Lecture-10: Kurtz's theorem: proof

1 Kurtz's theorem

Assumption 1.1. Consider a density-dependent family of CTMCs $((X^N : \Omega \to (\mathcal{Z}^N)^{\mathbb{R}}_+) : N \in \mathbb{N})$. For each N, state $x \in \mathcal{Z}^N$, empirical distribution of states $a(x) \in \mathcal{M}(\mathcal{Z})$, and $z, w \in \mathcal{Z}$, the transition rate $Q^{X^N_n}_{z,w} : \mathcal{M}(\mathcal{Z}) \to \mathbb{R}$ for a single particle n is Lipschitz continuous in the empirical distribution.

Lemma 1.2. *Let $X : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ be a random process and $\alpha \in \mathbb{R}^N$ are constants, then for any finite $T \in \mathbb{R}_+, \sup_{i\in[n]} |\alpha_i| \leqslant C$, and $x > 0$, we have*

$$\left\{\sup_{t\in[0,T]} \left\|\sum_{i=1}^n \alpha_i X_i(t)\right\| > x\right\} \subseteq \bigcup_{i=1}^n \left\{\sup_{t\in[0,T]} \|X_i(t)\| > \frac{x}{nC}\right\}.$$

Proof. Let $x > 0$, then we observe that $\sup_{t\in[0,T]} \|\sum_{i=1}^n \alpha_i X_i(t)\| \leqslant C \sum_{i=1}^n \sup_{t\in[0,T]} \|X_i(t)\|$. It follows that

$$\bigcap_{i=1}^n \left\{\sup_{t\in[0,T]} \|X_i(t)\| \leqslant \frac{x}{nC}\right\} \subseteq \left\{\sup_{t\in[0,T]} \left\|\sum_{i=1}^n \alpha_i X_i(t)\right\| \leqslant x\right\}.$$

The result follows from taking complements on both sides. □

Definition 1.3. We define $h : [-1, \infty) \to \mathbb{R}_+$ as $h(t) \triangleq (1 + t)\ln(1 + t) - t$ for all $t \geqslant -1$.

Lemma 1.4 (Gronwall inequality). *Consider a bounded function $u : [0, T] \to \mathbb{R}$ such that $u(t) \leqslant a + b\int_0^t u(s)ds$ for all $t \in [0, T]$. Then $u(t) \leqslant ae^{bt}$ for all $t \in [0, T]$.*

Proof. We define $v : [0, T] \to \mathbb{R}$ as $v(t) = e^{-bt}\int_0^t u(s)ds$ for all $t \in [0, T]$. Then, we have for all $t \in [0, T]$

$$\frac{d}{dt}v(t) = e^{-bt}\Big(u(t) - b\int_0^t u(s)ds\Big) \leqslant ae^{-bt}.$$

Recalling that $v(0) = 0$ and integrating on both sides, we get $v(t) \leqslant \frac{a}{b}(1-e^{-bt})$ for all $t \in [0, T]$. Multiplying with be^{bt} followed by adding a on both sides, we get the result. □

Theorem 1.5 (Kurtz). *Consider an N interacting particle CTMC $X^N : \Omega \to \mathcal{X}^{\mathbb{R}_+}$ with associated empirical distribution $A^N : \Omega \to \mathcal{M}(\mathcal{Z})^{\mathbb{R}_+}$ such that $(A^N : N \in \mathbb{N})$ is a density-dependent family of CTMCs with mean field model*

$$\dot{a}(t) = f(a(t)), \quad t \in \mathbb{R}_+.$$

If (a) Assumption 1.1 holds such that $f : \mathcal{M}(\mathcal{Z}) \to \mathbb{R}^{\mathcal{Z}}$ is M-Lipschitz and (b) $\lim_{N\to\infty} A^N(0) \to a(0)$ in distribution, then for a fixed $T, \epsilon > 0$ and $h : [-1, \infty) \to \mathbb{R}_+$ defined in Definition 1.3, we have

$$P\left\{d_T(A^N, a) > \epsilon\right\} \leqslant P\left\{\left\|A^N(0) - a(0)\right\| > \frac{\epsilon}{2e^{MT}}\right\} + Ce^{-NT\bar{Q}h(\frac{\epsilon}{C_1C\bar{Q}2e^{MT}})},$$

where $\sup_{w,z\in\mathcal{Z}} \|e_w - e_z\| \leqslant C_1$ and $|\mathcal{Z}|(|\mathcal{Z}| - 1) \leqslant C$.

Proof. We will show this using following steps.

Step 1. **Time change.** Let $N_{z,w} : \Omega \to \mathbb{Z}_+^{\mathbb{R}_+}$ be independent Poisson counting processes with unit rates for all $z, w \in \mathcal{Z}$. We define $\lambda_{z,w} : \mathbb{R}_+ \to \mathbb{R}_+$ as $\lambda_{zw}(s) \triangleq N A_z^N(s) Q_{zw}^{X_n^N}(A^N(s))$ as the instantaneous rate of transition from state $A^N(s)$ to $A^N(s) - e_z + e_w$. Then, the measure $\Lambda_{z,w}(0,t] \triangleq \int_0^t \lambda_{z,w}(s)ds$ defined for all $t \in \mathbb{R}_+$, is the mean number of transitions for particles in state z to state w. Recall that $f(a) = \sum_{z\in\mathcal{Z}} \sum_{w\neq z} \frac{1}{N}(e_w - e_z) N a_z Q_{zw}^{X_n^N}(a)$, and thus $\sum_{z\in\mathcal{Z}} \sum_{w\neq z} \frac{1}{N}(e_w - e_z)\lambda_{z,w}(s) = f(A_z^N(s))$. Thus, we can write

$$A^N(t) = A^N(0) + \int_0^t f(A_z^N(s))ds + \sum_{z\in\mathcal{Z}} \sum_{w\neq z} \frac{1}{N}(e_w - e_z)\Big[N_{zw}(\Lambda_{z,w}(0,t]) - \Lambda_{z,w}(0,t]\Big].$$

Step 2. **Lipschitz property.** Using triangle inequality and M-Lipschitz property of mean-field model f, we can write

$$\left\|A^N(t) - a(t)\right\| \leqslant \left\|A^N(0) - a(0)\right\| + M\int_0^t \left\|A_z^N(s)) - a(s)\right\| ds + \left\|\sum_{z\in\mathcal{Z}} \sum_{w\neq z} \frac{1}{N}(e_w - e_z)\bar{N}_{zw}(\Lambda_{z,w}(0,t])\right\|.$$

Step 3. **Bounding martingale error.** Let $r \triangleq |\mathcal{Z}|$ and $\|e_w - e_z\| \leqslant C_1$, then from Lemma 1.2 and union bound, we have

$$P\left\{\sup_{t\in[0,T]} \left\|\sum_{z\in\mathcal{Z}} \sum_{w\neq z} \frac{1}{N}(e_w - e_z)\bar{N}_{zw}(\Lambda_{z,w}(0,t])\right\| > \epsilon\right\} \leqslant \sum_{z\in\mathcal{Z}} \sum_{w\neq z} P\left\{\sup_{t\in[0,T]} \left\|\bar{N}_{zw}(\Lambda_{z,w}(0,t])\right\| > \frac{N\epsilon}{C_1 r(r-1)}\right\}. \tag{1}$$

From Assumption 1.1, it follows that $\max_{z,w,a} Q_{z,w}^{X_1^N} \leqslant \bar{Q}$. From the definition of Λ_{zw}, we get $\Lambda_{z,w}(0,t] = \int_0^t N A_z^N(s) Q_{z,w}^{X_1^N}(A^N(s))ds \leqslant N\bar{Q}t$. It follows that for all $w, z \in \mathcal{Z}$

$$\sup_{t\in[0,N\bar{Q}T]} \bar{N}_{zw}(t) \geqslant \sup_{t\in[0,T]} \bar{N}_{zw}(\Lambda_{zw}(0,t]).$$

Hence, $P\left\{\sup_{t\in[0,T]} \bar{N}_{zw}(\Lambda_{zw}(0,t]) > x\right\} \leqslant P\left\{\sup_{t\in[0,N\bar{Q}T]} \bar{N}_{zw}(t) > x\right\}$ for all $x > 0$. From Lemma **??**, we obtain

$$\sum_{z\in\mathcal{Z}} \sum_{w\neq z} P\left\{\sup_{t\in[0,N\bar{Q}T]} \left\|\bar{N}_{zw}(t)\right\| > \frac{N\epsilon}{C_1 r(r-1)}\right\} \leqslant 2e^{-N\bar{Q}Th(\frac{\epsilon}{r(r-1)C_1\bar{Q}T})}.$$

Step 4. **Applying Gronwall.** Applying union bound and (1), we observe that

$$\begin{aligned}
&P\left\{\sup_{t\in[0,T]} \left|\left\|A^N(t) - a(t)\right\| - M\int_0^t \left\|A^N(s)) - a(s)\right\| ds\right| > 2\epsilon\right\} \\
&\leqslant P\left\{\sup_{t\in[0,T]} \left|\left\|A^N(0) - a(0)\right\| + \left\|\sum_{z\in\mathcal{Z}} \sum_{w\neq z} \frac{1}{N}(e_w - e_z)\bar{N}_{zw}(\Lambda_{z,w}(0,t])\right\|\right| > 2\epsilon\right\} \\
&\leqslant P\left\{\left\|A^N(0) - a(0)\right\| > \epsilon\right\} + 2r(r-1)e^{-N\bar{Q}Th(\frac{\epsilon}{r(r-1)C_1\bar{Q}T})}.
\end{aligned}$$

Recall that $\left\|A^N(s) - a(s)\right\|$ is bounded on $[0,T]$, and thus we can apply Gronwall inequality to $u(t) \triangleq \left\|A^N(t) - a(t)\right\|$ for all $t \in \mathbb{R}_+$. In particular, if $\sup_{t\in[0,T]}(u(t) - M\int_0^t u(s)ds) \leqslant 2\epsilon$, then $\sup_{t\in[0,T]} u(t)e^{-MT} \leqslant 2\epsilon$. Therefore, we obtain

$$P\left\{\sup_{t\in[0,T]} \left\|A^N(t) - a(t)\right\| e^{-MT} > 2\epsilon\right\} \leqslant P\left\{\left\|A^N(0) - a(0)\right\| > \epsilon\right\} + 2r(r-1)e^{-N\bar{Q}Th(\frac{\epsilon}{r(r-1)C_1\bar{Q}T})}.$$

The result follows from the definition of $a(0)$ and taking $\epsilon' = 2\epsilon e^{MT}$. □

2 Asymptotic independence

Theorem 2.1 (DeFinetti). *If $X^N(0) : \Omega \to \mathcal{Z}^N$ is exchangeable for all $N \in \mathbb{N}$ and $\lim_{N\to\infty} A^N(0) = a(0)$, then for all finite $F \subseteq \mathbb{N}$ and $z \in \mathcal{Z}^F$, we have*

$$\lim_{N\to\infty} P \cap_{n\in F} \left\{X_n^N(0) = z_n\right\} = P \cap_{n\in F} \left\{X_n^\infty(0) = z_n\right\}.$$

Remark 1 (Propagation of chaos)*.* The above theorem is called *Boltzmann property* or *chaos*. The theorem states that initial chaos propagates.

Theorem 2.2. *Fix $t \in \mathbb{R}_+$ and $F \subseteq [N]$ finite. Let $X^N(0) : \Omega \to \mathcal{X} = Z^N$ be an exchangeable random vector, $\lim_{N\to\infty} A^N(0) \to a(0)$ in distribution, and Assumption 1.1 holds. Then $\lim_{N\to\infty}(X_n^N(t), n \in F) = (U_n(t), n \in F)$ in distribution for $U(t) : \Omega \to \mathbb{R}^N$* i.i.d. *with distribution $a(t)$.*

Proof. Since $X^N(0)$ is exchangeable and the marginal evolution of each particle is identical and depends only on the empirical distribution A^N, it follows that that $X^N(t)$ is exchangeable at all times $t \in \mathbb{R}_+$. Let $\Phi_n \in C_b(\mathcal{Z})$ be any bounded continuous function for all $n \in F$. Then it suffices to show that

$$\lim_{N\to\infty} \mathbb{E} \prod_{n\in F} \Phi_n(X_n^N(t)) = \prod_{n\in F} \mathbb{E}\Phi_n(U_n(t)) = \prod_{n\in F} \langle \Phi_n, a(t)\rangle,$$

where $\langle a, b\rangle \triangleq \sum_{z\in\mathcal{Z}} b_z z_z$ for all $b \in C_b(\mathcal{Z})$ and $a \in \mathcal{M}(\mathcal{Z})$. We will show this for $|F| \in \{1, 2\}$ and it follows for any finite F by induction.

$|F|$ =1. Without loss of generality, we take $F = \{1\}$. From the exchangeability of $X^N(t)$, we get

$$\mathbb{E}\Phi_1(X_1^N(t)) = \mathbb{E}\Big[\frac{1}{N}\sum_{n=1}^N \Phi_1(X_n^N(t))\Big] = \mathbb{E}\left\langle \Phi_1, A^N(t)\right\rangle. \tag{2}$$

Since $A^N(t) \to a(t)$ in distribution, we have $\lim_{N\to\infty} \mathbb{E}\left\langle \Phi_1, A^N(t)\right\rangle = \langle \Phi_1, a(t)\rangle = \mathbb{E}\Phi_1(U_1)$.

$|F|$ =2. Without loss of generality, we take $F = \{1, 2\}$. We look at the following difference

$$\mathbb{E}\prod_{n=1}^2 \Phi_n(X_n^N(t)) - \mathbb{E}\prod_{n=1}^2 \left\langle \Phi_n, A^N(t)\right\rangle + \mathbb{E}\prod_{n=1}^2 \left\langle \Phi_n, A^N(t)\right\rangle - \prod_{n=1}^2 \langle \Phi_n, a(t)\rangle.$$

We observe that the RHS of the above equation is sum of two difference. We use exchangeability property of X^N to write the following equation

$$\mathbb{E}[\Phi_1(X_1^N(t))\Phi_2(X_2^N(t))] = \frac{1}{N(N-1)}\mathbb{E}\sum_{n\neq m} \Phi_1(X_n^N(t))\Phi_2(X_m^N(t)). \tag{3}$$

From exchangeability property of X^N, adapting (2) for Φ_1, Φ_2, we also write the following,

$$\mathbb{E}\prod_{n=1}^2 \left\langle \Phi_n, A^N(t)\right\rangle = \frac{1}{N^2}\mathbb{E}\Big[\sum_{n=1}^N \Phi_1(X_n^N(t))\Phi_2(X_n^N(t)) + \sum_{n\neq m}^N \Phi_1(X_n^N(t))\Phi_2(X_m^N(t))\Big]. \tag{4}$$

Since Φ_1, Φ_2 are bounded, we define $v \triangleq \sup_{z\in\mathcal{Z}} |\Phi_1(z)| \vee |\Phi_2(z)| < \infty$. Taking the difference of (3) and (4), we upper bound the first difference as

$$\frac{v^2}{N} + \Big(1 - \frac{(N-1)}{N}\Big)\mathbb{E}\prod_{n=1}^2 \Phi_n(X_n^N(t)) \leqslant \frac{2v^2}{N}.$$

We can rewrite the second difference as

$$\sum_{z\in\mathcal{Z}} \Phi_1(z)\Phi_2(z)(\mathbb{E}(A_z^N(t))^2 - a_z(t)^2) + \sum_{w\neq z} \Phi_1(z)\Phi_2(w)(\mathbb{E}A_z^N(t)A_w^N(t) - a_z(t)a_w(t)).$$

For $w \neq z$, we can write the difference

$$\mathbb{E}A_z^N(t)A_w^N(t) - a_z(t)a_w(t) = \mathbb{E}\Big[(A_z^N(t) - a_z(t))A_w^N(t)\Big] + a_z(t)\mathbb{E}(A_w^N(t) - a_w(t)).$$

Since $\lim_{N\to\infty} A^N(t) = a(t)$ in distribution, $A^N(t), a(t) \in \mathcal{M}(\mathcal{Z})$, and v is the upper bound on $|\Phi_1|, |\Phi_2|$ we can upper bound the second difference as

$$\sum_{z\in\mathcal{Z}} v^2\mathbb{E}(A_z^N(t) - a_z(t))^2 + 2v^2 \sum_{z\in\mathcal{Z}} \mathbb{E}(A_z^N(t) - a_z(t)).$$

Taking limit $N \to \infty$, we get the result.

$\square$