IEEE Control Systems Letters paper presented at
2023 American Control Conference (ACC)
San Diego, CA, USA. May 31 - June 2, 2023

# On the global exponential stability of primal-dual dynamics for convex problems with linear equality constraints

Ibrahim K. Ozaslan and Mihailo R. Jovanović

***Abstract*— We examine global exponential stability of the primal-dual gradient flow dynamics for differentiable convex problems with linear equality constraints. We show that if the set of equilibrium points is affine, then, regardless of the initial conditions, trajectories of the gradient flow move in the direction that is perpendicular to equilibrium set. When the objective function is strongly convex, we utilize this structure to show that the primal-dual dynamics are globally exponentially stable even if the constraint matrix is not full-row rank. We also provide an explicit characterization of the exponential convergence rate in terms of the smallest nonzero singular value of the constraint matrix.**

***Index Terms*— Global exponential stability, gradient flow dynamics, Lyapunov functions, primal-dual methods.**

## I. Introduction

We study convex problems with linear equality constraints,

$$\begin{aligned} \underset{x}{\text{minimize}} \;\; & f(x) \\ \text{subject to} \;\; & Ex - q = 0 \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$ is the optimization variable, $f : \mathbb{R}^n \to \mathbb{R}$ is the convex objective function, $E \in \mathbb{R}^{m\times n}$ is the constraint matrix with $\operatorname{rank}(E) = r < \min(m, n)$, and $q \in \mathbb{R}^m$ is a given vector. In this work, we are interested in characterizing stability properties of primal-dual gradient flow dynamics,

$$\begin{aligned} \dot{x} &= -\nabla_x \mathcal{L}(x; y) \\ \dot{y} &= +\nabla_y \mathcal{L}(x; y) \end{aligned} \tag{2}$$

where $\mathcal{L}$ is the Lagrangian defined as

$$\mathcal{L}(x; y) = f(x) + \langle y,\, Ex - q\rangle \tag{3}$$

with $y \in \mathbb{R}^m$ being the vector of dual variables associated with the linear equality constraint.

Stability and convergence properties of primal-dual methods synthesized from gradient flows in the form of (2) have been continually studied under various scenarios for diverse problems since the introduction of dynamics (2) in the seminal paper [1]. Early results such as [2], [3] were concentrating on the global asymptotic stability of projected primal-dual dynamics derived from the Lagrangian of inequality constrained problems. Then in [4]–[7], asymptotic stability of (projected) primal-dual gradient flow was established for general saddle functions under different assumptions. In more recent works, the focus has started to shift toward the exponential stability. In [8]–[13] global exponential stability of the primal-dual gradient flow was proved for problem (1). Also [14] and [15] showed that gradient flow is contractive for problem (1).

However, all these results proving exponential stability assume that, in addition to $f$ being strongly convex, constraint matrix $E$ is full-row rank. This condition on $E$ may sound natural because one can think that redundant rows/constraints can be deleted before running the algorithm, but for large scale problems, obtaining a set of independent rows may not be a feasible operation due to high computational or communication cost. Consider for example solving undetermined linear systems of equations, which can be brought into the form of problem (1) by setting $f$ to $\|\cdot\|^2$, in a distributed architecture where constraint matrix $E$ is too large to fit in a single memory hence stored in a network of memory nodes. In that case, finding a set of independent rows would be cumbersome because of excessive communication cost. In such scenarios, from the current results, it is not clear whether we have exponential stability if we start with a rank deficient constraint matrix $E$. This motivates the following question: *Are the primal-dual gradient flow dynamics globally exponentially stable for problem* (1) *with strongly convex objective function even if the constraint matrix $E$ is rank deficient?*

Recently, [16] proved existence of a global exponential convergence rate for primal-dual gradient flow applied to a class of problems including (1) without having any assumption on the constraint matrix; but this result does not imply global exponential stability of (2). In [17], an explicit characterization of linear convergence rates were obtained for two different primal-dual methods that are derived from dynamics (2) via two different discretization schemes.

In this paper, we prove the global exponential stability, in the sense that there exists a decaying exponential upper bound on the distance to the set of equilibrium points, of *continuous-time* dynamics (2) for problem (1) without having any assumption on constraint matrix $E$. Our analysis also explicitly characterize the convergence rate. Besides exponential stability, as different from [17], our aim is to draw attention to an interesting phenomenon that we have not been aware, namely, if the set of equilibrium points is an affine set, then the trajectories of (2) move perpendicular to the equilibria. By using this fact, existing exponential stability analyses can be modified to remove the full row-

Financial support from the National Science Foundation under awards ECCS-1739210 and ECCS-1809833 is gratefully acknowledged.

I. K. Ozaslan and M. R. Jovanović are with the Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089. E-mails: ozaslan@usc.edu, mihailo@usc.edu.

rank assumption on $E$, and even strong convexity assumption can be relaxed for quadratic terms in the objective function. The latter is implied by the equivalence of asymptotic and exponential stability in linear systems [18, Theorem 4.11].

The rest of the paper is organized as follows. In Section II, we provide the background material. In Section III, we examine the aforementioned phenomenon and in Section IV utilize this phenomenon to prove the global exponential stability of (2) without having any assumptions on $E$. Lastly, we demonstrate validity of our findings through computational experiments in Section V and conclude the paper in Section VI with remarks.

*Notation:* $\|\cdot\|$ denotes the Euclidian norm. The range space and null space of matrix $A$ are denoted by $\mathcal{R}(A)$ and $\mathcal{N}(A)$, respectively. Scalars $\overline{\sigma}(A)$ and $\underline{\sigma}(A)$ represent the largest and smallest nonzero singular values of $A$, respectively.

## II. Background

We assume that $q \in \mathcal{R}(E)$, otherwise problem (1) is not feasible. Consequently, the Karush-Kuhn-Tucker conditions

$$\nabla f(x^\star) \;=\; -\,E^T y^\star, \quad E x^\star \;=\; q \tag{4}$$

are necessary and sufficient for $(x^\star, y^\star)$ to be a primal-dual solution pair to (1) [19]. When $f$ is strongly convex, problem (1) has a unique solution $x^\star$ and the set of optimal dual variables $\mathcal{Y}^\star$ is not a singleton but an affine set,

$$\begin{aligned}\mathcal{Y}^\star \;&=\; \{y \,\in\, \mathbb{R}^m \,|\, \nabla f(x^\star) \,=\, -E^T y\}\\ &=\; \{y \,\in\, \mathbb{R}^m \,|\, y \,=\, y_0^\star \,+\, U_2 w,\;\; w \,\in\, \mathbb{R}^{m-r}\}.\end{aligned}$$

Here, $y_0^\star$ is the unique vector in $\mathcal{R}(E)$ that satisfies $\nabla f(x^\star) = -E^T y_0^\star$, the columns of $U_2$ form an orthonormal basis of $\mathcal{N}(E^T)$, and $E$ has the singular value decomposition

$$E \;=\; \left[\begin{array}{cc} U_1 & U_2 \end{array}\right] \left[\begin{array}{cc} \Sigma & 0 \\ 0 & 0 \end{array}\right] \left[\begin{array}{c} V_1^T \\ V_2^T \end{array}\right]$$

where $\Sigma = \mathbf{diag}(\sigma_1, \ldots, \sigma_r)$ and $\sigma_1 \geq \cdots \geq \sigma_r > 0$.

To compute a solution pair $(x^\star, y^\star)$ for some $y^\star \in \mathcal{Y}^\star$, we cannot directly solve the system of equations (4) because of the nonlinear term $\nabla f$. Instead, we use an alternative characterization of the optimal primal-dual pairs: every solution $(x^\star, y^\star)$ of (4) is a saddle point of the Lagrangian satisfying

$$\mathcal{L}(x^\star; y) \;\leq\; \mathcal{L}(x^\star; y^\star) \;\leq\; \mathcal{L}(x; y^\star), \quad \forall x, y.$$

Based on this characterization, a solution to problem (1) can be computed by simultaneous minimization and maximization of the Lagrangian over primal variable $x$ and dual variable $y$, respectively. To this end, we deploy primal-dual gradient flow dynamics (2), which can be considered as a natural generalization of gradient flow dynamics $\dot{x} = \nabla f(x)$ to saddle functions,

$$\dot{x} \;=\; -\,\nabla_x \mathcal{L}(x; y) \;=\; -\,\nabla f(x) \,-\, E^T y \tag{5a}$$

$$\dot{y} \;=\; +\,\nabla_y \mathcal{L}(x; y) \;=\; E x \,-\, q \tag{5b}$$

where $x$: $[0, \infty) \to \mathbb{R}^n$ and $y$: $[0, \infty) \to \mathbb{R}^m$ with arbitrary initial conditions $x(0)$ and $y(0)$. The equilibrium points of these *primal-descent dual-ascent* dynamics, where $\dot{x} = 0$ and $\dot{y} = 0$, identify the saddle points of the Lagrangian that are characterized by conditions (4). Thus, we can use these dynamics to compute a solution to problem (1). In what follows, we examine the stability properties of these dynamics.

## III. Pointwise Convergence and Perpendicular Motion of Primal-Dual Dynamics

In this section, we show that the trajectories generated by (5) move perpendicular to the set of equilibria. This property is utilized in the remainder of the paper to prove that dynamics (5) are globally exponentially stable even for rank-deficient constraint matrix $E$.

*Assumption 1:* The function $f$ is $m_f$-strongly convex with an $L_f$-Lipschitz continuous gradient $\nabla f$.

*Theorem 1:* Let $q \in \mathcal{R}(E)$ and Assumption 1 hold. For any initial condition $y(0) \in \mathbb{R}^m$, the dual trajectory $y(t)$ is perpendicular to $\mathcal{Y}^\star$, i.e., for any $t \geq 0$

$$\dot{y}(t) \;\perp\; \{y \,-\, y^\star \,|\, y \in \mathcal{Y}^\star\} \text{ for all } y^\star \,\in\, \mathcal{Y}^\star.$$

Furthermore, $y(t)$ converges to $\bar{y}^\star \in \mathcal{Y}^\star$, where

$$\bar{y}^\star \;=\; \underset{y^\star \,\in\, \mathcal{Y}^\star}{\operatorname{argmin}} \;\|y(0) \,-\, y^\star\|. \tag{6}$$

*Proof:* The global asymptotic stability of (5) can be shown via LaSalle's invariance principle by assuming only strict[1] monotonicity of $\nabla f$ [2], [10], [20]. Moreover, if $\nabla f$ is Lipschitz continuous, then the gradient of Lagrangian (3) is also Lipschitz continuous. The global asymptotic stability together with Lipschitz continuity of the Lagrangian implies the pointwise convergence of (5), i.e., there exists a $\bar{y}^\star \in \mathcal{Y}^\star$ such that $\bar{y}^\star = \lim_{t\to\infty} y(t)$ [5, Lemma A.3].

We use fundamental theorem of calculus to prove (6)

$$\begin{aligned}\bar{y}^\star \;&=\; \lim_{t\,\to\,\infty} y(t) \;=\; y(0) \,+\, \int_0^\infty \dot{y}(\tau)\mathrm{d}\tau\\ &=\; U_1 U_1^T y(0) \,+\, U_2 U_2^T y(0) \,+\, \int_0^\infty \dot{y}(\tau)\mathrm{d}\tau\end{aligned} \tag{7}$$

where the second equality follows from the orthogonal decomposition of $y(0)$ over $\mathcal{R}(E)$ and $\mathcal{N}(E^T)$. Now, our feasibility assumption $q \in \mathcal{R}(E)$ together with (5b) implies $\dot{y}(t) \in \mathcal{R}(E)$, hence $\dot{y}(t) = U_1 U_1^T \dot{y}(t)$ for $t \geq 0$. Furthermore, since $U_2 U_2^T y(0) \perp \mathcal{R}(E)$, equation (7) together with $\bar{y}^\star \in \mathcal{Y}^\star$ implies that the component of $\bar{y}^\star$ in $\mathcal{N}(E^T)$ is given by $U_2 U_2^T y(0)$ and

$$y_0^\star \;=\; U_1 U_1^T y(0) \,+\, \int_0^\infty \dot{y}(\tau)\mathrm{d}\tau \tag{8}$$

thus we can use (8) to express $\bar{y}^\star$,

$$\bar{y}^\star \;=\; y_0^\star \,+\, U_2 U_2^T y(0) \tag{9}$$

[1] Solely monotonicity would suffice to invoke LaSalle's invariance principles if the dynamics were based on the augmented Lagrangian. In that case however, the dynamics may not be as convenient as (5) for implementation in distributed environments.

which is exactly the orthogonal projection of $y(0)$ onto affine set $\mathcal{Y}^\star$ as given in (6). Finally, the fact that $\dot{y} \in \mathcal{R}(E)$ combined with $\{y - y^\star \,|\, y \in \mathcal{Y}^\star\} = \mathcal{N}(E^T)$ for any $y^\star \in \mathcal{Y}^\star$ shows that the dual trajectories generated by dynamics (5) are perpendicular to affine set $\mathcal{Y}^\star$. ■

We note that the perpendicular motion is not result of solely linear dynamics. Even for linear dynamics, if the set of equilibrium points is not affine (e.g., union of affine sets when the objective function in (1) contains non-differentiable terms), then we would not have this property because we would not be able to use the orthogonal decomposition in (7).

## IV. Global Exponential Stability

In this section, we prove that the primal-dual dynamics are globally exponentially stable even if the constraint matrix $E$ is rank-deficient. For problems with twice differentiable $f$, we propose the Lyapunov function candidate

$$V(x,y) \;=\; \frac{1}{2}(\|\nabla\mathcal{L}(x;y)\|^2 + \|x - x^\star\|^2 + \mathbf{dist}^2(y,\ \mathcal{Y}^\star)) \tag{10}$$

where the distance function is defined as

$$\mathbf{dist}(y,\ \mathcal{Y}^\star) \;:=\; \underset{y^\star\,\in\,\mathcal{Y}^\star}{\text{minimize}}\, \|y \,-\, y^\star\|.$$

Norm of gradient is frequently used to characterize convergence rate of first-order optimization algorithms for strongly convex problems. In particular, the square of the norm of $\nabla\mathcal{L}$ was used in [13] as a Lyapunov function to prove exponential stability of a second order primal-dual method based on proximal augmented Lagrangian for non-smooth composite optimization problems. Here, we augment $\|\nabla\mathcal{L}(x;y)\|^2$ with a distance term in order to establish negative definiteness of the time derivative along the solutions of primal-dual gradient flow dynamics.

*Remark 1:* The twice differentiability assumption on $f$ can be relaxed by utilizing different Lyapunov functions; e.g., the quadratic function proposed in [8] or the nonlinear function used in [9] instead of (10). Nevertheless, in our study we use $V$ given in (10) to provide an alternative point of view and possibly novel insight about the synthesis of Lyapunov functions for general saddle function problems.

*Theorem 2:* Let Assumption 1 hold. Then, the primal-dual gradient flow dynamics (5) are globally exponentially stable with convergence rate

$$\rho \;=\; \frac{2m_f \min(m_f^2,\ \underline{\sigma}^2(E))}{(L_f^2 \,+\, \overline{\sigma}^2(E) \,+\, 1)(1 \,+\, 2m_f L_f).}$$

*Proof:* We start our analysis by noting that because of perpendicular motion of dual trajectories, the time dependence of the distance between $y(t)$ and $\mathcal{Y}^\star$ comes from the time dependence of solely $y(t)$ but not from the projection of $y(t)$ onto $\mathcal{Y}^\star$; as shown in Theorem 1, this projection is fixed. Thus, we have

$$\mathbf{dist}(y,\ \mathcal{Y}^\star) \;=\; \|y \,-\, \bar{y}^\star\|. \tag{11}$$

To prove the theorem, we show that for all $t \geq 0$

$$\dot{V}(t) \;\leq\; -\rho V(t) \tag{12}$$

which implies

$$\|x - x^\star\|^2 \,+\, \mathbf{dist}^2(y,\ \mathcal{Y}^\star) \;\leq\; 2V(x(0), y(0))\, \mathrm{e}^{-\rho t}.$$

To this end, we rewrite the dynamics and Lyapunov function in terms of errors $\tilde{x} := x - x^\star$, $\tilde{y} := y - \bar{y}^\star$, and $\widetilde{\nabla f} := \nabla f(x) - \nabla f(x^\star)$. Since $\nabla\mathcal{L}(x^\star; \bar{y}^\star) = 0$, the dynamics can be written as

$$\begin{aligned}\dot{x} &\;=\; \nabla_x\mathcal{L}(x;y) \,-\, \nabla_x\mathcal{L}(x^\star;\bar{y}^\star) \;=\; -\,\widetilde{\nabla f} \,-\, E^T\tilde{y}\\ \dot{y} &\;=\; \nabla_y\mathcal{L}(x;y) \,-\, \nabla_y\mathcal{L}(x^\star;\bar{y}^\star) \;=\; E\tilde{x}\end{aligned}$$

Similarly, the Lyapunov function reads

$$\begin{aligned}V(x;y) &\;=\; \tfrac{1}{2}(\|\dot{x}\|^2 \,+\, \|\dot{y}\|^2 \,+\, \|\tilde{x}\|^2 \,+\, \mathbf{dist}^2(y,\ \mathcal{Y}^\star))\\ &\;=\; \tfrac{1}{2}(\|\dot{x}\|^2 \,+\, \|\dot{y}\|^2 \,+\, \|\tilde{x}\|^2 \,+\, \|\tilde{y}\|^2)\end{aligned}$$

where the second line is a result of (11).

Using the chain rule, $\dot{V}$ can be written as

$$\begin{aligned}\dot{V} \;:=\; \tfrac{\mathrm{d}}{\mathrm{d}t}V &\;=\; \langle \tfrac{\mathrm{d}}{\mathrm{d}x}V,\ \dot{x}\rangle \,+\, \langle \tfrac{\mathrm{d}}{\mathrm{d}y}V,\ \dot{y}\rangle\\ &\;=\; \tfrac{1}{2}\langle \dot{x},\ \tfrac{\mathrm{d}}{\mathrm{d}x}\|\dot{x}\|^2\rangle \,+\, \tfrac{1}{2}\langle \dot{y},\ \tfrac{\mathrm{d}}{\mathrm{d}y}\|\dot{x}\|^2\rangle \,+\, \tfrac{1}{2}\langle \dot{x},\ \tfrac{\mathrm{d}}{\mathrm{d}x}\|\dot{y}\|^2\rangle\\ &\quad\;+\, \tfrac{1}{2}\langle \dot{x},\ \tfrac{\mathrm{d}}{\mathrm{d}x}\|\tilde{x}\|^2\rangle \,+\, \tfrac{1}{2}\langle \dot{y},\ \tfrac{\mathrm{d}}{\mathrm{d}y}\|\tilde{y}\|^2\rangle\\ &\;=\; -\,\langle \dot{x},\ (\nabla_x^2\mathcal{L})\dot{x}\rangle \,-\, \langle \dot{y},\ (\nabla_{xy}\mathcal{L})\dot{x}\rangle \,+\, \langle \dot{x},\ (\nabla_{yx}\mathcal{L})\dot{y}\rangle\\ &\quad\;+\, \langle \dot{x},\ \tilde{x}\rangle \,+\, \langle \dot{y},\ \tilde{y}\rangle\\ &\;=\; -\,\langle \dot{x},\ \nabla^2 f(x)\dot{x}\rangle \,-\, \langle \dot{y},\ E\dot{x}\rangle \,+\, \langle \dot{x},\ E^T\dot{y}\rangle\\ &\quad\;+\, \langle \dot{x},\ \tilde{x}\rangle \,+\, \langle \dot{y},\ \tilde{y}\rangle\\ &\;=\; -\,\langle \dot{x},\ \nabla^2 f(x)\dot{x}\rangle \,+\, \langle \dot{x},\ \tilde{x}\rangle \,+\, \langle \dot{y},\ \tilde{y}\rangle\\ &\;=\; -\,\langle \dot{x},\ \nabla^2 f(x)\dot{x}\rangle \,-\, \langle \widetilde{\nabla f},\ \tilde{x}\rangle.\end{aligned}$$

Since $f$ is an $m_f$-strongly convex function with an $L_f$-Lipschitz continuous gradient, at any $x$ there exists a symmetric and positive definite matrix $H_x$ such that $H_x\tilde{x} = \widetilde{\nabla f}$ and $m_f I \,\preceq\, H \,\preceq\, L_f I$ [13, Lemma 6]. Using $H_x$, time derivative $\dot{V}$ can be written as

$$\begin{aligned}\dot{V} &\;=\; -\,\langle H_x\tilde{x} + E^T\tilde{y},\ \nabla^2 f(x)(H_x\tilde{x} + E^T\tilde{y})\rangle \,-\, \langle H_x\tilde{x},\ \tilde{x}\rangle\\ &\;=\; -\begin{bmatrix}\tilde{x}\\ \tilde{y}\end{bmatrix}^T P \begin{bmatrix}\tilde{x}\\ \tilde{y}\end{bmatrix}\end{aligned}$$

where

$$P \;=\; \begin{bmatrix} H_x\nabla^2 f(x)H_x \,+\, H_x & H_x\nabla^2 f(x)E^T\\ E\nabla^2 f(x)H_x & E\nabla^2 f(x)E^T\end{bmatrix}.$$

To obtain (12), we need to show that the smallest singular value of $P$ is non-zero, but since $E$ is not full-row rank, $P$ is not strictly positive-definite. This is the main reason behind the full-row rank assumption on $E$ in many existing results. However, we can circumvent this issue without making any assumptions on $E$ by utilizing the fact that the trajectories are perpendicular to the equilibrium set. In particular, the orthogonal decomposition of $y$ together with (9) yields

$$\tilde{y} \;=\; U_1 U_1 y(0) \,-\, y_0^\star \;\in\; \mathcal{R}(E).$$

Let $\widetilde{w} := U_1^T \tilde{y}$. Since $\tilde{y} \in \mathcal{R}(E)$, $U_1 \widetilde{w} = \tilde{y}$. Substitution of $\widetilde{w}$ into $\dot{V}$ results in

$$\dot{V} \;=\; - \begin{bmatrix} \tilde{x} \\ \widetilde{w} \end{bmatrix}^T \widetilde{P} \begin{bmatrix} \tilde{x} \\ \widetilde{w} \end{bmatrix} \tag{13}$$

where

$$\widetilde{P} \;=\; \begin{bmatrix} H_x \nabla^2 f(x) H_x \;+\; H_x & H_x \nabla^2 f(x) V_1 \Sigma \\ \Sigma V_1^T \nabla^2 f(x) H_x & \Sigma V_1^T \nabla^2 f(x) V_1 \Sigma \end{bmatrix}.$$

We can derive a lower bound on the singular values of $\widetilde{P}$ as follows

$$\begin{aligned} \widetilde{P} \;&=\; \begin{bmatrix} H_x \\ \Sigma V_1^T \end{bmatrix} \nabla^2 f(x) \begin{bmatrix} H_x & V_1 \Sigma \end{bmatrix} \;+\; \begin{bmatrix} H_x & 0 \\ 0 & 0 \end{bmatrix} \\ &\succeq\; m_f \begin{bmatrix} H_x \\ \Sigma V_1^T \end{bmatrix} \begin{bmatrix} H_x & V_1 \Sigma \end{bmatrix} \;+\; \begin{bmatrix} H_x & 0 \\ 0 & 0 \end{bmatrix} \\ &=\; m_f \begin{bmatrix} H_x & 0 \\ 0 & \Sigma \end{bmatrix} \begin{bmatrix} I & V_1 \\ V_1^T & I \end{bmatrix} \begin{bmatrix} H_x & 0 \\ 0 & \Sigma \end{bmatrix} + \begin{bmatrix} H_x & 0 \\ 0 & 0 \end{bmatrix} \\ &=\; m_f \begin{bmatrix} H_x & 0 \\ 0 & \Sigma \end{bmatrix} \begin{bmatrix} I + (1/m_f) H_x^{-1} & V_1 \\ V_1^T & I \end{bmatrix} \begin{bmatrix} H_x & 0 \\ 0 & \Sigma \end{bmatrix}. \end{aligned}$$

where the inequality is obtained by using the implication of strong convexity on the Hessian, i.e., $\|\nabla^2 f(x)\| \succeq m_f I$. Since $H_x \preceq L_f I$, we have

$$\begin{aligned} \begin{bmatrix} I + (1/m_f) H_x^{-1} & V_1 \\ V_1^T & I \end{bmatrix} \;&\succeq\; \begin{bmatrix} (1 + (m_f L_f)^{-1}) I & V_1 \\ V_1^T & I \end{bmatrix} \\ &\succeq\; \frac{\alpha}{2 \;+\; \alpha} I \end{aligned}$$

where the second inequality is obtained by using Lemma 1 provided in Appendix with $\alpha := (L_f m_f)^{-1}$. Finally, Theorem 7.7.2 in [21] yields

$$\begin{aligned} \widetilde{P} \;&\succeq\; m_f \frac{\alpha}{2 + \alpha} \begin{bmatrix} H_x & 0 \\ 0 & \Sigma \end{bmatrix}^2 \\ &\succeq\; m_f \frac{\alpha}{2 + \alpha} \min \left( m_f^2,\; \underline{\sigma}^2(E) \right) I \\ &=\; c_1 I \end{aligned}$$

where $c_1 := m_f \min(m_f^2,\; \underline{\sigma}^2(E))/(1 + 2 m_f L_f)$.

Substitution of the lower bound on the singular values of $\widetilde{P}$ into (13) yields

$$\dot{V} \;\leq\; -\, c_1 \left( \|\tilde{x}\|^2 \;+\; \|\widetilde{w}\|^2 \right) \;=\; -\, c_1 \left( \|\tilde{x}\|^2 \;+\; \|\tilde{y}\|^2 \right) \tag{14}$$

where the equality follows from $U_1 \widetilde{w} = \tilde{y}$.

Lastly, by using (11) and the Lipschitz continuity of $\nabla \mathcal{L}$, which is proved in Lemma 2 provided in Appendix, we show that the Lyapunov function itself has the following quadratic upper bound

$$V(x; y) \;\leq\; c_2 \left( \|x \;-\; x^\star\|^2 \;+\; \|y \;-\; \bar{y}^\star\|^2 \right) \tag{15}$$

where $c_2 = (L_f^2 + \bar{\sigma}^2(E) + 1)/2$. Combining (14) with (15) yields (12) with $\rho = c_1/c_2$, which completes the proof. ■

## V. An Example

To demonstrate the global exponential stability of (5) for a rank-deficient $E$ and the perpendicular motion of the trajectories toward the equilibria, we examine the problem of finding the weighted least-norm solution to an underdetermined linear system of equations. We set $f(x)$ in (1) to $\frac{1}{2}\|Wx\|^2$ where the weight matrix $W \in \mathbb{R}^{10 \times 10}$ is drawn from the multivariate standard normal distribution. As for the constraint, we set $U_1$ to the first two columns of identity matrix of size $3 \times 3$, $\Sigma = \mathbf{diag}(1, 2)$, and construct $V_1 \in \mathbb{R}^{10 \times 2}$ by drawing an orthonormal set of two vectors from the multivariate standard normal distribution. Then, we set $E = U_1 \Sigma V_1^T$ which is in $\mathbb{R}^{3 \times 10}$. In this way, $\mathcal{Y}^\star$ is a line parallel to one of the coordinate axis in 3D plot. Finally, vector $q$ is a random vector in $\mathcal{R}(U_1)$.

Figure 1 shows that even if constraint matrix $E$ is not full row rank, the distance between the trajectory of dynamics (5) and the set of equilibrium points has an exponentially decaying upper bound. Moreover, Figure 2 shows that the trajectory moves perpendicular to the set of equilibrium points and converges to the projection of $y(0)$ onto affine set $\mathcal{Y}^\star$, which verifies the analysis provided in Section III.

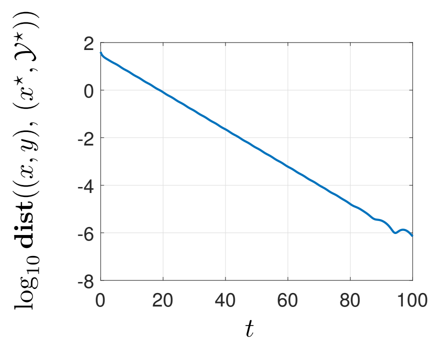


Fig. 1. Exponential convergence of the primal dual dynamics.

## VI. Concluding Remarks

We study stability properties of primal-dual gradient flow dynamics for differentiable convex problems with linear equality constraints. We show that if the set of equilibrium points is affine, then the trajectories exhibit a perpendicular motion toward the equilibria. For strongly convex objective functions, this structural property is utilized to prove global exponential stability of the primal-dual gradient flow dynamics without requiring full-row rank assumption on the constraint matrix. The same approach can also be used to relax the strong convexity assumption on the quadratic terms in the objective functions.

## Appendix

*Lemma 1:* Let $Q \in \mathbb{R}^{n \times r}$ be such that $Q^T Q = I_r$. Then for any $\alpha > 0$, the smallest singular value of

$$\Psi \;:=\; \begin{bmatrix} (1 \;+\; \alpha) I_n & Q \\ Q^T & I_r \end{bmatrix}$$

satisfies $\underline{\sigma}(\Psi) \geq \alpha/(\alpha + 2)$.

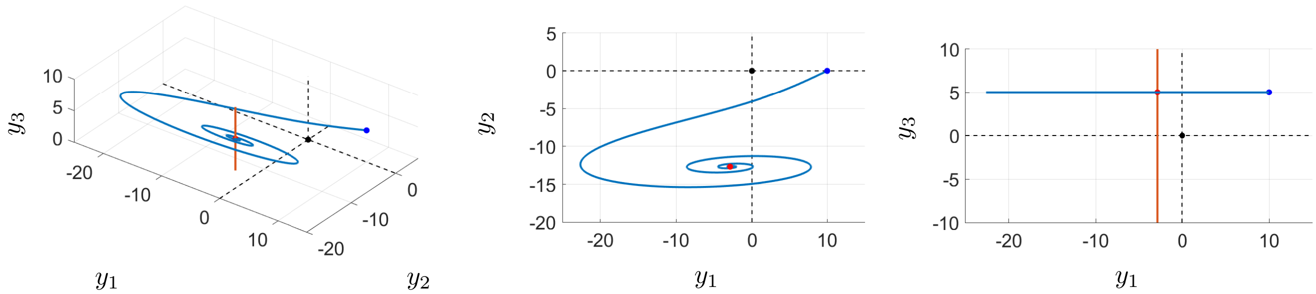


Fig. 2. The trajectory of the dual variables. The blue line is the trajectory of $y(t)$ for $t \geq 0$, the blue dot is initial point $y(0)$, the red dot is limit point $\bar{y}^\star$, the red line is affine set $\mathcal{Y}^\star$, the black dot is the origin and the dashed black lines are the canonical coordinates. As seen on $y_1 y_2$ and $y_1 y_3$ plots, the dual trajectory always stays in the hyperplane perpendicular to $\mathcal{Y}^\star$.

*Proof:* The smallest singular value of $\Psi$, which is nonzero since $\Psi$ is invertable, can be defined as

$$\begin{aligned}\underline{\sigma}(\Psi) \; &:= \min_{\|z_1\|^2 \,+\, \|z_2\|^2 \,=\, 1} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}^T \Psi \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \\ &= \min_{\|z_1\|^2 \,+\, \|z_2\|^2 \,=\, 1} \|z_1 \,+\, Q z_2\|^2 \,+\, \alpha \|z_1\|^2.\end{aligned}$$

By using the following basic inequality [22, Equation 3.7]

$$\|u \,+\, v\|^2 \;\geq\; \frac{1}{\beta+1}\|u\|^2 \,-\, \frac{1}{\beta}\|v\|^2, \quad \forall \beta > 0 \qquad (16)$$

we obtain

$$\begin{aligned}\|z_1 \,+\, Q z_2\|^2 \,+\, \alpha\|z_1\|^2 \;&\geq\; \tfrac{1}{\beta+1}\|z_2\|^2 \,+\, (\alpha - 1/\beta)\|z_1\|^2 \\ &\geq\; \tfrac{\alpha}{\alpha+2}(\|z_2\|^2 \,+\, \|z_1\|^2)\end{aligned}$$

where the second line is obtained by setting $\beta = 2/\alpha$. Substitution of this inequality into the definition of $\underline{\sigma}(\Psi)$ gives the desired result. ■

*Lemma 2:* The gradient of Lagrangian (3) is Lipschitz continuous with modulus $(L_f^2 \,+\, \overline{\sigma}^2(E))$.

*Proof:* For any pairs $(x, y)$ and $(x', y')$, the gradient of $\mathcal{L}$ satisfies

$$\begin{aligned}&\|\nabla\mathcal{L}(x; y) \,-\, \nabla\mathcal{L}(x'; y')\|^2 \\ &= \|\nabla f(x) - \nabla f(x') + E^T(y - y')\|^2 \,+\, \|E(x - x')\|^2 \\ &\leq \|\nabla f(x) - \nabla f(x')\|^2 \,+\, \|E^T(y - y')\|^2 \,+\, \|E(x - x')\|^2 \\ &\leq L_f^2\|x - x'\|^2 \,+\, \overline{\sigma}^2(E)\|y - y'\|^2 \,+\, \overline{\sigma}^2(E)\|x - x'\|^2 \\ &\leq (L_f^2 \,+\, \overline{\sigma}^2(E))(\|(x, y) \,-\, (x', y')\|^2.\end{aligned}$$

The first inequality is obtained by using triangle inequality and the second inequality follows from the assumption that $\nabla f$ is $L_f$-Lipschitz continuous. ■

## References


[1] K. Arrow, L. Hurwicz, and U. H., *Studies in Linear and Non-linear Programming*. Palo Alto CA: Stanford Univ. Press, 1958.

[2] D. Feijer and F. Paganini, “Stability of primal-dual gradient dynamics and applications to network optimization,” *Automatica*, vol. 46, no. 12, pp. 1974–1981, Dec. 2010.

[3] A. Cherukuri, E. Mallada, and J. Cortés, “Asymptotic convergence of constrained primal-dual dynamics,” *Systems Control Lett.*, vol. 87, pp. 10–15, Jan. 2016.

[4] R. Goebel, “Stability and robustness for saddle-point dynamics through monotone mappings,” *Systems Control Lett.*, vol. 108, pp. 16–22, Oct. 2017.

[5] A. Cherukuri, B. Gharesifard, and J. Cortés, “Saddle-point dynamics: conditions for asymptotic stability of saddle points,” *SIAM J. Control Optim.*, vol. 55, no. 1, pp. 486–511, Feb. 2017.

[6] A. Cherukuri, E. Mallada, S. Low, and J. Cortes, “The role of convexity on saddle-point dynamics: Lyapunov function and robustness,” *IEEE Trans. Automat. Control*, vol. 63, no. 8, pp. 2449–2464, Aug. 2018.

[7] T. Holding and I. Lestas, “Stability and instability in saddle point dynamics—part i,” *IEEE Trans. Automat. Control*, vol. 66, no. 7, pp. 2933–2944, Aug. 2021.

[8] G. Qu and N. Li, “On the exponential stability of primal-dual gradient dynamics,” *IEEE Contr. Syst. Lett.*, vol. 3, no. 1, pp. 43–48, 2019.

[9] J. Cortés and S. K. Niederländer, “Distributed coordination for nonsmooth convex optimization via saddle-point dynamics,” *J. Nonlinear Sci.*, vol. 29, no. 4, pp. 1247–1272, Aug. 2019.

[10] N. K. Dhingra, S. Z. Khong, and M. R. Jovanović, “The proximal augmented Lagrangian method for nonsmooth composite optimization,” *IEEE Trans. Automat. Control*, vol. 64, no. 7, pp. 2861–2868, Jul. 2019.

[11] X. Chen and N. Li, “Exponential stability of primal-dual gradient dynamics with non-strong convexity,” in *Amer. Control Conf.*, 2020, pp. 1612–1618.

[12] D. Ding and M. R. Jovanović, “Global exponential stability of primal-dual gradient flow dynamics based on the proximal augmented Lagrangian: A Lyapunov-based approach,” in *IEEE Conf. Decis. Control*, 2020, pp. 4836–4841.

[13] N. K. Dhingra, S. Z. Khong, and M. R. Jovanović, “A second order primal-dual method for nonsmooth convex composite optimization,” *IEEE Trans. Automat. Control*, vol. 67, no. 8, pp. 4061–4076, Aug. 2022.

[14] H. D. Nguyen, T. L. Vu, K. Turitsyn, and J.-J. Slotine, “Contraction and robustness of continuous time primal-dual dynamics,” *IEEE Contr. Syst. Lett.*, vol. 2, no. 4, pp. 755–760, Oct. 2018.

[15] P. Cisneros-Velarde, S. Jafarpour, and F. Bullo, “A contraction analysis of primal-dual dynamics in distributed and time-varying implementations,” *IEEE Tran. Autom. Control*, vol. 67, no. 7, pp. 3560–3566, Jul. 2022.

[16] I. K. Ozaslan and M. R. Jovanović, “Exponential convergence of primal-dual dynamics for multi-block problems under local error bound condition,” in *IEEE Conf. Decis. Control*, Cancun, Mexico, 2022, to appear.

[17] S. A. Alghunaim and A. H. Sayed, “Linear convergence of primal–dual gradient methods and their performance in distributed optimization,” *Automatica*, vol. 117, p. 109003, Jul. 2020.

[18] H. K. Khalil, *Nonlinear Systems*. Upper Saddle River, N.J.: Prentice Hall, 2002.

[19] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge U.K.: Cambridge Univ. Press, 2004.

[20] I. K. Ozaslan, S. Hassan-Moghaddam, and M. R. Jovanović, “On the asymptotic stability of proximal algorithms for convex optimization problems with multiple non-smooth regularizers,” in *Amer. Control Conf.*, 2022, pp. 132–137.

[21] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge U.K.: Cambridge Univ. Press, 2012.

[22] W. Deng and W. Yin, "On the global and linear convergence of the generalized alternating direction method of multipliers," *J. Sci. Comput.*, vol. 66, no. 3, pp. 889–916, May 2015.