

# Bucket Sort When You Know The Distribution.

David Ponnarovsky

January 20, 2023

## Abstract

We propose a new simple construction based on Tanner Codes, which yields a good LDPC code with testability query complexity of  $\Theta(n^{1-\varepsilon})$  for any  $\varepsilon > 0$ .

**The problem.** Let  $f : [0, 1] \rightarrow [0, 1]$  a fixed distribution function. Write an algorithm that sort  $n$  draws  $x_1 \dots x_n$  at linear expectation time.

**Solution.** We will define a partition of the input into a seira of  $n$  buckets  $\mathcal{B} = \{B_k = [t_k, t_{k+1}] : k \in [n]\}$  such that  $\Pr[x \in B_i] = \frac{1}{n}$  for any bucket.

$$\begin{aligned} \mathbf{E}[B_i^2] &= \mathbf{E}\left[\left(\sum_j X_{ij}\right)^2\right] \\ &= \mathbf{E}\left[\sum_{j,j'} X_{ij} X_{ij'}\right] = \sum_{j,j'} \mathbf{E}[X_{ij}] \mathbf{E}[X_{ij'}] \\ &= \sum_{j \neq j'} \mathbf{E}[X_{ij}] \mathbf{E}[X_{ij'}] + \sum_j \mathbf{E}[X_{ij}]^2 \\ &= \frac{1}{n} \binom{n}{2} + 1 = O(1) \end{aligned}$$