Single Letter Formulas for Quantized Compressed Sensing with Gaussian Codebooks

Alon Kipnis*, Galen Reeves† and Yonina C. Eldar‡
* Department of Statistics, Stanford University
† Department of Electrical and Computer Engineering, Department of Statistical Science, Duke University
‡ Department of Electrical Engineering, Technion – Israel Institute of Technology

***Abstract*—Theoretical and experimental results have shown that compressed sensing with quantization can perform well if the signal is very sparse, the noise is very low, and the bitrate is sufficiently large. However, a precise characterization of the fundamental tradeoffs between these quantities has remained elusive. In our previous work, we considered a quantization scheme that first computes the conditional expectation of the signal. In this paper, we focus on a different approach in which the measurements are encoded directly using Gaussian codebooks. We show that that mean-square error (MSE) distortion of this approach can be analyzed by studying a degraded measurement model without any bitrate constraints.**

Building upon ideas from statistical physics and random matrix theory, we then provide single-letter formulas for the reconstruction error associated with optimal decoding. These formulas provide an explicit characterization of the mean-squared error (MSE) as a function of: (1) the average quantization bitrate, (2) the prior distribution of the signal, and (3) the spectral distribution of the sensing matrix. These formulas provide upper bounds on the fundamental limits of compressed sensing with quantization. Interestingly, it is shown that in some problem regimes, this method achieves the best known performance, even though the encoding stage does not use any information about the signal distribution other than its mean and variance.

I. Introduction

This paper considers the problem of reconstructing a random signal vector $X^n = (X_1, \dots, X_n)$ from a quantized version of measurements $Y^m = (Y_1, \dots, Y_m)$ obtained via

$$Y^m = \mathbf{H}X^n + N^m, \tag{1}$$

where $\mathbf{H}$ is an $m \times n$ *sensing matrix* and N^m is i.i.d. standard Gaussian noise. The objective is to understand how the mean-square error (MSE) distortion depends on:

- The number of bits used to encode the measurements.
- The prior distribution of the signal.
- Properties of the sensing matrix, such as the number rows m and the average power of the entries.

Despite a significant amount of work, especially within the compressed sensing (CS) framework [1], the fundamental limits of this problem are still not fully understood. This problem is a special case of the remote (or indirect) source coding problem [2, Ch. 3.5].

Much of the work on CS with quantization has focused on approaches that apply scalar quantization directly to the measurements [3]–[6]. These approaches are straightforward to implement in practice because the quantization does not depend on the sensing matrix or the signal distribution. In our previous work [7], we studied a different approach that first estimates the signal from the measurements and then encodes the estimate using an i.i.d. codebook. By analyzing the asymptotic performance of this approach in the setting of i.i.d. Gaussian sensing matrices and optimal estimation, we showed that it can provide significant improvements, particularly in cases where the signal is very sparse and/or highly non-Gaussian.

Although our results from [7] shed light on what is possible using optimal encoding, one might argue that the idea of estimating the signal prior to quantization goes against the central mantra of CS, which is that the compression and sensing should be done simultaneously with the reconstruction taking place at a later stage. This motivates us to consider in this paper quantization schemes that require only minimal processing of the measurements prior to encoding.

The contribution of this paper is to analyze two new coding schemes, both of which use Gaussian codebooks [8] to encode the measurements:

- *Compress-and-Estimate (CE):* The measurements are encoded using an i.i.d. Gaussian codebook.
- *Linear Compress-and-Estimate (LCE):* The quantizer applies a linear transformation to the measurements and then encodes the transformed measurements using block-wise Gaussian codebooks with variable rates.

In both cases, the signal is reconstructed from the quantized measurements using information about the sensing matrix and signal distribution.

The CE quantization scheme is universal in the sense that the encoding is independent of the sensing matrix and the signal distribution. The LCE quantization scheme adapts to the sensing matrix by assigning a variable bitrates to a linear transformation of measurements. As a consequence, LCE can provide significant improvements, particularly in settings where there is large variation in the singular values of the sensing matrix.

Our theoretical results characterize the performance of the CE and LCE quantization schemes in the setting where the signal entries are drawn i.i.d. from a known signal distribution and the sensing matrix is drawn from a right-orthogonally invariant matrix distribution. In particular, we provide explicit upper bounds for the MSE distortion as a function of the average bitrate, the signal distribution, and the spectral distribution of the sensing matrix.

978-1-5386-4780-6/18/$31.00©2018 IEEE

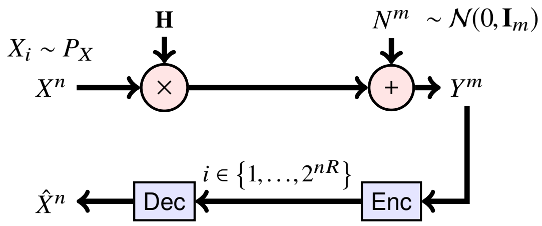

Fig. 1. Signal reconstruction from quantized compressed sensing measurements with $m \times n$ sensing matrix $\mathbf{H}$ and average bitrate R.

Our analysis provides a link between the CS problem with and without bitrate constants. This connection allows us to leverage work focusing on single letter formulas for the asymptotic MMSE. This includes postulated formulas obtained via the replica method from statistical physics [9]–[11] as well as some recent rigorous characterizations for the special case of i.i.d. Gaussian matrices [12], [13]. Results based on the analysis of specific algorithms, such as VAMP [14], can also be used to provide upper bounds.

Finally, we study the performance of our quantization schemes in the setting of a sparse signal drawn from an i.i.d. Bernoulli-Gaussian distribution. Our results show that LCE outperforms the previous quantization methods for certain regions of the problem parameters.

II. Problem Formulation and Related Work

A. Problem assumptions

We consider the measurement model (1) in the setting where the signal X^n is a random vector with i.i.d. entries and the sensing matrix $\mathbf{H}$ is a random matrix that is independent of the signal and the noise. We focus on the case of right-orthogonally invariant matrix distributions such that or any $n \times n$ orthogonal matrix O,

$$\mathbf{H} \overset{\text{dist}}{=} \mathbf{H}O.$$

The dependence on the singular values of the sensing matrix is characterized in terms of the *empirical spectral distribution* μ_n, which is the probability measure on $[0, \infty)$ assigning probability mass $1/n$ to each of the n eignevalues of $\mathbf{H}^T\mathbf{H}$. Note that the average power of the entries in $\mathbf{H}$ is equal to the mean of the spectral distribution:

$$\frac{1}{n}\sum_{i,j} H_{i,j}^2 = \int \lambda \, d\mu_n(\lambda).$$

We analyze the behavior for a sequence of CS problems in which both the signal dimension n and the the number of measurements m increase to infinity with $m/n \to \delta$ for some fixed measurement rate $\delta \in (0, \infty)$.

The problem assumptions are summarized as follows:

A1 The entries of the signal X^n are drawn i.i.d. from a scalar distribution P_X with mean zero and variance σ_X^2.

A2 For each problem of size n, the sensing matrix $\mathbf{H}$ is an $m \times n$ random matrix that is right-orthogonally invariant.

A3 The empirical spectral distribution μ_n converges almost surely to a nonrandom limit μ that is supported on a compact subset of $[0, \infty)$. Furthermore, the maximum eigenvalue is bounded almost surely:

$$\limsup_{n\to\infty} \lambda_{\max}(\mathbf{H}^T\mathbf{H}) \le \infty, \qquad \text{a.s.}$$

B. Reconstruction from quantized measurements

A rate R coding scheme for the CS problem consists of an *encoder* $f : \mathbb{R}^m \to \{1, \ldots, M\}$ that maps the measurements Y^m to one of $M = 2^{nR}$ possible values and a *decoder* $g : \{1, \ldots, M\} \to \mathbb{R}^n$ that produces a reconstruction $\hat{X}^n$ of the source signal. The *distortion* between the signal and its reconstruction is assessed in terms of the expected mean squared error (MSE)

$$\frac{1}{n}\mathbb{E}\left[\left\|X^n - \hat{X}^n\right\|^2\right],$$

where the expectation is with respect to the signal, the measurements, and the sensing matrix. The combined setting of CS measurement model with encoding and decoding at bitrate R is illustrated in Fig. 1.

Definition 1: The *distortion function* $D(P_X, \mu, R)$ is defined to be the infimum over all distortions $D > 0$ for which there exists a rate R coding scheme such that

$$\limsup_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\|X^n - \hat{X}^n\|^2\right] \le D,$$

for any sequence of problems satisfying assumptions A1-A3.

Some special cases of the distortion function $D(P_X, \mu, R)$ are worth mentioning. First, if the columns of the sensing matrix are orthogonal, that is $\mathbf{H}^T\mathbf{H} = \gamma I$, then the spectral distribution is the point-mass distribution δ_γ. In this case, the measurement model in (1) is statistically equivalent to a noisy observation of the signal:

$$Y_i = \sqrt{\gamma}X_i + N_i, \quad i = 1, \ldots, n.$$

Optimal quantization on this setting has been studied in [7]. Furthermore, taking the high SNR limit recovers the usual (Shannon's) distortion-rate function:

$$D^\infty(P_X, R) \triangleq \lim_{\gamma\to\infty} D(P_X, \delta_\gamma, R). \tag{2}$$

Alternatively, in the absence of any bitrate constraints, the optimal reconstruction of the signal under MSE distortion is provided by the conditional expectation. The resulting distortion, which is known as the minimum-mean squared error (MMSE), provides an important baseline for any finite bitrate coding scheme. The MMSE function is defined as:

$$\mathcal{M}(P_X, \mu) \triangleq \lim_{R\to\infty} D(P_X, \mu, R) \tag{3}$$

A great deal of recent work has focused on single-letter formulas the MMSE function [9]–[13], [15].

The results in this paper shows that the MMSE function (3) can be used to *upper bound* the distortion function (2), once an appropriate transformation has been applied to the spectral distribution.

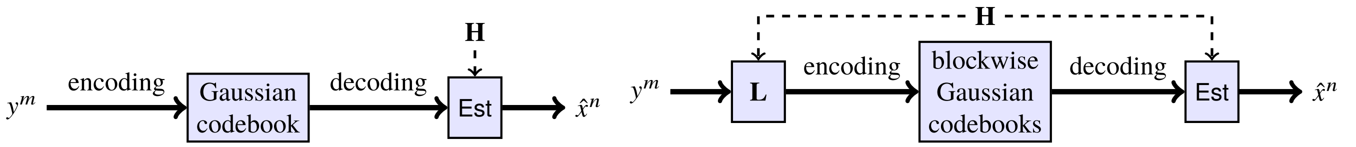

Fig. 2. Left: Compress-and-Estimate (CE) Right: Linear transform Compress-and-Estimate (LCE). Both schemes encode the measurements vector, or a linearly transformed version of it, using a Gaussian codebook. The estimator estimates the source signal from the encoded measurements.

In the ideal setting, both the encoder and decoder in Fig. 1 are designed with full knowledge of the sensing matrix. In more realistic settings, however, the encoder and/or decoder have only partial information about the sensing matrix. Our previous work [7] focused on the case where the *decoder* knows the distribution of the sensing matrix, but not the specific realization. By contrast, the current paper focuses on setting in which the *encoder* has limited information.

III. Achievability using Gaussian Codebooks

A. Compress-and-Estimate

The CE coding scheme consists of the following steps.

- *Gaussian codebook:* Generate a random codebook $\{\hat{y}^m(i))\}_{i=1}^M$ whose columns are i.i.d. Gaussian $\mathcal{N}(0, \tau^2 I_m)$. The codebook is revealed to the encoder and decoder.
- *Minimum distance encoding:* The encoder maps the measurements to the index of the codeword that minimizes the Euclidean distance to the measurements:
$$f(Y^m) = \arg\min_i \|Y^m - \hat{y}^m(i)\|.$$
- *Measurement decoding:* The reconstruction of the measurements is declared to be a the codeword that minimizes the distance to the measurements:
$$\hat{Y}^m = \hat{y}^m(f(Y^m)).$$
- *Signal estimation:* The reconstruction of the signal $\hat{X}^n$ is obtained from the reconstructed measurements $\tilde{Y}^m$ and the sensing matrix $\mathbf{H}$ using an estimator that is based on the measurement model
$$\hat{Y}^m = a\mathbf{H}X^n + bN^m.$$
for some numbers a and b.

In order to describe our results we consider the *push-forward* measure $T\mu$ of the spectral distribution μ with respect to a measurable mapping T, defined as
$$(T\mu)(B) = \mu(T^{-1}(B)) \qquad \text{for all subsets } B.$$

Theorem 1 (CE Achievability): Consider Assumptions 1-3. For every rate $R > 0$
$$D(P_X, \mu, R) \leq \overline{D}_{\mathrm{CE}}(P_X, \mu, R) \triangleq \mathcal{M}(P_X, T\mu), \tag{4}$$
where T is the scaling operator given by
$$T(\lambda) = \frac{1 - 2^{-2R/\delta}}{1 + (\gamma/\delta)\sigma_X^2 2^{-2R/\delta}}\,\lambda, \tag{5}$$
and $\gamma = \int \lambda\, d\mu(\lambda)$ is the mean of μ.

We note that $T\mu$ is merely a scaled version of μ by a constant, hence $\overline{D}_{CE}(P_X, \mu, R)$ is the asymptotic MMSE in estimating X^n from
$$Y^m = \frac{1 - 2^{-2R/\delta}}{1 + \gamma\sigma_X^2 2^{-2R/\delta}/\delta}\mathbf{H}X^n + N^m,$$
i.e., the MMSE under the original measurement model with attenuated SNR.

B. Linear transformation Compress-and-Estimate

The LCE coding scheme is summarized as follows:

- *Linear Transform:* Apply an $m \times m$ linear transformation $\mathbf{L}$ that *diagonalizes* the covariance of transformed measurements:
$$Z^m = \mathbf{L}Y^m, \qquad \mathsf{Cov}(Z^m) = \begin{bmatrix} \sigma_{Z_1}^2 & & \\ & \ddots & \\ & & \sigma_{Z_m}^2 \end{bmatrix}$$
- *Block-wise encoding using Gaussian codebooks:* Each block of the transformed measurements is quantized using an i.i.d. Gaussian codebook and minimum distance encoding. The bitrate allocated to each block is according to the water-filling principle [2].
- *Measurement decoding:* The reconstruction of the measurements is denoted by $\hat{Z}^m$.
- *Signal estimation:* The reconstruction of the signal $\hat{X}^n$ is obtained from the reconstructed measurements $\tilde{Y}^m$ and the sensing matrix $\mathbf{H}$ using an estimator that is based on the measurement model
$$\hat{Z}^m = A\mathbf{H}X^n + bN^m.$$
for some matrix A and scalar b.

An achievable distortion under LCE is as follows:

Theorem 2 (LCE Achievability): Consider Assumptions 1-3. For every rate $R > 0$,
$$D(P_X, \mu, R) \leq \overline{D}_{\mathrm{LCE}}(P_X, \mu, R) \triangleq \mathcal{M}(P_X, T_\theta\mu), \tag{6}$$
where T_θ is the non-linear operator given by
$$T_\theta(\lambda) = \lambda \left[\frac{\lambda\sigma_X^4}{1 + \lambda\sigma_X^2} - \theta\right]^+ \Bigg/ \left(\frac{\lambda\sigma_X^4}{1 + \lambda\sigma_X^2} + \theta\lambda\sigma_X^2\right),$$
$[x]^+ \triangleq \max\{0, x\}$, and θ is the unique positive solution to:
$$R = \frac{1}{2}\int_0^\infty \log^+\left(\frac{\lambda\sigma_X^4}{(1 + \lambda\sigma_X^2)\theta}\right) d\mu(\lambda).$$

C. Discussion

Theorems 1 and 2 show that distortion functions of the CE and LCE coding schemes can be upper bounded in terms of the MMSE function of degraded a measurement model. For the CE coding schemes, this degradation corresponds to a rescaling of SNR. For the LCE scheme, the degradation correspond to a nonlinear transformation of the spectral distribution.

The main advantage of the CE coding scheme is that the encoding depends only on the average power of the signal and the sensing matrix. Hence, the same codebook can be used for a wide variety of settings. The advantage of the LCE coding scheme is that the variable bitrate allocation provides better performance.

IV. Single-Letter Formulas

The results in the previous section provide a link between the CS problem with and without rate constraints. In this section, we leverage recent results focusing on the MMSE function to provide explicit single-letter formulas for the distortion functions associated with the CE and LCE coding schemes.

We will describe the formulas for MMSE function using the characterization given in [16]. Given a scalar signal distribution P_X, the single-letter mutual information and MMSE functions are defined as

$$I_X(s) \triangleq I(X; \sqrt{s}X + N) \tag{7}$$

$$M_X(s) \triangleq \mathbb{E}\left[(X - \mathbb{E}\left[\sqrt{s}X + N\right])^2\right], \tag{8}$$

where $N \sim \mathcal{N}(0,1)$ is independent Gaussian noise. The Legendre transform $I_X^*(u)$ is defined as

$$I_X^*(u) \triangleq \sup_{s \geq 0} \left\{ I_X(s) - \frac{1}{2} su \right\} \tag{9}$$

By the I-MMSE relationship [17], it follows that

$$\frac{d}{ds} I_X(s) = \frac{1}{2} M_X(s), \qquad \frac{d}{du} I_X^*(u) = -\frac{1}{2} M_X^{(-1)}(u)$$

where $M_X^{(-1)}(\cdot)$ denotes the functional inverse of $M_X(\cdot)$.

Given a probability measure μ on the real line, the Stieltjes transform and R transform are defined as

$$C_\mu(t) \triangleq \int \frac{\mu(\lambda)}{\lambda - t} d\lambda, \qquad R_\mu(z) \triangleq C_\mu^{(-1)}(-z) - \frac{1}{z}, \tag{10}$$

where $C_\mu^{(-1)}(\cdot)$ denotes the functional inverse of $C_\mu(\cdot)$,

Using this notation, the replica-symmetric (RS) formula for the MMSE function can be expressed as

$$\mathcal{M}^{\mathrm{RS}}(P_X, \mu) \triangleq \arg\min_u \left\{ I_X^*(u) + \frac{1}{2} \int_0^u R_\mu(-z)\, dz \right\} \tag{11}$$

for all μ such that the minimizer is unique [10], [11], [16]. By the I-MMSE relationship, it follows that every stationary point of the term inside the brackets in (11) is a solution to the fixed-point equation

$$R_\mu(-u) = M_X^{(-1)}(u) \tag{12}$$

In the context of compressed sensing, there are two special cases where the replica-symmetric formula has been proven rigorously: (1) the entries of the sensing matrix are i.i.d. Gaussian and the signal distribution satisfies a certain single-crossing property [18]. (2) The signal entries are i.i.d. Gaussian. These cases are considered in further detail below.

We also note that the vector approximate message passing algorithm (VAMP) of [14] provides an alternative asymptotic analysis for the limiting MSE in the model (1) that does not rely on the replica method. Specifically, the VAMP algorithm attains MSE equals to the largest fixed point of (12).

A. Gaussian Input Signal

Consider the case where P_X is the standard Gaussian distribution $\mathcal{N}(0,1)$. As we shall see, in this case $\overline{D}_{\mathrm{LCE}}(P_X, \mu, R)$ provides a lower bound for the distortion under *any* encoding and decoding scheme in Fig. 1, and not just under LCE. That is, LCE is optimal for the quantized CS problem of Fig. 1.

Since X^n and Y^m are jointly Gaussian given $\mathbf{H}$, the MMSE function is given by

$$\mathcal{M}(P_X, \mu) = \int \frac{\mu(d\lambda)}{1 + \lambda} = C_\mu(-1).$$

Similarly, the CE distortion function defined in Thm. 1 is

$$\overline{D}_{CE}(P_X, \mu, R) = \int \frac{\mu(d\lambda)}{1 + \eta\lambda}.$$

where $\eta \triangleq (1 - 2^{-2R/\delta})/(1 + \gamma 2^{-2R/\delta}/\delta)$.

Furthermore, the LCE distortion function defined in Thm. 2 can be expressed as

$$\overline{D}_{\mathrm{LCE}}(P_X, \mu, R_\theta) = \mathcal{M}(P_X, \mu) + \int \min\left\{\theta, \frac{\lambda}{1+\lambda}\right\} \mu(d\lambda),$$
$$R_\theta = \frac{1}{2} \int \log^+ \left[\frac{\lambda}{(1+\lambda)\theta} \right] \mu(d\lambda). \tag{13}$$

This last expression has a water-filling interpretation similar to the minimal distortion in the combined sampling and source coding problem of [19], [20]. In fact, as the following Theorem states, (13) is the minimal attainable distortion under any encoding and decoding scheme for Y^m when X^n is Gaussian:

Theorem 3 (quantized CS converse for a Gaussian input): Let $P_X = \mathcal{N}(0,1)$. For any encoder $f : \mathbb{R}^m \to \left\{1, \ldots, 2^{nR}\right\}$ and decoder $g : \left\{1, \ldots, 2^{nR}\right\} \to \mathbb{R}^n$,

$$\liminf_{n \to \infty} \frac{1}{n} \mathbb{E}\left[\|X^n - g(f(Y^m))\|^2 \right] \geq \overline{D}_{\mathrm{LCE}}(P_X, \mu, R),$$

for any sequence of matrices satisfying A3.

Proof: (13) is the limit of the of indirect rate-distortion function of X^n given Y^m, hence Thm. 3 follows from the converse side of the indirect source coding theorem [2], [21]. ■

B. IID Gaussian Sensing Matrix

Consider the case where each entry of $\mathbf{H}$ is drawn independently from $\mathcal{N}(0, \gamma/n)$. The resulting $\mathbf{H}$ is right-orthogonally invariant with spectral distribution $\mu = T_\gamma \mu_{\mathrm{MP}}$, where $T_\gamma(\lambda) =$

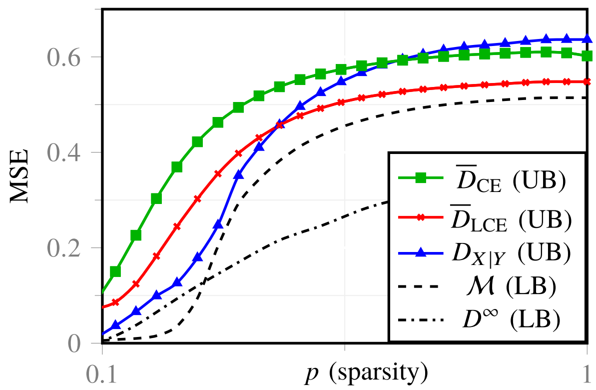

Fig. 3. Bounds on the distortion function $D(P_X,\mu,R)$ as a function of the sparsity p for the Bernoulli-Gaussian distribution (14), $R = 0.75$ bits per source dimension, i.i.d. Gaussian matrix with aspect ratio $\delta = 0.5$ and $\gamma = 50$.

$\gamma\lambda$ and μ_{MP} is the Marchenco-Pasture law. In this situation, (11) reduces to

$$\mathcal{M}^{\mathrm{RS}}(P_X, T_\gamma\mu_{\mathrm{MP}}) \triangleq \arg\min_u \left\{ I_X^*(u) + \frac{\delta}{2}\log(1+\gamma u) \right\},$$

which is equivalent to [9, Eq. 22]. If we further assume that P_X satisfies the single crossing property presented in [18], then $\overline{D}_{\mathrm{CE}}(P_X,\mu,R)$ and $\overline{D}_{\mathrm{LCE}}(P_X,\mu,R)$ are given by $\mathcal{M}^{RS}(P_X, T\mu)$ where T is determined by either Thm. 1 or Thm.2.

Finally, to illustrate the impact of the sparsity, we consider the special case of the Bernoulli-Gaussian signal prior:

$$P_X = (1-p)\delta_0 + p\mathcal{N}(0,1/p), \tag{14}$$

Here, the parameter p is the expected fraction of nonzero entries and variance of the nonzero component is scaled such P_X has unit variance for all p. For this P_X and $T_\gamma\mu_{\mathrm{MP}}$ we illustrate in Fig. 3 the following bounds to $D(P_X,\mu,R)$:

- $\overline{D}_{CE}$ – CE distortion function defined in Thm. 1 evaluated using the single-letter formula (11).
- $\overline{D}_{\mathrm{LCE}}$ – LCE distortion function defined in Thm. 2 evaluated using the single-letter formula (11).
- $D_{X|Y}$ – upper bound from [7].
- $\mathcal{M}$ – lower bound given by the MMSE function (3).
- D^{∞} – lower bound given by Shannon's distortion rate function (2).

Interestingly, it follows from Fig. 3 that the distortion under CE and LCE is smaller than $D_{X|Y}$ when the signal is not sufficiently sparse. This sub-optimality of $D_{X|Y}$ is the result of ignoring dependency between disjoint blocks of $\mathbb{E}[X^n|Y^m,\mathbf{H}]$. We investigate the conditions under which LCE outperform encoding with respect to the asymptotic posterior in our future work [22].

V. Conclusion

We considered the estimation of a signal from a quantized version of its noisy linear measurements. We proposed and analyzed two encoding schemes that are based on encoding the noisy measurements, or their linearly transformed version, using a Gaussian codebook that is independent of the projection matrix. Under a Gaussian signal prior and a particular transformation, this approach is optimal in the sense that it leads to the minimal distortion among all quantization schemes satisfying the bit constraint. For an arbitrary finite variance signal prior, we provided explicit expressions for the MSE distortion under these schemes given in terms of the signal distribution, the spectral density of the sensing matrix, and the bitrate.

References

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

[2] T. Berger, *Rate-distortion theory: A mathematical basis for data compression*. Englewood Cliffs, NJ: Prentice-Hall, 1971.

[3] L. Jacques, D. K. Hammond, and J. M. Fadili, "Dequantizing compressed sensing: When oversampling and non-Gaussian constraints combine," *IEEE Trans. Inform. Theory*, vol. 57, no. 1, pp. 559–571, 2011.

[4] U. Kamilov, V. K. Goyal, and S. Rangan, "Optimal quantization for compressive sensing under message passing reconstruction," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*. IEEE, 2011, pp. 459–463.

[5] W. Dai and O. Milenkovic, "Information theoretical and algorithmic approaches to quantized compressive sensing," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1857–1866, July 2011.

[6] Y. Xu, Y. Kabashima, and L. Zdeborová, "Bayesian signal reconstruction for 1-bit compressed sensing," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2014, no. 11, p. P11015, 2014.

[7] A. Kipnis, G. Reeves, Y. C. Eldar, and A. J. Goldsmith, "Fundamental limits of compressed sensing under optimal quantization," in *Information Theory (ISIT), 2017 IEEE International Symposium on*, 2017.

[8] A. Lapidoth, "On the role of mismatch in rate distortion theory," *IEEE Trans. Inform. Theory*, vol. 43, no. 1, pp. 38–47, 1997.

[9] D. Guo and S. Verdu, "Randomly spread CDMA: asymptotics via statistical physics," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 1983–2010, June 2005.

[10] K. Takeda, S. Uda, and Y. Kabashima, "Analysis of CDMA systems that are characterized by eigenvalue spectrum," *Europhysics Letters*, vol. 76, no. 6, pp. 1193– 1199, Dec. 2006.

[11] A. M. Tulino, G. Caire, S. Verdu, and S. Shamai, "Support recovery with sparsely sampled free random matrices," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4243–4271, 2013.

[12] G. Reeves and H. D. Pfister, "The replica-symmetric prediction for compressed sensing with Gaussian matrices is exact," in *Proc. IEEE Int. Symp. Inform. Theory*, Barcelona, Spain, Jul. 2016, pp. 665 – 669.

[13] J. Barbier, M. Dia, N. Macris, and F. Krzakala, "The mutual information in random linear estimation," in *Proc. Annual Allerton Conf. on Commun., Control, and Comp.*, Monticello, IL, 2016.

[14] S. Rangan, P. Schniter, and A. Fletcher, "Vector approximate message passing," *arXiv preprint arXiv:1610.03082*, 2016.

[15] J. Barbier, F. Krzakala, N. Macris, L. Miolane, and L. Zdeborová, "Phase transitions, optimal errors and optimality of message-passing in generalized linear models," *arXiv preprint arXiv:1708.03395*, 2017.

[16] G. Reeves, in *Proc. Annual Allerton Conf. on Commun., Control, and Comp.*, Monticello, IL, 2017.

[17] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1261–1282, 2005.

[18] G. Reeves and H. D. Pfister, "The replica-symmetric prediction for compressed sensing with Gaussian matrices is exact," *CoRR*, 2016. [Online]. Available: http://arxiv.org/abs/1607.02524

[19] A. Kipnis, A. J. Goldsmith, Y. C. Eldar, and T. Weissman, "Distortion rate function of sub-Nyquist sampled Gaussian sources," *IEEE Trans. Inform. Theory*, vol. 62, no. 1, pp. 401–429, Jan 2016.

[20] A. Kipnis, Y. C. Eldar, and A. J. Goldsmith, "Analog-to-digital compression: A new paradigm for converting signals to bits," *IEEE Signal Processing Magazine*, vol. 35, no. 3, pp. 16–39, May 2018.

[21] R. Dobrushin and B. Tsybakov, "Information transmission with additional noise," *IRE Trans. Inform. Theory*, vol. 8, no. 5, pp. 293–304, 1962.

[22] A. Kipnis, G. Reeves, and Y. C. Eldar, "Fundamental limits of compresses sensing under quantization," 2018, in preparation.