Weighted Parity-Check Codes

Chih Wei Ling **Yanxiao Liu** Cheuk Ting Li

Department of Information Engineering, The Chinese University of Hong Kong

Abstract

We introduce a new class of codes, called weighted parity-check codes, where each parity-check bit has a weight that indicates its likelihood to be one (instead of fixing each parity-check bit to be zero). It is applicable to a wide range of settings, e.g. asymmetric channels, channels with state and/or cost constraints, and the Wyner-Ziv problem, and can provably achieve the capacity. For the channel with state (Gelfand-Pinsker) setting, our code not only achieves the capacity of any channel with state, but also achieves a smaller error rate compared to the nested linear code.

Ideas and Advantages

The goal is to present a general code construction based on **weighted codebook** idea, but with a linear structure for efficient encoding and decoding.

- The codebook is a "fuzzy set", where each bit sequence has a weight that corresponds to the likelihood that the sequence is selected.
- By [1], weighted codebook eliminates the need of subcodebooks and gives sharper finite-blocklength and second-order error bounds.
- It applies to general(symmetric/asymmetric) channels with/without state.

Channels with State Information

Consider the channel has a state sequence $\mathbf{s} = [s_1, \ldots, s_n]$, where $s_i \in \mathcal{S}$ (not necessarily binary), $s_i \stackrel{iid}{\sim} P_S$, is available noncausally to the encoder. Given $\mathbf{s}$, the encoder encodes message $\mathbf{m} \in \mathbb{F}_2^k$ into $\mathbf{x} \in \mathbb{F}_2^n$, which is sent through a memoryless channel $P_{Y|S,X}(y|s,x)$. The decoder receives $\mathbf{y}$ and outputs $\hat{\mathbf{m}}$. The input may have a cost constraint $\mathbf{E}[\sum_{i=1}^n c(s_i, x_i)] \le nD$, where $c : \mathcal{S} \times \mathbb{F}_2 \to [0, \infty)$.

Binary-Hamming Information Embedding

Consider $s_i \stackrel{iid}{\sim} \mathrm{Bern}(1/2)$ and $X \to Y$ is $\mathrm{BSC}(\beta)$. We have an expected cost/distortion constraint $\mathbf{E}[|\{i : x_i \neq s_i\}|] \le nD$. For $0 \le \beta \le D \le 1/2$, the capacity is the upper concave envelope of $H(D) - H(\beta)$.

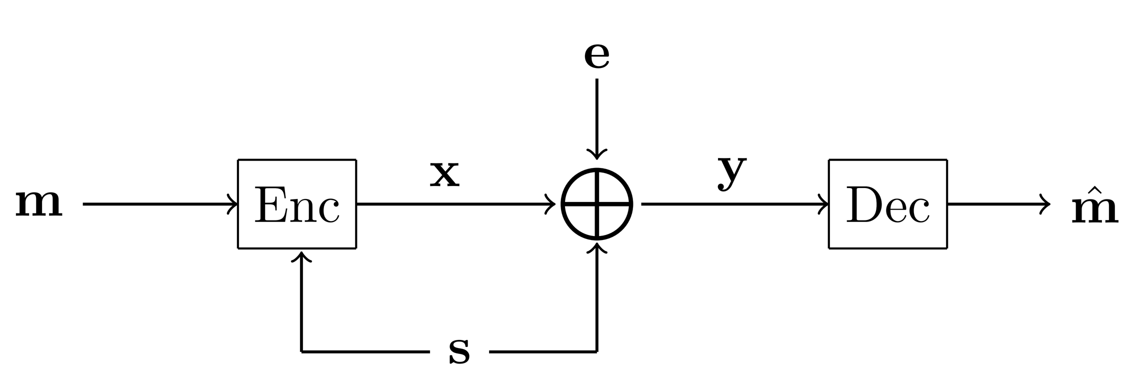

Figure 1. Block diagram of channel coding with state information. The encoder embeds $\mathbf{m}$ into the channel input $\mathbf{x}$, which is with a cost constraint. $\mathbf{e} \sim \mathrm{Ber}(\beta)$ is the channel noise.

References

[1] 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.

[2] Ram Zamir, Shlomo Shamai, and Uri Erez. Nested linear/lattice codes for structured multiterminal binning. *IEEE Transactions on Information Theory*, 48(6):1250–1276, 2002.

Weighted Parity-Check Codes (WPC)

In channel coding, the encoder encodes the message $\mathbf{m} \in \mathbb{F}_2^k$ into codeword $\mathbf{x} \in \mathbb{F}_2^n$. The decoder receives $\mathbf{y} \in \mathbb{F}_2^n$ and recovers $\hat{\mathbf{m}} \in \mathbb{F}_2^k$.

Randomly choose a full-rank parity check matrix $\mathbf{H} \in \mathbb{F}_2^{n \times n}$. 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

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

Intuitively, $w_{\mathbf{q}}(\mathbf{u})$ is the probability of $\mathbf{u}$ assuming the entries $u_i \sim \mathrm{Bern}(q_i)$ are independent across i.

Given the codeword/parity bias vectors $\mathbf{p}, \mathbf{q} \in [0,1]^n$, the *query function* is

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

The encoder has two parameters: the *encoder codeword bias function* $\mathbf{p}_e : \mathbb{F}_2^k \to [0,1]^n$ which maps the 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 actual encoding function is

$$\mathbf{m} \mapsto \mathbf{x} = f_{\mathbf{H}}(\mathbf{p}_e(\mathbf{m}), \mathbf{q}_e(\mathbf{m})).$$

The decoder likewise has two parameters: the *decoder codeword and parity bias functions* $\mathbf{p}_d, \mathbf{q}_d : \mathbb{F}_2^n \to [0,1]^n$. The decoding function is

$$\mathbf{y} \mapsto \hat{\mathbf{m}} = \left[(\hat{\mathbf{x}}\mathbf{H}^T)_1, \ldots, (\hat{\mathbf{x}}\mathbf{H}^T)_k \right], \tag{2}$$

where $\hat{\mathbf{x}} := f_{\mathbf{H}}(\mathbf{p}_d(\mathbf{y}), \mathbf{q}_d(\mathbf{y}))$.

Weighted Parity-Check Codes with State (WPCS)

The encoder observes $\mathbf{m}, \mathbf{s}$ and uses $\mathbf{p}_e(\mathbf{m}, \mathbf{s})$, $\mathbf{q}_e(\mathbf{m}, \mathbf{s})$ to obtain $\mathbf{x}$. The decoder uses $\mathbf{p}_d(\mathbf{y})$, $\mathbf{q}_d(\mathbf{y})$ to obtain $\hat{\mathbf{x}}$, and outputs $\hat{\mathbf{m}} = [(\hat{\mathbf{x}}\mathbf{H}^T)_1, \ldots, (\hat{\mathbf{x}}\mathbf{H}^T)_k]$.

$$\begin{aligned}
\mathbf{p}_e(\mathbf{m}, \mathbf{s}) &= [p_e(s_1), \ldots, p_e(s_n)], \\
\mathbf{q}_e(\mathbf{m}, \mathbf{s}) &= [\mathbf{m}, \mathbf{q}], \\
\mathbf{p}_d(\mathbf{y}) &= [p_d(y_1), \ldots, p_d(y_n)], \\
\mathbf{q}_d(\mathbf{y}) &= [\tfrac{1}{2}\mathbf{1}^k, \mathbf{q}],
\end{aligned}$$

such that $\mathbf{q} = [q_1, \ldots, q_{n-k}]$, where $q_i \sim P_Q$ i.i.d., and P_Q is a distribution over $[0,1]$ symmetric about $1/2$ (i.e., if $Q \sim P_Q$, then $1 - Q \sim P_Q$).

Parity Bias Distribution

The nested linear code [2] is a special case of WPCS, where there are $n-k-\tilde{k}$ parity-check bits fixed to zero (i.e., $q_i = 0$), and $\tilde{k}$ unused parity-check bits ($q_i = 1/2$), where $\tilde{k} \in \{0, \ldots, n-k\}$ is the dimension of each coset. It can be approximated by taking $P_Q(0) = P_Q(1) = (1-\gamma)/2$, $P_Q(1/2) = \gamma$, where $\gamma = \tilde{k}/(n-k)$, giving around $(n-k)P_Q(1/2) = \tilde{k}$ unused parity-check bits.

We construct our P_Q so that (7) holds, named *Threshold linear* P_Q using a cdf:

$$F_Q(t) := \begin{cases} 0 & \text{if } t < 0 \\ \max\{\theta/2, 0\} & \text{if } 0 \le t < |\theta|/2 \\ t & \text{if } |\theta|/2 \le t < 1 - |\theta|/2 \\ 1 - \max\{\theta/2, 0\} & \text{if } 1 - |\theta|/2 \le t < 1 \\ 1 & \text{if } t \ge 1, \end{cases} \tag{3}$$

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

Optimality for the Channels with States

Consider WPCS that $|\mathcal{S}|, |\mathcal{Y}| < \infty$, 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

$$\mathbf{E}\left[H_b(X, p_e(S))\right] + (1-R)\mathbf{E}\left[H_b(V, Q)\right], \tag{4}$$

where H_b is the binary cross entropy function, subject to

$$H(X|S) + (1-R)H(V|Q) \ge 1. \tag{5}$$

If the minimizer of (4) is unique, and for all $P_{\tilde{X}|Y}, P_{\tilde{V}|Q}$ satisfying

$$H(\tilde{X}|Y) + (1-R)H(\tilde{V}|Q) \ge 1 - R, \tag{6}$$

we have

$$\mathbf{E}[H_b(\tilde{X}, p_d(Y))] + (1-R)\mathbf{E}[H_b(\tilde{V}, Q)] > \mathbf{E}[H_b(X, p_d(Y))] + (1-R)\mathbf{E}[H_b(V, Q)]$$

and then as $n \to \infty$, the probability of error tends to 0.

Corollary

Let $|\mathcal{S}|, |\mathcal{Y}| < \infty$, fix $P_{X|S}$. Consider WPCS that $p_e(s) = P_{X|S}(1|s)$, $p_d(y) = P_{X|Y}(1|y)$, and P_Q is discrete and over $[0,1]$ with finite support satisfying

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

For any $R < I(X;Y) - I(X;S)$, as $n \to \infty$, the error probability goes to 0.

Performance Evaluation

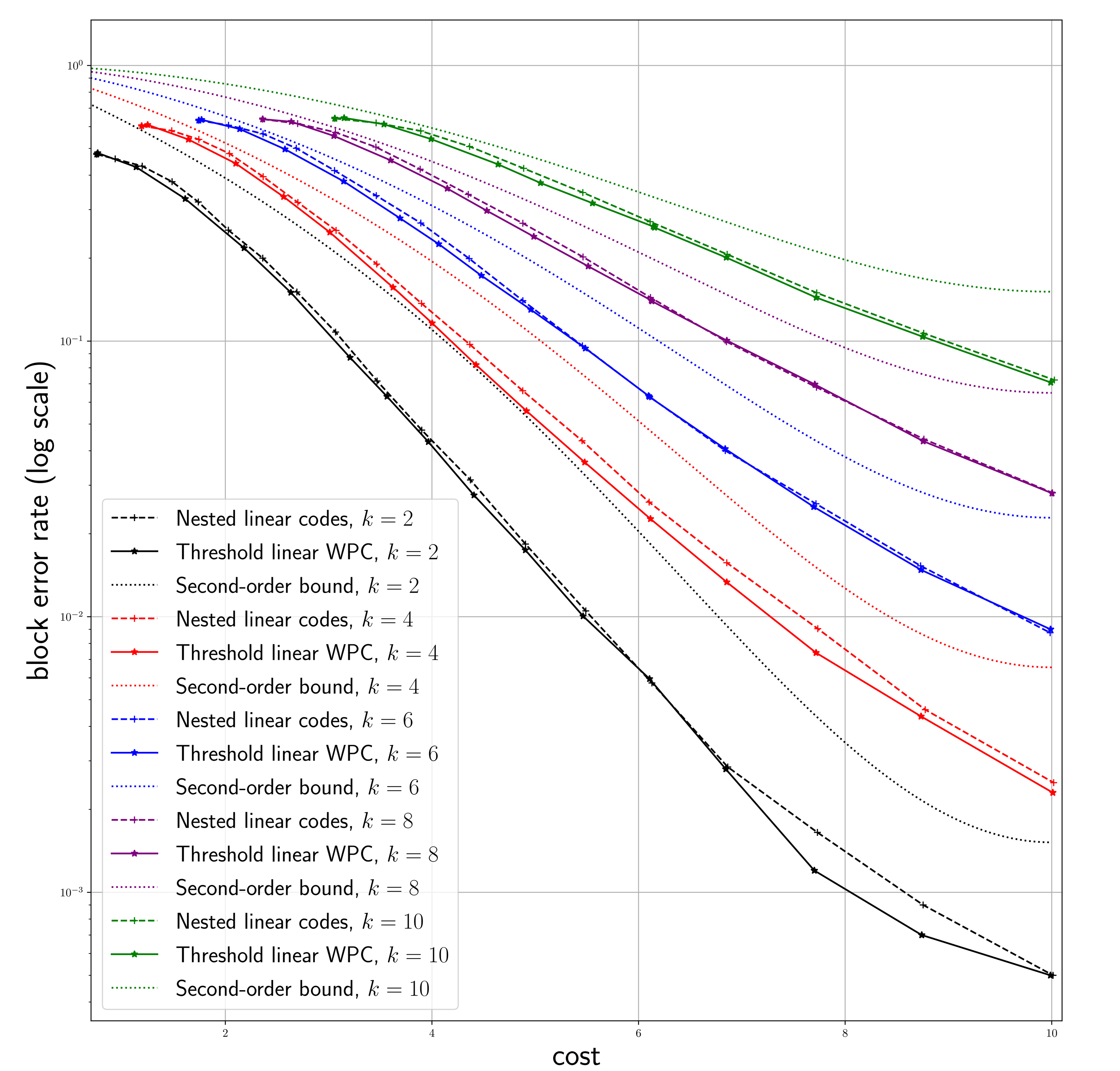

Figure 2. Performance of $n = 20$, BSC channel of crossover probability $\beta = 0.05$. Note we use $p_e(s) = \alpha^{1-s}(1-\alpha)^s$ for $S \to X$ be about $\mathrm{BSC}(\alpha)$, $p_d(y) = \beta^{1-y}(1-\beta)^y$ and (3) for P_Q.