COMPRESSIVE INFORMATION ACQUISITION WITH HARDWARE IMPAIRMENTS AND CONSTRAINTS: A CASE STUDY

S. Gopalakrishnan[⋆] *T. Moy*[†] *U. Madhow*[⋆] *N. Verma*[†]

[⋆] University of California, Santa Barbara, Department of ECE, Santa Barbara, CA 93106, USA
[†] Princeton University, Department of EE, Princeton, NJ 08544, USA

ABSTRACT

Compressive information acquisition is a natural approach for low-power hardware front ends, since most natural signals are sparse in *some* basis. Key design questions include the impact of hardware impairments (e.g., nonlinearities) and constraints (e.g., spatially localized computations) on the fidelity of information acquisition. Our goal in this paper is to obtain specific insights into such issues through modeling of a Large Area Electronics (LAE)-based image acquisition system. We show that compressive information acquisition is robust to stochastic nonlinearities, and that appropriately designed spatially localized computations are effective, by evaluating the performance of reconstruction and classification based on the information acquired.

***Index Terms*—** Compressed sensing, stochastic nonlinearities, image acquisition, low-power front ends, hardware impairments.

1. INTRODUCTION

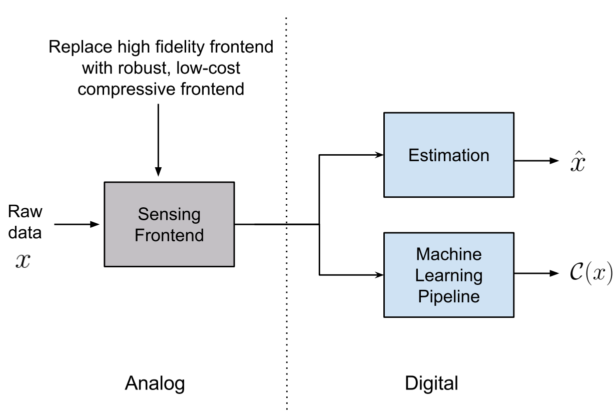


Fig. 1. Using the compressive framework as a front end for learning applications.

Since most natural signals are sparse in some basis, compressive projections [1] are a promising general-purpose approach for low-power front ends for acquisition of information for downstream estimation/learning tasks (Fig. 1). They can be realized using inner products with binary coefficients, which is attractive for hardware implementation, and are expected to be resilient to a broad class of impairments. Successful signal reconstruction has been demonstrated under theoretical models of impairments such as outliers [2] [3] and severe quantization [4] [5], and successful image classification was demonstrated in recent experiments on a Large Area Electronics (LAE)-based image acquisition platform [6], despite significant nonlinearities.

In this paper, we carry out a case study of the LAE-based system in [6] to obtain insight into tradeoffs in designing compressive hardware. To our knowledge there is no other prior work on compressive acquisition in such low-power, high-variability hardware. We model and evaluate the effect of the stochastic nonlinearities in this system, and use the model to explore alternative design choices [7]. While we consider a large-area system, we believe that a similar approach would be highly effective in the design of nanoscale hardware: as semiconductor processes are scaled down, significant stochastic impairments begin to appear in the computational fabric [8]. Our key results are as follows:

1. We develop a synthetic model for the effect of stochastic nonlinearities in the LAE-based system, which allows us to investigate the impact of potential modifications in hardware design via simulations. We provide insight into the effect of these impairments by further simplifying the synthetic model via a (slightly optimistic) Gaussian approximation.
2. The LAE-based system employs row-by-row compressive sensing, with the same matrix employed for all rows. Based on recent theory [7], we expect this to be suboptimal. However, we show that row-by-row compressive projections, which are far easier to implement than projections on the entire image, are competitive in performance, as long as the compressive matrices used are independent across rows.

Section 2 provides a brief review of the compressed sensing principle. Section 3 describes the LAE-based image acquisition system. Section 4 presents a synthetic model of the measurement process, followed by analysis of stochastic nonlinearities and row-wise computations. Finally in Section 5 we provide conclusions.

2. BACKGROUND

A generic approach to dimension reduction for high-dimensional data with an unknown low-dimensional structure is to use compressive transformations that pseudo-randomly project observations to a low-dimensional subspace. Consider a signal $x \in \mathbb{R}^N$ with a K-sparse representation in basis Ψ: $x = \Psi s$, where $\|s\|_0 \leq K$, and linear measurements $y \in \mathbb{R}^M$ (where $M < N$) of the form $y = \Phi x = \Phi \Psi s$.

Compressive sensing theory [1] states that if $M = \mathcal{O}(K \log(N/K))$ and $\Phi\Psi$ satisfies the Restricted Isometry Property (RIP), then pairwise distances are preserved, i.e, $\|x_i - x_j\| \approx \|y_i - y_j\|$ (where x_i, x_j are two instances of the signal x, and y_i, y_j are the corresponding measurements), and x can be reconstructed from y.

In particular, when elements of the sensing matrix Φ are chosen ± 1 with equal probability and basis Ψ is orthonormal, $\Phi\Psi$ satisfies the RIP [9]. Hence compressive signal representations can be obtained at low complexity using binary matrix transformations. As shown by Eftekhari *et al.* [7], it is also possible to satisfy the RIP with compressive projections which operate on portions of the original vector (i.e., with block diagonal Φ), which is attractive for hardware implementations and is directly relevant for our case study, as discussed later.

3. IMAGE ACQUISITION SYSTEM

A practical example of a low-power compressive front end is the LAE-based image acquisition platform [6] in Fig. 2. Images are projected onto an array of photoconductor sensors and illumination levels are sensed as voltages. The matrix of sensors is then scanned row-by-row, and each row x is fed into a compression block. The compression matrix, which is the same for each row, consists of ± 1 elements, implemented via highly variable, low-performance thin-film transistors (TFTs) (hardwired to either a $+1$ or a -1 line). The compressed signal y is collected for all rows and then digitized for classification and reconstruction tasks.

There are two sources of noise in the measurement process (Fig. 3): variations in the image sensors, and variations in $I_d - V_{gs}$ curves across TFTs in the compression block.

4. SYNTHETIC MODEL

We model the nonlinear measurement process as follows:

1. Transform each row of the image x to the range $[l, h]$ of photoconductor output voltages (with nominal values of $l = 10$V, $h = 22$V).

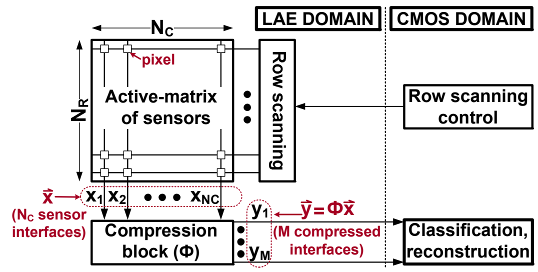


Fig. 2. Architecture of image acquisition system [6].

2. Add piecewise uniform noise (based on experimental data on image sensor variations):

$$\tilde{x} = \begin{cases} u\left(l, l+\Delta\right) & x \leq l + \Delta \\ u\left(x - \Delta/2, x + \Delta/2\right) & l + \Delta \leq x \leq h - \Delta \\ u\left(h - \Delta, h\right) & h - \Delta \leq x \end{cases}$$

(where $u(a, b)$ denotes a realization of the uniform random variable $\mathcal{U}(a, b)$, and the parameter Δ has a nominal value of 5V).

3. Perform random scaling from a set of 80 functions (representing $I_d - V_{gs}$ curve variations) and then compress:

$$y = \Phi f_R(\tilde{x})$$

Overall, this results in the following synthetic model for each row of the image:

$$y = \Phi f_R(x + n(x)) \tag{1}$$

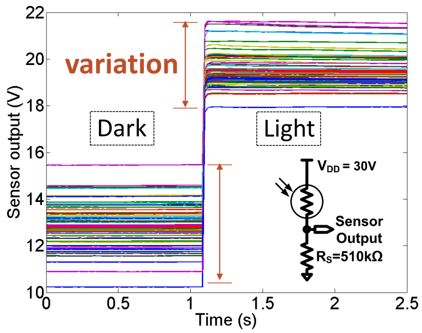
4.1. Verification with measured data

We first attempt to verify the synthetic model via comparisons of reconstruction and classification performance with measured data. Measurements are available for 1500 MNIST images, resized to 80×80. (For more details on the physical setup, see [6]). We perform classification on the measured and synthetic data via RBF-SVM with 1000 training and 500 test images, and reconstruction via the GPSR algorithm [10] with the assumption that the data is sparse in the 2D-DFT basis. Reconstruction performance (Fig. 4) is qualitatively similar for both measured and synthetic data, while classification performance (Fig. 5) indicates that the synthetic model is pessimistic.

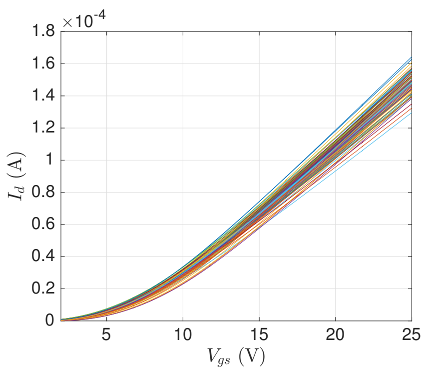
Now we proceed to analyze the impact of hardware impairments. We divide our analysis into two parts: impact of row-by-row compression, and stochastic nonlinearities.

4.2. Row-by-row compression

Eftekhari *et al.* [7] study block diagonal compressive matrices operating on portions of the signal, and show that RIP properties depend on the characteristics of the basis in which



(a)



(b)

Fig. 3. Sources of noise in the measurement process: (a) Variations in the output voltage of image sensors for low and high input illumination (effect of shot noise on photodetection not visible due to the 12V range in the y-axis), (b) I_d vs. V_{gs} curve for TFTs in the compression block, showing variation over 80 devices.

the signal is sparse. These results are directly relevant here, since we consider row-by-row compressive measurements for an image. Specifically, we consider 3 types of binary compression matrices (Fig. 6): (1) Compression on the full image; (2) row-by-row compression with independent matrices across rows; (3) row-by-row compression with the same compressive matrix for each row.

The results in [7] show that matrices of types 2 and 3 can satisfy RIP with high probability, but the minimum number of measurements required scales linearly with the *coherence* of the sparsifying basis for type 2, and with its *block coherence* for type 3. We skip definitions due to lack of space (see [7]).

While we do not exactly know the sparsifying basis for

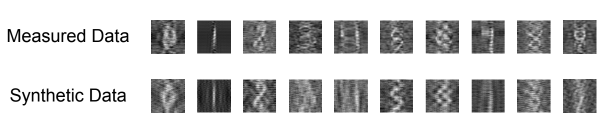


Fig. 4. Images reconstructed from 5x compression: synthetic vs. measured data.

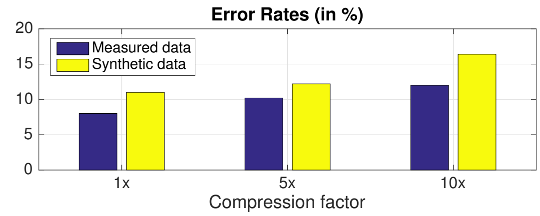


Fig. 5. Classification performance: synthetic vs. measured data.

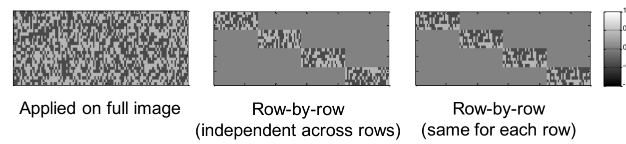


Fig. 6. Three types of compression matrices.

images, we can gain design intuition by assuming that images are sparse in the 2D-DFT basis. The 2D-DFT basis has small coherence and large block coherence (see Appendix), so that type 2 matrices (independent across rows) are expected to perform better than type 3 matrices (used in the experimental system). We test this hypothesis by plotting the CDF of $\|\Phi x\|/\|x\|$ over 60,000 MNIST images (Fig. 7). We see that $\|\Phi x\|/\|x\|$ deviates from 1 when the compressive matrix is the same for each row, but that use of independent matrices leads to near-RIP behavior.

Next we compare the 3 matrices with respect to GPSR reconstruction (Fig. 8) and linear SVM classification, using 60,000 images for training and 10,000 for testing (Fig. 9). Reconstruction and classification results show the same trend as for geometry preservation, especially at higher compression factors.

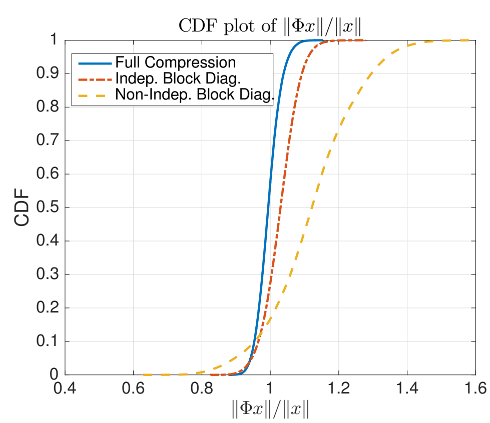


Fig. 7. CDF of $\|\Phi x\|/\|x\|$ over 60,000 MNIST images (upscaled to 80×80)

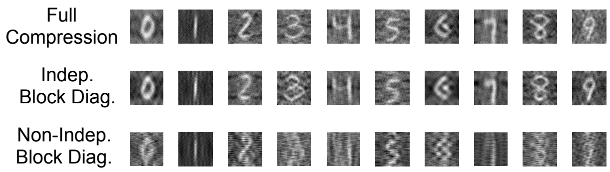


Fig. 8. Images reconstructed from 5x compression with different matrix types.

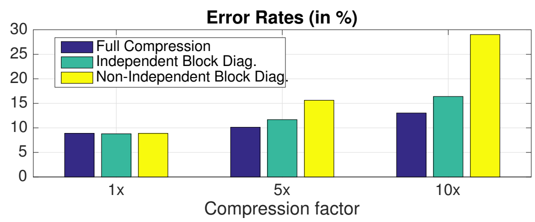


Fig. 9. Classification performance for different matrix types.

4.3. Stochastic nonlinearities

In order to derive insight into how the nonlinearities are impacting performance, we analyze the synthetic noise $z = \Phi f_R(x + n(x)) - \Phi f_0(x)$ by plotting a histogram over 60,000 MNIST images (Fig. 10). We observe the close match with a Gaussian, and therefore propose a simplified noise model: $y = \Phi f_0(x) + n_{\text{Gauss}}$. The variance of n_{Gauss} is estimated from the histogram.

To verify our simplified model, we compare reconstruction and classification performance (Figs. 11 and 12), via GPSR reconstruction and linear SVM classification with 60,000 MNIST images for training and 10,000 for testing. (Images are compressed using matrix type 2). Classification results indicate that the Gaussian model is slightly optimistic, but it is still a reasonable approximation that provides insight into the effect of stochastic impairments.

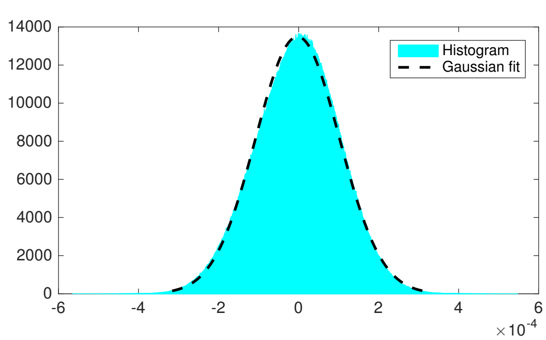


Fig. 10. Histogram of synthetic noise.

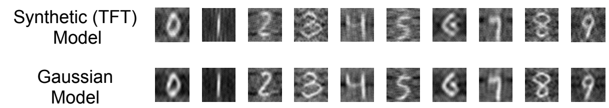


Fig. 11. Images reconstructed from 5x compression (matrix type 2): synthetic vs. Gaussian model.

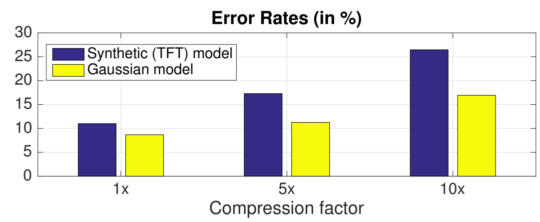


Fig. 12. Classification performance: synthetic vs. Gaussian model.

5. CONCLUSIONS

We show that compressive projections are robust to real-world hardware nonlinearities via our case study of an LAE-based image acquisition system. Our synthetic model, together with guidance from recent theory on block diagonal compressive matrices [7], enables us to explore design tradeoffs. Specifically, spatially localized compressive projections, which are easier to implement, can be highly effective, as long as the compressive matrices are appropriately designed (e.g., independent across rows of an image).

6. ACKNOWLEDGEMENT

This work was supported in part by Systems on Nanoscale Information fabriCs (SONIC), one of the six SRC STARnet Centers, sponsored by MARCO and DARPA.

Appendix: Coherence and block-coherence of the 2D Fourier basis

Consider an image $X \in \mathbb{R}^{L\times L}$. If W is the 1D-DFT matrix of size $L \times L$, the 2D-DFT of X is $\mathcal{F}_2(X) = WXW$. Vectorizing, we get $\mathrm{vec}(\mathcal{F}_2(X)) = \mathrm{vec}\,(WXW) = (W \otimes W)\,\mathrm{vec}\,(X)$ (using the property $\mathrm{vec}\,(ABC) = \left(C^T \otimes A\right)\mathrm{vec}(B)$, where $\otimes$ is the Kronecker product). Hence the 2D-DFT matrix of size $L^2 \times L^2$ is $W_2 = W \otimes W$.

Now we compute the coherence and block-coherence as defined in [7]. Assume that W_2 is of size $N \times N$ and there are J blocks in the compressive matrix. The coherence of W_2 is $\mu(W_2) = \sqrt{N}\max_{i,j}|W_2(i,j)| = 1$.

The block-coherence $\gamma(W_2)$ is defined as $\sqrt{J}$ times the maximal spectral norm when any column of W_2 is reshaped into an $(N/J) \times J$ matrix. Now the entries of the first column of W_2 all equal $1/\sqrt{N}$, so when reshaped into matrix form its spectral norm is 1. Hence the block-coherence of W_2 is $\gamma(W_2) = \sqrt{J}$.

7. REFERENCES

[1] Candès E. J. and Wakin M. B., "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[2] Razavi S. A., Ollila E., and Koivunen V., "Robust greedy algorithms for compressed sensing," in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, 2012, pp. 969–973.

[3] Krzakala F., Mézard M., and Zdeborová L., "Compressed sensing under matrix uncertainty: Optimum thresholds and robust approximate message passing," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 5519–5523.

[4] Jacques L., Degraux K., and De Vleeschouwer C., "Quantized iterative hard thresholding: Bridging 1-bit and high-resolution quantized compressed sensing," in *10th International Conference on Sampling Theory and Applications (SampTA 2013)*, Bremen, Germany, July 2013, pp. 105–108.

[5] Tsagkatakis G. and Tsakalides P., "Recovery of quantized compressed sensing measurements," in *2015 IS&T/SPIE Electronic Imaging Conference, Computational Imaging XIII*, San Francisco, CA, Feb. 2015, vol. 9401.

[6] Moy T., Rieutort-Louis W., Wagner S., Sturm J. C., and Verma N., "A thin-film, large-area sensing and compression system for image detection," *IEEE Transactions on Circuits and Systems I*, vol. 63, no. 11, pp. 1833–1844, Nov. 2016.

[7] Eftekhari A., Yap H. L., Rozell C. J., and Wakin M. B., "The restricted isometry property for random block diagonal matrices," *Applied and Computational Harmonic Analysis*, vol. 38, no. 1, pp. 1–31, 2015.

[8] Shanbhag N. R., Abdallah R. A., Kumar R., and Jones D. L., "Stochastic computation," in *Proceedings of the 47th Design Automation Conference*. ACM, 2010, pp. 859–864.

[9] Baraniuk R., Davenport M., DeVore R., and Wakin M., "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[10] Figueiredo M. A. T., Nowak R. D., and Wright S. J., "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Journal on Selected Topics in Signal Processing*, vol. 1, pp. 586–597, 2003.