Branching Diffusion Representations of Nonlinear PDEs

Christoph Belak
Technische Universität Berlin

Joint Work with **Daniel Hoffmann**, **Frank Seifried**, and **Lotte Schnell**.

Technische Universität Berlin

Stochastic Numerics and Statistical Learning Workshop
KAUST, May 29, 2023

Long Term Goal: efficient numerical methods for nonlinear, nonlocal PDEs such as

$$\max\{-\partial_t u - \mathcal{L}[u], u - \mathcal{M}[u]\} = 0 \qquad \text{on } [0,T) \times \mathcal{S}.$$

Long Term Goal: efficient numerical methods for nonlinear, nonlocal PDEs such as

$$\max\{-\partial_t u - \mathcal{L}[u], u - \mathcal{M}[u]\} = 0 \qquad \text{on } [0,T) \times \mathcal{S}.$$

This equation

- involves the infinitesimal generator $\mathcal{L}$ of a diffusion;

Long Term Goal: efficient numerical methods for nonlinear, nonlocal PDEs such as

$$\max\{-\partial_t u - \mathcal{L}[u], u - \mathcal{M}[u]\} = 0 \qquad \text{on } [0, T) \times \mathcal{S}.$$

This equation

- involves the infinitesimal generator $\mathcal{L}$ of a diffusion;
- involves a nonlocal operator $\mathcal{M}$;

Long Term Goal: efficient numerical methods for nonlinear, nonlocal PDEs such as

$$\max\{-\partial_t u - \mathcal{L}[u], u - \mathcal{M}[u]\} = 0 \qquad \text{on } [0, T) \times \mathcal{S}.$$

This equation

- involves the infinitesimal generator $\mathcal{L}$ of a diffusion;
- involves a nonlocal operator $\mathcal{M}$;
- is linked to stochastic impulse control problems;

Long Term Goal: efficient numerical methods for nonlinear, nonlocal PDEs such as

$$\max\{-\partial_t u - \mathcal{L}[u], u - \mathcal{M}[u]\} = 0 \qquad \text{on } [0,T) \times \mathcal{S}.$$

This equation

- involves the infinitesimal generator $\mathcal{L}$ of a diffusion;
- involves a nonlocal operator $\mathcal{M}$;
- is linked to stochastic impulse control problems;
- is formally fully nonlinear, nonlocal, has a free boundary, may have degeneracies in $\mathcal{L}$, possibly higher dimensional, ...;

Long Term Goal: efficient numerical methods for nonlinear, nonlocal PDEs such as

$$\max\{-\partial_t u - \mathcal{L}[u], u - \mathcal{M}[u]\} = 0 \qquad \text{on } [0,T) \times \mathcal{S}.$$

This equation

- ▷ involves the infinitesimal generator $\mathcal{L}$ of a diffusion;
- ▷ involves a nonlocal operator $\mathcal{M}$;
- ▷ is linked to stochastic impulse control problems;
- ▷ is formally fully nonlinear, nonlocal, has a free boundary, may have degeneracies in $\mathcal{L}$, possibly higher dimensional, ...;
- ▷ is to be approximated on (a subset of) $\mathcal{S}$ and not just a single point.

Long Term Goal: efficient numerical methods for nonlinear, nonlocal PDEs such as

$$\max\{-\partial_t u - \mathcal{L}[u], u - \mathcal{M}[u]\} = 0 \qquad \text{on } [0,T) \times \mathcal{S}.$$

This equation

- involves the infinitesimal generator $\mathcal{L}$ of a diffusion;
- involves a nonlocal operator $\mathcal{M}$;
- is linked to stochastic impulse control problems;
- is formally fully nonlinear, nonlocal, has a free boundary, may have degeneracies in $\mathcal{L}$, possibly higher dimensional, ...;
- is to be approximated on (a subset of) $\mathcal{S}$ and not just a single point.

The λ-regularized version takes the form

$$-\partial_t u - \mathcal{L}[u] + \lambda\big(u - \mathcal{M}[u]\big)_+ = 0$$

and allows us to reduce the problem to a nonlocal semilinear PDE.

How can we solve such PDEs numerically?

How can we solve such PDEs numerically?

- Finite Differences and Policy Iteration?

How can we solve such PDEs numerically?

- Finite Differences and Policy Iteration?
- Discrete-Time and Space Approximation of the Control Problem?

How can we solve such PDEs numerically?

- Finite Differences and Policy Iteration?
- Discrete-Time and Space Approximation of the Control Problem?
- Feynman-Kac Representation and Monte Carlo?

How can we solve such PDEs numerically?

- ▷ Finite Differences and Policy Iteration?
- ▷ Discrete-Time and Space Approximation of the Control Problem?
- ▷ Feynman-Kac Representation and Monte Carlo?
- ▷ Deep Learning or Reinforcement Learning?

How can we solve such PDEs numerically?

- ▷ Finite Differences and Policy Iteration?
- ▷ Discrete-Time and Space Approximation of the Control Problem?
- ▷ Feynman-Kac Representation and Monte Carlo?
- ▷ Deep Learning or Reinforcement Learning?
- ▷ ...

How can we solve such PDEs numerically?

- Finite Differences and Policy Iteration?
- Discrete-Time and Space Approximation of the Control Problem?
- Feynman-Kac Representation and Monte Carlo?
- Deep Learning or Reinforcement Learning?
- ...

In this talk, we present an approach based on a

Branching Diffusion Representation

of the solution of the PDE.

Overview

(1) Branching Diffusion Representations

(2) Main Theoretical Results

(3) A Numerical Example from Finance

(4) Outlook, Ideas, and Conclusion

Branching Diffusion Representations

In what follows, we consider a PDE of the form

$$\begin{aligned} \partial_t u(t,x) + \mathcal{L}[u](t,x) + f(t,x,u) &= 0, && (t,x) \in [0,T] \times \mathbb{R}^d, \\ u(T,x) &= g(x), && x \in \mathbb{R}^d. \end{aligned}$$

In what follows, we consider a PDE of the form

$$
\begin{aligned}
\partial_t u(t,x) + \mathcal{L}[u](t,x) + f(t,x,u) &= 0, && (t,x) \in [0,T] \times \mathbb{R}^d, \\
u(T,x) &= g(x), && x \in \mathbb{R}^d.
\end{aligned}
$$

Here $\mathcal{L}$ is an infinitesimal generator of the form

$$
\mathcal{L}[u](t,x) = \mu(t,x)^\top \mathrm{D}_x u(t,x) + \frac{1}{2} \operatorname{tr}\big[\sigma(t,x)\sigma(t,x)^\top \mathrm{D}_x^2 u(t,x)\big]
$$

In what follows, we consider a PDE of the form

$$\begin{aligned} \partial_t u(t,x) + \mathcal{L}[u](t,x) + f(t,x,u) &= 0, && (t,x) \in [0,T] \times \mathbb{R}^d, \\ u(T,x) &= g(x), && x \in \mathbb{R}^d. \end{aligned}$$

Here $\mathcal{L}$ is an infinitesimal generator of the form

$$\mathcal{L}[u](t,x) = \mu(t,x)^\top \mathrm{D}_x u(t,x) + \frac{1}{2} \operatorname{tr}\big[\sigma(t,x)\sigma(t,x)^\top \mathrm{D}_x^2 u(t,x)\big]$$

and the nonlocal nonlinearity is (for simplicity) of the form

$$f(t,x,u) = \sum_{i \in \mathcal{I}} c_i(t,x) \mathbb{E}\big[u(t,x+\Delta)^i\big]$$

for a random variable Δ and some $\mathcal{I} \subseteq \mathbb{N}_0$.

Towards the Feynman-Kac representation of this PDE, let

$$\mathrm{d}\overline{X}_t = \mu(t, \overline{X}_t)\mathrm{d}t + \sigma(t, \overline{X}_t)\mathrm{d}W_t, \qquad \overline{X}_0 = x.$$

Towards the Feynman-Kac representation of this PDE, let

$$\mathrm{d}\overline{X}_t = \mu(t, \overline{X}_t)\mathrm{d}t + \sigma(t, \overline{X}_t)\mathrm{d}W_t, \qquad \overline{X}_0 = x.$$

Then, at least informally,

$$u(0, x) = \mathbb{E}\Big[u\big(T, \overline{X}_T\big) + \int_0^T f\big(s, \overline{X}_s, u\big)\mathrm{d}s\Big]$$

Towards the Feynman-Kac representation of this PDE, let

$$\mathrm{d}\overline{X}_t = \mu(t, \overline{X}_t)\mathrm{d}t + \sigma(t, \overline{X}_t)\mathrm{d}W_t, \qquad \overline{X}_0 = x.$$

Then, at least informally,

$$\begin{aligned} u(0,x) &= \mathbb{E}\Big[u\big(T, \overline{X}_T\big) + \int_0^T f\big(s, \overline{X}_s, u\big)\mathrm{d}s\Big] \\ &= \mathbb{E}\Big[g\big(\overline{X}_T\big) + \int_0^T \sum_{i \in \mathcal{I}} c_i(s, \overline{X}_s) u(s, \overline{X}_s + \Delta)^i \mathrm{d}s\Big] \end{aligned}$$

if W is chosen independent of Δ.

$$u(0,x) = \mathbb{E}\Big[g\big(\overline{X}_T\big) + \int_0^T \sum_{i\in\mathcal{I}} c_i(s,\overline{X}_s)u(s,\overline{X}_s+\Delta)^i \mathrm{d}s\Big]$$

We randomize this representation by observing that

$$u(0,x) = \mathbb{E}\Big[g\big(\overline{X}_T\big) + \int_0^T \sum_{i\in\mathcal{I}} c_i(s,\overline{X}_s)u(s,\overline{X}_s+\Delta)^i \mathrm{d}s\Big]$$

We randomize this representation by observing that

$$
\begin{aligned}
u(0,x) &= \mathbb{E}\Big[g\big(\overline{X}_T\big) + \int_0^T \sum_{i\in\mathcal{I}} c_i(s,\overline{X}_s)u(s,\overline{X}_s+\Delta)^i \mathrm{d}s\Big] \\
&= \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T)}F(T) + \int_0^T \rho(s) \sum_{i\in\mathcal{I}} \frac{c_i(s,\overline{X}_s)}{\rho(s)}u(s,\overline{X}_s+\Delta)^i \mathrm{d}s\Big]
\end{aligned}
$$

where τ is an independent stopping time with strictly positive density ρ and

$$
F(T) \triangleq \int_T^\infty \rho(s)\mathrm{d}s.
$$

We randomize this representation by observing that

$$\begin{aligned}
u(0,x) &= \mathbb{E}\Big[g\big(\overline{X}_T\big) + \int_0^T \sum_{i\in\mathcal{I}} c_i(s,\overline{X}_s)u(s,\overline{X}_s+\Delta)^i \mathrm{d}s\Big] \\
&= \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T)}F(T) + \int_0^T \rho(s)\sum_{i\in\mathcal{I}} \frac{c_i(s,\overline{X}_s)}{\rho(s)}u(s,\overline{X}_s+\Delta)^i \mathrm{d}s\Big] \\
&= \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T)}\mathbb{1}_{\{\tau\geq T\}} + \sum_{i\in\mathcal{I}} \frac{c_i(\tau,\overline{X}_\tau)}{\rho(\tau)}u(\tau,\overline{X}_\tau+\Delta)^i \mathbb{1}_{\{\tau<T\}}\Big],
\end{aligned}$$

where τ is an independent stopping time with strictly positive density ρ and

$$F(T) \triangleq \int_T^\infty \rho(s)\mathrm{d}s.$$

Similarly, we can handle the sum since

$$u(0,x) = \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T)}\mathbb{1}_{\{\tau \geq T\}} + \sum_{i \in \mathcal{I}} \frac{c_i(\tau, \overline{X}_\tau)}{\rho(\tau)} u(\tau, \overline{X}_\tau + \Delta)^i \mathbb{1}_{\{\tau < T\}}\Big]$$

Similarly, we can handle the sum since

$$\begin{aligned} u(0,x) &= \mathbb{E}\Big[\frac{g(\overline{X}_T)}{F(T)}\mathbb{1}_{\{\tau\geq T\}} + \sum_{i\in\mathcal{I}} \frac{c_i(\tau,\overline{X}_\tau)}{\rho(\tau)} u(\tau,\overline{X}_\tau+\Delta)^i \mathbb{1}_{\{\tau<T\}}\Big] \\ &= \mathbb{E}\Big[\frac{g(\overline{X}_T)}{F(T)}\mathbb{1}_{\{\tau\geq T\}} + \sum_{i\in\mathcal{I}} p_i \frac{c_i(\tau,\overline{X}_\tau)}{\rho(\tau)p_i} u(\tau,\overline{X}_\tau+\Delta)^i \mathbb{1}_{\{\tau<T\}}\Big] \end{aligned}$$

where I is an independent random variable on $\mathcal{I}$ with strictly positive probability mass function $p = (p_i)_{i\in\mathcal{I}}$.

Similarly, we can handle the sum since

$$
\begin{aligned}
u(0,x) &= \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T)}\mathbb{1}_{\{\tau\geq T\}} + \sum_{i\in\mathcal{I}} \frac{c_i(\tau,\overline{X}_\tau)}{\rho(\tau)} u(\tau,\overline{X}_\tau+\Delta)^i \mathbb{1}_{\{\tau<T\}}\Big] \\
&= \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T)}\mathbb{1}_{\{\tau\geq T\}} + \sum_{i\in\mathcal{I}} p_i \frac{c_i(\tau,\overline{X}_\tau)}{\rho(\tau)p_i} u(\tau,\overline{X}_\tau+\Delta)^i \mathbb{1}_{\{\tau<T\}}\Big] \\
&= \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T)}\mathbb{1}_{\{\tau\geq T\}} + \frac{c_I(\tau,\overline{X}_\tau)}{\rho(\tau)p_I} u(\tau,\overline{X}_\tau+\Delta)^I \mathbb{1}_{\{\tau<T\}}\Big],
\end{aligned}
$$

where I is an independent random variable on $\mathcal{I}$ with strictly positive probability mass function $p = (p_i)_{i\in\mathcal{I}}$.

Similarly, we can handle the sum since

$$
\begin{aligned}
u(0,x) &= \mathbb{E}\Big[\frac{g(\overline{X}_T)}{F(T)}\mathbb{1}_{\{\tau\geq T\}} + \sum_{i\in\mathcal{I}} \frac{c_i(\tau,\overline{X}_\tau)}{\rho(\tau)} u(\tau,\overline{X}_\tau+\Delta)^i \mathbb{1}_{\{\tau<T\}}\Big] \\
&= \mathbb{E}\Big[\frac{g(\overline{X}_T)}{F(T)}\mathbb{1}_{\{\tau\geq T\}} + \sum_{i\in\mathcal{I}} p_i \frac{c_i(\tau,\overline{X}_\tau)}{\rho(\tau)p_i} u(\tau,\overline{X}_\tau+\Delta)^i \mathbb{1}_{\{\tau<T\}}\Big] \\
&= \mathbb{E}\Big[\frac{g(\overline{X}_T)}{F(T)}\mathbb{1}_{\{\tau\geq T\}} + \frac{c_I(\tau,\overline{X}_\tau)}{\rho(\tau)p_I} u(\tau,\overline{X}_\tau+\Delta)^I \mathbb{1}_{\{\tau<T\}}\Big],
\end{aligned}
$$

where I is an independent random variable on $\mathcal{I}$ with strictly positive probability mass function $p = (p_i)_{i\in\mathcal{I}}$.

The last expression is the branching diffusion representation in the first generation up to time $\tau \wedge T$.

$$u(0,x) = \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T)}\mathbb{1}_{\{\tau \geq T\}} + \frac{c_I(\tau, \overline{X}_\tau)}{\rho(\tau) p_I} u(\tau, \overline{X}_\tau + \Delta)^I \mathbb{1}_{\{\tau < T\}}\Big]$$

$$u(0,x) = \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T)}\mathbb{1}_{\{\tau \geq T\}} + \frac{c_I(\tau, \overline{X}_\tau)}{\rho(\tau)p_I} u(\tau, \overline{X}_\tau + \Delta)^I \mathbb{1}_{\{\tau < T\}}\Big]$$

For $\ell = 1, \dots, I$, let X^ℓ be an independent diffusions (driven by an independent Brownian motion W^ℓ) started in $(\tau, \overline{X}_\tau + \Delta)$. Then

$$\begin{aligned} u(0,x) &= \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T)}\mathbb{1}_{\{\tau \geq T\}} + \frac{c_I(\tau, \overline{X}_\tau)}{\rho(\tau)p_I} u(\tau, \overline{X}_\tau + \Delta)^I \mathbb{1}_{\{\tau < T\}}\Big] \\ &= \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T)}\mathbb{1}_{\{\tau \geq T\}} + \frac{c_I(\tau, \overline{X}_\tau)}{\rho(\tau)p_I} \mathbb{1}_{\{\tau < T\}} \prod_{\ell=1}^{I} u(\tau, X_\tau^\ell)\Big]. \end{aligned}$$

For $\ell = 1, \ldots, I$, let X^ℓ be an independent diffusions (driven by an independent Brownian motion W^ℓ) started in $(\tau, \overline{X}_\tau + \Delta)$. Then

$$\begin{aligned} u(0,x) &= \mathbb{E}\Big[\frac{g(\overline{X}_T)}{F(T)} \mathbb{1}_{\{\tau \geq T\}} + \frac{c_I(\tau, \overline{X}_\tau)}{\rho(\tau) p_I} u(\tau, \overline{X}_\tau + \Delta)^I \mathbb{1}_{\{\tau < T\}}\Big] \\ &= \mathbb{E}\Big[\frac{g(\overline{X}_T)}{F(T)} \mathbb{1}_{\{\tau \geq T\}} + \frac{c_I(\tau, \overline{X}_\tau)}{\rho(\tau) p_I} \mathbb{1}_{\{\tau < T\}} \prod_{\ell=1}^{I} u(\tau, X^\ell_\tau)\Big]. \end{aligned}$$

For each $\ell = 1, \ldots, I$, we have a another branching diffusion representation for $u(\tau, X^\ell_\tau)$ using

$$\text{independent copies } (\tau^\ell, \Delta^\ell, I^\ell, W^\ell) \qquad \text{of} \qquad (\tau, \Delta, I, W).$$

Plugging these into the expectation above yields the branching diffusion representation in the second generation up to time $(\tau + \tau^\ell) \wedge T$.

$$u(0,x) = \mathbb{E}\Big[\frac{g\big(\overline{X}_T\big)}{F(T)}\mathbb{1}_{\{\tau \geq T\}} + \frac{c_I(\tau, \overline{X}_\tau)}{\rho(\tau)p_I}\mathbb{1}_{\{\tau < T\}}\prod_{\ell=1}^{I} u(\tau, X_\tau^\ell)\Big]$$

Iterating this procedure gives the branching diffusion representation, which can be compactly written as

$$u(0,x) = \mathbb{E}[\Psi_{0,x}] \qquad \text{for} \qquad \Psi_{0,x} \triangleq \prod_{k\in\mathcal{K}} \frac{g\big(X_T^k\big)}{F(T-T^{k-})} \prod_{k\in\overline{\mathcal{K}}\setminus\mathcal{K}} \frac{c_{I^k}(T^k, X_{T^k}^k)}{\rho(T^k - T^{k-})p_{I^k}},$$

Iterating this procedure gives the branching diffusion representation, which can be compactly written as

$$u(0,x) = \mathbb{E}[\Psi_{0,x}] \qquad \text{for} \qquad \Psi_{0,x} \triangleq \prod_{k \in \mathcal{K}} \frac{g\big(X_T^k\big)}{F(T - T^{k-})} \prod_{k \in \overline{\mathcal{K}} \setminus \mathcal{K}} \frac{c_{I^k}(T^k, X_{T^k}^k)}{\rho(T^k - T^{k-}) p_{I^k}},$$

where

- $\mathcal{K}$ is the set of particles alive at time T;

Iterating this procedure gives the branching diffusion representation, which can be compactly written as

$$u(0,x) = \mathbb{E}[\Psi_{0,x}] \qquad \text{for} \qquad \Psi_{0,x} \triangleq \prod_{k\in\mathcal{K}} \frac{g\big(X_T^k\big)}{F(T-T^{k-})} \prod_{k\in\overline{\mathcal{K}}\setminus\mathcal{K}} \frac{c_{I^k}(T^k, X_{T^k}^k)}{\rho(T^k - T^{k-})p_{I^k}},$$

where

- $\mathcal{K}$ is the set of particles alive at time T;
- $\overline{\mathcal{K}}$ is the set of all particles alive before or at time T;

Iterating this procedure gives the branching diffusion representation, which can be compactly written as

$$u(0,x) = \mathbb{E}[\Psi_{0,x}] \qquad \text{for} \qquad \Psi_{0,x} \triangleq \prod_{k\in\mathcal{K}} \frac{g\left(X_T^k\right)}{F(T-T^{k-})} \prod_{k\in\overline{\mathcal{K}}\setminus\mathcal{K}} \frac{c_{I^k}(T^k, X_{T^k}^k)}{\rho(T^k - T^{k-})p_{I^k}},$$

where

- $\mathcal{K}$ is the set of particles alive at time T;
- $\overline{\mathcal{K}}$ is the set of all particles alive before or at time T;
- $k-$ denotes the parent of a particle k.

Iterating this procedure gives the branching diffusion representation, which can be compactly written as

$$u(0,x) = \mathbb{E}[\Psi_{0,x}] \qquad \text{for} \qquad \Psi_{0,x} \triangleq \prod_{k\in\mathcal{K}} \frac{g\big(X_T^k\big)}{F(T-T^{k-})} \prod_{k\in\overline{\mathcal{K}}\setminus\mathcal{K}} \frac{c_{I^k}(T^k, X_{T^k}^k)}{\rho(T^k-T^{k-})p_{I^k}},$$

where

- $\mathcal{K}$ is the set of particles alive at time T;
- $\overline{\mathcal{K}}$ is the set of all particles alive before or at time T;
- $k-$ denotes the parent of a particle k.

With this, we can compute $u(0,x)$ by (forward!) Monte Carlo.

Literature on Branching Diffusion Representations:

- McKean 1975

Literature on Branching Diffusion Representations:

- McKean 1975
- Zeroth-Order Case: Henry-Labordère 2012; Henry-Labordère, Tan, Touzi 2014

Literature on Branching Diffusion Representations:

- McKean 1975
- Zeroth-Order Case: Henry-Labordère 2012; Henry-Labordère, Tan, Touzi 2014
- First-Order Case: Henry-Labordère, Oudjane, Tan, Touzi, Warin 2019

Literature on Branching Diffusion Representations:

- McKean 1975
- Zeroth-Order Case: Henry-Labordère 2012; Henry-Labordère, Tan, Touzi 2014
- First-Order Case: Henry-Labordère, Oudjane, Tan, Touzi, Warin 2019
- Elliptic PDEs: Agarwal, Claisse 2020

Literature on Branching Diffusion Representations:

- McKean 1975
- Zeroth-Order Case: Henry-Labordère 2012; Henry-Labordère, Tan, Touzi 2014
- First-Order Case: Henry-Labordère, Oudjane, Tan, Touzi, Warin 2019
- Elliptic PDEs: Agarwal, Claisse 2020
- Locally Analytic, Lipschitz Case: Bouchard, Tan, Warin, Zou 2017; Bouchard, Tan, Warin 2019

Literature on Branching Diffusion Representations:

- McKean 1975
- Zeroth-Order Case: Henry-Labordère 2012; Henry-Labordère, Tan, Touzi 2014
- First-Order Case: Henry-Labordère, Oudjane, Tan, Touzi, Warin 2019
- Elliptic PDEs: Agarwal, Claisse 2020
- Locally Analytic, Lipschitz Case: Bouchard, Tan, Warin, Zou 2017; Bouchard, Tan, Warin 2019
- Nonlocal Case: Belak, Hoffmann, Seifried 2021

Literature on Branching Diffusion Representations:

- McKean 1975
- Zeroth-Order Case: Henry-Labordère 2012; Henry-Labordère, Tan, Touzi 2014
- First-Order Case: Henry-Labordère, Oudjane, Tan, Touzi, Warin 2019
- Elliptic PDEs: Agarwal, Claisse 2020
- Locally Analytic, Lipschitz Case: Bouchard, Tan, Warin, Zou 2017; Bouchard, Tan, Warin 2019
- Nonlocal Case: Belak, Hoffmann, Seifried 2021
- ... and more

Main Theoretical Results

B./Hoffmann/Seifried (2021)

Let u be a classical solution of the PDE. Under some natural integrability and growth assumptions, it holds that

$$u(0,x) = \mathbb{E}[\Psi_{0,x}] \qquad \text{for} \qquad \Psi_{0,x} \triangleq \prod_{k \in \mathcal{K}} \frac{g(X_T^k)}{F(T - T^{k-})} \prod_{k \in \overline{\mathcal{K}} \setminus \mathcal{K}} \frac{c_{I^k}(T^k, X_{T^k}^k)}{\rho(T^k - T^{k-}) p_{I^k}}.$$

B./Hoffmann/Seifried (2021)

Let u be a classical solution of the PDE. Under some natural integrability and growth assumptions, it holds that

$$u(0,x) = \mathbb{E}[\Psi_{0,x}] \qquad \text{for} \qquad \Psi_{0,x} \triangleq \prod_{k \in \mathcal{K}} \frac{g(X_T^k)}{F(T - T^{k-})} \prod_{k \in \overline{\mathcal{K}} \setminus \mathcal{K}} \frac{c_{I^k}(T^k, X_{T^k}^k)}{\rho(T^k - T^{k-}) p_{I^k}}.$$

Proof: Just make the arguments on the previous slides rigorous...

B./Hoffmann/Seifried (2021)

Define a function

$$u(t,x) \triangleq \mathbb{E}\big[\Psi_{t,x}\big], \qquad (t,x) \in [0,T] \times \mathbb{R}^d,$$

and, in addition to some natural integrability and growth assumptions, assume that

there exists $\varepsilon > 0$ such that $\{\Psi_{s,y} : |(s,y) - (t,x)| < \varepsilon\}$ is uniformly integrable.

Then u is a viscosity solution of the PDE.

B./Hoffmann/Seifried (2021)

Define a function

$$u(t,x) \triangleq \mathbb{E}\big[\Psi_{t,x}\big], \qquad (t,x) \in [0,T] \times \mathbb{R}^d,$$

and, in addition to some natural integrability and growth assumptions, assume that

there exists $\varepsilon > 0$ such that $\{\Psi_{s,y} : |(s,y) - (t,x)| < \varepsilon\}$ is uniformly integrable.

Then u is a viscosity solution of the PDE.

Proof: Go through the slides backwards to obtain the Feynman-Kac representation. Then use standard viscosity arguments.

B./Hoffmann/Seifried (2021)

Let $\kappa > 1$ and define

$$C_1 \triangleq \frac{\|g\|_\infty^\kappa}{F(T)^{\kappa-1}} \qquad \text{and} \qquad C_2 \triangleq \sup_{i\in\mathcal{I}, t\in[0,T]} \left(\frac{\|c_i\|_\infty}{\rho(t)p_i}\right)^{\kappa-1}$$

Then $\{\Psi_{t,x} : (t,x) \in [0,T] \times \mathbb{R}^d\}$ is bounded in $\mathrm{L}^\kappa(\mathbb{P})$, hence uniformly integrable, if either

(i) it holds that

$$\frac{C_1}{F(T)} \vee C_2 \leq 1;$$

(ii) the power series $\sum_{i\in\mathcal{I}} \|c_i\|_\infty x^i$ has infinite radius of convergence and it holds that

$$T < \int_{C_1}^{\infty} \Big(C_2 \sum_{i\in\mathcal{I}} \|c_i\|_\infty x^i\Big)^{-1} \mathrm{d}x.$$

Recall:

$$\textbf{(i)} \quad \frac{C_1}{F(T)} \vee C_2 \leq 1 \qquad \text{and} \qquad \textbf{(ii)} \quad T < \int_{C_1}^{\infty} \Big(C_2 \sum_{i \in \mathcal{I}} \|c_i\|_\infty x^i \Big)^{-1} \mathrm{d}x,$$

where

$$C_1 \triangleq \frac{\|g\|_\infty^\kappa}{F(T)^{\kappa-1}} \qquad \text{and} \qquad C_2 \triangleq \sup_{i \in \mathcal{I}, t \in [0,T]} \left(\frac{\|c_i\|_\infty}{\rho(t) p_i} \right)^{\kappa-1}.$$

Recall:

$$\text{(i)} \quad \frac{C_1}{F(T)} \vee C_2 \leq 1 \qquad \text{and} \qquad \text{(ii)} \quad T < \int_{C_1}^{\infty} \Big(C_2 \sum_{i \in \mathcal{I}} \|c_i\|_\infty x^i \Big)^{-1} \mathrm{d}x,$$

where

$$C_1 \triangleq \frac{\|g\|_\infty^\kappa}{F(T)^{\kappa-1}} \qquad \text{and} \qquad C_2 \triangleq \sup_{i \in \mathcal{I}, t \in [0,T]} \left(\frac{\|c_i\|_\infty}{\rho(t) p_i} \right)^{\kappa-1}.$$

Discussion of the previous result:

- In case of (i), in the proof one simply bounds each factor in the product making up $|\Psi_{t,x}|^\kappa$ to obtain the result.

Recall:

$$\textbf{(i)} \quad \frac{C_1}{F(T)} \vee C_2 \leq 1 \qquad \text{and} \qquad \textbf{(ii)} \quad T < \int_{C_1}^{\infty} \Big(C_2 \sum_{i \in \mathcal{I}} \|c_i\|_\infty x^i \Big)^{-1} \mathrm{d}x,$$

where

$$C_1 \triangleq \frac{\|g\|_\infty^\kappa}{F(T)^{\kappa-1}} \qquad \text{and} \qquad C_2 \triangleq \sup_{i \in \mathcal{I}, t \in [0,T]} \Big(\frac{\|c_i\|_\infty}{\rho(t) p_i} \Big)^{\kappa-1}.$$

Discussion of the previous result:

- In case of (i), in the proof one simply bounds each factor in the product making up $|\Psi_{t,x}|^\kappa$ to obtain the result.
- In case of (ii), one bounds $|\Psi_{t,x}|^\kappa$ by a random variable, which can be interpreted as a branching diffusion representation of an ODE. The bound in (ii) is necessary and sufficient to guarantee that the ODE does not blow up.

Recall:

$$\textbf{(i)} \quad \frac{C_1}{F(T)} \vee C_2 \leq 1 \qquad \text{and} \qquad \textbf{(ii)} \quad T < \int_{C_1}^{\infty} \Big(C_2 \sum_{i \in \mathcal{I}} \|c_i\|_\infty x^i \Big)^{-1} \mathrm{d}x,$$

where

$$C_1 \triangleq \frac{\|g\|_\infty^\kappa}{F(T)^{\kappa-1}} \qquad \text{and} \qquad C_2 \triangleq \sup_{i \in \mathcal{I}, t \in [0,T]} \left(\frac{\|c_i\|_\infty}{\rho(t) p_i} \right)^{\kappa-1}.$$

Discussion of the previous result:

- In case of (i), in the proof one simply bounds each factor in the product making up $|\Psi_{t,x}|^\kappa$ to obtain the result.
- In case of (ii), one bounds $|\Psi_{t,x}|^\kappa$ by a random variable, which can be interpreted as a branching diffusion representation of an ODE. The bound in (ii) is necessary and sufficient to guarantee that the ODE does not blow up.
- The results are a bit discouraging. For the branching representation to work, one has to impose quite strong boundedness assumptions and one needs a short time horizon.

Recall:

$$\textbf{(i)} \quad \frac{C_1}{F(T)} \vee C_2 \leq 1 \qquad \text{and} \qquad \textbf{(ii)} \quad T < \int_{C_1}^{\infty} \Big(C_2 \sum_{i \in \mathcal{I}} \|c_i\|_\infty x^i \Big)^{-1} \mathrm{d}x,$$

where

$$C_1 \triangleq \frac{\|g\|_\infty^\kappa}{F(T)^{\kappa-1}} \qquad \text{and} \qquad C_2 \triangleq \sup_{i \in \mathcal{I}, t \in [0,T]} \left(\frac{\|c_i\|_\infty}{\rho(t) p_i} \right)^{\kappa-1}.$$

Discussion of the previous result:

- In case of (i), in the proof one simply bounds each factor in the product making up $|\Psi_{t,x}|^\kappa$ to obtain the result.
- In case of (ii), one bounds $|\Psi_{t,x}|^\kappa$ by a random variable, which can be interpreted as a branching diffusion representation of an ODE. The bound in (ii) is necessary and sufficient to guarantee that the ODE does not blow up.
- The results are a bit discouraging. For the branching representation to work, one has to impose quite strong boundedness assumptions and one needs a short time horizon.
- This shows in practice! Ideas how to mitigate this problem are very much welcome!

A Numerical Example from Finance

Consider the following situation:

- Example inspired by Pykhtin, Sokol 2013; Mercurio, Li 2015; Brigo, Pede, Petrelli 2019

Consider the following situation:

- Example inspired by Pykhtin, Sokol 2013; Mercurio, Li 2015; Brigo, Pede, Petrelli 2019
- You hold a position in an option $g(X_T)$ written on an underlying X.

Consider the following situation:

- Example inspired by Pykhtin, Sokol 2013; Mercurio, Li 2015; Brigo, Pede, Petrelli 2019
- You hold a position in an option $g(X_T)$ written on an underlying X.
- Your counterparty, say a bank, is defaultable and systemically important.

Consider the following situation:

- Example inspired by Pykhtin, Sokol 2013; Mercurio, Li 2015; Brigo, Pede, Petrelli 2019
- You hold a position in an option $g(X_T)$ written on an underlying X.
- Your counterparty, say a bank, is defaultable and systemically important.
- This means: if the bank defaults, it has an impact on the prices of the underlyings.

Consider the following situation:

- Example inspired by Pykhtin, Sokol 2013; Mercurio, Li 2015; Brigo, Pede, Petrelli 2019
- You hold a position in an option $g(X_T)$ written on an underlying X.
- Your counterparty, say a bank, is defaultable and systemically important.
- This means: if the bank defaults, it has an impact on the prices of the underlyings.
- Thus, this has an impact on the value of the option.

Consider the following situation:

- Example inspired by Pykhtin, Sokol 2013; Mercurio, Li 2015; Brigo, Pede, Petrelli 2019
- You hold a position in an option $g(X_T)$ written on an underlying X.
- Your counterparty, say a bank, is defaultable and systemically important.
- This means: if the bank defaults, it has an impact on the prices of the underlyings.
- Thus, this has an impact on the value of the option.
- Risk neutral pricing yields a price of

$$\frac{V_t}{B_t} = \mathbb{E}\Big[\frac{g(X_T)}{B_T}\mathbb{1}_{\{\tau>T\}} + \frac{h(V_\tau)}{B_\tau}\mathbb{1}_{\{\tau\leq T\}}\Big] \qquad \text{on } \{\tau > T\}.$$

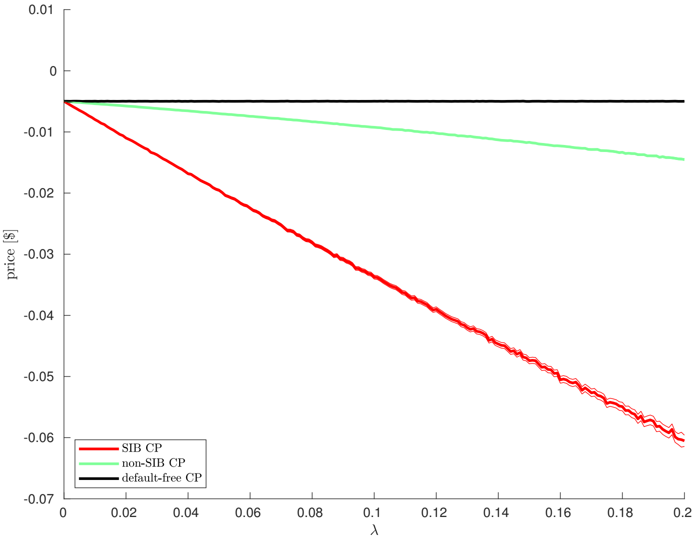
Here, τ is the default time, $h(v) = Rv_+ - v_-$ the recovery value, and V_τ is the time-τ mark-to-market value of an identical option with a systemically important counterparty which has not defaulted yet (otherwise identical and based on post-default market conditions).

Consider the following situation:

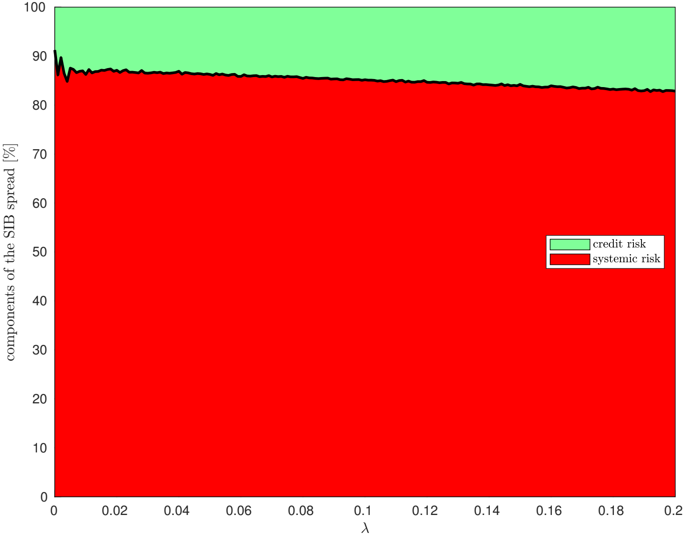
- Example inspired by Pykhtin, Sokol 2013; Mercurio, Li 2015; Brigo, Pede, Petrelli 2019
- You hold a position in an option $g(X_T)$ written on an underlying X.
- Your counterparty, say a bank, is defaultable and systemically important.
- This means: if the bank defaults, it has an impact on the prices of the underlyings.
- Thus, this has an impact on the value of the option.
- Risk neutral pricing yields a price of

$$\frac{V_t}{B_t} = \mathbb{E}\Big[\frac{g(X_T)}{B_T}\mathbb{1}_{\{\tau>T\}} + \frac{h(V_\tau)}{B_\tau}\mathbb{1}_{\{\tau\le T\}}\Big] \qquad \text{on } \{\tau > T\}.$$

  Here, τ is the default time, $h(v) = Rv_+ - v_-$ the recovery value, and V_τ is the time-τ mark-to-market value of an identical option with a systemically important counterparty which has not defaulted yet (otherwise identical and based on post-default market conditions).
- If one approximates h by a polynomial, the associated pricing PDE can be treated with the branching diffusion approach. We consider a put on a basket with about 500 underlyings.



0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
-0.06
-0.07
price [$]
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18
0.2
λ
SIB CP
non-SIB CP
default-free CP


100
90
80
70
60
50
40
30
20
10
0
components of the SIB spread [%]
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18
0.2
λ
credit risk
systemic risk

Outlook, Ideas, and Conclusion

How can we generalize the branching diffusion mechanism to work with less stringent assumptions?

- The underlying issue is that there is too much randomization: unbounded number of branches, unbounded particles per branch, randomness in the diffusion, ...

How can we generalize the branching diffusion mechanism to work with less stringent assumptions?

- The underlying issue is that there is too much randomization: unbounded number of branches, unbounded particles per branch, randomness in the diffusion, ...
- Idea: use a (monotone, stable, consistent) discrete approximation of the PDE, which still allows for a probabilistic representation (finite differences!).

How can we generalize the branching diffusion mechanism to work with less stringent assumptions?

- ▷ The underlying issue is that there is too much randomization: unbounded number of branches, unbounded particles per branch, randomness in the diffusion, ...
- ▷ Idea: use a (monotone, stable, consistent) discrete approximation of the PDE, which still allows for a probabilistic representation (finite differences!).
- ▷ First results indicate, that a careful balancing of spatial grid and time grid allows us to circumvent the strong boundedness / smallness of time horizon assumptions if the nonlinearity behaves sufficiently nice.

How can we generalize the branching diffusion mechanism to work with less stringent assumptions?

- ▷ The underlying issue is that there is too much randomization: unbounded number of branches, unbounded particles per branch, randomness in the diffusion, ...
- ▷ Idea: use a (monotone, stable, consistent) discrete approximation of the PDE, which still allows for a probabilistic representation (finite differences!).
- ▷ First results indicate, that a careful balancing of spatial grid and time grid allows us to circumvent the strong boundedness / smallness of time horizon assumptions if the nonlinearity behaves sufficiently nice.
- ▷ Interesting to note: despite working with finite differences, one does not have to impose artificial boundary conditions.

How can we generalize the branching diffusion mechanism to work with less stringent assumptions?

- ▷ The underlying issue is that there is too much randomization: unbounded number of branches, unbounded particles per branch, randomness in the diffusion, ...
- ▷ Idea: use a (monotone, stable, consistent) discrete approximation of the PDE, which still allows for a probabilistic representation (finite differences!).
- ▷ First results indicate, that a careful balancing of spatial grid and time grid allows us to circumvent the strong boundedness / smallness of time horizon assumptions if the nonlinearity behaves sufficiently nice.
- ▷ Interesting to note: despite working with finite differences, one does not have to impose artificial boundary conditions.
- ▷ Depending on the structure of the diffusion, this allows the approximation of finite difference schemes in high dimensions.

How can we generalize the branching diffusion mechanism to work with less stringent assumptions?

- The underlying issue is that there is too much randomization: unbounded number of branches, unbounded particles per branch, randomness in the diffusion, ...
- Idea: use a (monotone, stable, consistent) discrete approximation of the PDE, which still allows for a probabilistic representation (finite differences!).
- First results indicate, that a careful balancing of spatial grid and time grid allows us to circumvent the strong boundedness / smallness of time horizon assumptions if the nonlinearity behaves sufficiently nice.
- Interesting to note: despite working with finite differences, one does not have to impose artificial boundary conditions.
- Depending on the structure of the diffusion, this allows the approximation of finite difference schemes in high dimensions.
- We are still in the process of experimenting and writing up. Stay tuned!

Conclusion:

- Branching Diffusion Representations are an interesting take on probabilistic representations of PDEs.

Conclusion:

- Branching Diffusion Representations are an interesting take on probabilistic representations of PDEs.
- Under strong assumptions, they yield well performing numerical schemes.

Conclusion:

- Branching Diffusion Representations are an interesting take on probabilistic representations of PDEs.
- Under strong assumptions, they yield well performing numerical schemes.
- Despite these strong assumptions, there are cases in which branching diffusions can be considered the method of choice.

Conclusion:

- Branching Diffusion Representations are an interesting take on probabilistic representations of PDEs.
- Under strong assumptions, they yield well performing numerical schemes.
- Despite these strong assumptions, there are cases in which branching diffusions can be considered the method of choice.
- We are still thinking about extensions. Feel free to join us!

Conclusion:

- Branching Diffusion Representations are an interesting take on probabilistic representations of PDEs.
- Under strong assumptions, they yield well performing numerical schemes.
- Despite these strong assumptions, there are cases in which branching diffusions can be considered the method of choice.
- We are still thinking about extensions. Feel free to join us!

Thanks for your attention!