Online Thinning to Reduce Discrepancy

Raaz Dwivedi, Ohad Feldheim, Ori Gurel-Gurevich and Aaditya Ramdas

raaz.rsk@berkeley.edu, ohad.feldheim@mail.huji.ac.il, ori.gurel-gurevich@mail.huji.ac.il, aramdas@berkeley.e

Berkeley UNIVERSITY OF CALIFORNIA

האוניברסיטה העברית בירושלים THE HEBREW UNIVERSITY OF JERUSALEM

Notation

$X_i \in [0,1)^d, X^N = \{X_1, \ldots, X_N\}, X^\infty = \{X_1, X_2, \ldots\},$

$\mathcal{R}$ = Set of axis-aligned hyper-rectangles in $[0,1)^d$,

$\mathrm{vol}(\cdot)$ = Lebesgue Measure.

Quantity of interest: Rate of the discrepancy sequence

$$\mathrm{Dis}(X^N) = \sup_{R\in\mathcal{R}} \left| \frac{1}{N} |R \cap X^N| - \mathrm{vol}(R) \right|.$$

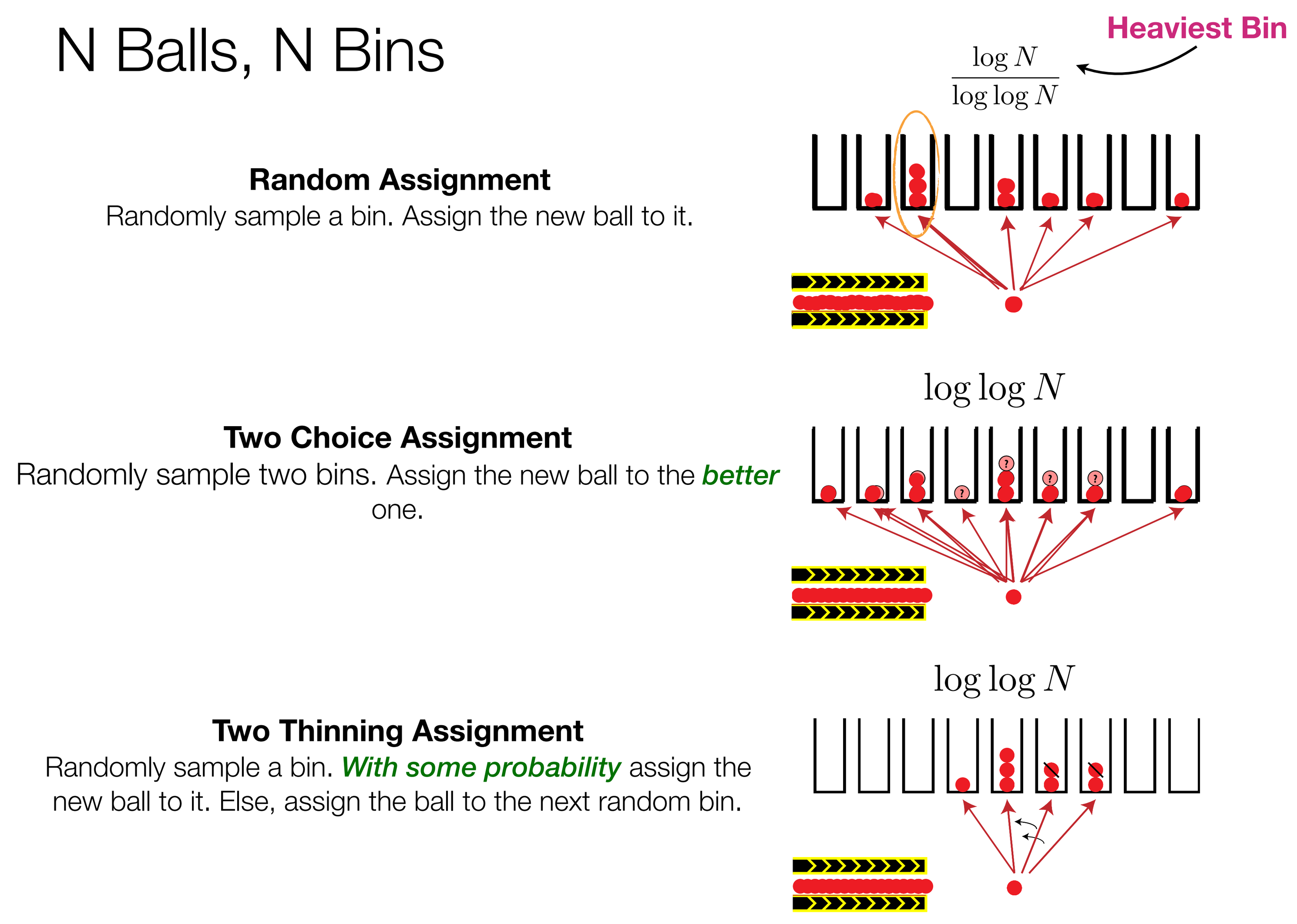
Set-up and Objectives

Goal: Design a strategy, which *selects a subset* Z^∞ from a streaming sequence X^∞ of i.i.d. $U[0,1]^d$ random variables such that $\mathrm{Dis}(Z^N) \ll \mathrm{Dis}(X^N)$.

Requirements: We can keep or reject any point in X^∞. The sequence Z^∞ should be **dense** in X^∞. The strategy should be **online** and **time and space efficient**.

Remark: $\mathrm{Dis}(X^N)$ = Monte Carlo Discrepancy $= \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$.

Inspiration: Power of Two Choice



Idea: Extend the two thinning strategy to our set-up.
Difficulty: The cardinality of $\mathcal{R}$ and heavy dependence across its elements.
Tool: Use of Haar Wavelet Basis.

Main Result

For a streaming sequence X^∞ of i.i.d. $U[0,1)^d$ random variables, *Haar 2-Thinning strategy* outputs a streaming sequence Z^∞ such that

$$\mathrm{Dis}(Z^N) = \mathcal{O}\left(\frac{d\log^{2d+1}(N)}{N}\right) \quad \forall N \in \mathbb{N}, \text{ almost surely.}$$

The strategy has $\mathcal{O}(d \log^d N)$ space-time complexity to output N points and keeps at least one out of every two consecutive points of the sequence X^∞.

Haar 2-Thinning Strategy

- Given Z^N and a new candidate point x, each Haar wavelet of order up-to $\lceil \log_2 N \rceil$ votes whether to keep x or reject it.
- Haar functions vote to maintain *balance of points* in their support.
- Based on votes, the point x is kept *randomly with probability*

$$\mathbb{P}(\text{keep } x) = \frac{1}{2} + \frac{1}{2} \cdot \frac{\#\text{votes in favor of } x - \#\text{votes against } x}{\#\text{total votes}};$$

if x is rejected, the next candidate point is kept.

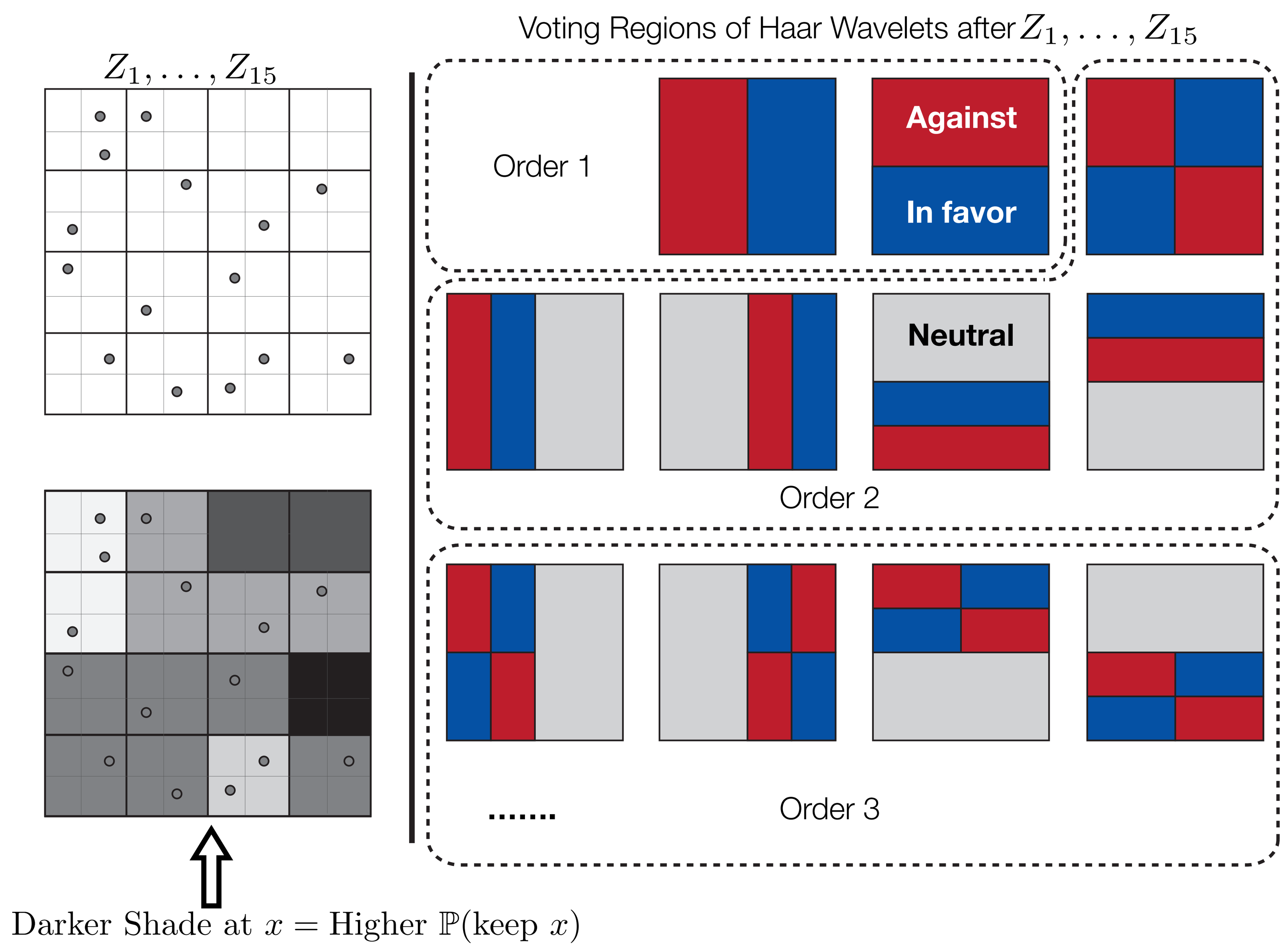


Figure 1: Illustration of Haar 2-thinning strategy after obtaining 15 samples in two dimensions. For clarity, $\mathbb{P}(\text{keep } x)$ is averaged on diadic squares of side 1/4. The strategy favors to keep points in regions that are *deficient* in samples so far.

Haar Greedy-Thinning Strategy

Conjecture: A simplified greedy strategy with $\mathcal{O}(d \log^d N)$ space-time complexity outputs a thinned sequence Z^∞ from a streaming sequence of i.i.d. $U[0,1)^d$ random variables, such that

$$\mathrm{Dis}(Z^N) = \mathcal{O}\left(\frac{d\log^{d} N}{N}\right) \quad \forall N \in \mathbb{N}, \text{ almost surely.}$$

Greedy Strategy: **Winner takes it all**. Keep x *deterministically* if # votes in favor of x > # votes against x. Reject otherwise.

Numerical Experiments

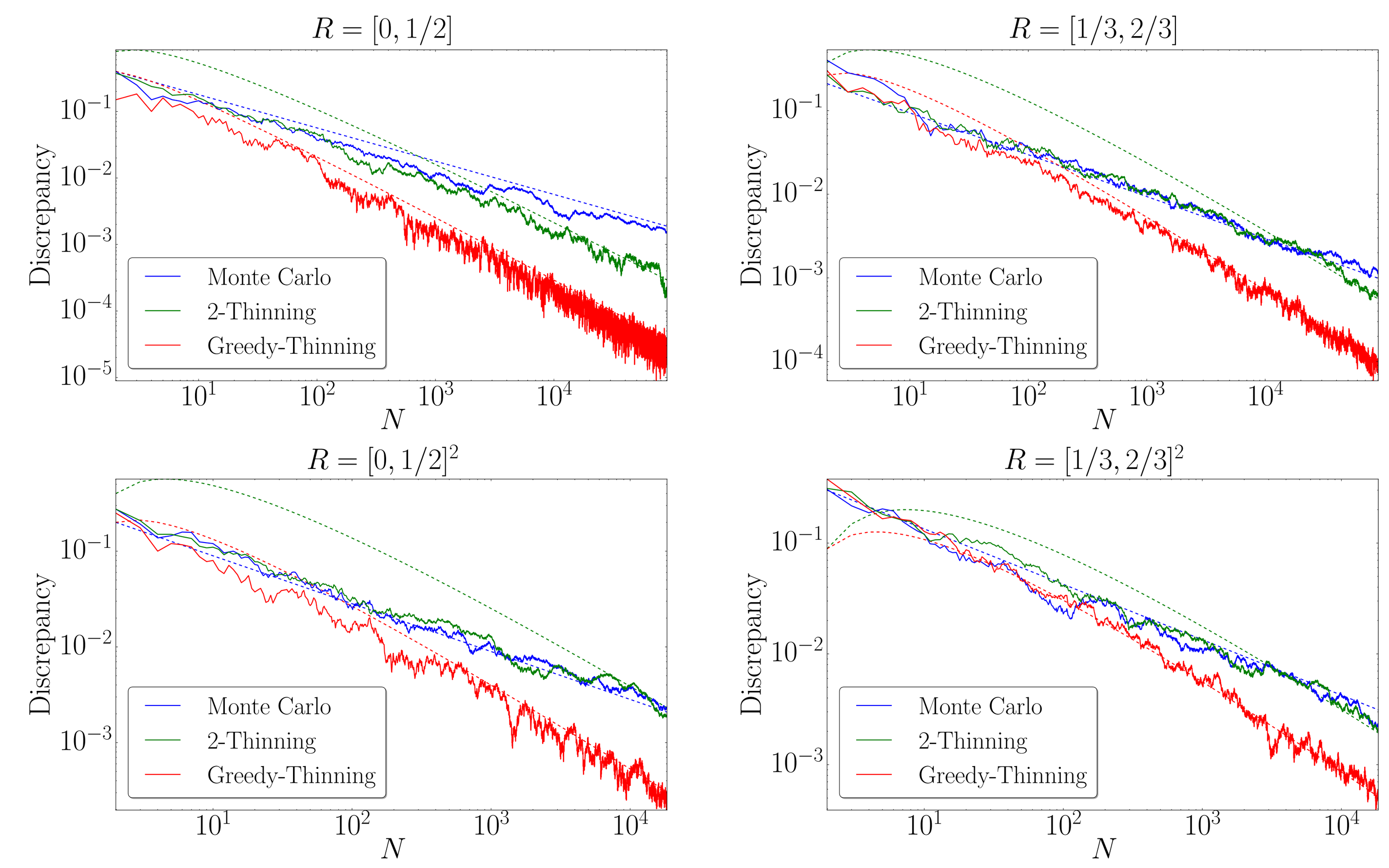


Figure 2: Plots of discrepancy for two different rectangles R in one and two dimensions (averaged over 20 experiments). Axes are in logarithmic scale.

Proof Techniques

- Decompose arbitrary rectangles in terms of "diadic rectangles"
- Express diadic rectangles in terms of "Haar wavelets"
- Exploit *exponential concentration* of self-regulating processes to maintain the balance of points in Haar wavelets up-to some resolution

References

[1] Y. Azar, A. Broder, A. Karlin and E. Upfal, *Balanced allocations.* SIAM Journal of Computing 29 no. 1 (1999).

[2] O. N. Feldheim and O. Gurel-Gurevich, *The power of one-retry in reducing discrepancy.* preprint arXiv:1608.02895 (2016).

[3] Chazelle, Bernard, *The discrepancy method: randomness and complexity*, Cambridge University Press, 2000.

[4] P. Maillard and E. Paquette, *Choices and intervals.* Israel Journal of Mathematics (2014).

[5] A. B. Owen, *Statistically efficient thinning of a Markov chain sampler*, preprint arXiv:1510.07727 [stat.CO].

[6] R. Dwivedi, O. N. Feldheim, O. Gurel-Gurevich and A. Ramdas, *Online thinning to reduce discrepancy.*, arXiv:version, coming soon.