Statistical estimation of Fokker-Planck equation at fixed time

Tien-Dat Nguyen

In collaboration with Mylène Maïda, Thanh Mai Pham Ngoc, Vincent Rivoirard and Viet Chi Tran.

Laboratoire de Mathématiques d'Orsay, Université Paris-Sud.

May 27^{th} 2020

Département de mathématiques d'Orsay

UNIVERSITÉ PARIS SUD Comprendre le monde

université PARIS-SACLAY

École doctorale de mathématiques Hadamard (EDMH)

Contents

Outline

Consider 1-d free Fokker-Planck equation

$$\frac{\partial}{\partial t}\mu_t = -\frac{\partial}{\partial x}\left[\mu_t\left(H\mu_t\right)\right], \tag{1}$$

where $\left(\mu_t\right)_{t\geq 0}$ a family of proba measures, and initial condition $\mu_0(x) = p_0(x)dx$ (unknown), and H is Hilbert transform defined by

$$H\mu_t(x) = p.v. \int \frac{d\mu_t(y)}{x-y} := \lim_{\varepsilon \searrow 0} \int_{\mathbb{R}\setminus[x-\varepsilon, x+\varepsilon]} \frac{1}{x-y} d\mu_t(y).$$

Consider 1-d free Fokker-Planck equation

$$\frac{\partial}{\partial t}\mu_t = -\frac{\partial}{\partial x}\left[\mu_t\left(H\mu_t\right)\right], \qquad (1)$$

where $\left(\mu_t\right)_{t\geq 0}$ a family of proba measures, and initial condition $\mu_0(x) = p_0(x)dx$ (unknown), and H is Hilbert transform defined by

$$H\mu_t(x) = p.v.\int \frac{d\mu_t(y)}{x-y} := \lim_{\varepsilon \searrow 0} \int_{\mathbb{R}\setminus[x-\varepsilon, x+\varepsilon]} \frac{1}{x-y} d\mu_t(y).$$

- Then, for $f \in C^1$:

$$\langle \mu_t, f\rangle = \langle \mu_0, f\rangle + \int_0^t \left(\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{f'(\lambda)}{\lambda - \widetilde{\lambda}} \mu_s(d\widetilde{\lambda})\mu_s(d\lambda)\right) ds, \quad t > 0.$$

- We have:

$$\mu_t = \sigma_t \boxplus \mu_0, \qquad (2)$$

with $\sigma_t(dx) = \frac{1}{2\pi t}\sqrt{4t - x^2}.\mathbb{1}_{\left[-2\sqrt{t}, 2\sqrt{t}\right]}(x)dx$, semi-circular distribution.

Observations

Consider

$$X_n(t) = X_n(0) + H_n(t), t \geq 0, \tag{3}$$

where $X_n(0)$ diagonal matrix with entries $\overset{iid}{\sim} \mu_0$, and $H_n(t)$ standard Hermitian Brownian motion.

- Denote $\mathcal{H}_n(\mathbb{C})$ the space of n-dim matrices H_n s.t. $(H_n)^* = H_n$.

Definition 1

Let $\left(B_{k,l}, \tilde{B}_{k,l}, 1 \leq k, l \leq n\right)$ be a collection of i.i.d. real valued standard Brownian motions, the Hermitian Brownian motion, denoted $H_n \in \mathcal{H}_n(\mathbb{C})$, is the random process with entries $\{(H_n(t))_{k,l}, t \geq 0, k \leq l\}$ equal to

$$(H_n)_{k,l} = \begin{cases} \frac{1}{\sqrt{2n}}\left(B_{k,l} + i\ \tilde{B}_{k,l}\right), & \text{if } k < l \\ \frac{1}{\sqrt{n}} B_{k,k} \quad , & \text{if } k = l \end{cases}$$

- Then $(\lambda_1^n(t), ..., \lambda_n^n(t))$ (eigenvalues of $X_n(t)$) solves

$$d\lambda_j^n(t) = \frac{1}{\sqrt{n}} d\beta_j(t) + \frac{1}{n} \sum_{k \neq j} \frac{dt}{\lambda_j^n(t) - \lambda_k^n(t)}, \tag{4}$$

where β_j i.i.d. standard Brownian motion.

- Then $(\lambda_1^n(t), \ldots, \lambda_n^n(t))$ (eigenvalues of $X_n(t)$) solves

$$d\lambda_j^n(t) = \frac{1}{\sqrt{n}} d\beta_j(t) + \frac{1}{n} \sum_{k \neq j} \frac{dt}{\lambda_j^n(t) - \lambda_k^n(t)}, \tag{4}$$

where β_j i.i.d. standard Brownian motion.

- For $t > 0$ define

$$\mu_t^n := \frac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_j^n(t)} \quad . \tag{5}$$

Proposition 1.1

$\lambda^n(0)$ satisfies $C_0 := \sup_{n \geq 1} \frac{1}{n} \log\left(\lambda_j^n(0)^2 + 1\right) < \infty$ a.s. .
Then, $(\mu_t^n)_{t \geq 0}$ converges weakly a.s. to $(\mu_t)_{t \geq 0}$ solving equation (1) .

Outline

Outline

Let μ proba measure on $\mathbb{R}$, and define the Cauchy transform of μ by

$$G_\mu(z) = \int_{\mathbb{R}} \frac{d\mu(x)}{z - x}, \quad z \in \mathbb{C} \backslash \mathbb{R}. \tag{6}$$

- Denote $\mathbb{C}^+ = \{z \in \mathbb{C} \mid \operatorname{Im}(z) > 0\}$, and $\mathbb{C}_\gamma := \{z \in \mathbb{C} \mid \operatorname{Im}(z) > \gamma\}$
- Define: $R_\mu(z) = G_\mu^{<-1>}(z) - \frac{1}{z}$.

Let μ proba measure on $\mathbb{R}$, and define the Cauchy transform of μ by

$$G_\mu(z) = \int_{\mathbb{R}} \frac{d\mu(x)}{z - x}, \quad z \in \mathbb{C} \backslash \mathbb{R}. \tag{6}$$

- Denote $\mathbb{C}^+ = \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}$, and $\mathbb{C}_\gamma := \{z \in \mathbb{C} \mid \mathrm{Im}(z) > \gamma\}$
- Define: $R_\mu(z) = G_\mu^{<-1>}(z) - \frac{1}{z}$.

Given μ_1 and μ_2 proba measures, $\exists!$ proba measure μ:

$$R_\mu = R_{\mu_1} + R_{\mu_2}$$

The measure $\mu := \mu_1 \boxplus \mu_2$, is called **free convolution** of μ_1 and μ_2.

- G_μ does NOT vanish on $\mathbb{C}^+ \Rightarrow$ define reciprocal Cauchy transform of μ by

$$F_\mu(z) = \frac{1}{G_\mu(z)}, \quad z \in \mathbb{C}^+.$$

Outline

For $t > 0$, $\mu_t = \mu_0 \boxplus \sigma_t \quad \Rightarrow \quad$ Problem: recover μ_0, knowledge on μ_t .

For $t > 0$, $\mu_t = \mu_0 \boxplus \sigma_t$ $\Rightarrow$ Problem: recover μ_0, knowledge on μ_t .

Theorem 2.1

∃! subordination functions $w_1, w_{fp} : \mathbb{C}_{2\sqrt{t}} \to \mathbb{C}^+$, *s.t. for* $z \in \mathbb{C}_{2\sqrt{t}}$:

(i) $Im(w_1(z)) \geq \frac{1}{2} Im(z)$, $Im(w_{fp}(z)) \geq \frac{1}{2} Im(z)$, *and*

$$\lim_{y \to +\infty} \frac{w_1(iy)}{iy} = \lim_{y \to +\infty} \frac{w_{fp}(iy)}{iy} = 1;$$

(ii) $F_{\mu_0}(z) = F_{\sigma_t}\big(w_1(z)\big) = F_{\mu_t}\big(w_{fp}(z)\big)$;

(iii) $w_{fp}(z) = z + w_1(z) - F_{\mu_0}(z)$;

For $t > 0$, $\mu_t = \mu_0 \boxplus \sigma_t$ $\Rightarrow$ Problem: recover μ_0, knowledge on μ_t .

Theorem 2.1

∃! subordination functions $w_1, w_{fp} : \mathbb{C}_{2\sqrt{t}} \to \mathbb{C}^+$, *s.t. for* $z \in \mathbb{C}_{2\sqrt{t}}$:

(i) $Im(w_1(z)) \geq \frac{1}{2} Im(z)$, $Im(w_{fp}(z)) \geq \frac{1}{2} Im(z)$, *and*

$$\lim_{y \to +\infty} \frac{w_1(iy)}{iy} = \lim_{y \to +\infty} \frac{w_{fp}(iy)}{iy} = 1;$$

(ii) $F_{\mu_0}(z) = F_{\sigma_t}\big(w_1(z)\big) = F_{\mu_t}\big(w_{fp}(z)\big)$;

(iii) $w_{fp}(z) = z + w_1(z) - F_{\mu_0}(z)$;

(iv) *Denote* $h_{\sigma_t}(w) = w - F_{\sigma_t}(w) = t.G_{\sigma_t}(w)$ *and* $\widetilde{h}_{\mu_t}(w) = w + F_{\mu_t}(w)$ *on* $\mathbb{C}^+$. *Moreover, define* $K_z(w) = h_{\sigma_t}\big(\widetilde{h}_{\mu_t}(w) - z\big) + z$. *Then,* $K_z\big(w_{fp}(z)\big) = w_{fp}(z)$, *and* $K_z^{\circ m}(w) \xrightarrow{m \to \infty} w_{fp}(z)$ *for any* $w \in \mathbb{C}_{\frac{1}{2} Im(z)}$.

(see also Arizmendi et al. [2] for general setting)

Lemma 2.2

For $z \in \mathbb{C}_{2\sqrt{t}}$, $G_{\mu_0}(z) = \frac{1}{t}\big(w_{fp}(z) - z\big) = G_{\mu_t}\big(w_{fp}(z)\big)$.
Consequently, $\big|w_{fp}(z) - z\big| \leq \sqrt{t}$.

- For $\gamma > 0$, $\mathcal{C}_\gamma$ denotes the centered Cauchy distribution, parameter γ.
- For any proba measure μ on $\mathbb{R}$:

$$f_{\mu * \mathcal{C}_\gamma}(x) = -\frac{1}{\pi}\mathrm{Im}\, G_\mu(x + i\gamma), \quad x \in \mathbb{R}. \tag{7}$$

Lemma 2.2

For $z \in \mathbb{C}_{2\sqrt{t}}$, $G_{\mu_0}(z) = \frac{1}{t}\left(w_{fp}(z) - z\right) = G_{\mu_t}\left(w_{fp}(z)\right)$.
Consequently, $\left|w_{fp}(z) - z\right| \leq \sqrt{t}$.

- For $\gamma > 0$, $\mathcal{C}_\gamma$ denotes the centered Cauchy distribution, parameter γ.
- For any proba measure μ on $\mathbb{R}$:

$$f_{\mu * \mathcal{C}_\gamma}(x) = -\frac{1}{\pi}\mathrm{Im}\, G_\mu(x + i\gamma), \quad x \in \mathbb{R}. \tag{7}$$

Then, with $\gamma > 2\sqrt{t}$, for $x \in \mathbb{R}$

$$f_{\mu_0 * \mathcal{C}_\gamma}(x) = \frac{1}{\pi t}\left[\gamma - \mathrm{Im}\, w_{fp}(x + i\gamma)\right]. \tag{8}$$

Outline

Observation: matrix $X_n(t)$, at time $t > 0$, given n.
Thus, an estimator of $G_{\mu_t}(z)$:

$$\widehat{G}_{\mu_t^n}(z) := \frac{1}{n}\sum_{j=1}^{n} \frac{1}{z - \lambda_j^n(t)} = \frac{1}{n}\mathrm{Tr}\Big(\big(zI_n - X_n(t)\big)^{-1}\Big). \tag{9}$$

Observation: matrix $X_n(t)$, at time $t > 0$, given n.
Thus, an estimator of $G_{\mu_t}(z)$:

$$\widehat{G}_{\mu_t^n}(z) := \frac{1}{n}\sum_{j=1}^{n}\frac{1}{z-\lambda_j^n(t)} = \frac{1}{n}\mathrm{Tr}\Big(\big(zI_n - X_n(t)\big)^{-1}\Big). \tag{9}$$

Theorem 3.1

$\exists!$ a fixed-point to the functional equation in $w(z)$, for $z \in \mathbb{C}_{2\sqrt{t}}$:

$$\frac{1}{t}\big(w(z)-z\big) = \widehat{G}_{\mu_t^n}\big(w(z)\big), \tag{10}$$

this fixed-point is denoted $\widehat{w}_{fp}^n(z)$. Moreover, $\left|\widehat{w}_{fp}^n(z) - z\right| \leq \sqrt{t}$.

- Fourier Transform (FT) of Cauchy distribution $\mathcal{C}_\gamma$: $f^\star_{\mathcal{C}_\gamma}(z) = e^{-\gamma|z|}$.

- Fourier Transform (FT) of Cauchy distribution $\mathcal{C}_\gamma$: $f^\star_{\mathcal{C}_\gamma}(z) = e^{-\gamma|z|}$.

- bandwidth $h > 0$ and a regularizing Kernel K (compactly supported in Fourier domain) $\Rightarrow$ estimator of p_0 via its FT:

$$\widehat{p_0^\star}(z) = e^{\gamma|z|}.K_h^\star(z).\frac{1}{\pi t}\left[\gamma - \mathrm{Im}\,\widehat{w}_{fp}^n(x+i\gamma)\right]^\star(z), \tag{11}$$

with $K_h(x) = \frac{1}{h}K(\frac{x}{h})$.

For instance, choose $K(x) = sinc(x)$, then $K^\star(z) = \mathbb{1}_{[-1,1]}(z)$.

Consistency of estimator

Proposition 3.2

Let $\gamma > 2\sqrt{t}$*:*

(i) *for any* $z \in \mathbb{C}_{2\sqrt{t}}$, $\widehat{w}_{fp}^{n}(z) \xrightarrow{a.s.} w_{fp}(z)$ *as* $n \to \infty$ *;*

(ii) *the convergence is uniform on* $\mathbb{C}_\gamma$ *;*

Consistency of estimator

Proposition 3.2

Let $\gamma > 2\sqrt{t}$*:*

(i) *for any* $z \in \mathbb{C}_{2\sqrt{t}}$, $\widehat{w}^n_{fp}(z) \xrightarrow{a.s.} w_{fp}(z)$ *as* $n \to \infty$;

(ii) *the convergence is uniform on* $\mathbb{C}_\gamma$;

(iii) *convergence rate on* $\mathbb{C}_\gamma$*:*

$$\sup_{n\in\mathbb{N}} \sup_{z\in\mathbb{C}_\gamma} \mathbb{E}\left[\left|\sqrt{n}\left(\widehat{w}^n_{fp}(z) - w_{fp}(z)\right)\right|^2\right] < +\infty. \tag{12}$$

Proof: We have

$$\left|\widehat{w}^n_{fp}(z) - w_{fp}(z)\right| \leq \left(\frac{t\gamma^2}{\gamma^2 - 4t}\right) \times \left|\widehat{G}_{\mu^n_t}\left(w_{fp}(z)\right) - G_{\mu_t}\left(w_{fp}(z)\right)\right|.$$

Using properties on $\widehat{G}_{\mu^n_t} \Rightarrow$ we get (i) and (ii).

Proof of Proposition 3.2-(iii)

Decompose by:

$$\begin{aligned}\widehat{G}_{\mu_t^n}(z) - G_{\mu_t}(z) =& \widehat{G}_{\mu_t^n}(z) - \mathbb{E}\left(\widehat{G}_{\mu_t^n}(z)\right) + \mathbb{E}\left(\widehat{G}_{\mu_t^n}(z)\right) - \mathbb{E}\left(G_{\mu_0^n \boxplus \sigma_t}(z)\right) \\ &+ \mathbb{E}\left(G_{\mu_0^n \boxplus \sigma_t}(z)\right) - G_{\mu_t}(z) \\ =& A_{n,1}(z) + A_{n,2}(z) + A_{n,3}(z).\end{aligned}$$

Proof of Proposition 3.2-(iii)

Decompose by:

$$\begin{aligned}\widehat{G}_{\mu_t^n}(z) - G_{\mu_t}(z) =& \widehat{G}_{\mu_t^n}(z) - \mathbb{E}\left(\widehat{G}_{\mu_t^n}(z)\right) + \mathbb{E}\left(\widehat{G}_{\mu_t^n}(z)\right) - \mathbb{E}\left(G_{\mu_0^n \boxplus \sigma_t}(z)\right) \\ &+ \mathbb{E}\left(G_{\mu_0^n \boxplus \sigma_t}(z)\right) - G_{\mu_t}(z) \\ =& A_{n,1}(z) + A_{n,2}(z) + A_{n,3}(z).\end{aligned}$$

Then,

■ Fluctuations of $\widehat{G}_{\mu_t^n}$ $\Rightarrow$ upper bounds for $nA_{n,1}(z)$ and $nA_{n,2}(z)$.

(see more in [6], S.Dallaporta and M.Février, 2019)

Remark: $X_n(0)$ is random.

Proof of Proposition 3.2-(iii)

Decompose by:

$$\begin{aligned}\widehat{G}_{\mu_t^n}(z) - G_{\mu_t}(z) =& \widehat{G}_{\mu_t^n}(z) - \mathbb{E}\left(\widehat{G}_{\mu_t^n}(z)\right) + \mathbb{E}\left(\widehat{G}_{\mu_t^n}(z)\right) - \mathbb{E}\left(G_{\mu_0^n \boxplus \sigma_t}(z)\right) \\ &+ \mathbb{E}\left(G_{\mu_0^n \boxplus \sigma_t}(z)\right) - G_{\mu_t}(z) \\ =& A_{n,1}(z) + A_{n,2}(z) + A_{n,3}(z).\end{aligned}$$

Then,

■ Fluctuations of $\widehat{G}_{\mu_t^n}$ $\Rightarrow$ upper bounds for $nA_{n,1}(z)$ and $nA_{n,2}(z)$.
(see more in [6], S.Dallaporta and M.Février, 2019)

Remark: $X_n(0)$ is random.

■ The third term $\mathbb{E}\left(G_{\mu_0^n \boxplus \sigma_t}(z)\right) - G_{\mu_t}(z)$ is associated to a C.L.T., so is of order $n^{1/2}$.

Outline

Some assumptions

Assumption 4.1 (H1)

p_0 belongs to space $S(a,r,L)$ defined for $a>0, L>0$ and $0<r\le 2$ by

$$S(a,r,L) := \left\{ f \ \textit{density s.t.} \ \int_{\mathbb{R}} \left|f^{\star}(z)\right|^2 . e^{2a|z|^r} dz \le L^2 \right\}. \tag{13}$$

Assumption 4.2 (H2)

For $\kappa>0$ sufficiently large, $\exists C>0$:

$$\mu_0\big((\kappa,+\infty)\big) \le \frac{C}{\kappa}. \tag{14}$$

By Parseval's equality:

$$\|\widehat{p}_0 - p_0\|^2 \leq \frac{1}{\pi}\|\widehat{p}_0^\star - K_h^\star . p_0^\star\|^2 + \frac{1}{\pi}\|K_h^\star . p_0^\star - p_0^\star\|^2 \tag{15}$$

Under Assumption (H1)-(H2):

$$MISE = \mathbb{E}\left(\|\widehat{p}_0 - p_0\|^2\right) \leq \frac{C_{var} . e^{\frac{2\gamma}{h}}}{\sqrt{n}} + C_{bias}(L)e^{-2ah^{-r}}. \tag{16}$$

By Parseval's equality:

$$\|\widehat{p}_0 - p_0\|^2 \leq \frac{1}{\pi}\|\widehat{p}_0^\star - K_h^\star . p_0^\star\|^2 + \frac{1}{\pi}\|K_h^\star . p_0^\star - p_0^\star\|^2 \tag{15}$$

Under Assumption (H1)-(H2):

$$MISE = \mathbb{E}\left(\|\widehat{p}_0 - p_0\|^2\right) \leq \frac{C_{var} . e^{\frac{2\gamma}{h}}}{\sqrt{n}} + C_{bias}(L) e^{-2ah^{-r}}. \tag{16}$$

Then, minimizing in h $\Rightarrow$ convergence rate for MISE.

$$MISE = \begin{cases} O\left(n^{-\frac{a}{2(a+\gamma)}}\right) & \text{if } r = 1 \\ O\left(\exp\left\{-\frac{2a}{(2\gamma)^r}(\ln\sqrt{n})^r + \sum_{i=1}^k b_i^* (\ln\sqrt{n})^{r+i(r-1)}\right\}\right) & \text{if } r < 1 \\ O\left(\frac{1}{\sqrt{n}}\exp\left\{\frac{2\gamma}{(2a)^{1/r}}(\ln\sqrt{n})^{1/r} - \sum_{i=1}^k d_i^* (\ln\sqrt{n})^{\frac{1}{r} - i\frac{r-1}{r}}\right\}\right) & \text{if } r > 1 \end{cases} \tag{17}$$

where b_i^* and d_i^* are some coefficients. (see more in [7] C.Lacour)

Outline

$\mu_0 \sim$ Cauchy(0,5) with $n = 2000$, at $t = 15$

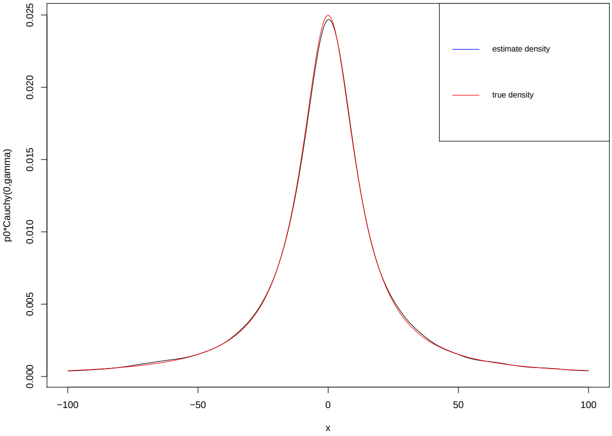

Figure: 1^{st} step: convolution between the estimate $\widehat{p}_0$ and $\mathcal{C}_\gamma$

$\mu_0 \sim$ Cauchy(0,5) with $n = 2000$, at $t = 15$

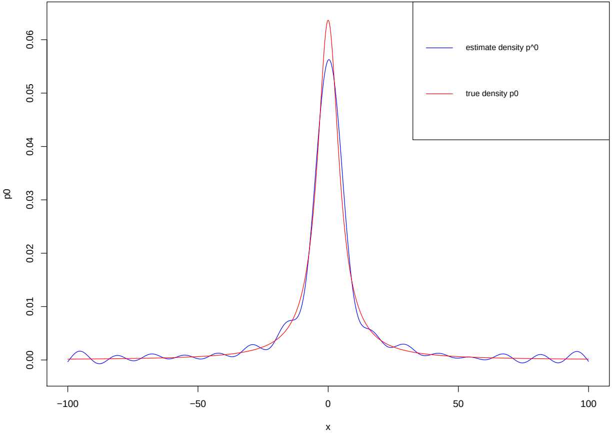

Figure: 2^{nd} step: After deconvolution with $\mathcal{C}_\gamma$.

$\mu_0 \sim$ Gaussian(0,5) with $n = 2000$, at $t = 15$

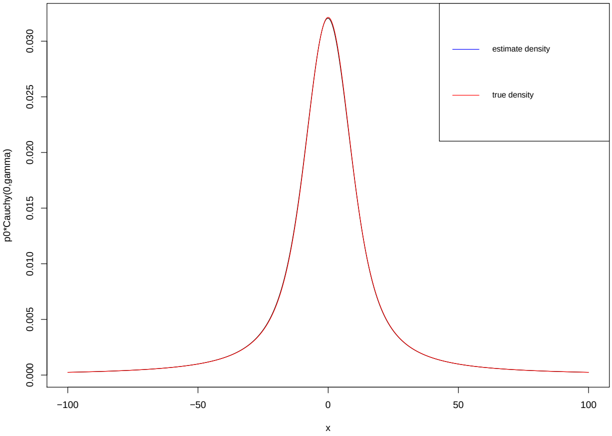

Figure: 1^{st} step: convolution between the estimate $\widehat{p}_0$ and $\mathcal{C}_\gamma$

$\mu_0 \sim$ Gaussian(0,5) with $n = 2000$, at $t = 15$

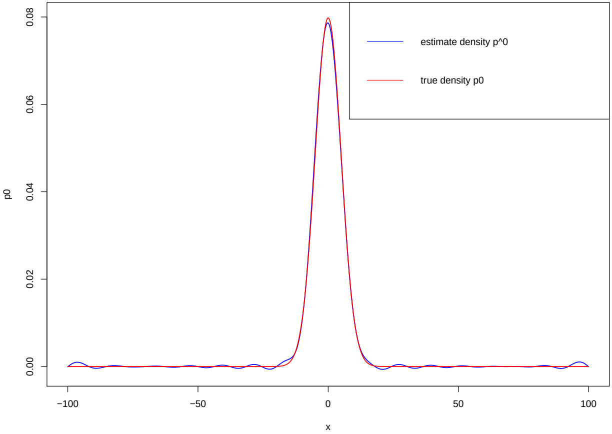

Figure: 2^{nd} step: After deconvolution with $\mathcal{C}_\gamma$

Outline

Conclusion

- Solving fixed-point equation results an estimator $\widehat{w}_{fp}^n$ for subordination function.
- The convergence of $\widehat{w}_{fp}^n(z)$ towards $w_{fp}(z)$ is uniform and has a rate of $n^{1/2}$.

Conclusion

- Solving fixed-point equation results an estimator $\widehat{w}_{fp}^n$ for subordination function.
- The convergence of $\widehat{w}_{fp}^n(z)$ towards $w_{fp}(z)$ is uniform and has a rate of $n^{1/2}$.
- Using Stieltjes-inversion-formula to recover the density function, remaining in a classical convolution with Cauchy distribution $\mathcal{C}_\gamma$.
- Using a regularizing Kernel to do the deconvolution with C_γ.

Conclusion

- Work in progress: to improve the rate of convergence.
- Using Cross-Validation to obtain a data-driven optimal value for bandwidth h.

Some of references

Greg W. Anderson, Alice Guionnet and Ofer Zeitouni, *An introduction to Random Matrices*, Cambridge University Press, 2010.

O. Arizmendi, P. Tarrago and C. Vargas, *Subordination methods for free deconvolution*, submitted, 2020.

S.Belinschi and H.Bercovici, *A new approach to subordination results in free probability*, Journal d'Analyse Mathematique, 101:357-365,2007.

P. Biane, *On the free convolution with a semi-circular distribution*, Indiana Uni. Math. J., page 705-718, 1997.

C. Butucea and A. B. Tsybakov, *Sharp optimality in density deconvolution with dominating bias*, I., Teor. Veroyatn. Primen., 52(1):111-128, 2007.

S. Dallaporta and M. Fevrier, *Fluctuations of linear spectral statistics of deformed Wigner matrices*, hal-02079313, 2019.

C. Lacour, *Rates of convergence for nonparametric deconvolution*, Comptes rendus de l'Acedémie des sciences. Série I, Mathématique, 342(11):877-882,2006.

and others.

Thank you for your attention !