Weighted Parity-Check Codes for Channels with State and Asymmetric Channels

Chih Wei Ling, Yanxiao Liu and Cheuk Ting Li
chihweiLing@link.cuhk.edu.hk
yanxiaoliu@link.cuhk.edu.hk
ctli@ie.cuhk.edu.hk

Department of Information Engineering, The Chinese University of Hong Kong

1st July 2022

Table of Contents

Introduction and Motivation

- Background: Find practical construction for the Gelfand-Pinsker setting with cost constraint and for the asymmetric channels

Introduction and Motivation

- Background: Find practical construction for the Gelfand-Pinsker setting with cost constraint and for the asymmetric channels
  - Barron et al. [2003] proposed nested linear code, but not sparse

Introduction and Motivation

- Background: Find practical construction for the Gelfand-Pinsker setting with cost constraint and for the asymmetric channels
  - Barron et al. [2003] proposed nested linear code, but not sparse
  - Martinian and Wainwright [2006] used sparse graphical code to generate nested linear code which is practical, but the construction induces two error events

Introduction and Motivation

- Background: Find practical construction for the Gelfand-Pinsker setting with cost constraint and for the asymmetric channels
  - Barron et al. [2003] proposed nested linear code, but not sparse
  - Martinian and Wainwright [2006] used sparse graphical code to generate nested linear code which is practical, but the construction induces two error events
  - Li and Anantharam [2021] proposed Poisson functional representation construction, with the best known second-order error bound (Scarlett [2015]) compared to other finite-blocklength schemes, but is impractical

Introduction and Motivation

- Background: Find practical construction for the Gelfand-Pinsker setting with cost constraint and for the asymmetric channels
  - Barron et al. [2003] proposed nested linear code, but not sparse
  - Martinian and Wainwright [2006] used sparse graphical code to generate nested linear code which is practical, but the construction induces two error events
  - Li and Anantharam [2021] proposed Poisson functional representation construction, with the best known second-order error bound (Scarlett [2015]) compared to other finite-blocklength schemes, but is impractical
- Our goal:
  - Construct code that is simultaneously practical and having error performance as good as (and sometimes better than) the construction in Barron et al. [2003]

Query Functions

- Let $\mathbf{H} \in \mathbb{F}_2^{n\times n}$ be a full-rank matrix, called the *full parity check matrix*

Query Functions

- Let $\mathbf{H} \in \mathbb{F}_2^{n \times n}$ be a full-rank matrix, called the *full parity check matrix*
  - $\mathbf{H}$ uniformly chosen random full-rank matrix

Query Functions

- Let $\mathbf{H} \in \mathbb{F}_2^{n \times n}$ be a full-rank matrix, called the *full parity check matrix*
    - $\mathbf{H}$ uniformly chosen random full-rank matrix
- For a *bias vector* $\mathbf{q} = [q_1, \ldots, q_n] \in [0,1]^n$, define the $\mathbf{q}$-*weight* of a vector $\mathbf{u} \in \mathbb{F}_2^n$ as

$$w_{\mathbf{q}}(\mathbf{u}) := \prod_{i=1}^{n} q_i^{u_i}(1-q_i)^{1-u_i} = 2^{-\sum_{i=1}^{n} H_b(u_i, q_i)}$$

Query Functions

- Let $\mathbf{H} \in \mathbb{F}_2^{n \times n}$ be a full-rank matrix, called the *full parity check matrix*
  - $\mathbf{H}$ uniformly chosen random full-rank matrix
- For a *bias vector* $\mathbf{q} = [q_1, \ldots, q_n] \in [0,1]^n$, define the $\mathbf{q}$-*weight* of a vector $\mathbf{u} \in \mathbb{F}_2^n$ as

$$w_{\mathbf{q}}(\mathbf{u}) := \prod_{i=1}^{n} q_i^{u_i}(1-q_i)^{1-u_i} = 2^{-\sum_{i=1}^{n} H_b(u_i, q_i)}$$

Definition

Given the bias vectors $\mathbf{p}, \mathbf{q} \in [0,1]^n$ (we call $\mathbf{p}$ the *codeword bias*, and $\mathbf{q}$ the *parity bias*), the *query function* with respect to $\mathbf{H}$ is given by

$$f_{\mathbf{H}}(\mathbf{p}, \mathbf{q}) := \mathrm{argmax}_{\mathbf{x} \in \mathbb{F}_2^n}\, w_{\mathbf{p}}(\mathbf{x}) w_{\mathbf{q}}(\mathbf{x}\mathbf{H}^T) \tag{1}$$

Weighted Parity-Check Codes (WPC)

Definition: Encoder

Given the *encoder codeword bias function* $\mathbf{p}_e : \mathbb{F}_2^k \to [0,1]^n$, which maps a message $\mathbf{m} \in \mathbb{F}_2^k$ (and other information available at the encoder) to a bias vector $\mathbf{p}_e(\mathbf{m})$, and the *encoder parity bias function* $\mathbf{q}_e : \mathbb{F}_2^k \to [0,1]^n$. The encoding function is

$$\mathbf{m} \mapsto \mathbf{x} = f_{\mathbf{H}}\left(\mathbf{p}_e(\mathbf{m}), \mathbf{q}_e(\mathbf{m})\right) \quad (2)$$

Weighted Parity-Check Codes (WPC)

Definition: Encoder

Given the *encoder codeword bias function* $\mathbf{p}_e : \mathbb{F}_2^k \to [0,1]^n$, which maps a message $\mathbf{m} \in \mathbb{F}_2^k$ (and other information available at the encoder) to a bias vector $\mathbf{p}_e(\mathbf{m})$, and the *encoder parity bias function* $\mathbf{q}_e : \mathbb{F}_2^k \to [0,1]^n$. The encoding function is

$$\mathbf{m} \mapsto \mathbf{x} = f_{\mathbf{H}}\left(\mathbf{p}_e(\mathbf{m}),\, \mathbf{q}_e(\mathbf{m})\right) \tag{2}$$

Definition: Decoder

Similarly, given the *decoder codeword and parity bias functions* $\mathbf{p}_d, \mathbf{q}_d : \mathbb{F}_2^n \to [0,1]^n$. For a corrupted version $\mathbf{y}$ of $\mathbf{x}$, the decoding function is

$$\mathbf{y} \mapsto \hat{\mathbf{m}} = \left[(\hat{\mathbf{x}}\mathbf{H}^T)_1, \ldots, (\hat{\mathbf{x}}\mathbf{H}^T)_k\right], \tag{3}$$

where

$$\hat{\mathbf{x}} := f_{\mathbf{H}}\left(\mathbf{p}_d(\mathbf{y}),\, \mathbf{q}_d(\mathbf{y})\right) \tag{4}$$

Recovering Conventional Linear Codes by WPC

- Given the binary symmetric channel with parameter β, i.e., $P(y_i|x_i)$ is $\mathrm{BSC}(\beta)$

Recovering Conventional Linear Codes by WPC

- Given the binary symmetric channel with parameter β, i.e., $P(y_i|x_i)$ is $\mathrm{BSC}(\beta)$
- To recover the conventional linear code, we take

$$\mathbf{p}_e(\mathbf{m}) = \frac{1}{2}\mathbf{1}^n, \qquad \mathbf{q}_e(\mathbf{m}) = [\mathbf{m},\, \mathbf{0}^{n-k}],$$
$$\mathbf{p}_d(\mathbf{y}) = \beta\mathbf{1}^n + (1-2\beta)\mathbf{y}, \qquad \mathbf{q}_d(\mathbf{y}) = \frac{1}{2}\mathbf{1}^n,$$

and substitute into Equations (2) and (4)

Recovering Conventional Linear Codes by WPC

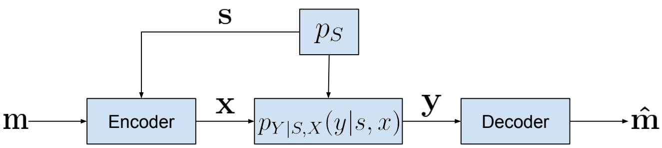
- Given the binary symmetric channel with parameter β, i.e., $P(y_i|x_i)$ is $\mathrm{BSC}(\beta)$
- To recover the conventional linear code, we take

$$\mathbf{p}_e(\mathbf{m}) = \frac{1}{2}\mathbf{1}^n, \qquad \mathbf{q}_e(\mathbf{m}) = [\mathbf{m},\, \mathbf{0}^{n-k}],$$

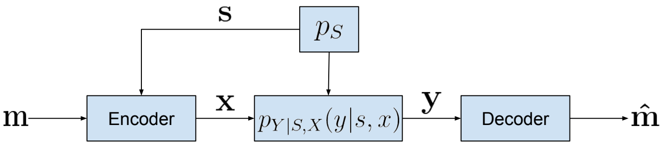
$$\mathbf{p}_d(\mathbf{y}) = \beta\mathbf{1}^n + (1-2\beta)\mathbf{y}, \qquad \mathbf{q}_d(\mathbf{y}) = \frac{1}{2}\mathbf{1}^n,$$

  and substitute into Equations (2) and (4)
- Note that $w_{\mathbf{p}_d(\mathbf{y})}(\mathbf{x}) = P(\mathbf{x}|\mathbf{y})$ is the posterior distribution of $\mathbf{x}$

WPC for Gelfand-Pinsker Setting



WPC for Gelfand-Pinsker Setting

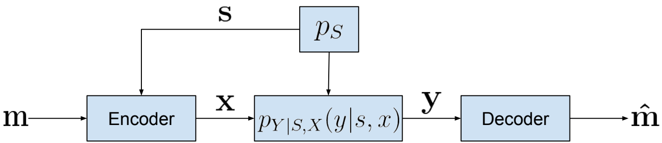


- We construct the *weighted parity-check codes with state* as follows
- Encoder: after observing **m** and **s**, takes

$$\mathbf{p}_e(\mathbf{m},\mathbf{s}) = [p_e(s_1), \ldots, p_e(s_n)], \quad \mathbf{q}_e(\mathbf{m},\mathbf{s}) = [\mathbf{m},\ \mathbf{q}], \tag{5}$$

and substitutes into (2) to obtain the codeword **x**

WPC for Gelfand-Pinsker Setting



- We construct the *weighted parity-check codes with state* as follows
- Encoder: after observing **m** and **s**, takes

$$\mathbf{p}_e(\mathbf{m},\mathbf{s}) = [p_e(s_1),\ldots,p_e(s_n)], \quad \mathbf{q}_e(\mathbf{m},\mathbf{s}) = [\mathbf{m},\ \mathbf{q}], \tag{5}$$

  and substitutes into (2) to obtain the codeword **x**
- Decoder: after observing **y**, takes

$$\mathbf{p}_d(\mathbf{y}) = [p_d(y_1),\ldots,p_d(y_n)], \quad \mathbf{q}_d(\mathbf{y}) = [\frac{1}{2}\mathbf{1}^k,\ \mathbf{q}], \tag{6}$$

  and outputs $\hat{\mathbf{m}} = [(\hat{\mathbf{x}}\mathbf{H}^T)_1,\ldots,(\hat{\mathbf{x}}\mathbf{H}^T)_k]$ after substituting into (4)

WPC is Capacity-achieving (1/3)

- We first state our main result as follows:

Theorem 1

- Assume $|\mathcal{S}|, |\mathcal{Y}| < \infty$. Fix any $P_{X|S}$, and let $S \sim P_S$, $X|S \sim P_{X|S}$, $Y|(S,X) \sim P_{Y|S,X}$

WPC is Capacity-achieving (1/3)

- We first state our main result as follows:

Theorem 1

- Assume $|\mathcal{S}|, |\mathcal{Y}| < \infty$. Fix any $P_{X|S}$, and let $S \sim P_S$, $X|S \sim P_{X|S}$, $Y|(S,X) \sim P_{Y|S,X}$
- Consider the weighted parity-check code with state, where $p_e(s) = P_{X|S}(1|s)$, $p_d(y) = P_{X|Y}(1|y)$, and P_Q is a discrete distribution over $[0,1]$ with finite support satisfying

$$\mathbf{E}[H_b(Q)] = (1 - H(X|S))/(1 - R) \tag{7}$$

WPC is Capacity-achieving (1/3)

- We first state our main result as follows:

Theorem 1

- Assume $|\mathcal{S}|, |\mathcal{Y}| < \infty$. Fix any $P_{X|S}$, and let $S \sim P_S$, $X|S \sim P_{X|S}$, $Y|(S,X) \sim P_{Y|S,X}$
- Consider the weighted parity-check code with state, where $p_e(s) = P_{X|S}(1|s)$, $p_d(y) = P_{X|Y}(1|y)$, and P_Q is a discrete distribution over $[0,1]$ with finite support satisfying

$$\mathbf{E}[H_b(Q)] = (1 - H(X|S))/(1 - R) \tag{7}$$

- Then, for any $R < I(X;Y) - I(X;S)$, the probability of error of the code tends to 0, and the empirical joint distribution of $\{(s_i, x_i)\}_{i=1,\ldots,n}$ tends to $P_S P_{X|S}$ in probability as $n \to \infty$

WPC is Capacity-achieving (2/3)

Lemma 1

- Consider the WPC with state, where $|\mathcal{S}|, |\mathcal{Y}| < \infty$, and P_Q is a discrete distribution over $[0, 1]$ with finite support

WPC is Capacity-achieving (2/3)

Lemma 1

- Consider the WPC with state, where $|\mathcal{S}|, |\mathcal{Y}| < \infty$, and P_Q is a discrete distribution over $[0,1]$ with finite support
- Let $S \sim P_S$, $X|S \sim P_{X|S}$, $Y|(S,X) \sim P_{Y|S,X}$, $Q \sim P_Q$, $V \in \{0,1\}$, $V|Q \sim P_{V|Q}$, where $(P_{X|S}, P_{V|Q})$ is the minimizer of

$$\begin{aligned} \min \quad & \mathbf{E}\left[H_b(X, p_e(S))\right] + (1-R)\mathbf{E}\left[H_b(V,Q)\right] \\ \text{s.t.} \quad & H(X|S) + (1-R)H(V|Q) \geq 1 \end{aligned} \tag{8}$$

WPC is Capacity-achieving (2/3)

Lemma 1

- Consider the WPC with state, where $|\mathcal{S}|, |\mathcal{Y}| < \infty$, and P_Q is a discrete distribution over $[0, 1]$ with finite support
- Let $S \sim P_S$, $X|S \sim P_{X|S}$, $Y|(S, X) \sim P_{Y|S,X}$, $Q \sim P_Q$, $V \in \{0, 1\}$, $V|Q \sim P_{V|Q}$, where $(P_{X|S}, P_{V|Q})$ is the minimizer of

$$\begin{aligned} \min \quad & \mathbf{E}\left[H_b(X, p_e(S))\right] + (1-R)\mathbf{E}\left[H_b(V, Q)\right] \\ \text{s.t.} \quad & H(X|S) + (1-R)H(V|Q) \geq 1 \end{aligned} \tag{8}$$

- If the minimizer of (8) is unique, and for all $P_{\tilde{X}|Y}$, $P_{\tilde{V}|Q}$ satisfying

$$H(\tilde{X}|Y) + (1-R)H(\tilde{V}|Q) \geq 1 - R, \tag{9}$$

  we have

$$\begin{aligned} & \mathbf{E}[H_b(\tilde{X}, p_d(Y))] + (1-R)\mathbf{E}[H_b(\tilde{V}, Q)] \\ & > \mathbf{E}[H_b(X, p_d(Y))] + (1-R)\mathbf{E}[H_b(V, Q)] \end{aligned} \tag{10}$$

WPC is Capacity-achieving (3/3)

Lemma 1 (Continued)

- Then the probability of error of the code tends to 0, and the empirical joint distribution of $\{(s_i, x_i)\}_{i=1,\ldots,n}$ tends to $P_S P_{X|S}$ in probability, as $n \to \infty$

WPC is Capacity-achieving (3/3)

Lemma 1 (Continued)

- Then the probability of error of the code tends to 0, and the empirical joint distribution of $\{(s_i, x_i)\}_{i=1,\ldots,n}$ tends to $P_S P_{X|S}$ in probability, as $n \to \infty$

Proof of Lemma 1

(Sketch): Use Sanov's theorem and robust typicality

WPC is Capacity-achieving (3/3)

Lemma 1 (Continued)

- Then the probability of error of the code tends to 0, and the empirical joint distribution of $\{(s_i, x_i)\}_{i=1,\ldots,n}$ tends to $P_S P_{X|S}$ in probability, as $n \to \infty$

Proof of Lemma 1

(Sketch): Use Sanov's theorem and robust typicality

Proof of Theorem 1

(Sketch):

- Use the Lagrange multiplier method to find a general expression for the minimizer for minimization problem (8) with the parameter $\lambda \geq 0$
- Set $\lambda = 1$, after some algebra manipulations with the facts in information theory and Lemma 1, we get the conclusion

How to Choose P_Q satisfying Equation (7)

- (Threshold) Take $P_Q(0) = P_Q(1) = (1-\gamma)/2$, $P_Q(1/2) = \gamma$, where $\gamma = (1 - H(X|S))/(1-R)$
  - Essentially equivalent to the nested linear code

How to Choose P_Q satisfying Equation (7)

- (Threshold) Take $P_Q(0) = P_Q(1) = (1-\gamma)/2$, $P_Q(1/2) = \gamma$, where $\gamma = (1 - H(X|S))/(1-R)$
  - Essentially equivalent to the nested linear code
- (Linear) Take P_Q to be the uniform distribution $\mathrm{Unif}[0,1]$
  - May not achieve the capacity
  - "Universal" – the decoder does not need to know P_S or p_e

How to Choose P_Q satisfying Equation (7)

- (Threshold) Take $P_Q(0) = P_Q(1) = (1-\gamma)/2$, $P_Q(1/2) = \gamma$, where $\gamma = (1 - H(X|S))/(1-R)$
  - Essentially equivalent to the nested linear code
- (Linear) Take P_Q to be the uniform distribution $\mathrm{Unif}[0,1]$
  - May not achieve the capacity
  - "Universal" – the decoder does not need to know P_S or p_e
- (Threshold linear) Construct P_Q using the cdf

$$F_Q(t) := \begin{cases} 0 & \text{if } t < 0 \\ \max\{\theta/2, 0\} & \text{if } 0 \le t < |\theta|/2 \\ t & \text{if } |\theta|/2 \le t < 1 - |\theta|/2 \\ 1 - \max\{\theta/2, 0\} & \text{if } 1 - |\theta|/2 \le t < 1 \\ 1 & \text{if } t \ge 1 \end{cases} \tag{11}$$

where $\theta \in [-1, 1]$ is chosen such that (7) holds

- Combines the linear method for t close to $1/2$, and the threshold method for smaller and larger t's

Simulation Result

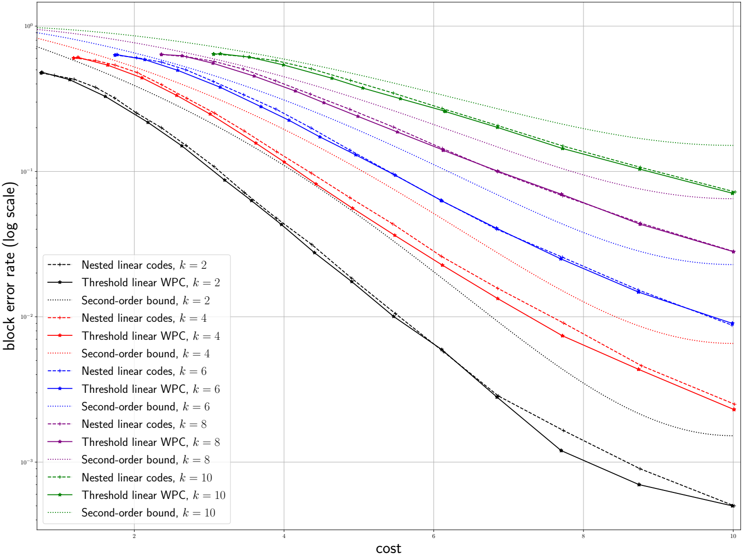


Figure: Performance evaluation with $n = 20$, $\beta = 0.05$

Reference

Richard J Barron, Brian Chen, and Gregory W Wornell. The duality between information embedding and source coding with side information and some applications. *IEEE Transactions on Information Theory*, 49 (5):1159–1180, 2003.

Cheuk Ting Li and Venkat Anantharam. A unified framework for one-shot achievability via the Poisson matching lemma. *IEEE Transactions on Information Theory*, 67(5):2624–2651, 2021.

Emin Martinian and Martin J Wainwright. Low-density constructions can achieve the Wyner-Ziv and Gelfand-Pinsker bounds. pages 484–488, 2006.

Jonathan Scarlett. On the dispersions of the Gel'fand–Pinsker channel and dirty paper coding. *IEEE Trans. Inf. Theory*, 61(9):4569–4586, 2015.