The Kolmogorov N-width for linear transport: Exact representation and the influence of the data

Florian Arbes[1*], Constantin Greif[2], Karsten Urban[2]

[1]Computational Materials Processing, IFE, Institute for Energy Technology, Instituttveien 18, Kjeller, 2007, Norway.
[2]Institute of Numerical Mathematics, Ulm University, Helmholtzstr. 20, Ulm, 89081, Germany.

Contributing authors: florian.arbes@ife.no; constantin.greif@uni-ulm.de; karsten.urban@uni-ulm.de;

The Kolmogorov N-width describes the best possible error one can achieve by elements of an N-dimensional linear space. Its decay has extensively been studied in Approximation Theory and for the solution of Partial Differential Equations (PDEs). Particular interest has occurred within Model Order Reduction (MOR) of parameterized PDEs e.g. by the Reduced Basis Method (RBM).

While it is known that the N-width decays exponentially fast (and thus admits efficient MOR) for certain problems, there are examples of the linear transport and the wave equation, where the decay rate deteriorates to $N^{-1/2}$. On the other hand, it is widely accepted that a smooth parameter dependence admits a fast decay of the N-width. However, a detailed analysis of the influence of properties of the data (such as regularity or slope) on the rate of the N-width seems to lack.

In this paper, we use techniques from Fourier Analysis to derive exact representations of the N-width in terms of initial and boundary conditions of the linear transport equation modeled by some function g for half-wave symmetric data. For arbitrary functions g, we derive bounds and prove that these bounds are sharp. In particular, we prove that the N-width decays as $c_r N^{-(r+1/2)}$ for functions in the Sobolev space, $g \in H^r$. Our theoretical investigations are complemented by numerical experiments which confirm the sharpness of our bounds and give additional quantitative insight.

[*]STIPINST funding [318024] from the Research Council of Norway is gratefully acknowledged.

1 Introduction

The Kolmogorov N-width describes the best possible error one can achieve by a linear approximation with $N \in \mathbb{N}$ degrees of freedom, i.e. by elements of the best possible N-dimensional linear space, [1]. The arising optimal space in the sense of Kolmogorov can often not explicitly be constructed, at least not in a reasonable (computing) time. On the other hand, however, the decay rate of the N-width tells us if a given set can be well-approximated by a linear method, or not. This is a classical question in Approximation Theory and has been widely studied in the literature, see e.g. [2–5], this list being far from being complete.

Particular interest has been devoted to the case when the set to be approximated is given by solutions of certain equations, e.g. Partial Differential Equations (PDEs) with different data (one might think of the domain, coefficients, right-hand side loadings, initial- and/or boundary conditions), which might be considered as parameters, [6–8]. In that direction, model order reduction of parametric PDEs (PPDE) has become a field of very intensive research, also with many very relevant real-life applications, [9–11]. A prominent example is the Reduced Basis Method (RBM), where a PPDE is aimed to be reduced to an N-dimensional linear space in order to allow multi-query (w.r.t. different parameter values) and/or realtime (embedded systems, cold computing) applications. The reduced N-dimensional system is determined in an offline training phase using sufficiently accurate detailed numerical solutions by any standard method. In this framework, the question arises, if a given PPDE can be well-reduced by means of the RBM or not. Since it has been proven in [12] that the offline reduced basis generation using a Greedy method realizes the same asymptotic rate of decay as the Kolmogorov N-width, one is left with the investigation of the decay for PPDEs to decide whether the RBM is suitable for a given PPDE, or not. Also in that direction, there is a significant amount of literature, e.g. [13–22], just to name a few. Roughly speaking, it was shown there that a PPDE admits a fast decay of the N-width if the solution depends smoothly on the parameter, which is, e.g. known for elliptic and parabolic problems which allow for a separation of the parameters from the physical variables. As a rule of thumb: "holomorphic dependence admits exponential decay".

However, when leaving the *nice* realm of such PPDEs, the situation becomes dramatically worse. It has for example been shown that the decay may drop down to N^{-s}, $0 < s < 1$, for the linear transport equation [20] and the wave equation [23]. However, the problems considered in the latter papers are quite specific examples yielding to a non-smooth dependence of the solution in terms of the parameter (the velocity in [20] and the wave speed in [23]). It was also demonstrated that the decay not only depends on the PDE, but also on the underlying physics, e.g. alloy compositions in case of solidification problems [24]. For problems of such type (transport, transport-dominated, hyperbolic), the above quoted rule of thumb remains true.

This is why we are interested in the *exact* dependence of the decay rate of the N-width in terms of the data / parameters of the problem. To our own surprise we could not find corresponding results in the literature. In [18, 21, 22], the fast decay is shown using techniques from interpolation proving that a Greedy-type selection selects the optimal nodes. The positive result in [20, Thm. 3.1] has been deduced by using the decay of the complex power series.

We consider the linear transport problem whose solution is given by the characteristics in terms of initial and boundary conditions. Hence, we can reduce ourselves to approximate the mapping $x \mapsto g(x-\mu)$, where μ is the parameter and $x \in \Omega$, which is the domain on which the PPDE is posed; g is the real-valued univariate function modeling initial and boundary conditions. To this end, we use the Fourier series approximation, which allows us to incorporate the parametric shift by μ into the approximation spaces. We derive exact representations of the N-width for certain classes (half-wave symmetric -HWS- functions) and give estimates in terms of the regularity and the slope of the function g.

This paper is organized as follows. In Section 2, we collect preliminaries on the linear transport equation, the N-width and some facts from Fourier analysis, which we shall need in the sequel. The main tool for our analysis is a shift-isometric orthogonal decomposition (Def. 3.1) based upon the Fourier series of g. This notion is introduced in Section 3. In Section 4, we use Def. 3.1 to construct approximation spaces with which we can derive exact representations for the N-width for HWS functions and sharp estimates in the general case. The specific influence of the regularity of g on the decay of the N-width is investigated in Section 5, where we prove that the N-width is bounded by $c_r N^{-(r+1/2)}$ for functions in the Sobolev space H^r. In Section 6, using specifically constructed piecewise functions, we show that our bounds are sharp. Some results of our extensive numerical results are presented in Section 7. The paper finishes with some conclusions in Section 8.

2 Preliminaries

2.1 The linear transport equation

We consider the univariate linear transport equation with velocity μ, which is interpreted as a parameter, i.e., we seek a function $\Phi(\cdot,\cdot;\mu) : I \times \Omega \to \mathbb{R}$ such that[1]

$$\partial_t \Phi(t,x;\mu) + \mu\, \partial_x \Phi(t,x;\mu) = 0, \qquad (t,x) \in I \times \Omega, \tag{2.1a}$$
$$\Phi(0,x;\mu) = g(x), \qquad x \in \Omega, \tag{2.1b}$$
$$\Phi(t,0;\mu) = g(-\mu\, t), \qquad t \in I, \tag{2.1c}$$

where $I := (0,1)$ is the time interval and $\Omega := (0,1)$ the spatial domain.[2] The velocity can be chosen in a compact interval $\mu \in \mathcal{P} := [0,1]$. The initial and boundary conditions, (2.1b) and (2.1c), respectively, are given in terms of a function $g : \Omega_{\mathcal{P}} \to \mathbb{R}$, whose properties will be relevant in the sequel. Here, $\Omega_{\mathcal{P}} := \Omega \ominus \mathcal{P} := \{x - \mu : x \in \Omega, \mu \in \mathcal{P}\} = (-1,1)$ is the domain on which g needs to be defined in order to obtain a well-posed problem (2.1) for every parameter. The solution of (2.1) is well-known to read $\Phi(t,x;\mu) = g(x-\mu\, t)$, $(t,x) \in I \times \Omega$. We are particularly interested in the solution at the final time $t=1$, i.e.,

$$u_\mu(x) := \Phi(1,x;\mu) = g(x-\mu), \qquad x \in \Omega, \tag{2.2}$$

[1] We restrict ourselves to the homogeneous case for simplicity. One could also consider a right-hand side $f(t,x;\mu) \not\equiv 0$, which would also impact the rate of approximation of the solution.

[2] Our analysis is restricted to the 1D-case, but some results can be extended to higher dimensions.

and consider the low regularity case, i.e., we only assume that $g \in L_2(\Omega_{\mathcal{P}})$, and therefore $u_\mu \in L_2(\Omega)$, no additional smoothness.

Remark 2.1. *Often, the time $t \in I$ is also seen as a parameter. But our considerations are not restricted to the final time, since then for a given $t \in I$ and $\mu \in \mathcal{P}$, we can define the new parameter $\tilde{\mu} := t\,\mu \in [0,1)$ and get $u_{\tilde{\mu}} = g(\cdot - \tilde{\mu}) = \Phi(t, \cdot; \mu)$.*

2.2 Linear approximation: The N-width

The specific focus of this paper is the approximation rate provided by linear subspaces. In particular, we are considering N-dimensional subspaces which are "optimal" to approximate u_μ for all parameters $\mu \in \mathcal{P}$ in an appropriate manner. The maybe most classical setting is the worst case scenario w.r.t. the parameter yielding the classical *Kolmogorov N-width* defined as, [1]

$$\begin{aligned} d_N(\mathcal{P}) &:= \inf_{\substack{V_N \subset L_2(\Omega) \\ \dim(V_N)=N}} \sup_{\mu \in \mathcal{P}} \inf_{\tilde{v}_N \in V_N} \|u_\mu - \tilde{v}_N\|_{L_2(\Omega)} \\ &= \inf_{\substack{V_N \subset L_2(\Omega) \\ \dim(V_N)=N}} \operatorname{dist}(V_N, \mathcal{U}_g(\mathcal{P}))_{L_\infty(\mathcal{P}; L_2(\Omega))}, \end{aligned} \tag{2.3}$$

where $\mathcal{U}_g(\mathcal{P}) := \{u_\mu \equiv g(\cdot - \mu) : \mu \in \mathcal{P}\} \subset L_2(\Omega)$ is also (slightly misleading) called *solution manifold*. The dependence on g will be crucial below.

Remark 2.2. *There are several results concerning the decay of $d_N(\mathcal{P})$ for the linear transport problem (2.1).*

(a) In [20] it was shown that $d_N(\mathcal{P}) \cong N^{-1/2}$, i.e., very slow, for the specific choice $g = \chi_{[0,1]}$, namely initial and boundary conditions involving a jump.

(b) On the other hand, one can achieve exponential decay, i.e., $d_N(\mathcal{P}) \lesssim e^{-\alpha N}$ for some $\alpha > 0$ if the function g is analytic. This can be seen by considering a truncated power series in the complex plane, [20, Thm. 3.1].

Our main focus in this paper is to study the decay of the N-width w.r.t. properties of the function g, in particular we want to detail the influence of the regularity of g on the decay of the N-width. In addition to the "worst-case in the parameter" Kolmogorov N-width $d_N(\mathcal{P})$, which measures the error in $L_\infty(\mathcal{P}; L_2(\Omega))$, we will also consider "mean-squared error in the parameter", i.e., $L_2(\mathcal{P}; L_2(\Omega))$ w.r.t. a probability measure (i.e., $d\mu(\mathcal{P}) = 1$), which we call *L_2-average N-width* defined as

$$\begin{aligned} \delta_N(\mathcal{P}) &:= \inf_{\substack{V_N \subset L_2(\Omega) \\ \dim(V_N)=N}} \left\{ \int_{\mathcal{P}} \inf_{\tilde{v}_N \in V_N} \|u_\mu - \tilde{v}_N\|^2_{L_2(\Omega)} \, d\mu \right\}^{1/2} \\ &= \inf_{\substack{V_N \subset L_2(\Omega) \\ \dim(V_N)=N}} \operatorname{dist}(V_N, \mathcal{U}_g(\mathcal{P}))_{L_2(\mathcal{P}; L_2(\Omega))}. \end{aligned} \tag{2.4}$$

Remark 2.3. *For later reference, we collect some equivalent expressions.*

(a) Let $P_N : L_2(\Omega) \to V_N$ denote the orthogonal projection onto V_N. Then,

$$d_N(\mathcal{P}) = \inf_{\substack{V_N \subset L_2(\Omega)\\ \dim(V_N)=N}} \|u_\mu - P_N u_\mu\|_{L_\infty(\mathcal{P};L_2(\Omega))}, \tag{2.5a}$$

$$\delta_N(\mathcal{P}) = \inf_{\substack{V_N \subset L_2(\Omega)\\ \dim(V_N)=N}} \|u_\mu - P_N u_\mu\|_{L_2(\mathcal{P};L_2(\Omega))}. \tag{2.5b}$$

(b) By $\|w\|_{L_2(\mathcal{P})} \le d\mu(\mathcal{P})\, \|w\|_{L_\infty(\mathcal{P})}$, $w \in L_\infty(\mathcal{P})$, we get $\delta_N(\mathcal{P}) \le d_N(\mathcal{P})$.

2.3 Fourier Analysis

Our major tool for determining the decay of the N-widths is Fourier Analysis. We collect the main ingredients needed for the sequel of this paper. Recall that for the above model problem, we have $I = (0, 1)$, $\Omega = (0, 1)$, $\mathcal{P} = [0, 1]$ and $\Omega_{\mathcal{P}} = (-1, 1)$, but the analysis is not restricted to that case. Here, we consider the space $L_2(\Omega_{\mathcal{P}})$ corresponding to signals of wave-length 2. Thus, the half-wave length is 1, which is used in the following definition, whose notion is well-known in electrical engineering (see e.g. [25]) and turns out to be crucial for the subsequent analysis.

Definition 2.4. *We call $g \in L_2(\Omega_{\mathcal{P}})$* even half-wave symmetric (even HWS, $g \in L_2^{\mathrm{evn}}$)*, if $g(x) = g(x + 1)$ for all $x \in [-1, 0]$ a.e., and* odd half-wave symmetric (odd HWS, $g \in L_2^{\mathrm{odd}}$)*, if $g(x) = -g(x + 1)$ for all $x \in [-1, 0]$ a.e. A function is called* half-wave symmetric (HWS, $g \in L_2^{\mathrm{hws}}$)*, if it is either even or odd HWS.*

We shall use the Fourier series of any $L_2(\Omega_{\mathcal{P}})$-function, namely

$$g(x) = \hat{A}_0 + \sum_{k=1}^{\infty} \Big[\hat{A}_k \cos(k\pi x) + \hat{B}_k \sin(k\pi x)\Big], \quad x \in \Omega_{\mathcal{P}} \text{ a.e.}, \tag{2.6}$$

where the Fourier coefficients are known as $\hat{A}_0 = \frac{1}{2}\int_{\Omega_{\mathcal{P}}} g(x)\, dx$ and for $k \ge 1$ they read $\hat{A}_k = \langle g, \cos(k\pi\cdot)\rangle_{L_2(\Omega_{\mathcal{P}})}$ as well as $\hat{B}_k = \langle g, \sin(k\pi\cdot)\rangle_{L_2(\Omega_{\mathcal{P}})}$. Using the Fourier expansion, it can readily be seen that any $g \in L_2(\Omega_{\mathcal{P}})$ can be decomposed into an even HWS and and odd HWS part, i.e., $g = g^{\mathrm{evn}} + g^{\mathrm{odd}}$, where

$$g^{\mathrm{evn}} := \hat{A}_0 + \sum_{k=1}^{\infty} \Big[\hat{A}_{2k} \cos(2k\pi\cdot) + \hat{B}_{2k} \sin(2k\pi\cdot)\Big] \in L_2^{\mathrm{evn}},$$

$$g^{\mathrm{odd}} := \sum_{k=1}^{\infty} \Big[\hat{A}_{2k-1} \cos((2k-1)\pi\cdot) + \hat{B}_{2k-1} \sin((2k-1)\pi\cdot)\Big] \in L_2^{\mathrm{odd}}.$$

We shall use this decomposition in order to determine the decay of the Kolmogorov N-width by splitting g into its even HWS and odd HWS part and then estimating the N-width for both of these parts. For later reference, we collect the facts

$$L_2(\Omega_{\mathcal{P}}) = L_2^{\mathrm{evn}} \oplus L_2^{\mathrm{odd}}, \qquad L_2^{\mathrm{hws}} = L_2^{\mathrm{evn}} \cup L_2^{\mathrm{odd}}, \qquad L_2^{\mathrm{hws}} \subsetneq L_2(\Omega_{\mathcal{P}}). \tag{2.7}$$

3 Shift-isometric orthogonal decompositions

We will construct spaces to bound or represent $d_N(\mathcal{P})$ and $\delta_N(\mathcal{P})$ for a given function g. It turns out that the cases $g \in L_2^{\text{evn}}$ and $g \in L_2^{\text{odd}}$ have to be treated separately. The arising shift-isometric orthogonal decompositions will depend on the given function g.

Definition 3.1. *Let $g \in L_2(\Omega_{\mathcal{P}})$; We call a family of subspaces $\mathcal{W}^g := \{W_k^g\}_{k\in\mathbb{N}} \subset L_2(\Omega)$ with the associated orthogonal projectors $\mathcal{Q}^g := \{Q_k^g\}_{k\in\mathbb{N}}$, $Q_k^g : L_2(\Omega) \to W_k^g$ a* shift-isometric orthogonal decomposition *(w.r.t. g) of $L_2(\Omega)$ if*

(i) $L_2(\Omega) = \bigoplus_{k=1}^{\infty} W_k^g$ and $W_k^g \perp W_j^g$ for $j \neq k$;

(ii) The orthogonal projectors are shift-isometric, *i.e., it holds*

$$\|Q_k^g g(\cdot - \mu)\|_{L_2(\Omega)} = \|Q_k^g g\|_{L_2(\Omega)} \quad \textit{for all } \mu \in \mathcal{P}. \tag{3.1}$$

Remark 3.2. *Note, that g needs to be defined on the larger space $\Omega_{\mathcal{P}}$ in order to apply $g(\cdot - \mu)$. However, the solution of the transport problem (2.1) is defined on the domain Ω. Whenever we take a norm $\|\cdot\|_{L_2(\Omega)}$ or apply Q_k^g, we consider implicitly the restriction $g_{|\Omega}$ or $g(\cdot - \mu)_{|\Omega}$ respectively.*

Now, we start by *assuming* that such a $\mathcal{W}^g$ exists and show that (3.1) is a key property. If $\mathcal{W}^g$ is a (shift-isometric) orthogonal decomposition of $L_2(\Omega)$, each $u_\mu \in \mathcal{U}_g(\mathcal{P}) \subset L_2(\Omega)$ has a unique decomposition $u_\mu = \sum_{k=1}^{\infty} Q_k^g u_\mu$. We define

$$V_N^g := \bigoplus_{k=1}^{M(N)} W_k^g, \quad \text{where } N = \dim(V_N^g) = \sum_{k=1}^{M(N)} \dim(W_k^g) \tag{3.2}$$

as a candidate for the best approximation space in the sense of Kolmogorov.[3] Clearly, the approximation $\tilde{v}_N := \sum_{k=1}^{M(N)} Q_k^g u_\mu$ converges to u_μ as $N \to \infty$, which implies that both $d_N(\mathcal{P})$ and $\delta_N(\mathcal{P})$ converge towards zero. Moreover, the orthogonality easily allows us to control the error.

Proposition 3.3. *Let $g \in L_2(\Omega_{\mathcal{P}})$, let $\mathcal{W}^g$ be a corresponding shift-isometric orthogonal decomposition of $L_2(\Omega)$ and let V_N^g be defined as in (3.2). Then,*

$$\operatorname{dist}(V_N^g, \mathcal{U}_g(\mathcal{P}))^2_{L_\infty(\mathcal{P};L_2(\Omega))} = \operatorname{dist}(V_N^g, \mathcal{U}_g(\mathcal{P}))^2_{L_2(\mathcal{P};L_2(\Omega))} = \sum_{k=M(N)+1}^{\infty} \|Q_k^g g\|^2_{L_2(\Omega)}. \tag{3.3}$$

Proof. The statement is a direct consequence of orthogonality, (3.1) and the fact that the orthogonal projection corresponds to the best approximation. Indeed, denoting the orthogonal projector onto V_N^g by P_N^g, we obtain

$$\inf_{\tilde{v}_N \in V_N^g} \|u_\mu - \tilde{v}_N\|^2_{L_2(\Omega)} = \|u_\mu - P_N^g u_\mu\|^2_{L_2(\Omega)} = \Big\| u_\mu - \sum_{k=1}^{M(N)} Q_k^g u_\mu \Big\|^2_{L_2(\Omega)}$$

[3] Note, that by fixing the dimensions of W_k^g a priori, it might happen that one might not be able to construct spaces V_N^g of any dimension $N \in \mathbb{N}$, since N must be a sum of the dimensions of the W_k^g, $k = 1, \ldots, M(N)$, see Remark 3.7. As an example, if $\dim(W_k^g) = 2$ for all $k \in \mathbb{N}$, N must be even.

$$= \Big\| \sum_{k=M(N)+1}^{\infty} Q_k^g u_\mu \Big\|_{L_2(\Omega)}^2 = \sum_{k=M(N)+1}^{\infty} \|Q_k^g u_\mu\|_{L_2(\Omega)}^2,$$

since $u_\mu \in \mathcal{U}_g(\mathcal{P}) \subset L_2(\Omega)$. Finally, by shift-isometry (Def. 3.1 (ii)), $\|Q_k^g u_\mu\|_{L_2(\Omega)} = \|Q_k^g g(\cdot - \mu)\|_{L_2(\Omega)} = \|Q_k^g g\|_{L_2(\Omega)}$ and we get that

$$\begin{aligned}
\operatorname{dist}(V_N^g, \mathcal{U}_g(\mathcal{P}))^2_{L_\infty(\mathcal{P};L_2(\Omega))} &= \sup_{\mu\in\mathcal{P}} \|u_\mu - P_N^g u_\mu\|_{L_2(\Omega)}^2 \\
&= \sup_{\mu\in\mathcal{P}} \sum_{k=M(N)+1}^{\infty} \|Q_k^g g\|_{L_2(\Omega)}^2 = \sum_{k=M(N)+1}^{\infty} \|Q_k^g g\|_{L_2(\Omega)}^2 \\
&= \int_{\mathcal{P}} \sum_{k=M(N)+1}^{\infty} \|Q_k^g g\|_{L_2(\Omega)}^2 d\mu = \operatorname{dist}(V_N^g, \mathcal{U}_g(\mathcal{P}))^2_{L_2(\mathcal{P};L_2(\Omega))},
\end{aligned}$$

since $\mathcal{P} = [0, 1]$, which completes the proof. □

Before continuing, we stress that for V_N^g both versions of the N-width coincide and that the right-hand side of (3.3) is independent of the parameter μ, which allows us to bound the N-width for all parameters. As the simple proof above shows, the shift-isometry (3.1) is the key property.

3.1 Optimal N-term approximation spaces

We will later construct shift-isometric decompositions by choosing appropriate bases for W_k^g for a given $g \in L_2^{\text{hws}}$. If an ON basis $\{v_1, v_2, ...\}$ of $L_2(\Omega)$ is known, one can consider one-dimensional spaces $W_k := \operatorname{span}\{v_k\}$ and then define $V_N := \oplus_{k=1}^N W_k$. So far, such W_k and V_N are independent of g. However, the following result shows that such spaces are optimal in the sense of Kolmogorov if the basis $\{v_1, v_2, ...\}$ is sorted in an appropriate manner. The sorting in fact depends on g; shift-isometry is not needed here. The subsequent statement will later be used for the shift-isometric case, then even implying $L_\infty(\mathcal{P})$-optimality.

Lemma 3.4 ($L_2(\mathcal{P})$-optimality)**.** *Let $g \in L_2(\Omega_{\mathcal{P}})$ and $\{W_k^g := \operatorname{span}\{v_k\}\}_{k\in\mathbb{N}}$ be an ON decomposition of $L_2(\Omega)$, which is* sorted[4]*, i.e.,*

$$C_{k+1}^g \le C_k^g, k \in \mathbb{N}, \quad \textit{for } C_k^g := \int_{\mathcal{P}} \|Q_k^g g(\cdot - \mu)\|_{L_2(\Omega)}^2 d\mu, \tag{3.4}$$

where $Q_k^g : L_2(\Omega) \to W_k^g$ is the orthogonal projector. Then, $V_N^g := \operatorname{span}\{v_1, ..., v_N\}$, $N \in \mathbb{N}$, is optimal, i.e. $\operatorname{dist}(V_N^g, \mathcal{U}_g(\mathcal{P}))_{L_2(\mathcal{P};L_2(\Omega))} = \delta_N(\mathcal{P})$.

Proof. Let $\{\tilde{v}_k\}_{k\in\mathbb{N}}$ be another ON basis of $L_2(\Omega)$; set $\tilde{W}_k := \operatorname{span}\{\tilde{v}_k\}$, let $\tilde{Q}_k : L_2(\Omega) \to \tilde{W}_k$ be the orthogonal projector and define $\tilde{V}_N := \operatorname{span}\{\tilde{v}_1, ..., \tilde{v}_N\}$. Then, each $\tilde{v}_i$, can be expanded in terms of $\{v_k\}_{k\in\mathbb{N}}$, i.e., $\tilde{v}_i = \sum_{k=1}^{\infty} \langle \tilde{v}_i, v_k\rangle_{L_2(\Omega)} v_k$, where

[4]Note, that the spaces W_k^g are independent of g, their sorting, however, depends on it.

$\sum_{k=1}^{\infty}\langle\tilde{v}_i, v_k\rangle^2_{L_2(\Omega)} = \|\tilde{v}_i\|^2_{L_2(\Omega)} = 1$ by orthonormality. Then, since the orthogonal projection is the best approximation and $u_\mu = g(\cdot - \mu)$,

$$\begin{aligned}
\operatorname{dist}(\tilde{V}_N, \mathcal{U}_g(\mathcal{P}))^2_{L_2(\mathcal{P};L_2(\Omega))} &= \sum_{k=N+1}^{\infty}\int_{\mathcal{P}}\|\tilde{Q}_k u_\mu\|^2 d\mu = \sum_{k=N+1}^{\infty}\int_{\mathcal{P}}\langle\tilde{v}_k, u_\mu\rangle^2_{L_2(\Omega)} d\mu \\
&= \sum_{k=N+1}^{\infty}\sum_{i=1}^{\infty}\langle\tilde{v}_k, v_i\rangle^2_{L_2(\Omega)}\int_{\mathcal{P}}\|Q_i^g u_\mu\|^2 d\mu \\
&= \sum_{k=N+1}^{\infty}\sum_{i=1}^{\infty} C_i^g\,\langle\tilde{v}_k, v_i\rangle^2_{L_2(\Omega)} = \sum_{i=1}^{\infty} C_i^g \sum_{k=N+1}^{\infty}\langle\tilde{v}_k, v_i\rangle^2_{L_2(\Omega)}.
\end{aligned}$$

Now, set $b_k^N := \sum_{j=1}^{N}\langle\tilde{v}_j, v_k\rangle^2_{L_2(\Omega)}$; then $0 \le b_k^N \le 1$ and by orthogonality $\sum_{k=1}^{\infty} b_k^N = N$ as well as $b_k^\infty = 1$. The assumption (3.4) yields

$$\begin{aligned}
0 &\le \sum_{k=1}^{N}(1-b_k^N)(C_k^g - C_{N+1}^g) = \sum_{k=1}^{N}(1-b_k^N)C_k^g - C_{N+1}^g\sum_{k=N+1}^{\infty} b_k^N \\
&\le \sum_{k=1}^{N}(1-b_k^N)C_k^g - \sum_{k=N+1}^{\infty} b_k^N C_k^g = \sum_{k=1}^{N} C_k^g - \sum_{k=1}^{\infty} b_k^N C_k^g.
\end{aligned}$$

Again, since the orthogonal projection is the best approximation, we thus get

$$\begin{aligned}
\operatorname{dist}(V_N^g, \mathcal{U}_g(\mathcal{P}))^2_{L_2(\mathcal{P};L_2(\Omega))} &= \sum_{k=N+1}^{\infty} C_k^g \\
\le \sum_{k=N+1}^{\infty} C_k^g + \sum_{k=1}^{N} C_k^g - \sum_{k=1}^{\infty} b_k^N C_k^g &= \sum_{k=1}^{\infty} C_k^g - \sum_{k=1}^{\infty} b_k^N C_k^g = \sum_{k=1}^{\infty}(1-b_k^N)C_k^g \\
= \sum_{k=1}^{\infty}(b_k^\infty - b_k^N)C_k^g &= \sum_{k=1}^{\infty} C_k^g \sum_{\ell=N+1}^{\infty}\langle\tilde{v}_\ell, v_k\rangle^2_{L_2(\Omega)} = \operatorname{dist}(\tilde{V}_N, \mathcal{U}_g(\mathcal{P}))^2_{L_2(\mathcal{P};L_2(\Omega))},
\end{aligned}$$

i.e., V_N^g is optimal in the sense of Kolmogorov. □

Remark 3.5. *As the proof shows, the optimality statement also holds if we replace the full sorting in (3.4) by $C_j^g \le C_k^g$ for all $k \in \{1, ..., N\}$ and $j \ge N+1$.*

Lemma 3.4 yields $L_2(\mathcal{P})$-optimal spaces V_N^g for one-dimensional spaces $W_k^g := \operatorname{span}\{v_k\}$ as long as those spaces are appropriately sorted w.r.t. g. However, we cannot guarantee that subspaces W_k^g are one-dimensional. However, as we shall see now, by adding an additional condition, the subspaces can be split into one-dimensional subspaces so that we still get the desired $L_2(\mathcal{P})$-optimality. The $L_\infty(\mathcal{P})$-optimality follows along with shift-isometry.

Corollary 3.6 (Optimality for shift-isometry)**.** *Let $g \in L_2(\Omega_{\mathcal{P}})$ and $\mathcal{W}^g$ be a corresponding shift-isometric orthogonal decomposition of $L_2(\Omega)$ with $W_k^g =$*

$\mathrm{span}\{w_{k,1},...,w_{k,m_k}\}$, $\dim(W_k^g) = m_k \in \mathbb{N}$, $k \in \mathbb{N}$, *spanned by orthonormal functions, such that the orthogonal projectors* $Q_k^g : L_2(\Omega) \to W_k^g$, $Q_{k,i}^g : L_2(\Omega) \to \mathrm{span}\{w_{k,i}\}$, $i = 1,...,m_k$, *satisfy a sorting condition (w.r.t. g) of the form*

$$\tfrac{1}{m_{k+1}}\|Q_{k+1}^g g\|_{L_2(\Omega)}^2 \le \tfrac{1}{m_k}\|Q_k^g g\|_{L_2(\Omega)}^2 \quad \text{for all } k \in \mathbb{N}, \tag{3.5}$$

$$\int_{\mathcal{P}} \|Q_{k,i}^g g(\cdot - \mu)\|_{L_2(\Omega)}^2 d\mu = \tfrac{1}{m_k}\|Q_k^g g\|_{L_2(\Omega)}^2, \quad \text{for all } i = 1,...,m_k \text{ and } k \in \mathbb{N}. \tag{3.6}$$

Then, V_N^g *defined by* (3.2) *satisfies* $\delta_N(\mathcal{P})^2 = \sum_{k=M(N)+1}^{\infty} \|Q_k^g g\|_{L_2(\Omega)}^2 = d_N(\mathcal{P})^2$.

Proof. We start by recalling (3.3) in Proposition 3.3. Next, we verify the conditions of Lemma 3.4. To this end, we need to associate a monotonous sequence $\{C_\ell^g\}_{\ell\in\mathbb{N}} \subset \mathbb{R}^+$ to the one-dimensional spaces spanned by the functions $w_{k,i}$. This requires to assign an index $\ell \in \mathbb{N}$ to any index pair (k,i), $k \in \mathbb{N}$, $i = 1,...,m_k$. We do so by setting $\ell = \ell(k,i) := i + \sum_{j=1}^{k-1} m_j$. Then, we define $C_\ell^g = C_{\ell(k,i)}^g := \int_{\mathcal{P}} \|Q_{k,i}^g g(\cdot - \mu)\|_{L_2(\Omega)}^2 d\mu$, $\ell = 1,2,3,...$ and observe by (3.5) and (3.6) that $C_1^g = \cdots = C_{m_1}^g \ge C_{m_1+1}^g = \cdots = C_{m_1+m_2}^g \ge C_{m_1+m_2+1}^g = \cdots$. Hence, the assumptions of Lemma 3.4 are satisfied, so that we can deduce $\delta_N(\mathcal{P}) = \mathrm{dist}(V_N^g, \mathcal{U}_g(\mathcal{P}))_{L_2(\mathcal{P};L_2(\Omega))}$. Finally, by $d\mu(\mathcal{P}) = 1$,

$$d_N(\mathcal{P})^2 \le \mathrm{dist}(V_N^g, \mathcal{U}_g(\mathcal{P}))_{L_\infty(\mathcal{P};L_2(\Omega))}^2 = \sum_{k=M(N)+1}^{\infty} \|Q_k^g g\|_{L_2(\Omega)}^2 = \delta_N(\mathcal{P})^2 \le d_N(\mathcal{P})^2,$$

which proves the assertion. □

Remark 3.7. *Constructing* V_N^g *as in* (3.2) *implies that the dimension* N *of* V_N^g *is always given by* $N = \sum_{k=1}^{M(N)} m_k$. *Hence, probably, we do not get spaces* V_N^g *for* all $N \in \mathbb{N}$ *in that way. However, we can extend* V_N^g *by* $\tilde{N} - N$ *basis functions from* $W_{M(N)+1}^g$ *and get* $\delta_{\tilde{N}}(\mathcal{P})$*-optimal spaces* $V_{\tilde{N}}^g$ *for any* $\mathbb{N} \ni \tilde{N} = \dim(V_{\tilde{N}}^g)$*, loosing shift-isometry, however. We do not go into details here as a precise description requires significant technicalities.*

4 N-widths for half-wave symmetric functions

We are now going to construct shift-isometric orthogonal decompositions for half-wave symmetric functions $g \in L_2(\Omega_{\mathcal{P}})$ in terms of trigonometric functions. This will allow us to define spaces $W_k^{\mathrm{hws}} \subset L_2(\Omega)$ which have an additional property, namely they are *shift-invariant*, i.e., if $w \in W_k^{\mathrm{hws}}$, then $w(\cdot - \mu) \in W_k^{\mathrm{hws}}$ for all $\mu \in \mathcal{P}$. We will need to consider even and odd HWS functions separately.

Even half-wave symmetric functions

Lemma 4.1. *Let* $k \in \mathbb{N}$, *set* $\varphi_{k,1}^{\mathrm{evn}}(x) := \sqrt{2}\sin\big(2k\pi x\big)$, $\varphi_{k,2}^{\mathrm{evn}}(x) := \sqrt{2}\cos\big(2k\pi x\big)$, $x \in \Omega$ *and* $W_k^{\mathrm{evn}} := \mathrm{span}\{\varphi_{k,1}^{\mathrm{evn}}, \varphi_{k,2}^{\mathrm{evn}}\}$. *Then,* $\{\varphi_{k,1}^{\mathrm{evn}}, \varphi_{k,2}^{\mathrm{evn}}\}$ *is an orthonormal basis for* W_k^{evn}, $\dim(W_k^{\mathrm{evn}}) = 2$ *and the spaces are shift-invariant.*

Proof. The statements concerning orthonormality and dimension are straightforward. Let $\mu \in \mathcal{P}$ and $x \in \Omega$, then $\varphi_{k,1}^{\text{evn}}(x-\mu) = \sqrt{2}\cos\big(2k\pi\mu\big)\sin\big(2k\pi x\big) - \sqrt{2}\sin\big(2k\pi\mu\big)\cos\big(2k\pi x\big) =: a_{k,1}(\mu)\,\varphi_{k,1}^{\text{evn}}(x) + a_{k,2}(\mu)\;\varphi_{k,2}^{\text{evn}}(x)$, therefore $\varphi_{k,1}^{\text{evn}}(\cdot-\mu) \in W_k^{\text{evn}}$. The same applies for $\varphi_{k,2}^{\text{evn}}$, which concludes the proof. $\square$

The simple proof shows that $W_k^{\text{evn}} = \{\alpha\,\sin(2k\pi\cdot+\beta) : \alpha,\beta \in \mathbb{R}\}$. For $k=0$, we set $W_0^{\text{even}} := \text{span}\{\chi_{[0,1)}\}, \dim(W_0^{\text{even}}) = 1$, i.e., the constant functions.

Remark 4.2. *The above definition is similar to Kolmogorov's paper in 1936 [1], where the best basis functions for all periodic $H^r(0,1)$-functions with $\|f^{(r)}\|_{L_2(0,1)} \le 1$ is shown to be $\{1,\ \sqrt{2}\sin(2\pi kx),\ \sqrt{2}\cos(2\pi kx),\ k=1,2,...,\frac{N-1}{2}\}$. For such classes of functions, Kolmogorov quantified a constant, which was later called* Kolmogorov N-width *in honor of his contributions and proved $d_N = (\pi N)^{-r}$.*

Odd half-wave symmetric functions

In a quite analogous manner, we get a similar result for the odd half-wave symmetric case. We skip the proof.

Lemma 4.3. *Let $k \ge 1$, define $\varphi_{k,1}^{\text{odd}}(x) := \sqrt{2}\sin\big((2k-1)\pi x\big)$, $\varphi_{k,2}^{\text{odd}}(x) := \sqrt{2}\cos\big((2k-1)\pi x\big)$, $x \in \Omega$ and $W_k^{\text{odd}} := \text{span}\{\varphi_{k,1}^{\text{odd}}, \varphi_{k,2}^{\text{odd}}\}$. Then, $\{\varphi_{k,1}^{\text{odd}}, \varphi_{k,2}^{\text{odd}}\}$ is an orthonormal basis for W_k^{odd}, $\dim(W_k^{\text{odd}}) = 2$ and the spaces are shift-invariant.* $\square$

Shift-isometric orthogonal decompositions

We shall now prove that the above construction yields shift-isometric orthogonal decompositions. It turns out that shift-isometry (Def. 3.1 (ii)) and half-wave symmetry allow us to use the same orthogonal decomposition of $L_2(\Omega)$ for all L_2^{evn}- and L_2^{odd}-functions, respectively. We do not need $\mathcal{W}^g$ for each individual function $g \in L_2^{\text{hws}}$.

Lemma 4.4. *(a) The family $\mathcal{W}^{\text{evn}} := \{W_k^{\text{evn}}\}_{k\in\mathbb{N}_0}$ is a shift-isometric orthogonal decomposition of $L_2(\Omega)$ w.r.t. all $g \in L_2^{\text{evn}}$. (b) The family $\mathcal{W}^{\text{odd}} := \{W_k^{\text{odd}}\}_{k\in\mathbb{N}}$ is a shift-isometric orthogonal decomposition of $L_2(\Omega)$ w.r.t. all $g \in L_2^{\text{odd}}$.*

Proof. We restrict ourselves to the odd case since the even case is analogous. We start by considering some $g^{\text{odd}} \in L_2^{\text{odd}}$, which, by standard Fourier-type arguments, can be expressed also on $\Omega_{\mathcal{P}}$ in terms of $\varphi_k^{\text{odd}} := \{2^{-1/2}\varphi_{k,1}^{\text{odd}}, 2^{-1/2}\varphi_{k,2}^{\text{odd}} : k \in \mathbb{N}\}$,[5] i.e., $g^{\text{odd}} = \sum_{\ell=1}^{\infty}\big(A_\ell^{\text{odd}}\varphi_{\ell,1}^{\text{odd}} + B_\ell^{\text{odd}}\varphi_{\ell,2}^{\text{odd}}\big)$,[6] where the expansion coefficients can be expressed as (recall the scaling)

$$\begin{aligned}
A_k^{\text{odd}} &= 2^{-1/2}\langle g^{\text{odd}}, 2^{-1/2}\varphi_{k,1}^{\text{odd}}\rangle_{L_2(\Omega_{\mathcal{P}})} = \tfrac{1}{2}\int_{-1}^{1} g^{\text{odd}}(x)\,\sqrt{2}\sin((2k-1)\pi x)\,dx \\
&= \tfrac{1}{2}\left\{\int_{-1}^{0} + \int_{0}^{1}\right\} g^{\text{odd}}(x)\,\sqrt{2}\sin((2k-1)\pi x)\,dx = \langle g^{\text{odd}}, \varphi_{k,1}^{\text{odd}}\rangle_{L_2(\Omega)}.
\end{aligned}$$

Here, we used a change of variables in the first integral and the fact that g is an odd HWS function. In a completely analogous fashion, we get $B_k^{\text{odd}} = \langle g^{\text{odd}}, \varphi_{k,2}^{\text{odd}}\rangle_{L_2(\Omega)}$.

---

[5] The scaling by $2^{-1/2}$ makes the functions orthonormal on the larger interval $\Omega_{\mathcal{P}}$.

[6] The coefficients have a different scaling than in (2.6): $A_k^{\text{odd}} = 2^{-1/2}\hat{A}_k$, $B_k^{\text{odd}} = 2^{-1/2}\hat{B}_k$.

We use this expansion for the shifted function $g^{\rm odd}(\cdot-\mu)$ on Ω, namely

$$g^{\rm odd}(\cdot-\mu)=\sum_{\ell=1}^{\infty}\big(A_\ell^{\rm odd}\varphi_{\ell,1}^{\rm odd}(\cdot-\mu)+B_\ell^{\rm odd}\varphi_{\ell,2}^{\rm odd}(\cdot-\mu)\big). \tag{4.1}$$

From this we can compute the orthogonal projection, namely

$$Q_k^{\rm odd}g^{\rm odd}(\cdot-\mu)=\langle g^{\rm odd}(\cdot-\mu),\varphi_{k,1}^{\rm odd}\rangle_{L_2(\Omega)}\varphi_{k,1}^{\rm odd}+\langle g^{\rm odd}(\cdot-\mu),\varphi_{k,2}^{\rm odd}\rangle_{L_2(\Omega)}\varphi_{k,2}^{\rm odd}. \tag{4.2}$$

Next, we insert (4.1) into (4.2) and use the following identities, which are easily verified

$$\begin{aligned}\langle\varphi_{\ell,1}^{\rm odd}(\cdot-\mu),\varphi_{k,1}^{\rm odd}\rangle_{L_2(\Omega)}&=\langle\varphi_{\ell,2}^{\rm odd}(\cdot-\mu),\varphi_{k,2}^{\rm odd}\rangle_{L_2(\Omega)}=\cos((2k-1)\pi\mu)\delta_{\ell,k},\\ -\langle\varphi_{\ell,1}^{\rm odd}(\cdot-\mu),\varphi_{k,2}^{\rm odd}\rangle_{L_2(\Omega)}&=\langle\varphi_{\ell,2}^{\rm odd}(\cdot-\mu),\varphi_{k,1}^{\rm odd}\rangle_{L_2(\Omega)}=\sin((2k-1)\pi\mu)\delta_{\ell,k},\end{aligned} \tag{4.3}$$

so that

$$\begin{aligned}Q_k^{\rm odd}g^{\rm odd}(\cdot-\mu)&=\langle A_k^{\rm odd}\varphi_{k,1}^{\rm odd}(\cdot-\mu)+B_k^{\rm odd}\varphi_{k,2}^{\rm odd}(\cdot-\mu),\varphi_{k,1}^{\rm odd}\rangle_{L_2(\Omega)}\varphi_{k,1}^{\rm odd}\\ &\quad+\langle A_k^{\rm odd}\varphi_{k,1}^{\rm odd}(\cdot-\mu)+B_k^{\rm odd}\varphi_{k,2}^{\rm odd}(\cdot-\mu),\varphi_{k,2}^{\rm odd}\rangle_{L_2(\Omega)}\varphi_{k,2}^{\rm odd}\\ &=[A_k^{\rm odd}\cos((2k-1)\pi\mu)+B_k^{\rm odd}\sin((2k-1)\pi\mu)]\varphi_{k,1}^{\rm odd}\\ &\quad+[-A_k^{\rm odd}\sin((2k-1)\pi\mu)+B_k^{\rm odd}\cos((2k-1)\pi\mu)]\varphi_{k,2}^{\rm odd}.\end{aligned}$$

Since $\{\varphi_{k,1}^{\rm odd},\varphi_{k,2}^{\rm odd}\}$ is an ON-basis for $W_k^{\rm odd}$, we get that $\|Q_k^{\rm odd}g^{\rm odd}(\cdot-\mu)\|^2_{L_2(\Omega)}=(A_k^{\rm odd})^2+(B_k^{\rm odd})^2=\|Q_k^{\rm odd}g^{\rm odd}\|^2_{L_2(\Omega)}$, i.e., the projectors are shift-isometric. Considering (4.3) for $\mu=0$ shows that the spaces $W_k^{\rm odd}$ are mutually orthogonal. Finally, $\mathrm{span}(\{\varphi_{k,1}^{\rm odd},\varphi_{k,2}^{\rm odd}\}_{k\in\mathbb{N}})$ is dense in $L_2(\Omega)$, so that $\mathcal{W}^{\rm odd}$ is an orthogonal decomposition of $L_2(\Omega)$, which completes the proof. □

The following result will be needed later in order to balance the dimensions for the optimal spaces in the sense of Kolmogorov in the even and odd HWS case.

Lemma 4.5. *For the orthogonal projectors* $Q_{k,i}^{\rm evn}:L_2(\Omega)\to\mathrm{span}\{\varphi_{k,i}^{\rm evn}\}$ *and* $Q_{k,i}^{\rm odd}:L_2(\Omega)\to\mathrm{span}\{\varphi_{k,i}^{\rm odd}\}$, $i=1,2$, $k\in\mathbb{N}$, *we have for all* $g\in L_2(\Omega_{\mathcal{P}})$

$$\begin{aligned}\int_{\mathcal{P}}\|Q_{k,i}^{\rm evn}g(\cdot-\mu)\|^2_{L_2(\Omega)}d\mu&=\tfrac12\|Q_k^{\rm evn}g\|^2_{L_2(\Omega)},\\ \int_{\mathcal{P}}\|Q_{k,i}^{\rm odd}g(\cdot-\mu)\|^2_{L_2(\Omega)}d\mu&=\tfrac12\|Q_k^{\rm odd}g\|^2_{L_2(\Omega)}.\end{aligned}$$

Proof. As above, we only give the proof for the second identity, i.e., the odd HWS case. Recalling (4.3), we have for $i=1$ the identity $Q_{k,1}^{\rm odd}g(\cdot-\mu)=[A_k^{\rm odd}\cos((2k-1)\pi\mu)+B_k^{\rm odd}\sin((2k-1)\pi\mu)]\,\varphi_{k,1}^{\rm odd}$, so that

$$\int_{\mathcal{P}}\|Q_{k,1}^{\rm odd}g(\cdot-\mu)\|^2_{L_2(\Omega)}d\mu=\int_0^1[A_k^{\rm odd}\cos((2k-1)\pi\mu)+B_k^{\rm odd}\sin((2k-1)\pi\mu)]^2d\mu$$

$$= \int_0^1 [A_k^{\mathrm{odd}} \tfrac{1}{\sqrt{2}} \varphi_{k,2}^{\mathrm{odd}}(\mu) + B_k^{\mathrm{odd}} \tfrac{1}{\sqrt{2}} \varphi_{k,1}^{\mathrm{odd}}(\mu)]^2 d\mu = \tfrac{1}{2}((A_k^{\mathrm{odd}})^2 + (B_k^{\mathrm{odd}})^2)$$
$$= \tfrac{1}{2} \|Q_k^{\mathrm{odd}} g\|^2_{L_2(\Omega)},$$

since $\sqrt{2}\sin((2k-1)\pi\mu)$, $\sqrt{2}\cos((2k-1)\pi\mu)$ are orthonormal in $L_2(\mathcal{P})$. The other case is similar. □

4.1 Optimal N-term approximation for half-wave symmetric functions

We are now in position to identify the spaces which are optimal in the sense of Kolmogorov. It turns out that we need to distinguish between even and odd HWS functions. Unfortunately, this requires some technicalities since the even case involves the one-dimensional space W_0^{evn} and all other W_k^{evn}, $k \geq 1$, are two-dimensional, whereas all odd HWS spaces W_k^{odd}, $k \geq 1$, are of dimension two. As before, the optimal space requires a sorting, which depends on the given function g[7] and will be done by permutations σ^g.

Corollary 4.6 (Optimal basis by sorting)**.** *Let $\mathcal{W}^{\mathrm{hws}}$ be a shift-isometric orthogonal decomposition according to Lemma 4.4 and $g \in L_2^{\mathrm{hws}}$. We assume that the spaces are* ordered *w.r.t. g, i.e.,*

$$\begin{aligned} \frac{\|Q^{\mathrm{evn}}_{\sigma^g(k+1)-1} g\|^2_{L_2(\Omega)}}{\min(2, \sigma^g(k+1))} &\leq \frac{\|Q^{\mathrm{evn}}_{\sigma^g(k)-1} g\|^2_{L_2(\Omega)}}{\min(2, \sigma^g(k))}, k \in \mathbb{N}, && \text{if } g \in L_2^{\mathrm{evn}}, \\ \tfrac{1}{2}\|Q^{\mathrm{odd}}_{\sigma^g(k+1)} g\|^2_{L_2(\Omega)} &\leq \tfrac{1}{2}\|Q^{\mathrm{odd}}_{\sigma^g(k)} g\|^2_{L_2(\Omega)}, k \in \mathbb{N}, && \text{if } g \in L_2^{\mathrm{odd}}, \end{aligned}$$

where $\sigma^g : \mathbb{N} \to \mathbb{N}$ is a permutation. As in (3.2), define V_N^g as the direct sum of the first $M(N)$ spaces $W^{\mathrm{evn}}_{\sigma^g(k)-1}$, $W^{\mathrm{odd}}_{\sigma^g(k)}$ such that $\dim(V_N^g) = N$. Then,

$$\delta_N(\mathcal{P})^2 = d_N(\mathcal{P})^2 = \sum_{k=M(N)+1}^{\infty} \begin{cases} \|Q^{\mathrm{evn}}_{\sigma^g(k)-1} g\|^2_{L_2(\Omega)}, & \text{if } g \in L_2^{\mathrm{evn}}, \\ \|Q^{\mathrm{odd}}_{\sigma^g(k)} g\|^2_{L_2(\Omega)}, & \text{if } g \in L_2^{\mathrm{odd}}. \end{cases}$$

Proof. Use Lemma 4.4 and Lemma 4.5 applied to Corollary 3.6. □

As mentioned in Remark 3.7, we can extend spaces V_N^g as constructed here by optimal sorting by additional basis functions from $W^g_{M(N)+1}$ to construct spaces $V^g_{\tilde{N}}$ for any $\tilde{N}$. We do not go into (the quite technical) details here.

Example 4.7 (Discontinuous jump)**.** *We detail the example already investigated in [20], where $d_N(\mathcal{P}) \geq \frac{1}{2} N^{-1/2}$ was shown for discontinuous initial and boundary conditions, i.e., $g = -\chi_{[-1,0)} + \chi_{[0,1]}$, which is easily seen to be odd HWS. Since $\langle g, \varphi^{\mathrm{odd}}_{k,1}\rangle_{L_2(\Omega)} = \frac{2\sqrt{2}}{(2k-1)\pi}$ and $\langle g, \varphi^{\mathrm{odd}}_{k,2}\rangle_{L_2(\Omega)} = 0$, we have $\|Q_k^{\mathrm{odd}} g\|^2_{L_2(\Omega)} = \frac{8}{(2k-1)^2\pi^2} \geq \frac{8}{(2k+1)^2\pi^2} = \|Q^{\mathrm{odd}}_{k+1} g\|^2_{L_2(\Omega)}$, so that the assumptions of Corollary 4.6 hold without*

[7] The spaces depend only on the HWS type, the sorting on g.

additional sorting. Hence, we get an exact representation of the N-width by

$$\begin{aligned}\delta_N(\mathcal{P})^2 &= \sum_{k=N+1}^{\infty} \tfrac{1}{2}\,\|Q^{\text{odd}}_{M(k)}g\|^2_{L_2(\Omega)} = \frac{1}{2}\sum_{k=N+1}^{\infty} \|Q^{\text{odd}}_{\lfloor\frac{k+1}{2}\rfloor}g\|^2_{L_2(\Omega)} \\ &= \frac{4}{\pi^2}\sum_{k=N+1}^{\infty}\Big(2\Big\lfloor\tfrac{k+1}{2}\Big\rfloor-1\Big)^{-2} = \frac{1}{\pi^2}\Psi^{(1)}\Big(\big\lfloor\tfrac{N}{2}\big\rfloor+\tfrac{1}{2}\Big)+\frac{1}{\pi^2}\Psi^{(1)}\Big(\big\lfloor\tfrac{N+1}{2}\big\rfloor+\tfrac{1}{2}\Big),\end{aligned} \tag{4.4}$$

where $\Psi^{(1)}(\cdot)$ is the first derivative of the Digamma function $\Psi(\cdot)$. Moreover, for even N, we have $\delta_N(\mathcal{P}) = d_N(\mathcal{P})$. △

4.2 Quasi-optimality

Hence, as a preparation for our subsequent analysis (see Theorem 5.3 below), we are now going to bound these norms.

Lemma 4.8. *With the definitions of $\mathcal{W}^{\text{evn}}$ in Lemma 4.1 and $\mathcal{W}^{\text{odd}}$ in Lemma 4.3 (without sorting) we have for all $g \in L_2(\Omega_{\mathcal{P}})$*[8] *the estimates for all $k \in \mathbb{N}$*

$$\|Q^{\text{evn}}_k g\|_{L_2(\Omega)} \le \frac{2}{k\pi}\|g\|_{L_2(\Omega)}, \qquad \|Q^{\text{odd}}_k g\|_{L_2(\Omega)} \le \frac{4}{(2k-1)\pi}\|g\|_{L_2(\Omega)}.$$

Proof. Again, we shall only give the proof for the decomposition $\mathcal{W}^{\text{odd}}$, but it works the same way for $\mathcal{W}^{\text{evn}}$. We express $g_{|\Omega} \in L_2(\Omega)$ as limit of step functions, namely $g = \lim_{k\to\infty} g_n$,[9] where g_n is a step function piecewise defined as $g_n(x) \equiv g_{n,i} := \int_{I_{n,i}} g(y)dy$ for $x \in I_{n,i}$ and $I_{n,i} = [\frac{i-1}{n}, \frac{i}{n})$, so that $\cup_{i=1}^n \bar{I}_{n,i} = [0,1]$. The Fourier coefficients of g_n converge towards those of $g \in L_2(\Omega)$, i.e., adopting the above notion of A_k and B_k, the Dominated Convergence Theorem yields

$$\begin{aligned}A_k &:= \langle g, \varphi^{\text{odd}}_{k,1}\rangle_{L_2(\Omega)} = \lim_{n\to\infty}\langle g_n, \varphi^{\text{odd}}_{k,1}\rangle_{L_2(\Omega)} = \sqrt{2}\lim_{n\to\infty}\sum_{i=1}^{n} g_{n,i}\int_{I_{n,i}}\sin((2k-1)\pi x)dx \\ &= \sqrt{2}\lim_{n\to\infty}\sum_{i=1}^{n} g_{n,i}\frac{-\cos((2k-1)\pi\frac{i}{n})+\cos((2k-1)\pi\frac{i-1}{n})}{(2k-1)\pi} \\ &\le \frac{2\sqrt{2}}{(2k-1)\pi}\lim_{n\to\infty}\sum_{i=1}^{n}|g_{n,i}| = \frac{\sqrt{8}}{(2k-1)\pi}\|g\|_{L_1(\Omega)} \le \frac{\sqrt{8}}{(2k-1)\pi}\|g\|_{L_2(\Omega)}.\end{aligned}$$

We can estimate $B_k := \langle g, \varphi^{\text{odd}}_{k,2}\rangle_{L_2(\Omega)}$ in a similar manner. Finally, we recall that $\|Q^{\text{odd}}_k g\|^2_{L_2(\Omega)} = A_k^2 + B_k^2$, which yields the desired estimate. □

Remark 4.9. *Even though we do not have a proof yet, we conjecture that $\|Q^{\text{hws}}_k u_\mu\|_{L_2(\Omega)} = \Theta(k^{-(r+1)})$ for $g \in H^r(\Omega_{\mathcal{P}}) \setminus H^{r+1}(\Omega_{\mathcal{P}})$, which would mean that the above estimates are in fact sharp. This is also backed by our numerical experiments*

[8] For this lemma, we do not need half-wave symmetry of g.
[9] The limit is taken in $L_2(\Omega)$.

in Section 7. If this conjecture holds true, one can show that quasi-optimal spaces give rise to the same rate as the optimal ones involving sorting. Hence, one can avoid sorting which we will do from now one, i.e., we choose σ^g as the identity.

Recalling the above settings, we need to distinguish not only between even and odd HWS functions, but also between even and odd dimensions. For convenience, we summarize the construction in a definition.

Definition 4.10. *Let $g \in L_2(\Omega_\mathcal{P})$ and $N \in \mathbb{N}$.*

1. *If $g \in L_2^{\mathrm{evn}}$, $M = \lfloor \frac{N}{2} \rfloor$, then set*
   (a) $V_N := \mathrm{span}\{\chi_{[0,1)}, \varphi_{1,1}^{\mathrm{odd}}, \varphi_{1,2}^{\mathrm{odd}}, ..., \varphi_{M,1}^{\mathrm{odd}}, \varphi_{M,2}^{\mathrm{odd}}\}$, *if $N = 2M+1$ is odd;*
   (b) $V_N := \mathrm{span}\{\chi_{[0,1)}, \varphi_{1,1}^{\mathrm{odd}}, \varphi_{1,2}^{\mathrm{odd}}, ..., \varphi_{M-1,1}^{\mathrm{odd}}, \varphi_{M,1}^{\mathrm{odd}}\}$, *if $N = 2M$ is even;*
2. *if $g \in L_2^{\mathrm{odd}}$, $M = \lfloor \frac{N+1}{2} \rfloor$, then set*
   (a) $V_N := \mathrm{span}\{\varphi_{1,1}^{\mathrm{evn}}, \varphi_{1,2}^{\mathrm{evn}}, ..., \varphi_{M-1,2}^{\mathrm{evn}}, \varphi_{M,1}^{\mathrm{evn}}, \varphi_{M,2}^{\mathrm{evn}}\}$, *if $N = 2M$ is even;*
   (b) $V_N := \mathrm{span}\{\varphi_{1,1}^{\mathrm{evn}}, \varphi_{1,2}^{\mathrm{evn}}, ..., \varphi_{M-1,1}^{\mathrm{evn}}, \varphi_{M-1,2}^{\mathrm{evn}}, \varphi_{M,1}^{\mathrm{evn}}\}$, *if $N = 2M-1$ is odd.*

Note that V_N is independent of the specific form of g, there is no sorting.

In all cases, we have $\dim(V_N) = N$. We use these spaces as candidates for optimal spaces in the sense of Kolmogorov and now relate the distance to the N-width.

Theorem 4.11. *Let $g \in L_2(\Omega_\mathcal{P})$, $\mathcal{U}_g(\mathcal{P}) = \{g(\cdot - \mu) : \mu \in \mathcal{P}\}$ and construct V_N according to Definition 3.4 for $M = M(N)$. Then,*

1. $\mathrm{dist}(V_N, \mathcal{U}_g(\mathcal{P}))^2_{L_2(\mathcal{P}; L_2(\Omega))} = \frac{1}{2} \sum_{k=N+1}^{\infty} \|Q^{\mathrm{evn}}_{\lfloor \frac{k}{2} \rfloor} g\|^2_{L_2(\Omega)}$, *if $g \in L_2^{\mathrm{evn}}$;*
2. $\mathrm{dist}(V_N, \mathcal{U}_g(\mathcal{P}))^2_{L_2(\mathcal{P}; L_2(\Omega))} = \frac{1}{2} \sum_{k=N+1}^{\infty} \|Q^{\mathrm{odd}}_{\lfloor \frac{k+1}{2} \rfloor} g\|^2_{L_2(\Omega)}$, *if $g \in L_2^{\mathrm{odd}}$.*

Proof. The assertion follows from the fact that $\mathcal{W}$ is a shift-isometric orthogonal decomposition in the respective space L_2^{evn} and L_2^{odd} (Lemma 4.4), the formula for the distances (Proposition 3.3) and the formula for the norms of the projections in Lemma 4.5. □

4.3 Non-half-wave symmetric functions

Our results can also be extended to non-half-wave symmetric functions. However, we were only able to derive an estimate for the N-width and not a representation as before. Recall from §2.3 that any $g \in L_2(\Omega_\mathcal{P})$ has a unique decomposition $g = g^{\mathrm{evn}} + g^{\mathrm{odd}}$ into its even and odd HWS part.

Corollary 4.12. *Let $N \in \mathbb{N}$ and $g \in L_2(\Omega_\mathcal{P})$. Then,*

$$\delta_N(\mathcal{P})^2 \le d_N(\mathcal{P})^2 \le \frac{1}{2} \sum_{k=N+1}^{\infty} (\|Q^{\mathrm{evn}}_{\lfloor \frac{k}{2} \rfloor} g^{\mathrm{evn}}\|^2_{L_2(\Omega)} + \|Q^{\mathrm{odd}}_{\lfloor \frac{k+1}{2} \rfloor} g^{\mathrm{odd}}\|^2_{L_2(\Omega)}).$$

Proof. Let $g = g^{\mathrm{evn}} + g^{\mathrm{odd}}$, with $g^{\mathrm{evn}} \in L_2^{\mathrm{evn}}, g^{\mathrm{odd}} \in L_2^{\mathrm{odd}}$, be uniquely decomposed. We set $\mathcal{U}_g^{\mathrm{evn}}(\mathcal{P}) := \{g^{\mathrm{evn}}(\cdot - \mu) : \mu \in \mathcal{P}\}$ and $\mathcal{U}_g^{\mathrm{odd}}(\mathcal{P}) := \{g^{\mathrm{odd}}(\cdot - \mu) : \mu \in \mathcal{P}\}$. Then, $\mathcal{U}_g(\mathcal{P}) = \{[g^{\mathrm{odd}} + g^{\mathrm{evn}}](\cdot - \mu) : \mu \in \mathcal{P}\}$, so that $\mathrm{dist}(V_N, \mathcal{U}_g(\mathcal{P}))_{L_\infty(\mathcal{P}; L_2(\Omega))} \le \mathrm{dist}(V_N, \mathcal{U}_g^{\mathrm{evn}}(\mathcal{P})_{L_\infty(\mathcal{P}; L_2(\Omega))} + \mathrm{dist}(V_N, \mathcal{U}_g^{\mathrm{odd}}(\mathcal{P})_{L_\infty(\mathcal{P}; L_2(\Omega))}$. For a fixed $N \in \mathbb{N}$, denote by V_N^{evn} the space defined in part (1) and by V_N^{odd} the space in part (2) of Definition

4.10. Then,

$$
\begin{aligned}
d_N(\mathcal{P}) &= \inf_{\substack{V_N \subset L_2(\Omega)\\ \dim(V_N)=N}} \operatorname{dist}(V_N, \mathcal{U}_g(\mathcal{P}))_{L_\infty(\mathcal{P};L_2(\Omega))}\\
&\le \inf_{\substack{V_N \subset L_2(\Omega)\\ \dim(V_N)=N}} \operatorname{dist}(V_N, \mathcal{U}_g^{\mathrm{odd}}(\mathcal{P}))_{L_\infty(\mathcal{P};L_2(\Omega))} + \inf_{\substack{V_N \subset L_2(\Omega)\\ \dim(V_N)=N}} \operatorname{dist}(V_N, \mathcal{U}_g^{\mathrm{evn}}(\mathcal{P}))_{L_\infty(\mathcal{P};L_2(\Omega))}\\
&\le \operatorname{dist}(V_N^{\mathrm{evn}}, \mathcal{U}_g^{\mathrm{evn}}(\mathcal{P}))_{L_\infty(\mathcal{P};L_2(\Omega))} + \operatorname{dist}(V_N^{\mathrm{odd}}, \mathcal{U}_g^{\mathrm{odd}}(\mathcal{P}))_{L_\infty(\mathcal{P};L_2(\Omega))},
\end{aligned}
$$

so that Theorem 4.11 proves the claim. □

5 The effect of smoothness on the N-width

So far, we did not use any specific properties of the function g modeling initial and boundary values of the linear transport equation. In this section, we shall investigate the influence of the regularity to the decay of the N-width. In particular, we use the above exact representation to derive formulae for the decay of the N-width.

5.1 Finite regularity

Let $g \in H^r(\Omega_{\mathcal{P}})$[10], $r \in \mathbb{N}_0$, be either even or odd HWS. Then, all weak derivatives $g^{(j)} \in H^{r-j}(\Omega_{\mathcal{P}})$, $j = 1, \dots, r$, are HWS of the same type. We express $g^{(r)} = \sum_{k=1}^{\infty} Q_k\, g^{(r)}$, where Q_k is either Q_k^{evn} or Q_k^{odd} depending on the half-wave symmetry of $g^{(r)}$. Hence, in order to determine the N-width, we estimate the norms of the projectors. We consider the shift-isometric orthogonal decomposition without additional sorting.

Theorem 5.1. *Let $g \in H^r(\Omega_{\mathcal{P}})$, $k \in \mathbb{N}$, then*

$$\|Q_k^{\mathrm{evn}} g\|_{L_2(\Omega)} = \left[\tfrac{1}{2k\pi}\right]^r \|Q_k^{\mathrm{evn}} g^{(r)}\|_{L_2(\Omega)} \le 4\left[\tfrac{1}{2k\pi}\right]^{r+1} \|g^{(r)}\|_{L_2(\Omega)}, \ \ \textit{if } g \in L_2^{\mathrm{evn}}, \tag{5.1}$$

$$\|Q_k^{\mathrm{odd}} g\|_{L_2(\Omega)} = \left[\tfrac{1}{(2k-1)\pi}\right]^r \|Q_k^{\mathrm{odd}} g^{(r)}\|_{L_2(\Omega)} \le 4\left[\tfrac{1}{(2k-1)\pi}\right]^{r+1} \|g^{(r)}\|_{L_2(\Omega)}, \ \ \textit{if } g \in L_2^{\mathrm{odd}}. \tag{5.2}$$

Proof. Again, we only prove the odd HWS case, i.e., we show (5.2), since (5.1) can be proven completely analogously. We recall the definition of the Fourier coefficients in Lemma 4.8, i.e., $A_k = \langle g, \varphi_{k,1}^{\mathrm{odd}} \rangle_{L_2(\Omega)}$ and $B_k = \langle g, \varphi_{k,2}^{\mathrm{odd}} \rangle_{L_2(\Omega)}$, so that $\|Q_k^{\mathrm{odd}} g\|_{L_2(\Omega)}^2 = A_k^2 + B_k^2$. Then, using integration by parts, we get

$$A_k = (-1)^{\lfloor \frac{r+1}{2} \rfloor} \left(\frac{(-1)}{(2k-1)\pi} \right)^r \begin{cases} \langle g^{(r)}, \varphi_{k,1}^{\mathrm{odd}} \rangle_{L_2(\Omega)}, & \text{if } r \text{ odd},\\ \langle g^{(r)}, \varphi_{k,2}^{\mathrm{odd}} \rangle_{L_2(\Omega)}, & \text{if } r \text{ even}, \end{cases}$$

$$B_k = (-1)^{\lfloor \frac{r}{2} \rfloor} \left(\frac{(-1)}{(2k-1)\pi} \right)^r \begin{cases} \langle g^{(r)}, \varphi_{k,2}^{\mathrm{odd}} \rangle_{L_2(\Omega)}, & \text{if } r \text{ odd},\\ \langle g^{(r)}, \varphi_{k,1}^{\mathrm{odd}} \rangle_{L_2(\Omega)}, & \text{if } r \text{ even}. \end{cases}$$

[10] The Sobolev space of r-times weakly differentiable functions.

Thus, $A_k^2 + B_k^2 = ((2k-1)\pi)^{-2r}\|Q_k^{\text{odd}} g^{(r)}\|^2_{L_2(\Omega)}$, which proves the first equality in (5.2) and the inequality follows from Lemma 4.8. $\square$

Remark 5.2 (No sorting required)**.** *We recall Remark 3.5, where the monotonicity requirement for $C_k^g := \int_{\mathcal{P}} \|Q_{v_k} g(\cdot - \mu)\|^2_{L_2(\Omega)} d\mu$ in Lemma 3.4 was weakened, i.e., we only need $C_k^g \geq C_\ell^g$ for $k \in \{1, ..., N\}, \ell \geq N+1$. Using Remark 4.9, (5.1) and (5.2), one can show that $C_k^g > C_\ell^g$ for $\ell > (2\pi)^{-1}\big(4\|g^{(r)}\|_{L_2(\Omega)}/C_k^g\big)^{1/(r+1)}$. Therefore, the spaces V_N are quasi-optimal* without sorting.

Theorem 5.3. *Let $g \in H^r(\Omega_{\mathcal{P}}) \cap L_2^{\text{hws}}$, $r \in \mathbb{N}_0$, such that $g^{(r)}$ is either even or odd HWS: For the spaces V_N constructed in Definition 4.10 we then have*

$$\delta_N(\mathcal{P}) \leq d_N(\mathcal{P}) \leq \operatorname{dist}(V_N, \mathcal{U}_g(\mathcal{P}))_{L_\infty(\mathcal{P}; L_2(\Omega))} \leq \tfrac{4}{\pi^{r+1}} \|g^{(r)}\|_{L_2(\Omega)}\, N^{-(r+1/2)}.$$

Proof. We just detail the odd HWS case. From Theorems 4.11 and 5.1 we get

$$\begin{aligned}
\operatorname{dist}(V_N, \mathcal{U}_g(\mathcal{P}))^2_{L_\infty(\mathcal{P}; L_2(\Omega))} &= \frac{1}{2} \sum_{k=N+1}^{\infty} \|Q^{\text{odd}}_{\lfloor \frac{k+1}{2} \rfloor} g\|^2_{L_2(\Omega)} \\
&\leq 8\|g^{(r)}\|^2_{L_2(\Omega)} \sum_{k=N+1}^{\infty} \big(2\left\lfloor \tfrac{k+1}{2} \right\rfloor - 1)\pi\big)^{-2r-2} \leq 8\|g^{(r)}\|^2_{L_2(\Omega)} \sum_{k=N+1}^{\infty} ((k-1)\pi)^{-2r-2} \\
&= \frac{16\|g^{(r)}\|^2_{L_2(\Omega)}}{\pi^{2r+2}} \sum_{k=N}^{\infty} \frac{k^{-2r-2}}{2} \leq \frac{16\|g^{(r)}\|^2_{L_2(\Omega)}}{\pi^{2r+2}} \sum_{k=N}^{\infty} \big(k^{-2r-1} - (k+1)^{-2r-1}\big) \\
&= \frac{16\|g^{(r)}\|^2_{L_2(\Omega)}}{\pi^{2r+2}} N^{-(2r+1)},
\end{aligned}$$

which proves the claim. $\square$

Remark 5.4. *The N-width for H^r-functions was already investigated by Kolmogoroff [1] and Pinkus [2] and is known to be of order $\mathcal{O}(N^{-r})$. Since $\mathcal{U}_g(\mathcal{P}) \subset H^r$, the result in [2] can be used as an upper bound. However, Theorem 5.3 results in a slightly improved rate $\mathcal{O}(N^{-r-1/2})$. We conjecture that this is even sharp, but we did not prove it here. It is remarkable that the error for approximating $\mathcal{U}_g(\mathcal{P})$ is only by a factor $N^{-1/2}$ smaller than the error for the full space H^r.*

Remark 5.5. *Theorem 5.3 can also be extended to functions that are not half-wave symmetric by using Corollary 4.12. Since this is a straightforward extension of Theorem 5.3, we do not detail this.*

5.2 Infinite regularity

It is known that smooth dependency yields exponential decay of the N-width, see e.g. [18, 20]. We can now detail the dependency of the rate w.r.t. the involved parameters by applying Theorem 5.1 in the limit case $r \to \infty$ for $g \in H^\infty$. It turns out that we need some complex analysis in this case.

Proposition 5.6. *Let $g \in H^\infty(\Omega_\mathcal{P}) \cap L_2^{\mathrm{hws}}$ with a complex analytic extension $\bar{g} : \mathcal{B}_1(0) \to \mathbb{C}$[11], $\bar{g}_{|\Omega_\mathcal{P}} = g$ and V_N defined by Definition 4.10. Then, there exists a constants $C = C(g) > 0$ such that $\delta_N(\mathcal{P}) \le d_N(\mathcal{P}) \le \mathrm{dist}(V_N, \mathcal{U}_g(\mathcal{P}))_{L_\infty(\mathcal{P})} \le C K d^{-N}$, where $d = \frac{\pi e}{2} \approx 4.27$ and $K = \sqrt{\frac{32e}{\pi}} \approx 5.26$.*

Proof. We consider the holomorphic extension $\bar{g}$ in $\mathcal{B}_1(0) \supset (-1,1)$. Let $x \in (-0.5, 0.5)$, i.e., $\mathcal{B}_{0.5}(x) \subset \mathcal{B}_1(0)$. Cauchy's integral formula gives $|g^{(r)}(x)| \le r!\, 2^r \max_{z \in \partial\mathcal{B}_{0.5}(x)} |g(z)| \le C\, r!\, 2^r$, where $C := \max_{z\in\mathcal{B}_1(0)} |g(z)|$, [26]. By half-wave symmetry, we get $\|g^{(r)}\|_{L_2(\Omega)} = \|g^{(r)}\|_{L_2(-0.5,0.5)} \le C\, r!\, 2^r$. Then, $d_N(\mathcal{P}) \le \mathrm{dist}(V_N, \mathcal{U}_g(\mathcal{P}))_{L_\infty(\mathcal{P})} \le \frac{4C}{\pi} (\frac{2}{\pi})^N N!\, N^{-N-1/2}$ follows by Theorem 5.3. Using Stirling's approximation $N! < \sqrt{2\pi e N}(N/e)^N$ for the factorial gives the desired result. □

6 Exact decay for piecewise functions

So far, we derived *upper bounds* for the decay of the N-widths for certain classes of functions. In this section, we investigate exact formula for the decay as well as *lower bounds*. As we shall see, the above estimates are in fact sharp in the sense that we are going to construct specific functions $g \in H^r(\Omega_\mathcal{P}) \setminus H^{r+1}(\Omega_\mathcal{P})$, for which lower bounds can be proven. These functions are such that $g^{(r)} = -\chi_{[-1,0)} + \chi_{[0,1]}$, i.e., the jump function considered in Example 4.7. This is done in a recursive manner by setting $g_{0,\text{left}} := -1$, $g_{0,\text{right}} := 1$ and $L_2(\Omega_\mathcal{P}) \ni g_0 := g_{0,\text{left}}\, \chi_{[-1,0)} + g_{0,\text{right}}\, \chi_{[0,1]}$. Then, for $r = 1, 2, 3, \ldots$, we set

$$g_r := g_{r,\text{left}}\, \chi_{[-1,0)} + g_{r,\text{right}}\, \chi_{[0,1]}, \quad \text{where} \tag{6.1a}$$

$$g_{r,\text{left}}(x) := -\int_x^0 g_{r-1,\text{left}}(y)\, dy - \tfrac{1}{2} \int_0^1 g_{r-1,\text{right}}(y)\, dy, \quad x \in [-1,0), \tag{6.1b}$$

$$g_{r,\text{right}}(x) := \int_0^x g_{r-1,\text{right}}(y)\, dy - \tfrac{1}{2} \int_0^1 g_{r-1,\text{right}}(y) dy, \quad x \in [0,1]. \tag{6.1c}$$

It is easily seen that

$$g_1(x) = \begin{cases} -x - \frac{1}{2}, & x \in [-1,0), \\ x - \frac{1}{2}, & x \in [0,1], \end{cases} \qquad g_2(x) = \begin{cases} -\frac{1}{2}x^2 - \frac{1}{2}x, & x \in [-1,0), \\ \frac{1}{2}x^2 - \frac{1}{2}x, & x \in [0,1]. \end{cases}$$

Lemma 6.1. *Let $r \in \mathbb{N}_0$ and g_r as defined in (6.1). Then, $g_r \in H^r(\Omega_\mathcal{P})$ is odd HWS, $g_r^{(r)} = g_0 = -\chi_{[-1,0)} + \chi_{[0,1]}$ and*

$$\|Q_k^{\text{odd}} g_r\|_{L_2(\Omega)} = \sqrt{8}((2k-1)\pi)^{-(r+1)}. \tag{6.2}$$

[11] For $z \in \mathbb{C}$ and $r > 0$, we denote the ball of radius r around z by $\mathcal{B}_r(z) := \{y \in \mathbb{C} : |y - z| \le r\}$.

Proof. The recursive definition immediately implies that $g_r \in H^r(\Omega_{\mathcal{P}})$. Moreover, $g_0 \in L_2^{\text{odd}}$. Assuming that $g_{r-1} \in L_2^{\text{odd}}$ for $r \in \mathbb{N}$, we get for any $0 < \varepsilon < 1$

$$\begin{aligned} -g_r(-\varepsilon) &= \int_{-\varepsilon}^{0} g_{r-1,\text{left}}(y)\,dy + \tfrac{1}{2}\int_0^1 g_{r-1,\text{right}}(y)\,dy \\ &= -\int_{-\varepsilon}^{0} g_{r-1,\text{right}}(1+y)\,dy + \int_0^1 g_{r-1,\text{right}}(y)dy - \tfrac{1}{2}\int_0^1 g_{r-1,\text{right}}(y)\,dy \\ &= -\int_{1-\varepsilon}^{1} g_{r-1,\text{right}}(y)\,dy + \int_0^1 g_{r-1,\text{right}}(y)dy - \tfrac{1}{2}\int_0^1 g_{r-1,\text{right}}(y)\,dy \\ &= \int_0^{1-\epsilon} g_{r-1,\text{right}}(y)\,dy - \tfrac{1}{2}\int_0^1 g_{r-1,\text{right}}(y)dy = g_r(1-\varepsilon), \end{aligned}$$

so that $g_r \in L_2^{\text{odd}}$ by induction. Finally, Theorem 5.1 and Example 4.7 yields

$$\|Q_k^{\text{odd}} g_r\|_{L_2(\Omega)} = ((2k-1)\pi)^{-r}\|Q_k^{\text{odd}} g_r^{(r)}\|_{L_2(\Omega)} = \sqrt{8}((2k-1)\pi)^{-(r+1)},$$

which proves (6.2) and finishes the proof. $\square$

Theorem 6.2. *Let $r \in \mathbb{N}_0$ and V_N by Definition 4.10 for $g = g_r$ odd HWS. Then,*

$$\delta_N(\mathcal{P})^2 = \sum_{k=N+1}^{\infty} \frac{4}{\left((2\lfloor\frac{k+1}{2}\rfloor - 1)\pi\right)^{2r+2}},$$

and, for even N, we have $\delta_N(\mathcal{P}) = d_N(\mathcal{P}) = \operatorname{dist}(V_N, \mathcal{U}_g(\mathcal{P}))_{L_\infty(\mathcal{P};L_2(\Omega))}$. Moreover $\frac{2}{(2r+1)^{1/2}\pi^{r+1}}(N+1)^{-(2r+1)/2} \le \delta_N(\mathcal{P}) \le \frac{\sqrt{8}}{\pi^{r+1}} N^{-(2r+1)/2}$ for all $N \in \mathbb{N}$.

Proof. The representation follows with Lemma 6.1 in combination with Theorem 4.11. It remains to prove the bounds. To prove the bounds, we deduce that

$$\operatorname{dist}(V_N, \mathcal{U}_g(\mathcal{P}))^2_{L_\infty(\mathcal{P};L_2(\Omega))} = \sum_{k=N+1}^{\infty} 4\left(\tfrac{1}{(2\lfloor\frac{k+1}{2}\rfloor-1)\pi}\right)^{2r+2}$$

$$\begin{cases} \le \frac{4}{\pi^{2r+2}} \sum_{k=N+1}^{\infty} \left(\frac{1}{k-1}\right)^{2r+2} & \le \frac{8}{\pi^{2r+2}} N^{-(2r+1)}, \\ \ge \frac{4}{\pi^{2r+2}(2r+1)} \sum_{k=N+1}^{\infty} \left(\frac{2r+1}{k^{2r+2}}\right) & \ge \frac{4}{\pi^{2r+2}(2r+1)} \sum_{k=N+1}^{\infty} \frac{1}{k^{2r+2}} \frac{\sum_{i=1}^{2r+1}\binom{2r+1}{i-1}k^i}{\sum_{i=0}^{2r+1}\binom{2r+1}{i}k^i} \\ & = \frac{4}{\pi^{2r+2}(2r+1)} \sum_{k=N+1}^{\infty} \frac{1}{k^{2r+2}} \frac{k(k+1)^{2r+1} - k^{2r+2}}{(k+1)^{2r+1}} \\ & = \frac{4}{\pi^{2r+2}(2r+1)} \sum_{k=N+1}^{\infty} \frac{1}{k^{2r+1}} - \frac{1}{(k+1)^{2r+1}} \\ & = \frac{4}{(2r+1)\,\pi^{2r+2}} (N+1)^{-(2r+1)}, \end{cases}$$

which are the desired bounds. Finally, (6.2) ensures the required sorting in Corollary 4.6, so that $\delta_N(\mathcal{P}) = \mathrm{dist}(V_N, \mathcal{U}_g(\mathcal{P}))_{L_\infty(\mathcal{P};L_2(\Omega))}$ (see Remark 3.7), which is equal to $d_N(\mathcal{P})$ for even N and completes the proof. □

Remark 6.3. *We have numerically validated the representation in Theorem 6.2 along the lines of Section 7 below. Moreover, we have seen in our experiments that the lower bound given in Theorem 6.2 seems to be the exact rate.*

7 Numerical experiments

We are now going to report results of some of our numerical experiments highlighting different quantitative aspects of our previous theoretical investigations. More details and the code can be found in https://github.com/flabowski/n_widths_for_transport.

7.1 Numerical approximation of the N-width

It is clear that we cannot compute $\delta_N(\mathcal{P})$ or $d_N(\mathcal{P})$ exactly, at least in general. Even for a given linear approximation space V_N, the distance of $\mathcal{U}_g(\mathcal{P})$ to V_N amounts computing an integral over $\mathcal{P}$ in the L_2-case or the determination of a supremum in the L_∞-framework. Both would only be possible exactly, if we would have a formula for the error $u_\mu - P_N u_\mu$ at hand.

Otherwise, we need a discretization in space Ω and for the parameter set $\mathcal{P}$ in such a way that the resulting numerical approximation is sufficiently accurate. In space, we fix a number $n_x \in \mathbb{N}$ of uniformly spaced quadrature or sampling points x_i, $i = 1, ..., n_x$ (wee choose $n_x = 2500$), by setting $\Delta x := 1/n_x$ and $x_i := (2i-1)/2\,\Delta x$. We collect these points in a vector $\boldsymbol{x} := (x_1, ..., x_{n_x})^\top \in \mathbb{R}^{n_x}$. We proceed in a similar manner for $\mathcal{P}$ by choosing $n_\mu \in \mathbb{N}$ of uniformly spaced points μ_j, $j = 1, ..., n_\mu$ (we choose $n_\mu = 2500$) by $\Delta\mu := 1/n_\mu$ and $\mu_j := (2j-1)/2\,\Delta\mu$, $\boldsymbol{\mu} := (\mu_1, ..., \mu_{n_\mu})^\top \in \mathbb{R}^{n_\mu}$. This corresponds to the midpoint rule for numerical integration.

POD-ROM

For some given parameter value μ_j and a given function g, we determine $X_{i,j} := g(x_i - \mu_j) = u_{\mu_j}(x_i)$ as a "snapshot" of (2.2). These values are collected in the *snapshot matrix* $\boldsymbol{X} := (X_{i,j})_{i=1,\dots,n_x; j=1,\dots,n_\mu} \in \mathbb{R}^{n_x \times n_\mu}$.

For the L_2-width, we perform a singular value decomposition (SVD) $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$, which is then truncated to dimension $N \in \mathbb{N}$ in order to obtain a reduced basis, which corresponds to the *Proper Orthogonal Decomposition (POD)*. Since it is known that the POD is the best approximation w.r.t. $L_2(\mathcal{P})$, we get the optimal spaces V_N in the $L_2(\mathcal{P})$-sense. The error and thus the δ_N-width can be computed from the singular values: $\delta_N(\mathcal{P})^2 \approx \sum_{k=N+1}^{\min(n_\mu, n_x)} \sigma_k^2$.

Optimal spaces

In some cases, we have constructed (in some cases optimal) spaces V_N, i.e., we know an ON-basis for V_N. In such a case, the SVD is given through the projection terms, i.e. no eigenvalue decomposition needs to be performed. In case we know that $d_N(\mathcal{P}) =$

$\delta_N(\mathcal{P})$, no further computations are needed. In case they are not equal, we can proceed with the basis V_N in order to compute the approximation error as described now in detail.

Computation of the distance/error

In order to determine the distance of V_N to the set $\mathcal{U}_g(\mathcal{P})$ of solutions, let $V_N = \text{span}\{\psi_1, ..., \psi_N\}$ for some ON-basis functions ψ_ℓ, $\ell = 1, ..., N$. The best approximation of some function u_μ onto V_N is the orthogonal projection, i.e., $P_N(u_\mu) = \sum_{\ell=1}^N \langle u_\mu, \psi_\ell\rangle_{L_2(\Omega)} \psi_\ell$. The inner products are approximated by the midpoint rule, i.e., $\langle u_\mu, \psi_\ell\rangle_{L_2(\Omega)} \approx \frac{1}{n_x} \sum_{i'=1}^{n_x} u_\mu(x_{i'})\, \psi_\ell(x_{i'})$, so that by orthormality

$$
\begin{aligned}
&\inf_{\tilde{v}_N \in V_N} \|u_{\mu_j} - \tilde{v}_N\|^2_{L_2(\Omega)} = \|u_{\mu_j} - P_N(u_{\mu_j})\|^2_{L_2(\Omega)} \\
&\approx \tfrac{1}{n_x} \sum_{i=1}^{n_x} \big(u_{\mu_j}(x_i) - (P_N(u_{\mu_j}))(x_i)\big)^2 = \tfrac{1}{n_x} \sum_{i=1}^{n_x} \left(u_{\mu_j}(x_i) - \sum_{\ell=1}^N \langle u_{\mu_j}, \psi_\ell\rangle_{L_2(\Omega)} \psi_\ell(x_i)\right)^2 \\
&\approx \tfrac{1}{n_x} \sum_{i=1}^{n_x} \left(u_{\mu_j}(x_i) - \sum_{\ell=1}^N \tfrac{1}{n_x} \sum_{i'=1}^{n_x} u_{\mu_j}(x_{i'})\, \psi_\ell(x_{i'}) \psi_\ell(x_i)\right)^2 \\
&= \tfrac{1}{n_x} \sum_{i=1}^{n_x} \left(X_{i,j} - \tfrac{1}{n_x} \sum_{\ell=1}^N \sum_{i'=1}^{n_x} X_{i',j}\, \Psi_{i',\ell} \Psi_{i,\ell}\right)^2 = \tfrac{1}{n_x} \sum_{i=1}^{n_x} \left(X_{i,j} - \tfrac{1}{n_x} (\boldsymbol{\Psi}_N \boldsymbol{\Psi}_N^\top \boldsymbol{X})_{i,j}\right)^2,
\end{aligned}
$$

where $\boldsymbol{\Psi}_N = (\psi_\ell(x_i))_{i,\ell} \in \mathbb{R}^{n_x \times N}$. Then, $\text{dist}(V_N, \mathcal{U}_g(\mathcal{P}))_{L_\infty(\mathcal{P}; L_2(\Omega))}$ is approximated by taking the maximum over $j = 1, ..., n_\mu$ of the latter quantity and then the square root. As for the L_2-distance $\text{dist}(V_N, \mathcal{U}_g(\mathcal{P}))_{L_2(\mathcal{P}; L_2(\Omega))}$, by (2.5b)

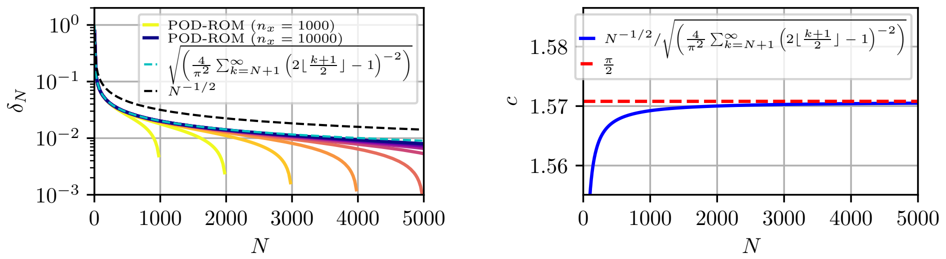
$$
\begin{aligned}
\text{dist}(V_N, \mathcal{U}_g(\mathcal{P}))^2_{L_2(\mathcal{P}; L_2(\Omega))} &= \|u_\mu - P_N u_\mu\|^2_{L_2(\mathcal{P}; L_2(\Omega))} = \int_{\mathcal{P}} \|u_\mu - P_N u_\mu\|^2_{L_2(\Omega)}\, d\mu \\
&\approx \tfrac{1}{n_\mu} \sum_{j=1}^{n_\mu} \|u_{\mu_j} - P_N u_{\mu_j}\|^2_{L_2(\Omega)} \approx \tfrac{1}{n_\mu} \tfrac{1}{n_x} \sum_{j=1}^{n_\mu} \sum_{i=1}^{n_x} \left(X_{i,j} - \tfrac{1}{n_x} (\boldsymbol{\Psi}_N \boldsymbol{\Psi}_N^\top \boldsymbol{X})_{i,j}\right)^2 \\
&= \tfrac{1}{n_\mu}\ \tfrac{1}{n_x} \|\boldsymbol{X} - \tfrac{1}{n_x} \boldsymbol{\Psi}_N \boldsymbol{\Psi}_N^\top \boldsymbol{X}\|^2_F =: \tfrac{1}{n_\mu}\ \tfrac{1}{n_x} \|\boldsymbol{G}_N\|^2_F,
\end{aligned}
$$

where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix.

7.2 Error bound for a jump discontinuity

For a discontinuous function, it is known that $d_N(\mathcal{P}) \leq cN^{-1/2}$, [20], the novel exact representation is given in (4.4). We compare a reduced order model determined by POD (termed POD-ROM here) with the exact rate, which allows us to numerically investigate the difference between the asymptotic order $N^{-1/2}$ and the exact rate. The results are shown in Figure 1. In the graph on the left, we show the decay for different sizes of n_x, i.e., various numbers of the original snapshots to build the POD (shown in different colors). As we see, they asymptotically reach the exact representation shown in cyan. This also confirms the known fact that POD is optimal w.r.t. the L_2-width. We also show the asymptotic order $N^{-1/2}$ in black.

The formula for the exact rate cannot immediately be re-interpreted as a simple asymptotic w.r.t. N. To this end, on the right-hand side of Figure 1b we plot the ratio of $N^{-1/2}$ and the exact form and see that it reaches $\frac{\pi}{2}$, which is interesting at least for two reasons: (i) the asymptotic rate $N^{-1/2}$ is sharp with a multiple factor of $\frac{\pi}{2}$; (ii) the exact formula has an asymptotic behavior as $N^{-1/2}$.



(a) POD-ROM vs. analytic error decay.

(b) Ratio exact form and asymptotic rate.

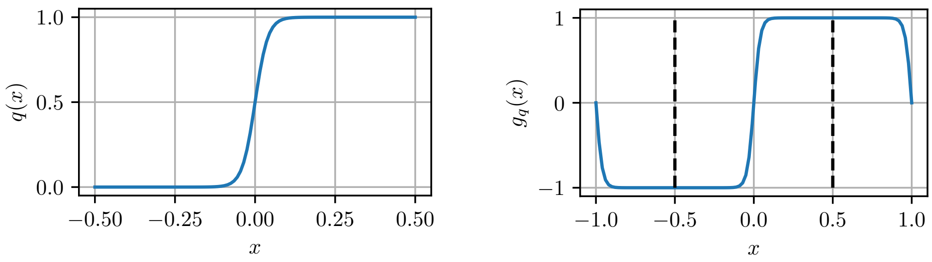
Fig. 1: Kolmorogov N-width $d_N(\mathcal{P})$ width for a discontinuous function – comparison of POD, exact form of $d_N(\mathcal{P})$ and known asymptotic rate.

7.3 Smooth steep functions

We are now considering smooth functions which are "close" to a jump in the sense that they have one or more steep ramps. To this end, we construct an odd half-wave symmetric function, shown in Figure 2b. The starting point is some smooth odd-symmetric function q on the interval $(-\frac{1}{2}, \frac{1}{2})$ (see Figure 2a). Based upon this, we define the odd HWS function $g = g_q$ by

$$g(x) := \begin{cases} 1 - 2q(x+1), & -1 < x \leq -\frac{1}{2}, \\ 2q(x) - 1, & -\frac{1}{2} < x \leq \frac{1}{2}, \\ 1 - 2q(x-1), & \frac{1}{2} < x < 1. \end{cases} \tag{7.1}$$

Following this idea, we can derive functions with arbitrary smoothness and arbitrarily steep ramps in order to be able to numerically investigate the dependence of the decay rate of $d_N(\mathcal{P})$ on the regularity and the shape of the function. To this end, we construct a whole family $\{q_k\}_{k\in\mathbb{N}_0}$ such that $q_k \in C^k$ but $q_k \notin C^{k+1}$ (so that k is the exact degree of regularity of q_k). We show an example of such functions $q_0, \ldots, q_5$ in (7.2). Starting from a linear function q_0, we successively increase the polynomial



(a) Smooth function $q : (-\frac{1}{2}, \frac{1}{2}) \to \mathbb{R}$ with a ramp at $x = 0$.

(b) Resulting od HWS $g = g_q$ with smooth seems (marked by dashed lines).

Fig. 2: Construction of an odd HWS initial condition from a smooth ramp.

degree. A parameter ϵ is used to control the steepness of the ramp.[12] Then, we get

$$q_0(x) := \tfrac{1}{\epsilon}x, \tag{7.2a}$$
$$q_1(x) := -\tfrac{2}{\epsilon^3}x^3 + \tfrac{3}{\epsilon^2}x^2, \tag{7.2b}$$
$$q_2(x) := \tfrac{6}{\epsilon^5}x^5 - \tfrac{15}{\epsilon^4}x^4 + \tfrac{10}{\epsilon^3}x^3, \tag{7.2c}$$
$$q_3(x) := -\tfrac{20}{\epsilon^7}x^7 + \tfrac{70}{\epsilon^6}x^6 - \tfrac{84}{\epsilon^5}x^5 + \tfrac{35}{\epsilon^4}x^4, \tag{7.2d}$$
$$q_4(x) := \tfrac{70}{\epsilon^9}x^9 - \tfrac{315}{\epsilon^8}x^8 + \tfrac{540}{\epsilon^7}x^7 - \tfrac{420}{\epsilon^6}x^6 + \tfrac{126}{\epsilon^5}x^5, \tag{7.2e}$$
$$q_5(x) := -\tfrac{252}{\epsilon^{11}}x^{11} + \tfrac{1386}{\epsilon^{10}}x^{10} - \tfrac{3080}{\epsilon^9}x^9 + \tfrac{3465}{\epsilon^8}x^8 - \tfrac{1980}{\epsilon^7}x^7 + \tfrac{462}{\epsilon^6}x^6. \tag{7.2f}$$

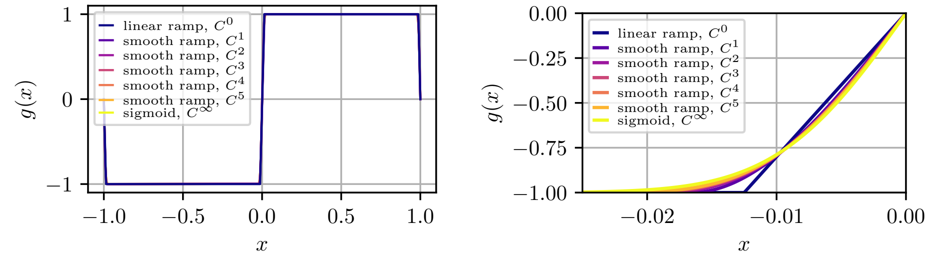
We can continue this process to obtain a C^∞-function, but do not go into details. In order to get a meaningful comparison for the dependency of the N-width in terms of the smoothness, we will use ϵ for such a fine-tuning. The aim is that all functions q_k should feature a similar steep jump from 0 to 1, but differ in their regularity, which of course causes different shapes of the functions, see Figure 3. Hence, we fit each resulting g_{q_k} to g_{q_0} and choose ϵ as the parameter resulting in the best fit. We indicate the resulting values for ϵ in Table 1. The resulting functions of different smoothness

regularity	C^0	C^1	C^2	C^3	C^4	C^5
ϵ	0.025	0.03316	0.04002	0.04592	0.05116	0.05592

Table 1: Values for 2ϵ for each g_k

are plotted in Figure 3. As we can see from the left graph in Figure 3a, the shape of all functions is quite similar. The main difference lies in the regularity as can be seen in the zoom in Figure 3b.

[12]We give all details for the sake of reproducible research.

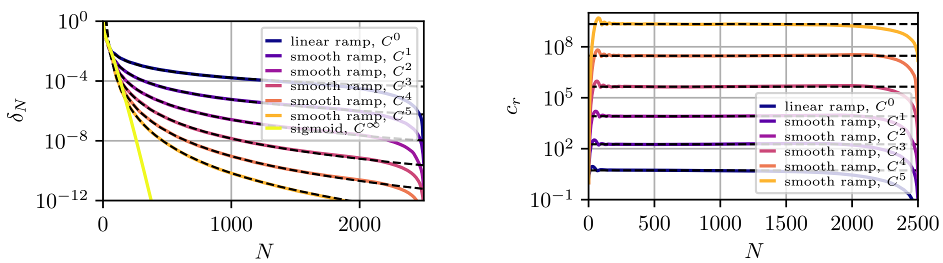


(a) Ramp functions on $\Omega_\mathcal{P}$.

(b) Zoom into subinterval $[-0.025, 0]$.

Fig. 3: Ramp functions with varying smoothness C^k.

The results concerning the N-width are shown in Figure 4. On the left, in Figure 4a, we compare the N-width $\delta_N(\mathcal{P})$ for $g_{q_{r-1}} \in C^{r-1}(\Omega)$, $r = 1, ..., 6$ and also for the C^∞-sigmoid function (yellow) with exponential decay. We also indicate the error bound from Theorem 5.3, i.e., $\tilde{c}_r N^{-(r+1/2)}$. As there is no difference visible, in Figure 4b, we plot the ratio of the numerically computed error and $c_r N^{-(r+1/2)}$ for a fitted c_r for every $r \in \mathbb{N}_0$. We see very good matches indicating that our bounds are sharp regarding N. As with the decay of the jump discontinuity (c.f. Figure 1a), the numerically computed decay for N close to $\min(n_x, n_\mu)$ suffers from inaccuracies that are related to the discretization error.



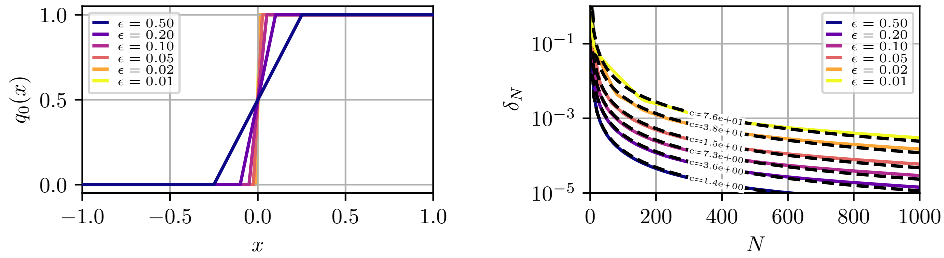
(a) δ_N and error estimation for C^{r-1}-functions.

(b) Estimation of c_r s.t. $\delta_N \approx c_r N^{-(r+1/2)}$.

Fig. 4: N-width for ramps with varying regularity.

7.4 The impact of the slope

In §7.3 we have investigated functions with an almost identical ramp but with different smoothness. Now, we fix the regularity and vary the slope, i.e., the maximal value of the derivative (or norm of the gradient in higher dimensions). From our theoretical findings, we expect that the asymptotic decay rate should not be influenced by the slope. However, all estimates involve a multiplicative factor, which might depend on the slope. In order to clarify this, we consider a continuous, piecewise linear function with varying steepness. We choose the function q_0 in (7.2a) for different values of ϵ, see Figure 5a. The results are displayed in Figure 5. We observe that the asymptotic rate is in fact identical, but the multiplicative factor grows with decreasing ϵ, the steeper the slope, the slower the decay of the N-width.



(a) Function q_0 for different ϵ.

(b) $\delta_N(\mathcal{P})$ for different ϵ.

Fig. 5: N-width depending on the slope of a continuous, piecewise linear function.

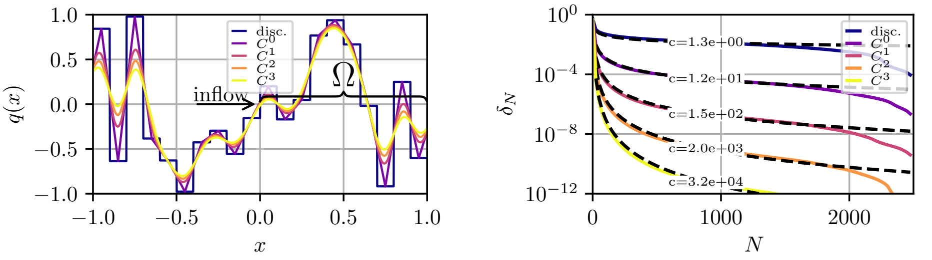
7.5 Beyond symmetry

Finally, we consider almost arbitrary functions g to define initial and boundary conditions for our original linear transport problem (2.1) in the sense that $g_{|[-1,0]}$ defines the inflow (i.e., the boundary condition) and $g_{|[0,1]}$ is the initial condition on Ω. Again, we focus on the influence of the regularity on the decay of the N-width. To this end, we start by a piecewise constant discontinuous function as displayed in Figure 6a (dark blue), where the height of the 20 steps are chosen at random. Smother versions are constructed by applying a convolution with a uniform box kernel, that is as wide as the distance between two discontinuities, see also Figure 6a.[13] The N-width is shown in Figure 6b, where -again- we clearly see the dependence of the decay on the regularity; the smoother, the better the rate. The rates are the same as in the previous case, but the constants (indicated in Figure 6b) differ. This experiment confirms our results also beyond half-wave symmetry, which we had to assume for the given proofs.

A 2D-example

All our analysis above was restricted to the 1D case $\Omega = (0, 1)$. However, from the presentation it should be clear that at least some of what has been presented can be

[13] A closer agreement between the original and the convoluted function as well as a faster error decay could be achieved through a convolution by a narrow Gaussian kernel. However, we aimed at highlighting the effect of regularity.

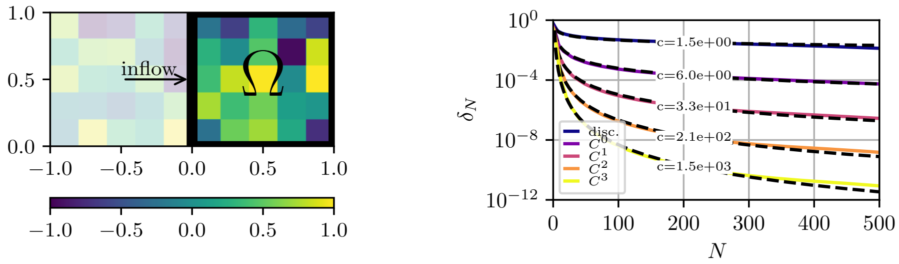


(a) Piecewise constant with 20 random steps (dark blue); increasing smoothness $C^k(\Omega_\mathcal{P})$ by $k+1$-fold convolution, $k = 0, ..., 3$.

(b) N-width decay for discontinuous function and $g \in C^k(\Omega_\mathcal{P})$, $k = 0, ..., 3$.

Fig. 6: $\delta_N(\mathcal{P})$ for random functions of different smoothness.

generalized to the higher-dimensional case by means of tensor products. In order to show this also numerically, we consider a linear transport problem on $\Omega = (0, 1)^2$, see Figure 7a. There, we indicate piecewise constant boundary conditions (on the left square yielding the inflow conditions) and initial conditions on Ω (on the right square). As before, we realize initial and boundary conditions of higher regularity by applying convolutions. The resulting N-widths are displayed in Figure 7b, where we see once more that the rate is correlated to the regularity.



(a) Piecewise constant initial and boundary conditions indicated by color boxes.

(b) Kolmorogov N-width for $C^k(\Omega_\mathcal{P})$-conditions, $k = 0, ..., 4$.

Fig. 7: 2D-transport problem: δ_N-width for initial- and boundary conditions of different regularity $C^k(\Omega_\mathcal{P})$, $k = 0, ..., 4$.

8 Conclusions

We have derived both exact representations as well as sharp bounds for the N-width for significant classes of functions used as initial and boundary values for the linear transport problem. The influence of the regularity on the decay has been rigorously investigated, namely $c_r N^{-(r+1/2)}$ for functions in the Sobolev space H^r. It became clear that a poor decay of the N-width is *only* a question of the smoothness of the solution in terms of the parameter, not of the problem itself. In other words, the decay does not necessarily depend on the PDE, but on the data such as initial and boundary values. We have also seen that the constant in the decay estimate depends on the slope of the function in a severe manner.

Our main tool is Fourier analysis and the notion of half-wave symmetric functions. This notion allowed us to construct linear spaces which we have shown to be optimal in the sense of Kolmogorov. Since any function can be written as a sum of even and odd HWS function, we derived a general *upper* estimate for the N-width. We have investigated both the $L_2(\mathcal{P})$-based N-width $\delta_N(\mathcal{P})$ and the $L_\infty(\mathcal{P})$-based (worst case) N-width $d_N(\mathcal{P})$ and we have proven $\delta_N(\mathcal{P}) = d_N(\mathcal{P})$ for all HWS functions (Cor. 4.6). We conjecture that this upper bound is also asymptotically sharp, but we do not have a proof yet. Finally, the presented approach could also be generalized and adapted for other kind of PPDEs.

References

[1] Kolmogorov, A.: Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse. Annals of Mathematics **37**(1), 107–110 (1936)

[2] Pinkus, A.: N-Widths in Approximation Theory. Springer, Berlin (1985)

[3] DeVore, R.A.: Nonlinear approximation. Acta Numerica **7**, 51–150 (1998)

[4] Floater, M.S., Manni, C., Sande, E., Speleers, H.: Best low-rank approximations and Kolmogorov n-widths. SIAM J. Matrix Anal. Appl. **42**(1), 330–350 (2021)

[5] Bressan, A., Floater, M.S., Sande, E.: On best constants in L2 approximation. IMA J. Numer. Anal. **41**(4), 2830–2840 (2021)

[6] Cohen, A., DeVore, R., Schwab, C.: Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs. Found. Comp. Math. **10**(6), 615–646 (2010)

[7] Cohen, A., DeVore, R.: Kolmogorov widths under holomorphic mappings. IMA J. Numer. Anal. **36**(1), 1–12 (2015)

[8] Melenk, M.: On n-widths for elliptic problems. J. Math. Anal. Appl. **247**(1), 272–289 (2000)

[9] Benner, P., Ohlberger, M., Cohen, A., Willcox, K.: Model Reduction and Approximation. Computational Science & Engineering. SIAM, Philadelphia (2017)

[10] Hesthaven, J.S., Rozza, G., Stamm, B.: Certified Reduced Basis Methods for Parametrized Partial Differential Equations. Springer, Cham (2015)

[11] Quarteroni, A., Manzoni, A., Negri, F.: Reduced Basis Methods for Partial Differential Equations: An Introduction. Springer, Cham (2015)

[12] Binev, P., Cohen, A., Dahmen, W., DeVore, R., Petrova, G., Wojtaszczyk, P.: Convergence Rates for Greedy Algorithms in Reduced Basis Methods. SIAM J. Math. Anal. **43**(3), 1457–1472 (2011)

[13] Bachmayr, M., Cohen, A.: Kolmogorov widths and low-rank approximations of parametric elliptic PDEs. Math. Comp. **86**(304), 701–724 (2017)

[14] DeVore, R.A.: The Theoretical Foundation of Reduced Basis Methods. Computational Science & Engineering, pp. 137–168. SIAM, Philadelphia (2017)

[15] Maday, Y., Patera, A.T., Turinici, G.: A priori convergence theory for reduced-basis approximations of single-parameter elliptic partial differential equations. J. Sci. Comp. **17**, 437–446 (2002)

[16] Maday, Y.: Reduced basis method for the rapid and reliable solution of partial differential equations. In: Proceedings of ICM, Madrid, EMS (2006)

[17] Rozza, G., Huynh, D.B.P., Patera, A.T.: Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics. Arch. Comput. Methods Eng. **15**(3), 229–275 (2008)

[18] Buffa, A., Maday, Y., Patera, A.T., Prud'homme, C., Turinici, G.: *A priori* convergence of the greedy algorithm for the parametrized reduced basis method. ESAIM Math. Model. Numer. Anal. **46**(3), 595–603 (2012)

[19] Lassila, T., Manzoni, A., Quarteroni, A., Rozza, G.: Generalized reduced basis methods and n-width estimates for the approximation of the solution manifold of parametric pdes. Boll. Un. Mat. Italiana **6**(1), 113–135 (2013)

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

[21] Maday, Y., Patera, A.T., Turinici, G.: A priori convergence theory for reduced-basis approximations of single-parameter elliptic partial differential equations. J. Sci. Comp. **17**(1), 437–446 (2002)

[22] Maday, Y., Patera, A.T., Turinici, G.: Global a priori convergence theory for reduced-basis approximations of single-parameter symmetric coercive elliptic partial differential equations. C.R. Math. **335**(3), 289–294 (2002)

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

[24] Arbes, F., Jensen, Ø., Mardal, K.-A., Dokken, J.: Model order reduction of solidification problems. ECCOMAS Congress (2022)

[25] Attenborough, M.P.: Mathematics for Electrical Engineering and Computing. Elsevier, Oxford (2003)

[26] Rudin, W.: Real and Complex Analysis, 3rd edn., p. 416. McGraw-Hill Book Co., New York (1987)