

Variance Reduction in Equation-free Newton-Krylov-Methods

July 12, 2016



Model Problem

Algorithm

Systemic risk

Research Plan

Other

Fokker-Planck equation

$$\frac{\partial \rho(x, t)}{\partial t} + \frac{\partial (\mu(x) \rho(x, t))}{\partial x} = \frac{\sigma^2}{2} \frac{\partial^2 \rho(x, t)}{\partial x^2}$$

Simulate an ensemble of N particles evolving according to the corresponding SDE

$$dX = \mu f(X)dt + \sigma dW_t.$$

Coarse time stepper

$$\rho(t + n\Delta t) = \Phi_T(\rho) = (\mathcal{R} \circ \mathcal{E}(n\Delta t, \omega) \circ \mathcal{L}(\omega))(\rho(t))$$

Coarse time stepper

$$\rho(t + n\Delta t) = \Phi_T(\rho) = (\mathcal{R} \circ \mathcal{E}(n\Delta t, \omega) \circ \mathcal{L}(\omega))(\rho(t))$$

- Lifting $\mathcal{L} : \rho \rightarrow \mathbf{x}$

Coarse time stepper

$$\rho(t + n\Delta t) = \Phi_T(\rho) = (\mathcal{R} \circ \mathcal{E}(n\Delta t, \omega) \circ \mathcal{L}(\omega))(\rho(t))$$

- Lifting $\mathcal{L} : \rho \rightarrow \mathbf{x}$

Coarse time stepper

$$\rho(t + n\Delta t) = \Phi_T(\rho) = (\mathcal{R} \circ \mathcal{E}(n\Delta t, \omega) \circ \mathcal{L}(\omega))(\rho(t))$$

- ▶ Lifting $\mathcal{L} : \rho \rightarrow \mathbf{x}$
- ▶ Evolution \mathcal{E} :
Simulation of the SDE for N particles over n timesteps.

$$\mathbf{X}^{n+1} = \mathbf{X}^n + \mu(\mathbf{X}^n)\Delta t + \sqrt{\sigma\Delta t} \cdot \xi^n$$

with

$$\xi^n(\omega) \sim \mathcal{N}(0, 1)$$

Coarse time stepper

$$\rho(t + n\Delta t) = \Phi_T(\rho) = (\mathcal{R} \circ \mathcal{E}(n\Delta t, \omega) \circ \mathcal{L}(\omega))(\rho(t))$$

- ▶ Lifting $\mathcal{L} : \rho \rightarrow \mathbf{x}$
- ▶ Evolution \mathcal{E} :
Simulation of the SDE for N particles over n timesteps.

$$\mathbf{X}^{n+1} = \mathbf{X}^n + \mu(\mathbf{X}^n)\Delta t + \sqrt{\sigma\Delta t} \cdot \xi^n$$

with

$$\xi^n(\omega) \sim \mathcal{N}(0, 1)$$

Coarse time stepper

$$\rho(t + n\Delta t) = \Phi_T(\rho) = (\mathcal{R} \circ \mathcal{E}(n\Delta t, \omega) \circ \mathcal{L}(\omega))(\rho(t))$$

- ▶ Lifting $\mathcal{L} : \rho \rightarrow \mathbf{x}$
- ▶ Evolution \mathcal{E} :
Simulation of the SDE for N particles over n timesteps.

$$\mathbf{X}^{n+1} = \mathbf{X}^n + \mu(\mathbf{X}^n)\Delta t + \sqrt{\sigma\Delta t} \cdot \xi^n$$

with

$$\xi^n(\omega) \sim \mathcal{N}(0, 1)$$

- ▶ Restriction $\mathcal{R} : \mathbf{x} \rightarrow \rho$

$$\frac{1}{N} \sum_{i=1}^N w_i \cdot \chi_{\Delta_j}(X_i) = \rho_j$$

Evaluating the bias on Jacobian-vector products

$$\text{Bias}(\hat{\mathbf{J}}\mathbf{v}, \mathbf{J}\mathbf{v}_{FP}) = \hat{\mathbb{E}}[\hat{\mathbf{J}}\mathbf{v}] - \mathbf{J}\mathbf{v}_{FP}$$

Evaluating the bias on Jacobian-vector products

$$\text{Bias}(\hat{\mathbf{J}}\mathbf{v}, \mathbf{J}\mathbf{v}_{FP}) = \hat{\mathbb{E}}[\hat{\mathbf{J}}\mathbf{v}] - \mathbf{J}\mathbf{v}_{FP}$$

with

$$\mathbf{J}\mathbf{v}_{FP} \approx \frac{\rho^\varepsilon(t + n\Delta t) - \rho(t + n\Delta t)}{\varepsilon}$$

Evaluating the bias on Jacobian-vector products

$$\text{Bias}(\hat{\mathbf{J}}\mathbf{v}, \mathbf{J}\mathbf{v}_{FP}) = \hat{\mathbb{E}}[\hat{\mathbf{J}}\mathbf{v}] - \mathbf{J}\mathbf{v}_{FP}$$

with

$$\mathbf{J}\mathbf{v}_{FP} \approx \frac{\rho^\varepsilon(t + n\Delta t) - \rho(t + n\Delta t)}{\varepsilon}$$

and ρ calculated by explicitly solving

$$\frac{\partial \rho(x, t)}{\partial t} + \frac{\partial(\mu(x)\rho(x, t))}{\partial x} = D \frac{\partial^2 \rho(x, t)}{\partial x^2}$$

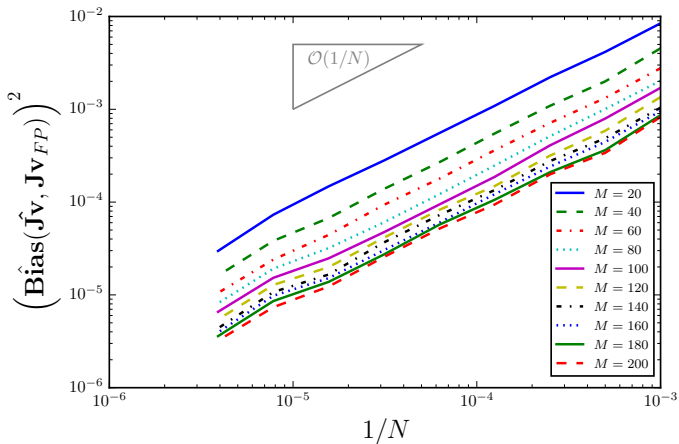
using

$$\rho_i^{n+1} = \rho_i^n + \Delta t \left(\frac{D}{\Delta x^2} (\rho_{i+1}^n - 2\rho_i^n + \rho_{i-1}^n) - \frac{a(x)}{\Delta x} (\rho_i^n - \rho_{i-1}^n) \right)$$

for the value of ρ at position $x = i\Delta x$ and time $t = (n+1)\Delta t$

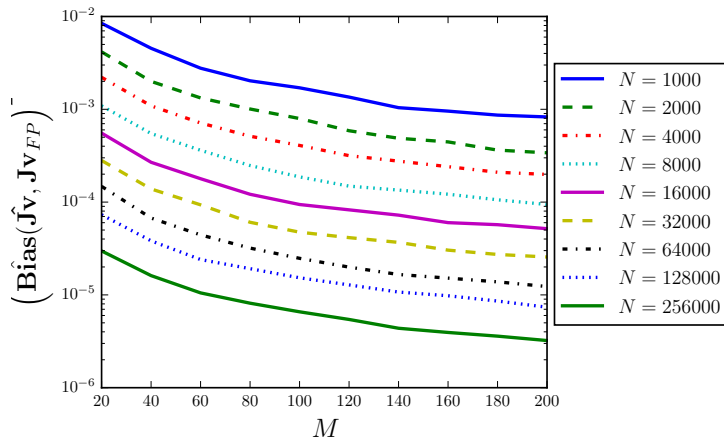
Evaluating the bias on Jacobian-vector products

$$\text{Bias}(\hat{\mathbf{J}}\mathbf{v}, \mathbf{J}\mathbf{v}_{FP}) = \hat{\mathbb{E}}[\hat{\mathbf{J}}\mathbf{v}] - \mathbf{J}\mathbf{v}_{FP}$$



Evaluating the bias on Jacobian-vector products

$$\text{Bias}(\hat{\mathbf{J}}\mathbf{v}, \mathbf{J}\mathbf{v}_{FP}) = \hat{\mathbb{E}}[\hat{\mathbf{J}}\mathbf{v}] - \mathbf{J}\mathbf{v}_{FP}$$



Newton-Krylov solver

Computing steady states ρ^* by solving the non-linear system

$$F(\rho^*) = \rho^* - \Phi_T(\rho^*) = 0$$

Newton-Krylov solver

Computing steady states ρ^* by solving the non-linear system

$$F(\rho^*) = \rho^* - \Phi_T(\rho^*) = 0$$

- ▶ Starting from an initial state ρ^0 , we iterate

$$\left\{ \begin{array}{l} \text{Solve } J(\rho^k)\delta_k = -F(\rho^k) \\ \text{Set } \rho^{k+1} = \rho^k + \delta_k \end{array} \right.$$

until convergence

Newton-Krylov solver

Computing steady states ρ^* by solving the non-linear system

$$F(\rho^*) = \rho^* - \Phi_T(\rho^*) = 0$$

- ▶ Starting from an initial state ρ^0 , we iterate

$$\begin{cases} \text{Solve } J(\rho^k)\delta_k = -F(\rho^k) \\ \text{Set } \rho^{k+1} = \rho^k + \delta_k \end{cases}$$

until convergence

- ▶ No explicit formula for $J(\Phi_T) \Rightarrow$ using iterative method (GMRES) that only requires Jacobian-vector products

Newton-Krylov solver

Computing steady states ρ^* by solving the non-linear system

$$F(\rho^*) = \rho^* - \Phi_T(\rho^*) = 0$$

- ▶ Starting from an initial state ρ^0 , we iterate

$$\begin{cases} \text{Solve } J(\rho^k)\delta_k = -F(\rho^k) \\ \text{Set } \rho^{k+1} = \rho^k + \delta_k \end{cases}$$

until convergence

- ▶ No explicit formula for $J(\Phi_T) \Rightarrow$ using iterative method (GMRES) that only requires Jacobian-vector products
- ▶ These are estimated by a finite difference approximation

$$J(\Phi_T) \cdot \mathbf{v} \approx \frac{\Phi_T(\rho + \varepsilon \mathbf{v}, \omega_1) - \Phi_T(\rho, \omega_2)}{\varepsilon}$$

Newton-Krylov solver

Computing steady states ρ^* by solving the non-linear system

$$F(\rho^*) = \rho^* - \Phi_T(\rho^*) = 0$$

- ▶ Starting from an initial state ρ^0 , we iterate

$$\begin{cases} \text{Solve } J(\rho^k)\delta_k = -F(\rho^k) \\ \text{Set } \rho^{k+1} = \rho^k + \delta_k \end{cases}$$

until convergence

- ▶ No explicit formula for $J(\Phi_T) \Rightarrow$ using iterative method (GMRES) that only requires Jacobian-vector products
- ▶ These are estimated by a finite difference approximation

$$J(\Phi_T) \cdot \mathbf{v} \approx \frac{\Phi_T(\rho + \varepsilon \mathbf{v}, \omega_1) - \Phi_T(\rho, \omega_2)}{\varepsilon}$$

Problem

$\text{Var}(\hat{\mathbf{J}}\mathbf{v}) \sim \mathcal{O}(1/(\varepsilon^2 N)) \Rightarrow$ Numerical noise for $\varepsilon \ll 1$

Variance reduction of Jacobian-vector products

$$\begin{aligned} \mathbf{J}\mathbf{v} = D(\Phi_T) \cdot \mathbf{v} &\approx \frac{\Phi_T(\rho + \varepsilon\mathbf{v}, \omega_1) - \Phi_T(\rho, \omega_2)}{\varepsilon} \\ &\approx \frac{\Phi_T(\rho, \omega_1) + \varepsilon D(\Phi_T)(\rho, \omega_1) \cdot \mathbf{v} - \Phi_T(\rho, \omega_2)}{\varepsilon} \end{aligned}$$

Variance reduction of Jacobian-vector products

$$\begin{aligned} \mathbf{J}\mathbf{v} = D(\Phi_T) \cdot \mathbf{v} &\approx \frac{\Phi_T(\rho + \varepsilon\mathbf{v}, \omega_1) - \Phi_T(\rho, \omega_2)}{\varepsilon} \\ &\approx \frac{\Phi_T(\rho, \omega_1) + \varepsilon D(\Phi_T)(\rho, \omega_1) \cdot \mathbf{v} - \Phi_T(\rho, \omega_2)}{\varepsilon} \end{aligned}$$

Solution:

Perturbations on the density \rightarrow perturbations in the weights

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N w_i \cdot \chi_{\Delta_j}(x_i) &= \rho_j \\ \frac{1}{N} \sum_{i=1}^N w_\varepsilon^i \cdot \chi_{\Delta_j}(x^i) &= \rho_j + \varepsilon v_j. \end{aligned}$$

For the perturbed density: compute the weight per bin as $w_\varepsilon^j = 1 + \frac{\varepsilon v_j}{\rho_j}$ and assign this value to each particle in Δ_j .

Variance reduction of Jacobian-vector products

$$\begin{aligned} \mathbf{J}\mathbf{v} = D(\Phi_T) \cdot \mathbf{v} &\approx \frac{\Phi_T(\rho + \varepsilon\mathbf{v}, \omega_1) - \Phi_T(\rho, \omega_2)}{\varepsilon} \\ &\approx \frac{\cancel{\Phi_T(\rho, \omega_1)} + \varepsilon D(\Phi_T)(\rho, \omega_1) \cdot \mathbf{v} - \cancel{\Phi_T(\rho, \omega_1)}}{\varepsilon} \end{aligned}$$

Solution:

Perturbations on the density \rightarrow perturbations in the weights

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N w_i \cdot \chi_{\Delta_j}(x_i) &= \rho_j \\ \frac{1}{N} \sum_{i=1}^N w_\varepsilon^i \cdot \chi_{\Delta_j}(x^i) &= \rho_j + \varepsilon v_j. \end{aligned}$$

For the perturbed density: compute the weight per bin as $w_\varepsilon^j = 1 + \frac{\varepsilon v_j}{\rho_j}$ and assign this value to each particle in Δ_j .

Variance reduction of Jacobian-vector products

$$\begin{aligned} \mathbf{J}\mathbf{v} = D(\Phi_T) \cdot \mathbf{v} &\approx \frac{\Phi_T(\rho + \varepsilon \mathbf{v}, \omega_1) - \Phi_T(\rho, \omega_2)}{\varepsilon} \\ &\approx \frac{\cancel{\Phi_T(\rho, \omega_1)} + \varepsilon D(\Phi_T)(\rho, \omega_1) \cdot \mathbf{v} - \cancel{\Phi_T(\rho, \omega_1)}}{\varepsilon} \end{aligned}$$

Solution:

Perturbations on the density \rightarrow perturbations in the weights

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N w_i \cdot \chi_{\Delta_j}(x_i) &= \rho_j \\ \frac{1}{N} \sum_{i=1}^N w_\varepsilon^i \cdot \chi_{\Delta_j}(x^i) &= \rho_j + \varepsilon v_j. \end{aligned}$$

For the perturbed density: compute the weight per bin as $w_\varepsilon^j = 1 + \frac{\varepsilon v_j}{\rho_j}$ and assign this value to each particle in Δ_j .

Variance reduction of Jacobian-vector products

$$\begin{aligned} \mathbf{J}\mathbf{v} = D(\Phi_T) \cdot \mathbf{v} &\approx \frac{\Phi_T(\rho + \varepsilon \mathbf{v}, \omega_1) - \Phi_T(\rho, \omega_2)}{\varepsilon} \\ &\approx \frac{\cancel{\Phi_T(\rho, \omega_1)} + \cancel{D(\Phi_T)(\rho, \omega_1) \cdot \mathbf{v}} - \cancel{\Phi_T(\rho, \omega_1)}}{\cancel{\varepsilon}} \end{aligned}$$

Solution:

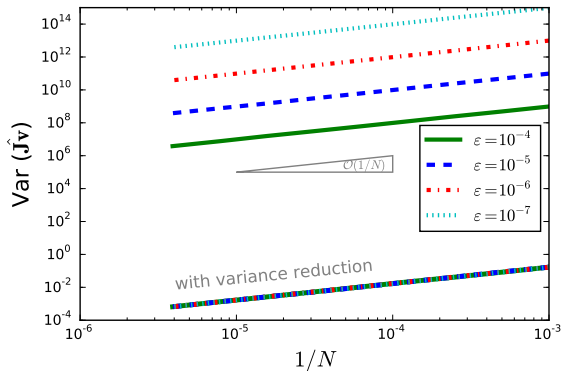
Perturbations on the density \rightarrow perturbations in the weights

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N w_i \cdot \chi_{\Delta_j}(x_i) &= \rho_j \\ \frac{1}{N} \sum_{i=1}^N w_\varepsilon^i \cdot \chi_{\Delta_j}(x^i) &= \rho_j + \varepsilon v_j. \end{aligned}$$

For the perturbed density: compute the weight per bin as $w_\varepsilon^j = 1 + \frac{\varepsilon v_j}{\rho_j}$ and assign this value to each particle in Δ_j .

Variance reduction of Jacobian-vector products

$$\text{Var}(\hat{\mathbf{J}}\mathbf{v}) = \hat{\mathbb{E}} \left[\left(\hat{\mathbf{J}}\mathbf{v} - \hat{\mathbb{E}}[\hat{\mathbf{J}}\mathbf{v}] \right)^2 \right] \sim \mathcal{O}(1/N)$$



Mean Field Model

Interaction between components

- ▶ Adding *mean field interaction* to model: each particle feels an attractive force towards the mean state (each agent tends to follow the state of the majority)

$$dX_j = \mu f(X_j)dt + \sigma dW_j + \alpha(\bar{X} - X_j)dt$$

- ▶ Interconnectedness between agents can affect the stability of the whole system

Mean Field Model

Interaction between components

- ▶ Adding *mean field interaction* to model: each particle feels an attractive force towards the mean state (each agent tends to follow the state of the majority)

$$dX_j = \mu f(X_j)dt + \sigma dW_j + \alpha(\bar{X} - X_j)dt$$

- ▶ Interconnectedness between agents can affect the stability of the whole system

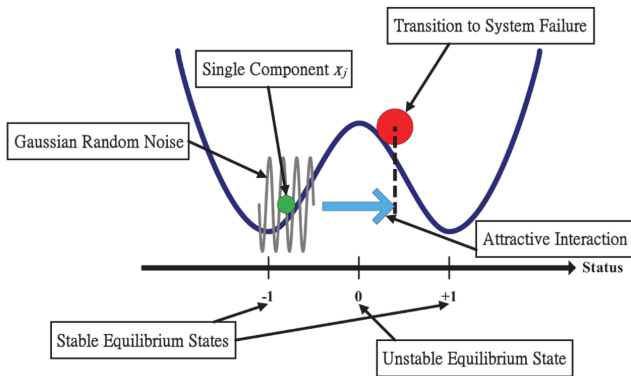
Application: Systemic Risk in Banking Systems

- ▶ Banks cooperate. By spreading the risk of credit shocks, they try to minimize their own risk.
- ▶ However, this increases the risk that they may all fail

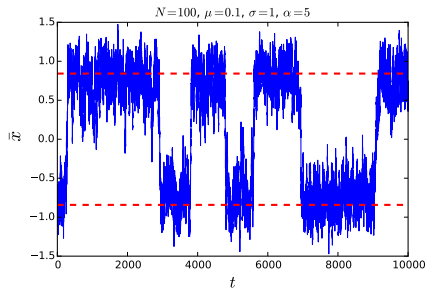
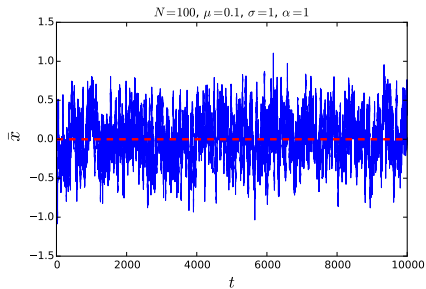
Mathematical Model for Systemic Risk

$$dX_j = \mu f(X_j)dt + \sigma dW_j + \alpha(\bar{X} - X_j)dt$$

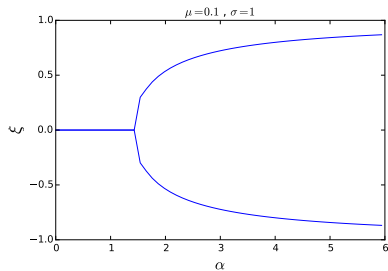
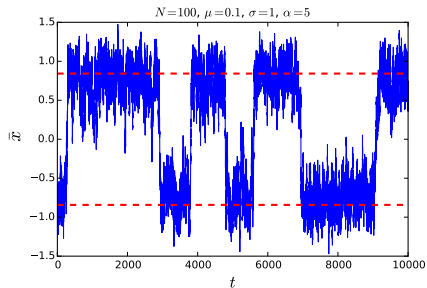
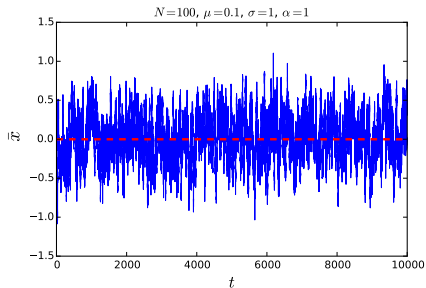
- ▶ μ The intrinsic stability of each component
- ▶ σ The strength of external random perturbations to the system
- ▶ α The degree of interconnectedness between agents



Metastable Coarse States



Metastable Coarse States



Analytic Solution for Equilibrium Distribution

$$\frac{\partial \rho}{\partial t} = -\mu \frac{\partial(f(x)\rho)}{\partial x} - \alpha \frac{\partial}{\partial x} \left[\left(\int x \rho dx - x \right) \rho \right] + \frac{\sigma^2}{2} \frac{\partial^2 \rho}{\partial x^2}.$$

Analytic Solution for Equilibrium Distribution

$$\frac{\partial \rho}{\partial t} = -\mu \frac{\partial(f(x)\rho)}{\partial x} - \alpha \frac{\partial}{\partial x} \left[\left(\int x \rho dx - x \right) \rho \right] + \frac{\sigma^2}{2} \frac{\partial^2 \rho}{\partial x^2}.$$

Assuming that $\xi = \lim_{t \rightarrow \infty} \int x \rho(x, t) dx$, an equilibrium solution satisfies

$$-\mu \frac{\partial(f(x)\rho_\xi)}{\partial x} - \alpha \frac{\partial}{\partial x} [(\xi - x)\rho_\xi] + \frac{\sigma^2}{2} \frac{\partial^2 \rho_\xi}{\partial x^2} = 0$$

Analytic Solution for Equilibrium Distribution

$$\frac{\partial \rho}{\partial t} = -\mu \frac{\partial(f(x)\rho)}{\partial x} - \alpha \frac{\partial}{\partial x} \left[\left(\int x \rho dx - x \right) \rho \right] + \frac{\sigma^2}{2} \frac{\partial^2 \rho}{\partial x^2}.$$

Assuming that $\xi = \lim_{t \rightarrow \infty} \int x \rho(x, t) dx$, an equilibrium solution satisfies

$$-\mu \frac{\partial(f(x)\rho_\xi)}{\partial x} - \alpha \frac{\partial}{\partial x} [(\xi - x)\rho_\xi] + \frac{\sigma^2}{2} \frac{\partial^2 \rho_\xi}{\partial x^2} = 0$$

The non-zero solutions $\pm \xi$ are

$$\xi = \pm \sqrt{1 - 3 \frac{\sigma^2}{2\alpha}} \left(1 + \mu \frac{6}{\sigma^2} \left(\frac{\sigma^2}{2\alpha} \right)^2 \frac{1 - 2 \frac{\sigma^2}{2\alpha}}{1 - 3 \frac{\sigma^2}{2\alpha}} \right) + \mathcal{O}(\mu^2)$$

Calculate fixed points by applying variance reduced Newton-Krylov-solver

$$F(\rho^*) = \rho^* - \Phi_T(\rho^*) = 0$$

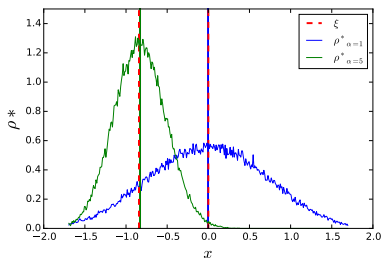
$$\left\{ \begin{array}{l} \text{Solve } J(\rho^k)\delta_k = -F(\rho^k) \\ \text{Set } \rho^{k+1} = \rho^k + \delta_k \end{array} \right.$$

Calculate fixed points by applying variance reduced Newton-Krylov-solver

$$F(\rho^*) = \rho^* - \Phi_T(\rho^*) = 0$$

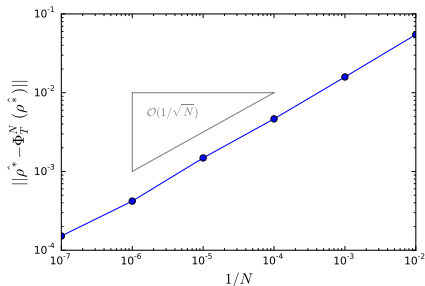
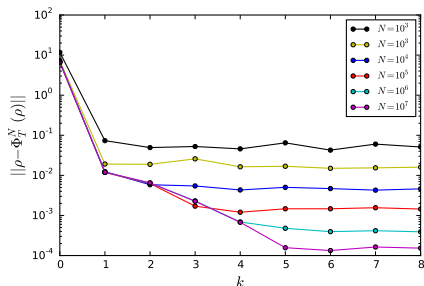
$$\left\{ \begin{array}{l} \text{Solve } J(\rho^k)\delta_k = -F(\rho^k) \\ \text{Set } \rho^{k+1} = \rho^k + \delta_k \end{array} \right.$$

- ▶ How to choose the Newton tolerance?
- ▶ Which time window to choose for the coarse time stepper Φ_T ?

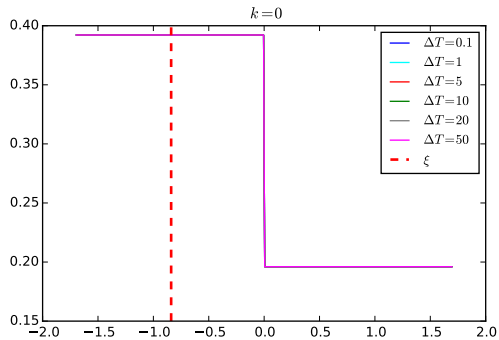


Estimating Stopping Criterion for the Non-linear Solver

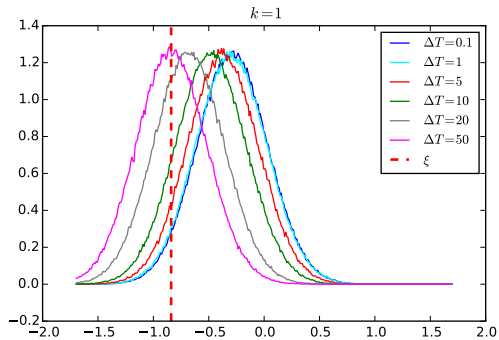
- ▶ The accuracy on the Newton-Krylov solution is inevitably limited by the noise on the stochastic coarse-time-stepper
- ▶ When the Newton-Krylov solution is converged, it stays oscillating around the true solution with a standard deviation depending on the number of particles.



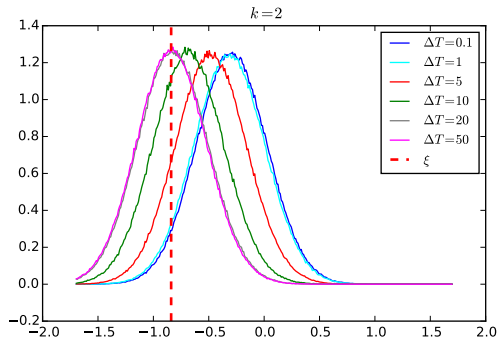
Estimating ΔT



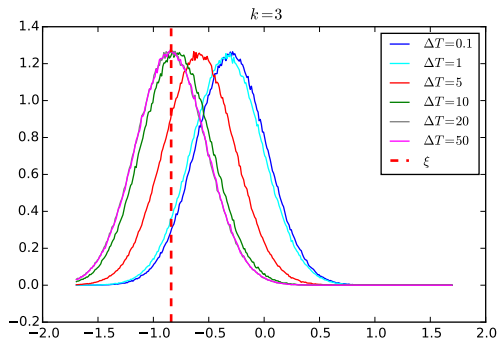
Estimating ΔT



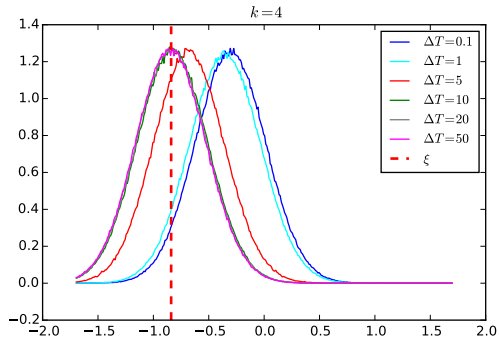
Estimating ΔT



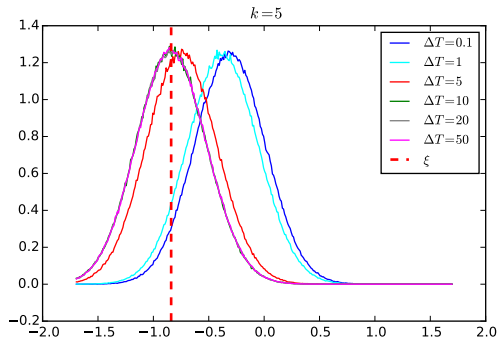
Estimating ΔT



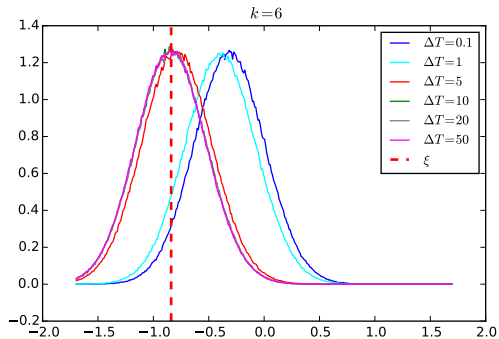
Estimating ΔT



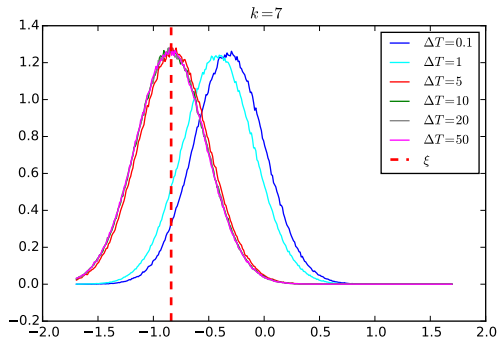
Estimating ΔT



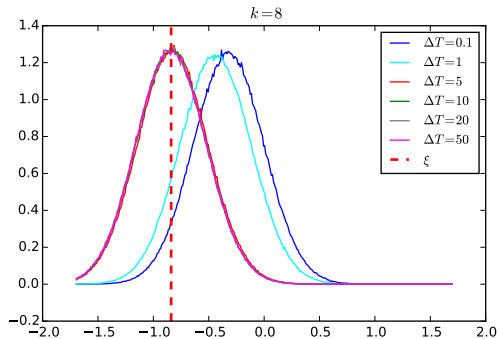
Estimating ΔT



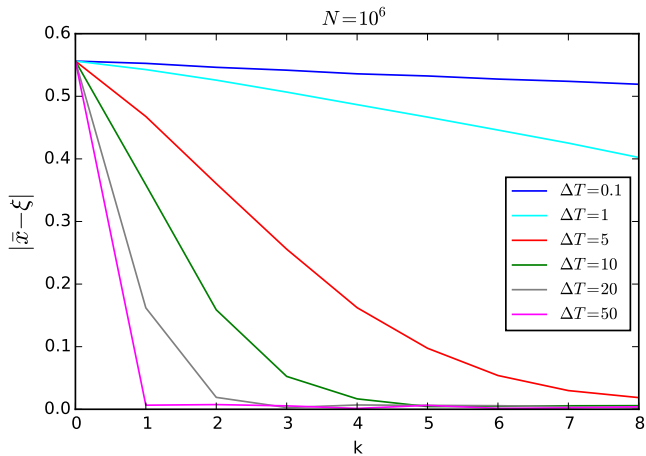
Estimating ΔT



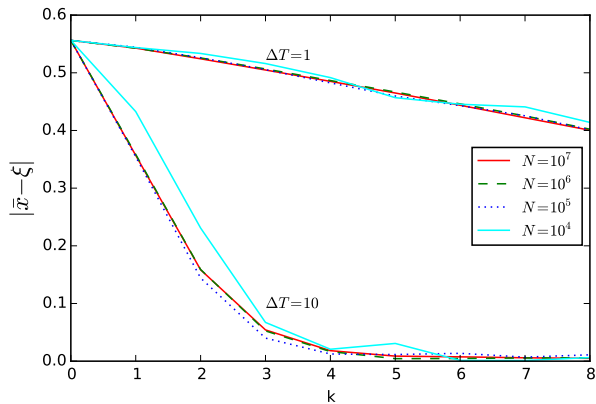
Estimating ΔT



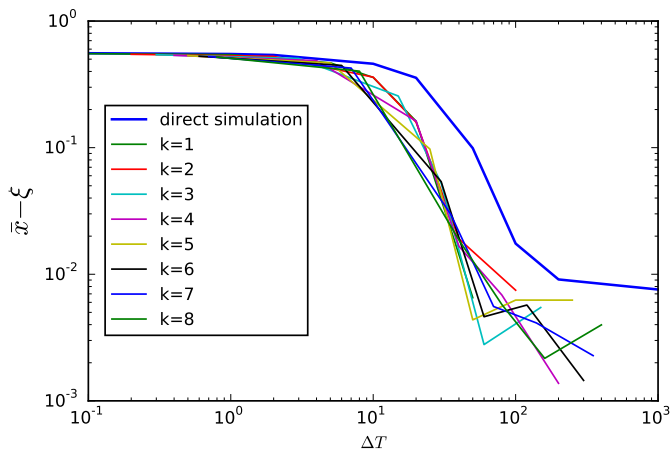
Estimating ΔT



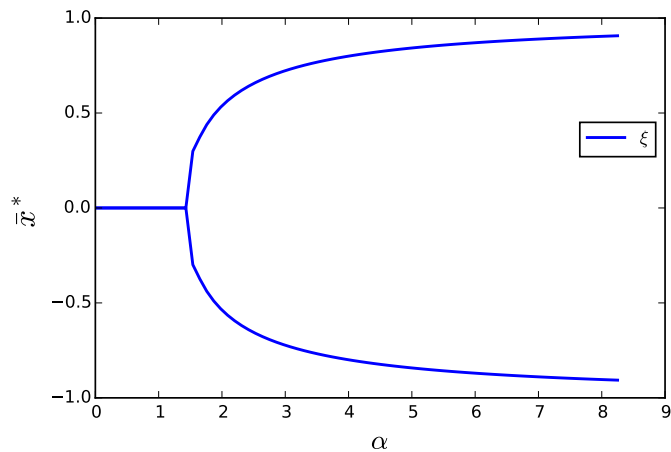
Estimating ΔT



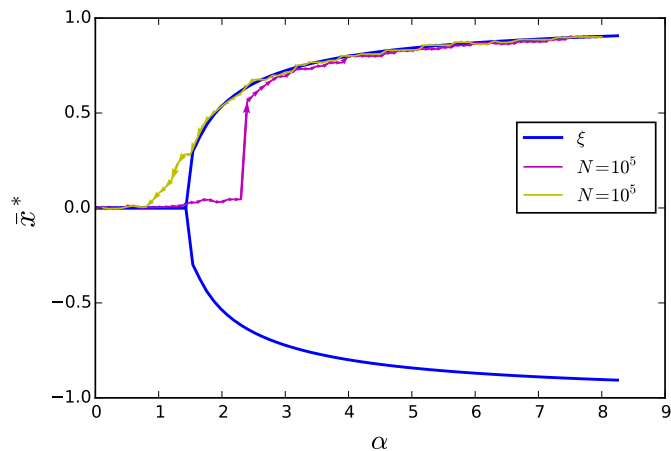
Efficiency compared with direct simulation



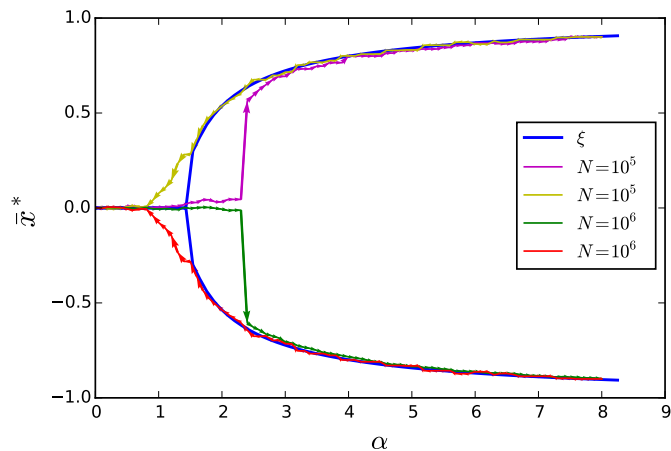
Bifurcation diagram



Bifurcation diagram



Bifurcation diagram



Courses

Contribution to Education

- ▶ Exercises Analysis 1
- ▶ Exercises Analysis 2

Doctoral Training Programme

- ▶ Seminar scientific integrity
- ▶ Teacher assistant training
- ▶ Science Communication and Outreach

Other courses

- ▶ Functional Analysis