Automatica 171 (2025) 111916

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Robust global optimization on smooth compact manifolds via hybrid gradient-free dynamics☆

Daniel E. Ochoa *, Jorge I. Poveda

Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA, 92161, United States of America

ARTICLE INFO

Article history:
Received 27 November 2022
Received in revised form 12 March 2024
Accepted 31 July 2024
Available online xxxx

ABSTRACT

It is well known that smooth autonomous dynamical systems modeled by ordinary differential equations (ODEs) cannot robustly and globally stabilize a point on compact, boundaryless manifolds. This obstruction, which is topological in nature, has significant implications for optimization problems, rendering traditional continuous-time algorithms incapable of robustly solving *global* optimization problems in such spaces. In turn, *gradient-free* optimization algorithms, which usually inherit their stability and convergence properties from their gradient-based counterparts, can also suffer from similar topological obstructions. For instance, this is the case in zeroth-order methods and perturbation-based techniques, where gradients and Hessian matrices are usually estimated in real-time via measurements or evaluations of the cost function. To address this problem, in this paper we introduce a novel class of *hybrid gradient-free optimization dynamics* that combine continuous-time and discrete-time feedback to overcome the obstructions that emerge in traditional ODE-based optimization algorithms evolving on smooth compact manifolds. The proposed hybrid dynamics switch between different gradient-free feedback-laws obtained by applying suitable exploratory *geodesic dithers* to a family of synergistic diffeomorphisms adapted to the cost function that defines the optimization problem. The use of geodesic dithers enables a suitable exploration of the manifold while simultaneously preserving its forward invariance, a property that is fundamental for many practical applications with physics-based constraints. The hybrid dynamics exploit the information obtained from the geodesic dithers to achieve robust global practical stability of the set of minimizers of the cost function. This stabilization is achieved without having direct access to the gradients of the cost functions, but rather using only real-time and continuous evaluations of the cost. Examples and numerical results are presented to illustrate the main ideas and advantages of the method.

1. Introduction

This paper studies algorithms for the *global* solution of optimization problems of the form

$$\min \ \phi(z) \quad \text{subject to} \quad z \in M, \tag{1}$$

where ϕ is a smooth cost function and (M, g) is an n-dimensional Riemannian manifold to be formally defined in Section 2. The mathematical form of ϕ and its derivatives is assumed to be *unknown*. It is only assumed that ϕ is available through *measurements* or *evaluations* on M. This class of problems arises in various practical applications, spanning from aerospace engineering (Hauser, 2002) to power systems (Absil, Mahony, & Sepulchre, 2009) and quantum control (Grivopoulos & Bamieh, 2003). One of the simplest and most successful algorithms for optimization is the gradient-descent method, which has been studied in the context of manifolds since at least the end of the last century (Gabay, 1982). Recently, these methods have gained considerable interest due to their potential applications in estimation, machine learning, and data science pipelines (Bottou, 2010). In the context of dynamical systems described by ordinary differential equations (ODEs), real-time optimization problems defined on a manifold M are common in robotics, mechanical systems, and aerospace control problems evolving under kinematic constraints. For example, controlling unicycles (Sontag, 1999, Sec. 2.2) or navigating in obstacle-occluded spaces (Poveda, Benosman, Teel, & Sanfelice, 2021). In such problems, the restriction to evolve on M limits the feasible directions that any onboard algorithm can exploit in real-time. For comprehensive introductions to ODE-based optimization algorithms on manifolds, we refer the reader to Absil et al. (2009), Helmke and Moore (2012).

☆ This work was supported in part by NSF ECCS CAREER 2305756 and AFOSR YIP: FA9550-22-1-0211. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Martin Guay under the direction of Editor Miroslav Krstic.

* Corresponding author.
E-mail addresses: dochoatamayo@ucsd.edu (D.E. Ochoa), poveda@ucsd.edu (J.I. Poveda).

https://doi.org/10.1016/j.automatica.2024.111916

One of the primary challenges in solving optimization problems on smooth (boundaryless) compact manifolds stems from the fact that in such spaces, a point cannot be *robustly globally* asymptotically stabilized using continuous feedback in ordinary differential equations (ODEs) (Bhat & Bernstein, 2000, Thm. 1). This result extends to compact Lie groups and non-contractible spaces in general, as shown in Sontag (2013, Thm. 21). This well-known property implies that standard gradient flows or Newton-like flows cannot achieve robust *global* convergence to the minimizer of a continuously differentiable cost function for every type of smooth compact manifold. The reason behind this incompatibility lies in the fundamental mismatch between the topological nature of the basin of attraction of a point under continuous dynamics, and the topological properties of a compact boundaryless manifold (Sontag, 2013, Thm. 21). Specifically, the basin of attraction of a point under continuous feedback is contractible, while a compact manifold is not. Many results in the literature overcome this issue by focusing on asymptotic stability properties that overlook measure-zero sets containing the critical points of the cost function that are not solutions to the optimization problem under study (Angeli, 2004; Efimov, 2012), such as local maximizers and saddle points. However, algorithms with *almost* global convergence certificates have been shown to be susceptible to arbitrarily small (adversarial) disturbances. Under such disturbances, the set of problematic initial conditions from which convergence is not achieved is not of measure zero anymore, but rather an open set. Examples illustrating this susceptibility can be found in Sontag (1999), Poveda et al. (2021, Ex. 1), and Mayhew and Teel (2011b).

Alternatively, other works have circumvented the obstruction via time-varying (Coron, 1992), or discontinuous feedback (Malisoff, Krichman, & Sontag, 2006), finding success in achieving global convergence in certain applications. However, as shown in Mayhew and Teel (2011a, Cor. 21), time-varying approaches can only circumvent the issue when the optimization dynamics operate in nominal conditions. In particular, when the system is subject to (even arbitrarily) small disturbances, *robust* and global stabilization of a point in compact boundaryless manifolds cannot be achieved by merely using discontinuous or time-varying feedback strategies. To address this issue, in Mayhew (2010) the authors introduced a *hybrid* controller that synergistically switches between different continuous vector fields, generated from a family of potential functions, to globally stabilize a point. Recent works have employed the synergistic framework to solve attitude stabilization problems in SO(3) (Berkane, Abdessameud, & Tayebi, 2017), stabilization by hybrid backstepping (Casau, Sanfelice, & Silvestre, 2019), and for the robust stabilization of trajectories in multi-rotor aerial vehicles (Casau, Mayhew, Sanfelice, & Silvestre, 2019). However, since these works address stabilization problems, where the point to be stabilized is known *a priori*, in general, they cannot be directly used for the solution of optimization problems where the set of optimizers is unknown, or in cases where the potential functions are only accessible via measurements or evaluations.

Optimization problems where the cost function is unknown and only accessible through measurements or evaluations are common across many applications. These problems have traditionally been studied using gradient-free methods, such as *zeroth-order* optimization algorithms. While the literature of continuous-time zeroth-order optimization dynamics, also known as extremum seeking, is quite rich (Dürr, Stanković, Johansson, & Ebenbauer, 2014; Krstić & Wang, 2000; Poveda & Teel, 2017), most of the algorithms applicable to smooth compact manifolds are characterized by smooth gradient-free dynamical systems that aim to emulate, via averaging or other "approximation" technique, the behavior of a target gradient-flow on the manifold (Dürr et al., 2014; Taringoo, Dower, Nesic, & Tan, 2018). In these settings, the stability properties of gradient-free dynamics are usually inherited from the stability properties of the target system being approximated. Therefore, the challenges of robust global optimization extend to the gradient-free counterparts whenever the target system is characterized by a smooth ODE. Moreover, existing results that achieve global optimization via switching algorithms (Strizic, Poveda, & Teel, 2017) do not necessarily preserve the forward invariance of the manifold due to the use of dither signals that do not evolve in the manifold's tangent space, a requirement that is relevant for practical applications where the evolution on manifolds is enforced by physical constraints, or in problems where the cost function is defined only on the manifold.

To address the above challenges and limitations of existing approaches, the main contribution of this paper is the introduction of a novel class of gradient-free algorithms for the *global* solution of optimization problems defined on compact boundaryless connected Riemannian manifolds. The algorithms are characterized by a family of hybrid gradient-free dynamics that switch between different zeroth-order feedback laws that implement exploratory geodesic dithers to extract suitable "descent directions" from the cost function ϕ. The switches in the algorithms are implemented in both the exploration and the exploitation components of the dynamics. In particular, to globally navigate and explore manifolds that are not parallelizable (e.g., $\mathbb{S}^2$), the exploratory geodesic dithers switch between different local frames using a hysteresis-based mechanism. To achieve *global* convergence to (a neighborhood of) the set of minimizers, the algorithms implement a class of switching diffeomorphisms adapted to the cost function of interest. Such diffeomorphisms can be constructed under mild qualitative assumptions on the cost functions and for different types of manifolds. Our main result establishes *robust global* practical asymptotic stability of the set of minimizers of the cost function for the proposed hybrid gradient-free dynamics. Compared to previous approaches for gradient-free optimization on manifolds, e.g. Dürr et al. (2014), Taringoo et al. (2018) and Suttner (2022), our convergence results are global rather than local or almost global. Compared to existing switching algorithms (Strizic et al., 2017), our gradient-free dynamics are designed to evolve on the manifold and preserve its invariance via geodesic dithering. The results presented in this paper are also applicable to a larger class of manifolds and optimization problems. Our results also provide an alternative approach to the solution of gradient-free optimization and extremum seeking problems with multiple critical points, typical in non-convex settings, a problem that has also been recently studied in Suttner and Krstić (2023) using other techniques.

The rest of this paper is organized as follows. Section 2 presents the preliminaries. Section 3 presents the main results, including the general hybrid gradient-free optimization dynamics and three specific examples of algorithms synthesized for different applications. Section 4 presents the proofs, and Section 5 ends with the conclusions.

2. Preliminaries

In this section, we introduce the notation used in the paper, as well as some mathematical preliminaries.

2.1. Notation

Given a compact set $\mathcal{A} \subset N$ in a metric space N, with metric $d : N \times N \to \mathbb{R}_{\geq 0}$, and an element $z \in N$, we use $|z|_{\mathcal{A}} := \min_{s \in \mathcal{A}} d(z, s)$ to denote the minimum distance of z to $\mathcal{A}$. We use $\mathbb{S}^n := \{z \in \mathbb{R}^{n+1} : \sum_{i=1}^{n+1} z_i^2 = 1\}$ to denote the nth dimensional sphere, with $\mathbb{S}^1$ representing the unit circle in $\mathbb{R}^2$.

We use $\mathbb{T}^n = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ to denote the nth Cartesian product of $\mathbb{S}^1$. We also use $r\mathbb{B}$ to denote a closed ball in the Euclidean space, of radius $r > 0$ and centered at the origin. We use $I_n \in \mathbb{R}^{n\times n}$ for the identity matrix, and $\mathbf{1}_A$ for the indicator function of the set A. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class $\mathcal{KL}$ if it is non-decreasing in its first argument, non-increasing in its second argument, $\lim_{r\to 0^+} \beta(r,s) = 0$ for each $s \in \mathbb{R}_{\geq 0}$, and $\lim_{s\to\infty} \beta(r,s) = 0$ for each $r \in \mathbb{R}_{\geq 0}$. We use $\pi_A : A \times B \to A$ to denote the natural projection from $A \times B$ to A, and gph J to denote the graph of a mapping J. The Kronecker delta is denoted as δ_{ij}.

2.2. Riemannian manifolds

We introduce the main differential geometric concepts used in the paper. For more details, we refer the reader to Lee (2013, 2018). The concept of smooth manifold will play an important role in this paper:

Smooth manifolds: An n-dimensional manifold is a second-countable Hausdorff topological space that is locally Euclidean of dimension n. A coordinate chart for M is a pair (U,φ) where $U \subset M$ is an open set and $\varphi : U \to \hat{U} \subset \mathbb{R}^n$ is a homeomorphism. Two coordinate charts (U,φ) and (V,ψ) are said to be smoothly compatible if the transitions maps $\psi \circ \varphi^{-1}$ and $\varphi \circ \psi^{-1}$ are diffeomorphisms. A smooth structure on M is a maximal collection of coordinate charts for which any two charts are smoothly compatible; a smooth coordinate chart is any chart that belongs to a smooth structure. Then, a *smooth manifold* is a manifold endowed with a particular smooth structure. Given a smooth manifold M, the set of all smooth real-valued functions $f : M \to \mathbb{R}$ is denoted by $C^\infty(M)$.

Tangent space and Vector Fields: Dynamical systems evolving on smooth manifolds are defined by vector fields that lie within their tangent spaces. For each $z \in M$, a tangent vector at z is a linear map $v : C^\infty(M) \to \mathbb{R}$ that satisfies $v(fh) = f(z)\cdot v(h) + h(z)\cdot v(f)$, for $f, h \in C^\infty(M)$. The set of all tangent vectors at z is denoted by T_zM and is called the *tangent space* of M at z. The tangent bundle TM is defined to be the disjoint union of the tangent spaces at all points in the manifold, i.e., $TM := \bigsqcup_{z\in M} T_zM$. A *smooth vector field* is a smooth map $X : M \to TM$ satisfying $X(z) \in T_zM$ for all $z \in M$. We use $\mathfrak{X}(M)$ to denote the set of all smooth vector fields on M.

The differential of a function $f \in C^\infty(M)$, denoted by $df : TM \to \mathbb{R}$, is a map defined pointwise by:

$$df_z(v) = v(f), \qquad \forall\, v \in T_zM. \tag{2}$$

Using the differential, we define the sets of *critical points and critical values* of $f \in C^\infty(M)$ as follows:

$$\text{Crit}\, f := \{z \in M \;:\; df_z = 0\}, \tag{3}$$

$$\text{Val}\, f := \{a \in \mathbb{R} \;:\; a = f(z),\ z \in \text{Crit}\, f\}. \tag{4}$$

A local frame for M is defined as a tuple of vector fields $(X_1,\ldots,X_n)$ defined on an open set $U \subset M$, that is linearly independent and spans T_zM at each $z \in M$. If this frame is defined in the entire manifold ($U = M$), it is called a global frame. When M admits a global frame, the manifold is said to be *parallelizable*. Parallelizability will play an important role in our algorithms.

Riemannian Manifolds: In this paper, we will focus on Riemannian manifolds. An n-dimensional *Riemannian manifold* is a pair (M,g), where M is an n-dimensional smooth manifold, and g is a Riemannian metric whose value at each point $z \in M$ is an inner product defined on T_zM. The Riemannian metric g enables the definition of the gradient of f, grad $f : M \to TM$, as the continuous map satisfying:

$$df_z(v) = g\,(\text{grad}\, f|_z, v), \quad \text{for all } z \in M, v \in T_zM, \tag{5}$$

where grad $f|_z \in T_zM$ denotes the value of the gradient of f at z.

To guarantee a suitable exploration of M, while preserving its invariance, we will work with algorithms that implement geodesic dithers:

Geodesics: Geodesics are defined as curves $\gamma : [a,b] \to M$ on a Riemannian manifold, satisfying

$$\nabla_{\dot\gamma(t)}\dot\gamma(t) = 0, \tag{6}$$

where $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ is the Levi-Civita connection (Lee, 2018, Ch. 5). To generate the dither signals used by the gradient-free optimization algorithms considered in this paper, we use the *restricted exponential map* $\exp_z : T_zM \to M$, defined by $\exp_z(v) = \gamma_v(1)$, where γ_v is the unique maximal geodesic satisfying $\gamma_v(0) = z$ and $\dot\gamma_v(0) = v$.

Throughout the paper, we make use of the following standing assumption.

Standing Assumption 2.1. *The Riemannian manifold (M,g) is compact, boundaryless, and connected.* □

In particular, Assumption 2.1 guarantees the existence of a path between any two points in M (Lee, 2018, Prop 2.50), which facilitates the definition of a notion of distance.

Riemannian Distance: The Riemannian distance, denoted by $d_g(z_1,z_2)$ is defined to be the infimum of the lengths of all admissible curves between a pair of points in the manifold (Lee, 2013, Ch 2.). Formally, the Riemannian distance $d_g : M \times M \to \mathbb{R}_{\geq 0}$ is defined by $d_g(z_1,z_2) := \inf_{\gamma\in\mathbb{A}(z_1,z_2)} \int_{t_1}^{t_2} \sqrt{g\,(\dot\gamma(t),\dot\gamma(t))}dt$, where $\mathbb{A}(z_1,z_2)$ represents the set of all admissible curves connecting z_1 and z_2, and $t_1, t_2 \in \mathbb{R}$ are such that $\gamma(t_1) = z_1$ and $\gamma(t_2) = z_2$ for $\gamma \in \mathbb{A}(z_1,z_2)$.

2.3. Hybrid dynamical systems and stability notions

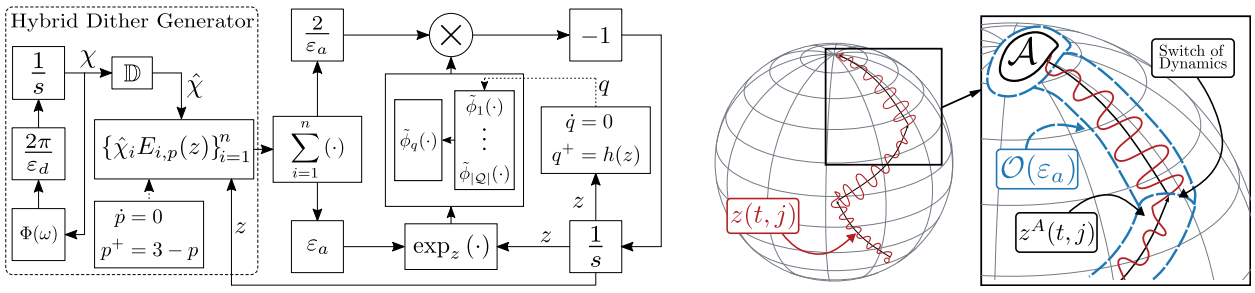
In this paper, we consider algorithms modeled as hybrid dynamical systems (HDS) (Goebel, Sanfelice, & Teel, 2012) of the form:

$$x \in C, \qquad \dot x = F(x) \tag{7a}$$

$$x \in D, \qquad x^+ \in G(x), \tag{7b}$$

where $x \in M \subset \mathbb{R}^p$ is the state, $F : M \to TM$ is called the flow map, and $G : M \rightrightarrows M$ is a set-valued map called the jump map. The sets C and D, called the flow set and the jump set, respectively, characterize the points in M where the system can flow or jump via Eqs. (7a) or (7b), respectively. Then, the HDS $\mathcal{H}$ is defined as the tuple $\mathcal{H} := \{C,F,D,G\}$. Systems of the form (7) generalize purely continuous-time systems and purely discrete-time systems. Namely, continuous-time dynamical systems (e.g., ODEs) can be seen as a HDS of the form (7) with $D = \emptyset$, while discrete-time dynamical systems (e.g. recursions) correspond to the case when $C = \emptyset$. Solutions to HDS of the form (7) are defined on hybrid time domains, i.e., they are parameterized by both a continuous-time index $t \in \mathbb{R}_{\geq 0}$, and a discrete-time index $j \in \mathbb{Z}_{\geq 0}$. Consequently, the notation $\dot x$ in (7a) represents the derivative of x with respect to time t, i.e., $\frac{dx(t,j)}{dt}$; and x^+ in (7b) represents the value of x after an instantaneous jump, i.e., $x(t,j+1)$. For a precise definition of hybrid time domains and solutions to HDS of the form (7) we refer the reader to Goebel et al. (2012, Ch.2). A HDS $\mathcal{H}$ is said to be well-posed if C and D are closed sets, $C \subset \text{dom}(F)$ and $D \subset \text{dom}(G)$, F is continuous in C, and G is outer-semicontinuous (Goebel et al., 2012, Def. 5.9) and locally bounded (Goebel et al., 2012, Def. 5.14) relative to D.

Stability notions: By endowing the manifold with the distance function d_g, M constitutes a metric space (Lee, 2018, Thm 2.55). Accordingly, we can use stability notions analogous to those studied in the Euclidean space.



(a) Block diagram of the Zeroth-Order Hybrid Dynamics $\mathcal{H}_0$. (b) Example of a trajectory z and its average z^A.

Fig. 1. Left: Block diagram of the proposed hybrid zeroth-order dynamics with geodesic dithering. Right: Cartoon of the trajectories of the system evolving on a manifold M.

Definition 2.1. The compact set $\mathcal{A} \subset C \cup D$ is said to be *uniformly globally asymptotically stable* (UGAS) for (7) if $\exists\, \beta \in \mathcal{KL}$ such that for all solutions x:

$$|x(t,j)|_{\mathcal{A}} \leq \beta(|x(0,0)|_{\mathcal{A}}, t+j), \tag{8}$$

$\forall\, (t,j) \in \text{dom}(x)$, where $|z|_{\mathcal{A}} = \min_{s\in\mathcal{A}} d_g(z,s)$. □

We also consider ε-parameterized HDS $\mathcal{H}_\varepsilon$ of the form

$$x \in C_\varepsilon, \quad \dot{x} = F_\varepsilon(x), \quad \text{and} \quad x \in D_\varepsilon, \ x^+ \in G_\varepsilon(x),$$

where $\varepsilon > 0$. For these systems, we will study *global practical stability* properties as $\varepsilon \to 0^+$.

Definition 2.2. The compact set $\mathcal{A} \subset C \cup D$ is said to be *Globally Practically Asymptotically Stable* (GP-AS) as $\varepsilon \to 0^+$ for system (7) if $\exists\, \beta \in \mathcal{KL}$ such that for each $\nu > 0$ there exists $\varepsilon^* > 0$ such that for all $\varepsilon \in (0, \varepsilon^*)$ and $x(0,0) \in M$, every solution of $\mathcal{H}_\varepsilon$ satisfies

$$|x(t,j)|_{\mathcal{A}} \leq \beta(|x(0,0)|_{\mathcal{A}}, t+j) + \nu, \tag{9}$$

$\forall\, (t,j) \in \text{dom}(x)$. □

The notion of GP-AS can be extended to systems that depend on two parameters $\varepsilon = (\varepsilon_1, \varepsilon_2)$. In this case, we say that $\mathcal{A}$ is GP-AS as $(\varepsilon_2, \varepsilon_1) \to 0^+$ where the parameters are tuned in order starting from ε_1.

3. Main results

Approaches for optimization in Euclidean spaces with global convergence certificates usually rely on convexity properties of ϕ. For Riemannian manifolds, convexity is characterized along geodesics. However, under Assumption 2.1 geodesic convexity has little utility since in compact Riemannian manifolds geodesically convex functions are necessarily constant (Udriste, 2013, Cor. 2.5). Given the limitations imposed by convexity in compact Riemannian manifolds, in this paper we alternatively rely on the following regularity assumption on ϕ, which is closely related to the decomposability of invariant sets introduced in Angeli and Efimov (2015, Assumption 1).

Standing Assumption 3.1. *The cost function ϕ has a finite amount of critical values, i.e., there exists $l \in \mathbb{N}$ such that Val $\phi = \{\phi_1, \phi_2, \ldots, \phi_l\}$, where $\underline{\phi} := \phi_1 < \phi_2 \leq \phi_i \leq \phi_l =: \overline{\phi}$, and $\underline{\phi} \leq \phi(z) \leq \overline{\phi}$ for all $z \in M$. Moreover, the critical points of ϕ are isolated, and ϕ has a unique minimizer.* □

Let $\mathcal{A} := \{z \in \text{Crit}\ \phi \ :\ \phi(z) = \underline{\phi}\}$ represent the minimizer of ϕ and define $\mathcal{B} := \text{Crit}\ \phi \setminus \mathcal{A}$. Since M is compact, the set $\mathcal{A}$ is also compact. Note that Assumption 3.1 does not rule out functions ϕ with multiple critical points. Indeed, in our problem setup, $\mathcal{B}$ is not empty since, by Morse theory, there exist at least two critical points for scalar-valued functions on compact boundaryless manifolds. Such critical points correspond to equilibria in traditional gradient flows, rendering them highly susceptible to even small (potentially adversarial) disturbances. This robustness issue, thoroughly discussed in Sontag (1999), Poveda et al. (2021, Ex. 1), and Mayhew and Teel (2011b), and illustrated later in Section 3.4 via numerical examples, is one of the main motivations for the development of robust hybrid algorithms. In our case, we design the hybrid algorithms to be gradient-free by leveraging tools from averaging theory for hybrid dynamical systems.

Remark 3.1. For the case when ϕ is a Morse function (Milnor, 2015, Definition 2.3), Assumption 3.1 is automatically satisfied. Moreover, since the set of Morse functions is an open dense set in the space of differentiable functions (Milnor, 2015, Theorem 2.7), we can dispense with Standing Assumption 3.1 by considering a surrogate approximate optimization problem to (1), whose solution is the minimizer of a Morse function sufficiently close to ϕ. □

Remark 3.2. When the set of minimizers $\mathcal{A}$ forms a submanifold rather than a singleton in M, the basin of attraction is diffeomorphic to a tubular neighborhood of $\mathcal{A}$ in M (Mayhew & Teel, 2011a, Cor. 21). This neighborhood may or may not be contractible. In this case, to assess the applicability of our approach, further assumptions regarding the topological characteristics of $\mathcal{A}$ and its tubular neighborhood are required. To simplify our presentation, we defer this problem to future research. □

3.1. Description of the proposed algorithms

To solve problem (1), the left plot of Fig. 1 shows a block diagram of the proposed dynamics. Before analyzing the mathematical properties of this system, we first briefly describe the main ideas behind the algorithms:

(a) A set of dynamic oscillators, with state χ and frequency proportional to $1/\varepsilon_d$, where $\varepsilon_d > 0$ is a small tunable parameter, is employed to generate exploratory signals defined in $\mathbb{T}^n$. The signals are then suitably combined with a local orthonormal frame $\{E_{i,p}\}_{i=1}^n$, $p \in \mathcal{P} \subset \mathbb{Z}_{\geq 1}$, to obtain a dithering vector field $\mathcal{D}_p$ that drives dithering geodesics along the manifold M. These geodesic dithers will be used for the purpose of *local (real-time) exploration*.

(b) To ensure a well-defined local exploration around every point $z \in M$ for all time, we introduce a logic state p. This state selects an orthonormal frame $\{E_{i,p}\}_{i=1}^n$ that locally spans the tangent space at a given point z. This logic state is updated using a *hybrid exploration supervisor* that hysterically switches between local frames. When the manifold is *parallelizable*, we can dispense with this logic state and its associated hybrid dynamics.

(c) The geodesic dithers, together with measurements or evaluations of the cost ϕ, are used to generate families of vector fields $\{\hat{f}_{q,p}(\cdot,\chi)\}_{q\in\mathcal{Q}}$, $p\in\mathcal{P}$, given by

$$\hat{f}_{q,p}(z,\chi) := \frac{2}{\varepsilon_a}\tilde{\phi}_q\left(\exp_z\left(\varepsilon_a\mathcal{D}_p(z)\right)\right)\mathcal{D}_p(z), \tag{10}$$

where $\varepsilon_a>0$ is a tunable gain and $\mathcal{Q}\subset\mathbb{Z}_{\geq 1}$. These vector fields, explained below, are used for the purpose of *exploitation* in the optimization dynamics.

(d) To define the vector fields $\{\hat{f}_{q,p}(\cdot,\chi)\}_{q\in\mathcal{Q}}$, we use a set of diffeomorphisms and generate a family of surrogate warped cost functions $\{\tilde{\phi}_q\}_{q\in\mathcal{Q}}$. The chosen diffeomorphisms shift the points that are not in a neighborhood of the minimizers of ϕ. In this manner, by appropriately *partitioning* the manifold M, for each $q\in\mathcal{Q}$ we can implement the vector field $\hat{f}_{q,p}(\cdot,\chi)$ in a "safe zone" where its average dynamics have no critical points other than $\mathcal{A}$. A *hybrid exploitation supervisor* is then used to switch the logic state q to globally steer the state z to $\mathcal{A}$. These partitions can be constructed under mild qualitative assumptions on the cost function.

(f) As we increase the frequency of the dithers (i.e., $\varepsilon_d\to 0^+$), the trajectories induced by the switching vector fields (10) will approximate the trajectories of a class of hybrid gradient flows that will be shown to achieve *robust global* asymptotic stability of $\mathcal{A}$ on M.

The above ideas suggest that the proposed algorithms are similar in spirit to synergistic hybrid controllers studied in the context of robust global *stabilization* problems Mayhew (2010), (Sanfelice, 2020, Ch. 7). However, the algorithms studied in this paper do not exactly fit the setting of synergistic hybrid control, since the family $\{\hat{f}_{q,p}(\cdot,\chi)\}_{q\in\mathcal{Q}}$ does not describe gradients of synergistic Lyapunov functions. In fact, unlike standard stabilization problems tackled via hybrid control, the main challenges in problem (1) are that the set $\mathcal{A}$ and the function ϕ are unknown. Therefore, to implement the gradient-free hybrid dynamics we need to characterize the family of cost functions ϕ and smooth manifolds (M,g) that admit suitable partitions and deformations to generate feasible adaptive switching rules that induce global stability of $\mathcal{A}$, in a gradient-free way.

3.2. Stability, convergence, and robustness results for parallelizable manifolds

To solve problem (1), we first focus on manifolds M that are parallelizable, which enables the use of a global orthonormal frame $\{E_i\}_{i=1}^n$. This facilitates the definition of a single *dithering vector field* $\mathcal{D}: M\to TM$ as $\mathcal{D}(z) := \sum_{i=1}^n \hat{\chi}_i E_i(z)$, where $\hat{\chi}$ corresponds to the vector that stacks the odd components of χ. This single vector field will drive the dithering geodesics, ensuring global exploration of M without the need of using additional logic states (i.e., with $p\equiv 1$). The study of the non-parallelizable scenario is postponed to Section 3.6.

The closed-loop system describing the gradient-free hybrid dynamics, shown in Fig. 1(a), has three main states: $(z,\ q,\ \chi)\in M\times\mathcal{Q}\times\mathbb{T}^n$, where z is an internal auxiliary state, $q\in\mathcal{Q}:=\{1,2,\ldots,N\}$, $N\in\mathbb{Z}_{\geq 2}$, is a logic decision variable, and χ is the state of the oscillator. The data of this hybrid system is denoted as:

$$\mathcal{H}_0 = \{C_0,\ F_0,\ D_0,\ G_0\}. \tag{11}$$

In this way, the continuous-time dynamics of $\mathcal{H}_0$, with state $y:=(z,\ q,\ \chi)$ are given by

$$y\in C_0,\quad \dot{y}=F_0(y) := \begin{pmatrix} -\hat{f}_q(z,\chi)\\ 0\\ \frac{2\pi}{\varepsilon_d}\Psi(\omega)\chi\end{pmatrix}, \tag{12}$$

where $\hat{f}_q: M\times\mathbb{T}^n\to TM$ is defined via (10) by omitting the state p, and $\Psi:\mathbb{R}^n\to\mathbb{R}^{2n\times 2n}$ is given by

$$\Psi(\omega) := \begin{pmatrix} \Omega(\omega_1) & 0 & \ldots & 0\\ 0 & \Omega(\omega_2) & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & \Omega(\omega_n)\end{pmatrix},\quad \Omega(\alpha) := \begin{pmatrix} 0 & \alpha\\ -\alpha & 0\end{pmatrix},$$

where $\alpha>0$. Here, ω_i is a positive rational number, and $\varepsilon_d\in\mathbb{R}_{>0}$ and $\varepsilon_a\in\mathbb{R}_{>0}$ are tunable gains. For every $q\in\mathcal{Q}$, the vector field $\hat{f}_q(z,\chi)$ is obtained by geodesically dithering the corresponding warped cost function $\tilde{\phi}_q$ (defined below in Definition 3.1) around the current point z. In particular, the dither is obtained along a geodesic γ, originating from z with an initial velocity parameterized by the dithering amplitudes, denoted by χ.

To model the switches between different vector fields, the discrete-time dynamics G_0 of $\mathcal{H}_0$ are given by the following constrained difference inclusion

$$y\in D_0,\quad y^+\in G_0(y) := \{z\}\times h(z)\times\{\chi\}, \tag{13}$$

where the set-valued map $h: M\rightrightarrows\mathcal{Q}$, is defined as

$$h(z) := \left\{q\in\mathcal{Q}\ :\ \tilde{\phi}_q(z)=m(z)\right\}, \tag{14}$$

and $m: M\to\mathbb{R}$ is defined as:

$$m(z) := \min_{q\in\mathcal{Q}}\tilde{\phi}_q(z). \tag{15}$$

Namely, $m(z)$ is the minimum value among all the warped cost functions $\tilde{\phi}_q$ at a given point z. To compute $m(z)$, the algorithm only needs measurements or evaluations of $\tilde{\phi}_q(z)$, which preserves the gradient-free nature of the hybrid dynamics. Moreover, the minimum in (15) is well-defined since $\mathcal{Q}$ is finite, and obtaining the value of m is not computationally expensive, since the complexity scales linearly with the cardinality of $\mathcal{Q}$.

The final elements needed for the characterization of the hybrid system $\mathcal{H}_0$ are the flow and jump sets C_0 and D_0, respectively. To define these sets, and since the warping induced by the diffeomorphisms is only useful if it modifies the points that are not in a neighborhood of the minimizers, we will use a threshold parameter $\gamma\in\mathbb{R}$ characterized by the following assumption:

Standing Assumption 3.2. *There exists a known threshold number* $\gamma\in(\underline{\phi},\phi_2)$. □

Remark 3.3. Knowledge of γ does not necessarily imply a precise knowledge of the minimizer or the exact mathematical form of ϕ. Instead, Assumption 3.2 requires only a mild qualitative understanding of the values of ϕ near its minimum. Such a qualitative characterization is often available in practical scenarios where the range of ϕ is known to lie within certain broad bounds. An example of this can be found in Lauand and Meyn (2023, pp. 131), where a known lower bound on the cost function is employed to design the gain of an exploratory signal for extremum seeking control. In the particular case when $\underline{\phi}=0$, the assumption holds for any sufficiently small $\gamma>0$.

Using γ, we can characterize a *synergistic family of diffeomorphisms* for the solution of problem (1).

Definition 3.1. Let M be a smooth manifold, and suppose $\phi\in C^\infty(M)$ satisfies Assumption 3.1. A family of functions $\mathcal{S}=\{S_q\}_{q\in\mathcal{Q}}$ is said to be a *δ-gap synergistic family of diffeomorphisms adapted to ϕ* if it satisfies:

(A_1) For every $q\in\mathcal{Q}$, $S_q: M\to M$ is a diffeomorphism.
(A_2) For every $q\in\mathcal{Q}$, $\phi(z)<\gamma \implies S_q(z)=z$.

(A3) There exists $\delta \in (0, \mu(\mathcal{S}))$, where

$$\mu(\mathcal{S}) := \min_{\substack{q\in\mathcal{Q} \\ z\in \text{Crit } \tilde{\phi}_q \setminus \mathcal{A}}} \left(\tilde{\phi}_q(z) - \min_{p\in\mathcal{Q}} \tilde{\phi}_p(z) \right),$$

and the *warped cost* $\tilde{\phi}_q : M \to \mathbb{R}$ is given by $\tilde{\phi}_q := \phi \circ S_q, \ \forall\, q \in \mathcal{Q}$. □

The family of functions $\mathcal{S}$ satisfying the above properties ensures there are enough ways to distort the manifold (M, g), allowing for the distinction of critical points other than the minimizers of ϕ using only cost *measurements or evaluations*. For each distortion of (M, g), a warped cost $\tilde{\phi}_q$ can be defined, leading to a family of N different vector fields in (12). Using Definition 3.1, we state our last main standing assumption

Standing Assumption 3.3. *There exists a δ-gap synergistic family of diffeomorphisms adapted to ϕ with finite index set $\mathcal{Q}$.* □

Remark 3.4. Verifying conditions (A1)-(A3) is clearly application-dependent, and different manifolds typically result in different warped costs. However, we stress that the constructions needed to implement the hybrid dynamics do not require explicit mathematical knowledge of the cost function ϕ, but only knowledge of qualitative properties that could be verified a priori via simple tests or experiments. Particular examples of pairs $(\phi, (M, g))$ that satisfy Standing Assumptions 3.1–3.3 will be presented in Section 3.5. □

The flow sets and jump sets of the zeroth-order hybrid dynamics $\mathcal{H}_0$, given by (11), are given by

$$C_0 := \left\{(z, q, \chi) \in M \times \mathcal{Q} \times \mathbb{T}^n : \left(\tilde{\phi}_q - m\right)(z) \leq \delta\right\}$$
$$D_0 := \{(z, q, \chi) \in M \times \mathcal{Q} \times \mathbb{T}^n : \left(\tilde{\phi}_q - m\right)(z) \geq \delta\}.$$

Based on the structure of the sets (C_0, D_0), switches of q (i.e., jumps) are allowed whenever the difference $\tilde{\phi}_q(z) - m(z)$ exceeds a δ-threshold. Flows following the vector field (12) are allowed when this difference is less than or equal to δ. When the difference is exactly equal to δ, flows and jumps are both allowed. This immediately indicates that solutions of $\mathcal{H}_0$ are not unique. However, the structure of the warped cost functions $\tilde{\phi}_q$ and the jump map will prevent the occurrence of infinite consecutive jumps by inducing a hysteresis-like behavior. In this manner, whenever a solution approaches a critical point of $\tilde{\phi}_q$ outside the set of minimizers $\mathcal{A}$, the dynamics will transition to a different vector field generated from a warped cost function $\tilde{\phi}_p$ with a lower value. The existence of such a warped cost function is guaranteed by the following technical Lemma. All proofs are presented in Section 4.

Lemma 3.1. *Suppose that ϕ satisfies Assumption 3.1, and let $\mathcal{S} = \{\mathcal{S}_q\}_{q\in\mathcal{Q}}$ be a family of functions satisfying (A_1) and (A_2) in Definition 3.1. If $\mathcal{S}$ satisfies (A_3), then, for all $q \in \mathcal{Q}$ and every $z \in \text{Crit } \tilde{\phi}_q \setminus \mathcal{A}$, there exists $p \in \mathcal{Q}$ such that:*

$$\tilde{\phi}_p(z) + \delta < \tilde{\phi}_q(z). \tag{16}$$

Conversely, if for all $q \in \mathcal{Q}$ and every $z \in \text{Crit } \tilde{\phi}_q \setminus \mathcal{A}$, there exists $p \in \mathcal{Q}$ such that (16) holds, then $\mathcal{S}$ satisfies (A_3), making it a δ-gap synergistic family of diffeomorphisms adapted to ϕ. □

We can now state the first main result of the paper.

Theorem 3.2. *Assume that the manifold M is parallelizable, and consider the hybrid zeroth-order dynamics $\mathcal{H}_0$ with data (11). Let the frequencies ω_i in (12) satisfy:*

$$\omega_i \neq \omega_j, \quad \omega_i \neq 2\omega_j, \ \omega_i \neq 3\omega_j, \quad \text{for all } i \neq j. \tag{17}$$

Then, the set $\mathcal{A} \times \mathcal{Q} \times \mathbb{T}^n$ is GP-AS as $(\varepsilon_d, \varepsilon_a) \to 0^+$, and $M \times \mathcal{Q} \times \mathbb{T}^n$ is strongly forward invariant. □

The result of Theorem 3.2 establishes *global* convergence of the trajectories z of $\mathcal{H}_0$ to an arbitrarily small neighborhood of the set of minimizers $\mathcal{A}$, while simultaneously evolving on (and exploring) the manifold M. This behavior is illustrated in Fig. 1(b). To our best knowledge, Theorem 3.2 is the first result in the literature that achieves *global* bounds of the form (9) in smooth boundaryless compact Riemannian manifolds via deterministic continuous-time zeroth-order optimization algorithms.

3.3. Approximation via 1st-order hybrid dynamics

The result of Theorem 3.2 relies on using averaging theory and perturbation theory (for hybrid systems) to show that, as $(\varepsilon_d, \varepsilon_a) \to 0^+$, the trajectories of $\mathcal{H}_0$ will approximate (on compact time domains) a solution of a first-order hybrid algorithm $\mathcal{H}_1$, with state $x = (z, q)$, continuous-time dynamics given by

$$x \in C_1, \ \dot{x} = F_1(x) := \begin{pmatrix} -\sum_{i=1}^n \nabla_{E_i}\tilde{\phi}_q(z)E_i(z) \\ 0 \end{pmatrix} \tag{18}$$

discrete-time dynamics given by

$$x \in D_1, \quad x^+ \in G_1(x) = \{z\} \times h(z), \tag{19}$$

and flow set and jump set given by

$$C_1 := \{(z, q) \in M \times \mathcal{Q} \ : \ (\tilde{\phi}_q - m)(z) \leq \delta\} \tag{20a}$$
$$D_1 := \{(z, q) \in M \times \mathcal{Q} \ : \ (\tilde{\phi}_q - m)(z) \geq \delta\}. \tag{20b}$$

Since system (18)–(20) makes use of first-order information of the warped costs $\tilde{\phi}_q$ via $\nabla_{E_i}\tilde{\phi}_q E_i(z)$, we will refer to this system as the *first-order hybrid dynamics* $\mathcal{H}_1 := \{C_1, F_1, D_1, G_1\}$. In this system, for every $q \in \mathcal{Q}$, the dynamics $\dot{z}$ in (18) represents a scaled version of grad $\tilde{\phi}_q$. Similar dynamics have been studied in the literature (Taringoo et al., 2018). They differ from the coordinate representation of grad $\tilde{\phi}_q$:

$$\text{grad } \tilde{\phi}_q(z) = \sum_{i,j=1}^n \zeta^{ij}(z)\nabla_{E_i}\tilde{\phi}_q(z)E_j(z), \tag{21}$$

by excluding the values $\zeta^{ij}(z) \in \mathbb{R}$ that represent the Riemannian metric g at a point $z \in M$, in terms of the basis $\{E_i(z)\}_{i=1}^n$. However, as shown in Lemma 3.3, such dynamics do not modify the set of critical points of the warped cost functions.

Lemma 3.3. *For all $q \in \mathcal{Q}$ we have that grad $\tilde{\phi}_q|_z = 0$ if and only if $\sum_{i=1}^n \nabla_{E_i}\tilde{\phi}_q(z)E_i(z) = 0$.* □

The following theorem provides a first-order version of Theorem 3.2 for the case when the vector field (18) can be explicitly computed or measured in real time, and all the standing assumptions hold.

Theorem 3.4. *The first-order hybrid dynamics $\mathcal{H}_1$ render the set $\mathcal{A} \times \mathcal{Q}$ UGAS, and the set $M \times \mathcal{Q}$ is strongly forward invariant.* □

Similar to Theorem 3.2, the main novelty of Theorem 3.4 is the ability to overcome topological obstructions to global optimization on smooth compact manifolds that emerge in ODEs. In particular, the asymptotic stability result is *global* rather than *almost* global, semi-global, or local. This result, combined with the well-posedness of the dynamics, will allow us to establish important robustness properties with respect to small (potentially adversarial) disturbances, which could also act on the hybrid zeroth-order dynamics $\mathcal{H}_0$.

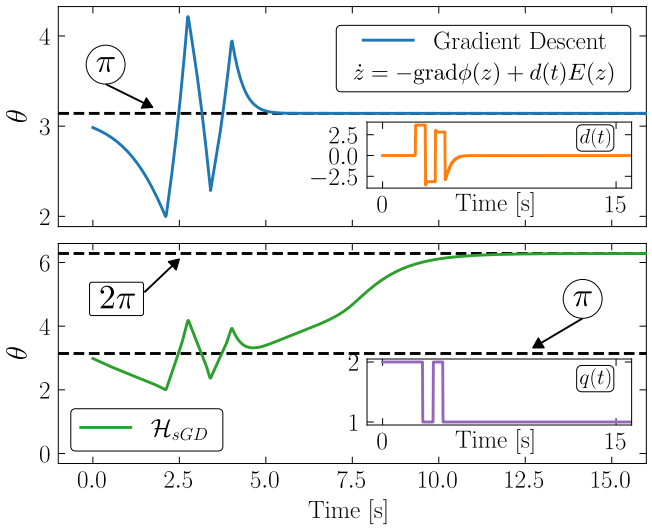


Fig. 2. Top: Trajectory of a gradient flow under a disturbance $d(t)E(z)$. Bottom: Evolution in time of the main state of $\mathcal{H}_1$ under the same perturbation applied to the z-component of the dynamics. See Example 3.6.

3.4. Robustness corollaries: Stability under adversarial disturbances

Crucially, the hybrid dynamics $\mathcal{H}_0$ and $\mathcal{H}_1$ satisfy the Basic Assumptions of Goebel et al. (2012, Ch. 6). Consequently, their stability properties are not drastically affected by small (potentially adversarial) additive disturbances acting on the states and data of the hybrid systems (Goebel et al., 2012, Thm. 7.20). This property is formalized in the following corollary:

Corollary 3.5. *Consider the perturbed first-order hybrid dynamics*

$$x + d_1 \in C_1, \quad \dot{x} = F_1(x + d_2) + d_3 \tag{22a}$$

$$x + d_4 \in D_1, \quad x^+ \in G_1(x + d_5) + d_6 \tag{22b}$$

where $\{C_1, F_1, D_1, G_1\}$ *is the data of* $\mathcal{H}_1$*, and the signals* $d_j : dom(x) \to C_1 \cup D_1$*, for all* $j \in \{1, 2, 4, 5, 6\}$*, and* $d_3 : dom(x) \to TC_1$*, are measurable functions satisfying* $\sup_{(t,j)\in dom(x)} |d_k(t, j)| \leq d^*$*, where* $d^* > 0$*, for all* $k \in \{1, 2, \ldots, 6\}$*. Then, system (22) renders the set* $\mathcal{A} \times \mathcal{Q}$ *GP-AS as* $d^* \to 0^+$*.* □

Robustness results, such as Corollary 3.5, are relevant for practical applications where measurement noise or numerical approximations induce unavoidable disturbances during implementations. They also hold with respect to adversarial perturbations designed to destabilize the set $\mathcal{A}$, or to stabilize spurious equilibria.

Example 3.6. Let $M = \mathbb{S}^1 \subset \mathbb{R}^2$ be the unit circle, which is a smooth, boundaryless compact parallelizable manifold. We consider the cost function $\phi : \mathbb{S}^1 \to \mathbb{R}$, $z \mapsto 1 - z_1$, where $z_i \in [-1, 1]$ represents the ith coordinate of $z \in \mathbb{S}^1$ expressed in regular Cartesian coordinates. The cost function ϕ has two critical points in $\mathbb{S}^1$ corresponding to the global minimizer given by (in polar coordinates) $\theta^* = 2\pi$, and a global maximizer, given by $\theta' = \pi$. To find the unknown minimizer of ϕ, we first implement the first-order dynamics

$$z \in M, \quad \dot{z} = -\nabla_{E(z)}\tilde{\phi}_q(z)E(z) + d(t)E(z), \tag{23}$$

where $E : \mathbb{S}^1 \subset \mathbb{R}^2 \to T\mathbb{S}^1$ is the vector field defined by $E(\cos(\theta), \sin(\theta)) = (-\sin(\theta), \cos(\theta))$ and θ denotes the polar coordinate on the circle. By Lee (2013, Example 8.10.d)), E constitutes a smooth global frame for $\mathbb{S}^1$. In (23), $d(t)E(z)$ is a time-varying perturbation that preserves the invariance of M. The amplitude of this perturbation $d(t)$ was generated by interconnecting (23) with an adversarial hybrid system to stabilize the *maximizer* θ'. As shown in Fig. 2, the adversarial perturbation is always bounded and it succeeds in stabilizing θ'. On the other hand, when this same adversarial signal $d(t)E(z)$ is added in open loop to $\mathcal{H}_1$, as in (22), the hybrid dynamics achieve global convergence to the minimizer θ^*, as shown in the bottom plot of Fig. 2. Finally, we show in Fig. 3 the performance of the hybrid system $\mathcal{H}_1$ when interconnected to the same adversarial dynamical system used to destabilize θ^* in (23). As observed, the hybrid dynamics still achieve convergence to θ^*. □

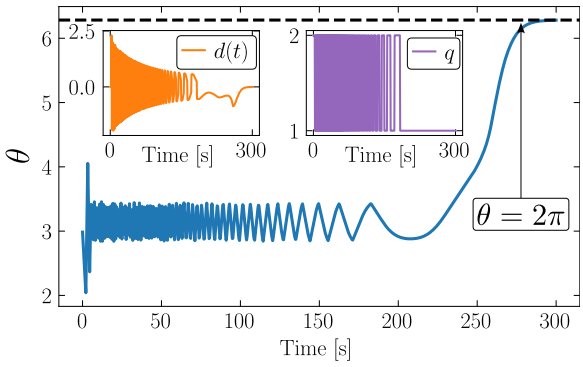


Fig. 3. Trajectories of $\mathcal{H}_1$, under a small adversarial disturbance generated by a dynamical system. The insets show the amplitude of the injected disturbance, as well as the evolution of the index state q in time. See Example 3.6.

We note that smooth gradient-free versions of (23), obtained via averaging theory, might encounter similar issues as those illustrated in Example 3.6. Specifically, if a small adversarial disturbance can locally stabilize the average dynamics of the system to a point outside $\mathcal{A}$, and if this stabilizing effect of the disturbance is preserved after averaging, then applying the same disturbance to the original dynamics may cause the system to locally converge to a neighborhood of that point, as predicted by standard averaging results for ODEs (e.g., Khalil (2002, Ch. 10)). An example of this behavior in obstacle avoidance problems was presented in Poveda et al. (2021, Ex. 1). The question of systematically constructing such adversarial signals in other manifolds remains application-dependent and is not further explored in this paper.

The following corollary parallels the results of Corollary 3.5 for the zeroth-order dynamics $\mathcal{H}_0$.

Corollary 3.7. *Consider the perturbed zeroth-order hybrid dynamics, given by*

$$y + d_1 \in C_0, \quad \dot{y}_0 = F_0(y + d_2) + d_3 \tag{24a}$$

$$y + d_4 \in D_0, \quad y_0^+ \in G_0(y + d_5) + d_6 \tag{24b}$$

where $\{C_0, F_0, D_0, G_0\}$ *is the data of* $\mathcal{H}_0$ *in (11), and the signals* $d_j : dom(y) \to C_0 \cup D_0$*, for all* $j \in \{1, 2, 4, 5, 6\}$*, and* $d_3 : dom(y) \to TC_0$*, are measurable functions satisfying* $\sup_{(t,j)\in dom(y)} |d_k(t, j)| \leq d^*$*, where* $d^* > 0$*, for all* $k \in \{1, 2, \ldots, 6\}$*. Then, system (24) renders the set* $\mathcal{A} \times \mathcal{Q} \times \mathbb{T}^n$ *GP-AS as* $(d^*, \varepsilon_2, \varepsilon_1) \to 0^+$*.* □

Remark 3.5. The class of problems for which smooth optimization dynamics cannot achieve robust global certificates on a compact boundaryless manifold M extends beyond the case where the cost has a unique minimizer. Indeed, as briefly stated in Remark 3.2, the basin of attraction of the set of minimizers $\mathcal{A}$ of a continuous cost ϕ under any outer-semicontinuous, convex-valued and locally bounded optimization dynamics F, is

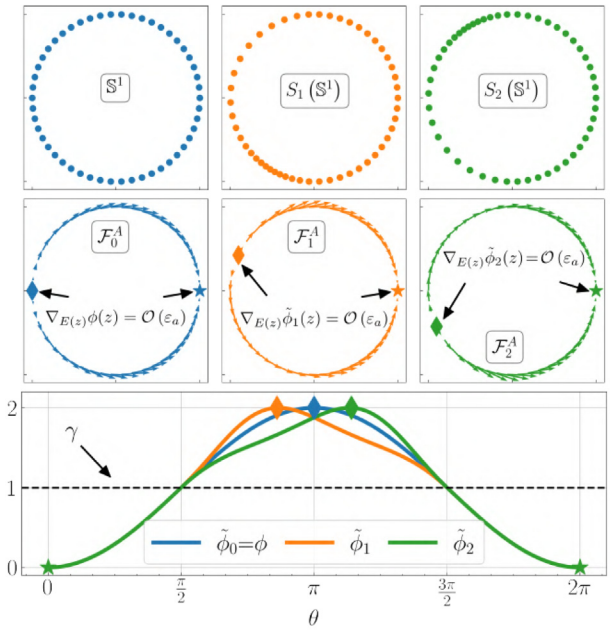


Fig. 4. Top: Visualization of diffeomorphisms on the circle. Middle: Average gradient-based vector fields derived from warped costs. Bottom: Original and warped costs obtained by precomposing with diffeomorphisms.

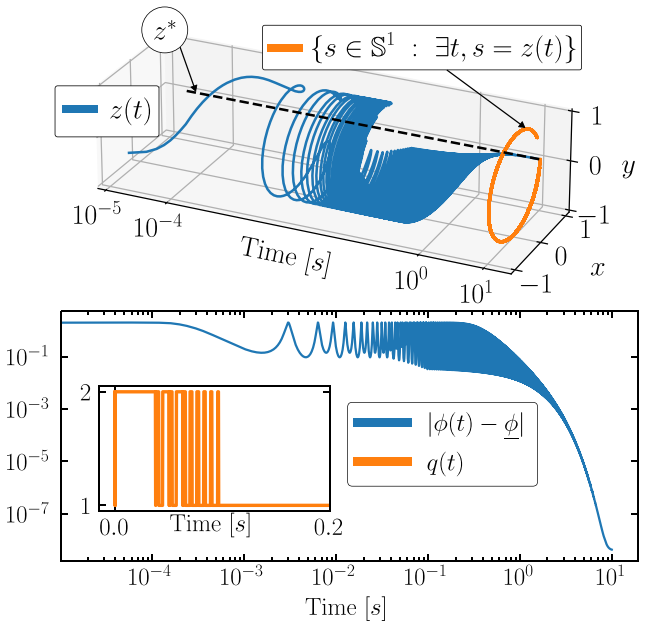


Fig. 5. Gradient-free global optimization via $\mathcal{H}_0$ on $\mathbb{S}^1$ using Geodesic Dithering.

diffeomorphic to an open tubular neighborhood of $\mathcal{A}$. In general, this neighborhood is not topologically compatible with M. For instance, when the cost has a finite set of global isolated minimizers $\mathcal{A} = \bigcup_{i\in I}\{\underline{x}_i\}$, the basin of attraction $\mathcal{B}_F(\mathcal{A}) = \{x\in M \;:\; d_g(x,\underline{x}_i) < \frac{1}{2}\min_{i\neq j} d_g(\underline{x}_j,\underline{x}_i)\}$ is not contractible. However, the results of Theorems 3.2 and 3.4 can be directly extended to overcome this type of topological obstruction. We omit this extension due to space limitations.

3.5. Applications: Synthesis of algorithms

In this section, we showcase the effectiveness of the proposed zeroth-order hybrid dynamics $\mathcal{H}_0$ for solving problems of the form (1) on two distinct compact parallelizable Riemannian manifolds. In particular, we show how to synthesize specific algorithms by generating a δ-gap family of diffeomorphisms adapted to smooth cost functions defined in the unitary circle $\mathbb{S}^1$, and in the special orthogonal group SO(3), and we use the hybrid algorithms to achieve *global* gradient-free (practical) optimization while preserving the forward invariance of the manifolds during the real-time exploration.

3.5.1. Gradient-free feedback optimization on $\mathbb{S}^1$

Consider the unitary circle $\mathbb{S}^1 = \{z\in\mathbb{R}^2 \;:\; |z|^2 = 1\}$. Given $k_q\in\mathbb{R}$, with q belonging to some index set $\mathcal{Q}$, we define the map $S_q^{(1)}:\mathbb{S}^1\to\mathbb{S}^1$ as follows:

$$S_q^{(1)}(z) := \mathbf{1}_{\{\phi(z)\leq\gamma\}} z + \mathbf{1}_{\{\phi(z)>\gamma\}} e^{k_q\alpha(\phi(z)-\gamma)\Psi} z, \tag{25a}$$

where $\Psi := e_2e_1^\top - e_1e_2^\top \in \mathbb{R}^{2\times 2}$, and $\alpha:\mathbb{R}\to\mathbb{R}$ is a continuously differentiable function satisfying: (B_1) $\alpha(0) = 0$; (B_2) $\alpha'(0) = 0$; (B_3) $\alpha'(r) > -1$, $\forall r \geq 0$. The conditions (B_1)-(B_3) ensure that $S_q^{(1)}$ is a continuously differentiable function that constitutes a suitable candidate for a diffeomorphism. In particular, by leveraging (Strizic et al., 2017, Thm 4.1), we have that if

$$|k_q| < \frac{1}{\max\left\{|\alpha'(\phi(z)-\gamma)\, d\phi_z(\Psi z)| \;:\; z\in\mathbb{S}^1,\ \phi(z)\geq\gamma\right\}},$$

then $S_q^{(1)}$ is a diffeomorphism. Although the value of the bound on k_q might not be known (since we do not know the cost function nor its differential) its existence is guaranteed by the continuity of α', ϕ, and $d\phi$, and the compactness of $\{z\in\mathbb{S}^1, \phi(z)\geq\gamma\}$. Estimates of the bound could be obtained by, e.g., a Monte Carlo method that uses measurements or evaluations of ϕ at different points of $z\in\mathbb{S}^1$.

Given a cost $\phi^{(1)}:\mathbb{S}^1\to\mathbb{R}$, and using gains $\{k_q\}_{q\in\mathcal{Q}}$ with corresponding diffeomorphisms defined by (25), it is possible to build a suitable δ-gap synergistic family of diffeomorphisms subordinate to $\phi^{(1)}$. To illustrate this process, similarly to Example 3.6, consider the cost function $\phi^{(1)}(z) := 1 - z_1$. Assume that only measurements or evaluations of $\phi^{(1)}$ are available for feedback design, but that the intermediate value $\gamma = 1\in(0,2) = \left(\underline{\phi}^{(1)}, \phi_2^{(1)}\right)$ and the number of critical points of $\phi^{(1)}$ are known in advance. Let $\alpha(r) = r^2$, and note that it satisfies conditions (B_1)-(B_3). Then, by choosing any two gains satisfying the bound on $|k_q|$, we can obtain a synergistic family of diffeomorphisms subordinate to $\phi^{(1)}$. Indeed, with $\mathcal{Q} = \{1,2\}$, $|k_q| < 1$, $q\in\mathcal{Q}, k_1\neq k_2$ the set $\mathcal{S}^{(1)} = \left\{S_q^{(1)}\right\}_{q\in\mathcal{Q}}$ is a δ-gap family of diffeomorphisms adapted to $\phi^{(1)}$ with gap $\delta < \mu\left(\mathcal{S}^{(1)}\right)$. In Fig. 4 we present a visualization of the diffeomorphisms in this family using the choice $k_1 = \frac{1}{2}, k_2 = -\frac{1}{2}$, and we show how these maps warp the original cost function. We also plot the gradient-based vector fields obtained from the warped cost functions which, as shown in Section 4.2, correspond to $\mathcal{O}(\varepsilon_a)$-perturbations of the flows of $\mathcal{H}_1$ in (18). In turn, the trajectories of $\mathcal{H}_0$ are shown in Fig. 5. As observed, the zeroth-order hybrid dynamics with geodesic dithering successfully converge (globally) to the minimizer of $\phi^{(1)}$, $z^* = (1,0)$, while escaping the other critical point $z' = (-1,0)$.

3.5.2. Gradient-free feedback optimization on SO(3)

As an additional application, we consider the special orthogonal group SO(3), i.e., the group of 3×3 orthogonal matrices with determinant equal to 1 and matrix multiplication as the group operation. By Hall and Hall (2013, Cor. 3.45), SO(3) forms a 3-dimensional compact Lie group. The tangent space at z is given by $T_z\mathrm{SO}(3) = \{zX \;:\; X\in\mathbb{R}^{3\times 3},\ X^\top = -X\}$, see Hall and Hall (2013, Def. 3.18)

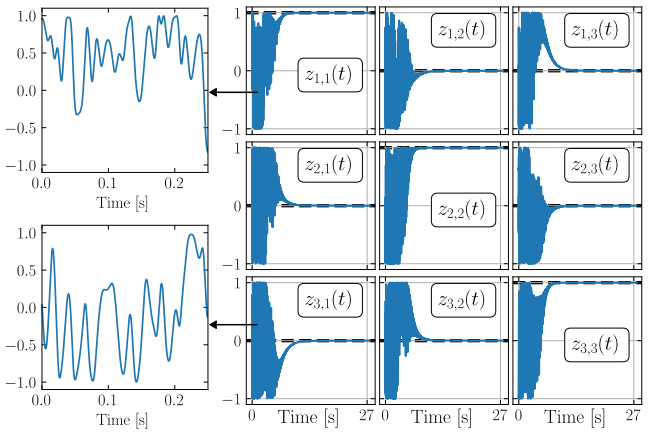


Fig. 6. Synergistic Gradient-free Optimization Seeking on SO(3) via Geodesic Dithering.

To equip SO(3) with a Riemannian structure, we consider the metric $\langle X, Y\rangle_z = \text{tr}(X^\top Y)$ for all $z \in \text{SO(3)}$, and all $X, Y \in T_z\text{SO(3)}$. Using this choice, the Riemannian exponential map can be written in terms of the matrix exponential $e^{(\cdot)}$ (Gallier & Quaintance, 2020, Prop. 21.20) as $\exp_z(X) = ze^{z^{-1}X}$ for $X \in T_z\text{SO(3)}$. Moreover, since SO(3) is a Lie group, it is parallelizable (Lee, 2013, Cor. 8.39). Indeed, for $i \in \{1, 2, 3\}$ let $E_i : \text{SO(3)} \to T\text{SO(3)}$ be the vector field defined as $E_i(z) = zb_i$, where

$$b_1 := \frac{1}{\sqrt{2}}(e_3e_2^\top - e_2e_3^\top), \quad b_2 := \frac{1}{\sqrt{2}}(e_1e_3^\top - e_3e_1^\top),$$
$$b_3 := \frac{1}{\sqrt{2}}(e_2e_1^\top - e_1e_2^\top).$$

It follows that for every $z \in \text{SO(3)}$, $T_z\text{SO(3)} = \text{span}\{E_i(z)\}_{i=1}^3$ and $\langle E_i(z), E_j(z)\rangle_z = \delta_{ij}$, which implies that $\{E_i\}_{i=1}^\infty$ constitutes an orthonormal global frame for SO(3). Using this global frame, we can implement the dithering vector field $\mathcal{D}(z) = \sum_{i=1}^n \hat{\chi}_i E_i(z)$ *everywhere* to extract suitable information from a cost function ϕ at every point in SO(3).

Given a cost $\phi \in C^\infty(\text{SO(3)})$, to establish a suitable family of diffeomorphisms consider the map $S_q^{(2)} : \text{SO(3)} \to \text{SO(3)}$, defined as

$$\mathcal{S}_q^{(2)}(z) = \mathbf{1}_{\{\phi(z)\le\gamma\}}z + \mathbf{1}_{\{\phi(z)>\gamma\}}e^{k_q\alpha(\phi(z)-\gamma)X}z, \tag{26}$$

where $k_q \in \mathbb{R}^n$ and $X \in T_I\text{SO(3)}$, $X \neq 0$ are tunable parameters, and $\alpha : \mathbb{R} \to \mathbb{R}$ satisfies the conditions (B_1)-(B_3) defined in Section 3.5.1 to ensure continuous differentiability of the map. The definition of the map $\mathcal{S}_q^{(2)}$, results from modifying the function introduced in Mayhew and Teel (2011b, Sec 3.4.3) for the angular warping of the two-dimensional sphere by using the function α, and letting the warping act only when ϕ exceeds the threshold γ. For this map we establish the following technical lemma:

Lemma 3.8. *Let k_q satisfy the bound $|k_q| < \overline{k}^{(2)}$, with:*

$$\overline{k}^{(2)} := \frac{\|X\|_F^{-1}}{\max_{z\in SO,\phi(z)\ge\gamma}|\alpha'(\phi(z)-\gamma)|\,\|grad\,\phi|_z\|_F}, \tag{27}$$

and $\|X\|_F = \sqrt{tr(X^\top X)}$. Then, $\mathcal{S}_q^{(2)}$ is a global diffeomorphism. □

To illustrate the application of the zeroth-order hybrid dynamics $\mathcal{H}_0$ in SO(3), we consider the cost function $\phi^{(2)}(z) = \text{tr}((I - z)A)$, where $A = \frac{3}{\sum_{i=1}^3 a_i}\text{diag}(a)$, and $a = (11, 12, 13)$. It follows that $\text{Crit}\phi^{(2)} = \{I\} \cup \bigcup_{i=1}^3\{I + 2[e_i]_\times^2\}$, where $e_i \in \mathbb{R}^3$ denotes the standard basis vector with a 1 in the ith position and zeros in the other entries, and where $[u]_\times : \mathbb{R}^3 \to \mathbb{R}^{3\times3}$ is defined as

$$[u]_\times := \sqrt{2}u_1b_1 + \sqrt{2}u_2b_2 + \sqrt{2}u_3b_3. \tag{28}$$

For this problem, we consider the threshold value $\gamma = 2 \in \left(\underline{\phi}^{(2)}, \phi_2^{(2)}\right)$, we select the gains $k_1 = 0.15$ and $k_2 = -0.15$, let $X = [a/|a|]_\times$, $\alpha(r) = \frac{r^2}{2}$, and consider the family of functions $\mathcal{S}^{(2)} := \{\mathcal{S}_q^{(2)}\}_{q\in\{1,2\}}$. With these choices, the value of the upper bound $\overline{k}^{(2)}$ in (27) is approximately 0.188, which means that $|k_q| < \overline{k}^{(2)}$, and hence, via Lemma 3.8, that the set $\mathcal{S}^{(2)}$ is a family of diffeomorphisms adapted to $\phi^{(2)}$. It can be computed that the family is δ-synergistic with gap $\delta = 0.0796$.

Using $\mathcal{S}^{(2)}$ and the global orthonormal frame $\{E_i\}_{i=1}^3$, we implement the HDS $\mathcal{H}_0$ and obtain the results shown in Fig. 6. The figure shows the trajectories of the entries of the state $z \in \text{SO(3)}$ converging (globally) to the optimal values z_{ij}^*, where $z^* = I$.

3.6. Extensions to non-parallelizable manifolds: Gradient-free feedback optimization on $\mathbb{S}^2$

In this section, we extend our results to manifolds M that are not parallelizable. In such cases, a unique global orthonormal frame is unavailable to define dithering vectors that are valid at every point on M. To address this issue, we employ local orthonormal frames and we introduce a suitable switching mechanism between them to cover M. This mechanism ensures that dithering vectors are always available for real-time exploration. Since the constructions of the dynamics in non-parallelizable manifolds are highly dependent on the manifold, we focus our attention on the 2-dimensional sphere $\mathbb{S}^2 := \{z \in \mathbb{R}^3 : z^\top z = 1\}$. However, we stress that the proposed methodology can be extended to other compact non-parallelizable manifolds.

First, we introduce two local orthonormal frames $\{E_{i,p}\}_{i=1}^2$, $p \in \mathcal{P} := \{1, 2\}$, which will later be used to generate suitable dithering vector fields. Specifically, inspired by Baradaran, Poveda, and Teel (2019), we use local coordinate frames established through the stereographic projection maps:

$$\varphi_1 : U_1 := \mathbb{S}^2 \setminus \{N\} \to \mathbb{R}^2,\ z \mapsto \frac{1}{1 - z_3}(z_1, z_2), \tag{29a}$$

$$\varphi_2 : U_2 := \mathbb{S}^2 \setminus \{S\} \to \mathbb{R}^2,\ z \mapsto \frac{1}{1 + z_3}(z_1, z_2), \tag{29b}$$

where $N := (0, 0, 1)$ and $S := (0, 0, -1)$, denote the north and south pole of $\mathbb{S}^2$, respectively. The stereographic projections constitute homeomorphisms onto their images, and their inverse functions are given by Gallier and Quaintance (2020, Ex. 4.1):

$$\varphi_1^{-1}(u_1, u_2) = \frac{1}{1 + |u|^2}(2u_1, 2u_2, |u|^2 - 1) \tag{30a}$$

$$\varphi_2^{-1}(u_1, u_2) = \frac{1}{1 + |u|^2}(2u_1, 2u_2, 1 - |u|^2). \tag{30b}$$

Using (29)–(30), we let $E_{i,p}(z) := d(\varphi_j^{-1})_{\varphi_j(z)}(e_i)$ for all $z \in U_j$, $i \in \{1, 2\}$, and $p \in \mathcal{P} := \{1, 2\}$, where e_i denotes the ith canonical basis vector in $\mathbb{R}^2$. Unwrapping definitions, we obtain:

$$E_{1,1}(z) = \begin{pmatrix} 1 - z_3 - z_1^2 \\ z_1z_2 \\ (1 - z_3)z_1 \end{pmatrix},\ E_{1,2}(z) = \begin{pmatrix} z_1z_2 \\ 1 - z_3 - z_2^2 \\ (1 - z_3)z_2 \end{pmatrix},$$

$$E_{2,1}(z) = \begin{pmatrix} 1 + z_3 - z_1^2 \\ -z_1z_2 \\ -z_1(1 + z_3) \end{pmatrix},\ E_{2,2}(z) = \begin{pmatrix} -z_1z_2 \\ 1 + z_3 - z_2^2 \\ -z_2(1 + z_3) \end{pmatrix}.$$

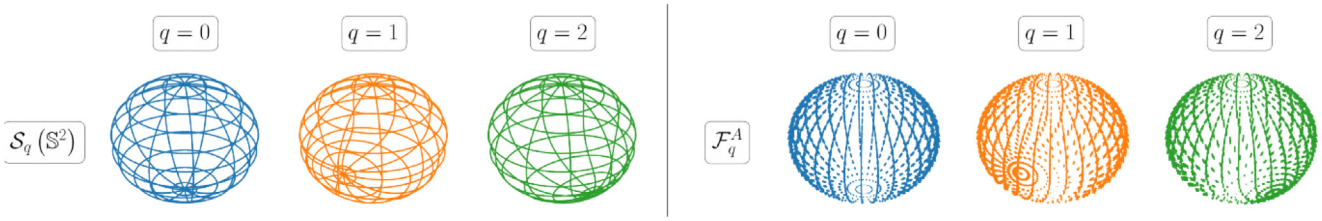


Fig. 7. Left: Visualization of diffeomorphisms on $\mathbb{S}^2$. Right: Average vector fields derived from warped cost functions.

Next, for each local orthonormal frame $\{E_{i,p}\}_{i=1}^2$, $p \in \mathcal{P}$, we define a corresponding dithering vector field used for the purpose of real-time exploration of M:

$$\mathcal{D}_p = \sum_{i=1}^{2} \hat{\chi}_i E_{i,p}. \tag{31}$$

Additionally, for each dithering vector field $\mathcal{D}_p$, and given a δ-gap synergistic family of diffeomorphisms $\mathcal{S} = \{S_q\}_{q\in\mathcal{Q}}$ adapted to a cost function ϕ, we define a family of vector fields $\{\hat{f}_{q,p}(\cdot,\chi)\}_{q\in\mathcal{Q}}$ suitable for exploitation of the information of ϕ learned during the exploration. Specifically, given $q \in \mathcal{Q}$ and $p \in \mathcal{P}$, we let

$$\hat{f}_{q,p}(z,\chi) := \frac{2}{\varepsilon_a}\tilde{\phi}_q\left(\exp_z\left(\varepsilon_a \mathcal{D}_p(z)\right)\right)\mathcal{D}_p(z), \tag{32}$$

where the vector of oscillating amplitudes $\chi \in \mathbb{T}^2$, and the warped cost function $\tilde{\phi}_q = \phi \circ S_q$ are as defined in Section 3.2. Finally, we modify the zeroth-order hybrid dynamics $\mathcal{H}_0$ to incorporate the switching between frames. The new zeroth-order hybrid system, termed $\tilde{\mathcal{H}}_0$, incorporates an additional logic state $p \in \mathcal{P}$ and implements a hysteresis-based switching mechanism dependent on z. The mechanism enables the robust transition between the families of vector fields $\{\hat{f}_{q,p}(z,\chi)\}_{q\in\mathcal{Q}}$, and ensures that the orthonormal frame associated with the selected family satisfies the condition $\text{span}\left(\{E_{i,p}\}_{i=1}^2\right) = T_z\mathbb{S}^2$ for the current value of z. To the best of our knowledge, this approach has not been studied before in the context of zeroth-order optimization and extremum-seeking.

To define the hysteresis-based switching, we first let $r > 1$, and define the open sets $C_p := \varphi_p^{-1}(r\mathbb{B}^\circ)$. By using the definitions of φ_i and φ_i^{-1}, it follows that $C_1 \cup C_2 = \mathbb{S}^2$, and that $\text{span}(\{E_{i,p}\}_{i=1}^2) = T_z\mathbb{S}^2$ for all $z \in \overline{C}_p$ and all $p \in \mathcal{P}$. Using these sets, we characterize the new dynamics $\tilde{\mathcal{H}}_0$, which describe the evolution of the state $\tilde{y} := (z, q, \chi, p) \in \mathbb{S}^2 \times \mathcal{Q} \times \mathbb{T}^2 \times \mathcal{P}$, and have data $\tilde{\mathcal{H}}_0 = \{\tilde{C}_0,\ \tilde{F}_0,\ \tilde{D}_0,\ \tilde{G}_0\}$, with continuous-time dynamics:

$$\tilde{y} \in \tilde{C}_0, \quad \dot{\tilde{y}} = \tilde{F}_0(\tilde{y}) := \begin{pmatrix} -\hat{f}_{q,p}(z,\chi) \\ 0 \\ \frac{2\pi}{\varepsilon_d}\Psi(\omega)\chi \\ 0 \end{pmatrix}, \tag{33}$$

where $\Psi(\omega) \in \mathbb{R}^{4\times 4}$ and $\omega \in \mathbb{R}^2$ are as defined in Section 3.2. The flow set is defined by $\tilde{C}_0 := \tilde{C}_{0,1} \cup \tilde{C}_{0,2}$, where, for all $p \in \mathcal{P}$, we let

$$\tilde{C}_{0,p} := \left\{(z,q,\chi) \in \overline{C}_p \times \mathcal{Q} \times \mathbb{T}^2 : \left(\tilde{\phi}_q - m\right)(z) \le \delta\right\} \times \{p\}.$$

The jump set $\tilde{\mathcal{D}}_0$ is constructed as the union of two sets: 1) $\tilde{D}_{0,d}$, which enables switching between the families of vector fields $\{\hat{f}_{q,p}(\cdot,\chi)\}_{q\in\mathcal{Q}}$, and 2) $\tilde{D}_{0,s}$, which enables the synergistic switching between vector fields within the selected family, akin to the methodology outlined in Section 3.2. Specifically, we let $\tilde{D}_0 := \tilde{D}_{0,d} \cup \tilde{D}_{0,s}$, where $\tilde{D}_{0,d} := \bigcup_{p\in\{1,2\}} \tilde{D}_{p,d}$, and

$$\tilde{D}_{p,d} := (\mathbb{S}^2 \setminus C_p) \times \mathcal{Q} \times \mathbb{T}^2 \times \{p\}, \quad \forall p \in \mathcal{P}$$
$$\tilde{D}_{0,s} := \{(z,q,\chi) \in \mathbb{S}^2 \times \mathcal{Q} \times \mathbb{T}^2 : \left(\tilde{\phi}_q - m\right)(z) \ge \delta\} \times \mathcal{P}.$$

The jump map describing the switches of p is given by $\tilde{G}_{0,d}(\tilde{y}) := (z, q, \chi, 3-p)$, $\forall \tilde{y} \in \tilde{D}_{0,d}$, which updates the current frame used for the purpose of dithering. On the other hand, the jump map describing the switches of q is given by $\tilde{G}_{0,s}(\tilde{y}) := \{z\} \times h(z) \times \{(\chi, p)\}$, $\forall \tilde{y} \in \tilde{D}_{0,s}$, where h is the set-valued map defined in (14). Using these maps, the overall jump map of the HDS $\tilde{\mathcal{H}}_0$ is given by:

$$\tilde{G}_0(\tilde{y}) := \begin{cases} \tilde{G}_{0,s}(\tilde{y}) & \forall \tilde{y} \in \tilde{D}_{0,s} \setminus \tilde{D}_{0,d} \\ \tilde{G}_{0,d}(\tilde{y}) & \forall \tilde{y} \in \tilde{D}_{0,d} \setminus \tilde{D}_{0,s} \\ \tilde{G}_{0,s}(\tilde{y}) \cup \tilde{G}_d(y) & \forall \tilde{y} \in \tilde{D}_{0,d} \cap \tilde{D}_{0,s}. \end{cases}$$

By leveraging our standing assumptions, the following theorem extends the global results of Theorem 3.2 to the non-parallelizable manifold $\mathbb{S}^2$.

Theorem 3.9. *Consider the zeroth-order hybrid dynamics $\tilde{\mathcal{H}}_0$ and let the vector of frequencies ω in (33) satisfy condition (17). Then, the set $\mathcal{A} \times \mathcal{Q} \times \mathbb{T}^2 \times \mathcal{P}$ is GP-AS as $(\varepsilon_d, \varepsilon_a) \to 0^+$.* □

To illustrate the performance of $\tilde{\mathcal{H}}_0$ in $\mathbb{S}^2$, we synthesize the algorithms by using the parameterized transformation $\mathcal{S}_q^{(3)} : \mathbb{S}^2 \to \mathbb{S}^2$ defined as:

$$\mathcal{S}_q^{(3)}(z) = \mathbf{1}_{\{\phi(z)\le\gamma\}} z + \mathbf{1}_{\{\phi(z)>\gamma\}} e^{k_q \alpha(\phi(z)-\gamma)[u]_\times} z, \tag{34}$$

with $k_q \in \mathbb{R}$, $X \in T_I\text{SO}(3)$ and α as defined in Section 3.5.2. Note that $\mathcal{S}_q^{(3)}$ is identical to $\mathcal{S}_q^{(2)}$, except for the fact that its domain and codomain are now $\mathbb{S}^2$ instead of SO(3). The following Lemma extends the result of Lemma 3.8 to $\mathbb{S}^2$.

Lemma 3.10. *Let k_q satisfy $|k_q| < \bar{k}^{(3)}$, where:*

$$\bar{k}^{(3)} := \frac{1}{\max\left\{\left|\alpha'(\phi(z)-\gamma)\, d\phi_z(Xz)\right| : z \in \mathbb{S}^2, \phi(z) \ge \gamma\right\}}.$$

Then, $S_q^{(3)}$ is a diffeomorphism. □

For numerical verification, we consider the cost function $\phi^{(3)} : \mathbb{S}^2 \to \mathbb{S}^2$ defined by $\phi^{(3)}(z) = 1 - z_3$. We choose the threshold value $\gamma = 1$, the gains $k_1 = \frac{1}{2}, k_2 = -\frac{1}{2}$, the matrix $X = [u]_\times \in \text{SO}(3)$, where $u = (0,1,0) \in \mathbb{R}^3$ and $[u]_\times$ is as defined in Section 3.5.2, and let $\alpha(r) = r^2$. With this data, we define the family of transformations $\mathcal{S}^{(3)} := \{S_q^{(3)}\}_{q\in\{1,2\}}$. Since $|k_q| < \bar{k}^{(3)} = 1$ for all $q \in \mathcal{Q} := \{1,2\}$, via Lemma 3.10, $\mathcal{S}^{(3)}$ is a family of diffeomorphisms. In fact, by Lemma 3.1, $\mathcal{S}^{(3)}$ constitutes a δ-synergistic family of diffeomorphisms adapted to ϕ with gap $\delta < \frac{1}{4}$. Fig. 7 shows a visualization of the diffeomorphisms in this family with the choice $k_1 = \frac{1}{2}, k_2 = -\frac{1}{2}$. The figure also shows the vector fields obtained from the warped cost functions. We stress that such diffeomorphisms can be constructed using only mild qualitative knowledge of ϕ, namely, under a suitable choice of γ, which can be seen as an additional tunable parameter of the algorithm. In Fig. 8, we show the trajectory of the coordinates of the state z and indicate when the local frame used for the dithering switches by showing the moments when the state p jumps. In Fig. 9 we show the trajectory evolving on the sphere.

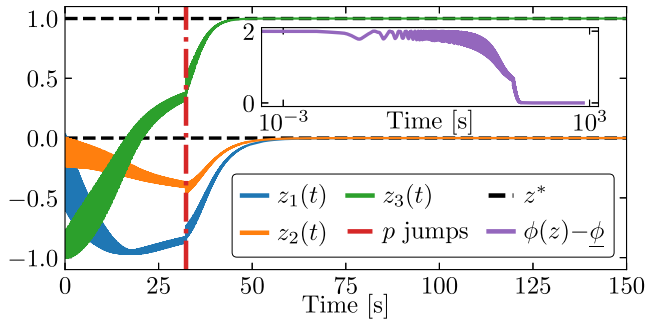


Fig. 8. Evolution of the coordinates of z under the Synergistic Gradient-Free Optimization Seeking dynamics on $\mathbb{S}^2$.

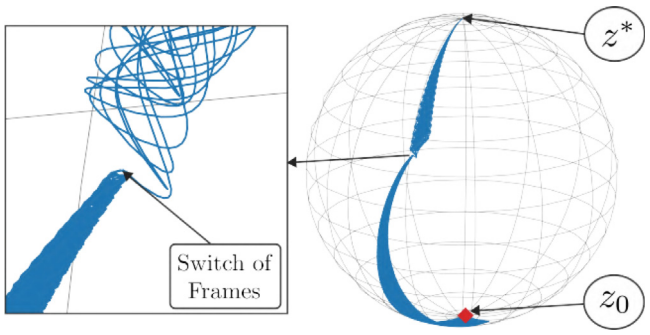


Fig. 9. Synergistic gradient-free optimization on $\mathbb{S}^2$ via geodesic dithering. The inset shows the moment when the system switches between one local frame to another.

As observed, the state z converges to the global minimizer $z^* = (0, 0, 1)$, while escaping the critical point $z' = (0, 0, -1)$.

4. Analysis and proofs

In this section, we present the proofs of our main results. Since the stability results of the zeroth-order hybrid dynamics $\mathcal{H}_0$ in Theorem 3.2 rely on the stability properties of the first-order dynamics $\mathcal{H}_1$, we first present the proof of Theorem 3.4.

4.1. Proof of Theorem 3.4

We begin by presenting the proof of our auxiliary lemmas.

Proof of Lemma 3.1. Suppose that $\mathcal{S}$ is a δ-gap synergistic family of diffeomorphisms adapted to ϕ. Then, we have that $\delta < \mu(\mathcal{S})$, meaning that $\delta < \left(\tilde{\phi}_q - \min_{p\in\mathcal{Q}} \tilde{\phi}_p\right)(z) \quad \forall q \in \mathcal{Q}$, and all $z \in \text{Crit}\, \tilde{\phi}_q \setminus \mathcal{A}$. Then, it follows that for all $q \in \mathcal{Q}$ and $z \in \text{Crit}\tilde{\phi}_q \setminus \mathcal{A}$, there exists $p \in \mathcal{Q}$ such that (16) is satisfied.

Conversely, assume that for every $q \in \mathcal{Q}$ and $z \in \text{Crit}\tilde{\phi}_q \setminus \mathcal{A}$, there exists $p \in \mathcal{Q}$ such that (16) is satisfied. In particular, for all $q \in \mathcal{Q}$ and $z \in \text{Crit}\tilde{\phi}_q \setminus \mathcal{A}$ it follows that $\delta < (\tilde{\phi}_q - \min_{p\in\mathcal{Q}} \tilde{\phi}_p)(z)$, which implies that

$$\delta < \min_{\substack{q\in\mathcal{Q} \\ z\in\text{Crit } \tilde{\phi}_q\setminus\mathcal{A}}} (\tilde{\phi}_q - \min_{p\in\mathcal{Q}} \tilde{\phi}_p)(z).$$

This concludes the proof. ∎

Lemma 4.1. *The HDS $\mathcal{H}_1$ is well-posed.*

Proof. We prove that $\mathcal{H}_1$ satisfies the hybrid-basic conditions (Sanfelice, 2020, Def. 2.20). First, note that the flow map F_1 is continuous, by the continuity of $\sum_{i=1}^n \nabla_{E_i}\tilde{\phi}_q(\cdot)E_i(\cdot)$ in M for all $q \in \mathcal{Q}$, and the fact that $\mathcal{Q}$ is a discrete set. Second, define the function $\tilde{\mu} : M \times \mathcal{Q} \to \mathbb{R}$ by letting $\tilde{\mu}(z, q) := (\tilde{\phi}_q - m)(z)$. Note that $\tilde{\mu}$ is continuous by following similar reasoning as in the continuity argument for F_1. Then, gph $h = \{(z, q) \in M \times \mathcal{Q} : z \in M,\ \tilde{\mu}(z, q) = 0\}$ is closed since $\tilde{\mu}$ is continuous. It follows that h and G_1 are outer-semicontinuous. Boundedness of G_1 follows by compactness of $M \times Q$ and outer-semicontinuity of G_1. The sets C_1 and D_1 are closed, since they are sublevel and superlevel sets, respectively, of the continuous function $\tilde{\mu}$. The result follows via Goebel et al. (2012, Thm. 6.30). ∎

Proof of Lemma 3.3. Let $\phi \in C^\infty(M)$ be arbitrary. Assume that grad $\phi|_z = 0$ at some $z \in M$. Then, by the representation of grad ϕ in terms of the global orthonormal frame $\{E_i\}_{i=1}^n$, it follows that: $\sum_{i,j=1}^n \zeta^{ij}(z)\nabla_{E_i}\phi(z)E_j(z) = 0$. Thus, since the matrix with coefficients $\zeta^{ij}(z)$ is nonsingular for all $z \in M$, as it provides a local representation of the Riemannian metric, and given that $\{E_i\}_{i=1}^n$ is a frame, we obtain:

$$\nabla_{E_i}\phi(z) = 0, \ \forall i \in \{1, \dots, n\}, \tag{35}$$

which implies that $\sum_{i=1}^n \nabla_{E_i}\phi(z)E_i(z) = 0$. Conversely, assume that $\sum_{i=1}^n \nabla_{E_i}\phi(z)E_i(z) = 0$. Then, Eq. (35) holds, and thus $0 = \sum_{i,j=1}^n \zeta^{ij}(z)\nabla_{E_i}\phi(z)E_j(z) = \text{grad}\phi|_z$. ∎

Now, we consider the set of critical points of the warped cost functions that are not minimizers of ϕ:

$$\mathcal{E} := \left\{(z, q) \in M \times \mathcal{Q} \ : \ z \in \text{Crit}\, \tilde{\phi}_q \setminus \mathcal{A}\right\}.$$

The following lemma shows that $\mathcal{E}$ is properly contained in D_1, thus enforcing jumps whenever $(z, q) \in \mathcal{E}$. This means that the HDS $\mathcal{H}_1$ must jump at critical points that are not minimizers of ϕ. In the following, we use A° to denote the topological interior of a set A.

Lemma 4.2. *Suppose that Assumption 3.3 is satisfied. Then $\mathcal{E} \subsetneq D_1^\circ$ and $G_1(\mathcal{E}) \subsetneq C_1^\circ$.*

Proof. First, the fact that $\delta > 0$ combined with the continuity and positive semidefiniteness of $(\tilde{\phi}_q - m)(\cdot)$ ensures the existence of open subsets of C_1 and D_1 where $0 < (\tilde{\phi}_q - m)(z) < \delta$ and $(\tilde{\phi}_q - m)(z) > \delta$, respectively, proving that C_1° and D_1° are nonempty. Second, consider $(z, q) \in \mathcal{E}$. Lemma 3.1 guarantees the existence of $p \in \mathcal{Q}$ such that $\tilde{\phi}_p(z) + \delta < \tilde{\phi}_q(z)$. Given that $m(z) \leq \tilde{\phi}_q(z)$ for all $p \in \mathcal{Q}$, we deduce: $m(z) + \delta < \tilde{\phi}_q(z)$, implying $(z, q) \in D_1$. Thus, $\mathcal{E} \subseteq D_1$. Now, consider $(z, q) \in D_1 \setminus D_1^\circ$, which implies $m(z) = \tilde{\phi}_q(z) - \delta$. Assume, for contradiction, that $(z, q) \in \mathcal{E}$. By Lemma 3.1, there exists $p \in \mathcal{Q}$ such that $\tilde{\phi}_p(z) < \tilde{\phi}_q(z) - \delta = m(z)$, contradicting $m(z) \leq \tilde{\phi}_q(z)$ for all $q \in \mathcal{Q}$. Hence, $(z, q) \notin \mathcal{E}$, proving that $D_1 \setminus D_1^\circ$ contains elements not in $\mathcal{E}$, and therefore that $\mathcal{E} \subsetneq D_1^\circ$. The fact that $G_1(\mathcal{E}) \subsetneq C_1^\circ$ follows by construction, since after a jump we have that $\tilde{\phi}_{q^+}(z^+) - m(z^+) = \tilde{\phi}_{q^+}(z) - m(z) = 0 < \delta$. ∎

By leveraging the results of the previous lemmas we can now prove the first main theorem of the paper.

Proof of Theorem 3.4. Consider the Lyapunov function:

$$V(x) := \tilde{\phi}_q(z) - \underline{\phi}, \tag{36}$$

which is continuous due to similar arguments to the ones used to prove the continuity of F_1 in Lemma 4.1. Since $\underline{\phi} < \phi(z)$ for all $z \notin \mathcal{A}$, together with (A_2) in Definition 3.1, we have that $\tilde{\phi}_q(z) - \underline{\phi} \geq 0$ for all $(z, q) \in M \times \mathcal{Q}$ and $\tilde{\phi}_q(z) - \underline{\phi} = 0$ if and

only if $(z,q) \in \mathcal{A} \times \mathcal{Q}$. Therefore, it follows that $V(x) \geq 0$ for all $(z,q) \in M \times \mathcal{Q}$ and $V(x) = 0$ if and only if $z \in \mathcal{A}$. Now, during the flows of $\mathcal{H}_1$, the Lie-derivative of V satisfies

$$\begin{aligned}\mathcal{L}_{F_1} V(x) &= -\sum_{i=1}^{n} \nabla_{E_i} \tilde{\phi}_q(z) E_i(z) \left(\tilde{\phi}_q - \underline{\phi}\right) \\ &= -\sum_{i=1}^{n} (E_i(z)\tilde{\phi}_q)^2 =: u_C(x), \ \ \forall x \in C_1, \end{aligned} \tag{37}$$

where we used the fact that $\mathcal{L}_X f(z) = \nabla_X f(z) = (X(z))(f)$ for all $X \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$, and that $v(c) = 0$ for all $v \in T_z M$, every constant function c, and all $z \in M$, via Lee (2013, Lemma 3.4). On the other hand, during the jumps of $\mathcal{H}_1$, using the definition of h and m in (14) and (15), it follows that:

$$\begin{aligned} V\left(x^+\right) - V(x) &= \left(\tilde{\phi}_{h(z)} - \tilde{\phi}_q\right)(z) \\ &= \left(m - \tilde{\phi}_q\right)(z) \leq -\delta =: u_D(x), \end{aligned} \tag{38}$$

for all $x \in D_1$. Since $u_C(x) \leq 0$ for all $z \in C_1$ and $u_D(x) < 0$ for all $x \in D_1$, it follows that $\mathcal{A}$ is stable under $\mathcal{H}_1$ via Sanfelice (2020, Thm. 3.19).

To show the attractivity of $\mathcal{A}$ we employ the hybrid invariance principle (Sanfelice, 2020, Thm. 3.23). Indeed, since $u_C(x) \leq 0$ for all $x \in C_1$ and $u_D(x) < 0$ for all $x \in D_1$, and using $u_D^{-1}(0) = \emptyset$, given $r \in V(\mathcal{A} \cup \mathcal{E}) \subset [0, \overline{\phi} - \underline{\phi}]$, solutions approach the largest weakly invariant set in $V^{-1}(r) \cap ((\mathcal{A} \cup \mathcal{E}) \times \mathcal{Q})$. Let Ω denote such an invariant set and assume that $r \neq 0$. By the definition of $\mathcal{E}$ and the synergistic family of diffeomorphisms, it follows that $\Omega \subset \mathcal{E}$. Additionally, by Lemma 4.2, we obtain that $\Omega \subset D_1^\circ$. Since $D_1^\circ \cap C_1 = \emptyset$ by construction, for Ω to be invariant under $\mathcal{H}_1$, we would need to have that $\Omega = G(\Omega)$, but this would imply, via Lemma 4.2, that $\Omega \subset C_1^\circ$, and thus that $\Omega \subset C_1^\circ \cap D_1^\circ = \emptyset \implies \Omega = \emptyset$. Therefore, we must have that $r = 0$, and thus that $\forall (z(0), q(0)) \in M \times \mathcal{Q}$ solutions approach the largest weakly invariant set in $V^{-1}(0) \cap ((\mathcal{A} \cup \mathcal{E}) \times \mathcal{Q}) = \mathcal{A} \times \mathcal{Q}$, which is $\mathcal{A} \times \mathcal{Q}$ it self. UGAS follows directly by the global attractivity and stability of $\mathcal{A}$. ■

4.2. Proof of Theorem 3.2

The proof uses tools recently developed for averaging on compact Riemannian manifolds (Taringoo et al., 2018) together with the framework for hybrid extremum seeking control introduced in Poveda and Teel (2017).

First, since M is compact, we can select $\varepsilon_a \in \mathbb{R}_{>0}$ such that $\exp_z(\varepsilon_a \mathcal{D}(z)) \in \imath(M)$, with $\imath(M)$ the injectivity radius of M (Taringoo et al., 2018, Lemma 3.2). This makes possible a Taylor expansion in normal coordinates along the geodesic dithers for every $\tilde{\phi}_q$, such that the average dynamics of $\mathcal{H}_0$ can be computed to be (see Poveda and Teel (2017)) $\mathcal{H}_0^A = \{C_1, F_0^A, D_1, G_1\}$, where C_1, D_1, G_1 are defined in (20a), (19) and (20b) respectively, and $F_0^A : M \times \mathcal{Q} \to TM \times \mathbb{N}$ is the *average flow map*, given by:

$$F_0^A(x) := \begin{pmatrix} -\sum_{i=1}^n \nabla_{E_i} \tilde{\phi}_q(z) E_i(z) + \sum_{i=1}^n \mathcal{O}(\varepsilon_a) E_i(z) \\ 0 \end{pmatrix}.$$

Hence, on closed subsets of M we have that

$$F_0^A(x) \in \overline{\text{con}}_z F_1(x + k\varepsilon_a \mathbb{B}, 0) + (k\varepsilon_a \mathbb{B}, 0)\,, \tag{39}$$

for some $k > 0$, where F_1 was defined in (18). Here, the convex hull affects the state z only, and the Minkowski additions $(z + k\varepsilon_a \mathbb{B})$ are defined in a suitable ambient Euclidean space which always exists due to the Whitney Embedding Theorem (Lee, 2013, Thm 6.15). By (39), any solution of the average dynamics $\mathcal{H}_0^A$ is also a solution of an inflated HDS generated from $\mathcal{H}_1$. Hence, and since $\mathcal{H}_1$ is a well-posed HDS via Lemma 4.1, by Goebel et al. (2012, Thm. 7.21) we conclude that system $\mathcal{H}_0^A$ renders the set $\mathcal{A}$ GP-AS in the ambient Euclidean space as $\varepsilon_a \to 0^+$. Since $\mathcal{H}_0^A$ and $\mathcal{H}_1$ are nominally well-posed, all conditions to apply (Poveda & Teel, 2017, Cor. 1) are satisfied. Therefore, together with the compactness of M, $\mathcal{H}_0$ renders the set $\mathcal{A} \times \mathbb{T}^n$ GP-AS in the ambient Euclidean space as $(\varepsilon_d, \varepsilon_a) \to 0^+$. Note that any solution z to $\mathcal{H}_0$ is constrained to M since the dithering is performed along geodesics on the manifold, and $\hat{f}_q(z, \chi) \in T_z M$ for all $(z, q, \chi) \in M \times \mathcal{Q} \times \mathbb{T}^n$. Thus, we obtain GP-AS of $\mathcal{A}$ in the sense of Definition 2.2. ■

4.3. Proof of Lemma 3.8

First, we compute the differential of $\mathcal{S}_q^{(2)}$ and, whenever (27) is satisfied, we show that it is full rank for all $z \in \text{SO(3)}$. When $\phi(z) \leq \gamma$, the differential is trivially full-rank since $d\left(\mathcal{S}_q^{(2)}\right)_z = I$. When $\phi(z) > \gamma$, we obtain: $d\left(\mathcal{S}_q^{(2)}\right)_z = e^{k_q \alpha(\phi(z)-\gamma)X} \left[k_q \alpha'(\phi(z)-\gamma)(X \cdot z)\, d\phi_z + I\right]$. Since the linear operator $v \mapsto e^{k_q \alpha(\phi(z)-\gamma)X} v$ is invertible, because $e^{rX} \in \text{SO(3)}$ for all $r \in \mathbb{R}$ and $X \in T_I \text{SO(3)}$, to prove that $d\left(\mathcal{S}_q^{(2)}\right)_z$ is full-rank it suffices to show that $(\Psi_z + I)$ is invertible, where $\Psi_z := k_q \alpha'(\phi(z) - \gamma)(X \cdot z)\, d\phi_z$. To this end, letting $\|\Psi_z\|_z$ denote the operator 2-norm induced by the inner product in the Hilbert space $\mathcal{V}_z := (T_z \text{SO(3)}, \langle \cdot, \cdot \rangle_z)$, we obtain that: $\|\Psi_z\|_z \leq |k_q \| \alpha'(\phi(z) - \gamma)| \|X\|_F \|\text{grad}\phi|_z\|_F$. Then, whenever (27) is satisfied, it follows that $\|\Psi_z\|_z < 1$, which implies that $(I + \Psi_z)$ is invertible, and hence that $d\left(\mathcal{S}_q^{(2)}\right)_z$ is full rank for all z such that $\phi(z) > \gamma$. By the inverse function theorem (Lee, 2013, Thm. 4.5), it follows that $\mathcal{S}_q^{(2)}$ is a local diffeomorphism everywhere. Now, note that $\mathcal{S}_q^{(2)}$ is a proper map[1] since it is continuous and SO(3) is a compact Hausdorff space. This fact, together with the compactness of SO(3), implies that $\mathcal{S}_q^{(2)}$ is surjective via Ho (1975, Lemma. 1). Injectivity of $\mathcal{S}_q^{(2)}$ follows from the arguments presented in Mayhew and Teel (2011b, Appendix, Proof Thm. 8), which we omit here for conciseness. Since $\mathcal{S}_q^{(2)}$ is bijective, as well as a local diffeomorphism everywhere, it follows that it is a global diffeomorphism. ■

4.4. Proof of Theorem 3.9

The proof employs the same concepts as the proof of Theorem 3.2. We provide some details for completeness. Specifically, we now consider the Taylor expansion of the flow-map $\tilde{F}_0$ in normal coordinates and we analyze the corresponding average hybrid dynamics $\tilde{\mathcal{H}}_0^A = \{\tilde{C}_1, \tilde{F}_0^A, \tilde{D}_1, \tilde{G}_1\}$, describing the evolution of the state $\tilde{x} := (z, q, p) \in \mathbb{S}^2 \times \mathcal{Q} \times \mathcal{P}$. In this case, the average flow-map $\tilde{F}_0^A(\cdot)$ is given by

$$\tilde{F}_0^A(\tilde{x}) = \left\{ \tilde{\mathcal{F}}_{1,qp}(z) + \sum_{i=1}^{2} \mathcal{O}(\varepsilon_a) E_{i,p}(z) \right\} \times \{0\} \times \{0\},$$

where $\tilde{\mathcal{F}}_{1,qp}(z) := -\sum_{i=1}^2 \nabla_{E_{i,p}} \tilde{\phi}_q(z) E_{i,p}(z)$. The flow set $\tilde{C}_1$, the jump set $\tilde{D}_1$, and the jump map $\tilde{G}_1$ are the same as the sets $\tilde{C}_0$, $\tilde{D}_0$, and them map $\tilde{G}_0$ defined in Section 3.6, but disregarding the state $\chi \in \mathbb{T}^2$ from the main state of the system. Using this construction, (39) becomes

$$\tilde{F}_0^A(\tilde{x}) \in \overline{\text{con}}_z \tilde{F}_1(\tilde{x} + k\varepsilon_a \mathbb{B}, 0) + (k\varepsilon_a \mathbb{B}, 0)\,, \tag{40}$$

where $k > 0$, and $\tilde{F}_1(\tilde{x}) := \{\tilde{F}_{1,qp}(z)\} \times \{0\} \times \{0\}$. Furthermore, let $\tilde{\mathcal{H}}_1$ be the first-order HDS with data $\tilde{\mathcal{H}}_1 = \{\tilde{C}_1, \tilde{F}_1, \tilde{D}_1, \tilde{G}_1\}$, and consider the same Lyapunov function of (36). During the flows of

[1] A map $f : A \to B$ is proper if the preimage of each compact subset of B is compact.

$\tilde{\mathcal{H}}_1$, it follows that $\mathcal{L}_{\tilde{F}_1} V(\tilde{x}) = -\sum_{i=1}^{n} (E_{i,p}(z)\tilde{\phi}_q)^2$ for all $\tilde{x} \in \tilde{C}_1$. During the jumps of $\tilde{\mathcal{H}}_1$, the change of the Lyapunov function $\Delta V(\tilde{x}) := V(\tilde{x}^+) - V(\tilde{x})$ satisfies: $\Delta V(\tilde{x}) = -\delta$ whenever $\tilde{x} \in \tilde{D}_{1,s} := \{(z,q) \in \mathbb{S}^2 \times \mathcal{Q} : \left(\tilde{\phi}_q - m\right)(z) \geq \delta\} \times \mathcal{P}$, and $\Delta V(\tilde{x}) = 0$ whenever $\tilde{x} \in \tilde{D}_1 \setminus \tilde{D}_{1,s}$. In words, the Lyapunov function decreases whenever there is a switch between warped cost functions, denoted by a change in q, and does not increase if the system only switches between families of vector fields, i.e., *only* when the state p changes. Now, by the structure of the flow and jump sets, after a jump that only changes p is triggered, the system can either exhibit a change in q, after which it necessarily flows, or directly flows. The converse is true if a jump that only changes q is triggered first. Then, leveraging the decrease of the Lyapunov function during flows and employing a similar reasoning as in the proof of Theorem 3.4, it follows that every solution of $\tilde{\mathcal{H}}_1$ converges to the largest weakest invariant set in $V^{-1}(0) \cap ((\mathcal{A} \cup \mathcal{E}) \times \mathcal{Q} \times \mathcal{P})$, which is $\tilde{\mathcal{A}} := \mathcal{A} \times \mathcal{Q} \times \mathcal{P}$ itself. It follows that $\tilde{\mathcal{A}}$ is UGAS under $\tilde{\mathcal{H}}_1$ via the hybrid invariance principle (Sanfelice, 2020, Thm. 3.23). The GP-AS of $\mathcal{A}$ under $\tilde{\mathcal{H}}_0$ is obtained by using (22), the well-posedness of $\tilde{\mathcal{H}}_1$ and $\tilde{\mathcal{H}}_0^A$, and applying the same arguments at the end of the proof of Theorem 3.2. ■

4.5. Proof of Lemma 3.10

Using the fact that $d\left(e^{A\eta(z)}\right)_z = e^{A\eta(z)}(I + Azd\eta_z)$ for $A \in \mathbb{R}^{n\times n}$ and $\eta : \mathbb{R}^n \to \mathbb{R}$ we get: $\det\left(d\left(S_q^{(3)}\right)_z\right) = 1 + k_q \alpha'(\phi(z) - \gamma)\, d\phi_z(Xz)$, for all z such that $\phi(z) > \gamma$. Thus, whenever $|k_q| < \overline{k}^{(3)}$, $\det(d(S_q^{(3)})_z) \neq 0$ for all $z \in \mathbb{S}^2$. Note that $S_q^{(3)}$ is proper, being both continuous and defined on the compact space $\mathbb{S}^2$. Then, by the fact that $\mathbb{S}^2$ is simply connected and $\det(dS_q^{(3)}) \neq 0$, it follows that $S_q^{(3)}$ is a diffeomorphism via Gordon (1972, Thm. B). ■

5. Conclusions and outlook

We introduced a novel class of zeroth-order hybrid algorithms for the global solution of gradient-free optimization problems on smooth, compact, and boundaryless manifolds. These algorithms combine continuous-time dynamics and discrete-time dynamics to achieve robust *global* practical stability of the optimizer of a smooth cost function accessible only via measurements or evaluations. The proposed approach overcomes topological obstructions that prevent the solution of this problem using algorithms modeled by smooth ODEs. We characterized the stability and robustness of the algorithms using tools from the theory of hybrid dynamic inclusions. Future research will explore tracking problems in time-varying optimization settings, as well as the incorporation of dynamic plants in the loop. A completely coordinate-free formulation of the hybrid algorithms, and the development of accelerated dynamics and single-point algorithms, are also future research directions.

Acknowledgments

The authors would like to thank Mahmoud Abdelgalil for fruitful discussions on non-parallelizable manifolds via first-order averaging, as well as the anonymous reviewers for multiple constructive comments.

References

Absil, P.-A., Mahony, R., & Sepulchre, R. (2009). *Optimization algorithms on matrix manifolds*. Princeton University Press.

Angeli, D. (2004). An almost global notion of input-to-state stability. *IEEE Transactions on Automatic Control*, *49*(6), 866–874.

Angeli, D., & Efimov, D. (2015). Characterizations of input-to-state stability for systems with multiple invariant sets. *IEEE Transactions on Automatic Control*, *60*(12), 3242–3256.

Baradaran, M., Poveda, J. I., & Teel, A. R. (2019). Global optimization on the sphere: A stochastic hybrid systems approach. *IFAC-PapersOnLine*, *52*(16), 96–101.

Berkane, S., Abdessameud, A., & Tayebi, A. (2017). Hybrid global exponential stabilization on SO(3). *Automatica*, *81*, 279–285.

Bhat, S. P., & Bernstein, D. S. (2000). A topological obstruction to continuous global stabilization of rotational motion and the unwinding phenomenon. *Systems & Control Letters*, *39*(1), 63–70.

Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010* (pp. 177–186). Springer.

Casau, P., Mayhew, C. G., Sanfelice, R. G., & Silvestre, C. (2019). Robust global exponential stabilization on the n-dimensional sphere with applications to trajectory tracking for quadrotors. *Automatica*, *110*, Article 108534.

Casau, P., Sanfelice, R. G., & Silvestre, C. (2019). Adaptive backstepping of synergistic hybrid feedbacks with application to obstacle avoidance. In *2019 American ctrl. conf.* (pp. 1730–1735).

Coron, J.-M. (1992). Global asymptotic stabilization for controllable systems without drift. *Mathematics of Control, Signals, and Systems*, *5*(3), 295–312.

Dürr, H.-B., Stanković, M. S., Johansson, K. H., & Ebenbauer, C. (2014). Extremum seeking on submanifolds in the Euclidian space. *Automatica*, *50*(10), 2591–2596.

Efimov, D. (2012). Global Lyapunov analysis of multistable nonlinear systems. *SIAM Journal on Control and Optimization*, *50*(5), 3132–3154.

Gabay, D. (1982). Minimizing a differentiable function over a differential manifold. *Journal of Optimization Theory and Applications*, *37*(2), 177–219.

Gallier, J., & Quaintance, J. (2020). *Differential geometry and Lie groups: A computational persepective*: *vol. 12*, Springer.

Goebel, R., Sanfelice, R. G., & Teel, A. R. (2012). *Hybrid dynamical systems: modeling, stability, and robustness*. Princeton University Press.

Gordon, W. B. (1972). On the diffeomorphisms of Euclidean space. *American Mathematical Monthly*, *79*(7), 755–759.

Grivopoulos, S., & Bamieh, B. (2003). Lyapunov-based control of quantum systems. In *42nd IEEE international conference on decision and control*: *vol. 1*, (pp. 434–438).

Hall, B. C., & Hall, B. C. (2013). *Lie groups, Lie algebras, and representations*. Springer.

Hauser, J. (2002). A projection operator approach to the optimization of trajectory functionals. *IFAC Proceedings Volumes*, *35*(1), 377–382.

Helmke, U., & Moore, J. B. (2012). *Optimization and dynamical systems*. Springer Science & Business Media.

Ho, C. W. (1975). A note on proper maps. *Proceedings of the American Mathematical Society*, *51*(1), 237–241.

Khalil, H. K. (2002). *Nonlinear systems* (3rd ed.). Prentice Hall.

Krstić, M., & Wang, H.-H. (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, *36*(4), 595–601.

Lauand, C. K., & Meyn, S. (2023). Quasi-stochastic approximation: Design principles with applications to extremum seeking control. *IEEE Control Systems Magazine*, *43*(5), 111–136.

Lee, J. M. (2013). *Introduction to smooth manifolds*. Springer.

Lee, J. M. (2018). *Introduction to Riemannian manifolds*. Springer.

Malisoff, M., Krichman, M., & Sontag, E. (2006). Global stabilization for systems evolving on manifolds. *Journal of Dynamical and Control Systems*, *12*(2), 161–184.

Mayhew, C. G. (2010). *Hybrid control for topologically constrained systems*. University of California, Santa Barbara.

Mayhew, C. G., & Teel, A. R. (2011a). On the topological structure of attraction basins for differential inclusions. *Systems & Control Letters*, *60*(12), 1045–1050.

Mayhew, C. G., & Teel, A. R. (2011b). Synergistic potential functions for hybrid control of rigid-body attitude. In *2011 American ctrl. conf.* (pp. 875–880). IEEE.

Milnor, J. (2015). *Lectures on the H-cobordism theorem*: *vol. 2258*, Princeton University Press.

Poveda, J. I., Benosman, M., Teel, A. R., & Sanfelice, R. G. (2021). Robust coordinated hybrid source seeking with obstacle avoidance in multivehicle autonomous systems. *IEEE Transactions on Automatic Control*, *67*(2), 706–721.

Poveda, J. I., & Teel, A. R. (2017). A framework for a class of hybrid extremum seeking controllers with dynamic inclusions. *Automatica*, *76*, 113–126.

Sanfelice, R. G. (2020). *Hybrid feedback control*. Princeton University Press.

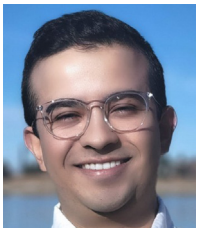
Sontag, E. D. (1999). Stability and stabilization: Discontinuities and the effect of disturbances. In *Nonlinear analysis, differential equations and control* (pp. 551–598). Springer.

Sontag, E. D. (2013). *Mathematical control theory: deterministic finite dimensional systems*: *vol. 6*, Springer Science & Business Media.

Strizic, T., Poveda, J. I., & Teel, A. R. (2017). Hybrid gradient descent for robust global optimization on the circle. In *56th conf. on decision and ctrl.* (pp. 2985–2990). IEEE.

Suttner, R. (2022). Extremum seeking control for a class of mechanical systems. *IEEE Transactions on Automatic Control*.
Suttner, R., & Krstić, M. (2023). Nonlocal nonholonomic source seeking despite local extrema. *IEEE Transactions on Automatic Control*.
Taringoo, F., Dower, P. M., Nesic, D., & Tan, Y. (2018). Optimization methods on Riemannian manifolds via extremum seeking algorithms. *SIAM Journal on Control and Optimization*, *56*(5), 3867–3892.
Udriste, C. (2013). *Convex functions and optimization methods on Riemannian Manifolds: vol. 297*, Springer Science & Business Media.

Daniel E. Ochoa is a Ph.D. candidate in the Department of Electrical and Computer Engineering at the University of California, San Diego. He received double B.S. degrees in Electronics Engineering (cum laude) and Physics with a minor in Computational Mathematics from the the University of Los Andes, Colombia. Subsequently, he received an M.Sc. degree in Electronics and Computer Engineering from the University of Los Andes, Colombia in 2019, and an M.Sc. degree in Electrical Engineering from the University of Colorado, Boulder in 2022. He has received the Young Author Award at the 8th IFAC Conference on Analysis and Design of Hybrid Systems.



Jorge I. Poveda is Assistant Professor in the Electrical and Computer Engineering Department at the University of California, San Diego. He received double B.S. degrees in Electronics and Mechanical Engineering, both from the University of Los Andes, Bogota, Colombia. Subsequently, he obtained his M.Sc. and Ph.D. degrees in Electrical and Computer Engineering from the University of California, Santa Barbara, in 2016 and 2018, respectively, and was a Postdoctoral Fellow at Harvard University in 2018. From 2019 to 2022, he was Assistant Professor at the University of Colorado, Boulder. He has received the NSF CRII (2020) and CAREER (2022) awards, the AFOSR (2022) and SHPE (2024) Young Investigator awards, the AACC Donald P. Eckman (2023) award, the IEEE Transactions on Control of Network Systems Best Paper (2023) award, and the CCDC Outstanding Scholar Fellowship (2013) award and Best Ph.D. Dissertation (2020) award from UCSB. He was also a finalist for the Best Student Paper award at the IEEE Conference on Decision and Control in 2017 (as student) and 2021 (as co-author), and is the advisor of students who have won the Young Author Award at the IFAC Conference on Analysis and Design of Hybrid Systems (2024) and the Best Student Paper Award finalist at the American Control Conference (2024). He is an Associate Editor for IEEE Control Systems Letters and the IFAC journal Nonlinear Analysis: Hybrid Systems.