

Noise Recycling

Alejandro Cohen*, Amit Solomon*, Ken R. Duffy†, and Muriel Médard*

*RLE, MIT Cambridge, MA 02139, USA, {cohenale,amitsol,medard}@mit.edu

†Hamilton Institute Maynooth University, Ireland, ken.duffy@mu.ie

Abstract—We introduce Noise Recycling, a method that enhances decoding performance of channels subject to correlated noise without joint decoding. The method can be used with any combination of codes, code-rates and decoding techniques. In the approach, a continuous realization of noise is estimated from a lead channel by subtracting its decoded output from its received signal. This estimate is then used to improve the accuracy of decoding of an orthogonal channel that is experiencing correlated noise. In this design, channels aid each other only through the provision of noise estimates post-decoding. In a Gauss-Markov model of correlated noise, we constructively establish that noise recycling employing a simple successive order enables higher rates than not recycling noise. Simulations illustrate noise recycling can be employed with any code and decoder, and that noise recycling shows Block Error Rate (BLER) benefits when applying the same predetermined order as used to enhance the rate region. Finally, for short codes we establish that an additional BLER improvement is possible through noise recycling with racing, where the lead channel is not pre-determined, but is chosen on the fly based on which decoder completes first.

Index Terms—Noise Recycling, Channel Decoding, Correlated Noise, Orthogonal Channels.

I. INTRODUCTION

The use of orthogonal channels is commonplace in applications from wired to wireless channels. Examples include the wide-spread use of orthogonal frequency division multiplexing (OFDM) [1], [2], and of orthogonal schemes in multiple access, such frequency division multiplexing access (FDMA), time-division multiple access (TDMA), or orthogonal code-division multiple access (CDMA), see for instance [3]. In OFDM or FDMA, channels separated by less than a coherence band will experience correlated noise. Joint decoding of all orthogonal channels can, in theory, make use of such correlation to improve performance, but it is challenging to implement efficiently in practice and, indeed, runs counter to the reason for seeking orthogonality in the first place.

As detailed in Section II, we consider a particular type of correlated noise model across orthogonal channels, Gauss-Markov (GM) [4] noise. The Gauss-Markov process has been used to model progressive decorrelation of noise with growing separation among channels [5]–[8] in time, frequency, or both. Within that model, we introduce a novel approach that embraces noise correlation to significantly improve decoding performance while maintaining separate decoding over orthogonal channels.

In interference cancellation in multiple access channels, decoded codewords are subtracted from the received signal to remove interference [3], [9]. In contrast, in noise recycling modulated decoded codewords are subtracted from the

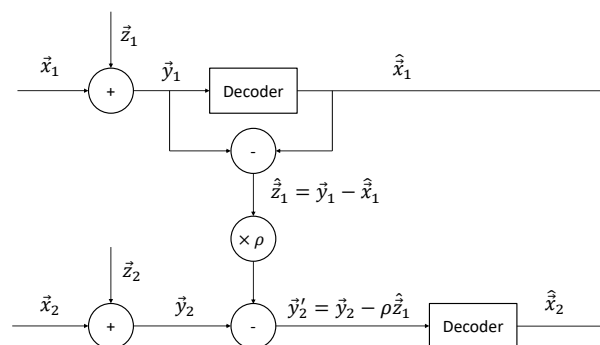


Fig. 1: In noise recycling, a noise estimate is created from a lead channel by subtracting its modulated decoding from the received signal. That estimate is used to reduce noise on a channel subject to correlated noise prior to decoding.

received signal to recover noise estimates. That noise estimate is a component of the noise in another as-yet undecoded orthogonal channel due to correlation across channels. A proportion of the estimate can, therefore, be subtracted from the received signal on the orthogonal channel before decoding, reducing the latter’s effective noise.

Figure 1 provides an illustration of the technique. It shows two independent channel inputs, (\vec{x}_1, \vec{x}_2) from potentially different codebooks, on orthogonal channels that are corrupted by correlated real-valued additive symmetric Gaussian noise (\vec{Z}_1, \vec{Z}_2) with correlation ρ . This results in correlated random real-valued channel outputs $(\vec{Y}_1, \vec{Y}_2) = (\vec{x}_1, \vec{x}_2) + (\vec{Z}_1, \vec{Z}_2)$. For a particular realization of outputs, (\vec{y}_1, \vec{y}_2) , on decoding the lead channel \vec{y}_1 to $\hat{\vec{x}}_1$, the decoder estimates the noise realization experienced on the lead channel by subtracting the decoded codeword from the received signal $\hat{\vec{z}}_1 = \vec{y}_1 - \hat{\vec{x}}_1$. The second receiver updates its channel output to $\vec{y}'_2 = \vec{y}_2 - \rho\hat{\vec{z}}_1$, eliminating part of the additive noise experienced on the second channel \vec{z}_2 , before decoding. This noise recycling results in the second channel output being a less noisy version of the channel input \vec{x}_2 , which in turn leads to improved decoding performance.

We consider both the rate gain and Block Error Rate (BLER) improvements yielded by noise recycling vis-à-vis independently decoding the channels. For rate gain, we provide a proof of achievability with an ordering for the successive decoding of orthogonal channels using noise recycling in Section III-B. We evaluate rate gains numerically, which improve both with correlation and with the number of orthogonal channels for a given correlation.

For BLER improvements, we illustrate that noise recycling can work with any codes at any rates using any decoders on any channels, since noise recycling only uses noise estimates. We consider two cases. The first, discussed in Section IV-A, is similar in spirit to the approach presented in Section III to achieve rate gains where a low rate code is reliably decoded on a lead channel, giving an accurate estimate of the noise on that channel, whereupon a second, higher rate, code can be more reliably decoded on an orthogonal channel subject to correlated noise.

The second case of BLER improvement, presented in Section IV-B, does not pre-determine which of the orthogonal channels is decoded first. Instead, the decoders of orthogonal channels are run in parallel, in effect racing each other. The first decoder to terminate provides the initial noise estimate for noise recycling. While this approach is not designed to provide rate gains, we show that it yields considerable BLER improvements for short codes. With the increasing demand for low-latency communications, high rate short codes are expected to become more common.

II. SYSTEM MODEL

A. Definition, Notation and System Model

Let x, \vec{x}, X, \vec{X} denote a scalar, vector, random variable, and random vector, respectively. All vectors are row vectors. A linear block code is characterized by a code-length, n , code-dimension, k , $[n, k]$, and rate $R = k/n$. The binary field is denoted by \mathbb{F}_2 . Mutual information between X, Y is denoted by $I(X; Y)$. We study an orthogonal channel system where $i \in \{1, \dots, m\}$ messages, $\vec{u}_i \in \mathbb{F}_2^{k_i}$, are encoded into codewords \vec{c}_i . The codewords are modulated into \vec{x}_i and sent over analog orthogonal channels subject to additive noise. Channel outputs are $\vec{Y}_i = \vec{x}_i + \vec{Z}_i$, where $Z_{i,j}$ is the j -th element of \vec{Z}_i . For each i , we call the orthogonal channels i and $i + 1$ *sequential orthogonal channels*. Noise is assumed to follow a GM model where the j -th element of the i -th noise vector \vec{Z}_i is generated in the following way: $Z_{i,j} = \rho Z_{i-1,j} + \Xi_{i,j}$ and the innovation processes, $\{\Xi_{i,j}\}_i$, are all mutually independent and identically distributed $\Xi_{i,j} \sim \mathcal{N}(0, (1 - \rho^2) \sigma^2)$, so that $Z_{i,j} \sim \mathcal{N}(0, \sigma^2)$, for $|\rho| < 1$. The rate of the i -th code is $R_i = k_i/n$, and the total rate is $R = \sum_{i=1}^m R_i$. Given $(\vec{Y}_1, \dots, \vec{Y}_m)$, the goal is to estimate $(\vec{c}_1, \dots, \vec{c}_m)$ using m distinct decoders.

III. NOISE RECYCLING RATE GAIN

We first determine the rate region that can be achieved by sending at a lower rate on one orthogonal channel, estimating the realization of the noise in that channel, and using that knowledge to reduce the impact of noise in other orthogonal channels.

Theorem III.1. *Assume a GM noise model with fixed m orthogonal channels, each with variance σ^2 , and correlation ρ between sequential orthogonal channels. For a given average*

power constraint, $\mathbb{E}(X_i^2) \leq P$, and any correlation factor $|\rho| < 1$, the following region is achievable:

$$R_1 < C\left(\frac{P}{\sigma^2}\right), \quad R_j < C\left(\frac{P}{(1 - \rho^2)\sigma^2}\right), \quad \text{for all } j > 1,$$

where $C(P/\sigma^2) = 0.5 \log(1 + P/\sigma^2)$ and so P/σ^2 is the the Signal to Noise Ratio (SNR).

A. Code construction for achievability

To establish constructively the theorem, our coding scheme carries the flavor of [9, Chapter 9.1]. We create m independent random codebooks such that the j -th codebook consists of 2^{nR_j} codewords independently drawn from $\vec{X}^j(1), \dots, \vec{X}^j(2^{nR_j}) \sim \mathcal{N}(0, P - \epsilon)$, where $R_1 < C(P/\sigma^2)$, $R_j < C(P/(1 - \rho^2)\sigma^2) \forall j \neq 1$, and a superscript j indicates that a codeword was chosen from the j -th codebook. For each channel, codebooks are known to transmitter and receiver, but each channel need only know its own codebook. The transmitters of the orthogonal channels send $\vec{X}^1(i_1), \dots, \vec{X}^m(i_m)$. The decoders operate sequentially from $j = 1$ to m as follows. With $\hat{\vec{Z}}_0 = 0$ and, $\hat{\vec{Z}}_{j-1}$ being the noise estimated from the decoding of the $j - 1$ -th channel, the j -th decoder subtracts $\rho \hat{\vec{Z}}_{j-1}$ from its channel output \vec{Y}_j , resulting in $\vec{Y}'_j = \vec{Y}_j - \rho \hat{\vec{Z}}_{j-1}$. It then identifies as the decoding the unique codeword that is jointly-typical with \vec{Y}'_j and satisfies the power constraints. If a codeword does not exist or is not unique, the j -th decoder decodes in error. In the mathematical treatment, an error occurs if any of the m decoders is in error, but in simulations it errors are accounted for separately on each channel. Figure 2 illustrates the rate that can be gained when compared to the case where channel decoders operate independently. It is evident that there is a gap between the single-channel capacity and the average rate that can be achieved by noise recycling decoders. In particular, there is a significant rate-gain even when the number of orthogonal channels, m , or the noise correlation, $|\rho|$, is low.

B. Achievability proof

Proof. We prove the result using techniques redolent of those in [9, Chapter 9.1]. We bound from below the probability that the jointly decoding $\vec{X}_1, \dots, \vec{X}_m$ successfully. The event that the first j decodings are successful is denoted by \mathcal{C}_j . The event of a decoding failure in the j -th decoder is denoted by \mathcal{E}_j . Define the following events: $E_{0,j} = \left\{ n^{-1} \sum_i (X_i^j)^2 > P \right\}$, $E_{i,j} = \left\{ \vec{X}_j(i), \vec{Y}'_j \text{ are jointly } \epsilon\text{-typical} \right\}$. Then $P(\mathcal{C}_m) = P(\mathcal{C}_1) \prod_{j=2}^m P(\mathcal{C}_j | \mathcal{C}_{j-1})$. From results for a single channel, we know that $P(\mathcal{C}_1) \geq 1 - 3\epsilon$ for n sufficiently large. Without loss of generality, assume that the j -th transmitter sends the

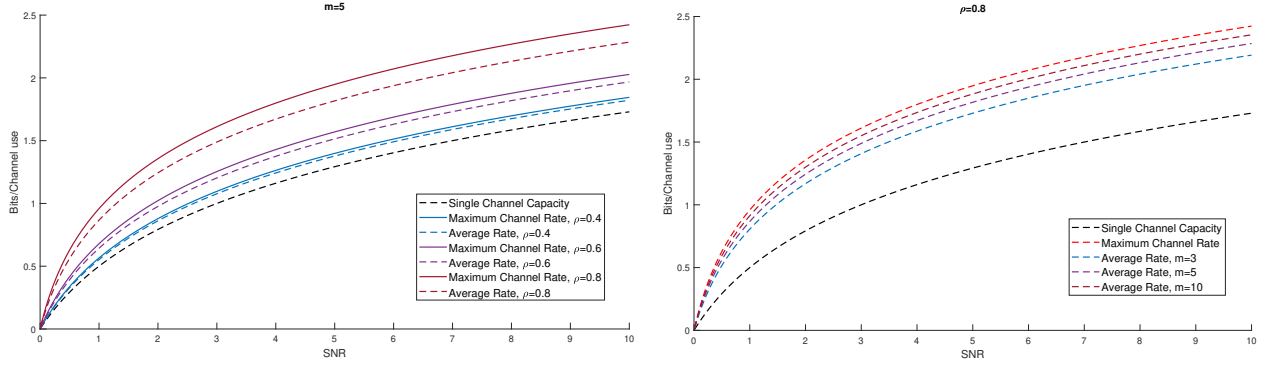


Fig. 2: Achievable regime for the decoder of Section III. Single Channel Capacity is $C_1 = C(P/\sigma^2) = 0.5 \log(1 + P/\sigma^2)$, the capacity of an orthogonal channel that does not use recycled noise, as is the case for the first channel, $j = 1$. Maximum Channel Rate is the rate of an orthogonal channel decoded with noise recycling $C_j = C(P/((1 - \rho^2)\sigma^2))$, $j > 1$. The average rate is the average rate per orthogonal channel, namely $(C_1 + (m - 1)C_2)/m$.

first codeword of the j -th codebook. We bound $P(\mathcal{E}_j | \mathcal{C}_{j-1})$ in a similar fashion to the single case:

$$P(\mathcal{E}_j | \mathcal{C}_{j-1}) \leq P(E_{0,j} | \mathcal{C}_{j-1}) + P(E_{1,j}^c | \mathcal{C}_{j-1}) + \sum_{i=2}^{2^{nR_j}} P(E_{i,j} | \mathcal{C}_{j-1}) \leq \epsilon + \epsilon + \sum_{i=2}^{2^{nR_j}} P(E_{i,j} | \mathcal{C}_{j-1}) = 2\epsilon + \sum_{i=2}^{2^{nR_j}} P(E_{i,j} | \mathcal{C}_{j-1}),$$

for n sufficiently large, where the first inequality follows from the union bound, and the second inequality from the law of large numbers and joint typicality. We bound $P(E_{i,j} | \mathcal{C}_{j-1})$, $i > 1$

$$P(E_{i,j} | \mathcal{C}_{j-1}) = P(E_{i,j} | X_1, Y_1, \dots, X_{j-1}, Y_{j-1}) = P(E_{i,j} | X_{j-1}, Y_{j-1}) \leq 2^{-n(I(X_j; Y_j | X_{j-1}, Y_{j-1}) - R_j - 3\epsilon)}$$

using the Markov property, and

$$\begin{aligned} I(X_j; Y_j | X_{j-1}, Y_{j-1}) &= I(X_j; Y_j | X_{j-1}, X_{j-1} + Z_{j-1}) = I(X_j; Y_j | Z_{j-1}) \\ &= h(X_j | Z_{j-1}) + h(Y_j | Z_{j-1}) - h(X_j, Y_j | Z_{j-1}) \\ &= h(X_j) + h(X_j + \rho Z_{j-1} + \Xi_j | Z_{j-1}) - h(X_j, X_j + \rho Z_{j-1} + \Xi_j | Z_{j-1}) \\ &= h(X_j) + h(x_j + \Xi_j) - h(X_j, X_j + \Xi_j) \\ &= I(X_j; X_j + \Xi_j) = I(X_j; Y_j') \end{aligned}$$

using the fact that $\vec{X}_j \perp \vec{Z}_{j-1}$. Therefore, $P(E_{i,j} | X_{j-1}, Y_{j-1}) \leq 2^{-n(I(X_j; Y_j') - R_j - 3\epsilon)}$. Picking $R_j < I(X_j; Y_j') - 3\epsilon$ yields $P(\mathcal{E}_{i,j} | \mathcal{C}_{j-1}) \leq 3\epsilon$. Ultimately, we get $P(\mathcal{C}_m) \geq (1 - 3\epsilon)^m$ which concludes the proof as ϵ can be made arbitrarily small. \square

IV. NOISE RECYCLING BLER IMPROVEMENT

In Section III we determined the rate-gains available from noise recycling through the use of random codebooks and joint typicality. Here we illustrate that BLER performance is enhanced by noise recycling for existing codes and decoders.

We demonstrate the technique with CRC-Aided Polar (CA-Polar) Codes [10], [11], which are Polar codes [12], [13] with an outer CRC code and will be in 5G NR control channel communications [14], and Random Linear Codes (RLCs) which are known to be capacity achieving [15], but have been little investigated owing to the absence of efficient decoders that can work at high rates until recent developments. For decoders, we use the state-of-the-art CA-Polar-specific CRC-Aided Successive Cancellation List decoder (CA-SCL) [10], [11], [16]–[18] as implemented in MATLAB's Communications Toolbox. We also use two soft-information variants [19], [20] of the recently introduced Guessing Random Additive Noise Decoder (GRAND) [21], [22], both of which can decode any block code and are well suited to short, high-rate codes. In simulations, we employ the GM channel described in Section II with Binary Phase Shift Keying (BPSK) modulation.

A. Predetermined decoding order approach

We first consider a sequential decoding scheme akin to the one described in Section III where a lead channel is decoded and a subsequent channel that has a higher rate is decoded using noise recycled information. Block-errors are counted separately on both the lead and subsequent channels as it is possible that the subsequent channel decodes correctly, even if the lead channel is in error.

In the first simulation, the lead channel encodes its data using a CA-Polar code [256, 170] with rate $R_1 \approx 2/3$. The second, orthogonal channel uses a higher rate CA-Polar code, either [256, 180] or [256, 190] giving $R_2 \approx 0.7$ or 0.74 respectively. The noise correlation is set to $\rho = 0.5$ and is known to the second decoder. Both channels are decoded with CA-SCL, with the second channel benefiting from noise recycling, where before decoding the noise estimate of the lead orthogonal channel is subtracted after multiplication by the noise correlation factor ρ .

Figure 3 reports BLER vs E_b/N_0 . The black dashed line corresponds to the lead channel, while the dashed blue and red lines give the performance curves should noise recycling not be used, corresponding to independent decoding of all

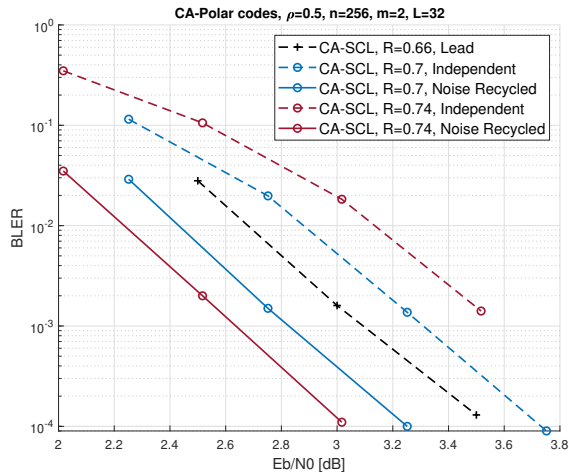


Fig. 3: BLER vs. E_b/N_0 for 256 bit CA-Polar codes decoded with CA-SCL, uses a list size of $L = 32$, with and without noise recycling. Dashed lines correspond to independent decoding, and solid lines to decoding after noise recycling. The lead orthogonal channel is encoded with a rate $2/3$ code. The second channel uses either a rate 0.7 or 0.74 code.

channels. As the second orthogonal channel runs at a higher rate than the lead channel, if decoded independently the second channel would experience higher BLER than the lead channel. The solid blue and red lines report the performance of the second decoder given noise recycling. Despite using a higher rate code than the lead channel, with noise recycling the second channel experiences better BLER vs E_b/N_0 performance. Notably, owing to the better E_b/N_0 (i.e. the energy per **information** bit used in the transmission) that comes from running a higher rate code, the rate 0.74 code provides better BLER than the rate 0.7 code. For a commonly used target BLER of 10^{-2} , noise recycling results in ≈ 1 dB gain for the [256, 190] code.

Figure 4 reports an analogous simulation, but where $\rho = 0.8$, the lead channel's code is a [128, 105] CA-Polar code, $R_1 = 0.82$, and the second channel is one of three RLCs ranging in rate from 0.85 to 0.98. Both channels are decoded with the recently proposed Ordered Reliability Bits Guessing Random Additive Noise Decoding (ORBGRAND) [20]. ORBGRAND is a soft detection decoder that has been reported to provide more accurate decodings of CA-Polar codes than CA-SCL for short codes. As with all the GRAND algorithms, it can decode any code, making it viable for use with RLCs. A similar phenomenology to the previous figure can be seen, where the impact of noise recycling is even more dramatic, allowing the second channel code to use reliably a much higher rate than the lead channel.

B. Additional gains for short codes by racing

While previous sections identified rate and BLER improvements that are available from running a pre-determined lead channel with a lower rate code so that an accurate inference of a noise realization could be obtained to aid the signal at

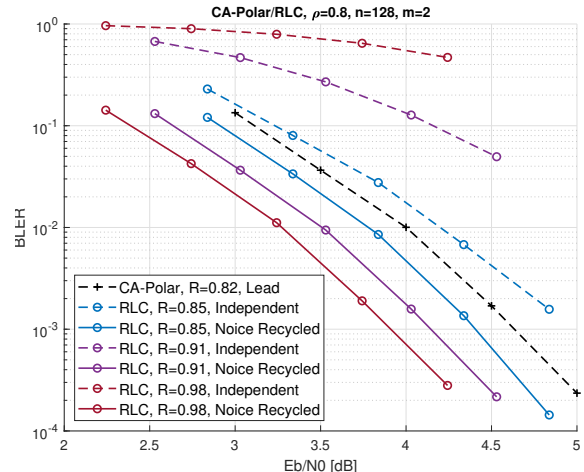


Fig. 4: BLER vs. E_b/N_0 for codes of length $n = 128$ decoded with ORBGRAND with and without noise recycling. Dashed lines correspond to independent decoding, and solid lines to decoding after noise recycling. Data on the lead orthogonal channel is encoded with a rate 0.82 CA-Polar code. The second channel uses rate 0.85, 0.91 or 0.98 RLCs.

a higher rate second channel, here we consider an alternate design that leverages a significant effect only observable with short codes. The principle behind noise recycling with racing is that all channels use the same code-rate and orthogonal channels initially attempt to decode their outputs contemporaneously. For certain types of decoder, speed of decoding provides a measure of confidence in the decoding accuracy and hence the precision of the noise estimate. Thus once one decoder has identified a codeword, it is determined to be the lead channel, other decoders cease their decodings, remove the recycled noise estimate from the lead channel from their received signal and decode. Noise recycling continues until all orthogonal channel outputs have been decoded.

This decoding procedure is described in Algorithm 1. For example, suppose there are $m = 3$ orthogonal channels. At the first step, all decode in parallel. If decoder 2 is the first to finish decoding, it is declared the winner of the race. The winner acts as lead channel and provides an estimate \hat{z}_2 to the decoders 1 and 3, which repeat the process. Mixing-and-matching of decoders, even at different stages of the race, is still possible. This offers, for example, the possibility of using at the race phase a decoder that is highly accurate, but with potentially poor in worst-case runtime. As the race winner will terminate generally quickly, a decoder with uncertain termination time may not necessarily be deleterious in the race phase. Substituting a different deciding algorithm after noise recycling is then a possibility.

We demonstrate the race approach using a recently proposed technique, Soft GRAND with ABandonment (SGRANDAB) [19], that has the required feature. SGRANDAB aims to identify the noise that corrupted a transmission from which the codeword can be inferred, rather than identifying the codeword directly. It does this by

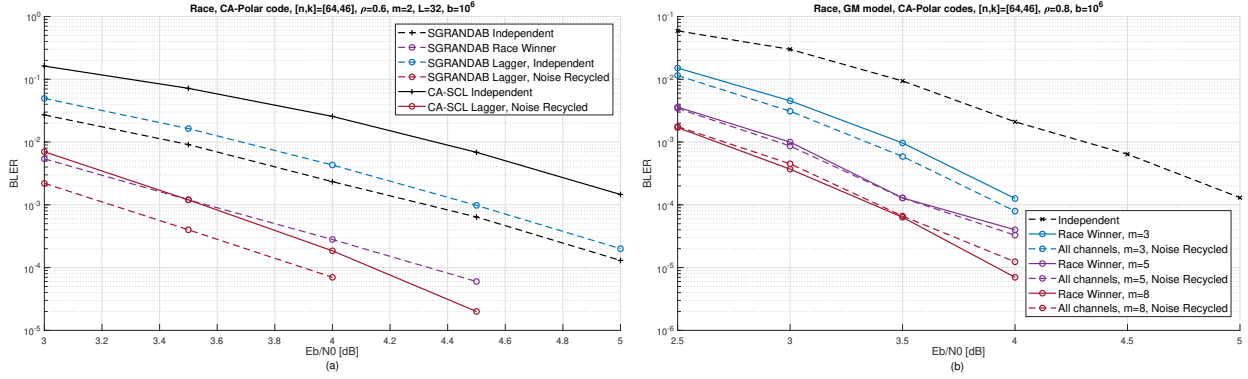


Fig. 5: BLER vs E_b/N_0 for $[64, 46]$ CA-Polar codes with racing. SGRANDAB, which uses an abandonment threshold of $b = 10^6$, is used for the race. In (a), either SGRANDAB or CA-SCL, which uses a list size of $L = 32$, are used to decode the lagggers after noise recycling. In (b), all decoding is done using SGRANDAB.

removing possible noise effects, from most likely to least likely as determined by soft information, from a received signal and querying whether what remains is in the codebook. The first instance that results in success is a maximum likelihood decoding. If no codeword is found before a given number of codebook queries, SGRANDAB abandons decoding and reports an error. The channel that is decoded with the fewest codebook queries by SGRAND wins the noise recycling race.

Algorithm 1 Noise recycling with racing

Input: $\vec{y}_1, \dots, \vec{y}_m$
Output: $\hat{c}_1, \dots, \hat{c}_m$

```

1: Decode orthogonal channel outputs
2:  $i \leftarrow$  # of decoder that won the decoding race
3:  $\hat{c}_i \leftarrow$   $i$ -th decoded codeword
4:  $\hat{x}_i \leftarrow$  encode of  $\hat{c}_i$ 
5:  $\hat{z}_i \leftarrow \vec{y}_i - \hat{x}_i$ 
6: for  $j = 1 \rightarrow \max\{m - i, i - 1\}$  do
7:   if  $i + j \leq m$  then
8:      $[\hat{c}_{i+j}, \hat{z}_{i+j}] \leftarrow$  DecodeAndEst.  $(\vec{y}_{i+j}, \hat{z}_{i+j-1})$ 
9:   end if
10:  if  $i - j \geq 1$  then
11:     $[\hat{c}_{i-j}, \hat{z}_{i-j}] \leftarrow$  DecodeAndEst.  $(\vec{y}_{i-j}, \hat{z}_{i-j+1})$ 
12:  end if
13: end for
14: return  $\hat{c}_1, \dots, \hat{c}_m$ 

procedure DECODEANDEST.  $(\vec{y}, \hat{z}, j)$ 
   $\hat{y} \leftarrow \vec{y} - \rho \hat{z}$ 
  Decode orthogonal channel  $j$  using  $\hat{y}$ 
   $\hat{c} \leftarrow$  decoded codeword
   $\hat{x} \leftarrow$  encode of  $\hat{c}$ 
   $\hat{z} \leftarrow \vec{y} - \hat{x}$ 
  return  $\hat{c}, \hat{z}$ 
end procedure

```

We simulated the noise recycling race in the presence of m GM channels using a $[64, 46]$ CA-Polar code. We first

consider the method on $m = 2$ orthogonal channels with $\rho = 0.6$, where the race, is done using SGRANDAB on all orthogonal channels, and the noise-recycled decoding is performed using either SGRANDAB or CA-SCL. Figure 5 (a) reports BLER performance. SGRANDAB outperforms CA-SCL without noise recycling, so it provides a better noise estimate following the decoding of the orthogonal channel that wins the race. The lagging channel benefits from noise recycling. It can continue to use SGRANDAB as during the race, or instead use CA-SCL. These results again shows a gain of more than 1 dB can be achieved, even for codes of the same rate, by racing noise recycling. Figure 5 (b) reports the BLER of a decoder without noise recycling for $\rho = 0.8$ and $m = 3, 5$ or 8 , the race winner with noise recycling, and all m decoders after noise recycling, where all decoders use SGRANDAB. This shows a significant improvement in BLER for all values of m . For example there is a gain of about 1.7 dB for $m = 8$ at a target BLER of 10^{-4} . While this race advantage disappears as a consequence of averaging for long codes, it can be seen to provide a significant advantage for short codes.

V. CONCLUSION AND DISCUSSION

We presented a novel way to recycle noise in orthogonal channels in order to improve communication performance for any combination of codes. The performance improvement is twofold, we proved its rate gain aspect and showed empirical evidence of its reliability improvement aspect. We analyzed orthogonal correlated channels, i.e. channels in which data that is sent on different channels is independent. A natural extension is considering the use of noise recycling in wireless communications, and the consequences of uncertainty in it. Noise recycling points to the benefit of correlation among orthogonal channels, opening an interesting vein of investigation where orthogonal channels, say in OFDM or TDMA, are chosen with a preference for noise correlation among them, with attendant effects in terms of rate and power allocation among orthogonal channels.

REFERENCES

- [1] R. v. Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*, 1st ed. USA: Artech House, Inc., 2000.
- [2] Hongwei Yang, "A road to future broadband wireless access: MIMO-OFDM-based air interface," *IEEE Commun. Mag.*, vol. 43, no. 1, pp. 53–60, Jan 2005.
- [3] R. Gallager, "A perspective on multiaccess channels," *IEEE Tran. Inf. Theory*, vol. 31, no. 2, pp. 124–142, March 1985.
- [4] B. D. Anderson and J. B. Moore, *Optimal filtering*. Courier Corporation, 2012.
- [5] I. Nevat and J. Yuan, "Channel tracking using pruning for mimo-ofdm systems over gauss-markov channels," in *ICASSP*, vol. 3, 2007, pp. 193–196.
- [6] I. Abou-Faycal, M. Médard, and U. Madhow, "Binary adaptive coded pilot symbol assisted modulation over rayleigh fading channels without feedback," *IEEE Tran. Commun.*, vol. 53, no. 6, pp. 1036–1046, 2005.
- [7] M. Médard and R. G. Gallager, "Bandwidth scaling for fading multipath channels," *IEEE Tran. Inf. Theory*, vol. 48, no. 4, pp. 840–852, 2002.
- [8] S. Akin and M. C. Gursoy, "Training optimization for Gauss-Markov Rayleigh fading channels," in *IEEE ICC*, 2007, pp. 5999–6004.
- [9] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
- [10] I. Tal and A. Vardy, "List decoding of polar codes," in *IEEE Int. Symp. Inf. Theory*, 2011, pp. 1–5.
- [11] —, "List decoding of polar codes," *IEEE Tran. Inf. Theory*, vol. 61, no. 5, pp. 2213–2226, 2015.
- [12] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes," in *IEEE Int. Symp. Inf. Theory*, 2008, pp. 1173–1177.
- [13] —, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Tran. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.
- [14] "3GPP TS 38.212, 5G-NR-Multiplexing and channel coding," https://www.etsi.org/deliver/etsi_ts/138200_138299/138212/15.02.00_60/ts_138212v150200p.pdf, [Online].
- [15] R. Gallager, "The random coding bound is tight for the average code (corresp.)," *IEEE Tran. Inf. Theory*, vol. 19, no. 2, pp. 244–246, 1973.
- [16] K. Niu and K. Chen, "CRC-aided decoding of polar codes," *IEEE Commun. Letters*, vol. 16, no. 10, pp. 1668–1671, 2012.
- [17] A. Balatsoukas-Stimming, M. B. Parizi, and A. Burg, "LLR-based successive cancellation list decoding of Polar codes," *IEEE Trans. Signal Process.*, vol. 63, no. 19, pp. 5165–5179, 2015.
- [18] X. Liang, J. Yang, C. Zhang, W. Song, and X. You, "Hardware efficient and low-latency ca-scl decoder based on distributed sorting," in *IEEE GLOBECOM*, 2016, pp. 1–6.
- [19] A. Solomon, K. R. Duffy, and M. Médard, "Soft Maximum Likelihood Decoding using GRAND," *preprint arXiv:2001.03089*, 2020.
- [20] K. R. Duffy, "Ordered Reliability Bits Guessing Random Additive Noise Decoding," *preprint arXiv:2001.00546*, 2020.
- [21] K. R. Duffy, J. Li, and M. Médard, "Guessing noise, not code-words," in *IEEE Int. Symp. Inf. Theory*, 2018, pp. 671–675.
- [22] —, "Capacity-achieving Guessing Random Additive Noise Decoding," *IEEE Tran. Inf. Theory*, vol. 65, no. 7, pp. 4023–4040, 2019.