MAXIMUM ENTROPY HADAMARD SENSING OF SPARSE AND LOCALIZED SIGNALS

Valerio Cambareri [1,2], *Riccardo Rovatti* [1,2], *Gianluca Setti* [2,3]

[1] Department of Electrical, Electronic and Information Engineering, University of Bologna, Italy
[2] Advanced Research Center on Electronic Systems, University of Bologna, Italy
[3] Engineering Department, University of Ferrara, Italy

ABSTRACT

The quest for optimal sensing matrices is crucial in the design of efficient Compressed Sensing architectures. In this paper we propose a maximum entropy criterion for the design of optimal Hadamard sensing matrices (and similar deterministic ensembles) when the signal being acquired is sparse and non-white. Since the resulting design strategy entails a combinatorial step, we devise a fast evolutionary algorithm to find sensing matrices that yield high-entropy measurements. Experimental results exploiting this strategy show quality gains when performing the recovery of optimally sensed small images and electrocardiographic signals.

Index Terms— Compressed Sensing, Walsh-Hadamard Transform, Maximum Entropy Principle, Sensing Matrix Design, Evolutionary Heuristics

1. INTRODUCTION

Compressed Sensing (CS) [1, 2] is a recent sampling paradigm applicable to signals whose intrinsic dimensionality is considerably smaller than suggested by their Nyquist rate. In a nutshell, the compressed acquisition process amounts to mapping the signal being acquired into an undersampled set of measurements by applying a suitable sensing matrix. Thus, CS allows one to balance the effort and resources needed to acquire a signal to its intrinsic dimensionality rather than its apparent one. This idea has fostered a growing literature pursuing the implementation of CS principles in actual sampling and imaging systems [3, 4].

In doing so, it is natural to expect that the generality of the theoretical approach may be dropped to meet the constraints of feasible hardware implementations. An exemplary simplification in this sense is the assumption that the sensing matrix applied by the acquisition process is antipodal-valued, i.e., it lies in $\{-1, +1\}^{m \times n}$ with $m < n$.

In this paper we investigate CS with antipodal-valued Hadamard sensing matrices (i.e., *Hadamard sensing*) when the signals being acquired are (a) sparse in a proper basis that is not necessarily incoherent [5] w.r.t. Hadamard sensing vectors and (b) localized [6], i.e., the random vector (RV) representing the signal has non-white covariance matrix. This theoretical background is summarized in Section 2.

Leveraging on these assumptions, in Section 3 we formulate a sensing matrix design criterion motivated by the maximum entropy principle [7] with the aim of selecting the optimal set of m measurements based on the analysis of their covariance. Since the exact solution of the resulting selection problem is NP-hard due to the nature of the maximum entropy sampling problem [8] we introduce a lightweight evolutionary algorithm to generate a pool of candidate sensing matrices yielding m near-maximum entropy measurements.

The criterion and its heuristic implementation allow us to devise a strategy for optimal Hadamard sensing, which is applied to some examples in Section 4. There, the results when applying our method to CS of small images and of electrocardiographic tracks (ECGs) show clear improvements in terms of signal recovery performances despite the non-minimum coherence between the sensing matrix and the chosen sparsity bases, thus partially overcoming the non-universality of the Hadamard matrix ensemble.[1]

1.1. Prior Work

The design of optimal sensing matrices is commonly tackled in absence of physical constraints [9, 10]. When the sensing vectors are fixed by the acquisition process this problem is known as variable density sampling [11, 12] although the *a priori* assumptions and objectives are radically different from our own. Our second-order statistics assumption is a very mild prior based on previous research on *rakeness* [6, 13, 14]. It is also worth noting that, independently, the works of Carson *et al.* [15] and Chen *et al.* [16, 17] address the problem of optimal sensing matrix design by leveraging on similar principles (power-constrained mutual information maximization in the presence of noise) but without the constraint of choosing the sensing matrix from a deterministic ensemble.

2. THEORETICAL BACKGROUND

2.1. Compressed Sensing with Deterministic Ensembles

The most common formulation of CS [1, 2] is in the discrete domain, where the signal being acquired is represented by a set of n samples collected in a vector $\mathbf{x} \in \mathbb{R}^n$, while the sampling operation is performed by a dimensionality-reducing *sensing matrix* $\bar{\mathbf{A}}_{m\times n}$ $(m < n)$ which produces the *measurement vector* $\bar{\mathbf{y}} = \bar{\mathbf{A}}\mathbf{x} \in \mathbb{R}^m$. Since $m < n$ the recovery

[1] The code to reproduce them is available at https://sites.google.com/site/ssigprocs/CS/maxenths

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

of $\mathbf{x}$ from $\bar{\mathbf{y}}$ is an ill-posed problem, whose solution is possible if there is a *sparsity basis* or *dictionary* $\mathbf{D}_{n\times n_d}, n_d \geq n$ such that every signal being acquired can be expressed as $\mathbf{x} = \mathbf{D}\mathbf{s}, \mathbf{s} \in \mathbb{R}^{n_d}$ where $\mathbf{s}$ is (at least approximately) k-*sparse*, i.e., the cardinality $\|\mathbf{s}\|_0 = |\operatorname{supp}(\mathbf{s})| = k$.

The original $\mathbf{s}$ is then recovered from the undersampled measurements $\bar{\mathbf{y}} = \bar{\mathbf{W}}\mathbf{s}$, $\bar{\mathbf{W}} = \bar{\mathbf{A}}\mathbf{D}$ by solving the combinatorial problem $\mathbf{s} = \arg\min_{\mathbf{z}\in\mathbb{R}^{n_d}} \|\mathbf{z}\|_0$ s.t. $\bar{\mathbf{y}} = \bar{\mathbf{W}}\mathbf{z}$ ($\min \ell_0$) which has been relaxed to a vast array of greedy decoding algorithms [18, Chap. 8]. The recovery quality achieved by these algorithms depends on $\bar{\mathbf{A}}$, which is safely designed by following celebrated theoretical guarantees [19, 20] relating m, n, k and the structure of $\bar{\mathbf{W}}$.

Using these guarantees two suitable hardware-friendly, antipodal-valued random matrix ensembles emerge: the *random Bernoulli ensemble* (RBE) and the *partial Hadamard ensemble* (PHE). The RBE is comprised of matrices $\bar{\mathbf{A}} \in \{-1, +1\}^{m\times n}$ having i.i.d. equiprobable antipodal entries. On the other hand, $\bar{\mathbf{A}}$ belonging to the PHE are constructed by choosing random subsets of rows from the Hadamard matrix of order n, $\mathbf{H}_n, n = 2^q, q \in \mathbb{N}$. While the RBE is universal [19] regardless of $\mathbf{D}$, the PHE is considered suitable for CS only when $\mathbf{D} = \mathbf{I}_n$ is the identity.

Deterministic ensembles such as the PHE become of practical interest when the choice of sensing vectors is fixed or constrained by the acquisition mechanism. Thus, we assume to be limited to an orthonormal *design space* of feasible sensing vectors $\{\mathbf{a}_j\}_{j=0}^{n-1}, \mathbf{a}_j \in \mathbb{R}^n$ collected in the rows of $\mathbf{A}_{n\times n}$.[2] The sensing matrix, which we now denote $\mathbf{A}_T$, is constructed by extracting m row vectors from $\mathbf{A}$ with indexes in the subset $T = \{j_0, \ldots, j_{m-1}\}, |T| = m$. In absence of other assumptions, the $\binom{n}{m}$ possible $\mathbf{A}_T$ are considered equally good candidates in the corresponding matrix ensemble.

$\mathbf{A}$ interacts with the columns of $\mathbf{D}$ as $\mathbf{W}_{n\times n_d} = \mathbf{A}\mathbf{D}$, their relationship being commonly quantified by *coherence* [5], i.e., by $\mu(\mathbf{A}, \mathbf{D}) = \max_{j,k} \frac{|\mathbf{a}_j^\dagger \mathbf{d}_k|}{\|\mathbf{a}_j\|_2 \|\mathbf{d}_k\|_2} \in [n^{-\frac{1}{2}}, 1]$. Although coherence should be as small as possible, it is suggested [21] that it only needs to be bounded in many cases. Since the design space is fixed by the properties of the acquisition mechanism and the sparsity basis is set by the nature of the signal being sampled, this will often be the case in practical CS applications.

In this paper we explore the Hadamard sensing case, where $\mathbf{A} = \frac{1}{\sqrt{n}}\mathbf{H}_n$ corresponds to the normalized Walsh-Hadamard transform [22]. This transform is particularly suitable to both analog, optical and digital implementations. Due to its recursive structure it can also be fully computed by divide and conquer in $O(n \log_2 n)$ instead of $mn = O(nk \log_2 n)$ additions/subtractions.

This appealing low-complexity property drives the idea of finding optimal Hadamard sensing matrices $\mathbf{A}_T$ for generic dictionaries $\mathbf{D}$ (here assumed to be orthonormal bases for the sake of simplicity), noting that $\mathbf{A}_T$ will exhibit non-minimum coherence with respect to $\mathbf{D} \neq \mathbf{I}_n$.

[2] We will denote by $\cdot_T$ the selection of m-cardinality row subsets T in a matrix/vector and by $\cdot^\star$, $\hat{\cdot}$ optimal and approximate values.

2.2. Localized Signals and Correlated Measurements

Standard CS makes no assumption on the statistical properties of $\mathbf{s}$, thus conforming to a white, worst-case scenario which does not harness the probabilistic structure in sparse/compressible representations of natural signals [23].

With the aim of exploiting such structure, in the following we consider a simple second-order description of the RV $\mathbf{s}$ by its mean $\mu_\mathbf{s}$ and non-white covariance matrix $\mathbf{K}_\mathbf{s} = \mathbb{E}[(\mathbf{s} - \mu_\mathbf{s})(\mathbf{s} - \mu_\mathbf{s})^\dagger]$. Such properties are straightforwardly estimated from available realizations or valid recoveries of $\mathbf{s}$. In this context the authors in [6] have introduced *localization* to quantify the non-whiteness of $\mathbf{s}$ as a deviation in the eigenvalues of $\mathbf{K}_\mathbf{s}$ from the white case with the same energy.

Given a RV $\mathbf{s}$ with covariance $\mathbf{K}_\mathbf{s}$ we acquire it by applying $\mathbf{W}_T = \mathbf{A}_T\mathbf{D}$ which yields the RV of measurements $\mathbf{y}_T = \mathbf{W}_T\mathbf{s}$ corresponding to subset T. Its covariance $\mathbf{K}_{\mathbf{y}_T} = \mathbf{W}_T\mathbf{K}_\mathbf{s}\mathbf{W}_T^\dagger$ will be non-white in general. Thus, localized RVs $\mathbf{s}$ generally imply localized measurements.

As an example, let $\mathbf{y}_T$ follow a non-white multivariate Gaussian distribution $\mathcal{N}(0, \mathbf{K}_{\mathbf{y}_T})$. Then its localization would directly indicate that the acquisition process represented by $\mathbf{W}_T$ is not maximizing the quantity of information embedded in the measurements, since by considering their differential entropy $h(\mathbf{y}_T)$ we have [24, Theorem 9.4.1]

$$h(\mathbf{y}_T) = \tfrac{1}{2}\log(2\pi e)^m \det \mathbf{K}_{\mathbf{y}_T} \leq \tfrac{1}{2}\log\left(2\pi e \frac{\mathcal{E}_{\mathbf{y}_T}}{m}\right)^m \quad (1)$$

which by [24, Theorem 16.8.4] attains the upper bound in (1) when $\mathbf{y}_T$ is white for a fixed energy $\mathcal{E}_{\mathbf{y}_T} = \operatorname{tr}\mathbf{K}_{\mathbf{y}_T}$.

3. AN OPTIMAL DESIGN CRITERION FOR HADAMARD SENSING MATRICES

The previous example indicates that the measurements $\mathbf{y}_T$ will not achieve the white-case entropy upper bound in (1) when the original signal is localized and $\mathbf{y}_T$ are obtained from sensing matrices $\mathbf{A}_T$ in the PHE (or similar deterministic ensembles). In such a constrained setting and inspired by the classic maximum entropy principle [7] we aim at finding the optimal $\mathbf{A}_T$ in the design space which conveys the maximum achievable quantity of information in the measurements $\mathbf{y}_T$.

3.1. A Maximum Entropy Problem

In general, we are searching for the m-cardinality subset $\mathbf{y}_{T^\star}$ of the full measurements' RV $\mathbf{y} = \mathbf{W}\mathbf{s}$ which attains the maximum differential entropy $h(\mathbf{y}_{T^\star})$, i.e., we solve

$$T^\star = \underset{T\subset[0,n-1]}{\arg\max}\ h(\mathbf{y}_T) \text{ s.t. } |T| = m \quad (2)$$

In a Gaussian context, let $\mathbf{y} \sim \mathcal{N}(0, \mathbf{K}_\mathbf{y})$ with $\mathbf{K}_\mathbf{y} = \mathbf{W}\mathbf{K}_\mathbf{s}\mathbf{W}^\dagger$. Then $h(\mathbf{y}_T)$ is given in (1) where $\mathbf{K}_{\mathbf{y}_T}$ is a principal minor of $\mathbf{K}_\mathbf{y}$ corresponding to subset T. Thus (2) amounts to solving

$$T^\star = \underset{T\subset[0,n-1]}{\arg\max}\ \log\det \mathbf{K}_{\mathbf{y}_T} \text{ s.t. } |T| = m \quad (3)$$

More realistically $\mathbf{y}$ will only be approximately Gaussian and substantially depend on the distribution of $\mathbf{s}$, thus requiring the more general solution of (2).

In a distribution-agnostic fashion we choose to solve problem (3) in its place: while the solution $T^\star$ might not achieve globally maximum entropy, the corrisponding $\mathbf{y}_{T^\star}$ will be the subset of measurements having least linear predictability (or equivalently maximum prediction error) from one another [25, Sec. 2.4.3 and 6.6].

Note that in Section 4 we report reassuring evidence that natural signals considered in our experiments produce approximately Gaussian $\mathbf{y}$, thus suggesting that (3) is well-posed for finding Hadamard sensing vectors that maximize the measurements' entropy.

3.2. Heuristic Solution to Optimal Sensing Matrix Design

On the computational side solving (3) amounts to finding the maximum determinant principal minor $\mathbf{K}_{\mathbf{y}_{T^\star}}$. When $\mathbf{K}_\mathbf{y}$ is diagonal this is straightforwardly solved by $T^\star = \arg\max_{T\subset[0,n-1]} \operatorname{tr} \mathbf{K}_{\mathbf{y}_T}$ s. t. $|T| = m$ (choosing the m largest-variance components of $\mathbf{y}$).

When this is not verified (3) is a well-known combinatorial problem. In [8] Ko *et al.* prove its hardness and propose an exact branch-and-bound algorithm. Since in natural signals fluctuations in $\mathbf{K}_\mathbf{s}$ will occur and eventually require an update in $T^\star$ we propose to use a lightweight genetic algorithm [26] to find a heuristic solution.

The evolutionary analogue is obtained by mapping the i-th subset $T_{(i)}, |T_{(i)}| = m$ into a length n, binary-valued chromosome $\tau_{(i)} = I_{T_{(i)}}$ (the indicator function of $T_{(i)}$), whose fitness function is simply $f_{(i)} = \log\det \mathbf{K}_{\mathbf{y}_{T_{(i)}}}$. Since covariance matrices are Hermitian non-negative definite, by using the Cholesky factorization $\log\det \mathbf{K}_{\mathbf{y}_{T_{(i)}}} = 2\log\det \mathbf{L}_{T_{(i)}}$ where $\mathbf{L}_{T_{(i)}}$ is lower triangular. This allows fast and accurate computation of the i-th fitness $f_{(i)} = 2\operatorname{tr}\log\operatorname{diag}\mathbf{L}_{T_{(i)}}$.

The algorithm is implemented as in Proc. 1 for a generic covariance matrix $\mathbf{K}$ and controlled by the global parameters N_{gen} (number of generations), N_{pop} (population size at each generation). We note the use of a warm start by including in the initial population $\mathbf{\Omega}^{(0)}$ the element $\tau_{(0)} = I_{T_{(0)}}$ initialized to the indexes of the m largest variances in $\mathbf{K}$.

At each generation, mating occurs between the $N_{\text{par}} = \frac{N_{\text{pop}}}{3}$ highest-fitness parent chromosomes such that their $N_{\text{chi}} = \frac{2N_{\text{pop}}}{3}$ children are m-cardinality subsets. An elitist policy grants survival to the parent chromosomes until they are replaced by fitter children. To avoid possible stagnation in the population we have introduced common genetic operators such as one-point random crossover, random mutation with probability P_{mut} and a final unicity check to avoid clones in the population.

Thus, the algorithm yields a near-optimal $T^\star$ depending on the chosen parameters, which we use to construct $\mathbf{W}_{T^\star} = \mathbf{A}_{T^\star}\mathbf{D}$. Setting large $(N_{\text{gen}}, N_{\text{pop}})$ increases the complexity of this procedure but typically leads to a larger fitness gap between any random index subset T and the final $T^\star$. Moreover, rather than using a single $T^\star$ one may choose T from the high-entropy final population $\mathbf{\Omega}^{(N_{\text{gen}}-1)}$, which we refer to as the MaxDet pool for the following experimental evaluation.

Procedure 1 Evolutionary Heuristic Solution of (3)

Require: $\mathbf{K}_{n\times n}$ (covariance matrix), m, N_{gen}, N_{pop}, N_{par}, N_{chi}, P_{mut}

$\tau_{(0)} = I_{T_{(0)}} \in \mathbf{\Omega}^{(0)}, T_{(0)} = \arg\max_{T\subset[0,n-1]} \operatorname{tr}\mathbf{K}_T$ s.t. $|T| = m$

for all $\tau_{(i)} \in \mathbf{\Omega}^{(0)}, i > 0$ **do** {Random initialization}

  Generate random $\tau_{(i)} = I_{T_{(i)}}, T_{(i)} \subset [0, n-1], |T_{(i)}| = m.$

end for

for $l = 0$ to $N_{\text{gen}} - 1$ **do** {Genetic search}

  for all $\tau_{(i)} \in \mathbf{\Omega}^{(l)}$ **do** {Fitness evaluation}

    Calculate the fitness $f_{(i)} = 2\operatorname{tr}\log\operatorname{diag}\mathbf{L}_{T_{(i)}}$

  end for

  Sort $\tau_{(i)} \in \mathbf{\Omega}^{(l)}$ in descending order w.r.t. their fitness $f_{(i)}$

  $\mathbf{\Omega}_{\text{par}} \leftarrow \{\tau_{(i)}\}_{i=0}^{N_{\text{par}}-1}$ {Parents selection}

  for $k = 0$ to $N_{\text{chi}} - 1$ **do** {Mating phase}

    Randomly pick $\tau_{(i)}, \tau_{(j)} \in \mathbf{\Omega}_{\text{par}}$

    $o \leftarrow$ random index in $[1, n-1]$

    $\tau_{(N_{\text{par}}+k)} \leftarrow \left[(\tau_{(i)})_{0,\dots,o-1} \quad (\tau_{(j)})_{o,\dots,n-1}\right]$ {Crossover}

    $r \leftarrow$ uniform random real in $[0, 1]$

    if $r < P_{\text{mut}}$ **then** {Mutation}

      Shuffle random $(0, 1)$ pairs in $\tau_{(N_{\text{par}}+k)}$

    end if

    if $|T_{(N_{\text{par}}+k)}| - m > 0$ **then** {Well-formed check}

      Remove $|T_{(i)}| - m$ exceeding ones in $(\tau_{(N_{\text{par}}+k)})_{o,\dots,n-1}$

    else

      Add $m - |T_{(N_{\text{par}}+k)}|$ missing ones in $(\tau_{(N_{\text{par}}+k)})_{o,\dots,n-1}$

    end if

  end for

  Eliminate duplicates and replenish the pool $\mathbf{\Omega}^{(l+1)}$ {Unicity check}

end for

return $T^\star \leftarrow \operatorname{supp}\left(\tau_{(0)}\right), \tau_{(0)} \in \mathbf{\Omega}^{(N_{\text{gen}}-1)}$

4. EXPERIMENTS

We choose to assess the near-optimality of the MaxDet pool against random PHE sensing matrices by observing the reconstruction signal-to-noise ratio, $\mathrm{RSNR}_{\mathrm{dB}} = 20\log_{10}\frac{\|\mathbf{s}\|_2}{\|\hat{\mathbf{s}}-\mathbf{s}\|_2}$ attained by basis pursuit (BP) [27] from the corresponding measurements $\mathbf{y}_T$. The signal classes tested here are natural images and ECGs taken from public-domain databases. To efficiently recover these signals by BP we used SPGL_1 [28] while identical results were obtained by linear programming in GUROBI [29].

The experiments are carried out by following the proposed Hadamard sensing design strategy as summarized in Proc. 2 with heuristic parameters $N_{\text{gen}} = 200, N_{\text{pop}} = 50, P_{\text{mut}} = 0.1$. Due to the importance of correctly inferring $\mathbf{K}_\mathbf{y}$ we note the use of the shrinkage covariance estimator $\tilde{\mathbf{K}}_\mathbf{y}$, which safely balances the sample covariance matrix (SCM) $\widehat{\mathbf{K}}_\mathbf{y}$ with the same-energy $\mathcal{E}_\mathbf{y}$ white case. This covariance estimator leads (with suitable ϵ) to a full-rank $\tilde{\mathbf{K}}_\mathbf{y}$ in the presence of additive measurement noise and from limited or linearly dependent observations.

Practical applications will also require an update in $T^\star$ to track statistically significant variations in the signal statistics. This update will be triggered whenever the recovered sparse coefficients are classified as outliers w.r.t. $\mathbf{K}_\mathbf{s}$. However, we leave the analysis of online updates to future investigations.

Procedure 2 Optimal Hadamard Sensing of Localized Signals

1: Estimate $\widehat{\mathbf{K}}_{\mathbf{y}}$ and $\mathcal{E}_{\mathbf{y}} = \mathrm{tr}\,\widehat{\mathbf{K}}_{\mathbf{y}}$ (either by direct observation or by setting $\widehat{\mathbf{K}}_{\mathbf{y}} = \mathbf{W}\widehat{\mathbf{K}}_{\mathbf{s}}\mathbf{W}^{\dagger}$ with $\widehat{\mathbf{K}}_{\mathbf{s}}$ the sparse coefficients' SCM)
2: Evaluate $\tilde{\mathbf{K}}_{\mathbf{y}} = (1-\epsilon)\widehat{\mathbf{K}}_{\mathbf{y}} + \epsilon\frac{\mathcal{E}_{\mathbf{y}}}{\sqrt{n}}\mathbf{I}_n$
3: Solve (3) by running Proc. 1 on $\tilde{\mathbf{K}}_{\mathbf{y}}$
4: Update the Hadamard sensing matrix $\mathbf{A}_{T^\star}$
5: **loop**
6: Acquire $\mathbf{y}_{T^\star} = \mathbf{A}_{T^\star}\mathbf{x}$
7: Recover $\hat{\mathbf{s}}$ by BP from $(\mathbf{A}_{T^\star}, \mathbf{y}_{T^\star})$
8: **end loop**

4.1. Handwritten digits

The first experiment is carried out on image samples from the USPS handwritten digits database [30]. We estimate $\widehat{\mathbf{K}}_{\mathbf{s}}$ on a training set of 2000 images resized to 64×64 pixel ($n = 4096$) and with $\mathbf{D}$ the two-dimensional discrete cosine basis (2D-DCT) on which on average $k = 467$ coefficients represent 95% of the original signal energy.

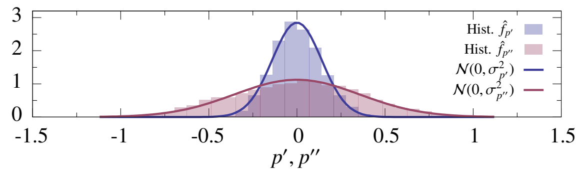
To show that the corresponding $\mathbf{y}$ can be considered Gaussian, we collect the full Hadamard transform coefficients $\mathbf{y}$ for 4000 images in the database and project them on two random orthonormal directions $p', p'' \in \mathbb{R}^n$. The resulting empirical densities of p', p'' are reported in Fig. 1a for 32 histogram bins and fit by standard normal distributions. Albeit with different variances, p', p'' are approximately Gaussian, thus suggesting that $\mathbf{y}$ is also approximately Gaussian.

By running Proc. 2 (1:-4:) with $\epsilon = 10^{-12}$ we obtain MaxDet pools $\mathbf{\Omega}^{(N_{\mathrm{gen}}-1)}$ and near-optimal solutions $T^\star$ yielding high-entropy measurements for $m = 1024, 1365$. For each of 20 test images, we simulate CS by three sets of sensing matrices: 25 PHE matrices from the MaxDet pool (including the optimal $\mathbf{A}_{T^\star}$), 25 randomly chosen PHE matrices and 50 RBE matrices. Then, signal recovery is performed by BP from these sets of measurements. The results in terms of average RSNR are reported in Fig. 1f, where measurements obtained by the MaxDet pool Hadamard sensing matrices outperform both randomly selected Hadamard and random Bernoulli sensing matrices. Fig. 1b-e illustrate this observable improvement in terms of typical RSNR performances for a sample digit in the dataset and $m = 1024$.

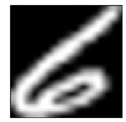
4.2. Electrocardiographic tracks

In this second experiment we illustrate Hadamard sensing of ECG tracks from the PhysioNet database [31]. $\widehat{\mathbf{K}}_{\mathbf{s}}$ is estimated on a training set of 180 ECG fragments of $n = 256$ equivalent Nyquist-rate samples and with $\mathbf{D}$ the Coiflet-3 orthonormal wavelet basis [32], on which on average $k = 39$ coefficients represent 95% of the original signal energy. Although we omit their histogram, $\mathbf{y}$ can also be considered approximately Gaussian in this case.

Proc. 2 (1:-4:) with $\epsilon = 10^{-12}$ yields MaxDet pools $\mathbf{\Omega}^{(N_{\mathrm{gen}}-1)}$ and near-optimal solutions $T^\star$ for $m = 64, 85$ (a strongly undersampled setting). Signal recovery is then performed in the same fashion of the first example for 50 sample ECGs. The results in terms of average RSNR are reported in Fig. 2d, while Fig.2a-c illustrate the typical RSNR performances for a sample ECG in the dataset and $m = 85$.

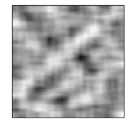


(a) Empirical PDF of p', p'' and Gaussian approximation with $\sigma^2_{p'} = 0.02, \sigma^2_{p''} = 0.13$

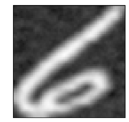


(b) Original image, $n = 4096$ pixel

(c) $T^\star$ PHE, RSNR = 36.82 dB



(d) Random PHE, RSNR = 1.88 dB

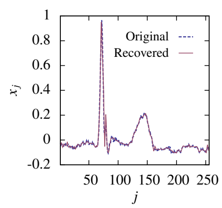


(e) RBE, RSNR = 19.85 dB

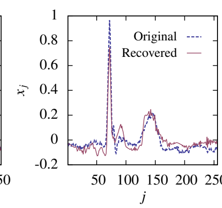
m	MaxDet pool PHE	Random PHE	RBE
1024 ($^n/_4$)	**36.57**	1.51	20.63
1365 ($\lfloor^n/_3\rfloor$)	**39.63**	2.89	26.08

(f) Average $\mathrm{RSNR}_{\mathrm{dB}}$ over 20 sample images, 25 MaxDet pool PHE, 25 Random PHE and 50 RBE sensing matrices.

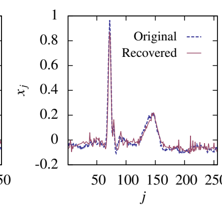
Fig. 1: Comparison of PHE and RBE sensing matrices on the USPS handwritten digits dataset: (a) empirical PDF of p', p'' (b)-(e) sample image and recovery performances for different antipodal-valued sensing matrices and $m = 1024$ (f) average signal recovery performances of PHE and RBE sensing matrices.



(a) $T^\star$ PHE, RSNR = 15.94 dB



(b) Random PHE, RSNR = 7.40 dB



(c) RBE, RSNR = 12.85 dB

m	MaxDet pool PHE	Random PHE	RBE
64 ($^n/_4$)	**15.12**	2.68	6.94
85 ($\lfloor^n/_3\rfloor$)	**17.20**	3.73	11.11

(d) Average $\mathrm{RSNR}_{\mathrm{dB}}$ over 50 sample ECG tracks, 25 MaxDet pool PHE, 25 Random PHE and 50 RBE sensing matrices.

Fig. 2: Comparison of PHE and RBE sensing matrices on the PhysioNet ECG dataset: (a)-(c) sample ECG and its recoveries for different antipodal-valued sensing matrices and $m = 85$ (d) average signal recovery performances of PHE and RBE sensing matrices.

5. CONCLUSION

We have presented a design criterion to select Hadamard sensing matrices when the signals being acquired by CS are localized, its optimization rationale being the maximization of the measurements' entropy in the assumption that they are approximately Gaussian for most natural signals.

Due to its computational hardness, we have applied a genetic algorithm to choose near-optimal Hadamard sensing matrices in the corresponding deterministic ensemble. Experiments on two signal classes have shown that measurements provided by such matrices yield positive signal recovery performance increments w.r.t. randomly selected matrix designs.

6. REFERENCES

[1] E. J. Candes and M. B. Wakin, "An Introduction to Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[2] D. L. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[3] J. Haboba, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A pragmatic look at some compressive sensing architectures with saturation and quantization," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, Sept. 2012.

[4] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 83–91, 2008.

[5] E. J. Candes and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse problems*, vol. 23, no. 3, pp. 969, 2007.

[6] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A rakeness-based design flow for analog-to-information conversion by compressive sensing," in *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*. IEEE, 2013, pp. 1360–1363.

[7] E. T. Jaynes, "On the rationale of maximum-entropy methods," *Proceedings of the IEEE*, vol. 70, no. 9, pp. 939–952, 1982.

[8] C. Ko, J. Lee, and M. Queyranne, "An exact algorithm for maximum entropy sampling," *Operations Research*, pp. 684–691, 1995.

[9] M. Elad, "Optimized projections for compressed sensing," *Signal Processing, IEEE Transactions on*, vol. 55, no. 12, pp. 5695–5702, 2007.

[10] J.M. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," *Image Processing, IEEE Transactions on*, vol. 18, no. 7, pp. 1395–1408, 2009.

[11] G. Puy, P. Vandergheynst, and Y. Wiaux, "On variable density compressive sampling," *Signal Processing Letters, IEEE*, vol. 18, no. 10, pp. 595–598, 2011.

[12] Z. Wang and G. R. Arce, "Variable density compressed image sampling," *Image Processing, IEEE Transactions on*, vol. 19, no. 1, pp. 264–270, 2010.

[13] M. Mangia, R. Rovatti, and G. Setti, "Rakeness in the Design of Analog-to-Information Conversion of Sparse and Localized Signals," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, no. 5, pp. 1001–1014, May 2012.

[14] J. Ranieri, R. Rovatti, and G. Setti, "Compressive sensing of localized signals: Application to analog-to-information conversion," in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2010, pp. 3513–3516.

[15] W. R. Carson, M. Chen, M. R. D. Rodrigues, R. Calderbank, and L. Carin, "Communications-inspired projection design with application to compressive sensing," *SIAM Journal on Imaging Sciences*, vol. 5, no. 4, pp. 1185–1212, 2012.

[16] W. Chen, M.R.D. Rodrigues, and I.J. Wassell, "Projection design for statistical compressive sensing: A tight frame based approach," *Signal Processing, IEEE Transactions on*, vol. 61, no. 8, pp. 2016–2029, 2013.

[17] F. Renna, M. R.D. Rodrigues, M. Chen, R. Calderbank, and L. Carin, "Compressive sensing for incoherent imaging systems with optical constraints," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, 2013, pp. 5484–5488.

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

[19] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *Information Theory, IEEE Transactions on*, vol. 52, no. 12, pp. 5406–5425, 2006.

[20] D. L. Donoho and J. Tanner, "Precise Undersampling Theorems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 913–924, June 2010.

[21] E. J. Candes, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59–73, 2011.

[22] B. J. Fino and V. Algazi, "Unified matrix treatment of the fast walsh-hadamard transform," *Computers, IEEE Transactions on*, vol. C-25, no. 11, pp. 1142–1146, 1976.

[23] R. G. Baraniuk, V. Cevher, and M. B. Wakin, "Low-dimensional models for dimensionality reduction and signal recovery: A geometric perspective," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 959–971, 2010.

[24] T. M. Cover and J. A. Thomas, *Elements of information theory*, John Wiley & Sons, 2012.

[25] P. P. Vaidyanathan, "The theory of linear prediction," *Synthesis Lectures on Signal Processing*, vol. 2, no. 1, pp. 1–184, 2007.

[26] M. Mitchell, *An Introduction to Genetic Algorithms (Complex Adaptive Systems)*, A Bradford Book, third printing edition, Feb. 1998.

[27] S. S. Chen, D. L. Donoho, and M. A Saunders, "Atomic decomposition by basis pursuit," *SIAM journal on scientific computing*, vol. 20, no. 1, pp. 33–61, 1998.

[28] E. van den Berg and M. P. Friedlander, "SPGL1: A solver for large-scale sparse reconstruction," June 2007, http://www.cs.ubc.ca/labs/scl/spgl1.

[29] Gurobi Optimization, Inc., "Gurobi optimizer reference manual," 2013.

[30] J. J. Hull, "A database for handwritten text recognition research," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 16, no. 5, pp. 550–554, 1994.

[31] A.L. Goldberger *et al.*, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215–220, 2000.

[32] S. Mallat, *A wavelet tour of signal processing*, Access Online via Elsevier, 1999.