Homogénéisation guidée par les données de l'équation de Langevin multi-échelle

Assyr Abdulle [1] Giacomo Garegnani [1] Grigorios A. Pavliotis [2]
Andrew M. Stuart [3] Andrea Zanoni [1]

CANUM 2020+2 – Évian-les-Bains – 13-17 Juin, 2022

[1]Institute of Mathematics, EPFL
[2]Department of Mathematics, Imperial College London
[3]Department of Computing and Mathematical Sciences, Caltech

Motivation

Applications

- ▷ Multiscale diffusions suitable model for
  - oceanography
  - finance
  - ...
- ▷ Infer from data effective simple models for complex phenomena

Motivation

Applications

- ▷ Multiscale diffusions suitable model for
  - oceanography
  - finance
  - ...
- ▷ Infer from data effective simple models for complex phenomena

Existing Techniques

- ▷ Pre-processing needed to obtain effective dynamics
- ▷ Subsampling widely employed, but
  - requires knowledge of scale-separation
  - strongly dependent on rate
  - throws a lot of data

Motivation

Applications

- ▷ Multiscale diffusions suitable model for
  - oceanography
  - finance
  - ...
- ▷ Infer from data effective simple models for complex phenomena

Existing Techniques

- ▷ Pre-processing needed to obtain effective dynamics
- ▷ Subsampling widely employed, but
  - requires knowledge of scale-separation
  - strongly dependent on rate
  - throws a lot of data

Necessity for new method

- ▷ Consistence w.r.t. homogenization theory
- ▷ Robustness w.r.t. subsampling techniques
- ▷ Ease of applicability

Outline

Setting – The model

Multiscale Langevin SDE

$$\mathrm{d}X_t^\varepsilon = -\alpha \cdot V'(X_t^\varepsilon)\,\mathrm{d}t - \frac{1}{\varepsilon} p'\left(\frac{X_t^\varepsilon}{\varepsilon}\right)\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t$$

Parameters	Is it known?
- drift coefficient $\alpha \in \mathbb{R}^M$	✗
- slow potential $V\colon \mathbb{R} \to \mathbb{R}^M$	✓
- L-periodic fast potential $p\colon \mathbb{R} \to \mathbb{R}$	✗
- diffusion coefficient $\sigma > 0$	✗
- multiscale parameter $\varepsilon > 0$	✗

Setting – The model

Multiscale Langevin SDE

$$\mathrm{d}X_t^\varepsilon = -\underbrace{\alpha \cdot V'(X_t^\varepsilon)}\,\mathrm{d}t - \frac{1}{\varepsilon}p'\left(\frac{X_t^\varepsilon}{\varepsilon}\right)\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t$$

$$\downarrow$$

<u>Remark:</u> we consider a semi-parametric framework

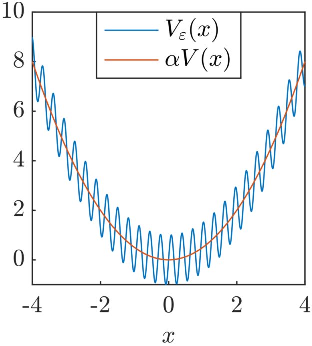
$$\alpha \cdot V(x) = \sum_{m=1}^{M} \alpha_m V_m(x)$$

indeed $\{V_m\}_{m=1}^M$ can be chosen as basis of appropriate function space

Setting – The model

Multiscale Langevin SDE

$$\mathrm{d}X_t^\varepsilon = -\alpha \cdot V'(X_t^\varepsilon)\,\mathrm{d}t - \frac{1}{\varepsilon}p'\left(\frac{X_t^\varepsilon}{\varepsilon}\right)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t$$



Motion in multiscale potential

$$V_\varepsilon(x) = \alpha V(x) + p(x/\varepsilon)$$

where:

- $V(x) = x^2/2$
- $\alpha = 1$
- $p(y) = \sin(y)$
- $\varepsilon = 0.05$

Setting – Homogenization

$X_t^\varepsilon \to X_t^0$ as $\varepsilon \to 0$ as random variables in $\mathcal{C}^0([0,T])$[4]

Homogenized Langevin SDE

$$\mathrm{d}X_t^0 = -A \cdot V'(X_t^0)\,\mathrm{d}t + \sqrt{2\Sigma}\,\mathrm{d}W_t$$

New parameters	Is it known?
- effective drift coefficient $A = K\alpha \in \mathbb{R}^M$	✗
- effective diffusion coefficient $\Sigma = K\sigma > 0$	✗

where $0 < K < 1$ dependent on p and σ and $K \to 1$ as $\sigma \to \infty$

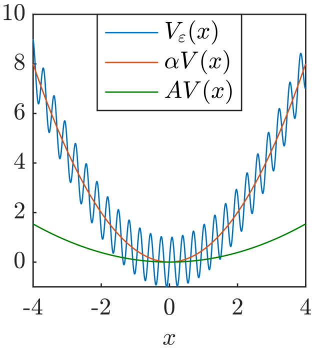
[4] Bensoussan et al. (1978)

Setting – Homogenization

$X_t^\varepsilon \to X_t^0$ as $\varepsilon \to 0$ as random variables in $\mathcal{C}^0([0,T])$[4]

Homogenized Langevin SDE

$$\mathrm{d}X_t^0 = -A \cdot V'(X_t^0)\,\mathrm{d}t + \sqrt{2\Sigma}\,\mathrm{d}W_t$$



Effect of σ on the homogenized potential $AV(x)$ with

- $\sigma = 0.5 \implies K = 0.19$
- $V(x) = x^2/2$
- $\alpha = 1$
- $p(y) = \sin(y)$
- $\varepsilon = 0.05$
- $V_\varepsilon(x) = \alpha V(x) + p(x/\varepsilon)$

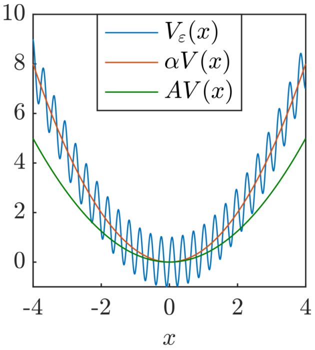
[4] Bensoussan et al. (1978)

Setting – Homogenization

$X_t^\varepsilon \to X_t^0$ as $\varepsilon \to 0$ as random variables in $\mathcal{C}^0([0,T])$[4]

Homogenized Langevin SDE

$$\mathrm{d}X_t^0 = -A \cdot V'(X_t^0)\,\mathrm{d}t + \sqrt{2\Sigma}\,\mathrm{d}W_t$$



Effect of σ on the homogenized potential $AV(x)$ with

- $\sigma = 1 \quad \Longrightarrow \quad K = 0.62$
- $V(x) = x^2/2$
- $\alpha = 1$
- $p(y) = \sin(y)$
- $\varepsilon = 0.05$
- $V_\varepsilon(x) = \alpha V(x) + p(x/\varepsilon)$

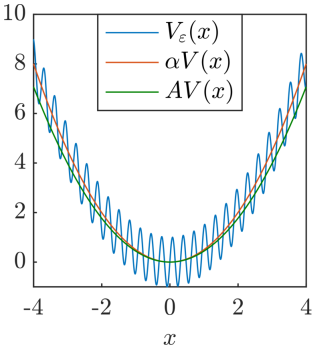
[4] Bensoussan et al. (1978)

Setting – Homogenization

$X_t^\varepsilon \to X_t^0$ as $\varepsilon \to 0$ as random variables in $\mathcal{C}^0([0,T])$[4]

Homogenized Langevin SDE

$$\mathrm{d}X_t^0 = -A \cdot V'(X_t^0)\,\mathrm{d}t + \sqrt{2\Sigma}\,\mathrm{d}W_t$$



Effect of σ on the homogenized potential $AV(x)$ with

- $\sigma = 2 \quad \Longrightarrow \quad K = 0.88$
- $V(x) = x^2/2$
- $\alpha = 1$
- $p(y) = \sin(y)$
- $\varepsilon = 0.05$
- $V_\varepsilon(x) = \alpha V(x) + p(x/\varepsilon)$

[4] Bensoussan et al. (1978)

Setting – Parameter inference

$$\mathrm{d}X_t^\varepsilon = -\alpha \cdot V'(X_t^\varepsilon)\,\mathrm{d}t - \frac{1}{\varepsilon} p'\left(\frac{X_t^\varepsilon}{\varepsilon}\right)\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t \qquad \longrightarrow \qquad \text{data}$$

$$\mathrm{d}X_t^0 = -A \cdot V'(X_t^0)\,\mathrm{d}t + \sqrt{2\Sigma}\,\mathrm{d}W_t \qquad \longrightarrow \qquad \text{model}$$

Goal: estimate $A \in \mathbb{R}^M$ and $\Sigma > 0$ given:

- continuous observations $X^\varepsilon = (X_t^\varepsilon, 0 \leq t \leq T)$
  - ▷ maximum likelihood estimator
  - ▷ quadratic variation
  - ▷ subsampling
  - ▷ filtering

Setting – Parameter inference

$$\mathrm{d}X_t^\varepsilon = -\alpha \cdot V'(X_t^\varepsilon)\,\mathrm{d}t - \frac{1}{\varepsilon} p'\left(\frac{X_t^\varepsilon}{\varepsilon}\right)\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t \qquad \longrightarrow \qquad \text{data}$$

$$\mathrm{d}X_t^0 = -A \cdot V'(X_t^0)\,\mathrm{d}t + \sqrt{2\Sigma}\,\mathrm{d}W_t \qquad \longrightarrow \qquad \text{model}$$

Goal: estimate $A \in \mathbb{R}^M$ and $\Sigma > 0$ given:

- continuous observations $X^\varepsilon = (X_t^\varepsilon, 0 \leq t \leq T)$
  - ▷ maximum likelihood estimator
  - ▷ quadratic variation
  - ▷ subsampling
  - ▷ filtering
- discrete observations $\widetilde{X}^\varepsilon = (\widetilde{X}_n^\varepsilon = X_{n\Delta}^\varepsilon, n = 0, \dots, N, \Delta = T/N)$
  - ▷ martingale estimating functions
  - ▷ eigenvalues and eigenfunctions of the generator

Outline

Classic estimators ($M = 1$)

Assume to know continuous observations $X^\varepsilon = (X_t^\varepsilon, 0 \le t \le T)$

Drift: maximum likelihood estimator (MLE)

$$\widehat{A}_{\mathrm{MLE}}(X^\varepsilon, T) = -\frac{\int_0^T V'(X_t^\varepsilon)\,\mathrm{d}X_t^\varepsilon}{\int_0^T V'(X_t^\varepsilon)^2\,\mathrm{d}t} \qquad \text{(Girsanov formula)}$$

Diffusion: quadratic variation (partition $P = \{0 = t_0 < \cdots < t_{K_P} = T\}$)

$$\widehat{\Sigma}_{\mathrm{QV}}(X^\varepsilon, T) = \frac{\langle X^\varepsilon \rangle_T}{2T}, \qquad \langle X^\varepsilon \rangle_T = \lim_{\|P\| \to 0} \sum_{k=0}^{K_P - 1} (X_{t_{k+1}}^\varepsilon - X_{t_k}^\varepsilon)^2$$

Classic estimators ($M = 1$)

Assume to know continuous observations $X^\varepsilon = (X_t^\varepsilon, 0 \leq t \leq T)$

Drift: maximum likelihood estimator (MLE)

$$\widehat{A}_{\mathrm{MLE}}(X^\varepsilon, T) = -\frac{\int_0^T V'(X_t^\varepsilon)\,\mathrm{d}X_t^\varepsilon}{\int_0^T V'(X_t^\varepsilon)^2\,\mathrm{d}t} \qquad \text{(Girsanov formula)}$$

Diffusion: quadratic variation (partition $P = \{0 = t_0 < \cdots < t_{K_P} = T\}$)

$$\widehat{\Sigma}_{\mathrm{QV}}(X^\varepsilon, T) = \frac{\langle X^\varepsilon \rangle_T}{2T}, \qquad \langle X^\varepsilon \rangle_T = \lim_{\|P\| \to 0} \sum_{k=0}^{K_P - 1} (X_{t_{k+1}}^\varepsilon - X_{t_k}^\varepsilon)^2$$

Issue: estimators are asymptotically biased[5]

$$\lim_{\varepsilon \to 0} \lim_{T \to \infty} \widehat{A}_{\mathrm{MLE}}(X^\varepsilon, T) = \alpha \qquad a.s.$$

$$\widehat{\Sigma}_{\mathrm{QV}}(X^\varepsilon, T) = \sigma$$

[5] Pavliotis and Stuart (2007)

Classic estimators ($M = 1$)

Assume to know continuous observations $X^\varepsilon = (X_t^\varepsilon, 0 \leq t \leq T)$

Drift: maximum likelihood estimator (MLE)

$$\widehat{A}_{\mathrm{MLE}}(X^\varepsilon, T) = -\frac{\int_0^T V'(X_t^\varepsilon)\,\mathrm{d}X_t^\varepsilon}{\int_0^T V'(X_t^\varepsilon)^2\,\mathrm{d}t} \qquad \text{(Girsanov formula)}$$

Diffusion: quadratic variation (partition $P = \{0 = t_0 < \cdots < t_{K_P} = T\}$)

$$\widehat{\Sigma}_{\mathrm{QV}}(X^\varepsilon, T) = \frac{\langle X^\varepsilon \rangle_T}{2T}, \qquad \langle X^\varepsilon \rangle_T = \lim_{\|P\| \to 0} \sum_{k=0}^{K_P - 1} (X_{t_{k+1}}^\varepsilon - X_{t_k}^\varepsilon)^2$$

Solution in literature: subsample the data with rate $\delta = \varepsilon^\zeta$, $\zeta \in (0,1)$[5]

$$\widehat{A}_{\mathrm{sub}}^\delta(X^\varepsilon, T) = -\frac{\sum_{n=0}^{N-1} V'(\widetilde{X}_n^\varepsilon)(\widetilde{X}_{n+1}^\varepsilon - \widetilde{X}_n^\varepsilon)}{\delta \sum_{n=0}^{N-1} V'(\widetilde{X}_n^\varepsilon)^2} \xrightarrow[\varepsilon \to 0]{T = \varepsilon^{-\gamma}} A \text{ in probability}$$

$$\widehat{\Sigma}_{\mathrm{sub}}^\delta(X^\varepsilon, T) = \frac{1}{2T} \sum_{n=0}^{N-1} (\widetilde{X}_{n+1}^\varepsilon - \widetilde{X}_n^\varepsilon)^2 \xrightarrow[\varepsilon \to 0]{T \text{ fixed}} \Sigma \text{ in probability}$$

[5] Pavliotis and Stuart (2007)

Classic estimators ($M = 1$)

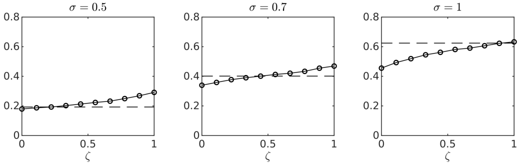
Solution in literature: subsample the data with rate $\delta = \varepsilon^\zeta$, $\zeta \in (0,1)$[5]

$$\widehat{A}^\delta_{\text{sub}}(X^\varepsilon, T) = -\frac{\sum_{n=0}^{N-1} V'(\widetilde{X}^\varepsilon_n)(\widetilde{X}^\varepsilon_{n+1} - \widetilde{X}^\varepsilon_n)}{\delta \sum_{n=0}^{N-1} V'(\widetilde{X}^\varepsilon_n)^2} \xrightarrow[\varepsilon \to 0]{T = \varepsilon^{-\gamma}} A \text{ in probability}$$

$$\widehat{\Sigma}^\delta_{\text{sub}}(X^\varepsilon, T) = \frac{1}{2T} \sum_{n=0}^{N-1} (\widetilde{X}^\varepsilon_{n+1} - \widetilde{X}^\varepsilon_n)^2 \xrightarrow[\varepsilon \to 0]{T \text{ fixed}} \Sigma \text{ in probability}$$

Warning: subsampling has several disadvantages:

- **not robust** with respect to subsampling rate δ
- knowledge of ε is required
- majority of data is wasted



[5] Pavliotis and Stuart (2007)

The filtered data approach

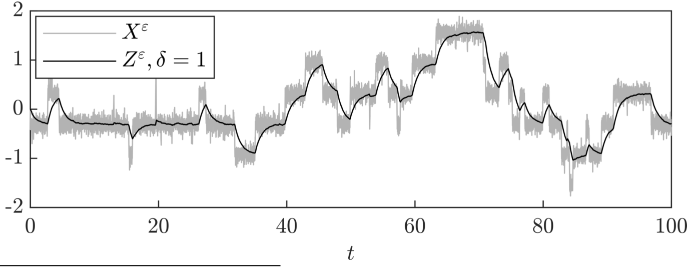
Idea: filter the data to cancel fast scale components[6]

$$Z_t^\varepsilon = \int_0^t k_{\exp}^{\delta,\beta}(t-s) X_s^\varepsilon \, \mathrm{d}s$$

where $\delta, \beta > 0$ and

$$k_{\exp}^{\delta,\beta}(r) = \frac{\beta}{\Gamma(1/\beta)\delta^{1/\beta}} e^{-\frac{r^\beta}{\delta}}$$

In the plot $\beta = 1$:



[6] Abdulle et al. (2021); Garegnani and Zanoni (2021)

The filtered data approach

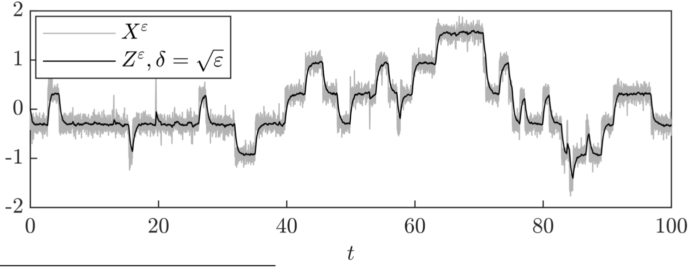
Idea: filter the data to cancel fast scale components[6]

$$Z_t^\varepsilon = \int_0^t k_{\exp}^{\delta,\beta}(t-s) X_s^\varepsilon \, \mathrm{d}s$$

where $\delta, \beta > 0$ and

$$k_{\exp}^{\delta,\beta}(r) = \frac{\beta}{\Gamma(1/\beta)\delta^{1/\beta}} e^{-\frac{r^\beta}{\delta}}$$

In the plot $\beta = 1$:



[6] Abdulle et al. (2021); Garegnani and Zanoni (2021)

The filtered data approach

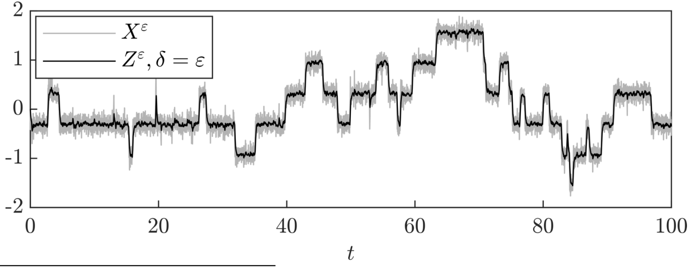
Idea: filter the data to cancel fast scale components[6]

$$Z_t^\varepsilon = \int_0^t k_{\exp}^{\delta,\beta}(t-s) X_s^\varepsilon \, \mathrm{d}s$$

where $\delta, \beta > 0$ and

$$k_{\exp}^{\delta,\beta}(r) = \frac{\beta}{\Gamma(1/\beta)\delta^{1/\beta}} e^{-\frac{r^\beta}{\delta}}$$

In the plot $\beta = 1$:



[6] Abdulle et al. (2021); Garegnani and Zanoni (2021)

The filtered data approach – The estimators

Drift: modification of the MLE

$$\widehat{A}^{\delta,\beta}_{\exp}(X^\varepsilon, T) = -\frac{\int_0^T V'(Z_t^\varepsilon)\,\mathrm{d}X_t^\varepsilon}{\int_0^T V'(Z_t^\varepsilon)V'(X_t^\varepsilon)\,\mathrm{d}t} \qquad (M=1)$$

Diffusion:

$$(i) \quad \widehat{\Sigma}^{\delta}_{\exp}(X^\varepsilon, T) = \frac{1}{\delta T}\int_0^T (X_t^\varepsilon - Z_t^\varepsilon)^2\,\mathrm{d}t \qquad (\beta = 1)$$

$$(ii) \quad \widetilde{\Sigma}^{\delta,\beta}_{\exp}(X^\varepsilon, T) = \frac{\widehat{A}^{\delta,\beta}_{\exp}(X^\varepsilon, T)}{\widehat{A}_{\mathrm{MLE}}(X^\varepsilon, T)}\widehat{\Sigma}_{\mathrm{QV}}(X^\varepsilon, T)$$

Why (ii)? $\quad \Sigma = K\sigma = \dfrac{A}{\alpha}\sigma$

The filtered data approach – Convergence analysis

Drift:

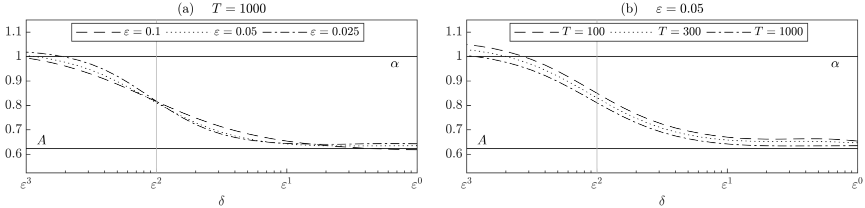
$$\widehat{A}^{\delta,\beta}_{\exp}(X^\varepsilon, T) = -\frac{\int_0^T V'(Z_t^\varepsilon)\,\mathrm{d}X_t^\varepsilon}{\int_0^T V'(Z_t^\varepsilon)V'(X_t^\varepsilon)\,\mathrm{d}t} \qquad (M=1)$$

Theorem ($\beta = 1$)

$$(i)\ \lim_{\varepsilon\to 0}\lim_{T\to\infty} \widehat{A}^{\delta,1}_{\exp}(X^\varepsilon, T) = A, \quad a.s., \quad \delta \textit{ independent of } \varepsilon$$

$$(ii)\ \lim_{\varepsilon\to 0}\lim_{T\to\infty} \widehat{A}^{\delta,1}_{\exp}(X^\varepsilon, T) = A, \quad a.s., \quad \delta = \varepsilon^\zeta,\ 0<\zeta<2$$

$$(iii)\ \lim_{\varepsilon\to 0}\lim_{T\to\infty} \widehat{A}^{\delta,1}_{\exp}(X^\varepsilon, T) = \alpha, \quad a.s., \quad \delta = \varepsilon^\zeta,\ \zeta>2$$



The filtered data approach – Convergence analysis

Diffusion:

$$\widehat{\Sigma}^{\delta}_{\text{exp}}(X^{\varepsilon}, T) = \frac{1}{\delta T} \int_0^T (X_t^{\varepsilon} - Z_t^{\varepsilon})^2 \, \mathrm{d}t$$

Theorem

$$\lim_{\varepsilon \to 0} \lim_{T \to \infty} \widehat{\Sigma}^{\delta}_{\text{exp}}(X^{\varepsilon}, T) = \Sigma, \quad a.s., \quad \delta = \varepsilon^{\zeta}, \; 0 < \zeta < 2$$

The filtered data approach – Convergence analysis

Diffusion:

$$\widehat{\Sigma}^{\delta}_{\exp}(X^{\varepsilon}, T) = \frac{1}{\delta T} \int_0^T (X_t^{\varepsilon} - Z_t^{\varepsilon})^2 \, \mathrm{d}t$$

Theorem

$$\lim_{\varepsilon \to 0} \lim_{T \to \infty} \widehat{\Sigma}^{\delta}_{\exp}(X^{\varepsilon}, T) = \Sigma, \quad a.s., \quad \delta = \varepsilon^{\zeta}, \; 0 < \zeta < 2$$

$$\widetilde{\Sigma}^{\delta,\beta}_{\exp}(X^{\varepsilon}, T) = \frac{\widehat{A}^{\delta,\beta}_{\exp}(X^{\varepsilon}, T)}{\widehat{A}_{\mathrm{MLE}}(X^{\varepsilon}, T)} \widehat{\Sigma}_{\mathrm{QV}}(X^{\varepsilon}, T)$$

Theorem ($\beta = 1$)

$$(i) \;\; \lim_{\varepsilon \to 0} \lim_{T \to \infty} \widetilde{\Sigma}^{\delta,1}_{\exp}(X^{\varepsilon}, T) = \Sigma, \quad a.s., \quad \delta \textit{ independent of } \varepsilon$$

$$(ii) \;\; \lim_{\varepsilon \to 0} \lim_{T \to \infty} \widetilde{\Sigma}^{\delta,1}_{\exp}(X^{\varepsilon}, T) = \Sigma, \quad a.s., \quad \delta = \varepsilon^{\zeta}, \; 0 < \zeta < 2$$

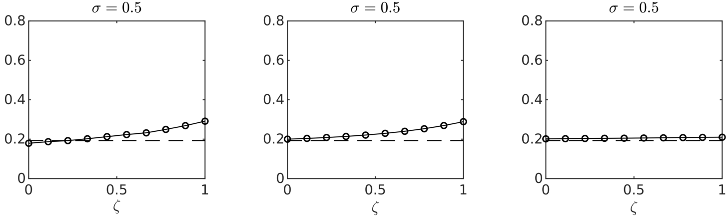
$$(iii) \;\; \lim_{\varepsilon \to 0} \lim_{T \to \infty} \widetilde{\Sigma}^{\delta,1}_{\exp}(X^{\varepsilon}, T) = \sigma, \quad a.s., \quad \delta = \varepsilon^{\zeta}, \; \zeta > 2$$

Numerical experiments – Filtering vs Subsampling

Setting: drift estimation for Ornstein–Uhlenbeck process

$$V(x) = \frac{x^2}{2},\ p(y) = \cos(y),\ \alpha = 1,\ \varepsilon = 0.1,\ T = 1000,\ \delta = \varepsilon^\zeta$$

Subsampling | Filtering, $\beta = 1$ | Filtering, $\beta = 5$

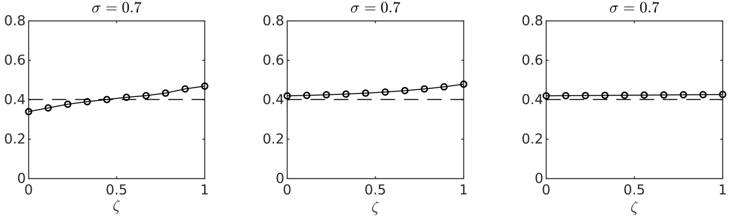


Numerical experiments – Filtering vs Subsampling

Setting: drift estimation for Ornstein–Uhlenbeck process

$$V(x) = \frac{x^2}{2},\ p(y) = \cos(y),\ \alpha = 1,\ \varepsilon = 0.1,\ T = 1000,\ \delta = \varepsilon^\zeta$$

Subsampling   Filtering, $\beta = 1$ Filtering, $\beta = 5$

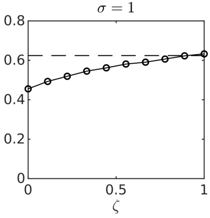


Numerical experiments – Filtering vs Subsampling

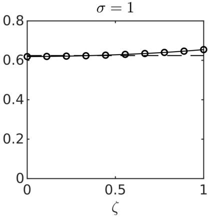
Setting: drift estimation for Ornstein–Uhlenbeck process

$$V(x) = \frac{x^2}{2},\ p(y) = \cos(y),\ \alpha = 1,\ \varepsilon = 0.1,\ T = 1000,\ \delta = \varepsilon^\zeta$$

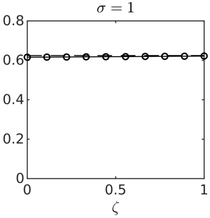
Subsampling



Filtering, $\beta = 1$



Filtering, $\beta = 5$

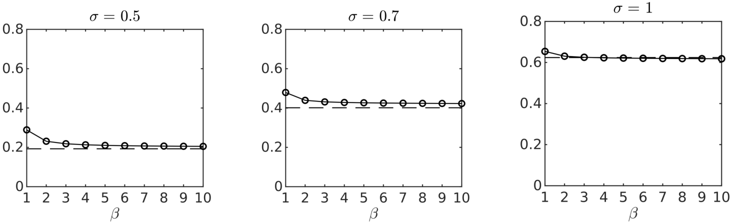


Numerical experiments – Filtering vs Subsampling

Setting: drift estimation for Ornstein–Uhlenbeck process

$$V(x) = \frac{x^2}{2},\ p(y) = \cos(y),\ \alpha = 1,\ \varepsilon = 0.1,\ T = 1000$$

Sensitivity analysis w.r.t. β ($\delta = \varepsilon$)

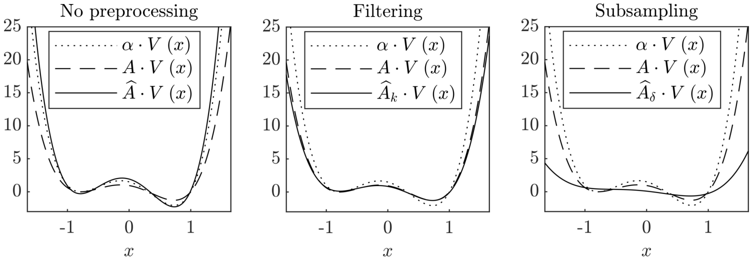


Numerical experiments – Multidimensional drift

Setting: let $T_m(x)$ be Chebyshev polynomials

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{m+1}(x) = 2xT_m(x) - T_{m-1}(x)$$

- Drift estimation for $V_i(x) = T_i(x)$, $i = 1, \dots, 4$ and $p(y) = \cos(y)$, $\varepsilon = 0.05$, $T = 1000$
- Filtering with $\delta = 1, \beta = 1$
- Subsampling with $\delta = \varepsilon^{2/3}$

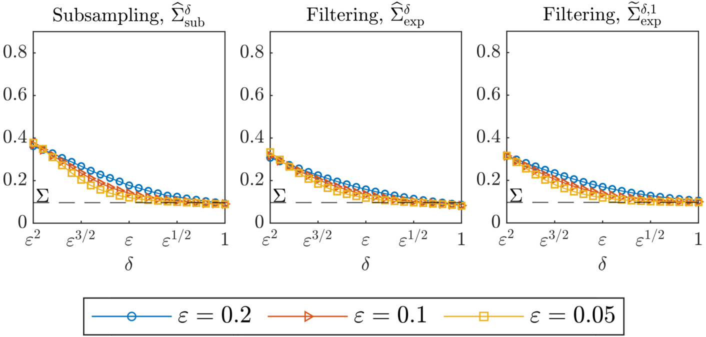


Numerical experiments – Diffusion coefficient

Setting: diffusion estimation for Ornstein–Uhlenbeck process

$$V(x) = \frac{x^2}{2},\ p(y) = \cos(y),\ \alpha = 1,\ T = 1000$$

$$\sigma = 0.5$$

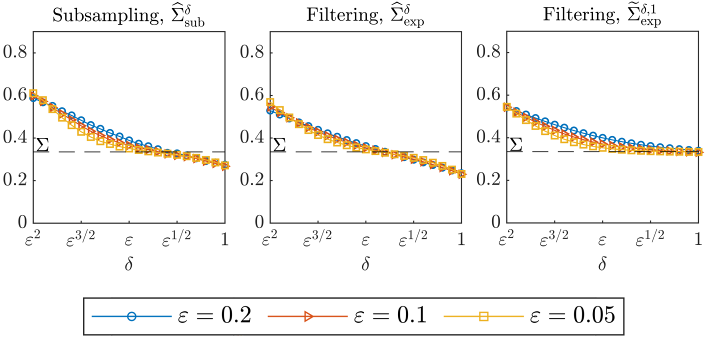


Numerical experiments – Diffusion coefficient

Setting: diffusion estimation for Ornstein–Uhlenbeck process

$$V(x) = \frac{x^2}{2},\ p(y) = \cos(y),\ \alpha = 1,\ T = 1000$$

$$\sigma = 0.75$$

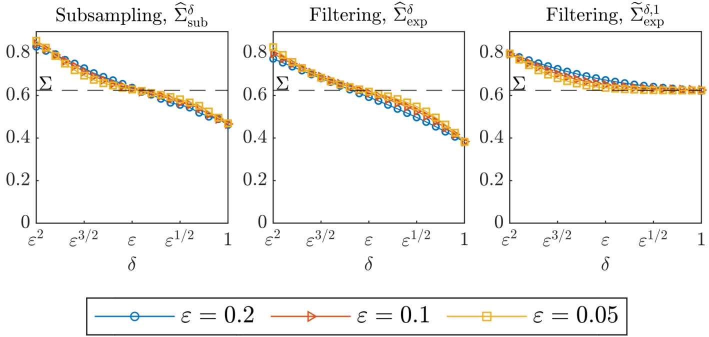


Numerical experiments – Diffusion coefficient

Setting: diffusion estimation for Ornstein–Uhlenbeck process

$$V(x) = \frac{x^2}{2},\ p(y) = \cos(y),\ \alpha = 1,\ T = 1000$$

$$\sigma = 1$$



Outline

Martingale estimating functions

Assume to know discrete observations $\{\widetilde{X}_n^\varepsilon\}_{n=0}^N, \widetilde{X}_n^\varepsilon = X_{n\Delta}^\varepsilon, \Delta = T/N$

Issue: **discretization** of MLE is **biased** if Δ independent of ε

Martingale estimating functions

Assume to know discrete observations $\{\widetilde{X}_n^\varepsilon\}_{n=0}^N, \widetilde{X}_n^\varepsilon = X_{n\Delta}^\varepsilon, \Delta = T/N$

Issue: **discretization** of MLE is **biased** if Δ independent of ε

Idea: employ martingale estimating functions based on eigenvalues and eigenfunctions of the generator of the homogenized dynamics[7]

Homogenized dynamics:	$\mathrm{d}X_t^0 = -a \cdot V'(X_t^0)\,\mathrm{d}t + \sqrt{2s}\mathrm{d}W_t$
Generator:	$\mathcal{L}_{(a,s)}u(x) = -a \cdot V'(x)u'(x) + su''(x)$
Eigenvalue problem:	$-\mathcal{L}_{(a,s)}\phi_j(x;a,s) = \lambda_j(a,s)\phi_j(x;a,s)$
Smooth functions:	$\beta_j(\cdot;a,s) \colon \mathbb{R} \to \mathbb{R}^{M+1}$

[7] Kessler and Sørensen (1999); Abdulle et al. (2022)

Martingale estimating functions

Assume to know discrete observations $\{\widetilde{X}_n^\varepsilon\}_{n=0}^N, \widetilde{X}_n^\varepsilon = X_{n\Delta}^\varepsilon, \Delta = T/N$

Issue: **discretization** of MLE is **biased** if Δ independent of ε

Idea: employ martingale estimating functions based on eigenvalues and eigenfunctions of the generator of the homogenized dynamics[7]

Homogenized dynamics:	$\mathrm{d}X_t^0 = -a \cdot V'(X_t^0)\,\mathrm{d}t + \sqrt{2s}\mathrm{d}W_t$
Generator:	$\mathcal{L}_{(a,s)}u(x) = -a \cdot V'(x)u'(x) + su''(x)$
Eigenvalue problem:	$-\mathcal{L}_{(a,s)}\phi_j(x;a,s) = \lambda_j(a,s)\phi_j(x;a,s)$
Smooth functions:	$\beta_j(\cdot;a,s)\colon \mathbb{R} \to \mathbb{R}^{M+1}$

If Δ small: combine previous filtering methodology

$$\widetilde{Z}_n^\varepsilon = \Delta \sum_{k=0}^{n-1} e^{-\Delta(n-k)} \widetilde{X}_k^\varepsilon \qquad (\delta = 1,\ \beta = 1)$$

[7] Kessler and Sørensen (1999); Abdulle et al. (2022)

Martingale estimating functions – The estimators

Estimator implicitly defined from nonlinear system of dimension $M+1$

Without filtered data: $\widehat{G}^{\varepsilon}_{N,J}(a,s) = 0 \;\Rightarrow\; \widehat{A}^{J}_{\text{eigen}}(\widetilde{X}^{\varepsilon}, N), \widehat{\Sigma}^{J}_{\text{eigen}}(\widetilde{X}^{\varepsilon}, N)$

$$\widehat{G}^{\varepsilon}_{N,J}(a,s) = \frac{1}{\Delta}\sum_{n=0}^{N-1}\sum_{j=1}^{J}\left\{\beta_j(\widetilde{X}^{\varepsilon}_n; a, s)\left(\phi_j(\widetilde{X}^{\varepsilon}_{n+1}; a, s) - e^{-\lambda_j(a,s)\Delta}\phi_j(\widetilde{X}^{\varepsilon}_n; a, s)\right)\right\}$$

With filtered data: $\widetilde{G}^{\varepsilon}_{N,J}(a,s) = 0 \;\Rightarrow\; \widetilde{A}^{J}_{\text{eigen}}(\widetilde{X}^{\varepsilon}, N), \widehat{\Sigma}^{J}_{\text{eigen}}(\widetilde{X}^{\varepsilon}, N)$

$$\widetilde{G}^{\varepsilon}_{N,J}(a,s) = \frac{1}{\Delta}\sum_{n=0}^{N-1}\sum_{j=1}^{J}\left\{\beta_j(\widetilde{Z}^{\varepsilon}_n; a, s)\left(\phi_j(\widetilde{X}^{\varepsilon}_{n+1}; a, s) - e^{-\lambda_j(a,s)\Delta}\phi_j(\widetilde{X}^{\varepsilon}_n; a, s)\right)\right\}$$

Martingale estimating functions – The estimators

Estimator implicitly defined from nonlinear system of dimension $M+1$

Without filtered data: $\widehat{G}^{\varepsilon}_{N,J}(a,s) = 0 \;\Rightarrow\; \widehat{A}^{J}_{\text{eigen}}(\widetilde{X}^{\varepsilon}, N), \widehat{\Sigma}^{J}_{\text{eigen}}(\widetilde{X}^{\varepsilon}, N)$

$$\widehat{G}^{\varepsilon}_{N,J}(a,s) = \frac{1}{\Delta}\sum_{n=0}^{N-1}\sum_{j=1}^{J}\left\{\beta_j(\widetilde{X}^{\varepsilon}_n; a,s)\left(\phi_j(\widetilde{X}^{\varepsilon}_{n+1}; a,s) - e^{-\lambda_j(a,s)\Delta}\phi_j(\widetilde{X}^{\varepsilon}_n; a,s)\right)\right\}$$

With filtered data: $\widetilde{G}^{\varepsilon}_{N,J}(a,s) = 0 \;\Rightarrow\; \widetilde{A}^{J}_{\text{eigen}}(\widetilde{X}^{\varepsilon}, N), \widehat{\Sigma}^{J}_{\text{eigen}}(\widetilde{X}^{\varepsilon}, N)$

$$\widetilde{G}^{\varepsilon}_{N,J}(a,s) = \frac{1}{\Delta}\sum_{n=0}^{N-1}\sum_{j=1}^{J}\left\{\beta_j(\widetilde{Z}^{\varepsilon}_n; a,s)\left(\phi_j(\widetilde{X}^{\varepsilon}_{n+1}; a,s) - e^{-\lambda_j(a,s)\Delta}\phi_j(\widetilde{X}^{\varepsilon}_n; a,s)\right)\right\}$$

Why this definition? $\quad \mathbb{E}\left[\phi_j(\widetilde{X}^{0}_{n+1}; A,\Sigma)\middle|\widetilde{X}^{0}_n\right] = e^{-\lambda_j(A,\Sigma)\Delta}\phi_j(\widetilde{X}^{0}_n; A,\Sigma)$

$$\Longrightarrow \mathbb{E}\left[\widehat{G}^{0}_{N,J}(A,\Sigma)\right] = \mathbb{E}\left[\widetilde{G}^{0}_{N,J}(A,\Sigma)\right] = 0$$

Martingale estimating functions – Convergence analysis

Theorem (Without filtered data)

If Δ is independent of ε or $\Delta = \varepsilon^{\zeta}$ with $0 < \zeta < 1$ then

$$(i)\ \lim_{\varepsilon \to 0} \lim_{N \to \infty} \widehat{A}^{J}_{\text{eigen}}(X^{\varepsilon}, N) = A, \quad \textit{in probability}$$

$$(ii)\ \lim_{\varepsilon \to 0} \lim_{N \to \infty} \widehat{\Sigma}^{J}_{\text{eigen}}(X^{\varepsilon}, N) = \Sigma, \quad \textit{in probability}$$

Martingale estimating functions – Convergence analysis

Theorem (Without filtered data)

If Δ *is independent of* ε *or* $\Delta = \varepsilon^\zeta$ *with* $0 < \zeta < 1$ *then*

$$(i) \lim_{\varepsilon \to 0} \lim_{N \to \infty} \widehat{A}^J_{\text{eigen}}(X^\varepsilon, N) = A, \quad \textit{in probability}$$

$$(ii) \lim_{\varepsilon \to 0} \lim_{N \to \infty} \widehat{\Sigma}^J_{\text{eigen}}(X^\varepsilon, N) = \Sigma, \quad \textit{in probability}$$

Theorem (With filtered data)

If Δ *is independent of* ε *or* $\Delta = \varepsilon^\zeta$ *with* $\zeta > 0, \zeta \neq 1, \zeta \neq 2$ *then*

$$(i) \lim_{\varepsilon \to 0} \lim_{N \to \infty} \widetilde{A}^J_{\text{eigen}}(X^\varepsilon, N) = A, \quad \textit{in probability}$$

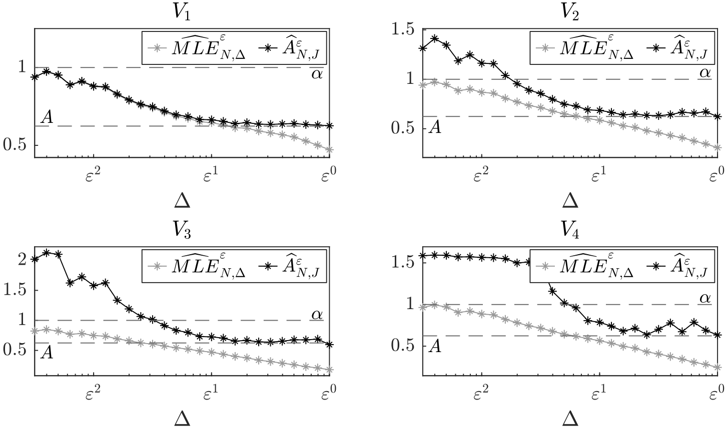
$$(ii) \lim_{\varepsilon \to 0} \lim_{N \to \infty} \widetilde{\Sigma}^J_{\text{eigen}}(X^\varepsilon, N) = \Sigma, \quad \textit{in probability}$$

Numerical experiments – Comparison estimators

Setting: drift estimation for polynomial potentials

$$V_1(x) = \frac{x^2}{2}, \quad V_2(x) = \frac{x^4}{4}, \quad V_3(x) = \frac{x^6}{6}, \quad V_4(x) = \frac{x^4}{4} - \frac{x^2}{2}$$

$$p(y) = \cos(y),\ \alpha = 1,\ \sigma = 1,\ \varepsilon = 0.1,\ T = 1000,\ J = 1$$

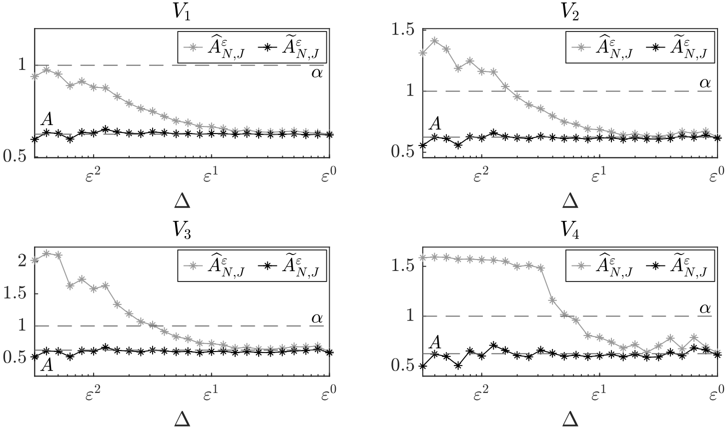


Numerical experiments – Comparison estimators

Setting: drift estimation for polynomial potentials

$$V_1(x) = \frac{x^2}{2}, \quad V_2(x) = \frac{x^4}{4}, \quad V_3(x) = \frac{x^6}{6}, \quad V_4(x) = \frac{x^4}{4} - \frac{x^2}{2}$$

$$p(y) = \cos(y), \ \alpha = 1, \ \sigma = 1, \ \varepsilon = 0.1, \ T = 1000, \ J = 1$$

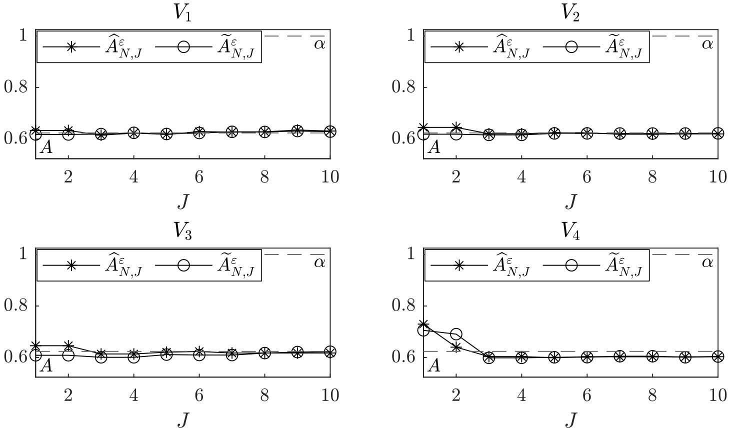


Numerical experiments – Sensitivity analysis w.r.t. J

Setting: drift estimation for polynomial potentials

$$V_1(x) = \frac{x^2}{2}, \quad V_2(x) = \frac{x^4}{4}, \quad V_3(x) = \frac{x^6}{6}, \quad V_4(x) = \frac{x^4}{4} - \frac{x^2}{2}$$

$$p(y) = \cos(y),\ \alpha = 1,\ \sigma = 1,\ \varepsilon = 0.1,\ T = 1000,\ \Delta = \varepsilon$$



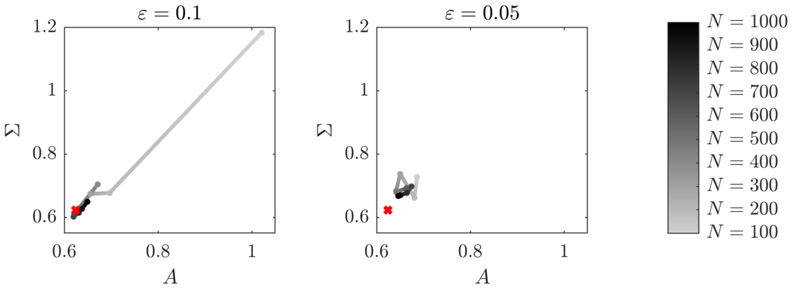
Numerical experiments – Multidimensional parameter

Setting: drift and diffusion estimation for Ornstein–Uhlenbeck process

$$V(x) = \frac{x^2}{2},\ p(y) = \cos(y),\ \alpha = 1,\ \sigma = 1$$

$$\varepsilon = 0.1,\ T = 1000,\ J = 2,\ \Delta = 1$$

Estimator $\left(\widehat{A}^J_{\text{eigen}}(X^\varepsilon, N), \widehat{\Sigma}^J_{\text{eigen}}(X^\varepsilon, N)\right)$ without filtered data



Conclusions

- We considered the problem of fitting effective dynamics from continuous and discrete multiscale data (Langevin SDE)

- We proposed novel estimators for the effective drift and diffusion coefficients based on filtered data and eigenpairs of the generator of the homogenized dynamics

- We proved the asymptotic unbiasedness of our estimators

- We showed numerically that our estimators outperform traditional techniques like subsampling both in terms of accuracy and robustness

References

Abdulle, A., Garegnani, G., Pavliotis, G. A., Stuart, A. M., and Zanoni, A. (2021). Drift estimation of multiscale diffusions based on filtered data. *Found. Comput. Math.*

Abdulle, A., Pavliotis, G. A., and Zanoni, A. (2022). Eigenfunction martingale estimating functions and filtered data for drift estimation of discretely observed multiscale diffusions. *Stat. Comput. (to appear)*.

Bensoussan, A., Lions, J.-L., and Papanicolaou, G. (1978). *Asymptotic analysis for periodic structures*. North-Holland Publishing Co., Amsterdam.

Garegnani, G. and Zanoni, A. (2021). Robust estimation of effective diffusions from multiscale data. arXiv preprint arXiv:2109.03132.

Kessler, M. and Sørensen, M. (1999). Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Bernoulli*, 5(2):299–314.

Pavliotis, G. A. and Stuart, A. M. (2007). Parameter estimation for multiscale diffusions. *J. Stat. Phys.*, 127(4):741–781.