Numerical approximation of singular FBSDEs: application to carbon markets

J-F Chassagneux (Université Paris Cité)
based on joint works with
M. Yang (Université Paris Cité)

9th Colloquium on BSDEs and MF system
Annecy, 2022, June 27 - July 1

Outline

Carbon markets

- Carbon dioxide (CO_2) emission have a negative impact on the environment.
- Carbon markets are implemented to 'price' this and hopefully carbon emission reduction could be achieved

Carbon markets

- Carbon dioxide (CO_2) emission have a negative impact on the environment.
- Carbon markets are implemented to 'price' this and hopefully carbon emission reduction could be achieved
- Since 2005, the EU has had its own emissions trading system (ETS): an example of *cap-and-trade scheme*
  - A central authority set a limit on pollutant emission during a given period. Allowances are allocated to participating installations (via auctioning).
  - The total amount of allowances is the aggregated *cap*.
  - At the end of the period, each participating installation has to surrender an allowance for each unit of emission or pay a penalty.
  - During the period, participants can trade the allowances.

Carbon markets

- Carbon dioxide (CO_2) emission have a negative impact on the environment.
- Carbon markets are implemented to 'price' this and hopefully carbon emission reduction could be achieved
- Since 2005, the EU has had its own emissions trading system (ETS): an example of *cap-and-trade scheme*
  - A central authority set a limit on pollutant emission during a given period. Allowances are allocated to participating installations (via auctioning).
  - The total amount of allowances is the aggregated *cap*.
  - At the end of the period, each participating installation has to surrender an allowance for each unit of emission or pay a penalty.
  - During the period, participants can trade the allowances.
- China, whose carbon emissions make up approximately one quarter of the global total, has launched a national emissions trading scheme in July 2021 (with various pilot schemes already running)

EUA price (*tradingeconomics.com*)

Euros per tCO_2 (compare with China ETS price: 8.4 euros/tCO_2 on 1 April 2022)

Main features

• Model based on FBSDEs see e.g. Carmona, Delarue, Espinosa & Touzi (2013), Carmona & Delarue (2013), Howison & Schwarz (2015), C.-Chotai-Crisan (2020)

Main features

• Model based on FBSDEs see e.g. Carmona, Delarue, Espinosa & Touzi (2013), Carmona & Delarue (2013), Howison & Schwarz (2015), C.-Chotai-Crisan (2020)
• Three main processes on one period $[0, T]$.

Main features

- Model based on FBSDEs see e.g. Carmona, Delarue, Espinosa & Touzi (2013), Carmona & Delarue (2013), Howison & Schwarz (2015), C.-Chotai-Crisan (2020)
- Three main processes on one period $[0, T]$.
  1. The spot allowance price Y: we assume that the market is frictionless and arbitrage-free and that there is a probability such that $(e^{-rt}Y_t)_{0 \le t \le T}$ is a martingale, namely

$$\mathrm{d}Y_t = rY_t\mathrm{d}t + Z_t\mathrm{d}W_t$$

     r is the interest rate, Z is a square integrable process.

Main features

- Model based on FBSDEs see e.g. Carmona, Delarue, Espinosa & Touzi (2013), Carmona & Delarue (2013), Howison & Schwarz (2015), C.-Chotai-Crisan (2020)
- Three main processes on one period $[0, T]$.

1. The spot allowance price Y: we assume that the market is frictionless and arbitrage-free and that there is a probability such that $(e^{-rt}Y_t)_{0\le t\le T}$ is a martingale, namely

$$\mathrm{d}Y_t = rY_t\mathrm{d}t + Z_t\mathrm{d}W_t$$

   r is the interest rate, Z is a square integrable process.

2. Auxiliary process P:

$$\mathrm{d}P_t = b(P_t)\mathrm{d}t + \sigma(P_t)\mathrm{d}W_t$$

   Represent state variables that trigger the emission process (Electricity price or demand & fuel prices etc.) Fundamentals that are linked to goods emitting CO_2.

Main features

- Model based on FBSDEs see e.g. Carmona, Delarue, Espinosa & Touzi (2013), Carmona & Delarue (2013), Howison & Schwarz (2015), C.-Chotai-Crisan (2020)
- Three main processes on one period $[0, T]$.

1. The spot allowance price Y: we assume that the market is frictionless and arbitrage-free and that there is a probability such that $(e^{-rt}Y_t)_{0\le t\le T}$ is a martingale, namely

$$\mathrm{d}Y_t = rY_t\mathrm{d}t + Z_t\mathrm{d}W_t$$

   r is the interest rate, Z is a square integrable process.

2. Auxiliary process P:

$$\mathrm{d}P_t = b(P_t)\mathrm{d}t + \sigma(P_t)\mathrm{d}W_t$$

   Represent state variables that trigger the emission process (Electricity price or demand & fuel prices etc.) Fundamentals that are linked to goods emitting CO_2.

3. Emission process E: cumulative process with impact from the allowance price

$$\mathrm{d}E_t = \mu(P_t, Y_t)\mathrm{d}t$$

   $\hookrightarrow$ μ is decreasing in Y to take into account feedback of the allowance price

Associated singular FBSDE

- System of Equations: $0 \leq t \leq T$

$$
\begin{aligned}
\mathrm{d}P_t &= b(P_t)\,\mathrm{d}t + \sigma(P_t)\,\mathrm{d}W_t, && \text{(forward)}\\
\mathrm{d}E_t &= \mu(P_t, Y_t)\,\mathrm{d}t, && \text{(forward coupled)}\\
\mathrm{d}Y_t &= rY_t\,\mathrm{d}t + Z_t\,\mathrm{d}W_t, && \text{(backward)}
\end{aligned}
$$

E_0, P_0 is known but Y_0 is unknown!

Associated singular FBSDE

- System of Equations: $0 \le t \le T$

$$\begin{aligned} \mathrm{d}P_t &= b(P_t)\,\mathrm{d}t + \sigma(P_t)\,\mathrm{d}W_t, && \text{(forward)} \\ \mathrm{d}E_t &= \mu(P_t, Y_t)\,\mathrm{d}t, && \text{(forward coupled)} \\ \mathrm{d}Y_t &= rY_t\,\mathrm{d}t + Z_t\,\mathrm{d}W_t, && \text{(backward)} \end{aligned}$$

  E_0, P_0 is known but Y_0 is unknown!

- The terminal condition for the Allowance price Y. There is a cap Λ on the total emission set by the regulator
  1. If non-compliance i.e. $E_T > \Lambda$ then the penalty ρ is paid so $Y_T = \rho$
  2. If compliance i.e. $E_T < \Lambda$ then the Allowance is worth nothing (Emission regulation stops at the end of the period) so $Y_T = 0$

  $\hookrightarrow$ $Y_T = \phi(E_T) := \rho \mathbf{1}_{\{E_T > \Lambda\}}$ and $Y_t = e^{-r(T-t)}\mathbb{E}[Y_T | \mathcal{F}_t]$

Results for one-period model

- Carmona and Delarue (2013), there exists a unique solution to:

$$
\begin{aligned}
\mathrm{d}P_t &= b(P_t)\,\mathrm{d}t + \sigma(P_t)\,\mathrm{d}W_t,\\
\mathrm{d}E_t &= \mu(P_t, Y_t)\,\mathrm{d}t,\\
\mathrm{d}Y_t &= rY_t\,\mathrm{d}t + Z_t\,\mathrm{d}W_t,
\end{aligned}
$$

with terminal condition: $\phi(E_T) = \rho \mathbf{1}_{\{E_T > \Lambda\}} \le Y_T \le \rho \mathbf{1}_{\{E_T \ge \Lambda\}} =: \phi_+(E_T)$.
There exists a decoupling field s.t. $Y_t = v(t, P_t, E_t)$ for $t < T$.

Results for one-period model

- Carmona and Delarue (2013), there exists a unique solution to:

$$\begin{aligned} \mathrm{d}P_t &= b(P_t)\,\mathrm{d}t + \sigma(P_t)\,\mathrm{d}W_t, \\ \mathrm{d}E_t &= \mu(P_t, Y_t)\,\mathrm{d}t, \\ \mathrm{d}Y_t &= rY_t\,\mathrm{d}t + Z_t\,\mathrm{d}W_t, \end{aligned}$$

  with terminal condition: $\phi(E_T) = \rho\mathbf{1}_{\{E_T > \Lambda\}} \leq Y_T \leq \rho\mathbf{1}_{\{E_T \geq \Lambda\}} =: \phi_+(E_T)$.
  There exists a decoupling field s.t. $Y_t = v(t, P_t, E_t)$ for $t < T$.

- The decoupling field v is the "entropy" solution to

$$\partial_t v + \mu(p, v)\partial_e v + \mathcal{L}_p v = rv, \quad \text{and } v(T, e, p) = \phi(e)$$

Results for one-period model

- Carmona and Delarue (2013), there exists a unique solution to:

$$\begin{aligned} \mathrm{d}P_t &= b(P_t)\,\mathrm{d}t + \sigma(P_t)\,\mathrm{d}W_t, \\ \mathrm{d}E_t &= \mu(P_t, Y_t)\,\mathrm{d}t, \\ \mathrm{d}Y_t &= rY_t\,\mathrm{d}t + Z_t\,\mathrm{d}W_t, \end{aligned}$$

  with terminal condition: $\phi(E_T) = \rho\mathbf{1}_{\{E_T > \Lambda\}} \leq Y_T \leq \rho\mathbf{1}_{\{E_T \geq \Lambda\}} =: \phi_+(E_T)$.
  There exists a decoupling field s.t. $Y_t = v(t, P_t, E_t)$ for $t < T$.

- The decoupling field v is the "entropy" solution to

$$\partial_t v + \mu(p, v)\partial_e v + \mathcal{L}_p v = rv, \quad \text{and } v(T, e, p) = \phi(e)$$

  $\hookrightarrow$ v is Lipschitz in p and non decreasing in e.

Results for one-period model

- Carmona and Delarue (2013), there exists a unique solution to:

$$\mathrm{d}P_t = b(P_t)\,\mathrm{d}t + \sigma(P_t)\,\mathrm{d}W_t,$$
$$\mathrm{d}E_t = \mu(P_t, Y_t)\,\mathrm{d}t,$$
$$\mathrm{d}Y_t = rY_t\,\mathrm{d}t + Z_t\,\mathrm{d}W_t,$$

with terminal condition: $\phi(E_T) = \rho\mathbf{1}_{\{E_T > \Lambda\}} \leq Y_T \leq \rho\mathbf{1}_{\{E_T \geq \Lambda\}} =: \phi_+(E_T)$.
There exists a decoupling field s.t. $Y_t = v(t, P_t, E_t)$ for $t < T$.

- The decoupling field v is the "entropy" solution to

$$\partial_t v + \mu(p, v)\partial_e v + \mathcal{L}_p v = rv, \quad \text{and } v(T, e, p) = \phi(e)$$

$\hookrightarrow$ v is Lipschitz in p and non decreasing in e.
$\hookrightarrow$ $\partial_e v$ explodes at T near Λ, we only know $|\partial_e v(t, p, e)| \leq \frac{C}{T-t}$

Results for one-period model

- Carmona and Delarue (2013), there exists a unique solution to:

$$\begin{aligned} \mathrm{d}P_t &= b(P_t)\,\mathrm{d}t + \sigma(P_t)\,\mathrm{d}W_t, \\ \mathrm{d}E_t &= \mu(P_t, Y_t)\,\mathrm{d}t, \\ \mathrm{d}Y_t &= rY_t\,\mathrm{d}t + Z_t\,\mathrm{d}W_t, \end{aligned}$$

  with terminal condition: $\phi(E_T) = \rho\mathbf{1}_{\{E_T > \Lambda\}} \le Y_T \le \rho\mathbf{1}_{\{E_T \ge \Lambda\}} =: \phi_+(E_T)$.
  There exists a decoupling field s.t. $Y_t = v(t, P_t, E_t)$ for $t < T$.
- The decoupling field v is the "entropy" solution to

$$\partial_t v + \mu(p, v)\partial_e v + \mathcal{L}_p v = rv, \quad \text{and } v(T, e, p) = \phi(e)$$

  $\hookrightarrow$ v is Lipschitz in p and non decreasing in e.
  $\hookrightarrow$ $\partial_e v$ explodes at T near Λ, we only know $|\partial_e v(t, p, e)| \le \frac{C}{T-t}$
  $\hookrightarrow$ Set $\mu(p, v) = -v$ and $\mathcal{L}_p = 0, r = 0$. One obtains a 'backward' inviscid Burgers equation...

Results for one-period model

- Carmona and Delarue (2013), there exists a unique solution to:

$$\begin{aligned} \mathrm{d}P_t &= b(P_t)\,\mathrm{d}t + \sigma(P_t)\,\mathrm{d}W_t, \\ \mathrm{d}E_t &= \mu(P_t, Y_t)\,\mathrm{d}t, \\ \mathrm{d}Y_t &= rY_t\,\mathrm{d}t + Z_t\,\mathrm{d}W_t, \end{aligned}$$

  with terminal condition: $\phi(E_T) = \rho\mathbf{1}_{\{E_T > \Lambda\}} \le Y_T \le \rho\mathbf{1}_{\{E_T \ge \Lambda\}} =: \phi_+(E_T)$.
  There exists a decoupling field s.t. $Y_t = v(t, P_t, E_t)$ for $t < T$.

- The decoupling field v is the "entropy" solution to

$$\partial_t v + \mu(p, v)\partial_e v + \mathcal{L}_p v = rv, \quad \text{and } v(T, e, p) = \phi(e)$$

  $\hookrightarrow$ v is Lipschitz in p and non decreasing in e.
  $\hookrightarrow$ $\partial_e v$ explodes at T near Λ, we only know $|\partial_e v(t,p,e)| \le \frac{C}{T-t}$
  $\hookrightarrow$ Set $\mu(p, v) = -v$ and $\mathcal{L}_p = 0$, $r = 0$. One obtains a 'backward' inviscid Burgers equation...

- Multi-period model (finite or infinite number of period): Dan Crisan's talk!

Outline

A numerical Toy model

- One-period Toy model ($r = 0$), dimension $d + 1$, $\sigma > 0$:

$$\mathrm{d}P_t = \sigma \mathrm{d}W_t,\ \mathrm{d}E_t = \left(\frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} P_t^\ell - Y_t \right) \mathrm{d}t,\ \mathrm{d}Y_t = Z_t \cdot \mathrm{d}W_t,$$

and "$Y_T = \mathbf{1}_{[1,\infty)}(E_T)$".

A numerical Toy model

- One-period Toy model ($r = 0$), dimension $d + 1$, $\sigma > 0$:
$$\mathrm{d}P_t = \sigma \mathrm{d}W_t,\ \mathrm{d}E_t = \left(\frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} P_t^\ell - Y_t\right) \mathrm{d}t,\ \mathrm{d}Y_t = Z_t \cdot \mathrm{d}W_t,$$
and "$Y_T = \mathbf{1}_{[1,\infty)}(E_T)$".
- The quasi-linear pde associated is:
$$\partial_t v + \left(\frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} p^\ell - v\right) \partial_e v + \frac{\sigma^2}{2} \sum_{\ell=1}^{d} \partial^2_{p_\ell p_\ell} v = 0$$

A numerical Toy model

- One-period Toy model ($r = 0$), dimension $d+1$, $\sigma > 0$:

$$\mathrm{d}P_t = \sigma \mathrm{d}W_t,\ \mathrm{d}E_t = \left(\frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} P_t^{\ell} - Y_t\right) \mathrm{d}t,\ \mathrm{d}Y_t = Z_t \cdot \mathrm{d}W_t,$$

and "$Y_T = \mathbf{1}_{[1,\infty)}(E_T)$".

- The quasi-linear pde associated is:

$$\partial_t v + \left(\frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} p^{\ell} - v\right) \partial_e v + \frac{\sigma^2}{2} \sum_{\ell=1}^{d} \partial^2_{p_\ell p_\ell} v = 0$$

- Reduced to one dimension via $v(t,p,e) = u(t, e + (T-t)\frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} p^{\ell})$ with

$$\partial_t u - u \partial_\xi u + \frac{\sigma^2 (T-t)^2}{2} \partial^2_{\xi\xi} u = 0 \text{ and } u(T,\xi) = \mathbf{1}_{\{\xi \geq 1\}}$$

A numerical Toy model

- One-period Toy model ($r = 0$), dimension $d + 1$, $\sigma > 0$:
$$\mathrm{d}P_t = \sigma \mathrm{d}W_t,\ \mathrm{d}E_t = \left(\frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} P_t^\ell - Y_t\right) \mathrm{d}t,\ \mathrm{d}Y_t = Z_t \cdot \mathrm{d}W_t,$$
and "$Y_T = \mathbf{1}_{[1,\infty)}(E_T)$".
- The quasi-linear pde associated is:
$$\partial_t v + \left(\frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} p^\ell - v\right) \partial_e v + \frac{\sigma^2}{2} \sum_{\ell=1}^{d} \partial^2_{p_\ell p_\ell} v = 0$$
- Reduced to one dimension via $v(t,p,e) = u(t, e + (T-t)\frac{1}{\sqrt{d}} \sum_{\ell=1}^{d} p^\ell)$ with
$$\partial_t u - u \partial_\xi u + \frac{\sigma^2 (T-t)^2}{2} \partial^2_{\xi\xi} u = 0 \text{ and } u(T,\xi) = \mathbf{1}_{\{\xi \geq 1\}}$$

$\hookrightarrow$ Particle method associated to scalar conservation law can be used (Bossy, Jourdain, Tallay...) to get a proxy for the true solution: $e \mapsto v(0,0,e)$.

Toward a probabilistic scheme

- When $d = 1$ (dimension of P), PDE methods can be used see e.g. Howison-Schwarz (2012), C.-Chotai-Crisan (2020).

Toward a probabilistic scheme

- When $d = 1$ (dimension of P), PDE methods can be used see e.g. Howison-Schwarz (2012), C.-Chotai-Crisan (2020).
- However in applications d is larger... ($d \geq 3$ e.g. for electricity generation sector producer: demand and two fuel prices).

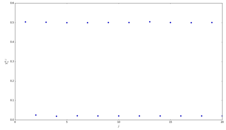
Toward a probabilistic scheme

- When $d = 1$ (dimension of P), PDE methods can be used see e.g. Howison-Schwarz (2012), C.-Chotai-Crisan (2020).
- However in applications d is larger... ($d \geq 3$ e.g. for electricity generation sector producer: demand and two fuel prices).
- We test "classical" FBSDE methods ($d = 1$):
  - Bender-Zhang Method (decoupling via Picard iteration+regression)
  - Delarue-Menozzi Method (probabilistic layer method, decoupling via predictor method)
  - Deep FBSDE solver (E-Han-Jentzen) learning method+DNN

Toward a probabilistic scheme

- When $d = 1$ (dimension of P), PDE methods can be used see e.g. Howison-Schwarz (2012), C.-Chotai-Crisan (2020).
- However in applications d is larger... ($d \geq 3$ e.g. for electricity generation sector producer: demand and two fuel prices).
- We test "classical" FBSDE methods ($d = 1$):
  - Bender-Zhang Method (decoupling via Picard iteration+regression)
  - Delarue-Menozzi Method (probabilistic layer method, decoupling via predictor method)
  - Deep FBSDE solver (E-Han-Jentzen) learning method+DNN

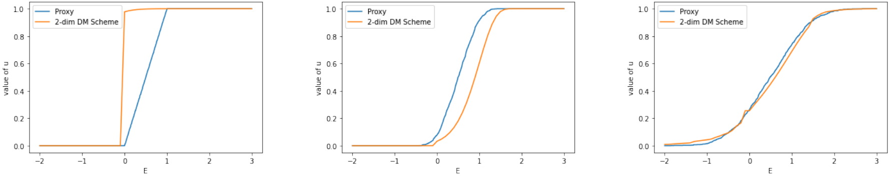
• Results for Bender-Zhang Method to compute Y_0:



Iterations do not converge (regularisation would help but difficult to tune)

Other methods

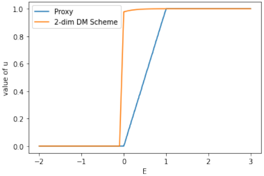
- Results for Delarue-Menozzi scheme:



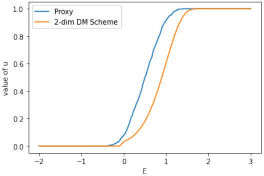
(a) $\sigma = 0.01$ (b) $\sigma = 0.3$ (c) $\sigma = 1.0$

Other methods

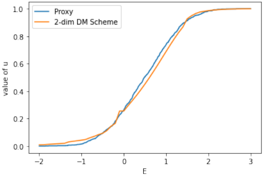
- Results for Delarue-Menozzi scheme:



(g) $\sigma = 0.01$

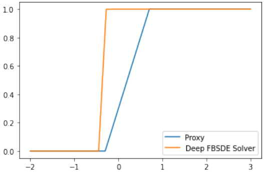


(h) $\sigma = 0.3$

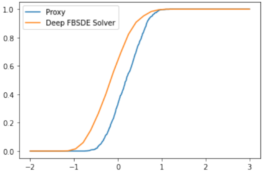


(i) $\sigma = 1.0$

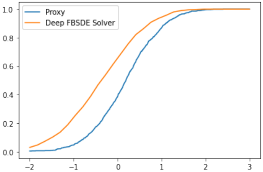
- Results for the deep FBSDE solver (learning error is small):



(j) $\sigma = 0.01$



(k) $\sigma = 0.3$



(l) $\sigma = 1.0$

A splitting scheme

- The numerical methods above fail to capture the correct weak solution.
- This comes from the degeneracy in e and the irregularity of the final condition. Many PDE methods would work, however the dimension of P is too 'big' in applications.

A splitting scheme

- The numerical methods above fail to capture the correct weak solution.
- This comes from the degeneracy in e and the irregularity of the final condition. Many PDE methods would work, however the dimension of P is too 'big' in applications.
- We use a splitting scheme to treat both problem: on a time grid $\pi = (t_n)_{0 \le n \le N}$ we iterate a *transport operator* (fixing p) and a *diffusion operator* (fixing e)

A splitting scheme

- The numerical methods above fail to capture the correct weak solution.
- This comes from the degeneracy in e and the irregularity of the final condition. Many PDE methods would work, however the dimension of P is too 'big' in applications.
- We use a splitting scheme to treat both problem: on a time grid $\pi = (t_n)_{0 \le n \le N}$ we iterate a *transport operator* (fixing p) and a *diffusion operator* (fixing e)
- The transport part is implemented using methods designed for discontinuous solution.

A splitting scheme

- The numerical methods above fail to capture the correct weak solution.
- This comes from the degeneracy in e and the irregularity of the final condition. Many PDE methods would work, however the dimension of P is too 'big' in applications.
- We use a splitting scheme to treat both problem: on a time grid $\pi = (t_n)_{0 \le n \le N}$ we iterate a *transport operator* (fixing p) and a *diffusion operator* (fixing e)
- The transport part is implemented using methods designed for discontinuous solution.
- Results:
  1. we prove the convergence of the splitting scheme with rate $\frac{1}{2}$, in the setting of existence and uniqueness for singular FBSDEs.
  2. we test the splitting scheme using various approximations of the transport operator and the diffusion part (regression).

Theoretical splitting

- Recall the pde "satisfied" by the decoupling field ($r = 0$):

$$\partial_t v + \mu(p, v)\partial_e v + \mathcal{L}_p v = 0, \quad \text{and } v(T, e, p) = \phi(e)$$

and we know $v \in \mathcal{K} := \{(p, e) \mapsto \psi(p, e) : |\partial_p \psi| \leq L, \partial_e \psi \geq 0\}$

Theoretical splitting

- Recall the pde "satisfied" by the decoupling field ($r = 0$):

$$\partial_t v + \mu(p, v)\partial_e v + \mathcal{L}_p v = 0, \quad \text{and } v(T, e, p) = \phi(e)$$

  and we know $v \in \mathcal{K} := \{(p, e) \mapsto \psi(p, e) : |\partial_p \psi| \leq L, \partial_e \psi \geq 0\}$
- We define operators from $(0, \infty) \times \mathcal{K} \ni (h, \psi) \mapsto \mathrm{op}_h \in \mathcal{K}$.

Theoretical splitting

- Recall the pde "satisfied" by the decoupling field ($r = 0$):

$$\partial_t v + \mu(p, v)\partial_e v + \mathcal{L}_p v = 0, \quad \text{and } v(T, e, p) = \phi(e)$$

and we know $v \in \mathcal{K} := \{(p, e) \mapsto \psi(p, e) : |\partial_p \psi| \leq L, \partial_e \psi \geq 0\}$

- We define operators from $(0, \infty) \times \mathcal{K} \ni (h, \psi) \mapsto \mathrm{op}_h \in \mathcal{K}$.
  - Transport step: $\mathcal{T}_h(\psi) = \tilde{v}(0, \cdot)$ with $\tilde{v}$ solution to

$$\partial_t w + \mu(w, p)\partial_e w = 0 \quad \forall p \in \mathbb{R}^d$$

$\hookrightarrow$ better to consider the scalar conservation law to avoid issues at T

Theoretical splitting

- Recall the pde "satisfied" by the decoupling field ($r = 0$):

$$\partial_t v + \mu(p, v)\partial_e v + \mathcal{L}_p v = 0, \quad \text{and } v(T, e, p) = \phi(e)$$

  and we know $v \in \mathcal{K} := \{(p, e) \mapsto \psi(p, e) : |\partial_p \psi| \le L, \partial_e \psi \ge 0\}$

- We define operators from $(0, \infty) \times \mathcal{K} \ni (h, \psi) \mapsto \mathrm{op}_h \in \mathcal{K}$.
  - Transport step: $\mathcal{T}_h(\psi) = \tilde{v}(0, \cdot)$ with $\tilde{v}$ solution to

$$\partial_t w + \mu(w, p)\partial_e w = 0 \quad \forall p \in \mathbb{R}^d$$

    $\hookrightarrow$ better to consider the scalar conservation law to avoid issues at T
  - Diffusion step: $\mathcal{D}_h(\psi) = \bar{v}(0, \cdot)$ with $\bar{v}(t, p, e) = \mathbb{E}\big[\psi(P_h^{t,p}, e)\big]$

Theoretical splitting

- Recall the pde "satisfied" by the decoupling field ($r = 0$):

$$\partial_t v + \mu(p,v)\partial_e v + \mathcal{L}_p v = 0, \quad \text{and } v(T,e,p) = \phi(e)$$

  and we know $v \in \mathcal{K} := \{(p,e) \mapsto \psi(p,e) : |\partial_p \psi| \leq L, \partial_e \psi \geq 0\}$

- We define operators from $(0,\infty) \times \mathcal{K} \ni (h,\psi) \mapsto \mathrm{op}_h \in \mathcal{K}$.
  - Transport step: $\mathcal{T}_h(\psi) = \tilde{v}(0,\cdot)$ with $\tilde{v}$ solution to

$$\partial_t w + \mu(w,p)\partial_e w = 0 \quad \forall p \in \mathbb{R}^d$$

    $\hookrightarrow$ better to consider the scalar conservation law to avoid issues at T
  - Diffusion step: $\mathcal{D}_h(\psi) = \bar{v}(0,\cdot)$ with $\bar{v}(t,p,e) = \mathbb{E}\big[\psi(P_h^{t,p},e)\big]$
  - Then one sets: $\mathcal{S}_h = \mathcal{T}_h \circ \mathcal{D}_h$

Theoretical splitting

- Recall the pde "satisfied" by the decoupling field ($r = 0$):
$$\partial_t v + \mu(p, v)\partial_e v + \mathcal{L}_p v = 0, \quad \text{and } v(T, e, p) = \phi(e)$$
and we know $v \in \mathcal{K} := \{(p, e) \mapsto \psi(p, e) : |\partial_p \psi| \leq L, \partial_e \psi \geq 0\}$
- We define operators from $(0, \infty) \times \mathcal{K} \ni (h, \psi) \mapsto \mathrm{op}_h \in \mathcal{K}$.
  - Transport step: $\mathcal{T}_h(\psi) = \tilde{v}(0, \cdot)$ with $\tilde{v}$ solution to
$$\partial_t w + \mu(w, p)\partial_e w = 0 \quad \forall p \in \mathbb{R}^d$$
$\hookrightarrow$ better to consider the scalar conservation law to avoid issues at T
  - Diffusion step: $\mathcal{D}_h(\psi) = \bar{v}(0, \cdot)$ with $\bar{v}(t, p, e) = \mathbb{E}\big[\psi(P_h^{t,p}, e)\big]$
  - Then one sets: $\mathcal{S}_h = \mathcal{T}_h \circ \mathcal{D}_h$
- Splitting scheme (u_n^π), solution to the backward induction on $\pi = (t_n)_{0 \leq n \leq N}$:
  - for $n = N$, set $u_N^\pi := \phi$,
  - for $n < N$, $u_n^\pi = \mathcal{S}_{t_{n+1} - t_n}(u_{n+1}^\pi)$.

Theoretical splitting

- Recall the pde "satisfied" by the decoupling field ($r = 0$):

$$\partial_t v + \mu(p, v)\partial_e v + \mathcal{L}_p v = 0, \quad \text{and } v(T, e, p) = \phi(e)$$

  and we know $v \in \mathcal{K} := \{(p, e) \mapsto \psi(p, e) : |\partial_p \psi| \le L, \partial_e \psi \ge 0\}$

- We define operators from $(0, \infty) \times \mathcal{K} \ni (h, \psi) \mapsto \mathrm{op}_h \in \mathcal{K}$.
  - Transport step: $\mathcal{T}_h(\psi) = \tilde{v}(0, \cdot)$ with $\tilde{v}$ solution to

$$\partial_t w + \mu(w, p)\partial_e w = 0 \quad \forall p \in \mathbb{R}^d$$

    $\hookrightarrow$ better to consider the scalar conservation law to avoid issues at T
  - Diffusion step: $\mathcal{D}_h(\psi) = \bar{v}(0, \cdot)$ with $\bar{v}(t, p, e) = \mathbb{E}\big[\psi(P_h^{t,p}, e)\big]$
  - Then one sets: $\mathcal{S}_h = \mathcal{T}_h \circ \mathcal{D}_h$

- Splitting scheme (u_n^π), solution to the backward induction on $\pi = (t_n)_{0 \le n \le N}$:
  - for $n = N$, set $u_N^\pi := \phi$,
  - for $n < N$, $u_n^\pi = \mathcal{S}_{t_{n+1} - t_n}(u_{n+1}^\pi)$.

- We obtain, under minimal Lipschitz assumption + structure conditions

$$\int_{\mathbb{R}} |v(0, p, e) - u_0^\pi(p, e)| \mathrm{d}e \le C(1 + |p|^2) N^{-\frac{1}{2}},$$

Convergence 'proof'

Classically, we study stability of the scheme and truncation error.

- Stability for the error $\mathbb{E}\big[\int_{\mathbb{R}} |v(t_n, P_{t_n,e}) - u_n^{\pi}(P_{t_n}, e)|\mathrm{d}e\big]$, ok because:

$$|\mathcal{T}_h(\psi) - \mathcal{T}_h(\psi')|_{L_1} \le |\psi - \psi'|_{L_1} \;\text{ and }\; |\mathcal{D}_h(\psi) - \mathcal{D}_h(\psi')|_{L_\infty} \le |\psi - \psi'|_{L_\infty}$$

Convergence 'proof'

Classically, we study stability of the scheme and truncation error.

- Stability for the error $\mathbb{E}\big[\int_{\mathbb{R}} |v(t_n, P_{t_n,e}) - u^{\pi}_n(P_{t_n}, e)|\mathrm{d}e\big]$, ok because:

$$|\mathcal{T}_h(\psi) - \mathcal{T}_h(\psi')|_{L_1} \le |\psi - \psi'|_{L_1} \text{ and } |\mathcal{D}_h(\psi) - \mathcal{D}_h(\psi')|_{L_\infty} \le |\psi - \psi'|_{L_\infty}$$

- Truncation error: compare $v(,\cdot)$ and one step of the scheme $\tilde{v}(0,\cdot)$

Convergence ‘proof’

Classically, we study stability of the scheme and truncation error.

- Stability for the error $\mathbb{E}\big[\int_{\mathbb{R}} |v(t_n, P_{t_n,e}) - u_n^\pi(P_{t_n}, e)|\mathrm{d}e\big]$, ok because:

$$|\mathcal{T}_h(\psi) - \mathcal{T}_h(\psi')|_{L_1} \le |\psi - \psi'|_{L_1} \text{ and } |\mathcal{D}_h(\psi) - \mathcal{D}_h(\psi')|_{L_\infty} \le |\psi - \psi'|_{L_\infty}$$

- Truncation error: compare $v(,\cdot)$ and one step of the scheme $\tilde{v}(0,\cdot)$
  - on $[0,h]$, $\tilde{v}(0,\cdot)$ given by $\bar{v}(t,p,e) = \mathbb{E}\big[\phi(P_h^{t,p}, e)\big]$ and $\tilde{v}(0,\cdot) = \mathcal{T}_h(\bar{v}(0,\cdot))$,

Convergence ‘proof’

Classically, we study stability of the scheme and truncation error.

- Stability for the error $\mathbb{E}\big[\int_{\mathbb{R}} |v(t_n, P_{t_n,e}) - u_n^\pi(P_{t_n}, e)|\mathrm{d}e\big]$, ok because:

$$|\mathcal{T}_h(\psi) - \mathcal{T}_h(\psi')|_{L_1} \le |\psi - \psi'|_{L_1} \;\text{ and }\; |\mathcal{D}_h(\psi) - \mathcal{D}_h(\psi')|_{L_\infty} \le |\psi - \psi'|_{L_\infty}$$

- Truncation error: compare $v(,\cdot)$ and one step of the scheme $\tilde{v}(0,\cdot)$
  - on $[0,h]$, $\tilde{v}(0,\cdot)$ given by $\bar{v}(t,p,e) = \mathbb{E}\big[\phi(P_h^{t,p}, e)\big]$ and $\tilde{v}(0,\cdot) = \mathcal{T}_h(\bar{v}(0,\cdot))$,
  - the associated FBSDE (the characteristics) is well defined: $\tilde{Y}_t := \tilde{v}(t,p,\tilde{E}_t)$

$$\mathrm{d}\tilde{E}_t = \mu(\tilde{Y}_t, p)\mathrm{d}t \;,\; \mathrm{d}\tilde{Y}_t = 0 \;,\; \tilde{Y}_h = \bar{v}(t,p,\tilde{E}_t)$$

Convergence 'proof'

Classically, we study stability of the scheme and truncation error.

- Stability for the error $\mathbb{E}\left[\int_{\mathbb{R}} |v(t_n, P_{t_n,e}) - u_n^{\pi}(P_{t_n}, e)| \mathrm{d}e\right]$, ok because:

$$|\mathcal{T}_h(\psi) - \mathcal{T}_h(\psi')|_{L_1} \le |\psi - \psi'|_{L_1} \text{ and } |\mathcal{D}_h(\psi) - \mathcal{D}_h(\psi')|_{L_\infty} \le |\psi - \psi'|_{L_\infty}$$

- Truncation error: compare $v(,\cdot)$ and one step of the scheme $\tilde{v}(0,\cdot)$
  - on $[0,h]$, $\tilde{v}(0,\cdot)$ given by $\bar{v}(t,p,e) = \mathbb{E}\left[\phi(P_h^{t,p}, e)\right]$ and $\tilde{v}(0,\cdot) = \mathcal{T}_h(\bar{v}(0,\cdot))$,
  - the associated FBSDE (the characteristics) is well defined: $\tilde{Y}_t := \tilde{v}(t,p,\tilde{E}_t)$

$$\mathrm{d}\tilde{E}_t = \mu(\tilde{Y}_t, p)\mathrm{d}t \ , \ \mathrm{d}\tilde{Y}_t = 0 \ , \ \tilde{Y}_h = \bar{v}(t,p,\tilde{E}_t)$$

  - Expand $V_t = \tilde{v}(t,p,\tilde{E}_t) - v(t, P_t, \tilde{E}_t)$, to get $|V_t|_\infty \le C\sqrt{h}$.

Convergence ‘proof’

Classically, we study stability of the scheme and truncation error.

- Stability for the error $\mathbb{E}\left[\int_{\mathbb{R}} |v(t_n, P_{t_n,e}) - u_n^\pi(P_{t_n}, e)| \mathrm{d}e\right]$, ok because:

$$|\mathcal{T}_h(\psi) - \mathcal{T}_h(\psi')|_{L_1} \le |\psi - \psi'|_{L_1} \text{ and } |\mathcal{D}_h(\psi) - \mathcal{D}_h(\psi')|_{L_\infty} \le |\psi - \psi'|_{L_\infty}$$

- Truncation error: compare $v(,\cdot)$ and one step of the scheme $\tilde{v}(0,\cdot)$
  - on $[0,h]$, $\tilde{v}(0,\cdot)$ given by $\bar{v}(t,p,e) = \mathbb{E}\left[\phi(P_h^{t,p}, e)\right]$ and $\tilde{v}(0,\cdot) = \mathcal{T}_h(\bar{v}(0,\cdot))$,
  - the associated FBSDE (the characteristics) is well defined: $\tilde{Y}_t := \tilde{v}(t,p,\tilde{E}_t)$

$$\mathrm{d}\tilde{E}_t = \mu(\tilde{Y}_t, p)\mathrm{d}t \ , \ \mathrm{d}\tilde{Y}_t = 0 \ , \ \tilde{Y}_h = \bar{v}(t,p,\tilde{E}_t)$$

  - Expand $V_t = \tilde{v}(t,p,\tilde{E}_t) - v(t,P_t,\tilde{E}_t)$, to get $|V_t|_\infty \le C\sqrt{h}$.
  - We want to control $\int |V_0| \mathrm{d}e$: study $t \mapsto \int |V_t \partial_e \tilde{E}_t| \mathrm{d}e$.
    $\hookrightarrow$ we get terms like

$$\int \partial_y \mu(\tilde{v}(t,p,\tilde{E}_t), p) \partial_e \tilde{v}() \partial_e \tilde{E} \mathrm{d}e = \int \partial_e M(\tilde{v}(t,p,\tilde{E}_t), p)) \mathrm{d}e$$

that we can control...

Outline

Backward scheme

- For the E-direction:
  - J be a positive integer and $\mathfrak{E} = (e_j)_{1 \le j \le J}$ a discrete grid of $\mathbb{R}$.
  - $\mathcal{T}_h^{\mathfrak{E}}$ an approximation of the transport operator on $\mathfrak{E}$:

$$\mathbb{R}^d \times \mathbb{R}^J \ni (p, \theta) \mapsto \mathcal{T}_h^{\mathfrak{E}}(p, \theta) \in \mathbb{R}^J$$

Backward scheme

- For the E-direction:
  - J be a positive integer and $\mathfrak{E} = (e_j)_{1 \le j \le J}$ a discrete grid of $\mathbb{R}$.
  - $\mathcal{T}_h^{\mathfrak{E}}$ an approximation of the transport operator on $\mathfrak{E}$:

$$\mathbb{R}^d \times \mathbb{R}^J \ni (p, \theta) \mapsto \mathcal{T}_h^{\mathfrak{E}}(p, \theta) \in \mathbb{R}^J$$

- Euler scheme associated to P on π, namely, for $n \ge 0$,

$$\widehat{P}^\pi_{t_{n+1}} = \widehat{P}^\pi_{t_n} + b(\widehat{P}^\pi_{t_n})h + \sigma(\widehat{P}^\pi_{t_n})\Delta \widehat{W}_n \text{ and } \widehat{P}^\pi_0 = p\,.$$

Backward scheme

- For the E-direction:
  - J be a positive integer and $\mathfrak{E} = (e_j)_{1\le j\le J}$ a discrete grid of $\mathbb{R}$.
  - $\mathcal{T}_h^{\mathfrak{E}}$ an approximation of the transport operator on $\mathfrak{E}$:

$$\mathbb{R}^d \times \mathbb{R}^J \ni (p,\theta) \mapsto \mathcal{T}_h^{\mathfrak{E}}(p,\theta) \in \mathbb{R}^J$$

- Euler scheme associated to P on π, namely, for $n \ge 0$,

$$\widehat{P}^{\pi}_{t_{n+1}} = \widehat{P}^{\pi}_{t_n} + b(\widehat{P}^{\pi}_{t_n})h + \sigma(\widehat{P}^{\pi}_{t_n})\Delta\widehat{W}_n \;\text{ and }\; \widehat{P}^{\pi}_0 = p\,.$$

- Scheme:
  1. For $n = N$, $\Gamma^j_N = \phi(\widehat{P}^{\pi}_{t_N}, e_j)$ for $1 \le j \le J$.
  2. Then, compute for $n < N$

$$\tilde{\Gamma}^j_n = \mathbb{E}\Big[\Gamma^j_{n+1} | \widehat{P}^{\pi}_{t_n}\Big] \;\text{ for all } 1 \le j \le J,$$
$$\Gamma_n = \mathcal{T}_h^{\mathfrak{E}}(\widehat{P}^{\pi}_{t_n}, \tilde{\Gamma}_n)\,.$$

Implementation

- The transport operator is implemented using finite difference schemes: Upwind scheme or Lax-Friedrichs scheme, with J steps in space.

Implementation

- The transport operator is implemented using finite difference schemes: Upwind scheme or Lax-Friedrichs scheme, with J steps in space.
- The regression to estimate functions from $\mathbb{R}^d \to \mathbb{R}^J$ is computed using NN. (simple version of HPW scheme)

Implementation

- The transport operator is implemented using finite difference schemes: Upwind scheme or Lax-Friedrichs scheme, with J steps in space.
- The regression to estimate functions from $\mathbb{R}^d \to \mathbb{R}^J$ is computed using NN. (simple version of HPW scheme)
- We develop also an alternative scheme: the regression is computed on a tree and transport operator approximated by a particles system. This works well for $d \leq 4$ and $P_t := f(t, W_t)$. (Convergence with a rate can be proven).

Implementation

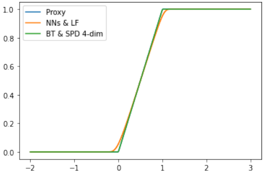
- The transport operator is implemented using finite difference schemes: Upwind scheme or Lax-Friedrichs scheme, with J steps in space.
- The regression to estimate functions from $\mathbb{R}^d \to \mathbb{R}^J$ is computed using NN. (simple version of HPW scheme)
- We develop also an alternative scheme: the regression is computed on a tree and transport operator approximated by a particles system. This works well for $d \leq 4$ and $P_t := f(t, W_t)$. (Convergence with a rate can be proven).
- We test also a multiplicative model:

$$\mathrm{d}P^\ell_t = \mu P^\ell_t \mathrm{d}t + \sigma P^\ell_t \mathrm{d}W^\ell_t,\ P^\ell_0 = 1,\ \text{and } \mathrm{d}E_t = \tilde{\mu}(Y_t, P_t)\mathrm{d}t$$

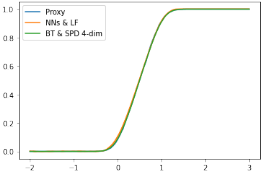
with $\tilde{\mu}(y,p) = \left(\prod_{\ell=1}^d p^\ell\right)^{\frac{1}{\sqrt{d}}} e^{-\theta y}$, for some $\theta > 0$ and $\phi(p,e) = \mathbf{1}_{\{e \geq 0\}}$.

$\hookrightarrow$ *it can be reduced to a 2-dimensional model*!

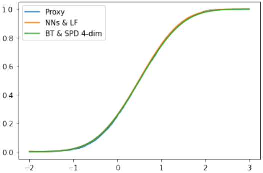
Some numerics on the Toy model



(m) $\sigma = 0.01$



(n) $\sigma = 0.3$

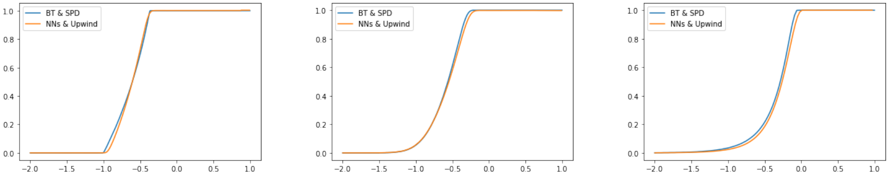


(o) $\sigma = 1.0$

Figure: Linear Toy Model: Comparison of the three methods:
- Neural Nets & Lax-Friedrichs (NN&LF) with $d = 10$
- an alternative scheme (BT&SPD) with $d = 4$
- The Proxy solution given by particle method.

Lax-Friedrichs scheme implemented with discretization of space $J = 1500, 1000, 500$, for $\sigma = 0.01, 0.3, 1$ respectively and number of time step $K = 30$. The number of time step for the splitting is $N = 64$. For $BT\&SPD$, the number of particles is $M = 3500$ and the number of time steps $N = 20$.

On the multiplicative model



(a) $\sigma = 0.01$ (b) $\sigma = 0.3$ (c) $\sigma = 1.0$

Figure: A multiplicative model in dimension $d = 10$. Comparison of two methods:
- Neural nets & Upwind scheme
- the alternative scheme on equivalent 4-dimensional model (BT&SPD).

The Upwind scheme used discretization of space $J = 100, 400, 500$ respectively for $\sigma = 1, 0.3, 0.01$ and number of time step $K = 20$. The number of time step for the splitting is $N = 32$. For $BT\&SPD$, the number of particles is $M = 3500$, and the number of time steps $N = 20$.