Deep Orthogonal Decomposition for surrogate modelling of time dependent PDEs

written and submitted by Dawid Kotowski

A thesis presented for the degree of
Master of Science

Supervisor: Prof. Dr. Mario Ohlberger
Faculty of Mathematics and Computer Science
University Münster
Germany, DE

submitted on June 29, 2025

Contents

Chapter 1

Introduction

Chapter 2

Theoretical Background

2.1 for PDE Solution Spaces

For this section we define $u(\mu_j, t_k) \in \mathbb{R}^{N_h}$ for $\mu_j \in \mathcal{P}$ and $t_k \in \mathcal{T}$ to be some vectors, later representing the solutions to a parametric PDE for a given parameter and time point in a given discrete subspace of the solution manifold, sampled uniformly and iid from their respective spaces. Here $\mathcal{P}$ represents and arbitrary compact subspace of $\mathbb{R}^{p+q}$, where $p, q > 0$ and $\mathcal{T} = [0, T]$ for some $T > 0$. Further we let $I \subset \mathbb{N}$ define some index set. Additionally we let $N_s := |I|$ and N_t denote the number of time samples t_k.

Here we will discuss prerequisites to the Proper Orthogonal Decomposition and related concepts, such as the Kolmogorov n-width, which is closely related to the partial sums of the eigenvalues of the POD matrix. This is among the results, this section will cover. We base this chapter on [QMN15].

Definition 2.1.1 (Proper Orthogonal Decomposition)**.** Assuming $N_{data} = N_s N_t < \infty$, we define by $U_j = [u(\mu_j, t_k)]_{k=1}^{N_t} \in \mathbb{R}^{N_h \times N_t}$ the snapshot matrix for a given $j \in I$ and consequently $U = [U_j]_{j \in I} \in \mathbb{R}^{N_h \times N_{data}}$ the collective snapshot matrix. Then the (discrete) correlation matrix is given by

$$K = \frac{|\mathcal{P} \times \mathcal{T}|}{N_{data}} U U^T \in \mathbb{R}^{N_h \times N_h}. \tag{2.1.1}$$

K is symmetric and semi-positive definite and thus has real and positive eigenvalues $\sigma_1^2 \geq \cdots \geq \sigma_r^2$. We further call ψ_k the eigenvectors of the correlation matrix K. With these, the corresponding POD basis of dimension $N < N_h$ is given via

$$\xi_k = \frac{1}{\sigma_k} K \psi_k, \quad 1 \leq k \leq N, \tag{2.1.2}$$

for the N largest eigenvalues and their corresponding eigenvectors. Further, if $N_s, N_t = \infty$, we define the continuous counterpart to the correlation matrix via

$$K_\infty = \int_{\mathcal{P} \times \mathcal{T}} u(\mu, t) u^T(\mu, t) d(\mu, t) \in \mathbb{R}^{N_h \times N_h}, \tag{2.1.3}$$

and its real eigenvalues by $\sigma_{k,\infty}^2$.

Proposition 2.1.2. Let K and K_∞ and their respective eigenvalues be defined as in (2.1.1), then letting $N_s, N_t \to \infty$, we get

$$\sum_{k>N} \sigma_k^2 \longrightarrow \sum_{k>N} \sigma_{k,\infty}^2, \qquad \text{a.s.}$$

Proof.

By some consideration there exists a matrix $V_\infty \in \mathbb{R}^{N_h \times N_h}$ such that

$$\begin{aligned} \sum_{k>N} \sigma_{k,\infty}^2 &= \int_{\mathcal{P} \times \mathcal{T}} \| u(\mu, t) - V_\infty V_\infty^T u(\mu, t) \|^2 d(\mu, t) \\ &= \min_{W \in \mathbb{R}^{N_h \times N} : W^T W = I} \int_{\mu \times \mathcal{T}} \| u(\mu, t) - W W^T u(\mu, t) \|^2 d(\mu, t). \end{aligned}$$

Thus, since this is the optimal choice of an approximation matrix for the continuous case and $V \in \mathbb{R}^{N_h \times N}$ is only optimal in the discrete setting and not necessarily the continuous case, we have

$$\sum_{k>N} \sigma_{k,\infty}^2 \leq \int_{\mathcal{P}\times\mathcal{T}} \| u(\mu,t) - VV^T u(\mu,t) \|^2 d(\mu,t) = \sum_{k>N} \sigma_k^2. \tag{2.1.4}$$

To properly handle the notion of $N_s, N_t \to \infty$ in a regularized fashion, we introduce some arbitrary samples $I_{train} \subset I$ and $\mathcal{T}_{train} \subset \mathcal{T}$, such that I_{train} and $\mathcal{T}_{train}$ are finite. Note, that in general this does rely on the axiom of choice in the theoretical setting, but since the test samples are already known in practice, this is rather an unnecessary remark. We then write entry-wise for $m, l = 1, \dots N_h$

$$[K_\infty - K]_{ml} = \int_{\mathcal{P}\times\mathcal{T}} [u(\mu,t)]_m [u(\mu,t)]_l d(\mu,t) - \frac{|\mathcal{P}\times\mathcal{T}|}{N_{data}} \sum_{j\in I_{train}} \sum_{k\in\mathcal{T}_{train}} [u(\mu_j,t_k)]_m [u(\mu_j,t_k)]_l.$$

By the uniform iid sampling assumed in the introduction to Section (2.1), we know that with the entry-wise boundedness of the PDE solutions, the resulting matrix is subject to the strong law of large numbers, and thus converges to zero, given the entry-wise matrix norm $\|\cdot\|_1$ as $N_t, N_s \to \infty$ almost surely.

Now using Bauer-Fike's Theorem with the 1-norm, we have upon ordering, that for each $\sigma_{k,\infty}^2$ there is a σ_k^2 in the spectrum of K, such that

$$|\sigma_{k,\infty}^2 - \sigma_k^2| \leq \kappa(V) \| K_\infty - K \|_1,$$

where $\kappa(V)$ denotes the condition number of the eigenvector matrix of K.

Now by convergence of $\| K_\infty - K \| \to 0$ almost surely for $N_s, N_t \to \infty$, we conclude the proof. □

Proposition 2.1.3. Let $\mathcal{V}_N = \{W \in \mathbb{R}^{N_h \times N} \,|\, W^T W = I_N\}$ be the set of all N-dimensional orthonormal bases, $N_{data} := N_t N_s$ and $V = [\xi]_{i=1}^N$ be the POD matrix, then

$$\begin{aligned} &\frac{|\mathcal{P}\times\mathcal{T}|}{N_{data}} \sum_{i\in I, k=1,\dots,N_t} \| u(\mu_j,t_k) - VV^T u(\mu_j,t_k) \|_2^2 \\ &\qquad = \min_{W\in\mathcal{V}_N} \frac{|\mathcal{P}\times\mathcal{T}|}{N_{data}} \sum_{i\in I, k=1,\dots,N_t} \| u(\mu_j,t_k) - WW^T u(\mu_j,t_k) \|_2^2 = \sum_{j=N+1}^{r} \sigma_k^2. \end{aligned}$$

Definition 2.1.4 (Linear Kolmogorov N-width)**.** Let $\mathcal{S}_{N_h} = \{u(\mu,t) \in \mathbb{R}^{N_h} \,|\, (\mu,t) \in \mathcal{P}\times\mathcal{T}\}$ be some manifold. The linear Kolmogorov N-width of $\mathcal{S}_{N_h}$ is defined as

$$d_N(S_{N_h}) = \inf_{V\in\mathbb{R}^{N_h\times N}} \sup_{u\in S_{N_h}} \| u - VV^T u \|.$$

Definition 2.1.5 (Nonlinear Kolmogorov n-width)**.** The nonlinear Kolmogorov n-width of the reduced manifold $\mathcal{S}_N = \{q(\mu,t) = V^T u(\mu,t) \,|\, (\mu,t) \in \mathcal{P}\times\mathcal{T}\}$ is defined as

$$\delta_n(\mathcal{S}_N) = \inf_{\psi\in C(\mathbb{R}^N,\mathbb{R}^n),\psi'\in C(\mathbb{R}^n,\mathbb{R}^N)} \sup_{u\in S_{N_h}} \| u - \psi(\psi'(u)) \|.$$

2.2 for Neural Networks

Chapter 3

Formulation of the General Problem

3.1 Stationary parametric PDE

3.2 Deep-learning based Reduced Order Models

3.3 Non-stationary parametric PDE

Let $\Theta \subset \mathbb{R}^p$ and $\Theta' \subset \mathbb{R}^q$ denote the geometric and physical parameter spaces respectively. Further we let $\Omega \subset \mathbb{R}^d$ be some Lipschitz domain and $\Gamma = [0, T]$ for some $T > 0$ be the trial space of a parametric time-dependent PDE, i.e. $u(t; \mu, \nu) = u(x, t; \mu, \nu)$ is the solution of the PDE, if

$$\begin{cases} \partial_t u(t;\mu,\nu) + \mathcal{L}(\mu,\nu)u(t;\mu,\nu) + \mathcal{N}(u(\mu,\nu),\mu,\nu) = f(t;\mu,\nu), & (x,t) \in \Omega \times \Gamma, \\ \mathcal{B}(\mu,\nu)u(t;\mu,\nu) = g_N(t;\mu,\nu) & (x,t) \in \partial\Omega \times \Gamma, \\ u(t;\mu,\nu) = u_0(\mu,\nu) & (x,t) \in \Omega \times \{0\}. \end{cases}$$

Assume we half-discretize the problem in the spacial domain Ω with N_h degrees of freedom. Note that we refer to this discrete grid space on Ω as Ω^h and its boundary grid points as $\partial\Omega^h$. Then the resulting full order model can be seen as the solution to

$$\begin{cases} M(\mu,\nu)\partial_t u_h(t;\mu,\nu) + A(\mu,\nu)u_h(t;\mu,\nu) + N(u_h(\mu,\nu),\mu,\nu) = f(t;\mu,\nu) & (x,t) \in \Omega^h \times \Gamma, \\ u_h(0;\mu,\nu) = u_0(\mu,\nu) & (x,t) \in \Omega^h \times \{0\}, \end{cases}$$

where we assume $M : \Theta \times \Theta' \to \mathbb{R}^{N_h \times N_h}$ to send the parameter set to a mass matrix, $A : \Theta \times \Theta' \to \mathbb{R}^{N_h \times N_h}$ to a symmetric positive definite stiffness matrix, $N : \mathbb{R}^{N_h} \to \mathbb{R}^{N_h}$ to be some non-linear term and $f : \Gamma \times \mathbb{R}^{N_h} \times \Theta \times \Theta' \to \mathbb{R}^{N_h}$ to be the corresponding source term of the equation. Additionally we assume to have initial conditions via the function $u_0 : \Theta \times \Theta' \to \Omega^{N_h}$.

Assumption 1. Let $\mathcal{S} := \{u(\mu,\nu,t) \,|\, (\mu,\nu,t) \in \Theta \times \Theta' \times \mathcal{T}\}$ denote the total solution manifold and $\mathcal{S}_{\mu,t} := \{u(\mu,\nu,t) \,|\, \nu \in \Theta'\}$ denote the solution submanifold for a fixed set $(\mu,t) \in \Theta \times \mathcal{T}$, then we assume that

1. the linear KnW of $\mathcal{S}$ decays slowly, i.e.

$$d_N(\mathcal{S}) \leq CN^{-\alpha},$$

   where $0 < \alpha \leq 1$ and $C > 0$ some constant.

2. the linear KnW of $\mathcal{S}_{\mu,t}$ decays quickly for all $\mu \in \Theta$ and $t \in \mathcal{T}$, i.e.

$$\sup_{(\mu,t)\in\Theta\times\mathcal{T}} d_N(\mathcal{S}_{\mu,t}) \leq C' N^{-\beta},$$

   where $\beta > 1$ and $C' > 0$ some constant.

3.4 Deep-learning based Reduced Order Models

A general Deep learning based reduced order model (DL-ROM) firstly appeared in architecture and training in the paper [FDM20]; to properly approximate the potentially non-linear solution manifold of the high fidelity FOM, we employ an Autoencoder (AE), and on the reduced manifold we employ some other Deep Learning architecture, like the proposed Deep Feed-Forward Neural Network (DFNN), to learn the latent dynamics of the parametric PDE. This concludes in the following approximation of the FOM solution u_h

$$\hat{u}_h(t;\mu,\nu) = \Psi_h(\Phi_n(t;\mu,\nu,\theta_{DF}),\theta_D) \in \mathbb{R}^{N_h}, \tag{3.4.1}$$

where we set $\Psi_h = f_h^D : \mathbb{R}^n \times \Theta_D \to \mathbb{R}^{N_h}$ for the decoder (and $f_n^E : \mathbb{R}^{N_h} \times \Theta_E \to \mathbb{R}^n$ the corresponding encoder) of a convolutional AE. Then f_n^E denotes the projection onto the reduced trial solution manifold and we assume $\Phi_n = \Phi_n^{DF} : \Gamma \times \Theta \times \Theta' \times \Theta_{DF} \to \mathbb{R}^n$ to be the DFNN mapping the dynamics of the parametric PDE onto the reduced trial solution manifold.

The subsequent training is then performed via a loss function of a convex combination of both, the projection error using the encoder and the dynamics error considering both the decoder and the feed forward network, i.e. for $\theta = (\theta_D, \theta_E, \theta_{DF})$ define the optimal weights via

$$\theta^* = \arg\min_{\theta} \frac{1}{N_s} \sum_{i,j=1}^{N_{train}} \sum_{k=1}^{N_t} \frac{\omega_h}{2} \parallel u_h(t^k;\mu_i,\nu_j) - f_h^D(\Phi_n^{DF}(t^k;\mu_i,\nu_j,\theta_{DF}),\theta_D),\theta_D) \parallel^2 \tag{3.4.2}$$

$$+ \frac{1-\omega_h}{2} \parallel f_n^E(u_h(t^k;\mu_i,\nu_j),\theta_E) - \Phi_n^{DF}(t^k;\mu_i,\nu_j,\theta_{DF}) \parallel^2. \tag{3.4.3}$$

3.4.1 Deep Orthogonal Decomposition based Reduction

For comparison, one can see, that the POD-DL-ROM works significantly more efficient than the DL-ROM by pre-reducing the dimension N_h to some $N < N_h$ in a linear POD-based fashion. This has the disadvantage of potentially losing fidelity, if the Kolmogorov N-width is rather large, which is the case for non-linear PDEs, relying on geometric circumstances. For this, we might instead of using snapshot-based rPOD for the matrix S (optionally the correlation matrix) rather directly employ a Deep Orthogonal Decomposition to find the reduction matrix in dependency of $\mu \in \Theta$ and $t \in \Gamma$.

Thus we have a common denominator for the DOD-DL-ROM approach. Defining $V : \Gamma \times \Theta \times \Theta_{DOD} \to \mathbb{R}^{N_h \times N}; V(t;\mu,\theta_{DOD}) \mapsto \mathbb{A}\tilde{V}_{N_A}(t;\mu,\theta_{DOD})$, where $\mathbb{A} \in \mathbb{R}^{N_h \times N_A}$ is the pre-reduction POD for $N < N_A < N_h$ employed to deal with exceptionally large inherent dimensions N_h and $\tilde{V}_{N_A}(t;\mu,\theta_{DOD}) \in \mathbb{R}^{N_A \times N}$ is called *inner module* of the DOD, yields the foundation of the reduction consideration. From hereon, there are a couple of different approaches, which lead to similar ideas; one is able to, after learning the DOD-map, either suppose, the remaining reduced solution manifold is "linear" enough to work directly with a coefficient finder network (Option A) or a further non-linear reduction, potentially by an Autoencoder, is needed (Option B). As an additional, maybe more ambiguous regarding error estimates, we want to introduce a CoLoRA-based ROM, which first tries to use the DOD-NN reductor for the stationary equivalent of the problem (if existent) and then employs a CoLoRA architecture to reduce the dimension up to the dynamics of the system (Option C).

For Options A and B: To use the same loss function as in the (POD-)DL-ROM, we do need to guarantee a previous construction of the inner module of the DOD. This first step is done with a separate training with a time dependent analogue to the DOD, i.e.

$$\theta^*_{DOD} = \arg\min_{\theta_{DOD}} \frac{1}{N_s} \sum_{i,j=1}^{N_{train}} \sum_{k=1}^{N_t} \parallel u_h(t^k;\mu_i,\nu_j) - V(t^k;\mu_i,\theta_{DOD})V^T(t^k;\mu_i,\theta_{DOD})\mathbb{G}u_h(t^k;\mu_i,\nu_j) \parallel^2. \tag{3.4.4}$$

3.4.2 DOD+DNN

Let $\Phi_n^{DLNN} : \Gamma \times \Theta \times \Theta' \times \Theta_{DLNN} \to \mathbb{R}^N$ be some Deep Feed-Forward Neural Network, then the overall workflow can be summed up by

$$\hat{u}_h(t;\mu,\nu,\theta_{DOD},\theta_{DFNN}) = \mathbb{A}\tilde{V}_{N_A}(t;\mu,\theta_{DOD}) \cdot \Phi_n^{DFNN}(t;\mu,\nu,\theta_{DFNN}). \tag{3.4.5}$$

This is a rather obvious extension of [Fra+24], but works surprisingly well. It will be the basic benchmark to overcome for the following algorithms. As for training a general comparison between the projection of a true solution and the prediction of the Feed Forward Network works well, i.e. for $\theta = \theta_{DFNN}$ let the optimal weights be given by

$$\theta^* = \arg\min_{\theta} \frac{1}{N_s} \sum_{i,j=1}^{N_{train}} \sum_{k=1}^{N_t} \| V^T(t^k;\mu_i,\theta^*_{DOD}) u_h(t^k;\mu_i,\nu_j) - \Phi_n^{DFNN}(t^k;\mu_i,\nu_j,\theta_{DFNN}) \|. \tag{3.4.6}$$

3.4.3 DOD-DL-ROM

Let f_N^D and f_n^E be the decoder and encoder of an AE respectively as above. Each of these non-linear DOD-DL-ROMs is subject to further dimensional reduction and thus is not comparable directly to the linear method above. Nonetheless in terms of reduction in dimensionality both methods can be comparable, if the latent dynamics dimension n is the same in both cases. The general workflow is given by

$$\hat{u}_h(t;\mu,\nu,\theta_{DOD},\theta_D,\theta_{DFNN}) = \mathbb{A}\tilde{V}_{N_A}(t;\mu,\theta_{DOD}) \cdot f_N^D(\Phi_n^{DFNN}(t;\mu,\nu,\theta_{DFNN}),\theta_D). \tag{3.4.7}$$

Since for this, we employ an AE, we optimize for two loss functions, since the employed decoder is tied strictly to its respective encoder being able to correctly project onto the low-dimensional subspace of the latent dynamics of the PDE. For $\theta = (\theta_D, \theta_E, \theta_{DFNN})$ define the optimal weights via

$$\begin{aligned}\theta^* = \arg\min_{\theta} \frac{1}{N_s} \sum_{i,j=1}^{N_{train}} \sum_{k=1}^{N_t} \\ &\frac{\omega_h}{2} \| u_h(t^k;\mu_i,\nu_j) - V(t^k;\mu_i,\theta^*_{DOD}) \cdot f_h^D(\Phi_n^{DFNN}(t^k;\mu_i,\nu_j,\theta_{DFNN}),\theta_D),\theta_D) \|^2 \\ &+ \frac{1-\omega_h}{2} \| V^T(t^k;\mu_i,\theta^*_{DOD}) \cdot f_n^E(u_h(t^k;\mu_i,\nu_j),\theta_E) - \Phi_n^{DFNN}(t^k;\mu_i,\nu_j,\theta_{DFNN}) \|^2.\end{aligned}$$

3.4.4 CoLoRA-DL-ROM

Here we aim at minimizing the dimension through two major reductions; the first being the reduction of the stationary DOD to the latent dimension of the μ-invariant solution manifold of the stationary equivalent to the problem (regarding the Dirichlet boundary conditions as the initial conditions on some a priori defined boundary), and the second concerning the actual reduction with CoLoRA-layered architecture to the latent dynamic dimension $n << N_h$. The general workflow is thus given by

$$\hat{u}_h(t;\mu,\nu,\theta^*_{sDOD},\theta_C) = \mathbb{A} \cdot \mathcal{C}_L(\ldots(\mathcal{C}_2(\mathcal{C}_1(V_0(\mu,\theta^*_{sDOD}) \cdot \phi_0(\mu,\nu,\theta^*_{sDOD})),\theta_{C_1}),\theta_{C_2}),\theta_{C_L}), \tag{3.4.8}$$

where $\mathcal{C}_i : \mathbb{R}^{N_A} \times \Theta_C \to \mathbb{R}^{N_A}; \mathcal{C}_i(x,\theta_{C_i}) = W_i x + A_i \operatorname{diag}(\alpha_i(\nu,t)) B_i x + b_i$ for all $1 \le i \le L$. Remark, that we will still employ a pre-reduction using the rPOD matrix $\mathbb{A}$ stemming from the same snapshot matrix S as in both Options before.

Chapter 4

Error estimates

The main goal of all error estimates can be generally hindered by two rather obvious circumstances: bad sampling and a completely rigid parameter-to-solution map. Thus two main assumptions will be assumed throughout the following section.

Assumption 2. Let $p, q > 0$, assume that $\Theta \subset \mathbb{R}^p$ and $\Theta' \subset \mathbb{R}^q$ are compact and denote $\mathcal{T} = [0, T]$ for some $T > 0$. We assume that all N_s data snapshots are sampled uniformly and iid in the parameter spaces Θ, Θ', while a uniform grid is employed for the time variable, $t \in \{\Delta t, 2\Delta t, \ldots, N_t \Delta t\}$, where $N_t \in \mathbb{N}_{\geq 2}$, N_{s_1} and N_{s_2} are the respective number of random samples in the parameter spaces and $N_s = N_{s_1} N_{s_2} \in \mathbb{N}$ the total sample size, and $\Delta t = T/N_t$.

We remark, that the equidistant spacing of the time samples is a matter of convenience, and could be loosened under some constraints.

Assumption 3. Let $\mathcal{G} : \Theta \times \Theta' \to \mathbb{R}^{N_h}$ be the parameter-to-solution map, mapping $(\mu, \nu, t) \mapsto u_h(\mu, \nu, t)$. Here we assume that

1. $m = \underset{(\mu,\nu,t) \in \Theta \times \Theta' \times \mathcal{T}}{\text{ess sup}} \| u_h(\mu, \nu, t) \| > 0$, $M = \underset{(\mu,\nu,t) \in \Theta \times \Theta' \times \mathcal{T}}{\text{ess inf}} \| u_h(\mu, \nu, t) \| < \infty$,
2. $\mathcal{G}$ is Lipschitz-continuous with the constant $L > 0$.

Assumption 4. We assume that for s, s' sufficiently large, there are two maps $\psi_* : \mathbb{R}^n \to \mathbb{R}^N$ and $\psi'_* : \mathbb{R}^N \to \mathbb{R}^n$ which attain the perfect nonlinear Kolmogorov n-width of the reduced solution manifold of the problem, as given in (2.1.5) for any $n \leq 2(p+q)+3$, i.e. for $\mathcal{S}_N := \{q(\mu, \nu, t) = V^T u(\mu, \nu, t) \,|\, (\mu, \nu, t) \in \Theta \times \Theta' \times \mathcal{T}\}$, where $V \in \mathbb{R}^{N_h \times N}$ is the POD matrix,

$$\delta_n(\mathcal{S}_N) = 0.$$

Proposition 4.0.1. Let $f \in L^2(\Theta \times \Theta' \times \mathcal{T})$. Under Assumption (2), it is true, that

$$\mathbb{E}\left[\left| \int_{\Theta \times \Theta' \times \mathcal{T}} f(\mu, \nu, t) d(\mu, \nu, t) - \frac{|\Theta \times \Theta' \times \mathcal{T}|}{N_{data}} \sum_{i=1}^{N_{s_1}} \sum_{j=1}^{N_{s_2}} \sum_{k=1}^{N_t} f(\mu_i, \nu_j, t_k) \right|\right] \leq \mathcal{O}(N_s^{-1/2} + N_t^{-1}).$$

4.1 for POD+DNNs

Theorem 4.1.1. *Let $\mathcal{G} : (\mu, \nu, t) \mapsto u_h(\mu, \nu, t)$ for all $\mu \in \Theta, \nu \in \Theta'$ and $t \in \mathcal{T}$ be the parameter-to-solution map. Here we consider a general POD+DNN structure, thus neglecting the potentially different nature of Θ and Θ', i.e. $\mathcal{G}(\mu, \nu, t) \approx V\hat{q}$, where $\hat{q} : \mathbb{R}^{p+q+1} \to \mathbb{R}^N$ is a neural network trained using data $(\mu_i, \nu_i, t_i)_{i=1}^{N_{data}}$ and V is the POD matrix, as described in (2.1.1). Then under the Assumptions (2) and (3), we have*

$$\mathcal{E}_R \leq \mathcal{E}_S + \mathcal{E}_{POD} + \mathcal{E}_{NN}, \tag{4.1.1}$$

where

1. *The sampling error*
$$\mathcal{E}_S = \mathcal{E}_S\left(\mathcal{G}, (\mu_i, \nu_i, \tau_i)_{i=1}^{N_{data}}, N\right)$$
satisfies $\mathcal{E}_S \to 0$ *as* $N_s, N_t \to \infty$, *with expectation*
$$\mathbb{E}[\mathcal{E}_S] = \mathcal{O}(N_s^{-1/4} + N_t^{-1}).$$

2. *The POD projection error*
$$\mathcal{E}_{POD} = \mathcal{E}_{POD}\left(\mathcal{G}, (\mu_i, \nu_i, \tau_i)_{i=1}^{N_{data}}, N\right)$$
satisfies $\mathcal{E}_{POD} \to \mathcal{E}_{POD,\infty}$ *almost surely as* $N_s, N_t \to \infty$, *where*
$$\mathcal{E}_{POD,\infty} = \mathcal{E}_{POD,\infty}(\mathcal{G}, N)$$
is independent of the data snapshots and proportional to the linear KnW of the manifold $\mathcal{S} = \{u(\mu,\nu,t) \,|\, (\mu,\nu,t) \in \Theta' \times \mathcal{T}\}$, *i.e.* $d_N(\mathcal{S})$.

3. *The neural network approximation error*
$$\mathcal{E}_{NN} = \mathcal{E}_{NN}(\mathcal{G}, N, \hat{q})$$
is arbitrarily small depending on the complexity and structure of the neural network $\hat{q}$.

Proof. With the knowledge of $\|\cdot\|_{L^2_\omega}$ being a norm, we know it must be subject to the triangle inequality, thus

$$\begin{aligned}\mathcal{E}_R &= \left(\int_{\Theta\times\Theta'\times\mathcal{T}} \frac{\| u(\mu,\nu,t) - V\hat{q}(\mu,\nu,t) \|}{\| u(\mu,\nu,t) \|} d(\mu,\nu,t)\right)^{1/2} && (4.1.2)\\ &\leq \| u(\mu,\nu,t) - VV^T u(\mu,\nu,t) \|_{L^2_\omega} + \| VV^T u(\mu,\nu,t) - V\hat{q}(\mu,\nu,t) \|_{L^2_\omega}. && (4.1.3)\end{aligned}$$

Let $q(\mu,\nu,t) := V^T u(\mu,\nu,t)$ be a reduced solution according to the POD basis. Then the natural definition of the neural network approximation error is

$$\mathcal{E}_{NN} = \left(\int_{\Theta\times\Theta'\times\mathcal{T}} \frac{\| Vq(\mu,\nu,t) - V\hat{q}(\mu,\nu,t) \|^2}{\| u(\mu,\nu,t) \|^2} d(\mu,\nu,t)\right)^{1/2}. \tag{4.1.4}$$

This marks the second part in (4.1.3). This is only dependent on the ability of the neural network to identifying the best choice on a given linear sub-manifold of the solution manifold. The first part is concerned with the error generated from the best *possible* approximation of such a linear sub-manifold.

For this, we can bound the norm $\|\cdot\|_{L^2_\omega} \leq m^{-1}\|\cdot\|$, where $m > 0$ is the constant stemming from Assumption (3). Further we remind ourselves of the definition of the discrete correlation matrix $K = |\Theta \times \Theta \times \mathcal{T}| N_{data}^{-1} UU^T$ and its eigenvalues σ_k in (2.1.1). Using $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for all

$a, b \geq 0$, we get

$$\begin{aligned}
&\| u(\mu,\nu,t) - VV^T u(\mu,\nu,t) \|_{L^2_\omega} \\
&\leq m^{-1} \left(\int_{\Theta\times\Theta'\times\mathcal{T}} \| u(\mu,\nu,t) - VV^T u(\mu,\nu,t) \|^2 d(\mu,\nu,t) \right)^{1/2} \\
&\leq m^{-1} \left(\int_{\Theta\times\Theta'\times\mathcal{T}} \| u(\mu,\nu,t) - VV^T u(\mu,\nu,t) \|^2 d(\mu,\nu,t) - \sum_{k>N} \sigma_k^2 + \sum_{k>N} \sigma_k^2 \right)^{1/2} \\
&\leq m^{-1} \left(\left| \int_{\Theta\times\Theta'\times\mathcal{T}} \| u(\mu,\nu,t) - VV^T u(\mu,\nu,t) \|^2 d(\mu,\nu,t) - \sum_{k>N} \sigma_k^2 \right| + \sum_{k>N} \sigma_k^2 \right)^{1/2} \\
&\leq \underbrace{m^{-1} \left| \int_{\Theta\times\Theta'\times\mathcal{T}} \| u(\mu,\nu,t) - VV^T u(\mu,\nu,t) \|^2 d(\mu,\nu,t) - \sum_{k>N} \sigma_k^2 \right|^{1/2}}_{=:\mathcal{E}_S} + \underbrace{m^{-1} \sqrt{\sum_{k>N} \sigma_k^2}}_{=:\mathcal{E}_{POD}}.
\end{aligned}$$

This concludes the estimate up to the facts about the convergence of the errors. Recalling (2.1.3), we can rewrite the sampling error as follows

$$\begin{aligned}
\mathcal{E}_S = m^{-1} \Bigg| \int_{\Theta\times\Theta'\times\mathcal{T}} &\| u(\mu,\nu,t) - VV^T u(\mu,\nu,t) \|^2 \, d(\mu,\nu,t) - \\
&\frac{|\Theta\times\Theta'\times\mathcal{T}|}{N_{data}} \sum_{j=1}^{N_{data}} \| u_j - VV^T u_j \|^2 \Bigg|^{1/2}
\end{aligned}$$

Furthermore, we define using the compactness postulated in Assumption (2) and the boundedness in Assumption (3), the L^2-error map via

$$f(\mu,\nu,t) := \| u(\mu,\nu,t) - VV^T u(\mu,\nu,t) \|^2 \leq M^2 \| I - VV^T \|^2 < \infty,$$

where we justify neglecting the dependence on u in the mapping, since $\mathcal{G}(\mu,\nu,t)$ is assumed to be known and can be projected onto the N_h-dimensional solution manifold. Thus with $f \in L^2(\Theta\times\Theta'\times\mathcal{T})$, we know that by Proposition (4.0.1), $\mathbb{E}[\mathcal{E}_S] \leq \mathcal{O}(N_s^{-1/4} + N_t^{-1/2})$.

By the strong law of large numbers, we can thus conclude, that $\mathcal{E}_S \to 0$ almost surely, as $N_t, N_s \to \infty$, and further by (2.1.2), we have that $\mathcal{E}_{POD} \to \mathcal{E}_{POD,\infty}$ as $N_t, N_s \to \infty$, where

$$\mathcal{E}_{POD,\infty} = m^{-1} \sqrt{\sum_{k>N} \sigma_{k,\infty}^2}. \tag{4.1.5}$$

□

Theorem 4.1.2. *Under the assumptions of Theorem (4.1.1), we have that*

$$\mathcal{E}_R \geq \frac{m}{M} \mathcal{E}_{DOD,\infty},$$

where $\mathcal{E}_{DOD,\infty} := m^{-1} \sqrt{\sum_{k>N} \sigma_{k,\infty}^2}$.

Proof. □

4.2 for DOD+DNNs

Lemma 4.2.1 (Existence of the DOD-Module)**.**
Let $\Theta \subset \mathbb{R}^p$ and $\Theta' \subset \mathbb{R}^q$ be two compact sets. Let

$$\mathcal{F} : \Theta \times \Theta' \to \mathbb{R}^{N_h}; \, (\boldsymbol{\mu}, \boldsymbol{\nu}) \mapsto \boldsymbol{u}_{\boldsymbol{\mu},\boldsymbol{\nu}}$$

be continuous. For each $\boldsymbol{\mu} \in \Theta$, let $\mathcal{S}_{\boldsymbol{\mu}} := \{\boldsymbol{u}_{\boldsymbol{\mu},\boldsymbol{\nu}}\}_{\boldsymbol{\nu}\in\Theta'} \subset \mathcal{F}(\Theta \times \Theta')$ be the $\boldsymbol{\mu}$-submanifold in the image of $\mathcal{F}$. Let $\mathbb{G}$ be the Gram matrix associated with an inner product in $\mathbb{R}^{N_h}$, and let $\| \cdot \|$ be its corresponding norm. Then, for every $\varepsilon > 0$ there is a ReLU matrix-valued deep neural network $\boldsymbol{V} : \mathbb{R}^p \to \mathbb{R}^{N_h \times n}$ such that

$$\mathbb{E}_{\boldsymbol{\mu},\boldsymbol{\nu}} \| \boldsymbol{u}_{\boldsymbol{\mu},\boldsymbol{\nu}} - \boldsymbol{V}_{\boldsymbol{\mu}} \boldsymbol{V}_{\boldsymbol{\mu}}{}^T \mathbb{G} \boldsymbol{u}_{\boldsymbol{\mu},\boldsymbol{\nu}} \| < \varepsilon + \mathbb{E}_{\boldsymbol{\mu}}[d_n(\mathcal{S}_{\boldsymbol{\mu}})]$$

where $\boldsymbol{V}_{\boldsymbol{\mu}} := \boldsymbol{V}(\boldsymbol{\mu})$.

Proof. Let $n \in \mathbb{N}$ be fixed. Assuming we have proven, that $|\mathbb{V}_{ij}| \leq 1$ w.l.o.g., we define $J : \Theta \times [-1,1]^{N_h \times n} \to \mathbb{R}$ as

$$J(\boldsymbol{\mu}, \mathbb{V}) := \mathbb{E}_{\boldsymbol{\nu}} \| \boldsymbol{u}_{\boldsymbol{\mu},\boldsymbol{\nu}} - \mathbb{V}\mathbb{V}^T \mathbb{G} \boldsymbol{u}_{\boldsymbol{\mu},\boldsymbol{\nu}} \|, \tag{4.2.1}$$

and let some $\varepsilon > 0$. Then the continuity of J implies uniform continuity by the compactness of its pre-image, i.e. there is $\delta > 0$ such that

$$|\boldsymbol{\mu} - \boldsymbol{\mu}'| + \| \mathbb{V} - \mathbb{V}' \|_\infty < \delta \implies |J(\boldsymbol{\mu}, \mathbb{V}) - J(\boldsymbol{\mu}', \mathbb{V}')| < \varepsilon.$$

Then there is a Borel measurable map $s : \Theta \to [-1,1]^{N_h \times n}$ given by

$$s : \boldsymbol{\mu} \mapsto \underset{\mathbb{V}\in[-1,1]^{N_h \times n}}{\operatorname{argmin}} J(\boldsymbol{\mu}, \mathbb{V}).$$

This map is bounded (given by continuity of J and compactness of its image) and hence $s \in L^1(\Theta; \mathbb{R}^{N_h \times n})$, i.e. $\| s \|_{L^1(\Theta;\mathbb{R}^{N_h \times n})} := \int_\Theta \| s(\boldsymbol{\mu}) \|_\infty d\mu(\boldsymbol{\mu}) < \infty$. This means that by Hornik's Theorem [Hor91] there is a deep ReLU neural network $\boldsymbol{V}_0 : \Theta \to \mathbb{R}^{N_h \times n}$ such that $\mathbb{E}_{\boldsymbol{\mu}} |\boldsymbol{V}_0(\boldsymbol{\mu}) - s(\boldsymbol{\mu})| < \delta$ for all $\delta > 0$. Now let ρ be a entry-wise ReLU activation function and we consider a three-layered network $L : \mathbb{R}^{N_h \times n} \to \mathbb{R}^{N_h \times n}$ given by

$$L(\boldsymbol{x}) := \boldsymbol{e} - \mathbb{I}_{N_h \times n} \rho(-\mathbb{I}_{N_h \times n} \rho(\mathbb{I}_{N_h \times n} \boldsymbol{x} + \boldsymbol{e}) + 2\boldsymbol{e}),$$

where $\mathbb{I}_{N_h \times n}$ is the identity tensor on $\mathbb{R}^{N_h \times n}$, and $\boldsymbol{e}$ is a $N_h \times n$ matrix with ones as entries. Thus for each entry $x_{i,j}$ of $\boldsymbol{x}$, we have $L(x_{i,j}) = 1 - \rho(2 - \rho(x_{i,j} + 1)) \in [-1,1]$. Thus, since L is 1-Lipschitz and L only effects values outside $[-1,1]$, and hence $L \circ s = s$, we have for $\boldsymbol{V(\mu)} := L \circ \boldsymbol{V}_0(\boldsymbol{\mu})$

$$\mathbb{E}|\boldsymbol{V(\mu)} - s(\boldsymbol{\mu})| = \mathbb{E}|L \circ \boldsymbol{V}_0(\boldsymbol{\mu}) - L \circ s(\boldsymbol{\mu})| \leq \mathbb{E}|\boldsymbol{V}_0(\boldsymbol{\mu}) - s(\boldsymbol{\mu})| < \delta.$$

With this we can take $\boldsymbol{V(\mu)}$ to be in the pre-image of J and therefore conforming to its uniform continuity, i.e. take $\varepsilon > 0$, s.t.

$$|J(\boldsymbol{\mu}, s(\boldsymbol{\mu})) - J(\boldsymbol{\mu}, \boldsymbol{V}(\boldsymbol{\mu}))| < \varepsilon.$$

In total we calculate with $\mathbb{E}_{\boldsymbol{\mu},\boldsymbol{\nu}} \| \boldsymbol{u}_{\boldsymbol{\mu},\boldsymbol{\nu}} - \boldsymbol{V}_{\boldsymbol{\mu}} \boldsymbol{V}_{\boldsymbol{\mu}}{}^T \mathbb{G} \boldsymbol{u}_{\boldsymbol{\mu},\boldsymbol{\nu}} \| = \mathbb{E}_{\boldsymbol{\mu}}[J(\boldsymbol{\mu}, \boldsymbol{V}(\boldsymbol{\mu})],$

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\mu},\boldsymbol{\nu}} \| \boldsymbol{u}_{\boldsymbol{\mu},\boldsymbol{\nu}} - \boldsymbol{V}_{\boldsymbol{\mu}} \boldsymbol{V}_{\boldsymbol{\mu}}{}^T \mathbb{G} \boldsymbol{u}_{\boldsymbol{\mu},\boldsymbol{\nu}} \| &\leq \mathbb{E}_{\boldsymbol{\mu}}[J(\boldsymbol{\mu}, \boldsymbol{V}(\boldsymbol{\mu})) - J(\boldsymbol{\mu}, s(\boldsymbol{\mu}))] + \mathbb{E}_{\boldsymbol{\mu}}[J(\boldsymbol{\mu}, s(\boldsymbol{\mu}))] \\
&< \varepsilon + \mathbb{E}_{\boldsymbol{\mu}} \left[\min_{\mathbb{V}} J(\boldsymbol{\mu}, \mathbb{V}) \right] \\
&\leq \varepsilon + \mathbb{E}_{\boldsymbol{\mu}} \left[\min_{\mathbb{V}} \sup_{\boldsymbol{\nu}} \| \boldsymbol{u}_{\boldsymbol{\mu},\boldsymbol{\nu}} - \mathbb{V}\mathbb{V}^T \mathbb{G} \boldsymbol{u}_{\boldsymbol{\mu},\boldsymbol{\nu}} \| \right]
\end{aligned}$$

□

Theorem 4.2.2. *Let $\mathcal{G} : (\mu, \nu, t) \mapsto u_h(\mu, \nu, t)$ for all $\mu \in \Theta, \nu \in \Theta'$ and $t \in \mathcal{T}$ be the parameter-to-solution map. Here we consider a general DOD+DNN structure, i.e. $\mathcal{G}(\mu, \nu, t) \approx V(\mu, t)\hat{q}(\mu, \nu, t)$, where $\hat{q} : \mathbb{R}^{p+q+1} \to \mathbb{R}^N$ is a neural network trained using data $(\mu_i, \nu_i, t_i)_{i=1}^{N_{data}}$ and $V : \mathbb{R}^{p+1} \to \mathbb{R}^{N_h \times N}$ is the DOD neural network, as described in (3.4.2). Then under the Assumptions (2) and (3), we have*

$$\mathcal{E}_R \leq \mathcal{E}_S + \mathcal{E}_{DOD} + \mathcal{E}_{NN}, \tag{4.2.2}$$

where

1. *The sampling error*
$$\mathcal{E}_S = \mathcal{E}_S \left(\mathcal{G}, (\mu_i, \nu_i, \tau_i)_{i=1}^{N_{data}}, N \right)$$
satisfies $\mathcal{E}_S \to 0$ as $N_s, N_t \to \infty$, with expectation
$$\mathbb{E}[\mathcal{E}_S] = \mathcal{O}(N_s^{-1/4} + N_t^{-1}).$$

2. *The DOD projection error*
$$\mathcal{E}_{DOD} = \mathcal{E}_{DOD} \left(\mathcal{G}, (\mu_i, \nu_i, \tau_i)_{i=1}^{N_{data}}, N \right)$$
satisfies $\mathcal{E}_{DOD} \to \mathcal{E}_{DOD,\infty}$ almost surely as $N_s, N_t \to \infty$, where
$$\mathcal{E}_{DOD,\infty} = \mathcal{E}_{DOD,\infty}(\mathcal{G}, N)$$
is independent of the data snapshots and arbitrarily close, with regards to the capabilities of the DOD neural network, to the expected linear KnW of the submanifold $\mathcal{S}_{\mu,t} = \{u(\mu, \nu, t) \,|\, \nu \in \Theta'\}$ up to a constant, i.e. $\mathbb{E}_{\mu,t}[d_N(\mathcal{S}_{\mu,t})]$.

3. *The neural network approximation error*
$$\mathcal{E}_{NN} = \mathcal{E}_{NN} (\mathcal{G}, N, \hat{q})$$
satisfies $\mathcal{E}_{NN} \to d_N(\mathcal{S}_\mu)$ in expectation, as $N_s, N_t \to \infty$.

Proof. In the core this proof is the same as the one of Theorem (4.1.1), we will therefore only consider the parts in which they are not interchangeable; the use of the triangle inequality gives us the neural network error in the same way, with only the additional factor being dependent on the choice and sophistication of the DOD map,

$$\mathcal{E}_{NN} = \left(\int_{\Theta \times \Theta' \times \mathcal{T}} \frac{\| V(\mu, t)q(\mu, \nu, t) - V(\mu, t)\hat{q}(\mu, \nu, t) \|^2}{\| u(\mu, \nu, t) \|^2} d(\mu, \nu, t) \right)^{1/2}. \tag{4.2.3}$$

Moreover, we can extrapolate the same procedure of splitting the remaining term into the respective errors

$$\| u(\mu, \nu, t) - V(\mu, t)V(\mu, t)^T u(\mu, \nu, t) \|_{L^2_\omega} \leq \mathcal{E}_S + \mathcal{E}_{\text{DOD}}, \tag{4.2.4}$$

where

$$\begin{aligned} \mathcal{E}_S = m^{-1} \Bigg| & \int_{\Theta \times \Theta' \times \mathcal{T}} \| u(\mu, \nu, t) - V(\mu, t)V(\mu, t)^T u(\mu, \nu, t) \|^2 d(\mu, \nu, t) \\ & - \frac{|\Theta \times \Theta' \times \mathcal{T}|}{N_{data}} \sum_{i=1}^{N_{s_1}} \sum_{j=1}^{N_{s_2}} \sum_{k=1}^{N_t} \| u(\mu_i, \nu_j, t_k) - V(\mu_i, t_k)V(\mu_i, t_k)^T u(\mu_i, \nu_j, t_k) \|^2 \Bigg|^{1/2} \end{aligned}$$

and

$$\mathcal{E}_{\text{DOD}} = \frac{m^{-1}|\Theta \times \Theta' \times \mathcal{T}|}{N_{data}} \left| \sum_{i=1}^{N_{s_1}} \sum_{j=1}^{N_{s_2}} \sum_{k=1}^{N_t} \| u(\mu_i, \nu_j, t_k) - V(\mu_i, t_k) V(\mu_i, t_k)^T u(\mu_i, \nu_j, t_k) \|^2 \right|^{1/2}.$$

This leaves us with the proof for the convergence statements. While $\mathcal{E}_S \to 0$ as $N_s, N_t \to \infty$ can be shown in the same manner as it has been in aforementioned proof above, the DOD error is more tricky.

Using the triangle inequality again and combining it with the findings of Lemma (4.2.1) we get for any $\varepsilon > 0$ and a suitable amount of complexity of the ReLU neural network $V : \mathbb{R}^{p+1} \to \mathbb{R}^{N_h \times N}$

$$\begin{aligned}
\mathcal{E}_{\text{DOD}} &\leq m^{-1} \left| \int_{\Theta \times \Theta' \times \mathcal{T}} \| u(\mu, \nu, t) - V(\mu, t) V(\mu, t)^T u(\mu, \nu, t) \|^2 d(\mu, \nu, t) \right|^{1/2} \\
&\quad + \mathcal{E}_{\text{DOD}} - m^{-1} \left| \int_{\Theta \times \Theta' \times \mathcal{T}} \| u(\mu, \nu, t) - V(\mu, t) V(\mu, t)^T u(\mu, \nu, t) \|^2 d(\mu, \nu, t) \right|^{1/2} \\
&\underset{N_s, N_t \to \infty}{\longrightarrow} m^{-1} \left| \int_{\Theta \times \Theta' \times \mathcal{T}} \| u(\mu, \nu, t) - V(\mu, t) V(\mu, t)^T u(\mu, \nu, t) \|^2 d(\mu, \nu, t) \right|^{1/2} \\
&\quad < m^{-1} \varepsilon + m^{-1} \int_{\Theta \times \mathcal{T}} d_N(\mathcal{S}_{\mu, t}) d(\mu, t)
\end{aligned}$$

□

Lemma 4.2.3 (Projection Error)**.** Let N_{s_2} be given under Assumption (2), then for all $\varepsilon > 0$ there is a ReLU neural network $V : \mathbb{R}^{p+1} \to \mathbb{R}^{N_h \times N}$, such that for $u_j := u(\mu, \nu_j, t)$ with given $\mu \in \Theta$ and $t \in \mathcal{T}$

$$\sup_{(\mu,t) \in \Theta \times \mathcal{T}} \left| \frac{1}{N_{s_2}} \sum_{j=1}^{N_{s_2}} \| u_j - V(\mu, t) V(\mu, t)^T u_j \|^2 - \min_{W \in \mathbb{R}^{N_h \times N}} \frac{1}{N_{s_2}} \sum_{j=1}^{N_{s_2}} \| u_j - W W^T u_j \|^2 \right|^{1/2} < \varepsilon.$$

Proof. The existence and identity of V^* s.t.

$$V^* = \underset{W \in \mathbb{R}^{N_h \times N}}{\operatorname{argmin}} \sum_{j=1}^{N_{s_2}} \| u_j - V^* (V^*)^T u_j \|^2$$

is given via the statement in Proposition (2.1.3) with the POD basis. Hence defining a norm via

$$\| (u_j)_{j=1}^{N_{s_2}} \|_{train} := \left| \sum_{j=1}^{N_{s_2}} \| u_j \|^2 \right|^{1/2}$$

gives us the following inequality firstly using the monotonicity of the square root and then using the triangle inequalities multiple times

$$\begin{aligned}
&N_{s_2}^{-1/2} \sup_{(\mu,t) \in \Theta \times \mathcal{T}} \left| \sum_{j=1}^{N_{s_2}} \| u_j - V(\mu, t) V(\mu, t)^T u_j \|^2 - \sum_{j=1}^{N_{s_2}} \| u_j - V^* (V^*)^T u_j \|^2 \right|^{1/2} \\
&\leq N_{s_2}^{-1/2} \sup_{(\mu,t) \in \Theta \times \mathcal{T}} \left| \sum_{j=1}^{N_{s_2}} \| V(\mu, t) V(\mu, t)^T u_j - V^* (V^*)^T u_j \|^2 \right|^{1/2},
\end{aligned}$$

then define $q_j := (V^*)^T u_j$ and $\hat{q}_j := V(\mu, t)^T u_j$ and their training vectors accordingly, i.e. $q :=$

$(q_j)_{j=1}^{N_{s_2}}, \hat{q} := (\hat{q}_j)_{j=1}^{N_{s_2}}$

$$
\begin{aligned}
&\leq N_{s_2}^{-1/2} \sup_{(\mu,t)\in\Theta\times\mathcal{T}} \| V(\mu,t)\hat{q} - V(\mu,t)q \|_{train} + \| V(\mu,t)q - V^* q \|_{train} \\
&\leq N_{s_2}^{-1/2} \sup_{(\mu,t)\in\Theta\times\mathcal{T}} \| V(\mu,t) \|_\infty \| \hat{q} - q \|_{train} + \| V(\mu,t) - V^* \|_\infty \| q \|_{train} \\
&= N_{s_2}^{-1/2} \sup_{(\mu,t)\in\Theta\times\mathcal{T}} \| V(\mu,t) \|_\infty \| V(\mu,t)^T u - (V^*)^T u \|_{train} + \| V(\mu,t) - V^* \|_\infty \| (V^*)^T u \|_{train} \\
&\leq N_{s_2}^{-1/2} \sup_{(\mu,t)\in\Theta\times\mathcal{T}} \| V(\mu,t) \|_\infty \| V(\mu,t) - V^* \|_\infty \| u \|_{train} + \| V(\mu,t) - V^* \|_\infty \| V^* \|_\infty \| u \|_{train} \\
&\leq N_{s_2}^{-1/2} \sup_{(\mu,t)\in\Theta\times\mathcal{T}} \| u \|_{train} (\| V(\mu,t) - V^* \|_\infty (\| V(\mu,t) \|_\infty + 1)) \\
&\leq 2M \sup_{(\mu,t)\in\Theta\times\mathcal{T}} \| V(\mu,t) - V^* \|_\infty,
\end{aligned}
$$

where we used $\| u \|_{train} \leq | \sum_{j=1}^{N_{s_2}} M^2 |^{1/2} = N_{s_2}^{1/2} M$ and additionally $\| V(\mu,t) \|_\infty = \| V^* \|_\infty = 1$ by orthonormality in the last arguments.

□

4.3 for POD-DL-ROMs

Theorem 4.3.1. *Let $\mathcal{G} : (\mu,\nu,t) \mapsto u(\mu,\nu,t)$ for any $(\mu,\nu,t) \in \Theta \times \Theta' \times \mathcal{T}$ be the parameter-to-solution map and let Assumption (2) and (3) be true. Let $1 > \delta > 0$ and $\varepsilon > 0$. Now assume we can collect freely a sufficient amount of training data $N_{data} = N_{data}(\delta,\varepsilon)$ and put it into the correlation matrix K as given in Definition (2.1.1) with σ_k^2 being its sorted eigenvalues. Now choose*

$$N = \operatorname{argmin} \left\{ j \in \mathbb{N} \,\middle|\, \sum_{k>j} \sigma_k^2 \leq \frac{m^2}{9}\varepsilon^2 \right\}.$$

Given the POD matrix $V \in \mathbb{R}^{N_h \times N}$ we now define the parameter-to-POD-coefficient map $\mathcal{Q} : (\mu,\nu,t) \mapsto q(\mu,\nu,t) := V^T u(\mu,\nu,t)$ for any $(\mu,\nu,t) \in \Theta \times \Theta' \times \mathcal{T}$. We assume that there exists $n > 0, \psi_* : \mathbb{R}^n \to \mathbb{R}^N, \psi'_* : \mathbb{R}^N \to \mathbb{R}^n$ that are respectively s-times and s'-times differentiable (with $s >> s' \geq 2$), such that they enjoy the perfect embedding Assumption (4), i.e.

$$\psi_*(\psi'_*(q(\mu,\nu,t))) = q(\mu,\nu,t), \quad \forall(\mu,\nu,t) \in \Theta \times \Theta' \times \mathcal{T}.$$

We further let

$$C_1 := \sup_{|\alpha|\leq s'} \sup_{v\in\mathbb{R}^N} | D^\alpha \psi'_*(v) |, \quad C_2 := \sup_{|\alpha|\leq s} \sup_{w\in\mathbb{R}^n} | D^\alpha \psi'_*(w) |.$$

Then there is a constant $c = c(\Theta,\Theta',\mathcal{T},L,C_1,C_2,p,q,n,s,s')$ and a POD-DL-ROM architecture $V\hat{q} = V\psi \circ \phi : \mathbb{R}^{p+q+1} \to \mathbb{R}^{N_h}$ composed of a decoder $\psi : \mathbb{R}^n \to \mathbb{R}^N$ having at most:

- $L_{n\to N} = c\log(\varepsilon^{-1})$ *layers,*
- $\omega_{n\to N} = cN\varepsilon^{-n/(s-1)}\log(\varepsilon^{-1})$ *active weights,*

and a reduced feed forward neural network $\phi : \mathbb{R}^{p+q+1} \to \mathbb{R}^n$ having at most:

- $L_{(p+q+1)\to n} = c\log(\varepsilon^{-1})$ *layers,*
- $\omega_{(p+q+1)\to n} = cn\varepsilon^{-(p+q+1)}\log(\varepsilon^{-1})$ *active weights,*

such that $\mathbb{P}[\mathcal{E}_R < \varepsilon] > 1 - \delta$.

Proof. We can directly bound $\mathcal{E}_S = \mathcal{E}_S(N_{s_1}, N_{s_2}, N_t)$ independently of N, under the Assumption (2) by the Weak Law of Large Numbers, i.e. for all $1 > \delta > 0$ and $\varepsilon > 0$ there are N_{s_1}, N_{s_2}, N_t, such that

$$\mathbb{P}[\mathcal{E}_S((N_{s_1}, N_{s_2}, N_t)) < \varepsilon/3] > 1 - \delta. \tag{4.3.1}$$

This result provides the "probability" part of the PAC (probably almost correct) bound, which we want to show. Bounding $\mathcal{E}_{\text{NN}}$ and $\mathcal{E}_{\text{POD}}$ by $\varepsilon/3$ respectively will then provide the wanted result. For this, by choosing N as in the assumption of the statement, we simply rewrite

$$\mathcal{E}_{\text{POD}} := m^{-1}\sqrt{\sum_{k>N} \sigma_k^2} \leq \frac{\varepsilon}{3}. \tag{4.3.2}$$

To now bound the neural network approximation error, we will need to simply rewrite the error in terms of only the norm of the network itself;

$$\begin{aligned}
\mathcal{E}_{\text{NN}} &= \left(\int_{\Theta\times\Theta'\times\mathcal{T}} \frac{\| Vq(\mu,\nu,t) - V\hat{q}(\mu,\nu,t) \|^2}{\| u(\mu,\nu,t) \|^2}\, d(\mu,\nu,t)\right)^{1/2} \\
&\leq m^{-1}\left(\int_{\Theta\times\Theta'\times\mathcal{T}} \| Vq(\mu,\nu,t) - V\hat{q}(\mu,\nu,t) \|^2\, d(\mu,\nu,t)\right)^{1/2} \\
&\leq m^{-1}\left(|\Theta\times\Theta'\times\mathcal{T}| \sup_{(\mu,\nu,t)\in\Theta\times\Theta'\times\mathcal{T}} \| Vq(\mu,\nu,t) - V\hat{q}(\mu,\nu,t) \|^2\right)^{1/2} \\
&\leq m^{-1}\left(|\Theta\times\Theta'\times\mathcal{T}|\, \| V \|_\infty^2 \sup_{(\mu,\nu,t)\in\Theta\times\Theta'\times\mathcal{T}} \| q(\mu,\nu,t) - \hat{q}(\mu,\nu,t) \|^2\right)^{1/2} \\
&= m^{-1}\, |\Theta\times\Theta'\times\mathcal{T}|^{1/2} \sup_{(\mu,\nu,t)\in\Theta\times\Theta'\times\mathcal{T}} \| q(\mu,\nu,t) - \hat{q}(\mu,\nu,t) \|.
\end{aligned} \tag{4.3.3}$$

whereas we split the latter error into both parts, since $\hat{q} = \psi \circ \phi$ per definition. This means employing two types of error estimations in the matter of finding the respective optimal maps for the feed forward estimation of the latent dynamics via $\phi_* : \mathbb{R}^{(p+q+1)} \to \mathbb{R}^n$ and the decoder lift via $\psi_* : \mathbb{R}^n \to \mathbb{R}^N$.

- Let $\mathcal{S}_N := \{q = \mathcal{Q}(\mu,\nu,t) \,|\, (\mu,\nu,t) \in \Theta\times\Theta'\times\mathcal{T}\}$, then the image of the optimal embedding $\mathcal{V}_n = \psi'_*(\mathcal{S}_N)$ is such that $\text{diam}(\mathcal{V}_N) \leq LC_1\text{diam}(\Theta\times\Theta'\times\mathcal{T})$, given the Assumption (3) by continuity.

  Thus, by Theorem Gühring, there is a ReLU Neural Network $\psi : \mathbb{R}^n \to \mathbb{R}^N$, such that

  $$\sup_{v\in\mathcal{V}_N} \| \psi(v) - \psi_*(v) \| < \frac{m}{6} |\, \Theta\times\Theta'\times\mathcal{T} \,|^{-1/2}\varepsilon, \tag{4.3.4}$$

  $$\text{ess}\sup_{v,v'\in\mathcal{V}_N} \frac{|\,(\psi-\psi_*)(v) - (\psi-\psi_*)(v')\,|}{|\, v - v' \,|} < \frac{m}{6} |\, \Theta\times\Theta'\times\mathcal{T} \,|^{-1/2}\varepsilon, \tag{4.3.5}$$

  with $L_{n\to N} = c\log(\varepsilon^{-1})$ layers and $\omega_{n\to N} = cN\varepsilon^{-n/(s-1)}\log(\varepsilon^{-1})$ active weights.

- Now let $\phi_*(\mu,\nu,t) = \psi'_*(q(\mu,\nu,t))$ for $(\mu,\nu,t) \in \Theta\times\Theta'\times\mathcal{T}$, which is Lipschitz continuous, with constant LC_1, and thus, by

  Theorem Yarotski, there exists a ReLU Neural Network $\phi : \mathbb{R}^{p+q+1} \to \mathbb{R}^n$, such that

  $$\sup_{(\mu,\nu,t)\in\Theta\times\Theta\times\mathcal{T}} \| \phi(\mu,\nu,t) - \phi_*(\mu,\nu,t) \| < \frac{m}{6C_3} |\, \Theta\times\Theta'\times\mathcal{T} \,|^{-1/2}\varepsilon, \tag{4.3.6}$$

  with $L_{(p+q+1)\to n} = c\log(\varepsilon^{-1})$ layers and $\omega_{(p+q+1)\to n} = cn\varepsilon^{-(p+q+1)}\log(\varepsilon^{-1})$ active weights.

Starting from the last inequality in (4.3.3) we can perform the split in a clean fashion using the triangle inequality and reminding ourselves of the Assumption (4), which guarantees $q = \psi_*(\psi'_*)$, and thus gives with both neural network estimates in Equations (4.3.4) and (4.3.6)

$$\begin{aligned}
&\sup_{(\mu,\nu,t)\in\Theta\times\Theta'\times\mathcal{T}} \| \, q(\mu,\nu,t) - \hat{q}(\mu,\nu,t) \, \| \\
\leq &\sup_{(\mu,\nu,t)\in\Theta\times\Theta'\times\mathcal{T}} (\, \| \, \psi_*(\psi'_*(\mu,\nu,t)) - \psi(\phi_*(\mu,\nu,t)) \, \| + \| \, \psi(\phi_*(\mu,\nu,t)) - \psi(\phi(\mu,\nu,t)) \, \| \,) \\
\leq &\sup_{v\in\mathcal{V}_N} \| \, \psi_*(v) - \psi(v) \, \| + C_3 \sup_{(\mu,\nu,t)\in\Theta\times\Theta'\times\mathcal{T}} (\, \| \, \psi(\phi_*(\mu,\nu,t)) - \psi(\phi(\mu,\nu,t)) \, \| \,) \\
< &\frac{m}{6} | \, \Theta\times\Theta'\times\mathcal{T} \, |^{-1/2}\varepsilon + C_3 \frac{m}{6C_3} | \, \Theta\times\Theta'\times\mathcal{T} \, |^{-1/2}\varepsilon = \frac{\varepsilon}{3} m \, | \, \Theta\times\Theta'\times\mathcal{T} \, |^{-1/2},
\end{aligned}$$

which cancels with $m^{-1} \, |\Theta\times\Theta'\times\mathcal{T}|^{1/2}$ to get

$$\mathcal{E}_{\text{NN}} < \frac{\varepsilon}{3}. \tag{4.3.7}$$

Now putting together all three estimates (4.3.1), (4.3.2) and (4.3.7) with the reasoning of Theorem (4.1.1) finally gives

$$\mathcal{E}_R \leq \mathcal{E}_S + \mathcal{E}_{\text{POD}} + \mathcal{E}_{\text{NN}} < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

with higher probability than $1-\delta$. □

4.4 for DOD-DL-ROMs

Theorem 4.4.1. *Let $\mathcal{G} : (\mu,\nu,t) \mapsto u(\mu,\nu,t)$ for any $(\mu,\nu,t)\in\Theta\times\Theta'\times\mathcal{T}$ be the parameter-to-solution map and let Assumption (2) and (3) be true. Let $1 > \delta > 0$ and $\varepsilon > 0$. Now assume we can collect freely a sufficient amount of training data $\Theta_{data}, \Theta'_{data}$ and $\mathcal{T}_{data}$ of size $N_{data} = N_{data}(\delta,\varepsilon)$ and put all samples of $u(\mu,\nu_j,t)_{j=1}^{N_{s_2}}$ into a correlation matrix $K_{\mu,t}$ for each sample $(\mu,t)\in\Theta_{data}\times\mathcal{T}_{data}$ as given in Definition (2.1.1) with $\sigma_k(\mu,t)_k^2$ being its sorted eigenvalues. Now choose*

$$N_{DOD} = \operatorname{argmin}\left\{ j\in\mathbb{N} \,\middle|\, \sup_{(\mu,t)\in\Theta_{data}\times\mathcal{T}_{data}} \sum_{k>j} \sigma(\mu,t)_k^2 \leq \frac{m^2}{16}\varepsilon^2 \right\}.$$

For each pair $(\mu,t)\in\Theta_{data}\times\mathcal{T}_{data}$ given the respective POD matrix, so the best possible projection given the data set (refer to (2.1.3)), $V^*(\mu,t)\in\mathbb{R}^{N_h\times N_{DOD}}$ we now define the parameter-to-optimal-coefficient map $\mathcal{Q} : (\mu,\nu,t)\mapsto q(\mu,\nu,t) := V^*(\mu,t)^T u(\mu,\nu,t)$ for any $(\mu,\nu,t)\in\Theta_{data}\times\Theta'\times\mathcal{T}_{data}$. We assume that there exists $n>0, \Psi_* : \mathbb{R}^n\to\mathbb{R}^{N_{DOD}}, \Psi'_* : \mathbb{R}^{N_{DOD}}\to\mathbb{R}^n$ that are respectively s-times and s'-times differentiable (with $s >> s' \geq 2$), such that they enjoy the perfect embedding Assumption (4), i.e.

$$\Psi_*(\Psi'_*(q(\mu,\nu,t))) = q(\mu,\nu,t), \quad \forall(\mu,\nu,t)\in\Theta\times\Theta'\times\mathcal{T}.$$

We further let

$$C'_1 := \sup_{|\alpha|\leq s'} \sup_{v\in\mathbb{R}^{N_{DOD}}} | \, D^\alpha\Psi'_*(v) \, |, \quad C'_2 := \sup_{|\alpha|\leq s} \sup_{w\in\mathbb{R}^n} | \, D^\alpha\Psi'_*(w) \, |.$$

Then there is a constant $c = c(\Theta,\Theta',\mathcal{T},L,C_1,C_2,p,q,n,s,s')$ and a DOD-DL-ROM architecture $V\hat{q} = V_{DOD}\Psi\circ\phi : \mathbb{R}^{p+q+1}\to\mathbb{R}^{N_h}$ composed of a decoder $\Psi : \mathbb{R}^n\to\mathbb{R}^{N_{DOD}}$ having at most:

- $L_{n\to N_{DOD}} = c\log(\varepsilon^{-1})$ *layers,*
- $\omega_{n\to N_{DOD}} = c\,N_{DOD}\,\varepsilon^{-n/(s-1)}\log(\varepsilon^{-1})$ *active weights,*

a reduced feed forward neural network $\phi : \mathbb{R}^{p+q+1} \to \mathbb{R}^n$ *having at most:*

- $L_{(p+q+1)\to n} = c\log(\varepsilon^{-1})$ *layers,*

- $\omega_{(p+q+1)\to n} = cn\varepsilon^{-(p+q+1)}\log(\varepsilon^{-1})$ *active weights,*

and a DOD Module $V_{DOD} : \mathbb{R}^{p+1} \to \mathbb{R}^{N_h \times N_{DOD}}$ *having at most:*

- $L_{(p+1)\to N_h \times N_{DOD}} = c\log(\varepsilon^{-1})$ *layers,*

- $\omega_{(p+1)\to N_h \times N_{DOD}} = c\, N_h\, N_{DOD}\, \varepsilon^{-(p+1)}\log(\varepsilon^{-1})$ *active weights,*

such that $\mathbb{P}[\mathcal{E}_R < \varepsilon] > 1 - \delta$.

Proof. Similar as to the proof of Theorem (4.3.1), we will try to bound each quantity of the statement in Theorem (4.2.2), but with the catch, that $\mathcal{E}_{\text{DOD}}$ is stochastic in nature, and thus we will employ Lemma (4.2.3) to find an auxilary term, which is independent of the data, to which the approximation of the neural net of the DOD is sufficiently "close".

The sampling error can be bounded in the same way as before, i.e. we can directly bound $\mathcal{E}_S = \mathcal{E}_S(N_{s_1}, N_{s_2}, N_t)$ independently of N_{DOD}, under the Assumption (2) by the Weak Law of Large Numbers, i.e. for all $1 > \delta > 0$ and $\varepsilon > 0$ there are N_{s_1}, N_{s_2}, N_t, such that

$$\mathbb{P}[\mathcal{E}_S((N_{s_1}, N_{s_2}, N_t)) < \varepsilon/4] > 1 - \delta. \tag{4.4.1}$$

Now recall the definition of $\mathcal{E}_{\text{DOD}}$ in the proof of Theorem (4.2.2) and further establish

$$\begin{aligned}\mathcal{E}_{\text{DOD}} &= \frac{m^{-1}|\Theta \times \Theta' \times \mathcal{T}|}{N_{data}^{1/2}} \left| \sum_{i=1}^{N_{s_1}} \sum_{j=1}^{N_{s_2}} \sum_{k=1}^{N_t} \| u(\mu_i, \nu_j, t_k) - V(\mu_i, t_k) V(\mu_i, t_k)^T u(\mu_i, \nu_j, t_k) \|^2 \right|^{1/2} \\ &\le m^{-1} \sup_{(\mu,t)\in\Theta\times\mathcal{T}} \left| \frac{|\Theta \times \Theta' \times \mathcal{T}|}{N_{s_2}} \sum_{j=1}^{N_{s_2}} \| u(\mu, \nu_j, t) - V(\mu, t) V(\mu, t)^T u(\mu, \nu_j, t) \|^2 \right|^{1/2},\end{aligned}$$

enabling us to use Proposition (2.1.3) in order to switch to the eigenvalues $\sigma(\mu,t)_k$ for $(\mu,t) \in \Theta_{\text{data}} \times \mathcal{T}_{\text{data}}$ of the theoretical POD matrix $V^*(\mu,t)$ by employing the Lemma (4.2.3) with the bound

$$\frac{\varepsilon\, m^2}{4|\Theta \times \Theta' \times \mathcal{T}|^{1/2}},$$

which directly gives us the approximation result via the defition of N_{DOD}

$$\mathcal{E}_{\text{DOD}} < \frac{\varepsilon}{4} + m^{-1} \sup_{(\mu,t)\in(\Theta_{\text{data}}\times\mathcal{T}_{\text{data}})} \sqrt{\sum_{k>N_{\text{DOD}}} \sigma(\mu,t)_k^2} \le \frac{\varepsilon}{2}. \tag{4.4.2}$$

Moreover we can use the exact same arguments as in the bound for $\mathcal{E}_{\text{NN}}$ for the POD-DL-ROM as in the proof of Theorem (4.3.1), which yields

$$\mathcal{E}_{\text{NN}} < \frac{\varepsilon}{4}. \tag{4.4.3}$$

Finally putting each Equation, i.e. (4.4.1), (4.4.2) and (4.4.3), together, we get

$$\mathcal{E}_R \le \mathcal{E}_S + \mathcal{E}_{\text{DOD}} + \mathcal{E}_{\text{NN}} < \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon,$$

with probability greater than $1 - \delta$ concluding our proof. □

Chapter 5

Numerical Experiments

Chapter 6

Conclusion

Bibliography

[FDM20] Stefania Fresca, Luca Dede, and Andrea Manzoni. *A comprehensive deep learning-based approach to reduced order modeling of nonlinear time-dependent parametrized PDEs*. 2020. arXiv: 2001.04001 [math.NA]. URL: https://arxiv.org/abs/2001.04001.

[Fra+24] Nicola Rares Franco et al. *Deep orthogonal decomposition: a continuously adaptive data-driven approach to model order reduction*. Preprint, arXiv:2404.18841 [math.NA] (2024). arXiv:2404.18841. 2024. URL: https://arxiv.org/abs/2404.18841.

[Hor91] Kurt Hornik. "Approximation capabilities of multilayer feedforward networks". In: *Neural Networks* 4.2 (1991), pp. 251–257. ISSN: 0893-6080. DOI: https://doi.org/10.1016/0893-6080(91)90009-T. URL: https://www.sciencedirect.com/science/article/pii/089360809190009T.

[QMN15] Alfio Quarteroni, Andrea Manzoni, and Federico Negri. *Reduced Basis Methods for Partial Differential Equations*. Springer, 2015. DOI: 10.1007/978-3-319-15431-2.