Stronger Polarization for the Deletion Channel

Dar Arava, Ido Tal

Technion

Big picture first

A polar coding scheme for the deletion channel where the:

- Deletion channel has *constant* deletion probability δ
- Fix a hidden-Markov input distribution[1]
- Code rate converges to information rate
- Achieves capacity
- Error probability decays like $2^{-\Lambda^{\gamma}}$, where $\gamma < \frac{1}{2}$ and Λ is the codeword length
- Prior art [TPFV][2]: Same, apart for $\gamma < \frac{1}{3}$

[1] i.e., a function of an aperiodic, irreducible, finite-state Markov chain

[2] I. Tal, H. D. Pfister, A. Fazeli, A. Vardy, "Polar Codes for the Deletion Channel: Weak and Strong Polarization"

Key ideas

Encoding

- [TPFV]: break codeword into blocks using guard bands
- We do as well, but with different parameters

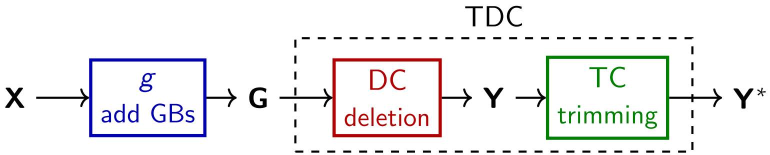
Decoding

- [TPFV]: use guard bands to break received word into blocks
- [TPFV]: build a trellis for each block
- We build a trellis for the whole received word

Analysis

- Use [TPFV] as "boot-strap"
- Use "walking-to-running" lemma

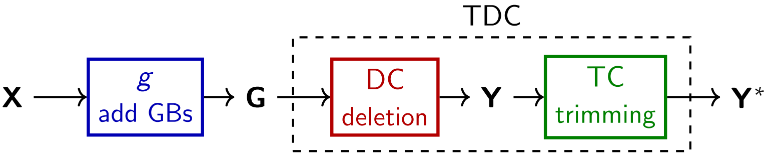
Our setting



- Trimming: $\mathbf{Y}^*$ removes leading and trailing 0's from $\mathbf{Y}$

$$(00110010)^* = (11001)$$

Our setting



- Trimming: $\mathbf{Y}^*$ removes leading and trailing 0's from $\mathbf{Y}$

$$(00110010)^* = (11001)$$

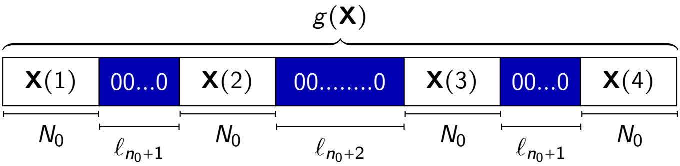
- Why add guard bands?
- Why trimming?

Guard bands

- We transform **X** into $g(\mathbf{X}) \triangleq g(\mathbf{X}, n_0, \xi)$
  - n_0 and $\xi > 0$ are fixed
  - **X** is of length $N = 2^n$
- The result: blocks of length $N_0 = 2^{n_0}$, interspaced by GBs

---

Example: $n = n_0 + 2$



Guard bands

- Recursion for $\mathbf{X} = \mathbf{X}_{\mathsf{I}} \odot \mathbf{X}_{\mathsf{II}}$:

$$n > n_0 \quad \Longrightarrow \quad g(\mathbf{X}) = \underbrace{g(\mathbf{X}_{\mathsf{I}})}_{\mathbf{G}_{\mathsf{I}}} \odot \underbrace{\overbrace{000\ldots00}^{\ell_n}}_{\triangleq \mathbf{G}_{\Delta}} \odot \underbrace{g(\mathbf{X}_{\mathsf{II}})}_{\triangleq \mathbf{G}_{\mathsf{II}}}$$

$$n \leq n_0 \quad \Longrightarrow \quad g(\mathbf{X}) = \mathbf{X} \text{ (stopping condition)}$$

Guard bands

- Recursion for $\mathbf{X} = \mathbf{X}_\text{I} \odot \mathbf{X}_\text{II}$:

$$n > n_0 \quad \Longrightarrow \quad g(\mathbf{X}) = \underbrace{g(\mathbf{X}_\text{I})}_{\mathbf{G}_\text{I}} \odot \underbrace{\overset{\ell_n}{\overbrace{000\dots00}}}_{\triangleq \mathbf{G}_\Delta} \odot \underbrace{g(\mathbf{X}_\text{II})}_{\triangleq \mathbf{G}_\text{II}}$$

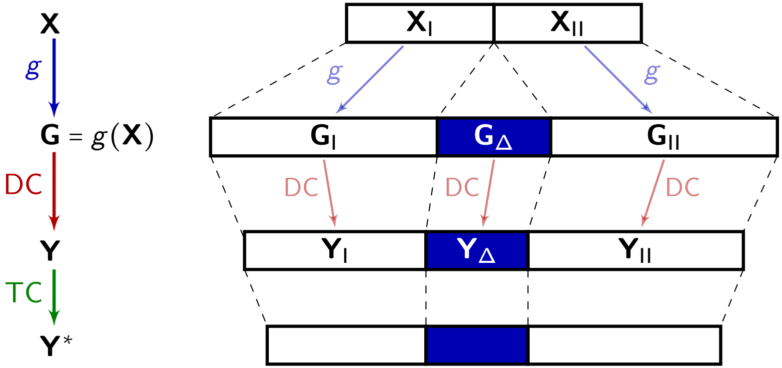
$$n \le n_0 \quad \Longrightarrow \quad g(\mathbf{X}) = \mathbf{X} \text{ (stopping condition)}$$

- Are not harmful: Middle GB length is $\ell_n \triangleq \left\lfloor 2^{(1-\xi)(n-1)} \right\rfloor$, s.t. the effect on the rate vanishes for a large enough n_0:

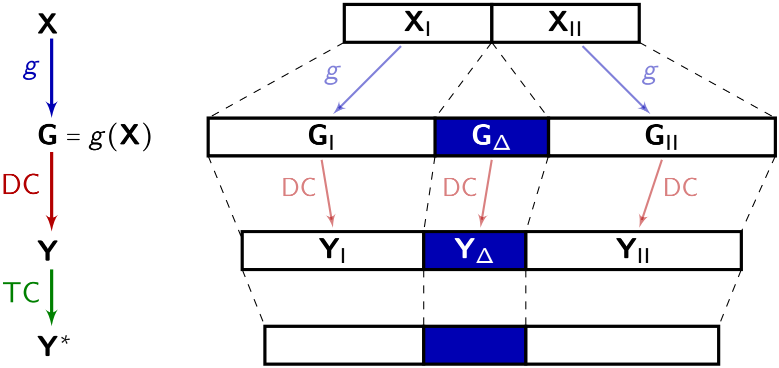
$$\frac{N}{\Lambda} \triangleq \frac{|\mathbf{X}|}{|g(\mathbf{X})|} \xrightarrow{n\to\infty} 1$$

- Will come in handy: With GBs, it is easier to separate the output to independent blocks.

Guard bands, deletion, and trimming



Guard bands, deletion, and trimming



More notation:

- The Arıkan transform of $\mathbf{X}$ is $\mathbf{U} = \mathcal{A}(\mathbf{X})$
- $\mathbf{V} \triangleq \mathcal{A}(\mathbf{X}_{\mathrm{I}})$ and $\mathbf{V}' \triangleq \mathcal{A}(\mathbf{X}_{\mathrm{II}})$
- $U_{2j-1} = V_j + V_j'$ ('−') and $U_{2j} = V_j'$ ('+')

Evolution of Z for the trim-deletion channel

First, a relation between these two Bhattacharyya parameters:

<ins>n polarization steps:</ins>

$$Z(U_i|U_1^{i-1},\mathbf{Y}^*)$$

$$\mathbf{X} = \mathcal{A}(\mathbf{U}) \xrightarrow{g} \mathbf{G} \xrightarrow{\text{DC}} \mathbf{Y} \xrightarrow{\text{TC}} \mathbf{Y}^*$$

<ins>$n-1$ polarization steps:</ins>

$$Z(V_j|V_1^{j-1},\mathbf{Y}_{\mathrm{I}}^*) = Z(V'_j|V'^{j-1}_1,\mathbf{Y}_{\mathrm{II}}^*)$$

$$\mathbf{X}_{\mathrm{I}} = \mathcal{A}(\mathbf{V}) \xrightarrow{g} \mathbf{G}_{\mathrm{I}} \xrightarrow{\text{DC}} \mathbf{Y}_{\mathrm{I}} \xrightarrow{\text{TC}} \mathbf{Y}_{\mathrm{I}}^*$$

Evolution of Z for the trim-deletion channel

Lemma (evolution of Z): Some fine print. There exist $m_0^{\mathrm{th}}(\xi)$ and $m^{\mathrm{th}}(\xi, \delta)$ s.t. for $n_0 \geq m_0^{\mathrm{th}}$ and all $n \geq \max\{m^{\mathrm{th}}, n_0 + 1\}$ the following holds. Let $1 \leq i \leq N$ and $j = \lfloor (i+1)/2 \rfloor$. Then,

$$Z(U_i | U_1^{i-1}, \mathbf{Y}^*) \leq \frac{3}{2} N \cdot Z(U_i | U_1^{i-1}, \mathbf{Y}_{\mathrm{I}}^*, \mathbf{Y}_{\mathrm{II}}^*) + 2^{-N^{\frac{2}{3}}}$$

$$\leq \begin{cases} \frac{3}{2} N \cdot 2 \cdot Z(V_j | V_1^{j-1}, \mathbf{Y}_{\mathrm{I}}^*) + 2^{-N^{\frac{2}{3}}} & \text{if } b_n = 0 \text{ (`}-\text{')} \\ \frac{3}{2} N \cdot Z(V_j | V_1^{j-1}, \mathbf{Y}_{\mathrm{I}}^*)^2 + 2^{-N^{\frac{2}{3}}} & \text{if } b_n = 1 \text{ (`}+\text{')} \end{cases}$$

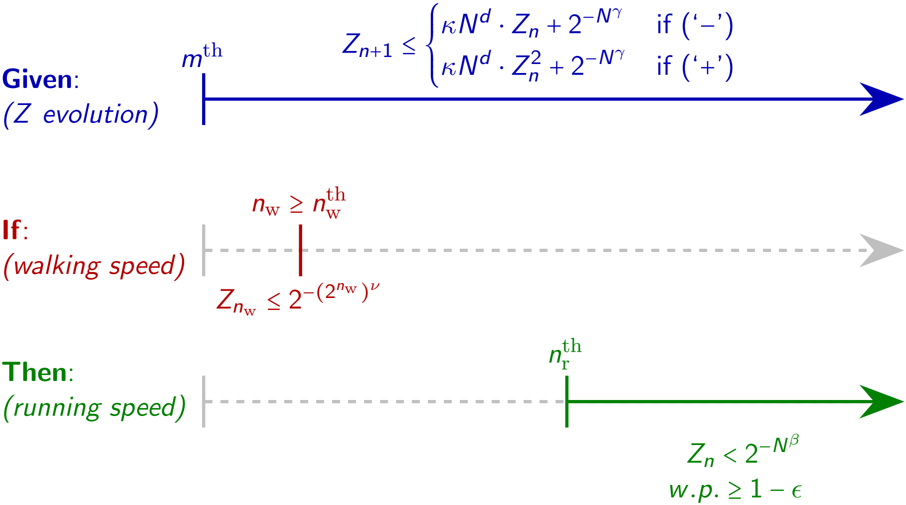
For binary $b_1, b_2, \ldots, b_n$, $\quad i = 1 + \sum_{k=1}^{n} b_k 2^{n-k}$.

A corresponding random process

Let $B_1, B_2, \ldots$ be i.i.d. uniformly distributed Bernoulli random variables. Fix constants $\kappa \geq 1, d \geq 0, \gamma > \frac{1}{2}$ and $m^{\mathrm{th}} > 0$. Let $Z_0, Z_1, Z_2, \ldots$ be a random process s.t. for all $n \geq m^{\mathrm{th}}$,

$$Z_{n+1} \leq \begin{cases} \kappa N^d \cdot Z_n + 2^{-N^\gamma} & \text{if } B_{n+1} = 0 \text{ (`$-$')} \\ \kappa N^d \cdot Z_n^2 + 2^{-N^\gamma} & \text{if } B_{n+1} = 1 \text{ (`$+$')} . \end{cases}$$

Walking-to-running lemma



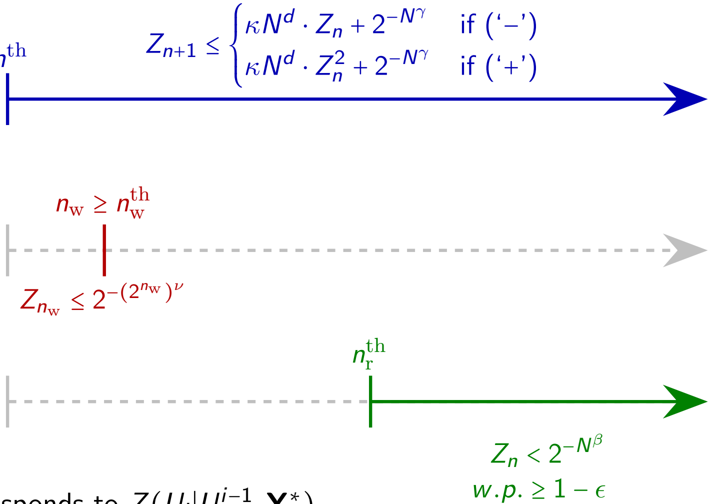
$\nu > 0$, $\beta \in (0, 1/2)$

Walking-to-running lemma

Given: *(Z evolution)*

If: *(walking speed)*

Then: *(running speed)*



In our case:

- Z_n corresponds to $Z(U_i|U_1^{i-1}, \mathbf{Y}^*)$.
- Walking speed is by [TPFV], and the evolution of $Z(U_i|U_1^{i-1}, \mathbf{Y}^*)$ was stated in the previous lemma.

Walking-to-running lemma

Lemma (walking-to-running): Let $Z_0, Z_1, Z_2, \ldots$ be a random process s.t. for all $n \geq m^{\mathrm{th}}$,

$$Z_{n+1} \leq \begin{cases} \kappa N^d \cdot Z_n + 2^{-N^\gamma} & \text{if } B_{n+1} = 0 \text{ (`}-\text{')} \\ \kappa N^d \cdot Z_n^2 + 2^{-N^\gamma} & \text{if } B_{n+1} = 1 \text{ (`}+\text{')} \end{cases} .$$

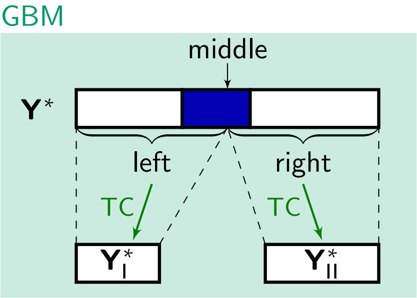
Fix $\beta \in (0, \frac{1}{2})$, the "running speed" parameter, and $\nu > 0$, the "walking speed" parameter. For all $\epsilon > 0$ there exists a threshold $n_{\mathrm{w}}^{\mathrm{th}} = n_{\mathrm{w}}^{\mathrm{th}}(\epsilon, \beta, \nu, \kappa, d, \gamma, m^{\mathrm{th}}) \geq m^{\mathrm{th}}$ such that if for some $n_{\mathrm{w}} \geq n_{\mathrm{w}}^{\mathrm{th}}$ we are assured "walking speed":

$$Z_{n_{\mathrm{w}}} \leq 2^{-(2^{n_{\mathrm{w}}})^\nu} ,$$

then there exists $n_{\mathrm{r}}^{\mathrm{th}} = n_{\mathrm{r}}^{\mathrm{th}}(\epsilon, \beta, \nu, \kappa, d, n_{\mathrm{w}}) > n_{\mathrm{w}}$ such that above this threshold, with high probability, we are indefinitely at "running speed":

$$\mathbb{P}\left(Z_n < 2^{-N^\beta}, \quad \forall n \geq n_{\mathrm{r}}^{\mathrm{th}}\right) \geq 1 - \epsilon .$$

The Guard Band in the Middle (GBM) event

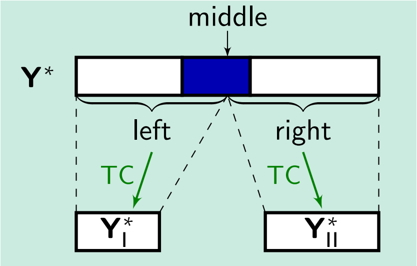


Under the GBM event:

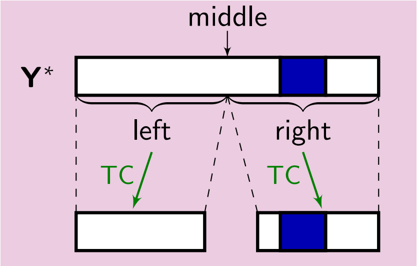
$$\mathbf{Y}^* = \mathbf{Y}_{\mathrm{I}}^* \odot \overbrace{000\ldots00}^{L_0} \odot \mathbf{Y}_{\mathrm{II}}^*$$

The Guard Band in the Middle (GBM) event

GBM



¬GBM



Under the GBM event:

$$\mathbf{Y}^* = \mathbf{Y}_{\mathrm{I}}^* \odot \overbrace{000\ldots00}^{L_0} \odot \mathbf{Y}_{\mathrm{II}}^*$$

this is a “bad” event...

Bounding Z

We will show:

$$Z(U_i|U_1^{i-1},\mathbf{Y}^*) \leq \underbrace{Z(U_i|U_1^{i-1},\mathbf{Y}^*,\mathrm{GBM})}_{\leq\ \frac{3}{2}N\cdot\ Z(U_i|U_1^{i-1},\mathbf{Y}^*_{\mathrm{I}},\mathbf{Y}^*_{\mathrm{II}})} + \underbrace{\sqrt{\mathbb{P}(\neg\mathrm{GBM})}}_{\leq\ 2^{-N^{2/3}}}$$

Bounding Z

$$Z(U_i|U_1^{i-1},\mathbf{Y}^*) = \sum_{u_1^{i-1},\mathbf{y}^*} \sqrt{\begin{array}{l}\mathbb{P}(U_i = 0, U_1^{i-1} = u_1^{i-1}, \mathbf{Y}^* = \mathbf{y}^*) \\ \times\, \mathbb{P}(U_i = 1, U_1^{i-1} = u_1^{i-1}, \mathbf{Y}^* = \mathbf{y}^*)\end{array}}$$

Bounding Z

$$Z(U_i | U_1^{i-1}, \mathbf{Y}^*) = \sum_{u_1^{i-1}, \mathbf{y}^*} \sqrt{\begin{array}{l} \mathbb{P}(U_i = 0, U_1^{i-1} = u_1^{i-1}, \mathbf{Y}^* = \mathbf{y}^*) \\ \times\, \mathbb{P}(U_i = 1, U_1^{i-1} = u_1^{i-1}, \mathbf{Y}^* = \mathbf{y}^*) \end{array}}$$

The law of total probability:

$$= \sum_{u_1^{i-1}, \mathbf{y}} \sqrt{\begin{array}{l} \left(\mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \text{GBM}) + \mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \neg\text{GBM})\right) \\ \times \left(\mathbb{P}(1, u_1^{i-1}, \mathbf{y}, \text{GBM}) + \mathbb{P}(1, u_1^{i-1}, \mathbf{y}, \neg\text{GBM})\right) \end{array}}$$

Rearranging:

$$\leq \sum_{u_1^{i-1}, \mathbf{y}} \sqrt{\mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \text{GBM}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{y}, \text{GBM})}$$

$$+ \sum_{u_1^{i-1}, \mathbf{y}} \sqrt{\begin{array}{l} \mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \text{GBM}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{y}, \neg\text{GBM}) \\ +\, \mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \neg\text{GBM}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{y}) \end{array}}$$

Bounding the first sum

$$\sum_{u_1^{i-1},\mathbf{y}} \sqrt{\mathbb{P}(0, u_1^{i-1}, \mathbf{Y}^* = \mathbf{y}, \text{GBM}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{Y}^* = \mathbf{y}, \text{GBM})}$$

Bounding the first sum

$$\sum_{u_1^{i-1},\mathbf{y}} \sqrt{\mathbb{P}(0, u_1^{i-1}, \mathbf{Y}^* = \mathbf{y}, \text{GBM}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{Y}^* = \mathbf{y}, \text{GBM})}$$

Under GBM, knowing $\mathbf{Y}^*$ is equivalent to knowing $\mathbf{Y}_\mathsf{I}^*, \mathbf{Y}_\mathsf{II}^*$ and L_0:

$$= \sum_{\ell=1}^{\frac{3}{2}N} \sum_{u_1^{i-1},\mathbf{y}',\mathbf{y}''} \sqrt{\begin{array}{l}\mathbb{P}(0, u_1^{i-1}, \mathbf{Y}_\mathsf{I}^* = \mathbf{y}', \mathbf{Y}_\mathsf{II}^* = \mathbf{y}'', L_0 = \ell, \text{GBM}) \\ \times \mathbb{P}(1, u_1^{i-1}, \mathbf{Y}_\mathsf{I}^* = \mathbf{y}', \mathbf{Y}_\mathsf{II}^* = \mathbf{y}'', L_0 = \ell, \text{GBM})\end{array}}$$

"Throwing away" $\{L_0 = \ell, \text{GBM}\}$:

$$\leq \frac{3}{2}N \cdot \underbrace{\sum_{u_1^{i-1},\mathbf{y}',\mathbf{y}''} \sqrt{\begin{array}{l}\mathbb{P}(0, u_1^{i-1}, \mathbf{Y}_\mathsf{I}^* = \mathbf{y}', \mathbf{Y}_\mathsf{II}^* = \mathbf{y}'') \\ \times \mathbb{P}(1, u_1^{i-1}, \mathbf{Y}_\mathsf{I}^* = \mathbf{y}', \mathbf{Y}_\mathsf{II}^* = \mathbf{y}'')\end{array}}}_{=Z(U_i|U_1^{i-1},\mathbf{Y}_\mathsf{I}^*,\mathbf{Y}_\mathsf{II}^*)}$$

Bounding the second sum

$$\sum_{u_1^{i-1},\mathbf{y}} \sqrt{\begin{array}{l}\mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \mathsf{GBM}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{y}, \neg\mathsf{GBM}) \\ + \mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \neg\mathsf{GBM}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{y})\end{array}}$$

Bounding the second sum

$$\sum_{u_1^{i-1},\mathbf{y}} \sqrt{\begin{array}{l}\mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \mathrm{GBM}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{y}, \neg\mathrm{GBM}) \\ + \mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \neg\mathrm{GBM}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{y})\end{array}}$$

Relaxing constraints:

$$\leq \sum_{u_1^{i-1},\mathbf{y}} \sqrt{\begin{array}{l}\mathbb{P}(u_1^{i-1}, \mathbf{y}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{y}, \neg\mathrm{GBM}) \\ + \mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \neg\mathrm{GBM}) \cdot \mathbb{P}(u_1^{i-1}, \mathbf{y})\end{array}}$$

Total probability:

$$= \sum_{u_1^{i-1},\mathbf{y}} \sqrt{\mathbb{P}(u_1^{i-1}, \mathbf{y}) \cdot \mathbb{P}(u_1^{i-1}, \mathbf{y}, \neg\mathrm{GBM})}$$

Jensen:

$$\leq \sqrt{\mathbb{P}(\neg\mathrm{GBM})}$$

Bounding the second sum

$$\sum_{u_1^{i-1},\mathbf{y}} \sqrt{\begin{array}{l}\mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \mathsf{GBM}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{y}, \neg\mathsf{GBM}) \\ +\, \mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \neg\mathsf{GBM}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{y})\end{array}}$$

Relaxing constraints:

$$\leq \sum_{u_1^{i-1},\mathbf{y}} \sqrt{\begin{array}{l}\mathbb{P}(u_1^{i-1}, \mathbf{y}) \cdot \mathbb{P}(1, u_1^{i-1}, \mathbf{y}, \neg\mathsf{GBM}) \\ +\, \mathbb{P}(0, u_1^{i-1}, \mathbf{y}, \neg\mathsf{GBM}) \cdot \mathbb{P}(u_1^{i-1}, \mathbf{y})\end{array}}$$

Total probability:

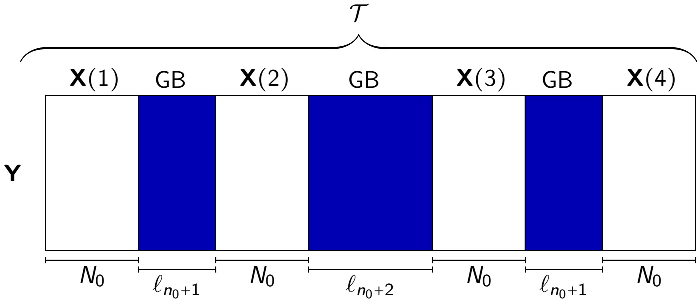
$$= \sum_{u_1^{i-1},\mathbf{y}} \sqrt{\mathbb{P}(u_1^{i-1}, \mathbf{y}) \cdot \mathbb{P}(u_1^{i-1}, \mathbf{y}, \neg\mathsf{GBM})}$$

Jensen: $\leq \sqrt{\mathbb{P}(\neg\mathsf{GBM})}$

this is small, since GBM is the 'typical' case: $\overset{\text{[TPFV]}}{\leq} 2^{-N^{\frac{2}{3}}}$

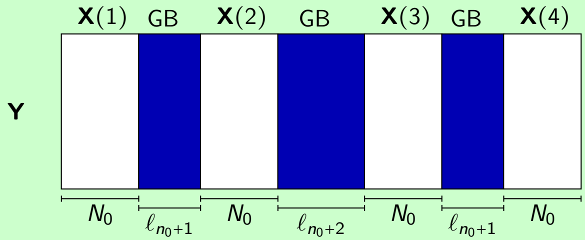
Decoding

- Decoding complexity $O(\Lambda^4)$
- Decoding is similar to [TPFV], but the trellis corresponds to the whole received word (including the guard bands):

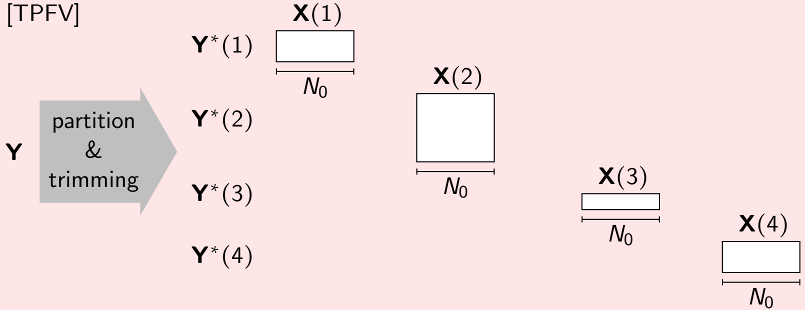


Decoding

Our decoder



[TPFV]



Simulation Results