Asymptotic Rates for Lattice-Based WOM Codes

Brian M. Kurkoski
kurkoski@ice.uec.ac.jp

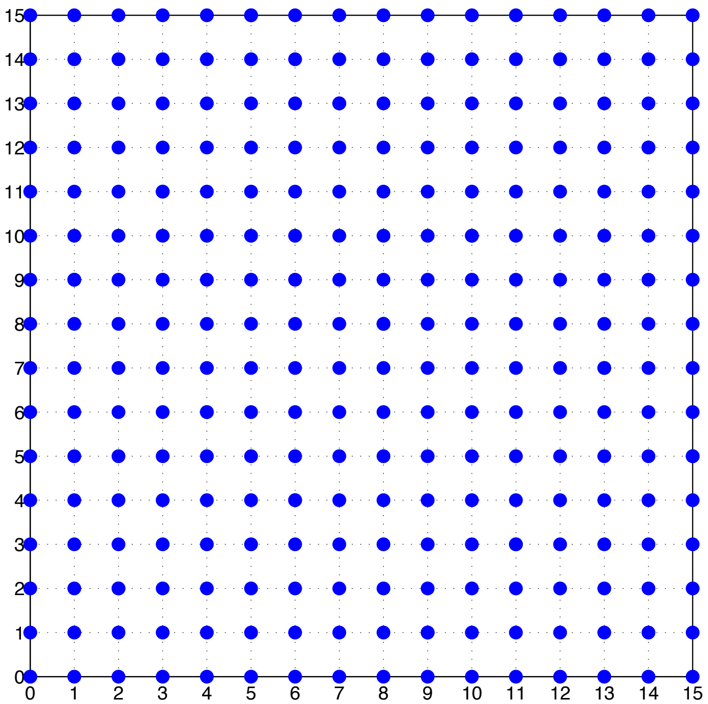
University of Electro-Communications
Tokyo, Japan

3rd Annual Non-Volatile Memories Workshop 2012
University of California, San Diego
March 5, 2012

Sphere Packing is an arrangement of non-overlapping spheres in space

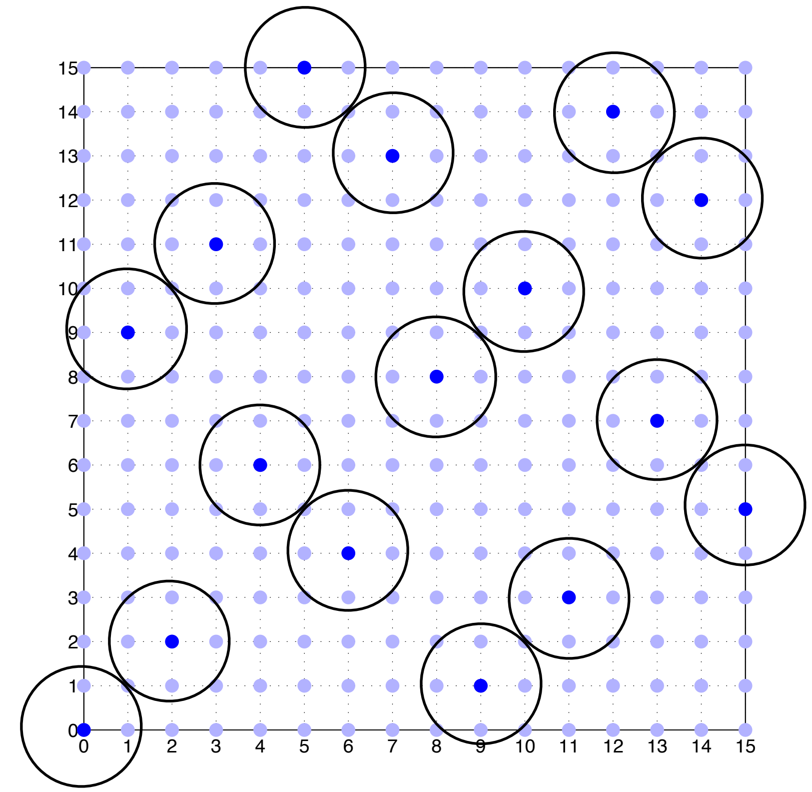
image © JJ Harrison
found on Wikipedia

Sphere Packings and Lattices Can Correct Errors



2 cells with $q = 16$

Gray mapping &
(8,4) Extended Hamming
$d_{\min} = 4$



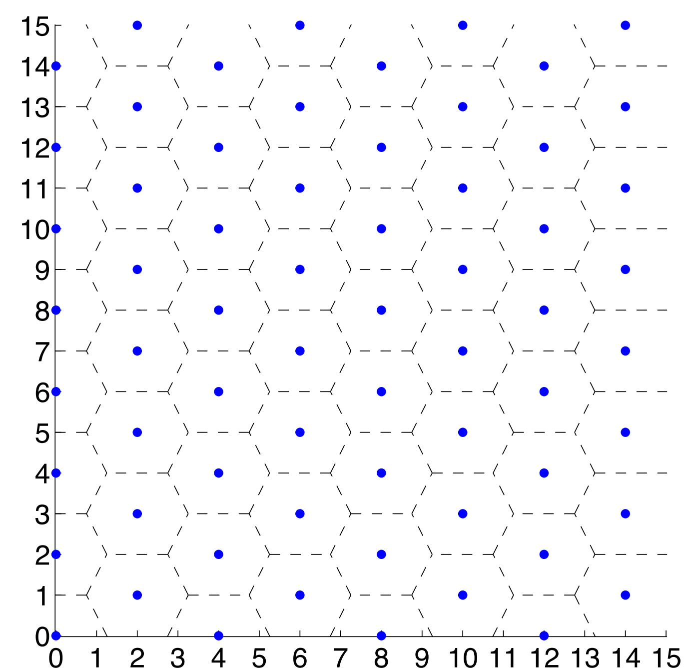
Sphere Packing (non-linear)

A **lattice** is a linear sphere packing

Lattices:

- Have a rich theory
- Can correct errors, achieve capacity

What about the WOM properties of lattices?



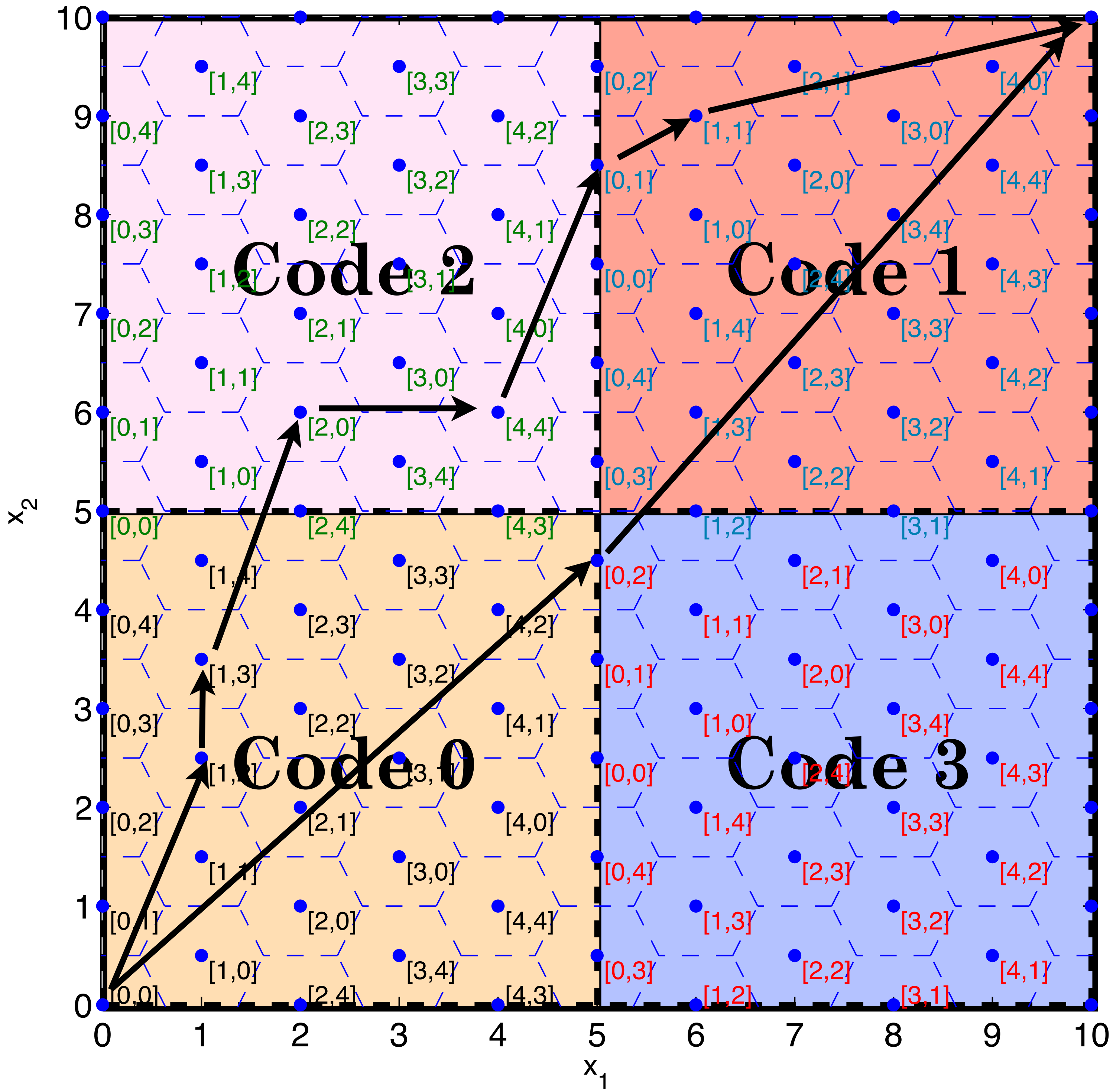
Lattice (linear)

Notes:
- *n*-dimensional lattice encoded in *n* flash cells
- pictures are *n*=2, but arbitrary *n* is possible

Overview of this talk

- WOM (write-once memory) codes are a coding-theoretic approach to extend the life of flash memories
- Consider WOM codes based on lattices
  - Most multilevel WOM codes are based on the integer lattice $\mathbf{Z}^n$
  - $\mathbf{Z}^n$ is a special case of this work
- Restrict our attention to 2-write WOM codes. Maximize rates.
  - Invoke Forney's continuous approximation for AWGN channels
  - Normalized coding rate
    - rate penalty separates **"shaping penalty"** and **"coding penalty"**
  - Idealized shaping region is a **hyperbola** in n dimensions
- If the rate for the two writes are equal:
  - The asymptotic shaping penalty is $\frac{2}{e \ln 2}$
  - Show an asymptotic gain of 0.4693 bit/dimension over "cubic construction"

Existing Construction: “Cubic Scheme”



Partition the lattice into four codes: Code 0,1,2,3

(Each code is one-to-one with information)

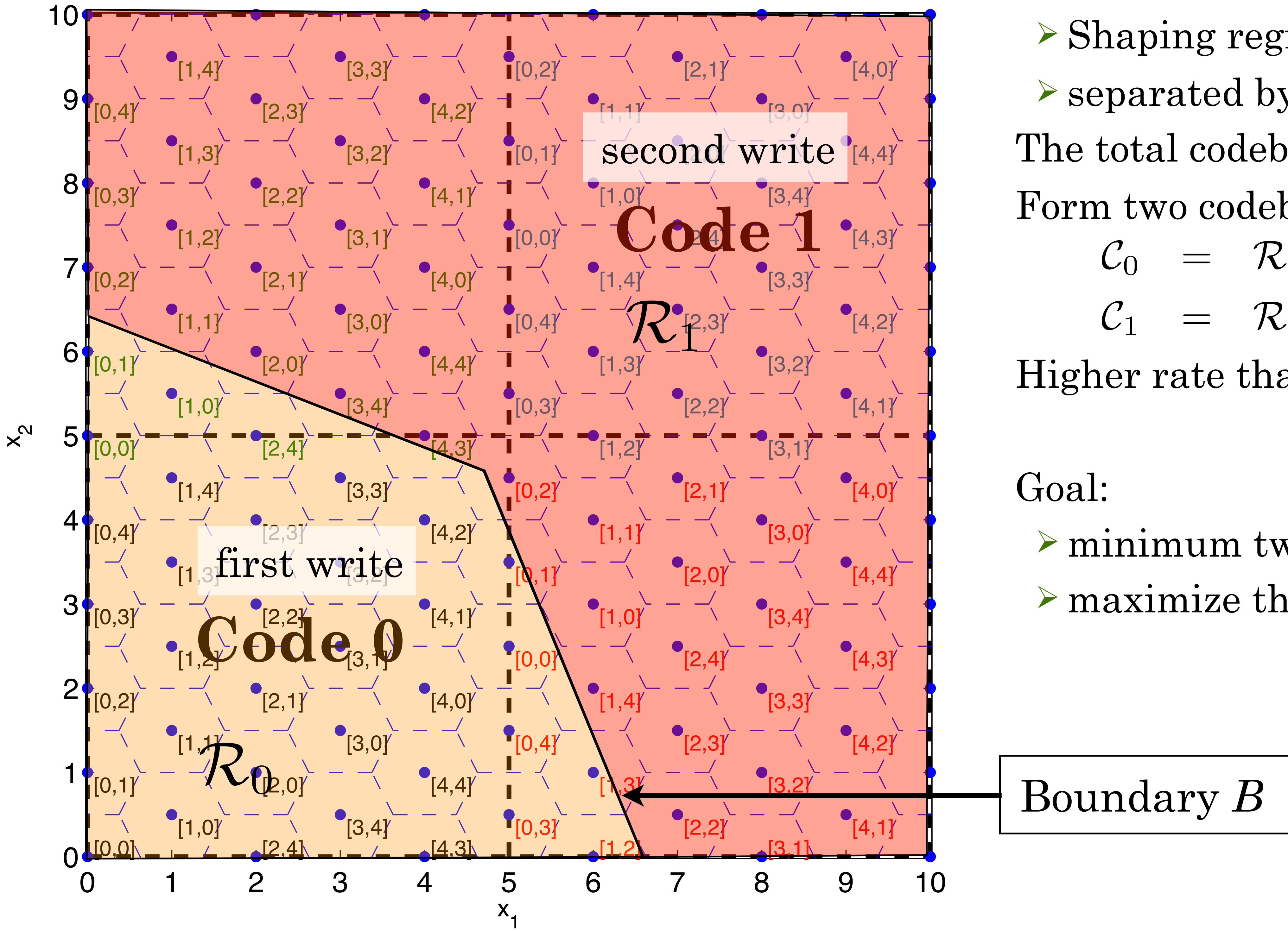
Previous work

- concentrated on average number of writes
- The minimum number of writes is 2.
- But clearly there are unused lattice points in Code 2 and Code 3.

This work:

- maximize the rate, when the minimum number of writes is two

A Higher Rate Construction — Guarantee 2 Writes



Restrict to two writes

- Shaping regions $\mathcal{R}_0, \mathcal{R}_1$
- separated by boundary B

The total codebook is $\mathcal{C}$

Form two codebooks

$$\begin{aligned} \mathcal{C}_0 &= \mathcal{R}_0 \cap \mathcal{C} \\ \mathcal{C}_1 &= \mathcal{R}_1 \cap \mathcal{C} \end{aligned}$$

Higher rate than cubic scheme

Goal:

- minimum two writes
- maximize the rate

Definition of Normalized Rate

Rates for 1st write:

Conventional Rate: $$R_0 = \frac{1}{n} \log_2 |\mathcal{C}_0|; \quad 0 \leq R_0 \leq \log_2 q$$

Normalized Rate: $$\widetilde{R}_0 = \frac{\log_2 |\mathcal{C}_0|}{\log_2 |\mathcal{C}|}; \quad 0 \leq \widetilde{R}_0 \leq 1$$

Since cell values only increase $\Rightarrow S(x)$ is a box

Rates for 2nd write:

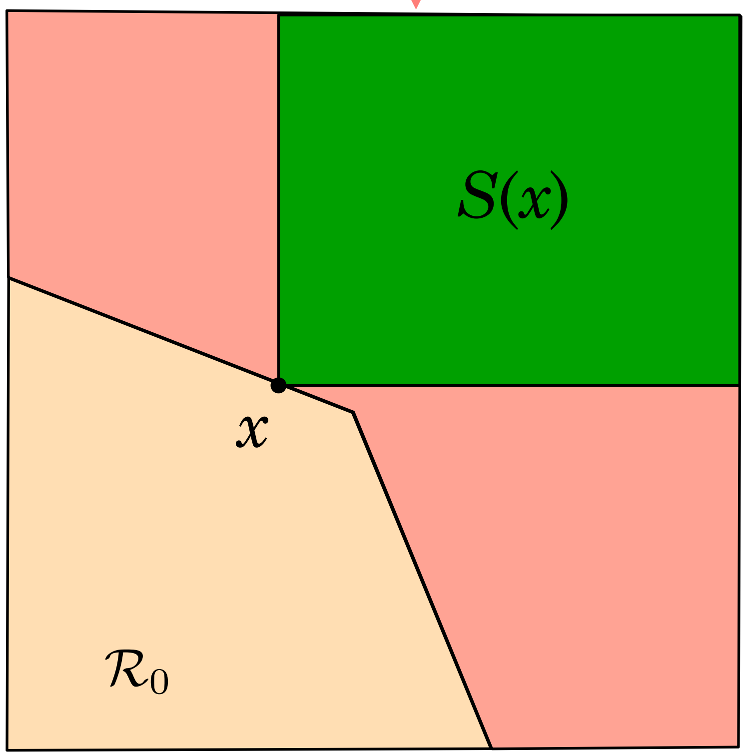
Define a region $S(x)$, the space "reached" from x

$S(x)$ is a box in n dimensions

$$S_1 = \min_{x \in \mathcal{C}_0} |S(x) \cap \Lambda|$$

$$R_1 = \frac{1}{n} \log_2 |S_1|; \quad 0 \leq R_1 \leq \log_2 q$$

$$\widetilde{R}_1 = \frac{\log_2 |S_1|}{\log_2 |\mathcal{C}|}; \quad 0 \leq \widetilde{R}_1 \leq 1$$



Continuous Approximation — AWGN Channels

- Assumption that code points are **uniformly** distributed over the shaping region

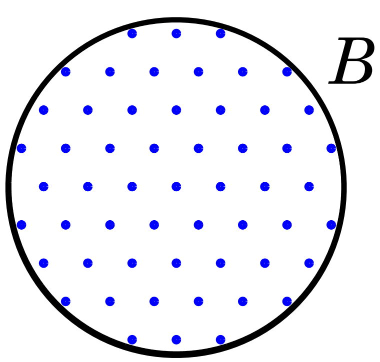
Separate lattice Λ and shaping region B contribution to signal power:

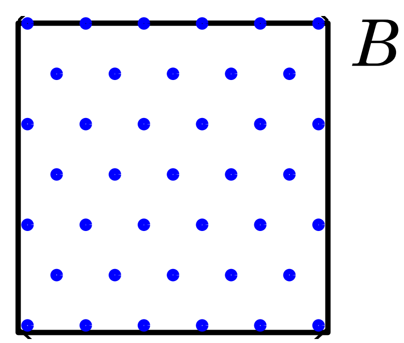
$$\text{Average Power} \approx \underbrace{\frac{\int_B ||\mathbf{x}||^2 d\mathbf{x}}{nV(B)^{\frac{2}{n}+1}}}_{G(B)} \cdot M^n V(\Lambda)$$

Depends only on *shape* of B
(normalized second moment)

Depends only
on coding lattice $\mathbf{\Lambda}$

Shaping Gain





$$\lim_{n\to\infty} \frac{G(\text{cube})}{G(n\text{-sphere})} = \frac{\pi e}{6} = 1.53 \text{ dB}$$

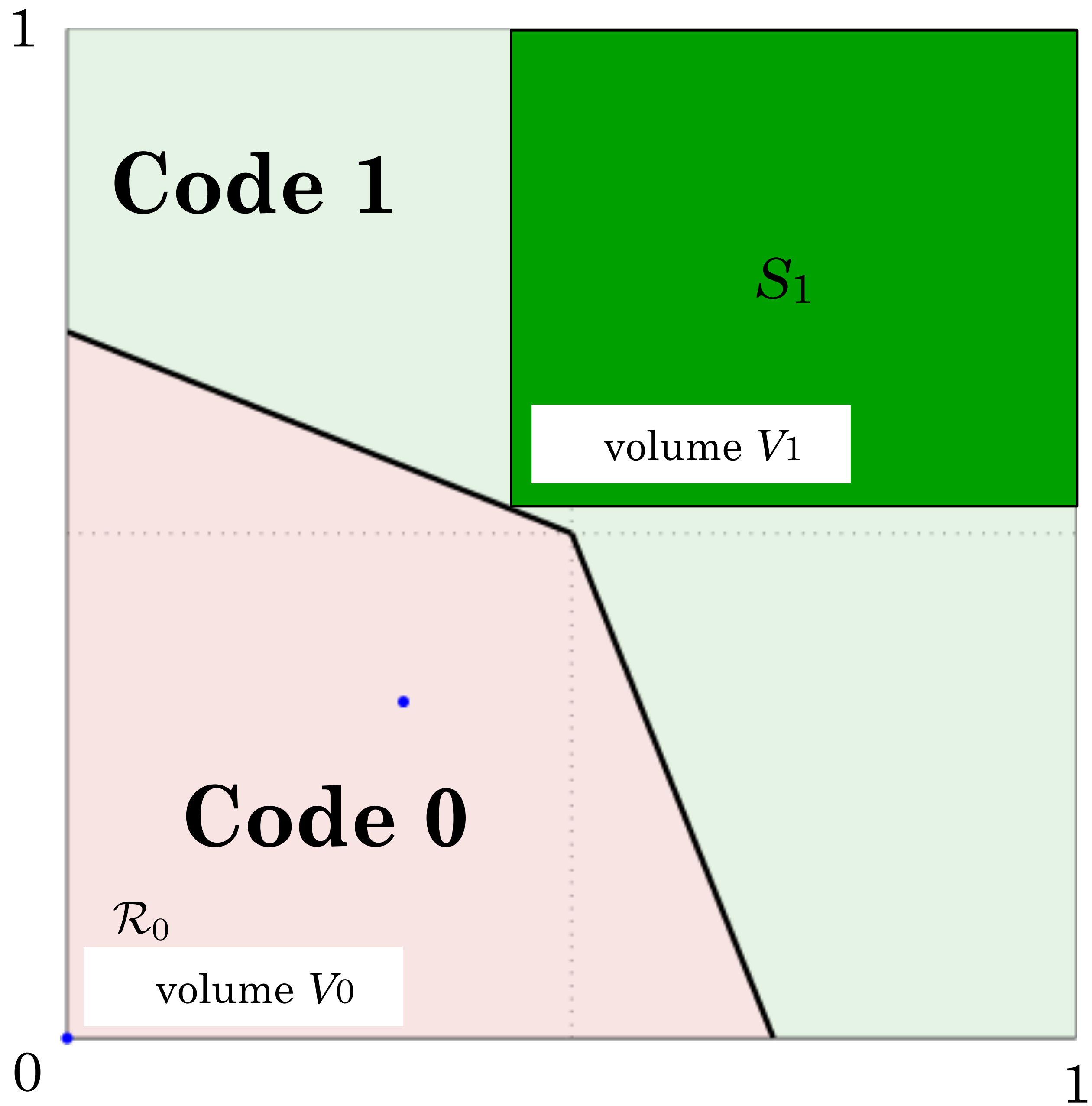
$$\lim_{n\to\infty} G(n\text{-sphere}) = \frac{1}{2\pi e}$$

$$G(\text{cube}) = \frac{1}{12}$$

Continuous Approximation — WOM Codes

Continuous Approximation:

$$\text{number of codewords} \quad |\mathcal{C}| \quad \approx \quad \frac{1}{\underbrace{V(\Lambda)}_{\text{volume of Voronoi cell}}}$$



Define volumes:

- Entire space: 1
- $\mathcal{R}_0$ Volume $= V_0$
- S_1 Volume $= V_1$

Continuous Approximation:

$$|\mathcal{C}_i| \quad \approx \quad \frac{V_i}{V(\Lambda)}$$

Normalized rate:

$$\widetilde{R}_i \quad = \quad 1 - \frac{\log_2 V_i}{\log_2 V(\Lambda)}$$

Separation of "Coding Penalty" and "Shaping Penalty"

Normalized rate:

$$\widetilde{R}_i = 1 - \underbrace{\frac{\log_2 V_i}{\log_2 V(\Lambda)}}_{\text{penalty } \gamma}$$

Penalty "gap from ideal:"

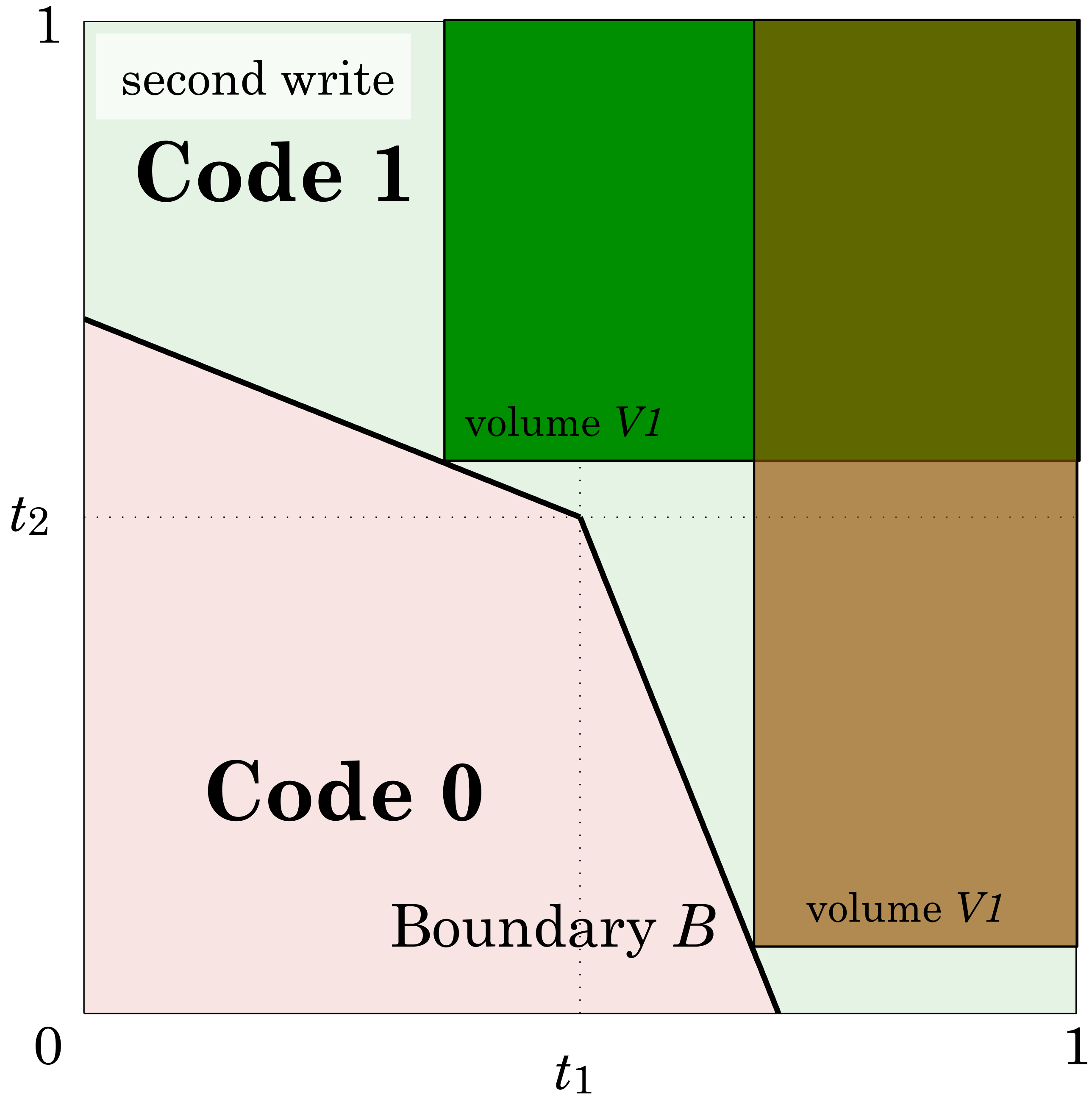
$$\begin{aligned} \gamma &= \frac{2}{n}\log_2 V_i \cdot \frac{1}{\log_2 V(\Lambda)^{2/n}} \\ \gamma &= \gamma_{\text{shape}} \cdot \gamma_{\text{code}} \end{aligned}$$

The rate penalty γ :

- γ_{shape} : V_i is a volume — depends on the "shaping" region
- γ_{code} : $V(\Lambda)$ is the Voronoi cell volume — depends on lattice/code

Can separate the rate penalty into a "shaping penalty" and "coding penalty"

Maximizing the Rate: B is a Hyperbola



$V_1(x)$: volume of space from x

Hypothesis For any $x \in B$, selecting $V(x)$ equal to a constant will maximize the rate.

For any point on B, the volume V_1 should be constant:

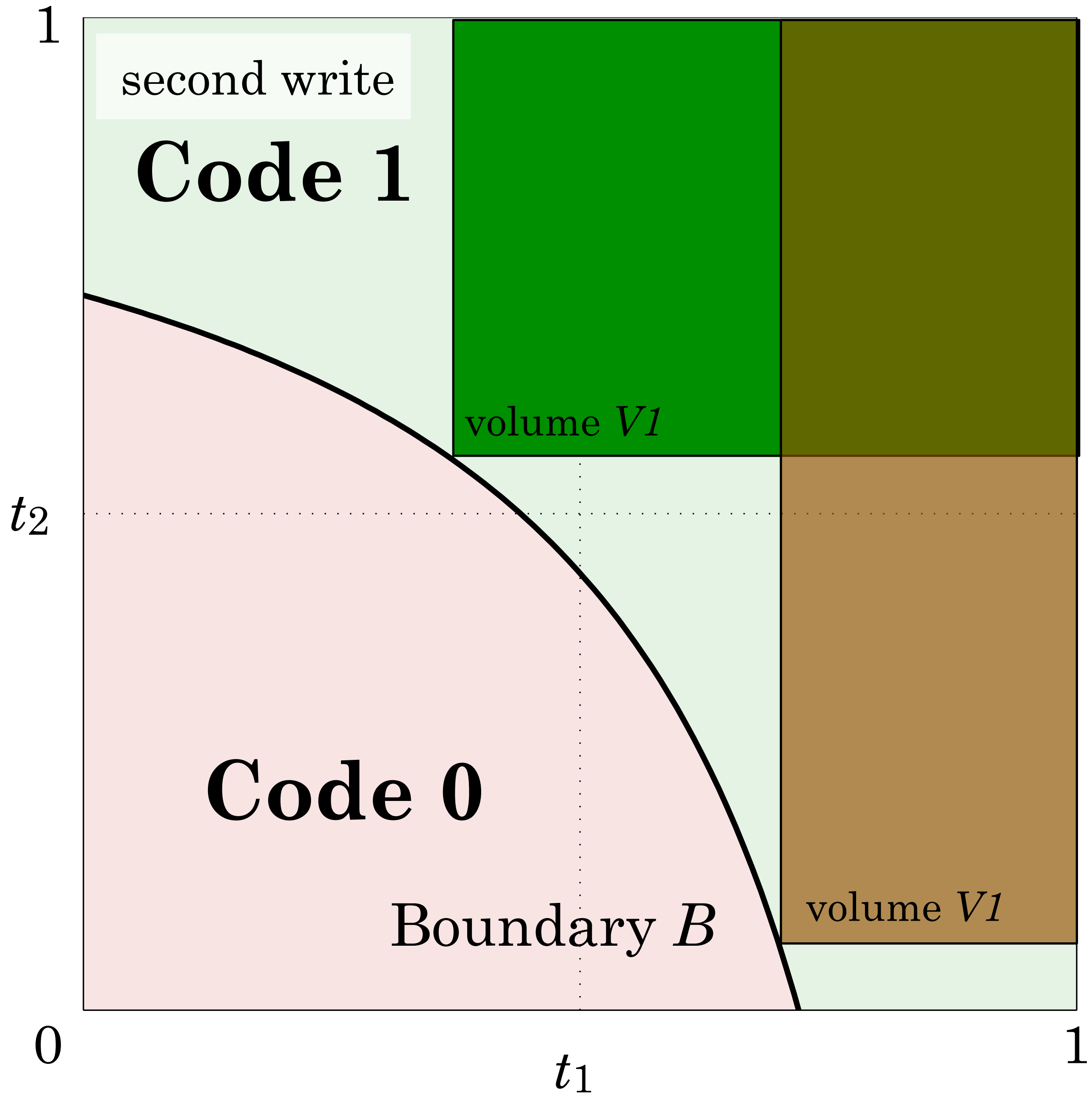
$$V_1 = (1 - t_1)(1 - t_2)$$

and in n dimensions:

$$V_1 = \prod_{i=1}^{n} (1 - t_i)$$

So, B is a **hyperbola**. We have a **hyperbolic shaping region**.

Maximizing the Rate: B is a Hyperbola



$V_1(x)$: volume of space from x

Hypothesis For any $x \in B$, selecting $V(x)$ equal to a constant will maximize the rate.

For any point on B, the volume V_1 should be constant:

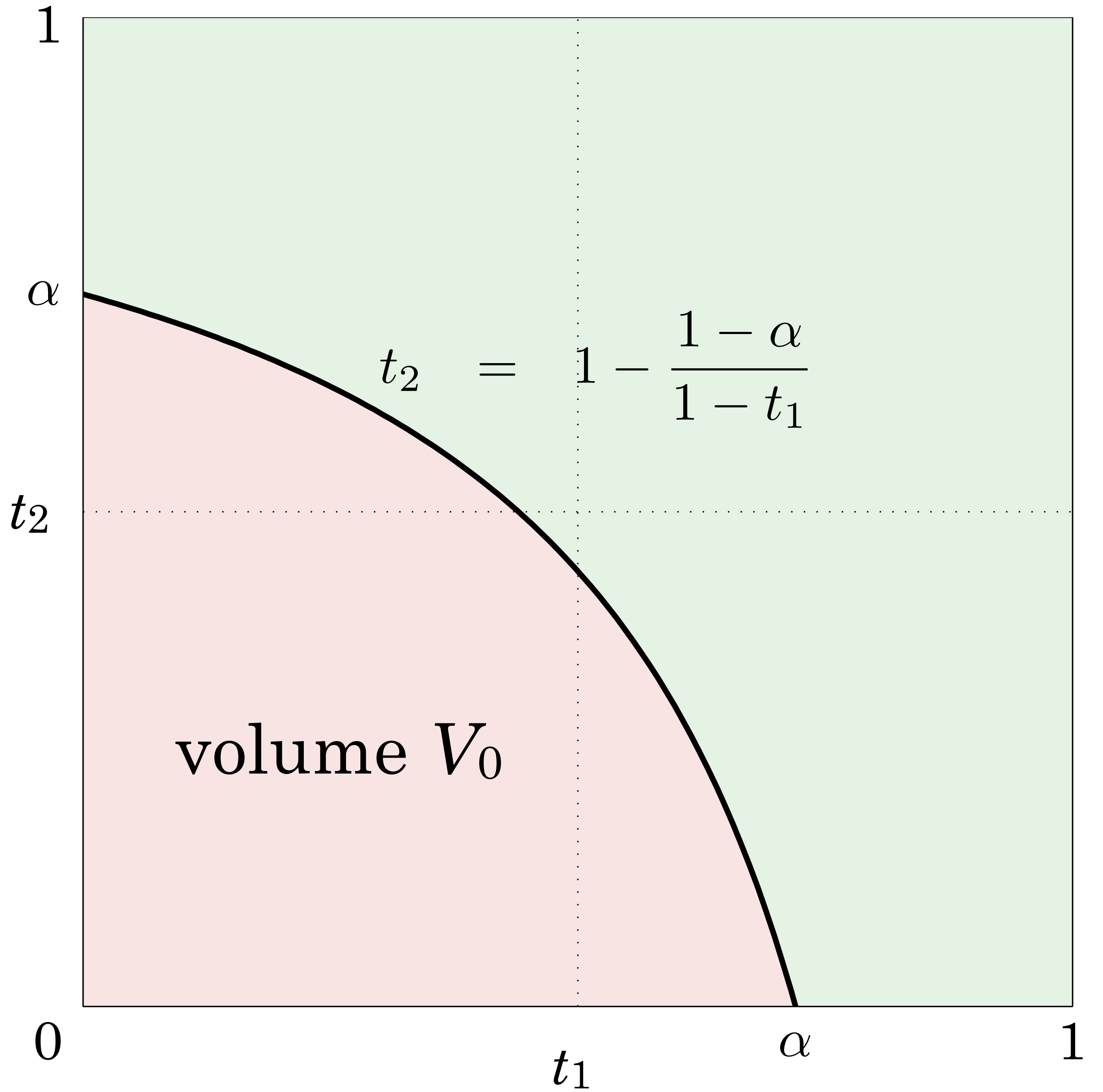
$$V_1 = (1 - t_1)(1 - t_2)$$

and in n dimensions:

$$V_1 = \prod_{i=1}^{n} (1 - t_i)$$

So, B is a **hyperbola**. We have a **hyperbolic shaping region**.

Computation of Volume V0



We can calculate the volume (and thus the rate).

For $n = 2$:

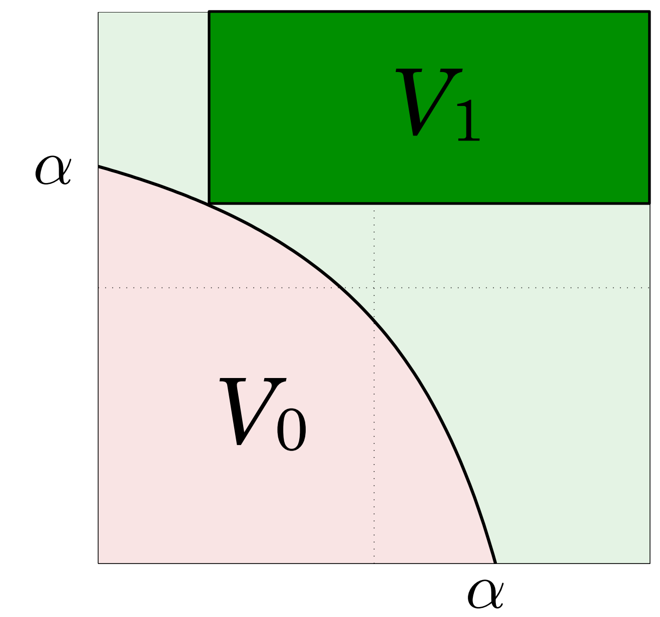
$$V_0 = 1 - (1-\alpha) + (1-\alpha)\log(1-\alpha)$$

For arbitrary n:

$$V_0 = 1 - e^{-z} \sum_{m=0}^{n-1} \frac{z^m}{m!},$$

where $z = -\log(1-\alpha)$

Equal Rates
Upper bound on shaping penalty



$z = -\log(1-\alpha)$

Assume equal rates for first and second writes:

$$
\begin{aligned}
V_0 &= V_1 \\
1 - e^{-z} \sum_{m=0}^{n} \frac{z^m}{m!} &= e^z - 1
\end{aligned}
$$

The solution z^* can only be found numerically. But can upper bound:

$$z^* \leq (n!)^{\frac{1}{n}}$$

And apply a Stirling-like bound:

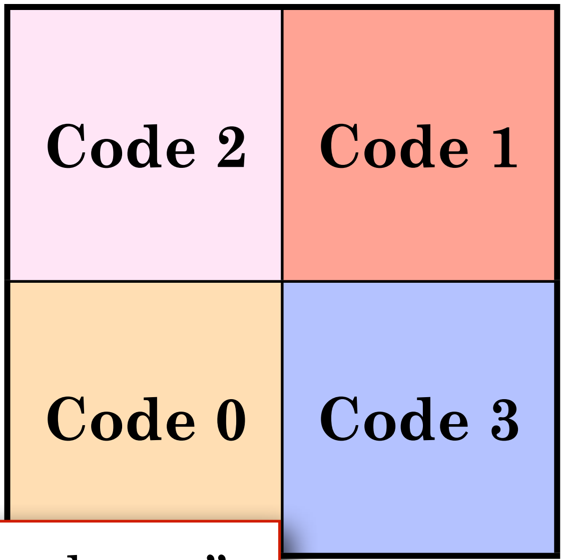
$$z^* \leq (ne)^{\frac{1}{n}} \frac{n}{e}$$

Highly preliminary! Recall:

$$\widetilde{R}_i = 1 - \gamma_{\text{shape}} \cdot \gamma_{\text{code}}$$

Now, the "shaping rate penalty" γ_{shape}:

$$
\begin{aligned}
\gamma_{\text{shape}} &= \frac{2}{n} \log_2 V_1 \leq \frac{2}{e \ln 2} (ne)^{\frac{1}{n}} \\
\lim_{n \to \infty} \gamma_{\text{shape}} &\leq \frac{2}{e \ln 2} \approx 1.0615
\end{aligned}
$$



Comparison with Cubic Scheme

$$\text{Gain} \quad = \quad \frac{1}{n}\log_2 \frac{V_0}{|\det G|} - \frac{1}{n}\log_2 \frac{(1/2)^n}{|\det G|}$$

Rate of hyperbolic shaping bits/dimension

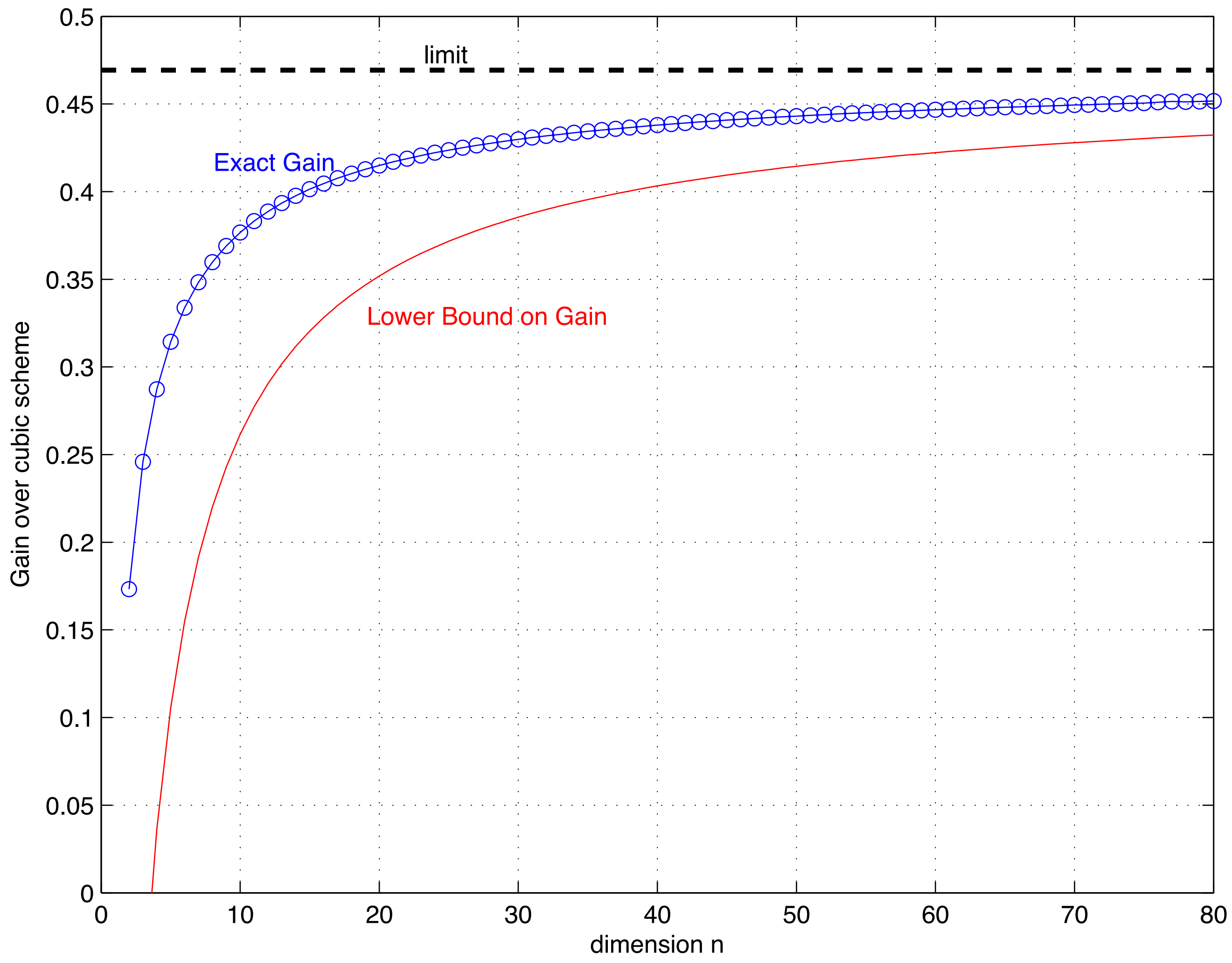
Rate of "cubic scheme" bits/dimension

Lower bound:

$$\text{Gain} \quad \geq \quad 1 - \frac{(ne)^{\frac{1}{n}}}{e \log 2}$$

And in the case $n \to \infty$:

$$\lim_{n\to\infty} \text{Gain} \quad \geq \quad 1 - \frac{1}{e\log 2} \approx 0.4693$$



Conclusion

Considered lattice-based construction of WOM codes for flash memories

- Used a "continuous approximation" similar to coding for AWGN channel
- converted a discrete problem into a continuous problem

As a result:

- Separation of the "shaping penalty" and "coding penalty"
- For two writes, shaping region is a **hyperbola** in n dimensions
- asymptotic shaping penalty is $\dfrac{2}{e \ln 2}$
- **asymptotic gain of 0.4693 bit/dimension** (for equal rates)

Discussion

Did not show the existence of a specific mapping

- Mapping from information to Code 1 must satisfy conditions