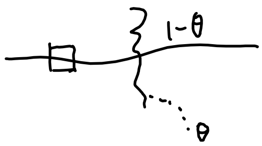
20240923 LT Code



Erasure Channel

LT Code Scheme

① There are packets • and encoded packets □

② Each encoded packet □ is a xor of packets ··•···

③ Each □ first select a degree d from degree distribution

Then uniformly choose •s to xor.

20240923 Lévy continuity theorem & Infinite Divisibility

Idea: convergence in distribution ⇐ pointwise convergence of char. functions

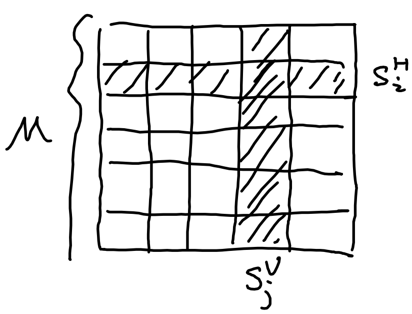
Stmt: $(X_n)_{n\in\mathbb{N}}$ $\quad \varphi_n(t) = \mathbb{E}(\exp(itX_n))$

$\forall t: \lim_{n\to\infty} \varphi_n(t) = \varphi(t)$

$\Rightarrow (X_n)_{n\in\mathbb{N}}$ converges in distribution to X

as long as $\lim_{x\to\infty} \sup_{n\in\mathbb{N}} \mathbb{P}\{|X_n| > x\} = 0$ (tightness)

Q: Given S, $(S_i^H)_i$ $(S_j^V)_j$ are independent randon vectors, what is μ?



$S = \sum_{i,j} M_{i,j}$ $\quad (M_{i,j})_{i,j} \overset{\text{i.i.d}}{\sim} \mu$

$S_i^H = \sum_j M_{i,j}$

$S_j^V = \sum_i M_{i,j}$

Take a look at Lévy process

20240923 Extend Research Topics

- More Applications
- Augment Input : like SeRF
- Restrict Input for Better Performance
- Change Your Machine : Distributed / Parallel

20240924 RECIPE Code Codebase

LT Code, but distributed

CPP: { Simulation, Search }

CS7260 Project

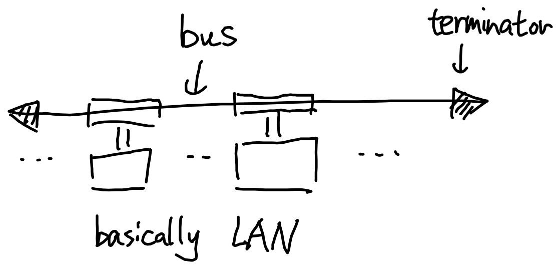
Python: Evaluation

20240924 Switching

Do something to actually unterstand internet!

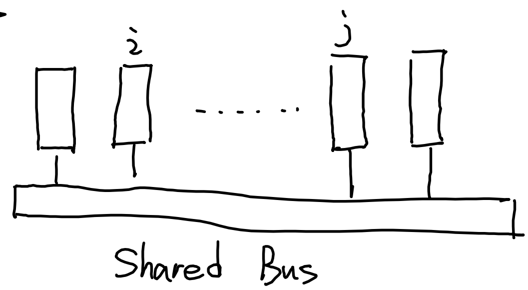
History of Routers

1984 | 1989

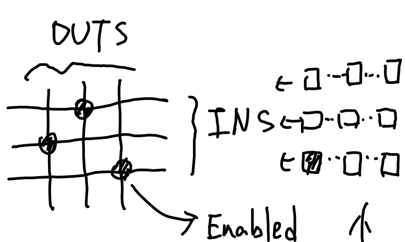


broadcast only

each end node performs selective pick-up



A pair of end nodes occupies the bus for a clock cycle.



schedule inputs and outputs

each enabled knot can transmit a packet per-clock cycle

each input port has a queue

If the first packet ▨ cannot go, the queue is stuck.

X_i is the output port of the packet at input port i

If $(X_i)_i$ is independent, the throughput is roughly $2-\sqrt{2}$.

VOQ: Virtual Output Queueing

Idea

"Take-a-ticket" Algo

Read the Book

20240924 Parallel Iterative Matching (PIM)

① An input port requests from all output ports to which the corresponding output if the VOQ is not empty.

② An output port randomly grants to one.

how?

③ An input port randomly accepts one.

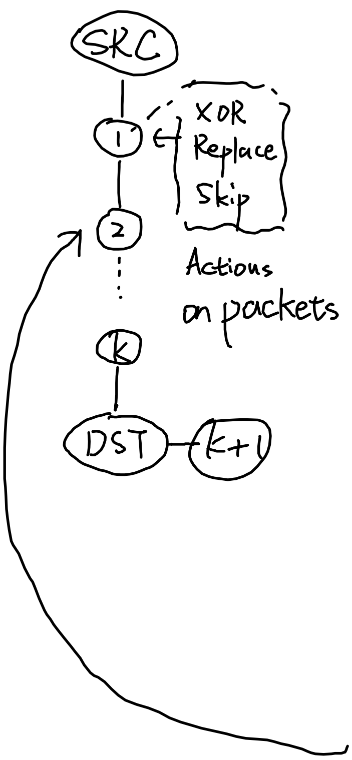
Tom Enderson

look up

O(lg n) iters

Bad: $O(n^2)$ requests for each iteration

20240925 Convert Path Tracing to Distribute LT Code



Assumption:

- Given SRC & DST, the path is always the same
- Each switch knows its distance from SRC

Goal:

- Get the list of switch IDs

Prev Work: PINT Code

- 1st code: reservoir sampling
- 2nd code: with probability p Xor

} randomly commit one of these actions

How to know the XOR set:

- The action at i is determined by packet content. Therefore, given the action table the decoder knows the action on each step.

$$q_i(d) = \{1, 2, 3\}$$
$$\{1, 3, 4\}$$

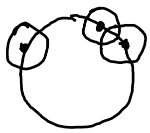
Check with Jingfan

Uniformity Condition:

$\{1\}$ . . . . $\{i-1\}$ $\{i\}$

$$q_i(\{i\}) = \sum_d r(i,d) \cdot \mu(i-1, d) = \sum_{d=1}^{i-1} \left(q_{i-1}(d) - q_i(d+1) - q_i(d)\right)$$

20240925 Approximate Data-normalized Softmax



Exponential decrease w.r.t. distance

Therefore, it should be ideal if we have an approximation for anything within $\log(1/\varepsilon)$ distance to any of these points.

Cut several layers and add values to random mesh

Idea : use more shapes

$O(D^2\sqrt{N})$

Independently maintain D versions of these structures.

20240925 Definition of ε-net on range space $(X, \mathcal{R}, A)$

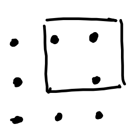
X : set of points, usually a subset of $\mathbb{R}^d$

$\mathcal{R}$: a family of subset of X

A : a finite subset of X

A set V is an ε-net of A if $\forall R \in \mathcal{R}$

$$\left| \frac{|A \cap r|}{|A|} - \frac{|V \cap r|}{|V|} \right| < \varepsilon$$

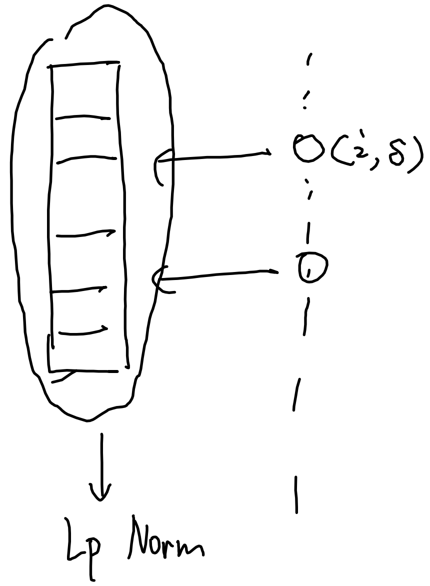


20240925 Cone and Hull

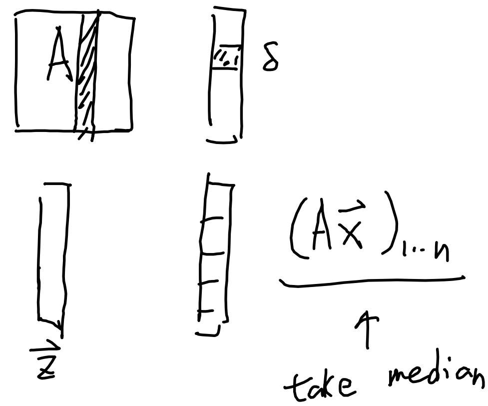
On a plane,

What am I doing here?

20240925 Universal Sketch



Presene a linear transformatio of the state.



20240925 LP Optimality via Duality

$$\begin{array}{ccc} \max\ c^T x & & \min\ b^T y \\ Ax \preceq b & \xleftrightarrow{\text{Dual}} & A^T y \succeq c \\ x \succeq 0 & & y \succeq 0 \end{array}$$

Weak LP Duality :

For any feasible solution y for dual LP $\qquad c^T x \le b^T y$

& ... x for original LP

Proof : $\sum_i x_i c_i \le \sum_j y_j \sum_i A_{ji} x_i \le \sum_j y_j b_j$

Strong LP Duality :

If optimal LP value for original LP is finite,

optimal dual LP value is the same.

Remark : If one side is infinite, then dual LP is infeasible.

Farka's Lemma

If $Ax = b$ and $x \ge 0$ is infeasible then

$\exists y : y^T A \ge 0$ and $y^T b < 0$

20240925 Chain of Thoughts

No big deal, just showed $O(n)$ circuits can be simulated by $O(n)$ inference step & $O(\ln(n))$ embedding size.

20240926 Ext. RECIPE

Dynamic information like queue size of each router.

20240926 Streaming ECC & Packet Swapping

If the queue of frames are head-blocked & orderly requirements are strict, the swapping matter.

20240926 Parallel Iterative Matching Cont. (iSLIP)

How to randomly select a packet?

Every output $\underline{\overset{o}{\text{port}}}$ has an grant pointer (G_o)

Every input $\underline{\overset{i}{}}$ . . . . . accept pointer (A_i)

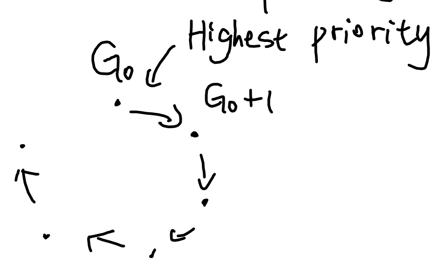
$G_o :=$ granted input port & accepted grant

$A_i :=$ ---

If $G_o = \emptyset$, use the last pointer. o.w. use $G_o + 1$.

↑ Only for the first iteration

This ensembles priority encoder.



20240926 SERENA for Switching

Sample & Compare Algorithm

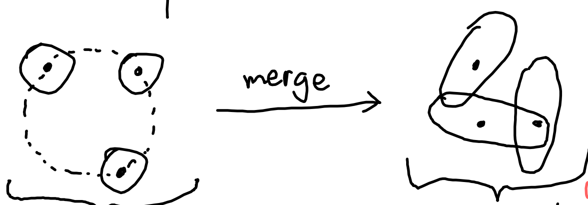
VOQ size as Weight $\Rightarrow$ Max Matching gives 100%

When VOQ change, find a random permutation.

If it is better then current matching, replace it.

20240926 Merging a Unimodal Mixture of Gaussians

Some Clean Up:



Unless: Gaussians are low dimensional

$N \times d$ for k

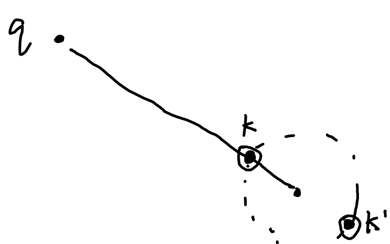
$N \times d$ for v

$f(N) \times \underline{d \times d}$ for k&v

$(c + f(N) \times d)$ for denominator

Other worries: very large $\|q\|_2$, very small denominator

However, the direction is almost certain!



20240926 Diffusion

Somewhat like Martingale

$X_1, \dots, X_n, \dots\dots X_{(\cdot)}$

$\mu(X_{n+1}, \dots, X_{(\cdot)} \mid X_1, \dots, X_n)$

20240929 GC-Free Dynamic Language

Ownership: In each instruction, an address is either:

1. owned
2. mutable borrow
3. immutable borrow

by one and only one symbol.

The earliest possible moment to drop a variable.

Drop cascades to fields.

20240929 Orthogonal & Far away

Merge two Gaussians

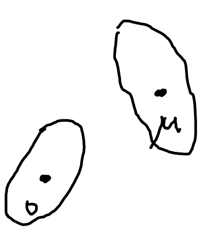
$$\lambda \mathcal{N}(x; \mu, \Sigma) + (1-\lambda)\mathcal{N}(x; \tilde{\mu}, \tilde{\Sigma})$$

$$\approx \mathcal{N}(x; \lambda\mu + (1-\lambda)\tilde{\mu}, \lambda\Sigma + (1-\lambda)\tilde{\Sigma} + \lambda(1-\lambda)(\mu-\tilde{\mu})(\mu-\tilde{\mu})^T)$$

The problem is how it goes when $\|x\|_2$ is large.

For simplicity, suppose $\tilde{\mu} = \vec{0}$ & $\tilde{\Sigma}$ is diagonal.

$$\frac{\lambda \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)/\sqrt{\cancel{(2\pi)^d}|\Sigma|} + (1-\lambda)\exp\left(-\frac{1}{2}x^T\tilde{\Sigma}^{-1}x\right)/\sqrt{\cancel{(2\pi)^d}|\tilde{\Sigma}|}}{\exp\left(-\frac{1}{2}(x-\lambda\mu)^T(\lambda\Sigma+(1-\lambda)\tilde{\Sigma})^{-1}(x-\lambda\mu)\right)/\sqrt{\cancel{(2\pi)^d}|\lambda\Sigma+(1-\lambda)\tilde{\Sigma}|}}$$



Too hard for hand calculation

20240929 Sherman-Morrison Update

$$(I + uv^T)^{-1} = I - \frac{uv^T}{1 + v^T w}$$

$$\mathbb{V}ar(X) = \mathbb{E}[\,\mathbb{E}[(X_I - \bar{\mu})(X_I - \bar{\mu})^T \mid I\,]]$$

$$\bar{\Sigma}_n = I \quad + \frac{1}{n}\sum_i (\mu_i - \bar{\mu}_n)(\mu_i - \bar{\mu}_n)^T$$

$$\bar{\mu}_n = \frac{1}{n}\sum_i \mu_i$$

$$\mathbb{E}[(X_i - \bar{\mu})(X_i - \bar{\mu})^T \mid I = i]$$
$$= \Sigma_i + (\mu_i - \bar{\mu})(\mu_i - \bar{\mu})^T$$

$$I + \sum_{i=1}^{n+1} (\mu_i - \bar{\mu}_{n+1})(\mu_i - \bar{\mu}_{n+1})^T = \sum_{i=1}^{n+1} (\mu_i - \bar{\mu}_n)(\mu_i - \bar{\mu}_n)^T + I$$
$$+ \sum_{i=1}^{n+1} (\bar{\mu}_n - \bar{\mu}_{n+1})(\mu_i - \bar{\mu}_n)^T$$
$$+ \sum_{i=1}^{n+1} (\mu_i - \bar{\mu}_n)(\bar{\mu}_n - \bar{\mu}_{n+1})^T$$
$$+ \sum_{i=1}^{n+1} (\bar{\mu}_n - \bar{\mu}_{n+1})(\bar{\mu}_n - \bar{\mu}_{n+1})^T$$
$$= (\bar{\mu}_n - \bar{\mu}_{n+1})\mu_{n+1}^T + \mu_{n+1}(\bar{\mu}_n - \bar{\mu}_{n+1})^T + \bar{\Sigma}_n$$
$$= \frac{1}{n+1}(\bar{\mu}_n - \mu_{n+1})\mu_{n+1}^T + \frac{1}{n+1}\mu_{n+1}(\bar{\mu}_n - \mu_{n+1})^T + \bar{\Sigma}_n$$
$$\left(\frac{1}{n+1}\right) \quad (\bar{\mu}_n - \mu_{n+1})(\bar{\mu}_n - \mu_{n+1})^T$$
$$= \frac{1}{n+1}\bar{\mu}_n\bar{\mu}_n^T - \frac{1}{n+1}\mu_{n+1}\mu_{n+1}^T + \bar{\Sigma}_n$$

$$\bar{\Sigma}_{n+1} - \bar{\Sigma}_n = \frac{1}{n+1}\bar{\mu}_n\bar{\mu}_n^T - \frac{1}{n+1}\mu_{n+1}\mu_{n+1}^T$$

$$\frac{1}{4}(1-0)(1-0) + \frac{1}{2} + \frac{1}{2}$$

20240929 Far-away Points Example

$$f(x) = \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) + \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-1)^2}{2}\right)$$

$$g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\frac{1}{2})^2}{2\sigma^2}\right) \qquad \sigma^2 = 1 + \frac{1}{2} \cdot \frac{1}{2} (1-0)(1-0)$$