Conditional Monte Carlo Learning to Avoid the Curse of Dimensionality

Who? Lokman ABBAS TURKI (Joint work with B. Diallo and G. Pagès)

When? February 2021

Plan

Nested MC regression

- From Monte Carlo (MC) regression to two steps Nested MC regression
- Increasing the learning depth in terms of regressions

Nested error control

- Learning depth and bias reduction
- Value at $t = 0$ and variance adjustment

Some numerical examples

Plan

For $s \in \{t_0, ..., t_{2^L}\}$

$$(f)\ U_s = E\left(\sum_{t_k \geq s}^{T} f(t_k, X_{t_k}, X_{t_{k+1}}) \middle| \mathcal{F}_s\right)$$

with

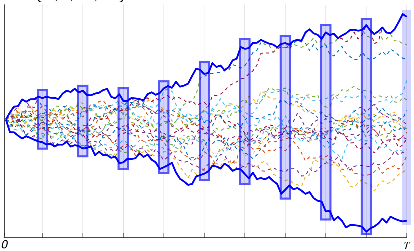
- $\{t_0, ..., t_{2^L}\} = \{0, \Delta_t, 2\Delta_t, ..., T\}$ sufficiently fine discretization set
- $X_{t_k} = \mathcal{E}_{t_{k-1}}(X_{t_{k-1}}, \xi_{t_k}) \in \mathbb{R}^{d_1}$, $X_{t_0}^{m_0} = x_0$ and $(\xi_{t_k})_{k=1,...,2^L}$ is the noise

  Basically $\quad \mathcal{E}_{t_k}(x, \theta) = x + \Delta_t b(t_k, x) + \sigma(t_k, x)\theta$
- Each $f(t_k, \cdot, \cdot)$ is $\mathcal{B}(\mathbb{R}^{d_1}) \otimes \mathcal{B}(\mathbb{R}^{d_1})$-measurable and satisfies $E(f^2(t_k, X_{t_k}, X_{t_{k+1}})) < \infty$ with $f(t_{2^L}, X_{t_{2^L}}, X_{t_{2^L+1}}) := f(t_{2^L}, X_{t_{2^L}})$
- Although some extensions are possible, we assume here f to be Lipschitz

Coarse discretization

- A decreasing time sequence $(s_k)_{k=0,...,2^L}$ taking its values in $\mathcal{S} \subset \{t_0, ..., t_{2^L}\}$ with $s_0 = T$ and $s_{2^L} = 0$
- A new operator $\delta(\cdot)$ on $(s_k)_{k \geq 0}$ that returns the next increment i.e. $\delta(s_k) = \min\{(s_l)_{l \in \{0,1,...,2^L\}}; s_k < s_l\}$ with default value $\delta(T) = T$

Standard regression methods and the curse of dimensionality



Two layers

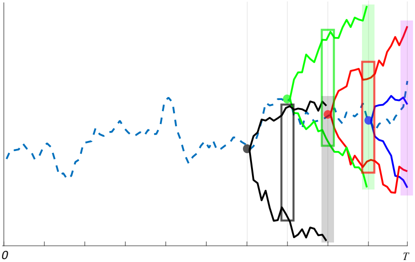
- First layer of M_0 outer trajectories $(X^{m_0}_{t_k})^{m_0=1,\dots,M_0}_{k=1,\dots,2^L}$ of the Markov process X that stays on the machine memory

$$X^{m_0}_{t_k} = \mathcal{E}_{t_{k-1}}(X^{m_0}_{t_{k-1}}, \xi^{m_0}_{t_k}) \text{ when } k \geq 1 \text{ and } X^{m_0}_{t_0} = x_0$$

- Second layer of M_1 inner trajectories starting at the coarse discretization

$$X^{m_0,m_1}_{s_j,t_k} = \mathcal{E}_{t_{k-1}}(X^{m_0,m_1}_{s_j,t_{k-1}}, \xi^{m_0,m_1}_{s_j,t_k}), \; k \geq 1 \text{ and } X^{m_0,m_1}_{s_j,s_j}\Big|_{m_1=1,\dots,M_1} = X^{m_0}_{s_j}$$

Conditional regression to avoid the curse of dimensionality



Indeed for $s_j \leq s < s_k$

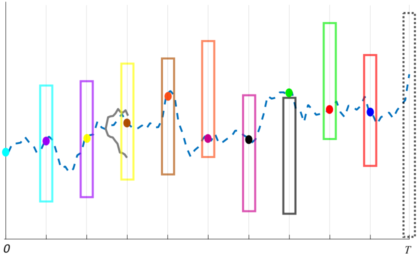
$$\begin{aligned} U_{s_k}(x) &= E\left(\sum_{t_l \geq s_k}^{T} f(t_l, X^{m_0,m_1}_{s_j,t_l}, X^{m_0,m_1}_{s_j,t_{l+1}}) \Big| X^{m_0,m_1}_{s_j,s_k} = x\right) \\ &= E\left(\sum_{t_l \geq s_k}^{T} f(t_l, X^{m_0,m_1}_{s,t_l}, X^{m_0,m_1}_{s,t_{l+1}}) \Big| X^{m_0,m_1}_{s,s_k} = x\right) \end{aligned} \quad (1)$$

Nested MC with one regression at each $s \in \mathcal{S}$

Coarse approximation

Given $t_k \in [s, \delta(s)[$ and denoting $\overline{u}^{m_0,\mathcal{S}}_{s,\delta(s)}(\cdot)$ the regressed representation of $U_{\delta(s)}(\cdot)$, we set the approximation $U_{\delta(s)}(X^{m_0,m_1}_{t_k,\delta(s)}) \approx \overline{u}^{m_0,\mathcal{S}}_{s,\delta(s)}(X^{m_0,m_1}_{t_k,\delta(s)})$

Two layers of approximation



Fine approximation

Given $t_k \in [s, \delta(s)[$, we approximate $U_{t_k}(X^{m_0}_{t_k})$ by

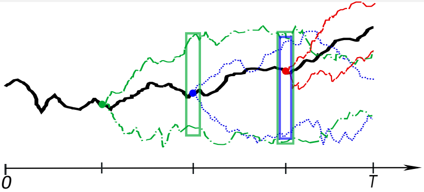
$$\widetilde{u}^{m_0,\mathcal{S}}_{t_k} = \frac{1}{M_1}\sum_{m_1=1}^{M_1}\left(\left[\sum_{t_{l+1}>t_k}^{\delta(s)} f(t_k, X^{m_0,m_1}_{t_k,t_l}, X^{m_0,m_1}_{t_k,t_{l+1}})\right] + \overline{u}^{m_0,\mathcal{S}}_{s,\delta(s)}(X^{m_0,m_1}_{t_k,\delta(s)})\right)$$

Benefits

- Nested MC regression more comparable to PDE methods
- Nested MC with outer trajectories that control the errors of inner ones
- Fine approximation of $V_s = E\left(f(U_{\delta(s)}(X_{\delta(s)}))\middle|\mathcal{F}_s\right)$

When the variance is not too large, small T or small dimension

Three steps depth



Linear regression arround X^{m_0}

For any $s_j < s_k$, a regression basis $\mathcal{T}^{m_0}_{s_j,s_k}(x - X^{m_0}_{s_k})$ with

$$\Lambda^{m_0}_{s_j,s_k} = E^{m_0}_{s_j}\left(\mathcal{T}^{m_0}_{s_j,s_k}(X^{m_0,m_1}_{s_j,s_k} - X^{m_0}_{s_k})^t\mathcal{T}^{m_0}_{s_j,s_k}(X^{m_0,m_1}_{s_j,s_k} - X^{m_0}_{s_k})\right)$$

and $\overline{\Lambda}^{m_0}_{s_j,s_k}$ its estimation using an offline independent sample

Coarse approximation

Around $X^{m_0}_{s_k}$ conditionally on $X^{m_0}_{s_j}$ $(s_k > s_j \in \mathcal{S})$, $U_{s_k}(x)$ approximated by

$$\overline{u}^{m_0,\mathcal{S}}_{s_j,s_k}(x) = \widetilde{u}^{m_0,\mathcal{S}}_{s_k} + {}^t\mathcal{T}^{m_0}_{s_j,s_k}(x - X^{m_0}_{s_k})H^{m_0,\mathcal{S}}_{s_j,s_k} \tag{3}$$

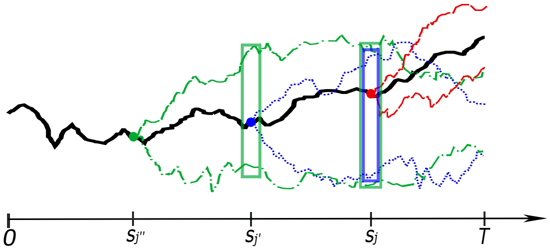
with

$H^{m_0,\mathcal{S}}_{s_j,s_k}$ resulting from the projection on $\mathcal{T}^{m_0}_{s_j,s_k}(x - X^{m_0}_{s_k})$ of

$$\overline{u}^{m_0,\mathcal{S}}_{s_j,\delta(s_k)}\left(X^{m_0,m_1}_{s_j,\delta(s_k)}\right) - \widetilde{u}^{m_0,\mathcal{S}}_{s_k} + \sum_{t_{l+1}>s_k}^{\delta(s_k)} f(t_l, X^{m_0,m_1}_{s_j,t_l}, X^{m_0,m_1}_{s_j,t_{l+1}})$$

Example $U_t = E\left(f(X_T)\middle|X_t\right)$, small T

$\mathcal{S}=\{0, s_{j''}, s_{j'}, s_j, T\}$



Fine approximation At $s \in \{s_{j''}, s_{j'}, s_j\}$, $\widetilde{u}_s^{m_0,\mathcal{S}} = \frac{1}{M_1}\sum_{m_1=1}^{M_1} \overline{u}_{s,\delta(s)}^{m_0,\mathcal{S}}(X_{s,\delta(s)}^{m_0,m_1})$

Coarse approximation For $s \in \{s_{j''}, s_{j'}, s_j\}$ and $r > s$,

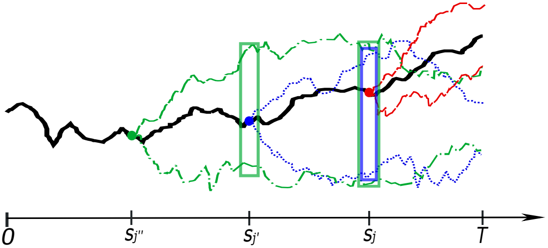
$$\overline{u}_{s,r}^{m_0,\mathcal{S}}(x) = \widetilde{u}_r^{m_0,\mathcal{S}} + {}^t\mathcal{T}_{s,r}^{m_0}(x - X_r^{m_0})H_{s,r}^{m_0,\mathcal{S}} \quad (4)$$

Terminal condition For $s \in \mathcal{S}$

$$\overline{u}_{s,T}^{m_0,\mathcal{S}}(x) = f(x) \quad (5)$$

Example $U_t = E\left(f(X_T)\middle|X_t\right)$

$\mathcal{S}=\{0, s_{j''}, s_{j'}, s_j, T\}$

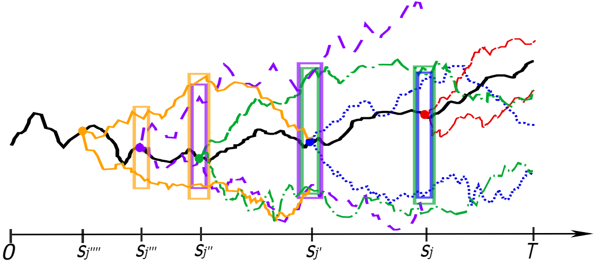


Dependence structure of computations

$$
\begin{array}{ccccccccccc}
\widetilde{u}^{m_0,\mathcal{S}}_{s_{j''}} & \to & \overline{u}^{m_0,\mathcal{S}}_{s_{j''},s_{j'}} & \to & \widetilde{u}^{m_0,\mathcal{S}}_{s_{j'}} & \to & \overline{u}^{m_0,\mathcal{S}}_{s_{j'},s_j} & \to & \widetilde{u}^{m_0,\mathcal{S}}_{s_j} & \to & \overline{u}^{m_0,\mathcal{S}}_{s_j,T} = f \\
 & & \searrow & & \uparrow & & & \searrow & \uparrow & & \\
 & & & & H^{m_0,\mathcal{S}}_{s_{j''},s_{j'}} & & & & H^{m_0,\mathcal{S}}_{s_{j'},s_j} & \to & \overline{u}^{m_0,\mathcal{S}}_{s_{j'},T} = f \\
 & & & & & \searrow & & & & & \\
 & & & & & & \overline{u}^{m_0,\mathcal{S}}_{s_{j''},s_j} & \to & \widetilde{u}^{m_0,\mathcal{S}}_{s_j} & \to & \overline{u}^{m_0,\mathcal{S}}_{s_j,T} = f \\
 & & & & & & & \searrow & \uparrow & & \\
 & & & & & & & & H^{m_0,\mathcal{S}}_{s_{j''},s_j} & \to & \overline{u}^{m_0,\mathcal{S}}_{s_{j''},T} = f
\end{array}
\tag{6}
$$

Plan

Example $U_t = E\left(f(X_T)\middle|X_t\right)$, T large



Dependence structure of computations

$$\widetilde{u}^{m_0,\mathcal{S}}_{s_{j'''},\boxed{s_j}} \rightarrow \overline{u}^{m_0,\mathcal{S}}_{s_{j'''},s_{j''}} \rightarrow \widetilde{u}^{m_0,\mathcal{S}}_{s_{j''},T} \rightarrow \overline{u}^{m_0,\mathcal{S}}_{s_{j''},s_{j'}} \rightarrow \widetilde{u}^{m_0,\mathcal{S}}_{s_{j'},T} \rightarrow \overline{u}^{m_0,\mathcal{S}}_{s_{j'},s_j} \rightarrow \widetilde{u}^{m_0,\mathcal{S}}_{s_j,T} \ldots f$$

$$\searrow \uparrow \quad H^{m_0}_{s_{j'''},s_{j''}} \qquad \searrow \uparrow \quad H^{m_0}_{s_{j''},s_{j'}} \qquad \searrow \uparrow \quad H^{m_0}_{s_{j'},s_j} \ldots f$$

$$\searrow \quad \overline{u}^{m_0,\mathcal{S}}_{s_{j'''},s_{j'}} \rightarrow \widetilde{u}^{m_0,\mathcal{S}}_{s_{j''},T} \rightarrow \overline{u}^{m_0,\mathcal{S}}_{s_{j''},s_j} \rightarrow \widetilde{u}^{m_0,\mathcal{S}}_{s_j,T} \ldots f \tag{7}$$

$$\searrow \uparrow \quad H^{m_0}_{s_{j'''},s_{j'}} \qquad \searrow \uparrow \quad H^{m_0}_{s_{j''},s_j} \ldots f$$

$$\searrow \quad \overline{u}^{m_0,\mathcal{S}}_{s_{j'''},s_j} = \overline{u}^{m_0,\mathcal{S}}_{s_{j''},s_j}$$

Decreasing complexity in terms of regressions

$$U_s = E\left(\sum_{t_l \geq s}^{T} f(t_l, X_{t_l}, X_{t_{l+1}}) \Big| \mathcal{F}_s\right)$$

Terminal condition truncation

$$\text{For } s_j \in \mathcal{S},\ \overline{u}^{m_0,\mathcal{S}}_{s_j,\overline{s_j}}(x) = \begin{cases} f(T,x) & \text{if } \overline{s_j} = T \\ \overline{u}^{m_0,\mathcal{S}}_{\delta(s_j),\overline{s_j}}(x) & \text{if } \overline{s_j} < T \end{cases} \tag{8}$$

$\overline{s_j}$ Choice

Since

$$E(U_{s_j}) = E\left[U_{\delta(s_j)} + \sum_{t_{l+1} > s_j}^{\delta(s_j)} f(t_l, X^{m_0}_{t_l}, X^{m_1}_{t_{l+1}})\right]$$

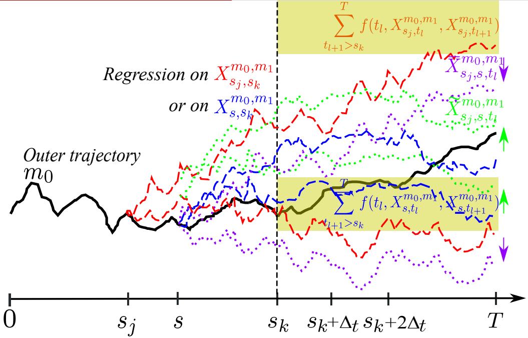
one can track the average bias by tracking the difference

$$D^{M_0}_{s_j} = \frac{1}{M_0}\sum_{m_0=1}^{M_0} \widetilde{u}^{m_0,\mathcal{S}}_{s_j,\overline{s_j}} - \frac{1}{M_0}\sum_{m_0=1}^{M_0}\left[\widetilde{u}^{m_0,\mathcal{S}}_{\delta(s_j),\overline{\delta(s_j)}} + \sum_{t_{l+1} > s_j}^{\delta(s_j)} f(t_l, X^{m_0}_{t_l}, X^{m_1}_{s_j,t_{l+1}})\right]$$

If M_0 is sufficiently large, one can improve this bias analysis since

$$E\left(U_{s_j} 1_{\{U_{s_j} \in [a,b]\}}\right) = E\left(1_{\{U_{s_j} \in [a,b]\}}\left[U_{\delta(s_j)} + \sum_{t_{l+1} > s_j}^{\delta(s_j)} f(t_l, X^{m_0}_{t_l}, X^{m_0}_{t_{l+1}})\right]\right) \tag{9}$$

Terminal condition truncation



Key idea on the control

$$\widetilde{X}^{m_0,m_1}_{s_j,s,s} = X^{m_0,m_1}_{s_j,s},$$
$$\widetilde{X}^{m_0,m_1}_{s_j,s,t_l} = \mathcal{E}_{t_{l-1}}(\mathcal{E}_{t_{l-2}}(\ldots\mathcal{E}_s(X^{m_0,m_1}_{s_j,s}, \xi^{m_0,m_1}_{s,s+\Delta_t}),\ldots\xi^{m_0,m_1}_{s,t_{l-1}}), \xi^{m_0,m_1}_{s,t_l})$$

$$E^{m_0}_s\left(\left[\overline{u}^{m_0,\mathcal{S}}_{s,s_k}(X^{m_0,m_1}_{s_j,s_k}) - U_{s_k}(X^{m_0,m_1}_{s_j,s_k})\right]^2\right) = E^{m_0}_s\left(\left[\overline{u}^{m_0,\mathcal{S}}_{s,s_k}(\widetilde{X}^{m_0,m_1}_{s_j,s,s_k}) - U_{s_k}(\widetilde{X}^{m_0,m_1}_{s_j,s,s_k})\right]^2\right)$$
$$\leq E^{m_0}_s\left(\left[\overline{u}^{m_0,\mathcal{S}}_{s,s_k}(\widetilde{X}^{m_0,m_1}_{s_j,s,s_k}) - \overline{u}^{m_0,\mathcal{S}}_{s,s_k}(X^{m_0,m_1}_{s,s_k})\right]^2\right)$$
$$+E^{m_0}_s\left(\left[\overline{u}^{m_0,\mathcal{S}}_{s,s_k}(X^{m_0,m_1}_{s,s_k}) - U_{s_k}(X^{m_0,m_1}_{s,s_k})\right]^2\right)$$
$$+E^{m_0}_s\left(\left[U_{s_k}(X^{m_0,m_1}_{s,s_k}) - U_{s_k}(\widetilde{X}^{m_0,m_1}_{s_j,s,s_k})\right]^2\right)$$

$$U_s = E\left(\sum_{t_l \geq s}^{T} f(t_l, X_{t_l}, X_{t_{l+1}}) \middle| \mathcal{F}_s\right)$$

Approximation at $s_{2^L} = t_0 = 0$

Two approximations are possible at 0

- The average U_0^{lear} of learned values

$$U_0^{\text{lear}} = \frac{1}{M_0}\sum_{m_0=1}^{M_0} \widetilde{u}_{0,\overline{0}}^{m_0,\mathcal{S}} \quad (10)$$

- The simulated value U_0^{sim}

$$U_0^{\text{sim}} = \frac{1}{M_0}\sum_{m_0=1}^{M_0}\left[\widetilde{u}_{\delta(0),\overline{\delta(0)}}^{m_0,\mathcal{S}} + \sum_{t_{l+1}>0}^{\delta(0)} f(t_l, X_{t_l}^{m_0}, X_{t_{l+1}}^{m_0})\right] \quad (11)$$

Risk measures using $\overline{u}$

Regressions produce $\text{Var}_{s_j}^{m_0}\left(\overline{u}_{s_j,s_k}^{m_0,\mathcal{S}}(X_{s_j,s_k}^{m_0,m_1})\right) \neq \text{Var}_{s_j}^{m_0}\left(U_{s_k}(X_{s_j,s_k}^{m_0,m_1})\right)$ and usually $\text{Var}_{s_j}^{m_0}\left(\overline{u}_{s_j,s_k}^{m_0,\mathcal{S}}(X_{s_j,s_k}^{m_0,m_1})\right) << \text{Var}_{s_j}^{m_0}\left(U_{s_k}(X_{s_j,s_k}^{m_0,m_1})\right)$

Fact to check using estimators

For $s_j < s_k$, we know that $E\left(\text{Var}_{s_j}[U_{s_k}]\right) = E\left(\left[U_{s_k} - E_{s_j}(U_{s_k})\right]^2\right)$

Important!

Bias reduction and variance adjustment possible because of the nested simulation

Plan

Original discrete BSDE Largely studied in papers by E. Gobet and his co-authors: J. P. Lemor, X. Warin and P. Turkedjiev

$$Y_T = f(T, X_T), \begin{cases} Y_{t_k} = E_{t_k}[Y_{t_{k+1}} + \Delta_t f(t_k, Y_{t_{k+1}}, Z_{t_k})] \\ Z_{t_k} = E_{t_k}[Y_{t_{k+1}}(W_{t_{k+1}} - W_{t_k})/\Delta_t] \end{cases}$$

Sub-discretization

$$\widehat{Y}_T = f(T, X_T), \begin{cases} \widehat{Y}_{s_k} = E_{s_k}[\widehat{Y}_{\delta(s_k)} + (\delta(s_k) - s_k) f(s_k, \widehat{Y}_{\delta(s_k)}, \widehat{Z}_{s_k})] \\ \widehat{Z}_{s_k} = E_{s_k}[\widehat{Y}_{\delta(s_k)}(W_{\delta(s_k)} - W_{s_k})/(\delta(s_k) - s_k)] \end{cases}$$

Original Snell envelope

$$V_T = g(X_T),\ \ V_{t_k} = g(X_{t_k}) \vee E_{t_k}[V_{t_{k+1}} + \Delta_t f(t_k, V_{t_{k+1}})]$$

Sub-discretization

$$\widehat{V}_T = g(X_T),\ \ \widehat{V}_{s_k} = g(X_{s_k}) \vee E_{s_k}[\widehat{V}_{\delta(s_k)} + (\delta(s_k) - s_k) f(s_k, \widehat{V}_{\delta(s_k)})]$$

Allen-Cahn equation: $f(t,x,y,z) = y - y^3$, $f(T,x) = \left[2 + \frac{2}{5}|x|^2_{d_1}\right]^{-1}$, $\mathcal{E}_{t_k}(x,w) = x + \sqrt{2}w$, $X_0 = 0$.

$u_b(0,0) = 0.0528$ $T = 0.3$, $d_1 = 100$, $M_0 = 2^4$ and $L = 4$.

M_1	Learned		Simulated		Runtime in sec. (10^{-3})
	Y_0^{learn}	std	Y_0^{sim}	std	
2^4	0.0454	$(\pm\ 0.0093)$	0.0455	$(\pm\ 0.0073)$	13
2^5	0.0513	$(\pm\ 0.0011)$	0.0517	$(\pm\ 0.0008)$	23
2^6	0.0523	$(\pm\ 0.0004)$	0.0518	$(\pm\ 0.0006)$	56
2^7	0.0526	$(\pm\ 0.0003)$	0.0515	$(\pm\ 0.0001)$	119
2^8	0.0525	$(\pm\ 0.0002)$	0.0517	$(\pm\ 0.0002)$	227
2^9	0.0527	$(\pm\ 0.0002)$	0.0515	$(\pm\ 0.0002)$	414

$u_b(0,0) = 0.0338$ $T = 1$, $d_1 = 100$, $M_0 = 2^5$ and $L = 6$.

M_1	Learned		Simulated		Runtime in sec.
	Y_0^{learn}	std	Y_0^{sim}	std	
2^5	0.0345	$(\pm\ 0.0008)$	0.0350	$(\pm\ 0.0021)$	2
2^6	0.0333	$(\pm\ 0.0003)$	0.0326	$(\pm\ 0.0004)$	4
2^7	0.0334	$(\pm\ 0.0002)$	0.0330	$(\pm\ 0.0003)$	7
2^8	0.0336	$(\pm\ 0.0002)$	0.0332	$(\pm\ 0.0002)$	12
2^9	0.0336	$(\pm\ 0.0001)$	0.0331	$(\pm\ 0.0001)$	27

Burgers-type PDE: $f(t,x,y,z) = \left(y - \frac{2+d_1}{2d_1}\right)\left(\sum_{i=1}^{d_1} z_i\right)$,

$$f(T,x) = \frac{e^{\left(T+\frac{1}{d_1}\sum_{i=1}^{d_1} x_i\right)}}{1+e^{\left(T+\frac{1}{d_1}\sum_{i=1}^{d_1} x_i\right)}},\ \mathcal{E}_{t_k}(x,w) = x + \frac{d_1}{\sqrt{2}}w,\ X_0 = 0.$$

$u_b(0,0) = 0.5000$ $\quad T = 0.2,\ d_1 = 100,\ M_0 = 2^6,\ L = 5.$

M_1	Learned		Simulated		Runtime in sec.
	Y_0^{learn}	std	Y_0^{sim}	std	
2^8	0.4785	(± 0.0428)	0.5170	(± 0.0431)	7
2^9	0.5113	(± 0.0450)	0.5108	(± 0.0450)	16
2^{10}	0.4966	(± 0.0448)	0.4912	(± 0.0447)	27
2^{11}	0.5022	(± 0.0421)	0.5012	(± 0.0435)	49

$u_b(0,0) = 0.5000$ $\quad T = 0.2,\ d_1 = 100,\ M_1 = 2^{11},\ L = 5.$

M_0	Learned		Simulated		Runtime in sec.
	Y_0^{learn}	std	Y_0^{sim}	std	
2^5	0.4953	(± 0.0618)	0.4941	(± 0.0615)	24
2^6	0.5022	(± 0.0424)	0.5013	(± 0.0435)	49
2^7	0.5079	(± 0.0346)	0.5066	(± 0.0342)	103
2^8	0.5158	(± 0.0221)	0.5151	(± 0.0221)	194
2^9	0.5023	(± 0.0164)	0.5029	(± 0.0164)	408

Time-dependent reaction-diffusion-type example:

$$f(t,x,y,z) = \min\left\{1, \left[y - \kappa - 1 - \sin\left(\lambda \sum_{i=1}^{d_1} x_i\right) e^{\frac{\lambda^2 d(t-T)}{2}}\right]^2\right\},$$

$$f(T,x) = 1 + \kappa + \sin\left(\lambda \sum_{i=1}^{d_1} x_i\right),\ \mathcal{E}_{t_k}(x,w) = x + w,\ X_0 = 0.$$

$u_b(0,0) = 1.6000$ $\quad T = 0.5,\ d_1 = 100,\ M_0 = 2^{10},\ L = 3,\ \kappa = 0.6,\ \lambda = \frac{1}{\sqrt{d_1}}$.

M_1	Learned		Simulated		Runtime in sec. (10^{-3})
	Y_0^{learn}	std	Y_0^{sim}	std	
2^5	1.8197	$(\pm\ 0.0386)$	1.7587	$(\pm\ 0.0287)$	244
2^6	1.7125	$(\pm\ 0.0104)$	1.6799	$(\pm\ 0.0116)$	311
2^7	1.6605	$(\pm\ 0.0037)$	1.6376	$(\pm\ 0.0091)$	466
2^8	1.6458	$(\pm\ 0.0023)$	1.6290	$(\pm\ 0.0089)$	817
2^9	1.6439	$(\pm\ 0.0019)$	1.6283	$(\pm\ 0.0061)$	1526

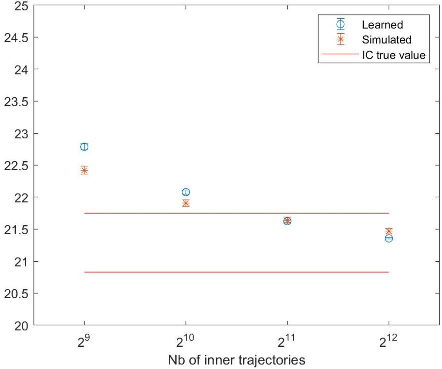
European derivatives with different interest rates for borrowing and lending:

$$f(t,x,y,z) = -R^l y - \frac{(\mu - R^l)}{\sigma}\sum_{i=1}^{d_1} z_i + (R^b - R^l)\max\left\{0, \frac{1}{\sigma}\sum_{i=1}^{d_1} z_i - y\right\},$$

$$f(T,x) = \max\{\max_{1\le i\le d_1} x_i - 120, 0\} - 2\max\{\max_{1\le i\le d_1} x_i - 150, 0\},$$

$$\mathcal{E}_{t_k}(x,w) = x\exp\left((\mu - \frac{\sigma^2}{2})\Delta_t + \sigma w\right),\ X_0 = 100.$$

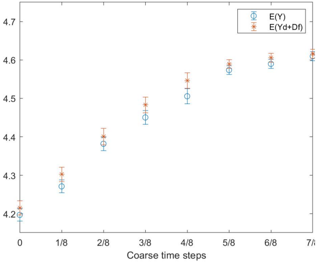
$$T = 0.5,\quad d_1 = 100,\quad M_0 = 2^7,\ L = 2,\quad \mu = \frac{6}{100},\ \sigma = \frac{2}{10},\quad R^l = \frac{4}{100},\quad R^b = \frac{6}{100}.$$

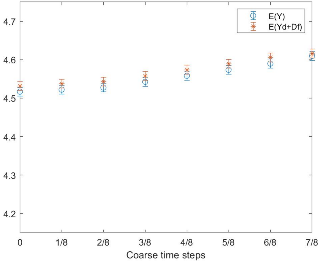


The runtime with 2^7 outer trajectories and 2^{11} inner trajectories is 53 seconds.

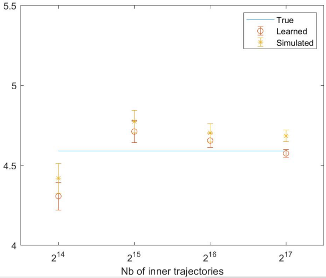
HJB equation: $f(t,x,y,z) = -|z|^2_{d_1}$, $f(T,x) = ln\left(\left[1+|x|^2_{d_1}\right]/2\right)$,
$\mathcal{E}_{t_k}(x,w) = x + \sqrt{2}w$, $X_0 = 0$, $T = 1$, $d_1 = 100$.

[Left] $\overline{s_k} = T$, [Right] $\overline{s_k} = \left(s_k + \frac{3}{8}\right) \wedge T$: $M_0 = 2^7$, $M_1 = 2^{15}$, $L = 3$.



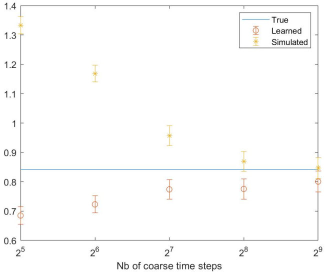


$M_0 = 2^7$, $L = 3$.

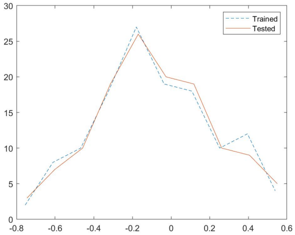


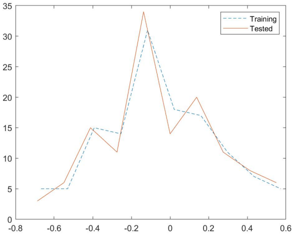
PDE with quadratically growing derivatives: $\psi(t,x) = \sin\left(\left[T - t + |x|^2_{d_1}\right]^\alpha\right)$, $f(t,x,y,z) = |z|^2_{d_1} - |\nabla_x \psi(t,x)|^2_{d_1} - \frac{\partial \psi}{\partial t}(t,x) - \frac{1}{2}(\Delta_x \psi)(t,x)$, $\alpha = 0.4$, $f(T,x) = \sin\left(|x|^{2\alpha}_{d_1}\right)$, $\mathcal{E}_{t_k}(x,w) = x + w$, $X_0 = 0$, $T = 1$, $d_1 = 100$.

$M_0 = 2^7$, $M_1 = 2^7$.
Execution time 620 seconds



$M_0 = 2^7$,
$M_1 = 2^{12}$, $L = 8$.



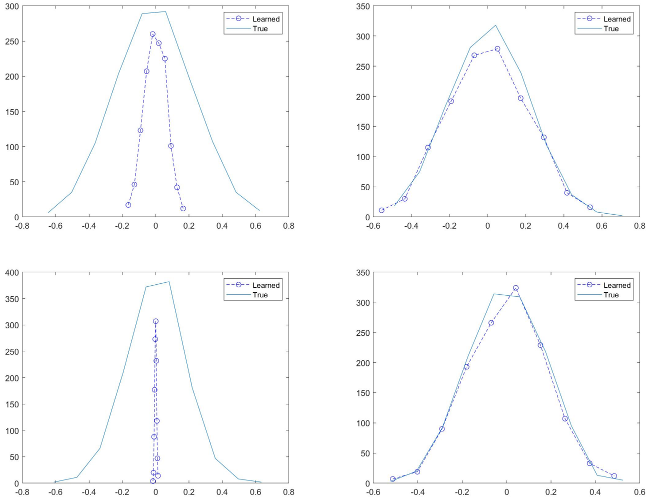


Initial margin: $\mathrm{IM}_{s_k} = \mathbb{ES}^a_{s_k}(L_{s_k,s_k+\delta})$, $L_{s_k,s_k+\delta} = V_{s_k+\delta} - V_{s_k}$, $a = 99\%$,
$V_{s_k} = \sum_{i=0}^{d_1} e^{-(T-s_k)r} E_{s_k}\left(\left[K - X^i_T\right]^+\right)$, $X_0 = K = 100$, $T = 1$, $d_1 = 100$,
$\mathcal{E}_{t_k}(x, w) = x \exp\left(\left(r - \frac{\sigma^2}{2}\right)\Delta_t + \sigma w\right)$, $r = 0.1$ and $\sigma = 0.2$.

Loss distribution [Left] Without variance adjustment,

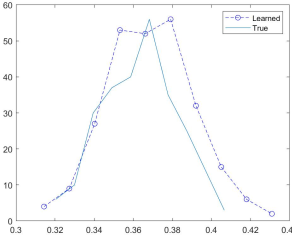
[Right] With variance adjustment;

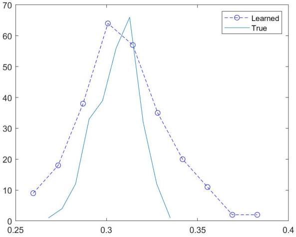
[top to bottom] $s_k \in \left\{\frac{29}{32}, \frac{9}{32}\right\}$; $M_0 = 2^8$, $M_1 = 2^8 * 5$.



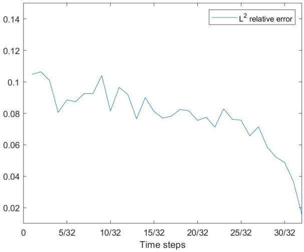
Initial margin: $\mathrm{IM}_{s_k} = \mathbb{ES}^a_{s_k}(L_{s_k,s_k+\delta})$, $L_{s_k,s_k+\delta} = V_{s_k+\delta} - V_{s_k}$, $a = 99\%$,
$V_{s_k} = \sum_{i=0}^{d_1} e^{-(T-s_k)r} E_{s_k}\left(\left[K - X_T^i\right]^+\right)$, $X_0 = K = 100$, $T = 1$, $d_1 = 100$,
$\mathcal{E}_{t_k}(x, w) = x \exp\left((r - \frac{\sigma^2}{2})\Delta_t + \sigma w\right)$, $r = 0.1$ and $\sigma = 0.2$.

IM distribution: [left to right] $s_k \in \{\frac{29}{32}, \frac{9}{32}\}$; $M_0 = 2^8$, $M_1 = 2^8 * 5$.



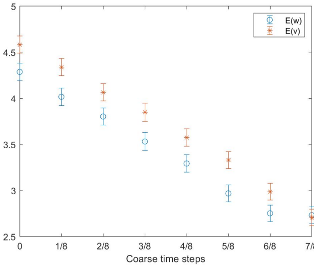


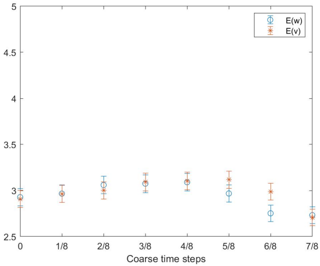
L2 relative error.



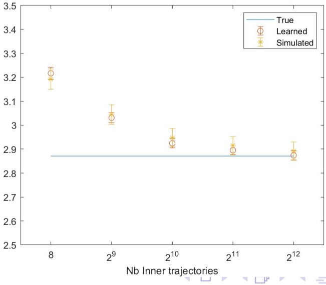
American option: $g(x) = \left[K - \prod_{i=1}^{d_1} (x_i)^{1/d_1}\right]^+$, $X_0 = K = 100$, $d_1 = 20$,
$\mathcal{E}_{t_k}(x, w) = x \exp\left((r - \frac{\sigma^2}{2})\Delta_t + \sigma w\right)$, $r = \log(1.1)$, $\sigma = 0.4$, $T = 1$.

[Left] $\overline{s_k} = T$, [Right] $\overline{s_k} = (s_k + \frac{1}{4}) \wedge T$: $M_0 = 2^{11}$, $M_1 = 2^{12}$ and $L = 3$.





$M_0 = 2^9$ and $L = 3$.



American option: $g(x) = \left[K - \prod_{i=1}^{d_1}(x_i)^{1/d_1}\right]^+$, $X_0 = K = 100$, $d_1 = 20$,
$\mathcal{E}_{t_k}(x, w) = x \exp\left((r - \frac{\sigma^2}{2})\Delta_t + \sigma w\right)$, $r = \log(1.1)$, $\sigma = 0.4$.

[BC] bias control
[VA] variance adjustment:
$M_0 = 2^{11}$,
$M_1 = 2^{12}$.

	L = 2 (T = 0.5)	L = 3 (T = 1)	L = 4 (T = 2)
[VA]	2.561 (± 0.035)	4.236 (± 0.042)	6.363 (± 0.054)
[BC]	2.493 (± 0.041)	3.734 (± 0.061)	5.130 (± 0.089)
[VA] + [BC]	2.291 (± 0.035)	2.890 (± 0.037)	3.961 (± 0.055)
Real Price	2.153	2.871	3.754