Estimating satellite collision probabilities via the adaptive splitting technique

From Application to Theory

Rudy PASTEL, ONERA/DCPS-Palaiseau
Supervisor: Jérôme MORIO, ONERA/DCPS-Palaiseau
Scientific director: François Le GLAND, INRIA de Rennes

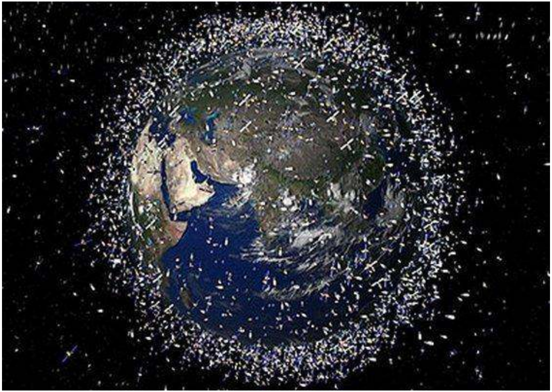
January 2011

Outline

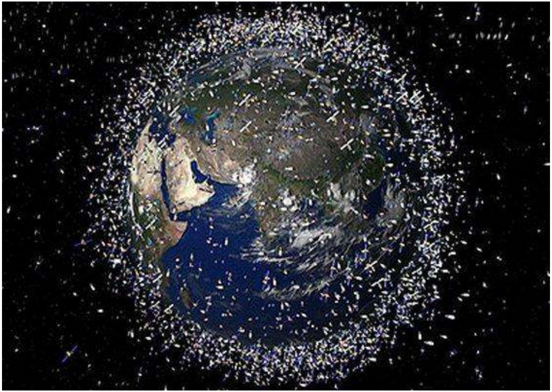
One might think of **an empty outer space**.

One might think of **an empty outer space**.

One would be wrong!



One would be wrong!



Space debris surround Earth and active orbiting objects.

On Tuesday February the 10^{th} 2009, satellites Iridium and Cosmos collided.

On Tuesday February the 10^{th} 2009, satellites Iridium and Cosmos collided.

What was the probability it happened?

Outline

Satellites sizing: Rough estimates

Satellites sizing: Rough estimates

- Real sizes are **confidential**.

Satellites sizing: Rough estimates

- Real sizes are **confidential**.
- For convenience sake, we assume **spherical satellites**.

Satellites sizing: Rough estimates

- Real sizes are **confidential**.
- For convenience sake, we assume **spherical satellites**.
- Radii decided based on **field experience**.

Satellites localisation: Two line element (TLE)

Satellites localisation: Two line element (TLE)

- Provided **by the NORAD and the NASA**.

Satellites localisation: Two line element (TLE)

- Provided **by the NORAD and the NASA**.
- Position and motion at measurement time **without accuracy details**.

Satellites localisation: Two line element (TLE)

- Provided **by the NORAD and the NASA**.
- Position and motion at measurement time **without accuracy details**.
- **Home made** means of the orbit parameters.

Satellites propagation: Keplerian dynamics

Satellites propagation: Keplerian dynamics

- **Fast** to compute.

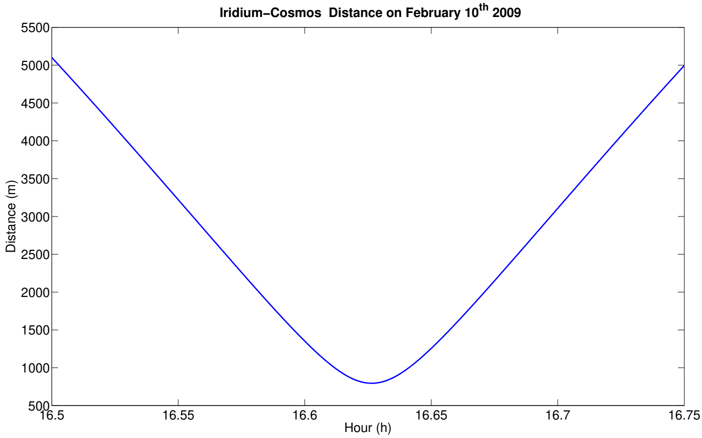
Satellites propagation: Keplerian dynamics

- **Fast** to compute.
- **Easy** to implement.

Satellites propagation: Keplerian dynamics

- **Fast** to compute.
- **Easy** to implement.
- **Less accurate** than the NASA model.

Iridium-Cosmos distance about collision time according to TLE



Randomness at hand: initial conditions

Randomness at hand: initial conditions

- Keplerian dynamics is **deterministic once initial conditions are known**.

Randomness at hand: initial conditions

- Keplerian dynamics is **deterministic once initial conditions are known**.
- **We noised the TLE** to cope with their unknown uncertainty.

Notations and Noise

Notations and Noise

- Initial conditions given by TLEs $(t_{m,i}, \vec{r}_{m,i}, \vec{v}_{m,i})$.

Notations and Noise

- Initial conditions given by TLEs $(t_{m,i}, \vec{r}_{m,i}, \vec{v}_{m,i})$.
- Gaussian additive i.i.d. noise $(\vec{\rho}_{m,i}, \vec{\nu}_{m,i}) \sim \mathcal{N}(0_6, \Sigma^t\Sigma)$.

Notations and Noise

- Initial conditions given by TLEs $(t_{m,i}, \vec{r}_{m,i}, \vec{v}_{m,i})$.
- Gaussian additive i.i.d. noise $(\vec{\rho}_{m,i}, \vec{\nu}_{m,i}) \sim \mathcal{N}(0_6, \Sigma^t\Sigma)$.
- Studied time interval $I = [16h30, 16h45]$.

Notations and Noise

- Initial conditions given by TLEs $(t_{m,i}, \vec{r}_{m,i}, \vec{v}_{m,i})$.
- Gaussian additive i.i.d. noise $(\vec{\rho}_{m,i}, \vec{\nu}_{m,i}) \sim \mathcal{N}(0_6, \Sigma^t\Sigma)$.
- Studied time interval $I = [16h30, 16h45]$.
- Keplerian deterministic dynamics
  $\forall t \in I, \vec{R}_i(t) = \vec{\phi}(t, t_{m,i}, (\vec{r}_{m,i}, \vec{v}_{m,i}) + (\vec{\rho}_{m,i}, \vec{\nu}_{m,i}))$.

Notations and Noise

- Initial conditions given by TLEs $(t_{m,i}, \vec{r}_{m,i}, \vec{v}_{m,i})$.
- Gaussian additive i.i.d. noise $(\vec{\rho}_{m,i}, \vec{\nu}_{m,i}) \sim \mathcal{N}(0_6, \Sigma^t \Sigma)$.
- Studied time interval $I = [16h30, 16h45]$.
- Keplerian deterministic dynamics
  $\forall t \in I, \vec{R}_i(t) = \vec{\phi}(t, t_{m,i}, (\vec{r}_{m,i}, \vec{v}_{m,i}) + (\vec{\rho}_{m,i}, \vec{\nu}_{m,i}))$.
- **Minimum distance during time interval**

$$\Delta\left(\vec{\rho}_{m,1}, \vec{\nu}_{m,1}, \vec{\rho}_{m,2}, \vec{\nu}_{m,2}\right) = \min_{t \in I}\{\|\vec{R}_1(t) - \vec{R}_2(t)\|\}$$

Notations and Noise

- Initial conditions given by TLEs $(t_{m,i}, \vec{r}_{m,i}, \vec{v}_{m,i})$.
- Gaussian additive i.i.d. noise $(\vec{\rho}_{m,i}, \vec{\nu}_{m,i}) \sim \mathcal{N}(0_6, \Sigma^t \Sigma)$.
- Studied time interval $I = [16h30, 16h45]$.
- Keplerian deterministic dynamics
  $\forall t \in I, \vec{R}_i(t) = \vec{\phi}(t, t_{m,i}, (\vec{r}_{m,i}, \vec{v}_{m,i}) + (\vec{\rho}_{m,i}, \vec{\nu}_{m,i}))$.
- **Minimum distance during time interval**

$$\Delta\left(\vec{\rho}_{m,1}, \vec{\nu}_{m,1}, \vec{\rho}_{m,2}, \vec{\nu}_{m,2}\right) = \min_{t \in I}\{\|\vec{R}_1(t) - \vec{R}_2(t)\|\}$$

$$\mathbb{P}[\Delta \leq d_{col}]?$$

Crude Monte Carlo estimation.
$d_{col} = 100$m,100 estimations based on 300000 throws.

Crude Monte Carlo estimation.
d_{col} = 100m,100 estimations based on 300000 throws.

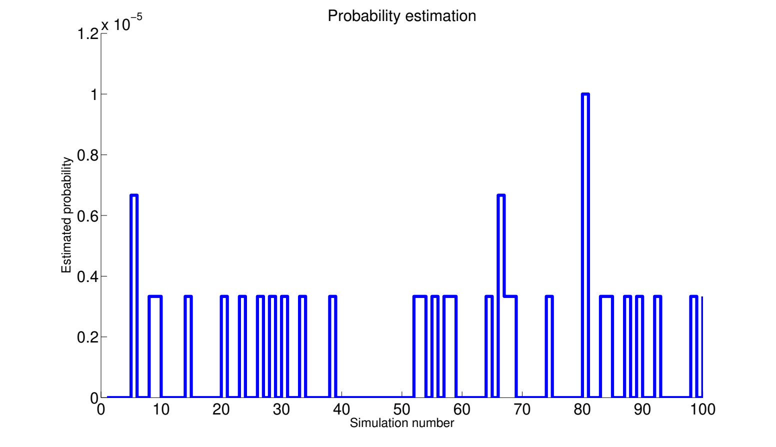


Figure: Probability estimations

Crude Monte Carlo estimation.
$d_{col} = 100$m,100 estimations based on 300000 throws.

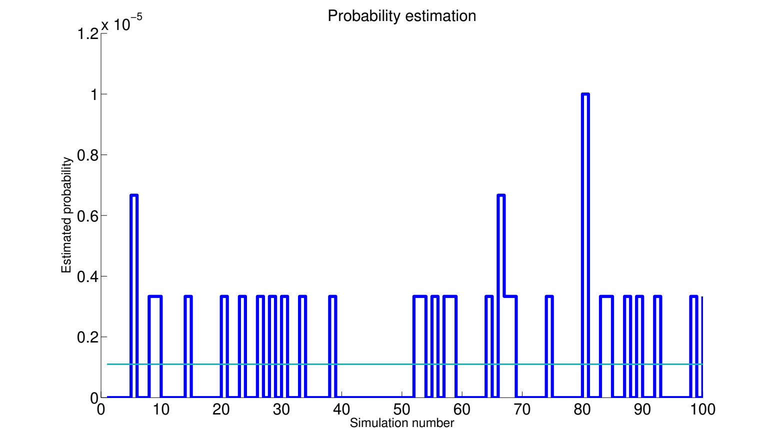


Figure: Probability estimations with their mean.

Crude Monte Carlo estimation. $d_{col} = 100$m,100 estimations based on 300000 throws.

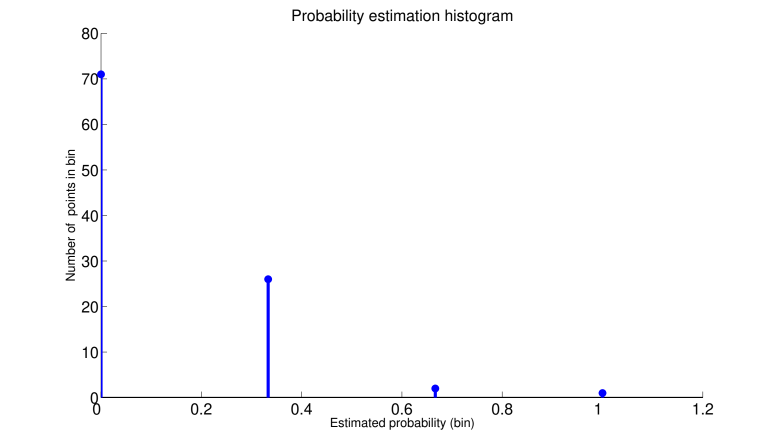
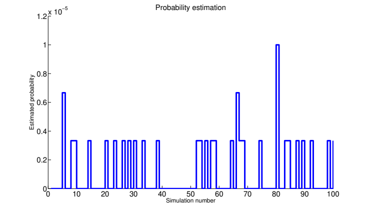
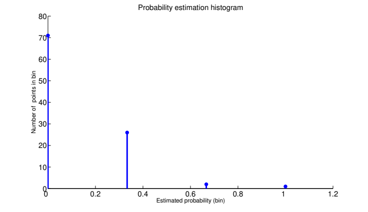


Figure: Probability estimation histogram

Crude Monte Carlo estimation. $d_{col} = 100\text{m}$



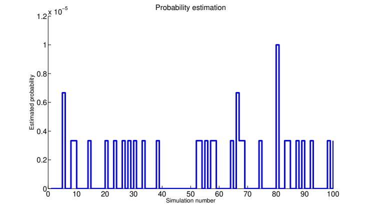
(a) Probability estimations



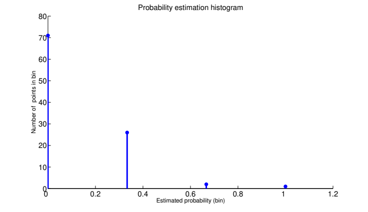
(b) Probability estimation histogram

Figure: 100 estimations based on 300000 exactly simulations.

Crude Monte Carlo estimation. $d_{col} = 100\text{m}$



(a) Probability estimations

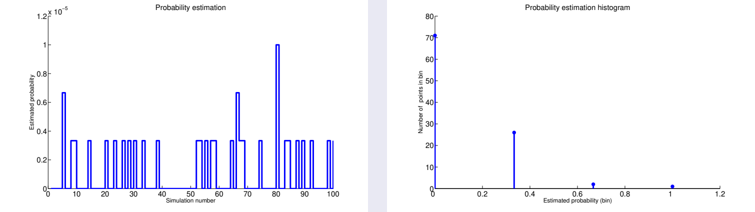


(b) Probability estimation histogram

Figure: 100 estimations based on 300000 exactly simulations.

- Mean estimate: $1.6 \cdot 10^{-6}$.
- Empirical relative deviation: 1.7258.

Crude Monte Carlo estimation. $d_{col} = 100\text{m}$



(a) Probability estimations

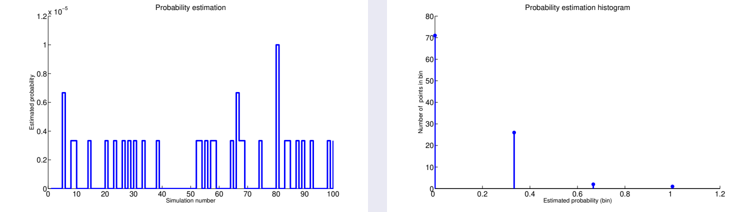
(b) Probability estimation histogram

Figure: 100 estimations based on 300000 exactly simulations.

- Mean estimate: $1.6 \cdot 10^{-6}$.
- Empirical relative deviation: 1.7258.

Unreliable!

Crude Monte Carlo estimation. $d_{col} = 100\text{m}$



(a) Probability estimations (b) Probability estimation histogram

Figure: 100 estimations based on 300000 exactly simulations.

- Mean estimate: $1.6 \cdot 10^{-6}$.
- Empirical relative deviation: 1.7258.

Unreliable!

Need for a well chosen rare event dedicated technique!

Rare event technique requirements

Rare event technique requirements

- **Applicability**: method must be able to operate on a black box mapping and without prior.

Rare event technique requirements

- **Applicability**: method must be able to operate on a black box mapping and without prior.
- **Affordability**: method simulation cost should be as small as possible.

Rare event technique requirements

- **Applicability**: method must be able to operate on a black box mapping and without prior.
- **Affordability**: method simulation cost should be as small as possible.
- **Engineer-friendliness**: method should be as self-tuning as possible.

Outline

A typical industrial probability estimation problem

A typical industrial probability estimation problem

- $h : \mathbb{X} \rightarrow \mathbb{R}^d$ is the transfer function and can be a black box.

A typical industrial probability estimation problem

- $h : \mathbb{X} \longrightarrow \mathbb{R}^d$ is the transfer function and can be a black box.
- $X \sim f_X$ is the input random variable on $\mathbb{X}$.

A typical industrial probability estimation problem

- $h : \mathbb{X} \longrightarrow \mathbb{R}^d$ is the transfer function and can be a black box.
- $X \sim f_X$ is the input random variable on $\mathbb{X}$.
- $g : \mathbb{R}^d \longrightarrow \mathbb{R}$ is the criterion function.

A typical industrial probability estimation problem

- $h : \mathbb{X} \longrightarrow \mathbb{R}^d$ is the transfer function and can be a black box.
- $X \sim f_X$ is the input random variable on $\mathbb{X}$.
- $g : \mathbb{R}^d \longrightarrow \mathbb{R}$ is the criterion function.
- $\xi \equiv s(X) = g \circ h(X)$ is the observed random score.

A typical industrial probability estimation problem

- $h : \mathbb{X} \to \mathbb{R}^d$ is the transfer function and can be a black box.
- $X \sim f_X$ is the input random variable on $\mathbb{X}$.
- $g : \mathbb{R}^d \to \mathbb{R}$ is the criterion function.
- $\xi \equiv s(X) = g \circ h(X)$ is the observed random score.
- $\mathcal{A} \subset \mathbb{R}$ is the target score set.

A typical industrial probability estimation problem

- $h : \mathbb{X} \longrightarrow \mathbb{R}^d$ is the transfer function and can be a black box.
- $X \sim f_X$ is the input random variable on $\mathbb{X}$.
- $g : \mathbb{R}^d \longrightarrow \mathbb{R}$ is the criterion function.
- $\xi \equiv s(X) = g \circ h(X)$ is the observed random score.
- $\mathcal{A} \subset \mathbb{R}$ is the target score set.

$$\mathbb{P}[\xi \in \mathcal{A}] = \mathbb{P}[X \in s^{-1}(\mathcal{A})]?$$

A typical industrial probability estimation problem

- $h : \mathbb{X} \longrightarrow \mathbb{R}^d$ is the transfer function and can be a black box.
- $X \sim f_X$ is the input random variable on $\mathbb{X}$.
- $g : \mathbb{R}^d \longrightarrow \mathbb{R}$ is the criterion function.
- $\xi \equiv s(X) = g \circ h(X)$ is the observed random score.
- $\mathcal{A} \subset \mathbb{R}$ is the target score set.

$$\mathbb{P}[\xi \in \mathcal{A}] = \mathbb{P}[X \in s^{-1}(\mathcal{A})]?$$

If $\mathbb{P}[\xi \in \mathcal{A}]$ is *small*, Crude Monte Carlo won't do.

The Splitting Technique in a nutshell

The Splitting Technique in a nutshell

1. Rephrase the sought probability

$$\begin{cases} \mathbb{R} = \mathcal{A}_0 \supset \cdots \supset \mathcal{A}_\kappa = \mathcal{A} \\ \mathbb{P}[\xi \in \mathcal{A}] = \prod_{i=1}^{\kappa} \mathbb{P}[\xi \in \mathcal{A}_i | \xi \in \mathcal{A}_{i-1}] \end{cases}$$

The Splitting Technique in a nutshell

1. Rephrase the sought probability

$$\begin{cases} \mathbb{R} = \mathcal{A}_0 \supset \cdots \supset \mathcal{A}_\kappa = \mathcal{A} \\ \mathbb{P}[\xi \in \mathcal{A}] = \prod_{i=1}^{\kappa} \mathbb{P}[\xi \in \mathcal{A}_i | \xi \in \mathcal{A}_{i-1}] \end{cases}$$

2. Estimate the conditional probabilities iteratively.

The Splitting Technique in a nutshell

1. Rephrase the sought probability

$$\begin{cases} \mathbb{R} = \mathcal{A}_0 \supset \cdots \supset \mathcal{A}_\kappa = \mathcal{A} \\ \mathbb{P}[\xi \in \mathcal{A}] = \prod_{i=1}^{\kappa} \mathbb{P}[\xi \in \mathcal{A}_i | \xi \in \mathcal{A}_{i-1}] \end{cases}$$

2. Estimate the conditional probabilities iteratively.

From now on, suppose $\mathcal{A} = [T, \infty[$

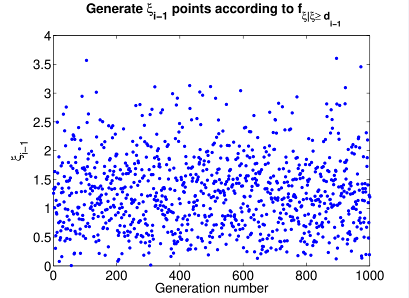
The Adaptive Splitting Technique step by step

The Adaptive Splitting Technique in images

The Adaptive Splitting Technique step by step

1. Generate $\{X_{k-1}^i\}$ points according to $f_{X|s(X)\geq\delta_{k-1}}$.

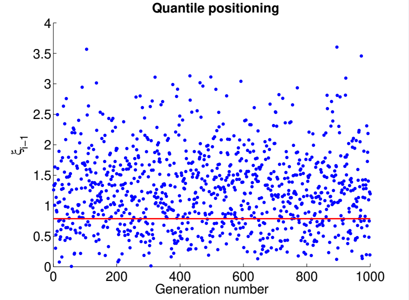
The Adaptive Splitting Technique in images



The Adaptive Splitting Technique step by step

1. Generate $\{X^i_{k-1}\}$ points according to $f_{X|s(X)\geq\delta_{k-1}}$.
2. Define δ_k as $\{s(X^i_{k-1})\}$'s empirical a_k-quantile and set $\mathcal{A}_k = [\delta_k, \infty[$.

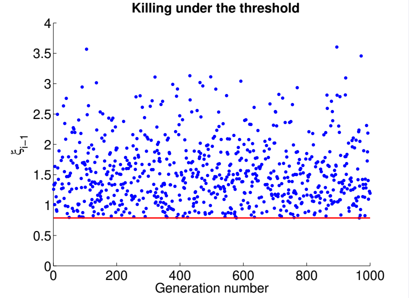
The Adaptive Splitting Technique in images



The Adaptive Splitting Technique step by step

1. Generate $\{X^i_{k-1}\}$ points according to $f_{X|s(X)\geq\delta_{k-1}}$.
2. Define δ_k as $\{s(X^i_{k-1})\}$'s empirical a_k-quantile and set $\mathcal{A}_k = [\delta_k, \infty[$.
3. $\mathbb{P}[s(X) \geq \delta_k | \mathcal{A}_{k-1}] \sim a_k$.

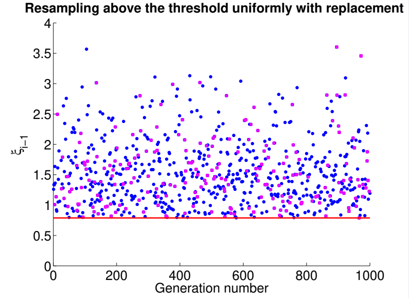
The Adaptive Splitting Technique in images



The Adaptive Splitting Technique step by step

1. Generate $\{X^i_{k-1}\}$ points according to $f_{X|s(X)\geq\delta_{k-1}}$.
2. Define δ_k as $\{s(X^i_{k-1})\}$'s empirical a_k-quantile and set $\mathcal{A}_k = [\delta_k, \infty[$.
3. $\mathbb{P}[s(X) \geq \delta_k | \mathcal{A}_{k-1}] \sim a_k$.
4. Replace points outside $\mathcal{A}_k$ by choosing uniformly with replacement among those in $\mathcal{A}_k$.

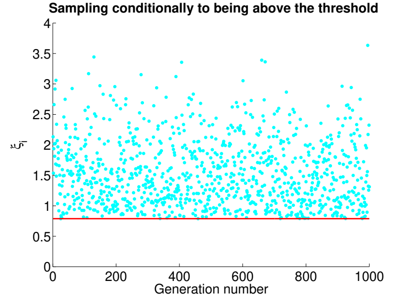
The Adaptive Splitting Technique in images



The Adaptive Splitting Technique step by step

1. Generate $\{X_{k-1}^i\}$ points according to $f_{X|s(X)\geq\delta_{k-1}}$.
2. Define δ_k as $\{s(X_{k-1}^i)\}$'s empirical a_k-quantile and set $\mathcal{A}_k = [\delta_k, \infty[$.
3. $\mathbb{P}[s(X) \geq \delta_k | \mathcal{A}_{k-1}] \sim a_k$.
4. Replace points outside $\mathcal{A}_k$ by choosing uniformly with replacement among those in $\mathcal{A}_k$.
5. Use all the points to sample according to $f_{X|s(X)\geq\delta_k}$.

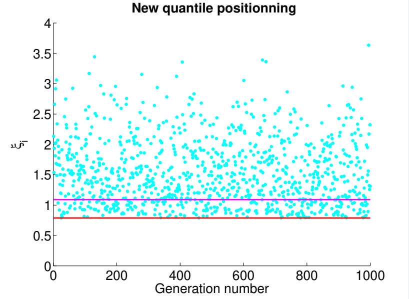
The Adaptive Splitting Technique in images



The Adaptive Splitting Technique step by step

1. Generate $\{X^i_{k-1}\}$ points according to $f_{X|s(X)\geq\delta_{k-1}}$.
2. Define δ_k as $\{s(X^i_{k-1})\}$'s empirical a_k-quantile and set $\mathcal{A}_k = [\delta_k, \infty[$.
3. $\mathbb{P}[s(X) \geq \delta_k | \mathcal{A}_{k-1}] \sim a_k$.
4. Replace points outside $\mathcal{A}_k$ by choosing uniformly with replacement among those in $\mathcal{A}_k$.
5. Use all the points to sample according to $f_{X|s(X)\geq\delta_k}$.
6. Back to step 2.

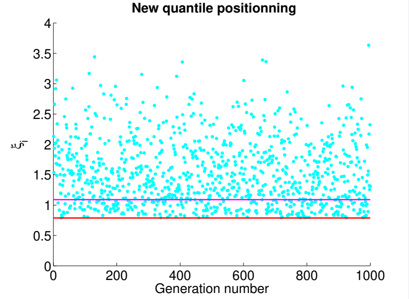
The Adaptive Splitting Technique in images



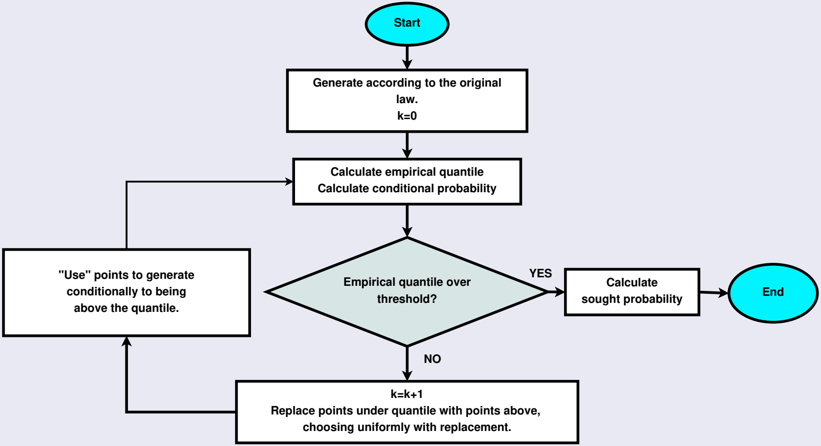
The Adaptive Splitting Technique step by step

1. Generate $\{X^i_{k-1}\}$ points according to $f_{X|s(X)\geq\delta_{k-1}}$.
2. Define δ_k as $\{s(X^i_{k-1})\}$'s empirical a_k-quantile and set $\mathcal{A}_k = [\delta_k, \infty[$.
3. $\mathbb{P}[s(X) \geq \delta_k | \mathcal{A}_{k-1}] \sim a_k$.
4. Replace points outside $\mathcal{A}_k$ by choosing uniformly with replacement among those in $\mathcal{A}_k$.
5. Use all the points to sample according to $f_{X|s(X)\geq\delta_k}$.
6. Back to step 2.

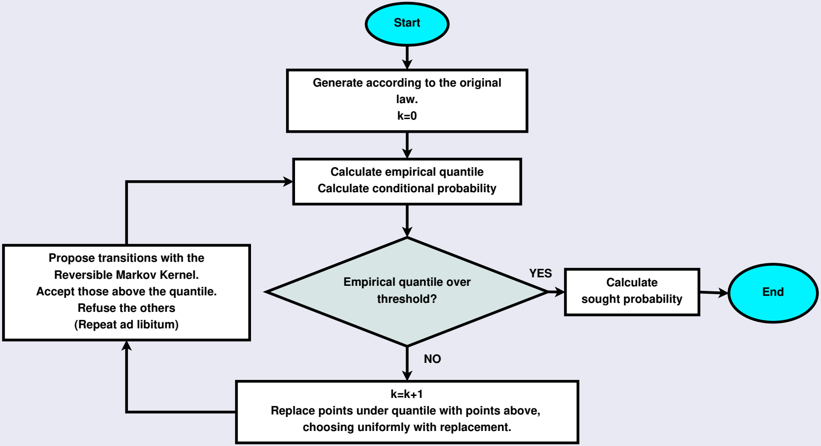
The Adaptive Splitting Technique in images



AST logigram



AST logigram with Markovian resampling



f_X-reversible Markov Kernel

f_X-reversible Markov Kernel

$M(\cdot,\cdot) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ is mapping such that

$$\forall x \in \mathbb{X}, \ \ M(x,\cdot) : \mathbb{X} \to \mathbb{R} \text{ is a density function.}$$

f_X-reversible Markov Kernel

$M(\cdot,\cdot) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ is mapping such that

$$\forall x \in \mathbb{X},\ \ M(x,\cdot) : \mathbb{X} \to \mathbb{R} \text{ is a density function.}$$

M is said to be a f_X-reversible Markov Kernel if

$$\forall (x,y) \in \mathbb{X} \times \mathbb{X},\ \ f_X(x)M(x,y) = f_X(y)M(y,x)$$

f_X-reversible Markov Kernel

$M(\cdot,\cdot) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ is mapping such that

$$\forall x \in \mathbb{X},\ \ M(x,\cdot) : \mathbb{X} \to \mathbb{R} \text{ is a density function.}$$

M is said to be a f_X-reversible Markov Kernel if

$$\forall (x,y) \in \mathbb{X} \times \mathbb{X},\ \ f_X(x)M(x,y) = f_X(y)M(y,x)$$

Meaning, if from a f_X set, you use M to generate another,

f_X-reversible Markov Kernel

$M(\cdot,\cdot) : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}$ is mapping such that

$$\forall x \in \mathbb{X},\ M(x,\cdot) : \mathbb{X} \rightarrow \mathbb{R} \text{ is a density function.}$$

M is said to be a f_X-reversible Markov Kernel if

$$\forall (x,y) \in \mathbb{X} \times \mathbb{X},\ f_X(x)M(x,y) = f_X(y)M(y,x)$$

Meaning, if from a f_X set, you use M to generate another,

- the new set is distributed according to f_X as well.

f_X-reversible Markov Kernel

$M(\cdot,\cdot) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ is mapping such that

$$\forall x \in \mathbb{X},\ \ M(x,\cdot) : \mathbb{X} \to \mathbb{R} \text{ is a density function.}$$

M is said to be a f_X-reversible Markov Kernel if

$$\forall (x,y) \in \mathbb{X} \times \mathbb{X},\ \ f_X(x)M(x,y) = f_X(y)M(y,x)$$

Meaning, if from a f_X set, you use M to generate another,

- the new set is distributed according to f_X as well.
- statistically, no one can say which set generated the other.

Sampling according to $f_{X|s(X)\geq\delta_k}$ using a f_X-reversible M.

Sampling according to $f_{X|s(X)\geq\delta_k}$ using a f_X-reversible M.

Given $X_k \sim f_{X|s(X)\geq\delta_k}$, set

Sampling according to $f_{X|s(X)\geq\delta_k}$ using a f_X-reversible M.

Given $X_k \sim f_{X|s(X)\geq\delta_k}$, set

1. Z according to $M(X_k, \cdot)$.

Sampling according to $f_{X|s(X)\geq\delta_k}$ using a f_X-reversible M.

Given $X_k \sim f_{X|s(X)\geq\delta_k}$, set

1. Z according to $M(X_k,\cdot)$.
2. $\tilde{X}_k = \begin{cases} Z & \text{if } s(Z) \geq \delta_k \\ X_k & \text{otherwise} \end{cases}$

Sampling according to $f_{X|s(X)\geq\delta_k}$ using a f_X-reversible M.

Given $X_k \sim f_{X|s(X)\geq\delta_k}$, set

1. Z according to $M(X_k,\cdot)$.
2. $\tilde{X}_k = \begin{cases} Z & \text{if } s(Z)\geq\delta_k \\ X_k & \text{otherwise} \end{cases}$

i.e. given X_k, $\tilde{X}_k \sim M_k(X_k, dy)$ where

$$\begin{aligned} M_k(X_k,dy) &= \mathbf{1}_{\mathcal{A}_k}(s(y))M(X_k,dy) + M(X_k,\mathcal{A}_k^c)\delta_{X_k}(dy) \\ \mathcal{A}_k &= [\delta_k,+\infty[\end{aligned}$$

Sampling according to $f_{X|s(X)\geq\delta_k}$ using a f_X-reversible M.

Given $X_k \sim f_{X|s(X)\geq\delta_k}$, set

1. Z according to $M(X_k, \cdot)$.
2. $\tilde{X}_k = \begin{cases} Z & \text{if } s(Z) \geq \delta_k \\ X_k & \text{otherwise} \end{cases}$

i.e. given X_k, $\tilde{X}_k \sim M_k(X_k, dy)$ where

$$\begin{aligned} M_k(X_k, dy) &= \mathbf{1}_{\mathcal{A}_k}(s(y))M(X_k, dy) + M(X_k, \mathcal{A}_k^c)\delta_{X_k}(dy) \\ \mathcal{A}_k &= [\delta_k, +\infty[\end{aligned}$$

$\tilde{X}_k$ is distributed according to $f_{X|s(X)\geq\delta_k}$ aswell!

Sampling according to $f_{X|s(X)\geq\delta_k}$ using a f_X-reversible M.

Given $X_k \sim f_{X|s(X)\geq\delta_k}$, set

1. Z according to $M(X_k, \cdot)$.
2. $\tilde{X}_k = \begin{cases} Z & \text{if } s(Z) \geq \delta_k \\ X_k & \text{otherwise} \end{cases}$

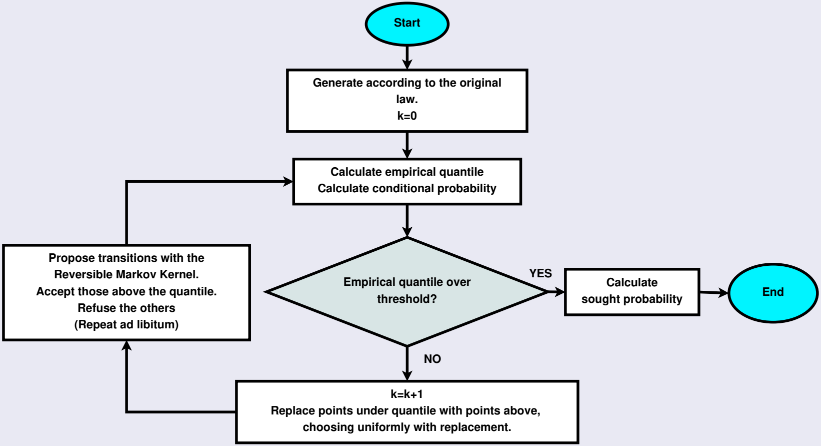
i.e. given X_k, $\tilde{X}_k \sim M_k(X_k, dy)$ where

$$\begin{aligned} M_k(X_k, dy) &= \mathbf{1}_{\mathcal{A}_k}(s(y))M(X_k, dy) + M(X_k, \mathcal{A}_k^c)\delta_{X_k}(dy) \\ \mathcal{A}_k &= [\delta_k, +\infty[\end{aligned}$$

$\tilde{X}_k$ is distributed according to $f_{X|s(X)\geq\delta_k}$ aswell!

$\Rightarrow$ One can use M to inflate a $f_{X|s(X)\geq\delta_k}$ set!

AST logigram with Markovian resampling



A global understanding of the AST with markovian resampling!

Notations and Noise

Notations and Noise

- Initial conditions given by TLEs $(t_{m,i}, \vec{r}_{m,i}, \vec{v}_{m,i})$.
- Gaussian additive i.i.d. noise $(\vec{\rho}_{m,i}, \vec{\nu}_{m,i}) \sim \mathcal{N}(0_6, \Sigma^t \Sigma)$.
- Studied time interval $I = [16h30, 16h45]$.
- Keplerian deterministic dynamics
  $\forall t \in I, \vec{R}_i(t) = \vec{\phi}(t, t_{m,i}, (\vec{r}_{m,i}, \vec{v}_{m,i}) + (\vec{\rho}_{m,i}, \vec{\nu}_{m,i}))$.
- **Minimum distance during time interval**

$$\Delta\left(\vec{\rho}_{m,1}, \vec{\nu}_{m,1}, \vec{\rho}_{m,2}, \vec{\nu}_{m,2}\right) = \min_{t \in I}\{\|\vec{R}_1(t) - \vec{R}_2(t)\|\}$$

Notations and Noise II

Notations and Noise II

- $X = (\vec{\rho}_{m,1}, \vec{\nu}_{m,1}, \vec{\rho}_{m,2}, \vec{\nu}_{m,2})$ is the Gaussian input.

Notations and Noise II

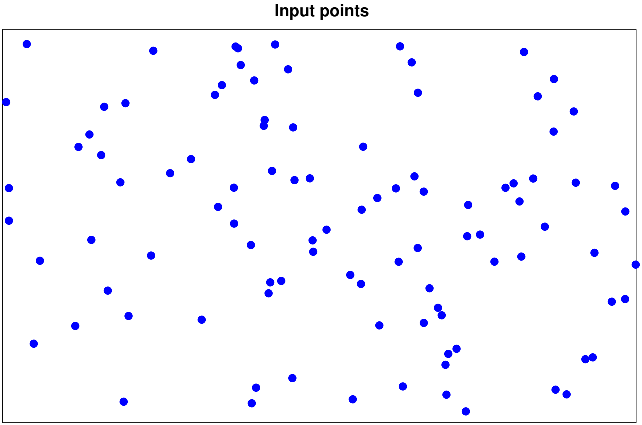
- $X = (\vec{\rho}_{m,1}, \vec{\nu}_{m,1}, \vec{\rho}_{m,2}, \vec{\nu}_{m,2})$ is the Gaussian input.
- $\xi = \Delta(X)$ is the random minimum distance during I.

Notations and Noise II

- $X = (\vec{\rho}_{m,1}, \vec{\nu}_{m,1}, \vec{\rho}_{m,2}, \vec{\nu}_{m,2})$ is the Gaussian input.
- $\xi = \Delta(X)$ is the random minimum distance during I.
- $\mathcal{A} = [0, d_{col}]$ is the target score set.

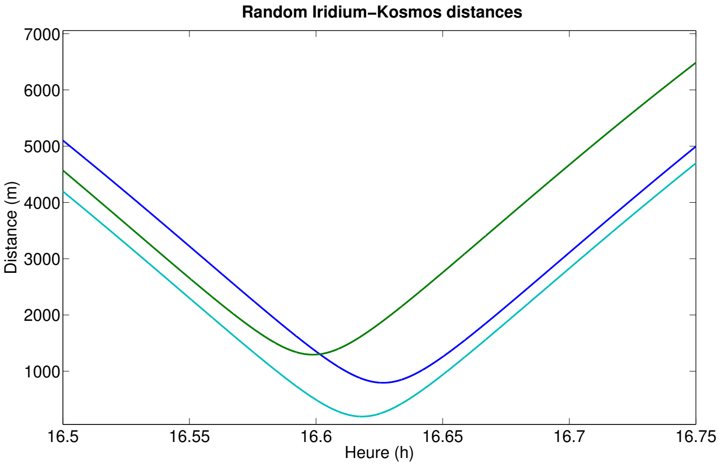
Applied AST in images

Applied AST in images



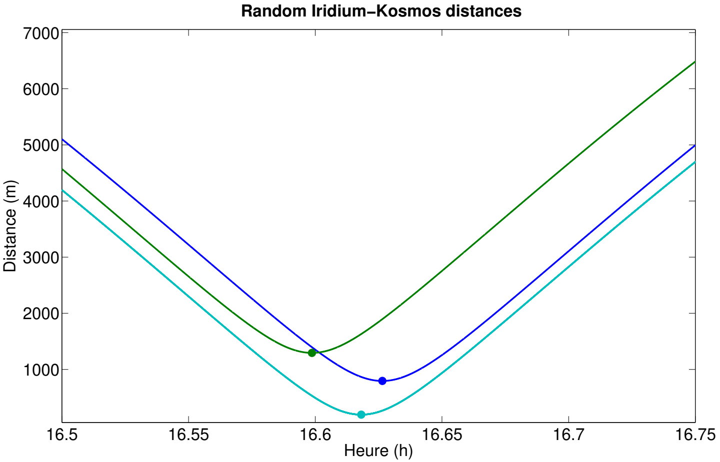
Generate random noises $(X_k^1, \cdots, X_k^n)$.

Applied AST in images



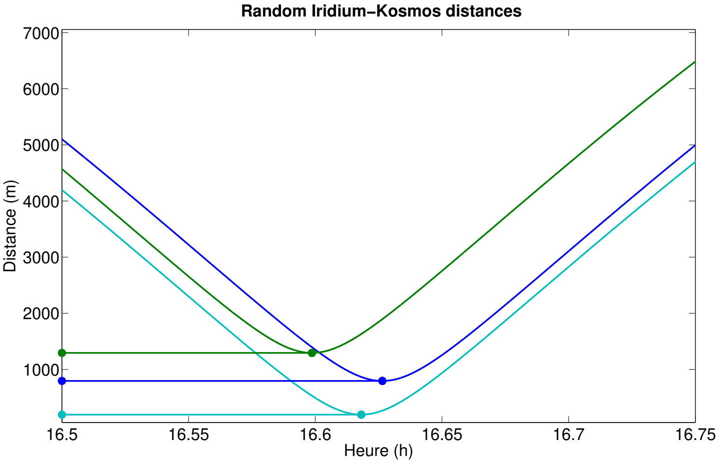
Calculate Iridium-Kosmos distances over time intervalle I.

Applied AST in images



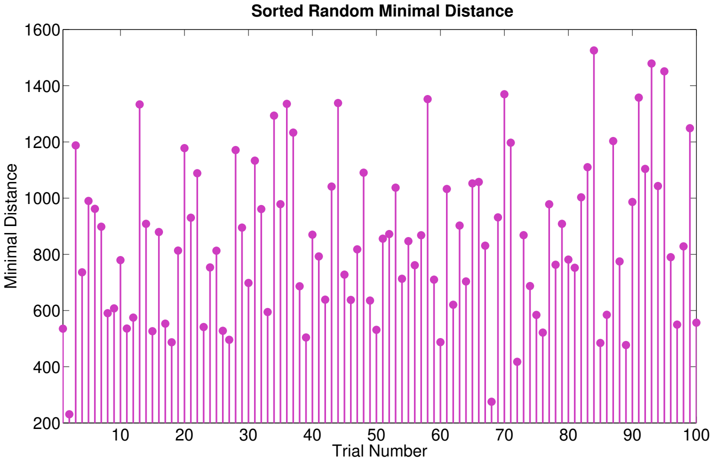
Find minima $(\Delta_k^1, \cdots, \Delta_k^n)$.

Applied AST in images



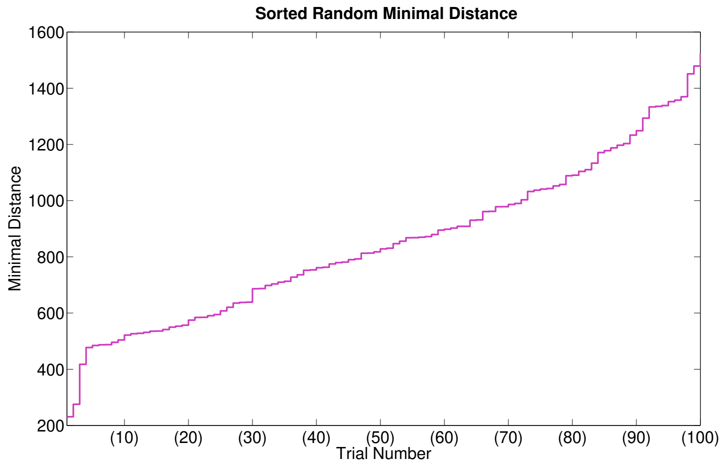
Find minima $\left(\Delta_k^1, \cdots, \Delta_k^n\right)$.

Applied AST in images



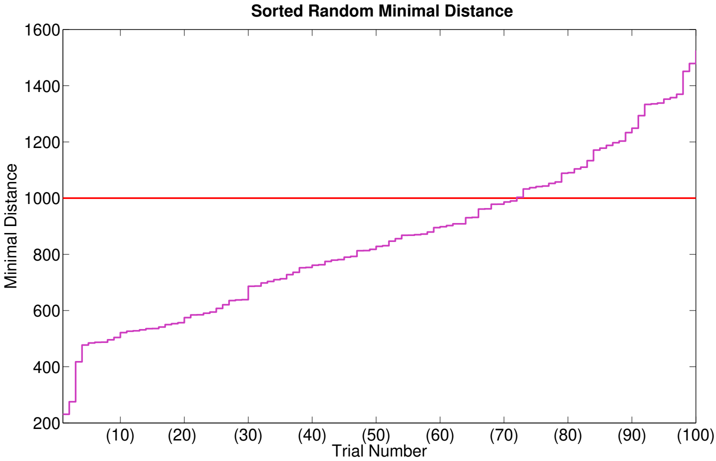
Find minima $\left(\Delta_k^1, \cdots, \Delta_k^n\right)$.

Applied AST in images



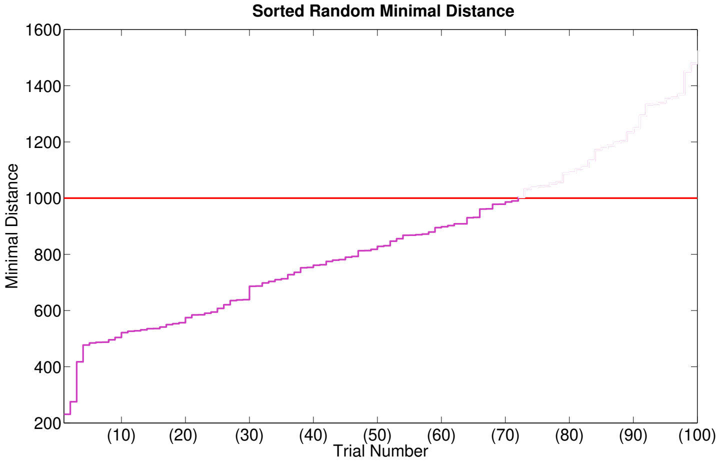
Sort minima to form $\left(\Delta_k^{(1)}, \cdots, \Delta_k^{(n)}\right)$.

Applied AST in images



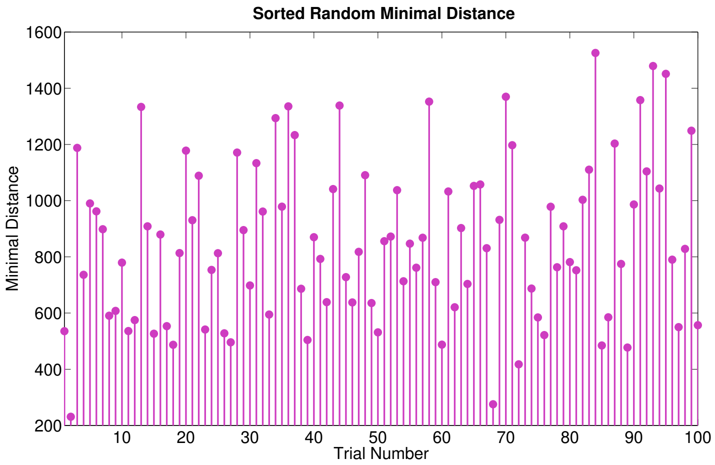
Define quantile δ_{k+1}.

Applied AST in images



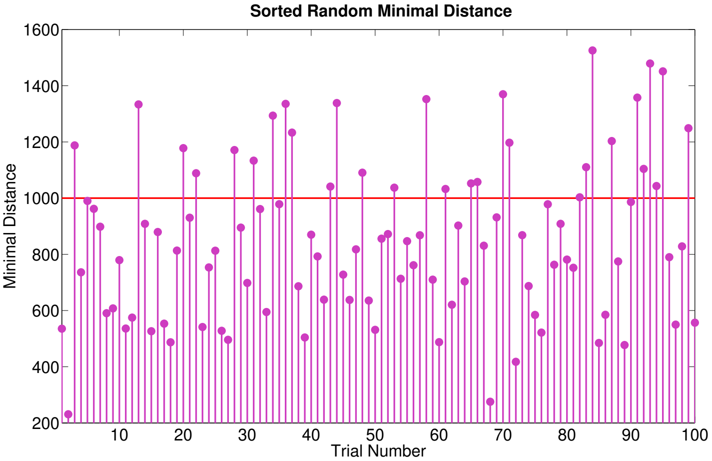
Discard points over quantile.

Applied AST in images



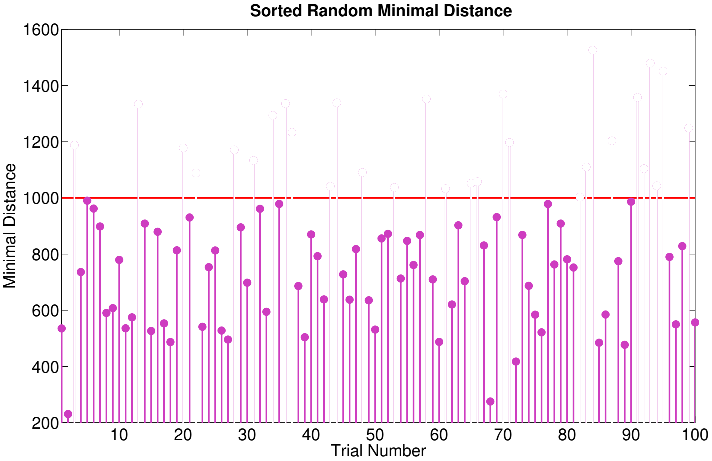
Discard points over quantile.

Applied AST in images



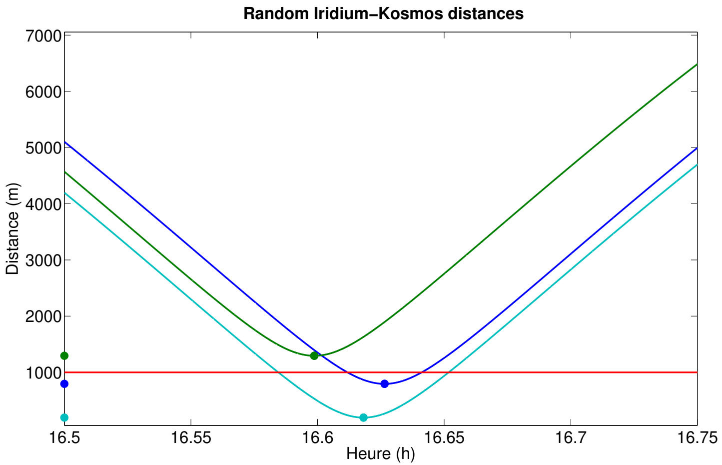
Discard points over quantile.

Applied AST in images



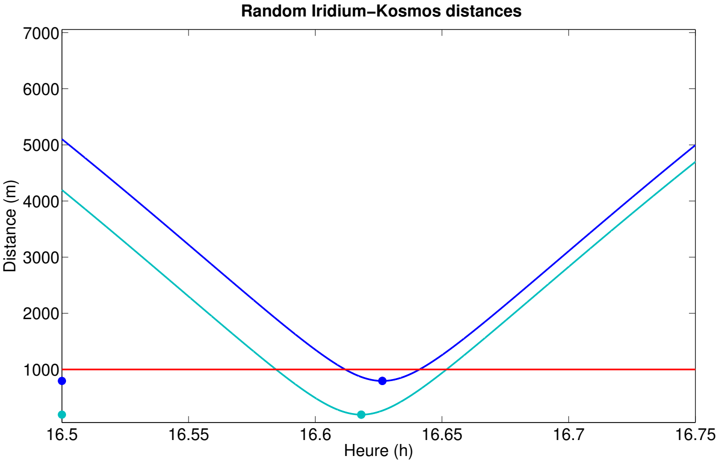
Discard points over quantile.

Applied AST in images



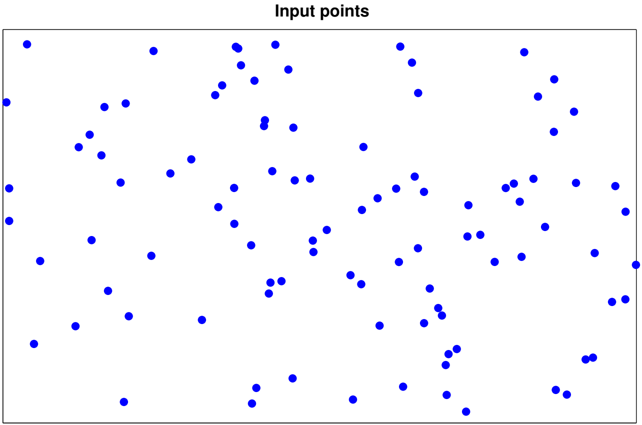
Discard points over quantile.

Applied AST in images



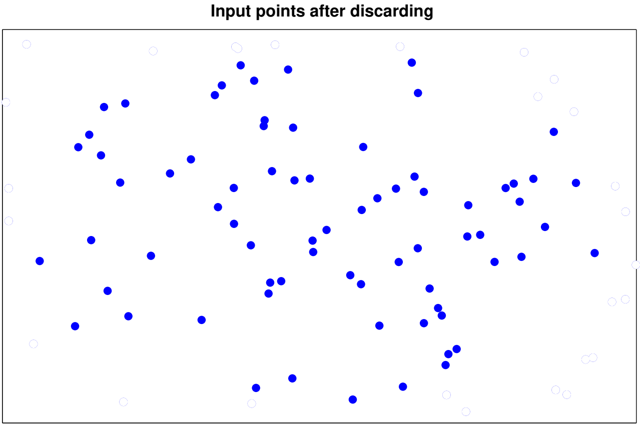
Discard points over quantile.

Applied AST in images



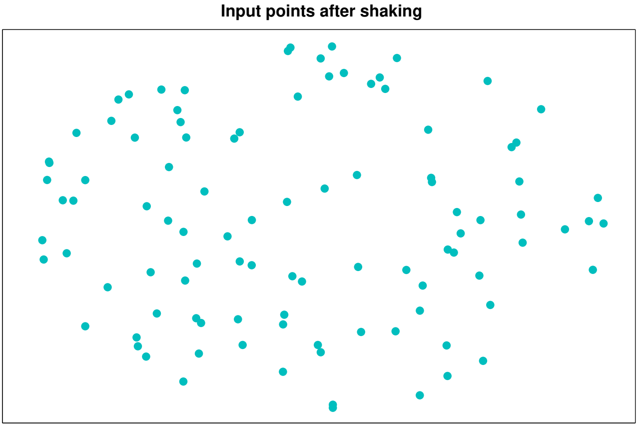
Discard points over quantile.

Applied AST in images



Discard points over quantile.

Applied AST in images



Use the markov kernel and selected points to regenerate sample.

AST estimation.
$d_{col} = 100$m, 100 estimations based on $309060(1 \pm 2\%)$ throws.

AST estimation.
$d_{col} = 100$m,100 estimations based on $309060(1 \pm 2\%)$ throws.

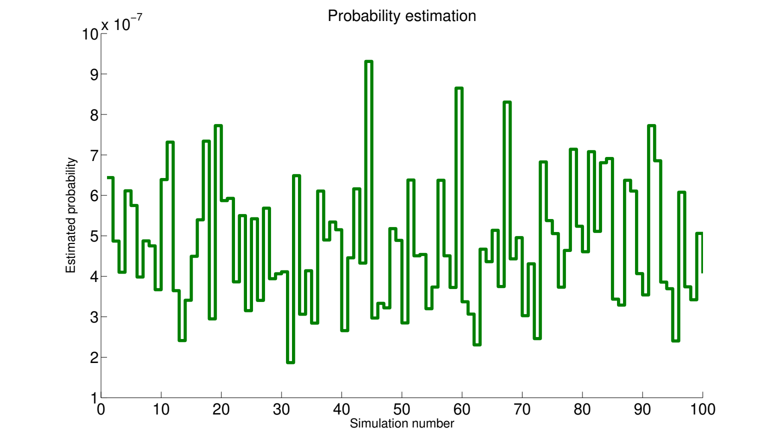


Figure: Probability estimations

AST estimation.

$d_{col} = 100$m, 100 estimations based on $309060(1 \pm 2\%)$ throws.

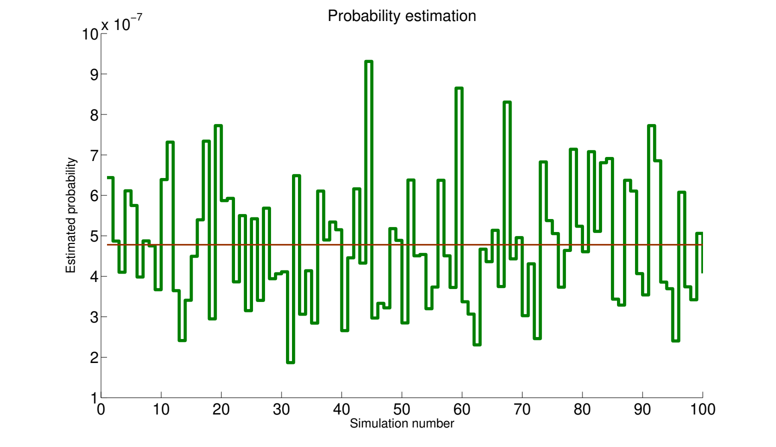


Figure: Probability estimations with their mean.

AST estimation.

$d_{col} = 100$m, 100 estimations based on $309060(1 \pm 2\%)$ throws.

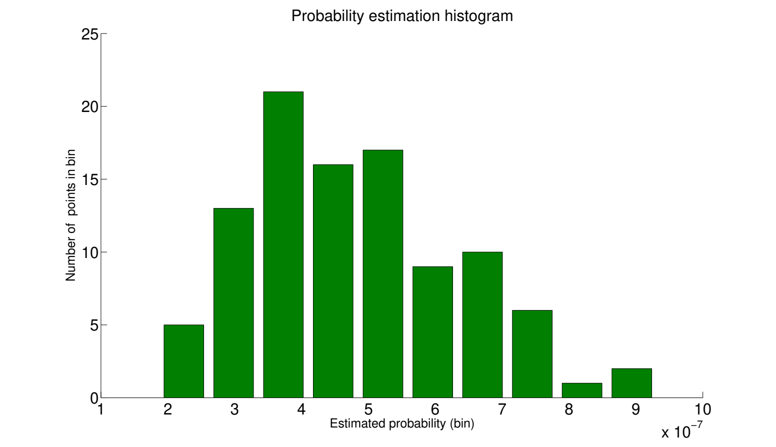
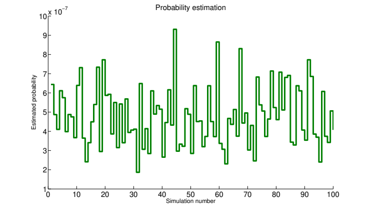
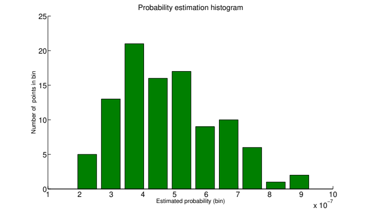


Figure: Probability estimation histogram

AST estimation. $d_{col} = 100$m



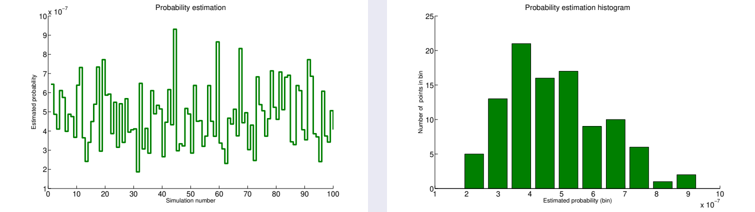
(a) Probability estimations



(b) Probability estimation histogram

Figure: 100 estimations based on $309060(1 \pm 2\%)$ simulations.

AST estimation. $d_{col} = 100\text{m}$

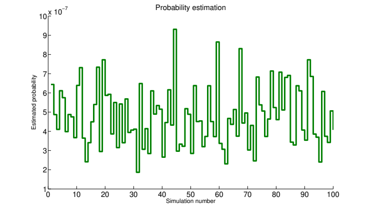


(a) Probability estimations (b) Probability estimation histogram

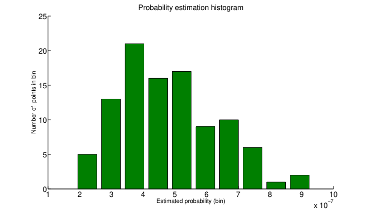
Figure: 100 estimations based on $309060(1 \pm 2\%)$ simulations.

- Mean estimate: $4.78 \cdot 10^{-7}$.
- Empirical relative deviation: 0.3232.

AST estimation. $d_{col} = 100$m



(a) Probability estimations



(b) Probability estimation histogram

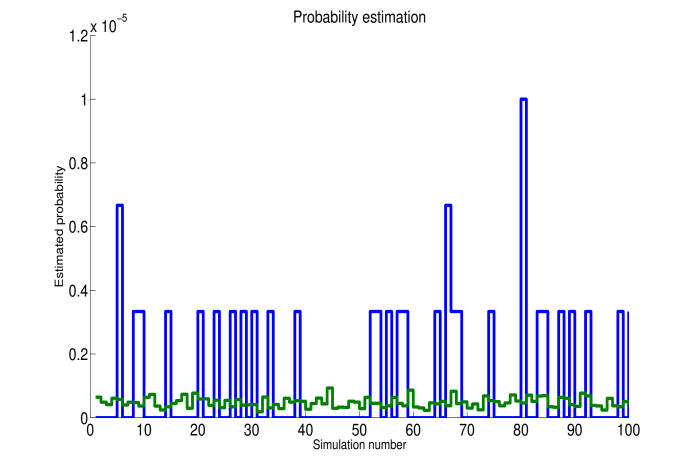
Figure: 100 estimations based on $309060(1 \pm 2\%)$ simulations.

- Mean estimate: $4.78 \cdot 10^{-7}$.
- Empirical relative deviation: 0.3232.

Quite Reliable!

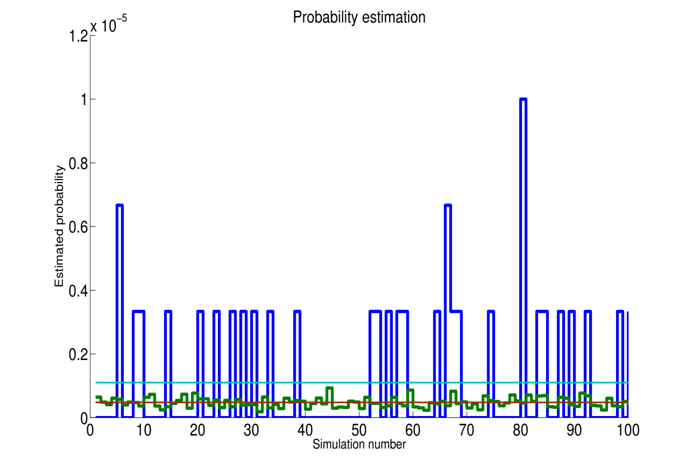
CMC Vs AST

Figure: Estimations



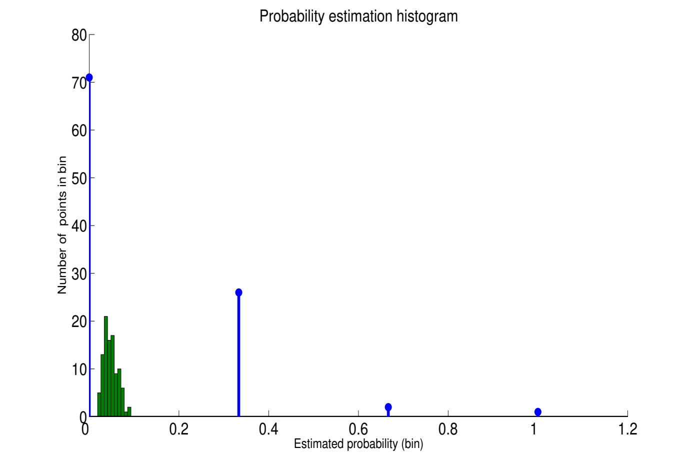
CMC Vs AST

Figure: Estimations and means



CMC Vs AST

Figure: Histograms



CMC Vs AST

Figure: Estimations

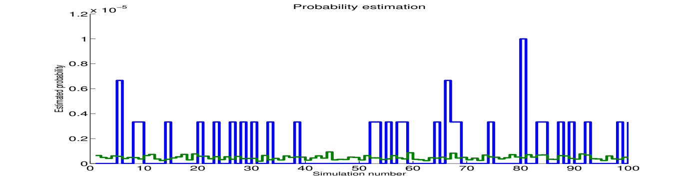
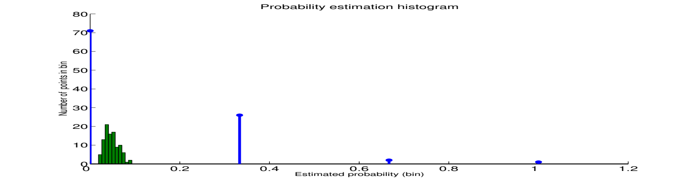
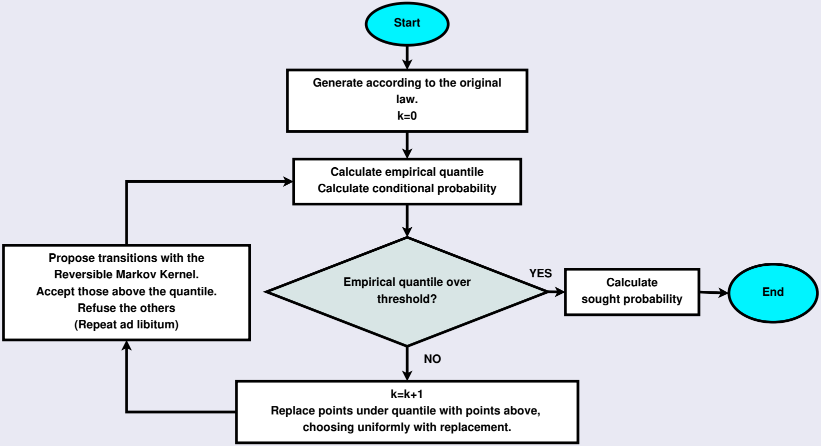


Figure: Histograms



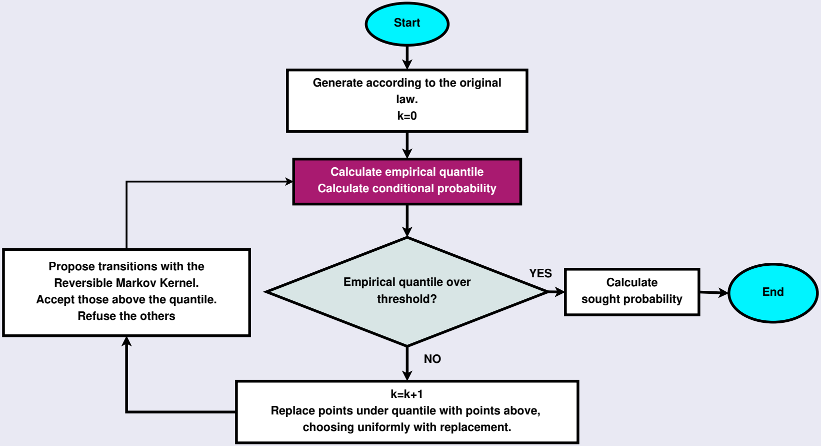
Outline

AST logigram with Markovian resampling



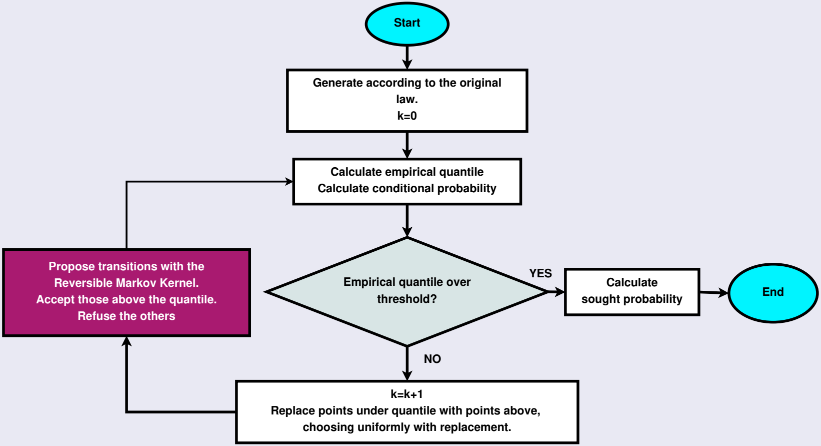
A global understanding of the AST with markovian resampling!

AST logigram with Markovian resampling and difficulties



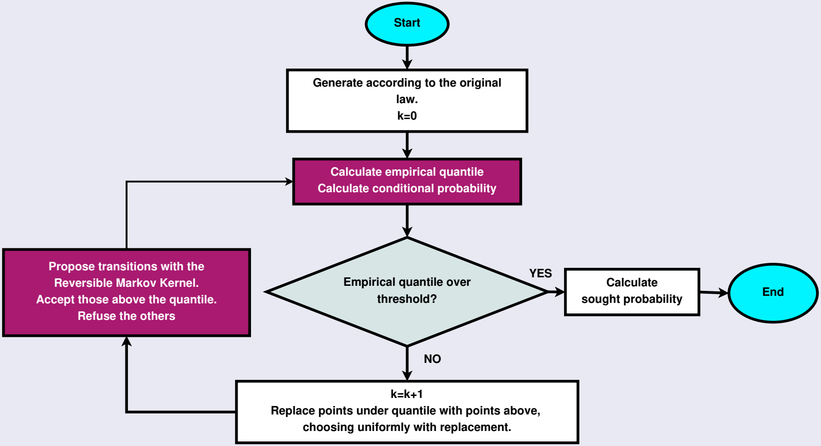
However, no rational quantile level choice...

AST logigram with Markovian resampling and difficulties



...nor Markovian choice and tuning!

AST logigram with Markovian resampling and difficulties



Two issues to tackle!

The Adaptive Splitting Technique is

The Adaptive Splitting Technique is

- a interacting particle technique [2]

The Adaptive Splitting Technique is

- a interacting particle technique [2]
- a variance reduction method

The Adaptive Splitting Technique is

- a interacting particle technique [2]
- a variance reduction method
- used as a rare event dedicated Monte Carlo method [5, 6].

Basic Ingredients [1]

Basic Ingredients [1]

	Given
f_X	Input Probability Density Function
T	Threshold to exceed
N	Simulation budget
	Known or designed
$M(\cdot,\cdot)$	f_X-reversible Markov Kernel
	Decided and deterministic
a_k	k^{th} Quantile level
b_k	k^{th} kernel parameter set
n_k	k^{th} step simulation budget
	Observed and random
δ_k	k^{th} empirical quantile
κ	Total number of quantiles

Gaussian-reversible Markov Kernel [1]

Gaussian-reversible Markov Kernel [1]

If $X \sim \mathcal{N}(0_d, I_d)$, then a possible kernel choice is

$$M(x, dy) \sim \mathcal{N}\left(\frac{x}{\sqrt{1+b^2}}, \frac{b^2}{1+b^2}\right)$$

Gaussian-reversible Markov Kernel [1]

If $X \sim \mathcal{N}(0_d, I_d)$, then a possible kernel choice is

$$M(x, dy) \sim \mathcal{N}\left(\frac{x}{\sqrt{1+b^2}}, \frac{b^2}{1+b^2}\right)$$

Up to now, parameter b is initiated **heuristically**

Gaussian-reversible Markov Kernel [1]

If $X \sim \mathcal{N}(0_d, I_d)$, then a possible kernel choice is

$$M(x, dy) \sim \mathcal{N}\left(\frac{x}{\sqrt{1+b^2}}, \frac{b^2}{1+b^2}\right)$$

Up to now, parameter b is initiated **heuristically** and optionally self tunes as a whole set dependent Markov chain.

Gaussian-reversible Markov Kernel [1]

If $X \sim \mathcal{N}(0_d, I_d)$, then a possible kernel choice is

$$M(x, dy) \sim \mathcal{N}\left(\frac{x}{\sqrt{1+b^2}}, \frac{b^2}{1+b^2}\right)$$

Up to now, parameter b is initiated **heuristically** and optionally self tunes as a whole set dependent Markov chain.

General case

Gaussian-reversible Markov Kernel [1]

If $X \sim \mathcal{N}(0_d, I_d)$, then a possible kernel choice is

$$M(x, dy) \sim \mathcal{N}\left(\frac{x}{\sqrt{1+b^2}}, \frac{b^2}{1+b^2}\right)$$

Up to now, parameter b is initiated **heuristically** and optionally self tunes as a whole set dependent Markov chain.

General case

In other cases, no off-the-shelf reversible kernel.

Gaussian-reversible Markov Kernel [1]

If $X \sim \mathcal{N}(0_d, I_d)$, then a possible kernel choice is

$$M(x, dy) \sim \mathcal{N}\left(\frac{x}{\sqrt{1+b^2}}, \frac{b^2}{1+b^2}\right)$$

Up to now, parameter b is initiated **heuristically** and optionally self tunes as a whole set dependent Markov chain.

General case

In other cases, no off-the-shelf reversible kernel.

Metropolis-Hasting algorithm can help [8].

Gaussian-reversible Markov Kernel [1]

If $X \sim \mathcal{N}(0_d, I_d)$, then a possible kernel choice is

$$M(x, dy) \sim \mathcal{N}\left(\frac{x}{\sqrt{1+b^2}}, \frac{b^2}{1+b^2}\right)$$

Up to now, parameter b is initiated **heuristically** and optionally self tunes as a whole set dependent Markov chain.

General case

In other cases, no off-the-shelf reversible kernel.

Metropolis-Hasting algorithm can help [8].

A wider array of density-reversible Markov kernel pair is needed.

Sample dependence

Sample dependence

Samples are **correlated and identically distributed** as they share a common genealogy.

Sample dependence

Samples are **correlated and identically distributed** as they share a common genealogy. This correlation can result into variance.

Sample dependence

Samples are **correlated and identically distributed** as they share a common genealogy. This correlation can result into variance. **Iterating** $M_k(\cdot, \cdot)$ never hurts and can reduce variance [8, 7].

Sample dependence

Samples are **correlated and identically distributed** as they share a common genealogy. This correlation can result into variance. **Iterating** $M_k(\cdot, \cdot)$ never hurts and can reduce variance [8, 7].

How to create independence?
How fast is independence reached?
How to create diversity?

Most efficient combination of quantile levels

Most efficient combination of quantile levels

According to [4], to minimise variance, under mild assumptions, all quantile levels a_k should be equal, say to a.

Most efficient combination of quantile levels

According to [4], to minimise variance, under mild assumptions, all quantile levels a_k should be equal, say to a.

Best level choice

Most efficient combination of quantile levels

According to [4], to minimise variance, under mild assumptions, all quantile levels a_k should be equal, say to a.

Best level choice

There is no idea now of an optimal choice for a !

Most efficient combination of quantile levels

According to [4], to minimise variance, under mild assumptions, all quantile levels a_k should be equal, say to a.

Best level choice

There is no idea now of an optimal choice for a !

- Extremely high value should be avoided as we want to avoid rare events.

Most efficient combination of quantile levels

According to [4], to minimise variance, under mild assumptions, all quantile levels a_k should be equal, say to a.

Best level choice

There is no idea now of an optimal choice for a !

- Extremely high value should be avoided as we want to avoid rare events.
- Experimentally, $a \in [.20, .25]$ works fine [1].

Most efficient combination of quantile levels

According to [4], to minimise variance, under mild assumptions, all quantile levels a_k should be equal, say to a.

Best level choice

There is no idea now of an optimal choice for a !

- Extremely high value should be avoided as we want to avoid rare events.
- Experimentally, $a \in [.20, .25]$ works fine [1].
- Very low a too, when dealt with carefullly [3].

Most efficient combination of quantile levels

According to [4], to minimise variance, under mild assumptions, all quantile levels a_k should be equal, say to a.

Best level choice

There is no idea now of an optimal choice for a !

- Extremely high value should be avoided as we want to avoid rare events.
- Experimentally, $a \in [.20, .25]$ works fine [1].
- Very low a too, when dealt with carefullly [3].

How to choose a to reduce variance?
How to choose a to respect the simulation budget?

Outline

Key ideas to remember

Key ideas to remember

- AST widely outperforms CMC when it comes to rare events .

Key ideas to remember

- AST widely outperforms CMC when it comes to rare events .
- Rule of the thumb tuning can do the trick...

Key ideas to remember

- AST widely outperforms CMC when it comes to rare events .
- Rule of the thumb tuning can do the trick...
- ...but theoretical understanding is needed.

Further work: theoretical mastery

Further work: theoretical mastery

- Applying it to a wider class of random input via Metropolis-Hasting and using empirical stopping criterion.

Further work: theoretical mastery

- Applying it to a wider class of random input via Metropolis-Hasting and using empirical stopping criterion.
- Using extra knowledge $\mathbb{P}\left[X \in \mathcal{A}\right] \in [p_-, p^+]$ to choose parameters.

Further work: theoretical mastery

- Applying it to a wider class of random input via Metropolis-Hasting and using empirical stopping criterion.
- Using extra knowledge $\mathbb{P}[X \in \mathcal{A}] \in [p_-, p^+]$ to choose parameters.
- Should we resample all the points or only the doubloon?

Please, let me answer your questions.

Thank you for your attention.

Outline

Conferences

1. Estimating satellite versus debris collision probabilities via the adaptive splitting technique: 8^{th} *international workshop on rare event simulation*, Poster, 20-23 June 2010 Cambridge (UK), R. Pastel, J. Morio & F. Le Gland.
2. Representing Spatial Distribution via the Adaptive Splitting technique and Isoquantile Curves : *European Meeting of Statisticians 2010*, Présentation Orale sur résumé, 17-22 August 2010 Piraeus (Greece), R. Pastel, J. Morio & F. Le Gland.
3. From satellite versus debris collision probabilities to the adaptive splitting technique From Application to Theory : *Rare Event Simulation Workshop*, Présentation Orale sur résumé, 28-29 October 2010 Bordeaux (France), R. Pastel, J. Morio & F. Le Gland.
4. Estimating satellite collision probabilities via the adaptive splitting technique : *International Conference in Computer Modeling and Simulation*, Présentation Orale sur article, 7-9 January 2011 Mumbai (India), R. Pastel, J. Morio & F. Le Gland.

Revues

1. Sampling Technique for launcher impact safety zone estimation : *Acta astronautica*, 66(5-6): 736-741, 2010, J. Morio & R. Pastel.
2. An overview or importance splitting for rare event simulation : *European Journal of Physics*, 31:1295-1303, 2010, J. Morio, R. Pastel & F. Le Gland.
3. Estimation De probabilités et de quantiles rares pour la caractérisation d'une zone de retombée d'un engin : *Journal de la Société Française de Statistique*, pending subject to modifications, J. Morio, R. Pastel & F. Le Gland.

Outline

[1] Frédéric Cérou, Pierre Del Moral, Teddy Furon, and Arnaud Guyader.
Rare event simulation for a static distribution.
Technical report,
http://hal.inria.fr/inria-00350762/fr/, 2009.

[2] Pierre Del MORAL.
Feynman-Kac Formulae: Genealogical and Interacting Particle Systems With Applications.
Springer, 2004.

[3] Arnaud Guyader, Nicolas W. Hengartner, and Eric Matzner-Løber.
Iterative Monte Carlo for extreme quantiles and extreme probabilities.
Technical report, March 2010.
Paper and slideshow and presentation available on line.

[4] Agnés Lagnoux.

Rare event simulation.
PEIS, 20(1):45–66, January 2006.
Available on line.

[5] Pierre L'Écuyer, François Le Gland, Pascal Lezaud, and Bruno Tuffin.
Splitting methods.
In Gerardo Rubino and Bruno Tuffin, editors, *Monte Carlo Methods for Rare Event Analysis*, chapter 3, pages 39–61. John Wiley & Sons, Chichester, 2009.

[6] François LEGLAND.
Filtrage bayésien optimal et approximation particulaire.
Les Presses de l'Ecole Nationale Supérieure de Techniques Avancées, 2008.
Available on line in French.

[7] S. P. Meyn and R. L. Tweedie.
Markow chains and stochastic stability.
Springer–Verlag, 1993.

[8] Luke TIERNEY.
Markov chains for exploring posterior distributions.
Annals of Statistics, 22:1701–1728, December 1994.
Available on line.