On the LAN and LAMN Properties for Mean-Field Model of Interacting Neurons

Aurélien Velleret

LaMME, UEVE, PARIS-SACLAY, Evry, Val-d'Essonne

aurelien.velleret@nsup.org

10/06/2026

Introduction

LAN and LAMN: general definition
- Definitions

Propagation of Chaos and LAN
- The finite model
- The limit model
- LAN

Efficiency of the MLE

Conditional Propagation of Chaos and LAMN
- The finite model
- The limit model
- LAMN

Parametric estimation of the interaction rate (of $\theta^* \in \mathbb{R}^d$)

Temporal Process : system of N neurons,

Interaction: spike of neuron $i \in [\![1, N]\!]$,

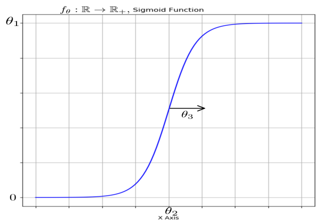
as a function of its current state $X_i^{N,\theta^*}(t)$

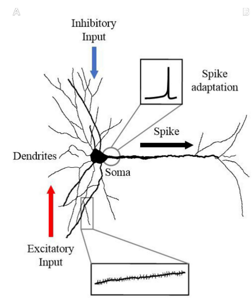
with continuous observation of $X^N = (X_1^{N,\theta^*}, \dots, X_N^{N,\theta^*})$ on $[0, t]$.

Asymptotic development as $N \to \infty$ of the likelihood:

LAN/LAMN (in the wave, new for jump systems)

- ⋆ punctual spike $(\pi_i)_i \perp\!\!\!\perp$ Poisson processes
- ⋆ spiking rate $(f_{\theta^*}(x))_{x \in \mathbb{R}_+}$





Efficiency of likelihood-based estimators

Focus on $\hat{\theta}^N_t = \text{argmax}_{\theta \in \Theta} L_t^{\theta/\theta^*}(X^N)$, Max. Likelihood Estimator (MLE)

⋆ Consistency:

convergence of $\hat{\theta}^N_t$ to θ^* under $\mathbb{P}^N_{\theta^*}$ as $N \to \infty$.

⋆ Asymptotic normality, with minimal variance, under LAN:

$$\sqrt{N}(\hat{\theta}^N_t - \theta^*) \xrightarrow{\mathcal{L}} (I_t^{\theta^*})^{-1/2}\mathcal{N}, \quad \mathcal{N} \sim \mathcal{N}(0, I_d) .$$

⋆ Local asymptotic minimax efficiency, under LAN:

$$\limsup_{N\to\infty} \sup_{|\theta-\theta^*|\leq\delta} \mathbb{E}_{\mathbb{P}^N_\theta}\left[w\left(N^{1/2}(I_t^{\theta^*})^{1/2}(\hat{\theta}^N_t - \theta)\right)\right] \to \mathbb{E}[w(\mathcal{N})]$$

with $\delta \to 0$, w polynomial loss function.

⋆ Global asymptotic minimax efficiency, under LAN:

$$\mathcal{R}^N_w(\hat{\theta}^N_t; \Theta_0) = \inf_{\tilde{\theta}_N} \mathcal{R}^N_w(\tilde{\theta}_N; \Theta_0)\,(1 + o(1)) \quad \text{as } N \to \infty\,,$$

with the risk $\mathcal{R}^N_w$ associated with w, defined as:

$$\mathcal{R}^N_w(\tilde{\theta}_N; \Theta_0) = \sup_{\theta\in\Theta_0} \mathbb{E}_{\mathbb{P}^N_\theta}\left[w\left(N^{1/2}(I_t^{\theta})^{1/2}(\tilde{\theta}_N - \theta)\right)\right] .$$

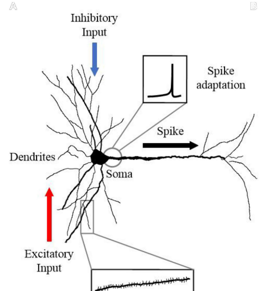
Internal dynamics $(X_t^{i,N})_{t\geq 0}$ of a specific **neuron** i,
$X_t^{i,N}$: its "**potential**" at time t:

$$X_t^{i,N} = X_0^{i,N} - \alpha \int_0^t X_s^{i,N} \mathrm{d}s + \iint_{[0,t]\times\mathbb{R}_+} \left[- X_{s-}^{i,N} \right] \mathbf{1}_{\{z \leq f_\theta(-X_{s-}^{i,N})\}} \pi_i(\mathrm{d}s, \mathrm{d}z) + \sum_{j\neq i} \iint_{[0,t]\times\mathbb{R}_+} \frac{m}{N} \mathbf{1}_{\{z \leq f_\theta(-X_{s-}^{j,N})\}} \pi_j(\mathrm{d}s, \mathrm{d}z).$$

and the associated reduced model, the **McKean-Vlasov** dynamics $\bar{\mu} = \mathcal{L}(\bar{X}(t))$

$$\bar{X}_t = \ldots + \int_0^t m \, \mathbb{E}[f_\theta(\bar{X}_s)] \mathrm{d}s$$

- ⋆ punctual spike — $(\pi_i)_i \perp\!\!\!\perp$ Poisson processes
- ⋆ spiking rate — $(f_{\theta^*}(x))_{x\in\mathbb{R}_+}$
- ⋆ interaction effect — m
- ⋆ exponential decay — $\alpha > 0$.



Outline

The talk is based on two ongoing joint works

- LAN property for mean field model of interacting neurons, with Aline Duarté and Dasha Loukianova.

  on arXiv:2605.03698

- LAMN property and conditional propagation of chaos for mean field model of interacting neurons, with Xavier Erny, Eva Löcherbach, Dasha Loukianova.

Introduction

LAN and LAMN: general definition
- Definitions

Propagation of Chaos and LAN
- The finite model
- The limit model
- LAN

Efficiency of the MLE

Conditional Propagation of Chaos and LAMN
- The finite model
- The limit model
- LAMN

Local Asymptotic Normality (LAN)

A sequence of statistical experiments $(\Omega^N, \mathcal{F}^N, \mathbb{P}^\theta_N)$, $\theta \in \Theta \subset \mathbb{R}^d$, is LAN at $\theta \in \Theta$ if

$$\log \frac{d\mathbb{P}^{\theta+r_N h}_N}{d\mathbb{P}^\theta_N} = h^\top \Delta^{N,\theta} - \frac{1}{2} h^\top I^{N,\theta} h + o_{\mathbb{P}^\theta_N}(1), \quad \forall h \in \mathbb{R}^d,$$

for a sequence of norming matrices $r_N \in \mathcal{M}^{d\times d}_+$, $\lim_{N\to\infty} \|r_N\| = 0$, of random vectors $\Delta^{N,\theta} \in \mathbb{R}^d$, matrices $I^{N,\theta} \in \mathcal{M}^{d\times d}_+$ such that

$I^{N,\theta} \to I^\theta$ in $\mathbb{P}^\theta_N$-probability

$\Delta^{N,\theta} \xrightarrow{\mathcal{L}} \mathcal{N}_d(0, I(\theta))$ under $\mathbb{P}^\theta_N$.

with $I(\theta) \in \mathcal{M}^{d\times d}_+$, *positive definite*, the Fisher information (deterministic).

The likelihood locally behaves like that of a Gaussian shift.

Gaussian shift: $X_1, \ldots X_n$ i.i.d. $\mathcal{N}(\theta, \sigma^2)$

$$\log \frac{\mathrm{d}\mathbb{P}_N^{\theta+\frac{h}{\sqrt{N}}}}{\mathrm{d}\mathbb{P}_N^{\theta}} = \frac{h}{\sqrt{N}} \sum_{i=1}^{N} \frac{X_i - \theta}{\sigma^2} - \frac{h^2}{2\sigma^2}.$$

$$I^{N,\theta} = I^{\theta} = \frac{1}{\sigma^2}, \quad \Delta^{N,\theta} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{X_i - \theta}{\sigma^2} \sim \mathcal{N}(0, I^{\theta}), \quad r_N = \frac{1}{\sqrt{N}}$$

Recall LAN:

$$\log \frac{d\mathbb{P}_N^{\theta + r_N h}}{d\mathbb{P}_N^{\theta}} = h^\top \Delta^{N,\theta} - \frac{1}{2} h^\top I^{N,\theta} h + o_{\mathbb{P}_N^\theta}(1), \quad \forall h \in \mathbb{R}^d,$$

where

$$I^{N,\theta} \to I^{\theta} \ \text{ in } \mathbb{P}_N^{\theta}; \quad \Delta^{N,\theta} \xrightarrow{\mathcal{L}} \mathcal{N}_d(0, I(\theta)) \ \text{ under } \mathbb{P}_N^{\theta}.$$

Local Asymptotic Mixed Normality (LAMN)

A sequence of statistical experiments $(\Omega^N, \mathcal{F}^N, \mathbb{P}_N^\theta)$, $\theta \in \Theta \subset \mathbb{R}^d$, is LAMN at $\theta \in \Theta$, if

$$\log \frac{d\mathbb{P}_N^{\theta + r_N h}}{d\mathbb{P}_N^\theta} = h^\top (I^{N,\theta})^{1/2} \mathcal{N}^{N,\theta} - \frac{1}{2} h^\top I^{N,\theta} h + o_{\mathbb{P}_N^\theta}(1), \quad \forall h \in \mathbb{R}^d,$$

for a sequence of norming matrices $r_N \in \mathcal{M}_+^{d\times d}$, $\lim_{N\to\infty} \|r_N\| = 0$, of random vectors $\mathcal{N}^{N,\theta} \in \mathbb{R}^d$, and matrices $I^{N,\theta} \in \mathcal{M}_+^{d\times d}$ such that

$(\mathcal{N}^{N,\theta}, I^{N,\theta}) \xrightarrow{\mathcal{L}} (\mathcal{N}, \; I^\theta)$ under $\mathbb{P}_N^\theta$.

$\mathcal{N} \sim \mathcal{N}_d(0, I_d)$ and $\mathcal{N} \perp\!\!\!\perp I^\theta$.

with the random Fisher information $I(\theta) \in \mathcal{M}_+^{d\times d}$,

The likelihood locally behaves like a Gaussian shift experiment, conditionaly on I^θ.

Example: Ornstein-Uhlenbeck

$$\mathrm{d}X_t^\theta = -\theta X_t^\theta \mathrm{d}t + \mathrm{d}W_t, \quad \theta \in \mathbb{R}$$

$$\log \frac{\mathrm{d}\mathbb{P}_T^{\theta'}}{\mathrm{d}\mathbb{P}_T^{\theta}}(X) = -(\theta' - \theta)\int_0^T X_t \mathrm{d}W_t - \frac{(\theta'-\theta)^2}{2}\int_0^T X_t^2 \mathrm{d}t$$

If $\theta' = \theta + \frac{h}{\sqrt{T}}$, then under $\mathbb{P}^\theta$

$$\log \frac{\mathrm{d}\mathbb{P}_T^{\theta'}}{\mathrm{d}\mathbb{P}_T^{\theta}}(X) = \frac{h}{\sqrt{T}}\int_0^T X_t \mathrm{d}W_t - \frac{h^2}{2T}\int_0^T X_t^2 \mathrm{d}t = h\Delta_T - \frac{h^2}{2} I_T$$

where

$$\Delta_T = \frac{1}{\sqrt{T}}\int_0^T X_t \mathrm{d}W_t \quad \text{and} \quad I_T = \frac{1}{T}\int_0^T X_t^2 \mathrm{d}t.$$

If $\theta > 0$, then X_t-ergodic with invariant measure $\mathcal{N}(0, \frac{1}{2\theta})$. LAN holds:

$$I_T = \frac{1}{T}\int_0^T X_t^2 \mathrm{d}t \to \frac{1}{\theta^2}\ a.s.;$$

$$\Delta_T = \frac{1}{\sqrt{T}}\int_0^T X_t \mathrm{d}W_t \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \frac{1}{\theta^2}\Big).$$

If $\theta < 0$, then X_t-transient. LAMN holds with $r_T = e^{-\theta T}$.

$$I_T = e^{-2\theta T}\int_0^T X_t^2 \mathrm{d}t \xrightarrow{\mathcal{L}} \mathcal{N}^2(x, \frac{1}{2\theta}) \times \frac{1}{2\theta} = I^\theta.$$

Convolution theorem

- ⋆ (Hájek, 1972) Under LAN each regular estimator $\hat{\theta}_N$ satisfies

$$r_N^{-1}(\hat{\theta}_N - \theta) \xrightarrow{\mathcal{L}} (I^\theta)^{-1/2}\mathcal{N} + Z$$

where $\mathcal{N} \sim \mathcal{N}_d(0, I_d)$, $\mathcal{N} \perp\!\!\!\perp Z$ (recall that I^θ is deterministic).

- ⋆ (Jeganathan, 1988) Under LAMN each regular estimator $\hat{\theta}_N$ satisfies

$$r_N^{-1}(\hat{\theta}_N - \theta) \xrightarrow{\mathcal{L}} (I^\theta)^{-1/2}\mathcal{N} + Z$$

where $\mathcal{N} \sim \mathcal{N}_d(0, I_d)$, $\mathcal{N} \perp\!\!\!\perp Z$ conditionally on I^θ (recall that I^θ is random).

Introduction

LAN and LAMN: general definition
- Definitions

Propagation of Chaos and LAN
- The finite model
- The limit model
- LAN

Efficiency of the MLE

Conditional Propagation of Chaos and LAMN
- The finite model
- The limit model
- LAMN

Propagation of Chaos (PoC) for interacting neurons

General formulation:

$$\mathrm{d}X_t^{i,N,\theta} = b(X_t^{i,N,\theta})\mathrm{d}t + \sum_{j\neq i}^{N} \int_{\mathbb{R}_+}\int_{\mathbb{R}} \frac{u}{N}\mathbf{1}_{\{z\leq f_\theta(X_{s-}^{j,N,\theta})\}}\pi^j(\mathrm{d}t,\mathrm{d}z,\mathrm{d}u)$$
$$+ \int_{\mathbb{R}_+}\int_{\mathbb{R}} \varphi(X_{s-}^{i,N,\theta})\,\mathbf{1}_{\{z\leq f_\theta(X_{s-}^{i,N,\theta})\}}\pi^i(\mathrm{d}t,\mathrm{d}z,\mathrm{d}u)\,,$$

- ⋆ spiking rate f_θ, depends on $\theta \in \mathring{\Theta}$, Θ a compact set of $\mathbb{R}^d$,
- ⋆ (π^j), $j \in \mathbb{N}^*$; family of i.i.d. Poisson Random Measures on $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ with intensity $\mathrm{d}s\,\mathrm{d}z\,\nu(\mathrm{d}u)$,
- ⋆ ν is a probability measure, $\int_{\mathbb{R}} u^2\nu(\mathrm{d}u) < \infty$, $m = \int_{\mathbb{R}} u\,\nu(\mathrm{d}u)$,
- ⋆ drift b, change of potential φ after the spike.

Recall the classical choice

$$b(x) = -\alpha x, \quad \phi(x) = -x, \quad \nu(\mathrm{d}u) = \delta_m(\mathrm{d}u).$$

$$\mathrm{d}X_t^{i,N,\theta} = b(X_t^{i,N,\theta})\mathrm{d}t + \frac{1}{N}\sum_{j\neq i}^{N}\int_{\mathbb{R}_+}\int_{\mathbb{R}} \frac{u}{N}\mathbf{1}_{\{z\leq f_\theta(X_{s-}^{j,N,\theta})\}}\pi^j(\mathrm{d}t,\mathrm{d}z,\mathrm{d}u)$$
$$+\int_{\mathbb{R}_+}\int_{\mathbb{R}} \varphi(X_{s-}^{i,N,\theta})\,\mathbf{1}_{\{z\leq f_\theta(X_{s-}^{i,N,\theta})\}}\pi^i(\mathrm{d}t,\mathrm{d}z,\mathrm{d}u)\,,$$

- $X_t^{i,N}$ meant to represent the membrane potential at time t of neuron i among the N neurons
- The spiking times of neuron i are encoded by π^i, where the jump time is accepted at rate $f_\theta(X_{t-}^{i,N,\theta})$, through the condition $z\leq f_\theta(X_{t-}^{i,N,\theta})$ on the auxiliary variable.
- When neuron i spikes, the potential of neuron i recieves $\varphi(X_{s-}^{i,N,\theta})$ and all other neuron potentials, $j\neq i$, receive an additional potential value $U/N,\ U\sim\nu$
- Deterministic behavior between successive spikes:

$$\mathrm{d}X_t^{i,N} = b(X_t^{i,N})\mathrm{d}t.$$

Corresponding McKean Vlasov dynamics:

$$\mathrm{d}\bar{X}_t^{i,\theta} = b(\bar{X}_t^{i,\theta})dt + m\,\bar{\mu}_t^{\theta}[f_\theta]\mathrm{d}t + \\ + \int_{\mathbb{R}_+}\int_{\mathbb{R}} \varphi(\bar{X}_{s-}^{i,\theta})\,\mathbf{1}_{\{z \le f_\theta(\bar{X}_{s-}^{i})\}}\pi^i(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u)\,, \quad (1)$$

with $\bar{\mu}^\theta$ the law of $(\bar{X}_t^\theta)_{t\ge 0}$, so $\bar{\mu}_t^\theta[f_\theta] = \mathbb{E}[f_\theta(\bar{X}_t^{i,\theta})]$.
Recall $m = \int_{\mathbb{R}} u\, v(du)$.

Propagation of chaos for the system of neurons

For any k, $(X^{1,N,\theta}, \ldots, X^{k,N,\theta})_N$ converges in law as $N \to \infty$ to the family of i.i.d. McKean-Vlasov dynamics $(\bar{X}^{1,\theta}, \ldots, \bar{X}^{k,\theta})$, in the product of Skorohod spaces $\mathcal{D}(\mathbb{R}_+)^{\otimes k}$.

$$PoC \quad \Longleftrightarrow \quad \mu^{N,\theta} \xrightarrow{\mathcal{L}} \bar{\mu}^\theta \quad \text{in} \quad \mathcal{P}(\mathcal{D}(\mathbb{R}_+;\mathbb{R})).$$

Andreis & Dai Pra & Fischer (2018)

Typical assumptions

- The functions $b : \mathbb{R} \to \mathbb{R}$, $\varphi : \mathbb{R} \to \mathbb{R}$ are Lipschitz
- For all $\theta \in \Theta$ the function $x \mapsto f_\theta(x)$ is Lipschitz continuous, bounded and strictly lowerbounded: $\inf_{x\in\mathbb{R}} f_\theta(x) > 0$.
- For all $x \in \mathbb{R}$, the function $\theta \mapsto f_\theta(x)$ is differentiable, and the derivative $\dot{f}_\theta$ satisfies $\sup_{x\in\mathbb{R}} |\dot{f}_\theta(x)| < \infty$.
- There exists $C > 0$ such that for all $\theta \in \Theta, \theta' \in \Theta$,

$$\left| \frac{f_{\theta'}}{f_\theta}(y) - 1 - (\theta' - \theta) \cdot \frac{\dot{f}_\theta}{f_\theta}(y) \right| \le C \left|\theta' - \theta\right|^2,$$

Define:

$$I_t^\theta = \int_0^t \mathbb{E}\Big[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}(\bar{X}_s^\theta)\Big]\,\mathrm{d}s = \int_0^t \bar{\mu}_s^\theta\Big[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}\Big]\,ds$$

LAN property for the spiking rate

The LAN property holds for any $\theta \in \Theta$ at rate $1/\sqrt{N}$ and with the Fischer information matrix I_t^θ.

$$\log \frac{d\mathbb{P}_N^{\theta+\frac{h}{\sqrt{n}}}}{d\mathbb{P}_N^\theta} = h^\top \Delta_t^{N,\theta} - \frac{1}{2}h^\top I_t^{N,\theta}\,h + o_{\mathbb{P}^\theta}(1), \quad \forall h \in \mathbb{R}^d,$$

and the following convergences hold under $\mathbb{P}^\theta$:

$$\Delta_t^{N,\theta} \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, I_t^\theta\Big), \quad I_t^{N,\theta} \xrightarrow{\mathbb{P}^{N,\theta}} I_t^\theta\,.$$

Continuous observation of the whole system on the time-interval $[0, t]$.

Proof: The likelihood ratio process for the observation $X^{N,\theta}_{[0,t]} = (X^{i,N,\theta}_{[0,t]})_{i\leq N}$ is given by

$$L_t^{N,\theta'/\theta}(X^{N,\theta}_{[0,t]}) = \frac{\mathrm{d}\mathbb{P}^{\theta'}_N}{\mathrm{d}\mathbb{P}^{\theta}_N}$$

$$= \prod_{i=1}^{N} \prod_{k\geq 1:\, T^i_k \leq t} \frac{f_{\theta'}}{f_\theta}\big(X^{i,N,\theta}_{T^i_k-}\big) \cdot \exp\Big[-\sum_{j=1}^{N}\int_0^t (f_{\theta'} - f_\theta)\big(X^{j,N,\theta}_s\big)\,ds\Big],$$

where $(T^i_k)_{k\geq 1}$ for $i \in \{1,\ldots,N\}$ is the ordered sequence of jump times of $X^{i,N,\theta}$.

We omit in the following the dependency on the realized trajectory $X^{N,\theta}_{[0,t]}$.

The likelihood satisfies:

$$\log \frac{\mathrm{d}\mathbb{P}_N^{\theta+\frac{h}{\sqrt{N}}}}{\mathrm{d}\mathbb{P}_N^{\theta}} = h^\top \Delta_t^{N,\theta} - \frac{1}{2} h^\top I_t^{N,\theta}\, h + o_{\mathbb{P}^\theta}(1)\,,$$

$$\Delta_t^{N,\theta} = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \iint_{[0,t]\times\mathbb{R}_+} \frac{\dot{f}_\theta}{f_\theta}\left(X_{s-}^{j,N,\theta}\right) \mathbf{1}_{\left\{z \le f_\theta\left(X_{s-}^{j,N,\theta}\right)\right\}} \tilde{\tilde{\pi}}^j(\mathrm{d}s, \mathrm{d}z)\,.$$

$$I_t^{N,\theta} =< \Delta^{N,\theta} >_t = \int_0^t \mu_s^{N,\theta}\left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}\right] \mathrm{d}s$$

$I_t^{N,\theta} \xrightarrow{\mathbb{P}^{N,\theta}} \int_0^t \bar{\mu}_s^\theta \left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}\right] \mathrm{d}s = I_t^\theta$, as a consequence of PoC: $\mu^{N,\theta} \xrightarrow{\mathcal{L}} \bar{\mu}^\theta$.

We treat $\Delta_t^{N,\theta}$ as a process in time to justify its convergence.

Jacod-Shiriaev theory: convergence of semi-martingales via the convergence of their characteristics $(B^N, \tilde{C}^N, \nu^N)$.

$$\Delta_t^{N,\theta} = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \iint_{[0,t]\times\mathbb{R}_+} \frac{\dot{f}_\theta}{f_\theta}(X_{s-}^{j,N,\theta}) \mathbf{1}_{\{z \leq f_\theta\left(X_{s-}^{j,N,\theta}\right)\}} \tilde{\tilde{\pi}}^j(\mathrm{d}s, \mathrm{d}z) .$$

$$\tilde{C}_t^N = < \Delta^{N,\theta} >_t = \int_0^t \mu_s^{N,\theta}\left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}\right] \mathrm{d}s \xrightarrow{\mathbb{P}^{N,\theta}} I_t^\theta$$

Since $(B^N, \tilde{C}^N, \nu^N) \to (0, I_t^\theta, 0)$:

$(\Delta_t^{N,\theta})_{t\geq 0} \xrightarrow{\mathcal{L}} (\int_0^t (\bar{\mu}_s\left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}\right])^{1/2} \mathrm{d}B_s)_{t\geq 0}$ in $\mathcal{D}(\mathbb{R}_+, \mathbb{R})$.

Introduction

LAN and LAMN: general definition
- Definitions

Propagation of Chaos and LAN
- The finite model
- The limit model
- LAN

Efficiency of the MLE

Conditional Propagation of Chaos and LAMN
- The finite model
- The limit model
- LAMN

Efficiency of the Maximum Likelihood Estimator

Under PoC with the LAN property, how efficient is the MLE

$$\hat{\theta}_t^N = \arg\max_{\theta\in\Theta} L_t^{N,\theta/\theta^*}$$

for estimating the unknown jump rate parameter θ^*?

Local asymptotic minimax lower-bound

Entailed under LAN similarly to Hájek convolution Theorem

For any sequence of estimators $(\tilde{\theta}_t^N)$ which estimation errors in θ^* are tight at rate $1/\sqrt{N}$:

$$\liminf_{C\to\infty} \liminf_{N\to\infty} \sup_{\sqrt{N}|\theta-\theta^*|\leq C} \mathbb{E}_{\mathbb{P}_N^\theta}\Big[w\Big((N\, I_t^{\theta^*})^{1/2}\cdot(\tilde{\theta}_t^N-\theta)\Big)\Big]$$

$$\geq \frac{1}{(\sqrt{2\pi})^d}\int_{\mathbb{R}^d} w(x)\, e^{-\frac{1}{2}\|x\|^2}\mathrm{d}x = \mathbb{E}[w(\mathcal{N})]\,,$$

where $w:\mathbb{R}^p\to[0,\infty)$ is any continuous, subconvex and bounded loss function.

Consistency

Treated specifically by considering the asymptotics of $\frac{1}{N} \log L_t^{N,\theta/\theta^*}$, which converges in probability.

Asymptotic normality and efficiency

Two theoretical approaches:

✯ Höpfner's approach (1988):

Asymptotic normality, local minimax optimality
General validity including neuron reset ($\varphi \neq 0$)

✯ Ibragimov & Hasminski's approach (1973):

Stronger results, up to global asymptotic minimax bounds

Currently limited to models without reset ($\varphi \equiv 0$)

Second order development and Höpfner's approach

The main step to exploit Höpfner's approach is to show

$$\Delta_t^{N,\theta} = \hat{I}_t^{N,\theta^*} \cdot \sqrt{N}(\hat{\theta}_t^N - \theta^*) \, ,$$

where $|\hat{I}_t^{N,\theta^*} - I_t^{N,\theta^*}| \to 0$ under $\mathbb{P}_N^{\theta^*}$,

so that $\hat{I}_t^{N,\theta^*}$ can replace I_t^{N,θ^*} in the LAN property.

Why is it?

Second order development and Höpfner's approach

The main step to exploit Höpfner's approach is to show

$$\Delta_t^{N,\theta} = \hat{I}_t^{N,\theta^*} \cdot \sqrt{N}(\hat{\theta}_t^N - \theta^*)\,,$$

where $|\hat{I}_t^{N,\theta^*} - I_t^{N,\theta^*}| \to 0$ under $\mathbb{P}_N^{\theta^*}$,

so that $\hat{I}_t^{N,\theta^*}$ can replace I_t^{N,θ^*} in the LAN property.

Why is it?

✯ $\Delta_t^{N,\theta} = N^{-1/2}\dot{\ell}^N[\theta^*]$, $\dot{\ell}^N[\hat{\theta}_t^N] = 0$, where $\ell_{\theta^*}^N[\theta] = \log \frac{\mathrm{d}\mathbb{P}_N^{\theta}}{\mathrm{d}\mathbb{P}_N^{\theta^*}}$

✯ Consistency of $\hat{\theta}_t^N$ in the interpolation by the second order derivative and control of fluctuations

$\Rightarrow$ uniform bound on the Hessian matrix of f assumed

With Höpfner's approach:

Asymptotic normality with minimal variance

$$\sqrt{N}(\hat{\theta}_t^N - \theta^*) \xrightarrow{\mathcal{L}} (I_t^{\theta^*})^{-1/2}\mathcal{N}, \quad \mathcal{N} \sim \mathcal{N}(0, I_d)\,.$$

Local asymptotic minimax efficiency

The MLE achieves Hájek's asymptotic minimax lower-bound:

$$\lim_{N\to\infty} \sup_{\sqrt{N}|\theta-\theta^*|\leq C} \mathbb{E}_{\mathbb{P}_N^\theta}\Big[w\Big((N\, I_t^{\theta^*})^{1/2} \cdot (\hat{\theta}_t^N - \theta)\Big)\Big] = \mathbb{E}[w(\mathcal{N})]\,,$$

for any $C > 0$, where $w : \mathbb{R}^p \to [0, \infty)$ is any continuous, subconvex and bounded loss function.

Fine controls of the likelihood

Two main steps to exploit Ibragimov & Hasminski's approach

Denote

$$\lambda^N_{\theta^*}[h] = \ell^N_{\theta^*}\left[\theta^* + \frac{h}{\sqrt{N}}\right] = \log \frac{\mathrm{d}\mathbb{P}_N^{\theta+\frac{h}{\sqrt{N}}}}{\mathrm{d}\mathbb{P}_N^{\theta}},$$

for any $h \in \mathbb{R}^d$ such that $\theta^* + h/\sqrt{N} \in \Theta$.

1- Fine regularity property

$$\mathbb{E}_{\mathbb{P}^N_{\theta^*}}\left[\left|\exp\left(\frac{\lambda^N_{\theta^*}[h]}{\mathbf{r}}\right) - \exp\left(\frac{\lambda^N_{\theta^*}[h']}{\mathbf{r}}\right)\right|^{\mathbf{r}}\right] \leq C \cdot \|h - h'\|^{\mathbf{r}}, \ \forall h, h'$$

for any $\mathbf{r} \geq d$, C possibly depending on r.

The argument appeared quite general and direct to obtain by exploiting a change of measure, inspired by Della Maestra & Hoffmann (2022).

2-Fine decay property

$$\mathbb{E}_{\mathbb{P}^N_{\theta^*}}\left[\exp\left(\frac{\lambda^N_{\theta^*}[h]}{2}\right)\right] \leq C\,\|h\|^{-\mathbf{r}}, \forall \theta^* \in K,\ h\,\text{s.t.}\,\theta^* + \frac{h}{\sqrt{N}} \in \Theta,$$

for any compact set K and any $\mathbf{r} > 0$, on which C may depend.

related to the following concentration of measure property:

Concentration of measure in Wasserstein distance

$$\sup_{s\in[0,t]} \mathbb{E}_{\mathbb{P}^N_\theta}[\mathcal{W}_1(\mu^{N,\theta}_s, \bar{\mu}^\theta_s)^{\mathbf{r}}\,] \leq \frac{C}{N^{\mathbf{r}/2}}$$

for any $\mathbf{r} > 2$, C depending only on r.

When $\varphi = 0$, we could adapt the approach of Della Maestra & Hoffmann (2022), based on the concentration result for i.i.d. sample of Fournier & Guillin (2015)

With large jumps (when $\varphi \neq 0$), we conjecture that such a strong result do not hold!

With Ibragimov & Hasminski's approach
($\varphi = 0$, no direct use of the second order derivative):

✯ Asymptotic normality with minimal variance

Local asymptotic minimax efficiency

The MLE achieves a more general bound:

$$\lim_{\delta\to 0} \lim_{N\to\infty} \sup_{|\theta-\theta^*|\le\delta} \mathbb{E}_{\mathbb{P}^\theta_N}\Big[w\Big((N\, I_t^{\theta^*})^{1/2}\cdot(\hat\theta_t^N-\theta)\Big)\Big] = \mathbb{E}[w(\mathcal{N})]\,,$$

where w is any continuous, subconvex and polynomial loss function.

Global asymptotic minimax efficiency

$$\mathcal{R}_w^N(\hat\theta_t^N;\Theta_0) = \inf_{\tilde\theta_N} \mathcal{R}_w^N(\tilde\theta_N;\Theta_0)\,(1+o(1)) \quad \text{as } N\to\infty\,,$$

where the risk is defined for any non-empty open set $\Theta_0 \subset \mathring{\Theta}$ as

$$\mathcal{R}_w^N(\tilde\theta_N;\Theta_0) = \sup_{\theta\in\Theta_0} \mathbb{E}_{\mathbb{P}^\theta}\Big[w\Big((N\, I_t^{\theta})^{1/2}\cdot(\tilde\theta_N-\theta)\Big)\Big]\,.$$

still with w any continuous, subconvex and polynomial loss function.

Introduction

LAN and LAMN: general definition
Definitions

Propagation of Chaos and LAN
The finite model
The limit model
LAN

Efficiency of the MLE

Conditional Propagation of Chaos and LAMN
The finite model
The limit model
LAMN

Finite system with the emergence of common noise

$$
\begin{aligned}
\mathrm{d}X_t^{i,N,\theta} = b(X_t^{i,N,\theta})\mathrm{d}t + \int_{\mathbb{R}_+\times\mathbb{R}} \varphi(X_{t-}^{i,N,\theta}) \mathbf{1}_{\{z \le f_\theta(X_{t-}^{i,N,\theta})\}} \pi^i(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u) \\
+ \sum_{j\neq i} \int_{\mathbb{R}_+\times\mathbb{R}} \frac{u}{\sqrt{N}} \mathbf{1}_{\{z \le f_\theta(X_{t-}^{j,N,\theta})\}} \pi^j(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u),
\end{aligned}
$$

- ⋆ spiking rate f_θ, depends on $\theta \in \mathring{\Theta}$, Θ a compact set of $\mathbb{R}^d$,
- ⋆ (π^j), $j \in \mathbb{N}^*$; family of i.i.d. Poisson Random Measures on $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ with intensity $\mathrm{d}s\,\mathrm{d}z\,\nu(\mathrm{d}u)$,
- ⋆ ν is a centered probability measure, $\int_{\mathbb{R}} u\,\nu(\mathrm{d}u) = 0$,

$$
\sigma^2 = \int_{\mathbb{R}} u^2 \nu(\mathrm{d}u) < \infty.
$$

- ⋆ drift b, change of potential φ after the spike.

Conditional McKean-Vlasov dynamics:

$$\mathrm{d}\bar{X}_t^{i,\theta} = b(\bar{X}_t^{i,\theta})\mathrm{d}t + \varphi(\bar{X}_{t-}^{i,\theta}) \int \mathbf{1}_{\{z \le f_\theta(\bar{X}_{t-}^{i,\theta})\}} \pi^i(\mathrm{d}t, \mathrm{d}z) + \sigma\sqrt{\bar{\mu}_t^\theta[f_\theta]}\mathrm{d}B_t, \quad i = 1, 2, \ldots$$

where $\bar{\mu}_t^\theta[f_\theta] = \mathbb{E}\left[f_\theta(\bar{X}_t^{1,\theta}) \mid \mathcal{B}_t\right]$.

Conditional propagation of chaos for the system of neurons

For any k, $(X^{1,N,\theta}, \ldots, X^{k,N,\theta})_N$ converges in law as $N \to \infty$ to the family of McKean-Vlasov dynamics $(\bar{X}^{1,\theta}, \ldots, \bar{X}^{k,\theta})$ (with common noise), in the product of Skorohod spaces $\mathcal{D}(\mathbb{R}_+)^{\otimes k}$.
X. Erny & E. Löcherbach & D. Loukianova (2021).

Define

$$I_t^\theta = \int_0^t \mathbb{E}\Big[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}(\bar{X}_s^\theta)|\mathcal{B}_t\Big]\, ds = \int_0^t \bar{\mu}_s^\theta\Big[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}\Big]\, ds$$

LAMN for the spiking rate

The LAMN property holds for any $\theta \in \Theta$ at rate $1/\sqrt{N}$ and the Fischer information matrix I_t^θ.

$$\log \frac{d\mathbb{P}_N^{\theta+\frac{h}{\sqrt{n}}}}{d\mathbb{P}_N^\theta} = h^\top (I_t^{N,\theta})^{1/2}\mathcal{N}_t^{N,\theta} - \frac{1}{2}h^\top I_t^{N,\theta}\, h + o_{\mathbb{P}^\theta}(1), \quad \forall h \in \mathbb{R}^d.$$

Moreover, $(\mathcal{N}_t^{N,\theta}, I_t^{N,\theta}) \xrightarrow{\mathcal{L}} (\mathcal{N}, I_t^\theta)$,
where $\mathcal{N}$ is a standard Gaussian variable, independent of B and hence of I_t^θ.

Idea of the proof

If $\theta' = \theta + h/\sqrt{N}$, the likelihood admits the representation:

$$\log L_t^{N,\theta'/\theta} = h^\top \Delta_t^{N,\theta} - \frac{1}{2} h^\top I_t^{N,\theta} h + o_{\mathbb{P}_N^\theta}(1).$$

$$\Delta_t^{N,\theta} = \frac{1}{\sqrt{N}} \sum_{j=1}^N \iint_{[0,t]\times\mathbb{R}_+} \frac{\dot{f}_\theta}{f_\theta}(X_{s-}^{j,N,\theta}) \mathbf{1}_{\left\{z \le f_\theta\left(X_{s-}^{j,N,\theta}\right)\right\}} \tilde{\overline{\pi}}^j(\mathrm{d}s, \mathrm{d}z).$$

$$I_t^{N,\theta} = \langle \Delta^{N,\theta} \rangle_t = \int_0^t \mu_s^{N,\theta} \left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta} \right] \mathrm{d}s.$$

Recall

$$I_t^\theta = \int_0^t \bar{\mu}_s^\theta \left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta} \right] \mathrm{d}s = \int_0^t \mathbb{E}\left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}(X_s^\theta) \middle| \mathcal{B}_s \right] \mathrm{d}s.$$

We want to show: $(\Delta_t^{N,\theta}, I_t^{N,\theta}) \xrightarrow{\mathcal{L}} \left(\int_0^t (\bar{\mu}_s^\theta [\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}])^{1/2} \mathrm{d}W_s, I_t^\theta \right)$,

where B and W are independent.

By Conditional Propagation of Chaos (CPoC):
$\mu^{N,\theta} \xrightarrow{\mathcal{L}} \bar{\mu}^{\theta} = \mathcal{L}(X^{\theta} \mid \mathcal{B})$.
Hence, we already know that:

$$I_t^{N,\theta} = \langle \Delta^{N,\theta} \rangle_t = \int_0^t \mu_s^{N,\theta} \left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta} \right] \mathrm{d}s \xrightarrow{\mathcal{L}} \int_0^t \bar{\mu}_s^{\theta} \left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta} \right] \mathrm{d}s = I_t^{\theta}.$$

Considering $(\Delta_t^{N,\theta}, I_t^{N,\theta})$ could give convergence in law to $\left(\int_0^t (\bar{\mu}_s^{\theta} \left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta} \right])^{1/2} \mathrm{d}W_s, I_t^{\theta} \right)$, with some Brownian motion W, but would not clarify the independence between W and B.

W is created by $(\widetilde{\pi}_j)_{j \in \mathbb{N}^*}$. So where does B come from?

By Conditional Propagation of Chaos (CPoC):
$\mu^{N,\theta} \xrightarrow{\mathcal{L}} \bar{\mu}^{\theta} = \mathcal{L}(X^{\theta} \mid \mathcal{B})$.
Hence, we already know that:

$$I_t^{N,\theta} = \langle \Delta^{N,\theta} \rangle_t = \int_0^t \mu_s^{N,\theta} \left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta} \right] \mathrm{d}s \xrightarrow{\mathcal{L}} \int_0^t \bar{\mu}_s^{\theta} \left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta} \right] \mathrm{d}s = I_t^{\theta} .$$

Considering $(\Delta_t^{N,\theta}, I_t^{N,\theta})$ could give convergence in law to $\left(\int_0^t (\bar{\mu}_s^{\theta} \left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta} \right])^{1/2} \mathrm{d}W_s, I_t^{\theta} \right)$, with some Brownian motion W, but would not clarify the independence between W and B.

W is created by $(\widetilde{\pi}_j)_{j \in \mathbb{N}^*}$. So where does B come from?

$$\sum_{j \neq i} \int_{\mathbb{R}_+ \times \mathbb{R}} \frac{u}{\sqrt{N}} \mathbf{1}_{\{z \leq f_\theta(X_{t-}^{j,N,\theta})\}} \pi^j(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u) \xrightarrow{\mathcal{L}} \sigma \sqrt{\bar{\mu}_t^{\theta}[f_\theta]} \mathrm{d}B_t .$$

Consider the three-dimensional semi-martingale $\zeta^{N,\theta} := (J^{N,\theta}, \Delta^{N,\theta}, I^{N,\theta})$:

$$J_t^{N,\theta} = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \int_{\mathbb{R}_+ \times \mathbb{R}} u \cdot \mathbf{1}_{\{z \leq f_\theta(X_{t-}^{j,N,\theta})\}} \pi_j(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u),$$

$$\Delta_t^{N,\theta} = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \int_{[0,t] \times \mathbb{R}_+} \frac{\dot{f}_\theta}{f_\theta}(X_{s-}^{j,N,\theta}) \mathbf{1}_{\{z \leq f_\theta(X_{t-}^{j,N,\theta})\}} \widetilde{\bar{\pi}}_j(\mathrm{d}t, \mathrm{d}z),$$

$$I_t^{N,\theta} = \langle \Delta^{N,\theta} \rangle_t = \int_0^t \mu_s^{N,\theta} \Big[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta} \Big] \mathrm{d}s .$$

$\zeta^{N,\theta} = (J^{N,\theta}, \Delta^{N,\theta}, I^{N,\theta})$ is a three-dimensional semi-martingale.
We calculate its characteristics: $\widetilde{C}_t^{N,\theta} = \langle \text{Mart part}, \text{Mart part} \rangle_t$.
The modified second characteristics is defined as:

$$\widetilde{C}_t^{N,\theta} = \begin{pmatrix} \langle J^{N,\theta} \rangle_t & \langle J^{N,\theta}, \Delta^{N,\theta} \rangle_t & \langle J^{N,\theta}, 0 \rangle_t \\ \langle \Delta^{N,\theta}, J^{N,\theta} \rangle_t & \langle \Delta^{N,\theta} \rangle_t & \langle \Delta^{N,\theta}, 0 \rangle_t \\ \langle 0, J^{N,\theta} \rangle_t & \langle 0, \Delta^{N,\theta} \rangle_t & 0 \end{pmatrix}$$

As $I^{N,\theta}$ is of bounded variation, the last line and column are zero.
$\int_{\mathbb{R}} u\, \nu(\mathrm{d}u) = 0$ entails $\langle J_t^{N,\theta}, \Delta_t^{N,\theta} \rangle = 0$, so the independence between W and B.
Moreover:

$$\langle \Delta^{N,\theta} \rangle_t = \int_0^t \mu_s^{N,\theta} \left(\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta} \right) \mathrm{d}s.$$

And $\zeta^{N,\theta} = (J^{N,\theta}, \Delta^{N,\theta}, I^{N,\theta}) \xrightarrow{\mathcal{L}} \zeta = (J^{\theta}, \Delta^{\theta}, I^{\theta})$,

$$J_t^\theta = \sigma \int_0^t \sqrt{\bar{\mu}_s^\theta(f_\theta)} \mathrm{d}B_s = \sigma \int_0^t \sqrt{\mathbb{E}\left[f_\theta(X_s^\theta) \mid \mathcal{B}_s\right]} \mathrm{d}B_s,$$

$$\Delta_t^\theta = \int_0^t \left[\bar{\mu}_s^\theta\Big(\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}\Big)\right]^{1/2} \mathrm{d}W_s = \int_0^t \mathbb{E}\left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}(X_s^\theta) \mid \mathcal{B}_s\right]^{1/2} \mathrm{d}W_s\,,$$

$$I_t^\theta = \int_0^t \bar{\mu}_s^\theta\left(\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}\right) \mathrm{d}s = \int_0^t \mathbb{E}\left[\frac{(\dot{f}_\theta)^{\otimes 2}}{f_\theta}(X_s^\theta) \mid \mathcal{B}_s\right] \mathrm{d}s\,.$$

with W, B independent Brownian motions.

In particular, $(\Delta^{N,\theta}, I^{N,\theta}) \xrightarrow{\mathcal{L}} (\Delta^{\theta}, I^{\theta})$.

Synthesis of Main Results

Mean-Field Limit and Statistical Inference

- Propagation of Chaos (PoC) and its conditional counterpart: As $N \to \infty$, the finite system converges to a McKean-Vlasov limit
- LAN Property: when the interaction effects scale as $1/N$, the likelihood admits a Gaussian shift representation
- LAMN Property: with interaction effects that are centered and scale as $1/\sqrt{N}$, we obtain a conditional Gaussian shift

MLE Efficiency Results, under LAN

- Consistency: $\hat{\theta}_t^N \to \theta^*$ in probability under $\mathbb{P}_N^{\theta^*}$
- Asymptotic Normality: $\sqrt{N}(\hat{\theta}_t^N - \theta^*) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, (I_t^{\theta^*})^{-1}\right)$
- Asymptotic optimality:
  - Local minimax (Höpfner's approach): For all models, also with reset
  - Global minimax (Ibragimov & Hasminski): For models without reset

Perspectives

- Interactions along a random heterogeneous w-graph, and beyond exchangeability
- Initial condition after relaxation, already at quasi-stationary regime
- Restricted observation scheme:
  observation of spike trains with many neurons,
  but only a few for which we can follow the membrane potential
- Test the emergence of noise?
- Adaptation to ecological models (of dispersion e.g.)

Thank you for youre attention

and your questions