# Continuous-Discrete Filtering using the Duncan-Mortensen-Zakai (DMZ) Equation: Smooth Likelihood Surface

Hermann Singer
Department of Economics
FernUniversität in Hagen

Advanced Techniques for Longitudinal Data Analysis
in Social Science
Bielefeld, March 2023

March 9, 2023

# Continuous-Discrete Filtering

## Nonlinear bistable diffusion: Ginzburg-Landau model

$$
\begin{aligned}
dY &= -[\alpha Y + \beta Y^3]dt + \sigma dW(t) \\
&= -\nabla \Phi(Y) + \sigma dW(t) \\
z_i &= Y(t_i) + \epsilon_i
\end{aligned}
$$

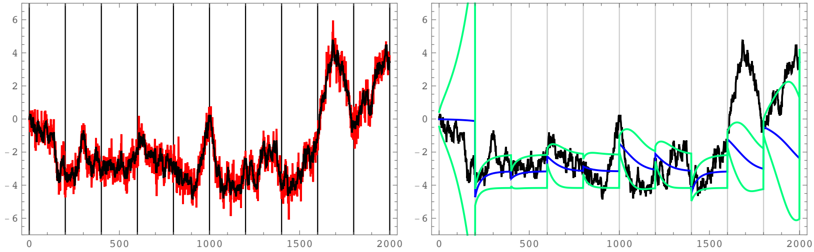


Figure: Simulated data (left) and extended Kalman filter (right).

# Applications

- phase transitions, superconductivity: Ginzburg and Landau (1950)
- economic equilibrium model: Herings (1996)
- oscillatory finger motions: Molenaar and Newell (2003)

Potentials (order parameter $y$)

$$\begin{aligned} \Phi(y) &\sim \frac{\alpha}{2} y^2 + \frac{\beta}{4} y^4 \\ V(y) &= -a\cos(y) - b\cos(2y) \\ &= (a-b) + \left(2b - \tfrac{a}{2}\right) y^2 + \tfrac{1}{24}(a - 16b)\, y^4 + O\left(y^5\right) \end{aligned}$$

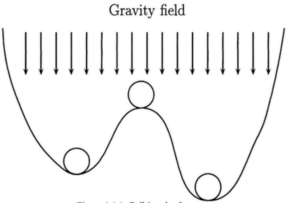


Figure 1.1.1. Ball in a landscape.

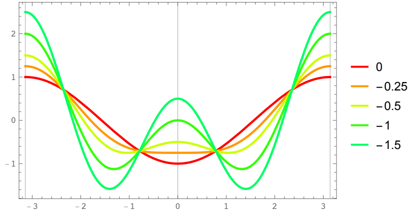


Figure: Equilibrium model (left), potential $V(y)$ (right) of a model for human hand movement. The value $b/a$ controls the number of stable minima. Below $b/a = -0.25$, only one minimum occurs.

# Particle filter: likelihood surface is not smooth

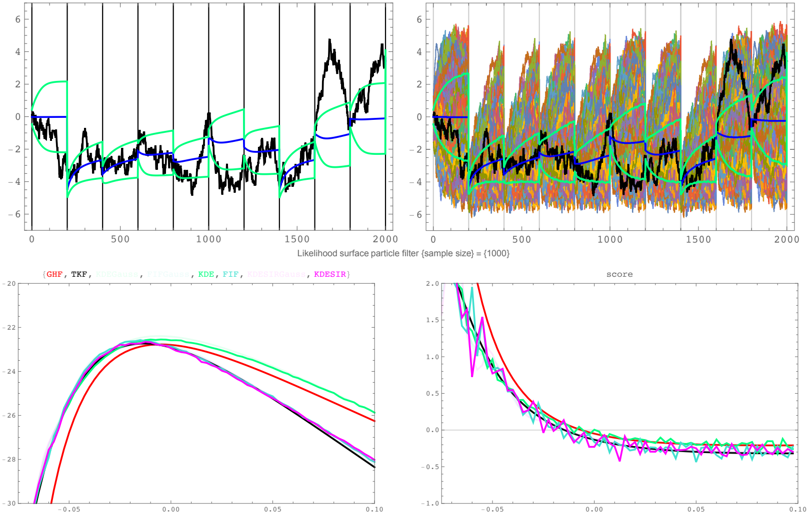


Figure: GHF (top, left). SIR particle filter (mean, SD and trajectories) (top, right), likelihood and score for $\beta$ (bottom). TKF (black line): matrix representation of Fokker-Planck operator Increment $d\beta = 0.0025$.

# Forward and backward simulation: Particle and Zakai filter

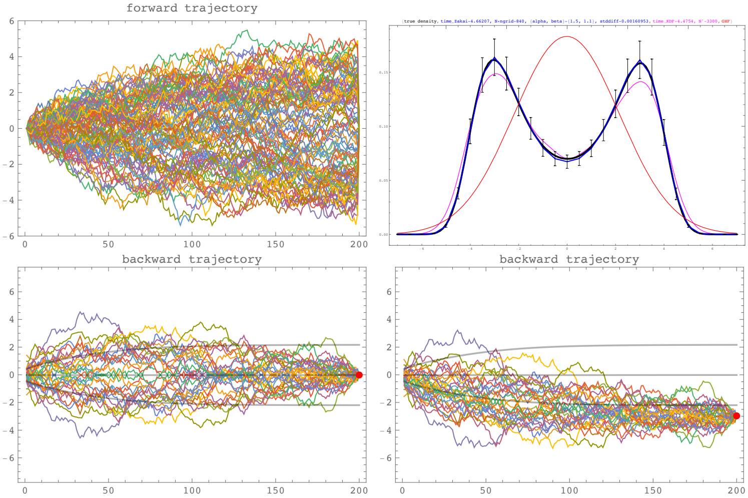


Figure: Forward simulation (top, left, $N = 3300$), estimated filter density (ZKF, SIR, GHF, top, right), backward simulation of DMZ equation with importance sampling ($N' * \text{ngrid} = 840$, bottom).

# Forward and backward simulation: Particle and Zakai filter

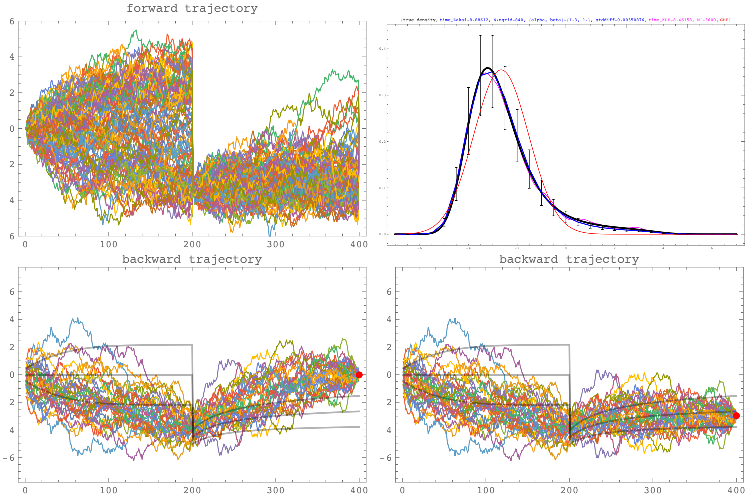


Figure: Forward simulation (top, left, $N = 3300$), estimated filter density (ZKF, SIR, GHF, top, right), backward simulation of DMZ equation with importance sampling ($N' * \text{ngrid} = 840$, bottom).

# Likelihood: Particle filter (top) and Zakai filter (bottom)

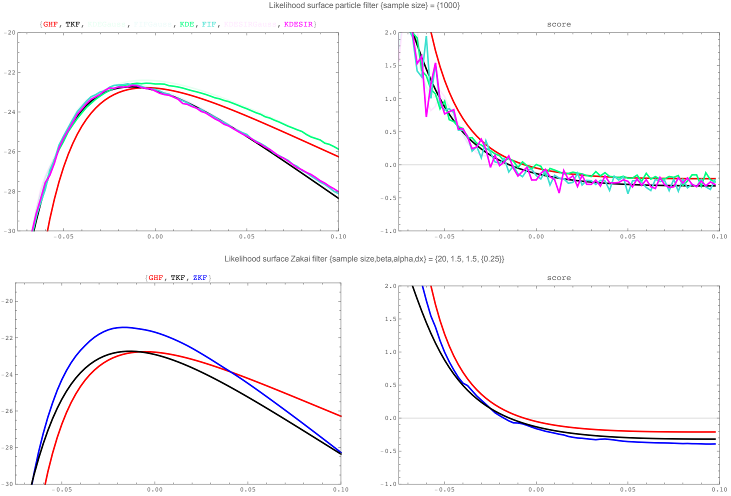


Figure: Likelihood and score for SIR particle filter ($N = 1000$, top) and ZKF (Riemann sample points, $N = 20$, bottom), GHF, TKF as reference. Increment $d\beta = 0.0025$.

# Likelihood: Zakai filter (UT sample points)

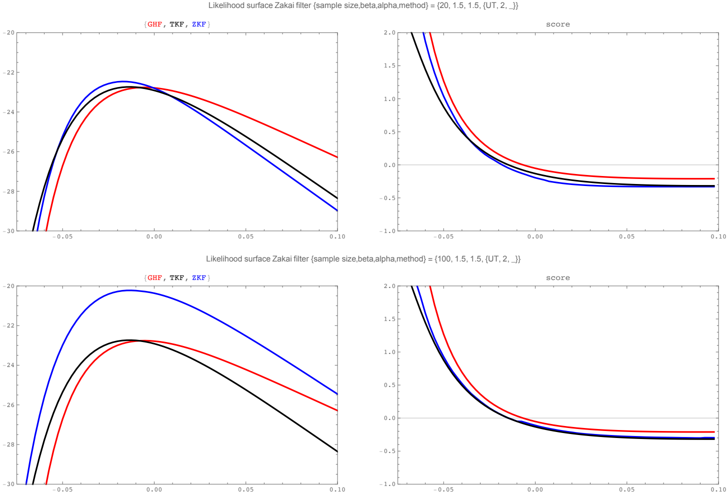


Figure: Likelihood and score for ZKF (unscented transform UT sample points), GHF, TKF. Sample size $N = 20$ (top), $N = 100$ (bottom). Increment $d\beta = 0.0025$.

# Continuous Time State Space Model

$$\begin{aligned} dY(t) &= f(Y,t)dt + G(Y,t)dW(t) \\ dZ(t) &= h(Y,t)dt + dV(t) \end{aligned}$$

sampled measurements:

$$z_i = (dZ/dt)(t_i) = h(Y(t_i), t_i) + \epsilon_i$$

- Goal: Optimal Filtering and Maximum Likelihood Estimation
- Wiener process $W(t), V(t)$
- Itô stochastic differential equations
- measurement error $\mathrm{Var}(dV(t)) = \rho(t)dt, \mathrm{Var}(\epsilon_i) = R(t_i)$
- scaling $\rho(t)/dt = R(t)$

# State estimation: Continuous-discrete Kalman Filter

- filter density $p(y,t|Z^t)$
- Fokker-Planck operator $F(y,t) = -\partial_\alpha f_\alpha + \frac{1}{2}\partial_\alpha\partial_\beta\Omega_{\alpha\beta}$

time update: Fokker-Planck equation $(t_i \leq t < t_{i+1})$

$$\partial_t p(y,t|Z^i) \quad = \quad F(y,t)p(y,t|Z^i)$$

measurement update: Bayes formula (new information $z_{i+1}$)

$$p(y,t_{i+1}|z_{i+1},Z^i) \quad = \quad \frac{p(z_{i+1},t_{i+1}|y)p(y,t_{i+1}|Z^i)}{p(z_{i+1},t_{i+1}|Z^i)}$$

# (some) Solution Methods

- Sequential (Kalman Filtering)
  - Moment based methods
    - Taylor expansion: EKF, SNF, HNF
    - Numerical integration: UKF, GHF, Smolyak sparse grid
  - PDE based methods: Stratonovich-Kushner and Duncan-Mortensen-Zakai (DMZ) equation
  - Exact filters: Daum, Benes
  - Particle Filters:
    Sequential Monte Carlo
- Non-Sequential
  - Simulated likelihood
  - Bayesian approaches

# SIR Particle Filter

time update: $t_i \leq t < t_{i+1}$

$$\begin{aligned} p(y,t|Z^i) &\approx \sum_n N^{-1}\delta(y - Y_n(t)) \\ dY_n(t) &= f(Y_n,t)dt + G(Y_n,t)dW_n(t) \\ Y_n(t_i) &\sim p(y,t_i|Z^i) \end{aligned}$$

measurement update: $t = t_{i+1}$, $\alpha_n := p(z_{i+1}, t_{i+1}|Y_n(t_{i+1}))$

$$\begin{aligned} u(y,t_{i+1}|z_{i+1},Z^i) &\approx \sum_n \alpha_n \delta(y - Y_n(t_{i+1})) \quad \text{unnormalized} \\ p(z_{i+1},t_{i+1}|Z^i;\psi) &\approx \int u(y)dy = \sum_n \alpha_n \quad \text{likelihood} \end{aligned}$$

- Multinomial resampling with weights $\alpha_n / \sum \alpha_n$

# Forward simulation: Particle filter

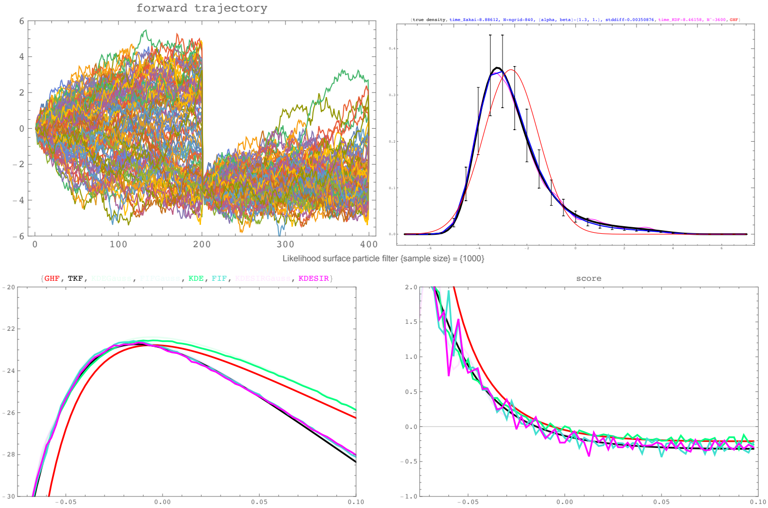


Figure: Forward simulation (top, left), estimated filter density (top, right), likelihood and score (bottom).

# Idea: compute likelihood with numerical integration

$$
\begin{aligned}
p(z_{i+1}, t_{i+1} | Z^i; \psi) &= \int p(z_{i+1}, t_{i+1} | y) p(y, t_{i+1} | Z^i) \, dy \\
&:= \int u(y, t_{i+1} | z_{i+1}, Z^i) \, dy \\
&\approx \sum_l w_l \; u_l
\end{aligned}
$$

- unnormalized filter density $u(y, t | Z^t)$
- numerical integration using quadrature formulas
- sample points $y_l := y_{i+1,l}$, $u_l := u(y_l | z_{i+1}, Z^i)$, weights $w_l$
- measurements up to time $t_{i+1}$: $Z^{i+1} = \{Z(s) \mid s \leq t_{i+1}\}$

# Compute $u$: Continuous time filtering: DMZ equation

SPDE: Zakai (1969)

$$\partial_t u(y,t|Z^t) = [F + h'\rho^{-1}(\dot{Z} - h/2)] \circ u(y,t|Z^t)$$
$$= [F(y,t) + M(y,t)] \circ u(y,t|Z^t)$$

- measurement precision $\rho^{-1}(y,t) = 0, t \neq t_i$
- Gaussian measurement density

$$p(dZ(t)|y,Z^t) = \phi(dZ, hdt, \rho dt)$$
$$\propto \exp h'\rho^{-1}(\dot{Z} - h/2)$$

- $dZ \circ u$ : symmetrized product: Stratonovich calculus

# Stochastic Representation: Feynman-Kac Formula

Lie -Trotter formula

$$
\begin{aligned}
u(y,t|Z^t) &= \lim_{\delta\tau\to 0} \prod_{j=0}^{J-1} e^{[F(y,\tau_j)+M(y,\tau_j)]\delta\tau} u(y,t_0|Z^{t_0}) \\
&= E\left[ e^{\int_{t_0}^{t} M(Y(\tau),\tau)d\tau} \delta(y - Y(t)) \,\middle|\, Z^t \right]
\end{aligned}
$$

- Zassenhaus formula $e^{(F_j+M_j)\delta\tau} \approx e^{M_j\delta\tau} e^{F_j\delta\tau}$
- integral operator (transition probability kernel)

$$
\begin{aligned}
e^{F(y,\tau_j)\delta\tau} h(y) &= \int e^{F(y,\tau_j)\delta\tau} \delta(y - y_j) h(y_j) dy_j \\
&= \int p(y,\tau_j + \delta\tau | y_j, \tau_j) h(y_j) dy_j
\end{aligned}
$$

# Importance Sampling: Backward DMZ Equation

analogous to option pricing in finance

time reversal $c(x, s) = u(x, T - s), s \leq T$

$$\begin{aligned} \partial_s c + Lc + (M + v)c &= 0 \\ \text{terminal condition } c(x, T) &= h(x) = u(x, 0) \end{aligned}$$

- rewrite $F = -\partial_\alpha f_\alpha + \frac{1}{2}\partial_\alpha \partial_\beta \Omega_{\alpha\beta} := L + v$
- backward operator $L = [-f_\alpha + (\partial_\beta \Omega_{\alpha\beta})]\partial_\alpha + \frac{1}{2}\Omega_{\alpha\beta}\partial_\alpha\partial_\beta$
- scalar potential $v = -(\partial_\alpha f_\alpha) + \frac{1}{2}(\partial_\alpha \partial_\beta \Omega_{\alpha\beta})$

# Simulation of Backward DMZ Equation

Stochastic representation

$$\begin{aligned} c(x,s) &= E\left[e^{\int_s^T (M+v)(X(\tau),\tau)d\tau} h(X(T)) \,\middle|\, X(s)=x\right] \\ dX(\tau) &= \tilde{f}(X, T-\tau)dt + G(X, T-\tau)dW(\tau) \\ X(s) &= x \end{aligned}$$

- Importance sampling: drift correction (Milstein; 1995)

  $$\tilde{f} + \Omega(X, T-\tau)\nabla \log u(X, T-\tau)$$

- use approximate filter solution $\hat{u}(X, T-\tau)$ instead of $u$:
- EKF, GHF, UKF or particle filter

# Forward and backward simulation: Particle and Zakai filter

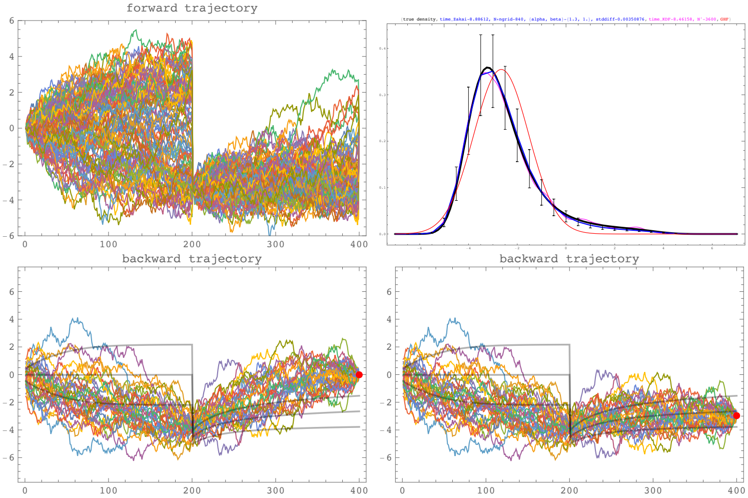


Figure: Forward simulation (top, left), estimated filter density (top, right), backward simulation with importance sampling (bottom).

# Likelihood: Particle filter (top) and Zakai filter (bottom)

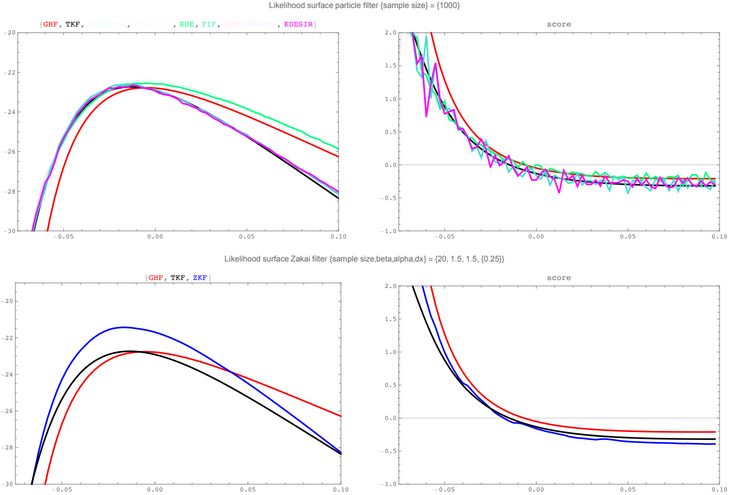


Figure: Likelihood and score for SIR particle filter, $N = 1000$ (top) and ZKF (Riemann sample points, $N = 20$, bottom), GHF, TKF as reference. Increment $d\beta = 0.0025$.

# Likelihood: Zakai filter (UT sample points)

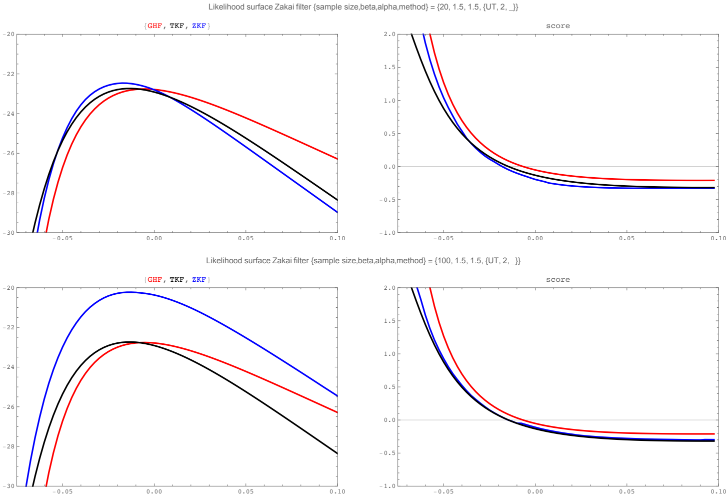


Figure: Likelihood and score for ZKF (unscented transform UT sample points), GHF, TKF. Sample size $N = 20$ (top), $N = 100$ (bottom). Increment $d\beta = 0.0025$.

# Conclusions

- Use stochastic analysis for continuous time models
- Continuous-discrete filtering
  with continuous time measurement equation
- Feynman-Kac representation
  of backward Zakai equation
- Variance reduced simulation
  of unnormalized filter density
  at supporting points
- No resampling required
- Smooth likelihood approximation
  using quadrature formulas at supporting points

# References


Blankenship, G. and Baras, J. (1981). Accurate evaluation of stochastic Wiener integrals with applications to scattering in random media and to nonlinear filtering, *SIAM Journal on Applied Mathematics* **41**(3): 518–552.

Daum, F. and Huang, J. (2016). A plethora of open problems in particle flow research for nonlinear filters, Bayesian decisions, Bayesian learning, and transport, *Signal Processing, Sensor/Information Fusion, and Target Recognition XXV*, Vol. 9842, International Society for Optics and Photonics, p. 98420I.

Ginzburg, V. and Landau, L. (1950). On the theory of superconductivity, *Collected papers of LD Landau*, Elsevier, pp. 546–568.

Herings, J. (1996). *Static and Dynamic Aspects of General Disequilibrium Theory*, Kluwer, Boston, London, Dordrecht.

Hürzeler, M. and Künsch, H. R. (2001). Approximating and maximising the likelihood for a general state-space model, *Sequential Monte Carlo methods in practice*, Springer, pp. 159–175.

Kantas, N., Doucet, A., Singh, S. S., Maciejowski, J., Chopin, N. et al. (2015). On Particle Methods for Parameter Estimation in State-Space Models, *Statistical Science* **30**(3): 328–351.

Lemos, J. M., Costa, B. A. and Rocha, C. (2018). A Fokker-Planck approach to joint state-parameter estimation, *IFAC-PapersOnLine* **51**(15): 389–394.

Malik, S. and Pitt, M. K. (2011). Particle filters for continuous likelihood evaluation and maximisation, *Journal of Econometrics* **165**(2): 190–209.

Milstein, G. N. (1995). *Numerical integration of stochastic differential equations*, Vol. 313, Springer Science & Business Media (1988 in Russian).

Mitter, S. K. (1982). Nonlinear Filtering of Diffusion Processes: A Guided Tour, *Advances in Filtering and Optimal Stochastic Control*, Springer, pp. 256–266.

Molenaar, P. and Newell, K. M. (2003). Direct fit of a theoretical model of phase transition in oscillatory finger motions, *British Journal of Mathematical and Statistical Psychology* **56**(2): 199–214.

Singer, H. (2014). Importance Sampling for Kolmogorov Backward Equations, *Advances in Statistical Analysis* **98**: 345–369.

Zakai, M. (1969). On the optimal filtering of diffusion processes, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **11**(3): 230–243.

# Operator splitting

Lie –Trotter formula

$$\lim_{n\to\infty}[e^{At/n}e^{Bt/n}]^n \quad = \quad e^{(A+B)t}$$

Zassenhaus formula

$$\begin{aligned} e^{\lambda(A+B)} &= e^{\lambda A}e^{\lambda B}e^{\lambda^2 C_2}e^{\lambda^3 C_3}\ldots. \\ C_2 &= \tfrac{1}{2}[B, A] \\ C_3 &= \tfrac{1}{3}[C_2, A+2B] \end{aligned}$$

$$e^{(A+B)} \approx \left[e^{A/n}e^{B/n}e^{C_2/n^2}e^{C_3/n^3}\ldots e^{C_m/n^m}\right]^n$$

# Stratonovich calculus

$$dZ(t)u(y,t) = dZ(t) \circ u(y,t) - \tfrac{1}{2} h(y,t)u(y,t)dt$$

DMZ equation in Itô-form

$$du(y,t|Z^t) = [F(y,t)dt + h'(y,t)\rho^{-1}(t)dZ(t)]u(y,t|Z^t)$$

symmetrized product

$$\begin{aligned} dZ(t) \circ u(y,t) &:= dZ(t)\bar{u}(y,t) \\ \bar{u}(y,t) &:= \tfrac{1}{2}[u(y,t) + u(y,t+dt)] \\ u(y,t) &= \bar{u}(y,t) - \tfrac{1}{2} du(y,t) \end{aligned}$$

# Potential $\Phi(y) = \frac{\alpha}{2}y^2 + \frac{\beta}{4}y^4$, drift $f(y) = -\nabla\Phi$

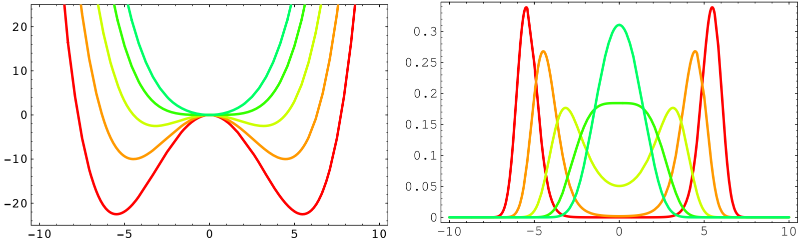


Figure: Left: Potential as a function of $y$ for parameter values $\alpha = -3, -2, ..., 1$. Right: Stationary density $p_{stat} \propto \exp[-(2/\sigma^2)\Phi(y)]$.

# Importance sampling: Kolmogorov Backward Equation

$$\begin{aligned} \partial_s c(x,s) + L(x,s)c(x,s) + v(x,s)c(x,s) &= 0 \\ \text{terminal condition } c(x,T) &= h(x) \end{aligned}$$

solution

$$c(x,s) = E\left[ e^{\int_s^T v(Y(\tau),\tau)d\tau} h(X(T)) \,\middle|\, X(s) = x \right]$$

- $dX(t) = f(X,t)dt + G(X,t)dW(t)$, $X(s) = x$
- importance sampling: drift correction $\Omega(x,s)\nabla \log c(x,s)$ (Milstein; 1995)
- backward operator $L = f_\alpha \partial_\alpha + \frac{1}{2}\Omega_{\alpha\beta}\partial_\alpha\partial_\beta$