

A deep solver for backward stochastic Volterra integral equations

Kristoffer Andersson

University of Verona

`kristofferherbert.andersson@univr.it`

Joint work with Alessandro Gnoatto and Camilo Andrés García Trillos

The XIII Bachelier World Congress

June 29, 2026

Table of contents

- 1 What is a Backward Stochastic Volterra Integral Equation?
- 2 Proposed algorithm
- 3 Convergence analysis
- 4 Numerical results

What is a Backward Stochastic Volterra Integral Equation?

Classical stochastic control

Consider the stochastic control problem

$$V(t, x) = \sup_{\alpha} \mathbb{E} \left[\int_t^T f(s, X_s^{\alpha}, \alpha_s) ds + g(X_T^{\alpha}) \mid X_t^{\alpha} = x \right].$$

- The objective does not depend explicitly on the evaluation time t .
- Bellman's principle of optimality holds.
- Dynamic programming and BSDEs provide the natural mathematical framework.

Time-inconsistent stochastic control

Consider instead

$$V(t, x) = \sup_{\alpha} \mathbb{E} \left[\int_t^T f(t, s, X_t^{\alpha}, X_s^{\alpha}, \alpha_s) ds + g(t, X_t^{\alpha}, X_T^{\alpha}) \mid X_t^{\alpha} = x \right].$$

- The objective now depends on the **evaluation time** t .
- The strategy chosen at time 0 need not remain optimal at time $t > 0$.
- Bellman's principle of optimality breaks down.

Examples:

- Mean–variance portfolio optimization
- Hyperbolic (non-exponential) discounting
- Dynamic risk measures
- Recursive utilities

FSDE–BSVIE formulation

Many time-inconsistent stochastic control problems can be characterized by the following coupled FSDE–BSVIE:

$$\begin{cases} X_t = x_0 + \int_0^t b(s, X_s, Y_s, Z_{s,s}) ds + \int_0^t \sigma(s, X_s, Y_s) dW_s, \\ Y_t = g(t, X_t, X_T) + \int_t^T f(t, s, X_t, X_s, Y_s, Z_{t,s}, Z_{s,s}) ds - \int_t^T Z_{t,s} dW_s. \end{cases}$$

- X is the forward state process.
- Y represents the equilibrium value process.
- The additional time index in $Z_{t,s}$ reflects the dependence on the evaluation time.

Proposed algorithm

Contribution of this paper

Given an FSDE–BSVIE, we ask:

How can we solve it numerically in high dimensions?

Main contributions:

- First **end-to-end solver** for BSVIEs.
- A neural-network formulation that directly approximates the solution fields.
- Applicable to both decoupled and coupled FSDE–BSVIEs.
- Non-asymptotic error analysis in the decoupled setting.

Deep BSDE: variational formulation

Starting from the variational formulation of FBSDEs (Han, Jentzen & E, PNAS, 2018)

$$\left\{ \begin{array}{l} \underset{y_0, Z}{\text{minimize}} \quad \mathbb{E} \left[\left| Y_T^{y_0, Z} - g(X_T^{y_0, Z}) \right|^2 \right], \\ X_t^{y_0, Z} = x_0 + \int_0^t b(s, X_s^{y_0, Z}, Y_s^{y_0, Z}, Z_s) \, ds + \int_0^t \sigma(s, X_s^{y_0, Z}, Y_s^{y_0, Z}, Z_s) \, dW_s, \\ Y_t^{y_0, Z} = y_0 - \int_0^t f(s, X_s^{y_0, Z}, Y_s^{y_0, Z}, Z_s) \, ds + \int_0^t (Z_s^{y_0, Z})^\top \, dW_s. \end{array} \right.$$

Can we extend this philosophy to BSVIEs?

Variational formulation for FSDE–BSVIEs

We extend the Deep BSDE philosophy to FSDE–BSVIEs.

$$\left\{ \begin{array}{l} \underset{Y,Z}{\text{minimize}} \quad \int_0^T \mathbb{E} \left[\left| Y_T^{Y,Z}(t) - g(t, X_t^{Y,Z}, X_T^{Y,Z}) \right|^2 \right] dt, \\ X_t^{Y,Z} = x_0 + \int_0^t b(s, X_s^{Y,Z}, Y_s, Z_{s,s}) ds + \int_0^t \sigma(s, X_s^{Y,Z}, Y_s) dW_s, \\ Y_T^{Y,Z}(t) = Y_t - \int_t^T f(t, s, X_s^{Y,Z}, Y_s, Z_{t,s}, Z_{s,s}) ds + \int_t^T Z_{t,s}^\top dW_s. \end{array} \right.$$

A single global optimization problem over the solution fields Y and Z .

Time-discrete variational formulation

Let $0 = t_0 < t_1 < \dots < t_N = T$ and $h = T/N$.

$$\left\{ \begin{array}{l} \underset{\theta^Y, \theta^Z}{\text{minimize}} \quad \sum_{n=0}^{N-1} \mathbb{E} [|Y_N(n) - g(t_n, X_n, X_N)|^2] h, \\ \\ X_n = x_0 + \sum_{k=0}^{n-1} b(t_k, X_k, \phi_k^Y, \phi_{k,k}^Z) h + \sum_{k=0}^{n-1} \sigma(t_k, X_k, \phi_k^Y) \Delta W_k, \\ \\ Y_N(n) = \phi_n^Y - \sum_{k=n}^{N-1} f(t_n, t_k, X_k, \phi_k^Y, \phi_{n,k}^Z, \phi_{k,k}^Z) h + \sum_{k=n}^{N-1} (\phi_{n,k}^Z)^\top \Delta W_k, \end{array} \right.$$

where

$$\phi_n^Y = \phi^Y(t_n, X_n; \theta^Y), \quad \phi_{n,k}^Z = \phi^Z(t_n, t_k, X_n, X_k; \theta^Z).$$

Convergence analysis

Main convergence result

Theorem

Under suitable regularity assumptions, there exists a constant $C > 0$ such that

$$\begin{aligned} & \int_0^T \mathbb{E} |Y_t - Y_{\pi(t)}^\pi|^2 dt + \int_0^T \int_t^T \mathbb{E} |Z_{t,s} - Z_{\pi(t), \pi(s)}^\pi|^2 ds dt \\ & \leq C \left(h + \sum_{n=0}^{N-1} \mathbb{E} |Y_N^\pi(n) - g(t_n, X_N^\pi)|^2 h \right). \end{aligned}$$

The proof consists of four main steps:

- 1 Approximate the BSVIE by a family of coupled BSDEs.
- 2 Discretize the BSDE family using an Euler–Maruyama scheme.
- 3 Derive a stability estimate comparing the BSDE and BSVIE discretizations.
- 4 Combine the convergence and stability estimates to obtain the final error bound.

Numerical results

Numerical example: decoupled FSDE–BSVIE

Let $d = 5$, $T = 1$, and define

$$\bar{X}_t = \frac{1}{d} \sum_{i=1}^d X_t^i.$$

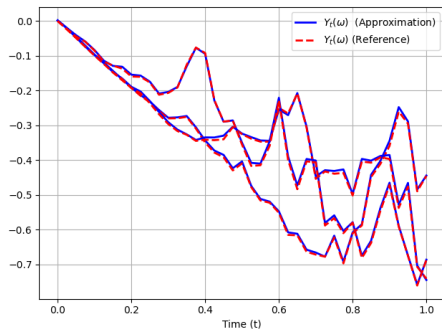
We consider

$$\begin{cases} X_t = x_0 + \int_0^t \text{diag}(\mu) X_s \, ds + \int_0^t \text{diag}(\sigma X_s) \, dW_s, \\ Y_t = t \sin(k \bar{X}_T) + \int_t^T \left(\frac{tk^2}{2d^2} \sin(k \bar{X}_s) \|\sigma X_s\|^2 - \mu^\top \sigma^{-1} Z_{t,s} \right) ds - \int_t^T Z_{t,s}^\top dW_s. \end{cases}$$

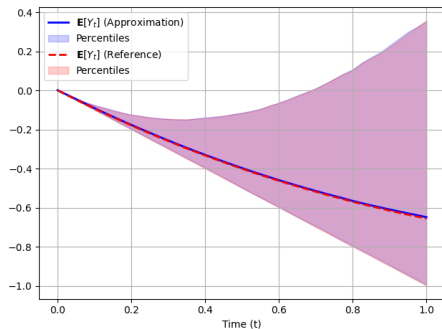
The closed-form solution is

$$Y_t = t \sin(k \bar{X}_t), \quad Z_{t,s} = \frac{tk}{d} \cos(k \bar{X}_s) \sigma X_s.$$

Numerical results: approximation of Y

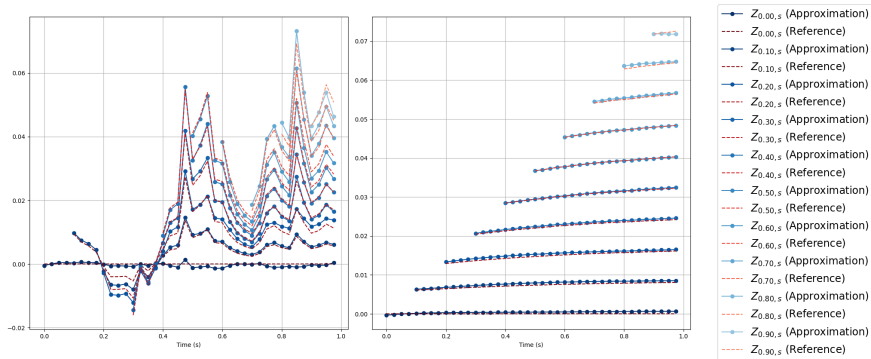


Three representative sample paths.



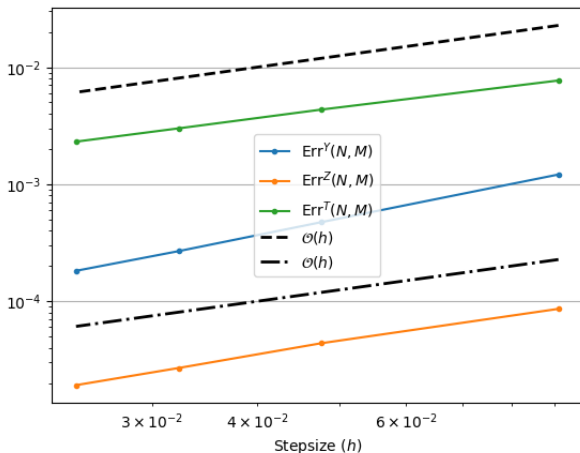
Sample mean with 5th and 95th percentiles.

Numerical results: approximation of Z



- The network accurately recovers the two-time process $Z_{t,s}$.
- Excellent agreement with the analytical reference solution for several evaluation times.

Empirical order of convergence



- The optimization loss converges with empirical order 1.
- The approximation errors for Y and Z converge with empirical order 0.5.
- This agrees with the theoretical error bound and the strong order of the Euler-Maruyama discretization.

Numerical example: coupled FSDE–BSVIE

Let $d = 5$, $T = 1$, and define

$$\bar{X}_t = \frac{1}{d} \sum_{i=1}^d X_t^i.$$

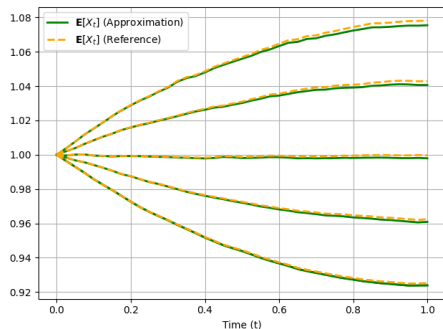
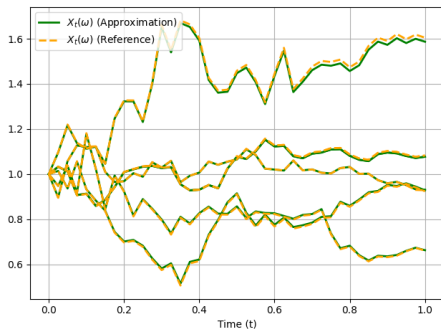
We consider the coupled system

$$\begin{cases} X_t = x_0 + \int_0^t (a + bZ_{s,s}) ds + \int_0^t (c + Y_s) \sigma dW_s, \\ Y_t = t \sin(k\bar{X}_T) \\ \quad + \int_t^T \left(-\frac{tk^2}{2d^2} \sin(k\bar{X}_s)(c + Y_s)^2 \|\mathbf{1}^\top \sigma\|^2 - \cos(k\bar{X}_s)(\mathbf{1}^\top a + s\mathbf{1}^\top b Z_{t,s}) \right) ds \\ \quad - \int_t^T Z_{t,s}^\top dW_s. \end{cases}$$

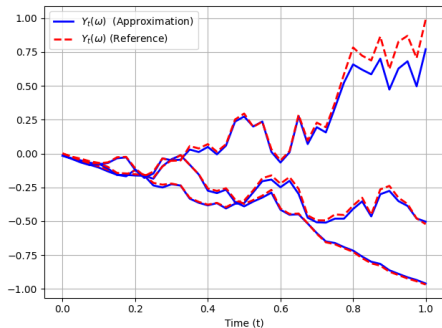
The closed-form solution is

$$Y_t = t \sin(k\bar{X}_t), \quad Z_{t,s} = t \cos(k\bar{X}_s)(c + s \sin(k\bar{X}_s)) \mathbf{1}^\top \sigma.$$

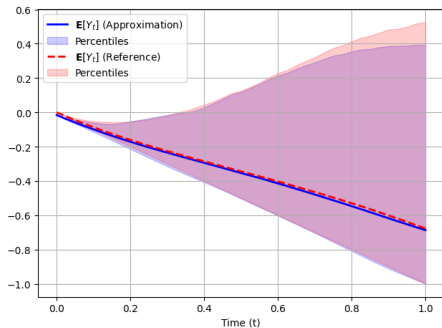
Numerical results: approximation of X



Numerical results: approximation of Y

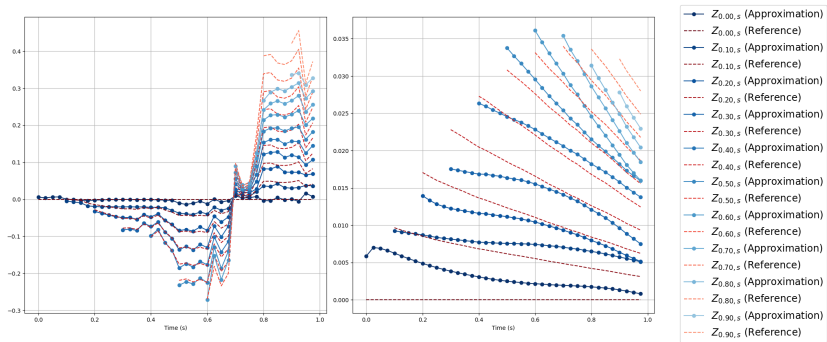


Three representative sample paths.



Sample mean with 5th and 95th percentiles.

Numerical results: coupled FSDE–BSVIE



- Accurate approximation of the two-time process $Z_{t,s}$.
- The same solver extends naturally to coupled FSDE–BSVIEs.

Scalability

We finally investigate how the method scales with the state dimension.

d	$\text{Err}(Y)$	$\text{Err}(Z)$	Runtime (s)
1	7.7×10^{-5}	8.2×10^{-5}	430
5	9.6×10^{-5}	2.9×10^{-5}	480
20	1.8×10^{-4}	7.9×10^{-6}	480
100	1.3×10^{-3}	4.5×10^{-6}	480
500	8.0×10^{-3}	2.9×10^{-6}	480

- Accurate approximations up to 500 dimensions.
- GPU batching keeps the wall-clock time essentially constant.

Take-away messages

- First end-to-end solver for BSVIEs.
- Non-asymptotic error analysis in the decoupled setting.
- Accurate approximations for coupled systems up to 500 dimensions.

Thank you!