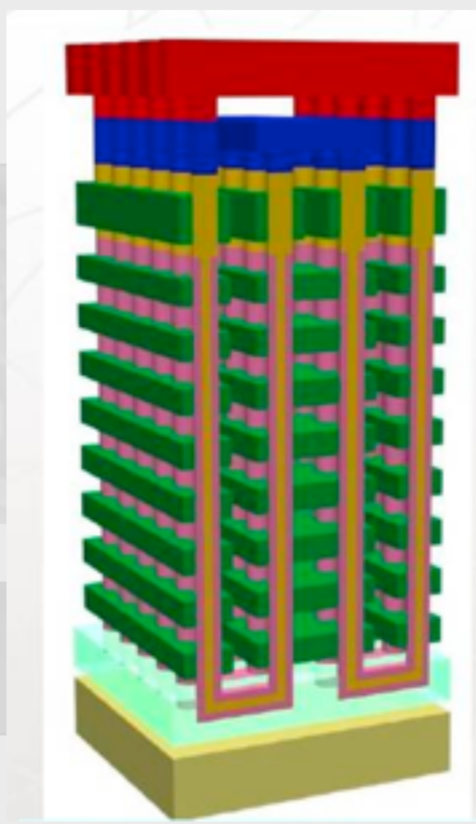
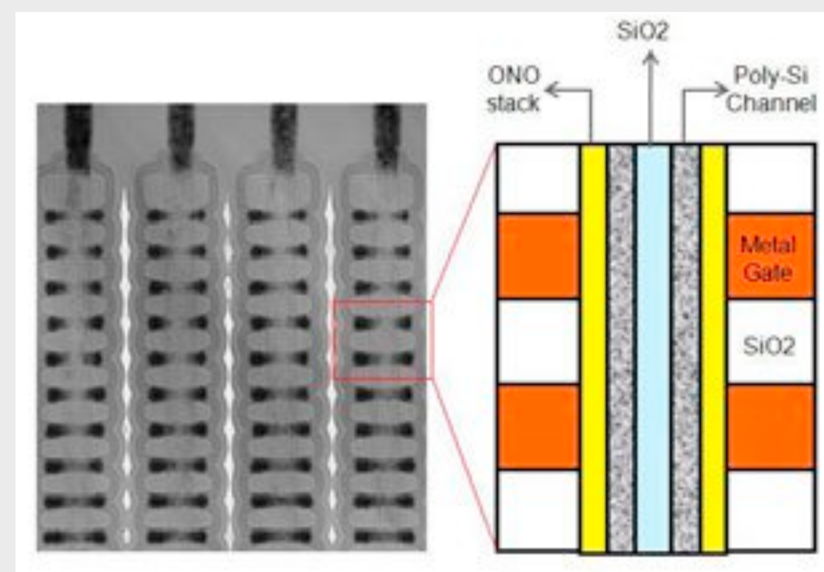
Flash Math - FTL Algorithms and Performance

Peter Desnoyers
Northeastern University

Why do we care?

Flash is on the way out, isn't it?





3D NAND

Why Math?

$$\frac{d\vec{\mu}(t)}{dt} = e_b +$$

$$A_{Greedy}(\alpha, r, f) \approx \frac{A_{LRW}((1+\frac{1}{2N_p})\alpha, r, f)}{1+\frac{1}{2N_p}}$$

Table IV. LRW vs. Windowed Greedy (simulated)

S_f	Greedy	LRW	Windowed Greedy	95% CI

Because we don't understand what's going on in the middle of our ~~systems~~

simulations

$$+m_{i+1} \sum_{j=1, j\neq b-(i+1)} p_{b-j}(\vec{m}) B_1(j, (i+1)/b\rho N)$$

parameters: U=50000, N_p=64, w=500.

$$A = \frac{N_p}{N_p - (X_0 - 1)}$$

$$+p_{i+1}(\vec{m}) \left(m_{i+1} - \frac{1}{N}\right) B_1(b-(i+1), (i+1)/b\rho N)$$

$$-m_i \sum_{j=1, j\neq b-i}^{b} p_{b-j}(\vec{m}) B_1(j, i/b\rho N)$$

$$-p_i(\vec{m}) \left(m_i - \frac{1}{N}\right) B_1(b-i, i/b\rho N) + o(1/N), \quad (6)$$

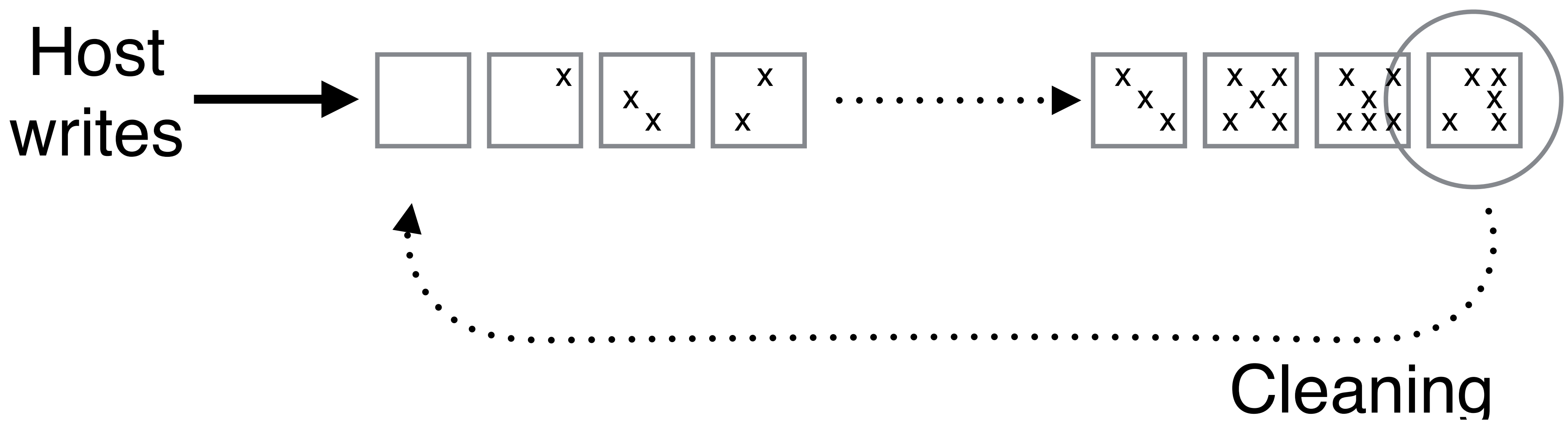
nipulation we can express this as:

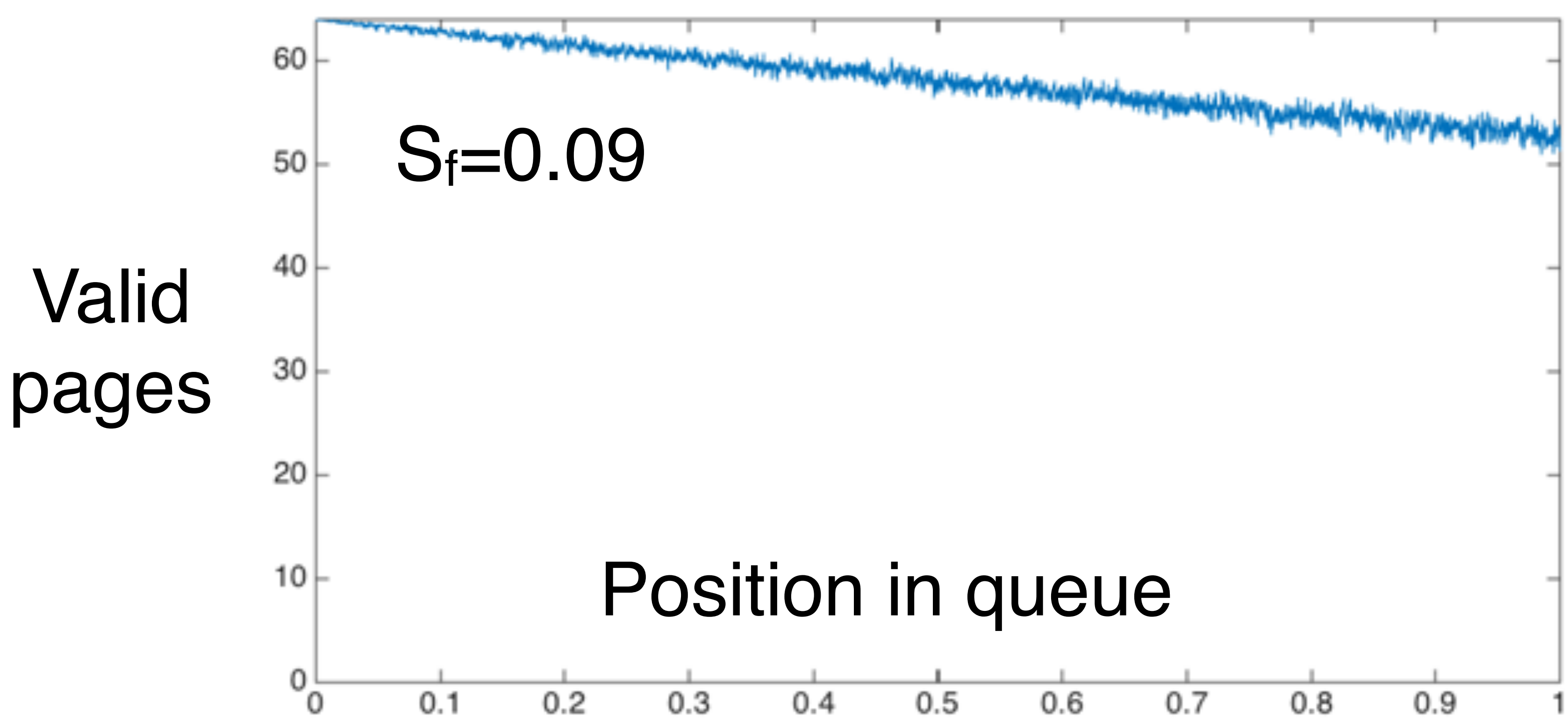
$$A = \frac{A_{LRW}(\alpha')}{1+\frac{1}{2N_p}}$$

Definitions

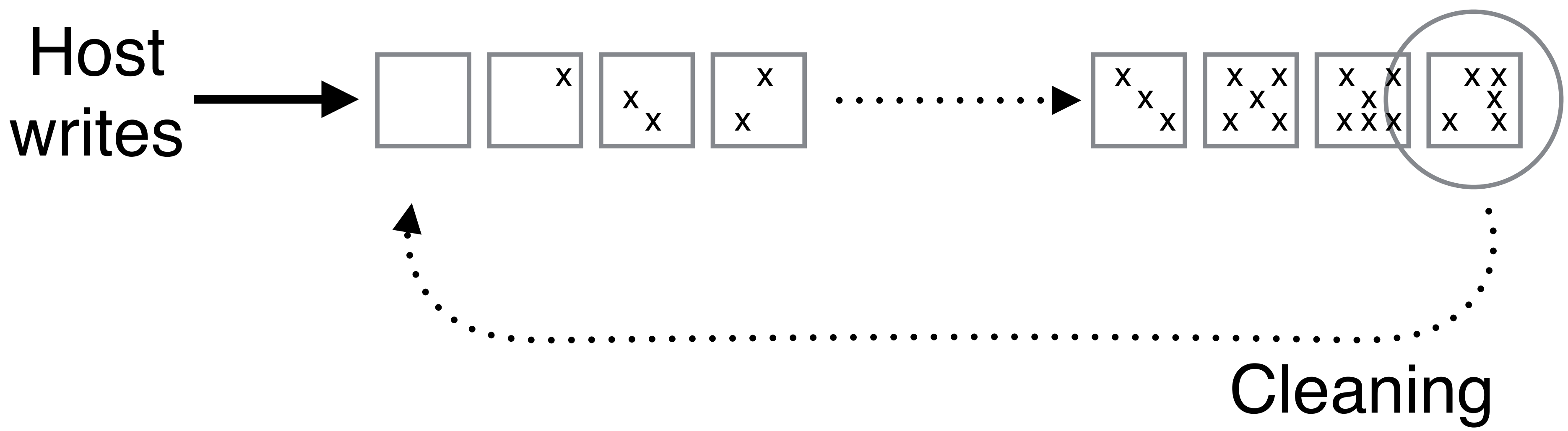
- Physical storage: **T** erase units, $\mathbf{N_p}$ pages each
- $\mathbf{U \cdot N_p}$ (**U** < **T**) logical pages, independently mapped
- Uniform random 1-page writes over LBA space
- Over-provisioning $\alpha = \frac{T}{U}$
- Spare factor $S_f = \frac{T-U}{T}$
- Single channel

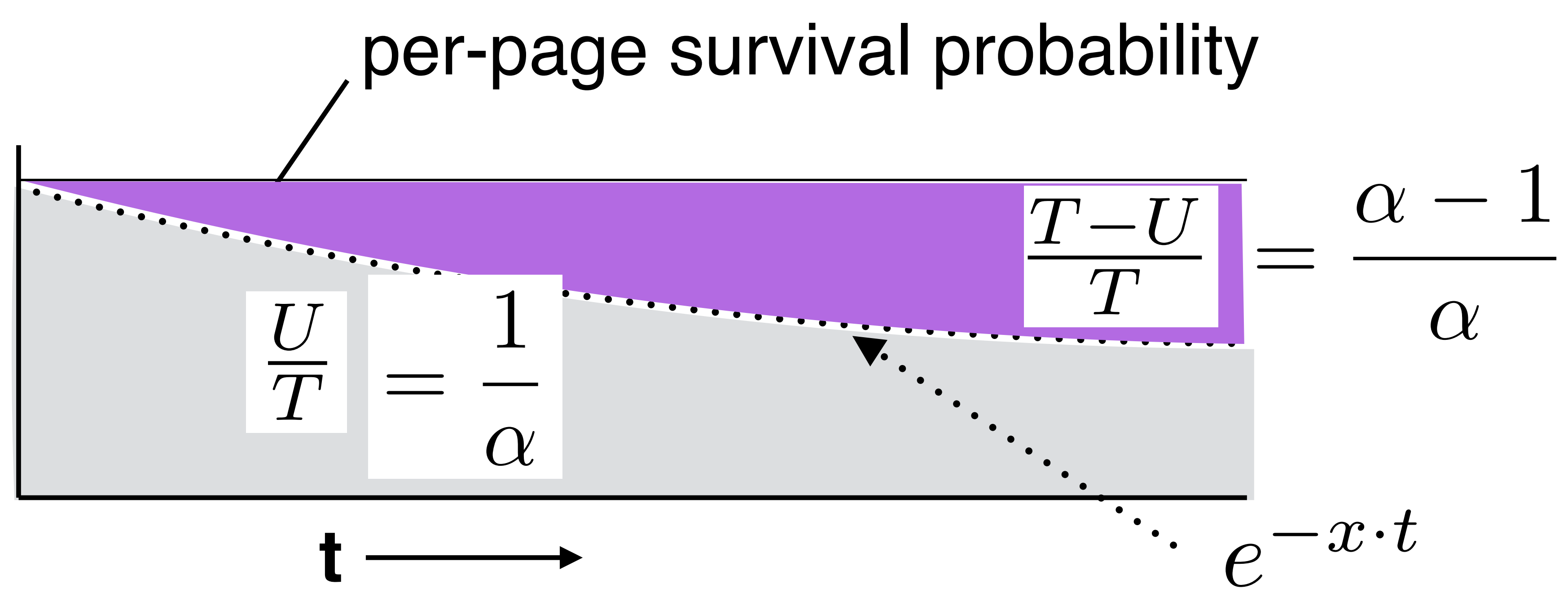
“LRU” cleaning



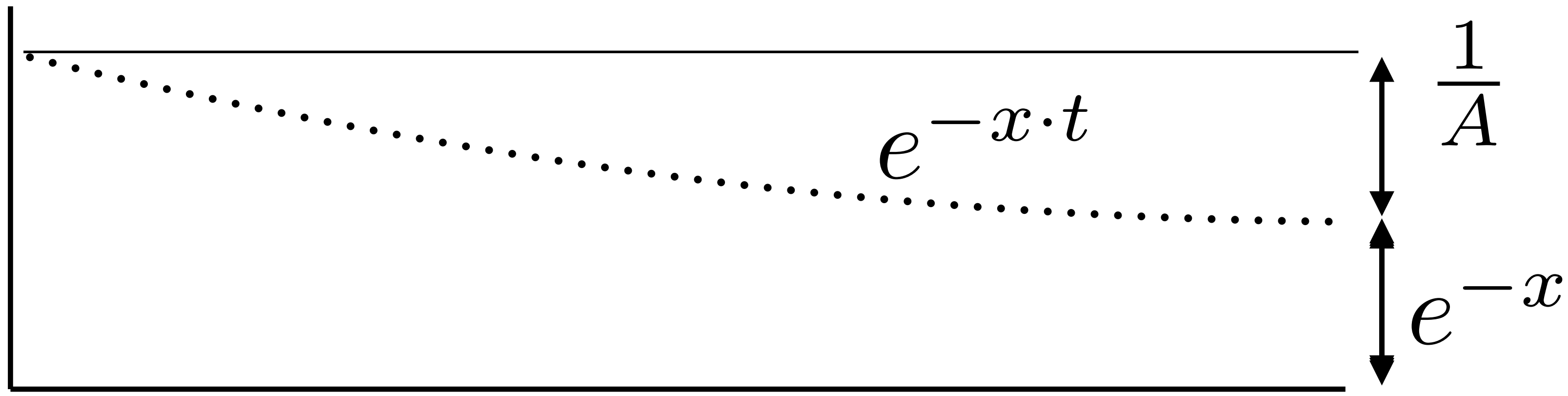


“LRU” cleaning





A bit of math...




```
In[2]:= Solve[Integrate[Exp[-x * t], {t, 0, 1}] == 1 / α,
          x] // TraditionalForm
```

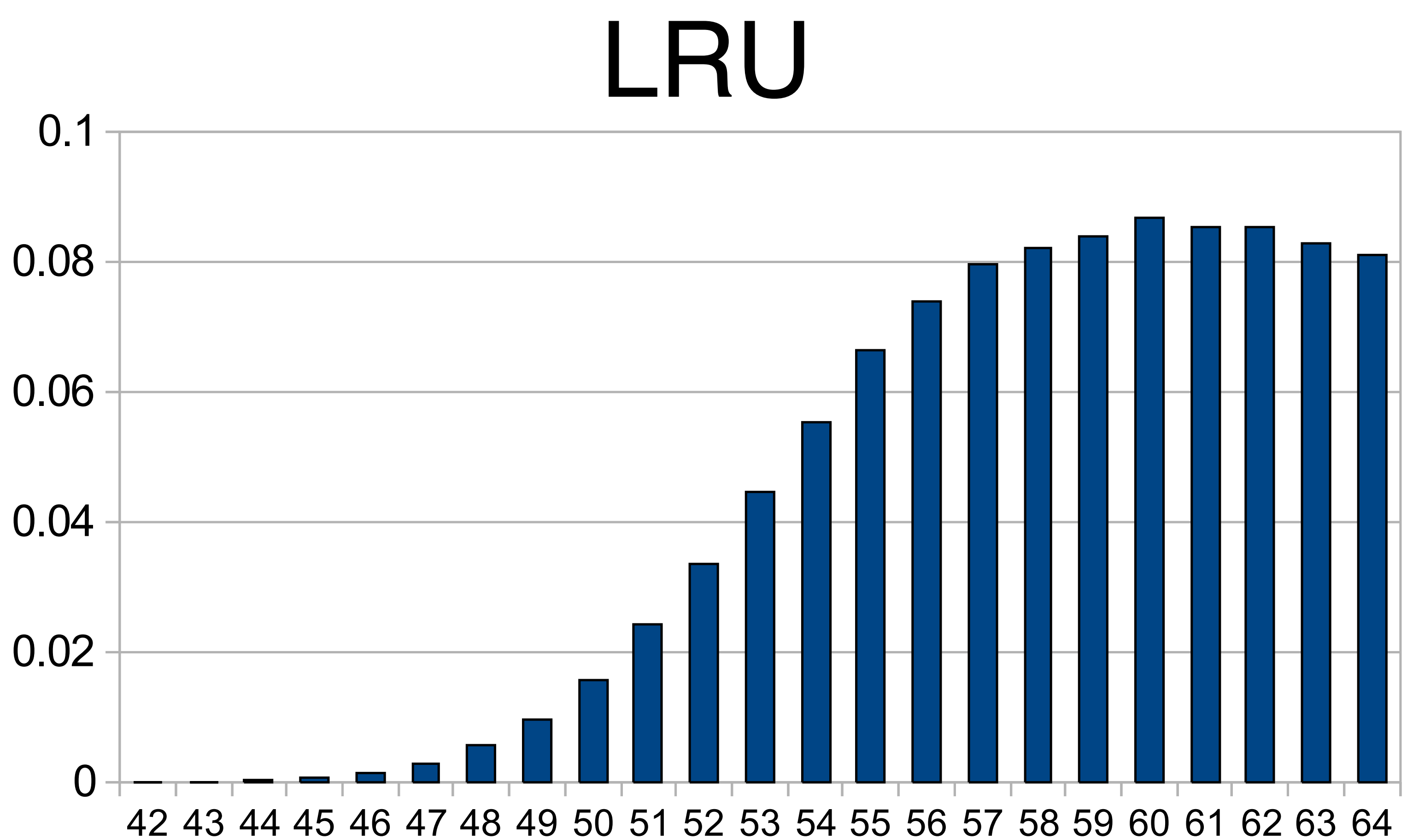
Out[2]//TraditionalForm=

$$\{\{x \to \alpha + W(-e^{-\alpha}\,\alpha)\}\}$$

$$A = \frac{\alpha}{\alpha + W\left(-\alpha e^{-\alpha}\right)}$$

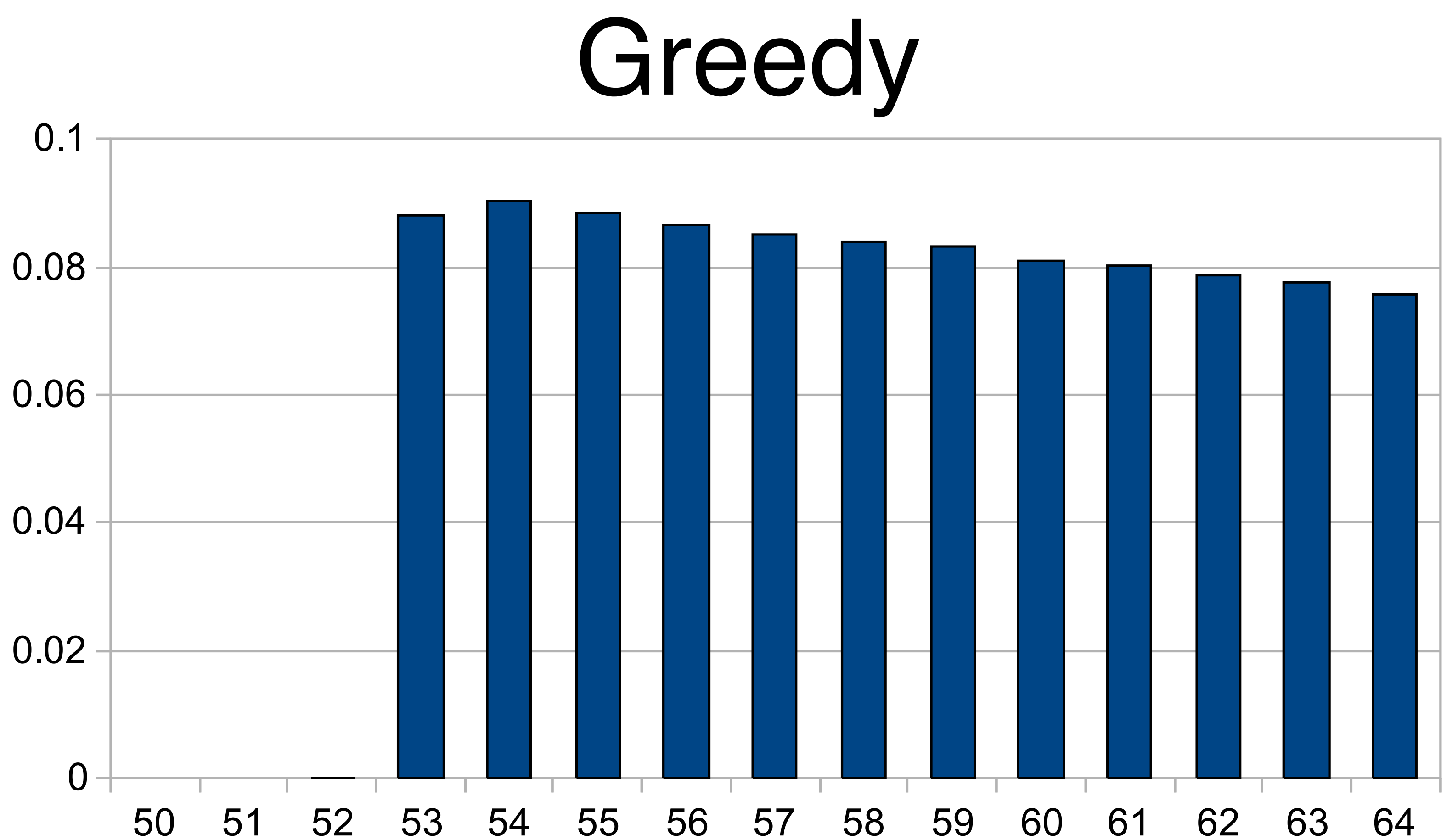
- Robinson ’96
- Menon & Stockmeyer ’98
- Xiang & Kurkoski ’12
- Desnoyers ‘12

What about Greedy?



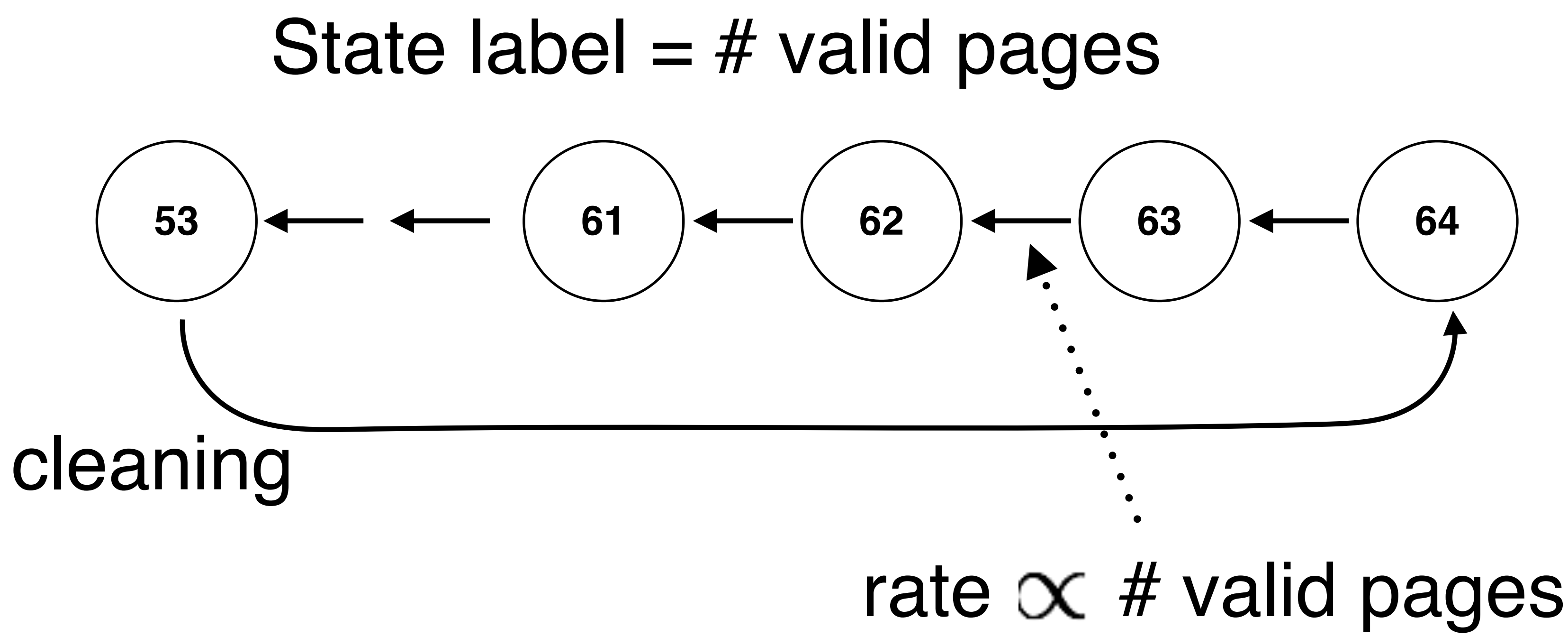
Valid page statistics (N_p=64, S_f=0.09)

What about Greedy?



Valid page statistics (N_p=64, S_f=0.09)

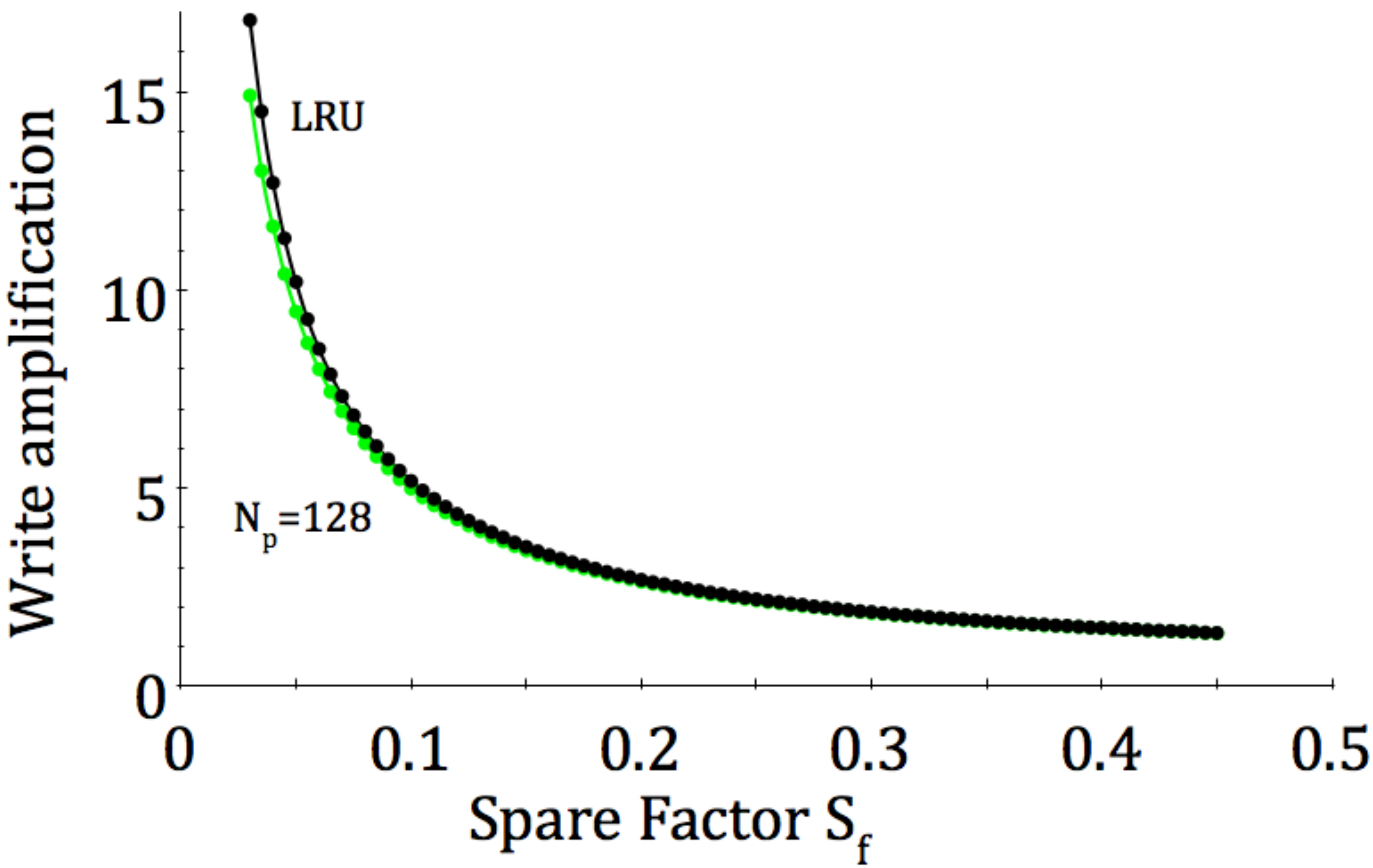
Per-block Markov



$$\alpha' = \left(1 + \frac{1}{2N_p}\right)\alpha \qquad A \approx \frac{A_{LRU}(\alpha')}{1 + \frac{1}{2N_p}}$$

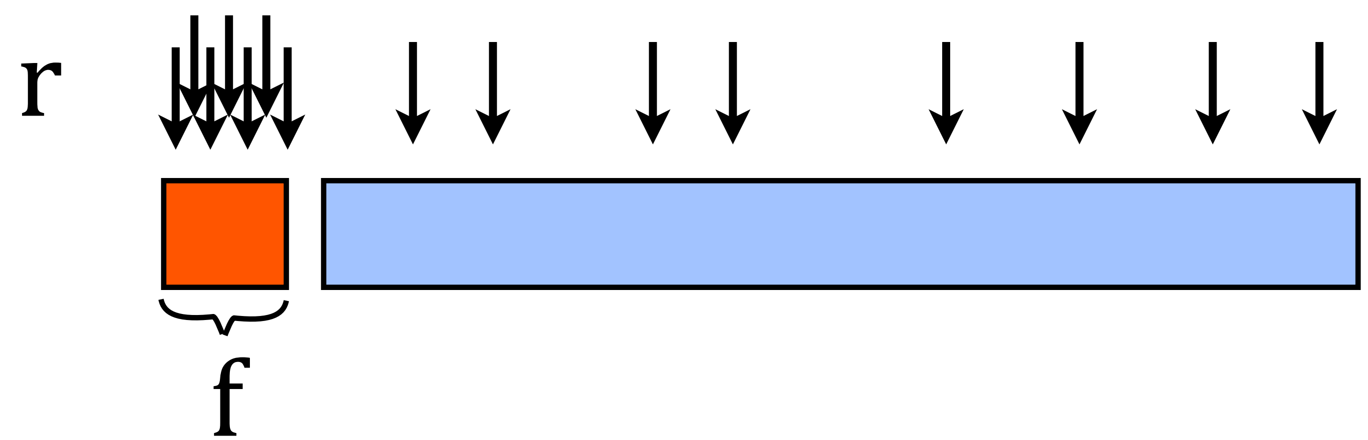
- Bux & Iliadis ‘10
- Desnoyers ‘12

Greedy vs. LRU



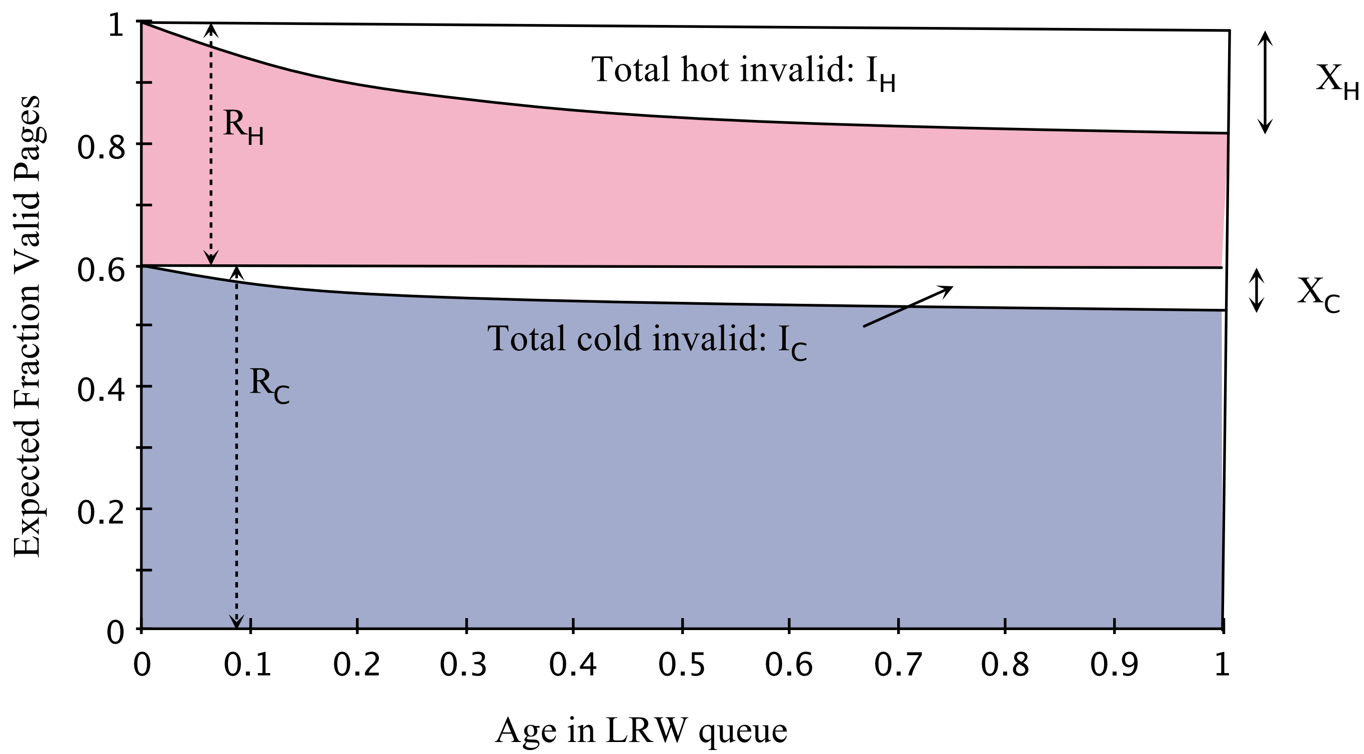
Hot/Cold data

- Rosenblum's model:
  fraction ***r*** of writes to ***f*** of LBA space



- E.g. 90% of writes to 10% of LBAs

LRU Cleaning

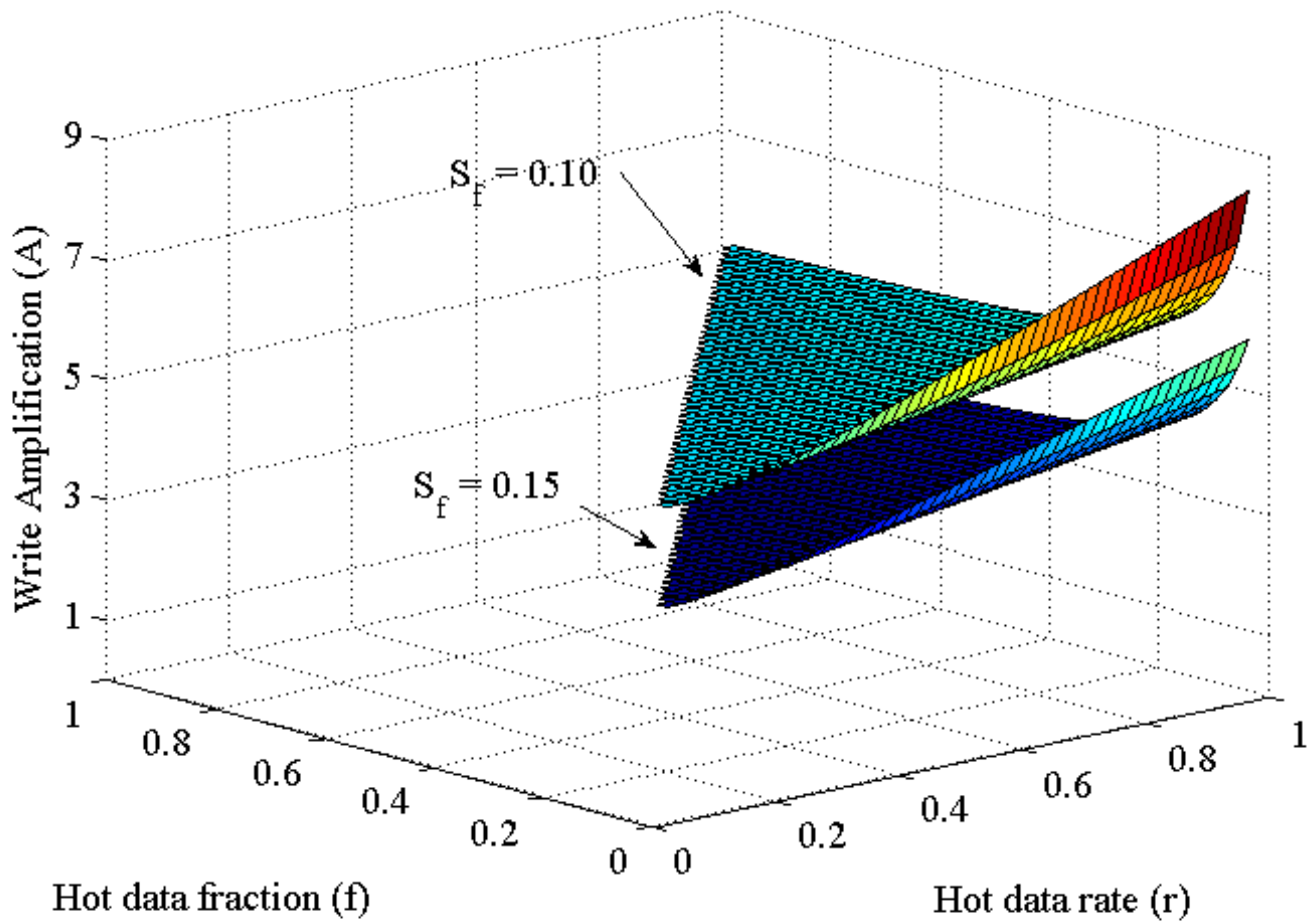


$$A = 1 + \frac{r}{e^{\frac{r}{f}\frac{\alpha}{A}} - 1} + \frac{1-r}{e^{\frac{(1-r)}{(1-f)}\frac{\alpha}{A}} - 1}$$

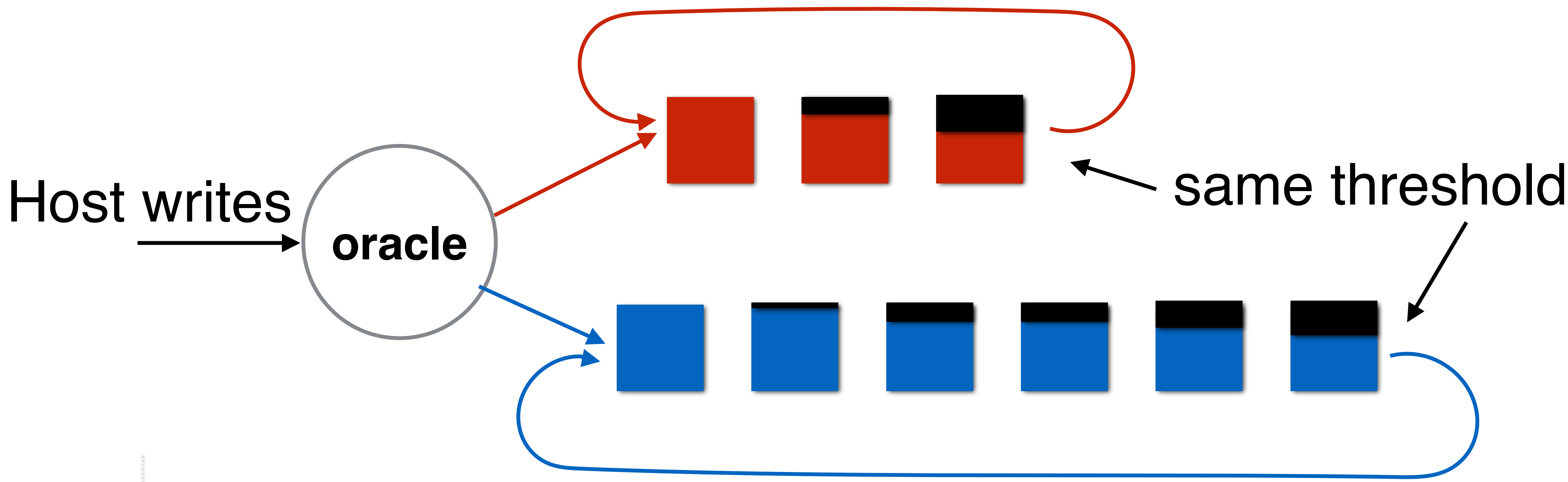
- solve numerically

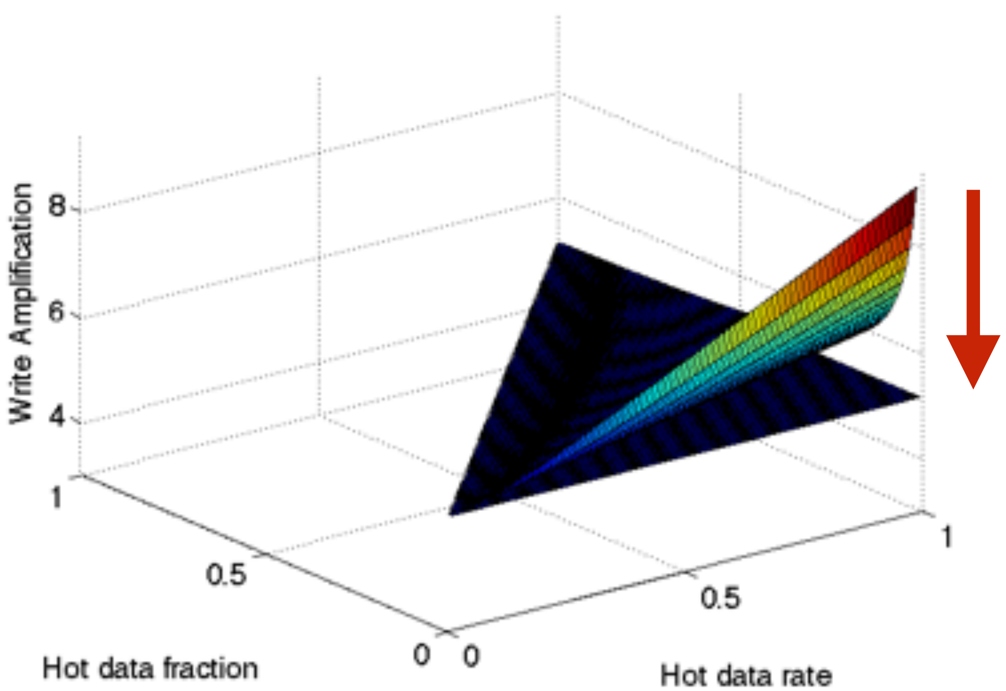
- Menon & Stockmeyer '98
- Desnoyers '12

LRU hot/cold performance

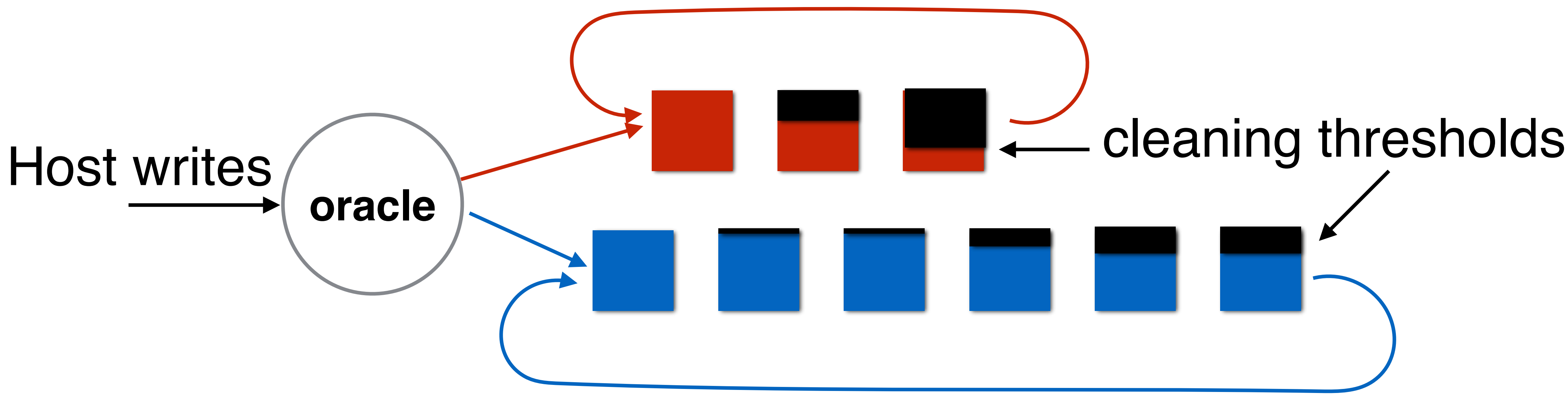


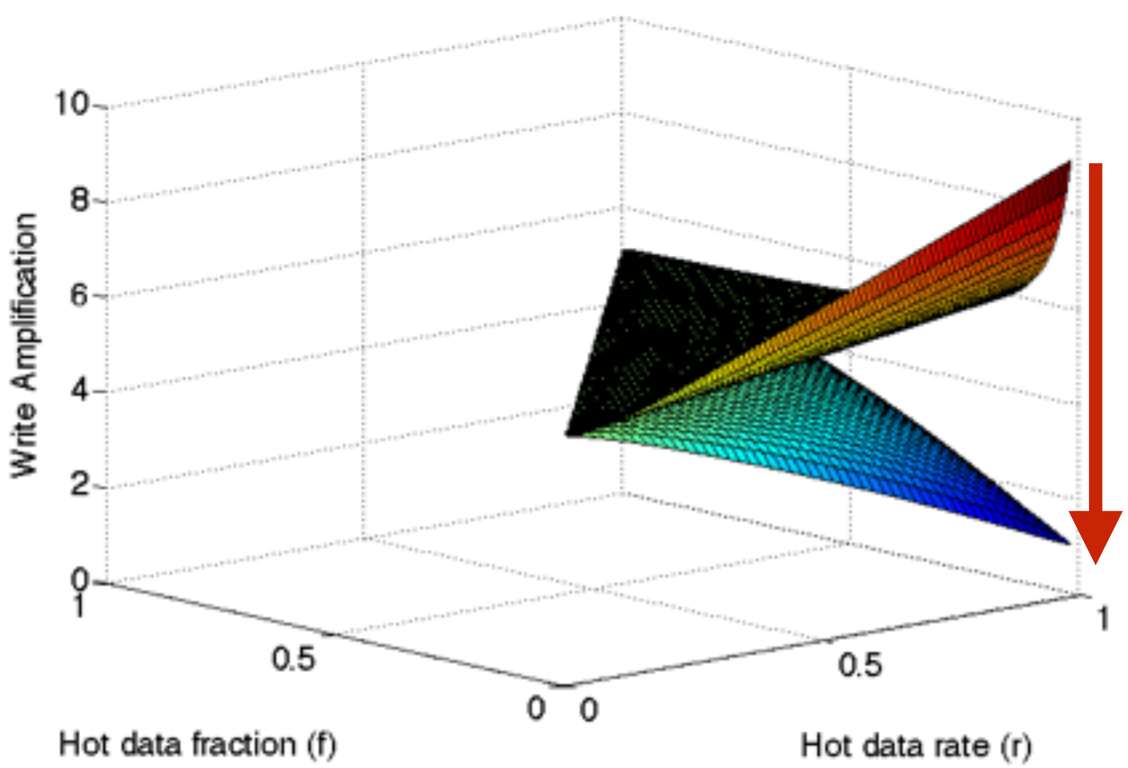
Hot/Cold Separation

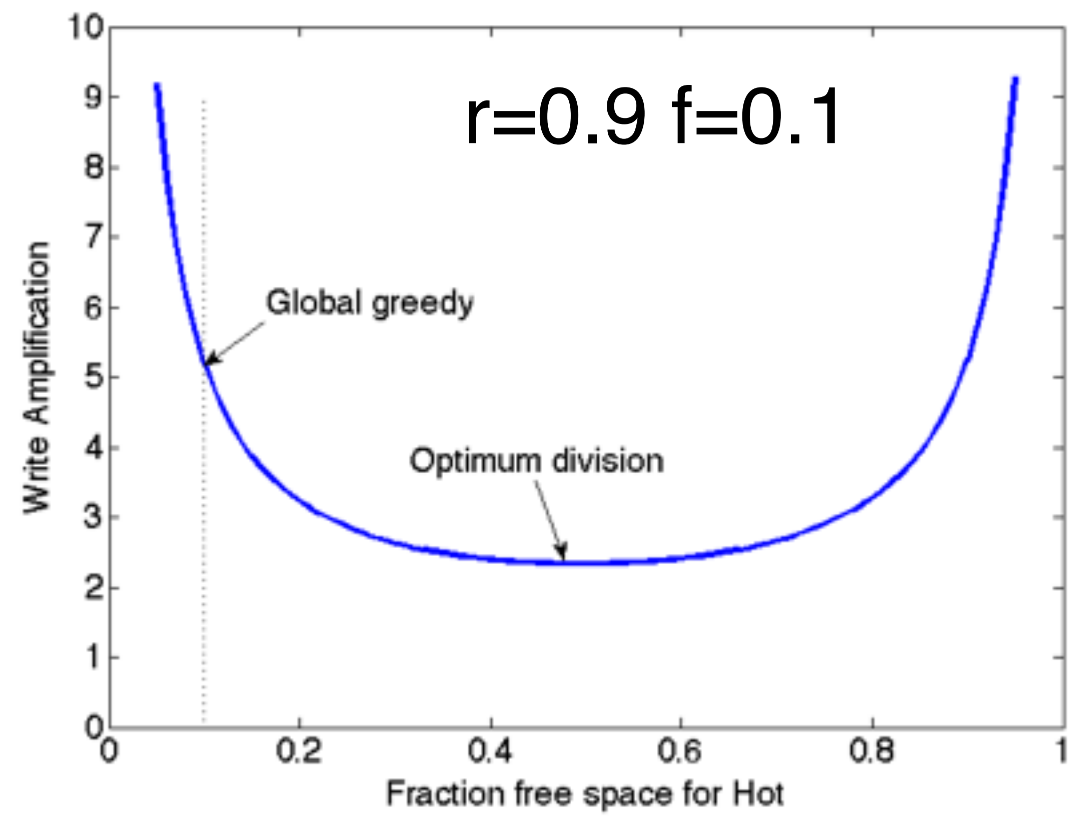




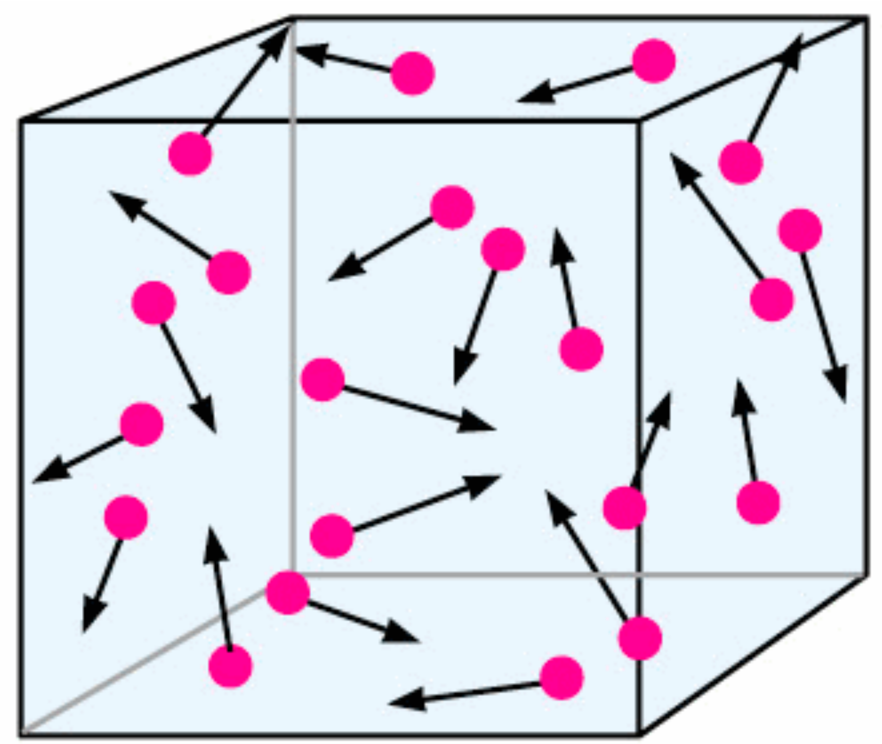
Split cleaning thresholds

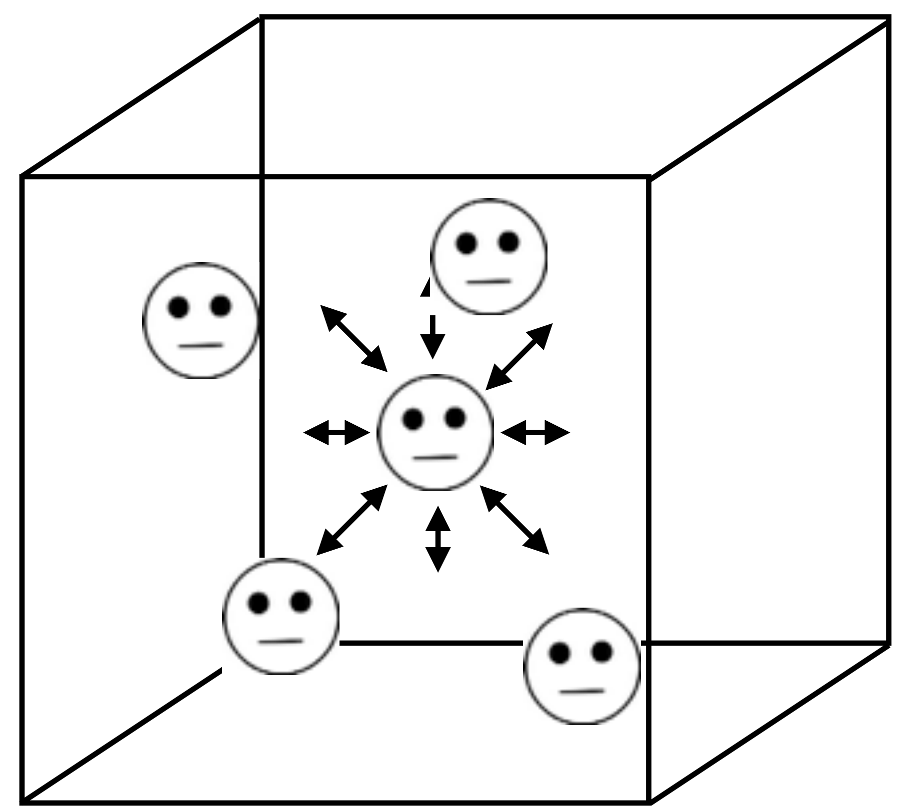






Mean field methods

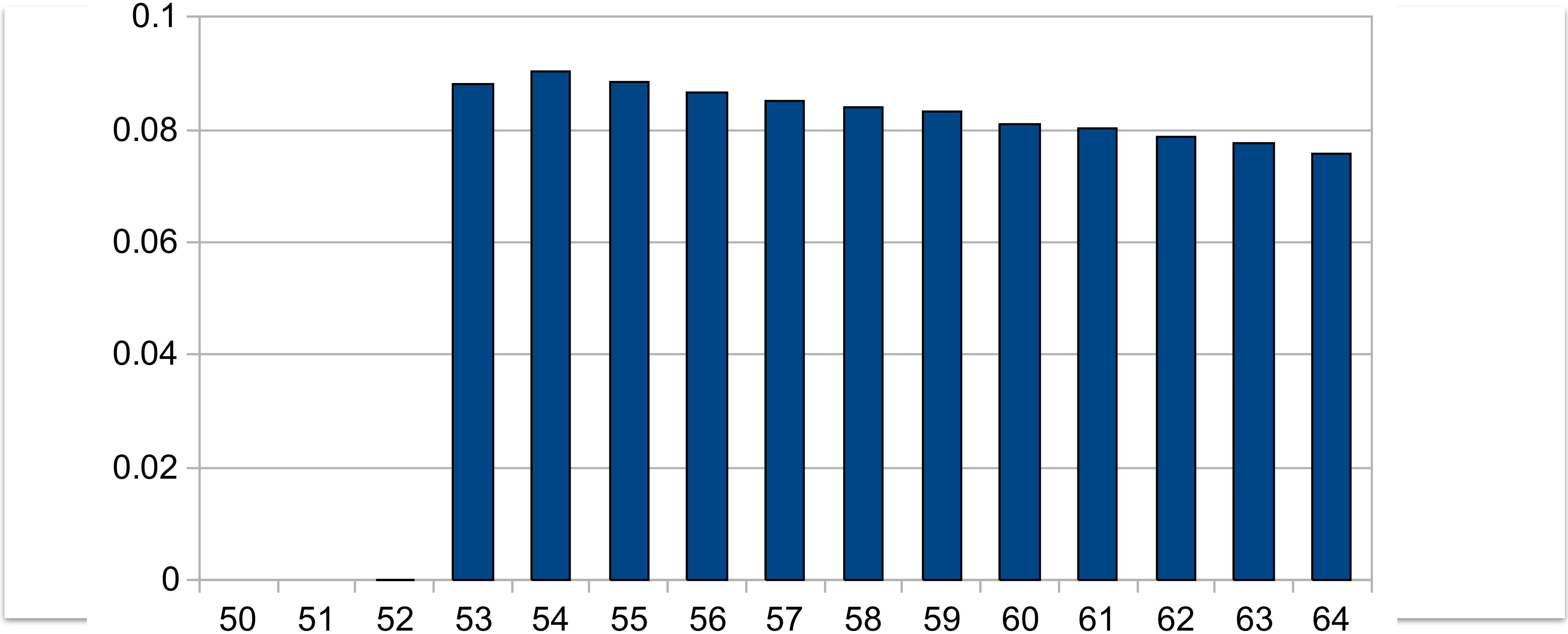




- Van Houdt 2013

Mean Field Analysis

Requirements for mean-field analysis:

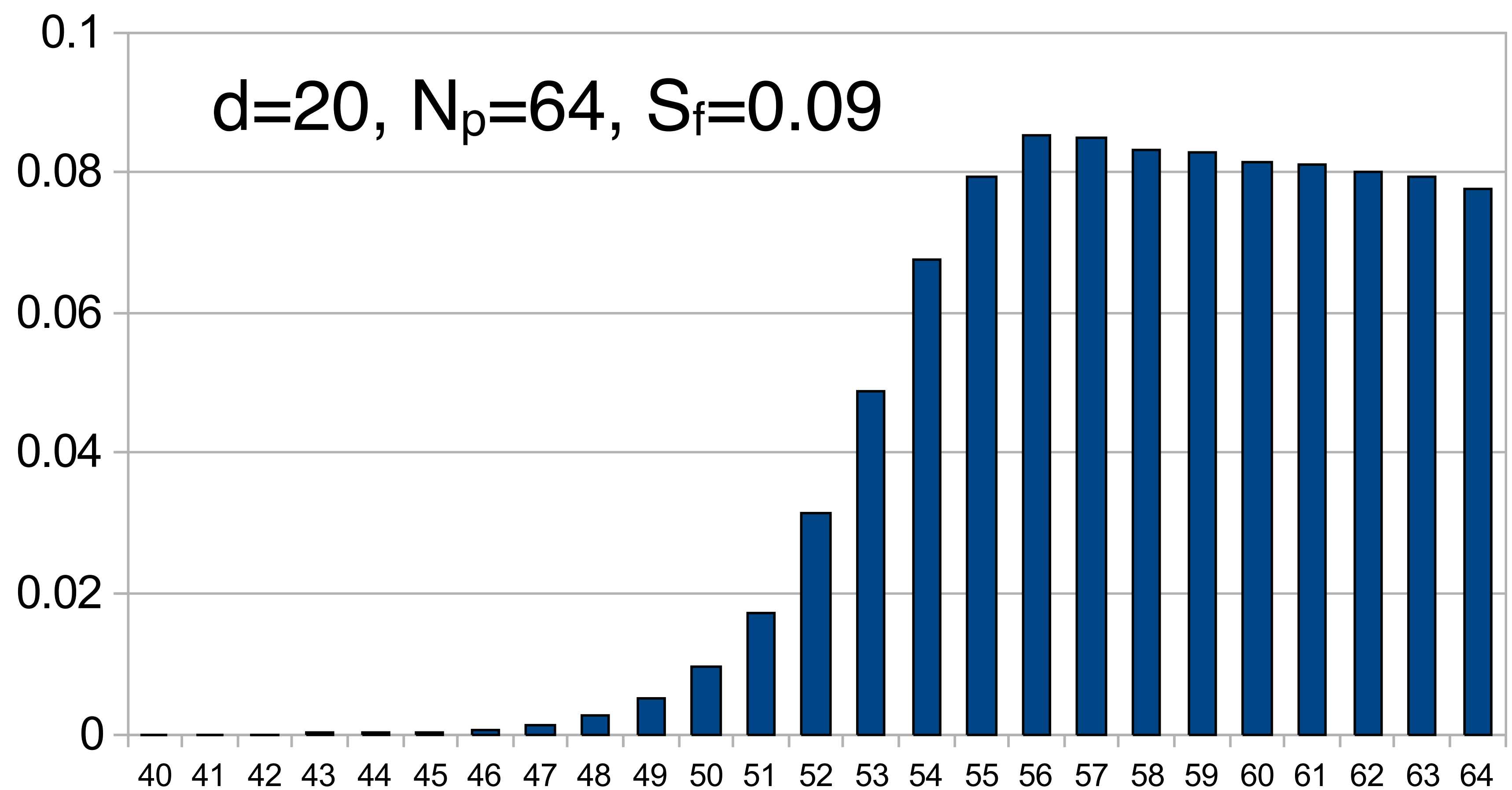


2. $\mathbf{p_j(m)}$ is *smooth* in $\mathbf{m}$ (e.g. <u>not</u> greedy)

- SIGMETRICS 2013

d-CHOICES Cleaning

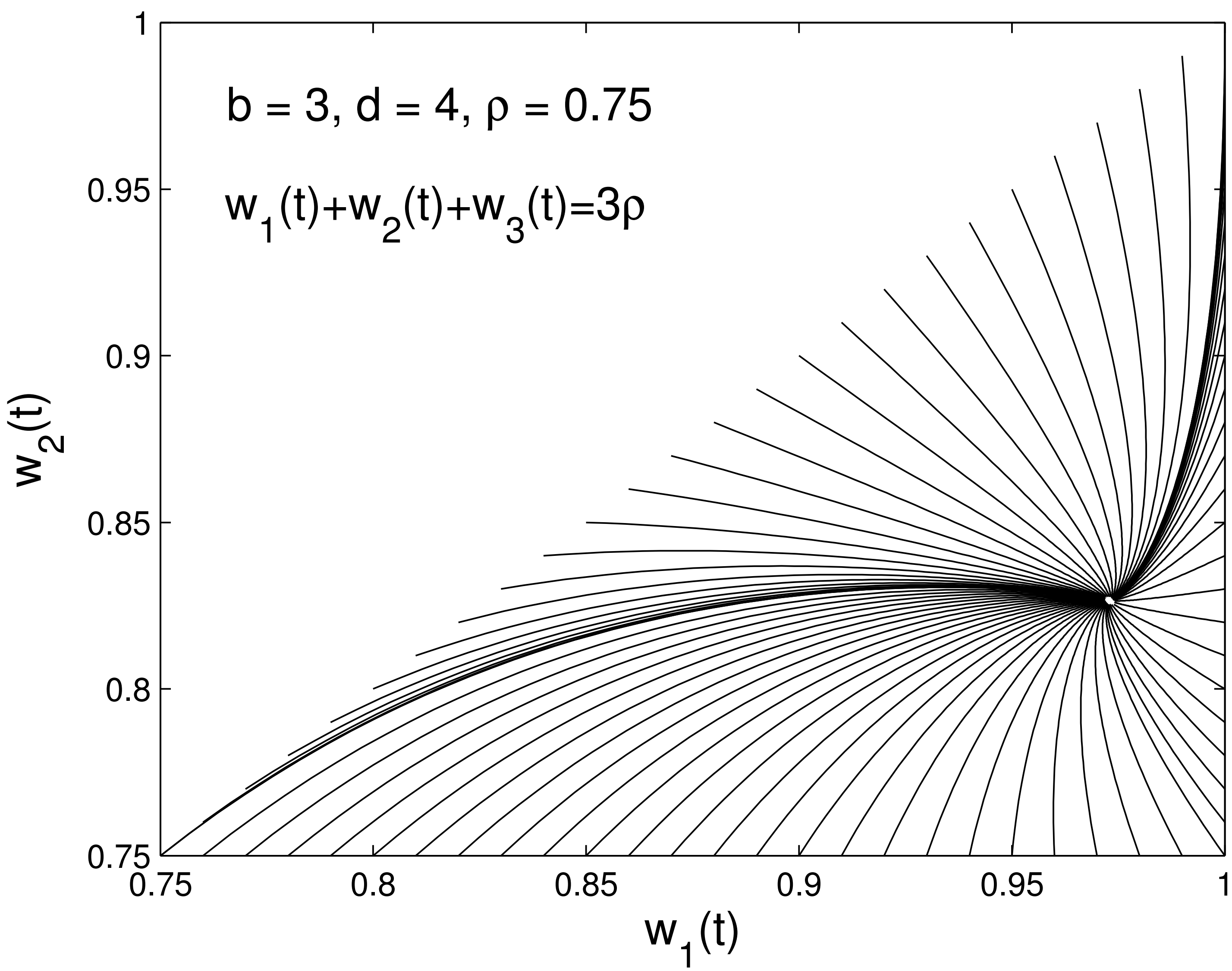
- d-CHOICES:
  - Randomly select ***d*** erase units **{b_1…b_d}**
  - Choose **b_i** with minimal valid pages



Mean field solution

- Define *drift* ***f*** as change in global state $\mathbf{M^N(t)}$ over a single cleaning cycle.
- Let ***μ*(t)** be defined by ODE: $\frac{d\mu}{dt} = f(\mu(t))$
- as t →∞, $\|M_N(t) - \mu(t)\| \to 0$
- and converges if the ODE has a fixed point that is a global attractor.

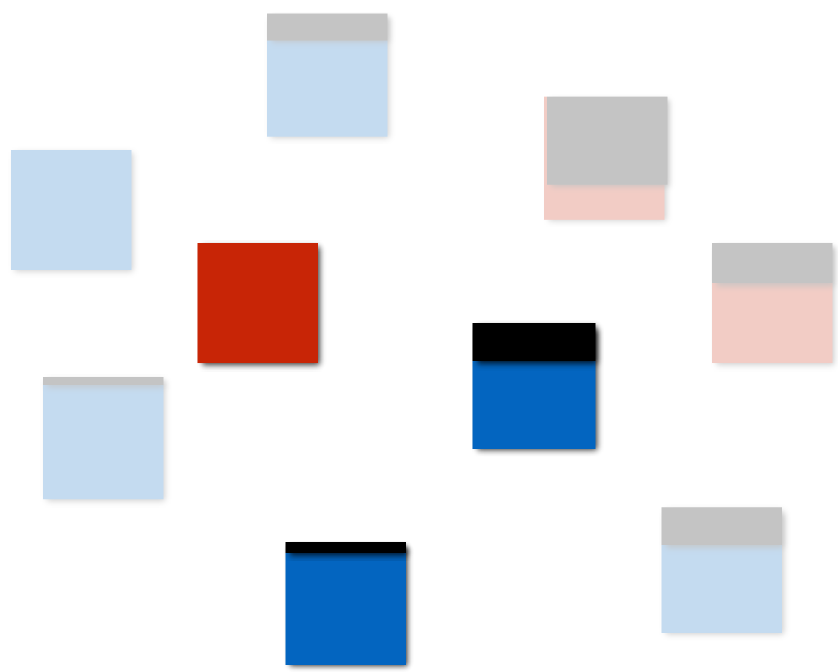
Solve numerically…

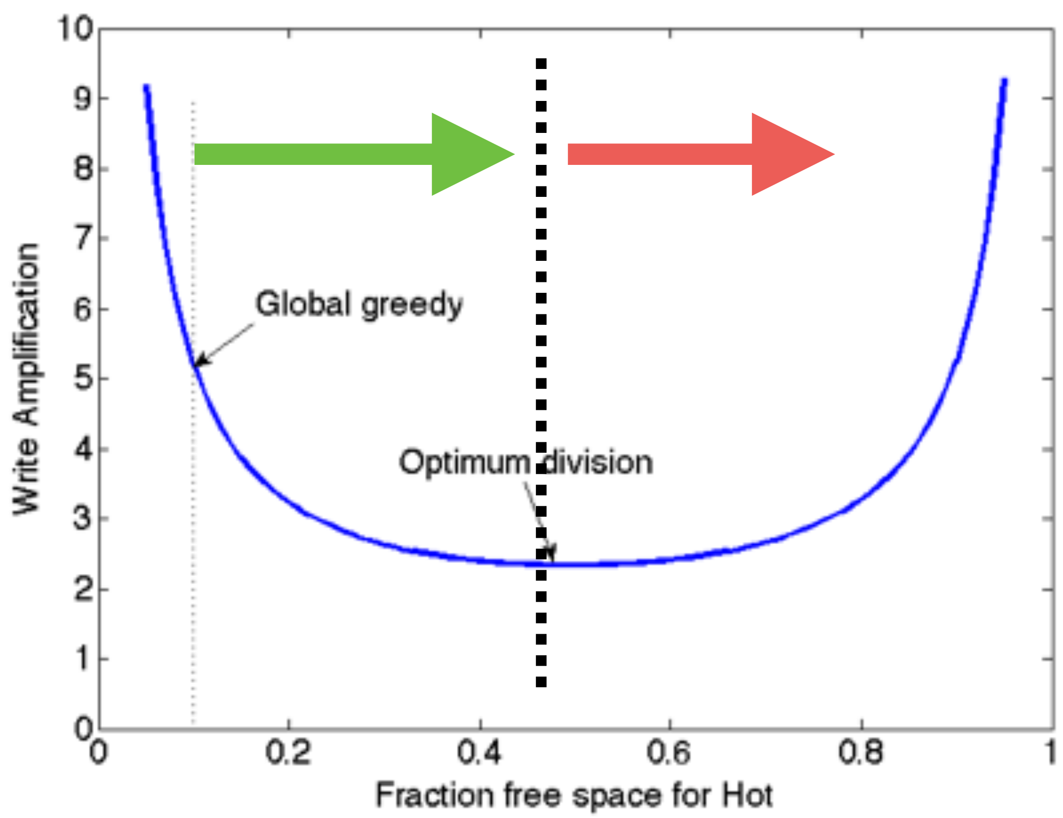


Extensions

- Greedy (limit as $d \to \infty$)
  - SIGMETRICS 2013
- hot/cold data and single write frontier
- hot/cold data and dual write frontier (i.e. with separation)
  - IFIP Performance 2013

d-CHOICES for Hot/Cold

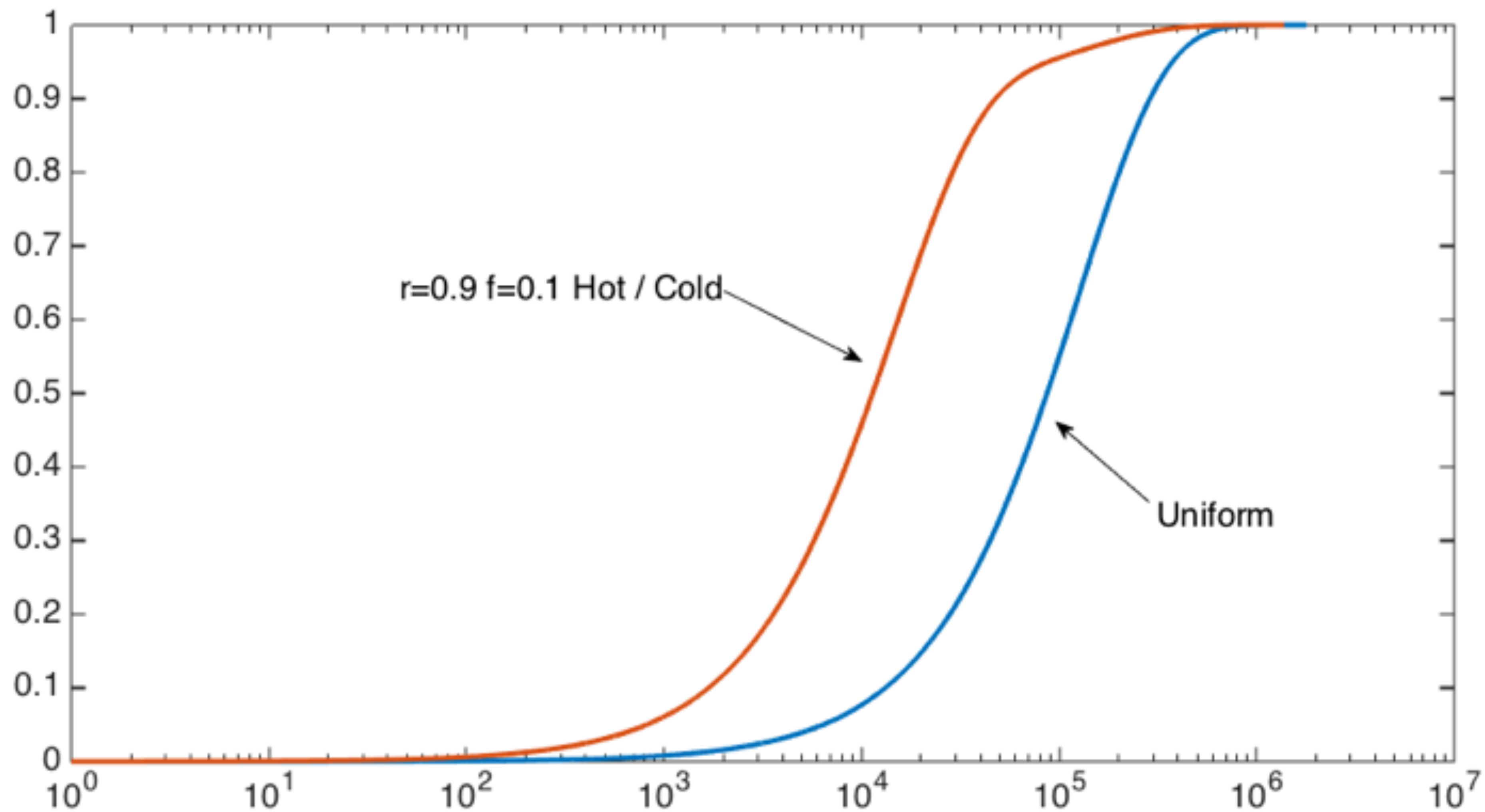




More on locality

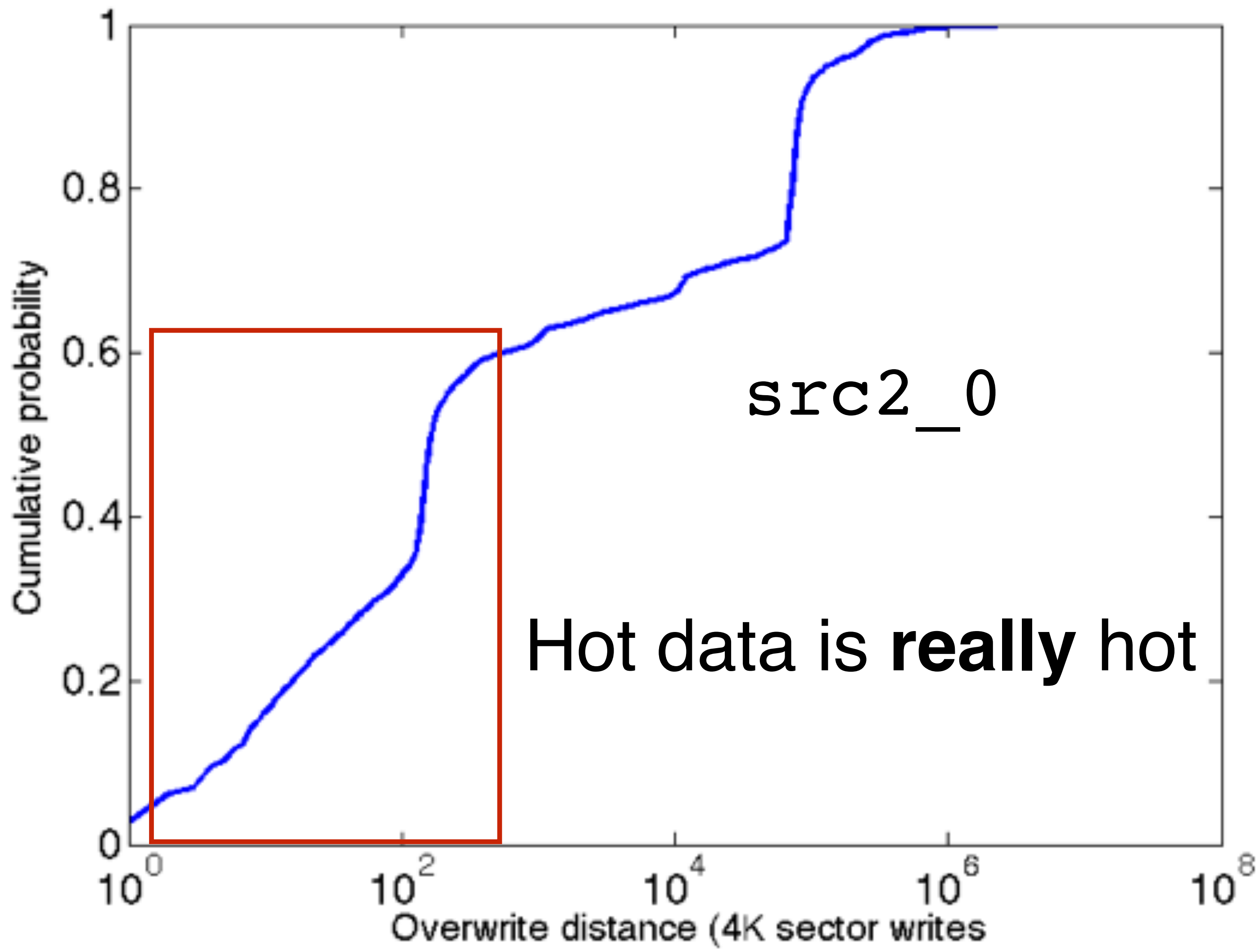
- What does hot / cold data mean?
    - Expected time to over-write
- Can it be predicted?
    - Overwrite vs. create/delete
- What about spatial locality?

Looking at the data



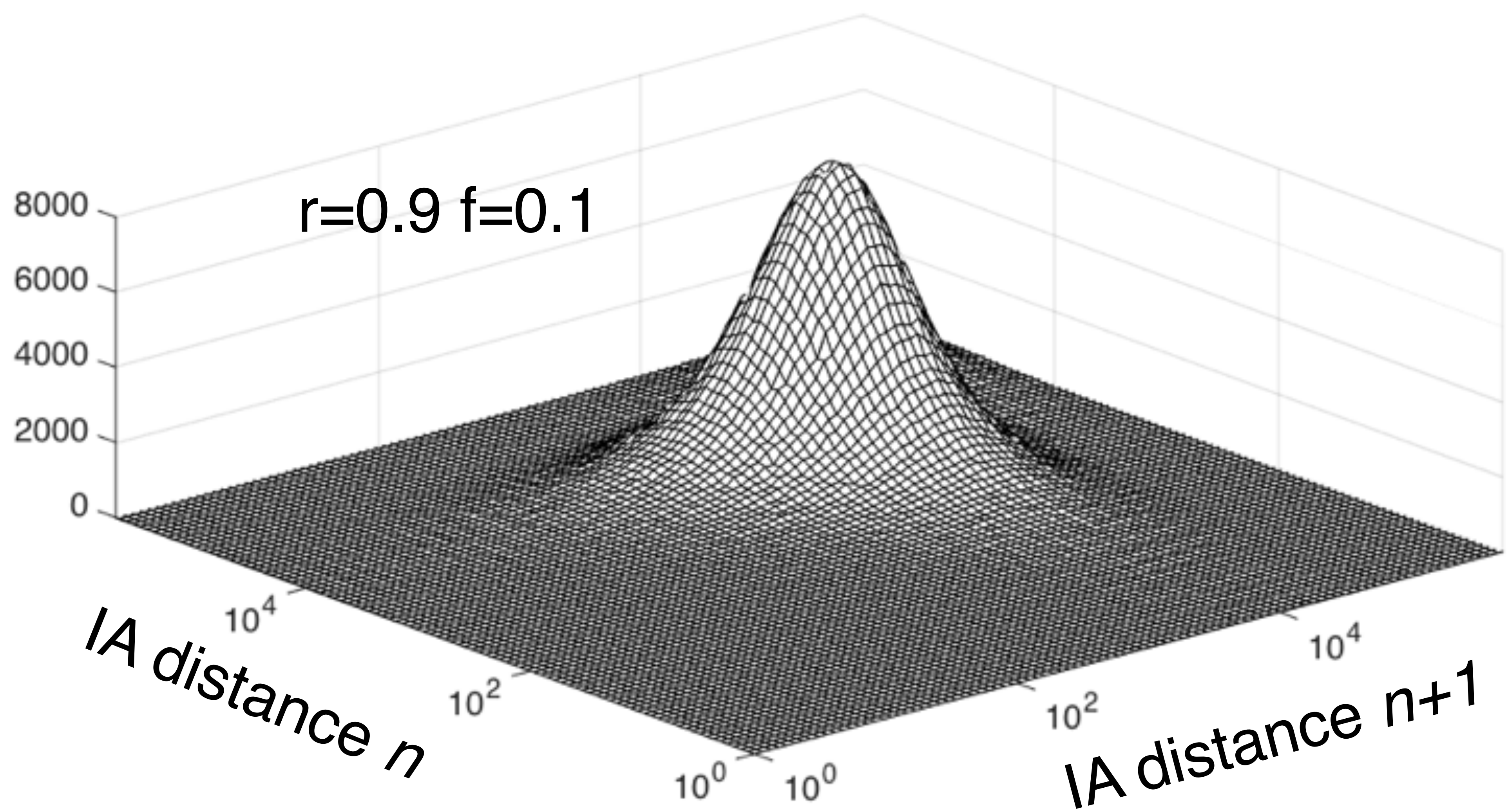
Synthetic

Looking at the data



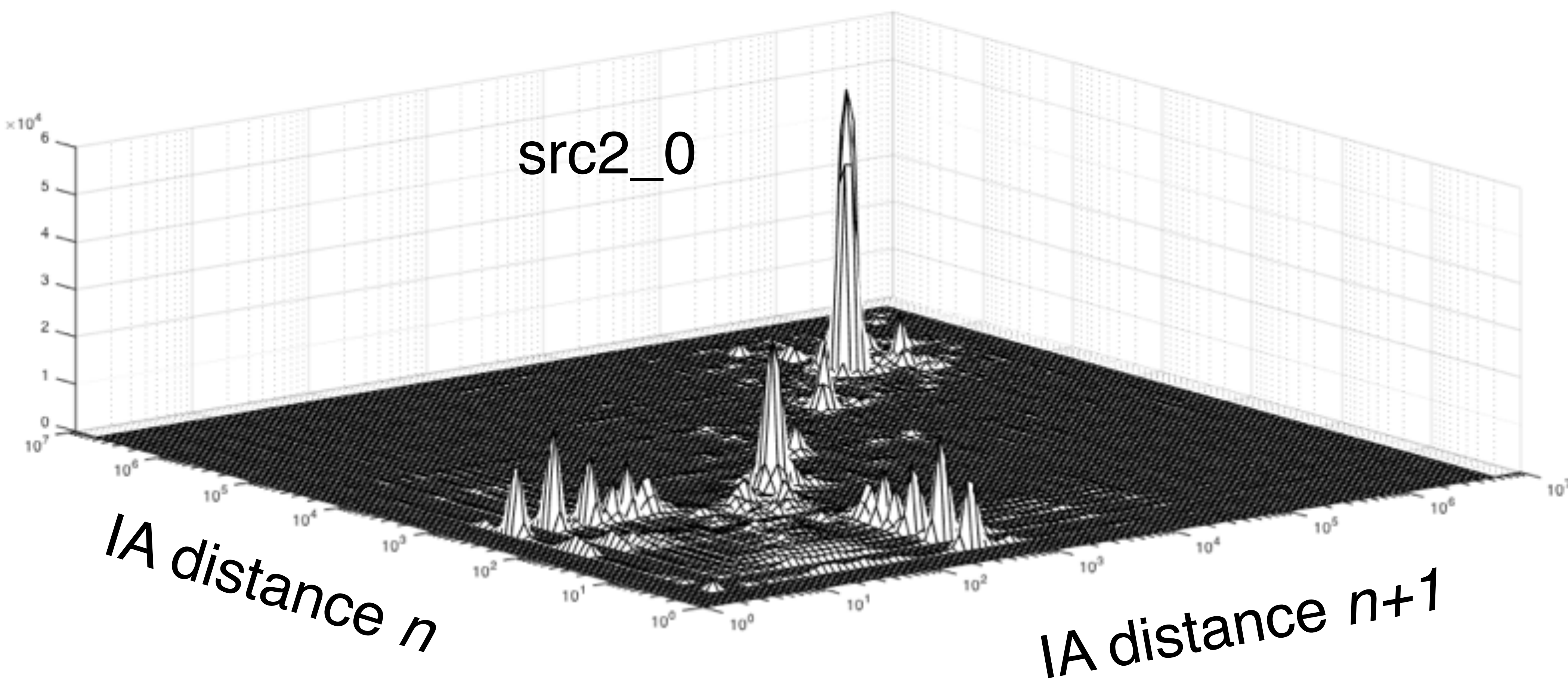
Real

Predicting lifetime



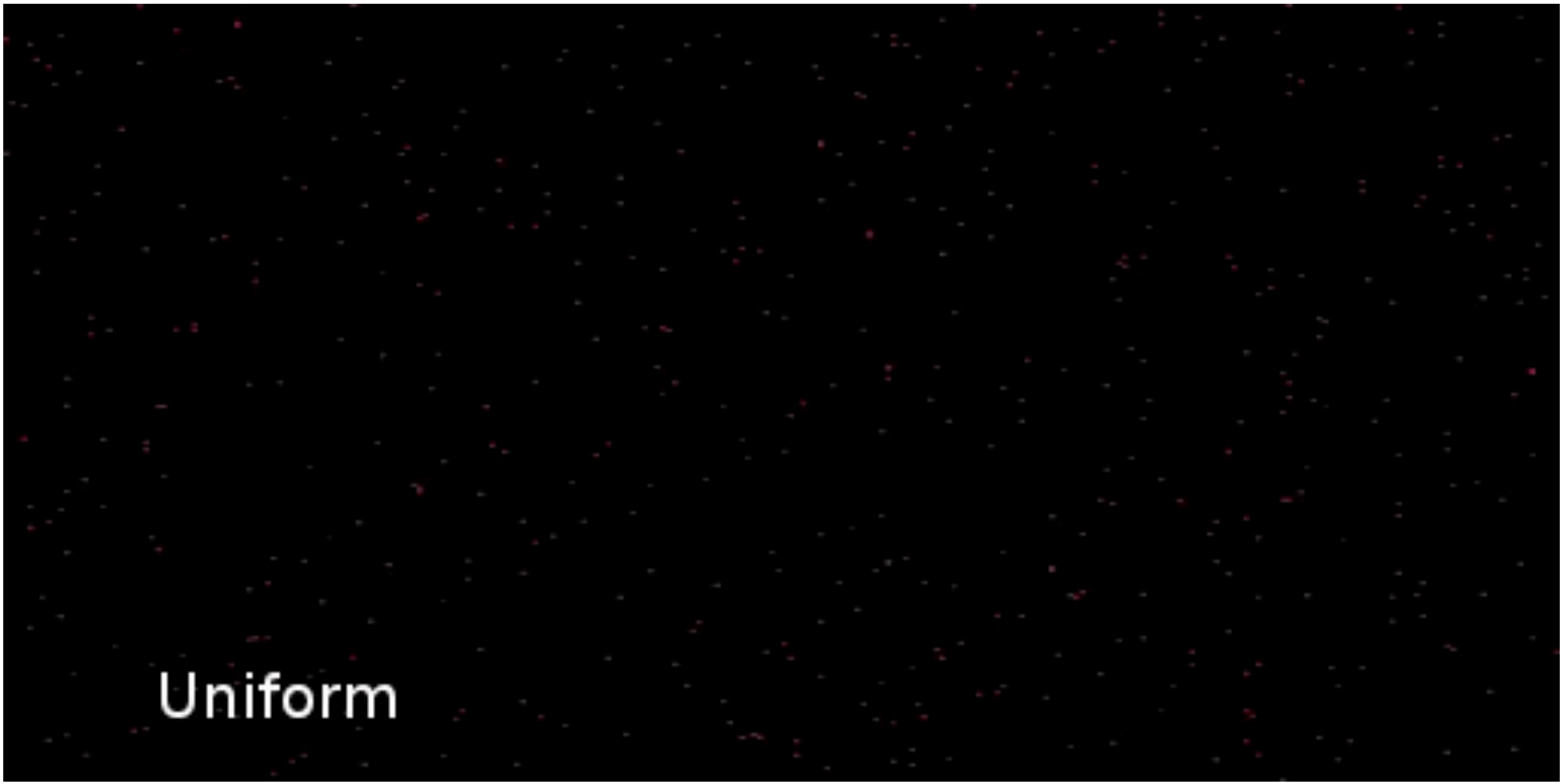
Synthetic

Predicting lifetime



Real

Models ≠ data



What have we learned?

- How FTLs perform for synthetic data
- What hot/cold data means
  - expected time until re-write
- How Greedy helps random data (vs. LRU)
  - ½ page per cleaning, ½ page/block free space
- For best performance, clean hot blocks at lower utilization than cold ones. (but how much lower?)

What don't we know?

- Other models of hot/cold data
  - mutate vs. create/delete
- Spatial locality
  - no model
  - no metric
- Log-on-Log?
- Is there an optimal FTL?