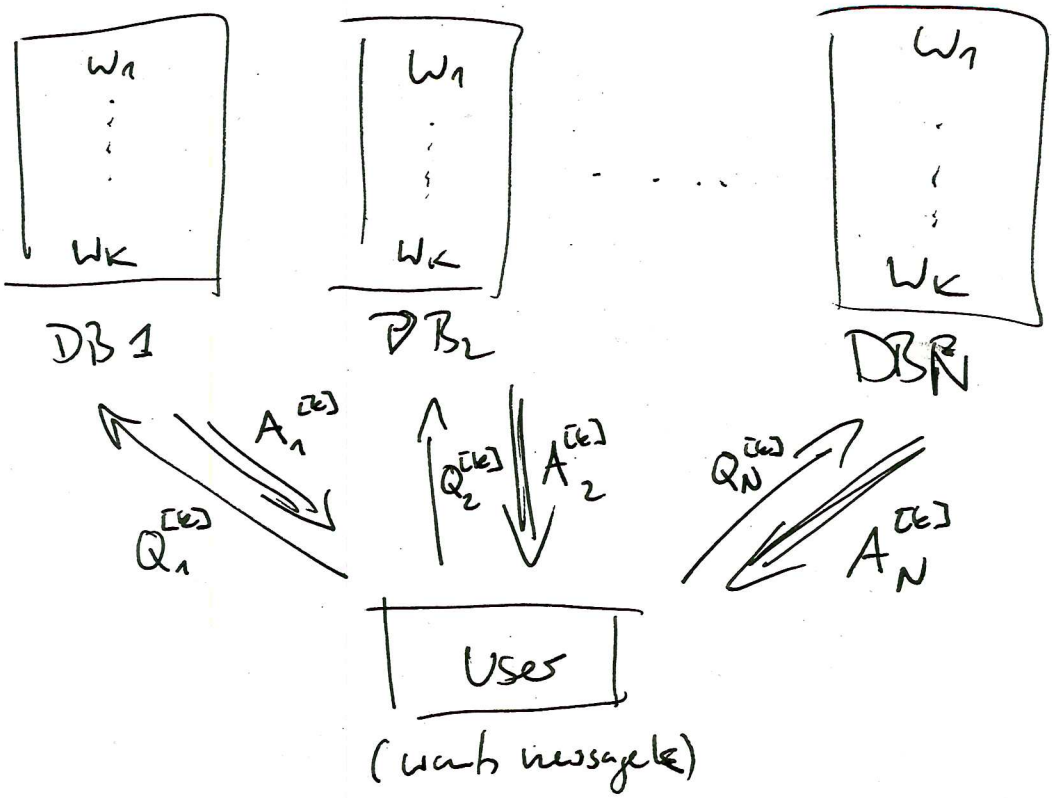
Multi-Server PIR



$$C = \frac{\text{Capacity of Multi-Server PIR}}{} = \frac{\#\text{ of useful bits}}{\text{Total \# of downloaded bits}}$$

Theorem: Capacity

$$C = \frac{1}{1 + \frac{1}{N} + \frac{1}{N^2} + \dots + \frac{1}{N^{K-1}}} > \frac{1}{K}$$

$$= \frac{1}{1 - \left(\frac{1}{N}\right)^K}$$

↑ (download-equally trivial scheme)

- Idea: Use coding to reduce communication overhead

Warmup Example:

$N=2$ databases $K=2$ messages

Message 1 $\Rightarrow A=\{a_1, a_2, a_3, a_4\}$

Message 2 $\Rightarrow B=\{b_1, b_2, b_3, b_4\}$

$L=4$ bits (size of each message)

DB1: $a_1\ a_2\ a_3\ a_4$ / $b_1\ b_2\ b_3\ b_4$

DB2: $a_1\ a_2\ a_3\ a_4$ / $b_1\ b_2\ b_3\ b_4$

User wants A: from DB1 a_1, b_1, a_3+b_2; from DB2 a_2, b_2, a_4+b_1

download (a_1, b_1, a_3+b_2) from DB1
download (a_2, b_2, a_4+b_1) from DB2
$\Rightarrow$ recover $a_1\ a_2\ a_3\ a_4$

use of complementary side-information!

User wants B: a_1, b_1, a_3+b_2 — same 3 bits from DB1; a_3, b_3, a_1+b_4 — 3 different bits downloaded from DB2

$\Rightarrow$ recovers $b_1\ b_2\ b_3\ b_4$

$$C \geq \frac{4}{6} = \frac{2}{3}$$

(4 $\sim$ # of desired bits; 6 $\rightarrow$ # of downloaded bits)

we will later show $C=\frac{2}{2}$

Example 2 $N = 2$ databases

$K = 3$ messages

Message $A = \{a_1, a_2, a_3, \dots, a_8\}$

Message $B = \{b_1, b_2, b_3, \dots, b_8\}$

Message $C = \{c_1, c_2, c_3, \dots, c_8\}$

(more generally size of message $= N^K$)

($\hat{=} 2^3 = 8$ bits in this example)

DB 1

$a_1 \dots a_8$
$b_1 \dots b_8$
$c_1 \dots c_8$

DB 2

$a_1 \dots a_8$
$b_1 \dots b_8$
$c_1 \dots c_8$

use code A

$\downarrow$

enforce symmetry across messages: a_1, b_1, c_1

$\downarrow$

enforce symmetry across messages: a_2, b_2, c_2

use (b_2, c_2) as side info from DB 2: $a_3 + b_2$, $a_4 + c_2$

$\downarrow$

(b_1, c_1) side-info from DB 1: $a_5 + b_1$, $a_6 + c_1$

enforce symmetry: $b_3 + c_3$ $\leftarrow$ enforce symmetry $\leftarrow$ $b_4 + c_4$

$\downarrow$

$a_7 + b_4 + c_4$ SI from DB 2

$a_8 + b_3 + c_3$ SI from DB 1

$\Rightarrow$ user can recover $A = \{a_1 \dots a_8\}$!

$$C \geq \frac{\#\text{useful bits}}{\#\text{downloaded bits}} = \frac{8}{7+7} = \frac{8}{14} = \frac{4}{7}$$

(check:

$$\frac{1}{1 + \frac{1}{N} + \dots + \frac{1}{N^{K-1}}} = \frac{1}{1 + \frac{1}{2} + \frac{1}{2^2}} = \frac{1}{1 + \frac{1}{2} + \frac{1}{4}} = \frac{4}{7} \text{ !})$$

Example 3 $N = 3$ DBs, $K = 3$ messages

$A = \{a_1, a_2, \dots, a_{27}\}$

$B = \{b_1, b_2, \dots, b_{27}\}$

$C = \{c_1, c_2, \dots, c_{27}\}$

$N^K = 3^3 = 27$ bits

DB1	DB2	DB3
a_1	a_2	a_3
b_1	b_2	b_3
c_1	c_2	c_3

$a_4 + b_2$	$a_8 + b_1$	$a_{12} + b_1$
$a_5 + c_2$	$a_9 + c_1$	$a_{13} + c_1$
$a_6 + b_3$	$a_{10} + b_3$	$a_{14} + b_2$
$a_7 + c_3$	$a_{11} + c_3$	$a_{15} + c_2$
SI from $DB_3(2,3)$	SI from $DB_2(1,3)$	SI from $DB_1(1,2)$
$b_4 + c_4$	$b_6 + c_6$	$b_8 + c_8$
$b_5 + c_5$	$b_7 + c_7$	$b_9 + c_9$
		enforce symmetry

$a_{16} + b_6 + c_6$	$a_{20} + b_4 + c_4$	$a_{24} + b_4 + c_4$
$a_{17} + b_7 + c_7$	$a_{21} + b_5 + c_5$	$a_{25} + b_5 + c_5$
$a_{18} + b_8 + c_8$	$a_{22} + b_8 + c_8$	$a_{26} + b_6 + c_6$
$a_{19} + b_9 + c_9$	$a_{23} + b_9 + c_9$	$a_{27} + b_7 + c_7$
		exploit SI from other DBs

$$C = \frac{27}{3 \times [3+6+4]} = \frac{9}{13}$$

Check:

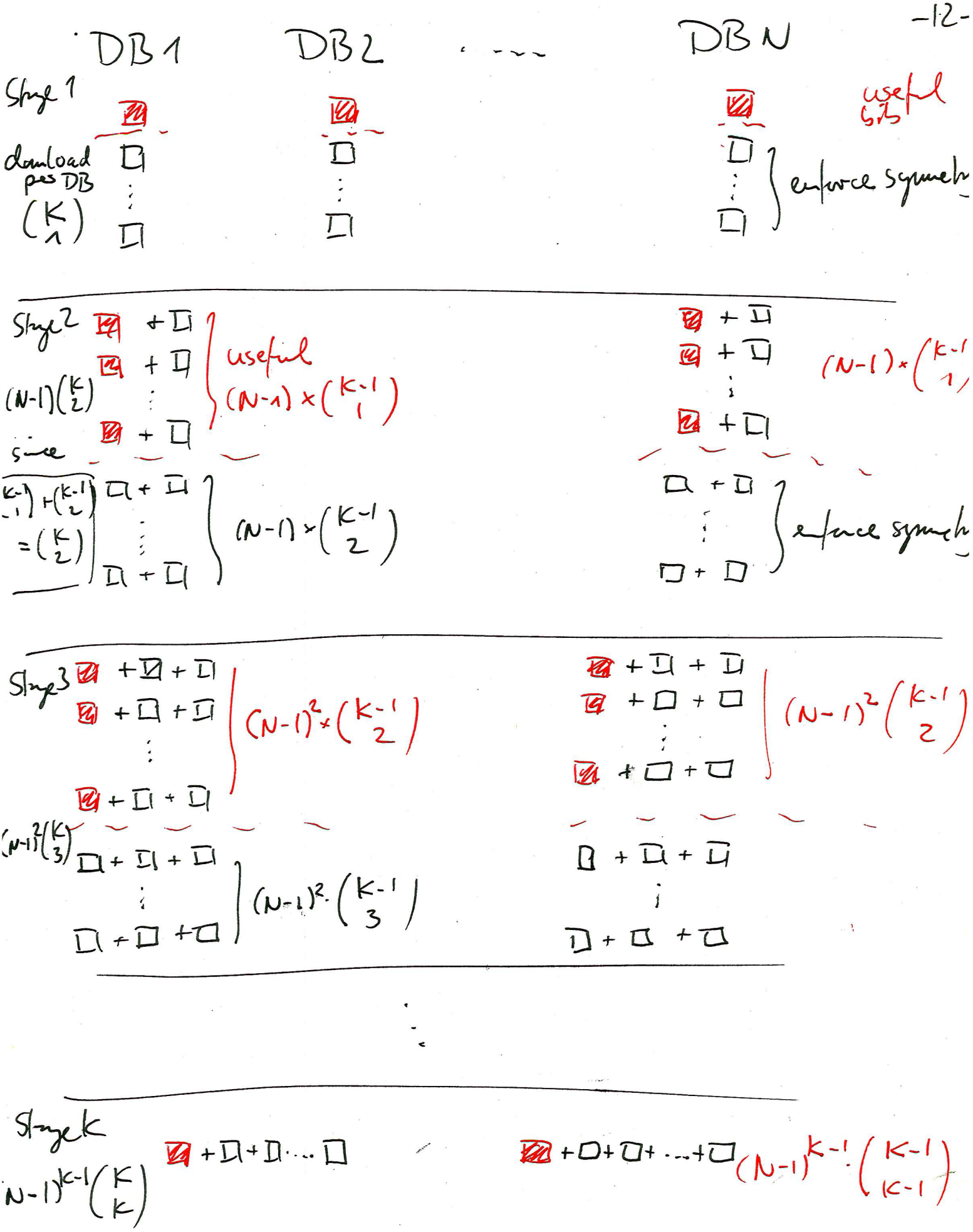
$$\frac{1}{1 + \frac{1}{N} + \dots \frac{1}{N^{K-1}}} = \frac{1}{1 + \frac{1}{3} + \frac{1}{3^2}} = \frac{1}{1 + \frac{1}{3} + \frac{1}{9}} = \frac{9}{13} !$$

General Scheme:

$$C = \frac{1}{1 + \frac{1}{N} + \frac{1}{N^2} + \cdots + \frac{1}{N^{K-1}}}$$

- We have K messages
- Split each message into N^K bits (the reason for this will become clear)

[Note: this is not necessarily fundamental & an artifact of the construction, i.e. it is possible to achieve capacity with fewer bits/message. The general answer is however unknown ....


DB 1
DB 2
DB N
Stage 1
download per DB
$\binom{K}{1}$
useful GTs
enforce synch
Stage 2
useful
$(N-1)\times\binom{K-1}{1}$
since
$\binom{K-1}{1}+\binom{K-1}{2}=\binom{K}{2}$
$(N-1)\binom{K}{2}$
$(N-1)\times\binom{K-1}{2}$
$(N-1)\times\binom{K-1}{1}$
enforce synch
Stage 3
$(N-1)^2\times\binom{K-1}{2}$
$(N-1)^2\binom{K}{3}$
$(N-1)^2\cdot\binom{K-1}{3}$
$(N-1)^2\binom{K-1}{2}$
Stage k
$(N-1)^{k-1}\binom{K}{k}$
$(N-1)^{k-1}\cdot\binom{K-1}{k-1}$

$$C \geq \frac{\#\text{useful bits}}{\#\text{downloaded bits}}$$

$$= \frac{N \times (\#\text{useful bits per DB})}{N \times (\#\text{downloaded bits per DB})}$$

$$= \frac{\binom{K-1}{0} + (N-1)\binom{K-1}{1} + (N-1)^2\binom{K-1}{2} + \cdots + (N-1)^{K-1}\binom{K-1}{K-1}}{\binom{K}{1} + (N-1)\binom{K}{2} + \cdots (N-1)^2\binom{K}{3} + \cdots + (N-1)^{K-1}\binom{K}{K}}$$

Binomial identity: $(x+y)^{K-1} = \sum_{k=0}^{K-1} x^k \cdot y^{(K-1-k)} \binom{K-1}{k}$

Numerator:

$$\binom{K-1}{0} + (N-1)\binom{K-1}{1} \cdots = \sum_{k=0}^{K-1} (N-1)^k \cdot 1^{(K-1-k)} \binom{K-1}{k}$$

$$= \left((N-1)+1\right)^{K-1}$$

$$= N^{K-1}$$

(Remark)
Total size of each file: $N \times N^{K-1} = N^K$

Denominator:

$$\binom{K}{1} + (N-1)\binom{K}{2} + \cdots = \frac{1}{N-1}\left[(N-1)\binom{K}{1} + \cdots + (N-1)^K\binom{K}{K}\right]$$

$$= \frac{1}{N-1} \times \left[\sum_{k=0}^{K} (N-1)^k \binom{K}{k} - 1\right]$$

$$= \frac{1}{N-1} \times \left[(N-1+1)^K - 1\right] = \frac{N^K - 1}{(N-1)}$$

$$C \geq \frac{N^{K-1}}{\frac{N^K - 1}{N-1}}$$

$$= \frac{N^{K-1}}{\frac{N^K\left[1 - \frac{1}{N^K}\right]}{N-1}}$$

$$= \frac{1}{\frac{N\left(1 - \frac{1}{N^K}\right)}{N-1}} = \frac{1}{\frac{\left(1 - \frac{1}{N^K}\right)}{1 - \frac{1}{N}}}$$

Geometric series:

$$S = 1 + r + r^2 + \ldots r^{k-1}$$

$$rS = r + r^2 + r^3 + \ldots r^k$$

$$S - rS = 1 - r^k$$

$$\Rightarrow S = \frac{1 - r^k}{1 - r} \quad \text{set } r = \frac{1}{N}$$

$$\Rightarrow 1 + \frac{1}{N} + \ldots + \frac{1}{N^{k-1}} = \frac{1 - \frac{1}{N^k}}{1 - \frac{1}{N}}$$

$$\Rightarrow C \geq \frac{1}{\left(\frac{1 - \frac{1}{N^k}}{1 - \frac{1}{N}}\right)} = \frac{1}{1 + \frac{1}{N} + \ldots + \frac{1}{N^{k-1}}}$$

PZ

Converse $C \leq \frac{1}{1+\frac{1}{N}+\cdots+\frac{1}{N^{K-1}}}$

$$C = \max\left(\frac{H(A_1^{[c]}) + H(A_2^{[c]}) + \cdots + H(A_N^{[c]})}{L} \right)^{-1}$$

D = downloaded data (in bits)
L = useful # of bits

Suppose user requests message #1.
$\Rightarrow$ Servers send: $A_1^{[1]}, A_2^{[1]}, \ldots, A_N^{[1]} \Rightarrow H(W_1 | A_1^{[1]}, A_2^{[1]}, \ldots, A_N^{[1]})_2$

Consider:

$$H(W_1, A_1^{[1]}, A_2^{[1]}, \ldots, A_N^{[1]}) = \underbrace{H(W_1)}_{=L} + H(A_1^{[1]}, A_2^{[1]}, \ldots, A_N^{[1]})$$
$$= L + H(A_1^{[1]}, A_2^{[1]}, \ldots, A_N^{[1]})$$

Expanding it another way:

$$H(W_1, A_1^{[1]}, A_2^{[1]}, \ldots, A_N^{[1]}) = H(A_1^{[1]}, \ldots, A_N^{[1]}) + \underbrace{H(W_1 | A_1^{[1]} \ldots A_N^{[1]})}_{=0}$$
$$\leq \sum_{i=1}^{N} H(A_i^{[1]})$$
$$\leq D$$

$\Rightarrow$ ~~data~~ we have

$$D \geq L + H(A_1^{[1]} \ldots A_N^{[1]} | W_1)$$

- So far we have not invoked the privacy constraint.
- Consides

$$H(A_1^{[1]} \dots A_N^{[1]} | W_1)$$

$$= H(A_1^{[1]} | W_1) + H(A_2^{[1]} \dots A_N^{[1]} | W_1 A_1^{[1]})$$

$$\geq H(A_1^{[1]} | W_1)$$

$$= H(A_1^{[2]} | W_1)$$

↓ by privacy constraint for DB 2.

similarly

$$H(A_1^{[1]} \dots A_N^{[1]} | W_1) \geq H(A_2^{[2]} | W_1)$$

$$\vdots \qquad \vdots$$

$$\geq H(A_N^{[2]} | W_1)$$

} N inequalities

$$N \cdot H(A_1^{[1]} \dots A_N^{[1]} | W_1) \geq H(A_1^{[2]} | W_1) + \dots + H(A_N^{[2]} | W_1)$$

$$\geq H(A_1^{[2]} \dots A_N^{[2]} | W_1)$$

$$= H(W_2 A_1^{[2]} \dots A_N^{[2]} | W_1)$$

$$- \underbrace{H(W_2 | A_1^{[2]} \dots A_N^{[2]} W_1)}_{=0 \text{ decodability for } W_2}$$

$$= H(W_2 | W_1) + H(A_1^{[2]} \dots A_N^{[2]} | W_1 W_2)$$

$$= H(W_2) + \text{——} \quad \text{——}$$

$$= L + H(A_1^{[2]} \dots A_N^{[2]} | W_1 W_2)$$

$$\Rightarrow H(A_1^{[1]} \dots A_N^{[1]} \mid W_1) \geq \frac{L}{N} + \frac{1}{N} \times \left\{ H(A_1^{[2]} \dots A_N^{[2]} \mid W_1 W_2) \right\}$$

· We can now apply the same "trick" a

$$H(A_1^{[2]} \dots A_N^{[2]} \mid W_1 W_2) \geq \frac{L}{N} + \frac{1}{N} \times \left\{ H(A_1^{[3]} \dots A_N^{[3]} \mid W_1 W_2 W_3) \right\}$$

$$\vdots$$

$$H(A_1^{[K-1]} \dots A_N^{[K-1]} \mid W_1 \dots W_{K-1}) \geq \frac{L}{N} + \frac{1}{N} \times \underbrace{\left\{ H(A_1^{[K]} \dots A_N^{[K]} \mid W_1 \dots W_K) \right\}}_{=0}$$

$\Rightarrow$

$$D \geq L + H(A_1^{[1]} \dots A_N^{[1]} \mid W_1)$$

$$\geq L + \frac{L}{N} + \frac{1}{N} H(A_1^{[2]} \dots A_N^{[2]} \mid W_1 W_2)$$

$$\geq L + \frac{L}{N} + \frac{L}{N^2} + \frac{1}{N^2} H(A_1^{[3]} \dots A_N^{[3]} \mid W_1 W_2 W_3)$$

$$\vdots$$

$$\geq L + \frac{L}{N} + \frac{L}{N^2} + \frac{L}{N^3} \dots + \frac{L}{N^{K-1}}$$

$$\Rightarrow \frac{D}{L} \geq 1 + \frac{1}{N} + \dots + \frac{1}{N^{K-1}}$$

$$\Rightarrow \boxed{C = \frac{L}{D} \leq \frac{1}{1 + \frac{1}{N} + \dots + \frac{1}{N^{K-1}}}}$$

RG