# Constructing Polar Codes for Non-Binary Alphabets and MACs

**Ido Tal**

UCSD

**Artyom Sharov**

Technion

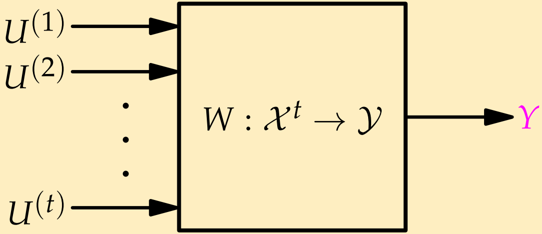
**Alexander Vardy**

UCSD

# MAC channels and their polarization

## $t$-user MAC

Let $W : \mathcal{X}^t \to \mathcal{Y}$ be a $t$-user MAC

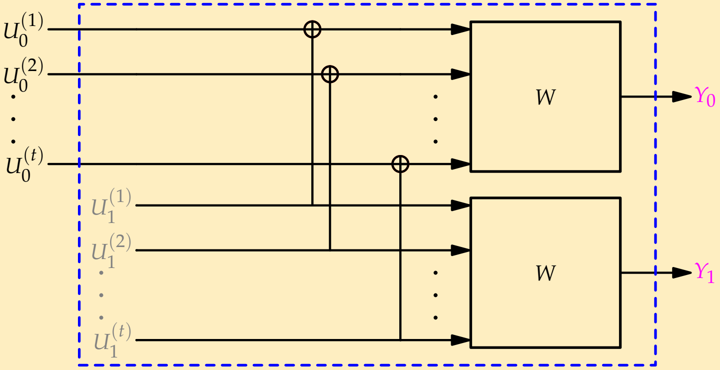



- Input alphabet $\mathcal{X} = \{0, 1, \ldots, p-1\}$, where $p$ prime.
- Output alphabet $\mathcal{Y}$, finite.

# Arıkan "-" transform

## $W^-$ channel

Define $W^-$ : $\mathcal{X}^t \to \mathcal{Y}^2$ as follows:

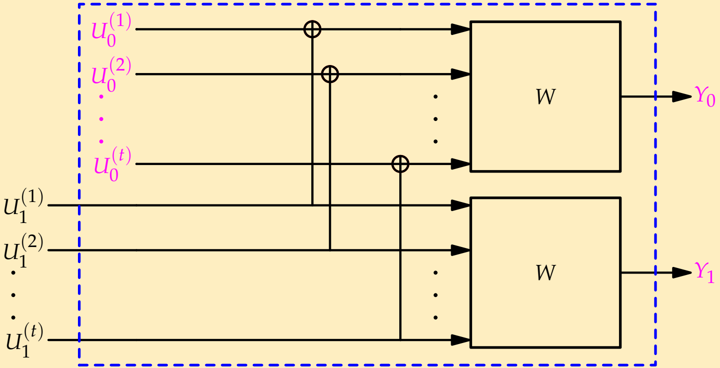



$$W^-(y_0, y_1 | \mathbf{u}_0) = \sum_{\mathbf{u}_1 \in \mathcal{X}^t} \frac{1}{p^t} W(y_0 | \mathbf{u}_0 \oplus_p \mathbf{u}_1) \cdot W(y_1 | \mathbf{u}_1) \ .$$

# Arıkan "+" transform

## $W^+$ channel

Define $W^+ : \mathcal{X}^t \to \mathcal{Y}^2 \times \mathcal{X}^t$ as follows:




$$W^+(y_0, y_1, \mathbf{u}_0 | \mathbf{u}_1) = \frac{1}{p^t} W(y_0 | \mathbf{u}_0 \oplus_p \mathbf{u}_1) \cdot W(y_1 | \mathbf{u}_1) \ .$$

# Evolving MACs

## Recursive definition

Let the underlying MAC be

$$\mathcal{W}_0^{(0)} = \mathbb{W}$$

For $n = 2^m$ and $0 \leq i < n$, recursively define

$$\mathcal{W}_{2i}^{(m+1)} = \left(\mathcal{W}_i^{(m)}\right)^- \;, \quad \mathcal{W}_{2i+1}^{(m+1)} = \left(\mathcal{W}_i^{(m)}\right)^+$$

## Theorem [Şaşoğlu,Telatar,Yeh], [Abbe,Telatar]

As $m \to \infty$, almost all MACs

$$\mathcal{W}_i^{(m)} \;, \quad 0 \leq i < n = 2^m$$

"polarize". Thus, a polar-coding scheme can be implemented*.

*See [Şaşoğlu,Telatar,Yeh: Appendix A] for a simpler implementation.

# The problem

**Output alphabet grows exponentially in $n$**

Recall that if $W : \mathcal{X}^t \to \mathcal{Y}$, then

$$W^- : \mathcal{X}^t \to \mathcal{Y}^2 , \quad W^+ : \mathcal{X}^t \to \mathcal{Y}^2 \times \mathcal{X}^t .$$

Thus, the size of the output alphabet of $\mathcal{W}_i^{(m)}$ is at least $|\mathcal{Y}|^{2^m} = |\mathcal{Y}|^n$.

**Solution**

- Instead of calculating $\mathcal{W}_i^{(m)}$ exactly, calculate an approximation
- Approximate by a channel having a bounded output alphabet size
- Prove that the approximation is tight

## Comparison to previous [Tal,Vardy] method

| Parameter | Previous | New |
|---|---|---|
| Input alphabet $\mathcal{X}$ | $\{0, 1\}$ | $\{0, 1, \ldots, p-1\}$ |
| Users | single user | $t$ users |
| Running time, $n$ | $O(n)$ | $O(n)$ |
| Running time, $q = p^t$ | — | exponential in $q$ |
| Need $W$ symmetric? | yes | no |

## Main idea in previous method

- Find two "closest" output letters
- Merge these two letters into one
- Continue until alphabet is small enough

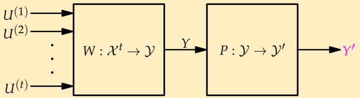
## Main idea in new method

- Place output letters in "bins"
- Merge all letters in the same "bin"

# Degradation

## MAC degradation

$Q : \mathcal{X}^t \to \mathcal{Y}'$ is degraded with respect to $W : \mathcal{X}^t \to \mathcal{Y}$ if there exists a single-user channel $P : \mathcal{Y} \to \mathcal{Y}'$ such that




$$Q(y'|\mathbf{u}) = \sum_{y \in \mathcal{Y}} W(y|\mathbf{u}) \cdot P(y'|y) \; .$$

We denote this as $Q \preceq W$.

## Lemma [Korada]: Arıkan transforms preserve degradation

Let $Q \preceq W$. Then,

$$Q^{-} \preceq W^{-} \quad \text{and} \quad Q^{+} \preceq W^{+} \; .$$

# Sum-rate as figure of merit

## Sum-rate definition

- Let $\mathbf{U} = (U^{(i)})_{i=1}^{t}$ be uniformly distributed over $\mathcal{X}^t$
- Let $Y$ be the the output of $W : \mathcal{X}^t \to \mathcal{Y}$ when the input is $\mathbf{U}$.
- Define

$$R(W) = I(\mathbf{U}; Y) \ .$$

## Lemma

Let $Q \preceq W$. Define $Y'$ as the output of $Q$ when the input is $\mathbf{U}$. Let $A, B \subseteq \{1, 2, \ldots, t\}$, where $A \cap B = \emptyset$. Denote

$$\mathbf{U}_A = (U^{(i)})_{i \in A} \quad \text{and} \quad \mathbf{U}_B = (U^{(i)})_{i \in B} \ .$$

Then,

$$R(Q) \geq R(W) - \varepsilon \implies I(\mathbf{U}_A; \mathbf{U}_B, Y') \geq I(\mathbf{U}_A; \mathbf{U}_B, Y) - \varepsilon \ .$$

# A bit of notation

## The channel

- $W : \mathcal{X}^t \to \mathcal{Y}$
- $\mathbf{U} = (U^{(i)})_{i=1}^t$ uniform on $\mathcal{X}^t$, input to $W$
- $Y$ output of $W$

## Probabilities

- $\varphi(\mathbf{u}|y) = \mathbb{P}(\mathbf{U} = \mathbf{u}|Y = y)$
- $\varphi(y) = \mathbb{P}(Y = y)$

## The function $\eta$

Let

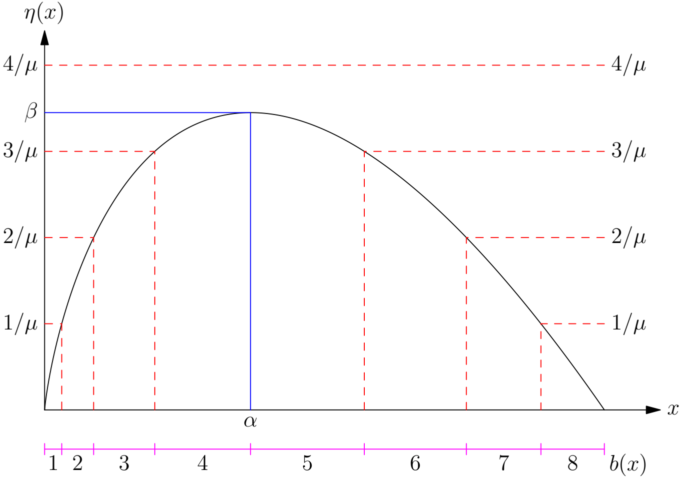
$$\eta(x) = -x \cdot \log_2 x \ .$$

Thus,

$$R(W) = t \log_2 p - \sum_{y \in \mathcal{Y}} \varphi(y) \sum_{\mathbf{u} \in \mathcal{X}^t} \eta(\varphi(\mathbf{u}|y)) \ .$$

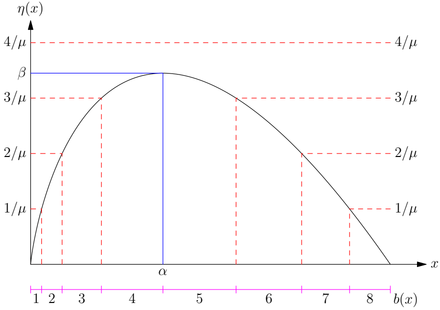
# Quantizing $\eta$

Let $\mu$ be a fidelity criterion, and let $\widehat{\mu} = \lceil \beta \cdot \mu \rceil$. Define the function $b : [0,1] \to \{1, 2, \ldots, 2\widehat{\mu}\}$ as follows.



# Quantizing $\eta$

Let $\mu$ be a fidelity criterion, and let $\widehat{\mu} = \lceil \beta \cdot \mu \rceil$. Define the function $b : [0,1] \to \{1, 2, \ldots, 2\widehat{\mu}\}$ as follows.




**Lemma**

Let $0 \leq x \leq 1$ and $0 \leq x' \leq 1$ be such that $b(x) = b(x')$. Then,

$$\left|\eta(x) - \eta(x')\right| \leq \frac{1}{\mu} .$$

# Constructing $Q \preceq W$

## Output letters in the same bin

We say that two output letters $y_1, y_2 \in \mathcal{Y}$ are in the same bin if for all $\mathbf{u} \in \mathcal{X}^t$ we have

$$b(\varphi(\mathbf{u}|y_1)) = b(\varphi(\mathbf{u}|y_2)) .$$

## Constructing $Q$

- Degrade $W$: rename all the letters $y_1, y_2, \ldots$ in the same bin to $y'$.

## Lemma

Let $y \in \mathcal{Y}$ be renamed to $y' \in \mathcal{Y}'$. Then, for all $\mathbf{u} \in \mathcal{X}^t$,

$$b(\varphi_W(\mathbf{u}|y)) = b(\varphi_Q(\mathbf{u}|y')) .$$

# Degrading bound

**Theorem**

Let $W$ be a $t$-user MAC with $\mathcal{X} = \{0, 1, \ldots, p-1\}$. Degrade $W$ to $Q$, using fidelity criterion $\mu$. Then,

$$R(Q) \geq R(W) - \frac{p^t}{\mu} \,.$$

**Proof**

$$\begin{aligned}
R(W) - R(Q) &= \sum_{y' \in \mathcal{Y}'} \sum_{y \in \mathcal{B}(y')} \varphi(y) \sum_{\mathbf{u} \in \mathcal{X}^t} \left[ \eta(\varphi_Q(\mathbf{u}|y')) - \eta(\varphi_W(\mathbf{u}|y)) \right] \\
&\leq \sum_{y' \in \mathcal{Y}'} \sum_{y \in \mathcal{B}(y')} \varphi(y) \sum_{\mathbf{u} \in \mathcal{X}^t} \frac{1}{\mu} \\
&= \sum_{y' \in \mathcal{Y}'} \sum_{y \in \mathcal{B}(y')} \varphi(y) \cdot \frac{p^t}{\mu} = \frac{p^t}{\mu} \,.
\end{aligned}$$

# Bounding the output alphabet size

**Lemma**

Let $W$ be a $t$-user MAC with $\mathcal{X} = \{0, 1, \ldots, p-1\}$. Degrade $W : \mathcal{X}^t \to \mathcal{Y}$ to $Q : \mathcal{X}^t \to \mathcal{Y}'$, using fidelity criterion $\mu$. Denote $q = p^t$. Then,

$$\left|\mathcal{Y}'\right| \leq (2\widehat{\mu})^q \leq (2\mu)^q \; .$$

**Proof**

$(2\widehat{\mu})^q$ is an upper-bound on the number of non-empty bins.

# Repeated application of our method

---

**Algorithm A**: A high level description of the degrading procedure

---

**input** : An underlying MAC $\mathbb{W}$, a fidelity parameter $\mu$, an index $i = \langle b_1, b_2, \ldots, b_m \rangle_2$.

**output**: A MAC that is degraded with respect to $\mathcal{W}_i^{(m)}$.

```
Q ← degrading_merge(𝕎, μ);
for j = 1, 2, ..., m do
    if b_j = 0 then
        W ← (Q)⁻
    else
        W ← (Q)⁺
    Q ← degrading_merge(W, μ);
return Q;
```

# Average error

**Theorem**

Let an underlying $t$-user MAC $\mathbb{W} : \mathcal{X}^t \to \mathcal{Y}$ be given, where $\mathcal{X} = \{0, 1, \ldots, p-1\}$ and $p$ is prime. Denote by $\mathcal{Q}_i^{(m)}$ the channel returned by running Algorithm A with parameters $i$ and $\mu$. Then,

$$\frac{1}{n} \sum_{0 \le i < n} \left( R(\mathcal{W}_i^{(m)}) - R(\mathcal{Q}_i^{(m)}) \right) \le \frac{m \cdot p^t}{\mu} \ .$$

**Proof sketch**

Follows easily from the error bound for a single round, and from the fact that

$$2R(W) = R(W^-) + R(W^+) \ .$$

# Can we do better?

**Re-grouping $R(W) - R(Q)$**

$$R(W) - R(Q) =$$

$$\sum_{y' \in \mathcal{Y}'} \varphi_Q(y') \sum_{\mathbf{u} \in \mathcal{X}^t} \left( \eta \left[ \sum_{y \in \mathcal{B}(y')} \frac{\varphi_W(y)}{\varphi_Q(y')} \cdot \varphi_W(\mathbf{u}|y) \right] - \left[ \sum_{y \in \mathcal{B}(y')} \frac{\varphi_W(y)}{\varphi_Q(y')} \eta(\varphi_W(\mathbf{u}|y)) \right] \right) .$$

For a given $y' \in \mathcal{Y}'$ and $\mathbf{u} \in \mathcal{X}^t$, the value of $b(\eta(\varphi_W(\mathbf{u}|y))$ is the same for all $y \in \mathcal{B}(y')$. Denote the interval that gets mapped to this value as

$$I_{y'} = \{x \ : \ b(x) = b(\varphi_W(\mathbf{u}|y)\} \ , \quad \text{where } y \in \mathcal{B}(y') \ .$$

# Can we do better?

## Lemma

Let $a = \inf I_{y'}$, $b = \sup I_{y'}$. Then,

$$\eta\left[\sum_{y\in\mathcal{B}(y')}\frac{\varphi_W(y)}{\varphi_Q(y')}\cdot\varphi_W(\mathbf{u}|y)\right]-\left[\sum_{y\in\mathcal{B}(y')}\frac{\varphi_W(y)}{\varphi_Q(y')}\eta(\varphi_W(\mathbf{u}|y))\right]$$

is at most

$$\max_{0\leq\theta\leq 1}\left\{\eta\left[\theta\cdot a+(1-\theta)\cdot b\right]-\left[\theta\cdot\eta(a)+(1-\theta)\cdot\eta(b)\right]\right\},$$

where

$$\theta_{\max}=\frac{b-\frac{1}{e}\cdot 2^{\frac{-(\eta(b)-\eta(a))}{b-a}}}{b-a}.$$

# Can we do better?