Compressive Sampling and Least Squares based Reconstruction of Correlated Signals

Ali Ahmed and Fahad Shamshad
Department of Electrical Engineering, Information Technology University, Lahore, Pakistan.
Email: {ali.ahmed, fahad.shamshad}@itu.edu.pk

***Abstract*—This paper presents a novel sampling scheme for the acquisition of an ensemble of *correlated* (lying in an a priori unknown subspace) signals at a sub-Nyquist rate. We propose an implementable sampling architecture that acquires structured samples of the signals. We then show that a much fewer of these samples compared to what is dictated by the Shannon-Nyquist sampling theorem suffice for exact signal reconstruction. Quantitatively, we show that an ensemble of M correlated signals each of which is bandlimited to $W/2$ and can be expressed as the linear combination of R underlying signals can be acquired at roughly RW (to with log factors) samples per second. This is a considerable gain in sampling rate compared to MW samples required by Shannon-Nyquist sampling in the case when $M \gg R$.**

We propose a simple least squares program for the reconstruction of the correlated signal ensemble. This result is in stark contrast with the previous work, where a prohibitively computationally expensive semidefinite program is required for signal reconstruction.

I. Introduction

This paper presents an implementable sampling architecture for the acquisition of an ensemble of correlated signals at a sub-Nyquist rate. We augment it with a sampling theorem showing that the sampling rate does not scale with the number of signals in the ensemble but only with the *degree of correlation* among them. We further show that signals can be reconstructed by solving a simple least squares program. This is in stark contrast to earlier work [1]–[5] on the compressive sampling of correlated signals that suggests solving a computationally prohibitive semidefinite program for signal reconstruction.

We consider ensemble of M signals, each of which is bandlimited to frequencies below $W/2$ (see Figure I). The M signals lie in an a priori unknown R-dimensional subspace meaning that each of the signals can be decomposed as the linear combination of R underlying signals. Shannon-Nyquist sampling theorem dictates that the ensemble can be reconstructed from rate MW uniformly spaced samples; W samples for each signal in the ensemble. In comparison, we present a sampling strategy that can acquire the correlated ensemble at a rate $\approx RW$ that scales with the inherent number R of the signals rather than with the total number M of signals. We want to emphasize here that no a priori knowledge of signal subspace is assumed.

The sampling architecture, shown in Figure I, is composed of implementable components including analog vector-matrix multiplier (AVMM), modulators, low pass filters and ADCs. The main strategy is to preprocess the signal before acquiring the samples. Intuitively, the preprocessing step ensures that the information in the signal ensemble is dispersed across the M sensors and in time so that each sample in some sense contains a global information about the ensemble. This allows us to take a few samples and still be able to reconstruct the signals exactly.

We now briefly state the task of each of the component in the sampling architecture. The AVMM is input with M signals in the ensemble and it produces a fewer number Δ of signals at the output. Each of the output signal is generated

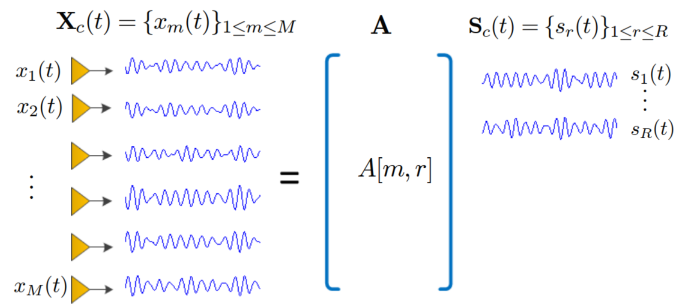

Fig. 1. Acquire an ensemble of M signals, each bandlimited to $W/2$ radians per second. The signals are correlated — M signals can be well approximated by the linear combination of R underlying signals. Quantitatively, we can write M signals in ensemble $\boldsymbol{X}_c(t)$ (on the left) as a tall matrix (a correlation structure) multiplied by an ensemble of R underlying independent signals.

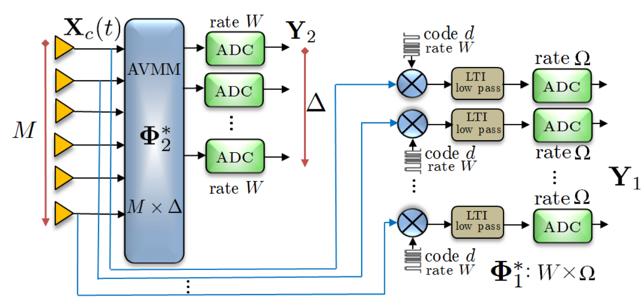

Fig. 2. Sampling architecture for multiple signals lying in a subspace: M signals, bandlimited to $W/2$ are preprocessed in analog using an analog vector-matrix multiplier (AVMM) to produce Δ signals, each of which is then sampled at W samples per second. In addition, each of the M input signal is processed by a modulator, and a low-pass filter. The resultant signal is then sampled uniformly at a rate Ω samples per second. The analog preprocessing is designed to perform the row and the column operation on $\boldsymbol{X}$ so that a simple least-squares program can be used for decoding. The net sampling rate is $\Omega M + \Delta W$ samples per second.

by the random linear combinations of the M input analog signals. Each of the signal is then sampled at rate W. On the other hand, the modulators multiply in analog each of the M input signals with a random ± 1 binary waveform that switches signs at most W times a second. A modulator is implemented using a simple switching circuit that changes the polarity of the input signal randomly form instant to instant. Low-pass filter are integrators that smooth out high frequency variations in the signals, which are later sampled at a sub-Nyquist rate $\Omega < W$. Our main contribution is to show that the net sampling rate $\Delta W + M\Omega$ is roughly of the order of RW (assuming $W > M$ w.l.o.g.).

Our motivation to study the correlated signal ensemble are the classical problems in array processing, where multiple sources emanate narrow-band signals modulated with a high carrier frequency that arrive on array elements at different spatial locations. The signals received at these array elements are typically heavily spatially correlated and can be very well approximated by a much fewer number of underlying signals. Most of the array processing tasks such as beam forming, interference removal, direction-of-arrival estimation, and multiple source separation exploit this correlation among the signals. All we are doing is to exploit the same correlation structure to obtain gains in the sampling rate. In addition, correlated signals arise in many applications using spatially collocated micro-sensor arrays [6]–[9]; for example, in neural recordings of brain tissue, and in pressure sensing in robotics, thousands of tiny sensor arrays record heavily correlated neural data. Sampling naively at the Nyquist rate would require ADCs with very high rate sampling capacities resulting in much more expensive and not as precise devices. In addition, a huge storage capacity would be required to store several giga bits per second of data generated in typical applications. A compressive acquisition scheme on the other hand can significantly dilute the requirements by orders of magnitude when the signals are heavily correlated. For a more detailed description of the applications, we refer the interested reader to [1], [2], and references therein.

Most of the theoretical results presented are based on ideas derived from the existing literature on randomized SVD; see, for example, [10], [11]. Briefly, we frame the correlated ensemble reconstruction from compressive samples as a low-rank matrix recovery problem. The sampling architecture is designed in a way such that the samples are delivering the row, and column space measurements of the unknown low-rank matrix. After discovering the row and column space, the underlying low-rank matrix can be recovered using a simple least squares program as will be demonstrated in Section IV. The fact that we can use the computationally inexpensive least squares program for signal reconstruction sets this work apart from the earlier work that proposed a prohibitively computationally expensive semidefinite program for this purpose [1]–[5]. A more detailed version of this paper has already appeared in [12].

The layout of the remaining manuscript is as follows. We start with the the signal model in Section II. Section III shows that the samples acquired by the ADCs are basically row, and column space measurements of an underlying unknown low-rank matrix. The sampling theorem along with a brief sketch of its proof is given in Section VI, and VII.

II. Signal Model

We will denote the continuous correlated input ensemble by $\boldsymbol{X}_c(t) = \{x_m(t)\}_{1\le m\le M}$, it can be thought of as a set of M elements $x_m(t)$ each of which is a signal bandlimited [1] to frequencies $W/2$. We take $\boldsymbol{X}_c(t)$ to be a correlated signal ensemble, which means that

$$x_m(t) \approx \sum_{r=1}^{R} A[m,r]s_r(t), 1 \le m \le M,$$

for some scalar $A[m,r]$, and underlying R signals $s_r(t)$ that are also bandlimited to $W/2$. Let $A[m,r]$ be the entries of $\boldsymbol{A}$, and $S_c(t) := \{s_r(t)\}_{1\le r\le R}$. We can concisely write then

$$\boldsymbol{X}_c(t) \approx \boldsymbol{A}\boldsymbol{S}_c(t). \tag{1}$$

We can capture $x_m(t)$ perfectly by taking $W = 2B+1$ equally spaced samples per row. Let $\boldsymbol{X} \in \mathbb{R}^{M\times W}$ be the matrix that contains as its rows the Nyquist rate samples of the signals $x_m(t)$ in $t \in [0,1)$. Let $\boldsymbol{F}$ be a $W \times W$ normalized DFT matrix with entries

$$F[\omega, n] = \frac{1}{\sqrt{W}} e^{-j2\pi\omega n/W}, \quad (\omega, n) \in (\mathcal{W}, \mathcal{N}),$$

where $\mathcal{W} := \{0, \pm 1, \ldots, \pm(W/2-1), W/2\}$, and $\mathcal{N} := \{0, 1, \ldots, W-1\}$. We can write

$$\boldsymbol{X} = \boldsymbol{C}\boldsymbol{F}, \tag{2}$$

where $\boldsymbol{C}$ is a $M \times W$ matrix whose rows contain Fourier series coefficients for the signals in $\boldsymbol{X}_c(t)$. The signals $x_m(t)$ in the ensemble $\boldsymbol{X}_c(t)$ only approximately lie in an R dimensional subspace, and hence the matrices $\boldsymbol{X}$, and $\boldsymbol{C}$ are compressible, rank-R (only top R singular values are significant).

III. System in Discrete Form

In this section, we show that the samples acquired by the proposed sampling architecture in Figure I are the linear measurements of an underlying unknown low-rank matrix. We will show that the recovering this matrix is tantamount to the reconstruction of signal ensemble $\boldsymbol{X}_c(t)$.

We begin by expressing the samples taken by the operating at rate Ω as a linear transform of an underlying unknown low-rank matrix. These ADCs sample modulated and low-pass filtered signal of the ensemble. A modulator simply takes the analog signals $x_m(t)$ and returns the pointwise multiplication $x_m(t) \cdot d(t)$. We will take $d(t)$ to be a binary number ± 1 waveform that is constant over time intervals of a certain length $1/W$. The sign changes of the binary waveforms in each of these intervals occur randomly, and independently. Qualitatively, the random shifts in signal polarity help disperse

[1] To avoid clutter, the signals are also considered to be periodic at this point, mainly to reduce the clutter in mathematics to follow, however, the discussion can also be extended to more general aperiodic signals, see [1]

the energy across the bandwidth of the signal. The low-pass operation will be carried out by integrating the input continuous time signals over an interval $t \in [(n-1)/\Omega, n/\Omega) \subset [0,1)$, where $n \in \{1,\ldots,\Omega\}$. Each of the resultant signal is then sampled at a rate $\Omega < W$.

Since the integration commutes with the modulation, we begin by integrating signals $x_m(t)$ over the interval of width $1/W$, which is also the rate at which modulation occurs, to form a matrix $\boldsymbol{H}$ with entries

$$H[m,n] = \int_{(n-1)/W}^{n/W} x_m(t)dt.$$

Let $x_m(t) = \sum_{\omega\in\mathcal{W}} C[m,\omega]e^{-\iota 2\pi\omega t}, \quad t \in [0,1)$ be the Fourier series expansion of $x_m(t)$, where $C[m,\omega]$; the entries of $\boldsymbol{C}$, denote the Fourier coefficient of $x_m(t)$ at frequency ω. Plugging the Fourier series relation in the integral above and carrying out the integration, we obtain

$$H[m,n] = \sum_{\omega\in\mathcal{W}} C[m,\omega]\left[\frac{e^{\iota 2\pi\omega/W}-1}{\iota 2\pi\omega}\right]e^{-\iota 2\pi\omega n/W}, \quad (3)$$

where the bracketed term depicts the action of integration over an interval of $1/W$ in the frequency domain. Define a $W\times W$ diagonal matrix $\boldsymbol{L}$ with entries

$$L[\omega,\omega] = \left[\frac{e^{\iota 2\pi\omega/W}-1}{\iota 2\pi\omega}\right], \text{ for every } \omega \in \mathcal{W}.$$

Note that $\boldsymbol{L}$ is invertible, and well conditioned. With the DFT matrix $\boldsymbol{F}$ defined earlier, (3) can be expressed in matrix form as $\boldsymbol{H} = \boldsymbol{CLF}$.

Now let $\boldsymbol{D}$ be a $W\times W$ random diagonal matrix containing the sign patterns of $d(t)$ in $t \in [0,1)$, and $\boldsymbol{P} : \Omega\times W$ that contains ones in locations $(\alpha,\beta) \in (j,\mathcal{B}_j)$. For every $j = 1,\cdots,\Omega$, the set $\mathcal{B}_j$ is defined as $\mathcal{B}_j = \{(j-1)W/\Omega+1,(j-1)W/\Omega+2,\ldots,jW/\Omega\}$. Since we have already carried out the integration step, it is easy to see that the uniform Ω rate samples of the signals at the output of low-pass filters are simply $\boldsymbol{Y}_1 = \boldsymbol{HD}^*\boldsymbol{P}^*$. Let us denote $\boldsymbol{\Phi}_1 = \boldsymbol{PD}$ then the measurements $\boldsymbol{Y}_1$ are simply

$$\boldsymbol{Y}_1 = \boldsymbol{H}\boldsymbol{\Phi}_1^*, \quad (4)$$

which will be interpreted as the column space measurements of $\boldsymbol{H}$.

We now express the samples taken by the batch of ADCs operating at rate W as the linear measurements of $\boldsymbol{H}$. Let $\boldsymbol{\Phi}_1$ be a $\Delta\times M$ standard Gaussian random matrix. When AVMM is fed with the ensemble $\boldsymbol{X}_c(t)$ it produces $\boldsymbol{\Phi}_2\boldsymbol{X}_c(t)$. Each rows of $\boldsymbol{\Phi}_2$ contains the random gains to mix the M signals. In discrete time, the action of the AVMM becomes $\boldsymbol{\Phi}_2\boldsymbol{X}$, where $\boldsymbol{X}$ is defined in (2). The Δ rows of $\boldsymbol{\Phi}_2\boldsymbol{X}$ are simply the samples taken by the Δ ADCs after the AVMM in $t \in [0,1)$. We will take $\boldsymbol{\Phi}_2\boldsymbol{X}$ weighted by a known and invertible matrix $\boldsymbol{F}^*\boldsymbol{LF}$ as the observations:

$$\boldsymbol{Y}_2 = \boldsymbol{\Phi}_2\boldsymbol{XF}^*\boldsymbol{LF} = \boldsymbol{\Phi}_2\boldsymbol{CLF} = \boldsymbol{\Phi}_2\boldsymbol{H}, \quad (5)$$

which can be interpreted as the row measurements of $\boldsymbol{H}$.

As mentioned earlier the matrix of Fourier coefficients $\boldsymbol{C}$ is compressible, rank-R. Since $\boldsymbol{F}$ is orthogonal, and $\boldsymbol{L}$ is diagonal, this implies that $\boldsymbol{H} = \boldsymbol{CLF}$ is also a compressible, rank-R matrix. In addition, as $\boldsymbol{LF}$ is square and invertible, the recovery of $\boldsymbol{H}$ implies that we have $\boldsymbol{C}$, and $\boldsymbol{X}$. The signal ensemble $\boldsymbol{X}_c(t)$ can then be reconstructed using sinc interpolation of the samples in $\boldsymbol{X}$.

IV. Recovery Algorithm: Least Squares

In this section, we will layout our strategy to recover the unknown matrix $\boldsymbol{H}$ from $\boldsymbol{Y}_1$, and $\boldsymbol{Y}_2$. Assume that $\Delta,\Omega > R$. Since $\boldsymbol{H}$ is compressible, rank-R, so rank of matrices $\boldsymbol{Y}_1$, and $\boldsymbol{Y}_2$ can also be greater than R. We compute the best rank-R approximations of $\boldsymbol{Y}_1$, and $\boldsymbol{Y}_2$ using the SVDs as follows

$$\begin{aligned}\boldsymbol{Y}_1 &\approx \boldsymbol{U}_1\boldsymbol{\Sigma}_1\boldsymbol{V}_1^*\\ \boldsymbol{Y}_2 &\approx \boldsymbol{U}_2\boldsymbol{\Sigma}_2\boldsymbol{V}_2^*,\end{aligned} \quad (6)$$

where $\boldsymbol{U}_1 : \Delta\times R$, $\boldsymbol{\Sigma}_1 : R\times R$, $\boldsymbol{V}_1 : W\times R$, $\boldsymbol{U}_2 : M\times R$, $\boldsymbol{\Sigma}_2 : R\times R$, and $\boldsymbol{V}_2 : \Omega\times R$. We will take $\boldsymbol{U}_1$, and $\boldsymbol{V}_2$ to be the estimates of the row, and column space of $\boldsymbol{H}$. Given this information, we can immediately construct the rank-R estimate of $\boldsymbol{H}$ to within an unknown $R\times R$ matrix $\boldsymbol{Q}$. We propose finding $\boldsymbol{Q}$ by just fitting the proposed solution $\boldsymbol{U}_1\boldsymbol{Q}\boldsymbol{V}_2^*$ to the measurements $\boldsymbol{Y}_1$, and $\boldsymbol{Y}_2$ using a quadratic loss function below

$$\begin{aligned}\hat{\boldsymbol{Q}} &= \underset{\boldsymbol{Q}}{\operatorname{argmin}} \;\; \|\boldsymbol{U}_1\boldsymbol{Q}\boldsymbol{V}_2^*\boldsymbol{\Phi}_1^* - \boldsymbol{Y}_1\|_{\mathrm{F}}^2 + \|\boldsymbol{\Phi}_2\boldsymbol{U}_1\boldsymbol{Q}\boldsymbol{V}_2^* - \boldsymbol{Y}_2\|_{\mathrm{F}}^2\\ &= \underset{\boldsymbol{Q}}{\operatorname{argmin}} \|\boldsymbol{Q}\boldsymbol{V}_2^*\boldsymbol{\Phi}_1^* - \boldsymbol{U}_1^*\boldsymbol{Y}_1\|_{\mathrm{F}}^2 + \|\boldsymbol{\Phi}_2\boldsymbol{U}_1\boldsymbol{Q} - \boldsymbol{Y}_2\boldsymbol{V}_2\|_{\mathrm{F}}^2\end{aligned} \quad (7)$$

which is a matrix least square program. The normal equations for the least-squares program above are

$$\hat{\boldsymbol{Q}}\boldsymbol{V}_2^*\boldsymbol{\Phi}_1^*\boldsymbol{\Phi}_1\boldsymbol{V}_2 + \boldsymbol{U}_1^*\boldsymbol{\Phi}_2^*\boldsymbol{\Phi}_2\boldsymbol{U}_1\hat{\boldsymbol{Q}} = \boldsymbol{U}_1^*\boldsymbol{Y}_1\boldsymbol{\Phi}_1\boldsymbol{V}_2 + \boldsymbol{U}_1^*\boldsymbol{\Phi}_2^*\boldsymbol{Y}_2\boldsymbol{V}_2,$$

or $\hat{\boldsymbol{Q}}\boldsymbol{G}_1 + \boldsymbol{G}_2\hat{\boldsymbol{Q}} = \boldsymbol{S}$, where $\boldsymbol{G}_1 = \boldsymbol{V}_2^*\boldsymbol{\Phi}_1^*\boldsymbol{\Phi}_1\boldsymbol{V}_2$, $\boldsymbol{G}_2 = \boldsymbol{U}_1^*\boldsymbol{\Phi}_2^*\boldsymbol{\Phi}_2\boldsymbol{U}_1$, and $\boldsymbol{S} = \boldsymbol{U}_1^*\boldsymbol{Y}_1\boldsymbol{\Phi}_1\boldsymbol{V}_2 + \boldsymbol{U}_1^*\boldsymbol{\Phi}_2^*\boldsymbol{Y}_2\boldsymbol{V}_2$. We can write this $R^2\times R^2$ system of equations in vectorized form (columns stacked on one another) as $\boldsymbol{K}\hat{\boldsymbol{q}} = \boldsymbol{s}$, where $\hat{\boldsymbol{q}} = \operatorname{vec}(\hat{\boldsymbol{Q}})$, $\boldsymbol{s} = \operatorname{vec}(\boldsymbol{S})$, and $\boldsymbol{K} = \boldsymbol{I}\otimes\boldsymbol{G}_1 + \boldsymbol{G}_2\otimes\boldsymbol{I}$. Explicitly,

$$\boldsymbol{K} = \begin{bmatrix}\boldsymbol{G}_2 & & \\ & \ddots & \\ & & \boldsymbol{G}_2\end{bmatrix} + \begin{bmatrix}G_1[1,1]\boldsymbol{I} & \ldots & G_1[R,1]\boldsymbol{I}\\ \vdots & \ddots & \vdots\\ G_1[1,R]\boldsymbol{I} & \ldots & G_1[R,R]\boldsymbol{I}\end{bmatrix}.$$

So we take $\hat{\boldsymbol{q}} = \boldsymbol{K}^{-1}\boldsymbol{s}$, and unstack the columns to get $\hat{\boldsymbol{Q}}$. Given $\hat{\boldsymbol{Q}}$, we take our estimate $\hat{\boldsymbol{H}}$ of $\boldsymbol{H}$ to be

$$\hat{\boldsymbol{H}} = \boldsymbol{U}_1\hat{\boldsymbol{Q}}\boldsymbol{V}_2^*. \quad (8)$$

Theorem 1 below will show that the estimate $\hat{\boldsymbol{H}}$ exactly equals a rank-R matrix $\boldsymbol{H}$ for a large enough sampling rate. Moreover, in the case when $\boldsymbol{H}$ is not exactly rank-R the estimate deteriorates linearly with $\|\boldsymbol{H}-\boldsymbol{H}_R\|_{\mathrm{F}}$, where $\boldsymbol{H}_R$ is the best rank-R approximation of $\boldsymbol{H}$.

V. Incoherence

Before stating the sampling theorem, we also need to introduce a coherence parameter μ_0^2 that arises in our theoretical results. Let $\boldsymbol{H}_R = \boldsymbol{U}_R\boldsymbol{\Sigma}_R\boldsymbol{V}_R^*$ be the best rank-R approximation of $\boldsymbol{H}$. The coherence is defined as

$$\mu_0^2 := \frac{W}{R}\|\boldsymbol{V}_R\|_{2\to\infty}^2 = \frac{W}{R}\left(\max_k \|\boldsymbol{V}_R^{(k)}\|_2^2\right), \qquad (9)$$

where $\boldsymbol{V}_R^{(k)}$ is the kth row of $\boldsymbol{V}_R$. It can easily be checked that $1 \leq \mu_0^2 \leq W/R$. Matrix $\boldsymbol{V}_R$ spans the rows of $\boldsymbol{H}$, thus the coherence parameter controls the peak value of the rows of $\boldsymbol{H}$. Our main result in Theorem 1 shows that the sampling rate scales linearly with μ_0^2. Thus it is better to have a small coherence. Which means that in the ideal case, we want the rows of $\boldsymbol{H}$ to have more or less even energy distribution. In continuous time, this dictates that the signals after going through the low-pass filter should be sufficiently diverse.

VI. Sampling Theorem

We are now ready to state our main result that dictates how well the estimate $\hat{\boldsymbol{H}}$ in (8) agrees with the true solution $\boldsymbol{H}$. We will denote by $\boldsymbol{H}_R$, the best rank-R approximation of $\boldsymbol{H}$.

Theorem 1. *[12] Fix $\beta \geq 1$. Given the compressive samples $\boldsymbol{Y}_1$, and $\boldsymbol{Y}_2$ in (4), and (5) of the correlated ensemble $\boldsymbol{X}_c(t)$ in $t \in [0,1)$. The least squares estimate $\hat{\boldsymbol{H}}$ of $\boldsymbol{H}$ obeys*

$$\|\hat{\boldsymbol{H}} - \boldsymbol{H}\|_{\mathrm{F}}^2 \leq c\left(1 + \frac{M}{\Delta} + \frac{W}{\Omega}\right)\|\boldsymbol{H} - \boldsymbol{H}_R\|_{\mathrm{F}}^2$$

with probability at least $1 - \mathcal{O}(W^{-\beta})$, when[2]

$$\Delta \geq cR, \quad \Omega \geq c\beta\mu_0^2 R\log^2 W,$$

and hence the sampling rate obeys

$$\Delta W + M\Omega \gtrsim cRW + c\beta\mu_0^2 RM\log^2 W.$$

The result states that an incoherent ensemble of M bandlimited (to frequencies less than $W/2$), and correlated signals lying in an a priori unknown R dimensional subspace such that $W > M \gg R$ can be acquired by sampling only at a rate $\sim RW\log^2 W$ samples per second. This could be a huge reduction in the sampling rate if we compare it with the rate of MW dictated by the Nyquist sampling theorem.

VII. Proof Sketch of Theorem 1

We now give a sketch of the proof of Theorem 1; the detailed proof is deferred to a manuscript under preparation.

Proof. Let $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^*$ denote the singular value decomposition of the matrix $\boldsymbol{H}$. Let

$$\boldsymbol{U} = [\boldsymbol{U}_R\ \tilde{\boldsymbol{U}}], \quad \text{and} \quad \boldsymbol{V} = [\boldsymbol{V}_R\ \tilde{\boldsymbol{V}}] \qquad (10)$$

where $\boldsymbol{U}_R$ be the matrix of first R columns of $\boldsymbol{U}$, and $\tilde{\boldsymbol{U}}$ be the remaining columns of $\boldsymbol{U}$. and similarly define $\boldsymbol{V}_R$, and $\tilde{\boldsymbol{V}}$.

Bounds on the singular values

[2]The symbol c refers to a numerical constant that may refer to a different number every time it is used.

Recall that $\boldsymbol{\Phi}_1$, and $\boldsymbol{\Phi}_2$ be random matrices as defined earlier. The first technical challenge is to obtain a lower bound on the smallest singular values of $\boldsymbol{\Phi}_2\boldsymbol{U}_R$, and $\boldsymbol{\Phi}_1\boldsymbol{V}_R$, and an upper bound on the highest singular values of $\boldsymbol{\Phi}_2\tilde{\boldsymbol{U}}$, and $\boldsymbol{\Phi}_1\tilde{\boldsymbol{V}}$.

Start by noting that the $\Delta \times R$ matrix $\boldsymbol{\Phi}_2\boldsymbol{U}_R$ is also a standard Gaussian matrix. Computing bounds on the singular values of a Gaussian matrix is fairly standard; see, for example [13] to find that

$$\sqrt{\Delta} - \sqrt{R} \approx \sigma_{\min}(\boldsymbol{\Phi}_2\boldsymbol{U}_R) \leq \sigma_{\max}(\boldsymbol{\Phi}_2\boldsymbol{U}_R) \approx \sqrt{\Delta} + \sqrt{R}$$

with probability at least $1 - \mathrm{e}^{-\Delta}$. Choosing $\Delta \geq cR$ for a sufficiently large constant c ensures that $\sigma_{\min}(\boldsymbol{\Phi}_2\boldsymbol{U}_R) \approx \sqrt{\Delta/2}$, and $\sigma_{\min}(\boldsymbol{\Phi}_2\boldsymbol{U}_R) \approx \sqrt{2\Delta}$, and $\sigma_{\min}(\boldsymbol{\Phi}_2\boldsymbol{U}_R) \approx \sqrt{2\Delta}$. In a similar manner, $\sigma_{\max}(\boldsymbol{\Phi}_2\tilde{\boldsymbol{U}}) \approx \sqrt{M}$ with probability at least $1 - \mathrm{e}^{-M}$.

As far as the structured random matrix $\boldsymbol{\Phi}_1\boldsymbol{V}_R$ goes, an application of matrix Bernstein inequality shows that $\sigma_{\min}(\boldsymbol{\Phi}_1\boldsymbol{V}_R) \geq (\sqrt{2})^{-1}$, and $\sigma_{\max}(\boldsymbol{\Phi}_1\boldsymbol{V}_R) \leq \sqrt{3/2}$ with probability at least $1 - \mathcal{O}(W^{-\beta})$, whenever $\Omega \geq c\beta\mu_0^2 R\log^2 W$ for some fixed $\beta \geq 1$. Moreover, the inequality $\sigma_{\max}(\boldsymbol{\Phi}_1\tilde{\boldsymbol{V}}) \leq c\sqrt{W/\Omega}$ holds deterministically.

Recall that $\boldsymbol{U}_1, \boldsymbol{U}_2, \boldsymbol{V}_1$, and $\boldsymbol{V}_2$ come from the SVD of row, and column space measurements in (6). We now state two inequalities that are obtained directly from Theorem 9.1 in [10].

$$\|\boldsymbol{H} - \boldsymbol{U}_1\boldsymbol{U}_1^*\boldsymbol{H}\|_{\mathrm{F}}^2 \leq \|\boldsymbol{H} - \boldsymbol{H}_R\|_{\mathrm{F}}^2\left(1 + \frac{\sigma_{\max}^2(\boldsymbol{\Phi}_1\tilde{\boldsymbol{V}})}{\sigma_{\min}^2(\boldsymbol{\Phi}_1\boldsymbol{V}_R)}\right), \qquad (11)$$

and similarly,

$$\|\boldsymbol{H} - \boldsymbol{H}\boldsymbol{V}_2\boldsymbol{V}_2^*\|_{\mathrm{F}}^2 \leq \|\boldsymbol{H} - \boldsymbol{H}_R\|_{\mathrm{F}}^2\left(1 + \frac{\sigma_{\max}^2(\boldsymbol{\Phi}_2\tilde{\boldsymbol{U}})}{\sigma_{\min}^2(\boldsymbol{\Phi}_2\boldsymbol{U}_R)}\right). \qquad (12)$$

Lemma 1. *Suppose that the radom matrices $\boldsymbol{\Phi}_1$, and $\boldsymbol{\Phi}_2$ preserve geometry, i.e., $\sigma_{\min}(\boldsymbol{\Phi}_2\boldsymbol{U}_R) \geq \sqrt{2/\Delta}$, $\sigma_{\max}(\boldsymbol{\Phi}_2\boldsymbol{U}_R) \leq \sqrt{2\Delta}$; and $\sigma_{\min}(\boldsymbol{\Phi}_1\boldsymbol{V}_R) \geq 1/\sqrt{2}$, $\sigma_{\max}(\boldsymbol{\Phi}_1\boldsymbol{V}_R) \leq \sqrt{3/2}$. The solution $\hat{\boldsymbol{H}}$ in (8) of the least squares program (7) obeys*

$$\|\boldsymbol{H} - \hat{\boldsymbol{H}}\|_{\mathrm{F}}^2 \leq c\left(\|\boldsymbol{H} - \boldsymbol{U}_1\boldsymbol{U}_1^*\boldsymbol{H}\|_{\mathrm{F}}^2 + \|\boldsymbol{H} - \boldsymbol{H}\boldsymbol{V}_2\boldsymbol{V}_2^*\|_{\mathrm{F}}^2\right)$$

for an absolute constant c.

Proof. Proof of this theorem follows the template of the proof in [14], which establishes a similar result for a simpler least squares program compared to (7). □

Combining this result of this lemma with (11), and (12), and plugging in the bounds for the singular values □

VIII. Related Work and Our Contribution

An in depth study of the compressive sampling schemes of ensembles of correlated signals first started appearing in [1]–[5]. Multiple implementable architectures are presented along with sampling theorems that dictate the sampling rate sufficient for exact recovery in each architecture. However, one common theme among all this work is that the signal reconstruction problem involves solving a semidefinite program, which is prohibitively computationally expensive. In comparison, we propose a sampling architecture that allows us to use the computationally much less expensive least-squares approach for signal reconstruction.

Sub-Nyquist sampling strategies for sparse signals using the ideas of compressed sensing have been pursued; see, for example, [15]–[17] in the last decade after the advent of compressed sensing. The work presented here differs from this existing body of work fundamentally as the underlying structure in not sparsity but correlation (signals reside in an a priori unknown subspace). The work presented in this manuscript is more general as in sparse signal reconstruction one needs to know the sparsifying basis in advance and then a subset of the basis functions that contribute in signal are discovered. Comparatively, in this paper, we do not have any knowledge about the subspace in which the signals reside and it is completely discovered implicitly during the reconstruction process.

Theorem 1 is basically a matrix reconstruction result from the row and column space measurements of the matrix. A complete body of work; see, for example, [10], [11] and references therein, is available that shows how one can employ least squares program to recover a matrix from its row, and column space measurements. The main difference of our theoretical result from these previous results is that we are working with a very structured random matrix $\mathbf{\Phi}_1$ with a very limited randomness, and the results in the literature are only available for random sampling or at most Gaussian random projections. To derive all these results for $\mathbf{\Phi}_1$ required more sophisticated tools from the theory of random matrices. A more detailed version of this paper has already appeared in [12].

IX. Acknowledgements

The author would like to acknowledge Justin Romberg for several indepth discussions on this topic.

References

[1] A. Ahmed and J. Romberg, "Compressive multiplexing of correlated signals," *IEEE Trans. Inform. Theory*, vol. 1, pp. 479–498, 2015.

[2] ——, "Compressive sampling of ensembles of correlated signals," *arXiv preprint arXiv:1501.06654*, 2015.

[3] ——, "Compressive multiplexers for correlated signals," in *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 2012, pp. 963–967.

[4] ——, "Compressive sampling of correlated signals," in *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 2011, pp. 1188–1192.

[5] ——, "Compressive sampling in array processing," in *2013 IEEE 5th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*. IEEE, 2013, pp. 192–195.

[6] U. Frey, C. Sanchez-Bustamante, T. Ugniwenko, F. Heer, J. Sedivy, S. Hafizovic, B. Roscic, M. Fussenegger, A. Blau, U. Egert *et al.*, "Cell recordings with a CMOS high-density microelectrode array," in *Conf. Proc. IEEE Eng. Med. Biol. Soc. (EMBS)*, 2007, pp. 167–170.

[7] K. Imfeld, S. Neukom, A. Maccione, Y. Bornat, S. Martinoia, P. Farine, M. Koudelka-Hep, and L. Berdondini, "Large-scale, high-resolution data acquisition system for extracellular recording of electrophysiological activity," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 8, pp. 2064–2073, 2008.

[8] A. Haas, "Programmable high density cmos microelectrode array," in *Proc. IEEE Conf. Sensors*, Lecce, Italy, 2008, pp. 890–893.

[9] D. Gray, J. Tan, J. Voldman, and C. Chen, "Dielectrophoretic registration of living cells to a microelectrode array," *J. Biosens. and Bioelectron.*, vol. 19, no. 7, pp. 771–780, 2004.

[10] N. Halko, P.-G. Martinsson, and J. A. Tropp, "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions," *SIAM review*, vol. 53, no. 2, pp. 217–288, 2011.

[11] E. Liberty, F. Woolfe, P.-G. Martinsson, V. Rokhlin, and M. Tygert, "Randomized algorithms for the low-rank approximation of matrices," *Proceedings of the National Academy of Sciences*, vol. 104, no. 51, pp. 20 167–20 172, 2007.

[12] A. Ahmed, "Compressive acquisition and least-squares reconstruction of correlated signals," *IEEE Signal Processing Letters*, vol. 24, no. 7, pp. 933–937, 2017.

[13] R. Vershynin, *Compressed sensing: theory and applications*, Y. C. Eldar and G. Kutyniok, Eds. Cambridge University Press, 2012.

[14] F. Woolfe, E. Liberty, V. Rokhlin, and M. Tygert, "A fast randomized algorithm for the approximation of matrices," *Applied and Computational Harmonic Analysis*, vol. 25, no. 3, pp. 335–366, 2008.

[15] J. Tropp and J. Laska and M. Duarte and J. Romberg, and R. Baraniuk, "Beyond nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Trans. Inform. Theory*, vol. 56, no. 1, pp. 520–544, 2010.

[16] J. Laska and S. Kirilos and M. Duarte and T. Raghed and R. Baraniuk and Y. Massoud, "Theory and implementation of an analog-to-information converter using random demodulation," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2007, pp. 1959–1962.

[17] M. Mishali and Y. Eldar and O. Dounaevsky and E. Shoshan, "Xampling: Analog to digital at sub-Nyquist rates," *IET Circuits Devices Syst.*, vol. 5, no. 1, pp. 8–20, 2011.