AIAA AVIATION Forum
5-9 June 2017, Denver, Colorado
23rd AIAA Computational Fluid Dynamics Conference

Energetically Optimal Flapping Flight via a Fully Discrete Adjoint Method with Explicit Treatment of Flapping Frequency

Jingyi Wang*, Matthew J. Zahr† and Per-Olof Persson‡
Lawrence Berkeley National Laboratory,
University of California, Berkeley,
Berkeley, CA 94720-3840, U.S.A.

This work introduces a fully discrete, high-order numerical framework for solving PDE-constrained optimization problems using gradient-based methods in the case where one or more of the optimization parameters affects the time domain; a canonical example being optimization of the frequency of a flapping wing. In a fully discrete setting, this effective parametrization of the time domain leads to a parametrization of the time discretization, e.g., to maintain a fixed number of timesteps per period, the timestep size is parameter-dependent. Gradients of quantities of interest in this work are computed using the adjoint method, which must take into account the parametric dependence of the time discretization. As this work considers energetically optimal flight, a globally high-order discretization of conservation laws on deforming domains is employed: an Arbitrary Lagrangian-Eulerian formulation maps the conservation law to a fixed reference domain and a high-order discontinuous Galerkin method and diagonally implicit Runge-Kutta method are used for the spatial and temporal discretizations. This framework is applied to study energetically optimal flapping subject to a minimum required thrust, including frequency and pitching/heaving trajectories as optimization parameters. This marks a distinct departure from other adjoint-based approaches to optimal flapping that fix the frequency.

I. Introduction

Flapping flight has been a subject of intense interest and research over the past several decades due to its relevance in designing Micro Aerial Vehicles (MAVs) – unmanned aerial vehicles measuring no more than 15 cm in any dimension, envisioned in a number of civilian and military applications, including surveillance and reconnaissance [1, 2] – and in the understanding of biological systems. The basic goal of any system, whether biological or manmade, that relies on flapping propulsion is to adjust the kinematics of the flapping wing, and possibly its shape, to minimize the energy required to complete a given mission, i.e., satisfy lift and thrust constraints. The problem of determining the flapping kinematics that leads to an energetically optimal motion, while satisfying various mission constraints, leads to a nonlinearly constrained PDE-constrained optimization problem.

It is well-documented [3, 4] that the Strouhal number is an important factor in efficiently generating thrust and should therefore be included in the parametrization of the flapping motion when considering energetically optimal flapping flight. Most work on this subject [5–9] only optimizes for the trajectory of the heaving and pitching motion body *at a fixed frequency*, which implies the Strouhal number is only parametrized through the parametrization of the amplitude of the heaving motion. This approach potentially excludes more efficient flapping motions that could be obtained by flapping faster or slower and may generate

*Graduate Student, Department of Mechanical Engineering, University of California, Berkeley, Berkeley CA 94720-3840. E-mail: jingtian@berkeley.edu.
†Postdoctoral Fellow, Department of Mathematics, Lawrence Berkeley National Laboratory, University of California, Berkeley, Berkeley, CA 94720. E-mail: mjzahr@lbl.gov.
‡Associate Professor, Department of Mathematics, University of California, Berkeley, Berkeley CA 94720-3840. E-mail: persson@berkeley.edu.

infeasible optimization problems if there are severe physical or mechanical limitations on the amplitude of pitching or heaving.

Therefore, the goal of this work is to develop a high-fidelity numerical framework for solving PDE-constrained optimization problems that include parameters that influence the time domain, such as flapping frequency, using gradient-based optimization methods. To ensure the gradients of the optimization functionals are discretely consistent [9–11] and a large number of parameters can be handled efficiently, a fully discrete adjoint method is introduced for the gradient computations. A fundamental issue that arises when incorporating any parameter that directly influences the time domain itself in a discrete setting is the time discretization becomes parameter-dependent. For example, if the timestep size and number of timesteps are fixed and the flapping frequency changes, a non-integral number of periods would be considered. This will lead to a completely irrelevant problem if less than a single period is considered or require special care when computing time-averaged quantities of interest. Another issue with this approach of fixing the size and number of timesteps is there will be fewer timesteps per period as frequency increases, causing the accuracy of the numerical method to degrade. The approach taken here is to fix the number of timesteps per period, causing the size of each timestep to become parameter-dependent, which ensures sufficient accuracy in the time discretization across any range of frequencies. This work describes a primal solver and the corresponding adjoint method for computing gradients of quantities of interest that takes this dependency into account. Other work that has considered optimization with frequency parameters in the context of optimal flapping flight are found in [12, 13]; however they did not parametrize the time discretization by e.g., fixing the number of timesteps per period.

Following the authors previous work [9, 14], the compressible Navier-Stokes equations will be recast on a fixed reference domain using an Arbitrary Lagrangian Eulerian formulation and discretized using a high-order discontinuous Galerkin method. The temporal discretization employs a high-order diagonally implicit Runge-Kutta scheme. The quantities of interest—space-time integrals that depend on the flow solution—that comprise the optimization objective and constraints are discretized using the same high-order spatial (DG) and temporal (RK) schemes used for the conservation law. This ensures the discretization is globally high-order. The fully discrete adjoint equations derived from this discretization preserve the diagonally implicit nature of the temporal integrator and are therefore amenable to an efficient implementation. This primal and adjoint CFD framework is used to determine the energetically optimal flapping motion of a two-dimensional airfoil in compressible, viscous flow, subject to mission requirements, i.e., constraint on the time-averaged thrust produced by the flapping motion. In this setup, the flapping frequency as well as the trajectory of the pitching and heaving motions are included as parameters. The energetically optimal flapping motions are studied as a function of the required thrust and it is shown that the optimal flapping energy varies linearly with the required thrust while the optimal frequency and pitch/heave amplitudes exhibit nonlinear trends.

The remainder of this document is organized as follows. Section II introduces the governing conservation law considered in this work, the isentropic Navier-Stokes equations, and an Arbitrary Lagrangian-Eulerian method that transforms it from a deforming, parametrized domain to a fixed one. Section III introduces the high-order discretization of the conservation law and its quantities of interest, with special attention paid to the parametrization of the time discretization that results from, e.g., parametrizing frequency in a fully discrete setting with a fixed number of timesteps per period. Section III.D introduces the fully discrete adjoint equations and corresponding adjoint method for computing gradients of quantities of interest that take into account the parametrization of the time discretization. Finally, Section IV applies this high-order simulation and optimization framework to study energetically optimal flapping motions as a function of the minimum required thrust and Section V offers conclusions.

II. Governing equations

This section is devoted to the treatment of conservation laws on a *parametrized, deforming domain* using an Arbitrary Lagrangian-Eulerian (ALE) formulation. Given that this work is concerned with energetically optimal flapping flight, the compressible Navier-Stokes equations are taken as the governing equations.

II.A. Compressible Navier-Stokes equations

The compressible Navier-Stokes equations are written as:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}(\rho u_i) = 0, \tag{1}$$

$$\frac{\partial}{\partial t}(\rho u_i) + \frac{\partial}{\partial x_i}(\rho u_i u_j + p) = +\frac{\partial \tau_{ij}}{\partial x_j} \quad \text{for } i = 1, 2, 3, \tag{2}$$

$$\frac{\partial}{\partial t}(\rho E) + \frac{\partial}{\partial x_i}\left(u_j(\rho E + p)\right) = -\frac{\partial q_j}{\partial x_j} + \frac{\partial}{\partial x_j}(u_j \tau_{ij}), \tag{3}$$

where ρ is the fluid density, u_1, u_2, u_3 are the velocity components, and E is the total energy. The viscous stress tensor and heat flux are given by

$$\tau_{ij} = \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}\frac{\partial u_k}{\partial x_k}\delta_{ij}\right) \qquad \text{and} \qquad q_j = -\frac{\mu}{\text{Pr}}\frac{\partial}{\partial x_j}\left(E + \frac{p}{\rho} - \frac{1}{2}u_k u_k\right). \tag{4}$$

Here, μ is the viscosity coefficient and $\text{Pr} = 0.72$ is the Prandtl number which we assume to be constant. For an ideal gas, the pressure p has the form

$$p = (\gamma - 1)\rho\left(E - \frac{1}{2}u_k u_k\right), \tag{5}$$

where γ is the adiabatic gas constant. In this work, the entropy is assumed constant, that is to say the flow is adiabatic and reversible. This makes the energy equation redundant and effectively reduces the square system of PDEs of size $n_{sd}+2$ to one of size $n_{sd}+1$, where n_{sd} is the number of spatial dimensions. It can be shown, under suitable assumptions, that the solution of the isentropic approximation of the Navier-Stokes equations converges to the solution of the incompressible Navier-Stokes equations as the Mach number goes to 0 [15–17].

II.B. Arbitrary Lagrangian-Eulerian formulation of conservation laws

Consider a general system of conservation laws defined on a parametrized, deforming domain, $v(\boldsymbol{\mu}, t)$,

$$\frac{\partial \boldsymbol{U}}{\partial t} + \nabla \cdot \boldsymbol{F}(\boldsymbol{U},\ \nabla \boldsymbol{U}) = 0 \quad \text{in} \quad v(\boldsymbol{\mu}, t) \tag{6}$$

where the physical flux is decomposed into an inviscid and a viscous part $\boldsymbol{F}(\boldsymbol{U},\ \nabla \boldsymbol{U}) = \boldsymbol{F}^{inv}(\boldsymbol{U}) + \boldsymbol{F}^{vis}(\boldsymbol{U},\ \nabla \boldsymbol{U})$, $\boldsymbol{U}(\boldsymbol{x}, \boldsymbol{\mu}, t)$ is the solution of the system of conservation laws, $t \in (0, T)$ represents time, and $\boldsymbol{\mu} \in \mathbb{R}^{N_{\boldsymbol{\mu}}}$ is a vector of parameters. This work will focus on the case where the *domain* is parametrized by $\boldsymbol{\mu}$.

The conservation law on the physical, deforming domain $v(\boldsymbol{\mu}, t) \subset \mathbb{R}^{n_{sd}}$ is transformed into one on a *fixed* reference domain $V \subset \mathbb{R}^{n_{sd}}$ through the introduction of a time-dependent diffeomorphism between the physical and reference domains: $\boldsymbol{x}(\boldsymbol{X}, \boldsymbol{\mu}, t) = \mathcal{G}(\boldsymbol{X}, \boldsymbol{\mu}, t)$. In this setting, n_{sd} is the number of spatial dimensions, $\boldsymbol{X} \in V$ is a point in the reference domain and $\boldsymbol{x}(\boldsymbol{X}, \boldsymbol{\mu}, t) \in v(\boldsymbol{\mu}, t)$ is the corresponding point in the physical domain at time t and parameter configuration $\boldsymbol{\mu}$. The transformed system of conservation laws takes the form

$$\left.\frac{\partial \boldsymbol{U}_{\boldsymbol{X}}}{\partial t}\right|_{\boldsymbol{X}} + \nabla_{\boldsymbol{X}} \cdot \boldsymbol{F}_{\boldsymbol{X}}(\boldsymbol{U}_{\boldsymbol{X}},\ \nabla_{\boldsymbol{X}} \boldsymbol{U}_{\boldsymbol{X}}) = 0 \tag{7}$$

where $\nabla_{\boldsymbol{X}}$ denotes spatial derivatives with respect to the reference variables, $\boldsymbol{X}$. The transformed state vector, $\boldsymbol{U}_{\boldsymbol{X}}$, and its corresponding spatial gradient with respect to the reference configuration take the form

$$\boldsymbol{U}_{\boldsymbol{X}} = g\boldsymbol{U}, \qquad \nabla_{\boldsymbol{X}} \boldsymbol{U}_{\boldsymbol{X}} = g^{-1}\boldsymbol{U}_{\boldsymbol{X}}\frac{\partial g}{\partial \boldsymbol{X}} + g \nabla \boldsymbol{U} \cdot \boldsymbol{G}, \tag{8}$$

where $\boldsymbol{G} = \nabla_{\boldsymbol{X}} \mathcal{G}$, $g = \det(\boldsymbol{G})$, $\boldsymbol{v}_{\boldsymbol{G}} = \dfrac{\partial \boldsymbol{x}}{\partial t} = \dfrac{\partial \mathcal{G}}{\partial t}$. The transformed fluxes are

$$\begin{aligned}
\boldsymbol{F}_{\boldsymbol{X}}(\boldsymbol{U}_{\boldsymbol{X}}, \nabla_{\boldsymbol{X}} \boldsymbol{U}_{\boldsymbol{X}}) &= \boldsymbol{F}^{inv}_{\boldsymbol{X}}(\boldsymbol{U}_{\boldsymbol{X}}) + \boldsymbol{F}^{vis}_{\boldsymbol{X}}(\boldsymbol{U}_{\boldsymbol{X}}, \nabla_{\boldsymbol{X}} \boldsymbol{U}_{\boldsymbol{X}}), \\
\boldsymbol{F}^{inv}_{\boldsymbol{X}}(\boldsymbol{U}_{\boldsymbol{X}}) &= g\boldsymbol{F}^{inv}(g^{-1}\boldsymbol{U}_{\boldsymbol{X}})\boldsymbol{G}^{-T} - \boldsymbol{U}_{\boldsymbol{X}} \otimes \boldsymbol{G}^{-1}\boldsymbol{v}_{\boldsymbol{G}}, \\
\boldsymbol{F}^{vis}_{\boldsymbol{X}}(\boldsymbol{U}_{\boldsymbol{X}}, \nabla_{\boldsymbol{X}} \boldsymbol{U}_{\boldsymbol{X}}) &= g\boldsymbol{F}^{vis}\left(g^{-1}\boldsymbol{U}_{\boldsymbol{X}}, g^{-1}\left[\nabla_{\boldsymbol{X}} \boldsymbol{U}_{\boldsymbol{X}} - g^{-1}\boldsymbol{U}_{\boldsymbol{X}}\frac{\partial g}{\partial \boldsymbol{X}}\right]\boldsymbol{G}^{-1}\right)\boldsymbol{G}^{-T}.
\end{aligned} \tag{9}$$

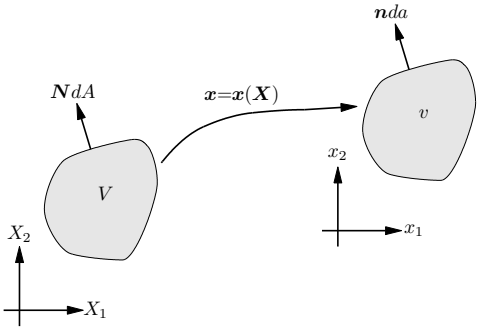

Figure 1: Time-dependent mapping between reference and physical domains.

For details regarding the derivation of the transformed equations, the reader is referred to [18].

When integrated using inexact numerical schemes, this ALE formulation does not satisfy the Geometric Conservation Law (GCL) [18,19]. This is overcome by introduction of an auxiliary variable $\bar{g}$, defined as the solution of

$$\frac{\partial \bar{g}}{\partial t} - \nabla_{\boldsymbol{X}} \cdot \left(g \boldsymbol{G}^{-1} \boldsymbol{v}_{\boldsymbol{G}} \right) = 0. \tag{10}$$

The auxiliary variable, $\bar{g}$ is used to modify the *transformed* conservation law according to

$$\left. \frac{\partial \boldsymbol{U}_{\bar{\boldsymbol{X}}}}{\partial t} \right|_{\boldsymbol{X}} + \nabla_{\boldsymbol{X}} \cdot \boldsymbol{F}_{\bar{\boldsymbol{X}}}(\boldsymbol{U}_{\bar{\boldsymbol{X}}},\ \nabla_{\boldsymbol{X}} \boldsymbol{U}_{\bar{\boldsymbol{X}}}) = 0 \tag{11}$$

where the GCL-transformed state variables are

$$\boldsymbol{U}_{\bar{\boldsymbol{X}}} = \bar{g} \boldsymbol{U}, \qquad \nabla_{\boldsymbol{X}} \boldsymbol{U}_{\bar{\boldsymbol{X}}} = \bar{g}^{-1} \boldsymbol{U}_{\bar{\boldsymbol{X}}} \frac{\partial \bar{g}}{\partial \boldsymbol{X}} + \bar{g} \nabla \boldsymbol{U} \cdot \boldsymbol{G} \tag{12}$$

and the corresponding fluxes

$$\begin{aligned} \boldsymbol{F}_{\bar{\boldsymbol{X}}}(\boldsymbol{U}_{\bar{\boldsymbol{X}}}, \nabla_{\boldsymbol{X}} \boldsymbol{U}_{\bar{\boldsymbol{X}}}) &= \boldsymbol{F}^{inv}_{\bar{\boldsymbol{X}}}(\boldsymbol{U}_{\bar{\boldsymbol{X}}}) + \boldsymbol{F}^{vis}_{\bar{\boldsymbol{X}}}(\boldsymbol{U}_{\bar{\boldsymbol{X}}}, \nabla_{\boldsymbol{X}} \boldsymbol{U}_{\bar{\boldsymbol{X}}}), \\ \boldsymbol{F}^{inv}_{\bar{\boldsymbol{X}}}(\boldsymbol{U}_{\bar{\boldsymbol{X}}}) &= g \boldsymbol{F}^{inv}(\bar{g}^{-1} \boldsymbol{U}_{\bar{\boldsymbol{X}}}) \boldsymbol{G}^{-T} - \boldsymbol{U}_{\bar{\boldsymbol{X}}} \otimes \boldsymbol{G}^{-1} \boldsymbol{v}_{\boldsymbol{G}}, \\ \boldsymbol{F}^{vis}_{\bar{\boldsymbol{X}}}(\boldsymbol{U}_{\bar{\boldsymbol{X}}}, \nabla_{\boldsymbol{X}} \boldsymbol{U}_{\bar{\boldsymbol{X}}}) &= g \boldsymbol{F}^{vis} \left(\bar{g}^{-1} \boldsymbol{U}_{\bar{\boldsymbol{X}}}, \bar{g}^{-1} \left[\nabla_{\boldsymbol{X}} \boldsymbol{U}_{\bar{\boldsymbol{X}}} - \bar{g}^{-1} \boldsymbol{U}_{\bar{\boldsymbol{X}}} \frac{\partial \bar{g}}{\partial \boldsymbol{X}} \right] \boldsymbol{G}^{-1} \right) \boldsymbol{G}^{-T}. \end{aligned} \tag{13}$$

It was shown in [18] that the transformed equations (11) satisfy the GCL.

III. Numerical discretization and the fully discrete adjoint equations

This section details the high-order numerical discretization used for the ALE formulation of the conservation law in (11). Particular attention is given to the case where the *time discretization* is parametrized as this is a consequence of parametrizing frequency (or period) with a fixed number of timesteps per period. The adjoint equations and corresponding adjoint method for computing gradients of quantities of interest from this type of discretization will be introduced. The section will conclude with important implementation details as the adjoint method corresponding to these parametrized time discretizations requires some atypical terms from the spatial discretization, namely the velocity of the spatial residual and quantity of interest.

III.A. Spatial discretization: discontinuous Galerkin method

Following the work in [20], the second-order system of partial differential equations in (11) is converted to first-order form

$$\begin{aligned} \left. \frac{\partial \bar{g}}{\partial t} \right|_{\boldsymbol{X}} + \nabla_{\boldsymbol{X}} \cdot \left(g \boldsymbol{G}^{-1} \boldsymbol{v}_{\boldsymbol{G}} \right) &= 0 \\ \left. \frac{\partial \boldsymbol{U}_{\bar{\boldsymbol{X}}}}{\partial t} \right|_{\boldsymbol{X}} + \nabla_{\boldsymbol{X}} \cdot \boldsymbol{F}_{\bar{\boldsymbol{X}}}(\boldsymbol{U}_{\bar{\boldsymbol{X}}},\ \boldsymbol{Q}_{\bar{\boldsymbol{X}}}) &= 0 \\ \boldsymbol{Q}_{\bar{\boldsymbol{X}}} - \nabla_{\boldsymbol{X}} \boldsymbol{U}_{\bar{\boldsymbol{X}}} &= 0, \end{aligned} \tag{14}$$

where $\boldsymbol{Q}_{\bar{\boldsymbol{X}}}$ is introduced as an auxiliary variable to represent the spatial gradient of the $\boldsymbol{U}_{\bar{\boldsymbol{X}}}$. Equation (14) is discretized using a standard nodal discontinuous Galerkin finite element method [21], which, after local elimination of the auxiliary variables $\boldsymbol{Q}_{\bar{\boldsymbol{X}}}$, leads to the following system of ODEs

$$\boldsymbol{M}\frac{\partial \boldsymbol{u}}{\partial t} = \boldsymbol{r}(\boldsymbol{u}, \boldsymbol{\mu}, t), \tag{15}$$

where $\boldsymbol{M}$ is the block-diagonal, symmetric, *fixed* mass matrix (state- and parameter-independent), $\boldsymbol{u}$ is the vectorization of $\begin{bmatrix}\boldsymbol{U}_{\bar{\boldsymbol{X}}}^T & \bar{g}\end{bmatrix}^T$ at all nodes in the mesh, and $\boldsymbol{r}$ is the nonlinear function defining the DG discretization of the inviscid and viscous fluxes. See [9] for an efficient treatment of $\bar{g}$ that does not lead to an *enlarged* system of ODEs.

III.B. Temporal discretization: diagonally implicit Runge-Kutta

The system of ODEs in (15) are discretized in time using Diagonally Implicit Runge-Kutta (DIRK) schemes. These schemes are capable of achieving high-order accuracy with the desired stability properties, without requiring the solution of an enlarged system of equations. The temporal domain, $[0, N_p\ T]$ is determined by the number of periods N_p and the period T, and discretized into $N_p N_t$ segments with endpoints $\{t_0, t_1, \ldots, t_{N_p N_t}\}$, with the nth segment having length $\Delta t_n = t_n - t_{n-1}$ for $n = 1, \ldots, N_p N_t$. Thus for each period, a fixed number of time steps N_t are used, which implies Δt_n depends on the parameter, i.e., $\Delta t_n = \Delta t_n(\boldsymbol{\mu})$. DIRK schemes are defined by a *lower triangular* Butcher tableau (Table 1) and take the following form when applied to (15)

$$\begin{aligned}
\boldsymbol{u}_0 &= \bar{\boldsymbol{u}}(\boldsymbol{\mu}) \\
\boldsymbol{u}_n &= \boldsymbol{u}_{n-1} + \sum_{i=1}^{s} b_i \boldsymbol{k}_{n,i} \\
\boldsymbol{M}\boldsymbol{k}_{n,i} &= \Delta t_n(\boldsymbol{\mu}) \boldsymbol{r}(\boldsymbol{u}_{n,i}, \boldsymbol{\mu}, t_{n,i}(\boldsymbol{\mu})),
\end{aligned} \tag{16}$$

for $n = 1, \ldots, N_p N_t$ and $i = 1, \ldots, s$, where s is the number of stages in the DIRK scheme. Additionally, in (16), *pstguni* is used to denote the approximation of $\boldsymbol{u}_n$ at the ith stage of time step n

$$\boldsymbol{u}_{n,i} = \boldsymbol{u}_{n,i}(\boldsymbol{u}_{n-1}, \boldsymbol{k}_{n,1}, \ldots, \boldsymbol{k}_{n,s}) = \boldsymbol{u}_{n-1} + \sum_{j=1}^{i} a_{ij} \boldsymbol{k}_{n,j} \tag{17}$$

and $t_{n,i}(\boldsymbol{\mu}) = t_{n-1}(\boldsymbol{\mu}) + c_i \Delta t_n(\boldsymbol{\mu})$ denotes the corresponding time of the stage. From (16), a complete time step requires the solution of a sequence of s nonlinear systems of equation of size $N_{\boldsymbol{u}}$.

c_1	a_{11}			
c_2	a_{21}	a_{22}		
$\vdots$	$\vdots$	$\vdots$	$\ddots$	
c_s	a_{s1}	a_{s2}	$\cdots$	a_{ss}
	b_1	b_2	$\cdots$	b_s

Table 1: Butcher Tableau for s-stage diagonally implicit Runge-Kutta scheme

III.C. Solver-consistent discretization of quantities of interest

Quantities of interest that take the form of space-time integrals of nonlinear functions that depend on the solution of the conservation law are discretized in a solver-consistent manner [9], i.e., using the same spatial and temporal discretization used for the conservation law. This ensures the truncation error of the quantities of interest exactly match that of the governing equations.

Consider a quantity of interest of the form

$$\mathcal{F}(\boldsymbol{U}, \boldsymbol{\mu}, t) = \int_0^t \int_{\boldsymbol{\Gamma}} w(\boldsymbol{x}, \tau) f(\boldsymbol{U}, \boldsymbol{\mu}, \tau)\ dS\ d\tau. \tag{18}$$

In general, $\mathcal{F}$ corresponds to either the objective or a constraint function of the optimization problem of interest. Define f_h as the approximation of $\int_{\boldsymbol{\Gamma}} w(\boldsymbol{x},\, t) f(\boldsymbol{U},\, \boldsymbol{\mu},\, t)\; dS$ using the DG shape functions from the spatial discretization of the governing equations. The solver-consistent spatial discretization of (18) becomes

$$\mathcal{F}^h(\boldsymbol{u},\, \boldsymbol{\mu},\, t) = \int_0^t f^h(\boldsymbol{u},\, \boldsymbol{\mu},\, \tau)\; d\tau, \tag{19}$$

which ensures the spatial integration error in the quantity of interest exactly matches that of the governing equations. Solver-consistent temporal discretization requires the semi-discrete functional in (19) be converted to an ODE, which is accomplished via differentiation of (19) with respect to t

$$\dot{\mathcal{F}}_h(\boldsymbol{u},\, \boldsymbol{\mu},\, t) = f_h(\boldsymbol{u},\, \boldsymbol{\mu},\, t). \tag{20}$$

Augmenting the semi-discrete governing equations with this ODE (20) yields the system of ODEs

$$\begin{bmatrix} \boldsymbol{M} & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{u}} \\ \dot{\mathcal{F}}^h \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}(\boldsymbol{u},\, \boldsymbol{\mu},\, t) \\ f^h(\boldsymbol{u},\, \boldsymbol{\mu},\, t) \end{bmatrix}. \tag{21}$$

Application of the DIRK temporal discretization introduced in Section III.B yields the fully discrete governing equations and corresponding solver-consistent discretization of the quantity of interest (18)

$$\begin{aligned} \boldsymbol{u}_n &= \boldsymbol{u}_{n-1} + \sum_{i=1}^{s} b_i \boldsymbol{k}_{n,i} \\ \mathcal{F}^h_n &= \mathcal{F}^h_{n-1} + \Delta t_n(\boldsymbol{\mu}) \sum_{i=1}^{s} b_i f^h(\boldsymbol{u}_{n,i},\; \boldsymbol{\mu},\; t_{n,i}(\boldsymbol{\mu})) \\ \boldsymbol{M}\boldsymbol{k}_{n,i} &= \Delta t_n(\boldsymbol{\mu}) \boldsymbol{r}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}(\boldsymbol{\mu})), \end{aligned} \tag{22}$$

for $n = 1, \ldots, N_p N_t$, $i = 1, \ldots, s$, and $\boldsymbol{u}_{n,i}$ is defined in (17). The output functionals considered in this paper are time-averaged quantities, which can be achieved by either integrating the quantity of interest over the entire time domain and scaling by the appropriate factor $(N_p T(\boldsymbol{\mu}))$ or choosing the weighting function such that the quantity of interest is only integrated over a single period and normalize by $T(\boldsymbol{\mu})$. The latter approach is taken in this work as there are non-physical transients associated with initializing unsteady flow simulations that would pollute the approximation of the quantity of interest; see [22] for an effective way to completely eliminate nonlinear transients associated with periodic motions. Either of these approaches fits into this framework and the fully discrete quantity of interest corresponds to $\mathcal{F}^h_{N_p N_t}$ to yield the solver-consistent approximation of $\mathcal{F}(\boldsymbol{u},\, \boldsymbol{\mu},\, T)$

$$F(\boldsymbol{u}_0,\, \ldots,\, \boldsymbol{u}_{N_p N_t},\, \boldsymbol{k}_{1,1},\, \ldots,\, \boldsymbol{k}_{N_p N_t, s},\, \boldsymbol{\mu}) = \mathcal{F}^h_{N_p N_t} = \sum_{n=1}^{N_p N_t} \Delta t_n(\boldsymbol{\mu}) \sum_{i=1}^{s} b_i f^h(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}(\boldsymbol{\mu})). \tag{23}$$

For brevity, the dependence of $t_{n,i}$ and Δt_n on $\boldsymbol{\mu}$ will be dropped in the remainder.

III.D. Fully Discrete, Time-Dependent Adjoint Equations

This section develops the adjoint equations corresponding to the fully discrete system of conservation laws in (16) and the adjoint method for computing the total derivative of the fully discrete quantity of interest F in (23) without requiring solution sensitivities, $\partial_{\boldsymbol{\mu}} \boldsymbol{u}_n$ and $\partial_{\boldsymbol{\mu}} \boldsymbol{k}_{n,i}$. It is emphasized that F represents *any* quantity of interest whose gradient is desired, such as the optimization objective function or a constraint. A distinguishing feature of this work is the fully discrete adjoint method must account for the parametrization of *time*, i.e., $t_{n,i}(\boldsymbol{\mu})$ and $\Delta t_n(\boldsymbol{\mu})$, for the gradient $\nabla_{\boldsymbol{\mu}} F$ to be consistent. The parametrized time discretization does not modify the adjoint equations themselves, only the formula to *reconstruct* $\nabla_{\boldsymbol{\mu}} F$ from the adjoint solutions.

The following definitions are introduced for the Runge-Kutta stage equations and state updates

$$\begin{aligned} \tilde{\boldsymbol{r}}_0(\boldsymbol{u}_0,\, \boldsymbol{\mu}) &= \; \boldsymbol{u}_0 - \bar{\boldsymbol{u}}_0(\boldsymbol{\mu}) = 0 \\ \tilde{\boldsymbol{r}}_n(\boldsymbol{u}_{n-1},\, \boldsymbol{u}_n,\, \boldsymbol{k}_{n,1},\, \ldots,\, \boldsymbol{k}_{n,s}) &= \; \boldsymbol{u}_n - \boldsymbol{u}_{n-1} - \sum_{i=1}^{s} b_i \boldsymbol{k}_{n,i} = 0 \\ \boldsymbol{R}_{n,i}(\boldsymbol{u}_{n-1},\, \boldsymbol{k}_{n,1},\, \ldots,\, \boldsymbol{k}_{n,i},\, \boldsymbol{\mu}) &= \; \boldsymbol{M}\boldsymbol{k}_{n,i} - \Delta t_n \boldsymbol{r}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) = 0. \end{aligned} \tag{24}$$

It was shown in [9] that the adjoint equations corresponding to a primal timestepping scheme with the same structure as (24) are

$$\begin{aligned}
\frac{\partial \tilde{\boldsymbol{r}}_{N_pN_t}}{\partial \boldsymbol{u}_{N_pN_t}}^T \boldsymbol{\lambda}_{N_pN_t} &= \frac{\partial F}{\partial \boldsymbol{u}_{N_pN_t}}^T \\
\frac{\partial \tilde{\boldsymbol{r}}_n}{\partial \boldsymbol{u}_{n-1}}^T \boldsymbol{\lambda}_n + \frac{\partial \tilde{\boldsymbol{r}}_{n-1}}{\partial \boldsymbol{u}_{n-1}}^T \boldsymbol{\lambda}_{n-1} &= \frac{\partial F}{\partial \boldsymbol{u}_{n-1}}^T - \sum_{i=1}^{s} \frac{\partial \boldsymbol{R}_{n,i}}{\partial \boldsymbol{u}_{n-1}}^T \boldsymbol{\kappa}_{n,i} \\
\sum_{j=i}^{s} \frac{\partial \boldsymbol{R}_{n,j}}{\partial \boldsymbol{k}_{n,i}}^T \boldsymbol{\kappa}_{n,j} &= \frac{\partial F}{\partial \boldsymbol{k}_{n,i}}^T - \frac{\partial \tilde{\boldsymbol{r}}_n}{\partial \boldsymbol{k}_{n,i}}^T \boldsymbol{\lambda}_n
\end{aligned} \tag{25}$$

for $n = 1, \ldots, N_pN_t$ and $i = 1, \ldots, s$, where $\boldsymbol{\lambda}_n$ and $\boldsymbol{\kappa}_{n,i}$ are the adjoint state and stage variables, respectively. Furthermore, it was shown in [9] that the gradient computation reduces to

$$\frac{\mathrm{d}F}{\mathrm{d}\boldsymbol{\mu}} = \frac{\partial F}{\partial \boldsymbol{\mu}} - \boldsymbol{\lambda}_0^T \frac{\partial \tilde{\boldsymbol{r}}_n}{\partial \boldsymbol{\mu}} - \sum_{n=1}^{N_pN_t} \sum_{p=1}^{s} \boldsymbol{\kappa}_{n,p}^T \frac{\partial \boldsymbol{R}_{n,p}}{\partial \boldsymbol{\mu}} \tag{26}$$

which is *independent* of the state sensitivities. Elimination of the auxiliary terms, $\tilde{\boldsymbol{r}}_n$ and $\boldsymbol{R}_{n,i}$, in equations (25) through differentiation of their expressions in (24) gives rise to the exact adjoint equations from [9]

$$\begin{aligned}
\boldsymbol{\lambda}_{N_pN_t} &= \frac{\partial F}{\partial \boldsymbol{u}_{N_pN_t}}^T \\
\boldsymbol{\lambda}_{n-1} &= \boldsymbol{\lambda}_n + \frac{\partial F}{\partial \boldsymbol{u}_{n-1}}^T + \sum_{i=1}^{s} \Delta t_n \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{u}}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i})^T \boldsymbol{\kappa}_{n,i} \\
\boldsymbol{M}^T \boldsymbol{\kappa}_{n,i} &= \frac{\partial F}{\partial \boldsymbol{k}_{n,i}}^T + b_i \boldsymbol{\lambda}_n + \sum_{j=i}^{s} a_{ji} \Delta t_n \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{u}}(\boldsymbol{u}_{n,j},\, \boldsymbol{\mu},\, t_{n,j})^T \boldsymbol{\kappa}_{n,j}
\end{aligned} \tag{27}$$

for $n = 1, \ldots, N_pN_t$ and $i = 1, \ldots, s$. The expression for $\nabla_{\boldsymbol{\mu}} F$, independent of state sensitivities, is

$$\begin{aligned}
\frac{\mathrm{d}F}{\mathrm{d}\boldsymbol{\mu}} = \boldsymbol{\lambda}_0^T \frac{\partial \bar{\boldsymbol{u}}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}) &+ \sum_{n=1}^{N_pN_t} \Delta t_n \sum_{i=1}^{s} b_i \left[\frac{\partial f^h}{\partial \boldsymbol{\mu}}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) + \frac{\partial f^h}{\partial t}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) \frac{\partial t_{n,i}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}) \right] \\
&+ \sum_{n=1}^{N_pN_t} \Delta t_n \sum_{i=1}^{s} \boldsymbol{\kappa}_{n,i}^T \left[\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\mu}}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) + \frac{\partial \boldsymbol{r}}{\partial t}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) \frac{\partial t_{n,i}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}) \right] \\
&+ \sum_{n=1}^{N_pN_t} \sum_{i=1}^{s} b_i f^h(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) \frac{\partial \Delta t_n}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}) + \sum_{n=1}^{N_pN_t} \sum_{i=1}^{s} \boldsymbol{\kappa}_{n,i}^T \boldsymbol{r}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) \frac{\partial \Delta t_n}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu})
\end{aligned} \tag{28}$$

where the following identities were used

$$\begin{aligned}
\frac{\partial F}{\partial \boldsymbol{\mu}} &= \sum_{n=1}^{N_pN_t} \Delta t_n \sum_{i=1}^{s} b_i \left[\frac{\partial f^h}{\partial \boldsymbol{\mu}}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) + \frac{\partial f^h}{\partial t}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) \frac{\partial t_{n,i}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}) \right] + \sum_{n=1}^{N_pN_t} \sum_{i=1}^{s} b_i f^h(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) \frac{\partial \Delta t_n}{\partial \boldsymbol{\mu}} \\
\frac{\partial \boldsymbol{R}_{n,i}}{\partial \boldsymbol{\mu}} &= \sum_{n=1}^{N_pN_t} \Delta t_n \sum_{i=1}^{s} b_i \left[\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\mu}}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) + \frac{\partial \boldsymbol{r}}{\partial t}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) \frac{\partial t_{n,i}}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}) \right] + \sum_{n=1}^{N_pN_t} \sum_{i=1}^{s} b_i \boldsymbol{r}(\boldsymbol{u}_{n,i},\, \boldsymbol{\mu},\, t_{n,i}) \frac{\partial \Delta t_n}{\partial \boldsymbol{\mu}},
\end{aligned} \tag{29}$$

i.e., the partial derivatives with respect to $\boldsymbol{\mu}$ must account the the explicit dependence of f^h and $\boldsymbol{r}$ on $\boldsymbol{\mu}$ as well as the dependence of the time discretization on $\boldsymbol{\mu}$. The only sensitivity required to compute $\nabla_{\boldsymbol{\mu}} F(\boldsymbol{\mu})$ is the sensitivity with respect to the initial condition. This term is zero for the application considered in this work since the initial condition is uniform flow, regardless of $\boldsymbol{\mu}$. For more general initial conditions, the product $\boldsymbol{\lambda}_0^T \partial_{\boldsymbol{\mu}} \bar{\boldsymbol{u}}$ can be computed efficiently; see [9] for details.

III.E. Implementation

Implementation of the fully discrete adjoint method introduced in Section III.D relies on the computation of the following terms from the spatial discretization

$$\boldsymbol{M}, \boldsymbol{r}, \frac{\partial \boldsymbol{r}}{\partial t}, \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{u}}, \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\mu}}, f^h, \frac{\partial f^h}{\partial \boldsymbol{u}}, \frac{\partial f^h}{\partial \boldsymbol{\mu}}, \frac{\partial f^h}{\partial t}, \tag{30}$$

as well as terms from the time discretization

$$t_{n,i}, \Delta t_n, \frac{\partial t_{n,i}}{\partial \boldsymbol{\mu}}, \frac{\partial \Delta t_n}{\partial \boldsymbol{\mu}}. \tag{31}$$

The mass matrix and derivatives with respect to the state $\boldsymbol{u}$ are standard terms required by an implicit solver and will not be considered further. From the form of the governing equations in (11), the residual $\boldsymbol{r}$ and quantity of interest f_h solely depend on the state vector, domain position $\boldsymbol{x}$, and domain velocity $\dot{\boldsymbol{x}}$ since the deformation gradient $\boldsymbol{G}$ and its determinant g can be written solely in terms of $\boldsymbol{x}$; see [23] for a detailed description. Therefore, the residual and quantity of interest can be written as

$$\begin{aligned} \boldsymbol{r}(\boldsymbol{u}, \boldsymbol{\mu}, t) &= \boldsymbol{r}(\boldsymbol{u}, \boldsymbol{x}(\boldsymbol{\mu}, t), \dot{\boldsymbol{x}}(\boldsymbol{\mu}, t)) \\ f^h(\boldsymbol{u}, \boldsymbol{\mu}, t) &= f^h(\boldsymbol{u}, \boldsymbol{x}(\boldsymbol{\mu}, t), \dot{\boldsymbol{x}}(\boldsymbol{\mu}, t)). \end{aligned}$$

Then the derivatives of $\boldsymbol{r}$ and f^h with respect to parameter, $\boldsymbol{\mu}$, and time, t, that are required to construct $\nabla_{\boldsymbol{\mu}} F$ from the adjoint solution are computed as

$$\begin{aligned} \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\mu}} &= \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\mu}} + \frac{\partial \boldsymbol{r}}{\partial \dot{\boldsymbol{x}}}\frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{\mu}} & \frac{\partial f_h}{\partial \boldsymbol{\mu}} &= \frac{\partial f_h}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\mu}} + \frac{\partial f_h}{\partial \dot{\boldsymbol{x}}}\frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{\mu}} \\ \frac{\partial \boldsymbol{r}}{\partial t} &= \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial t} + \frac{\partial \boldsymbol{r}}{\partial \dot{\boldsymbol{x}}}\frac{\partial \dot{\boldsymbol{x}}}{\partial t} & \frac{\partial f_h}{\partial t} &= \frac{\partial f_h}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial t} + \frac{\partial f_h}{\partial \dot{\boldsymbol{x}}}\frac{\partial \dot{\boldsymbol{x}}}{\partial t} \end{aligned} \tag{32}$$

where the terms

$$\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{x}}, \frac{\partial \boldsymbol{r}}{\partial \dot{\boldsymbol{x}}}, \frac{\partial f_h}{\partial \boldsymbol{x}}, \frac{\partial f_h}{\partial \dot{\boldsymbol{x}}} \tag{33}$$

are determined from the form of the governing equations and spatial discretization outlined in [9]. The remaining terms

$$\frac{\partial \boldsymbol{x}}{\partial t}, \frac{\partial \dot{\boldsymbol{x}}}{\partial t}, \tag{34}$$

are the velocity and acceleration of the motion of the domain, respectively, and

$$\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\mu}}, \frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{\mu}}, \tag{35}$$

are the sensitivities of the domain position and velocity, respectively. These terms can be computed from the specific form of the domain mapping (Section IV). The parametrization of the temporal discretization is problem-specific and a detailed description will be deferred to Section IV.

IV. Application to energetically optimal flapping flight

In this section, the high-order numerical discretization of the isentropic, compressible Navier-Stokes equations and corresponding adjoint method are applied to determine the energetically optimal flapping motion, with a requirement on the generated thrust, of a 2D NACA0012 airfoil (Figure 2) using gradient-based optimization. A novel component of this work is the inclusion of the flapping frequency as an optimization parameter. As discussed in Section III.B, assuming a fixed number of timesteps per period are desired, this implies the timestep is parameter-dependent $\Delta t_n(\boldsymbol{\mu})$.

IV.A. Geometry and kinematics of flapping airfoil

The motion of the airfoil is given as a combination of pitching and heaving, which each take the form of a single harmonic function parametrized by frequency, amplitude, and phase shift

$$\begin{aligned} y(\boldsymbol{\mu},\, t) &= y_{\max} \sin(2\pi f t) \\ \theta(\boldsymbol{\mu},\, t) &= \theta_{\max} \sin(2\pi f t + \phi) \end{aligned} \tag{36}$$

where $\boldsymbol{\mu} = [f,\, y_{\max},\, \theta_{\max},\, \phi]$. In this work, the parameters are restricted to the following intervals: $f \in [0.05,\, 0.5]$, $y_{\max} \in [0,\, 2]$, $\theta_{\max} \in [0°,\, 60°]$, $\phi \in [-90°,\, 90°]$. From the rigid body motion of the airfoil, the

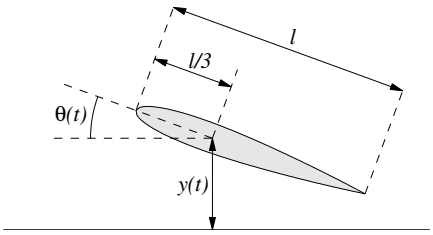

Figure 2: Geometry and kinematics of NACA0012 airfoil used in energetically optimal flapping study

deforming fluid domain is determined analytically by blending the rigid body motion of the entire fluid domain with the reference mesh [18]. Let $\boldsymbol{X}' = [X_1, X_2]^T$ denote the coordinates of the fluid domain after application of the rigid body motion corresponding to $y(\boldsymbol{\mu},\, t)$ and $\theta(\boldsymbol{\mu},\, t)$

$$\begin{aligned} X_1' &= X_1 \cos\theta(\boldsymbol{\mu},\, t) + X_2 \sin\theta(\boldsymbol{\mu},\, t) \\ X_2' &= y(\boldsymbol{\mu},\, t) - X_1 \sin\theta(\boldsymbol{\mu},\, t) + X_2 \cos\theta(\boldsymbol{\mu},\, t). \end{aligned} \tag{37}$$

Then the coordinates of the physical fluid domain, $\boldsymbol{x}'$, are defined by smoothly blending the domain from the desired rigid body motion near the airfoil to the reference domain as

$$\boldsymbol{x}' = (1 - b(\boldsymbol{X}))\boldsymbol{X}' + b(\boldsymbol{X})\boldsymbol{X} \tag{38}$$

where the blending maps from [9, 18] take the form

$$b(\boldsymbol{X}) = \begin{cases} 0 & d(\boldsymbol{X}) \le r_1 \\ 1 & d(\boldsymbol{X}) \ge r_1 + r_2 \\ q\left(\frac{d(\boldsymbol{X}) - r_2}{r_1}\right) & \text{otherwise} \end{cases} \tag{39}$$

and $q(s) = 3s^2 - 2s^3$. In this work, the blending radii are set to $r_1 = 2.0$, $r_2 = 4.0$.

The expression for the deformed domain, $\boldsymbol{x}'(\boldsymbol{X},\, \boldsymbol{\mu},\, t)$, in (38) will have a non-trivial deformation and velocity at $t = 0$. This may cause difficulty in initializing the simulation from uniform flow as violent transients will result that may cause the simulation to crash. For this reason, following the work in [8, 9], the deformation is smoothly blended to zero at $t = 0$ using the infinitely differentiable blending

$$b_t(t) = e^{-(t/T_c)^2}. \tag{40}$$

Temporal blendings have also been used in experimental studies involving flapping wings [24], where a quintic blending was used. The final form of the deformed domain is

$$\boldsymbol{x}(\boldsymbol{X},\, \boldsymbol{\mu},\, t) \coloneqq (1 - b_t(t))\boldsymbol{x}'(\boldsymbol{X},\, \boldsymbol{\mu},\, t) + b_t(t)\boldsymbol{X} \tag{41}$$

and the domain velocity $\dot{\boldsymbol{x}}(\boldsymbol{X},\, \boldsymbol{\mu},\, t)$ can be computed analytically. It can easily be verified that this temporal blending guarantees $\boldsymbol{x}(\boldsymbol{X},\, \boldsymbol{\mu},\, 0) = \boldsymbol{X}$ and $\dot{\boldsymbol{x}}(\boldsymbol{X},\, \boldsymbol{\mu},\, 0) = \boldsymbol{0}$. In this work, $T_c = 0.4$, which is 10% of the fastest flapping motion considered in this work and ensures $\boldsymbol{x}$, $\dot{\boldsymbol{x}}$ are effectively equal to $\boldsymbol{x}'$, $\dot{\boldsymbol{x}}'$ (within 0.1%) within 1/2 a period. This blending limits the transients that result from initializing the flow with incompatible boundary conditions at the viscous wall and implies the sensitivity of the initial condition is zero, i.e., $\frac{\partial \boldsymbol{u}_0}{\partial \boldsymbol{\mu}} = 0$, since $\boldsymbol{x}(\boldsymbol{X},\, \boldsymbol{\mu},\, 0) = \boldsymbol{X}$.

IV.B. Energetically optimal flapping under thrust constraint

In this section, the high-order numerical discretization of the isentropic, compressible Navier-Stokes equations and corresponding adjoint method are applied to determine the energetically optimal flapping motion of the geometry introduced in the previous section using gradient-based optimization techniques. For a physically relevant mission, requirements are placed on the time-averaged thrust leading to an optimization problem with a single nonlinear constraint. As a result, two adjoint equations must be solved at each optimization iteration to compute the gradient of the objective function and nonlinear constraint. The reconstruction of these gradients from the adjoint solution includes the contributions from the parameter-dependent timestep, detailed in (28).

The DG-ALE scheme introduced in Section II is used for the spatial discretization of the system of conservation laws with polynomial order $p = 3$ (971 elements). The Reynolds number of the flow (with respect to the chord length and freestream velocity) is set to $Re = 1000$ and the Mach number is set to $M = 0.02$, which makes the flow nearly incompressible. The DG-ALE scheme uses the Roe flux [25] for the inviscid numerical flux and the Compact DG flux [26] for the viscous numerical flux. The diagonally implicit Runge-Kutta scheme detailed in Section III.B is used for the temporal discretization with $N_t = 100$ uniform, three-stage, third-order timesteps used per period; see Table 2 for the Butcher tableau. Since the timesteps

α	α			
$\frac{1+\alpha}{2}$	$\frac{1+\alpha}{2} - \alpha$	α		
1	γ		ω	α
	γ		ω	α

Table 2: Butcher Tableau for 3-stage, 3rd order DIRK scheme [27]
$\alpha = 0.435866521508459$, $\gamma = -\frac{6\alpha^2 - 16\alpha + 1}{4}$, $\omega = \frac{6\alpha^2 - 20\alpha + 5}{4}$.

are taken to be uniform within a period, an explicit formula for $\Delta t_n(\boldsymbol{\mu})$ is obtained as

$$\Delta t_n(\boldsymbol{\mu}) = \Delta t(\boldsymbol{\mu}) = \frac{T(\boldsymbol{\mu})}{N_t} = \frac{1}{f(\boldsymbol{\mu}) N_t} = \frac{1}{\mu_1 N_t}. \tag{42}$$

From this, the time at the ith DIRK stage, $t_{n,i}$, is parametrized as

$$t_{n,i}(\boldsymbol{\mu}) = t_{n-1}(\boldsymbol{\mu}) + c_i \Delta t_n(\boldsymbol{\mu}) = (n - 1 + c_i)\Delta t(\boldsymbol{\mu}) = \frac{n - 1 + c_i}{\mu_1 N_t}. \tag{43}$$

The quantities of interest for the optimization problem are: the total work done by the foil on the fluid $\mathcal{W}(\boldsymbol{U}, \boldsymbol{\mu})$ and the average thrust, $\mathcal{T}_x(\boldsymbol{U}, \boldsymbol{\mu})$, generated over one flapping period. These quantities are defined as

$$\begin{aligned} \mathcal{W}(\boldsymbol{U}, \boldsymbol{\mu}) &= -\frac{1}{T(\boldsymbol{\mu})} \int_{T(\boldsymbol{\mu})}^{2T(\boldsymbol{\mu})} \int_{\boldsymbol{\Gamma}} \boldsymbol{f}(\boldsymbol{U}, \boldsymbol{\mu}) \cdot \dot{\boldsymbol{x}} \; dS \; dt \\ \mathcal{T}_x(\boldsymbol{U}, \boldsymbol{\mu}) &= -\frac{1}{T(\boldsymbol{\mu})} \int_{T(\boldsymbol{\mu})}^{2T(\boldsymbol{\mu})} \int_{\boldsymbol{\Gamma}} \boldsymbol{f}(\boldsymbol{U}, \boldsymbol{\mu}) \cdot \boldsymbol{e}_1 \; dS \; dt \end{aligned} \tag{44}$$

where $\boldsymbol{\Gamma}$ is the surface of the foil, $\boldsymbol{f} \in \mathbb{R}^3$ is the force imparted by the fluid on the body, $\boldsymbol{e}_i \in \mathbb{R}^3$ is the ith canonical basis vector, and $\dot{\boldsymbol{x}}$ is the velocity of each point on $\boldsymbol{\Gamma}$. The fully discrete, high-order approximation of the integrated quantities of interest (DG in space, DIRK in time) will be denoted with the corresponding Roman symbol, e.g., $W(\boldsymbol{u}_0, \ldots, \boldsymbol{u}_{N_p N_t}, \boldsymbol{k}_{1,1}, \ldots, \boldsymbol{k}_{N_p N_t, s}, \boldsymbol{\mu})$ is the fully discrete approximation of $\mathcal{W}(\boldsymbol{U}, \boldsymbol{\mu})$

and similarly for T_x. With these definitions, the fully discrete optimization problem takes the form

$$\begin{aligned}
&\underset{\substack{\boldsymbol{u}_0,\ldots,\boldsymbol{u}_{N_pN_t}\in\mathbb{R}^{N_{\boldsymbol{u}}},\\ \boldsymbol{k}_{1,1},\ldots,\boldsymbol{k}_{N_pN_t,s}\in\mathbb{R}^{N_{\boldsymbol{u}}},\\ \boldsymbol{\mu}\in\mathbb{R}^{N_{\boldsymbol{\mu}}}}}{\text{minimize}} && W(\boldsymbol{u}_0,\ldots,\boldsymbol{u}_{N_pN_t},\boldsymbol{k}_{1,1},\ldots,\boldsymbol{k}_{N_pN_t,s},\boldsymbol{\mu})\\
&\text{subject to} && T_x(\boldsymbol{u}_0,\ldots,\boldsymbol{u}_{N_pN_t},\boldsymbol{k}_{1,1},\ldots,\boldsymbol{k}_{N_pN_t,s},\boldsymbol{\mu})=\bar{T}_x\\
& && \boldsymbol{u}_0=\bar{\boldsymbol{u}}\\
& && \boldsymbol{u}_n=\boldsymbol{u}_{n-1}+\sum_{i=1}^{s}b_i\boldsymbol{k}_{n,i}\\
& && \boldsymbol{M}\boldsymbol{k}_{n,i}=\Delta t_n(\boldsymbol{\mu})\boldsymbol{r}(\boldsymbol{u}_{n,i},\boldsymbol{\mu},t_{n,i}(\boldsymbol{\mu})).
\end{aligned} \tag{45}$$

In this work, the optimization problem in (45) is solved using IPOPT [28].

First, the convergence of the total work and thrust as a function of optimization iteration for a fixed $\bar{T}_x = 1.0$ is considered (Figure 3). Since the initial guess ($f = 0.2$, $y_{\max} = 0.0$, $\theta_{\max} = 0.0$, $\phi = 90°$) is far from the optimal solution ($f = 0.227$, $y_{\max} = 1.52$, $\theta_{\max} = 49.4°$, $\phi = 85.8°$), several iterations are required as the optimizer searches for parameters that satisfy the thrust constraint and minimize the energy required to flap. After 20 iterations the optimizer has found a solution that is feasible and energetically optimal to a tolerance of 10^{-4} and only a few additional iterations are required to fine tune the motion such that the optimality condition is satisfied to 10^{-8}. The remainder of the optimization problems considered in this document, for various values of $\bar{T}_x$, are initialized from the optimal solution corresponding to $\bar{T}_x = 1$ or the optimal solution corresponding to another value of $\bar{T}_x$ that is available.

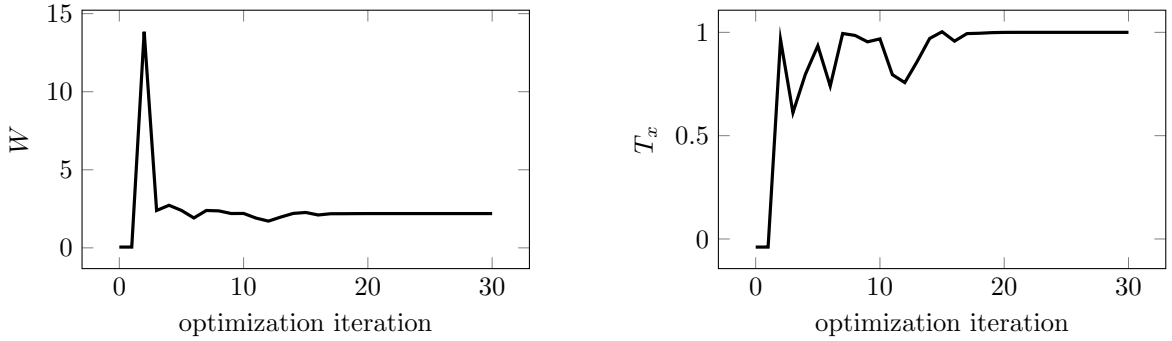

Figure 3: Convergence of required flapping energy and thrust as a function of optimization iteration corresponding to the thrust constraint $\bar{T}_x = 1$. The final values are $W^* = 2.1961022$ and $T_x^* = 0.9999999$ and the first-order optimality conditions are satisfied to a tolerance of 10^{-8}. For a convergence tolerance of 10^{-4}, the optimization iterations could have been terminated after 20 iterations.

In the remainder of this document, we study the behavior of the energetically optimal flapping motions as a function of the constraint $\bar{T}_x$. The optimal pitching and heaving motions corresponding to various values of $\bar{T}_x$ are provided in Figure 4. Additionally, the energy (W^*) and parameters (f^*, $y^*_{\max}$, $\theta^*_{\max}$, ϕ^*) corresponding to the energetically optimal flapping motions corresponding to these values of the thrust constraint $\bar{T}_x$ are provided in Figure 5.

At this point a number of observations can be made regarding these energetically optimal flapping motions as a function of $\bar{T}_x$. First, the energy required to flap increases linear with $\bar{T}_x$, at least in the regime considered. The optimal frequency also increases as $\bar{T}_x$ increases, but the increase is not linear. For large values of $\bar{T}_x$ the benefit of flapping faster, i.e., increasing frequency, diminishes and larger values of pitch and heave are preferred. This suggests that increasing flapping frequency is not necessary the most energetically optimal way to generate additional thrust. The trends for the maximum pitch and heave amplitude are less clear due to outliers at $\bar{T}_x = 0.75$ and $\bar{T}_x = 2.5$ where $y^*_{\max}$ and $\theta^*_{\max}$ hit their upper bounds. It is likely that a local minima was obtained and a better starting point would cause these quantities to exhibit smoother trends. It is interesting to note that despite these outliers in the optimal trajectory, the optimal work (W^*) still exhibits a nearly linear behavior as a function of $\bar{T}_x$.

The motion of the airfoil and vorticity of the surrounding flow are shown in Figure 6 (optimal motion corresponding to $\bar{T}_x = 0$), Figure 7 (optimal motion corresponding to $\bar{T}_x = 1$), and Figure 8 (optimal

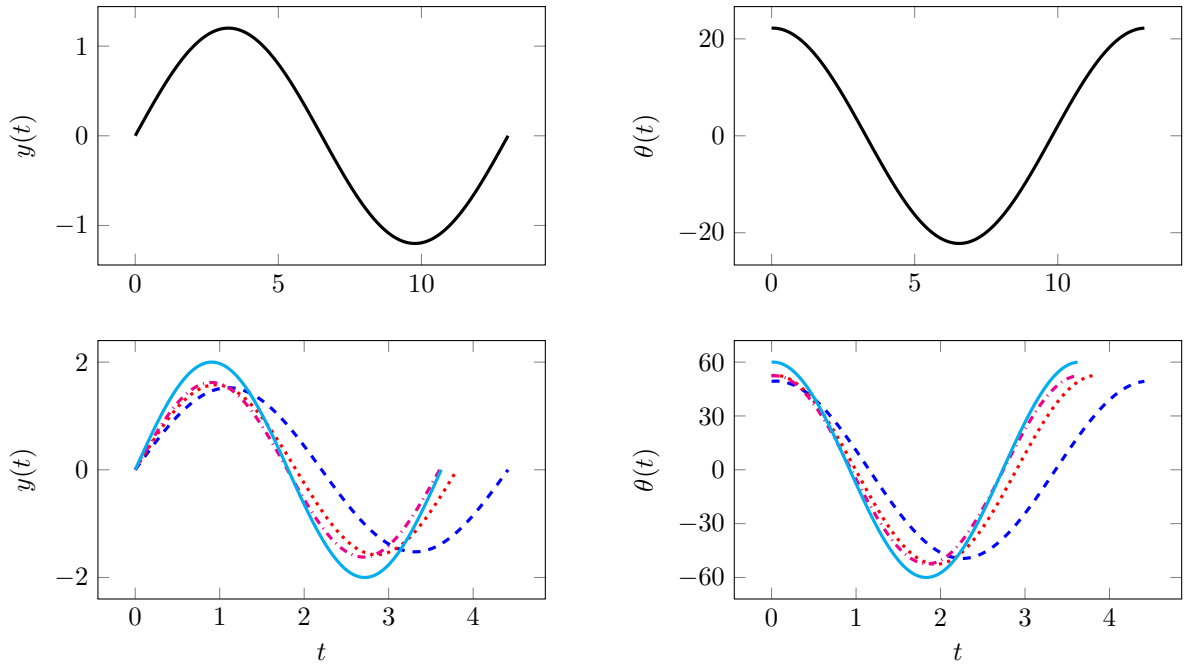

Figure 4: Optimal trajectories of $y(t)$ and $\theta(t)$ for various value of the thrust constraint: $\bar{T}_x = 0.0$ (——), $\bar{T}_x = 1.0$ (– – –), $\bar{T}_x = 1.5$ (·····), $\bar{T}_x = 2.0$ (–·–·–), $\bar{T}_x = 2.5$ (——).

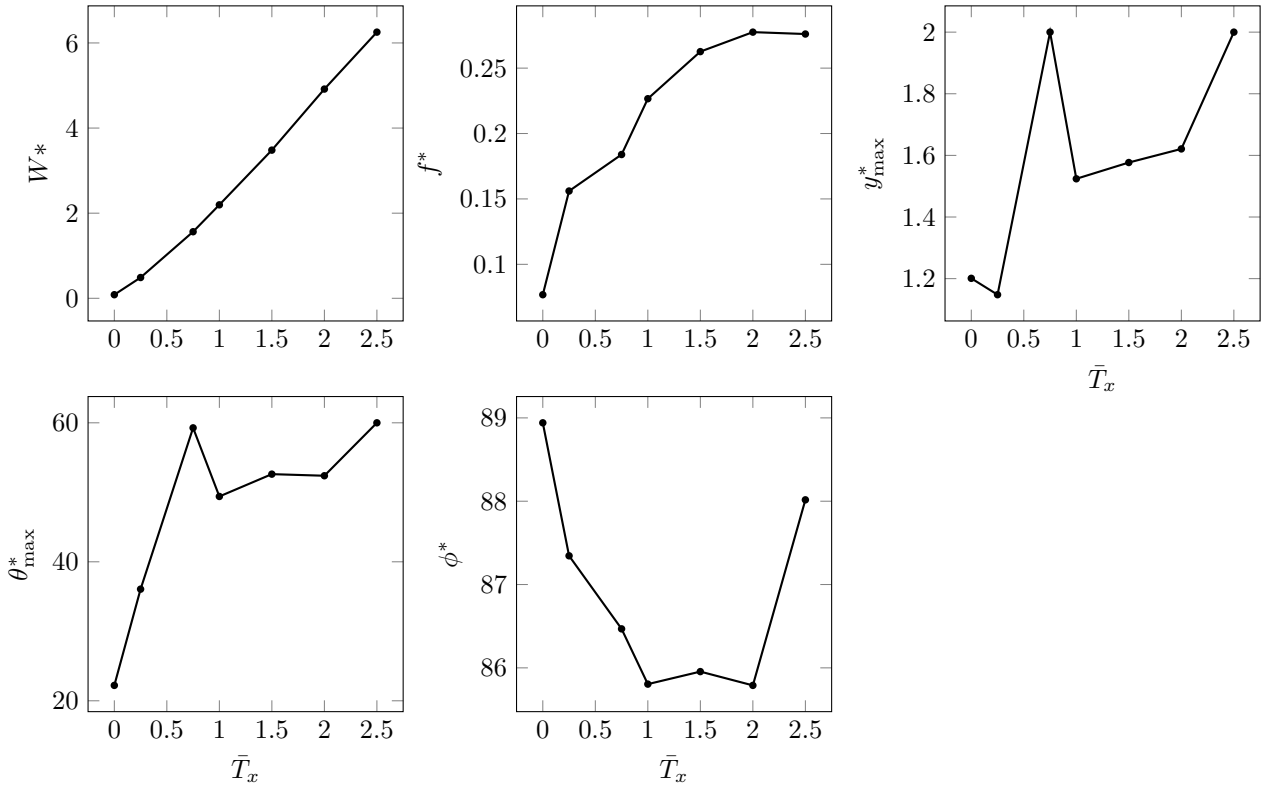

Figure 5: The optimal flapping energy (W^*), frequency (f^*), maximum heaving amplitude ($y^*_{\max}$), maximum pitching amplitude ($\theta^*_{\max}$), and phase between pitching and heaving (ϕ^*) as a function of the thrust constraint $\bar{T}_x$.

motion corresponding to $\bar{T}_x = 2.5$). As the desired thrust increases, the amplitude of the pitch and heave increase significantly and the flapping frequency increases causing vortices to shed when the airfoil attains its maximum/minimum heaving amplitude. Despite these complex interactions occurring in the flow, the work required to perform these optimal trajectories increases (roughly) linearly as $\bar{T}_x$ increases.

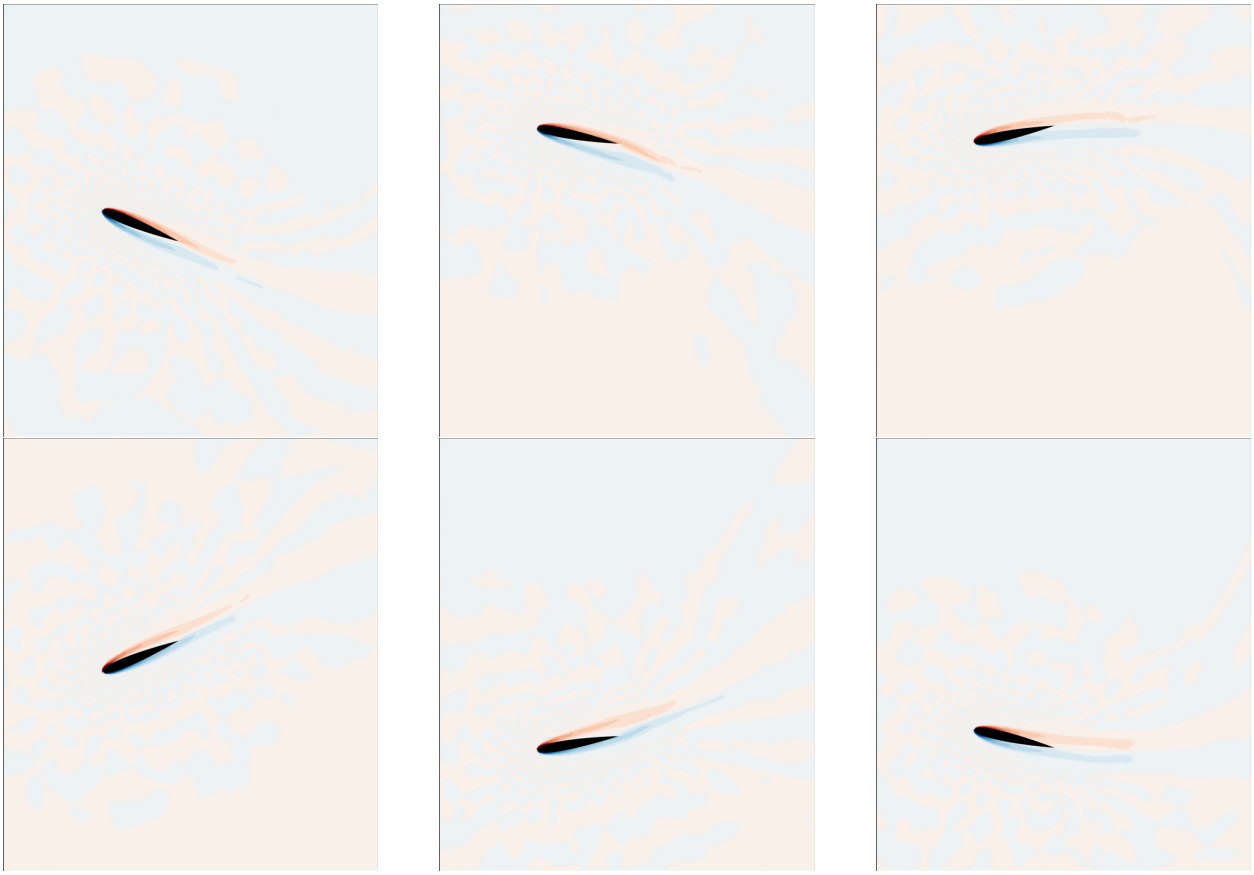

Figure 6: Trajectory of airfoil and flow vorticity at energetically optimal, $\bar{T}_x = 0.0$ flapping motion (see Figure 4 for flapping trajectory). Snapshots taken at times 6 equally spaced instances throughout the period ($T(\boldsymbol{\mu}) = 13.0$): $t = 0.0,\ 2.17,\ 4.34,\ 6.51,\ 8.68,\ 10.9$.

V. Conclusion

This work introduces a fully discrete, high-order numerical framework for solving PDE-constrained optimization problems using gradient-based methods in the case where one or more of the optimization parameters affects the time domain. This leads to a fully discrete system whose time discretization is parametrized in addition to the parameter-dependence of the spatial quantities. The adjoint method corresponding to this system was introduced and shown that the adjoint equations themselves are identical to the case with a fixed time discretization, but the formula to reconstruct the gradient of a quantity of interest from the adjoint solution involves additional terms. These additional terms involve the velocity of the quantity of interest and spatial residual as well as the sensitivities of the time discretization with respect to the parameters.

The fully discrete primal and adjoint equations were employed to study energetically optimal flapping subject to a minimum required thrust, where flapping frequency and trajectory were included as optimization parameters. It was shown that convergence to an energetically optimal motion for a desired thrust of $\bar{T}_x = 1$ was obtained in only 20 optimization iterations. Additionally, a study was performed to gain insight to the optimal flapping energy, frequency, and trajectory as a function of the desired thrust, $\bar{T}_x$. It was shown that the optimal flapping energy increases roughly linearly with $\bar{T}_x$ while trends in the other quantities are less clear, possibly due to some outliers at $\bar{T}_x = 0.75$ and $\bar{T}_x = 2.5$ where optimization parameters hit their bounds and could be the result of sub-optimal local minima. The linear behavior in the optimal flapping

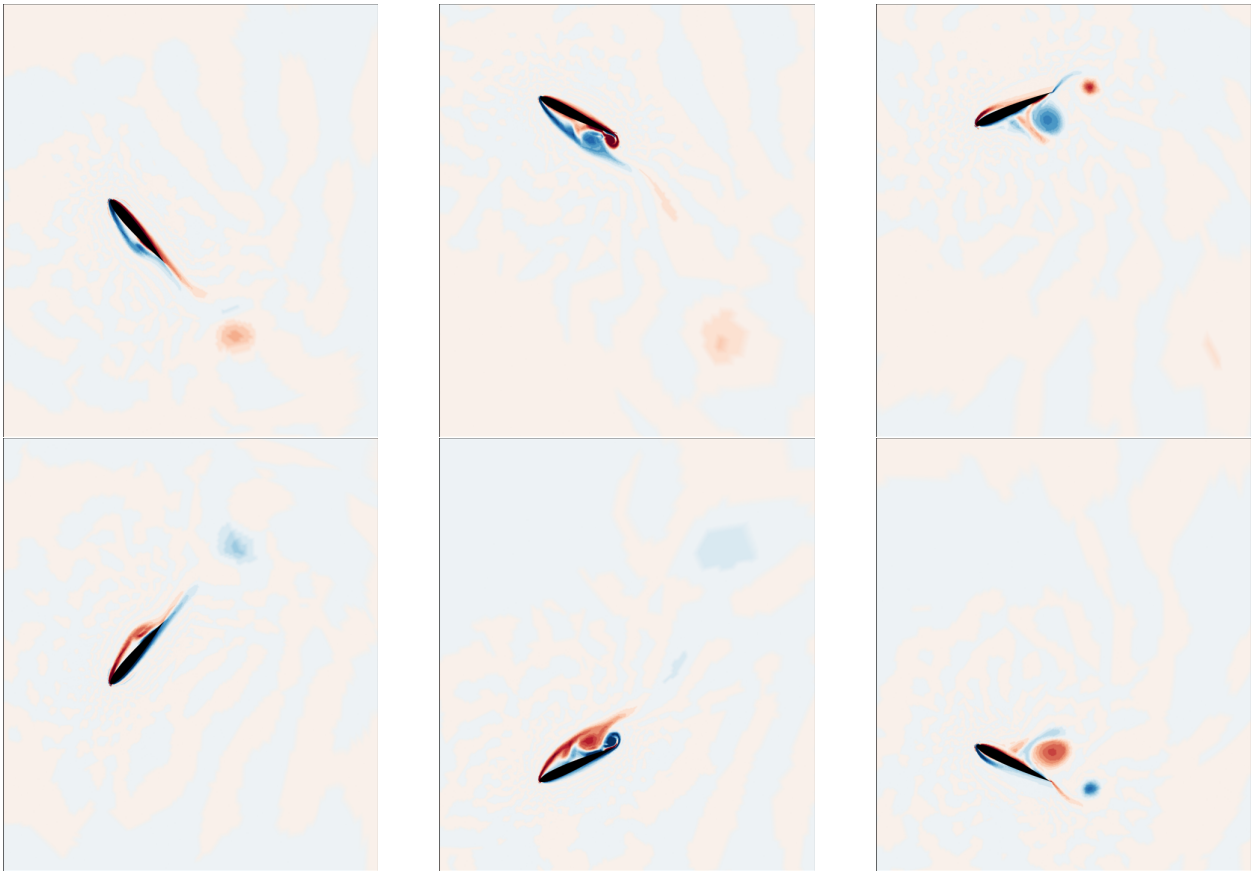

Figure 7: Trajectory of airfoil and flow vorticity at energetically optimal, $\bar{T}_x = 1.0$ flapping motion (see Figure 4 for flapping trajectory). Snapshots taken at times 6 equally spaced instances throughout the period ($T(\boldsymbol{\mu}) = 4.41$): $t = 0.0,\ 0.735,\ 1.47,\ 2.21,\ 2.94,\ 3.68$.

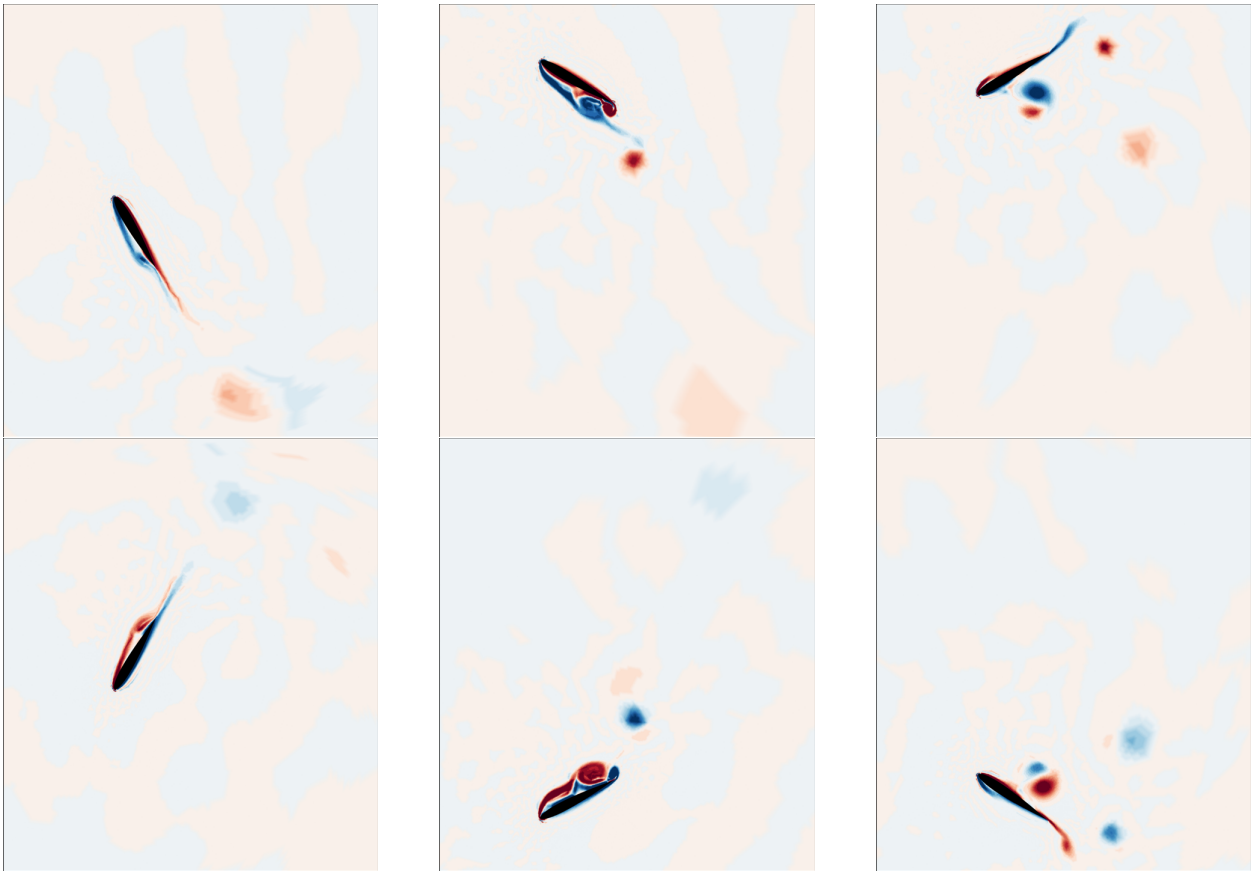

Figure 8: Trajectory of airfoil and flow vorticity at energetically optimal, $\bar{T}_x = 2.5$ flapping motion (see Figure 4 for flapping trajectory). Snapshots taken at times 6 equally spaced instances throughout the period ($T(\boldsymbol{\mu}) = 3.62$): $t = 0.0,\ 0.604,\ 1.21,\ 1.81,\ 2.41,\ 3.02$.

energy is seen despite these potential outliers.

Future work will focus on applying this method to three-dimensional wing geometries with a richer parametrization of the flapping trajectory, i.e., periodic splines or multiple harmonic terms instead of a single harmonic term, as well as shape morphing parameters [9] to gain additional insight into flapping physics.

Acknowledgments

This work was supported in part by the Luis Alvarez Postdoctoral Fellowship by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 (MZ), by the Director, Office of Science, Computational and Technology Research, U.S. Department of Energy under contract number DE-AC02-05CH11231 (PP), and by the AFOSR Computational Mathematics program under grant FA9550-15-1-0010 (PP). The content of this publication does not necessarily reflect the position or policy of any of these supporters, and no official endorsement should be inferred.

References

[1]J. M. McMichael and M. S. Francis, "Micro air vehicles - towards a new dimension in flight," tech. rep., DARPA, August 7 1997.

[2]W. Shyy, M. Berg, and D. Ljungqvist, "Flapping and flexible wings for biological and micro air vehicles," *Progress in aerospace sciences*, vol. 35, no. 5, pp. 455–505, 1999.

[3]R. Ramamurti and W. Sandberg, "Simulation of flow about flapping airfoils using finite element incompressible flow solver," *AIAA Journal*, vol. 39, no. 2, pp. 253–260, 2001.

[4]M. F. Platzer, K. D. Jones, J. Young, and J. S. Lai, "Flapping wing aerodynamics: progress and challenges," *AIAA Journal*, vol. 46, no. 9, pp. 2136–2149, 2008.

[5]I. H. Tuncer and M. Kaya, "Optimization of flapping airfoils for maximum thrust and propulsive efficiency," *AIAA Journal*, vol. 43, no. 11, pp. 2329–2336, 2005.

[6]M. Kaya and I. H. Tuncer, "Nonsinusoidal path optimization of a flapping airfoil," *AIAA journal*, vol. 45, pp. 2075–2082, August 2007.

[7]C. Chabalko, R. D. Snyder, P. S. Beran, and G. Parker, "The physics of an optimized flapping wing micro air vehicle," in *Proceedings of the 47th AIAA Aerospace Science Meeting Including The New Horizons Forum and Aerospace Exposition, AIAA Paper*, no. 2009-801, (Orlando, Florida), January 5 – 8 2009.

[8]M. P. van Schrojenstein Lantman and K. Fidkowski, "Adjoint-based optimization of flapping kinematics in viscous flows," in *21st AIAA Computaional Fluid Dynamics Conference*, 2013.

[9]M. J. Zahr and P.-O. Persson, "An adjoint method for a high-order discretization of deforming domain conservation laws for optimization of flow problems," *Journal of Computational Physics*, In review, 2016.

[10]C. A. Mader, J. RA Martins, J. J. Alonso, and E. V. Der Weide, "ADjoint: An approach for the rapid development of discrete adjoint solvers," *AIAA Journal*, vol. 46, no. 4, pp. 863–873, 2008.

[11]N. Yamaleev, B. Diskin, and E. Nielsen, "Adjoint-based methodology for time-dependent optimization," in *12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, American Institute of Aeronautics and Astronautics, 2008.

[12]M. Culbreth, Y. Allaneau, and A. Jameson, "High-fidelity optimization of flapping airfoils and wings," in *29th AIAA Applied Aerodynamics Conference*, (Honolulu, Hawaii), June 27 – 30 2011.

[13]M. Jones and N. K. Yamaleev, "Adjoint based shape and kinematics optimization of flapping wing propulsive efficiency," 43rd AIAA Fluid Dynamics Conference. San Diego, CA, 2013. AIAA 2013-2472, 2013.

[14]M. J. Zahr and P.-O. Persson, "High-order, time-dependent aerodynamic optimization using a discontinuous Galerkin discretization of the Navier-Stokes equations," in *AIAA Science and Technology Forum and Exposition*, (San Diego, CA), 2016.

[15]C.-K. Lin, "On the incompressible limit of the compressible Navier-Stokes equations," *Communications in partial differential equations*, vol. 20, no. 3-4, pp. 677–707, 1995.

[16]B. Desjardins, E. Grenier, P.-L. Lions, and N. Masmoudi, "Incompressible limit for solutions of the isentropic navier–stokes equationswith dirichlet boundary conditions," *Journal de Mathématiques Pures et Appliquées*, vol. 78, no. 5, pp. 461–471, 1999.

[17]B. M. Froehle, *High-order discontinuous Galerkin fluid-structure interaction methods*. University of California, Berkeley, 2013.

[18]P.-O. Persson, J. Bonet, and J. Peraire, "Discontinuous Galerkin solution of the Navier–Stokes equations on deformable domains," *Computer Methods in Applied Mechanics and Engineering*, vol. 198, no. 17, pp. 1585–1595, 2009.

[19]C. Farhat, P. Geuzaine, and C. Grandmont, "The discrete geometric conservation law and the nonlinear stability of ale schemes for the solution of flow problems on moving grids," *Journal of Computational Physics*, vol. 174, no. 2, pp. 669–694, 2001.

[20]D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, "Unified analysis of discontinuous Galerkin methods for elliptic problems," *SIAM Journal on Numerical Analysis*, vol. 39, no. 5, pp. 1749–1779, 2002.

[21]B. Cockburn and C.-W. Shu, "Runge-Kutta discontinuous Galerkin methods for convection-dominated problems," *J. Sci. Comput.*, vol. 16, no. 3, pp. 173–261, 2001.

[22]M. J. Zahr, P.-O. Persson, and J. Wilkening, "A fully discrete adjoint method for optimization of flow problems on deforming domains with time-periodicity constraints," *Computers & Fluids*, vol. Special Issue on USNCCM13 International Symposium on Spectral and High-Order Methods, 2016.

[23]M. J. Zahr and P.-O. Persson, "Energetically optimal flapping wing motions via adjoint-based optimization and high-order discretizations," in *Frontiers in PDE-Constrained Optimization*, Springer, 2017.

[24]R. B. George, M. B. Colton, C. A. Mattson, and S. L. Thomson, "A differentially driven flapping wing mechanism for force analysis and trajectory optimization," *International Journal of Micro Air Vehicles*, vol. 4, no. 1, pp. 31–49, 2012.

[25]P. L. Roe, "Approximate Riemann solvers, parameter vectors, and difference schemes," *Journal of Computational Physics*, vol. 43, no. 2, pp. 357–372, 1981.

[26]J. Peraire and P.-O. Persson, "The Compact Discontinuous Galerkin (CDG) method for elliptic problems," *SIAM Journal on Scientific Computing*, vol. 30, no. 4, pp. 1806–1824, 2008.

[27]R. Alexander, "Diagonally implicit Runge-Kutta methods for stiff ODEs," *SIAM J. Numer. Anal.*, vol. 14, no. 6, pp. 1006–1021, 1977.

[28]A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.