Large population stochastic control: analysis and numerical solution to the master equation

J-F Chassagneux (Université Paris Diderot, LPMA)
joint work with D. Crisan (Imperial College London)
and F. Delarue (Université de Nice)

MAD-Stat seminar,
Toulouse, 20 October 2016

Introduction
- Motivation
- Well-posedness of the MKVFBSDE in small time
- MKVFBSDEs and Master Equation

Numerical approximation for small T
- Introduction
- A first scheme
- Convergence results

Numerical approximation for arbitrary T
- The `solver[]()` algorithm
- Review of convergence
- Numerical results

Large population stochastic control

- n players: personal state of player i

$$\mathrm{d}X_t^i = b(t, X_t^i, \mu_t^n, \alpha_t^i)\mathrm{d}t + \sigma \mathrm{d}W_t^i$$

  (W^i) indep. Brownian Motion, $\mu_t^n = \frac{1}{n}\sum_i \delta_{X_t^i}$, α^i control of player
- Cost to minimise for player i:

$$J^i(\alpha) = \mathbb{E}\Big[g(X_T^i, \mu_T^n) + \int_0^T f(t, X_t^i, \mu_t^n, \alpha_t^i)\mathrm{d}t\Big]$$

- Asymptotic description of equilibrium, hopefully "easier" to handle.
- Simplification: at the optimum symmetric feedback control i.e. $\alpha^i = \phi(t, X_t^i)$.

Example - Mean Field Games

Lasry-Lions (06) / Huang-Caines-Malhamé (06)

- "Individual" strategies, looking for Nash-equilibrium $\bar{\alpha}$?

$$J^i(\dots,\bar{\alpha}^{i-1},\alpha^i,\bar{\alpha}^{i-1},\dots) \geq J^i(\dots,\bar{\alpha}^{i-1},\bar{\alpha}^i,\bar{\alpha}^{i-1},\dots)$$

- Optimisation problem for a player: given a flow of measure $(\mu_t)_{t\in[0,T]}$

$$\bar{\phi} = \mathrm{argmin}_{\phi}\mathbb{E}\Big[g(X_T^{\mu},\mu_t) + \int_0^T f(t,X_t^{\mu},\mu_t,\phi(t,X_t^{\mu}))\mathrm{d}t\Big]$$

  with $\mathrm{d}X_t^{\mu} = b(t,X_t^{\mu},\mu_t,\phi(t,X_t))\mathrm{d}t + \sigma\mathrm{d}W_t$.
- Asymptotic $n\to\infty$ yields $\bar{\mu}_t = \mathcal{L}(X_t^{\bar{\mu}})$ (matching problem)
- Conclusion: MFG = optimise first then pass to the limit

Getting the FBSDE

notation: $\mu_t = \mathcal{L}(X_t)$.

- Direct approach: optimum described by $(X_t, Y_t, Z_t)_{t \le T}$:

$$X_t = X_0 + \int_0^t b(s, X_s, \mu_s, \bar{\phi}(s, X_s, Z_s, \mu_s))\mathrm{d}s + \sigma W_t,$$

$$Y_t = g(X_T, \mu_T) + \int_t^T f(s, X_s, \mu_s, \bar{\phi}(s, X_s, Z_s, \mu_s))\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s$$

  (PDE: Lasry-Lions)

- Variational approach (Stochastic Pontryagin Principle)

$$X_t = X_0 + \int_0^t b(s, X_s, \mu_s, \bar{\phi}(s, X_s, Y_s, \mu_s))\mathrm{d}s + \sigma W_t$$

$$Y_t = \partial_x g(X_T, \mu_T) + \int_t^T \partial_x H(s, X_s, Y_s, \mu_s, \bar{\phi}(s, X_s, Y_s, \mu_s))\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s$$

  where $H(\cdot) = b(\cdot)y + f(\cdot)$ and $\bar{\phi}(\cdot) = \mathrm{argmin}_\phi H(\cdot, \phi)$

Example - Control of MKV

- "Cooperative" equilibrium, when the strategy of one player changes, the strategy of all the player changes
  $\hookrightarrow$ Impact the statistical distribution of the system μ^n
- Asymptotic $n \to \infty$ "yields"

$$\mathrm{d}X_t = b(t, X_t, \mathcal{L}(X_t), \alpha_t)\mathrm{d}t + \sigma \mathrm{d}W_t$$

  and for the cost

$$J(\alpha) = \mathbb{E}\Big[g(X_T, \mathcal{L}(X_T)) + \int_0^T f(t, X_t, \mathcal{L}(X_t), \alpha_t)\mathrm{d}t\Big]$$

- then optimise $J(\alpha)$
- conclusion: control of MKV $=$ pass to the limit then optimise
- Coupled FBSDE arises when using stoch. max. principle (Carmona-Delarue) or DPP (Pham)

Contraction approach

- Let us consider

$$\left\{\begin{array}{lcl} \mathrm{d}X_t & = & b(Y_t)\mathrm{d}t + \sigma \mathrm{d}W_t\,,\ X_0 = \xi \\ \mathrm{d}Y_t & = & Z_t\mathrm{d}W_t\,,\ Y_T = g(X_T, \mathcal{L}(X_T)) \end{array}\right.$$

- in a Lipschitz setting

$$|b(y) - b(y')| \leq K|y - y'|\,,$$
$$|g(x,\mu) - g(x',\mu')| \leq K\left(|x - x'| + W_2(\mu,\mu')\right)\,,$$

where $W_2(\mu,\mu') = \inf_{X\sim\mu, X'\sim\mu'} \mathbb{E}\left[|X - X'|^2\right]^{\frac{1}{2}}$.

- For $T \leq c(K)$, existence and uniqueness (via contraction).

Decoupling field

- Non MKV case:

$$\begin{cases} \mathrm{d}X_t &= b(Y_t)\mathrm{d}t + \sigma \mathrm{d}W_t \,,\ X_0 = \xi \\ \mathrm{d}Y_t &= Z_t \mathrm{d}W_t \,,\ Y_T = g(X_T) \end{cases}$$

One can show $Y_t = U(t, X_t)$.

- PDE for U?

On one hand

$$\mathrm{d}U(t, X_t) = \left(\partial_t U + b(Y)\partial_x U + \frac{1}{2}\sigma^2 \partial^2_{xx} U\right)\mathrm{d}t + \mathrm{d}(\mathrm{mart})$$

Moreover $\mathrm{d}U(t, X_t) = \mathrm{d}Y_t = \mathrm{d}(\mathrm{mart})$ and so

$$\partial_t U(t,x) + b(U(t,x))\partial_x U(t,x) + \frac{1}{2}\sigma^2 \partial^2_{xx} U(t,x) = 0 \,.$$

Decoupling field in the MKV case

- For e.g.

$$\begin{cases} \mathrm{d}X_t &= b(Y_t, \mathcal{L}(X_t))\mathrm{d}t + \sigma \mathrm{d}W_t\,,\ X_0 = \xi \\ \mathrm{d}Y_t &= Z_t \mathrm{d}W_t\,,\ Y_T = g(X_T) \end{cases}$$

  One has: $Y_t = U(t, X_t, \mathcal{L}(X_t))$ and U is defined on $[0, T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R})$.

- U satisfies a PDE ?
  $\hookrightarrow$ Need a chain rule to expand U in the measure argument
  $\hookrightarrow$ Need some smoothness also...

Differential Calculus on $\mathcal{P}_2(\mathbb{R})$

- Lions' approach:

$$\text{“Lift” to } L^2: \quad U(\mu) \to \mathcal{U}(\xi) := U(\mathcal{L}(\xi))\,;$$

- U differentiable at μ if $\mathcal{U}$ Frechet differentiable at ξ.
- Moreover, if $\mathcal{U}$ is $\mathcal{C}^1$ then

$$D\mathcal{U}(\xi)\cdot\chi = \mathbb{E}[\partial_\mu U(\mu)(\xi)\chi]\ .$$

$\hookrightarrow \partial_\mu U(\mu)(\cdot) \in L^2(\mathbb{R},\mu)$ derivative of U at μ.

- Example: $U(\mu) = \int \phi(x)\mathrm{d}\mu(x)$

$$\partial_\mu U(\mu)(\upsilon) = \phi'(\upsilon)$$

- Order 2:

$$\partial^2_\mu U(\mu)(\upsilon,\upsilon') \quad \text{and} \quad \partial_\upsilon\partial_\mu U(\mu)(\upsilon)$$

Finite dimensional projection

$u(x) = u(x_1, \dots, x_n) := U(\mu^n_x)$ with $\mu^n_x = \frac{1}{n}\sum_i \delta_{x_i}$.

- First order derivative

$$\partial_{x_i} u(x) = \frac{1}{n}\partial_\mu U(\mu^n_x)(x_i)$$

Proof. ϑ unif. distributed in $\{1, \dots, n\}$, $h = (h_i)$ small perturbation:

$$\begin{aligned} u(x+h) = U(\mathcal{L}(x_\vartheta + h_\vartheta)) &= U(\mathcal{L}(x_\vartheta)) + \mathbb{E}[\partial_\mu U(\mathcal{L}(x_\vartheta))(x_\vartheta)h_\vartheta] + o(|h|)\ , \\ &= U(\mathcal{L}(x_\vartheta)) + \sum_i \frac{1}{n}\partial_\mu U(\mu^n_x)(x_i)h_i + o(|h|). \end{aligned}$$

- second order derivative

$$\partial^2_{x_i x_j} u(x) = \frac{1}{n}\partial_\upsilon \partial_\mu U(\mu^n_x)(x_i)\mathbf{1}_{i=j} + \frac{1}{n^2}\partial^2_\mu U(\mu^n_x)(x_i, x_j)$$

Chain Rule

For a flow a measure $(\mu_t)_{t\in[0,T]}$ where $\mu_t = \mathcal{L}(X_t)$:

$$\mathrm{d}X_t = b_t \mathrm{d}t + \sigma_t \mathrm{d}W_t\,.$$

- The chain rule

$$U(\mu_T) = U(\mu_0) + \int_0^T \mathbb{E}\Big[b_t \partial_\mu U(\mu_t)(X_t) + \frac{1}{2}\partial_\upsilon \partial_\mu U(\mu_t)(X_t)\sigma_t^2\Big]\,\mathrm{d}t$$

proof.
Particle system: (X^i) i.i.d. copies of X, $\mu^n_X = \frac{1}{n}\sum_i \delta_{X^i_t} \rightarrow_{n\infty} \mu_t$.
Apply Ito's formula to $u(X^1_t, \dots, X^n_t)$ and let n goes to ∞:

$$\mathrm{d}u(X^1_t, \dots, X^n_t) = \frac{1}{n}\sum_i \partial_\mu U(\mu^n_{X_t})(X^i_t) b^i_t \mathrm{d}t + \mathrm{dmart}$$

$$+ \sum_i (\sigma^i_t)^2 \left(\frac{1}{2n}\partial_\upsilon \partial_\mu U(\mu^n_{X_t})(X^i_t) + \frac{1}{2n^2}\partial^2_\mu U(\mu^n_{X_t})(X^i_t, X^i_t)\right)\mathrm{d}t$$

Master equation - PDE for U

Consider

$$\begin{cases} \mathrm{d}X_t = b(Y_t, \mathcal{L}(X_t))\mathrm{d}t + \mathrm{d}W_t \,, \; X_0 = \xi \\ \mathrm{d}Y_t = -f(Z_t)\mathrm{d}t + Z_t \mathrm{d}W_t \,, \; Y_T = g(X_T, \mathcal{L}(X_T)) \end{cases}$$

U s.t. $Y_t = U(t, X_t, \mathcal{L}(X_t))$ satisfies $U(T, x, \mu) = g(x, \mu)$ and

$$\partial_t U(\cdot) + b(U(\cdot), \mu)\partial_x U(\cdot) + \frac{1}{2}\partial^2_{xx} U(\cdot) + f(\partial_x U(\cdot))$$
$$+ \mathbb{E}\Big[b(U(t,\xi,\mu),\mu)\partial_\mu U(t,x,\mu)(\xi) + \frac{1}{2}\partial_\upsilon \partial_\mu U(t,x,\mu)(\xi)\Big] = 0$$

$\hookrightarrow$ We prove existence and uniqueness of a "classical" solution in small time to the above PDE written on $[0, T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R})$.

Arbitrary T - difficulties

Consider the following system of FBSDEs

$$\begin{cases} \mathrm{d}Y_t &= -\mathbb{E}[X_t]\,\mathrm{d}t + Z_t \mathrm{d}W_t \text{ and } Y_T = -X_T\,, \\ \mathrm{d}X_t &= Y_t \mathrm{d}t + \sigma(X_t)\mathrm{d}W_t \text{ and } X_0 = x\,. \end{cases} \tag{1}$$

where $T = \frac{3\pi}{4}$ and σ is a Lipschitz function.

If $x \neq 0$, there is no solution in $\mathcal{S}^2 \times \mathcal{S}^2 \times \mathcal{H}^2$ to the above equation.

proof Note $m_X(t) := \mathbb{E}[X_T]$ and $m_Y(t) := \mathbb{E}[Y_T]$ satisfies

$$\begin{cases} \mathrm{d}m_Y(t) &= -m_X(t)\mathrm{d}t \text{ and } m_Y(T) = -m_X(T)\,, \\ \mathrm{d}m_X(t) &= m_Y(t)\mathrm{d}t \text{ and } m_X(0) = x\,. \end{cases} \tag{2}$$

The above system has no solution for $x \neq 0$. Observe that
$m_X(t) = x\cos(t) + \mu\sin(t)$, $m_Y(t) = -x\sin(t) + \mu\cos(t)$
so that $m_Y(T) + m_X(T) = -x\sqrt{2}$.

Positive results in the "classical" case

$\hookrightarrow$ No MKV interaction

- σ is non degenerate, coefficients are bounded (Delarue)
- Existence and uniqueness also for some singular FBSDEs (Carmona-Delarue).

In any case, need a control on the solution's gradient.

Generic method

- Recursive method by splitting the time interval
- Possible only if control of Lipschitz constant of U, obtained from the estimate

$$\mathbb{E}\big[|U(t,\xi,\mathcal{L}(\xi)) - U(t,\xi',\mathcal{L}(\xi'))|^2\big]^{\frac{1}{2}} \le \Lambda\mathbb{E}\big[|\xi-\xi'|^2\big]^{\frac{1}{2}} \ . \tag{3}$$

- Structural condition on the coefficient allows to obtain previous estimate both in the MFG and control of MKV setting.

Objective and difficulties

- Goal: Numerical Approximation of $U(0,\xi,\mathcal{L}(\xi))$, U decoupling field for

$$\begin{cases} X_t &= \xi + \int_0^t b(Y_r, \mathbb{E}[X_r])\mathrm{d}r + \sigma W_t, \\ Y_t &= g(X_T) + \int_t^T f(Z_r)\mathrm{d}r - \int_t^T Z_r \mathrm{d}W_r, \end{cases}$$

  in particular: $Y_0 = U(0,\xi,\mathcal{L}(\xi))$.

- Method: Adaptating grid method for coupled FBSDE is difficult... $Y_t = U(t, X_t, \mathcal{L}(X_t))$.
  $\hookrightarrow$ back to basics: we use a binomial tree and a Picard iteration scheme (Need T small!)

Dealing with the coupling

- Picard Iteration, $(\tilde{X}^j, \tilde{Y}^j, \tilde{Z}^j)_{0 \le j}$:

$$\left\{ \begin{array}{rcl} \tilde{X}^j_t &=& \xi + \int_0^t b(\tilde{Y}^j_r, \mathbb{E}\Big[\tilde{X}^j_r\Big])\mathrm{d}r + W_t, \\ Y^j_t &=& g(\tilde{X}^{j-1}_T) + \int_t^T f(\tilde{Z}^j_r)\mathrm{d}r - \int_t^T \tilde{Z}^j_r \mathrm{d}W_r, \end{array} \right. \tag{4}$$

  with $\tilde{X}^0 = \xi$ (and $\tilde{Y}^0 = \tilde{Z}^0 = 0$).
- Easily shown: $(\tilde{X}^j, \tilde{Y}^j, \tilde{Z}^j) \to (X, Y, Z)$
- Stopped after J iteration: output is $Y^J_0 \leftrightarrow U(0, \xi, \mathcal{L}(\xi))$
- In practice, one cannot solve perfectly (4)

Discrete approximation

- A discrete time grid $\pi = \{t_0, \ldots, t_n\}$ with mesh size $|\pi| := h$.
- Use a Binomial Tree for Brownian Motion: $\bar{\mathbb{P}}(\Delta W_i = \pm\sqrt{h}) = \frac{1}{2}$.
- "Classical" BTZ scheme:

$$\bar{X}_{t_{i+1}} = \bar{X}_{t_i} + b(\bar{Y}_{t_i}, \bar{\mathbb{E}}[\bar{X}_{t_i}])h + \sigma \Delta \bar{W}_i \,,$$

$$\bar{Y}_{t_i} = \bar{\mathbb{E}}_{t_i}\left[\bar{Y}_{t_{i+1}} + hf(\bar{Z}_{t_i})\right] \text{ with } \bar{Z}_{t_i} = \bar{\mathbb{E}}_{t_i}\left[\frac{\Delta W_i}{h}\bar{Y}_{t_{i+1}}\right]$$

with $\bar{X}_0 = \xi$ and $\bar{Y}_{t_n} = g(\bar{X}_T)$.

Note: For the X-part, classical Explicit Euler scheme...

Deriving the scheme (1/2) - Y part

On the equidistant grid $\pi = \{0 = t_0 < ... < t_i < ... < t_n = T\}$, with $h = T/n$.

- Start with:

$$Y_{t_i} + \int_{t_i}^{t_{i+1}} Z_s \mathrm{d}W_s = Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f(Z_s)\mathrm{d}s \quad (1)$$

Deriving the scheme (1/2) - Y part

On the equidistant grid $\pi = \{0 = t_0 < \ldots < t_i < \ldots < t_n = T\}$, with $h = T/n$.

- Start with:

$$Y_{t_i} + \int_{t_i}^{t_{i+1}} Z_s \mathrm{d}W_s \simeq Y_{t_{i+1}} + hf(Z_{t_i}) \quad (1)$$

Deriving the scheme (1/2) - Y part

On the equidistant grid $\pi = \{0 = t_0 < ... < t_i < ... < t_n = T\}$, with $h = T/n$.

- Start with:

$$Y_{t_i} + \int_{t_i}^{t_{i+1}} Z_s \mathrm{d}W_s \simeq Y_{t_{i+1}} + hf(Z_{t_i}) \quad (1)$$

- *For the Y-part:*

Deriving the scheme (1/2) - Y part

On the equidistant grid $\pi = \{0 = t_0 < ... < t_i < ... < t_n = T\}$, with $h = T/n$.

- Start with:

$$Y_{t_i} + \int_{t_i}^{t_{i+1}} Z_s \mathrm{d}W_s \simeq Y_{t_{i+1}} + hf(Z_{t_i}) \quad (1)$$

- *For the Y-part:*
  Take conditional expectation,

$$Y_{t_i} \simeq \mathbb{E}_{t_i}\big[Y_{t_{i+1}} + hf(Z_{t_i})\big]$$

Deriving the scheme (1/2) - Y part

On the equidistant grid $\pi = \{0 = t_0 < ... < t_i < ... < t_n = T\}$, with $h = T/n$.

- Start with:

$$Y_{t_i} + \int_{t_i}^{t_{i+1}} Z_s \mathrm{d}W_s \simeq Y_{t_{i+1}} + hf(Z_{t_i}) \quad (1)$$

- *For the Y-part:*
  Take conditional expectation,

$$Y_{t_i} \simeq \mathbb{E}_{t_i}\big[Y_{t_{i+1}} + hf(Z_{t_i})\big]$$

$$\hookrightarrow \quad \bar{Y}_{t_i} := \bar{\mathbb{E}}_{t_i}\left[\bar{Y}_{t_{i+1}} + hf(\bar{Z}_{t_i})\right]$$

Deriving the scheme (2/2) - Z part

- Start with:

$$Y_{t_i} + \int_{t_i}^{t_{i+1}} Z_s \mathrm{d}W_s \simeq Y_{t_{i+1}} + hf(Z_{t_i}) \quad (1)$$

Deriving the scheme (2/2) - Z part

- Start with:

$$Y_{t_i} + \int_{t_i}^{t_{i+1}} Z_s \mathrm{d}W_s \simeq Y_{t_{i+1}} + hf(Z_{t_i}) \quad (1)$$

- *For the Z-part:*

Deriving the scheme (2/2) - Z part

- Start with:

$$Y_{t_i} + \int_{t_i}^{t_{i+1}} Z_s \mathrm{d}W_s \simeq Y_{t_{i+1}} + hf(Z_{t_i}) \quad (1)$$

- *For the Z-part:*
  Multiply (1) by $\Delta W_i := W_{t_{i+1}} - W_{t_i}$, take conditional expectation:

$$\mathbb{E}_{t_i}\left[\int_{t_i}^{t_{i+1}} Z_s \mathrm{d}s\right] \simeq \mathbb{E}_{t_i}\left[\Delta W_i Y_{t_{i+1}}\right]$$

Deriving the scheme (2/2) - Z part

- Start with:

$$Y_{t_i} + \int_{t_i}^{t_{i+1}} Z_s \mathrm{d}W_s \simeq Y_{t_{i+1}} + hf(Z_{t_i}) \quad (1)$$

- *For the Z-part:*
  Multiply (1) by $\Delta W_i := W_{t_{i+1}} - W_{t_i}$, take conditional expectation:

$$hZ_{t_i} \simeq \mathbb{E}_{t_i}\left[\int_{t_i}^{t_{i+1}} Z_s \mathrm{d}s\right] \simeq \mathbb{E}_{t_i}\left[\Delta W_i Y_{t_{i+1}}\right]$$

Deriving the scheme (2/2) - Z part

- Start with:

$$Y_{t_i} + \int_{t_i}^{t_{i+1}} Z_s \mathrm{d}W_s \simeq Y_{t_{i+1}} + hf(Z_{t_i}) \quad (1)$$

- *For the Z-part:*
Multiply (1) by $\Delta W_i := W_{t_{i+1}} - W_{t_i}$, take conditional expectation:

$$hZ_{t_i} \simeq \mathbb{E}_{t_i}\Big[\int_{t_i}^{t_{i+1}} Z_s \mathrm{d}s\Big] \simeq \mathbb{E}_{t_i}\big[\Delta W_i Y_{t_{i+1}}\big]$$

$$\hookrightarrow \quad \bar{Z}_{t_i} := \bar{\mathbb{E}}_{t_i}\left[h^{-1}\Delta W_i \bar{Y}_{t_{i+1}}\right] .$$

Convergence "analysis"

- Errors:
  1. Due to the Picard Iteration: $\leq CT^J$
  2. Due to the discretisation: $\leq C\sqrt{h}$
- To prove
  1. Compare $\tilde{Y}^j_t$ and $U(t, \tilde{X}^j_t, \mathcal{L}(\tilde{X}^j_t))$
     $\hookrightarrow$ use "extended" Ito formula + smoothness.
  2. Compare $\bar{Y}_{t_i}$ and $U(t_i, \bar{X}_{t_i}, \mathcal{L}(\bar{X}_{t_i}))$
     $\hookrightarrow$ use a "discrete" Ito formula.

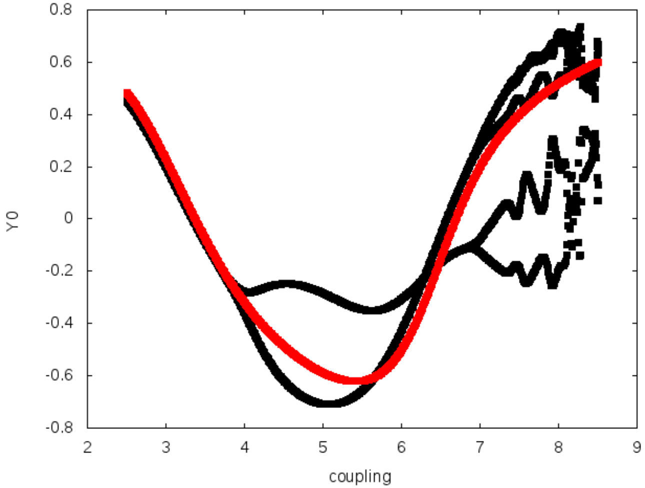
Numerical result: a model with no MKV interaction

- The model:

$$
\begin{aligned}
\mathrm{d}X_t &= \rho \cos(Y_t)\mathrm{d}t + \sigma \mathrm{d}W_t \text{ and } X_0 = x \in \mathbb{R} \,, \\
dY_t &= Z_t \mathrm{d}W_t \text{ and } Y_T = \sin(X_T) \,.
\end{aligned}
$$

- The important parameter is the coupling parameter ρ that will vary in $[2.5, 8.5]$.
- Parameters for the simulation: 25 Picard iterations, 15 time steps, $T = \sigma = 1$

Numerical result: output



Continuation method

Divide $[0, T]$ in small intervals of size $\delta = \frac{T}{N}$.

- **Continuation Method**:

  - We know that $Y_0 = U(0, \xi, \mathcal{L}(\xi))$ with (X, Y, Z) solution to

$$\begin{cases} X_t &= \xi + \int_0^t b(Y_r, \mathbb{E}[X_r])\mathrm{d}r + W_t, \\ Y_t &= U(\delta, X_\delta, \mathcal{L}(X_\delta)) + \int_t^\delta f(Z_r)\mathrm{d}r - \int_t^\delta Z_r \mathrm{d}W_r, \end{cases}$$

Continuation method

Divide $[0, T]$ in small intervals of size $\delta = \frac{T}{N}$.

- **Continuation Method**:
  - We know that $Y_0 = U(0, \xi, \mathcal{L}(\xi))$ with (X, Y, Z) solution to

$$\begin{cases} X_t = \xi + \int_0^t b(Y_r, \mathbb{E}[X_r])\mathrm{d}r + W_t, \\ Y_t = U(\delta, X_\delta, \mathcal{L}(X_\delta)) + \int_t^\delta f(Z_r)\mathrm{d}r - \int_t^\delta Z_r \mathrm{d}W_r, \end{cases}$$

  - which can be approximated by Picard Iteration

$$\begin{cases} X_t^j = \xi + \int_0^t b(Y_r^j, \mathbb{E}[X_r^j])\mathrm{d}r + W_t, \\ Y_t^j = U(\delta, X_\delta^{j-1}, \mathcal{L}(X_\delta^{j-1})) + \int_t^\delta f(Z_r^j)\mathrm{d}r - \int_t^\delta Z_r^j \mathrm{d}W_r, \end{cases}$$

Continuation method

Divide $[0, T]$ in small intervals of size $\delta = \frac{T}{N}$.

- **Continuation Method**:

  \- We know that $Y_0 = U(0, \xi, \mathcal{L}(\xi))$ with (X, Y, Z) solution to

$$\begin{cases} X_t &= \xi + \int_0^t b(Y_r, \mathbb{E}[X_r])\mathrm{d}r + W_t, \\ Y_t &= U(\delta, X_\delta, \mathcal{L}(X_\delta)) + \int_t^\delta f(Z_r)\mathrm{d}r - \int_t^\delta Z_r \mathrm{d}W_r, \end{cases}$$

  \- which can be approximated by Picard Iteration

$$\begin{cases} X_t^j &= \xi + \int_0^t b(Y_r^j, \mathbb{E}[X_r^j])\mathrm{d}r + W_t, \\ Y_t^j &= U(\delta, X_\delta^{j-1}, \mathcal{L}(X_\delta^{j-1})) + \int_t^\delta f(Z_r^j)\mathrm{d}r - \int_t^\delta Z_r^j \mathrm{d}W_r, \end{cases}$$

- Problem: U is required and this is what we want to compute...
  $\hookrightarrow$ We use a recursive algorithm, assuming that

$$U(\delta, \xi, \mathcal{L}(\xi)) \simeq \texttt{solver[1]}(\xi)$$

Recursive Method

For any "level", $0 \leq k < N-1$

- we compute on $[r_k, r_{k+1}]$ with $r_k := k\delta$

$$\begin{cases} X_t^j &= \xi + \int_{r_k}^t b(Y_r^j, \mathbb{E}[X_r^j])\mathrm{d}r + W_t - W_{r_k}, \\ Y_t^j &= \texttt{solver[}k+1\texttt{](}X_{r_{k+1}}^{j-1}\texttt{)} + \int_t^{r_{k+1}} f(Z_r^j)\mathrm{d}r - \int_t^{r_{k+1}} Z_r^j \mathrm{d}W_r, \end{cases}$$

- we stop at Picard Iteration J and set

$$\texttt{solver[}k\texttt{](}\xi\texttt{)} := Y_{r_k}^J \,.$$

Recursive Method

For any "level", $0 \le k < N-1$

- we compute on $[r_k, r_{k+1}]$ with $r_k := k\delta$

$$\begin{cases} X_t^j &= \xi + \int_{r_k}^t b(Y_r^j, \mathbb{E}[X_r^j])\mathrm{d}r + W_t - W_{r_k}, \\ Y_t^j &= \texttt{solver}[k+1](X_{r_{k+1}}^{j-1}) + \int_t^{r_{k+1}} f(Z_r^j)\mathrm{d}r - \int_t^{r_{k+1}} Z_r^j \mathrm{d}W_r, \end{cases}$$

- we stop at Picard Iteration J and set

$$\texttt{solver}[k](\xi) := Y_{r_k}^J .$$

At Level $N-1$, we have

- $\texttt{solver}[N-1](\xi) := Y_{r_{N-1}}^J$ where, for $j \le J$,

$$\begin{cases} X_t^j &= \xi + \int_{r_{N-1}}^t b(Y_r^j, \mathbb{E}[X_r^j])\mathrm{d}r + W_t - W_{r_{N-1}}, \\ Y_t^j &= g(X_T^{j-1}) + \int_t^T f(Z_r^j)\mathrm{d}r - \int_t^T Z_r^j \mathrm{d}W_r, \end{cases}$$

- In particular, $\texttt{solver}[N](\cdot) = g(\cdot)$, No error...

Full algorithm

- One cannot solve the following BSDE perfectly on $[r_k, r_{k+1}]$:

$$\begin{cases} X_t & = & \xi + \int_{r_k}^{t} b(Y_r, \mathbb{E}[X_r])\mathrm{d}r + W_t - W_{r_k}, \\ Y_t & = & \chi + \int_t^{r_{k+1}} f(Z_r)\mathrm{d}r - \int_t^{r_{k+1}} Z_r \mathrm{d}W_r, \end{cases}$$

- the solution is approximated by $(\bar{X}_t, \bar{Y}_t, \bar{Z}_t)_{t\in\pi^k}$ on a subgrid π^k with $|\pi^k| = h$ via a generic solver:

$$(\bar{X}_t, \bar{Y}_t)_{t\in\pi^k} := \overline{\texttt{solver}}[k](\xi, \chi)$$

Full algorithm

- One cannot solve the following BSDE perfectly on $[r_k, r_{k+1}]$:

$$\begin{cases} X_t &= \xi + \int_{r_k}^t b(Y_r, \mathbb{E}[X_r])\mathrm{d}r + W_t - W_{r_k}, \\ Y_t &= \chi + \int_t^{r_{k+1}} f(Z_r)\mathrm{d}r - \int_t^{r_{k+1}} Z_r \mathrm{d}W_r, \end{cases}$$

- the solution is approximated by $(\bar{X}_t, \bar{Y}_t, \bar{Z}_t)_{t\in\pi^k}$ on a subgrid π^k with $|\pi^k| = h$ via a generic solver:

$$(\bar{X}_t, \bar{Y}_t)_{t\in\pi^k} := \overline{\texttt{solver}}\texttt{[}k\texttt{](}\xi\texttt{,}\chi\texttt{)}$$

- A level k, to compute $\texttt{solver[}k\texttt{](}\xi\texttt{)}$:
  1. initialisation at $\bar{X}_t^{0,k} = \xi$ and $\bar{Y}_t^{0,k} = 0$ for $t \in \pi_k$
  2. for $j \leq J$
     - 2.1 compute $\bar{Y}^{j,k}_{r_{k+1}} = \texttt{solver[}k+1\texttt{](}\bar{X}^{j-1,k}_{r_{k+1}}\texttt{)}$.
     - 2.2 compute $(\bar{Y}^{j,k}, \bar{X}^{j,k}) = \overline{\texttt{solver}}\texttt{[}k\texttt{](}\xi\texttt{,}\bar{Y}^{j,k}_{r_{k+1}}\texttt{)}$
  3. return $\bar{Y}^{J,k}_{r_{k+1}}$.

Definition of $\overline{\texttt{solver}}$[](,)

In practice, we use the classical BTZ scheme e.g. for level k:

$$\bar{X}_{t_{i+1}} = \bar{X}_{t_i} + b(\bar{Y}_{t_i}, \bar{\mathbb{E}}[\bar{X}_{t_i}])h + \sigma(\bar{X}_{t_i})\Delta \bar{W}_i \,,$$

$$\bar{Y}_{t_i} = \bar{\mathbb{E}}_{t_i}\left[\bar{Y}_{t_{i+1}} + hf(\bar{Z}_{t_i})\right] \text{ with } \bar{Z}_{t_i} = \bar{\mathbb{E}}_{t_i}\left[\frac{\Delta W_i}{h}\bar{Y}_{t_{i+1}}\right]$$

with $\bar{X}_{r_k} = \xi$ and $\bar{Y}_{r_{k+1}} = \eta$.

Errors and convergence

- At each level, local error comes from
  1. Stopping the Picard Iteration
  2. Discretising the BSDE.
- Global error: Propagation of local error through the levels?
  1. When no error is made on $\overline{\texttt{solver}}$`[](,)`: $\mathrm{err} \leq C\delta^{J-1}$.
  2. When ζ error made: $\mathrm{err} \leq C(\delta^{J-1} + N\zeta)$.
- Result:

$$\mathrm{err} \leq C(\delta^{J-1} + \sqrt{h})$$

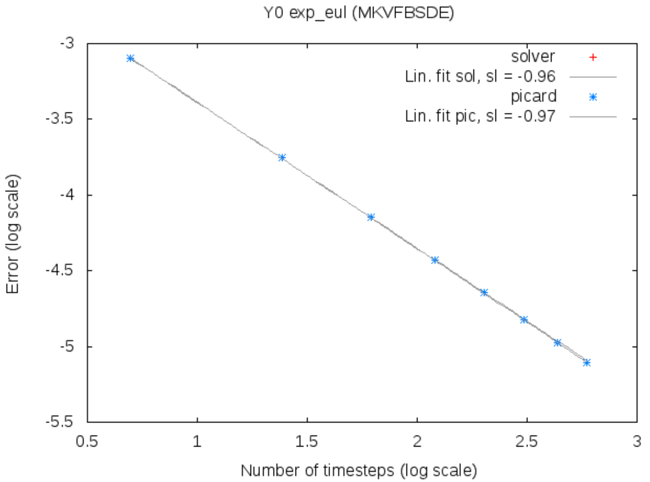
Safety check: A linear model

- The model:

$$
\begin{aligned}
\mathrm{d}X_t &= -\rho\mathbb{E}[Y]_t\,\mathrm{d}t + \sigma\mathrm{d}W_t\ ,\ X_0 = x\,,\\
\mathrm{d}Y_t &= -aY_s\mathrm{d}s + Z_s\mathrm{d}W_s\ \text{ and }\ Y_T = X_T\,.
\end{aligned}
$$

- The coupling parameter is fixed.
- We study the convergence of the discretisation error for both method
  1. Picard Iteration (25 iterations)
  2. `solver[]()` with two levels (5 Picard iterations each)

Numerical result for the linear model



Non-linear example with MKV interaction

- The model

$$
\begin{aligned}
&\mathrm{d}X_t = -\rho Y_t \mathrm{d}t + \mathrm{d}W_t\,,\ X_0 = x\,,\\
&\mathrm{d}Y_t = \mathrm{atan}(\mathbb{E}[X_t])\mathrm{d}t + Z_t \mathrm{d}W_t \text{ and } Y_T = G'(X_T) := \mathrm{atan}(X_T)
\end{aligned}
$$

- coming from Pontryagin principle applied to MFG

$$
\inf_{\alpha} \mathbb{E}\left[G(X_t^\alpha) + \int_0^T \left(\frac{1}{2\rho}\alpha_t^2 + X_t^\alpha \mathrm{atan}(\mathbb{E}[X_t^\alpha]) \right) \mathrm{d}t \right]
$$

with $\mathrm{d}X_t^\alpha = \alpha_t \mathrm{d}t + \mathrm{d}W_t$.

- numerics
  1. Picard Iterations (25) - in *blue*
  2. `solver[](,)` with two levels (5 iterations per level) - in *black*

Output