Achieving the Capacity of any DMC using only Polar Codes

David Sutter¹, Joseph M. Renes², Frédéric Dupuis², and Renato Renner²

1) Department for Information Technology and Electrical Engineering, ETH Zurich, Switzerland, suhdadav@student.ethz.ch
2) Institute for Theoretical Physics, ETH Zurich, Switzerland, {renes,.dupuis,renner}@phys.ethz.ch

Abstract
We construct a channel coding scheme to achieve the capacity of any discrete memoryless channel based solely on the techniques of polar coding. In particular, we show how source polarization and randomness extraction via polarization can be employed to “shape” uniformly-distributed i.i.d. random variables into approximate i.i.d. random variables distributed according to the capacity-achieving distribution. An application to the AWGN channel is discussed.

Background and Results
Polar codes achieve the symmetric capacity of any DMC W having O(N log N) encoding and decoding complexity, where N denotes the blocklength [1].

Using Gallager’s shaper idea, polar codes can be used to achieve the capacity of any DMC at the cost of a slightly higher encoding and decoding complexity [2].

Polar Coding Scheme
Construct the shaper using the Arıkan transform G_L; combination with W gives the super-channel W_{KL}: 

\[ \hat{U}_K \quad \tilde{X}_L \quad W_{KL} \quad Y_L \]

Outer layer: Polar codes are used to provide reliable communication over W_{KL}, by using the multilevel coding method.

Inner layer: Polarization is used to shape the uniform inputs from the outer layer into a good approximation to X_L for transmission over W.

Capacity, Reliability & Efficiency
The rate R of the above scheme is given by:

\[ R := \lim_{L \to \infty} \frac{1}{L} \mathbb{E} \left[ |\delta X| - H(U^{L}|\delta X)|Y^L) \right] \]

Now make use of the polarization phenomenon. Define the two ordered sets [3]

\[ \mathcal{R}_c := \{ i \in [L] : H(U^{i-1}|U^{i-1}) > 1 - \epsilon \} \]

\[ \mathcal{D}_c := \{ i \in [L] : H(U^{i-1}|U^{i-1}) < \epsilon \} \]

for which it can be shown that \(|\mathcal{R}_c| = LH(X) - o(L)\) and \(\epsilon = O(L^{-\gamma})\) with \(\gamma < 1/2\). Choosing \(\delta K = \mathcal{R}_c\) with \(K = |\mathcal{R}_c|\) we have for \(\beta < 1/2\)

\[ P_{err} = O(L^{-\beta}) \]

Theorem: Reliability
The error probability of the coding scheme satisfies

\[ P_{err} = O(L^{2-\beta M} + L^{2-\beta \epsilon}) \]

for \(\beta, \beta' > 1/2\). Derandomizing the shaper leads an error probability

\[ P'_{err} = P_{err} \left( 1 + O \left( L \left( 1 - 2^{-\epsilon \gamma} \right) \right) \right) \]

Proposition: Efficiency
The encoder and decoder have both complexity \(O(N \log N)\).

AWGN Capacity Approximation
For Gallager’s method we use a dyadic m-term approximation [4], whereas for the new scheme we can use a general m-term approximation which can be for example a Gauss quadrature [5].

Convergence Rates
(i) Using Gallager’s method we achieve an exponential convergence rate

\[ \text{SNR}^2 \rightarrow m. \]

(ii) Using the new scheme with Gauss quadrature we achieve a double-exponential convergence rate

\[ 4(1 + \text{SNR}^2) \left( \frac{\text{SNR}}{1 + \text{SNR}} \right)^{2^{-\gamma}}. \]

References

David Sutter, Joseph M. Renes, Frédéric Dupuis, and Renato Renner, ITW Lausanne, September 3-7, 2012