Lower Bounds on the Probability of Error of Polar Codes

Boaz Shuval and Ido Tal

Andrew and Erna Viterbi Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa 32000, Israel

June 2017

Introduction

- Arıkan's Polar Codes asymptotically achieve capacity
- Analysis based on upper bounds on P_e
- How tight is the upper bound?
- Existing lower bounds on P_e are trivial
- **In this work:** Improved lower bounds

Preliminaries

- BMS Channel $W(y|u)$
- Polar Construction $\rightarrow N = 2^n$ synthetic channels

$$W_i(y_1^N, u_1^{i-1}|u_i)$$

- Polarize to "good" ($\mathcal{A}$) and "bad" channels
- Transmit frozen bits on "bad" channels
- Successive Cancellation Decoding

$$\hat{U}_i(y_1^N, \hat{u}_1^{i-1}) = \begin{cases} u_i & i \in \mathcal{A}^c \\ \arg\max_{u_i} W_i(y_1^N, \hat{u}_1^{i-1}|u_i) & i \in \mathcal{A} \end{cases}$$

SC Probability of Error

Let

$$\mathcal{E}_i = \text{event that } W_i \text{ errs}$$

SC probability of error:

$$P_e^{\mathrm{SC}} = \mathbb{P}\left\{\bigcup_{i \in \mathcal{A}} \mathcal{E}_i\right\}$$

Bounds:

$$\max_{i \in \mathcal{A}} \mathbb{P}\{\mathcal{E}_i\} \leq P_e^{\mathrm{SC}} \leq \sum_{i \in \mathcal{A}} \mathbb{P}\{\mathcal{E}_i\}$$

Question

How do we improve the lower bound?

Improving the Lower Bound

Two ingredients:

- If $\mathcal{A}' \subseteq \mathcal{A}$ then

$$\mathbb{P}\left\{\bigcup_{i\in\mathcal{A}} \mathcal{E}_i\right\} \geq \mathbb{P}\left\{\bigcup_{i\in\mathcal{A}'} \mathcal{E}_i\right\}$$

- Bonferroni bound:

$$\mathbb{P}\left\{\bigcup_{i\in\mathcal{A}} \mathcal{E}_i\right\} \geq \sum_{i\in\mathcal{A}} \mathbb{P}\left\{\mathcal{E}_i\right\} - \sum_{\substack{i,j\in\mathcal{A},\\ i<j}} \mathbb{P}\left\{\mathcal{E}_i \cap \mathcal{E}_j\right\}$$

Recall $\mathbb{P}\left\{\mathcal{E}_i \cap \mathcal{E}_j\right\} = \mathbb{P}\left\{\mathcal{E}_i\right\} + \mathbb{P}\left\{\mathcal{E}_j\right\} - \mathbb{P}\left\{\mathcal{E}_i \cup \mathcal{E}_j\right\}$

Approach

Lower bounds on $\mathbb{P}\left\{\mathcal{E}_i \cup \mathcal{E}_j\right\} \Longrightarrow$ *better lower bounds on* P_e^{SC}

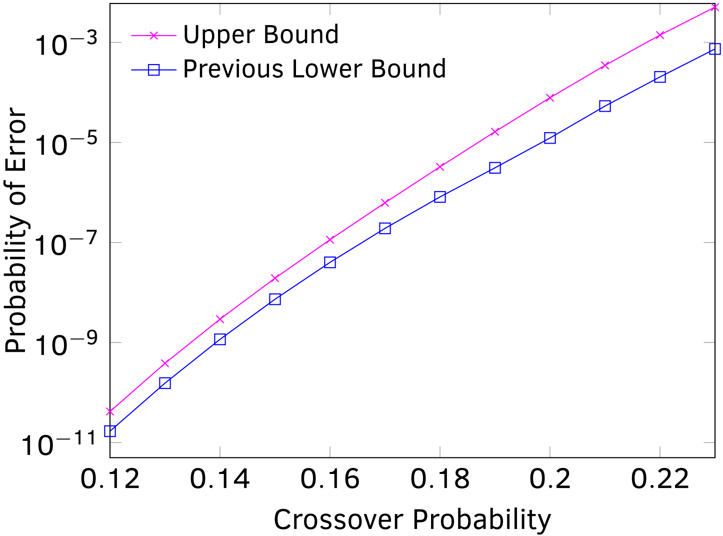
Previous Work

- Mori & Tanaka [2009]
  - Density evolution to approximate joint distribution $\Rightarrow \mathbb{P}\left\{\mathcal{E}_i \cup \mathcal{E}_j\right\}$
  - Exact for BEC
- Parizi & Telatar [2013]
  - Only for BEC
  - Track correlation between erasure events
  - Showed: union bound asymptotically tight for BEC

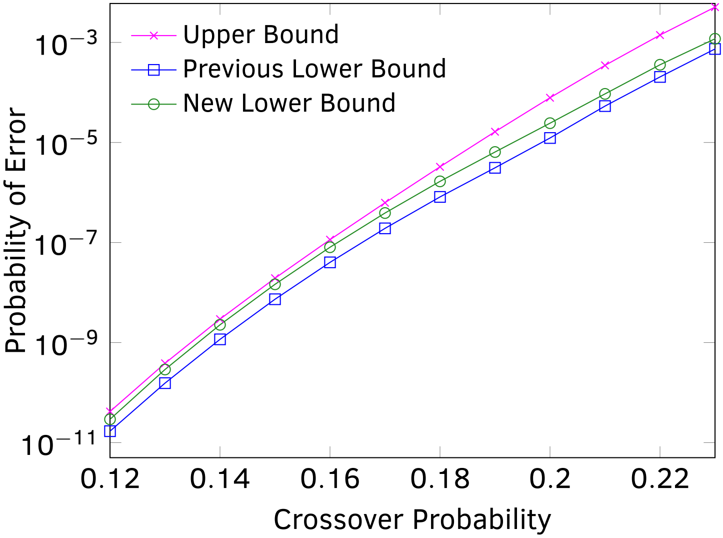
New Lower Bound

- Works for any initial BMS channel W
- **Provable** lower bound on P_e^{SC}
- Approximates joint distribution of two synthetic channels $W_{a,b}$
- Controls output alphabet sizes
- Coincides with lower bounds for BEC
- Better than existing lower bound for general BMS channels

Numerical Results



Numerical Results



Conceptual Algorithm

Input:

- BMS channel W
- a-channel transform list $\alpha_1, \alpha_2, \ldots, \alpha_n$
- b-channel transform list $\beta_1, \beta_2, \ldots, \beta_n$

$\alpha, \beta \in \{-, +\}$

Output: Lower bound on $P_e^{\mathrm{SC}}(W_{a_n,b_n})$

Steps:

1. Initialize: $W_{0,0} = W$
2. For $i = 1, \ldots, n$, do:
   - $W_{a_i,b_i} \hookleftarrow \texttt{JointlyPolarize}_{\alpha_i,\beta_i}(W_{a_{i-1},b_{i-1}})$
3. Compute: $\texttt{LowerBound}(P_e^{\mathrm{SC}}(W_{a_n,b_n}))$

Conceptual Algorithm

Input:

- BMS channel W
- a-channel transform list $\alpha_1, \alpha_2, \ldots, \alpha_n$
- b-channel transform list $\beta_1, \beta_2, \ldots, \beta_n$

$\alpha, \beta \in \{-, +\}$

Output: Lower bound on $P_e^{\mathrm{SC}}(W_{a_n,b_n})$

Steps:

1. Initialize: $W_{0,0} = W$
2. For $i = 1, \ldots, n$, do:
   - $W_{a_i,b_i} \leftarrow \mathtt{JointlyPolarize}_{\alpha_i,\beta_i}(W_{a_{i-1},b_{i-1}})$
   - $W_{a_i,b_i} \leftarrow \mathtt{JointlyUpgrade}(W_{a_i,b_i})$
3. Compute: $\mathtt{LowerBound}(P_e^{\mathrm{SC}}(W_{a_n,b_n}))$

alphabet size grows

control alphabet size

SC Decoding – suboptimal

(u_a, u_b)	(y_a, y_b)			
	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(0,0)$	0.30	0.04	0.04	0.62
$(0,1)$	0.44	0.46	0.01	0.09
$(1,0)$	0.22	0.49	0.24	0.05
$(1,1)$	0.05	0.54	0.32	0.09

SC Decoding – suboptimal

(u_a, u_b)	(y_a, y_b)			
	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(0,0)$	0.30	0.04	0.04	0.62
$(0,1)$	0.44	0.46	0.01	0.09
$(1,0)$	0.22	0.49	0.24	0.05
$(1,1)$	0.05	0.54	0.32	0.09

- Optimal decoder: $P_e = 0.52$

SC Decoding – suboptimal

(u_a, u_b)	(y_a, y_b)			
	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(0,0)$	0.30	0.04	0.04	0.62
$(0,1)$	0.44	0.46	0.01	0.09
$(1,0)$	0.22	0.49	0.24	0.05
$(1,1)$	0.05	0.54	0.32	0.09

- Optimal decoder: $P_e = 0.52$
- SC decoder: $P_e^{SC} = 0.7075$

SC Decoding – suboptimal

Degrade:

$$(0,0),(1,1) \to (0',0')$$
$$(0,1),(1,0) \to (1',1')$$

(u_a,u_b)	(y_a,y_b)			
	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(0,0)$	0.30	0.04	0.04	0.62
$(0,1)$	0.44	0.46	0.01	0.09
$(1,0)$	0.22	0.49	0.24	0.05
$(1,1)$	0.05	0.54	0.32	0.09

(u_a,u_b)	(y_a,y_b)	
	$(0',0')$	$(1',1')$
$(0,0)$	0.92	0.08
$(0,1)$	0.53	0.47
$(1,0)$	0.27	0.73
$(1,1)$	0.14	0.86

- Optimal decoder: $P_e = 0.52$
- SC decoder: $P_e^{SC} = 0.7075$

SC Decoding – suboptimal

Degrade:

$$(0,0),(1,1) \to (0',0')$$
$$(0,1),(1,0) \to (1',1')$$

(u_a, u_b)	(y_a, y_b)			
	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(0,0)$	0.30	0.04	0.04	0.62
$(0,1)$	0.44	0.46	0.01	0.09
$(1,0)$	0.22	0.49	0.24	0.05
$(1,1)$	0.05	0.54	0.32	0.09

- Optimal decoder: $P_e = 0.52$
- SC decoder: $P_e^{\text{SC}} = 0.7075$

(u_a, u_b)	(y_a, y_b)	
	$(0',0')$	$(1',1')$
$(0,0)$	0.92	0.08
$(0,1)$	0.53	0.47
$(1,0)$	0.27	0.73
$(1,1)$	0.14	0.86

- Optimal decoder: $P_e = 0.555$
- SC decoder: $P_e^{\text{SC}} = 0.555$

SC Decoding – suboptimal

Degrade:

$$(0,0),(1,1) \to (0',0')$$
$$(0,1),(1,0) \to (1',1')$$

(u_a,u_b)	(y_a,y_b)			
	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(0,0)$	0.30	0.04	0.04	0.62
$(0,1)$	0.44	0.46	0.01	0.09
$(1,0)$	0.22	0.49	0.24	0.05
$(1,1)$	0.05	0.54	0.32	0.09

- Optimal decoder: $P_e = 0.52$
- SC decoder: $P_e^{SC} = 0.7075$

(u_a,u_b)	(y_a,y_b)	
	$(0',0')$	$(1',1')$
$(0,0)$	0.92	0.08
$(0,1)$	0.53	0.47
$(1,0)$	0.27	0.73
$(1,1)$	0.14	0.86

- Optimal decoder: $P_e = 0.555$
- SC decoder: $P_e^{SC} = 0.555$

Conclusion

SC decoder is not ordered by degradation

New Decoder

Joint channel: $W_{a,b}(y_a, y_b | u_a, u_b)$

- New Decoder: minimize $\mathbb{P}\{\mathcal{E}_a \cup \mathcal{E}_b\}$ using

$$\hat{u}_a = \phi_a(y_a)$$
$$\hat{u}_b = \phi_b(y_b)$$

- Notation: P_e^*
- Generally requires exhaustive search

New Decoder

Joint channel: $W_{a,b}(y_a, y_b | u_a, u_b)$

- New Decoder: minimize $\mathbb{P}\left\{\mathcal{E}_a \cup \mathcal{E}_b\right\}$ using

$$\hat{u}_a = \phi_a(y_a)$$
$$\hat{u}_b = \phi_b(y_b)$$

- Notation: P_e^*

$$P_e^{SC}(W_{a,b}) \geq P_e^*(W_{a,b})$$

- Generally requires exhaustive search

New Decoder

Joint channel: $W_{a,b}(y_a, y_b | u_a, u_b)$

- New Decoder: minimize $\mathbb{P}\{\mathcal{E}_a \cup \mathcal{E}_b\}$ using

$$\hat{u}_a = \phi_a(y_a)$$
$$\hat{u}_b = \phi_b(y_b)$$

- Notation: P_e^*

$$P_e^{SC}(W_{a,b}) \geq P_e^*(W_{a,b})$$

- Generally requires exhaustive search
- For polar codes:
  - easily found
  - ordered by (proper) joint degradation:

$$Q_{a,b} \overset{p}{\succcurlyeq} W_{a,b} \Rightarrow P_e^*(W_{a,b}) \geq P_e^*(Q_{a,b})$$

Upgrading Procedures Overview

Goal:

- Find $Q_{a,b} \overset{p}{\succcurlyeq} W_{a,b}$
- Reduce output alphabet of one marginal
- Leave other marginal unchanged

New joint channel upgrading procedures:

- A-channel upgrade
- B-channel upgrade

Joint Synthetic Channels – *D*-value Representation

General form of Joint channel:

$$W_{a,b}(\underbrace{y_a, u_a, y_r}_{y_b} | u_a, u_b)$$

D-values for BMS channel:

$$d(y) = \frac{W(y|0) - W(y|1)}{W(y|0) + W(y|1)}$$

May switch to ***D*-value representation:**

$$W_{a,b}(y_a, u_a, d_b | u_a, u_b)$$

Lemma

$$W_{a,b}(y_a, u_a, y_r | u_a, u_b) \equiv W_{a,b}(y_a, u_a, d_b | u_a, u_b)$$

Symmetry

Question

For $W_b(y_a, u_a, d_b|u_b)$, what is $\overline{(y_a, u_a, d_b)}$?

Symmetry

Question

For $W_b(y_a, u_a, d_b | u_b)$, what is $\overline{(y_a, u_a, d_b)}$?

$\exists\, y_a^{(b)}$ such that

$$W_b(y_a, u_a, d_b | u_b) = W_b(y_a^{(b)}, u_a, -d_b | \bar{u}_b)$$

Symmetry

Question

For $W_b(y_a, u_a, d_b|u_b)$, what is $\overline{(y_a, u_a, d_b)}$?

$\exists\, y_a^{(b)}$ such that

$$W_b(y_a, u_a, d_b|u_b) = W_b(y_a^{(b)}, u_a, -d_b|\bar{u}_b)$$

New decoder decision the same for (y_a, u_a, d_b) and $(y_a^{(b)}, u_a, d_b)$

Symmetrization

- **Symmetrized** joint synthetic channel:

$$\mathring{y}_a \triangleq \{y_a, y_a^{(b)}\}$$

$$W_{a,b}(y_a, u_a, d_b | u_a, u_b)$$

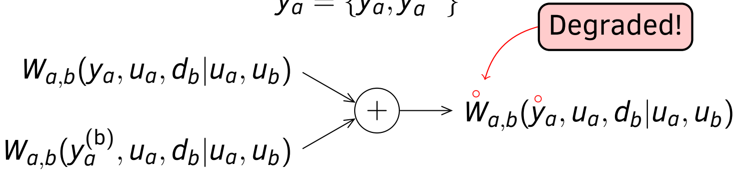
$$W_{a,b}(y_a^{(b)}, u_a, d_b | u_a, u_b)$$

$$\oplus \rightarrow \mathring{W}_{a,b}(\mathring{y}_a, u_a, d_b | u_a, u_b)$$

Symmetrization

- **Symmetrized** joint synthetic channel:

$$\mathring{y}_a \triangleq \{y_a, y_a^{(b)}\}$$



Symmetrization

- **Symmetrized** joint synthetic channel:

$$\mathring{y}_a \triangleq \{y_a, y_a^{(b)}\}$$

$$W_{a,b}(y_a, u_a, d_b | u_a, u_b)$$

$$W_{a,b}(y_a^{(b)}, u_a, d_b | u_a, u_b)$$

$$\oplus \rightarrow \mathring{W}_{a,b}(\mathring{y}_a, u_a, d_b | u_a, u_b)$$

Degraded!

- $P_e(\text{symmetrized}) = P_e(\text{non-symmetrized})$
- Symmetrization preserved under polarization/upgrading
  - Need to perform only once!

- **Symmetrized** joint synthetic channel:

$$\mathring{y}_a \triangleq \{y_a, y_a^{(b)}\}$$

$$W_{a,b}(y_a, u_a, d_b | u_a, u_b) \;\; \oplus \;\; W_{a,b}(y_a^{(b)}, u_a, d_b | u_a, u_b) \rightarrow \mathring{W}_{a,b}(\mathring{y}_a, u_a, d_b | u_a, u_b)$$

Degraded!

- $P_e(\text{symmetrized}) = P_e(\text{non-symmetrized})$
- Symmetrization preserved under polarization/upgrading
  - Need to perform only once!

- Decoupling Decomposition:

$$\mathring{W}_{a,b}(\mathring{y}_a, u_a, d_b | u_a, u_b) = W_a(\mathring{y}_a | u_a) W_2(d_b | u_b; \mathring{y}_a, u_a)$$

A-channel Upgrade

Decoupling decompositions:

$$W_{a,b}(y_a, u_a, d_b | u_a, u_b) = W_a(y_a | u_a) \cdot W_2(d_b | u_b; y_a, u_a)$$

$$Q_{a,b}(z_a, u_a, z_b | u_a, u_b) = Q_a(z_a | u_a) \cdot Q_2(z_b | u_b; z_a, u_a)$$

Theorem

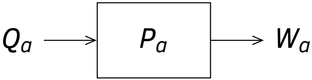
$Q_{a,b} \overset{p}{\succcurlyeq} W_{a,b}$ *if*

1. $Q_a(z_a | u_a) \succcurlyeq W_a(y_a | u_a)$ *with degrading channel* $P_a(y_a | z_a)$
2. $Q_2(z_b | u_b; z_a, u_a) \succcurlyeq W_2(d_b | u_b; y_a, u_a)$ *whenever* $P_a(y_a | z_a) > 0$

Per state, $Q_2 \succcurlyeq$ a family of channels

Why this isn't Enough

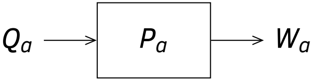
Step 1: $Q_a \succcurlyeq W_a$



$$P_a(y_a|z_{a0}) = \begin{cases} 0.25 & y_a = y_{a1} \\ 0.75 & y_a = y_{a2} \\ 0 & \text{otherwise} \end{cases}$$

Why this isn't Enough

Step 1: $Q_a \succcurlyeq W_a$



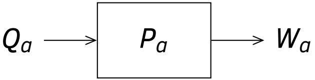
$$P_a(y_a|z_{a0}) = \begin{cases} 0.25 & y_a = y_{a1} \\ 0.75 & y_a = y_{a2} \\ 0 & \text{otherwise} \end{cases}$$

Assume:

$$W_2(d_b|u_b; y_{a1}, u_a) \leftarrow \text{BSC}(0.4)$$
$$W_2(d_b|u_b; y_{a2}, u_a) \leftarrow \text{BSC}(0.01)$$

Why this isn't Enough

Step 1: $Q_a \succcurlyeq W_a$



$$P_a(y_a|z_{a0}) = \begin{cases} 0.25 & y_a = y_{a1} \\ 0.75 & y_a = y_{a2} \\ 0 & \text{otherwise} \end{cases}$$

Assume:

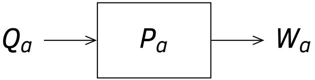
$$W_2(d_b|u_b; y_{a1}, u_a) \leftarrow \text{BSC}(0.4)$$
$$W_2(d_b|u_b; y_{a2}, u_a) \leftarrow \text{BSC}(0.01)$$

Step 2:

$$Q_2(z_b|u_b; z_{a0}, u_a) \leftarrow \text{BSC}(0.01)$$

Why this isn't Enough

Step 1: $Q_a \succcurlyeq W_a$



$$P_a(y_a|z_{a0}) = \begin{cases} 0.25 & y_a = y_{a1} \\ 0.75 & y_a = y_{a2} \\ 0 & \text{otherwise} \end{cases}$$

Assume:

$$W_2(d_b|u_b; y_{a1}, u_a) \leftarrow \text{BSC}(0.4)$$
$$W_2(d_b|u_b; y_{a2}, u_a) \leftarrow \text{BSC}(0.01)$$

Step 2:

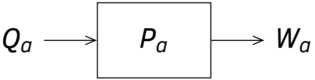
$$Q_2(z_b|u_b; z_{a0}, u_a) \leftarrow \text{BSC}(0.01)$$

Step 3:

$$Q_{a,b}(z_{a0}, u_a, z_b|u_a, u_b) = Q_a(z_{a0}|u_a) \cdot \underbrace{Q_2(z_b|u_b; z_{a0}, u_a)}_{\text{BSC}(0.01)}$$

Why this isn't Enough

Step 1: $Q_a \succcurlyeq W_a$



$$P_a(y_a|z_{a0}) = \begin{cases} 0.25 & y_a = y_{a1} \\ 0.75 & y_a = y_{a2} \\ 0 & \text{otherwise} \end{cases}$$

Assume:

$$W_2(d_b|u_b; y_{a1}, u_a) \leftarrow \text{BSC}(0.4)$$
$$W_2(d_b|u_b; y_{a2}, u_a) \leftarrow \text{BSC}(0.01)$$

Step 2:

$$Q_2(z_b|u_b; z_{a0}, u_a) \leftarrow \text{BSC}(0.01)$$

Step 3:

$$Q_{a,b}(z_{a0}, u_a, z_b|u_a, u_b) = Q_a(z_{a0}|u_a) \cdot \underbrace{Q_2(z_b|u_b; z_{a0}, u_a)}_{\text{BSC}(0.01)}$$

Problem due to different W_2

Upgrade-couple Transform

- Split a-channel symbols $y_a \to y_a^{i,j}$
- Such that:

$$y_a^{i,j} \Longrightarrow d_b = \begin{cases} \pm d_{bi} & u_a = 0 \\ \pm d_{bj} & u_a = 1 \end{cases}$$

- Upgrade-couple transform $\Rightarrow W_2$ the same for fixed i, j

Upgrade-couple Transform

- Split a-channel symbols $y_a \to y_a^{i,j}$
- Such that:

$$y_a^{i,j} \Longrightarrow d_b = \begin{cases} \pm d_{bi} & u_a = 0 \\ \pm d_{bj} & u_a = 1 \end{cases}$$

- Upgrade-couple transform $\Rightarrow W_2$ the same for fixed i, j

Theorem

1. *Upgrade-couple the channel*
2. *Confine upgrades to fixed i, j*

$\Rightarrow$ *Jointly upgrade $W_{a,b}$, and*

- *upgrade a-channel*
- *not change b-channel*

B-channel Upgrade

Joint channel: $W_{a,b}(y_a, u_a, d_b | u_a, u_b)$

Canonical b-channel marginal: $W_b^*(d_b | u_b)$

Theorem

$Q_b^*(z_b|u_b) \succcurlyeq W_b^*(d_b|u_b) \Rightarrow Q_{a,b}(y_a, u_a, z_b|u_a, u_b) \overset{p}{\succcurlyeq} W_{a,b}(y_a, u_a, d_b|u_a, u_b)$

with:

- *unchanged a-channel*
- *Same canonical b-channel marginal*

Full Algorithm

Input:

- BMS channel W
- a-channel transform list $\alpha_1, \alpha_2, \ldots, \alpha_n$
- b-channel transform list $\beta_1, \beta_2, \ldots, \beta_n$

$\alpha, \beta \in \{-, +\}$

Output: Lower bound on $P_e^{\mathrm{SC}}(W_{a_n,b_n})$

Steps:

1. Initialize: $W_{0,0} = W$
2. For $i = 1, \ldots, n$, do:
   - $W_{a_i,b_i} \leftarrow \texttt{JointlyPolarize}^{\circ}_{\alpha_i,\beta_i}(W_{a_{i-1},b_{i-1}})$
   - $W_{a_i,b_i} \leftarrow \texttt{B-channelUpgrade}(W_{a_i,b_i})$
   - $W_{a_i,b_i} \leftarrow \texttt{A-channelUpgrade}(W_{a_i,b_i})$
3. Compute: $P_e^*(W_{a_n,b_n})$

Limits b-channel alphabet size

Uses upgrade-couple

Limits a-channel alphabet size