

Asymptotic Escape of Spurious Critical Points on the Low-rank Matrix Manifold

Thomas Y. Hou*, Zhenzhen Li*, Ziyun Zhang*

June 13, 2022

Abstract. We show that on the manifold of fixed-rank and symmetric positive semi-definite matrices, the Riemannian gradient descent algorithm almost surely escapes some spurious critical points on the boundary of the manifold. Our result is the first to partially overcome the incompleteness of the low-rank matrix manifold without changing the vanilla Riemannian gradient descent algorithm. The spurious critical points are some rank-deficient matrices that capture only part of the eigen components of the ground truth. Unlike classical strict saddle points, they exhibit very singular behavior. We show that using the dynamical low-rank approximation and a rescaled gradient flow, some of the spurious critical points can be converted to classical strict saddle points in the parameterized domain, which leads to the desired result. Numerical experiments are provided to support our theoretical findings.

Keywords. Low-rank matrix manifold, Riemannian gradient descent, spurious critical points, strict saddles

AMS subject classifications. 58D17, 65F10, 90C26, 15A23

1 Introduction

Low-rank matrix recovery problems are prevalent in modern data science, artificial intelligence and related technological fields. The low-rank property of matrices is widely exploited to extract the hidden low-complexity structure in massive datasets from machine learning, signal processing, imaging science, advanced statistics, information theory and quantum mechanics, just to name a few.

The low-rank matrix manifold [10, 11] has gained popularity in recent years since it gives a neat description of low-rank matrices. The set of matrices with the same size m by n and a fixed rank r forms a smooth manifold \mathcal{M}_r , which is a nonconvex set that is locally isomorphic to the Euclidean space. Many nonconvex optimization techniques can be transferred to \mathcal{M}_r without much difficulty. Among them, the Riemannian gradient descent, the manifold version of the vanilla gradient descent, demonstrates nearly optimal convergence rate and practical flexibility in a number of problems, see e.g. [4, 5, 13, 27, 30, 31].

A fundamental problem has yet remained open in the global analysis of optimization on the low-rank matrix manifold. This comes from the fact that \mathcal{M}_r is *not* a complete set. The boundary of \mathcal{M}_r consists of matrices with rank smaller than r , which are not in \mathcal{M}_r themselves. In other words, $\overline{\mathcal{M}_r} \setminus \mathcal{M}_r = \cup_{s=0}^{r-1} \mathcal{M}_s \not\subset \mathcal{M}_r$. There is no guarantee that the limit point of an

*ACM, Caltech. Correspondence to: Ziyun Zhang (zyzhang@caltech.edu).

iterative sequence will converge to a rank- r ground truth instead of being stuck at some lower-rank *spurious critical points*.

In our previous work [13], it has been proved that under certain assumptions, the converging set of the spurious points has very small measure. This means that starting from a randomly sampled initialization, the iterative sequence avoids these spurious critical points *with high probability*. However, practical applications seem to imply an even stronger result. In fact, we observe that from random initializations, Riemannian gradient descent *almost surely* avoids these spurious critical points. This motivates us to conjecture that their converging sets actually have zero measure.

To understand this phenomenon, it helps to compare it with the asymptotic escape of strict saddle points by gradient descent [18]. The two are remarkably similar, except that the spurious critical points in our context are *not* strict saddles. Instead, the spurious critical points are singular points with negative infinity Hessian directions. More advanced techniques are needed to deal with their singularity.

In this paper, we give a partially confirmatory answer to the aforementioned conjecture. We show that the Riemannian gradient flow and the Riemannian gradient descent with varying stepsize asymptotically escape the rank- $(r - 1)$ spurious critical points on the rank- r symmetric positive semi-definite (SPSD) manifold. We propose to use the dynamical low-rank approximation [15] to describe the gradient flow on the low-rank matrix manifold. We then introduce a rescaled gradient flow to remove the singularity of the ODE system. After rescaling, classical saddle escape theorems can be applied to derive the desired result.

Below is an example which illustrates that the spurious critical points can be the limit points of the Riemannian gradient descent algorithm, but the required initialization is so special that it is almost impossible under random initialization.

Example 1.1. Assume that $n = 3$, $r = 2$. We use the vanilla Riemannian gradient descent (Riemannian GD) algorithm $Z_{k+1} = R(Z_k - \alpha \cdot P_{T_{Z_k}}(Z_k - X))$ to minimize the least squares loss function $f(Z) = \frac{1}{2}\|Z - X\|_F^2$ on the manifold $\mathcal{M}_2 = \{Z : Z \in \mathbb{R}^{3 \times 3}, \text{rank}(Z) = 2\}$. Here $P_{T_{Z_k}}(\cdot)$ is the projection onto the tangent space of \mathcal{M}_2 at Z_k and $R(\cdot)$ is the retraction, cf. Section 2.1. Let

$$X = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z_0 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let the step size $\alpha \in (0, 1)$. Then the sequence $\{Z_k\}_{k=0}^{\infty}$ generated by the Riemannian GD and its limit point are given by

$$Z_k = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & (1 - \alpha)^k \end{pmatrix}, \quad Z_{\#} := \lim_{k \rightarrow \infty} Z_k = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We see that $Z_{\#}$ is a spurious critical point. Note that even though $Z_k \in \mathcal{M}_2$ for any k , $Z_{\#} \notin \mathcal{M}_2$. Instead, $Z_{\#} \in \mathcal{M}_1$.

However, from a slightly perturbed initial point

$$Z_0 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & \epsilon^2 & \epsilon \\ 0 & \epsilon & 1 \end{pmatrix},$$

with arbitrarily small $\epsilon > 0$, one can always show that $\lim_{k \rightarrow \infty} Z_k = X$, i.e., the limit point is not a spurious point.

Figure 1 is a visualization of the gradient $\|P_{T_Z}(Z - X)\|_F$ in the neighborhood of a spurious $Z_\#$. We can see that the gradient is singular near $Z_\#$. There is only one direction in which the sequence converges to $Z_\#$. Along other directions, the Riemannian gradient remains large and the sequence does not converge to $Z_\#$.

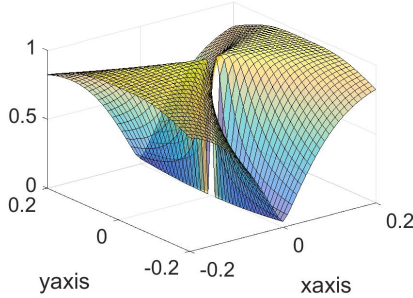


Figure 1: Magnitude of the gradient in the neighborhood of a spurious critical point

1.1 Related work

Incompleteness of the low-rank matrix manifold. The fact that \mathcal{M}_r is not a complete set is first reported in [30] in the context of matrix completion, and later in [16] with low Tucker-rank tensor completion. To guarantee that the iterative sequence of the proposed algorithm stays inside a compact subset of \mathcal{M}_r , the author of [30] proposes to add a regularization term to the objective function $f(Z)$:

$$g(Z) = f(Z) + \mu^2(\|Z\|_F^2 + \|Z^\dagger\|_F^2),$$

where Z^\dagger is the pseudo-inverse of Z , and μ is a parameter. In particular, the term $\mu^2\|Z^\dagger\|_F^2$ guarantees that $\|Z^\dagger\|_F$ will not go to infinity, i.e. Z will not go to rank lower than r .

However, the author also comments that μ^2 can be chosen very small, in fact as small as 10^{-16} . In numerical experiments, one can simply neglect this term and use the original function $f(Z)$ instead of the regularized function $g(Z)$. In other words, the author observes that the iterative sequence of the vanilla Riemannian gradient descent almost surely avoids the rank-deficient points and stays inside \mathcal{M}_r .

Apocalypses from a geometric point of view. Concurrent with our paper, the authors of [19] propose a similar concept. They use the term *apocalypse* to describe the event where the sequence of iterative points is in \mathcal{M}_r but the limit point has rank less than r . This is exactly what happens in Example 1.1. They observe that apocalypse occurs when the tangent cone at the limit is not contained in the limit of the tangent cones. A more detailed discussion on the relation between tangent cones and optimality conditions can be found in [20].

Along this line of research, two remedies have been proposed to fix the apocalypse. The first is a second-order algorithm [19], which uses a smooth lift (essentially the Burer-Monteiro fac-

torization $X = UV^\top$) and the trust-region method. Another is a first-order algorithm proposed in [23], which uses the numerical rank to perform suitable rank reductions.

We remark that although both approaches could avoid spurious points, they require major modification to the gradient descent algorithm. In contrast, we focus on explaining why gradient descent needs no modification in practice.

Asymptotic escape of classical strict saddle points. Gradient descent with random initialization almost surely escapes strict saddle points and converges to minimizers. Such phenomenon has been well studied in the literature. The seminal works in [18] and [17] deal with isolated strict saddles in the Euclidean space. Later the result is extended to non-isolated saddles in [24], to the Riemannian manifold in [6] and [29], and to the strict critical submanifold in [12]. We cite it in its most general form in Theorem 2.13. In addition to the results with fixed step size, later work also extends the result to a diminishing step size [25].

Our central observation is that the spurious critical points in $\overline{\mathcal{M}_r} \setminus \mathcal{M}_r$ are fundamentally different from, but subtly related to, the classical strict saddle points. The spurious critical points have singular local neighborhoods as illustrated in Figure 1. Their asymptotic escape behavior cannot be directly explained by Theorem 2.13. However, using a rescaled gradient flow, we can eliminate the singularity, and apply the saddle escape results to the rescaled system.

Implicit regularization in low-rank matrix factorization. The concept of *implicit regularization* is often used to describe the emergence of favorable structures without explicit regularization terms. In deep matrix factorization and deep neural networks, this describes a tendency towards low-rank solutions and better generalization [2]. In statistical estimation, this could mean a tendency to promote incoherence and accelerate convergence [5][21]. As we have seen, the phenomenon that iterative sequences on the incomplete manifold \mathcal{M}_r stay inside the manifold does not rely on an explicit regularization term $\mu^2 \|Z^\dagger\|_F^2$. Thus it can also be seen as a form of implicit regularization.

Matrix decomposition and its continuity. Our analysis crucially relies on finding a low-rank decomposition that is sufficiently continuous along the whole gradient flow trajectory. The dynamical low-rank approximation (DLRA), first proposed in [15], is a decomposition that suits our purpose. In contrast, the singular value decomposition will lose its differentiability whenever singular values coalesce [8]. A variant called the *analytic SVD* [3] could fix this issue, but it requires analyticity of the gradient function, which cannot be satisfied by Riemannian gradients on \mathcal{M}_r . We remark that the success of DLRA is still limited to the rank- $(r-1)$ spurious critical points. Extension of the current analysis to general spurious critical points is left for future work.

1.2 Organization of this paper

The rest of this paper is organized as follows. In Section 2 we introduce some preliminary results to set the stage. In Section 3 we present and prove the main result of this paper, which is the asymptotic escape of the rank- $(r-1)$ spurious critical points by the gradient flow. Specifically, we introduce the rescaled gradient flow, prove its C^0 - and C^1 -extension to the rank- $(r-1)$ spurious critical points, and show that these points are strict saddles under the rescaled flow. In Section 4 we present the corresponding result for the gradient descent. In Section 5, some

numerical experiments are performed to illustrate our theoretical results. Finally, Section 6 is devoted to some discussions.

2 Preliminaries

Notations. Unless otherwise specified, upper case letters stand for matrices, lower case letters stand for vectors or scalars, and calligraphic letters stand for manifolds or sets. The field \mathbb{F} can be either \mathbb{R} or \mathbb{C} . The low-rank matrix manifold is denoted as \mathcal{M}_r . The Hermitian transpose is denoted as $(\cdot)^*$. The set of $n \times n$ Hermitian matrices is denoted as \mathbb{S}_n . The Stiefel manifold is $\text{St}(n, r) = \{U \in \mathbb{F}^{n \times r} : U^*U = I_r\}$. The orthogonal group is $\text{SO}(n) = \text{St}(n, n)$. The subscript $(\cdot)_\#$ is reserved for the spurious critical points. We use grad and Hess to denote the Riemannian gradient and Hessian, and ∇ to denote the Euclidean derivative.

2.1 Manifold setting

Let \mathcal{M}_r ($r \in \mathbb{N}$) denote the fixed rank manifold $\{Z \in \mathbb{F}^{m \times n} : \text{rank}(Z) = r\}$, where $\mathbb{F} = \mathbb{R}$ or \mathbb{C} . Let $\overline{\mathcal{M}_r}$ be its closure. We summarize some basic properties of \mathcal{M}_r below. A more detailed introduction can be found in [12].

Lemma 2.1. *Let $\mathcal{M}_r = \{Z \in \mathbb{F}^{m \times n} : \text{rank}(Z) = r\}$, where $\mathbb{F} = \mathbb{R}$ or \mathbb{C} . Then $\overline{\mathcal{M}_r} = \{Z \in \mathbb{F}^{m \times n} : \text{rank}(Z) \leq r\}$. Furthermore, we have the following:*

- (1) \mathcal{M}_r is dense in $\overline{\mathcal{M}_r}$.
- (2) For general $\mathcal{M}_r \subset \mathbb{F}^{m \times n}$ of non-Hermitian matrices, \mathcal{M}_r is connected. If restricted to $m = n$, $\mathcal{M}_r \subset \mathbb{S}_n$ Hermitian, then \mathcal{M}_r has $r + 1$ disjoint branches and each branch is connected.
- (3) The local dimension of \mathcal{M}_r is

$$\dim(\mathcal{M}_r) = \begin{cases} (m + n - r)r, & \mathbb{F} = \mathbb{R}, \text{ non-Hermitian}; \\ (2m + 2n - r)r, & \mathbb{F} = \mathbb{C}, \text{ non-Hermitian}; \\ \frac{(2m-r+1)r}{2}, & \mathbb{F} = \mathbb{R}, \text{ Hermitian}; \\ \frac{(4m-r+1)r}{2}, & \mathbb{F} = \mathbb{C}, \text{ Hermitian}. \end{cases}$$

- (4) The boundary of \mathcal{M}_r is $\overline{\mathcal{M}_r} \setminus \mathcal{M}_r = \cup_{s=0}^{r-1} \mathcal{M}_s$.

Lemma 2.2 (Tangent space of \mathcal{M}_r). *Let $X \in \mathcal{M}_r$, $X = U\Sigma V^*$. Let $\mathcal{U} = \text{Col}(U)$, $\mathcal{V} = \text{Col}(V)$ be the column spaces of U and V respectively. Then the tangent space of \mathcal{M}_r at X is*

$$T_X \mathcal{M}_r = (\mathcal{U} \otimes \mathcal{V}) \oplus (\mathcal{U} \otimes \mathcal{V}^\perp) \oplus (\mathcal{U}^\perp \otimes \mathcal{V}).$$

We use the abbreviation T_X when the manifold \mathcal{M}_r is clear from context. The projection operator onto the tangent space can be characterized as

$$P_{T_X}(Y) = P_U \cdot Y + Y \cdot P_V - P_U \cdot Y \cdot P_V.$$

Definition 2.3 (Retraction). Let $X \in \mathcal{M}_r$ and $\xi \in T_X$. We define the natural retraction on \mathcal{M}_r as

$$R(X + \xi) = \arg \min_{Z \in \overline{\mathcal{M}_r}} \|X + \xi - Z\|_F.$$

2.2 Existence of spurious critical points

Our study of spurious critical points is motivated by the study in [13], which reveals a previously unknown property of the fixed rank matrix manifold \mathcal{M}_r . Namely, when minimizing the least squares loss function $f(Z) = \frac{1}{2}\|Z - X\|_F^2$ on \mathcal{M}_r with $\text{rank}(X) = r$, there exist some points with rank smaller than r which could also serve as the limit points of minimizing sequences. This phenomenon was first reported in [11] and later attracted more research interest. We summarize it in the following lemma.

Lemma 2.4 ([13, Lemma 3.8]). *Consider using the Riemannian gradient descent algorithm*

$$Z_{k+1} = R\left(Z_k - \alpha_k P_{T_{Z_k}}(\text{grad} f(Z_k))\right)$$

to minimize the least squares objective function

$$f(Z) = \frac{1}{2}\|Z - X\|_F^2.$$

Let the step size be $\alpha_k \equiv \alpha$. Assume $X = U_x D_x V_x^*$ is a singular value decomposition of X , where $D \in \mathbb{R}^{r \times r}$ is a non-singular diagonal matrix and $U \in \mathbb{F}^{m \times r}$, $V \in \mathbb{F}^{n \times r}$. Then,

1) There are two types of fixed points: one is the ground truth $Z = X$, and the other is the set

$$\mathcal{S}_\# := \{Z_\# : Z_\# = U_1 D_1 V_1^*, \text{ where } U_1, D_1, V_1 \text{ are submatrices of } U_x, D_x, V_x, \text{ satisfying } U_x = (U_1, U_2), D_x = \text{diag}\{D_1, D_2\}, V_x = (V_1, V_2) \text{ respectively, and } Z_\# \neq X\}.$$

2) Specifically, if X has distinct singular values, i.e. all the eigenvalues of X have algebraic multiplicity equal to 1, then $\mathcal{S}_\#$ has cardinality $|\mathcal{S}_\#| = 2^r - 1$. Assume that $X = \sum_{i=1}^r d_i u_i v_i^*$, then $\mathcal{S}_\# = \{Z_\# = \sum_{i=1}^r d_i \eta_i u_i v_i^*, \text{ where } \eta \in \{0, 1\}^r \text{ and } \eta \neq (1, 1, \dots, 1)^*\}$.

Motivated by the above lemma, we introduce the formal definition of the spurious critical points.

Definition 2.5 (Spurious critical points). Assume that $X = U_x D_x V_x^*$ is a singular value decomposition of X . Then the set of **spurious critical points** with respect to $f(Z) = \frac{1}{2}\|Z - X\|_F^2$ on \mathcal{M}_r is $\mathcal{S}_\# = \cup_{s=0}^{r-1} \mathcal{S}_s$, where each \mathcal{S}_s can be characterized as

$$\begin{aligned} \mathcal{S}_s &:= \{Z_\# : Z_\# \in \mathcal{S}_\#, \text{rank}(Z_\#) = s\} \\ &= \{Z_\# : Z_\# = U_1 D_1 V_1^*, U_1 \in \mathbb{F}^{m \times s}, V_1 \in \mathbb{F}^{n \times s}, D_1 \in \mathbb{F}^{s \times s}\}. \end{aligned}$$

Here $U = (U_1, U_2)$, $U_1 \in \mathbb{F}^{m \times s}$, $U_2 \in \mathbb{F}^{m \times (r-s)}$ is a block decomposition of U ; similarly for V and D .

A simple example of a minimizing sequence converging to a spurious critical point instead of the ground truth X has been given in Example 1.1. Observing such phenomenon, one naturally asks how common this happens in practice. Interestingly, when the initial point Z_0 is sampled on \mathcal{M}_r according to some general random sampling scheme, we observe that convergence to spurious critical points almost never happens. The goal of this work is thus to investigate the mechanism behind it.

2.3 Dynamical low-rank approximation

The dynamical low-rank approximation was first proposed in [15] and soon gained popularity as a discretization method for the computation of low-rank evolution systems. It gives a neat description of the column space and core matrix of the low-rank matrix along the evolution. The decomposition enjoys better smoothness than SVD and other classical decompositions. While a smooth version of SVD is only available when the gradient function is analytic, the dynamical low-rank approximation always preserves the smoothness of the gradient function. Thus, it well suits our purpose.

Lemma 2.6 (Dynamical low-rank approximation¹, [15]). *Consider the gradient flow of a function $f(Z) : \mathcal{M}_r \rightarrow \mathbb{R}$. Assume $Z = USV^*$, where $U, V \in \mathbb{F}^{n \times r}$ are orthonormal, $S \in \mathbb{R}^{r \times r}$ nonsingular. Let $M := -\text{grad}f(Z) = -P_{T_Z}(\nabla f(Z))$ denote the negative Riemannian gradient of $f(Z)$ on \mathcal{M}_r . Impose the constraints $\dot{U}^*U = \dot{V}^*V = 0$. Then the gradient flow of $f(Z)$ can be described by the following ODE system:*

$$\begin{cases} \dot{U} = P_U^\perp M V S^{-1}, \\ \dot{V} = P_V^\perp M^* U (S^{-1})^*, \\ \dot{S} = U^* M V. \end{cases} \quad (1)$$

Here, $P_U^\perp = I - UU^*$ and $P_V^\perp = I - VV^*$.

The dynamical low-rank approximation introduces a multiple-to-one mapping as a parameterization of \mathcal{M}_r . Let $\text{St}(n, r)$ denote the n by r Stiefel manifold, i.e. $\text{St}(n, r) = \{U \in \mathbb{F}^{n \times r} : U^*U = I_r\}$. Then we have that for S nonsingular,

$$\begin{aligned} \text{St}(m, r) \oplus \text{St}(n, r) \oplus \mathbb{F}^{r \times r} &\rightarrow \mathcal{M}_r \\ (U, V, S) &\mapsto Z = USV^*. \end{aligned}$$

Since S is not required to be diagonal, there are infinitely many tuples of (U, V, S) corresponding to the same Z , and these tuples are not equivalent under permutations. However, after we impose the constraints $\dot{U}^*U = \dot{V}^*V = 0$, from any initial tuple (U_0, V_0, S_0) there is a *unique* path in $\text{St}(m, r) \oplus \text{St}(n, r) \oplus \mathbb{F}^{r \times r}$ that describes the gradient flow of $f(Z)$ using a similar argument as [15]. In other words, as long as the initial decomposition $Z_0 = U_0 S_0 V_0^*$ is given, the decomposition that satisfies the dynamical low-rank relation is uniquely determined along the whole trajectory.

The advantage of the dynamical low-rank approximation (1) lies in the fact that the ODE system generically stays continuous. This is especially remarkable for the singular vector matrices U and V . As a comparison, SVD might enjoy uniqueness to some extent, but it is known to lose its differentiability when singular values coalesce [8], and that could only be fixed with the unrealistic assumption of analyticity [3].

More specifically, in the SPSD setting, for the least squares function $f(Z) = \frac{1}{2}\|Z - X\|_F^2$, we have the following result.

¹Strictly speaking, our ODE system is not an ‘‘approximation’’ but an exact characterization of the gradient flow. We stick to this terminology for ease of reference.

Lemma 2.7 (Existence of gradient flow). *Consider the manifold of symmetric positive semi-definite (SPSD) matrices, i.e., $m = n$, and $\mathcal{M}_r = \{Z \in \mathbb{S}_n, Z \succcurlyeq 0, \text{rank}(Z) = r\}$. Consider the least squares objective function $f(Z) = \frac{1}{2}\|Z - X\|_F^2$. Let $M := -\text{grad} f(Z)$ denote its negative Riemannian gradient. Let $Z_0 \in \mathcal{M}_r$ be the initialization of the gradient flow at time $T = 0$, and $U_0 \in \text{St}(n, r)$, $S_0 \in \mathbb{S}_r$ nonsingular such that $Z_0 = U_0 S_0 U_0^\top$. Then there exists a unique gradient flow satisfying*

$$\begin{cases} \dot{U} = P_U^\perp M U S^{-1}, \\ \dot{S} = U^* M U. \end{cases}$$

for all $0 \leq T < \infty$.

Proof. The Riemannian gradient of the objective function $f(Z) = \frac{1}{2}\|Z - X\|_F^2$ is $P_{T_Z}(Z - X)$. Plugging in $M = -P_{T_Z}(Z - X)$, and noticing that $P_U^\perp Z = 0$ and $P_U U = U$, we get the following ODE system:

$$\begin{cases} \dot{U} = P_U^\perp X U S^{-1}, \\ \dot{S} = -S + U^* X U. \end{cases}$$

It suffices to show that the ODE system does not blow up in finite time. We prove that for any $T_1 > 0$, $\sigma_{\min}(S)$ is bounded from below for all $T \in [0, T_1]$, where $\sigma_{\min}(S)$ is the smallest eigenvalue of $S \in S_r$.

At $T = 0$, we have $\sigma_{\min}(S) > 0$. At a given time T , let the multiplicity of $\sigma_{\min}(S)$ be j , i.e., $\sigma_{r-j}(S) > \sigma_{r-j+1}(S) = \dots = \sigma_r(S)$. Denote $P_{U_{(r-j+1) \text{ to } r}}$ as the projection onto the corresponding eigen subspace. Using a similar argument as in [22], one can show that

$$\frac{d}{dt} \left(\sum_{l=r-j+1}^r \sigma_l(S) \right) = \text{tr} \left(P_{U_{(r-j+1) \text{ to } r}} \cdot \frac{d}{dt} S \right).$$

In particular, when $\sigma_r(S)$ is a simple eigenvalue and u_r is its eigenvector, this reduces to the classical result

$$\frac{d}{dt} \sigma_r(S) = u_r^* \left(\frac{d}{dt} S \right) u_r.$$

Note that $\frac{d}{dt} S = -S + U^* X U$ and X is positive semi-definite. Thus $\frac{d}{dt} S \succcurlyeq -S$, and we have

$$\frac{d}{dt} \left(\sum_{l=r-j+1}^r \sigma_l(S) \right) \geq \text{tr} \left(P_{U_{(r-j+1) \text{ to } r}} \cdot (-S) \right) = - \sum_{l=r-j+1}^r \sigma_l(S).$$

In particular, when $\sigma_r(S)$ is a simple eigenvalue, one has

$$\frac{d}{dt} \sigma_r(S) \geq -\sigma_r(S).$$

By Grönwall's inequality, $\sigma_{\min}(S)$ decays no faster than exponentially fast. Thus it is bounded from below in any finite time interval. \square

Under the above parameterization, any isolated critical point $Z_\#$ on \mathcal{M}_r corresponds to a critical set on $\text{St}(m, r) \oplus \text{St}(n, r) \oplus \mathbb{F}^{r \times r}$ consisting of infinitely many points, denoted as $\mathcal{N}_{Z_\#}$:

$$\mathcal{N}_{Z_\#} := \{(U_\#, V_\#, S_\#) : U_\# S_\# V_\#^* = Z_\#\}.$$

Some constraints need to be imposed on the above decomposition to make it a valid parameterization for a spurious critical point. We will discuss it in more detail in Section 3.1.

We do not distinguish between the parameterized gradient flow on $\text{St}(m, r) \oplus \text{St}(n, r) \oplus \mathbb{F}^{r \times r}$ and the original gradient flow on \mathcal{M}_r , when there is no confusion. To prove the asymptotic escape of spurious critical points on \mathcal{M}_r , then, is to prove the asymptotic escape of spurious critical submanifolds on $\text{St}(m, r) \oplus \text{St}(n, r) \oplus \mathbb{F}^{r \times r}$.

2.4 Asymptotic escape of classical strict saddles

In this subsection, we introduce the classical results for the asymptotic escape of strict saddle points by gradient descent. Note that we only intend to include the results for the vanilla gradient descent. We do not cover the perturbed or stochastic gradient descent, as they are less relevant to our problem.

We emphasize that the spurious critical points in Definition 2.5, the subject of this study, are *not* classical strict saddle points. It is because the Riemannian Hessian at the spurious critical points is singular, as will be revealed in subsequent sections. Therefore, the theorems and lemmas in this subsection are *not* directly applicable to the spurious critical points. Nevertheless, these theorems and lemmas will be used in an indirect manner, on a rescaled system where the singularity is removed.

The first theorem is a result on the stable and unstable manifolds of the gradient flow at a hyperbolic point.

Theorem 2.8 ([26, The Center Manifold Theorem]). *Let $f \in C^r(E)$ where E is an open subset of \mathbb{R}^n containing the origin and $r \geq 1$. Let $x(t) = \phi_t(x_0)$ be the gradient flow of the system $\dot{x} = f(x)$. Suppose that $f(0) = 0$ and that $Df(0)$ has k eigenvalues with negative real part, j eigenvalues with positive real part, and $m = n - k - j$ eigenvalues with zero real part. Then there exist*

- (1) *A k -dimensional stable manifold $W^s(0)$ of class C^r tangent to the stable subspace E^s at 0, where for all $x_0 \in W^s(0)$,*

$$\lim_{t \rightarrow +\infty} \phi_t(x_0) = 0;$$

- (2) *A j -dimensional unstable manifold $W^u(0)$ of class C^r tangent to the unstable subspace E^u at 0, where for all $x_0 \in W^u(0)$,*

$$\lim_{t \rightarrow -\infty} \phi_t(x_0) = 0;$$

- (3) *And an m -dimensional center manifold $W^c(0)$ of class C^r tangent to the center subspace E^c at 0.*

Furthermore, $W^c(0)$, $W^s(0)$ and $W^u(0)$ are invariant under the gradient flow.

Next, we introduce the counterpart of the previous results for the gradient descent. Specifically, instead of $\dot{x} = f(x)$, we consider $f(x) = \varphi(x)$ where $\varphi(x)$ is the iteration function of the gradient descent algorithm. For example, when minimizing the least squares loss function on \mathcal{M}_r , the iteration function is $\varphi(Z) = R(Z - \alpha P_{T_Z}(Z - X))$. The strict saddle point is defined as follows. It basically says that a strict saddle point is a hyperbolic point of the iteration function.

Definition 2.9 (Strict saddle point). Consider a function $f(\cdot) : \mathcal{M} \rightarrow \mathbb{R}$ defined on a manifold \mathcal{M} . We call $Z \in \mathcal{M}$ a *strict saddle point* of f , if

- (1) $\text{grad}f(Z) = 0$;
- (2) $\text{Hess} f(Z)$ has at least one negative eigenvalue.

We then have the following theorem for the asymptotic escape of isolated saddle points.

Theorem 2.10 ([12, Theorem 2.15]). *Let $f(\cdot) : \mathcal{M} \rightarrow \mathbb{R}$ be a C^2 function on \mathcal{M} . Suppose that $f(\cdot) : \mathcal{M} \rightarrow \mathbb{R}$ has either finitely many saddle points, or countably many saddle points in a compact submanifold of \mathcal{M} , and all saddle points of f are strict saddles as is defined in Definition 2.9. Let \mathcal{A} denote the set of strict saddles. Then we have*

$$\text{Prob}(\lim_{k \rightarrow \infty} Z_k \in \mathcal{A}) = 0$$

The proof of the theorem is based on [28, Theorem III.7], which is very similar to Theorem 2.8 but focuses on the contraction/expansion of the iteration function. We omit the details here.

As is mentioned in the previous subsection, using the parameterization $\text{St}(m, r) \oplus \text{St}(n, r) \oplus \mathbb{F}^{r \times r} \rightarrow \mathcal{M}_r$, each single critical point $Z_{\#}$ corresponds to a submanifold $\mathcal{N}_{Z_{\#}}$. We need the following definitions of the analogy of strict saddle points for submanifolds.

Definition 2.11 (Critical submanifold). For $f : \mathcal{M} \rightarrow \mathbb{R}$, a connected submanifold $\mathcal{N} \subset \mathcal{M}$ is called a *critical submanifold* of f if every point Z in \mathcal{N} is a critical point of f , i.e. $\text{grad}f(Z) = 0$ for any $Z \in \mathcal{N}$.

Definition 2.12 (Strict critical submanifold). A critical submanifold \mathcal{N} of f is called a *strict critical submanifold*, if $\forall Z \in \mathcal{N}$,

$$\lambda_{\min}(\text{Hess} f(Z)) \leq c < 0,$$

where $\lambda_{\min}(\cdot)$ takes the smallest eigenvalue, and $c = c(\mathcal{N})$ is a uniform constant for all $Z \in \mathcal{N}$ depending only on \mathcal{N} .

Using the above definitions, we have the following theorems on the asymptotic escape of strict critical submanifolds for gradient descent.

Theorem 2.13 ([12, Theorem 2.19]). *Let $f(\cdot) : \mathcal{M} \rightarrow \mathbb{R}$ be a C^2 function on \mathcal{M} . Suppose that $f(\cdot) : \mathcal{M} \rightarrow \mathbb{R}$ has either finitely many critical submanifolds, or countably many critical submanifolds in a compact region of \mathcal{M} , and all of them are strict critical submanifolds as defined in Definition 2.12. Let \mathcal{A} denote the union of strict critical submanifolds. Then we have*

$$\text{Prob}(\lim_{k \rightarrow \infty} Z_k \in \mathcal{A}) = 0.$$

We remark that the results on the asymptotic escape of saddle points in the Euclidean space, e.g., the results in [18], can be seen as special cases of Theorem 2.13.

3 Main result

From this section on, we focus on the set of symmetric positive semi-definite (SPSD) matrices, i.e., $m = n$, and $\mathcal{M}_r = \{Z \in \mathbb{S}_n, Z \succcurlyeq 0, \text{rank}(Z) = r\}$. The ground truth X is also a rank- r SPSPD matrix. Recall that by Definition 2.5, the set of spurious critical points is $\mathcal{S}_\# = \cup_{s=0}^{r-1} \mathcal{S}_s$, where each \mathcal{S}_s ($0 \leq s \leq r-1$) is the set of rank- s spurious critical points, i.e.,

$$\begin{aligned} \mathcal{S}_s &:= \{Z_\# : Z_\# \in \mathcal{S}_\#, \text{rank}(Z_\#) = s\} \\ &= \{Z_\# : Z_\# = U_1 D_1 U_1^*, U_1 \in \mathbb{F}^{n \times s}, D_1 \in \mathbb{F}^{s \times s}\}. \end{aligned}$$

The first main result of this paper is as follows.

Theorem 3.1 (Asymptotic escape of \mathcal{S}_{r-1} : gradient flow). *Let $f(Z) = \frac{1}{2}\|Z - X\|_F^2$, where $X \in \mathcal{M}_r$ has distinct eigenvalues. Let $Z_t : t \geq 0$ be the gradient flow of $f(Z)$ on \mathcal{M}_r starting from a random initialization Z_0 . Then we have that $Z_t \in \mathcal{M}_r \forall 0 \leq t < +\infty$, and*

$$\text{Prob}(\lim_{t \rightarrow \infty} Z_t \in \mathcal{S}_{r-1}) = 0.$$

The rest of this section is devoted to the proof of Theorem 3.1.

3.1 Parameterization of $\overline{\mathcal{M}_r}$

In order to use the dynamical low-rank approximation from Section 2.3, we decompose a rank- r matrix $Z \in \mathbb{S}_n$ into $Z = USU^*$, where $U \in \text{St}(n, r)$ and $S \in \mathbb{S}_r$. This decomposition differs from the eigenvalue decomposition in that S is not necessarily a diagonal matrix.

Consider a spurious critical point $Z_\# = U_1 D_1 U_1^* \in \mathcal{M}_s \subset \overline{\mathcal{M}_r} \setminus \mathcal{M}_r$, where $U_1 \in \mathbb{F}^{n \times s}$ represents the s eigenvectors that are also eigenvectors of X . We would like to determine a submanifold $\mathcal{N}_{Z_\#} \subset \text{St}(n, r) \oplus \mathbb{S}_r$ that corresponds to $Z_\#$. Assume that

$$Z_\# = U_\# S_\# U_\#^*,$$

where

$$S_\# = P_\# \Sigma_\# P_\#^*$$

is the eigenvalue decomposition of $S_\#$. Then there exists $U_3 \perp U_1$, such that

$$U_\# = (U_1, U_3) P_\#^*, \quad \Sigma_\# = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix}.$$

In addition, for $Z_\#$ to be a critical point of $f(Z) = \frac{1}{2}\|Z - X\|_F^2$, we need $P_{T_{Z_\#}}(Z_\# - X) = 0$. One can show that this gives

$$U_3 \perp U_x = (U_1, U_2).$$

In other words, U_3 , the $n \times (r-s)$ matrix that makes up for the missing rank, should be chosen to be perpendicular to the missing component U_2 . This also gives us $\lim_{Z \rightarrow Z_\#} \text{grad} f(Z) = 0$, a property that will be useful in upcoming computations.

To sum up, a spurious critical point $Z_{\#} \in \mathcal{S}_{\#}$ can be parameterized as

$$\mathcal{N}_{Z_{\#}} = \left\{ (U_{\#}, S_{\#}) : U_{\#} = (U_1, U_3)P_{\#}^*, S_{\#} = P_{\#} \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix} P_{\#}^*, U_3 \perp U_x \right\},$$

where $P_{\#} \in \text{SO}(r)$ is an orthonormal matrix.

Lemma 3.2. $\mathcal{N}_{Z_{\#}}$ is an embedded submanifold of the manifold $\mathcal{M} := \text{St}(n, r) \oplus \mathbb{S}_r$.

Proof. See Appendix B. □

3.2 Rescaled gradient flow

Consider the dynamical low-rank description of the gradient flow for the objective function $f(Z) = \frac{1}{2}\|Z - X\|_F^2$. Impose the constraint $\dot{U}^*U = 0$ as required by Lemma 2.6. Plug the Riemannian gradient $P_{T_Z}(Z - X)$ into (1), and notice that $P_U^\perp Z = 0$ and $P_U U = U$, we obtain the following ODE system:

$$\begin{cases} \dot{U} = F(U, S) := P_U^\perp X U S^{-1}, \\ \dot{S} = H(U, S) := -S + U^* X U. \end{cases} \quad (\text{DLRA})$$

The main tool for the proof of asymptotic escape is the following *rescaled gradient flow* ODE system:

$$\begin{cases} \dot{U} = \tilde{F}(U, S) := P_U^\perp X U S^{-1} \cdot \sigma_{\min}(S), \\ \dot{S} = \tilde{H}(U, S) := (-S + U^* X U) \cdot \sigma_{\min}(S). \end{cases} \quad (\text{DLRA}^*)$$

Here $\sigma_{\min}(S)$ denotes the smallest eigenvalue of the $r \times r$ matrix S . In other words, the rescaled system (DLRA*) is just the original system (DLRA) times a scalar $\sigma_{\min}(S)$.

We first show that the rescaled system (DLRA*) is well-defined.

Lemma 3.3 (Continuity). *The functions $\tilde{F}(U, S)$ and $\tilde{H}(U, S)$ are C^0 in \mathcal{M}_r .*

Proof. Inside \mathcal{M}_r , the matrix inverse S^{-1} is well-defined, so are the functions $F(U, S)$ and $H(U, S)$. Then use the fact that the smallest eigenvalue $\sigma_{\min}(S)$ is C^0 with respect to S . □

Lemma 3.4 (C^0 -extension). *The functions $\tilde{F}(U, S)$ and $\tilde{H}(U, S)$ can be extended continuously to \mathcal{S}_{r-1} .*

Proof. Take any $Z_{\#} \in \mathcal{S}_{r-1}$ with parameterization $Z_{\#} = U_{\#} S_{\#} U_{\#}^*$. It suffices to show that $\lim_{Z \rightarrow Z_{\#}} \tilde{F}(U, S)$ and $\tilde{H}(U, S)$ exist, and are independent of the specific choices of parameterization.

Let $S = P \Sigma P^*$ and $S_{\#} = P_{\#} \Sigma_{\#} P_{\#}^*$ be the eigenvalue decompositions of S and $S_{\#}$ respectively. Denote $p_i = P(:, i)$, and $p_{\#, i} = P_{\#}(:, i)$. Assume that $X = U_x D_x U_x^* = \sum_{i=1}^r d_i u_i u_i^*$. Since $Z_{\#} \in \mathcal{S}_{r-1}$, from the previous subsection, we know that

$$\Sigma_{\#} = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix},$$

where D_1 is an $(r-1) \times (r-1)$ diagonal matrix, $D_1 = \text{diag}\{d_1, \dots, d_{r-1}\}$. Moreover, when $\|S - S_\#\|_F < \epsilon$ for small enough ϵ , by the $\sin\Theta$ theorem (Lemma A.1), we have

$$\Sigma = \text{diag}\{\sigma_1, \dots, \sigma_{r-1}, \sigma_r\},$$

where

$$\begin{aligned} \sigma_j &> \min\{d_1, \dots, d_{r-1}\} - \epsilon, \quad 1 \leq j \leq r-1; \\ 0 &\leq \sigma_r < \epsilon. \end{aligned}$$

In other words, σ_r and the rest of the eigenvalues of S are well-separated. Thus, when ϵ is small enough, we always have $\sigma_{\min}(S) = \sigma_r$.

Consider $\varphi(S) := S^{-1}\sigma_{\min}(S)$. When $\|S - S_\#\|_F < \epsilon$, we have

$$\begin{aligned} \varphi(S) &= P \cdot \text{diag}\{\sigma_1^{-1}, \dots, \sigma_{r-1}^{-1}, \sigma_r^{-1}\} \cdot P^* \cdot \sigma_r \\ &= P \cdot \text{diag}\left\{\frac{\sigma_r}{\sigma_1}, \dots, \frac{\sigma_r}{\sigma_{r-1}}, 1\right\} \cdot P^* \\ &= P \cdot \text{diag}\left\{\frac{\sigma_r}{\sigma_1}, \dots, \frac{\sigma_r}{\sigma_{r-1}}, 0\right\} \cdot P^* + p_r p_r^*. \end{aligned}$$

Thus,

$$\begin{aligned} \lim_{S \rightarrow S_\#} \varphi(S) &= \lim_{S \rightarrow S_\#} \left(P \cdot \text{diag}\left\{\frac{\sigma_r}{\sigma_1}, \dots, \frac{\sigma_r}{\sigma_{r-1}}, 0\right\} \cdot P^* + p_r p_r^* \right) \\ &= 0 + p_{\#,r} p_{\#,r}^* \\ &= p_{\#,r} p_{\#,r}^*. \end{aligned}$$

In other words, $\varphi(S)$ can be continuously extended to $S_\#$.

We can now compute the limits of \tilde{F} and \tilde{H} . Note that

$$\tilde{F}(U, S) = P_U^\perp X U \cdot \varphi(S).$$

Using the parameterization

$$Z_\# = U_\# S_\# U_\#^* : \quad U_\# = (U_1, U_3) P_\#^*, \quad S_\# = P_\# \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix} P_\#^*,$$

we have

$$\begin{aligned} \lim_{(U, S) \rightarrow (U_\#, S_\#)} \tilde{F}(U, S) &= P_{U_\#}^\perp X U_\# \cdot \lim_{S \rightarrow S_\#} \varphi(S) \\ &= P_{U_\#}^\perp X U_\# \cdot p_{\#,r} p_{\#,r}^* \\ &= (I - P_{U_1} - P_{U_3}) \cdot (U_1 D_1 U_1^* + U_2 D_2 U_2^*) \cdot (U_1, U_3) P_\#^* \cdot p_{\#,r} p_{\#,r}^* \\ &= U_2 D_2 U_2^* \cdot (U_1, U_3) \cdot P_\#^* \cdot p_{\#,r} p_{\#,r}^* \\ &= 0 \end{aligned}$$

As for $\tilde{H}(U, S)$, since $H(U, S)$ is bounded and $\sigma_{\min}(S)$ converges to zero, we have

$$\lim_{(U,S) \rightarrow (U_{\#}, S_{\#})} \tilde{H}(U, S) = \lim_{(U,S) \rightarrow (U_{\#}, S_{\#})} H(U, V, S) \cdot \sigma_{\min}(S) = 0.$$

Thus, $\tilde{F}(U, S)$ and $\tilde{H}(U, S)$ can both be extended continuously to S_{r-1} , independent of the parameterization. \square

3.3 Critical points of the rescaled system

In this section, we show that the ODE systems (DLRA) and (DLRA*) have the same critical points.

Lemma 3.5 (Existence of rescaled gradient flow). *Consider the rescaled ODE system (DLRA*). Let $Z_0 \in \mathcal{M}_r$ be the initialization of the gradient flow at time $T = 0$, and $U_0 \in \text{St}(n, r)$, $S_0 \in \mathbb{S}_r$ nonsingular such that $Z_0 = U_0 S_0 U_0^\top$. Then there exists a unique gradient flow that satisfies (DLRA*) for all $T \in [0, \infty)$.*

Proof. The proof follows the same idea as that of Lemma 2.7. We show that within finite time, (U, S) remains in a region where \tilde{F} and \tilde{H} are Lipschitz continuous. Note that $\nabla_{S_{ij}}(S^{-1}) = -S^{-1}E_{ij}S^{-1}$ where E_{ij} is the indicator matrix of the (i, j) -entry. Note also that the smallest eigenvalue $\sigma_{\min}(S)$ is Lipschitz continuous with respect to S [14]. Thus the Lipschitz continuity of \tilde{F} and \tilde{H} holds if S^{-1} is bounded. This is true if $\sigma_{\min}(S)$ is bounded from below.

At a given time T , let the multiplicity of $\sigma_{\min}(S)$ be j , i.e., $\sigma_{r-j}(S) > \sigma_{r-j+1}(S) = \dots = \sigma_r(S)$. Denote $P_{U_{(r-j+1) \text{ to } r}}$ as the projection onto the corresponding eigen subspace. Using a similar argument as in [22], we now have

$$\begin{aligned} \frac{d}{dt} \left(\sum_{l=r-j+1}^r \sigma_l(S) \right) &= \text{tr} \left(P_{U_{(r-j+1) \text{ to } r}} \cdot \frac{d}{dt} S \right) \\ &= \text{tr} \left(P_{U_{(r-j+1) \text{ to } r}} \cdot (-S + U^* X U) \cdot \sigma_{\min}(S) \right) \\ &\geq \text{tr} \left(P_{U_{(r-j+1) \text{ to } r}} \cdot (-S) \right) \cdot \sigma_{\min}(S) \\ &= - \left(\sum_{l=r-j+1}^r \sigma_l(S) \right) \cdot \sigma_{\min}(S). \end{aligned}$$

In particular, when $\sigma_r(S)$ is a simple eigenvalue, this reduces to

$$\frac{d}{dt} \sigma_r(S) \geq -\sigma_r(S)^2.$$

Thus $\sigma_{\min}(S)$ decays no faster than geometrically due to Grönwall's inequality. Thus it is bounded from below in any finite time interval. \square

Lemma 3.6 (Limit points). *Let $Z_0 \in \mathcal{M}_r$. Then the critical points of the ODE system (DLRA*) are the same as those of (DLRA). Moreover, the gradient flows starting from the same initial point always converge to the same limit point.*

Proof. Observe that the rescaled system (DLRA*) is just the original system (DLRA) multiplied by a scalar:

$$\begin{aligned}\tilde{F}(U, S) &= F(U, S) \cdot \sigma_{\min}(S), \\ \tilde{H}(U, S) &= H(U, S) \cdot \sigma_{\min}(S).\end{aligned}$$

Thus the gradient flow of the rescaled system follows the same path as the original system. In other words, let Z_t ($t \geq 0$) and \tilde{Z}_t ($t \geq 0$) be the solutions of (DLRA) and (DLRA*) starting from the same initial point Z_0 , then for any time $t \geq 0$, there exists a corresponding time $w \geq 0$ such that $\tilde{Z}_t = Z_w$.

When the time goes to infinity, both flows have limit points because both are minimizing flows of a coercive and lower-bounded function $f(Z)$. Denote them as \tilde{Z}_∞ and Z_∞ respectively. Then either $\tilde{Z}_\infty = Z_\infty$, or there exists a finite T such that $\tilde{Z}_\infty = Z_T$.

We now argue that only $\tilde{Z}_\infty = Z_\infty$ is possible. Looking at the ODE system (DLRA*), a critical point has to satisfy either $F(U, S) = H(U, S) = 0$, or $\sigma_{\min}(S) = 0$. In the former case, $F(U, S) = H(U, S) = 0$ means such (U, S) is stationary for (DLRA), so $\tilde{Z}_\infty = Z_\infty$. In the latter case, such (U, S) has to be Z_∞ because we know that $\sigma_{\min}(S)$ cannot be zero at any finite time from Lemma 2.7. So either way, $\tilde{Z}_\infty = Z_\infty$. Therefore, the critical points of (DLRA*) could only be those of (DLRA). \square

By Lemma 3.6, if we can prove that gradient flows of (DLRA*) starting from random initializations almost surely avoids the spurious critical points in \mathcal{S}_{r-1} , we immediately have that the same results apply to (DLRA). In the next subsection, we will show that this is much easier to prove for the rescaled system than for the original system, because the points in \mathcal{S}_{r-1} are now strict saddle points in the classical sense.

3.4 Landscape around the critical points

We now analyze the landscape around the critical points of (DLRA*). In fact, we will show that the C^0 -extension that we proved in Lemma 3.4 can be improved to a C^1 -extension.

Lemma 3.7 (C^1 -extension). *Assume that the eigenvalues of the ground truth matrix X are all distinct. The functions $\tilde{F}(U, S)$ and $\tilde{H}(U, S)$ can be C^1 -extended to \mathcal{S}_{r-1} .*

Proof. We first compute ∇F and ∇H in the interior of \mathcal{M}_r . In this region, S is non-singular, and all the derivatives are well defined. Let $\xi = (\xi_1, \xi_2)$ be a placeholder for the directional derivative, where ξ_1 and ξ_2 correspond to the direction of U and S respectively. Direct computation gives

$$\begin{aligned}\nabla F(U, S)[\xi] &= \begin{pmatrix} -(U\xi_1^* + \xi_1 U^*)XUS^{-1} + P_U^\perp X\xi_1(S^{-1})^* \\ -P_U^\perp XUS^{-1}\xi_2 S^{-1} \end{pmatrix}, \\ \nabla H(U, S)[\xi] &= \begin{pmatrix} \xi_1^* XU + U^* X\xi_1 \\ -\xi_2 \end{pmatrix}.\end{aligned}$$

To extend ∇F and ∇H themselves to \mathcal{S}_{r-1} is impossible: S^{-1} is singular near \mathcal{S}_{r-1} , causing the derivatives to explode. We aim to show that it becomes possible with the rescaled system (DLRA*).

For this purpose, we define the following function, which is the directional derivative of $\varphi(S)$ along the direction η :

$$\psi(S, \eta) := \nabla\varphi(S)[\eta] = \nabla(S^{-1}\sigma_{\min}(S))[\eta].$$

We follow the same notations as before. Direct computation gives

$$\lim_{S \rightarrow S_{\#}} \psi(S, \eta) = \lim_{S \rightarrow S_{\#}} \left(-S^{-1}\eta S^{-1}\sigma_{\min}(S) + S^{-1} \cdot \nabla\sigma_{\min}(S)[\eta] \right).$$

We know from the proof of Lemma 3.4 that when $\|S - S_{\#}\|_F < \epsilon$ for small enough ϵ , the larger eigenvalues σ_1 to σ_{r-1} and the smallest eigenvalue σ_r are well-separated. In fact, assuming that the eigenvalues of X are distinct, for small enough ϵ , all the eigenvalues of S are well-separated, and the corresponding eigenvectors are continuous with respect to the change of S . In this case, we know from [22] that

$$\nabla\sigma_r(S)[\eta] = p_r^* \eta p_r.$$

Thus, we have

$$\lim_{S \rightarrow S_{\#}} \psi(S, \eta) = \lim_{S \rightarrow S_{\#}} \left(-S^{-1}\eta S^{-1}\sigma_r + S^{-1} p_r^* \eta p_r \right).$$

For simplicity, we focus on the case $\mathbb{F} = \mathbb{R}$. Since $\{p_i p_j^*\}_{i,j=1}^r$ form a complete orthogonal basis of $\mathbb{R}^{r \times r}$, we can write

$$\eta = \sum_{1 \leq i, j \leq r} c_{ij} p_i p_j^*.$$

Such decomposition is continuous around $S_{\#}$, since all σ_i 's are well-separated and all p_i 's are continuous with respect to the change of S .

It now suffices to compute $\lim_{S \rightarrow S_{\#}} \psi(S, \eta)$ for $\eta = p_i p_j^*$, as $\psi(S, \eta)$ is linear in η . This comes in the following cases:

(1) If $i, j < r$:

$$\begin{aligned} \lim_{S \rightarrow S_{\#}} \psi(S, p_i p_j^*) &= \lim_{S \rightarrow S_{\#}} \left(-S^{-1} p_i p_j^* S^{-1} \sigma_r + S^{-1} p_r^* p_i p_j^* p_r \right) \\ &= \lim_{S \rightarrow S_{\#}} \left(-P \Sigma^{-1} e_i e_j^* \Sigma^{-1} \sigma_r P^* + S^{-1} \cdot 0 \right) \\ &= \lim_{S \rightarrow S_{\#}} (-P \cdot 0 \cdot P^* + 0) \\ &= 0. \end{aligned}$$

(2) If $i < r, j = r$:

$$\begin{aligned} \lim_{S \rightarrow S_{\#}} \psi(S, p_i p_r^*) &= \lim_{S \rightarrow S_{\#}} \left(-S^{-1} p_i p_r^* S^{-1} \sigma_r + S^{-1} p_r^* p_i p_r^* p_r \right) \\ &= \lim_{S \rightarrow S_{\#}} \left(-P \Sigma^{-1} e_i e_r^* \Sigma^{-1} \sigma_r P^* + S^{-1} \cdot 0 \right) \\ &= d_i^{-1} p_i p_r^*. \end{aligned}$$

(3) If $i = r, j < r$: Similar to the previous case,

$$\lim_{S \rightarrow S_{\#}} \psi(S, p_r p_j^*) = d_j^{-1} p_r p_j^*.$$

(4) If $i = j = r$:

$$\begin{aligned} \lim_{S \rightarrow S_{\#}} \psi(S, p_r p_r^*) &= \lim_{S \rightarrow S_{\#}} \left(-S^{-1} p_r p_r^* S^{-1} \sigma_r + S^{-1} p_r^* p_r p_r^* v_r \right) \\ &= \lim_{S \rightarrow S_{\#}} \left(-P \Sigma^{-1} e_r e_r^* \Sigma^{-1} \sigma_r P^* + S^{-1} \right) \\ &= \lim_{S \rightarrow S_{\#}} \left(P \cdot \text{diag} \{ \sigma_1^{-1}, \dots, \sigma_{r-1}^{-1}, -\sigma_r^{-1} + \sigma_r^{-1} \} \cdot P^* \right) \\ &= \lim_{S \rightarrow S_{\#}} \left(P \cdot \text{diag} \{ \sigma_1^{-1}, \dots, \sigma_{r-1}^{-1}, 0 \} \cdot P^* \right) \\ &= P_{\#} \cdot \text{diag} \{ \sigma_1^{-1}, \dots, \sigma_{r-1}^{-1}, 0 \} \cdot P_{\#}^*. \end{aligned}$$

Therefore, $\psi(S, \eta)$ can be continuously extended to $S_{\#}$ for any η .

We now compute the derivatives of \tilde{F} and \tilde{H} at $Z_{\#}$. The directional derivative in U only involves $\varphi(S_{\#})$, and we have

$$\begin{aligned} \lim_{Z \rightarrow Z_{\#}} \nabla_U \tilde{F}(U, S)[\xi_1] &= \lim_{Z \rightarrow Z_{\#}} \left(-(U \xi_1^* + \xi_1 U^*) X U S^{-1} \sigma_{\min}(S) + P_{U_{\#}}^{\perp} X \xi_1 S^{-1} \sigma_{\min}(S) \right) \\ &= -(U_{\#} \xi_1^* + \xi_1 (U_{\#})^*) X U_{\#} \cdot \varphi(S_{\#}) + P_{U_{\#}}^{\perp} X \xi_1 \cdot \varphi(S_{\#}) \\ &= -(U_{\#} \xi_1^* + \xi_1 (U_{\#})^*) X U_{\#} \cdot p_r p_r^* + P_{U_{\#}}^{\perp} X \xi_1 p_r p_r^*. \end{aligned}$$

As for the directional derivative in S , we now make use of $\psi(S_{\#}, \eta)$:

$$\begin{aligned} \lim_{Z \rightarrow Z_{\#}} \nabla_S \tilde{F}(U, S)[\xi_2] &= \lim_{Z \rightarrow Z_{\#}} \nabla_S \left(P_{U_{\#}}^{\perp} X U S^{-1} \right) [\xi_2] \\ &= \lim_{Z \rightarrow Z_{\#}} \left(P_{U_{\#}}^{\perp} X U \psi(S, \xi_2) \right) \\ &= P_{U_{\#}}^{\perp} X U_{\#} \cdot \psi(S_{\#}, \xi_2). \end{aligned}$$

Since $P_{U_{\#}}^{\perp} X U_{\#} = 0$ and $\psi(S_{\#}, \xi_2)$ is bounded, we have

$$\lim_{Z \rightarrow Z_{\#}} \nabla_S \tilde{F}(U, S)[\xi_2] = 0 \cdot \psi(S_{\#}, \xi_2) = 0.$$

Thus, the derivatives of \tilde{F} can be extended continuously to $Z_{\#}$, and we have

$$\lim_{Z \rightarrow Z_{\#}} \nabla \tilde{F}(U, S) = \begin{pmatrix} -(U_{\#} \xi_1^* + \xi_1 (U_{\#})^*) X U_{\#} \cdot p_r p_r^* + P_{U_{\#}}^{\perp} X \xi_1 \cdot p_r p_r^* \\ 0 \end{pmatrix}.$$

As for the derivative of \tilde{H} , we have

$$\lim_{Z \rightarrow Z_{\#}} \nabla \tilde{H}(U, S)[\xi] = \lim_{Z \rightarrow Z_{\#}} \begin{pmatrix} (\xi_1^* X U + U^* X \xi_1) \cdot \sigma_{\min}(S) \\ -\xi_2 \cdot \sigma_{\min}(S) + (-S + U^* X U) \nabla_S \sigma_{\min}(S)[\xi_2] \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Thus, we have shown that the derivatives of $\tilde{F}(U, S)$ and $\tilde{H}(U, S)$ can both be extended continuously to such $Z_{\#}$, which is equivalent to saying that the functions themselves can be C^1 -extended to such $Z_{\#}$. \square

The C^1 -extension is crucial to the landscape analysis of the system (DLRA*) at the submanifolds corresponding to the rank- $(r - 1)$ spurious critical points. It enables us to compute the Jacobian right at those submanifolds, and determine its eigenvalues. We now show that those submanifolds are actually strict critical submanifolds of the system (DLRA*).

Lemma 3.8 (Strict critical submanifold). *Assume that the eigenvalues of the ground truth matrix X are all distinct. Given a point $Z_\# \in \mathcal{S}_{r-1}$, let $\mathcal{N}_{Z_\#} = \{(U_\#, S_\#) : U_\# S_\# U_\#^* = Z_\#\}$ be the submanifold after parameterization that corresponds to $Z_\#$. Then $\mathcal{N}_{Z_\#}$ is a strict critical submanifold of the system (DLRA*).*

Proof. The goal is to show that for any $(U_\#, S_\#) \in \mathcal{N}_{Z_\#}$, it is a hyperbolic point of the gradient flow with at least one escape direction, and all these points in $\mathcal{N}_{Z_\#}$ share a common escape direction perpendicular to the submanifold itself with a uniformly bounded eigenvalue. We will determine this escape direction by construction, using the results from the proof of Lemma 3.7.

Recall that $S = P\Sigma P^*$, $S_\# = P_\#\Sigma_\#P_\#^*$, and $X = U_x D_x U_x^* = \sum_{i=1}^r d_i u_i u_i^*$. Let

$$\xi = (\xi_1, \xi_2), \quad \xi_1 = u_r p_{\#,r}^*, \quad \xi_2 = 0.$$

Note that $X = U_1 D_1 U_1^* + U_2 D_2 U_2^*$, where $U_2 = u_r$, $D_2 = d_r$, and $U_2 D_2 U_2^*$ is the missing component in this spurious critical point $Z_\#$. In other words, we construct ξ exactly along the direction of this missing component. Using this property, we have that

$$\begin{aligned} & \nabla_U \widetilde{F}(U, S)[\xi_1] \big|_{(U,S)=(U_\#,S_\#)} \\ &= -(U_\# \xi_1^* + \xi_1 U_\#^*) X U_\# \cdot p_{\#,r} p_{\#,r}^* + P_{U_\#}^\perp X \xi_1 \cdot p_{\#,r} p_{\#,r}^* \\ &= -(U_\# p_{\#,r} u_r^* + u_r p_{\#,r}^* U_\#^*) X U_\# \cdot p_{\#,r} p_{\#,r}^* + (I - P_{U_1}^\perp - P_{U_3}^\perp) X \cdot u_r p_{\#,r}^* \cdot p_{\#,r} p_{\#,r}^* \\ &= -((0, U_3) u_r^* + u_r (0, U_3)^*) (U_1 D_1 U_1^* + U_2 D_2 U_2^*) \cdot U_\# p_{\#,r} p_{\#,r}^* + U_2 D_2 U_2^* \cdot u_r p_{\#,r}^* \\ &= 0 + d_r u_r u_r^* \cdot u_r p_{\#,r}^* \\ &= d_r u_r p_{\#,r}^*, \end{aligned}$$

and

$$\nabla_S \widetilde{F}(U, S)[\xi_2] \big|_{(U,S)=(U_\#,S_\#)} = 0.$$

Thus,

$$\nabla \widetilde{F}(U, S)[\xi] \big|_{(U,S)=(U_\#,S_\#)} = d_r u_r p_{\#,r}^* + 0 = d_r u_r p_{\#,r}^*.$$

Meanwhile,

$$\nabla \widetilde{H}(U, S)[\xi] \big|_{(U,S)=(U_\#,S_\#)} = 0.$$

Putting everything together, we have

$$\begin{aligned} \nabla(\widetilde{F}, \widetilde{H})[\xi] &= d_r \cdot (u_r p_{\#,r}^*, 0) \\ &= d_r \cdot \xi. \end{aligned}$$

This means that $\xi = (u_r p_{\#,r}^*, 0)$ is an eigenvector of the Jacobian $\nabla(\widetilde{F}, \widetilde{H})$ with eigenvalue d_r , which is positive.

Thus, for every tuple $(U_\#, S_\#)$ in $\mathcal{N}_{Z_\#}$, we have found an escape direction with uniform eigenvalue. So $\mathcal{N}_{Z_\#}$ is a strict critical submanifold as desired. \square

3.5 Proof of the main result

We now prove Theorem 3.1 using the results from previous subsections.

Proof of Theorem 3.1. By Lemma 2.4, there are only finitely many spurious critical points in \mathcal{S}_{r-1} . By Lemma 3.8, for each $Z_{\#} \in \mathcal{S}_{r-1}$, in the parameterized domain $\text{St}(n, r) \oplus \mathbb{S}_r$, the corresponding submanifold $\mathcal{N}_{Z_{\#}}$ is a strict critical submanifold for the rescaled gradient flow. Since there are only finitely many of them, we can apply Theorem 2.8. This implies that the rescaled gradient flow in the parameterized domain almost never converges to $\cup_{Z_{\#} \in \mathcal{S}_{r-1}} \mathcal{N}_{Z_{\#}}$. Thus the rescaled gradient flow in the original domain \mathcal{M}_r also almost never converges to \mathcal{S}_{r-1} . By Lemma 3.6, the original gradient flow has the same limit as the rescaled gradient flow. Thus the original gradient flow enjoys the same result, i.e., $\text{Prob}(\lim_{t \rightarrow \infty} Z_t \in \mathcal{S}_{r-1}) = 0$. \square

4 Main result for the gradient descent

The previous section has focused on the gradient flow. In this section we derive the result for the gradient descent, namely the asymptotic escape of the Riemannian gradient descent algorithm from the spurious critical points in \mathcal{S}_{r-1} .

Lemma 4.1 (Asymptotic escape of \mathcal{S}_{r-1} : gradient descent). *Let \mathcal{M}_r be the rank- r SPSD matrix manifold. Consider $f(Z) = \frac{1}{2} \|Z - X\|_F^2$ where $X \in \mathcal{M}_r$ has distinct eigenvalues. Let $Z_0 \in \mathcal{M}_r$ be a random initialization, and $\{Z_k\}_{k=0}^{\infty}$ be the sequence generated by the following Riemannian gradient descent algorithm with varying step size:*

$$Z_{k+1} = R\left(Z_k - \alpha \cdot \sigma_r(Z_k) \cdot P_{T_{Z_k}}(\nabla f(Z_k))\right), \quad (2)$$

i.e. $\alpha_k = \alpha \cdot \sigma_r(Z_k)$, where $\sigma_r(Z_k)$ is the r -th eigenvalue of Z_k , and $\alpha > 0$. Assume that $Z_k \in \mathcal{M}_r$ for any $k < +\infty$, i.e., the sequence stays inside \mathcal{M}_r at any finite step. Then we have

$$\text{Prob}(\lim_{k \rightarrow \infty} Z_k \in \mathcal{S}_{r-1}) = 0.$$

In particular, this holds true for arbitrarily large $\alpha > 0$.

Remark 4.2. A few remarks are in order.

- (1) The stepsize $\alpha_k = \alpha \cdot \sigma_r(Z_k)$ is varying but not necessarily diminishing. It is important to note that there is no upper bound on the constant α . Thus even though $\sigma_r(Z_k) \rightarrow 0$ as $Z_k \rightarrow Z_{\#}$, the constant α can be chosen accordingly so that $\{\alpha_k\}$ can be arbitrarily close to non-diminishing stepsize.
- (2) The reason for the choice $\alpha_k = \alpha \cdot \sigma_r(Z_k)$ is similar to the rescaling of the ODE system (DLRA*) in the previous section. Namely, this makes the Jacobian of the iteration function C^1 -extendable to the rank- $(r - 1)$ spurious critical points in \mathcal{S}_{r-1} , using the same techniques as in the proof of Lemma 3.7.

Proof of Lemma 4.1. We use the same notations as before, namely $Z = USU^*$, $S = P\Sigma P^*$, $S_\# = P_\#\Sigma_\#P_\#^*$, and $X = U_x D_x U_x^* = \sum_{i=1}^r d_i u_i u_i^*$. We also let $Z = U_z \Sigma U_z^*$ denote the SVD of Z , which implies $U_z = U \cdot P^*$. We let $\tilde{U} \in \text{St}(n, n-r)$ be the orthogonal complement of U . It is also the orthogonal complement of U_z . Since $U = (U_1, U_3)$, where $U_3 \perp U_2$, we know that $\text{span}\{U_2\} \subset \text{span}\{\tilde{U}\}$. Without loss of generality, we let U_2 be the first column of \tilde{U} .

Consider the iteration function

$$\begin{aligned}\phi(Z) &= R(Z - \alpha \cdot \sigma_r(Z) \cdot P_{T_Z}(\nabla f(Z))) \\ &= R(Z - \alpha \cdot \sigma_r(Z) \cdot \text{grad}f(Z)).\end{aligned}\tag{3}$$

Here $\text{grad}f(Z)$ is the Riemannian gradient. The Jacobian of the iteration function is

$$D\phi(Z) = I - \alpha \cdot (\sigma_r(Z) \cdot \text{Hess}f(Z) + D\sigma_r(Z) \cdot \text{grad}f(Z)).$$

It has been shown in [30] that

$$\text{Hess}f(Z)[\xi] = \xi + P_{U_z}^\perp(Z - X)\tilde{U}N\Sigma^{-1}U_z^* + U_z\Sigma^{-1}N^*\tilde{U}^*(Z - X)P_{U_z}^\perp,\tag{4}$$

where the vector ξ is parameterized as

$$\xi = U_z M U_z^* + U_z N \tilde{U}^* + \tilde{U} N^* U_z^*, \quad M \in \mathbb{F}^{r \times r}, \quad N \in \mathbb{F}^{r \times (n-r)}.$$

In particular, when $\mathbb{F} = \mathbb{R}$, the degree of freedom of ξ is $\frac{r(2n-r+1)}{2}$. It is equal to the dimension of the tangent space that ξ lies in, which is the same as the dimension of the manifold.

Consider $\lim_{Z \rightarrow Z_\#} D\phi(Z)$ for $Z_\# \in \mathcal{S}_{r-1}$. Note that the parameterization from Section 3.1 ensures that $\text{span}\{U_3\} \perp \text{span}\{U_1, U_2\}$, so that $Z_\#$ is a valid critical point, i.e. $\text{grad}f(Z_\#) = 0$. Plugging Equation (4) into Equation (3), we have

$$\begin{aligned}D\phi(Z_\#)[\xi] &:= \lim_{Z \rightarrow Z_\#} D\phi(Z)[\xi] \\ &= \xi - \alpha \cdot \left(\lim_{Z \rightarrow Z_\#} (\sigma_r(Z) \cdot \text{Hess}f(Z)[\xi]) + D\sigma_r(Z)[\xi] \cdot \text{grad}f(Z_\#) \right) \\ &= \xi - \alpha \cdot \left(\lim_{Z \rightarrow Z_\#} (\sigma_r(Z) \cdot \text{Hess}f(Z)[\xi]) \right) \\ &= \xi - \alpha \cdot \left(\lim_{Z \rightarrow Z_\#} \left(\sigma_r(Z) \cdot \xi - P_{U_z}^\perp(Z - X)\tilde{U}N\Sigma^{-1}U_z^* - U_z\Sigma^{-1}N^*\tilde{U}^*(Z - X)P_{U_z}^\perp \right) \right) \\ &= \xi - \alpha \cdot \left(0 \cdot \xi - U_2 D_2 U_2^\top \tilde{U} N_2 \left(\lim_{\Sigma \rightarrow \Sigma_\#} \Sigma^{-1} \sigma_r \right) U_\#^* - U_\# \left(\lim_{\Sigma \rightarrow \Sigma_\#} \Sigma^{-1} \sigma_r \right) N^* \tilde{U}^* U_2 D_2 U_2^\top \right) \\ &= \xi + \alpha \cdot \left(U_2 D_2 U_2^\top \tilde{U} N_2 \left(\lim_{\Sigma \rightarrow \Sigma_\#} \Sigma^{-1} \sigma_r \right) U_\#^* + U_\# \left(\lim_{\Sigma \rightarrow \Sigma_\#} \Sigma^{-1} \sigma_r \right) N^* \tilde{U}^* U_2 D_2 U_2^\top \right).\end{aligned}$$

Here, similar to the proof of Lemma 3.4, we have

$$\lim_{\Sigma \rightarrow \Sigma_\#} \Sigma^{-1} \sigma_r = \text{diag}\{0, \dots, 0, 1\} = e_r e_r^*.$$

Thus, it follows that

$$D\phi(Z_{\#})[\xi] = \xi + \alpha \cdot \left(U_2 D_2 U_2^\top \tilde{U} N e_r e_r^* U_{\#}^* + U_{\#} e_r e_r^* N^* \tilde{U}^* U_2 D_2 U_2^\top \right).$$

Note that without loss of generality, we have let U_2 be the first column of \tilde{U} . Thus we have

$$\begin{aligned} D\phi(Z_{\#})[\xi] &= \xi + \alpha \cdot \left(U_2 D_2 (1, 0, \dots, 0) (N e_r) \begin{pmatrix} U_2^\top \\ 0 \\ \vdots \\ 0 \end{pmatrix} + (U_2, 0, \dots, 0) (e_r^* N^*) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} D_2 U_2^\top \right) \\ &= \xi + \alpha \cdot 2N(1, 1) \cdot U_2 D_2 U_2^\top. \end{aligned}$$

We can immediately read the eigenvalues and eigenvectors of $D\phi(Z_{\#})$ from the above expression. Specifically, when $\mathbb{F} = \mathbb{R}$, $D\phi(Z_{\#})$ has

$$(1) \text{ One eigenvector } \xi = UN\tilde{U}^* + \tilde{U}NU^* \text{ with } N = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \text{ whose corresponding eigen-}$$

value is $\lambda = 1 + 2\alpha \cdot D_2 > 1$;

$$(2) \left(\frac{r(2n-r+1)}{2} - 1 \right) \text{ eigenvectors with eigenvalues } \lambda = 1.$$

The case $\mathbb{F} = \mathbb{C}$ is similar except that the dimensionality is different.

Now that $D\phi(Z_{\#})$ has one eigenvalue greater than 1, while the rest of the eigenvalues are equal to 1. By [28, Theorem III.7], there is an unstable manifold and a center manifold in the neighborhood of $Z_{\#}$, which can be extended globally. The existence of the unstable manifold ensures that $Z_{\#}$ is an asymptotic unstable fixed point of the iteration function $\phi(Z_{\#})$. Thus, the Riemannian gradient descent algorithm with varying step size (2) almost surely escapes \mathcal{S}_{r-1} .

In particular, $D\phi(Z_{\#})$ is always a local diffeomorphism independent of the choice of α , as its only eigenvalues are 1 and $1 + \alpha D_2$. Therefore, the result of Lemma 4.1 holds true for arbitrarily large $\alpha > 0$. \square

5 Numerical experiments

In this section, we present some numerical experiments to illustrate our theoretical results in Theorem 3.1 and 4.1. We also provide some evidence in support of conjectures beyond the previous theorem and lemma.

In all experiments, we let $\mathbb{F} = \mathbb{R}$, $m = n = 100$, $r = 5$, and we use the same ground truth matrix $X \in \mathcal{M}_r$ with distinct singular values. We use the Riemannian gradient descent algorithm to minimize $f(Z) = \frac{1}{2} \|Z - X\|_{\mathbb{F}}^2$. The experiments only differ by the sampling rule and the choice of the step sizes α_k . Each figure is generated by repeating the experiment 100 times. The shaded area represents the range of the data and the solid line represents the median.

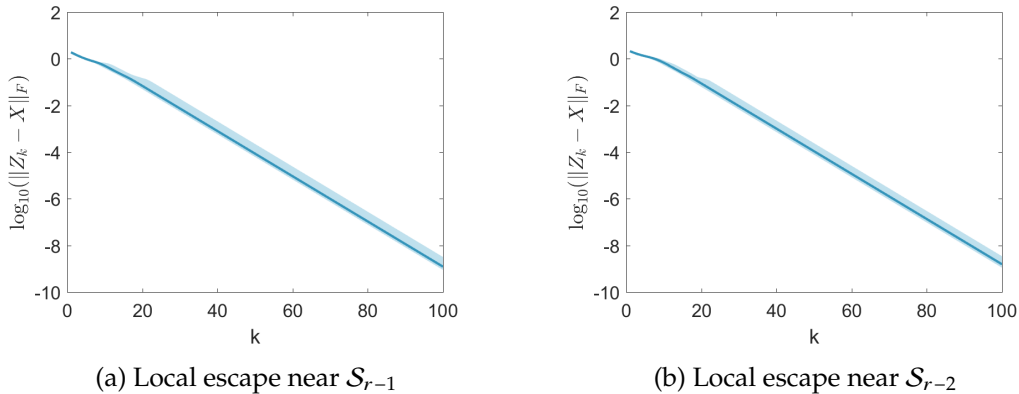


Figure 2: Escape of spurious critical points

The first experiment is performed near a rank- $(r-1)$ spurious critical point $Z_{\#}^{(1)} \in \mathcal{S}_{r-1}$. The initial points are randomly sampled in the local neighborhood of $Z_{\#}^{(1)}$. The stepsize is fixed to be $\alpha_k \equiv \alpha = 0.2$. Figure 2a shows the \log_{10} distance between Z_k and X . It can be seen that in all the repeated experiments, the sequence always succeeds to escape $Z_{\#}^{(1)}$ and converge to X .

To verify whether \mathcal{S}_s ($s < r-1$) incurs the same behavior, we repeat the experiment with $Z_{\#}^{(2)} \in \mathcal{S}_{r-2}$. It can be seen in Figure 2b that the phenomenon is indeed the same. Thus we conjecture that a similar result as Theorem 3.1 holds for those \mathcal{S}_s with $s < r-1$ as well. Proof of such result is left for future work.

Next, we investigate Lemma 4.1 and the varying step size $\alpha_k = \alpha \cdot \sigma_r(Z_k)$. Figures 3a and 3b are the results with a fixed stepsize $\alpha_k \equiv 0.2$. Figures 3c and 3d are the results with varying stepsizes $\alpha_k = 2\sigma_r(Z_k)$. The left are the distances to the ground truth X . The right are the log values of $\sigma_r(Z_k)$ along the iterative path. We can see that first of all, the iterative sequences always escape all spurious critical points and converge to the ground truth. Moreover, the value of $\sigma_r(Z_k)$ is never too small, but soon converges to the smallest singular value of X . This helps illustrate that the varying stepsize $\alpha_k = \alpha \cdot \sigma_r(Z_k)$ is *not a diminishing stepsize* in practice, but is rather always above a certain value.

6 Discussion

In this paper, we discuss the asymptotic escape of the spurious critical points on the low-rank matrix manifold. The goal is to shed some light on the incompleteness of the low-rank matrix manifold \mathcal{M}_r and justify the global use of Riemannian gradient descent on the manifold. To this end, we first point out the existence of a set of spurious critical points $\mathcal{S}_{\#} \subset \overline{\mathcal{M}_r} \setminus \mathcal{M}_r$ and discuss its singularity. We then use a rescaled gradient flow combined with the dynamical low-rank approximation to describe the local landscape, which enables us to eliminate the singularity and prove the asymptotic escape result. We also present a corresponding result for the gradient descent. Numerical experiments are provided to illustrate the theoretical results.

Though this study is focused on \mathcal{S}_{r-1} , the asymptotic escape is empirically observed for \mathcal{S}_s

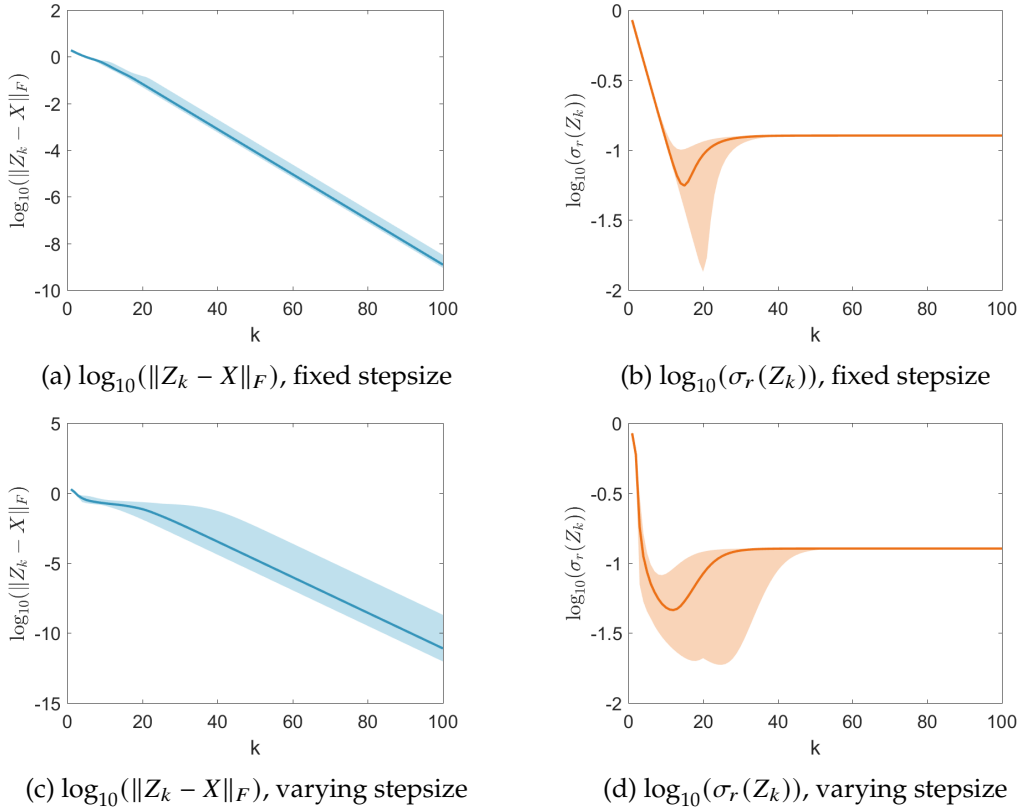


Figure 3: Comparison of fixed and varying step sizes

with $s \leq r - 2$ as well. In fact, all spurious critical points in $\mathcal{S}_\#$ are observed to be asymptotically unstable in practice, which can be seen from the numerical experiments. The current rescaled gradient flow (DLRA*) loses both C^0 - and C^1 -extensions at \mathcal{S}_s with $s \leq r - 2$. This is because the continuity of eigenvalues and eigenvectors are only possible when only one of the eigenvalues is approaching zero. Extension of the result to the case $s \leq r - 2$ is left for future work. On the other hand, the assumption that the eigenvalues of X are distinct is not an essential assumption, and can easily be removed.

Even though the result for the gradient descent calls for a step size $\alpha \cdot \sigma_r(Z)$, there is no upper bound on the constant α from the asymptotic escape analysis. It is because the isomorphism requirement will not be violated even for arbitrarily large α . Thus the step size criterion is not more stringent than that in classical saddle escape results, where there is usually an upper bound on the step size.

In addition to the asymptotic result in this paper, a non-asymptotic result on the number of steps needed to escape the spurious critical points can be found in [13]. There it is shown that the converging set of the spurious critical points can be upper bounded by a small positive measure. With high probability, one has nearly linear convergence rate towards the ground truth. The two sides of the story complement each other and provide a wholesome picture of the unique structure of the low-rank matrix manifold.

References

- [1] P-A Absil, Robert Mahony, and Rodolphe Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.
- [2] Sanjeev Arora, Nadav Cohen, Wei Hu, and Yuping Luo. Implicit regularization in deep matrix factorization. *arXiv preprint arXiv:1905.13655*, 2019.
- [3] Angelika Bunse-Gerstner, Ralph Byers, Volker Mehrmann, and Nancy K Nichols. Numerical computation of an analytic singular value decomposition of a matrix valued function. *Numerische Mathematik*, 60(1):1–39, 1991.
- [4] Léopold Cambier and P-A Absil. Robust low-rank matrix completion by riemannian optimization. *SIAM Journal on Scientific Computing*, 38(5):S440–S460, 2016.
- [5] Yuejie Chi, Yue M Lu, and Yuxin Chen. Nonconvex optimization meets low-rank matrix factorization: An overview. *IEEE Transactions on Signal Processing*, 67(20):5239–5269, 2019.
- [6] Christopher Criscitiello and Nicolas Boumal. Efficiently escaping saddle points on manifolds. *Advances in Neural Information Processing Systems*, 32:5987–5997, 2019.
- [7] Chandler Davis and William Morton Kahan. The rotation of eigenvectors by a perturbation. III. *SIAM Journal on Numerical Analysis*, 7(1):1–46, 1970.
- [8] Luca Dieci and Timo Eirola. On smooth decompositions of matrices. *SIAM Journal on Matrix Analysis and Applications*, 20(3):800–819, 1999.
- [9] Catherine Fraikin, K Hüper, and P Van Dooren. Optimization over the stiefel manifold. In *PAMM: Proceedings in Applied Mathematics and Mechanics*, volume 7, pages 1062205–1062206. Wiley Online Library, 2007.
- [10] Uwe Helmke and John B Moore. *Optimization and dynamical systems*. Springer Science & Business Media, 2012.
- [11] Uwe Helmke and Mark A Shayman. Critical points of matrix least squares distance functions. *Linear Algebra and its Applications*, 215:1–19, 1995.
- [12] Thomas Y Hou, Zhenzhen Li, and Ziyun Zhang. Analysis of asymptotic escape of strict saddle sets in manifold optimization. *SIAM Journal on Mathematics of Data Science*, 2(3):840–871, 2020.
- [13] Thomas Y Hou, Zhenzhen Li, and Ziyun Zhang. Fast global convergence for low-rank matrix recovery via Riemannian gradient descent with random initialization. *arXiv preprint arXiv:2012.15467*, 2020.
- [14] Fatih Kangal, Karl Meerbergen, Emre Mengi, and Wim Michiels. A subspace method for large-scale eigenvalue optimization. *SIAM Journal on Matrix Analysis and Applications*, 39(1):48–82, 2018.

- [15] Othmar Koch and Christian Lubich. Dynamical low-rank approximation. *SIAM Journal on Matrix Analysis and Applications*, 29(2):434–454, 2007.
- [16] Daniel Kressner, Michael Steinlechner, and Bart Vandereycken. Low-rank tensor completion by Riemannian optimization. *BIT Numerical Mathematics*, 54(2):447–468, 2014.
- [17] Jason D Lee, Ioannis Panageas, Georgios Piliouras, Max Simchowitz, Michael I Jordan, and Benjamin Recht. First-order methods almost always avoid strict saddle points. *Mathematical programming*, 176(1):311–337, 2019.
- [18] Jason D Lee, Max Simchowitz, Michael I Jordan, and Benjamin Recht. Gradient descent converges to minimizers. *arXiv preprint arXiv:1602.04915*, 2016.
- [19] Eitan Levin, Joe Kileel, and Nicolas Boumal. Finding stationary points on bounded-rank matrices: A geometric hurdle and a smooth remedy. *arXiv preprint arXiv:2107.03877*, 2021.
- [20] Xinrong Li, Naihua Xiu, and Ziyang Luo. Low-rank matrix optimization over affine set. *arXiv preprint arXiv:1912.03029*, 2019.
- [21] Cong Ma, Kaizheng Wang, Yuejie Chi, and Yuxin Chen. Implicit regularization in nonconvex statistical estimation: Gradient descent converges linearly for phase retrieval and matrix completion. In *International Conference on Machine Learning*, pages 3345–3354. PMLR, 2018.
- [22] Jan R Magnus. On differentiating eigenvalues and eigenvectors. *Econometric Theory*, pages 179–191, 1985.
- [23] Guillaume Ollier, Kyle A Gallivan, and P-A Absil. An apocalypse-free first-order low-rank optimization algorithm. *arXiv preprint arXiv:2201.03962*, 2022.
- [24] Ioannis Panageas and Georgios Piliouras. Gradient descent only converges to minimizers: Non-isolated critical points and invariant regions. *arXiv preprint arXiv:1605.00405*, 2016.
- [25] Ioannis Panageas, Georgios Piliouras, and Xiao Wang. First-order methods almost always avoid saddle points: The case of vanishing step-sizes. In *Advances in Neural Information Processing Systems*, pages 6474–6483, 2019.
- [26] Lawrence Perko. *Differential equations and dynamical systems*, volume 7. Springer Science & Business Media, 2013.
- [27] Reinhold Schneider and André Uschmajew. Convergence results for projected line-search methods on varieties of low-rank matrices via Łojasiewicz inequality. *SIAM Journal on Optimization*, 25(1):622–646, 2015.
- [28] Michael Shub. *Global stability of dynamical systems*. Springer Science & Business Media, 2013.
- [29] Yue Sun, Nicolas Flammarion, and Maryam Fazel. Escaping from saddle points on Riemannian manifolds. *Advances in Neural Information Processing Systems*, 32:7276–7286, 2019.

- [30] Bart Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM Journal on Optimization*, 23(2):1214–1236, 2013.
- [31] Ke Wei, Jian-Feng Cai, Tony F Chan, and Shingyu Leung. Guarantees of riemannian optimization for low rank matrix recovery. *SIAM Journal on Matrix Analysis and Applications*, 37(3):1198–1222, 2016.

A Auxiliary lemmas

Lemma A.1 (The $\sin\Theta$ Theorem, [7]). *Let A be a Hermitian operator. Assume that*

$$A = \begin{pmatrix} E_0 & E_1 \end{pmatrix} \begin{pmatrix} A_0 & 0 \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} E_0^* \\ E_1^* \end{pmatrix}$$

is an invariant subspace decomposition (i.e., a generalized eigenvalue decomposition) of A . Let

$$B = A + \Delta, \quad B = \begin{pmatrix} F_0 & F_1 \end{pmatrix} \begin{pmatrix} B_0 & 0 \\ 0 & B_1 \end{pmatrix} \begin{pmatrix} F_0^* \\ F_1^* \end{pmatrix}.$$

Let Θ_0 be the angle matrix between subspaces E_0 and F_0 . Define the residual as

$$R := BE_0 - E_0A_0.$$

If there is an interval $[\beta, \alpha]$ and $\delta > 0$, such that the spectrum of A_0 lies entirely in $[\beta, \alpha]$, while that of B_1 lies entirely in $(-\infty, \beta - \delta] \cup [\alpha + \delta, +\infty)$, then for every unitary-invariant norm $\|\cdot\|$, we have

$$\delta \|\sin\Theta_0\| \leq \|R\|.$$

In particular, this holds true for the matrix 2-norm and the Frobenius norm.

B Proof of Lemma 3.2

We recall the lemma from the main text.

Lemma 3.2. Define

$$\mathcal{N}_{Z_\#} := \left\{ (U_\#, S_\#) : U_\# = (U_1, U_3)P_\#^*, S_\# = P_\# \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix} P_\#^*, U_3 \perp U_x \right\}.$$

Then $\mathcal{N}_{Z_\#}$ is an embedded submanifold of the manifold $\mathcal{M} := \text{St}(n, r) \oplus \mathbb{S}_r$.

The intuition behind Lemma 3.2 is that the set $\mathcal{N}_{Z_\#}$ is a subset of \mathcal{M} characterized by some algebraic constraints, namely $U_\#S_\#U_\#^* = Z_\#$ and $U_3 \perp U_x$. As is often the case, one would expect such algebraic constraints to give an embedded submanifold. We will make this intuition rigorous in this section.

We note that traditionally, embedded submanifold is proved by the submersion theorem, i.e., by showing that the set is the preimage of a regular value of a submersive mapping. But this approach does not work here because $Z_\#$ is not a regular value. Instead, we need to go back to the definition of a submanifold and construct chart functions on $\mathcal{N}_{Z_\#}$ directly.

Below are some auxiliary results from the literature.

Lemma B.1 ([1, Proposition 3.3.2]). *A subset \mathcal{N} of a manifold \mathcal{M} is a d -dimensional embedded submanifold of \mathcal{M} if and only if, around each point $x \in \mathcal{N}$, there exists a chart (\mathcal{U}, φ) of \mathcal{M} such that $\mathcal{N} \cap \mathcal{U}$ is a φ -coordinate slice of \mathcal{U} , i.e.,*

$$\mathcal{N} \cap \mathcal{U} = \{x \in \mathcal{U} : \varphi(x) \in \mathbb{R}^d \times \mathbf{0}\}.$$

In this case, the chart $(\mathcal{N} \cap \mathcal{U}, \varphi)$, where φ is seen as a mapping into \mathbb{R}^d , is a chart of the embedded submanifold \mathcal{N} .

By Lemma B.1, if we can construct an atlas of \mathcal{M} and an atlas of $\mathcal{N}_{Z\#}$, such that the charts in the latter atlas are coordinate slices of the charts in the former atlas, then $\mathcal{N}_{Z\#}$ is an embedded submanifold of \mathcal{M} . This approach is less common than the traditional submersion theorem approach, but is necessary for our problem.

Lemma B.2 ([9]). *For the real Stiefel manifold $St(n, k)$, there exists an atlas $\cup_Q(\mathcal{U}_Q, \varphi_Q)$ of the Stiefel manifold. Namely, for each chart $(\mathcal{U}_Q, \varphi_Q)$, Q is a matrix in $St(n, k)$, and the function φ_Q can be expressed as*

$$\begin{aligned} \varphi_Q : \quad \mathcal{U}_Q &\rightarrow \text{Skew}(k) \oplus \mathbb{R}^{(n-k) \times k}, \\ U &\mapsto (\Omega_{11}, \Omega_{21}), \end{aligned}$$

where

$$\begin{aligned} \Omega_{11} &= (U_1^\top + Q_1^\top)^{-1} (Q_1^\top U_1 + U_2^\top Q_2 - U_1^\top Q_1 - Q_2^\top U_2) (U_1 + Q_1)^{-1}, \quad \Omega_{11} = -\Omega_{11}^\top, \\ \Omega_{21} &= (U_2 - Q_2)(U_1 + Q_1)^{-1}, \end{aligned}$$

and $U = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$, $Q = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}$ are the block forms of U and Q respectively. Such chart function is defined on the subset $\mathcal{U}_Q \subset St(n, k)$ which covers all of the manifold $St(n, k)$ except a zero-measure set.

In particular, if $Q = \begin{pmatrix} I_k \\ \mathbf{0} \end{pmatrix}$, then

$$\begin{aligned} \Omega_{11} &= (U_1^\top + I_k)^{-1} (U_1 - U_1^\top) (U_1 + I_k)^{-1}, \\ \Omega_{21} &= U_2 (U_1 + I_k)^{-1}. \end{aligned}$$

Lemma B.2 provides a neat construction of charts on the Stiefel manifold. In fact, we only need two charts to cover the whole manifold, if we choose any two Q 's that do not share any left singular vector. We will use this construction frequently in the proof of Lemma 3.2.

We are now ready to prove Lemma 3.2.

Proof of Lemma 3.2. We restrict our attention to the case $\mathbb{F} = \mathbb{R}$. The case $\mathbb{F} = \mathbb{C}$ is very similar except that the dimensionalities of some manifolds in the subsequent proof are slightly different.

We aim to construct explicit charts of $\mathcal{M} = St(n, r) \oplus \mathbb{S}_r$, and explicit charts of $\mathcal{N}_{Z\#}$, such that the latter are the coordinate slices of the former. For clarity, we will first write out the charts of $\mathcal{N}_{Z\#}$, and then express them as coordinate slices of charts of \mathcal{M} .

Step 1: Construct charts of $\mathcal{N}_{Z\#}$.

For any $(U, S) \in \mathcal{N}_{Z\#}$, we rewrite U and S as the following:

$$\begin{aligned} U &= U_1 P_1^\top + U_3 P_2^\top, & S &= P_1 D_1 P_1^\top, \\ \text{where } P_1 &\in \mathbb{R}^{r \times s}, & P_2 &\in \mathbb{R}^{r \times (r-s)}, & P &= (P_1, P_2) \in \text{SO}(r). \end{aligned}$$

We argue that there exists a mapping from every P_1 to a unique P_2 . An intuitive explanation is that P_2 can always be uniquely determined by a Gram-Schmidt process starting from the identity matrix. Thus we can write $P_2 = \mathcal{P}_2(P_1)$ where $\mathcal{P}_2 : \mathbb{R}^{r \times s} \rightarrow \mathbb{R}^{r \times (r-s)}$ is a function. Therefore, any $(U, S) \in \mathcal{N}_{Z\#}$ can be re-parameterized using only (P_1, U_3) . We write this re-parameterization as a function f :

$$\begin{aligned} f : \mathcal{N}_{Z\#} &\rightarrow \text{St}(r, s) \oplus \widetilde{\text{St}}(n, r-s; U_x); \\ (U, S) &\mapsto (P_1, U_3). \end{aligned}$$

Here $\text{St}(r, s)$ is a Stiefel manifold, and $\widetilde{\text{St}}(n, r-s; U_x)$ is a constrained Stiefel manifold:

$$\widetilde{\text{St}}(n, r-s; U_x) := \{U_3 : U_3 \in \text{St}(n, r-s), U_3 \perp U_x\}, \quad \text{where } U_x = (U_1, U_2).$$

We now construct charts for P_1 and U_3 respectively. The domain of P_1 is the Stiefel manifold $\text{St}(r, s)$. By Lemma B.2, there exists an atlas where every chart function maps to $\text{Skew}(s) \oplus \mathbb{R}^{(r-s) \times s}$. Let $g^{(1)}$ be one such chart function:

$$\begin{aligned} g^{(1)} : \text{St}(r, s) &\rightarrow \text{Skew}(s) \oplus \mathbb{R}^{(r-s) \times s} \\ P_1 &\mapsto (\Omega_{11}, \Omega_{21}). \end{aligned}$$

The domain of U_3 is the constrained Stiefel manifold $\widetilde{\text{St}}(n, r-s; U_x)$. Here $U_x \in \text{St}(n, r)$ is the eigenvectors matrix of the ground truth X , which is fixed. To construct a chart of $\widetilde{\text{St}}(n, r-s; U_x)$, we first construct a mapping $g^{(2)}$ according to Lemma B.2, such that

$$\begin{aligned} g^{(2)} : \widetilde{\text{St}}(n, r-s; U_x) &\rightarrow \text{Skew}(r-s) \oplus \widetilde{\mathbb{R}}^{(n-(r-s)) \times (r-s)}, \\ U_3 &\mapsto (\Lambda_{11}, \Lambda_{21}). \end{aligned}$$

The domain of Λ_{21} is $\widetilde{\mathbb{R}}^{(n-(r-s)) \times (r-s)}$, which is a constrained set. To express the constraints $U_3 \perp U_x$ in terms of constraints on Λ_{21} , we write $U_3 = \begin{pmatrix} U_{3,1} \\ U_{3,2} \end{pmatrix}$ and $U_x = \begin{pmatrix} U_{x,1} \\ U_{x,2} \end{pmatrix}$. Assume without loss of generality that $g^{(2)}$ is constructed by picking $Q = (I_{r-s}, \mathbf{0})^\top$ in Lemma B.2. Then

$$\Lambda_{21} = U_{3,2}(U_{3,1} + I_{r-s})^{-1}.$$

Since $U_3 \perp U_x$, we have

$$U_x^\top U_3 = U_{x,1}^\top U_{3,1} + U_{x,2}^\top U_{3,2} = 0.$$

Thus,

$$U_{x,2}^\top U_{3,2} = -U_{x,1}^\top U_{3,1}.$$

This gives us

$$U_{x,2}^\top \Lambda_{21} = -U_{x,1}^\top U_{3,1} (U_{3,1} + I_{r-s})^{-1}.$$

These are linear constraints on Λ_{21} .

Let g be the concatenation of $g^{(1)}$ and $g^{(2)}$, then we have a re-parameterization of (P_1, U_3) as follows:

$$\begin{aligned} g : \quad \text{St}(r, s) \oplus \widetilde{\text{St}}(n, r-s) &\rightarrow \text{Skew}(s) \oplus \mathbb{R}^{(r-s) \times s} \oplus \text{Skew}(r-s) \oplus \widetilde{\mathbb{R}}^{(n-(r-s)) \times (r-s)}; \\ (P_1, U_3) &\mapsto (\Omega_{11}, \Omega_{21}, \Lambda_{11}, \Lambda_{21}). \end{aligned}$$

Here $\widetilde{\mathbb{R}}^{(n-(r-s)) \times (r-s)}$ is the submanifold of $\mathbb{R}^{(n-(r-s)) \times (r-s)}$ defined by the linear constraints that we derived:

$$\widetilde{\mathbb{R}}^{(n-(r-s)) \times (r-s)} := \left\{ \Lambda_{21} \in \mathbb{R}^{(n-(r-s)) \times (r-s)} : U_{x,2}^\top \Lambda_{21} = -U_{x,1}^\top U_{3,1} (U_{3,1} + I_{r-s})^{-1} \right\}.$$

Let Λ_{21}° be an arbitrary solution to the equation $U_{x,2}^\top \Lambda_{21} = -U_{x,1}^\top U_{3,1} (U_{3,1} + I_{r-s})^{-1}$. Then

$$\widetilde{\mathbb{R}}^{(n-(r-s)) \times (r-s)} = \Lambda_{21}^\circ + \text{Ker}(U_{x,2}^\top).$$

By finding an orthogonal basis for $\text{Ker}(U_{x,2}^\top)$, it is easy to construct a chart function

$$\begin{aligned} h : \quad \widetilde{\mathbb{R}}^{(n-(r-s)) \times (r-s)} &\rightarrow \mathbb{R}^{(n-(r-s))(r-s)-r(r-s)} \\ \Lambda_{21} &\mapsto \Gamma. \end{aligned}$$

Putting everything together, we have that

$$\begin{aligned} \varphi := (\text{id}, h) \circ g \circ f : \quad \mathcal{N}_{Z\#} &\rightarrow \text{Skew}(s) \oplus \mathbb{R}^{(r-s) \times s} \oplus \text{Skew}(r-s) \oplus \mathbb{R}^{(n-(r-s))(r-s)-r(r-s)}; \\ (U, S) &\mapsto (\Omega_{11}, \Omega_{21}, \Lambda_{11}, \Gamma). \end{aligned}$$

This is a chart function for the whole $\mathcal{N}_{Z\#}$ except a zero-measure set. Varying $g^{(1)}$ and $g^{(2)}$ as needed and we have the atlas for the whole $\mathcal{N}_{Z\#}$.

Step 2: Express the charts of $\mathcal{N}_{Z\#}$ as coordinate slices of charts of \mathcal{M} .

To express things into coordinate slices, we will work the other way around: we extend the chart function φ into a chart function $\widetilde{\varphi}$ defined on $\mathcal{M} = \text{St}(n, r) \oplus \mathbb{S}_r$.

For any $(U, S) \in \mathcal{M} = \text{St}(n, r) \oplus \mathbb{S}_r$, we construct a re-parameterization as follows:

$$\begin{aligned} U &= \left((U_1 R_1, 0) + (M_4, U_3 R_2) \right) \begin{pmatrix} P_1^\top \\ P_2^\top \end{pmatrix}, & S &= (P_1, P_2) \widetilde{S} \begin{pmatrix} P_1^\top \\ P_2^\top \end{pmatrix}, \\ \text{where } P_1 &\in \mathbb{R}^{r \times s}, \quad P_2 \in \mathbb{R}^{r \times (r-s)}, \quad P = (P_1, P_2) \in \text{SO}(r), \\ U_3 &\in \widetilde{\text{St}}(n, r-s; U_1), \quad M_4 \in \widetilde{\mathbb{R}}^{n \times s}, \quad \widetilde{S} \in \mathbb{S}_r, \\ R_1 &\in \text{upper}(s, s), \quad R_2 \in \text{upper}(r-s, r-s). \end{aligned}$$

The domain of P_1 is $\text{St}(r, s)$. P_2 is still uniquely determined by P_1 as before. The domain of U_3 is the constrained Stiefel manifold $\widetilde{\text{St}}(n, r-s; U_1) := \{U_3 : U_3 \in \text{St}(n, r-s), U_3 \perp U_1\}$. Note that

the constraints are only in terms of U_1 instead of $U_x = (U_1, U_2)$. The domain of M_4 is the linearly constrained subspace $\widetilde{\mathbb{R}}^{n \times s} := \{M_4 \in \mathbb{R}^{n \times s}, M_4 \perp U_1\}$. The domain of \widetilde{S} is \mathbb{S}_r . The domain of R_1 is the subspace of $s \times s$ upper triangular matrices. The domain of R_2 is the subspace of $(r-s) \times (r-s)$ upper triangular matrices, but with some constraints that will be specified later. We define the following mapping:

$$\begin{aligned} \widetilde{f}: \quad \text{St}(n, r) \oplus \mathbb{S}_r &\rightarrow \text{St}(r, s) \oplus \widetilde{\text{St}}(n, r-s; U_1) \oplus \widetilde{\mathbb{R}}(n, s) \oplus \overline{\text{upper}}(r-s, r-s) \oplus \mathbb{S}_r \oplus \text{upper}(s, s); \\ (U, S) &\mapsto \left(P_1, U_3, M_4, R_2 - I_{r-s}, \widetilde{S} - \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix}, R_1 - I_s \right). \end{aligned}$$

The mapping \widetilde{f} is written in such a way because, if $(U, S) \in \mathcal{N}_{Z\#}$, then the last few components are all zero, and f is just a coordinate slice of \widetilde{f} :

$$\widetilde{f}(U, S) = (P_1, U_3, 0, 0, 0, 0).$$

For the first part of the image of \widetilde{f} , we apply g as before:

$$\begin{aligned} g: \quad \text{St}(r, s) \oplus \widetilde{\text{St}}(n, r-s) &\rightarrow \text{Skew}(s) \oplus \mathbb{R}^{(r-s) \times s} \oplus \text{Skew}(r-s) \oplus \widehat{\mathbb{R}}^{(n-(r-s)) \times (r-s)}; \\ (P_1, U_3) &\mapsto (\Omega_{11}, \Omega_{21}, \Lambda_{11}, \Lambda_{21}). \end{aligned}$$

However, the set $\widehat{\mathbb{R}}^{(n-(r-s)) \times (r-s)}$ is different from the $\widetilde{\mathbb{R}}^{(n-(r-s)) \times (r-s)}$ before, because the constraints only contain U_1 but does not contain U_2 . Fewer constraints mean a larger subspace, and we have

$$\begin{aligned} \widehat{\mathbb{R}}^{(n-(r-s)) \times (r-s)} &= \left\{ \Lambda_{21} \in \mathbb{R}^{(n-(r-s)) \times (r-s)} : U_{1,2}^\top \Lambda_{21} = -U_{1,1}^\top U_{3,1} (U_{3,1} + I_{r-s})^{-1} \right\} \\ &= \widetilde{\mathbb{R}}^{(n-(r-s)) \times (r-s)} + \left(\text{Ker}(U_{1,2}^\top) \setminus \text{Ker}(U_{x,2}^\top) \right) \\ &= \Lambda_{21}^\circ + \text{Ker}(U_{x,2}^\top) + \left(\text{Ker}(U_{1,2}^\top) \setminus \text{Ker}(U_{x,2}^\top) \right). \end{aligned}$$

Let $h^{(2)}$ be the chart function for the extra subspace, then

$$\begin{aligned} (h, h^{(2)}) : \quad \widehat{\mathbb{R}}^{(n-(r-s)) \times (r-s)} &\rightarrow \mathbb{R}^{(n-(r-s))(r-s) - r(r-s)} \oplus \mathbb{R}^{(r-s)(r-s)}, \\ \Lambda_{21} &\mapsto (\Gamma, \Gamma^{(2)}). \end{aligned}$$

Putting them together, we have

$$(\text{id}, h, h^{(2)}) \circ g \circ f : (P_1, U_3) \mapsto (\Omega_{11}, \Omega_{21}, \Lambda_{11}, \Gamma, \Gamma^{(2)}).$$

The chart function φ is a coordinate slice of the above mapping.

It suffices to find the chart functions for the remaining components of $\widetilde{f}(U, S)$, i.e., the components $M_4, \widetilde{S}, R_1, R_2$. For $\widetilde{S} \in \mathbb{S}_r$ and $R_1 \in \text{upper}(s, s)$, the domains are Euclidean spaces with natural bases. We now look at M_4 and R_2 .

Decompose M_4 into parts that are parallel to and perpendicular to the subspace of U_3 :

$$M_4 = M_4^\parallel + M_4^\perp, \quad \text{where } M_4^\parallel = P_{U_3} M_4, \quad M_4^\perp = P_{U_3}^\perp M_4.$$

Let $M_4^\perp = U_4 R_4$ be the QR decomposition of M_4^\perp . Then the whole M_4 could be written as

$$M_4 = (U_3, U_4) \begin{pmatrix} R_3 \\ R_4 \end{pmatrix}, \quad R_3 \in \mathbb{R}^{(r-s) \times (r-s)}, \quad R_4 \in \text{upper}(s, s).$$

In this way, we can re-parameterize (M_4, R_2) using (U_4, R_2, R_3, R_4) :

$$p : (M_4, R_2 - I_{r-s}) \mapsto (U_4, R_2, R_3, R_4).$$

The domain of U_4 is the constrained Stiefel manifold $\widetilde{\text{St}}(n, s; U_1, U_3)$. Just as before, we can construct a composite function for this constrained Stiefel manifold:

$$\begin{aligned} g^{(3)} : \quad & \widetilde{\text{St}}(n, s; U_1, U_3) \rightarrow \text{Skew}(s) \oplus \widetilde{\mathbb{R}}^{(n-s) \times s}, \\ & U_4 \mapsto (\Pi_{11}, \Pi_{21}); \\ h^{(3)} : \quad & \widetilde{\mathbb{R}}^{(n-s) \times s} \rightarrow \mathbb{R}^{(n-s)s-rs}, \\ & \Pi_{21} \mapsto \Xi; \\ (\text{id}, h^{(3)}) \circ g^{(3)} : \quad & \widetilde{\text{St}}(n, s; U_1, U_3) \rightarrow \text{Skew}(s) \oplus \mathbb{R}^{(n-s)s-rs}, \\ & U_4 \mapsto (\Pi_{11}, \Xi). \end{aligned}$$

The remaining components are R_2, R_3 , and R_4 . The constraints for them come from the requirement that U as a whole is in $\text{St}(n, r)$. This gives

$$\begin{aligned} U^\top U &= \left((U_1 R_1, 0) + (M_4, U_3 R_2) \right)^\top \left((U_1 R_1, 0) + (M_4, U_3 R_2) \right) \\ &= \begin{pmatrix} R_1^\top U_1^\top U_1 R_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \left((U_4, U_3) \begin{pmatrix} R_4 & 0 \\ R_3 & R_2 \end{pmatrix} \right)^\top \left((U_4, U_3) \begin{pmatrix} R_4 & 0 \\ R_3 & R_2 \end{pmatrix} \right) \\ &= \begin{pmatrix} R_1^\top R_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \begin{pmatrix} R_4 & 0 \\ R_3 & R_2 \end{pmatrix}^\top \begin{pmatrix} R_4 & 0 \\ R_3 & R_2 \end{pmatrix} \\ &= \begin{pmatrix} R_1^\top R_1 + R_3^\top R_3 + R_4^\top R_4 & R_3^\top R_2 \\ R_2^\top R_3 & R_2^\top R_2 \end{pmatrix} = I_r \end{aligned}$$

Denote

$$R_0 := \begin{pmatrix} R_2 & R_3 \\ \mathbf{0} & R_4 \end{pmatrix} \in \text{upper}(r, r).$$

Then the $r \times r$ upper-triangular matrix R_0 should satisfy

$$R_0^\top R_0 = \begin{pmatrix} I_{r-s} & \mathbf{0} \\ \mathbf{0} & I_s - R_1^\top R_1 \end{pmatrix}.$$

Such R_0 is uniquely determined. Therefore, we get the following chart function for the components (M_4, R_2) :

$$\begin{aligned} (\text{id}, h^{(3)}) \circ g^{(3)} \circ p : \quad & \widetilde{\mathbb{R}}(n, s) \oplus \widetilde{\text{upper}}(r-s, r-s) \rightarrow \text{Skew}(s) \oplus \mathbb{R}^{(n-s)s-rs}, \\ & (M_4, R_2 - I_{r-s}) \mapsto (\Pi_{11}, \Xi). \end{aligned}$$

Putting everything together, we get the following chart function for the manifold \mathcal{M} :

$$\begin{aligned} \tilde{\varphi} &:= \left((\text{id}, h, h^{(2)}) \circ g, (\text{id}, h^{(3)}) \circ g^{(3)} \circ p, \text{id} \right) \circ \tilde{f} : \\ \mathcal{M} &\rightarrow \left(\text{Skew}(s) \oplus \mathbb{R}^{(r-s)s} \oplus \text{Skew}(r-s) \oplus \mathbb{R}^{(n-2r+s)(r-s)} \oplus \mathbb{R}^{(r-s)(r-s)} \right) \\ &\quad \oplus \left(\text{Skew}(s) \oplus \mathbb{R}^{(n-s-r)s} \right) \oplus \mathbb{S}_r \oplus \text{upper}(s, s), \\ (U, S) &\mapsto \left(\left(\Omega_{11}, \Omega_{21}, \Lambda_{11}, \Gamma, \Gamma^{(2)} \right), \left(\Pi_{11}, \Xi \right), \tilde{S} - \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix}, R_1 - I_s \right). \end{aligned}$$

The chart function φ is a coordinate slice of the chart function $\tilde{\varphi}$. Hence, $\mathcal{N}_{Z\#}$ is an embedded submanifold of \mathcal{M} . □