Attaining Capacity with iterated $(U|U+V)$ codes based on AG codes and Koetter-Vardy soft decoding

Irene Marquez-Corbella, Jean-Pierre Tillich

Attaining Capacity with iterated $(U|U+V)$ codes based on AG codes and Koetter-Vardy soft decoding

Irene Márquez-Corbella
Dept. of Mathematics, Statistics and O. R.,
University of La Laguna, 38200, Spain
Email: irene.marquez.corbella@ull.es

Jean-Pierre Tillich
Inria, Paris 75012, France
Email: jean-pierre.tillich@inria.fr

***Abstract*—In this paper we show how to attain the capacity of discrete symmetric channels with polynomial time decoding complexity by considering iterated $(U \mid U+V)$ constructions with algebraic geometry (AG) code components. These codes are decoded with a recursive computation of the *a posteriori* probabilities of the code symbols together with decoding the AG components with the Koetter-Vardy algorithm. We show that, when the number of levels of the iterated $(U \mid U+V)$ construction tends to infinity, we attain the capacity of any discrete symmetric channel. Moreover the error probability decays quasi-exponentially with the codelength in the case of Reed-Solomon code constituents and exponentially with Tsfasman-Vlăduts-Zink code constituents.**

I. Introduction

The purpose of this paper is to explore a coding/decoding strategy which mixes polar codes and algebraic geometry (AG) codes decoded with the Koetter-Vardy soft decoder [3]. The codes we deal with here are iterated $(U \mid U+V)$ codes with AG codes, meaning a $(U \mid U+V)$ code where U and V are themselves $(U \mid U+V)$ codes up to some depth, and the indecomposable codes at the last level are then AG codes. From now on a linear code of length n, dimension k and distance d over a finite field $\mathbb{F}_q$ is referred to as an $[n,k,d]_q$ code. We first recall the definition of a $(U \mid U+V)$ code.

Definition 1 ($(U \mid U+V)$ code)**.** Let U be an $[n,k_U,d_U]_q$ code and let V be an $[n,k_V,d_V]_q$ code. We define the $(U \mid U+V)$-construction as the $[2n,k,d]_q$ code:

$$(U \mid U+V) = \{(\mathbf{u} \mid \mathbf{u}+\mathbf{v}); \mathbf{u} \in U \text{ and } \mathbf{v} \in V\}$$

with $k = k_U + k_V$ and $d = \min(2d_U, d_V)$.

The codes we are interested in are then defined by

Definition 2 (iterated $(U \mid U+V)$-construction of depth ℓ)**.** An iterated $(U \mid U+V)$-code U_ϵ of depth ℓ is defined from a set of 2^ℓ codes $\left\{U_{\mathbf{x}}; \mathbf{x} \in \{0,1\}^\ell\right\}$ which have all the same length and are defined over the same finite field $\mathbb{F}_q$ by using the recursive definition

$$U_\epsilon \stackrel{\text{def}}{=} (U_0 \mid U_0 + U_1)$$
$$U_{\mathbf{x}} \stackrel{\text{def}}{=} (U_{\mathbf{x}|0} \mid U_{\mathbf{x}|0} + U_{\mathbf{x}|1}) \;\; \mathbf{x} \in \{0,1\}^i, i \in \{1,\ldots,\ell-1\}$$

The codes $U_{\mathbf{x}}$ for $\mathbf{x} \in \{0,1\}^\ell$ are called the *constituent codes* of the construction.

These codes can be decoded by first computing recursively the *a posteriori probabilities* of each symbol of the constituent code in the same way as for polar codes and then using a soft decoder for the constituent codes. Indeed, the point of choosing the constituent codes to be AG codes is that they admit a soft decoder that has polynomial complexity, namely the Koetter-Vardy decoder. The difference with polar codes is that at the last level we decode a whole code and not just a symbol as for polar codes.

When the depth of this iterative $(U \mid U+V)$-construction tends to infinity we have the same phenomenon as for polar codes, namely the channels faced by the constituent codes polarize: they become either very noisy channels or very clean channels of capacity close to 1. The polarization phenomenon together with a result proving that the Koetter-Vardy decoder is able to operate sucessfully at rates close to 1 for channels of capacity close to 1 can be used to show that it is possible to attain the capacity of symmetric channels by choosing appropriately the rates of the constituent AG codes. We have in our case an additional freedom when compared to polar codes, we can indeed choose freely the length of the constituent codes. Furthermore, we can use Reed-Solomon codes by grouping symbols together and viewing them as symbols living in an extension field.

This allows to choose long enough constituent codes and by choosing the length/type/and alphabet size of the constituent codes appropriately we obtain a very sharp decay of the probability of error after decoding. We will indeed show that if we insist on using Reed-Solomon codes in the code construction we obtain a quasi-exponential decay of the probability of error in terms of the codelength (i.e. exponential if we forget about the logarithmc terms in the exponent) and an exponential decay if we use the right AG codes (namely Tsfasman-Vlăduts-Zink codes [9]). This improves very significantly upon polar codes. In essence, this sharp decay of the probability of error after decoding is due to a result of this paper (see Theorems 5 and 6) showing that even if the Koetter-Vardy decoder is not able to attain the capacity with a probability of error going to zero as the codelength goes to infinity its probability of error decays like $2^{-K\epsilon^2 n}$ where n is the codelength and ϵ is the difference between a quantity which is strictly smaller than the capacity of the channel and the code-rate.

Note: All the results of this paper are given without proofs.

They can be found in the full version [6].

Notation. Throughout the paper we will use the following notation.

• For a vector $\mathbf{x}$ we either denote by $x(i)$ or by x_i the i-th coordinate of $\mathbf{x}$. We use the first notation when the subscript is already used for other purposes or when there is already a superscript for $\mathbf{x}$.

• By some abuse of terminology, we also view a discrete memoryless channel W with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ as an $\mathcal{X} \times \mathcal{Y}$ matrix whose (x, y) entry is denoted by $W(y|x)$ which is defined as the probability of receiving y given that x was sent.

II. Link with polar codes

Recursive soft decoding of an iterated $(U \mid U+V)$-code. As explained in the introduction, our approach is to use the same decoding strategy as for Arıkan polar codes (that is his successive cancellation decoder) but by using now constituent codes which are much longer than single symbols. This decoder was actually considered before the invention of polar codes and has been considered for decoding for instance Reed-Muller codes based on the fact that they are $(U \mid U+V)$ codes [2].

Let us recall how such a $(U \mid U+V)$-decoder works. Suppose we transmit the codeword $(\mathbf{u} \mid \mathbf{u}+\mathbf{v}) \in (U \mid U+V)$ over a noisy channel and we receive the vector: $\mathbf{y} = (\mathbf{y}_1 \mid \mathbf{y}_2)$. We denote by $p(b \mid a)$ the probability of receiving b when a was sent and assume a memoryless channel here. We also assume that all the codeword symbols u_i and v_i are uniformly distributed.

Step 1: We first decode V. We compute the probabilities

$$\text{prob}(v_i{=}\alpha|y_1(i), y_2(i)){=}\sum_{\beta\in\mathbb{F}_q} p(y_1(i)|\beta)p(y_2(i)|\alpha+\beta)$$

for all positions i and all α in $\mathbb{F}_q$.

Step 2: We then use Arıkan's successive decoding approach and assume that the V decoder was correct and thus we have recovered $\mathbf{v}$. We then compute the probabilities

$$\text{prob}(u_i{=}\alpha|y_1(i), y_2(i), v(i)){=}\frac{p(y_1(i)|\alpha)p(y_2(i) \mid \alpha+v_i)}{\sum_{\beta\in\mathbb{F}_q} p(y_1(i)|\beta)p(y_2(i)|\beta+v_i)}$$

for all $\alpha \in \mathbb{F}_q$ and all coordinates i. This can be considered as soft-information on $\mathbf{u}$ which can be used by a soft information decoder for U.

This decoder can then be used recursively for decoding an iterated $(U \mid U+V)$-code. When the constituent codes of this recursive $(U \mid U+V)$ construction are just codes of length 1, it is readily seen that this decoding simply amounts to the successive cancellation decoder of Arıkan. But, we will be interested in the case where these constituent codes are longer than this: we will use as constituent codes, codes for which we have an efficient but possibly suboptimal decoder which can make use of soft information. Reed-Solomon codes or AG codes with the Koetter Vardy decoder are precisely codes with this kind of property.

Polarization. The probability computations made during the $(U \mid U+V)$ decoding correspond in a natural way to changing the channel model for the codes U and V. These two channels correspond to the two channel combining models considered for polar codes. More precisely, if we consider a memoryless channel of input alphabet $\mathbb{F}_q$ and output alphabet $\mathcal{Y}$ defined by a transition matrix $W = (W(y|u))_{\substack{u\in\mathbb{F}_q \\ y\in\mathbb{F}_q}}$, then the channel viewed by the U decoder, respectively the V decoder is a memoryless channel with transition matrix W^0 and W^1 respectively, which are given by

$$\begin{aligned} W^0(y_1, y_2, u_2|u_1) &\stackrel{\text{def}}{=} \frac{1}{q}W(y_1|u_1)W(y_2|u_1\oplus u_2) \\ W^1(y_1, y_2|u_2) &\stackrel{\text{def}}{=} \frac{1}{q}\sum_{u_1\in\mathbb{F}_q} W(y_1|u_1)W(y_2|u_1\oplus u_2) \end{aligned}$$

Here the y_i''s belong to $\mathcal{Y}$ and the u_i's belong to $\mathbb{F}_q$.

If we define the channel W^x for $x = (x_1 \dots x_n) \in \{0,1\}^n$ recursively by $W^{x_1\dots x_{n-1}x_n} = (W^{x_1\dots x_{n-1}})^{x_n}$ then the channel viewed by the decoder for one of the constituent codes $U_{x_1\dots x_n}$ of an iterated $(U \mid U+V)$ code of depth n (with the notation of Definition 2) is nothing but the channel $W^{x_1\dots x_n}$.

The key result used for showing that polar codes attain the capacity is that these channels polarize in the following sense

Theorem 1 ([8, Theorem 1] and [7, Theorem 4.10])**.** *Let q be an arbitrary prime. Then for a discrete q-ary input channel W of symmetric capacity [1] C we have for all $0 < \beta < \frac{1}{2}$*

$$\lim_{\ell\to\infty} \frac{1}{n}\left|i \in \{0,1\}^\ell : \mathcal{Z}\left(W^i\right) \le 2^{-n^\beta}\right| = C,$$

where $n \stackrel{def}{=} 2^\ell$ and $\mathcal{Z}(W)$ denotes the Bhattacharyya of W[2].

III. Soft decoding of AG codes with the Koetter-Vardy decoding algorithm

It has been a long standing open problem to obtain an efficient soft-decision decoding algorithm for Reed-Solomon codes until Koetter and Vardy showed in [3] how to modify appropriately the Guruswami-Sudan decoding algorithm in order to achieve this purpose. The complexity of this algorithm is polynomial and we will show here that the probability of error decreases exponentially in the codelength when the noise level is below a certain threshold. Let us first review a few basic facts about this decoding algorithm.

The reliability matrix. The Koetter-Vardy decoder [3] is based on a *reliability matrix* $\mathbf{\Pi}_\mathbf{y}$ of the codeword symbols $x_1, \dots, x_n$ computed from the knowledge of the received word $\mathbf{y}$ and which is defined by

$$\mathbf{\Pi}_\mathbf{y} \stackrel{\text{def}}{=} \left(\text{prob}(x_j = \alpha|y_j)\right)_{\substack{\alpha\in\mathbb{F}_q \\ 1\le j\le n}}.$$

[1] Recall that the symmetric capacity of such a channel is defined as the mutual information between a uniform input and the corresponding output of the channel, that is $C \stackrel{\text{def}}{=} \frac{1}{q}\sum_{\alpha\in\mathbb{F}_q}\sum_{y\in\mathcal{Y}} W(y|\alpha)\log_q \frac{W(y|\alpha)}{\sum_{\beta\in\mathbb{F}_q}\frac{1}{q}W(y|\beta)}$, where $\mathcal{Y}$ denotes the output alphabet of the channel.

[2] The Bhattacharyya parameter of W, denoted by $\mathcal{Z}(W)$ is given by

$$\mathcal{Z}(W) \stackrel{\text{def}}{=} \frac{1}{q(q-1)}\sum_{x,x'\in\mathbb{F}_q, x'\neq x}\sum_{y\in\mathcal{Y}}\sqrt{W(y|x)W(y|x')}$$

This parameter quantifies the amount of noise in the channel. It is close to 0 for channels with very low noise (i.e. channels of capacity close to 1) whereas it is close to 1 for very noisy channels (i.e. channels of capacity close to 0).

The j-th column of this matrix, denoted by $\mathbf{\Pi}_{\mathbf{y}}^{j}$, gives the a posteriori probabilities (APP) that the j-th codeword symbol is equal to α where α ranges over $\mathbb{F}_q$.

The reliability matrix is used by the Koetter-Vardy decoder to compute a multiplicity matrix that serves as the input to its soft interpolation step.

Algebraic geometry (AG) codes. The problem with Reed-Solomon codes is that their length is limited by the alphabet size. To overcome this limitation it is possible to proceed as in [4] and use instead AG codes[3].

Theorem 2 ([9])**.** *For any number $R \in [0,1]$ and any square prime power q there exists an infinite family of AG codes over $\mathbb{F}_q$ of rate $\geq R$ of increasing length n such that the normalized genus $\gamma \stackrel{def}{=} \frac{g}{n}$ of the underlying curve satisfies $\gamma \leq \frac{1}{\sqrt{q}-1}$.*

We will call such codes *Tsfasman-Vlăduts-Zink* AG codes.

When does the Koetter-Vardy decoding algorithm succeed? The Koetter-Vardy soft decoder [3] starts by computing with Algorithm A of [3, p.2814] from the knowledge of the reliability matrix $\mathbf{\Pi}$ and an integer parameter s (the total number of interpolation points counted with multiplicity) a $q \times n$ nonnegative integer matrix $\boldsymbol{M}(s)$ whose entries sum up to s. When s goes to infinity $\boldsymbol{M}(s)$ becomes proportional to $\mathbf{\Pi}$. The cost of this matrix (we will drop the dependency in s) $C(\boldsymbol{M})$ is defined as

$$C(\boldsymbol{M}) \stackrel{\text{def}}{=} \frac{1}{2}\sum_{i=1}^{q}\sum_{j=1}^{n} m_{ij}(m_{ij}+1) = \frac{1}{2}\left(\langle \boldsymbol{M}, \boldsymbol{M}\rangle + \langle \boldsymbol{M}, \mathbf{1}\rangle\right)$$

where m_{ij} denotes the entry of $\boldsymbol{M}$ at row i and column j and $\mathbf{1}$ is the all-one matrix. The complexity of the Koetter-Vardy decoding algorithm is dominated by solving a system of $C(\boldsymbol{M})$ linear equations. Then, the number of codewords on the list produced by the Koetter-Vardy decoder for a given multiplicity matrix $\boldsymbol{M}$ does not exceed (this is a corollary of results in [4], for more details see [6])

$$\mathcal{L}(\boldsymbol{M}) \stackrel{\text{def}}{=} \frac{g + \sqrt{2m(C(\boldsymbol{M})+g) + g^2}}{m}$$

It is straightforward to obtain from these considerations a soft-decision list decoder with a list which does not exceed some prescribed quantity L. Indeed it suffices to increase the value of s in [3, Algorithm A] until getting a matrix $\boldsymbol{M}$ which is such that $L \leq \mathcal{L}(\boldsymbol{M}) < L+1$. and to use this multiplicity matrix $\boldsymbol{M}$ in the Koetter-Vardy decoding algorithm. We refer to this decoding procedure as *algebraic soft-decoding with list size limited to L.*

[3]An AG code is constructed from a triple $(\mathcal{X}, \mathcal{P}, mQ)$ where $\mathcal{X}$ denotes an algebraic curve over $\mathbb{F}_q$, $\mathcal{P} = \{P_1, \ldots, P_n\}$ denotes an n-tuple of pairwise distinct $\mathbb{F}_q$-rational points of $\mathcal{X}$ and mQ denotes a divisor of $\mathcal{X}$ consisting of a nonnegative integer m and an $\mathbb{F}_q$-rational point Q of $\mathcal{X}$ which is not in $\mathcal{P}$. Then, the AG code $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, mQ)$ of length n over $\mathbb{F}_q$ is defined by

$$\mathcal{C}_L(\mathcal{X}, \mathcal{P}, mQ) \stackrel{\text{def}}{=} \{(f(P_1), \ldots, f(P_n)) \mid f \in \mathcal{L}(mQ)\}$$

where $\mathcal{L}(mQ)$ denotes the corresponding Riemann Roch space of mQ. If $n > m \geq 2g-1$, then $\mathcal{C}_L(\mathcal{X}, \mathcal{P}, mQ)$ has dimension $k = m - g + 1$ and distance $d \geq n - m$, where g is the genus of $\mathcal{X}$.

The following theorem which is as straightforward generalization of [Theorem 17, [3]] to the case of AG codes gives a sufficient condition ensuring that decoding produces a list containing the right codeword, for more details see [6].

Theorem 3. *Algebraic soft-decoding for AG codes with list-size limited to L produces a list that contains a codeword $\mathbf{c} \in \mathcal{C}_L(\mathcal{X}, \mathcal{P}, mQ)$ if*

$$\frac{\langle \mathbf{\Pi}, \lfloor \mathbf{c} \rfloor \rangle}{\sqrt{\langle \mathbf{\Pi}, \mathbf{\Pi} \rangle}} \quad = \quad \sqrt{m}\left(1 + \mathcal{O}\left(\frac{1}{L}\right)\right)$$

where $\tilde{R} = \frac{m}{n}$, $\tilde{\gamma} = \frac{g}{m}$ and $\mathcal{O}(\cdot)$ depends only on $\tilde{R}$, $\tilde{\gamma}$ and q.

Decoding capability of the Koetter-Vardy decoder when the channel is symmetric. The previous formula does not explain directly under which condition on the rate of the Reed-Solomon code decoding typically succeeds. We will derive now such a result, it will be convenient to restrict the channel to be *weakly symmetric*[4]. The idea underlying this restriction is to make the behavior of the quantity $\langle \mathbf{\Pi}, \lfloor \mathbf{c} \rfloor \rangle$ independent of the codeword $\mathbf{c}$ which is sent.

Notation 4. *We denote for such a channel and for a given output y by $\pi_y = (\pi(\alpha))_{\alpha \in \mathbb{F}_q}$ the associated APP vector, that is $\pi(\alpha) = \text{prob}(x = \alpha | y)$ where we denote by x the input symbol to the channel.*

From now on we assume the input of the communication channel is assumed to be uniformly distributed over $\mathbb{F}_q$.

The quantity $\mathbb{E}\left(\|\pi\|^2\right)$ turns out to be the limit of the rate for which the Koetter-Vardy decoder succeeds in decoding when the alphabet when the alphabet becomes large.

Definition 3 (Koetter-Vardy capacity)**.** Consider a weakly symmetric channel and denote by π the associated probability vector. The Koetter-Vardy capacity of this channel, which we denote by C_{KV}, is defined by

$$C_{\text{KV}} \stackrel{\text{def}}{=} \mathbb{E}(\|\pi\|^2).$$

Indeed for Reed-Solomon codes we have

Theorem 5. *Consider a weakly symmetric q-ary input channel of Koetter-Vardy capacity C_{KV}. Consider a Reed-Solomon code over $\mathbb{F}_q$ of length n, dimension k such that its rate $R = \frac{k}{n}$ satisfies $R < C_{KV}$. Let $\delta \stackrel{def}{=} \frac{C_{KV} - R}{R}$, $R^* \stackrel{def}{=} \frac{k-1}{n}$ and*

$$L \stackrel{def}{=} \left\lceil \frac{3\left(\frac{1}{R^*} + \frac{\sqrt{q}}{2\sqrt{R^*}}\right)\left(1 + \frac{\delta}{3}\right)}{\delta} \right\rceil.$$

The probability that the Koetter-Vardy decoder with list size bounded by L does not output in its list the right codeword is upper-bounded by $\mathcal{O}\left(e^{-K\delta^2 n}\right)$ for some constant K.

[4]A discrete memoryless W with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ is said to be weakly symmetric if and only if there is a partition of the output alphabet $\mathcal{Y} = Y_1 \cup \cdots \cup Y_n$ such that all the submatrices $W_i \stackrel{\text{def}}{=} (W(y|x))_{\substack{x \in \mathcal{X} \\ y \in Y_i}}$ are symmetric. A matrix is said to be symmetric if all if its rows are permutations of each other, and all its columns are permutations of each other.

Note that this result also allows to recover [5, Theorem 4]. For *Tsfasman-Vlăduts-Zink* AG codes we have the followong result

Theorem 6. *Consider a weakly symmetric q-ary input channel of Koetter-Vardy capacity C_{KV} where q is a square prime power. Consider a Tsfasman-Vlăduts-Zink AG code over $\mathbb{F}_q$ of length n, dimension k such that its rate $R = \frac{k}{n}$ satisfies $R < C_{KV} - \gamma$ where $\gamma \stackrel{def}{=} \frac{1}{\sqrt{q}-1}$.*

Let $\delta \stackrel{def}{=} \frac{C_{KV}-R-\gamma}{R}$, $\tilde{R} \stackrel{def}{=} \frac{m}{n}$, $\tilde{\gamma} \stackrel{def}{=} \frac{g}{m}$, $L \stackrel{def}{=} f^{-1}\left(1+\frac{\delta}{3}\right)$ and

$$f(\ell) \stackrel{def}{=} \frac{1+\frac{\tilde{\gamma}+\sqrt{2\tilde{\gamma}}}{\ell\sqrt{1-\frac{2\tilde{\gamma}}{\ell}\left(1+\frac{2}{\ell}\right)}}}{1-\frac{1}{\ell\sqrt{1-\frac{2\tilde{\gamma}}{\ell}\left(1+\frac{2}{\ell}\right)}}\left(\frac{\sqrt{q}}{2\sqrt{\tilde{R}}}+\frac{1}{\tilde{R}}\right)}.$$

The probability that the Koetter-Vardy decoder with list size bounded by L does not output in its list the right codeword is upper-bounded by $\mathcal{O}\left(e^{-K\delta^2 n}\right)$ for some constant K. Moreover $L = \Theta(1/\delta)$ as δ tends to zero.

This result applies to a wider family of codes than Theorem 5, however it does not allow to recover Theorem 5 in the case of (generalized) Reed-Solomon codes which corresponds to $g = 0$.

IV. Correcting errors beyond the Guruswami-Sudan bound

The asymptotic Koetter-Vardy capacity of a family of q-ary symmetric channels of crossover probability p (or q-SC$_p$, for short) is equal to $(1-p)^2$. It turns out that this is also the maximum crossover probability that the Guruswami-Sudan decoder is able to sustain when the alphabet and the length go to infinity. We have proved in [5] that the $(U \mid U+V)$ construction with Reed-Solomon components already performs a bit better than $(1-p)^2$ when the rate is small enough (as soon as $R < 0.17$). Now, by using iterated $(U \mid U+V)$ constructions, even for a moderate number of levels, we will be able to improve significantly the performances.

To study an iterated $(U \mid U+V)$ code of depth ℓ it will be helpful to bring in the quantity $f(W,\ell) \stackrel{\text{def}}{=} \frac{1}{2^\ell}\sum_{i=0}^{2^\ell-1} C_{\text{KV}}(W^i)$ where W^i is the channel viewed by the constituent U_i code for a given noisy channel W. We will consider the case where W is a q-ary symmetric channel of crossover probability p and all constituent codes are Reed-Solomon codes. It is straighforward to prove in this case [6] that all the channels W^i are weakly symmetric. When q tends to infinity it is then clear that from Theorem 5 the asymptotic rate up to which succesful decoding is possible with probability $1-o(1)$ is $\lim_{q\to\infty} f(W,\ell)$.

Figure 1 summarizes the performances of the iterated $(U \mid U+V)$-construction of depth 2 and 3 (for the proof we refer to [6]). From this figure we see that if we apply the iterated $(U \mid U+V)$-construction of depth 2 we get better performance than decoding a classical Reed-Solomon code with the Guruswami-Sudan decoder for low rate codes, specifically for $R < 0.325$. Moreover, if we apply the iterated $(U \mid U+V)$-construction of depth 3 we get even better results, we beat the Guruswami-Sudan for codes of rate $R < 0.475$.

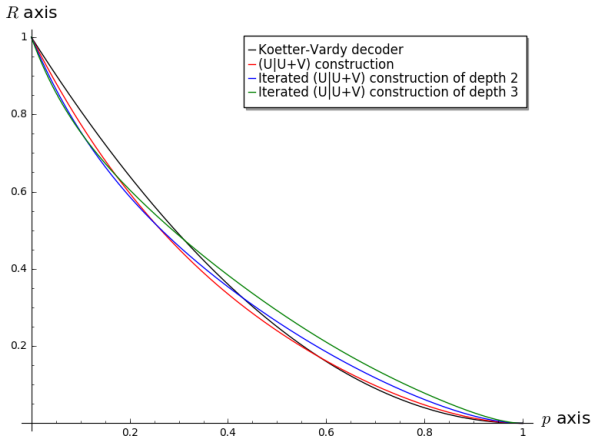

Figure 1. Asymptotic rate for which decoding succeeds with probability $1-o(1)$ plotted against the crossover error probability p for four code-constructions. The black line refers to standard Reed-Solomon codes decoded by the Guruswami-Sudan algorithm, the red line to the $(U \mid U+V)$-construction, the blue line to the iterated $(U \mid U+V)$-construction of depth 2 and the green line to the iterated $(U \mid U+V)$-construction of depth 3. All constituent codes are Reed-Solomon codes.

Even if for finite alphabet size q the Koetter-Vardy capacity cannot be understood as a capacity in the usual sense, it is still insightful to consider $f(W,\ell) \stackrel{\text{def}}{=} \frac{1}{2^\ell}\sum_{i=0}^{2^\ell-1} C_{\text{KV}}(W^i)$ where W^i is the channel viewed by the constituent U_i code for an iterated-$(U \mid U+V)$ construction of depth ℓ and for a given noisy channel. This could be considered as the limit for which we can not hope to have small probabilities of error after decoding when using Reed-Solomon constituent codes and the Koetter-Vardy decoding algorithm. We have plotted these functions in Figure 2 for $q = 256$ and $\ell = 0$ up to $\ell = 6$ and a q-SC$_p$. It can be seen that for $\ell = 5, 6$ we get rather close to the actual capacity of the channel in this way.

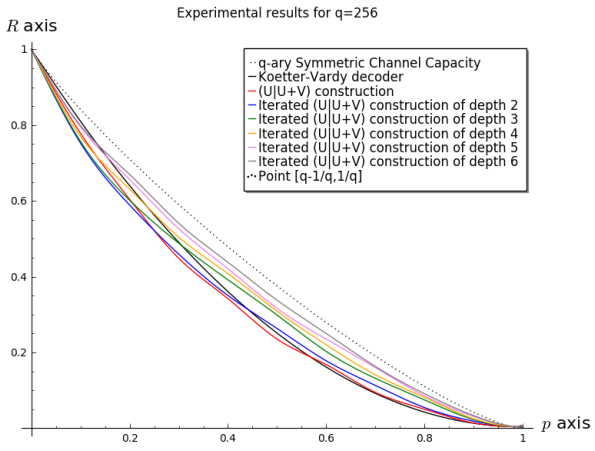

Figure 2. average Koetter-Vardy capacity plotted against the crossover error probability p for seven code constructions. The noise model is a a q-SC$_p$. .

V. Attaining the capacity with an iterated $(U \mid U+V)$ construction

When the number of levels for which we iterate this construction tends to infinity, we attain the capacity of any q-ary symmetric channel at least when the cardinality q is prime. Moreover the probability of error after decoding can be made to be (almost) exponentially small with respect to the overall codelength. More precisely we have the following results about the probability of error.

Theorem 7. *Let W be a cyclic-symmetric[5] q-ary channel where q is prime. Let C be the capacity of this channel. For any β in the range $(0, 1/2)$ and sufficiently small positive $\epsilon > 0$, there exists a sequence of iterated $(U \mid U+V)$ codes with Reed-Solomon constituent codes of arbitrarily large length over some extension field $\mathbb{F}_{q^m}$ which have rate $\geq C - \epsilon$ when the codelength is sufficiently large and whose probability of error P_e is upper bounded by*

$$P_e \leq n e^{-\frac{K(\epsilon,\beta)N}{n \log N}}$$

when decoded with the iterated $(U \mid U+V)$ decoder based on decoding the constituent codes with the Koetter-Vardy decoder with listsize bounded by $\mathcal{O}\left(\frac{1}{\epsilon}\right)$ and where N is the codelength (over $\mathbb{F}_q$), $n = O(\log\log N)^{1/\beta}$, the alphabet size q^m is related to n and N by $N = nmq^m$ and $K(\epsilon, \beta)$ is some positive function of ϵ and β.

Remark 1. For more details the reader is referred to [6]. The channel is with q-ary inputs, but the Reed-Solomon codes that are decoded are defined over $\mathbb{F}_{q^m}$. The depth of the iterated $(U \mid U+V)$ construction is $\ell \stackrel{\text{def}}{=} \log_2 n$ and the constituent codes have length (over $\mathbb{F}_q^m$) $q^m = \frac{N}{nm} \sim \frac{N}{n \log_q N}$ which is of order $O\left(\frac{N}{\log_q N (\log\log N)^{1/\beta}}\right)$. This accounts for the exponent $-\frac{K(\epsilon,\beta)N}{n \log N}$ of the probability of error. The alphabet size q^m has to scale almost linearly with the codelength and this is somewhat unsatisfactory. We will avoid this in what follows by using AG codes.

For the iterated $(U \mid U+V)$-construction with algebraic geometry codes as constituent codes we obtain an even stronger result which is

Theorem 8. *Let W be a cyclic-symmetric q-ary channel where q is prime. Let C be the capacity of this channel. For any sufficiently small positive ϵ, there exists a sequence of iterated $(U \mid U+V)$ codes of arbitrarily large length with AG defining codes of rate $\geq C - \epsilon$ when the codelength is sufficiently large and whose probability of error P_e is upper bounded by*

$$P_e \leq e^{-K(\epsilon)N}$$

when decoded with the iterated $(U \mid U+V)$ decoder based on the Koetter-Vardy algorithm with listsize bounded by $\mathcal{O}\left(\frac{1}{\epsilon}\right)$ and where K is some positive function of ϵ.

[5] By following [1] we denote for a vector $\mathbf{y} = (y_i)_{i \in \mathbb{F}_q}$ with coordinates indexed by a finite field $\mathbb{F}_q$ by $\mathbf{y}^{+g}$ the vector $\mathbf{y}^{+g} = (y_{i+g})_{i \in \mathbb{F}_q}$, by $n(\mathbf{y})$ the number of g's in $\mathbb{F}_q$ such that $\mathbf{y}^{+g} = \mathbf{y}$ and by $\mathbf{y}^*$ the set $\{\mathbf{y}^{+g}, g \in \mathbb{F}_q\}$. A q-ary input channel is cyclic-symmetric channel if and only there exists a probability function Q defined over the sets of possible π^* such that for any $i \in \mathbb{F}_q$ we have $\text{prob}(\pi = \mathbf{y} | x = i) = y_i n(\mathbf{y}) Q(\mathbf{y}^*)$.

VI. Conclusion

A variation on polar codes that is much more flexible. We have given here a variation of polar codes that allows to attain capacity with a polynomial-time decoding complexity in a more flexible way than standard polar codes. It consists in taking an iterated-$(U \mid U+V)$ construction based on Reed-Solomon codes or more generally AG codes. Decoding consists in computing the APP of each position in the same way as polar codes and then to decode the constituent codes with a soft information decoder, the Koetter-Vardy list decoder in our case.

An exponentially small probability of error. This allows to control the rate and the error probability in a much finer way as for standard polar codes. In our case, this error probability can be decreased significantly by choosing a long enough code and a rate below the noise value that our decoder is able to sustain (which is more or less the Koetter-Vardy capacity of the noisy channel in our case).

Practical constructions. What is suggested by the experimental evidence shown in Section IV is that these codes do not only have some theoretical significance, but that they should also yield interesting codes for practical applications. Indeed Figure 2 shows that it should be possible to get very close to the channel capacity by using only a construction with a small depth, say $5 - 6$ together with constituent codes of moderate length that can be chosen to be Reed-Solomon codes (say codes of length a hundred/a few hundred at most).

References

[1] A. Bennatan and D. Burshtein. Design and analysis of nonbinary LDPC codes over arbitrary discrete-memoryless channels. *IEEE Trans. Inform. Theory*, 52(2):549–583, Feb. 2006.

[2] I. Dumer. Soft-decision decoding of Reed-Muller codes: a simplified algorithm. *IEEE Trans. Inform. Theory*, 52(3):954–963, 2006.

[3] R. Koetter and A. Vardy. Algebraic soft-decision decoding of Reed-Solomon codes. *IEEE Trans. Inform. Theory*, 49(11):2809–2825, 2003.

[4] R. Koetter and A. Vardy. Algebraic soft-decision decoding of Reed-Solomon codes. 2003.

[5] I. Márquez-Corbella and J.-P. Tillich. Using Reed-Solomon codes in the $(u|u+v)$ construction and an application to cryptography. In *Proc. IEEE Int. Symposium Inf. Theory - ISIT*, pages 930–934, 2016.

[6] I. Márquez-Corbella and J.-P. Tillich. Attaining capacity with algebraic geometry codes through the $(U|U+V)$ construction and Koetter-Vardy soft decoding. preprint, arXiv:1701.07112, 2017.

[7] E. Şaşoğlu. Polarization and polar codes. *Foundations and Trends in Communications and Information Theory*, 8(4):259–381, 2011.

[8] E. Şaşoğlu, E. Telatar, and E. Arıkan. Polarization for arbitrary discrete memoryless channels. In *Proc. IEEE Inf. Theory Workshop- ITW*, pages 144–149, Oct. 2009.

[9] M. A. Tsfasman, S. Vlăduts, and T. Zink. Modular curves, Shimura curves, and Goppa codes, better than Varshamov-Gilbert bound. *Mathematische Nachrichten*, 109(1):21–28, 1982.