EXPLICIT DESIGN OF PROVABLY COVERT CHANNEL CODES

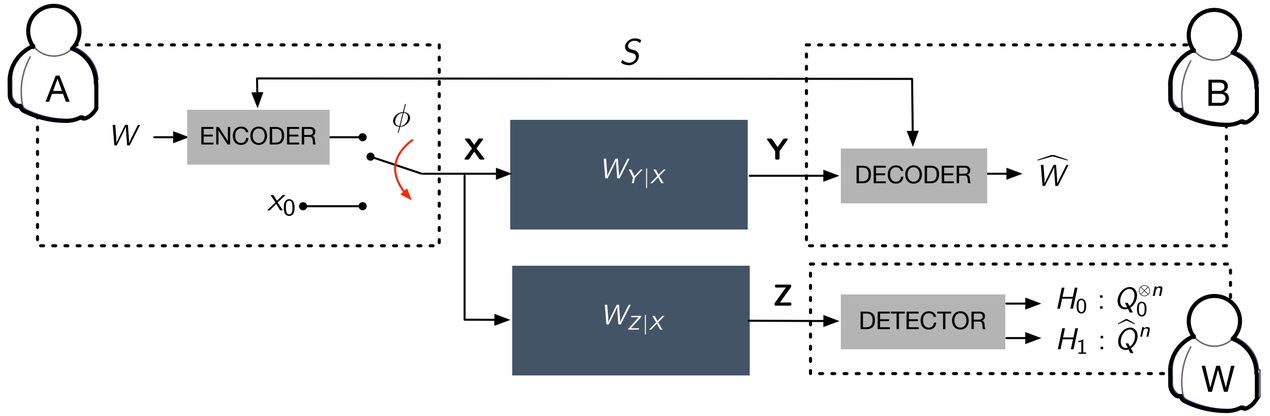
IEEE International Symposium on Information Theory ▪ July 2021

Shi-Yuan Wang and Matthieu R. Bloch
School of Electrical and Computer Engineering

- **OBJECTIVE**
  - Construct a **provably covert** channel code while using secret key bits **efficiently**
  - Based on MLC (Multi-level Coded)-PPM (Pulse-Position Modulation) **[Kadampot-Tahmasbi-Bloch'20]**
  - MLC-PPM + Polar code **[Arikan'10, Chou-Bloch'16]** + Invertible extractor **[Bellare-Tessaro'12, Chou'18]**

- **PRACTICAL CHALLENGES**
  - Perform source polarization to identify bit-sources efficiently
  - Carefully design invertible extractor to control trade-off between secret key usage and covertness

- **CONTRIBUTIONS**
  - Explicit polar codes for joint reliability and channel resolvability
  - Source polarization algorithm with upgrading and degrading
  - Channel resolvability codes based on invertible extractor
  - Explicit example **achieving** covert communication with **efficient use of secret key bits**



- Switch $\phi \in \{0, 1\}$ controls transmission state at Alice
- **Innocent symbol** x_0 **:** *absence of communications*
  - Induces distributions $P_0 \triangleq W_{Y|X=x_0}$ and $Q_0 \triangleq W_{Z|X=x_0}$
- **Code** $\mathcal{C}$ **:** *occurrence of communications*
  - induces distribution $\widehat{Q}^n$ at Willie

- **RELIABILITY:** Bob reliably recovers the message

- **DETECTION:** Willie (passive warden) distinguishes
  - Hypothesis H_0 : $Q_0^{\otimes n}$ (*absence of communications*)
  - Hypothesis H_1 : $\widehat{Q}^n$ (*occurrence of communications*)

- **GOAL:** make sure Willie's test close to **blind test**

- **RELIABILITY METRIC**
  - Maximal average probability of error $P_e \triangleq \max_s \mathbb{P}\left(W \neq \widehat{W} | S = s, \phi = 1\right)$

- **COVERTNESS METRIC**
  - For any test of Willie $T(\mathbf{Z}^n)$, the probabilities of false alarm α and missed detection β satisfy

$$1 \geqslant \alpha + \beta \geqslant 1 - \mathbb{V}(\widehat{Q}^n, Q_0^{\otimes n}) \geqslant 1 - \sqrt{\mathbb{D}(\widehat{Q}^n \,\|\, Q_0^{\otimes n})}$$

  - Any test is close to **blind test** when $\mathbb{V}(\widehat{Q}^n, Q_0^{\otimes n})$ or $\mathbb{D}(\widehat{Q}^n \,\|\, Q_0^{\otimes n})$ is negligible

- **OPTIMAL THROUGHPUT: square-root law**
  - Covert code of length n such that $\lim_{n\to\infty} P_e = 0$ and $\mathbb{V}(\widehat{Q}^n, Q_0^{\otimes n}) \leqslant \delta$ characterized by number of message bits $M(n, \delta)$ and secret key bits $K(n, \delta)$

> **THEOREM: COVERT CAPACITY OF BI-DMC [Bloch'16, Tahmasbi-Bloch'19]**
>
> Optimal number of message bits
>
> $$\lim_{n\to\infty} \frac{\log M^*(n, \delta)}{\sqrt{n}} = \frac{2\mathbb{D}(P_1 \,\|\, P_0)}{\sqrt{\chi_2(Q_1 \| Q_0)}} Q^{-1}\left(\frac{1-\delta}{2}\right)$$
>
> is achievable with number of secret key bits
>
> $$\lim_{n\to\infty} \frac{\log K(n, \delta)}{\sqrt{n}} = \frac{2}{\sqrt{\chi_2(Q_1 \| Q_0)}} Q^{-1}\left(\frac{1-\delta}{2}\right) [\mathbb{D}(Q_1 \,\|\, Q_0) - \mathbb{D}(P_1 \,\|\, P_0)]^+$$

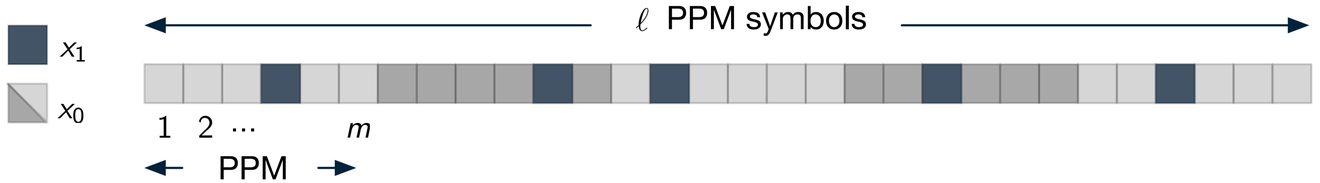
- **FROM INFORMATION-THEORETIC ANALYSIS: sparse signal**
  - Codebook should consist of "low-weight" codewords with fraction $\mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$ of non-zero symbols

- Concatenated codes for DMCs **[Zhang et al.'16]**
  - Low-complexity construction exists with Reed-Solomon outer code
  - Construction of inner code is **not explicit**

- Polar codes for asynchronous covert communication **[Freche-Bloch-Barret'17]**
  - Construct a polar code with joint reliability and resolvability
  - Low-weight nature of codewords **affects speed of polarization**

- Concatenated codes with graph-based non-linear codes **[Larmarca-Matas'19]**
  - Explicit construction of both outer and inner code with graph-based decoding
  - **No theoretical guarantee** to achieve optimality

- **IDEA:** ensure covertness through signal modulation

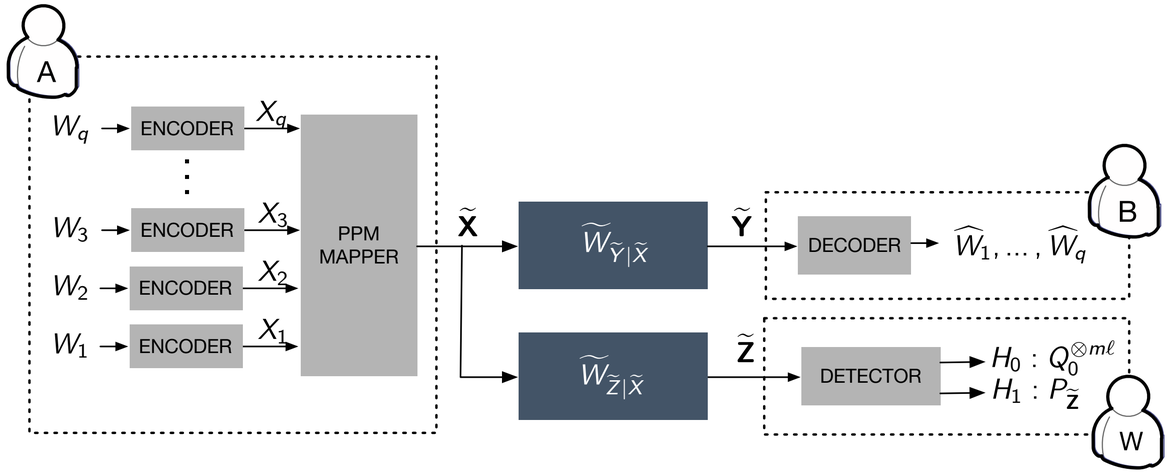


- PPM symbols of order m: $\{\tilde{x}_i\}_{i=1}^m$ with $x_i = x_1$ and $x_j = x_0$ if $j \neq i$
  - Induces "super-channel" $(\widetilde{\mathcal{X}}, \widetilde{W}_{\widetilde{Y}|\widetilde{X}}, \widetilde{\mathcal{Y}}, \widetilde{W}_{\widetilde{Z}|\widetilde{X}}, \widetilde{\mathcal{Z}})$, where $\widetilde{\mathcal{X}} \triangleq \{\tilde{x}_i\}_{i=1}^m$, $\widetilde{\mathcal{Y}} \triangleq \mathcal{Y}^m$, $\widetilde{\mathcal{Z}} \triangleq \mathcal{Z}^m$, $\widetilde{W}_{\widetilde{Y}|\widetilde{X}} \triangleq W_{Y|X}^{\otimes m}$ and $\widetilde{W}_{\widetilde{Z}|\widetilde{X}} \triangleq W_{Z|X}^{\otimes m}$
  - i.i.d PPM process: $Q_{\mathrm{PPM}}^m(z^m) \triangleq \dfrac{1}{m}\displaystyle\sum_{i=1}^m \widetilde{W}_{\widetilde{Z}|\widetilde{X}}(z^m|\tilde{x}_i)$

> **THEOREM:** COVERT CAPACITY WITH PPM [Bloch-Guha'17]
>
> By choosing $\ell \triangleq \left\lceil \dfrac{2\delta}{\chi_2\left(Q_1 \| Q_0\right)} m \right\rceil$,
>
> there exists a code with ℓ super-channel uses over PPM of order m such that
>
> $$\mathbb{D}((Q_{\mathrm{PPM}}^m)^{\otimes \ell} \,\|\, Q_0^{\otimes m\ell}) \leqslant \delta + \mathcal{O}(\frac{1}{m})$$

  - Overall block length $n \triangleq m\ell$

- Design a code of length ℓ to achieve **reliability** and **resolvability** jointly
  - For a fixed key value S, encoder-decoder pair forms reliable code
  - Randomization over key and message ensures resolvability for Willie's channel
    - Code $\mathcal{C}$ induces output distribution at Willie's terminal $P_{\widetilde{\mathbf{Z}}}(\mathbf{z}) \triangleq \frac{1}{|\mathcal{C}|} \sum_{\mathbf{c} \in \mathcal{C}} \widetilde{W}_{\widetilde{Z}|\widetilde{X}}^{\otimes \ell}(\mathbf{z}|\mathbf{c})$
    - Resolvability code ensures $\mathbb{V}(P_{\widetilde{\mathbf{Z}}}, (Q_{\mathrm{PPM}}^{m})^{\otimes \ell}) \leqslant \bar{\delta}$
    - Combined with $\mathbb{V}((Q_{\mathrm{PPM}}^{m})^{\otimes \ell}, Q_0^{\otimes m\ell}) \leqslant \sqrt{\mathbb{D}((Q_{\mathrm{PPM}}^{m})^{\otimes \ell} \,\|\, Q_0^{\otimes m\ell})\,/2}$ to ensure covertness
- **CHALLENGE**
  - Alphabet size of $\widetilde{\mathcal{X}}$ grows linearly with block length ℓ
  - Hard to obtain a **fixed-rate** design
- **SOLUTION:** multi-level coding



	1	2	3	4	5	6	7	8
X_1	0	1	0	1	0	1	0	1
X_2	0	0	1	1	0	0	1	1
X_3	0	0	0	0	1	1	1	1

- Number of levels $q = \log_2 m$ and index PPM symbol by q bits
- Use simple labeling strategy and multistage decoding (MSD)
- Decompose PPM super-channel into q binary-input equivalent channels (multiple-access channel)
- **OBSERVATION:** number of level grows logarithmically with block length ℓ

- **Channel invariance**

> **LEMMA:** CHANNEL INVARIANCE UNDER MSD [Kadampot-Tahmasbi-Bloch'20]
>
> Under multistage decoding starting from level q, the channel at level j is given by
>
> $$W_{Y|X}^{(j)}(y^{2^j}|x_j) = \frac{1}{2^{j-1}} \sum_{k=2^{j-1}x_j+1}^{2^{j-1}(x_j+1)} P_0^{\otimes 2^j}(y^{2^j}) \frac{P_1(y_k)}{P_0(y_k)}$$

  - Each equivalent channel is **fixed** when block length ℓ and number of levels q varies
  - Enable **fixed-rate** code design

- Achieving reliability and resolvability **at each level** guarantees overall performance
  - Design channel reliability and resolvability code for each level **independently** (component code design)

- Capacity concentrates at **lower levels**

> **LEMMA:** CHANNEL CAPACITY OF EACH LEVEL [Kadampot-Tahmasbi-Bloch'20]
> Capacity of j -th equivalent channel satisfies
> $$\mathbb{I}\left(X_j; \widetilde{Y} | X_{j+1:q}\right) \leqslant 2^{-j} \chi_2(P_1 \| P_0) + \mathcal{O}(2^{-2j})$$

  - Use lower levels to transmit message (joint reliability and resolvability)
  - Use higher levels to achieve resolvability

- **TAKE AWAY:** design **fixed-rate** component code for each level **independently**

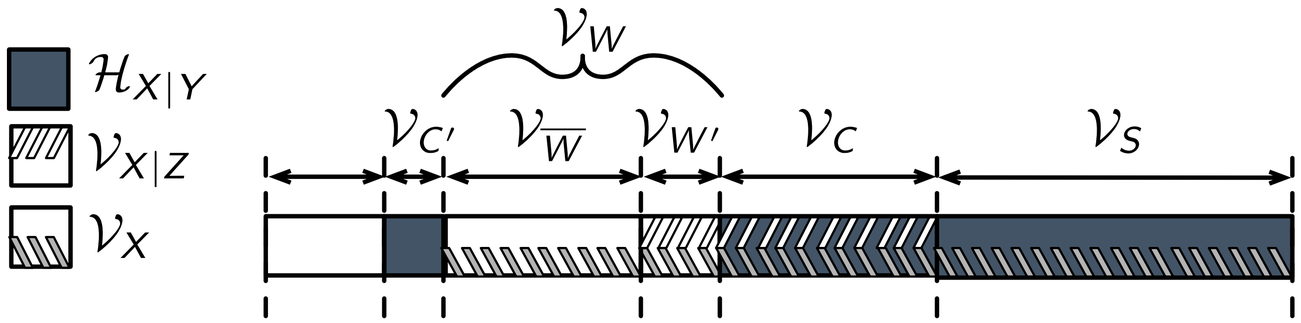
- Polar codes with **binning [Chou-Bloch'16]**
  - Based on source polarization **[Arikan'10]**

- Define DMSs as each equivalent channel interacting with input distribution $(\mathcal{X}, \mathcal{Y}, P_X W_{Y|X}^{(j)})$ and$(\mathcal{X}, \mathcal{Z}, P_X W_{Z|X}^{(j)})$

- **POLARIZATION SET:** Let $U^\ell \triangleq X^\ell G_\ell$ be polar transform of X^ℓ with polar transform G_ℓ . Set $\eta_\ell \triangleq 2^{-\ell^\beta}$ where $\beta \in (0, 1/2)$.
  - $\mathcal{V}_X \triangleq \{i \in [\![1, \ell]\!] : \mathbb{H}(U_i|U^{i-1}) \geqslant 1 - \eta_\ell\}$ (Very high-entropy set)
  - $\mathcal{H}_{X|Y} \triangleq \{i \in [\![1, \ell]\!] : \mathbb{H}(U_i|U^{i-1}Y^\ell) \geqslant \eta_\ell\}$ (High-entropy set)
  - $\mathcal{V}_{X|Z} \triangleq \{i \in [\![1, \ell]\!] : \mathbb{H}(U_i|U^{i-1}Z^\ell) \geqslant 1 - \eta_\ell\}$
  - Polarization sets are building blocks for polar codes with binning

- **CHALLENGE:** intractable complexity
  - Identifying these sets requires knowledge of $P_\ell^i(u_i, u^{i-1}y^\ell)$ **(bit source)**
  - Alphabet size for i-th bit source is $2^i\,|\mathcal{Y}|^\ell$

- **SOLUTION:** source upgrading and degrading
  - Modified from channel upgrading and degrading **[Tal-Vardy'13,Kartowsky-Tal'19]**
  - Properly quantize alphabet during polarization
  - Estimate conditional entropy



- Construct joint reliability and resolvability polar code by combining polarization sets **[Freche-Bloch-Barret'17]**
- Covert message set $\mathcal{V}_W \triangleq \mathcal{V}_{W'} \cup \mathcal{V}_{\overline{W}}$
  - $\mathcal{V}_{W'} \triangleq \mathcal{H}^c_{X|Y} \cap \mathcal{V}_{X|Z}$ carries message bits secret from Willie
- $\mathcal{H}_{X|Y}$ Plays similar role of "frozen bits" in conventional polar channel codes
  - Require secret key bits to secure $\mathcal{V}_{C'}$ and $\mathcal{V}_S$
  - $\mathcal{V}_C \triangleq \mathcal{H}_{X|Y} \cap \mathcal{V}_{X|Z}$ is secret from Willie

- Even if Willie's channel is **degraded**, we still need secret key bits
  - Theoretical analysis shows **existence** of scheme requiring **asymptotically negligible** amount of secret key bits
  - **Chaining [Sasoglu-Vardy'13, Chou-Bloch'16]:** reusing part of $\mathcal{V}_{W'}$ in previous transmission as secret key bits
  - **Efficient** scheme generates more covert and secret message bits than secret key bits it consumes

- We analyze efficient schemes when only single transmission occurs in the present work

- Multiple-block transmission
  - $\mathcal{V}_C$ and $\mathcal{V}_{W'}$ are not **perfectly secret**
  - Information leakage $\mathbb{I}\left(\widetilde{Z}; W'C\right)$ is non-zero and introduce **dependency between transmissions**
  - Covertness analysis must account for this **dependency**

- **TAKE AWAY:** for single transmission, our code design does not contribute to covertness metric

- **RECALL:** Higher levels are noisy and we use these levels for resolvability
- All higher levels are equivalent to one **symmetric** channel

> **LEMMA:** EQUIVALENT CHANNEL FOR ALL HIGHER LEVELS [Kadampot-Tahmasbi-Bloch'20]
> Equivalent channel from level $u+1$ to q is given by
> $$W_{\widetilde{Z}|X}^{u+1:q}(\widetilde{z}|x_{u+1:q}) = \frac{1}{2^u} \sum_{k \in \mathcal{A}^q(x_{u+1:q})} Q_0^{\otimes 2^q}(\widetilde{z}) \frac{Q_1(z_k)}{Q_0(z_k)}$$

  - Design one resolvability code for this channel

- Inverter-Extractor pair: let Ext be two-universal with seed $\mathbf{s} \in \mathbb{S}$ and bin index $\mathbf{b} \in \mathbb{F}_2^{(q-u)\ell-k}$

$$\text{Ext} : \mathbb{S} \times \mathbb{F}_2^{(q-u)\ell} \to \mathbb{F}_2^{(q-u)\ell-k} : (\mathbf{s}, \mathbf{x}_{u+1:q}) \mapsto \mathbf{b}$$

$$\text{Inv} : \mathbb{S} \times \mathbb{F}_2^{(q-u)\ell-k} \times \mathbb{F}_2^{k} \to \mathbb{F}_2^{(q-u)\ell} : (\mathbf{s}, \mathbf{b}, \mathbf{r}) \mapsto \mathbf{x}_{u+1:q}$$

  - Let $\mathcal{P}_{\mathbf{s},\mathbf{b}} \triangleq \{\mathbf{x}_{u+1:q} : \text{Ext}(\mathbf{s}, \mathbf{x}_{u+1:q} = \mathbf{b})\}$ be pre-image under $\mathbf{s}$ and $\mathbf{b}$
  - If $\mathbf{r} \in \mathbb{F}_2^k$ is uniform randomness bits, Inv equivalently selects $\mathbf{x}_{u+1:q}$ from $\mathcal{P}_{\mathbf{s},\mathbf{b}}$ **uniformly at random**

- **ENCODER:** $\phi : \mathbb{F}_2^k \to \mathbb{F}_2^{(q-u)\ell} : \mathbf{r} \mapsto \mathsf{Inv}(\mathbf{s}, \mathbf{b}, \mathbf{r})$

> **LEMMA:** LEFTOVER HASH LEMMA FOR RESOLVABILITY CODE
>
> Encoder ϕ with secret key rate $R_{u+1:q} \triangleq k/\ell$ satisfies
>
> $$\mathbb{E}_{\mathbf{S}}\{\mathbb{V}(P_{\widetilde{\mathbf{Z}}|\mathbf{S},\mathbf{B}=\mathbf{b}}, Q_{\widetilde{\mathbf{Z}}})\} \leqslant 2^{-\frac{\ell\varepsilon^2}{2\log_2^2(2^{q-u}+3)}} + \sqrt{2^{-\ell(R_{u+1:q}-\mathbb{I}\left(X_{u+1:q};\widetilde{Z}\right)-\varepsilon)}}/2$$
>
> where $P_{\widetilde{\mathbf{Z}}|\mathbf{s},\mathbf{b}}$ is induced by ϕ and $Q_{\widetilde{\mathbf{Z}}}$ is induced by uniform input,
> if $R_{u+1:q} > \mathbb{I}\left(X_{u+1:q}; \widetilde{Z}\right) + \varepsilon$

- Invertible extractor can be implemented by **finite-field multiplication** efficiently

- **CHALLENGE:** construction of finite field matching desired block length of code
  - Find irreducible polynomial to construct field

- **CANDIDATE**
  - Trinomials with Mersenne exponents **[Van Assche'06]**
  - Finite-field Arithmetic using Circulant Matrices (FACM) **[Hayashi-Tsurumaru'16]**

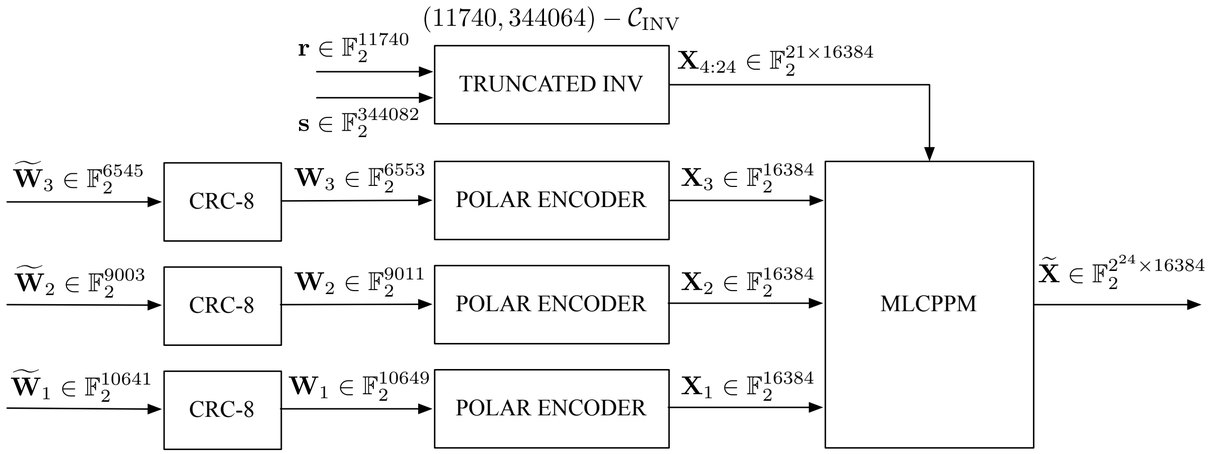
- Available size of field **may not match** desired length of invertible extractor

- **TRUNCATION:** generate resolvability code with longer block length then truncate it

> **LEMMA: TRUNCATION DOES NOT INCREASE VARIATIONAL DISTANCE**
> For $n' > n$, any $\mathbf{s}$ and $\mathbf{b}$,
> $$\mathbb{V}(P_{Z^n}, Q_{Z^n}) \leqslant \mathbb{V}(P_{Z^{n'}}, Q_{Z^{n'}})$$

  - Truncated inverter still ensures resolvability
  - Require more secret key bits due to increase of block length

- **TAKE AWAY:** constructing finite field with **mismatching length** requires **more secret key bits** to ensure same resolvability performance



- Combine all components in our design, we have upper-bound for covertness metric (total covertness metric)

$$\mathbb{E}_{\mathbf{S}}\{\mathbb{V}(P_{\widetilde{Z}}, Q_0^{\otimes n})\} \leqslant 2^{-\frac{\ell\varepsilon^2}{2\log_2^2(2^{q-u}+3)}} + \sqrt{2^{-\ell(R_{u+1:q}-\mathbb{I}(X_{u+1:q};\widetilde{Z})-\varepsilon)}}/2 + \sqrt{\delta/2} \triangleq \delta_t$$

$W_{Y\|X}$	$W_{Z\|X}$	Block length	δ	Number of levels	Covert message bits (throughput)	Covert and secret message bits	Secret key bits	Covert capacity	Total covertness metric
BSC(0.1)	BSC(0.42)	16384	0.0002	24	26189 (3.372)	19510	11764	15.65	0.0182

- This scheme is **efficient**: 19510 > 11764 (more covert and secret message bits than consumed secret key bits)
- **[Bash'13]** requires roughly 10^6 secret key bits to enable covert communication

- Present **joint reliability and resolvability** polar code based on **source polarization**
- Extension to multiple transmissions for **asymptotically negligible secret key rate** is possible but requires extra care
- **Construction of finite field** is challenging and crucial in **secret key bits usage**

Limits of Reliable Communication with Low Probability of Detection on AWGN Channels
Boulat A. Bash, Dennis Goeckel, Don Towsley
IEEE Journal on Selected Areas in Communications, vol. 31, no. 9, Sep. 2013

Multilevel-Coded Pulse-Position Modulation for Covert Communications Over Binary-Input Discrete Memoryless Channel
Ishaque A. Kadampot, Mehrdad Tahmasbi, Matthieu R. Bloch
IEEE Transactions on Information Theory, vol. 66, no. 10, Oct. 2020

Source Polarization
Erdal Arıkan
Proc. of IEEE International Symposium on Information Theory, Austin, TX, Jun. 2010

Polar Coding for the Broadcast Channel with Confidential Messages: A Random Binning Analogy
Rémi A. Chou, Matthieu R. Bloch
IEEE Transactions on Information Theory, vol. 62, no. 5, May 2016

Polynomial-Time, Semantically-Secure Encryption Achieving the Secrecy Capacity
Mihir Bellare, Stefano Tessaro
arXiv preprint, vol. 1201.3160, Jan. 2012

Explicit Codes for the Wiretap Channel with Uncertainty on the Eavesdropper's Channel
Rémi A. Chou
Proc. of IEEE International Symposium on Information Theory, Vail, CO, Aug. 2018

Covert Communication over Noisy Channels: A Resolvability Perspective
Matthieu R. Bloch
IEEE Transactions on Information Theory, vol. 62, no. 5, May 2016

First- and Second-Order Asymptotics in Covert Communication
Mehrdad Tahmasbi, Matthieu R. Bloch
IEEE Transactions on Information Theory, vol. 65, no. 4, Apr. 2019

Covert Communication with Polynomial Computational Complexity
Qiaosheng Zhang, Mayank Bakshi, Sidharth Jaggi
IEEE Transactions on Information Theory, vol. 66, no. 3, Mar. 2020

Polar Codes for Covert Communications over Asynchronous Discrete Memoryless Channels
Guillaume Frèche, Matthieu R. Bloch, Michel Barret
Entropy, vol. 20, no. 1, Dec. 2017

A Non-linear Channel Code for Covert Communications
Meritxell Lamarca, David Matas
Proc. of IEEE Wireless Communications and Networking Conference (WCNC), Marrakesh, Morocco, Apr. 2019

Optimal Covert Communications using Pulse-Position Modulation
Matthieu R. Bloch, Saikat Guha
Proc. of IEEE International Symposium on Information Theory, Aachen, Germany, Jun. 2017

How to Construct Polar Codes
Ido Tal, Alexander Vardy
IEEETransactions on Information Theory, vol. 59, no. 10, Oct. 2013

Greedy-Merge Degrading has Optimal Power-Law
Assaf Kartowsky, Ido Tal
IEEE Transactions on Information Theory, vol. 65, no. 2, Feb. 2019

A New Polar Coding Scheme for Strong Security on Wiretap Channels
Eren Şaşoğlu, Alexander Vardy
Proc. of IEEE International Symposium on Information Theory, Istanbul, Turkey, Jul. 2013

Quantum Cryptography and Secret-Key Distillation
Gilles Van Assche
Cambridge University Press, 2006

More Efficient Privacy Amplification with Less Random Seeds via Dual Universal Hash Function
Masahito Hayashi, Toyohiro Tsurumaru
IEEE Transactions on Information Theory, vol. 62, no. 4, Apr. 2016