Overcoming the curse of dimensionality: from nonlinear Monte Carlo to deep artificial neural networks

Arnulf Jentzen (ETH Zurich, Switzerland)

Joint works with

Christian Beck (ETH Zurich, Switzerland),

Sebastian Becker (ZENAI AG, Switzerland),

Julius Berner (University of Vienna, Austria),

Patrick Cheridito (ETH Zurich, Switzerland),

Weinan E (Princeton University, USA),

Dennis Elbrächter (University of Vienna, Austria),

Philipp Grohs (University of Vienna, Austria),

Jiequn Han (Princeton University, USA),

Fabian Hornung (ETH Zurich, Switzerland, & KIT, Germany),

Martin Hutzenthaler (University of Duisburg-Essen, Germany),

Nor Jaafari (ZENAI AG, Switzerland),

Thomas Kurse (University of Giessen, Germany),

Tuan Anh Nguyen (University of Duisburg-Essen, Germany),

Diyora Salimova (ETH Zurich, Switzerland),

Christoph Schwab (ETH Zurich, Switzerland),

Timo Welti (ETH Zurich, Switzerland), and

Philippe von Wurstemberger (ETH Zurich, Switzerland)

Computational problems from

- **Financial Engineering** (Evaluations of risks and financial products, XVA, optimal stopping),
- **Operations Research** (Optimal control, robots, game intelligence, optimal use of resources, formation of prices),
- **Filtering** (Chemical engineering, Kushner and Zakai equations)

often require approximations for high-dimensional functions such as $u\colon [0,1]^d \to \mathbb{R}$ for $d \in \mathbb{N}$ large.

Approximations methods such as finite element methods, finite differences, sparse grids suffer under *the curse of dimensionality* (Bellman 1957).

Monte Carlo method based on Feynman-Kac formula:
high-dimensional linear partial differential equations (PDEs)

Deep BSDE method:
Han, J, E 2017 *PNAS*,
E, Han, J 2017 *Comm. Math. Stat.*,
. . .

Theorem (Hutzenthaler, J, Kruse, Nguyen 2019)

Let $T, p, \kappa > 0$, *let* $f\colon \mathbb{R} \to \mathbb{R}$ *be Lipschitz,* $\forall\, d \in \mathbb{N}$ *let* $g_d \in C(\mathbb{R}^d, \mathbb{R})$ *and* $u_d\colon [0, T] \times \mathbb{R}^d \to \mathbb{R}$ *be an at most poly. grow. solution of*

$$\tfrac{\partial u_d}{\partial t} = \Delta_x u_d + f(u_d) \qquad \text{with} \qquad u_d(0, \cdot) = g_d,$$

assume $|g_d(x)| \leq \kappa d^\kappa (1 + \|x\|^\kappa)$, *let* $\mathcal{A}_l\colon \mathbb{R}^l \to \mathbb{R}^l$, $l \in \mathbb{N}$, *satisfy* $\mathcal{A}_l(x_1, \ldots, x_l) = (\max\{x_1, 0\}, \ldots, \max\{x_l, 0\})$, *let*

$$\mathbf{N} = \cup_{L\in\mathbb{N}} \cup_{l_0,\ldots,l_L\in\mathbb{N}} \left(\times_{n=1}^{L} (\mathbb{R}^{l_n \times l_{n-1}} \times \mathbb{R}^{l_n})\right),$$

let $\mathcal{R}\colon \mathbf{N} \to \cup_{a,b=1}^{\infty} C(\mathbb{R}^a, \mathbb{R}^b)$ *satisfy for all* $L \in \mathbb{N}$, $l_0, \ldots, l_L \in \mathbb{N}$, $\Phi = ((W_1, B_1), \ldots, (W_L, B_L)) \in \times_{n=1}^{L} (\mathbb{R}^{l_n \times l_{n-1}} \times \mathbb{R}^{l_n})$, $x_0 \in \mathbb{R}^{l_0}, \ldots, x_{L-1} \in \mathbb{R}^{l_{L-1}}$ *with* $\forall\, n \in \mathbb{N} \cap (0, L)\colon x_n = \mathcal{A}_{l_n}(W_n x_{n-1} + B_n)$ *that*

$$(\mathcal{R}\Phi)(x_0) = W_L x_{L-1} + B_L,$$

let $\mathcal{P}\colon \mathbf{N} \to \mathbb{N}$ *be the number of parameters, and let* $(G_{d,\varepsilon})_{d\in\mathbb{N},\, \varepsilon\in(0,1]} \subseteq \mathbf{N}$ *satisfy* $\mathcal{P}(G_{d,\varepsilon}) \leq \kappa d^\kappa \varepsilon^{-\kappa}$ *and* $|g_d(x) - (\mathcal{R}G_{d,\varepsilon})(x)| \leq \varepsilon \kappa d^\kappa (1 + \|x\|^\kappa)$. *Then* $\exists\, (U_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,1]} \subseteq \mathbf{N}$, $c > 0\colon \forall\, d \in \mathbb{N}$, $\varepsilon \in (0, 1]\colon$

$$\left[\int_{[0,T]\times[0,1]^d} |u_d(y) - (\mathcal{R}U_{d,\varepsilon})(y)|^p \, dy\right]^{1/p} \leq \varepsilon \quad \text{and} \quad \mathcal{P}(U_{d,\varepsilon}) \leq c\, d^c \varepsilon^{-c}.$$

Linear PDEs: Grohs, Hornung, J, von Wurstemberger 2018; Berner, Grohs, J 2018; Elbrächter, Grohs, J, Schwab 2018; J, Salimova, Welti 2018

Full history recursive Multilevel-Picard method: Let $T > 0, L, p \geq 0, \Theta = \cup_{n=1}^{\infty} \mathbb{Z}^n$, $\forall\, d \in \mathbb{N}$ let $g_d \in C(\mathbb{R}^d, \mathbb{R})$ satisfy $\forall\, x \in \mathbb{R}^d\colon\ |g_d(x)| \leq L(1 + \|x\|^p)$, let $f\colon \mathbb{R} \to \mathbb{R}$ be Lipschitz, let $(\Omega, \mathcal{F}, \mathbb{P})$ probab. sp., let $W^{d,\theta}\colon [0, T] \times \Omega \to \mathbb{R}^d$, $d \in \mathbb{N}$, $\theta \in \Theta$, be i.i.d. Brownian motions, let $S^\theta\colon [0, T] \times \Omega \to \mathbb{R}$, $\theta \in \Theta$, i.i.d. continuous satisfying $\forall\, t \in [0, T]$, $\theta \in \Theta$ that S^θ_t is $\mathcal{U}_{[t,T]}$-distributed, assume that $(S^\theta)_{\theta\in\Theta}$ and $(W^{d,\theta})_{\theta\in\Theta, d\in\mathbb{N}}$ are independent, let $U^{d,\theta}_{n,M}\colon [0, T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$, $n, M \in \mathbb{Z}$, $\theta \in \Theta$, $d \in \mathbb{N}$, satisfy $\forall\, d, n, M \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0, T]$, $x \in \mathbb{R}^d$: $U^{d,\theta}_{-1,M}(t, x) = U^{d,\theta}_{0,M}(t, x) = 0$ and

$$
\begin{aligned}
U^{d,\theta}_{n,M}(t, x) &= \sum_{m=1}^{M^n} \frac{g_d\big(x + W^{d,(\theta,0,-m)}_{T-t}\big)}{M^n} \\
&\quad + \Bigg[\sum_{l=0}^{n-1} \frac{(T-t)}{M^{n-l}} \sum_{m=1}^{M^{n-l}} f\Big(U^{d,(\theta,l,m)}_{l,M}\big(S^{(\theta,l,m)}_t, x + W^{d,(\theta,l,m)}_{S^{(\theta,l,m)}_t - t}\big)\Big) \\
&\qquad\qquad - \mathbb{1}_{\mathbb{N}}(l)\, f\Big(U^{d,(\theta,l,m)}_{l-1,M}\big(S^{(\theta,l,m)}_t, x + W^{d,(\theta,l,m)}_{S^{(\theta,l,m)}_t - t}\big)\Big)\Bigg]
\end{aligned}
$$

and $\forall\, d, n \in \mathbb{N}$ let $\mathrm{Cost}_{d,n} \in \mathbb{N}$ be the comp. cost of $U^{d,0}_{n,n}(0, 0)$.

Then

i) $\forall\, d \in \mathbb{N}$: there exists $u_d\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$ at most polyn. grow. solution of

$$\frac{\partial u_d}{\partial t} + \tfrac{1}{2}\Delta_x u_d + f(u_d) = 0 \qquad \text{with} \qquad u_d(T,\cdot) = g_d$$

and

ii) $\forall\, \delta > 0$: there exist $n\colon \mathbb{N} \times (0,\infty) \to \mathbb{N}$ and $C > 0$: $\forall\, d \in \mathbb{N}, \varepsilon > 0$:

$$\left(\mathbb{E}\left[\left|u_d(0,0) - U^{d,0}_{n_{d,\varepsilon},n_{d,\varepsilon}}(0,0)\right|^2\right]\right)^{1/2} \le \varepsilon$$

and

$$\mathrm{Cost}_{d,n_{d,\varepsilon}} \le C d^{1+p(1+\delta)} \varepsilon^{-(2+\delta)}.$$

Extensions: Algorithms/Simulations/Proofs: Fully nonlinear PDEs (Beck, E, J 2018 *JNS*), Optimal stopping (Becker, Cheridito, J 2018 *JMLR*), Uniform errors (Beck, Becker, Grohs, Jaafari, J 2018), Semilinear PDEs/CVA (Hutzenthaler, J, von Wurstemberger 2019), . . .