QC-LDPC codes for mmWave

Borja Peleato

Purdue University

Introduction

The high speeds and low latency expected in mmWave wireless communications will require strong error correcting codes and efficient decoders. Channel conditions can change widely due to factors such as beam alignment, blockage, and interference; so it will be necessary to probe the channel frequently and adjust the modulation, coding rate, and decoder to the current conditions. Furthermore, some of the common assumptions might no longer be valid, difficulting the generation and quantization of Log-Likelihood Ratios (LLR) for the decoder.

To maximize throughput, the decoder will use all the information available, even corrupted data in packets that failed to decode. We propose incremental redundancy (IR) instead of automatic repeat request (ARQ) for error control. The choice of IR bits will be based on receiver feedback when possible.

Low Density Parity Check (LDPC) codes offer good performance over a wide range of rates and efficient parallel decoding. Quasi-cyclic LDPC codes (QC-LDPC) are used in practice since they can be encoded and decoded using feedback shift-registers [1]. The quantization of the LLR values has not been studied when it is to be combined with IR and practical QC-LDPC codes. Our research focuses on the design of LLR quantization and IR schemes for QC-LDPC codes tailored towards mmWave conditions. This will require collaboration across multiple research areas: channel characterization for the LLR generation, mmWave hardware for the decoder implementation, and networking for the feedback and incremental redundancy protocols.

Background

Two main areas will be addressed in this poster: quantization and IR.

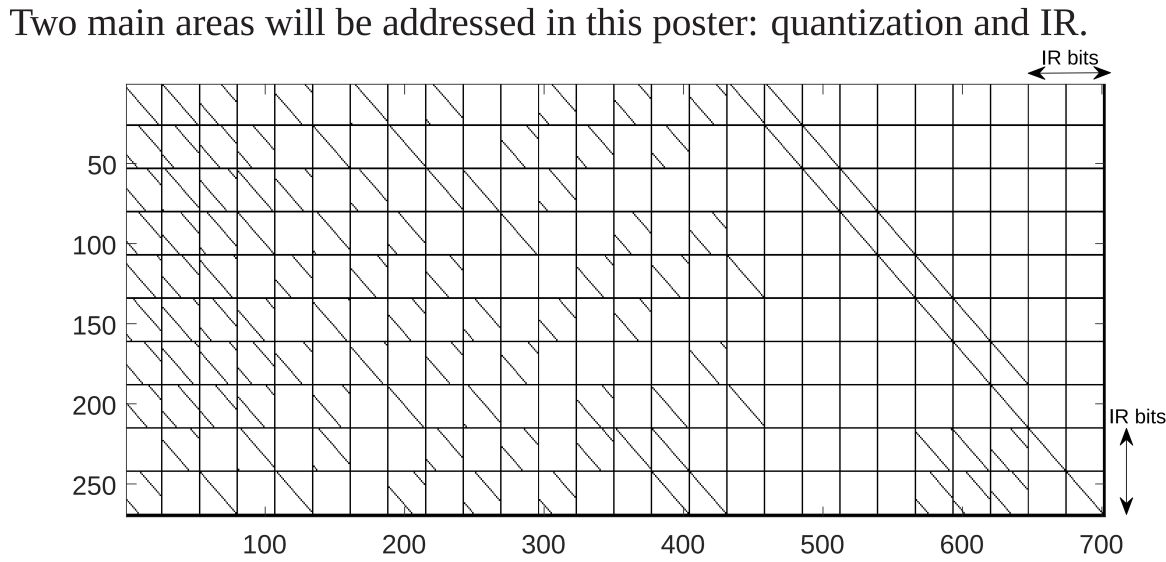


Figure 1: QC-LDPC code parity-check (H) matrix 216x648 + 54 IR.

Feedback does not increase channel capacity, with random codes of infinite length. In practice, however, feedback can help significantly. To understand this, consider a BI-AWGN channel with $\sigma_N^2 = 0.25$ and a code of length 100 capable of correcting 5 errors. Since the channel BER is 2%, we have $FER = 5\%$. When the decoding fails, the wrong bits are likely to have small LLRs. Five incremental redundancy bits will barely increase the error correction capability (Hamming bound), but if we retransmit the bits in the center region (3 bits in average), we are likely to correct the existing errors.

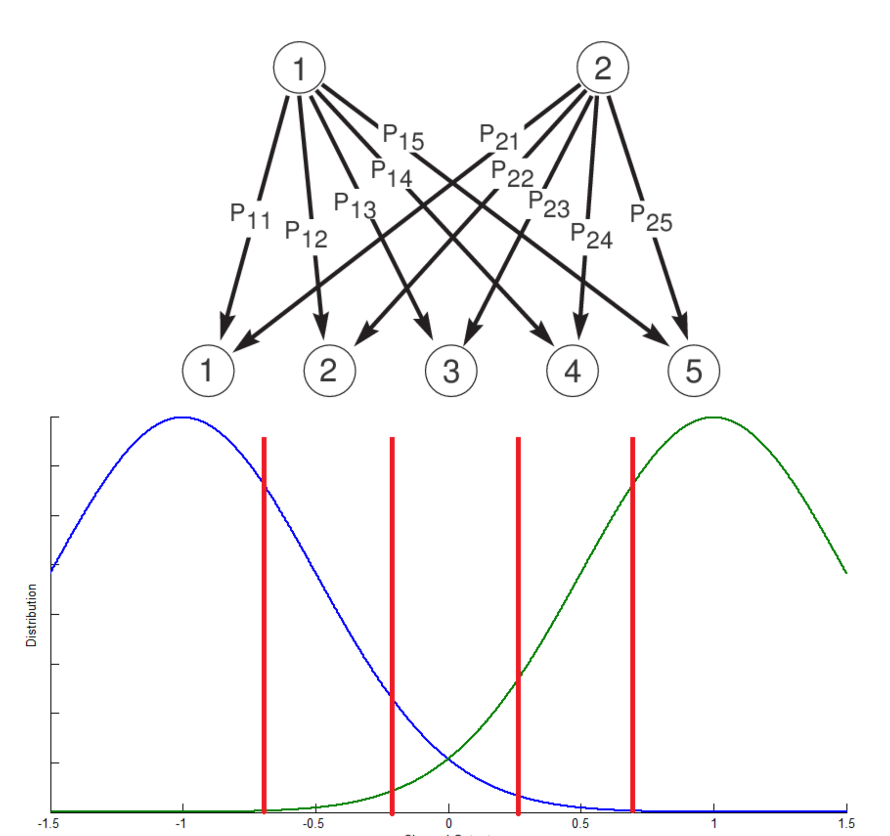


Figure 2: Quantization equivalence to a Discrete Memoryless Channel.

Abstract

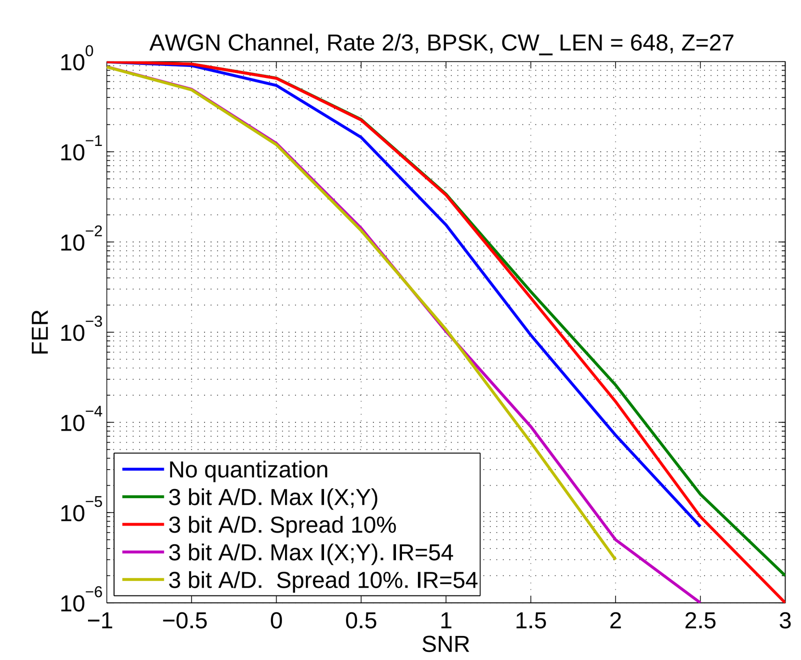
Quasi-Cyclic Low Density Parity Check (QC-LDPC) codes offer very good asymptotic performance over a wide range of code rates and efficient parallel decoding, which makes them one of the top candidates for error correction in mmWave communications. However, it is necessary to investigate how practical limitations will impact their performance, and research methods to overcome any degradation that might appear. This poster focuses on the quantization of the LLR values at the decoder input and the effect of incremental redundancy with or without feedback.

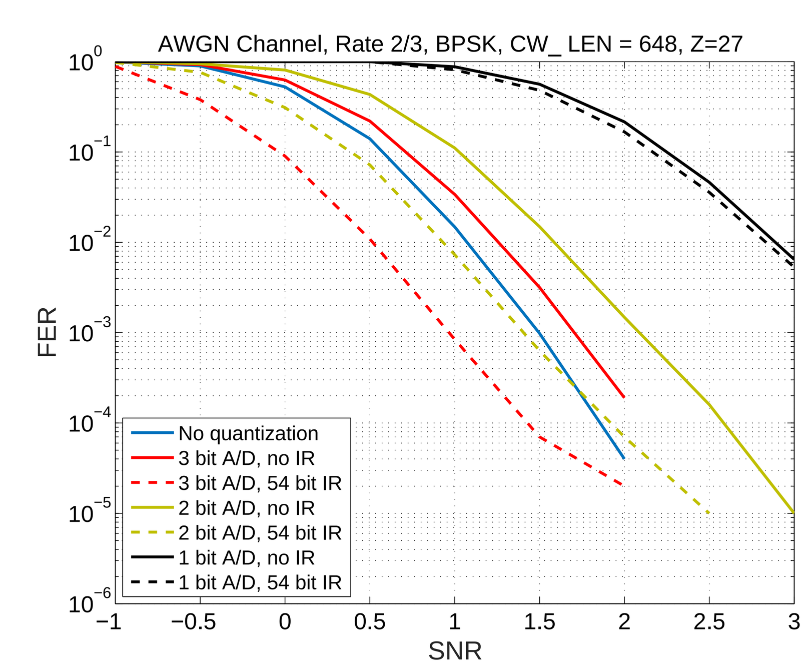
The main results in this poster are: 1) spreading the quantization thresholds beyond those maximizing the channel capacity can lower the Frame Error Rate (FER) in the high-SNR regime, 2) QC-LDPC decoders correct error bursts better than random errors, 3) Zeroing small LLRs when the decoding fails can help in the high SNR regime, 4) SNR can be estimated adaptively based on channel output distribution or number of failed checks, 5) Feedback significantly enhances the benefits of Incremental Redundancy (IR) bits.

Quantization

The quantization of LLR values needs to strike a balance between speed and precision. In theory the thresholds should maximize the capacity of the DMC, but these are not always optimal for a Min-Sum decoder with limited precision and iterations. The left plot below shows that spreading these quantization thresholds by 10% reduces FER both with and without IR. This could be related to the assignment of LLR values or with the scaling of extrinsic information in the Min-Sum decoder.

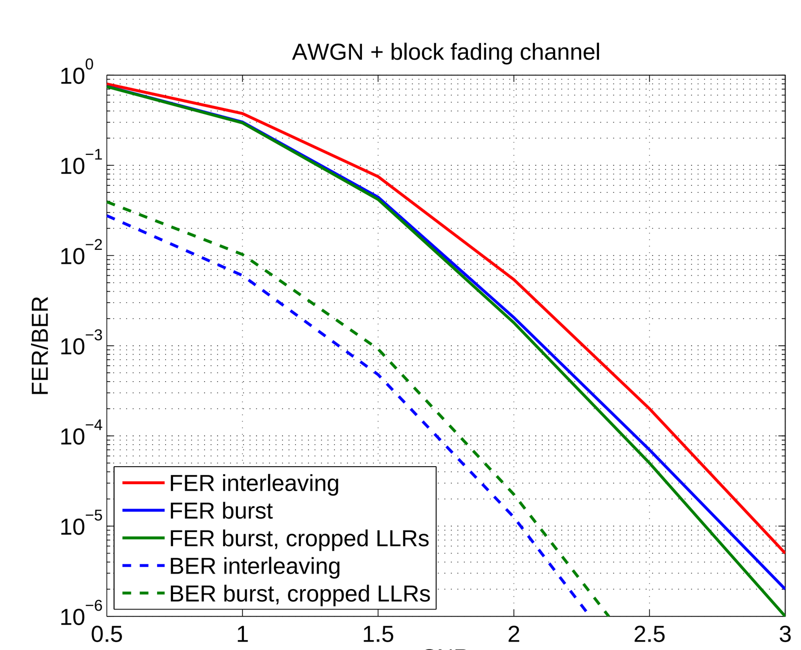
Increasing the number of quantization bits provides diminishing returns in raw BER. Incremental redundancy can help bridge the gap left by a coarse A/D, as shown in the right plot below.

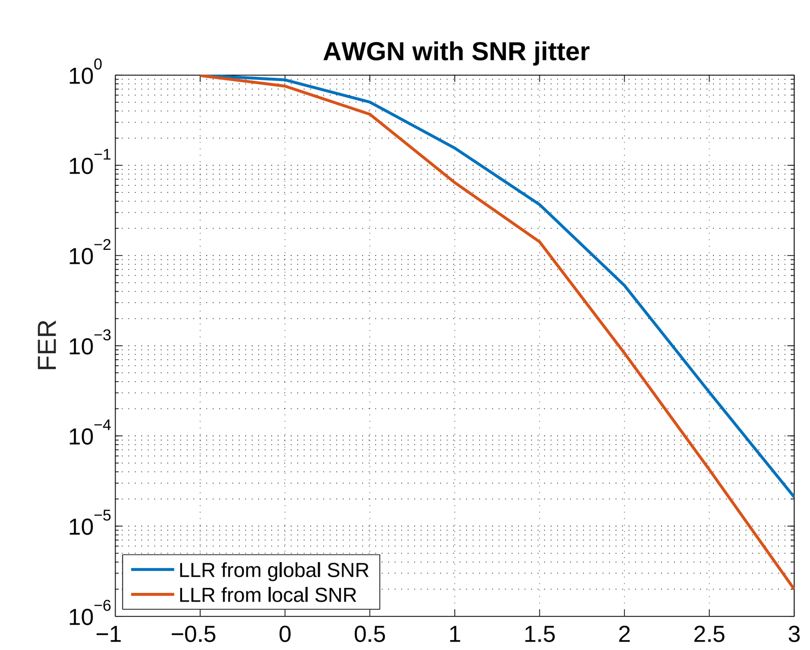




Bits in the same circulant are far in the Tanner graph, so it is better to avoid interleaving and keep error bursts in the same circulant. Zeroing small LLRs when decoding fails can also help: BER increases, but FER decreases. Message passing decoders are sensible to the sign of the LLR values, specially with few iterations. The left plot below illustrates this using an AWGN channel with 50 bit bursts of -3dB fading.

When SNR varies within a codeword, a sliding window local SNR estimation based on channel output distribution could improve performance. The right plot below shows the results for an AWGN channel where each circulant has uniform $\pm$ 2dB noise.





The quantization of the messages inside the decoder (also LLRs) has also been studied, but it goes beyond the scope of this poster. We found that 3 bits are enough to obtain reasonable performance; the performance with 4 bit messages is identical to that with floating point messages.

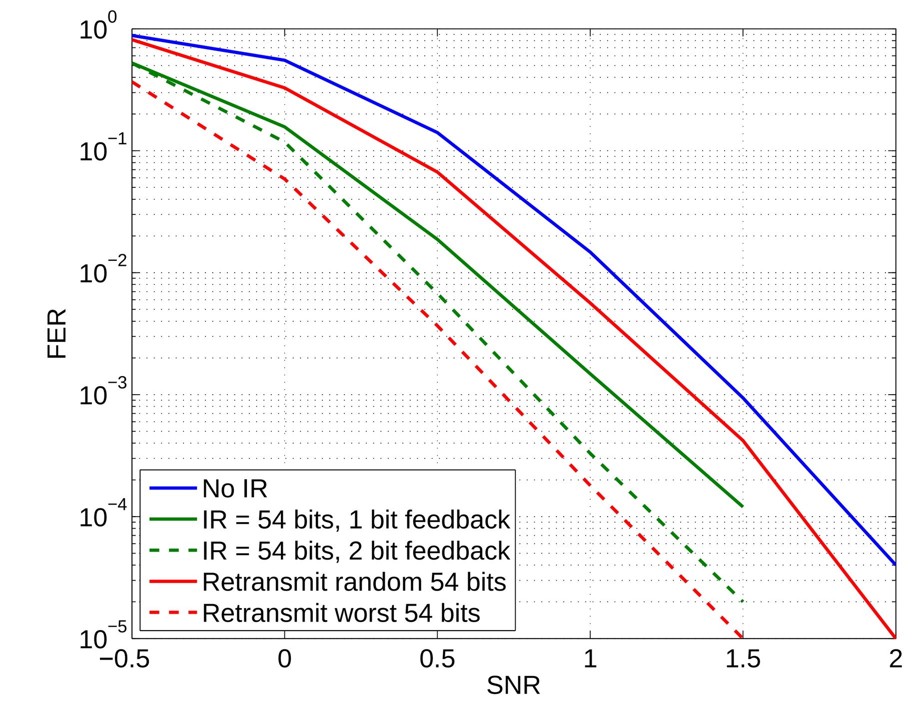
Incremental redundancy

When a codeword fails to decode, we can request additional parity bits (IR) that will aid the decoder succeed. Codes are often punctured in the first transmission [2] (then punctured bits can be sent as IR), or extended when needed [5]. We extend the H matrix as

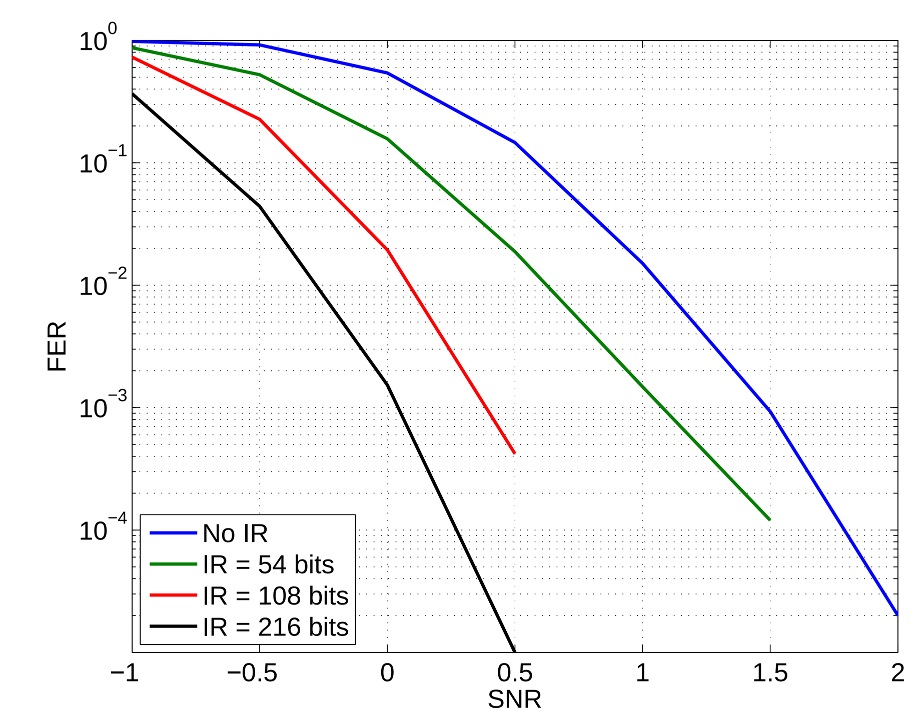
$$H_{\text{ext}} \begin{bmatrix} D & P & 0 \\ A & 0 & I \end{bmatrix} \qquad H = [D\ P],$$

where 0 is a ZxZ block of zeros, I is the identity matrix, and A is a cyclic shift of D so as to reuse the implementation used for the first transmission [4]. Other extension and puncturing techniques can be found in [3].

Under the traditional ARQ paradigm there is a single bit of feedback between the receiver and the transmitter (ACK/NACK). With more feedback, the decoder can specify which codeword segments are least reliable and the transmitter can focus the IR on those segments. This approach provides significant gains, as shown in the figure below. Retransmitting the same bits (chase combining) performs worse than sending new redundancy, unless the retransmitted bits were the least reliable in the first transmission.



The number of IR bits required depends on the codeword SNR, as shown below. It can be estimated based on the number of failed checks in H, if quantization is too coarse to use the channel outputs.



Future research

Large antenna arrays in mmWave provide narrow beams with strong gains, but beams could get misaligned and it is not possible to have fine A/D converters for each antenna element. Hence, we will seek trade-offs between beamforming, spatial multiplexing, sampling rates, and A/D quantizers.

References

[1] Dariush Divsalar, Sam Dolinar, Christopher R Jones, and Kenneth Andrews. Capacity-approaching protograph codes. *IEEE Journal on Selected Areas in Communications*, 27(6):876–888, 2009.

[2] Mostafa El-Khamy, Jilei Hou, and Naga Bhushan. Design of rate-compatible structured ldpc codes for hybrid arq applications. *IEEE Journal on Selected Areas in Communications*, 27(6):965–973, 2009.

[3] Jingjing Liu and Rodrigo C de Lamare. Rate-compatible ldpc codes with short block lengths based on puncturing and extension techniques. *AEU-International Journal of Electronics and Communications*, 69(11):1582–1589, 2015.

[4] Alcatel-Lucent Shanghai Bell Nokia. Ir harq schemes for ldpc. *3GPP TSG-RAN WG1 #86bis, Lisbon, Portugal*, 2016.

[5] Thuy Van Nguyen, Aria Nosratinia, and Dariush Divsalar. The design of rate-compatible protograph ldpc codes. *IEEE Transactions on communications*, 60(10):2841–2850, 2012.