Beyond infinite time scale separation

Edgeworth approximations for subgrid-scale parameterization

Jeroen Wouters [○,△] and Georg Gottwald [△]

[○] Meteorological Institute, University of Hamburg

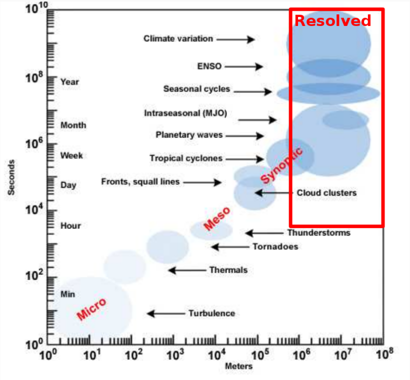
[△] School of Mathematics and Statistics, University of Sydney

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

THE UNIVERSITY OF SYDNEY

The model reduction problem

Many systems of scientific interest are to complex to simulate numerically.



E.g. climate models can resolve only part of the relevant processes of the climate system.

Can a dynamical system of lower dimensionality be determined that approximates the full system?

Approach: Model reduction through time scale separation

- Assume a time scale separation between slow variables x and fast variables y

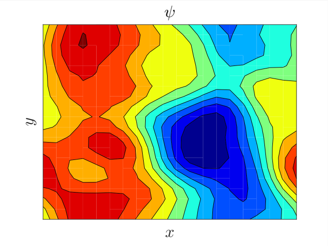
$$
\begin{cases}
\mathrm{d}x = \frac{1}{\varepsilon} f_0(x, y)\,\mathrm{d}t + f_1(x, y)\,\mathrm{d}t & \text{(resolved/slow/“climate”)} \\
\mathrm{d}y = \frac{1}{\varepsilon^2} g(x, y)\,\mathrm{d}t + \frac{1}{\varepsilon} \sigma(x, y)\,\mathrm{d}W & \text{(unresolved/fast/“weather”)}
\end{cases}
$$

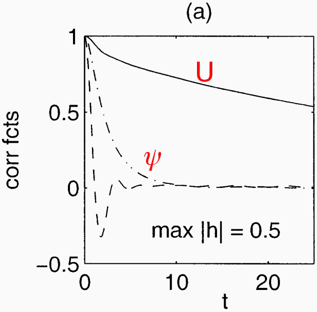
- As $\varepsilon \to 0$ the fast y variable decorrelates ever faster and acts as a Gaussian white noise on the slow variables and the slow x variable converges weakly to an SDE.
- This idea can be made mathematically rigourous by the method of homogenization

stochastic: Khasminsky ’66, Kurtz ’73, Papanicolaou ’76

deterministic: Melbourne & Stuart ’11, Gottwald & Melbourne ’13, Melbourne & Kelly ’15, De Simoi & Liverani ’14

Slow-fast systems in geophysics





Majda et al. (2003) JAS 60(14), 1705

Barotropic vorticity equation with topography

$$\begin{cases} \frac{\mathrm{d}U}{\mathrm{d}t} = \frac{1}{4\pi^2} \int h \frac{\partial \psi}{\partial x} \, \mathrm{d}x \, \mathrm{d}y \\ \frac{\partial q}{\partial t} + \nabla^{\perp}\psi \cdot \nabla q + U \frac{\partial \psi}{\partial x} + \beta \frac{\partial \psi}{\partial x} = 0 \\ q = \Delta\psi + h \end{cases}$$

The zonal mean flow U evolves slower than the fast Fourier modes $\psi_{i,j}$ of the stream function

This can be modeled by a system with a time scale separation parameter ε

$$\begin{cases} \frac{\mathrm{d}U}{\mathrm{d}t} & = \frac{1}{\varepsilon} f_1(\psi) \\ \frac{\mathrm{d}\psi_{i,j}}{\mathrm{d}t} & = \frac{1}{\varepsilon^2} g_2(\psi) + \frac{1}{\varepsilon} g_1(U) \end{cases}$$

Reduces through homogenization, assuming infinite time scale separation to

$$\mathrm{d}U = \alpha(U)\,\mathrm{d}t + \sigma(U)\,\mathrm{d}W$$

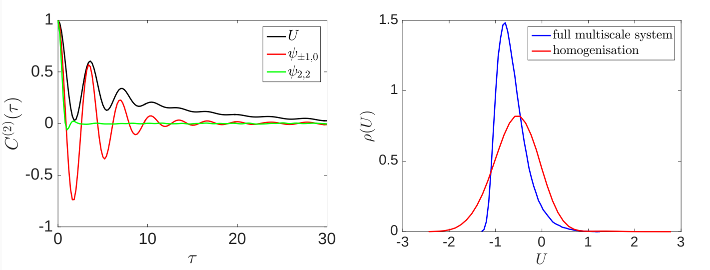
This can be modeled by a system with a time scale separation parameter ε

$$\begin{cases} \frac{\mathrm{d}U}{\mathrm{d}t} & = \frac{1}{\varepsilon} f_1(\psi) \\ \frac{\mathrm{d}\psi_{i,j}}{\mathrm{d}t} & = \frac{1}{\varepsilon^2} g_2(\psi) + \frac{1}{\varepsilon} g_1(U) \end{cases}$$

Reduces through homogenization, assuming infinite time scale separation to

$$\mathrm{d}U = \alpha(U)\,\mathrm{d}t + \sigma(U)\,\mathrm{d}W$$

Fails when only a moderate time scale separation



The Central Limit Theorem

Assume X_i are i.i.d. random variables

$$S_n := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (X_j - \mu) \rightarrow_d \mathcal{N}(0, \sigma^2)$$

where $\mu = \mathbb{E}[X_i]$ and $\sigma^2 = \mathbb{E}[X_i^2]$

For finite n, there are *deviations* to the CLT

These are described by the Edgeworth expansion

$$\rho_n(x) = \Phi_{0,\sigma^2}(x) \times \left(1 + \frac{1}{6\sqrt{n}} \frac{\gamma}{\sigma^3} H_3(x/\sigma)\right) + o(1/\sqrt{n})$$

where H_3 is the third Hermite polynomial and $\gamma = \mathbb{E}[X_i^3]$

Feller (1957) "An introduction to probability theory and its applications"

The CLT and the Edgeworth expansion (dependent version)

The Central Limit Theorem

Assume X_i are stationary weakly dependent random variables

$$S_n := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (X_j - \mu) \to_d \mathcal{N}(0, \sigma^2)$$

where $\mu = \mathbb{E}[X_i]$ and $\sigma^2 = \mathbb{E}[X_1^2] + 2\sum_{j=1}^{\infty} \mathbb{E}[X_1 X_{j+1}]$

For finite n, there are *deviations* to the CLT

These deviations are described by the Edgeworth expansion

$$\rho_n(x) = \Phi_{0,\sigma^2 + \delta\sigma^2/n}(x) \times \left(1 + \frac{1}{\sqrt{n}} \delta\kappa H_3(x/\sigma)\right) + o(1/\sqrt{n})$$

where H_3 is the third Hermite polynomial and $\delta\sigma^2$ and $\delta\kappa$ are sums over correlation functions of X_i

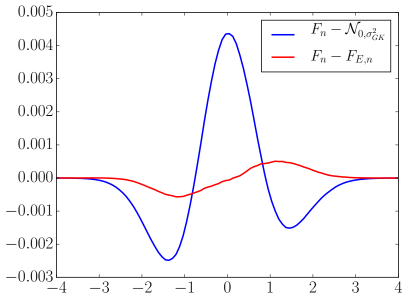
Götze and Hipp (1983) Z Wahrscheinlichkeit, 64, 211

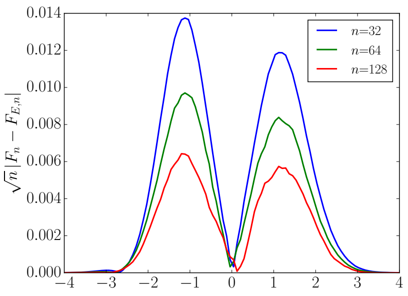
Edgeworth expansion in action: deterministic processes

Example: deterministic mod process

$$x_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} A(y_j) \qquad \text{with} \qquad \begin{aligned} y_{j+1} &= p y_j \mod 1 \\ A(y) &= y^5 + y^4 + y^3 + y^2 + y - c \end{aligned}$$

We can calculate σ^2, $\delta\sigma^2$ and $\delta\kappa$ *explicitly*





Edgeworth $F_{E,n}$ is closer to the truth F_n than Gaussian $\mathcal{N}_{0,\sigma^2_{GK}}$

$n = 32, p = 3$

$|F_n - F_{E,n}|$ is $o(1/\sqrt{n})$

Edgeworth expansion for slow/fast systems

For slow-fast systems

$$\begin{cases} \dot{x} &= \frac{1}{\varepsilon} f_0(x,y) + f_1(x,y) \\ \dot{y} &= \frac{1}{\varepsilon^2} g_0(y) + \frac{1}{\varepsilon} g_1(x,y) \end{cases}$$

we have that $\frac{x(\varepsilon)-x(0)}{\sqrt{\varepsilon}}$ converges to a Gaussian as $\varepsilon \to 0$.

$$\rho_t(x(t)|x(0) = x_0) = \int \mathrm{d}y \, e^{\mathcal{L}t} \delta_{x_0}(x) \mu(\mathrm{d}y)$$

where $\mathcal{L} = \frac{1}{\varepsilon^2}\mathcal{L}_0 + \frac{1}{\varepsilon}\mathcal{L}_1 + \mathcal{L}_2$ with $\mathcal{L}_0 \rho = -\partial_y(g(y)\rho)$, $\mathcal{L}_1 \rho = -\partial_x(f_0(x,y)\rho)$ and $\mathcal{L}_2 \rho = -\partial_x(f_1(x,y)\rho)$ are generators

Edgeworth corrections to ρ_t can be calculated from a Dyson series for the transfer operator

$$e^{\mathcal{L}t} = e^{\mathcal{L}_0 t/\varepsilon^2} + \int_0^t ds \, e^{\mathcal{L}_0 (t-s)/\varepsilon^2} (\frac{1}{\varepsilon}\mathcal{L}_1 + \mathcal{L}_0) e^{\mathcal{L}_0 s/\varepsilon^2} + \dots$$

Stochastic parameterization using the Edgeworth expansion

Given a slow-fast dynamical system

$$\begin{cases} \dot{x} &= \frac{1}{\varepsilon} f_0(y) + f_1(x, y) \\ \dot{y} &= \frac{1}{\varepsilon^2} g(y) \end{cases}$$

1. determine the Edgeworth expansion coefficients $\sigma^2_{\mathrm{GK}}, \delta\kappa$ associated with $f_0(x, y)$
2. model x of the multi-scale system by X of a surrogate stochastic process

$$\begin{cases} \dot{X} &= \frac{1}{\varepsilon} A(\eta) + F(x) \\ \mathrm{d}\eta &= -\frac{\gamma}{\varepsilon^2}\,\mathrm{d}t + \frac{1}{\varepsilon}\,\mathrm{d}W \end{cases}$$

with $A(\eta) = a\eta^2 + b\eta + c$, where a, b, c, γ are determined such that the Edgeworth expansion coefficients of $A(\eta)$ match those of f_0 in the true system.

Application: parameterization of a discrete-time multiscale system

$$\begin{cases} x_{j+1}^{(\varepsilon)} &= x_j^{(\varepsilon)} + \varepsilon f_0(y_j) + \varepsilon^2 f_1(x_j^{(\varepsilon)}) \\ y_{j+1} &= p y_j \mod 1 \end{cases}$$

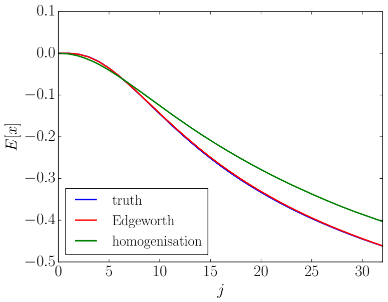
homogenization: converges for $\varepsilon \to 0$ to a diffusion (Gottwald & Melbourne (2013))

$$\mathrm{d}X = f_1(X)\,\mathrm{d}t + \sigma_{GK}\,\mathrm{d}W$$

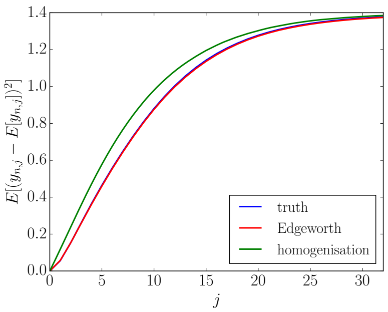
Edgeworth: replace fast mod map y by an AR1 process η

$$\begin{cases} X_{j+1}^{(\varepsilon)} &= X_j^{(\varepsilon)} + \varepsilon f_{0,s}(\eta_j) + \varepsilon^2 f_1(X_j^{(\varepsilon)}) \\ \eta_{j+1} &= \phi \eta_j + N_j \end{cases}$$

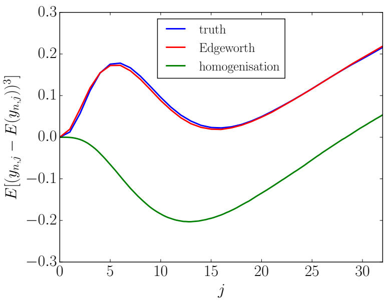
with $f_{0,s}(\eta) = a_3\eta^3 + a_2\eta^2 + a_1\eta + a_0$ and parameters a_i tuned to match Edgeworth corrections



mean



variance



third moment

Parameterization of a continuous-time multiscale system

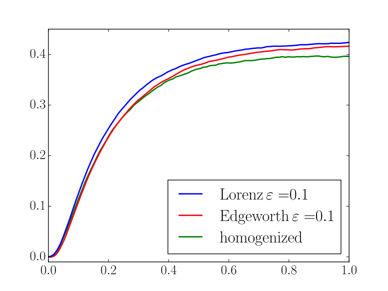
$$\begin{cases} \dot{x} &= \frac{1}{\varepsilon} f_0(y) + f_1(x) \\ \dot{y} &= \frac{1}{\varepsilon^2} g(y) \end{cases}$$

where $f_1(x) = -\nabla V(x)$, with $V(X)$ an assymetric double well potential and $\dot{y} = g(y)$ is the standard Lorenz '63 system.

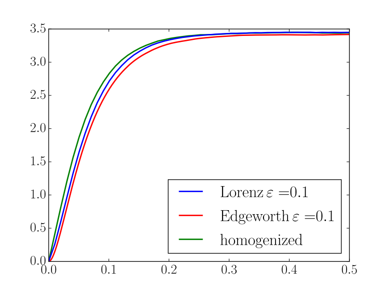
Edgeworth: replace fast Lorenz system y by an Ornstein-Uhlenbeck process η

$$\begin{cases} \dot{X} = \frac{1}{\varepsilon} f_{0,s}(\eta) + f_1(X) \\ \mathrm{d}\eta = -\frac{1}{\varepsilon^2} \gamma \eta \, \mathrm{d}t + \frac{1}{\varepsilon} \, \mathrm{d}W \end{cases}$$

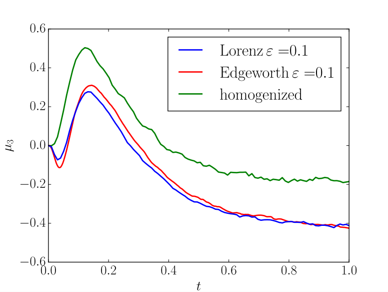
with $f_{0,s}(\eta) = a_3 \eta^3 + a_2 \eta^2 + a_1 \eta + a_0$ and parameters a_i tuned to match Edgeworth corrections



mean



variance



third moment

Summary

- We have used the Edgeworth expansion to extend the range of time scale separation over which slow-fast sytems can be approximated
- The fast variables are replaced by a stochastic surrogate process, the parameters of which are tuned to *match the Edgeworth expansion*
- We have shown good agreement when reducing deterministic discrete and continuous time systems
- To do: Apply Edgeworth based reduction to the barotropic vorticity equation

Summary

- We have used the Edgeworth expansion to extend the range of time scale separation over which slow-fast sytems can be approximated
- The fast variables are replaced by a stochastic surrogate process, the parameters of which are tuned to *match the Edgeworth expansion*
- We have shown good agreement when reducing deterministic discrete and continuous time systems
- To do: Apply Edgeworth based reduction to the barotropic vorticity equation

Thank you for your attention!

homogenization extends the Central Limit Theorem

$$\begin{cases} dx^{(\varepsilon)} = \frac{1}{\varepsilon} f_0(y^{(\varepsilon)})\, dt & \textit{resolved/slow} \\ dy^{(\varepsilon)} = \frac{1}{\varepsilon^2} g(y^{(\varepsilon)})\, dt + \frac{1}{\varepsilon}\sigma\, dW & \textit{unresolved/fast} \end{cases}$$

The slow variable x *integrates* the fast variable y

$$\begin{aligned} x^{(\varepsilon)}(t) - x^{(\varepsilon)}(0) &= \frac{1}{\varepsilon}\int_0^t f_0(y^{(\varepsilon)}(s))\, ds \\ &= \varepsilon \int_0^{t/\varepsilon^2} f_0(y^{(\varepsilon=1)}(s))\, ds \\ &= \frac{1}{\sqrt{n}} \int_0^{tn} f_0(y^{(\varepsilon=1)}(s))\, ds \end{aligned}$$

Invoking the CLT, $x(t)$ converges weakly to $dX = \sigma\, dW$ where $\sigma^2 = 2\int_0^\infty \mathbb{E}[f_0(y^{(1)}(0)) f_0(y^{(1)}(s))]\, ds$

homogenization

$$\begin{cases} \mathrm{d}x = \frac{1}{\varepsilon} f_0(x,y)\,\mathrm{d}t + f_1(x,y)\,\mathrm{d}t & \textit{resolved/slow} \\ \mathrm{d}y = \frac{1}{\varepsilon^2} g(x,y)\,\mathrm{d}t + \frac{1}{\varepsilon}\sigma(x,y)\,\mathrm{d}W & \textit{unresolved/fast} \end{cases}$$

Assumptions:

- fast y-process is ergodic with measure μ_x
- $\int f_0(x,y)\,\mathrm{d}\mu_x = 0$

In the limit $\varepsilon \to 0$, the slow x-dynamics is approximated by

$$\mathrm{d}X = F(X)\,\mathrm{d}t + \Sigma(X)\,\mathrm{d}W$$

where

$$\Sigma\Sigma^T = 2\int_0^\infty \mathbb{E}^{\mu_x}\left[f_0(x,y) f_0(x,y(s))\right]\,\mathrm{d}s$$

$$F(X) = \int f_1(x,y)\;\mathrm{d}\mu_x + \int_0^\infty \int \nabla_x f_0(x,y(s)) f_0(x,y)\;\mathrm{d}\mu_x\;\mathrm{d}s$$

stochastic: Khasminsky '66, Kurtz '73, Papanicolaou '76

deterministic: Melbourne & Stuart '11, Gottwald & Melbourne '13, Melbourne & Kelly '15

“Proof” of the Edgeworth expansion

Expand the characteristic function of $X/\sqrt{n}$ (assuming $\mu = 0$, $\sigma = 1$):

$$\begin{aligned}
\mathbb{E}[e^{\imath tX/\sqrt{n}}] &= \mathbb{E}[1 + \frac{\imath tX}{\sqrt{n}} + \frac{(\imath t)^2 X^2}{2n} + \frac{(\imath t)^3 X^3}{6n\sqrt{n}} + \ldots] \\
&= (1 - \frac{t^2}{2n}) + \frac{(\imath t)^3}{6n\sqrt{n}}\mathbb{E}[X^3] + \ldots
\end{aligned}$$

The characteristic function of $\sum_{j=1}^{n} X_j/\sqrt{n}$

$$\begin{aligned}
\mathbb{E}[e^{\imath tX/\sqrt{n}}]^n &= (1 - \frac{t^2}{2n})^n + (1 - \frac{t^2}{2n})^{(n-1)}\frac{(\imath t)^3\gamma}{6\sqrt{n}} + \ldots \\
&= e^{-t^2/2}\left(1 + \frac{(\imath t)^3\gamma}{6\sqrt{n}}\right) + O(\frac{1}{n})
\end{aligned}$$

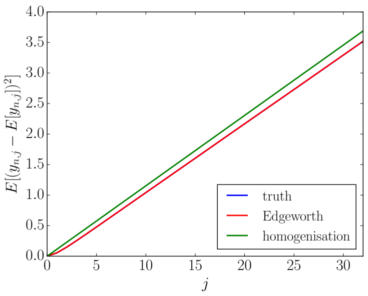
Application: stochastic approximation of a deterministic map

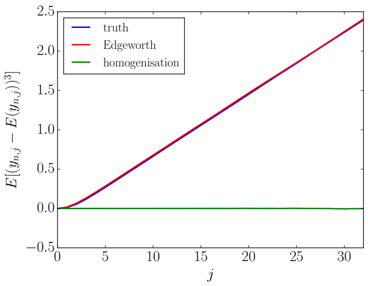
Replace $\begin{cases} x_{j+1} &= x_j + \varepsilon A(y_j) \\ y_{j+1} &= p y_j \mod 1 \end{cases}$ with $A(y) = y^5 + y^4 + y^3 + y^2 + y - c$

by a surrogate AR1 process

$$\begin{cases} X_{j+1} &= X_j + \varepsilon B(\eta_j) \\ \eta_{j+1} &= \phi \eta_j + N_j \end{cases} \text{ with } B(y) = a_s \eta^2 + b_s \eta + c_s$$

such that σ^2_{GK} (homogenization), as well as $\delta \kappa_3$ (1st Edgeworth term) match



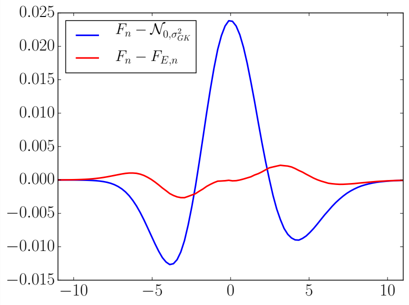


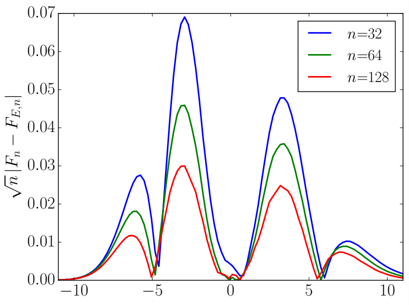
Edgeworth expansion in action: stochastic processes

Example 1: AR1 process

$$\begin{cases} x_{j+1} &= x_j + \varepsilon A(\eta_j) \\ \eta_{j+1} &= \phi \eta_j + N_j \end{cases} \quad \text{with} \quad \begin{cases} A(\eta) = a\eta^2 + b\eta + c \\ N_j \sim \mathcal{N}(0,1) \end{cases}$$

We can calculate σ^2, $\delta\sigma^2$ and $\delta\kappa$ *explicitly* (everything is Gaussian)





Edgeworth $F_{E,n}$ is closer to the truth F_n than Gaussian $\mathcal{N}_{0,\sigma^2_{GK}}$

$|F_n - F_{E,n}|$ is $o(1/\sqrt{n})$

$n = 32$, $\phi = 1/3$, $a = b = 1$

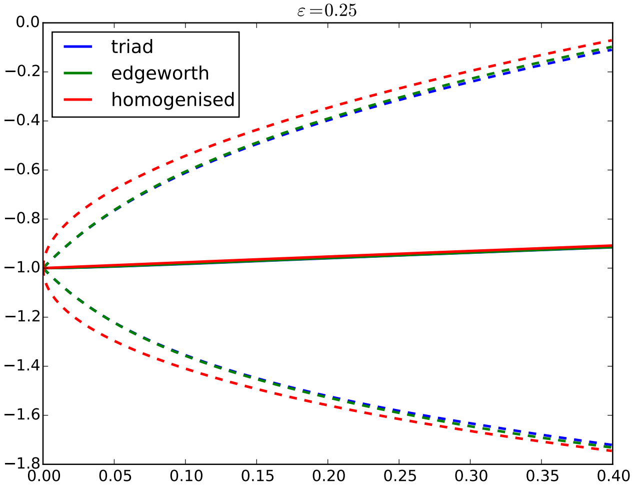
Triad system Majda et al (2001)

$$\begin{cases} \dot{x} &= \frac{1}{\varepsilon} B_0 y_1 y_2 \\ \dot{y}_1 &= \frac{1}{\varepsilon} B_1 y_2 x - \frac{1}{\varepsilon^2}\gamma_1 y_1 - \frac{1}{\varepsilon}\sigma_1 \dot{W}_1 \\ \dot{y}_2 &= \frac{1}{\varepsilon} B_2 x y_1 - \frac{1}{\varepsilon^2}\gamma_2 y_2 - \frac{1}{\varepsilon}\sigma_2 \dot{W}_2 \end{cases}$$

Edgeworth:

$$\begin{cases} \dot{X} &= \frac{1}{\varepsilon} A(\eta) \\ \dot{\eta} &= \frac{1}{\varepsilon}\alpha X - \frac{1}{\varepsilon^2}\eta - \frac{1}{\varepsilon}\sigma \dot{W} \end{cases}$$

with $A(\eta) = a_s \eta^2 + b_s \eta + c_s$



$\varepsilon = 0.25$
triad
edgeworth
homogenised
0.0
−0.2
−0.4
−0.6
−0.8
−1.0
−1.2
−1.4
−1.6
−1.8
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40