Saddle Flow Dynamics: Observable Certificates and Separable Regularization

Pengcheng You and Enrique Mallada

Abstract— This paper proposes a certificate, rooted in observability, for asymptotic convergence of saddle flow dynamics of convex-concave functions to a saddle point. This observable certificate directly bridges the gap between the invariant set and the equilibrium set in a LaSalle argument, and generalizes conventional conditions such as strict convexity-concavity and proximal regularization. We further build upon this certificate to propose a separable regularization method for saddle flow dynamics that makes minimal requirements on convexity-concavity and yet still guarantees asymptotic convergence to a saddle point. Our results generalize to saddle flow dynamics with projections on the vector field and have an immediate application as a distributed solution to linear programs.

I. INTRODUCTION

There has been increasing interest in studying optimization algorithms from a dynamical systems view-point as means to understand their stability [1], [2], rate of convergence [3]–[5], and robustness [5]–[8]. For example, in the basic case of gradient descent dynamics for unconstrained convex optimization, the objective function monotonically decreases along trajectories towards the optimum, naturally rendering a Lyapunov function [9]. Such realization later on leads to multiple extensions, including finite-time convergence [10], [11], acceleration [4], [5], and time-varying optimization [12], [13].

One prominent branch of this line of work is the study of saddle flow dynamics, i.e., dynamics in the gradient descent direction on a sub-set of variables and the gradient ascent direction on the rest. Designed for locating min-max saddle points, saddle flow dynamics are particularly suited for solving constrained optimization problems via primal-dual methods [14], and finding Nash equilibria of zero-sum games [15], which lead to a broad application spectrum, including power systems [16], communication networks [17], and cloud computing [18].

The study of saddle flow dynamics is rooted in the seminal work [19], which first considers asymptotic behavior of saddle flows in the context of primal-dual algorithms. This has received renewed attention over the last decade. In particular, conventional conditions for asymptotic convergence are re-validated with advanced analytical tools in more general setups. For instance, [14] revisits the condition of strict convexity-concavity in the case of discontinuous vector fields, using LaSalle's invariance principle for discontinuous Caratheodory systems.

More recent analysis has been centered around asymptotic convergence of saddle flow dynamics under weaker conditions. A set of literature establishes stronger results that require only local strong convexity-concavity [8], convexity-linearity, or strong quasiconvexity-quasiconcavity [20]. Further, asymptotic oscillatory behavior of saddle flow dynamics for general convex-concave functions and their Ω limit sets are explicitly characterized in [1], [2].

Another active topic is to circumvent the above conditions via regularization, which is an appealing approach to handling the Lagrangian of constrained convex/linear optimization. This includes various penalty terms on equality constraints or even projected inequality constraints [7], [21], as well as the proximal method [18], [22]. The rationale of regularization is further interpreted in [23] in the frequency domain. Despite the merit of regularization that relaxes conditions for convergence, the extra penalty terms commonly introduce couplings that may require additional computation and communication overheads in order to realize distributed implementation.

Our work complements the existing literature by first obtaining a sufficient certificate for asymptotic convergence of saddle flow dynamics of a convex-concave function. The underlying rationale still adopts the standard quadratic Lyapunov function and LaSalle's invariance principle, but directly connects the invariant set and the equilibrium set through the existence of a certificate with certain observability properties. In this way, the asymptotic convergence to an equilibrium, i.e., a saddle point, is immediately available. Our observable certificate is weaker than conventional conditions, e.g., strict convexity-concavity and proximal regularization.

We further exploit this certificate to develop a novel separable regularization method where only minimal convexity-concavity is required to establish convergence, a significantly milder condition than most existing ones in the literature, that includes bi-linear saddle functions as a special case. More importantly, the introduced regularization terms are separable and local, thus consistently preserving any distributed structure that original systems may have. The certificate can be generalized to accommodate projections on the vector field and, as a result, allows us to apply our separable regularization to designing distributed solvers for linear programs.

The remainder of the paper is organized as follows. Section II introduces the problem formulation with basic definitions and assumptions, followed by the key results on asymptotic convergence of saddle flow dynamics in Section III. We further generalize the results to the projected cases in Section IV. Section V provides simulation validations and Section VI concludes.

Notation: Let $\mathbb{R}$ and $\mathbb{R}_{\geq 0}$ be the sets of real and non-negative real numbers, respectively. Given two vectors $x, y \in \mathbb{R}^n$, x_i and y_i denote their i^{th} entries, respectively; and $x \leq y$ holds if $x_i \leq y_i$ holds for $\forall i$. Given a continuously differentiable function $S(x, y) \in \mathcal{C}^1$ with $S : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, we use $\frac{\partial}{\partial x} S(x, y) \in \mathbb{R}^{1 \times n}$ and $\frac{\partial}{\partial y} S(x, y) \in \mathbb{R}^{1 \times m}$ to denote the partial derivatives with respect to x and y, respectively. We further define $\nabla_x S(x, y) = \left[\frac{\partial}{\partial x} S(x, y)\right]^T$.

II. PROBLEM FORMULATION

We consider a function $S(x, y)$ with $S : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$. Our goal is to study different dynamic laws that seek to converge to some saddle point $(x_\star, y_\star)$ of $S(x, y)$.

Definition 1 (Saddle Point). *A point $(x_\star, y_\star) \in \mathbb{R}^{n \times m}$ is a saddle point of a function $S(x, y)$ if*

$$\begin{cases} \nabla_x S(x_\star, y_\star) = 0 \\ \nabla_y S(x_\star, y_\star) = 0 \end{cases} \tag{1}$$

holds, and $S(x_\star, y_\star)$ is not a local extremum.

While in general, such questions could be asked on a setting without any further restrictions, neither the existence of saddle points nor convergence towards them is easy to guarantee. For the purpose of this paper, we focus our attention on functions $S(x, y)$ that are *convex-concave*.

Definition 2 (Convex-Concave Functions). *A function $S(x, y)$ is convex-concave if and only if $S(\cdot, y)$ is convex for $\forall y \in R^m$ and $S(x, \cdot)$ is concave for $\forall x \in \mathbb{R}^n$. A function $S(x, y)$ is strictly convex-concave if and only if $S(x, y)$ is convex-concave and either $S(\cdot, y)$ is strictly convex for $\forall y \in R^m$ or $S(x, \cdot)$ is strictly concave for $\forall x \in \mathbb{R}^n$.*

Due to the convexity-concavity of $S(x, y)$, (1) implies that indeed any of its saddle points $(x_\star, y_\star)$ can be characterized by

$$S(x_\star, y) \leq S(x_\star, y_\star) \leq S(x, y_\star), \tag{2}$$

for $\forall x \in \mathbb{R}^n$ and $\forall y \in \mathbb{R}^m$. Therefore, we are specifically interested in minimizing $S(x, y)$ over x and meanwhile maximizing $S(x, y)$ over y.

Throughout this work, we will assume that such a point $(x_\star, y_\star)$ does exist and that $S(x, y)$ is continuously differentiable, i.e., $S(x, y) \in \mathcal{C}^1$, as formally summarize below.

Assumption 1. *$S(x, y)$ is convex-concave, continuously differentiable, and there exists at least one saddle point $(x_\star, y_\star)$ satisfying (2).*

The continuous differentiability in Assumption 1 is introduced to simplify the exposition. It does not significantly limit the scope of the results as one can always derive a continuously differentiable surrogate of a continuous convex/concave function by means of the Moreau Envelope [24].

Given a convex-concave function $S(x, y)$ satisfying Assumption 1, we refer to the following dynamic law

$$\dot{x} = -\nabla_x S(x, y), \tag{3a}$$
$$\dot{y} = +\nabla_y S(x, y), \tag{3b}$$

as the saddle flow dynamics of $S(x, y)$. Due to convexity-concavity, a point is an equilibrium of (3) if and only if it is a saddle point of $S(x, y)$, and the dynamic law drives the system towards such points in the directions of gradient descent and ascent, respectively, for x and y. We will mainly work with this standard form of saddle flow dynamics to locate a saddle point of $S(x, y)$.

III. ASYMPTOTIC CONVERGENCE

In this section we present an observable certificate that guarantees asymptotic convergence of the saddle flows dynamics (3) to a saddle point of $S(x, y)$. We show that two conventional conditions of strict convexity-concavity and proximal regularization satisfy this certificate as special cases. We further build on this certificate to develop a separable regularization method that entails minimal convexity-concavity requirements on $S(x, y)$ for saddle flow dynamics to asymptotically converge to a saddle point.

A. Observable Certificates

We now describe the proposed observable certificate for the saddle flow dynamics (3) to asymptotically converge to a saddle point of $S(x, y)$.

Definition 3 (Observable Certificate). *A function $h(x, y)$ with $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^2_{\geq 0}$ is an observable certificate of $S(x, y)$, if and only if there exists a saddle point $(x_\star, y_\star)$ such that*

$$\begin{bmatrix} S(x_\star, y_\star) - S(x_\star, y) \\ S(x, y_\star) - S(x_\star, y_\star) \end{bmatrix} \geq h(x, y) \geq 0 \tag{4}$$

holds and for any trajectory $(x(t), y(t))$ of (3) that satisfies $h(x(t), y(t)) \equiv 0$, we have $\dot{x}, \dot{y} \equiv 0$.

Remark. *We call $h(x, y)$ an observable certificate, due the second property of Definition 3, which is akin to (3) having $h(x, y)$ as an observable output. It is exactly this observability property that will allow us to connect invariant sets with saddle-points.*

Assumption 2. *$S(x, y)$ has an observable certificate $h(x, y)$ as given by Definition 3.*

Checking whether Assumption 2 holds basically requires hunting for a qualified observable certificate $h(x, y)$ of $S(x, y)$. Under this assumption, asymptotic convergence of the saddle flow dynamics (3) is formally stated below.

Theorem 1 (Sufficiency of Observable Certificates). *Let Assumptions 1 and 2 hold. Then the saddle flow dynamics (3) asymptotically converge to some saddle point $(x_\star, y_\star)$ of $S(x, y)$.*

Proof. The proof follows from applying LaSalle's invariance principle [9] to the following candidate Lyapunov function

$$V(x, y) = \frac{1}{2}\|x - x_\star\|^2 + \frac{1}{2}\|y - y_\star\|^2, \tag{5}$$

where $(x_\star, y_\star)$ is the saddle point identified in Definition 3. Taking the Lie derivative of (5) along the trajectory

$(x(t), y(t))$ of (3) gives

$$\begin{aligned}
\dot{V} &= (x - x_\star)^T \dot{x} + (y - y_\star)^T \dot{y} \\
&= (x - x_\star)^T \left[-\nabla_x S(x, y) \right] + (y - y_\star)^T \left[+\nabla_y S(x, y) \right] \\
&= (x_\star - x)^T \nabla_x S(x, y) - (y_\star - y)^T \nabla_y S(x, y) \\
&\leq S(x_\star, y) - S(x, y) - (S(x, y_\star) - S(x, y)) \\
&= S(x_\star, y) - S(x, y_\star) \\
&= \underbrace{S(x_\star, y) - S(x_\star, y_\star)}_{\leq 0} + \underbrace{S(x_\star, y_\star) - S(x, y_\star)}_{\leq 0},
\end{aligned}$$

where the second equality plugs in (3), the first inequality applies the convexity-concavity of $S(x, y)$, and the last inequality follows from the saddle property (2) of $(x_\star, y_\star)$.

Since (5) is radially unbounded, every sub-level set of it is compact. From above, it follows that the trajectories of (3) are bounded and contained in an invariant domain

$$D_0(x_0, y_0) := \{ (x, y) \mid V(x, y) \leq V(x_0, y_0) \}, \quad (6)$$

where (x_0, y_0) is any given initial point. LaSalle's invariance principle then implies that any trajectory of (3) should converge to the largest invariant set

$$\mathbb{S} := D_0 \cap \left\{ (x, y) \mid \dot{V}(x(t), y(t)) \equiv 0 \right\}. \quad (7)$$

Given Assumption 2, (4) implies that $\mathbb{S}$ is indeed a subset of

$$\{ (x, y) \mid h(x(t), y(t)) \equiv 0 \}, \quad (8)$$

which is further a subset of the equilibrium set of (3), denoted as

$$\mathbb{E} := \{ (x, y) \mid \dot{x}(t), \dot{y}(t) \equiv 0 \}, \quad (9)$$

i.e., $\mathbb{S} \subset \mathbb{E}$.

It follows that the invariant set $\mathbb{S}$ contains only equilibrium points. If $\mathbb{S}$ were to be composed of isolated points – only possible when there is a unique saddle point – this would be sufficient to prove convergence to the (unique) saddle point. However, in general LaSalle's invariance principle only shows asymptotic convergence to the invariant set, without guaranteeing convergence to a point within it, even in the case where the set is composed of equilibrium points.

This issue is circumvented by the fact that all the equilibria within $\mathbb{S}$ are stable. See, e.g., [25, Corollary 5.2]. Alternatively, notice that $\mathbb{S}$ is compact, and as a result any trajectory within the Ω limit set of (3) has a convergent subsequence. Let $(\bar{x}, \bar{y})$ be the limit point of such a sequence. Due to $(\bar{x}, \bar{y}) \in \mathbb{S}$, it is also a saddle point. By changing $(x_\star, y_\star)$ specifically to $(\bar{x}, \bar{y})$ in the definition of $V(x, y)$, it follows that $0 \leq V(x(t), y(t)) \to 0$ holds, which implies $(x(t), y(t)) \to (\bar{x}, \bar{y})$. ∎

The existence and characterization of such observable certificates $h(x, y)$ may still be vague from only Definition 3. We next discuss how they can be identified through concrete examples.

B. Revisiting Existing Conditions

We show that the observable certificate is indeed a weaker condition underneath some of the conventional ones required for asymptotic convergence of the saddle flow dynamics (3).

1) Strict Convexity-Concavity: The most common condition is arguably the strict convexity-concavity of $S(x, y)$ [14]. We formalize its connection with our observable certificate as below.

Assumption 3. $S(x, y)$ *is strictly convex-concave.*

Proposition 2 (Strict Convexity-Concavity). *Let Assumptions 1 and 3 hold. Then the function*

$$h(x, y) := \begin{bmatrix} S(x_\star, y_\star) - S(x_\star, y) \\ S(x, y_\star) - S(x_\star, y_\star) \end{bmatrix}, \quad (10)$$

with $(x_\star, y_\star)$ being an arbitrary saddle point of $S(x, y)$, is an observable certificate of $S(x, y)$.

Proof. Indeed, the upper bound of $h(x, y)$ in (4) itself is naturally a qualified observable certificate. Without loss of generality, we assume $S(x, y)$ is only strictly convex in x for $\forall y$ and just concave in y for $\forall x$. Due to the strict convexity in x, $S(x, y_\star) \equiv S(x_\star, y_\star)$ implies $x(t) \equiv x_\star$ uniquely. Meanwhile, $\nabla_y S(x_\star, y) \equiv 0$ follows from $S(x_\star, y) \equiv S(x_\star, y_\star)$, which leads to $\dot{y} = \nabla_y S(x, y) \equiv \nabla_y S(x_\star, y) \equiv 0$. Therefore, $\dot{x}, \dot{y} \equiv 0$ is guaranteed from $h(x, y) \equiv 0$. ∎

Asymptotic convergence of the saddle flow dynamics (3) then immediately follows from Theorem 1.

Corollary 3. *Let Assumptions 1 and 3 hold. Then the saddle flow dynamics* (3) *asymptotically converge to some saddle point $(x_\star, y_\star)$ of $S(x, y)$.*

2) Proximal Regularization: In the particular form of saddle flow dynamics known as primal-dual dynamics [14], a proximal regularization method is proposed in [18], [22] to guarantee asymptotic convergence of the regularized saddle flow dynamics, even in the absence of strict convexity-concavity. Specifically, a surrogate differentiable convex-concave function

$$\bar{S}(z, y) := \min_x \left\{ S(x, y) + \frac{1}{2} \|x - z\|^2 \right\}$$

is defined from $S(x, y)$ that maintains the same saddle points [22]. Then the following regularized saddle flow dynamics

$$\dot{z} = -\nabla_z \bar{S}(z, y), \quad (11a)$$
$$\dot{y} = +\nabla_y \bar{S}(z, y), \quad (11b)$$

suffice to locate a saddle point. We formalize the connection of this method with our observable certificate as follows.

Proposition 4 (Proximal Regularization). *Let $S(x, y)$ be a Lagrangian function for some constrained convex program and Assumption 1 hold. Then the function*

$$h(z, y) := \begin{bmatrix} \bar{S}(z_\star, y_\star) - \bar{S}(z_\star, y) \\ \frac{1}{2} \|\bar{x}(z, y_\star) - z\|^2 \end{bmatrix}, \quad (12)$$

with $\bar{x}(z, y_\star) := \arg\min_x \left\{ S(x, y_\star) + \frac{1}{2}\|x - z\|^2 \right\}$ and $(z_\star, y_\star)$ being an arbitrary saddle point of $\bar{S}(z, y)$, is an observable certificate of $\bar{S}(z, y)$.

Details of the proof are omitted here and readers are referred to [22] for more insights. We remark that the identification of this observable certificate (12) does not significantly alleviate the analysis overheads since the complementary equilibrium properties of proximal regularization on the Lagrangian $S(x, y)$ are still crucial to validating the observable certificate (12) and establishing convergence.

Anyhow, the existence of an observable certificate satisfies Assumption 2 for $\bar{S}(z, y)$ and thus asymptotic convergence of the saddle flow dynamics (11) follows immediately from Theorem 1.

Corollary 5. *Let $S(x, y)$ be a Lagrangian function for some constrained convex program and Assumption 1 hold. Then the regularized saddle flow dynamics (11) asymptotically converge to some saddle point $(z_\star, y_\star)$ of $\bar{S}(z, y)$, with $(x_\star = z_\star, y_\star)$ being a saddle point of $S(x, y)$.*

In fact, even the differentiability in Assumption 1 is not required since the surrogate $\bar{S}(z, y)$ can be continuously differentiable regardless. However, this proximal regularization method is only limited to the particular form of primal-dual dynamics.

C. Separable Regularization

We further design a novel separable regularization method that exploits our observable certificate and only requires Assumption 1 for a regularized version of saddle flow dynamics to asymptotically converge to a saddle point. The key of this method is to augment the domain of $S(x, y)$ and introduce regularization terms without altering the positions of the original saddle points. In particular, we propose a regularized surrogate for $S(x, y)$ via the following augmentation

$$S(x, z, y, w) := \frac{1}{2\rho}\|x - z\|^2 + S(x, y) - \frac{1}{2\rho}\|y - w\|^2, \quad (13)$$

where $z \in \mathbb{R}^n$ and $w \in \mathbb{R}^m$ serve as two new sets of virtual variables and $\rho > 0$ is a constant regularization coefficient. It is straightforward to verify the fixed positions of saddle points between $S(x, y)$ and $S(x, z, y, w)$ with virtual variables aligned with original variables.

Lemma 6 (Saddle Point Invariance). *Let Assumption 1 hold. Then a point $(x_\star, y_\star)$ is a saddle point of $S(x, y)$ if and only if $(x_\star, z_\star, y_\star, w_\star)$ is a saddle point of $S(x, z, y, w)$, with*

$$x_\star = z_\star \quad and \quad y_\star = w_\star. \quad (14)$$

Proof. Recall the saddle property (2) of a saddle point, this theorem follows immediately from

$$S(x_\star, z_\star, y, w) \leq S(x_\star, x_\star, y_\star, y_\star) \leq S(x, z, y_\star, w_\star)$$

$$\iff S(x_\star, z_\star, y, w) \leq S(x_\star, y_\star) \leq S(x, z, y_\star, w_\star)$$

$$\iff S(x_\star, y) - \frac{\|y - w\|^2}{2\rho} \leq S(x_\star, y_\star) \leq S(x, y_\star) + \frac{\|x - z\|^2}{2\rho}$$

$$\iff S(x_\star, y) \leq S(x_\star, y_\star) \leq S(x, y_\star),$$

where the first and second steps build upon the definition (13) of $S(x, z, y, w)$, and the third step uses norm non-negativity. ∎

The regularized function $S(x, z, y, w)$ is convex in (x, z), concave in (y, w), and continuously differentiable with at least one saddle point, by its definition in (13) and Lemma 6. Therefore, Assumption 1 also holds for $S(x, z, y, w)$. Lemma 6 ensures that whenever we locate a saddle point of $S(x, z, y, w)$, a saddle point of $S(x, y)$ satisfying (2) is attained simultaneously. This motivates us to instead look at the saddle flow dynamics of $S(x, z, y, w)$.

Following (3), this regularized version of saddle flow dynamics are given by

$$\dot{x} = -\nabla_x S(x, y) - \frac{1}{\rho}(x - z), \quad (15a)$$

$$\dot{z} = \frac{1}{\rho}(x - z), \quad (15b)$$

$$\dot{y} = +\nabla_y S(x, y) - \frac{1}{\rho}(y - w), \quad (15c)$$

$$\dot{w} = \frac{1}{\rho}(y - w). \quad (15d)$$

Although this dynamic law has twice as many state variables as its prototype (3), it is important to notice that, unlike the proximal gradient algorithm [18], [22], [24] and the equality constrained regularization [7], [21], (15) still preserves the same distributed structure that (3) may have. As a result, it can be implemented in a fully distributed fashion.

We are now ready to provide the key result that the regularized saddle flow dynamics (15) asymptotically reach a saddle point as long as the minimal convexity-concavity holds for $S(x, y)$.

Proposition 7 (Separable Regularization). *Let Assumption 1 hold. Then the function*

$$h(x, z, y, w) := \begin{bmatrix} \frac{1}{2\rho}\|y - w\|^2 \\ \frac{1}{2\rho}\|x - z\|^2 \end{bmatrix} \quad (16)$$

is an observable certificate of $S(x, z, y, w)$.

Proof. The above observable certificate $h(x, z, y, w)$ satisfies (4) in light of the following calculation:

$$\begin{bmatrix} S(x_\star, z_\star, y_\star, w_\star) - S(x_\star, z_\star, y, w) \\ S(x, z, y_\star, w_\star) - S(x_\star, z_\star, y_\star, w_\star) \end{bmatrix}$$

$$\geq \begin{bmatrix} \underbrace{S(x_\star, y_\star) - S(x_\star, y)}_{\geq 0} + \frac{1}{2\rho}\|y - w\|^2 \\ \underbrace{S(x, y_\star) - S(x_\star, y_\star)}_{\geq 0} + \frac{1}{2\rho}\|x - z\|^2 \end{bmatrix}$$

$$\geq \begin{bmatrix} \frac{1}{2\rho}\|y - w\|^2 \\ \frac{1}{2\rho}\|x - z\|^2 \end{bmatrix}$$

$$\geq 0.$$

The fact that $h(x, z, y, w) \equiv 0$ implies $x(t) \equiv z(t)$ and $y(t) \equiv w(t)$ enforces $\dot{z}, \dot{w} \equiv 0$ according to (15b), (15d), and then $\dot{x}, \dot{y} \equiv 0$ is simultaneously guaranteed. ∎

Assumption 2 holds for the regularized function $S(x, z, y, w)$ and asymptotic convergence of the regularized

saddle flow dynamics (15) follows immediately from Theorem 1.

Corollary 8. *Let Assumption 1 hold. Then the regularized saddle flow dynamics* (15) *asymptotically converge to some saddle point* $(x_\star, z_\star, y_\star, w_\star)$ *of* $S(x, z, y, w)$, *with* $(x_\star, y_\star)$ *being a saddle point of* $S(x, y)$.

Remark. *Proposition 7 indicates that only the convexity-concavity of* $S(x, y)$ *is required to asymptotically arrive at a saddle point through the regularized saddle flow dynamics* (15). *This condition is significantly milder than most existing ones in the literature, and is in some sense* minimal, *as it includes bi-linear saddle functions as a special case. Unlike the aforementioned proximal regularization method in Section III-B, our separable regularization method applies to saddle flow dynamics of general convex-concave functions.*

IV. PROJECTED SADDLE FLOW DYNAMICS

In this section we generalize the results in Section III to account for projections on the vector field of the saddle flow dynamics (3) that are commonly introduced in the case of solving inequality constrained optimization problems.

Specifically, we look at a projected version of saddle flow dynamics of a convex-concave function $S(x, y)$ as below:

$$\dot{x} = -\nabla_x S(x, y), \tag{17a}$$
$$\dot{y} = [+\nabla_y S(x, y)]_y^+, \tag{17b}$$

where, without loss of generality, we define the element-wise projection $[\cdot]_y^+$ only on part of the vector field regarding y as

$$[\nabla_{y_i} S(x, y)]_{y_i}^+ := \begin{cases} \nabla_{y_i} S(x, y), & \text{if } y_i > 0, \\ \max\{\nabla_{y_i} S(x, y), 0\}, & \text{otherwise.} \end{cases} \tag{18}$$

With this projection, $y(t)$ is constrained to be non-negative as long as it starts with a non-negative initial point. Accordingly, we slightly modify Assumption 1 to guarantee the existence of such saddle points.

Assumption 4. $S(x, y)$ *is convex-concave, continuously differentiable, and there exists at least one saddle point* $(x_\star, y_\star \geq 0)$ *satisfying* (2).

In this context, saddle points are restrained to ones in the non-negative orthant of y. Therefore, any observable certificate of $S(x, y)$ will be defined on a saddle point $(x_\star, y_\star \geq 0)$ in Definition 3. By symmetry, an analogous projection may also be imposed on the other part of the vector field regarding x. Next we formally generalize the sufficiency of observable certificates developed in Section III-A.

A. Observable Certificates for Projected Flows

The generalization of Theorem 1 for asymptotic convergence of the projected saddle flow dynamics (17) to a saddle point of $S(x, y)$ is summarized as follows.

Theorem 9 (Sufficiency of Observable Certificates for Projected Flows)**.** *Let Assumptions 2 and 4 hold. Then the*

projected saddle flow dynamics (17) *asymptotically converge to some saddle point* $(x_\star, y_\star \geq 0)$ *of* $S(x, y)$.

The proof requires a lemma regarding the projection $[\cdot]_y^+$.

Lemma 10. *Given any arbitrary* $y, y_\star \in \mathbb{R}_{\geq 0}^m$ *and* $\nu \in \mathbb{R}^m$,

$$(y - y_\star)^T \left([\nu]_y^+ - \nu\right) \leq 0$$

holds.

Proof. The proof of this lemma follows from the fact that element-wise, $[\nu_i]_{y_i}^+$ differs from ν_i only in the case of $\nu_i < 0$ and $y_i = 0$ where the projection is active, which implies

$$(y - y_\star)^T \left([\nu]_y^+ - \nu\right) = \sum_{i:\nu_i < 0, y_i = 0} (0 - y_{i*})(0 - \nu_i) \leq 0.$$

∎

Using this lemma, the proof of Theorem 9 essentially follows from that of Theorem 1 as follows.

Proof. Consider the same quadratic Lyapunov function (5). Taking its Lie derivative along the trajectory $(x(t), y(t))$ of (17) yields

$$\begin{aligned} \dot{V} &= (x - x_\star)^T \dot{x} + (y - y_\star)^T \dot{y} \\ &= (x - x_\star)^T [-\nabla_x S(x, y)] + (y - y_\star)^T [+\nabla_y S(x, y)]_y^+ \\ &= (x_\star - x)^T \nabla_x S(x, y) - (y_\star - y)^T \nabla_y S(x, y) \\ &\quad + \underbrace{(y - y_\star)^T \left([\nabla_y S(x, y)]_y^+ - \nabla_y S(x, y)\right)}_{\leq 0} \\ &\leq S(x_\star, y) - S(x, y) - (S(x, y_\star) - S(x, y)) \\ &= S(x_\star, y) - S(x, y_\star) \\ &= \underbrace{S(x_\star, y) - S(x_\star, y_\star)}_{\leq 0} + \underbrace{S(x_\star, y_\star) - S(x, y_\star)}_{\leq 0}, \end{aligned}$$

where the key step is to use Lemma 10 in the first inequality. The rest of the proof remains almost the same except that the largest invariant set is defined between the on-off switches of the projection. From above, $\dot{V}(x, y) \equiv 0$ additionally implies

$$y(t) \equiv y_\star$$

or

$$[\nabla_y S(x, y)]_y^+ \equiv \nabla_y S(x, y),$$

and an invariance principle for Caratheodory systems [26] can be applied to account for the discontinuities in the vector field due to the projection. ∎

B. Application: Distributed Solution to Linear Program

Theorem 9 enables the separable regularization method in Section III-C to apply to projected saddle flow dynamics as well since we can still identify the same observable certificate

$$h(x, z, y, w) := \begin{bmatrix} \frac{1}{2\rho} \|y - w\|^2 \\ \frac{1}{2\rho} \|x - z\|^2 \end{bmatrix}$$

to satisfy Assumption 2. One of its straightforward applications involves solving inequality constrained linear programs

(a) Control inputs. (b) State variables. (c) Dual variables. (d) System evolution.
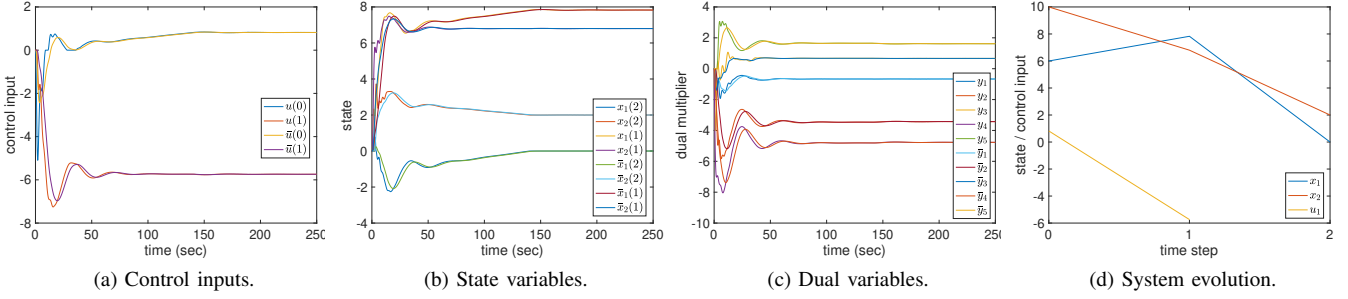
Fig. 1: A two-state linear optimal control problem solved by saddle flow dynamics with separable regularization.

in a distributed fashion with guaranteed asymptotic convergence to an optimal solution.

Consider the following problem:

$$\min_{x \in \mathbb{R}^n} \quad c^T x \tag{19a}$$

$$\text{s.t.} \quad Ax - b \leq 0 \; : \; y \in \mathbb{R}_{\geq 0}^m \tag{19b}$$

which corresponds to a bi-linear Lagrangian

$$S(x,y) := c^T x + y^T (Ax - b).$$

We introduce virtual variables $z \in \mathbb{R}^n$, $w \in \mathbb{R}^m$ and a constant $\rho > 0$ to define

$$S(x,z,y,w) := \frac{1}{2\rho}\|x-z\|^2 + c^T x + y^T(Ax-b) - \frac{1}{2\rho}\|y-w\|^2$$

to be its augmented Lagrangian. Lemma 6 implies that $(x_\star, y_\star \geq 0)$ is a saddle point of $S(x,y)$, i.e., one optimal solution to (19), if and only if $(x_\star, z_\star = x_\star, y_\star \geq 0, w_\star = y_\star)$ is a saddle point of $S(x,z,y,w)$.

Then an algorithm to optimally solve a linear program of the form (19) follows immediately from asymptotic convergence of the following projected and regularized saddle flow dynamics:

$$\dot{x} = -c - A^T y - \frac{1}{\rho}(x - z), \tag{20a}$$

$$\dot{z} = \frac{1}{\rho}(x - z), \tag{20b}$$

$$\dot{y} = \left[Ax - b - \frac{1}{\rho}(y - w)\right]_y^+, \tag{20c}$$

$$\dot{w} = \frac{1}{\rho}(y - w), \tag{20d}$$

which maintains the distributed structure where each agent $i = 1, 2, \ldots, n$ may locally manage

$$\dot{x}_i = -c_i - A_i^T y - \frac{1}{\rho}(x_i - z_i), \tag{21a}$$

$$\dot{z}_i = \frac{1}{\rho}(x_i - z_i), \tag{21b}$$

and/or each dual agent $j = 1, 2, \ldots, m$ may locally manage

$$\dot{y}_j = \left[A_j x - b_j - \frac{1}{\rho}(y_j - w_j)\right]_{y_j}^+, \tag{21c}$$

$$\dot{w}_j = \frac{1}{\rho}(y_j - w_j), \tag{21d}$$

with A_i and A_j being the ith column and the jth row of A, respectively.

V. SIMULATION RESULTS

We illustrate asymptotic convergence of the distributed algorithm (21) for linear programs, as guaranteed by our observable certificate, through a finite-horizon optimal control problem over a group of agents with coupled dynamics [7]. The goal is to use minimal aggregate control effort to maintain small system states subject to certain final performance requirements. In particular, consider the following problem defined on the time horizon $\mathcal{T} := \{0, 1, \ldots, T\}$ and the set $\mathcal{N} := \{1, 2, \ldots, N\}$ of agents:

$$\min \quad \sum_{t \in \mathcal{T}} \|x(t+1)\|_1 + \|u(t)\|_1 \tag{22a}$$

$$\text{s.t.} \quad x(t+1) = Gx(t) + Hu(t), \; t \in \mathcal{T} \tag{22b}$$

$$Dx(T+1) - d \leq 0 \tag{22c}$$

Here $x(t) \in \mathbb{R}^N$ and $u(t) \in \mathbb{R}^N$ are respectively vectors of system states and control inputs at time t. $G \in \mathbb{R}^{N \times N}$ and $H \in \mathbb{R}^{N \times N}$ are the time-invariant dynamics matrix and control matrix, respectively. (22c) specifies the constraint on the final state with given $D \in \mathbb{R}^{M \times N}$ and $d \in \mathbb{R}^M$. The initial state $x(0)$ is fixed. This problem can be cast into the standard form of a linear program by splitting each variable of state and control input into two, representing its positive and negative components, e.g., $x(t) = x^+(t) - x^-(t)$ with $x^+(t), x^-(t) \geq 0$:

$$\min \quad \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} x_n^+(t+1) + x_n^-(t+1) + u_n^+(t) + u_n^-(t)$$

$$\text{s.t.} \quad x^+(t+1) - x^-(t+1) = G\left(x^+(t) - x^-(t)\right)$$
$$+ H\left(u^+(t) - u^-(t)\right), \; t \in \mathcal{T}$$
$$D\left(x^+(T+1) - x^-(T+1)\right) - d \leq 0$$
$$x^+(t+1), x^-(t+1), u^+(t), u^-(t) \geq 0, \; t \in \mathcal{T}$$

It can be verified that at the optimum the split twin components of any variable cannot be simultaneously positive [27]. We apply (21) to solve this linear program with dual variables $y \in \mathbb{R}^{NT+M}$ and virtual variables $\bar{x}(t) \in \mathbb{R}^N$, $\bar{u}(t) \in \mathbb{R}^N$, $\bar{y} \in \mathbb{R}^{NT+M}$. Each agent n is responsible for $4T$ variables, $2T$ for x_n and $2T$ for u_n, while up to $4NT$ variables with any consensus-based distributed algorithm. Note that the dual

variable updates may also be locally managed by up to $NT + M$ dual agents, if necessary.

We test a specific case of $N = 2$ agents in a time horizon of $T = 2$ time slots with the initial state $x(0) = [6\ 10]^T$. The system is driven by a single control input on the first agent's state. More specifically, we set the regularization coefficient ρ to be 3, and the problem parameters are given by

$$G := \begin{bmatrix} 1.1 & 0 \\ -0.7 & 1.1 \end{bmatrix}, H := \begin{bmatrix} 1.5 & 0 \\ 0 & 0 \end{bmatrix}, D := \begin{bmatrix} 1 \\ 1.5 \end{bmatrix}^T, d := 3.$$

The evolution of the continuous-time saddle flow dynamics for the control inputs, system states, dual multipliers as well as their virtual counterparts is displayed in Figures 1(a)-1(c), respectively. For each primal variable, we have already combined its two split components back together for conciseness.

Apparently, the regularization renders a variable bound tightly with its virtual counterpart in transient, and they asymptotically converge to the same value of an optimal solution in the limit. Notably, y_5, corresponding to (22c), in Figure 1(c) is the only dual variable subject to the projection that strictly stays non-negative. Once all the variables are determined, the control inputs can be implemented to attain the system evolution with the minimal aggregate control effort and system states (in the sense of 1-norm), as depicted in Figure 1(d). This is consistent with the optimal solution acquired from any centralized linear programming solver:

$$u_\star(0) = \begin{bmatrix} 0.8190 \\ 0 \end{bmatrix}, \ u_\star(1) = \begin{bmatrix} -5.7410 \\ 0 \end{bmatrix},$$
$$x_\star(1) = \begin{bmatrix} 7.8286 \\ 6.8000 \end{bmatrix}, \ x_\star(2) = \begin{bmatrix} 0 \\ 2.0000 \end{bmatrix}.$$

VI. Conclusion

This paper proposes an observable certificate that directly establishes connection between the invariant set and the equilibrium set for saddle flow dynamics of a convex-concave function such that the asymptotic convergence to a saddle point can be guaranteed. The certificate is rooted in observability, and we identify the existence of such observable certificates in the presence of conventional conditions, e.g., strict convexity-concavity and proximal regularization, as well as the proposed separable regularization method. Therefore, our observable certificate is a weaker condition, and it further generalizes to situations with projections on the vector field of saddle flow dynamics. Besides, the novel separable regularization method that builds on our observable certificate requires only minimal convexity-concavity to establish convergence and enjoys a separable structure for potential distributed implementation, which is demonstrated through an application to distributed linear programming.

References

[1] T. J. Holding and I. Lestas, "Stability and instability in saddle point dynamics - Part I," *IEEE Transactions on Automatic Control*, pp. 1–1, 2020.

[2] T. J. Holding and I. Lestas, "Stability and instability in saddle point dynamics - Part II: The subgradient method," *IEEE Transactions on Automatic Control*, pp. 1–1, 2020.

[3] A. Wibisono, A. C. Wilson, and M. I. Jordan, "A variational perspective on accelerated methods in optimization," *proceedings of the National Academy of Sciences*, vol. 113, no. 47, pp. E7351–E7358, 2016.

[4] G. França, D. P. Robinson, and R. Vidal, "Gradient flows and accelerated proximal splitting methods," *arXiv preprint arXiv:1908.00865*, 2019.

[5] H. Mohammadi, M. Razaviyayn, and M. R. Jovanovic, "Robustness of accelerated first-order algorithms for strongly convex optimization problems," *IEEE Transactions on Automatic Control*, 2020.

[6] L. Lessard, B. Recht, and A. Packard, "Analysis and design of optimization algorithms via integral quadratic constraints," *SIAM Journal on Optimization*, vol. 26, no. 1, pp. 57–95, 2016.

[7] D. Richert and J. Cortés, "Robust distributed linear programming," *IEEE Transactions on Automatic Control*, vol. 60, no. 10, pp. 2567–2582, 2015.

[8] A. Cherukuri, E. Mallada, S. Low, and J. Cortes, "The role of convexity in saddle-point dynamics: Lyapunov function and robustness," *IEEE Transactions on Automatic Control*, vol. 63, no. 8, pp. 2449–2464, 2017.

[9] H. K. Khalil and J. W. Grizzle, *Nonlinear systems*. Prentice hall Upper Saddle River, NJ, 2002, vol. 3.

[10] J. CortéS, "Finite-time convergent gradient flows with applications to network consensus," *Automatica*, vol. 42, no. 11, pp. 1993–2000, 2006.

[11] K. Garg and D. Panagou, "Fixed-time stable gradient flows: Applications to continuous-time optimization," *IEEE Transactions on Automatic Control*, 2020.

[12] S. Rahili and W. Ren, "Distributed continuous-time convex optimization with time-varying cost functions," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1590–1605, 2016.

[13] M. Fazlyab, S. Paternain, V. M. Preciado, and A. Ribeiro, "Prediction-correction interior-point method for time-varying convex optimization," *IEEE Transactions on Automatic Control*, vol. 63, no. 7, pp. 1973–1986, 2017.

[14] A. Cherukuri, E. Mallada, and J. Cortés, "Asymptotic convergence of constrained primal–dual dynamics," *Systems & Control Letters*, vol. 87, pp. 10–15, 2016.

[15] B. Gharesifard and J. Cortes, "Distributed convergence to Nash equilibria in two-network zero-sum games," *Automatica*, vol. 49, no. 6, pp. 1683–1692, 2013.

[16] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1177–1189, 2014.

[17] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as optimization decomposition: A mathematical theory of network architectures," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 255–312, 2007.

[18] D. Goldsztajn, F. Paganini, and A. Ferragut, "Proximal optimization for resource allocation in distributed computing systems with data locality," in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2019, pp. 773–780.

[19] K. J. Arrow, H. Azawa, L. Hurwicz, and H. Uzawa, *Studies in linear and non-linear programming*. Stanford University Press, 1958, vol. 2.

[20] A. Cherukuri, B. Gharesifard, and J. Cortes, "Saddle-point dynamics: Conditions for asymptotic stability of saddle points," *SIAM Journal on Control and Optimization*, vol. 55, no. 1, pp. 486–511, 2017.

[21] A. Cherukuri, A. D. Domínguez-García, and J. Cortés, "Distributed coordination of power generators for a linearized optimal power flow problem," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 3962–3967.

[22] D. Goldsztajn and F. Paganini, "Proximal regularization for the saddle point gradient dynamics," *Working Paper*, 2020.

[23] S. Yamashita, T. Hatanaka, J. Yamauchi, and M. Fujita, "Passivity-based generalization of primal–dual dynamics for non-strictly convex cost functions," *Automatica*, vol. 112, p. 108712, 2020.

[24] N. Parikh and S. Boyd, "Proximal algorithms," *Foundations and Trends in optimization*, vol. 1, no. 3, pp. 127–239, 2014.

[25] S. P. Bhat and D. S. Bernstein, "Nontangency-based Lyapunov tests for convergence and stability in systems having a continuum of equilibria," *SIAM Journal on Control and Optimization*, vol. 42, no. 5, pp. 1745–1775, 2003.

[26] A. Bacciotti and F. Ceragioli, "Nonpathological Lyapunov functions and discontinuous Carathéodory systems," *Automatica*, vol. 42, no. 3, pp. 453–458, 2006.

[27] G. B. Dantzig and M. N. Thapa, *Linear programming I: Introduction*. Springer Science & Business Media, 2006.