Discontinuous Galerkin and Trefftz Methods for Model Reduction of Wave Phenomena

Tobias Born[1,*] and **Karsten Urban**[1]

[1] Institute of Numerical Mathematics, Ulm University, 89069 Ulm, Germany

The space-time discontinuous Galerkin (dG)-Trefftz is known to be a highly efficient numerical scheme for solving linear hyperbolic problems. We investigate to what extend such a dG-Trefftz method can be used as a basis for a model reduction method for a traveling wave problem using the wave speed as a parameter. Such problems are known to be tough for linear model reduction techniques as the error decay is slow with increasing size of the reduced model (by the Kolmogorov n-width).

The presented dG-Trefftz method yields a nonlinear model reduction technique as the reduced trial space is parameter-dependent. We present results of numerical experiments which show a convergence rate which is better than the known worst-case rate for linear schemes, but which is still polynomially. We compare the dG-Trefftz method with a nonlinear model technique based upon trained autoencoders using neural networks. It turns out that these methods are not able to outperform the dG-Trefftz scheme.

1 Introduction

Transport and wave-type problems are well-known to be hard problems for model reduction. In fact, the Kolmogorov n-width being the best possible error achieved by linear model reduction with $n \in \mathbb{N}$ degrees of freedom decays only at a polynomial rate. Instead, elliptic and parabolic problems allow for an exponential rate of convergence, [4, 9, 12].

On the other hand, Trefftz discontinuous Galerkin (dG-Trefftz) methods are known to yield very efficient numerical solvers for linear hyperbolic problems, [2]. Hence, such schemes are natural candidates to serve as a backbone for model reduction, for example one might think to determine the snapshots forming the reduced model in an offline training phase by using dG-Trefftz methods. Moreover, dG-Trefftz discretizations of parameterized problems lead to a variational form, where the parameter (we use the wave speed) is contained in the bilinear and linear forms and cannot be separated from the primitive variables space and time. On the one hand, this is challenge as such a separation (known as "affine decomposition") is a crucial property to ensure the efficiency of model order reduction schemes such as the Reduced Basis Method (RBM), [1, 10, 12]. On the other hand, however, this causes the fact that a dG-Trefftz method yields reduced trial spaces, which are parameter-dependent, which means that the resulting model reduction technique is nonlinear and might not suffer from the poor decay of the Kolmogorov n-width.

The aim of this paper is to investigate if and/or to which extent a dG-Trefftz method is able to yield a nonlinear model reduction that outperforms existing linear ones. To this end, we consider a seemingly simple model problem of a traveling wave in one space dimension using the wave speed as a parameter. This parameter is chosen in such a way that a reflection on the right-end boundary occurs for certain parameters and for others, the wave stops before the right endpoint at the terminal time. This problem has been described and investigated in [6], where also a decay of $n^{-3.5}$ was observed numerically.

The remainder of this paper is organized as follows. Section 2 contains the considered model problem, a brief introduction to the RBM and a justification that a dG-Trefftz method can lead to a nonlinear model reduction. Section 3 is devoted to the description of the space-time dG-Trefftz formulation of the traveling wave model problem and we introduce two variants of corresponding model reduction methods, one linear and one nonlinear one. Numerical results for the dG-Trefftz model reduction methods are presented in Section 4. Since the performance turned out to be far off the exponential decay known from elliptic and parabolic problems, we also consider a different nonlinear model reduction technique based upon autoencoders known from [5, 8]. To a certain extent, this is in fact a benchmark as it is known (see e.g. [3]) that appropriate combinations of nonlinear decoders and encoders yield an optimal compression. However, we observe that such techniques based upon neural networks cannot outperform the dG-Trefftz method (which also shows that the chosen model problem is in fact tough). We end this paper by some conclusions and an outlook in Section 6.

2 Model reduction for the wave equation

In this section, we introduce the considered model problem and review the main facts of (linear) model reduction by means of the Reduced Basis Method (RBM).

* Corresponding author: tobias.born@uni-ulm.de

2.1 A traveling wave model problem

Following [6], we consider the following model of a traveling wave problem, namely finding $u : I \times \Omega \times \mathcal{P} \to \mathbb{R}$ such that

$$\begin{aligned}
\mu^{-2}\left(\frac{\partial}{\partial t}\right)^2 u(t,x;\mu) - \left(\frac{\partial}{\partial x}\right)^2 u(t,x;\mu) &= 0 && \text{for } (t,x) \in I\times\Omega,\\
u(t,0;\mu) &= \tanh(5t)^3 && \forall t \in I,\\
u(t,1;\mu) &= 0 && \forall t \in I,\\
u(0,x;\mu) &= 0 && \forall x \in \Omega,\\
\frac{\partial}{\partial t} u(0,x;\mu) &= 0 && \forall x \in \Omega,
\end{aligned} \tag{1}$$

is fulfilled in a weak sense, where the time domain is $I := (0,1)$, the spatial domain reads $\Omega := (0,1)$ and the parameter $\mu \in [0.3, 2] =: \mathcal{P}$ denotes the wave speed. This corresponds to a wave traveling from $x = 0$ towards $x = 1$ within a time period of 1. For $\mu = 1$, the wave reaches $x = 1$ exactly at time $t = 1$. For $\mu < 1$, the wave is too slow to reach $x = 1$ in the given time period, and for $\mu > 1$, the wave is reflected at $x = 1$. These different scenarios are visualized in Figure 1 for three different parameter values.

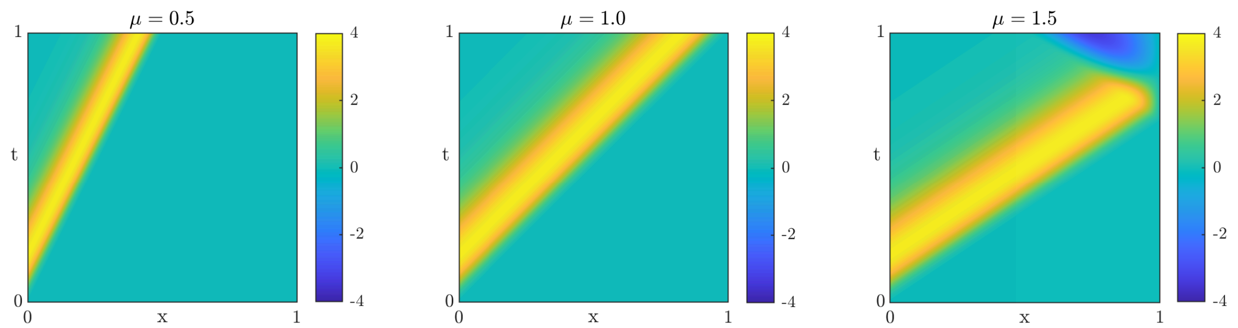

Fig. 1: Plot of $\frac{\partial}{\partial x}u$ on $I\times\Omega$ for different wave speeds μ. The different behaviors of the wave are clearly visible.

Even though seemingly simple, (1) contains all relevant challenges for model reduction, namely a traveling wave phenomenon and completely different scenarios (due to the reflection) for different ranges of the parameter.

2.2 Model reduction by the Reduced Basis Method (RBM)

The above model problem (1) is an example of a *parameterized* partial differential equation (PPDE). Within a multiquery, realtime or cold computing scenario, one often wishes to solve the PPDE for different values of the parameter very often, extremely fast or on heavily restricted computing devices. In such situations, model reduction is a must.

Roughly speaking, the recipe of the RBM can be summarized as follows, [1, 10, 12]:

1. Construct a well-posed (variational) formulation of the PPDE of the form: find $u(\mu) \in U$ such that $a(u(\mu), v; \mu) = f(v; \mu)$ for all $v \in U$, where U is an appropriate trial and test space[1], $a(\cdot, \cdot; \mu) : U \times U \to \mathbb{R}$ is a parameter-dependent bilinear form and $f(\cdot; \mu) : U \to \mathbb{R}$ a parameter-dependent right-hand side (e.g. a force). Well-posedness includes existence, uniqueness and stability of solutions.
2. The second ingredient is a sufficiently detailed numerical solution, e.g. by a finite element or finite volume method on a sufficiently fine grid. The arising approximation $u^N(\mu)$ of $u(\mu)$ is assumed to be sufficiently accurate and can be computed in complexity $\mathcal{O}(N)$ with $N \in \mathbb{N}$ large.
3. In an offline training phase, a subset of parameter samples $\mu^{(1)}, ..., \mu^{(n)} \subset \mathcal{P}$, $n \ll N$, is selected (by a greedy scheme w.r.t. the error estimator Δ_n, described below, on a finite training set $\mathcal{P}_{\text{train}} \subset \mathcal{P}$) and the reduced space $U_n := \operatorname{span}\{u^N(\mu^{(1)}), ..., u^N(\mu^{(n)})\}$ is determined by the "snapshots" $u^N(\mu^{(i)})$ computed through the detailed solver.
4. In the online phase, given a new parameter value $\mu \in \mathcal{P}$, the reduced approximation $u_n(\mu)$ is computed by the Galerkin approximation of the variational problem onto U_n.
5. In order that all this works efficiently, one needs two major ingredients, namely
   - an error estimator $\Delta_n(\mu)$, which is computable in a complexity depending on n, but not on N ("online efficient") such that $\|u^N(\mu) - u_n(\mu)\|_U \le \Delta_n(\mu)$ and

[1] We omit the case of different trial and test spaces for the sake of brevity.

- that parameters can be separated from the primal variables (here t and x) in the sense that

$$a(u,v;\mu) = \sum_{q=1}^{Q^a} \vartheta_q^a(\mu)\, a_q(u,v),$$

also known as “affine decomposition”, where $Q^a \in \mathbb{N}$, $\vartheta_q^a : \mathcal{P} \to \mathbb{R}$ and $a_q : U \times U \to \mathbb{R}$ (and similar for $f(\cdot;\mu)$).

Apparently, this is a *linear* model reduction technique as we use a linear subspace U_n of dimension $n \in \mathbb{N}$. The best possible error of such a linear method for all parameters is given by the *Kolmogorov n-width*

$$d_n(\mathcal{P}) := \inf_{\substack{U_n\subset U\\ \dim(U_n)=n}} \sup_{\mu\in\mathcal{P}} \inf_{w_n\in U_n} \left\| u^N(\mu) - w_n \right\|_U . \tag{2}$$

It is known that $d_n(\mathcal{P}) \lesssim e^{-cN/Q^a}$ for elliptic and parabolic problems, [9, 12]. On the other hand, it is also known that $d_n(\mathcal{P}) \cong N^{-1/4}$ has to be expected for wave-type problems, [4]. For the above problem (1), it was observed numerically in [6] that $d_n(\mathcal{P}) \cong N^{-7/2}$.

The strong greedy method. As we are mainly interested in the question if a dG-Trefftz method can be used to improve the poor decay of the Kolmogorov n-width, we do not use any a posteriori error estimator $\Delta_n(\mu)$. Instead, we use the error $\|u^N(\mu) - u_n(\mu)\|_U$ itself, which is if course computationally demanding. Putting this into the offline greedy method maximizing over a finite training set $\mathcal{P}_{\text{train}} \subset \mathcal{P}$ is called *strong greedy method*, whereas using an appropriate estimator $\Delta_n(\mu)$ is known as *week greedy method*.

2.3 Parameter-dependent trial and test spaces

The poor decay of the Kolmogorov n-width is the motivation to consider a dG-Trefftz method for (1). Following the same recipe for the RBM as above yields parameter-dependent trial and test spaces, i.e. $U_n(\mu)$, so that $d_n(\mathcal{P})$ is no longer the best possible error. Our aim is to investigate to which extent such a *nonlinear* dG-Trefftz model reduction is able to overcome the shortcomings of linear model reduction.

3 A space-time discontinuous Galerkin-Trefftz formulation

We start by a weak formulation of the wave equation that was introduced in [2], which refers to the first-order system wave equation with homogeneous right-hand side. On a space-time domain $Q = (0,T)\times\Omega$ with an open, bounded Lipschitz polytope $\Omega \subset \mathbb{R}$, using the wave speed $\mu > 0$ and initial and boundary conditions in terms of functions

$$v_0 : \Omega \to \mathbb{R}\,, \quad \sigma_0 : \Omega \to \mathbb{R} \quad \text{and} \quad g_D : (0,T)\times\partial\Omega \to \mathbb{R},$$

the problem is to find $v(\mu) : \overline{Q} \to \mathbb{R}$ and $\sigma(\mu) : \overline{Q} \to \mathbb{R}$ that satisfy (in a weak sense)

$$\begin{aligned}
&\frac{\partial}{\partial t}\sigma(\mu) + \frac{\partial}{\partial x}v(\mu) = 0 && \text{in } Q,\\
&\mu^{-2}\frac{\partial}{\partial t}v(\mu) + \frac{\partial}{\partial x}\sigma(\mu) = 0 && \text{in } Q,\\
&v(0,\cdot;\mu) = v_0\,,\ \sigma(0,\cdot;\mu) = \sigma_0 && \text{on } \Omega,\\
&v(\mu) = g_D && \text{on } (0,T)\times\partial\Omega\,.
\end{aligned} \tag{3}$$

This includes the second-order wave equation, since (1) can be reformulated as (3) with $v(\mu) = \frac{\partial}{\partial t}u(\mu)$ and $\sigma(\mu) = -\frac{\partial}{\partial x}u(\mu)$.

3.1 Trefftz spaces

A discretization of (3) is constructed on a mesh $\mathcal{T}_h$, on which a *local* (parameter-dependent) Trefftz space on a single element $K \in \mathcal{T}_h$ is defined as

$$\boldsymbol{T}(K;\mu) := \Big\{ (\omega,\tau) \in L^2(K)^2 \;\Big|\; \frac{\partial\tau}{\partial t} + \frac{\partial\omega}{\partial x} = 0, \quad \mu^{-2}\frac{\partial\omega}{\partial t} + \frac{\partial\tau}{\partial x} = 0 \\ \text{and} \quad \tau|_{\partial K} \in L^2(\partial K),\ \frac{\partial\omega}{\partial t}, \frac{\partial\tau}{\partial t}, \frac{\partial\tau}{\partial x}, \frac{\partial\omega}{\partial x} \in L^2(K) \Big\}$$

and the *global* Trefftz space is then defined as

$$\boldsymbol{T}(\mathcal{T}_h;\mu) := \Big\{ (\omega,\tau) \in L^2(Q)^2 \;\Big|\; (\omega|_K, \tau|_K) \in \boldsymbol{T}(K;\mu)\ \forall K \in \mathcal{T}_h \Big\}\,.$$

Local Trefftz functions have a certain regularity and solve the wave equation locally exactly (in a weak sense), which is called the "Trefftz property". Globally, they do not even have to be continuous. Here, we will consider piecewise-polynomial Trefftz functions: For a polynomial degree $p \in \mathbb{N}_0$, the respective space reads

$$\mathbb{T}^p(\mathcal{T}_h;\mu) := \prod_{K\in\mathcal{T}_h} \mathbb{T}^p(K;\mu) \quad \text{with} \quad \mathbb{T}^p(K;\mu) := \boldsymbol{T}(K;\mu) \cap \mathbb{P}^p(K)^2 \ ,$$

where $\mathbb{P}^p(K)$ denotes the space of polynomials on K with degree at most p. Clearly, $\mathbb{T}^p(\mathcal{T}_h;\mu) \subset \boldsymbol{T}(\mathcal{T}_h;\mu)$.

3.2 Weak dG-Trefftz formulation

Testing trial functions $(v,\sigma) \in \mathbb{T}^p(K;\mu)$ on a mesh element $K \in \mathcal{T}_h$ with test functions $(\omega,\tau) \in \mathbb{T}^p(K;\mu)$ yields

$$\int_K \left[\Big(\frac{\partial \sigma}{\partial t} + \frac{\partial v}{\partial x} \Big) \tau \;+\; \Big(\mu^{-2}\frac{\partial v}{\partial t} + \frac{\partial \sigma}{\partial x} \Big) \omega \right] dK \;=\; 0 \ ,$$

due to the Trefftz property. With the unit outer normal vector $(n_K^t, n_K^x) \in \mathbb{R}^2$, integration by parts gives

$$\int_{\partial K} \left[\left(\sigma\tau + \mu^{-2} v\omega\right) n_K^t \;+\; \left(v\tau + \omega\sigma\right) n_K^x \right] d(\partial K) \;=\; 0 \ ,$$

as the second volume term also vanishes because of the Trefftz property. Summing over all elements results in the dG-Trefftz weak formulation.

As Trefftz functions do not have to be continuous across element boundaries, one replaces the unknowns on the element boundaries by numerical fluxes, also in order to ensure global continuity of the approximation. To this end, the mesh skeleton $\mathcal{F}_h$ is subdivided into the internal faces $\mathcal{F}_h^{\text{time}}$ parallel to the time axis, all other internal faces $\mathcal{F}_h^{\text{space}}$, the temporal outer faces $\mathcal{F}_h^0$ at $t=0$, the temporal outer faces $\mathcal{F}_h^{\text{T}}$ at the terminal time $t=T$ and the Dirichlet faces $\mathcal{F}_h^{\text{D}}$. With the spatial outer normal vector n_Ω^x, the average $\{\cdot\}$, the temporal normal jump $[[\cdot]]_t$, the spatial normal jump $[[\cdot]]_x$ and the trace of the function v at earlier times v^-, the chosen fluxes read

$$\begin{array}{ll}
\left(v^-,\ \sigma^-\right) & \text{on } \mathcal{F}_h^{\text{space}}, \\
\left(\{v\} + \tfrac{1}{2}[[\sigma]]_x,\ \{\sigma\} + \tfrac{1}{2}[[v]]_x\right) & \text{on } \mathcal{F}_h^{\text{time}}, \\
(v_0,\ \sigma_0) & \text{on } \mathcal{F}_h^{0}, \\
(v,\ \sigma) & \text{on } \mathcal{F}_h^{\text{T}}, \\
\left(g_D,\ \sigma + \tfrac{1}{2}(v - g_D) n_\Omega^x\right) & \text{on } \mathcal{F}_h^{\text{D}}\ .
\end{array}$$

These terms correspond to upwind fluxes on $\mathcal{F}_h^{\text{space}}$ and centered fluxes with jump penalization on $\mathcal{F}_h^{\text{time}}$. Eventually, the space-time dG-Trefftz problem then amounts finding $(v^N(\mu),\ \sigma^N(\mu)) \in U^N(\mu) := \mathbb{T}^p(\mathcal{T}_h;\mu)$, where $N = \dim(\mathbb{T}^p(\mathcal{T}_h;\mu))$, such that

$$\mathcal{A}\left((v^N(\mu),\ \sigma^N(\mu)),\ (\omega,\ \tau);\mu\right) \;=\; \ell\left((\omega,\ \tau);\mu\right) \quad \forall\, (\omega,\ \tau) \in U^N(\mu), \tag{4}$$

with

$$\begin{aligned}
\mathcal{A}\left((v^N(\mu),\ \sigma^N(\mu)),\ (\omega,\ \tau);\mu\right) = & \int_{\mathcal{F}_h^{\text{space}}} \Big[\left(v^N(\mu)\right)^- [[\tau]]_x + \mu^{-2} \left(v^N(\mu)\right)^- [[\omega]]_t + \left(\sigma^N(\mu)\right)^- [[\omega]]_x + \left(\sigma^N(\mu)\right)^- [[\tau]]_t \Big]\, d\mathcal{F}_h \\
& + \int_{\mathcal{F}_h^{\text{time}}} \left[\left\{v^N(\mu)\right\} [[\tau]]_x + \tfrac{1}{2} \left[\left[\sigma^N(\mu)\right]\right]_x [[\tau]]_x + \left\{\sigma^N(\mu)\right\} [[\omega]]_x + \tfrac{1}{2} \left[\left[v^N(\mu)\right]\right]_x [[\omega]]_x \right] d\mathcal{F}_h \\
& + \int_{\mathcal{F}_h^{\text{T}}} \left[\mu^{-2} v^N(\mu)\omega + \sigma^N(\mu)\tau\right] d\mathcal{F}_h \;+\; \int_{\mathcal{F}_h^{\text{D}}} \left(\tfrac{1}{2} v^N(\mu) + \sigma^N(\mu) n_\Omega^x\right) \omega \, d\mathcal{F}_h
\end{aligned}$$

and

$$\ell\left((\omega,\ \tau);\mu\right) \;=\; \int_{\mathcal{F}_h^{0}} \left[\mu^{-2} v_0 \omega + \sigma_0 \tau\right] d\mathcal{F}_h \;+\; \int_{\mathcal{F}_h^{\text{D}}} g_D \left(\tfrac{1}{2}\omega - \tau n_\Omega^x\right) d\mathcal{F}_h \ .$$

Well-posedness of (4) is proven in [2] with respect to a specific Trefftz space and a dG-norm depending on the wave speed μ. Note that in (4), not only the bilinear form and the right-hand side, but also the spaces are parameter-dependent, i.e. $U^N(\mu)$ instead of U^N in the above described RBM framework.

3.3 Linear dG-Trefftz model order reduction

The starting point is that $\dim(\mathbb{T}^p(\mathcal{T}_h;\mu)) =: N$ becomes large with finer discretization $\mathcal{T}_h$ and larger polynomial degree p. Then, (4) is a high-dimensional problem that is costly to solve. If we would follow the standard RBM recipe mentioned in §2 above, one would determine $n \ll N$ wave speed parameters $\mu^{(1)}, \dots, \mu^{(n)} \in \mathcal{P}$ by maximizing an a posteriori error estimator Δ_n over a finite training set $\mathcal{P}_{\text{train}} \subset \mathcal{P}$ and determining the snapshots by solving the high-dimensional problem (4) with respect to the sample values $\mu^{(i)}$, providing

$$\xi_1 := \big(v^N(\mu^{(1)}),\, \sigma^N(\mu^{(1)})\big) \in \mathbb{T}^p(\mathcal{T}_h;\mu^{(1)}),\, \dots,\, \xi_n := \big(v^N(\mu^{(n)}),\, \sigma^N(\mu^{(n)})\big) \in \mathbb{T}^p(\mathcal{T}_h;\mu^{(n)}). \tag{5}$$

These n high-dimensional (detailed) solutions are used as a basis for a reduced space $U_n = \mathrm{span}\{\xi_1, ..., \xi_n\}$ spanned by them. Apparently, even though we used the dG-Trefftz method on parameter-dependent spaces $U^N(\mu^{(i)})$, we would obtain a parameter-independent (i.e., linear) reduced model. Hence, we have to expect to fall back into the poor decay of the Kolmogorov n-width. In fact, for any $\mu \in \mathcal{P}$, this reduced problem takes the form

$$(v_n(\mu),\, \sigma_n(\mu)) \;=\; \sum_{i=1}^{n} \eta_i(\mu)\, \big(v^N(\mu^{(i)}),\, \sigma^N(\mu^{(i)})\big) \quad \text{with} \quad \eta_1(\mu), \dots, \eta_N(\mu) \in \mathbb{R}\,. \tag{6}$$

However, doing so, the reduced approximation is searched in $U_n = \mathrm{span}\{\xi_1, ..., \xi_n\} \subset \mathbb{T}^p(\mathcal{T}_h;\mu^{(1)}) \oplus \cdots \oplus \mathbb{T}^p(\mathcal{T}_h;\mu^{(n)})$, which is in general not contained in $\mathbb{T}^p(\mathcal{T}_h;\mu)$, i.e. $u_n(\mu) \notin \mathbb{T}^p(\mathcal{T}_h;\mu)$ for an online given $\mu \in \mathcal{P}$! Therefore, (4) does not hold true in the reduced space. In order to still obtain a reasonable reduced approximation, we optimize the $L^2(Q)^2$-norm distance of (6) to the high-dimensional solution, i.e. for $\mu \in \mathcal{P}$ we set

$$(v_n(\mu),\, \sigma_n(\mu)) \;:=\; \operatorname*{argmin}_{\eta_1(\mu),\dots,\eta_n(\mu)\in\mathbb{R}} \Big\| \big(v^N(\mu),\, \sigma^N(\mu)\big) - \sum_{i=1}^{n} \eta_i(\mu)\, \xi_i \Big\|_{L^2(Q)^2}\,. \tag{7}$$

As (7) is equivalent to a system of linear equations consisting of the Gramian matrix of (5) with respect to the $L^2(Q)^2$-norm, it has a unique solution.

The Trefftz space specific dG-norm depends on the wave speed and is therefore only meaningful for functions from the Trefftz space corresponding to that parameter value. Clearly, (7) is only sort of a "workaround" and is used for comparison, also since it requires the high-dimensional solution.

3.4 Nonlinear model reduction via exchanging bases

As an alternative, we consider a nonlinear model reduction by using parameter-dependent reduced spaces. To this end, for $\mu \in \mathcal{P}$ let

$$\mathrm{span}\Big\{ \big(\varphi^{(v)}_{\mu,1},\, \varphi^{(\sigma)}_{\mu,1}\big)\,,\; \dots\,,\; \big(\varphi^{(v)}_{\mu,N},\, \varphi^{(\sigma)}_{\mu,N}\big) \Big\} \;=\; \mathbb{T}^p(\mathcal{T}_h;\mu)$$

be a basis of the N-dimensional μ-dependent space $\mathbb{T}^p(\mathcal{T}_h;\mu)$. The representation of the high-dimensional solution with respect to μ then reads

$$\big(v^N(\mu),\, \sigma^N(\mu)\big) \;=\; \sum_{i=1}^{N} \zeta_i(\mu)\, \big(\varphi^{(v)}_{\mu,i},\, \varphi^{(\sigma)}_{\mu,i}\big) \quad \text{with} \quad \zeta_1(\mu), \dots, \zeta_N(\mu) \in \mathbb{R}\,.$$

Let us now assume that we are given a new parameter $\tilde{\mu}$ for which we aim at determining an approximation. The idea is to transform the detailed solution with respect to $\mu \in \mathcal{P}$ from $\mathbb{T}^p(\mathcal{T}_h;\mu)$ into $\mathbb{T}^p(\mathcal{T}_h;\tilde{\mu})$ by keeping the expansion coefficients but replacing the basis elements. The resulting function

$$\big(v^N,\, \sigma^N\big)_{\mu\to\tilde{\mu}} \;:=\; \sum_{i=1}^{N} \zeta_i(\mu)\, \big(\varphi^{(v)}_{\tilde{\mu},i},\, \varphi^{(\sigma)}_{\tilde{\mu},i}\big) \;\in \mathbb{T}^p(\mathcal{T}_h;\tilde{\mu})$$

is of course different from $\big(v^N(\tilde{\mu}),\, \sigma^N(\tilde{\mu})\big)$, but we could hope that it is a good approximation of the latter.

For $\mu \in \mathcal{P}$, a reduced solution is now defined as the superposition of transferred detailed solutions with regard to given sample values $\mu^{(1)}, \dots, \mu^{(n)} \in \mathcal{P}$, i.e.,

$$(v_n(\mu),\, \sigma_n(\mu)) \;=\; \sum_{i=1}^{n} \eta_i(\mu)\, \big(v^N,\, \sigma^N\big)_{\mu^{(i)}\to\mu} \quad \text{with} \quad \eta_1(\mu), \dots, \eta_n(\mu) \in \mathbb{R}\,. \tag{8}$$

Doing so, we obtain an approximation from a reduced space defined as

$$U_n(\mu) := \operatorname{span}\left\{ \left(v^N, \sigma^N\right)_{\mu^{(1)}\to\mu}, \ldots, \left(v^N, \sigma^N\right)_{\mu^{(n)}\to\mu} \right\} \subset \mathbb{T}^p(\mathcal{T}_h; \mu),$$

so that the high-dimensional problem (4) can be shifted into the low-dimensional space $U_n(\mu)$. The corresponding reduced approximation is obtained by finding $(v_n(\mu), \sigma_n(\mu)) \in U_n(\mu)$ such that

$$\mathcal{A}\left((v_n(\mu), \sigma_n(\mu)), (\omega, \tau); \mu\right) = \ell\left((\omega, \tau); \mu\right) \quad \forall (\omega, \tau) \in U_n(\mu)$$

and can easily be solved for small sizes n. Note, that for every $\mu \in \mathcal{P}$ the reduced approximation is determined in a different (μ-dependent) reduced space. Hence, the Kolmogorov n-width is no longer a meaningful benchmark.

As above, we use the strong greedy method to determine the samples $\mu^{(i)}$ and the corresponding snapshots. The error is determined in $L^2(Q)^2$ as different μ-dependent norms cannot be compared.

4 Numerical results

We describe results of some numerical experiments in order to examine the decay rate for the introduced model reduction techniques as n increases. To this end, we consider two cases.

Case 1. We determine a reduced approximation for the entire parameter set $\mathcal{P} = [0.3, 2]$, incorporating distinct behaviors of the problem (with and without reflection of the wave).

Case 2. In order to take the different behaviors into account, we follow [7] to partition the parameter set. We investigate a split of $\mathcal{P}$ into $\mathcal{P}_1 = [1.3, 1]$ (the pure transport behavior) and $\mathcal{P}_2 = [1, 2]$, where the wave is reflected.

All experiments were carried out on a Cartesian mesh with mesh width $h_{\text{mesh}} = 0.005$. All three parameter sets were discretized by 60 equidistant points. The polynomial degree of the Trefftz spaces was set to $p = 3$. The high-dimensional problems were solved with NGSolve using the add-on NGSTrefftz, [11].

Figure 2 provides the results. The exchanging bases approach demonstrates a convergence behavior that is comparable to the linear one, it does not perform better than the linear approach in terms of its rate of approximation. This applies to the full problem as well as to the ones with split parameter set.

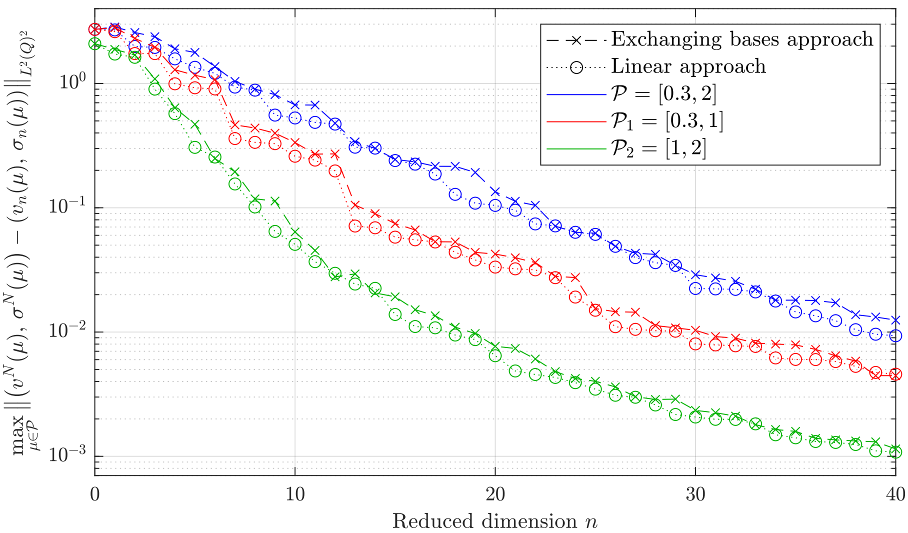

Fig. 2: Strong greedy error of the linear and the exchanging bases approach regarding parameter sets that cover various problem scenarios.

The strong greedy error of the nonlinear transforming-bases approach applied to the full parameter set $\mathcal{P} = [0.3, 2]$ seems to first form a plateau and then converges at a polynomial rate. In fact, Figure 3 suggests that the convergence rate is between n^{-3} and $n^{-3.5}$. This is supported by the fact that the least-squares rate for n between 10 and 40 is $n^{-3.07}$ and for n between 20 and 40 is $n^{-3.44}$. This is in line with the results in [6], where a rate of $n^{-7/2}$ has been observed numerically.

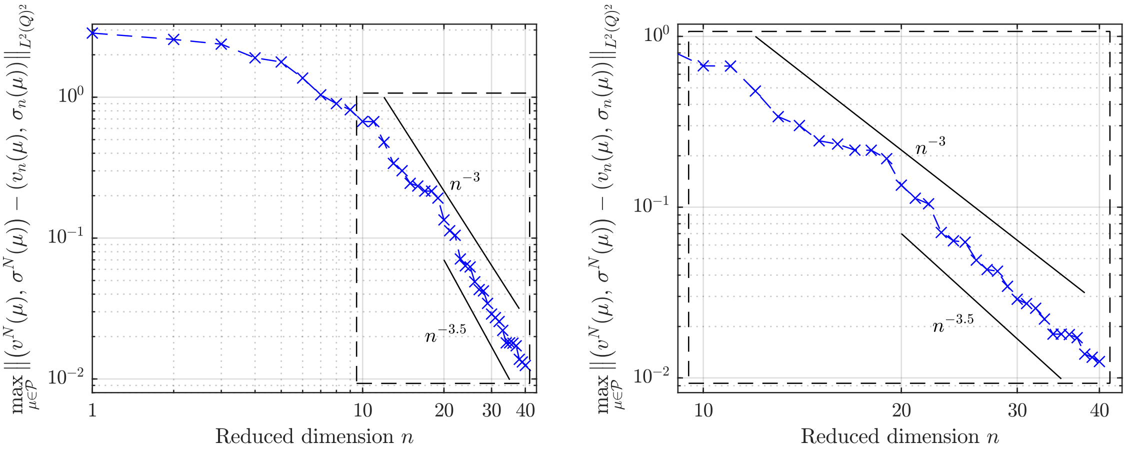

Fig. 3: Strong greedy error of the transforming-bases approach applied on $\mathcal{P} = [0.3, 2]$. The plot on the right shows the indicated cutout of the plot on the left.

5 Nonlinear model reduction using an autoencoder

In order to compare the results of the dG-Trefftz model reduction, we also pursued a second nonlinear model reduction strategy, which has been shown in the literature to yield good results. The starting point is that every model order reduction scheme can be formulated on an abstract level as a composition of an encoder and a decoder function, see [3]. Moreover, in that paper it was also shown that combinations of encoder and decoder exist such that the resulting nonlinear model reduction can outperform the Kolmogorov n-width.

5.1 Model reduction through encoder and decoder

An encoder is a mapping $E : \mathbb{R}^N \to \mathbb{R}^n$ so that $E(u^N(\mu))$ can be described in terms of $n \ll N$ degrees of freedom forming the reduced model. The decoder $D : \mathbb{R}^n \to \mathbb{R}^N$ maps a reduced approximation back to the high-dimensional space. Hence, one would want that $u^N(\mu) \approx D \circ E(u^N(\mu))$ and the reduced approximation reads $u_n(\mu) := E(u^N(\mu))$. In the online phase, one would evaluate the mapping $\mu \mapsto E(u^N(\mu))$ without going to the high-dimensional representation.

A corresponding model order reduction framework was introduced in [5] and modified later in [8], that aims to project dynamical systems onto low-dimensional nonlinear manifolds and approximate the corresponding encoder and decoder functions using deep convolutional autoencoders. In [5], the considered state is the solution of the underlying problem, evaluated in the spatial domain Ω at a specific point in time and for a specific parameter value. In that setting, N is the number of spatial discretization points. Then, a convolutional neural network (CNN) is trained to derive encoder and decoder. This framework was modified in [8] for problems in space and time. For this purpose, a deep feedforward neural network $M : I \times \mathcal{P} \to \mathbb{R}^n$ w.r.t. time and parameter has been used to reflect the dynamics on the low-dimensional manifold.

We adapt this to the problem (1) we are considering here. To this end, let as before $(v^N(\mu), \sigma^N(\mu))$ be the high-dimensional solution of the wave problem w.r.t. $\mu \in \mathcal{P}$. We used two different instances of the CNNs yielding E_v, M_v, D_v for $v^N(\mu)$ and E_σ, M_σ, D_σ w.r.t. $\sigma^N(\mu)$. We detail our proceeding for $v^N(\mu)$ and remark that $\sigma^N(\mu)$ can be treated in an analogous fashion.

With discretizations of the time interval $\{t_1, \dots, t_{N_t}\} \subset [0, T]$ and the parameter set $\{\mu_1, \dots, \mu_{N_\mathcal{P}}\} \subset \mathcal{P}$, the loss function for the training reads

$$\sum_{i=1}^{N_t}\sum_{j=1}^{N_\mathcal{P}} \left\| v^N(t_i, \cdot; \mu_j) - D_v\left(E_v\left(v^N(t_i, \cdot; \mu_j)\right)\right)\right\|_2^2 + \left\|E_v\left(v^N(t_i, \cdot; \mu_j)\right) - M_v(t_i, \mu_j)\right\|_2^2 , \tag{9}$$

where $v^N(t_i, \cdot; \mu_j) \in \mathbb{R}^N$ is the evaluation of the high-dimensional solution on the entire spatial discretization at time t_i for the parameter value μ_j and $\|\cdot\|_2$ denotes the Euclidean norm on $\mathbb{R}^N$. The idea behind the loss function (9) is to train the autoencoder property of $D_v(E_v(\cdot))$ and at the same time also the representation of the dynamics on the manifold by M_v.

In the online phase, only the map $D_v(M_v) : [0, T] \times \mathcal{P} \to \mathbb{R}^N$ defined by $(t, \mu) \mapsto D_v(M_v(t, \mu))$ is evaluated and this provides the evaluation of the reduced solution w.r.t. the parameter value μ on the full spatial discretization at time t.

5.2 Numerical results

We adopted the architecture of the NNs, the learning strategy and their implementation from [8]. The spatial domain $\Omega := (0, 1)$ and the time domain $I := [0, 1]$ were both discretized with 256 equidistant points, so that $N = N_t = 256$. For the training, the parameter set $\mathcal{P} = [0.3, 2]$ was divided into 20 equidistant points, i.e. $N_\mathcal{P} = 20$.

The training data was chosen to consist of $N_t \cdot N_{\mathcal{P}} = 5120$ triples $\big(t,\, \mu,\, v^N(t,\cdot;\mu)\big)$ for the first component of the solution and $\big(t,\, \mu,\, \sigma^N(t,\cdot;\mu)\big) \in \{t_1, \ldots, t_{N_t}\} \times \{\mu_1, \ldots, \mu_{N_{\mathcal{P}}}\} \times \mathbb{R}^N$ for the second one. We chose a learning rate of 0.0001 either until 10000 learning steps had been completed or 500 consecutive learning steps had not resulted in a change in the solution to the minimization problem. For the validation, 21 equidistant parameter values between 0.35 and 1.95 lying in-between the parameter values for the training were selected.

Figure 4 shows the $L^2(Q)^2$-errors at the validation points for instances of the framework with different reduced dimensions n. Note, that the training involves randomized steps and the displayed errors refer to instances of the framework that were only trained once.

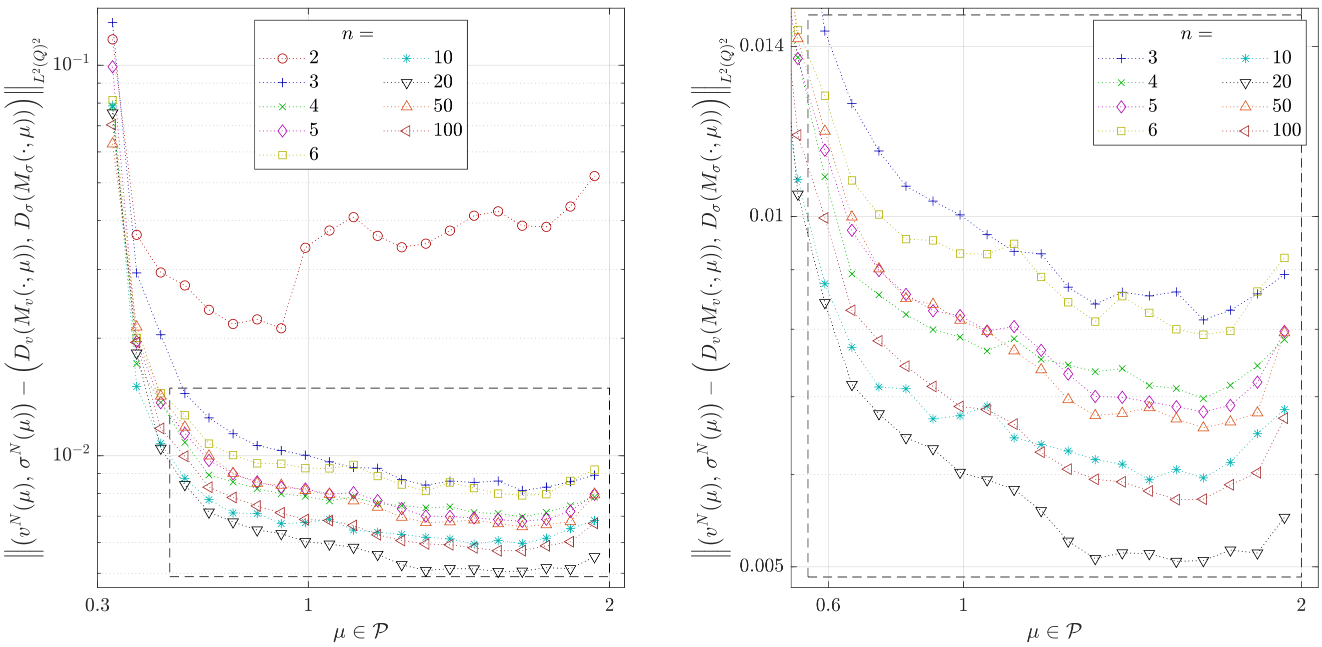

Fig. 4: Reduction error of the autoencoder model reduction for different sizes n of the reduced system, evaluated at the validation points. The plot on the right shows the indicated cutout of the plot on the left.

It can be seen that the approach struggles with the pure transport behavior, whereas for the reflective behavior the error is consistently better. The reduction error only improves slowly for increasing reduced dimension n and occasionally even gets worse. Moreover, we see that the method does not converge monotonically as n grows. Finally, comparing the achieved accuracies with the previous approaches, we observe that the nonlinear autoencoder method does not outperform the dG-Trefftz methods.

6 Conclusions and outlook

In this paper, we introduce linear and nonlinear model reduction techniques based upon a dG-Trefftz discretization for a model problem of a traveling wave. We show results of numerical methods and also compare the dG-Trefftz approach with a nonlinear model reduction technique based upon autoencoders. We summarize some of our observations:

- The nonlinear dG-Trefftz model order reduction approach via exchanging bases shows a convergence comparable to linear approaches. This applies for the whole parameter set (representing waves with and without reflection) as well as subsets. Hence, using dG-Trefftz alone cannot overcome the obstructions of the slow decay of the Kolmogorov n-width.
- We do observe faster convergence than predicted by the (worst-case) analysis, but still only a polynomial decay.
- The autoencoder-based model reduction does not reach better accuracies than the dG-Trefftz method. This also confirms that the considered model problem is in fact a challenging one.
- The autoencoder approach displays a considerable reduction power for very small reduced dimensions. However, for larger reduced dimensions the reduction errors decay slowly and soon appear to reach a level at which they no longer improve substantially.

We conclude that both approaches, just as they are, are not suitable for model reduction of linear hyperbolic problems. Moreover, we also see that using a dG-Trefftz method “just” as detailed solution does not even allow for a subsequent model reduction by a standard approach as the parameter is deeply involved in the discretization and cannot be separated easily from the primitive variables space and time.

On the other hand, it still seems to be promising to use dG-Trefftz methods within the model reduction chain as they are known to yield highly efficient solvers for linear hyperbolic problems. One approach might be a combination of the autoencoder and the dG-Trefftz method. In fact, so far, the autoencoder approach is purely data-based and does not exploit any knowledge about the structure of the underlying problem. However, if one chooses the reduced solution to be a Trefftz function and employs a similar autoencoder scheme to the coefficients of its basis representation, one might be able to take advantage of the information about the underlying problem contained in the Trefftz property.

Moreover, we will investigate the question if a dG-Trefftz method can be used to learn the nonlinear mapping $u^N(\mu) \mapsto u^N(\tilde{\mu})$, which would yield a possibly efficient model reduction technique as it is known that the solution of the traveling wave problem follows the characteristics of the equation.

References

[1] J. S. Hesthaven, G. Rozza, and B. Stamm. *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*. Springer, 2016.

[2] A. Moiola and I. Perugia, A space-time Trefftz discontinuous Galerkin method for the acoustic wave equation in first-order formulation, Numer. Math. **138**, pp. 389–435 (2018).

[3] A. Cohen, C. Farhat, Y. Maday, and A. Somacal, Nonlinear compressive reduced basis approximation for PDE's, C.R. Mécanique **351**, pp. 357–374 (2024).

[4] C. Greif and K. Urban, Decay of the Kolmogorov N-width for wave problems. *Appl. Math. Lett.* **96**, pp. 216–222 (2019).

[5] K. Lee, and K. T. Carlberg, Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders, J. Comput. Phys. **404** (2020).

[6] S. Glas, A. T. Patera, and K. Urban, A reduced basis method for the wave equation, Int. J. Comput. Fluid Dyn. **34**, pp. 139–146 (2020).

[7] J. L. Eftang, *Reduced basis methods for parametrized partial differential equations*, Doctoral Thesis, Norwegian University of Science and Technology Trondheim (2011).

[8] S. Fresca, L. Dede', and A. Manzoni, A Comprehensive Deep Learning-Based Approach to Reduced Order Modeling of Nonlinear Time-Dependent Parametrized PDEs, J. Sci. Comput. **87** (2021).

[9] M. Ohlberger and S. Rave, Reduced basis methods: Success, limitations and future challenges. *Proceedings of the Conference Algoritmy*, pp. 1–12 (2016).

[10] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced basis methods for partial differential equations: An introduction*. Springer, 2016.

[11] P. Stocker, 'NGSTrefftz': Add-on to NGSolve for Trefftz methods, J. Open Source Software **7** (2022).

[12] K. Urban, The Reduced Basis Method in Space and Time: Challenges, Limits and Perspectives, in: Model Order Reduction and Applications, M. Falcone and G. Rozza (eds.), Cetraro 2021, pp. 1–73, C.I.M.E. Series, Springer, 2023.