Texas Hold 'Em Algorithms for Distributed Compressive Sensing

Stephen R. Schnelle, Jason N. Laska, Chinmay Hegde, Marco F. Duarte, Mark A. Davenport, Richard G. Baraniuk

Department of Electrical and Computer Engineering, Rice University, Houston, TX 77005
Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544

ABSTRACT

This paper develops a new class of algorithms for signal recovery in the distributed compressive sensing (DCS) framework. DCS exploits both intra-signal and inter-signal correlations through the concept of joint sparsity to further reduce the number of measurements required for recovery. DCS is well-suited for sensor network applications due to its universality, computational asymmetry, tolerance to quantization and noise, and robustness to measurement loss. In this paper we propose recovery algorithms for the sparse common and innovation joint sparsity model. Our approach leads to a class of efficient algorithms, the Texas Hold 'Em algorithms, which are scalable both in terms of communication bandwidth and computational complexity.

Index Terms— Signal reconstruction, multisensor systems, data compression

1. INTRODUCTION

Compressive sensing (CS) is an emerging framework for the acquisition of signals $x \in \mathbb{R}^N$ that are exactly or approximately $K$-sparse in a known basis [1, 2]. In CS, we acquire $x$ via a length-$M$ measurement vector $y = \Phi x$, where $\Phi$ is a measurement matrix, usually sporting randomly drawn independent entries. When $K$ is sufficiently small, the number of measurements $M$ can be much less than the dimension of the signal $N$. A compelling example of this type of applications is sensor networks, where there has been significant interest in CS [3–7]. In distributed compressive sensing (DCS) we aim to recover the data acquired from a group of $J$ sensors from compressive measurements that are obtained either independently [3, 4] or collaboratively [5–7]. Most existing adaptations of CS to sensor networks require the collection of up to $O(J)$ measurements per sensor at a centralized location. The resulting communication scheme leads to congestion collapse when the network size becomes significant. Furthermore, the computational complexity of the algorithms used for signal recovery are typically at least quadratic in the number of measurements, further diminishing the scalability of CS to large sensor network sizes.

In this paper, we introduce a new algorithm, dubbed the Texas Hold 'Em algorithm, tailored for the DCS scenario in which each sensor observes the combination of a common sparse component and a unique innovation component [3, 4]. The algorithm is the first practical recovery method proposed for this setting. Our procedure separates the recovery of the common and innovation components into two stages, which in some cases allows us to significantly reduce the number of measurements and improve on the lowest signal-to-noise ratio (SNR) that still enables accurate recovery of the signal ensemble. The measurements used to recover the common component are obtained by either averaging or concatenating measurements from all the sensors in the network, which makes the amount of communication per sensor independent from the network size $J$. Furthermore, each innovation component is recovered locally at the corresponding sensor, making the recovery computational complexity only linear in $J$.

This paper is organized as follows. In Section 2 we provide the necessary background on CS and DCS, and in Section 3 we describe the Texas Hold 'Em algorithms. In Section 4 we provide some preliminary experimental results illustrating the performance of our algorithms. Finally, Section 5 concludes with a brief discussion.

2. BACKGROUND

2.1. Compressive Sensing (CS)

We first provide a brief overview of the CS framework. To begin, we acquire a signal $x \in \mathbb{R}^N$ via the linear measurements

$$y = \Phi x + e,$$

where $\Phi$ is an $M \times N$ measurement matrix modeling the sampling system, $y \in \mathbb{R}^M$ is the vector of samples acquired, and $e$ is an $M \times 1$ vector that represents measurement errors. If
where $x_j \in \mathbb{R}^N$. We assume that there exists a known sparse basis $\Psi$ for $\mathbb{R}^N$ in which $x_j$ can be sparsely represented. JSM-1 assumes that all signals share a common sparse component while each individual signal contains a sparse innovations component:

$$x_j = z_c + z_j, \quad j \in \{1, 2, \ldots, J\}$$

with $z_c = \Psi_0 \alpha_c, \|\alpha_c\|_0 = K_c$, and $z_j = \Psi_\alpha_j, \|\alpha_j\|_0 = K_j$.

Thus, the signal $z_c$ is common to all of the $x_j$ and has sparsity $K_c$ in the basis $\Psi$. The signals $z_j$ are the unique portions of the $x_j$ and have sparsity $K_j$ in the basis $\Psi$.

A practical situation well-suited to this model is a group of sensors measuring temperatures at a number of locations throughout the day. The temperature readings $x_j$ have both temporal (intra-signal) and spatial (inter-signal) correlations. Global factors, such as the sun and prevailing winds, could have an effect $z_c$ that is both common to all sensors and structured enough to permit sparse representation. More local factors such as shade, water, or animals could contribute localized innovations $z_j$ that are also structured (and hence sparse). A similar scenario could be imagined for a sensor network recording light intensities, air pressure, or other phenomena. All of these scenarios correspond to measuring properties of physical processes that change smoothly in time and space and thus are highly correlated.

### 3. THE TEXAS HOLD 'EM ALGORITHM

#### 3.1. Averaged Community Texas Hold 'Em

Before we state our proposed algorithm, we must first set some notation. We begin by noting that our measurements are obtained via $y_j = \Phi_j x_j$. We then decompose each $y_j$ into two components, $[y_j^c \ y_j^h]$. We let $M_c$ and $M_h$ denote the length of $y_j^c$ and $y_j^h$, so that $M = M_c + M_h$. The measurements $y_j^c$ are the community measurements. These measurements are shared among all the sensors. The measurements $y_j^h$ are the hold measurements, which are not transmitted but instead are retained by each sensor. Similarly, we let $\Phi_j^c$ and $\Phi_j^h$ denote the matrices obtained by selecting the appropriate rows of $\Phi_j$.

The Texas Hold 'Em algorithm provides an approach for recovering each $x_j$. The algorithm proceeds by first fusing the community measurements to obtain an estimate of $z_c$ so that each sensor is then able to use its hold measurements to recover $z_j$. While we will describe a number of simple variations on this algorithm below, the variant we analyze in this work assumes that $\Phi_j^c$ is fixed for all $j$. We then define

$$\hat{y} = \sum_{j=1}^J y_j^c / J.$$ 

We also let $\hat{\Phi} = \Phi_j^c$. Observe that $\hat{y}$ simply averages the community measurements, and thus we dub this variation of the algorithm Averaged Community.
We describe how the Texas Hold 'Em algorithm recovers a particular $z_j$ in Algorithm 1. First, we attempt to estimate the common component $z_c$ from $\bar{y}$. This is denoted via the notation $\text{RECOVER}(\Phi, y, K)$, which simply denotes any sparse recovery algorithm, taking as inputs a dictionary $\Phi$, a measurement vector $y$, and a sparsity $K$, and returning a $K$-sparse signal estimate $\hat{\alpha}$ for which a guarantee such as (3) holds. We then remove the contribution of the common component to all available measurements by forming the updated measurement vectors $\bar{y}_j = y - \Phi_j \hat{\alpha}$. The algorithm then attempts to recover the different innovation components from the vectors $\bar{y}_j$.

Note that Algorithm 1 requires two phases of communication. First, we must compute $\hat{y}$; we can do this by collecting all $y_j^c$ at a central location in a multicast network architecture or by using more intelligent network protocols in a multihop network architecture. Second, we relay the estimate $\hat{\alpha}_c$ to all sensors either via a broadcast or multihop network.

3.2. Performance bounds

We now provide some intuition behind the estimation step for $z_c$ featured in the Texas Hold 'Em algorithm.\(^1\) Observe that
\[
\bar{y} = \bar{\Phi} (z_c + \bar{z}),
\]
where $\bar{z} = \sum_{j=1}^J z_j / J$. For large values of $J$, provided that the innovation components are not overly coherent, $\bar{z}$ tends to zero; thus, we can expect that $\bar{y}$ preserves the common component while cancelling out the contribution of the innovations. This is made precise in the following theorem.

**Theorem 3.1.** Let $K = \max(2K_c, 2K_j)$. Suppose that $\Phi_j$ is fixed for all $j$ to be an $M \times N$ random matrix with $M_c = \mathcal{O}(K \log(N/K))$. Suppose also that $\|z_c\|_0 = K_c$, $\|z_j\|_0 = K_j$, and $\|z_j\|_2 = \kappa$ for each $j$, and the $z_j$ are pairwise orthogonal. Then with high probability
\[
\|\bar{x}_j - x_j\|_2 \leq C' \kappa / \sqrt{J},
\]
where $C'$ depends only on the constant from (3).

**Proof.** We begin by noting that
\[
\|\bar{x}_j - x_j\|_2 \leq \|\bar{z}_j - z_j\|_2 + \|\bar{z}_c - z_c\|_2.
\]

From the orthogonality and norm assumptions on $z_j$, we can easily show that $\|\bar{z}\|_2 = \kappa / \sqrt{J}$. From the assumptions on $M$, we have that with high probability [8], $\bar{\Phi}$ satisfies the restricted isometry property (RIP) of order $K$ with constant $\delta$, $\Phi_j \bar{\Phi}$ satisfies the RIP of order $K'$ with constant $\delta$, and
\[
\|\bar{\Phi} z\|_2 \leq \sqrt{1 + \delta} \|z\|_2 = \sqrt{1 + \delta} / \sqrt{J}.
\]
By assumption, we also have that the recovery algorithm used to estimate $\bar{z}_c$ satisfies (3), and thus from (4) we have that
\[
\|\bar{z}_c - z_c\|_2 \leq C \sqrt{1 + \delta \kappa / \sqrt{J}}.
\]
Furthermore, since $\bar{z}_c - z_c$ is $2K_c$-sparse and $\Phi_j \bar{\Phi}$ satisfies the RIP of order at least $2K_c$, we have
\[
\|\Phi_j (\bar{z}_c - z_c)\|_2 \leq C(1 + \delta) \kappa / \sqrt{J}.
\]
From this we observe that since $\bar{y}_j = \Phi_j z_j + \Phi_j (z_c - \bar{z}_c)$, we may again apply (3) to obtain
\[
\|\bar{z}_j - z_j\|_2 \leq C^2 (1 + \delta) \kappa / \sqrt{J}.
\]

3.3. Variations

Above we assumed that $\Phi_j$ was fixed for all $j$. This is not actually necessary. To modify the algorithm for the more general case where each $\Phi_j$ can be selected independently, we simply need to redefine $\bar{y}$ and $\bar{\Phi}$ according to
\[
\bar{y} = [(y_1^T, y_2^T, \ldots, y_J^T)^T]^T,
\bar{\Phi} = [(\Phi_1^T, \Phi_2^T, \ldots, \Phi_J^T)^T]^T.
\]
This method, which we call Combined Community Texas Hold 'Em to distinguish it from Algorithm 1, will require increased computation and communication costs. However, it may result in better performance as the estimate of the common component should be more accurate. A second variation which we do not explore here replaces the cancellation step with the compressive-domain interference cancellation approach of [13]. Since the estimate $\bar{z}_c$ is likely noisy, this approach would completely remove the contribution of $z_c$, provided that we have correctly identified the support of $\alpha_c$.

4. EXPERIMENTS

We now present preliminary experimental results illustrating the performance of the Texas Hold 'Em algorithm. In Figure 1 we show the normalized mean-squared error (MSE) for Algorithm 1 as we increase the number of measurements per sensor as a multiple of the total signal sparsity $K = K_c + K_j$. In our example, we set $N = 512$, $K_c = 10$, and $K_j = 30$. The common component nonzero coefficients are 4 times as large as the innovative components, enough to ensure that the innovative components can be averaged away. We perform
the experiment for several different percentages of community measurements per sensor – 20%, 50%, or 75% – and we compare against traditional reconstruction on each sensor independently, i.e., 0% community. For lower values of $M$ sharing measurements results in lower error. For values of $M$ that provide feasible recovery (i.e at least $M/K \approx 2$), we see the expected decay in MSE; additionally, sharing a larger fraction of measurements improves recovery performance. For larger number of measurements $M$, the advantage of sharing most measurements vanishes.

5. DISCUSSION

In a possible extension to our framework, one could consider the case where the common component is no longer sparse in a known basis; in mathematical terms,

$$x_j = z_c + z_j, \quad j \in \{1, 2, \ldots, J\}$$

$$z_j = \Psi \alpha_j, \quad \|\alpha_j\|_0 = K_j,$$

but $z_c$ is not necessarily sparse in the basis $\Psi$. This model, referred to as JSM-3 [4], is relevant in situations where several sources are recorded by different sensors together with a background signal that is not sparse in any basis. Consider, for example, a computer vision-based verification system in a device production plant. Cameras acquire snapshots of components in the production line; a computer system then checks for failures in the devices for quality control purposes. While each image could be extremely complicated, the ensemble of images will be highly correlated, since each camera observes the same device with minor (sparse) variations.

Our approach can be immediately applied to this scenario with only two main differences. First, the total number of community measurements must be sufficiently large to enable recovery of $z_c$, and second, we must have an alternative method for recovery of $z_c$. In the total absence of any prior knowledge concerning $z_c$, a reasonable approach is to require at least $N$ community measurements so that a least-squares approach to estimating $z_c$ can be reasonably accurate.

6. REFERENCES