# **Re**-sampled **Sca**lable **L**angevin **E**xact Method

Divakar Kumar

OxWaSP, University of Warwick

October, 14, 2016

# Overview

1. Motivation
2. The methodology
   - Simulating from the quasi-stationary density - Glynn and Blanchet's approach
   - ReScaLE algorithm
3. A toy example
   - Non-Uniform rebirth strategies

# Motivation

# Motivation

**The problem**

How to simulate from an intractable distribution $\pi$?

# Motivation

## The problem

How to simulate from an intractable distribution $\pi$?

## Standard MCMC

- Dynamics of a Markov process $\longrightarrow$ invariant density is $\pi$.
- The transition density $p_{(0,t)}(0, \cdot) \longrightarrow \pi$.
- Empirical density $\hat{\pi} \longrightarrow$ target density $\pi$.

# Motivation

## The problem

How to simulate from an intractable distribution $\pi$?

## Standard MCMC

- Dynamics of a Markov process $\longrightarrow$ invariant density is $\pi$.
- The transition density $p_{(0,t)}(0,\cdot) \longrightarrow \pi$.
- Empirical density $\hat{\pi} \longrightarrow$ target density $\pi$.
- Very few $\pi$ can be expressed as an interesting algebraic expression.
- Multi-dimensional settings might not produce desired result.

# Quasi-Stationary Monte Carlo

## Quasi-stationary MC

$$q_{(0,t)}(0,x) := \frac{P(X_t \in dx \,|\, \zeta > t, X_0 = 0)}{dx} \tag{1}$$

- Dynamics of a Markov process $\longrightarrow$ quasi-stationary density is $\pi$.
- The quasi-stationary density $q_{(0,t)}(0,x) \longrightarrow \pi$.
- Empirical quasi-stationary density $\hat{\pi} \longrightarrow$ target density $\pi$.
- Pollock *et al.* [2016] showed that for a suitably chosen process $q_{(0,t)}(0,x) \longrightarrow \pi$.

# The methodology

# The transition density

## Langevin diffusion

$$dX_t = \tfrac{1}{2}\nabla \log \pi(X_t)dt + dB_t$$

- The invariant distribution is given by $\pi$.
- Simulate the trajectories of Langevin diffusion and look at the occupation measure.

$$p_{0,t}(0,x) \propto \exp\left\{-\frac{(x)^2}{2t}\right\}\pi(x)^{\frac{1}{2}}\mathbf{E}_{\mathbb{W}_{|X_t=x}}\left(\exp\left\{-\int_0^t \phi(X_s)ds\right\}\right)$$

- It is still difficult to draw according to $p_{0,t}(0,x)$.
- Drop $\pi(x)^{\frac{1}{2}}$ and converge to wrong density $\pi(x)^{\frac{1}{2}}$.
- Double the drift! $\longrightarrow$ Drop $\pi(x)$ again and converge to correct density $\pi(x)$ !

# Killed Brownian Motion

## Theorem (Transition density of killed brownian motion)

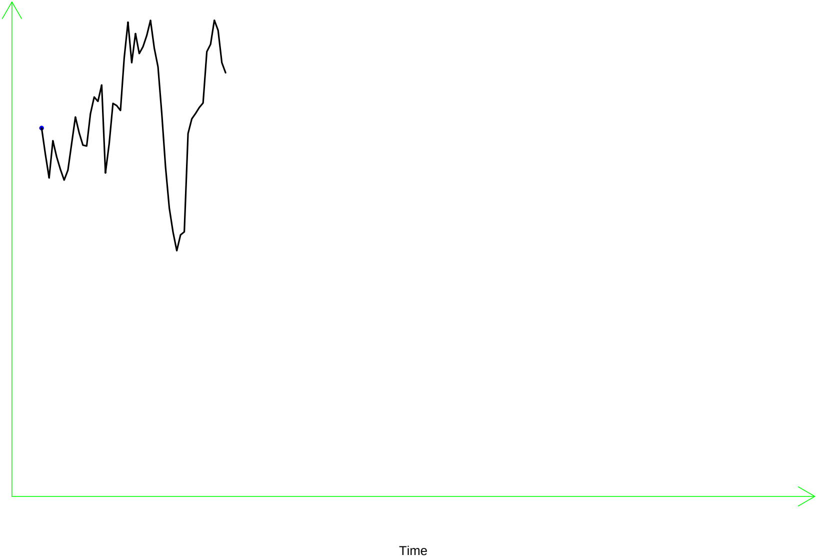
*Consider a standard Brownian motion $\{X_t : t \geq 0\}$ which is killed at $X_s$ with a state-dependent 'killing-rate' $\phi(X_s)$. Then the stationary density of this 'killed Brownian motion' conditional on its survival until time $t$ is given by*

$$q_{0,t}(0,x) \propto \exp\left\{-\frac{x^2}{2t}\right\} \mathbb{E}_{x_0,x}\left(\exp\left\{-\int_0^t \phi(X_s)ds\right\}\right). \tag{2}$$

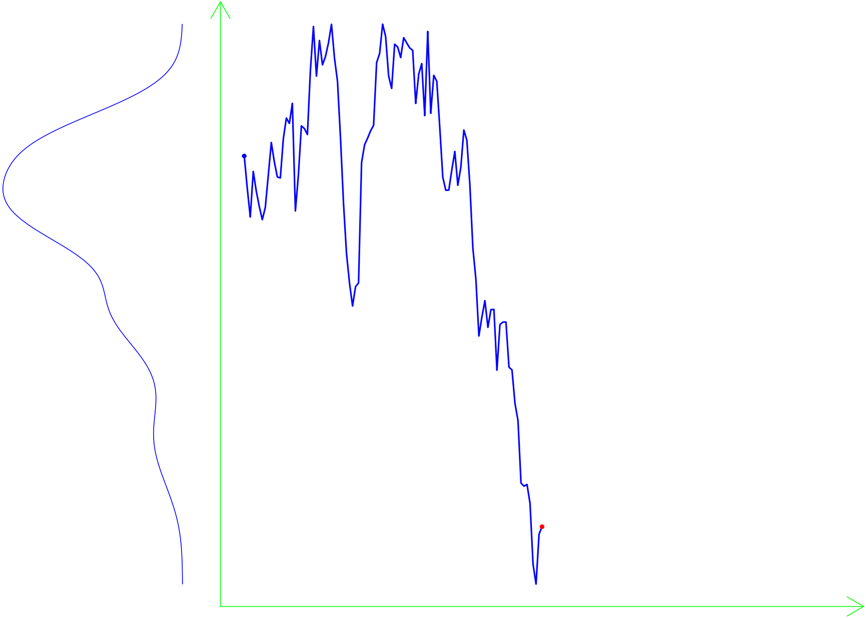
- Simulate from $\pi \longrightarrow$ simulate the quasi-stationary density of KBM.
    - The ScaLE method uses SMC-based approach to simulate the quasi-stationary density of KBM.

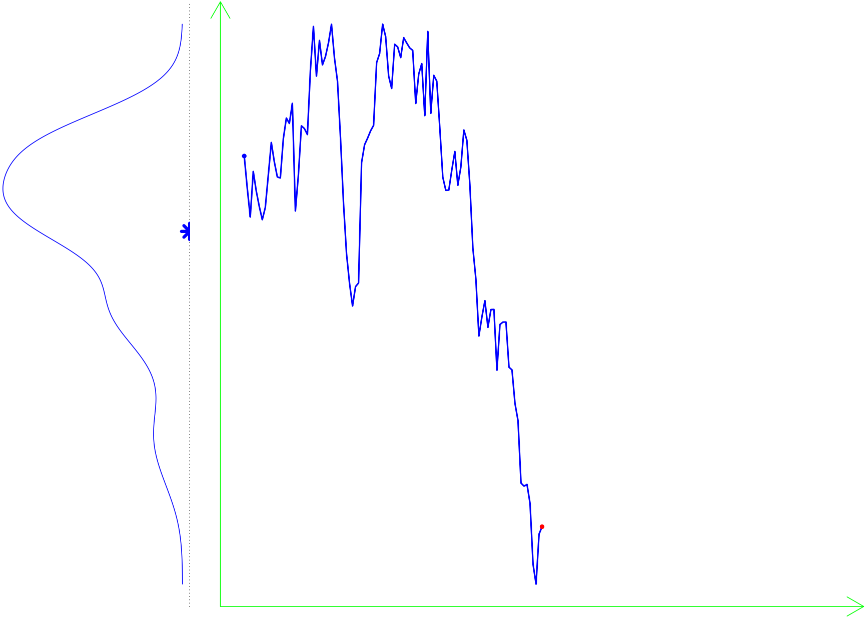
# Simulating from the quasi-stationary density - Glynn and Blanchet's approach

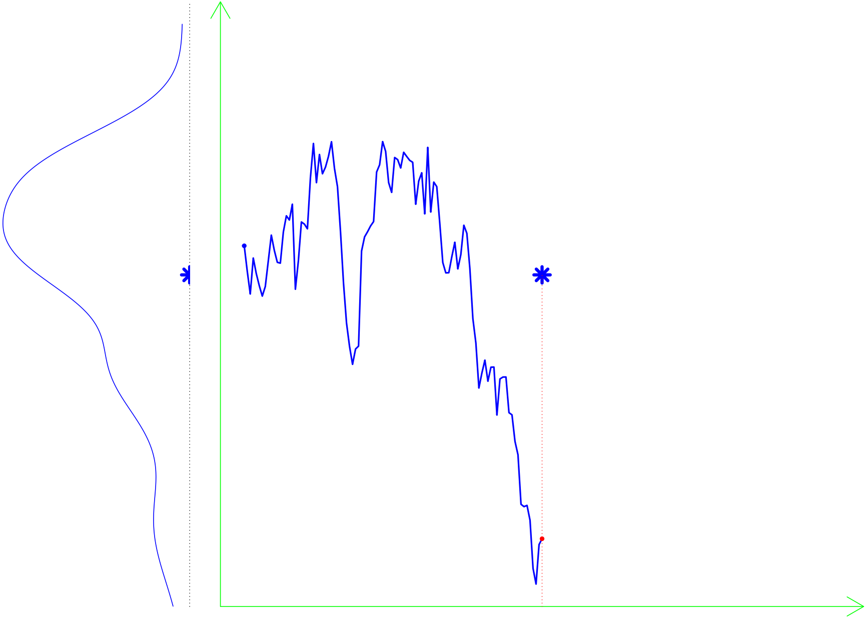

Time

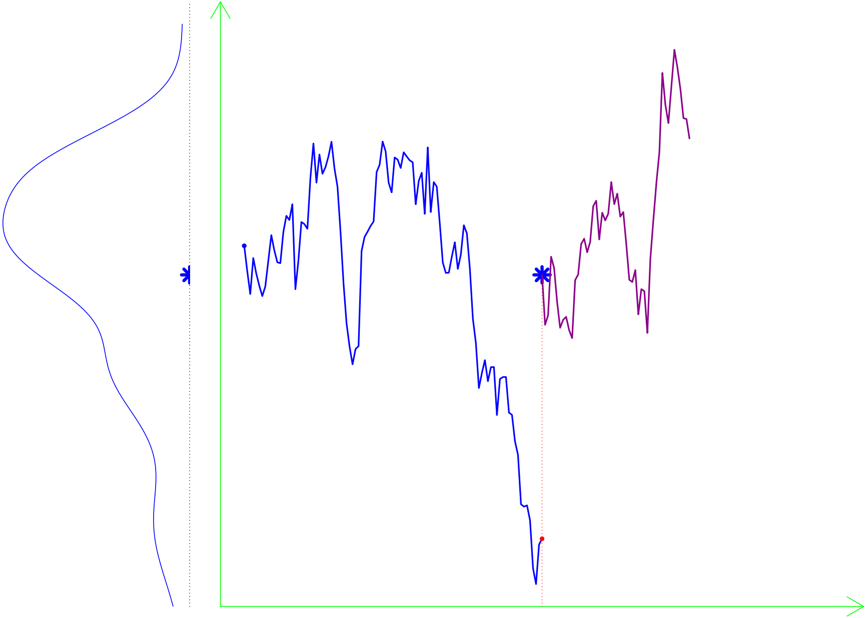

Time

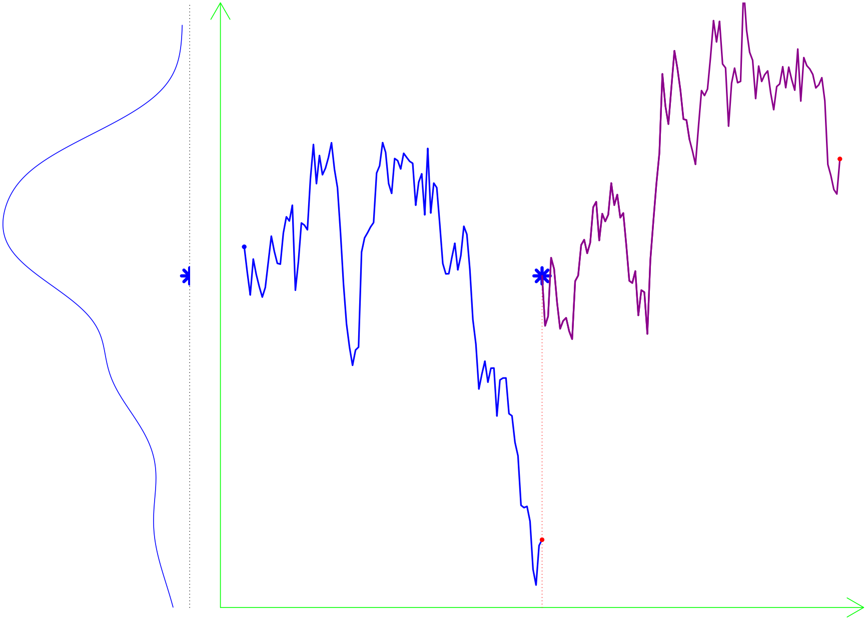

Time

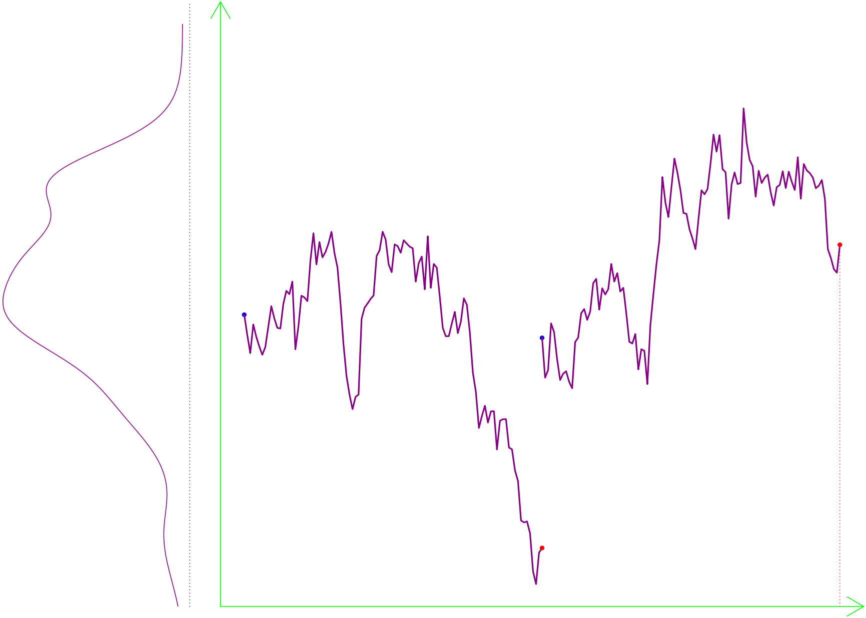












# How do we sample trajectories exactly?

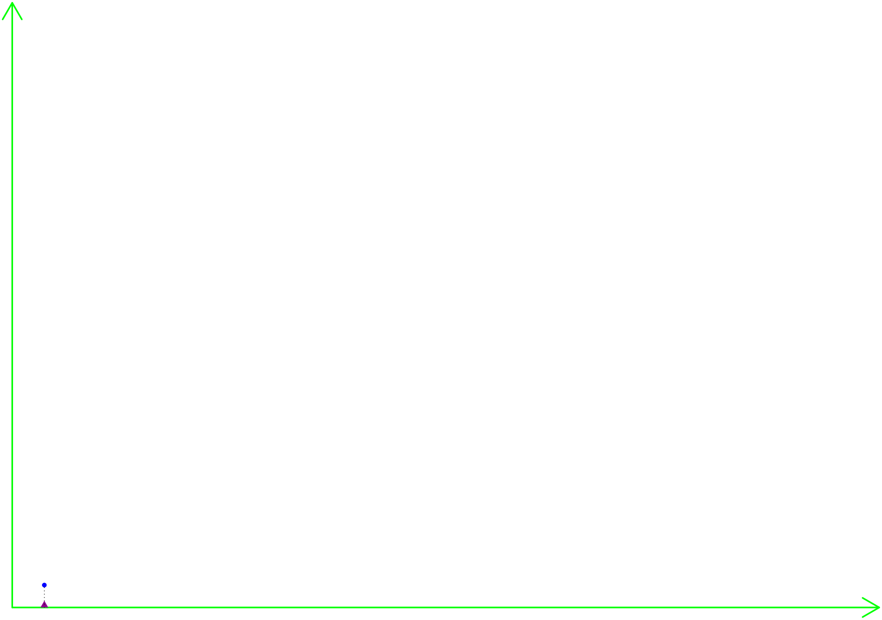
**Theorem (Colouring Scheme)**

*Let $\tau_1, ..., \tau_k$ be the Poisson Process with rate $M$ where $M$ is such that $\sup_x \phi(x) \leq M$. Let $X_{\tau_1}, ..., X_{\tau_k}$ be the realised skeleton of a Brownian motion $\{X_t : t \geq 0\}$ at times $\tau_1, ..., \tau_k$. If process is killed at $\tau_j$ with probability $\frac{\phi(X_{\tau_j})}{M}$. Then,*

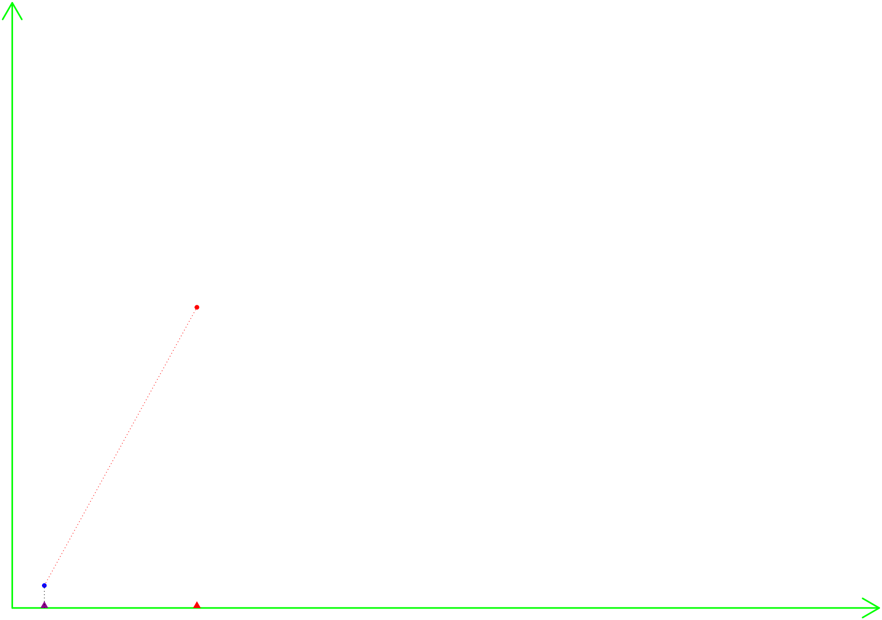
$$\mathbb{P}(\textit{Process survived until time}\, t) = \mathbb{E}_{0,x}\left(\exp\left\{-\int_0^t \phi(X_s)ds\right\}\right)$$

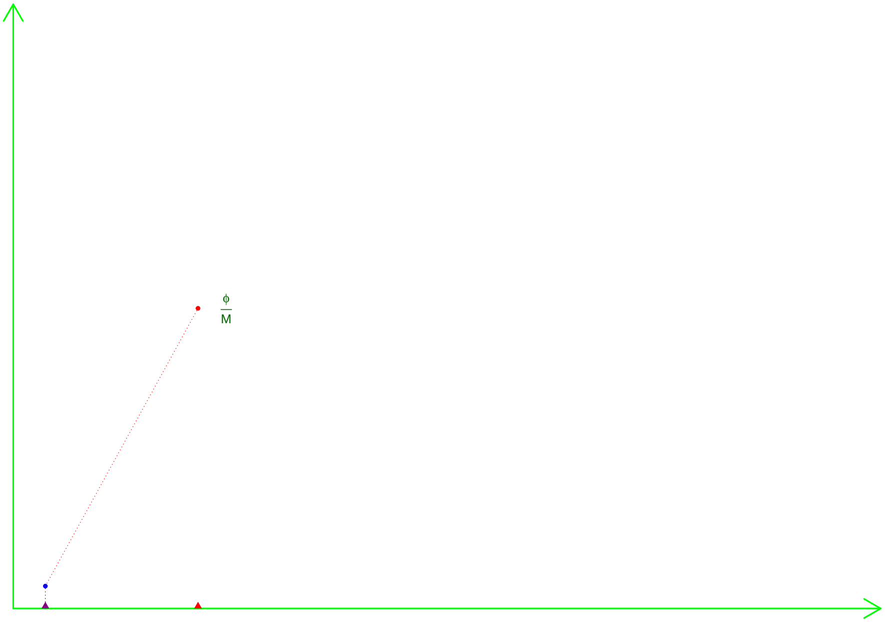
- Suggests to simulate $\tau_1, ..., \tau_k$ from homogeneous Poisson Process of rate $M$ and decide to kill the process at time of event $\tau_j$ with probability $\frac{\phi(X_{\tau_j})}{M}$.

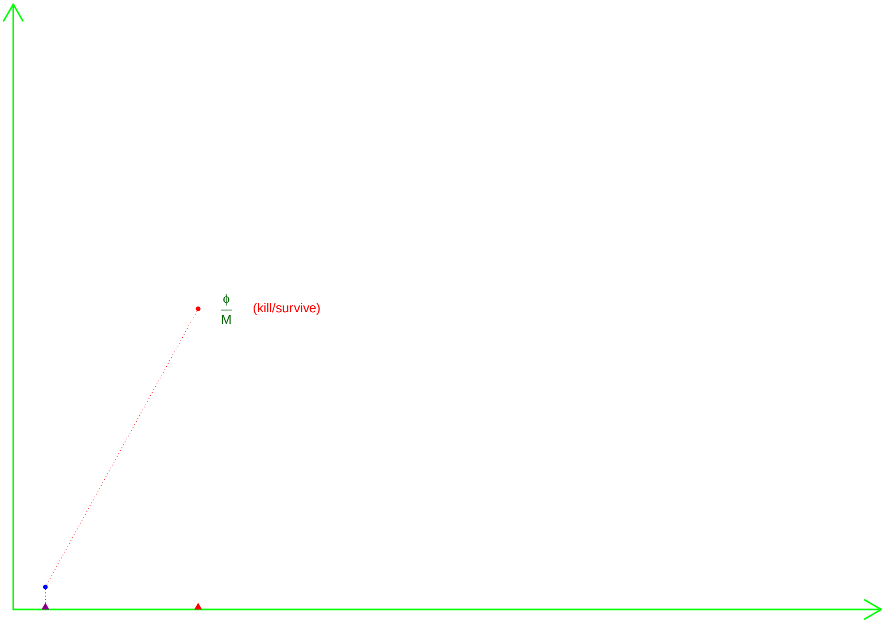
# ReScaLE algorithm

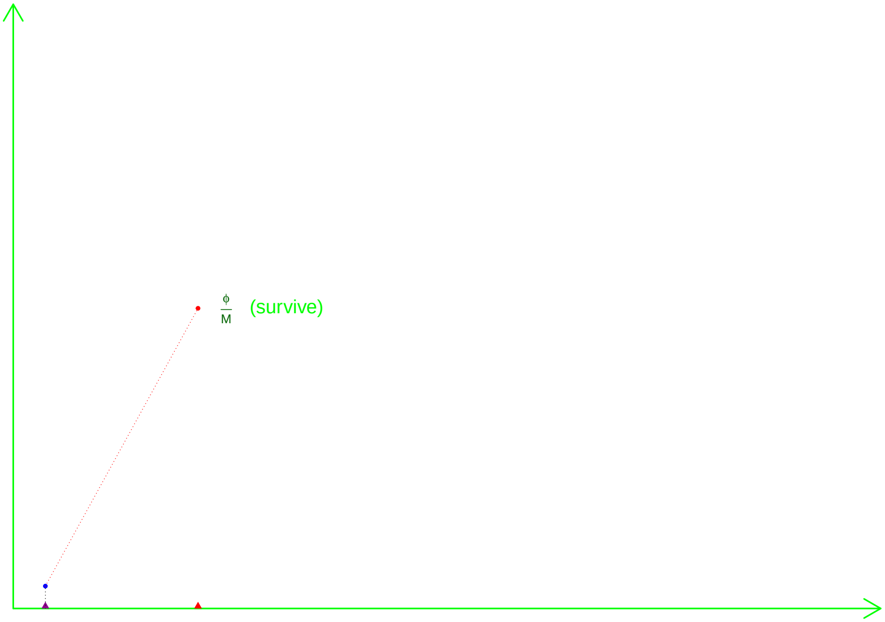


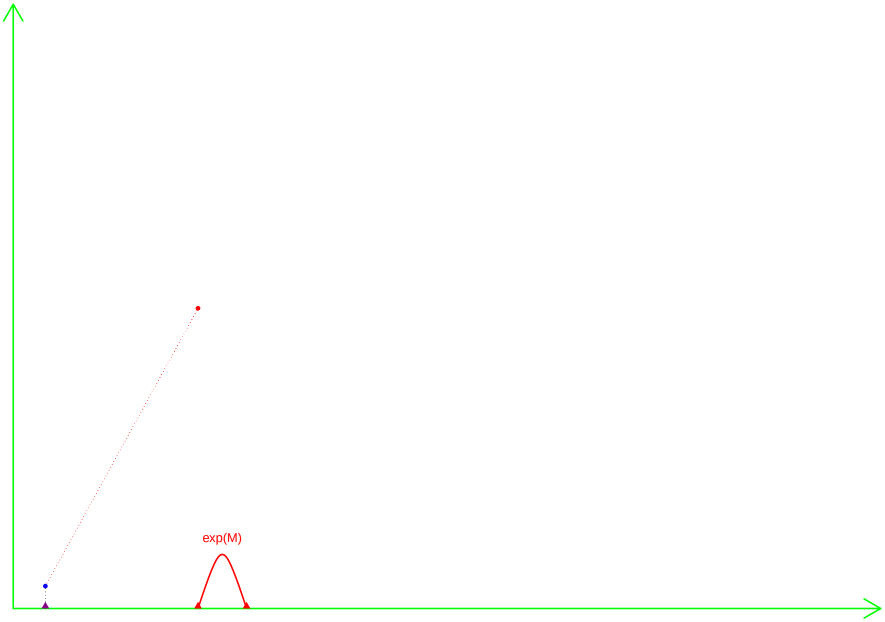

exp(M)

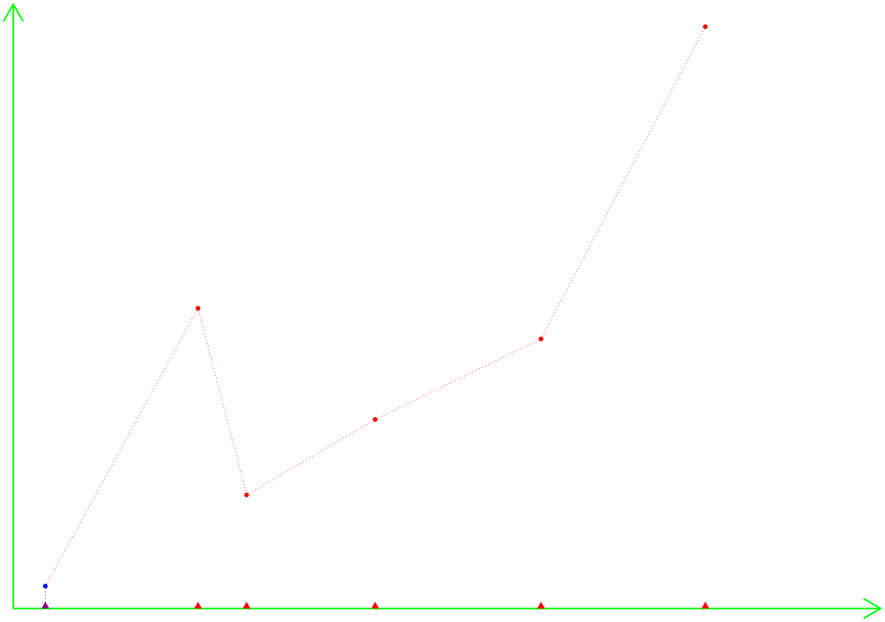


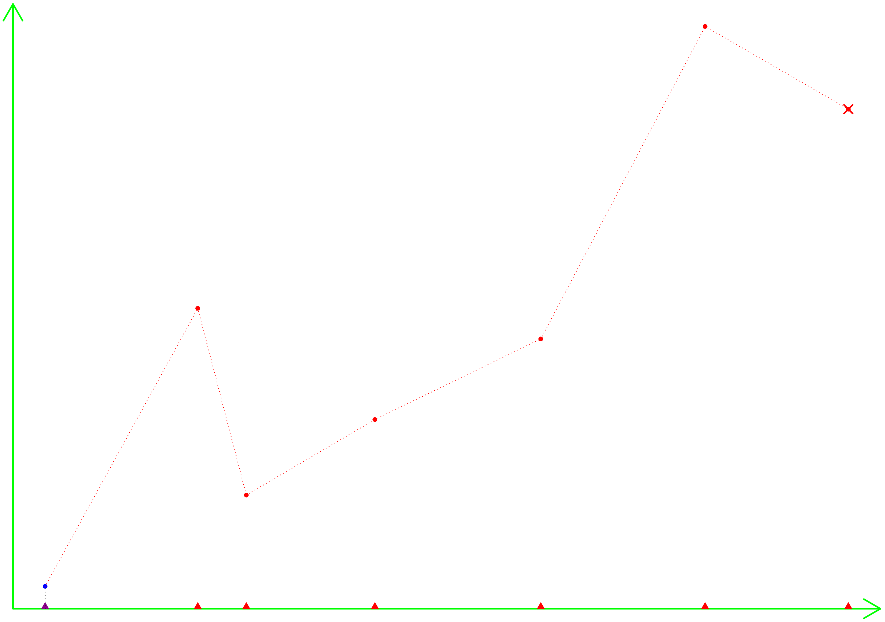

$\frac{\phi}{M}$

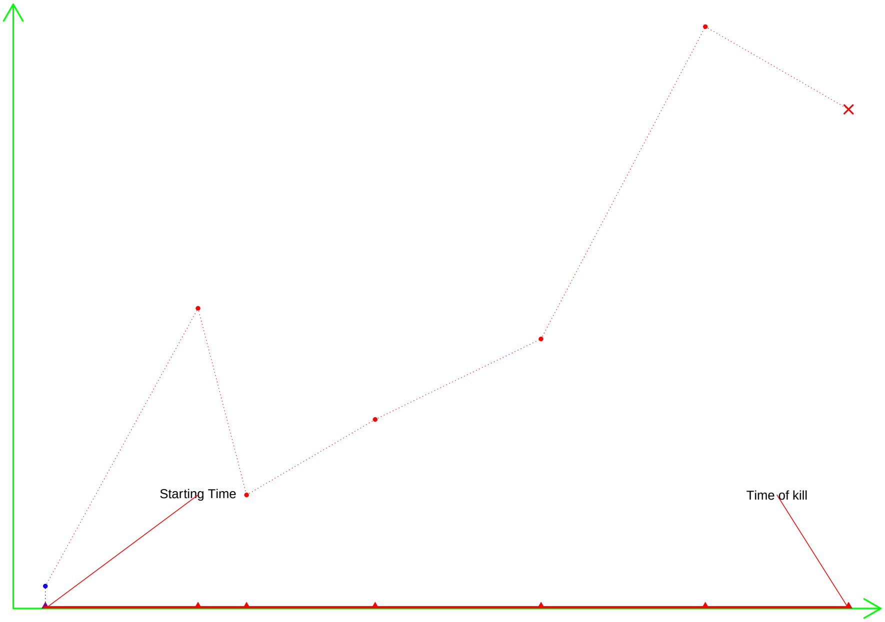

$\frac{\phi}{M}$
(kill/survive)

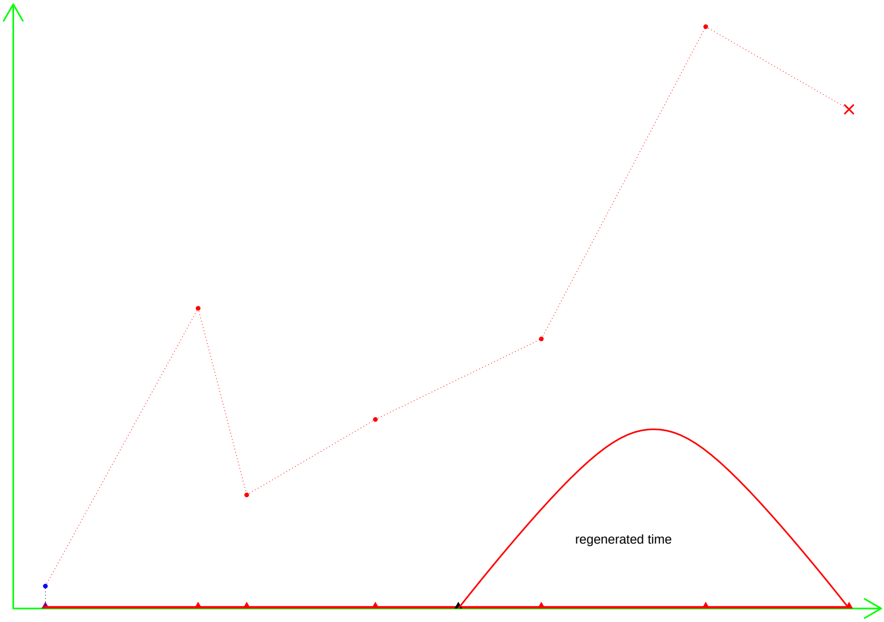

$\frac{\phi}{M}$
(survive)

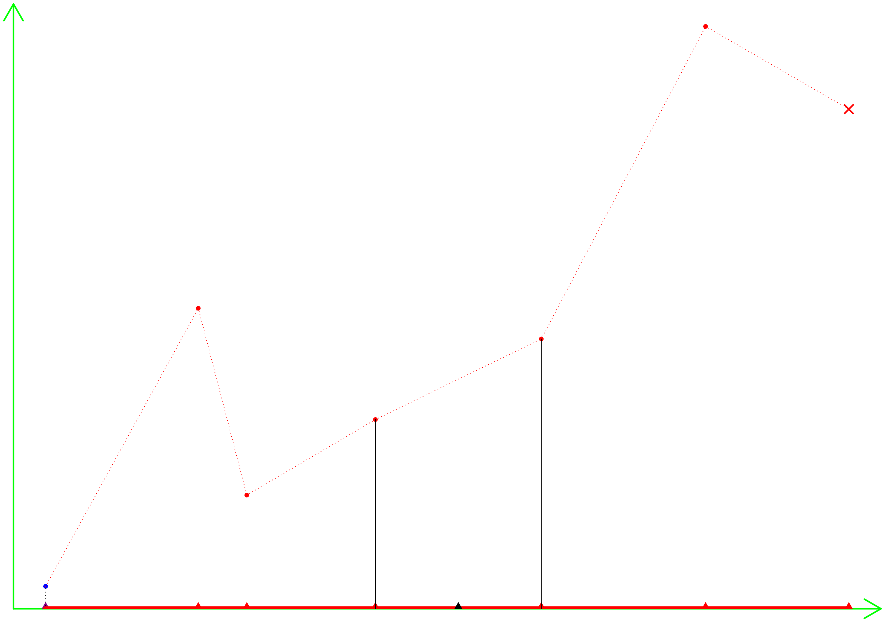

exp(M)

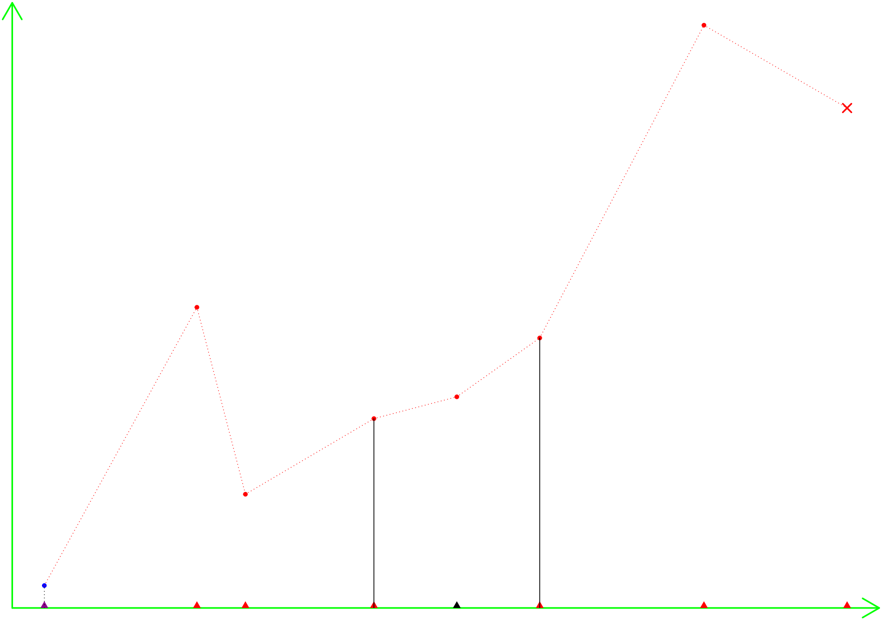


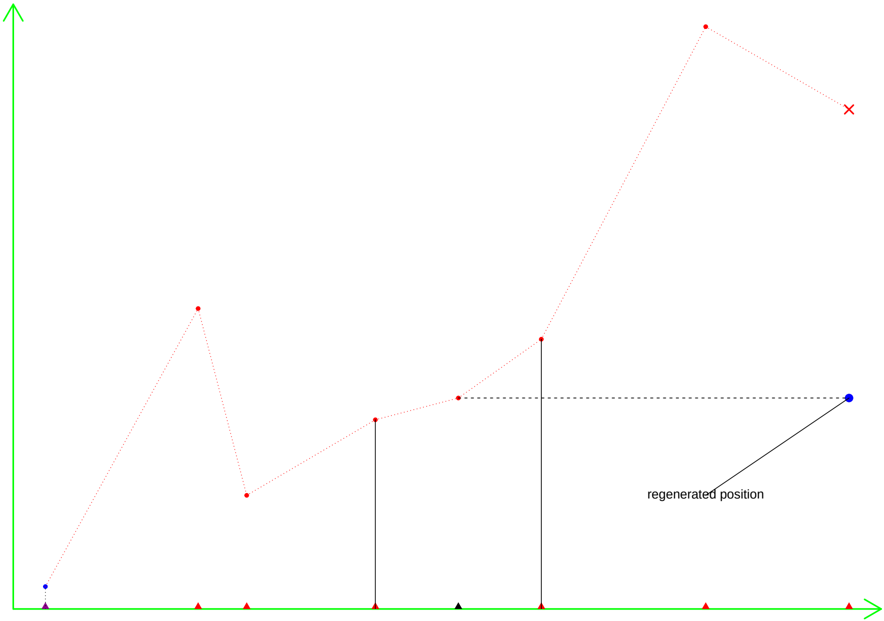


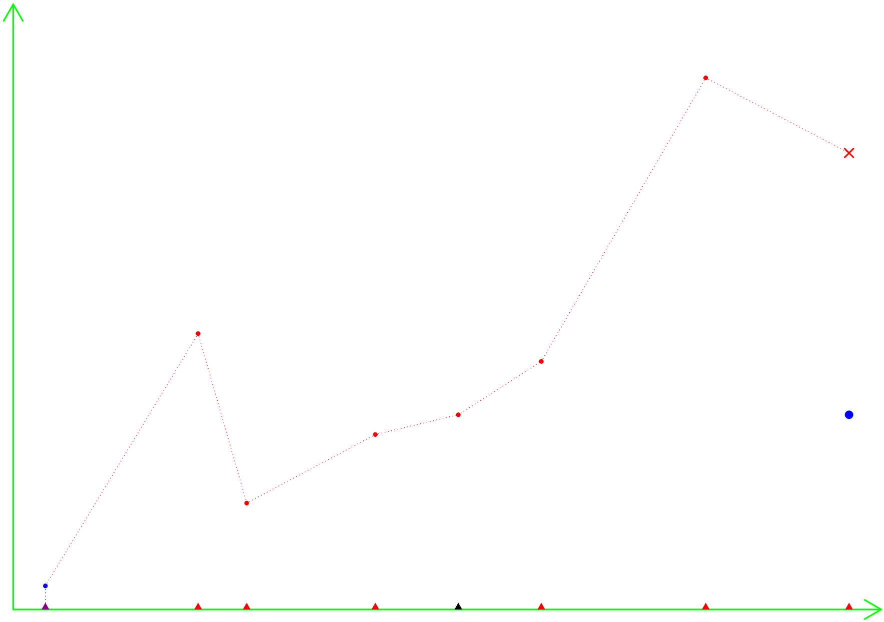

Starting Time
Time of kill


regenerated time






regenerated position



# A toy example

# An Example

We consider a toy example consisting of $5$ data points $y_i$ which are simulated from a density

$$\pi(y|x) \propto \frac{1}{1+(y-x)^2}. \tag{3}$$

The simulated data (when $x = 1$) is given as the following:

```
> y
[1] 2.65226687 1.27648783 1.61011759 1.27433040
    0.08721209.
```

Aim is to sample from the posterior distribution

$$\pi(x|\mathbf{y} = (y_1, ..., y_5)) \propto \frac{1}{1+x^2} \prod_{i=1}^{5} \frac{1}{1+(y_i-x)^2}. \tag{4}$$

# Chain Diagnostics

## Integrated Autocorrelation Time

$$\tau_{int} = 1 + 2\sum_{i=1}^{\infty} \rho_i.$$

# Chain Diagnostics

## Integrated Autocorrelation Time

$$\tau_{int} = 1 + 2\sum_{i=1}^{\infty} \rho_i.$$

- $(\tau_{int} = 5.11892)$.

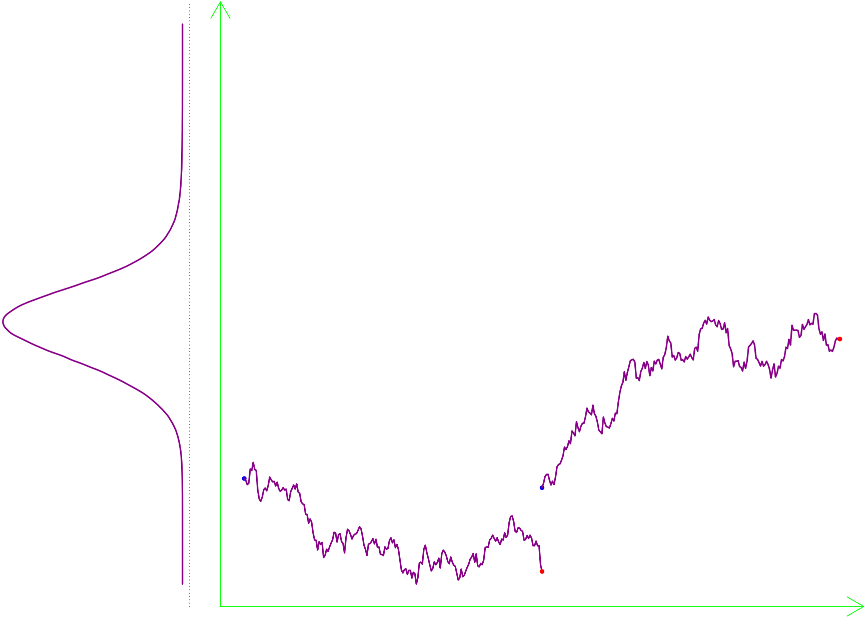
# Chain Diagnostics

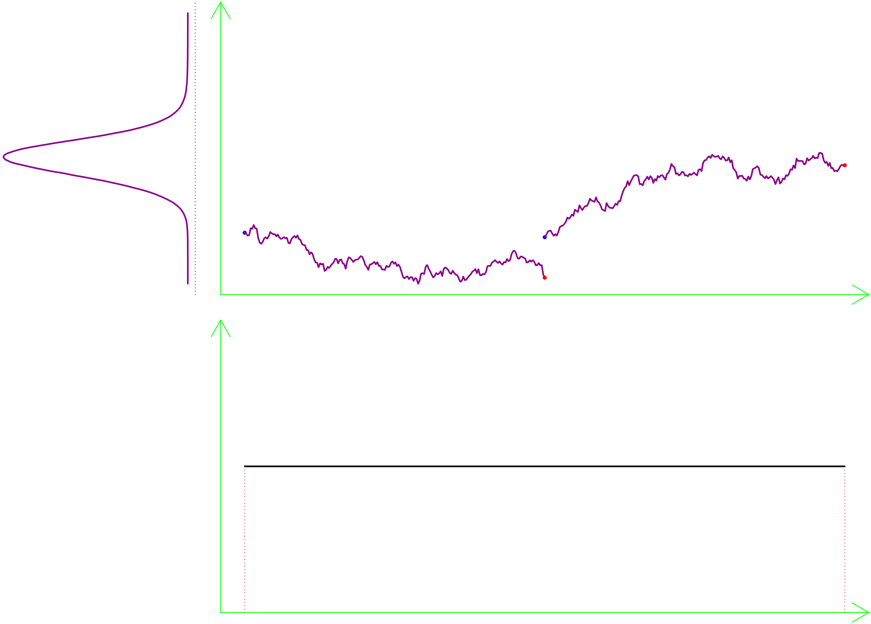
## Integrated Autocorrelation Time

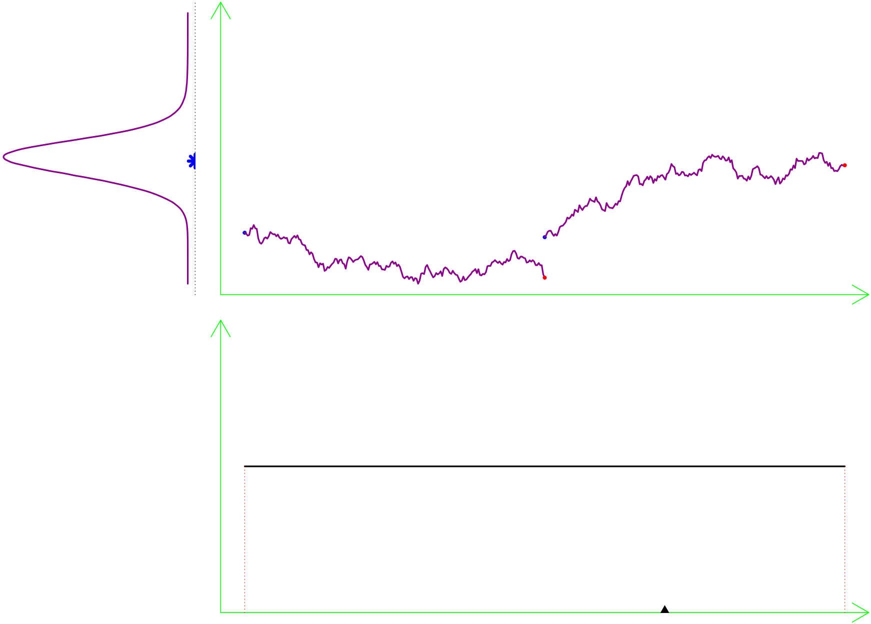
$$\tau_{int} = 1 + 2\sum_{i=1}^{\infty} \rho_i.$$

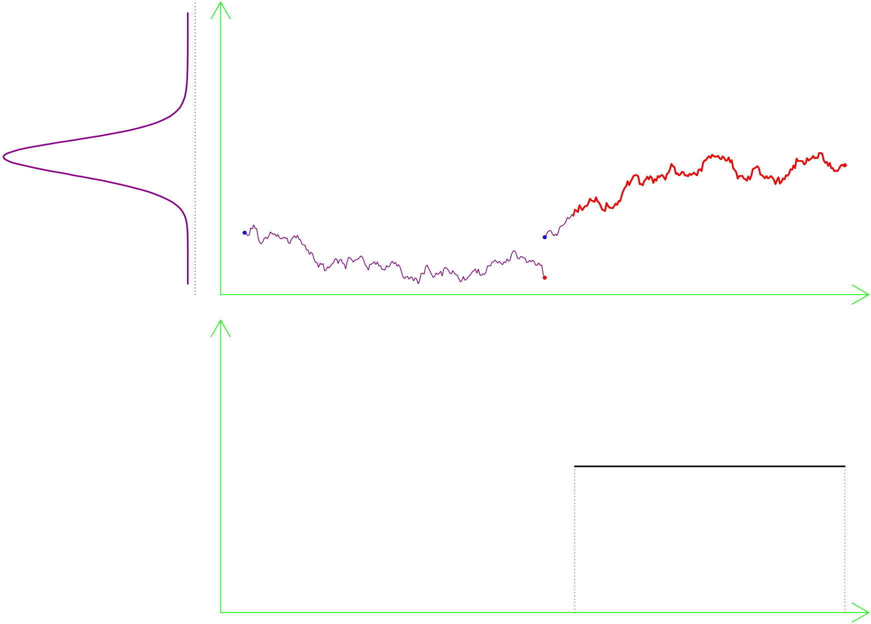
- $(\tau_{int} = 5.11892)$.
- K-S p-value = 0.490

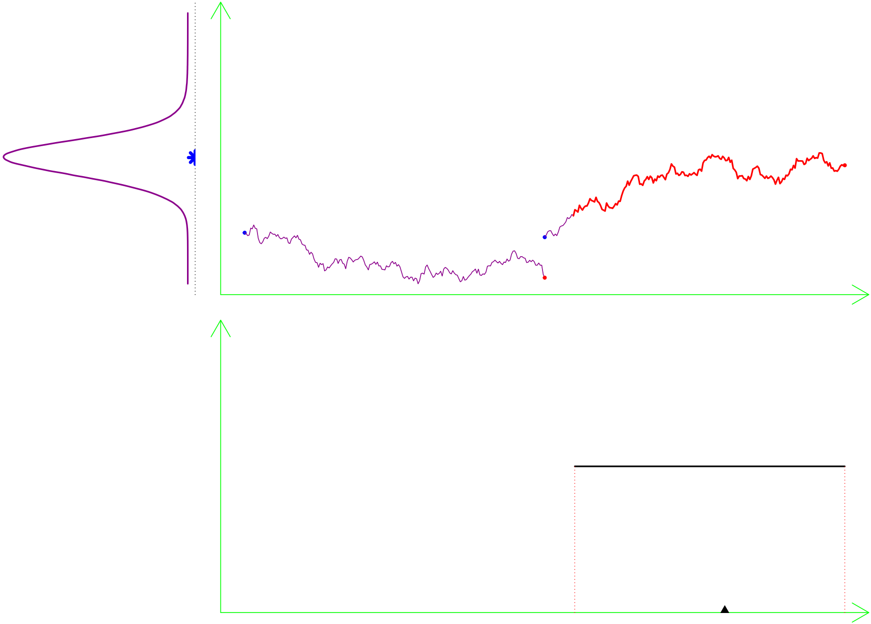
# Non-Uniform rebirth strategies

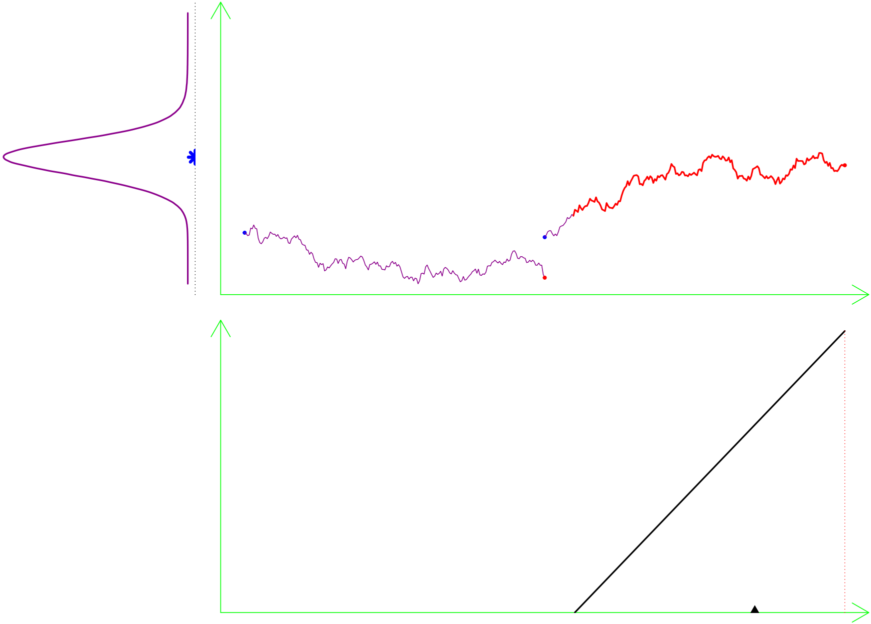


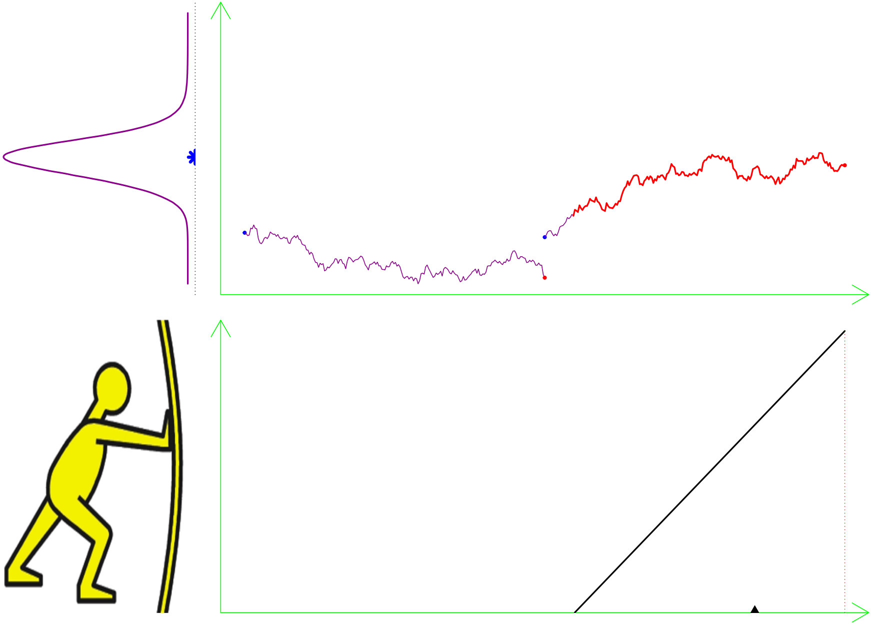


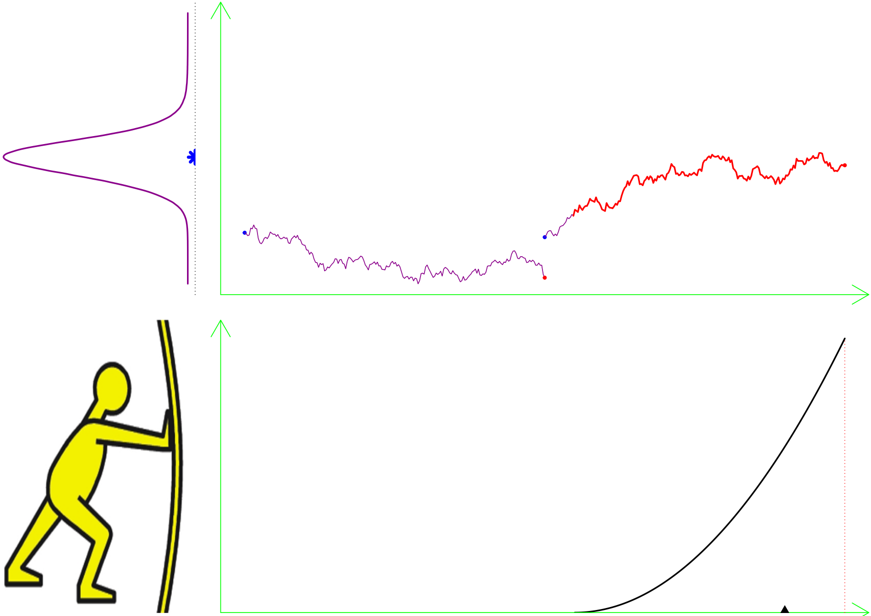




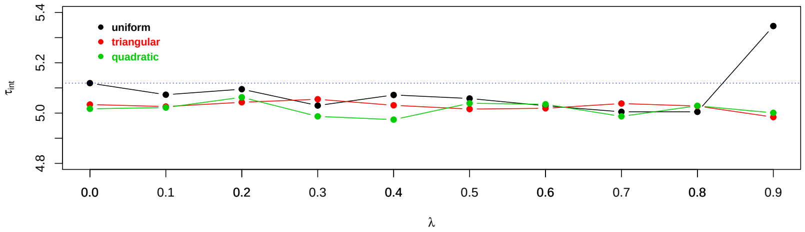


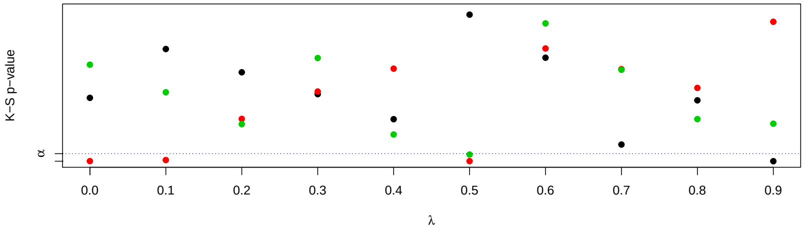






# Non-uniform rebirth strategy...



# Sub-sampling

$$\pi(x|\mathbf{y} = (y_1, ..., y_N)) \propto p(x) \prod_{i=1}^{N} \pi(y_i|x). \tag{5}$$

# Sub-sampling

$$\pi(x|\mathbf{y} = (y_1, ..., y_N)) \propto p(x) \prod_{i=1}^{N} \pi(y_i|x). \tag{5}$$

Then, the rate of kill function $\phi$ evaluates to

$$\phi(x) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(f_i'(x)f_j'(x) + \frac{1}{N}f_i''(x)\right)}{2} - l \tag{6}$$

$$f_i(x) = \log(\pi(y_i|x)) + \frac{1}{N}\log(p(x)) \tag{7}$$

# Sub-sampling

$$\pi(x|\mathbf{y} = (y_1, ..., y_N)) \propto p(x) \prod_{i=1}^{N} \pi(y_i|x). \tag{5}$$

Then, the rate of kill function $\phi$ evaluates to

$$\phi(x) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \left( f_i'(x) f_j'(x) + \frac{1}{N} f_i''(x) \right)}{2} - l \tag{6}$$

$$f_i(x) = \log(\pi(y_i|x)) + \frac{1}{N} \log(p(x)) \tag{7}$$

We choose an estimator as

$$\phi_{I,J}(x) := \frac{N^2 f_I'(x) f_J'(x) + N f_K''(x)}{2} - l. \tag{8}$$

Then

$$\mathbb{E}(\phi_{I,J}(x)) = \phi(x). \tag{9}$$

# Bibliography

Pollock, Murray, Fearnhead, Paul, Johansen, Adam M, & Roberts, G. O. 2016. The Scalable Langevin Exact Algorithm : Bayesian Inference for Big Data. *arXiv preprint:160903436*, 1–45.