Overcoming the curse of dimensionality: from nonlinear Monte Carlo to deep learning

Arnulf Jentzen (CUHK Shenzhen, China & University of Münster, Germany)
Funded by *German Research Foundation* through Cluster of Excellence *Mathematics Münster*

Joint works with

Patrick Cheridito (ETH Zurich, Switzerland)
Martin Hutzenthaler (University of Duisburg-Essen, Germany)
Benno Kuckuck (University of Münster, Germany)
Thomas Kruse (University of Giessen, Germany)
Tuan Anh Nguyen (University of Duisburg-Essen, Germany)
Joshua Lee Padgett (University of Arkansas, USA)
Florian Rossmannek (ETH Zurich, Switzerland)
Adrian Riekert (University of Münster, Germany)

14:50, Friday, June 25th, 2021

Theorem (Hutzenthaler, J, Kruse, Nguyen 2019 *SN PDE*; Kuckuck, J, Padgett 2021)

Let $T, p, \kappa > 0$, *let* $f\colon \mathbb{R} \to \mathbb{R}$ *be Lipschitz,* $\forall\, d \in \mathbb{N}$ *let* $g_d \in C^1(\mathbb{R}^d, \mathbb{R})$ *and* $u_d\colon [0, T] \times \mathbb{R}^d \to \mathbb{R}$ *be an at most poly. grow. solution of*

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

assume $|g_d(x)| + \|(\nabla g_d)(x)\| \leq \kappa d^\kappa (1 + \|x\|^\kappa)$, *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* $L \in \mathbb{N}$, $l_0, \ldots, l_L \in \mathbb{N}$, $\Phi = ((W_1, B_1), \ldots, (W_L, B_L)) \in \times_{n=1}^{L} (\mathbb{R}^{l_n \times l_{n-1}} \times \mathbb{R}^{l_n})$, $x_0 \in \mathbb{R}^{l_0}, \ldots, x_L \in \mathbb{R}^{l_L}$ *with* $\forall\, n \in \{1, \ldots, L\}\colon x_n = \mathcal{A}_{l_n}(W_n x_{n-1} + B_n)$ *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 *Mem. AMS*; Berner, Grohs, J 2018 *SIAM J. Math. Data Sci.*; Elbrächter, Grohs, J, Schwab 2018; . . .

Let $d, H, \mathcal{P} \in \mathbb{N}$, $\mathscr{a} \in \mathbb{R}$, $\mathscr{b} > \mathscr{a}$, $u \in C([\mathscr{a}, \mathscr{b}]^d, \mathbb{R})$ satisfy $\mathcal{P} = dH + 2H + 1$, let $\mu\colon \mathcal{B}([\mathscr{a}, \mathscr{b}]^d) \to [0, \infty)$ be a finite measure, let $\mathcal{L}\colon \mathbb{R}^{\mathcal{P}} \to \mathbb{R}$ satisfy $\forall\, \theta \in \mathbb{R}^{\mathcal{P}}$:

$$\mathcal{L}(\theta) = \int_{[\mathscr{a},\mathscr{b}]^d} \Big[u(x) - \theta_{\mathcal{P}} - \textstyle\sum_{i=1}^{H} \theta_{H(d+1)+i} \max\{\theta_{Hd+i} + \sum_{j=1}^{d} \theta_{(i-1)d+j} x_j, 0\}\Big]^2 \mu(\mathrm{d}x),$$

let $\mathcal{G}\colon \mathbb{R}^{\mathcal{P}} \to \mathbb{R}^{\mathcal{P}}$ be an appropriately generalized gradient of $\mathcal{L}$, and let $\Theta \in C([0, \infty), \mathbb{R}^{\mathcal{P}})$ satisfy for all $t \in [0, \infty)$ that $\Theta_t = \Theta_0 - \int_0^t \mathcal{G}(\Theta_s)\, \mathrm{d}s$.

Theorem (Cheridito, J, Riekert, Rossmannek 2021; J, Riekert 2021)

Assume for all $x, y \in [\mathscr{a}, \mathscr{b}]^d$ that $u(x) = u(y)$. Then there exist no non-global local minima and no saddle points of $\mathcal{L}$ and $\lim_{t\to\infty} \mathcal{L}(\Theta_t) = 0$.

Theorem (Cheridito, J, Rossmannek 2021; J, Riekert 2021)

Assume $\mu = \lambda_{[\mathscr{a},\mathscr{b}]}$, let $\alpha, \beta \in \mathbb{R}$ satisfy $u(x) = \alpha x + \beta$, and assume $\sup_{t\in[0,\infty)} \|\Theta_t\| < \infty$ and $\mathcal{L}(\Theta_0) < \frac{\alpha^2(\mathscr{b}-\mathscr{a})^3}{12(2\lfloor H/2\rfloor+1)^4}$. Then there exist ∞-many non-global local minima and ∞-many saddle points of $\mathcal{L}$ and $\lim_{t\to\infty} \mathcal{L}(\Theta_t) = 0$.

Theorem (J, Riekert 2021)

Assume $\mu \ll \lambda_{[\mathscr{a},\mathscr{b}]^d}$ and $\sup_{t\in[0,\infty)} \|\Theta_t\| < \infty$. Then there exists $\vartheta \in \mathcal{G}^{-1}(\{0\})$ such that $\lim_{t\to\infty} \mathcal{L}(\Theta_t) = \mathcal{L}(\vartheta)$.