A mathematician’s view on Asimov’s psychohistory

Dan Crisan

Imperial College London
Collaborators:

Alan Bain  
Jean-Francois Chassagneux  
Francois Delarue  
Arnaud Doucet  
Malte Grunwald  
Nick Kantas  
Christian Litterer  
Konstantinos Manolarakis  
Olasunkanmi Obanubi  
Szymon Peszat  
Nizar Touzi  
Alexandros Beskos  
Martin Clark  
Josha Diehl  
Peter Friz  
Kari Heine  
Michael Kouritzin  
Terry Lyons  
Joaquin Miguez  
Harald Oberhauser  
Boris Rozovskii  
Nick Whiteley  
Tom Cass  
Pierre Del Moral  
Peter Djuric  
Jessica Gaines  
Ajay Jasra  
Tom Kurtz  
Kai Li  
Colm Nee  
Yoshiki Otobe  
Sylvain Rubenthaler  
Jie Xiong

Phd students:

Saadia Ghazali  
Colm Nee  
Wang Han  
Kollias-Liapis Spyridon  
Marcel Ogrodnik  
Konstantinos Manolarakis  
Kai Li  
Ivo Mihaylov  
Adam Persing  
Olasunkanmi Obanubi  
Sean Violante  
Eamon McMurray  
Andrea Granelli

My family:

Alexander, Andrea and Cosette
A variety of large scale phenomena can be analyzed through their small scale counterpart. A “simple” example:

**Heat transfer**

**Macroscopic scale:** Heat transfer through mechanisms grouped in three broad categories: conduction, convection and radiation. Conduction is the transfer of heat between substances that are in direct contact. The flow of heat always occurs from a region of high temperature to another region of lower temperature.

**Heat flow in paper**

**Microscopic scale:** Thermal energy is related to the kinetic energy of molecules. The greater a material’s temperature, the greater the thermal agitation of its constituent molecules (manifested both in linear motion and vibrational modes). Kinetic energy is transferred from regions with greater molecular kinetic energy towards regions with less kinetic energy.

**Simulation of molecular kinetic energy evolution**
Feynman-Kac formulae

The Heat Equations

Microscopic level

Brownian motion

\[ W = \{ W_t, t \geq 0 \} \]

Macroscopic level

Heat equation

\[
\begin{align*}
\partial_t u_t &= \frac{1}{2} \Delta u_t \\
u_0 &= \Phi
\end{align*}
\]

Feynman-Kac Formula:

\[ u_t(x) = E[\Phi(x + W_t)] \]
Feynman-Kac Formula:

\[ u_t(x) = E[\Phi(x + W_t)]. \]

<table>
<thead>
<tr>
<th>Microscopic level</th>
<th>Macroscopic level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numerical approximation:</td>
<td>Exact solution:</td>
</tr>
<tr>
<td>[ u_t(x) \simeq \sum_{i=1}^{n} \Phi(x + W_t^i). ]</td>
<td>[ u_t(x) = \int \Phi(y) \frac{1}{\sqrt{2\pi t}} e^{-|y-x|^2/2t} , dy. ]</td>
</tr>
<tr>
<td>Malliavin differentiation (d=1):</td>
<td>Standard differentiation (d=1):</td>
</tr>
<tr>
<td>[ \frac{du_t(x)}{dx} = \frac{1}{t} E[\phi(x + W_t) W_t]. ]</td>
<td>[ \frac{du_t(x)}{dx} = \frac{1}{t} \int \Phi(y) \frac{x-y}{\sqrt{2\pi t}} e^{-|y-x|^2/2t} , dy. ]</td>
</tr>
</tbody>
</table>
Can do this for a more general equation:

\[
\begin{cases}
\partial_t u = V_0 u + \frac{1}{2} \sum_{i=1}^{d} V_i^2 u + f(t, x, u, V_1 u, ..., V_d u), & t \in (0, T], x \in \mathbb{R}^m \\
u(0, x) = \Phi(x), & x \in \mathbb{R}^m
\end{cases}
\]

1. $V_i u, i = 0, 1..., d$ is the directional derivative of $u$ in the direction $V_i$ that satisfy the UFG condition.

2. $\mathcal{W}$ the $C_b^\infty(\mathbb{R}^d)$-module generated by the vector fields $\{V_i, i = 1, ..., N\}$ within the Lie algebra generated by $\{V_i, i = 1, ..., N\}$ is finitely generated as a vector space and $\{V_\alpha, \alpha \in \mathcal{A}_0(m)\}$ is a finite set of generators.
Feynman-Kac Formula:

\[ u(t, x) = E[\Lambda_{t,x}(W)] = \int_{\omega \in C([0, \infty), \mathbb{R}^d)} \Lambda_{t,x}(\omega) dP_W(\omega). \]

<table>
<thead>
<tr>
<th>Microscopic level</th>
<th>Macroscopic level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numerical approximation:</td>
<td>No exact solution</td>
</tr>
<tr>
<td>• replace ( P_W ) with ( P_{\tilde{W}} = \frac{1}{n} \sum_{i=1}^{n} \delta_{\omega_i} )</td>
<td>No standard differentiation</td>
</tr>
<tr>
<td>• approximate ( \Lambda_{t,x} ) with a simple version ( \tilde{\Lambda}_{t,x} )</td>
<td></td>
</tr>
<tr>
<td>• control the computational effort</td>
<td></td>
</tr>
</tbody>
</table>

Malliavin differentiation

\[ V_i u_t(x) = \frac{1}{t} E[\Lambda_{t,x}^i(W)] \]

\[ \mathcal{V}_0 u = \frac{1}{2} \sum_{i=1}^{d} V_i^2 u + f(t, x, u, V_1 u, \ldots, V_d u) \]

\( \mathcal{V}_0 := \partial_t - \mathcal{V}_0 \)

• \( u \) not necessarily time differentiable
• temporal variable cannot be distinguished from the space one
• \( u \) remains differentiable in the direction \( \mathcal{V}_0, V_1, \ldots, V_d \).
Let \( u \in C^{1,2}([0, T] \times \mathbb{R}^4) \) be the solution of the Cauchy problem

\[
\begin{aligned}
\partial_t u &= \sum_{i=1}^{4} \left( \mu_i x_i \frac{du}{dx_i} + \frac{\sigma_i^2 x_i^2}{2} \frac{d^2 u}{dx_i^2} \right) + f(t, x, u, \Delta u), \quad t \in (0, T], \ x \in \mathbb{R}^4, \\
u(0, x) &= \Phi(x), \quad x \in \mathbb{R}^4
\end{aligned}
\]

where

\begin{itemize}
\item \( f(t, x, y, z) = -ry - \sum_{i=1}^{4} (\mu_i - r)z_i, \ x, z \in \mathbb{R}^4. \)
\item \( \Phi(x) = (x_1 x_2 x_3 x_4 - k)_+. \)
\end{itemize}
Common feature of many PDEs: their solutions can be represented as integrals of certain nonlinear functionals with respect to the Wiener measure.

**Feynman-Kac formula**

\[ u(t, x) = E[\Lambda_t, x(W)] = \int_{\omega \in C([0, \infty), \mathbb{R}^d)} \Lambda_t, x(\omega) dP_W(\omega) \]

<table>
<thead>
<tr>
<th>Microscopic level</th>
<th>Macroscopic level</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Brownian motion</strong></td>
<td>Heat equation</td>
<td>Feynman 1948 Kac 1949</td>
</tr>
</tbody>
</table>
| \( W = \{ W_t, t \geq 0 \} \) | \[ \begin{aligned} 
\partial_t u_t &= \frac{1}{2} \Delta u_t \\
\quad u_0 &= \Phi 
\end{aligned} \] | Duncan, Mortensen, Zakai 1970 |
| Zakai equation | \[ \begin{aligned} 
du_t &= Lu_t + hu_t dY_t \\
\partial_t u_t &= \sum_{i,j=1}^{d} a_{ij}(u_t) \partial_i \partial_j u_t \\
&\quad + \sum_{i=1}^{d} b_i(u_t) \partial_i u_t + c(u_t) u_t 
\end{aligned} \] | Gärtner 1988 |
| McKean-Vlasov PDEs | Semilinear PDEs | Pardoux & Peng 1990, 1992 |
| | \[ \begin{aligned} 
\partial_t u_t &= Lu_t + f(t, x, u_t, \nabla u_t) \\
\quad u_0 &= \Phi 
\end{aligned} \] | Soner, Touzi & Victoir 2007 |
| | Fully Nonlinear PDEs | |
| | \[ F(t, x, u_t, \nabla u_t, \Delta u_t) = 0 \] | |
| | 3 — d incompressible | Constantin & Iyer 2008 |
| | Navier — Stokes equation | \[ \begin{aligned} 
\partial_t u_t + (u_t \cdot \nabla) u_t - \nu \Delta u_t + \nabla p &= 0 \\
\nabla \cdot u_t &= 0 
\end{aligned} \] | |
| | K-S equation | Crisan & Xiong 2009 |
| | \[ du_t = Lu_t dt + u_t(\bar{h})dY_t - u_t(\bar{h})u_t(h) dt \] | Novikov & Iyer 2010 |
| viscous Burgers equation | \[ \partial_t u_t + u_t \partial_x u_t - \nu \partial_x^2 u_t = 0 \] | |
The Stochastic filtering Problem

Stochastic Filtering

\[(X, Y) = \{(X_t, Y_t), t \geq 0\}\]

- \(X\) the signal process - “hidden component”
- \(Y\) the observation process - “the data” - \(Y_t = f_t(X, “noise”).\)

The filtering problem: Find the conditional distribution of the signal \(X_t\) given \(Y_t = \sigma(Y_s, s \in [0, t]), i.e.,\)

\[\vartheta_t(\varphi) = \mathbb{E}[\varphi(X_t)|Y_t], \quad t \geq 0, \quad \varphi \in \mathcal{B}(\mathbb{R}^d).\]

The model

\[
\begin{align*}
X_t &= X_0 + \int_0^t b(X_s) \, ds + \int_0^t \sigma(X_s) \, dV_s \\
Y_t &= \int_0^t \gamma(X_s) \, ds + U_t,
\end{align*}
\]
The Stochastic filtering Problem

The SPDE

The Kallianpur-Striebel formula

\[ \vartheta_t(\phi) = \frac{\rho_t(\phi)}{\rho_t(1)}, \]

\( \rho = \{\rho_t, t \geq 0\} \) satisfies the Duncan-Mortensen-Zakai equation:

\[ d\rho_t(x) = A^* \rho_t(x) dt + \rho_t(x) \left( \sum_{k=1}^{m} \gamma_k(x) dY_t^k \right) \quad (2) \]

where

\[ A\phi(x) = \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \partial_i \partial_j \phi(x) + \sum_{i=1}^{d} b_i(x) \partial_i \phi(x) \quad a = \sigma \sigma^\top \]

The Feynman-Kac formula

\[ \rho_t(\phi) = \mathbb{E} \left[ \phi(X_t) \exp \left( \int_0^t \sum_{k=1}^{m} \gamma_k(X_s) dY_s^k - \frac{1}{2} \int_0^t \sum_{k=1}^{m} \gamma_k(X_s)^2 ds \right) \mid Y_t \right] \quad (3) \]
Car tracking: Extracting information from video images

Signal \((\dot{v}_t, v_t, \theta_t, x_t, y_t, \tau_t)\):

- \((x_t, y_t)\) be the coordinates of the car (the center of the rear axle),
- \(v_t\) be its tangential velocity (\(\dot{v}_t\) is the rate of change of \(v_t\)),
- \(\tau_t\) its orientation (the angle between the direction of the car and the \(x_t\)-axis)
- \(\theta_t\) the steering angle (the angle between the direction of the car and the direction of the front wheels).

\[
\begin{align*}
    d\dot{v}_t &= \alpha dW^1_t \\
    dv_t &= v_t dt \\
    d\theta_t &= -\beta \theta_t dt + \gamma dW^2_t \\
    dx_t &= v_t \cos(\tau_t) dt \\
    dy_t &= v_t \sin(\tau_t) dt \\
    d\tau_t &= a^{-1} v_t \theta_t dt
\end{align*}
\]  

(4)

Observation \((\bar{x}_t, \bar{y}_t, \bar{\tau}_t)\) rough estimations of \((x_t, y_t, \tau_t)\) obtained by a visual searching algorithm.

\[
\begin{align*}
    d\bar{x}_t &= v_t \cos(\tau_t) dt + \delta_3 dW^3_t \\
    d\bar{y}_t &= v_t \sin(\tau_t) dt + \delta_4 dW^4_t \\
    d\bar{\tau}_t &= a^{-1} v_t \theta_t dt + \delta_5 dW^5_t
\end{align*}
\]  

(5)
2D Stochastic Navier-Stokes equation

Torus $\mathbb{T}^2 \triangleq [0, L) \times [0, L)$ with periodic boundary conditions:

$$\frac{\partial u}{\partial t} - \nu \Delta u + u \cdot \nabla u + \nabla p = f + W(t, x)$$

for all $(x, t) \in \mathbb{T}^2 \times (0, \infty)$, \hspace{1cm} (6)

$$\nabla \cdot u = 0$$

for all $(x, t) \in \mathbb{T}^2 \times (0, \infty)$, \hspace{1cm} (6)

$$u(x, 0) = u_0(x)$$

for all $x \in \mathbb{T}^2$. \hspace{1cm} (6)
- $u : \mathbb{T}^2 \times [0, \infty) \rightarrow \mathbb{R}^2$ - the velocity
- $p : \mathbb{T}^2 \times [0, \infty) \rightarrow \mathbb{R}^2$ - the pressure
- $f : \mathbb{T}^2 \rightarrow \mathbb{R}^2$ - the forcing
- $W(t, x)$ - noise

\[
H \triangleq \left\{ L - \text{periodic trig. pol. } u : [0, L)^2 \rightarrow \mathbb{R}^2 \mid \nabla \cdot u = 0, \int_{\mathbb{T}^2} u(x) \, dx = 0 \right\}^{L^2(\mathbb{T}^2))^2}
\]

$P : (L^2(\mathbb{T}^2))^2 \rightarrow H$ - the Leray-Helmholtz orthogonal projector. An orthonormal basis for $H$ is given by

\[
\psi_k(x) \triangleq \frac{k_{\perp}}{|k|} \exp \left( \frac{2\pi i k \cdot x}{L} \right) \quad k = (k_1, k_2)^T \in \mathbb{Z}^2 \setminus \{0\} \quad k_{\perp} = (k_2, -k_1)^T.
\]
For \( u \in H \cdot \)

\[
  u = \sum_{k \in \mathbb{Z}^2 \setminus \{0\}} u_k(t) \psi_k(x).
\]

\[
  W(t, x) = \sum_{k \in \mathbb{Z}^2 \setminus \{0\}} \varepsilon_k \psi_t(x) W_t^k \in H.
\]

\[
  \{ W_t^k \}_{(t \geq 0, \ k \in \mathbb{Z}^2 \setminus \{0\})} \ 	ext{i.i.d. Brownian motions and}
\]

\[
  \sum_{k \in \mathbb{Z}^2 \setminus \{0\}} (4\pi^2 |k|^2)^s \varepsilon_k^2 < \infty \quad \text{for } s \in \mathbb{R},
\]

and then \( W(t, \cdot) \in H. \)

The stochastic Navier-Stokes equation can be written as

\[
  \frac{du}{dt} + \nu A u + B(u, u) = f + W(t, x). \tag{7}
\]

• \( A = -P\Delta \) is the Stokes operator
• \( B(u, u) = P(u \cdot \nabla u) \)
• \( f \) is the original forcing projected into \( H. \)
The equations for the modes:

\[ du_k(t) = \left( -\nu \lambda_k u_k(t) - \alpha_{k}^{l,j} \sum_{l+j=k} u_l(t)u_j(t) + f_k \right) dt + \varepsilon_k dW_t^k. \]

\[ \alpha_{k}^{l,j} = \begin{cases} \frac{2\pi i(l_2j_1-l_1j_2)(k_1j_1+k_2j_2)}{L|k||l||j|} & \text{if } k = l + j, \\ 0 & \text{otherwise}; \end{cases} \]

Define the projection operators \( P_{\lambda} : H \rightarrow H \) and \( Q_{\lambda} : H \rightarrow H \) by

\[ P_{\lambda} u = \sum_{k \in \mathbb{Z}^2 \setminus \{0\}} u_k(t)\psi_k(x), \quad Q_{\lambda} = I - P_{\lambda}; \]

and consider the projected eigenvalues, we obtain the following evolution equation for the approximation of \( u_k(t) \), which is denoted by \( \tilde{u}_k(t) \), for each \( k \in \mathbb{Z} \setminus \{0\} \) with \(|2\pi k|^2 < \lambda L^2\):

\[ d\tilde{u}_k(t) = \left( -\nu \lambda_k \tilde{u}_k(t) - \alpha_{k}^{l,j} \sum_{\Gamma} \tilde{u}_l(t)\tilde{u}_j(t) + f_k \right) dt + \varepsilon_k dW_t^k; \]  

(8)

where the set \( \Gamma \triangleq \{(l, j) \mid l + j = k \text{ and } |2\pi l|^2 < \lambda L^2 \text{ and } |2\pi j|^2 < \lambda L^2 \} \).
Model parameters

- we use $k_1, k_2 = -32, \ldots, 0, \ldots 32$ (i.e. a $64^2$ grid for the discrete fourier components).
- Smoothing problem approximate $p(x_0|y_{1:5})$ where each $y_i$ is a 4x4 grid on the torus and
  \[ y_i(j) = u(x_j, t_i) + N(0, 0.2). \]
- the dynamics are initialised by a random sample from the prior $N(0, \delta A^\alpha)$
- for the prior, $\delta = 5$ and $\alpha = 2.2$.
- torus size is $2\pi$.
- forcing is $\nabla \cos(\kappa \cdot x)$ with $\kappa = (1, 1)$ for the stationary regime and $\kappa = (5, 5)$ for the chaotic
- $\nu$ is $1/50$ for chaotic and $1/10$ for stationary

MCMC plot: computation cost roughly $10^5$ iterations per day, i.e. with the slow mixing need more than 10 days for a decent but not super-reliable answer.

SMC plots: computational cost around 18 hours for $N=500$ particles and 5 intermediate steps.
Asimov’s psychohistory

“Psychohistory: Branch of mathematics which deals with the reactions of human conglomerates to fixed social and economic stimuli. With refined knowledge of group psychology, the future can be extrapolated from the present. Psychohistory cannot work with individuals over any lengths of time any more than you can apply the kinetic theory of gases to single molecules.”

Nobel laureate Paul Robin Krugman said that his interest in economics began with Asimov’s Foundation novels, in which the social scientists of the future use ‘Psychohistory’ to attempt to save civilization. Since ‘Psychohistory’ in Asimov’s sense of the word does not exist, Krugman turned to economics, which he considered the next best thing.
It would be a wonderful achievement to be able to study the present and then accurately predict what will happen in the future.

Mathematics has come of age:

- good understanding of how to model noise/system uncertainty: stochastic processes (Brownian motion, Levy process), rough controls
- models for “noisy” dynamical systems: SDEs, SPDEs, RDEs
- rigorous analysis beyond classical results: stochastic calculus, Malliavin calculus, stochastic geometry, theory of rough paths
- averaging over individual evolutions to give the overall picture: strong law of large numbers, central limit theorems, large deviations
- can incorporate observed behaviour: stochastic filtering
- can model interactive behaviour individual-individual, individual-society: stochastic networks, McKean-Vlasov processes
- can handle multi-scale (slow/fast) phenomena: homogenization theory
- ...
A mathematics model for the evolution of the human society:

- how do we model individual/family/nation behaviour evolution
- what are the human characteristics that influence society
- how do we distinguish between common traits and unique qualities
- how do we identify social and political patterns
- ..... 

Not a mathematics-only problem: Understanding the complexities of human society requires major input from many other disciplines: sociology, history, economics, politics, etc.

The ingredients are available, all that we need now is the recipe.