Deep Learning and Elicitability for McKean–Vlasov FBSDEs with Common Noise

Picard iterations, elicitable statistics, and neural decoupling fields

Felipe J. P. Antunes **Yuri F. Saporito** Sebastian Jaimungal

FGV EMAp
ESCOLA DE MATEMÁTICA APLICADA

Workshop on Stochastic Analysis: Irregularity, Mean-Fields and Related Topics

August 13th, 2026

The target is a fully coupled MV-FBSDE with common noise

$$
\begin{aligned}
\mathrm{d}X_t &= \mu(t, X_t, Y_t, Z_t, Z_t^0, S_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t + \sigma_0(t, X_t)\,\mathrm{d}W_t^0, \qquad X_0 = \xi, \\
\mathrm{d}Y_t &= -f(t, X_t, Y_t, Z_t, Z_t^0, S_t)\,\mathrm{d}t + Z_t^\top\,\mathrm{d}W_t + (Z_t^0)^\top\,\mathrm{d}W_t^0, \qquad Y_T = G(X_T, S_T).
\end{aligned}
$$

- $X_t \in \mathbb{R}^d$, $Y_t \in \mathbb{R}$, $Z_t \in \mathbb{R}^d$, $Z_t^0 \in \mathbb{R}^{d_0}$, and $S_t \in \mathbb{R}^p$;
- $W \in \mathbb{R}^d$ and $W^0 \in \mathbb{R}^{d_0}$ are independent Brownian motions.

Random environment

$$
S_t = \mathcal{T}\Big(\mathcal{L}(X_t \mid \mathcal{F}_t^0)\Big),
$$

where $\mathcal{T}$ may be a conditional mean, moment, quantile, or another elicitable statistic.

Goal: simulate (X, Y, Z, Z^0, S) on the whole horizon without reconstructing the conditional law.

MV-FBSDEs encode equilibria and optimal controls

- The forward state describes a representative agent or controlled system.
- The backward variables encode marginal values, adjoint processes, or equilibrium feedback.
- Mean-field interaction makes the coefficients depend on the population distribution.
- Common shocks turn that distribution into a stochastic process observed by every agent.

Typical applications: mean-field games, mean-field control, systemic risk, economic growth, and interacting financial systems.

Common noise makes the population state random

Idiosyncratic noise W

- moves one representative agent;
- averages out across the population;
- remains relevant for the agent's state X_t.

Common noise W^0

- moves all agents simultaneously;
- survives in the mean-field limit;
- makes $\mathcal{L}(X_t \mid \mathcal{F}_t^0)$ a stochastic measure flow.

The state variable is not only X_t: it includes the random, path-dependent population statistic S_t.

The information structure determines admissible approximations

$$\mathcal{F}_t^0 = \sigma(W_s^0 : s \le t), \quad \mathcal{F}_t = \sigma(W_s : s \le t), \quad \mathcal{G}_t = \mathcal{F}_t \vee \mathcal{F}_t^0 \vee \sigma(\xi).$$

- The statistic network uses only $(t, W^0_{[0,t]})$, so S_t is $\mathcal{F}_t^0$-measurable and cannot see X_t, W, or future increments.
- The decoupling and martingale networks use $(t, X_t, W^0_{[0,t]})$: the current state summarizes the agent-specific information, while a causal recurrent encoder summarizes the common-noise history.
- Sharing the same time-causal inputs across all paths enforces adaptation at every grid point.

Adaptation is enforced by architecture, not added as a penalty.

Coefficients need only a conditional-law statistic

The classical state variable is the random measure

$$\nu_t = \mathcal{L}(X_t \mid \mathcal{F}_t^0).$$

But many models depend on ν_t only through a lower-dimensional statistic:

$$S_t = \mathcal{T}(\nu_t) \in \mathbb{R}^p.$$

Examples

- conditional mean or moments;
- conditional quantiles;
- jointly elicitable tail statistics.

Numerical consequence

- avoid density estimation;
- avoid a high-dimensional measure embedding;
- learn the exact object consumed by μ, f, G.

Two fixed points create the numerical difficulty

1. **Forward–backward coupling.** The forward drift needs (Y, Z, Z^0), while the backward equation needs the future path of X.
2. **Conditional-law coupling.** The coefficients need S_t, but S_t is determined by the law of the state generated by those same coefficients.

Classical route		**Our route**
conditional-law approximation $+$ nested conditional simulation	$\Longrightarrow$	whole-horizon Picard iteration $+$ elicitable pathwise losses

The forward solve has its own inner Picard loop

With $(Y^k, Z^k, Z^{0,k}, S^k)$ frozen, initialize $X^{k+1,0} = X^k$ and iterate

$$\begin{aligned} X_t^{k+1,n+1} = \xi &+ \int_0^t \mu(u, X_u^{k+1,n}, Y_u^k, Z_u^k, Z_u^{0,k}, S_u^k)\,\mathrm{d}u \\ &+ \int_0^t \sigma(u, X_u^{k+1,n})\,\mathrm{d}W_u + \int_0^t \sigma_0(u, X_u^{k+1,n})\,\mathrm{d}W_u^0. \end{aligned}$$

Stop when

$$\|X^{k+1,n+1} - X^{k+1,n}\|_{L^2(\Omega\times[0,T])}$$

falls below tolerance, then set $X^{k+1} = X^{k+1,n+1}$.

Direct conditional simulation scales poorly

For a fixed common-noise path w^0, a direct estimate is

$$\mathbb{E}[X_t \mid W^0_{[0,t]} = w^0] \approx \frac{1}{M} \sum_{m=1}^{M} X_t^{(m)}(w^0).$$

- Each outer common path requires a fresh cloud of idiosyncratic paths.
- The procedure must be repeated across time and Picard iterations.
- General conditional statistics require additional conditional estimation machinery.

The common-noise dimension turns an ordinary Monte Carlo average into a nested simulation problem.

Elicitability turns a conditional statistic into a loss

Strict consistency means equality identifies the target uniquely. Conditioning then gives

$$S_t = \operatorname*{arg\,min}_{s \in L^2(\mathcal{F}_t^0)} \mathbb{E}[\mathfrak{S}(s, X_t)] .$$

- Restricting the minimizer to $\mathcal{F}_t^0$-measurable variables produces the conditional statistic.
- Empirical risk minimization replaces conditional simulation.
- A score $\mathfrak{S}$ is consistent for $\mathcal{T}$ when, for every distribution F,

$$\mathbb{E}_F[\mathfrak{S}(\mathcal{T}(F), X)] \le \mathbb{E}_F[\mathfrak{S}(s, X)] \qquad \text{for every } s.$$

The statistical target is specified by the score; the information set is specified by the network input.

Different interactions correspond to different scores

Statistic	Score or score family	Conditional minimizer
Mean	$(x-s)^2$	$s^* = \mathbb{E}[X_t \mid \mathcal{F}_t^0]$
α**-quantile**	$(\mathbf{1}_{\{x\le s\}} - \alpha)(s-x)$	$s^* = q_\alpha(X_t \mid \mathcal{F}_t^0)$
Quantile and ES	strictly consistent joint scores	$(q_\alpha, \mathrm{ES}_\alpha)$

For lower-tail ES, one Fissler–Ziegel score is

$$\mathfrak{S}(q,e;x) = (\mathbf{1}_{\{x\le q\}} - \alpha)G_1(q) - \mathbf{1}_{\{x\le q\}}G_1(x) + G_2(e)\left(e - q + \frac{\mathbf{1}_{\{x\le q\}}}{\alpha}(q-x)\right) - \mathcal{G}_2(e),$$

where G_1 is increasing and $G_2 = \mathcal{G}_2'$ is positive and strictly increasing.

The solver is not tied to moments: any elicitable statistic can define the random environment.

The common statistic is learned over all paths and times

Parameterize the adapted statistic by a recurrent network

$$S_t \approx S_\theta(t, W^0_{[0,t]}).$$

For samples indexed by m and a grid $\{t_n\}_{n=0}^N$, minimize

$$\mathcal{L}_S(\theta) = \frac{1}{M(N+1)} \sum_{m=1}^{M} \sum_{n=0}^{N} \mathfrak{S}\Big(S_\theta(t_n, W^{0,(m)}_{[0,t_n]}), X^{(m)}_{t_n}\Big).$$

- one training problem shares information across common-noise paths;
- recurrence captures genuinely non-Markovian dependence;
- the same architecture supports different scores.

The backward process admits a path-dependent decoupling field

Under standard Lipschitz assumptions, the usual measure-dependent representation is

$$Y_t = V(t, X_t, \nu_t), \qquad \nu_t = \mathcal{L}(X_t \mid \mathcal{F}_t^0).$$

Because ν_t is determined by the common-noise history, we can reparameterize it as

$$Y_t = U(t, X_t, W^0_{[0,t]}).$$

- The current state X_t carries the Markovian information of the representative agent.
- A causal encoder of $W^0_{[0,t]}$ carries the random environment without explicitly feeding the random measure ν_t into the network.

The decoupling field is a function on state space and common-noise path space.

The backward process is another conditional regression

Recall the integrated BSDE

$$Y_t = G(X_T, S_T) + \int_t^T f_u \, \mathrm{d}u - \int_t^T Z_u^\top \, \mathrm{d}W_u - \int_t^T (Z_u^0)^\top \, \mathrm{d}W_u^0.$$

With the previous Picard iterate frozen in the driver, define

$$H_t = G(X_T, S_T) + \int_t^T f_u \, \mathrm{d}u.$$

Then

$$Y_t = \mathbb{E}[H_t \mid \mathcal{G}_t] = \operatorname*{arg\,min}_{\phi \in L^2(\mathcal{G}_t)} \mathbb{E}\left[|\phi - H_t|^2\right].$$

- Parameterize ϕ by $U_\theta(t, X_t, W^0_{[0,t]})$.
- Train simultaneously over the full time grid.

One regression trains Y across the whole horizon

For weights p_n and $P = \sum_n p_n$, the empirical objective is

$$\mathcal{L}_Y(\theta) = \frac{1}{MP} \sum_{m=1}^{M} \sum_{n=0}^{N} p_n \left| U_\theta(t_n, X_{t_n}^{(m)}, W_{[0,t_n]}^{0,(m)}) - G(X_T^{(m)}, S_T^{(m)}) - \sum_{j=n}^{N-1} f_{t_j}^{(m)} \Delta t \right|^2 .$$

- Typically $p_N = N/2$ and $p_n = 1$ before maturity.
- One terminal value and cumulative driver sums provide pathwise targets for every t_n.
- A single optimization fits all dates at once; there is no separate regression trained sequentially at $t_{N-1}, t_{N-2}, \ldots, t_0$.

Z and Z^0 are learned from local martingale increments

On $[t, t+\Delta t]$,

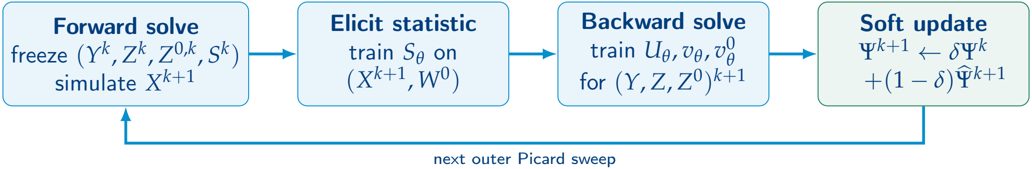
$$\Delta Y_t = -f_t \Delta t + Z_t \Delta W_t + Z_t^0 \Delta W_t^0 + o(\Delta t).$$

Orthogonality of independent increments gives the approximate targets

$$Z_t \approx \mathbb{E}\left[\left(\frac{\Delta Y_t}{\Delta t} + f_t\right) \Delta W_t \mid \mathcal{G}_t\right], \quad Z_t^0 \approx \mathbb{E}\left[\left(\frac{\Delta Y_t}{\Delta t} + f_t\right) \Delta W_t^0 \mid \mathcal{G}_t\right].$$

Both are learned with quadratic scores using recurrent fields $v_\theta(t, X_t, W^0_{[0,t]})$ and $v^0_\theta(t, X_t, W^0_{[0,t]})$.

One outer sweep alternates forward simulation and elicitation



- Each sweep acts on the **entire time horizon**.
- Damping stabilizes the fixed-point iteration and mirrors relaxed fictitious play.

The architecture separates state information from path dependence

Common statistic

$$S_\theta(t, W^0_{[0,t]})$$

- one GRU branch;
- adapted to common noise alone;
- learns non-Markovian conditional statistics.

Decoupling and martingale fields

$$(U_\theta, v_\theta, v^0_\theta)(t, X_t, W^0_{[0,t]})$$

- feedforward branch for (t, X_t);
- GRU branch for the common-noise path;
- concatenate, then four residual blocks.

Damping and repeated training stabilize the outer iteration

Soft update

$$\Psi^{k+1} = \delta\Psi^{k} + (1-\delta)\widehat{\Psi}^{k+1}, \qquad \delta = 0.5,$$

for $\Psi \in \{\Upsilon, Z, Z^0, S\}$.

Training inside each sweep

- 2,000 gradient steps;
- batch size 8,192;
- AdamW, learning rate 5×10^{-4};
- learning-rate decay every five steps.

The numerical experiment uses a large shared path sample

Discretization and sampling

- $N = 101$ time points;
- $M = 50{,}000$ simulated paths;
- $K = 20$ outer Picard sweeps;
- shared Brownian samples across updates.

Observed cost

- roughly 8 minutes 12 seconds per sweep;
- T4 GPU in Google Colab;
- small hidden width relative to path count.

State error controls the conditional measure-flow error

- The coefficients have linear growth and are Lipschitz in (x,y,z,z^0).
- The driver and terminal condition are Lipschitz in the statistic s.
- The solution admits a Lipschitz decoupling field in the state variable.
- Synchronous coupling controls conditional-law error:

$$\mathbb{E}\left[\mathcal{W}_2(\nu_t^k,\nu_t^\star)^2\right] \le \mathbb{E}|X_t^k - X_t^\star|^2.$$

State error therefore controls the error in the conditional measure flow seen by the coefficients.

One idealized Picard sweep contracts on short intervals

On a subinterval $I_i = [t_i, t_{i+1}]$ of length $h = T/N$, define

$$\mathcal{E}_i^k = \mathbb{E}\left[\sup_{t \in I_i} |X_t^k - X_t^\star|^2\right].$$

The idealized sweep satisfies

$$\mathcal{E}_i^k \leq \underbrace{\Gamma L^2 h(1 + 2\Gamma h)}_{\lambda(h)} \mathcal{E}_i^{k-1} + \Gamma \delta_i^k + \Gamma^2 h \tau_i^{k-1}.$$

- δ_i^k: mismatch at the forward initial boundary;
- τ_i^{k-1}: mismatch at the backward terminal boundary;
- $\lambda(h) = O(h)$, so $\lambda(h) < 1$ for a sufficiently fine grid.

The global bound separates iteration and boundary errors

Summing the local recursion across the partition gives

$$\mathbb{E}\left[\sup_{t\le T}|X_t^k - X_t^\star|^2\right] \le \lambda(h)^k \sum_{i=0}^{N-1} \mathcal{E}_i^0 + \sum_{j=1}^{k} \lambda(h)^{k-j}\left(\Gamma D^j + \Gamma^2 h\,\mathcal{T}^{j-1}\right).$$

Here

$$D^j = \sum_i \delta_i^j, \qquad \mathcal{T}^j = \sum_i \tau_i^j.$$

The initial iteration error decays geometrically; boundary errors are carried and geometrically damped.

Inexact neural solvers create a persistent error floor

Assume the learned statistic and backward fields have uniform error at most ε_{net}. Then

$$\mathbb{E}\left[\sup_{t\le T}|X_t^k - X_t^\star|^2\right] \le \underbrace{\lambda(h)^k\mathcal{E}^0}_{\text{Picard error}} + \underbrace{\text{carried boundary errors}}_{\text{time coupling}} + \underbrace{\frac{C\,T}{1-\lambda(h)}\varepsilon_{\text{net}}}_{\text{network floor}}.$$

- refining the grid improves the contraction factor;
- increasing k cannot remove training and capacity error;
- practical accuracy is governed by numerical and statistical approximation.

Three experiments test accuracy, flexibility, and scope

1. **Systemic risk with conditional mean interaction:** analytic benchmark for (X, Y, Z, Z^0, S).
2. **Systemic risk with quantile interaction:** same solver, different elicitable statistic.
3. **Economic growth with endogenous interest rates:** nonlinear mean-field equilibrium without a closed form.

The experiments progress from verification to a genuinely new equilibrium computation.

Systemic-risk FBSDE from an interbank control problem

Each bank chooses its borrowing/lending rate α to minimize

$$J(\alpha) = \mathbb{E}\left[\int_0^T \left\{ \frac{\alpha_t^2}{2} - q\alpha_t(S_t - X_t) + \frac{\epsilon}{2}(S_t - X_t)^2 \right\} \mathrm{d}t + \frac{c}{2}(S_T - X_T)^2 \right]$$

subject to

$$\mathrm{d}X_t = \left[a(S_t - X_t) + \alpha_t\right] \mathrm{d}t + \sigma\, \mathrm{d}B_t, \quad S_t = \mathbb{E}[X_t \mid \mathcal{F}_t^0], \quad B_t = \rho W_t^0 + \sqrt{1-\rho^2}\, W_t.$$

The first-order condition $\alpha_t^\star = q(S_t - X_t) - Y_t$ yields

$$\begin{aligned}
\mathrm{d}X_t &= \left[(a+q)(S_t - X_t) - Y_t\right] \mathrm{d}t + \sigma\, \mathrm{d}B_t,\\
\mathrm{d}Y_t &= \left[(a+q)Y_t + (\epsilon - q^2)(S_t - X_t)\right] \mathrm{d}t + Z_t\, \mathrm{d}W_t + Z_t^0\, \mathrm{d}W_t^0,\\
Y_T &= c(X_T - S_T).
\end{aligned}$$

The linear–quadratic structure yields a closed form

The backward variable has the form

$$Y_t = -\eta(t)(S_t - X_t),$$

where η solves the Riccati equation

$$\dot{\eta}(t) = 2(a+q)\eta(t) + \eta(t)^2 - (\epsilon - q^2).$$

Consequently,

$$S_t = \mathbb{E}[\xi] + \rho\sigma W_t^0, \qquad Z_t = \sigma\eta(t), \qquad Z_t^0 = 0.$$

Every output of the algorithm can be compared against an analytic target.

Benchmark parameters yield strong coupling

The experiment uses

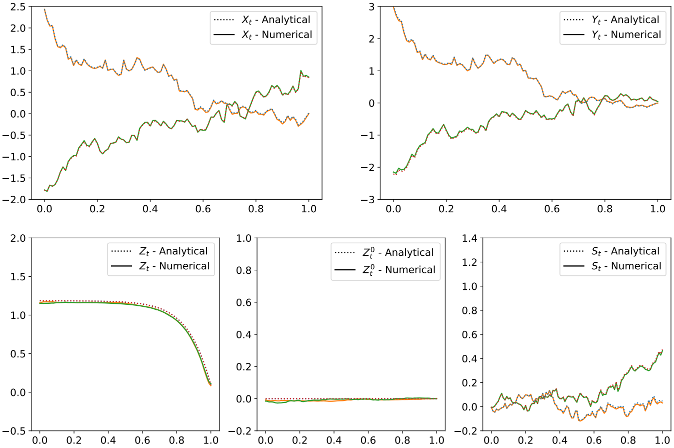
$$a = q = c = \sigma = 1, \qquad \epsilon = 10, \qquad \rho = 0.3,$$

with

$$\xi \sim \mathcal{N}(0, 4).$$

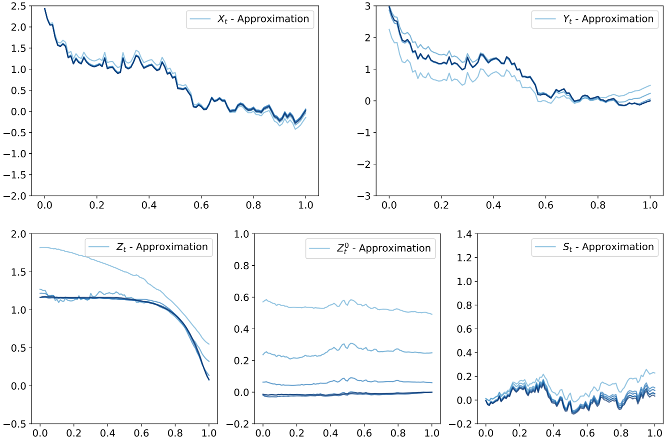
- ρ controls exposure to the common shock.
- $\epsilon - q^2$ determines the strength of the running interaction.
- The terminal penalty couples each bank to the population state at T.

The learned paths recover the analytic solution



Solid: numerical solution. Dotted: analytical solution. Two common/idiosyncratic path realizations.

Outer Picard sweeps progressively stabilize every component



One path at sweeps $1, 2, 4, 8, 16, 32$; darker curves correspond to later sweeps.

Accuracy degrades gradually as dimension increases

FGV EMAp ESCOLA DE MATEMÁTICA APLICADA

Mean Euclidean error at sweep $k = 30$

d	X	Y	Z	Z^0	S
1	.015	.025	.036	.062	.009
5	.020	.049	.056	.054	.022
10	.098	.174	.175	.219	.146

$$\mathrm{MEE}(a,b) = \frac{1}{M_{\mathrm{eval}}} \sum_{m=1}^{M_{\mathrm{eval}}} \left(\frac{1}{d|\mathcal{T}|} \sum_{t \in \mathcal{T}} \| a_t^{(m)} - b_t^{(m)} \|^2 \right)^{1/2}.$$

Root-mean-square error is computed along each path, then averaged over evaluation paths.

- Each dimension is an independent copy of the systemic-risk model.
- The architecture and training protocol remain otherwise comparable.
- The ten-dimensional case remains stable but exposes a clear capacity challenge.

Quantile interaction changes the equilibrium target

Replace the conditional mean by the conditional 60% quantile:

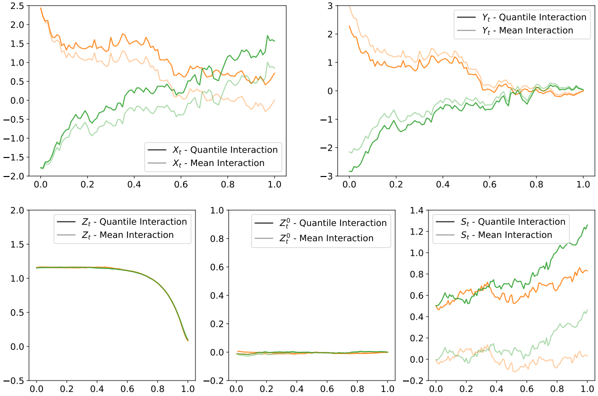
$$S_t = q_{0.60}(X_t \mid \mathcal{F}_t^0).$$

Train S_θ with the pinball score

$$\mathfrak{S}_{0.60}(s, x) = \big(\mathbf{1}_{\{x \le s\}} - 0.60\big)(s - x).$$

- the forward and backward equations are unchanged;
- the network architecture is unchanged;
- only the score defining the population statistic changes.

Quantile interaction shifts the population path upward



Dark paths: 60%-quantile interaction. Light paths: mean interaction under the same noises.

When agents target a higher population quantile, the learned measure flow rises relative to mean interaction.

The growth model couples consumption and aggregate capital

Agent i chooses consumption c^i to maximize

$$\mathbb{E}\left[\int_0^T u(c_t^i)\,\mathrm{d}t + \psi(K_T^i)\right]$$

subject to

$$\mathrm{d}K_t^i = \left[(r_t - \delta)K_t^i - c_t^i\right]\,\mathrm{d}t + \sigma\left(\rho\,\mathrm{d}W_t^0 + \sqrt{1-\rho^2}\,\mathrm{d}W_t^i\right).$$

- aggregate production depends on average capital;
- the equilibrium interest rate is the marginal product of capital;
- each agent's saving decision responds to the population.

The mean-field limit produces a nonlinear FBSDE

Choose

$$u(c) = \log c, \qquad \psi(K) = -\tfrac{1}{2}K^2, \qquad F(K) = \tfrac{C}{2}K^2.$$

Then

$$S_t = \mathbb{E}[K_t \mid \mathcal{F}_t^0], \qquad r_t = CS_t, \qquad c_t^\star = \frac{1}{Y_t}.$$

The equilibrium system is

$$\begin{aligned} \mathrm{d}K_t &= \left[(r_t - \delta)K_t - \frac{1}{Y_t}\right]\mathrm{d}t + \sigma\left(\rho\,\mathrm{d}W_t^0 + \sqrt{1-\rho^2}\,\mathrm{d}W_t\right), \\ \mathrm{d}Y_t &= -(r_t - \delta)Y_t\,\mathrm{d}t + Z_t\,\mathrm{d}W_t + Z_t^0\,\mathrm{d}W_t^0, \qquad Y_T = -K_T. \end{aligned}$$

Growth experiment with endogenous common fluctuations

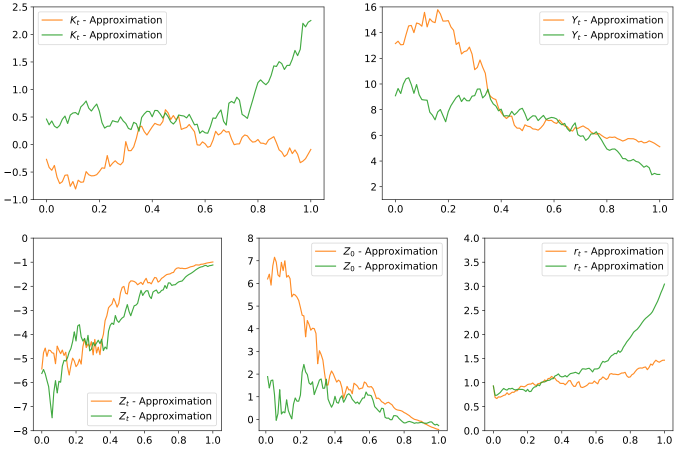
The initial capital and parameters are

$$K_0 \sim \mathcal{N}(0.5, 0.5^2), \qquad C = 1.5, \qquad \delta = 0.1, \qquad \sigma = 0.1.$$

- The common noise changes conditional aggregate capital S_t.
- The interest rate $r_t = CS_t$ becomes path-dependent on the common-noise history $W^0_{[0,t]}$.
- Optimal consumption feeds back into individual and aggregate capital.
- No analytical solution is available for comparison.

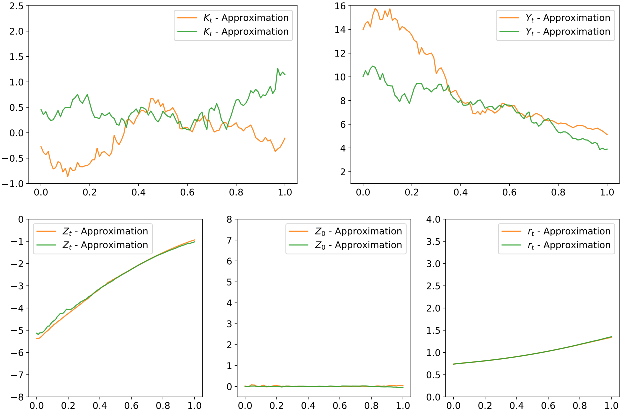
The solver must learn the equilibrium environment and optimal response simultaneously.

Common shocks move the entire equilibrium environment



Two paths of capital, adjoint process, martingale terms, and the endogenous rate $r_t = CS_t$.

Without common noise, the random environment collapses



The common martingale term satisfies $Z^0 = 0$, and r_t is deterministic.

The architecture learns that the common-noise channel is irrelevant when $\rho = 0$.

The learned fields yield a policy derivative

Because

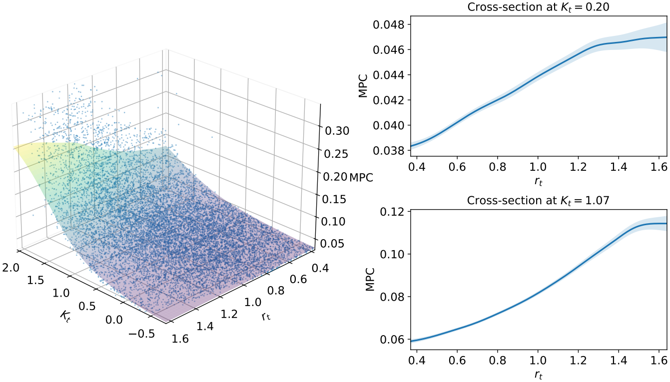
$$c_t^\star = \frac{1}{Y_t},$$

the marginal propensity to consume is

$$\partial_K c_t^\star = -\frac{1}{Y_t^2}\,\partial_K Y_t = -\frac{Z_t}{\sigma Y_t^2}.$$

- MPC measures how an additional unit of capital changes optimal consumption.
- The learned decoupling and martingale fields provide this derivative.
- We can inspect MPC as a function of (K_t, r_t), not only along sampled paths.

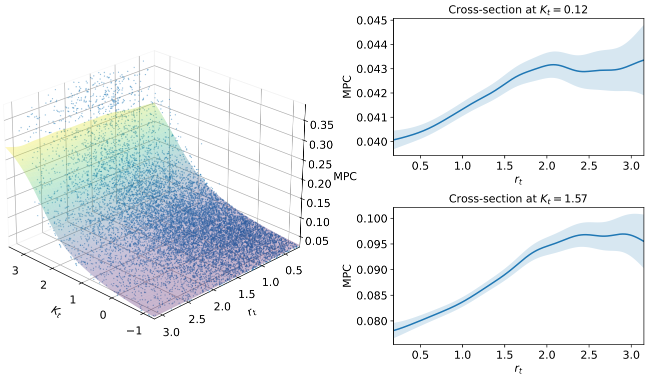
Mid-horizon policy responds strongly to capital and rates



Marginal propensity to consume at $t = 0.5$ as a function of capital K_t and interest rate r_t.

At mid-horizon the learned equilibrium response varies materially across both state variables.

Near maturity, the interest-rate response flattens



Marginal propensity to consume at $t = 0.9$. The response to interest rates flattens near the horizon.

The terminal condition increasingly dominates the remaining saving-versus-consumption trade-off.

What the method contributes

1. **Common noise without nested simulation:** conditional statistics are learned by proper scores.
2. **General interactions:** means, moments, and quantiles enter through the same solver interface.
3. **Whole-horizon coupling:** Picard sweeps alternate forward simulation and neural backward regression.
4. **A useful error picture:** geometric iteration error persists only down to boundary, discretization, and network floors.

Elicit the statistic the coefficients need; do not reconstruct more of the conditional law than the problem uses.

Questions?

Joint work with Felipe J. P. Antunes and Sebastian Jaimungal

github.com/fjpAntunes/mean-field-tools