma_K \subset \mathbb{R}^N$, which is a union of $\binom{N}{K}$ subspaces of dimension K. Other than its K-sparsity, there are no further constraints on the support or values of its coefficients. A *union-of-subspaces signal model* (a *signal model* in the sequel for brevity) endows the K-sparse signal x with additional structure that allows certain K-dimensional subspaces in Σ_K and disallows others [9, 10]. More formally, 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. A signal model $\mathcal{M}_K$ is then defined as the union of m_K canonical K-dimensional subspaces

$$\mathcal{M}_K = \bigcup_{m=1}^{m_K} \mathcal{X}_m, \text{ s.t. } \mathcal{X}_m := \{x : x|_{\Omega_m} \in \mathbb{R}^K, x|_{\Omega_m^C} = 0\};$$

each subspace $\mathcal{X}_m$ contains all signals x with $\mathrm{supp}(x) \in \Omega_m$. Thus, the signal model $\mathcal{M}_K$ is defined by the set of possible supports $\{\Omega_1, \ldots, \Omega_{m_K}\}$. Signals from $\mathcal{M}_K$ are called K-*model sparse*. Clearly, $\mathcal{M}_K \subseteq \Sigma_K$ and contains $m_K \leq \binom{N}{K}$ subspaces. In the sequel, we will use an algorithm $\mathbb{M}(x, K)$ that returns the best K-term approximation of the signal x under the signal model $\mathcal{M}_K$.

If we know that the signal x being acquired is K-model sparse, then we can relax the RIP constraint on the CS measurement matrix Φ and still achieve stable recovery from the compressive measurements $y = \Phi x$. A *model-based* RIP requires that (3) holds *only* for signals $x \in \mathcal{M}_K$ [9, 10]; we denote this new property as $\mathcal{M}_K$-RIP to specify the dependence on the chosen signal model, and change the model-based RIP constant from δ_K to $\delta_{\mathcal{M}_K}$ for clarity. Blumensath and Davies [9] have quantified 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

$$M \geq \frac{2}{c\delta_{\mathcal{M}_K}^2} \left(\ln(2m_K) + K \ln \frac{12}{\delta_{\mathcal{M}_K}} + t \right). \quad (6)$$

This bound can be used to recover the conventional CS result by substituting $m_K = \binom{N}{K} \approx (Ne/K)^K$. The $\mathcal{M}_K$-RIP property is sufficient for robust recovery of model-sparse signals using recovery algorithms like model-based CoSaMP [8].

B. Model-compressible signals

Just as compressible signals are "nearly K-sparse" and thus live close to the union of subspaces Σ_K in $\mathbb{R}^N$, *model-compressible signals* are "nearly K-model sparse" and live close to the restricted union of subspaces $\mathcal{M}_K$. To make this new concept rigorous, recall from (2) that we defined compressible signals in terms of the decay of their K-term approximation error. The ℓ_2 error incurred by approximating $x \in \mathbb{R}^N$ by the best model-based approximation in $\mathcal{M}_K$ is given by

$$\sigma_{\mathcal{M}_K}(x) := \inf_{\bar{x} \in \mathcal{M}_K} \|x - \bar{x}\|_2 = \|x - \mathbb{M}(x, K)\|_2.$$

The decay of this approximation error defines the model-compressibility of a signal; we define the set of *s-model-compressible signals* as

$$\mathfrak{M}_s = \{x \in \mathbb{R}^N \ : \ \sigma_{\mathcal{M}_K}(x) \leq S K^{-1/s},$$
$$1 \leq K \leq N, S < \infty\},$$

i.e., those signals whose model-based approximation error has a power-law decay with exponent s as K increases.

IV. THE RESTRICTED AMPLIFICATION PROPERTY (RAMP)

In conventional CS, the same requirement (RIP) is a sufficient condition for the stable recovery of both sparse and compressible signals. In model-based recovery, however, the class of compressible signals is much larger than that of sparse signals, since the set of subspaces containing model-sparse signals does not span all K-dimensional subspaces. Therefore, we need to introduce some additional tools to develop a sufficient condition for the stable recovery of model-compressible signals.

We will pay particular attention to signal models $\mathcal{M}_K$ that generate *nested approximations*, since they are more amenable to analysis. In words, a signal model generates nested approximations if the support of the best K'-term model-based approximation contains the support of the best K-term model-based approximation for all $K < K'$. An important example of a signal model that generates nested approximations is the standard compressible signal model of (2).

When a signal model generates nested approximations, the support of the difference between the best jK-term model-based approximation and the best $(j+1)K$-term model-based approximation of a signal can be shown to lie in a small union of K-dimensional subspaces, thanks to the structure enforced by the signal model. This structure is captured by the set of *residual subspaces* that are included in each subsequent approximation. We group them under the notation

$$\begin{aligned}
\mathcal{R}_{j,K}(\mathcal{M}) \quad = \quad & \{u \in \mathbb{R}^N \text{ s.t. for some } x \in \mathbb{R}^N, \\
& u = \mathbb{M}(x, jK) - \mathbb{M}(x, (j-1)K)\},
\end{aligned}$$

for $j = 1, \ldots, \lceil N/K \rceil$. Each signal x in a signal model can be partitioned into its best K-term approximation x_{T_1}, the additional components present in the best $2K$-term approximation x_{T_2}, and so on, with $x = \sum_{j=1}^{\lceil N/K \rceil} x_{T_j}$ and $x_{T_j} \in \mathcal{R}_{j,K}(\mathcal{M})$ for each j. Each signal partition x_{T_j} is a K-sparse signal, and thus $\mathcal{R}_{j,K}(\mathcal{M})$ is a union of subspaces of dimension K. We will denote by R_j the number of subspaces that compose $\mathcal{R}_{j,K}(\mathcal{M})$, omitting the dependence on $\mathcal{M}$ for brevity.

For exactly K-model-sparse signals, we discussed in Section III that the number of measurements M required for a random matrix to have the $\mathcal{M}_K$-RIP is determined by the number of canonical subspaces m_K via (6). Unfortunately, such model-sparse concepts and results do not immediately extend to model-compressible signals. Thus, we develop a generalization of the $\mathcal{M}_K$-RIP that we will use to quantify the stability of recovery for model-compressible signals.

To analyze the robustness of compressible signal recovery in conventional CS, we can consider the tail of the signal outside its K-term approximation as contributing additional "noise" to the measurements of size $\|\Phi(x-x_K)\|_2$ [6, 12]. Consequently, the conventional K-sparse recovery performance result can be applied with the augmented noise $n + \Phi(x - x_K)$. This technique can also be used to quantify the robustness of model-compressible signal recovery. The key quantity we must control is the amplification of the model-based approximation residual through Φ. The following property is a new generalization of the RIP and model-based RIP.

Definition 2: A matrix Φ has the (ϵ_K, r)-*restricted amplification property* (RAmP) for the residual subspaces $\mathcal{R}_{j,K}$ of the signal model $\mathcal{M}$ if

$$\|\Phi u\|_2^2 \leq (1 + \epsilon_K) j^{2r} \|u\|_2^2$$

for any $u \in \mathcal{R}_{j,K}$ and for each $1 \leq j \leq \lceil N/K \rceil$. Intuitively, the norms of the partitions $\|x_{T_j}\|_2$ decay as j increase for signals that are compressible under the signal model. This observation is instrumental in relaxing the isometry restrictions on the measurement matrix Φ and bounding the recovery error for s-model-compressible signals. The regularity parameter $r > 0$ caps the growth rate of the amplification of $u \in \mathcal{R}_{j,K}$ as a function of j. Its value can be chosen so that the growth in amplification with j balances the decay of the norm in each residual subspace $\mathcal{R}_{j,K}$ with j.

We can quantify the number of rows M required for a random measurement matrix Φ to have the RAmP with high probability; we prove the following in [8].

Theorem 1: Let Φ be an $M \times N$ matrix with i.i.d. subgaussian entries and let the set of residual subspaces $\mathcal{R}_{j,K}$ of the signal model $\mathcal{M}$ contain R_j subspaces of dimension K for each $1 \leq j \leq \lceil N/K \rceil$. If

$$M \geq \max_{1 \leq j \leq \lceil N/K \rceil} \frac{2K + 4 \ln \frac{R_j N}{K} + 2t}{\left(j^r \sqrt{1 + \epsilon_K} - 1\right)^2},$$

then Φ has the (ϵ_K, r)-RAmP with probability $1 - e^{-t}$. The crux of the theorem is that if the sequence $\{R_j\}$ has slow growth, then a small number of measurements will suffice for robust recovery of model-compressible signals. The order of the bound of Theorem 1 is lower than $\mathcal{O}(K \log(N/K))$ as long as the number of subspaces R_j grows slower than N^K. Armed with the RaMP, we can state the following result, which will provide robustness for the recovery of model-compressible signals; see [8] for the proof.

Theorem 2: Let $x \in \mathfrak{M}_s$ be an s-model compressible signal under a signal model $\mathcal{M}$ that generates nested approximations. If Φ has the (ϵ_K, r)-RAmP and $r = s - 1$, then we have

$$\|\Phi(x - \mathbb{M}(x, K))\|_2 \leq \sqrt{1 + \epsilon_K} S K^{-s} \ln \left\lceil \frac{N}{K} \right\rceil. \quad (7)$$

This result, combined with (4), provides us with a recovery guarantee for compressible signals:

$$\|x - \widehat{x}\| \leq \frac{C_1 S}{K^{-s}} + C_2 \left(\|n\|_2 + \sqrt{1 + \epsilon_K} S K^{-s} \ln \left\lceil \frac{N}{K} \right\rceil \right),$$

with C_1 and C_2 denoting constants. For matrices having the RAmP, the model-based recovery algorithms of [8] provide guarantees for model-compressible signals that are similar to that of standard algorithms in (5).

V. EXAMPLE: WAVELET TREE MODEL

Wavelet decompositions have found wide application in the analysis, processing, and compression of smooth and piecewise smooth signals because these signals are K-sparse and compressible, respectively [17]. Moreover, the wavelet coefficients can be naturally organized into a tree structure, and for many kinds of natural and manmade signals the largest coefficients cluster along the branches of this tree. This motivates a connected tree model for the wavelet coefficients. We first describe tree sparsity in the context of sparse wavelet decompositions. We focus on one-dimensional signals and binary wavelet trees, but all of our results extend directly to d-dimensional signals and 2^d-ary wavelet trees.

Consider a signal x of length $N = 2^I$, for an integer value of I. The wavelet representation of x is given by

$$x = v_0 \nu + \sum_{i=0}^{I-1} \sum_{j=0}^{2^i-1} w_{i,j} \psi_{i,j},$$

where ν is the scaling function and $\psi_{i,j}$ is the wavelet function at scale i and offset j. The wavelet transform consists of the scaling coefficient v_0 and wavelet coefficients $w_{i,j}$ at scale i, $0 \leq i \leq I-1$, and position j, $0 \leq j \leq 2^i - 1$. In terms of our earlier matrix notation, x has the representation $x = \Psi \alpha$, where Ψ is a matrix containing the scaling and wavelet functions as columns, and $\alpha = [v_0 \ w_{0,0} \ w_{1,0} \ w_{1,1} \ w_{2,0} \ldots]^T$ is the vector of scaling and wavelet coefficients. We are, of course, interested in sparse and compressible α.

The nested supports of the wavelets at different scales create a parent/child relationship between wavelet coefficients at different scales. We say that $w_{i-1,\lfloor j/2 \rfloor}$ is the *parent* of $w_{i,j}$ and that $w_{i+1,2j}$ and $w_{i+1,2j+1}$ are the *children* of $w_{i,j}$. Wavelet functions act as local discontinuity detectors; using the nested support property of wavelets at different scales, it is straightforward to see that a signal discontinuity will give rise to a chain of large wavelet coefficients along a branch of the wavelet tree from a leaf to the root. Moreover, smooth signal regions will give rise to regions of small wavelet coefficients. This "connected tree" property has been well-exploited in a number of wavelet-based processing and compression algorithms [18, 19]. In this section, we will specialize the theory from Section III to a connected tree model $\mathcal{T}$.

A set of wavelet coefficients Ω forms a *connected subtree* if, whenever a coefficient $w_{i,j} \in \Omega$, then its parent $w_{i-1,\lfloor j/2 \rfloor} \in \Omega$ as well. Each such set Ω defines a subspace of signals whose support is contained in Ω; that is, all wavelet coefficients outside Ω are zero. In this way, we define the signal model $\mathcal{T}_K$ as the union of all K-dimensional subspaces corresponding to supports Ω that form connected subtrees.

For tree-compressible signals, we apply Theorem 1 to find that a subgaussian random matrix has the (ϵ_K, s)-RAmP for the signal model $\mathcal{T}$ and all $s > 0.5$, with probability $1 - e^{-t}$, if the number of measurements obeys

$$M \geq \frac{20K + 4\ln\frac{601N}{K^3} + 2t}{\left(\sqrt{1+\epsilon_K} - 1\right)^2} \tag{8}$$

when $K \geq \log_2 N$, or

$$M \geq \frac{20K + 4\ln\frac{N}{2K^3+3K^2+K} + 2t}{\left(\sqrt{1+\epsilon_K} - 1\right)^2} \tag{9}$$

when $K < \log_2 N$; see [8] for the proofs. Both cases give a simplified bound on the number of measurements required as $M = \mathcal{O}(K)$, which is a substantial improvement over the $M = \mathcal{O}(K \log(N/K))$ required by conventional CS recovery methods. Thus, when Φ satisfies (8-9), we have the guarantee (7) for s-tree-compressible signals.

VI. EXPERIMENTS

We now present the results of a number of numerical experiments that illustrate the effectiveness of model-based recovery. Our consistent observation is that explicit incorporation of the tree model during recovery significantly improves its performance for a given number of measurements.

We study one-dimensional piecewise polynomial signals that match the connected wavelet-tree model described above. We employ the CoSaMP recovery algorithm of [6] and the model-based CoSaMP recovery algorithm of [8] using CSSA, a tree-based approximation algorithm [20], in all experiments.

Standard CS recovery algorithms require that the overmeasuring factor M/K is approximately logarithmic in N. In contrast, (8) suggests that stable CS recovery of wavelet tree-compressible signals can be performed with only $M = \mathcal{O}(K)$ measurements; thus, the overmeasuring factor does not increase with the signal length N. Figure 1 shows the growth of the overmeasuring factor with the signal length N for conventional CS and model-based recovery. We generated 50 sample piecewise cubic signals and numerically computed the minimum number of measurements M required for the recovery error $\|x - \widehat{x}\|_2 \leq 2.5\sigma_{\mathcal{T}_K}(x)$, the *best* tree-approximation error, for every sample signal. The figure shows that while doubling the signal length increases the number of measurements required by standard recovery by K, the number of measurements required by model-based recovery is constant for all N. These experimental
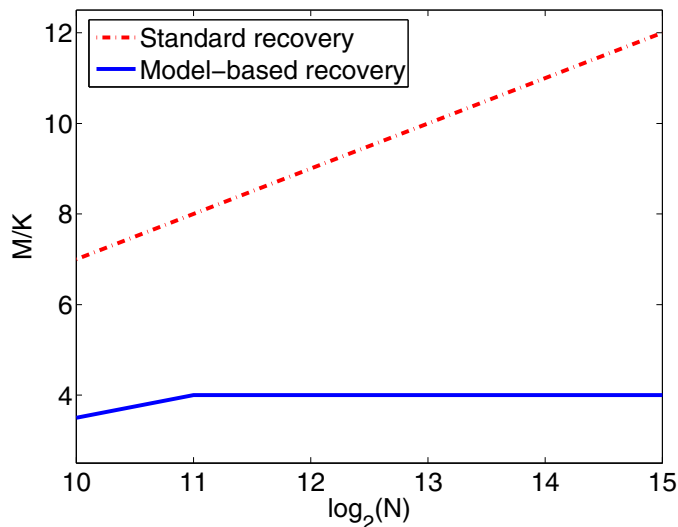
Fig. 1. Required overmeasuring factor M/K to achieve a target recovery error $\|x - \widehat{x}\|_2 \leq 2.5\sigma_{T_K}(x)$ as a function of the signal length N for standard and model-based recovery of piecewise smooth signals. While standard recovery requires M to increase logarithmically with N, the required M is essentially constant for model-based recovery.

results verify the theoretical performance described in Section V.

VII. Conclusions

In this paper, we have shown that significant performance gains can be obtained by exploiting more realistic and richer signal models beyond the simplistic sparse and compressible models that dominate the CS literature. Building on the unions of subspaces results of [9–11], we have taken some of the first steps towards what promises to be a general theory for model-based CS by introducing the notion of a model-compressible signal and the associated restricted amplification property (RAmP) condition it imposes on the measurement matrix Φ. For the volumes of natural and manmade signals and images that are wavelet-sparse or compressible, model-based CS offers performance that significantly exceeds today's state-of-the-art, requiring only $M = \mathcal{O}(K)$ rather than $M = \mathcal{O}(K \log(N/K))$ random measurements. Such savings become significant for high-dimensional signals such as high resolution images.

References

[1] E. J. Candès, "Compressive sampling," in *Proc. International Congress of Mathematicians*, vol. 3, (Madrid, Spain), pp. 1433–1452, 2006.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, pp. 1289–1306, Sept. 2006.

[3] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Mag.*, vol. 24, no. 4, pp. 118–120, 124, July 2007.

[4] J. Tropp and A. C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Trans. Info. Theory*, vol. 53, pp. 4655–4666, Dec. 2007.

[5] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Trans. Info. Theory*, vol. 52, pp. 4036–4048, Sept. 2006.

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

[7] V. Cevher, M. F. Duarte, C. Hegde, and R. G. Baraniuk, "Sparse signal recovery using Markov Random Fields," in *Proc. Workshop on Neural Info. Proc. Sys. (NIPS)*, (Vancouver, Canada), Dec. 2008.

[8] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," 2008. Preprint. Available at http://dsp.rice.edu/cs.

[9] T. Blumensath and M. E. Davies, "Sampling theorems for signals from the union of finite-dimensional linear subspaces," *IEEE Trans. Info. Theory*, Dec. 2008. To be published.

[10] Y. M. Lu and M. N. Do, "Sampling signals from a union of subspaces," *IEEE Signal Processing Mag.*, vol. 25, pp. 41–47, Mar. 2008.

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

[12] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Compte Rendus de l'Academie des Sciences, Series I*, vol. 346, pp. 589–592, May 2008.

[13] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," Mar. 2008. Preprint.

[14] 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.

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

[16] M. F. Duarte, M. B. Wakin, and R. G. Baraniuk, "Wavelet-domain compressive signal reconstruction using a hidden Markov tree model," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, (Las Vegas, NV), pp. 5137–5140, April 2008.

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

[18] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using Hidden Markov Models," *IEEE Trans. Signal Processing*, vol. 46, pp. 886–902, Apr. 1998.

[19] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Proc.*, vol. 41, pp. 3445–3462, Dec. 1993.

[20] R. G. Baraniuk, "Optimal tree approximation with wavelets," in *Wavelet Applications in Signal and Image Processing VII*, (Denver, CO), pp. 196–207, July 1999.