

C8.2 Stochastic Analysis and PDEs project

Branching processes and semilinear parabolic PDEs

Matthieu Meunier

April 16, 2024

Representing solutions of PDEs with probabilistic objects, such as diffusion processes, presents many advantages both from the theoretical and practical point of view. Among others, such a probabilistic representation allows to use stochastic algorithms (e.g. Monte Carlo methods) to numerically approximate solutions of high-dimensional PDEs without necessarily incurring an exponential cost in the dimension of the problem (the so-called curse of dimensionality).

This report proposes to take a deeper dive into the relationship between semilinear parabolic PDEs and branching processes, thus extending section 7.4 of the lecture notes. Our study will mostly rely on work by Labordere et al. [1]. In a nutshell, they are able to show an extended version of the classical representation for KPP equations (e.g. Theorem 7.32 of the lecture notes) to the case of gradient non-linearities, i.e. equations of the form

$$-\partial_t u - \mathcal{L}u = f(u, \nabla_x u), \quad u(T, x) = g(x), \quad x \in \mathbb{R}^d, t \in [0, T]$$

1 An introductory example

We present the main idea of the article in the context of a simple 1-dimensional equation

$$-\partial_t u - \frac{1}{2} \partial_{xx} u = \frac{1}{2} (u^2 + u \partial_x u), \quad u(T, x) = g(x), \quad x \in \mathbb{R}, t \in [0, T] \quad (1)$$

The Feynman-Kac formula allows to write the solution at initial time as follows

$$u(0, x) = \mathbb{E} \left[g(W_T) + \int_0^T f(u, \partial_x u)(t, W_t) dt \mid W_0 = x \right]$$

where (W_t) is a Brownian motion (not necessarily started at 0). Now, consider a non-negative random variable τ^1 with density $\rho > 0$ on \mathbb{R}_+ , as well as a Brownian motion W^1 independent of τ^1 . Let $\bar{F}(t) = \int_t^\infty \rho(s) ds$. Then one can rewrite the Feynman-Kac formula in terms of the distribution of τ^1

$$u(0, x) = \mathbb{E}_{0,x} \left[\bar{F}(T) \frac{g(W_T^1)}{\bar{F}(T)} + \int_0^T \frac{f(u, \partial_x u)(t, W_t^1)}{\rho(t)} \rho(t) dt \right] = \mathbb{E}_{0,x} [\phi(0, \tau^1 \wedge T, W_{\tau^1 \wedge T}^1)]$$

where $\phi(s, t, y) := \frac{1_{t \geq T}}{\bar{F}(t-s)} g(y) + \frac{1_{t < T}}{\rho(t-s)} (u \partial_x^{I_1} u)(t, y)$ and I_1 is a Bernoulli random variable independent of τ^1 and W^1 . Here $\partial_x^{I_1} u = \partial_x u$ and $\partial_x^0 u = u$.

By the Markov property, when $I_1 = 0$, we get

$$\begin{aligned} u \partial_x^{I_1} u(t, y) &= u(t, y)^2 = \mathbb{E}_{t,y} \left[\phi(t, (t + \tau^1) \wedge T, W_{(t+\tau^1) \wedge T}^1) \right]^2 \\ &= \mathbb{E}_{t,y} \left[\phi(t, \tau_t^{1,1}, W_{\tau_t^{1,1}}^{1,1}) \phi(t, \tau_t^{1,2}, W_{\tau_t^{1,2}}^{1,2}) \right] \end{aligned}$$

where $W^{1,i}, \tau_t^{1,i} := (t + \tau^{1,i}) \wedge T$ are all independent, $\tau^{1,i}$ are distributed according to ρ and $W^{1,i}$ are Brownian motions for $i = 1, 2$. Here, $\mathbb{E}_{t,y}$ means expectation conditioned on $W_t^{1,1} = W_t^{1,2} = y$.

When $I_1 = 1$, we get

$$u \partial_x^{I_1} u(t, y) = \mathbb{E}_{t,y} \left[\phi(t, \tau_t^{1,1}, W_{\tau_t^{1,1}}^{1,1}) \right] \partial_y \mathbb{E}_{t,y} \left[\phi(t, \tau_t^{1,2}, W_{\tau_t^{1,2}}^{1,2}) \right]$$

Expanding the second conditional expectation and differentiating inside the heat kernel gives

$$\partial_y \mathbb{E}_{t,y} \left[\phi(t, \tau_t^{1,2}, W_{\tau_t^{1,2}}^{1,2}) \right] = \mathbb{E}_{t,y} \left[\frac{W_{\tau_t^{1,2}}^{1,2} - W_t^{1,2}}{\tau_t^{1,2} - t} \phi(t, \tau_t^{1,2}, W_{\tau_t^{1,2}}^{1,2}) \right]$$

Introducing $\mathcal{W}^1 := \mathbf{1}_{I_1=0} + \mathbf{1}_{I_1=1} \frac{\Delta W_{T_{(1,2)}}^{1,2}}{\Delta T_{(1,2)}}$, where $\Delta W_{T_{(1,2)}}^{1,2} := W_{T_{(1,2)}}^{1,2} - W_{T_{(1)}}^{1,2}$, $\Delta T_{(1,2)} = T_{(1,2)} - T_{(1)}$ and $T_{(1,i)} = (\tau^1 + \tau^{1,i}) \wedge T$, yields the following representation of the solution of (1)

$$u(0, x) = \mathbb{E}_{0,x} \left[\mathbf{1}_{T_{(1)}=T} \frac{g(W_T)}{F(T)} + \mathbf{1}_{T_{(1)}<T} \mathcal{W}^1 \prod_{i=1}^2 \left(\mathbf{1}_{T_{(1,i)}=T} \frac{g(W_T^{1,i})}{F(\Delta T_{(1,i)})} + \mathbf{1}_{T_{(1,i)}<T} \frac{u \partial_x^{I^{1,i}} u(T_{(1,i)}, W_{T_{(1,i)}}^{1,i})}{\rho(\Delta T_{(1,i)})} \right) \right] \quad (2)$$

One might wonder why this representation is interesting. Essentially, we can repeat the trick of expressing the non-linearity $u \partial_x^I u$ using random times $T_{(1,i,j)} = (\tau^1 + \tau^{1,i} + \tau^{1,i,j}) \wedge T$, and so on. As such, the term carrying the non-linearity becomes negligible on average, since the events $\{T_{(1,i_2,\dots,i_k)} < T\}$ become less likely. If \mathcal{W} can be computed easily and has nice integrability properties, then we can reasonably hope that this representation is helpful.

We will now make this argument more general, considering more general polynomial nonlinearities and diffusion processes. The setting of our study is presented in Section 2. To formalize this splitting trick, we turn to age-dependent marked branching processes in Section 3, and then state the representation result in Section 4. Some numerical experiments are carried out in Section 5.

2 Semilinear parabolic PDEs with polynomial gradient non-linearity

Let $d \geq 1$, $\mu : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$. For $u : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$, we denote by ∇u , $\nabla^2 u$ the gradient and Hessian respectively of u with respect to the space variable x . Let $m \geq 0$, and consider a subset $L \subset \mathbb{N}^{m+1}$ along with a sequence of functions $(c_l)_{l \in L}$ and $(b_i)_{i \in \{1, \dots, m\}}$ where $c_l : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ and $b_i : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$. For $l \in L$ let $|l| = \sum_{i=1}^m l_i$. We define the generator function $f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$ by

$$f(t, x, y, z) := \sum_{l \in L} c_l(t, x) y^{l_0} \prod_{i=1}^m (b_i(t, x) \cdot z)^{l_i} \quad (3)$$

and we consider, for such a generator, the following PDE

$$\partial_t u + \mu \cdot \nabla u + \frac{1}{2} \text{Trace}(\sigma \sigma^T \nabla^2 u) + f(\cdot, u, \nabla u), \quad (t, x) \in [0, T] \times \mathbb{R}^d, \quad u(T, \cdot) = g \quad (4)$$

for some terminal condition $g : \mathbb{R}^d \rightarrow \mathbb{R}$ assumed to be Lipschitz and bounded.

3 Age-dependent marked branching processes.

3.1 Background and notations

Consider a fixed finite time horizon $0 < T < \infty$. We consider a branching time characterized by a splitting time density ρ and a probability mass function $(p_l)_{l \in L}$ where $L \subset \mathbb{N}^{m+1}$. For $l \in L$, let $|l| = \sum_i l_i$. $m \in \mathbb{N}_0$ characterizes the different marks assigned to particles. At the splitting time, the particle branches into $|l|$ offsprings with probability p_l , among which the first l_0 carry the mark 0, the next l_1 carry the mark 1 and so forth. Then, regardless of its mark, each particle performs the same independent branching process. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space carrying a sequence of iid positive random times $(\tau^{n,q})_{n,q \geq 1}$ of density $\rho : \mathbb{R} \rightarrow \mathbb{R}_+$ supported on \mathbb{R}_+ , and a sequence of iid L -valued random variables $(I^{n,q})_{n,q \geq 1}$ such that $\mathbb{P}(I^{n,q} = l) = p_l$ for all $l \in L$. We assume that the families $(\tau^{n,q})$ and $(I^{n,q})$ are independent.

We now construct an age-dependent branching process, where each particle is assigned a unique tuple label $k \in \bigcup_n \mathbb{N}^n$. Consider for each n an injection $\pi_n : \mathbb{N}^n \rightarrow \mathbb{N}$.

- The first particle is marked by 0, labelled by (1) , and is said to be of generation 1. Its splitting time is given by $T_{(1)} := \tau^{1,1} \wedge T$;
- Given a particle of generation n labelled by $k = (k_1, k_2, \dots, k_n) \in \mathbb{N}^n$, and with splitting time T_k , we let $I_k = I^{n, \pi_n(k)}$. If $T_k < T$, the particle is split into $|I_k|$ offsprings, and each are assigned a label (k_1, \dots, k_n, i) for $i = 1, \dots, |I_k|$, otherwise it simply dies at T .
- Given $I_k = l = (l_0, \dots, l_m)$, the first l_0 are assigned the mark 0, the next l_1 are assigned the mark 1, up until the last l_m which are assigned the mark m .
- For a particle $k = (k_1, \dots, k_{n+1})$ of generation $(n+1)$, we denote by $k- = (k_1, \dots, k_n)$ its parent, hence the splitting time of k is given by $T_k = (T_{k-} + \tau^{n+1, \pi_{n+1}(k)}) \wedge T$.
- In particular, for the initial particle, $k- = \emptyset$ and $T_{k-} = 0$.

Finally, we introduce $\theta_k \{0, 1, \dots, m\}$ to be the mark of k and

$$\mathcal{K}_t^n := \begin{cases} \{k \text{ of generation } n : T_{k-} \leq t < T_k\}, & \text{when } t \in [0, T), \\ \{k \text{ of generation } n : T_k = T\} & \text{when } t = T, \end{cases}$$

as well as $\bar{\mathcal{K}}_t^n := \bigcup_{s \leq t} \mathcal{K}_s^n$, $\mathcal{K}_t = \bigcup_n \mathcal{K}_t^n$, and $\bar{\mathcal{K}}_t := \bigcup_n \bar{\mathcal{K}}_t^n$. Clearly, \mathcal{K}_t (resp. \mathcal{K}_t^n) denotes the set of all living particles (resp. of generation n) in the system at time t , and $\bar{\mathcal{K}}_t$ (resp. $\bar{\mathcal{K}}_t^n$) denotes the set of all particles (resp. of generation n) which are alive at time t or have been alive before t .

It can be shown that under the assumption that $\sum_l |l| p_l < \infty$ and that the splitting time distribution does not have any mass at 0 (which is implied by our definition with a density function), then $\bar{\mathcal{K}}_t$ is finite almost-surely for all t .

3.2 The marked branching diffusion

Consider a family of d -dimensional Brownian motions $(W^{n,q})$ which are independent of $(\tau^{n,q}, I^{n,q})$. Define $W_t^{(1)} = \Delta W_t^{(1)} := W_t^{1,1}$ for $t \in [0, T^{(1)}]$, and for $k = (k_1, \dots, k_n) \in \bar{\mathcal{K}}_T \setminus \{(1)\}$, define

$$W_t^k := W_{T_{k-}}^{k-} + \Delta W_{t-T_{k-}}^k, \quad \text{where } \Delta W_{t-T_{k-}}^k := W_{t-T_{k-}}^{n, \pi_n(k)}, \quad \text{for all } t \in [T_{k-}, T_k]$$

For each $k \in \bar{\mathcal{K}}_T$ we define an associated diffusion process

$$X_t^k = X_{T_{k-}}^{k-} + \int_{T_{k-}}^t \mu(s, X_s^k) ds + \int_{T_{k-}}^t \sigma(s, X_s^k) dW_s^k, \quad t \in [T_{k-}, T_k], \mathbb{P}\text{-a.s.} \quad (5)$$

where the initial condition for (1) is $X_0^{(1)} = x_0 \in \mathbb{R}^d$. Finally, we introduce the σ -fields $\mathcal{F}_0 := \sigma\{\tau^{n,q}, I^{n,q} : n, q \geq 1\}$ and $\mathcal{F}_n := \sigma\{W^{i,q}, \tau^{i,q}, I^{i,q} : q \geq 1, i \leq n\}$ for all $n \geq 1$.

4 The representation result

We work under the following assumptions:

- (p_l) satisfies $p_l > 0$ for all $l \in L$ and $\sum_l |l| p_l < \infty$. ρ is continuous and strictly positive on $[0, T]$, and such that $\bar{F}(T) = \int_T^\infty \rho(t) dt > 0$.
- The coefficients of the diffusion μ, σ are bounded continuous and Lipschitz in space.
- c_l, b_i are bounded continuous for all $l \in L, i \in \{0, \dots, m\}$. In particular, f defined by (3) is continuous.
- There is a measurable functional $\bar{W}(t, s, x, (W_r - W_t)_{r \in [t, s]})$ such that $(t, x) \mapsto \bar{W}(t, s, x, (W_r - W_t)_{r \in [t, s]})$ is continuous, and for any $s \in [t, T]$ and bounded measurable function $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$, one has

$$\partial_x \mathbb{E} \left[\phi(\bar{X}_s^{t,x}) \right] = \mathbb{E} \left[\phi(\bar{X}_s^{t,x}) \bar{W}(t, s, x, (W_r - W_t)_{r \in [t, s]}) \right] \quad (6)$$

The functional \bar{W} is referred to as a Malliavin automatic differentiation function, and is obtained in practice through Malliavin calculus techniques. Note that in the case of μ, σ being constants, we can use

$$\bar{W} \left(t, s, x, (W_r - W_t)_{r \in [t, s]} \right) := \left(\sigma_0^\top \right)^{-1} \frac{W_s - W_t}{s - t}.$$

By analogy with the example of Section 1, we introduce

$$\mathcal{W}_k := \mathbf{1}_{\theta_k=0} + \mathbf{1}_{\theta_k \neq 0} b_{\theta_k}(T_{k-}, X_{T_{k-}}^k) \cdot \bar{W}(T_{k-}, T_k, X_{T_{k-}}^k, \Delta W^k).$$

for all particles $k \in \bar{\mathcal{K}}_T$. For $u \in C^{1,2}([0, T] \times \mathbb{R}^d)$ we consider the following n -th generation decomposition

$$\begin{aligned} \psi_n := & \left[\prod_{k \in \bigcup_{j=1}^n \mathcal{K}_T^j} \frac{g(X_T^k) - g(X_{T_{k-}}^k) \mathbf{1}_{\theta_k \neq 0}}{\bar{F}(\Delta T_k)} \mathcal{W}_k \right] \left[\prod_{k \in \bigcup_{j=1}^n (\bar{\mathcal{K}}_T^j \setminus \mathcal{K}_T^j)} \frac{c_{I_k}(T_k, X_{T_k}^k)}{p_{I_k}} \frac{\mathcal{W}_k}{\rho(\Delta T_k)} \right] \\ & \times \left[\prod_{k \in \bar{\mathcal{K}}_T^{n+1}} \left(\mathbf{1}_{\theta_k=0} u + \sum_{i=1}^m \mathbf{1}_{\theta_k=i} b_i \cdot \nabla u \right) (T_{k-}, X_{T_{k-}}^k) \right], \end{aligned} \quad (7)$$

and the limit

$$\psi := \left[\prod_{k \in \mathcal{K}_T} \frac{g(X_T^k) - g(X_{T_{k-}}^k) \mathbf{1}_{\theta_k \neq 0}}{\bar{F}(\Delta T_k)} \mathcal{W}_k \right] \left[\prod_{k \in \bar{\mathcal{K}}_T \setminus \mathcal{K}_T} \frac{c_{I_k}(T_k, X_{T_k}^k)}{p_{I_k}} \frac{\mathcal{W}_k}{\rho(\Delta T_k)} \right] \quad (8)$$

We similarly define $\psi_n^{t,x}, \psi^{t,x}$ when the branching diffusion is started at $X_t^{(1)} = x$ for $(t, x) \in [0, T] \times \mathbb{R}^d$. We now state a version of the representation result (Proposition 3.4 in [1]).

Theorem 1. *Suppose the above assumptions hold true, and that the PDE (4) has a classical solution $u \in C^{1,2}$ with $\mathbb{E} \left[\int_t^T |f(s, \bar{X}_s^{t,x}, u(s, \bar{X}_s^{t,x}), \nabla u(s, \bar{X}_s^{t,x}))| ds \right] < \infty$ for some initial condition $(t, x) \in [0, T] \times \mathbb{R}^d$. Assume further that the sequence $(\psi_n^{t,x})_n$ is uniformly integrable. Then $\psi^{t,x}$ is integrable and $u(t, x) = \mathbb{E}[\psi^{t,x}]$.*

The proof is given by an adaptation of the reasoning carried out in Section 1. A more explicit version is given by Theorem 3.5 in [1], where the conditions for the validity of the representation are more transparent, but it does not provide more insights.

5 Numerical experiments

We are now interested in applying this result to a concrete example. We will compare the choice of the authors for the splitting distribution against other ones. The authors propose to choose a gamma distribution $\Gamma(\kappa, \theta)$ with $\kappa \leq 1/2$, the density is given by

$$\rho_0(t) = \frac{1}{\Gamma(k)\theta^k} t^{k-1} e^{-t/\theta} \mathbf{1}_{t>0}$$

with $\Gamma(k) = \int_0^\infty s^{k-1} e^{-s} ds$. In particular, we have

$$\bar{F}(\Delta T_k) = 1 - \frac{\gamma(\kappa, \Delta T_k/\theta)}{\Gamma(k)}, \quad \gamma(\kappa, t) := \int_0^t s^{\kappa-1} e^{-s} ds$$

As a first step, we implement a naive solver based on Monte-Carlo sampling of (8) and use it on the toy problem presented in Section 5 of [1]. We consider the case of $\mu = 0, \sigma = d^{-1/2} I_d$ and $f(t, x, y, z) = k(t, x) + cy(b \cdot z)$ where

$$b = \frac{1}{d} \left(1 + \frac{1}{d}, 1 + \frac{2}{d}, \dots, 2 \right)$$

$$k(t, x) = \cos(x_1 + \dots + x_d) \left(\alpha + \frac{\sigma^2}{2} + c \sin(x_1 + \dots + x_d) \frac{3d+1}{2d} e^{\alpha(T-t)} \right) e^{\alpha(T-t)}$$

with terminal condition $g(x) = \cos(x_1 + \dots + x_d)$. Then the explicit solution of (4) is given by

$$u(t, x) = \cos(x_1 + \dots + x_d) e^{\alpha(T-t)}$$

Note that in this case we have $L = \{(0, 0), (1, 1)\}$.

References

- [1] Pierre Henry-Labordère, Nadia Oudjane, Xiaolu Tan, Nizar Touzi, and Xavier Warin. Branching diffusion representation of semilinear PDEs and Monte Carlo approximation. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 55(1):184 – 210, 2019.