Tight Motion Planning by Riemannian Optimization for Sliding and Rolling with Finite Number of Contact Points - Supplementary Material

Dror Livnat Michael M. Bilevich Dan Halperin

A Analysis of Critical Manifolds in $SE(3)$

We recall that for tight assemblies, sliding and rolling motions might be required for separating the parts. In this section, we present our formal analysis of such motions.

We will define sub-manifolds, which we will refer to as the *critical manifolds*, in the configuration space $\mathcal{C} = SE(3)$ such that any path of motion that lies on those sub-manifolds will have a corresponding sliding and rolling motion in the workspace. Our key motivation for dealing with these sub-manifolds is to force the exploration of configurations at which the robot shares one or more contact points with the obstacles.

We will present a rigorous analysis of those critical manifolds, and of the way our algorithm produces sequences of desired points on these manifolds. For the analysis we need to impose certain conditions on these manifolds. These conditions are not necessarily met in our scenarios. However, our analysis here was essential in guiding us through the development of our algorithm, which in turn works well in practice.

First, we will look at the case of fixing a finite set S of (contact) points on the robot, and show how the resulting critical manifolds in the configuration space $\mathcal{C} = SE(3)$ can be used to perform a *sliding* motion of the points S on the obstacles. By sliding motion we mean that the points in S are touching the boundary of M_2 throughout the motion, and other than these points there is no intersection between the robot and the obstacles. Then, we will show a relaxation of this technique, to allow the contact points on the robot to continuously change during the sliding motion, allowing for what we refer to as a *rolling* motion; see below for details.

We recall that $M_1, M_2 \subseteq \mathbb{R}^3$ are (not necessarily connected[1]) 2-manifolds, referred to as the *obstacles* and *robot*, respectively. We denote by $F_{M_1} : \mathbb{R}^3 \to \mathbb{R}$ the SDF function (See Section II-A) for the obstacles M_1. In this section we will also assume that F_{M_1} is differentiable, and that we can easily evaluate both $F_{M_1}(p)$ and $\nabla_p F_{M_1}$ for any $p \in \mathbb{R}^3$.

A.1 Single-Contact Critical Manifolds

The main ingredient in our analysis, is the *contact-SDF* of some point $p \in M_2$, which is the signed distance of that point p transformed by $q \in SE(3)$, formally defined as:

[1] It was natural to assume, as we do, that the robot is connected; however, the analysis here does not rely on this fact.

$$F^p_{M_1} : SE(3) \to \mathbb{R} \ , \tag{1}$$

$$F^p_{M_1}(q) \coloneqq F_{M_1}(\varphi(q) \cdot p) \ , \tag{2}$$

with $\varphi(q) = A_q$ the corresponding 4×4 matrix. We identify $p \in \mathbb{R}^3$ with $(p, 1) \in \mathbb{R}^4$.

Note that we use the notation of $SE(3)$ as a Lie-group, namely a manifold with a group multiplication. While we could do our entire analysis under that framework (similar to [1]), we will simplify our analysis by using calculus in Euclidean spaces[2]. We identify $SE(3) \simeq \mathbb{R}^6$ with coordinates $(x, y, z, \theta_x, \theta_y, \theta_z)$, as presented in the main paper (Section III). We then look at $\varphi(q) \cdot p = g(q, p)$ with $g : \mathbb{R}^6 \times \mathbb{R}^3 \to \mathbb{R}^3$ a smooth function.

Under this identification, we get that the function $F^p_{M_1}$ is $F^p_{M_1} : \mathbb{R}^6 \to \mathbb{R}$. Then, using the implicit function theorem [2], for any q such that $\nabla_q F^p_{M_1} \neq 0$, there is some small neighborhood $U \subseteq \mathbb{R}^6$ for which $(F^p_{M_1})^{-1}(\{0\}) \cap U$ is a 5-manifold. For the simplicity of our analysis, we will assume that $\forall q : \nabla_q F^p_{M_1} \neq 0$ and thus the set

$$M^F_p \coloneqq (F^p_{M_1})^{-1}(\{0\}) \tag{3}$$

is a 5-manifold. We call the set M^F_p a *(single-contact) critical manifold*, and the point $p \in M_2$ is its corresponding *contact point*. In other words, configurations $q \in M^F_p$ are such that by placing the body M_2 in configuration q, the point $p \in M_2$ will now touch the body M_1. Note that currently, there might be penetrations.

Intuitively, generating the desired sliding motion requires moving in small steps along the tangent space at some configuration. The details are presented in Section A.3. To compute the tangent space, we in fact only need the gradient of $F^p_{M_1}$:

Lemma A.1. *For any configuration $q \in M^F_p$, the gradient to the contact-SDF is orthogonal to the tangent space of M^F_p when embedded in $SE(3) \simeq \mathbb{R}^6$:*

$$\forall q \in M^F_p : \nabla_q F^p_{M_1} \perp T_q M^F_p \ . \tag{4}$$

Proof. Let v denote some tangent vector in $T_q M^F_p$. Then there exists a path $\gamma : (-1, 1) \to M^F_p$ with $\gamma(0) = q$ and $\dot\gamma(0) = v$. By the definition of the critical manifold, $F \circ \gamma : (-1, 1) \to \mathbb{R}$ has $F \circ \gamma \equiv 0$. On the other hand, from the chain rule, we have:

$$\frac{d}{dt} (F \circ \gamma)(0) = \nabla_{\gamma(0)} F^p_{M_1} \cdot \dot\gamma(0) = \langle \nabla_q F^p_{M_1}, v \rangle = 0 \ . \tag{5}$$

Since Equation 5 is true for any tangent vector, we get that $\nabla_q F^p_{M_1} \perp T_q M^F_p$. $\square$

A.2 Multi-Contact Critical Manifolds

Let us now consider the case of multiple contact points. We define the *contact-SDF* of the points $p_1, \ldots, p_k \in M_2$ as:

[2] We reduce our dynamics and analysis to Euclidean space, instead of the manifold $SE(3)$ both for simplicity of analysis, and for simplicity in the implementation of our algorithm. Throughout this work, we will refer to $SE(3)$ as its Euclidean representation.

$$F_{M_1}^{p_1,\dots,p_k} : SE(3) \to \mathbb{R}^k \ , \tag{6}$$

$$F_{M_1}^{p_1,\dots,p_k}(q) \coloneqq \left(F_{M_1}^{p_1}(q), \dots, F_{M_1}^{p_k}(q)\right) \ . \tag{7}$$

Similarly, if we assume that the Jacobian matrix of $F_{M_1}^{p_1,\dots,p_k}$ has full rank, we get that the set

$$M_{p_1,\dots,p_k}^F \coloneqq (F_{M_1}^{p_1,\dots,p_k})^{-1}\left(\{\vec{0}\}\right) \tag{8}$$

is a $(6-k)$-manifold. We then call this set $M_{p_1,\dots,p_k}^F$ a *(multi-contact) critial manifold.*

Note that by definition, the Jacobian matrix of $F_{M_1}^{p_1,\dots,p_k}$ is :

$$D_q F_{M_1}^{p_1,\dots,p_k} = \begin{pmatrix} -\nabla_q F_{M_1}^{p_1}- \\ \vdots \\ -\nabla_q F_{M_1}^{p_k}- \end{pmatrix} \in \mathbb{R}^{k\times 6} \tag{9}$$

Once again, we are able to compute the tangent space of the critical manifold as asserted in the following lemma.

Lemma A.2. *For any configuration $q \in M_{p_1,\dots,p_k}^F$, any tangent vector $v \in T_q M_{p_1,\dots,p_k}^F$ is orthogonal to any row of $D_q F_{M_1}^{p_1,\dots,p_k}$:*

$$\forall q \in M_{p_1,\dots,p_k}^F, \forall v \in T_q M_{p_1,\dots,p_k}^F : v \perp \nabla_q F_{M_1}^{p_1}, \dots, \nabla_q F_{M_1}^{p_k} \tag{10}$$

Proof. Similar to the proof of Lemma A.1. □

A.3 Sliding on the Critical Manifold: First Attempt

Recall our motivation: We would like the robot M_2 to slide over the obstacles M_1. In other words, say they are already initially touching at k contact points. Then we would like to find a path of motion for M_2 such that we keep these k contact points throughout the motion.

Formally, given some initial configuration $q_0 \in SE(3)$ of M_2, such that q_0 is on the critical manifold $M_{p_1,\dots,p_k}^F$, we would like to find a sequence of poses $q_1, q_2, \dots \in M_{p_1,\dots,p_k}^F$ which also lie on the critical manifold. We will refer to this sequence as *sliding on the critical manifold*, which in turn will represent the points $p_1, \dots, p_k$ sliding on the body M_1. The output is a sequence of poses, where we assume that consecutive pairs are sufficiently close, such that a collision-free path between each pair (q_i, q_{i+1}) can be easily computed. Note that in the following analysis we do not assume anything on that path connecting the two configurations.

One might look at a simple algorithm, borrowed from Riemannian optimization[3] [3], given some direction $v \in \mathbb{R}^6 \simeq SE(3)$, as shown in Algorithm A1.

The operator Project(v, W) takes a vector v and projects it onto a vector sub-space W. The operator $\mathcal{R}_{M_{p_1,\dots,p_k}^F}$ is the *retraction* operator, which takes a tangent vector of a manifold, and projects it onto the manifold: $\mathcal{R}_{M_{p_1,\dots,p_k}^F} : TM_{p_1,\dots,p_k}^F \to M_{p_1,\dots,p_k}^F$.

[3] Although our sliding problem is not an optimization problem, we can frame it as such: let some point $q \in SE(3)$ such that $q - q_0 = v$. Then we would like to optimize the function $h : \mathbb{R}^6 \to \mathbb{R}$, $h(q) = \frac{1}{2}||q - q_0||_2^2$, restricted on our critical manifold. We then use Riemannian gradient descent to derive Algorithm A1.

Algorithm A1 Sliding on the critical manifold $M^F_{p_1,\ldots,p_k}$

Require: $q_0 \in M^F_{p_1,\ldots,p_k}$, $v \in \mathbb{R}^6$, $\eta > 0$
Ensure: $q_t \in M^F_{p_1,\ldots,p_k}$
  while true **do**
    $v_t \leftarrow \text{Project}(v, T_{q_t} M^F_{p_1,\ldots,p_k})$
    $q'_{t+1} \leftarrow q_t + \eta \cdot v_t$
    $q_t \leftarrow \mathcal{R}_{M^F_{p_1,\ldots,p_k}}(q'_{t+1})$
  end while

There are many ways to implement this retraction operator, some are described in [4]. We present our implementation in Section A.6 below.

Notice that even though we indeed slide on the critical manifold, under various assumptions (which need not hold), in fact this motion does not describe our intended "sliding" motion of M_2 on M_1. We next explain this issue and then propose a solution.

A.4 The Problem: Rolling Contact Points

Thus far, by our definition of critical manifolds, we require that any initial contact point $p \in M_2$ will be the one that remains in contact with M_1.

To illustrate why this is undesired, let us consider a simple two-dimensional example, where $M_1, M_2 \subseteq \mathbb{R}^2$ and our configurations are in $SE(2)$, i.e., translation and rotation in two-dimensions. Let M_1 be a rectangle (which we refer to as the *floor*) and M_2 is a circle (which we refer to as the *wheel*).

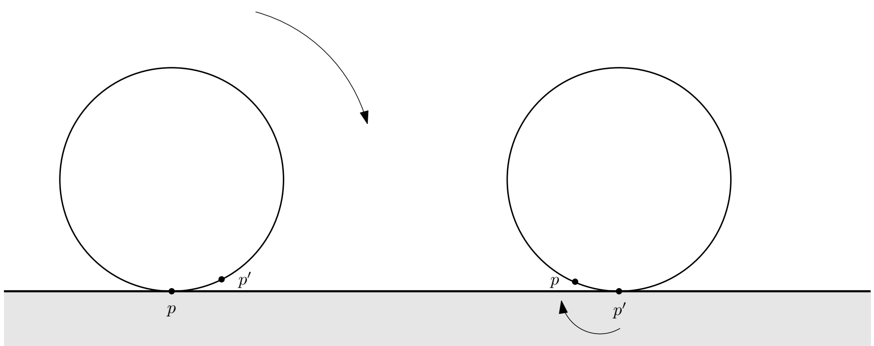

Figure 1: For any rotation of the wheel, the contact point with the floor changes.

Initially, there is only one contact point p, at the bottom of the wheel. If we rotate the wheel clockwise, in any rotation θ, the contact point with the floor will then become p', which is a different point on the original M_2. Similarly, by taking a box as M_1, and a sphere as M_2 we can generalize this example to three dimensions.

Hence we seek a more general definition of critical manifolds, which will also account for this case, which we refer to as *rolling*.

A.5 The Solution: Weak-Contact Critical Manifolds

Let us assume that at some pose $q_0 \in SE(3)$ of M_2, the points $p_1, \ldots, p_k$ are contact points. For any point $p_i \in M_2$, let us define a small open neighborhood $U(p_i) \subseteq M_2$ of p_i. We then can define the *(single) weak-contact-SDF* for any point $p_i \in M_2$:

$$F_{M_1}^{U(p_i)} : SE(3) \to \mathbb{R} \ , \tag{11}$$

$$F_{M_1}^{U(p_i)}(q) \coloneqq \min_{p' \in U(p_i)} F_{M_1}^{p'}(q) \ , \tag{12}$$

and the *(multiple) weak-contact-SDF* as:

$$F_{M_1}^{U(p_1),\ldots,U(p_k)} : SE(3) \to \mathbb{R}^k \ , \tag{13}$$

$$F_{M_1}^{U(p_1),\ldots,U(p_k)}(q) \coloneqq \left(F_{M_1}^{U(p_1)}(q), \ldots, F_{M_1}^{U(p_k)}(q) \right) \ . \tag{14}$$

Then the *weak-contact critical manifold* is the set:

$$M^F_{U(p_1),\ldots,U(p_k)} \coloneqq (F_{M_1}^{U(p_1),\ldots,U(p_k)})^{-1} \left(\{\vec{0}\} \right) \ . \tag{15}$$

We note that our weak-contact-SDF function is not necessarily differentiable. Making it differentiable requires assumptions on the dependency on the parameter $p_i \in M_2$. The exact conditions are presented in Section A.7. For the sake of our analysis, we will assume that they are in fact differentiable.

Also note that now a pose $q \in M^F_{U(p_1),\ldots,U(p_k)}$ means that for each $i = 1, \ldots, k$, some point $p'_i \in U(p_i)$ (after transforming by q) lies on M_1. Since it is also the minimum, we get the byproduct that for $U(p_i)$ there are no penetrations into M_1. Hence paths of motion in $M^F_{U(p_1),\ldots,U(p_k)}$ can also describe the sliding and rolling motions, presented in Section A.4, as we require throughout the motion that for each neighborhood $U(p_i)$ there will be some point in that neighborhood which touches the obstacle. By lifting the strict requirement that p_i will be the one touching the obstacle, we are now allowed to continuously change the contact point, enabling the rolling motion. Note that for configurations where p_i remains the contact point, if such exist, the configurations are contained in our weak-contact critical manifold.

Then we can repeat our algorithm to slide on the weak-contact critical manifolds, as shown in Algorithm A2.

Algorithm A2 Sliding on the weak-contact critical manifold $M^F_{U(p_1),\ldots,U(p_k)}$

Require: $q_0 \in M^F_{U(p_1),\ldots,U(p_k)}$, $v \in \mathbb{R}^6$, $\eta > 0$
Ensure: $q_t \in M^F_{U(p_1),\ldots,U(p_k)}$
  while true **do**
    $v_t \leftarrow \text{Project}(v, T_{q_t} M^F_{U(p_1),\ldots,U(p_k)})$
    $q'_{t+1} \leftarrow q_t + \eta \cdot v_t$
    $q_t \leftarrow \mathcal{R}_{M^F_{U(p_1),\ldots,U(p_k)}}(q'_{t+1})$
  end while

Note that we now project the points onto the weak-contact critical manifolds, effectively allowing to change the current contact points at each iteration.

In order to project onto the tangent-space of the weak-contact critical manifold, we need to compute its gradient. To do that, we use the following lemma to compute the gradient of a single weak-contact SDF (and then the multiple contacts follows as before):

Lemma A.3. *Let $q \in M^F_{U(p_i)}$, and assume that $p'_i = \arg\min_{p'} F^{p'}_{M_1}(q)$. Then:*

$$\nabla_q F^{U(p_i)}_{M_1} = \nabla_q F^{p'}_{M_1} \ . \tag{16}$$

The proof for Lemma A.3 follows immediately from the following general observation from calculus:

Lemma A.4. *Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ differentiable. Let $g : \mathbb{R}^m \to \mathbb{R}$ be $g(y) = \min_{x \in \mathbb{R}^n} f(x, y)$.*

Define $x^* : \mathbb{R}^m \to \mathbb{R}^n$ be $x^*(y) = \arg\min_{x \in \mathbb{R}^n} f(x, y)$. Then if x^* is differentiable, we have:

$$\frac{\partial g}{\partial y}(y_0) = \frac{\partial f}{\partial y}\left(x^*(y_0), y_0\right) \ . \tag{17}$$

Proof. Follows from the chain rule. Let us define the function $\psi : \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^n$, as:

$$\psi(y) = (x^*(y), y) \tag{18}$$

Then by definition, $g = f \circ \psi$. Hence:

$$\frac{\partial g}{\partial y}(y_0) = \frac{\partial f}{\partial \psi}(\psi(y_0)) \cdot \frac{\partial \psi}{\partial y}(y_0) \ , \tag{19}$$

and we have:

$$\frac{\partial f}{\partial \psi} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \ ; \ \frac{\partial \psi}{\partial y} = \begin{pmatrix} \frac{\partial x^*}{\partial y} \\ 1 \end{pmatrix} \ . \tag{20}$$

Combining the above:

$$\frac{\partial g}{\partial y}(y_0) = \left(\frac{\partial f}{\partial x}(\psi(y_0)) \cdot \frac{\partial x^*}{\partial y}(y_0) \right) + \frac{\partial f}{\partial y}(\psi(y_0)) \ . \tag{21}$$

Finally, we notice that since $x^*(y_0)$ is a local minima for f with respect to the variable x, $\frac{\partial f}{\partial x}(x^*(y_0), y_0) = 0$, and we get the desired result. □

Hence to compute the gradient, all we need to do is to re-compute the contact points while generating the sequence q_t. This is the method described in Section V in the paper, and its pseudo-code is also shown in Algorithm A3.

Algorithm A3 Project$(v, T_{q_t} M^F_{U(p_1),\ldots,U(p_k)})$

---

for $i = 1, \ldots, k$ **do**
  $p'_i \leftarrow \arg\min_{p' \in U(p_i)} F^{p'}_{M_1}(q_t)$
  $g_i \leftarrow \nabla_{q_t} F^{p'_i}_{M_1}$
end for
$g_1, \ldots, g_k \leftarrow \mathrm{GramSchmidt}(g_1, \ldots, g_k)$
$v' \leftarrow v$
for $i = 1, \ldots, k$ **do**
  $v' \leftarrow v' - \frac{\langle v', g_i \rangle}{||g_i||^2} \cdot g_i$
end for
return v'

A.6 Using Weak-Contact Critical Manifolds for Separation

In this section we describe our chosen method for retracting a configuration $q \in SE(3)$ onto the manifold $M^F_{U(p_1),\ldots,U(p_k)}$ [4]. We note that the retraction operator should get a tangent vector and project it onto the manifold. Since our chosen q will be of the form $q_0 + v$, with $q_0 \in M^F_{U(p_1),\ldots,U(p_k)}$ and $v \in T_{q_0} M^F_{U(p_1),\ldots,U(p_k)}$, it will suffice to describe a method that projects a configuration (and not a tangent vector) onto the manifold.

Recall that a configuration q is in $M^F_{U(p_1),\ldots,U(p_k)}$ if and only if for all p_i (for $i = 1, \ldots, k$), we have $F^{U(p_i)}_{M_1}(q) = 0$. Let us define the following function $\mathcal{L}_{U(p_1),\ldots,U(p_k)} : SE(3) \to \mathbb{R}$:

$$\mathcal{L}_{U(p_1),\ldots,U(p_k)}(q) = \sum_{i=1}^{k} \left(F^{U(p_i)}_{M_1}(q) \right)^2 . \tag{22}$$

As we see from the following lemma, finding a (global) minimum for $\mathcal{L}_{U(p_1),\ldots,U(p_k)}$ yields a configuration on the manifold $M^F_{U(p_1),\ldots,U(p_k)}$.

Lemma A.5. *Suppose that the manifold $M^F_{U(p_1),\ldots,U(p_k)}$ is not empty. Then for any configuration $q \in SE(3)$, q is a global minima of $\mathcal{L}_{U(p_1),\ldots,U(p_k)}$ if and only if is on the manifold $q \in M^F_{U(p_1),\ldots,U(p_k)}$.*

Proof. We note that $\mathcal{L}_{U(p_1),\ldots,U(p_k)}(q) \geq 0$ for any $q \in SE(3)$. Hence $\min \mathcal{L}_{U(p_1),\ldots,U(p_k)} \geq 0$. Since we assume that there exists some configuration $q' \in M^F_{U(p_1),\ldots,U(p_k)}$, we get that $\mathcal{L}_{U(p_1),\ldots,U(p_k)}(q') = 0$ and thus the value 0 is indeed a global minima of $\mathcal{L}_{U(p_1),\ldots,U(p_k)}$. Thus a configuration q is a global minima if and only if $\mathcal{L}_{U(p_1),\ldots,U(p_k)}(q) = 0 \iff q \in M^F_{U(p_1),\ldots,U(p_k)}$. □

Hence for our retraction operator, it suffices to apply any optimization algorithm on the function $\mathcal{L}_{U(p_1),\ldots,U(p_k)}$. Since our functions $F^{U(p_i)}_{M_1}$ are not necessarily convex, and so does $\mathcal{L}_{U(p_1),\ldots,U(p_k)}$, we cannot guarantee that we will find a global minima. Still, we hope that if we start close enough to the manifold, we will indeed get a value that will remain close enough to that manifold. We note that non-convex optimization is extensively researched, e.g., [5],[6],[7],[8], and as such, we defer to state of the art methods for our retraction operator. In our work, we use the Adam optimizer [9] for this function $\mathcal{L}_{U(p_1),\ldots,U(p_k)}$, with $x_0 = q'_{t+1}$.

A.7 The Differentiability of the $\arg\min$ Function

In Section A.5 we assume that the function for the points $p_i \in \mathbb{R}^3$, the function $F^{U(p_i)}_{M_1}$ is differentiable. This would be true if the $\arg\min$ function, defined as

$$p^* : SE(3) \to \mathbb{R}^3; \tag{23}$$

$$p^*(q) := \arg\min_{p' \in U(p_i)} F^{p'}_{M_1}(q) , \tag{24}$$

is differentiable. That is not necessarily true. We present one (sufficient) condition for which this claim holds, which is shown in [10].

[4] We can also apply the same method for projection onto the manifold $M^F_{p_1,\ldots,p_k}$

The function $G_{M_1} : \mathbb{R}^3 \times SE(3) \rightarrow \mathbb{R}$, defined by $G_{M_1}(p, q) = F_{M_1}(\varphi(q)\cdot p)$ is differentiable as a composition of differentiable functions, and then the function $F_{M_1}^{U(p_i)}$ is a composition of differentiable functions:

$$F_{M_1}^{U(p_i)}(q) = G_{M_1}(p^*(q), q) \ . \tag{25}$$

Theorem A.6. *Suppose there exists some $\kappa \geq 0$ and $\sigma > 0$ such that the function*

$$G_{M_1}(p, q) + \frac{\kappa}{2}||p||^2 - \frac{\sigma}{2}||q||^2 \tag{26}$$

is convex. Also assume that for any $q \in SE(3)$, $p^*(q)$ exists and is unique. Then the function p^* is differentiable almost everywhere.

References

[1] E. Gallo, "The SO(3) and SE(3) Lie algebras of rigid body rotations and motions and their application to discrete integration, gradient descent optimization, and state estimation," *arXiv preprint arXiv:2205.12572*, 2022.

[2] S. G. Krantz and H. R. Parks, "The implicit function theorem. modern birkhäuser classics," *History, theory, and applications, Reprint of the 2003 edition. Birkhäuser/Springer, New York, pp. xiv*, vol. 163, 2013.

[3] Y. Fei, X. Wei, Y. Liu, Z. Li, and M. Chen, "A survey of geometric optimization for deep learning: From euclidean space to riemannian manifold," 2023.

[4] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2008. [Online]. Available: https://doi.org/10.1515/9781400830244

[5] P. Jain, P. Kar *et al.*, "Non-convex optimization for machine learning," *Foundations and Trends® in Machine Learning*, vol. 10, no. 3-4, pp. 142–363, 2017.

[6] Y. Carmon, J. C. Duchi, O. Hinder, and A. Sidford, "Accelerated methods for non-convex optimization," *SIAM Journal on Optimization*, vol. 28, no. 2, pp. 1751–1772, 2018.

[7] M. Zaheer, S. Reddi, D. Sachan, S. Kale, and S. Kumar, "Adaptive methods for non-convex optimization," *Advances in neural information processing systems*, vol. 31, 2018.

[8] M. Danilova, P. Dvurechensky, A. Gasnikov, E. Gorbunov, S. Guminov, D. Kamzolov, and I. Shibaev, "Recent theoretical advances in non-convex optimization," in *High-Dimensional Optimization and Probability: With a View Towards Data Science*. Springer, 2022, pp. 79–163.

[9] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[10] J. Ross and D. W. Nyström, "Differentiability of the argmin function and a minimum principle for semiconcave subsolutions," 2019.