Overcoming the course of dimensionality with DNNs: Theoretical approximation results for PDEs

Philippe von Wurstemberger (ETH Zurich, Switzerland)

Joint works with

Arnulf Jentzen (ETH Zurich, Switzerland),
Fabian Hornung (ETH Zurich, Switzerland & Karlsruhe Institute of Technology, Germany),
Philipp Grohs (University of Vienna, Austria),
Martin Hutzenthaler (University of Duisburg-Essen, Germany)
Thomas Kruse (University of Duisburg-Essen, Germany)
Tuan Anh Nguyen (University of Duisburg-Essen, Germany)

SMAI 2019, Guidel Plages

13 May 2019

Goal: Approximate solutions of PDEs numerically, e.g., Black-Scholes PDE:

$$\left(\tfrac{\partial u}{\partial t}\right)(t,x) + \sum_{i=1}^{d} \tfrac{|\beta_i|^2 |x_i|^2}{2} \left(\tfrac{\partial^2 u}{\partial (x_i)^2}\right)(t,x) + \alpha_i x_i \left(\tfrac{\partial u}{\partial x_i}\right)(t,x) = 0,$$

$$u(T,x) = g(x),$$

where $u\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$, $g\colon \mathbb{R}^d \to \mathbb{R}$, $(\alpha_i)_{i \in \{1,\dots,d\}}, (\beta_i)_{i \in \{1,\dots,d\}} \subseteq \mathbb{R}$, $T > 0$, $t \in [0,T]$, and $x \in \mathbb{R}^d$. Dimension $d \in \mathbb{N}$ corresponds to # assets in model.

Major issue for $d \gg 1$:

Classical approximations methods such as finite differences, finite element methods, sparse grids suffer under the **curse of dimensionality**.

Monte Carlo methods based on Feynman-Kac formula :
approximations at a fixed point without curse of dimensionality.

Methods based on Deep learning:

- E, Han, & Jentzen, *Solving high-dimensional partial differential equations using deep learning. arXiv 2017. Proc. Natl. Acad. Sci. U.S.A. 2018*
- E, Han, & Jentzen, *Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. arXiv 2017. Comm. Math. Stat. 2017.*
- many more: Beck, Becker, Grohs, Jaafari, Jentzen 2018; Fujii, Takahashi, Takahashi 2017; E, Yu 2017; Sirignano, Spiliopoulos 2017; Berg, Nyström 2017; Henry-Labordere 2017; Raissi 2018; Chan-Wai-Nam, Mikael, Warin 2018; Farahmand, Nabi, Nikovski 2017; Goudenege, Molent, Zanette 2019; Huré, Pham, Warin 2019; . . .

Empirical observation:

DNNs seem to be able to **overcome** the **curse of dimensionality**

$\Rightarrow$ Theoretical evidence for this observation?

Definition (Artificial neural networks (ANNs), Parameters, Realizations of ANNs)

For all $d \in \mathbb{N}$, $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ let $\mathbf{A}_d \in C(\mathbb{R}^d, \mathbb{R}^d)$ satisfy $\mathbf{A}_d(x) = (\max\{x_1, 0\}, \ldots, \max\{x_d, 0\})$. We denote by $\mathcal{N}$ the set given by

$$\mathcal{N} = \cup_{L \in \{2,3,\ldots\}} \cup_{(l_0, l_1, \ldots, l_L) \in \mathbb{N}^{L+1}} \left(\times_{k=1}^{L} (\mathbb{R}^{l_k \times l_{k-1}} \times \mathbb{R}^{l_k}) \right)$$

and we denote by $\mathcal{P} \colon \mathcal{N} \to \mathbb{N}$ and $\mathcal{R} \colon \mathcal{N} \to \cup_{k,l=1}^{\infty} C(\mathbb{R}^k, \mathbb{R}^l)$ the functions satisfying $\forall\, L \in \{2, 3, \ldots\}$, $(l_0, l_1, \ldots, l_L) \in \mathbb{N}^{L+1}$, $\Phi = ((W_1, B_1), \ldots, (W_L, B_L)) \in \left(\times_{k=1}^{L} (\mathbb{R}^{l_k \times l_{k-1}} \times \mathbb{R}^{l_k}) \right)$, $x \in \mathbb{R}^{l_0}$ that

$$\mathcal{P}(\Phi) = \textstyle\sum_{k=1}^{L} l_k(l_{k-1} + 1), \qquad \mathcal{R}(\Phi) \in C(\mathbb{R}^{l_0}, \mathbb{R}^{l_L}),$$

$$(\mathcal{R}(\Phi))(x) = W_L(\mathbf{A}_{l_{L-1}}(W_{L-1}(\ldots \mathbf{A}_{l_1}(W_1 x + B_1) \ldots) + B_{L-1})) + B_L.$$

ANN have capacity to overcome curse of dim. for Black-Scholes PDEs

Theorem (Grohs, Hornung, Jentzen, vW 2018)

Let $T, \kappa > 0$, $(\alpha_{d,i})_{i\in\{1,\ldots,d\},d\in\mathbb{N}}, (\beta_{d,i})_{i\in\{1,\ldots,d\},d\in\mathbb{N}} \subseteq \mathbb{R}$ *satisfy that* $\sup_{i\in\{1,\ldots,d\},d\in\mathbb{N}}(|\alpha_{d,i}| + |\beta_{d,i}|) < \infty$, $\forall\, d \in \mathbb{N}$ *let* $g_d \in C(\mathbb{R}^d, \mathbb{R})$ *and let* $u_d \in C^{1,2}([0,T] \times \mathbb{R}^d, \mathbb{R})$ *be an at most poly. growing solution of*

$$\left(\tfrac{\partial u_d}{\partial t}\right)(t,x) + \sum_{i=1}^{d} \tfrac{|\beta_{d,i}|^2 |x_i|^2}{2} \left(\tfrac{\partial^2 u_d}{\partial (x_i)^2}\right)(t,x) + \alpha_{d,i} x_i \left(\tfrac{\partial u_d}{\partial x_i}\right)(t,x) = 0,$$

$$u_d(T,x) = g_d(x),$$

for $(t,x) \in (0,T) \times \mathbb{R}^d$, *and let* $(\phi_{d,\delta})_{d\in\mathbb{N},\,\delta\in(0,1]} \subseteq \mathcal{N}$ *satisfy* $\forall\, d \in \mathbb{N}, \delta \in (0,1]$, $x \in \mathbb{R}^d$ *that* $\mathcal{R}(\phi_{d,\delta}) \in C(\mathbb{R}^d, \mathbb{R})$, $|(\mathcal{R}(\phi_{d,\delta}))(x)| \leq \kappa\, d^\kappa (1 + \|x\|_{\mathbb{R}^d}^\kappa)$, $\mathcal{P}(\phi_{d,\delta}) \leq \kappa\, d^\kappa \delta^{-\kappa}$, *and* $|g_d(x) - (\mathcal{R}(\phi_{d,\delta}))(x)| \leq \kappa\, d^\kappa\, \delta\, (1 + \|x\|_{\mathbb{R}^d}^\kappa)$.
Then $\exists\, c > 0$, $(\psi_{d,\varepsilon})_{d\in\mathbb{N},\,\varepsilon\in(0,1]} \subseteq \mathcal{N}$ *such that* $\forall\, d \in \mathbb{N}, \varepsilon \in (0,1]$:
$\mathcal{R}(\psi_{d,\varepsilon}) \in C(\mathbb{R}^d, \mathbb{R})$, $\mathcal{P}(\psi_{d,\varepsilon}) \leq c\, d^c\, \varepsilon^{-c}$, *and*

$$\left[\int_{[0,1]^d} |u_d(0,x) - (\mathcal{R}(\psi_{d,\varepsilon}))(x)|^2 \, dx\right]^{1/2} \leq \varepsilon.$$

Example: $g_d(x_1, \ldots, x_d) = \max\left\{\frac{x_1 + \ldots + x_d}{d} - K, 0\right\}$ for some $K > 0$.

Main idea and comments about the proof

- Statement purely *deterministic*. Proof based on *probabilistic arguments*.
- Fix $d \in \mathbb{N}$, $\varepsilon \in (0,1]$. Goal: find $\psi \in \mathcal{N}$ such that $\mathcal{P}(\psi) \leq c\, d^c\, \varepsilon^{-c}$ and $\left[\int_{[0,1]^d} |u_d(0,x) - (\mathcal{R}(\psi))(x)|^2 \, dx\right]^{1/2} \leq \varepsilon$.
- Main steps:
  - Construct random field $\mathcal{X} \colon \Omega \times \mathbb{R}^d \to \mathbb{R}$ with $\forall\, x \in \mathbb{R}^d \colon \mathcal{X}(x) \colon \Omega \to \mathbb{R} \sim$ Monte Carlo type approx of $u_d(0,x)$.
  - $\forall\, \omega \in \Omega \colon \mathcal{X}(\omega, \cdot) = \mathcal{R}(\Psi_\omega)$ is realization of ANN $\Psi_\omega \in \mathcal{N}$.
  - $\mathcal{X}$ is close to $u_d(0,\cdot)$: $\mathbb{E}\left[\left[\int_{[0,1]^d} |u_d(0,x) - \mathcal{X}(x)|^2 \, dx\right]^{1/2}\right] \leq \varepsilon$.
  - $\Rightarrow \exists\, \boldsymbol{\omega} \in \Omega \colon \left[\int_{[0,1]^d} |u_d(0,x) - \underbrace{\mathcal{X}(\boldsymbol{\omega}, x)}_{(\mathcal{R}(\Psi_{\boldsymbol{\omega}}))(x)} |^2 \, dx\right]^{1/2} \leq \varepsilon$.
  - count parameters $\mathcal{P}(\Psi_{\boldsymbol{\omega}}) \leq c\, d^c\, \varepsilon^{-c}$.
  - define $\psi = \Psi_{\boldsymbol{\omega}}$.

Nonlinear equations?

Let $T > 0$, $\alpha, \beta \in \mathbb{R}$, $f\colon \mathbb{R} \to \mathbb{R}$ Lip. cont., $\forall\, d \in \mathbb{N}$ let $g_d \in C(\mathbb{R}^d, \mathbb{R})$ and let $u_d \in C([0, T] \times \mathbb{R}^d, \mathbb{R})$ an at most poly. growing visc. solution of

$$\tfrac{\partial u_d}{\partial t} + \left[\sum_{i=1}^{d} \tfrac{|\beta|^2 |x_i|^2}{2} \left(\tfrac{\partial^2 u_d}{\partial (x_i)^2}\right) + \alpha x_i \left(\tfrac{\partial u_d}{\partial x_i}\right)\right] + f(u_d) = 0,$$
$$u_d(T, x) = g_d(x),$$

for $(t, x) \in (0, T) \times \mathbb{R}^d$.

Application: e.g. pricing models with default risk (cf., e.g., Henry-Labordere 2012)

Full-history Recursive Multilevel Picard Algorithm (MLP)

(Hutzenthaler, Jentzen, Kruse, Nguyen, vW 2018)

Let $\kappa > 0$, $\Theta = \cup_{n=1}^{\infty} \mathbb{Z}^n$, $\forall\, d \in \mathbb{N}, x \in \mathbb{R}^d$ assume $|g_d(x)| \leq \kappa d^{\kappa}(1 + \|x\|_{\mathbb{R}^d}^{\kappa})$, let $(\Omega, \mathcal{F}, \mathbb{P})$ prob. space, let $W^{d,\theta} \colon [0,T] \times \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, $d \in \mathbb{N}$, be i.i.d. Brownian motion, let $R^{\theta} \colon [0,T] \times \Omega \to [0,T]$, $\theta \in \Theta$, be i.i.d. cont. satisfying $\forall\, t \in [0,T], \theta \in \Theta \colon\ R_t^{\theta}$ is $\mathcal{U}_{[t,T]}$-distributed, assume $(W^{d,\theta})_{d \in \mathbb{N}, \theta \in \Theta}$ and $(R^{\theta})_{\theta \in \Theta}$ are indep., $\forall\, d \in \mathbb{N}, \theta \in \Theta, t \in [0,T], s \in [t,T], x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ let $X_{t,s}^{d,\theta,x} \colon \Omega \to \mathbb{R}^d$ satisfy

$$X_{t,s}^{d,\theta,x} = \Big[x_i \exp\Big(\big(\alpha - \tfrac{\beta^2}{2}\big)(s-t) + \beta\big(W_s^{d,\theta,i} - W_t^{d,\theta,i}\big)\Big)\Big]_{i \in \{1,\ldots,d\}},$$

and let $V_{M,n}^{d,\theta} \colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$, $M, n \in \mathbb{N}_0$, $\theta \in \Theta$, $d \in \mathbb{N}$, satisfy $\forall\, d, M, n \in \mathbb{N}, \theta \in \Theta, t \in [0,T], x \in \mathbb{R}^d$ that $V_{M,0}^{d,\theta}(t,x) = 0$ and

$$V_{M,n}^{d,\theta}(t,x) = \sum_{k=0}^{n-1} \frac{(T-t)}{M^{n-k}} \Bigg[\sum_{m=1}^{M^{n-k}} f\Big(V_{M,k}^{d,(\theta,k,m)}\big(R_t^{(\theta,k,m)}, X_{t,R_t^{(\theta,k,m)}}^{d,(\theta,k,m),x}\big)\Big) - \mathbb{1}_{\mathbb{N}}(k) f\Big(V_{M,k-1}^{d,(\theta,k,-m)}\big(R_t^{(\theta,k,m)}, X_{t,R_t^{(\theta,k,m)}}^{d,(\theta,k,m),x}\big)\Big)\Bigg] + \Bigg[\sum_{m=1}^{M^n} \frac{g_d(X_{t,T}^{d,(\theta,n,-m),x})}{M^n}\Bigg].$$

Then $\exists\, c > 0$, $N = (N_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,1]} \colon \mathbb{N} \times (0,1] \to \mathbb{N}$ such that $\forall\, d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $x \in [0,1]^d$:

$$\big(\mathbb{E}\big[|u_d(0,x) - V^{d,0}_{N_{d,\varepsilon},N_{d,\varepsilon}}(0,x)|^2\big]\big)^{1/2} \leq \varepsilon$$

and

$$\mathrm{Cost}\big(V^{d,0}_{N_{d,\varepsilon},N_{d,\varepsilon}}(0,x)\big) \leq c\, d^c \varepsilon^{-c}.$$

Corresponding approximation result for DNNs and semilinear PDEs:

- Hutzenthaler, Jentzen, Kruse, Nguyen, *A proof that rectified deep neural networks overcome the curse of dimensionality in the numerical approximation of semilinear heat equations. arXiv 2019.*

DNNs for Kolmogorov PDEs with nonlinear drift: Jentzen, Salimova, Welti 2018.

Recap

- Deep learning methods for high-dim. PDEs seem to overcome curse of dim.
- A theoretical explanation: ANNs have capacity to overcome curse of dim.
- Proof method: Stochastic algorithm overcoming curse of dim. $\Rightarrow$ ANN result
- Recent breakthrough: MLP algorithm for semilinear PDEs

Thank you!

Appendix

- Construction of random ANN for proof of approximation result
- Approximation result for nonlinear PDEs
- Motivation for the multilevel Picard algorithm
- Precise complexity for the multilevel Picard algorithm
- DNN Approximation result for nonlinear drift coefficient

Appendix: Construction of random ANN for proof of approximation result

Let $W^k\colon [0,T]\times\Omega\to\mathbb{R}^d$, $k\in\mathbb{N}$, be independent Brownian motions, let $X^k\colon \mathbb{R}^d\times\Omega\to\mathbb{R}^d$, $k\in\mathbb{N}$, be the random fields which satisfy $\forall\, x=(x_1,\ldots,x_d)\in\mathbb{R}^d$:

$$X^k(x)=\Big[x_i\exp\Big(\Big(\alpha_{d,i}-\tfrac{|\beta_{d,i}|^2}{2}\Big)T+\beta_{d,i}W_T^{d,\theta,i}\Big)\Big]_{i\in\{1,\ldots,d\}}.$$

Observe that $\forall\, x\in\mathbb{R}^d\colon\ u_d(0,x)=\mathbb{E}[g_d(X^1(x))]$.

We define $\mathcal{X}\colon \mathbb{R}^d\times\Omega\to\mathbb{R}$ by

$$\forall\, x\in\mathbb{R}^d\colon\qquad \mathcal{X}(x)=\frac{1}{M}\sum_{k=1}^{M}(\mathcal{R}(\phi_{d,\delta}))(X^k(x)),$$

for suitable $M\in\mathbb{N}$, $\delta\in(0,1]$.

Appendix: Approximation result for nonlinear PDEs

Theorem (Hutzenthaler, Jentzen, Kruse, Nguyen)

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

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

and let $(\phi_{d,\varepsilon})_{d\in\mathbb{N},\, \varepsilon\in(0,1]} \subseteq \mathcal{N}$ *satisfy* $\forall\, d \in \mathbb{N}, \varepsilon \in (0,1], x \in \mathbb{R}^d$: $\mathcal{R}(\phi_{d,\varepsilon}) \in C(\mathbb{R}^d, \mathbb{R})$, $\mathcal{P}(\phi_{d,\varepsilon}) \leq \kappa d^{\kappa} \varepsilon^{-\kappa}$, $|(\mathcal{R}(\phi_{d,\varepsilon}))(x)| \leq \kappa d^{\kappa}(1 + \|x\|^{\kappa})$, *and*

$$|g_d(x) - (\mathcal{R}(\phi_{d,\varepsilon}))(x)| \leq \varepsilon \kappa d^{\kappa}(1 + \|x\|_{\mathbb{R}^d}^{\kappa}).$$

Then $\exists\, c \in (0,\infty)$, $(\psi_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,1]} \subseteq \mathcal{N}$ *such that* $\forall\, d \in \mathbb{N}, \varepsilon \in (0,1]$: $\mathcal{R}(\psi_{d,\varepsilon}) \in C(\mathbb{R}^d, \mathbb{R})$, $\mathcal{P}(\psi_{d,\varepsilon}) \leq c\, d^c \varepsilon^{-c}$, *and*

$$\left[\int_{[0,1]^d} |u_d(0,x) - (\mathcal{R}(\psi_{d,\varepsilon}))(x)|^p \, dx\right]^{1/p} \leq \varepsilon.$$

Appendix: Motivation for the multilevel Picard algorithm

The Feynman-Kac formula ensures that $u_d \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ satisfies the fixed point equation $\forall\, t \in [0,T], x \in \mathbb{R}^d$:

$$\begin{aligned} u_d(t,x) &= \mathbb{E}\Big[g_d\big(X_{t,T}^{d,0,x}\big) + \int_t^T f\big(u_d(r, X_{t,r}^{d,0,x})\big)\, dr \Big] \\ &= \mathbb{E}\Big[g_d\big(X_{t,T}^{d,0,x}\big) + (T-t) f\big(u_d(R_t^0, X_{t,R_t^0}^{d,0,x})\big)\Big]. \end{aligned} \tag{1}$$

Note that $\forall\, d, M, n \in \mathbb{N}, \theta \in \Theta, t \in [0,T], x \in \mathbb{R}^d$:

$$\begin{aligned} \mathbb{E}\big[V_{M,n}^{d,\theta}(t,x)\big] &= \sum_{k=0}^{n-1} (T-t)\mathbb{E}\Big[f\Big(V_{M,k}^{d,(\theta,1)}\big(R_t^{(\theta,1)}, X_{t,R_t^{(\theta,1)}}^{d,(\theta,1),x}\big)\Big) \\ &\quad - \mathbb{1}_{\mathbb{N}}(k) f\Big(V_{M,k-1}^{d,(\theta,-1)}\big(R_t^{(\theta,1)}, X_{t,R_t^{(\theta,1)}}^{d,(\theta,1),x}\big)\Big)\Big] + \mathbb{E}\big[g_d(X_{t,T}^{d,\theta,x})\big] \\ &= \mathbb{E}\Big[g_d(X_{t,T}^{d,\theta,x}) + (T-t) f\Big(V_{M,n-1}^{d,\theta}\big(R_t^\theta, X_{t,R_t^\theta}^{d,\theta,x}\big)\Big)\Big]. \end{aligned} \tag{2}$$

- Hence $V_{M,n}^{d,\theta} \colon \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}$, $n \in \mathbb{N}_0$, behave, in expectation, like Picard iterations for (1).
- Expectation in (2) is approximated with multilevel Monte Carlo over *full history* of previous approximations.

Appendix: Precise complexity for the multilevel Picard algorithm

Assume the setting of the MLP statement, let $p, \mathfrak{P}, q \in [0, \infty)$, recall that $g_d \in C(\mathbb{R}^d, \mathbb{R})$, $d \in \mathbb{N}$, are the terminal conditions of the considered PDEs, let $\xi_d \in \mathbb{R}^d$, $d \in \mathbb{N}$, and assume that $\sup_{d\in\mathbb{N}, x\in\mathbb{R}^d} \left(\frac{|g_d(x)|}{d^{\mathfrak{P}}(1+\|x\|_{\mathbb{R}^d}^p)} + \frac{\|\xi_d\|_{\mathbb{R}^d}}{d^q} \right) < \infty$. Then $\forall\, \delta > 0$ there exist $c > 0$ and $N\colon \mathbb{N} \times (0, 1] \to \mathbb{N}$ such that $\forall\, d \in \mathbb{N}$, $\varepsilon \in (0, 1]$ it holds that

$$\left(\mathbb{E}\left[\left|u_d(0, \xi_d) - V^{d,0}_{N_{d,\varepsilon}, N_{d,\varepsilon}}(0, \xi_d)\right|^2\right]\right)^{1/2} \leq \varepsilon$$

and

$$\mathrm{Cost}_{d, N_{d,\varepsilon}} \leq c\, d^{1+(\mathfrak{P}+qp)(2+\delta)} \varepsilon^{-(2+\delta)}.$$

Appendix: DNN Approximation result for nonlinear drift coefficient

Theorem (Jentzen, Salimova, & Welti 2018)

Let $T, \kappa, p > 0$, *let* $(a_{d,i,j})_{(i,j)\in\{1,\dots,d\}^2} \in \mathbb{R}^{d\times d}$, $d \in \mathbb{N}$, *be sym. positive semi-def., let* $g_d\colon \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, *and* $f_d\colon \mathbb{R}^d \to \mathbb{R}^d$, $d \in \mathbb{N}$, *be functions,* $\forall\, d \in \mathbb{N}$ *let* $u_d\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$ *be at most poly. growing visc. solution of*

$$(\tfrac{\partial}{\partial t}u_d)(t,x) = (\tfrac{\partial}{\partial x}u_d)(t,x)\, f_d(x) + \sum_{i,j=1}^{d} a_{d,i,j}\,(\tfrac{\partial^2}{\partial x_i \partial x_j}u_d)(t,x), \; u_d(0,x) = g_d(x)$$

for $(t,x) \in (0,T) \times \mathbb{R}^d$, *and let* $(\phi^{m,d}_{\varepsilon})_{m\in\{0,1\},d\in\mathbb{N},\varepsilon\in(0,1]} \subseteq \mathcal{N}$ *satisfy* $\forall\, d \in \mathbb{N}, x, y \in \mathbb{R}^d\colon\; \mathcal{R}(\phi^{0,d}_{\varepsilon}) \in C(\mathbb{R}^d,\mathbb{R}), \mathcal{R}(\phi^{1,d}_{\varepsilon}) \in C(\mathbb{R}^d,\mathbb{R}^d)$,

$$|g_d(x)| + \textstyle\sum_{i,j=1}^{d} |a_{d,i,j}| \le \kappa d^{\kappa}(1 + \|x\|^{\kappa}_{\mathbb{R}^d}),$$

$$\|f_d(x) - f_d(y)\|_{\mathbb{R}^d} \le \kappa\|x-y\|_{\mathbb{R}^d}, \quad \|(\mathcal{R}\phi^{1,d}_{\varepsilon})(x)\|_{\mathbb{R}^d} \le \kappa(d^{\kappa} + \|x\|_{\mathbb{R}^d}),$$

$$|(\mathcal{R}\phi^{0,d}_{\varepsilon})(x) - (\mathcal{R}\phi^{0,d}_{\varepsilon})(y)| \le \kappa d^{\kappa}(1 + \|x\|^{\kappa}_{\mathbb{R}^d} + \|y\|^{\kappa}_{\mathbb{R}^d})\|x-y\|_{\mathbb{R}^d},$$

$$|g_d(x) - (\mathcal{R}\phi^{0,d}_{\varepsilon})(x)| + \|f_d(x) - (\mathcal{R}\phi^{1,d}_{\varepsilon})(x)\|_{\mathbb{R}^d} \le \varepsilon\kappa d^{\kappa}(1 + \|x\|^{\kappa}_{\mathbb{R}^d}),$$

and $\sum_{m=0}^{1} \mathcal{P}(\phi^{m,d}_{\varepsilon}) \le \kappa d^{\kappa}\varepsilon^{-\kappa}$. *Then* $\exists\, c \in (0,\infty)$, $(\psi_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,1]} \subseteq \mathcal{N}$ *such that* $\forall\, d \in \mathbb{N}, \varepsilon \in (0,1]\colon\; \mathcal{R}(\psi_{d,\varepsilon}) \in C(\mathbb{R}^d,\mathbb{R}), \mathcal{P}(\psi_{d,\varepsilon}) \le c\, d^c \varepsilon^{-c}$, *and*

$$\left[\int_{[0,1]^d} |u_d(0,x) - (\mathcal{R}(\psi_{d,\varepsilon}))(x)|^p \, dy\right]^{1/p} \le \varepsilon.$$