

CONTINUOUS DATA ASSIMILATION FOR HYDRODYNAMICS: CONSISTENT DISCRETIZATION AND APPLICATION TO MOMENT RECOVERY

JINGCHENG LU, KUNLUN QI, LI WANG, AND JEFF CALDER

ABSTRACT. Motivated by the challenge of moment recovery in hydrodynamic approximation in kinetic theory, we propose a data-driven approach for the hydrodynamic models. Inspired by continuous data assimilation, our method introduces a relaxation-based nudging system coupled with a novel discretization technique. This approach facilitates the simultaneous recovery of both the force term and a high-resolution solution from sparsely observed data. To address potential numerical artifacts, we use kernel regression to fit the observed data. We also analyze the convergence of the proposed nudging system under both full and partial data scenarios. When applied to moment systems, the source term involves the derivative of higher-order moments, and our approach serves as a crucial step for data preparation in machine-learning based moment closure models. Multiple numerical experiments demonstrate the effectiveness of our algorithm, and we discuss its potential extension to high-dimensional systems.

CONTENTS

1. Introduction	2
2. Our approach: convergence analysis and numerical discretization	5
2.1. Convergence analysis	5
2.2. Discretization of nudged system	11
2.3. Smooth fitting with kernel regression	11
3. Numerical experiments I – states and force recovery	14
3.1. Scalar test case	14
3.2. Nonlinear system	16
3.3. Multi-dimensional system	17
3.4. Noisy measurement and sparse observation	19
4. Moment recovery with incomplete observations	22
4.1. Problem set up	23
4.2. Moment recovery via continuous data assimilation	25
4.3. Numerical experiments II – moment recovery	27
5. Extension to higher dimension	32

2020 *Mathematics Subject Classification.* Primary 93C20, 70S10, 65Dxx, 42A15, 82C40; Secondary 65Zxx, 35Q30.

Key words and phrases. Data assimilation, Hydrodynamic models, Moment recovery, Relaxation-based nudging system.

JL and JC were supported by NSF-CCF:2212318, and JC was additionally supported by an Albert and Dorothy Marden Professorship and a McKnight Presidential Fellowship. LW is partially supported by NSF grant DMS-1846854 and UMN DSI-SSG-4886888864. KQ is partially supported by direct funding from School of Mathematics at the University of Minnesota–Twin Cities.

1. INTRODUCTION

When modeling complicated physics processes, an accurate description can be provided at the mesoscopic scale using kinetic-type governing equations:

$$(1.1) \quad \frac{\partial}{\partial t} f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + F \cdot \nabla_{\mathbf{v}} f = \mathcal{Q}(f, f), \quad (\mathbf{x}, \mathbf{v}, t) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}_+.$$

Here $f(\mathbf{x}, \mathbf{v}, t)$ is the distribution function that describes particles within a physical system, where \mathbf{x} is the position, \mathbf{v} is the velocity and t is the time. F is the external force, which, for example, can take the form $F = q(\mathbb{E} + \mathbf{v} \times \mathbb{B})$ when modeling charged particles in plasma. Here q is the charge per unit mass of the particles, while \mathbb{E} and \mathbb{B} denote the strengths of the electric and magnetic fields, respectively. $\mathcal{Q}(f, f)$ is the collision operator that encapsulates the interactions between particles.

While significant progress has been made in developing computational methods for solving (1.1), such as finite difference or finite element methods, in industrial applications like fusion energy simulations, these grid-based methods are still cursed by dimensionality. To reduce the computational cost while maintaining important physical features, an attractive approach is to simulate the dynamics at the hydrodynamic scale, which is the primary motivation behind the moment method. To explain, by taking the zeroth and the first \mathbf{v} -moments of the equation (1.1), one obtains the following hydrodynamic system:

$$(1.2) \quad \begin{cases} \frac{\partial}{\partial t} \rho + \nabla \cdot (\rho \mathbf{u}) = 0, \\ \frac{\partial}{\partial t} \rho \mathbf{u} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) = -\nabla \cdot \mathbb{P} + \rho q(\mathbb{E} + \mathbf{u} \times \mathbb{B}), \end{cases}$$

where the density ρ , the momentum $\rho \mathbf{u}$, and the pressure tensor \mathbb{P} are given by

$$\begin{aligned} \rho(\mathbf{x}, t) &:= \int f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v}, \\ \rho \mathbf{u}(\mathbf{x}, t) &:= \int \mathbf{v} f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v}, \\ \mathbb{P}(\mathbf{x}, t) &:= \int (\mathbf{v} - \mathbf{u}(\mathbf{x}, t)) \otimes (\mathbf{v} - \mathbf{u}(\mathbf{x}, t)) f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v}. \end{aligned}$$

Clearly, the hydrodynamic system (1.2) resides in a 4-dimensional space: 3 for \mathbf{x} and 1 for t , making it more computationally friendly than the original 7-dimensional kinetic equation (1.1). However, this reduction comes at a price—the system is not closed. Specifically, \mathbb{P} depends on the second moments of f , which cannot be written solely as a function of the zero-th and first moment of f , i.e., ρ and $\rho \mathbf{u}$. One way to close this system is to assume f is near the Maxwellian, i.e., $f \propto \rho \exp -\frac{|\mathbf{v} - \mathbf{u}|^2}{2T}$. However, this approximation holds true only in certain regimes where collisional effects dominate [5]. In general, one can derive a larger moment system that evolves more moments, aiming to approximate the original kinetic equation more closely, but the closure problem persists [8, 9].

In this paper, we aim to address this problem with the help of observational data. In particular, we reformulate this closure problem into a force-identification problem:

$$(1.3) \quad \begin{cases} \frac{\partial}{\partial t} U(\mathbf{x}, t) + \nabla \cdot F(U(\mathbf{x}, t)) = G(\mathbf{x}, t) + S(U, \mathbf{x}, t), \\ U(\mathbf{x}, 0) = U_0(\mathbf{x}), \end{cases} \quad \mathbf{x} \in \Omega \subseteq \mathbb{R}^d, \quad t \geq 0.$$

Here $U(\mathbf{x}, t) = (U^0(\mathbf{x}, t), U^1(\mathbf{x}, t), \dots, U^N(\mathbf{x}, t))^T \in \mathbb{R}^N$ are the solution states. The flux, $F(U)$, and the source term, $S(U, \mathbf{x}, t)$, are assumed to be known from the physics, whereas the forcing term $G(\mathbf{x}, t)$ is to be determined. In the context of moment recovery, the force function takes the divergence form $G = -\nabla \cdot \mathbb{M}$, where the tensor $\mathbb{M} \in \mathbb{R}^{d \times N}$ contains the information of the moment one order higher than U^N . Indeed, the system (1.2) can be written in the form of (1.3) with

$$U = \begin{bmatrix} \rho \\ \rho \mathbf{u} \end{bmatrix}, \quad F = \begin{bmatrix} \rho \mathbf{u} \\ \rho \mathbf{u} \otimes \mathbf{u} \end{bmatrix}, \quad G = -\nabla \cdot \begin{bmatrix} \mathbf{0} \\ \mathbb{P} \end{bmatrix}, \quad S = \begin{bmatrix} \mathbf{0} \\ \rho q(\mathbb{E} + \mathbf{u} \otimes \mathbb{B}) \end{bmatrix}.$$

If the solution state $U(\mathbf{x}, t)$ and its gradients are known continuously in space and time, then the forcing term can be computed with $G = \frac{\partial U}{\partial t} + \nabla \cdot F - S$. However, in many real-world applications such as weather prediction, one may only have access to a limited amount of true data at sparse and/or non-uniform observation points, \mathbf{x}_j^{ob} , $j = 1, 2, \dots, N_{ob}$, which may not be sufficient to recover the solution gradients with the desired accuracy. Additionally, due to the limited observations, the governing equation (1.3) may not be initialized accurately. This limitation can sometimes lead to a severe accumulation of errors when making future predictions. All of these challenges underscore the pressing task:

Given only sparse observations $\{U(\mathbf{x}_j^{ob}, t)\}$, can we simultaneously recover G and obtain an enhanced resolution of the solution $U(\mathbf{x}, t)$ in (1.3)?

To fulfill this task, we propose the following nudged system:

$$(1.4a) \quad \begin{cases} \frac{\partial}{\partial t} V(\mathbf{x}, t) + \nabla \cdot F(V(\mathbf{x}, t)) = \tilde{G}(\mathbf{x}, t) + S(V, \mathbf{x}, t) + \mu(I_h(U) - I_h(V)), \\ V(\mathbf{x}, 0) = V_0(\mathbf{x}). \end{cases}$$

Here $I_h(\cdot)$ is the *interpolation operator*, which is a bounded linear operator that interpolates the observed data into a smooth function, i.e., $I_h : \{(\mathbf{x}_j^{ob}, U(\mathbf{x}_j^{ob}))\} \mapsto I_h(U)(\mathbf{x}) \in C^\infty(\Omega)$. h represents the mesh size of observation grids $\{\mathbf{x}_j^{ob}\}$. The approximate force \tilde{G} is constructed with

$$(1.4b) \quad \tilde{G} = \frac{\partial}{\partial t} I_h(U) + I_h(\nabla \cdot F(\tilde{U}) - S(\tilde{U})), \quad \tilde{U} = I_h(U) + (I - I_h)(V).$$

In the particular case where I_h is a projection, we have $I_h^2 = I_h$. Note that $\partial_t I_h = I_h \partial_t$ and $I_h(\tilde{U}) = I_h(U)$, (1.4b) can also be written as

$$\tilde{G} = I_h\left(\frac{\partial}{\partial t} \tilde{U} + \nabla \cdot F(\tilde{U}) - S(\tilde{U})\right).$$

Compared with the exact dynamics (1.3), the nudged equation (1.4a) includes an additional relaxation term, $\mu(I_h(U) - I_h(V))$, which serves as a *feedback control*

from the observations $I_h(U)$. The user-tuning parameter $\mu > 0$ is known as the *nudging coefficient*. With an appropriate choice of I_h , the approximate solutions V and \tilde{G} are expected to be time-asymptotically accurate, namely, $V \xrightarrow{t \rightarrow \infty} U$ and $\tilde{G} \xrightarrow{t \rightarrow \infty} G$ in certain norms.

We note that the concept of introducing feedback control into the system (1.4a) is inspired by the continuous data assimilation methods discussed in [34, 16], which aim to recover the unknown body force in the incompressible Navier-Stokes equations from low-mode observations. These methods trace back to the earlier work by Azouani, Olson and Titi [3, 19], which focuses on velocity field recovery for the two-dimensional Navier-Stokes equations. Extension to the continuous data assimilation for dynamics with a set of unidentified parameters can be found in [13, 10, 11, 33, 36]. Investigations have also been carried out for state recovery in various of hydrodynamic settings, including Rayleigh-Benard convection [15, 17, 2, 14], geophysical fluids [27, 1, 37, 26], and dispersive equations [28, 29].

A major advantage of the Azouani-Olson-Titi nudging framework is that it allows simultaneous reconstruction of both the solution state U and the unknown forcing term G . This capability is particularly important in our context, as G encapsulates the information of truncated higher-order moments, which are the critical components for accurate system recovery. In comparison, other conventional data assimilation methods, such as the Kalman filter [30, 40], particle filter [31, 12], and variational assimilation [32, 41, 18], typically assume fully known model dynamics and are thus primarily designed for state estimation. These methods do not inherently support the recovery of unknown or partially known dynamics, and it remains open to explore if they can enable more complicated data recovery tasks with affordable modifications.

To ensure the convergence of (1.4), we impose the following conditions on I_h :

$$(1.5a) \quad \|I_h(\partial^\alpha \phi)\|_2 \leq c_m h^{-m} \|\phi\|_2,$$

$$(1.5b) \quad \|\phi - I_h(\phi)\|_2 \leq C_m h^m \|\phi\|_{H^m},$$

for any multi-index $|\alpha| = m$ with $m \geq 0$. The constants $c_m, C_m > 0$ are independent of h . For example, the spectral projection onto Fourier modes with wave numbers $|k| \leq 1/h$ would satisfy the above conditions. In this special case, I_h serves as a low-pass filter that removes the high-frequency fluctuations, and the intermediate solution \tilde{U} is obtained with the low-mode observations, $I_h(U)$, corrected by the high-mode discrepancy, $(I - I_h)(V)$.

Remark 1.1. In the particular case where the function ϕ has weak oscillations such that

$$\sup_{t \geq 0} \frac{\|\nabla \phi\|_2}{\|\phi\|_2} \leq C < \infty,$$

the error bound (1.5a) can be modified to

$$(1.6) \quad \|I_h(\nabla \phi)\|_2 \leq c_0 \|\nabla \phi\|_2 \leq c_0 C \|\phi\|_2.$$

The h^{-1} dependence in the error bound for $m = 1$ is then removed.

Upon the recovery of (the divergence of) the highest moment, an approximate closure relation may be obtained with machine learning techniques. In particular,

consider the moment systems of (1.1):

$$(1.7) \quad \begin{aligned} \frac{\partial \mathbf{m}_N}{\partial t} + A_N \frac{\partial \mathbf{m}_N}{\partial x} &= \mathbf{g}_{N+1} + S_N \mathbf{m}_N, \\ \mathbf{g}_{N+1} &= -\frac{N+1}{2N+1} \partial_x m_{N+1} \mathbf{e}_{N+1}, \end{aligned}$$

where $\mathbf{m}_N = (m_0, m_1, \dots, m_N)^\top$, $\mathbf{e}_{N+1} = (0, \dots, 0, 1)^\top$ and

$$m_k(x, t) = \frac{1}{2} \int_{-1}^1 f(x, v, t) P_k(v) dv,$$

with P_k being the k -th order polynomial. A_N and S_N are defined in (4.5) and (4.6). Given observations of the first few moments, e.g., m_1, \dots, m_l , the proposed method allows for the recovery of the higher moments m_{l+1}, \dots, m_{N+1} . By using machine learning techniques, such as neural networks, to approximate the relationship between the gradient of the highest moments and those of the lower-order moments, one can close the system described by (1.7). For a simple kinetic equation, such a relation has been learned in a series of works by Huang et. al. [23, 25, 24] by running an expensive kinetic simulation to generate data (i.e., f), computing its moments, and then training the neural network on these moments. In contrast, our proposed method eliminates the need to compute the full kinetic system, instead providing a cost-efficient mechanism for generating data for the moment closure task.

The rest of the paper is organized as follows. In Section 2, we first introduce the main contents of our proposed method, including the convergence analysis of the nudge system, a novel numerical discretization and a kernel regression approach for interpolation. Extensive numerical experiments are presented for the validation of the effectiveness of our algorithm in Section 3. Furthermore, we apply the proposed approach to the moment system in kinetic theory, aiming to recover higher-order moments with incomplete observations in Section 4. Finally, in Section 5, we discuss the potential strategies and challenges in extending the method to higher-dimensional systems.

2. OUR APPROACH: CONVERGENCE ANALYSIS AND NUMERICAL DISCRETIZATION

This section presents the main approach of this paper, which centers around (1.4a). In the next subsection, we first analyze the convergence behavior of (1.4a) under both fully observed data (i.e., when (2.3) is satisfied) and partial data scenarios (see Proposition 2.3). Subsection 2.2 will then propose an efficient discretization of (1.4a). The choice of the interpolation operator I_h is then discussed in section 2.3.

2.1. Convergence analysis.

Proposition 2.1. (Convergence of the state) Let (U, G) be a strong solution of the equation

$$(2.1) \quad \partial_t U + \partial_x U = G + S(U, x, t),$$

associated with Lipschitz continuous source term,

$$|S(U, x, t) - S(V, x, t)| \leq L_s |U - V|, \quad L_s > 0.$$

Let (V, \tilde{G}) be a strong solution of the nudged system

$$(2.2) \quad \begin{cases} \partial_t V + \partial_x V = \tilde{G} + S(V) + \mu(I_h(U) - I_h(V)), \\ \tilde{G} = \partial_t I_h(U) + I_h(\partial_x \tilde{U} - S(\tilde{U})), \quad \tilde{U} = I_h(U) + (I - I_h)(V), \end{cases}$$

where I_h satisfies conditions (1.5). Assume that the data are sufficiently observed, meaning that

$$(2.3a) \quad \|W(t) - I_h(W(t))\|_2 \leq C_h e^{-\varepsilon t}, \quad W(t) = U(t) - V(t),$$

$$(2.3b) \quad I_h(G(t)) = G(t),$$

where $C_h, \varepsilon > 0$ are positive constants. Then the nudged solution $V(x, t)$ recovers the true data $U(x, t)$ at an exponential rate in time,

$$(2.4) \quad \|U(\cdot, t) - V(\cdot, t)\|_2 \leq \|U(\cdot, 0) - V(\cdot, 0)\|_2 e^{-(\mu - L_s)t} + C_{\mu, h} e^{-\varepsilon t},$$

with $C_{\mu, h} = (c_1 h^{-1} + C_0 L_s + \mu)C_h / (\mu - L_s - \varepsilon)$.

Remark 2.1. From (2.4), it appears that a larger μ results in faster convergence. However, in practice, a larger μ imposes stricter constraints on the time step due to stability considerations.

Proof. Let $W = U - V$, the difference of (2.1) and (2.2) yields

$$\frac{\partial W}{\partial t} + \frac{\partial W}{\partial x} = G - \tilde{G} + S(U) - S(V) + \mu(W - I_h(W)) - \mu W.$$

Take L^2 -inner product with W on both sides of the equation, we have

$$(2.5) \quad \begin{aligned} \frac{1}{2} \frac{d}{dt} \|W\|_2^2 &= \langle G - \tilde{G}, W \rangle + \langle S(U) - S(V), W \rangle + \mu \langle W - I_h(W), W \rangle - \mu \|W\|_2^2 \\ &\leq \|G - \tilde{G}\|_2 \|W\|_2 + \|S(U) - S(V)\|_2 \|W\|_2 + \mu \|W - I_h(W)\|_2 \|W\|_2 - \mu \|W\|_2^2 \\ &\leq \|G - \tilde{G}\|_2 \|W\|_2 + L_s \|W\|_2^2 + \mu \|W - I_h(W)\|_2 \|W\|_2 - \mu \|W\|_2^2 \\ &= \|G - \tilde{G}\|_2 \|W\|_2 + \mu \|W - I_h(W)\|_2 \|W\|_2 - (\mu - L_s) \|W\|_2^2. \end{aligned}$$

To bound the error $\|G - \tilde{G}\|_2$, we employ the condition (2.3b) and write

$$(2.6) \quad \begin{aligned} G - \tilde{G} &= I_h(G) - \tilde{G} \\ &= I_h(\partial_t U + \partial_x U - S(U)) - (\partial_t I_h(U) + I_h(\partial_x \tilde{U} - S(\tilde{U}))) \\ &= I_h(\partial_x U - \partial_x \tilde{U}) - I_h(S(U) - S(\tilde{U})). \end{aligned}$$

The temporal derivatives are canceled due to $I_h \partial_t(\cdot) = \partial_t I_h(\cdot)$. Subsequently, using conditions (1.5a), (1.5b) leads to

$$\begin{aligned} \|G - \tilde{G}\|_2 &\leq \|I_h(\partial_x(U - \tilde{U}))\|_2 + \|I_h(S(U) - S(\tilde{U}))\|_2 \\ &\leq (c_1 h^{-1} + C_0 L_s) \|U - \tilde{U}\|_2 \\ &= (c_1 h^{-1} + C_0 L_s) \|W - I_h(W)\|_2. \end{aligned}$$

By substituting the above inequality into (2.5) and rearranging, we obtain

$$(2.7) \quad \frac{d}{dt} \|W\|_2 + (\mu - L_s) \|W\|_2 \leq (c_1 h^{-1} + C_0 L_s + \mu) \|W - I_h(W)\|_2.$$

Applying Gronwall's inequality and employing condition (2.3a) yields

$$\begin{aligned} \|W(t)\|_2 &\leq \left\{ \|W(0)\|_2 + (c_1 h^{-1} + C_0 L_s + \mu) \int_0^t \|W(\tau) - I_h(W(\tau))\|_2 e^{(\mu - L_s)\tau} d\tau \right\} e^{-(\mu - L_s)t} \\ &\leq \|W(0)\|_2 e^{-(\mu - L_s)t} + \frac{(c_1 h^{-1} + C_0 L_s + \mu) C_h}{\mu - L_s - \varepsilon} e^{-\varepsilon t}. \end{aligned}$$

The desired estimate (2.4) is recovered. \square

Remark 2.2. To get a better understanding of the convergence of the algorithm, especially to understand the condition (2.3a), it is helpful to look at a simpler example:

$$(2.8a) \quad \partial_t U(x, t) = G(x, t) + sU(x, t),$$

and the corresponding nudged system becomes:

$$(2.8b) \quad \begin{cases} \partial_t V = \tilde{G} + sV + \mu(I_h(U) - I_h(V)), \\ \tilde{G}(t) = I_h(\partial_t U) - sI_h(\tilde{U}), \quad \tilde{U} = I_h(U) + (I - I_h)(V), \end{cases}$$

where $s \in \mathbb{R}$ is a constant. We consider the setting where I_h is a compact symmetric operator on $L^2(\mathbb{R}^d)$. Then there is an orthonormal basis of $L^2(\mathbb{R}^d)$ consisting of eigenfunctions, p_1, p_2, \dots , of I_h , with corresponding eigenvalues $\lambda_1, \lambda_2, \dots$. Let us assume $c_0 = 1$ in (1.5a) and I_h is semi-definite, so that $0 \leq \lambda_i \leq 1$.

Now, we can write all quantities using an eigenfunction basis as

$$U(x, t) = \sum_i u_i(t) p_i(x), \quad G(t) = \sum_i g_i(t) p_i(x), \quad V(t) = \sum_i v_i(t) p_i(x), \quad \tilde{G}(t) = \sum_i \tilde{g}_i(t) p_i(x).$$

Projection of equations (2.8) onto the eigenfunction modes yields

$$\begin{cases} u'_i(t) = g_i(t) + s u_i(t), \\ v'_i(t) = \tilde{g}_i(t) + s v_i(t) + \mu \lambda_i (u_i(t) - v_i(t)), \\ \tilde{u}_i(t) = \lambda_i u_i(t) + (1 - \lambda_i) v_i(t), \\ \tilde{g}_i(t) = \lambda_i (g_i + s u_i(t)) - s \lambda_i^2 u_i(t) - s \lambda_i (1 - \lambda_i) v_i(t). \end{cases}$$

Taking subtraction $u'_i(t) - v'_i(t)$ and using that $\lambda_i g_i = g_i$ (which follows from (2.3b)) yields

$$\frac{d}{dt} (v_i(t) - u_i(t)) = (s - \mu \lambda_i - s \lambda_i (1 - \lambda_i)) (v_i(t) - u_i(t)).$$

Therefore,

$$v_i(t) - u_i(t) = e^{(s - \mu \lambda_i - s \lambda_i (1 - \lambda_i))t} (v_i(0) - u_i(0)),$$

and so

$$(2.9) \quad V(t) = U(t) + \sum_{i=1}^{\infty} e^{(s - \mu \lambda_i - s \lambda_i (1 - \lambda_i))t} (v_i(0) - u_i(0)) p_i.$$

This implies an interplay between the structure of source terms and the conditions for convergence.

If $s > 0$, to guarantee exponential convergence for any initial condition we cannot have any $\lambda_i = 0$, or if we do, then we must ask that $u_i(0) = v_i(0)$. This is reasonable

to expect – we cannot damp any of the error modes that belong to the kernel of I_h . The lower bound of μ becomes

$$(2.10) \quad \mu \geq \frac{s + \varepsilon}{\lambda_i}, \quad \varepsilon > 0,$$

for all i such that $u_i(0) \neq v_i(0)$, from which we obtain $e^{-\varepsilon t}$ convergence. This is connected to condition (2.3a), since we can compute

$$\|W(t) - I_h(W(t))\|_2^2 = \sum_{i=1}^{\infty} (\lambda_i - 1)^2 (v_i(0) - u_i(0))^2 e^{2(s - \mu\lambda_i - s\lambda_i(1 - \lambda_i))t}.$$

Suppose that (2.10) holds whenever $0 < \lambda_i < 1$ and $u_i(0) \neq v_i(0)$, and $u_i(0) = v_i(0)$ whenever $\lambda_i = 0$. Then we recover

$$\|W(t) - I_h(W(t))\|_2 \leq \|W(0)\|_2 e^{-\varepsilon t}.$$

In other words, condition (2.3a) may be regarded as an implication from an appropriate initial state and a sufficiently large μ .

The situation becomes quite different when $s < 0$. In this scenario, the source term has stabilizing effect on the dynamics and helps dissipate the error. Note that in (2.9) the exponents are bounded by

$$s - \mu\lambda_i - s\lambda_i(1 - \lambda_i) \leq s - \lambda_i(1 - \lambda_i)s \leq \frac{3}{4}s < 0.$$

This implies unconditional convergence regardless of μ , and (2.3a) is automatically recovered with $\varepsilon = -\frac{3}{4}s > 0$. However, if the spatial transport is also considered, the dissipation of the source term is not necessarily sufficient to absorb the error from the convection term. Observational feedback control is still needed to enhance the convergence in general.

The convergence of state further implies the convergence of force, as stated in the following proposition.

Proposition 2.2. (Convergence of the force) Let (U, G) and (V, \tilde{G}) be the strong solutions of the hydrodynamics (1.3) and the nudged system (1.4) associated with I_h satisfying conditions (1.5). Assume that the convection flux $F(\cdot)$ and the source term $S(\cdot, \mathbf{x}, t)$ are Lipschitz continuous with Lipschitz constants $L_f, L_s > 0$. If the nudged solution $V(\mathbf{x}, t)$ recovers the true data $U(\mathbf{x}, t)$ asymptotically in time, $\|U(t) - V(t)\|_2 \rightarrow 0$ as $t \rightarrow \infty$, then

$$\|I_h(G(t)) - \tilde{G}(t)\|_2 \rightarrow 0, \quad t \rightarrow \infty.$$

In particular, if the force function is completely observable, $I_h(G) = G$, then

$$\|G(t) - \tilde{G}(t)\|_2 \rightarrow 0, \quad t \rightarrow \infty.$$

Proof. We have

$$\tilde{G} = \partial_t I_h(U) + I_h(\nabla \cdot F(\tilde{U}) - S(\tilde{U})), \quad \tilde{U} = I_h(U) + (I - I_h)(V),$$

and

$$I_h(G) = \partial_t I_h(U) + I_h(\nabla \cdot F(U) - S(U)).$$

The difference between the above equations yields

$$I_h(G) - \tilde{G} = I_h(\nabla \cdot (F(U) - F(\tilde{U}))) - I_h(S(U) - S(\tilde{U})).$$

Using the estimates (1.5a), (1.5b), it follows that

$$\begin{aligned}
(2.11) \quad \|I_h(G) - \tilde{G}\|_2 &\leq \|I_h(\nabla \cdot (F(U) - F(\tilde{U})))\|_2 + \|I_h(S(U) - S(\tilde{U}))\|_2 \\
&\leq c_1 h^{-1} \|F(U) - F(\tilde{U})\|_2 + C_1 h \|S(U) - S(\tilde{U})\|_2 \\
&\leq (c_1 h^{-1} L_f + C_1 h L_s) \|U - \tilde{U}\|_2 \\
&= (c_1 h^{-1} L_f + C_1 h L_s) \|(U - V) - I_h(U - V)\|_2 \\
&\leq (c_1 h^{-1} L_f + C_1 h L_s) C_0 \|U - V\|_2.
\end{aligned}$$

Hence, the convergence of the state, $\|U(t) - V(t)\|_2 \rightarrow 0$, implies the convergence of the reconstructed force, $\|I_h(G(t)) - \tilde{G}(t)\|_2 \rightarrow 0$. \square

In contrast to the earlier discussions on the incompressible Navier-Stokes equations [3, 19], our analysis in Proposition 2.1 employs the extra condition (2.3a). Essentially, the purpose is to ensure that the relaxation term $\mu(I_h(U) - I_h(V)) \xrightarrow{t \rightarrow \infty} \mu(U - V)$. The difficulty in weakening this condition arises from the lack of viscosity, which, if exists, would help dissipate the interpolation error. For instance, by adding viscous terms $\nu \partial_x^2 U$ and $\nu \partial_x^2 V$, with $\nu > 0$, to equations (2.1) and (2.2), we can derive the following inequality along the lines of (2.5) and (2.6),

$$\begin{aligned}
\frac{1}{2} \frac{d}{dt} \|W\|_2^2 + (\mu - L_s) \|W\|_2^2 &\lesssim \mu \|W - I_h(W)\|_2 \|W\|_2 - \nu \|\partial_x W\|_2^2 \\
&\lesssim c_1 \mu h \|w\|_{H^1} \|w\|_2 - \nu \|\partial_x w\|_2^2 \\
&\lesssim \frac{\mu}{2} \|w\|_2^2 + \frac{\mu c_1^2 h^2}{2} \|\partial_x w\|_2^2 - \nu \|\partial_x w\|_2^2.
\end{aligned}$$

The interpolation error is absorbed by the viscosity for $L_s \lesssim \mu \lesssim \nu/h^2$, thereby ensuring the exponential convergence. This is similar to the results in [3, 19].

Back to the setting of inviscid hydrodynamics, (1.5b) implies $\|w - I_h(w)\|_2 \sim h \|\partial_x w\|_2$, hence the condition (2.3a) amounts to $\|\partial_x w(t)\|_2 \lesssim \frac{1}{h} e^{-\varepsilon t}$. This indicates that when the data has strong oscillations (large \dot{H}^1 -norm), the analytical convergence may not be guaranteed under insufficient observations. Here we prove the worst-case scenario with respect to generic bounded data.

Proposition 2.3. (Error estimate under insufficient observations) Let (U, G) be a strong solution of the hydrodynamics (1.3) driven by the Lipschitz continuous convection flux and source term. Let (V, \tilde{G}) be a strong solution of the nudged system (1.4), where the associated I_h satisfies conditions (1.5). Assume that the data are uniformly bounded in $H^1(\mathbb{R}^d)$, i.e.,

$$\|G(t)\|_{H^1}, \|\tilde{G}(t)\|_{H^1} \leq M_g, \quad \|U(t)\|_{H^1}, \|V(t)\|_{H^1} \leq M_u,$$

for all $t > 0$, with $M_u, M_g > 0$. Then the nudged solutions approach the true solution time-asymptotically:

$$(2.12a) \quad \overline{\lim}_{t \rightarrow \infty} \|U(t) - V(t)\|_2 \leq \frac{\kappa_1}{\mu - L_s},$$

$$(2.12b) \quad \overline{\lim}_{t \rightarrow \infty} \|G(t) - \tilde{G}(t)\|_2 \leq \frac{\kappa_2}{\mu - L_s} + C_1 M_g h,$$

where $\kappa_1 = 2(L_f M_u + M_g + C_1 M_u h)$, $\kappa_2 = (c_1 h^{-1} L_f + C_1 h L_s) C_0 \kappa_1$.

Proof. Denote $W = U - V$, the difference of (1.3) and (1.4a) yields

$$\frac{\partial W}{\partial t} + \nabla \cdot (F(U) - F(V)) = G - \tilde{G} + S(U) - S(V) + \mu(W - I_h(W)) - \mu W.$$

Taking L^2 -inner product with W leads to

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|W\|_2^2 &= -\langle W, \nabla \cdot (F(U) - F(V)) \rangle + \langle W, G - \tilde{G} \rangle + \langle W, S(U) - S(V) \rangle + \langle W, W - I_h(W) \rangle - \mu \|W\|_2^2 \\ &\leq \int_{\mathbb{R}^d} (F(U) - F(V)) : \nabla W dx + \|G - \tilde{G}\|_2 \|W\|_2 + L_s \|W\|_2^2 + \|W\|_2 \|W - I_h(W)\|_2 - \mu \|W\|_2^2 \\ &\leq L_f \|W\|_2 \|W\|_{H^1} + 2M_g \|W\|_2 + L_s \|W\|_2^2 + C_1 h \|W\|_2 \|W\|_{H^1} - \mu \|W\|_2^2 \\ &\leq 2(L_f M_u + M_g + C_1 M_u h) \|W\|_2 - (\mu - L_s) \|W\|_2^2. \end{aligned}$$

Rearrange the above inequality, we obtain

$$\frac{d}{dt} \|W\|_2 + (\mu - L_s) \|W\|_2 \leq 2(L_f M_u + M_g + C_1 M_u h).$$

Integration yields

$$\|W(t)\|_2 \leq \|W(0)\|_2 e^{-(\mu - L_s)t} + \frac{2(L_f M_u + M_g + C_1 M_u h)}{\mu - L_s} (1 - e^{-(\mu - L_s)t}),$$

the bound (2.12a) is recovered by passing the limit $t \rightarrow \infty$. Furthermore,

$$\begin{aligned} \|G(t) - \tilde{G}(t)\|_2 &\leq \|I_h(G(t)) - \tilde{G}(t)\|_2 + \|G(t) - I_h(G(t))\|_2 \\ &\leq (c_1 h^{-1} L_f + C_1 h L_s) C_0 \|U(t) - V(t)\|_2 + C_1 h \|G(t)\|_{H^1} \\ &\leq (c_1 h^{-1} L_f + C_1 h L_s) C_0 \|W(t)\|_2 + C_1 M_g h. \end{aligned}$$

The second line is obtained using the inequality (2.11). The error bound of force recovery (2.12b) is a direct consequence of the large-time limit. \square

Remark 2.3. It is noticed that the error bounds in Propositions 2.1, 2.2, and 2.3 involve h^{-1} dependence. Indeed, along the lines of (2.11), the h^{-1} term occurs when estimating the error of force recovery, $\|\tilde{G} - I_h(G)\|_2$, where the usage of interpolation bound (1.5a) yields

$$\|I_h(\nabla \cdot (F(U) - F(\tilde{U})))\|_2 \leq c_1 h^{-1} \|F(U) - F(\tilde{U})\|_2.$$

Nevertheless, assuming the data admits sufficiently weak oscillations such that

$$\sup_{t \geq 0} \frac{\|\nabla \cdot (F(U) - F(\tilde{U}))\|_2}{\|F(U) - F(\tilde{U})\|_2} \leq C < \infty,$$

and the above estimate can be modified to

$$\|I_h(\nabla \cdot (F(U) - F(\tilde{U})))\|_2 \leq c_0 C \|F(U) - F(\tilde{U})\|_2,$$

following the lines of (1.6). Subsequently, the h^{-1} dependence in the error bounds will be removed.

Despite the difficulty in deriving an optimal error bound under less ideal assumptions, in actual computations, we should still expect a more accurate solution under a finer observation. Meanwhile, though the error bound in

Proposition 2.3 does not offer an explicit criterion for determining the minimum observation density required to meet a specific error threshold, we will demonstrate that our proposed algorithm maintains reasonable accuracy even with relatively sparse data through numerical experiments in Section 3.4.

2.2. Discretization of nudged system. To solve the nudged system (1.4), since both the state V and the force \tilde{G} are evolving in time, we introduce the following discretization. Given the solutions $V^n = \{V_j^n\}_{j=1}^N$ and $\tilde{G}^n = \{\tilde{G}_j^n\}_{j=1}^N$ at time level $t = t^n$, the solution at $x = x_j$ and $t = t^{n+1}$ are computed through:

$$\begin{cases} \frac{V_j^{n+1} - V_j^n}{\Delta t} + \delta_x^{\Delta x} F(V^n)_j = \tilde{G}_j^n + S(V^n, x_j, t^n), \\ \tilde{G}_j^{n+1} = \partial_t I_h(U(t^{n+1}))_j + I_h(\delta_x^{\Delta x} F(\tilde{U}^{n+1}) - S(\tilde{U}^{n+1}))_j, \\ \tilde{U}^{n+1} = I_h(U(t^{n+1})) + (I - I_h)(V^{n+1}). \end{cases}$$

Here Δt is the time step, and $\delta_x^{\Delta x} F$ is the approximate flux derivative obtained with non-oscillatory discretization such as ENO/WENO schemes with spatial mesh size Δx . We assume that the observation $I_h(U(t))$ is recorded continuously in time, whereas $\partial_t I_h(U(t^{n+1}))$ can be computed with high order finite difference methods. As such, the above forward Euler discretization can be directly replaced by higher order strong stability preserving (SSP) Runge-Kutta methods [20].

It is worth noting that the numerical scheme described above yields nodal values, $\{V_j\}$, over the computational grids, $\{x_j\}$. The interpolation $I_h(\cdot)$, however, is constructed based on the data over the observation grids, $\{x_j^{ob}\}$, which may not coincide with x_j in general. Therefore, appropriate interpolation techniques (such as linear interpolation or cubic spline) may be necessary to approximate nodal values (x_j^{ob}, V_j^{ob}) . These approximations are then used to compute the interpolations $(x_j, I_h(V)_j)$.

We also emphasize the simplicity of our discretization. In the previous works [34, 16], it was proposed to generate a *sequence* of approximations, $(V_k(\mathbf{x}, t), G_k(\mathbf{x}, t))|_{t \geq t_k}$, $t_k \gg t_{k-1}$, $k = 0, 1, 2, \dots$. The solution $V_k(\mathbf{x}, t)$ was obtained by solving the nudged equations with the re-initialization, $V_k(\mathbf{x}, t_k) = V_{k-1}(\mathbf{x}, t_k)$, coupled with the frozen force, $G_{k-1}(\mathbf{x}, t)|_{t \geq t_k}$, over the time interval $t \in [t_k, \infty)$. Although this algorithm guarantees exponential convergence when applied to the incompressible Navier-Stokes equations, the need to repeatedly solve the equation can be infeasible in industrial applications. In contrast, our discretization method updates the states and the force simultaneously *on the fly*, hence significantly reducing the computational cost.

2.3. Smooth fitting with kernel regression. In practical computations, the accuracy of approximate solutions can depend on the selection of I_h , particularly in scenarios with restricted observational data. Common methods for achieving a smooth interpolation of observed data involve using Lagrangian interpolation with nodal values distributed uniformly at a mesh size h , or projecting onto truncated Fourier series with a cut-off frequency of $1/h$. However, these approaches using high-order polynomials or trigonometric polynomials may encounter significant challenges such as numerical oscillations at boundaries (known as Runge's phenomenon) or sharp extrema (referred to as Gibbs' phenomenon).

To reduce the undesired numerical artifacts, we propose to *fit* the observed data with kernel regression,

$$(2.13) \quad I_h(U)(\mathbf{x}) = \frac{\sum_{j=1}^{N_{ob}} K_\sigma(|\mathbf{x} - \mathbf{x}_j^{ob}|) U_j^{ob}}{\sum_{j=1}^{N_{ob}} K_\sigma(|\mathbf{x} - \mathbf{x}_j^{ob}|)},$$

where U_j^{ob} is the observation at $\mathbf{x}_j^{ob} \in \mathbb{R}^d$, $K_\sigma(\cdot)$ is the Gaussian kernel with bandwidth σ ,

$$(2.14) \quad K_\sigma(|\mathbf{x}|) = \frac{1}{\sigma^d} \exp\left\{-\frac{|\mathbf{x}|^2}{2\sigma^2}\right\}.$$

A larger σ leads to stronger kernel smoothing, whereas a smaller σ generates a theoretically more accurate fitting. However, taking σ too small in actual computations can result in severe arithmetic errors. Hence, to achieve a balance between solution sharpness and computational stability, we suggest taking $\frac{\sqrt{d}}{2}h < \sigma \leq h$. It's important to note that the approximate data $I_h(U)$ is derived from weighted averages of nearby observations, which generally do not preserve nodal values at interpolation points. In practical applications where observations may be noisy due to measurement errors, using kernel regression instead of classical interpolation methods can better mitigate over-fitting issues by appropriately adjusting the kernel bandwidth.

A common challenge when using kernel regression (2.13) is the larger error observed at the boundary of the compact support. This is due to the boundary conditions of the data $U(\mathbf{x})$ not being incorporated, leading to higher accuracy in the interior of the domain compared to near the boundary. To address this, we impose a simple boundary correction via ghost cell extensions. For the convenience of demonstration, we consider uniformly distributed observation grids $\{x_j^{ob}\}_{j=1}^{N_{ob}}$ over the computational interval $[a, b]$ with $x_1^{ob} = a$ and $x_{N_{ob}}^{ob} = b$. We modify the kernel regression with

$$(2.15) \quad I_h(U)(x) = \frac{\sum_{j=-(m-1)}^{N_{ob}+m} K_\sigma(|x - x_j^{ob}|) U_j^{ob}}{\sum_{j=-(m-1)}^{N_{ob}+m} K_\sigma(|x - x_j^{ob}|)}.$$

The ghost nodes are given by

$$x_{1-j}^{ob} = a - jh, \quad x_{N_{ob}+j}^{ob} = b + jh, \quad j = 1, 2, \dots, m.$$

This way, boundary conditions are encoded implicitly in the extrapolation of $\{U_j^{ob}\}_{j \leq 0}$ and $\{U_j^{ob}\}_{j \geq N_{ob}+1}$. For instance, the extended data at the left boundary are set to be

$$U_{1-j}^{ob} = \begin{cases} U_{N_{ob}-j}^{ob} & , \text{ periodic boundary} \\ 2U_a - U_{1+j}^{ob} & , \text{ Dirichlet boundary } U(a) = u_a \quad , \quad j = 1, 2, \dots, m. \\ U_{1+j}^{ob} & , \text{ Neumann boundary } U'(a) = 0 \end{cases}$$

The data can be extrapolated at the right boundary in a similar manner. Although the kernel regression does not meet the conditions specified in (1.5), it does satisfy estimates similar to those in (1.5b), as is verified in the following result.

Theorem 2.4. *Let $U_j^{ob} = U(\mathbf{x}_j^{ob})$, where $U : \Omega \rightarrow \mathbb{R}$, and assume that $K_\sigma(\mathbf{x}) = 0$ for $|\mathbf{x}| \geq C\sigma$. Then the following hold.*

(i) If U is Lipschitz continuous with Lipschitz constant L , then

$$\|I_h(U) - U\|_\infty \leq CL\sigma.$$

(ii) If ∇U is Lipschitz continuous with Lipschitz constant L , then

$$\|I_h(U) - U\|_2 \leq C\sigma\|U\|_{H^1} + LC^2\sigma^2|\Omega|^{1/2}.$$

Proof. For simplicity, let us write

$$d_\sigma(x) = \sum_{j=1}^{N_{ob}} K_\sigma(|\mathbf{x} - \mathbf{x}_j^{ob}|).$$

Then we have

$$\begin{aligned} |I_h(U)(\mathbf{x}) - U(\mathbf{x})| &= \left| \frac{1}{d_\sigma(x)} \sum_{j=1}^{N_{ob}} K_\sigma(|\mathbf{x} - \mathbf{x}_j^{ob}|) U(\mathbf{x}_j^{ob}) - U(\mathbf{x}) \right| \\ &\leq \frac{1}{d_\sigma(x)} \sum_{j=1}^{N_{ob}} K_\sigma(|\mathbf{x} - \mathbf{x}_j^{ob}|) |U(\mathbf{x}_j^{ob}) - U(\mathbf{x})| \\ &\leq \frac{L}{d_\sigma(x)} \sum_{j=1}^{N_{ob}} K_\sigma(|\mathbf{x} - \mathbf{x}_j^{ob}|) |\mathbf{x}_j^{ob} - \mathbf{x}| \\ &\leq \frac{CL\sigma}{d_\sigma(x)} \sum_{j=1}^{N_{ob}} K_\sigma(|\mathbf{x} - \mathbf{x}_j^{ob}|) = CL\sigma, \end{aligned}$$

which establishes (i).

For (ii), we continue the computation above to obtain

$$\begin{aligned} |I_h(U)(\mathbf{x}) - U(\mathbf{x})| &\leq \frac{1}{d_\sigma(x)} \sum_{j=1}^{N_{ob}} K_\sigma(|\mathbf{x} - \mathbf{x}_j^{ob}|) |U(\mathbf{x}_j^{ob}) - U(\mathbf{x})| \\ &\leq \frac{1}{d_\sigma(x)} \sum_{j=1}^{N_{ob}} K_\sigma(|\mathbf{x} - \mathbf{x}_j^{ob}|) (|\nabla U(\mathbf{x}) \cdot (\mathbf{x}_j^{ob} - \mathbf{x})| + L|\mathbf{x}_j^{ob} - \mathbf{x}|^2) \\ &\leq \frac{1}{d_\sigma(x)} \sum_{j=1}^{N_{ob}} K_\sigma(|\mathbf{x} - \mathbf{x}_j^{ob}|) (|\nabla U(\mathbf{x})| C\sigma + LC^2\sigma^2) \\ &\leq |\nabla U(\mathbf{x})| C\sigma + LC^2\sigma^2. \end{aligned}$$

The proof is completed by taking the L^2 norm on both sides. \square

Remark 2.5. The estimates in Theorem 2.4 are similar to (1.5b) with $m = 1$, provided we use that $\sigma \sim h$, though they are not exactly the same. Unfortunately, it is not possible to prove estimates similar to (1.5a); instead, we can prove estimates on the derivatives of the interpolated function, that is the quantity $\|\partial^\alpha I_h(U)\|$ is straightforward to estimate, since in the case that $K_\sigma(\mathbf{x}) = \frac{1}{\sigma^d} K_1(\mathbf{x}/\sigma)$ we have

$$(2.16) \quad \partial^\alpha I_h(U)(\mathbf{x}) = \frac{\sum_{j=1}^{N_{ob}} \sigma^{-|\alpha|} \partial^\alpha K_\sigma(|\mathbf{x} - \mathbf{x}_j^{ob}|) U_j^{ob}}{\sum_{j=1}^{N_{ob}} K_\sigma(|\mathbf{x} - \mathbf{x}_j^{ob}|)}.$$

However, such estimates are not helpful in our analysis, so we do not pursue them further. Even though kernel regression does not fully satisfy (1.5), the method still

produces satisfactory numerical solutions, as demonstrated by the examples in the next section.

In summary, our construction of interpolant operator via kernel regression provides significant practical benefits in the following aspects:

- (i) Convenient implementation of flexible boundary conditions: kernel regression allows for straightforward handling of complex and non-standard boundary conditions, which is often a crucial challenge suffered by other interpolants, e.g. spectral projections or piecewise linear interpolation;
- (ii) Robust processing of noisy observations: our approach demonstrates inherent resilience to noisy observational data, facilitating more reliable interpolation in real-world scenarios where measurements are often imperfect. This is explicitly shown in the tests with noisy data presented in the following section 3.4.

These advantages are critical for potential extensions to industrial applications, such as reactor simulations with complex geometry and possibly inaccurate measured data. In simpler computational settings, alternative choices of the interpolant operator may also provide satisfactory results. As the main focus of our paper is to validate the nudging-based feedback control framework as applied to generic inviscid hydrodynamics, particularly when recovering moments from incomplete observations (see Section 4), further investigations of other interpolants will be left for future study.

3. NUMERICAL EXPERIMENTS I – STATES AND FORCE RECOVERY

We verify the effectiveness of the algorithm (1.4) in both scalar and system test cases. In all experiments, unless stated otherwise, the operator I_h is derived using the kernel regression (2.15) with $m = 3$ ghost nodes applied at each boundary. The kernel bandwidth in equation (2.14) is set to $\sigma = h$. The observation grids $\{x_j^{ob}\}$ are uniformly placed over the computational domain. The nudged equations are discretized with the 5th-order WENO finite difference scheme in space, and time integration is performed with an explicit SSPRK3 method with $CFL = 0.7$. Reference solutions are obtained by discretizing the exact dynamics with WENO5-SSPRK3 on fine meshes.

3.1. Scalar test case. Consider the following exact dynamics:

$$(3.1) \quad \partial_t u + \partial_x u = -\partial_x p + S(u), \quad x \in [0, 2\pi],$$

with

$$p = \frac{u^2}{6}, \quad S(u) = 0.2\sqrt{1 + u^2}.$$

The initial data consists of sine and cosine waves with different frequencies,

$$u_0(x) = -0.8 \sin(x) + 0.4 \sin(2x) + 0.02 \cos(10x).$$

Periodic boundary conditions are applied at both endpoints. In this test case, the pressure gradient $g = -\partial_x p$ generates high-frequencies and sharp extrema over time due to the non-linearity.

Assume that the source function $S(\cdot)$ is known, our goal is to recover the force term, $g = -\partial_x(\frac{u^2}{6})$, by evolving the nudged system

$$(3.2) \quad \begin{cases} \partial_t v + \partial_x v = \tilde{g} + 0.2\sqrt{1+v^2} + \mu(I_h(u) - I_h(v)), \\ \tilde{g} = \partial_t I_h(u) + I_h(\partial_x \tilde{u} - 0.2\sqrt{1+\tilde{u}^2}), \quad \tilde{u} = I_h(u) + (I - I_h)(v) \end{cases},$$

with zero initial data,

$$v_0(x) = 0, \quad \tilde{g}_0(x) = 0.$$

The nudging coefficient μ is set to be 3. We intend to recover the solution state u and the force function g from $N_{ob} = 150$ spatial observation points. The nudged system is discretized with $N = 800$ equidistant grid points.

Figures 3.1, 3.2 display the evolution of approximate solutions up to $T = 1.5$. The computed solutions v and \tilde{g} converge rapidly to the reference solutions u and g over time. Although the smoothing effect of kernel regression causes some clipping at local extrema in the reconstructed force \tilde{g} , the overall trend of the exact data is captured effectively. The L^1 -relative error at terminal time is $\frac{\|g(T) - \tilde{g}(T)\|_1}{\|g(T)\|_1} = 3.37 \times 10^{-2}$. The history of state error, $\|u(t) - v(t)\|_1$, is shown in Figure 3.3. The result confirms the exponential convergence rate, as expected from convergence analysis.

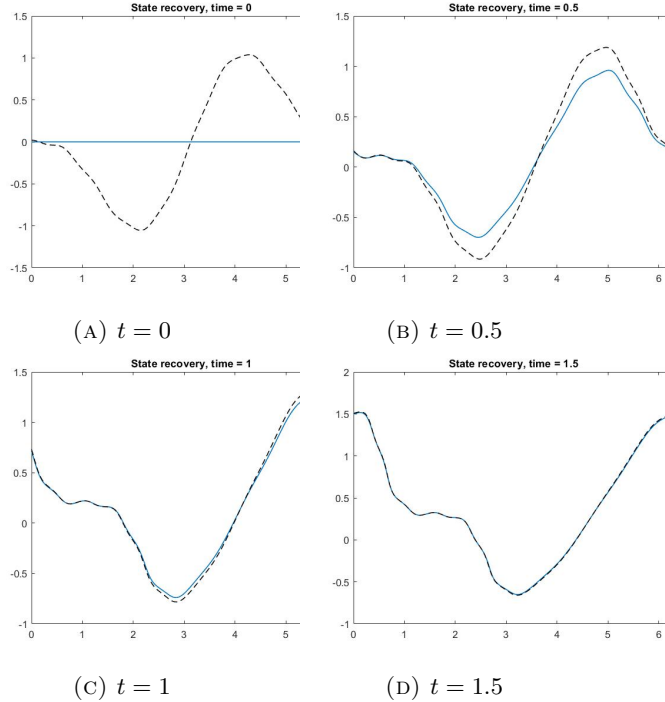


FIGURE 3.1. Scalar dynamics. State recovery. Nudged solution (blue lines) vs. reference solution (black dash lines).

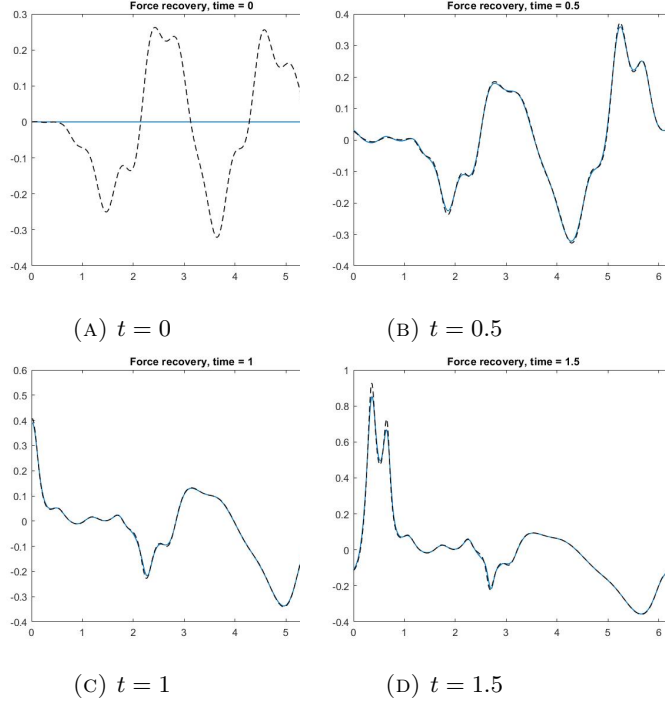


FIGURE 3.2. Scalar dynamics. Force recovery. Nudged solution (blue lines) vs. reference data (black dash lines).

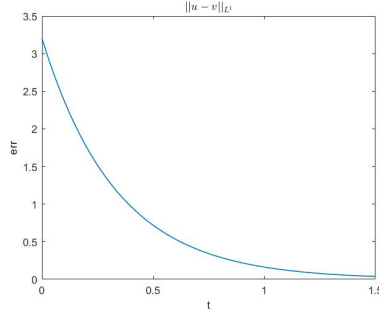


FIGURE 3.3. Scalar dynamics. History of L^1 state error.

3.2. Nonlinear system. We now examine the performance of the algorithm when applied to the one-dimensional isentropic Euler system for gas dynamics:

$$\frac{\partial U}{\partial t} + \frac{\partial}{\partial x} F(U) = G,$$

with

$$U = \begin{bmatrix} \rho \\ \rho u \end{bmatrix}, \quad F(U) = \begin{bmatrix} \rho u \\ \rho u^2 \end{bmatrix}, \quad G = -\frac{\partial}{\partial x} \begin{bmatrix} 0 \\ p \end{bmatrix}.$$

Here ρ is the density, $m = \rho u$ is the momentum, and p is the pressure given by the power law

$$p = \kappa \rho^\gamma, \quad \kappa = 1, \gamma = 1.4.$$

The exact data is generated with the initial condition

$$\rho_0(x) = 1 + 0.2 \sin(\pi x), \quad u_0(x) = 1,$$

over the computational domain $[0, 4]$. Periodic boundary conditions are applied at both endpoints. Assuming sparse observations are available for the density and the momentum, we expect to recover the true data with the nudged equations,

$$\begin{cases} \frac{\partial V}{\partial t} + \frac{\partial}{\partial x} F(V) = \tilde{G} + \mu(I_h(U) - I_h(V)), \\ \tilde{G} = \frac{\partial}{\partial t} I_h(U) + I_h\left(\frac{\partial}{\partial x} F(\tilde{U})\right), \quad \tilde{U} = I_h(U) + (I - I_h)(V). \end{cases}$$

The two components of V are the approximate density $\hat{\rho}$ and momentum $\hat{m} = \hat{\rho} \hat{u}$. The reconstructed pressure gradient, $\tilde{\partial_x p}$, is given by the second component of $-\tilde{G}$. The nudged system is initialized with

$$\hat{\rho}_0(x) = 1, \quad \hat{u}_0(x) = 0.5, \quad \tilde{G}_0(x) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The nudging coefficient is set to $\mu = 5$. Numerical discretization is implemented over $N = 600$ mesh points.

Figures 3.4 – 3.6 present the approximations to each state component and the pressure gradient based on $N_{ob} = 150$ equidistant observation grids up to $T = 1.5$. The results have shown a stable convergence of computed solutions towards the exact data. The recovered pressure gradient, although slightly clipped at sharp extrema, shows a small terminal relative error of $\frac{\|\partial_x p(T) - \tilde{\partial_x p}(T)\|_1}{\|\partial_x p(T)\|_1} = 3.56 \times 10^{-2}$. The state errors of density and momentum are shown in Figure 3.7, which again verify the exponential convergence rates.

3.3. Multi-dimensional system. To validate our algorithm in multi-dimensional problems, we consider the two-dimensional isentropic Euler equations

$$\frac{\partial U}{\partial t} + \frac{\partial}{\partial x} F_x(U) + \frac{\partial}{\partial y} F_y(U) = G$$

with

$$U = \begin{bmatrix} \rho \\ \rho u \\ \rho v \end{bmatrix}, \quad F_x(U) = \begin{bmatrix} \rho u \\ \rho u^2 \\ \rho uv \end{bmatrix}, \quad F_y(U) = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 \end{bmatrix}, \quad G = - \begin{bmatrix} 0 \\ \partial_x p \\ \partial_y p \end{bmatrix}.$$

The system is closed by the following constitutive relation:

$$p = \kappa \rho^\gamma, \quad \kappa = 0.5, \gamma = 1.2.$$

The exact data is generated with the initial conditions

$$\rho_0(x, y) = 1 + 0.4 \sin(\pi x) \cos(\pi y), \quad u_0(x, y) = 1, \quad v_0(x, y) = 0.5, \quad x, y \in [-1, 1]^2.$$

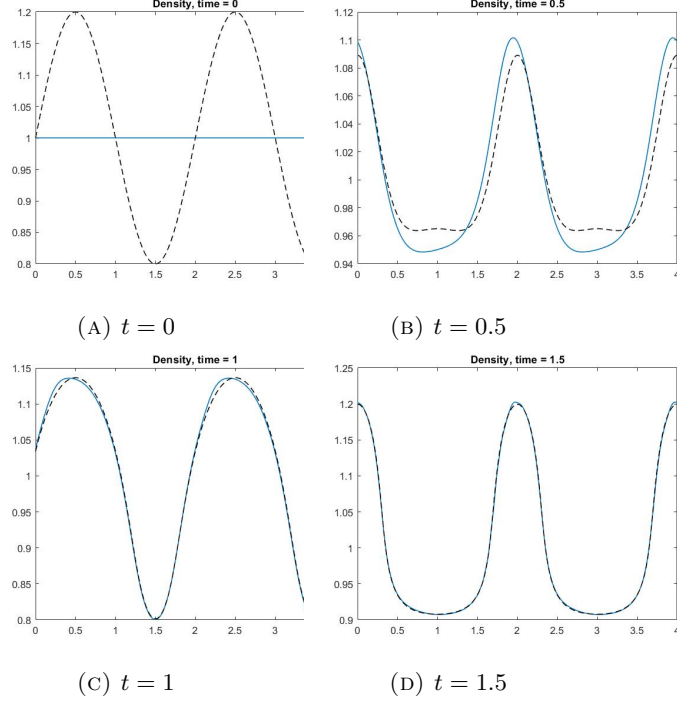


FIGURE 3.4. 1D isentropic Euler. Density recovery. Nudged solution (blue lines) vs. reference data (black dash lines).

Periodic boundary conditions are applied at the four boundaries. The nudged solutions are generated with the dynamics

$$\begin{cases} \frac{\partial V}{\partial t} + \frac{\partial}{\partial x} F_x(V) + \frac{\partial}{\partial y} F_y(V) = \tilde{G} + \mu(I_h(U) - I_h(V)) , \\ \tilde{G} = \frac{\partial}{\partial t} I_h(U) + I_h\left(\frac{\partial}{\partial x} F_x(\tilde{U}) + \frac{\partial}{\partial y} F_y(\tilde{U})\right), \quad \tilde{U} = I_h(U) + (I - I_h)(V) . \end{cases}$$

The three components of V are the approximate density $\hat{\rho}$, x -momentum $\hat{m}_x = \hat{\rho}\hat{u}$, and y -momentum $\hat{m}_y = \hat{\rho}\hat{v}$. The reconstructed pressure derivatives, $\partial_x p$ and $\partial_y p$, are given by the second and the third components of $-\tilde{G}$. We initialize the nudged system with

$$\hat{\rho}_0(x, y) = 1, \quad \hat{u}_0(x, y) = 1, \quad \hat{v}_0(x, y) = 0.5, \quad \tilde{G} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The nudging coefficient is set to $\mu = 8$. The kernel regression operator I_h is obtained with bandwidth $\sigma = \max\{h_x^{ob}, h_y^{ob}\}$, where h_x and h_y are the mesh sizes of observation grids along x - and y -directions. The exact data are observed over the 60×60 observation mesh. The nudged equations are discretized with WENO5-SSPRK3 under 180×180 computational mesh.

Figure 3.8 presents the computed solutions along $y = 0$ up to $T = 1$. The results verify the quick convergence of the approximate density and momentum towards

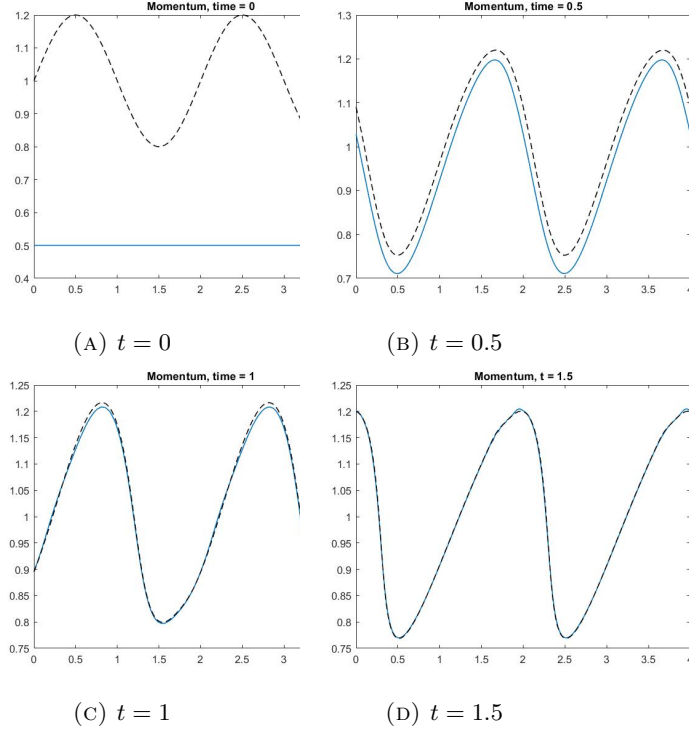


FIGURE 3.5. 1D isentropic Euler. Momentum recovery. Nudged solution (blue lines) vs. reference data (black dash lines).

the reference solutions. The computed pressure gradients closely characterize the variation in the exact data. The terminal relative errors are $\frac{\|\partial_x p(T) - \partial_x p(T)\|_1}{\|\partial_x p(T)\|_1} = 1.80 \times 10^{-2}$ and $\frac{\|\partial_y p(T) - \partial_y p(T)\|_1}{\|\partial_y p(T)\|_1} = 1.79 \times 10^{-2}$. The history of state errors, as shown in Figure 3.9, verifies the exponential convergence in multi-dimensions.

3.4. Noisy measurement and sparse observation. In practical implementations of data assimilation algorithms, challenges often arise due to noisy and/or sparse observations, which can degrade the accuracy of data recovery. In this section, we primarily investigate the performance of our algorithm under noisy observational data and varying observation density.

For convenience of demonstration, we consider the scalar dynamics (3.1) and the corresponding nudging system (3.2). At first, we examine the performance of our data assimilation algorithm under noisy observations, namely, the feedback control term is constructed with $I_h(u^\varepsilon(t))$ where

$$u^\varepsilon(x_j^{ob}, t) = u(x_j^{ob}, t) + \varepsilon \mathcal{N}_{j,t},$$

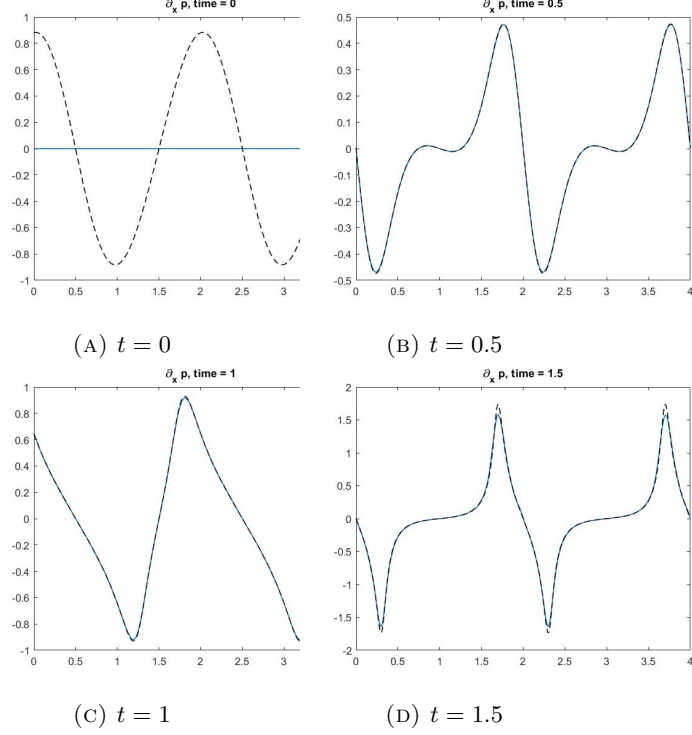


FIGURE 3.6. 1D isentropic Euler. Pressure gradient recovery. Nudged solution (blue lines) vs. reference data (black dash lines).

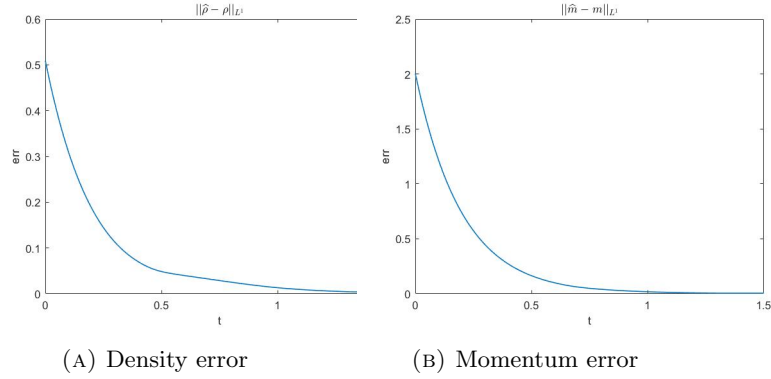


FIGURE 3.7. 1D isentropic Euler. History of state errors.

with $\{\mathcal{N}_{j,t}\}$ independently and identically distributed with respect to the standard normal distribution. Following the computational settings in Section 3.1, we still take the nudging coefficient $\mu = 3$ and assume that the measurement of u is conducted over $N_{ob} = 150$ grids, i.e., the mesh size of observation grids is $h = 2\pi/150$. We take the magnitude of noise to $\varepsilon = 10^{-4}$. As demonstrated in Section 2.3, a key advantage of using kernel regression as an observable interpolant lies

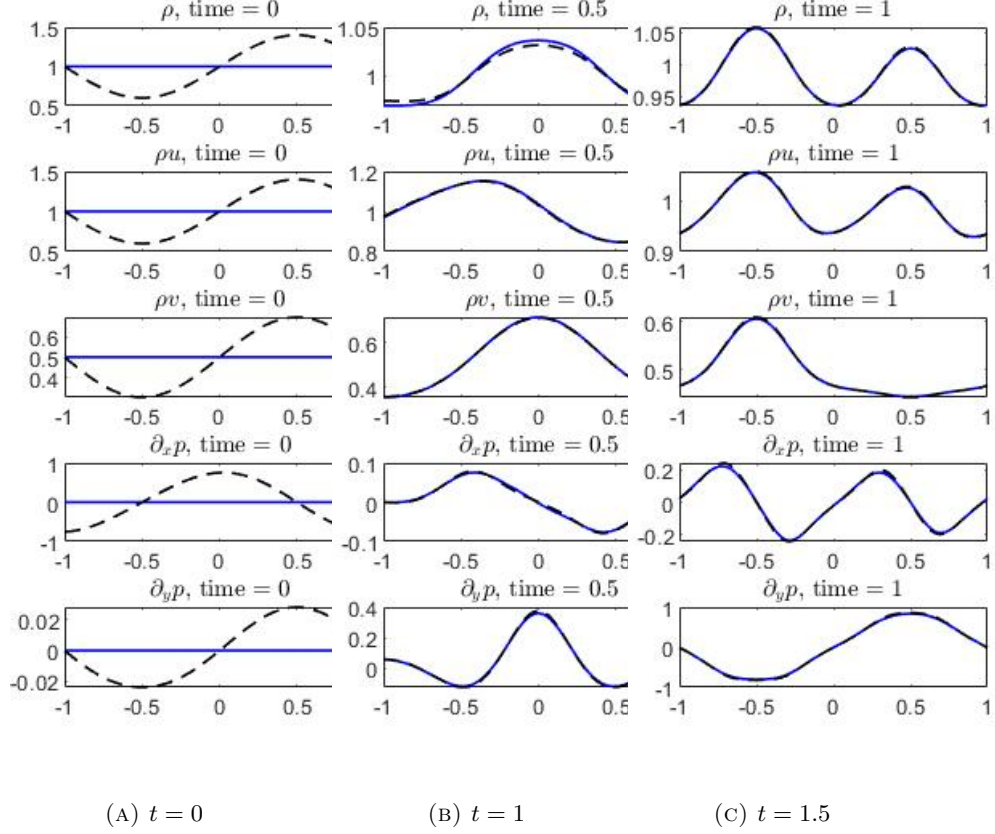


FIGURE 3.8. 2D Euler. Nudged solutions (blue lines) vs. exact data (black dash lines).

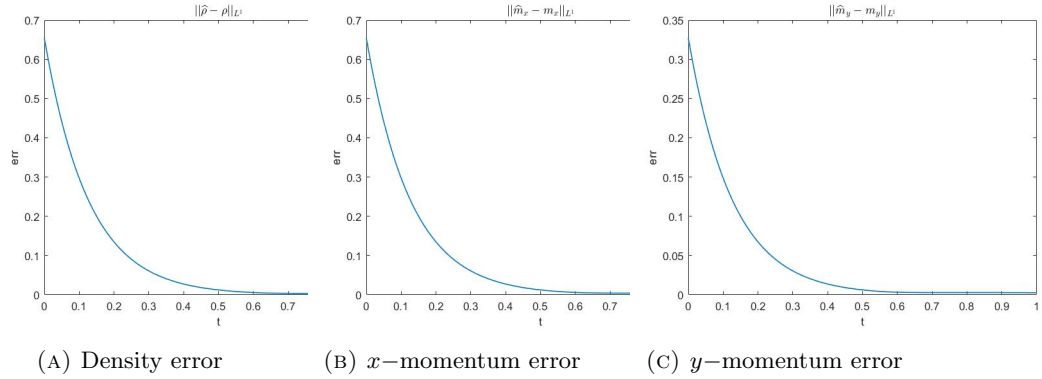


FIGURE 3.9. 2D Euler. History of state errors.

in its convenience and robustness in processing noisy data. To illustrate this, we

display the computational results in Figure 3.10. It is observed that, while the recovery of the solution state u remains relatively unaffected by small levels of noise, the recovery of the unknown forcing term $g = -\partial_x p$ is significantly more sensitive to inaccuracies in the observations. By increasing the bandwidth of the regression kernel, the recovered data can be effectively smoothed, although at the cost of more dissipative local extrema. In practical applications, the value of σ can be tuned to strike a balance between preserving sharp features and achieving desired smoothness in the reconstructed solutions.

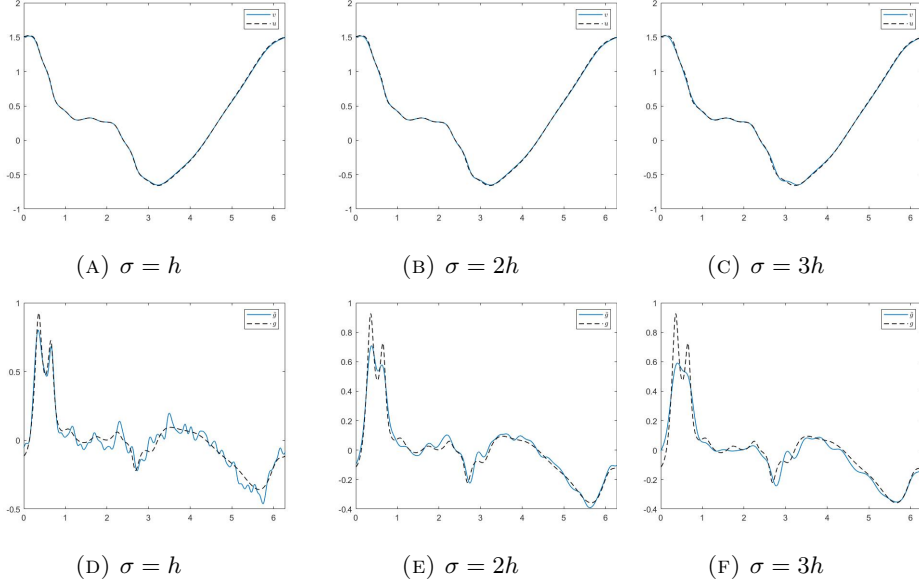


FIGURE 3.10. Scalar dynamics with noisy observation. Nudged solutions (blue lines) vs. reference solutions (black dash lines) at $t = 1.5$ under varying regression bandwidth. The first row presents the results for state recovery, and the second row presents the results for force recovery.

We also evaluate the performance of our algorithm under sparse observational data. In each test case, the bandwidth of the regression kernel is set to $\sigma = h = 2\pi/N_{ob}$, with N_{ob} being the number of observation grids. As shown in Figure 3.11, it can be found that reducing the number of observation grids leads to less clear details in the recovered solution. Nevertheless, the overall variation of the solutions can still be reasonably captured, even when the observations are very sparse. In practical scenarios where the number of observation points (sensors) is limited, local refinement of the observation grids in the regions with complex interactions may further improve the reconstruction accuracy. A detailed investigation of such adaptive strategies is left for our future work.

4. MOMENT RECOVERY WITH INCOMPLETE OBSERVATIONS

The previous discussion has established the validity of algorithm (1.4) when observational data for all state components are available, a situation we referred

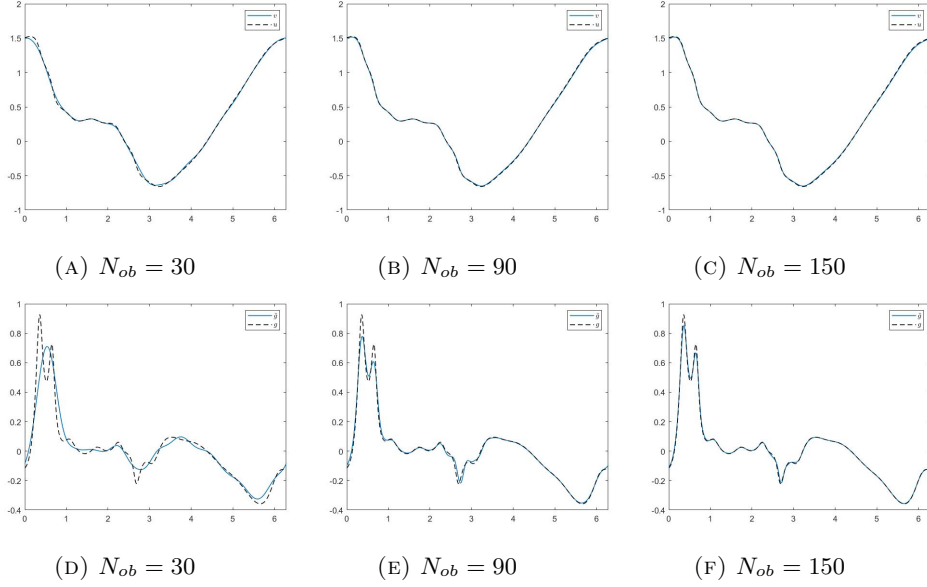


FIGURE 3.11. Scalar dynamics with varying observation density. Nudged solutions (blue lines) vs. reference solutions (black dash lines) at $t = 1.5$ under varying observation density. The first row presents the results for state recovery, and the second row presents the results for force recovery.

to as *complete* observation. However, in some applications, only some of the state components may be observable. In such cases, where data for some state components are missing, we refer to the observation as *incomplete*. This scenario is particularly relevant in kinetic theory, where one aims to derive a moment system as a reduced-order model. In this section, we demonstrate that the method proposed in Section 2 can be extended to handle incomplete observations, providing an efficient mechanism for moment recovery.

4.1. Problem set up. As a concrete example, we consider the one-dimensional radiative transfer equation (RTE):

$$(4.1) \quad \partial_t f + v \partial_x f = \sigma_s \left(\frac{1}{2} \int_{-1}^1 f dv - f \right) - \sigma_a f, \quad x \in [0, 1], \quad v \in [-1, 1].$$

Here $f(x, v, t)$ is the specific intensity of radiation, $v \in [-1, 1]$ is the cosine of the angle between the photon velocity and the x -axis. $\sigma_s(x) \geq 0$ and $\sigma_a(x) \geq 0$ are the scattering and absorption coefficients, respectively.

Let $P_k(v)$ be the Legendre polynomial of k degrees. Integrating the equation (4.1) against $P_k(v)$ over $v \in [-1, 1]$ and using Bonnet's recursion formula, one obtains the moment equations up to the N -th order:

$$\begin{aligned}
(4.2) \quad & \partial_t m_0 + \partial_x m_1 = -\sigma_a m_0, \\
& \partial_t m_1 + \frac{1}{3} \partial_x m_0 + \frac{2}{3} \partial_x m_2 = -(\sigma_s + \sigma_a) m_1, \\
& \dots \\
& \partial_t m_N + \frac{N}{2N+1} \partial_x m_{N-1} + \frac{N+1}{2N+1} \partial_x m_{N+1} = -(\sigma_a + \sigma_s) m_N,
\end{aligned}$$

where the k -th moment is defined as

$$m_k(x, t) = \frac{1}{2} \int_{-1}^1 f(x, v, t) P_k(v) dv.$$

The system (4.2) is deemed a model reduction for (4.1), but it is not closed as the evolution of m_N depends on the unknown moment m_{N+1} . One approach to closure involves using a machine learning-based moment closure [23, 25, 24]. Given sufficient data of moments m_1 to m_N , which can be obtained by solving (4.1), one attempts to learn a relationship such as:

$$(4.3) \quad \partial_x m_{N+1} = \sum_{k=0}^N \mathcal{N}_k(m_0, m_1, \dots, m_N) \partial_x m_k.$$

However, direct simulation of the kinetic equation (4.1) can be expensive, and instead we aim to reconstruct $\partial_x m_{N+1}$ from the observed data of lower-order moments. To this end, we write the moment equations in the condensed form:

$$\begin{aligned}
(4.4) \quad & \frac{\partial \mathbf{m}_N}{\partial t} + A_N \frac{\partial \mathbf{m}_N}{\partial x} = \mathbf{g}_{N+1} + S_N \mathbf{m}_N, \\
& \mathbf{g}_{N+1} = -\frac{N+1}{2N+1} \partial_x m_{N+1} \mathbf{e}_{N+1},
\end{aligned}$$

with $\mathbf{m}_N = (m_0, m_1, \dots, m_N)^\top$, $\mathbf{e}_{N+1} = (0, \dots, 0, 1)^\top$, and the coefficient matrices:

$$(4.5) \quad A_N = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 & \dots & 0 \\ 0 & \frac{2}{5} & 0 & \frac{3}{5} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{N-1}{2N-1} & 0 & \frac{N}{2N-1} \\ 0 & 0 & \dots & 0 & \frac{N}{2N+1} & 0 \end{bmatrix} \in \mathbb{R}^{(N+1) \times (N+1)},$$

$$(4.6) \quad S_N = \text{diag}(-\sigma_a, -(\sigma_s + \sigma_a), \dots, -(\sigma_s + \sigma_a)) \in \mathbb{R}^{(N+1) \times (N+1)}.$$

Ideally, if all components of \mathbf{m}_N are measurable at sparse observation points, one could asymptotically reconstruct the continuous solution fields $\mathbf{m}_N(x, t)$ and $\partial_x m_{N+1}(x, t)$ over time, following the methodology described in (1.4). These reconstructed data could then be utilized to train neural networks to model the relationship in (4.3). However, obtaining data for all moments is impractical in real-world applications. Indeed, higher-order moments beyond the first two—angle-averaged intensity m_0 and Eddington flux m_1 —often lack direct physical interpretation and are not readily available for measurement. Nonetheless, [23] highlights the necessity of using a sufficient number of moments ($N \geq 5$) to ensure accurate results in both optically-thin and optically-thick

regimes. Therefore, it is crucial to obtain data for m_2, m_3, \dots, m_{N+1} . To avoid the challenges associated with microscopic or full measurements of the particle distribution f in complex applications, we propose using the continuous data assimilation mechanism described earlier.

4.2. Moment recovery via continuous data assimilation. Back to the moment system (4.2) (or (4.4)), we distinguish the notations of the ground truth data, $m_k(x, t)$, and the numerical solutions, $\widehat{m}_k(x, t)$. Our purpose is to recover the continuous data fields of all moments, $\mathbf{m}_{N+1} = (m_0, m_1, \dots, m_{N+1})^\top$, while using the expensive observations of higher-order moments to a minimal extent. To this end, we assume that the full measurements of f , and thereby complete observations of all moments of interest, are *only conducted at the boundaries*. At the interior of the solution domain, we employ incomplete observations with the macroscopically observable moments up to the n -th order, $\mathbf{m}_n = (m_0, m_1, \dots, m_n)^\top$, $n < N$. Moreover, the gradients $\partial_x m_{n+1}, \partial_x m_{n+2}, \dots, \partial_x m_{N+1}$ are assumed to be completely observable with respect to the interpolant operator I_h , i.e.,

$$I_h(\partial_x m_k) = \partial_x m_k, \quad k = n+1, n+2, \dots, N.$$

We propose the following moment recovery algorithm for the one-dimensional RTE.

Step 1. Reconstruct the moments $\widehat{\mathbf{m}}_n(x, t) = (\widehat{m}_0(x, t), \widehat{m}_1(x, t), \dots, \widehat{m}_n(x, t))^\top$ by solving the nudged system with observations \mathbf{m}_n :

$$\begin{cases} \frac{\partial \widehat{\mathbf{m}}_n}{\partial t} + A_n \frac{\partial \widehat{\mathbf{m}}_n}{\partial x} = \widetilde{\mathbf{g}}_{n+1} + S_n \widehat{\mathbf{m}}_n + \mu(I_h(\mathbf{m}_n) - I_h(\widehat{\mathbf{m}}_n)), \\ \widetilde{\mathbf{g}}_{n+1} = \frac{\partial}{\partial t} I_h(\mathbf{m}_n) + I_h(A_n \frac{\partial \widetilde{\mathbf{m}}_n}{\partial x} - S_n \widetilde{\mathbf{m}}_n), \quad \widetilde{\mathbf{m}}_n = I_h(\mathbf{m}_n) + (I - I_h)(\widehat{\mathbf{m}}_n). \end{cases}$$

Step 2. Obtain the reconstructed gradient, $\widetilde{\partial_x m_{n+1}}$, from the last component of $-\frac{2n+1}{n+1} \widetilde{\mathbf{g}}_{n+1}$. Then given the boundary measurement $m_{n+1}(0, t)$ and $m_{n+1}(1, t)$, the approximation to m_{n+1} is then obtained with integration:

$$\widehat{m}_{n+1}(x, t) = m_{n+1}(0, t) + \int_0^x (\widetilde{\partial_x m_{n+1}} + c_{n+1}) dx.$$

Here c_{n+1} is the gradient correction

$$c_{n+1} = - \int_0^1 \widetilde{\partial_x m_{n+1}} dx + m_{n+1}(1, t) - m_{n+1}(0, t),$$

such that the boundary conditions, $\widehat{m}_{n+1}(0, t) = m_{n+1}(0, t)$ and $\widehat{m}_{n+1}(1, t) = m_{n+1}(1, t)$, are ensured.

Step 3. For $k = n+1, n+2, \dots, N$, suppose the nudged solutions $\widehat{m}_0(x, t), \dots, \widehat{m}_k(x, t)$ have been obtained, to get $\widehat{m}_{k+1}(x, t)$, we first compute its approximate gradient with

$$\widetilde{\partial_x m_{k+1}} = - \frac{2k+1}{k+1} I_h(\partial_t \widehat{m}_k + \frac{k}{2k+1} \partial_x \widehat{m}_{k-1} + (\sigma_a + \sigma_s) \widehat{m}_k).$$

Then we recover \widehat{m}_{k+1} through integration

$$\widehat{m}_{k+1}(x, t) = m_{k+1}(0, t) + \int_0^x (\widetilde{\partial_x m_{k+1}} + c_{k+1}) dx$$

with

$$c_{k+1} = - \int_0^1 \widetilde{\partial_x m_{k+1}} dx + m_{k+1}(1, t) - m_{k+1}(0, t).$$

The reconstructed moment satisfies the given boundary measurements: $\widehat{m}_{k+1}(0, t) = m_{k+1}(0, t)$ and $\widehat{m}_{k+1}(1, t) = m_{k+1}(1, t)$.

We now comment on the availability of the boundary data $m_k(0, t)$ and $m_k(1, t)$. While obtaining measurements of f throughout the entire interior domain can be impractical, measurements at the boundaries are more feasible. Specifically, one can impose a user-specified incoming boundary condition, i.e., $f(0, v > 0)$ and $f(1, v < 0)$, and then perform velocity-resolved boundary measurements, i.e., $f(0, v < 0)$ and $f(1, v > 0)$. This provides complete information about f at the boundaries, allowing for the calculation of moments:

$$m_k(0) = \frac{1}{2} \int_{-1}^1 f(x=0, v) P_k(v) dv, \quad m_k(1) = \frac{1}{2} \int_{-1}^1 f(x=1, v) P_k(v) dv.$$

This input-output data pair is common in inverse problem setups, and their relationship is often referred to as the Albedo operator. For more details, see [4].

It is worth mentioning that the above moment recovery procedure can be naturally extended to general 1-D moment systems that admit the structure:

$$\partial_t m_k + \sum_{i=0}^{k+1} a_{ki}(m_0, m_1, \dots, m_k) \partial_x m_i = S_k(m_0, m_1, \dots, m_k), \quad a_{k,k+1} \neq 0, \quad k = 0, 1, 2, \dots$$

For instance, we briefly mention the one-dimensional Boltzmann equation with BGK collision [6]:

$$\frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} = \frac{1}{\tau} (f_{\mathcal{M}} - f), \quad x \in [0, 1], \quad v \in \mathbb{R}.$$

Here $\tau > 0$ is the relaxation time, and $f_{\mathcal{M}} = \frac{\rho}{\sqrt{2\pi\theta}} \exp(-\frac{|v-u|^2}{2\theta})$ is the Maxwellian. We define the moments as $\mathbf{m}_N = (\rho, u, \theta, f_3, f_4, \dots, f_N)^\top \in \mathbb{R}^{N+1}$, where

$$\rho(x, t) = \int_{\mathbb{R}} f(x, v, t) dv, \quad \rho u(x, t) = \int_{\mathbb{R}} v f(x, v, t) dv, \quad \frac{1}{2} \rho \theta(x, t) = \int_{\mathbb{R}} \frac{(v-u)^2}{2} f(x, v, t) dv,$$

and f_k is the coefficient in the Hermite expansion [8, 9]:

$$f(x, v, t) = \sum_{k=0}^{\infty} f_k(x, t) \mathcal{H}_{\theta(x,t),k} \left(\frac{v-u(x,t)}{\sqrt{\theta(x,t)}} \right), \quad \mathcal{H}_{\theta,k}(v) = \frac{1}{\sqrt{2\pi}} \theta^{-\frac{k+1}{2}} He_k(v) \exp(-\frac{v^2}{2}),$$

where He_k being the k -th Hermite polynomial. In particular, it can be shown that $f_0 = \rho$ and $f_1 = f_2 = 0$. Inspired by [7], the corresponding moments system are given by

$$\frac{\partial \mathbf{m}_N}{\partial t} + A_N \frac{\partial \mathbf{m}_N}{\partial x} = -(N+1) \frac{\partial f_{N+1}}{\partial x} \mathbf{e}_{N+1} - \frac{1}{\tau} P_N \mathbf{m}_N,$$

where

$$A_N = \begin{bmatrix} u & \rho & 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \frac{\theta}{\rho} & u & 1 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 2\theta & u & \frac{\theta}{\rho} & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 4f_3 & \frac{\rho\theta}{2} & u & 4 & 0 & \cdots & \cdots & \cdots & 0 \\ -\frac{\theta f_3}{\rho} & 5f_4 & \frac{3f_3}{2} & \theta & u & 5 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -\frac{\theta f_{N-2}}{\rho} & Nf_{N-1} & \frac{1}{2}[(N-2)f_{N-2} + \theta f_{N-4}] & -\frac{3f_{N-3}}{\rho} & 0 & \cdots & 0 & \theta & u & N \\ -\frac{\theta f_{N-1}}{\rho} & (N+1)f_N & \frac{1}{2}[(N-1)f_{N-1} + \theta f_{N-3}] & -\frac{3f_{N-2}}{\rho} & 0 & \cdots & \cdots & 0 & \theta & u \end{bmatrix},$$

and

$$P_N = \text{diag}(0, 0, 0, 1, \cdots, 1).$$

Despite the non-constant coefficients, A_N remains an unreduced lower Hessenberg matrix for a physical density, $\rho > 0$, thereby allowing the recovery of higher moments through integration.

4.3. Numerical experiments II – moment recovery.

4.3.1. One-dimensional RTE. We first apply our moment recovery algorithm to the moment system (4.2) derived from the one-dimensional RTE. The reference solution is generated by solving the equation (4.1) using the WENO5-SSPRK3 with initial data

$$f_0(x, v) = 0.5 + \sum_{k=1}^5 \frac{1}{k^2} \sin(2k\pi x), \quad x \in [0, 1].$$

Periodic boundary conditions are applied at both endpoints. The number of grid points in space is set to $N = 600$. The velocity space is discretized with 15 Gauss-Legendre quadrature points. The time step is taken to $\Delta t = 0.5\Delta x$ and the terminal time is $t = 1$.

We consider the data under different absorption coefficients σ_a and σ_s . In all test cases, the nudged equations are discretized with WENO5-SSPRK3 under $N = 300$ mesh points. The nudging coefficient is set to $\mu = 6$. Complete observations are assumed to be available at both boundaries. We intend to recover the higher-order moments based on the observed data of m_0 and m_1 over $N_{ob} = 60$ uniform observation grids at the interior. The approximate moments are initialized with $\hat{m}_0 = 0.5$, $\hat{m}_2 = 0$, $\hat{m}_3 = 0$, $\hat{m}_4 = 0$, $\hat{m}_5 = 0$.

Test 1 ($\sigma_a = \sigma_s = 1$). We recover the moments up to the fifth order in the optically-thin regime with small absorption coefficients $\sigma_a = \sigma_s = 1$. The nudged solutions \hat{m}_k at terminal time are displayed in Figure 4.1. It is observed that all the recovered moments match the exact data accurately. The history of moment errors, as shown in Figure 4.2, verifies the convergence of the algorithm, where the relatively large error at early times is primarily due to the intentionally inaccurate initialization used in our test cases. This setup was chosen to demonstrate the robustness of the nudging-based feedback control, particularly its ability to recover accurate solutions even from poor initial guesses.

Test 2 ($\sigma_a = \sigma_s = 10$). Now we consider the data under a larger optical depth. Due to the stronger dissipation introduced by the source term, all the moments admit overall smaller scales, which can be observed from Figure 4.3. It is noticed that the algorithm generates accurate approximations to the first four moments, whereas the deviations in m_4 and m_5 appear significantly larger. We argue that the seemingly large errors are due to the excessively small moment magnitudes ($10^{-8} \sim 10^{-7}$), in which case the accumulated discretization errors from lower moments recovery might become dominating. In fact, the history of errors (Figure 4.4) indicates the stable convergence of the approximate moments toward the exact solutions.

We also mention that as reported in [23], the closure relation is relatively easy to capture in the intermediate and optically thick regimes, and the model (4.3) can achieve a good accuracy by only using m_0 , m_1 , and m_2 , which are accurately recovered in the computations. Hence, despite the difficulty in recovering high-order moments with extremely small magnitudes, we can still expect our algorithm to be effective when combined with actual machine-learning closure models.

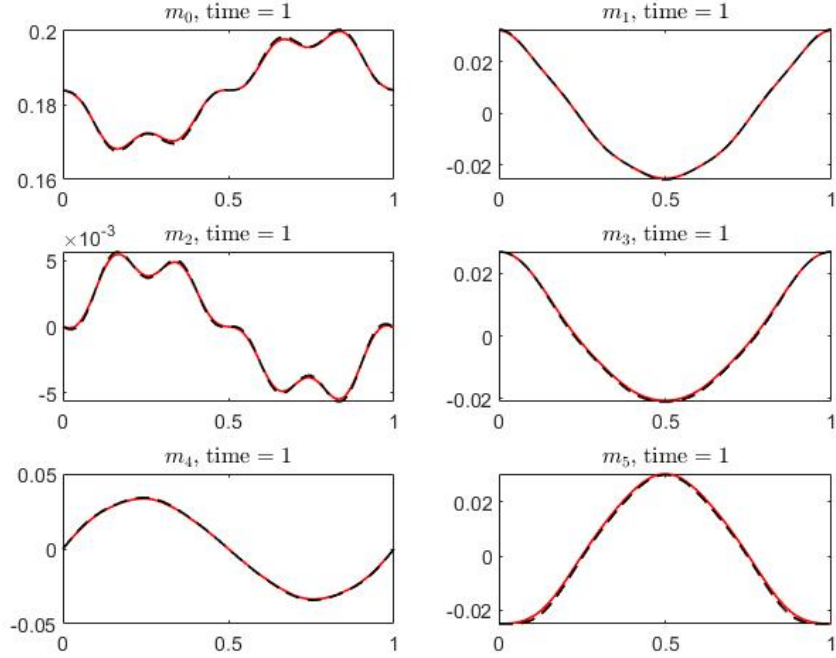
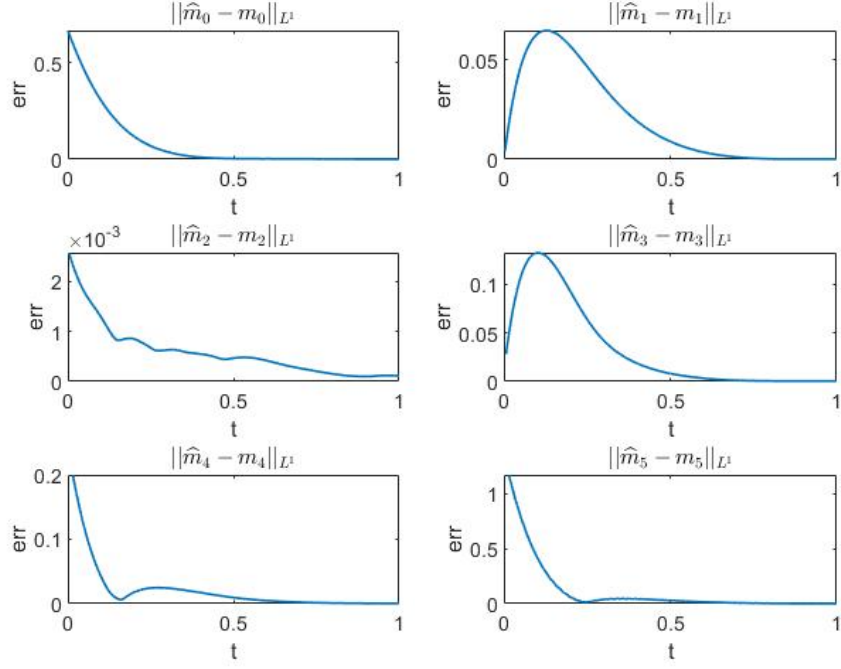


FIGURE 4.1. Moment recovery for 1D RTE. Nudged solutions (red lines) vs. reference data (black dash lines). $\sigma_a = \sigma_s = 1$. $t = 1$.

4.3.2. Boltzmann equation for channel flow. In addition to the one-dimensional radiation transport, we also test our algorithm on particle flow through a closed channel. The flow is governed by a reduced Boltzmann equation with one dimension

FIGURE 4.2. 1D RTE. History of moment errors. $\sigma_a = \sigma_s = 1$.

in space and two dimensions in velocity:

$$(4.7) \quad \frac{\partial f}{\partial t} + v_1 \partial_x f = \iint_{\mathbb{R}^2 \times \mathbb{S}^1} B(\mathbf{v} - \mathbf{v}_*, \sigma) (f' f'_* - f f_*) d\mathbf{v}_* d\sigma, \quad x \in \mathbb{R}, \quad \mathbf{v} = (v_1, v_2) \in \mathbb{R}^2,$$

where

$$f = f(x, \mathbf{v}, t), \quad f_* = f(x, \mathbf{v}_*, t), \quad f' = f(x, \mathbf{v}', t), \quad f'_* = f(x, \mathbf{v}'_*, t),$$

$$\mathbf{v}' = \frac{\mathbf{v} + \mathbf{v}_*}{2} + \frac{|\mathbf{v} - \mathbf{v}_*|}{2} \sigma, \quad \mathbf{v}'_* = \frac{\mathbf{v} + \mathbf{v}_*}{2} - \frac{|\mathbf{v} - \mathbf{v}_*|}{2} \sigma.$$

The collision kernel $B(\mathbf{v} - \mathbf{v}_*, \sigma)$ measures the intensity of particle collisions. In the case of the inverse power law, B can be separated as the kinetic part Φ and angular part b ,

$$B(\mathbf{v} - \mathbf{v}_*, \sigma) = b(\cos \theta) \Phi(|\mathbf{v} - \mathbf{v}_*|), \quad \text{with } \cos \theta = \sigma \cdot \frac{\mathbf{v} - \mathbf{v}_*}{|\mathbf{v} - \mathbf{v}_*|},$$

where kinetic collision part $\Phi(|\mathbf{v} - \mathbf{v}_*|) = |\mathbf{v} - \mathbf{v}_*|^\gamma$ includes hard potential ($\gamma > 0$), Maxwellian molecule ($\gamma = 0$) and soft potential ($\gamma < 0$), and angular part $b(\cos \theta)$ is often regarded to satisfy the Grad's cutoff assumption, i.e., $\int_{\mathbb{S}^2} b(\cos \theta) d\sigma < \infty$. For more details on the collision kernel, see [39]. For convenience of discussion, we consider the Maxwellian molecule with constant collision kernel $B = 0.2$. The

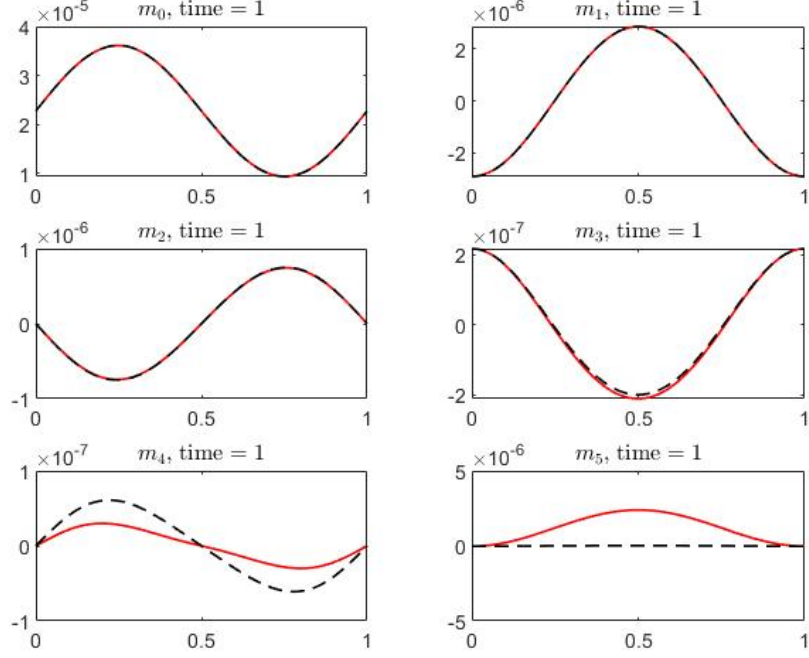


FIGURE 4.3. Moment recovery for 1D RTE. Nudged solutions (red lines) vs. reference data (black dash lines). $\sigma_a = \sigma_s = 10$. $t = 1$.

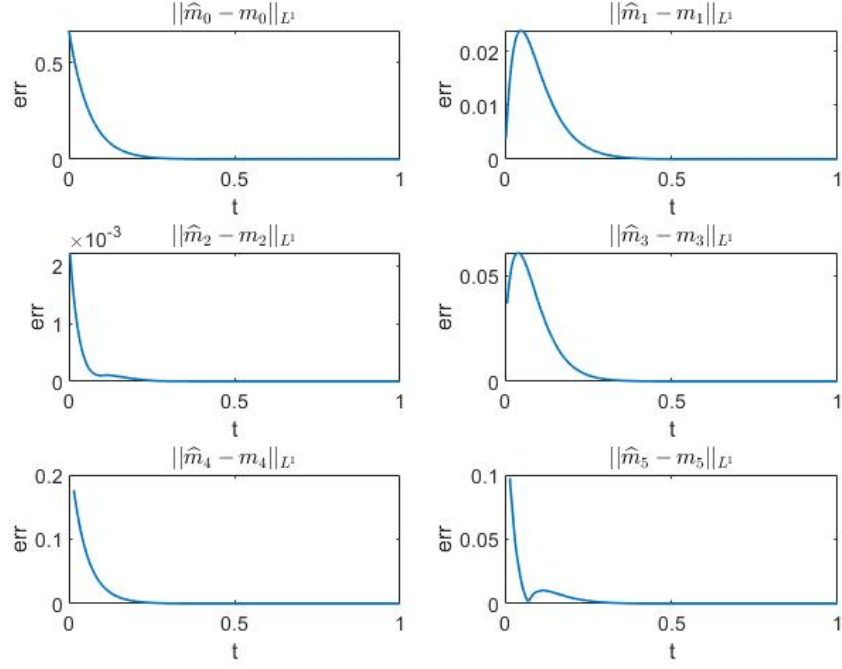
initial condition is set to:

$$f_0(x, v_1, v_2) = \rho(x)f_1(v_1)f_2(v_2), \quad x \in [0, 1], \quad (v_1, v_2) \in \mathbb{R}^2,$$

with

$$\begin{aligned} \rho(x) &= 1 + 0.2 \sin(2\pi x), \\ f_1(v_1) &= \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{(v_1 - 1)^2}{2\sigma_1^2}\right), \quad v_1 \in \mathbb{R}, \quad \sigma_1 = 0.5, \\ f_2(v_2) &= \frac{1}{\sigma_2 \sqrt{2\pi}} \exp\left(-\frac{(v_2 - 0.1)^2}{2\sigma_2^2}\right), \quad v_2 \in \mathbb{R}, \quad \sigma_2 = 0.05. \end{aligned}$$

Periodic boundary conditions are applied at both endpoints of the space interval. To generate reference data, the transport part is solved with WENO5-SSPRK3, while the collision part is calculated with the fast spectral method [35, 21, 22].

FIGURE 4.4. 1D RTE. History of moment errors. $\sigma_a = \sigma_s = 10$.

We define the moments of particle distribution f up to the third order,

$$\begin{aligned}
 \rho(x, t) &:= \int_{\mathbb{R}^2} f(x, \mathbf{v}, t) d\mathbf{v}, \\
 \rho \mathbf{u}(x, t) &:= \int_{\mathbb{R}^2} \mathbf{v} f(x, \mathbf{v}, t) d\mathbf{v}, \quad \mathbf{u} = (u, v), \\
 \mathbb{P}(x, t) &:= \int_{\mathbb{R}^2} (\mathbf{v} - \mathbf{u}) \otimes (\mathbf{v} - \mathbf{u}) f(x, \mathbf{v}, t) d\mathbf{v} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}, \quad p_{12} = p_{21}, \\
 \rho e(x, t) &:= \int_{\mathbb{R}^2} \frac{|\mathbf{v} - \mathbf{u}|^2}{2} f(x, \mathbf{v}, t) d\mathbf{v} = \frac{1}{2} \text{trace}(\mathbb{P}), \\
 \mathbf{q}(x, t) &:= \int_{\mathbb{R}^2} \frac{|\mathbf{v} - \mathbf{u}|^2}{2} (\mathbf{v} - \mathbf{u}) f(x, \mathbf{v}, t) d\mathbf{v} = [q_1, q_2]^\top.
 \end{aligned}$$

Integrating the equation (4.7) against 1, \mathbf{v} , and $\frac{|\mathbf{v}|^2}{2}$ yields the moment equations

$$(4.8) \quad \begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2) + \partial_x p_{11} = 0, \\ \partial_t(\rho v) + \partial_x(\rho uv) + \partial_x p_{12} = 0, \\ \partial_t(\rho e) + \partial_x(\rho eu) + \partial_x q_1 + p_{11} \partial_x u + p_{12} \partial_x v = 0. \end{cases}$$

Note that the internal energy is given by $\rho e = \frac{1}{2}(p_{11} + p_{22})$, while the evolution of p_{22} is not described by the moment equations (4.8). Nevertheless, due to the small variance of f in y -velocity, the contribution of p_{22} can be neglected in our test case, which yields $\rho e = \frac{1}{2}p_{11}$. Hence, the last equation in (4.8) reduces to

$$\partial_t p_{11} + \partial_x(p_{11}u) + 2\partial_x q_1 + 2p_{11}\partial_x u + 2p_{12}\partial_x v = 0.$$

Assume the complete observations of all moments are available at the boundaries. At the interior of the solution domain, we employ the partial observations with ρ , ρu , and ρv over $N_{ob} = 60$ uniform observation grids. To recover the continuous data fields asymptotically in time, the solution procedures are briefly described as follows: at first, following the constructions of the nudged system (1.4), we reconstruct the approximate moments $\widehat{\rho}$, $\widehat{\rho u}$, $\widehat{\rho v}$, \widehat{p}_{11} , \widehat{p}_{12} under $\mu = 10$. Then the approximate gradient of q_1 is obtained by

$$\widetilde{\partial_x q_1} = -\frac{1}{2}I_h(\partial_t \widehat{p}_{11} + \partial_x(\widehat{p}_{11}\widehat{u}) + 2\widehat{p}_{11}\partial_x \widehat{u} + 2\widehat{p}_{12}\partial_x \widehat{v}), \quad \widehat{u} = \frac{\widehat{\rho u}}{\widehat{\rho}}, \quad \widehat{v} = \frac{\widehat{\rho v}}{\widehat{\rho}}.$$

Subsequently, the reconstruction of the third-order moment, \widehat{q}_1 , is given by

$$\widehat{q}_1 = q_1(0) + \int_0^x (\widetilde{\partial_x q_1} + c_q) dx, \quad c_q = -\int_0^1 \widetilde{\partial_x q_1} dx + q_1(1) - q_1(0).$$

The nudged equations are initialized with

$$\widehat{\rho}_0(x) = 1, \quad \widehat{u}_0(x) = 1, \quad \widehat{v}_0(x) = 0.1, \quad \widehat{\mathbb{P}}_0(x) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

and discretized with WENO5-SSPRK3 under $N = 300$ space mesh points. Figure 4.5 displays the computational results at $t = 0.7$. The data of ρ , ρu , ρv , p_{11} and q_1 , which have major contributions to the system, are recovered with satisfactory accuracy. The seemingly large error of \widehat{p}_{12} is, again, due to the extremely small magnitude of the exact p_{12} ($\approx 10^{-9}$), hence the influence on the overall dynamics is minimal.

5. EXTENSION TO HIGHER DIMENSION

In the preceding discussions, we have demonstrated the efficacy of our moment recovery algorithm in one-dimensional space. Essentially, leveraging the sparsely observed data of lower-order moments enables the retrieval of gradients (variations) of higher-order moments at the interior of the domain. In addition, when complemented with sufficient ground truth data at the boundaries, the integration of the moment equations facilitates the recovery of all relevant moments.

However, extending the aforementioned procedure to recover higher-order moments in multiple dimensions is highly nontrivial. This complexity arises because the divergence term now involves a mixture of directional derivatives, which can lead to non-unique recovery of the quantity before differentiation. To illustrate this challenge and suggest a possible approach for addressing it, we

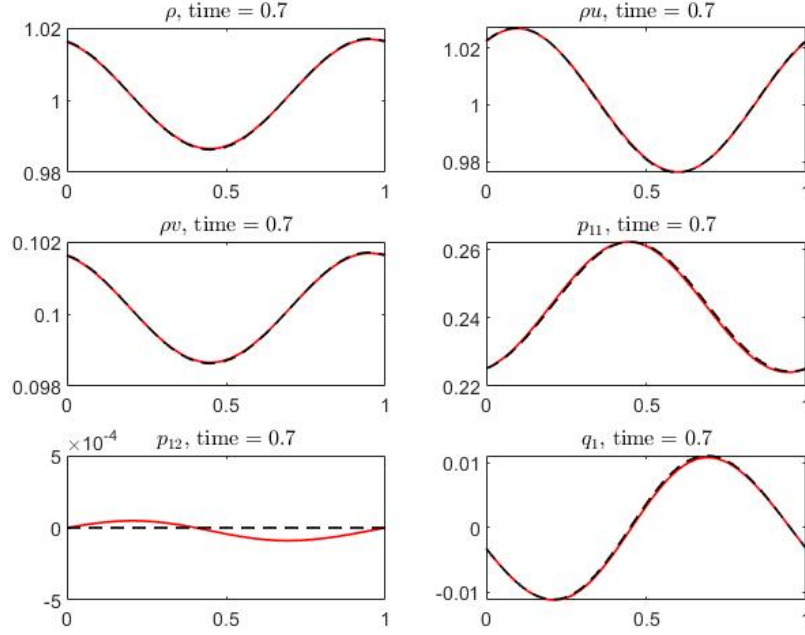


FIGURE 4.5. Moment recovery for Boltzmann. Nudged solutions (red lines) vs. reference data (black dash lines). $t = 0.7$.

briefly discuss the two-dimensional Eulerian system:

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho \mathbf{u}) = 0, \\ \partial_t \rho \mathbf{u} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla \cdot \mathbb{P} = 0, \\ \partial_t \rho e + \nabla \cdot (\rho e \mathbf{u}) + \nabla \cdot \mathbf{q} + \text{trace}(\mathbb{P} \nabla \mathbf{u}) = 0, \end{cases} \quad \mathbf{x} \in \Omega \in \mathbb{R}^2, \quad t \geq 0.$$

Assume the exact data of all solution variables are accessible at the boundary $\partial\Omega$, whereas the (sparse) interior observations are only available for ρ and $\rho \mathbf{u}$. We aim to recover the data of the pressure tensor, $\mathbb{P} \in \mathbb{R}^{2 \times 2}$, and subsequently the divergence of heat flux, $\nabla \cdot \mathbf{q}$, using the continuous data assimilation. By evolving the corresponding nudged system along the lines of (1.4), one recovers the *divergence* of the pressure, $g = -\nabla \cdot \mathbb{P}$, which contains the ‘mixture’ of x - and y -derivatives. Hence, the recovery of \mathbb{P} requires solving the divergence equations:

$$\begin{cases} -(\partial_x p_{11} + \partial_y p_{21}) = g_1 \\ -(\partial_x p_{12} + \partial_y p_{22}) = g_2 \end{cases}, \quad p_{12} = p_{21},$$

which, in general, are under-determined.

Indeed, denote $\mathbb{P} = [\mathbb{P}_1, \mathbb{P}_2]$ with $\mathbb{P}_j = [p_{1j}, p_{2j}]^\top$, $j = 1, 2$, the pressure term can be represented via the Helmholtz decomposition:

$$(5.1) \quad \mathbb{P}_j = \nabla \phi_j + \mathbf{R}_j, \quad \mathbf{R}_j = [R_{1j}, R_{2j}]^\top,$$

where

$$(5.2a) \quad \begin{cases} -\Delta\phi_j = g_j, & \mathbf{x} \in \Omega, \\ \phi_j = 0, & \mathbf{x} \in \partial\Omega; \end{cases}$$

and

$$(5.2b) \quad \begin{cases} \nabla \cdot \mathbf{R}_j = 0, & \mathbf{x} \in \Omega, \\ \mathbf{R}_j \cdot \mathbf{n} = \mathbb{P}_j \cdot \mathbf{n} - \nabla\phi_j \cdot \mathbf{n}, & \mathbf{x} \in \partial\Omega. \end{cases}$$

Here $\mathbf{n} = [n_1, n_2]^\top$ is the unit outer normal vector at the boundary. Moreover, the symmetry condition $p_{12} = p_{21}$ yields

$$(5.3) \quad R_{12} = R_{21} + \partial_y\phi_1 - \partial_x\phi_2.$$

Note that the Poisson equations (5.2a) for the potentials ϕ_1 and ϕ_2 are well-posed. Therefore, the non-uniqueness of the pressure arises from the non-uniqueness of the solenoidal vector fields \mathbf{R}_j .

For practical applications, a possible option is to generate approximate solutions with the physics-informed neural networks (PINNs) [38]. Due to the simple linear structure of the divergence equations, the PINNs can be expected to work effectively. As an example, consider the exact pressure field \mathbb{P} with

$$\begin{cases} p_{11} = 1 + 0.5 \sin(x) \cos(y), \\ p_{12} = p_{21} = 0.2 \sin(x + y), \\ p_{22} = 1 + 0.3e^{-0.1y} \cos(x) \sin(y), \end{cases} \quad (x, y) \in \Omega = [0, 2\pi]^2.$$

Based on the divergence at the interior, $\nabla \cdot \mathbb{P}_1 = \partial_x p_{11} + \partial_y p_{12}$ and $\nabla \cdot \mathbb{P}_2 = \partial_x p_{12} + \partial_y p_{22}$, and the ground truth boundary data, $\mathbb{P}|_{\partial\Omega}$, we reconstruct the pressure using the multi-outputs neural network,

$$\mathcal{N}(x, y; \theta) = [p_{11}^{NN}, p_{12}^{NN}, p_{22}^{NN}]^\top,$$

where $\mathcal{N}(\cdot; \theta)$ is a fully connected network with 3 hidden layers parameterized by θ . Each hidden layer has 20 neurons and the output is activated by the hyperbolic tangent function. The parameters of the neural network are optimized by minimizing the mean square loss,

$$\begin{aligned} MSE(\theta) = & \frac{1}{N} \sum_{j=1}^N |\partial_x p_{11}^{NN}(x_j, y_j) + \partial_y p_{12}^{NN}(x_j, y_j) - \nabla \cdot \mathbb{P}_1(x_j, y_j)|^2 \\ & + \frac{1}{N} \sum_{j=1}^N |\partial_x p_{12}^{NN}(x_j, y_j) + \partial_y p_{22}^{NN}(x_j, y_j) - \nabla \cdot \mathbb{P}_2(x_j, y_j)|^2 \\ & + \frac{1}{N_b} \sum_{m,n=1}^2 \sum_{j=1}^{N_b} |p_{mn}^{NN}(x_j^b, y_j^b) - p_{mn}(x_j^b, y_j^b)|^2, \end{aligned}$$

where $\{(x_j, y_j)\}_{j=1}^N$ and $\{(x_j^b, y_j^b)\}_{j=1}^{N_b}$ are training points at the interior and the boundary. In our computation, we apply 30×30 uniformly distributed training points at the interior. At each of the four boundaries, 30 training points are employed to enforce the boundary conditions. We train the network using LBFGS in Pytorch until the relative change in the loss is smaller than 10^{-4} . As observed

from the results displayed in Figure 5.1, the neural network solutions attain a reasonably good accuracy. The relatively errors are $\frac{\|p_{11}^{NN} - p_{11}\|_1}{\|p_{11}\|_1} = 1.54 \times 10^{-3}$, $\frac{\|p_{12}^{NN} - p_{12}\|_1}{\|p_{12}\|_1} = 7.05 \times 10^{-3}$, $\frac{\|p_{11}^{NN} - p_{11}\|_1}{\|p_{11}\|_1} = 1.51 \times 10^{-3}$.

In general, to resolve the potential non-uniqueness issue, a regularization term can be added that favors the solution with special property. For instance, the solution with the smallest oscillation is favored by adding a penalty term $\lambda \|\nabla \mathbb{P}\|_2^2$, where $\lambda > 0$ is a shrinking regularization parameter; by introducing the penalty $\lambda \|\mathbf{R}\|_2^2$, where $\mathbf{R} = [\mathbf{R}_1; \mathbf{R}_2]$ are the divergence-free vector fields in the Helmholtz decomposition (5.1), one minimizes the projection of pressure onto $\text{Im}(\nabla)^\perp = \ker(\nabla \cdot) = \{\mathbb{V} \in \mathbb{R}^{2 \times 2} : \nabla \cdot \mathbb{V} = 0, \mathbb{V} \cdot \mathbf{n}|_{\partial\Omega} = \mathbb{P} \cdot \mathbf{n}|_\Omega\}$.

Although using the PINNs provides a possible solution approach, one should be aware of the caveat in terms of the expensive computational cost. Further investigations are needed for more efficient data reconstruction from divergence information.

REFERENCES

1. Débora AF Albanez and Maicon J Benvenuti, *Continuous data assimilation algorithm for simplified bardina model*, Evolution Equations and Control Theory **7** (2018), no. 1, 33–52.
2. M. U. Altaf, E. S. Titi, T. Gebrael, O. M. Knio, L. Zhao, M. F. McCabe, and I. Hoteit, *Downscaling the 2d b nard convection equations using continuous data assimilation*, Computational Geosciences **21** (2017), 393–410.
3. Abderrahim Azouani, Eric Olson, and Edriss S Titi, *Continuous data assimilation using general interpolant observables*, Journal of Nonlinear Science **24** (2014), 277–304.
4. Guillaume Bal, *Inverse transport theory and applications*, Inverse Problems **25** (2009), no. 5, 053001.
5. Claude Bardos, Fran ois Golse, and C David Levermore, *Fluid dynamic limits of kinetic equations ii convergence proofs for the boltzmann equation*, Communications on pure and applied mathematics **46** (1993), no. 5, 667–753.
6. Prabhu Lal Bhatnagar, Eugene P Gross, and Max Krook, *A model for collision processes in gases. i. small amplitude processes in charged and neutral one-component systems*, Physical review **94** (1954), no. 3, 511.
7. Zhenning Cai, Yuwei Fan, and Ruo Li, *Globally hyperbolic regularization of grad’s moment system in one-dimensional space*, Communications in Mathematical Sciences **11** (2013), no. 2, 547–571.
8. Zhenning Cai and Ruo Li, *Numerical regularized moment method of arbitrary order for boltzmann-bgk equation*, SIAM Journal on Scientific Computing **32** (2010), no. 5, 2875–2907.
9. Zhenning Cai, Ruo Li, and Yanli Wang, *Numerical regularized moment method for high mach number flow*, Communications in Computational Physics **11** (2012), no. 5, 1415–1438.
10. Elizabeth Carlson, Joshua Hudson, and Adam Larios, *Parameter recovery for the 2 dimensional navier–stokes equations via continuous data assimilation*, SIAM Journal on Scientific Computing **42** (2020), no. 1, A250–A270.
11. Elizabeth Carlson, Joshua Hudson, Adam Larios, Vincent R. Martinez, Eunice Ng, and Jared P. Whitehead, *Dynamically learning the parameters of a chaotic system using partial observations*, Discrete and Continuous Dynamical Systems. Series A **42** (2022), no. 8, 3809–3839. MR 4447559
12. James Carpenter, Peter Clifford, and Paul Fearnhead, *Improved particle filter for nonlinear problems*, IEE Proceedings-Radar, Sonar and Navigation **146** (1999), no. 1, 2–7.
13. Patricio Clark Di Leoni, Andrea Mazzino, and Luca Biferale, *Inferring flow parameters and turbulent configuration with physics-informed data assimilation and spectral nudging*, Physical Review Fluids **3** (2018), no. 10, 104604.

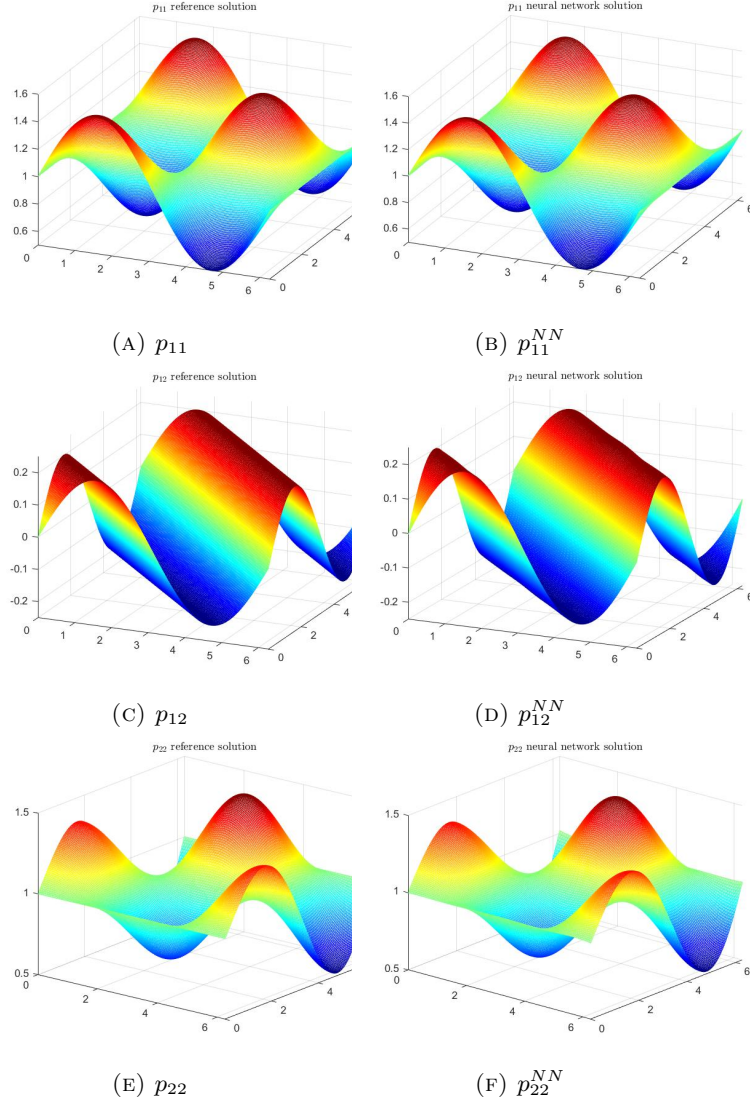


FIGURE 5.1. Pressure recovery from divergence. Neural network solutions vs. reference solutions.

14. Aseel Farhat, Nathan E Glatt-Holtz, Vincent R Martinez, Shane A McQuarrie, and Jared P Whitehead, *Data assimilation in large prandtl rayleigh-benard convection from thermal measurements*, SIAM Journal on Applied Dynamical Systems **19** (2020), no. 1, 510–540.
15. Aseel Farhat, Michael S Jolly, and Edriss S Titi, *Continuous data assimilation for the 2d benard convection through velocity measurements alone*, Physica D: Nonlinear Phenomena **303** (2015), 59–66.
16. Aseel Farhat, Adam Larios, Vincent R Martinez, and Jared P Whitehead, *Identifying the body force from partial observations of a two-dimensional incompressible velocity field*, Physical Review Fluids **9** (2024), no. 5, 054602.

17. Aseel Farhat, Evelyn Lunasin, and Edriss S Titi, *Data assimilation algorithm for 3d b  nard convection in porous media employing only temperature measurements*, Journal of Mathematical Analysis and Applications **438** (2016), no. 1, 492–506.
18. Claude Fischer, Thibaut Montmerle, Lo  k Berre, Ludovic Auger, and Simona Ecaterina   tef  nescu, *An overview of the variational assimilation in the aladin/france numerical weather-prediction system*, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography **131** (2005), no. 613, 3477–3492.
19. Masakazu Gesho, Eric Olson, and Edriss S Titi, *A computational study of a data assimilation algorithm for the two-dimensional navier-stokes equations*, Communications in Computational Physics **19** (2016), no. 4, 1094–1110.
20. Sigal Gottlieb, Chi-Wang Shu, and Eitan Tadmor, *Strong stability-preserving high-order time discretization methods*, SIAM review **43** (2001), no. 1, 89–112.
21. Jingwei Hu and Kunlun Qi, *A fast Fourier spectral method for the homogeneous Boltzmann equation with non-cutoff collision kernels*, J. Comput. Phys. **423** (2020), 109806, 21. MR 4156938
22. Jingwei Hu, Kunlun Qi, and Tong Yang, *A new stability and convergence proof of the Fourier-Galerkin spectral method for the spatially homogeneous Boltzmann equation*, SIAM J. Numer. Anal. **59** (2021), no. 2, 613–633. MR 4226997
23. Juntao Huang, Yingda Cheng, Andrew J Christlieb, and Luke F Roberts, *Machine learning moment closure models for the radiative transfer equation i: directly learning a gradient based closure*, Journal of Computational Physics **453** (2022), 110941.
24. ———, *Machine learning moment closure models for the radiative transfer equation iii: enforcing hyperbolicity and physical characteristic speeds*, Journal of Scientific Computing **94** (2023), no. 1, 7.
25. Juntao Huang, Yingda Cheng, Andrew J Christlieb, Luke F Roberts, and Wen-An Yong, *Machine learning moment closure models for the radiative transfer equation ii: Enforcing global hyperbolicity in gradient-based closures*, Multiscale Modeling & Simulation **21** (2023), no. 2, 489–512.
26. Michael S Jolly, Vincent R Martinez, Eric J Olson, and Edriss S Titi, *Continuous data assimilation with blurred-in-time measurements of the surface quasi-geostrophic equation*, Chinese Annals of Mathematics, Series B **40** (2019), no. 5, 721–764.
27. Michael S Jolly, Vincent R Martinez, and Edriss S Titi, *A data assimilation algorithm for the subcritical surface quasi-geostrophic equation*, Advanced Nonlinear Studies **17** (2017), no. 1, 167–192.
28. Michael S Jolly, Tural Sadigov, and Edriss S Titi, *A determining form for the damped driven nonlinear schr  dinger equation—fourier modes case*, Journal of Differential Equations **258** (2015), no. 8, 2711–2744.
29. ———, *Determining form and data assimilation algorithm for weakly damped and driven korteweg–de vries equation—fourier modes case*, Nonlinear Analysis: Real World Applications **36** (2017), 287–317.
30. Rudolph Emil Kalman, *A new approach to linear filtering and prediction problems*, (1960).
31. Genshiro Kitagawa, *Monte carlo filter and smoother for non-gaussian nonlinear state space models*, Journal of computational and graphical statistics **5** (1996), no. 1, 1–25.
32. Fran  ois-Xavier Le Dimet and Olivier Talagrand, *Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects*, Tellus A: Dynamic Meteorology and Oceanography **38** (1986), no. 2, 97–110.
33. Vincent R Martinez, *Convergence analysis of a viscosity parameter recovery algorithm for the 2d navier–stokes equations*, Nonlinearity **35** (2022), no. 5, 2241.
34. Vincent R. Martinez, *On the reconstruction of unknown driving forces from low-mode observations in the 2d navier–stokes equations*, Proceedings of the Royal Society of Edinburgh: Section A Mathematics (2024), 1–24.
35. Cl  ment Mouhot and Lorenzo Pareschi, *Fast algorithms for computing the boltzmann collision operator*, Mathematics of computation **75** (2006), no. 256, 1833–1852.
36. Benjamin Pachev, Jared P Whitehead, and Shane A McQuarrie, *Concurrent multiparameter learning demonstrated on the kuramoto–sivashinsky equation*, SIAM Journal on Scientific Computing **44** (2022), no. 5, A2974–A2990.

37. Yuan Pei, *Continuous data assimilation for the 3d primitive equations of the ocean*, arXiv preprint arXiv:1805.06007 (2018), 1–25.
38. Maziar Raissi, Paris Perdikaris, and George E Karniadakis, *Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*, Journal of Computational physics **378** (2019), 686–707.
39. Cédric Villani, *A review of mathematical topics in collisional kinetic theory*, Handbook of Mathematical Fluid Mechanics (S. Friedlander and D. Serre, eds.), vol. I, North-Holland, 2002, pp. 71–305.
40. Greg Welch and Gary Bishop, *An introduction to the kalman filter*, (1995).
41. Dušanka Županski and Fedor Mesinger, *Four-dimensional variational assimilation of precipitation data*, Monthly Weather Review **123** (1995), no. 4, 1112–1127.

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA–TWIN CITIES, MINNEAPOLIS,
MINNESOTA 55455

Email address: lu000688@umn.edu

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA–TWIN CITIES, MINNEAPOLIS,
MINNESOTA 55455

Email address: kqi@umn.edu, kunlunqi.math@gmail.com

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA–TWIN CITIES, MINNEAPOLIS,
MINNESOTA 55455

Email address: liwang@umn.edu

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA–TWIN CITIES, MINNEAPOLIS,
MINNESOTA 55455

Email address: jwcalder@umn.edu