Zero-Delay Joint Source-Channel Coding with a 1-Bit ADC Front End and Receiver Side Information

Morteza Varasteh[†], Borzoo Rassouli[†], Osvaldo Simeone[*] and Deniz Gündüz[†]
[†] Department of Electrical and Electronic Engineering, Imperial College London, London, U.K.
[*] CWCSPR, ECE Department, New Jersey Institute of Technology, NJ, USA.
{m.varasteh12; b.rassouli12; d.gunduz}@imperial.ac.uk, osvaldo.simeone@njit.edu.

***Abstract*—Zero-delay transmission of a Gaussian source over an additive white Gaussian noise (AWGN) channel with a 1-bit analog-to-digital converter (ADC) front end is investigated in the presence of correlated side information at the receiver. The design of the optimal encoder is considered for the mean squared error (MSE) distortion criterion under an average power constraint on the channel input. A necessary condition for the optimality of the encoder is derived. A numerically optimized encoder (NOE) is then obtained that aims that enforcing the necessary condition. It is observed that, due to the availability of receiver side information, the optimal encoder mapping is periodic, with its period depending on the correlation coefficient between the source and the side information. We then propose two parameterized encoder mappings, referred to as periodic linear transmission (PLT) and periodic BPSK transmission (PBT), which trade-off optimality for reduced complexity as compared to the NOE solution. We observe via numerical results that PBT performs close to the NOE in the high signal-to-noise ratio (SNR) regime, while PLT approaches the NOE performance in the low SNR regime.**

***Index Terms*-Joint source channel coding, zero-delay transmission, 1-bit ADC, correlated side information.**

I. Introduction

A key component of the front end of any digital receiver is the analog-to-digital converter (ADC) that is typically connected to each receiving antenna. The energy consumption of an ADC (in Joules/sample) increases exponentially with its resolution (in bits/sample) [1]. This leads to a growing concern regarding the energy consumption of digital receivers, either due to the increasing number of receiving antennas, e.g., for massive multiple-input multiple-output (MIMO) transceivers [2], or due to the limited availability of energy, e.g., in energy harvesting terminals [3]. An energy-efficient operation of digital receivers may hence impose constraints on the resolution of the ADCs that can be employed for each receiving antenna.

Motivated by communication among energy- and complexity-limited sensor nodes, we study zero-delay transmission of analog sensor measurements to a receiving sensor equipped with a 1-bit ADC front end. In keeping with the scenario of a network of sensors, we further assume that the receiving sensor nodes has its own correlated measurement of the transmitted source sample. Focusing on mean squared error (MSE) distortion criterion, our goal is to gain insights into the structure and the performance of optimal encoder and decoder functions when the source sample and the side information are jointly Gaussian.

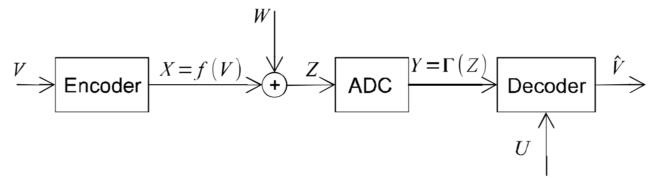

Figure 1. System model for the transmission of a Gaussian source over an AWGN channel with a 1-bit ADC receiver front end and correlated side information at the receiver.

This work contributes to a line of research that endeavors to understand the impact of front end ADC limitations on the performance limits of communication systems. The capacity analysis of a real discrete-time AWGN channel with a K-level ADC front end is studied in [4], proving the sufficiency of $K+1$ constellation points at the encoder. Furthermore, it is shown in [4] that BPSK modulation achieves the capacity when the receiver front end is limited to a 1-bit ADC. In [5], the authors show that, in the low signal-to-noise ratio (SNR) regime, the symmetric threshold 1-bit ADC is suboptimal, while asymmetric threshold quantizers and asymmetric signalling constellations are needed to obtain the optimal performance. Generalization of the analysis from single-antenna AWGN channels to MIMO fading systems are put forth in [6], and, more recently, to massive MIMO systems in [2] and [7]. In [8] some of the authors of this work considered the set-up analyzed here, but in the absence of correlated side information at the receiver. It is noted that the zero-delay constraint prevents the application of the mentioned channel capacity results to this set-up, and that, as it will be seen, the presence of correlated side information at the receiver significantly modifies the optimal design problem.

In this work, we first derive a necessary condition for the optimality of the encoder mapping using calculus of variations. We then develop a gradient-based numerically optimized encoder (NOE). From numerical section, we observe that, similarly to the case with an infinite resolution receiver front end studied in [9]–[11], the optimal encoder mapping is periodic. Furthermore, the period of this function depends solely on the correlation coefficient between the source and the side information, and is independent of the input power constraint, or equivalently the channel SNR. Motivated by the structure of the NOE, we also propose two simple parameterized mappings, which, although being suboptimal, approach the performance of NOE in low and high SNR regimes.

The rest of the paper is organized as follows. In Sec. II, the system model is explained. In Sec. III we present preliminaries and review the previous results, while Sec. IV focuses on the design of encoders and decoders. In Sec. V, numerical results are provided, and Sec. VI concludes the paper. The proofs of the propositions in this work are not included due to space limitations.

Notations: The standard Gaussian distribution is denoted by $\mathcal{N}(0,1)$ with probability density function $\Phi(\cdot)$ and complementary cumulative function $Q(\cdot)$. Unless stated otherwise, the integration intervals are $(-\infty,+\infty)$. The conditional density for standard bivariate Gaussian variables is denoted as

$$\Phi\left(v|u\right)=\frac{1}{\sqrt{2\pi(1-r^2)}}\exp\left\{-\frac{\left(v-ru\right)^2}{2(1-r^2)}\right\}. \qquad (1)$$

II. System Model

A single source sample $V\sim\mathcal{N}(0,\sigma_v^2)$ is transmitted over a single use of an AWGN channel followed by a 1-bit ADC (see Fig. 1). The receiver is provided with side information $U\sim\mathcal{N}(0,\sigma_u^2)$, which follows a bivariate Gaussian distribution with the source sample V with correlation coefficient r.

The channel input is denoted by $X=f(V)$, where $f:\mathbb{R}\to\mathbb{R}$ is the encoder mapping function, which must satisfy the average power constraint $\mathbb{E}[f(V)^2]\leq P$. The received noisy signal at the ADC is

$$Z=f(V)+W, \qquad (2)$$

where the Gaussian noise term $W\sim\mathcal{N}(0,\sigma_w^2)$ is independent of V. The decoder is fed by the output of the ADC given by

$$Y=\Gamma(Z)=\begin{cases}1 & Z<0\\ -1 & Z\geq 0\end{cases}. \qquad (3)$$

The SNR is defined as $\mathrm{SNR}=\frac{P}{\sigma_w^2}$. Having observed Y and U, the decoder produces an estimate $\hat{V}=\hat{v}_Y(U)$ of V.

The goal is to find the optimal encoder function $f(\cdot)$ and the optimal decoder functions $\hat{v}_y(\cdot)$, with $y\in\{-1,1\}$, which jointly minimize the MSE distortion defined as

$$\bar{D}=\mathbb{E}[(V-\hat{V})^2]. \qquad (4)$$

III. Preliminaries

In this section, we consider for reference the scenario in which both the encoder and the decoder have access to the side information U. In this case, without loss of optimality, the encoder can encode the error

$$E=V-\frac{\sigma_v}{\sigma_u}rU, \qquad (5)$$

where the random variable $\sigma_v rU/\sigma_u$ is the minimum MSE (MMSE) estimate of V from U, which can be computed at both encoder and decoder. Since the random variable E, which is distributed as $\mathcal{N}(0,\sigma_e^2)$, with $\sigma_e^2=\sigma_v^2(1-r^2)$, is independent of the side information U, the encoder can directly encode the error E via a mapping function $\tilde{f}(e)$ of the error E by neglecting the presence of the side information U at the receiver. Therefore, the problem reduces to that studied in [8].

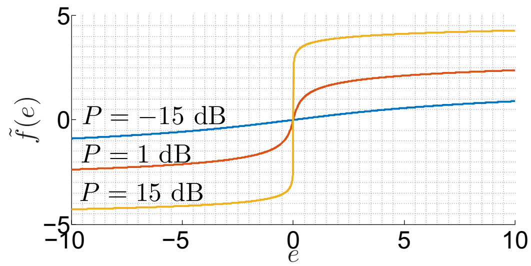

Figure 2. Illustration of the optimal encoder mapping where there is no side information at the receiver ($\sigma_1=\sigma_w=1$). The application if this mapping to the error (5) is optimal if the side information U is also known a the encoder.

In [8, Proposition. I] it is shown that the optimal zero-delay encoder mapping in the absence of side information (assuming the mapping function is odd) is obtained from the implicit equation

$$\tilde{f}(e)\exp\left\{\frac{\tilde{f}(e)^2}{2\sigma_w^2}\right\}=\frac{e}{\sqrt{2\pi}\sigma_w\lambda}, \qquad (6)$$

where $\lambda\geq 0$ is chosen such that the power constraint is satisfied. Examples of the optimal mapping are shown in Fig. 2. It is observed that, in the high SNR regime, that the optimal mapping tends to digital 2-level antipodal signaling, whereas, in the low SNR regime it tends to linear mapping.

In Sec. V, we will use the resulting optimal performance in the presence of side information at both encoder and decoder as a lower bound on the performance of the set-up under study in which the side information is solely available at the receiver.

IV. Transceiver Design

In this section we tackle the design problem introduced in the previous section. We first observe that, for any encoding function, the optimal decoder is always the MMSE estimator; therefore, in this section we focus on the design of the encoder mapping. Our first design is based on derivation of a necessary optimality condition and by an iterative gradient-based numerical optimization algorithm. Due to the relatively high computational complexity of this approach, we also propose two simple yet suboptimal encoder designs.

A. Optimal Encoder and Decoder

The design goal is to minimize the MSE distortion under an average power constraint with respect to the encoder mapping $f(v)$ and the decoding function $\hat{v}_y(u)$. Therefore, we consider the following optimization problem

$$\underset{f,\hat{v}_y}{\text{minimize}}\quad \bar{D}+\lambda\mathbb{E}[f(V)^2], \qquad (7)$$

where $\lambda\geq 0$ is the Lagrange multiplier. With an MMSE estimator at the receiver, the optimal reconstruction function is given by

$$\hat{v}_y(u)\triangleq\mathbb{E}[V|Y=y,U=u] \qquad (8a)$$

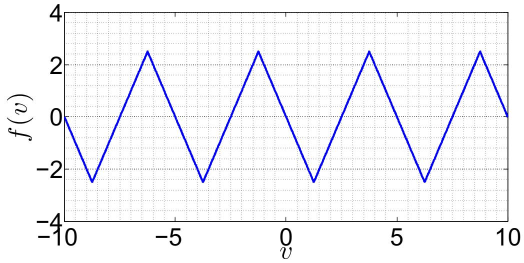

Figure 3. Illustration of the PLT encoder mapping for $\alpha = 2,\ \beta = 2.5$.

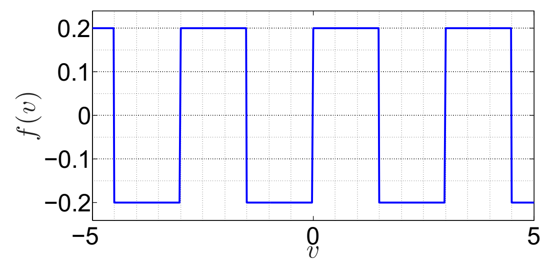

Figure 4. Illustration of the PBT encoder mapping for $\gamma = 0.2,\ \delta = 3$.

$$= \frac{\int v\Phi\left(\frac{v}{\sigma_v}\middle|\frac{u}{\sigma_u}\right) Q\left(\frac{yf(v)}{\sigma_w}\right) dv}{\int \Phi\left(\frac{v}{\sigma_v}\middle|\frac{u}{\sigma_u}\right) Q\left(\frac{yf(v)}{\sigma_w}\right) dv}. \tag{8b}$$

The following proposition provides a necessary condition for the optimal encoder mapping.

Proposition IV.1. *The optimal encoder mapping $f(\cdot)$ for problem (7) must satisfy the implicit equation*

$$2\sqrt{2\pi}\sigma_w\sigma_u\lambda f(v)\exp\left\{\frac{f(v)^2}{2\sigma_w^2}\right\} = 2vA(v) - B(v), \tag{9}$$

where λ is chosen such that the average power constraint P is satisfied. The functions $A(v)$ and $B(v)$ are defined as

$$A(v) \triangleq \int \Phi\left(\frac{u}{\sigma_u}\middle|\frac{v}{\sigma_v}\right)\left(\hat{v}_{-1}(u) - \hat{v}_1(u)\right) du, \tag{10a}$$

$$B(v) \triangleq \int \Phi\left(\frac{u}{\sigma_u}\middle|\frac{v}{\sigma_v}\right)\left(\hat{v}_{-1}(u)^2 - \hat{v}_1(u)^2\right) du, \tag{10b}$$

with $\hat{v}_y(u)$ as in (8).

Proof: The optimization objective function in (7) is continuous and coercive over $f(\cdot)$ (see [12, Sec. A.2] for definitions). This guarantees that (7) is Gateaux differentiable [13, Sec. 7.1]. Using the Theorem I in [13, Sec. 7.4], the gradient of the Lagrangian, denoted by $\nabla_f L$ is obtained as

$$\nabla_f L = -\frac{1}{\sigma_v}\Phi\left(\frac{v}{\sigma_v}\right)\left(2\lambda f(v) - \frac{\exp\left\{-\frac{f(v)^2}{2\sigma_w^2}\right\}}{\sigma_w\sigma_u\sqrt{2\pi}}\cdot(2vA(v) - B(v))\right). \tag{11}$$

Enforcing that the gradient of the Lagrangian in (11) be zero yields the necessary condition in (9) [13, Sec. 7.4]. □

Remark IV.1. *In Sec. V, it will be seen that the application of a gradient descent based optimization procedure that uses (11), yields periodic NOE mappings, whose periods are dependent on the correlation coefficient r. The periodic behaviour of the NOE mappings can be explained with reference to the optimal solution discussed in Sec III, for the scenario in which U is also known at the encoder. In fact, in that case, it was argued that a mapping $f(v) = \tilde{f}(v - \sigma_v ru/\sigma_u)$ is optimal, where $\tilde{f}(\cdot)$ is shown in Fig. 2. Therefore, the optimal mapping is centred on the MMSE estimate $\sigma_v ru/\sigma_u$. When the latter is not available at the encoder, the NOE turns out to consist of periodic replicas of a basic mapping that behaves in a manner similar to $\tilde{f}(\cdot)$ in Fig. 2. As further discussed in Sec. V, the period increases with the variance of the MMSE estimate of V given U, namely $\sigma_v^2(1 - r^2)$.*

To elaborate on the necessary condition (9), we consider the two extreme values of the correlation coefficient r.

- *Uncorrelated sources* ($r = 0$): When the correlation coefficient is zero, the necessary condition (9) reduces to

$$2\sqrt{2\pi}\sigma_w\lambda f(v)\exp\left\{\frac{f(v)^2}{2\sigma_w^2}\right\} = 2v\left(\hat{v}_{-1} - \hat{v}_1\right) - \left(\hat{v}_{-1}^2 - \hat{v}_1^2\right), \tag{12}$$

where $\hat{v}_y$, for $y = 1, -1$ is defined as

$$\hat{v}_y \triangleq \mathbb{E}[V|Y = y] = \frac{\int v\Phi\left(\frac{v}{\sigma_v}\right) Q\left(\frac{yf(v)}{\sigma_w}\right) dv}{\int \Phi\left(\frac{v}{\sigma_v}\right) Q\left(\frac{yf(v)}{\sigma_w}\right) dv}. \tag{13}$$

It can be shown that, in the absence of side information at the receiver, the optimal mapping is an odd function. Therefore, it can be easily verified that we can let $\hat{v}_0 = -\hat{v}_1$ with no loss of optimality. Hence, the equality (12) can be further simplified as

$$f(v)\exp\left\{\frac{f(v)^2}{2\sigma_w^2}\right\} = \frac{v}{\sqrt{2\pi}\sigma_w\lambda}, \tag{14}$$

which is the result (6) obtained in [8, Proposition III.1].
- *Identical sources* ($r = 1$): In this case, we have $\Phi\left(u/\sigma_u|v/\sigma_v\right) = \frac{\sigma_u}{\sigma_v}\delta(u - v)$, where $\delta(\cdot)$ is the Dirac delta function. Therefore, it can be easily verified from (9) that the optimal mapping is $f(v) = 0$, as expected.

Remark IV.2. *Due to the symmetry of the quantizer at the receiver and the symmetry of the noise distribution, we conjecture that the optimal encoder is an odd function of v. While this argument is strengthened by our numerical observations (see Sec. V), we leave the proof of the validity of this conjecture as an open problem for future work.*

In Sec. V, we will present NOE mappings obtained by using a gradient descent approach using (11). As mentioned in Remark IV.1 due to the correlated receiver side information, the resulting mappings are periodic, with a period that depends on the correlation coefficient r. Motivated by this observation, and related results for the case with an infinite-resolution front end in [9], we propose two simple parameterized encoder mappings. Their performance will be compared with that of NOE in Sec. V.

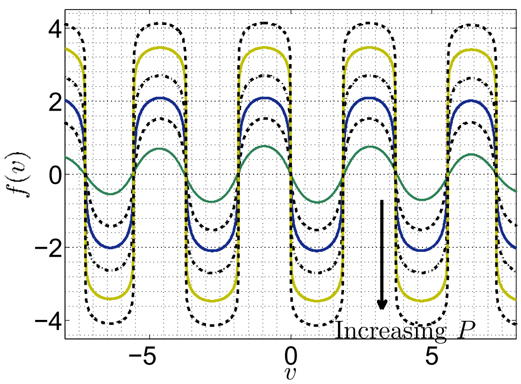

Figure 5. MSE encoder mappings $f(v)$ with different average power values and $r = 0.85$, $(\sigma_v = \sigma_w = 1)$. Increasing the power constraint P has no impact on the period of the NOE mapping.

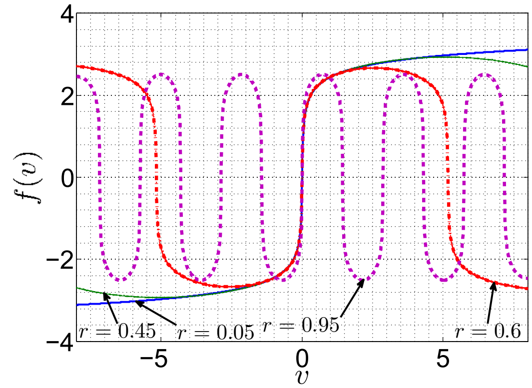

Figure 6. Encoder mappings that minimize MSE distortion for different correlation coefficients r and an average power constraint of $P = 5$ $(\sigma_v = \sigma_w = 1)$.

B. Periodic Linear Transmission (PLT)

The first proposed encoder mapping is a function, which is a periodic linear function with period 2β and slope α within each period. The encoder function is defined as

$$f(v) = \alpha(-1)^{\left\lfloor \frac{v}{\beta}+\frac{1}{2}\right\rfloor} \cdot \left(\beta \left\lfloor \frac{v}{\beta} + \frac{1}{2} \right\rfloor - v\right), \quad (15)$$

where $\lfloor x \rfloor$ is the largest integer less than or equal to x. In Fig. 3, an illustration of this mapping for $\alpha = 2$, $\beta = 2.5$ is shown. In (15) we optimize the parameters α and β under a given average power constraint in order to minimize the MSE distortion.

C. Periodic BPSK Transmission (PBT)

The second proposed encoder mapping, unlike NOE and PLT, adopts digital modulation with two levels, namely, γ and $-\gamma$, with a period of δ. The mapping is defined as

$$f(v) = \gamma \left(1 + 2\Gamma(v) \cdot \text{mod}\left(\left\lfloor \frac{2v}{\delta} \right\rfloor\right)_2\right), \quad (16)$$

where $\text{mod}(\cdot)_2$ is the argument in modulo 2. In Fig. 4, an illustration of this mapping for $\gamma = 0.2$, and $\delta = 2.5$ is shown. Due to the average power constraint, we set $\gamma = \sqrt{P}$, and parameter δ is optimized to minimize the MSE.

D. Shannon Lower Bound (SLB)

A lower bound on the MSE distortion can be obtained by relaxing the zero-delay constraint, and using the Shannon source-channel separation theorem. In [4], it is shown that the capacity of the AWGN channel with a 1-bit ADC in (3) is given by

$$C = 1 - h\left(Q\left(\sqrt{\text{SNR}}\right)\right), \quad (17)$$

where $h(\cdot)$ is the binary entropy function defined as $h(p) \triangleq -p\log_2 p - (1-p)\log_2(1-p)$. Furthermore, the rate-distortion function of a Gaussian source with correlated Gaussian side information at the receiver is given by the Wyner-Ziv rate-distortion function [14]

$$R(\bar{D}) = \frac{1}{2}\left[\log_2 \frac{\sigma_v^2(1-r^2)}{\bar{D}}\right]^+, \quad (18)$$

where $[x]^+ = \max(0, x)$. Combining (17) and (18) a lower bound on the MSE distortion $\bar{D}$ is obtained as

$$\bar{D}_{\text{lower}} = (1-r^2)\sigma_v^2 2^{-2\left(1-h\left(Q\left(\sqrt{\text{SNR}}\right)\right)\right)}. \quad (19)$$

V. Numerical Results

In this section, we present numerical results with the aim of assessing the performance of the encoder/ decoder pairs proposed in the previous sections. In order to derive NOE mapping functions we apply a gradient descent-based algorithm. The algorithm performs a gradient descent search in the direction of the derivative of the Lagrangian (11) with respect to the encoder mapping $f(\cdot)$. The update is done as

$$f_{i+1}(v) = f_i(v) - \mu\nabla_f L, \quad (20)$$

where i is the iteration index, $\nabla_f L$ is defined in (11) and $\mu > 0$ is the step size. The algorithm is initialized with an arbitrary mapping, e.g., linear mapping. It is noted that the algorithm is not guaranteed to converge to a global optimal solution. We also remark that different power constraints are imposed by means of a linear search over the Lagrange multiplier λ.

In Fig. 5, NOE mappings are plotted for different average power constraints, for a correlation coefficient of $r = 0.85$. We note the periodic structure of the mapping, which is due to the available side information at the receiver as discussed in Remark (IV.1). In contrast, the optimal mapping obtained in [8] for $r = 0$ is a monotonically increasing function (see Fig. 2). We also observe that the average power constraint, does not affect the period of the mapping. In Fig. 6, NOE mappings for an average power of $P = 5$ are plotted for different correlation coefficients. We see that the period of the mapping indeed depends on r: the higher the correlation coefficient r the smaller the period of the mapping.

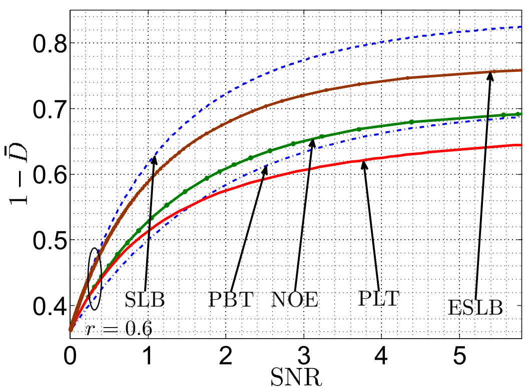

Figure 7. Complementary MSE distortion vs. SNR for $r = 0.6$ ($\sigma_v = \sigma_w = 1$).

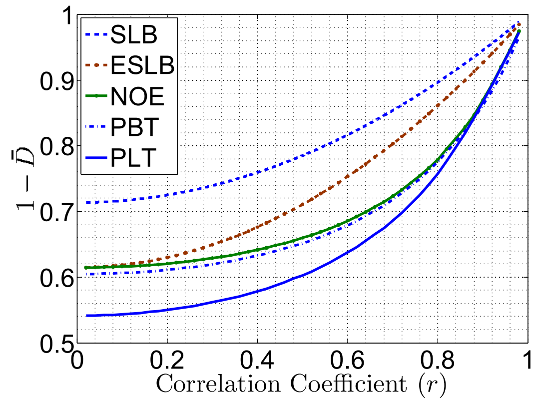

Figure 8. Complementary MSE distortion versus correlation coefficient under the average power constraint $P = 5$ ($\sigma_v = \sigma_w = 1$).

In Fig. 7, we plot the complementary MSE distortion $(1-\bar{D})$ versus SNR for NOE, as well as for the PLT and PBT schemes, for the correlation coefficients $r = 0.6$. The SLB and the MSE distortion for the reference case reviewed in Sec. III in which both encoder and decoder have access to the side information U, which is referred to as encoder side lower bound (ESLB), are also included for comparison. We observe that the performance of PBT is close to that of NOE at high SNR values. On the other hand, for low SNRs, PLT outperforms PBT and approaches the NOE performance. The results are aligned with the extreme-SNR behaviour of the mapping in Fig. 2 in the case of no side information (see Sec. III).

In Fig. 8, the complementary MSE distortion $(1 - \bar{D})$ is plotted versus the correlation coefficient r for a fixed average power constraint of $P = 5$. We see that, for this SNR value, PBT performs very close to NOE for a wide range of r values. However, as r approaches 1, PLT outperforms PBT, and approaches the performance of NOE. This can be explained based on the observation in Fig. 5 that, as average power constraint decreases, the NOE mapping functions resembles the PLT mapping.

We finally observe from the comparison of the SLB and ESLB bounds in both Fig. 7 and Fig. 8, that the zero-delay constraint entails a significant loss with respect to the case in which block processing is allowed.

VI. Conclusion

We have studied the problem of zero-delay transmission of a Gaussian source over an AWGN channel followed by a 1-bit ADC front end, in the presence of correlated side information at the receiver. We have adopted the MSE distortion criterion with average power constraint at the transmitter. We first derived a necessary condition for the optimality of an encoder function, and then, based on this condition, and using gradient descent algorithm, we obtained a numerically optimized encoder mapping. We observed that this encoder mapping is periodic, with a period that depends on the correlation coefficient of the side information. This motivated us to propose two new periodic parameterized encoding schemes, referred to as PLT and PBT. Finally, we have shown through numerical simulations that, PLT and PBT perform close to the NOE in the low and high SNR regimes, respectively.

References

[1] B. Murmann, "ADC performance survey," *CoRR*, vol. abs/1404.7736, 1997-2014. [Online]. Available: http://web.stanford.edu/~murmann/adcsurvey.html
[2] C. Risi, D. Persson, and E. G. Larsson, "Massive MIMO with 1-bit ADC," *CoRR*, vol. abs/1404.7736, 2014. [Online]. Available: http://arxiv.org/abs/1404.7736
[3] D. Gunduz, K. Stamatiou, N. Michelusi, and M. Zorzi, "Designing intelligent energy harvesting communication systems," *IEEE commun. magazine*, vol. 52, pp. 210–216, 2014.
[4] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3629–3639, Dec. 2009.
[5] T. Koch and A. Lapidoth, "At low snr, asymmetric quantizers are better," *IEEE Trans. on Inf. Theory*, vol. 59, no. 9, pp. 5421–5445, Sept 2013.
[6] A. Mezghani and J. Nossek, "On ultra-wideband MIMO systems with 1-bit quantized outputs: Performance analysis and input optimization," *IEEE Int. Sym. Inf. Theory*, pp. 1286–1289, Jun. 2007.
[7] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "One-bit massive MIMO: Channel estimation and high-order modulations," *CoRR*, vol. abs/1504.04540, 2015. [Online]. Available: http://arxiv.org/abs/1504.04540
[8] M. Varasteh, O. Simeone, and D. Gündüz, "Joint source-channel coding with one-bit ADC front end," *IEEE Int. Sym. Inf. Theory*, Jul 2016.
[9] E. Akyol, K. B. Viswanatha, K. Rose, and T. A. Ramstad, "On zero-delay source-channel coding," *IEEE Trans. on Inf. Theory*, vol. 60, no. 12, pp. 7473–7489, Dec 2014.
[10] M. S. Mehmetoglu, E. Akyol, and K. Rose, "A deterministic annealing approach to optimization of zero-delay source-channel codes," *CoRR*, vol. abs/1304.6969, 2013. [Online]. Available: http://arxiv.org/abs/1304.6969
[11] X. Chen and E. Tuncel, "Zero-delay joint source-channel coding for the Gaussian Wyner-Ziv problem," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, July 2011, pp. 2929–2933.
[12] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.
[13] D. Luenberger, *Optimization by Vector Space Methods*. New York: John Wiley & Sons, Inc, 1969.
[14] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *submitted to IEEE Trans. Information Theory*, vol. 22, no. 1, pp. 1–10, January 1976.