MATHEMATICS OF COMPUTATION
Volume 87, Number 312, July 2018, Pages 1757–1781
http://dx.doi.org/10.1090/mcom/3265
Article electronically published on October 26, 2017

A KERNEL-BASED DISCRETISATION METHOD FOR FIRST ORDER PARTIAL DIFFERENTIAL EQUATIONS

TOBIAS RAMMING AND HOLGER WENDLAND

Abstract. We derive a new discretisation method for first order PDEs of arbitrary spatial dimension, which is based upon a meshfree spatial approximation. This spatial approximation is similar to the SPH (smoothed particle hydrodynamics) technique and is a typical kernel-based method. It differs, however, significantly from the SPH method since it employs an Eulerian and not a Lagrangian approach. We prove stability and convergence for the resulting semi-discrete scheme under certain smoothness assumptions on the defining function of the PDE. The approximation order depends on the underlying kernel and the smoothness of the solution. Hence, we also review an easy way of constructing smooth kernels yielding arbitrary convergence orders. Finally, we give a numerical example by testing our method in the case of a one-dimensional Burgers equation.

1. Introduction

In this paper, we will derive and analyse a new discretisation method for a large class of first order evolution equations, i.e., we are interested in finding approximate solutions to initial value problems of the form

$$\partial_t \rho + f(t,x,\rho,\nabla\rho) = 0 \qquad \text{on } (0,\infty)\times\mathbb{R}^n, \tag{1.1}$$

$$\rho(0,\cdot) = \rho_0 \qquad \text{on } \mathbb{R}^n. \tag{1.2}$$

Here, $f:\mathbb{R}^{2n+2}\to\mathbb{R}$ is a given, twice continuously differentiable mapping, $\rho_0 \in C^r(\mathbb{R}^n)$ is the given initial condition and $\nabla\rho$ denotes as usual the vector of first order spatial derivatives of ρ. The function $\rho:[0,\infty)\times\mathbb{R}^n\to\mathbb{R}$ is the solution we want to compute and we will assume that the above problem has a solution $\rho\in C^{1,r}([0,\infty)\times\mathbb{R}^n)$, i.e., a solution which has at least first order continuous derivatives in time and r-th order continuous derivatives in space with $r\geq 1$.

These somewhat strong conditions are required for our error analysis. The numerical scheme itself can be set up under much milder conditions. Nonetheless, as usual the efficiency of the scheme is based upon the assumption that we have a strong solution to the problem.

First order problems of the above type occur in many different situations, often when modelling physical phenomena. The most simple example is given by the well-known *linear transport problems*

$$\partial_t\rho + u\cdot\nabla\rho = g$$

with a given drift $u:[0,\infty)\times\mathbb{R}^n\to\mathbb{R}^n$ and a source term $g:[0,\infty)\times\mathbb{R}^n\to\mathbb{R}$.

Received by the editor November 28, 2015, and, in revised form, December 22, 2016, and February 8, 2017.

2010 *Mathematics Subject Classification.* Primary 65M12, 65M15, 35F50.

Further examples range from the *nonlinear evolution problems of Hamilton-Jacobi type*, given by equations of the form

$$\partial_t \rho + H(x, \nabla \rho) = 0$$

with given Hamiltonian function $H : \mathbb{R}^{2n} \to \mathbb{R}$ to problems from *optimal control and dynamical programming*, where the control usually also has to satisfy a given (often physically motivated) partial differential equation.

One specific possible application we have in mind is the determination of the basin of attraction of a system of ordinary differential equations $\dot{x}(t) = f(t, x(t))$. Here, a feasible approach is to set up a first order partial differential equation for the Lyapunov function $L(x, t)$, which then has to satisfy

$$\nabla L(t, x) \cdot f(t, x) + \partial_t L(t, x) = -g(t, x)$$

with a given function g. For a recent review see [18]. Since the function g can be chosen by the user, it is possible to have a smooth solution such that this particular type of problem is indeed covered by our convergence analysis below.

Numerically, equations of the form (1.1), (1.2) are typically solved by classical finite differences, semi-Lagrangian approximation schemes, level set methods or by finite elements using an approach based upon viscosity solutions [16, 17].

However, some of these problems do not exhibit classical smooth solutions. This is particularly the case when scalar, non-linear conservation laws of the form

$$\partial_t \rho + \nabla \cdot f(\rho) = 0$$

with a flux function $f : \mathbb{R} \to \mathbb{R}^n$ are considered. Here it is well-known that, even for smooth initial data, the solutions can develop shocks in finite time (cf. [4]) or an even worse non-smooth behaviour.

The numerical simulation of such conservation laws and the handling of these non-smooth solutions has attracted considerable attention within the last few years. In particular, weighted essentially non-oscillatory (WENO) schemes and discontinuous Galerkin schemes have proven to work quite well—at least in one spatial dimension (cf. [20, 30, 31] and the references therein). However, there are still considerable problems when it comes to treating non-smooth data in the multivariate setting, since most of the known schemes are more or less extensions of the one-dimensional schemes and show spurious effects if the shocks and other discontinuities are not aligned with the mesh [15]. As a consequence, there is a need for suitable, genuinely multivariate methods. While the method we want to analyse in this paper is still far from tackling these kinds of problems, the method can hopefully be extended in such a direction in the future.

In this paper, we propose a new spatial discretisation method, which, in a certain way, uses techniques from classical particle methods such as SPH (smoothed particle hydrodynamics); see for example [22–25, 34, 35]. For this reason, our analysis will partially employ ideas of earlier works such as [10, 26]. However, our method differs from these methods significantly since we do not use a Lagrangian approach as it is usually done in this context. Instead, we employ an Eulerian approach. Hence, in a certain way, our method can also be seen as a generalised finite difference method. Moreover, from an approximation theory point of view, our approximation scheme uses and extends ideas from quasi-interpolation, particularly those with radial basis functions. Our scheme to approximate a function ρ which depends on time and space is based on the following ingredients. We fix a function $\zeta : \mathbb{R}^n \to \mathbb{R}$, a sampling size $h > 0$ and a scaling parameter $\varepsilon > 0$. Then, the function ρ would be approximated by

$$[\rho](t,x) := \left(\frac{h}{\varepsilon}\right)^n \sum_{j\in\mathbb{Z}^n} \rho(t,jh)\zeta\left(\frac{x-jh}{\varepsilon}\right) \tag{1.3}$$

if we were able to sample ρ at the points $\rho(t,jh)$, $t \geq 0$, $j \in \mathbb{Z}^n$. In approximation theory, an expression of the form (1.3) is referred to as a quasi-interpolant, though the functions approximated there usually do not depend on time. Moreover, usually some conditions on the function ζ are required, which ensure that certain polynomials are reproduced. Widely used choices for ζ are given by radial basis functions; see for example [2, 5–9, 28]. There are, however, fundamental differences to our approach. As mentioned above, in classical quasi-interpolation, polynomial reproduction is required, which is, due to a result by Strang and Fix [33], equivalent to the Strang-Fix conditions. We will make a similar but still different assumption on our function ζ; see Definition 2.1 below. The main difference, however, is that in quasi-interpolation, in part due to earlier results on principal shift-invariant spaces (see [11, 12, 27]), the underlying basis function is often globally supported. Obviously, a globally supported ζ might cause convergence problems in (1.3). Hence, classical quasi-interpolation actually employs a finite linear combination of a globally supported radial basis function as the function ζ. This linear combination is chosen such that ζ decays exponentially and polynomials are reproduced. In this paper, we will omit this step and work directly with compactly supported functions, which makes the evaluation of (1.3) extremely efficient. Yet another difference is that in quasi-interpolation the scale parameter ε and the sampling parameter are usually connected by $\varepsilon = h$. This is referred to as a *stationary setting* and the Strang-Fix conditions then yield approximation orders depending on the order of the Strang-Fix condition. Our situation is different since we will have to assume that h is significantly less than ε, i.e., that $h/\varepsilon \to 0$ with $h, \varepsilon \to 0$. This is comparable to, but not exactly the same as, the *non-stationary setting* in [27].

The main difference, however, is that we actually do not have the samples $\rho(t,jh)$ but need to approximate them by solving a system of ODEs. It might be possible to employ the machinery of quasi-interpolation in this context, as well, but we are not aware of such an approach.

In deriving the ODE just mentioned we will employ (1.3) to approximate spatial derivatives, in particular, $\nabla\rho$ of the unknown function ρ. This is simply achieved by differentiating $[\rho]$, i.e., by forming $\nabla[\rho]$. This means that we are reconstructing the derivative of a function from its sampled function values. Such a procedure is commonly called a generalised finite difference approximation. The difference to classical finite differences is, of course, that in finite differences a local interpolating polynomial is differentiated, while here we differentiate an expression of the form (1.3). Interestingly, in a series of papers, a connection between radial basis function approximation and polynomial approximation had been established. In [13], it was shown that univariate interpolation with increasingly flat Gaussians yields univariate polynomial interpolants. In [29] flat limits of analytical and radial functions were studied and it was shown that the limiting interpolant is again, under certain assumptions, a polynomial interpolant. Further results in this direction are, for example, in [19, 32]. Though all of these results are mainly for interpolation they could also be used to derive approximations for derivatives, which then would resemble classical finite differences in the flat limit. Our approach, however, is quite different. While in the flat limit case the number of data sites is fixed and the scaling ε tends to infinity (ε in those papers corresponds to $1/\varepsilon$ here), we have a possibly infinite number of data sites. Moreover, we change the scaling and data sites simultaniously and let both the scaling and the mesh size tend to zero in a very specific way. Hence, we cannot expect to end up with a polynomial interpolant in our case and the similarity to classical interpolation is restricted to the fact that we reconstruct a derivative from function values.

While our analysis is based upon the fact that our spatial discretisation points form a regular grid, it is our goal to extend these results to arbitrary point sets later on.

The paper is organised as follows. In the next section, we will derive our discretisation scheme and state our main convergence result. The third section is devoted to some approximation results which extend results from [26]. The fourth section then deals with the proof of our main convergence result. In the fifth section, we give a general scheme to construct high order kernels. The final section is devoted to a numerical example.

In this paper, we will only consider problems on all of $\mathbb{R}^n$ as the spatial domain. From a practical point of view, this means that we will, where necessary, assume that for a fixed time interval $[0,T]$ there is a compact set $\Omega \subseteq \mathbb{R}^n$ such that the solution $\rho(t) = \rho(t,\cdot)$ of (1.1), (1.2) has support contained in Ω for all $t \in [0,T]$. We will choose Ω large enough such that no boundary effects occur. For the purpose of our analysis, we will without restriction also assume that Ω is convex. In this situation, we obviously only need to know the defining function f on the set

$$M := \{(t, x, \rho(t,x), \nabla\rho(t,x)) \in \mathbb{R}^{2n+2} : t \in [0,T], x \in \Omega\} \subseteq \mathbb{R}^{2n+2}, \tag{1.4}$$

which is also compact since ρ is supposed to be a C^1 function. Again, for the purpose of our analysis, we will extend this set M to a bigger, convex set

$$\widetilde{M} := [0,T] \times \Omega_1 \times \Pi. \tag{1.5}$$

Here Ω_1 is the convex hull of $\bigcup_{x\in\Omega} B_1(x)$, where $B_1(x)$ is the ball of radius 1 and centre x. Moreover, Π is a convex, compact super set of $\{(\rho(t,x), \nabla\rho(t,x)) : t \in [0,T], x \in \Omega\} \subseteq \mathbb{R}^{n+1}$.

Then we may assume that f is smoothly defined on all of $\mathbb{R}^n$ but has support in $\widetilde{M}$. Obviously, this is no restriction under the given assumptions and we can modify any given f that does not satisfy this condition accordingly.

As usual, $W_p^k(\mathbb{R}^n)$ will denote the space of all functions having weak derivatives up to order k in L_p. Its norm will be denoted by $\|\cdot\|_{W_p^k(\mathbb{R}^n)}$ and the semi-norm consisting only of the derivatives of order $|\alpha| = k$ will be denoted by $|\cdot|_{W_p^k(\mathbb{R}^n)}$.

2. The discretisation scheme and its convergence

Our discretisation scheme is a classical kernel-based approximation scheme. We will employ kernels of the following form.

Definition 2.1. A continuous and bounded function $\zeta : \mathbb{R}^n \to \mathbb{R}$ is called a kernel of order $k \geq 1$ if

(i) $\zeta \in L_1(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \zeta(x)dx = 1$,
(ii) $\int_{\mathbb{R}^n} x^\alpha \zeta(x)dx = 0$ for all $\alpha \in \mathbb{N}_0^n$ with $1 \leq |\alpha| \leq k-1$,
(iii) $\int_{\mathbb{R}^n} |x|^k |\zeta(x)|dx < \infty$.

For $k, s \in \mathbb{N}_0$ with $k \geq 1$, we define $\mathcal{K}^{k,s}$ to be the set of all such kernels $\zeta \in C^s(\mathbb{R}^n) \cap W_1^s(\mathbb{R}^n)$ of order k and smoothness s. If $s = 0$ we require $\zeta \in C(\mathbb{R}^n) \cap L_1(\mathbb{R}^n)$. Moreover, we define $\mathcal{K}_c^{k,s}$ to be the subset of such kernels having compact support.

Note that the second condition in Definition 2.1 will only become active if $k \geq 2$.

The smoothness s and the order k of the kernel ζ are not related. However, the order k is related to the smoothness of the Fourier transform of ζ. In particular, the second condition in Definition 2.1 can be expressed as $D^\alpha \widehat{\zeta}(0) = 0$ for all $1 \leq |\alpha| \leq k - 1$. As this is of no importance to us, we will not pursue this any further.

So far, ζ is rather a function than a kernel. It is named a kernel since it is used as a convolution kernel. To be more precise, let $\varepsilon > 0$ and a kernel ζ as before be given and let us define the scaled version $\zeta_\varepsilon(x) = \varepsilon^{-n}\zeta(x/\varepsilon)$, $x \in \mathbb{R}^n$. Then, any function $\rho : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}$ can be approximated by a convolution of the form

$$\rho^\varepsilon(t, x) = (\rho * \zeta_\varepsilon)(t, x) = \int_{\mathbb{R}^n} \rho(t, y)\zeta_\varepsilon(x - y)dy, \qquad (t, x) \in [0, \infty) \times \mathbb{R}^n, \tag{2.1}$$

provided that the integral exists. Note that the convolution is only taken with respect to the spatial variable and not with respect to the time variable.

Next, in particle methods, the approximation ρ^ε is further approximated using a quadrature rule. From now on, we will use the notation

- $x_i = ih$, $i \in \mathbb{Z}^n$,
- $\rho_{x_i} = \rho_{ih} = \rho(\cdot, x_i)$, $i \in \mathbb{Z}^n$,

where $h > 0$ is a given discretisation parameter. Thus, applying a simple composite rectangular rule to the integral in (2.1) yields the new approximation

$$[\rho]_x(t) := [\rho](t, x) := h^n \sum_{j \in \mathbb{Z}^n} \rho_{jh}(t)\zeta_\varepsilon(x - x_j), \qquad (t, x) \in [0, \infty) \times \mathbb{R}^n. \tag{2.2}$$

For simplicity, we will assume here that ρ and ζ are chosen such that the series is well defined. This is, for example, the case if ρ satisfies sufficient decay conditions or if ζ is compactly supported. In this case, $[\rho]$ defines a smooth function in space where the smoothness is determined by the smoothness of ζ. It also defines a smooth function in time, where the smoothness is now determined by the smoothness of ρ in time. The function ζ can also be a fast decaying function like a Gaussian but the analysis becomes more complicated in such a situation.

Note that if ζ is a compactly supported function with support in the unit ball, then ζ_ε has support in the ball about zero with radius ε. Hence, in this situation $\varepsilon > h$ is required since otherwise the sum in (2.2) would reduce to one term for each spatial point x yielding a very efficient but also poor approximation.

Note that $[\rho]$ is already defined if only a countable number of time-dependent (or even constant) functions $\rho_j : [0, \infty) \to \mathbb{R}$, $j \in \mathbb{Z}^n$, is given and ρ_{jh} is replaced by ρ_j. This obvious observation can be used to compute an approximate solution to (1.1) by reducing the original problem to the problem of finding approximate coefficients $\rho_j^{\varepsilon h} : [0, \infty) \to \mathbb{R}$.

Hence, we may define a function depending on time and space by

$$[\rho^{\varepsilon h}]_x(t) = [\rho^{\varepsilon h}](t, x) = h^n \sum_{j \in \mathbb{Z}^n} \rho_j^{\varepsilon h}(t)\zeta_\varepsilon(x - x_j), \quad (t, x) \in [0, \infty) \times \mathbb{R}^n. \tag{2.3}$$

Provided that the kernel ζ is at least $|\alpha|$-times continuously differentiable, spatial derivatives of this function are simply given by

$$\partial^\alpha[\rho^{\varepsilon h}](t,x) = h^n \sum_{j\in\mathbb{Z}^n} \rho_j^{\varepsilon h}(t)\partial^\alpha\zeta_\varepsilon(x-x_j), \qquad (t,x)\in[0,\infty)\times\mathbb{R}^n.$$

In particular, we have

$$\nabla[\rho^{\varepsilon h}]_x(t) = \nabla[\rho^{\varepsilon h}](t,x) = h^n \sum_{j\in\mathbb{Z}^n} \rho_j^{\varepsilon h}(t)\nabla\zeta_\varepsilon(x-x_j), \qquad (t,x)\in[0,\infty)\times\mathbb{R}^n. \tag{2.4}$$

In the next section, we will discuss the approximation properties of this approximation process. The results derived there are essential for the proof of our main result; they are mainly technical improvements of earlier results which can be found in [26].

Before that, we will state our approximation method and our main convergence result.

If we restrict (1.1) to the spatial points $x_i = ih$, $i\in\mathbb{Z}^n$, the initial partial differential equation reduces to the system of differential equations

$$\dot\rho_{ih}(t) = -f(t, ih, \rho(t,ih), \nabla\rho(t,ih)), \qquad (t,i)\in(0,\infty)\times\mathbb{Z}^n.$$

We can now solve the latter equation by approximating the expressions ρ and $\nabla\rho$ on the right-hand side by their approximations $[\rho^{\varepsilon h}]$ and $\nabla[\rho^{\varepsilon h}]$, respectively.

Theorem 2.2. *Let $T>0$ and $I:=[0,T]$. Assume that the solution of problem (1.1), (1.2) for $f\in C^2(I\times\mathbb{R}^{2n+1})$ and $\rho_0\in C_c^r(\mathbb{R}^n)$ satisfies $\rho\in C_c^{2,r}(I\times\mathbb{R}^n)$. Let $\zeta\in\mathcal{K}_c^{k,s}$ be a kernel of order $k\geq 1$.*

For $\varepsilon, h>0$ let $\{\rho_i^{\varepsilon h}\}_{i\in\mathbb{Z}^n}$ be the solution of the initial value problem

$$\begin{aligned} \dot\rho_i^{\varepsilon h} &= -f(.,ih,[\rho^{\varepsilon h}]_{ih}, \nabla[\rho^{\varepsilon h}]_{ih}),\\ \rho_i^{\varepsilon h}(0) &= \rho_0(ih) \end{aligned}$$

with $[\rho^{\varepsilon h}]_{ih}$ and $\nabla[\rho^{\varepsilon h}]_{ih}$ as in (2.3) and (2.4), respectively. If the parameters obey the relations

$$r\geq\max\{k+1,\ell\},\quad k\geq 2+\frac{n}{2},\quad s>\ell>n \quad \text{and} \quad h\leq\varepsilon^{1+(3+\frac{n}{2})/\ell},$$

then there is a constant $C=C(f,\rho,T,\zeta)>0$ independent of ε and h such that for all $\alpha\in\mathbb{N}_0^n$ with $|\alpha|\leq 1$ it holds that

$$\|\partial^\alpha\rho-\partial^\alpha[\rho^{\varepsilon h}]\|_{L_\infty(\mathbb{R}^n)} \leq C\varepsilon^{-|\alpha|-\frac{n}{2}}\left(\varepsilon^k+\frac{h^\ell}{\varepsilon^{\ell+1}}\right)\leq C\varepsilon^{k-|\alpha|-\frac{n}{2}}$$

uniformly on $[0,T]$.

We will postpone the proof of this theorem. However, since its proof and the proof of certain auxiliary results will use a discrete L_p-norm, which is quite standard in this context, we will introduce this norm now.

Definition 2.3. Let $1\leq p\leq\infty$ and $h>0$. For a given sequence $(\rho_i)_{i\in\mathbb{Z}^n}\in\ell_p$ we define the h-dependent ℓ_p-norm by

$$\|\rho\|_{p,h} := \begin{cases} \left(h^n\sum_{i\in\mathbb{Z}^n}|\rho_i|^p\right)^{1/p} & \text{for } 1\leq p<\infty,\\ \sup_{i\in\mathbb{Z}^n}|\rho_i| & \text{for } p=\infty. \end{cases} \tag{2.5}$$

Obviously, we are particularly interested in the situation $\rho_i=\rho(ih)$ in which $\|\cdot\|_{p,h}$ becomes an approximation to the continuous $L_p(\mathbb{R}^n)$ norm.

3. Auxiliary results on quasi-interpolation

The main result of this section is essential for proving our convergence theorem, Theorem 2.2. It specifies the approximation power of our discretisation technique for approximating functions. It generalises an earlier result of [26], particularly by also providing estimates for derivatives.

To proof this theorem, we require two auxiliary results, Theorem 3.1 and Lemma 4.4 from [26].

The first result analyses the quadrature error of the composite rectangular rule and is quite standard. Its proof is based on the Bramble-Hilbert lemma.

Lemma 3.1. *Let $\ell \in \mathbb{N}$ with $\ell > n \geq 1$. Then there exists a constant $C > 0$ independent of h such that for all functions $\rho \in W_1^\ell(\mathbb{R}^n)$ we have*

$$\left| \int_{\mathbb{R}^n} \rho(y)dy - h^n \sum_{j \in \mathbb{Z}^n} \rho(jh) \right| \leq Ch^\ell |\rho|_{W_1^\ell(\mathbb{R}^n)}.$$

The second result analyses the convolution error. Its proof is mainly based upon the Taylor expansion.

Lemma 3.2. *Let $\zeta \in \mathcal{K}^{k,s}$ be a kernel function of order $k \geq 1$ and $\rho \in C^r(\mathbb{R}^n) \cap W_p^r(\mathbb{R}^n)$ for suitable constants $1 \leq p \leq \infty$ and $k \leq r \in \mathbb{N}$. Then, there is a constant $C = C(\zeta) > 0$ such that for any $\varepsilon > 0$ and $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq r - k$ the following holds:*

$$\|\partial^\alpha \rho - \partial^\alpha \rho^\varepsilon\|_{L_p(\mathbb{R}^n)} = \|\partial^\alpha \rho - \rho * \partial^\alpha \zeta_\varepsilon\|_{L_p(\mathbb{R}^n)} \leq C\varepsilon^k |\rho|_{W_p^{k+|\alpha|}(\mathbb{R}^n)}.$$

The kernel itself satisfies the following norm estimates, which we will also require later on.

Lemma 3.3. *Let $1 \leq p, q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, $s \in \mathbb{N}$ and $\zeta \in W_p^s(\mathbb{R}^n)$. For $\varepsilon > 0$ let $\zeta_\varepsilon(x) = \varepsilon^{-n}\zeta(x/\varepsilon)$, $x \in \mathbb{R}^n$. Let $\alpha \in \mathbb{N}_0^n$.*

(a) *If $|\alpha| \leq s$, then*

$$\|\partial^\alpha \zeta_\varepsilon\|_{L_p(\mathbb{R}^n)} = \varepsilon^{-\frac{n}{q}-|\alpha|} \|\partial^\alpha \zeta\|_{L_p(\mathbb{R}^n)}.$$

(b) *If $\zeta \in W_2^s(\mathbb{R}^n) \cap W_\infty^s(\mathbb{R}^n)$ and if $|\alpha| < s - n$, then for every $p \in [2, \infty]$ there is a constant $C = C(\zeta, \alpha, p) > 0$ such that for every $\varepsilon > h > 0$ we have*

$$\|\partial^\alpha \zeta_\varepsilon\|_{p,h} \leq C\varepsilon^{-\frac{n}{q}-|\alpha|}. \tag{3.1}$$

(c) *If $\zeta \in \mathcal{K}_c^{k,s}$ for any $k \in \mathbb{N}$, then (3.1) even holds for all $p \in [1, \infty]$.*

Proof. The first statement is obvious for $p = \infty$ and follows from the transformation formula for all other p.

The second statement also immediately follows for $p = \infty$. For $p = 2$, we note that Lemma 3.1 yields

$$\begin{aligned} \|\partial^\alpha \zeta_\varepsilon\|_{2,h}^2 &\leq \left| h^n \sum_{i \in \mathbb{Z}^n} |\partial^\alpha \zeta_\varepsilon(ih)|^2 - \int_{\mathbb{R}^n} |\partial^\alpha \zeta_\varepsilon(x)|^2 dx \right| + \|\partial^\alpha \zeta_\varepsilon\|_{L_2(\mathbb{R}^n)}^2 \\ &\leq Ch^\ell |g|_{W_1^\ell(\mathbb{R}^n)} + \|\partial^\alpha \zeta_\varepsilon\|_{L_2(\mathbb{R}^n)}^2, \end{aligned} \tag{3.2}$$

as long as $g = |\partial^\alpha \zeta_\varepsilon|^2 \in W_1^\ell(\mathbb{R}^n)$ with $n < \ell \in \mathbb{N}$ which is guaranteed by our assumptions because of $\ell := s - |\alpha| > n$ and

$$\begin{aligned}
|g|_{W_1^\ell(\mathbb{R}^n)} &= \sum_{|\beta|=\ell} \int_{\mathbb{R}^n} |\partial^\beta (\partial^\alpha \zeta_\varepsilon(x))^2| dx \\
&\le \sum_{|\beta|+|\gamma|=\ell} C \int_{\mathbb{R}^n} |\partial^{\alpha+\beta}\zeta_\varepsilon(x)||\partial^{\alpha+\gamma}\zeta_\varepsilon(x)| dx \\
&\le \sum_{|\beta|+|\gamma|=\ell} C \|\partial^{\alpha+\beta}\zeta_\varepsilon\|_{L_2(\mathbb{R}^n)} \|\partial^{\alpha+\gamma}\zeta_\varepsilon\|_{L_2(\mathbb{R}^n)} \\
&\le C\varepsilon^{-n-2|\alpha|-\ell}.
\end{aligned}$$

This gives, together with $0 < h < \varepsilon$ and (3.2),

$$\|\partial^\alpha \zeta_\varepsilon\|_{2,h}^2 \le C(h^\ell \varepsilon^{-n-2|\alpha|-\ell} + \varepsilon^{n-2|\alpha|}) \le C\varepsilon^{-n-2|\alpha|}.$$

Since we also know that $\|\partial^\alpha \zeta_\varepsilon\|_{\infty,h} \le C\varepsilon^{-n-|\alpha|}$, we can conclude via interpolation that

$$\begin{aligned}
\|\partial^\alpha \zeta_\varepsilon\|_{p,h} &\le \|\partial^\alpha \zeta_\varepsilon\|_{2,h}^{\frac{2}{p}} \|\partial^\alpha \zeta_\varepsilon\|_{\infty,h}^{1-\frac{2}{p}} \\
&\le C\varepsilon^{-\left(\frac{n}{2}+|\alpha|\right)\frac{2}{p}} \varepsilon^{-(n+|\alpha|)\left(1-\frac{2}{p}\right)} \\
&= C\varepsilon^{-\frac{n}{q}-|\alpha|}.
\end{aligned}$$

Finally, if $\zeta \in \mathcal{K}_c^{k,s}$, then we automatically have $\zeta \in W_1^s(\mathbb{R}^n) \cap W_\infty^s(\mathbb{R}^n)$. Furthermore, since the number of $i \in \mathbb{Z}^n$ with $i\frac{h}{\varepsilon}$ in the support of ζ is bounded by a constant times $(\varepsilon/h)^n$, we see that

$$\|\partial^\alpha \zeta_\varepsilon\|_{1,h} = h^n \sum_{i\in\mathbb{Z}^n} |\partial^\alpha \zeta(ih/\varepsilon)| \le C\varepsilon^{-|\alpha|}.$$

For general p the result then follows by interpolation between $p = 1$ and $p = \infty$ again. $\square$

After these preparations, we can state and prove a result concerning the approximation power of such quasi-interpolants. Note, that (2.3) can also be defined for a function ρ, which does not depend on time. In this case ρ_j is just a constant for each $j \in \mathbb{Z}^n$. For simplicity, we state the next result under this assumption. However, if ρ depends on $t \in [0, T]$, then the result obviously holds pointwise for all $t \in [0, T]$.

Theorem 3.4. *Let $1 \le p \le \infty$. Let $\rho \in C^r(\mathbb{R}^n) \cap W_p^r(\mathbb{R}^n)$ and $\zeta \in \mathcal{K}^{k,s}$ with $r, k, s \in \mathbb{N}$ be given. Assume that $r \ge k$ and $r, s \ge \ell > n$ with $\ell \in \mathbb{N}$. Finally, let $\varepsilon > h > 0$. Then, there is a constant $C = C(\zeta) > 0$ independent of ε and h such that for every $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \le \min\{r - k, s - \ell\}$ the error between ρ and its approximation $[\rho] = h^n \sum_i \rho(ih)\zeta_\varepsilon(\cdot - ih)$ can be bounded by*

$$\|\partial^\alpha \rho - \partial^\alpha [\rho]\|_{L_p(\mathbb{R}^n)} \le C\left(\varepsilon^k \|\rho\|_{W_p^{k+|\alpha|}(\mathbb{R}^n)} + \frac{h^\ell}{\varepsilon^{\ell+|\alpha|}} \|\rho\|_{W_p^\ell(\mathbb{R}^n)}\right). \tag{3.3}$$

Proof. We can split the error on the left-hand side of (3.3) into a convolution and a quadrature error:

$$\|\partial^\alpha \rho - \partial^\alpha[\rho]\|_{L_p(\mathbb{R}^n)} \le \|\partial^\alpha \rho - \partial^\alpha \rho^\varepsilon\|_{L_p(\mathbb{R}^n)} + \|\partial^\alpha \rho^\varepsilon - \partial^\alpha[\rho]\|_{L_p(\mathbb{R}^n)}.$$

Using Lemma 3.2, we see that the first term on the right-hand side can be bounded by

$$\|\partial^\alpha \rho - \partial^\alpha \rho^\varepsilon\|_{L_p(\mathbb{R}^n)} \le C\varepsilon^k |\rho|_{W_p^{k+|\alpha|}(\mathbb{R}^n)},$$

provided that $r \ge k + |\alpha|$. For the second term, we first note that we have for $x \in \mathbb{R}^n$ fixed that

$$|\partial^\alpha \rho^\varepsilon(x) - \partial^\alpha[\rho](x)| = \left| \int_{\mathbb{R}^n} \rho(y) \partial^\alpha \zeta_\varepsilon(x-y)dy - \sum_{j\in\mathbb{Z}^n} h^n \rho(jh) \partial^\alpha \zeta_\varepsilon(x - jh) \right|.$$

Hence, if $r \ge \ell$ and $s - |\alpha| \ge \ell$ for $\ell > n$, then we can use Lemma 3.1 to derive

$$\begin{aligned} |\partial^\alpha \rho^\varepsilon(x) - \partial^\alpha[\rho](x)| &\le Ch^\ell \, |\rho\, \partial^\alpha \zeta_\varepsilon(x - \cdot)|_{W_1^\ell(\mathbb{R}^n)} \\ &\le Ch^\ell \sum_{|\beta|=\ell} \int_{\mathbb{R}^n} |\partial_y^\beta(\rho(y)\partial^\alpha \zeta_\varepsilon(x-y))|dy \\ &\le Ch^\ell \sum_{|\beta|,|\gamma|\le\ell} \int_{\mathbb{R}^n} |\partial^\beta \rho(y)|\, |\partial^{\alpha+\gamma}\zeta_\varepsilon(x-y)|dy \\ &= Ch^\ell \sum_{|\beta|,|\gamma|\le\ell} |\partial^\beta \rho| * |\partial^{\alpha+\gamma}\zeta_\varepsilon|(x). \end{aligned}$$

Young's inequality finally yields

$$\|\partial^\alpha \rho^\varepsilon - \partial^\alpha[\rho]\|_{L_p(\mathbb{R}^n)} \le Ch^\ell \|\rho\|_{W_p^\ell(\mathbb{R}^n)} \|\zeta_\varepsilon\|_{W_1^{\ell+|\alpha|}(\mathbb{R}^n)} \le C(\zeta)\frac{h^\ell}{\varepsilon^{\ell+|\alpha|}}\|\rho\|_{W_p^\ell(\mathbb{R}^n)},$$

where we have also used Lemma 3.3 in the last step. □

Remark 3.5. We will use this result in particular to bound estimates on first order derivatives. Hence, the smoothness r and s of ρ and ζ, respectively, have to satisfy $r \ge \max\{\ell, k+1\}$ and $s \ge \ell + 1$ with $\ell > n$.

Finally, we want to state and prove a result, which is interesting on its own. It shows that the kernel and its derivatives provide a Bessel sequence. We use it in our proof to establish bounds on the L_∞ errors of our approximation based on error bounds for the coefficients in the discrete L_2-norm. It seems worthwhile to mention that for this purpose property (ii) of the kernel definition, i.e., the property that defines the order of the kernel, is not required.

Theorem 3.6. *Let $1 \le p, q \le \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Let $(a_i)_{i\in\mathbb{Z}^n} \in \ell_q(\mathbb{R})$ and let ζ be a kernel as in Lemma 3.3, i.e., either $\zeta \in W_2^s(\mathbb{R}^n) \cap W_\infty^s(\mathbb{R}^n)$ for $p \in [2,\infty]$ or $\zeta \in \mathcal{K}_c^{k,s}$ for $p \in [1,\infty]$. If $\alpha \in \mathbb{N}_0^n$ satisfies $|\alpha| < s - n$, then there is a constant $C > 0$ such that for $\varepsilon > h$ we have*

$$\|\partial^\alpha[a]\|_{\infty,h} = \sup_{i\in\mathbb{Z}^n} \left| h^n \sum_{j\in\mathbb{Z}^n} a_j \partial^\alpha \zeta_\varepsilon(ih - jh) \right| \le C\|a\|_{q,h}\varepsilon^{-\frac{n}{q}-|\alpha|}.$$

Moreover, if $\zeta \in \mathcal{K}_c^{k,s}$ *and* $|\alpha| < s - n$*, then for* $p \in [1, \infty]$ *there is a constant* $C > 0$ *such that*

$$\|\partial^\alpha[a]\|_{L_\infty(\mathbb{R}^n)} \le C\|a\|_{q,h}\varepsilon^{-\frac{n}{q}-|\alpha|}. \tag{3.4}$$

Proof. For ζ and p as specified in the theorem the first statement follows easily by using Hölder's inequality and Lemma 3.3, since for each $i \in \mathbb{Z}^n$ we have

$$\begin{aligned}\left|h^n \sum_{j\in\mathbb{Z}^n} a_j \partial^\alpha \zeta_\varepsilon(ih - jh)\right| &\le h^n \sum_{j\in\mathbb{Z}^n} |a_j||\partial^\alpha \zeta_\varepsilon(ih - jh)| \\ &\le \|a\|_{q,h}\|\partial^\alpha\zeta_\varepsilon(ih - .)\|_{p,h} \\ &\le \|a\|_{q,h}\|\partial^\alpha\zeta_\varepsilon\|_{p,h} \\ &\le C\|a\|_{q,h}\varepsilon^{-\frac{n}{q}-|\alpha|}.\end{aligned}$$

The second statement follows similarly. For $1 \le p < \infty$ it holds that

$$\begin{aligned}\|\partial^\alpha[a]\|_{L_\infty(\mathbb{R}^n)} &= \sup_{x\in\mathbb{R}^n}\left|h^n \sum_{j\in\mathbb{Z}^n} a_j \partial^\alpha\zeta_\varepsilon(x - jh)\right| \\ &\le \|a\|_{q,h} \sup_{x\in\mathbb{R}^n}\left(h^n \sum_{j\in\mathbb{Z}^n} |\partial^\alpha\zeta_\varepsilon(x - jh)|^p\right)^{1/p}.\end{aligned}$$

Using the compact support of ζ_ε shows once again that the sum is only a sum over at most $C(\varepsilon/h)^n$ terms so that we can continue with the estimate

$$\|\partial^\alpha[a]\|_{L_\infty(\mathbb{R}^n)} \le C\|a\|_{q,h}\left(h^n(\varepsilon/h)^n\varepsilon^{-(n+|\alpha|)p}\right)^{1/p} = C\|a\|_{q,h}\varepsilon^{-\frac{n}{q}-|\alpha|}.$$

The remaining case $p = \infty$ is trivial. ☐

4. Proof of Convergence

In this section, we will prove Theorem 2.2. As usual, to simplify the notation, we will suppress the time variable whenever possible.

Since our spatial discretisation technique immediately leads to an infinite system of ordinary differential equations with solutions $\{\rho_i^{\varepsilon h}\}$, it is natural to use the discrete norms defined in Definition 2.3 and to bound the error

$$e_i = \rho_{ih} - \rho_i^{\varepsilon h}, \qquad i \in \mathbb{Z}^n,$$

using these norms. In this context, it is also quite natural to split the error into a consistency and a stability error. We have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|e\|_{2,h}^2 &= \frac{1}{2}\frac{d}{dt}h^n\sum_{i\in\mathbb{Z}^n} e_i^2 = h^n \sum_{i\in\mathbb{Z}^n} e_i\dot{e}_i \\ &= -h^n\sum_{i\in\mathbb{Z}^n} e_i\left(f(t, ih, \rho_{ih}, \nabla\rho_{ih}) - f(t, ih, [\rho^{\varepsilon h}]_{ih}, \nabla[\rho^{\varepsilon h}]_{ih})\right) \\ &= -h^n\sum_{i\in\mathbb{Z}^n} e_i\left(f(t, ih, \rho_{ih}, \nabla\rho_{ih}) - f(t, ih, [\rho]_{ih}, \nabla[\rho]_{ih})\right) \\ &\quad - h^n\sum_{i\in\mathbb{Z}^n} e_i\left(f(t, ih, [\rho]_{ih}, \nabla[\rho]_{ih}) - f(t, ih, [\rho^{\varepsilon h}]_{ih}, \nabla[\rho^{\varepsilon h}]_{ih})\right) \\ &=: -(e_c + e_s).\end{aligned} \tag{4.1}$$

The first term in the last expression,

$$e_c = e_c(t) := h^n \sum_{i\in\mathbb{Z}^n} e_i \left(f(t, ih, \rho_{ih}, \nabla\rho_{ih}) - f(t, ih, [\rho]_{ih}, \nabla[\rho]_{ih})\right), \tag{4.2}$$

represents the consistency error of the method while the second term

$$e_s = e_s(t) := h^n \sum_{i\in\mathbb{Z}^n} e_i \left(f(t, ih, [\rho]_{ih}, \nabla[\rho]_{ih}) - f(t, ih, [\rho^{\varepsilon h}]_{ih}, \nabla[\rho^{\varepsilon h}]_{ih})\right) \tag{4.3}$$

represents the stability error. We will now bound both errors separately, starting with the consistency error.

For the convenience of the reader, we recall our initial assumptions. We assume that the support of the solution $\rho(t,\cdot)$ is contained in a compact, convex set $\Omega \subseteq \mathbb{R}^n$ for all $t \in [0,T]$. We have defined the set M and $\widetilde{M}$ in (1.4) and (1.5), respectively, to be

$$\begin{aligned} &M := \{(t, x, \rho(t,x), \nabla\rho(t,x)) \in \mathbb{R}^{2n+2} : t \in [0,T], x \in \Omega\} \subseteq \mathbb{R}^{2n+2},\\ &\widetilde{M} := [0,T] \times \Omega_1 \times \Pi. \end{aligned}$$

Here Ω_1 is the convex hull of $\bigcup_{x\in\Omega} B_1(x)$, where $B_1(x)$ is the ball of radius 1 and centre x. Moreover, Π is a convex, compact super set of $\{(\rho(t,x), \nabla\rho(t,x)) : t \in [0,T], x \in \Omega\} \subseteq \mathbb{R}^{n+1}$.

We have also assumed that the defining function f is sufficiently smooth and has compact support in the compact and convex set $\widetilde{M} \supseteq M$. As usual, we will consider f to be defined on all of $\mathbb{R}^n$ with zero value outside $\widetilde{M}$.

The reason for this technical definition is the following one. Suppose that our kernel ζ has support in the unit ball and that $\rho(t,\cdot)$ has support in Ω, then, obviously

$$[\rho] = \sum_{j\in\mathbb{Z}^d} \rho_{jh}(t)\zeta_\varepsilon(\cdot - jh)$$

has support in Ω_1 for all $0 < \varepsilon \le 1$. Moreover, the convexity of $\widetilde{M}$ guarantees that all connecting line segments between two points in $\widetilde{M}$ are also contained in $\widetilde{M}$.

Proposition 4.1. *Let $T > 0$. Let $\zeta \in \mathcal{K}_c^{k,s}$ with $k \in \mathbb{N}$ and $s > \ell > n$ for an $\ell \in \mathbb{N}$ and with support in the unit ball. Let ρ be the solution of* (1.1),(1.2) *Assume that $\rho(t,\cdot) \in C_c^r(\mathbb{R}^n)$ with $r \ge \max\{k+1, \ell\}$ and that the support of $\rho(t,\cdot)$ is contained in the compact set $\Omega \subseteq \mathbb{R}^n$ for all $t \in [0,T]$. Suppose further that $f \in C_c^1(\widetilde{M})$, with $\widetilde{M}$ defined in* (1.5)*. Then, there is a constant $C > 0$ depending on f, ζ, ρ and M such that the consistency error e_c from* (4.2) *can be bounded by*

$$|e_c(t)| \le C\left(\varepsilon^k + \frac{h^\ell}{\varepsilon^{\ell+1}}\right)^2 + \frac{1}{2}\|e(t)\|_{2,h}^2, \qquad t \in [0,T],$$

for all $0 < h \le \varepsilon^{1+\frac{2}{\ell}}$ sufficiently small.

Proof. Our assumption on ρ immediately shows $\rho \in W^r_\infty(\mathbb{R}^n)$. Thus, Theorem 3.4 gives, in particular, for each $i \in \mathbb{Z}^n$,

$$|\rho_{ih} - [\rho]_{ih}| \le C\left(\varepsilon^k + \frac{h^\ell}{\varepsilon^\ell}\right) \tag{4.4}$$

and, since $r \ge \max\{k+1, \ell\}$ and $s \ge \ell + 1$,

$$|\nabla\rho_{ih} - \nabla[\rho]_{ih}| \le C\left(\varepsilon^k + \frac{h^\ell}{\varepsilon^{\ell+1}}\right), \tag{4.5}$$

where the constant $C > 0$ depends on ρ and ζ. Next, let

$$I_{\varepsilon h} := \{i \in \mathbb{Z}^n : ih \in \Omega_\varepsilon\}$$

and note that the cardinality of $I_{\varepsilon h}$ can be bounded by a constant times h^{-n} since Ω is compact. The explanation given above shows that for $i \notin I_{\varepsilon h}$ we have $\rho_{ih} = [\rho]_{ih} = 0$.

Using the Cauchy-Schwarz inequality and the mean value theorem, there are $\eta_i \in \mathbb{R}$ on the line segment between ρ_{ih} and $[\rho]_{ih}$ and $\xi_i \in \mathbb{R}^n$ on the line segment between $\nabla\rho_{ih}$ and $\nabla[\rho]_{ih}$ such that

$$\begin{aligned}
|e_c| &= \left|h^n \sum_{i\in\mathbb{Z}^n} e_i\left(f(t, ih, \rho_{ih}, \nabla\rho_{ih}) - f(t, ih, [\rho]_{ih}, \nabla[\rho]_{ih})\right)\right| \\
&= \left|h^n \sum_{i\in I_{\varepsilon h}} e_i Df(t, ih, \cdot, \cdot)|_{(\eta_i,\xi_i)} \cdot (\rho_{ih} - [\rho]_{ih}, \nabla\rho_{ih} - \nabla[\rho]_{ih})\right| \\
&\le |f|_{W^1_\infty(\widetilde{M})} \|e\|_{2,h}\left(h^n \sum_{i\in I_{\varepsilon h}} |\rho_{ih} - [\rho]_{ih}|^2 + |\nabla\rho_{ih} - \nabla[\rho]_{ih}|^2\right)^{\frac{1}{2}} \\
&\le C(f, \zeta, \rho)\left(\varepsilon^k + \frac{h^\ell}{\varepsilon^{\ell+1}}\right)\|e\|_{2,h} \\
&\le C(f, \zeta, \rho)\left(\varepsilon^k + \frac{h^\ell}{\varepsilon^{\ell+1}}\right)^2 + \frac{1}{2}\|e\|^2_{2,h},
\end{aligned}$$

where we have also used (4.4) and (4.5) as well as the fact that $(\eta_i, \xi_i) \in M_1$ for $\varepsilon > 0$ small. ∎

The next step in our error analysis is to bound the stability error (4.3). Here, the proof is more demanding and requires a bootstrap argument, i.e., we need to make an assumption on the total error to prove the following bound, which we have to verify later on.

Theorem 4.2. *Let $T > 0$. Let $\zeta \in \mathcal{K}^{k,s}_c$ with $s > \ell > n$ for an $\ell \in \mathbb{N}$ being even and with support in the unit ball. Let ρ be the solution of* (1.1),(1.2). *Assume that $\rho(t, \cdot) \in C^r_c(\mathbb{R}^n)$ with $r > k \ge 1$ and that the support of $\rho(t, \cdot)$ is contained in the compact set $\Omega \subseteq \mathbb{R}^n$ for all $t \in [0, T]$. Suppose further that $f \in C^2(\widetilde{M})$, where $\widetilde{M}$ is defined in* (1.5). *Assume that $h \le \varepsilon^{1+2/\ell}$. Assume finally, that the error satisfies*

$$\|e(t)\|_{2,h} \le C_1\varepsilon^{2+\frac{n}{2}}, \qquad t \in [0, T], \tag{4.6}$$

with a constant $C_1 > 0$ independent of h and ε. Then, there is a constant $C > 0$ independent of $\varepsilon > 0$ and $h > 0$ such that the stability error e_s from (4.3) *can be bounded by*

$$|e_s(t)| \leq C\|e(t)\|_{2,h}^2, \qquad t \in [0,T],$$

provided $\varepsilon > 0$ is sufficiently small.

Proof. We start by further splitting the error e_s into

$$\begin{aligned} e_s =& - h^n \sum_{i\in\mathbb{Z}^n} e_i \left(f(t, ih, [\rho]_{ih}, \nabla[\rho]_{ih}) - f(t, ih, [\rho^{\varepsilon h}]_{ih}, \nabla[\rho^{\varepsilon h}]_{ih})\right) \\ =:& e_{s1} + e_{s2} \end{aligned}$$

with

$$e_{s1} := - h^n \sum_{i\in\mathbb{Z}^n} e_i \left(f(t, ih, [\rho]_{ih}, \nabla[\rho]_{ih}) - f(t, ih, [\rho]_{ih}, \nabla[\rho^{\varepsilon h}]_{ih})\right), \tag{4.7}$$

$$e_{s2} := - h^n \sum_{i\in\mathbb{Z}^n} e_i \left(f(t, ih, [\rho]_{ih}, \nabla[\rho^{\varepsilon h}]_{ih}) - f(t, ih, [\rho^{\varepsilon h}]_{ih}, \nabla[\rho^{\varepsilon h}]_{ih})\right). \tag{4.8}$$

Note that our assumption on the support of f means that all the sums are actually finite sums, summing at most over those indices $i \in \mathbb{Z}^n$ with $ih \in \Omega_1$.

We will now bound each term separately, starting with (4.7). Using once again the mean value theorem, this time only with respect to the last argument of f, yields positions $\xi_i \in \mathbb{R}^n$ on the line segment connecting $\nabla[\rho]_{ih}$ and $\nabla[\rho^{\varepsilon h}]_{ih}$ such that

$$\begin{aligned} e_{s1} &= -h^n \sum_{i\in\mathbb{Z}^n} e_i \left(f(t, ih, [\rho]_{ih}, \nabla[\rho]_{ih}) - f(t, ih, [\rho]_{ih}, \nabla[\rho^{\varepsilon h}]_{ih})\right) \\ &= -h^n \sum_{i\in\mathbb{Z}^n} e_i Df(t, ih, [\rho]_{ih}, \cdot)|_{\xi_i} \cdot \left(\nabla[\rho]_{ih} - \nabla[\rho^{\varepsilon h}]_{ih}\right) \\ &= -h^n \sum_{i\in\mathbb{Z}^n} e_i Df(t, ih, [\rho]_{ih}, \cdot)|_{\xi_i} \\ &\qquad \cdot \left(h^n \sum_{j\in\mathbb{Z}^n} (\rho_j - \rho_j^{\varepsilon h}) \nabla\zeta_\varepsilon(ih - jh)\right) \\ &= -h^{2n} \sum_{i,j\in\mathbb{Z}^n} e_i e_j Df(t, ih, [\rho]_{ih}, \cdot)|_{\xi_i} \cdot \nabla\zeta_\varepsilon(ih - jh). \end{aligned} \tag{4.9}$$

Next, note that the assumption that the kernel ζ is an even function means, in particular, that $\nabla\zeta_\varepsilon(-\cdot) = -\nabla\zeta_\varepsilon(\cdot)$ and hence $\nabla\zeta_\varepsilon(0) = 0$. To make use of this anti-symmetry, we partition $\mathbb{Z}^n \times \mathbb{Z}^n$ disjointly into $\mathbb{Z}^n \times \mathbb{Z}^n = \Lambda \cup \overline{\Lambda} \cup \{(i,i) : i \in \mathbb{Z}^n\}$, where $\Lambda, \overline{\Lambda} \subseteq \mathbb{Z}^n \times \mathbb{Z}^n$ are such that for every $i \neq j$ we have $(i,j) \in \Lambda$ if and only if $(j,i) \in \overline{\Lambda}$.

Since $\nabla\zeta_\varepsilon(0)=0$, we see that we can ignore all entries in the sum (4.9) corresponding to indices (i,i), $i\in\mathbb{Z}^n$. Hence, we can continue to rewrite e_{s1} by

(4.10)

$$\begin{aligned}
e_{s1} &= -h^{2n}\sum_{i,j\in\mathbb{Z}^n} e_ie_jDf(t,ih,[\rho]_{ih},\cdot)|_{\xi_i}\cdot\nabla\zeta_\varepsilon(ih-jh)\\
&= -h^{2n}\left(\sum_{(i,j)\in\Lambda}+\sum_{(i,j)\in\overline{\Lambda}}\right) e_ie_jDf(t,ih,[\rho]_{ih},\cdot)|_{\xi_i}\cdot\nabla\zeta_\varepsilon(ih-jh)\\
&= -h^{2n}\sum_{(i,j)\in\Lambda} e_ie_jDf(t,ih,[\rho]_{ih},\cdot)|_{\xi_i}\cdot\nabla\zeta_\varepsilon(ih-jh)\\
&\quad - h^{2n}\sum_{(j,i)\in\overline{\Lambda}} e_je_iDf(t,jh,[\rho]_{jh},\cdot)|_{\xi_j}\cdot\nabla\zeta_\varepsilon(jh-ih)\\
&= -h^{2n}\sum_{(i,j)\in\Lambda} e_ie_jDf(t,ih,[\rho]_{ih},\cdot)|_{\xi_i}\cdot\nabla\zeta_\varepsilon(ih-jh)\\
&\quad - h^{2n}\sum_{(i,j)\in\Lambda} e_je_iDf(t,jh,[\rho]_{jh},\cdot)|_{\xi_j}\cdot(-\nabla\zeta_\varepsilon(ih-jh))\\
&= -h^{2n}\sum_{(i,j)\in\Lambda} e_ie_j\left(Df(t,ih,[\rho]_{ih},\cdot)|_{\xi_i}-Df(t,jh,[\rho]_{jh},\cdot)|_{\xi_j}\right)\cdot\nabla\zeta_\varepsilon(ih-jh)\\
&= -h^{2n}\sum_{(i,j)\in\Lambda} e_ie_j\Delta_{ij}\cdot\nabla\zeta_\varepsilon(ih-jh),
\end{aligned}$$

where we have introduced the notation

$$\Delta_{ij}:=Df(t,ih,[\rho]_{ih},\cdot)|_{\xi_i}-Df(t,jh,[\rho]_{jh},\cdot)|_{\xi_j},$$

which obviously satisfies $\Delta_{ij}=-\Delta_{ji}$. Moreover, we can simply set $\Delta_{ij}=0$ for $(i,j)\in\Lambda$ satisfying $|ih-jh|\ge\varepsilon$, since in this situation our compactly supported kernel makes sure that the corresponding term in (4.10) is already zero.

Having this in mind, we can now derive the following bound on e_{s1}:

$$\begin{aligned}
|e_{s1}| &\le h^{2n}\sum_{(i,j)\in\Lambda}\frac{1}{2}(e_i^2+e_j^2)|\Delta_{ij}|\,|\nabla\zeta_\varepsilon(ih-jh)|\\
&\le h^{2n}\sum_{i,j\in\mathbb{Z}^n} e_i^2|\Delta_{ij}|\,|\nabla\zeta_\varepsilon(ih-jh)|\\
&= h^n\sum_{i\in\mathbb{Z}^n}e_i^2\left(h^n\sum_{j\in\mathbb{Z}^n}|\Delta_{ij}|\,|\nabla\zeta_\varepsilon(ih-jh)|\right)\\
&\le \|e\|_{2,h}^2\max_{i\in\mathbb{Z}^n}h^n\sum_{j\in\mathbb{Z}^n}|\Delta_{ij}|\,|\nabla\zeta_\varepsilon(ih-jh)|. \qquad (4.11)
\end{aligned}$$

To bound this further, we need a thorough estimate for the term $|\Delta_{ij}|$, which, in particular, has to compensate the $1/\varepsilon$ factor coming from $|\nabla\zeta_\varepsilon(ih-jh)|$. We have

$$\begin{aligned}
|\Delta_{ij}| &= |Df(t,ih,[\rho]_{ih},\xi_i)-Df(t,jh,[\rho]_{jh},\xi_j)|\\
&\le \|f\|_{W^2_\infty(\widetilde{M})}\left(|ih-jh|+|[\rho]_{ih}-[\rho]_{jh}|+|\xi_i-\xi_j|\right)\\
&\le \|f\|_{W^2_\infty(\widetilde{M})}\left(\varepsilon+|[\rho]_{ih}-[\rho]_{jh}|+|\xi_i-\xi_j|\right), \qquad (4.12)
\end{aligned}$$

where we used the fact that we only have to consider indices (i,j) with $|ih-jh| \le \varepsilon$. To continue our estimate, we use

$$|\xi_i - \xi_j| \le |\xi_i - \nabla[\rho]_{ih}| + |\nabla[\rho]_{ih} - \nabla[\rho]_{jh}| + |\nabla[\rho]_{jh} - \xi_j|. \tag{4.13}$$

Here, the second term on the right-hand side, as well as the term $|[\rho]_{ih} - [\rho]_{jh}|$ in (4.12) can be bounded as follows. For $\alpha \in \mathbb{N}_0^n$ with $|\alpha| = 0$ or $|\alpha| = 1$ we have, using Theorem 3.4,

$$\begin{aligned}
|\partial^\alpha[\rho]_{ih} - \partial^\alpha[\rho]_{jh}| &\le |\partial^\alpha[\rho]_{ih} - \partial^\alpha\rho_{ih}| + |\partial^\alpha\rho_{ih} - \partial^\alpha\rho_{jh}| + |\partial^\alpha\rho_{jh} - \partial^\alpha[\rho]_{jh}| \\
&\le C(\rho)\left(\varepsilon^k + \frac{h^\ell}{\varepsilon^{\ell+|\alpha|}}\right) + \|\rho\|_{W^1_\infty(\Omega)}|ih - jh| \\
&\le C(\rho)\varepsilon,
\end{aligned}$$

since $k \ge 1$ and $h \le \varepsilon^{1+2/\ell}$. For the first and last term in (4.13) we use the fact that ξ_i lies on the line segment connecting $\nabla[\rho]_{ih}$ and $\nabla[\rho^{\varepsilon h}]_{ih}$ such that we can derive the estimate

$$\begin{aligned}
|\Delta_{ij}| &\le \|f\|_{W^2_\infty(\widetilde{M})}\left(\varepsilon + |[\rho]_{ih} - [\rho]_{jh}| + |\xi_i - \xi_j|\right) \\
&\le C(f,\rho)\left(\varepsilon + \|\nabla[\rho]_\cdot - \nabla[\rho^{\varepsilon h}]_\cdot\|_{\infty,h}\right).
\end{aligned}$$

If we plug this into (4.11) and use Lemma 3.3 and Theorem 3.6 (which is possible since we have $s > n+1$) we find

$$\begin{aligned}
|e_{s1}| &\le \|e\|^2_{2,h} C(f,\rho)\left(\varepsilon + \|\nabla[\rho] - \nabla[\rho^{\varepsilon h}]\|_{\infty,h}\right)\|\nabla\zeta_\varepsilon\|_{1,h} \\
&\le \|e\|^2_{2,h} C(f,\rho)\left(\varepsilon + C\varepsilon^{-1-\frac{n}{2}}\|\rho - \rho^{\varepsilon h}\|_{2,h}\right)\varepsilon^{-1} \\
&\le \|e\|^2_{2,h} C(f,\rho)\left(\varepsilon + C\varepsilon^{-1-\frac{n}{2}}\|e\|_{2,h}\right)\varepsilon^{-1} \\
&= C(f,\rho)\left(1 + C\varepsilon^{-2-\frac{n}{2}}\|e\|_{2,h}\right)\|e\|^2_{2,h}.
\end{aligned} \tag{4.14}$$

Obviously, if condition (4.6) holds, i.e., if we have $\|e\|_{2,h} \le C_1\varepsilon^{2+\frac{n}{2}}$, then it immediately follows from (4.14) that we also have

$$|e_{s1}| \le C(f,\rho)\|e\|^2_{2,h}.$$

Finally, we need to bound the second part (4.8) of the stability error. As before, we will use the mean value theorem and denote the intermediate positions by ξ_i again. This time, we have

$$\begin{aligned}
e_{s2} &= -h^n \sum_{i\in\mathbb{Z}^n} e_i \left(f(t, ih, [\rho]_{ih}, \nabla[\rho^{\varepsilon h}]_{ih}) - f(t, ih, [\rho^{\varepsilon h}]_{ih}, \nabla[\rho^{\varepsilon h}]_{ih})\right) \\
&= -h^n \sum_{i\in\mathbb{Z}^n} e_i Df(t, ih, \cdot, \nabla[\rho^{\varepsilon h}]_{ih})|_{\xi_i}\left([\rho]_{ih} - [\rho^{\varepsilon h}]_{ih}\right) \\
&= -h^n \sum_{i\in\mathbb{Z}^n} e_i Df(t, ih, \cdot, \nabla[\rho^{\varepsilon h}]_{ih})|_{\xi_i}| \left(h^n \sum_{j\in\mathbb{Z}^n} (\rho_j - \rho_j^{\varepsilon h})\zeta_\varepsilon(ih - jh)\right) \\
&= -h^{2n} \sum_{i,j\in\mathbb{Z}^n} e_i e_j Df(t, ih, \cdot, \nabla[\rho^{\varepsilon h}]_{ih})|_{\xi_i}|\zeta_\varepsilon(ih - jh).
\end{aligned}$$

This gives the bound

$$
\begin{aligned}
|e_{s2}| &\le h^{2n}\|f\|_{W^1_\infty(\widetilde{M})} \sum_{i,j\in\mathbb{Z}^n} |e_i e_j||\zeta_\varepsilon(ih-jh)| \\
&\le C(f)h^n \sum_{i\in\mathbb{Z}} e_i^2 \; h^n \sum_{j\in\mathbb{Z}^n} |\zeta_\varepsilon(ih-jh)| \\
&= C(f)\left(h^n \sum_{i\in\mathbb{Z}^n} e_i^2\right)\left(h^n \sum_{j\in\mathbb{Z}^n} \zeta_\varepsilon(jh)\right) \\
&\le C(f)\|e\|_{2,h}^2,
\end{aligned} \tag{4.15}
$$

which, together with (4.14), proves the statement of the theorem. □

Proof of Theorem 2.2. Taking the results of Proposition 4.1 and Theorem 4.2 together and assuming (4.6), we see that

$$\frac{1}{2}\frac{d}{dt}\|e\|_{2,h}^2 \le C(f,\rho)\left(\varepsilon^k + \frac{h^\ell}{\varepsilon^{\ell+1}}\right)^2 + C(f,\rho)\|e\|_{2,h}^2.$$

Hence, applying Gronwall's inequality to this, yields

$$\|e\|_{2,h} \le C(f,\rho,T)\left(\varepsilon^k + \frac{h^\ell}{\varepsilon^{\ell+1}}\right) \tag{4.16}$$

uniformly for all $t \in [0,T]$. With this, we can justify the assumption (4.6) since we have

$$\varepsilon^k + \frac{h^\ell}{\varepsilon^{\ell+1}} \le C\varepsilon^{2+\frac{n}{2}},$$

provided that $k \ge 2 + n/2$ and $h \le \varepsilon^{1+(3+\frac{n}{2})/\ell}$.

Next, we split the $L_\infty(\mathbb{R}^n)$ error as follows:

$$\|\partial^\alpha\rho - \partial^\alpha[\rho^{\varepsilon h}]\|_{L_\infty(\mathbb{R}^n)} \le \|\partial^\alpha\rho - \partial^\alpha[\rho]\|_{L_\infty(\mathbb{R}^n)} + \|\partial^\alpha[\rho] - \partial^\alpha[\rho^{\varepsilon h}]\|_{L_\infty(\mathbb{R}^n)}. \tag{4.17}$$

We can bound the first term on the right-hand side using Theorem 3.4 by

$$\|\partial^\alpha\rho - \partial^\alpha[\rho]\|_{L_\infty(\mathbb{R}^n)} \le C(\rho)\left(\varepsilon^k + \frac{h^\ell}{\varepsilon^{\ell+|\alpha|}}\right)$$

Since this expression will be dominated by the second term in (4.17), we can ignore it. The second term is finally bounded by

$$
\begin{aligned}
\|\partial^\alpha[\rho] - \partial^\alpha[\rho^{\varepsilon h}]\|_{L_\infty(\mathbb{R}^n)} &\le C\varepsilon^{-\frac{n}{2}-|\alpha|}\|\rho - \rho^{\varepsilon h}\|_{2,h} \\
&\le C\varepsilon^{-\frac{n}{2}-|\alpha|}\left(\varepsilon^k + \frac{h^\ell}{\varepsilon^{\ell+1}}\right)
\end{aligned}
$$

using (3.4) from Theorem 3.6 and (4.16). □

5. Construction of high order kernels

A key ingredient of the method is the availability of high order kernels. There are some ways of constructing such kernels, and here, we will follow ideas from [1], see also [21], to construct compactly supported kernels of any prescribed order.

To this end, we will employ radial kernels, i.e., kernels of the form $\zeta(x) = \eta(|x|)$, $x \in \mathbb{R}^n$, with a univariate function $\eta : [0, \infty) \to \mathbb{R}$. For such radial kernels we can easily rewrite the conditions from Definition 2.1:

$$\begin{aligned}\int_{\mathbb{R}^n} x^\alpha \zeta(x)dx &= \int_0^\infty \int_{|x|=1} (xr)^\alpha \zeta(rx) r^{n-1} dS(x)dr \\ &= \int_{|x|=1} x^\alpha dS(x) \int_0^\infty r^{n-1+|\alpha|}\eta(r)dr.\end{aligned}$$

Obviously the first integral over the unit sphere S^{n-1} in $\mathbb{R}^n$ vanishes if $|\alpha|$ is odd. Thus the conditions of Definition 2.1 can be rewritten as follows. The kernel $\zeta = \eta(|\cdot|)$ is of order $k = 2\ell$ if it satisfies the following three conditions:

$$\int_0^\infty \eta(r) r^{n-1} dr = \frac{1}{\omega_{n-1}}, \tag{5.1}$$

$$\int_0^\infty \eta(r) r^{n+2j-1} dr = 0, \qquad 1 \le j \le \ell - 1, \tag{5.2}$$

$$\int_0^\infty \eta(r) r^{n-1+k} dr < \infty, \tag{5.3}$$

where ω_{n-1} denotes the surface area of the unit sphere S^{n-1} in $\mathbb{R}^n$.

We will use this to construct such kernels. To this end assume that we have fixed, pairwise distinct values $a_j > 0$ for $0 \le j \le \ell - 1$ and an even, continuous, non-negative univariate function $\phi : \mathbb{R} \to \mathbb{R}$ with compact support and $\|\phi(|\cdot|)\|_{L_1(\mathbb{R}^n)} = 1$.

Then, we want to pick real numbers $\lambda_0, \dots, \lambda_{\ell-1}$ such that

$$\eta(r) = \sum_{j=0}^{\ell-1} \lambda_j \phi(r/a_j), \tag{5.4}$$

defines a kernel of order $k = 2\ell$.

Since η also has compact support and is continuous, condition (5.3) is automatically satisfied. Conditions (5.1) and (5.2) can be summarised as follows. For $0 \le i \le \ell - 1$ we need that

$$\begin{aligned}\frac{1}{\omega_{n-1}}\delta_{0i} &= \sum_{j=0}^{\ell-1} \lambda_j \int_0^\infty \phi(r/a_j) r^{n-1+2i} dr \\ &= \sum_{j=0}^{\ell-1} \lambda_j a_j^{n+2i} \int_0^\infty \phi(s) s^{n-1+2i} ds.\end{aligned}$$

Noting that the integral on the right-hand side is independent of the summation index j and that the left-hand side is zero except for $i = 0$ we see that we can rewrite this system as

$$\delta_{0i} = \sum_{j=0}^{\ell-1} \widetilde{\lambda}_j a_j^{2i}, \qquad 0 \le i \le \ell - 1,$$

where we have set $\widetilde{\lambda}_j := \lambda_j a_j^n$. This means that the solution vector $\widetilde{\lambda} \in \mathbb{R}^\ell$ is the first column of the inverse of the matrix $A = (a_j^{2i})$, which simply is the transpose of a Vandermonde matrix in $a_0^2, \dots, a_{\ell-1}^2$. This guarantees solvability. To be more precise, we have the following result.

Proposition 5.1. *Let $a_j > 0$, $0 \le j \le \ell - 1$, be pairwise distinct and let $\phi : \mathbb{R} \to \mathbb{R}$ be non-negative, even, continuous, with compact support and satisfying $\|\phi(|\cdot|)\|_{L_1(\mathbb{R}^n)} = 1$. Then, there is exactly one radial kernel $\zeta(x) = \eta(|x|)$, $x \in \mathbb{R}^n$, of order $k = 2\ell$ with η of the form* (5.4)*. The coefficients are given by*

$$\begin{aligned} \lambda_j &= \frac{(-1)^j a_0^2 \cdots a_{j-1}^2 a_{j+1}^2 \cdots a_{\ell-1}^2}{a_j^n (a_j^2 - a_0^2) \cdots (a_j^2 - a_{j-1}^2)(a_{j+1}^2 - a_j^2) \cdots (a_{\ell-1}^2 - a_j^2)} \\ &= \frac{1}{a_j^n} \prod_{\substack{i=0 \\ i \ne j}}^{\ell-1} \frac{a_i^2}{a_i^2 - a_j^2}, \qquad 0 \le j \le \ell - 1. \end{aligned}$$

Proof. Using the notation from the paragraphs above, Cramer's rule shows that $\widetilde{\lambda_j} = \det A_j / \det A$, where A_j is the matrix resulting from A by replacing the j-th column with the first unit vector. Since, A is the transpose of a Vandermonde matrix, its determinant is given by

$$\det(A) = \prod_{0 \le \mu < \nu \le \ell-1} (a_\nu^2 - a_\mu^2).$$

Moreover, in our situation the determinant of A_j is easily determined. Setting $b_j = a_j^2$, we have

$$\begin{aligned} \det(A_j) &= \det \begin{pmatrix} 1 & \dots & 1 & 1 & 1 & \dots & 1 \\ b_0 & \dots & b_{j-1} & 0 & b_{j+1} & \dots & b_{\ell-1} \\ b_0^2 & \dots & b_{j-1}^2 & 0 & b_{j+1}^2 & \dots & b_{\ell-1}^2 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ b_0^{\ell-1} & \dots & b_{j-1}^{\ell-1} & 0 & b_{j+1}^{\ell-1} & \dots & b_{\ell-1}^{\ell-1} \end{pmatrix} \\ &= (-1)^j \det \begin{pmatrix} b_0 & \dots & b_{j-1} & b_{j+1} & \dots & b_{\ell-1} \\ b_0^2 & \dots & b_{j-1}^2 & b_{j+1}^2 & \dots & b_{\ell-1}^2 \\ \vdots & & \vdots & \vdots & & \vdots \\ b_0^{\ell-1} & \dots & b_{j-1}^{\ell-1} & b_{j+1}^{\ell-1} & \dots & b_{\ell-1}^{\ell-1} \end{pmatrix} \\ &= (-1)^j \left(\prod_{\substack{i=0 \\ i \ne j}}^{\ell-1} b_i \right) \det \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 \\ b_0 & \dots & b_{j-1} & b_{j+1} & \dots & b_{\ell-1} \\ b_0^2 & \dots & b_{j-1}^2 & b_{j+1}^2 & \dots & b_{\ell-1}^2 \\ \vdots & & \vdots & \vdots & & \vdots \\ b_0^{\ell-2} & \dots & b_{j-1}^{\ell-2} & b_{j+1}^{\ell-2} & \dots & b_{\ell-1}^{\ell-2} \end{pmatrix} \\ &= (-1)^j \left(\prod_{\substack{i=0 \\ i \ne j}}^{\ell-1} b_i \right) \left(\prod_{\substack{0 \le \mu < \nu \le \ell-1 \\ \mu,\nu \ne j}} (b_\nu - b_\mu) \right) \end{aligned}$$

which gives the stated form. ∎

TABLE 1. Radial functions leading to second order kernels.

$n=1$	$\phi_{1,0}(r) = (1-r)_+$	C^0
	$\phi_{1,1}(r) = \frac{5}{4}(1-r)_+^3(3r+1)$	C^2
	$\phi_{1,2}(r) = \frac{3}{2}(1-r)_+^5(8r^2+5r+1)$	C^4
	$\phi_{1,3}(r) = \frac{55}{32}(1-r)_+^7(21r^3+19r^2+7r+1)$	C^6
$n=2$	$\phi_{2,0}(r) = \frac{6}{\pi}(1-r)_+^2$	C^0
	$\phi_{2,1}(r) = \frac{7}{\pi}(1-r)_+^4(4r+1)$	C^2
	$\phi_{2,2}(r) = \frac{3}{\pi}(1-r)_+^6(35r^2+18r+3)$	C^4
	$\phi_{2,3}(r) = \frac{78}{7\pi}(1-r)_+^8(32r^3+25r^2+8r+1)$	C^6
$n=3$	$\phi_{3,0}(r) = \frac{15}{2\pi}(1-r)_+^2$	C^0
	$\phi_{3,1}(r) = \frac{21}{2\pi}(1-r)_+^4(4r+1)$	C^2
	$\phi_{3,2}(r) = \frac{165}{32\pi}(1-r)_+^6(35r^2+18r+3)$	C^4
	$\phi_{3,3}(r) = \frac{1365}{64\pi}(1-r)_+^8(32r^3+25r^2+8r+1)$	C^6

TABLE 2. Possible coefficients and weights for fourth and sixth order kernels in $\mathbb{R}^n$.

Order	a_0	a_1	a_2	λ_0	λ_1	λ_2
4	1.0	4/5		-16/9	$(5/4)^n \cdot (25/9)$	
6	1.0	4/5	3/5	1.0	$-(5/4)^n \cdot (25/7)$	$(5/3)^n \cdot 25/7$

This simple proof also follows from more general results on Vandermonde matrices; see for example [3, 14] and the references therein.

Here, we propose to use the following compactly supported kernels. We start with radial kernels $\phi(r) = \phi_{n,\ell}(r) = c_{n,\ell}(1-r)_+^{d(n,\ell)} p_{n,\ell}(r)$ from [36, 37]; see also Table 1. These kernels are known to have smoothness $C^{2\ell}(\mathbb{R}^n)$ and are as non-negative, radial kernels of order 2. The cases $n = 2$ and $n = 3$ only differ in the constant $c_{n,l}$, which has to be chosen such that the kernels satisfy (5.1).

For higher order kernels we employ the construction from Proposition 5.1. Here, we are free to choose the parameters a_j. It might be interesting to discuss the optimal choice of these parameters with respect to, for example, stability of the evaluation of the kernels. Here, however, we made the choice given in Table 2, which also contains the corresponding weights λ_j.

Note that these coefficients can be used for *any* radial kernel with compact support. In the following, we will denote a kernel of smoothness s and order k by $\eta = \eta^{k,s}$, i.e., for $\zeta^{k,s}(\cdot) = \eta^{k,s}(|\cdot|)$ we have $\zeta^{k,s} \in \mathcal{K}^{k,s}$. Hence, typical examples are

$$\begin{aligned}
\eta^{2,2}(r) &= \phi_{n,1}(r), \\
\eta^{4,4}(r) &= -\frac{16}{9}\phi_{n,2}(r) + \left(\frac{5}{4}\right)^n \frac{25}{9}\phi_{n,2}(5r/4), \\
\eta^{4,6}(r) &= -\frac{16}{9}\phi_{n,3}(r) + \left(\frac{5}{4}\right)^n \frac{25}{9}\phi_{n,3}(5r/4), \\
\eta^{6,6}(r) &= \phi_{n,3}(r) - \left(\frac{5}{4}\right)^n \frac{25}{7}\phi_{n,3}(5r/4) + \left(\frac{5}{3}\right)^n \frac{25}{7}\phi_{3,3}(5r/3).
\end{aligned}$$

For $n = 1$, some of these kernels are also depicted in Figure 1.

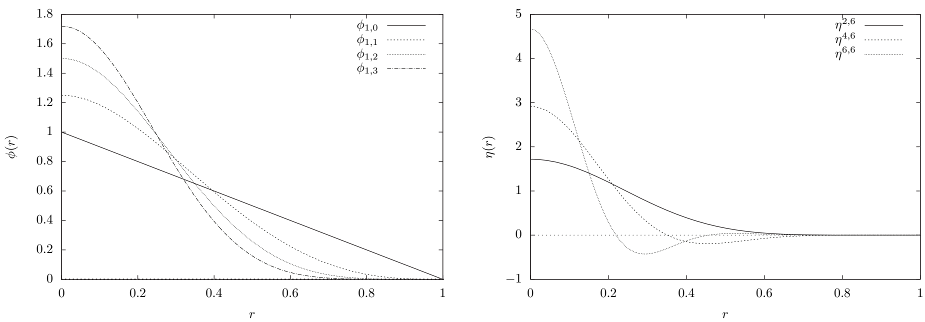

FIGURE 1. Some of the kernels for approximation in one dimension. The picture on the left shows the classical Wendland kernels $\phi_{1,\ell}$ for $\ell = 0, 1, 2, 3$. The picture on the right shows higher order kernels $\eta^{2k,6} \in C^6(\mathbb{R})$, $k = 1, 2, 3$, built from $\phi_{1,3}$.

6. Numerical example

We are now going to test our method by looking at a numerical example in one space dimension. To be more precise, we consider the following Burgers equation. For a given $T > 0$ and $\psi \in C^r(\mathbb{R})$ with compact support, we are looking for the solution ρ of

$$\begin{aligned}
\partial_t\rho - \rho\partial_x\rho &= [(1-\psi)\,\psi'](x+t), && \text{on } (0,T]\times\mathbb{R}, \\
\rho|_{t=0} &= \psi, && \text{on } \mathbb{R}.
\end{aligned}$$

Thus, in our initial setting, the defining function f is given by

$$f(t,x,\rho,\partial_x\rho) = \rho(t,x)\partial_x\rho(t,x) + [(1-\psi)\,\psi']\,(x+t), \qquad (t,x) \in [0,T]\times\mathbb{R}.$$

It is easy to see that $\rho(t,x) = \psi(x+t)$ solves the problem above. Hence, we can control the smoothness of the solution. Moreover, if we pick the initial data ψ with compact support in the one-dimensional *ball* $[-\delta,\delta]$, then the support of $\rho(t,\cdot)$ is in $[-t-\delta, -t+\delta]$ and hence, for a fixed $T > 0$, the support of $\rho(t,\cdot)$, $t \in [0,T]$, is a subset of the interval $\Omega = [-\delta - T, \delta]$. Thus, the assumption on the solution from Theorem 2.2 are easily satisfied and we can try to verify the convergence orders claimed therein. To this end, we carried out the following two test series.

(A) In the first series we have chosen the kernel and the solution such that Theorem 2.2 is applicable. To be more precise, we used $\zeta = \eta^{4,4}(|\cdot|) \in \mathcal{K}^{4,4}$ as the kernel for our scheme. The kernel has been built as described in the last section using the underlying function $\phi_{1,2}$ and the coefficients given in Table 2.

The initial data was given by $\psi = \delta^{-1}\tilde{\eta}^{4,4}(\cdot/\delta)$ with $\delta = 0.5$. To avoid any unwanted, positive side effects, which might result from choosing the same function as the approximation kernel and the initial data, we have used $\phi_{2,2}$ as the underlying kernel to built ψ. Of course, we have chosen $c_{2,2}$ as $\frac{9}{16}$ for the kernel so that $\|\phi_{2,2}\|_{L_1(\mathbb{R})} = 1$ is satisfied.

TABLE 3. Discrete L_∞ errors for series (A) for various discretisation parameters $h = 2^{\nu_h}$ and $\varepsilon = 2^{\nu_\varepsilon}$.

	ν_ε							
ν_h	-6	-7	-8	-9	-10	-11	-12	-13
-8								
-9	4.83e-3							
-10	1.47e-3	7.09e-3						
-11	1.03e-4	2.27e-3	1.05e-2					
-12	1.18e-4	3.39e-5	3.31e-3	1.62e-2				
-13	1.18e-4	1.71e-5	6.72e-5	5.28e-3	2.47e-2			
-14	1.18e-4	1.55e-5	3.90e-6	1.11e-4	8.53e-3	3.48e-2		
-15	1.18e-4	1.54e-5	1.32e-6	4.60e-6	1.86e-4	1.33e-2	4.26e-2	
-16		1.54e-5	1.27e-6	2.29e-7	7.52e-6	3.19e-4	1.89e-2	4.67e-2
-17			1.27e-6	1.49e-7	1.47e-7	1.29e-5	5.42e-4	2.33e-2
-18				1.53e-7	1.57e-8	2.27e-7	2.21e-5	9.09e-4

TABLE 4. Discrete L_∞ errors for series (B) for various discretisation parameters $h = 2^{\nu_h}$ and $\varepsilon = 2^{\nu_\varepsilon}$.

	ν_ε							
ν_h	-6	-7	-8	-9	-10	-11	-12	-13
-8								
-9	3.12e-2							
-10	2.71e-2	1.92e-2						
-11	2.69e-2	1.15e-2	1.94e-2					
-12	2.68e-2	1.11e-2	5.24e-3	2.44e-2				
-13	2.68e-2	1.11e-2	4.54e-3	3.17e-3	3.20e-2			
-14	2.68e-2	1.11e-2	4.50e-3	1.89e-3	2.86e-3	4.03e-2		
-15	2.68e-2	1.11e-2	4.50e-3	1.82e-3	8.48e-4	3.27e-3	4.70e-2	
-16		1.11e-2	4.50e-3	1.81e-3	7.34e-4	4.70e-4	3.96e-3	5.14e-2
-17			4.50e-3	1.81e-3	7.27e-4	3.00e-4	3.47e-4	4.62e-3
-18				1.81e-3	7.27e-4	2.90e-4	1.23e-4	3.26e-4

(B) The purpose of the second series was to investigate whether our scheme also shows convergence in situations with less regularity and order of the kernel than required by our main result.

To this end, we used $\zeta = \eta^{2,2}(|\cdot|) = \phi_{1,1}(|\cdot|) \in \mathcal{K}^{2,2}$ as the kernel and $\psi = \delta^{-1}\phi_{2,1}(\cdot/\delta)$ as the initial data with $c_{2,1} = \frac{3}{2}$ to achieve $\|\phi_{2,1}\|_{L_1(\mathbb{R})} = 1$ again.

In both cases we have chosen $\delta = T = 0.5$, meaning $\Omega = [-1.0, 0.5]$. Our computation was then restricted to those data sites in $\Omega_1 = [-1.25, 0.75]$, which, for computational reasons, we have chosen slightly smaller than the one we have used in Theorem 2.2.

The computations have been done for various values of ε and h, both of the form $2^{-\nu}$. For the time discretisation we have chosen an explicit Runge-Kutta method of order 4. The numerical results indicated that there is, as expected, a CFL condition. The theoretical analysis of the time discretisation will be the subject of a subsequent paper. Here, we simply have chosen the time discretisation sufficiently small such that its error was negligible.

The results can be found in Tables 3 and 4 for series (A) and (B), respectively. The conditions of Theorem 2.2 require $h \le \varepsilon^{1+3.5/\ell}$, which is why the tables only contain entries for $h < \varepsilon$.

In the context of particle methods it is an often encountered assumption that convergence is achieved in the so-called *stationary setting*, i.e., if the ratio between h and ε is fixed, meaning that approximately the same number of data sites lies in the support of each kernel. This assumption is in particular made in almost all application papers, though it is well-known by now (see for example [26]) that this is not true. Our results corroborate this since we can see divergence in this situation by looking at the diagonal entries of the Tables 3 and 4.

As for the general dependence of the errors on the discretisation parameters h and ε, the data seem to verify our findings in the following way: If we look at a fixed column of either table, which corresponds to a fixed ε, we see that the error becomes stationary, i.e., further refinement of the grid does not lead to convergence. If we look at the rows of the table, which corresponds to an ε-refinement while h is kept fixed, we see that the error eventually grows.

Finally, we have tried to estimate the crucial constants and exponents in the error estimate using a least-squares approach.

(A) Here, Theorem 2.2 can be applied with parameters $r, s = 4$, $k, \ell = 3$. In this situation, Theorem 2.2 yields an error bound of the form

$$C_1\varepsilon^{k-0.5} + C_2\frac{h^\ell}{\varepsilon^{\ell+1.5}} = C_1\varepsilon^{2.5} + C_2\frac{h^3}{\varepsilon^{4.5}} \tag{6.1}$$

for the approximate solution. The least-squares approximation yields the better estimate

$$68.7\varepsilon^{3.2} + 1.8\frac{h^{3.8}}{\varepsilon^{4.4}}.$$

(B) Our main result does not apply in this situation. Nonetheless, the data suggest that we still have convergence. The rows of Table 4 indicate that for a fixed, small h the approximation for ε to zero converges super-linear. A least-squares approach yields the estimate $k - 0.5 = 1.32$. To estimate the exponents in the second term of (6.1) is not feasible because there are too few values for a reasonable least-squares approximation.

In both cases, the exponents for the first term are not too far off from the value $(k - \frac{1}{2})$ that we can expect from Lemma 2.1 and Theorem 3.6 when ignoring the requirements on the regularity of the solution and the kernel. This could be because the kernel and the solution are smooth except for a finite number of points, so that errors that stem from the reduced regularity at these points are comparatively small. The same might be the reason why the parameters of the second term turn out to be better than expected.

In any case, this already indicates that the scheme should also be useful when approximating solutions with lower regularity. This would be particularly valuable in higher spatial dimensions where only very few good methods are available to the present day. Some simple experiments with shock formation seem to be promising and a detailed investigation is planned for the future.

References

[1] J. T. Beale and A. Majda, *High order accurate vortex methods with explicit velocity kernels*, J. Comput. Phys. **58** (1985), 188–208.

[2] R. K. Beatson and M. J. D. Powell, *Univariate multiquadric approximation: quasi-interpolation to scattered data*, Constr. Approx. **8** (1992), no. 3, 275–288, DOI 10.1007/BF01279020. MR1164070

[3] A. Björck and V. Pereyra, *Solution of Vandermonde systems of equations*, Math. Comp. **24** (1970), 893–903, DOI 10.2307/2004623. MR0290541

[4] A. Bressan, *Hyperbolic conservation laws: an illustrated tutorial*, Modelling and Optimisation of Flows on Networks, Lecture Notes in Math., vol. 2062, Springer, Heidelberg, 2013, pp. 157–245, DOI 10.1007/978-3-642-32160-3_2. MR3050281

[5] M. D. Buhmann, *Convergence of univariate quasi-interpolation using multiquadrics*, IMA J. Numer. Anal. **8** (1988), no. 3, 365–383, DOI 10.1093/imanum/8.3.365. MR968100

[6] M. D. Buhmann, *On quasi-interpolation with radial basis functions*, J. Approx. Theory **72** (1993), no. 1, 103–130, DOI 10.1006/jath.1993.1009. MR1198376

[7] M. D. Buhmann, *Radial basis functions: theory and implementations*, Cambridge Monographs on Applied and Computational Mathematics, vol. 12, Cambridge University Press, Cambridge, 2003. MR1997878

[8] M. D. Buhmann, N. Dyn, and D. Levin, *On quasi-interpolation by radial basis functions with scattered centres*, Constr. Approx. **11** (1995), no. 2, 239–254, DOI 10.1007/BF01203417. MR1342386

[9] M. D. Buhmann and F. Dai, *Pointwise approximation with quasi-interpolation by radial basis functions*, J. Approx. Theory **192** (2015), 156–192, DOI 10.1016/j.jat.2014.11.005. MR3313479

[10] R. Cortez, *Convergence of high-order deterministic particle methods for the convection-diffusion equation*, Comm. Pure Appl. Math. **50** (1997), no. 12, 1235–1260, DOI 10.1002/(SICI)1097-0312(199712)50:12⟨1235::AID-CPA2⟩3.3.CO;2-Z. MR1476314

[11] C. de Boor, R. A. DeVore, and A. Ron, *Approximation from shift-invariant subspaces of $L_2(\mathbf{R}^d)$*, Trans. Amer. Math. Soc. **341** (1994), no. 2, 787–806, DOI 10.2307/2154583. MR1195508

[12] C. de Boor and A. Ron, *Fourier analysis of the approximation power of principal shift-invariant spaces*, Constr. Approx. **8** (1992), no. 4, 427–462, DOI 10.1007/BF01203462. MR1194028

[13] T. A. Driscoll and B. Fornberg, *Interpolation in the limit of increasingly flat radial basis functions*, Comput. Math. Appl. **43** (2002), no. 3-5, 413–422, DOI 10.1016/S0898-1221(01)00295-4. Radial basis functions and partial differential equations. MR1883576

[14] A. Eisinberg and G. Fedele, *On the inversion of the Vandermonde matrix*, Appl. Math. Comput. **174** (2006), no. 2, 1384–1397, DOI 10.1016/j.amc.2005.06.014. MR2220623

[15] T. A. Eymann and P. L. Roe, *Multidimensional active flux schemes*, 21st AIAA Computational Fluid Dynamics Conference, 2013.

[16] M. Falcone and R. Ferretti, *Convergence analysis for a class of high-order semi-Lagrangian advection schemes*, SIAM J. Numer. Anal. **35** (1998), no. 3, 909–940, DOI 10.1137/S0036142994273513. MR1619910

[17] M. Falcone and R. Ferretti, *Semi-Lagrangian approximation schemes for linear and Hamilton-Jacobi equations*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2014. MR3341715

[18] P. Giesl and S. Hafstein, *Review on computational methods for Lyapunov functions*, Discrete Contin. Dyn. Syst. Ser. B **20** (2015), no. 8, 2291–2331, DOI 10.3934/dcdsb.2015.20.2291. MR3423237

[19] M. Kindelan, M. Moscoso, and P. González-Rodríguez, *Radial basis function interpolation in the limit of increasingly flat basis functions*, J. Comput. Phys. **307** (2016), 225–242, DOI 10.1016/j.jcp.2015.12.015. MR3448206

[20] R. J. LeVeque, *Numerical methods for conservation laws*, 2nd ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1992. MR1153252

[21] A. J. Majda and A. L. Bertozzi, *Vorticity and Incompressible Flow*, Cambridge Texts in Applied Mathematics, vol. 27, Cambridge University Press, Cambridge, 2002. MR1867882

[22] J. J. Monaghan, *Smoothed particle hydrodynamics*, Rep. Progr. Phys. **68** (2005), no. 8, 1703–1759, DOI 10.1088/0034-4885/68/8/R01. MR2158506

[23] J. J. Monaghan, *Smoothed Particle Hydrodynamics and Its Diverse Applications*, Annu. Rev. Fluid Mech., vol. 44, Annual Reviews, Palo Alto, CA, 2012, pp. 323–346, DOI 10.1146/annurev-fluid-120710-101220. MR2882600

[24] B. Ben Moussa, *On the convergence of SPH method for scalar conservation laws with boundary conditions*, Methods Appl. Anal. **13** (2006), no. 1, 29–61, DOI 10.4310/MAA.2006.v13.n1.a3. MR2275871

[25] B. Ben Moussa and J. P. Vila, *Convergence of SPH method for scalar nonlinear conservation laws*, SIAM J. Numer. Anal. **37** (2000), no. 3, 863–887, DOI 10.1137/S0036142996307119. MR1740385

[26] P.-A. Raviart, *An analysis of particle methods*, Numerical methods in fluid dynamics (Como, 1983), Lecture Notes in Math., vol. 1127, Springer, Berlin, 1985, pp. 243–324, DOI 10.1007/BFb0074532. MR802214

[27] A. Ron, *The L_2-approximation orders of principal shift-invariant spaces generated by a radial basis function*, Numerical Methods in Approximation Theory. Vol. 9: Proceedings of the conference held in Oberwolfach, Germany, November 24-30, 1991 (Basel) (D. Braess et al., eds.), Int. Ser. Numer. Math., vol. 105, Birkhäuser, 1992, pp. 245–268.

[28] R. Schaback and Z. Wu, *Construction techniques for highly accurate quasi-interpolation operators*, J. Approx. Theory **91** (1997), no. 3, 320–331, DOI 10.1006/jath.1996.3101. MR1486471

[29] R. Schaback, *Limit problems for interpolation by analytic radial basis functions*, J. Comput. Appl. Math. **212** (2008), no. 2, 127–149, DOI 10.1016/j.cam.2006.11.023. MR2383010

[30] C.-W. Shu, *High order weighted essentially nonoscillatory schemes for convection dominated problems*, SIAM Rev. **51** (2009), no. 1, 82–126, DOI 10.1137/070679065. MR2481112

[31] Chi-Wang Shu, *A brief survey on discontinuous galerkin methods in computational fluid dynamics*, Advances in Mechanics **43** (2013), 541–554.

[32] G. Song, J. Riddle, G. E. Fasshauer, and F. J. Hickernell, *Multivariate interpolation with increasingly flat radial basis functions of finite smoothness*, Adv. Comput. Math. **36** (2012), no. 3, 485–501, DOI 10.1007/s10444-011-9192-5. MR2893476

[33] G. Strang and G. Fix, *A Fourier analysis of the finite element variational method*, C.I.M.E. II, Ciclo 1971, Constructive Aspects of Functional Analysis (G. Geymonat, ed.), 1973, pp. 793–840.

[34] J. P. Vila, *On particle weighted methods and smooth particle hydrodynamics*, Math. Models Methods Appl. Sci. **9** (1999), no. 2, 161–209, DOI 10.1142/S0218202599000117. MR1674540

[35] D. Violeau, *Fluid Mechanics and the SPH Method: Theory and Applications*, Oxford University Press, Oxford, 2012. MR3222339

[36] H. Wendland, *Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree*, Adv. Comput. Math. **4** (1995), no. 4, 389–396, DOI 10.1007/BF02123482. MR1366510

[37] H. Wendland, *Scattered Data Approximation*, Cambridge Monographs on Applied and Computational Mathematics, vol. 17, Cambridge University Press, Cambridge, 2005. MR2131724

Department of Mathematics, University of Bayreuth, D-95440 Bayreuth, Germany
E-mail address: tobias.ramming@uni-bayreuth.de

Department of Mathematics, University of Bayreuth, D-95440 Bayreuth, Germany
E-mail address: holger.wendland@uni-bayreuth.de