Rare event simulation related to financial risks: Efficient estimation and sensitivity analysis

Ankush Agarwal
joint work with
Stefano De Marco, Emmanuel Gobet, Gang Liu

Centre de Mathématiques Appliquées, Ecole Polytechnique

International Conference on Monte Carlo Techniques, July 5-8 2016

With the support of La FONDATION du RISQUE

Motivation

Compute

$$p = \mathbb{P}(X \in A), \quad \alpha = \mathbb{E}(\varphi(X)\mathbf{1}_{X\in A}), \quad \text{when } p < 10^{-4}$$

Motivation

Compute

$$p = \mathbb{P}(X \in A), \quad \alpha = \mathbb{E}(\varphi(X)\mathbf{1}_{X\in A}), \quad \text{when } p < 10^{-4}$$

Goal: Compute p

Motivation

Compute

$$p = \mathbb{P}(X \in A), \quad \alpha = \mathbb{E}(\varphi(X)\mathbf{1}_{X \in A}), \quad \text{when } p < 10^{-4}$$

<u>Goal</u>: Compute p

☑ Generate $(X_i)_{1 \leqslant i \leqslant N}$ i.i.d. copies of X & set $\bar{S}_N = (N)^{-1} \sum_{i=1}^{N} \mathbf{1}_{X_i \in A}$

Motivation

Compute

$$p = \mathbb{P}(X \in A), \quad \alpha = \mathbb{E}(\varphi(X)\mathbf{1}_{X\in A}), \quad \text{when } p < 10^{-4}$$

Goal: Compute p

- ☑ Generate $(X_i)_{1\leqslant i\leqslant N}$ i.i.d. copies of X & set $\bar{S}_N = (N)^{-1}\sum_{i=1}^{N}\mathbf{1}_{X_i\in A}$
- ☑ By the Central Limit Theorem (CLT)

$$\sqrt{N}(\bar{S}_N - \mathbb{P}(X\in A)) \to \mathcal{N}(0, p(1-p))$$

Motivation

Compute

$$p = \mathbb{P}(X \in A), \quad \alpha = \mathbb{E}(\varphi(X)\mathbf{1}_{X \in A}), \quad \text{when } p < 10^{-4}$$

Goal: Compute p

- ☑ Generate $(X_i)_{1 \leqslant i \leqslant N}$ i.i.d. copies of X & set $\bar{S}_N = (N)^{-1} \sum_{i=1}^{N} \mathbf{1}_{X_i \in A}$
- ☑ By the Central Limit Theorem (CLT)

$$\sqrt{N}(\bar{S}_N - \mathbb{P}(X \in A)) \to \mathcal{N}(0, p(1-p))$$

- ☑ 95% confidence interval

$$\left(\bar{S}_N - 1.96\sqrt{\frac{p(1-p)}{N}}, \bar{S}_N + 1.96\sqrt{\frac{p(1-p)}{N}}\right)$$

Motivation

Compute

$$p = \mathbb{P}(X \in A), \quad \alpha = \mathbb{E}(\varphi(X)\mathbf{1}_{X\in A}), \quad \text{when } p < 10^{-4}$$

Goal: Compute p

- ☑ Generate $(X_i)_{1\leqslant i\leqslant N}$ i.i.d. copies of X & set $\bar{S}_N = (N)^{-1}\sum_{i=1}^{N}\mathbf{1}_{X_i\in A}$
- ☑ By the Central Limit Theorem (CLT)

$$\sqrt{N}(\bar{S}_N - \mathbb{P}(X\in A)) \to \mathcal{N}(0, p(1-p))$$

- ☑ 95% confidence interval

$$\left(\bar{S}_N - 1.96\sqrt{\frac{p(1-p)}{N}}, \bar{S}_N + 1.96\sqrt{\frac{p(1-p)}{N}}\right)$$

- ☑ Relative error: $\sqrt{p(1-p)}/(p\sqrt{N}) \approx (\sqrt{Np})^{-1}$ is large for small p

Examples

☑ Estimation of default probabilities in pricing of Credit Default Spreads:

N_0	firms
$(X_i)_{1\leqslant i\leqslant N_0}$	capital reserve
B	default threshold
T	time horizon

$$\mathbb{P}\Big(\sum_{i=1}^{N_0} \mathbf{1}_{\{\inf\left\{t: X_i(t)\leqslant B\right\}\leqslant T\}} > L\Big), \quad 0 \leqslant L \leqslant N_0 - 1$$

☑ Pricing Deep out-of-the-money options:

K	strike
N_0	assets
$(X_i)_{1\leqslant i\leqslant N_0}$	asset price
T	maturity

$$\mathbb{E}\Big((K - {N_0}^{-1} \sum_{i=1}^{N_0} X_i(T))\mathbf{1}_{{N_0}^{-1}\sum_{i=1}^{N_0} X_i(T)<K}\Big), K > 0$$

Motivation

Sensitivity

$$\Delta_\theta = \partial\alpha/\partial\theta = \partial\mathbb{E}(\varphi(X^\theta)\mathbf{1}_{X^\theta\in A})/\partial\theta$$

Motivation

Sensitivity

$$\Delta_\theta = \partial \alpha / \partial \theta = \partial \mathbb{E}(\varphi(X^\theta)\mathbf{1}_{X^\theta \in A}) / \partial \theta$$

☑ Generate $(X_i^\theta)_{1 \leqslant i \leqslant N}$ i.i.d. copies of X^θ

Motivation

Sensitivity

$$\Delta_\theta = \partial\alpha/\partial\theta = \partial\mathbb{E}(\varphi(X^\theta)\mathbf{1}_{X^\theta \in A})/\partial\theta$$

- ☑ Generate $(X_i^\theta)_{1 \leqslant i \leqslant N}$ i.i.d. copies of X^θ
- ☑ Generate another set of N i.i.d. copies of $X^{\theta+\Delta\theta}$ where $\theta + \Delta\theta$ is perturbed value of the parameter

Motivation

Sensitivity

$$\Delta_\theta = \partial\alpha/\partial\theta = \partial\mathbb{E}(\varphi(X^\theta)\mathbf{1}_{X^\theta\in A})/\partial\theta$$

- Generate $(X_i^\theta)_{1\leqslant i\leqslant N}$ i.i.d. copies of X^θ
- Generate another set of N i.i.d. copies of $X^{\theta+\Delta\theta}$ where $\theta+\Delta\theta$ is perturbed value of the parameter
- Use finite differences to estimate the sensitivity

$$\frac{1}{\Delta\theta}\left(N^{-1}\sum_{i=1}^{N}\varphi(X_i^{\theta+\Delta\theta})\mathbf{1}_{X^{\theta+\Delta\theta}\in A} - N^{-1}\sum_{i=1}^{N}\varphi(X_i^\theta)\mathbf{1}_{X^\theta\in A}\right)$$

Motivation

Sensitivity

$$\Delta_\theta = \partial\alpha/\partial\theta = \partial\mathbb{E}(\varphi(X^\theta)\mathbf{1}_{X^\theta\in A})/\partial\theta$$

- Generate $(X_i^\theta)_{1\leqslant i\leqslant N}$ i.i.d. copies of X^θ
- Generate another set of N i.i.d. copies of $X^{\theta+\Delta\theta}$ where $\theta+\Delta\theta$ is perturbed value of the parameter
- Use finite differences to estimate the sensitivity

$$\frac{1}{\Delta\theta}\left(N^{-1}\sum_{i=1}^{N}\varphi(X_i^{\theta+\Delta\theta})\mathbf{1}_{X^{\theta+\Delta\theta}\in A} - N^{-1}\sum_{i=1}^{N}\varphi(X_i^{\theta})\mathbf{1}_{X^{\theta}\in A}\right)$$

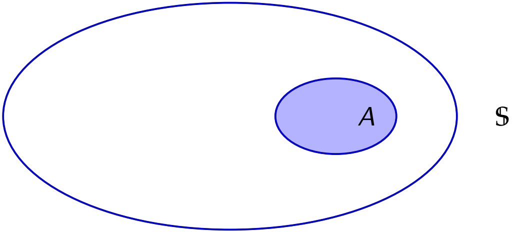
- Unstable estimates due to the presence of indicator function

General Rare Event Simulation Approaches

☑ Reformulate p using conditional probabilities

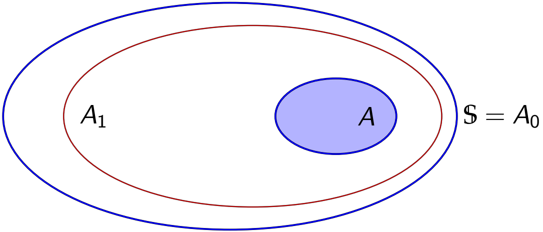
General Rare Event Simulation Approaches

☑ Reformulate p using conditional probabilities



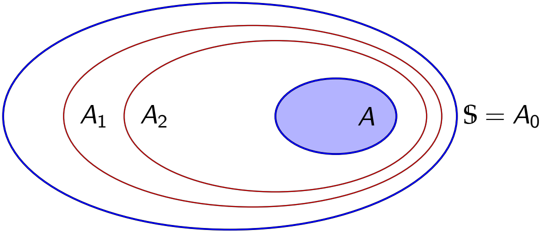
General Rare Event Simulation Approaches

- ☑ Reformulate p using conditional probabilities



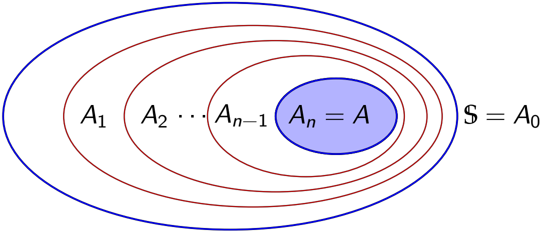
General Rare Event Simulation Approaches

☑ Reformulate p using conditional probabilities



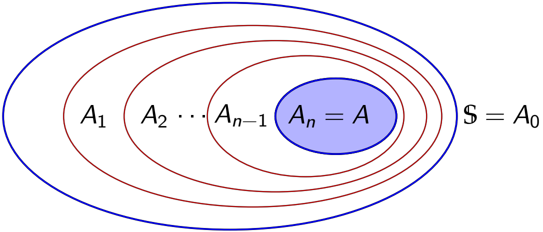
General Rare Event Simulation Approaches

☑ Reformulate p using conditional probabilities



General Rare Event Simulation Approaches

☑ Reformulate p using conditional probabilities



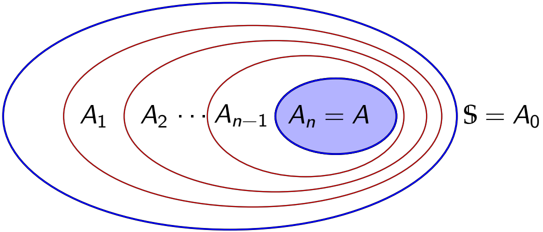
☑ Define a series of nested subsets on probability space $\mathbb{S}$

$$\mathbb{S} := A_0 \supset \cdots \supset A_k \supset \cdots \supset A_n := A$$

$$\mathbb{P}(X \in A) = \prod_{k=1}^{n} \mathbb{P}(X \in A_k | X \in A_{k-1})$$

General Rare Event Simulation Approaches

☑ Reformulate p using conditional probabilities



☑ Define a series of nested subsets on probability space $\mathbb{S}$

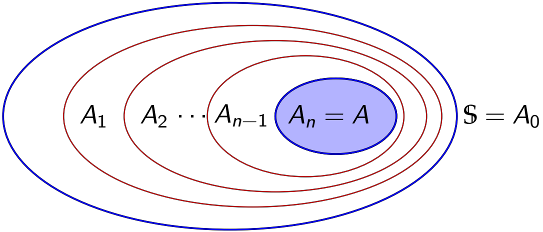
$$\mathbb{S} := A_0 \supset \cdots \supset A_k \supset \cdots \supset A_n := A$$

$$\mathbb{P}(X \in A) = \prod_{k=1}^{n} \mathbb{P}(X \in A_k | X \in A_{k-1})$$

☑ How to estimate $\mathbb{P}(X \in A_k | X \in A_{k-1})$?

General Rare Event Simulation Approaches

☑ Reformulate p using conditional probabilities



☑ Define a series of nested subsets on probability space $\mathbb{S}$

$$\mathbb{S} := A_0 \supset \cdots \supset A_k \supset \cdots \supset A_n := A$$

$$\mathbb{P}(X \in A) = \prod_{k=1}^{n} \mathbb{P}(X \in A_k | X \in A_{k-1})$$

☑ How to estimate $\mathbb{P}(X \in A_k | X \in A_{k-1})$?

☑ Existing methods: splitting/restart, Interacting Particles System (IPS)

Shaking Transformation

☑ New method (Parallel-One-Path) for rare event simulation using the ergodicity of Markov chain [Gobet and Liu, 2015]

Shaking Transformation

- ☑ New method (Parallel-One-Path) for rare event simulation using the ergodicity of Markov chain [Gobet and Liu, 2015]
- ☑ Based on the idea of Shaking Transformation for random objects

Shaking Transformation

- ☑ New method (Parallel-One-Path) for rare event simulation using the ergodicity of Markov chain [Gobet and Liu, 2015]
- ☑ Based on the idea of Shaking Transformation for random objects

Given $X, \mathcal{K}(\cdot)$ is a reversible shaking transformation for X if:

$$(X, \mathcal{K}(X)) \stackrel{\mathrm{d}}{=} (\mathcal{K}(X), X)$$

Also, $\mathcal{K}(X) := K(X, Y)$ where K is deterministic and Y is independent of X

Shaking Transformation

- ☑ New method (Parallel-One-Path) for rare event simulation using the ergodicity of Markov chain [Gobet and Liu, 2015]

- ☑ Based on the idea of Shaking Transformation for random objects

Given $X, \mathcal{K}(\cdot)$ is a reversible shaking transformation for X if:

$$(X, \mathcal{K}(X)) \overset{\mathrm{d}}{=} (\mathcal{K}(X), X)$$

Also, $\mathcal{K}(X) := K(X, Y)$ where K is deterministic and Y is independent of X

- ☑ If X is a standard Brownian motion (SBM), X' independent SBM

$$K(X, X') = \left(\int_0^t \rho_s dX_s + \int_0^t \sqrt{1-\rho_s^2} dX'_s\right)_{0 \leqslant t \leqslant T}$$

Shaking Transformation

- ☑ New method (Parallel-One-Path) for rare event simulation using the ergodicity of Markov chain [Gobet and Liu, 2015]
- ☑ Based on the idea of Shaking Transformation for random objects

Given $X, \mathcal{K}(\cdot)$ is a reversible shaking transformation for X if:

$$(X, \mathcal{K}(X)) \stackrel{\mathrm{d}}{=} (\mathcal{K}(X), X)$$

Also, $\mathcal{K}(X) := K(X, Y)$ where K is deterministic and Y is independent of X

- ☑ If X is a standard Brownian motion (SBM), X' independent SBM

$$K(X, X') = \left(\int_0^t \rho_s dX_s + \int_0^t \sqrt{1-\rho_s^2} dX'_s\right)_{0 \leqslant t \leqslant T}$$

- ☑ If X is a Poisson Process

$$K(X, [\mathrm{Bin}(X, 1-p), \mathrm{Poisson}(p\lambda)]) = \mathrm{Bin}(X, 1-p) + \mathrm{Poisson}(p\lambda)$$

Shaking Normal Random Variable

$$K(X, \mathcal{N}(0,1)) = \rho X + \sqrt{1-\rho^2}\mathcal{N}(0,1), -1 \leqslant \rho \leqslant 1$$

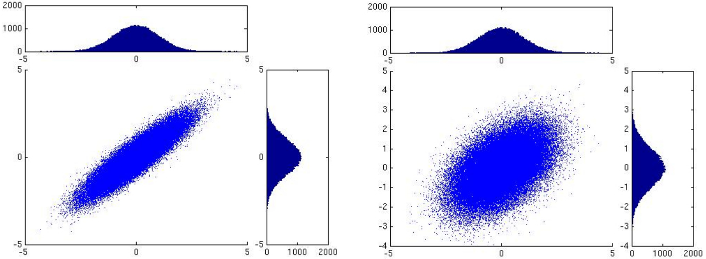


Figure: Plots of $(X, \mathcal{K}(X))$ with $\rho = 0.9$ and $\rho = 0.5$

Shaking with Rejection

Let $k \in \{0, 1, \cdots, n-1\}$, define

$$\mathcal{M}_k^{\mathcal{K}}(X) := \begin{cases} \mathcal{K}(X) & \text{if } \mathcal{K}(X) \in A_k \\ X & \text{if } \mathcal{K}(X) \notin A_k. \end{cases}$$

Shaking with Rejection

Let $k \in \{0, 1, \cdots, n-1\}$, define

$$\mathcal{M}_k^{\mathcal{K}}(X) := \begin{cases} \mathcal{K}(X) & \text{if } \mathcal{K}(X) \in A_k \\ X & \text{if } \mathcal{K}(X) \notin A_k. \end{cases}$$

Proposition (Conditional Invariance (Gobet and Liu, '15))

Let $k \in \{0, 1, \cdots, n-1\}$. The distribution of X conditionally on $\{X \in A_k\}$ is invariant w.r.t. the random transformation $\mathcal{M}_k^{\mathcal{K}}$: i.e. for any bounded (random) measurable $\varphi : \mathbb{S} \to \mathbb{R}$, we have

$$\mathbb{E}\big(\varphi(\mathcal{M}_k^{\mathcal{K}}(X))|X \in A_k\big) = \mathbb{E}\big(\varphi(X)|X \in A_k\big)$$

Shaking with Rejection

Let $k \in \{0, 1, \cdots, n-1\}$, define

$$\mathcal{M}_k^{\mathcal{K}}(X) := \begin{cases} \mathcal{K}(X) & \text{if } \mathcal{K}(X) \in A_k \\ X & \text{if } \mathcal{K}(X) \notin A_k. \end{cases}$$

Proposition (Conditional Invariance (Gobet and Liu, '15))

Let $k \in \{0, 1, \cdots, n-1\}$. The distribution of X conditionally on $\{X \in A_k\}$ is invariant w.r.t. the random transformation $\mathcal{M}_k^{\mathcal{K}}$: i.e. for any bounded (random) measurable $\varphi : \mathbb{S} \to \mathbb{R}$, we have

$$\mathbb{E}\big(\varphi(\mathcal{M}_k^{\mathcal{K}}(X))|X \in A_k\big) = \mathbb{E}\big(\varphi(X)|X \in A_k\big)$$

Proof.

$$\begin{aligned}\mathbb{E}\big(\varphi(\mathcal{M}_k^{\mathcal{K}}(X))\mathbf{1}_{X\in A_k}\big) &= \mathbb{E}\big(\varphi(K(X))\mathbf{1}_{X\in A_k}\mathbf{1}_{K(X)\in A_k}\big) + \mathbb{E}\big(\varphi(X)\mathbf{1}_{X\in A_k}\mathbf{1}_{K(X)\notin A_k}\big) \\ &= \mathbb{E}\big(\varphi(X)\mathbf{1}_{X\in A_k}\mathbf{1}_{K(X)\in A_k}\big) + \mathbb{E}\big(\varphi(X)\mathbf{1}_{X\in A_k}\mathbf{1}_{K(X)\notin A_k}\big)\end{aligned}$$

□

Parallel-One-Path (POP) Method

Parallel-One-Path (POP) Method

☑ Recall: How to estimate $\mathbb{P}(X \in A_{k+1} | X \in A_k)$?

Parallel-One-Path (POP) Method

☑ Recall: How to estimate $\mathbb{P}(X \in A_{k+1} | X \in A_k)$?

☑ Key idea:

- ⋆ Use $\mathcal{M}_k^{\mathcal{K}}$ (shaking with rejection) to create Markov chains directly in a path space
- ⋆ Conditional invariance of $\mathcal{M}_k^{\mathcal{K}}$ with respect to $X | X \in A_k$ enables to use the ergodic property of Markov chain

Parallel-One-Path (POP) Method

- ☑ Recall: How to estimate $\mathbb{P}(X \in A_{k+1} | X \in A_k)$?
- ☑ Key idea:
  - ⋆ Use $\mathcal{M}_k^{\mathcal{K}}$ (shaking with rejection) to create Markov chains directly in a path space
  - ⋆ Conditional invariance of $\mathcal{M}_k^{\mathcal{K}}$ with respect to $X | X \in A_k$ enables to use the ergodic property of Markov chain

    Birkhoff's Ergodic Theorem:

    For ergodic Markov chain $(X_i)_{i \geqslant 0}$ with a unique invariant distribution π and measurable f:

$$\frac{1}{N} \sum_{i=0}^{N-1} f(X_i) \underset{N \to +\infty}{\longrightarrow} \int f \mathrm{d}\pi \qquad a.s.$$

Parallel-One-Path (POP) Method

- Recall: How to estimate $\mathbb{P}(X \in A_{k+1} | X \in A_k)$?
- Key idea:
  - ⋆ Use $\mathcal{M}_k^{\mathcal{K}}$ (shaking with rejection) to create Markov chains directly in a path space
  - ⋆ Conditional invariance of $\mathcal{M}_k^{\mathcal{K}}$ with respect to $X | X \in A_k$ enables to use the ergodic property of Markov chain

  Birkhoff's Ergodic Theorem:

  For ergodic Markov chain $(X_i)_{i \geqslant 0}$ with a unique invariant distribution π and measurable f:

$$\frac{1}{N} \sum_{i=0}^{N-1} f(X_i) \underset{N \to +\infty}{\longrightarrow} \int f \mathrm{d}\pi \qquad a.s.$$

Given an initial position $X_{k,0} \in A_k$, define $X_{k,i} := \mathcal{M}_k^{\mathcal{K}}(X_{k,i-1})$

$$\mathbb{E}(\varphi(X) | X \in A_k) \approx \frac{1}{N} \sum_{i=0}^{N-1} \varphi(X_{k,i})$$

Parallel-One-Path (POP) Method

- ☑ Recall: How to estimate $\mathbb{P}(X \in A_{k+1} | X \in A_k)$?
- ☑ Key idea:
  - ⋆ Use $\mathcal{M}_k^{\mathcal{K}}$ (shaking with rejection) to create Markov chains directly in a path space
  - ⋆ Conditional invariance of $\mathcal{M}_k^{\mathcal{K}}$ with respect to $X | X \in A_k$ enables to use the ergodic property of Markov chain

  Birkhoff's Ergodic Theorem:

  For ergodic Markov chain $(X_i)_{i \geqslant 0}$ with a unique invariant distribution π and measurable f:

$$\frac{1}{N} \sum_{i=0}^{N-1} f(X_i) \underset{N \to +\infty}{\longrightarrow} \int f \mathrm{d}\pi \qquad a.s.$$

Given an initial position $X_{k,0} \in A_k$, define $X_{k,i} := \mathcal{M}_k^{\mathcal{K}}(X_{k,i-1})$

$$\mathbb{E}(\varphi(X) | X \in A_k) \approx \frac{1}{N} \sum_{i=0}^{N-1} \varphi(X_{k,i})$$

Estimators for each $P(X \in A_{k+1} | X \in A_k)$ can be made independent!

Adaptive POP Method

☑ Number of levels in POP depend on order of p

Adaptive POP Method

- ☑ Number of levels in POP depend on order of p
- ☑ Create levels adaptively for fixed conditional probability $q \approx 10^{-1}$

Adaptive POP Method

- ☑ Number of levels in POP depend on order of p
- ☑ Create levels adaptively for fixed conditional probability $q \approx 10^{-1}$
- ☑ Shaking with rejection for each level depends on the previous levels

Adaptive POP Method

- ☑ Number of levels in POP depend on order of p
- ☑ Create levels adaptively for fixed conditional probability $q \approx 10^{-1}$
- ☑ Shaking with rejection for each level depends on the previous levels
- ☑ True value $p := r(Q_q^{L^*})q^{L^*}$, estimator given as $\hat{p}_N := \hat{r}_N(\hat{Q}_{N,q}^{L_N})q^{L_N}$ where Q_q^l and $\hat{Q}_{N,q}^l$ are the quantile levels and their estimators

Adaptive POP Method

- ☑ Number of levels in POP depend on order of p
- ☑ Create levels adaptively for fixed conditional probability $q \approx 10^{-1}$
- ☑ Shaking with rejection for each level depends on the previous levels
- ☑ True value $p := r(Q_q^{L^*})q^{L^*}$, estimator given as $\hat{p}_N := \hat{r}_N(\hat{Q}_{N,q}^{L_N})q^{L_N}$ where Q_q^l and $\hat{Q}_{N,q}^l$ are the quantile levels and their estimators
- ☑ Use conditioning arguments and locally uniform probabilistic error bounds on quantile estimators

Adaptive POP Method

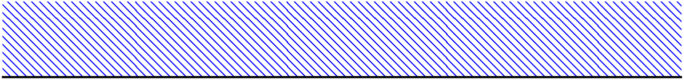
- ☑ Number of levels in POP depend on order of p
- ☑ Create levels adaptively for fixed conditional probability $q \approx 10^{-1}$
- ☑ Shaking with rejection for each level depends on the previous levels
- ☑ True value $p := r(Q_q^{L^*})q^{L^*}$, estimator given as $\hat{p}_N := \hat{r}_N(\hat{Q}_{N,q}^{L_N})q^{L_N}$ where Q_q^l and $\hat{Q}_{N,q}^l$ are the quantile levels and their estimators
- ☑ Use conditioning arguments and locally uniform probabilistic error bounds on quantile estimators

Theorem

Under some assumptions on conditional quantile estimator error bounds, for $N \to \infty$,

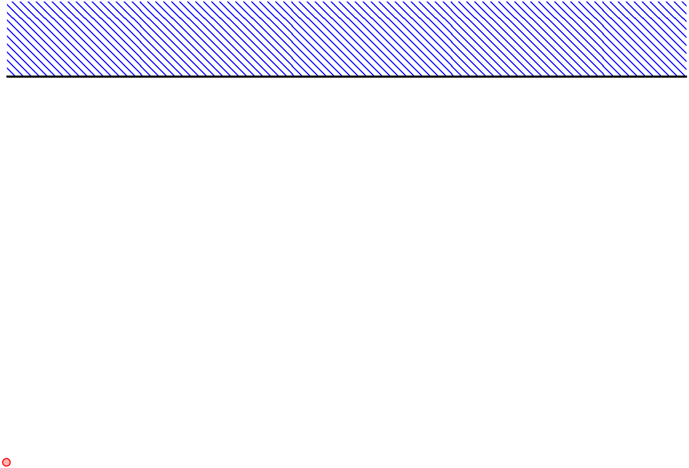
$$\hat{p}_N \to p \quad a.s.$$

Adaptive Method



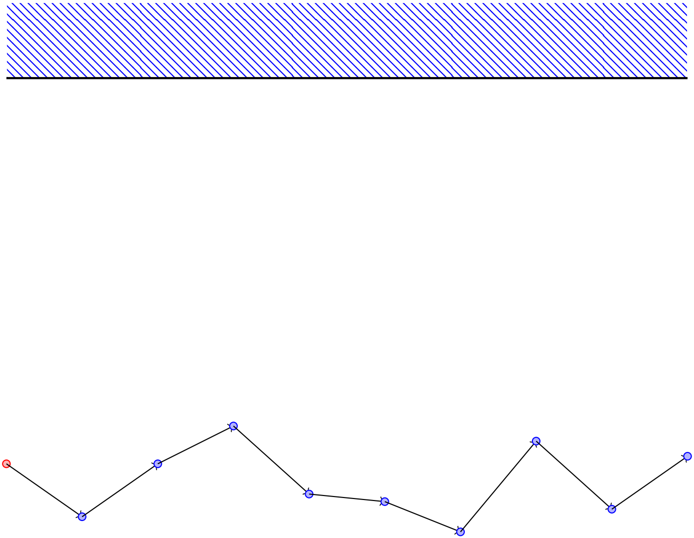
Rare event level

Adaptive Method



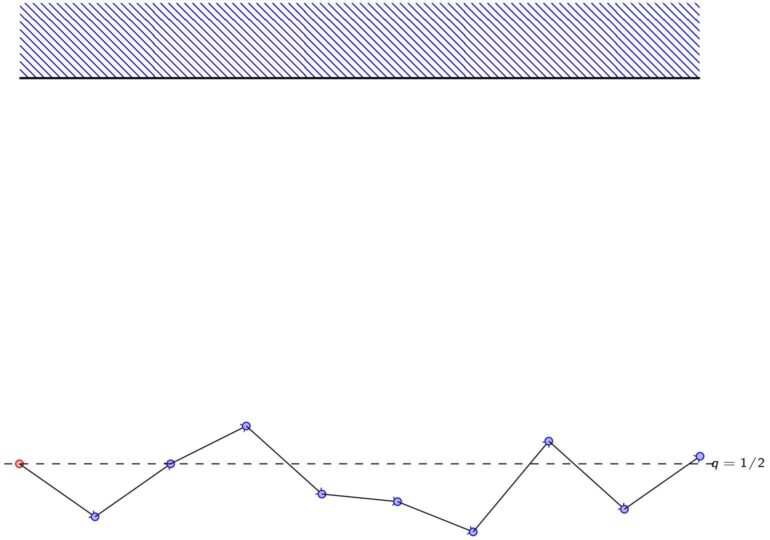
$A_0 = \$$

Adaptive Method



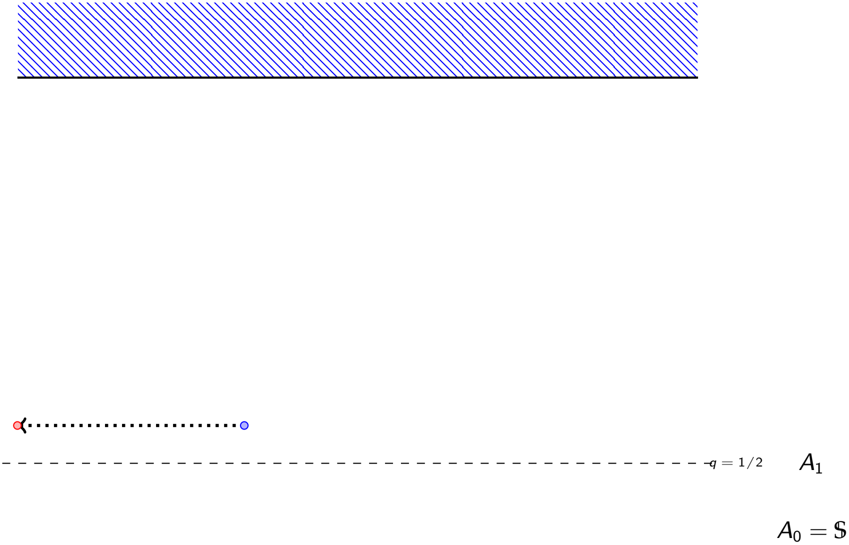
$A_0 = \$$

Adaptive Method

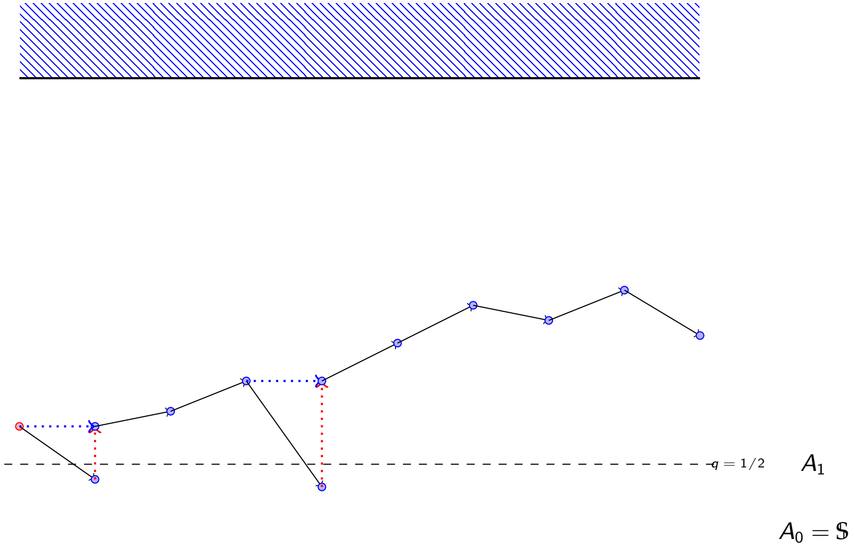


$A_0 = \$$

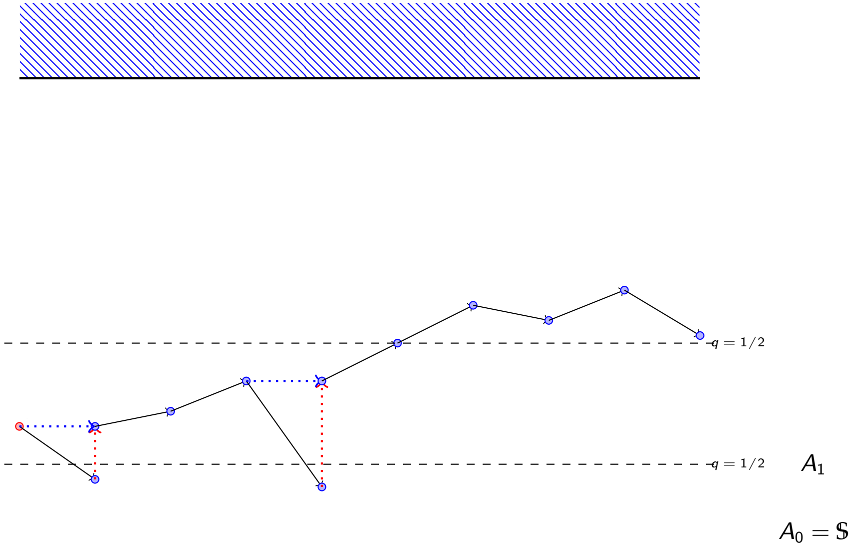
Adaptive Method



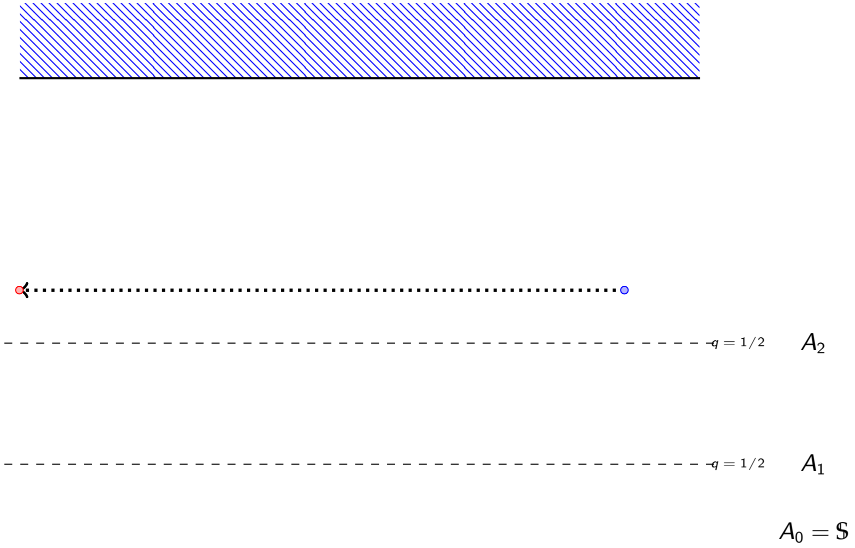
Adaptive Method



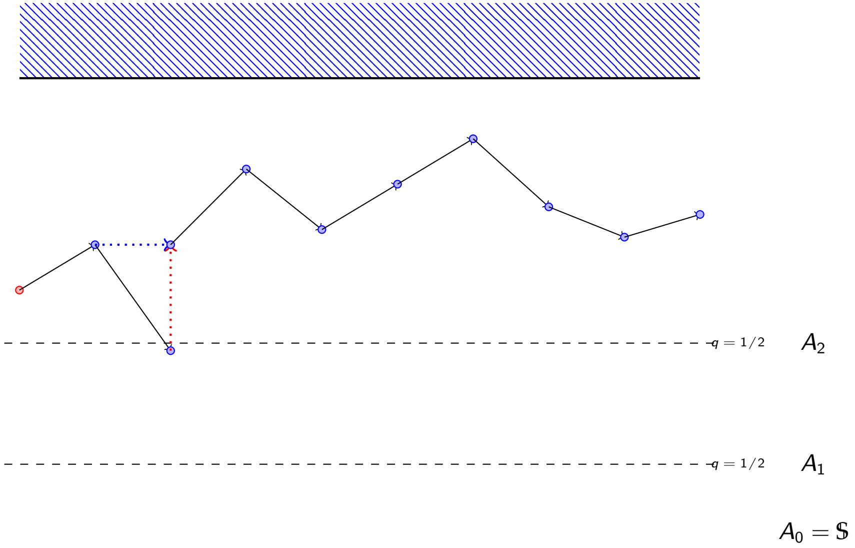
Adaptive Method



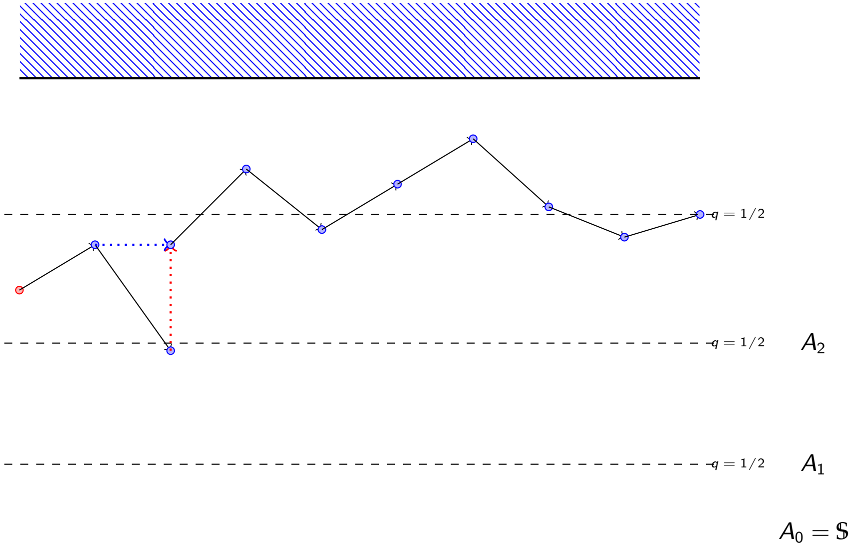
Adaptive Method



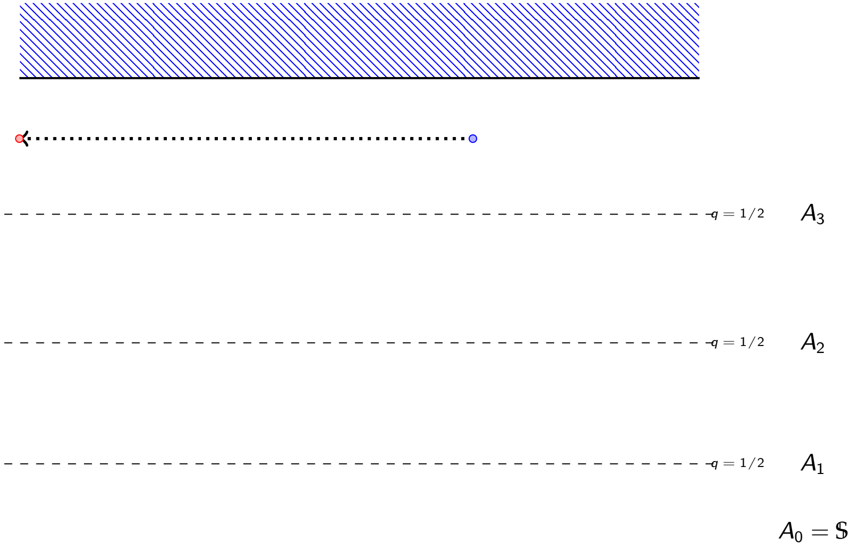
Adaptive Method



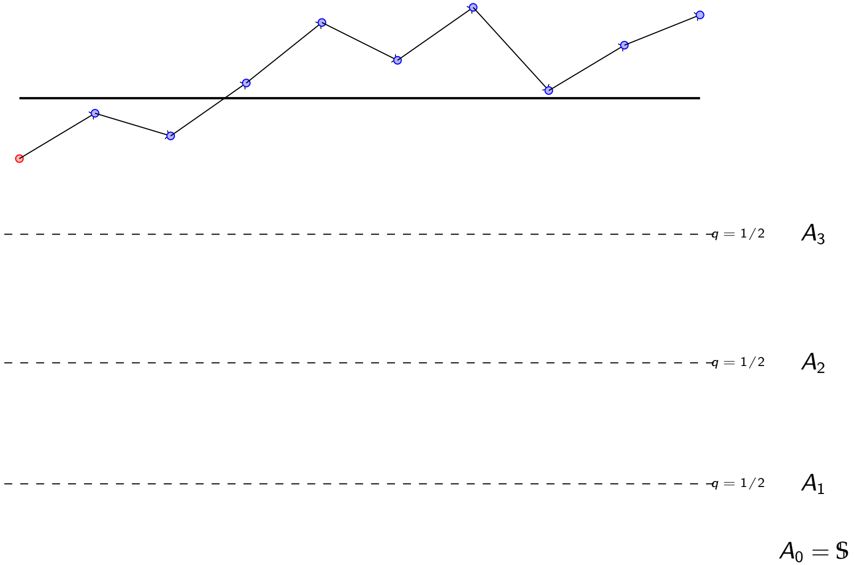
Adaptive Method



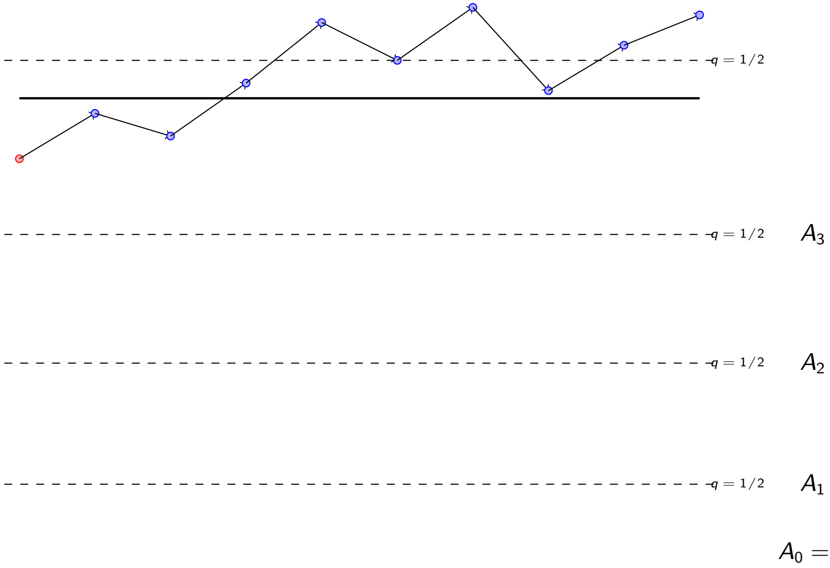
Adaptive Method



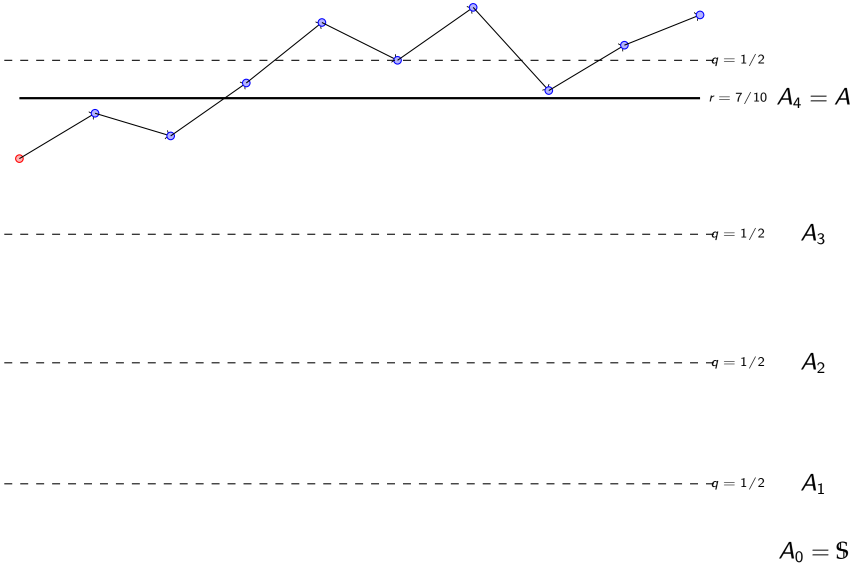
Adaptive Method



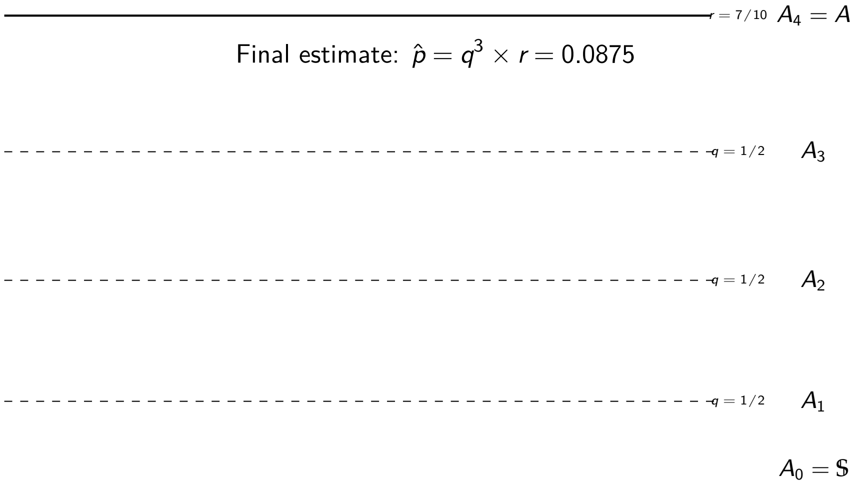
Adaptive Method



Adaptive Method



Adaptive Method



Convergence of Adaptive POP Method

- For any $l \in \{1, \ldots, L^* + 1\}$ and any $\varepsilon > 0$, we have

$$\mathbb{P}\left(|\hat{Q}^l_{N,q} - Q^l_q| > \varepsilon\right) < b(q, N, \varepsilon)$$

such that locally uniformly $\sum_N b(q, N, \varepsilon) < \infty$ which allows to show

$$L_N \to L^* \text{ a.s.} \qquad \text{as } N \to \infty$$

Convergence of Adaptive POP Method

☑ For any $l \in \{1, \dots, L^* + 1\}$ and any $\varepsilon > 0$, we have

$$\mathbb{P}\left(|\hat{Q}^l_{N,q} - Q^l_q| > \varepsilon\right) < b(q, N, \varepsilon)$$

such that locally uniformly $\sum_N b(q, N, \varepsilon) < \infty$ which allows to show

$$L_N \to L^* \text{ a.s.} \qquad \text{as } N \to \infty$$

☑ For any $l \in \{L^* - 1, L^*\}$ and any $\varepsilon > 0$, we have

$$\mathbb{P}\left(|\hat{r}_N(Q^l_q) - r(Q^l_q)| > \varepsilon\right) < c(q, N, \varepsilon)$$

such that locally uniformly $\sum_N c(q, N, \varepsilon) < \infty$ which allows to show

$$\hat{r}_N(\hat{Q}^{L_N}_{N,q}) \to r(Q^{L^*}_q) \text{ a.s.} \qquad \text{as } N \to \infty$$

Sensitivity Analysis I

- Need to be careful in the tails. For instance, if $G_\sigma \stackrel{\mathrm{d}}{=} \mathcal{N}(0, \sigma^2)$ then

$$\lim_{x \to +\infty} \frac{\mathbb{P}\left(G_\sigma \geqslant x\right)}{\mathbb{P}\left(G_{\sigma'} \geqslant x\right)} = \begin{cases} 0 & \text{if } 0 < \sigma < \sigma' \\ +\infty & \text{if } \sigma > \sigma' > 0 \end{cases},$$

Sensitivity Analysis I

- Need to be careful in the tails. For instance, if $G_\sigma \overset{\mathrm{d}}{=} \mathcal{N}(0, \sigma^2)$ then

$$\lim_{x \to +\infty} \frac{\mathbb{P}\left(G_\sigma \geqslant x\right)}{\mathbb{P}\left(G_{\sigma'} \geqslant x\right)} = \begin{cases} 0 & \text{if } 0 < \sigma < \sigma' \\ +\infty & \text{if } \sigma > \sigma' > 0 \end{cases},$$

- Embed the computation of rare event statistic in the isonormal Gaussian process framework

Sensitivity Analysis I

- Need to be careful in the tails. For instance, if $G_\sigma \overset{d}{=} \mathcal{N}(0, \sigma^2)$ then

$$\lim_{x \to +\infty} \frac{\mathbb{P}(G_\sigma \geqslant x)}{\mathbb{P}(G_{\sigma'} \geqslant x)} = \begin{cases} 0 & \text{if } 0 < \sigma < \sigma' \\ +\infty & \text{if } \sigma > \sigma' > 0 \end{cases},$$

- Embed the computation of rare event statistic in the isonormal Gaussian process framework
- Assume **(IBP)**: $\partial_\theta \mathbb{E}\left[\Phi^\theta \mathbf{1}_{X^\theta \in A}\right] = \mathbb{E}\left[\mathcal{I}(X^\theta, \Phi^\theta)\mathbf{1}_{X^\theta \in A}\right]$

Sensitivity Analysis I

- Need to be careful in the tails. For instance, if $G_\sigma \overset{d}{=} \mathcal{N}(0, \sigma^2)$ then

$$\lim_{x\to+\infty} \frac{\mathbb{P}\left(G_\sigma \geqslant x\right)}{\mathbb{P}\left(G_{\sigma'} \geqslant x\right)} = \begin{cases} 0 & \text{if } 0 < \sigma < \sigma' \\ +\infty & \text{if } \sigma > \sigma' > 0 \end{cases},$$

- Embed the computation of rare event statistic in the isonormal Gaussian process framework
- Assume **(IBP)**: $\partial_\theta \mathbb{E}\left[\Phi^\theta \mathbf{1}_{X^\theta \in A}\right] = \mathbb{E}\left[\mathcal{J}(X^\theta, \Phi^\theta)\mathbf{1}_{X^\theta \in A}\right]$
- Makes more sense to calculate relative sensitivity in rare event settings

$$\frac{\partial_\theta \mathbb{E}\left[\Phi^\theta \mathbf{1}_{X^\theta \in A}\right]}{\mathbb{E}\left[\Phi^\theta \mathbf{1}_{X^\theta \in A}\right]} = \frac{\mathbb{E}\left[\mathcal{J}(X^\theta, \Phi^\theta)|X^\theta \in A\right]}{\mathbb{E}\left[\Phi^\theta|X^\theta \in A\right]}$$

Sensitivity Analysis I

- Need to be careful in the tails. For instance, if $G_\sigma \stackrel{d}{=} \mathcal{N}(0, \sigma^2)$ then

$$\lim_{x \to +\infty} \frac{\mathbb{P}\left(G_\sigma \geqslant x\right)}{\mathbb{P}\left(G_{\sigma'} \geqslant x\right)} = \begin{cases} 0 & \text{if } 0 < \sigma < \sigma' \\ +\infty & \text{if } \sigma > \sigma' > 0 \end{cases},$$

- Embed the computation of rare event statistic in the isonormal Gaussian process framework
- Assume **(IBP)**: $\partial_\theta \mathbb{E}\left[\Phi^\theta \mathbf{1}_{X^\theta \in A}\right] = \mathbb{E}\left[\mathcal{J}(X^\theta, \Phi^\theta)\mathbf{1}_{X^\theta \in A}\right]$
- Makes more sense to calculate relative sensitivity in rare event settings

$$\frac{\partial_\theta \mathbb{E}\left[\Phi^\theta \mathbf{1}_{X^\theta \in A}\right]}{\mathbb{E}\left[\Phi^\theta \mathbf{1}_{X^\theta \in A}\right]} = \frac{\mathbb{E}\left[\mathcal{J}(X^\theta, \Phi^\theta) | X^\theta \in A\right]}{\mathbb{E}\left[\Phi^\theta | X^\theta \in A\right]}$$

- Only need to create Markov chain at the last level

Sensitivity Analysis II

Sketch of proof for Integration by Parts (IBP) result in $\mathbb{R}^d$:

- ☑ Procedure inspired by the work of Fournié et al. '99 which helps to provide formula for $\mathcal{I}(X^\theta, \Phi^\theta)$

Sensitivity Analysis II

Sketch of proof for Integration by Parts (IBP) result in $\mathbb{R}^d$:

- ☑ Procedure inspired by the work of Fournié et al. '99 which helps to provide formula for $\mathcal{I}(X^\theta, \Phi^\theta)$

- ☑ Careful analysis to properly mollify $\mathbf{1}_{X^\theta \in A}$

Sensitivity Analysis II

Sketch of proof for Integration by Parts (IBP) result in $\mathbb{R}^d$:

- ☑ Procedure inspired by the work of Fournié et al. '99 which helps to provide formula for $\mathcal{I}(X^\theta, \Phi^\theta)$
- ☑ Careful analysis to properly mollify $\mathbf{1}_{X^\theta \in A}$
- ☑ Choose a finite measure $\bar{\mu}(\mathrm{d}x) = (1 + |x|)^{-q}\mathrm{d}z$ on $\mathbb{R}^d$ with $q > d$
- ☑ $\mathbf{1}_{X^\theta \in A} \in \mathrm{L}^4(\bar{\mu})$ which gives a limiting sequence of smooth functions with compact support $(\xi_k)_{k\geqslant 1}$

$$u(\theta) := \mathbb{E}[\Phi^\theta \mathbf{1}_{X^\theta \in A}], \qquad u_k(\theta) := \mathbb{E}[\Phi^\theta \xi_k(X^\theta)],$$

$$v(\theta) := \mathbb{E}[\mathcal{I}(X^\theta, \Phi^\theta)\mathbf{1}_{X^\theta \in A}], \qquad v_k(\theta) := \mathbb{E}[\mathcal{I}(X^\theta, \Phi^\theta)\xi_k(X^\theta)],$$

Sensitivity Analysis II

Sketch of proof for Integration by Parts (IBP) result in $\mathbb{R}^d$:

- ☑ Procedure inspired by the work of Fournié et al. '99 which helps to provide formula for $\mathcal{I}(X^\theta, \Phi^\theta)$
- ☑ Careful analysis to properly mollify $\mathbf{1}_{X^\theta \in A}$
- ☑ Choose a finite measure $\bar{\mu}(\mathrm{d}x) = (1+|x|)^{-q}\mathrm{d}z$ on $\mathbb{R}^d$ with $q > d$
- ☑ $\mathbf{1}_{X^\theta \in A} \in \mathrm{L}^4(\bar{\mu})$ which gives a limiting sequence of smooth functions with compact support $(\xi_k)_{k\geqslant 1}$

$$u(\theta) := \mathbb{E}[\Phi^\theta \mathbf{1}_{X^\theta \in A}], \qquad u_k(\theta) := \mathbb{E}[\Phi^\theta \xi_k(X^\theta)],$$
$$v(\theta) := \mathbb{E}[\mathcal{I}(X^\theta, \Phi^\theta)\mathbf{1}_{X^\theta \in A}], \qquad v_k(\theta) := \mathbb{E}[\mathcal{I}(X^\theta, \Phi^\theta)\xi_k(X^\theta)],$$

- ☑ Conclude by showing

$$\text{(i) } u_k'(\theta) = v_k(\theta) \quad \text{(ii) } u_k(\theta) \underset{k\to\infty}{\to} u(\theta) \quad \text{(iii) } v_k(\theta) \underset{k\to\infty}{\to} v(\theta)$$

Measuring Default Probabilities in Credit Portfolios

☑ Dynamics of asset values:

$$dS_i(t) = rS_i(t)dt + \sigma(t)S_i(t)dW_i(t), \quad i = 1, \ldots, N_0,$$
$$d\sigma(t) = \kappa\big(\bar{\sigma} - \sigma(t)\big)dt + \gamma\sqrt{\sigma(t)}dW_t,$$
$$d\langle W_i, W_j\rangle = \rho_0 dt, i \neq j, \quad d\langle W_i, W\rangle = \rho_\sigma dt$$

Measuring Default Probabilities in Credit Portfolios

- Dynamics of asset values:

$$dS_i(t) = rS_i(t)dt + \sigma(t)S_i(t)dW_i(t), \quad i = 1, \ldots, N_0,$$
$$d\sigma(t) = \kappa\big(\bar{\sigma} - \sigma(t)\big)dt + \gamma\sqrt{\sigma(t)}dW_t,$$
$$d\langle W_i, W_j\rangle = \rho_0 dt, i \neq j, \quad d\langle W_i, W\rangle = \rho_\sigma dt$$

- Compute $\mathbb{P}\big(\sum_{i=1}^{N_0} \mathbf{1}_{\{\inf\{t:S_i(t)\leqslant B\}\leqslant T\}} > L\big), \quad 0 < L < N_0$

Measuring Default Probabilities in Credit Portfolios

- Dynamics of asset values:

$$dS_i(t) = rS_i(t)dt + \sigma(t)S_i(t)dW_i(t), \quad i = 1, \dots, N_0,$$
$$d\sigma(t) = \kappa\big(\bar{\sigma} - \sigma(t)\big)dt + \gamma\sqrt{\sigma(t)}dW_t,$$
$$d\langle W_i, W_j\rangle = \rho_0 dt, i \neq j, \quad d\langle W_i, W\rangle = \rho_\sigma dt$$

- Compute $\mathbb{P}\Big(\sum_{i=1}^{N_0} \mathbf{1}_{\{\inf\left\{t:S_i(t)\leqslant B\right\}\leqslant T\}} > L\Big), \quad 0 < L < N_0$

- $A_k := \left\{\sum_{i=1}^{N_0} \mathbf{1}_{\inf S_i(t)\leqslant S_i(0)-\frac{k}{n}(S_i(0)-B)} > L\right\}, \left\{\sum_{i=1}^{N_0} \mathbf{1}_{\inf S_i(t)\leqslant B} > \frac{k}{n}L\right\}$

Measuring Default Probabilities in Credit Portfolios

- Dynamics of asset values:

$$dS_i(t) = rS_i(t)dt + \sigma(t)S_i(t)dW_i(t), \quad i = 1, \ldots, N_0,$$
$$d\sigma(t) = \kappa\big(\bar{\sigma} - \sigma(t)\big)dt + \gamma\sqrt{\sigma(t)}dW_t,$$
$$d\langle W_i, W_j\rangle = \rho_0 dt, i \neq j, \quad d\langle W_i, W\rangle = \rho_\sigma dt$$

- Compute $\mathbb{P}\big(\sum_{i=1}^{N_0} \mathbf{1}_{\{\inf\{t:S_i(t)\leqslant B\}\leqslant T\}} > L\big), \quad 0 < L < N_0$
- $A_k := \left\{\sum_{i=1}^{N_0} \mathbf{1}_{\inf S_i(t)\leqslant S_i(0)-\frac{k}{n}(S_i(0)-B)} > L\right\}, \left\{\sum_{i=1}^{N_0} \mathbf{1}_{\inf S_i(t)\leqslant B} > \frac{k}{n}L\right\}$
- MC estimator with 3×10^9 sample paths: $[\mathbf{4.92}, \mathbf{5.13}] \times \mathbf{10^{-6}}$

Measuring Default Probabilities in Credit Portfolios

- Dynamics of asset values:

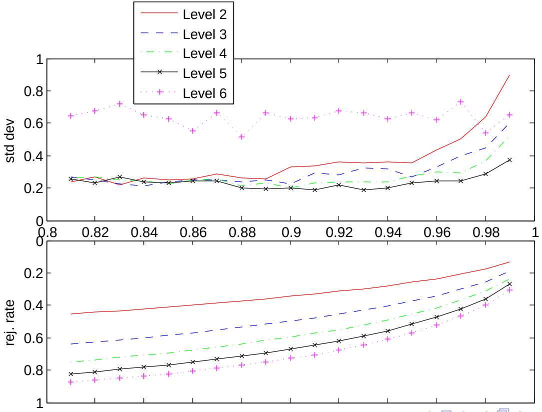
$$dS_i(t) = rS_i(t)dt + \sigma(t)S_i(t)dW_i(t), \quad i = 1, \ldots, N_0,$$
$$d\sigma(t) = \kappa\big(\bar{\sigma} - \sigma(t)\big)dt + \gamma\sqrt{\sigma(t)}dW_t,$$
$$d\langle W_i, W_j\rangle = \rho_0 dt,\, i \neq j, \quad d\langle W_i, W\rangle = \rho_\sigma dt$$

- Compute $\mathbb{P}\big(\sum_{i=1}^{N_0} \mathbf{1}_{\{\inf\{t: S_i(t) \leqslant B\} \leqslant T\}} > L\big), \quad 0 < L < N_0$
- $A_k := \left\{\sum_{i=1}^{N_0} \mathbf{1}_{\inf S_i(t) \leqslant S_i(0) - \frac{k}{n}(S_i(0) - B)} > L\right\}, \left\{\sum_{i=1}^{N_0} \mathbf{1}_{\inf S_i(t) \leqslant B} > \frac{k}{n} L\right\}$
- MC estimator with 3×10^9 sample paths: $[\mathbf{4.92}, \mathbf{5.13}] \times \mathbf{10^{-6}}$
- POP and IPS with $M = N = 10^4$

	IPS			POP		
	mean ($\times 10^{-6}$)	std. ($\times 10^{-6}$)	std./mean	mean ($\times 10^{-6}$)	std. ($\times 10^{-6}$)	std./mean
$\rho = 0.9$	5.82	4.37	0.75	5.01	0.80	0.16
$\rho = 0.7$	4.92	1.56	0.32	4.99	1.02	0.20
$\rho = 0.5$	4.79	3.80	0.79	5.02	1.94	0.39

Adaptive POP & Other Aspects

	Adaptive IPS			Adaptive POP		
	mean ($\times 10^{-6}$)	std. ($\times 10^{-6}$)	std./mean	mean ($\times 10^{-6}$)	std. ($\times 10^{-6}$)	std./mean
$\rho = 0.9$	4.93	1.91	0.39	5.16	0.85	0.16
$\rho = 0.7$	5.42	1.58	0.29	4.98	1.02	0.20
$\rho = 0.5$	6.40	5.00	0.78	5.35	2.05	0.38



Sensitivity Estimation

☑ Consider d-dimensional Black-Scholes model $\frac{dS_t^i}{S_t^i} = \mu^i dt + \sigma^i d(LW_t)^i$

Sensitivity Estimation

- ☑ Consider d-dimensional Black-Scholes model $\frac{dS_t^i}{S_t^i} = \mu^i dt + \sigma^i d(LW_t)^i$
- ☑ $LL^* = C$. Assume C (therefore L) is invertible

Sensitivity Estimation

- ☑ Consider d-dimensional Black-Scholes model $\frac{dS_t^i}{S_t^i} = \mu^i dt + \sigma^i d(LW_t)^i$
- ☑ $LL^* = C$. Assume C (therefore L) is invertible
- ☑ Log price $Z_T^i = Z_0^i + \left(\mu^i - \frac{1}{2}(\sigma^i)^2\right) T + \sigma^i (LW_T)^i$

Sensitivity Estimation

- ☑ Consider d-dimensional Black-Scholes model $\frac{dS_t^i}{S_t^i} = \mu^i dt + \sigma^i d(LW_t)^i$
- ☑ $LL^* = C$. Assume C (therefore L) is invertible
- ☑ Log price $Z_T^i = Z_0^i + \left(\mu^i - \frac{1}{2}(\sigma^i)^2\right) T + \sigma^i (LW_T)^i$
- ☑ Consider basket option-style payoff $\mathcal{P} := \mathbb{E}\left[\max\left(-\varphi(Z_T, \bar{a}), 0\right)\right]$ where

$$\varphi(z, \bar{a}) := \sum_{i=1}^{d} \varepsilon_i p_i e^{z^i} - \bar{a}, \quad p_i > 0, \varepsilon_i \in \{-1, 1\}, \bar{a} \in \mathbb{R}$$

Sensitivity Estimation

- ☑ Consider d-dimensional Black-Scholes model $\frac{dS_t^i}{S_t^i} = \mu^i dt + \sigma^i d(LW_t)^i$
- ☑ $LL^* = C$. Assume C (therefore L) is invertible
- ☑ Log price $Z_T^i = Z_0^i + \left(\mu^i - \frac{1}{2}(\sigma^i)^2\right) T + \sigma^i (LW_T)^i$
- ☑ Consider basket option-style payoff $\mathcal{P} := \mathbb{E}\left[\max\left(-\varphi(Z_T, \bar{a}), 0\right)\right]$ where

$$\varphi(z, \bar{a}) := \sum_{i=1}^{d} \varepsilon_i p_i e^{z^i} - \bar{a}, \quad p_i > 0, \varepsilon_i \in \{-1, 1\}, \bar{a} \in \mathbb{R}$$

- ☑ For $d = 2$, take $\sigma^2 = C_{1,2}\sigma^1$ (Motivated by the singular case in **Gulisashvili & Tankov, Bernoulli '16**)

Sensitivity ($d = 2$) w.r.t.	p_1	σ^1	$C_{1,2}$
POP method (10^6) (mean/std)	-0.7155 (0.0046)	24.0078 (0.1760)	3.1058 (0.0253)
Finite difference (10^6) (mean/std)	-0.7120 (0.0157)	23.9252 (0.4838)	3.0866 (0.1128)
Finite difference (10^9) (99% conf. interval)	$(-0.7155, -0.7129)$	(23.9285, 24.0108)	(3.0801, 3.0990)

Conclusion & Remarks

- ☑ Adaptive version of POP method and its convergence
- ☑ Sensitivity estimation for rare event statistics
- ☑ Applicable to wide class of rare event problems due to path space Markov chains
- ☑ Convergence of reversible shaking with rejection transformation in infinite dimension
- ☑ Central Limit Theorem for POP method and adaptive level estimator

Thank you for your attention!

$$\mathcal{I}(Z^\theta, \Phi^\theta) := \dot{\Phi}^\theta + \delta\left(\Phi^\theta \sum_{j=1}^{d} (\gamma_{Z^\theta}^{-1} \dot{Z}^\theta)_j D.Z_j^\theta\right).$$

$$\gamma_{Z^\theta} := (\langle D.Z_i^\theta, D.Z_j^\theta \rangle_{\mathcal{H}})_{1 \leqslant i,j \leqslant d}$$