Variational formulation based on duality to solve partial differential equations: Use of B-splines and machine learning approximants

N. Sukumar[a,*], Amit Acharya[b]

[a]*Department of Civil and Environmental Engineering, One Shields Avenue, University of California, Davis, CA 95616, USA*
[b]*Department of Civil and Environmental Engineering, and Center for Nonlinear Analysis, Carnegie Mellon University, Pittsburgh, PA 15213, USA*

Abstract

Many partial differential equations (PDEs) such as Navier–Stokes equations in fluid mechanics, inelastic deformation in solids, and transient parabolic and hyperbolic equations do not have an exact, primal variational structure. Recently, a variational principle based on the dual (Lagrange multiplier) field was proposed. The essential idea in this approach is to treat the given PDE as constraints, and to invoke an arbitrarily chosen auxiliary potential with strong convexity properties to be optimized. This leads to requiring a convex dual functional to be minimized subject to Dirichlet boundary conditions on dual variables, with the guarantee that even PDEs that do not possess a variational structure in primal form can be solved via a variational principle. The vanishing of the first variation of the dual functional is, up to Dirichlet boundary conditions on dual fields, the weak form of the primal PDE problem with the dual-to-primal change of variables incorporated. We derive the dual weak form for the linear, one-dimensional, transient convection-diffusion equation. A Galerkin discretization is used to obtain the discrete equations, with the trial and test functions chosen as linear combination of either RePU activation functions (shallow neural network) or B-spline basis functions; the corresponding stiffness matrix is symmetric. For transient problems, a space-time Galerkin implementation is used with tensor-product B-splines as approximating functions. Numerical results are presented for the steady-state and transient convection-diffusion equation, and transient heat conduction. The proposed method delivers sound accuracy for ODEs and PDEs and rates of convergence are established in the L^2 norm and H^1 seminorm for the steady-state convection-diffusion problem.

Keywords: dual variational principles, convex duality, weak formulation, space-time Galerkin method, B-splines, RePU neural networks

1. Introduction

Many classes of ordinary differential equations (ODEs) and partial differential equations (PDEs) do not possess a natural, exact, variational structure:[1] convection-diffusion and Navier–Stokes equations in fluid mechanics (cf. [1, 2]),

*Corresponding author
Email addresses: nsukumar@ucdavis.edu (N. Sukumar), acharyaamit@cmu.edu (Amit Acharya)

[1] In the sense of being derived, along with all associated boundary and initial conditions, as Euler–Lagrange equations of some functional of fields defined on space-(time) without approximation.

inelastic deformation of solids (cf. [3–6]), and time-dependent parabolic (heat) and hyperbolic (wave) [7] problems, to name just a few; an insightful discussion of the issues involved in the context of the equations of continuum mechanics and classical field theories can be found in [8, Sec. 1]. For such problems, it is desirable to provide a variational principle that can lead to alternate solution strategies with the finite element method, meshfree methods, virtual element methods, B-spline approximations, and neural networks. Recently, a scheme for generating dual variational principles for such problems was proposed in [9] which is cast in terms of dual (Lagrange multiplier) fields. In [9–14], that scheme has been placed within the broader context of the prior efforts for developing the variational principles mentioned above, including the method of least squares. There exists a superficial similarity between the dual formulation to solve differential-algebraic systems and the adjoint method [15, 16], the latter used in constrained optimization to compute the derivatives of a function with respect to unknown parameters. Their similarities and differences are noted in Appendix A.

The methodology developed and demonstrated in [9–14, 17, 18] applies to quasistatic and dynamic, conservative and dissipative PDEs that arise in continuum mechanics. In a notable prior work [8], a strategy for deriving variational principles for a limited set[2] of PDE systems from continuum mechanics was devised; the approach was successfully applied to nonlinear elastodynamics in the spatial setting, as well as the compressible Euler equations, Maxwell's equations, and those of the collisionless plasma. The essential idea, as introduced in [9], is to treat the given PDE as constraints for an arbitrarily chosen auxiliary potential with strong convexity properties that is optimized. The Lagrange multipliers associated with the constraints are referred to as the dual fields. On requiring the vanishing of the gradient of the Lagrangian with respect to the primal variables, a dual-to-primal (DtP) mapping from the dual to the primal fields is obtained. This leads to requiring a convex dual functional to be minimized subject to Dirichlet boundary conditions on dual variables, with the guarantee that even PDEs that do not possess a variational structure can be solved as Euler–Lagrange equations of the dual functional. The first variation (set to zero) of the dual functional is, up to Dirichlet boundary conditions on dual fields, the weak form of the primal PDE problem with the dual-to-primal change of variables incorporated. The Euler–Lagrange equations of the dual functional are shown to be locally degenerate elliptic, regardless of the properties of the primal system [11] (for elliptic primal problems, the dual problem is elliptic). Our ideas are related to those of Brenier [19, 20], and generalizes [20] in various ways [13].

When posed in primal form, the weak formulation for convection-diffusion problems and incompressible Navier–Stokes equations require special treatment via the use of upwinding and/or stabilization schemes. The streamline upwind/Petrov–Galerkin (SUPG) finite element method produces a variationally consistent scheme that delivers stable and accurate solutions [21]. In the Galerkin/least squares (GLS) method [22, 23], a given PDE (linearized, when not linear) is squared and the first variation of this functional, after scaling with a tuning parameter τ, is added to the

[2] In the words of Seliger and Whitham describing their work [8]: "We are still short, however, of a general theorem stating the conditions under which a variational principle can be found for any given system of equations and of an automatic fool-proof method of producing it."

standard Galerkin weak form. This enhances the stability of the standard Galerkin method. However, the weak formulation in GLS recovers the strong form only when $\tau \to 0$. The derivation of variational methods for multiscale phenomena and the connection of subgrid models therein to enriched bubble functions and to stabilized methods is presented in [24].

An alternative approach is taken in the least-squares finite element method (LSFEM) [25]: if $Au = 0$ is a PDE system (A can be a nonlinear operator), then the LSFEM functional $\|Au\|^2$ in an appropriate norm (e.g., L^2) is minimized. On setting the first variation of this functional to zero leads to the variational equations. A positive attribute of LSFEM is that it results in a symmetric stiffness matrix. However, though the solution of the strong form minimizes the LSFEM functional, notably for nonlinear problems it is not guaranteed that all solutions to the Euler–Lagrange equations of the LSFEM functional are solutions of the original PDE.[3] In addition, there is no systematic procedure to incorporate boundary conditions.

From a computational viewpoint, the availability of a variational principle is beneficial; for instance, regardless of nonlinearities, discretizations based on a variational principle result in symmetric stiffness/Jacobian matrices. To solve nonlinear PDEs (e.g., Allen–Cahn and Burgers' equations), transient wave phenomena as well as chaotic dynamical systems, the strong form is used in physics-informed neural networks (PINNs) [26–28], Kolomogorov–Arnold networks (KANs) [29, 30], and neural networks based on the Kolmogorov superposition theorem [31]. Use of duality principles can provide alternate routes to solutions via functional minimization with neural networks to better capture the underlying physics. For instance, deep Ritz [32] provide certain advantages to solving variational formulations with neural networks: only Dirichlet boundary conditions are required to be imposed (advances in [33] can be leveraged), and as a consequence, admissible neural network approximations ensure that the potential energy functional for the Poisson equation or in nonlinear elasticity is bounded, and since only first-order derivatives are computed, the computational overhead is less than when using the strong form for second-order PDEs. With the dual approach, smooth approximants for the dual fields are desirable, which are readily constructed using neural network approximants. In addition, since imposing homogeneous Dirichlet boundary conditions on dual fields suffices, it is easier to construct (a priori) ansatz that are kinematically admissible. Furthermore, a transient problem in the dual approach is posed as a boundary-value problem with a terminal Dirichlet boundary condition that can be solved using a space-time neural network discretization in finite time slices (see [17] for time-slicing in the dual context).

In previous studies, linear C^0 finite elements have been used in the dual variational formulation to solve linear and nonlinear PDEs and an ODE: transient heat conduction, linear transport and Euler's nonlinear ODE system for the motion of a rigid body [17]; nonconvex one-dimensional elastostatics and elastodynamics [12]; and (inviscid) Burgers equation [18]. Since in general the primal fields depend on the derivative of dual fields, an L^2 projection of

[3] This was understood, and explained in [8, Sec. 1, p. 2].

the DtP generated primal fields onto C^0 finite element spaces was used to reconstruct continuous primal fields. In this paper, we adopt smooth neural network approximations and high-order B-spline approximating functions in the dual variational formulation. In the numerical implementation, we show that it is desirable to use smooth, higher order approximations for the dual fields if the primal field is C^0. However, for the pure transport problem, a dual field that is C^0, obtained by adding C^0 basis functions to the C^k ($k \geq 1$) approximants already employed may be beneficial since it accommodates discontinuous solutions in the primal field. In one dimension, shallow neural networks with Rectified Power Unit (RePU) activation function [34, 35] and univariate B-splines [36] are used, and bivariate tensor-product B-splines are adopted in two dimensions.

The remainder of this paper is structured as follows. In Section 2, we introduce the general formalism of the dual variational principle and derive the dual function for a system of nonlinear equations. The solution procedure via the duality approach is presented for least squares minimization, solving a system of two quadratic equations, and to obtain nonnegative solutions for an underdetermined system of linear equations via entropy maximization. In Section 3, we first derive the dual formulation for an initial-value problem; the weak form of the associated dual second-order boundary-value problem is also derived, with the latter further elaborated in Appendix B. Then we present the dual formulation for the transient convection-diffusion equation. Numerical discretization of the variational form, with expressions for the stiffness matrix and force vector, are provided in Section 4. In Section 5, numerical results are presented for the Laplace equation, one-dimensional steady-state convection-diffusion, transient convection-diffusion, and transient heat equations. Accurate results are obtained with sound convergence, and we summarize the main findings from this work in Section 6.

2. General formalism for the finite-dimensional case

Following [11, Sec. 1], let $\boldsymbol{G}(\boldsymbol{U}) = \boldsymbol{0}$ represent a system of linear equations, nonlinear equations, ODEs or PDEs, where $\boldsymbol{G} : \mathbb{R}^n \to \mathbb{R}^m$. To fix ideas, we first present the dual formalism to solve a system of nonlinear equations. To formulate an optimization problem that is preferably convex, consider the objective function $\boldsymbol{U} \mapsto H(\boldsymbol{U})$, where $H(\boldsymbol{U})$ is a convex (choice is flexible) auxiliary potential that is tailored to dominate the nonlinearities in the primal problem.

Let $\boldsymbol{\lambda} \in \mathbb{R}^m$ be the Lagrange multiplier vector that is associated with the constraints. We now optimize (minimize) H subject to the constraint $\boldsymbol{G}(\boldsymbol{U}) = \boldsymbol{0}$, and can write the Lagrangian and the stationary conditions as:

$$L_H(\boldsymbol{U}, \boldsymbol{\lambda}) := H(\boldsymbol{U}) + \boldsymbol{\lambda} \cdot \boldsymbol{G}(\boldsymbol{U}), \tag{1a}$$

$$\nabla_{\boldsymbol{U}} L_H(\boldsymbol{U}, \boldsymbol{\lambda}) = \nabla H(\boldsymbol{U}) + \boldsymbol{\lambda} \cdot \nabla \boldsymbol{G}(\boldsymbol{U}) = \boldsymbol{0}, \tag{1b}$$

$$\nabla_{\boldsymbol{\lambda}} L_H(\boldsymbol{U}, \boldsymbol{\lambda}) = \boldsymbol{G}(\boldsymbol{U}) = \boldsymbol{0}. \tag{1c}$$

In (1), $\boldsymbol{U} \in \mathbb{R}^n$ are the primal variables and $\boldsymbol{\lambda} \in \mathbb{R}^m$ are the dual variables. In the critical point formulation, which shows the consistency of our scheme as generating solutions to the primal problem, one solves (1b) for $\boldsymbol{U}$ as a function

of $\boldsymbol{\lambda}$, the dual-to-primal mapping $\boldsymbol{U}_H(\boldsymbol{\lambda})$, and then look for $\boldsymbol{\lambda}^*$ that solves (1c) in the form $\boldsymbol{G}(\boldsymbol{U}_H(\boldsymbol{\lambda}^*)) = \boldsymbol{0}$. Defining the function $\boldsymbol{U}_H$ is facilitated by the choice of H. In the related convex optimization formulation, one solves the maximization problem

$$\sup_{\boldsymbol{\lambda}} \inf_{\boldsymbol{U}} L_H(\boldsymbol{U}, \boldsymbol{\lambda}) = \sup_{\boldsymbol{\lambda}} S(\boldsymbol{\lambda}), \tag{2}$$

where $S(\boldsymbol{\lambda})$ is the dual function, which is concave, since it is the pointwise infimum of a family of affine functions, regardless of the nonlinearity of $\boldsymbol{G}$. A maximizer (and a critical point) $\boldsymbol{\lambda}^*$ of the dual function $S(\boldsymbol{\lambda})$ is defined as

$$\boldsymbol{\lambda}^* = \operatorname*{argsup}_{\boldsymbol{\lambda}} S(\boldsymbol{\lambda}). \tag{3}$$

Note that the convex optimization problem does not need a notion of a DtP mapping (in fact, a concave dual problem can be obtained even with $H = 0$). However, for a meaningful scheme that addresses the question of finding solutions to the primal problem, a choice of a strictly convex function H facilitates the definition of a DtP map. The maximizer of the dual problem in this case, with some consideration on the smoothness of S at the maximizer $\boldsymbol{\lambda}^*$ as well as the smoothness of the DtP map, defines a primal solution as

$$\boldsymbol{U}^* = \boldsymbol{U}_H(\boldsymbol{\lambda}^*).$$

In this connection, we note that while the hypotheses of the Implicit Function theorem, primarily related to the invertibility of the Hessian $\nabla^2_U L_H(\cdot, \boldsymbol{\lambda})$ in $O_n \times O_m \subset \mathbb{R}^n \times \mathbb{R}^m$, ensures the existence of the function $\boldsymbol{U}_H : O^m \to \mathbb{R}^n$ (at least locally), the existence of such a function does not imply $\inf_{\boldsymbol{U}} L_H(\boldsymbol{U}, \boldsymbol{\lambda}) = L_H(\boldsymbol{U}_H(\boldsymbol{\lambda}), \boldsymbol{\lambda})$. Strict convexity of $L_H(\cdot, \boldsymbol{\lambda})$ in O_n for each $\boldsymbol{\lambda} \in O_m$ is a guarantee of this fact.

To verify that $\boldsymbol{U}^*$ satisfies (1c), that is $\boldsymbol{G}(\boldsymbol{U}^*) = \boldsymbol{0}$, assume that in a neighborhood of $\boldsymbol{\lambda}^*$, $S(\boldsymbol{\lambda}) = \inf_{\boldsymbol{U}} L_H(\boldsymbol{U}, \boldsymbol{\lambda}) = L_H(\boldsymbol{U}_H(\boldsymbol{\lambda}), \boldsymbol{\lambda})$ holds, and observe that solving the optimization problem posed in (3) by setting the gradient of S to vanish and using (1b) yields

$$\boldsymbol{0} = \left.\frac{\partial S}{\partial \boldsymbol{\lambda}}\right|_{(\boldsymbol{\lambda} = \boldsymbol{\lambda}^*)} = \left.\frac{\partial L_H}{\partial \boldsymbol{\lambda}}\right|_{(\boldsymbol{U} = \boldsymbol{U}_H(\boldsymbol{\lambda}^*), \boldsymbol{\lambda} = \boldsymbol{\lambda}^*)} = \boldsymbol{G}(\boldsymbol{U}_H(\boldsymbol{\lambda}^*)) = \boldsymbol{G}(\boldsymbol{U}^*),$$

which establishes that the primal problem is solved via dual maximization.

We now apply the duality principle to three test cases: system of linear equations, solution of two quadratic equations and nonnegative solutions for an underdetermined system of linear equations.

2.1. Linear system of equations

Following [10, Sec. 2], we seek $\boldsymbol{x} \in \mathbb{R}^n$ that solves the system of linear equations

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{4}$$

by the dual methodology, where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. If $m = n$, then (4) has a unique solution if $\boldsymbol{A}$ has full rank of n. If $n > m$ ($n < m$), the linear system is underdetermined (overdetermined).

For the dual approach, we choose $H(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top\boldsymbol{x}$ as the convex potential (function). Let $\boldsymbol{\lambda} \in \mathbb{R}^m$ be the Lagrange multiplier vector. We form the Lagrangian $L(\cdot,\cdot)$, and pose the optimization problem as:

$$\max_{\boldsymbol{\lambda}} \min_{\boldsymbol{x}} \Big[L(\boldsymbol{x},\boldsymbol{\lambda}) := H(\boldsymbol{x}) + \boldsymbol{\lambda}^\top(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})\Big]. \tag{5}$$

On setting $\frac{\partial L}{\partial \boldsymbol{x}} = \boldsymbol{0}$, we obtain the DtP map:

$$\boldsymbol{x} := \boldsymbol{x}_H = \boldsymbol{A}^\top\boldsymbol{\lambda}, \tag{6}$$

which on substituting in (5) yields the dual function:

$$S(\boldsymbol{\lambda}) = \min_{\boldsymbol{x}} L(\boldsymbol{x},\boldsymbol{\lambda}) = L(\boldsymbol{x}_H,\boldsymbol{\lambda}) = -\frac{1}{2}\boldsymbol{\lambda}^\top\boldsymbol{A}\boldsymbol{A}^\top\boldsymbol{\lambda} + \boldsymbol{\lambda}^\top\boldsymbol{b}, \tag{7}$$

which is a quadratic form in $\boldsymbol{\lambda}$. Now, the dual maximization problem is:

$$\boldsymbol{\lambda}^* = \operatorname*{argmax}_{\boldsymbol{\lambda}} S(\boldsymbol{\lambda}), \tag{8}$$

which is solved using the necessary (stationary) condition $\frac{\partial S}{\partial \boldsymbol{\lambda}} = \boldsymbol{0}$ to yield

$$\boldsymbol{\lambda}^* = (\boldsymbol{A}\boldsymbol{A}^\top)^{-1}\boldsymbol{b}. \tag{9}$$

Hence, on using (6), the solution for $\boldsymbol{x}$ is:

$$\boldsymbol{x}_H = \boldsymbol{A}^\top(\boldsymbol{A}\boldsymbol{A}^\top)^{-1}\boldsymbol{b}. \tag{10}$$

Note that the complete least squares solution of (4) is:

$$\boldsymbol{x}_{\text{LS}} = (\boldsymbol{A}^\top\boldsymbol{A})^\dagger\boldsymbol{A}^\top\boldsymbol{b}, \tag{11}$$

where $(\cdot)^\dagger$ is the Moore–Penrose pseudoinverse. We point out that there are distinctions between the solutions in (10) and (11) that are obtained by the duality approach and least squares minimization, respectively: least squares approach delivers a solution even if the linear system in (4) is not consistent, i.e., $\boldsymbol{b}$ is not in the column space of $\boldsymbol{A}$, whereas a solution via the duality approach is realized if and only if the linear system in (4) has at least one solution.

2.2. System of quadratic equations

Consider the following system of quadratic equations:

$$x^2 + y^2 = 3, \qquad x^2 - y^2 = 1, \tag{12a}$$

with exact solution sets:

$$(x, y) = (\pm\sqrt{2}, \pm 1). \tag{12b}$$

Since the solution to (12a) is nonunique, we choose $H(x, y; \bar{x}, \bar{y}, \beta) = \frac{\beta}{2}\left[(x-\bar{x})^2 + (y-\bar{y})^2\right]$ as the convex potential, where $(\bar{x}, \bar{y})$ is referred to as a base state and $\beta \in \mathbb{R}$ is a constant. Judicious choice of the base state and β allows one to target specific solutions in the primal problem [12]. For an example comprising a system of a quadratic and a linear algebraic equation, see [18, App. B]. The Lagrangian saddle-point problem is:

$$\max_{\boldsymbol{\lambda}} \min_{x,y} \left[L(\boldsymbol{x}, \boldsymbol{\lambda}) := H(x, y; \bar{x}, \bar{y}, \beta) + \lambda_1(3 - x^2 - y^2) + \lambda_2(1 - x^2 + y^2)\right]. \tag{13}$$

On setting $\frac{\partial L}{\partial \boldsymbol{x}} = \boldsymbol{0}$, we obtain the DtP map:

$$x := x_H = \frac{\beta\bar{x}}{\beta - \lambda_1 - \lambda_2}, \quad y := y_H = \frac{\beta\bar{y}}{\beta + \lambda_1 - \lambda_2}. \tag{14}$$

On substituting x_H and y_H in (13), we obtain the dual function:

$$S(\boldsymbol{\lambda}) = L(x_H(\boldsymbol{\lambda}), y_H(\boldsymbol{\lambda}), \boldsymbol{\lambda}). \tag{15}$$

For $\beta = 10$, and the base state $(\bar{x}, \bar{y}) = (1, 1)$, the plot of $S(\boldsymbol{\lambda})$ is presented in Fig. 1. We observe that g is maximized at $\lambda_1 = \lambda_2 \approx 1.467$, and hence the primal solution from (14) is: $x = x_H = \sqrt{2}$ and $y = y_H = 1$, which is in agreement with one of the exact solutions given in (12b).

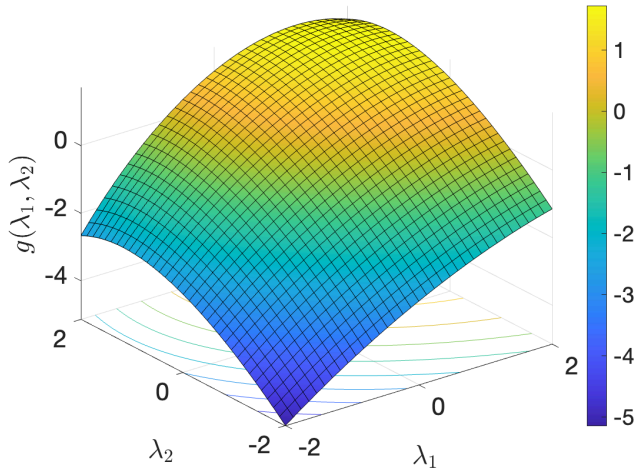


Figure 1: Plot of the dual function $S(\boldsymbol{\lambda})$ for the system of two quadratic equations.

2.3. Entropy shape functions for a polygon

Consider a convex polygon $P \subset \mathbb{R}^2$ with nodal (vertex) coordinates $\{\boldsymbol{x}_i\}_{i=1}^n$, where $\boldsymbol{x}_i \equiv (x_i, y_i)$. Let $\boldsymbol{\phi}(\boldsymbol{x}) : P \to \mathbb{R}^n_+$ be the n nonnegative generalized barycentric coordinates (shape functions) for a polygon [37]. For each fixed $\boldsymbol{x} \in P$,

these nonnegative shape functions ($\phi_i(\boldsymbol{x}) \geq 0$) satisfy the constant and linear reproducing conditions, namely

$$\sum_{i=1}^{n} \phi_i(\boldsymbol{x}) = 1, \tag{16a}$$

$$\sum_{i=1}^{n} \phi_i(\boldsymbol{x}) \boldsymbol{x}_i = \boldsymbol{x}. \tag{16b}$$

If $n > 3$ (polygon with more than three vertices), the solution to (16) is nonunique. A feasible solution using the convex potential H as the Shannon entropy [38, 39] is proposed in [40]. The primal optimization problem is posed as:

$$\max_{\boldsymbol{\phi} \in \mathbb{R}^n_+} \left[H(\boldsymbol{\phi}) := -\sum_{i=1}^{n} \phi_i \ln \phi_i \right], \tag{17}$$

subject to the linear constraints in (16).

If λ_0 and $\boldsymbol{\lambda} \in \mathbb{R}^2$ are the Lagrange multipliers associated with the constraints in (16a) and (16b), respectively, then the Lagrangian saddle-point problem is:

$$\min_{\lambda_0, \boldsymbol{\lambda}} \max_{\boldsymbol{\phi} \in \mathbb{R}^n_+} \left[L(\boldsymbol{\phi}, \lambda_0, \boldsymbol{\lambda}) = -\sum_{i=1}^{n} \phi_i \ln \phi_i + \lambda_0 \left(1 - \sum_{i=1}^{n} \phi_i \right) + \boldsymbol{\lambda} \cdot \sum_{i=1}^{n} \phi_i (\boldsymbol{x} - \boldsymbol{x}_i) \right]. \tag{18}$$

On using the stationary conditions, we obtain

$$\frac{\partial L}{\partial \boldsymbol{\phi}} = \boldsymbol{0} \implies -\ln \phi_i - \ln Z - \boldsymbol{\lambda} \cdot (\boldsymbol{x}_i - \boldsymbol{x}) = 0 \quad (i = 1, 2, \ldots n), \tag{19}$$

where $\lambda_0 = \ln Z - 1$ (Z is the partition function). On simplifying (19) and using (16a) leads us to the DtP map:

$$\phi_i(\boldsymbol{\lambda}) := \phi_i^H(\boldsymbol{\lambda}) = \frac{Z_i(\boldsymbol{\lambda})}{Z(\boldsymbol{\lambda})}, \quad Z(\boldsymbol{\lambda}) = \sum_{j=1}^{n} Z_j(\boldsymbol{\lambda}), \quad Z_i(\boldsymbol{\lambda}) = \exp(-\boldsymbol{\lambda} \cdot (\boldsymbol{x}_i - \boldsymbol{x})). \tag{20}$$

On substituting $\phi_i^H(\boldsymbol{\lambda})$ from (20) in (18) yields the dual function:

$$S(\boldsymbol{\lambda}) = \max_{\boldsymbol{\phi} \in \mathbb{R}^n_+} L(\boldsymbol{\phi}, \lambda_0, \boldsymbol{\lambda}) = L(\boldsymbol{\phi}^H(\boldsymbol{\lambda}), \lambda_0, \boldsymbol{\lambda}) = \ln Z(\boldsymbol{\lambda}), \tag{21}$$

and the dual variational principle can now be stated as:

$$\boldsymbol{\lambda}^* = \operatorname*{argmin}_{\boldsymbol{\lambda} \in \mathbb{R}^2} S(\boldsymbol{\lambda}), \tag{22}$$

which is an unconstrained convex optimization problem. Note that the dual problem has two unknowns, whereas the primal problem has $n \geq 3$ unknowns. Finally, we point the reader to a related primal-dual connection in meshfree methods. Moving least squares basis functions are computed via a dual formulation (minimizing a quadratic form [41]), whose primal formulation is the minimization of a quadratic form subject to the linear constraints in (16) [42].

3. Dual formulation for differential equations

We present the derivation of the dual functional and the dual variational principle for an ODE and then for PDEs. First, as an introduction to the duality method, we select an initial-value problem. Then we consider the transient convection-diffusion equation in one spatial dimension, so that as special cases we can obtain the dual functional for the Laplace equation, one-dimensional steady-state convection-diffusion equation, transient heat equation, and the one-dimensional transport equation.

3.1. Initial-value problem

We begin by illuminating the scheme in the simplest possible setting, which also explicitly demonstrates how an initial-value problem (IVP) can be robustly solved using a boundary-value problem in time. While simple, we are not aware of any existing variational principle for this problem. For $a, u_0 \in \mathbb{R}$, consider the following IVP:

$$\dot{u}(t) = au(t), \quad t \in (0, T), \tag{23a}$$

with the initial condition

$$u(0) = u_0. \tag{23b}$$

If $a < 0$ in (23a), the problem is dissipative.

The goal is to design a variational principle whose Euler–Lagrange equation is the ODE in an appropriate sense. To do so, we treat the ODE in (23a) as a constraint and optimize the arbitrarily chosen auxiliary convex potential (density) function, $H(u) = \frac{1}{2}u^2$, by the method of Lagrange multipliers (dual variables), with an important distinction – instead of looking for a solution in the larger (u, λ) space, we treat the optimality condition with respect to the primal field u as generating a functional relationship for the primal field u in terms of the dual field λ – and in doing so, we solve the problem in the smaller space of just the dual variables (this idea is general, and applies to the considerations of Sec. 2). A crucial ingredient that enables this approach is the free choice of the auxiliary potential. We begin by creating a functional in the fields (u, λ) from (23) with the dual field λ treated like a test function and the auxiliary function H appended to it:

$$\begin{aligned}
&\int_0^1 [H(u) + \lambda(\dot{u} - au)]\, dt = \int_0^1 \left[\frac{u^2}{2} + \lambda(\dot{u} - au)\right] dt \\
&= u(T)\lambda(T) - u(0)\lambda(0) + \int_0^1 \left[\frac{u^2}{2} - u(\dot{\lambda} + a\lambda)\right] dt \\
&= u(T)\lambda(T) - u_0\lambda(0) + \int_0^1 \left[\frac{u^2}{2} - u(\dot{\lambda} + a\lambda)\right] dt,
\end{aligned} \tag{24}$$

where integration by parts has been used and the initial condition has been incorporated to arrive at the last equality. We now omit the term $u(T)\lambda(T)$ – this choice is motivated by the sole design consideration that the Euler–Lagrange equation of the variational principle that we propose should recover the primal problem (see discussion around (29))

– and refer to the functional that results as L:

$$L[u,\lambda] = \int_0^T \mathcal{L}(u,\mathcal{D})\,dt - u_0\lambda(0), \quad \mathcal{L}(u,\mathcal{D}) = \frac{u^2}{2} - u(\dot{\lambda} + a\lambda), \quad \mathcal{D} = (\lambda,\dot{\lambda}), \tag{25}$$

where $\mathcal{L}$ is the Lagrangian density for the primal-dual problem.

Next we require that

$$\frac{\partial \mathcal{L}}{\partial u}(u,\mathcal{D}) = 0 \implies u - \dot{\lambda} - a\lambda = 0 \text{ for all } \mathcal{D}, \tag{26}$$

which generates the dual-to-primal (DtP) mapping $\mathcal{D} \mapsto u_H(\mathcal{D})$:

$$u := u_H(\mathcal{D}) = \dot{\lambda} + a\lambda. \tag{27}$$

This solvability requirement serves as the main design specification for the choice of the auxiliary potential. Substituting u_H from (27) in (25) generates the required dual functional ('action'):

$$S[\lambda] := L[u_H(\mathcal{D}),\lambda] = -\frac{1}{2}\int_0^T [u_H(\mathcal{D})]^2\,dt - u_0\lambda(0) = \int_0^T \mathcal{L}(u_H(\mathcal{D}),\mathcal{D})\,dt - u_0\lambda(0), \tag{28}$$

with the dual Lagrangian density given simply by $\mathcal{L}(u_H(\mathcal{D}),\mathcal{D})$. To obtain the strong form posed in (23), we begin by setting $\delta S[\lambda;\delta\lambda] = 0$ for all variations, and use integration by parts and the DtP map given in (27) along with the condition that $\delta\lambda$ vanishes at the terminal time (defining the set of admissible variations) to obtain

$$\dot{u}_H(t) = a u_H(t), \quad u_H(0) = u_0 \tag{29}$$

(where for ease of notation we use $u_H(t) := u_H(\mathcal{D}(t))$. This last requirement ($\delta\lambda(T) = 0$) indicates that one needs to impose a Dirichlet boundary condition on $\lambda(T)$ (not necessarily vanishing), which was the sole motivation to omit the term $u(T)\lambda(T)$ in (24). The key enabling feature of the above consistency of the dual functional with the primal problem is the condition $\frac{\partial \mathcal{L}}{\partial u} = 0$:

$$\begin{aligned}
\delta S[\lambda;\delta\lambda] &= \int_0^T \frac{\partial \mathcal{L}}{\partial \mathcal{D}}(u_H,\mathcal{D})\,\delta\mathcal{D}\,dt - u_0\delta\lambda(0) \\
&= -\int_0^T \left(u_H\delta\dot{\lambda} + a u_H\delta\lambda\right) dt - u_0\delta\lambda(0) = \int_0^T (\dot{u}_H - a u_H)\,\delta\lambda\,dt + (u_H(0) - u_0)\delta\lambda(0), \\
&\text{so that} \quad 0 = \delta S[\lambda;\delta\lambda] \quad \forall \delta\lambda,\ \delta\lambda(T) = 0 \iff (29),
\end{aligned} \tag{30}$$

on using the DtP mapping condition (26) and the affine dependence of $\mathcal{L}$ on $\mathcal{D}$ along with the Dirichlet boundary condition on $\lambda(T)$, and this persists regardless of the nonlinearities of $\mathcal{L}(u,\mathcal{D})$ in u. Substituting (26) in (30) along with using the arbitrariness of $\delta\lambda$ constrained by $\delta\lambda(T) = 0$ makes contact with the demonstration in Appendix B, where the weak form (the third expression in the chain of equalities in (30) set to zero) is worked out explicitly in terms of dual variables.

On substituting the DtP map from (27) in (29), the dual Euler–Lagrange (second-order) boundary-value problem

is given by:

$$\ddot{\lambda} - a^2\lambda = 0, \tag{31a}$$

$$\dot{\lambda}(0) + a\lambda(0) = u_0, \tag{31b}$$

$$\lambda(T) = \lambda_T \quad \text{(arbitrarily chosen)}. \tag{31c}$$

This is an *elliptic* boundary-value-problem in $(0, T)$ for the function λ. The reason why the condition at the terminal time $t = T$ in (31c) can be imposed, even though the primal problem (23) is an IVP that does not admit a 'final-time' boundary condition, is that

$$u_H(T) = \dot{\lambda}(T) + a\lambda(T), \tag{32}$$

and even with $\lambda(T)$ specified, $\dot{\lambda}(T)$ is free to adjust to recover the unique value of $u_H(T)$. Here, $\dot{\lambda}(T) = u_0 e^{aT} - a\lambda_T$ has to be satisfied, which depends on both u_0 and λ_T.

Clearly, the solution $t \mapsto u_H(t) = u_H(\mathcal{D}(t))$ cannot depend on λ_T by uniqueness of the primal solution, even though $\mathcal{D}(t) = (\lambda(t), \dot{\lambda}(t))$ depends on λ_T in a non-trivial manner (see (33) and (34) below), and we verify it in this specific case since it is perhaps not obvious from visual inspection. On writing the dual solution as

$$\lambda(t) = c_1 e^{|a|t} + c_2 e^{-|a|t} \tag{33}$$

for constants c_1, c_2, the boundary conditions in (31) yield

$$\begin{bmatrix} (|a| + a) & -(|a| - a) \\ e^{|a|T} & e^{-|a|T} \end{bmatrix} \begin{Bmatrix} c_1 \\ c_2 \end{Bmatrix} = \begin{Bmatrix} u_0 \\ \lambda_T \end{Bmatrix}, \tag{34}$$

and on using (27), $u_H(t)$ can be expressed as

$$\begin{aligned} u_H(t) &= \begin{Bmatrix} (|a| + a)e^{|a|t} & -(|a| - a)e^{-|a|t} \end{Bmatrix} \begin{Bmatrix} c_1 \\ c_2 \end{Bmatrix} \\ &= \frac{1}{\left((|a| + a)e^{-|a|T} + (|a| - a)e^{|a|T}\right)} \begin{Bmatrix} (|a| + a)e^{|a|t} & -(|a| - a)e^{-|a|t} \end{Bmatrix} \begin{Bmatrix} u_0 e^{-|a|T} + \lambda_T(|a| - a) \\ -u_0 e^{|a|T} + \lambda_T(|a| + a) \end{Bmatrix}. \end{aligned}$$

The term involving λ_T in $u_H(t)$ evaluates to

$$\lambda_T(|a|^2 - a^2)e^{|a|t} - \lambda_T(|a|^2 - a^2)e^{-|a|t} = 0 \quad (!),$$

and it can be verified that the rest of the terms evaluate to

$$u_H(t) = u_0 e^{at}.$$

In closing this section, we mention that the dual variational scheme applies seamlessly to nonlinear systems of ODEs, as demonstrated in [17] for Euler's system for the motion of a rigid body about a fixed point with and without damping.

3.2. Transient convection-diffusion equation

Consider the strong form of the transient convection-diffusion model problem:

$$\kappa \frac{\partial^2 u}{\partial x^2} - \alpha \frac{\partial u}{\partial x} = \frac{\partial u}{\partial t} \quad \text{in } \Omega = \Omega_0 \times \Omega_t = (0,1) \times (0,1), \tag{36a}$$

$$u(0,t) = \bar{u}_1, \quad u(1,t) = \bar{u}_2, \tag{36b}$$

$$u(x,0) = u_0(x), \tag{36c}$$

where $\kappa \geq 0$ is the diffusion coefficient, α is the convection coefficient, $\bar{u}_1$ and $\bar{u}_2$ are boundary data (constants) and $u_0(x)$ is the prescribed initial condition. In primal form, the convection-diffusion equation does not possess a variational structure, and therefore weak formulations have to be constructed using the strong form. In the strongly convective regime, standard Galerkin methods produce spurious oscillations. To remedy this deficiency, upwinding and/or stabilized methods such as SUPG [21] and GLS [22] have been proposed, but they involve a user-specified tuning parameter and do not possess a direct correspondence with a variational principle. Our goal is to devise a variational scheme that is based on dual fields, so that standard Galerkin discretization can be directly used (without upwinding or stabilization, nor concerns about the inf-sup condition) to compute accurate solutions.

On introducing the field $q = \partial u/\partial x$, we recast (36a) as a system of two first-order PDEs:

$$\kappa \frac{\partial q}{\partial x} - \alpha q = \frac{\partial u}{\partial t}, \tag{37a}$$

$$q = \frac{\partial u}{\partial x}. \tag{37b}$$

We refer to (u, q) as the primal fields. Let λ and μ be the dual (Lagrange multiplier) fields that correspond to the two constraints in (37). Now, we choose the convex potential (density) function $H(u, q)$ as:

$$H(u,q) = \frac{1}{2}(u^2 + q^2), \tag{38}$$

and consider the functional

$$\widehat{L}[u,q,\lambda,\mu] = \int_0^1 \int_0^1 H(u,q)\,dx\,dt \;+ \int_0^1 \int_0^1 \left[\frac{\partial u}{\partial t} - \kappa \frac{\partial q}{\partial x} + \alpha q\right] \lambda\,dx\,dt - \int_0^1 \int_0^1 \left[\frac{\partial u}{\partial x} - q\right] \mu\,dx\,dt. \tag{39a}$$

On using the divergence theorem on terms that involve the partial derivatives of u and q in $\widehat{L}(u,q,\lambda,\mu)$, we obtain

$$\begin{aligned}\widehat{L}[u,q,\lambda,\mu] = &\int_0^1\int_0^1 \frac{u^2+q^2}{2}dx\,dt + \int_0^1\int_0^1 \left[\kappa q\frac{\partial\lambda}{\partial x} + \alpha q\lambda - u\frac{\partial\lambda}{\partial t} - q\mu - u\frac{\partial\mu}{\partial x}\right]dx\,dt\\ &+ \int_0^1 [u\mu - \kappa q\lambda]_{x=0}^{x=1}\,dt + \int_0^1 [u\lambda]_{t=0}^{t=1}\,dx.\end{aligned} \tag{40}$$

There are six boundary terms in (40): initial and Dirichlet boundary conditions on u are present in three of them (these appear as natural boundary conditions in $\widehat{L}$), and the remaining three terms involve λ. On following the treatment of the boundary terms for the IVP in Section 3.1, we can use any Dirichlet data (zero or nonzero) on λ for these boundary terms, and the corresponding terms are dropped from $\widehat{L}$. We refer to the functional that results as L. Now, on incorporating the boundary conditions from (36b) and the initial condition from (36c), we obtain

$$\begin{aligned}L[u,q,\lambda,\mu] = &\int_0^1\int_0^1 \frac{u^2+q^2}{2}dx\,dt + \int_0^1\int_0^1 \left[-u\left(\frac{\partial\lambda}{\partial t}+\frac{\partial\mu}{\partial x}\right) - q\left(-\kappa\frac{\partial\lambda}{\partial x} - \alpha\lambda + \mu\right)\right]dx\,dt\\ &+ \int_0^1 [\bar{u}_2\mu(1,t) - \bar{u}_1\mu(0,t)]\,dt - \int_0^1 u_0(x)\lambda(x,0)\,dx,\end{aligned} \tag{41}$$

which remains valid for nonhomogeneous Dirichlet data for λ (see the numerical example in Section 5.1). The functional L in (41) is an example of a pre-dual functional as introduced in prior works [9–14, 17, 18]. We now pose the primal-dual/mixed variational problem as:

$$\sup_{\lambda,\mu}\inf_{u,q}\; L[u,q,\lambda,\mu]. \tag{42}$$

Define

$$\mathcal{L}(u,q,\mathcal{D}) := \frac{u^2+q^2}{2} - u\left(\frac{\partial\lambda}{\partial t}+\frac{\partial\mu}{\partial x}\right) - q\left(-\kappa\frac{\partial\lambda}{\partial x} - \alpha\lambda + \mu\right), \quad \mathcal{D} = \left(\lambda, \frac{\partial\lambda}{\partial x}, \dot{\lambda}, \mu, \frac{\partial\mu}{\partial x}\right), \tag{43}$$

which is the space-time integrand that appears in (41), and corresponds to the *Lagrangian density* of variational field theories and particle mechanics. On setting

$$\frac{\partial\mathcal{L}}{\partial u} = 0, \quad \frac{\partial\mathcal{L}}{\partial q} = 0, \tag{44a}$$

we obtain the DtP map:

$$u := u_H(\mathcal{D}) = \frac{\partial\lambda}{\partial t} + \frac{\partial\mu}{\partial x}, \quad q := q_H(\mathcal{D}) = \mu - \alpha\lambda - \kappa\frac{\partial\lambda}{\partial x}. \tag{44b}$$

Using (44b) in (42), we have (in this quadratic setting)

$$\inf_{u,q} L[u,q,\lambda,\mu] = L[u_H(\mathcal{D}), q_H(\mathcal{D})] =: S[\lambda,\mu], \tag{45}$$

where $S[\lambda,\mu]$ is the dual functional (*action integral*), and now we state the dual variational principle:

$$\sup_{\lambda\in\mathsf{S}_\lambda,\ \mu\in\mathsf{S}_\mu} S[\lambda,\mu], \tag{46a}$$

$$S[\lambda,\mu] = -\frac{1}{2}\int_0^1\int_0^1 \left[(u_H(\mathcal{D}))^2 + (q_H(\mathcal{D}))^2\right] dx\,dt + \int_0^1 \left[\bar{u}_2\mu(1,t) - \bar{u}_1\mu(0,t)\right] dt - \int_0^1 u_0(x)\lambda(x,0)\,dx, \tag{46b}$$

$$\mathsf{S}_\lambda = \{\lambda : \lambda \in H^1(\Omega),\ \lambda(0,t) = \lambda(1,t) = \lambda(x,1) = 0\}, \quad \mathsf{S}_\mu = \{\mu : \mu \in H^1(\Omega)\}, \tag{46c}$$

where $H^1(\Omega)$ is the Sobolev space that contains functions in Ω with square-integrable derivatives up to order 1, and $u_H(\mathcal{D})$ and $q_H(\mathcal{D})$ are given in (44b). In (46c), we have chosen homogeneous Dirichlet boundary conditions on λ: $\lambda(0,t) = \lambda(1,t) = \lambda(x,1) = 0$. To show that (46) implies the Euler–Lagrange equations of the primal problem and the imposed initial/boundary conditions (strong form) given in (36), one proceeds along familiar lines [11]: start with (46) and set $\delta S[\lambda,\mu;\delta\lambda] = 0$ and $\delta S[\lambda,\mu;\delta\mu] = 0$, apply the divergence theorem, and use (44b) and the fundamental lemma of calculus of variations to arrive at the desired result.

The direct way to see the above consistency check is to realize that the dual functional, $S[\lambda,\mu]$, up to boundary terms, is the space-time integral of the Lagrangian $\mathcal{L}(u_H(\mathcal{D}), q_H(\mathcal{D}))$, and to note that $\frac{\partial\mathcal{L}}{\partial u} = 0$ and that $\mathcal{L}$ is necessarily affine in $\mathcal{D}$ with the coefficient of $\mathcal{D}$ formed from the primal strong form through integration by parts. These steps continue to hold for nonlinear primal PDEs, and form the core idea of the duality-based method presented in this article.

3.3. Laplace equation

Consider the Laplace equation in $\Omega = (0,1)$ with Dirichlet boundary conditions $u(0) = \bar{u}_1$ and $u(1) = \bar{u}_2$. We choose the primal fields u and q that satisfy

$$q' = 0, \qquad q = u', \tag{47}$$

where $(\cdot)' := d(\cdot)/dx$. On using (44b) and (46), setting $\alpha = 0$ and $\kappa = 1$, and dropping the time dependence, the DtP map and the dual functional for this Laplace boundary-value problem are:

$$u := u_H(\mathcal{D}) = \mu', \quad q := q_H(\mathcal{D}) = \mu - \lambda', \quad \mathcal{D} = (\lambda, \lambda', \mu, \mu'), \tag{48a}$$

$$S[\lambda,\mu] = -\frac{1}{2}\int_0^1 \left[(u_H(\mathcal{D}))^2 + (q_H(\mathcal{D}))^2\right] dx + \bar{u}_2\mu(1) - \bar{u}_1\mu(0), \quad \lambda \in \mathsf{S}_\lambda,\ \mu \in \mathsf{S}_\mu, \tag{48b}$$

$$\mathsf{S}_\lambda = \{\lambda : \lambda \in H^1(0,1),\ \lambda(0) = \lambda(1) = 0\}, \quad \mathsf{S}_\mu = \{\mu : \mu \in H^1(0,1)\}. \tag{48c}$$

Let $\bar{u}_1 = 0$ and $\bar{u}_2 = 1$ be the Dirichlet data so that $\mu'(0) = u(0) = 0$ and $\mu'(1) = u(1) = 1$. On substituting (48a) in (47) and noting that the Dirichlet boundary data for λ can be arbitrary, we find that the dual fields satisfy

$$\mu' - \lambda'' = 0 \ \text{ and } \ \mu - \lambda' = \mu'' \implies \lambda'''' = 0 \ \text{ and } \ \mu''' = 0, \tag{49a}$$

$$\mu'(0) = 0, \ \ \mu'(1) = 1, \ \ \lambda(0) = \lambda_0, \ \ \lambda(1) = \lambda_1, \tag{49b}$$

where λ_0 and λ_1 are arbitrary. Equation (49a) provides the Euler–Lagrange equations for the dual fields. First, let us select $\lambda_0 = \lambda_1 = 0$, the choice we made earlier to define S_λ in (48c). Now, on using (49b), the exact solution for λ (cubic function) and μ (quadratic function) can be written as: $\mu(x) = a_0 + x^2/2$ and $\lambda(x) = x(1-x)(b_0 + b_1 x)$, where a_0, b_0 and b_1 are constants. Using the DtP map in (48a), the primal fields are:

$$u(x) = x, \quad q(x) = (a_0 - b_0) + (2b_0 - 2b_1)x + \left(\frac{1}{2} + 3b_1\right)x^2,$$

and to recover the exact solution $q(x) = 1$ we must have $b_1 = -1/6$, $b_0 = b_1 = -1/6$ and $a_0 = 1 + b_0 = 5/6$. Adding any fixed cubic function to $\lambda(x)$ will not affect the recovery of the exact solution, which is consistent with the fact that the boundary data for λ can be arbitrary (need not be homogeneous) in the set S_λ given in (48c).

3.4. One-dimensional, steady state, convection-diffusion equation

On choosing $\kappa = 1$ and dropping the time dependence in (36), we obtain the steady-state one-dimensional convection-diffusion problem: $u'' - \alpha u' = 0$, with boundary conditions $u(0) = \bar{u}_1$, $u(1) = \bar{u}_2$. On using (44b) and (46), the DtP map and the dual functional are:

$$u := u_H(\mathcal{D}) = \mu', \quad q := q_H(\mathcal{D}) = \mu - \alpha\lambda - \lambda', \quad \mathcal{D} = (\lambda, \lambda', \mu, \mu'), \tag{50a}$$

$$S[\lambda, \mu] = -\frac{1}{2}\int_0^1 \left[(u_H(\mathcal{D}))^2 + (q_H(\mathcal{D}))^2\right] dx + \bar{u}_2\mu(1) - \bar{u}_1\mu(0), \quad \lambda \in \mathsf{S}_\lambda,\ \mu \in \mathsf{S}_\mu, \tag{50b}$$

$$\mathsf{S}_\lambda = \{\lambda : \lambda \in H^1(0,1),\ \lambda(0) = \lambda(1) = 0\}, \quad \mathsf{S}_\mu = \{\mu : \mu \in H^1(0,1)\}. \tag{50c}$$

3.5. Transient heat equation

Consider the initial-boundary value problem (IBVP) of heat conduction, where we choose both Dirichlet and Neumann boundary conditions. Setting $\alpha = 0$ in (36) and changing the second Dirichlet boundary condition to a zero flux boundary condition, the IVBP reads:

$$\kappa\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t} \quad \text{in } \Omega = \Omega_0 \times \Omega_t = (0,1) \times (0,1), \tag{51a}$$

$$u(0,t) = \bar{u}_1, \quad \kappa\frac{\partial u}{\partial x}(1,t) = 0, \tag{51b}$$

$$u(x,0) = u_0(x), \tag{51c}$$

where $\kappa \geq 0$ is the thermal conductivity coefficient.

The dual formulation for this problem is derived in [17]. We proceed by following the steps carried out to derive the dual functional in Section 3.2. Referring to (40), we now note that the boundary term $\kappa q(1,t)\lambda(1,t)$ vanishes since $\kappa q(1,t)$ is zero due to the Neumann boundary condition. Hence, $\lambda(1,t)$ is unconstrained. In addition, since the boundary term $u(1,t)\mu(1,t)$ is free, we choose to impose the Dirichlet boundary condition $\mu(1,t) = 0$. With these modifications in place, we use (44b) and (46) to obtain the DtP map and the dual functional for the transient heat

conduction problem:

$$u := u_H(\mathcal{D}) = \frac{\partial \lambda}{\partial t} + \frac{\partial \mu}{\partial x}, \quad q := q_H(\mathcal{D}) = \mu - \kappa \frac{\partial \lambda}{\partial x}, \quad \mathcal{D} = \left(\lambda, \frac{\partial \lambda}{\partial x}, \dot{\lambda}, \mu, \frac{\partial u}{\partial x}\right), \tag{52a}$$

$$S[\lambda, \mu] = -\frac{1}{2} \int_0^1 \int_0^1 \left[(u_H(\mathcal{D}))^2 + (q_H(\mathcal{D}))^2 \right] dx\, dt - \int_0^1 \bar{u}_1 \mu(0,t)\, dt - \int_0^1 u_0(x) \lambda(x,0)\, dx, \tag{52b}$$

$$\mathsf{S}_\lambda = \{\lambda : \lambda \in H^1(\Omega),\ \lambda(0,t) = \lambda(x,1) = 0\}, \quad \mathsf{S}_\mu = \{\mu : \mu \in H^1(\Omega),\ \mu(1,t) = 0\}. \tag{52c}$$

4. Variational formulation and discrete equations

The numerical formulation and implementation of the dual variational form for the transient convection-diffusion problem (see Section 3.2) is presented. Let $\boldsymbol{D} := (\lambda, \mu)$ denote the dual fields. On setting the first variation of the dual functional in (46) to zero, we obtain:

$$\delta S[\lambda, \mu; \delta\lambda] = -\int_0^1 \int_0^1 \left[u_H \delta u_H(\lambda, \mu; \delta\lambda) + q_H \delta q_H(\lambda, \mu; \delta\lambda) \right] dx\, dt - \int_0^1 u_0(x) \delta\lambda(x,0)\, dx = 0, \tag{53a}$$

$$\delta S[\lambda, \mu; \delta\mu] = -\int_0^1 \int_0^1 \left[u_H \delta u_H(\lambda, \mu; \delta\mu) + q_H \delta q_H(\lambda, \mu; \delta\mu) \right] dx\, dt + \int_0^1 \left[\bar{u}_2 \delta\mu(1,t) - \bar{u}_1 \delta\mu(0,t) \right] dt = 0, \tag{53b}$$

and on taking the variation of the DtP map in (44b) and substituting it in (53), we obtain the variational equations

$$-\int_0^1 \int_0^1 \left\{ \left[\frac{\partial \lambda}{\partial t} + \frac{\partial \mu}{\partial x} \right] \frac{\partial (\delta\lambda)}{\partial t} + \left[\mu - \alpha\lambda - \kappa \frac{\partial \lambda}{\partial x} \right] \left[-\alpha\delta\lambda - \kappa \frac{\partial (\delta\lambda)}{\partial x} \right] \right\} dx\, dt - \int_0^1 u_0(x) \delta\lambda(x,0)\, dx = 0, \tag{54a}$$

$$-\int_0^1 \int_0^1 \left\{ \left[\frac{\partial \lambda}{\partial t} + \frac{\partial \mu}{\partial x} \right] \frac{\partial (\delta\mu)}{\partial x} + \left[\mu - \alpha\lambda - \kappa \frac{\partial \lambda}{\partial x} \right] \delta\mu \right\} dx\, dt + \int_0^1 \left[\bar{u}_2 \delta\mu(1,t) - \bar{u}_1 \delta\mu(0,t) \right] dt = 0, \tag{54b}$$

which after rearranging leads to the statement of the dual variational form. Find $\boldsymbol{D} \in \mathsf{S}_\lambda \times \mathsf{S}_\mu$, such that

$$\begin{aligned} &a_{11}(\lambda, \delta\lambda) + a_{12}(\mu, \delta\lambda) = \ell_1(\delta\lambda) \ \ \forall \delta\lambda \in \mathsf{S}_\lambda, \\ &a_{21}(\lambda, \delta\mu) + a_{22}(\mu, \delta\mu) = \ell_2(\delta\mu) \ \ \forall \delta\mu \in \mathsf{S}_\mu, \end{aligned} \tag{55a}$$

where the bilinear forms $a_{ij}(\cdot, \cdot)$, and the linear forms $\ell_1(\cdot$ and $\ell_2(\cdot)$ are given by

$$\begin{aligned} a_{11}(\lambda, \delta\lambda) &= \int_0^1 \int_0^1 \left[\frac{\partial \lambda}{\partial t} \frac{\partial (\delta\lambda)}{\partial t} + \left(\alpha\lambda + \kappa \frac{\partial \lambda}{\partial x} \right) \left(\alpha\delta\lambda + \kappa \frac{\partial (\delta\lambda)}{\partial x} \right) \right] dx\, dt, \\ a_{12}(\mu, \delta\lambda) &= \int_0^1 \int_0^1 \left[\frac{\partial \mu}{\partial x} \frac{\partial (\delta\lambda)}{\partial t} - \mu \left(\alpha\delta\lambda + \kappa \frac{\partial (\delta\lambda)}{\partial x} \right) \right] dx\, dt, \\ a_{21}(\lambda, \delta\mu) &= \int_0^1 \int_0^1 \left[\frac{\partial \lambda}{\partial t} \frac{\partial (\delta\mu)}{\partial x} - \left(\alpha\lambda + \kappa \frac{\partial \lambda}{\partial x} \right) \delta\mu \right] dx\, dt, \\ a_{22}(\mu, \delta\mu) &= \int_0^1 \int_0^1 \left[\frac{\partial \mu}{\partial x} \frac{\partial (\delta\mu)}{\partial x} + \mu\delta\mu \right] dx\, dt, \\ \ell_1(\delta\lambda) &= -\int_0^1 u_0(x) \delta\lambda(x,0)\, dx, \quad \ell_2(\delta\mu) = \int_0^1 \left[\bar{u}_2 \delta\mu(1,t) - \bar{u}_1 \delta\mu(0,t) \right] dt, \end{aligned} \tag{55b, 55c}$$

and

$$\mathsf{S}_\lambda = \{\lambda : \lambda \in H^1(\Omega),\ \lambda(0,t) = \lambda(1,t) = \lambda(x,1) = 0\}, \quad \mathsf{S}_\mu = \{\mu : \mu \in H^1(\Omega)\}. \tag{55d}$$

Homogeneous Dirichlet boundary conditions are chosen for λ in (55d). However, we remind the reader that λ admits non-zero Dirichlet boundary conditions on the left, right and top boundaries.

4.1. Uniqueness of solutions for the dual variational system

To establish uniqueness of solutions to the dual variational equations, we begin by considering two solutions (λ_1, μ_1) and (λ_2, μ_2) of (55), and denote their difference as

$$\lambda_1 - \lambda_2 =: \lambda_d, \qquad \mu_1 - \mu_2 =: \mu_d.$$

Clearly, (λ_d, μ_d) is a test function belonging to $\mathsf{S}_\lambda \times \mathsf{S}_\mu$, which also satisfies (55):

$$\begin{aligned} & a_{11}(\lambda_d, \lambda_d) + a_{12}(\mu_d, \lambda_d) + a_{21}(\lambda_d, \mu_d) + a_{22}(\mu_d, \mu_d) = 0 \\ \implies & \int_0^1 \int_0^1 \left[\left(\frac{\partial \lambda_d}{\partial t} + \frac{\partial \mu_d}{\partial x} \right)^2 + \left(\mu_d - \kappa \frac{\partial \lambda_d}{\partial x} - \alpha \lambda_d \right)^2 \right] dx\, dt = 0. \end{aligned} \tag{56}$$

Thus,

$$\frac{\partial \lambda_d}{\partial t} + \frac{\partial \mu_d}{\partial x} = 0 \ \text{ a.e. in } \Omega \tag{57}$$

and

$$\mu_d - \kappa \frac{\partial \lambda_d}{\partial x} - \alpha \lambda_d = 0 \ \text{ a.e. in } \Omega. \tag{58}$$

Noting the boundary conditions $\lambda_d(0,t) = \lambda_d(1,t) = 0$, on multiplying (57) by λ_d and integrating over Ω yields

$$\int_0^1 \int_t^1 \left[\lambda_d \frac{\partial \lambda_d}{\partial t} - \mu_d \frac{\partial \lambda_d}{\partial x} \right] dx\, dt = 0, \qquad 0 \le t \le 1,$$

and substituting for μ_d from (58) one obtains

$$\begin{aligned} & \int_0^1 \int_t^1 \left[\frac{1}{2} \frac{\partial}{\partial t} \lambda_d^2 - \frac{\partial \lambda_d}{\partial x} \left(\kappa \frac{\partial \lambda_d}{\partial x} + \alpha \lambda_d \right) \right] dx\, dt = 0 \\ \implies & \int_0^1 \int_t^1 \left[\frac{1}{2} \frac{\partial}{\partial t} \lambda_d^2 - \kappa \left(\frac{\partial \lambda_d}{\partial x} \right)^2 - \alpha \frac{1}{2} \frac{\partial}{\partial x} \lambda_d^2 \right] dx\, dt = 0 \\ & \implies \int_t^1 \frac{\partial}{\partial t} \int_0^1 \frac{1}{2} \lambda_d^2\, dx\, dt = \int_0^1 \int_t^1 \kappa \left(\frac{\partial \lambda_d}{\partial x} \right)^2 dx\, dt \ge 0 \end{aligned}$$

due to the boundary conditions satisfied by λ_d and $\kappa \ge 0$. We thus have, with the definition

$$A(t) := \int_0^1 \frac{1}{2} \lambda_d^2(x,t)\, dx \ge 0,$$

that

$$A(1) - A(t) \ge 0 \implies 0 = A(1) \ge A(t) \ge 0$$

on using the fact that $\lambda_d(x,1) = 0$, so that $A(t) = 0$, which implies that

$$\lambda_d = 0 \ \text{ a.e. in } \Omega,$$

and from (58) that

$$\mu_d = 0 \ \text{ a.e. in } \Omega,$$

due to the arbitrariness of $t \in [0,1]$, and uniqueness follows.

Remark: We note that for $\kappa = 0$, the primal system reduces to the linear transport equation, which is known to admit unique solutions that transport discontinuities in u if present in initial or boundary data. It is interesting that the DtP mapping involves gradients of the dual fields to define the primal solution and therefore the dual second-order boundary-value problem in space-time cannot be elliptic, in order to represent such discontinuous primal solutions. Indeed, this is the case, as shown in [17, Secs. 2, 3]. Our uniqueness proof above is also a generalization of similar proofs, done separately for the heat and linear transport equations in [17], where the degenerate ellipticity of the corresponding dual boundary-value problems is also shown. Since the highest order derivatives of the primal operator for the convection-diffusion equation considered here correspond to the heat ($\kappa > 0$) and linear transport ($\kappa = 0$) problems, the dual problem herein may be considered to be degenerate elliptic as well (and certainly for the special cases when either κ or α vanish). It is known that degenerate elliptic problems can have unique solutions [43].

4.2. Discrete equations

We discretize the variational form using a standard Galerkin method. On borrowing notation from neural networks (subscript θ points to the unknown weights and biases in a network), the numerical approximation for the dual fields are written as:

$$\lambda_\theta(x,t) = \sum_{j=1}^{N_\lambda} \phi_j^\lambda(x,t)\, d_j := \boldsymbol{N}_\lambda(x,t)\, \boldsymbol{d}_\lambda \in \mathcal{S}_\lambda, \quad \mu_\theta(x,t) = \sum_{j=1}^{N_\mu} \phi_j^\mu(x,t)\, d_j^\mu := \boldsymbol{N}_\mu(x,t)\, \boldsymbol{d}_\mu \ \in \mathcal{S}_\mu, \tag{59}$$

where the sets $\{\phi_j^\lambda\}_{j=1}^{N_\lambda}$ and $\{\phi_j^\mu\}_{j=1}^{N_\mu}$ are approximating functions for λ and μ, respectively. For convenience, we use $\lambda_\theta(x,t)$ and $\mu_\theta(x,t)$ to denote the trial functions in a neural network- or B-spline-based Galerkin method. We ensure that by construction, $\lambda_\theta(x,t)$ a priori satisfies the homogeneous Dirichlet boundary conditions in (55d). Let $\boldsymbol{d} := \{\boldsymbol{d}_\lambda \ \boldsymbol{d}_\mu\}^\top$ denote the vector that contains the unknown coefficients. On selecting trial functions of the form (59) and choosing admissible variations $\delta\lambda = \phi_i^\lambda$ and $\delta\mu = \phi_i^\mu$, and substituting them in (55), we obtain the following system of linear

equations:

$$\boldsymbol{K}\boldsymbol{d} = \boldsymbol{f}, \quad \boldsymbol{K} = \begin{bmatrix} \boldsymbol{K}_{\lambda\lambda} & \boldsymbol{K}_{\lambda\mu} \\ \boldsymbol{K}_{\mu\lambda} & \boldsymbol{K}_{\mu\mu} \end{bmatrix}, \quad \boldsymbol{f} = \begin{Bmatrix} \boldsymbol{f}_{\lambda} \\ \boldsymbol{f}_{\mu} \end{Bmatrix}, \tag{60a}$$

$$\begin{aligned}
K^{ij}_{\lambda\lambda} &= \int_0^1 \int_0^1 \left[\frac{\partial \phi^{\lambda}_j}{\partial t} \frac{\partial \phi^{\lambda}_i}{\partial t} + \left(\alpha \phi^{\lambda}_j + \kappa \frac{\partial \phi^{\lambda}_j}{\partial x} \right) \left(\alpha \phi^{\lambda}_i + \kappa \frac{\partial \phi^{\lambda}_i}{\partial x} \right) \right] dx\, dt \Rightarrow \boldsymbol{K}_{\lambda\lambda} = \int_0^1 \int_0^1 \left[\frac{\partial \boldsymbol{N}^{\top}_{\lambda}}{\partial t} \frac{\partial \boldsymbol{N}_{\lambda}}{\partial t} + \boldsymbol{C}^{\top}_{\lambda} \boldsymbol{C}_{\lambda} \right] dx\, dt, \\
K^{ij}_{\lambda\mu} &= \int_0^1 \int_0^1 \left[\frac{\partial \phi^{\mu}_j}{\partial x} \frac{\partial \phi^{\lambda}_i}{\partial t} - \phi^{\mu}_j \left(\alpha \phi^{\lambda}_i + \kappa \frac{\partial \phi^{\lambda}_i}{\partial x} \right) \right] dx\, dt \Rightarrow \boldsymbol{K}_{\lambda\mu} = \int_0^1 \int_0^1 \left[\frac{\partial \boldsymbol{N}^{\top}_{\lambda}}{\partial t} \frac{\partial \boldsymbol{N}_{\mu}}{\partial x} - \boldsymbol{C}^{\top}_{\lambda} \boldsymbol{N}_{\mu} \right] dx\, dt, \\
K^{ij}_{\mu\lambda} &= \int_0^1 \int_0^1 \left[\frac{\partial \phi^{\lambda}_j}{\partial t} \frac{\partial \phi^{\mu}_i}{\partial x} - \left(\alpha \phi^{\lambda}_j + \kappa \frac{\partial \phi^{\lambda}_j}{\partial x} \right) \phi^{\mu}_i \right] dx\, dt \Rightarrow \boldsymbol{K}_{\mu\lambda} = \int_0^1 \int_0^1 \left[\frac{\partial \boldsymbol{N}^{\top}_{\mu}}{\partial x} \frac{\partial \boldsymbol{N}_{\lambda}}{\partial t} - \boldsymbol{N}^{\top}_{\mu} \boldsymbol{C}_{\lambda} \right] dx\, dt, \\
K^{ij}_{\mu\mu} &= \int_0^1 \int_0^1 \left[\frac{\partial \phi^{\mu}_j}{\partial x} \frac{\partial \phi^{\mu}_i}{\partial x} + \phi^{\mu}_j \phi^{\mu}_i \right] dx\, dt \Rightarrow \boldsymbol{K}_{\mu\mu} = \int_0^1 \int_0^1 \left[\frac{\partial \boldsymbol{N}^{\top}_{\mu}}{\partial x} \frac{\partial \boldsymbol{N}_{\mu}}{\partial x} + \boldsymbol{N}^{\top}_{\mu} \boldsymbol{N}_{\mu} \right] dx\, dt,
\end{aligned} \tag{60b}$$

$$\begin{aligned}
f^i_{\lambda} &= -\int_0^1 u_0(x) \phi^{\lambda}_i(x,0)\, dx \Rightarrow \boldsymbol{f}_{\lambda} = -\int_0^1 u_0(x) \boldsymbol{N}^{\top}_{\lambda}(x,0)\, dx, \\
f^i_{\mu} &= \int_0^1 \left[\bar{u}_2 \phi^{\mu}_i(1,t) - \bar{u}_1 \phi^{\mu}_i(0,t) \right] dt \Rightarrow \boldsymbol{f}_{\mu} = \int_0^1 \left[\bar{u}_2 \boldsymbol{N}^{\top}_{\mu}(1,t) - \bar{u}_1 \boldsymbol{N}^{\top}_{\mu}(0,t) \right] dt,
\end{aligned} \tag{60c}$$

where

$$\boldsymbol{C}_{\lambda} = \alpha \boldsymbol{N}_{\lambda} + \kappa \frac{\partial \boldsymbol{N}_{\lambda}}{\partial x}. \tag{60d}$$

Observe that $\boldsymbol{K}$ is symmetric since $\boldsymbol{K}_{\mu\lambda} = \boldsymbol{K}^{\top}_{\lambda\mu}$. Since the admissible trial functions for the dual fields in (59) consist of linearly independent basis functions, it can be seen that the uniqueness proof in Section 4.1 translates to uniqueness of solutions for the discrete problem. Hence, $\boldsymbol{K}$ is invertible and the system of linear equations, $\boldsymbol{K}\boldsymbol{d} = \boldsymbol{f}$, has a unique solution. It is interesting to observe that because of this invertibility of $\boldsymbol{K}$, the quadratic form, $\boldsymbol{d}^{\top}\boldsymbol{K}\boldsymbol{d}$, of the discrete problem obtained by the substitution $(\lambda_d, \mu_d) \to (\lambda_\theta, \mu_\theta)$ on the left-hand side of (56) is positive definite, even though the corresponding PDE is not elliptic, in general.

5. Numerical examples

We apply the dual variational principle to solve the Laplace equation, steady-state one-dimensional convection-diffusion equations, transient convection-diffusion equation, and the transient heat equation. The transient problems are solved as a space-time Galerkin method with a terminal boundary condition. In previous applications of the dual formulation [12, 17, 18], the dual variables have been approximated using linear C^0 finite elements to solve linear and nonlinear PDEs. From (44b), since the primal field u depends on the space-time derivatives of (λ, μ), an L^2 projection of the primal fields u_H onto linear finite elements was carried out to obtain a C^0-continuous solution. Herein, we adopt smooth approximations for the dual fields so that they directly provide at least C^0 primal fields. In one dimension, RePU activation function on a shallow network and univariate B-splines are adopted as approximating functions, whereas tensor-product B-splines are used for two-dimensional (space-time) problems. The numerical implementation

is carried out in Matlab™ (R2024a Release).

5.1. Laplace equation

Consider the Laplace equation, $u'' = 0$ in $\Omega = (0, 1)$, with boundary conditions $u(0) = 0$, $(1) = 1$. The exact solution is: $u(x) = x$. From (49a), we deduce that a quadratic approximation for $\mu(x)$ and a cubic approximation for $\lambda(x)$ should recover the exact solution. As a verification test, we choose the following ansatz:

$$\mu_\theta(x) = a_0 + a_1 x + a_2 x^2, \quad \lambda_\theta(x) = x(1-x)(b_0 + b_1 x),$$

which are kinematically admissible since $\lambda(x)$ vanishes at $x = 0$ and $x = 1$.

On setting $\kappa = 1$, $\alpha = 0$, $\bar{u}_1 = 0$, $\bar{u}_2 = 1$, and dropping the time dependence in (60), we obtain the expressions for the stiffness matrix and force vector. The system of linear equations and its solution are:

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & 0 & 0 \\ \frac{1}{2} & \frac{4}{3} & \frac{5}{4} & \frac{1}{6} & \frac{1}{12} \\ \frac{1}{3} & \frac{5}{4} & \frac{23}{15} & \frac{1}{6} & \frac{1}{10} \\ 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ 0 & \frac{1}{12} & \frac{1}{10} & \frac{1}{6} & \frac{2}{15} \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \\ b_0 \\ b_1 \end{Bmatrix} = \begin{Bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{Bmatrix} \quad \Rightarrow \quad \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \\ b_0 \\ b_1 \end{Bmatrix} = \begin{Bmatrix} \frac{5}{6} \\ 0 \\ \frac{1}{2} \\ -\frac{1}{6} \\ -\frac{1}{6} \end{Bmatrix}.$$

The solution for the dual fields are:

$$\mu_\theta(x) = \frac{5}{6} + \frac{1}{2}x^2, \quad \lambda_\theta(x) = -\frac{1}{6}x(1-x^2),$$

and the solution for the primal fields are:

$$u_\theta(x) = \mu'_\theta(x) = x, \quad q_\theta(x) = \mu_\theta(x) - \lambda'_\theta(x) = 1,$$

and therefore the exact solution for u and $q = u'$ are recovered.

To demonstrate that changing the boundary conditions on $\lambda(x)$ does not affect the numerical solution, we add an arbitrary cubic function to the earlier choice for $\lambda(x)$. Here, the ansatz for the dual fields are:

$$\mu_\theta(x) = a_0 + a_1 x + a_2 x^2. \quad \lambda(x) = x(1-x)(b_0 + b_1 x) + 1 - 3x + x^2 + 3x^3.$$

The dual field $\lambda(x)$ satisfies the Dirichlet boundary conditions $\lambda(0) = 1$ and $\lambda(1) = 2$. Now, the system of linear

equations and its solution are:

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & 0 & 0 \\ \frac{1}{2} & \frac{4}{3} & \frac{5}{4} & \frac{1}{6} & \frac{1}{12} \\ \frac{1}{3} & \frac{5}{4} & \frac{23}{15} & \frac{1}{6} & \frac{1}{10} \\ 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ 0 & \frac{1}{12} & \frac{1}{10} & \frac{1}{6} & \frac{2}{15} \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \\ b_0 \\ b_1 \end{Bmatrix} = \begin{Bmatrix} 2 \\ \frac{29}{12} \\ \frac{23}{10} \\ \frac{11}{6} \\ \frac{16}{15} \end{Bmatrix} \quad \Rightarrow \quad \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \\ b_0 \\ b_1 \end{Bmatrix} = \begin{Bmatrix} \frac{11}{6} \\ 0 \\ \frac{1}{2} \\ \frac{23}{6} \\ \frac{17}{6} \end{Bmatrix}.$$

The solution for the dual fields are:

$$\mu_\theta(x) = \frac{11}{6} + \frac{1}{2}x^2, \quad \lambda_\theta(x) = \frac{x(1-x)(23+17x)}{6} + 1 - 3x + x^2 + 3x^3,$$

and the solution for the primal fields are:

$$u_\theta(x) = \mu'_\theta(x) = x, \quad q_\theta(x) = \mu_\theta(x) - \lambda'_\theta(x) = 1,$$

and hence once again the exact solutions for u and u' are recovered.

5.2. Steady-state one-dimensional convection-diffusion problem

Consider the following one-dimensional convection-diffusion model problem [33]:

$$u'' - \alpha u' = 0 \ \text{ in } \Omega = (0,1) \tag{61a}$$

$$u(0) = 0, \ \ u(1) = 1, \tag{61b}$$

with exact solution

$$u(x) = \frac{e^{\alpha x} - 1}{e^{\alpha} - 1}, \tag{61c}$$

where α is the Peclet number (ratio of convective rate to diffusive rate). As the Peclet number increases, a boundary layer develops near the right end $x = 1$.

Remark: As noted earlier, in the strongly convective regime, standard Galerkin methods produce spurious oscillations [44]. On using the dual variational form, we can use a standard Galerkin method to compute accurate solutions.

The dual functional for this problem is given in (50). On setting $\kappa = 0$ and dropping the time dependence in (60),

we can write the stiffness matrix (symmetric) and the force vector as:

$$
\begin{aligned}
\boldsymbol{K} &= \begin{bmatrix} \boldsymbol{K}_{\lambda\lambda} & \boldsymbol{K}_{\lambda\mu} \\ \boldsymbol{K}_{\mu\lambda} & \boldsymbol{K}_{\mu\mu} \end{bmatrix}, \quad \boldsymbol{f} = \begin{Bmatrix} \boldsymbol{f}_{\lambda} \\ \boldsymbol{f}_{\mu} \end{Bmatrix}, \\
\boldsymbol{K}_{\lambda\lambda} &= \int_0^1 \alpha^2 \boldsymbol{N}_{\lambda}^{\top} \boldsymbol{N}_{\lambda}\, dx, \quad \boldsymbol{K}_{\lambda\mu} = -\int_0^1 \alpha \boldsymbol{N}_{\lambda}^{\top} \boldsymbol{N}_{\mu}\, dx, \\
\boldsymbol{K}_{\mu\lambda} &= -\int_0^1 \alpha \boldsymbol{N}_{\mu}^{\top} \boldsymbol{N}_{\lambda}\, dx, \quad \boldsymbol{K}_{\mu\mu} = \int_0^1 \left[\frac{\partial \boldsymbol{N}_{\mu}^{\top}}{\partial x} \frac{\partial \boldsymbol{N}_{\mu}}{\partial x} + \boldsymbol{N}_{\mu}^{\top} \boldsymbol{N}_{\mu} \right] dx, \\
\boldsymbol{f}_{\lambda} &= \boldsymbol{0}, \quad \boldsymbol{f}_{\mu} = \boldsymbol{N}_{\mu}^{\top}(1),
\end{aligned}
\tag{62}
$$

where $\boldsymbol{N}_{\lambda}$ and $\boldsymbol{N}_{\mu}$ define row vectors that contain the basis used to form $\lambda_{\theta}(x)$ and $\mu_{\theta}(x)$, respectively. We present numerical solutions using a shallow neural network with RePU activation function and univariate B-splines.

5.2.1. Neural network solutions

The RePU (ReLU^k) activation function [34] in a neural network is given by

$$
\sigma(x; p) = \texttt{RePU}(x; p) = \texttt{ReLU}^p(x) = [\max(0, x)]^p. \tag{63}
$$

It is known that deep neural networks with ReLU and ReLU^2 activation functions can represent Lagrange finite elements of any order in arbitrary dimensions [35]. Here we consider a shallow neural network (single hidden layer) with $2n$ neurons. The interval $[0, 1]$ is discretized with the knot sequence:

$$
\Xi := [x_0,\ x_1, \ldots,\ x_n], \tag{64}
$$

where $x_0 = 0$ and $x_n = 1$. The bias and weight are fixed in the hidden layer. A linear output is used to define the two dual fields $\lambda_{\theta}(x)$ and $\mu_{\theta}(x)$.

Smooth approximations are chosen for the dual fields $\mu_{\theta}(x)$ (degree p) and $\lambda_{\theta}(x)$ (degree q), which are written in the form:

$$
\mu_{\theta}(x) = \sum_{i=0}^{n-1} \sigma(x - x_i; p) a_i^{+} + \sum_{i=1}^{n} \sigma(x_i - x; p) a_i^{-} := \sum_{k=1}^{2n} \phi_k^{\mu}(x)\, \boldsymbol{a}_k \ \in C^{p-1}(\Omega), \tag{65a}
$$

$$
\begin{aligned}
\lambda_{\theta}(x) &= (1 - x)\lambda_{\theta}^{+}(x) + x\lambda_{\theta}^{-}(x) \\
&= (1 - x)\left[\sum_{i=0}^{n-1} \sigma(x - x_i; q)\, b_i^{+} \right] + x\left[\sum_{i=1}^{n} \sigma(x_i - x; q)\, b_i^{-} \right] := \sum_{k=1}^{2n} \phi_k^{\lambda}(x)\, \boldsymbol{b}_k \ \in C^{q-1}(\Omega) \ \ (q \geq 2).
\end{aligned}
\tag{65b}
$$

The shallow neural network approximations that appear in (65) are atypical in the neural network literature. Instead of nonlinear approximations with both bias and weights that are unknowns, we fix the bias and weight ($z = wx + b$, $w = \pm 1$, $b = \mp x_i$) in the hidden layer so that the approximations are linear in the unknowns (weights) from the output layer. The two RePU functions with arguments $x - x_i$ and $x_i - x$ that are associated with all interior knot locations $\{x_i\}_{i=1}^{n-1}$ are clearly linearly independent. For $q \geq 2$ in (65b), the function associated with $i = 0$ in the first term and

the function associated with $i = n$ in the second term are both positive in $(0, 1)$, and linearly independent.[4] As a consequence, use of the trial functions (65) in a Ritz minimization procedure lead to a system of linearly independent equations. Observe that $\lambda_\theta(x)$ in (65b) satisfies the Dirichlet boundary conditions since $\lambda_\theta(0) = \lambda_\theta(1) = 0$. Plots of RePU functions that are used to construct $\mu_\theta(x)$ ($p = 2$) and $\lambda_\theta(x)$ ($q = 3$) are presented in Fig. 2. These functions are polynomials on either side of the knot location x_i. The functions $\text{ReLU}^2(x - x_i)$ and $\text{ReLU}^2(x_i - x)$ in Figs. 2a and 2b, respectively, contribute to $\mu_\theta(x)$. A barycentric convex combination of the contributions due to $\text{ReLU}^3(x - x_i)$ and $\text{ReLU}^3(x_i - x)$ in Figs. 2c and 2d, respectively, are used to form $\lambda_\theta(x)$.

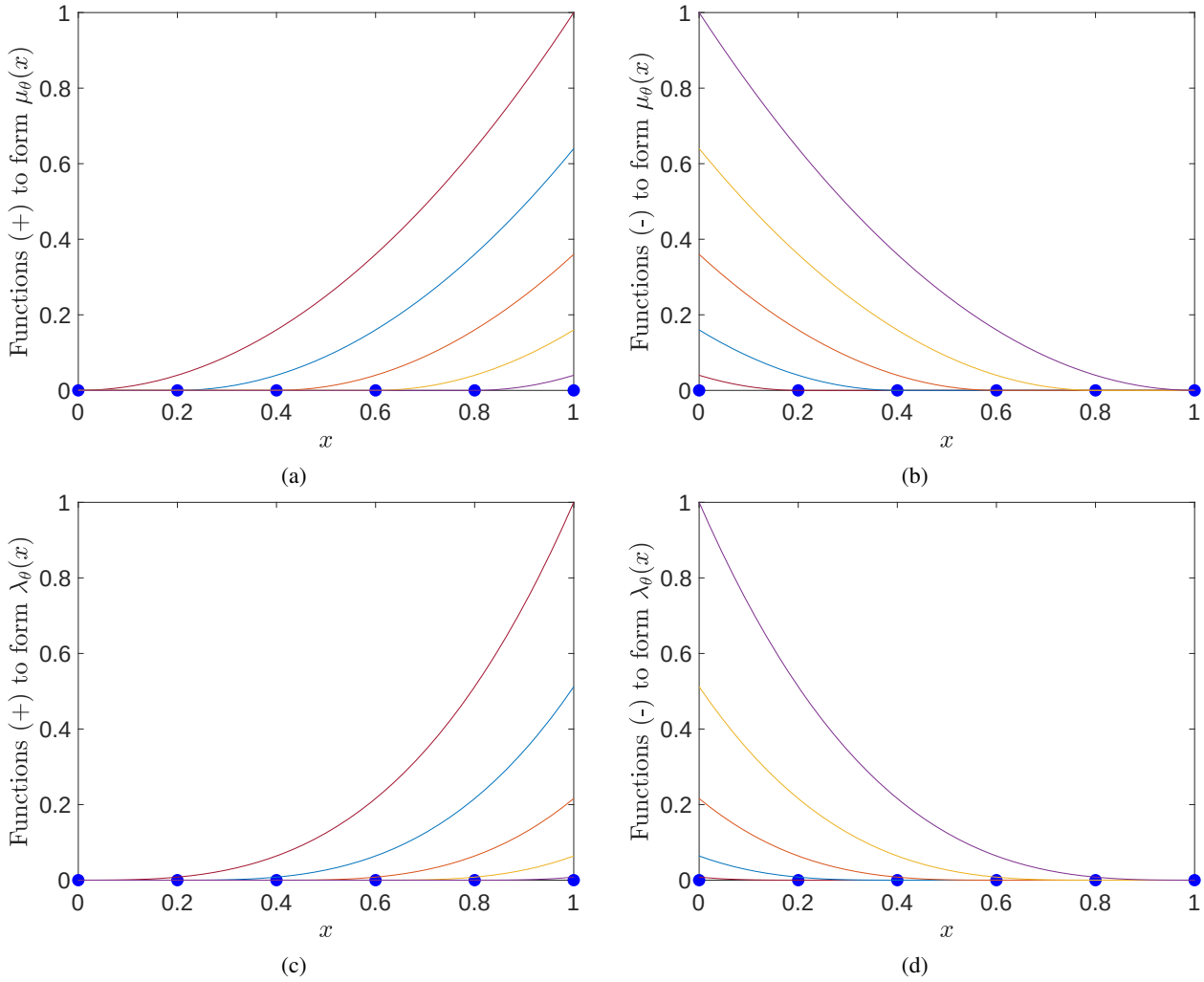


Figure 2: Plots of RePU functions that are used to form the dual fields. Filled circles in blue denote the location of the knots, x_i ($i = 0, 1, \ldots, 5$). Knots are uniformly spaced. Linear combination of the functions (a) $\sigma(x - x_i; p = 2)$ and (b) $\sigma(x_i - x; p = 2)$ are used to form $\mu_\theta(x)$. Linear combination of the functions (c) $\sigma(x - x_i; q = 3)$ are used to form $\lambda_\theta^+(x)$ and linear combination of the functions (d) $\sigma(x_i - x; q = 3)$ are used to form $\lambda_\theta^-(x)$. The dual field $\lambda_\theta(x) = (1 - x)\lambda_\theta^+(x) + x\lambda_\theta^-(x)$.

For $\alpha = 1$, the neural network solution is compared to the exact solution in Fig. 3. In the neural network compu-

[4]For $q = 1$, the two functions in question are identical.

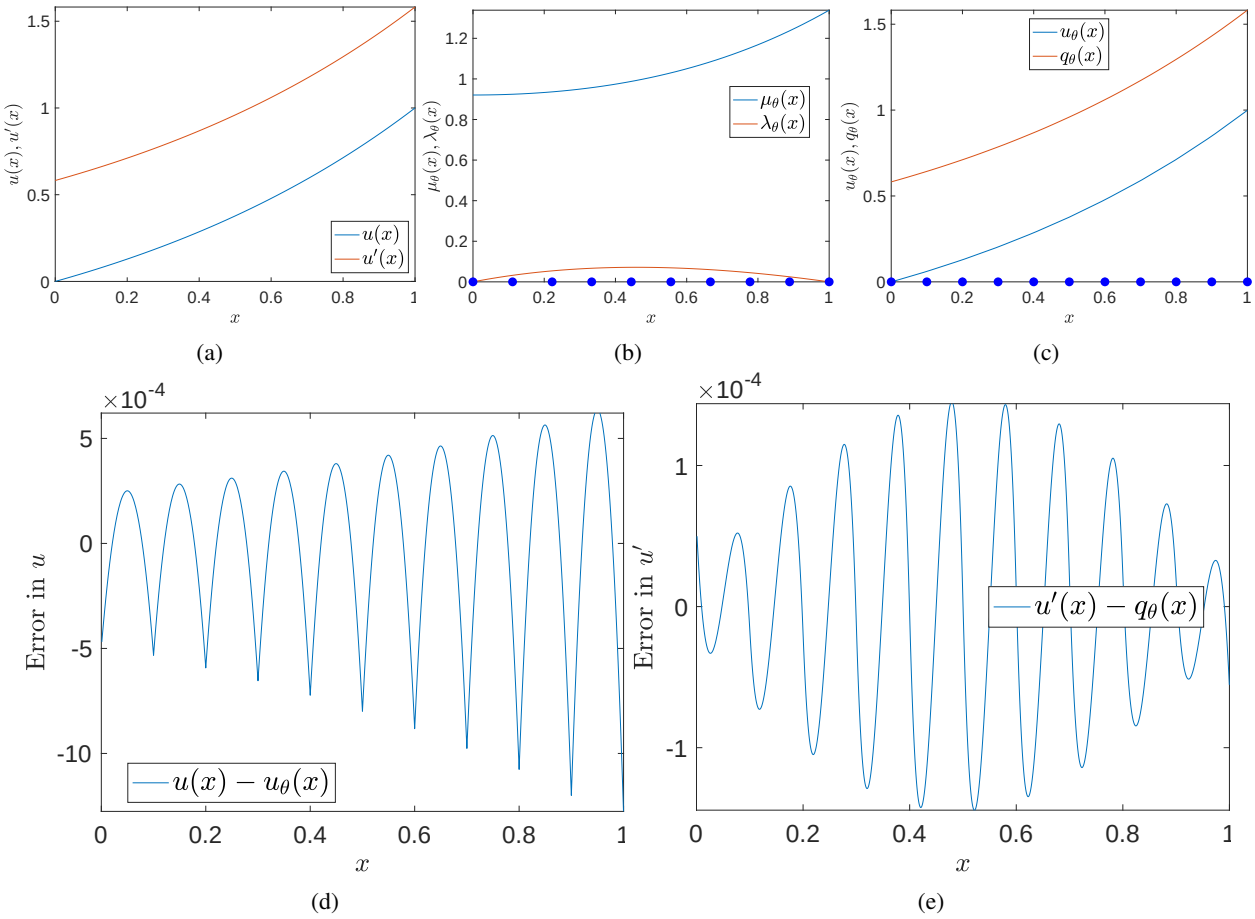
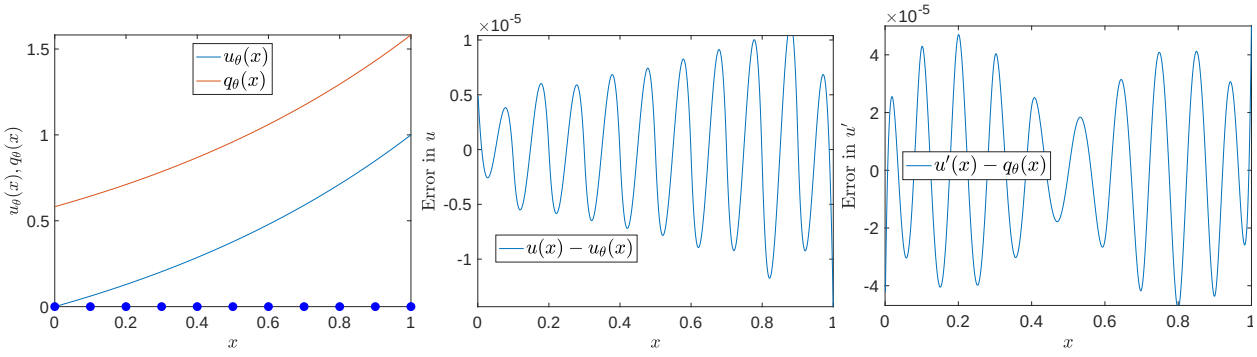


Figure 3: Neural network solution for the steady-state convection-diffusion problem ($\alpha = 1$) using $n = 10$, $p = 2$ and $q = 3$. (a) Exact solutions for u and u'; (b) Dual fields, $\mu_\theta(x)$ and $\lambda_\theta(x)$; (c) Primal fields, $u_\theta(x)$ and $q_\theta(x)$; (d) Error in u; and (e) Error in u'.

tations, eleven distinct knot locations are used ($n = 10$). Degrees $p = 2$ are $q = 3$ are chosen to form $\mu_\theta(x)$ and $\lambda_\theta(x)$, respectively. The dual fields are presented in Fig. 3b and the primal fields, which are computed from the dual fields using the DtP map in (50a), are shown in Fig. 3c. We observe that the maximum errors in Figs. 3d and 3e for u and u' are about 10^{-3} and 10^{-4}, respectively. Similar trends are observed in the results presented in Fig. 4 for $\mu_\theta(x)$ ($p = 3$) and $\lambda_\theta(x)$ ($q = 4$). The maximum errors in Figs. 4b and 4c for u and u' are 10^{-5} and 4×10^{-5}, respectively. In the results shown in both Figs. 3 and 4, the accuracy of $q_\theta = u'_\theta(x)$ is seen to be comparable to that of $u_\theta(x)$. This is so since the degree of the piecewise polynomial $\lambda_\theta(x)$ (contributes to $q_\theta = u'_\theta$, which is the constraint equation that is weakly enforced) is one greater than that of $\mu_\theta(x)$ ($u_\theta = \mu'_\theta$). Next we set $\alpha = 10$ and choose $n = 30$ in the numerical discretization. The neural network solution is compared to the exact solution in Fig. 5 for $\mu_\theta(x)$ ($p = 2$) and $\lambda_\theta(x)$ ($q = 3$). The exact solution displays a sharp ascent close to $x = 1$ in Fig. 5a. The maximum errors in Figs. 5c and 5d for u and u' are about 6×10^{-3} and 7×10^{-5}, respectively. A convergence study is conducted using $n = [2, 4, 8, 16, 32, 64]$ and three different choices for p and q: $p = 2$, $q = 3$; $p = 2$, $q = 4$; and $p = 3$, $q = 4$. The relative L^2 norm and relative H^1



(a) (b) (c)

Figure 4: Neural network solution for the steady-state convection-diffusion problem ($\alpha = 1$) using $n = 10$, $p = 3$, and $q = 4$. (a) $u_\theta(x)$, $q_\theta(x)$; (b) $u - u_\theta$; and (c) $u' - q_\theta$.

seminorm of the error $u - u_\theta$ are defined as:

$$E_u = \sqrt{\frac{\int_0^1 (u - u_\theta)^2 \, dx}{\int_0^1 u^2 \, dx}}, \qquad E_q = \sqrt{\frac{\int_0^1 (u' - q_\theta)^2 \, dx}{\int_0^1 (u')^2 \, dx}}.$$

Plots of the relative errors versus the number of degrees of freedom are presented in Fig. 6. The rates of convergence (slopes) in (u, u') are $(2, 3)$, $(2, 3.2)$ and $(2.9, 4)$.

Finally, we set $\alpha = 50$, and the solution u develops a sharp boundary layer in the vicinity of $x = 1$. The neural network solution is compared to the exact solution in Fig. 7. We choose $n = 50$, and $\mu_\theta(x)$ ($p = 2$ and $\lambda_\theta(x)$ ($q = 3$) are selected as piecewise quadratic and piecewise cubic polynomials, respectively. The maximum errors in Figs. 7c and 7d for u and u' are about 6×10^{-2} and 1.5×10^{-2}, respectively. The larger errors in u and u' are due to the steep gradient of u in the vicinity of $x = 1$, which can be captured by either increasing n (h-refinement) or using a high-order approximation (p-refinement). Observe that $u \in (0, 1]$ and $u' \in (0, 50]$ for $x \in [0.8, 1]$. We revisit this problem using B-splines in the next section.

5.2.2. Univariate B-spline solutions

We adopt a univariate B-spline approximation [36] to represent the dual fields $\mu(x)$ and $\lambda(x)$. Unlike the RePU functions used in the previous section, B-splines are compactly-supported and yield better-conditioned system matrices. An open uniform knot vector is used:

$$\Xi := [\underbrace{0, 0, \ldots, 0}_{p+1}, x_1, \ldots, x_{n-1}, \underbrace{1, 1, \ldots, 1}_{p+1}], \tag{66}$$

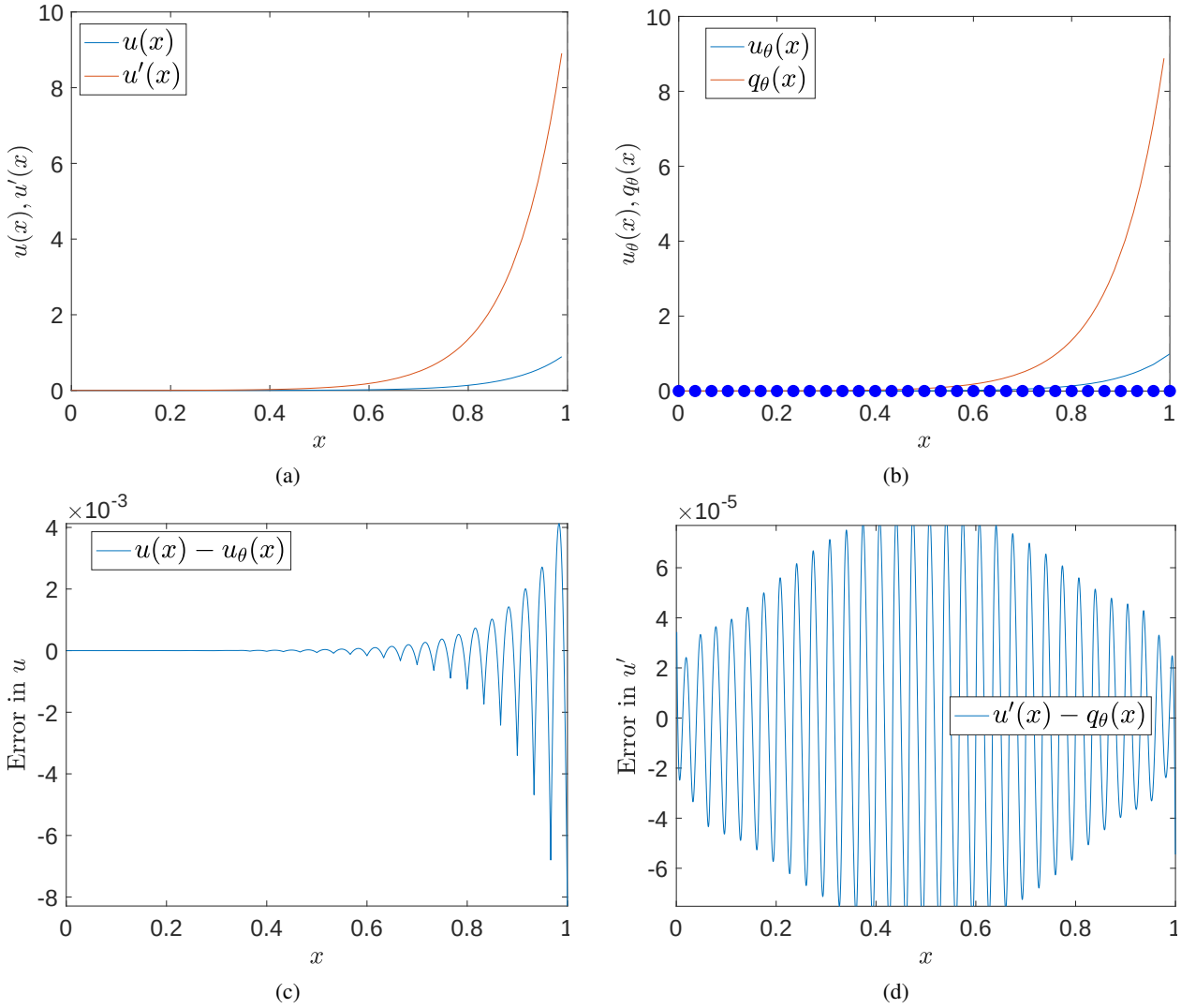


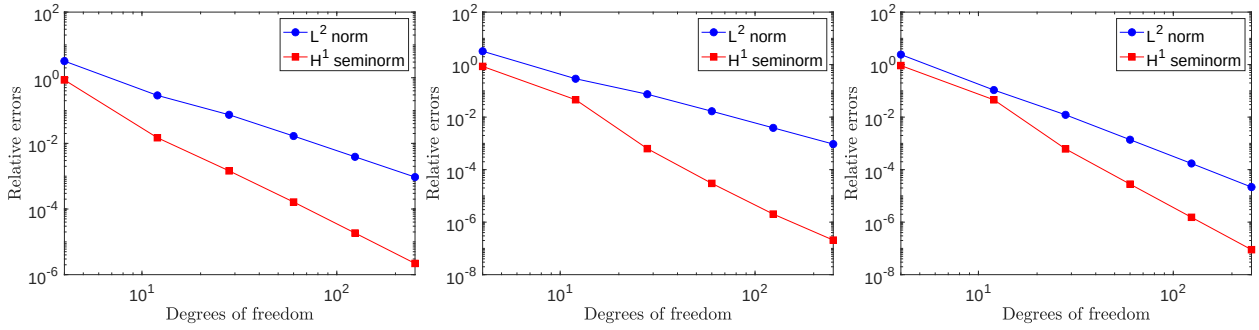
Figure 5: Neural network solution for the steady-state convection-diffusion problem ($\alpha = 10$) using $n = 30$, $p = 2$ and $q = 3$. (a) Exact solutions; (b) Neural network solutions; (c) $u - u_\theta$; and (d) $u' - q_\theta$.

where the end knot points are repeated $p + 1$ times so that C^0 continuity and interpolation is ensured at $x = 0$ and $x = 1$. B-splines are used to construct smooth approximations for the dual fields $\mu(x)$ and $\lambda(x)$:

$$\mu_\theta(x) = \sum_{i=1}^{n+p} B_i^p(x) a_i \in C^{p-1}(\Omega), \tag{67a}$$

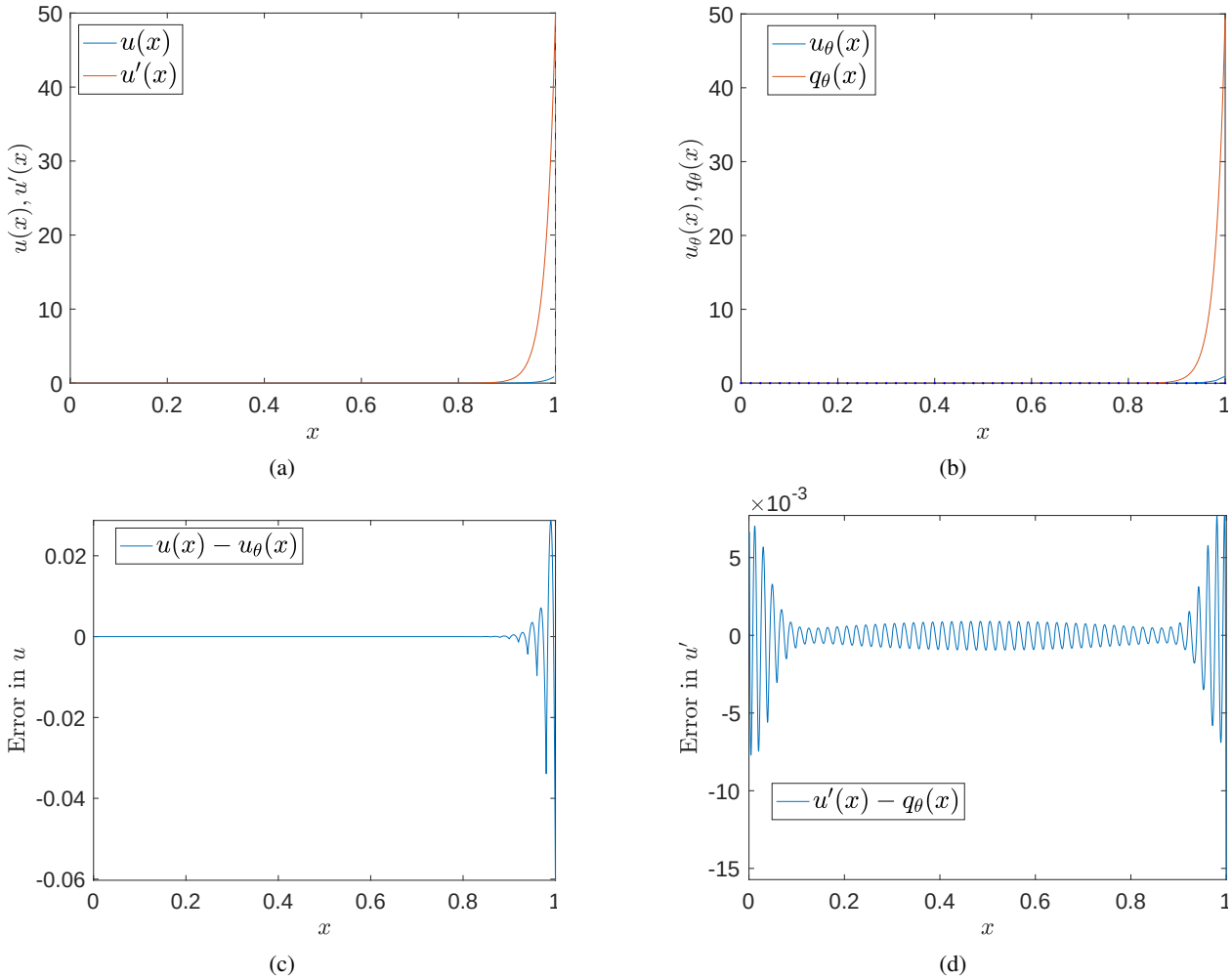
$$\lambda_\theta(x) = \sum_{i=1}^{n+q} B_i^q(x) b_i \in C^{q-1}(\Omega), \tag{67b}$$

where $B_i^p(x)$ is the p-th degree B-spline basis function with control points ($\boldsymbol{a}$ and $\boldsymbol{b}$) as the unknowns. B-spline computations are performed using the Cox–de Boor algorithm. Note that on setting $b_1 = b_{n+q} = 0$ in (67b), we have $\lambda_\theta(0) = \lambda_\theta(1) = 0$, which ensures that the trial functions are admissible. In one dimension, the approximations in (67) can be viewed as a shallow network within the KAN architecture [29].



(a) (b) (c)

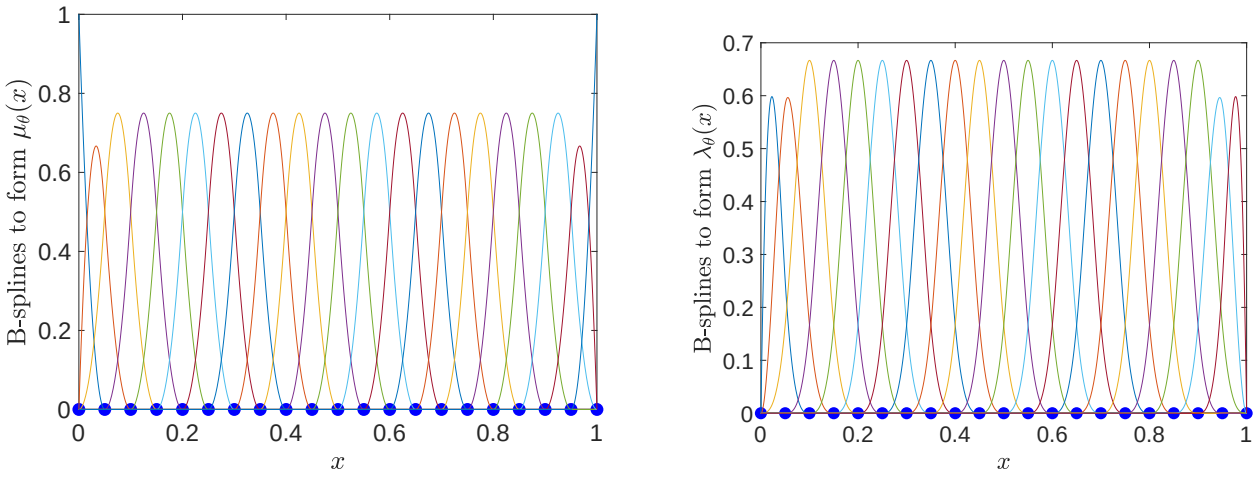
Figure 6: Convergence study with neural network approximants for the steady-state convection-diffusion problem ($\alpha = 10$). The dual field $\mu_\theta(x)$ is approximated using RePU functions of degree p and the dual field $\lambda_\theta(x)$ is formed by barycentric convex combination of RePU functions of degree q. (a) $p = 2$ and $q = 3$; (b) $p = 2$ and $q = 4$; and (c) $p = 3$ and $q = 4$. The rates of convergence in u and u' are of order p and $p + 1$, respectively.



(a) (b) (c) (d)

Figure 7: Neural network solution for the steady-state convection-diffusion problem ($\alpha = 50$) using $n = 50$, $p = 2$ and $q = 3$. (a) Exact solutions; (b) Neural network solutions; (c) $u - u_\theta$; and (d) $u' - q_\theta$.

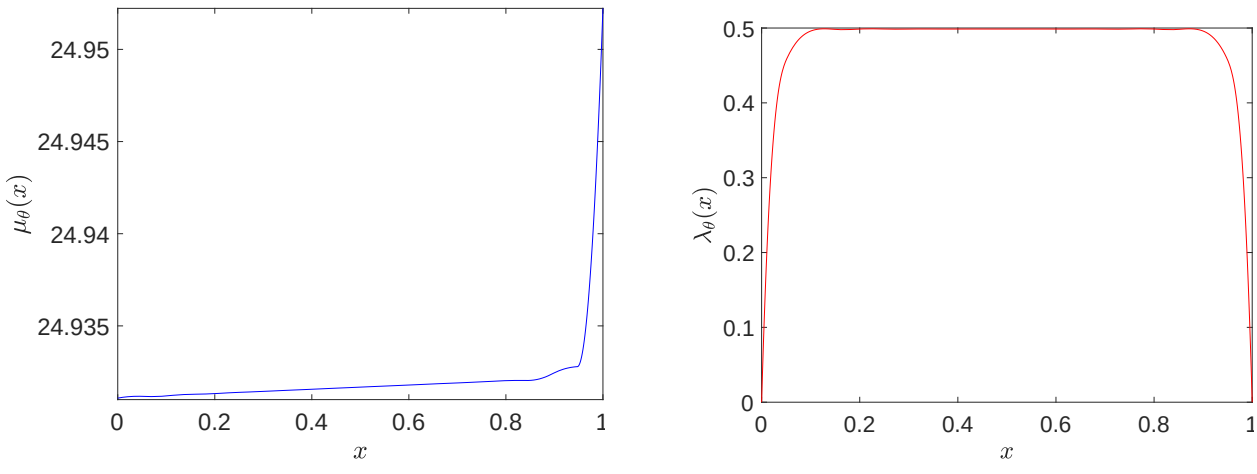
For $n = 20$, the nonzero B-spline basis functions that are used to form $\mu_\theta(x)$ ($p = 2$) and $\lambda_\theta(x)$ ($q = 3$) are presented in Fig. 8. We set the Peclet number α to 50. For the numerical computations, we choose $n = 20$, and vary p (degree of μ_θ) and q (degree of λ_θ). The B-spline basis functions to form $\mu_\theta(x)$ ($p = 2$) and $\lambda_\theta(x)$ ($q = 3$) are shown in Figs. 8a and 8b, respectively. For $p = 2$ and $q = 3$, the dual fields are presented in Fig. 9. The primal fields are computed from



(a) (b)

Figure 8: Plots of B-spline basis functions that are used to form the dual fields for $n = 20$. (a) $\mu_\theta(x)$ ($p = 2$) and (b) $\lambda_\theta(x)$ ($q = 3$).

the dual fields using the DtP map in (50a). The exact solution for u and u' is presented in Fig. 7a. The errors in the B-spline solution for two choices of p and q are presented in Fig. 10. For $p = 2$ and $q = 3$, we see from Figs. 10a and 10b that the maximum error in u and u' are about 0.2 and 2, respectively. For $p = 5$ and $q = 6$, the maximum error in u and u' from Figs. 10c and 10d are 4×10^{-3} and 8×10^{-3}, respectively. For $n = 20$, B-spline basis functions to form



(a) (b)

Figure 9: B-spline solution (dual fields) of the steady-state convection-diffusion equation ($\alpha = 50$) for $n = 20$. (a) $\mu_\theta(x)$ ($p = 2$) and (b) $\lambda_\theta(x)$ ($q = 3$).

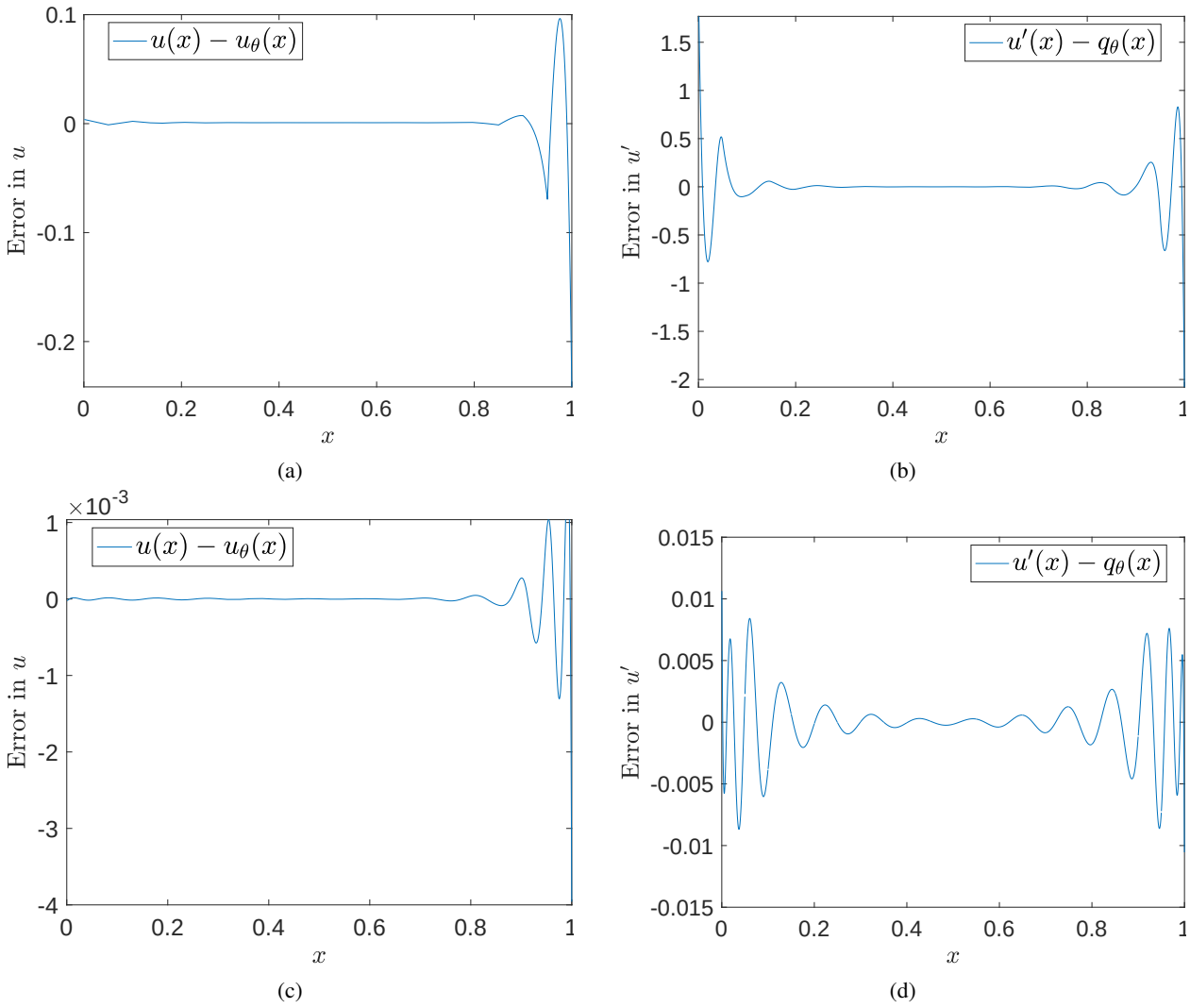


Figure 10: B-spline solution for the steady-state convection-diffusion problem ($\alpha = 50$). For $n = 20,\ p = 2,\ q = 3$: (a) $u - u_\theta$ and (d) $u' - q_\theta$. For $n = 20,\ p = 5,\ q = 6$: (c) $u - u_\theta$ and (d) $u' - q_\theta$.

$\mu_\theta(x)$ ($p = 7$) and $\lambda_\theta(x)$ ($q = 8$) are shown in Fig. 11a and Fig. 11b, respectively. The errors in u and u' are presented in Figs. 11c and 11d, respectively. The maximum errors in u and u' are 1.25×10^{-4} and 1.25×10^{-3}, respectively. Since $\|u\|_\infty = 1$ and $\|u'\|_\infty = 50$, the relative errors for u and u' that B-splines deliver are proximal.

We perform a convergence study with B-splines to assess the rate of convergence of the method. The dual fields $\mu_\theta(x) \in C^{p-1}(\Omega)$ and $\lambda_\theta(x) \in C^{q-1}(\Omega)$ are approximated using B-spline basis functions of degree p and q, respectively. The primal fields are obtained from the DtP map in (50a):

$$u_\theta(x) = \mu'_\theta(x), \quad q_\theta(x) = \mu_\theta(x) - \alpha\lambda_\theta(x) - \lambda'_\theta(x).$$

From the DtP map and the fact that B-spline approximations (degree p) possess p-th degree polynomial completeness, one can deduce that the error in u should decay as $O(h^p)$ (h is the mesh spacing), and the error in the flux u' should

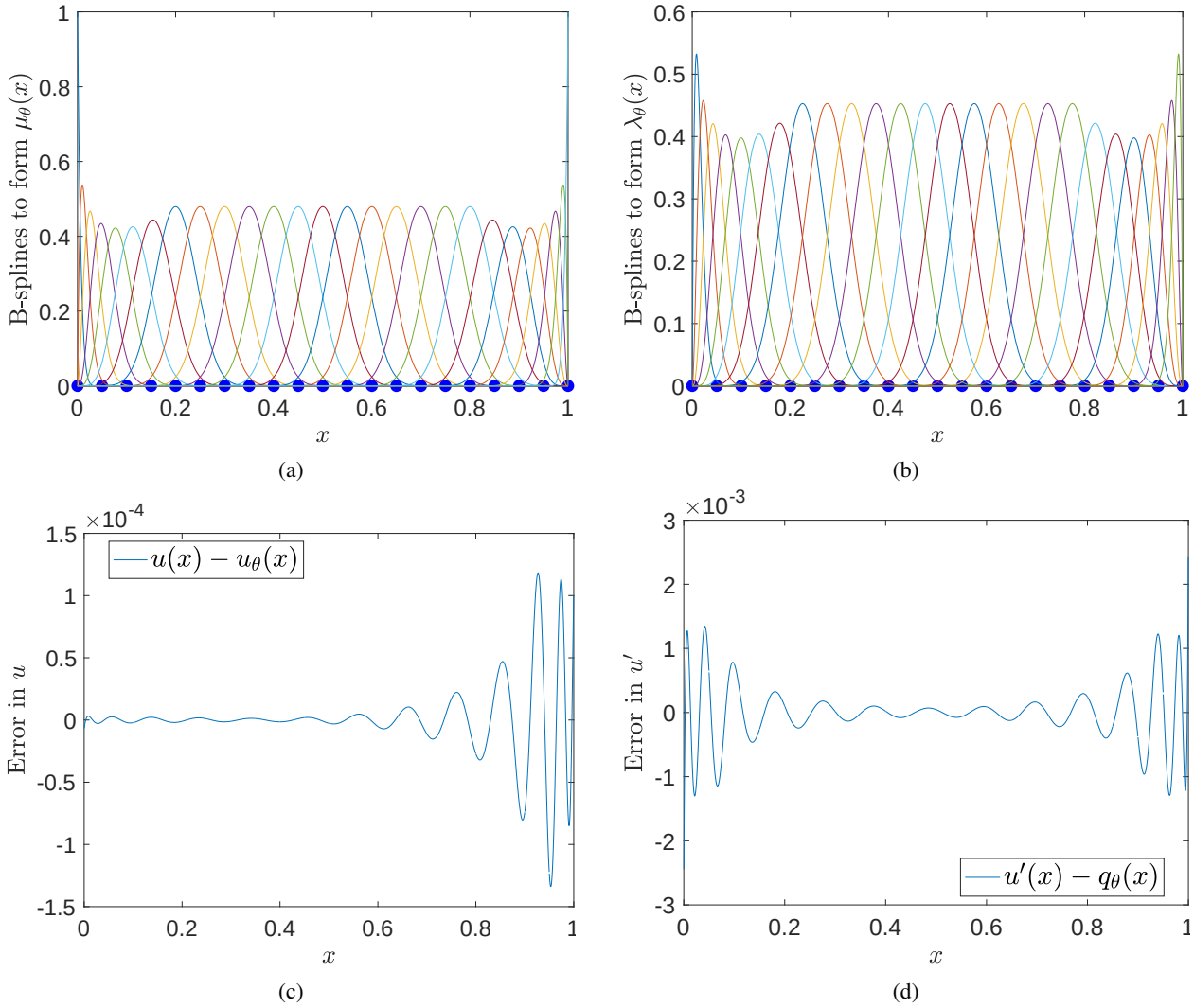


Figure 11: B-spline computations to solve the steady-state convection-diffusion problem ($\alpha = 50$). For $n = 20$, B-spline basis functions to form (a) $\mu_\theta(x)$ ($p = 7$) and (b) $\lambda_\theta(x)$ ($q = 8$); (c) $u - u_\theta$; and (d) $u' - q_\theta$.

decay as $O(h^p)$ if $q = p$ and $O(h^{p+1})$ if $q = p + 1$. A sequence of uniformly refined meshes with $n = [4, 8, 16, 32, 64]$ is selected. For $\alpha = 10$ and $\alpha = 50$, four different choices for p and q are considered: $p = q = 1$; $p = 1, q = 2$; $p = 2, q = 3$; and $p = 3, q = 4$. The total number of degrees of freedom is $2n + p + q - 2$. In Figs. 12 and 13, the plots of the relative errors versus the number of degrees of freedom are presented. Note that the convergence plots in Figs. 12a and 13a are for linear C^0 finite elements ($p = q = 1$), where primal fields that are discontinuous at the knot locations are used in the error norm computations. For the convergence plots shown in Figs. 12a–12d, the rates of convergence in (u, u') are found to be $(1, 1)$, $(1, 2)$, $(2.1, 3)$, $(3.1, 4.1)$, which is in agreement with a priori error estimates of $O(h^p)$ and $O(h^{p+1})$ ($q = p + 1$) in u and u', respectively. For the convergence plots shown in Figs. 13a–13d, the rates of convergence in (u, u') are found to be $(1.2, 0.9)$, $(0.9, 2)$, $(2, 3)$ and $(3, 3.7)$. We consistently observe that the B-spline solution for the derivative of u is more accurate than the B-spline solution for u. Due to the five-fold increase in α,

the relative errors in Fig. 13 are about two orders of magnitude less accurate than the corresponding results shown in Fig. 12.

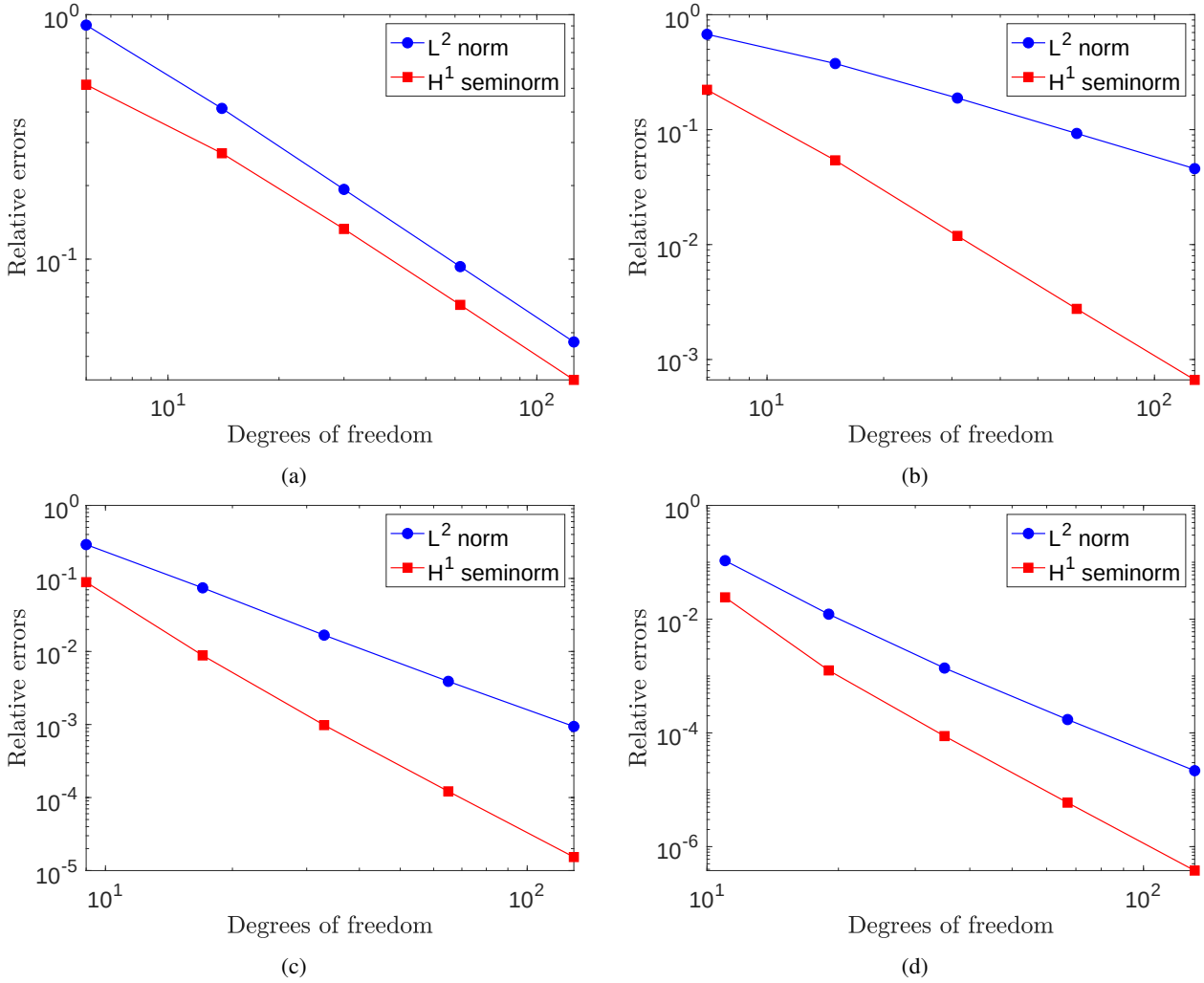


Figure 12: Convergence study with B-splines for the steady-state convection-diffusion problem ($\alpha = 10$). The dual fields $\mu_\theta(x)$ and $\lambda_\theta(x)$ are approximated using polynomials of degree (a) $p = 1$ and $q = 1$; (b) $p = 1$ and $q = 2$; (c) $p = 2$ and $q = 3$; and (d) $p = 3$ and $q = 4$. The rate of convergence in u and u' is p in (a), and are p and $p + 1$ in (b)–(d).

5.3. Transient convection-diffusion equation with B-splines

Consider the following transient convection-diffusion model problem:

$$\kappa \frac{\partial^2 u}{\partial x^2} - \alpha \frac{\partial u}{\partial x} = \frac{\partial u}{\partial t} \quad \text{in } \Omega = \Omega_0 \times \Omega_t = (0,1) \times (0,1), \tag{68a}$$

$$u(0,t) = 0, \quad u(1,t) = 0, \tag{68b}$$

$$u(x,0) = u_0(x), \tag{68c}$$

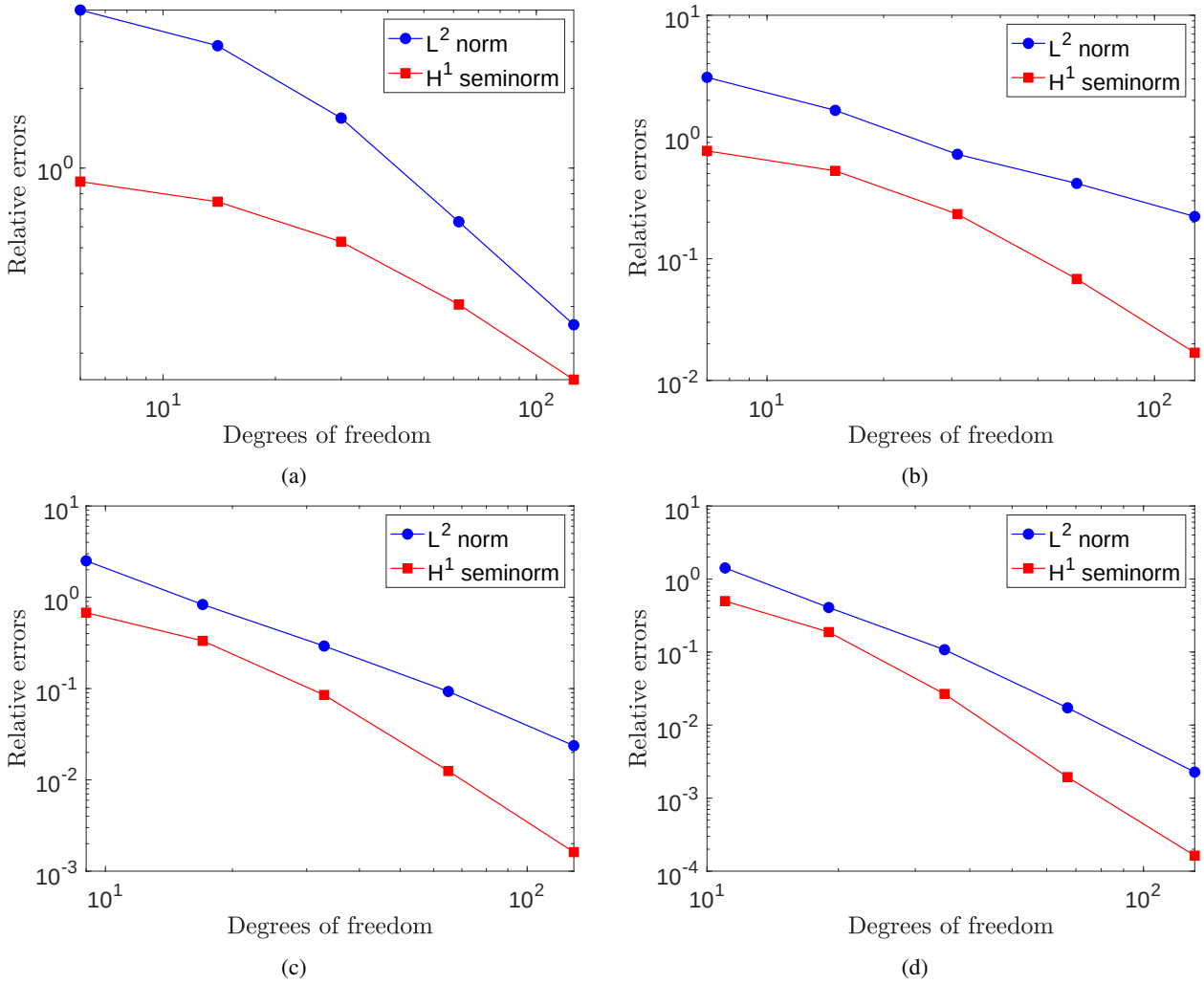


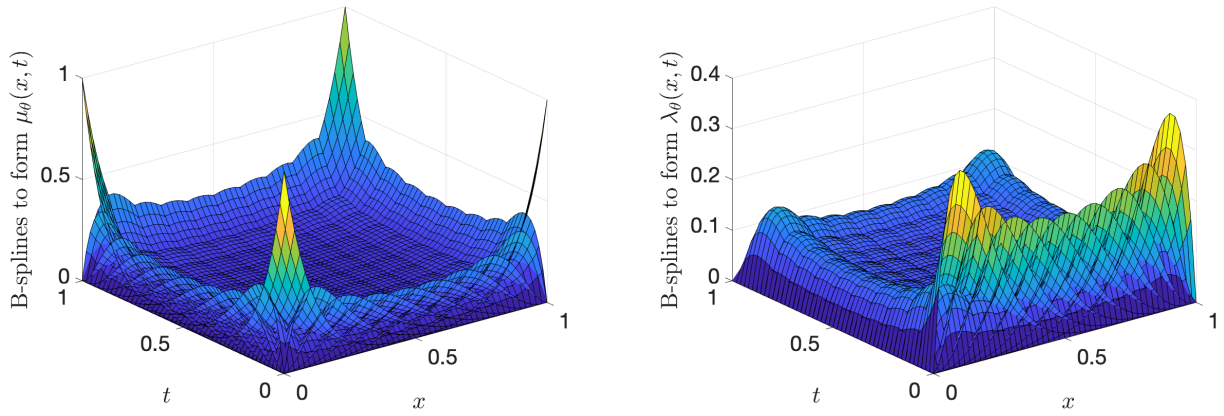
Figure 13: Convergence study with B-splines for the steady-state convection-diffusion problem ($\alpha = 50$). The dual fields $\mu_\theta(x)$ and $\lambda_\theta(x)$ are approximated using polynomials of degree (a) $p = 1$ and $q = 1$; (b) $p = 1$ and $q = 2$; (c) $p = 2$ and $q = 3$; and (d) $p = 3$ and $q = 4$. The rate of convergence in u and u' is p in (a), and are p and $p + 1$ in (b)–(d).

with the exact solution [45]:

$$\begin{aligned} u(x,t) &= \exp\left(-\frac{\alpha^2}{4\kappa}t\right)\exp\left(\frac{\alpha}{2\kappa}x\right)\sum_{n=1}^{\infty} b_n \sin(n\pi x)\exp\left(-\kappa n^2\pi^2 t\right), \\ b_n &= 2\int_0^1 \exp\left(-\frac{\alpha}{2\kappa}x\right) u_0(x)\,\sin(n\pi x)\,dx. \end{aligned} \tag{68d}$$

The ansatz for $\mu_\theta(x,t)$ and $\lambda_\theta(x,t)$ are formed by tensor products of univariate B-splines given in (67). The bivariate open knot vector is given by

$$\Xi \times \Xi, \quad \Xi := [\underbrace{0,0,\ldots,0}_{p+1}, x_1, x_2, \ldots, x_{n-1}, \underbrace{1,1,\ldots,1}_{p+1}]. \tag{69}$$



(a) (b)

Figure 14: Plots of two-dimensional B-spline basis functions to form the dual fields (a) $\mu_\theta(x,t)$ ($p = 9$) and (b) $\lambda_\theta(x,t)$ ($q = 10$) to solve the transient convection-diffusion problem. The bivariate open knot vector ($n = 1$) that is given in (69) is used.

where $x_0 = 0$ and $x_n = 1$, and the trial functions are

$$\mu_\theta(x,t) = \sum_{i=1}^{n+p}\sum_{j=1}^{n+p} B_i^p(x)\, B_j^p(t)\, a_{ij}, \qquad \lambda_\theta(x,t) = \sum_{i=1}^{n+q}\sum_{j=1}^{n+q} B_i^q(x)\, B_j^q(t)\, b_{ij}. \tag{70}$$

Appropriate coefficients are set to zero so that the Dirichlet boundary conditions on λ and μ are met. On setting $\bar{u}_1 = 0$ and $\bar{u}_2 = 0$ to meet the Dirichlet boundary conditions in (68b) and choosing $u_0(x) = \sin(2\pi x)$ for the initial condition in (68c), the expression for the stiffness matrix and force vector are obtained from (60).

In the numerical computations, we choose $\kappa = 0.01$ and $\alpha = 0.1$, so that advection dominates diffusion. Since we performed a single solve over the entire space-time domain, $\Omega = (0,1)^2$, we judiciously selected values for these constants so that u does not vary sharply in time and can be captured to modest accuracy by a high-order B-spline approximation. The exact solution in (68d) is computed with 1000 terms ($n = 1, 2, \ldots, 1000$), which satisfies the initial condition in (68c) to $O(10^{-7})$ accuracy.

The dual fields are constructed using bivariate tensor-product B-spline approximations. Plots of the two-dimensional basis functions for $\mu_\theta(x,t)$ (degree $p = 9$; 100 basis functions) and $\lambda_\theta(x,t)$ (degree $q = 10$; 90 basis functions) are presented in Fig. 14. Observe that $\lambda_\theta(0,t) = \lambda_\theta(x,1) = \lambda_\theta(x,1) = 0$ so that $\lambda_\theta \in S_\lambda$ and $\mu_\theta \in S_\mu$ satisfy (46c), and are therefore kinematically admissible. The B-spline solution for the dual fields, $\mu_\theta(x,t)$ and $\lambda_\theta(x,t)$, are presented in Fig. 15. The B-spline solution for u and $q := \partial u/\partial x$ are computed from the dual fields using the DtP map given in (44). These primal B-spline solutions are compared to the exact solutions in Fig. 16. We find that the B-spline solutions for u and q are in fair agreement with the exact solutions. Since $\|u\|_\infty = 1$ and $|\partial u/\partial x|_\infty \approx 10$, the relative maximum error in u and q are 0.06 and 0.1, respectively (see Figs. 16c and 16f). In addition, we observe that the errors grow larger close to the terminal time, $t = 1$. This is evident from the time history plots shown in Figs. 16g to 16j.

Notably the maximum error in u for $t \in [0, 1]$ (see Fig. 16g) is about six times that for $t \in [0, 0.9]$ (see Fig. 16h). The same trend is also noticed for the error in q from Figs. 16i and 16j. We attribute the larger errors that concentrate in the vicinity of $t = 1$ (terminal time) to the interactions between the boundary conditions in the primal problem and the terminal boundary condition in the dual problem, which may be understood via the explanation that follows.

5.3.1. Understanding the solution behavior near the terminal time

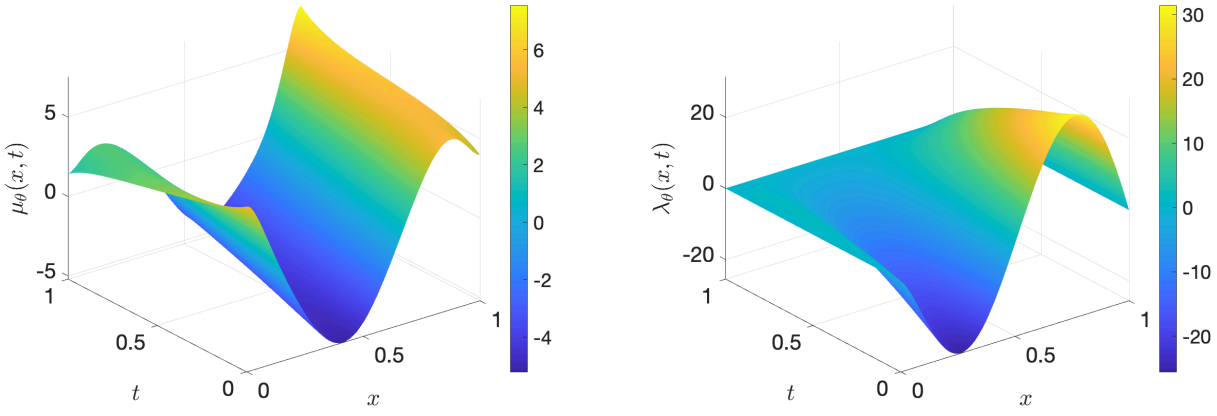
Consider the DtP mapping equations (44b), which are reproduced below:

$$\begin{aligned} u_H &= \frac{\partial \lambda}{\partial t} + \frac{\partial \mu}{\partial x} =: \partial_t \lambda + \partial_x \mu, \\ q_H &= \mu - \alpha\lambda - \kappa \frac{\partial \lambda}{\partial x} =: \mu - \alpha\lambda - \kappa\, \partial_x \lambda. \end{aligned} \tag{71}$$

Let T be the terminal time ($T = 1$ herein) and recall that $q(x, t) = \partial_x u(x, t)$. Now, for the sake of this argument, assume the functions $u(x, T)$, $q(x, T)$ for $x \in [0, 1]$ are known from the unique solution to the primal problem with specified initial and boundary conditions, which the discrete solution aspires to reproduce. Given the prescribed Dirichlet boundary condition on the top edge of the space-time domain, say $\lambda(x, T) = \lambda_{\text{top}}(x)$, it is clear from the expression for q_H in (71) that $\mu(x, T)$ for $x \in [0, 1]$ on the top edge become fixed, which implies that $\partial_x \mu(x, T)$ is also fixed. From the expression for u_H in (71), this also means that $\partial_t \lambda(x, T)$ is fixed on the top edge, and in particular $\partial_t \lambda(1, 1) =: d_{\text{top}}$. However, there is a Dirichlet boundary condition that is specified on the right edge of the domain, say $\lambda(1, t) = \lambda_{\text{right}}(t)$, thus fixing $\partial_t \lambda(1, \cdot)$, and in particular $\partial_t \lambda(1, T) =: d_{\text{right}}$. In our computations, we chose the functions $\lambda_{\text{top}} = 0, \lambda_{\text{right}} = 0$, but note that the value d_{top} is not entirely determined from the function λ_{top}. Now, if u is continuous at $(1, 1)$ but $d_{\text{top}} \neq d_{\text{right}}$, values that depend on the arbitrarily specified Dirichlet specifications of λ_{top} and λ_{right}, then approximating such a dual solution (this is not an impediment for a weak formulation) with smooth interpolation can result in higher errors at the $(1, 1)$ corner of the domain. On applying similar arguments to the behavior of the solutions on the left and top edges lead us to also draw the same inference at the $(0, 1)$ corner. Another way to interpret this result is that if (u_H, q_H) have to equal the exact solution, say (u, q), of the primal problem on the top edge, then it is possible that in general d_{top} may not equal d_{right}, and similarly for the left top corner. This implies that if such discontinuities are not allowed in the dual solution, then either $u_H \neq u$ or $q_H \neq q$ (or both) on the top edge. We note that μ has no boundary conditions specified for this problem (46c), and hence has substantially more freedom to accommodate the demands of continuity in the primal solution at the top corners of the domain. In general, even if an outright discontinuity like $d_{\text{top}} \neq d_{\text{right}}$ does not occur, the demands of the dual Dirichlet boundary conditions and the primal problem to be solved can set up boundary layers in the dual solution (see the dual solutions shown in Fig. 15). These features were demonstrated for the heat equation and pure convection case in [17, Secs. 5.1.1, 5.2.1] that led to some degradation in solution accuracy (as in our calculations here), but which can be alleviated with refinement. We point out that while the 'corner' issue does not arise in the solution of ODEs by the dual methodology, a boundary layer in the dual solution near the final time can arise. However, these final-time issues can be robustly resolved, as explained

in the following.

The degradation of accuracy of the solution on the top edge is easily dealt with, as demonstrated in [17, 18]. Let the primal initial-boundary value problem, posed in the time interval $[0, T]$, have a unique solution (and even otherwise). Then, one poses and solves the dual problem in an interval $[0, T + \delta]$ in time, with the boundary conditions of the primal (original) problem imposed in the interval $[0, T]$, and appending an (arbitrary) continuous extension of these primal boundary conditions in the interval $[T, T+\delta]$. Recovering the primal fields from the dual solution, it can be seen from the logic of the scheme that, in principle, the correct primal problem would have been solved in the time interval $[0, T]$. Thus, approaching this augmented primal problem with a 'buffer-zone' in the time-like direction with the dual technique and discarding the solution in the space-time strip $[0, L] \times (T, T + \delta]$ (δ preferably small), eliminates all issues with dual solution accuracy at top corners/edges in the truncated solution. Furthermore, since the dual solution scheme involves solving boundary-value problems in space-time domains, solution costs can be prohibitive for long-time simulations. For such situations, a time-slicing strategy can be utilized where the problem is solved in a finite number of stages in time, the truncated solution at the end of one stage serving as the primal initial condition for the successive stage; these procedures have been successfully demonstrated in [17, 18].



(a) (b)

Figure 15: B-spline solution (dual fields) of the transient convection-diffusion equation ($\kappa = 0.01, \alpha = 0.1$). (a) $\mu_\theta(x, t)$ ($p = 9$) and (b) $\lambda_\theta(x, t)$ ($q = 10$).

5.4. Transient heat equation with B-splines

Consider the following transient heat conduction model problem:

$$\kappa \frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t} \quad \text{in } \Omega = \Omega_0 \times \Omega_t = (0, 1) \times (0, 1), \tag{72a}$$

$$u(0, t) = 1, \quad \kappa \frac{\partial u}{\partial x}(1, t) = 0, \tag{72b}$$

$$u(x, 0) = u_0(x) = 1 + \sin\left(\frac{\pi x}{2}\right), \tag{72c}$$

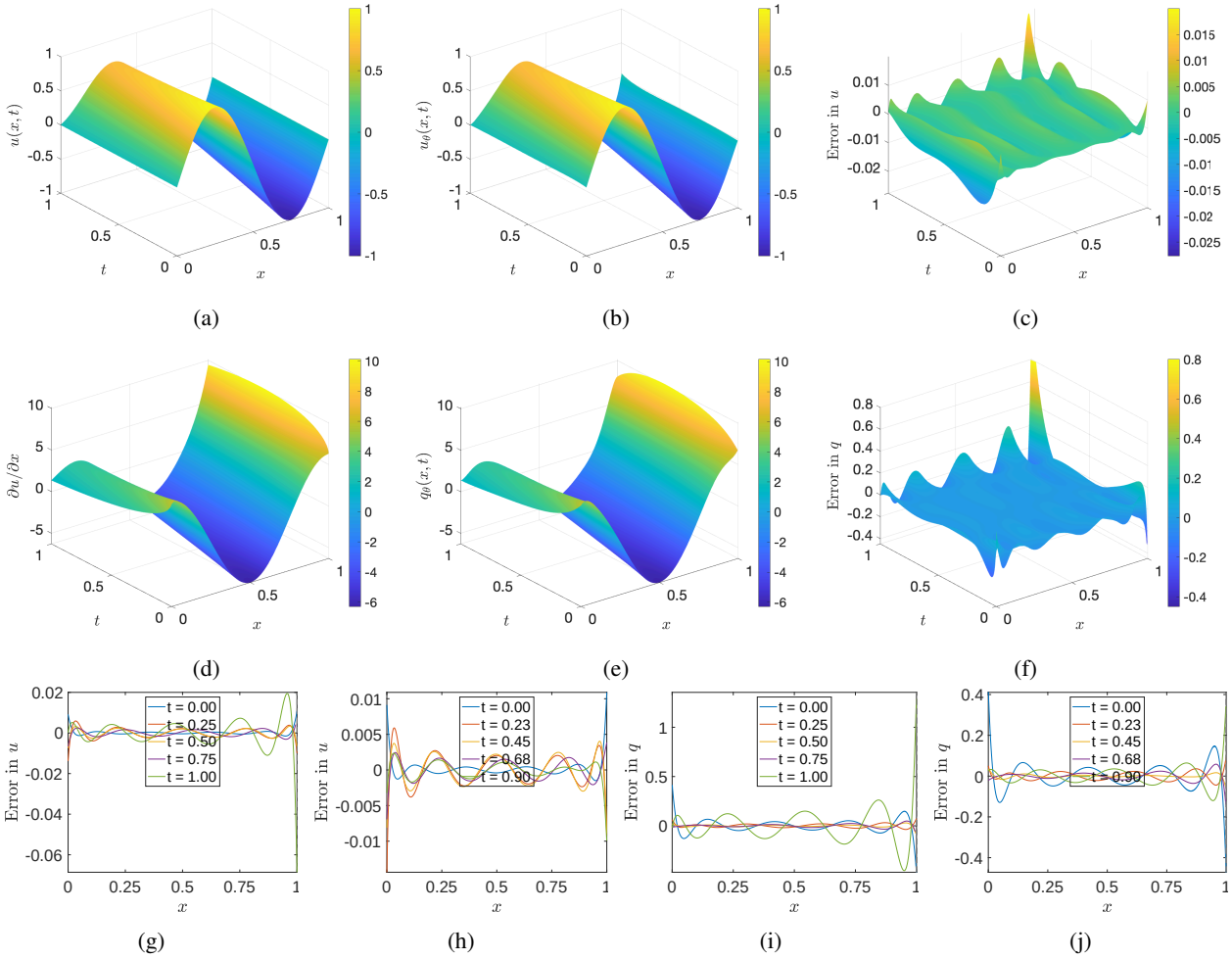


Figure 16: Space-time B-splines to solve the transient convection-diffusion problem. The conductivity coefficient $\kappa = 0.01$ and the convection coefficient $\alpha = 0.1$. The dual fields $\mu_\theta(x,t)$ and $\lambda_\theta(x,t)$ are composed of tensor-product B-splines of degree $p = 9$ and $q = 10$, respectively. (a) Exact solution, u; (b) u_θ; (c) Error, $u - u_\theta$; (d) Exact $q = \partial u/\partial x$; (e) q_θ; and (f) Error, $q_\theta - \partial u/\partial x$. Plots of the time history for the error in u at (g) $t \in [0,1]$ and (h) $t \in [0,0.9]$. Plots of the time history for the error in q for (i) $t \in [0,1]$ and (j) $t \in [0,0.9]$.

with the exact solution:

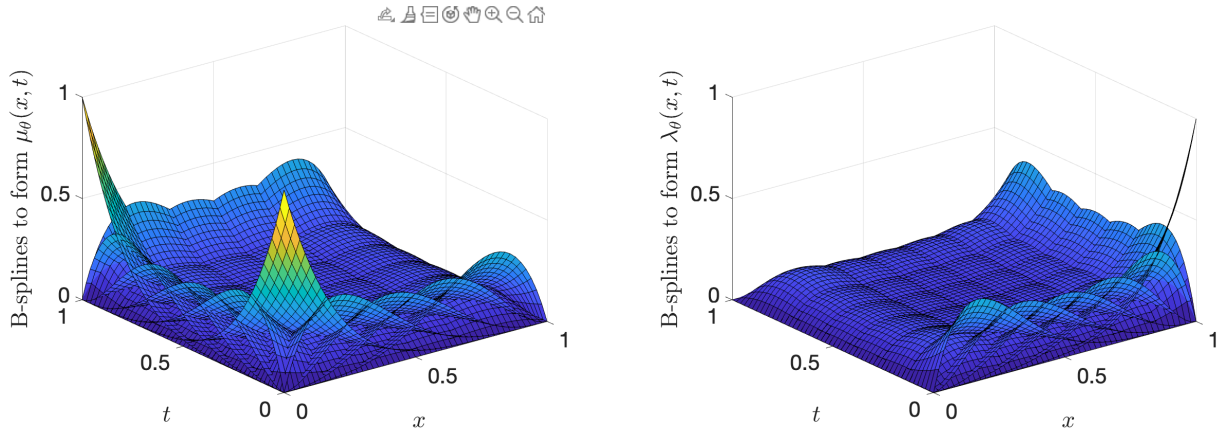
$$u(x,t) = 1 + \sin\left(\frac{\pi x}{2}\right)\exp\left(-\frac{\pi^2\kappa}{4}t\right). \tag{72d}$$

The dual functional for this IBVP with $\bar{u}_1 = 1$ is given in (52). We set its first variation to zero and follow the steps presented for the transient convection-diffusion problem. The stiffness matrix is obtained by setting $\alpha = 0$ in (60b) and the force vector is:

$$\boldsymbol{f}_\lambda = -\int_0^1 u_0(x)\boldsymbol{N}_\lambda^\top(x,0)\,dx, \quad \boldsymbol{f}_\mu = -\int_0^1 \boldsymbol{N}_\mu^\top(0,t)\,dt. \tag{73}$$

The trial functions are chosen to be of the form (70) with $n = 1$ in the bivariate open knot vector (69). Plots of the two-dimensional basis functions for $\mu_\theta(x,t)$ ($p = 5$) and for $\lambda_\theta(x,t)$ ($q = 6$) are shown in Fig. 17. Observe that

$\mu_\theta(1,t) = 0$ and $\lambda_\theta(0,t) = \lambda_\theta(x,1) = 0$ so that $\lambda_\theta \in S_\lambda$ and $\mu_\theta \in S_\mu$ meet (52c), and are therefore kinematically admissible. We conduct numerical computations for $\kappa = 1$. The B-spline solution for the dual fields, $\mu_\theta(x,t)$ and



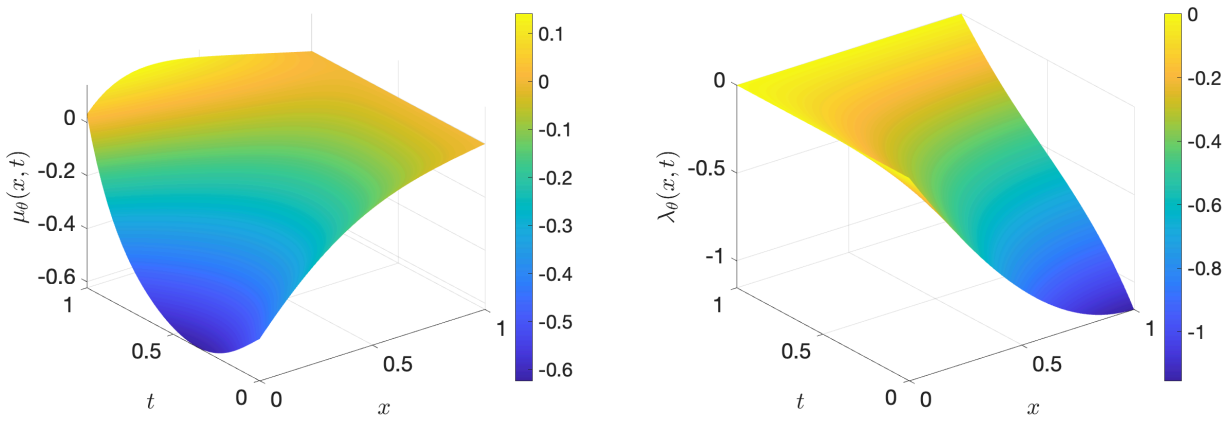
(a) (b)

Figure 17: Plots of two-dimensional B-spline basis functions to form the dual fields (a) $\mu_\theta(x,t)$ ($p = 5$) and (b) $\lambda_\theta(x,t)$ ($q = 6$) to solve the heat conduction problem. The bivariate open knot vector ($n = 1$) that is given in (69) is used.

$\lambda_\theta(x,t)$, are presented in Fig. 18. The primal fields are computed from the dual field using the DtP map in (52a). Comparisons of the exact solution to the B-spline solution for the temperature and the flux are provided in Fig. 19. Even on this coarse discretization, the space-time B-spline solutions are fairly accurate: maximum pointwise errors in u and $q = \partial u/\partial x$ (flux) are 4×10^{-3} and 9×10^{-2}, respectively. One can observe from Figs. 19i and 19j that the accuracy for notably $q = u'$ worsens as one approaches the terminal time $t = 1$. This behavior in the vicinity of $t = 1$ is also observed (albeit more severe) in Fig. 16 for the convection-diffusion problem; refer to the discussion in Section 5.3.1 on the behavior of the discrete solution near $t = 1$.

6. Conclusions

In this paper, we applied a recently proposed dual variational principle to solve partial differential equations that do not have a variational structure in primal form. In this variational approach, the primal partial differential equation is treated as a constraint and an arbitrarily chosen convex auxiliary potential is minimized. This leads to requiring a concave dual functional to be maximized subject to Dirichlet boundary conditions on dual variables. Prescribed Dirichlet boundary conditions in the primal problem appear as natural boundary conditions in the dual functional. The overarching goal of the duality approach is to solve ordinary/partial differential equations using a variational strategy. Beyond its use in the present work and previous studies [9–14], one can also solve for critical points of primal functionals (for example, the Chern–Simons functional [46]) that are not bounded below or above. We presented the general formalism and illustrated it for a system of linear equations, a quadratic system of two equations, and to obtain nonnegative solutions for an underdetermined system of linear equations. It was also shown how a causal initial-value problem



(a) (b)

Figure 18: B-spline solution (dual fields) of the transient heat equation ($\kappa = 1$). (a) $\mu_\theta(x,t)$ ($p = 5$) and (b) $\lambda_\theta(x,t)$ ($q = 6$).

(ODE) can be exactly solved (and accurately approximated) as an acausal boundary-value problem in time. The dual weak form was derived for the transient convection-diffusion equation, which was discretized using a Galerkin method with finite-dimensional trial and test functions that were formed using machine learning approximants and B-splines. Uniqueness of solutions of the dual variational equations for the transient convection-diffusion equation was established. Shallow neural network with smooth RePU activation function and univariate B-splines delivered accurate solutions for the steady-state convection-diffusion equation. Smooth B-spline approximations of degrees p and $p+1$ for the dual fields delivered more accurate solutions for the flux ($q = u'$) than u; the rates of convergence in the L^2 norm and H^1 seminorm of the primal field u were of order p and $p+1$, respectively. One-dimensional transient (heat and convection-diffusion) problems were successfully solved as a space-time Galerkin method with tensor-product B-spline basis functions. It was shown that the errors were small but they tended to become relatively larger near the terminal time $t = T$, and an explanation of this behavior and solution strategies for its remedy were put forth. As part of future work, we plan to pursue the application of the dual variational scheme to solve challenging linear and nonlinear systems of ordinary and partial differential equations.

Appendix A. Adjoint method for constrained-ODE

Consider the ODE in (23a), and denote the initial condition in (23b) as $u(0) = u_0 = p$ (p is now considered as a parameter), so that the exact solution is:

$$u(t,p) = pe^{at}. \tag{A.1}$$

Consider the objective function [47]

$$F(u,p) = \int_0^T u(t,p)\,dt. \tag{A.2}$$

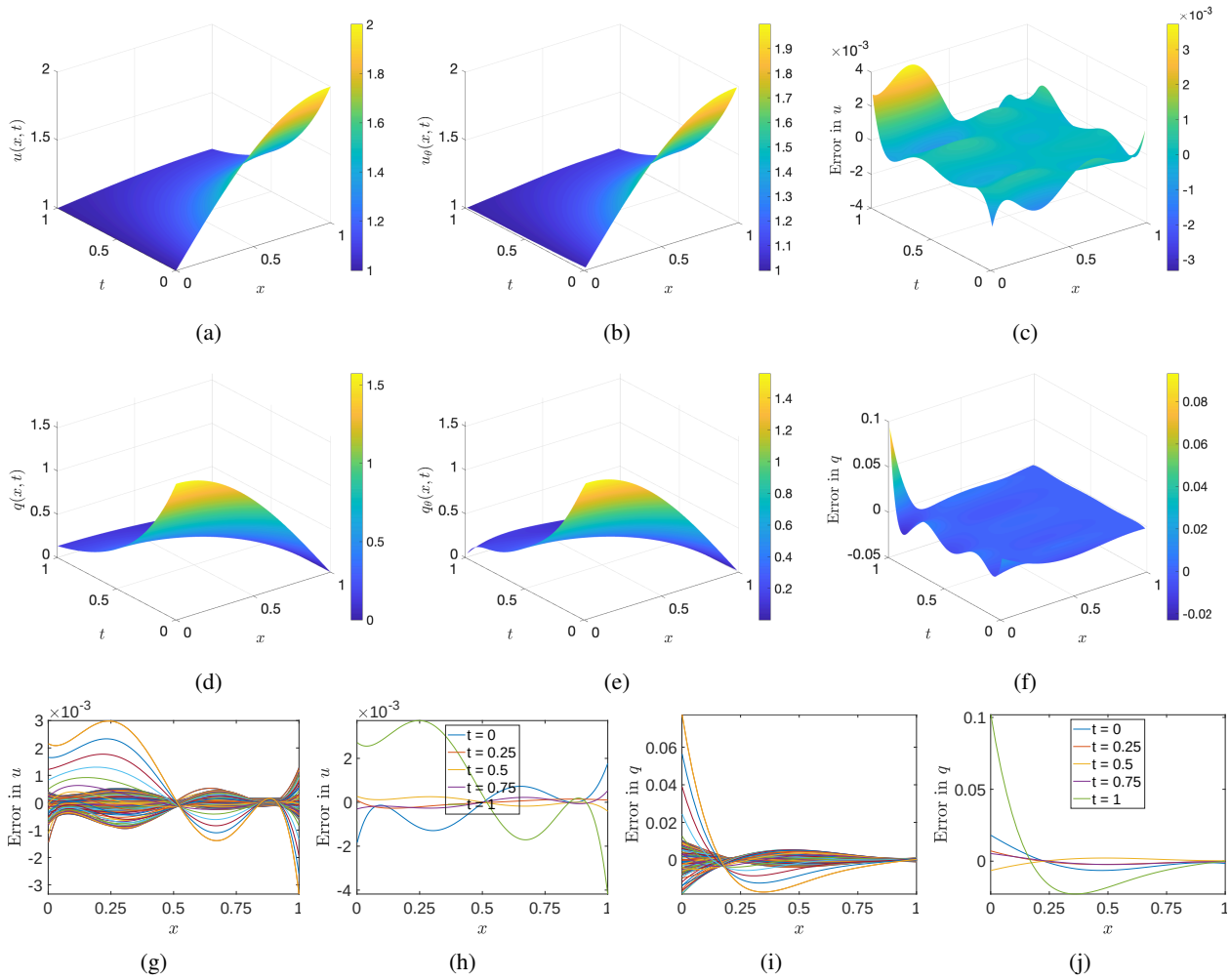


Figure 19: Space-time B-splines to solve the transient heat conduction problem ($\kappa = 1$). The dual fields $\mu_\theta(x,t)$ and $\lambda_\theta(x,t)$ are composed of tensor-product B-splines of degree $p = 5$ and $q = 6$, respectively. (a) Exact solution, u; (b) u_θ; (c) Error, $u - u_\theta$; (d) Exact $q = \partial u/\partial x$; (e) q_θ; and (f) Error, $q_\theta - \partial u/\partial x$. Plots of the time history for the error in u at (g) $t = [0, 0.01, 0.02, \ldots, 1]$ and (h) $t = [0, 0.25, 0.5, 0.75, 1]$. Plots of the time history for the error in q at (i) $t = [0, 0.01, 0.02, \ldots, 1]$ and (j) $t = [0, 0.25, 0.5, 0.75, 1]$.

The sensitivity of F with respect to p, dF/dp, is sought. On using the method of Lagrange multipliers, we form the Lagrangian:

$$L(u, \lambda) = F(u, p) - \int_0^T \lambda\,(\dot{u} - au)\,dt, \tag{A.3}$$

where λ is known as the adjoint variable. Since $au - \dot{u} = 0$, then

$$\begin{aligned}
\frac{dF}{dp} = \frac{dL}{dp} &= \int_0^T \frac{\partial u}{\partial p}\,dt + \int_0^T \lambda\left(a\frac{\partial u}{\partial p} - \frac{\partial \dot{u}}{\partial p}\right)dt \\
&= \int_0^T \frac{\partial u}{\partial p}\,dt - \left[\lambda\frac{\partial u}{\partial p}\right]_{t=0}^{t=T} + \int_0^T (\dot{\lambda} + a\lambda)\frac{\partial u}{\partial p}\,dt, \\
&= -\left[\lambda\frac{\partial u}{\partial p}\right]_{t=0}^{t=T} + \int_0^T (\dot{\lambda} + a\lambda + 1)\frac{\partial u}{\partial p}\,dt.
\end{aligned} \tag{A.4}$$

Now we choose $\lambda(T) = 0$ so that the boundary term at $t = T$ vanishes, and we eliminate the integral by requiring λ to satisfy the ODE (adjoint equation) $\dot{\lambda} + a\lambda + 1 = 0$. In addition, since $du(0)/dp = 1$ for the boundary term at $t = 0$, (A.4) simplifies to

$$\frac{dF}{dp} = \lambda(0). \tag{A.5}$$

Now, λ must satisfy the adjoint system:

$$\dot{\lambda} + a\lambda + 1 = 0, \quad \lambda(T) = 0, \tag{A.6}$$

which admits the exact solution

$$\lambda(t) = \frac{e^{a(T-t)} - 1}{a}. \tag{A.7}$$

On using this in (A.5), the solution from the adjoint method is:

$$\frac{dF}{dp} = \frac{e^{aT} - 1}{a}, \tag{A.8}$$

which matches the result for dF/dp using the exact solution in (A.1).

In the optimization of parametric differential-algebraic equation (DAE) systems, a given objective function is required to be optimized with respect to parameters $\boldsymbol{p}$, subject to satisfying an initial-boundary value problem (forward or primal problem) that depends on the parameters. It is assumed that solvers are available for the forward problem. In the dual approach for differential-algebraic systems, a user-specified convex potential is optimized subject to treating the system as a constraint. The potential is easily designed using the knowledge of the DAE system to be solved. Both methods use Lagrange multipliers (referred to as *adjoint variable* and *dual variable* in the two approaches) to form the Lagrangian of the respective problems. The goal of the adjoint method is to provide an efficient method to compute the sensitivity of the objective function with respect to $\boldsymbol{p}$. The goal of the duality approach is to solve ordinary/partial differential-algebraic systems using a variational strategy. This eases the burden of solving problems with nonstandard structure [10] or those without existence-of-solution guarantees when viewed through the lens of current methods for the primal problem [12]. The adjoint method in constrained optimization does not concern itself with solution strategies for the forward problem, the availability of a robust technique for which is an essential ingredient of executing the method. The dual variational approach to DAEs concerns itself with a strategy to solve such (primal) systems, but does not address the question of optimization of an objective function with respect to parameters defining the primal problem.

There is a common step (albeit superficial) in the two approaches. For parametric constrained function optimiza-

tion, the scheme in the adjoint method is as follows [47]:

$$\begin{aligned} \text{Given } f(\boldsymbol{x}, p), \text{ subject to } \boldsymbol{g}(\boldsymbol{x}, p) = 0, \quad & \text{find } \frac{df}{dp}. \\ \text{Crucial step is the definition of the adjoint } \boldsymbol{\lambda} \text{ from} \quad & \frac{\partial f}{\partial \boldsymbol{x}} + \boldsymbol{\lambda}^\top \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}} = \boldsymbol{0}. \\ \text{Evaluate} \quad & \frac{df}{dp} = \boldsymbol{\lambda}^\top \frac{\partial \boldsymbol{g}}{\partial p} + \frac{\partial f}{\partial p}. \end{aligned} \tag{A.9}$$

With the replacement $f \to H$, $\boldsymbol{g} \to \boldsymbol{G}$, and $\boldsymbol{x} \to \boldsymbol{U}$ (see Section 2), the equation for the adjoint in (A.9) is identical to the DtP mapping, but with a very different goal and interpretation. The adjoint method defines the adjoint (dual variable $\boldsymbol{\lambda}$) using the knowledge of the primal $\boldsymbol{x}$; the dual approach does exactly the opposite. In constrained optimization using the adjoint method, the solution of the primal problem is required to solve for the adjoint variable. The dual approach defines the dual problem, which provides the dual solution, and the DtP (dual-to-primal) mapping yields the primal solution that solves the primal problem.

Appendix B. Weak form of the dual BVP for the IVP

We consider the IVP posed in (23), with the dual functional $S[\lambda]$ given in (28) and the strong form of the dual BVP presented in (31). On setting $\delta S[\lambda; \delta\lambda] = 0$ and using (27), we obtain

$$\begin{aligned} \delta S[\lambda; \delta\lambda] &= -\int_0^T (\dot{\lambda} + a\lambda)(\dot{\delta\lambda} + a\delta\lambda)\, dt - u_0 \delta\lambda(0) \\ &= -\int_0^T [\dot{\lambda}\dot{\delta\lambda} + a^2\lambda\delta\lambda]\, dt - \int_0^T a[\dot{\lambda}\delta\lambda + \lambda\dot{\delta\lambda}]\, dt - u_0\delta\lambda(0) \\ &= 0 \quad \forall \delta\lambda \in \mathsf{S}_\lambda, \end{aligned} \tag{B.1}$$

where $\mathsf{S}_\lambda = \{\delta\lambda \in H^1(0,T),\ \delta\lambda(T) = 0\}$ is the space of admissible variations. We find that

$$\begin{aligned} \int_0^T a(\dot{\lambda}\delta\lambda + \lambda\dot{\delta\lambda})\, dt = \int_0^T a\delta\left[\frac{d}{dt}\left(\frac{1}{2}\lambda^2\right)\right] dt = \delta\left[\int_0^T a\frac{d}{dt}\left(\frac{1}{2}\lambda^2\right) dt\right] = \delta\left[a\frac{\lambda^2}{2}\right]_{t=0}^{t=T} \\ = a\lambda(T)\delta\lambda(T) - a\lambda(0)\delta\lambda(0) = -a\lambda(0)\delta\lambda(0), \end{aligned}$$

since $\delta\lambda(T) = 0$. On using the above in (B.1), we can write the variational form as:

$$\int_0^T [\dot{\lambda}\dot{\delta\lambda} + a^2\lambda\delta\lambda]\, dt - a\lambda(0)\delta\lambda(0) = -u_0\delta\lambda(0) \quad \forall \delta\lambda \in \mathsf{S}_\lambda. \tag{B.2}$$

To derive the weak form, we multiply the dual BVP in (31) by the test function (virtual field) $\delta\lambda$ and perform integration by parts to obtain

$$-\int_0^T [\dot{\lambda}\dot{\delta\lambda} + a^2\lambda\delta\lambda]\, dt + \dot{\lambda}(T)\delta\lambda(T) - \dot{\lambda}(0)\delta\lambda(0) = 0 \quad \forall \delta\lambda \in \mathsf{S}_\lambda.$$

On substituting $\delta\lambda(T) = 0$ and the Robin boundary condition from (31b) to replace $\dot{\lambda}(0)$ in the above equation leads us to the weak form, which is identical to the variational form given in (B.2).

References

[1] N. Ghoussoub, A. Moameni, Anti-symmetric Hamiltonians (II): Variational resolutions for Navier–Stokes and other nonlinear evolutions, Annales de l'Institut Henri Poincaré C 26 (1) (2009) 223–255.
[2] M. Ortiz, B. Schmidt, U. Stefanelli, A variational approach to Navier–Stokes, Nonlinearity 31 (12) (2018) 5664.
[3] M. Ortiz, E. A. Repetto, Nonconvex energy minimization and dislocation structures in ductile single crystals, Journal of the Mechanics and Physics of Solids 47 (2) (1999) 397–462.
[4] M. Ortiz, L. Stainier, The variational formulation of viscoplastic constitutive updates, Computer Methods in Applied Mechanics and Engineering 171 (3-4) (1999) 419–444.
[5] H. Petryk, A quasi-extremal energy principle for non-potential problems in rate-independent plasticity, Journal of the Mechanics and Physics of Solids 136 (2020) 103691.
[6] C. Carstensen, K. Hackl, A. Mielke, Non–convex potentials and microstructures in finite–strain plasticity, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 458 (2018) (2002) 299–317.
[7] M. Gurtin, Variational principles for linear initial-value problems, Quarterly of Applied Mathematics 22 (3) (1964) 252–256.
[8] R. L. Seliger, G. B. Whitham, Variational principles in continuum mechanics, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences 305 (1480) (1968) 1–25.
[9] A. Acharya, Variational principle for nonlinear PDE systems via duality, Quarterly of Applied Mathematics 81 (2023) 127–140.
[10] A. Acharya, A dual variational principle for nonlinear dislocation dynamics, Journal of Elasticity 154 (1) (2023) 383–395.
[11] A. Acharya, A hidden convexity in continuum mechanics, with application to classical, continuous-time, rate-(in) dependent plasticity, Mathematics and Mechanics of Solids (2024). doi:https://doi.org/10.1177/10812865241258154.
[12] S. Singh, J. Ginster, A. Acharya, A hidden convexity of nonlinear elasticity, Journal of Elasticity 156 (2024) 975–1014.
[13] A. Acharya, B. Stroffolini, A. Zarnescu, Variational dual solutions for incompressible fluids (2024). arXiv:2409.04911.
[14] A. Acharya, A. N. Sengupta, Action principles for dissipative, non-holonomic Newtonian mechanics, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 480 (2024) 20240113.
[15] R. M. Errico, What is an adjoint model?, Bulletin of the American Meteorological Society 78 (11) (1997) 2577–2591.
[16] Y. Cao, S. Li, L. Petzold, R. Serban, Adjoint sensitivity analysis for differential-algebraic equations: The adjoint DAE system and its numerical solution, SIAM Journal on Scientific Computing 24 (3) (2003) 1076–1089.
[17] U. Kouskiya, A. Acharya, Hidden convexity in the heat, linear transport, and Euler's rigid body equations: A computational approach, Quarterly of Applied Mathematics 82 (2024) 673–703.
[18] U. Kouskiya, A. Acharya, Inviscid Burgers as a degenerate elliptic problem, published online, Quarterly of Applied Mathematics (https://arxiv.org/abs/2401.08814, 2024).
[19] Y. Brenier, Hidden convexity in some nonlinear PDEs from geomety and physics, Journal of Convex Analysis 17 (3&4) (2010) 945–959.
[20] Y. Brenier, The initial value problem for the Euler equations of incompressible fluids viewed as a concave maximization problem, Communications in Mathematical Physics 364 (2018) 579–605.
[21] A. N. Brooks, T. J. R. Hughes, Streamline upwind/Petrov–Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier–Stokes equations, Computer Methods in Applied Mechanics and Engineering 32 (1-3) (1982) 199–259.
[22] T. J. R. Hughes, L. P. Franca, P. Leopoldo, G. M. Hulbert, A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/least-squares method for advective-diffusive equations, Computer Methods in Applied Mechanics and Engineering 73 (2) (1989) 173–189.
[23] T. E. Tezduyar, Stabilized finite element formulations for incompressible flow computations, Advances in Applied Mechanics 28 (1991) 1–44.
[24] T. J. R. Hughes, Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods, Computer Methods in Applied Mechanics and Engineering 127 (1-4) (1995) 387–401.
[25] P. B. Bochev, M. D. Gunzburger, Least-Squares Finite Element Methods, Vol. 166, Springer Science & Business Media, 2009.
[26] M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for forward and inverse problems involving nonlinear partial differential equations, Journal of Computational Physics 378 (2019) 686–707.
[27] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics-informed scientific machine learning, Nature Review Physics 3 (6) (2021) 422–440.
[28] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, F. Piccialli, Scientific machine learning through physics-informed neural networks: Where we are and what's next, Journal of Scientific Computing 92 (3) (2022) 88.
[29] Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljačić, T. Y. Hou, M. Tegmark, KAN: Kolmogorov–Arnold networks (2024). arXiv:2404.19756.
[30] B. C. Koenig, S. Kim, S. Deng, KAN-ODEs: Kolmogorov–Arnold network ordinary differential equations for learning dynamical systems and hidden physics (2024). arXiv:2407.04192.
[31] L. F. Guilhoto, P. Perdikaris, Deep learning alternatives of the Kolmogorov superposition theorem (2024). arXiv:2410.01990.
[32] W. E, B. Yu, The deep Ritz method: A deep learning-based numerical algorithm for solving variational problems, Communications in Mathematics and Statistics 6 (1) (2018) 1–12.
[33] N. Sukumar, A. Srivastava, Exact imposition of boundary conditions with distance functions in physics-informed deep neural networks, Computer Methods in Applied Mechanics and Engineering 389 (2022) 114333.
[34] J. He, J. Xu, Expressivity and approximation properties of deep neural networks with ReLU^k activation (2024). arXiv:2312.16483.
[35] J. He, J. Xu, Deep neural networks and finite elements of any order on arbitrary dimensions (2024). arXiv:2312.14276.
[36] C. de Boor, A Practical Guide to Splines, Applied Mathematical Sciences, Springer, New York, 2001.
[37] M. S. Floater, Generalized barycentric coordinates and applications, Acta Numerica 24 (2015) 161–214.
[38] C. E. Shannon, A mathematical theory of communication, The Bell System Technical Journal 27 (1948) 379–423.
[39] E. T. Jaynes, Information theory and statistical mechanics, Physical Review 106 (4) (1957) 620–630.
[40] N. Sukumar, Construction of polygonal interpolants: a maximum entropy approach, International Journal for Numerical Methods in Engineering 61 (12) (2004) 2159–2181.
[41] T. Belytschko, Y. Krongauz, D. Organ, M. Fleming, P. Krysl, Meshless methods: An overview and recent developments, Computer Methods in Applied Mechanics and Engineering 139 (1996) 3–47.

[42] T. M. Black, Mesh-free Applications to Fracture Mechanics and An Analysis of the Corrected Derivative Method, Ph.D. thesis, Theoretical and Applied Mechanics, Northwestern University, Evanston, IL, U.S.A. (December 1999).
[43] F. Punzo, A. Tesei, Uniqueness of solutions to degenerate elliptic problems with unbounded coefficients, Annales de l'Institut Henri Poincaré C 26 (5) (2009) 2001–2024.
[44] T. J. R. Hughes, The Finite Element Method: Linear Static and Dynamic Finite Element Analysis, Prentice-Hall, Englewood Cliffs, N.J., 1987.
[45] M. B. Glinowiecka-Cox, Analytic Solution of 1D Diffusion-Convection Equation with Varying Boundary Conditions, B.S. Thesis, Department of Mathematics, Portland State University, Portland, OR 97207, USA (2022).
[46] A. Acharya, J. Ginster, A. N. Sengupta, Variational dual solutions of Chern-Simons theory (2024). arXiv:2411.17635.
[47] A. M. Bradley, A tutorial on the adjoint method (2024).
URL https://cs.stanford.edu/~ambrad/adjoint_tutorial.pdf