UNIFORM SORTING NETWORKS

SÁRA SZEPESSY

Supervisor: Balázs Maga

Abstract. A sorting network is a minimal-length sequence of adjacent swaps that transforms the identity permutation $1, 2, \ldots, n$ into the reverse permutation $n, n-1, \ldots, 1$. In my project I focused on the first step of such networks. In particular, I studied a theorem of Angel, Holroyd, Romik, and Virág (see [1]) that describes the configuration immediately after the first swap and explored the combinatorial structure underlying this result.

1. Introduction

Let S_n be the symmetric group of all n-element permutations $\sigma = (\sigma(1), \ldots, \sigma(n))$ and let $\rho = (n, n-1, \ldots, 1)$ denote the reverse permutation. For $1 \leq s \leq n-1$ we call the adjecent transposition $\tau_s = (s\ s+1) = (1, 2, \ldots, s-1, s+1, s, s+2, \ldots, n) \in S_n$ a swap at location s. An n-element sorting network is a sequence of swaps $\omega = (s_1, \ldots, s_N)$ such that $\tau_{s_1} \cdots \tau_{s_N} = \rho$, where $N = \binom{n}{2}$. Notice that N is the minimum number of swaps whose composition gives the reverse permutation ρ. We denote the set of all n-element sorting networks Ω_n.

We call the permutation $\sigma_k = \tau_{s_1} \cdots \tau_{s_k}$ the configuration at time k and $\sigma_k^{-1}(i)$ the location of particle i at time k. The function $k \mapsto \sigma_k^{-1}(i)$ is the trajectory of particle i. We refer to $s_k = s_k(\omega)$ as the kth swap location.

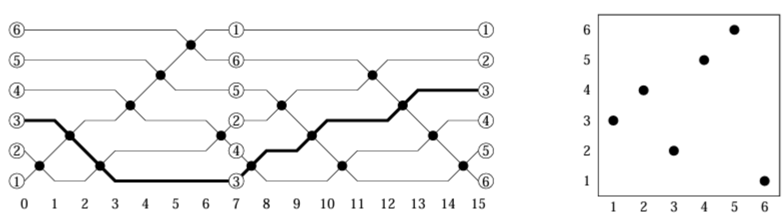

Figure 1. *Left*: the "wiring diagram" of the 6-element sorting network $\omega = (1, 2, 1, 3, 4, 5, 2, 1, 3, 2, 1, 4, 3, 2, 1)$. The swap locations are shown by the black discs. *Right*: the permutation matrix of $\sigma_7 = [3, 4, 2, 5, 6, 1]$. Figures from [1].

Let $\mathbb{P}_U = \mathbb{P}_U^n$ be the uniform probability measure on Ω_n, that is the probability measure assigning probability $\frac{1}{|\Omega_n|}$ to each n-element sorting network. We call a random sorting network $\omega_n \in \Omega_n$ choosen according to $\mathbb{P}_U$ a uniform sorting network.

As we will see later, describing the system after the first step is already a non-trivial task. In my project, I aimed to understand a neat result concerning the configuration after this first step and to explore the combinatorial structure underlying it. This investigation was motivated by a remarkable result in the area, which was only conjectured by [1]. While the result describes the system at any time, not just the first step, it served as an inspiration for my work. We begin by presenting this motivational result.

2. Motivation

A strong result in the topic was proved by Duncan Dauvergne in 2021 [2]. In his work he describes the asymptotic behavior of a uniform sorting network. For a uniform n-element sorting network one can define the global trajectory

$$T_i(t) = \frac{2\sigma^{-1}(i, \lfloor Nt \rfloor)}{n} - 1, \text{ for } t \in [0,1].$$

Here the function $T_i(t) : [0,1] \to [-1,1]$ is the rescaled trajectory of particle i, so that the discrete network is interpreted on a continuous timeline and the trajectory stays in $[-1,1]$.

Since convergence can be most conveniently expressed in terms of measures, we introduce the following random measure to encode the positions of the particles at time t:

$$\mu_t = \mu_t(\omega_n) = \frac{1}{n}\sum_{i=1}^{n} \delta(T_i(0), T_i(t)),$$

where $\delta(x,y)$ is the point measure on $\mathbb{R}^2$. The measure μ_t is called the scaled configuration at time t. The scaled configuration places atoms of weight $\frac{1}{n}$ on the positions of the ones in the rescaled permutation matrix, making a direct link between the permutation matrix and the measure μ_t.

To state the above mentioned theorem, we need to first define the time-t Archimedian measure: the measure Arch_t on the square $[-1,1]^2$ is a measure with probability function

$$f_t(x,y) := \frac{1}{2\pi}\Big(\sin^2(\pi t) + 2xy\cos(\pi t) - x^2 - y^2\Big)_+^{-1/2}.$$

Now we can state the theorem regarding the permutation matrix limits:

Theorem 1. *[D. Dauvergne [2]] Let ω_n be an n-element uniform sorting network. For all $t \in [0,1]$, $\mu_t(\omega_n) \to \mathrm{Arch}_t$ weakly in probability as $n \to \infty$.*

This result gives a strong description of the distribution of particle positions at each time t, showing that the rescaled positions follow a deterministic limit as $n \to \infty$ (see Figure 2). It is worth observing, that at time $t = \frac{1}{2}$, $f_{\frac{1}{2}}(x,y) = \frac{1}{\sqrt{1-x^2-y^2}} \cdot \chi(C)$, where $\chi(C)$ denotes the characteristic function of the circular disc $x^2+y^2 < 1$. In fact, $\mathrm{Arch}_{\frac{1}{2}}$ is the unique measure obtained by projecting the surface area measure of the unit sphere in $\mathbb{R}^3$ orthogonally onto $\mathbb{R}^2$.

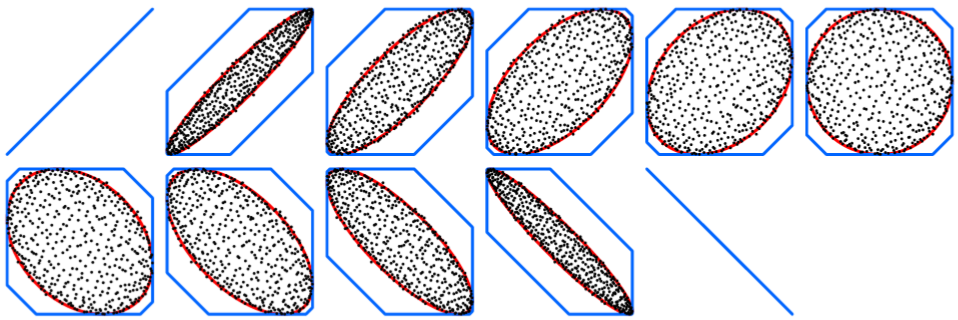

FIGURE 2. The measures $\{\mu_t : t \in \{0, 1/10, 2/10, \ldots, 1\}\}$ in a 500-element sorting network. The red ellipses are the supports of Arch_t. Figure from [1].

As a brief outlook, it is worth mentioning the connection of Theorem 1 to permutons. A permuton is a probability measure on the unit square $[0, 1]^2$ with uniform marginals. They arise as the natural limit objects of permutation sequences when $n \to \infty$. In our case, instead of introducing the measures μ_t as sums of point masses, one can construct permutons associated to the permutations at time t as follows: rather than placing atoms of weight $1/n$ on the positions of the ones in the rescaled permutation matrix, we place the uniform distribution on each of the corresponding $1/n \times 1/n$ squares.

This viewpoint also allows us to introduce a notion of convergence that is closer to the combinatorial behavior of permutations than to the traditional framework of measure convergence. Given a permuton μ, we obtain a pattern τ by sampling k points according the measure μ and reading off the order of their second coordinates from left to right. This produces the random permutation τ_μ. We denote the probability that this sampling procedure yields the pattern τ by $t(\tau, \mu) = \mathbb{P}(\tau_\mu = \tau)$. Using these pattern densities, we can define convergence for permutons: we say that $\mu_n \to \mu$ if $t(\tau, \mu_n) \longrightarrow t(\tau, \mu)$ for every pattern τ. Hoppen et al. proved that this pattern-based notion of convergence is in fact equivalent to weak convergence of measures. In particular, if we consider a random permutation of size n observed at time Nt in a uniform sorting network for each n, then this sequence of random permutations converges in probability, in the permuton sense, to the deterministic limit permuton Arch_t.

3. Configuration after the first step

In this section we state the theorem describing the configuration after the first swap.

Theorem 2 (Stationarity and semicircle law, [1]). *Let ω_n be a uniform n-element sorting network.*

(i) *The random sequence $(s_1, \dots, s_N)$ of swap locations is stationary; that is,*

$$(s_1, \dots, s_{N-1}) \stackrel{d}{=} (s_2, \dots, s_N) \quad \text{under } \mathbb{P}_U.$$

(ii) *The first swap location s_1 satisfies the convergence in distribution*

$$\frac{2s_1(\omega_n)}{n} - 1 \;\to\; Z,\; n \to \infty$$

where Z is a random variable with the semicircle law; that is, with probability density function $f_Z(y) = \frac{2}{\pi}\sqrt{1-y^2},\; y \in (-1,1)$.

Part (i) follows from a straightforward bijection on swap sequences. Part (ii), however, requires some background on Young tableaux and uses a more elaborate bijection between standard Young tableaux and the set of n-element sorting networks Ω_n. We will not present the full proof of part (ii) here. Instead, we introduce the formal definition of Young tableaux and outline the main ideas of the poof.

First we start by introducing some basic notation. Let $N \in \mathbb{N}_+$ and $\lambda = (\lambda_1, \dots, \lambda_k) \in \mathbb{N}_+^k$, such that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$ and $|\lambda| := \sum_{i=1}^k \lambda_i = N$. We call such λ a partition of N and denote it by $\lambda \in \mathrm{Part}(N)$. The Young diagram associated to λ is the set $\{(i,j) \in \mathbb{N}_+^2 : 1 \leq i \leq k,\ 1 \leq j \leq \lambda_i\}$. We represent each pair (i,j) by a cell, placing $(1,1)$ in the upper-left corner and letting the first and second coordinates increase to the right and downward respectively. The staircase diagram $(n-1, \dots, 1)$ is denoted by $\triangle_n$.

We can now define the Young tableaux as follows. Let $\lambda \in \mathrm{Part}(N)$ be a partition of N, then the Young tableau of shape λ is an assignment of positive integers, called entries, to the cells of λ, such that every row and column of the diagram contains increasing sequences of numbers. A standard Young tableau (SYT) is a Young tableau in which the numbers assigned to all cells are $1, 2, \dots, N$. We denote the set of SYT of shape λ by $SYT(\lambda)$ and its cardinality by $d(\lambda) = |SYT(\lambda)|$.

The number of SYT of shape λ can be explicitly computed by using the following Hook-formula:

Theorem 3 (Hook-formula). *Let $\lambda \in Part(N)$ be a partition of N. For each cell $(i,j) \in \lambda$, let $h_{i,j}(\lambda) := \lambda_i - j + \lambda'_j - i + 1$ be the hook-length of (i,j) in λ, where λ'_j denotes the length of the j-th column of λ. Then the number of standard Young tableaux of shape λ is*

$$d(\lambda) = \frac{|\lambda|!}{\prod_{(i,j)\in\lambda} h_{i,j}(\lambda)}.$$

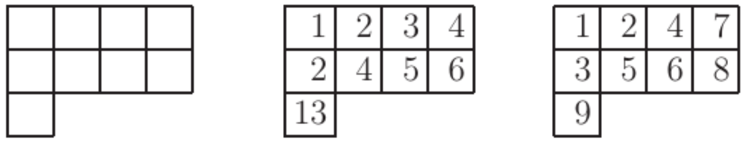

Figure 3. The Young diagram $(4, 4, 1)$, a Young tableau and a standard Young tableau. Figure from [1].

The connection between Young tableaux and sorting networks lies in a remarkable result of Stanley, who proved

$$\#\Omega_n = d(\triangle_n) = \#SYT(\triangle_n),$$

that is, the number of n-element sorting networks equals the number of standard Young tableaux of shape $\triangle_n$. Stanley was able to compute these quantities and recognize that they coincide, however, a direct bijection between the elements of the two sets was only constructed later by Edelman and Greene. This bijection is a key tool in the proof of part (ii), as it reduces the convergence problem to the combinatorial structure of Young tableaux.

References

[1] O. Angel, A. Holroyd, D. Romik, and B. Virág (2007), *Random sorting networks*, Adv. Math., 215(2):839–868.
https://arxiv.org/pdf/math/0609538.pdf

[2] D. Dauvergne (2022), *The Archimedean limit of random sorting networks*, J. Amer. Math. Soc., 35:1215–1267.
https://arxiv.org/pdf/1802.08934.pdf

Declaration of AI usage

I used the AI language model ChatGPT to rephrase and polish some of the text in this project. Its use was limited to improving clarity and wording.