Universally Decodable Matrices for Distributed Matrix-Vector Multiplication

Pascal O. Vontobel
(joint work with Aditya Ramamoorthy and Li Tang)

Department of Information Engineering
The Chinese University of Hong Kong

WPI 2019, Hong Kong, August 19, 2019

Motivation

Motivation

Given:

- a matrix $\mathbf{A}$ of size $m \times n$ over the reals;
- a vector $\mathbf{x}$ of length n over the reals.

Task: compute the vector $\mathbf{y}$ of length m over the reals, where

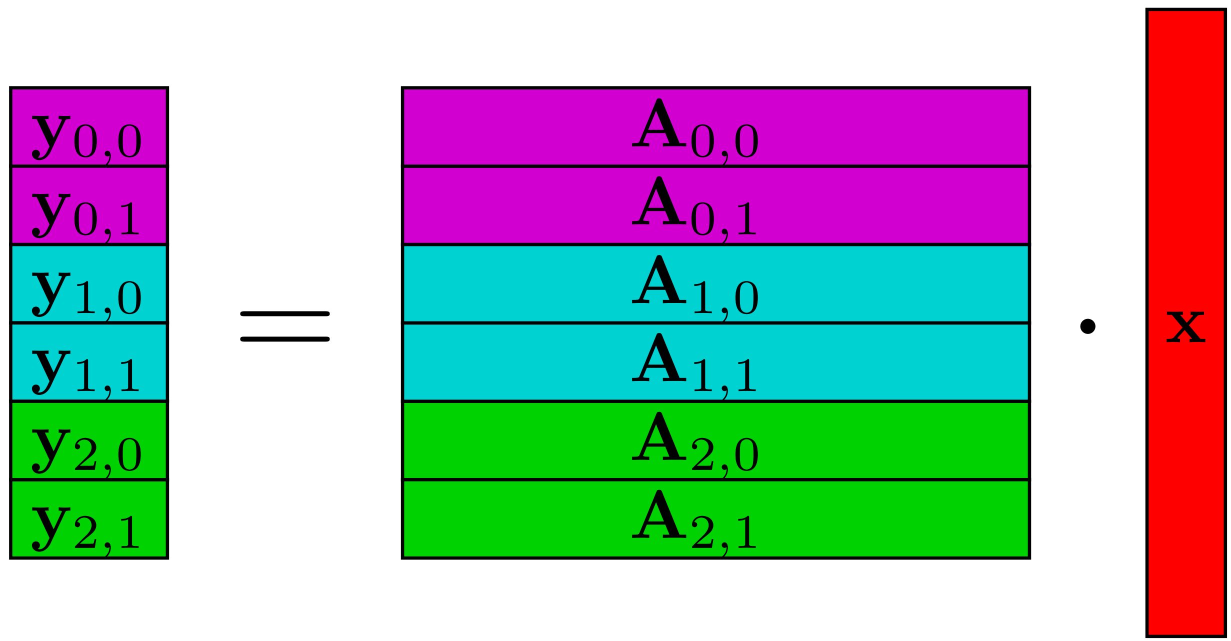
$$\mathbf{y} \triangleq \mathbf{A} \cdot \mathbf{x} \,.$$

Explicitly:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \triangleq \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \cdots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} .$$

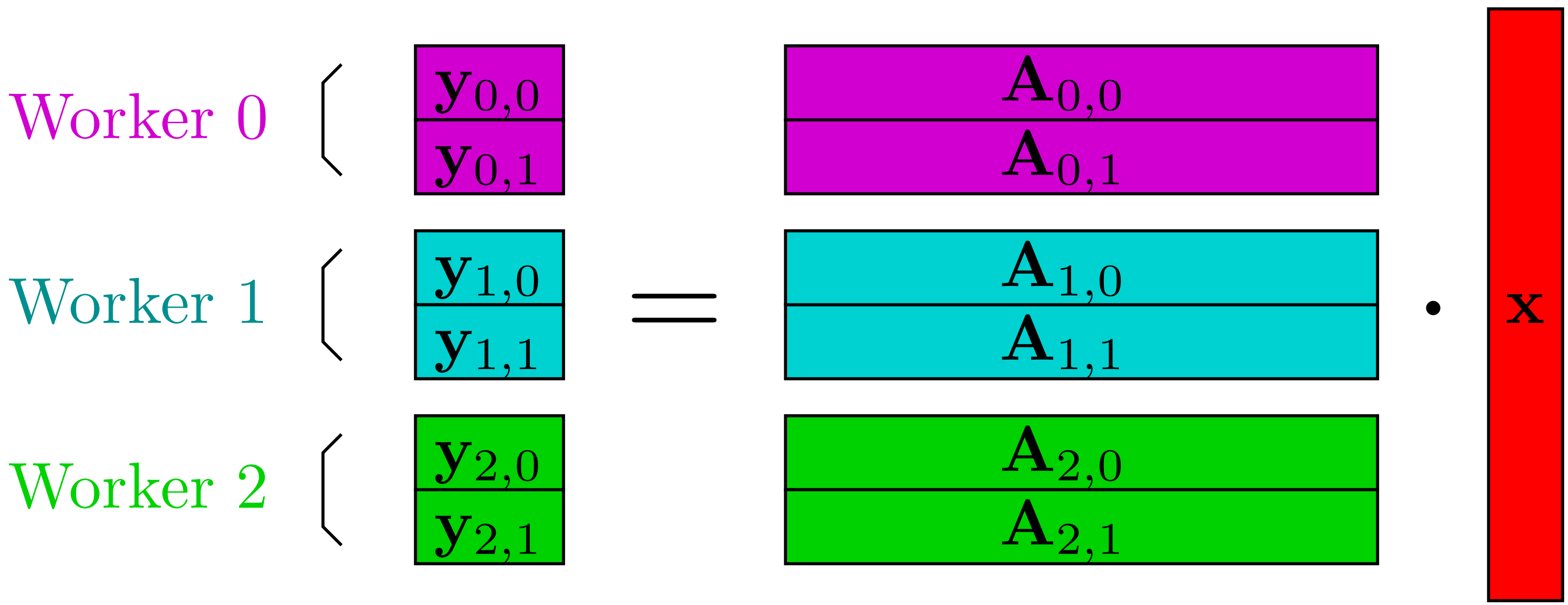
Motivation

We can split up the task into several submatrix-vector-multiplication tasks:

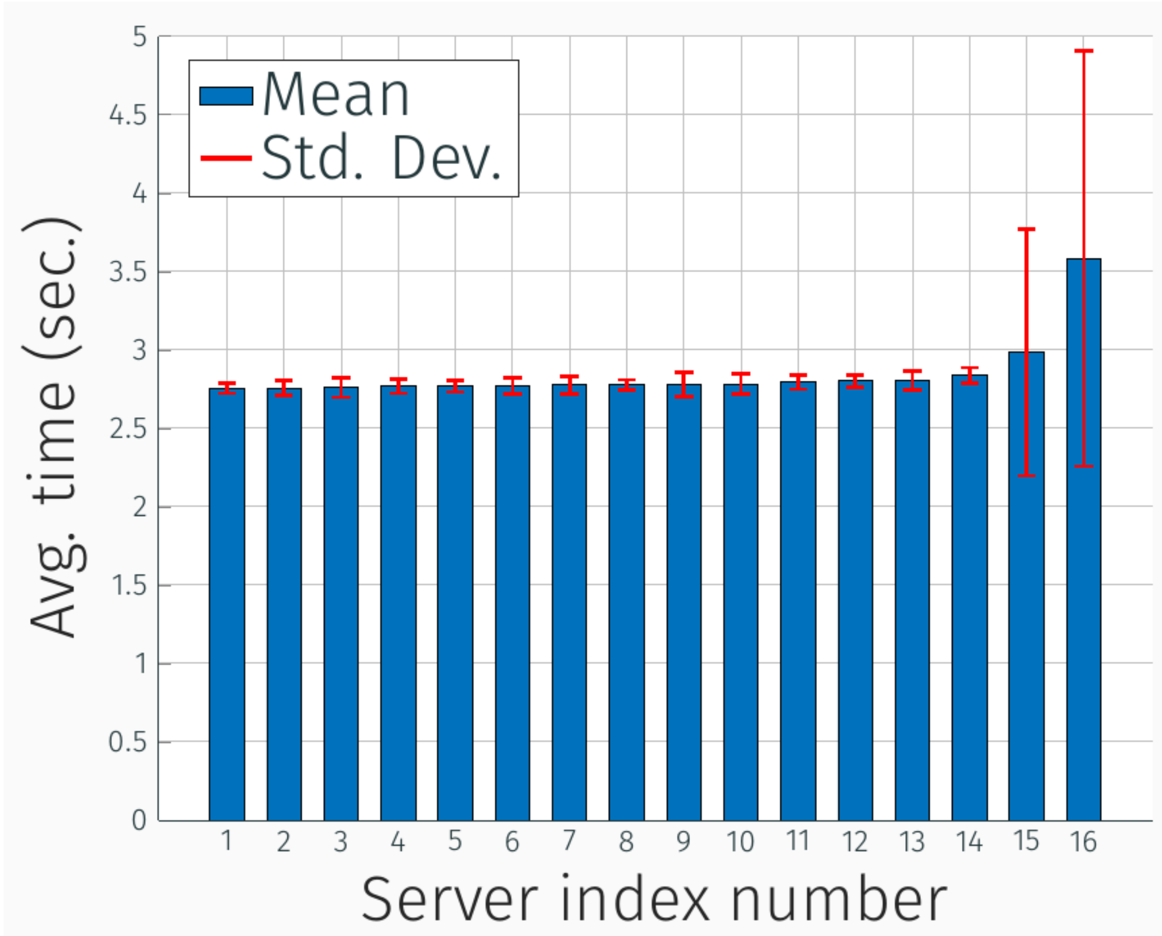


Motivation

We can split up the task into several submatrix-vector-multiplication tasks:

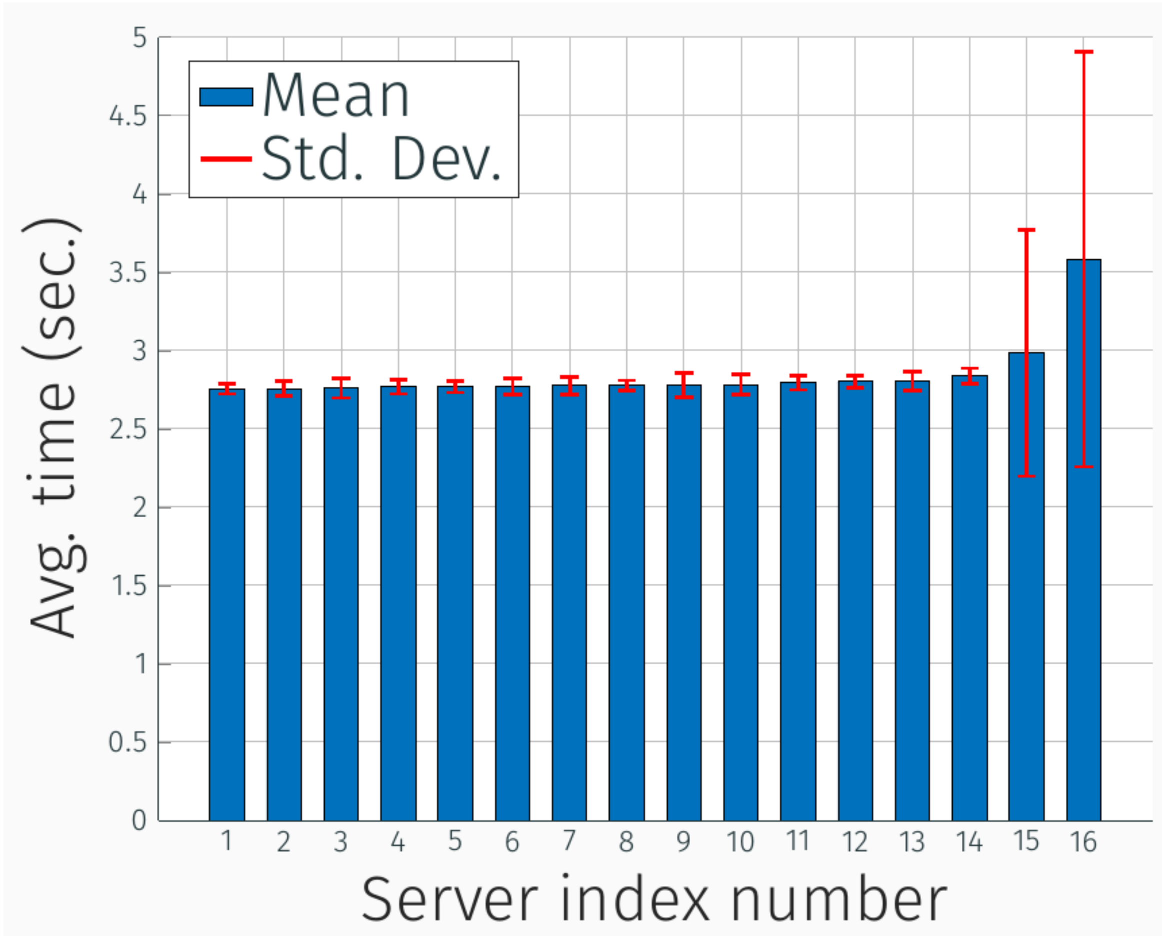


Motivation



Mean
Std. Dev.
Avg. time (sec.)
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Server index number

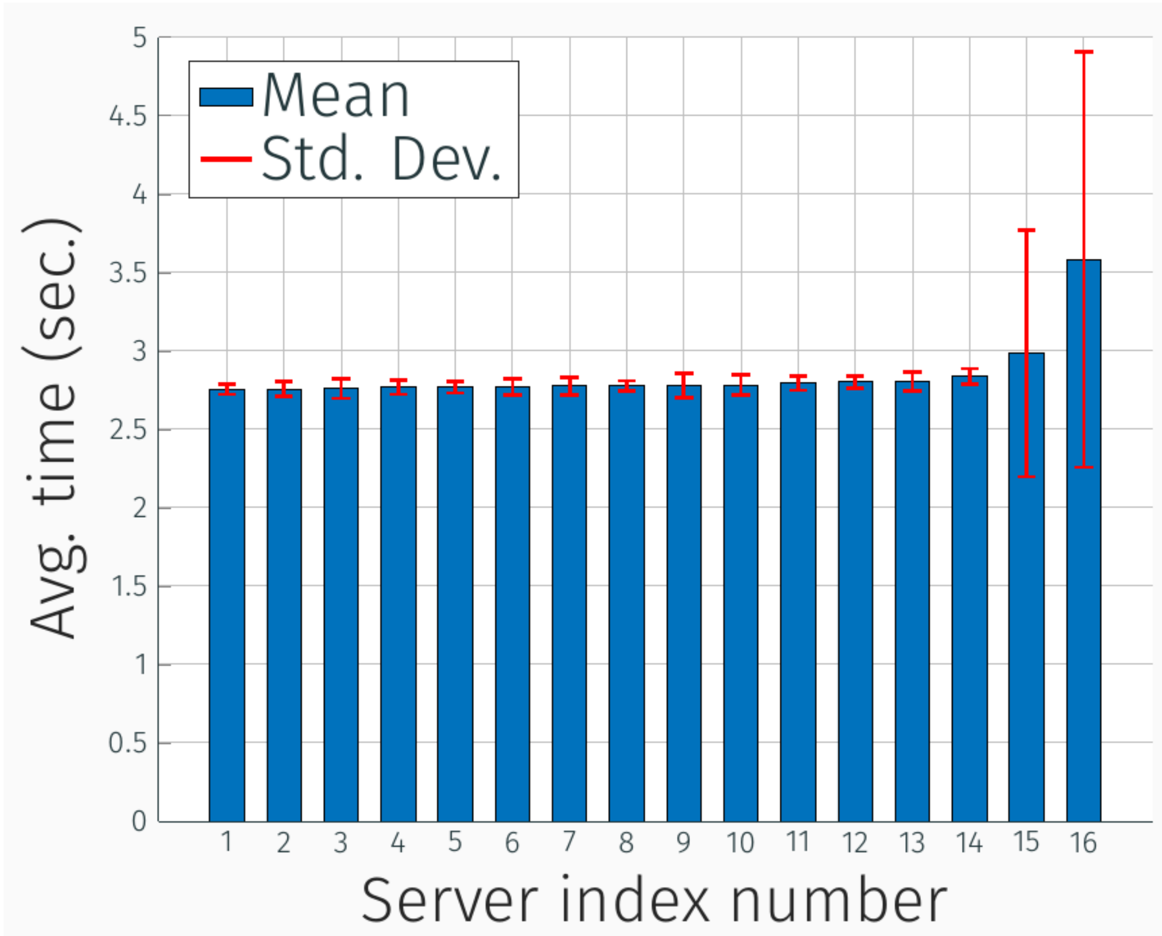
Motivation



Idea:

- Use **coding theory** to alleviate delay issues because of **stragglers**.

Motivation

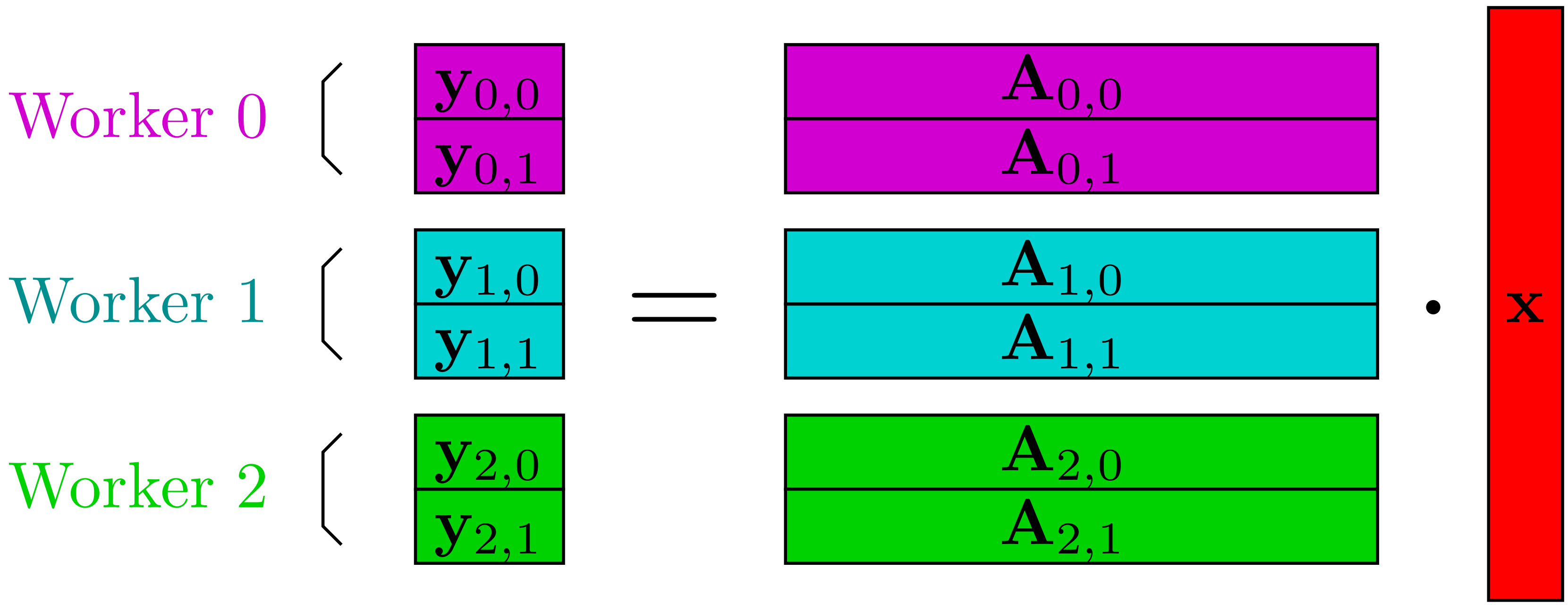


Idea:

- Use **coding theory** to alleviate delay issues because of **stragglers**.
- **Unavailable partial results** can be seen as **erasures**.

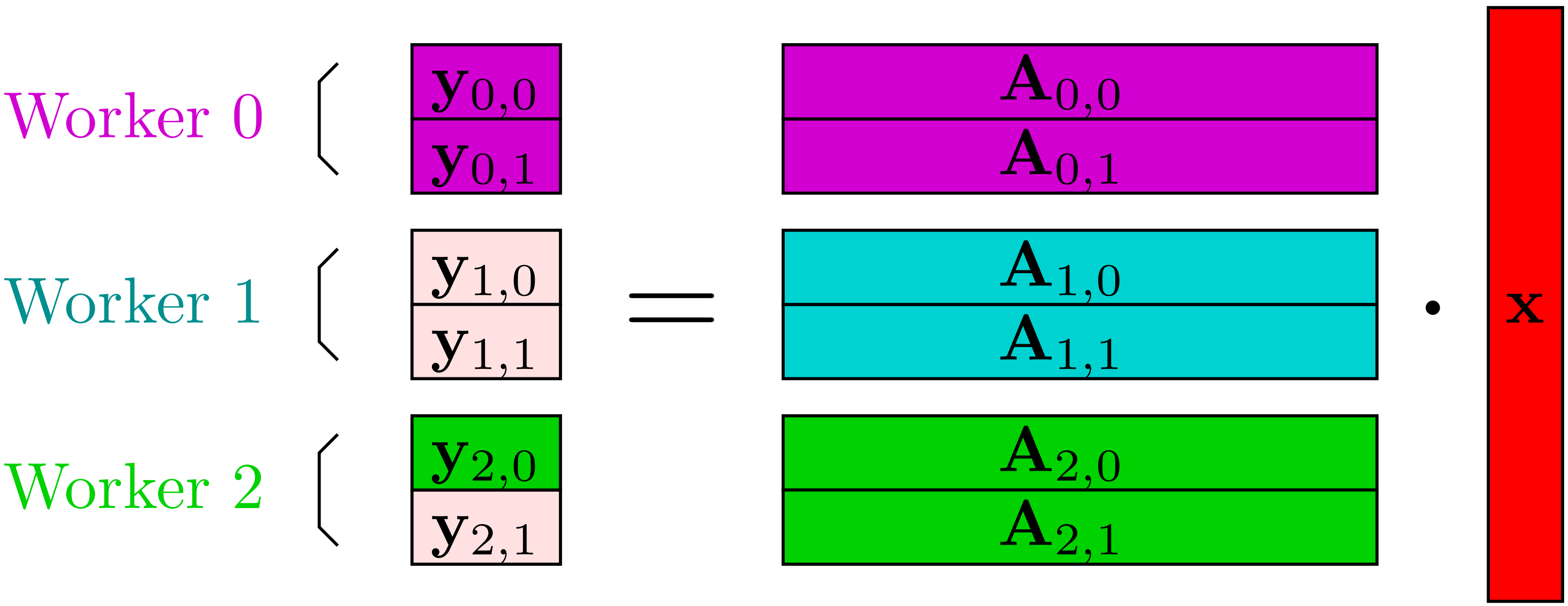
Motivation

We can split up the task into several submatrix-vector-multiplication tasks:



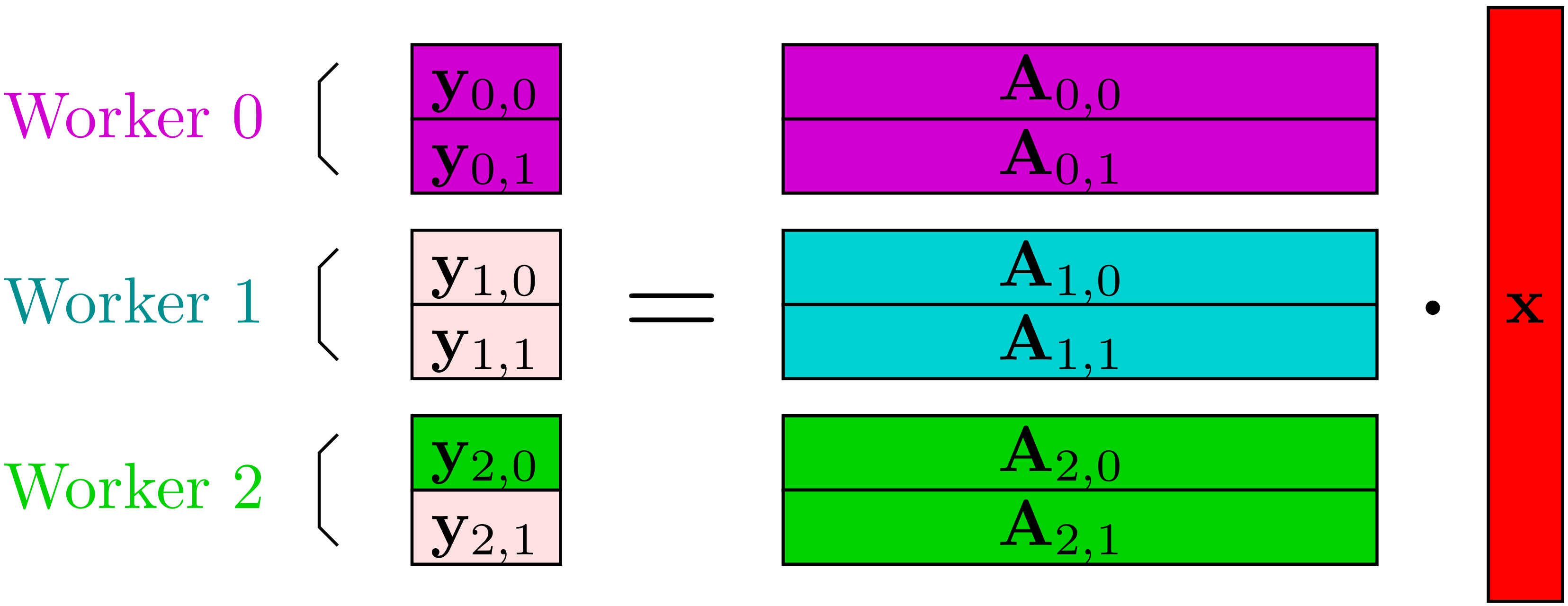
Motivation

We can split up the task into several submatrix-vector-multiplication tasks:



Motivation

We can split up the task into several submatrix-vector-multiplication tasks:



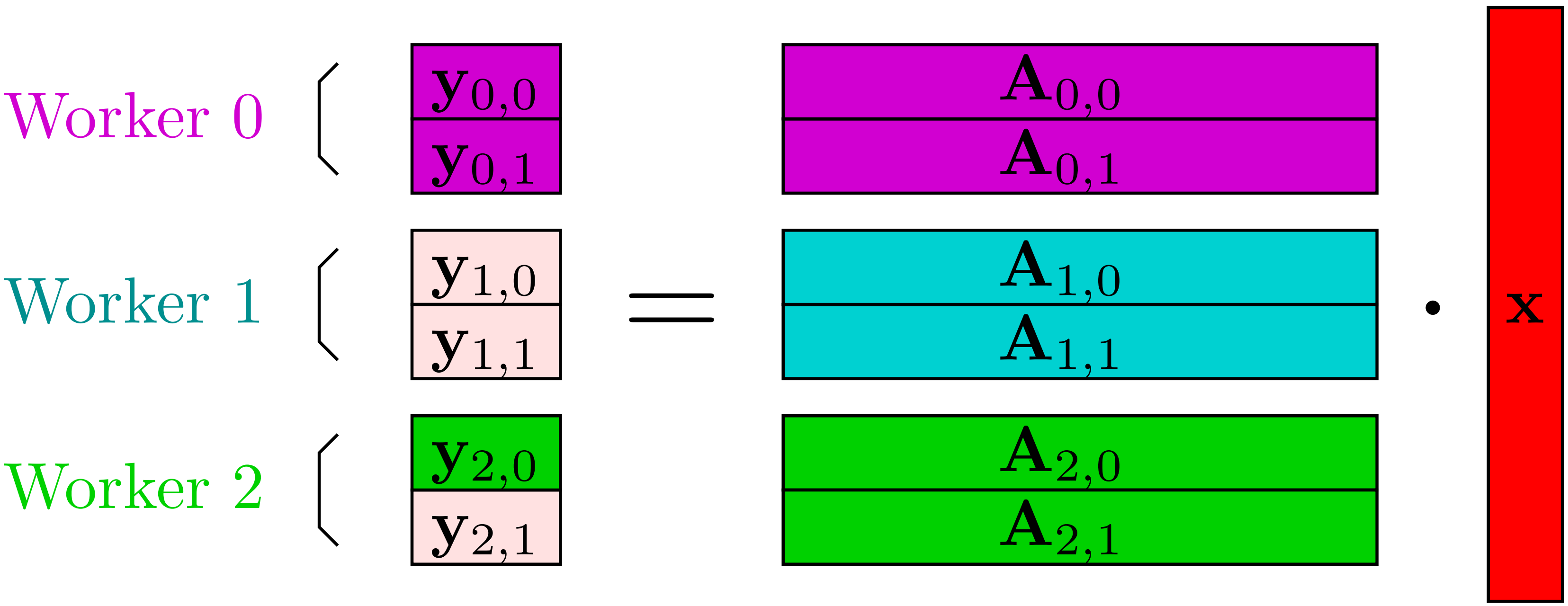
Idea:

- Coding scheme should take advantage of the fact that **erasures are correlated**.

Erasures are correlated because
if a partial result by one of the workers is not available,
then **all subsequent results by the same worker** are not available either.

Motivation

We can split up the task into several submatrix-vector-multiplication tasks:

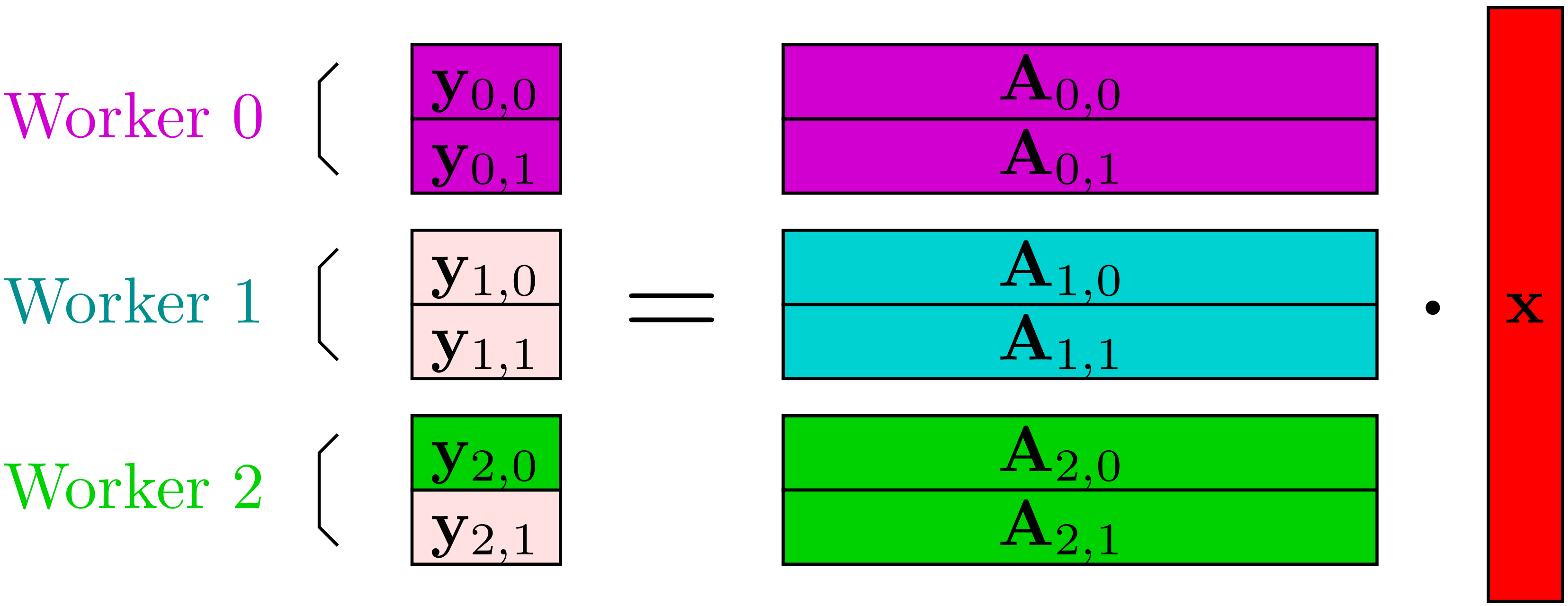


Idea:

- Base coding scheme on so-called **universally decodable matrices (UDMs)**.

Motivation

We can split up the task into several submatrix-vector-multiplication tasks:

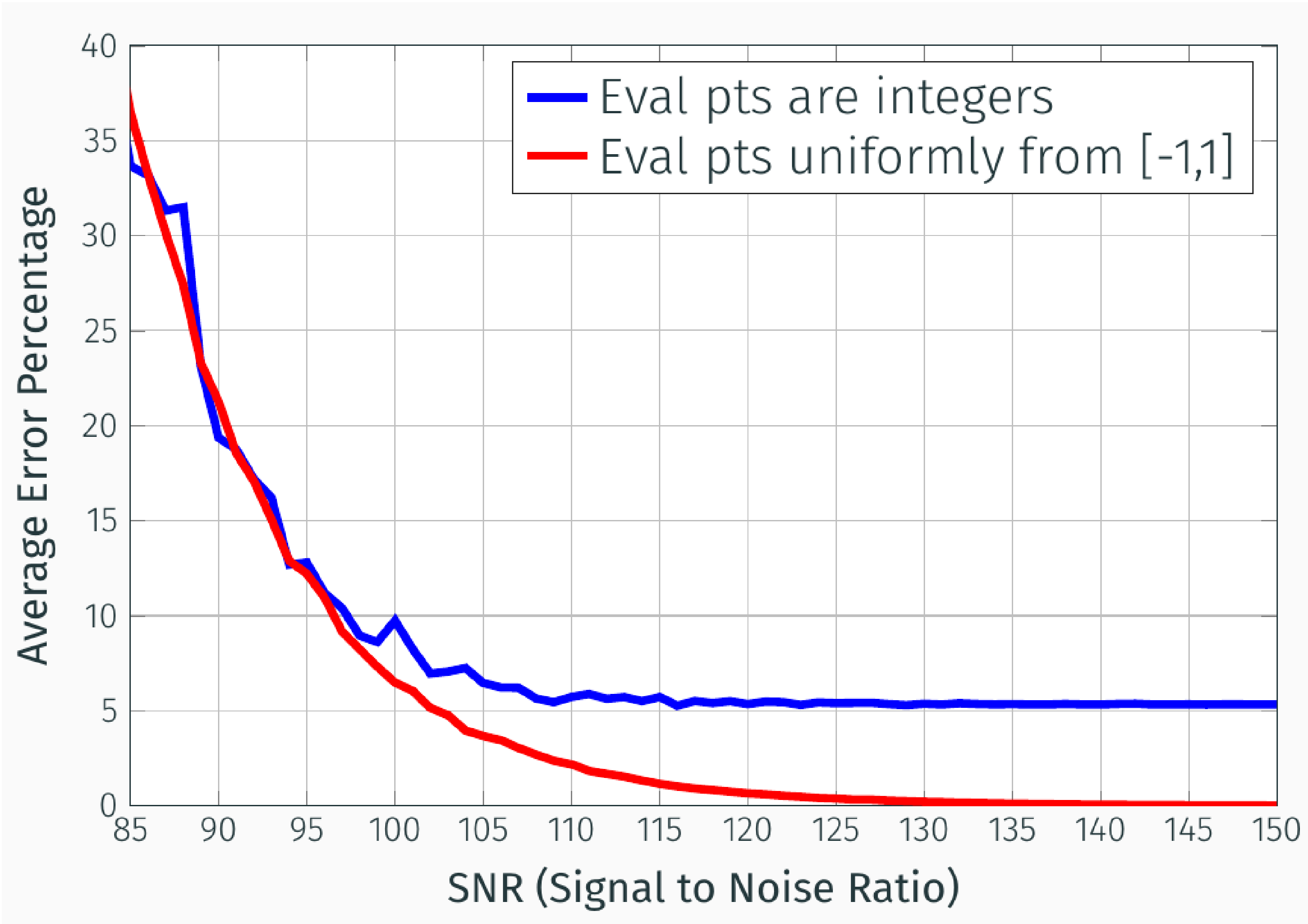


Idea:

- Base coding scheme on so-called **universally decodable matrices (UDMs)**.
- Use **companion matrices** in order to **reduce issues with condition numbers** when adapting a coding scheme over some finite field to a coding scheme over the reals.

Motivation

Bad condition number of unsuitably chosen encoding matrices is an issue.



Context (Part 1/2)

- Q. Yu, M. Maddah-Ali, and S. Avestimehr, "Polynomial codes: an optimal design for high-dimensional coded matrix multiplication," in Proc. of Adv. in Neural Inf. Proc. Sys. (NIPS), 2017, pp. 4403–4413.
- L. Tang, K. Konstantinidis, and A. Ramamoorthy, "Erasure coding for distributed matrix multiplication for matrices with bounded entries," IEEE Comm. Lett., vol. 23, no. 1, pp. 8–11, 2019.
- K. Lee, C. Suh, and K. Ramchandran, "High-dimensional coded matrix multiplication," in IEEE Int. Symp. Inf. Theory, 2017, pp. 2418–2422.
- K. Lee, M. Lam, R. Pedarsani, D. Papailiopoulos, and K. Ramchandran, "Speeding up distributed machine learning using codes," IEEE Trans. Inf. Theory, vol. 64, no. 3, pp. 1514–1529, 2018.
- S. Dutta, V. Cadambe, and P. Grover, "Short-dot: Computing large linear transforms distributedly using coded short dot products," in Proc. of Adv. in Neural Inf. Proc. Sys. (NIPS), 2016, pp. 2100–2108.

Context (Part 2/2)

- A. Mallick, M. Chaudhari, and G. Joshi, "Rateless codes for near-perfect load balancing in distributed matrix-vector multiplication," preprint, 2018. arXiv: 1804.10331.
- S. Wang, J. Liu, and N. B. Shroff, "Coded sparse matrix multiplication," in Proc. 35th Int. Conf. Mach. Learning, ICML, 2018, pp. 5139–5147.
- S. Kiani, N. Ferdinand, and S. C. Draper, "Exploitation of stragglers in coded computation," in IEEE Int. Symp. Inf. Theory, 2018, pp. 1988–1992.
- A. B. Das, L. Tang, and A. Ramamoorthy, "C^3LES: Codes for coded computation that leverage stragglers," in IEEE Inf. Th. Workshop, 2018, pp. 1–5.
- N. Raviv, Y. Cassuto, R. Cohen, and M. Schwartz, "Erasure correction of scalar codes in the presence of stragglers," in IEEE Int. Symp. Inf. Theory, 2018, pp. 1983–1987.
- N. Raviv, Q. Yu, J. Bruck, and S. Avestimehr, "Download and access tradeoffs in Lagrange coded computing," in IEEE Int. Symp. Inf. Theory, 2019.

Overview

Overview

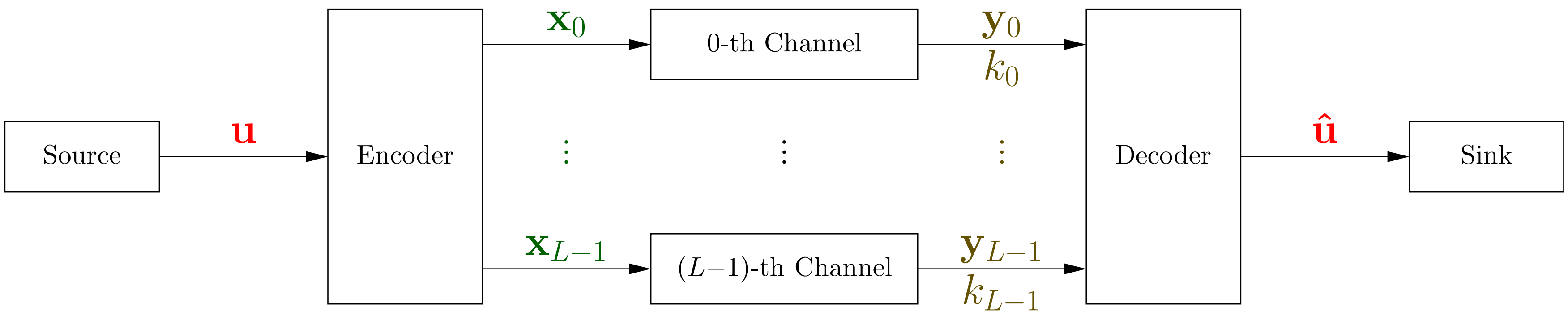
- Motivation
- A communication system with L parallel channels
  $\Rightarrow$ Coding for this system using universally decodable matrices
- Embedding into the reals
  $\Rightarrow$ Companion matrices

For more details:

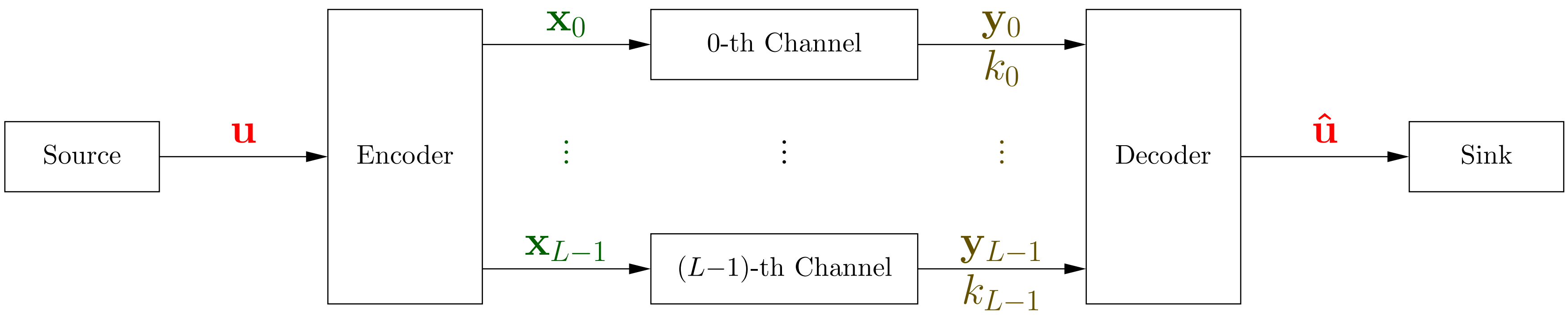
- A. Ramamoorthy, L. Tang, and P. O. Vontobel, "Universally decodable matrices for distributed matrix-vector multiplication," Proc. IEEE Int. Symp. Inf. Theory, Paris, France, pp. 1777-1781, July 2019.
- arXiv: 1901.10674

Communication system with L parallel channels

Comm. System with L Parallel Channels



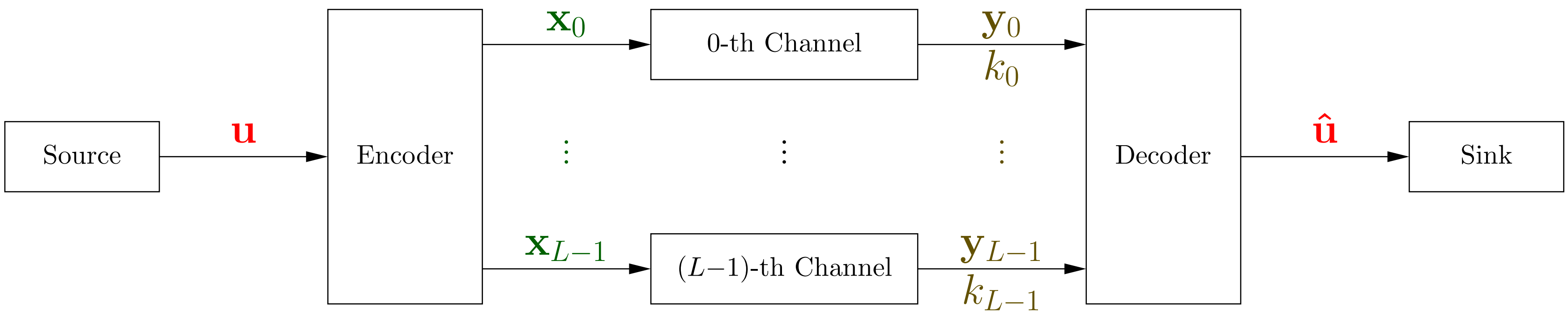
Comm. System with L Parallel Channels



$$\begin{pmatrix} u_0 & \cdots & u_{n-1} \end{pmatrix} \mapsto \begin{pmatrix} x_{0,0} & \cdots & x_{0,n-1} \\ \vdots & \vdots & \vdots \\ x_{L-1,0} & \cdots & x_{L-1,n-1} \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} y_{0,0} & \cdots & y_{0,n-1} \\ \vdots & \vdots & \vdots \\ y_{L-1,0} & \cdots & y_{L-1,n-1} \end{pmatrix} \Rightarrow \begin{pmatrix} \hat{u}_0 & \cdots & \hat{u}_{n-1} \end{pmatrix}$$

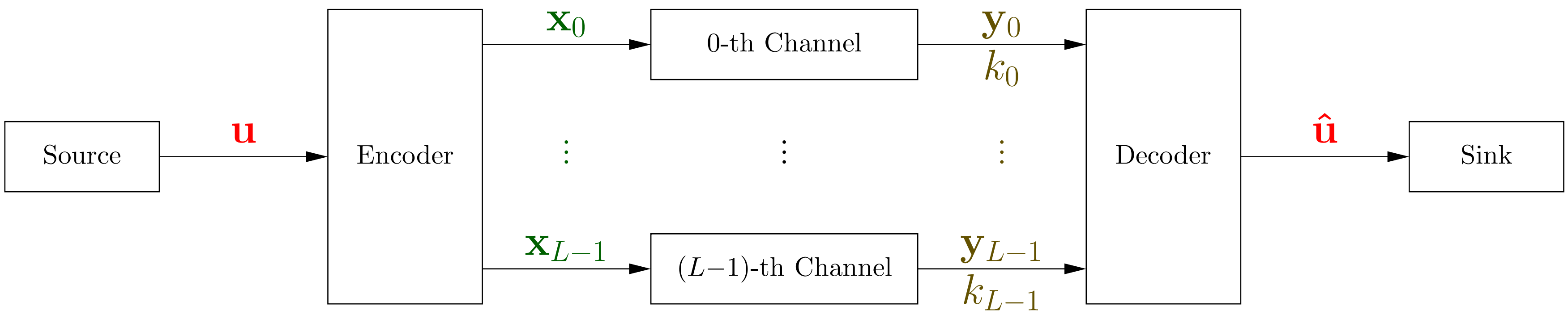
Comm. System with L Parallel Channels



E.g. $L = 4$, $n = 3$.

$$\begin{pmatrix} u_0 & u_1 & u_2 \end{pmatrix} \mapsto \begin{pmatrix} x_{0,0} & x_{0,1} & x_{0,2} \\ x_{1,0} & x_{1,1} & x_{1,2} \\ x_{2,0} & x_{2,1} & x_{2,2} \\ x_{3,0} & x_{3,1} & x_{3,2} \end{pmatrix} \Rightarrow \begin{pmatrix} y_{0,0} & y_{0,1} & y_{0,2} \\ y_{1,0} & y_{1,1} & y_{1,2} \\ y_{2,0} & y_{2,1} & y_{2,2} \\ y_{3,0} & y_{3,1} & y_{3,2} \end{pmatrix} \Rightarrow \begin{pmatrix} \hat{u}_0 & \hat{u}_1 & \hat{u}_2 \end{pmatrix}$$

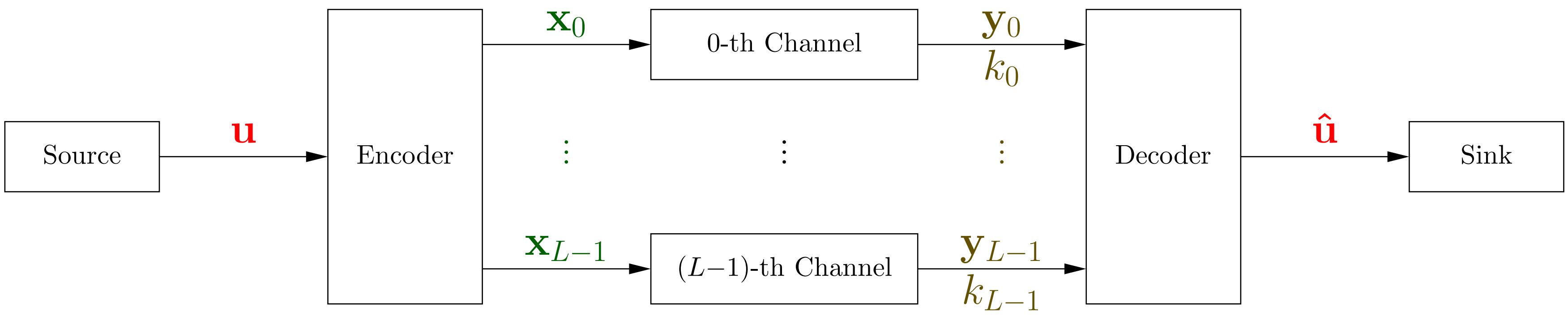
Comm. System with L Parallel Channels



E.g. $L = 4$, $n = 3$, $q = 3$.

$$\begin{pmatrix} 1 & 1 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 2 \\ 2 & 0 & 2 \end{pmatrix} \Rightarrow \begin{pmatrix} ? & ? & ? \\ \mathbf{2} & ? & ? \\ ? & ? & ? \\ \mathbf{2} & \mathbf{0} & ? \end{pmatrix} \Rightarrow \begin{pmatrix} \hat{u}_0 & \hat{u}_1 & \hat{u}_2 \end{pmatrix}$$

Comm. System with L Parallel Channels

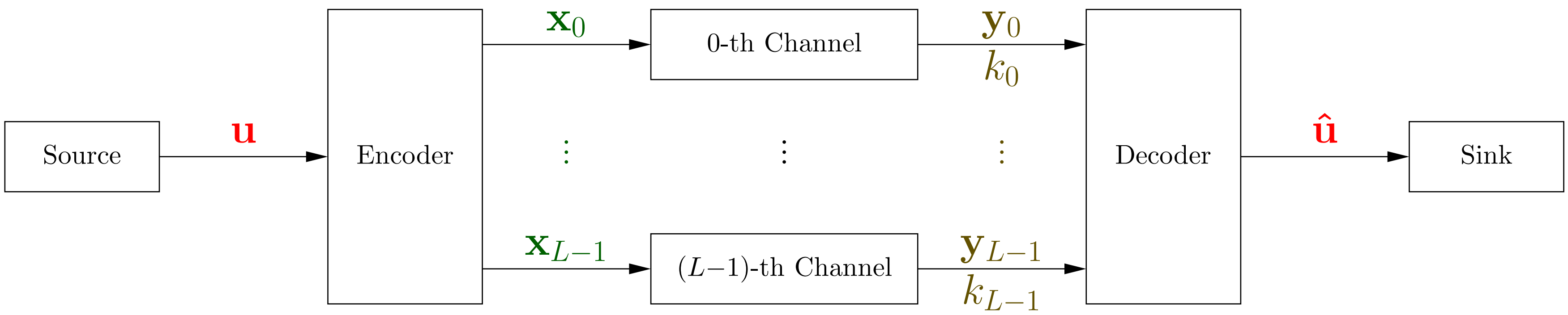


E.g. $L = 4$, $n = 3$, $q = 3$.

$$\begin{pmatrix} 1 & 1 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 2 \\ 2 & 0 & 2 \end{pmatrix} \Rightarrow \begin{pmatrix} ? & ? & ? \\ \mathbf{2} & ? & ? \\ ? & ? & ? \\ \mathbf{2} & \mathbf{0} & ? \end{pmatrix} \Rightarrow \begin{pmatrix} \hat{u}_0 & \hat{u}_1 & \hat{u}_2 \end{pmatrix}$$

The channels are such that if $y_{\ell,t}$ is erased then also $y_{\ell,t'}$ is erased for all $t' > t$.

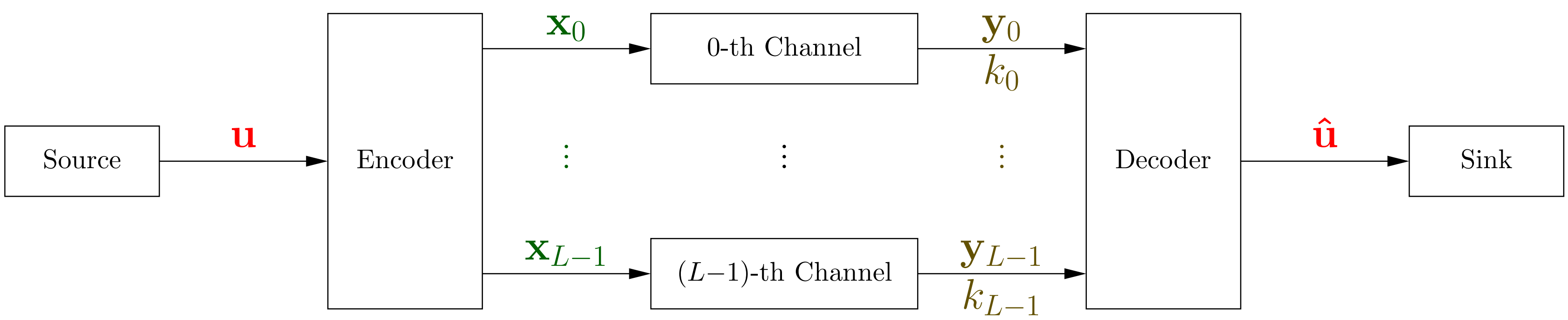
Comm. System with L Parallel Channels



E.g. $L = 4$, $n = 3$, $q = 3$.

$$\begin{pmatrix} 1 & 1 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 2 \\ 2 & 0 & 2 \end{pmatrix} \Rightarrow \begin{pmatrix} ? & ? & ? \\ \mathbf{2} & ? & ? \\ ? & ? & ? \\ \mathbf{2} & \mathbf{0} & ? \end{pmatrix} \begin{matrix} k_0 = 0 \\ k_1 = 1 \\ k_2 = 0 \\ k_3 = 2 \end{matrix} \Rightarrow \begin{pmatrix} \hat{u}_0 & \hat{u}_1 & \hat{u}_2 \end{pmatrix}$$

Comm. System with L Parallel Channels

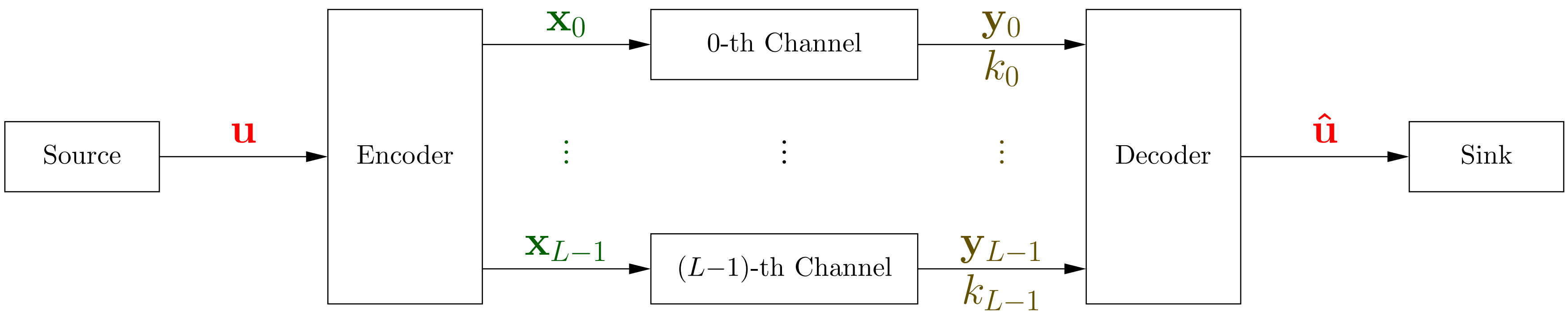


E.g. $L = 4$, $n = 3$, $q = 3$.

$$\begin{pmatrix} 1 & 1 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 2 \\ 2 & 0 & 2 \end{pmatrix} \Rightarrow \begin{pmatrix} ? & ? & ? \\ \mathbf{2} & ? & ? \\ ? & ? & ? \\ \mathbf{2} & \mathbf{0} & ? \end{pmatrix} \begin{matrix} k_0 = 0 \\ k_1 = 1 \\ k_2 = 0 \\ k_3 = 2 \end{matrix} \Rightarrow \begin{pmatrix} \hat{u}_0 & \hat{u}_1 & \hat{u}_2 \end{pmatrix}$$

We want unique decodability as long as $\sum_{\ell \in [L]} k_\ell \geq n$,

Comm. System with L Parallel Channels

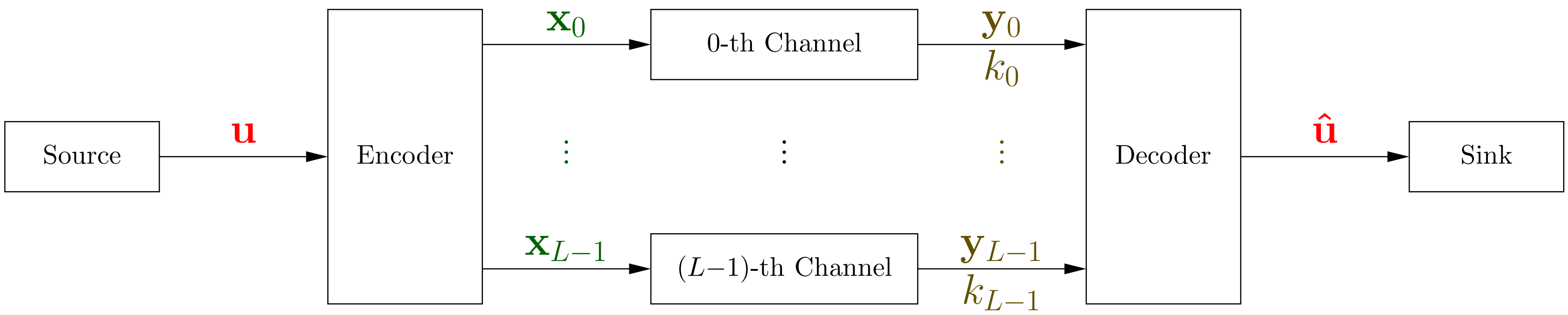


E.g. $L = 4$, $n = 3$, $q = 3$.

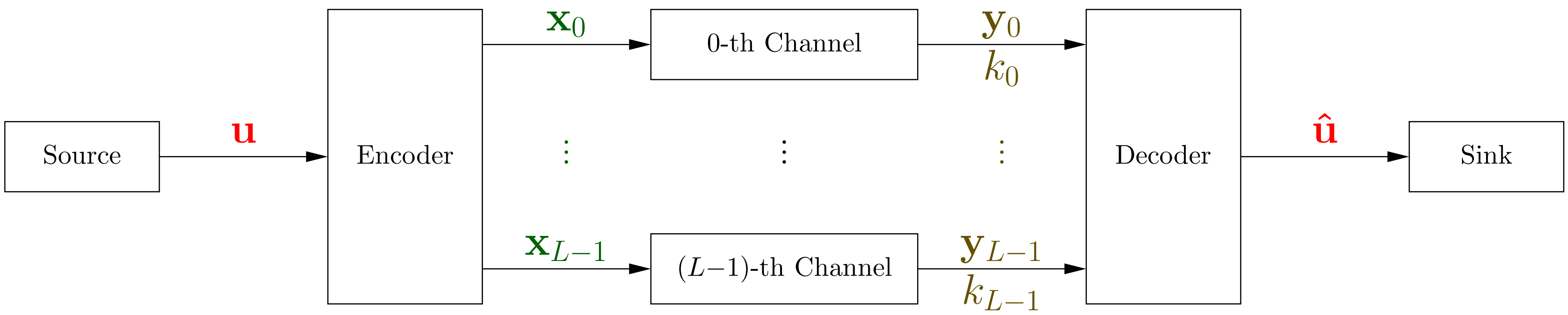
$$\begin{pmatrix} 1 & 1 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 2 \\ 2 & 0 & 2 \end{pmatrix} \Rightarrow \begin{pmatrix} ? & ? & ? \\ \mathbf{2} & ? & ? \\ ? & ? & ? \\ \mathbf{2} & \mathbf{0} & ? \end{pmatrix} \begin{matrix} k_0 = 0 \\ k_1 = 1 \\ k_2 = 0 \\ k_3 = 2 \end{matrix} \Rightarrow \begin{pmatrix} \hat{u}_0 & \hat{u}_1 & \hat{u}_2 \end{pmatrix}$$

We want unique decodability as long as $\sum_{\ell \in [L]} k_\ell \geq n$,
here: $k_0 + k_1 + k_2 + k_3 \geq 3$.

Comm. System with L Parallel Channels



Comm. System with L Parallel Channels

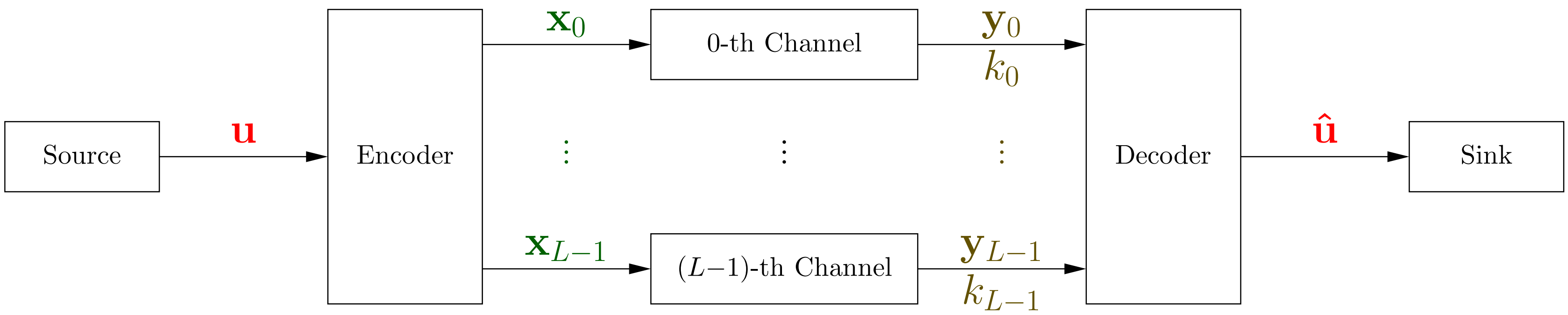


For reasons of simplicity, we would like the encoding to be linear:

$$\mathbf{x}_0 = \mathbf{u} \cdot \mathbf{G}_0, \quad \ldots, \quad \mathbf{x}_{L-1} = \mathbf{u} \cdot \mathbf{G}_{L-1},$$

where $\mathbf{G}_0, \ldots, \mathbf{G}_{L-1}$ are $n \times n$ matrices.

Comm. System with L Parallel Channels



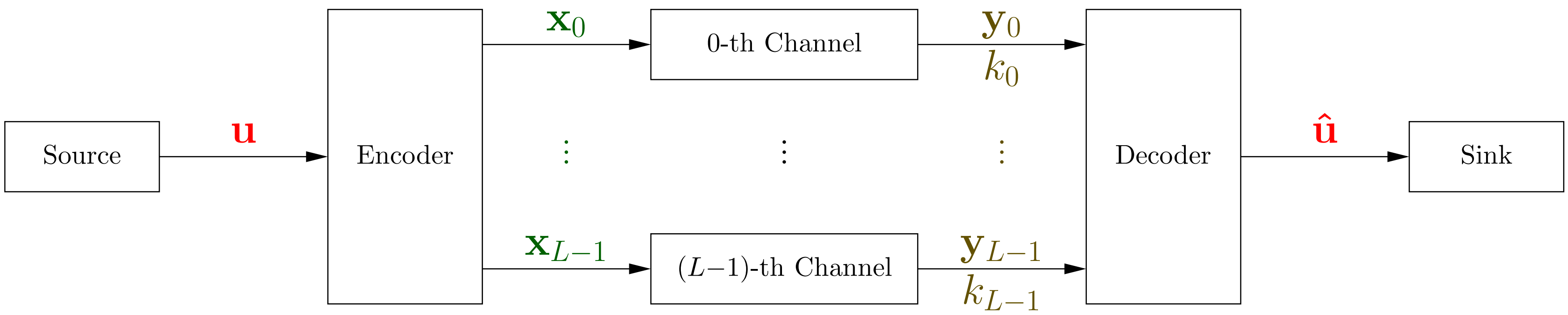
For reasons of simplicity, we would like the encoding to be linear:

$$\mathbf{x}_0 = \mathbf{u} \cdot \mathbf{G}_0, \quad \ldots, \quad \mathbf{x}_{L-1} = \mathbf{u} \cdot \mathbf{G}_{L-1},$$

where $\mathbf{G}_0, \ldots, \mathbf{G}_{L-1}$ are $n \times n$ matrices.

Definition: If the above matrices lead to unique decodability for any $k_0, \ldots, k_{L-1}$ with $\sum_{\ell \in [L]} k_\ell \geq n$, then we call these matrices **universally decodable matrices (UDMs)**.

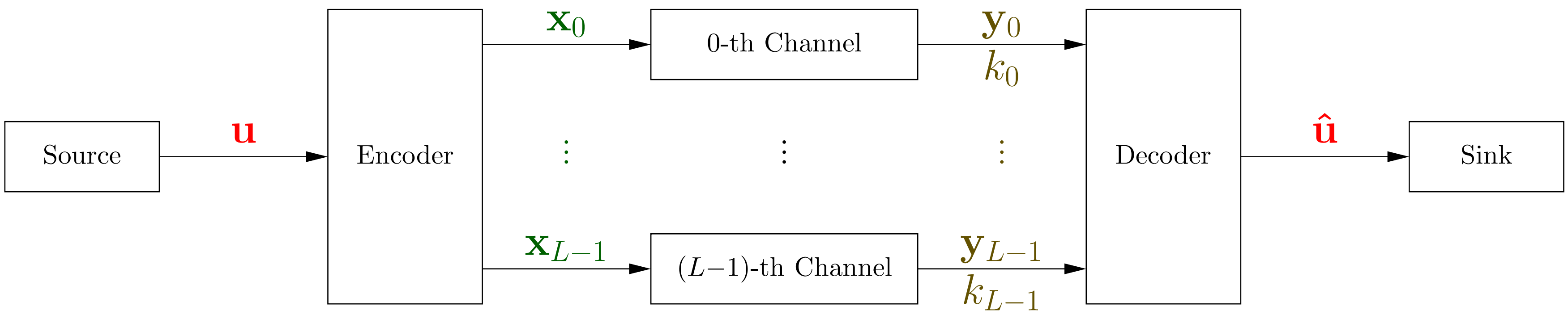
Comm. System with L Parallel Channels



E.g. $L = 2$, $n = 5$, any q. The matrices $\mathbf{G}_0$ and $\mathbf{G}_1$ are UDMs:

$$\mathbf{G}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Comm. System with L Parallel Channels

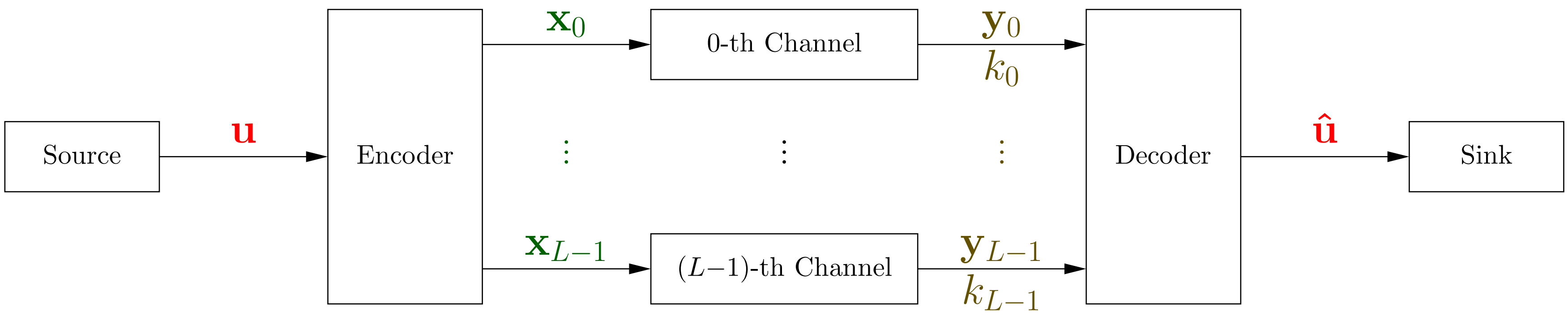


E.g. $L = 2$, $n = 5$, any q. The matrices $\mathbf{G}_0$ and $\mathbf{G}_1$ are UDMs:

$$\mathbf{G}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$\mathbf{u} = \begin{pmatrix} u_0 & u_1 & u_2 & u_3 & u_4 \end{pmatrix},$$

Comm. System with L Parallel Channels

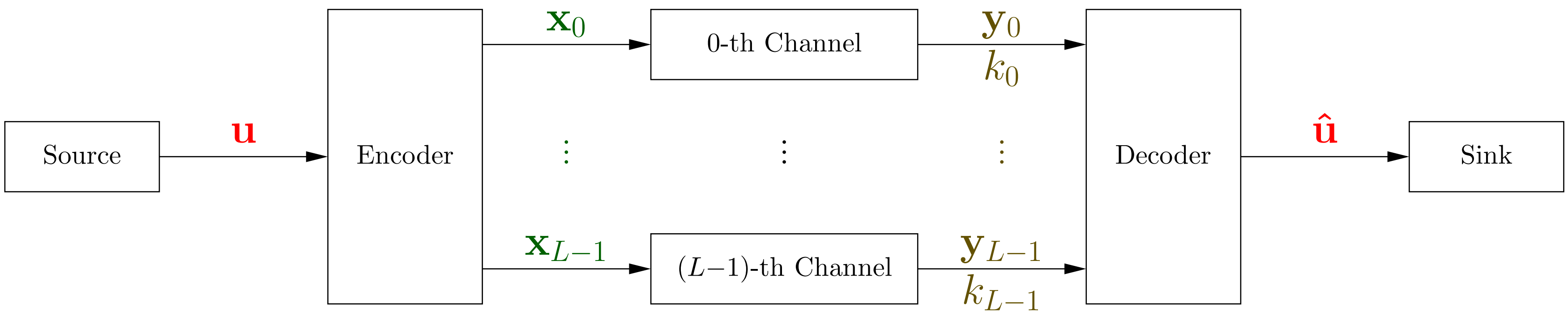


E.g. $L = 2$, $n = 5$, any q. The matrices $\mathbf{G}_0$ and $\mathbf{G}_1$ are UDMs:

$$\mathbf{G}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$\mathbf{u} = \begin{pmatrix} u_0 & u_1 & u_2 & u_3 & u_4 \end{pmatrix}, \qquad \mathbf{x}_0 = \begin{pmatrix} u_0 & u_1 & u_2 & u_3 & u_4 \end{pmatrix},$$

Comm. System with L Parallel Channels

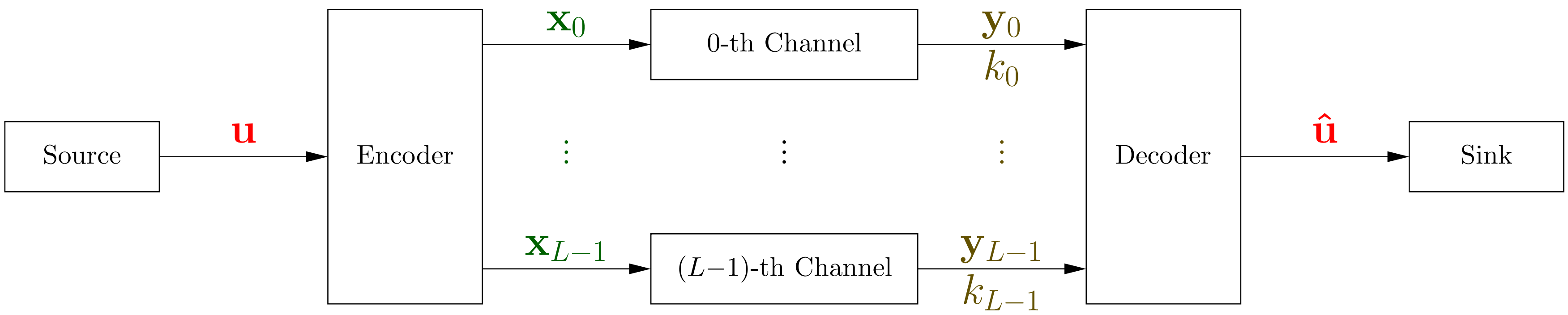


E.g. $L = 2$, $n = 5$, any q. The matrices $\mathbf{G}_0$ and $\mathbf{G}_1$ are UDMs:

$$\mathbf{G}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \ \mathbf{G}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$\mathbf{u} = \begin{pmatrix} u_0 & u_1 & u_2 & u_3 & u_4 \end{pmatrix}, \quad \mathbf{x}_0 = \begin{pmatrix} u_0 & u_1 & u_2 & u_3 & u_4 \end{pmatrix}, \ \mathbf{x}_1 = \begin{pmatrix} u_4 & u_3 & u_2 & u_1 & u_0 \end{pmatrix}.$$

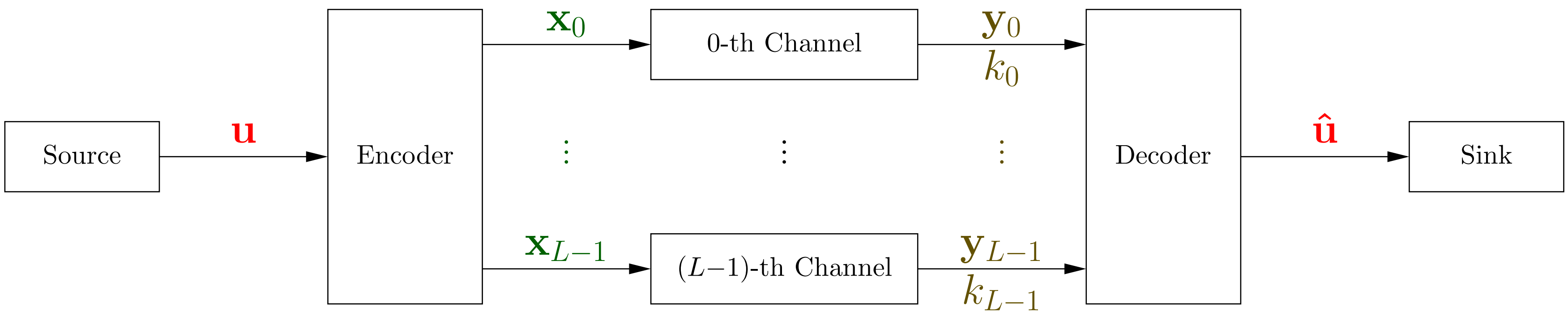
Comm. System with L Parallel Channels



E.g. $L = 2$, $n = 5$, any q. The matrices $\mathbf{G}_0$ and $\mathbf{G}_1$ are UDMs:

$$\mathbf{G}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Comm. System with L Parallel Channels

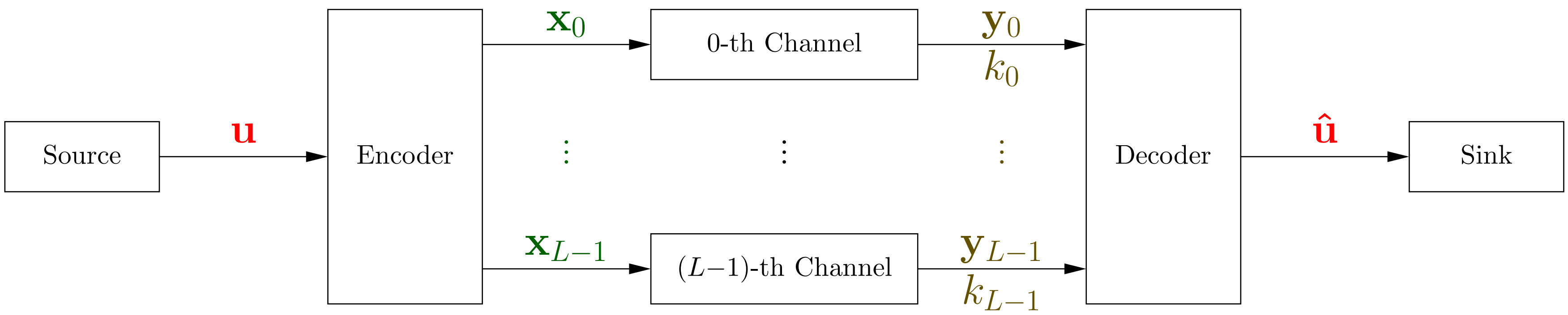


E.g. $L = 2$, $n = 5$, any q. The matrices $\mathbf{G}_0$ and $\mathbf{G}_1$ are UDMs:

$$\mathbf{G}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$\mathbf{u} = \begin{pmatrix} u_0 & u_1 & u_2 & u_3 & u_4 \end{pmatrix},$$

Comm. System with L Parallel Channels

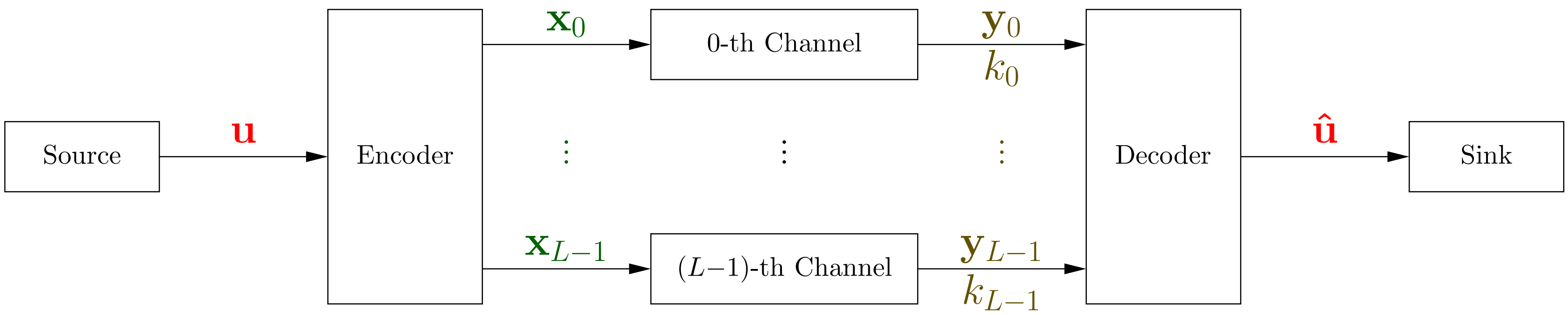


E.g. $L = 2$, $n = 5$, any q. The matrices $\mathbf{G}_0$ and $\mathbf{G}_1$ are UDMs:

$$\mathbf{G}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$\mathbf{u} = \begin{pmatrix} u_0 & u_1 & u_2 & u_3 & u_4 \end{pmatrix}, \quad \mathbf{y}_0 = \begin{pmatrix} u_0 & u_1 & u_2 & ? & ? \end{pmatrix},$$

Comm. System with L Parallel Channels

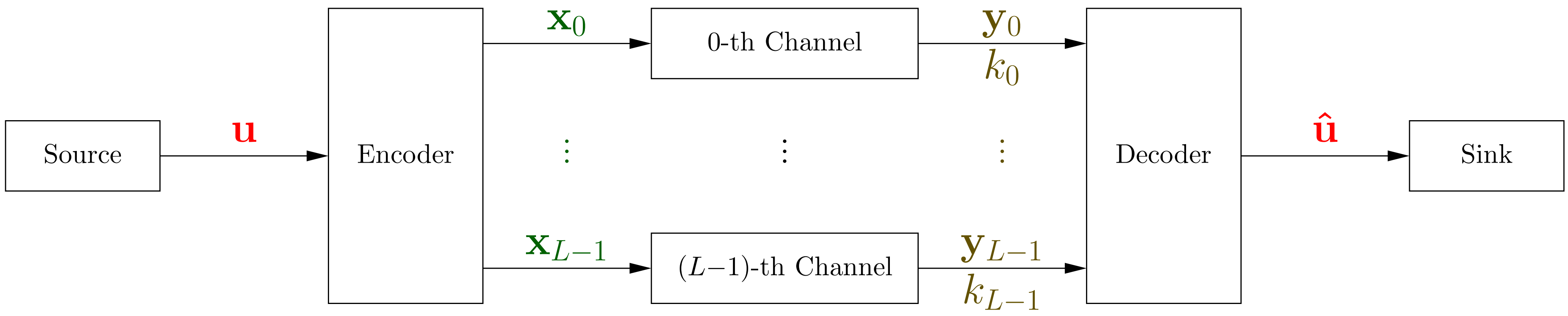


E.g. $L = 2$, $n = 5$, any q. The matrices $\mathbf{G}_0$ and $\mathbf{G}_1$ are UDMs:

$$\mathbf{G}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$\mathbf{u} = \begin{pmatrix} u_0 & u_1 & u_2 & u_3 & u_4 \end{pmatrix}, \quad \mathbf{y}_0 = \begin{pmatrix} u_0 & u_1 & u_2 & ? & ? \end{pmatrix}, \quad \mathbf{x}_1 = \begin{pmatrix} u_4 & u_3 & ? & ? & ? \end{pmatrix}.$$

Comm. System with L Parallel Channels



E.g. $L = 4$, $n = 3$, $q = 3$. The matrices $\mathbf{G}_0$, $\mathbf{G}_1$, $\mathbf{G}_2$, $\mathbf{G}_3$ are UDMs:

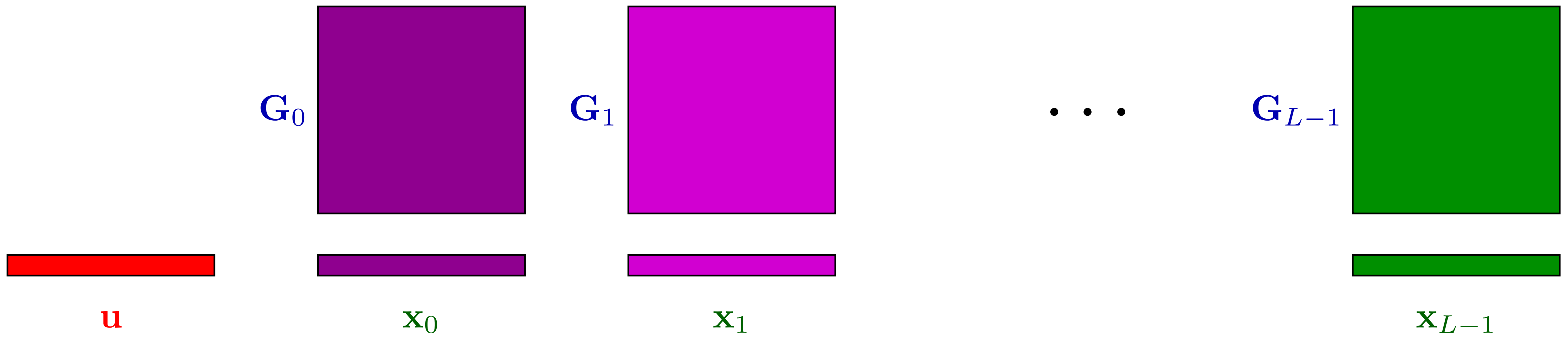
$$\mathbf{G}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{G}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \quad \mathbf{G}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

Comm. System with L Parallel Channels

What does unique decodability imply for the matrices $\mathbf{G}_0, \ldots, \mathbf{G}_{L-1}$?

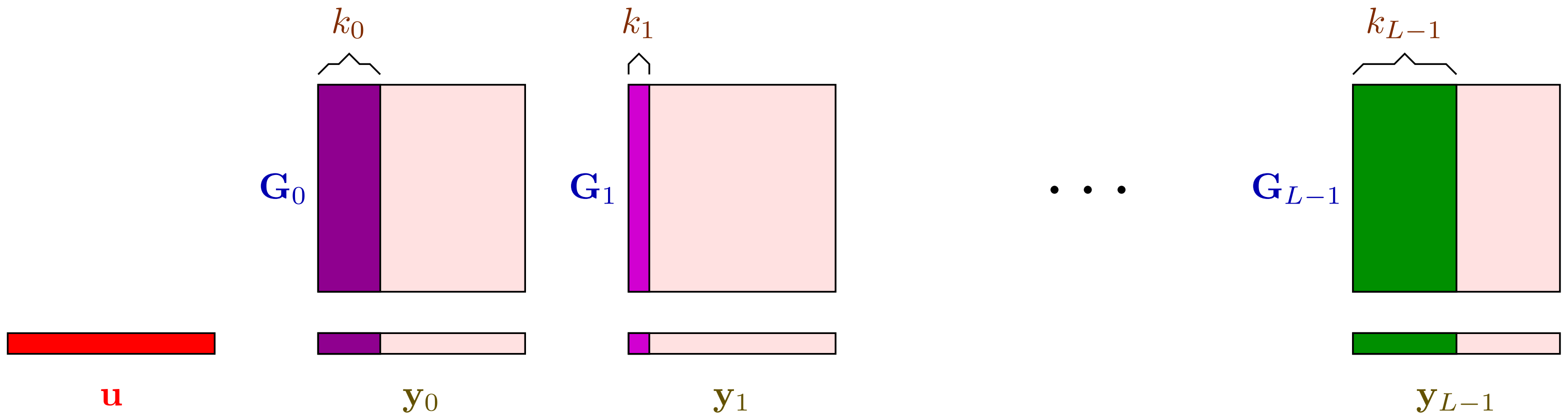
Comm. System with L Parallel Channels

What does unique decodability imply for the matrices $\mathbf{G}_0, \ldots, \mathbf{G}_{L-1}$?



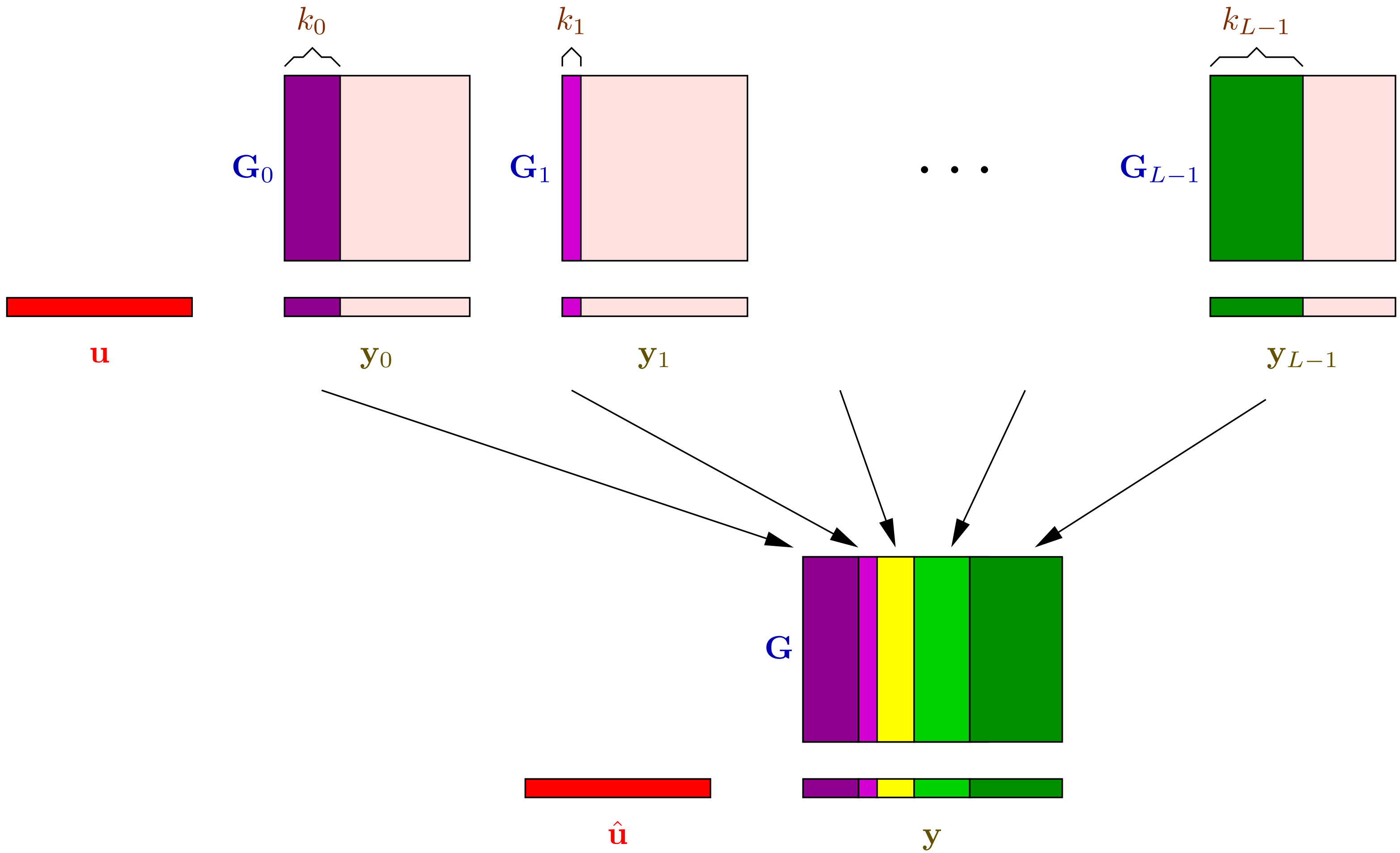
Comm. System with L Parallel Channels

What does unique decodability imply for the matrices $\mathbf{G}_0, \ldots, \mathbf{G}_{L-1}$?



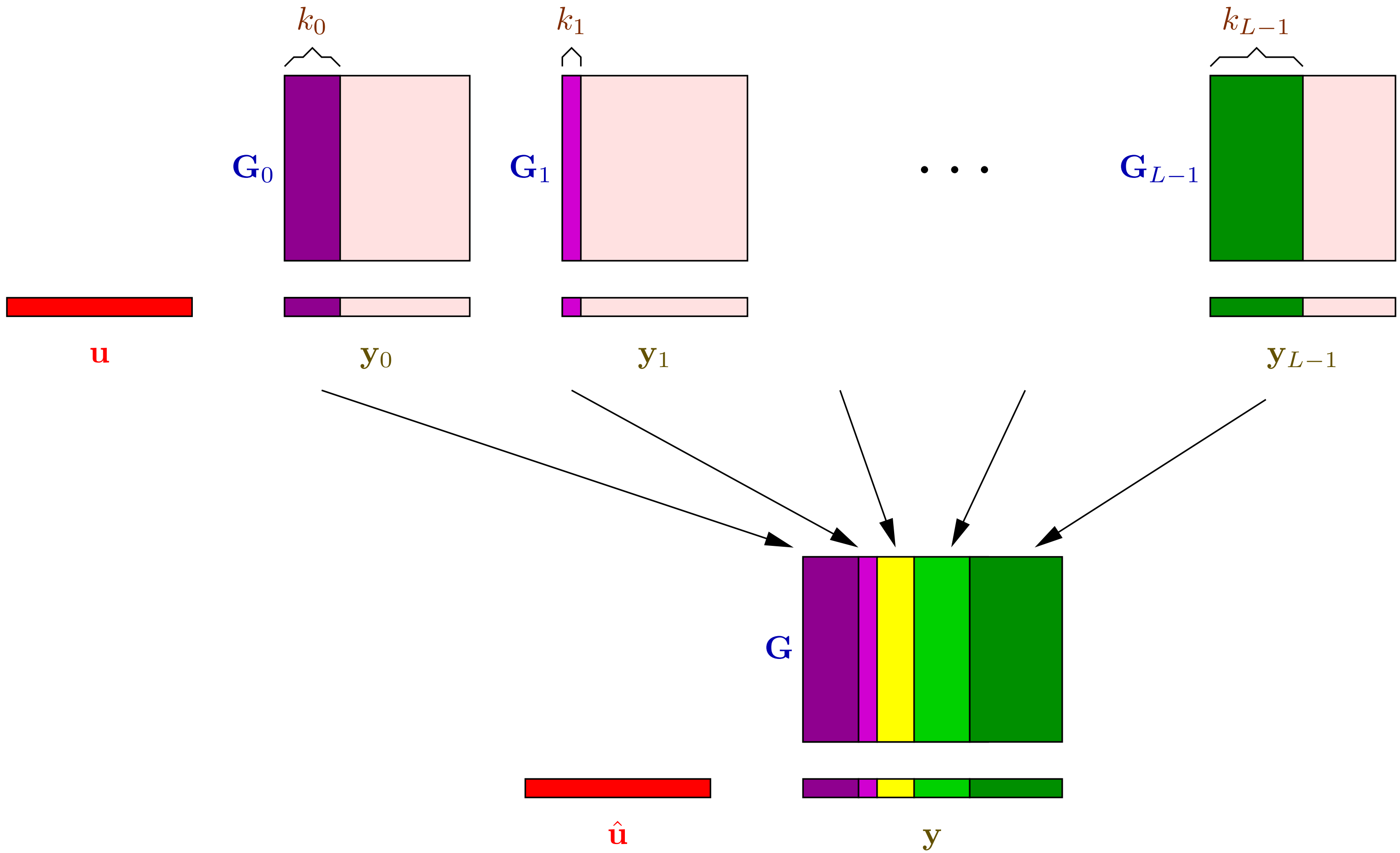
Comm. System with L Parallel Channels

What does unique decodability imply for the matrices $\mathbf{G}_0, \ldots, \mathbf{G}_{L-1}$?



Comm. System with L Parallel Channels

What does unique decodability imply for the matrices $\mathbf{G}_0, \ldots, \mathbf{G}_{L-1}$?



For any $k_0, \ldots, k_{L-1}$ with $\sum_{\ell \in [L]} k_\ell \geq n$ the matrix $\mathbf{G}$ must have full rank.

Comm. System with L Parallel Channels

- Another motivation for this channel model: paper by Tavildar and Viswanath, “Approximately universal codes over slow fading channels”, IEEE Trans. Inf. Theory, IT–52, no. 7, pp. 3233–3258, July 2006.

Comm. System with L Parallel Channels

- Another motivation for this channel model: paper by Tavildar and Viswanath, "Approximately universal codes over slow fading channels", IEEE Trans. Inf. Theory, IT–52, no. 7, pp. 3233–3258, July 2006.

- Consider slow-fading (point-to-point) MIMO channel

$$\mathbf{y}[m] = \mathbf{H} \cdot \mathbf{x}[m] + \mathbf{w}[m].$$

The complex matrix of fading gains $\mathbf{H}$ stays constant over the time-scale of communication; we suppose the exact characterization of $\mathbf{H}$ is known to the receiver while the transmitter has only access to its statistical characterization.

Comm. System with L Parallel Channels

- Another motivation for this channel model: paper by Tavildar and Viswanath, “Approximately universal codes over slow fading channels”, IEEE Trans. Inf. Theory, IT–52, no. 7, pp. 3233–3258, July 2006.

- Consider slow-fading (point-to-point) MIMO channel

$$\mathbf{y}[m] = \mathbf{H} \cdot \mathbf{x}[m] + \mathbf{w}[m].$$

  The complex matrix of fading gains $\mathbf{H}$ stays constant over the time-scale of communication; we suppose the exact characterization of $\mathbf{H}$ is known to the receiver while the transmitter has only access to its statistical characterization.

- The focus in the paper is on the high-SNR regime.

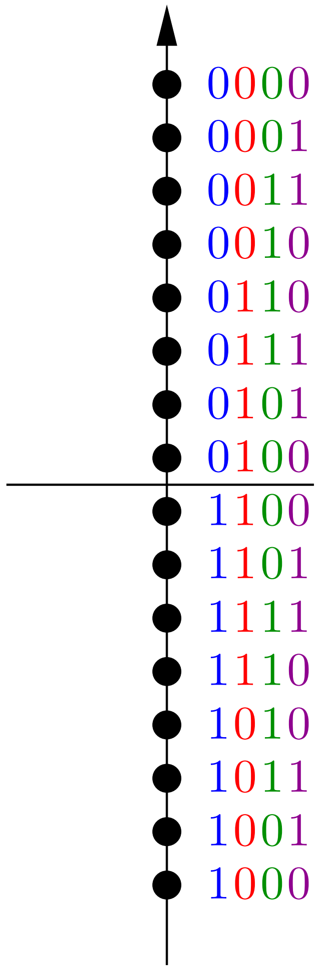
Comm. System with L Parallel Channels

- Another motivation for this channel model: paper by Tavildar and Viswanath, "Approximately universal codes over slow fading channels", IEEE Trans. Inf. Theory, IT–52, no. 7, pp. 3233–3258, July 2006.
- Consider slow-fading (point-to-point) MIMO channel

$$\mathbf{y}[m] = \mathbf{H} \cdot \mathbf{x}[m] + \mathbf{w}[m].$$

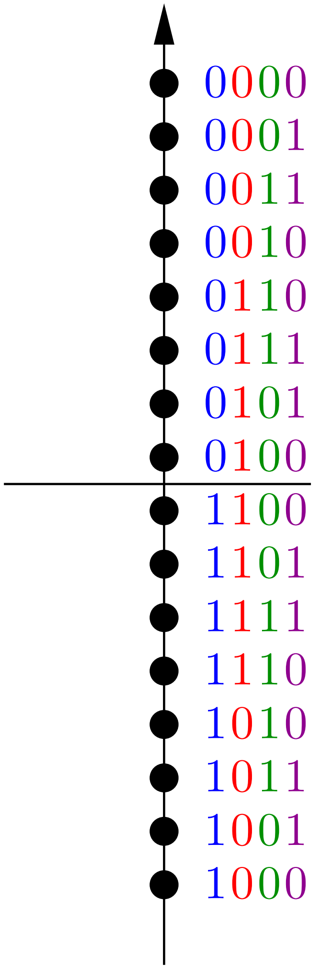
  The complex matrix of fading gains $\mathbf{H}$ stays constant over the time-scale of communication; we suppose the exact characterization of $\mathbf{H}$ is known to the receiver while the transmitter has only access to its statistical characterization.
- The focus in the paper is on the high-SNR regime.
- Coding for this channel can be seen as space-time coding.

Comm. System with L Parallel Channels



- Depending on what h_ℓ is, we can recover more or fewer of the most-significant bits.

Comm. System with L Parallel Channels

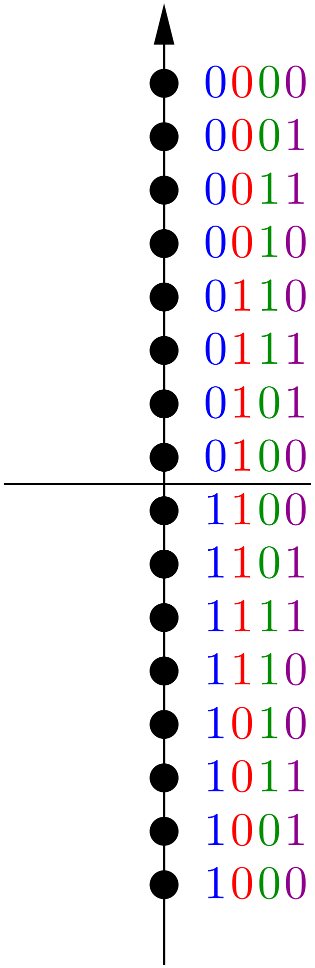


- Depending on what h_ℓ is, we can recover more or fewer of the most-significant bits.

- Assume $L = 2$: channel is not in outage if

$$\log\left(1 + |h_0|^2 \mathsf{SNR}\right) + \log\left(1 + |h_1|^2 \mathsf{SNR}\right) > 2R.$$

Comm. System with L Parallel Channels



- Depending on what h_ℓ is, we can recover more or fewer of the most-significant bits.
- Assume $L = 2$: channel is not in outage if

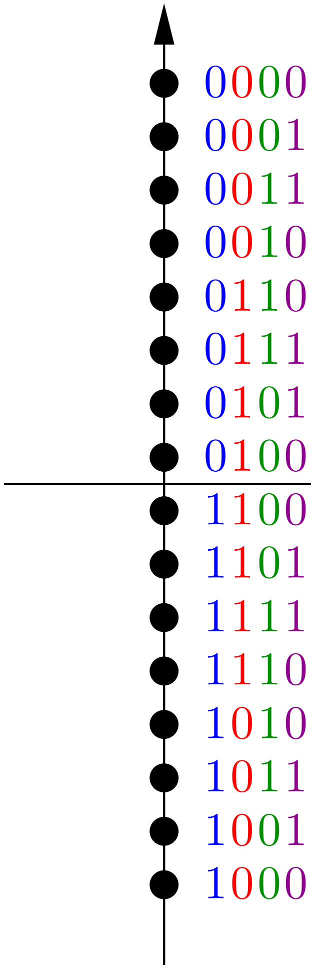
$$\log\left(1 + |h_0|^2 \mathsf{SNR}\right) + \log\left(1 + |h_1|^2 \mathsf{SNR}\right) > 2R.$$

- Assume that h_0 and h_1 are such that

$$\log\left(1 + |h_0|^2 \mathsf{SNR}\right) > 2k_0, \quad \log\left(1 + |h_1|^2 \mathsf{SNR}\right) > 2k_1.$$

for some k_0 and k_1, i.e. we can recover k_0 bits from the zeroth channel and k_1 bits from the first channel.

Comm. System with L Parallel Channels



- Depending on what h_ℓ is, we can recover more or fewer of the most-significant bits.

- Assume $L = 2$: channel is not in outage if

$$\log\left(1 + |h_0|^2\mathsf{SNR}\right) + \log\left(1 + |h_1|^2\mathsf{SNR}\right) > 2R.$$

- Assume that h_0 and h_1 are such that

$$\log\left(1 + |h_0|^2\mathsf{SNR}\right) > 2k_0, \quad \log\left(1 + |h_1|^2\mathsf{SNR}\right) > 2k_1.$$

  for some k_0 and k_1, i.e. we can recover k_0 bits from the zeroth channel and k_1 bits from the first channel.

- Not being in outage means that $k_0 + k_1 \geq R$.

Coding via Evaluation

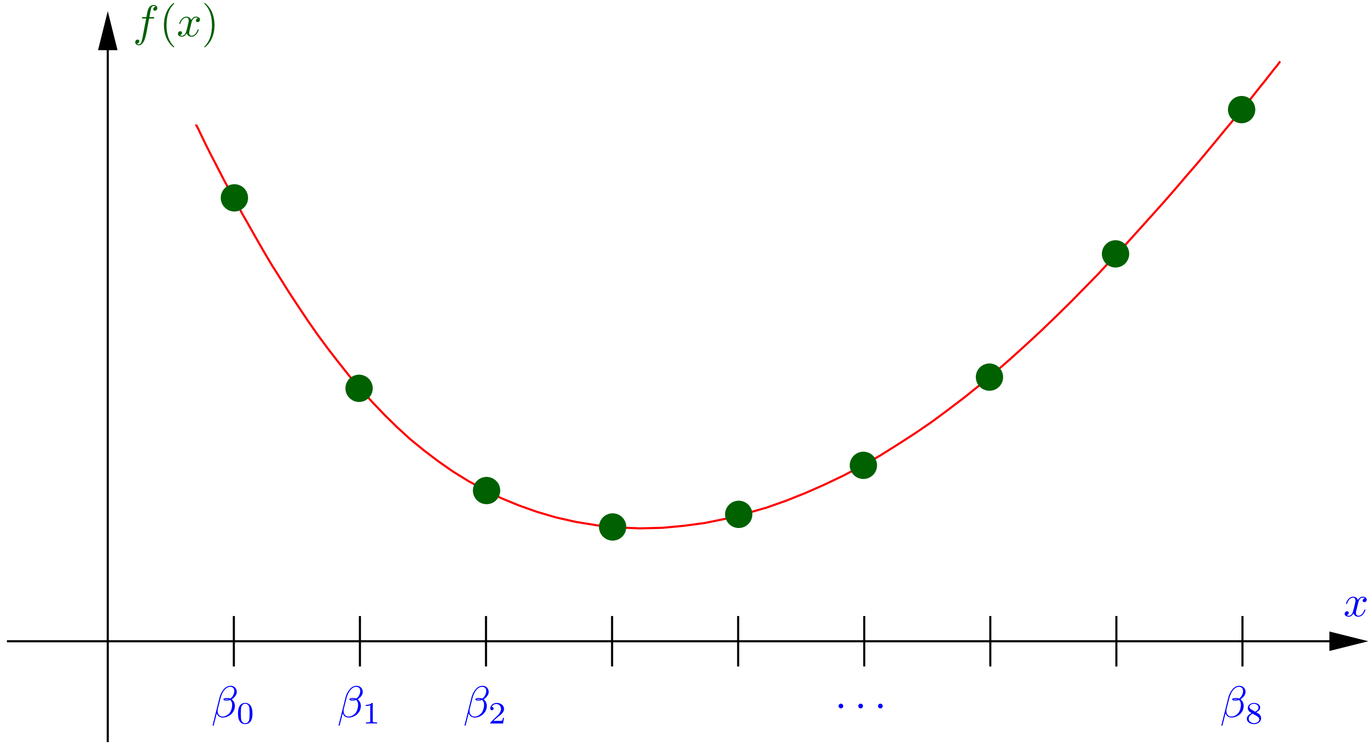
Coding via Evaluation (First Setup)

Encoding map (evaluation map):

$$(u_0, u_1, u_2) \mapsto \big(f(\beta_0), f(\beta_1), f(\beta_2), f(\beta_3), f(\beta_4), f(\beta_5), f(\beta_6), f(\beta_7), f(\beta_8)\big),$$

where $f(x) = u_0 x^0 + u_1 x^1 + u_2 x^2$.

Coding via Evaluation (First Setup)

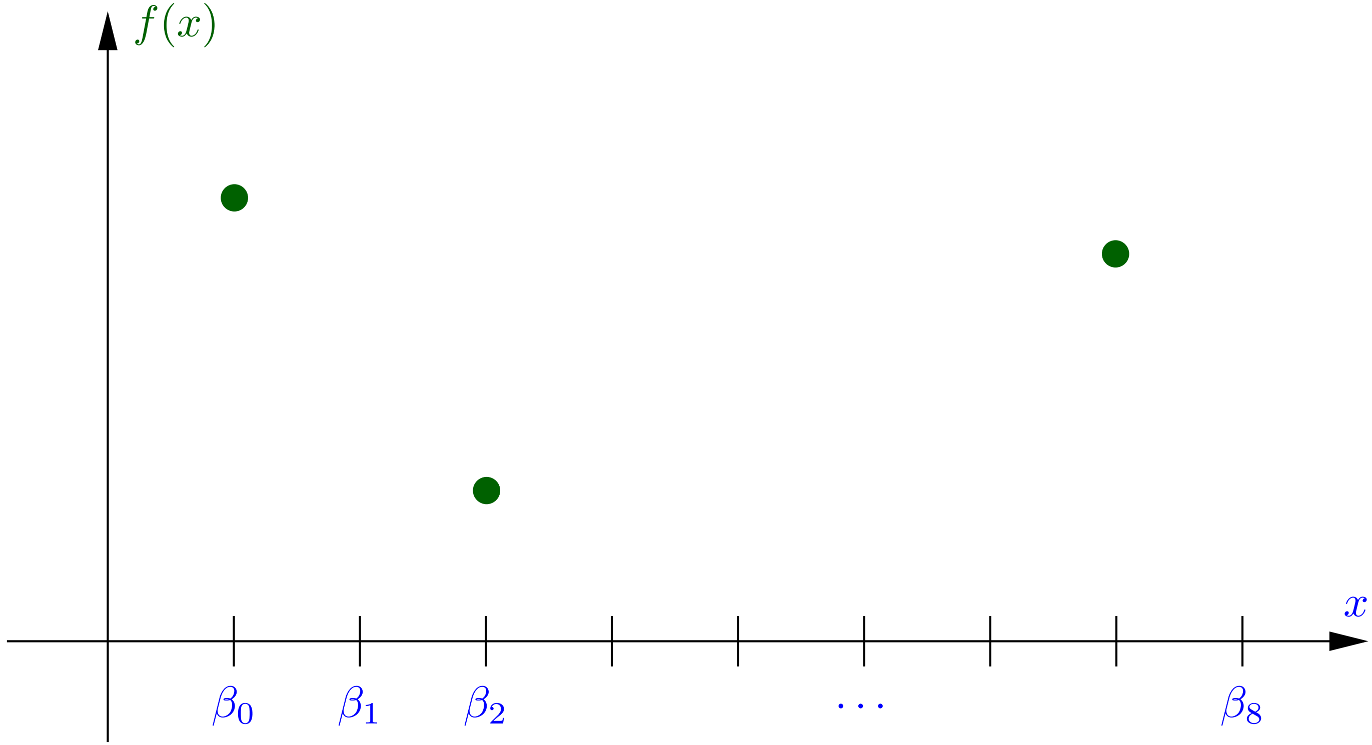


Encoding map (evaluation map):

$$(u_0, u_1, u_2) \mapsto \big(f(\beta_0), f(\beta_1), f(\beta_2), f(\beta_3), f(\beta_4), f(\beta_5), f(\beta_6), f(\beta_7), f(\beta_8)\big),$$

where $f(x) = u_0 x^0 + u_1 x^1 + u_2 x^2$.

Coding via Evaluation (First Setup)

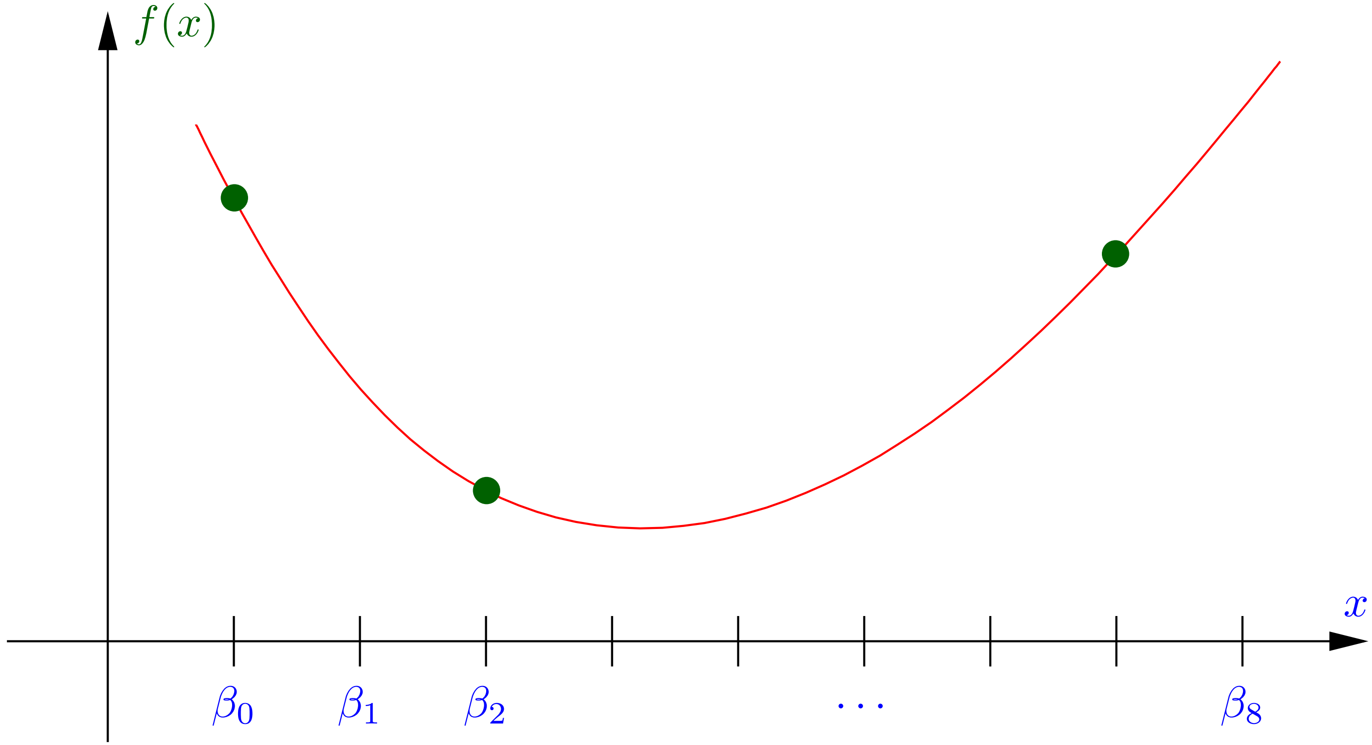


Encoding map (evaluation map):

$$(u_0, u_1, u_2) \mapsto \big(f(\beta_0), f(\beta_1), f(\beta_2), f(\beta_3), f(\beta_4), f(\beta_5), f(\beta_6), f(\beta_7), f(\beta_8)\big),$$

where $f(x) = u_0x^0 + u_1x^1 + u_2x^2$.

Coding via Evaluation (First Setup)

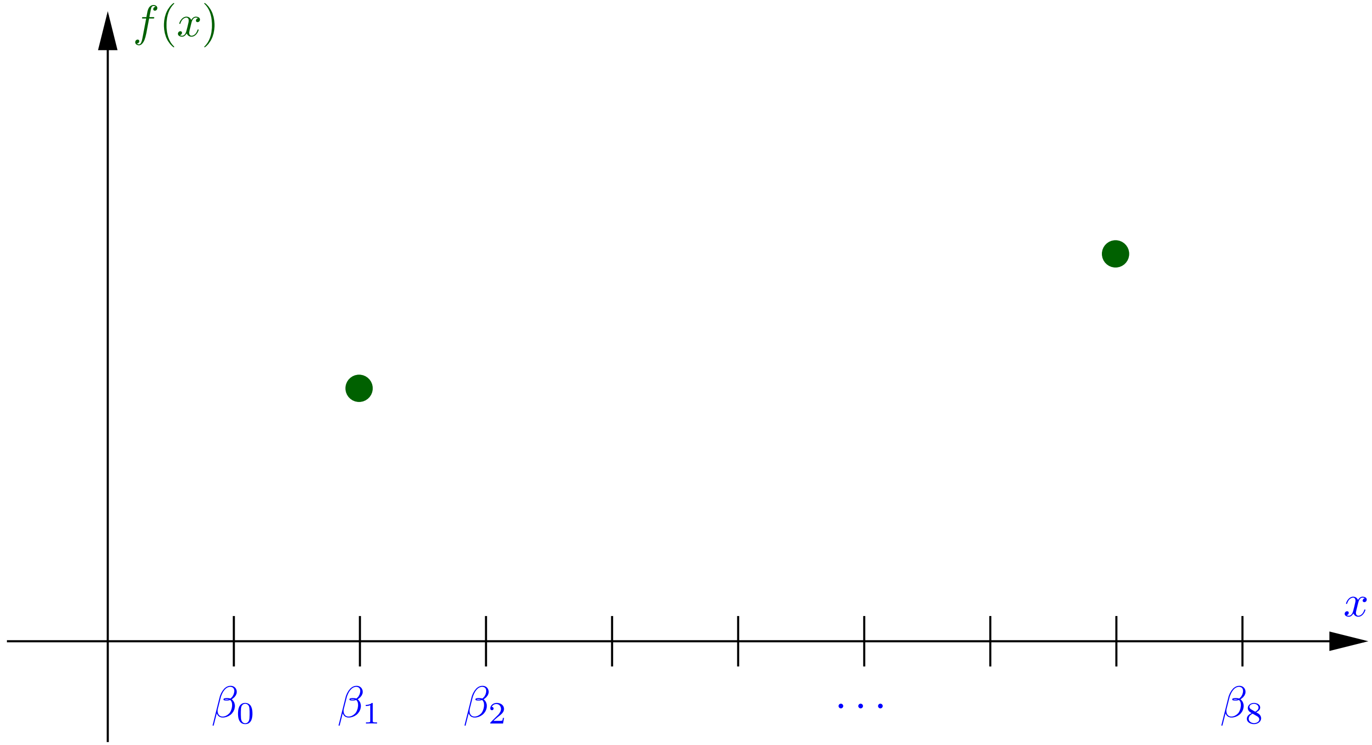


Encoding map (evaluation map):

$$(u_0, u_1, u_2) \mapsto \big(f(\beta_0), f(\beta_1), f(\beta_2), f(\beta_3), f(\beta_4), f(\beta_5), f(\beta_6), f(\beta_7), f(\beta_8)\big),$$

where $f(x) = u_0x^0 + u_1x^1 + u_2x^2$.

Coding via Evaluation (First Setup)

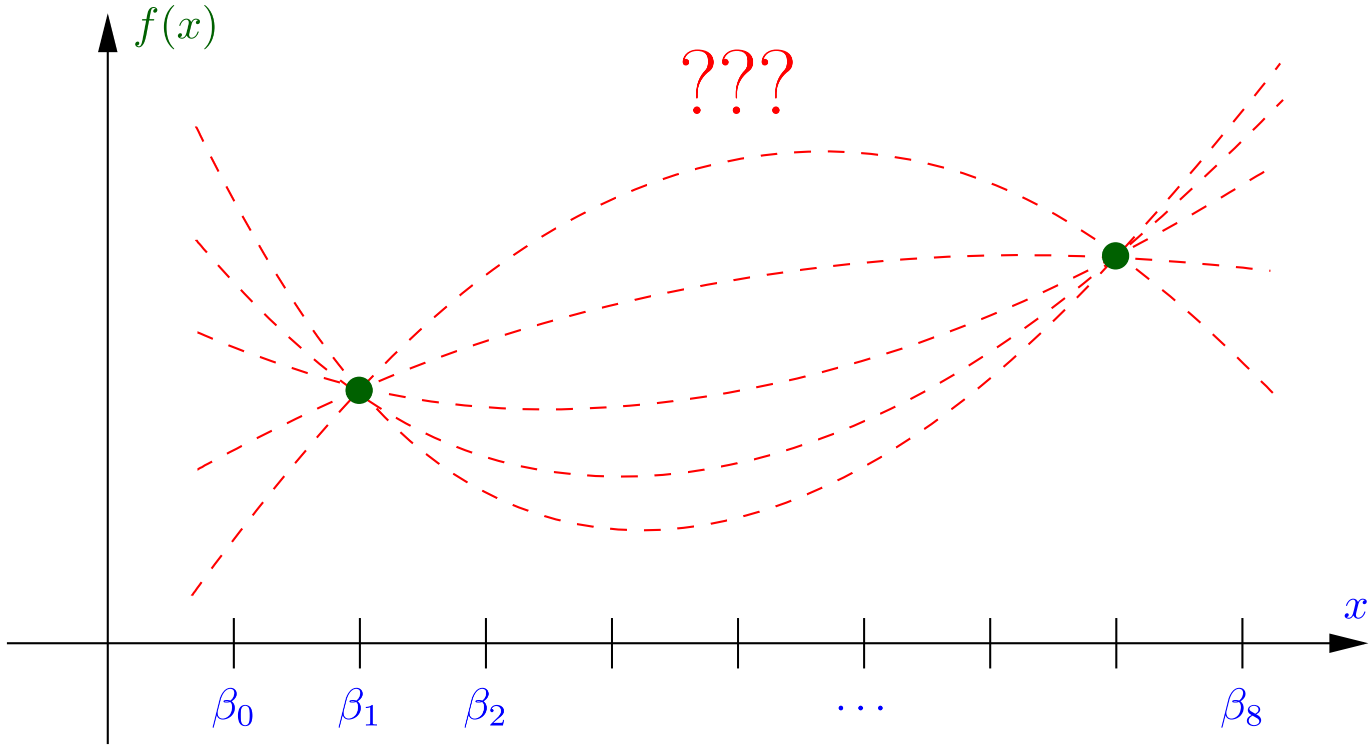


Encoding map (evaluation map):

$$(u_0, u_1, u_2) \mapsto \big(f(\beta_0), f(\beta_1), f(\beta_2), f(\beta_3), f(\beta_4), f(\beta_5), f(\beta_6), f(\beta_7), f(\beta_8)\big),$$

where $f(x) = u_0 x^0 + u_1 x^1 + u_2 x^2$.

Coding via Evaluation (First Setup)



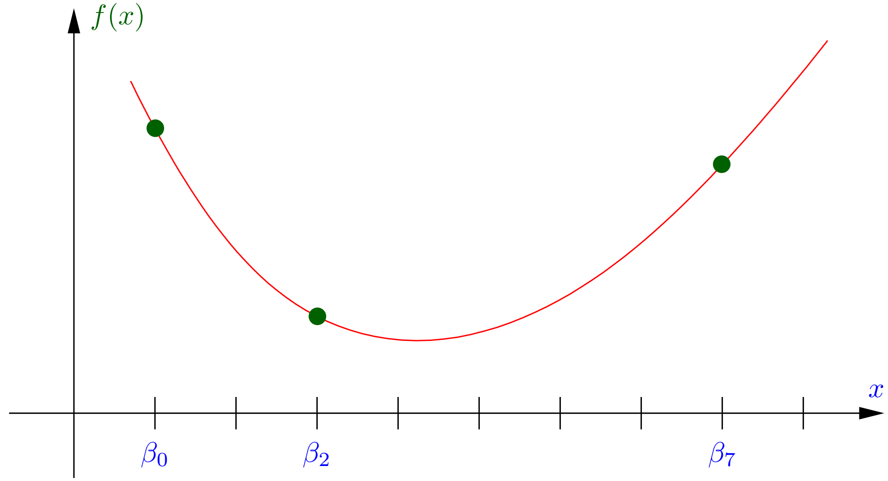
Encoding map (evaluation map):

$$(u_0, u_1, u_2) \mapsto \big(f(\beta_0), f(\beta_1), f(\beta_2), f(\beta_3), f(\beta_4), f(\beta_5), f(\beta_6), f(\beta_7), f(\beta_8)\big),$$

where $f(x) = u_0 x^0 + u_1 x^1 + u_2 x^2$.

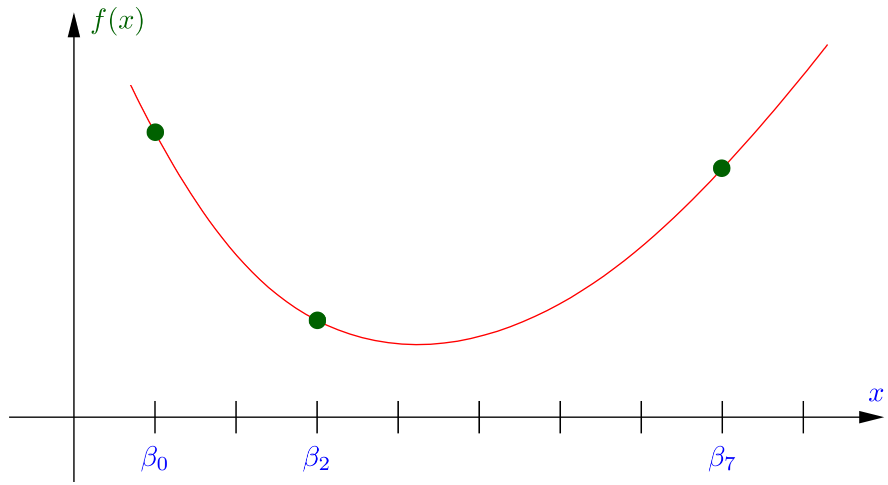
Coding via Evaluation (First Setup)

Assume that we only receive the function values for $x = \beta_0, \beta_2, \beta_7$.



Coding via Evaluation (First Setup)

Assume that we only receive the function values for $x = \beta_0, \beta_2, \beta_7$.



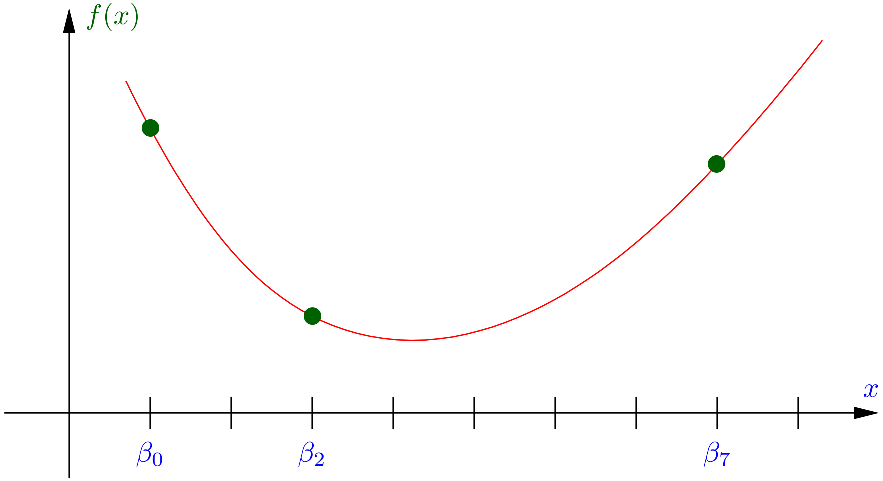
- We have to show that the mapping

$$(u_0, u_1, u_2) \mapsto (f(\beta_0), f(\beta_2), f(\beta_7))$$

is injective.

Coding via Evaluation (First Setup)

Assume that we only receive the function values for $x = \beta_0, \beta_2, \beta_7$.



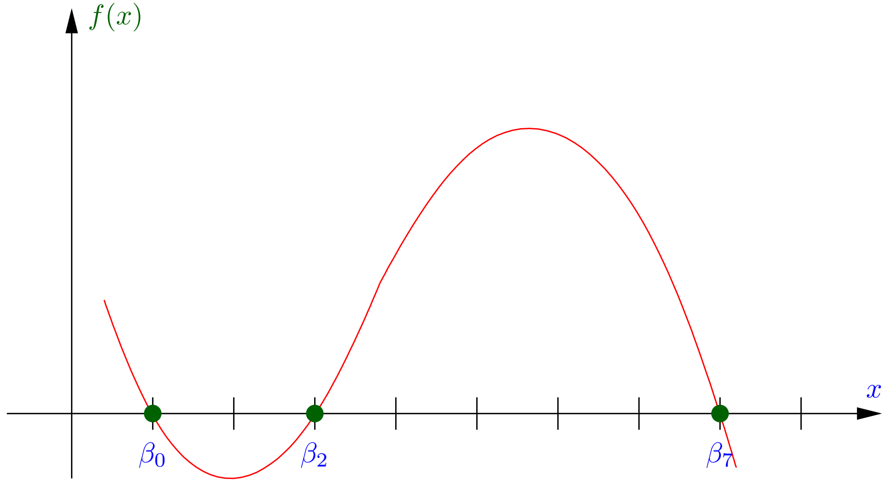
- We have to show that the mapping

$$\left(u_0, u_1, u_2\right) \mapsto \left(f(\beta_0), f(\beta_2), f(\beta_7)\right)$$

  is injective.

- Because the above mapping is linear it is sufficient to show that the kernel is trivial.

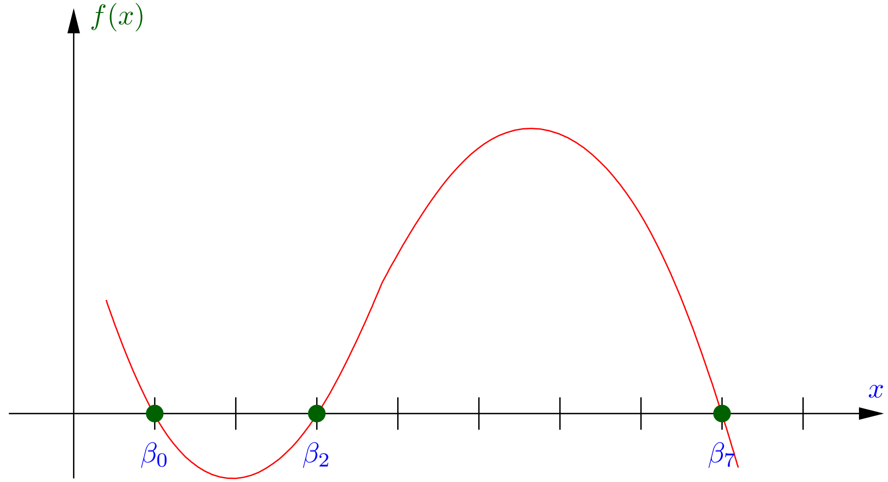
Coding via Evaluation (First Setup)



Case 2:

$f(x) \neq 0$ with at least three zeros.

Coding via Evaluation (First Setup)

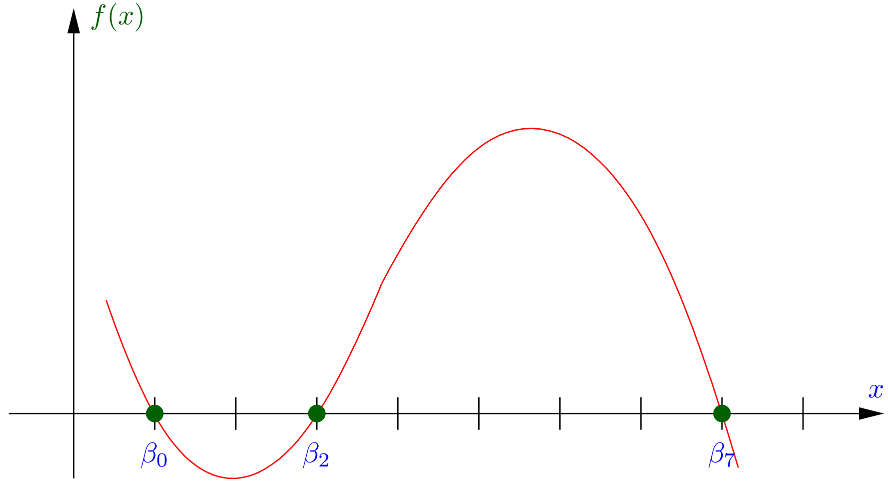


Case 2:

$f(x) \neq 0$ with at least three zeros.

The fundamental theorem of algebra implies that $\deg(f(x)) \geq 3$.

Coding via Evaluation (First Setup)

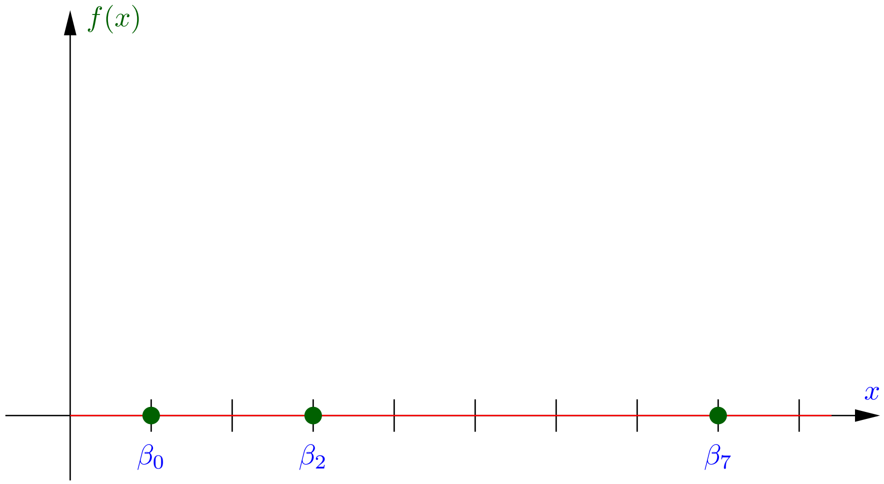


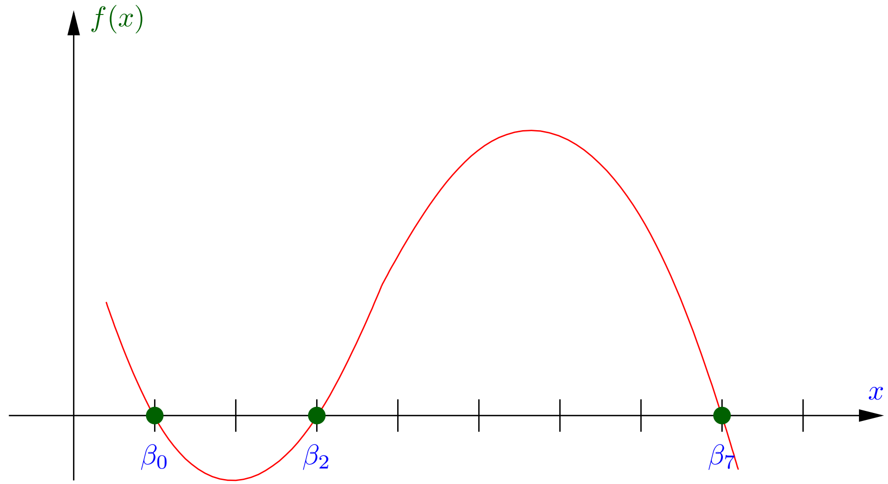
Case 2:

$f(x) \neq 0$ with at least three zeros.

The fundamental theorem of algebra implies that $\deg(f(x)) \geq 3$. However, no quadratic function can have more than two zeros.

Coding via Evaluation (First Setup)





Case 1:

$$f(x) = 0,$$

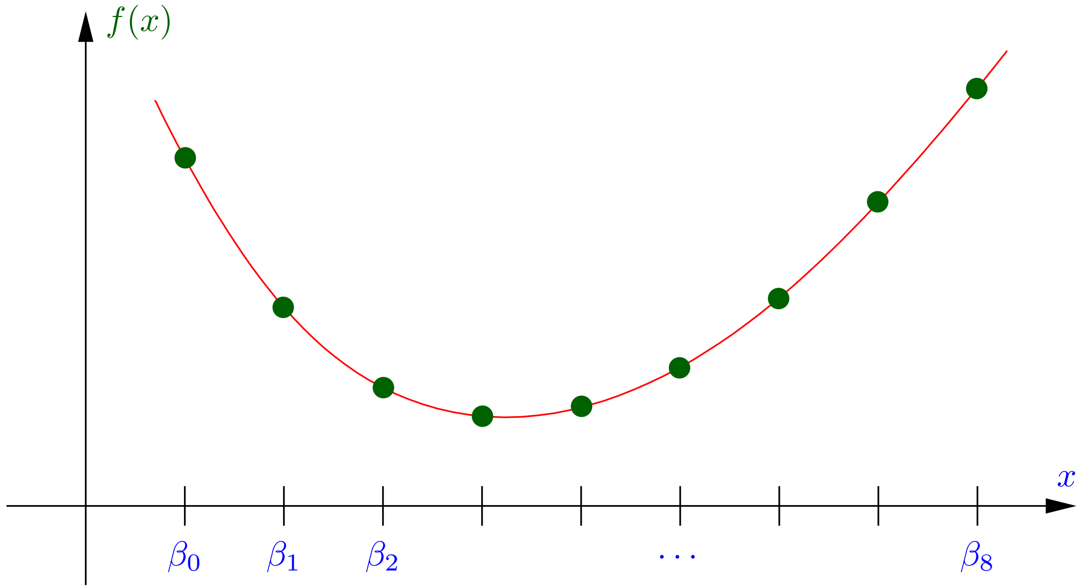
$$\Rightarrow (u_0, u_1, u_2) = (0, 0, 0).$$

Case 2:

$f(x) \neq 0$ with at least three zeros.

The fundamental theorem of algebra implies that $\deg(f(x)) \geq 3$. However, no quadratic function can have more than two zeros.

Coding via Evaluation (First Setup)

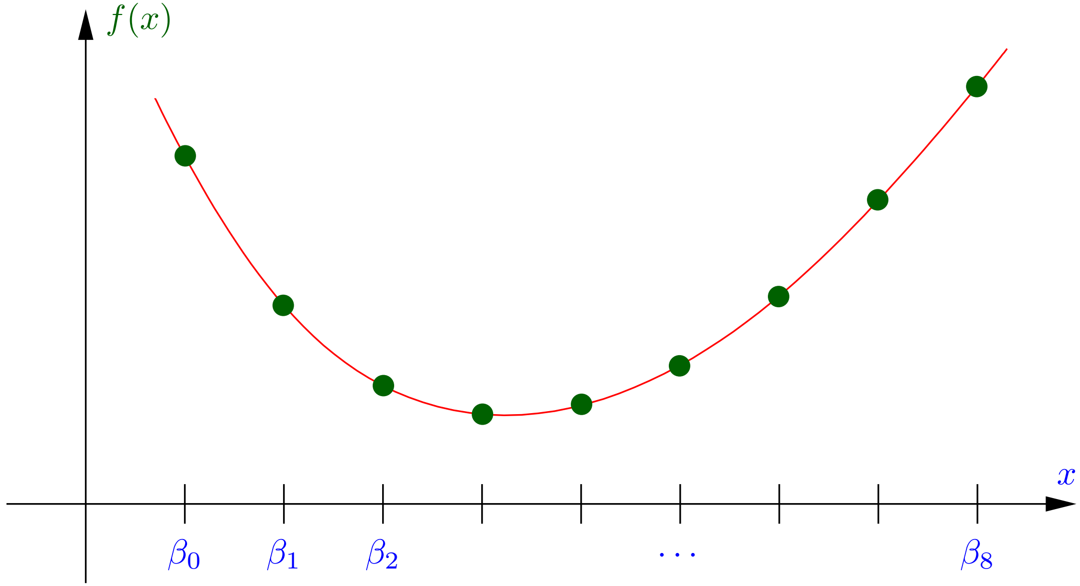


Encoding map (evaluation map):

$$(u_0, u_1, u_2) \mapsto \big(f(\beta_0), f(\beta_1), f(\beta_2), f(\beta_3), f(\beta_4), f(\beta_5), f(\beta_6), f(\beta_7), f(\beta_8)\big),$$

where $f(x) = u_0 x^0 + u_1 x^1 + u_2 x^2$.

Coding via Evaluation (First Setup)



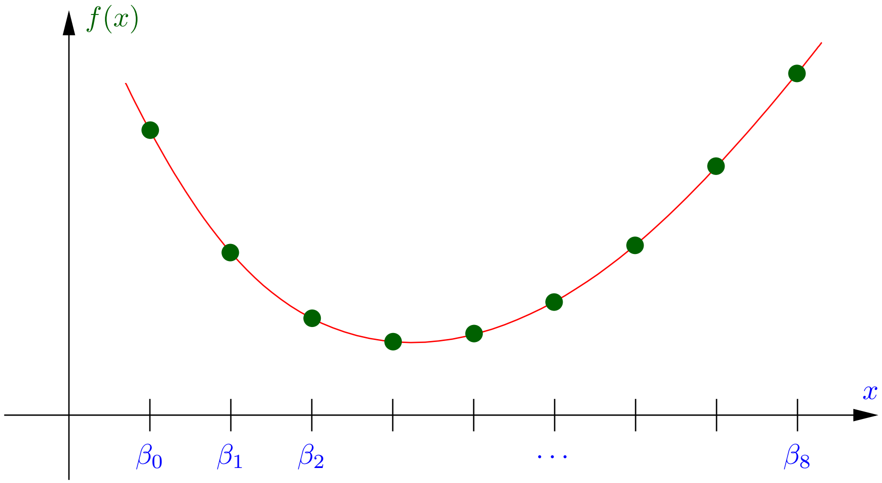
Encoding map (evaluation map):

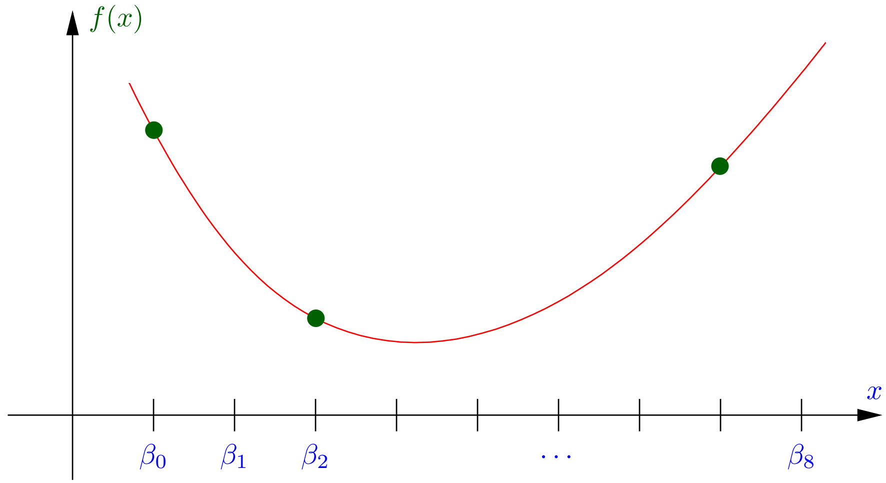
$$(u_0, u_1, u_2) \mapsto \big(f(\beta_0), f(\beta_1), f(\beta_2), f(\beta_3), f(\beta_4), f(\beta_5), f(\beta_6), f(\beta_7), f(\beta_8)\big),$$

where $f(x) = u_0 x^0 + u_1 x^1 + u_2 x^2$.

Note: the codes that result from this evaluation map are the well-known Reed-Solomon codes.

Coding via Evaluation (First Setup)



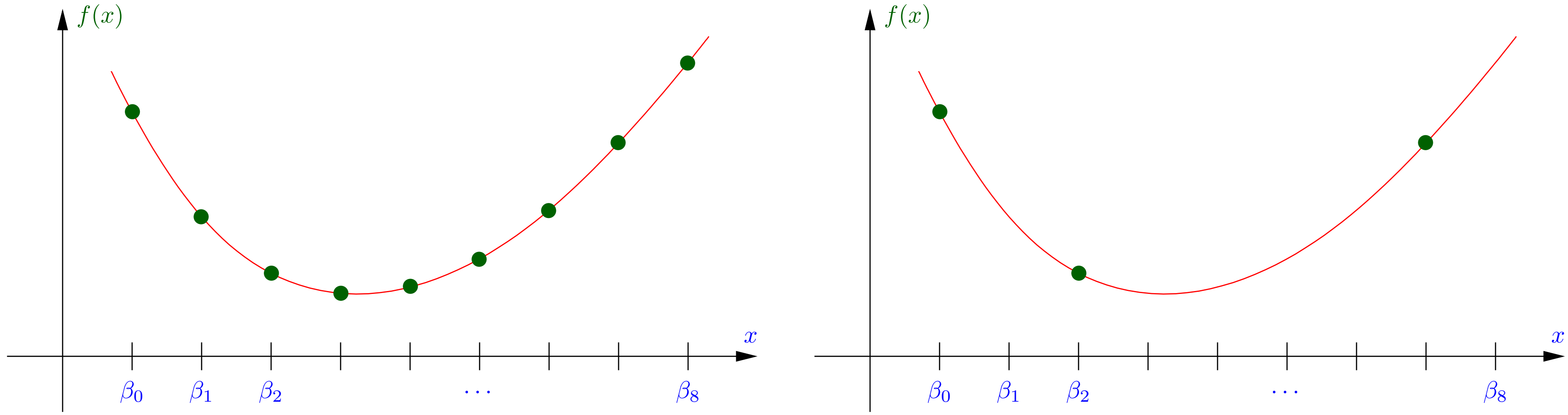


Encoding map (evaluation map):

$$(u_0, u_1, u_2) \mapsto \big(f(\beta_0), f(\beta_1), f(\beta_2), f(\beta_3), f(\beta_4), f(\beta_5), f(\beta_6), f(\beta_7), f(\beta_8)\big),$$

where $f(x) = u_0 x^0 + u_1 x^1 + u_2 x^2$.

Coding via Evaluation (First Setup)



Encoding map (evaluation map):

$$(u_0, u_1, u_2) \mapsto \big(f(\beta_0), f(\beta_1), f(\beta_2), f(\beta_3), f(\beta_4), f(\beta_5), f(\beta_6), f(\beta_7), f(\beta_8)\big),$$

where $f(x) = u_0 x^0 + u_1 x^1 + u_2 x^2$.

A way to find (u_0, u_1, u_2) is to specify at least three function values.

Coding via Evaluation (Second Setup)

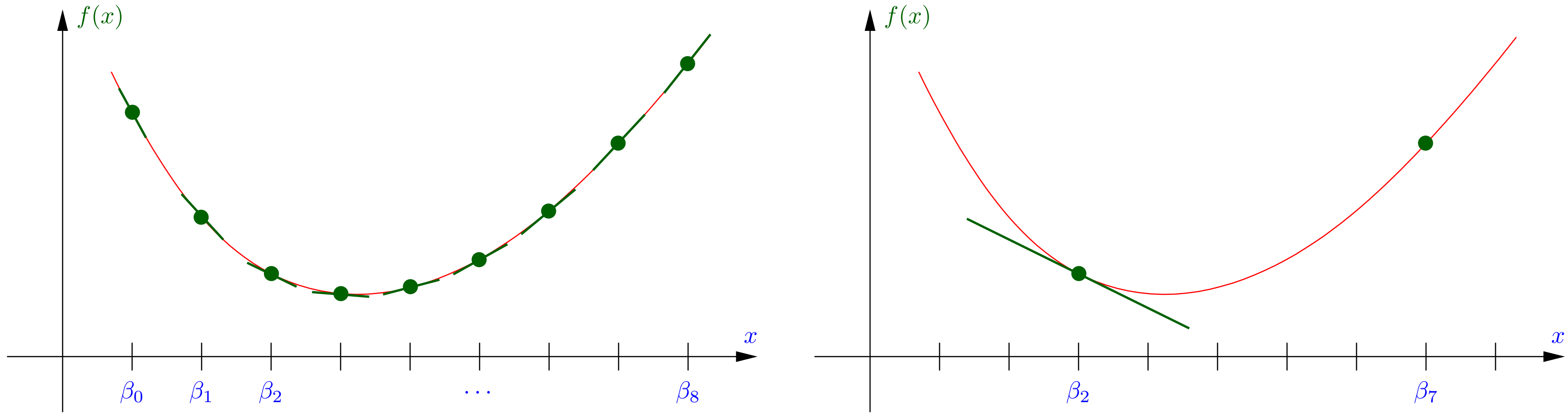
However, there are also other quantities that we can specify

so that we can find out (u_0, u_1, u_2).

Coding via Evaluation (Second Setup)

However, there are also other quantities that we can specify

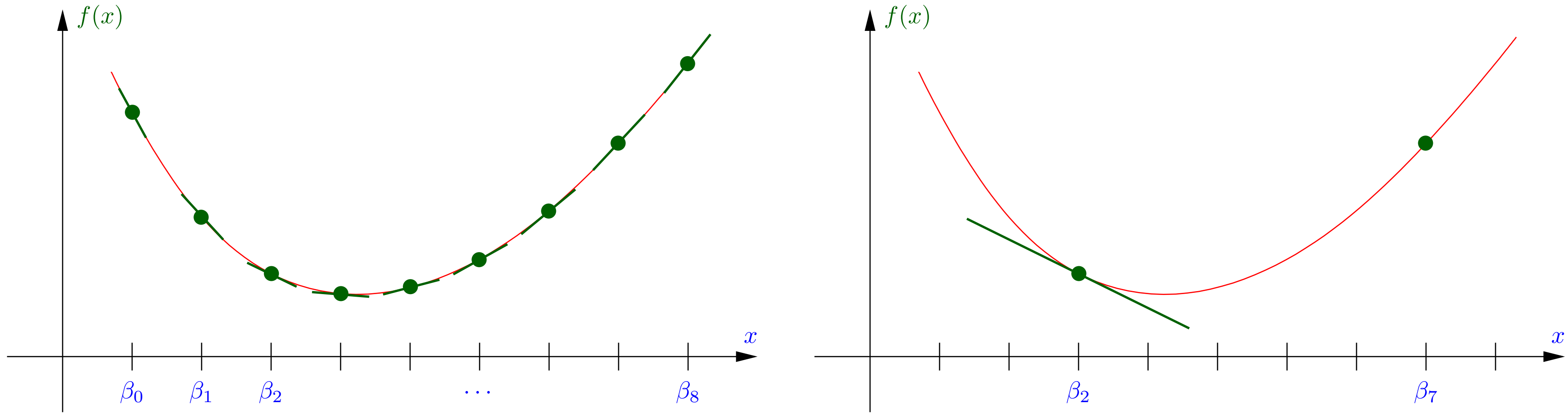
so that we can find out (u_0, u_1, u_2).



Coding via Evaluation (Second Setup)

However, there are also other quantities that we can specify

so that we can find out (u_0, u_1, u_2).



For example, knowing

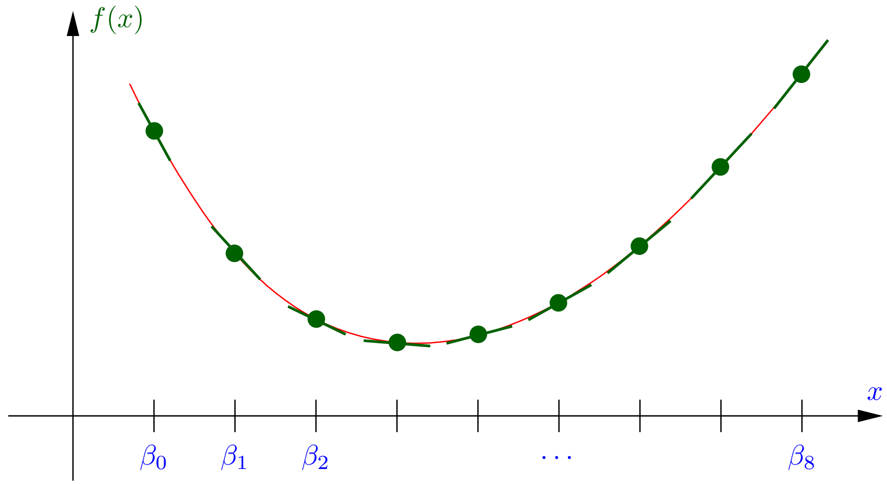
- the function value plus the value of the function derivative for one place and
- the function value at another place,

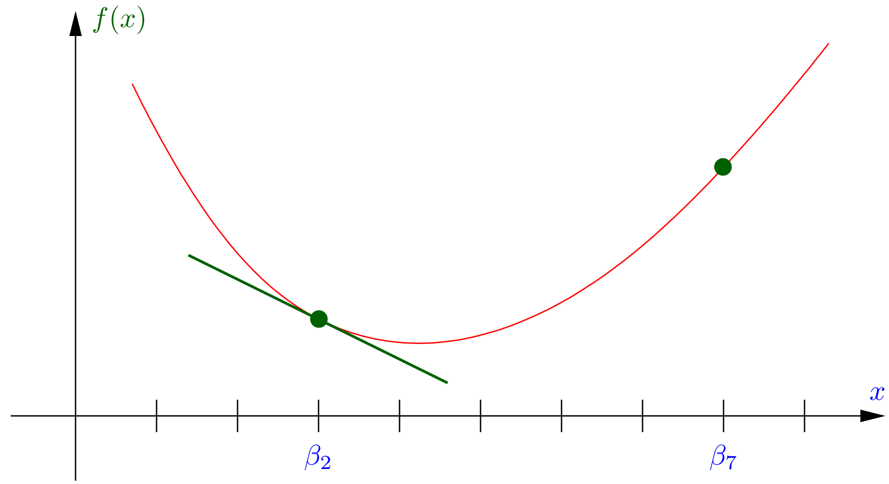
is sufficient to find (u_0, u_1, u_2).

Coding via Evaluation (Second Setup)

However, there are also other quantities that we can specify

so that we can find out (u_0, u_1, u_2).





Consider the following new evaluation map:

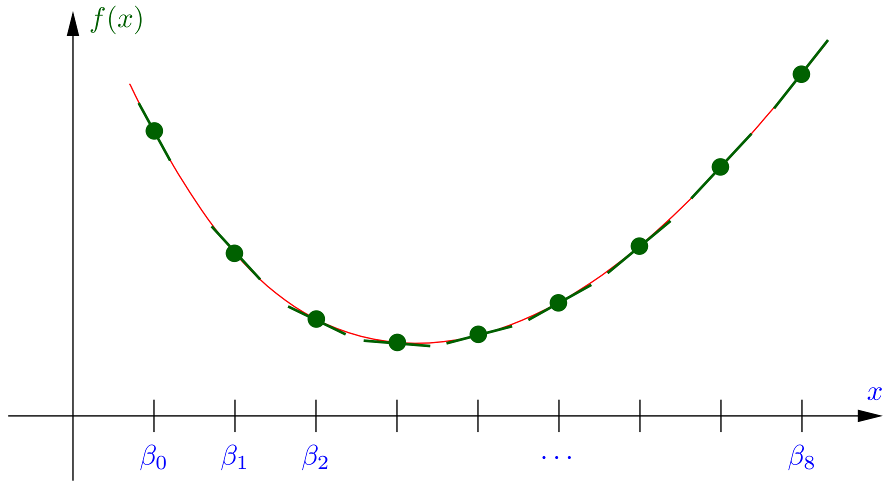
$$\begin{pmatrix} u_0 & u_1 & u_2 \end{pmatrix} \mapsto \begin{pmatrix} f(\beta_0) & f'(\beta_0) \\ \vdots & \vdots \\ f(\beta_8) & f'(\beta_8) \end{pmatrix}$$

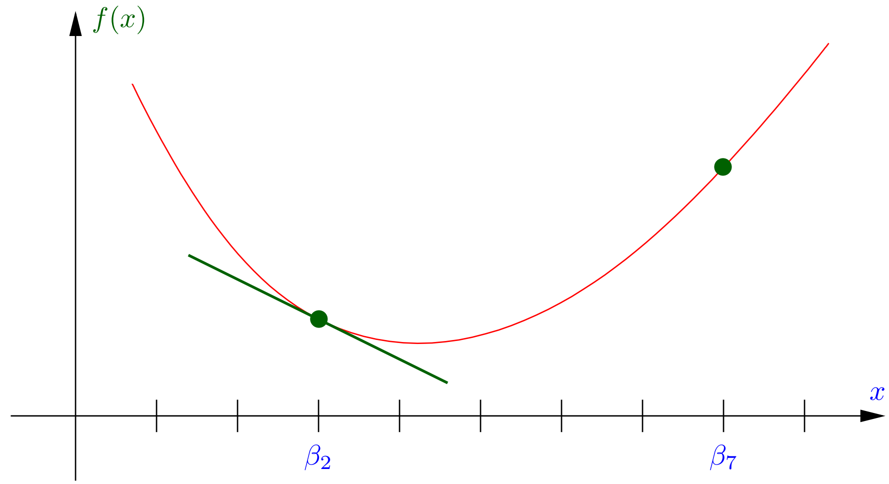
where $f(x) = u_0 x^0 + u_1 x^1 + u_2 x^2$ and $f'(x) = u_1 x^0 + 2u_2 x^1$.

Coding via Evaluation (Second Setup)

However, there are also other quantities that we can specify

so that we can find out (u_0, u_1, u_2).





General formula for the evaluation map:

$$\begin{pmatrix} u_0 & \cdots & u_{n-1} \end{pmatrix} \mapsto \begin{pmatrix} f^{(0)}(\beta_0) & f^{(1)}(\beta_0) & \cdots & f^{(n-1)}(\beta_0) \\ \vdots & \vdots & \vdots & \vdots \\ f^{(0)}(\beta_{L-1}) & f^{(1)}(\beta_{L-1}) & \cdots & f^{(n-1)}(\beta_{L-1}) \end{pmatrix}$$

where $f^{(i)}(x) = \sum_{t=0}^{n-1} \frac{t!}{(t-i)!} u_t x^t$ for $0 \le i \le n-1$.

Coding via Evaluation (Second Setup)

General formula for the evaluation map:

$$\begin{pmatrix} u_0 & \cdots & u_{n-1} \end{pmatrix} \mapsto \begin{pmatrix} f^{(0)}(\beta_0) & f^{(1)}(\beta_0) & \cdots & f^{(n-1)}(\beta_0) \\ \vdots & \vdots & \vdots & \vdots \\ f^{(0)}(\beta_{L-1}) & f^{(1)}(\beta_{L-1}) & \cdots & f^{(n-1)}(\beta_{L-1}) \end{pmatrix}$$

where we used the formal derivatives

$$f^{(i)}(x) = \sum_{t=0}^{n-1} \frac{t!}{(t-i)!} u_t x^t \quad \text{for } 0 \le i \le n-1.$$

Coding via Evaluation (Second Setup)

General formula for the evaluation map:

$$\begin{pmatrix} u_0 & \cdots & u_{n-1} \end{pmatrix} \mapsto \begin{pmatrix} f^{(0)}(\beta_0) & f^{(1)}(\beta_0) & \cdots & f^{(n-1)}(\beta_0) \\ \vdots & \vdots & \vdots & \vdots \\ f^{(0)}(\beta_{L-1}) & f^{(1)}(\beta_{L-1}) & \cdots & f^{(n-1)}(\beta_{L-1}) \end{pmatrix}$$

where we used the formal derivatives

$$f^{(i)}(x) = \sum_{t=0}^{n-1} \frac{t!}{(t-i)!} u_t x^t \quad \text{for } 0 \leq i \leq n-1.$$

There is a problem if we want to use this approach when we work over finite fields: if p is the characteristic of $\mathbb{F}_q$ then the i-th formal derivative is zero for $i \geq p$ and the corresponding channel symbols do not carry any information.

Coding via Evaluation (Second Setup)

General formula for the evaluation map:

$$\begin{pmatrix} u_0 & \cdots & u_{n-1} \end{pmatrix} \mapsto \begin{pmatrix} f^{(0)}(\beta_0) & f^{(1)}(\beta_0) & \cdots & f^{(n-1)}(\beta_0) \\ \vdots & \vdots & \vdots & \vdots \\ f^{(0)}(\beta_{L-1}) & f^{(1)}(\beta_{L-1}) & \cdots & f^{(n-1)}(\beta_{L-1}) \end{pmatrix}$$

where we used the formal derivatives

$$f^{(i)}(x) = \sum_{t=0}^{n-1} \frac{t!}{(t-i)!} u_t x^t \quad \text{for } 0 \leq i \leq n-1.$$

There is a problem if we want to use this approach when we work over finite fields: if p is the characteristic of $\mathbb{F}_q$ then the i-th formal derivative is zero for $i \geq p$ and the corresponding channel symbols do not carry any information.
However, replacing the formal derivative by the Hasse derivative, this approach works!

Coding via Evaluation (Second Setup)

General formula for the evaluation map:

$$\begin{pmatrix} u_0 & u_1 & u_2 \end{pmatrix} \mapsto \begin{pmatrix} \tilde{f}^{(0)}(\beta_0) & \tilde{f}^{(1)}(\beta_0) & \cdots & \tilde{f}^{(n-1)}(\beta_0) \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{f}^{(0)}(\beta_{L-1}) & \tilde{f}^{(1)}(\beta_{L-1}) & \cdots & \tilde{f}^{(n-1)}(\beta_{L-1}) \end{pmatrix}$$

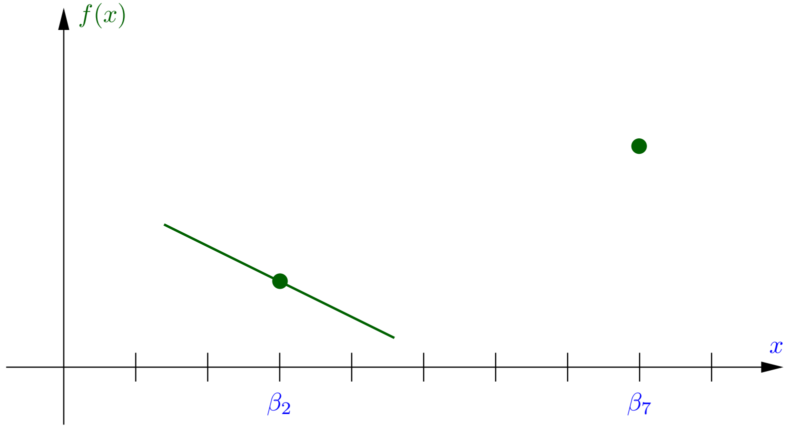
where we used the Hasse derivatives

$$\tilde{f}^{(i)}(x) = \sum_{t=0}^{n-1} \binom{t}{i} u_t x^t = \sum_{t=0}^{n-1} \frac{t!}{i!(t-i)!} u_t x^t \quad \text{for } 0 \leq i \leq n-1.$$

Coding via Evaluation (Second Setup)

Assume that we only receive

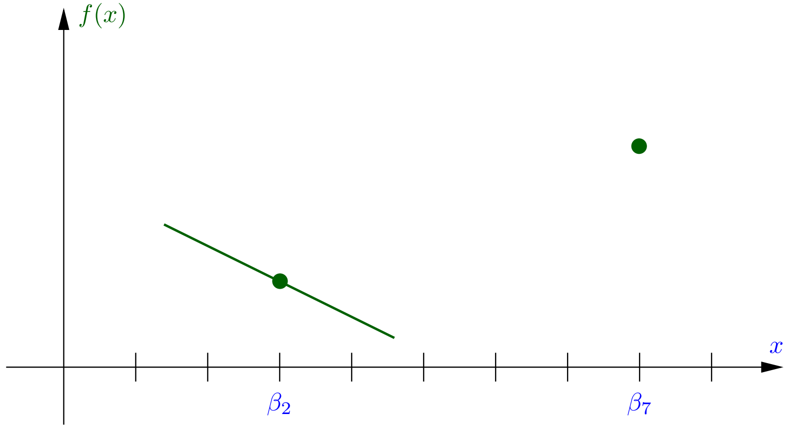
- the function value and the derivative for $x = \beta_2$ and
- the function value for $x = \beta_7$.



Coding via Evaluation (Second Setup)

Assume that we only receive

- the function value and the derivative for $x = \beta_2$ and
- the function value for $x = \beta_7$.

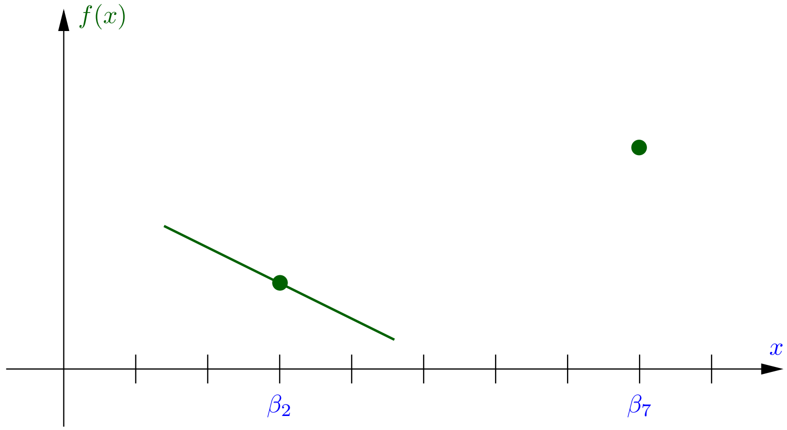


- We have to show that the mapping $\left(u_0, u_1, u_2\right) \mapsto \left(\tilde{f}^{(0)}(\beta_2), \tilde{f}^{(1)}(\beta_2), \tilde{f}^{(0)}(\beta_7)\right)$ is injective.

Coding via Evaluation (Second Setup)

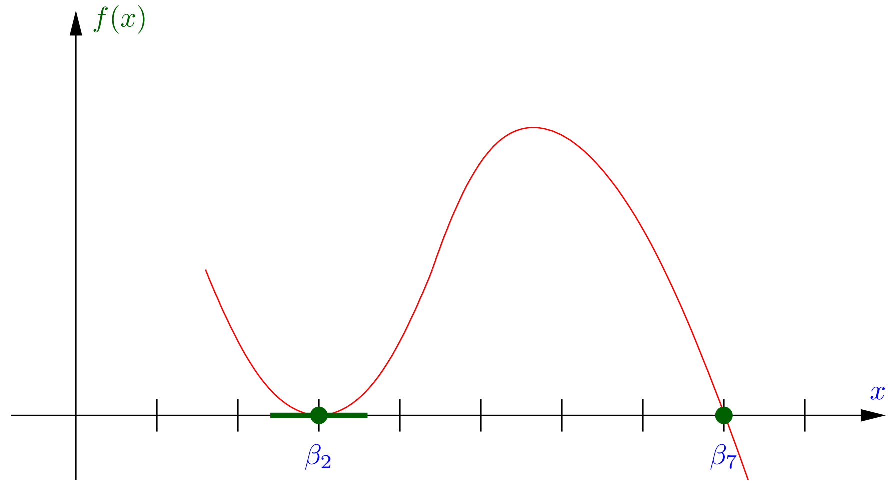
Assume that we only receive

- the function value and the derivative for $x = \beta_2$ and
- the function value for $x = \beta_7$.



- We have to show that the mapping $\left(u_0, u_1, u_2\right) \mapsto \left(\tilde{f}^{(0)}(\beta_2), \tilde{f}^{(1)}(\beta_2), \tilde{f}^{(0)}(\beta_7)\right)$ is injective.
- Because the above mapping is linear it is sufficient to show that the kernel is trivial.

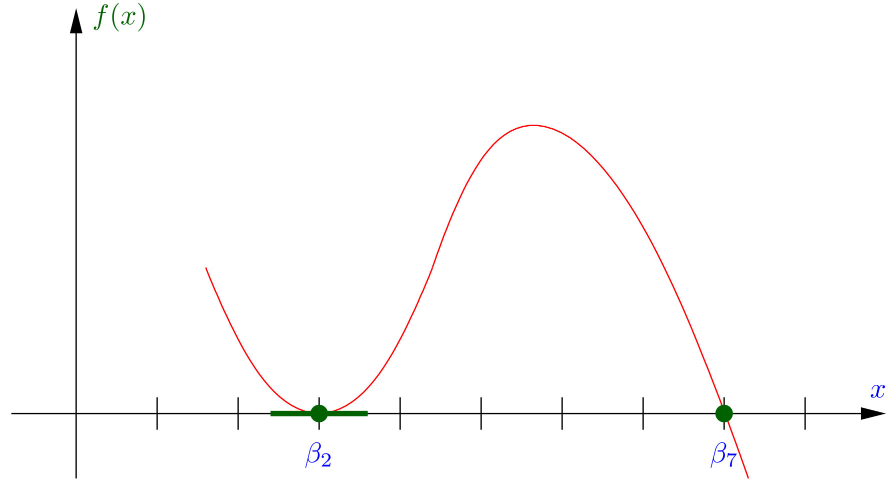
Coding via Evaluation (Second Setup)



Case 2:

$f(x) \neq 0$ with at least three zeros (counting with multiplicities).

Coding via Evaluation (Second Setup)

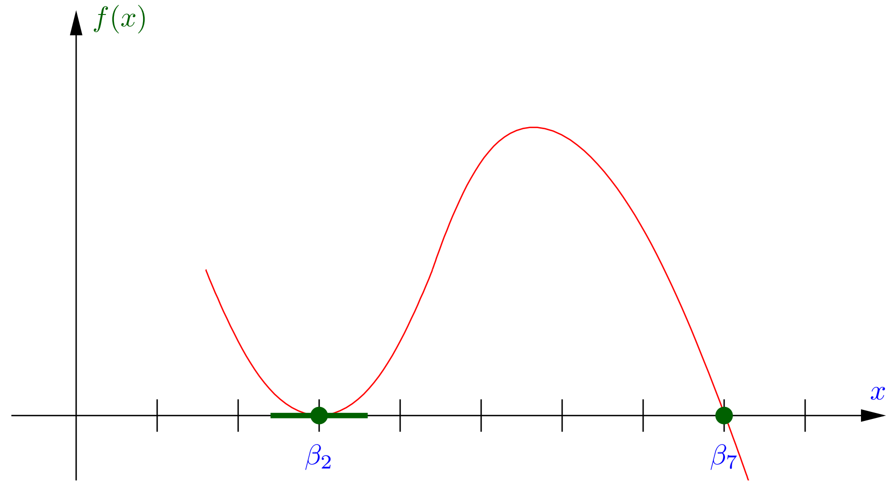


Case 2:

$f(x) \neq 0$ with at least three zeros (counting with multiplicities).

The fundamental theorem of algebra implies that $\deg(f(x)) \geq 3$.

Coding via Evaluation (Second Setup)

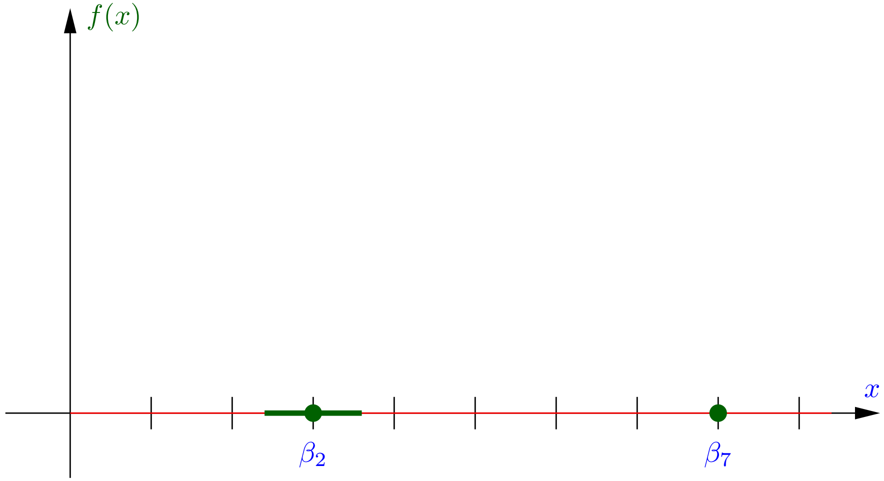


Case 2:

$f(x) \neq 0$ with at least three zeros (counting with multiplicities).

The fundamental theorem of algebra implies that $\deg(f(x)) \geq 3$. However, no quadratic function can have more than two zeros.

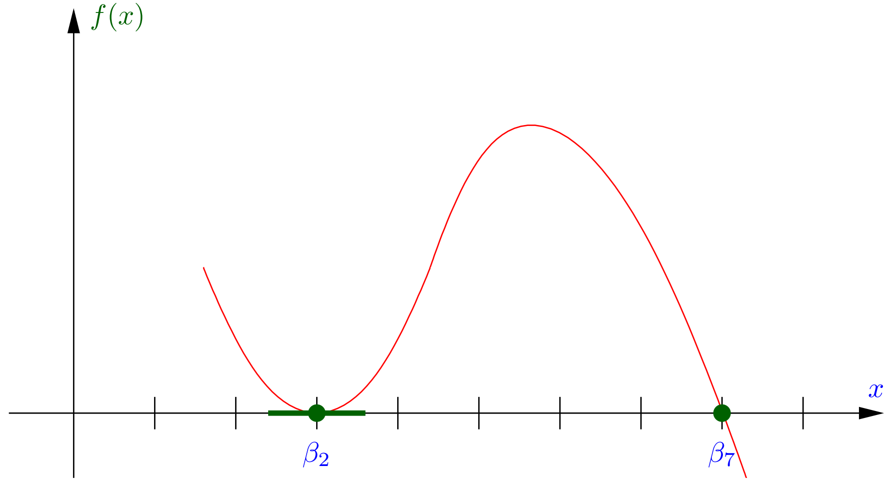
Coding via Evaluation (Second Setup)



Case 1:

$$f(x) = 0,$$

$$\Rightarrow \big(u_0, u_1, u_2\big) = \big(0, 0, 0\big).$$



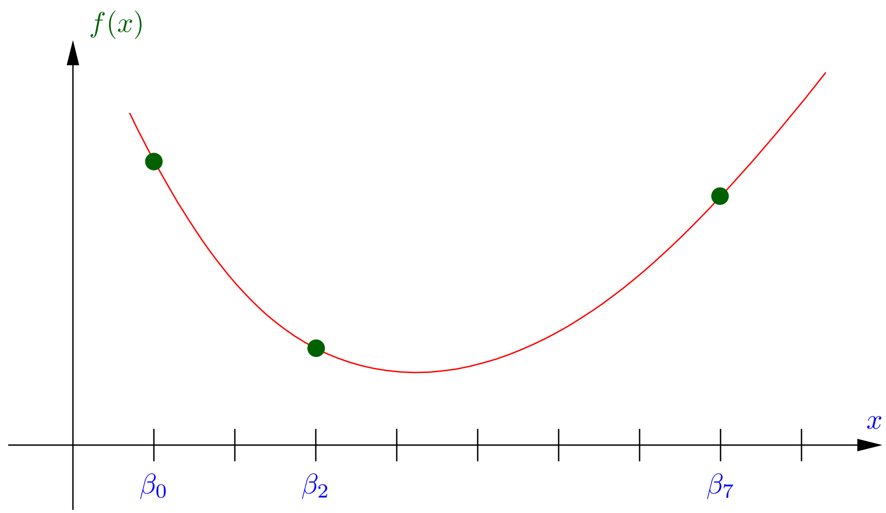
Case 2:

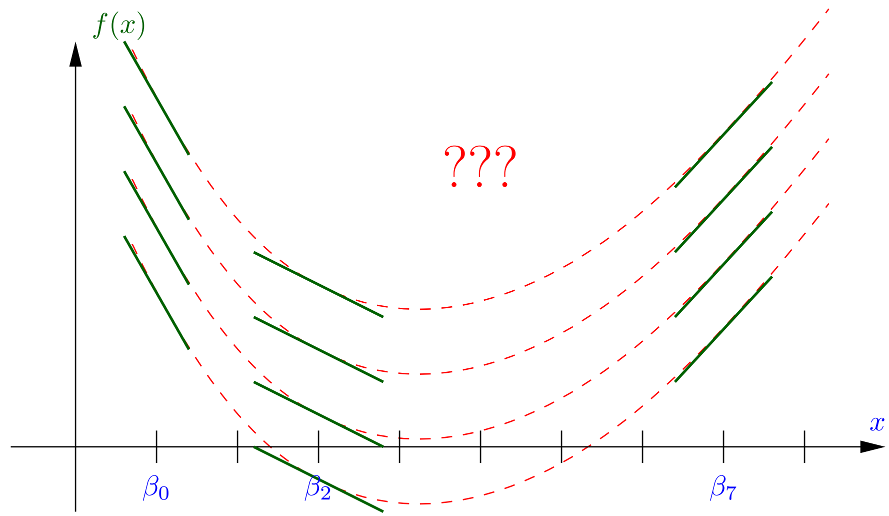
$f(x) \neq 0$ with at least three zeros (counting with multiplicities).

The fundamental theorem of algebra implies that $\deg(f(x)) \geq 3$. However, no quadratic function can have more than two zeros.

Coding via Evaluation (Second Setup)

Note that this second interpolation setup is not simply a special case of the first interpolation setup:





Knowing three points where a parabola goes through is sufficient to find out the parameters of the parabola.

Knowing e.g. the derivatives at three points of a parabola is not sufficient to find out the parameters of the parabola.

Universally decodable matrices (UDMs)

Universally Decodable Matrices

Proposition

- Let n be some positive integer, let q be some prime power.

Universally Decodable Matrices

Proposition

- Let n be some positive integer, let q be some prime power.
- Let α be a primitive element in $\mathbb{F}_q$.
  (I.e. α is an $(q-1)$-th primitive root of unity.)

Universally Decodable Matrices

Proposition

- Let n be some positive integer, let q be some prime power.
- Let α be a primitive element in $\mathbb{F}_q$.
  (I.e. α is an $(q-1)$-th primitive root of unity.)
- If $L \leq q+1$ then the following L matrices over $\mathbb{F}_q$ of size $n \times n$ are (L, n, q)-UDMs:

$$\mathbf{G}_0 \triangleq \mathbf{I}_n \,, \qquad \mathbf{G}_1 \triangleq \mathbf{J}_n \,, \qquad \mathbf{G}_2 \,, \qquad \ldots \,, \qquad \mathbf{G}_{L-1} \,,$$

  where
  - $\mathbf{J}_n$ is an $n \times n$ matrix with ones in the anti-diagonal and zeros otherwise;
  - $[\mathbf{G}_{\ell+2}]_{t,i} \triangleq \binom{t}{i} \alpha^{\ell(t-i)}, \ (\ell, t, i) \in [L-2] \times [n] \times [n]$.

Universally Decodable Matrices

E.g. $L = 4$, $n = 3$, $q = 3$.

$$\mathbf{G}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$\mathbf{G}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \quad \mathbf{G}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

Universally Decodable Matrices

E.g. $L = 4$, $n = 3$, $q = 3$.

$$\mathbf{G}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$\mathbf{G}_2 = \begin{pmatrix} \mathbf{1} & 0 & 0 \\ \mathbf{1} & \mathbf{1} & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{1} \end{pmatrix}, \quad \mathbf{G}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

Note that $[\mathbf{G}_2]_{t,i} \triangleq \binom{t}{i}$, therefore Pascal's triangle plays an important role when constructing these matrices.

Comments

- An explicit construction of UDMs was given in [Vontobel and Ganesan, 2006], [Ganesan and Vontobel, 2007].

Comments

- An explicit construction of UDMs was given in [Vontobel and Ganesan, 2006], [Ganesan and Vontobel, 2007].
- Earlier on, and in a different context, the resulting codes have also been defined by Rosenbloom and Tsfasman (PPI, 1997).

Comments

- An explicit construction of UDMs was given in [Vontobel and Ganesan, 2006], [Ganesan and Vontobel, 2007].
- Earlier on, and in a different context, the resulting codes have also been defined by Rosenbloom and Tsfasman (PPI, 1997).
- In the last ten years, the resulting codes have also appeared under the name “multiplicity codes” in the theoretical computer science literature.

Comments

- An explicit construction of UDMs was given in [Vontobel and Ganesan, 2006], [Ganesan and Vontobel, 2007].

- Earlier on, and in a different context, the resulting codes have also been defined by Rosenbloom and Tsfasman (PPI, 1997).

- In the last ten years, the resulting codes have also appeared under the name “multiplicity codes” in the theoretical computer science literature.

- The mathematics that is needed is very similar to the mathematics that is needed when studying so-called repeated-root cyclic codes [Castagnoli et al., 1991].

Comments

- An explicit construction of UDMs was given in [Vontobel and Ganesan, 2006], [Ganesan and Vontobel, 2007].
- Earlier on, and in a different context, the resulting codes have also been defined by Rosenbloom and Tsfasman (PPI, 1997).
- In the last ten years, the resulting codes have also appeared under the name "multiplicity codes" in the theoretical computer science literature.
- The mathematics that is needed is very similar to the mathematics that is needed when studying so-called repeated-root cyclic codes [Castagnoli et al., 1991].
- Are there other constructions of UDMs that are not simply reformulations of the above UDMs? Note that one can show that the given construction is in a certain sense a unique extension of Reed-Solomon codes [Vontobel and Ganesan, 2006].

Efficient Decoding

- Decoding means that we have to solve the system of linear equations

$$\mathbf{y} = \mathbf{u} \cdot \mathbf{G}.$$

Efficient Decoding

- Decoding means that we have to solve the system of linear equations

$$\mathbf{y} \quad = \quad \mathbf{u} \cdot \mathbf{G}.$$

Using Gaussian elimination, the decoding complexity is $O(n^3)$.

Efficient Decoding

- Decoding means that we have to solve the system of linear equations

$$\mathbf{y} = \mathbf{u} \cdot \mathbf{G}.$$

  Using Gaussian elimination, the decoding complexity is $O(n^3)$.

- However, decoding is obviously related to finding an interpolation polynomial: the problem at hand can be solved with a variant of Newton's interpolation algorithm.

Efficient Decoding

- Decoding means that we have to solve the system of linear equations

$$\mathbf{y} = \mathbf{u} \cdot \mathbf{G}.$$

  Using Gaussian elimination, the decoding complexity is $O(n^3)$.

- However, decoding is obviously related to finding an interpolation polynomial: the problem at hand can be solved with a variant of Newton's interpolation algorithm. This results in a decoding complexity of $O(n^2)$.

Generalizations (Part 1/2)

- Remember the encoding that we are using

$$\mathbf{u} \in \mathbb{F}_q^n \quad \mapsto \quad \mathbf{x}_\ell \in \mathbb{F}_q^n,\ \ell \in [L].$$

Generalizations (Part 1/2)

- Remember the encoding that we are using

$$\mathbf{u} \in \mathbb{F}_q^n \quad \mapsto \quad \mathbf{x}_\ell \in \mathbb{F}_q^n,\ \ell \in [L].$$

- Generalization: for any $1 \leq n' \leq n$ we can also send vectors of length n':

$$\mathbf{u} \in \mathbb{F}_q^n \quad \mapsto \quad \mathbf{x}_\ell \in \mathbb{F}_q^{n'},\ \ell \in [L].$$

Generalizations (Part 1/2)

- Remember the encoding that we are using

$$\mathbf{u} \in \mathbb{F}_q^n \quad \mapsto \quad \mathbf{x}_\ell \in \mathbb{F}_q^n,\ \ell \in [L].$$

- Generalization: for any $1 \leq n' \leq n$ we can also send vectors of length n':

$$\mathbf{u} \in \mathbb{F}_q^n \quad \mapsto \quad \mathbf{x}_\ell \in \mathbb{F}_q^{n'},\ \ell \in [L].$$

$\Rightarrow$ The above construction of UDMs can be extended straightforwardly to this new setup.

Generalizations (Part 2/2)

- Remember that for any $k_0, \ldots, k_{L-1}$ with $\sum_{k \in [L]} k_\ell \geq n$ we required that we can decode uniquely.

Generalizations (Part 2/2)

- Remember that for any $k_0, \ldots, k_{L-1}$ with $\sum_{k \in [L]} k_\ell \geq n$ we required that we can decode uniquely.
- Generalization: for any $k_0, \ldots, k_{L-1}$ with $\sum_{k \in [L]} k_\ell \geq n + g$ we require that we can decode uniquely for some $g \geq 0$.

Generalizations (Part 2/2)

- Remember that for any $k_0, \ldots, k_{L-1}$ with $\sum_{k \in [L]} k_\ell \geq n$ we required that we can decode uniquely.
- Generalization: for any $k_0, \ldots, k_{L-1}$ with $\sum_{k \in [L]} k_\ell \geq n + g$ we require that we can decode uniquely for some $g \geq 0$.

$\Rightarrow$ In the same way as Goppa codes / algebraic-geometry codes are generalizations of Reed-Solomon codes, one can construct UDMs that are generalizations of the above UDMs.

Generalizations (Part 2/2)

- Remember that for any $k_0, \ldots, k_{L-1}$ with $\sum_{k \in [L]} k_\ell \geq n$ we required that we can decode uniquely.
- Generalization: for any $k_0, \ldots, k_{L-1}$ with $\sum_{k \in [L]} k_\ell \geq n + g$ we require that we can decode uniquely for some $g \geq 0$.

$\Rightarrow$ In the same way as Goppa codes / algebraic-geometry codes are generalizations of Reed-Solomon codes, one can construct UDMs that are generalizations of the above UDMs.

- Riemann-Roch theorem gives new proof.

Generalizations (Part 2/2)

- Remember that for any $k_0, \ldots, k_{L-1}$ with $\sum_{k \in [L]} k_\ell \geq n$ we required that we can decode uniquely.
- Generalization: for any $k_0, \ldots, k_{L-1}$ with $\sum_{k \in [L]} k_\ell \geq n + g$ we require that we can decode uniquely for some $g \geq 0$.

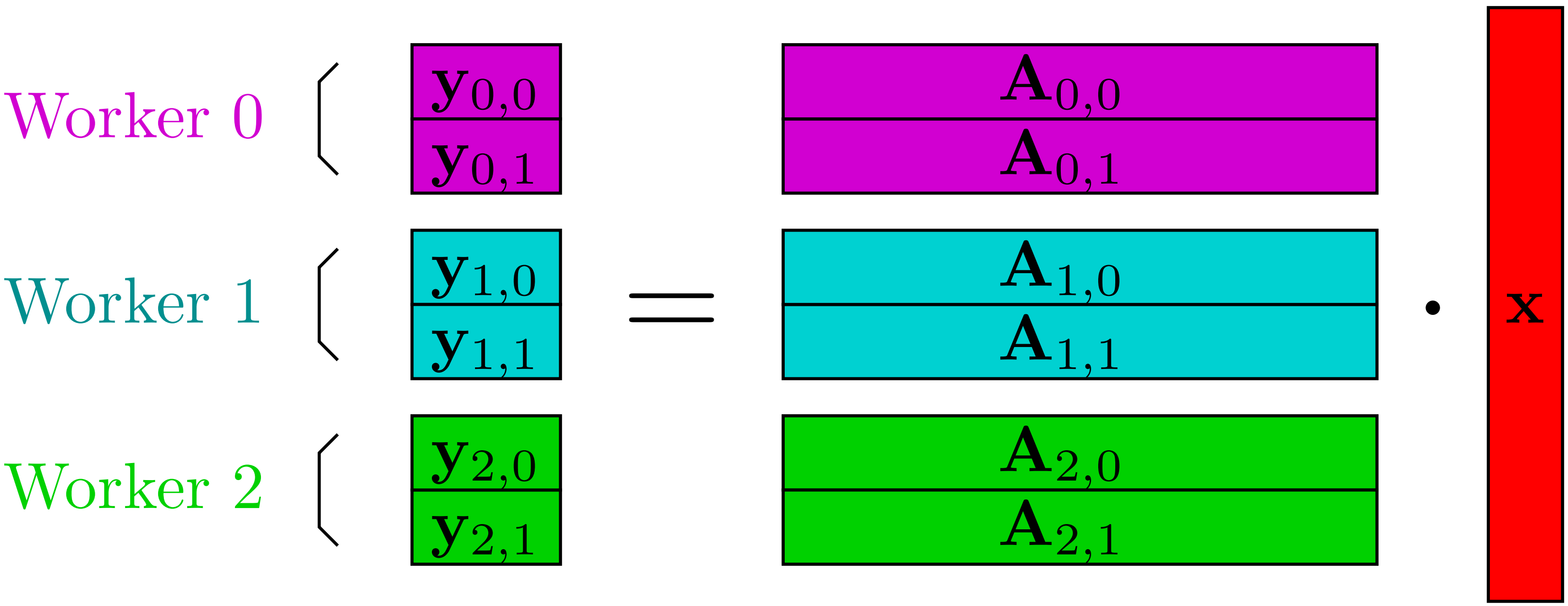
$\Rightarrow$ In the same way as Goppa codes / algebraic-geometry codes are generalizations of Reed-Solomon codes, one can construct UDMs that are generalizations of the above UDMs.

- Riemann-Roch theorem gives new proof.
- Hasse-Weil-Serre bound can be used to give new necessary conditions for L.

Back to the setup of interest

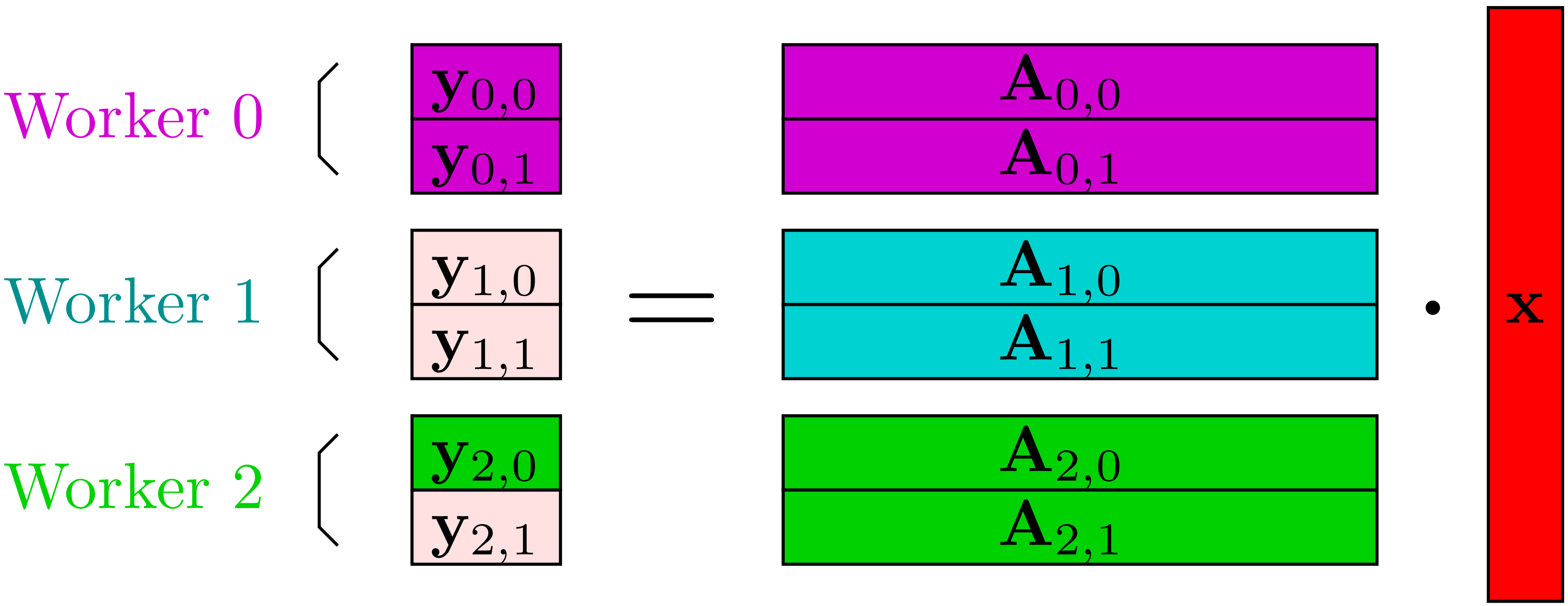
Motivation

We can split up the task into several submatrix-vector-multiplication tasks:



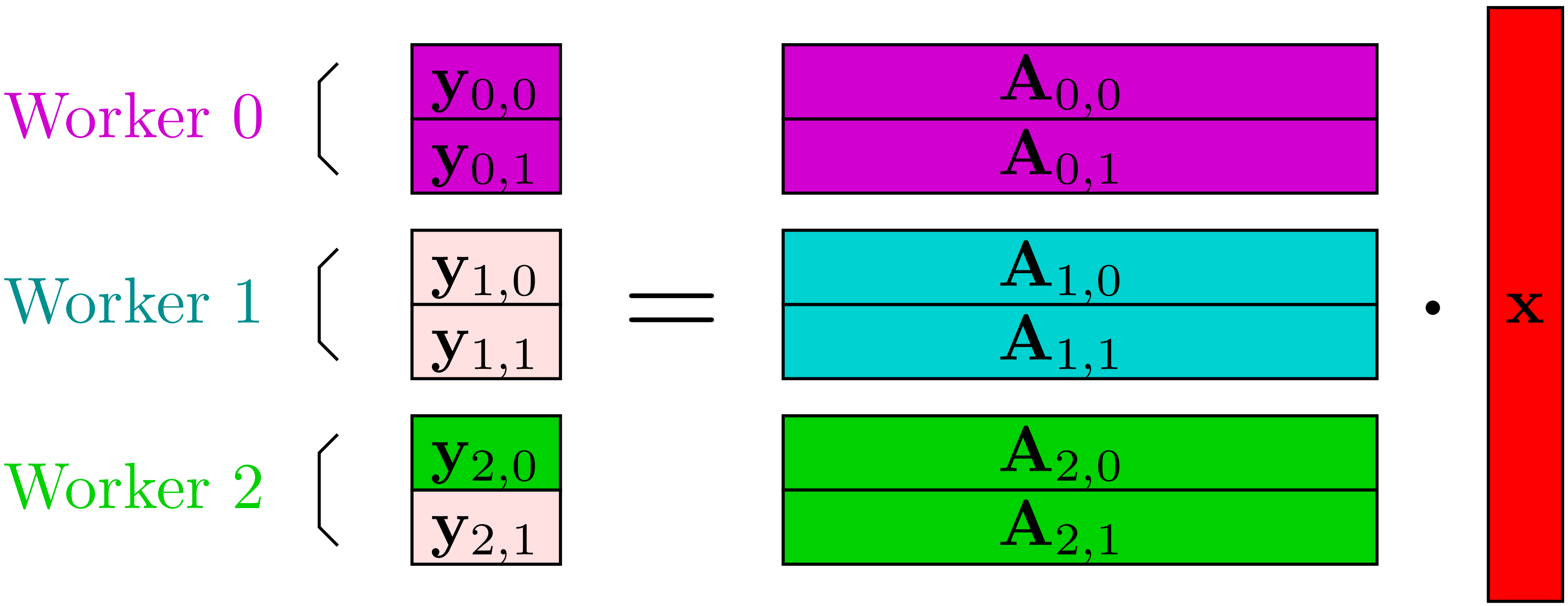
Motivation

We can split up the task into several submatrix-vector-multiplication tasks:



Motivation

We can split up the task into several submatrix-vector-multiplication tasks:



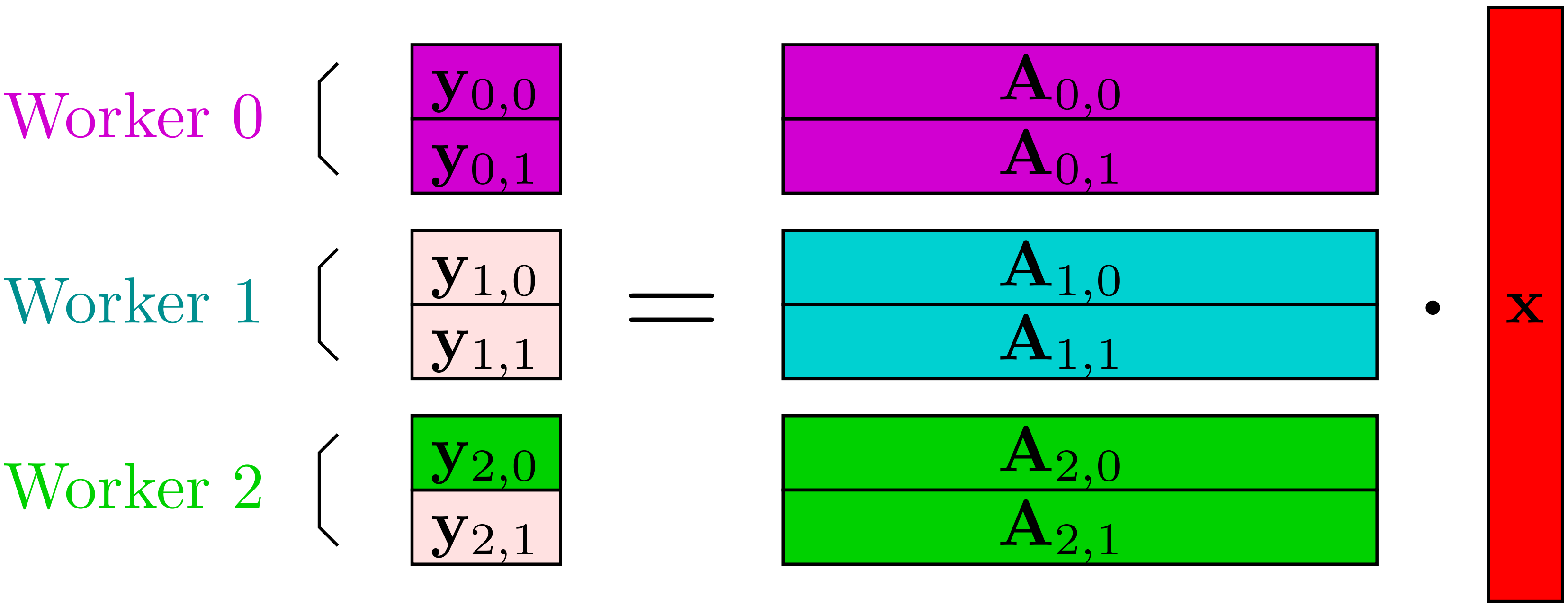
Idea:

- Coding scheme should take advantage of the fact that **erasures are correlated**.

Erasures are correlated because
if a partial result by one of the workers is not available,
then **all subsequent results by the same worker** are not available either.

Motivation

We can split up the task into several submatrix-vector-multiplication tasks:

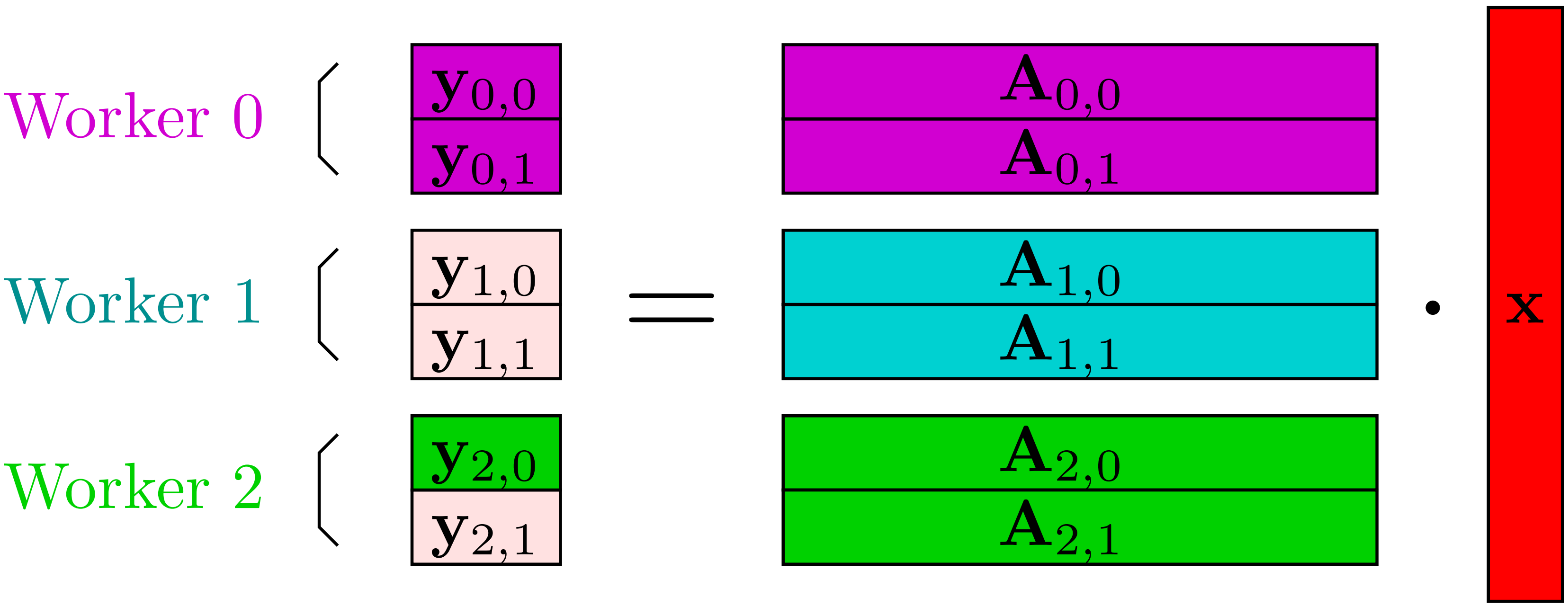


Idea:

- Base coding scheme on so-called **universally decodable matrices (UDMs)**.

Motivation

We can split up the task into several submatrix-vector-multiplication tasks:



Idea:

- Base coding scheme on so-called **universally decodable matrices (UDMs)**.
- Use **companion matrices** in order to **reduce issues with condition numbers** when adapting a coding scheme over some finite field to a coding scheme over the reals.

Embedding into the reals:

companion matrices

Companion Matrices

Assume that the field $\langle \mathbb{F}_{p^s}, +, \cdot \rangle$ is constructed based on the primitive polynomial

$$\pi(\mathsf{X}) \;=\; \mathsf{X}^s + \pi_{s-1}\mathsf{X}^{s-1} + \cdots + \pi_1\mathsf{X} + \pi_0 \;\in\; \mathbb{F}_p[\mathsf{X}]\,.$$

The companion matrix associated with $\pi(\mathsf{X})$ is defined to be the following matrix of size $s \times s$ over $\mathbb{F}_p$:

$$\mathbf{C} \triangleq \begin{pmatrix} 0 & 0 & \cdots & 0 & -\pi_0 \\ 1 & 0 & \cdots & 0 & -\pi_1 \\ 0 & 1 & \cdots & 0 & -\pi_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\pi_{s-1} \end{pmatrix}.$$

This matrix yields the following field isomorphism:

$$\langle \mathbb{F}_{p^s}, +, \cdot \rangle \;\cong\; \left\langle \left\{\mathbf{0}, \mathbf{C}, \mathbf{C}^2, \mathbf{C}^3, \ldots, \mathbf{C}^{p^s-1}\right\}, +, \cdot \right\rangle.$$

Companion Matrices

Lemma: let M be a square matrix with entries in $\mathbb{Z}$.

If M satisfies

$$\det(\mathbf{M}) \neq 0 \quad \text{(mod } p\text{)} ,$$

then also

$$\det(\mathbf{M}) \neq 0 \quad \text{(in } \mathbb{Z}\text{)} ,$$

and with that

$$\det(\mathbf{M}) \neq 0 \quad \text{(in } \mathbb{R}\text{)} .$$

Companion Matrices

Lemma: let M be a square matrix with entries in $\mathbb{Z}$.

If M satisfies

$$\det(\mathbf{M}) \neq 0 \quad \text{(mod } p\text{)}\,,$$

then also

$$\det(\mathbf{M}) \neq 0 \quad \text{(in } \mathbb{Z}\text{)}\,,$$

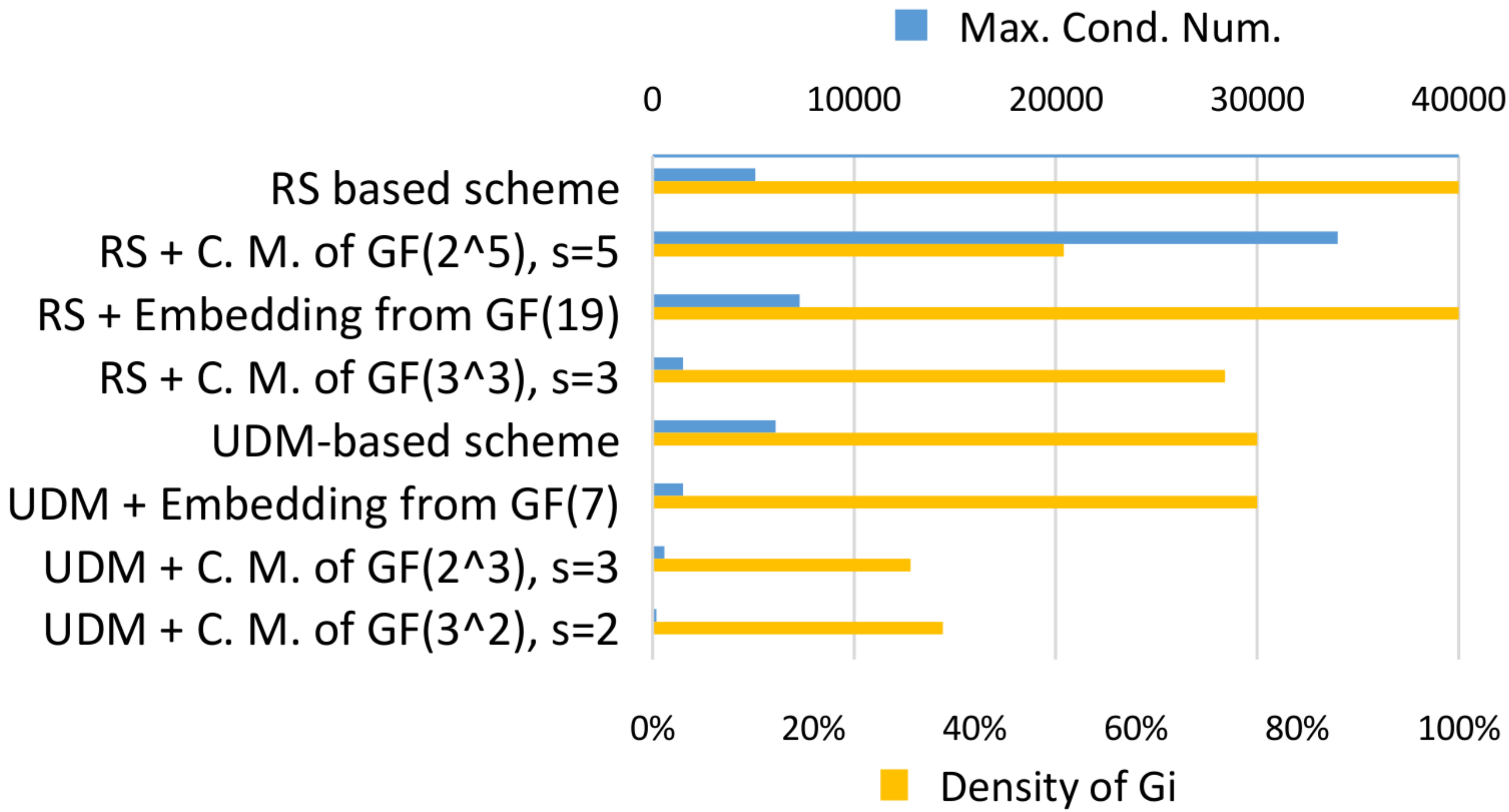
and with that

$$\det(\mathbf{M}) \neq 0 \quad \text{(in } \mathbb{R}\text{)}\,.$$

The above observations can be used to embed matrices over $\mathbb{F}_{p^s}$ into $\mathbb{R}$, and then give guarantees on them.

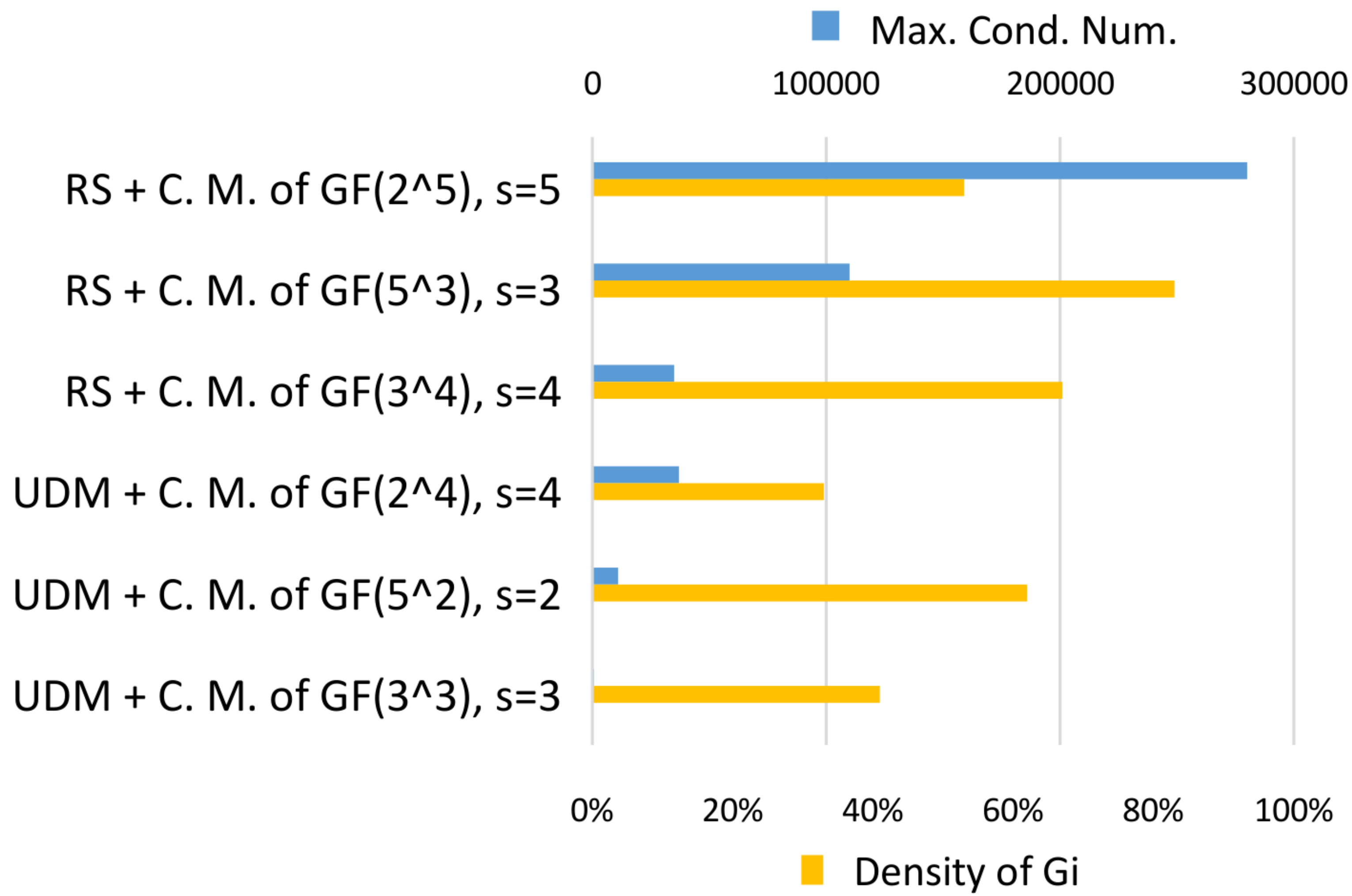
Performance comparison

Performance Comparison (Part 1/2)



Setup: $N = 6, \gamma = 3/4$, and $Q_{\mathrm{b}} = 4$.

Performance Comparison (Part 2/2)



Setup: $N = 15, \gamma = 1/2$, **and** $Q_{\mathrm{b}} = 4$.

References

References

A. Ramamoorthy, L. Tang, and P. O. Vontobel, "Universally decodable matrices for distributed matrix-vector multiplication," Proc. IEEE Int. Symp. Inf. Theory, Paris, France, pp. 1777-1781, July 2019. arXiv:1901.10674.

- M. Y. Rosenbloom and M. A. Tsfasman, "Codes for the m-metric," Probl. Inf. Transm., vol. 33, no. 1, pp. 45–52, 1997.
- Tavildar and Viswanath, "Approximately universal codes over slow fading channels," IEEE Trans. Inf. Theory, IT–52, no. 7, pp. 3233–3258, July 2006.
- P. O. Vontobel and A. Ganesan, "On universally decodable matrices for space-time coding", Designs, Codes, and Cryptography, Nov. 2006.
- A. Ganesan and P. O. Vontobel, "On the existence of universally decodable matrices," IEEE Trans. on Inf. Theory, vol. 53, no. 7, pp. 2572–2575, 2007.


Thank you!