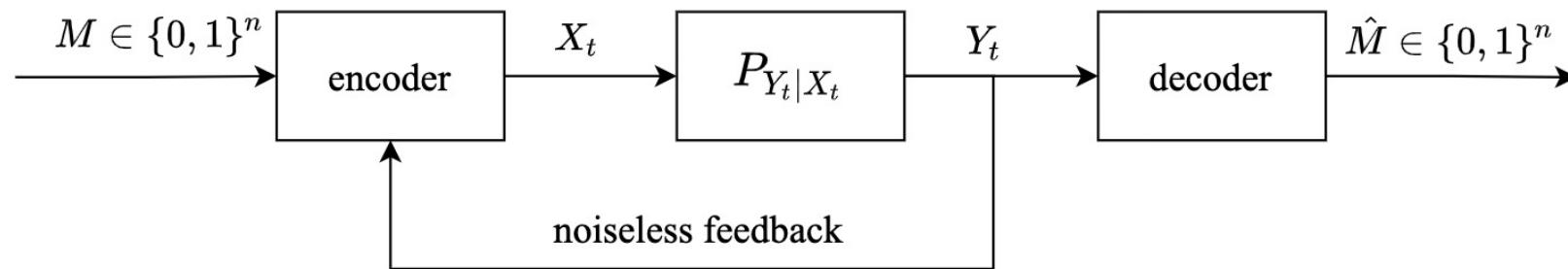


Instantaneous small-enough difference coding over a DMC

Nian Guo and Victoria Kostina
California Institute of Technology
ISIT 2021

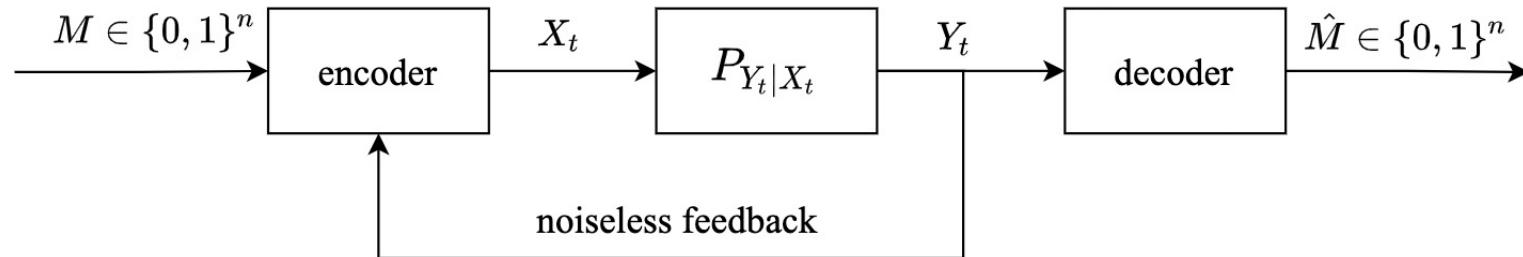
Introduction and Motivation

Communication over a channel with feedback (classical)



- encoder: observes n bits + sequential transmissions
- memoryless channel with feedback
- decoder: estimates the message

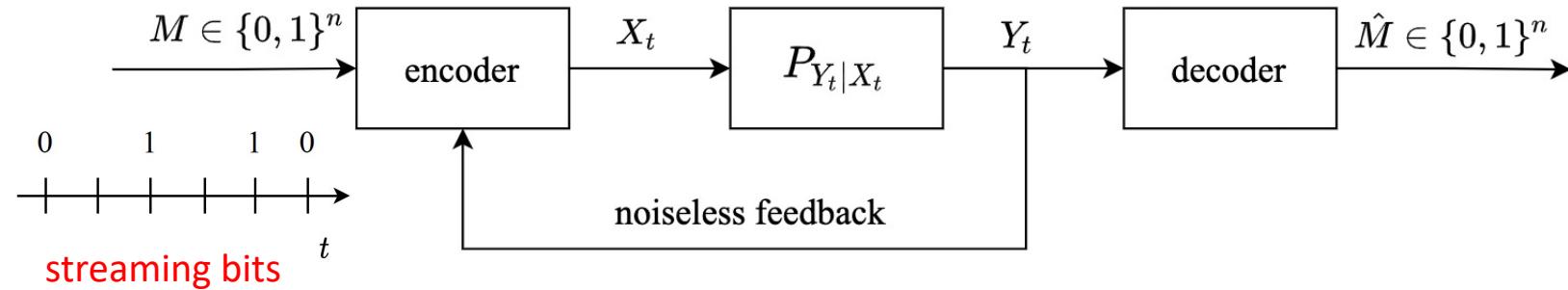
Introduction and Motivation



Prior literature:

- Shannon (1956): feedback cannot increase capacity
- Horstein (1963): **BSC** → Schalkwijk (1971), Schalkwijk and Post (1973), Waber et al. (2013)
- Schalkwijk and Kailath (1966): **AWGN**
- Burnashev (1975), Yamamoto and Itoh (1979), Ooi and Wornell (1998), Naghshvar et al. (2012): **DMC, reliability function**
- Shayevitz and Feder (2011): **memoryless channel**

Introduction and Motivation



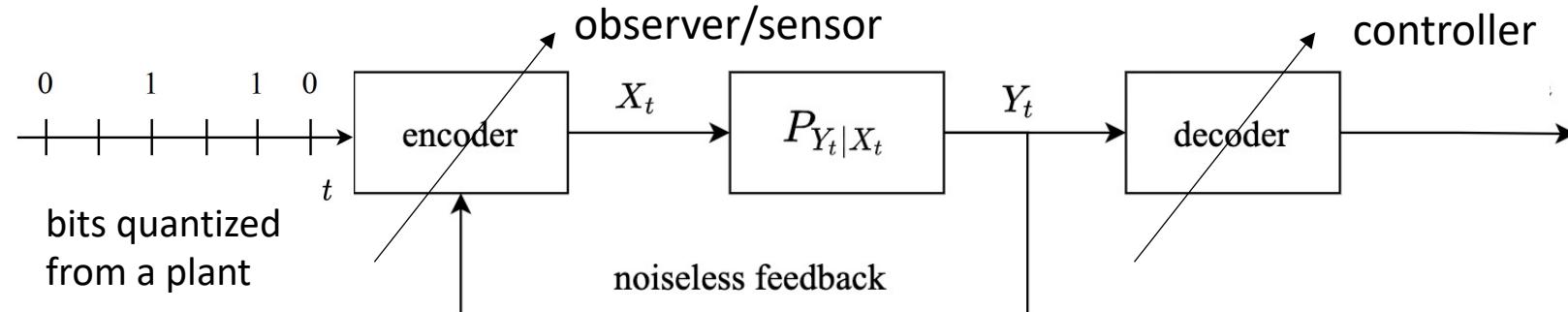
- **Limitation:** block encoding → **delay**
- **Recent literature:** causal/instantaneous encoding

Lalitha et al. (2019): BSC, decoder knows the bit arrival times

Antonini et al. (2021, ISIT): BSC, fixed bit arrival rate

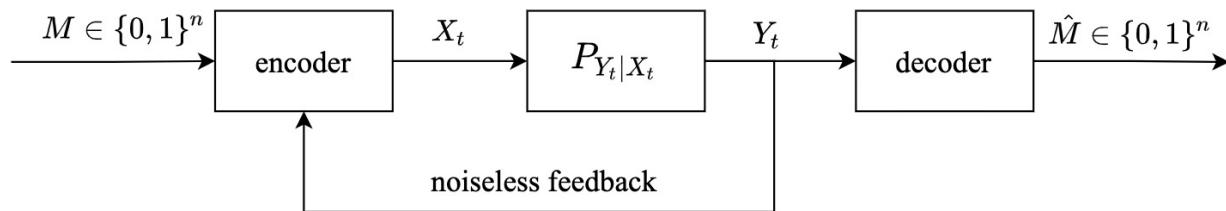
} contrast:
DMC, bit arrival distribution

Introduction and Motivation



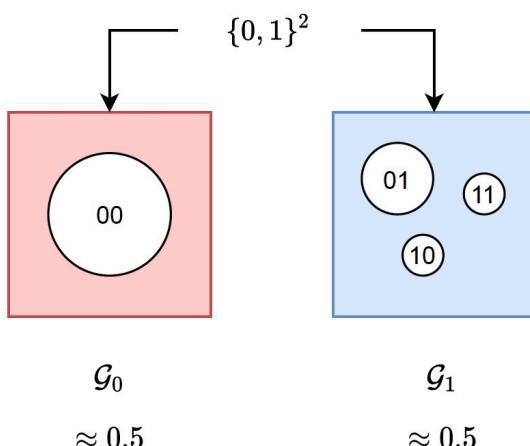
- Sahai and Mitter (2006), Sukhavasi and Hassibi (2016), Khina et al. (2016): **distributed/networked control** control plant over a noisy channel with feedback
- **Contribution:** A novel code – **instantaneous small-enough difference (SED) code**
 - DMC, bit streaming
 - Decoder only knows the bit arrival distribution
 - Outperforms existing schemes
- Name of “instantaneous SED code”
 - instantaneous = start to transmit as soon as the first bit arrives
 - What is SED?

SED rule



Naghshvar et al. (2012): block encoding

- Setting: 2-input DMC, capacity-achieving input distribution $P_X^*(0) = P_X^*(1) = 0.5$
- Algorithm:
 - track posteriors $P_{M|Y^t}(i|y^t) \ i \in \{0,1\}^n$
 - form channel input: partition $\{0,1\}^n$ using SED rule

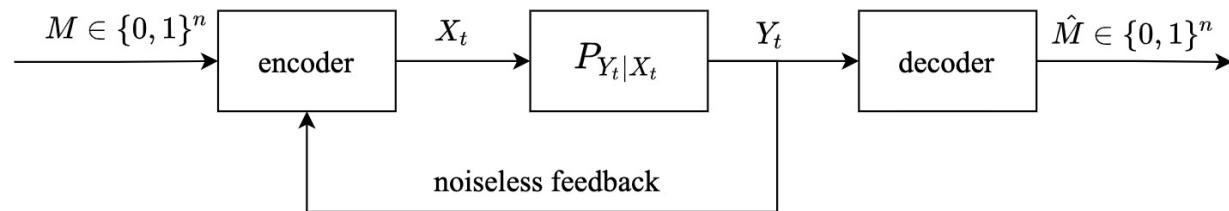


SED partitioning rule at time t

$$P_{M|Y^{t-1}}(\mathcal{G}_0(t)|y^{t-1}) \geq P_{M|Y^{t-1}}(\mathcal{G}_1(t)|y^{t-1})$$

$$\begin{aligned} & \sum_{x \in \{0,1\}} |P_{M|Y^{t-1}}(\mathcal{G}_x(t)|y^{t-1}) - P_X^*(x)| \\ & \leq \sum_{x \in \{0,1\}} |P_{M|Y^{t-1}}(\mathcal{G}'_x(t)|y^{t-1}) - P_X^*(x)| \end{aligned}$$

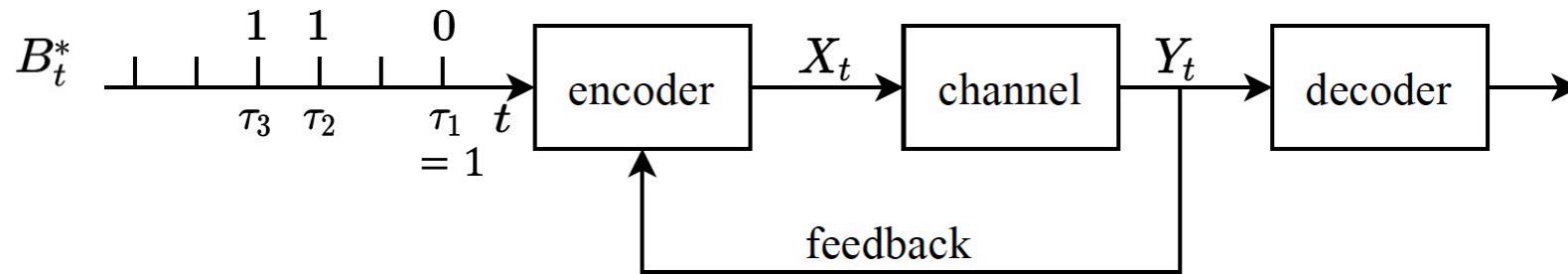
SED rule



Naghshvar et al. (2012): block encoding

- Setting: 2-input DMC, capacity-achieving input distribution $P_X^*(0) = P_X^*(1) = 0.5$
- Algorithm:
 - track posteriors $P_{M|Y^t}(i|y^t)$ $i \in \{0, 1\}^n$
 - form channel input: partition $\{0, 1\}^n$ using SED rule
- Complexity: double exponential in n
 - Partition problem (computer science): NP-hard
 - Antonini et al. (2020) reduce the complexity, block encoding
 - **We reduce the complexity to polynomial (type-set SED rule)**

Problem formulation

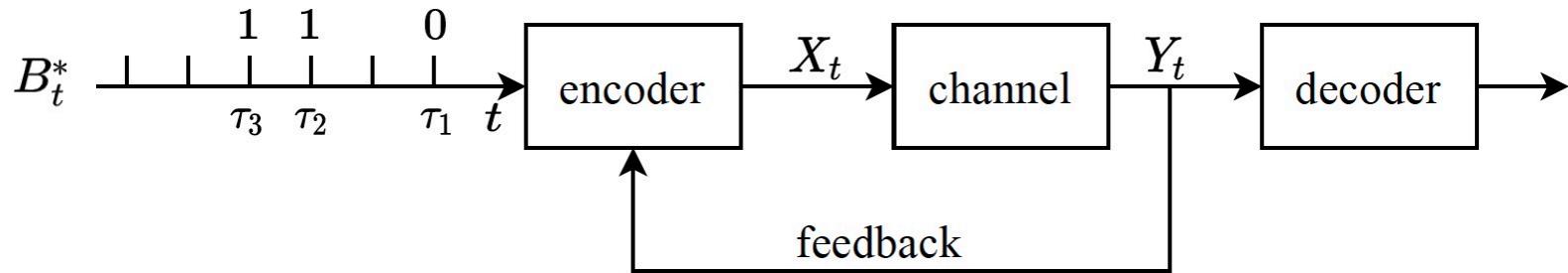


Streaming message bits:

- n -th bit arrival time τ_n
- B_t^* : message bits that arrived up to time t , $1 \leq \text{length}(B_t^*) \leq t$
- bit arrival distribution

$$P_{B_{t+1}^*|B_t^*} \left\{ \begin{array}{l} P_{B_{t+1}^*|B_t^*}(s \boxplus 1|s) \\ P_{B_{t+1}^*|B_t^*}(s \boxplus 0|s) \\ P_{B_{t+1}^*|B_t^*}(s|s) \end{array} \right.$$

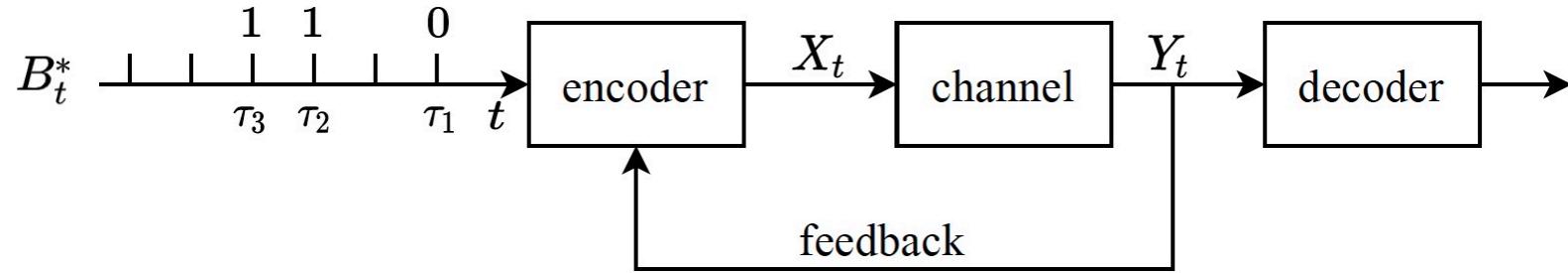
Problem formulation



Feedback code with instantaneous encoding at $t = 1, 2, \dots$

- Encoder B_t^* and $Y^{t-1} \rightarrow X_t$
- Decoder $Y^t \rightarrow$ estimate of n message bits
- Stopping rule: output estimate

Instantaneous SED code: algorithm



At time t

- Prior update $P_{B_t^*|Y^{t-1}}(i|y^{t-1})$ $1 \leq \text{length}(i) \leq t$
posterior $P_{B_{t-1}^*|Y^{t-1}}(j|y^{t-1})$ + bit arrival distribution $P_{B_t^*|B_{t-1}^*}(i|j)$

Instantaneous SED code: algorithm

At time t

- Group partition (multiple-input SED)

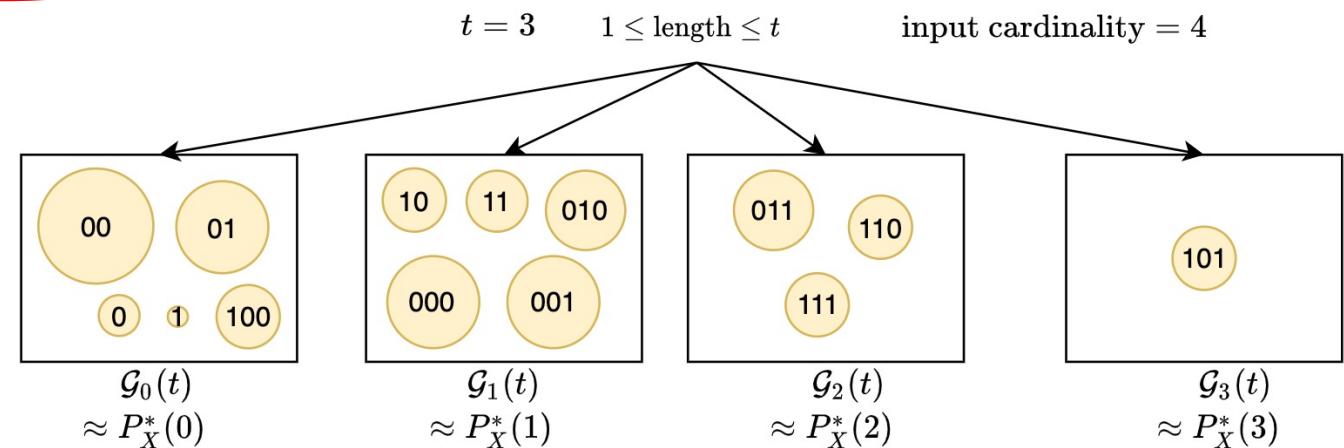
• partition all possible strings $1 \leq \text{length}(i) \leq t$

• minimum L-1 distance

$$\sum_{x \in \mathcal{X}} |P_{B_t^*|Y^{t-1}}(\mathcal{G}_x(t)|y^{t-1}) - P_X^*(x)| \\ \leq \sum_{x \in \mathcal{X}} |P_{B_t^*|Y^{t-1}}(\mathcal{G}'_x(t)|y^{t-1}) - P_X^*(x)|$$

• transmit group index

e.g., $B_3^* = 00, X_3 = 0$



Differences: our code vs. block/causal encoding schemes with known bit arrival times

Track the probabilities of all possible binary strings

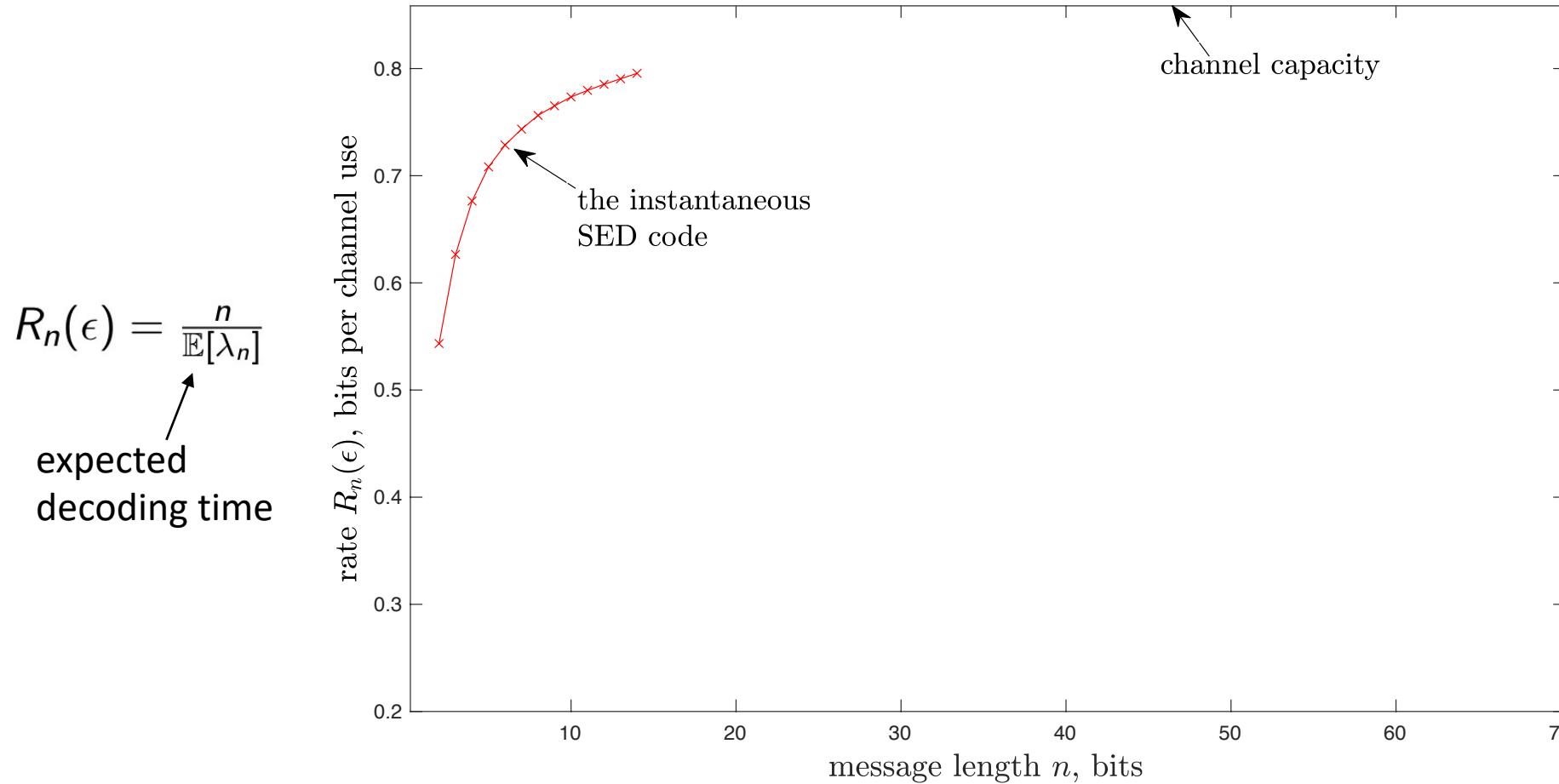
Use prior $P_{B_t^*|Y^{t-1}}$ (posterior + bit arrival distribution) in addition to posterior $P_{B_{t-1}^*|Y^{t-1}}$

Instantaneous SED code: algorithm

At time t

- Posterior update $P_{B_t^*|Y^t}(i|y^t)$ $1 \leq \text{length}(i) \leq t$
prior $P_{B_t^*|Y^{t-1}}(i|y^{t-1})$, channel transition probability, and channel output Y_t
- Stopping and decoding
 - aim: decode n bits with error probability $\leq \epsilon$
 - rule: $\exists i \in \{0,1\}^n$ s.t. the posterior of all binary strings with prefix “ i ” $\geq 1 - \epsilon$

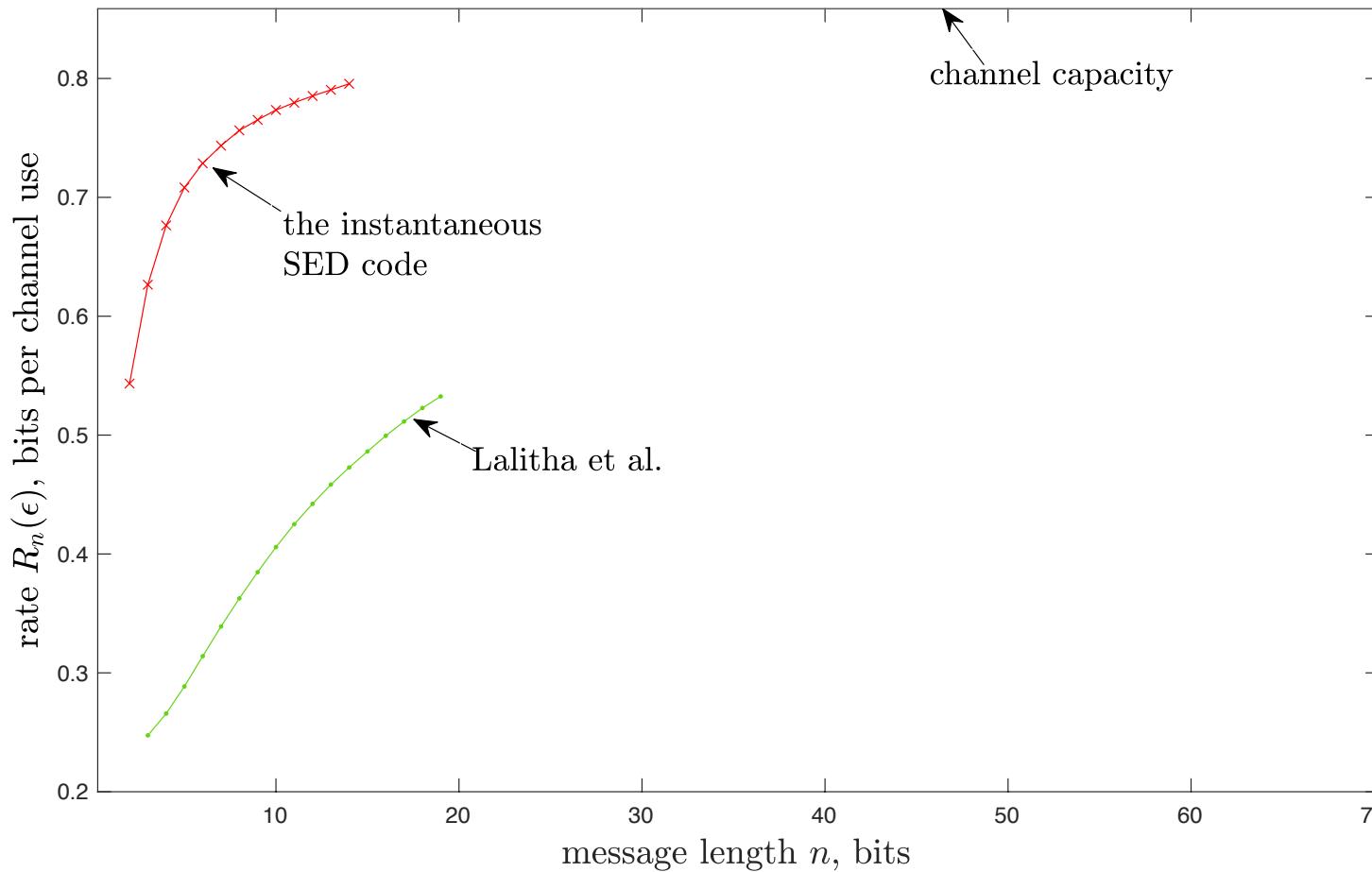
Instantaneous SED code: performance



BSC(0.02), $\epsilon = 10^{-3}$. At each time $t = 1, 2, \dots, n$, a new Bernoulli($\frac{1}{2}$) bit arrives at the encoder.

Instantaneous SED code: performance

$$R_n(\epsilon) = \frac{n}{\mathbb{E}[\lambda_n]}$$

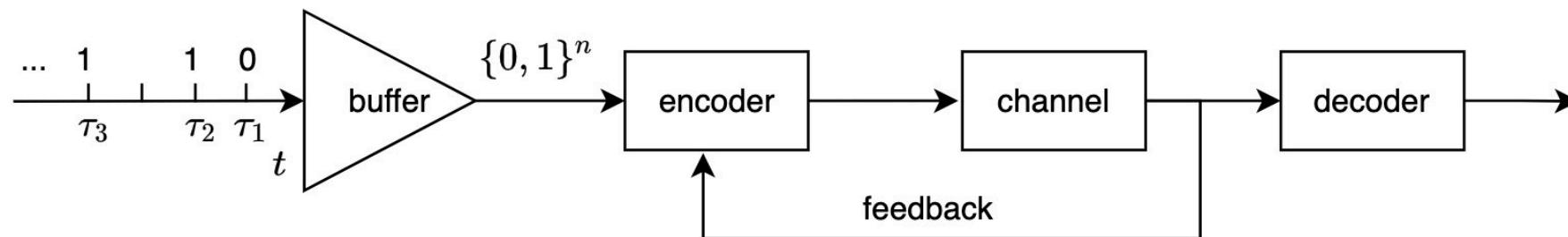


BSC(0.02), $\epsilon = 10^{-3}$. At each time $t = 1, 2, \dots, n$, a new Bernoulli($\frac{1}{2}$) bit arrives at the encoder.

Instantaneous SED code: performance

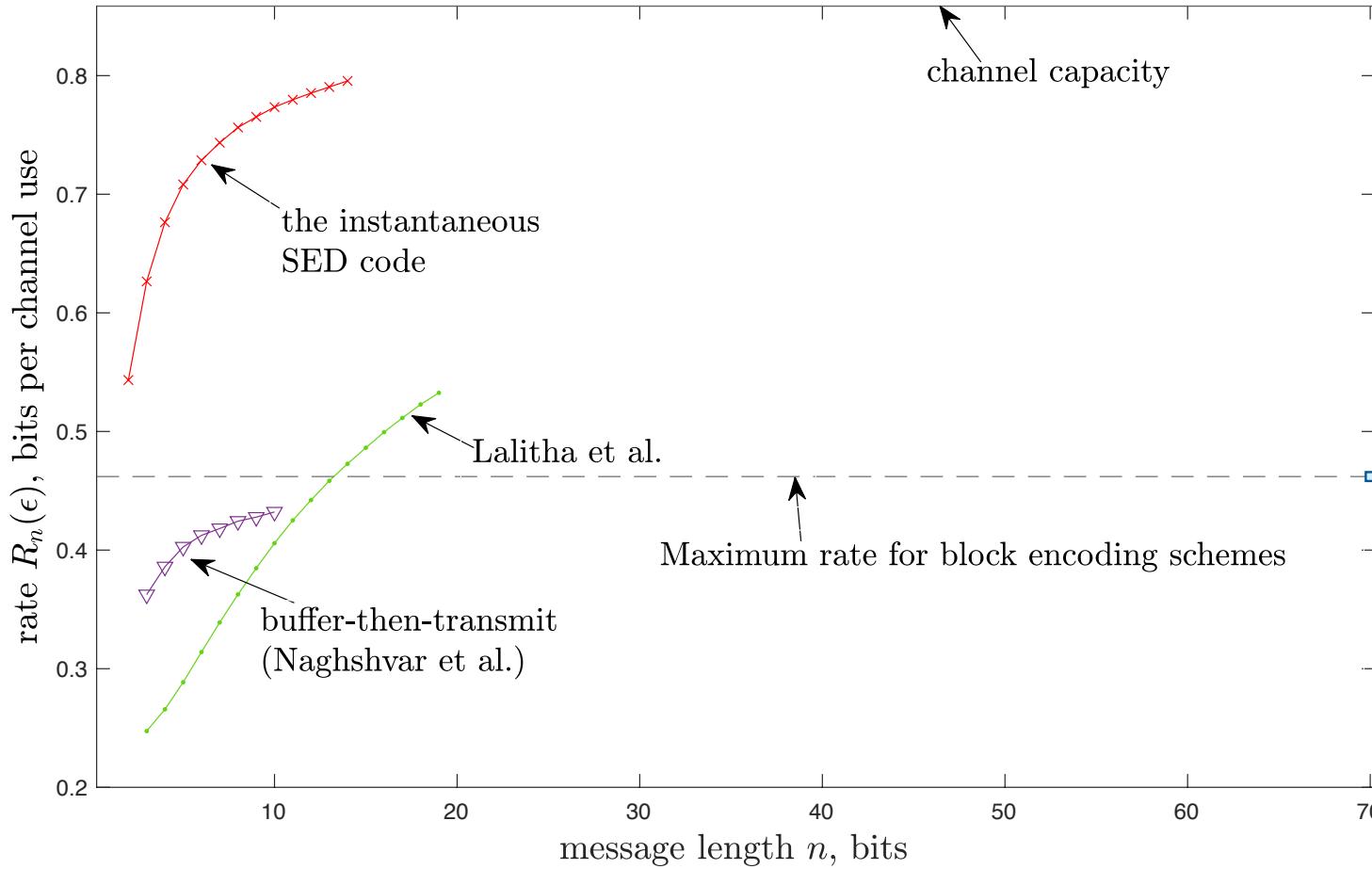
“Buffer-then-transmit” = a buffer + block encoding scheme

Introduce a delay τ_n



Instantaneous SED code: performance

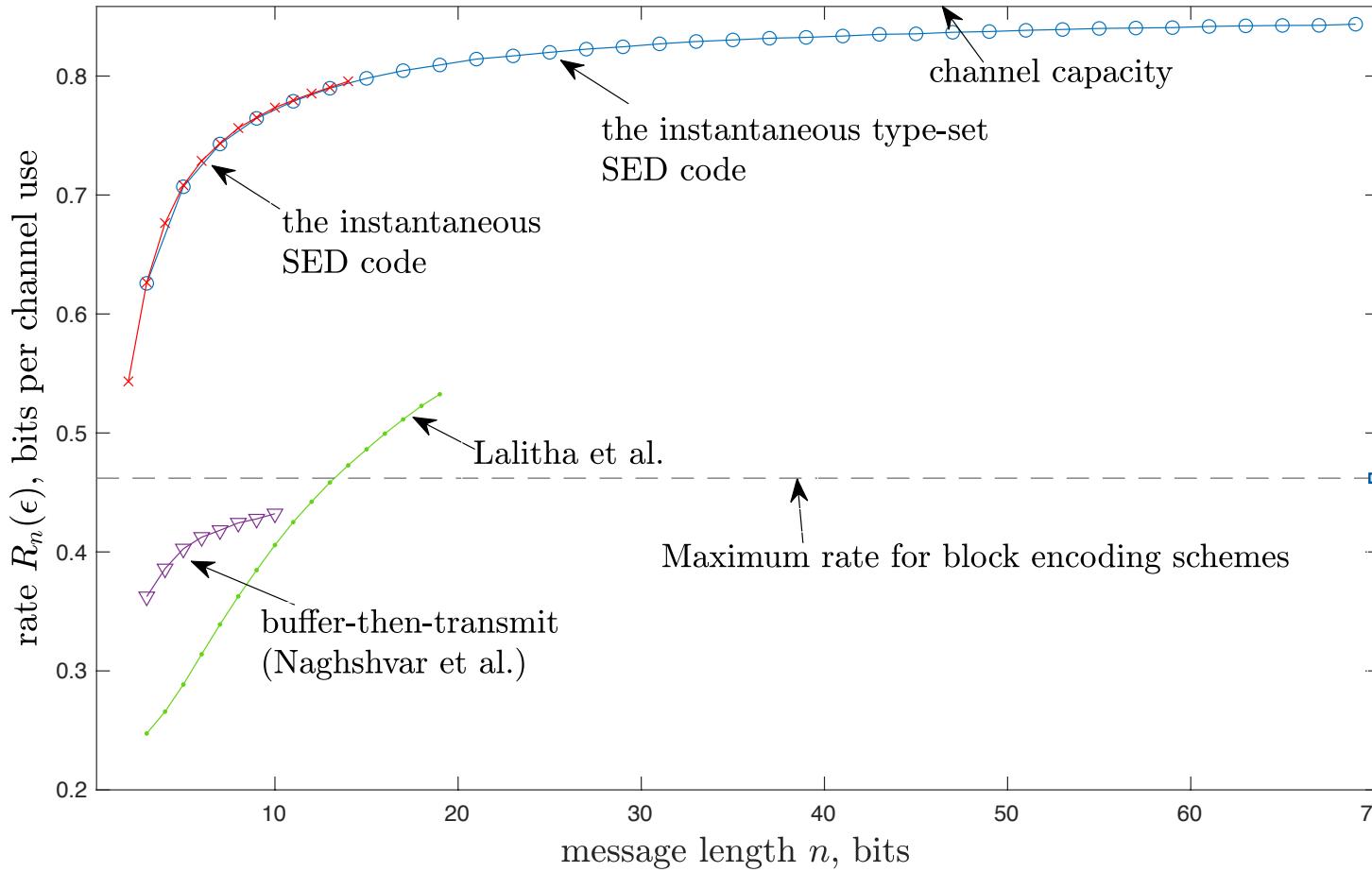
$$R_n(\epsilon) = \frac{n}{\mathbb{E}[\lambda_n]}$$



BSC(0.02), $\epsilon = 10^{-3}$. At each time $t = 1, 2, \dots, n$, a new Bernoulli($\frac{1}{2}$) bit arrives at the encoder.

Instantaneous SED code: performance

$$R_n(\epsilon) = \frac{n}{\mathbb{E}[\lambda_n]}$$



$\text{BSC}(0.02)$, $\epsilon = 10^{-3}$. At each time $t = 1, 2, \dots, n$, a new Bernoulli($\frac{1}{2}$) bit arrives at the encoder.

Instantaneous type-set SED code: setting

- Setting: BSC; n message bits;

$$\text{At time } t = 1 \quad P_{B_1^*}(0) = P_{B_1^*}(1) = 0.5$$

$$\text{At time } 2 \leq t \leq \tau_n \quad P_{B_t^*|B_{t-1}^*}(s \boxplus 1|s) = P_{B_t^*|B_{t-1}^*}(s \boxplus 0|s) = \frac{q}{2}$$

$$P_{B_t^*|B_{t-1}^*}(s|s) = 1 - q$$

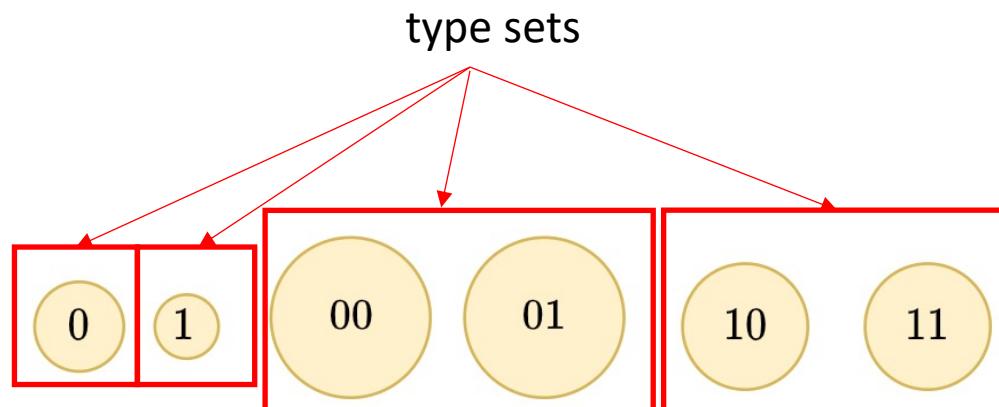
$$\text{At time } t > \tau_n \quad P_{B_t^*|B_{t-1}^*}(s \boxplus 1|s) = P_{B_t^*|B_{t-1}^*}(s \boxplus 0|s) = 0$$

$$P_{B_t^*|B_{t-1}^*}(s|s) = 1$$

key assumption: 0 and 1 arrive with the same probability

Instantaneous type-set SED code: complexity

- Instantaneous type-set SED code: double exponential → **polynomial** $O(t^4)$
- Type sets: $\mathcal{S}_1, \mathcal{S}_2, \dots$ contains strings
Strings within a type set: same length, same prior, same posterior
- Key to complexity reduction
 - type sets $O(t^2)$ → individual strings $O(2^t)$
 - type-set SED rule (approximate) → the SED rule (optimal)



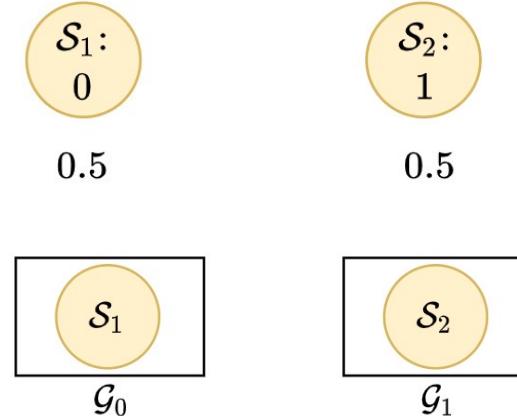
Instantaneous type-set SED code: algorithm

- example: BSC(0.4), $q = 0.8$, uniform capacity-achieving input distribution

$$t = 1, B_1^* = 0$$

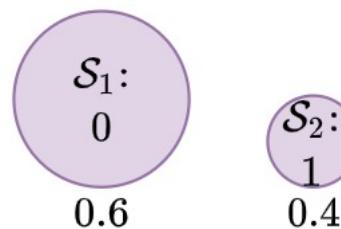
Initialization

- create type sets
- prior update
(using bit arrival distribution)
- group partition



- channel input 0, channel output 0

- posterior update
(using prior + channel output)

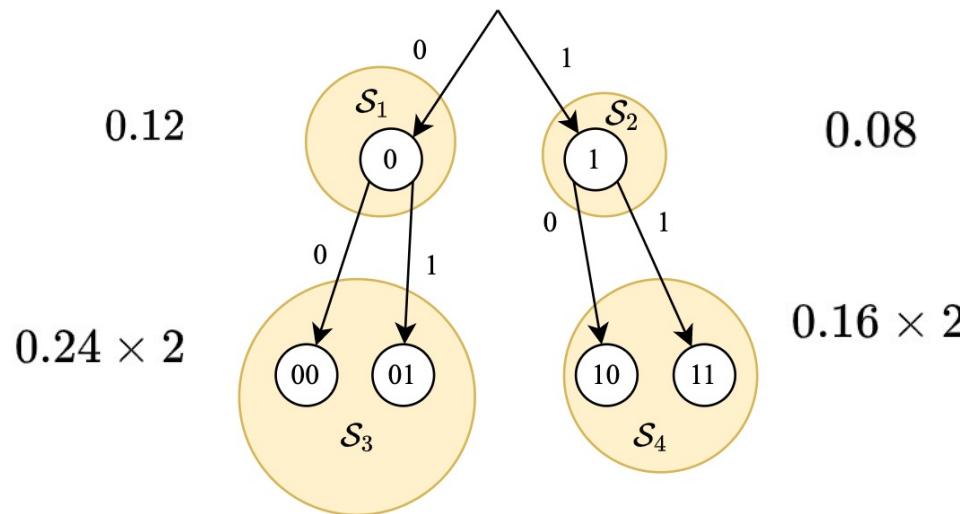


Instantaneous type-set SED code: algorithm

- example: BSC(0.4), $q = 0.8$

$$t = 2, B_2^* = 00$$

- create type sets
- prior update
(using posterior + bit arrival dist.)



Creating rules:

- Why create type sets:** accommodate the new strings $1 \leq t \leq n$
- Why generate type sets like this:** 0,1 equiprobable arrivals \rightarrow same prior

Instantaneous type-set SED code: algorithm

- example: BSC(0.4), $q = 0.8$

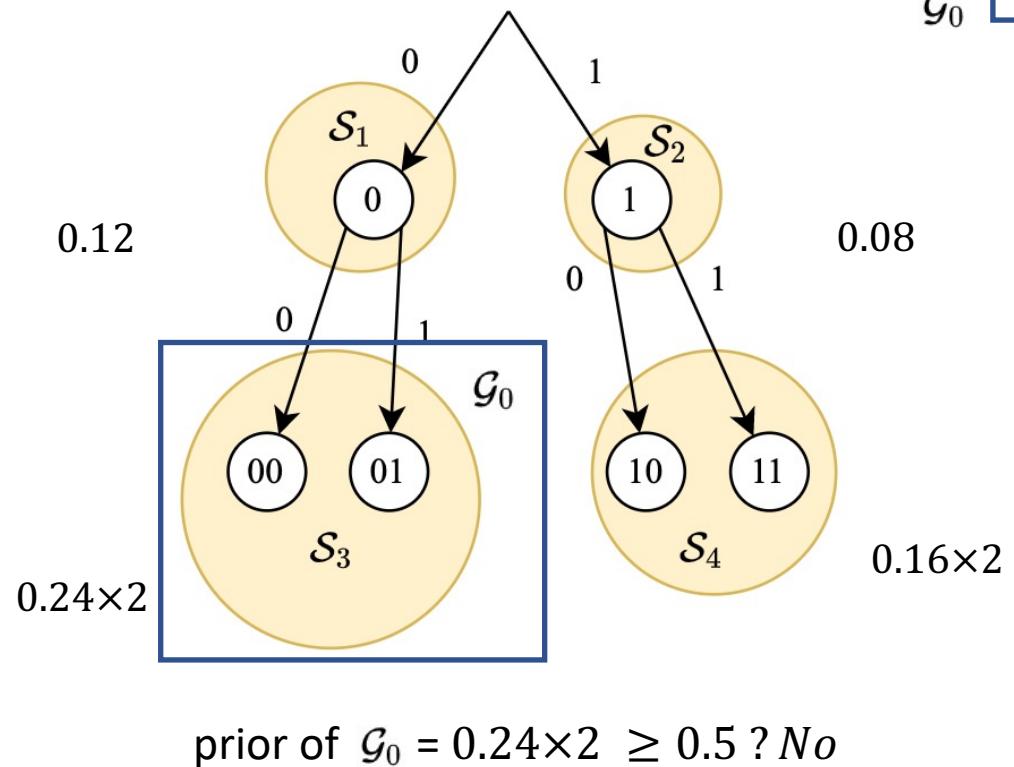
$$t = 2, B_2^* = 00$$

- create type sets
- prior update
- group partition

$$\text{sort: } \mathcal{S}_3 \geq \mathcal{S}_4 \geq \mathcal{S}_1 \geq \mathcal{S}_2$$

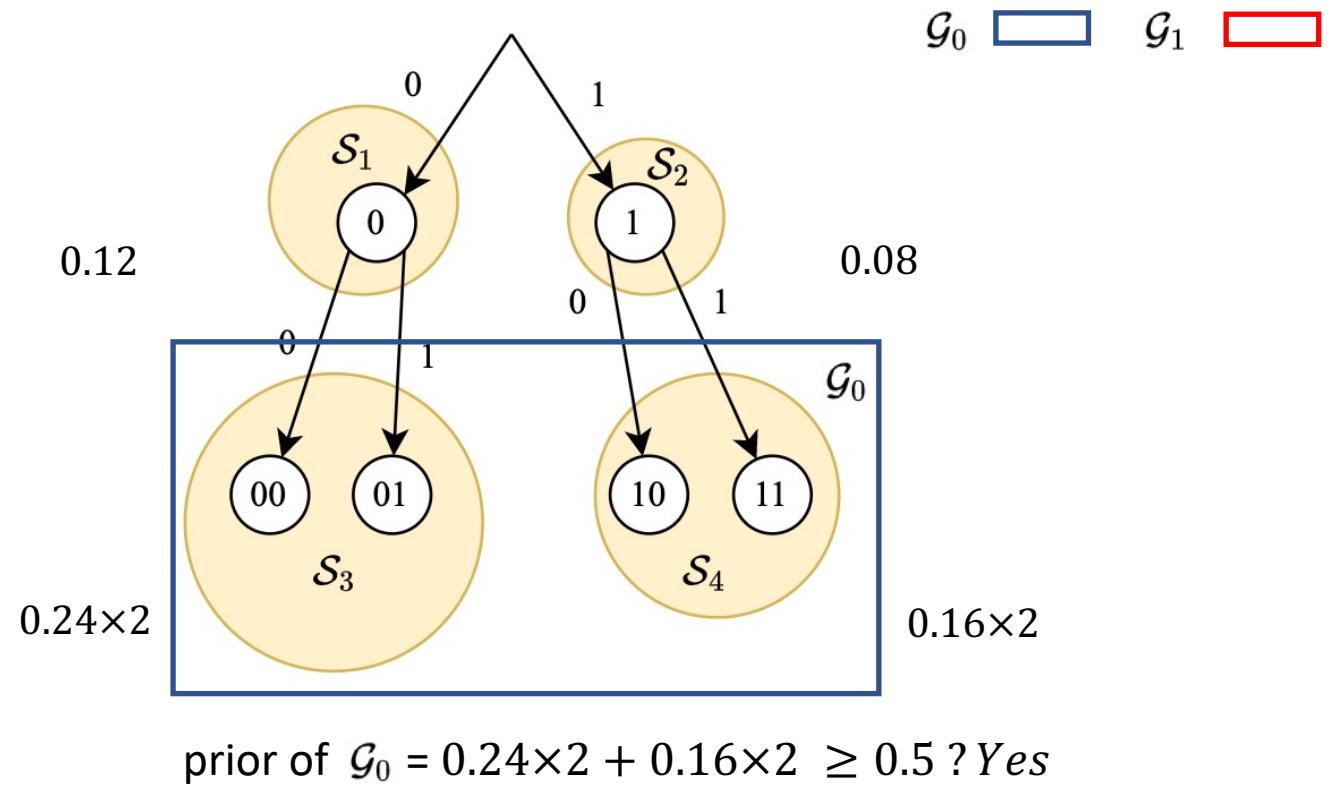
move: until prior of $\mathcal{G}_0 \geq 0.5$

\mathcal{G}_0  \mathcal{G}_1 



Instantaneous type-set SED code: algorithm

- example: BSC(0.4), $q = 0.8$
- $t = 2, B_2^* = 00$
 - create type sets
 - prior update
 - group partition
- sort: $\mathcal{S}_3 \geq \mathcal{S}_4 \geq \mathcal{S}_1 \geq \mathcal{S}_2$
- move: until prior of $\mathcal{G}_0 \geq 0.5$



Instantaneous type-set SED code: algorithm

- example: BSC(0.4), $q = 0.8$

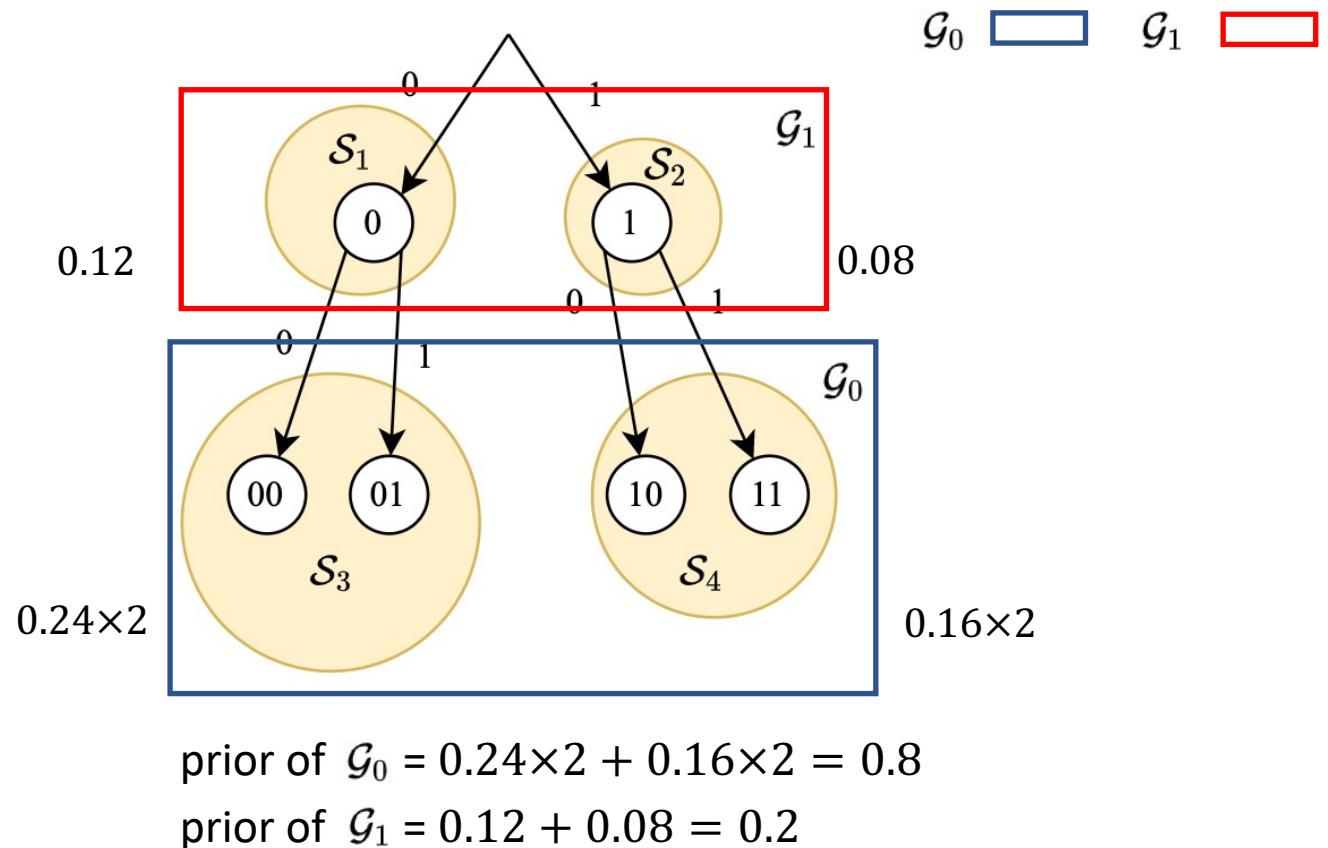
$$t = 2, B_2^* = 00$$

- create type sets
- prior update
- group partition

sort: $\mathcal{S}_3 \geq \mathcal{S}_4 \geq \mathcal{S}_1 \geq \mathcal{S}_2$

move: until prior of $\mathcal{G}_0 \geq 0.5$

split: priors closer to 0.5 and 0.5

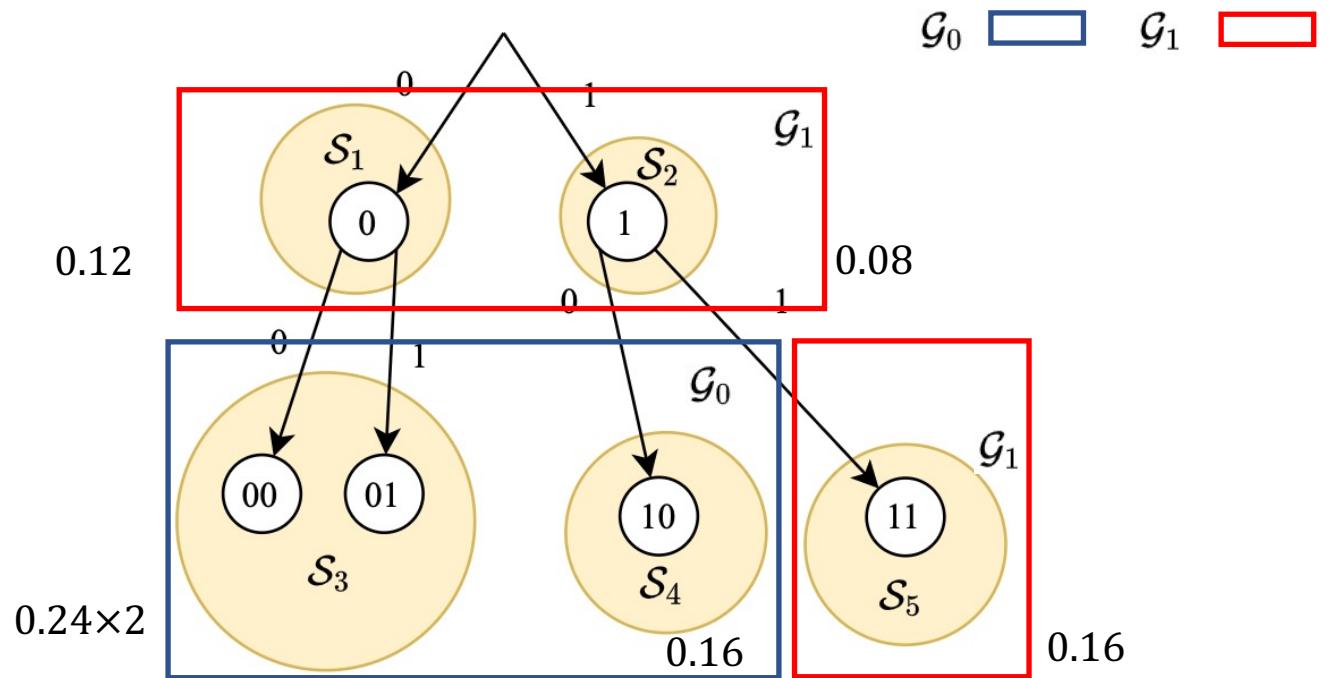


prior of $\mathcal{G}_0 = 0.24 \times 2 + 0.16 \times 2 = 0.8$

prior of $\mathcal{G}_1 = 0.12 + 0.08 = 0.2$

Instantaneous type-set SED code: algorithm

- example: BSC(0.4), $q = 0.8$
- $t = 2, B_2^* = 00$
 - create type sets
 - prior update
 - group partition
- sort: $\mathcal{S}_3 \geq \mathcal{S}_4 \geq \mathcal{S}_1 \geq \mathcal{S}_2$
- move: until prior of $\mathcal{G}_0 \geq 0.5$
- split: priors closer to 0.5 and 0.5



$$\text{prior of } \mathcal{G}_0 = 0.24 \times 2 + 0.16 = 0.64$$

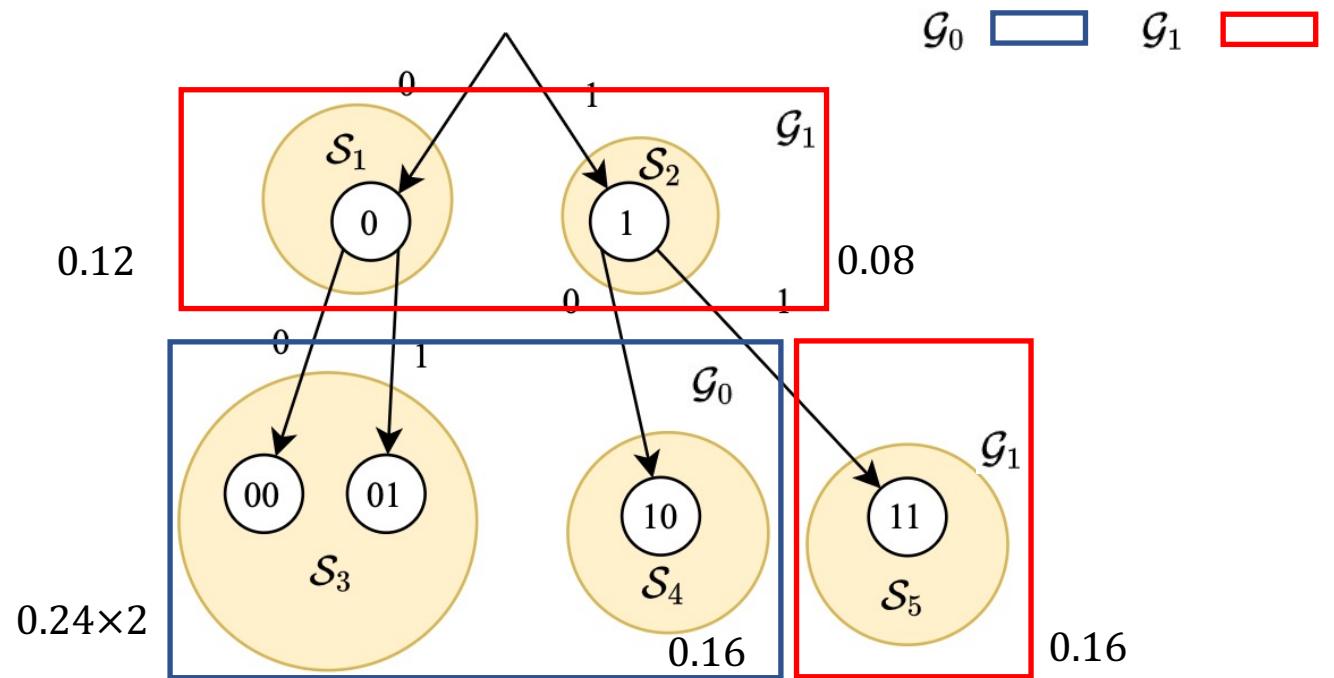
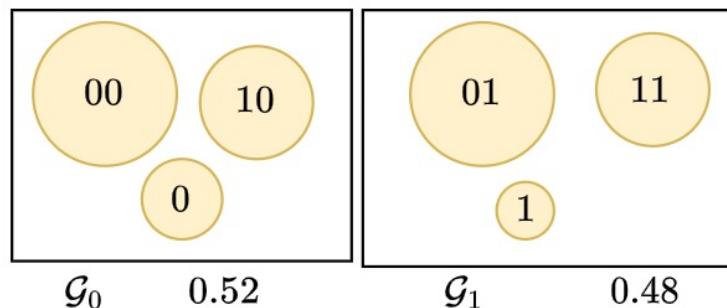
$$\text{prior of } \mathcal{G}_1 = 0.12 + 0.08 + 0.16 = 0.36$$

Instantaneous type-set SED code: algorithm

- example: BSC(0.4), $q = 0.8$

$$t = 2, B_2^* = 00$$

- create type sets
- prior update
- group partition



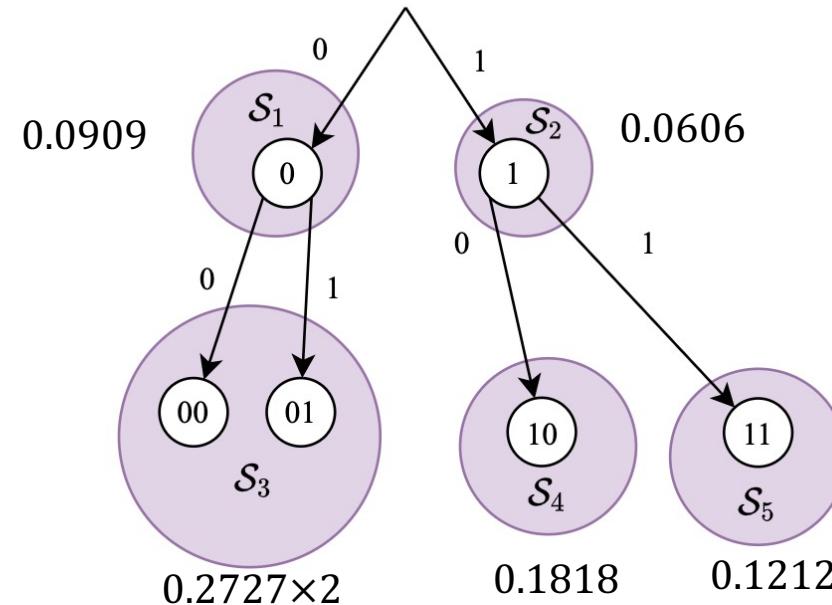
prior of $\mathcal{G}_0 = 0.24 \times 2 + 0.16 = 0.64$

prior of $\mathcal{G}_1 = 0.12 + 0.08 + 0.16 = 0.36$

type-set SED rule: potentially suboptimal

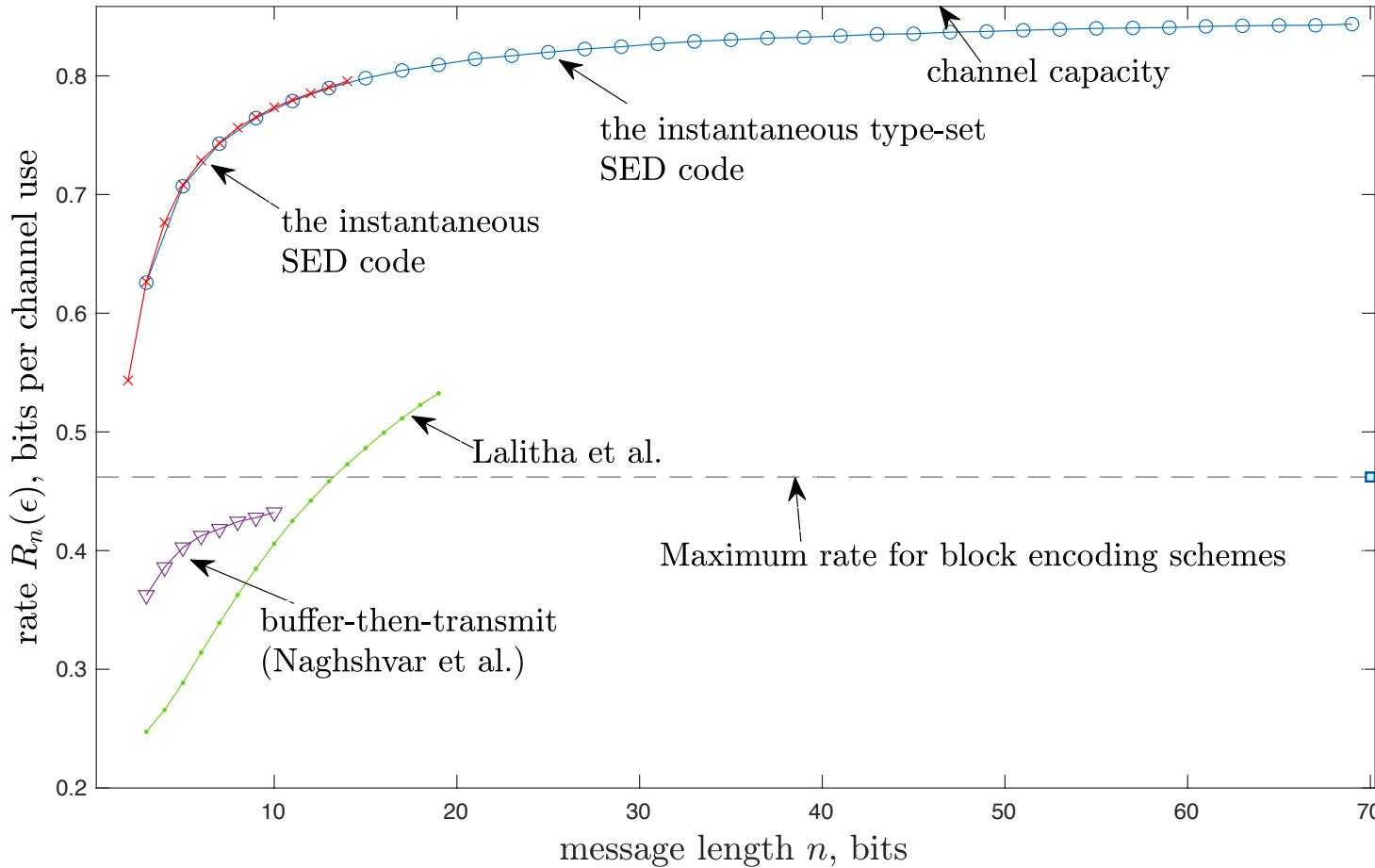
Instantaneous type-set SED code: algorithm

- example: BSC(0.4), $q = 0.8$
 - $t = 2, B_2^* = 00$
 - channel input 0, channel output 0
 - posterior update
(using prior + channel output)
 - iterations stop: posterior $\geq 1 - \epsilon$
 - cardinality = 1
 - length = n



Instantaneous type-set SED code: performance

$$R_n(\epsilon) = \frac{n}{\mathbb{E}[\lambda_n]}$$



BSC(0.02), $\epsilon = 10^{-3}$. At each time $t = 1, 2, \dots, n$, a new Bernoulli($\frac{1}{2}$) bit arrives at the encoder.

Instantaneous SED code: reliability function

$$\epsilon_n^*(R) \triangleq \min \left\{ \epsilon: \exists \left(n, \frac{n}{R}, \epsilon \right) \text{ feedback code with instantaneous encoding} \right\} \text{ (minimum error probability)}$$

of message bits expected decoding time error probability

$$E(R) \triangleq \lim_{n \rightarrow \infty} \frac{R}{n} \log_2 \frac{1}{\epsilon_n^*(R)} \quad \text{(reliability function)}$$

Instantaneous SED code: reliability function

Define the divergence $C_1 \triangleq \max_{x_1, x_2 \in \{0,1\}} D(P_{Y|X=x_1} || P_{Y|X=x_2})$

Theorem: Fix a bit arrival distribution that satisfies $\tau_n \leq \mathbb{E}[\tau_n] + o(n)$, a.s. and fix a 2-input DMC such that the channel capacity C is achieved by $P_X^*(0) = P_X^*(1) = 0.5$; and under some other regularity conditions, the reliability function is lower bounded by

$$E(R) \geq C_1 \left(1 - \lim_{n \rightarrow \infty} \left(\frac{H(B_{\mathbb{E}[\tau_n]+o(n)}^* | Y^{\mathbb{E}[\tau_n]+o(n)})}{nC} + \frac{\mathbb{E}[\tau_n]}{n} \right) R \right) \quad \text{better}$$

where Y_1, Y_2, \dots are the channel outputs in response to the channel inputs generated by the encoder of the instantaneous SED code.

Buffer-then-transmit

$$E(R) \geq C_1 \left(1 - \lim_{n \rightarrow \infty} \left(\frac{1}{C} + \frac{\mathbb{E}[\tau_n]}{n} \right) R \right)$$

Conclusion

- **Novel codes:** instantaneous SED code and instantaneous type-set SED code (low complexity)
- **First work:** decoder does not know the bit arrival times
- **Significant gain:** outperforms existing schemes
- **Future research:**
 - reliability function over a general DMC/ converse bound
 - distributed control systems
 - joint source-channel coding
 - Gaussian channels/channel with memory
 - age of information
 - ...