Random Permutations using Switching Networks

by Artur Czumaj

Presented by Jan Musílek

Definitions

A **switching network** $\mathfrak{N}$ of depth $\mathfrak{d}$ is a layered network with $\mathfrak{d}+1$ layers, each layer having n nodes. The nodes between consecutive layers are connected by disjoint *switches*.

A **switch** between two input nodes at layer l and two output nodes at layer $l+1$ takes the two inputs and either transposes them (if the switch is active) or leaves them unchanged (if the switch is inactive).

A **matching** of $\{1,\ldots,n\}$ is a set of pairs $\{i,j\} \subseteq \{1,\ldots,n\}$ with $i \neq j$ such that no element from $\{1,\ldots,n\}$ appears in more than one pair. A **perfect matching** of $\{1,\ldots,n\}$ is a matching of $\{1,\ldots,n\}$ of size exactly $\frac{n}{2}$.

Let $\mathfrak{M}^{\mathfrak{d}}$ be the set of all sequences $(M_0,\ldots,M_{\mathfrak{d}-1})$ such that each M_t is a perfect matching of $\{1,\ldots,n\}$. For a given $M^{\mathfrak{d}} = (M_0,\ldots,M_{\mathfrak{d}-1}) \in \mathfrak{M}^{\mathfrak{d}}$, we define a switching network $\mathfrak{N}$ of depth $\mathfrak{d}$ so that the switches between layers i and $i+1$ are determined by the pairs of matching M_i. Every layered network $\mathfrak{N}$ corresponding to $M^{\mathfrak{d}}$, $M^{\mathfrak{d}} = (M_0,\ldots,M_{\mathfrak{d}-1}) \in \mathfrak{M}^{\mathfrak{d}}$ defines in a natural way a stochastic process (Markov chain) $(\mathcal{Q})_{t=0}^{\mathfrak{d}}$ on state space of all permutation of $\{1,\ldots,n\}$.

A **k-partial n-permutation** is any sequence $\langle x_0,\ldots,x_{n-1}\rangle$ consisting of k 0s and $n-k$ distinct elements from $1,\ldots,n-k$. The set of all k-partial n-permutations is denoted by $\mathbb{S}_{n,k}$. Observe that $|\mathbb{S}_{n,k}| = \frac{n!}{k!}$.

Let $\mathfrak{MC}$ be a discrete-time Markov chain with a finite state space Ω and a unique stationary distribution $\mu_{\mathfrak{MC}}$. For any random variable X, let $\mathcal{L}(X)$ denote the probability distribution of X, and let $\mathcal{L}(\mathcal{Q}_t|\mathcal{Q}_0=\omega)$ denote the probability distribution of $\mathcal{Q}_t$ given that $\mathcal{Q}_0=\omega$. The **total variation distance** between two probability distributions $\mathcal{X}$ and $\mathcal{Y}$ over the same finite domain Ω is defined as:

$$d_{TV}(\mathcal{X},\mathcal{Y}) = \max_{S\subseteq\Omega} |\mathbf{Pr}_{\mathcal{X}}[S] - \mathbf{Pr}_{\mathcal{Y}}[S]| = \frac{1}{2}\sum_{\omega\in\Omega} |\mathbf{Pr}_{\mathcal{X}}[\omega] - \mathbf{Pr}_{\mathcal{Y}}[\omega]|$$

We define the total variation distance after t steps of $\mathfrak{MC}$ with respect to initial state $\omega \in \Omega$ as $\Delta_{\omega}^{\mathfrak{MC}}(t) = d_{TV}(\mathcal{L}(\mathcal{Q}_t|\mathcal{Q}_0=\omega), \mu_{\mathfrak{MC}})$. Then, the standard measure of the convergence of a Markov chain $\mathfrak{MC}$ to its stationary distribution $\mu_{\mathfrak{MC}}$ is the *mixing time*, denoted by $\tau_{\mathfrak{MC}}(\varepsilon)$, which is defined as $\tau_{\mathfrak{MC}}(\varepsilon) = \min\{T \in \mathbb{N} : \forall\omega\in\Omega \forall t \geq T \Delta_{\omega}^{\mathfrak{MC}}(t) \leq \varepsilon\}$.

A **coupling** for a Markov chain $\mathfrak{MC} = (\mathcal{Q}_t)_{t\in\mathbb{N}}$ on state space Ω is a stochastic process $(\mathcal{X}_t,\mathcal{Y}_t)_{t\in\mathbb{N}}$ on $\Omega\times\Omega$ such that each of $(\mathcal{X}_t)_{t\in\mathbb{N}}$, $(\mathcal{Y}_t)_{t\in\mathbb{N}}$ considered independently, is a faithful copy of $\mathfrak{MC}$.

Lemma 2.1 (Delayed Path Coupling Lemma): Let $\mathfrak{MC} = (\mathcal{X}_t)_{t\in\mathbb{N}}$ be a discrete-time Markov chain with a finite state space Ω. Let Γ be any subset of $\Omega\times\Omega$. Suppose that there is an integer D such that for every $(\mathcal{X},\mathcal{Y}) \in \Omega\times\Omega$ there exists a sequence $\mathcal{X} = \Lambda_0,\Lambda_1,\ldots,\Lambda_r = \mathcal{Y}$, where $(\Lambda_i,\Lambda_{i+1}) \in \Gamma$ for $0 \leq i < r$, and $r \leq D$. If there exists a coupling $(\mathcal{X}_t,\mathcal{Y}_t)_{t\in\mathbb{N}}$ for $\mathfrak{MC}$ such that for some $T\in\mathbb{N}$, for all $(\mathcal{X},\mathcal{Y})\in\Gamma$, it holds that $\mathbf{Pr}[\mathcal{X}_T \neq \mathcal{Y}_T|(\mathcal{X}_0,\mathcal{Y}_0) = (\mathcal{X},\mathcal{Y})] \leq \frac{\varepsilon}{D}$, then

$$||\mathcal{L}(\mathcal{X}_T|\mathcal{X}_0=\mathcal{X}) - \mathcal{L}(\mathcal{Y}_T|\mathcal{Y}_0=\mathcal{Y})|| \leq \varepsilon$$

for every $(\mathcal{X},\mathcal{Y})\in\Omega\times\Omega$. In particular, $\tau_{\mathfrak{MC}}(\varepsilon/2) \leq T$.

Random walks on expanders

Let us consider a switching network $\mathfrak{N}$ of depth $\mathfrak{d}$ that corresponds to $M^{\mathfrak{d}} = (M_0,\ldots,M_{\mathfrak{d}-1} \in \mathfrak{M}^{\mathfrak{d}}$. Define an **$\langle l,r\rangle$-truncate** of $\mathfrak{N}$ to be the multigraph $G=(V,E)$ on vertex set $V=\{1,\ldots,n\}$ with the edge set E consisting of all pairs (i,j) for which there is a path from i to j in the netwok induced by

$M_l, M_{l+1}, \ldots, M_{l+r-1}$; if there are s paths from i to j then we have s edges (i,j) in E. Notice that G is 2^r-regular and it has selfloops.

Lemma 2.2: For every $r \geq 4$, there is a constant a, $0 < a < 1$, such that for almost every switching network $\mathfrak{N}$ (all but at most a $\frac{1}{n^2}$ fraction), for every $0 \leq l \leq \mathfrak{d} - r$, the $\langle l, r\rangle$-truncate G of $\mathfrak{N}$ is an $(1-a)$-expander.

Let us call a switching network $\mathfrak{N}$ to be $\mathfrak{good}$ if there is a constant r and another positive constant a such that every $\langle i \cdot r, r\rangle$-truncate is a $(1-a)$-expander, $0 \leq i < \mathfrak{d}/r$.

Proposition 2.3: Almost all (all but a $\frac{1}{n^2}$ fraction) switching networks of logarithmic depth are $\mathfrak{good}$.

Proposition 2.4: One can explicitly construct a $\mathfrak{good}$ switching network $\mathfrak{N}$.

Lemma 3.1: There is a constant c such that if we run the random shuffling process for $c \log_2 n$ steps with all switches set at random then the probability that two fundamental trees will be build is at least $1 - n^{-3}$.

Lemma 3.2: Let us fix any two sets of ρ disjoint positions for the leaves of the fundamental trees. There is a constant c such that if we run the random shuffling process for $c \log_2 n$ steps with all switches set at random then the probability that there is a fundamental matching is at least $1 - n^{-3}$.

Main results

Theorem 3.4: Let $k = \Omega(n)$. Let $\mathfrak{N}$ be a $\mathfrak{good}$ switching network of depth $\mathfrak{d}$ with $\mathfrak{d} \geq c \log n$, for a sufficiently large constant c. Then $\mathfrak{N}$ generates random k-partial n-permutations almost uniformly. That is, for any positive constant c_1, if $\pi \in \mathbb{S}_{n,k}$, is the permutation generated by switching network $\mathfrak{N}$ on an arbitrary input from $\mathbb{S}_{n,k}$ and μ is the uniform distribution over $\mathbb{S}_{n,k}$, then $d_{TV}(\mathcal{L}(\pi), \mu) \leq \mathcal{O}(n^{-c_1})$.

Theorem 3.5: For any $\varepsilon > 0$, almost every (all but a $\mathcal{O}(n^{-2})$ fraction) switching network $\mathfrak{N}$ of depth $\mathfrak{d}$ ($\mathfrak{d} \geq c \log n$) almost randomly permutes any set of $(1-\varepsilon)n$ elements.

Theorem 3.6: For any $\varepsilon > 0$, there is an explicit switching network $\mathfrak{N}$ of depth $\mathfrak{d}$ ($\mathfrak{d} \geq c \log n$) that almost randomly permutes any set of $(1-\varepsilon)n$ elements.

Theorem 3.8: Let c_2 be an arbitrary constant. There is an explicit switching network $\mathfrak{N}$ of depth $\mathcal{O}(\log^2 n)$ and with $\mathcal{O}(n \log n)$ switches such that if $\pi \in \mathbb{S}_n$ denotes the permutation generated by $\mathfrak{N}$ and μ is the uniform distribution over $\mathbb{S}_n$, then $d_{TV}(\mathcal{L}(\pi), \mu) \leq \mathcal{O}(n^{-c_2})$.