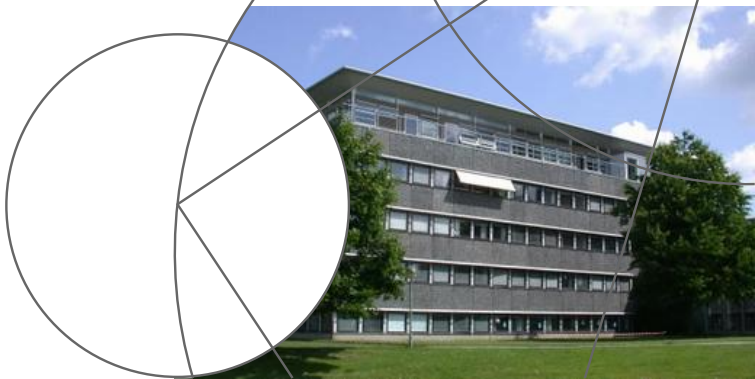
An MCMC approach to parameter estimation in the FitzHugh-Nagumo model

Anders Christian Jensen
Department of Mathematical Sciences



Outline

The FitzHugh-Nagumo (FHN) model

- Modeling of neuronal spike generation in axons
- The FHN model is a prototype for more complicated models, like the Hodgkin-Huxley, or Morris-Lecar model

$$\frac{\mathrm{d}}{\mathrm{d}t} x_t = \frac{1}{\varepsilon}\left(x_t - x_t^3 - y_t + s\right)$$

$$\frac{\mathrm{d}}{\mathrm{d}t} y_t = \gamma x_t - y_t + \beta$$

- Analyze the model
- Assume we observe x_t at discrete time points
- Subsampling, noisy, high frequency, ...
- Can we infer, and/or estimate the parameters?
- Parameters from other models
- Analyze the data, the model, and the inference

The FitzHugh-Nagumo (FHN) model

- Modeling of neuronal spike generation in axons
- The FHN model is a prototype for more complicated models, like the Hodgkin-Huxley, or Morris-Lecar model

$$\frac{\mathrm{d}}{\mathrm{d}t} x_t = \frac{1}{\varepsilon}\left(x_t - x_t^3 - y_t + s\right)$$
$$\frac{\mathrm{d}}{\mathrm{d}t} y_t = \gamma x_t - y_t + \beta$$

- x_t describes the membrane potential
- y_t is a recovery variable (sodium channels)
- s is a stimulus (current, voltage, chemical, ...)
- ε is a time scale and γ is a scaling factor
- β is a parameter from which the system switches between a stable fixed point and a limit cycle

The FitzHugh-Nagumo (FHN) model

- Modeling of neuronal spike generation in axons
- The FHN model is a prototype for more complicated models, like the Hodgkin-Huxley, or Morris-Lecar model

$$\frac{\mathrm{d}}{\mathrm{d}t} x_t = \frac{1}{\varepsilon}\left(x_t - x_t^3 - y_t + s\right)$$
$$\frac{\mathrm{d}}{\mathrm{d}t} y_t = \gamma x_t - y_t + \beta$$

- x describes the membrane potential
- y describes variables such as sodium channels
- s is a stimulus, e.g. a sudden injection of current
- ε is a time scale, and γ is the ratio between the time scales of x and y
- β determines the point of intersection between the two nullclines

The FitzHugh-Nagumo (FHN) model

- Modeling of neuronal spike generation in axons
- The FHN model is a prototype for more complicated models, like the Hodgkin-Huxley, or Morris-Lecar model

$$\frac{\mathrm{d}}{\mathrm{d}t} x_t = \frac{1}{\varepsilon}\left(x_t - x_t^3 - y_t + s\right)$$
$$\frac{\mathrm{d}}{\mathrm{d}t} y_t = \gamma x_t - y_t + \beta$$

- x describes the membrane potential
- y is a recovery variable modeling channel kinetics
- $\varepsilon > 0$ is a time scale separator, typically smaller than one
  - x is the fast, and y is the slow variable
- s denotes the input current
- γ and β determines the location of the fixed point(s)

The FitzHugh-Nagumo (FHN) model

- Modeling of neuronal spike generation in axons
- The FHN model is a prototype for more complicated models, like the Hodgkin-Huxley, or Morris-Lecar model

$$\frac{\mathrm{d}}{\mathrm{d}t}x_t = \frac{1}{\varepsilon}\left(x_t - x_t^3 - y_t + s\right)$$
$$\frac{\mathrm{d}}{\mathrm{d}t}y_t = \gamma x_t - y_t + \beta$$

- x describes the membrane potential
- y is a recovery variable modeling channel kinetics
- $\varepsilon > 0$ is a time scale separator, typically smaller than one - x is the fast, and y is the slow variable
- s denotes the input current
- γ and β determines the location of the fixed point(s)

The FitzHugh-Nagumo (FHN) model

- Modeling of neuronal spike generation in axons
- The FHN model is a prototype for more complicated models, like the Hodgkin-Huxley, or Morris-Lecar model

$$\frac{\mathrm{d}}{\mathrm{d}t}x_t = \frac{1}{\varepsilon}\left(x_t - x_t^3 - y_t + s\right)$$
$$\frac{\mathrm{d}}{\mathrm{d}t}y_t = \gamma x_t - y_t + \beta$$

- x describes the membrane potential
- y is a recovery variable modeling channel kinetics
- $\varepsilon > 0$ is a time scale separator, typically smaller than one - x is the fast, and y is the slow variable
- s denotes the input current
- γ and β determines the location of the fixed point(s)

The FitzHugh-Nagumo (FHN) model

- Modeling of neuronal spike generation in axons
- The FHN model is a prototype for more complicated models, like the Hodgkin-Huxley, or Morris-Lecar model

$$\frac{\mathrm{d}}{\mathrm{d}t}x_t = \frac{1}{\varepsilon}\left(x_t - x_t^3 - y_t + s\right)$$
$$\frac{\mathrm{d}}{\mathrm{d}t}y_t = \gamma x_t - y_t + \beta$$

- x describes the membrane potential
- y is a recovery variable modeling channel kinetics
- $\varepsilon > 0$ is a time scale separator, typically smaller than one - x is the fast, and y is the slow variable
- s denotes the input current
- γ and β determines the location of the fixed point(s)

The FitzHugh-Nagumo (FHN) model

- Modeling of neuronal spike generation in axons
- The FHN model is a prototype for more complicated models, like the Hodgkin-Huxley, or Morris-Lecar model

$$\frac{\mathrm{d}}{\mathrm{d}t}x_t = \frac{1}{\varepsilon}\left(x_t - x_t^3 - y_t + s\right)$$
$$\frac{\mathrm{d}}{\mathrm{d}t}y_t = \gamma x_t - y_t + \beta$$

- x describes the membrane potential
- y is a recovery variable modeling channel kinetics
- $\varepsilon > 0$ is a time scale separator, typically smaller than one - x is the fast, and y is the slow variable
- s denotes the input current
- γ and β determines the location of the fixed point(s)

The FitzHugh-Nagumo (FHN) model

- Modeling of neuronal spike generation in axons
- The FHN model is a prototype for more complicated models, like the Hodgkin-Huxley, or Morris-Lecar model

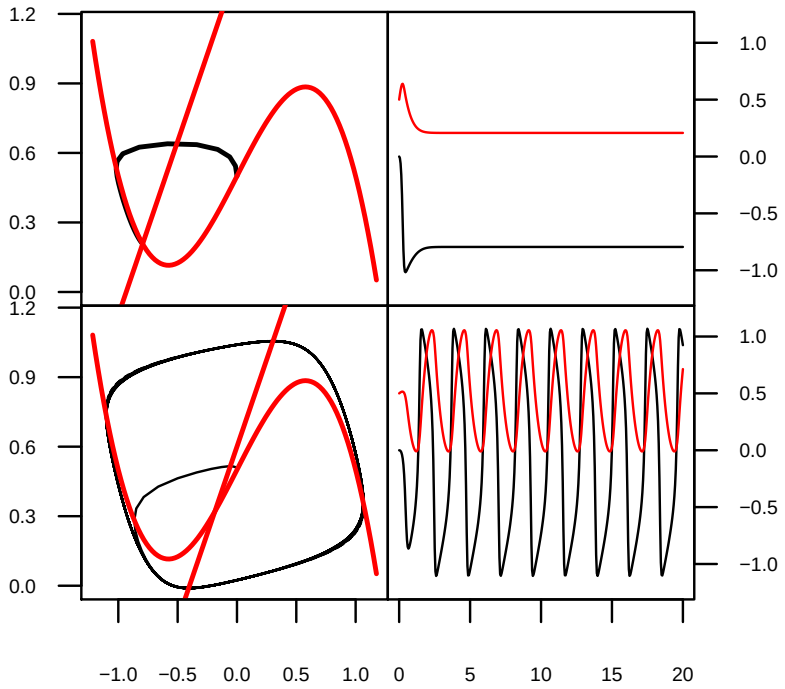
$$\frac{\mathrm{d}}{\mathrm{d}t}x_t = \frac{1}{\varepsilon}\left(x_t - x_t^3 - y_t + s\right)$$
$$\frac{\mathrm{d}}{\mathrm{d}t}y_t = \gamma x_t - y_t + \beta$$

- x describes the membrane potential
- y is a recovery variable modeling channel kinetics
- $\varepsilon > 0$ is a time scale separator, typically smaller than one - x is the fast, and y is the slow variable
- s denotes the input current
- γ and β determines the location of the fixed point(s)

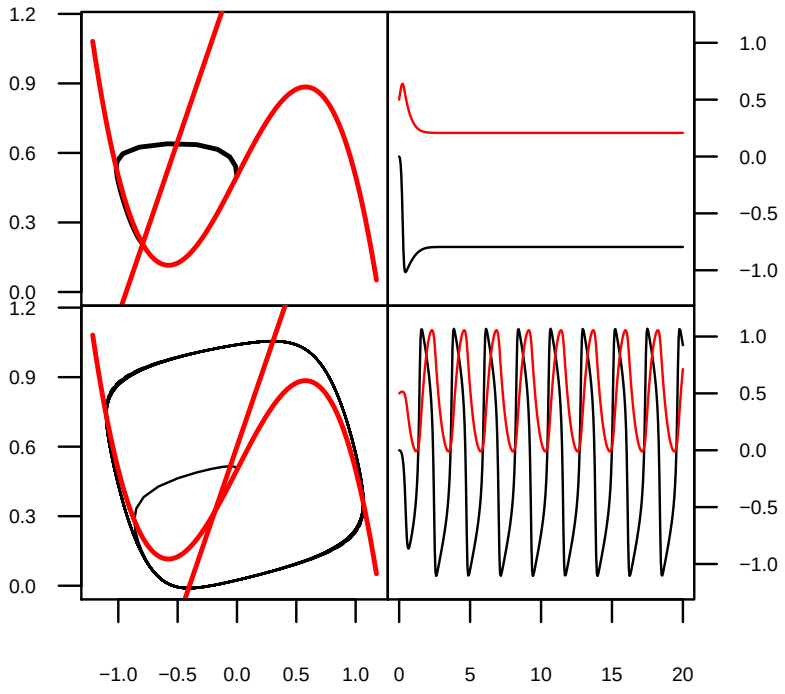
- Properties of fixed points determine the behavior of the model

- Properties of fixed points determine the behavior of the model
- Stable/unstable fixed points

- Properties of fixed points determine the behavior of the model
- Stable/unstable fixed points

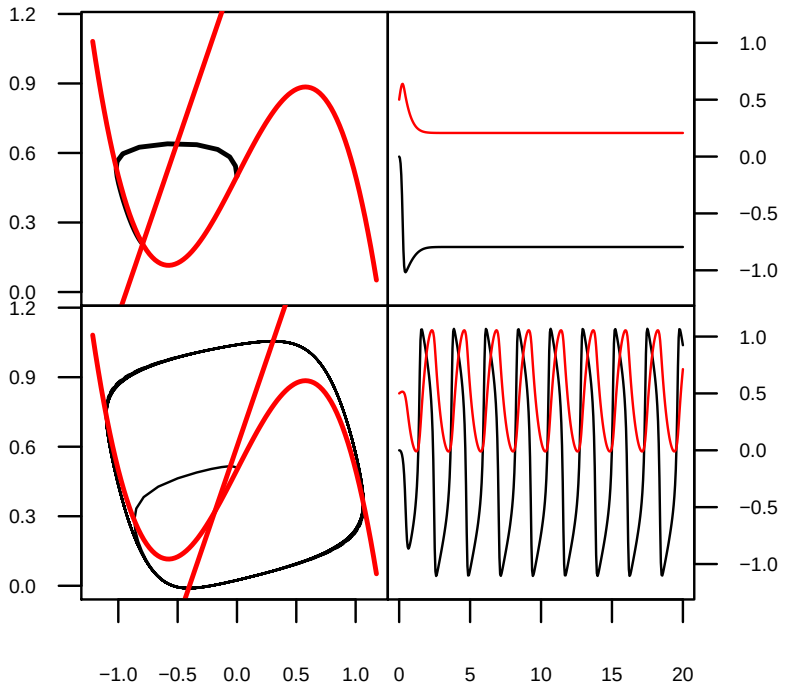


- Properties of fixed points determine the behavior of the model
- Stable/unstable fixed points



- Dynamics in the FHN model are well known

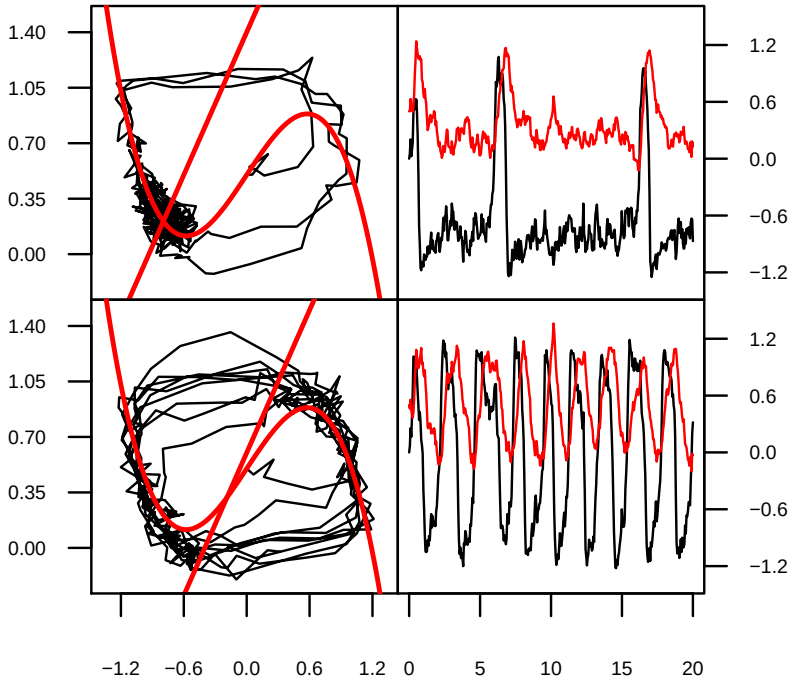
Noise terms account for various sources of noise: Random opening/closing of ion channels or noisy presynaptic currents



$$\mathrm{d}X_t = \frac{1}{\varepsilon}\left(X_t - X_t^3 - Y_t + s\right)\,\mathrm{d}t$$

$$\mathrm{d}Y_t = \left(\gamma X_t - Y_t + \beta\right)\,\mathrm{d}t$$

Noise terms account for various sources of noise: Random opening/closing of ion channels or noisy presynaptic currents



$$\mathrm{d}X_t = \frac{1}{\varepsilon}\left(X_t - X_t^3 - Y_t + s\right)\,\mathrm{d}t + \sigma_1\,\mathrm{d}B_t^{(1)}$$

$$\mathrm{d}Y_t = \left(\gamma X_t - Y_t + \beta\right)\,\mathrm{d}t + \sigma_2\,\mathrm{d}B_t^{(2)}$$

Setup

Consider a d-dimensional diffusion

$$\begin{aligned} \mathrm{d}V_t &= b(V_t, \theta)\,\mathrm{d}t + \Sigma\,\mathrm{d}B_t, \\ V_0 &= v_0 \quad 0 \leq t \leq T \end{aligned}$$

$$b : \mathbb{R}^d \times \mathbb{R}^p \mapsto \mathbb{R}^d, \quad \Sigma \in \mathcal{M}(d \times d)$$

B_t : A d dimensional Brownian motion

- Observations $D_n := \{V_0, V_1, \ldots, V_n\}$, with $V_i := V_{t_i}$ and $t_i - t_{i-1} = \Delta$
- Aim: Estimate the parameter θ based on D_n

Setup

Consider a d-dimensional diffusion

$$\begin{aligned} \mathrm{d}V_t &= b(V_t, \theta)\,\mathrm{d}t + \Sigma\,\mathrm{d}B_t, \\ V_0 &= v_0 \quad 0 \leq t \leq T \end{aligned}$$

$$b : \mathbb{R}^d \times \mathbb{R}^p \mapsto \mathbb{R}^d, \quad \Sigma \in \mathcal{M}(d \times d)$$

$$B_t : \text{A } d \text{ dimensional Brownian motion}$$

- Observations $D_n := \{V_0, V_1, \ldots, V_n\}$, with $V_i := V_{t_i}$ and $t_i - t_{i-1} = \Delta$
- Initially we assume Σ known
- Aim: Bayesian inference for $\theta \in \mathbb{R}^p$ conditional on D_n

Setup

Consider a d-dimensional diffusion

$$\begin{aligned} \mathrm{d}V_t &= b(V_t, \theta)\,\mathrm{d}t + \Sigma\,\mathrm{d}B_t, \\ V_0 &= v_0 \quad 0 \leq t \leq T \end{aligned}$$

$$b : \mathbb{R}^d \times \mathbb{R}^p \mapsto \mathbb{R}^d, \quad \Sigma \in \mathcal{M}(d \times d)$$

$$B_t : \text{A } d \text{ dimensional Brownian motion}$$

- Observations $D_n := \{V_0, V_1, \ldots, V_n\}$, with $V_i := V_{t_i}$ and $t_i - t_{i-1} = \Delta$
- Initially we assume Σ known
- Aim: Bayesian inference for $\theta \in \mathbb{R}^p$ conditional on D_n

Setup

Consider a d-dimensional diffusion

$$\begin{aligned} \mathrm{d}V_t &= b(V_t, \theta)\,\mathrm{d}t + \Sigma\,\mathrm{d}B_t, \\ V_0 &= v_0 \quad 0 \leq t \leq T \end{aligned}$$

$$b : \mathbb{R}^d \times \mathbb{R}^p \mapsto \mathbb{R}^d, \quad \Sigma \in \mathcal{M}(d \times d)$$

$$B_t : \text{A } d \text{ dimensional Brownian motion}$$

- Observations $D_n := \{V_0, V_1, \ldots, V_n\}$, with $V_i := V_{t_i}$ and $t_i - t_{i-1} = \Delta$
- Initially we assume Σ known
- Aim: Bayesian inference for $\theta \in \mathbb{R}^p$ conditional on D_n

Setup

Consider a d-dimensional diffusion

$$\begin{aligned} \mathrm{d}V_t &= b(V_t, \theta)\,\mathrm{d}t + \Sigma\,\mathrm{d}B_t, \\ V_0 &= v_0 \quad 0 \leq t \leq T \end{aligned}$$

$$b : \mathbb{R}^d \times \mathbb{R}^p \mapsto \mathbb{R}^d, \quad \Sigma \in \mathcal{M}(d \times d)$$

$$B_t : \text{A } d \text{ dimensional Brownian motion}$$

- Observations $D_n := \{V_0, V_1, \ldots, V_n\}$, with $V_i := V_{t_i}$ and $t_i - t_{i-1} = \Delta$
- Initially we assume Σ known
- Aim: Bayesian inference for $\theta \in \mathbb{R}^p$ conditional on D_n

The posterior distribution

$$
\begin{aligned}
p(\theta \mid D_n) &= \frac{p(\theta)p(D_n \mid \theta)}{\int p(D_n \mid \theta)\,\mathrm{d}\theta} \\
&\propto_\theta p(\theta)p(D_n \mid \theta) \\
&= p(\theta)\prod_{i=1}^{n} p(V_i \mid V_{i-1}, \theta)
\end{aligned}
$$

- $p(\theta)$ is chosen from a priori knowledge, but $p(V_i \mid V_{i-1}, \theta)$ is typically unknown
- Quick and easy solution: Approximate $p(V_i \mid V_{i-1}, \theta)$, using Euler-Maruyama:

$$V_{t+1} \approx V_t + b(t, V_t)\Delta + \sigma(t, V_t)\sqrt{\Delta}W_j, \quad W_j \sim \mathcal{N}(0,1)$$

- Inaccurate if Δ is too large...

The posterior distribution

$$
\begin{aligned}
p(\theta \mid D_n) &= \frac{p(\theta)p(D_n \mid \theta)}{\int p(D_n \mid \theta)\,\mathrm{d}\theta} \\
&\propto_\theta p(\theta)p(D_n \mid \theta) \\
&= p(\theta)\prod_{i=1}^{n} p(V_i \mid V_{i-1}, \theta)
\end{aligned}
$$

- $p(\theta)$ is chosen from a priori knowledge, but $p(V_i \mid V_{i-1}, \theta)$ is typically unknown
- Quick and easy solution: Approximate $p(V_i \mid V_{i-1}, \theta)$, using Euler-Maruyama:

$$
V_{t+1} \approx V_t + b(t, V_t)\Delta + \sigma(t, V_t)\sqrt{\Delta}W_j, \quad W_j \sim \mathcal{N}(0,1)
$$

- Inaccurate if Δ is too large...

The posterior distribution

$$
\begin{aligned}
p(\theta \mid D_n) &= \frac{p(\theta)p(D_n \mid \theta)}{\int p(D_n \mid \theta)\,\mathrm{d}\theta} \\
&\propto_\theta p(\theta)p(D_n \mid \theta) \\
&= p(\theta)\prod_{i=1}^{n} p(V_i \mid V_{i-1}, \theta)
\end{aligned}
$$

- $p(\theta)$ is chosen from a priori knowledge, but $p(V_i \mid V_{i-1}, \theta)$ is typically unknown
- Quick and easy solution: Approximate $p(V_i \mid V_{i-1}, \theta)$, using Euler-Maruyama:

$$V_{t+1} \approx V_t + b(t, V_t)\Delta + \sigma(t, V_t)\sqrt{\Delta}W_j, \quad W_j \sim \mathcal{N}(0,1)$$

- Inaccurate if Δ is too large...

The posterior distribution

$$
\begin{aligned}
p(\theta \mid D_n) &= \frac{p(\theta)p(D_n \mid \theta)}{\int p(D_n \mid \theta)\,\mathrm{d}\theta} \\
&\propto_\theta p(\theta)p(D_n \mid \theta) \\
&= p(\theta)\prod_{i=1}^{n} p(V_i \mid V_{i-1}, \theta)
\end{aligned}
$$

- $p(\theta)$ is chosen from a priori knowledge, but $p(V_i \mid V_{i-1}, \theta)$ is typically unknown
- Quick and easy solution: Approximate $p(V_i \mid V_{i-1}, \theta)$, using Euler-Maruyama:

$$V_{t+1} \approx V_t + b(t, V_t)\Delta + \sigma(t, V_t)\sqrt{\Delta}W_j, \quad W_j \sim \mathcal{N}(0,1)$$

- Inaccurate if Δ is too large...

The posterior distribution

$$
\begin{aligned}
p(\theta \mid D_n) &= \frac{p(\theta)p(D_n \mid \theta)}{\int p(D_n \mid \theta)\,\mathrm{d}\theta} \\
&\propto_\theta p(\theta)p(D_n \mid \theta) \\
&= p(\theta)\prod_{i=1}^{n} p(V_i \mid V_{i-1}, \theta)
\end{aligned}
$$

- $p(\theta)$ is chosen from a priori knowledge, but $p(V_i \mid V_{i-1}, \theta)$ is typically unknown
- Quick and easy solution: Approximate $p(V_i \mid V_{i-1}, \theta)$, using Euler-Maruyama:

$$
V_{t+1} \approx V_t + b(t, V_t)\Delta + \sigma(t, V_t)\sqrt{\Delta}W_j, \quad W_j \sim \mathcal{N}(0,1)
$$

- Inaccurate if Δ is too large...

The posterior distribution

$$
\begin{aligned}
p(\theta \mid D_n) &= \frac{p(\theta)p(D_n \mid \theta)}{\int p(D_n \mid \theta)\,\mathrm{d}\theta} \\
&\propto_\theta p(\theta)p(D_n \mid \theta) \\
&= p(\theta)\prod_{i=1}^{n} p(V_i \mid V_{i-1}, \theta)
\end{aligned}
$$

- $p(\theta)$ is chosen from a priori knowledge, but $p(V_i \mid V_{i-1}, \theta)$ is typically unknown
- Quick and easy solution: Approximate $p(V_i \mid V_{i-1}, \theta)$, using Euler-Maruyama:

$$V_{t+1} \approx V_t + b(t, V_t)\Delta + \sigma(t, V_t)\sqrt{\Delta}W_j, \quad W_j \sim \mathcal{N}(0,1)$$

- Inaccurate if Δ is too large...

Outline of estimation procedure

- Data augmentation and Gibbs sampling
  - Let $\bar{V}$ denote the latent path of V and apply Gibbs sampler on $(\theta, \bar{V})$ conditional on D_n
  - Alternatively
  - $\bar{V}$ is sampled using an independent Metropolis-Hastings sampler using Brownian bridge proposal
- Note: Imputed paths $\bar{V}$ are in practise finite dimensional
- Conditioning on D_n, means simulation of diffusion bridges

Outline of estimation procedure

- Data augmentation and Gibbs sampling
  - Let $\bar{V}$ denote the latent path of V and apply Gibbs sampler on $(\theta, \bar{V})$ conditional on D_n
  - θ is sampled directly
  - $\bar{V}$ is sampled using an independent Metropolis-Hastings step using Brownian bridge proposals
- Note: Imputed paths $\bar{V}$ are in practise finite dimensional
- Conditioning on D_n, means simulation of diffusion bridges

Outline of estimation procedure

- Data augmentation and Gibbs sampling
  - Let $\bar{V}$ denote the latent path of V and apply Gibbs sampler on $(\theta, \bar{V})$ conditional on D_n
  - θ is sampled directly
  - $\bar{V}$ is sampled using an independent Metropolis-Hastings step using Brownian bridge proposals
- Note: Imputed paths $\bar{V}$ are in practise finite dimensional
- Conditioning on D_n, means simulation of diffusion bridges

Outline of estimation procedure

- Data augmentation and Gibbs sampling
  - Let $\bar{V}$ denote the latent path of V and apply Gibbs sampler on $(\theta, \bar{V})$ conditional on D_n
  - θ is sampled directly
  - $\bar{V}$ is sampled using an independent Metropolis-Hastings step using Brownian bridge proposals
- Note: Imputed paths $\bar{V}$ are in practise finite dimensional
- Conditioning on D_n, means simulation of diffusion bridges

Outline of estimation procedure

- Data augmentation and Gibbs sampling
  - Let $\bar{V}$ denote the latent path of V and apply Gibbs sampler on $(\theta, \bar{V})$ conditional on D_n
  - θ is sampled directly
  - $\bar{V}$ is sampled using an independent Metropolis-Hastings step using Brownian bridge proposals
- Note: Imputed paths $\bar{V}$ are in practise finite dimensional
- Conditioning on D_n, means simulation of diffusion bridges

Outline of estimation procedure

- Data augmentation and Gibbs sampling
  - Let $\bar{V}$ denote the latent path of V and apply Gibbs sampler on $(\theta, \bar{V})$ conditional on D_n
  - θ is sampled directly
  - $\bar{V}$ is sampled using an independent Metropolis-Hastings step using Brownian bridge proposals
- Note: Imputed paths $\bar{V}$ are in practise finite dimensional
- Conditioning on D_n, means simulation of diffusion bridges

Outline of estimation procedure

- Data augmentation and Gibbs sampling
  - Let $\bar{V}$ denote the latent path of V and apply Gibbs sampler on $(\theta, \bar{V})$ conditional on D_n
  - θ is sampled directly
  - $\bar{V}$ is sampled using an independent Metropolis-Hastings step using Brownian bridge proposals
- Note: Imputed paths $\bar{V}$ are in practise finite dimensional
- Conditioning on D_n, means simulation of diffusion bridges

The Gibbs sampler

- Let $\bar{V}$ denote the latent path of V

Initialize

1) Initialize $\theta^{(0)}$ and imputed data $\bar{V}^{(0)}$

Iterate

At iteration k:

2a) Sample $\bar{V}^{(k)}$ from $p(V \mid \theta^{(k-1)}, D_n)$

2b) Sample $\theta^{(k)}$ from $p(\theta \mid \bar{V}^{(k)}, D_n)$

- Alternate between sampling $\bar{V}$ and θ from the full conditionals $p(\bar{V} \mid \theta, D_n)$ and $p(\theta \mid \bar{V}, D_n)$
- Always one-to-one between $\bar{V}$ and D_n

The Gibbs sampler

- Let $\bar{V}$ denote the latent path of V

Initialize

1) Initialize $\theta^{(0)}$ and imputed data $\bar{V}^{(0)}$

Iterate

At iteration k:

2a) Sample $\bar{V}^{(k)}$ from $p(V \mid \theta^{(k-1)}, D_n)$

2b) Sample $\theta^{(k)}$ from $p(\theta \mid \bar{V}^{(k)}, D_n)$

- Always one-step-ahead $p(\bar{V}_t \mid \bar{V}_{t-1}, \theta, D_n)$

The Gibbs sampler

- Let $\bar{V}$ denote the latent path of V

Initialize

1) Initialize $\theta^{(0)}$ and imputed data $\bar{V}^{(0)}$

Iterate

At iteration k:

2a) Sample $\bar{V}^{(k)}$ from $p(V \mid \theta^{(k-1)}, D_n)$

2b) Sample $\theta^{(k)}$ from $p(\theta \mid \bar{V}^{(k)}, D_n)$

- After burn in period, $(\theta^{(k)}, \bar{V}^{(k)})_k$ resembles draws from $p(\theta, \bar{V} \mid D_n)$
- Always one-to-one transition $p(\theta \mid \bar{V}, D_n)$

The Gibbs sampler

- Let $\bar{V}$ denote the latent path of V

Initialize

1) Initialize $\theta^{(0)}$ and imputed data $\bar{V}^{(0)}$

Iterate

At iteration k:

2a) Sample $\bar{V}^{(k)}$ from $p(V \mid \theta^{(k-1)}, D_n)$

2b) Sample $\theta^{(k)}$ from $p(\theta \mid \bar{V}^{(k)}, D_n)$

- After burn in period, $(\theta^{(k)}, \bar{V}^{(k)})_k$ resembles draws from $p(\theta, \bar{V} \mid D_n)$
- Allows one to infer about $p(\theta \mid D_n)$

The Gibbs sampler

- Let $\bar{V}$ denote the latent path of V

Initialize

1) Initialize $\theta^{(0)}$ and imputed data $\bar{V}^{(0)}$

Iterate

At iteration k:

2a) Sample $\bar{V}^{(k)}$ from $p(V \mid \theta^{(k-1)}, D_n)$

2b) Sample $\theta^{(k)}$ from $p(\theta \mid \bar{V}^{(k)}, D_n)$

- After burn in period, $(\theta^{(k)}, \bar{V}^{(k)})_k$ resembles draws from $p(\theta, \bar{V} \mid D_n)$
- Allows one to infer about $p(\theta \mid D_n)$

The Gibbs sampler

- Let $\bar{V}$ denote the latent path of V

Initialize

1) Initialize $\theta^{(0)}$ and imputed data $\bar{V}^{(0)}$

Iterate

At iteration k:

2a) Sample $\bar{V}^{(k)}$ from $p(V \mid \theta^{(k-1)}, D_n)$

2b) Sample $\theta^{(k)}$ from $p(\theta \mid \bar{V}^{(k)}, D_n)$

- After burn in period, $(\theta^{(k)}, \bar{V}^{(k)})_k$ resembles draws from $p(\theta, \bar{V} \mid D_n)$
- Allows one to infer about $p(\theta \mid D_n)$

The Gibbs sampler

- Let $\bar{V}$ denote the latent path of V

Initialize

1) Initialize $\theta^{(0)}$ and imputed data $\bar{V}^{(0)}$

Iterate

At iteration k:

2a) Sample $\bar{V}^{(k)}$ from $p(V \mid \theta^{(k-1)}, D_n)$

2b) Sample $\theta^{(k)}$ from $p(\theta \mid \bar{V}^{(k)}, D_n)$

- After burn in period, $(\theta^{(k)}, \bar{V}^{(k)})_k$ resembles draws from $p(\theta, \bar{V} \mid D_n)$
- Allows one to infer about $p(\theta \mid D_n)$

Sampling from $p(\bar{V} \mid \theta, D_n)$

- Due to the Markov property

$$p(\bar{V} \mid \theta, D_n) = \prod_{i=1}^{n} p(\bar{V}_i \mid V_{t_{i-1}}, V_{t_i}, \theta),$$

  where $\bar{V}_i$ denotes the continuous path V on (t_{i-1}, t_i).
- Paths $\bar{V}_i$ may be sampled independently
- Sampling directly from the distribution of a diffusion bridge is in general not easy
- Idea: Identify the Radon-Nikodym derivative of the desired distribution with respect to another distribution from which sampling is feasible
- Without loss of generality, focus on a single term of the form $p(V|V_0, V_T, \theta)$

Sampling from $p(\bar{V} \mid \theta, D_n)$

- Due to the Markov property

$$p(\bar{V} \mid \theta, D_n) = \prod_{i=1}^{n} p(\bar{V}_i \mid V_{t_{i-1}}, V_{t_i}, \theta),$$

  where $\bar{V}_i$ denotes the continuous path V on (t_{i-1}, t_i).
- Paths $\bar{V}_i$ may be sampled independently
- Sampling directly from the distribution of a diffusion bridge is in general not easy
- Idea: Identify the Radon-Nikodym derivative of the desired distribution with respect to another distribution from which sampling is feasible
- Without loss of generality, focus on a single term of the form $p(V|V_0, V_T, \theta)$

Sampling from $p(\bar{V} \mid \theta, D_n)$

- Due to the Markov property

$$p(\bar{V} \mid \theta, D_n) = \prod_{i=1}^{n} p(\bar{V}_i \mid V_{t_{i-1}}, V_{t_i}, \theta),$$

  where $\bar{V}_i$ denotes the continuous path V on (t_{i-1}, t_i).
- Paths $\bar{V}_i$ may be sampled independently
- Sampling directly from the distribution of a diffusion bridge is in general not easy
- Idea: Identify the Radon-Nikodym derivative of the desired distribution with respect to another distribution from which sampling is feasible
- Without loss of generality, focus on a single term of the form $p(V|V_0, V_T, \theta)$

Sampling from $p(\bar{V} \mid \theta, D_n)$

- Due to the Markov property

$$p(\bar{V} \mid \theta, D_n) = \prod_{i=1}^{n} p(\bar{V}_i \mid V_{t_{i-1}}, V_{t_i}, \theta),$$

  where $\bar{V}_i$ denotes the continuous path V on (t_{i-1}, t_i).
- Paths $\bar{V}_i$ may be sampled independently
- Sampling directly from the distribution of a diffusion bridge is in general not easy
- Idea: Identify the Radon-Nikodym derivative of the desired distribution with respect to another distribution from which sampling is feasible
- Without loss of generality, focus on a single term of the form $p(V|V_0, V_T, \theta)$

Sampling from $p(\bar{V} \mid \theta, D_n)$

- Due to the Markov property

$$p(\bar{V} \mid \theta, D_n) = \prod_{i=1}^{n} p(\bar{V}_i \mid V_{t_{i-1}}, V_{t_i}, \theta),$$

  where $\bar{V}_i$ denotes the continuous path V on (t_{i-1}, t_i).
- Paths $\bar{V}_i$ may be sampled independently
- Sampling directly from the distribution of a diffusion bridge is in general not easy
- Idea: Identify the Radon-Nikodym derivative of the desired distribution with respect to another distribution from which sampling is feasible
- Without loss of generality, focus on a single term of the form $p(V|V_0, V_T, \theta)$

Sampling from $p(\bar{V} \mid \theta, D_n)$

- Due to the Markov property

$$p(\bar{V} \mid \theta, D_n) = \prod_{i=1}^{n} p(\bar{V}_i \mid V_{t_{i-1}}, V_{t_i}, \theta),$$

  where $\bar{V}_i$ denotes the continuous path V on (t_{i-1}, t_i).
- Paths $\bar{V}_i$ may be sampled independently
- Sampling directly from the distribution of a diffusion bridge is in general not easy
- Idea: Identify the Radon-Nikodym derivative of the desired distribution with respect to another distribution from which sampling is feasible
- Without loss of generality, focus on a single term of the form $p(V|V_0, V_T, \theta)$

Consider the following processes and distributions on the interval $[0, T]$:

$$P_b : \ \mathrm{d}V_t = b(V_t, \theta)\,\mathrm{d}t + \Sigma\,\mathrm{d}B_t,$$
$$P_0 : \ \mathrm{d}V_t = \Sigma\,\mathrm{d}B_t, \quad V_0 = v_0$$

Consider the following processes and distributions on the interval $[0, T]$:

$$P_b : \ \mathrm{d}V_t = b(V_t, \theta)\,\mathrm{d}t + \Sigma\,\mathrm{d}B_t,$$
$$P_0 : \ \mathrm{d}V_t = \Sigma\,\mathrm{d}B_t, \quad V_0 = v_0$$

$$P_b^* : \ \mathrm{d}V_t = b(V_t, \theta)\,\mathrm{d}t + \Sigma\,\mathrm{d}B_t,$$
$$P_0^* : \ \mathrm{d}V_t = \Sigma\,\mathrm{d}B_t, \quad V_0 = v_0, V_T = v_T$$

Consider the following processes and distributions on the interval $[0, T]$:

$$P_b : \ \mathrm{d}V_t = b(V_t, \theta)\,\mathrm{d}t + \Sigma\,\mathrm{d}B_t,$$
$$P_0 : \ \mathrm{d}V_t = \Sigma\,\mathrm{d}B_t, \quad V_0 = v_0$$

$$P_b^* : \ \mathrm{d}V_t = b(V_t, \theta)\,\mathrm{d}t + \Sigma\,\mathrm{d}B_t,$$
$$P_0^* : \ \mathrm{d}V_t = \Sigma\,\mathrm{d}B_t, \quad V_0 = v_0, V_T = v_T$$

$$\begin{aligned}
\frac{\mathrm{d}P_b^*}{\mathrm{d}P_0^*}(V \mid \theta) &= \frac{\varphi(V_T \mid V_0; \Sigma\Sigma^T T)}{p_{0,T}(V_T \mid V_0, \theta)} \frac{\mathrm{d}P_b(V \mid \theta)}{\mathrm{d}P_0(V \mid \theta)} \\
&\propto_V \frac{\mathrm{d}P_b(V \mid \theta)}{\mathrm{d}P_0(V \mid \theta)} \\
&= e^{\int_0^T b(V_u,\theta)^T \Gamma^{-1}\,\mathrm{d}V_u - \frac{1}{2}\int_0^T b(V_u,\theta)^T \Gamma^{-1} b(V_u,\theta)\,\mathrm{d}u} \\
&:= \phi(V, \theta), \qquad (\Gamma = \Sigma\Sigma^T, \text{ invertible})
\end{aligned}$$

The algorithm for simulating the bridge $\bar{V}$ under P_b^*

Initialize

The algorithm for simulating the bridge $\bar{V}$ under P_b^*

Initialize

1) Initialize a skeleton path $(\bar{V}^M)_0$ according to P_0^*, and compute an approximation of the Radon-Nikodym derivative $\phi((\bar{V}^M)_0, \theta)$, w_0

Iterate

The algorithm for simulating the bridge $\bar{V}$ under P_b^*

Initialize

1) Initialize a skeleton path $(\bar{V}^M)_0$ according to P_0^*, and compute an approximation of the Radon-Nikodym derivative $\phi((\bar{V}^M)_0, \theta)$, w_0

Iterate

2) Generate a proposal skeleton path, $\tilde{V}^M$ according to the distribution P_0^* and compute an approximation of the Radon-Nikodym derivative $\phi(\tilde{V}^M, \theta)$, $\tilde{w}$

3) Let $(\bar{V}^M)_{k+1} = \begin{cases} \tilde{V}^M & \text{with prob. } \min\left(1, \frac{\tilde{w}}{w_k}\right) \\ (\bar{V}^M)_k & \text{otherwise} \end{cases}$.

The Gibbs sampler

- Let $\bar{V}$ denote the latent path of V

Initialize

1) Initialize $\theta^{(0)}$ and imputed data $\bar{V}^{(0)}$

Iterate

At iteration k:

2a) Sample $\bar{V}^{(k)}$ from $p(V \mid \theta^{(k-1)}, D_n)$

2b) Sample $\theta^{(k)}$ from $p(\theta \mid \bar{V}^{(k)}, D_n)$

- After burn in period, $(\theta^{(k)}, \bar{V}^{(k)})_k$ resembles draws from $p(\theta, \bar{V} \mid D_n)$
- Allows one to infer about $p(\theta \mid D_n)$

The Gibbs sampler

- Let $\bar{V}$ denote the latent path of V

Initialize

1) Initialize $\theta^{(0)}$ and imputed data $\bar{V}^{(0)}$

Iterate

At iteration k:

2a) Sample $\bar{V}^{(k)}$ from $p(V \mid \theta^{(k-1)}, D_n)$

2b) Sample $\theta^{(k)}$ from $p(\theta \mid \bar{V}^{(k)}, D_n)$

- After burn in period, $(\theta^{(k)}, \bar{V}^{(k)})_k$ resembles draws from $p(\theta, \bar{V} \mid D_n)$
- Allows one to infer about $p(\theta \mid D_n)$

Sampling from $p(\theta \mid \bar{V}, D_n)$

$$p(\theta \mid \bar{V}, D_n) \propto_\theta p(\theta)\phi(\bar{V}, \theta)$$

- $p(\theta \mid V, D_n)$ may be sampled directly if the drift is an affine transformation of θ:

$$b(V_t, \theta) = f_0(V_t) + \sum_{i=1}^{p} \theta_i f_i(V_t),$$

  where f_i is a $d \times 1$ vector
- $\phi(V, \theta)$ is exponentially quadratic in θ
- A Gaussian prior for θ yields a Gaussian posterior

Sampling from $p(\theta \mid \bar{V}, D_n)$

$$p(\theta \mid \bar{V}, D_n) \propto_\theta p(\theta)\phi(\bar{V}, \theta)$$

- $p(\theta \mid V, D_n)$ may be sampled directly if the drift is an affine transformation of θ:

$$b(V_t, \theta) = f_0(V_t) + \sum_{i=1}^{p} \theta_i f_i(V_t),$$

  where f_i is a $d \times 1$ vector
- $\phi(V, \theta)$ is exponentially quadratic in θ
- A Gaussian prior for θ yields a Gaussian posterior

Sampling from $p(\theta \mid \bar{V}, D_n)$

$$p(\theta \mid \bar{V}, D_n) \propto_\theta p(\theta)\phi(\bar{V}, \theta)$$

- $p(\theta \mid V, D_n)$ may be sampled directly if the drift is an affine transformation of θ:

$$b(V_t, \theta) = f_0(V_t) + \sum_{i=1}^{p} \theta_i f_i(V_t),$$

  where f_i is a $d \times 1$ vector
- $\phi(V, \theta)$ is exponentially quadratic in θ
- A Gaussian prior for θ yields a Gaussian posterior

Identifying the Gaussian posterior

- Assume $p(\theta) = \varphi(\theta \mid \mu; \Phi)$

Identifying the Gaussian posterior

- Assume $p(\theta) = \varphi(\theta \mid \mu; \Phi)$
- $p(\theta \mid \bar{V}, D_n) \sim \varphi(\theta \mid (R + \Phi^{-1})^{-1}(F + \Phi^{-1}\mu); (R + \Phi^{-1})^{-1})$, with

$$
\begin{aligned}
R_{ij} &= \int_0^T f_i(V_u)^T \Gamma^{-1} f_j(V_u) \, \mathrm{d}u, \\
I_i &= \int_0^T f_i(V_u)^T \Gamma^{-1} \, \mathrm{d}V_u, \\
F_i &= I_i - R_{i0}
\end{aligned}
$$

Identifying the Gaussian posterior

- Assume $p(\theta) = \varphi(\theta \mid \mu; \Phi)$
- $p(\theta \mid \bar{V}, D_n) \sim \varphi(\theta \mid (R + \Phi^{-1})^{-1}(F + \Phi^{-1}\mu); (R + \Phi^{-1})^{-1})$, with

$$
\begin{aligned}
R_{ij} &= \int_0^T f_i(V_u)^T \Gamma^{-1} f_j(V_u) \, \mathrm{d}u, \\
I_i &= \int_0^T f_i(V_u)^T \Gamma^{-1} \, \mathrm{d}V_u, \\
F_i &= I_i - R_{i0}
\end{aligned}
$$

- No tuning of parameters is required!

Sampling from $p(\theta \mid \bar{V}, D_n)$ - MH step

$$p(\theta \mid \bar{V}, D_n) \propto_\theta p(\theta) p(\bar{V}, D_n \mid \theta)$$

- $p(\theta \mid \bar{V}, D_n)$ may be sampled using an MH-step
- Project path $\bar{V}$ onto discrete subset
- Use Euler-Maruyama on small time scale
- Approximate samples from $p(\theta \mid \bar{V}, D_n)$ using Metropolis-Hastings algorithm

Sampling from $p(\theta \mid \bar{V}, D_n)$ - MH step

$$p(\theta \mid \bar{V}, D_n) \propto_\theta p(\theta) p(\bar{V}, D_n \mid \theta)$$

- $p(\theta \mid \bar{V}, D_n)$ may be sampled using an MH-step
- Project path $\bar{V}$ onto discrete subset
- Use Euler-Maruyama on small time scale
- Approximate samples from $p(\theta \mid \bar{V}, D_n)$ using Metropolis-Hastings algorithm

Sampling from $p(\theta \mid \bar{V}, D_n)$ - MH step

$$p(\theta \mid \bar{V}, D_n) \propto_\theta p(\theta) p(\bar{V}, D_n \mid \theta)$$

- $p(\theta \mid \bar{V}, D_n)$ may be sampled using an MH-step
- Project path $\bar{V}$ onto discrete subset
- Use Euler-Maruyama on small time scale
- Approximate samples from $p(\theta \mid \bar{V}, D_n)$ using Metropolis-Hastings algorithm

Sampling from $p(\theta \mid \bar{V}, D_n)$ - MH step

$$p(\theta \mid \bar{V}, D_n) \propto_\theta p(\theta) p(\bar{V}, D_n \mid \theta)$$

- $p(\theta \mid \bar{V}, D_n)$ may be sampled using an MH-step
- Project path $\bar{V}$ onto discrete subset
- Use Euler-Maruyama on small time scale
- Approximate samples from $p(\theta \mid \bar{V}, D_n)$ using Metropolis-Hastings algorithm

Unknown diffusion

- Drop assumption about Σ known. Instead, assume $\Sigma(\sigma)$, invertible
- For any $t \in [0, T]$

$$\lim_{M\to\infty} \sum_{i=1}^{M} (V_{ti/M} - V_{t(i-1)/M})(V_{ti/M} - V_{t(i-1)/M})^T$$
$$= t\Sigma\Sigma^T(\sigma) \text{ in probability}$$

- Gibbs sampler is reducible when sampling σ conditional on the path
- Solution: Apply the one-to-one transformation $x \mapsto \Sigma^{-1}(\sigma)x$ to V, to obtain

$$\mathrm{d}Z_t = \alpha(Z_t, \theta, \sigma)\,\mathrm{d}t + \mathrm{d}B_t, \quad Z_0 = \Sigma^{-1}(\sigma)V_0,$$

where $\alpha(Z_t, \theta, \sigma) = \Sigma^{-1}(\sigma)b(\Sigma(\sigma)Z_t, \theta)$.

Unknown diffusion

- Drop assumption about Σ known. Instead, assume $\Sigma(\sigma)$, invertible
- For any $t \in [0, T]$

$$\lim_{M\to\infty} \sum_{i=1}^{M} (V_{ti/M} - V_{t(i-1)/M})(V_{ti/M} - V_{t(i-1)/M})^T$$
$$= t\Sigma\Sigma^T(\sigma) \text{ in probability}$$

- Gibbs sampler is reducible when sampling σ conditional on the path
- Solution: Apply the one-to-one transformation $x \mapsto \Sigma^{-1}(\sigma)x$ to V, to obtain

$$\mathrm{d}Z_t = \alpha(Z_t, \theta, \sigma)\,\mathrm{d}t + \mathrm{d}B_t, \quad Z_0 = \Sigma^{-1}(\sigma)V_0,$$

where $\alpha(Z_t, \theta, \sigma) = \Sigma^{-1}(\sigma)b(\Sigma(\sigma)Z_t, \theta)$.

Unknown diffusion

- Drop assumption about Σ known. Instead, assume $\Sigma(\sigma)$, invertible
- For any $t \in [0, T]$

$$\lim_{M\to\infty} \sum_{i=1}^{M} (V_{ti/M} - V_{t(i-1)/M})(V_{ti/M} - V_{t(i-1)/M})^T$$
$$= t\Sigma\Sigma^T(\sigma) \text{ in probability}$$

- Gibbs sampler is reducible when sampling σ conditional on the path
- Solution: Apply the one-to-one transformation $x \mapsto \Sigma^{-1}(\sigma)x$ to V, to obtain

$$\mathrm{d}Z_t = \alpha(Z_t, \theta, \sigma)\,\mathrm{d}t + \mathrm{d}B_t, \quad Z_0 = \Sigma^{-1}(\sigma)V_0,$$

where $\alpha(Z_t, \theta, \sigma) = \Sigma^{-1}(\sigma)b(\Sigma(\sigma)Z_t, \theta)$.

Unknown diffusion

- Drop assumption about Σ known. Instead, assume $\Sigma(\sigma)$, invertible
- For any $t \in [0, T]$

$$\lim_{M\to\infty} \sum_{i=1}^{M} (V_{ti/M} - V_{t(i-1)/M})(V_{ti/M} - V_{t(i-1)/M})^T$$
$$= t\Sigma\Sigma^T(\sigma) \text{ in probability}$$

- Gibbs sampler is reducible when sampling σ conditional on the path
- Solution: Apply the one-to-one transformation $x \mapsto \Sigma^{-1}(\sigma)x$ to V, to obtain

$$\mathrm{d}Z_t = \alpha(Z_t, \theta, \sigma)\,\mathrm{d}t + \mathrm{d}B_t, \quad Z_0 = \Sigma^{-1}(\sigma)V_0,$$

where $\alpha(Z_t, \theta, \sigma) = \Sigma^{-1}(\sigma)b(\Sigma(\sigma)Z_t, \theta)$.

Imputation of latent data

- Focus on single term, say $p(\bar{V} \mid V_0, V_T, \theta, \sigma)$
- Sampling $\bar{V}$ is equivalent to sampling Z conditionally on $Z_0 = \Sigma^{-1}(\sigma)V_0$ and $Z_T = \Sigma^{-1}(\sigma)V_T$
- The quadratic variation of Z is now independent of σ
- Unfortunately, still perfect dependence between Z and σ via the endpoints of Z

Imputation of latent data

- Focus on single term, say $p(\bar{V} \mid V_0, V_T, \theta, \sigma)$
- Sampling $\bar{V}$ is equivalent to sampling Z conditionally on $Z_0 = \Sigma^{-1}(\sigma)V_0$ and $Z_T = \Sigma^{-1}(\sigma)V_T$
- The quadratic variation of Z is now independent of σ
- Unfortunately, still perfect dependence between Z and σ via the endpoints of Z

Imputation of latent data

- Focus on single term, say $p(\bar{V} \mid V_0, V_T, \theta, \sigma)$
- Sampling $\bar{V}$ is equivalent to sampling Z conditionally on $Z_0 = \Sigma^{-1}(\sigma)V_0$ and $Z_T = \Sigma^{-1}(\sigma)V_T$
- The quadratic variation of Z is now independent of σ
- Unfortunately, still perfect dependence between Z and σ via the endpoints of Z

Imputation of latent data

- Focus on single term, say $p(\bar{V} \mid V_0, V_T, \theta, \sigma)$
- Sampling $\bar{V}$ is equivalent to sampling Z conditionally on $Z_0 = \Sigma^{-1}(\sigma)V_0$ and $Z_T = \Sigma^{-1}(\sigma)V_T$
- The quadratic variation of Z is now independent of σ
- Unfortunately, still perfect dependence between Z and σ via the endpoints of Z

$$\mathrm{d}Z_t = \alpha(Z_t, \theta, \sigma)\,\mathrm{d}t + \mathrm{d}B_t, \quad Z_0 = \Sigma^{-1}(\sigma)V_0,$$

$$\mathrm{d}Z_t = \alpha(Z_t, \theta, \sigma)\,\mathrm{d}t + \mathrm{d}B_t, \quad Z_0 = \Sigma^{-1}(\sigma)V_0,$$

- Define $\tilde{V}$ as

$$\tilde{V}_t = Z_t - \left(1 - \frac{t}{T}\right) Z_0 - \frac{t}{T} Z_T$$

for $0 \leq t \leq T$.

$$\mathrm{d}Z_t = \alpha(Z_t, \theta, \sigma)\,\mathrm{d}t + \mathrm{d}B_t, \quad Z_0 = \Sigma^{-1}(\sigma)V_0,$$

- Define $\tilde{V}$ as

$$\tilde{V}_t = Z_t - \left(1 - \frac{t}{T}\right) Z_0 - \frac{t}{T} Z_T$$

  for $0 \leq t \leq T$.
- Consider the measure of Z conditionally on endpoints with $\alpha = 0$: Z is a Brownian bridge starting and ending at Z_0 and Z_T

$$\mathrm{d}Z_t = \alpha(Z_t, \theta, \sigma)\,\mathrm{d}t + \mathrm{d}B_t, \quad Z_0 = \Sigma^{-1}(\sigma)V_0,$$

- Define $\tilde{V}$ as

$$\tilde{V}_t = Z_t - \left(1 - \frac{t}{T}\right) Z_0 - \frac{t}{T} Z_T$$

  for $0 \leq t \leq T$.
- Consider the measure of Z conditionally on endpoints with $\alpha = 0$: Z is a Brownian bridge starting and ending at Z_0 and Z_T
- Using Girsanov

$$\frac{\mathrm{d}P^*_\alpha}{\mathrm{d}P^*_0}(Z \mid \theta, \sigma, Z_0, Z_T) \propto_Z \phi(Z, \theta, \sigma)$$

  where

$$\phi(Z, \theta, \sigma) = e^{\int_0^T \alpha(Z_u,\theta,\sigma)^T\,\mathrm{d}Z_u - \frac{1}{2}\int_0^T \alpha(Z_u,\theta,\sigma)^T \alpha(Z_u,\theta,\sigma)\,\mathrm{d}u}$$

- Gibbs sampler works on $(\theta, \sigma, \tilde{V})$

- Gibbs sampler works on $(\theta, \sigma, \tilde{V})$
- Sampling from $p(\tilde{V} \mid D_n, \theta, \sigma)$:

- Gibbs sampler works on $(\theta, \sigma, \tilde{V})$
- Sampling from $p(\tilde{V} \mid D_n, \theta, \sigma)$:

Initialize

1) Initialize a skeleton path $(\tilde{V}^M)_0$ according to P_0^*. Compute Z and approximate the Radon-Nikodym derivative $\phi(Z, \theta, \sigma)$, w_0.

Iterate

- Gibbs sampler works on $(\theta, \sigma, \tilde{V})$
- Sampling from $p(\tilde{V} \mid D_n, \theta, \sigma)$:

Initialize

1) Initialize a skeleton path $(\tilde{V}^M)_0$ according to P_0^*. Compute Z and approximate the Radon-Nikodym derivative $\phi(Z, \theta, \sigma)$, w_0.

Iterate

2) Generate a proposal skeleton path, $\tilde{V}_P^M$ according to P_0^*. Compute $\tilde{Z}$ and approximate the Radon-Nikodym derivative $\phi(\tilde{Z}, \theta, \sigma)$, $\tilde{w}$.

- Gibbs sampler works on $(\theta, \sigma, \tilde{V})$
- Sampling from $p(\tilde{V} \mid D_n, \theta, \sigma)$:

Initialize

1) Initialize a skeleton path $(\tilde{V}^M)_0$ according to P_0^*. Compute Z and approximate the Radon-Nikodym derivative $\phi(Z, \theta, \sigma)$, w_0.

Iterate

2) Generate a proposal skeleton path, $\tilde{V}_P^M$ according to P_0^*. Compute $\tilde{Z}$ and approximate the Radon-Nikodym derivative $\phi(\tilde{Z}, \theta, \sigma)$, $\tilde{w}$.

3) Let $(\tilde{V}^M)_{k+1} = \begin{cases} \tilde{V}_P^M & \text{with prob. } \min\left(1, \frac{\tilde{w}}{w_k}\right) \\ (\tilde{V}^M)_k & \text{otherwise} \end{cases}$.

Sampling from $p(\sigma \mid \theta, \tilde{V}, D_n)$

It can be shown that

$$p(\sigma \mid \tilde{V}, D_n, \theta)$$

$$\propto_\sigma p(\sigma) |\det(\Sigma(\sigma)^{-1})|^n \cdot \prod_{i=1}^{n} \varphi(\Sigma(\sigma)^{-1} V_{t_i} \mid \Sigma(\sigma)^{-1} V_{t_{i-1}}; \Delta I_2) \phi_i(\tilde{V}, \theta, \sigma)$$

$$:= h(\sigma),$$

Sampling from $p(\sigma \mid \theta, \tilde{V}, D_n)$

It can be shown that

$$p(\sigma \mid \tilde{V}, D_n, \theta)$$

$$\propto_\sigma p(\sigma) |\det(\Sigma(\sigma)^{-1})|^n \cdot \prod_{i=1}^{n} \varphi(\Sigma(\sigma)^{-1} V_{t_i} \mid \Sigma(\sigma)^{-1} V_{t_{i-1}}; \Delta I_2) \phi_i(\tilde{V}, \theta, \sigma)$$

$$:= h(\sigma),$$

One can use a Gaussian random walk on transformed variable $\bar{\sigma} = \log(\sigma)$.

Sampling from $p(\sigma \mid \theta, \tilde{V}, D_n)$

It can be shown that

$$p(\sigma \mid \tilde{V}, D_n, \theta)$$

$$\propto_\sigma p(\sigma) |\det(\Sigma(\sigma)^{-1})|^n \cdot \prod_{i=1}^{n} \varphi(\Sigma(\sigma)^{-1} V_{t_i} \mid \Sigma(\sigma)^{-1} V_{t_{i-1}}; \Delta I_2) \phi_i(\tilde{V}, \theta, \sigma)$$

$$:= h(\sigma),$$

One can use a Gaussian random walk on transformed variable $\bar{\sigma} = \log(\sigma)$.

Initialize

1) Initialize $\bar{\sigma}^{(0)}$.

Iterate

Sampling from $p(\sigma \mid \theta, \tilde{V}, D_n)$

It can be shown that

$$p(\sigma \mid \tilde{V}, D_n, \theta)$$

$$\propto_\sigma p(\sigma) |\det(\Sigma(\sigma)^{-1})|^n \cdot \prod_{i=1}^{n} \varphi(\Sigma(\sigma)^{-1} V_{t_i} \mid \Sigma(\sigma)^{-1} V_{t_{i-1}}; \Delta I_2) \phi_i(\tilde{V}, \theta, \sigma)$$

$$:= h(\sigma),$$

One can use a Gaussian random walk on transformed variable $\bar{\sigma} = \log(\sigma)$.

Initialize

1) Initialize $\bar{\sigma}^{(0)}$.

Iterate

2a) Generate a proposal $\tilde{\sigma}$ from $\varphi(\cdot \mid \bar{\sigma}^{(k)}; \Omega)$.

2b) Let $\bar{\sigma}^{(k+1)} = \begin{cases} \tilde{\sigma} & \text{with prob. } \min\left(1, \frac{h(\tilde{\sigma})}{h(\bar{\sigma}^{(k)})}\right) \\ \bar{\sigma}^{(k)} & \text{otherwise} \end{cases}$.

Simulations

	ε	s	γ	β	σ_1	σ_2
Oscillatory	0.1	0.5	1.5	0.6	0.5	0.3
Excitatory	0.1	0.5	1.5	1.4	0.5	0.3

Simulations

	ε	s	γ	β	σ_1	σ_2
Oscillatory	0.1	0.5	1.5	0.6	0.5	0.3
Excitatory	0.1	0.5	1.5	1.4	0.5	0.3

- Model was re-parameterized:

$$(\varepsilon, s, \gamma, \beta) \mapsto (1/\varepsilon, s/\varepsilon, \gamma, \beta)$$

Simulations

	ε	s	γ	β	σ_1	σ_2
Oscillatory	0.1	0.5	1.5	0.6	0.5	0.3
Excitatory	0.1	0.5	1.5	1.4	0.5	0.3

- Model was re-parameterized:

$$(\varepsilon, s, \gamma, \beta) \mapsto (1/\varepsilon, s/\varepsilon, \gamma, \beta)$$

- $\varepsilon > 0$ and thus a further transformation is needed to guarantee positivity of $\tilde{\varepsilon}$. However, for small ε the problem can be neglected for practical implementations.

Simulations

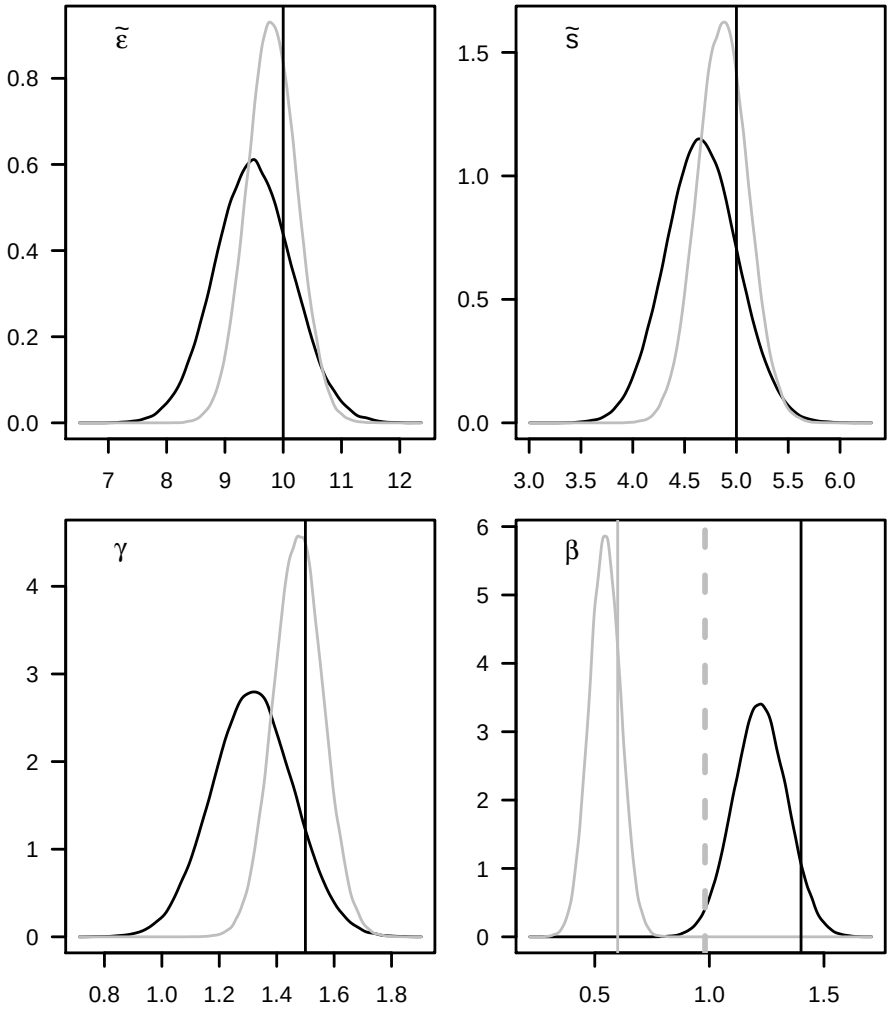
	ε	s	γ	β	σ_1	σ_2
Oscillatory	0.1	0.5	1.5	0.6	0.5	0.3
Excitatory	0.1	0.5	1.5	1.4	0.5	0.3

- Model was re-parameterized:

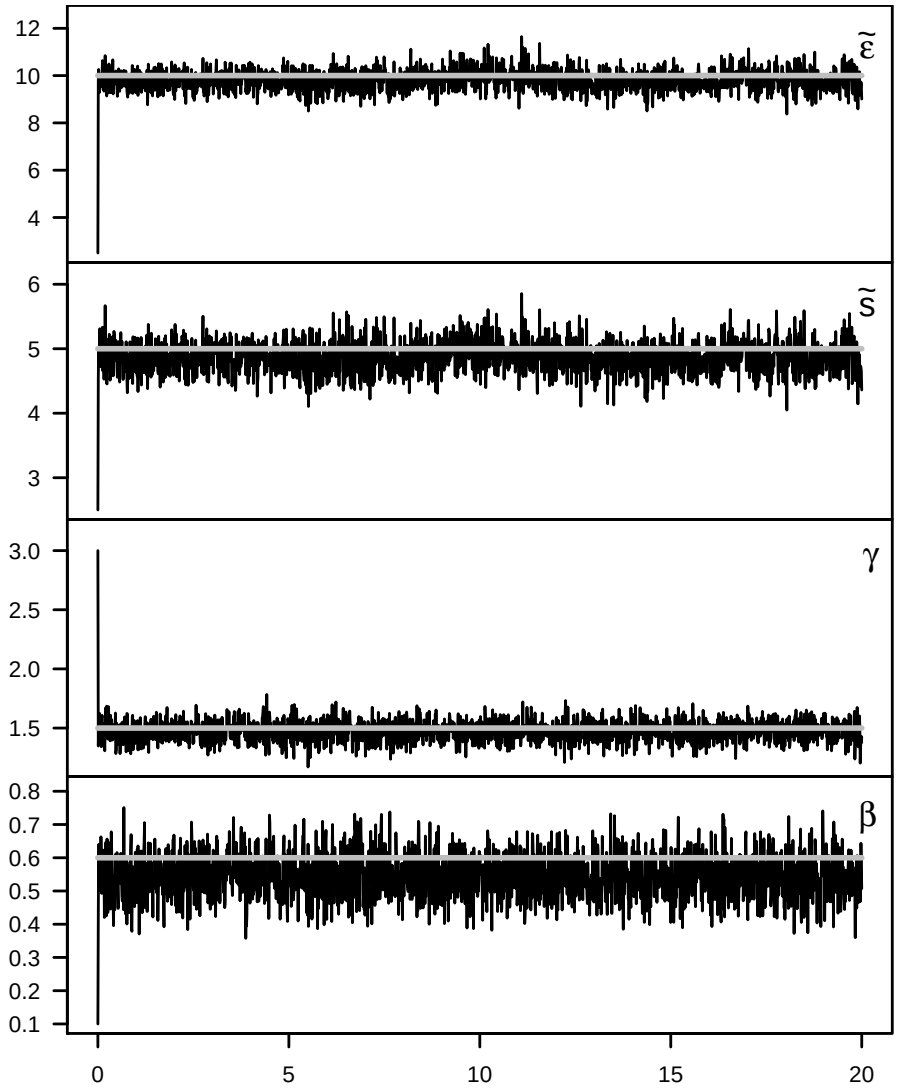
$$(\varepsilon, s, \gamma, \beta) \mapsto (1/\varepsilon, s/\varepsilon, \gamma, \beta)$$

- $\varepsilon > 0$ and thus a further transformation is needed to guarantee positivity of $\tilde{\varepsilon}$. However, for small ε the problem can be neglected for practical implementations.
- Radon Nikodym derivatives are evaluated using Riemann approximations

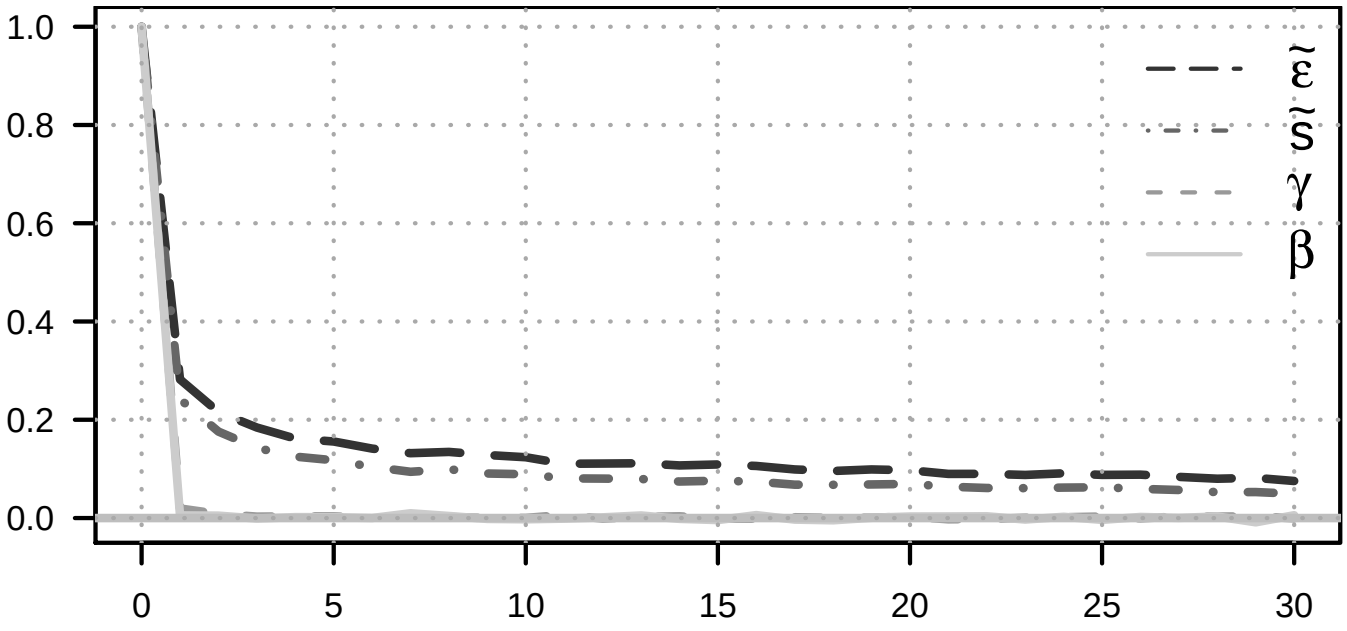
- $n = 200$, $\Delta = 0.1$
- Black lines: excitatory data, Gray: oscillatory data



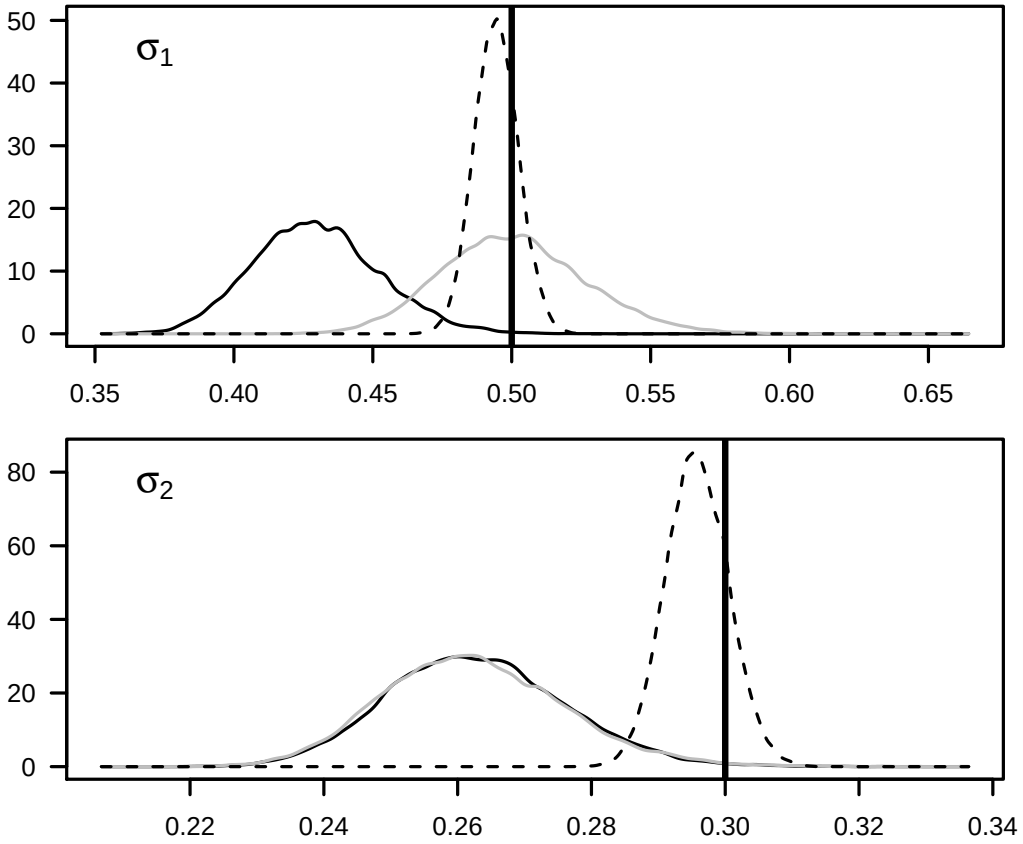
- $n = 200$, $\Delta = 0.1$, oscillatory data
- Shows only every 50th iteration



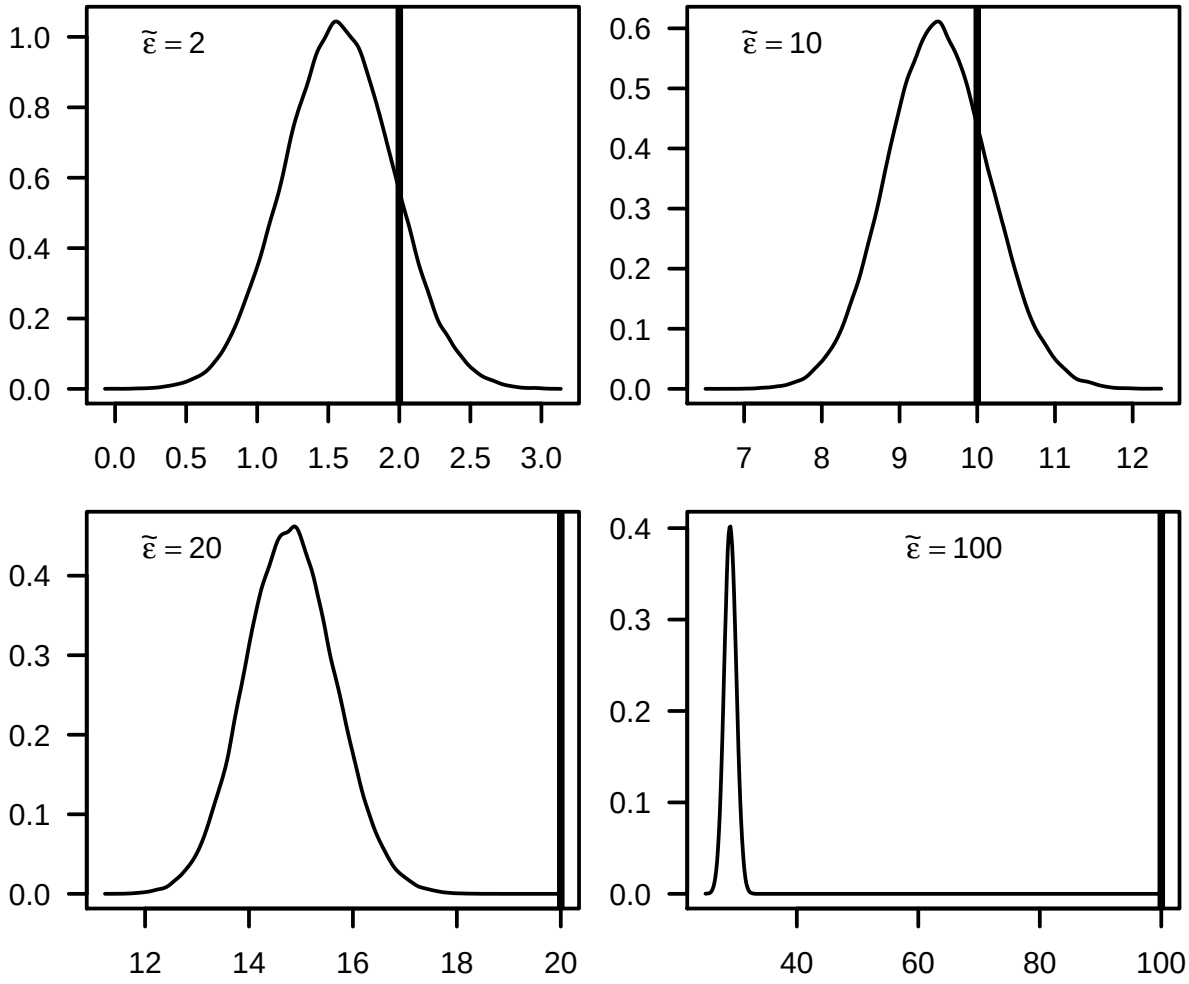
- $n = 200$, $\Delta = 0.1$, oscillatory data
- Burn in: 1000 iterations



- Solid black line: excitatory data, solid gray line: oscillatory data ($n = 200$, $\Delta = 0.1$)
- Dashed black line: excitatory data ($n = 2000$, $\Delta = 0.01$)



- $n = 200$, $\Delta = 0.1$, excitatory data



Future works

- Larger simulation study, with parameters designed to stress the procedure
- Optimize implementation and construct R-package
- Propose paths from bridge with linear drift
- Generalize to partial observation of only one coordinate

Future works

- Larger simulation study, with parameters designed to stress the procedure
- Optimize implementation and construct R-package
- Propose paths from bridge with linear drift
- Generalize to partial observation of only one coordinate

Future works

- Larger simulation study, with parameters designed to stress the procedure
- Optimize implementation and construct R-package
- Propose paths from bridge with linear drift
- Generalize to partial observation of only one coordinate

Future works

- Larger simulation study, with parameters designed to stress the procedure
- Optimize implementation and construct R-package
- Propose paths from bridge with linear drift
- Generalize to partial observation of only one coordinate