# Signature-based models: theory and calibration

Sara Svaluto-Ferro

(based on joint works with C. Cuchiero and G. Gazzani)

UNIVERSITÀ di VERONA

Vienna, November 2, 2022

## Signatures...why?

Because the (time extended) signature of a continuous semimartingale uniquely determines its path...

...and because every polynomial on the signature has a linear representative.

$\longrightarrow$ If $S_T = F((X_t)_{t\in[0,T]})$ for some continuous map $F$, then

$$S_T \approx L(\widehat{\mathbb{X}}_T)$$

for some linear map $L$, where $\widehat{\mathbb{X}}$ denotes the signature of $t \mapsto (t, X_t)$.

$\longrightarrow$ Linear regressions, affine and polynomial technology, and other useful machinery can be applied!

# Signature: definition and properties

## Examples of the signature $\mathbb{X}$ of $X$

### Example

Set $X_t = t$. Then

$$\mathbb{X}_t = (1, t, \tfrac{t^2}{2}, \tfrac{t^3}{6}, \ldots, \tfrac{t^k}{k!}, \ldots).$$

### Example

Let $X$ be a one dimensional continuous semimartingale with $X_0 = 0$. Then

$$\mathbb{X}_t = (1, X_t, \tfrac{X_t^2}{2}, \tfrac{X_t^3}{6}, \ldots, \tfrac{X_t^k}{k!}, \ldots).$$

### Example

Consider $\widehat{X}_t = (t, X_t)$, where $X$ is a one dimensional continuous semimartingale with $X_0 = 0$. Then

$$\widehat{\mathbb{X}}_t = (1, t, X_t, \tfrac{t^2}{2}, \int_0^t s dX_s, \int_0^t X_s ds, \tfrac{X_t^2}{2}, \tfrac{t^3}{6}, \ldots).$$

# Signature of a $d$ dimensional continuous semimartingale

The signature $(\mathbb{X}_t)_{t\in[0,T]}$ of a $d$-dimensional continuous semimartingale $(X_t)_{t\in[0,T]}$ is the process given by

$$\mathbb{X}_t = (\langle e_\emptyset, \mathbb{X}_t\rangle, \langle e_1, \mathbb{X}_t\rangle, \ldots, \langle e_d, \mathbb{X}_t\rangle, \langle e_1 \otimes e_1, \mathbb{X}_t\rangle, \langle e_1 \otimes e_2, \mathbb{X}_t\rangle, \ldots),$$

for $\langle e_\emptyset, \mathbb{X}_t\rangle = 1$ and

$$\langle e_{i_1} \otimes \cdots \otimes e_{i_n}, \mathbb{X}_t\rangle = \int_0^t \langle e_{i_1} \otimes \cdots \otimes e_{i_{n-1}}, \mathbb{X}_s\rangle \circ dX_s^{i_n},$$

where $\circ$ denotes the Stratonivoch integral:

$$\int_0^t Y_t \circ dZ_t = \int_0^t Y_t dZ_t + \frac{1}{2}[Y, Z]_t.$$

Notation: we write $\langle e_I, \mathbb{X}_T\rangle$ for $\langle e_{i_1} \otimes \cdots \otimes e_{i_n}, \mathbb{X}_T\rangle$, where $I = (i_1, \ldots, i_n)$.

# Nice properties

- Linearity: for each $I$, $J$ there is a linear combination of indices $I \sqcup J$ such that

$$\langle e_I, \mathbb{X}_t\rangle\langle e_J, \mathbb{X}_t\rangle = \underbrace{\langle e_I \sqcup e_J, \mathbb{X}_t\rangle}_{\text{linear combination of } \mathbb{X}_t\text{'s elements!}}.$$

  Every polynomial in the signature has a linear representation! Example:

$$\langle e_1, \mathbb{X}_t\rangle^2 = (X_t)^2 \overset{\text{Itô}}{=} 2\int_0^t X_s dX_s + [X]_t = 2\int_0^t X_s \circ dX_s = 2\langle e_1 \otimes e_1, \mathbb{X}_t\rangle.$$

- Uniqueness: the value of the signature of $\widehat{X}_t := (t, X_t)$ at time $T$ uniquely determines the trajectories of $(X_t - X_0)_{t\in[0,T]}$.
  Welcome back Markovianity :).

- Universal approximation theorem: For $K$ compact, $f : K \to \mathbb{R}$ continuous, and $\varepsilon > 0$, there is a finite set $\mathcal{I}$ and $\lambda_I \in \mathbb{R}$ such that

$$|f(\widehat{\mathbb{X}}^2) - \sum_{I\in\mathcal{I}} \lambda_I \langle e_I, \widehat{\mathbb{X}}_T\rangle| 1_{\{\widehat{\mathbb{X}}^2 \in K\}} < \varepsilon,$$

  almost surely.
  Door open for linear approximations!

# Outline

- The model
- A first example
- Calibration to time series data and discussion of the performance
- Calibration to option prices and discussion of the performance
- Conclusion and outlooks

# The model

## The model

Goal: provide a *good* model for a set of *traded assets* $S = (S^1, \dots, S^D)$.

$\rightarrow$*good* = universal, tractable, and easy to calibrate.

Main ingredient: the *market's primary (underlying) process* $\widehat{X}_t := (t, X_t)$.

Requested properties:

- The realizations of $\widehat{X}$ are available in form of time series data and/or the law of $\widehat{X}$ under the pricing measure is known.
- It is reasonable to assume that:
  - $X$ is $d$-dimensional continuous semimartingale.
  - $\widehat{X}$ encodes all the randomness of $S$ in a good way, meaning that the paths of $S$ are continuous maps of the paths of $\widehat{X}$.

The model: $S_n(\ell)_t = (S_n^1(\ell^1)_t, \dots, S_n^D(\ell^D)_t)$, where

$$S_n^j(\ell^j)_t := \ell_\emptyset^j + \sum_{0<|I|\leq n} \ell_I^j \langle e_I, \widehat{\mathbb{X}}_t \rangle,$$

- $\widehat{\mathbb{X}}$ is the signature of $\widehat{X}$,
- $n \in \mathbb{N}$ is the degree of truncation,
- $\ell_\emptyset^j, \ell_I^j \in \mathbb{R}$ are the deterministic coefficients to be found.

See also Perez Arribas, Salvi, Szpruch ('20).

In one sentence: the model

$$S_n^j(\ell^j)_t := \ell^j_\emptyset + \sum_{0<|I|\le n} \ell^j_I \langle e_I, \widehat{\mathbb{X}}_t \rangle,$$

is a linear model whose parameters are $\ell^j_I$ and whose building blocks are

$$\langle e_I, \widehat{\mathbb{X}}_t \rangle = \int_0^t \int_0^{t_n} \cdots \int_0^{t_2} 1 \, d\widehat{X}^{i_1}_{t_1} \cdots d\widehat{X}^{i_n}_{t_n}$$

for some continuous semimartingale $\widehat{X} = (\widehat{X}^0, \widehat{X}^1, \ldots, \widehat{X}^d)$.

## The model: $S_n(\ell)_t := \ell_\emptyset + \sum_{0<|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{X}}_t \rangle$ $(D = 1)$

Flexibility: From the UAT $S$ can be approximated by $S_n(\ell)$.

Universality: Any classical model driven by Brownian motions can be arbitrarily well approximated. Extensions to Lévy driven models are possible (joint work with F. Primavera).

Classical requirements: No arbitrage can easily be guaranteed.

Tractability: Time extended signature of $S_n(\ell)$ can be written as map of $(\ell, \widehat{\mathbb{X}})$.
$\longrightarrow$ Knowing $\mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_t]$, computing an approximation of the price of (path-dependent) options reduces to evaluating a polynomial. Mathematically:

$$\mathbb{E}_{\mathbb{Q}}[F((S_n(\ell)_t)_{t\in[0,T]})] \approx P(\ell, \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_t]),$$

for some some $P$ such that $P(\,\cdot\,, \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_T])$ is polynomial.

$\longrightarrow$ Formulas for the computations of $\mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_t]$ are available if $X$ is a sufficiently regular Markov (or non Markov) diffusion.

Two short excursus

# The model: $S_n(\ell)_t := \ell_\emptyset + \sum_{0<|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{X}}_t \rangle$ $(D = 1)$

Flexibility: $S$ can be approximated by $S_n(\ell)$.

Universality: Any classical model driven by Brownian motions can be arbitrarily well approximated. Extensions to Lévy driven models are possible (joint work with F. Primavera).

Classical requirements: No arbitrage can be guaranteed.

Tractability: Time extended signature of $S_n(\ell)$ can be written as map of $(\ell, \widehat{\mathbb{X}})$.
$\longrightarrow$ Knowing $\mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_t]$, computing an approximation of the price of (path-dependent) options reduces to evaluating a polynomial. Mathematically:

$$\mathbb{E}_{\mathbb{Q}}[F((S_n(\ell)_t)_{t\in[0,T]})] \approx P(\ell, \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_t]),$$

for some some $P$ such that $P(\,\cdot\,, \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_T])$ is polynomial.

$\longrightarrow$ Formulas for the computations of $\mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_t]$ are available if $X$ is a sufficiently regular Markov (or non Markov) diffusion.

# Expected signature: what about the good old polynomial processes?

In $\mathbb{R}^d$, if $dX_t = b(X_t)dt + \sqrt{a(X_t)}dW_t$ for some good $a$ and some linear $b$, then

$$\mathbb{E}[X_t] = \exp(tG)X_0,$$

for some matrix $G$.

Fix $dX_t = b_t dt + \sqrt{a_t}dW_t$. Then

$$d\langle e_I, \mathbb{X}_t\rangle = \left(\langle e_{I'}, \mathbb{X}_t\rangle b_t^{i_n} + \frac{1}{2}\langle e_{I''}, \mathbb{X}_t\rangle a_t^{i_{n-1}i_n}\right)dt + \sigma_t dW_t.$$

If $b_t^i = \langle \mathbf{b}^{(i)}, \mathbb{X}_t\rangle$ and $a_t^{ij} = \langle \mathbf{a}^{(ij)}, \mathbb{X}_t\rangle$ then

$$d\langle e_I, \mathbb{X}_t\rangle = \underbrace{\langle e_{I'} \sqcup\!\sqcup\, \mathbf{b}^{(i_n)} + \frac{1}{2} e_{I''} \sqcup\!\sqcup\, \mathbf{a}^{(i_{n-1}i_n)}, \mathbb{X}_t\rangle}_{\text{Linear map of } \mathbb{X}_t!} dt + \sigma_t dW_t,$$

for some $\sigma$. Hence, under some technical conditions,

$$\mathbb{E}[\mathbb{X}_t] = \exp(tG)\mathbb{X}_0,$$

for some (potentially infinite dimensional) matrix $G$.

## The model: $S_n(\ell)_t := \ell_\emptyset + \sum_{0<|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{X}}_t \rangle$ $(D = 1)$

Flexibility: $S$ can be approximated by $S_n(\ell)$.

Universality: Any classical model driven by Brownian motions can be arbitrarily well approximated. Extensions to Lévy driven models are possible (joint work with F. Primavera).

Classical requirements: No arbitrage can be guaranteed.

Tractability: Time extended signature of $S_n(\ell)$ can be written as map of $(\ell, \widehat{\mathbb{X}})$.
$\longrightarrow$ Knowing $\mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_t]$, computing an approximation of the price of (path-dependent) options reduces to evaluating a polynomial. Mathematically:

$$\mathbb{E}_{\mathbb{Q}}[F((S_n(\ell)_t)_{t\in[0,T]})] \approx P(\ell, \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_t]),$$

for some some $P$ such that $P(\,\cdot\,, \mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_T])$ is polynomial.

$\longrightarrow$ Formulas for the computations of $\mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}_t]$ are available if $X$ is a sufficiently regular Markov (or non Markov) diffusion.

## The model $S_n(\ell)_t := \ell_\emptyset + \sum_{0<|I|\leq n} \ell_I \langle e_I, \widehat{\mathbb{X}}_t \rangle$: what about Taylor?

Let $X$ be a Brownian motion and $S$ be a nice stochastic process driven by $X$.

Then there are $\ell$ such that

$$\mathbb{E}[|S_t - S_n(\ell)_t|] = o(t^{n/2}).$$

How? Observe that

$$\begin{aligned} S_t &= S_0 + \int_0^t \underbrace{c_0(s)}_{=c_0(0)+\int_0^s c_{00}(r)dr+\int_0^s c_{10}(r)dX_r} ds + \int_0^t \underbrace{c_1(s)}_{=c_1(0)+\int_0^s c_{01}(r)dr+\int_0^s c_{11}(r)dX_r} dX_s \\ &= \underbrace{S_0 + c_0(0)t + c_1(0)X_t}_{=S_1(\ell)_t} + \underbrace{(\text{linear combination of double integrals})}_{=o(t^{1/2})} \end{aligned}$$

for $\ell_\emptyset = S_0, \ell_0 = c_0(0)$, and $\ell_1 = c_1(0)$.

Moments, characteristic function, and every possible property of $X$ and its signature is very well understood.

One can study the asymptotic properties of $S$ using the asymptotic properties of $\widehat{\mathbb{X}}$! (Joint ongoing work with F. Bandi and R. Renò)

## The model $S_n(\ell)_t := \ell_\emptyset + \sum_{0<|I|\le n} \ell_I \langle e_I, \widehat{\mathbb{X}}_t \rangle$: what about Taylor?

Example: Edgeworth expansion for the normalized characteristic function

$$
\begin{aligned}
&\mathbb{E}\Big[\exp\Big(iu\frac{S_t - c_0 t}{c_1\sqrt{t}}\Big)\Big] e^{\frac{u^2}{2}} \\
&\quad = 1 + \Big[-\frac{c_{11}}{c_1}\frac{i}{2}u^3\Big]\sqrt{t} \\
&\qquad + \frac{1}{2}\Big[-\Big(\frac{c_{01}}{c_1} + \frac{c_{10}}{c_1}\Big)u^2 + \Big(\frac{c_{11}}{c_1}\Big)^2\Big(-\frac{1}{2}u^2 + u^4 - \frac{1}{4}u^6\Big)\Big]t \\
&\qquad + \frac{1}{2}\Big[\Big(\frac{c_{21}}{c_1}\Big)^2\Big(\frac{1}{3}u^4 - \frac{1}{2}u^2\Big) - \frac{c_{111}}{c_1}\frac{i}{6}u^3\Big]t \\
&\qquad + o(t)
\end{aligned}
$$

See for instance Todorov ('21) or Bandi, Renò (to appear).

Calibration to time-series data

# Calibration to time-series data

Model: $S_{n+1}(\ell)_t := S_{n+1}(\ell)_0 + \ell_\emptyset \langle \tilde{e}_\emptyset, \widehat{\mathbb{X}}_t \rangle + \sum_{0<|I|\leq n} \ell_I \langle \tilde{e}_I, \widehat{\mathbb{X}}_t \rangle$, for

$$\langle \tilde{e}_I, \widehat{\mathbb{X}}_t \rangle = \int_0^t \langle e_I, \widehat{\mathbb{X}}_s \rangle dX_s.$$

Scenario: The realizations of the *market's primary (underlying) process* $\widehat{X}$ are available in form of time series data: $\widehat{X}_{t_1}, \ldots, \widehat{X}_{t_N}$.

Procedure:

- Compute the paths of the signature $\widehat{\mathbb{X}}$ (e.g. using `iisignature` in python).
- Use the paths of $\widehat{\mathbb{X}}$ as linear regression basis to find $\ell$ matching the prices, i.e. minimizing the expression:

$$\sum_{i=1}^{N} (S_{n+1}(\ell)_{t_i} - S_{t_i})^2 = \sum_{i=1}^{N} \left( S_0 + \ell_\emptyset \langle \tilde{e}_\emptyset, \widehat{\mathbb{X}}_{t_i} \rangle + \sum_{0<|I|<n} \ell_I \langle \tilde{e}_I, \widehat{\mathbb{X}}_{t_i} \rangle - S_{t_i} \right)^2.$$

# Out of sample result for a Heston market model

- Consider a Heston model (d=2, D=1):

$$dS_t = \mu S_t dt + S_t \sqrt{V_t} dB_t^{\mathbb{P}}$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma \sqrt{V_t} dW_t^{\mathbb{P}},$$

- Goal: approximate $S$ with $S_3(\ell^*)$, using the estimated $\mathbb{Q}$-Brownian motions as primary underlying process ($\ell^* \in \mathbb{R}^{13}$).
- Test: Simulate a new trajectory of the Heston model, extract the corresponding $\mathbb{Q}$-Brownian motions, use them to compute the trajectory of $S_3(\ell^*)$, compare the obtained process with $S$.

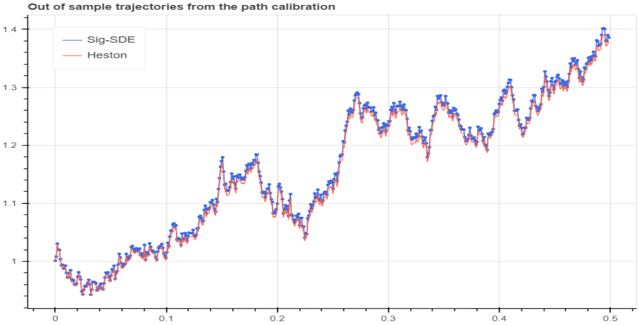



Figure: Out of sample performance over 0,5 years with $\sigma = 0.25$.

# Calibration to option prices

## Calibration to option prices

Model: $S_{n+1}(\ell)_t := S_{n+1}(\ell)_0 + \ell_\emptyset \langle \tilde{e}_\emptyset, \widehat{\mathbb{X}}_t \rangle + \sum_{0<|I|\leq n} \ell_I \langle \tilde{e}_I, \widehat{\mathbb{X}}_t \rangle$. $(D = d = 1)$

Scenario: The following quantities are available:

- Prices of options on $S$.
- The law of the *market's primary (underlying) process* $\widehat{X}$ under the pricing measure $\mathbb{Q}$.

Cool idea: Since computing the approximated price of an (even path dependent) option with the proposed model reduces to evaluating a polynomial, calibration on (even path dependent) option prices could be done in a simple and efficient way.

$\longrightarrow$ ...cool but dangerous! The given approximation has to be good enough in each optimization's step!

Alternative idea: Use Monte Carlo pricing (with variance reduction). Note that there is no need of new simulations in the optimization procedure.

## Calibration to option prices: procedure

Scenario: The following quantities are available:

- Prices $\pi_1, \ldots, \pi_N$ of $N$ options with payoffs

$$F_1((S_t)_{t\in[0,T_1]}), \ldots, F_N((S_t)_{t\in[0,T_N]}).$$

Procedure:

- Look for $\ell$ matching the corresponding option prices, i.e. minimizing the expression

$$\sum_{i=1}^{N} w^i \left(P_i^{MC}(\ell) - \pi^i\right)^2,$$

for some weights $w^i$, where $P^{MC}(\ell)$ denotes the empirical mean of

$$F_i\left((S_n(\ell)_t)_{t\in[0,T_i]})\right).$$

Important observation: the linearity of the model makes this procedure very quick. Trajectories of $\widehat{\mathbb{X}}$ could be simulated just once in advance and stored. A coefficients update reduces to a scalar product.

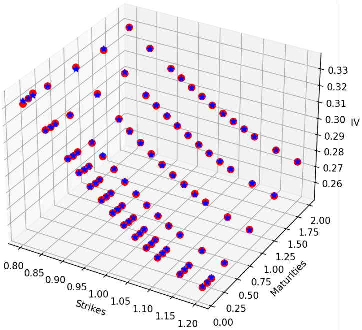
# Calibration to option prices: the Heston model

- Consider a Heston model (d=2, D=1):

$$dS_t = \mu S_t dt + S_t \sqrt{V_t} dB_t^{\mathbb{P}}$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t^{\mathbb{P}},$$

- Goal: approximate $S$ with $S_3(\ell^*)$, using two $\mathbb{Q}$-Brownian motions as primary underlying process ($\ell^* \in \mathbb{R}^{13}$).
- Test: Compute the implied volatility surface (using Monte Carlo) under $S_3(\ell^*)$ (red) and compare it with the Heston's one (blue).



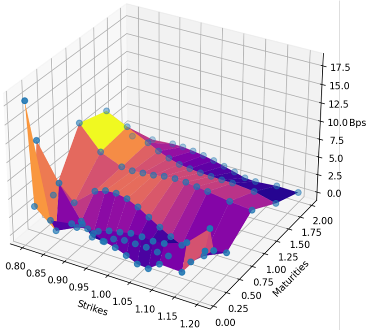




Figure: IVSs and corresponding absolute error (7 maturities from 30 days to 2 years).

# Calibration to option prices: S&P 500 17.03.2021

- Let $S$ be the stochastic process describing the price of S&P 500 starting at day 17.03.2021.
- Goal: approximate $S$ with $S_4(\ell^*)$, using two $\mathbb{Q}$-Brownian motions as primary underlying process ($\ell^* \in \mathbb{R}^{121}$).
- Test: Compute the implied volatility surface (using Monte Carlo) under $S_4(\ell^*)$ and compare it with the market's one.

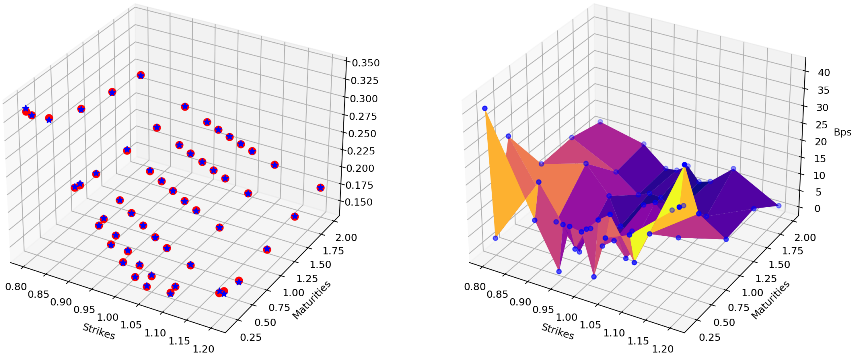



Figure: IVSs and corresponding absolute error (6 maturities within 60 days and 2 years).

## Remarks on the previous example

- The result is obtained using a closer-to-sup-norm loss function:

$$\sum_{i=1}^{N} \alpha \varepsilon^i \left( P_i^{MC}(\ell) - \pi^i \right)^p,$$

where $\varepsilon_i$ is the absolute error for the $i$-th price in a previous calibration and $\alpha$ and $p$ are big.

- The calibrated model produces a reasonable implied volatility surface also for out of sample strikes and maturities.

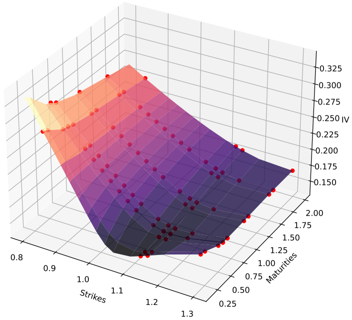



Figure: IVS of the calibrated model (6 maturities within 60 days and 2 years). Out of sample represented as red dots.

# Conclusions

# Conclusions

- We saw that from a mathematical point of view signatures have some extremely interesting properties and deserve to be used in a modeling context.

⇒ $F((X_t)_{t\in[0,T]}) \approx L(\widehat{\mathbb{X}}_T)$ for some linear map $L$.

- We introduced a linear model based on the signature of an underlying process.

⇒ Flexible: classical models can be approximated arbitrarily well.

⇒ Tractable: since as soon as $\mathbb{E}_{\mathbb{Q}}[\widehat{\mathbb{X}}]$ is known, estimators for different quantities are available in closed form.

- We illustrated two calibration methods showing the corresponding performances on simulated and real data.

**Thank you for your attention!**