Joint Design of Codes and Decoders for Minimization of Transmit, Decoding, and A/D Power

Karthik Ganesan†, Pulkit Grover‡, Jan Rabaey†, Andrea Goldsmith‡
†-University of California, Berkeley ‡-Stanford University
karthik.ganesan@berkeley.edu, pulkit@stanford.edu, jan@eecs.berkeley.edu, andrea@wsl.stanford.edu

I. Introduction

The design of error correcting codes and their decoders is usually done in isolation. The code is often designed first with the goal of minimizing the gap from Shannon capacity [1] and attaining the target error probability. To reflect the concerns of implementation, the code is usually chosen from a family of codes that can be decoded with low "complexity[1]" [3]. On the implementation side, decoders are carefully designed (see e.g. [4]) for the chosen code with the goal of consuming low power while achieving the required decoding throughput[2]. This "division of labor" has been extremely successful and forms the backbone of many modern long-distance communication systems.

However, this approach can be suboptimal for short-distance communication, where the power consumed in processing can dominate transmit power [5], [6]. For instance, while *irregular* LDPC codes approach capacity for many channels (e.g. [7]), the 10 GBASE-T standard for communication within a data-center [8] uses *regular* LDPC codes because they require less decoding power (due to reasons explained in [5]). In short-distance wireless communication systems (e.g. wireless LAN, or the 60 GHz band [6]), it has been proposed to use *uncoded transmission* in order to do away with the decoding altogether! Is that the "optimal" strategy? How do we design codes and decoders so that the total power of the communication link is minimized?

Shannon-theoretic limits, complemented by modern coding-theoretic constructions [3], have provided codes that are provably good for minimizing transmit power. However, significant sources of processing power in transceivers (notably the decoder, encoder, ADCs, DACs, and Equalization) are often heavily influenced by the choice of the code and are not accounted for in such an approach. Can we develop a parallel approach in order to minimize the total system power? With simplistic encoding/decoding models, the issue of fundamental limits on transmit + encoding + decoding power has been addressed in some recent works [5], [9], [10], [11], [12]. These fundamental limits abstract power consumed in computational nodes [5], [11], [12] and wiring [10], [9] in the encoder/decoder implementation and can provide insights into the choice of the code and the decoding algorithm.

While such theoretical insights can serve to guide the choice of the code family, the simplicity of these theoretical models, which (to an extent) is needed[3] in order to be able to obtain fundamental results, also limits their applicability. Even if the models are refined further, the large-deviations techniques used [5], [9] are usually tight only in asymptopia (even though the obtained results are non-asymptotic). Thus, at reasonably high error probability (e.g. 10^{-6}) and small distances (e.g. less than five meters), it is unlikely that the bounds themselves can be used to give precise answers on what codes to use.

In this work, we therefore take a middle path that mixes theory and practice. First, observing the order-optimality of regular LDPC codes in some theoretical models [5] (namely those where computational nodes consume all the power), we restrict our attention to regular LDPC codes[4]. In order to be able to make an educated choice of degrees of the LDPC code and the code girth, we make use of circuit models of decoding power consumption for simple decoding algorithms and regular LDPC codes which are detailed in [13, Section IV]. The models are developed by rigorously simulating (post-layout) power consumption for some simple codes and decoders, breaking down the circuit power into its constituents (e.g. power consumed in computation at nodes, wires, etc.), and generalizing these constituents of power to all structurally similar codes.

These models allow for an exploration of the decoding power for different codes without requiring an implementation. However, because analog and RF circuits are often the dominant sinks of power consumption in transceivers (see e.g. [14], [15]) they should also be taken into account in an optimization. As a start, we take into account the ADC power at the receiver[5]. Based on the observation that the required signaling constellation size for a fixed data-rate and bandwidth will vary with the code-rate, the number and resolution of ADCs at the receiver can be chosen based on the code (see Section [**?**]). We then make use of circuit models for Nyquist ADCs developed in [17] to determine

Because this paper straddles theory and practice, some terms used commonly in circuits are introduced via footnotes.

[1]This complexity is often measured in the "number of operations" in an approximate order sense. In practice, these complexity notions often do not correlate well with the power consumed by the decoder (see e.g. [2]).

[2]"Decoding throughput" is the rate of decoding, measured in information bits per second.

[3]In a nut-shell, the models assume that any synchronous VLSI circuit is a set of computational nodes connected to each other using wires. The simplicity of the models is by necessity: they have to be general and yet analyzable.

[4]While this is a start, we believe that at small blocklengths it is important to investigate other (more classical) coding techniques such as RS codes and BCH codes.

[5]Since ADC power is often believed to dominate that of other components [16]).

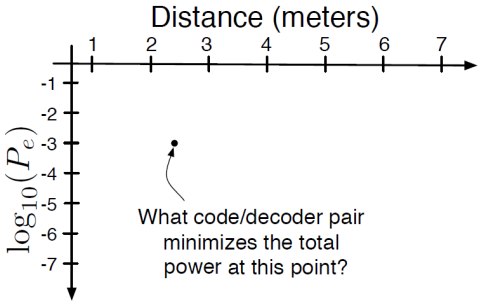

Fig. 1: The question this work addresses: what is the most power-efficient code-decoder pairing for a given distance and error probability P_e? Shannon-theory provides the answer at large distances, ignoring processing power. Including decoding and A/D power in optimization brings out another dimension of the problem: the path-loss of communication. The question therefore ties in the distance of communication with the choice of the code-decoder pair.

the required A/D power for a given code.

In traditional transmit-power-centric exploration, the results are plotted as "waterfall" curves (with corresponding "error-floors," see e.g. [18]) demonstrating how close the code performs to the ideal Shannon limit. There, the channel path-loss can usually be ignored because it shows up as a scaling factor for the term to be optimized, namely, the transmit power, thereby not affecting the optimizing code. Since we are interested in *total* power, the path-loss affects the code choice. For simplicity of understanding, we translate path-loss into the more relatable metric of communication distance using a simple model of path-loss. The resulting question is crystallized in Fig. 1.

In Section IV, we present some optimization results for this question for the limited class of codes and decoders we consider. The results show a graceful increase in the complexity of the suggested code and decoding algorithm as the communication distance is increased, or as the target error probability is lowered. We then introduce multiplicative factors for the analog and digital circuit power and consider a scenario where A/D conversion is the dominant sink of the total processing power. The results suggest that in this scenario, uncoded transmission becomes even more favorable at high error probabilities and short distances. In addition even in regimes where coding is suggested, highly-complicated codes are most favorable.

Our hope is that this method of presenting optimizing codes and decoders can help designers choose codes and decoders with stronger guidance. While the results here are merely suggestive of what the true optimal code/decoder pair would look like, the larger goal is to simulate and model a larger number of codes and decoders before implementing them in a full system.

II. Problem statement

Suppose we want to design a point-to-point communication system that operates over a given channel. We are given a target error probability P_e, communication distance r, and system data-rate R_{data} that the link must operate at. Our general goal is to find the code and decoding algorithm $(\hat{c}, \hat{d})$ such that

$$P_{TX}(\hat{c},\hat{d}) + P_{Dec}(\hat{c},\hat{d}) + P_{ADC}(\hat{c})$$
$$= \min_{c\in\mathcal{C}, d\in\mathcal{D}} \left(P_{TX}(c,d) + P_{Dec}(c,d) + P_{ADC}(c)\right)$$

Where $P_{TX}(c,d) + P_{Dec}(c,d) + P_{ADC}(c)$ is the minimum required transmit + decoding + A/D power for a coded system using (c,d) to satisfy the error-probability, distance, and data-rate requirements. In this work, we consider $\mathcal{C}$ to be the set of regular, binary LDPC codes of variable-node (VN) degrees $3 \le d_v \le 5$, check-node (CN) degrees $4 \le d_c \le 13$, and code girths $g \in \{6, 8, 10, 12, 14\}$. Because of our focus on simple decoders, we consider $\mathcal{D}$ to be a set consisting of only two iterative message-passing decoding algorithms that pass one-bit (Gallager-A [19]) or two-bit (as proposed in [20]) messages.

III. System model, assumptions, and notation

The channel is assumed to be binary-input AWGN with flat-fading, with noise variance $\sigma_z^2 = kTW$, where k is the Boltzmann constant (1.38×10^{-23} J/K), T is the temperature (300 K) and W is the passband bandwidth of the channel. The power is assumed to decay with the path-loss model of $1/r^{\alpha}$, where α is the path-loss coefficient. The transmission strategy uses BPSK or square-QAM modulation, mapping codeword bits to constellation symbols. We assume the transmitter signals at a symbol-rate of W symbols/s and that the minimum square constellation size (M) which satisfies the system data-rate requirement is chosen. Explicitly, M is always the smallest square of an even integer for which:

$$M \ge 2^{R_{data}/(W\times\rho_{code})}$$

Here, ρ_{code} is the rate of the code: $\left(1-\frac{d_v}{d_c}\right)$ for regular LDPC codes. The decoder is assumed to perform a hard-decision on the observed channel outputs before starting the decoding process, thereby first recovering noisy codeword bits transmitted through a Binary Symmetric Channel (BSC) before decoding them.

The received SNR $\left(\frac{E_b}{N0}\right)$ for our system is obtained from a modified Friis transmission equation [21] as a function of system parameters and transmit power P_{TX}:

$$\frac{E_b}{N0} = \left(\frac{P_{TX}}{kTW\left(\frac{r}{\lambda}\right)^{\alpha}\log_2(M)}\right)$$

where λ is the wavelength of transmission at center frequency f_c in Hz, and $\lambda = 3\times10^8/f_c$. The channel error probability for BPSK transmissions is simply:

$$p_{ch} = \mathbb{Q}\left(\sqrt{\frac{E_b}{N0}}\right)$$

And, the channel error probability for M-ary square QAM

is [22]:

$$p_{ch} = \frac{1}{\log_2(\sqrt{M})} \sum_{k=1}^{\log_2(\sqrt{M})} \sum_{j=0}^{(1-2^{-k})\sqrt{M}-1} \left[(-1)^{\left\lfloor \frac{j \times 2^{k-1}}{\sqrt{M}} \right\rfloor} \times \left(2^{k-1} - \left\lfloor \frac{j \times 2^{k-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor \right) \times 2\mathbb{Q}\left((2j+1) \sqrt{\frac{3\log_2(M) \times \frac{E_b}{N0}}{(M-1)}} \right) \right]$$

Where $\mathbb{Q}$ is the right-tail cumulative density function of the standard normal distribution, $\mathbb{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{\frac{-u^2}{2}} du$. Hence, p_{ch} can be determined from M, P_{TX}, and r. For a target P_e, the required channel error probability p_{ch} can be determined for each decoding algorithm (see Figure **??**), which can be used to find the minimum required P_{TX} to meet all the system-level specifications when using a chosen code-decoder pair.

When a QAM constellation is used, we assume that receiver first isolates the I and Q channels and uses a separate ADC for each before passing the digital output bits to the channel decoder. For QAM, we assume the number of bits of each of the ADCs is $log_2(\sqrt{M})$. If BPSK modulation is used we assume only a single comparator is used before the digital baseband at the receiver. We assume all ADCs always sample at the Nyquist rate (in our setup, W samples/s). When calculating the error-probability of the system, we ignore the noise figure of the circuits in the analog front-end of the receiver.

For presenting our results in Section IV, we assume $R_{data} = 7$ Gb/s which is required to be equal to or smaller than the decoding throughput. We assume a channel center frequency of $f_c = 60$ GHz and passband bandwidth of $W = 7$ GHz. The communication distances considered in this work are significantly larger than the wavelength of transmission (≈ 0.5 cm) so that the "far-field approximation" applies.

A. Modeling the decoder implementation

The circuit model for the decoder used in our analysis is presented and explained in [13, Section IV]. The model assumes a synchronous, fully-parallel LDPC decoding architecture which is implemented in 90 nm CMOS. The expression for the overall power consumption of the decoder ($P_{Dec}(b, g, d_v, d_c, R_{data})$) is:

$$\begin{aligned} P_{Dec}(b, g, d_v, d_c) = {} & N(g, d_v, d_c) \times P_{VN}(b, g, d_v, d_c, R_{data}) \\ & + N(g, d_v, d_c) \times \frac{d_v}{d_c} \times P_{CN}(b, g, d_v, d_c, R_{data}) \\ & + N(g, d_v, d_c) \times d_v \\ & \times b \times P_{interconnect}(b, g, d_v, d_c, R_{data}) \end{aligned}$$

where b is the number of message-passing bits used in the decoder, g is the girth of the code, d_v is the VN degree, d_c is the CN degree, and R_{data} is the system data-rate which we assume to be equal to the decoding throughput. $N(g, d_v, d_c)$ is the minimum blocklength (of codes found in [23],[24], and [25]) for a code with parameters g, d_v, and d_c. $P_{VN}(b, g, d_v, d_c, R_{data})$ and $P_{CN}(b, g, d_v, d_c, R_{data})$ are the modeled power consumption of a single VN and CN, respectively, in the decoder, and $P_{interconnect}(b, g, d_v, d_c, R_{data})$ is the modeled power consumption of a single message-passing interconnect in the decoder.

Note that $N(g, d_v, d_c)$ is equal to the total number of VNs in the decoder, $N(g, d_v, d_c) \times \frac{d_v}{d_c}$ is equal to the total number of CNs, and $N(g, d_v, d_c) \times d_v \times b$ is equal to the total number of message-passing interconnects. Hence, this expression is simply a sum of the power consumed by all computation nodes and interconnects in the circuit. The modeling and justification of VN and CN power is detailed in [13, Appendix III] and the modeling of interconnect power is detailed in [13, Appendix IV].

B. Modeling the ADC implementation

The circuit model for the Nyquist ADC is presented and explained in [17]. The derivation of the model is independent of the exact CMOS process, but the models include some technology-dependent parameters. The authors provide models for both pipeline and flash ADC architectures (see [26]) under scenarios where the power consumption is limited by noise, the CMOS process, and device mismatches within the circuit. In this work, we focus on the set of models which take into account noise and process constraints which are detailed in [17, Section IV.C].

The authors first assume that the size of the sampling capacitor C_s of the ADC is chosen such that the quantization noise of the ADC is equal to the sampling noise[6]. This results in a sampling capacitor size of:

$$C_s = 12kT \frac{2^{2n}}{VDD^2}$$

Where k is again the Boltzmann constant (1.38×10^{-23} J/K), T is the temperature (300 K), n is the number of bits of the ADC (in our case, $log_2(\sqrt{M})$), and VDD is the nominal supply voltage of the 90nm CMOS process (we assume it to be 1.2 V).

A minimum bound on the power required for the ADC to sample at a rate of f_s samples/s is given as:

$$P_s = 24kT f_s 2^{2n}$$

Where the sampling rate f_s used in this work is again W samples/s.

IV. Joint optimization over code-decoder pairs

Acknowledgements

This research was supported in part by the Interconnect Focus Center of the Semiconductor Research Corporation, and in part by the NSF Center for Science of Information (CSoI). We thank Anant Sahai, Yang Wen for stimulating discussions, and Jose Moura for providing code constructions that we started our simulations with. We thank the students, faculty, staff and sponsors of the Berkeley Wireless Research

[6] A common choice for Nyquist ADCs [26]

Center and the Wireless Foundations group at Berkeley. In particular, Brian Richards assisted with the simulation flow and Tsung-Te Liu advised us on modeling circuits. We also acknowledge STMicroelectronics for making available the technology information that was used in our simulations.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. Jour.*, vol. 27, pp. 379–423, 623–656, Jul./Oct. 1948.

[2] K. Ganesan, P. Grover, and J. Rabaey, "The power cost of over-designing codes," in *Proc. of the 2011 IEEE Workshop on Signal Processing Systems (SiPS)*, Oct. 2011, pp. 128 –133.

[3] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2007.

[4] Z. Zhang, V. Anantharam, M. Wainwright, and B. Nikolic, "An efficient 10 GBASE-T ethernet LDPC decoder design with low error floors," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 4, pp. 843 –855, Apr. 2010.

[5] P. Grover, K. Woyach, and A. Sahai, "Towards a communication-theoretic understanding of system-level power consumption," *IEEE J. Select. Areas Commun.*, vol. 29, no. 8, pp. 1744 – 1755, Sept. 2011.

[6] C. Marcu, D. Chowdhury, C. Thakkar, J.-D. Park, L.-K. Kong, M. Tabesh, Y. Wang, B. Afshar, A. Gupta, A. Arbabian, S. Gambini, R. Zamani, E. Alon, and A. Niknejad, "A 90 nm CMOS low-power 60 GHz transceiver with integrated baseband circuitry," *IEEE J. Solid-State Circuits*, vol. 44, no. 12, pp. 3434 – 3447, 2009.

[7] S.-Y. Chung, J. Forney, G.D., T. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 db of the shannon limit," *IEEE Commun. Lett.*, vol. 5, no. 2, pp. 58 –60, Feb. 2001.

[8] *IEEE Std. 802.3an-2006: "Physical Layer and Management Parameters for 10 Gb/s Operation, Type 10GBASE-T," amendment to IEEE Std. 802.3-2005*. IEEE, Sept. 2006.

[9] P. Grover, A. Goldsmith, and A. Sahai, "Fundamental limits on complexity and power consumption in coded communication," in *extended version of paper presented at ISIT'12*, Feb. 2012,, http://www.stanford.edu/~pulkit/files/ISIT12FullProofs.pdf.

[10] P. Grover and A. Sahai, "Fundamental bounds on the interconnect complexity of decoder implementations," in *Proc. of the 45th Annual Conference on Information Sciences and Systems (CISS)*, March 2011, pp. 1 – 6.

[11] ——, "Green codes: Energy-efficient short-range communication," in *Proceedings of the 2008 IEEE Symposium on Information Theory*, Toronto, Canada, Jul. 2008.

[12] ——, "Time-division multiplexing for green broadcasting," in *Proceedings of the 2009 IEEE Symposium on Information Theory*, Seoul, South Korea, Jul. 2009.

[13] K. Ganesan, Y. Wen, P. Grover, A. Goldsmith, and J. Rabaey, "Optimizing code-decoder pairs for "green" code design," in *extended version of paper to appear in Proceedings of IEEE GLOBECOM 2012*, March 2012, http://inst.eecs.berkeley.edu/~karthikg/globecom2012full.pdf.

[14] S Cui, AJ Goldsmith and A Bahai, "Energy Constrained Modulation Optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 1–11, 2005.

[15] Y. Li, B. Bakkaloglu, and C. Chakrabarti, "A system level energy model and energy-quality evaluation for integrated transceiver front-ends," *IEEE Trans. VLSI Syst.*, vol. 15, no. 1, pp. 90–103, Jan. 2007.

[16] J. Singh, S. Ponnuru, and U. Madhow, "Multi-gigabit communication: the adc bottleneck," in *Proc. of the 2009 IEEE International Conference on Ultra-Wideband (ICUWB)*, Sept. 2009, pp. 293 –302.

[17] T. Sundstrom, B. Murmann, and C. Svensson, "Power dissipation bounds for high-speed nyquist analog-to-digital converters," *IEEE Trans. Circuits Syst. I*, vol. 56, no. 3, pp. 509 – 518, May 2009.

[18] Z. Zhang, "Design of LDPC decoders for improved low error rate performance," Ph.D. dissertation, UC Berkeley, Berkeley, CA, 2009.

[19] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.

[20] S. Chilappagari, D. Declercq, L. Sassatelli, and B. Vasic, "Two-bit message passing decoders for LDPC codes over the binary symmetric channel," in *IEEE International Symposium on Information Theory (ISIT)*, July 2009, pp. 2156 –2160.

[21] H. Friis, "A note on a simple transmission formula," *Proceedings of the IRE*, vol. 34, no. 5, pp. 254 – 256, may 1946.

[22] K. Cho and D. Yoon, "On the general BER expression of one and two-dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1074 – 1080, July 2002.

[23] Y. Wang, J. S. Yedidia, and S. C. Draper, "Construction of high-girth qc-LDPC codes," Mitsubishi Electric Research Laboratories, Tech. Rep. TR2008-061, Sept. 2008.

[24] T. Zhang and K. Parhi, "Joint (3,k)-regular ldpc code and decoder/encoder design," *IEEE Trans. Signal Processing*, vol. 52, no. 4, pp. 1065 – 1079, April 2004.

[25] I. Bocharova, F. Hug, R. Johannesson, B. Kudryashov, and R. Satyukov, "Some voltage graph-based ldpc tailbiting codes with large girth," in *Proc. of the 2011 IEEE International Symposium on Information Theory (ISIT)*, Aug. 2011, pp. 732 –736.

[26] R. Van de Plassche, *CMOS Integrated Analog-to-Digital and Digital-to-Analog Converters*. Kluwer Academic Publishers, 2010.