Mapping of coherent structures in parameterized flows by learning optimal transportation with Gaussian models

Angelo Iollo and Tommaso Taddei

September 21, 2021

Abstract

We present a general (i.e., independent of the underlying model) interpolation technique based on optimal transportation of Gaussian models for parametric advection-dominated problems. The approach relies on a scalar testing function to identify the coherent structure we wish to track; a maximum likelihood estimator to identify a Gaussian model of the coherent structure; and a nonlinear interpolation strategy that relies on optimal transportation maps between Gaussian distributions. We show that well-known self-similar solutions can be recast in the frame of optimal transportation by appropriate rescaling; we further present several numerical examples to motivate our proposal and to assess strengths and limitations; finally, we discuss an extension to deal with more complex problems.

1 Introduction

In science and engineering, it is important to identify low-rank approximations valid over a range of configurations (corresponding to different physical properties, different geometries or operational configurations). Low-rank approximations are of paramount importance in parameterized model order reduction (pMOR, [7, 22, 14, 36, 45]) to speed up model evaluations in the limit of many queries, but also in optimization and uncertainty quantification to efficiently generate samples from the solution manifold. In this paper, given two snapshots of the solution manifold $U_0, U_1 : \Omega \subset \mathbb{R}^n \to \mathbb{R}^d$, we wish to determine an interpolation $\widehat{U} : [0,1] \times \Omega \to \mathbb{R}^d$ such that $\widehat{U}(0,\cdot) = U_0$ and $\widehat{U}(1,\cdot) = U_1$: this task is one of the key enablers towards the implementation of approximation strategies for parameterized systems. Our emphasis is on the development of a general (i.e., independent of the underlying parametric model), interpretable methodology that allows simple (i.e., non-intrusive) integration with high-fidelity codes and that is robust also for small datasets.

The vast majority of data compression methods aims to determine linear low-rank approximations. If we denote by $U(x,\mu)$ the solution field, where $x = (x_1, \ldots, x_n)$ denotes the spatial variable and $\mu = (\mu_1, \ldots, \mu_p)$ denotes the vector of parameters, linear approaches consider approximations of the form

$$\widehat{U}(x,\mu) = \sum_{i=1}^{r} \widehat{\alpha}_i(\mu) \zeta_i(x).$$

Here, $\widehat{\alpha}_1, \ldots, \widehat{\alpha}_r$ are parameter-dependent coefficients that can be obtained by solving a reduced-order model (ROM), while $\zeta_1, \ldots, \zeta_r$ are a reduced-order basis (ROB) that is computed by exploring the parameter domain. Linear models can be interpreted as a generalization of convex interpolations of two snapshots U_0, U_1, that is

$$\widehat{U}^{\rm co}(s,x) \,=\, (1-s)U_0(x) \,+\, sU_1(x) \quad s \in [0,1], x \in \mathbb{R}^n.$$

The use of linear methods relies on the assumption that the problem of interest exhibits linear coherent structures.

There exists a broad class of problems for which linear methods are effective. To provide concrete examples, the presence of coherent structures in turbulent flows provides physical foundations for the use of linear methods in numerous applications in flow control and design [3]; evanescence of high-frequency modes for diffusion-dominated problems [29] is at the foundation of component-mode synthesis [9] and more recently of component-based MOR strategies [15]. Despite the successes of linear methods, there exists a broad class of problems of interest in engineering for which linear methods are highly inaccurate: this motivates the development of nonlinear methods.

We present a general approach that relies on optimal transportation [44] to perform accurate nonlinear interpolations between solution snapshots. First, we rely on a scalar testing function (cf. section 4.1.2) to derive a Gaussian model $g[U]$ of the solution field; then, we rely on well-known results for optimal transportation of

Gaussian distributions to determine the optimal transport mapping T_g from $g[U_0]$ and $g[U_1]$ and the optimal transport mapping R_g from $g[U_1]$ and $g[U_0]$; finally, we define the nonlinear interpolation

$$\widehat{U}(s,x) \;=\; (1-s)U_0 \circ W_g(s,x) \;+\; sU_1 \circ T_g(1-s,x), \quad s \in [0,1], x \in \mathbb{R}^n,$$

where $W_g(s,x) = (1-s)x + sR_g(x)$ and $T_g(s,x) = (1-s)x + sT_g(x)$. We refer to $\widehat{U}$ as *convex displacement interpolation* due to the analogy with displacement interpolation (cf. section 2). We present several numerical examples to motivate our proposal; furthermore, we show in section 3 that well-known self-similar solutions can be recast in the frame of optimal transportation by appropriate rescaling.

The use of optimal transportation theory to devise nonlinear interpolation has been considered in several works in the pMOR literature, [4, 10, 16]. Here, we apply optimal transportation to a model of the solution field: as a result, the parametric field of interest U does not have to be neither scalar nor positive and does not have to fulfill conservation of mass over the parameter domain.

The proposed approach shares relevant features with Lagrangian or registration-based approaches to pMOR ([27, 31, 40]) and also with the works on Lagrangian coherent structures (LCS, [13, 32]) in the field of nonlinear dynamics. In particular, the feature-based Gaussian model of the field is similar in scope to the registration sensor introduced in [42] and also to shock-capturing sensors used for high-order schemes ([30, 33]). On the other hand, we remark that, while our emphasis is on the development of predictive models for parametric systems, LCS literature mainly focuses on the physical understanding — and subsequently the control — of chaotic systems.

As extensively discussed in the example of section 4.2.3, our approach might suffer from (i) the presence of boundaries, and (ii) the presence of multiple coherent structures that we wish to track. In section 5, we discuss how to extend the approach to deal with more complex problems: the key ideas are to replace the Gaussian model with a mixture of Gaussian models and to replace the affine-in-s maps W_g, T_g with suitable nonlinear maps.

The outline of the paper is as follows. In section 2, we present a short introduction to optimal transportation: we provide a number of references and we introduce relevant notation. In section 3, we illustrate the connection between self-similarity and optimal transportation through the vehicle of several examples. In section 4, we present our method in its elementary form and we provide numerical investigations. In section 5, we discuss the extension of the elementary approach to deal with more complex problems. Section 6 concludes the paper by offering a short summary and several perspectives.

2 Optimal transportation

As discussed in the introduction, classical model reduction techniques approximate the solution to a physical model by affine combinations of empirical modes. However, the nature of the parametric dependence of the PDE can be such that the main features of the solution are displaced (or advected) in space and eventually time. For example, coherent structures like advected vortices in a wake or displacing shock structures in supersonic flows are not amenable to linear compression methods. In order to identify and model the displacement of coherent structures with respect to parameter variation, we employ optimal transportation theory [44], as it was already done with different approaches in [16] and [4].

Let π and ν be two continuous non-negative real-valued functions defined on $\mathbb{R}^n$ with finite second-order moments such that

$$\int_{\mathbb{R}^n} \pi(\xi) d\xi = \int_{\mathbb{R}^n} \nu(x) dx = c;$$

without loss of generality, we assume $c = 1$ and we refer to π and ν as probability density functions (pdfs). Consider a one-to-one vector-valued function $X(\xi)$: $\mathbb{R}^n \to \mathbb{R}^n$; ensuring local mass conservation, we find the constraint:

$$\pi(\xi) d\xi = \nu(x) dx \tag{1}$$

with $x = X(\xi)$ and $dx = \det(\nabla_\xi X) d\xi$. Then, there exists a unique convex potential $\Psi(\xi)$: $\mathbb{R}^n \to \mathbb{R}$, such that the mapping $X(\xi) = \nabla_\xi \Psi$ minimizes:

$$I(\pi,\nu) = \int_{\mathbb{R}^n} \left(X(\xi) - \xi\right)^2 \pi(\xi)\, d\xi \tag{2a}$$

subject to the mass conservation constraint:

$$\pi(\xi) = \nu(\nabla_\xi \Psi) \det\left(\nabla^2_\xi \Psi\right). \tag{2b}$$

This result follows from a duality principle in convex optimization introduced in [17] that is linked to the polar factorization and monotone rearrangement of vector-valued functions [6]: the full proof is based on the convexity of the potential minimizing eq. (2a) and the existence and uniqueness of monotone measure-preserving maps [24]. Let J be the minimum of $I(\pi, \nu)$ subject to eq. (2b). Existence and uniqueness of Ψ implies that

$$W_2(\pi, \nu) = \sqrt{J}$$

is a distance function between probability measures; W_2 is known as the Wasserstein metric.

The Wasserstein distance is a rigorous proxy of the notion of displacement; it can thus be exploited to introduce a notion of similarity among snapshots that is well-adapted to transport phenomena. This feature of the Wasserstein distance has motivated its application to various supervised and unsupervised learning tasks including clustering (e.g., [38]) and interpolation. In this work, we exclusively focus on interpolation.

We introduce the displacement (or geodesic) interpolant $\widehat{\rho}_s^+$ between π and ν such that

$$\widehat{\rho}_s^+(\xi) = \nu(T(s,\xi)) \det\left(\nabla_\xi T(s,\xi)\right), \tag{3a}$$

ρ_s^+ is also known as McCann interpolation; here, the mapping $T(s,\xi) : [0,1] \times \mathbb{R}^n \to \mathbb{R}^n$ is defined as

$$T(s,\xi) = (1-s)\,\xi + s\,\nabla_\xi \Psi(\xi). \tag{3b}$$

Note that $\widehat{\rho}_s^+ = \nu$ for $s = 0$ and $\widehat{\rho}_s^+ = \pi$ for $s = 1$. In view of the discussion below, we further introduce the inverse map $R : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$ such that

$$R(s,\cdot) = T^{-1}(s,\cdot), \quad \forall\, s \in [0,1], \tag{4a}$$

and the reverse McCann interpolation, which is obtained by inverting the role of π and ν in (3a):

$$\widehat{\rho_s^-}(x) = \pi(R(s,x)) \det\left(\nabla_x R(s,x)\right). \tag{4b}$$

3 Motivating examples

Self-similarity plays a fundamental role in physics and we show in the next examples that well-known self-similar solutions to PDEs can be recast in the frame of optimal transportation by appropriate rescaling. In the remainder, we repeatedly use the Brenier's theorem ([6]), which states that given two densities π and ν there exists a unique function X that is the gradient of a convex function and transports π onto ν.

3.1 The heat kernel

3.1.1 Convex potential

Let us consider the heat kernel $K : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}_+$, $(t,x) \mapsto K(t,x)$,

$$K(t,x) = \frac{1}{(4\pi t)^{n/2}}\, e^{-\frac{\|x\|_2^2}{4t}} \tag{5}$$

that satisfies the initial value problem:

$$\begin{cases} \dfrac{\partial K}{\partial t} = \Delta K & \text{in } \mathbb{R}^+ \times \mathbb{R}^n, \\ K(0,x) = \delta_0(x) & \text{in } \mathbb{R}^n, \end{cases} \tag{6}$$

where δ_0 is the Dirac mass concentrated in $x = 0$. Of course, we have for all $t_0, t_1 \in \mathbb{R}^+$:

$$\int_{\mathbb{R}^n} K(t_0,\xi) d\xi = \int_{\mathbb{R}^n} K(t_1,x) dx = 1.$$

Define now the convex potential $\Psi : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}_+$,

$$\Psi(t_0,t_1,\xi) = \frac{1}{2}\sqrt{\frac{t_1}{t_0}}\, \|\xi\|_2^2.$$

Note that $\nabla_\xi \Psi = \sqrt{\dfrac{t_1}{t_0}}\xi$ and $\det(\nabla_\xi^2 \Psi) = \left(\dfrac{t_1}{t_0}\right)^{n/2}$. It is thus easy to verify that

$$K(t_0,\cdot) = K(t_1, \nabla_\xi \Psi) \det\left(\nabla_\xi^2 \Psi\right) \quad \forall\, t_0, t_1 \in \mathbb{R}_+,$$

which corresponds to (2b) for $\pi = K(t_0,\cdot)$ and $\nu = K(t_1,\cdot)$. Since Ψ is a convex function that satisfies (2b), $X = \nabla_\xi \Psi$ must be the unique solution to (2).

3.1.2 Displacement interpolation

We introduce the forward mapping:

$$T(s,\xi) = (1-s)\,\xi + s\sqrt{\frac{t_1}{t_0}}\,\xi = \left(\frac{\sqrt{t_0}+(\sqrt{t_1}-\sqrt{t_0})s}{\sqrt{t_0}}\right)\xi,$$

and the corresponding backward map

$$R(s,x) = \left(\frac{\sqrt{t_0}+(\sqrt{t_1}-\sqrt{t_0})s}{\sqrt{t_0}}\right)^{-1} x.$$

Exploiting (4), we find the displacement interpolation between $K(t_0,\cdot)$ and $K(t_1,\cdot)$:

$$\widehat{K}_s = K\left(t_0, R(s,\cdot)\right)\,\left(\frac{\sqrt{t_0}+(\sqrt{t_1}-\sqrt{t_0})s}{\sqrt{t_0}}\right)^{-n}.$$

If we introduce the parameter re-scaling $s : [t_0,t_1] \to [0,1]$ such that

$$s(t) = \frac{\sqrt{t}-\sqrt{t_0}}{\sqrt{t_1}-\sqrt{t_0}},$$

we obtain $\dfrac{\sqrt{t_0}+(\sqrt{t_1}-\sqrt{t_0})s(t)}{\sqrt{t_0}} = \sqrt{\dfrac{t}{t_0}}$ and thus $\widehat{K}_{s(t)}$ satisfies:

$$\widehat{K}_{s(t)}(x) = K\left(t_0, \sqrt{\frac{t_0}{t}}x\right)\left(\sqrt{\frac{t_0}{t}}\right)^{n} = K(t,x).$$

Note that the displacement interpolation with rescaling $\widehat{K}_{s(t)}$ is the exact solution to the heat equation for all $t \in [t_0,t_1]$. In model-order reduction vocabulary, we have that, given two snapshots of the solution, their displacement interpolation is an exact solution to the PDE, provided that the parameter is appropriately re-scaled.

We further observe that the mapped heat kernel $\widetilde{K} : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ such that $\widetilde{K}(t,x) = K\left(t, R(s(t),x)\right)$, satisfies

$$\widetilde{K}(t,x) = \frac{1}{(4\pi t)^{n/2}} e^{-\frac{\|x\|_2^2}{4t_1}}.$$

Note that time dependency enters only through a multiplicative constant: as a result, the manifold $\{\widetilde{K}(t,\cdot) : t \in \mathbb{R}_+\}$ is contained in a linear space of dimension one. From an approximation standpoint, application of optimal transportation has the effect of linearizing the manifold of interest. More formally, the Kolmogorov N-width ([35]) associated with the mapped solution manifold is equal to zero for all $N \geq 1$, while the Kolmogorov N-width associated with the original manifold is strictly positive for all $N \geq 0$.

3.2 Nonlinear diffusion

3.2.1 Convex potential

A suitable model for diffusion of heat in hot plasma, very intense thermal waves or diffusion in porous media [2] is the following nonlinear diffusion equation:

$$\frac{\partial \theta}{\partial t} = \Delta \theta^m \tag{7}$$

where $\theta : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}_+$, $(t,x) \mapsto \theta(t,x)$, and $m \in \mathbb{N}$ with $m > 1$. For example, for an instantaneous release of heat at time $t = 0$ and concentrated at the origin, this equation admits the so called ZKB [2] solution profile:

$$B(t,x) = t^{-\alpha}\left(\left(C - k\|x\|_2^2 t^{-2\beta}\right)^+\right)^{1/(m-1)} \tag{8}$$

where $\alpha = \dfrac{n}{n(m-1)+2}$, $\beta = \dfrac{\alpha}{n}$, $k = \dfrac{(m-1)\alpha}{2mn}$, C is a positive constant and $z^+ = \max(z,0)$. The L^1 norm of the ZKB profile, $\displaystyle\int_{\mathbb{R}^n} B(t,x)\,dx$, is time invariant and equal to the “initial heat” released.

We introduce the convex potential $\Psi : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}_+$,

$$\Psi(t_0, t_1, \xi) = \frac{1}{2} \left(\frac{t_1}{t_0}\right)^{\beta} \|\xi\|_2^2 \,,$$

Note that $\nabla_\xi \Psi = (t_1/t_0)^\beta \, \xi$ and $\det(\nabla_\xi^2 \Psi) = (t_1/t_0)^\alpha$: we thus obtain

$$B(t_0, \xi) = B\left(t_1, \xi \left(\frac{t_1}{t_0}\right)^{\beta}\right) \left(\frac{t_1}{t_0}\right)^{\alpha},$$

which corresponds to (2b) for $\pi(\xi) = B(t_0, \xi)$ and $\nu(x) = B(t_1, x)$. As in the previous case, since Ψ is convex and satisfies (2b), we must have that $\nabla_\xi \Psi$ is the unique optimal forward mass transportation between the solutions at t_0 and t_1.

3.2.2 Displacement interpolation

As in the heat kernel case, we use the optimal mapping to define the displacement interpolation. Define the forward interpolation mapping:

$$T(s, \xi) = (1-s)\,\xi + s\left(\frac{t_1}{t_0}\right)^{\beta} \xi = \left(1 + \left(\left(\frac{t_1}{t_0}\right)^{\beta} - 1\right) s\right) \xi$$

and the corresponding backward mapping:

$$R(s, x) = \frac{x}{1 + \left(\left(\frac{t_1}{t_0}\right)^{\beta} - 1\right) s}$$

where $s \in [0,1]$ and $x, \xi \in \mathbb{R}^n$. Exploiting (4b), we find that the displacement interpolation of the ZKB profile between t_0 and t_1 is:

$$\widehat{B}_s(x) = B\left(t_0, R(s,x)\right) \left(1 + \left(\left(\frac{t_1}{t_0}\right)^{\beta} - 1\right) s\right)^{-n}. \tag{9}$$

If we introduce the rescaling

$$s(t) = \frac{t^\beta - t_0^\beta}{t_1^\beta - t_0^\beta},$$

we obtain $1 + \left(\left(\frac{t_1}{t_0}\right)^{\beta} - 1\right) s(t) = \left(\frac{t}{t_0}\right)^{\beta} = \left(\frac{t}{t_0}\right)^{\alpha/n}$ and $R(t,x) = \left(\frac{t}{t_0}\right)^{\alpha/n} x$; substituting in (9), we obtain

$$\widehat{B}_{s(t)}(x) = B\left(t_0, \left(\frac{t}{t_0}\right)^{\beta} x\right) \left(\frac{t}{t_0}\right)^{-\alpha} = B(t, x).$$

Again, in model-order reduction vocabulary, we have that given two snapshots of the solution for different parameter values, their displacement interpolation is an exact solution to the PDE, provided that the parameter is appropriately rescaled.

3.3 Conservation laws

The Euler equations for an inviscid compressible ideal fluid flow in one dimension are given by:

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} = 0$$

where $U = (\rho, \rho u, E)$ is the vector of conserved variables, $F = \big(\rho u, \rho u^2 + p, (E+p)u\big)$ is the flux, ρ is the density, p is the pressure, E is the total internal energy per unit volume and u is the velocity. We denote by $\gamma > 0$ the ratio of specific heats and we denote by $a = \sqrt{\frac{\gamma p}{\rho}}$ the speed of sound.

3.3.1 Displacement interpolation of simple wave solutions

We assume that the flow is isentropic and does not contain shock waves for all $t \in (0, t^\star) = A$: under this assumption, the equation of state reduces to $p = C\rho^\gamma$; the Cauchy problem is well-posed in $A \times \mathbb{R}$ and all physical quantities can be traced back to the value at time $t = 0$. It is possible to show that the Euler system admits the two Riemann invariants $R^\pm = \frac{2}{\gamma - 1} a \pm u$ that satisfy the equation:

$$\frac{\partial R^\pm}{\partial t} + (u \pm a) \frac{\partial R^\pm}{\partial x} = 0 \quad \text{in } A \times \mathbb{R}. \tag{10}$$

In this section, we consider flows for which one of the two invariants, say R^-, is constant and equal to $c \in \mathbb{R}$ at time $t = 0$; the corresponding solution to the Euler equations is known as simple wave ([19]).

We define the characteristics $X^\pm : A \times \mathbb{R} \to \mathbb{R}$ such that

$$\begin{cases} \frac{dX^\pm}{dt}(t, \xi) = u(t, X^\pm(t, \xi)) \pm a(t, X^\pm(t, \xi)) & t \in A \\ X^\pm(0, \xi) = \xi \end{cases} \tag{11}$$

Combining (10) with (11), we find that $\frac{d}{dt} R^\pm(t, X^\pm(t, \xi)) = 0$, which implies that R^+ (resp. R^-) is constant on the characteristic X^+ (resp. X^-). Since R^- is constant at $t = 0$, we must have that $R^-(t, x) = c$ for all $(t, x) \in A \times \mathbb{R}$. This implies that $u(t, x) = c + \frac{2}{\gamma - 1} a(t, x)$ and then

$$\begin{aligned} & a(t, X^+(t, \xi)) = R^+(t, X^+(t, \xi)) - u(t, X^+(t, \xi)) = R_0^+(\xi) - \frac{2}{\gamma - 1} a(t, X^+(t, \xi)) \\ & \Rightarrow a(t, X^+(t, \xi)) = (\gamma - 1) \frac{R_0^+(\xi) - c}{\gamma + 1}, \end{aligned}$$

and in particular a (and thus all state variables) are constant on the X^+ characteristic. Furthermore, since $u + a$ in constant on X^+, we must have that $\frac{dX^+}{dt}$ is constant and thus the X^+ characteristics are straight lines (cf. Figure 1) and satisfy $X^+(t, \xi) = \xi + (u_0(\xi) + a_0(\xi))\, t$. Due to the assumption on the smoothness of the flow, characteristics do not coalesce for $t \in A$: as a result, we find that $X^+(t, \cdot)$ is bijective in $\mathbb{R}$ for all $t \in A$.

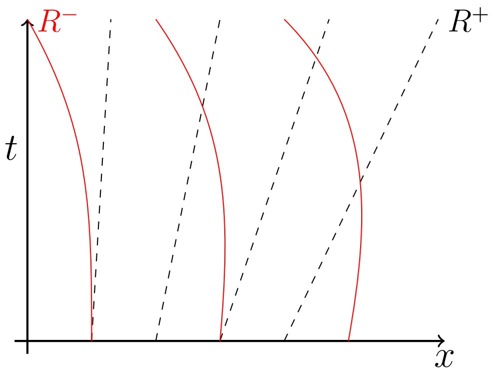


Figure 1: Simple wave solution: the right-going characteristics are straight lines. $X^+(t, \xi) = \xi + (u_0(\xi) + a_0(\xi))\, t$ where u_0 and a_0 are the initial condition for the state variables.

We denote by Y^+ the inverse map of X^+ such that $X^+(t, Y^+(t, x)) = x$ for all $(t, x) \in A \times \mathbb{R}$ (Y^+ is the mapping that associates (t, x) to the foot of the corresponding X^+ characteristic); we have that the pressure p (and any other state variable) satisfies

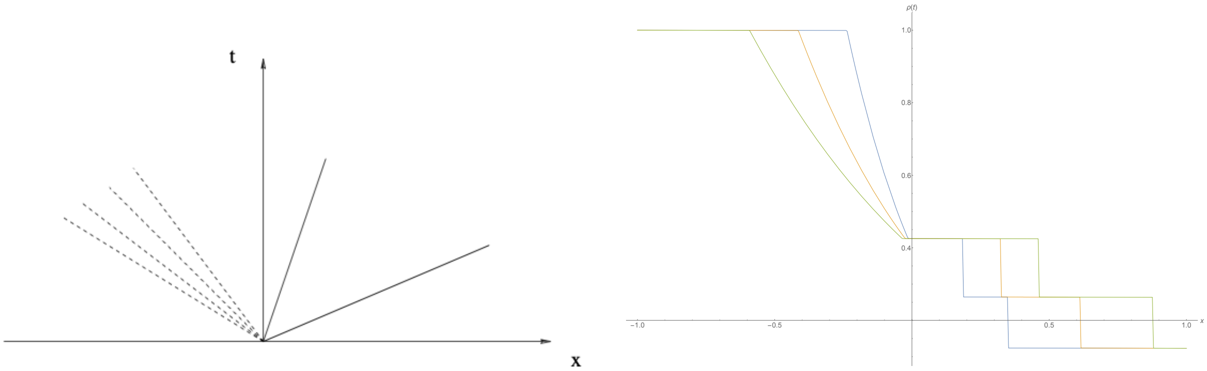
$$p(t, x) = p_0(Y^+(t, x)) = p_0(\xi)$$

and thus

$$\frac{\partial p}{\partial x} dx = \frac{\partial p_0}{\partial \xi} \frac{\partial Y^+}{\partial x} dx = \frac{\partial p_0}{\partial \xi} d\xi,$$

and also

$$\frac{\partial p_0}{\partial \xi} d\xi = \frac{\partial p}{\partial x} \frac{\partial X^+}{\partial \xi} d\xi = \frac{\partial p}{\partial x} dx. \tag{12}$$



(a) (b)

Figure 2: Sod problem. (a) shock tube solution pattern: a receding expansion fan and two forward waves corresponding to a contact discontinuity and a shock. The solution is self-similar with respect to the variable $\eta = x/t$. (b) shock tube solution: density at $t = t_0$ (blue), density at $t = t_1$ (green) and density displacement interpolant at $t = (t_0 + t_1)/2$ (yellow). The displacement interpolant coincides with the the exact solution.

We define the scalar field $\Psi : A \times \mathbb{R} \to \mathbb{R}$ such that

$$\Psi(t, \xi) \;=\; \int_0^{\xi} X^+(t, \xi')\, d\xi'.$$

Since $X^+(t, \cdot)$ is bijective in $\mathbb{R}$ for all $t \in A$, we must have $\dfrac{\partial X^+}{\partial \xi} > 0$: as a result, Ψ is convex in the second argument; furthermore, we obtain that Ψ satisfies (2b) with $\pi = \dfrac{\partial p_0}{\partial \xi}$ and $\nu = \dfrac{\partial p(t, \cdot)}{\partial x}$, provided that $\dfrac{\partial p_0}{\partial \xi} \in L^1(\mathbb{R})$ and $\dfrac{\partial p_0}{\partial \xi} > 0$[1]. In conclusion, for any $t \in A$, $X^+(t, \cdot)$ is the unique optimal forward mass transportation between $\dfrac{\partial p_0}{\partial \xi}$ and $\dfrac{\partial p(t, \cdot)}{\partial x}$.

We observe that in this example optimal transportation can only be rigorously applied to a derived quantity — the pressure gradient $\dfrac{\partial p}{\partial x}$ — as opposed to the solution itself. Nevertheless, we find that the mapped solution field satisfies

$$\widetilde{p}(t, \cdot) := p(t, \cdot) \circ X^+(t, \cdot) = p_0,$$

and is thus independent of time — we recall that $X^+(t, \xi) = \xi + (u_0(\xi) + a_0(\xi))t$. Application of optimal transportation thus detects the appropriate self-similarity transformation associated with the problem.

3.3.2 Displacement interpolation for Riemann problems

The Riemann problem for conservation laws is a Cauchy problem with piece-wise constant initial data where a single discontinuity is placed at $x = 0$ in the domain of interest. The problem is essentially one-dimensional and the solution is self-similar with respect to the self-similarity variable $\eta = x/t$ (see, e.g., [43, Chapter 3]). To fix the ideas, we consider the Sod shock tube problem, which corresponds to impose that the ideal gas is at rest ($u = 0$) with high pressure and density for $x < 0$ and low pressure and density for $x > 0$. There are three waves emerging from the initial discontinuity: a receding expansion fan and two forward waves corresponding to a contact discontinuity and a shock, see Figures 2(a) and 2(b).

Density is a monotonically-decreasing function in x and we have that $\rho(t, x) = \hat{\rho}(x/t) = \hat{\rho}(\eta)$, $\forall t > 0$. As before, we can write:

$$\frac{\partial \rho}{\partial x} dx = \frac{\partial \hat{\rho}}{\partial \eta} \frac{\partial \eta}{\partial x} dx = \frac{\partial \hat{\rho}}{\partial \eta} d\eta \tag{13}$$

where $d\eta = \dfrac{\partial \eta}{\partial x} dx$. This means that, up to a change of sign, the density spatial derivative satisfies optimal transportation between any two times t_0 and t_1, see eq. (1) with $\Psi(t_0, t_1, \xi) = \frac{1}{2} \xi^2 \frac{t_1}{t_0}$. If we consider a linear scaling $s(t) = \frac{t - t_0}{t_1 - t_0}$, we obtain $T(s(t), \xi) = \frac{t}{t_0} \xi$, which is the appropriate self-similarity transformation associated with the problem.

[1]The result can be trivially extended to the more general case by taking the absolute value of $\frac{\partial p_0}{\partial \xi}$.

4 Nonlinear interpolation based on Gaussian models

Although for other notable cases (such as steady boundary layers) the exact displacement interpolant coincides with the exact solution to the physical model, in general it is not possible to readily identify an extensive scalar physical quantity that is representative of the whole solution field and for which the optimal transport map is available in closed form.

In the last few decades, there has been a growing interest in determining effective algorithms to approximate the solution to optimal transportation problems for arbitrary choices of the densities π and ν in (2), [34]. In this work, we pursue a different approach: first, we identify a Gaussian model $g[U]$ of the solution U to the PDE; then, we exploit the knowledge in closed-form of the forward mapping T between two standard Gaussians to define the interpolation operator. To illustrate the many features of our approach and its limitations that motivate the extensions of section 5, we present extensive numerical investigations for one-dimensional and two-dimensional test problems with exact or numerical solutions

4.1 Methodology

4.1.1 Optimal transportation of multivariate normal density distributions

We briefly review the solution to (2) for multivariate Gaussian densities; we refer to [23] for the proofs. We define the normal density distribution ϕ with mean $\mu \in \mathbb{R}^n$ and symmetric positive definite covariance $\Sigma \in \mathbb{R}^{n \times n}$:

$$\phi(x;\, \mu, \Sigma) = \frac{1}{(2\pi)^{n/2}\, |\Sigma|^{1/2}}\; e^{-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)}. \tag{14}$$

Given the densities $\pi = \phi(\cdot;\, \mu_0, \Sigma_0)$ and $\nu = \phi(\cdot;\, \mu_1, \Sigma_1)$, we find that the displacement interpolant $\widehat{\phi}_s$ is Gaussian with mean and covariance given by

$$\mu_s = (1-s)\,\mu_0 + s\,\mu_1, \quad \Sigma_s = \Sigma_0^{-1/2}\left((1-s)\,\Sigma_0 + s\left(\Sigma_0^{1/2}\Sigma_1\Sigma_0^{1/2}\right)^{1/2}\right)^2 \Sigma_0^{-1/2}, \tag{15a}$$

for all $s \in [0,1]$. The forward mapping T is also available in closed form:

$$T(s,\xi) = (1-s)\,\xi + s\left(\mu_1 + \Sigma_0^{-1/2}\left(\Sigma_0^{1/2}\Sigma_1\Sigma_0^{1/2}\right)^{1/2}\Sigma_0^{-1/2}\,(\xi - \mu_0)\right). \tag{15b}$$

Finally, the Wasserstein distance between Gaussian density distributions is given by:

$$W_2\left(\phi\left(\mu_0,\Sigma_0\right), \phi\left(\mu_1,\Sigma_1\right)\right) = \sqrt{\|\mu_1 - \mu_0\|_2^2 \,+\, \mathrm{Tr}\left(\Sigma_0 + \Sigma_1 - 2\left(\Sigma_0^{1/2}\Sigma_1\Sigma_0^{1/2}\right)^{1/2}\right)}. \tag{15c}$$

As a final remark, we note that the optimal mapping between Gaussian distributions is always well-defined, affine and can be obtained at negligible computational cost.

4.1.2 Gaussian models of coherent structures

Given the field $U: \mathbb{R}^n \to \mathbb{R}^d$, we define the scalar testing function $\mathcal{T}(\cdot; U): \mathbb{R}^n \to \mathbb{R}$ and the set

$$\mathcal{C}_{\mathcal{T}}(U) := \{x \in \mathbb{R}^n \,:\, \mathcal{T}(x;U) > 0\}, \tag{16}$$

which identifies the coherent structure associated with the criterion $\mathcal{T}$. To provide a concrete example, if U is the velocity field, we might define $\mathcal{T}(x;U) = \|\nabla \times U(x)\|_2 - \tau$ with $\tau > 0$: in this case, $\mathcal{C}_{\mathcal{T}}(U)$ identifies the region of the domain where the enstrophy exceeds a user-defined threshold.

In order to fit a Gaussian model to $\mathcal{C}_{\mathcal{T}}(U)$ in (16), we define a finite-dimensional discretization of the domain of interest $P_{\rm hf} = \{x_i\}_{i=1}^{N_{\rm hf}}$ and we define

$$P_{\rm hf}^+ := \{x \in P_{\rm hf} \,:\, \mathcal{T}(x;U) > 0\} = \{y_j\}_{j=1}^{N_{\rm hf}^+} \tag{17}$$

Then, the statistical parametric model of the coherent structure is obtained by assuming that $\{y_j\}_j$ are independent identically distributed (iid) realizations of a multivariate Gaussian distribution, and then resorting to maximum likelihood estimation (MLE, see, e.g., [37, Chapter 8]) to estimate mean and variance:

$$g(x;U) := \phi\left(x;\mu_{\rm mle}[U], \Sigma_{\rm mle}[U]\right), \quad \text{where} \quad \begin{cases} \mu_{\rm mle}[U] = \dfrac{1}{N_{\rm hf}^+} \displaystyle\sum_{j=1}^{N_{\rm hf}^+} y_j, \\ \Sigma_{\rm mle}[U] = \dfrac{1}{N_{\rm hf}^+} \displaystyle\sum_{j=1}^{N_{\rm hf}^+} \left(y_j - \mu_{\rm mle}[U]\right)\left(y_j - \mu_{\rm mle}[U]\right)^T. \end{cases} \tag{18}$$

Some comments are in order.

- The scalar testing function $\mathcal{T}$ identifies flow features that we wish to track. From an approximation standpoint, it is natural to identify and then track high-gradient regions of the flow, which correspond to shock waves or contact discontinuities. In this respect, the testing function $\mathcal{T}$ is related in scope to shock-capturing sensors that are used in high-order methods to activate numerical dissipation where needed, and also to error indicators used for mesh adaptation and refinement. In the framework of model reduction, we observe that we might also interpret the Gaussian model $U \mapsto g(\cdot; U)$ as a registration sensor: similarly to [42], $g(\cdot; U)$ is indeed used to learn a suitable parametric mapping T that is ultimately used to approximate the parametric field of interest.

- The use of Gaussian distributions allows to readily define the optimal mapping T and ultimately the displacement interpolant, possibly at the price of inaccurate representations of the coherent structure of interest. The choice of the distribution model should be a compromise between *learnability* and *expressivity*: here, learnability can be measured in terms of the degree of difficulty of solving the subsequent optimal transportation problem, while expressivity is related to the difference in performance between the displacement interpolant based on Gaussian models and the displacement interpolant obtained by transporting the indicator function of $\mathcal{C}_{\mathcal{T}}(U)$. Note also that, for practical high-fidelity data, estimates of $\mathcal{C}_{\mathcal{T}}(U)$ might be noisy: our Gaussian model might thus also filter raw data and ultimately prevent over-fitting.

4.1.3 Convex displacement interpolation

Given the parametric field $U : [0,1] \times \mathbb{R}^n \to \mathbb{R}^d$, we consider the problem of constructing (nonlinear) interpolations between $U_0 = U(0,\cdot)$ and $U_1 = U(1,\cdot)$. Towards this end, we use the procedure in section 4.1.2 to generate the Gaussian models g_0, g_1 and we use (15b) to compute the forward map T_g and its inverse $R_g = T_g^{-1}$ — the latter is simply obtained by interchanging U_0 with U_1. Then, we define the *convex displacement interpolation* $\widehat{U} : [0,1] \times \mathbb{R}^n \to \mathbb{R}^d$ such that

$$\widehat{U}(s,x) \;=\; (1-s)U_0 \circ W_g(s,x) \;+\; sU_1 \circ T_g(1-s,x), \quad s \in [0,1], x \in \mathbb{R}^n, \tag{19}$$

where $W_g(s,x) = (1-s)x + sR_g(1,x)$.

Note that $\widehat{U}(0,\cdot) = U_0$ and $\widehat{U}(1,\cdot) = U_1$ (*interpolation*); furthermore, if we interchange U_0 with U_1, we find an interpolant $\widehat{U}'$ that satisfies $\widehat{U}'(s,\cdot) = \widehat{U}(1-s,\cdot)$ for all $s \in [0,1]$ (*symmetry*). Since optimal transportation is applied to the Gaussian models, our displacement interpolation procedure is well-defined regardless of the sign of U_0 and U_1 and can also deal with vector-valued fields. We further observe that our strategy does not depend on the particular way in which the map T_g is computed and can thus be coupled with other registration techniques.

If multiple snapshots of U are available for $0 = t_0 < \ldots < t_K = 1$, we can improve the accuracy of (19) by considering piecewise approximations in the intervals $A_k := (t_k, t_{k+1})$, for $k = 0, \ldots, K-1$, or by learning a more accurate rescaling function $s : [0,1] \to [0,1]$. Re the latter, we might (i) compute $s_1, \ldots, s_K$ such that

$$s_k \in \arg\min_{s\in[0,1]} \|\widehat{U}(s,\cdot) - U(t_k,\cdot)\|_\star, \quad \text{such that } 0 = s_0 < s_1 \ldots < s_K = 1, \tag{20a}$$

where $\|\cdot\|_\star$ is a functional norm of interest; (ii) compute a bijective rescaling $\widehat{s} : [0,1] \to [0,1]$ based on the dataset $\{(t_k, s_k)\}_{k=0}^K$ using a standard regression algorithm; and (iii) define the interpolant:

$$\widehat{U}'(t,x) \;=\; \widehat{U}(\widehat{s}(t),x), \quad t \in [0,1],\; x \in \mathbb{R}^n. \tag{20b}$$

In the numerical examples, we show that optimizing the rescaling function s might have a significant impact on performance; furthermore, it might unveil relevant features of the coherent structure of interest.

Remark 4.1. ***Lagrangian interpolation.*** *Convex displacement interpolation can be related to the following nonlinear interpolation:*

$$\widehat{U}(s,x) \;=\; ((1-s)U_0 \;+\; sU_1 \circ T_g(1,\cdot)) \circ R_g(s,x), \quad s \in [0,1], x \in \mathbb{R}^n. \tag{21}$$

The latter performs linear interpolation of the mapped field $\{\widetilde{U}(t,\cdot) = U(t,\cdot) \circ T_g(t,\cdot) \,:\, t \in [0,1]\}$ *and can thus be referred to as Lagrangian interpolation: it is indeed consistent with Lagrangian (or registration-based) approaches presented in the MOR literature, (e.g. [16, 40, 42]). . Similarly to* (19)*, the nonlinear interpolation* (21) *satisfies* $\widehat{U}(0,\cdot) = U_0$ *and* $\widehat{U}(1,\cdot) = U_1$*; however,* $\widehat{U}$ *is not symmetric: interchanging* U_0 *with* U_1 *leads to a different nonlinear interpolant.*

Lagrangian interpolation can be generalized to the case of multiple snapshots: the strategy is equivalent to the one discussed in [40, 42] and is here outlined for completeness. Given snapshots $\{(t_k, U^k := U(t_k, \cdot)\}_{k=0}^K$, *(i) we compute the forward maps* T_g^k *between* U_0 *and* U^k *for* $k = 0, \ldots, K$*; (ii) we infer the parametric map* $t \in [0,1] \mapsto \widehat{T}_g(t, \cdot)$*; (iii) we define the mapped snapshots* $\widetilde{U}^k := U^k \circ \widehat{T}_g(t_k, \cdot)$ *for* $k = 0, \ldots, K$*; (iv) we define the Lagrangian interpolant*

$$\widehat{U}(t,x) \;=\; \left(\sum_{k=0}^{K} \alpha_k(t)\,\widetilde{U}^k\right) \circ \widehat{T}_g^{-1}(t,x), \quad t \in [0,1], x \in \mathbb{R}^n, \tag{22}$$

where the coefficients $\alpha_0, \ldots, \alpha_K : [0,1] \to \mathbb{R}$ *should also be learned based on the available training data, or based on a mathematical physical model. Note that inference of the parametric map* $\widehat{T}$ *at step (ii) should preserve bijectivity: we refer to the above-mentioned literature for a discussion on this issue.*

4.2 Numerical examples

4.2.1 Simple wave field

We study the problem described in section 3.3.1. In this case, the parametric evolution of the solution is considered with respect to time. As discussed in the previous sections, optimal transportation exactly maps the initial condition to subsequent solution profiles. Here, we compare the empirical similarity transform determined based on two solution snapshots at t_0 and t_1, to the exact time-dependent solution.

Let the initial condition for the speed of sound be $a(0,x) = 2 + \tanh((x+1)/0.2)$, the initial condition for the left-going Riemann invariant $R^- = 1$ and $\gamma = 7/5$. The right-going characteristics are hence straight lines and the solution is an expansion fan traveling rightward. As an example, in Figure 3 we show two snapshots of the velocity field $u(t,x)$ at times $t_0 = 0.05$ and $t_1 = 0.4$.

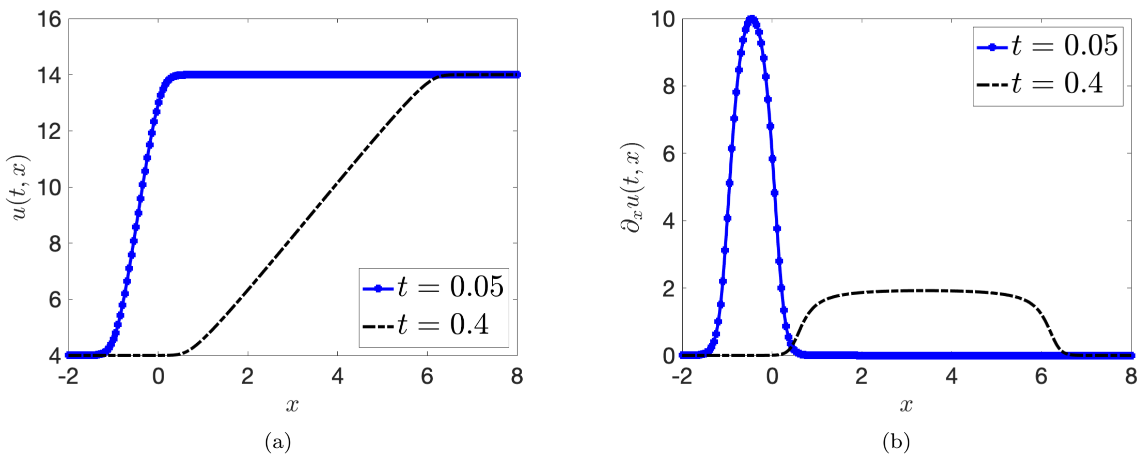
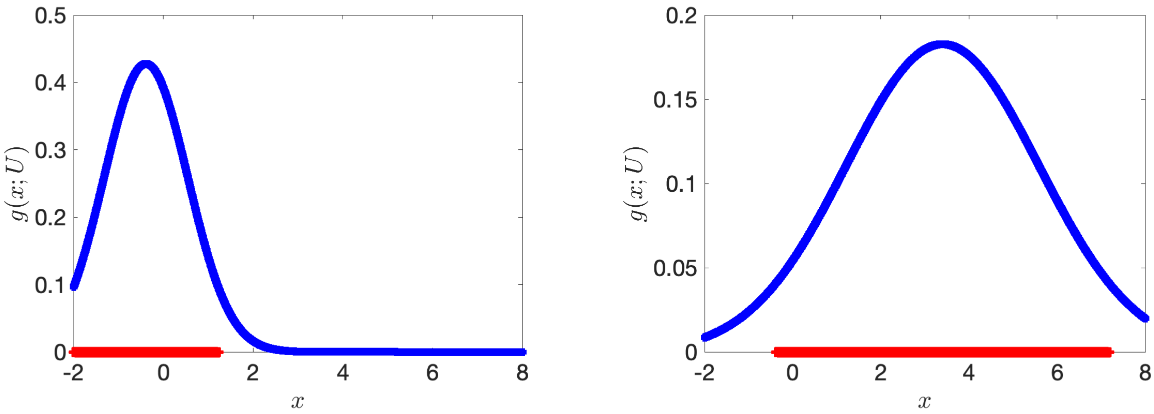


Figure 3: simple wave field. (a) solution velocity snapshots $u(t,x)$ at time $t_0 = 0.05$ (blue) and $t_1 = 0.4$ (dashed black). (b) space derivative of the velocity at time $t_0 = 0.05$ (blue) and $t_1 = 0.4$ (dashed black).

We consider the scalar testing function

$$\mathcal{T}(x;U) := \left|\frac{\partial u}{\partial x}(x)\right| - \epsilon, \tag{23}$$

with $\epsilon = 10^{-4}$: a point $x \in \mathbb{R}$ belongs to $C_{\mathcal{T}}$ if the absolute value of the space derivative of the velocity field $u(x)$ is larger than ϵ. Figure 4 shows the MLE Gaussian density distributions (blue) at times $t_0 = 0.05$ and $t_1 = 0.4$; the red points indicate the elements of the sets P_{hf}^+ (17).



(a) $t_0 = 0.05$

(b) $t_1 = 0.4$

Figure 4: simple wave field. (a)-(b) MLE Gaussian density distributions (blue) at times $t_0 = 0.05$ and $t_1 = 0.4$. Red points indicate the elements of $P_{\rm hf}^{+}$.

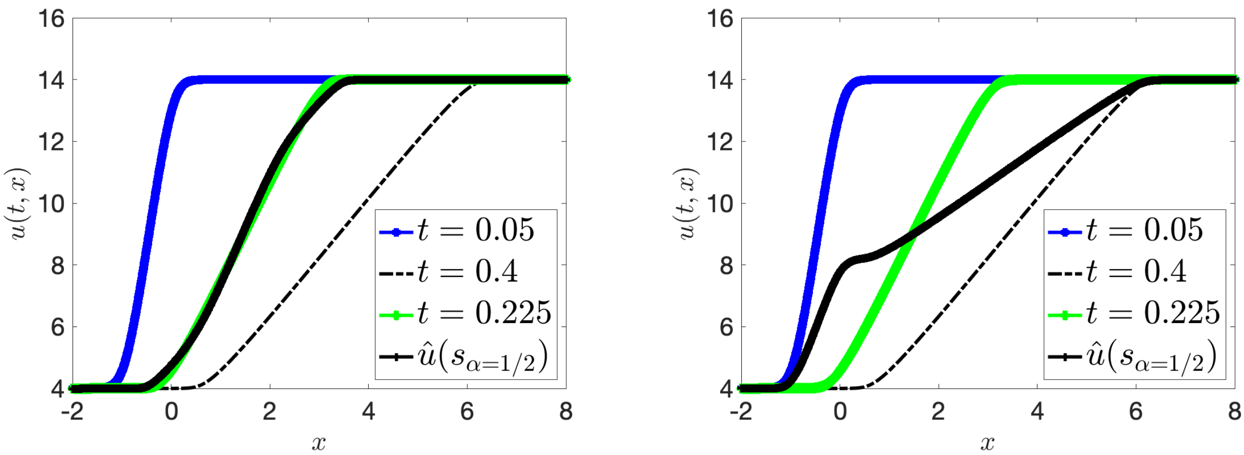
We compare the exact velocity solutions at $t \in [0.05, 0.4]$ to their L^2 projection in the manifold of displacement interpolants $\{\widehat{u}_s : s \in [0,1]\}$, cf. eq. (19). More precisely, given $\alpha \in [0,1]$, we define s_α such that

$$s_\alpha := \arg\min_s \int_{\mathbb{R}} \left(\widehat{u}_s(x) - u((1-\alpha)t_0 + \alpha t_1, x)\right)^2 \, dx, \tag{24}$$

and the projection $\widehat{u}_\alpha^{\rm proj} = \widehat{u}_{s_\alpha}$. Similarly, we project the exact solution in the convex set spanned by the two solutions at time t_0 and t_1 and we define

$$\begin{aligned} & s_\alpha^{\rm co} := \arg\min_s \int_{\mathbb{R}} \left(\widehat{u}_s^{\rm co}(x) - u(x, (1-\alpha)t_0 + \alpha t_1)\right)^2 \, dx, \\ & \text{with } \; \widehat{u}_s^{\rm co}(x) = (1-s)u(t_0, x) + s\, u(t_1, x). \end{aligned} \tag{25}$$

In Figure 5, we compare the exact velocity profile for $\alpha = 0.5$ (i.e., $t = 0.225$) to the optimal displacement interpolant $\widehat{u}_{s_\alpha}$ and to the convex interpolation $\widehat{u}^{\rm co}_{s_\alpha^{\rm co}}$. We find that $s_\alpha = 0.70$ and $s_\alpha^{\rm co} = 0.58$. The displacement interpolant captures the essential features of the solution while the convex projection in the convex set of the initial and final snapshots is completely inaccurate.

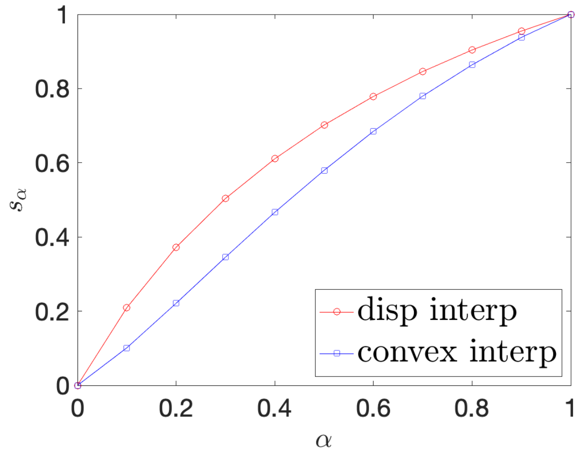


(a)

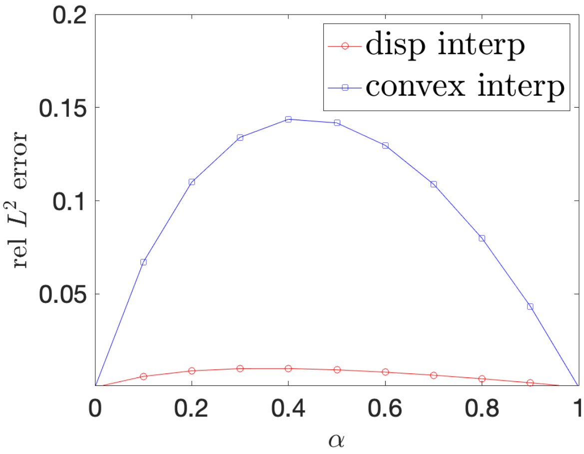
(b)

Figure 5: simple wave field. (a) velocity profile at t_0 (blue), t_1 (dashed black), exact solution for $\frac{t_0+t_1}{2}$ (green) and convex displacement interpolant $\widehat{u}_s$ for $s = 0.7$ (black dots). (b) convex interpolant $\widehat{u}_s^{\rm co}$ for $s = 0.58$ (black dots).

In Figure 6, we show the behaviors of s_α and $s_\alpha^{\rm co}$ with respect to α. As expected from the motivating examples in section 3, s_α is not necessarily linear with respect to α. In the same figure, we show the relative L^2 projection error of the exact solution with respect to the displacement interpolant $\widehat{u}$ and with respect to the convex interpolant $\widehat{u}^{\rm co}$, for several values of α. These results show that even for a smooth solution with a non-compact support, convex displacement interpolation systematically improves the approximations with respect to the convex projection in the space of the snapshots.



(a)



(b)

Figure 6: simple wave field. (a) behavior of s_α (disp interp) and $s_\alpha^{\rm co}$ (convex interp) with respect to α. (b) behavior of the relative L^2 error in $(-2, 8)$.

4.2.2 Supersonic flow past a wedge

We consider a two-dimensional compressible Euler flow of air ($\gamma = 7/5$) past a wedge. The upstream flow is supersonic and it induces a steady attached shock wave that develops from the leading edge if the upstream Mach number is within a given range, which depends on the wedge angle. In this test case we let the solution vary with respect to the upstream Mach number $M^{\rm u}$ and the wedge angle δ. We compare below the convex displacement interpolant obtained by the empirical similarity transform to the exact solution. As an example, we study the interpolation between $M_0^{\rm u} = 5, \delta_0 = 28.275$ and $M_1^{\rm u} = 8, \delta_1 = 22.80$.

We denote by $\Omega_{\rm f}$ the physical domain, and we denote by $(x_1, x_2) \mapsto M(x_1, x_2)$ the Mach number. In view of the discussion, we introduce the reference domain $\Omega_{\rm r} = (-0.5, 1) \times (0, 1)$ and the geometric transformation $\Lambda : \Omega_{\rm r} \times (-\pi/2, \pi/2) \to \Omega_{\rm f} \subset \mathbb{R}^2$, $(x_1, x_2, \delta) \mapsto \Lambda(x_1, x_2, \delta)$, such that

$$\Lambda(x_1, x_2, \delta) = \begin{cases} (x_1, x_2) & x_1 < 0 \\ (x_1, x_1 \tan(\delta) + (1 - x_1 \tan(\delta)) x_2) & x_1 \geq 0. \end{cases} \tag{26a}$$

and its inverse

$$\Theta(x_1, x_2, \delta) = \begin{cases} (x_1, x_2) & x_1 < 0 \\ \left(x_1, \dfrac{x_2 - x_1 \tan(\delta)}{1 - x_1 \tan(\delta)} \right) & x_1 \geq 0. \end{cases} \tag{26b}$$

Finally, we define the mapping $(x_1, x_2, \delta) \mapsto \Phi(x_1, x_2, \delta)$ such that

$$\Phi(x_1, x_2, \delta) = \Lambda\left(\Theta\left(x_1, x_2, \bar{\delta}\right), \delta\right), \qquad \bar{\delta} = \frac{\delta_0 + \delta_1}{2}, \tag{26c}$$

which maps $\Omega_{\rm f}(\bar{\delta})$ into $\Omega_{\rm f}(\delta)$.

It is possible to show that the exact solution is piecewise-constant and exhibits a straight shock discontinuity that is generated at the wedge leading edge. If we denote by θ the shock angle and by $M^{\rm d}$ the downstream Mach number, we obtain the expression for M:

$$M(x_1, x_2) = \begin{cases} M^{\rm u} & \text{if } x_2 > x_1 \tan(\theta) \\ M^{\rm d} & \text{if } x_2 < x_1 \tan(\theta) \end{cases} \tag{27a}$$

Given M^{u} and δ, we can employ the relationships (cf. [1])

$$
\begin{aligned}
\cot(\delta) \;&=\; \tan(\theta)\left(\frac{(\gamma+1)(M^{\mathrm{u}})^2}{2(M^{\mathrm{u}}\sin(\theta))^2-1}\;-\;1\right), \\
\left(M^{\mathrm{d}}\sin(\delta-\theta)\right)^2 \;&=\; \frac{(\gamma-1)\left(M^{\mathrm{u}}\sin(\theta)\right)^2+2}{2\gamma\left(M^{\mathrm{u}}\sin(\theta)\right)^2-(\gamma-1)},
\end{aligned}
\tag{27b}
$$

to find the downstream Mach number and the shock angle.

In order to deal with geometry variations, we pursue two different strategies.

1. *Extension:* we extend the Mach number M to $\mathbb{R}^2$ for all parameters.
2. *Geometric registration:* we apply the interpolation procedure to the mapped field $\widetilde{M}(x_1, x_2; M^{\mathrm{u}}, \delta) = M(\Phi(x_1, x_2; \delta); M^{\mathrm{u}}, \delta)$, which is defined in $\Omega_{\mathrm{f}}(\bar{\delta})$ for all values of the parameters M^{u}, δ.

Note that since the proposed displacement interpolation strategy does not preserve boundaries, extension outside Ω_{f} is necessary for both techniques; note also that, since the solution is piecewise-constant and the shock curve is linear, the extension is straightforward.

We pursue the first approach based on extension. Towards this end, we consider a regular 151×101 grid in the rectangle Ω_{r} and we consider the scalar testing function

$$
\mathcal{T}(x; U) = \frac{1}{\epsilon}\left|M(x_1+\epsilon, x_2) - M(x_1, x_2)\right| \;-\; 1,
\tag{28}
$$

where $\epsilon = 10^{-2}$ is equal to the size of the grid. Figure 7 shows the results: in Figure 7(a), we show the selected points for three choices of the parameter pair $\mu_\alpha := (1-\alpha)(M_1^{\mathrm{u}}, \delta_1) + \alpha(M_1^{\mathrm{u}}, \delta_1)$, $\alpha \in \{0, 1/2, 1\}$; in Figure 7(b), we compare the exact Mach profile for $x_2 = 0.3$ and $x_1 \in (0,1)$ for $\mu_{\alpha=1/2}$ with the optimal convex displacement interpolant $\widehat{M}_{s_\alpha}$ and the convex interpolation $\widehat{M}_{s_\alpha}^{\mathrm{co}}$; in Figure 7(c), we show the behavior of s_α and s_α^{co} defined as in (24) and (25), respectively; in Figure 7(d), we show the behavior of the L^2 relative projection error. Note that the optimal value of s is a linear function of α; note also that displacement interpolation offers extremely accurate results compared to the more standard convex interpolation.

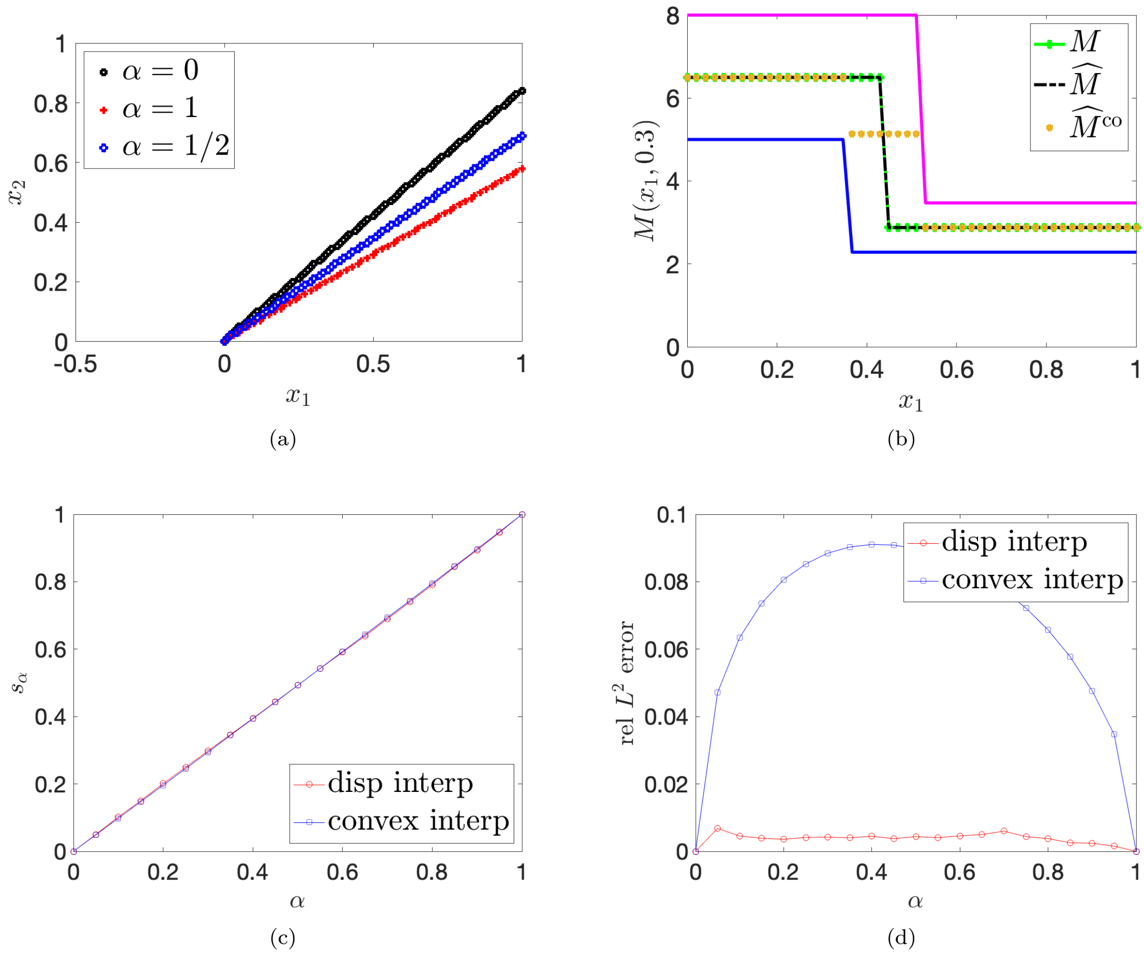


Figure 7: supersonic flow past a wedge; approach based on extension. (a) selected points $P_{\rm hf}^{+}$ for three choices of the parameter pair μ_α; (b) Mach profile for $x_2 = 0.3$ at μ_0 (blue), μ_1 (violet), $\mu_{1/2}$ (green), and convex displacement interpolant $\widehat{M}$ and convex interpolant $\widehat{M}^{\rm co}$ for $s = 1/2$; (c) behavior of s_α and $s_\alpha^{\rm co}$ in (24) and (25); (d) behavior of the relative L^2 projection error.

We also pursue the second approach based on geometric registration. Towards this end, we consider the same regular 151×101 grid in the rectangle $\Omega_{\rm r}$, but we discard points outside $\Omega_{\rm f}(\bar{\delta})$; then, we consider the scalar testing function

$$\mathcal{T}'(x; U) = \frac{1}{\epsilon}\left|\widetilde{M}(x_1 + \epsilon, x_2) - \widetilde{M}(x_1, x_2)\right| - 1, \quad \epsilon = 10^{-2}. \tag{29}$$

Figure 8 replicates the same tests considered for the other strategy: as for the previous approach, displacement interpolation significantly outperforms linear interpolation. We note, however, that the geometric mapping has a beneficial effect on the performance of the linear approach, while it is slightly detrimental for displacement interpolation.

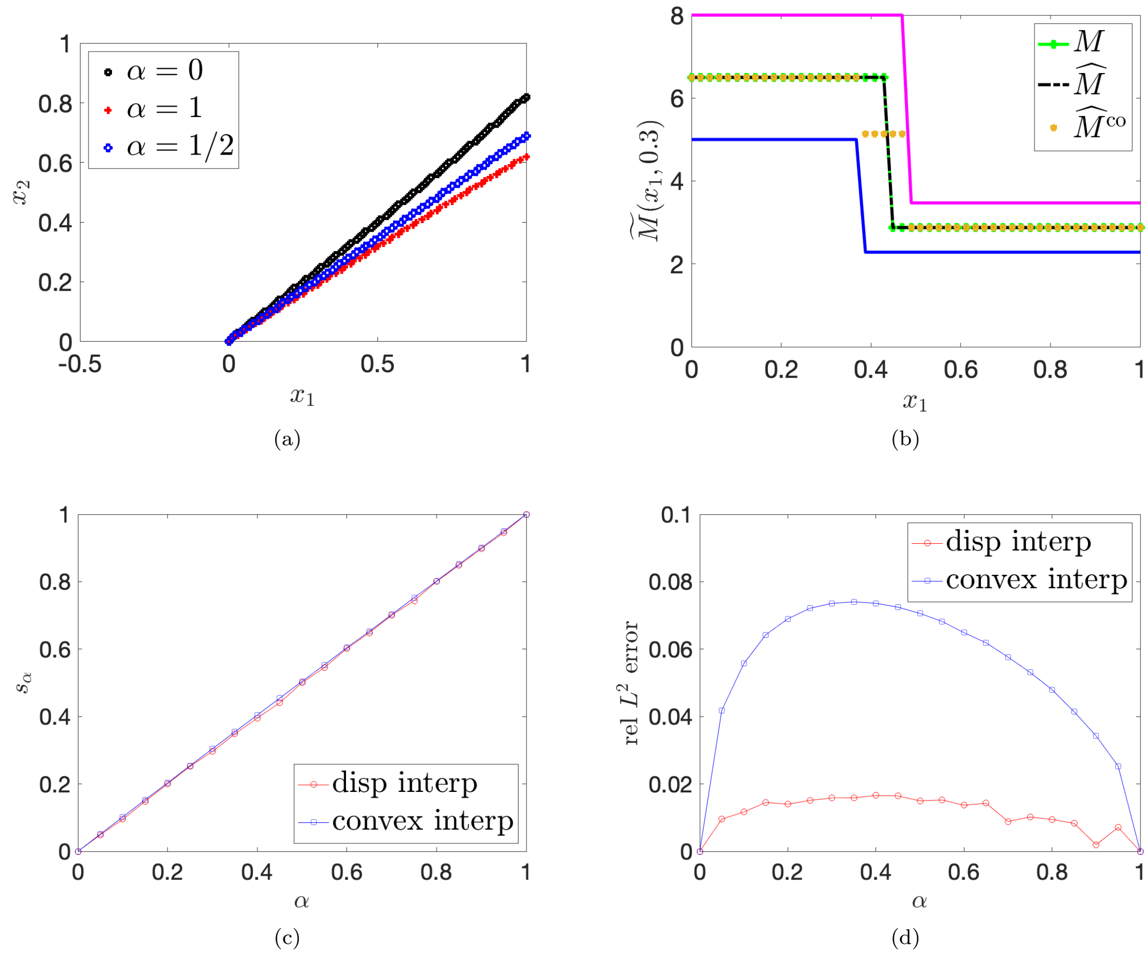


Figure 8: supersonic flow past a wedge; approach based on registration. (a) selected points $P_{\rm hf}^+$ for three choices of the parameter pair μ_α; (b) Mach profile for $x_2 = 0.3$ at μ_0 (blue), μ_1 (violet), $\mu_{1/2}$ (green), and convex displacement interpolant $\widehat{M}$ and convex interpolant $\widehat{M}^{\rm co}$ for $s = 1/2$; (c) behavior of s_α and $s_\alpha^{\rm co}$ in (24) and (25); (d) behavior of the relative L^2 projection error.

4.2.3 Transonic flow past an airfoil

We consider a two-dimensional transonic flow past a NACA 0012 airfoil at angle of attack $\alpha = -4^o$; we let the solution vary with respect to the free-stream Mach number $M \in [0.77, 0.83]$; we study the interpolation between $M_0 = 0.77$ and $M_1 = 0.83$. Related examples are considered in [39, 41]. We resort to a discontinuous Galerkin (DG) discretization with artificial viscosity to estimate the solution field; computations are performed in the domain $\Omega = (-4, 10) \times (-10, 10) \setminus \Omega_{\rm naca}$ where $\Omega_{\rm naca}$ is the domain associated with the airfoil. Figure 9 shows the behavior of the flow density for $M = 0.77$ and $M = 0.83$: note that the solution develops a shock on the lower boundary of the airfoil that is extremely sensitive to the value of the Mach number.

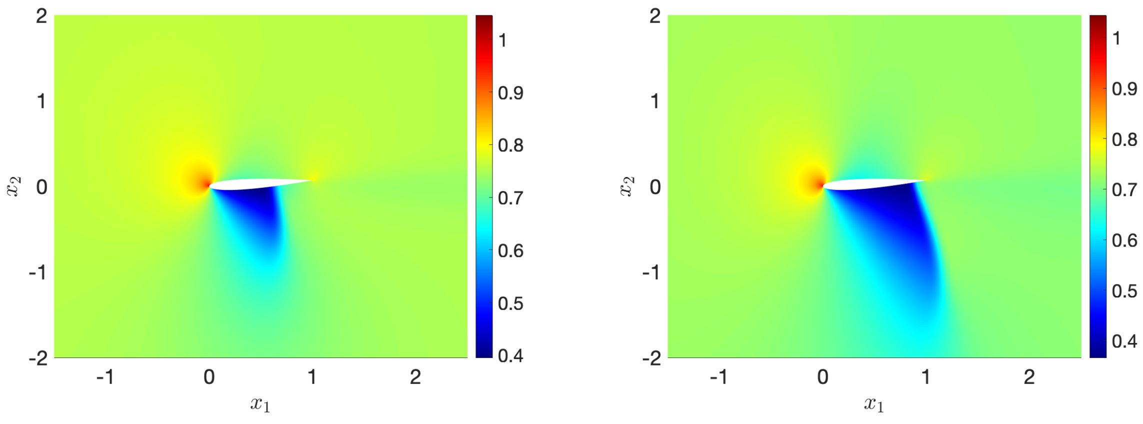


Figure 9: transonic flow past an airfoil. Flow density for two values of the Mach number.

Figure 10(a) shows the computational mesh used for DG calculations. For simplicity of implementation, we here apply our interpolation procedure in the mapped domain corresponding to angle of attack $\alpha = 0^o$; furthermore, for efficiency reasons, interpolation is performed on the structured mesh in Figure 10(b). Since the proposed approach does not ensure bijectivity in the domain Ω, it is necessary to extend the solution field inside the airfoil: we here build the extension based on the solution to a Laplace problem in the interior of the airfoil.

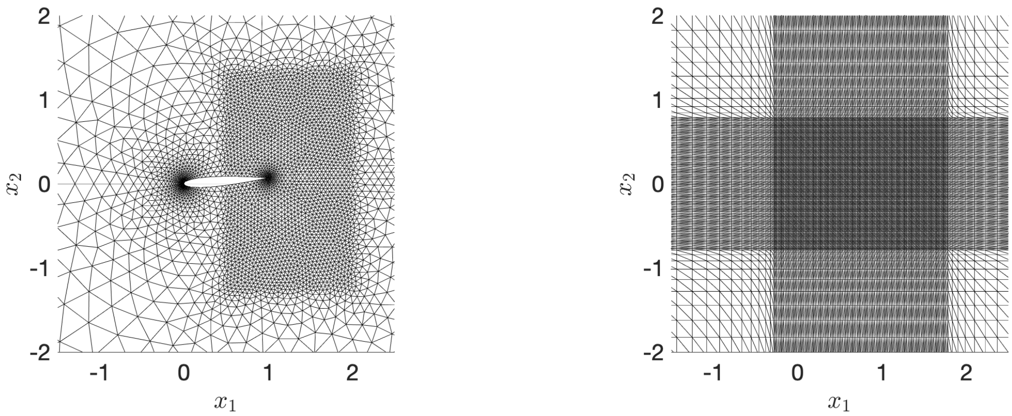
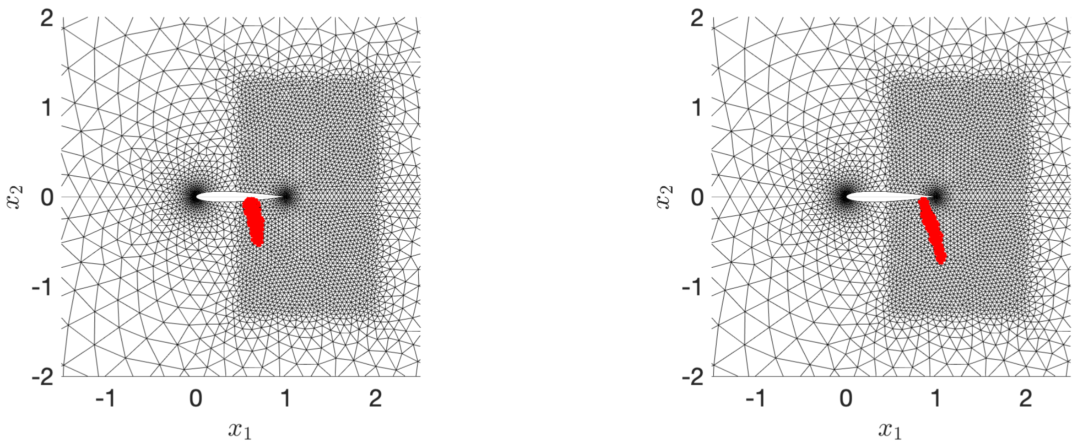


Figure 10: transonic flow past an airfoil. (a) mesh used for DG calculations; (b) structured mesh used for interpolation.

The definition of an effective scalar testing function that detects the presence of shock discontinuities is challenging due to numerical dissipation. We here proceed as follows: first, we define $P_{\rm hf}$ as the set of elements' centers and we compute the indicator

$$d_k = \max_{x \in \mathtt{D}_k} \left| \phi(x; U) \right|, \quad \phi(x; U) := \frac{(-\nabla \cdot u)^+}{\sqrt{(\nabla \cdot u)^2 + \|\nabla \times u\|_2^2 + a^2}} \frac{\|\nabla p\|_2}{p + \epsilon} \|u\|_2 \tag{30}$$

where $\mathtt{D}_k$ denotes the k-th element of the DG mesh, $k = 1, \ldots, N_{\rm e}$, and $\epsilon = 10^{-4}$; then, we define $P_{\rm hf}^{+}$ as the set associated with the largest 0.5% values of the indicator $\{d_k\}_{k=1}^{N_{\rm e}}$. We observe that the first term in $\phi(x; U)$ is a modified Ducros sensor (see [26, 30]) that identifies strong compressions of the flow, the second term identifies regions characterized by large pressure gradients and the third term is intended to discard regions where the velocity is small — such as the leading edge. We further remark that the indicator (30) is used in [11] to define the artificial viscosity for high-order DG discretizations of inviscid flows. Figure 11 shows the selected points for two values of the Mach number.

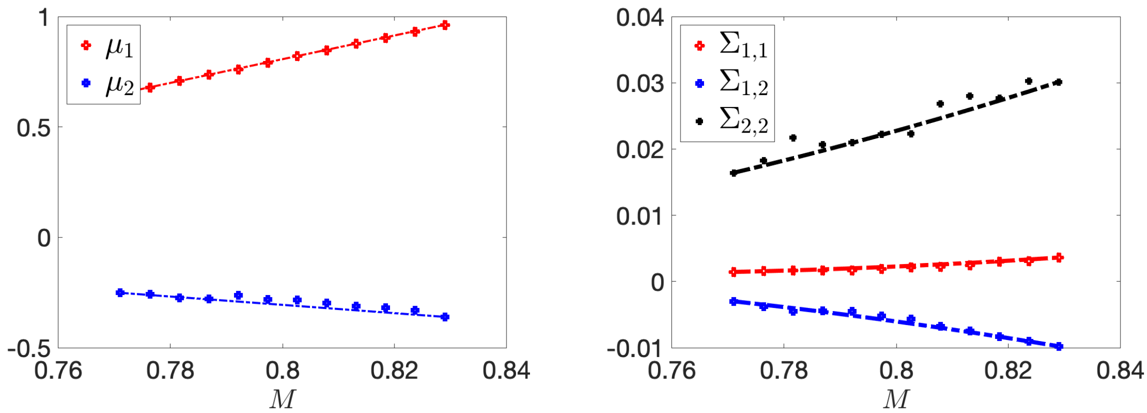


(a) $M = 0.77$

(b) $M = 0.83$

Figure 11: transonic flow past an airfoil. Elements of $P_{\rm hf}^{+}$ for two values of the Mach number.

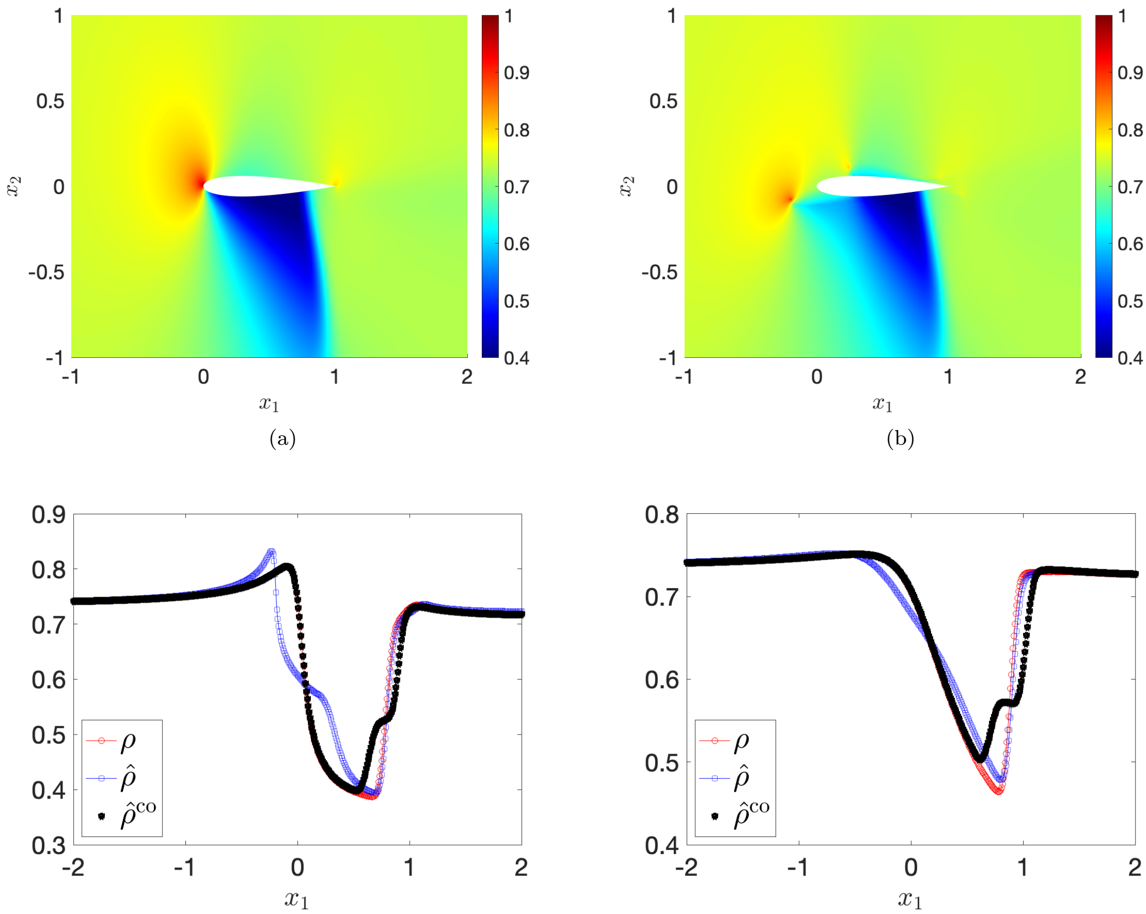
In Figure 12, we investigate the behavior of the MLE estimates $\mu_{\rm mle}$ and $\Sigma_{\rm mle}$ in (18) with respect to the Mach number. We observe that both mean and variance are smooth functions of the parameter. In Figure 13, we compare the density field for $M = 0.8$ ($\alpha = 0.5$) with the convex displacement interpolant (19) with $s = \alpha$; we further provide horizontal slices of the DG solution (in red), the convex displacement interpolant (in blue) and the convex interpolant (in black) for two values of x_2. We observe that the convex displacement interpolant is extremely accurate in the proximity of the shock, while it is highly inaccurate far from the shock, especially in the proximity of the airfoil.



(a)

(b)

Figure 12: transonic flow past an airfoil. Behavior of the MLE estimates $\mu_{\rm mle}$ and $\Sigma_{\rm mle}$ (18) with respect to the Mach number.



(a) (b)

(c) $x_2 = -0.1$ (d) $x_2 = -0.5$

Figure 13: transonic flow past an airfoil. (a)-(b) behavior of the density field for $M = 0.8$ and of the convex displacement interpolant (19) for $s = 1/2$. (c)-(d) two horizontal slices of truth and predicted density profiles.

5 Extension: boundary-aware transportation of Gaussian models

The examples of the previous section show that convex displacement interpolation (19) based on optimal transportation of Gaussian models is effective if (i) boundaries are not present (cf. section 4.2.1) or the extension of the solution outside the domain is trivial (cf. section 4.2.2), and (ii) the solution field presents a single coherent structure that is well-approximated by an ellipsoid. Inaccuracy of the displacement interpolation for the example in section 4.2.3 is the consequence of two factors. First, since displacement interpolation does not preserve the boundaries of the domain, interpolation might be highly inaccurate in the neighborhood of the airfoil (cf. Figure 13(c)). Second, the Gaussian model considered is not able to take into account the coherent structures that develop at leading and trailing edges: it is thus a too simplistic representation of the solution field.

Based on these considerations, we propose here an extension of the approach in section 4: convex displacement interpolation based on boundary-aware (BA) transportation of multiple Gaussian models. We investigate performance of our approach for the transonic flow test case introduced in section 4.2.3. For simplicity, in the remainder we assume that the domain Ω is a Lipschitz two-dimensional domain.

5.1 Methodology

Given $M \in \mathbb{N}$, we introduce the approximation map $\mathcal{N} : \Omega \times \mathbb{R}^M \to \mathbb{R}^2$ and the set $\mathcal{A}_{\mathrm{bj}} \subset \mathbb{R}^M$ such that $\mathcal{N}(\cdot, \mathbf{a})$ is a bijection in Ω for all $\mathbf{a} \in \mathcal{A}_{\mathrm{bj}}$. We pursue the approach in [41] to define $\mathcal{N}$: we refer to Appendix A for further details. We denote by $P_{\mathrm{hf}}^{+,0} = \{y_j^0\}_{j=1}^{N_{\mathrm{hf},0}^{+}}$ and $P_{\mathrm{hf}}^{+,1} = \{y_j^1\}_{j=1}^{N_{\mathrm{hf},0}^{+}}$ the selected points for U_0 and U_1, and we denote by T_g and R_g the optimal maps obtained using (15b).

Then, we define $\widehat{\mathbf{a}}_{0,1} \in \mathbb{R}^M$ to minimize

$$\sum_{j=1}^{N_{\rm hf,0}^+} \| T_g(y_j^0) - \mathcal{N}(y_j^0, \mathbf{a}) \|_2^2 + \mathfrak{P}(\mathbf{a}), \quad \text{subject to } \mathfrak{C}(\mathbf{a}) \leq 0, \tag{31a}$$

where $\mathfrak{P}$ is a suitable regularization that penalizes the H^2 seminorm of the mapping and $\mathfrak{C}$ is a bijectivity constraint that, combined with $\mathfrak{P}$, enforces that $\widehat{\mathbf{a}}_{0,1}$ belongs to $\mathcal{A}_{\rm bj}$: we refer to [41] for the details. Similarly, we define $\widehat{\mathbf{a}}_{1,0} \in \mathbb{R}^M$ to minimize

$$\sum_{j=1}^{N_{\rm hf,1}^+} \| R_g(y_j^1) - \mathcal{N}(y_j^1, \mathbf{a}) \|_2^2 + \mathfrak{P}(\mathbf{a}), \quad \text{subject to } \mathfrak{C}(\mathbf{a}) \leq 0. \tag{31b}$$

In conclusion, we introduce the boundary-aware (BA) convex displacement interpolation as

$$\widehat{U}(s,x) = (1-s) U_0 \circ \widetilde{W}_g(s,x) + s U_1 \circ \widetilde{T}_g(1-s,x), \quad s \in [0,1], \ x \in \Omega, \tag{32}$$

where $\widetilde{T}_g(s,x) = \mathcal{N}(x, s \cdot \widehat{\mathbf{a}}_{0,1})$ and $\widetilde{W}_g(s,x) = \mathcal{N}(x, s \cdot \widehat{\mathbf{a}}_{1,0})$. In the next section, we investigate performance of (32) for the transonic flow test case.

Some comments are in order. First, registration provides sub-optimal — in the sense of optimal transportation — bijective-in-Ω approximations of the Gaussian maps obtained using (15b): we can thus interpret $\widetilde{T}_g$ and $\widetilde{W}_g$ as approximate projections of the actions on marked points of the optimal transport maps T_g and W_g onto the space of bijective maps in Ω. Second, we cannot in general guarantee that $\widetilde{T}_g$ and $\widetilde{W}_g$ are bijections in Ω for all $s \in [0,1]$: nevertheless, in our experience, provided that the distance — in the sense of Wasserstein — between the Gaussian models associated with $P_{\rm hf}^{+,0}$ and $P_{\rm hf}^{+,1}$ is moderate, the approach leads to bijective maps for all $s \in [0,1]$.

Following [8], it is straightforward to extend (31)-(32) to track multiple structures. First, given the points

$$\left\{ y_j^{0,k} \ : \ j = 1, \ldots, N_{\rm hf,0,k}^+, k = 1, \ldots, N_{\rm g} \right\}, \quad \left\{ y_j^{1,k} \ : \ j = 1, \ldots, N_{\rm hf,1,k}^+, k = 1, \ldots, N_{\rm g} \right\}$$

we first compute the mappings $\{T_{\rm g}^{(k,k')}\}_{k,k'}$ and the associated Wasserstein distances $\{W_2^{(k,k')}\}_{k,k'}$ using the identities in section 4.1.2. Then, we identify the permutation I of $\{1, \ldots, N_{\rm g}\}$ that minimizes

$$\sum_{k=1}^{N_{\rm g}} W_2^{k, I_k}$$

over all possible permutations. Finally, we compute $\widehat{\mathbf{a}}_{0,1} \in \mathbb{R}^M$ to minimize

$$\sum_{k=1}^{N_{\rm g}} \sum_{j=1}^{N_{\rm hf,0,k}^+} \| T_g^{k,I_k}(y_j^{0,k}) - \mathcal{N}(y_j^{0,k}, \mathbf{a}) \|_2^2 + \mathfrak{P}(\mathbf{a}), \quad \text{subject to } \mathfrak{C}(\mathbf{a}) \leq 0, \tag{33a}$$

and we compute $\widehat{\mathbf{a}}_{1,0} \in \mathbb{R}^M$ to minimize

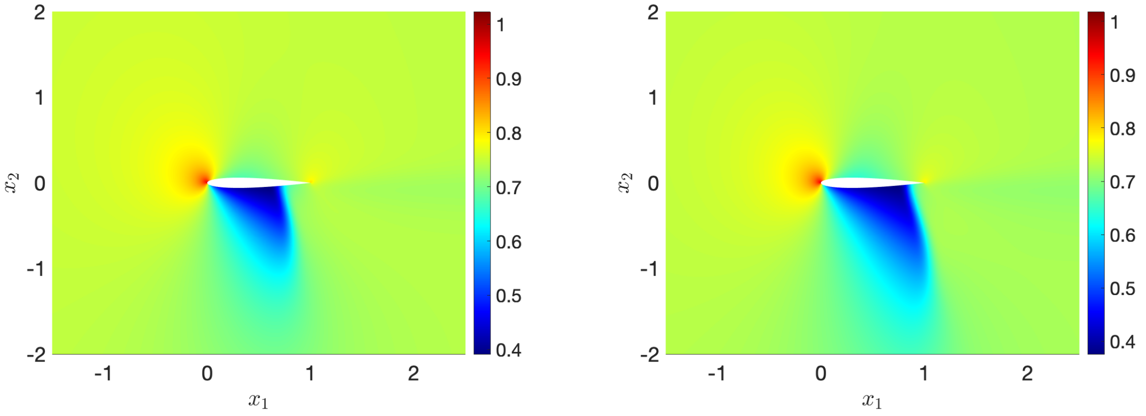
$$\sum_{k=1}^{N_{\rm g}} \sum_{j=1}^{N_{\rm hf,1,k}^+} \| R_g^{k,I_k}(y_j^{1,k}) - \mathcal{N}(y_j^{1,k}, \mathbf{a}) \|_2^2 + \mathfrak{P}(\mathbf{a}), \quad \text{subject to } \mathfrak{C}(\mathbf{a}) \leq 0, \tag{33b}$$

which are of the same form as (31). We validate this approach through the vehicle of a transonic flow with two shocks.

5.2 Numerical results

5.2.1 Transonic flow at angle of attack 4^o

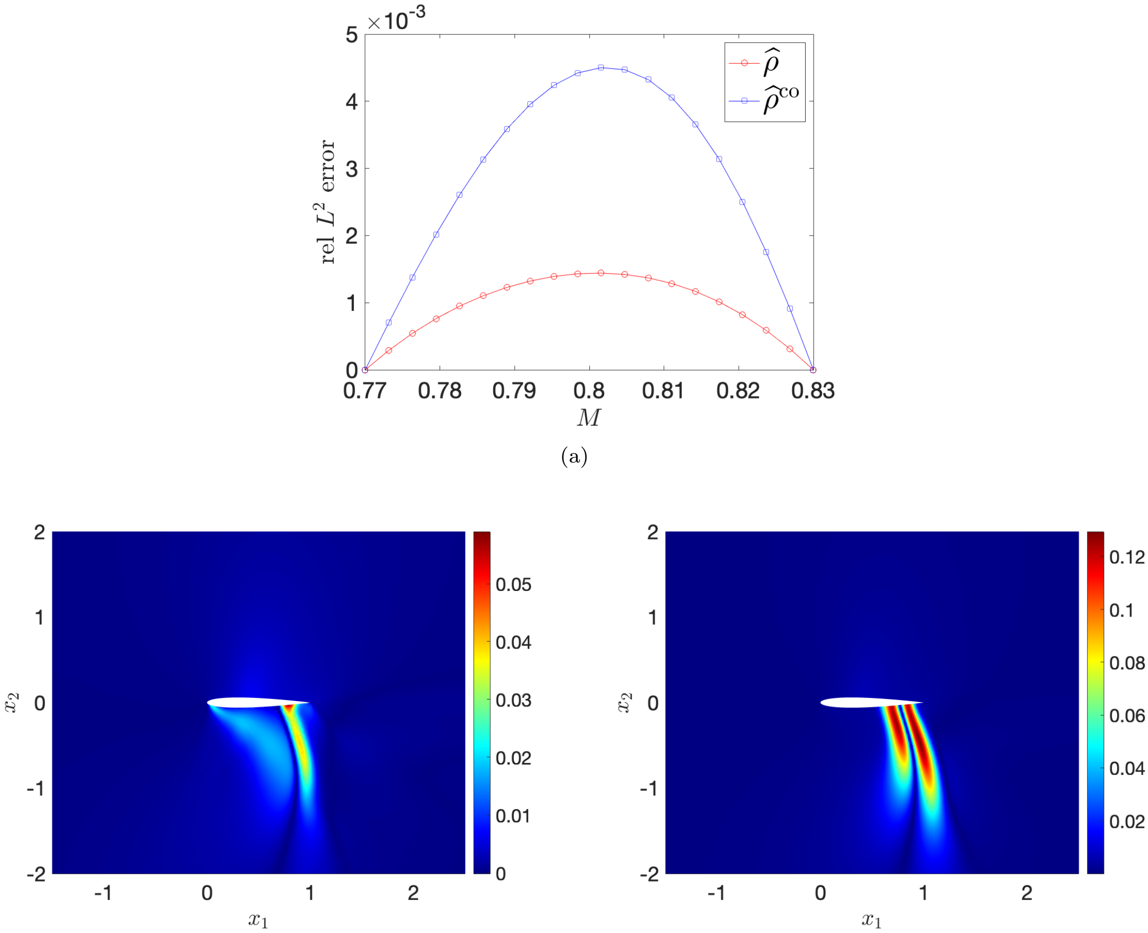
Figure 14 shows the behavior of the BA convex displacement interpolation (32) for two values of $s \in [0,1]$: we observe that the proposed interpolation preserves the structures at leading and trailing edges and is able to smoothly deform the shock attached to the airfoil.



(a) $s = 0.25$ (b) $s = 0.75$

Figure 14: boundary-aware transportation of Gaussian models (angle of attack 4^o). BA convex displacement interpolation (32) for two values of s.

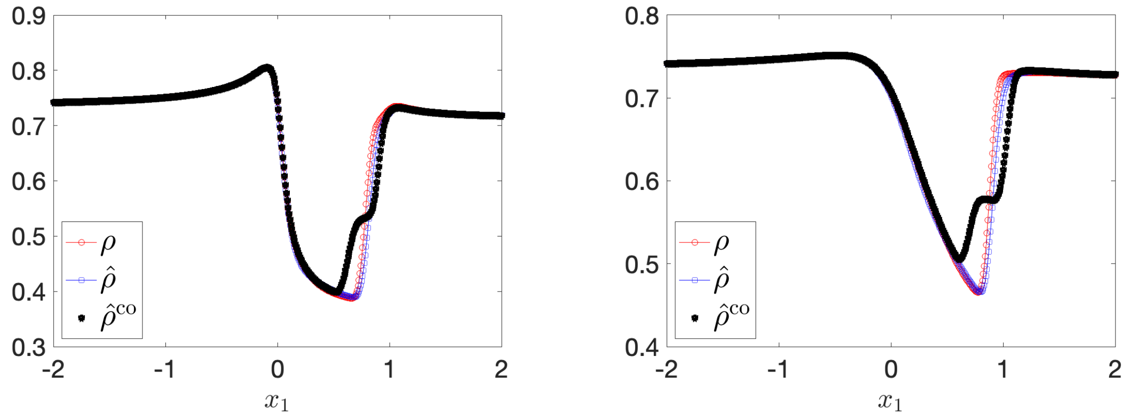
Figure 15 compares performance of the nonlinear interpolation (32) with the linear convex interpolation for $M \in [0.77, 0.83]$ — we here consider $s_\alpha = \alpha$ for both linear and nonlinear interpolation. Similarly, Figure 16 compares horizontal slices of the truth density profile with horizontal slices of linear and nonlinear interpolation (32), for $M = 0.8$. We observe that nonlinear interpolation leads to more accurate performance in terms of relative L^2 error, and in particular is more accurate in the proximity of the shock.



(a)

(b) (c)

Figure 15: boundary-aware transportation of Gaussian models (angle of attack 4^o). (a) behavior of the relative L^2 error for BA convex displacement interpolation. b)-(c) behavior of the error fields $|\widehat{\rho}(s,x) - \rho(x; M)|$ and $|\widehat{\rho}^{\rm co}(s,x) - \rho(x; M)|$ for $M = 0.8$ and $s = 0.5$.



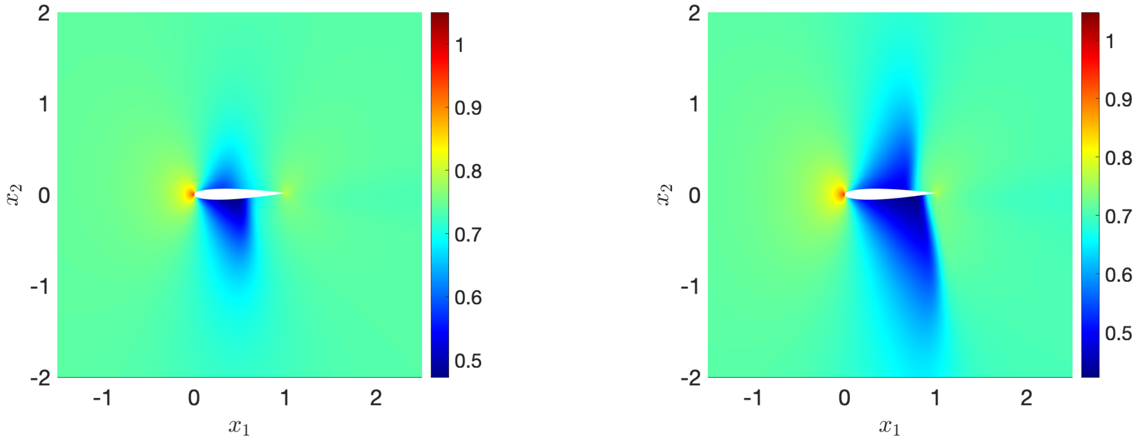
(a) $x_2 = -0.1$

(b) $x_2 = -0.5$

Figure 16: boundary-aware transportation of Gaussian models (angle of attack 4^o). (a)-(b) horizontal slices of truth and predicted density profiles for $M = 0.8$.

5.2.2 Transonic flow at angle of attack 1^o

We consider a two-dimensional transonic flow past a NACA 0012 airfoil at angle of attack $\alpha = -1^o$; we let the solution vary with respect to the free-stream Mach number $M \in [0.8, 0.86]$. As shown in Figure 17, the flow density exhibits two shocks that are very sensitive to the value of the parameter.

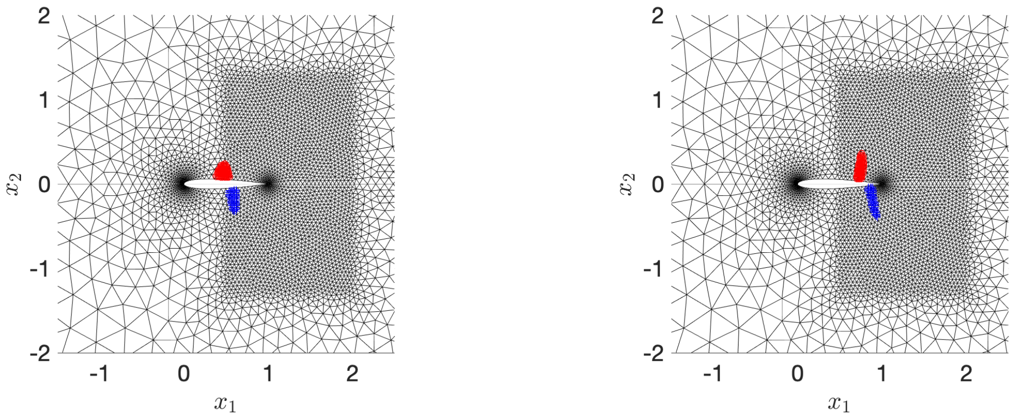


(a) $M = 0.8$

(b) $M = 0.86$

Figure 17: transonic flow past an airfoil at angle of attack 1^o. Flow density for two values of the Mach number.

Figure 18 shows the selected points $P_{\rm hf}^{+}$ for two values of the Mach number. We resort to the same indicator introduced in (30) to identify the set $P_{\rm hf}^{+}$. To facilitate the interpolation task we discard points outside $(0,1)\times\mathbb{R}$; furthermore, we separate *a priori* the two clouds of points by discriminating between positive and negative heights. Note that the latter expedient allows us to apply the procedure described in section 4.1.2 to build the Gaussian models for the upper and lower shocks, and ultimately robustifies the identification task. In the future, we wish to investigate performance of automated detection algorithms for Gaussian mixtures, [25].

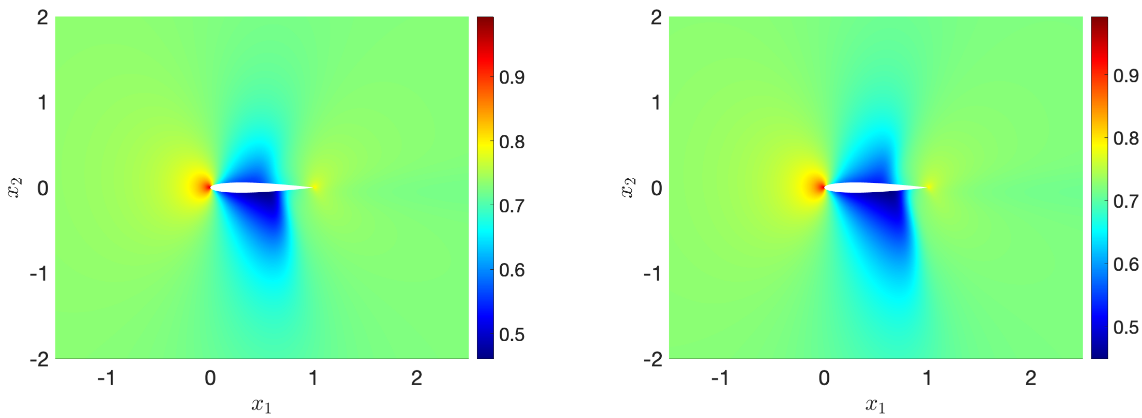


(a) $M = 0.8$

(b) $M = 0.86$

Figure 18: transonic flow past an airfoil at angle of attack 1^o. Elements of $P_{\rm hf}^{+}$ for two values of the Mach number.

Figures 19, 20 and 21 show performance of our nonlinear interpolation procedure. Figure 19, shows the BA convex displacement interpolation for two values of the parameter s: we observe that the interpolation procedure is able to generate physically-meaningful interpolations. Figure 20 compares the behavior of the relative L^2 error for BA convex displacement interpolation and linear convex interpolation: similarly, Figure 21 shows horizontal slices of truth and predicted density profiles for $M = 0.83$ — we set $s = 1/2$ for both linear and nonlinear interpolation. Note that the shock on the upper part of the airfoil is not tracked as accurately as the lower shock by our nonlinear interpolation: this might be due to the inaccuracy of the Gaussian model and might also be due to the fact that the optimal value of s is not necessarily a linear function of $\alpha = \frac{M-M_0}{M_1-M_0}$ (see discussion in section 3 and results in Figure 6).



(a) $s = 0.25$

(b) $s = 0.75$

Figure 19: boundary-aware transportation of Gaussian models (angle of attack 1^o). BA convex displacement interpolation (32) for two values of s.

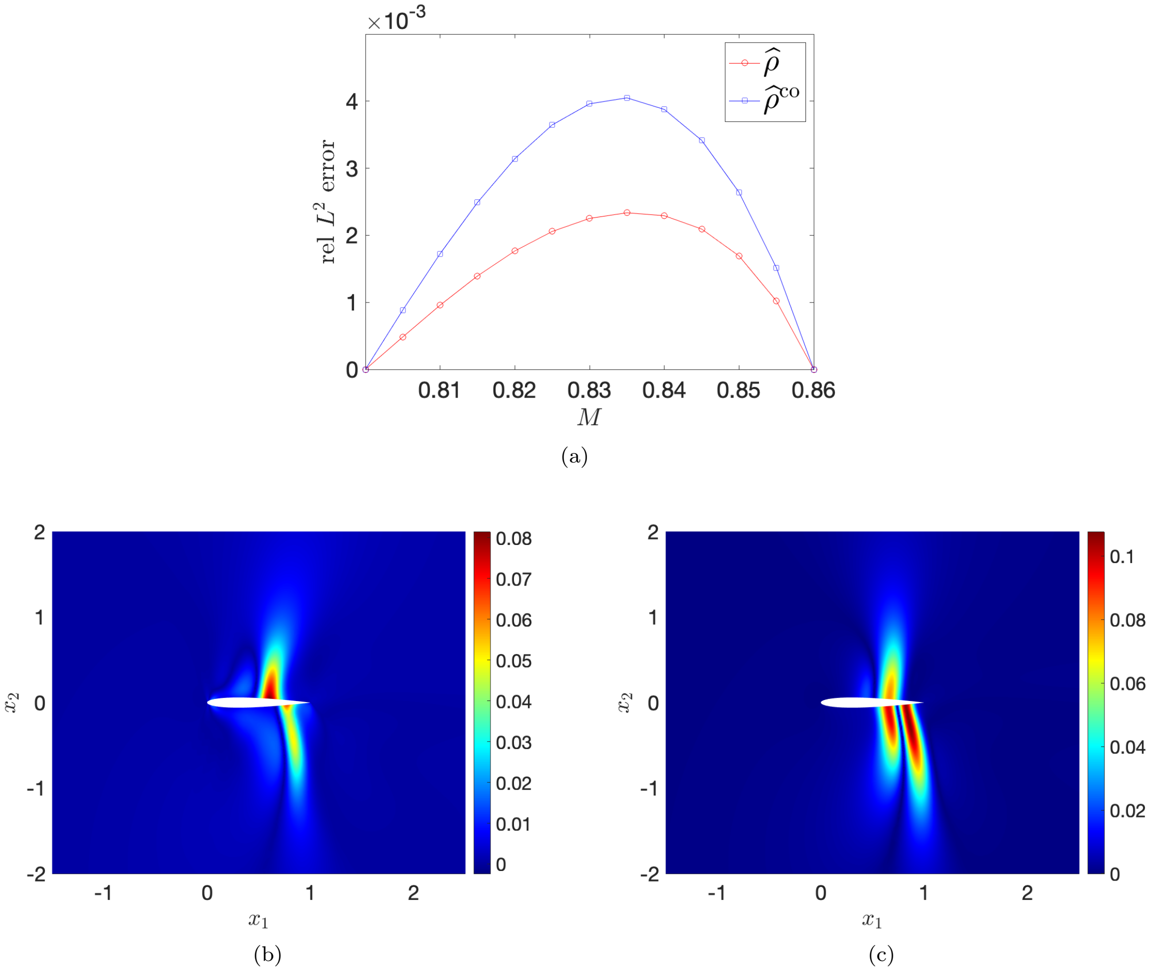


Figure 20: boundary-aware transportation of Gaussian models (angle of attack 1^o). (a) behavior of the relative L^2 error for BA convex displacement interpolation. (b)-(c) behavior of the error fields $|\widehat{\rho}(s,x) - \rho(x;M)|$ and $|\widehat{\rho}^{\text{co}}(s,x) - \rho(x;M)|$ for $M = 0.83$ and $s = 0.5$.

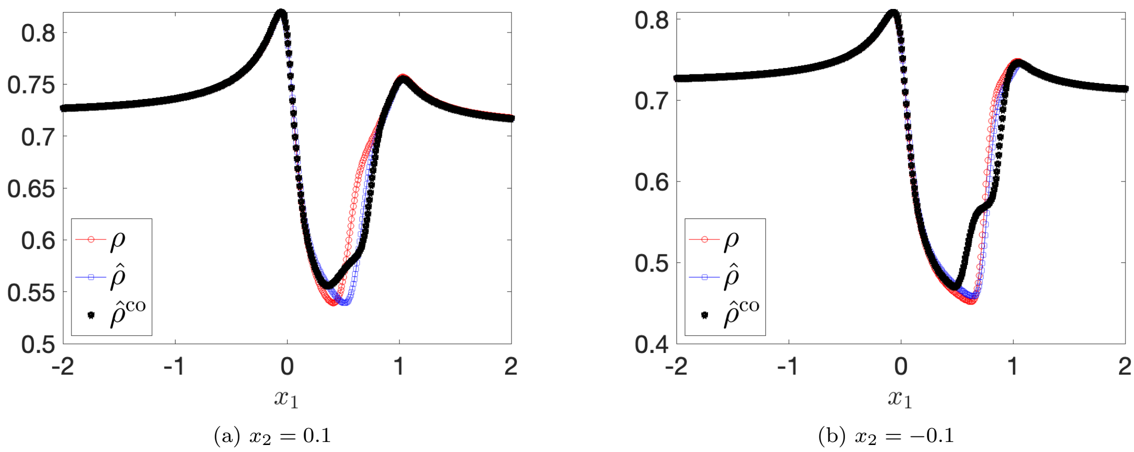


Figure 21: boundary-aware transportation of Gaussian models (angle of attack 1^o). (a)-(b) horizontal slices of truth and predicted density profiles for $M = 0.83$.

6 Conclusions

We presented a general interpolation technique based on optimal transportation of Gaussian models for parametric advection-dominated problems. Application of optimal transportation to a Gaussian model of the solution field, for which the transport map is known explicitly, simplifies the implementation of the method and allows to deal with fields that are neither scalar nor positive and that do not satisfy conservation of mass over the parameter domain. We presented several examples to establish the connection between self-similarity and optimal transportation, which is at the foundation of the proposed technique. Furthermore, we presented several numerical investigations to illustrate the many features of the approach and assess strengths and weaknesses.

As discussed in section 4.1.2 — and shown numerically in the example of section 4.2.3 — the choice of the Gaussian distribution model might not suffice to properly track relevant coherent structures of the flow; furthermore, the approach is not suited to accurately represent the flow in the proximity of the boundaries. To address these issues, we proposed in section 5 a more sophisticated interpolation procedure that combines Gaussian mixture models with a nonlinear registration procedure.

The key elements of our approach are (i) a scalar testing function $\mathcal{T}(\cdot;U):\mathbb{R}^n\to\mathbb{R}$, (ii) a (non-necessarily conforming) mapping technique for the construction of T_g, W_g, and (iii) a registration (or mesh morphing) procedure to project the mappings T_g, W_g onto a suitable subspace of bijective maps in Ω. In this work, we proposed simple yet effective strategies based on (i) physics-informed scalar testing functions, (ii) Gaussian models and optimal transportation maps between Gaussian distributions, and (if needed) (iii) the registration approach proposed in [41]. In the future, we aim to design more accurate strategies for each of the three steps, and discuss the application to a broad class of problems in computational mechanics.

The aim of this work is to devise a nonlinear interpolation procedure for continuum mechanics applications that is simple to implement, interpretable, and robust for small training sets. In the past decade, the spectacular successes of deep learning methods [20] for data science applications have motivated the development of deep convolutational architectures in model reduction [5, 12, 18, 21, 28]: in our experience, these approaches require large training sets and are difficult to interpret; furthermore, for small datasets, convergence to local minima might impact their robustness and generalization properties. In this regard, we observe that our approach might be interpreted as an Eulerian and non-intrusive counterpart of the registration-based approach in [40, 41]: we remark that the latter relies on the introduction of a template space and thus cannot deal with datasets of very modest size $n_{\text{train}}=\mathcal{O}(1)$. In the future, we aim to devise strategies to optimally combine the many available linear and nonlinear reduction strategies for a wide range of offline computational budgets. In this respect, similarly to [4], we wish to apply the proposed technique in the framework of projection-based schemes, to augment the dataset of snapshots used to generate the reduced-order basis.

Acknowledgements

The authors acknowledge the support by European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Actions, grant agreement 872442 (ARIA). Tommaso Taddei also acknowledges the support of IdEx Bordeaux (projet EMERGENCE 2019).

A Construction of the approximation map $\mathcal{N}$

Let $\widehat{\Omega}=(0,1)^2$ be the unit square, let $\mathbb{P}_J$ be the space of one-dimensional polynomials of degree lower or equal to J, and let $\mathbb{Q}_J$ be the space of two-dimensional tensorized polynomials

$$\mathbb{Q}_J=\operatorname{span}\left\{\varphi(x)=\ell_1(x_1)\ell_2(x_2)e_d : \ell_1,\ell_2\in\mathbb{P}_J,\ d\in\{1,2\}\right\},$$

where e_1,e_2 are the canonical basis of $\mathbb{R}^2$. Given the domain $\Omega\subset\mathbb{R}^2$, we define the non-overlapping partition $\{\Omega_q\}_{q=1}^{N_{\text{dd}}}$ such that each element is isomorphic to the unit square; we denote by $\Psi_1,\ldots,\Psi_{N_{\text{dd}}}$ Gordon-Hall maps from $\widehat{\Omega}$ to $\Omega_1,\ldots,\Omega_{N_{\text{dd}}}$, respectively: we recall that Gordon-Hall maps are uniquely defined based on the parameterizations of the partition interfaces. To provide a concrete example, for the problem in section 4.2.3 we consider the partition depicted in Figure 22.

Given the set of polynomials $\overrightarrow{\varphi}:=[\varphi_1,\ldots,\varphi_{N_{\text{dd}}}]\in\bigotimes_{q=1}^{N_{\text{dd}}}\mathbb{Q}_J$, we define

$$\widetilde{\mathcal{N}}(x;\overrightarrow{\varphi}) := \sum_{q=1}^{N_{\text{dd}}}\Psi_q\circ\Phi_q\circ\Psi_q^{-1}(x)\mathbb{1}_{\Omega_q}(x)\ \text{ where }\ \Phi_q(x)=x+\varphi_q(x)\ q=1,\ldots,N_{\text{dd}}. \tag{34}$$

It is possible to verify that the space

$$\mathcal{W}_0=\left\{\overrightarrow{\varphi}=[\varphi_1,\ldots,\varphi_{N_{\text{dd}}}],\in\bigotimes_{q=1}^{N_{\text{dd}}}\mathbb{Q}_J\ :\ \varphi_q\cdot n|_{\partial\widehat{\Omega}}=0,\ \widetilde{\mathcal{N}}(\cdot;\overrightarrow{\varphi})\in C(\Omega)\right\} \tag{35}$$

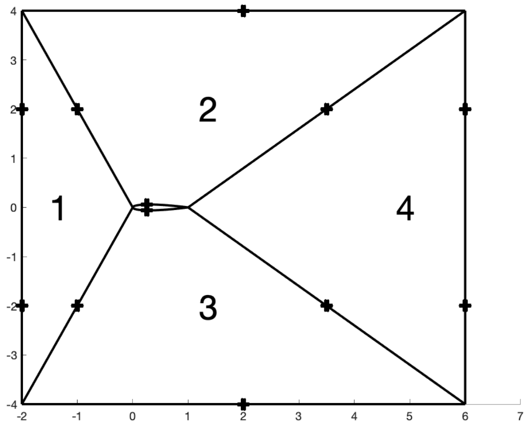


Figure 22: partition considered for the model problem in section 4.2.3.

is a linear space of size $M < 2(J+1)^2 N_{\mathrm{dd}}$; we denote by $\{\overrightarrow{\varphi}_m\}_{m=1}^M$ a basis of $\mathcal{W}_0$. Finally, we define the approximation map

$$\mathcal{N}\left(x; \overrightarrow{\varphi}\right) := \widetilde{\mathcal{N}}\left(x; \sum_{m=1}^{M} (\mathbf{a})_m \overrightarrow{\varphi}_m\right), \quad x \in \Omega, \quad \mathbf{a} \in \mathbb{R}^M. \tag{36}$$

References

[1] N. A. C. F. Aeronautics. Equations, tables, and charts for compressible flow. *Report*, 1135, 1953.

[2] G. I. Barenblatt. *Scaling, self-similarity, and intermediate asymptotics: dimensional analysis and intermediate asymptotics*. Number 14. Cambridge University Press, 1996.

[3] G. Berkooz, P. Holmes, and J. L. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual review of fluid mechanics*, 25(1):539–575, 1993.

[4] F. Bernard, A. Iollo, and S. Riffaud. Reduced-order model for the BGK equation based on POD and optimal transport. *Journal of Computational Physics*, 373:545–570, 2018.

[5] S. Bhatnagar, Y. Afshar, S. Pan, K. Duraisamy, and S. Kaushik. Prediction of aerodynamic flow fields using convolutional neural networks. *Computational Mechanics*, 64(2):525–545, 2019.

[6] Y. Brenier. Polar decomposition and increasing rearrangement of vector-fields. *Comptes rendus de l'academie des sciences serie I - Mathematique*, 305(19):805–808, 1987.

[7] S. L. Brunton, B. R. Noack, and P. Koumoutsakos. Machine learning for fluid mechanics. *Annual Review of Fluid Mechanics*, 52:477–508, 2020.

[8] Y. Chen, T. T. Georgiou, and A. Tannenbaum. Optimal transport for gaussian mixture models. *IEEE Access*, 7:6269–6278, 2018.

[9] R. R. Craig Jr and M. C. Bampton. Coupling of substructures for dynamic analyses. *AIAA journal*, 6(7):1313–1319, 1968.

[10] V. Ehrlacher, D. Lombardi, O. Mula, and F.-X. Vialard. Nonlinear model reduction on metric spaces. Application to one-dimensional conservative PDEs in Wasserstein spaces. *ESAIM. Mathematical Modelling and Numerical Analysis*, 54, 2020.

[11] A. Ferrero and D. D'Ambrosio. An hybrid numerical flux for supersonic flows with application to rocket nozzles. In *AIP Conference Proceedings*, volume 2293, page 200004. AIP Publishing LLC, 2020.

[12] S. Fresca, A. Manzoni, L. Dedè, and A. Quarteroni. Deep learning-based reduced order models in cardiac electrophysiology. *PloS one*, 15(10):e0239416, 2020.

[13] G. Haller. Lagrangian coherent structures. *Annual Review of Fluid Mechanics*, 47:137–162, 2015.

[14] J. S. Hesthaven, G. Rozza, and B. Stamm. *Certified reduced basis methods for parametrized partial differential equations*. Springer, 2016.

[15] D. B. P. Huynh, D. J. Knezevic, and A. T. Patera. A static condensation reduced basis element method: approximation and a posteriori error estimation. *ESAIM: Mathematical Modelling and Numerical Analysis*, 47(1):213–251, 2013.

[16] A. Iollo and D. Lombardi. Advection modes by optimal mass transfer. *Physical Review E*, 89(2):022923, 2014.

[17] L. V. Kantorovich. On the translocation of masses. In *Dokl. Akad. Nauk. USSR (NS)*, volume 37, pages 199–201, 1942.

[18] Y. Kim, Y. Choi, D. Widemann, and T. Zohdi. Efficient nonlinear manifold reduced order model. *arXiv preprint arXiv:2011.07727*, 2020.

[19] L. D. Landau and E. M. Lifshitz. *Course of theoretical physics VI: fluid mechanics*. Elsevier, 2013.

[20] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

[21] K. Lee and K. T. Carlberg. Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders. *Journal of Computational Physics*, 404:108973, 2020.

[22] J. L. Lumley. The structure of inhomogeneous turbulent flows. *Atmospheric turbulence and radio wave propagation*, 1967.

[23] R. J. McCann. A convexity principle for interacting gases. *Advances in mathematics*, 128(1):153–179, 1997.

[24] R. J. McCann et al. Existence and uniqueness of monotone measure-preserving maps. *Duke Mathematical Journal*, 80(2):309–324, 1995.

[25] G. J. McLachlan, S. X. Lee, and S. I. Rathnayake. Finite mixture models. *Annual review of statistics and its application*, 6:355–378, 2019.

[26] D. Modesti and S. Pirozzoli. A low-dissipative solver for turbulent compressible flows on unstructured meshes, with openfoam implementation. *Computers & Fluids*, 152:14–23, 2017.

[27] R. Mojgani and M. Balajewicz. Arbitrary lagrangian eulerian framework for efficient projection-based reduction of convection dominated nonlinear flows. In *APS Division of Fluid Dynamics Meeting Abstracts*, pages M1–008, 2017.

[28] R. Mojgani and M. Balajewicz. Physics-aware registration based auto-encoder for convection dominated pdes. *arXiv preprint arXiv:2006.15655*, 2020.

[29] M. L. Munjal. *Acoustics of ducts and mufflers with application to exhaust and ventilation system design*. John Wiley & Sons, 1987.

[30] F. Nicoud and F. Ducros. Subgrid-scale stress modelling based on the square of the velocity gradient tensor. *Flow, turbulence and Combustion*, 62(3):183–200, 1999.

[31] M. Ohlberger and S. Rave. Nonlinear reduced basis approximation of parameterized evolution equations via the method of freezing. *Comptes Rendus Mathematique*, 351(23-24):901–906, 2013.

[32] T. Peacock and G. Haller. Lagrangian coherent structures: The hidden skeleton of fluid flows. *Physics today*, 66(2):41, 2013.

[33] P.-O. Persson and J. Peraire. Sub-cell shock capturing for discontinuous Galerkin methods. In *44th AIAA Aerospace Sciences Meeting and Exhibit*, page 112, 2006.

[34] G. Peyré, M. Cuturi, et al. Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.

[35] A. Pinkus. *N-widths in Approximation Theory*, volume 7. Springer Science & Business Media, 2012.

[36] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced basis methods for partial differential equations: an introduction*, volume 92. Springer, 2015.

[37] J. A. Rice. *Mathematical statistics and data analysis*. Cengage Learning, 2006.

[38] S. Riffaud. *Reduced-order models: convergence between scientific computing and data for fluid mechanics.* PhD thesis, Université de Bordeaux, 2020.

[39] S. Riffaud, M. Bergmann, C. Farhat, S. Grimberg, and A. Iollo. The DGDD method for reduced-order modeling of conservation laws. *Journal of Computational Physics*, 437:110336, 2021.

[40] T. Taddei. A registration method for model order reduction: data compression and geometry reduction. *SIAM Journal on Scientific Computing*, 42(2):A997–A1027, 2020.

[41] T. Taddei and L. Zhang. Registration-based model reduction in complex two-dimensional geometries. *Journal of Scientific Computing*, 88(3):79, 2021.

[42] T. Taddei and L. Zhang. Space-time registration-based model reduction of parameterized one-dimensional hyperbolic pdes. *ESAIM: M2AN*, 55(1):99–130, 2021.

[43] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics: a practical introduction.* Springer Science & Business Media, 2013.

[44] C. Villani. *Topics in optimal transportation.* Number 58. American Mathematical Soc., 2003.

[45] K. Willcox and J. Peraire. Balanced model reduction via the proper orthogonal decomposition. *AIAA journal*, 40(11):2323–2330, 2002.