SUBSPACE LEARNING FOR DATA ARISING FROM A UNION OF SUBSPACES OF HIGH RELATIVE DIMENSION

by

Tianyu Ding

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
June, 2021

Abstract

As we have witnessed the rapid growth of statistical machine learning over the past decades, the ability of processing big and corrupted data becomes increasingly important. One of the major challenges is that structured data, such as images, videos and 3D point clouds, involved in many application scenarios are high-dimensional. Conventional techniques usually approximate the high-dimensional data with low-dimensional structures by fitting the data with one or more linear subspaces. However, their theory and algorithms are restricted to the setting in which the underlying subspaces have a low relative dimension compared to the ambient space.

This thesis attempts to advance the understanding of subspace learning for data arising from subspaces of high relative dimension, as well as develop efficient algorithms for handling big and corrupted data. The first major contribution of this thesis is a theoretical analysis that extends Dual Principal Component Pursuit (DPCP), a non-convex approach that learns a *hyperplane* in the presence of *noiseless* data, to learn a *subspace* of any dimension with *noisy* data. We provide geometric and probabilistic analyses to characterize how the principal angles between the global solution and the orthogonal complement of the subspace behave as a function of the noise level. Moreover, we improve the DPCP theory in multi-hyperplane case with a more interpretable geometric analysis and a new statistical analysis.

The second major contribution of this thesis is the development of a linearly convergent method for non-convex optimization on the Grassmannian. We show that if the objective function satisfies a certain Riemannian regularity condition (RRC) with respect to some point in the Grassmannian, then a Projected Riemannian Sub-Gradient Method (PRSGM) converges at a linear rate to that point. In particular, we prove that the DPCP problem for learning a single subspace satisfies the RRC and PRSGM converges linearly to a neighborhood of the orthogonal complement of the subspace with error proportional to the noise level. We also extend the RRC to DPCP for a union of hyperplanes and prove the linear convergence of PRSGM to a specific hyperplane. Finally, both synthetic and real experiments demonstrate the superiority of the proposed method.

Primary Reader and Advisor: Daniel P. Robinson

Secondary Reader: René Vidal

Dedicated to my parents, Yuyuan Ma and Xueqing Ding,

and my wife, Tong Jin,

for their unconditional acceptance and love.

Acknowledgments

First of all, I would like to express my greatest appreciation to my advisor, Professor Daniel P. Robinson, for his endless understanding and support during my entire graduate life. Daniel's profound knowledge in the optimization field is a shining example to me for being an outstanding researcher, and his extraordinary scientific presentation skills influence me a lot whenever I am writing and communicating with others. I am also thankful to him for bringing me into the project of non-convex subspace learning and clustering. Without such a wonderful opportunity, this thesis would not be the one it is.

I am also very grateful to Professor René Vidal, who is the co-investigator of the aforementioned project and is the person I am most honored to collaborate with, for his powerful guidance and help along the road of research. Have been working with René for the past few years, I was exposed to the marvelous world of computer vision, which opens a new research and career direction for me. This thesis would not be possible without his expertise in machine learning and broad vision in research.

I am very thankful to Professor Zhihui Zhu from the University of Denver for always being supportive and encouraging in my research and life, and serving in my dissertation defense panel. Zhihui is both a great mentor and a close friend. He helped me a lot when I was at the trough of my doctorate program, which I deeply appreciate.

I am also thankful to Professor Amitabh Basu for serving in my Graduate Board Oral committee and dissertation defense panel. I really appreciate his time and effort! Then I would express my gratitude to Prof. Donniell Fishkind and Prof. Nicolas Charon for serving in my Candidacy Exam committee, to Prof. Michael Kazhdan and Prof. Michael Dinitz for serving in my Graduate Board Oral committee, and to Prof. Avanti Athreya for recommending me during my graduate application.

I also thank Professor Manolis C. Tsakiris from ShanghaiTech University for many fruitful discussions during our group meetings. I would also thank Tianjiao Ding and Yunchen Yang, who are students of Manolis, helped me a lot in conducting the real experiments in many of our collaborations.

At Hopkins, I thank the Department of Applied Mathematics and Statistic and the Department of Computer Science for providing me with abundant support and gathering so many wonderful people together. I thank the Center for Imaging Science (CIS) for offering me numerous computational resources. I also want to thank the National Science Foundation grant 1704458 for supporting my graduate study and research life here.

Finally and most importantly, this thesis would not have been possible without the company of my family. I thank my parents and my wife in China for their eternal trust and love. They give a meaning to every single unit of me.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Many real world applications in machine learning, computer vision, and signal processing aim to discover certain structures from a large amount of collected data that are usually of high dimensions. For example, irrelevant information removal from web image search results [19, 71, 91, 135] involve distinguishing the query intent among all the retrieved images, whose dimension can be hundreds to millions according to the number of pixels (resolution) of an image by viewing it as a long vector. Another closely related example is video abnormal event discovery [27, 98] in which the data are of even higher dimensions since a video is treated as a sequence of images. For these applications, however, directly working in the high-dimensional raw space is both inefficient and unnecessary. It is expected that a certain hidden structure associated with the data can be well-represented using features with a lower dimension due to the fact that such structures impose additional constraints on the data, and thus the problem is transformed into learning a compact representation of the dataset.

Among the various techniques for modeling specific structures of high-dimensional data, perhaps the simplest one is as linear subspaces, which assumes that the data points are drawn from one or more linear subspaces with dimension fewer than that of the ambient (raw) space. Despite its simplicity, it has been shown effective in a broad range of application scenarios, such as dimensionality reduction [53, 117], human face identification and clustering [7, 37], motion segmentation [125, 127], multiple view geometry [2] and so on. We organize this chapter as follows. In Section 1.1, we introduce the data modeling for learning linear subspaces from corrupted data as well as its example applications. Next, in Section 1.2, we discuss the challenges when the underlying subspaces are of high relative dimension, namely the subspace dimension is high relative to the ambient dimension, which is no longer appropriately tackled by the prevalent methods designed for the low relative dimension regime. We finally summarize the main contributions of this thesis in Section 1.3, and provide the notation used throughout the thesis in Section 1.4.

1.1 Learning linear subspaces from corrupted data

1.1.1 Data modeling

A rule of thumb in dealing with high-dimensional data is that we aim to find ways to interpret them with fewer degree of freedoms. This not only facilitates the success of many real world applications, but also provides insights on characteristics of the underlying structure of the datasets. In particular, we are interested in fitting one or more linear subspaces to data points, depending on the specific task and dataset.

Inliers are the data points that exactly lie in the underlying subspaces that we aim to identify, while *outliers* are the data points that do not exhibit the linear structure. The existence of outliers corrupts the datasets and adversely affects the results of data analysis methods. Besides outliers, another form of corruption in real world data is *noise*, which means that the inliers are perturbed so that they no longer exactly lie on but close to the subspaces, i.e., they are *noisy inliers*. Note that noise usually comes from systematic errors [132, 150], e.g., measurement and sensor error, and is difficult to be eliminated in the data gathering stage. The above two forms of corruption in real world datasets pose significant challenges to the subspace recovery task.

Single subspace learning. In the simplest case, inlier points are drawn from a single subspace, so that the problem is to robustly learn the underlying subspace in the presence of both outliers and noise. For example, it is well-known [7] that images of a human face under different lighting conditions approximately lie in a 9-dimensional linear subspace, and thus screening the face pictures of an individual from other irrelevant images appears as an outlier removal task. Another example in computer vision is the robust homography estimation from image correspondences across multiple views [2, 29], which can be cast as robust subspace learning with dimension 8 and 26 for two and three views, respectively. In fact, Principal Component Analysis (PCA) [52, 55, 84] is a classical solution for learning such a linear subspace from data, and it enjoys a closed form solution via the Singular Value Decomposition (SVD). PCA works well even when the data is noisy, however, the least square loss employed in PCA causes its performance to be sensitive to outliers, and thus limits its performance to a large extent. However, there are many robust PCA methods [9, 11, 13, 28, 61, 75, 76, 94, 119, 136] that have been developed over the past decade.

Multiple subspaces clustering. In may cases, it is inappropriate to model the dataset with a single subspace; instead, inlier points are assumed to be drawn from a union of subspaces, and the goal is to estimate the underlying subspaces and cluster the data points into their respective groups. For example, a dataset consisting of face images from more than one human subject is naturally treated as a union of subspaces model [7]. As another example, point trajectories corresponding to the motions of multiple rigid bodies in a video lie approximately in a union of 3-dimensional affine subspaces [104, 105]. Many other real world applications involve exploring such multi-subspace structure, including document clustering [97], motion segmentation [125, 127], 3D point cloud analysis [90, 92], gene expressions [79, 115] and so on. Similar to single subspace learning, this is an unsupervised problem so that the hidden structures need to be automatically learned from data. Nevertheless, unlike the former, clustering multiple subspaces is more difficult due to the potentially complicated relative arrangement of the underlying subspaces. Although numerous techniques have been developed in this area, the most well-known approaches are based on sparse or low-rank representations of the data [35, 36, 37, 69, 70, 72, 124, 141, 142, 143], and we refer the reader to [123, 127] for more details.

1.1.2 Example applications

We now introduce two examples of robustly learning linear subspaces from 3D point cloud data, namely 3D roadplane detection (Section 1.1.2.1) and 3D plane clustering from indoor scenes (Section 1.1.2.2).

1.1.2.1 3D roadplane detection

In the task of 3D road plane detection, we are given a 3D point cloud of a road scene and the goal is to learn an affine plane $\mathcal{A} = \mathcal{H} + \boldsymbol{t} \subset \mathbb{R}^3$ as a model for the road. This is important in autonomous driving applications. Here $\mathcal{H}$ is a plane through the origin with normal vector $\boldsymbol{b}$ and $\boldsymbol{t}$ is its translation with respect to the origin; this latter is the center of the laser sensor. Hence the task is to estimate $\boldsymbol{b}$ and $\boldsymbol{t}$, which are taken to be co-linear in order to resolve the inherent ambiguity in estimating $\boldsymbol{t}$. In turn, this can be converted to a linear subspace learning problem by working in homogeneous coordinates, i.e., by embedding $\mathcal{A}$ into the linear hyperplane $\bar{\mathcal{H}} \subset \mathbb{R}^4$ with normal vector $\bar{\boldsymbol{b}} = [\boldsymbol{b}^\top \;\; -\boldsymbol{t}^\top \boldsymbol{b}]^\top$, through the mapping $\boldsymbol{x} \mapsto [\boldsymbol{x}^\top \;\; 1]^\top$. Figure 1.1 gives an illustration of the road detection challenge of the KITTI dataset [40], in which the image data and the depth information for each pixel are collected by a laser scanner. The depth data can then be used to reconstruct a 3D point cloud corresponding to the scene. Note that this is exactly a real world application of robust single subspace learning since the 3D point cloud datasets are usually noisy and corrupted by gross outliers due to the imperfect depth estimation of the laser sensor.

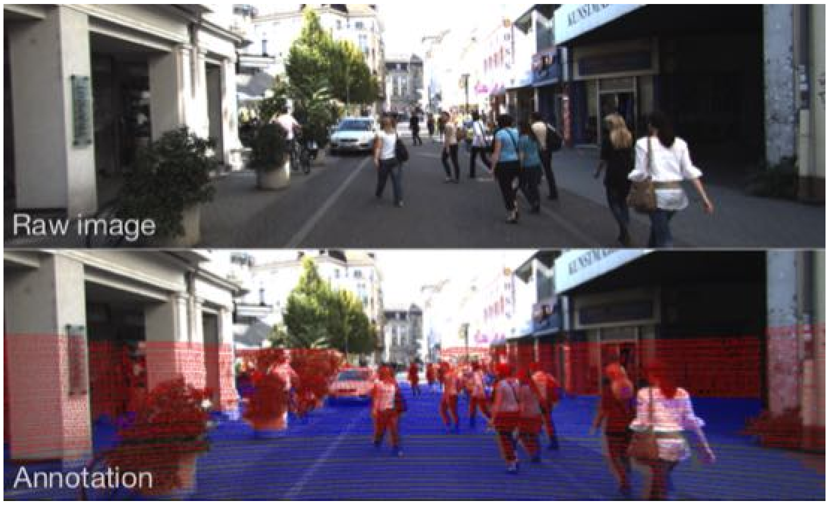


Figure 1.1. An illustration of the 3D roadplane detection problem. The raw image is from KITTY-CITY-71 [40]. We annotate the frame such that the (noisy) inlier points associated with the roadplane are in blue and outlier points are in red. The goal is to identify the underlying roadplane.

1.1.2.2 3D plane clustering from indoor scenes

An interesting problem is that of fitting multiple planes to 3D indoor scene data, which usually appears in robotics applications where a robot navigates an indoor environment, e.g., kitchens and bedrooms, and reasons about the interior building structures, e.g., desktops and walls. Although the planes associated with an indoor scene are affine in $\mathbb{R}^3$, we work in homogeneous coordinates by adding a 1 as a fourth coordinate, which is similar to the practice used for single roadplane detection (see Section 1.1.2.1), and the task is then transformed into a multi-hyperplane clustering problem in $\mathbb{R}^4$. Figure 1.2 gives an illustration with frames from the real dataset NYUdepthV2 [92], for which the indoor RGB images with depth information are collected by a Microsoft kinect sensor. Again, this problem is challenging not only because it involves the interplay of more than one underlying subspace but also because of the considerable amount of noise introduced during the collection of the real data.

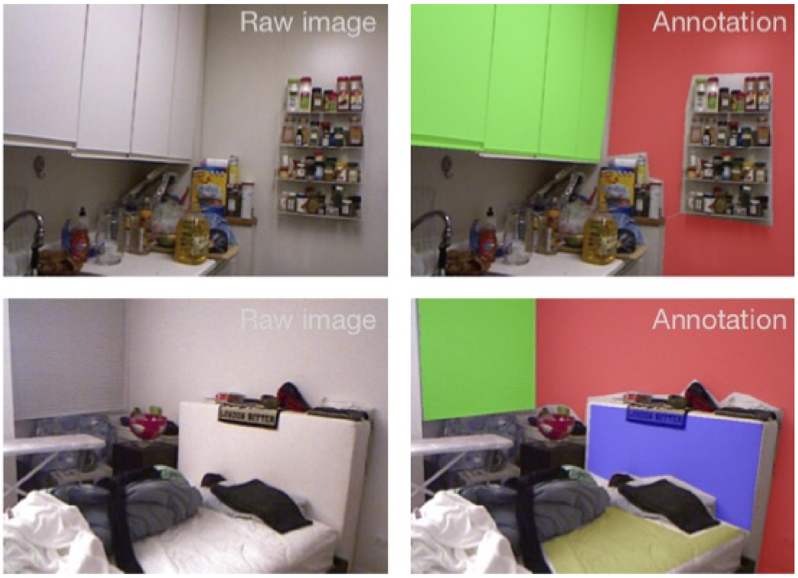


Figure 1.2. An illustration of the 3D plane clustering from an indoor scene. The raw images are from NYUdepthV2 [92]. We annotate the frames such that the (noisy) inlier points associated with each plane are in the same color. The goal is to learn the underlying plane arrangement given the 3D point cloud data.

1.2 High relative dimension challenge

Although the problem of fitting one or more subspaces to a dataset has a long history (plus numerous robust subspace recovery methods [62] and subspace clustering methods [123, 127] have emerged over the past twenty years), the existing methods typically assume that high-dimensional data can be well-approximated by low-dimensional structures. In other words, they require the dimension of the underlying subspaces to be relatively low compared to the dimension of the ambient space. This assumption advances the derivation of strong theoretical results and the development of efficient implementations since inliers of a subspace with low-dimensional structure are more well-separated from inliers of other subspaces as well as outliers in a fixed ambient space. For example, the success of sparse or self-expressive subspace clustering approaches [35, 36, 37, 72, 124, 141, 142, 143] rely on the property that each data point can be represented by a linear combination of a few other points belonging to the same subspace. However, this property is no longer valid in the *high relative dimension* regime, where an underlying subspace itself is high-dimensional, e.g., a hyperplane, since it is difficult to find a sparse representation of the inlier points.

As already mentioned before, many computer vision applications involve learning a single hyperplane (e.g., pose estimation in multi-view geometry [2], detection of planar structures in 3D point clouds [40, 92]), or clustering multiple hyperplanes (e.g., motion segmentation [105, 125, 128, 130], hybrid system identification [4, 129], sparse component analysis [41, 50, 137]). For these scenarios, simply applying the methods designed for the low relative dimension setting is ineffective because the theory and algorithms do not fit the hyperplane case. There exist methods, such as K-subspaces [1, 10, 146], that work reasonably well in the high relative dimension regime, while their theoretical support is limited due to the non-convex nature of the objective problem. On the other hand, methods like Algebraic Subspace Clustering (ASC) [126, 127, 129] admit strong theoretical guarantees, but they suffer from an inherent combinatorial complexity that prohibits them from being applied to high-dimensional datasets. Indeed, there is relatively littler work in the literature that directly tackles the high relative dimension regime and provides justifiable theory and convergent algorithms that scale well to the data size.

1.3 Thesis contributions

In this thesis, we develop theory and algorithms for learning subspaces of high relative dimension. In particular, we extend and improve the existing results of Dual Principal Component Pursuit (DPCP) [106, 109, 111, 112, 113, 152, 153], a state-of-the-art non-convex optimization based method primarily designed for learning hyperplanes. To the best of our knowledge, DPCP is the only method that directly focuses on the high relative dimension regime. In the following, we summarize the main contributions of this thesis. We remark that the bulk of this work comes from [30, 31, 32, 151].

1.3.1 Geometric and probabilistic analysis of noisy DPCP

DPCP is originally designed for learning a single hyperplane containing the inliers in the presence of outliers [106, 111, 112]. It is formulated as a non-convex ℓ_1 optimization problem on the sphere, which searches for a basis element of the orthogonal complement of the subspace, i.e., one normal vector to the underlying subspace. The main theoretical advantage of DPCP that distinguishes it from existing robust subspace recovery methods is that it can tolerate as many outliers as the *square* of the number of inliers [152, 153], while other methods can only provably handle a number of outliers on the same order of the number of inliers. However, the analyses of DPCP assume outliers are the only form of corruption, and its behavior is unclear when data is further contaminated by noise as is the case in real data sets.

In Chapter 3, we establish a global optimality theory for noisy DPCP that holds when the inlier points are only assumed to lie close to the underlying subspace $\mathcal{S}$

due to the existence of noise. We provide a geometric analysis that reveals that the global minimizers of the non-convex noisy DPCP problem are perturbed away from the orthogonal complement of the inlier subspace (i.e., $\mathcal{S}^\perp$) by an amount proportional to the noise level, hence generalizing the results of DPCP in the noiseless case. We also give a probabilistic analysis that further interprets the results and shows that the DPCP approach is still able to handle $O((\#\text{inliers})^2)$ outliers even for noisy data. Finally, we show that the global optimality conditions for noisy DPCP are much tighter compared to those required for other closely related state-of-the-art methods.

1.3.2 Extension of DPCP for learning a subspace with codimension larger than one

As already mentioned, the DPCP approach is based on an optimization problem over the sphere that aims at finding a normal vector to a single hyperplane that contains the inliers. When the codimension of the underlying subspace is larger than 1, i.e., not a hyperplane, one could consider computing the subspace as the intersection of many orthogonal hyperplanes learned by DPCP in a recursive fashion. In practice, this approach sequentially finds a new basis element of the space orthogonal to the subspace, which is computationally expensive and lacks any theoretical guarantees.

In Chapter 3, we extend the DPCP approach to the case of learning a subspace $\mathcal{S}$ with codimension larger than 1 by *simultaneously* computing the entire basis of the orthogonal complementary subspace (we call this a holistic approach) by solving a non-convex optimization problem over the Grassmannian [34]. For this new approach, we provide geometric and probabilistic analyses related to global optimality in both noiseless and noisy settings. For noiseless data, under certain conditions, we show that any global solution of the holistic DPCP optimization problem is an orthonormal basis of $\mathcal{S}^\perp$. If the dataset contains noise, we show that the subspace angle between the global solution and $\mathcal{S}^\perp$ is upper bounded by an amount that is proportional to the noise level. In both cases, we derive probabilistic arguments showing that the holistic DPCP approach can tolerate $O((\#\text{inliers})^2)$ outliers, which is superior to other existing methods that can handle at best $O(\#\text{inliers})$ outliers in theory.

1.3.3 Efficient algorithms for subspace learning with DPCP

The existing scalable and provably convergent algorithms for solving DPCP are based on a Projected Sub-Gradient Method (DPCP-PSGM) [152, 153], which enjoy a linear rate of convergence if piecewise geometrically diminishing step sizes are used. Nevertheless, it is only developed for learning a *single* basis element of the orthogonal complement of the underlying subspace $\mathcal{S}$ under the *noiseless* setting. Since in Chapter 3 we extend the original DPCP approach to learn the *entire* basis and prove its effectiveness under *both* noiseless and noisy settings, it is desired to develop a unified algorithmic framework that is able to efficiently solve the DPCP problem for all of these cases.

In Chapter 4, we propose a Projected Riemannian Sub-Gradient Method (PRSGM) for minimizing non-smooth non-convex functions over the Grassmannian. We show that if the objective function satisfies a certain Riemannian regularity condition (RRC) with respect to some point in the Grassmannian, then PRSGM with appropriate initialization and geometrically diminishing step size converges at a linear rate to that point. In particular, we show that the optimization problem associated with the holistic DPCP approach under noiseless setting satisfies the RRC, which allows us to apply the generic result and conclude that the PRSGM converges linearly to a basis for $\mathcal{S}^\perp$. We remark that, even for subspaces of codimension 1 (i.e., hyperplanes), PRSGM improves upon DPCP-PGSM by allowing for a much simpler step size selection strategy and a weaker condition on the initialization. Furthermore, with noisy data we show that the holistic DPCP problem satisfies the RRC in a neighborhood of $\mathcal{S}^\perp$, leading to a linear convergence of PRSGM to a neighborhood of $\mathcal{S}^\perp$ whose radius is proportional to the noise level. Experiments on synthetic data demonstrate the superiority of the holistic DPCP approach implemented by PRSGM relative to the state-of-the-art in learning a single subspace of high relative dimension. An experiment on road plane detection with real 3D data further strengthens the view that DPCP performs favorably against other methods in the high relative dimension regime.

1.3.4 Improved analysis and algorithms of DPCP for learning a union of hyperplanes

Besides the theory and algorithms of DPCP for learning a single subspace, it is known [109, 113] that DPCP can also be applied to the case when data points are drawn from a union of hyperplanes (UoH), for which the DPCP problem admits a unique global minimizer equal to the normal vector of the most dominant hyperplane and thus it proves to be a useful tool in clustering hyperplanes. However, existing analyses of DPCP in the multi-hyperplane case focus on the recovery of the hyperplane with the largest number of points, while lacking a precise characterization of the data distribution and involving quantities that are difficult to interpret. It is natural to ask if one can derive a more transparent analysis that allows for a probabilistic interpretation. Also, the provably convergent algorithm in [109, 113] for solving DPCP based on recursive linear programming is not efficient. It is unclear whether the PRSGM proposed in Chapter 4 can be extended to solve DPCP under a UoH model.

In Chapter 5, we introduce a new notion of geometric dominance for determining the hyperplane that is learned by DPCP under a UoH model, which explicitly captures the distribution of the data and the geometric relationships among the hyperplanes, and derive both geometric and probabilistic conditions under which a global solution to DPCP for a UoH is a normal vector to the geometrically dominant hyperplane. We then prove that the DPCP problem for a UoH satisfies a RRC, and use this result to show that the PRSGM exhibits linear convergence to a normal vector of the geometrically dominant hyperplane. Finally, we integrate DPCP into K-subspaces [1, 10] (DPCP-KSS) by using DPCP to estimate the geometrically dominant hyperplane for each cluster, and leverage an ensemble of DPCP-KSS via the frameworks of K-ensembles [42, 58]. Experiments show that by using DPCP we are able to achieve superior or competitive performance over the state-of-the-art in clustering hyperplanes.

1.4 Notation

We introduce some general notation used throughout this thesis. We let $\mathbb{R}$ denote the set of real numbers, and $\mathbb{R}^D$ denote the D-dimensional linear vector space. We use $\mathbb{S}^{D-1}$ to denote the unit sphere of $\mathbb{R}^D$. Letters that are not bolded denote scalars, such as $x \in \mathbb{R}$ and $K \in \mathbb{R}$, lowercase boldface letters denote vectors, such as $\boldsymbol{x} \in \mathbb{R}^D$, and uppercase (calligraphic) boldface letters denote matrices, such as $\boldsymbol{B} \in \mathbb{R}^{D\times c}$ and $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{D\times N}$. The transpose of a matrix $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{D\times N}$ is denoted as $\boldsymbol{\mathcal{X}}^\top \in \mathbb{R}^{N\times D}$. We also treat a matrix as a set with all of its columns as its elements, i.e., $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$ means $\boldsymbol{x} \in \mathbb{R}^D$ is a column of $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{D\times N}$. Similarly, if $\boldsymbol{\mathcal{O}} \in \mathbb{R}^{D\times M}$, then $\boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{O}}$ consists of the points of $\mathbb{R}^D$ that are common columns of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{O}}$. If $\mathcal{S}$ is a subspace of $\mathbb{R}^D$, then $\dim(\mathcal{S})$ denotes the dimension of $\mathcal{S}$. For a matrix $\boldsymbol{U} \in \mathbb{R}^{D\times c}$, we denote by $\mathrm{Span}(\boldsymbol{U})$ the subspace of $\mathbb{R}^D$ spanned by the columns of $\boldsymbol{U}$. For a subspace $\mathcal{S}$ with $\dim(\mathcal{S}) = d < D$, its orthogonal complement subspace is denoted as $\mathcal{S}^\perp$ with codimension $\dim(\mathcal{S}^\perp) = D - d$. If $\boldsymbol{S} \in \mathbb{R}^{D\times d}$ is the orthonormal basis of $\mathcal{S}$, then we use $\boldsymbol{S}^\perp \in \mathbb{R}^{D\times(D-d)}$ to denote the orthonormal basis of $\mathcal{S}^\perp$. Also, the shorthand RHS (respectively, LHS) stands for *Right-Hand-Side* (respectively, *Left-Hand-Side*). For any real valued convex function $f(\cdot)$, we use $\partial f(\cdot)$ to denote its subdifferential. For any vector $\boldsymbol{x} = [x_1, \cdots, x_D]^\top \in \mathbb{R}^D$ and $p \geq 1$, the ℓ_p norm is defined as $\|\boldsymbol{x}\|_p := \left(\sum_{i=1}^D |x_i|^p\right)^{\frac{1}{p}}$. Unless stated otherwise, we also write $\|\cdot\|$ for the ℓ_2 norm. Additionally, we define $\|\boldsymbol{x}\|_0$ as the number of non-zero entries in $\boldsymbol{x}$. Finally, for any matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ with entries a_{ij}, we define the Frobenius norm as $\|\boldsymbol{A}\|_F := \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2\right)^{\frac{1}{2}}$.

Chapter 2

Dual Principal Component Pursuit (DPCP)

In the context of learning linear structures from corrupted data, to the best of our knowledge *Dual Principal Component Pursuit* (DPCP) [106, 109, 111, 112, 113, 152, 153]—a method designed for robust subspace learning and clustering—is the only method that directly aims at recovering subspaces in the high relative dimension regime (see Section 1.2). All other existing methods assume the underlying structure can be well captured by low-dimensional subspaces, thus making DPCP unique.

In this chapter, we first introduce the existing work of DPCP in Section 2.1. Then in Section 2.2, we briefly review the work closely related to DPCP or highly popular methods for robust subspace learning and clustering. Finally, in Section 2.3 we discuss several open problems of DPCP in terms of its theory and algorithms, which also helps us highlight the main contributions of this thesis.

2.1 Existing work of DPCP

DPCP was originally proposed as a single subspace learning method [111, 112] in the high relative dimension regime where $d/D \approx 1$ with d and D the underlying subspace dimension and ambient dimension, respectively. DPCP is a natural choice for this setting since it learns a subspace by estimating a basis for its orthogonal complement, which is supposed to be of low dimension. Moreover, DPCP has been extended to learn a hyperplane arrangement [109, 113] when the data come from a union of hyperplanes, indicating that it can also be helpful in clustering subspaces of high relative dimension. We review the existing work of DPCP for learning a single subspace with outliers [106, 111, 112, 152, 153] in Section 2.1.1 and learning a union of hyperplanes [106, 109, 113] in Section 2.1.2.

2.1.1 Learning a single subspace with outliers

Consider learning a single subspace from data corrupted with outliers. Suppose we are given the ℓ_2 column-normalized dataset $\widetilde{\boldsymbol{\mathcal{X}}} = [\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{O}}]\boldsymbol{\Gamma} \in \mathbb{R}^{D \times L}$, where $\boldsymbol{\mathcal{X}} = [\boldsymbol{x}_1, \cdots .\boldsymbol{x}_N] \in \mathbb{R}^{D \times N}$ are N inlier points within a d-dimensional subspace $\mathcal{S}$ of $\mathbb{R}^D$ with $1 \leq d \leq D-1$, $\boldsymbol{\mathcal{O}} = [\boldsymbol{o}_1, \cdots, \boldsymbol{o}_M] \in \mathbb{R}^{D \times M}$ are M outlier points that lie on the unit sphere $\mathbb{S}^{D-1}$ in $\mathbb{R}^D$ that do not exhibit linear structure, $L = N + M$ is the total number of points, and $\boldsymbol{\Gamma}$ is an unknown permutation matrix. The goal of DPCP is to recover the underlying subspace $\mathcal{S}$ from the corrupted data $\widetilde{\boldsymbol{\mathcal{X}}}$. Since we might not necessarily know the subspace dimension d in many cases, DPCP resorts to computing a maximal hyperplane of $\mathbb{R}^D$ that contains all the inliers $\boldsymbol{\mathcal{X}}$ as the first step, which can be used to eliminate the vast majority of outliers. Then, one may either utilize popular outlier detection methods such as RANSAC [39] on the reduced dataset for identifying the remaining outliers, or, if d is known, sequentially proceed to compute $\mathcal{S}$ as the intersection of $D-d$ orthogonal hyperplanes that contain $\boldsymbol{\mathcal{X}}$. As the key ingredient, DPCP proposes to search for a maximal hyperplane that contains all the inliers by estimating its normal vector from the following problem:

$$\min_{\boldsymbol{b}\in\mathbb{R}^D} \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_0 \ \text{s.t.}\ \boldsymbol{b}\neq\boldsymbol{0}, \tag{2.1}$$

where $\|\boldsymbol{a}\|_0$ denotes the number of non-zero elements in the vector $\boldsymbol{a}$. Problem (2.1) seeks a normal vector $\boldsymbol{b}$ (thought of as a normal to a hyperplane) that is orthogonal to as many points in $\widetilde{\boldsymbol{\mathcal{X}}}$ as possible. It has been shown in [106] that with mild assumptions such as $N\geq d+1$ and $M\geq D-d$, then every solution $\boldsymbol{b}^*$ to (2.1) is a normal vector of a hyperplane that contains all the inliers $\boldsymbol{\mathcal{X}}$, or equivalently, $\boldsymbol{b}^*$ is orthogonal to $\mathcal{S}$.

Although problem (2.1) is intuitive and theoretically feasible, its combinatorial nature makes it prohibitive in practice. It is reasonable to consider its relaxation [106, 111, 112] that replaces the ℓ_0 function in the objective of (2.1) with an ℓ_1 norm:

$$\min_{\boldsymbol{b}\in\mathbb{R}^D} \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_1 \ \text{s.t.}\ \|\boldsymbol{b}\|_2=1, \tag{2.2}$$

which is refer to as *Dual Principal Component Pursuit* (DPCP). Problem (2.2) is non-smooth and non-convex due to the objective function and the unit sphere constraint. Note that the same problem has appeared before, as early as in [95], and in the context of dictionary recovery [85, 96, 99, 100, 101, 102].

There are two major questions concerning the DPCP problem (2.2): (i) under what conditions is every global minimizer of (2.2) orthogonal to the underlying inlier subspace $\mathcal{S}$; and (ii) how to efficiently compute the global minimum of the non-convex problem (2.2) with theoretical guarantees. In [106, 111, 112], it is shown that if the outliers $\boldsymbol{\mathcal{O}}$ are well-distributed on the unit sphere $\mathbb{S}^{D-1}$ and the inliers $\boldsymbol{\mathcal{X}}$ are well-distributed on $\mathcal{S} \cap \mathbb{S}^{D-1}$, then it is guaranteed that global solutions of (2.2) are orthogonal to $\mathcal{S}$. However, the analysis is deterministic in nature and difficult to interpret. In [152, 153], the deterministic analysis is refined to have interpretable and tighter geometric quantities, and provides a new probabilistic analysis that for the first time shows that the DPCP problem (2.2) can tolerate $M = O(N^2)$ outliers, thus improving upon the existing provably convergent robust PCA methods that can only handle $M = O(N)$ outliers. On the other hand, [106, 111, 112] propose to solve (2.2) through a recursion of convex problems based on linear programs (LPs), which is guaranteed to converge to a vector orthogonal to $\mathcal{S}$ in a finite number of steps. Nevertheless, this approach is computationally expensive. Alternatively, [106, 111, 112] recommend an Iteratively Reweighted Least Squares (IRLS) method [15, 16, 24], which is more efficient than solving a sequence of LPs, but does not have convergence guarantees in this case. To address this dilemma, [152, 153] propose a scalable Projected Sub-Gradient Method with piecewise geometrically diminishing step sizes (DPCP-PSGM) whose main computational cost each iteration is matrix-vector multiplications; this method has a linear convergence rate, thus enhancing its usability.

2.1.2 Learning a union of hyperplanes

Interestingly, although DPCP was originally proposed as a robust single subspace learning method, it is shown in [106, 109, 113] that DPCP can also be used to learn a hyperplane arrangement, which can be attributed to its ability to learn a specific hyperplane from a union of hyperplanes (UoH). Note that the data modeling for a UoH is fundamentally different than a single subspace case: when we treat the data points from one specific hyperplane as inliers, the points from the other hyperplanes play a similar role as "outliers" but exhibit additional linear structure, which we refer to as *structured outliers*[1], making the problem even more challenging.

Consider the ℓ_2 column-normalized dataset $\widetilde{\boldsymbol{\mathcal{X}}} = \boldsymbol{\mathcal{X}}\boldsymbol{\Gamma} \in \mathbb{R}^{D\times N}$, where $\boldsymbol{\mathcal{X}} = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N] \in \mathbb{R}^{D\times N}$ are N inlier points that lie in a union of K hyperplanes $\mathcal{H}_1, \cdots, \mathcal{H}_K$ of $\mathbb{R}^D$ with unit normal vectors $\boldsymbol{n}_1, \cdots, \boldsymbol{n}_K$, respectively, and $\boldsymbol{\Gamma}$ is an unknown permutation matrix. We assume that for every $k \in [K] := \{1, \cdots, K\}$, there are precisely N_k inlier points, denoted by $\boldsymbol{\mathcal{X}}_k \subset \boldsymbol{\mathcal{X}}$, that belong to $\mathcal{H}_k$, so that $\sum_{k=1}^{K} N_k = N$ and we can write $\widetilde{\boldsymbol{\mathcal{X}}} = [\boldsymbol{\mathcal{X}}_1, \cdots, \boldsymbol{\mathcal{X}}_K]\boldsymbol{\Gamma}$. Given this model, the goal of *hyperplane clustering* is to estimate the underlying hyperplanes $\{\mathcal{H}_k\}$ from the data $\widetilde{\boldsymbol{\mathcal{X}}} = \bigcup_{k=1}^{K} \boldsymbol{\mathcal{X}}_k$, as well as cluster the data points according to their membership.

Although such a UoH model is distinct from the one introduced in Section 2.1.1 for learning a single subspace, the DPCP optimization problem of interest to us has the same formulation as (2.2). For the ease of analysis, [106, 109, 113] further assume an

[1] In learning a single hyperplane from data under a UoH model, the *structured* outliers are the data points that come from the remaining hyperplanes; *regular* outliers are uniformly distributed in the ambient space. Throughout this thesis, unless stated otherwise, outliers refer to the regular kind.

ordering $N_1 \geq N_2 \geq \cdots \geq N_K$, and refer to $\mathcal{H}_1$ as *the dominant hyperplane*, namely one with the largest number of points. Then [106, 109, 113] show that as long as $\mathcal{H}_1$ is sufficiently dominant, the data points are well-distributed inside their associated hyperplanes, and the other hyperplanes are sufficiently separated from each other, the normal vector of $\mathcal{H}_1$, i.e., $\boldsymbol{n}_1$, is the unique (up to sign) global minimizer of the DPCP problem (2.2). Algorithmically, [106, 109, 113] recommend solving (2.2) with a standard IRLS method applied to the ℓ_1 minimization problem, as suggested in [111, 112], albeit the convergence analysis is left as future work. Finally, when applied to the task of hyperplane clustering, [106, 109, 113] embed DPCP into a K-subspaces (KSS) [1, 10] scheme that alternates between assigning data points to clusters and estimating a hyperplane for each cluster using DPCP, and show that this strategy works very well in practice.

2.2 Related work

Fitting linear subspaces to data is a fundamental problem in statistical machine learning that has a long history going back more than a century. As a conventional method, Principal Component Analysis (PCA) [13, 52, 55, 84] learns a subspace by minimizing the reconstruction error when projecting the data points to a lower dimensional space, measured by the mean squared distance between the data and their projections [84]. Alternatively, it can be viewed as maximizing the variance of the projected data points [52]. Although PCA enjoys a closed form solution given by the span of the top eigenvectors of the data covariance matrix and it works well even when the data is noisy, it is limited when the dataset is corrupted by outliers since the ℓ_2-based loss in PCA is sensitive to outliers. Another classical approach is Random Sample Consensus (RANSAC) [39], which is very popular in many computer vision applications such as camera calibration [81, 149], metric rectification [66, 134] and so on. Despite its effectiveness in practice, it is sensitive to the settings of many interrelated hyperparameters, and admits limited theoretical guarantees. In the past decade, many robust subspace recovery (RSR) methods have been proposed [62, 119] with the assumption that high-dimensional data can be well-approximated by low-dimensional structures; representatives include robust PCA [9, 11, 75, 120], low-rank matrix methods [88, 136], and approaches based on least absolute deviations [61, 76, 145], which are normally solved using a convex optimization approach. However, their guarantees for theory and algorithms are developed for a low-dimensional underlying structure, which may be violated in the high relative dimension regime. In Section 2.2.1, we will briefly review popular representative methods for robust single subspace learning that are highly related to the rationale behind DPCP.

On the other hand, in many applications the data points are drawn from a union of subspaces instead of a single one, such as motion segmentation [105, 128, 130], hybrid system identification [4, 129] and so on. This is known as the problem of *subspace clustering* [123], which is a general case of the hyperplane clustering introduced in Section 2.1.2 in that the underlying subspaces may have different dimensions and their codimensions are not necessarily equal to one. Most existing subspace clustering methods require the underlying subspaces to be of low relative dimension compared to the ambient space in order to enjoy strong theoretical guarantees together with efficient implementations, which have been heavily researched in the past decade. For example, the self-expressive approaches [35, 37, 70, 72, 124, 141, 143] assume each data point can be expressed as a sparse linear combination of other data points from the same subspace, which is rarely the case in the high relative dimension regime. There are many other categories of subspace clustering methods, including algebraic methods [21, 107, 108, 110, 127], iterative methods [1, 10, 114, 146], statistical methods [48, 89, 103], and spectral clustering-based methods [17, 36, 37, 44, 70, 138, 147]. We provide a brief review of these classes of methods in Section 2.2.2.

2.2.1 Learning a single subspace

RANSAC. Since its inception almost 40 years ago, the Random Sampling And Consensus (RANSAC) [39] algorithm has been one of the most popular methods in computer vision. RANSAC alternates between fitting a subspace from d randomly sampled points (recall that d is the dimension of the underlying subspace) and then using the number of data points close to the subspace as a measure of the quality of the estimation. The interplay between four factors governs when RANSAC is successful: the ambient dimension D, the outlier ratio, the thresholding parameter for determining when points are considered close to a subspace, and the allocated time budget. RANSAC can be extremely effective when the probability of sampling outlier-free samples inside the allocated time budget is large, although its exponential complexity limits its impact in the high relative dimension regime. There are also many derivatives of the standard RANSAC developed in recent years [6, 20, 87].

R1-PCA. Rotational invariant ℓ_1-norm PCA (R1PCA) [28] is a natural extension of PCA that is more robust, whose solution is comprised of the principal eigenvectors of a robust covariance matrix. In particular, it solves the following problem:

$$\min_{\boldsymbol{U} \in R^{D \times d}, \boldsymbol{V} \in \mathbb{R}^{d \times L}} \left\| \widetilde{\boldsymbol{\mathcal{X}}} - \boldsymbol{U}\boldsymbol{V} \right\|_{2,1} = \sum_{j=1}^{L} \|\widetilde{\boldsymbol{x}}_j - \boldsymbol{U}\boldsymbol{v}_j\| \text{ s.t. } \boldsymbol{U}^\top \boldsymbol{U} = \mathbf{I} \tag{2.3}$$

where $\widetilde{\boldsymbol{\mathcal{X}}} \in \mathbb{R}^{D \times L}$ is the data matrix, $\widetilde{\boldsymbol{x}}_j$ is the j-th column of $\widetilde{\boldsymbol{\mathcal{X}}}$, $\boldsymbol{U}$ is the orthonormal basis of the estimated underlying subspace, $\boldsymbol{v}_j$ is the j-th column of $\boldsymbol{V}$, and $\boldsymbol{V}$ is the representation matrix whose columns correspond to the coordinates of the data points represented by $\boldsymbol{U}$. The original R1PCA approach proposes to solve problem (2.3) via alternating minimization that involves some form of the power method [45]. However, it lacks both a theoretical guarantee for subspace recovery and any convergence guarantee to the global optimal solution for the non-convex problem (2.3).

CoP. Coherence Pursuit (CoP) [88] is a non-iterative robust PCA method for recovering a low-dimensional subspace that assumes that the inlier points are likely to have stronger mutual coherence with a large number of inliers compared with the unstructured outliers. It measures the mutual coherences according to the column magnitudes of a gram matrix formed from the dataset, and computes the subspace as the span of the d data points with largest coherence. CoP is fast due to its non-iterative nature, especially when the dataset is small. Although CoP can provably handle outliers and additive noise, it can only tolerate $M = O(N)$ outliers and requires $d < \sqrt{D}$ in theory, making it not well-suited for the high relative dimension regime.

REAPER. Similar to DPCP, the REAPER method [61] computes the subspace by aiming to minimize the sum of the distances between all points in the dataset and the subspace. Specifically, it tackles the following problem:

$$\min_{\mathbf{\Pi} \in \mathbb{R}^{D \times D}} \left\| \widetilde{\boldsymbol{\mathcal{X}}}^{\top} (\mathbf{I} - \mathbf{\Pi}) \right\|_{1,2} = \sum_{j=1}^{L} \| (\mathbf{I} - \mathbf{\Pi}) \widetilde{\boldsymbol{x}}_j \|_2 \ \text{ s.t. } \mathbf{\Pi} \text{ is an orthoprojector} \quad \text{and } \operatorname{trace}(\mathbf{\Pi}) = d, \tag{2.4}$$

where $\mathbf{\Pi}$ can be thought of the orthoprojector that projects data to the d-dimensional inlier subspace $\mathcal{S}$. Since problem (2.4) is non-convex due to the orthoprojectors do not form a convex set, [61] turns to solve a tight convex relaxation that robustly estimates the orthoprojector onto $\mathcal{S}$, which is referred to as REAPER:

$$\min_{\boldsymbol{P} \in \mathbb{R}^{D \times D}} \left\| \widetilde{\boldsymbol{\mathcal{X}}}^{\top} (\mathbf{I} - \boldsymbol{P}) \right\|_{1,2} \ \text{ s.t. } \mathbf{0} \preccurlyeq \boldsymbol{P} \preccurlyeq \mathbf{I} \text{ and } \operatorname{trace}(\boldsymbol{P}) = d, \tag{2.5}$$

and the underlying subspace $\mathcal{S}$ is then computed as the top d eigenvectors of $\boldsymbol{P}^*$ with $\boldsymbol{P}^*$ the global solution of problem (2.5). [61] establishes the theory for recovering $\mathcal{S}$ from (2.5) under both noiseless and noisy settings. Nevertheless, its theoretical guarantees require $d < (D-1)/2$, thus excluding the high relative dimensional setting, and can still handle only $M = O(N)$ outliers. Algorithmically, since the original semi-definite program (2.5) may be prohibitively expensive to solve, [61] proposes to solve it via an IRLS scheme [144, 145] with a guarantee that the iterates converge to a point whose value is close to the optimal objective value of (2.5), but it does not provide the rate of convergence nor how the iterates relate to the recovery of $\mathcal{S}$.

GGD. Recently, [76] improves upon REAPER with a Geodesic Gradient Descent (GGD) method for solving the non-convex least absolute deviations problem without any relaxation. The underlying optimization problem it considers is

$$\min_{\boldsymbol{V} \in \mathbb{R}^{D \times d}} \left\| \widetilde{\boldsymbol{\mathcal{X}}}^\top (\mathbf{I} - \boldsymbol{V}\boldsymbol{V}^\top) \right\|_{1,2} \text{ s.t. } \boldsymbol{V}^\top \boldsymbol{V} = \mathbf{I}. \tag{2.6}$$

Ideally, the global solution $\boldsymbol{V}^*$ to problem (2.6) consists of an orthonormal basis for the underlying subspace $\mathcal{S}$. Note that (2.6) is an optimization problem on the Grassmannian $\mathbb{G}(D, d)$ [34], i.e., the set of d-dimensional subspaces in $\mathbb{R}^D$. [76] provides conditions under which any orthonormal basis of $\mathcal{S}$ is a local minimizer of (2.6) for both noiseless and noisy settings. Additionally, an intrinsic GGD algorithm, for which the iterates move along a geodesic in $\mathbb{G}(D, d)$, is proposed to solve (2.6) with a guarantee of linear convergence to the local minimizer, if properly initialized. One advantage of GGD with respect to CoP and REAPER, is that its theoretical analysis does not have restrictions on the inlier dimension d, hence it can be used in the high relative dimension regime in theory. On the other hand, like CoP and REAPER, GGD can only provably handle $M = O(N)$ outliers. Moreover, [76] only provides a local optimality analysis that characterizes the geometry of the critical points of (2.6), while a global optimality condition for (2.6) remains an open question.

2.2.2 Clustering multiple subspaces

RANSAC. As a classical method, RANSAC not only can robustly learn a single subspace in the presence of unstructured outliers, but it can also cluster data points according to their memberships when they are drawn from multiple subspaces. Heuristically, it fits one subspace at a time using PCA from d randomly sampled points, in which it treats the points from other subspaces as structured outliers; this process is repeated after the points identified as belonging to the previously selected subspaces are removed. However, as in the single subspace learning case introduced in Section 2.2.1, its performance is highly sensitive to various factors, e.g., the thresholding parameter, and it suffers from an exponential complexity as the number of subspaces grows since the probability of drawing exactly d inlier points from a subspace drops exponentially with the number of subspaces.

***K*-subspaces.** K-subspaces (KSS) [1, 10, 114] is a simple but effective method for subspace clustering, which alternates between assigning data points to clusters and estimating a subspace for each cluster using PCA. KSS is scalable in practice, but it can easily get stuck near a local minimum due to its non-convex nature, and it is not robust to outliers. The suboptimality issue can be alleviated by running the method multiple times with diverse initializations and then selecting the best, or leveraging ensembles of multiple KSS results [58, 67]. The lack of robustness stems from the fact that the squared ℓ_2 loss used in PCA is incapable of handling outliers during the subspace estimation step of KSS where most of the data points in any cluster come from one underlying subspace (serve as inliers) and the rest are points from the union of other $K-1$ subspaces (serve as structured outliers). In order to improve its robustness, Median K-Flats (MKF) [146] replaces the squared ℓ_2 objective in KSS with an unsquared one, but it lacks competitive performance as observed by [42].

Alternatively, [42] proposes to substitute CoP [88] for the PCA step in KSS, but CoP is only able to deal with low-rank structured outliers, as introduced in Section 2.2.1.

Self expressive methods. Self expressive methods belong to one of the most effective approaches for clustering low-dimensional subspaces. The fundamental idea is that a point from one subspace with dimension d can always be expressed as a linear combination of d linear independent points from the same subspace. This means, if we consider the outlier-free noise-free case and a data matrix $\widetilde{\boldsymbol{\mathcal{X}}} \in \mathbb{R}^{D\times N}$, that $\widetilde{\boldsymbol{x}}_j = \widetilde{\boldsymbol{\mathcal{X}}}\boldsymbol{c}_j$, where $\widetilde{\boldsymbol{x}}_j$ is a point (column) of $\widetilde{\boldsymbol{\mathcal{X}}}$ and $\boldsymbol{c}_j \in \mathbb{R}^{N\times 1}$ is its coefficient representation in terms of the other $(N-1)$ data points in $\widetilde{\boldsymbol{\mathcal{X}}}$. Normally, we have $N \gg d$, and thus $\boldsymbol{c}_j$ is presumably a sparse vector. In matrix form, we can write $\widetilde{\boldsymbol{\mathcal{X}}} = \widetilde{\boldsymbol{\mathcal{X}}}\boldsymbol{C}$ where $\boldsymbol{C} \in \mathbb{R}^{N\times N}$ is a sparse coefficient matrix with $\mathrm{diag}(\boldsymbol{C}) = \boldsymbol{0}$, and the self-expressive methods seek to solve a convex optimization problem of the form

$$\min_{\boldsymbol{C}\in\mathbb{R}^{N\times N}} \lambda \left\|\widetilde{\boldsymbol{\mathcal{X}}} - \widetilde{\boldsymbol{\mathcal{X}}}\boldsymbol{C}\right\|_F^2 + \Phi(\boldsymbol{C}) \tag{2.7}$$

where $\lambda > 0$ is the coefficient parameter, $\Phi(\cdot)$ is a regularization function, and different choices of $\Phi(\cdot)$ result in different categories of methods: sparse subspace clustering (SSC) [35, 36, 37] uses $\Phi(\cdot) = \|\cdot\|_1$, least-squares regression (LSR) [72] uses $\Phi(\cdot) = \|\cdot\|_2$, low-rank subspace clustering [38, 69, 70, 124, 133, 142] uses $\Phi(\cdot) = \|\cdot\|_*$, and elastic net subspace clustering (EnSC) [54, 83, 141, 143] uses $\Phi(\cdot)$. Given a solution $\boldsymbol{C}^*$ to problem (2.7), a pairwise affinity matrix $\boldsymbol{A}$ is built by $\boldsymbol{A} = |\boldsymbol{C}^*| + |\boldsymbol{C}^{*\top}|$, and finally a spectral clustering technique [131] is applied to obtain the segmentation. With mild modifications, (2.7) can be extended to the dataset contaminated with outliers and noise. However, the construction of the $N \times N$ coefficient matrix $\boldsymbol{C}^*$ is expensive with large-scale data, and the theoretical guarantees for the self-expressive methods require the underlying subspaces to be low-dimensional, preventing its impact in clustering high-dimensional subspaces.

Spectral Curvature Clustering. The spectral curvature clustering (SCC) method [17] is a multi-way spectral clustering technique [46], which is well-suited for clustering affine subspaces with the same dimension d. In particular, given the data matrix $\widetilde{\boldsymbol{\mathcal{X}}} \in \mathbb{R}^{D \times L}$, it constructs a multi-way affinity $\mathcal{A}(i_1, \cdots, i_{d+2})$ for any $d+2$ points in $\widetilde{\boldsymbol{\mathcal{X}}}$ based on a certain polar curvature, which is zero when points are in the same subspace. The $(d+2)$-way tensor $\mathcal{A}$ of size $L \times L \times \cdots \times L$ is then unfolded to build an affinity matrix $\boldsymbol{A} \in \mathbb{R}^{L \times L^{d+1}}$, which is then followed by the use of standard spectral clustering. Considering the storage and the expense of computing $\boldsymbol{A}$, [17] proposes an iterative sampling procedure to significantly improve the performance. Nevertheless, the combinatorial nature of SCC prohibits its application in clustering high-dimensional subspaces in practice.

Algebraic Subspace Clustering. Algebraic subspace clustering (ASC) [74, 107, 108, 110, 127, 129] is a class of purely algebraic algorithms designed for subspace clustering. The main idea is that a union of K subspaces can be associated with a set of polynomials of degree K whose derivatives at an inlier point are orthogonal to the subspace that the point lies in; the clustering is based on the grouping of these normal vectors. More formally, suppose data are drawn from a union of K subspaces, i.e., $\bigcup_{k=1}^{K} \mathcal{S}_k$, with $\boldsymbol{b}_k$ a normal to $\mathcal{S}_k$, then one can represent the data with polynomials of degree K of the form $p(\boldsymbol{x}) = (\boldsymbol{b}_1^\top \boldsymbol{x}) \cdots (\boldsymbol{b}_K^\top \boldsymbol{x}) = 0$, and the coefficients of the polynomials can be computed by solving a linear system. ASC can also be extended to handle noisy data by adding an additional treatment of the involved linear systems [82, 139]. Although ASC enjoys strong theoretical guarantees, it is sensitive to outliers and suffers from the combinatorial computational cost in aspects of the number of underlying subspaces and the ambient dimension.

2.3 Open problems

Despite the advances made by DPCP in robustly learning and clustering high-dimensional subspaces, there are still many open problems related to DPCP in terms of both theory and algorithms. We will discuss them in this section and provide our solutions in the rest of the remainder of the thesis.

2.3.1 Single subspace learning theory with DPCP

2.3.1.1 DPCP in the presence of noisy inliers

As introduced in Section 2.1.1, DPCP uses a non-convex optimization problem for learning subspaces of high relative dimension from *noiseless* datasets contaminated by as many outliers as the square of the number of inliers [152, 153]. Although the theoretical features of DPCP are appealing, they have only been established for the idealized case when inliers perfectly lie in the subspace. Experimentally, DPCP has proved to be robust to noise and outperform the popular RANSAC algorithm on

3D vision tasks such as road plane detection and relative pose estimation from three views [29]. Therefore, it is reasonable to ask whether similar theoretical guarantees hold when there is noise in the data.

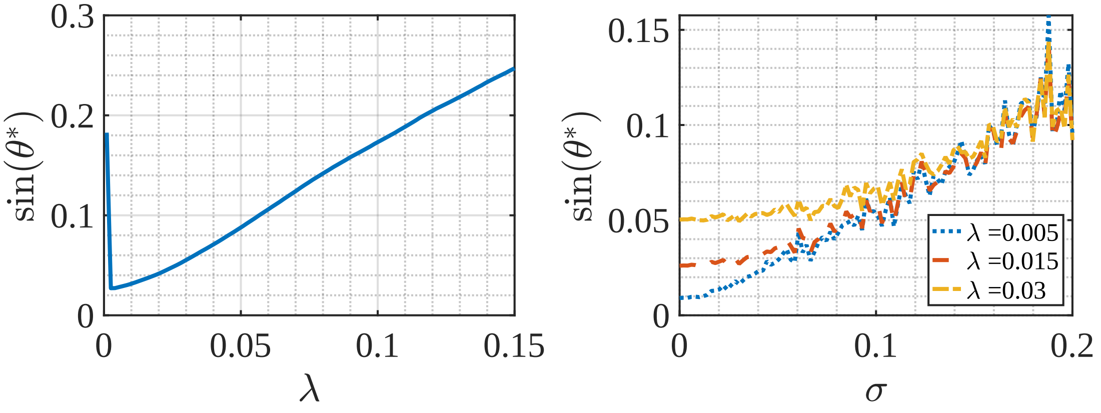
A more realistic data modeling strategy is to consider the corruption of inliers by noise. If $\boldsymbol{\mathcal{E}} = [\boldsymbol{\epsilon}_1, \cdots, \boldsymbol{\epsilon}_N] \in \mathbb{R}^{D\times N}$ denotes the additive noise on inliers $\boldsymbol{\mathcal{X}}$, then the data matrix now has the form $\widetilde{\boldsymbol{\mathcal{X}}} = [\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}}\ \boldsymbol{\mathcal{O}}]\boldsymbol{\Gamma} \in \mathbb{R}^{D\times L}$, and the goal of DPCP is to estimate the underlying subspace $\mathcal{S}$ from noisy data $\widetilde{\boldsymbol{\mathcal{X}}}$. Recall that when there is no noise (i.e., $\boldsymbol{\mathcal{E}} = \mathbf{0}$), the vectors $\boldsymbol{b}$ that make $\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}$ as sparse as possible are precisely those satisfying $\boldsymbol{b} \perp \mathcal{S}$; this is the motivation for problem (2.2). As an analogy, in the noisy case, we expect $\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}$ to be close to a sparse vector $\boldsymbol{y}$ in the Euclidean sense, whenever $\boldsymbol{b}$ is close to a normal vector of $\mathcal{S}$. This motivates [112] to consider the following denoised version of the DPCP problem[2] in (2.2):

$$\min_{\boldsymbol{b}\in\mathbb{S}^{D-1}, \boldsymbol{y}\in\mathbb{R}^L} \lambda \left\|\boldsymbol{y}\right\|_1 + \frac{1}{2}\left\|\boldsymbol{y} - \widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_2^2 \tag{2.8}$$

for some $\lambda > 0$. However, the performance of (2.8) depends crucially on the parameter λ, as illustrated in Figure 2.1, where we solve (2.8) by alternating minimization, which empirically converges fast even though no convergence theory is known. Figure 2.1a shows that the regularization parameter λ should be chosen very carefully in order to achieve an optimal result, and Figure 2.1b shows this when the noise level varies.

Comparing the denoised DPCP problem (2.8) with its original formulation (2.2), a natural question to ask is that whether (2.2) can also be extended to the noisy case

[2] Problem (2.8) has also appeared in the context of dictionary learning; see [86].



(a) Sensitivity to λ when $\sigma = 0.05$ **(b)** Performance of (2.8) when varying σ

Figure 2.1. Illustration of the performance for the denoised DPCP problem. We generate the data according to a certain random spherical model. In particular, $\sigma > 0$ denotes the standard deviation of the Gaussian noise added to the inliers, and we evaluate the performance of the denoised DPCP problem (2.8) by computing the principal angle θ^* of its solution to $\mathcal{S}^{\perp}$. Here, we fix the ambient dimension as $D = 30$, the subspace dimension as $d = 29$, the number of inliers as $N = 500$, and the outlier ratio as $M/(M + N) = 0.7$. (a) Sensitivity to different choices of λ for fixed noise level. (b) Performance of the denoised DPCP problem (2.8) for a large range of noise levels with specific choices of λ.

such that we can get rid of choosing the extra hyperparameter λ in (2.8). Moreover, it is unclear what kind of theoretical guarantees we can obtain. Although we expect that the global solution of the noisy DPCP problem will be perturbed away from $\mathcal{S}^{\perp}$ with an amount bounded by an increasing function of the noise level, a precise characterization of such relationship is of interest.

2.3.1.2 DPCP for learning a subspace with codimension larger than 1

In addition to the drawback that the existing analyses of DPCP for learning a single subspace are restricted to the case of no noise, another drawback is that the current analyses mainly focus on finding a normal to a *single* hyperplane that contains the inliers by solving (2.2). Extending these ideas to the recovery of a *subspace* with codimension $c = D - d > 1$ requires the recursive application of (2.2) c times, with each time finding a normal vector to $\mathcal{S}$ that is also orthogonal to previously computed normal vectors. This procedure is computationally expensive and lacks a convergence analysis. Moreover, the error accumulated during the recursion makes its behavior difficult to analyze in theory.

A reasonable extension to the current formulation (2.2) of DPCP that learns an element of a basis of the orthogonal complement subspace $\mathcal{S}^\perp$ is that we *simultaneously* estimate the entire basis of $\mathcal{S}^\perp$ by solving the problem

$$\min_{\boldsymbol{B} \in \mathbb{R}^{D \times c}} \left\| \widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B} \right\|_{1,2} = \sum_{j=1}^{L} \left\| \widetilde{\boldsymbol{x}}_j^\top \boldsymbol{B} \right\|_2 \text{ s.t. } \boldsymbol{B}^\top \boldsymbol{B} = \mathbf{I}. \tag{2.9}$$

We call problem (2.9) a *holistic approach* as compared with the recursive approach with problem (2.2). Note that (2.9) extends (2.2) in that it seeks a matrix $\boldsymbol{B}$ with orthonormal columns that are orthogonal to as many data points as possible. We remark that (2.9) has a close relationship to the formulation (2.6) of GGD [76]: the former considers recovering the orthogonal complement subspace $\mathcal{S}^\perp$ while the latter focuses on estimating the actual subspace $\mathcal{S}$. The fundamental reason for this difference is that DPCP aims at the high relative dimension regime where $d/D \approx 1$, thus making it more efficient to operate on the dual space. Similar to (2.6), (2.9) is an optimization problem on the Grassmannian $\mathbb{G}(D, c)$ [34], i.e., the set of c-dimensional subspaces in $\mathbb{R}^D$, and is inherently non-convex. An open question that we answer in this thesis is to establish conditions under which every global solution $\boldsymbol{B}^*$ of (2.9) is an orthonormal basis of $\mathcal{S}^\perp$ when no noise is present, and how the principal angles between $\mathrm{Span}(\boldsymbol{B}^*)$

and $\mathcal{S}^\perp$ behave as a function of the noise level when the data is noisy.

2.3.2 Efficient algorithms for learning a single subspace with DPCP

When DPCP was proposed in [111, 112], the core nonconvex optimization problem was solved by solving a recursion of convex problems based on linear programs (LPs), and convergence guarantees were established. Nevertheless, the LP-based approach lacks scalability for big-data applications. To alleviate the issue, [111, 112] recommend solving (2.2) with an IRLS method that is more efficient but lacks a theoretical convergence guarantee. Fortunately, [152, 153] take one large step forward by utilizing a scalable Projected Sub-Gradient Method (DPCP-PSGM), which is proven to have a linear convergence rate for solving the non-convex problem (2.2) and is orders of magnitude faster than the LP-based method and IRLS scheme.

Two major limitations of the above algorithms are that, in accordance with Section 2.3.1, *none* of them can provably handle the DPCP problem in the noisy case or can be extended from codimension one to higher codimensions. On the one hand, although DPCP-PSGM works well for road plane detection from 3D point cloud data using the KITTI dataset [40], which is real (and hence noisy) data, it is unknown whether it provably converges to a neighborhood of $\mathcal{S}^\perp$ in the presence of noise. On the other hand, we are in need of designing an algorithm for solving the holistic DPCP problem (2.9) that efficiently finds the entire orthogonal basis directly on $\mathbb{G}(D, c)$, as opposed to the less efficient approach of solving a sequence of c problems on $\mathbb{G}(D, 1)$.

In fact, noticing that optimization problems on the Grassmannian $\mathbb{G}(D, c)$ commonly appear in a wide variety of applications, not only including robust subspace recovery or clustering, but also dictionary learning [86, 100, 101, 102], subspace tracking [5], system identification [116], action recognition [93], object categorization [49], and blind deconvolution [148], it would be even more interesting to develop a generic optimization technique over the Grassmannian with a particular application to the DPCP problem. However, a key challenge is that the Grassmannian is a non-convex set, making the associated results difficult to be established in terms of the theoretical guarantees or the rate of convergence. We face this challenge in this thesis.

2.3.3 Learning a union of hyperplanes with DPCP

One nice thing about DPCP is that problem (2.2) can not only handle regular outliers from the ambient space such as those appearing in the robust single hyperplane learning case (Section 2.1.1) but also structured outliers coming from other hyperplanes when the data points are drawn from a union of hyperplanes (Section 2.1.2) [109, 113]. It is not known, however, whether DPCP can learn a normal to one of the hyperplanes in the presence of *both* structured and regular outliers. In particular, [109, 113] define the notion of a dominant hyperplane that depends only on the number of inlier points in each group, while the global optimum also depends on geometric quantities related to their distribution. In other words, its global optimality analysis lacks a precise characterization of the distribution of the data. Moreover, the analysis in [109, 113] is deterministic in nature and involves quantities that are difficult to interpret. It is desirable to leverage more transparent geometric quantities such as those introduced in [152, 153] to derive a probabilistic analysis for the DPCP problem under a UoH model. Finally, similar to the dilemma in the single subspace learning case, there does not exist a scalable algorithm that ensures global convergence for learning a single hyperplane for a UoH. Even more interestingly, provided a generic optimization algorithm over the Grassmannian is developed (as discussed in Section 2.3.2) that can be applied to DPCP for data drawn from a single subspace with outliers, is it possible to extend the algorithm to a UoH setting while enjoy similar convergence properties. Although the above challenges associated with DPCP under a UoH model do not have clear solutions, we will address all of them in this thesis.

Chapter 3

Single Subspace Learning Theory with DPCP

In this chapter, we establish the theory for learning a single subspace with DPCP. In particular, we provide geometric and probabilistic analyses for learning a subspace of any codimension under both noiseless and noisy settings, which largely extends the existing analysis of DPCP that has only been derived for learning a hyperplane with noiseless data. In Section 3.1, we present the noisy analysis of the DPCP problem (2.2) for learning a hyperplane. Then, in Section 3.2 we extend the method to simultaneously learn the entire basis of the orthogonal complement subspace by solving the holistic DPCP problem (2.9). Comparison with the theory of other closely related methods is given in Section 3.3.

3.1 Learning a hyperplane (codimension equal to one)

3.1.1 Review of the existing analysis with noiseless data

We now briefly review the existing analysis [152, 153] of the DPCP problem (2.2) for learning a hyperplane that contains noiseless inliers. Specifically, we will introduce some useful geometric quantities that are also leveraged in our subsequent analysis. For convenience, we repeat the optimization problem (2.2) here:

$$\min_{\boldsymbol{b}\in\mathbb{R}^D} \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_1 \ \text{s.t.}\ \|\boldsymbol{b}\|_2 = 1. \tag{3.1}$$

Here $\widetilde{\boldsymbol{\mathcal{X}}} = [\boldsymbol{\mathcal{X}},\ \boldsymbol{\mathcal{O}}]\boldsymbol{\Gamma} \in \mathbb{R}^{D\times L}$ is a (column-wise) unit ℓ_2 norm dataset, where $\boldsymbol{\mathcal{X}} = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N] \in \mathbb{R}^{D\times N}$ are N inlier points spanning a single d-dimensional subspace $\mathcal{S}$ of $\mathbb{R}^D$, $\boldsymbol{\mathcal{O}} = [\boldsymbol{o}_1, \cdots, \boldsymbol{o}_M] \in \mathbb{R}^{D\times M}$ are M outlier points, and $\boldsymbol{\Gamma}$ is an unknown permutation matrix.

Since the objective in (3.1) is not continuously differentiable, we need to deal with its subdifferential. Denote the sign function by

$$\operatorname{sign}(a) = \begin{cases} a/|a|, & a \neq 0, \\ 0, & a = 0, \end{cases} \tag{3.2}$$

and denote the subdifferential of the absolute value function $|a|$ by

$$\operatorname{Sgn}(a) = \begin{cases} \operatorname{sign}(a), & a \neq 0, \\ [-1, 1], & a = 0. \end{cases} \tag{3.3}$$

We also apply sign and Sgn element-wise to vectors. With this notation, [152, 153] first characterize the distribution of the inliers by

$$c_{\boldsymbol{\mathcal{X}},\min} := \frac{1}{N} \min_{\boldsymbol{b} \in \mathcal{S} \cap \mathbb{S}^{D-1}} \left\| \boldsymbol{\mathcal{X}}^\top \boldsymbol{b} \right\|_1 . \tag{3.4}$$

Note that $c_{\boldsymbol{\mathcal{X}},\min}$ is also the *permeance statistic* defined in [61]. Well-distributed inliers $\boldsymbol{\mathcal{X}}$ leads to a relatively large value of $c_{\boldsymbol{\mathcal{X}},\min}$ since it is difficult to find a single direction $\boldsymbol{b}$ that is orthogonal to many points in $\boldsymbol{\mathcal{X}}$. Next, to characterize the distribution of the outliers, similar to $c_{\boldsymbol{\mathcal{X}},\min}$, [152, 153] define quantities

$$\begin{aligned} c_{\boldsymbol{\mathcal{O}},\min} &:= \frac{1}{M} \min_{\boldsymbol{b} \in \mathbb{S}^{D-1}} \left\| \boldsymbol{\mathcal{O}}^\top \boldsymbol{b} \right\|_1 \quad \text{and} \\ c_{\boldsymbol{\mathcal{O}},\max} &:= \frac{1}{M} \max_{\boldsymbol{b} \in \mathbb{S}^{D-1}} \left\| \boldsymbol{\mathcal{O}}^\top \boldsymbol{b} \right\|_1 . \end{aligned} \tag{3.5}$$

For well-distributed outliers $\boldsymbol{\mathcal{O}}$, the permeance statistics $c_{\boldsymbol{\mathcal{O}},\min}$ and $c_{\boldsymbol{\mathcal{O}},\max}$ are bounded away from small and large values, respectively, since there is not a single direction that sufficiently captures the distribution of $\boldsymbol{\mathcal{O}}$. Moreover, $c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min} \to 0$ as $M \to \infty$ for well-distributed outliers [152, Lemma 4]. Finally, besides $c_{\boldsymbol{\mathcal{O}},\min}$ and $c_{\boldsymbol{\mathcal{O}},\max}$, [152, 153] additionally characterize the distribution of the outliers using the quantity

$$\eta_{\boldsymbol{\mathcal{O}}} := \frac{1}{M} \max_{\boldsymbol{b} \in \mathbb{S}^{D-1}} \left\| (\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top) \boldsymbol{\mathcal{O}} \operatorname{sign}(\boldsymbol{\mathcal{O}}^\top \boldsymbol{b}) \right\|_2, \tag{3.6}$$

which can be viewed as the maximum norm Riemannian subgradient of the function $\frac{1}{M} \left\| \boldsymbol{\mathcal{O}}^\top \boldsymbol{b} \right\|_1$. More uniformly distributed outliers lead to smaller values of $\eta_{\boldsymbol{\mathcal{O}}}$. This follows since if $M \to \infty$ and $\boldsymbol{\mathcal{O}}$ is well distributed, then $\frac{1}{M}\boldsymbol{\mathcal{O}} \operatorname{sign}(\boldsymbol{\mathcal{O}}^\top \boldsymbol{b})$ approaches the direction of $\boldsymbol{b}$, which leads to $\eta_{\boldsymbol{\mathcal{O}}} \to 0$ [152, 153].

With the above geometric quantities, the following lemma characterizes the geometry of critical points of the DPCP problem (3.1).

Lemma 1. ([152, Lemma 1])**.** *Any critical point $\boldsymbol{b}$ of problem* (3.1) *must either be a normal vector of $\mathcal{S}$, or have a principal angle θ from $\mathcal{S}^\perp$ larger than or equal to* $\arccos\left(M\overline{\eta}_{\boldsymbol{\mathcal{O}}}/Nc_{\boldsymbol{\mathcal{X}},\min}\right)$*, where*

$$\overline{\eta}_{\boldsymbol{\mathcal{O}}} := \eta_{\boldsymbol{\mathcal{O}}} + \frac{D}{M}. \tag{3.7}$$

In other words, Lemma 1 indicates that any critical point of (3.1) is either orthogonal to the inlier subspace $\mathcal{S}$, or is close to $\mathcal{S}$, with its principal angle θ from $\mathcal{S}^\perp$ being larger for well-distributed data points and smaller M/N. The following theorem provides global optimality conditions for (3.1), under which any global minimizer of (3.1) must be a normal vector to $\mathcal{S}$.

Theorem 1. ([152, Theorem 1])**.** *Any global solution $\boldsymbol{b}^*$ to problem* (3.1) *must be* *orthogonal to the inlier subspace $\mathcal{S}$ as long as*

$$\frac{M}{N} \cdot \frac{\sqrt{\overline{\eta}_{\boldsymbol{\mathcal{O}}}^2 + (c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min})^2}}{c_{\boldsymbol{\mathcal{X}},\min}} < 1. \tag{3.8}$$

One can see that if the data points are well-distributed and we have more and more inliers and outliers while keeping M/N fixed, then condition (3.8) is more likely to be satisfied so that any global solution to (3.1) must be orthogonal to $\mathcal{S}$. In order to better interpret the result in Theorem 1, [152, 153] provide a probabilistic analysis that characterizes the number of outliers that the DPCP problem (3.1) can tolerate. Towards that end, they derive concentration bounds for the associated geometric quantities under a random spherical model as we present next.

Lemma 2. ([152, Lemma 4])**.** *Consider a random spherical model where the columns of $\boldsymbol{\mathcal{O}}$ and $\boldsymbol{\mathcal{X}}$ are drawn independently and uniformly at random from the unit sphere $\mathbb{S}^{D-1}$ and the intersection of the unit sphere and a subspace $\mathcal{S}$ of dimension $d < D$, respectively. Fix a number $t > 0$, then*

$$\begin{aligned} &\mathbb{P}\left[c_{\boldsymbol{\mathcal{X}},\min} \geq \sqrt{\frac{2}{\pi d}} - \left(2 + \frac{t}{2}\right)/\sqrt{N}\right] \geq 1 - 2e^{-\frac{t^2}{2}}, \\ &\mathbb{P}\left[\eta_{\boldsymbol{\mathcal{O}}} \leq C_0\left(\sqrt{D}\log D + t\right)/\sqrt{M}\right] \geq 1 - 2e^{-\frac{t^2}{2}}, \\ &\mathbb{P}\left[c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min} \leq (4+t)/\sqrt{M}\right] \geq 1 - 2e^{-\frac{t^2}{2}}, \end{aligned} \tag{3.9}$$

where C_0 is a universal constant that is independent of N, M, D, d and t.

We remark that the concentration bounds in (3.9) give us a better understanding of those geometric quantities. For example, it shows that $c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min} \to 0$ as

$M \to \infty$ for well-distributed outliers. Furthermore, it tells us that $Nc_{\boldsymbol{\mathcal{X}},\min}$ scales as $O(N)$ while $M\eta_{\boldsymbol{\mathcal{O}}}$ only scales as $O(\sqrt{M})$ with high probability. As we substitute the geometric quantities in (3.8) with their concentration bounds, it leads to the following probabilistic theorem.

Theorem 2. ([152, Theorem 2]). *Consider the random spherical model described in Lemma 2. Then for any positive scalar $t < 2\left(\sqrt{\frac{2N}{\pi d}} - 2\right)$, with probability at least $1 - 6e^{-\frac{t^2}{2}}$, any global solution of* (3.1) *is orthogonal to $\mathcal{S}$ as long as*

$$(4+t)^2 M + C_0\left(\sqrt{D}\log D + t\right)^2 M \leq \left(\sqrt{\frac{2}{\pi d}}N - \left(2 + \frac{t}{2}\right)\sqrt{N}\right)^2,$$

where C_0 is a universal constant that is independent of N, M, D, d and t.

Theorem 2 suggests that the DPCP problem (3.1) can tolerate $M = O\left(\frac{1}{dD\log^2 D}N^2\right)$ outliers, and in particular can tolerate $M = O(N^2)$ for fixed D and d, which is in sharp contrast with many existing robust PCA methods (see [62] for an overview) that can only handle $M = O(N)$ outliers in theory. They attribute this advantageous theoretical property of DPCP to the tighter geometric quantities used for the analysis. Specifically, as shown in (3.8), the scalings of $M\eta_{\boldsymbol{\mathcal{O}}}$ as $O(\sqrt{M})$ and $Nc_{\boldsymbol{\mathcal{X}},\min}$ as $O(N)$ make it possible to tolerate as many outliers as the square of the number of inliers.

3.1.2 Analysis with noisy data

Although the theoretical features of the DPCP problem (3.1) developed in [152, 153] are appealing, they have only been established for the idealized case when the inliers perfectly lie in the subspace. Yet, DPCP has proved to be competitive on noisy real datasets, so that it is reasonable to ask whether similar theoretical guarantees hold when there is noise in the data. As the first contribution of the thesis, we bridge that gap by extending the analysis of (3.1) to the noisy setting.

Based on the discussion in Section 2.3.1, we consider the same formulation of the DPCP problem as (3.1) but with noisy data. Let us repeat the problem here:

$$\min_{\boldsymbol{b}\in\mathbb{R}^D} \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_1 \text{ s.t. } \|\boldsymbol{b}\|_2 = 1,$$

where $\widetilde{\boldsymbol{\mathcal{X}}} = [\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}}, \boldsymbol{\mathcal{O}}]\boldsymbol{\Gamma} \in \mathbb{R}^{D\times L}$ is a unit ℓ_2 norm dataset that contains noisy inliers, namely $\boldsymbol{\mathcal{E}} = [\boldsymbol{\epsilon}_1, \cdots, \boldsymbol{\epsilon}_N] \in \mathbb{R}^{D\times N}$ is additive noise for the inliers $\boldsymbol{\mathcal{X}}$. Since noiseless DPCP is a special case of the noisy problem with $\boldsymbol{\mathcal{E}} = \mathbf{0}$, in the rest of the section, unless stated otherwise, the dataset $\widetilde{\boldsymbol{\mathcal{X}}}$ refers to the one containing noisy inliers.

Towards analyzing the noisy DPCP problem, we first define the random spherical model under which the data points for all the simulations in this chapter are generated.

Definition 1 (Random spherical model for a single subspace)**.** *Consider a random spherical model where the columns of $\boldsymbol{\mathcal{O}}$ are drawn uniformly from the sphere $\mathbb{S}^{D-1}$, the columns of noisy inliers $\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}}$ are drawn by first independently generating inliers from $\mathcal{N}\left(\mathbf{0}, \frac{1}{d}\mathcal{P}_{\mathcal{S}}\right)$ and noise from $\mathcal{N}\left(\mathbf{0}, \frac{\sigma^2}{D}\mathbf{I}_D\right)$, and then projecting their sum onto $\mathbb{S}^{D-1}$, where $d = \dim(\mathcal{S})$, $\mathcal{P}_{\mathcal{S}}$ is the ortho-projector onto $\mathcal{S}$, and $\sigma \geq 0$ controls the amount of noise present in the inliers; under this model, the SNR is $\mathbb{E}[\|\boldsymbol{\mathcal{X}}\|_F]/\,\mathbb{E}[\|\boldsymbol{\mathcal{E}}\|_F] = 1/\sigma$. In the analysis presented in this thesis, we always assume $\sigma < 1$.*

We aim to provide a global optimality analysis for the noisy DPCP problem (3.1). Note that any global solution $\boldsymbol{b}$ to (3.1) must be a critical point, i.e., there exists $\boldsymbol{v} \in \partial \left\| \widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b} \right\|_1$ such that $(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\boldsymbol{v} = \mathbf{0}$, where

$$\partial \left\| \widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b} \right\|_1 = (\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}}) \operatorname{Sgn}\left((\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}})^\top \boldsymbol{b} \right) + \boldsymbol{\mathcal{O}} \operatorname{Sgn}\left(\boldsymbol{\mathcal{O}}^\top \boldsymbol{b} \right).$$

When noise is not present (i.e., $\boldsymbol{\mathcal{E}} = \mathbf{0}$), the term $\operatorname{Sgn}\left((\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}})^\top \boldsymbol{b} \right) = \operatorname{Sgn}(\boldsymbol{\mathcal{X}}^\top \boldsymbol{b})$ is simple since it only relates to inliers. In the noisy case, however, it is much more complicated to deal with this term. For example, since the function sign is discontinuous, $\operatorname{Sgn}\left((\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}})^\top \boldsymbol{b} \right)$ cannot easily be separated into two parts with one part only involving the inliers and the other part only involving the noise. As a consequence, compared to the noiseless case, a significantly more technical analysis is required to analyze the effect of noise.

Geometric quantities. We now introduce several helpful geometric quantities for analyzing the noisy DPCP problem. Since the noise dose not affect the outlier term $\boldsymbol{\mathcal{O}}$ in the dataset $\widetilde{\boldsymbol{\mathcal{X}}}$, we borrow the quantities $c_{\boldsymbol{\mathcal{O}},\max}, c_{\boldsymbol{\mathcal{O}},\min}$ and $\eta_{\boldsymbol{\mathcal{O}}}$ (see Section 3.1.1) from the noiseless analysis [152, 153] to characterize the distribution of the outlier points. As for noisy inliers, to facilitate an analysis, we decompose the noise as $\boldsymbol{\mathcal{E}} = \boldsymbol{\mathcal{E}}_{\boldsymbol{s}} + \boldsymbol{\mathcal{E}}_{\boldsymbol{n}}$, where $\boldsymbol{\mathcal{E}}_{\boldsymbol{s}}$ is the projection of the noise onto $\mathcal{S}$ and $\boldsymbol{\mathcal{E}}_{\boldsymbol{n}}$ is the projection of the noise onto $\mathcal{S}^\perp$. Denote $\widehat{\boldsymbol{\mathcal{X}}} := \boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}}_{\boldsymbol{s}}$ and $\widehat{\boldsymbol{\mathcal{E}}} := \boldsymbol{\mathcal{E}}_{\boldsymbol{n}}$ so that the columns $\{\widehat{\boldsymbol{x}}_j\}$ of $\widehat{\boldsymbol{\mathcal{X}}}$ lie in $\mathcal{S}$ and the columns $\{\widehat{\boldsymbol{\epsilon}}_j\}$ of $\widehat{\boldsymbol{\mathcal{E}}}$ lie in $\mathcal{S}^\perp$ for $j = 1, \cdots, N$. $\widehat{\boldsymbol{\mathcal{X}}}$ can be viewed as *effective inliers* since they lie in $\mathcal{S}$, whereas $\widehat{\boldsymbol{\mathcal{E}}}$ can be interpreted as *effective noise* because it perturbs $\widehat{\boldsymbol{\mathcal{X}}}$ away from $\mathcal{S}$. Similar to $c_{\boldsymbol{\mathcal{X}},\min}$ in (3.4), we define the permeance statistic [61] associate with the effective inliers as

$$c_{\widehat{\boldsymbol{\mathcal{X}}},\min} := \frac{1}{N} \min_{\boldsymbol{b} \in \mathcal{S} \cap \mathbb{S}^{D-1}} \left\| \widehat{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b} \right\|_1, \tag{3.10}$$

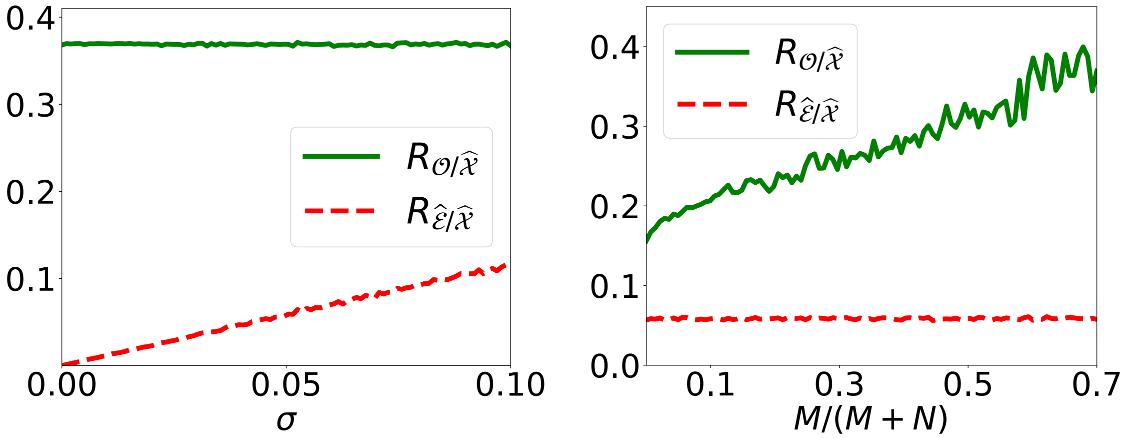
which attains larger values for better distributed inliers. Note that $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$ involves a mixture of inliers and components of noise projected onto $\mathcal{S}$. This particular integration of inliers and noise leads to tighter deterministic bounds in the deterministic phase of our analysis. Next, we capture the effective noise $\widehat{\boldsymbol{\mathcal{E}}}$ via the quantity

$$c_{\widehat{\boldsymbol{\mathcal{E}}},\max} := \frac{1}{N} \max_{\boldsymbol{b} \in \mathcal{S}^\perp \cap \mathbb{S}^{D-1}} \left\| \widehat{\boldsymbol{\mathcal{E}}}^\top \boldsymbol{b} \right\|_1, \tag{3.11}$$

which is closely related to the *total inlier residual* $\mathscr{R}(\mathcal{S}) := \frac{1}{N} \sum_{j=1}^{N} \|\widehat{\boldsymbol{\epsilon}}_j\|_2$ used by [61] to measure the level of the effective noise. By the Cauchy-Schwartz inequality $\left|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b}\right| \leq \|\widehat{\boldsymbol{\epsilon}}_j\|_2 \|\boldsymbol{b}\|_2$, it is clear that $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$ is a lower bound of $\mathscr{R}(\mathcal{S})$ since $\|\boldsymbol{b}\|_2 = 1$. Indeed, $\mathscr{R}(\mathcal{S})$ only depends on the energy of $\widehat{\boldsymbol{\mathcal{E}}}$, whereas $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$ also depends on the distribution of $\widehat{\boldsymbol{\mathcal{E}}}$: the more uniformly distributed $\widehat{\boldsymbol{\mathcal{E}}}$ is in $\mathcal{S}^\perp$, the smaller $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$ becomes. Thus, $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$ leads to a tighter result in our analysis than if one used $\mathscr{R}(\mathcal{S})$. Finally, two more definitions are needed for our analysis:

$$R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} := \frac{M}{N} \frac{\overline{\eta}_{\boldsymbol{\mathcal{O}}}}{c_{\widehat{\boldsymbol{\mathcal{X}}},\min}} \quad \text{and} \quad R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} := \frac{c_{\widehat{\boldsymbol{\mathcal{E}}},\max}}{c_{\widehat{\boldsymbol{\mathcal{X}}},\min}}. \tag{3.12}$$

$R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}$ and $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ can be simply viewed as *outlier-to-inlier* and *noise-to-inlier* types of ratios, respectively. When we fix the inliers and outliers, $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ is proportional to the noise level (see Figure 3.1a). Similarly, when we fix the inliers and noise level, $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}$



(a) Varying σ

(b) Varying outlier ratio

Figure 3.1. Plots of $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}$ and $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ as a function of (a) σ and (b) outlier ratio. Here we fix $D = 30, d = 29, N = 1500$, and $M/(M+N) = 0.7$ in (a), and $\sigma = 0.05$ in (b).

is proportional to the number of outliers (see Figure 3.1b).

3.1.2.1 Geometry of the critical points

For the rest of the analysis, let $\theta \in [0, \pi/2]$ be the principal angle of a vector $\boldsymbol{b} \in \mathbb{S}^{D-1}$ from the orthogonal complement subspace $\mathcal{S}^{\perp}$. Thus, $\boldsymbol{b}$ is normal to $\mathcal{S}$ if and only if $\theta = 0$. Using $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}$ and $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ defined in (3.12), we can now characterize the geometry of the critical points of the noisy DPCP problem (3.1).

Lemma 3. *Assume* $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} < 1$ *and*

$$\frac{32 R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}}{\left(\sqrt{R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} + 8} - 3R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}\right)^{\frac{3}{2}} \left(\sqrt{R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} + 8} + R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}\right)^{\frac{1}{2}}} < 1, \tag{3.13}$$

then any critical point $\boldsymbol{b}$ *of problem* (3.1) *has its principal angle* θ *from* $\mathcal{S}^\perp$ *satisfy*

$$\theta \leq \sin^{-1}(t_1) \quad \textit{or} \quad \theta \geq \sin^{-1}(t_2), \tag{3.14}$$

where $0 \leq t_1 \leq t_2 \leq 1$ *are the two nonnegative roots of the quartic equation*

$$t^4 + \left(R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 1\right) t^2 + 4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} t + 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} = 0. \tag{3.15}$$

Proof. As the first part of the proof, we prove a useful result.

Sublemma 1. *Given* $0 \leq \alpha < 1$ *and* $\beta > 0$*, the equation*

$$h(\phi) := \sin(\phi)\cos(\phi) - \alpha \sin(\phi) - 2\beta = 0$$

has two roots in $(0, \pi/2)$ *if and only if*

$$\frac{32\beta}{\left(\sqrt{\alpha^2+8} - 3\alpha\right)^{3/2} \left(\sqrt{\alpha^2+8} + \alpha\right)^{1/2}} < 1.$$

Proof. Note that $h(0) = -2\beta < 0$ and $h(\pi/2) = -\alpha - 2\beta < 0$. One can compute its derivative as

$$h'(\phi) = 2\cos^2(\phi) - \alpha\cos(\phi) - 1$$

and $h'(0) = 1 - \alpha > 0$ and $h'(\pi/2) = -1 < 0$, which means $h(\phi)$ is increasing at $\phi = 0$

and is decreasing at $\phi = \pi/2$. By solving $h'(\bar{\phi}) = 0$, we obtain

$$\cos(\bar{\phi}) = \frac{\alpha + \sqrt{\alpha^2 + 8}}{4} > 0 \quad \text{or} \quad \cos(\bar{\phi}) = \frac{\alpha - \sqrt{\alpha^2 + 8}}{4} < 0.$$

Since we are only interested in the domain $[0, \pi/2]$, the second solution is discarded. Moreover, $\alpha \in [0, 1)$ implies $(\alpha + \sqrt{\alpha^2 + 8})/4 \in [\sqrt{2}/2, 1)$, so $\bar{\phi} = \arccos((\alpha + \sqrt{\alpha^2 + 8})/4)$ is indeed a extreme point of $h(\cdot)$ in $[0, \pi/2]$. Combining the facts that $h'(\phi) > 0$ for $\phi \in [0, \arccos(\bar{\phi}))$, $h'(\phi) < 0$ for $\phi \in (\arccos(\bar{\phi}), \pi/2]$, and there is only one extreme point $\bar{\phi}$ in $[0, \pi/2]$, we know that $\bar{\phi}$ is a maximizer. Therefore, $h(\phi)$ has two roots in $(0, \pi/2)$ if and only if $h(\bar{\phi}) > 0$, which is further equivalent to

$$\begin{aligned}
& \sin(\bar{\phi})\cos(\bar{\phi}) - \alpha\sin(\bar{\phi}) - 2\beta > 0 \\
\Leftrightarrow \quad & (1 - \cos^2(\bar{\phi}))(\cos(\bar{\phi}) - \alpha)^2 > 4\beta^2 \\
\Leftrightarrow \quad & \left(8 - 2\alpha^2 - 2\alpha\sqrt{\alpha^2 + 8}\right)\left(\sqrt{\alpha^2 + 8} - 3\alpha\right)^2 > (32\beta)^2 \\
\Leftrightarrow \quad & \left(\sqrt{\alpha^2 + 8} + \alpha\right)\left(\sqrt{\alpha^2 + 8} - 3\alpha\right)^3 > (32\beta)^2 \\
\Leftrightarrow \quad & \left(\sqrt{\alpha^2 + 8} - 3\alpha\right)^{3/2}\left(\sqrt{\alpha^2 + 8} + \alpha\right)^{1/2} > 32\beta,
\end{aligned}$$

thus completing the proof of the sublemma. □

Continuing with the proof of Lemma 3, we show that as long as $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} < 1$ and (3.13) holds, the quartic equation (3.15) must have exactly two roots in $[0, 1]$. We consider two cases: $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} = 0$ and $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} > 0$. If $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} = 0$, then the quartic equation

(3.15) reduces to

$$t^4 + (R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 1)t^2 = 0, \tag{3.16}$$

which has two roots in $[0,1]$ $\Big($namely 0 and $\sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}\Big)$. On the other hand, if $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} > 0$, consider the following equation for $\phi \in (0, \pi/2)$:

$$h(\phi) := \sin(\phi)\cos(\phi) - \sin(\phi) R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} = 0. \tag{3.17}$$

Letting $t := \sin(\phi) \in (0,1)$, we have

$$t\sqrt{1-t^2} = tR_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} + 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}},$$

which is equivalent to the quartic equation (3.15):

$$t^4 + (R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 1)t^2 + 4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} t + 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} = 0. \tag{3.18}$$

This tells us that each root $\phi \in (0, \pi/2)$ of $h(\cdot)$ corresponds to a root $t = \sin(\phi) \in (0,1)$ of (3.15). It follows from Sublemma 1 that condition (3.13), $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} \in [0,1]$ and $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} > 0$ together ensure that the equation in (3.17) has two roots $\phi_1, \phi_2 \in (0, \pi/2)$. Then, from the above discussion, we know that $\sin(\phi_1)$ and $\sin(\phi_2)$ are two positive roots in $[0,1]$ of the quartic equation (3.15). Moreover, according to Descartes' rule of signs, (3.15) has zero or two positive roots, so that there are no other positive roots. Therefore, we conclude that if $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} < 1$ and (3.13) holds, the quartic equation (3.15) must have exactly two roots in $[0, 1]$, and we denote them as $0 \leq t_1 \leq t_2 \leq 1$.

Next, let us consider the geometry of the critical points of problem (3.1). Similarly, we consider two cases: $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} = 0$ and $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} > 0$. If $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} = 0$, the problem reduces to the noiseless case as analyzed in [152, 153] with dataset $[\widehat{\boldsymbol{\mathcal{X}}}\ \boldsymbol{\mathcal{O}}]$ (recall that the points in $\widehat{\boldsymbol{\mathcal{X}}}$ lie perfectly in the inlier subspace $\mathcal{S}$). According to Lemma 1, we have

$$\sin(\theta) = 0 \quad \text{or} \quad \sin(\theta) \geq \sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}. \tag{3.19}$$

Moreover, according to (3.16), in this case we solve for the two roots of (3.15) as $t_1 = 0$ and $t_2 = \sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}$. Combine this with (3.19), we obtain

$$\sin(\theta) = t_1 \quad \text{or} \quad \sin(\theta) \geq t_2. \tag{3.20}$$

In the remainder of the analysis, we consider the case $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} > 0$. For any critical point $\boldsymbol{b}$ of problem (3.1), we decompose it as $\boldsymbol{b} = \sin(\theta)\boldsymbol{s} + \cos(\theta)\boldsymbol{n}$, where $\theta \in [0, \pi/2]$ is the principal angle of $\boldsymbol{b}$ from $\mathcal{S}^\perp$, $\boldsymbol{s} \in \mathcal{S}, and \boldsymbol{n} \in \mathcal{S}^\perp$ with $\|\boldsymbol{s}\|_2 = \|\boldsymbol{n}\|_2 = 1$. Note that if $\theta = 0$ or $\theta = \pi/2$, then (3.14) trivially holds. Hence, for the remainder, assume that $\theta \in (0, \pi/2)$. For any critical point $\boldsymbol{b}$, there exists a vector $\boldsymbol{v} \in \partial \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_1$ so that

$$\left(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top\right)\boldsymbol{v} = \mathbf{0}. \tag{3.21}$$

We further explicitly write down the subdifferential of $\left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_1$:

$$\partial \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_1 = \sum_{j=1}^{N} \operatorname{Sgn}\left((\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{b}\right)(\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j) + \sum_{j=1}^{M} \operatorname{Sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{b}\right)\boldsymbol{o}_j,$$

and thus the $\boldsymbol{v}$ as defined above in (3.21) can be written as

$$\boldsymbol{v} = \sum_{j=1}^{N} \operatorname{sgn}\left((\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{b}\right)(\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j) + \sum_{j=1}^{M} \operatorname{sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{b}\right)\boldsymbol{o}_j, \tag{3.22}$$

where $\operatorname{sgn}(x)$ denotes a specific element that belongs to the subdifferential $\operatorname{Sgn}(x)$.

Define $\boldsymbol{y} := -\cos(\theta)\boldsymbol{s} + \sin(\theta)\boldsymbol{n}$, which is orthogonal to $\boldsymbol{b}$. Then, we have that

$$\begin{aligned}
0 &= \left|\left\langle \left(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top\right)\boldsymbol{v}, \boldsymbol{y}\right\rangle\right| = \left|\boldsymbol{v}^\top \left(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top\right)\boldsymbol{y}\right| = \left|\boldsymbol{v}^\top \boldsymbol{y}\right| \\
&= \left|\sum_{j=1}^{N} \operatorname{sgn}\left((\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{b}\right)(\widehat{\boldsymbol{x}}_j^\top + \widehat{\boldsymbol{\epsilon}}_j^\top)\boldsymbol{y} + \sum_{j=1}^{M} \operatorname{sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{b}\right)\boldsymbol{o}_j^\top \boldsymbol{y}\right| \\
&= \left|\sum_{j=1}^{N} \operatorname{sgn}\left(\sin(\theta)\widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} + \cos(\theta)\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n}\right)\left(-\cos(\theta)\widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} + \sin(\theta)\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n}\right)\right. \qquad (3.23) \\
&\qquad \left. + \sum_{j=1}^{M} \operatorname{sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{b}\right)\boldsymbol{o}_j^\top \boldsymbol{y}\right| \\
&= \left|-\cos(\theta)\sum_{j=1}^{N} \operatorname{sgn}\left(\widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} + \cot(\theta)\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n}\right)\left(\widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} - \tan(\theta)\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n}\right) + \sum_{j=1}^{M} \operatorname{sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{b}\right)\boldsymbol{o}_j^\top \boldsymbol{y}\right| \\
&= \left|-\cos(\theta)\sum_{j=1}^{N} \operatorname{sgn}\left(\widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} + \cot(\theta)\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n}\right)\left(\widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} + \cot(\theta)\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n}\right)\right. \\
&\qquad \left. + \cos(\theta)\sum_{j=1}^{N} \operatorname{sgn}\left(\widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} + \cot(\theta)\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n}\right)\left(\cot(\theta) + \tan(\theta)\right)\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n} + \sum_{j=1}^{M} \operatorname{sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{b}\right)\boldsymbol{o}_j^\top \boldsymbol{y}\right| \\
&\qquad\qquad (3.24) \\
&\geq \cos(\theta)\sum_{j=1}^{N} \left|\widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} + \cot(\theta)\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n}\right| - \left|\sum_{j=1}^{M} \operatorname{sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{b}\right)\boldsymbol{o}_j^\top \boldsymbol{y}\right|
\end{aligned}$$

$$
\begin{aligned}
&\quad - \cos(\theta) \left| \sum_{j=1}^{N} \operatorname{sgn}\left(\widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} + \cot(\theta) \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n} \right) \left(\cot(\theta) + \tan(\theta) \right) \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n} \right| \\
&\geq \cos(\theta) \sum_{j=1}^{N} \left| \widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} + \cot(\theta) \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n} \right| - \cos(\theta) \sum_{j=1}^{N} \left(\cot(\theta) + \tan(\theta) \right) \left| \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n} \right| \\
&\quad - \left| \sum_{j=1}^{M} \operatorname{sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{b} \right) \boldsymbol{o}_j^\top \boldsymbol{y} \right| \\
&\geq \cos(\theta) \sum_{j=1}^{N} \left| \widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} \right| - \left(2\cos(\theta)\cot(\theta) + \sin(\theta) \right) \sum_{j=1}^{N} \left| \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n} \right| - \left| \sum_{j=1}^{M} \operatorname{sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{b} \right) \boldsymbol{o}_j^\top \boldsymbol{y} \right| \\
&> \cos(\theta) N c_{\widehat{\boldsymbol{\mathcal{X}}},\min} - \frac{2}{\sin(\theta)} N c_{\widehat{\boldsymbol{\mathcal{E}}},\max} - M \overline{\eta}_{\boldsymbol{\mathcal{O}}}, \qquad (3.25)
\end{aligned}
$$

where (3.23) follows from the decomposition of $\boldsymbol{b}$ and $\boldsymbol{y}$ plus the fact that $\widehat{\boldsymbol{x}}_j \perp \boldsymbol{n}, \widehat{\boldsymbol{\epsilon}}_j \perp \boldsymbol{s}$, (3.24) follows from

$$
\widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} - \tan(\theta) \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n} = \left(\widehat{\boldsymbol{x}}_j^\top \boldsymbol{s} + \cot(\theta) \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n} \right) - \left(\cot(\theta) \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n} + \tan(\theta) \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n} \right),
$$

(3.25) uses the definition of $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$ in (3.10), the definition of $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$ in (3.11),

$$
-(2\cos(\theta)\cot(\theta) + \sin(\theta)) = -\frac{2\cos^2(\theta) + \sin^2(\theta)}{\sin(\theta)} = -\frac{2 - \sin^2(\theta)}{\sin(\theta)} > -\frac{2}{\sin(\theta)},
$$

and the fact that the general position [152, 153] of data ensures that $\boldsymbol{b}$ can be orthogonal to at most D columns of $\boldsymbol{\mathcal{O}}$:

$$\begin{aligned}\left|\sum_{j=1}^{M} \operatorname{sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{b}\right) \boldsymbol{o}_j^\top \boldsymbol{y}\right| &= \left|\sum_{\{j:\boldsymbol{o}_j^\top \boldsymbol{b}\neq 0\}} \operatorname{sign}\left(\boldsymbol{o}_j^\top \boldsymbol{b}\right) \boldsymbol{o}_j^\top \boldsymbol{y} + \sum_{\{j:\boldsymbol{o}_j^\top \boldsymbol{b}=0\}} \operatorname{sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{b}\right) \boldsymbol{o}_j^\top \boldsymbol{y}\right| \\ &\le \left|\sum_{j=1}^{M} \operatorname{sign}\left(\boldsymbol{o}_j^\top \boldsymbol{b}\right) \boldsymbol{o}_j^\top \boldsymbol{y}\right| + \left|\sum_{\{j:\boldsymbol{o}_j^\top \boldsymbol{b}=0\}} \operatorname{sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{b}\right) \boldsymbol{o}_j^\top \boldsymbol{y}\right| \\ &\le \max_{\boldsymbol{g},\boldsymbol{b}\in\mathbb{S}^{D-1},\boldsymbol{g}\perp\boldsymbol{b}} \left|\boldsymbol{g}^\top \boldsymbol{\mathcal{O}} \operatorname{sign}(\boldsymbol{\mathcal{O}}^\top \boldsymbol{b})\right| + D \\ &= \max_{\boldsymbol{b}\in\mathbb{S}^{D-1}} \left\|(\boldsymbol{I} - \boldsymbol{b}\boldsymbol{b}^\top)\boldsymbol{\mathcal{O}} \operatorname{sign}(\boldsymbol{\mathcal{O}}^\top \boldsymbol{b})\right\|_2 + D \\ &= M(\eta_{\boldsymbol{\mathcal{O}}} + D/M) = M\overline{\eta}_{\boldsymbol{\mathcal{O}}}.\end{aligned}$$

Therefore, we obtain

$$0 > \cos(\theta) N c_{\widehat{\boldsymbol{\mathcal{X}}},\min} - \frac{2}{\sin(\theta)} N c_{\widehat{\boldsymbol{\mathcal{E}}},\max} - M\overline{\eta}_{\boldsymbol{\mathcal{O}}}, \tag{3.26}$$

which is equivalent to

$$\cos(\theta) - \frac{2}{\sin(\theta)} \frac{N c_{\widehat{\boldsymbol{\mathcal{E}}},\max}}{N c_{\widehat{\boldsymbol{\mathcal{X}}},\min}} - \frac{M\overline{\eta}_{\boldsymbol{\mathcal{O}}}}{N c_{\widehat{\boldsymbol{\mathcal{X}}},\min}} < 0 \tag{3.27}$$

$$\Leftrightarrow \quad \cos(\theta) - \frac{2}{\sin(\theta)} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} - R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} < 0 \tag{3.28}$$

$$\Leftrightarrow \quad \sin(\theta)\cos(\theta) - \sin(\theta) R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} < 0 \tag{3.29}$$

where we use the fact that $N c_{\widehat{\boldsymbol{\mathcal{X}}},\min} \neq 0$ since $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} < 1$, and the definitions of $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ and $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}$ from (3.12).

From (3.29), Sublemma 1, $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \in [0, 1)$, $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} > 0$ and condition (3.13) we know that $h(\theta) := \sin(\theta)\cos(\theta) - \sin(\theta) R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ has two zeros $\theta_1, \theta_2 \in [0, \pi/2]$

(suppose $\theta_1 < \theta_2$). Based on the proof of Sublemma 1, we know that $h(\theta) < 0$ for $\theta \in (0, \theta_1) \cup (\theta_2, \pi/2)$. In other words, the solution for inequality (3.29) satisfies either

$$0 < \theta < \theta_1 \quad \text{or} \quad \theta_2 < \theta < \pi/2. \tag{3.30}$$

From the previous discussions on (3.17) and (3.18), we know that zeros $\theta_1, \theta_2 \in [0, \pi/2]$ of $h(\cdot)$ correspond to roots $t_1 = \sin(\theta_1)$ and $t_2 = \sin(\theta_2)$ of the quartic equation (3.15). Combining this fact with (3.30), we have either

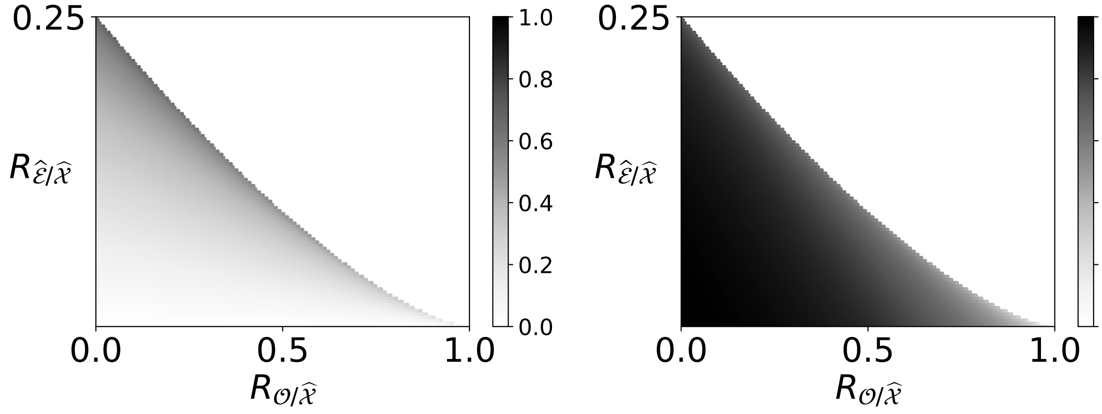
$$\sin(\theta) < t_1 \quad \text{or} \quad \sin(\theta) > t_2. \tag{3.31}$$

Finally, (3.20) and (3.31) together imply that any critical point $\boldsymbol{b}$ of problem (3.1) must have its principal angle θ from $\mathcal{S}^\perp$ satisfy either

$$\sin(\theta) \leq t_1 \quad \text{or} \quad \sin(\theta) \geq t_2$$

where $0 \leq t_1 \leq t_2 \leq 1$ are the two nonnegative roots of the quartic equation (3.15). □

Discussion of Lemma 3. First note that $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} < 1$ ensures that the denominator of the LHS in (3.13) is well-defined. Since the function $a \mapsto f(a) = \left(\sqrt{a^2+8} - 3a\right)^{\frac{3}{2}} \left(\sqrt{a^2+8} + a\right)^{\frac{1}{2}}$ is decreasing between $[0, 1]$ with $f(0) = 8$ and $f(1) = 0$, (3.13) implies that larger noise levels lead to smaller numbers of outliers that DPCP can tolerate. With (3.13), it can be shown that (3.15) has two nonnegative roots $0 \leq t_1 \leq t_2 \leq 1$, and (3.14) implies that none of the critical points



(a) Value of t_1 **(b)** Value of t_2

Figure 3.2. Plot of (a) t_1 and (b) t_2 while varying $R_{\mathcal{O}/\widehat{\mathcal{X}}}$ and $R_{\widehat{\mathcal{E}}/\widehat{\mathcal{X}}}$. In each plot, condition (3.13) holds only below the curve, which corresponds to valid pairs of $\left(R_{\mathcal{O}/\widehat{\mathcal{X}}}, R_{\widehat{\mathcal{E}}/\widehat{\mathcal{X}}}\right)$.

have principal angle in $(\sin^{-1}(t_1), \sin^{-1}(t_2))$. Figure 3.2 displays t_1 and t_2 while varying $R_{\mathcal{O}/\widehat{\mathcal{X}}}$ and $R_{\widehat{\mathcal{E}}/\widehat{\mathcal{X}}}$ under condition (3.13). One can observe that smaller percentages of outliers and noise levels lead to t_1 being closer to 0 and t_2 being closer to 1, which means that critical points of (3.1) either lie in a neighborhood of $\mathcal{S}^\perp$ or close to $\mathcal{S}$.

When there is no noise ($\boldsymbol{\mathcal{E}} = \mathbf{0}$), Lemma 3 reduces to Lemma 1 [152, 153]: $R_{\widehat{\mathcal{E}}/\widehat{\mathcal{X}}} = 0$ and $R_{\mathcal{O}/\widehat{\mathcal{X}}} = \overline{\eta}_{\mathcal{O}}/c_{\mathcal{X},\min}$, so that (3.13) always holds and (3.15) becomes $t^4 + ((\overline{\eta}_{\mathcal{O}}/c_{\mathcal{X},\min})^2 - 1)t^2 = 0$, which implies $t_1 = 0$ and $t_2 = \sqrt{1 - (\overline{\eta}_{\mathcal{O}}/c_{\mathcal{X},\min})^2}$. Nevertheless, we stress that the proof for Lemma 3 is far more complicated than for the noiseless case, partly because of the need to deal with $\operatorname{Sgn}\left((\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}})^\top \boldsymbol{b}\right)$.

Proposition 1. *Assume $R_{\mathcal{O}/\widehat{\mathcal{X}}} < 1$ and condition (3.13) holds. Let $0 \le t_1 \le t_2 \le 1$ be the two nonnegative roots of the quartic equation (3.15), then t_1 is upper bounded by*

$$t_1 \le \frac{25}{\left(1 - R_{\mathcal{O}/\widehat{\mathcal{X}}}\right)^2} \cdot R_{\widehat{\mathcal{E}}/\widehat{\mathcal{X}}}. \tag{3.32}$$

Proof. First of all, according to the proof of Lemma 3, $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} < 1$ and condition (3.13) ensure that the quartic equation (3.15) must exactly have two roots in $[0, 1]$, and we denote them by t_1 and t_2 with $0 \leq t_1 \leq t_2 \leq 1$. Let t_3 and t_4 be the other two roots of (3.15). Based on the relationship between coefficients and roots (Vieta's Formulas),we know that

$$\sum_{i=1}^{4} t_i = 0 \quad \text{and} \quad \sum_{i,j} t_i t_j = R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 1, \tag{3.33}$$

and thus

$$\sum_{i=1}^{4} t_i^2 = \left(\sum_{i=1}^{4} t_i\right)^2 - 2\sum_{i,j} t_i t_j = 2\left(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}\right). \tag{3.34}$$

Setting $\nu := \sqrt{2\left(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}\right)}$, it then follows that

$$\nu^2 = \sum_{i=1}^{4} t_i^2 \geq t_j^2 \quad \text{or} \quad \nu \geq t_j \quad \text{for } j = 1, \cdots, 4. \tag{3.35}$$

We claim that the roots $t_i, i = 1, \cdots, 4$, must satisfy the following inequality:

$$t^4 + \left(R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 1\right) t^2 + 4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \nu + 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \geq 0, \tag{3.36}$$

where the original linear term $4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} t$ in (3.15) reduces to a constant term $4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \nu$ in (3.36), thus allowing (3.36) to be solved as a quadratic inequality. In fact, for any $t_i, i = 1, \cdots, 4$, we have

$$\begin{aligned}
&t_i^4 + \left(R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 1\right) t_i^2 + 4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \nu + 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \\
=& \left(t_i^4 + (R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 1)t_i^2 + 4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} t_i + 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}\right) + 4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \nu - 4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} t_i \\
=& \, 0 + 4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} (\nu - t_i) \geq 0,
\end{aligned}$$

where we used (3.35) and that $t_i, i = 1, \cdots, 4$, are solutions to equation (3.15).

By viewing (3.36) as a quadratic inequality of $u := t^2$, we have

$$u^2 + \left(R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 1\right) u + 4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \nu + 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \geq 0, \tag{3.37}$$

and the associated quadratic function has the following two zeros:

$$\begin{aligned}
u_{1,2} &= \frac{(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}) \pm \sqrt{(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})^2 - 16R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} - 16R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \nu}}{2} \\
&= \tfrac{1}{2}(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}) \pm \sqrt{\left(\tfrac{1}{2}(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})\right)^2 - 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} - 4\sqrt{2} R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}}.
\end{aligned} \tag{3.38}$$

Note that $(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})^2 - 16R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} - 16R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \nu \geq 0$ is guaranteed by (3.13), and $u_{1,2} \in [0, 1]$. Therefore, (3.36) implies that there exists $t' > 0$ such that t_2 satisfies

$$\begin{aligned}
t_2^2 \geq (t')^2 :=& \tfrac{1}{2}(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}) \\
&+ \sqrt{\left(\tfrac{1}{2}(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})\right)^2 - 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} - 4\sqrt{2} R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}}
\end{aligned} \tag{3.39}$$

In addition to the first and second order symmetric sums (3.33) for the roots of (3.15), we also have the third and fourth order relationships

$$t_1t_2t_3 + t_1t_2t_4 + t_2t_3t_4 + t_1t_3t_4 = -4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \quad \text{and} \quad t_1t_2t_3t_4 = 4R^2_{\boldsymbol{\mathcal{E}}/\widehat{\boldsymbol{\mathcal{X}}}}. \tag{3.40}$$

Reorganizing the first equation in (3.40), we have

$$t_1t_2(t_3 + t_4) + (t_1 + t_2)t_3t_4 = -4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}. \tag{3.41}$$

Since $t_3 + t_4 = -(t_1 + t_2)$ from (3.33), we rewrite (3.41) as

$$(t_1 + t_2)(t_3t_4 - t_1t_2) = -4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}},$$

which implies $t_3t_4 \leq t_1t_2$. Combine this with the second equation in (3.40), we have $t_3t_4 \leq 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$. Noticing that

$$2t_3t_4 = (t_3 + t_4)^2 - t_3^2 - t_4^2 = (t_1 + t_2)^2 - 2(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}) + t_1^2 + t_2^2$$

follows from (3.33) and (3.34), we find together with $t_3t_4 \leq 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ that

$$t_1^2 + t_2t_1 + t_2^2 - 2R_{\boldsymbol{\mathcal{E}}/\widehat{\boldsymbol{\mathcal{X}}}} - (1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}) \leq 0. \tag{3.42}$$

Viewing this as a quadratic inequality with respect to t_1, we solve (3.42) for t_1:

$$t_1 \leq \frac{1}{2}\left(-t_2 + \sqrt{4(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}) + 8R_{\boldsymbol{\mathcal{E}}/\widehat{\boldsymbol{\mathcal{X}}}} - 3t_2^2}\right). \tag{3.43}$$

We now validate that the square root in (3.43) is well-defined. In fact, we have

$$\begin{aligned}
4(1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})+8R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}-3t_2^2 &= 2(t_1^2+t_2^2+t_3^2+t_4^2)+8R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}-3t_2^2\\
&\geq 2(t_1^2+t_2^2+t_3^2+t_4^2)+4t_3t_4-3t_2^2\\
&= 2t_1^2-t_2^2+2(t_3+t_4)^2\\
&= 2t_1^2-t_2^2+2(t_1+t_2)^2\\
&= 4t_1^2+t_2^2+4t_1t_2>0,
\end{aligned}$$

where the first equality follows from (3.34), the inequality follows from $t_3t_4 \leq 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ and the third equality follows from (3.33).

Combine (3.43) with $t_2 \geq t' > 0$, we have

$$t_1 \leq \frac{1}{2}\left(-t'+\sqrt{4(1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})+8R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}-3(t')^2}\right). \tag{3.44}$$

Recalling the definition of t' in (3.39), we have

$$\begin{aligned}
&1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}-(t')^2\\
&=\tfrac{1}{2}(1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})-\sqrt{\left(\tfrac{1}{2}(1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})\right)^2-4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}-4\sqrt{2}R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}\sqrt{1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}}\\
&=\frac{4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}+4\sqrt{2}R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}\sqrt{1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}}{\tfrac{1}{2}(1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})+\sqrt{\left(\tfrac{1}{2}(1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})\right)^2-4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}-4\sqrt{2}R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}\sqrt{1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}}}\\
&\leq\frac{4R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}+2\sqrt{2}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}}{\tfrac{1}{2}(1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})}=\frac{8+4\sqrt{2}}{1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}
\end{aligned}$$

which follows from $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} < 1$ and $a\sqrt{1-a^2} \leq 1/2$ for any $0 \leq a < 1$. Then we obtain

$$-(t')^2 \leq -\left(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}\right) + \frac{8+4\sqrt{2}}{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}. \tag{3.45}$$

Now we are able to bound the RHS of (3.44):

$$\begin{aligned}
t_1 &\leq \frac{1}{2}\left(-t' + \sqrt{4(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}) + 8R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} - 3(t')^2}\right) = \frac{1}{2}\frac{4(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}) + 8R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} - 3(t')^2 - (t')^2}{\sqrt{4(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}) + 8R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} - 3(t')^2} + t'} \\
&\leq \frac{4R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} + 2(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}) - 2(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}) + \frac{16+8\sqrt{2}}{1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}}{\sqrt{4(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}) + 8R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}} - \sqrt{3(t')^2} + t'} \\
&\leq \frac{4R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} + \frac{16+8\sqrt{2}}{1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}}{\sqrt{4(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})} + (1-\sqrt{3})\sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}} \\
&= \frac{4R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} + \frac{16+8\sqrt{2}}{1-R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}}{(3-\sqrt{3})\sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}} = \frac{20 + 8\sqrt{2} - 4R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}{(3-\sqrt{3})(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})^{3/2}} \cdot R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \leq \frac{25}{(1 - R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})^2} \cdot R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}
\end{aligned}$$

where the second inequality follows from (3.45) and $\sqrt{a-b} \geq \sqrt{a} - \sqrt{b}$ for $a \geq b \geq 0$, the third inequality follows because the denominator is an increasing function of $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ (notice t' is itself a decreasing function of $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$) so we substitute $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ with 0 in the denominator to get an upper bound. This completes the proof. □

Discussion of Proposition 1. The upper bound for t_1 in (3.32) helps in further interpreting Lemma 3. In particular, this means that t_1 is close to 0 when $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ and $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}$ are small. More generally, for fixed $\boldsymbol{\mathcal{O}}$ and $\widehat{\boldsymbol{\mathcal{X}}}$, (3.32) guarantees that t_1 is perturbed away from 0 by at most the effective noise level, which is intuitive.

3.1.2.2 Geometry of the global solutions

We are now ready to provide deterministic conditions under which any global solution to the noisy DPCP problem (3.1) lies in a neighborhood of $\mathcal{S}^\perp$.

Theorem 3. *If* $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} < 1$, (3.13) *holds, and*

$$\frac{M}{N}\frac{c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min}}{c_{\widehat{\boldsymbol{\mathcal{X}}},\min}} < t_2 - 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}, \tag{3.46}$$

then any global minimizer $\boldsymbol{b}^*$ *of* (3.1) *must have its principal angle* θ^* *from* $\mathcal{S}^\perp$ *satisfy*

$$\theta^* \le \sin^{-1}(t_1), \tag{3.47}$$

where $0 \le t_1 \le t_2 \le 1$ *are the nonnegative roots of* (3.15)*.*

Proof. Since $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} < 1$ and (3.13) holds, we can apply Lemma 3 to obtain that any critical point $\boldsymbol{b}$ of problem (3.1) must have its principal angle θ from $\mathcal{S}^\perp$ satisfy

$$\sin(\theta) \le t_1 \quad \text{or} \quad \sin(\theta) \ge t_2,$$

where $0 \le t_1 \le t_2$ are the two nonnegative roots of (3.15). Since a global minimizer $\boldsymbol{b}^*$ must be a critical point, for the sake of contradiction, let us assume (3.47) does not hold, which allows us to conclude that

$$\sin(\theta^*) \ge t_2. \tag{3.48}$$

Moreover, for any global minimizer $\boldsymbol{b}^*$, we decompose it as $\boldsymbol{b}^* = \sin(\theta^*)\boldsymbol{s} + \cos(\theta^*)\boldsymbol{n}$, where $\boldsymbol{s} \in \mathcal{S}, \boldsymbol{n} \in \mathcal{S}^\perp$ and both $\boldsymbol{s}$ and $\boldsymbol{n}$ are unit vectors. Observing that

$$\begin{aligned}
\left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}^*\right\|_1 &= \min_{\boldsymbol{b}\in\mathbb{S}^{D-1}} \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_1 \leq \min_{\boldsymbol{b}\in\mathbb{S}^{D-1}\cap\mathcal{S}^\perp} \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_1 \\
&= \min_{\boldsymbol{b}\in\mathbb{S}^{D-1}\cap\mathcal{S}^\perp} \left\{\left\|\widehat{\boldsymbol{\mathcal{E}}}^\top \boldsymbol{b}\right\|_1 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{b}\right\|_1\right\} \leq Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max} + Mc_{\boldsymbol{\mathcal{O}},\max}
\end{aligned}$$

and

$$\begin{aligned}
\left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}^*\right\|_1 &= \sum_{j=1}^{N} \left|(\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{b}^*\right| + \sum_{j=1}^{M} \left|\boldsymbol{o}_j^\top \boldsymbol{b}^*\right| \\
&\geq \sum_{j=1}^{N} \left|\widehat{\boldsymbol{x}}_j^\top \boldsymbol{b}^*\right| - \sum_{j=1}^{N} \left|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b}^*\right| + \sum_{j=1}^{M} \left|\boldsymbol{o}_j^\top \boldsymbol{b}^*\right| \\
&= \sin(\theta^*) \sum_{j=1}^{N} \left|\widehat{\boldsymbol{x}}_j^\top \boldsymbol{s}\right| - \cos(\theta^*) \sum_{j=1}^{N} \left|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{n}\right| + \sum_{j=1}^{M} \left|\boldsymbol{o}_j^\top \boldsymbol{b}^*\right| \\
&\geq \sin(\theta^*) Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min} - Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max} + Mc_{\boldsymbol{\mathcal{O}},\min},
\end{aligned}$$

it follows that

$$\sin(\theta^*) Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min} - Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max} + Mc_{\boldsymbol{\mathcal{O}},\min} \leq Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max} + Mc_{\boldsymbol{\mathcal{O}},\max}$$

or, equivalently, that

$$\sin(\theta^*) \leq \frac{Mc_{\boldsymbol{\mathcal{O}},\max} - Mc_{\boldsymbol{\mathcal{O}},\min} + 2Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max}}{Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min}}. \tag{3.49}$$

Combine (3.48) and (3.49), we have

$$\frac{Mc_{\boldsymbol{\mathcal{O}},\max} - Mc_{\boldsymbol{\mathcal{O}},\min} + 2Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max}}{Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min}} \geq t_2,$$

and after rearranging and using the definition for $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ in (3.12), we have

$$\frac{M}{N}\frac{c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min}}{c_{\widehat{\boldsymbol{\mathcal{X}}},\min}} \geq t_2 - 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}},$$

which contradicts the condition (3.46), and thus completes the proof. □

Discussion of Theorem 3. Theorem 3 builds upon Lemma 3, with the intuition that critical points that are close to the subspace $\boldsymbol{\mathcal{S}}$ (i.e., for which $\theta^* \geq \sin^{-1}(t_2)$) cannot be global minimizers as they result in large objective values. As long as data points are well-distributed (small $c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min}$, large $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$, large t_2) and effective noise is mild (small $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$), (3.46) will be satisfied and global minimizers must be close to $\boldsymbol{\mathcal{S}}^\perp$. When $\boldsymbol{\mathcal{E}} = \mathbf{0}$, we have already remarked that $t_1 = 0$ and $t_2 = \sqrt{1 - (\overline{\eta}_{\boldsymbol{\mathcal{O}}}/c_{\boldsymbol{\mathcal{X}},\min})^2}$, which together with (3.46) and (3.47) imply that global minimizers are orthogonal to $\boldsymbol{\mathcal{S}}$ when

$$\frac{M}{N}\frac{c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min}}{c_{\boldsymbol{\mathcal{X}},\min}} < \sqrt{1 - (\overline{\eta}_{\boldsymbol{\mathcal{O}}}/c_{\boldsymbol{\mathcal{X}},\min})^2},$$

which is precisely Theorem 1 of [152, 153] under the noiseless setting.

Corollary 1. *Assume $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} < 1$. If it holds that*

$$\frac{M}{N}\frac{c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min}}{c_{\widehat{\boldsymbol{\mathcal{X}}},\min}} < t' - 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}, \tag{3.50}$$

then any global minimizer $\boldsymbol{b}^*$ of problem (3.1) *must have its principal angle θ^* from* *subspace* $\mathcal{S}^\perp$ *satisfy*

$$\sin(\theta^*) \leq \frac{25}{(1 - R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})^2} \cdot R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \tag{3.51}$$

where t' *is defined in* (3.39)*.*

Proof. We first show that if t' in (3.39) is well-defined, then (3.13) holds. In other words, we show that if

$$\left(\tfrac{1}{2}(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})\right)^2 - 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} - 4\sqrt{2}R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}\sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}} \geq 0 \tag{3.52}$$

then (3.13) holds. For the sake of contradiction, assume that (3.13) does not hold, so that according to the proof of Sublemma 1, we have that

$$h(\phi) := \sin(\phi)\cos(\phi) - R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}\sin(\phi) - 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \leq 0$$

for any $\phi \in [0, \pi/2]$. Let $t := \sin(\phi) \in [0, 1]$ so that

$$t\sqrt{1 - t^2} \leq R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}t + 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}},$$

which leads to

$$\widetilde{h}(t) := t^4 + (R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 1)t^2 + 4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}t + 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \geq 0 \tag{3.53}$$

for any $t \in [0,1]$. Now let us consider the quadratic function

$$\bar{h}(u) := u^2 + (R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 1)u + 4\sqrt{2}R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}\sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} + 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}},$$

whose discriminant $\Delta := (1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})^2 - 16R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} - 16\sqrt{2}R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}\sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}} \geq 0$, which is exactly the assumption (3.52). Notice that the minimizer of $\bar{h}(u)$ is $u^* = (1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}})/2 \in (0,1)$ for any u we have

$$\begin{aligned}
&\bar{h}(u) \geq \bar{h}(u^*) \\
&\quad = \left(\frac{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}{2}\right)^2 + (R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 1)\frac{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}{2} + 4\sqrt{2}R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}\sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} + 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \\
&\quad > \left(\frac{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}{2}\right)^2 + (R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} - 1)\frac{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}{2} + 4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}\sqrt{\frac{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}{2}}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} + 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} \\
&\quad = \widetilde{h}\left(\sqrt{\frac{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}}{2}}\right) \geq 0
\end{aligned}$$

where the second inequality follows from $4\sqrt{2} > 4/\sqrt{2}$ and the last inequality follows from (3.53). Since $\bar{h}(u) > 0$ for any u, which means $\Delta < 0$, we reach a contradiction to (3.52), i.e., $\Delta \geq 0$.

We proved that the existence of t' in (3.39) implies condition (3.13). Together with $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}} < 1$ and $t' \geq t_2$, we find that (3.46) holds. Then (3.51) directly follows from the results of Theorem 3 and Proposition 1. □

Discussion of Corollary 1. Corollary 1 is more interpretable than Theorem 3 in characterizing how global solutions to the noisy DPCP problem (3.1) are perturbed away from $\mathcal{S}^{\perp}$. Concretely, it removes the obscure quartic equation (3.15) and only relies on geometric quantities such as $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}$ and $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$. Still, it is of the deterministic type, and we are also interested in the statistical behavior of the geometric quantities in the noisy case that allows for a better understanding of the problem.

3.1.2.3 Probabilistic analysis

In this section, we provide a probabilistic characterization of global optimality of the noisy DPCP problem (3.1). Since the associated geometric quantities play critical roles in the deterministic analysis, understanding their statistical behavior is key to a probabilistic analysis. Towards that end, we first give their concentration bounds. Note that the outlier-related geometric quantities, namely $c_{\boldsymbol{\mathcal{O}},\max}, c_{\boldsymbol{\mathcal{O}},\min}$ and $\eta_{\boldsymbol{\mathcal{O}}}$, are the same as in the noiseless case under the random spherical model (see Definition 1), and their concentration bounds are already given in (3.9). We only derive concentration bounds for $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$ and $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$ that are newly introduced under the noisy setting.

We start by presenting some useful preliminary results in statistics.

Lemma 4. (McDiarmid's Inequality, [78]). *Let $Z_1, \ldots, Z_n$ be real-valued independent random variables, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function that satisfies*

$$\sup_{z_1,\cdots,z_n,z_i'} \left| f(z_1, \cdots, z_{i-1}, z_i, z_{i+1}, \cdots, z_n) - f(z_1, \ldots, z_{i-1}, z_i', z_{i+1}, \cdots, z_n) \right| \leq c_i,$$

for every $i = 1, \cdots, n$. Then

$$\mathbb{P}\left[\left|f(Z_1, \cdots, Z_n) - \mathbb{E}\left[f(Z_1, \cdots, Z_n)\right]\right| \geq \epsilon\right] \leq 2\exp\left(-\frac{2\epsilon^2}{\sum_{i=1}^n c_i^2}\right).$$

Lemma 5. (Rademacher Comparison, [59, Equation (4.20)])**.** *Let $F : \mathbb{R} \to \mathbb{R}$ be a convex and increasing function, $\varphi_i : \mathbb{R} \to \mathbb{R}$ for $1 \leq i \leq N$ be 1-Lipschitz functions such that $\varphi_i(0) = 0$, and ε_i for $1 \leq i \leq N$ be Rademacher random variables. Then, for any bounded subset $\mathcal{T}$ in $\mathbb{R}^N$, we have*

$$\mathbb{E}\left[F\left(\sup_{(t_1,\cdots,t_N)\in\mathcal{T}} \sum_{i=1}^N \varepsilon_i \varphi_i(t_i)\right)\right] \leq \mathbb{E}\left[F\left(\sup_{(t_1,\cdots,t_N)\in\mathcal{T}} \sum_{i=1}^N \varepsilon_i t_i\right)\right].$$

Lemma 6. (Rademacher Symmetrization, [56])**.** *Let $\mathcal{F}$ be a class of functions $f : \mathbb{R} \to \mathbb{R}$ such that $0 \leq f(z) \leq 1$, and ε_i for $1 \leq i \leq n$ be Rademacher random variables. Then for independent and identically distributed random variables $Z_1, \cdots, Z_n$, we have*

$$\mathbb{E}\left[\sup_{f\in\mathcal{F}}\left(\frac{1}{n}\sum_{i=1}^n f(Z_i) - \mathbb{E}[f(Z)]\right)\right] \leq 2\mathbb{E}\left[\sup_{f\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^n \varepsilon_i f(Z_i)\right] \quad \textit{and}$$
$$\mathbb{E}\left[\sup_{f\in\mathcal{F}}\left(\mathbb{E}[f(Z)] - \frac{1}{n}\sum_{i=1}^n f(Z_i)\right)\right] \leq 2\mathbb{E}\left[\sup_{f\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^n \varepsilon_i f(Z_i)\right].$$

Moreover, the result also holds for multivariate random variables $Z_1, \cdots, Z_n$ and Z.

Bounding $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$. In the following, we present the concentration bound for $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$ under the random spherical model specified in Definition 1.

Lemma 7. *Consider the random spherical model in Definition 1. For a fixed number*

$t > 0$, we have

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{X}}},\min} \geq \sqrt{\frac{2}{\pi d}}\rho(\sigma) - \left(2 + \frac{t}{2}\right)\frac{1}{\sqrt{N}}\right] \geq 1 - 2e^{-\frac{t^2}{2}} \tag{3.54}$$

where

$$\rho(\sigma) := (1-\sigma)F_{D-d,d}\left(\frac{1}{\sigma}\right) \tag{3.55}$$

with $F_{d_1,d_2}(\cdot)$ the cumulative density function (CDF) of F-distribution with $F_{d_1,d_2}(0) = 0$ and $F_{d_1,d_2}(\infty) = 1$. Moreover, we have $\rho(\sigma) = 1 - O(\sigma + \sigma^{\frac{d}{2}})$.

Proof. According to the generative model in Definition 1, let $\overline{\boldsymbol{x}}_1, \cdots, \overline{\boldsymbol{x}}_N \sim \mathcal{N}\left(\mathbf{0}, \frac{1}{d}\mathcal{P}_{\mathcal{S}}\right)$, and $\overline{\boldsymbol{\epsilon}}_1, \cdots, \overline{\boldsymbol{\epsilon}}_N \sim \mathcal{N}\left(\mathbf{0}, \frac{\sigma^2}{D}\mathbf{I}\right)$ so that the normalized noisy inliers $\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}}$ are given as

$$\begin{aligned} \boldsymbol{\mathcal{X}} &= \left[\frac{\overline{\boldsymbol{x}}_1}{\|\overline{\boldsymbol{x}}_1 + \overline{\boldsymbol{\epsilon}}_1\|_2}, \cdots, \frac{\overline{\boldsymbol{x}}_N}{\|\overline{\boldsymbol{x}}_N + \overline{\boldsymbol{\epsilon}}_N\|_2}\right] \quad \text{and} \\ \boldsymbol{\mathcal{E}} &= \left[\frac{\overline{\boldsymbol{\epsilon}}_1}{\|\overline{\boldsymbol{x}}_1 + \overline{\boldsymbol{\epsilon}}_1\|_2}, \cdots, \frac{\overline{\boldsymbol{\epsilon}}_N}{\|\overline{\boldsymbol{x}}_N + \overline{\boldsymbol{\epsilon}}_N\|_2}\right]. \end{aligned} \tag{3.56}$$

We denote

$$\begin{aligned} \boldsymbol{\mathcal{E}_s} &:= \left[\frac{\overline{\boldsymbol{\epsilon}}_1^{\boldsymbol{s}}}{\|\overline{\boldsymbol{x}}_1 + \overline{\boldsymbol{\epsilon}}_1\|_2}, \cdots, \frac{\overline{\boldsymbol{\epsilon}}_N^{\boldsymbol{s}}}{\|\overline{\boldsymbol{x}}_N + \overline{\boldsymbol{\epsilon}}_N\|_2}\right] \quad \text{and} \\ \boldsymbol{\mathcal{E}_n} &:= \left[\frac{\overline{\boldsymbol{\epsilon}}_1^{\boldsymbol{n}}}{\|\overline{\boldsymbol{x}}_1 + \overline{\boldsymbol{\epsilon}}_1\|_2}, \cdots, \frac{\overline{\boldsymbol{\epsilon}}_N^{\boldsymbol{n}}}{\|\overline{\boldsymbol{x}}_N + \overline{\boldsymbol{\epsilon}}_N\|_2}\right], \end{aligned} \tag{3.57}$$

where $\boldsymbol{\mathcal{E}_s} = \mathcal{P}_{\mathcal{S}}\boldsymbol{\mathcal{E}} \subset \mathcal{S}$ and $\boldsymbol{\mathcal{E}_n} = (\mathbf{I} - \mathcal{P}_{\mathcal{S}})\boldsymbol{\mathcal{E}} \subset \mathcal{S}^{\perp}$. By definition, $\widehat{\boldsymbol{\mathcal{X}}} := \boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}_s}$, and thus

$$
\begin{aligned}
c_{\widehat{\boldsymbol{\mathcal{X}}},\min} &= \min_{\boldsymbol{b}\in\mathcal{S}\cap\mathbb{S}^{D-1}} \frac{1}{N}\sum_{j=1}^{N}\left|\widehat{\boldsymbol{x}}_j^\top \boldsymbol{b}\right| \geq \min_{\|\boldsymbol{b}\|=1} \frac{1}{N}\sum_{j=1}^{N}\left|\widehat{\boldsymbol{x}}_j^\top \boldsymbol{b}\right| \\
&= \min_{\|\boldsymbol{b}\|=1} \frac{1}{N}\sum_{j=1}^{N}\left|\frac{(\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{s}})^\top \boldsymbol{b}}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2}\right| = \min_{\|\boldsymbol{b}\|=1} \frac{1}{N}\sum_{j=1}^{N}\frac{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{s}}\|_2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2}\left|\boldsymbol{v}_j^\top \boldsymbol{b}\right|,
\end{aligned}
$$

where $\boldsymbol{v}_j$ is the direction vector of $\widehat{\boldsymbol{x}}_j$ such that $\|\boldsymbol{v}_j\|_2 = 1$. Denote the scaling factor of $\widehat{\boldsymbol{x}}_j$ by

$$
R_j := \frac{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{s}}\|_2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2}, \tag{3.58}
$$

we are interested in $\mathbb{E}\left[R_j\right]$. Consider the following random variables

$$
\begin{aligned}
Y_j :=& R_j^2 = \frac{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{s}}\|_2^2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2^2} = \frac{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{s}}\|_2^2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{s}}\|_2^2 + \|\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}}\|_2^2} \quad \text{and} \\
Z_j :=& \frac{\|\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}}\|_2^2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{s}}\|_2^2} = \frac{\frac{\sigma^2}{D}\mathcal{X}_{D-d}^2}{(\frac{1}{d}+\frac{\sigma^2}{D})\mathcal{X}_d^2} = \frac{(D-d)\sigma^2}{D+d\sigma^2}\frac{\mathcal{X}_{D-d}^2/(D-d)}{\mathcal{X}_d^2/d} \\
\sim& \frac{(D-d)\sigma^2}{D+d\sigma^2}F(D-d,d)
\end{aligned}
$$

where $\mathcal{X}_d$ is the chi-square distribution with d degrees of freedom, and $F(d_1, d_2)$ is the F-distribution with parameters d_1 and d_2. Note that

$$
\frac{1}{Y_j} = 1 + Z_j
$$

and that, for any $a \in [0,1]$, we have

$$\begin{aligned}
\mathbb{P}\left[R_j \leq a\right] &= \mathbb{P}[\sqrt{Y_j} \leq a] = \mathbb{P}[Y_j \leq a^2] = \mathbb{P}\left[\frac{1}{Y_j} \geq \frac{1}{a^2}\right] = 1 - \mathbb{P}\left[\frac{1}{Y_j} < \frac{1}{a^2}\right] \\
&= 1 - \mathbb{P}\left[1 + Z_j < \frac{1}{a^2}\right] = 1 - \mathbb{P}\left[Z_j < \frac{1}{a^2} - 1\right] \\
&= 1 - \mathbb{P}\left[\frac{D+d\sigma^2}{(D-d)\sigma^2} Z_j < \frac{D+d\sigma^2}{(D-d)\sigma^2}\left(\frac{1}{a^2} - 1\right)\right] \\
&= 1 - \int_0^{\frac{D+d\sigma^2}{(D-d)\sigma^2}\left(\frac{1}{a^2}-1\right)} f_F(t; D-d, d)\,\mathrm{d}\,t
\end{aligned}$$

where f_F is the probability density function of F-distribution. Hence

$$\begin{aligned}
\mathbb{E}\left[R_j\right] &= \int_0^1 a f_{R_j}(a)\,\mathrm{d}\,a = \int_0^1 a\,\mathrm{d}\,F_{R_j}(a) = aF_{R_j}(a)\Big|_0^1 - \int_0^1 F_{R_j}(a)\,\mathrm{d}\,a \\
&= \int_0^1 \int_0^{\frac{D+d\sigma^2}{(D-d)\sigma^2}\left(\frac{1}{a^2}-1\right)} f_F(t; D-d, d)\,\mathrm{d}\,t\,\mathrm{d}\,a \\
&\geq \int_0^{1-\sigma} \int_0^{\frac{D+d\sigma^2}{(D-d)\sigma^2}\left(\frac{1}{(1-\sigma)^2}-1\right)} f_F(t; D-d, d)\,\mathrm{d}\,t\,\mathrm{d}\,a \\
&\geq \int_0^{1-\sigma} \int_0^{\frac{D+d\sigma^2}{(D-d)\sigma^2}(2\sigma)} f_F(t; D-d, d)\,\mathrm{d}\,t\,\mathrm{d}\,a \\
&= (1-\sigma) F_{D-d,d}\left(2\frac{D/\sigma + d\sigma}{D-d}\right),
\end{aligned} \tag{3.59}$$

where the second inequality follows from $1/(1-\sigma)^2 = \sum_{i=0}^{\infty}(i+1)\sigma^i$.

For any $\boldsymbol{b} \in \mathbb{S}^{D-1}$, we define the function $f_{\boldsymbol{b}} : \mathbb{S}^{D-1} \cap \mathcal{S} \times [0,1] \to \mathbb{R}$ by

$$f_{\boldsymbol{b}}(\boldsymbol{v}_j, R_j) = \frac{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{s}}\|_2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2}\left|\boldsymbol{b}^\top \boldsymbol{v}_j\right| = R_j\left|\boldsymbol{b}^\top \boldsymbol{v}_j\right|.$$

We let $\mu_{[0,1]}$ and $\mu_{\mathbb{S}^{D-1}\cap\mathcal{S}}$ denote the uniform measures on $[0,1]$ and $\mathbb{S}^{D-1} \cap \mathcal{S}$, respectively. Then, it follows that

$$\begin{aligned}\mathbb{E}\left[\left|\widehat{\boldsymbol{x}}_j^\top \boldsymbol{b}\right|\right] &= \int_{\boldsymbol{v}_j\in\mathbb{S}^{D-1}\cap\mathcal{S}}\int_0^1 f_{\boldsymbol{b}}(\boldsymbol{v}_j, R_j)\,\mathrm{d}\,\mu_{[0,1]}\,\mathrm{d}\,\mu_{\mathbb{S}^{D-1}\cap\mathcal{S}} \\ &= \int_{\boldsymbol{v}_j\in\mathbb{S}^{D-1}\cap\mathcal{S}}\int_0^1 R_j\left|\boldsymbol{b}^\top\boldsymbol{v}_j\right|\mathrm{d}\,\mu_{[0,1]}\,\mathrm{d}\,\mu_{\mathbb{S}^{D-1}\cap\mathcal{S}} \\ &= \int_0^1 R_j\,\mathrm{d}\,\mu_{[0,1]}\int_{\boldsymbol{v}_j\in\mathbb{S}^{D-1}\cap\mathcal{S}}\left|\boldsymbol{b}^\top\boldsymbol{v}_j\right|\mathrm{d}\,\mu_{\mathbb{S}^{D-1}\cap\mathcal{S}} \\ &= c_d\mathbb{E}\left[R_j\right] \\ &\geq \sqrt{\frac{2}{\pi d}}(1-\sigma)F_{D-d,d}\left(2\frac{D/\sigma+d\sigma}{D-d}\right)\end{aligned}$$

where c_d is the average height of the unit hemisphere of $\mathbb{R}^d$, the last equality follows from [112, Equation (59)], and the last inequality follows from (3.59) and $c_d \geq \sqrt{2/(\pi d)}$ (as shown in [152, Footnote 9]). Therefore, we obtain

$$\mathbb{E}_0 := \mathbb{E}\left[\left|\widehat{\boldsymbol{x}}_j^\top \boldsymbol{b}\right|\right] \geq \sqrt{\frac{2}{\pi d}}(1-\sigma)F_{D-d,d}\left(2\frac{D/\sigma+d\sigma}{D-d}\right),\ \forall j. \tag{3.60}$$

We are now ready to bound $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$. Note that

$$\begin{aligned}c_{\widehat{\boldsymbol{\mathcal{X}}},\min} &\geq \inf_{\|\boldsymbol{b}\|_2=1}\frac{1}{N}\sum_{j=1}^N\left|\widehat{\boldsymbol{x}}_j^\top\boldsymbol{b}\right| = \inf_{\|\boldsymbol{b}\|_2=1}\frac{1}{N}\sum_{j=1}^N\left|\widehat{\boldsymbol{x}}_j^\top\boldsymbol{b}\right| - \mathbb{E}_0 + \mathbb{E}_0 \\ &= \mathbb{E}_0 - \sup_{\|\boldsymbol{b}\|_2=1}\left(\mathbb{E}_0 - \frac{1}{N}\sum_{j=1}^N\left|\widehat{\boldsymbol{x}}_j^\top\boldsymbol{b}\right|\right).\end{aligned} \tag{3.61}$$

Since $\mathbb{S}^{D-1}$ is compact, there exists $\boldsymbol{b}^+ \in \mathbb{S}^{D-1}$ that achieves the supremum in (3.61).

Then for any $\widehat{\boldsymbol{x}}_1, \widehat{\boldsymbol{x}}_2, \cdots, \widehat{\boldsymbol{x}}_N, \widehat{\boldsymbol{x}}_k'$, we have

$$\begin{aligned}
&\left| \sup_{\|\boldsymbol{b}\|_2=1} \left(\mathbb{E}_0 - \frac{1}{N} \sum_{j=1}^N \left| \widehat{\boldsymbol{x}}_j^\top \boldsymbol{b} \right| \right) - \sup_{\|\boldsymbol{b}\|_2=1} \left(\mathbb{E}_0 - \frac{1}{N} \sum_{j\neq k} \left(\left| \widehat{\boldsymbol{x}}_j^\top \boldsymbol{b} \right| + \left| \widehat{\boldsymbol{x}}_k'^\top \boldsymbol{b} \right| \right) \right) \right| \\
\leq &\left| \mathbb{E}_0 - \frac{1}{N} \sum_{j=1}^N \left| \widehat{\boldsymbol{x}}_j^\top \boldsymbol{b}^+ \right| - \left(\mathbb{E}_0 - \frac{1}{N} \sum_{j\neq k} \left(\left| \widehat{\boldsymbol{x}}_j^\top \boldsymbol{b}^+ \right| + \left| \widehat{\boldsymbol{x}}_k'^\top \boldsymbol{b}^+ \right| \right) \right) \right| \\
= &\left| \frac{1}{N} \left(\left| \widehat{\boldsymbol{x}}_k'^\top \boldsymbol{b}^+ \right| - \left| \widehat{\boldsymbol{x}}_k^\top \boldsymbol{b}^+ \right| \right) \right| \leq \frac{1}{N}.
\end{aligned}$$

Applying Lemma 4 with $c_k = 1/N$, we have

$$\mathbb{P}\left[\left| \sup_{\|\boldsymbol{b}\|_2=1} \left(\mathbb{E}_0 - \frac{1}{N} \sum_{j=1}^N \left| \widehat{\boldsymbol{x}}_j^\top \boldsymbol{b} \right| \right) - \mathbb{E}\left[\sup_{\|\boldsymbol{b}\|_2=1} \left(\mathbb{E}_0 - \frac{1}{N} \sum_{j=1}^N \left| \widehat{\boldsymbol{x}}_j^\top \boldsymbol{b} \right| \right) \right] \right| \geq \epsilon \right] \leq 2e^{-2\epsilon^2 N}. \tag{3.62}$$

Moreover, we have

$$\begin{aligned}
&\mathbb{E}\left[\sup_{\|\boldsymbol{b}\|_2=1} \left(\mathbb{E}_0 - \frac{1}{N} \sum_{j=1}^N \left| \widehat{\boldsymbol{x}}_j^\top \boldsymbol{b} \right| \right) \right] \\
\leq &\, 2\mathbb{E}\left[\sup_{\|\boldsymbol{b}\|_2=1} \frac{1}{N} \sum_{j=1}^N \varepsilon_j \left| \widehat{\boldsymbol{x}}_j^\top \boldsymbol{b} \right| \right] \leq 2\mathbb{E}\left[\sup_{\|\boldsymbol{b}\|_2=1} \frac{1}{N} \sum_{j=1}^N \varepsilon_j \widehat{\boldsymbol{x}}_j^\top \boldsymbol{b} \right] \\
= &\, \frac{2}{N}\mathbb{E}\left[\sup_{\|\boldsymbol{b}\|_2=1} \left\langle \boldsymbol{b}, \sum_{j=1}^N \varepsilon_j \widehat{\boldsymbol{x}}_j \right\rangle \right] = \frac{2}{N}\mathbb{E}\left[\left\| \sum_{j=1}^N \varepsilon_j \widehat{\boldsymbol{x}}_j \right\|_2 \right] \\
\leq &\, \frac{2}{N}\sqrt{\mathbb{E}\left[\left\| \sum_{j=1}^N \varepsilon_j \widehat{\boldsymbol{x}}_j \right\|_2^2 \right]} \leq \frac{2}{N}\sqrt{\mathbb{E}\left[N + \sum_{i\neq j} \varepsilon_i \varepsilon_j \widehat{\boldsymbol{x}}_i^\top \widehat{\boldsymbol{x}}_j \right]} = \frac{2}{\sqrt{N}},
\end{aligned} \tag{3.63}$$

where the first inequality follows from Lemma 6, the second inequality follows from Lemma 5 by letting $\varphi_i(\cdot) = |\cdot|$, and the third inequality comes from Jensen's Inequality. Applying (3.63) to (3.62), we obtain

$$\mathbb{P}\left[\sup_{\|\boldsymbol{b}\|_2=1} \left(\mathbb{E}_0 - \frac{1}{N} \sum_{j=1}^N \left| \widehat{\boldsymbol{x}}_j^\top \boldsymbol{b} \right| \right) \geq \frac{2}{\sqrt{N}} + \epsilon \right] \leq 2e^{-2\epsilon^2 N}.$$

Therefore

$$\mathbb{P}\left[\mathbb{E}_0 - \sup_{\|\boldsymbol{b}\|_2=1}\left(\mathbb{E}_0 - \frac{1}{N}\sum_{j=1}^{N}\left|\widehat{\boldsymbol{x}}_j^\top \boldsymbol{b}\right|\right) \leq \mathbb{E}_0 - \frac{2}{\sqrt{N}} - \epsilon\right] \leq 2e^{-2\epsilon^2 N}.$$

From (3.61) we have

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{X}}},\min} \leq \mathbb{E}_0 - \frac{2}{\sqrt{N}} - \epsilon\right] \leq 2e^{-2\epsilon^2 N}.$$

Applying the lower bound for $\mathbb{E}_0$ in (3.60), we obtain

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{X}}},\min} \leq \sqrt{\frac{2}{\pi d}}(1-\sigma)F_{D-d,d}\left(2\frac{D/\sigma + d\sigma}{D-d}\right) - \frac{2}{\sqrt{N}} - \epsilon\right] \leq 2e^{-2\epsilon^2 N},$$

and by setting $\epsilon = \frac{t}{2\sqrt{N}}$, we have

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{X}}},\min} \leq \sqrt{\frac{2}{\pi d}}(1-\sigma)F_{D-d,d}\left(2\frac{D/\sigma + d\sigma}{D-d}\right) - \left(2 + \frac{t}{2}\right)\frac{1}{\sqrt{N}}\right] \leq 2e^{-\frac{t^2}{2}}.$$

From $\frac{1}{\sigma} \leq 2\frac{D/\sigma+d\sigma}{D-d}$ and the fact that all the CDFs are nondecreasing, we get

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{X}}},\min} \leq \sqrt{\frac{2}{\pi d}}(1-\sigma)F_{D-d,d}\left(\frac{1}{\sigma}\right) - \left(2 + \frac{t}{2}\right)\frac{1}{\sqrt{N}}\right] \leq 2e^{-\frac{t^2}{2}},$$

which completes the proof of (3.54).

Finally, by expanding the CDF formula of the F-distribution, we have

$$
\begin{aligned}
\rho(\sigma) &= (1-\sigma)F_{D-d,d}(1/\sigma) \\
&= (1-\sigma)\left(1 - \int_{1/\sigma}^{\infty} f_F(x; D-d, d)\,\mathrm{d}\,x\right) \\
&= 1-\sigma-(1-\sigma)\int_{1/\sigma}^{\infty} f_F(x; D-d, d)\,\mathrm{d}\,x \\
&\geq 1-\sigma-\int_{1/\sigma}^{\infty} f_F(x; D-d, d)\,\mathrm{d}\,x \\
&= 1-\sigma-\frac{1}{\mathrm{B}\left(\frac{D-d}{2},\frac{d}{2}\right)}\left(\frac{D-d}{d}\right)^{\frac{D-d}{2}}\int_{1/\sigma}^{\infty} x^{\frac{D-d}{2}-1}\left(1+\frac{D-d}{d}x\right)^{-\frac{D}{2}}\mathrm{d}\,x \\
&= 1-\sigma-\frac{1}{\mathrm{B}\left(\frac{D-d}{2},\frac{d}{2}\right)}\left(\frac{D-d}{d}\right)^{\frac{D-d}{2}}\int_{1/\sigma}^{\infty} x^{-\frac{d}{2}-1}\left[\frac{x}{1+\frac{D-d}{d}x}\right]^{\frac{D}{2}}\mathrm{d}\,x \\
&\geq 1-\sigma-\frac{1}{\mathrm{B}\left(\frac{D-d}{2},\frac{d}{2}\right)}\left(\frac{D-d}{d}\right)^{\frac{D-d}{2}}\int_{1/\sigma}^{\infty} x^{-\frac{d}{2}-1}\left(\frac{d}{D-d}\right)^{\frac{D}{2}}\mathrm{d}\,x \\
&= 1-\sigma-\frac{1}{\mathrm{B}\left(\frac{D-d}{2},\frac{d}{2}\right)}\left(\frac{d}{D-d}\right)^{\frac{d}{2}}\frac{2}{d}\sigma^{\frac{d}{2}} \\
&= 1-\sigma-\left[\frac{2/d}{\mathrm{B}\left(\frac{D-d}{2},\frac{d}{2}\right)}\left(\frac{d}{D-d}\right)^{\frac{d}{2}}\right]\cdot\sigma^{\frac{d}{2}},
\end{aligned}
$$

where $\mathrm{B}(\cdot,\cdot)$ is the Beta function. □

Bounding $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$. Next, we present the concentration bound for $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$ under the random spherical model specified in Definition 1.

Lemma 8. *Consider the random spherical model in Definition 1. For a fixed number $t > 0$, we have that*

$$
\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{E}}},\max} \leq \left(1+\frac{2}{\sqrt{N}}\right)\delta(\sigma)+\frac{t}{2\sqrt{N}}\right] \geq 1-2e^{-\frac{t^2}{2}} \tag{3.64}
$$

where

$$\delta(\sigma) := \sqrt{\sigma} + \sqrt{1-\sigma}\sqrt{F_{d,D-d}(\sigma)}, \tag{3.65}$$

and $F_{d_1,d_2}(\cdot)$ *is the cumulative density function (CDF) of the F-distribution with* $F_{d_1,d_2}(0) = 0$ *and* $F_{d_1,d_2}(\infty) = 1$*. Moreover, we have* $\delta(\sigma) = O(\sigma^{d/4} + \sigma^{1/2})$*.*

Proof. According to the generative model in Definition 1, let $\overline{\boldsymbol{x}}_1, \cdots, \overline{\boldsymbol{x}}_N \sim \mathcal{N}\left(\mathbf{0}, \frac{1}{d}\mathcal{P}_{\mathcal{S}}\right)$, and $\overline{\boldsymbol{\epsilon}}_1, \cdots, \overline{\boldsymbol{\epsilon}}_N \sim \mathcal{N}\left(\mathbf{0}, \frac{\sigma^2}{D}\mathbf{I}\right)$, so that from (3.56), (3.57) and $\widehat{\boldsymbol{\mathcal{E}}} = \boldsymbol{\mathcal{E}}_{\boldsymbol{n}}$, we have

$$\begin{aligned} c_{\widehat{\boldsymbol{\mathcal{E}}},\max} &= \max_{\boldsymbol{b}\in\mathcal{S}^\perp\cap\mathbb{S}^{D-1}} \frac{1}{N}\sum_{j=1}^N \left|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b}\right| \le \max_{\|\boldsymbol{b}\|=1} \frac{1}{N}\sum_{j=1}^N \left|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b}\right| \\ &= \max_{\|\boldsymbol{b}\|=1} \frac{1}{N}\sum_{j=1}^N \left|\frac{\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}\top}\boldsymbol{b}}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2}\right| = \max_{\|\boldsymbol{b}\|=1} \frac{1}{N}\sum_{j=1}^N \frac{\|\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}}\|_2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2}\left|\boldsymbol{v}_j^\top \boldsymbol{b}\right|, \end{aligned}$$

where $\boldsymbol{v}_j$ is the direction vector of $\widehat{\boldsymbol{\epsilon}}_j$ such that $\|\boldsymbol{v}_j\|_2 = 1$. Denote the scaling factor of $\widehat{\boldsymbol{\epsilon}}_j$ by

$$R_j := \frac{\|\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}}\|_2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2},$$

we are interested in $\mathbb{E}\left[R_j\right]$. Consider the following random variables

$$\begin{aligned} Y_j :=& R_j^2 = \frac{\|\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}}\|_2^2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2^2} = \frac{\|\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}}\|_2^2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{s}}\|_2^2 + \|\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}}\|_2^2} \quad \text{and} \\ Z_j :=& \frac{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{s}}\|_2^2}{\|\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}}\|_2^2} = \frac{(\frac{1}{d} + \frac{\sigma^2}{D})\mathcal{X}_d^2}{\frac{\sigma^2}{D}\mathcal{X}_{D-d}^2} = \frac{D + d\sigma^2}{(D-d)\sigma^2}\frac{\mathcal{X}_d^2/d}{\mathcal{X}_{D-d}^2/(D-d)} \\ \sim& \frac{D + d\sigma^2}{(D-d)\sigma^2}F(d, D-d), \end{aligned}$$

where $\mathcal{X}_d$ is the chi-square distribution with d degrees of freedom, and $F(d_1, d_2)$ is the F-distribution with parameters d_1 and d_2. Note that

$$\frac{1}{Y_j} = 1 + Z_j,$$

and that for any $a \in [0, 1]$, we have

$$\begin{aligned}
\mathbb{P}\left[R_j \leq a\right] = \mathbb{P}[\sqrt{Y_j} \leq a] = \mathbb{P}[Y_j \leq a^2] = \mathbb{P}\left[\frac{1}{Y_j} \geq \frac{1}{a^2}\right] &= 1 - \mathbb{P}\left[\frac{1}{Y_j} < \frac{1}{a^2}\right] \\
&= 1 - \mathbb{P}\left[1 + Z_j < \frac{1}{a^2}\right] = 1 - \mathbb{P}\left[Z_j < \frac{1}{a^2} - 1\right] \\
&= 1 - \mathbb{P}\left[\frac{(D-d)\sigma^2}{D + d\sigma^2} Z_j < \frac{(D-d)\sigma^2}{D + d\sigma^2}\left(\frac{1}{a^2} - 1\right)\right] \\
&= 1 - \int_0^{\frac{(D-d)\sigma^2}{D+d\sigma^2}\left(\frac{1}{a^2}-1\right)} f_F(t; d, D-d)\,\mathrm{d}\,t
\end{aligned}$$

where f_F is the probability density function of the F-distribution. Therefore,

$$\begin{aligned}
\mathbb{E}\left[R_j\right] &= \int_0^1 a f_{R_j}(a)\,\mathrm{d}\,a = \int_0^1 a\,\mathrm{d}\,F_{R_j}(a) = aF_{R_j}(a)\Big|_0^1 - \int_0^1 F_{R_j}(a)\,\mathrm{d}\,a \\
&= \int_0^1 \int_0^{\frac{(D-d)\sigma^2}{D+d\sigma^2}\left(\frac{1}{a^2}-1\right)} f_F(t; d, D-d)\,\mathrm{d}\,t\,\mathrm{d}\,a \\
&= \int_0^{\sqrt{\sigma}} \int_0^{\frac{(D-d)\sigma^2}{D+d\sigma^2}\left(\frac{1}{a^2}-1\right)} f_F(t; d, D-d)\,\mathrm{d}\,t\,\mathrm{d}\,a \\
&\qquad + \int_{\sqrt{\sigma}}^1 \int_0^{\frac{(D-d)\sigma^2}{D+d\sigma^2}\left(\frac{1}{a^2}-1\right)} f_F(t; d, D-d)\,\mathrm{d}\,t\,\mathrm{d}\,a \\
&\leq \int_0^{\sqrt{\sigma}} 1\,\mathrm{d}\,a + \int_{\sqrt{\sigma}}^1 \int_0^{\frac{(D-d)\sigma^2}{D+d\sigma^2}\left(\frac{1}{\sigma}\right)} f_F(t; d, D-d)\,\mathrm{d}\,t\,\mathrm{d}\,a \\
&= \sqrt{\sigma} + (1 - \sqrt{\sigma}) F_{d,D-d}\left(\frac{(D-d)\sigma}{D + d\sigma^2}\right) \leq \sqrt{\sigma} + \sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D + d\sigma^2}\right)}.
\end{aligned} \tag{3.66}$$

For any $\boldsymbol{b} \in \mathbb{S}^{D-1}$, we define the function $f_{\boldsymbol{b}} : \mathbb{S}^{D-1} \cap \mathcal{S}^{\perp} \times [0,1] \to \mathbb{R}$ by

$$f_{\boldsymbol{b}}(\boldsymbol{v}_j, R_j) = \frac{\|\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}}\|_2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2} \left|\boldsymbol{b}^{\top}\boldsymbol{v}_j\right| = R_j \left|\boldsymbol{b}^{\top}\boldsymbol{v}_j\right|.$$

Letting $\mu_{[0,1]}$ and $\mu_{\mathbb{S}^{D-1}\cap\mathcal{S}^{\perp}}$ denote the uniform measures on $[0,1]$ and $\mathbb{S}^{D-1} \cap \mathcal{S}^{\perp}$, respectively, it follows that

$$\begin{aligned}
\mathbb{E}\left[\left|\widehat{\boldsymbol{\epsilon}}_j^{\top}\boldsymbol{b}\right|\right] &= \int_{\boldsymbol{v}_j \in \mathbb{S}^{D-1}\cap\mathcal{S}^{\perp}} \int_0^1 f_{\boldsymbol{b}}(\boldsymbol{v}_j, R_j)\,\mathrm{d}\,\mu_{[0,1]}\,\mathrm{d}\,\mu_{\mathbb{S}^{D-1}\cap\mathcal{S}^{\perp}} \\
&= \int_{\boldsymbol{v}_j \in \mathbb{S}^{D-1}\cap\mathcal{S}^{\perp}} \int_0^1 R_j \left|\boldsymbol{b}^{\top}\boldsymbol{v}_j\right| \mathrm{d}\,\mu_{[0,1]}\,\mathrm{d}\,\mu_{\mathbb{S}^{D-1}\cap\mathcal{S}^{\perp}} \\
&= \int_0^1 R_j\,\mathrm{d}\,\mu_{[0,1]} \int_{\boldsymbol{v}_j \in \mathbb{S}^{D-1}\cap\mathcal{S}^{\perp}} \left|\boldsymbol{b}^{\top}\boldsymbol{v}_j\right| \mathrm{d}\,\mu_{\mathbb{S}^{D-1}\cap\mathcal{S}^{\perp}} \qquad (3.67) \\
&= c_{D-d}\mathbb{E}\left[R_j\right] \\
&\le \sqrt{\sigma} + \sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)},
\end{aligned}$$

where c_{D-d} is the average height of the unit hemisphere of $\mathbb{R}^{D-d}$, the last equality follows from [112, Equation (59)], and the last inequality follows from (3.66) and $c_{D-d} \le 1$ [112, Equation (23)]. Therefore, we obtain

$$\mathbb{E}_0 := \mathbb{E}\left[\left|\widehat{\boldsymbol{\epsilon}}_j^{\top}\boldsymbol{b}\right|\right] \le \sqrt{\sigma} + \sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)},\ \forall j. \qquad (3.68)$$

We are now ready to bound $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$. Note that

$$c_{\widehat{\boldsymbol{\mathcal{E}}},\max} \le \sup_{\|\boldsymbol{b}\|_2=1} \frac{1}{N}\sum_{j=1}^{N}\left|\widehat{\boldsymbol{\epsilon}}_j^{\top}\boldsymbol{b}\right| = \sup_{\|\boldsymbol{b}\|_2=1}\left(\frac{1}{N}\sum_{j=1}^{N}\left|\widehat{\boldsymbol{\epsilon}}_j^{\top}\boldsymbol{b}\right| - \mathbb{E}_0\right) + \mathbb{E}_0. \qquad (3.69)$$

Since $\mathbb{S}^{D-1}$ is compact, there exists $\boldsymbol{b}^+ \in \mathbb{S}^{D-1}$ that achieves the supremum in (3.69). Therefore, for any $\widehat{\boldsymbol{\epsilon}}_1, \widehat{\boldsymbol{\epsilon}}_2, \cdots, \widehat{\boldsymbol{\epsilon}}_N, \widehat{\boldsymbol{\epsilon}}_k'$, we have

$$\begin{aligned}
&\left| \sup_{\|\boldsymbol{b}\|_2=1} \left(\frac{1}{N} \sum_{j=1}^{N} \left| \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b} \right| - \mathbb{E}_0 \right) - \sup_{\|\boldsymbol{b}\|_2=1} \left(\frac{1}{N} \sum_{j \neq k} \left(\left| \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b} \right| + \left| \widehat{\boldsymbol{\epsilon}}_k'^\top \boldsymbol{b} \right| \right) - \mathbb{E}_0 \right) \right| \\
\leq &\left| \frac{1}{N} \sum_{j=1}^{N} \left| \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b}^+ \right| - \mathbb{E}_0 - \left(\frac{1}{N} \sum_{j \neq k} \left(\left| \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b}^+ \right| + \left| \widehat{\boldsymbol{\epsilon}}_k'^\top \boldsymbol{b}^+ \right| \right) - \mathbb{E}_0 \right) \right| \\
= &\left| \frac{1}{N} \left(\left| \widehat{\boldsymbol{\epsilon}}_k^\top \boldsymbol{b}^+ \right| - \left| \widehat{\boldsymbol{\epsilon}}_k'^\top \boldsymbol{b}^+ \right| \right) \right| \leq \frac{1}{N}.
\end{aligned} \tag{3.70}$$

Applying Lemma 4 with $c_k = \frac{1}{N}$, we have

$$\mathbb{P} \left[\left| \sup_{\|\boldsymbol{b}\|_2=1} \left(\frac{1}{N} \sum_{j=1}^{N} \left| \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b} \right| - \mathbb{E}_0 \right) - \mathbb{E} \left[\sup_{\|\boldsymbol{b}\|_2=1} \left(\frac{1}{N} \sum_{j=1}^{N} \left| \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b} \right| - \mathbb{E}_0 \right) \right] \right| \geq \epsilon \right] \leq 2e^{-2\epsilon^2 N}. \tag{3.71}$$

Moreover, we have

$$\begin{aligned}
&\mathbb{E} \left[\sup_{\|\boldsymbol{b}\|_2=1} \left(\frac{1}{N} \sum_{j=1}^{N} \left| \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b} \right| - \mathbb{E}_0 \right) \right] \\
&\leq 2\mathbb{E} \left[\sup_{\|\boldsymbol{b}\|_2=1} \frac{1}{N} \sum_{j=1}^{N} \varepsilon_j \left| \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b} \right| \right] \leq 2\mathbb{E} \left[\sup_{\|\boldsymbol{b}\|_2=1} \frac{1}{N} \sum_{j=1}^{N} \varepsilon_j \widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b} \right] \\
&= \frac{2}{N} \mathbb{E} \left[\sup_{\|\boldsymbol{b}\|_2=1} \left\langle \boldsymbol{b}, \sum_{j=1}^{N} \varepsilon_j \widehat{\boldsymbol{\epsilon}}_j \right\rangle \right] \leq \frac{2}{N} \mathbb{E} \left[\left\| \sum_{j=1}^{N} \varepsilon_j \widehat{\boldsymbol{\epsilon}}_j \right\|_2 \right] \\
&\leq \frac{2}{N} \sqrt{ \mathbb{E} \left[\left\| \sum_{j=1}^{N} \varepsilon_j \widehat{\boldsymbol{\epsilon}}_j \right\|_2^2 \right] } = \frac{2}{N} \sqrt{ \mathbb{E} \left[\sum_{j=1}^{N} \|\widehat{\boldsymbol{\epsilon}}_j\|_2^2 + \sum_{i \neq j} \varepsilon_i \varepsilon_j \widehat{\boldsymbol{\epsilon}}_i^\top \widehat{\boldsymbol{\epsilon}}_j \right] } \\
&\leq \frac{2}{\sqrt{N}} \sqrt{ \sigma + (1-\sigma) F_{d,D-d} \left(\frac{(D-d)\sigma}{D + d\sigma^2} \right) } \\
&\leq \frac{2}{\sqrt{N}} \left(\sqrt{\sigma} + \sqrt{1-\sigma} \sqrt{ F_{d,D-d} \left(\frac{(D-d)\sigma}{D + d\sigma^2} \right) } \right),
\end{aligned} \tag{3.72}$$

where the first inequality follows from Lemma 6, the second inequality follows from Lemma 5 by letting $\varphi_i(\cdot) = |\cdot|$, the fourth inequality comes from the Jensen's Inequality, and the fifth inequality follows from an upper bound for $\mathbb{E}[\|\widehat{\boldsymbol{\epsilon}}_j\|_2^2] = \mathbb{E}[R_j^2]$ that is similar to (3.66). Applying (3.72) to (3.71), we obtain

$$\mathbb{P}\left[\sup_{\|\boldsymbol{b}\|_2=1}\left(\frac{1}{N}\sum_{j=1}^{N}\left|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{b}\right| - \mathbb{E}_0\right) \geq \frac{2}{\sqrt{N}}\left(\sqrt{\sigma} + \sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)}\right) + \epsilon\right]$$
$$\leq 2e^{-2\epsilon^2 N}.$$

Therefore, from (3.69), we have

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{E}}},\max} \geq \mathbb{E}_0 + \frac{2}{\sqrt{N}}\left(\sqrt{\sigma} + \sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)}\right) + \epsilon\right] \leq 2e^{-2\epsilon^2 N}.$$

Applying the upper bound for $\mathbb{E}_0$ in (3.68), we obtain

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{E}}},\max} \geq \left(1 + \frac{2}{\sqrt{N}}\right)\left(\sqrt{\sigma} + \sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)}\right) + \epsilon\right] \leq 2e^{-2\epsilon^2 N},$$

and by setting $\epsilon = \frac{t}{2\sqrt{N}}$ we have

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{E}}},\max} \geq \left(1 + \frac{2}{\sqrt{N}}\right)\left(\sqrt{\sigma} + \sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)}\right) + \frac{t}{2\sqrt{N}}\right] \leq 2e^{-\frac{t^2}{2}}.$$

Note that $\frac{(D-d)\sigma}{D+d\sigma^2} \leq \sigma$ and all the CDFs are nondecreasing, we get

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{E}}},\max} \geq \left(1 + \frac{2}{\sqrt{N}}\right)\left(\sqrt{\sigma} + \sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\sigma\right)}\right) + \frac{t}{2\sqrt{N}}\right] \leq 2e^{-\frac{t^2}{2}},$$

which completes the proof of (3.64).

Finally, by expanding the CDF formula of the F-distribution, we have

$$\begin{aligned}
F_{d,D-d}(\sigma) &= \frac{1}{\mathrm{B}\left(\frac{d}{2},\frac{D-d}{2}\right)}\left(\frac{d}{D-d}\right)^{\frac{d}{2}}\int_0^\sigma x^{\frac{d}{2}-1}\left(1+\frac{d}{D-d}x\right)^{-\frac{D}{2}}\mathrm{d}\,x \\
&\leq \frac{1}{\mathrm{B}\left(\frac{d}{2},\frac{D-d}{2}\right)}\left(\frac{d}{D-d}\right)^{\frac{d}{2}}\int_0^\sigma x^{\frac{d}{2}-1}\mathrm{d}\,x \\
&= \frac{1}{\mathrm{B}\left(\frac{d}{2},\frac{D-d}{2}\right)}\left(\frac{d}{D-d}\right)^{\frac{d}{2}}\frac{2}{d}\cdot\sigma^{\frac{d}{2}}
\end{aligned}$$

where $\mathrm{B}(\cdot,\cdot)$ is the Beta function. Hence

$$\begin{aligned}
\delta(\sigma) &= \sqrt{\sigma}+\sqrt{(1-\sigma)F_{d,D-d}(\sigma)} \\
&\leq \sqrt{\sigma}+\sqrt{F_{d,D-d}(\sigma)} \\
&\leq \sqrt{\sigma}+\left[\frac{2/d}{\mathrm{B}\left(\frac{d}{2},\frac{D-d}{2}\right)}\left(\frac{d}{D-d}\right)^{\frac{d}{2}}\right]^{\frac{1}{2}}\cdot\sigma^{\frac{d}{4}}.
\end{aligned}$$

which completes the proof. □

Discussion of Lemma 7 and Lemma 8. First note that the concentration bound for $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$ reduces to the one for $c_{\boldsymbol{\mathcal{X}},\min}$ in (3.9) when $\boldsymbol{\mathcal{E}} = \mathbf{0}$ (or $\sigma = 0$). In particular, since $\rho(\sigma) = 1 - O(\sigma + \sigma^{\frac{d}{2}})$, $\rho(\sigma)$ tends to be large (close to 1) for small σ. Compared with $c_{\boldsymbol{\mathcal{X}},\min}$, one of the major challenges for deriving the concentration for $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$ in the noisy case is that under the random spherical model (Definition 1), the columns of $\widehat{\boldsymbol{\mathcal{X}}}$ now lie *inside* the unit sphere due to the effect of the additive noise, making it difficult to analyze their statistical behavior. On the other hand, since $\delta(\sigma) = O(\sigma^{\frac{d}{4}} + \sigma^{\frac{1}{2}})$, the concentration for $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$ in (3.64) essentially implies that $c_{\widehat{\boldsymbol{\mathcal{E}}},\max} = O(\sigma^{\frac{d}{4}} + \sigma^{\frac{1}{2}})$ with high probability. However, we remark that when

$\sigma = 0$, (3.64) does not immediately lead to $c_{\widehat{\boldsymbol{\mathcal{E}}},\max} = 0$ because of the existence of the additional small term $\frac{t}{2\sqrt{N}}$, which is an artifact of the proof technique used; we believe that the upper bound for $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$ can be improved to a quantity proportional to σ by a more sophisticated analysis.

We are now ready to give the probabilistic characterization of the global optimality for the noisy DPCP problem (3.1).

Theorem 4. *Consider the random spherical model of Definition 1. If* $0 < t < 2\left(\sqrt{\frac{2N}{\pi d}}\rho(\sigma) - 2\right)$*, then with probability at least* $1 - 8e^{-t^2/2}$*, any global solution to the noisy DPCP problem* (3.1) *must have its principal angle* θ^* *from* $\mathcal{S}^\perp$ *satisfy*

$$\sin(\theta^*) \leq \frac{C_1\delta(\sigma) + \frac{t}{2\sqrt{N}}}{\sqrt{\frac{2}{\pi d}}\rho(\sigma) - C_2\frac{t\sqrt{M}+\sqrt{DM}\log D}{N} - \frac{4+t}{\sqrt{N}}} \tag{3.73}$$

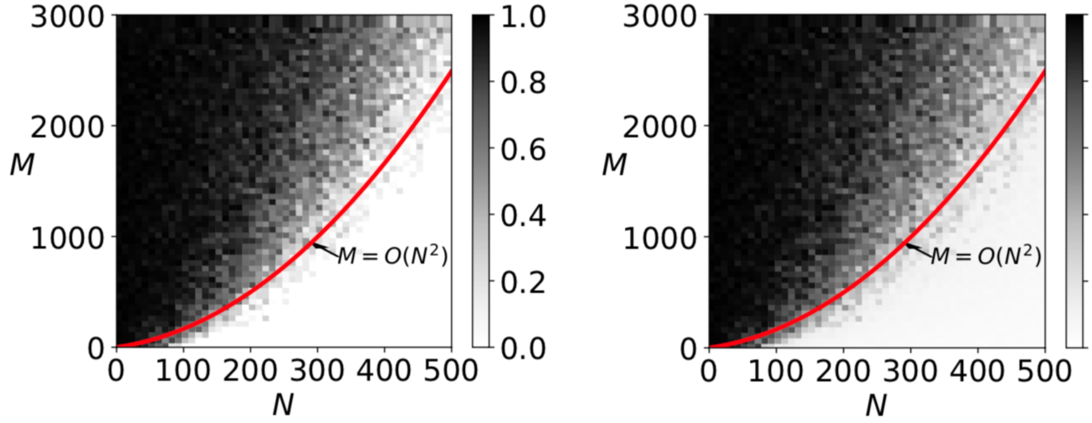
as long as

$$M\left((4\sqrt{2}+\sqrt{2}t)^2 + C_3(\sqrt{D}\log D + t)^2\right) \leq N^2\left(\frac{1}{\sqrt{\pi d}}\rho(\sigma) - C_4\delta(\sigma) - \frac{4+3t}{2\sqrt{2N}}\right)^2 \tag{3.74}$$

where C_1, C_2, C_3, C_4 *are universal constants that are independent of* N, M, D, d, t, σ.

Proof. Theorem 4 follows directly from Corollary 1 by plugging the concentration bounds for $c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min}$ and $\eta_{\boldsymbol{\mathcal{O}}}$ from (3.9), $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$ from (3.54), and $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$ from (3.64) into (3.50) and (3.51). □

Discussion of Theorem 4. The effect of the noise in perturbing the global solution away from $\mathcal{S}^\perp$ is captured by (3.73), where the RHS approaches 0 when



(a) $\sigma = 0$

(b) $\sigma = 0.1$

Figure 3.3. Plot of $\sin(\theta^*)$ where θ^* is the principal angle between the computed solution $\boldsymbol{b}^*$ to the noisy DPCP problem (3.1) and $\mathcal{S}^{\perp}$ when varying N and M for (a) noise level $\sigma = 0$ and (b) noise level $\sigma = 0.1$. Here $D = 30$ and $d = 29$.

$\sigma \to 0$, except for the small term $\frac{t}{2\sqrt{N}}$, which we commented earlier is (we believe) due to the proof technique used. Moreover, (3.73) together with $\delta(\sigma) = O(\sigma^{d/4} + \sigma^{\frac{1}{2}})$ and $\rho(\sigma) = 1 - O(\sigma + \sigma^{d/2})$ imply that $\sin(\theta^*) = O((\sigma^{d/4}) + \sigma^{\frac{1}{2}})$ when σ is small. The inequality (3.74) suggests that, unlike existing state-of-the-art $O(N)$ outlier bounds as reviewed in [62], DPCP can tolerate $O(N^2)$ outliers even for noisy data. Figure 3.3 verifies this point by plotting $\sin(\theta^*)$.

3.2 Learning a subspace with codimension larger than one

So far, all of the analyses of DPCP for learning a single subspace have been restricted to finding a normal vector to a maximal *hyperplane* that contains the inliers by solving (3.1), regardless of whether the data is contaminated by noise or not. Although this approach can be extended to a *subspace* of higher codimension through a recursive approach that sequentially finds a new basis element of the space orthogonal to the subspace, the procedure is computationally expensive and lacks theoretical support.

In this section, we consider a more powerful approach for learning a d-dimensional subspace $\mathcal{S}$ in $\mathbb{R}^D$ with codimension $c = D - d$ larger than 1 by *simultaneously* estimating the entire basis of the orthogonal complement subspace $\mathcal{S}^\perp$. We term this as a *holistic approach*, which is stated in (2.9), and for convenience we repeat here:

$$\min_{\boldsymbol{B}\in\mathbb{R}^{D\times c}} \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}\right\|_{1,2} = \sum_{j=1}^{L} \left\|\widetilde{\boldsymbol{x}}_j^\top \boldsymbol{B}\right\|_2 \quad \text{s.t.} \quad \boldsymbol{B}^\top \boldsymbol{B} = \mathbf{I} \tag{3.75}$$

where $\widetilde{\boldsymbol{\mathcal{X}}} = [\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}}, \boldsymbol{\mathcal{O}}]\boldsymbol{\Gamma}$ is the dataset that has the same form as in the previous section. Intuitively, in the noiseless case ($\boldsymbol{\mathcal{E}} = \mathbf{0}$), if $\boldsymbol{B}$ is an orthonormal basis of $\mathcal{S}^\perp$, then the objective in (3.75) only depends on the outliers and is insensitive to the choice of $\boldsymbol{B}$ since outliers are unstructured, which motivates the formulation. Although it naturally extends the original DPCP problem (3.1) by seeking a matrix $\boldsymbol{B}$ with orthonormal columns that are orthogonal to as many data points as possible, its theoretical guarantees for recovering an orthonormal basis of $\mathcal{S}^\perp$ under both noiseless and noisy settings remain open questions.

3.2.1 Background

Towards analyzing the holistic DPCP problem (3.75), we first introduce some background knowledge. Observe that (3.75) is an optimization problem on the Grassmannian $\mathbb{G}(D, c)$ [34], i.e., the set of c-dimensional subspaces in $\mathbb{R}^D$, we parameterize $\mathbb{G}(D, c)$ with orthonormal matrices in the set $\mathbb{O}(D, c) := \{\boldsymbol{B} \in \mathbb{R}^{D\times c} : \boldsymbol{B}^\top \boldsymbol{B} = \mathbf{I}\}$. In particular, when $c = 1$, we also use $\mathbb{S}^{D-1}$, i.e., the unit sphere, as a substitute for $\mathbb{O}(D, 1)$, for which the problem reduces to (3.1). In addition, we denote $\mathbb{O}(c, c)$ by $\mathbb{O}(c)$ for simplicity. Let $\boldsymbol{S}^\perp \in \mathbb{O}(D, c)$ be an orthonormal basis of $\mathcal{S}^\perp$. Since the objective function in (3.75) is rotational invariant, we consider equivalence classes of matrices. In particular, for $\boldsymbol{U}, \boldsymbol{V} \in \mathbb{G}(D, c)$ we say $\boldsymbol{U}$ is equivalent to $\boldsymbol{V}$ if $\operatorname{Span}(\boldsymbol{U}) = \operatorname{Span}(\boldsymbol{V})$, and use $\boldsymbol{U}$ to represent the equivalence class $[\boldsymbol{U}] := \{\boldsymbol{U}\boldsymbol{R} : \boldsymbol{R} \in \mathbb{O}(c)\}$.

As the dataset is contaminated with noise, a solution $\boldsymbol{B}^*$ to (3.75) is expected to be perturbed away from $\boldsymbol{S}^\perp$, which can be measured geometrically by the *principal angles* between two subspaces, which we now define.

Definition 2 ([57])**.** *Let $\boldsymbol{U}, \boldsymbol{V} \in \mathbb{R}^{D\times c}$ be orthonormal matrices. The principal angles between* $\operatorname{Span}(\boldsymbol{U})$ *and* $\operatorname{Span}(\boldsymbol{V})$ *are defined as*

$$\theta_i(\boldsymbol{U}, \boldsymbol{V}) = \arccos\Big(\sigma_i(\boldsymbol{U}^\top \boldsymbol{V})\Big) \tag{3.76}$$

for all $i \in \{1, 2, \dots, c\}$, where $\sigma_i(\cdot)$ denotes the i-th largest singular value. The largest principal angle $\theta_c(\boldsymbol{U}, \boldsymbol{V})$ defines the subspace angle *between* $\operatorname{Span}(\boldsymbol{U})$ *and* $\operatorname{Span}(\boldsymbol{V})$*.*

With Definition 2, we can compute how close $\operatorname{Span}(\boldsymbol{B}^*)$ and $\operatorname{Span}(\boldsymbol{S}^\perp) = \mathcal{S}^\perp$ are to one another. In particular, when $\operatorname{Span}(\boldsymbol{B}^*) = \mathcal{S}^\perp$, we have $\theta_1(\boldsymbol{B}^*, \boldsymbol{S}^\perp) = \cdots = \theta_c(\boldsymbol{B}^*, \boldsymbol{S}^\perp) = 0$ so that their subspace angle is zero, thus justifying the definition.

Since the objective in (3.75) involves the sum of ℓ_2 norms, with a mild abuse of notation on the subdifferential of the absolute value function defined in (3.3), we denote the subdifferential of $\|\boldsymbol{a}\|_2$ for any $\boldsymbol{a} \in \mathbb{R}^c$ by

$$\operatorname{Sgn}(\boldsymbol{a}) = \begin{cases} \{\boldsymbol{a}/\|\boldsymbol{a}\|_2\}, & \boldsymbol{a} \neq \boldsymbol{0}, \\ \{\boldsymbol{d} \in \mathbb{R}^c : \|\boldsymbol{d}\| \leq 1\}, & \boldsymbol{a} = \boldsymbol{0}. \end{cases} \tag{3.77}$$

Within this context, an element of the set $\operatorname{Sgn}(\boldsymbol{a})$ of particular interest will be

$$\operatorname{sign}(\boldsymbol{a}) = \begin{cases} \boldsymbol{a}/\|\boldsymbol{a}\|_2, & \boldsymbol{a} \neq \boldsymbol{0}, \\ \boldsymbol{0}, & \boldsymbol{a} = \boldsymbol{0}. \end{cases} \tag{3.78}$$

In this section, unless stated otherwise, $\operatorname{Sgn}(\boldsymbol{a})$ and $\operatorname{sign}(\boldsymbol{a})$ refer to the above generalized definitions for analyzing the holistic DPCP problem (3.75). Finally, we can write the subdifferential of $\left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}\right\|_{1,2}$ at $\boldsymbol{B}$ as

$$\begin{aligned} \partial \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}\right\|_{1,2} &= \sum_{j=1}^{L} \widetilde{\boldsymbol{x}}_j \operatorname{Sgn}\left(\widetilde{\boldsymbol{x}}_j^\top \boldsymbol{B}\right) \\ &= \sum_{j=1}^{N} (\boldsymbol{x}_j + \boldsymbol{\epsilon}_j) \operatorname{Sgn}\left((\boldsymbol{x}_j + \boldsymbol{\epsilon}_j)^\top \boldsymbol{B}\right) + \sum_{j=1}^{M} \boldsymbol{o}_j \operatorname{Sgn}\left(\boldsymbol{o}_j^\top \boldsymbol{B}\right). \end{aligned}$$

3.2.2 Analysis with noiseless data

We first analyze the holistic DPCP problem (3.75) in the noiseless setting where $\boldsymbol{\mathcal{E}} = \boldsymbol{0}$. We consider the same random spherical model (see Definition 1) for the underlying dataset as in analyzing (3.1) since the problem formulation is the only difference.

Geometric quantities. For inliers, we adopt the same permeance statistic $c_{\boldsymbol{\mathcal{X}},\min}$ in (3.4) from [152, 153]. For outliers, we extend the $\eta_{\boldsymbol{\mathcal{O}}}$ quantity in (3.6) for codimension $c = 1$, to the more general case of $c \geq 1$ by defining

$$\eta_{\mathcal{O},c} := \frac{1}{M} \max_{\boldsymbol{B} \in \mathbb{O}(D,c)} \left\| (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top) \sum_{j=1}^{M} \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B}) \right\|_F \tag{3.79}$$

which is the maximum norm of a Riemannian subgradient of $\frac{1}{M}\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}\|_{1,2}$. As an analogy to $\eta_{\mathcal{O}}$, the $\eta_{\mathcal{O},c}$ characterizes how well the outliers are distributed in the ambient space, with more uniformly distributed outliers leading to smaller $\eta_{\mathcal{O},c}$. We remark that $\eta_{\mathcal{O},c} \equiv \eta_{\mathcal{O}}$ when $c = 1$. Besides $\eta_{\mathcal{O},c}$, we also use another two quantities to describe the distribution of outliers, namely, we extend the definitions for $c_{\mathcal{O},\min}$ and $c_{\mathcal{O},\max}$ in (3.5) for $c = 1$ to the following:

$$\begin{aligned} c_{\mathcal{O},\min,c} &:= \frac{1}{M} \min_{\boldsymbol{B} \in \mathbb{O}(D,c)} \sum_{j=1}^{M} \|\boldsymbol{o}_j^\top \boldsymbol{B}\|_2 \quad \text{and} \\ c_{\mathcal{O},\max,c} &:= \frac{1}{M} \max_{\boldsymbol{B} \in \mathbb{O}(D,c)} \sum_{j=1}^{M} \|\boldsymbol{o}_j^\top \boldsymbol{B}\|_2. \end{aligned} \tag{3.80}$$

Well-distributed outliers lead to larger values for $c_{\mathcal{O},\min,c}$ and smaller values for $c_{\mathcal{O},\max,c}$, and a small gap between $c_{\mathcal{O},\max,c}$ and $c_{\mathcal{O},\min,c}$.

3.2.2.1 Geometry of the critical points

Using the above geometric quantities, we have the following lemma, which characterizes the geometry of the critical points of (3.75) in a deterministic sense.

Lemma 9. *Suppose $\boldsymbol{\mathcal{E}} = \mathbf{0}$. Then, any critical point $\boldsymbol{B}$ of problem* (3.75) *must either be an orthonormal basis for $\mathcal{S}^\perp$, or span a subspace that has an angle θ from $\mathcal{S}^\perp$ larger* *than or equal to* $\arccos(M\overline{\eta}_{\boldsymbol{\mathcal{O}},\mathrm{c}}/Nc_{\boldsymbol{\mathcal{X}},\min})$ *where*

$$\overline{\eta}_{\boldsymbol{\mathcal{O}},\mathrm{c}} := \eta_{\boldsymbol{\mathcal{O}},\mathrm{c}} + \frac{D}{M}.$$

Proof. As the first step of the proof, we prove the following useful result.

Sublemma 2. *Suppose* $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n \backslash \{\boldsymbol{0}\}$*,* $v_i \geq 0, \forall i$*, and* $v_n \geq v_i, \forall i \neq n$*. Then*

$$\frac{\langle \boldsymbol{u}, \boldsymbol{u}^\top \operatorname{diag}(\boldsymbol{v})\rangle}{\|\boldsymbol{u}^\top \operatorname{diag}(\boldsymbol{v})\|} \geq |u_n|.$$

Proof. We are going to prove that

$$\frac{\langle \boldsymbol{u}^\top \operatorname{diag}(\boldsymbol{v}), \boldsymbol{u}\rangle}{\|\boldsymbol{u}^\top \operatorname{diag}(\boldsymbol{v})\|} = \frac{\sum_{i=1}^n u_i^2 v_i}{\sqrt{\sum_{i=1}^n u_i^2 v_i^2}} \geq |u_n|.$$

Squaring both sides and then rearranging yields

$$\left(\sum_{i=1}^{n-1} u_i^2 v_i + u_n^2 v_n\right)^2 \geq u_n^2 \left(\sum_{i=1}^{n-1} u_i^2 v_i^2 + u_n^2 v_n^2\right),$$

which is equivalent to

$$\left(\sum_{i=1}^{n-1} u_i^2 v_i\right)^2 + 2u_n^2 v_n \sum_{i=1}^{n-1} u_i^2 v_i \geq u_n^2 \sum_{i=1}^{n-1} u_i^2 v_i^2. \tag{3.81}$$

Since $v_i \geq 0, \forall i$ and $v_n \geq v_i, \forall i \neq n$, we always have $u_n^2 v_n \sum_{i=1}^{n-1} u_i^2 v_i \geq u_n^2 \sum_{i=1}^{n-1} u_i^2 v_i^2$, and thus (3.81) always holds, which completes the proof. □

We now proceed by proving that any critical point $\boldsymbol{B}$ that is not an orthonormal basis for $\mathcal{S}^{\perp}$ must span a subspace that is far from $\mathcal{S}^{\perp}$. Let $\boldsymbol{S} \in \mathbb{R}^{D\times d}$ be an orthonormal basis of the subspace $\mathcal{S}$ and let $\boldsymbol{S}^{\perp} \in \mathbb{R}^{D\times c}$ be an orthonormal basis of the orthogonal complement $\mathcal{S}^{\perp}$. We rewrite $\boldsymbol{B}$ as

$$\boldsymbol{B} = \boldsymbol{S}\boldsymbol{S}^{\top}\boldsymbol{B} + \boldsymbol{S}^{\perp}(\boldsymbol{S}^{\perp})^{\top}\boldsymbol{B}, \tag{3.82}$$

where $\boldsymbol{S}\boldsymbol{S}^{\top}\boldsymbol{B}$ represents the projection of $\boldsymbol{B}$ onto the subspace $\mathcal{S}$, while the other term $\boldsymbol{S}^{\perp}(\boldsymbol{S}^{\perp})^{\top}\boldsymbol{B}$ represents the projection of $\boldsymbol{B}$ onto the complement $\mathcal{S}^{\perp}$. Let $(\boldsymbol{S}^{\perp})^{\top}\boldsymbol{B} = \boldsymbol{U}\cos(\boldsymbol{\Theta})\boldsymbol{R}^{\top}$ be the canonical SVD of $(\boldsymbol{S}^{\perp})^{\top}\boldsymbol{B}$, where $\cos(\boldsymbol{\Theta})$ is a diagonal matrix with $\cos(\theta_1), \ldots, \cos(\theta_c)$ along its diagonal, $\boldsymbol{U} \in \mathbb{R}^{c\times c}, \boldsymbol{R} \in \mathbb{R}^{c\times c}$ are orthonormal matrices. Here θ_i is the i-th principal angle between $\operatorname{Span}(\boldsymbol{B})$ and $\mathcal{S}^{\perp}$. When $\theta_1 = \cdots = \theta_c = 0$, it implies that $\boldsymbol{B} \in [\boldsymbol{S}^{\perp}]$, i.e., $\boldsymbol{B}$ is equivalent to $\boldsymbol{S}^{\perp}$. Since we assume that $\boldsymbol{B}$ is not orthogonal to $\mathcal{S}$, we always have $\theta_c > 0$ (recall that $\theta_c \equiv \theta_{\max}(\boldsymbol{B}, \boldsymbol{S}^{\perp})$).

Next, we will prove Lemma 9 when $c \leq d$, and the case $c > d$ can be proved in a similar way. If $c \leq d$, we rewrite $\boldsymbol{S}^{\top}\boldsymbol{B} = \boldsymbol{V}\sin(\boldsymbol{\Theta})\boldsymbol{R}^{\top}$, where $\boldsymbol{V} \in \mathbb{R}^{d\times c}$ is an orthonormal matrix. Thus, we have

$$\boldsymbol{B} = \boldsymbol{S}\boldsymbol{V}\sin(\boldsymbol{\Theta})\boldsymbol{R}^{\top} + \boldsymbol{S}^{\perp}\boldsymbol{U}\cos(\boldsymbol{\Theta})\boldsymbol{R}^{\top}. \tag{3.83}$$

Without loss of generality, we consider $\boldsymbol{R} = \mathbf{I}$ since the objective function of (3.75) is rotation invariant. Letting $\boldsymbol{P} = \boldsymbol{S}\boldsymbol{V}$ and $\boldsymbol{Q} = \boldsymbol{S}^{\perp}\boldsymbol{U}$, we have

$$\boldsymbol{B} = \boldsymbol{P}\sin(\boldsymbol{\Theta}) + \boldsymbol{Q}\cos(\boldsymbol{\Theta}), \tag{3.84}$$

where $\boldsymbol{P} \in \mathbb{R}^{D\times c}$ and $\boldsymbol{Q} \in \mathbb{R}^{D\times c}$ are orthonormal matrices satisfying $\mathrm{Span}(\boldsymbol{P}) \subseteq \mathcal{S}$ and $\mathrm{Span}(\boldsymbol{Q}) \subseteq \mathcal{S}^\perp$. As a result, $\boldsymbol{P}$ is orthogonal to $\boldsymbol{Q}$ and $\boldsymbol{B}$ is orthonormal. Next, we define

$$\boldsymbol{G} = \boldsymbol{P}\cos(\boldsymbol{\Theta}) - \boldsymbol{Q}\sin(\boldsymbol{\Theta}) \tag{3.85}$$

so that $\boldsymbol{G}$ is an orthonormal matrix and orthogonal to $\boldsymbol{B}$.

Let $f(\boldsymbol{B}) := \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}\right\|_{1,2}$. For any critical point $\boldsymbol{B}$ of problem (3.75), there exists $\boldsymbol{W} \in \partial f(\boldsymbol{B})$ such that $(\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\boldsymbol{W} = \mathbf{0}$. Due to the general position [152, 153] of the data and $\boldsymbol{B} \notin [\boldsymbol{S}^\perp]$, $\boldsymbol{B}$ can be orthogonal to $K \leq D - c$ columns of $\widetilde{\boldsymbol{\mathcal{X}}}$, and

$$\begin{aligned}\mathbf{0} &= (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\boldsymbol{W}\\ &= (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\left(\sum_{j=1}^N \boldsymbol{x}_j \operatorname{sign}\left(\boldsymbol{x}_j^\top \boldsymbol{B}\right) + \sum_{j=1}^M \boldsymbol{o}_j \operatorname{sign}\left(\boldsymbol{o}_j^\top \boldsymbol{B}\right) + \boldsymbol{\xi}\right),\end{aligned}$$

where $\boldsymbol{\xi} = \sum_{k=1}^K \widetilde{\boldsymbol{x}}_{j_k}\boldsymbol{\alpha}_{j_k}$ with $\widetilde{\boldsymbol{x}}_{j_1}, \cdots, \widetilde{\boldsymbol{x}}_{j_K}$ the columns of $\widetilde{\boldsymbol{\mathcal{X}}}$ orthogonal to $\boldsymbol{B}$, and $\{\|\boldsymbol{\alpha}_{j_1}\|, \cdots, \|\boldsymbol{\alpha}_{j_K}\|\} \in [-1, 1]$. We then have

$$\begin{aligned}0 &= \left|\left\langle (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\boldsymbol{W}, \boldsymbol{G}\right\rangle\right| = \left|\left\langle \boldsymbol{W}, (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\boldsymbol{G}\right\rangle\right| = |\langle \boldsymbol{W}, \boldsymbol{G}\rangle|\\ &= \left|\left\langle \sum_{j=1}^N \boldsymbol{x}_j \operatorname{sign}(\boldsymbol{x}_j^\top \boldsymbol{B}), \boldsymbol{G}\right\rangle + \left\langle \sum_{j=1}^M \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B}), \boldsymbol{G}\right\rangle + \langle \boldsymbol{\xi}, \boldsymbol{G}\rangle\right|\\ &= \left|\sum_{j=1}^N \left\langle \boldsymbol{x}_j^\top \boldsymbol{G}, \operatorname{sign}(\boldsymbol{x}_j^\top \boldsymbol{B})\right\rangle + \left\langle (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\sum_{j=1}^M \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B}), \boldsymbol{G}\right\rangle + \langle \boldsymbol{\xi}, \boldsymbol{G}\rangle\right|\\ &\geq \left|\sum_{j=1}^N \left\langle \boldsymbol{x}_j^\top \boldsymbol{G}, \operatorname{sign}(\boldsymbol{x}_j^\top \boldsymbol{B})\right\rangle\right| - \left|\left\langle (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\sum_{j=1}^M \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B}), \boldsymbol{G}\right\rangle\right| - |\langle \boldsymbol{\xi}, \boldsymbol{G}\rangle|.\end{aligned} \tag{3.86}$$

The first term in (3.86) can be written as

$$\begin{aligned}
\left|\sum_{j=1}^{N}\left\langle \boldsymbol{x}_j^\top \boldsymbol{G}, \operatorname{sign}(\boldsymbol{x}_j^\top \boldsymbol{B})\right\rangle\right| &= \left|\sum_{j=1}^{N}\left\langle \boldsymbol{x}_j^\top \boldsymbol{P}\cos(\boldsymbol{\Theta}), \operatorname{sign}(\boldsymbol{x}_j^\top \boldsymbol{P}\sin(\boldsymbol{\Theta}))\right\rangle\right| \\
&= \left|\sum_{j=1}^{N}\frac{\left\langle \boldsymbol{x}_j^\top \boldsymbol{P}\cos(\boldsymbol{\Theta}), \boldsymbol{x}_j^\top \boldsymbol{P}\sin(\boldsymbol{\Theta})\right\rangle}{\|\boldsymbol{x}_j^\top \boldsymbol{P}\sin(\boldsymbol{\Theta})\|}\right| \\
&\geq \cos(\theta_c)\sum_{j=1}^{N}\frac{\left\langle \boldsymbol{x}_j^\top \boldsymbol{P}, \boldsymbol{x}_j^\top \boldsymbol{P}\sin(\boldsymbol{\Theta})\right\rangle}{\|\boldsymbol{x}_j^\top \boldsymbol{P}\sin(\boldsymbol{\Theta})\|} \\
&\geq \cos(\theta_c)\sum_{j=1}^{N}|\boldsymbol{x}_j^\top \boldsymbol{p}_c| \geq \cos(\theta_c)Nc_{\boldsymbol{\mathcal{X}},\min},
\end{aligned} \tag{3.87}$$

where the first inequality utilizes the fact that $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_c$, and the second inequality follows from Sublemma 2 where $\boldsymbol{p}_c$ is the cth column of $\boldsymbol{P}$. Plugging this result into (3.86), and using the definition of $\eta_{\boldsymbol{\mathcal{O}},c}$, we have

$$0 \geq \cos(\theta_c)Nc_{\boldsymbol{\mathcal{X}},\min} - M\eta_{\boldsymbol{\mathcal{O}},c} - D.$$

This tells us that if $\boldsymbol{B} \notin [\boldsymbol{S}^\perp]$, then it is far from $[\boldsymbol{S}^\perp]$ in the sense that the largest principal angle $\theta_c \equiv \theta_{\max}(\boldsymbol{B}, \boldsymbol{S}^\perp)$ satisfies

$$\cos(\theta_c) \leq \frac{M\overline{\eta}_{\boldsymbol{\mathcal{O}},c}}{Nc_{\boldsymbol{\mathcal{X}},\min}},$$

thus completing the proof when $c \leq d$.

On the other hand, if $c > d$, there are only d principal angles between the subspaces spanned by $\boldsymbol{S} \in \mathbb{R}^{D\times d}$ and $\boldsymbol{B} \in \mathbb{R}^{D\times c}$. Since $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_c$ are the principle angles between $\operatorname{Span}(\boldsymbol{S}^\perp)$ and $\operatorname{Span}(\boldsymbol{B})$, according to [57], the principal angles between $\operatorname{Span}(\boldsymbol{S})$ and $\operatorname{Span}(\boldsymbol{B})$ are $\frac{\pi}{2} - \theta_c, \cdots, \frac{\pi}{2} - \theta_{c-d+1}$. Similar to the case of $c \leq d$, we rewrite $\boldsymbol{S}^\top \boldsymbol{B} = \boldsymbol{V}\sin(\boldsymbol{\Theta})\boldsymbol{R}^\top$, where $\boldsymbol{V} = \begin{bmatrix}\boldsymbol{0} \ \overline{\boldsymbol{V}}\end{bmatrix}$ with $\overline{\boldsymbol{V}} \in \mathbb{R}^{d\times d}$ an orthonormal matrix. Thus again, we have

$$\boldsymbol{B} = \boldsymbol{S}\boldsymbol{V}\sin(\boldsymbol{\Theta})\boldsymbol{R}^\top + \boldsymbol{S}^\perp \boldsymbol{U}\cos(\boldsymbol{\Theta})\boldsymbol{R}^\top. \tag{3.88}$$

The rest of the proof is now the same as before after one replaces $\boldsymbol{V}$ by $\overline{\boldsymbol{V}}$ and $\boldsymbol{\Theta}$ by $\overline{\boldsymbol{\Theta}} = \operatorname{diag}(\theta_{c-d+1}, \cdots, \theta_c)$. This completes the proof. □

Discussion of Lemma 9. Lemma 9 generalizes the special case $c = 1$ in Lemma 1. It says that, with noiseless data, any critical point of the holistic DPCP problem (3.75) either spans $\mathcal{S}^\perp$ or spans a subspace that is far from $\mathcal{S}^\perp$. Note that for well-distributed inliers and outliers (M/N and c fixed), the geometric location of $\boldsymbol{B}$ becomes more restricted. Moreover, any critical point $\boldsymbol{B}$ such that $\operatorname{Span}(\boldsymbol{B})$ is sufficiently close to $\mathcal{S}^\perp$ (angle smaller than $\arccos(M\overline{\eta}_{\boldsymbol{\mathcal{O}},c}/Nc_{\boldsymbol{\mathcal{X}},\min})$) must satisfy $\operatorname{Span}(\boldsymbol{B}) = \mathcal{S}^\perp$, which motivates the next result on the geometry of the global minimizers.

3.2.2.2 Geometry of the global solutions

Theorem 5. *Suppose $\boldsymbol{\mathcal{E}} = \boldsymbol{0}$. Then, any global solution $\boldsymbol{B}^*$ to problem* (3.75) *must be an orthonormal basis for $\mathcal{S}^\perp$ as long as*

$$\frac{M}{N} \cdot \frac{\sqrt{\overline{\eta}^2_{\boldsymbol{\mathcal{O}},c} + \left(c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c}\right)^2}}{c_{\boldsymbol{\mathcal{X}},\min}} < 1. \tag{3.89}$$

Proof. Let $\boldsymbol{B}^*$ be a global optimal solution of (3.75). To reach a contradiction, suppose that $\boldsymbol{B}^*$ is not an orthonormal basis for $\mathcal{S}^\perp$. It then follows from Lemma 9 that

$$\cos(\theta_c) \le \frac{M\overline{\eta}_{\boldsymbol{\mathcal{O}},c}}{Nc_{\boldsymbol{\mathcal{X}},\min}}, \tag{3.90}$$

where θ_c is the subspace angle between $\operatorname{Span}(\boldsymbol{B}^*)$ and $\mathcal{S}^\perp$. Utilizing the fact that $\boldsymbol{B}^*$ is a global solution, we have

$$\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}^*\|_{1,2} \le \min_{\boldsymbol{B}\in\mathbb{O}(D,c),\boldsymbol{B}\perp\mathcal{S}} \|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}\|_{1,2} = \min_{\boldsymbol{B}\in\mathbb{O}(D,c),\boldsymbol{B}\perp\mathcal{S}} \|\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}\|_{1,2} \le Mc_{\boldsymbol{\mathcal{O}},\max,c}. \tag{3.91}$$

On the other hand, by utilizing a similar decomposition of $\boldsymbol{B}^*$ as in (3.84), we can write $\boldsymbol{B}^* = \boldsymbol{P}\sin(\boldsymbol{\Theta}) + \boldsymbol{Q}\cos(\boldsymbol{\Theta})$, where $\boldsymbol{P} \in \mathbb{R}^{D\times c}$ and $\boldsymbol{Q} \in \mathbb{R}^{D\times c}$ are orthonormal matrices satisfying $\operatorname{Span}(\boldsymbol{P}) \subseteq \mathcal{S}$ and $\operatorname{Span}(\boldsymbol{Q}) \subseteq \mathcal{S}^\perp$, and $\boldsymbol{\Theta}$ is the diagonal matrix whose diagonal entries $\theta_1 \le \theta_2 \le \cdots \le \theta_c$ are the principal angles between $\operatorname{Span}(\boldsymbol{B}^*)$ and $\mathcal{S}^\perp$. Then we have

$$\begin{aligned}
\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}^*\|_{1,2} &= \sum_{j=1}^N \|\boldsymbol{x}_j^\top \boldsymbol{B}^*\|_2 + \sum_{j=1}^M \|\boldsymbol{o}_j^\top \boldsymbol{B}^*\|_2 \\
&= \sum_{j=1}^N \|\boldsymbol{x}_j^\top \boldsymbol{P}\sin(\boldsymbol{\Theta})\|_2 + \sum_{j=1}^M \|\boldsymbol{o}_j^\top \boldsymbol{B}^*\|_2 \\
&= \sum_{j=1}^N \sqrt{\sum_{k=1}^c \sin^2(\theta_k)(\boldsymbol{x}_j^\top \boldsymbol{p}_k)^2} + \sum_{j=1}^M \|\boldsymbol{o}_j^\top \boldsymbol{B}^*\|_2 \\
&\ge \sum_{j=1}^N \sin(\theta_c)\left|\boldsymbol{x}_j^\top \boldsymbol{p}_c\right| + \sum_{j=1}^M \|\boldsymbol{o}_j^\top \boldsymbol{B}^*\|_2 \\
&\ge \sin(\theta_c)Nc_{\boldsymbol{\mathcal{X}},\min} + Mc_{\boldsymbol{\mathcal{O}},\min,c},
\end{aligned}$$

which together with (3.91) gives

$$\sin(\theta_c) \leq \frac{M(c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c})}{N c_{\boldsymbol{\mathcal{X}},\min}}. \tag{3.92}$$

Combining (3.92) and (3.90), we obtain

$$1 = \sin^2(\theta_c) + \cos^2(\theta_c) \leq \frac{M^2(\overline{\eta}^2_{\boldsymbol{\mathcal{O}},c} + (c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c})^2)}{N^2 c^2_{\boldsymbol{\mathcal{X}},\min}},$$

which contradicts (3.89), thus completing the proof. □

Discussion of Theorem 5. Theorem 5 is an extension of Theorem 1 for the hyperplane case. Condition (3.89) tells us that, with fixed M/N and c, as we obtain more and more data points that are well-distributed, (3.89) is easier to be satisfied and thus any global solution to problem (3.75) spans $\mathcal{S}^\perp$. We remark that a similar theorem appeared in [29, Proposition 3], where they analyzed a group-DPCP formulation different from (3.75) that was designed specifically for homography estimation.

3.2.2.3 Probabilistic analysis

We now derive a probabilistic result that characterizes global optimality for problem (3.75) with noiseless data that is more interpretable. As a first step, we derive concentration bounds for the generalized geometric quantities $\eta_{\boldsymbol{\mathcal{O}},c}$ and $c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c}$ appearing in the deterministic Theorem 5. We begin with basic results in statistics.

Suppose $Z_1, \ldots, Z_n$ are n independent and identically distributed (i.i.d.) random observations from a probability measure P on a measurable space $(\Omega, \mathcal{A})$. Given a measurable function $f : \Omega \to \mathbb{R}$, the *empirical process* evaluated at f is defined as

$$\mathbb{G}_n f := \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^{n} f(Z_i) - \int f \,\mathrm{d}\, P \right), \tag{3.93}$$

where $\int f \,\mathrm{d}\, P$ is the expectation of f under P and $\frac{1}{n} \sum_{i=1}^{n} f(Z_i)$ is called the *empirical distribution.* Define an *envelope function* $F : \Omega \to \mathbb{R}$ such that $|f| \leq F$ for every $f \in \mathcal{F}$, where $\mathcal{F}$ is a given class of measurable functions. The $L_r(P)$-norm is defined as $\|f\|_{L_r(P)} = (\int |f|^r \,\mathrm{d}\, P)^{1/r}$. Given two functions l and u, the *bracket* $[l, u]$ is the set of all functions f with $l \leq f \leq u$. An ϵ-bracket in $L_r(P)$ is a bracket $[l, u]$ with $\int (u - l)^r \,\mathrm{d}\, P \leq \epsilon^r$ (since $l \leq u$, it is equivalent to say $\|u - l\|_{L_r(P)} \leq \epsilon$). The bracket number $N_{[]}(\epsilon, \mathcal{F}, L_2(P))$ is the minimum number of ϵ-brackets needed to cover $\mathcal{F}$, which can be viewed as a metric for characterizing the size of the class of functions $\mathcal{F}$.

Lemma 10. ([118, Corollary 19.35]). *For any class $\mathcal{F}$ of measurable functions and associated envelope function F, we have*

$$\mathbb{E}\left[\sup_{f \in \mathcal{F}} |\mathbb{G}_n f| \right] \lesssim J_{[]}(\|F\|_{P,2}, \mathcal{F}, L_2(P)), \tag{3.94}$$

where $J_{[]}(\|F\|_{P,2}, \mathcal{F}, L_2(P))$ is called the bracketing integral and defined as

$$J_{[]}(\|F\|_{L_2(P)}, \mathcal{F}, L_2(P)) = \int_0^{\|F\|_{L_2(P)}} \sqrt{\log\left(N_{[]}(\epsilon, \mathcal{F}, L_2(P)) \right)} \,\mathrm{d}\, \epsilon.$$

Lemma 11 (Vector-valued Comparison Inequality for Rademacher Process, [77]). *Let $\mathcal{F}$ be a class of functions $f : \mathbb{R}^D \to \mathbb{R}^c$ and let $h_i : \mathbb{R}^c \to \mathbb{R}$ for $i = 1, \ldots, N$ be*

1-Lipschitz functions. Then, for any $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N \in \mathbb{R}^D$, we have

$$\mathbb{E}\left[\sup_{f \in \mathcal{F}} \sum_{i=1}^{N} \varepsilon_i h_i(f(\boldsymbol{v}_i))\right] \le \sqrt{2}\mathbb{E}\left[\sup_{f \in \mathcal{F}} \sum_{i=1}^{N} \boldsymbol{\varepsilon}_i^\top f(\boldsymbol{v}_i)\right],$$

where ε_i are independent Rademacher random variables, and each $\boldsymbol{\varepsilon}_i \in \mathbb{R}^c$ is independent with each component an independent Rademacher random variable.

Bounding $\eta_{\boldsymbol{\mathcal{O}},c}$. In the following, we present the concentration bound for $\eta_{\boldsymbol{\mathcal{O}},c}$ under the random spherical model specified in Definition 1.

Lemma 12. *Consider the random spherical model in Definition 1. Fix a number $t > 0$, it follows that*

$$\mathbb{P}\left[\eta_{\boldsymbol{\mathcal{O}},c} \le C_0 \frac{\sqrt{cD}\log D + t}{\sqrt{M}}\right] \ge 1 - 2e^{-\frac{t^2}{2}}, \tag{3.95}$$

where C_0 is a universal constant independent of N, M, D, d, c and t.

Proof. First note that

$$\eta_{\boldsymbol{\mathcal{O}},c} = \frac{1}{M} \max_{\boldsymbol{B},\boldsymbol{G} \in \mathbb{O}(D,c), \boldsymbol{G} \perp \boldsymbol{B}} \left| \sum_{j=1}^{M} \left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}_j), \boldsymbol{G}^\top \boldsymbol{o}_j \right\rangle \right|,$$

and that we are going to show that

$$\mathbb{E}\left[\sup_{\boldsymbol{B},\boldsymbol{G} \in \mathbb{O}(D,c), \boldsymbol{G} \perp \boldsymbol{B}} \left| \sum_{j=1}^{M} \left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}_j), \boldsymbol{G}^\top \boldsymbol{o}_j \right\rangle \right| \right] \lesssim \sqrt{cD} \log(D) \sqrt{M},$$

where $\lesssim$ means smaller than up to a universal constant. By defining the set $\mathbb{F} :=$

$\{(\boldsymbol{B}, \boldsymbol{G}) : \boldsymbol{B}, \boldsymbol{G} \in \mathbb{O}(D, c), \boldsymbol{G} \perp \boldsymbol{B}\}$, and the parameterized function $f_{\boldsymbol{B},\boldsymbol{G}}(\boldsymbol{o}) := \left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}), \boldsymbol{G}^\top \boldsymbol{o} \right\rangle$, the class of functions we are interested in is $\mathcal{F} := \{f_{\boldsymbol{B},\boldsymbol{G}} : (\boldsymbol{B}, \boldsymbol{G}) \in \mathbb{F}\}$. Note that for any $f_{\boldsymbol{B},\boldsymbol{G}} \in \mathcal{F}$, we have $\mathbb{E}[f_{\boldsymbol{B},\boldsymbol{G}}(\boldsymbol{o})] = \mathbb{E}\left[\left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}), \boldsymbol{G}^\top \boldsymbol{o} \right\rangle\right] = 0$. Then, by viewing

$$\frac{1}{\sqrt{M}} \sup_{\boldsymbol{B},\boldsymbol{G} \in \mathbb{O}(D,c), \boldsymbol{G} \perp \boldsymbol{B}} \left| \sum_{j=1}^{M} \left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}_j), \boldsymbol{G}^\top \boldsymbol{o}_j \right\rangle \right|$$

as an empirical process, together with (3.93), this indicates that

$$\sum_{j=1}^{M} \left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}_j), \boldsymbol{G}^\top \boldsymbol{o}_j \right\rangle = \sqrt{M} \mathbb{G}_M f_{\boldsymbol{B},\boldsymbol{G}},$$

where $\mathbb{G}_M f_{\boldsymbol{B},\boldsymbol{G}}$ is the empirical process of $f_{\boldsymbol{B},\boldsymbol{G}}$.

To utilize Lemma 10, we show that the corresponding bracketing integral is finite for our problem. Since $|f_{\boldsymbol{B},\boldsymbol{G}}(\boldsymbol{o})| \leq \|\boldsymbol{o}\|_2$ for any $(\boldsymbol{B}, \boldsymbol{G}) \in \mathbb{F}$, we know $F(\boldsymbol{o}) = \|\boldsymbol{o}\|_2$ is the envelope function of $\mathcal{F}$ and $\|F\|_{P,2} = 1$. Thus, we only need to consider the bracket integral $J_{[]}(1, \mathcal{F}, L_2(P))$, where P is the corresponding probability measure. To that end, we compute the bracket number $N_{[]}(\epsilon, \mathcal{F}, L_2(P))$.

Since our function $f_{\boldsymbol{B},\boldsymbol{G}}$ is parameterized by $(\boldsymbol{B}, \boldsymbol{G})$, covering the class of functions $\mathcal{F}$ is related to covering the set $\mathbb{F}$. For any fixed $(\boldsymbol{B}, \boldsymbol{G}) \in \mathbb{F}$, define the set of points near $(\boldsymbol{B}, \boldsymbol{G})$ as

$$\mathbb{B}((\boldsymbol{B}, \boldsymbol{G}), \epsilon_1) := \left\{(\boldsymbol{B}', \boldsymbol{G}') \in \mathbb{F} : \sqrt{\|\boldsymbol{B} - \boldsymbol{B}'\|_F^2 + \|\boldsymbol{G} - \boldsymbol{G}'\|_F^2} \leq \epsilon_1\right\},$$

and define

$$\mathbb{A} := \left\{\boldsymbol{o} \in \mathbb{S}^{D-1} : \left\|\operatorname{sign}(\boldsymbol{o}^\top \boldsymbol{B}) - \operatorname{sign}(\boldsymbol{o}^\top \boldsymbol{B}')\right\| \leq \epsilon_2,\ \forall\ (\boldsymbol{B}', \boldsymbol{G}') \in \mathbb{B}((\boldsymbol{B}, \boldsymbol{G}), \epsilon_1)\right\}.$$

If $\boldsymbol{o} \in \mathbb{A}$, then for any $(\boldsymbol{B}', \boldsymbol{G}') \in \mathbb{B}((\boldsymbol{B}, \boldsymbol{G}), \epsilon_1)$ we have

$$\begin{aligned}
&|f_{\boldsymbol{B},\boldsymbol{G}}(\boldsymbol{o}) - f_{\boldsymbol{B}',\boldsymbol{G}'}(\boldsymbol{o})| \\
&= \left|\left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}), \boldsymbol{G}^\top \boldsymbol{o}\right\rangle - \left\langle \operatorname{sign}(\boldsymbol{B}'^\top \boldsymbol{o}), \boldsymbol{G}'^\top \boldsymbol{o}\right\rangle\right| \\
&= \left|\left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}), (\boldsymbol{G} - \boldsymbol{G}')^\top \boldsymbol{o}\right\rangle - \left\langle \left(\operatorname{sign}(\boldsymbol{B}'^\top \boldsymbol{o}) - \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o})\right), \boldsymbol{G}'^\top \boldsymbol{o}\right\rangle\right| \\
&\leq \|\boldsymbol{G} - \boldsymbol{G}'\| + \left\|\operatorname{sign}(\boldsymbol{B}'^\top \boldsymbol{o}) - \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o})\right\| \\
&\leq \epsilon_1 + \epsilon_2.
\end{aligned}$$

On the other hand, if $\boldsymbol{o} \in \mathbb{A}^c$, then for any $(\boldsymbol{B}', \boldsymbol{G}') \in \mathbb{B}((\boldsymbol{B}, \boldsymbol{G}), \epsilon_1)$ we have

$$|f_{\boldsymbol{B},\boldsymbol{G}}(\boldsymbol{o}) - f_{\boldsymbol{B}',\boldsymbol{G}'}(\boldsymbol{o})| = \left|\left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}), \boldsymbol{G}^\top \boldsymbol{o}\right\rangle\right| + \left|\left\langle \operatorname{sign}(\boldsymbol{B}'^\top \boldsymbol{o}), \boldsymbol{G}'^\top \boldsymbol{o}\right\rangle\right| \leq 2.$$

In summary, we have

$$|f_{\boldsymbol{B},\boldsymbol{G}}(\boldsymbol{o}) - f_{\boldsymbol{B}',\boldsymbol{G}'}(\boldsymbol{o})| \leq \epsilon_1 \mathbf{1}_{\mathbb{A}}(\boldsymbol{o}) + 2\mathbf{1}_{\mathbb{A}^c}(\boldsymbol{o}),\ \forall\ (\boldsymbol{B}', \boldsymbol{G}') \in \mathbb{B}((\boldsymbol{B}, \boldsymbol{G}), \epsilon_1), \tag{3.96}$$

where the indicator function $\mathbf{1}_{\mathbb{A}}(\boldsymbol{o})$ is defined as $\mathbf{1}_{\mathbb{A}}(\boldsymbol{o}) = \begin{cases} 1, & \boldsymbol{o} \in \mathbb{A} \\ 0, & \boldsymbol{o} \in \mathbb{A}^c \end{cases}.$

In order to bound $\mathbb{P}\left[\boldsymbol{o} \in \mathbb{A}^c\right]$, we note that

$$\begin{aligned}
&\left\|\operatorname{sign}(\boldsymbol{o}^\top \boldsymbol{B}) - \operatorname{sign}(\boldsymbol{o}^\top \boldsymbol{B}')\right\| \\
&= \left\| \frac{\boldsymbol{o}^\top \boldsymbol{B}}{\|\boldsymbol{o}^\top \boldsymbol{B}\|} - \frac{\boldsymbol{o}^\top \boldsymbol{B}'}{\|\boldsymbol{o}^\top \boldsymbol{B}'\|} \right\| = \left\| \frac{\left\|\boldsymbol{o}^\top \boldsymbol{B}'\right\| \boldsymbol{o}^\top \boldsymbol{B} - \left\|\boldsymbol{o}^\top \boldsymbol{B}\right\| \boldsymbol{o}^\top \boldsymbol{B}'}{\|\boldsymbol{o}^\top \boldsymbol{B}\| \, \|\boldsymbol{o}^\top \boldsymbol{B}'\|} \right\| \\
&= \left\| \frac{\left\|\boldsymbol{o}^\top \boldsymbol{B}'\right\| \boldsymbol{o}^\top (\boldsymbol{B} - \boldsymbol{B}') - \left(\left\|\boldsymbol{o}^\top \boldsymbol{B}\right\| - \left\|\boldsymbol{o}^\top \boldsymbol{B}'\right\| \right) \boldsymbol{o}^\top \boldsymbol{B}'}{\|\boldsymbol{o}^\top \boldsymbol{B}\| \, \|\boldsymbol{o}^\top \boldsymbol{B}'\|} \right\| \\
&\leq \frac{\|\boldsymbol{B} - \boldsymbol{B}'\|}{\|\boldsymbol{o}^\top \boldsymbol{B}\|} + \frac{\left| \left\|\boldsymbol{o}^\top \boldsymbol{B}\right\| - \left\|\boldsymbol{o}^\top \boldsymbol{B}'\right\| \right|}{\|\boldsymbol{o}^\top \boldsymbol{B}\|} \\
&\leq 2 \frac{\|\boldsymbol{B} - \boldsymbol{B}'\|}{\|\boldsymbol{o}^\top \boldsymbol{B}\|} \leq 2 \frac{\epsilon_1}{\|\boldsymbol{o}^\top \boldsymbol{B}\|}.
\end{aligned}$$

Thus, as long as $\left\|\boldsymbol{o}^\top \boldsymbol{B}\right\| \geq \frac{\epsilon_1}{2\epsilon_2}$, we have $\left\|\operatorname{sign}(\boldsymbol{o}^\top \boldsymbol{B}) - \operatorname{sign}(\boldsymbol{o}^\top \boldsymbol{B}')\right\| \leq \epsilon_2$. Hence

$$\mathbb{P}\left[\boldsymbol{o} \in \mathbb{A}^c\right] \leq \mathbb{P}\left[\left\|\boldsymbol{o}^\top \boldsymbol{B}\right\| \leq \frac{\epsilon_1}{2\epsilon_2}\right] \leq \mathbb{P}\left[o_1 \leq \frac{\epsilon_1}{2\epsilon_2}\right] \lesssim D \frac{\epsilon_1^2}{\epsilon_2^2}, \tag{3.97}$$

where o_1 is the first entry in $\boldsymbol{o}$, and the last inequality follows from [153, Lemma 12].

We now define a bracket $[l, u]$ by

$$\begin{aligned}
&l(\boldsymbol{o}) = f_{\boldsymbol{B},\boldsymbol{G}}(\boldsymbol{o}) - (\epsilon_1 + \epsilon_2)\mathbf{1}_{\mathbb{A}}(\boldsymbol{o}) - 2\mathbf{1}_{\mathbb{A}^c}(\boldsymbol{o}) \quad \text{and} \\
&u(\boldsymbol{o}) = f_{\boldsymbol{B},\boldsymbol{G}}(\boldsymbol{o}) + (\epsilon_1 + \epsilon_2)\mathbf{1}_{\mathbb{A}}(\boldsymbol{o}) + 2\mathbf{1}_{\mathbb{A}^c}(\boldsymbol{o}).
\end{aligned}$$

Due to (3.96), we have $f_{\boldsymbol{B}',\boldsymbol{G}'} \in [l, u]$ for all $(\boldsymbol{B}', \boldsymbol{G}') \in \mathbb{B}((\boldsymbol{B}, \boldsymbol{G}), \epsilon_1)$. Also,

$$\begin{aligned} \|u - l\|_{L_2(P)} &= \|2(\epsilon_1 + \epsilon_2)\mathbf{1}_{\mathbb{A}}(\boldsymbol{o}) + 4\mathbf{1}_{\mathbb{A}^c}(\boldsymbol{o})\|_{L_2(P)} \\ &= \sqrt{4(\epsilon_1 + \epsilon_2)^2 \mathbb{P}[\boldsymbol{o} \in \mathbb{A}] + 16\mathbb{P}[\boldsymbol{o} \in \mathbb{A}^c]} \\ &< 2(\epsilon_1 + \epsilon_2) + 4\sqrt{\mathbb{P}[\boldsymbol{o} \in \mathbb{A}^c]} \\ &\le 2(\epsilon_1 + \epsilon_2) + 4\sqrt{C_1 D \frac{\epsilon_1}{\epsilon_2}} \end{aligned} \tag{3.98}$$

where the last inequality follows from (3.97) with C_1 a universal constant. Therefore, the number of brackets to cover $\mathcal{F}$ is equal to the number of such balls $\mathbb{B}((\boldsymbol{B}, \boldsymbol{G}), \epsilon_1)$ that cover $\mathbb{F}$. According to [121, Lemma 5.2], the covering number for $\mathbb{F}$ is

$$\mathcal{N}(\mathbb{F}, \epsilon_1) \le \left(1 + \frac{2\sqrt{2}}{\epsilon_1}\right)^{2cD}. \tag{3.99}$$

Recall that the bracket number $N_{[]}(\epsilon, \mathcal{F}, L_2(P))$ is the minimum number of ϵ-brackets needed to cover $\mathcal{F}$, where an ϵ-bracket in $L_2(P)$ is a bracket $[l, u]$ with $\|u - l\|_{L_2(P)} \le \epsilon$. Thus, by letting $\epsilon_2 = \sqrt{\epsilon_1}$, $2(\epsilon_1 + \sqrt{\epsilon_1}) + 4\sqrt{C_1 D}\sqrt{\epsilon_1} = \epsilon$ and plugging this into (3.99), we obtain the bracket number

$$N_{[]}(\epsilon, \mathcal{F}, L_2(P)) \le \left(1 + C_2 \frac{D}{\epsilon^2}\right)^{2cD},$$

where C_2 is a universal constant. Now from Lemma 10, we have

$$\begin{aligned} &\frac{1}{\sqrt{M}} \mathbb{E}\left[\sup_{\boldsymbol{B}, \boldsymbol{G} \in \mathbb{O}(D,c), \boldsymbol{G} \perp \boldsymbol{B}} \left| \sum_{j=1}^M \left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}_j), \boldsymbol{G}^\top \boldsymbol{o}_j \right\rangle \right| \right] \\ &\lesssim \int_0^1 \sqrt{\left(1 + C_2 \frac{D}{\epsilon^2}\right)^{2cD}} \,\mathrm{d}\,\epsilon \lesssim \sqrt{cD} \log D. \end{aligned} \tag{3.100}$$

Since the product of compact spaces is compact, there exist $\boldsymbol{B}^+, \boldsymbol{G}^+ \in \mathbb{O}(D,c)$ for which the supremum in (3.100) is achieved. Then, for any $\boldsymbol{o}'_k \in \mathbb{S}^{D-1}$, we have

$$\begin{aligned}
&\Bigg| \sup_{\boldsymbol{B},\boldsymbol{G}\in\mathbb{O}(D,c),\boldsymbol{G}\perp\boldsymbol{B}} \left| \sum_{j=1}^{M} \left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}_j), \boldsymbol{G}^\top \boldsymbol{o}_j \right\rangle \right| \\
&\qquad - \sup_{\boldsymbol{B},\boldsymbol{G}\in\mathbb{O}(D,c),\boldsymbol{G}\perp\boldsymbol{B}} \left| \sum_{j\neq k} \left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}_j), \boldsymbol{G}^\top \boldsymbol{o}_j \right\rangle + \left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}'_k), \boldsymbol{G}^\top \boldsymbol{o}'_k \right\rangle \right| \Bigg| \\
&\leq \left| \left| \sum_{j=1}^{M} \left\langle \operatorname{sign}(\boldsymbol{B}^{+\top} \boldsymbol{o}_j), \boldsymbol{G}^{+\top} \boldsymbol{o}_j \right\rangle \right| - \left| \sum_{j\neq k} \left\langle \operatorname{sign}(\boldsymbol{B}^{+\top} \boldsymbol{o}_j), \boldsymbol{G}^{+\top} \boldsymbol{o}_j \right\rangle + \left\langle \operatorname{sign}(\boldsymbol{B}^{+\top} \boldsymbol{o}'_k), \boldsymbol{G}^{+\top} \boldsymbol{o}'_k \right\rangle \right| \right| \\
&\leq \left| \left\langle \operatorname{sign}(\boldsymbol{B}^{+\top} \boldsymbol{o}_k), \boldsymbol{G}^{+\top} \boldsymbol{o}_k \right\rangle - \left\langle \operatorname{sign}(\boldsymbol{B}^{+\top} \boldsymbol{o}_k), \boldsymbol{G}^{+\top} \boldsymbol{o}'_k \right\rangle \right| \leq 2,
\end{aligned}$$

where the second inequality follows from the reverse triangle inequality. Applying Lemma 4 with $c_k = 2$ and using (3.100), we obtain

$$\mathbb{P}\left[\sup_{\boldsymbol{B},\boldsymbol{G}\in\mathbb{O}(D,c),\boldsymbol{G}\perp\boldsymbol{B}} \left| \sum_{j=1}^{M} \left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}_j), \boldsymbol{G}^\top \boldsymbol{o}_j \right\rangle \right| \gtrsim \sqrt{M}\sqrt{cD}\log D + \epsilon \right] \leq 2e^{-\frac{2\epsilon^2}{4M}}.$$

Setting $\epsilon = t\sqrt{M}$, we have

$$\mathbb{P}\left[\sup_{\boldsymbol{B},\boldsymbol{G}\in\mathbb{O}(D,c),\boldsymbol{G}\perp\boldsymbol{B}} \left| \sum_{j=1}^{M} \left\langle \operatorname{sign}(\boldsymbol{B}^\top \boldsymbol{o}_j), \boldsymbol{G}^\top \boldsymbol{o}_j \right\rangle \right| \gtrsim \left(\sqrt{cD}\log D + t\right)\sqrt{M} \right] \leq 2e^{-\frac{t^2}{2}}.$$

Plugging back into the definition of $\eta_{\boldsymbol{\mathcal{O}},c}$, we get

$$\mathbb{P}\left[\eta_{\boldsymbol{\mathcal{O}},c} \gtrsim \frac{\sqrt{cD}\log D + t}{\sqrt{M}} \right] \leq 2e^{-\frac{t^2}{2}},$$

thus completing the proof. □

Bounding $c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c}$. Next, we present the concentration bound for

$c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c}$ under the random spherical model specified in Definition 1.

Lemma 13. *Consider the random spherical model in Definition 1. For a fixed number* $t > 0$, *we have that*

$$\mathbb{P}\left[c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c} \leq \frac{4\sqrt{2c}+t}{\sqrt{M}}\right] \geq 1 - 2e^{-\frac{t^2}{2}}. \tag{3.101}$$

Proof. First note that

$$\begin{aligned} & c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c} \\ & = \sup_{\boldsymbol{B}\in\mathbb{O}(D,c)} \frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\| - \inf_{\boldsymbol{B}\in\mathbb{O}(D,c)} \frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\| \\ & = \sup_{\boldsymbol{B}\in\mathbb{O}(D,c)} \left(\frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\| - \kappa\right) + \sup_{\boldsymbol{B}\in\mathbb{O}(D,c)} \left(\kappa - \frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\|\right) \end{aligned} \tag{3.102}$$

where $\kappa := \mathbb{E}_{\boldsymbol{o}\sim\mathbb{S}^{D-1}}\left\|\boldsymbol{B}^\top\boldsymbol{o}\right\|$. Applying Lemma 6, we have

$$\begin{aligned} \mathbb{E}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)} \left(\frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\| - \kappa\right)\right] & \leq \frac{2}{M}\mathbb{E}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)} \sum_{j=1}^{M}\varepsilon_j\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\|\right], \\ \mathbb{E}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)} \left(\kappa - \frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\|\right)\right] & \leq \frac{2}{M}\mathbb{E}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)} \sum_{j=1}^{M}\varepsilon_j\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\|\right], \end{aligned} \tag{3.103}$$

where ε_i are independent Rademacher random variables. We then apply Lemma 11, and get

$$\begin{aligned}
&\frac{2}{M}\mathbb{E}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)}\sum_{j=1}^{M}\varepsilon_j\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\|\right]\\
&\leq\frac{2\sqrt{2}}{M}\mathbb{E}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)}\sum_{j=1}^{M}\boldsymbol{\varepsilon}_j^\top\boldsymbol{B}^\top\boldsymbol{o}_j\right]\\
&=\frac{2\sqrt{2}}{M}\mathbb{E}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)}\left\langle\boldsymbol{B},\sum_{j=1}^{M}\boldsymbol{o}_j\boldsymbol{\varepsilon}_j^\top\right\rangle\right]\\
&\leq\frac{2\sqrt{2c}}{M}\mathbb{E}\left[\left\|\sum_{j=1}^{M}\boldsymbol{o}_j\boldsymbol{\varepsilon}_j^\top\right\|_F\right]\leq\frac{2\sqrt{2c}}{M}\sqrt{\mathbb{E}\left[\left\|\sum_{j=1}^{M}\boldsymbol{o}_j\boldsymbol{\varepsilon}_j^\top\right\|_F^2\right]}\\
&=\frac{2\sqrt{2c}}{M}\sqrt{\mathbb{E}\left[M+\sum_{i\neq j}\boldsymbol{\varepsilon}_i^\top\boldsymbol{\varepsilon}_j\boldsymbol{o}_i^T\boldsymbol{o}_j\right]}=\frac{2\sqrt{2c}}{M}\sqrt{M}=\frac{2\sqrt{2c}}{\sqrt{M}},
\end{aligned}$$

where $\boldsymbol{\varepsilon}_i\in\mathbb{R}^c$ contains independent Rademacher random variables, the second inequality utilizes the Cauchy-Schwartz inequality that $\langle\boldsymbol{B},\boldsymbol{A}\rangle\leq\|\boldsymbol{B}\|_F\|\boldsymbol{A}\|_F$, and the last inequality follows from Jensen's Inequality. Together with (3.102) and (3.103), this leads to

$$\mathbb{E}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)}\frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\|-\inf_{\boldsymbol{B}\in\mathbb{O}(D,c)}\frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\|\right]\leq\frac{4\sqrt{2c}}{\sqrt{M}}.\tag{3.104}$$

Furthermore, notice that for any $\boldsymbol{o}_k'\in\mathbb{S}^{D-1}$, we have

$$\begin{aligned}
&\Bigg|\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)}\frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\|-\inf_{\boldsymbol{B}\in\mathbb{O}(D,c)}\frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\|\\
&\quad-\frac{1}{M}\left(\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)}\Big(\sum_{j\neq k}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\|+\left\|\boldsymbol{B}^\top\boldsymbol{o}_k'\right\|\Big)-\inf_{\boldsymbol{B}\in\mathbb{O}(D,c)}\Big(\sum_{j\neq k}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\|+\left\|\boldsymbol{B}^\top\boldsymbol{o}_k'\right\|\Big)\right)\Bigg|\leq\frac{2}{M},
\end{aligned}$$

which after applying Lemma 4 and using (3.104) leads to

$$\mathbb{P}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)}\frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\| - \inf_{\boldsymbol{B}\in\mathbb{O}(D,c)}\frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\| \geq \frac{4\sqrt{2c}}{\sqrt{M}}+\epsilon\right] \leq 2e^{-\frac{\epsilon^2 M}{2}}.$$

Finally, by setting $\epsilon=\frac{t}{\sqrt{M}}$, we get

$$\mathbb{P}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)}\frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\| - \inf_{\boldsymbol{B}\in\mathbb{O}(D,c)}\frac{1}{M}\sum_{j=1}^{M}\left\|\boldsymbol{B}^\top\boldsymbol{o}_j\right\| \geq \frac{1}{\sqrt{M}}(4\sqrt{2c}+t)\right] \leq 2e^{-\frac{t^2}{2}}.$$

Plugging this back into the definitions of $c_{\boldsymbol{\mathcal{O}},\max,c}$ and $c_{\boldsymbol{\mathcal{O}},\min,c}$, we have

$$\mathbb{P}\left[c_{\boldsymbol{\mathcal{O}},\max,c}-c_{\boldsymbol{\mathcal{O}},\min,c}\geq\frac{4\sqrt{2c}+t}{\sqrt{M}}\right]\leq 2e^{-\frac{t^2}{2}},$$

thus completing the proof. □

Discussion of Lemma 12 and Lemma 13. First note that, similar to the concentrations of $\eta_{\boldsymbol{\mathcal{O}}}$ and $c_{\boldsymbol{\mathcal{O}},\max}-c_{\boldsymbol{\mathcal{O}},\min}$ in (3.9), both $\eta_{\boldsymbol{\mathcal{O}},\mathrm{c}}$ and $c_{\boldsymbol{\mathcal{O}},\mathrm{max,c}}-c_{\boldsymbol{\mathcal{O}},\mathrm{min,c}}$ scale as $O(1/\sqrt{M})$. Moreover, the role of c can be seen clearly from (3.95) and (3.101): as c increases, both $\eta_{\boldsymbol{\mathcal{O}},c}$ and $c_{\boldsymbol{\mathcal{O}},\max,c}-c_{\boldsymbol{\mathcal{O}},\min,c}$ tend to be larger. Together with the sufficient condition (3.89) for a global solution to (3.75) to span $\mathcal{S}^\perp$, this implies that (3.89) is more difficult to be satisfied for larger values of c.

We are now ready to give the probabilistic characterization of the global optimality for the holistic DPCP problem (3.75) for the noiseless setting.

Theorem 6. *Consider the random spherical model in Definition 1 with $\sigma=0$. Fix any $0<t<2\left(\sqrt{\frac{2N}{\pi d}}-2\right)$. With probability at least $1-6e^{-\frac{t^2}{2}}$, any global solution $\boldsymbol{B}^*$*

to problem (3.75) *must be an orthonormal basis for* $\mathcal{S}^\perp$ *if*

$$M\left((4\sqrt{c}+t)^2 + C_0(\sqrt{cD}\log D + t)^2\right) \leq N^2\left(\sqrt{\frac{2}{\pi d}} - \left(2+\frac{t}{2}\right)\frac{1}{\sqrt{N}}\right)^2, \qquad (3.105)$$

where C_0 *is a universal constant that is independent of* N, M, D, d, c *and* t*.*

Proof. Theorem 6 follows directly from Theorem 5 by plugging the concentrations for $c_{\boldsymbol{\mathcal{X}},\min}$ from (3.9), $\eta_{\boldsymbol{\mathcal{O}},c}$ from (3.95), and $c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c}$ from (3.101) into (3.89). □

Discussion of Theorem 6. Condition (3.105) interprets the global optimality condition (3.89) of Theorem 5 with natural quantities such as N, M, D, d and c. Most importantly, it validates that the new formulation (3.75) of DPCP on the Grassmannian $\mathbb{G}(D, c)$ is still able to tolerate $O(N^2)$ outliers for recovering the entire orthonormal basis of $\mathcal{S}^\perp$. Also, note that for fixed N, M, D, and d, the smaller c becomes, the easier it is for condition (3.105) to be satisfied. For the hyperplane case $c = 1$, Theorem 6 reduces to Theorem 2 that analyzes the original DPCP problem (3.1) without noise.

3.2.3 Analysis with noisy data

We now consider the holistic DPCP problem (3.75) under the scenario when inliers $\boldsymbol{\mathcal{X}}$ are further contaminated with noise, i.e., $\sigma > 0$ and $\boldsymbol{\mathcal{E}} \neq \mathbf{0}$ in Definition 1. As with the analysis for the noisy setting in Section 3.1.2, we decompose the noise term as $\boldsymbol{\mathcal{E}} = \boldsymbol{\mathcal{E}_s} + \boldsymbol{\mathcal{E}_n}$, where $\boldsymbol{\mathcal{E}_s}$ is the projection of $\boldsymbol{\mathcal{E}}$ onto $\mathcal{S}$ and $\boldsymbol{\mathcal{E}_n}$ is the projection onto $\mathcal{S}^\perp$. Observe that the term $\boldsymbol{\mathcal{E}_s}$ plays the same role as inliers since its columns lie exactly in $\mathcal{S}$, and that the component $\boldsymbol{\mathcal{E}_n}$ is the effective noise that influences the global solution to problem (3.75), making it different from the noiseless case. As before, we separate them by denoting $\widehat{\boldsymbol{\mathcal{X}}} := \boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}}_{\boldsymbol{s}}$ with $\operatorname{Span}(\widehat{\boldsymbol{\mathcal{X}}}) \subseteq \mathcal{S}$ and $\widehat{\boldsymbol{\mathcal{E}}} := \boldsymbol{\mathcal{E}}_{\boldsymbol{n}}$ with $\operatorname{Span}(\widehat{\boldsymbol{\mathcal{E}}}) \subseteq \mathcal{S}^{\perp}$. Since we have $\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}} = \widehat{\boldsymbol{\mathcal{X}}} + \widehat{\boldsymbol{\mathcal{E}}}$, and we can rewrite the objective in (3.75) as

$$f(\boldsymbol{B}) = \sum_{j=1}^{N} \|(\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^{\top}\boldsymbol{B}\|_2 + \sum_{j=1}^{M} \|\boldsymbol{o}_j^{\top}\boldsymbol{B}\|_2,$$

with $\widehat{\boldsymbol{x}}_j$ and $\widehat{\boldsymbol{\epsilon}}_j$ the j-th columns of $\widehat{\boldsymbol{\mathcal{X}}}$ and $\widehat{\boldsymbol{\mathcal{E}}}$, respectively.

Geometric quantities. First note that the previous quantities related to outliers, i.e., $c_{\boldsymbol{\mathcal{O}},\max,c}, c_{\boldsymbol{\mathcal{O}},\min,c}$ and $\eta_{\boldsymbol{\mathcal{O}},c}$, remain the same. For noisy inliers, we adopt the $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$ defined in (3.10) to characterize the distribution of the mixture of inliers and components of noise projected onto the inlier subspace. Additionally, we have one extra quantity with respect to $\widehat{\boldsymbol{\mathcal{E}}}$, namely

$$c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} := \frac{1}{N} \max_{\boldsymbol{B}\in\mathbb{O}(D,c)} \sum_{j=1}^{N} \|\widehat{\boldsymbol{\epsilon}}_j^{\top}\boldsymbol{B}\|_2, \tag{3.106}$$

which generalizes $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$ defined in (3.11) for $c = 1$, and quantifies the effective noise level. Note that $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} \leq \frac{1}{N}\sum_{j=1}^{N} \|\widehat{\boldsymbol{\epsilon}}_j\|_2$, which is the *total inlier residual* used in [61], but $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}$ also considers the geometry of the effective noise. To simplify the presentation of the remaining analysis, let

$$R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} := \frac{M}{N}\frac{\overline{\eta}_{\boldsymbol{\mathcal{O}},c}}{c_{\widehat{\boldsymbol{\mathcal{X}}},\min}} \quad \text{and} \quad R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c} := \frac{c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}}{c_{\widehat{\boldsymbol{\mathcal{X}}},\min}}, \tag{3.107}$$

which are analogous to $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}}}$ and $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}}$ defined in (3.12) and can be viewed as outlier-to-inlier and noise-to-inlier type of ratios, respectively.

3.2.3.1 Geometry of the critical points

We are now ready to characterize the distribution of the critical points of problem (3.75) when the dataset is further contaminated with noise.

Lemma 14. *Assume* $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} < 1$ *and*

$$\frac{32R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}}{\left(\sqrt{R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}+8}-3R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}\right)^{\frac{3}{2}}\left(\sqrt{R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}+8}+R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}\right)^{\frac{1}{2}}} < 1. \tag{3.108}$$

Then, every critical point $\boldsymbol{B}$ *of problem* (3.75) *spans a subspace that has an angle* θ *from* $\mathcal{S}^\perp$ *satisfying*

$$\theta \leq \sin^{-1}(t_1) \quad \textit{or} \quad \theta \geq \sin^{-1}(t_2) \tag{3.109}$$

where $0 \leq t_1 \leq t_2 \leq 1$ *with* t_1 *be the smallest nonnegative root of the quartic equation*

$$t^4 + (R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}-1)t^2 + 4R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}t + 4R^2_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c} = 0 \tag{3.110}$$

and

$$t_2 := \sqrt{1-\frac{1}{4}\left(R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}+\sqrt{R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}+8R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}}\right)^2}. \tag{3.111}$$

Proof. We note that the condition (3.108) and the quartic equation (3.110) have the same formulation as the condition (3.13) and the quartic equation (3.15) in Lemma 3, respectively. Then, according to the first part of the proof of Lemma 3, we know that if $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} < 1$ and (3.108) holds, the quartic equation (3.110) must have exactly two roots in $[0, 1]$, and we denote the smallest one by t_1.

Next, let us consider the geometry of the critical points of problem (3.75). There are two cases: $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c} = 0$ and $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c} > 0$. If $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c} = 0$, we can compute that $t_1 = 0$ from (3.110) and that $t_2 = \sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}}$ from (3.111), and can note that problem reduces to a noiseless one with dataset $[\widehat{\boldsymbol{\mathcal{X}}}\ \boldsymbol{\mathcal{O}}]$ (recall that the points in $\widehat{\boldsymbol{\mathcal{X}}}$ lie perfectly in the inlier subspace $\mathcal{S}$). According to Lemma 9, we have

$$\sin(\theta) = 0 \quad \text{or} \quad \sin(\theta) \geq \sqrt{1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}}, \tag{3.112}$$

which justifies the correctness of (3.109).

It remains to consider the case when $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c} > 0$. For any critical point $\boldsymbol{B}$ of problem (3.75), we utilize a similar decomposition of $\boldsymbol{B}$ as in (3.84), namely

$$\boldsymbol{B} = \boldsymbol{P}\sin(\boldsymbol{\Theta}) + \boldsymbol{Q}\cos(\boldsymbol{\Theta}), \tag{3.113}$$

where $\boldsymbol{P} \in \mathbb{R}^{D\times c}$ and $\boldsymbol{Q} \in \mathbb{R}^{D\times c}$ are orthonormal matrices satisfying $\text{Span}(\boldsymbol{P}) \subseteq \mathcal{S}$ and $\text{Span}(\boldsymbol{Q}) \subseteq \mathcal{S}^\perp$,and $\boldsymbol{\Theta}$ is the diagonal matrix whose diagonal entries $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_c$ are the principal angles between $\text{Span}(\boldsymbol{B})$ and $\mathcal{S}^\perp$ ($\theta_c \equiv \theta$ is also the subspace angle between $\text{Span}(\boldsymbol{B})$ and $\mathcal{S}^\perp$). As a result, $\boldsymbol{P}$ is orthogonal to $\boldsymbol{Q}$ and $\boldsymbol{B}$ is orthonormal. Note that if $\theta_c = 0$ ($\boldsymbol{B}$ is orthogonal to $\mathcal{S}$) or $\theta_c = \pi/2$, then (3.109) is trivial. Hence, for the rest of our discussion, we assume that $\theta_c \in (0, \pi/2)$.

Let $f(\boldsymbol{B}) := \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}\right\|_{1,2}$. For any critical point $\boldsymbol{B}$ of problem (3.75), there exists $\boldsymbol{W} \in \partial f(\boldsymbol{B})$ such that $(\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\boldsymbol{W} = \mathbf{0}$. Due to the general position [152, 153] of the data and the fact that $\boldsymbol{B}$ is not orthogonal to $\mathcal{S}$, $\boldsymbol{B}$ can be orthogonal to $K \leq D - c$ columns of $\widetilde{\boldsymbol{\mathcal{X}}}$, and therefore we have that

$$\begin{aligned}
\mathbf{0} &= (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\boldsymbol{W} \\
&= (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\left(\sum_{j=1}^{N}(\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)\operatorname{sign}((\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{B}) + \sum_{j=1}^{M}\boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B}) + \boldsymbol{\xi}\right)
\end{aligned}$$

where $\boldsymbol{\xi} = \sum_{k=1}^{K} \widetilde{\boldsymbol{x}}_{j_k}\boldsymbol{\alpha}_{j_k}$ with $\widetilde{\boldsymbol{x}}_{j_1}, \cdots, \widetilde{\boldsymbol{x}}_{j_K}$ the columns of $\widetilde{\boldsymbol{\mathcal{X}}}$ orthogonal to $\boldsymbol{B}$, and $\{\|\boldsymbol{\alpha}_{j_1}\|, \cdots, \|\boldsymbol{\alpha}_{j_K}\|\} \in [-1, 1]$. We further define

$$\boldsymbol{G} = \boldsymbol{P}\cos(\boldsymbol{\Theta}) - \boldsymbol{Q}\sin(\boldsymbol{\Theta}),$$

which is also an orthonormal matrix that is orthogonal to $\boldsymbol{B}$. Then, we have

$$\begin{aligned}
0 &= \left|\left\langle (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\boldsymbol{W}, \boldsymbol{G}\right\rangle\right| = \left|\left\langle \boldsymbol{W}, (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\boldsymbol{G}\right\rangle\right| = |\langle \boldsymbol{W}, \boldsymbol{G}\rangle| \\
&= \left|\left\langle \sum_{j=1}^{N}(\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)\operatorname{sign}((\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{B}), \boldsymbol{G}\right\rangle + \left\langle \sum_{j=1}^{M}\boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B}), \boldsymbol{G}\right\rangle + \langle \boldsymbol{\xi}, \boldsymbol{G}\rangle\right| \\
&= \left|\sum_{j=1}^{N}\left\langle (\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{G}, \operatorname{sign}((\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{B})\right\rangle\right. \\
&\qquad\qquad \left. + \left\langle (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\sum_{j=1}^{M}\boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B}), \boldsymbol{G}\right\rangle + \langle \boldsymbol{\xi}, \boldsymbol{G}\rangle\right| \qquad (3.114) \\
&\geq \left|\sum_{j=1}^{N}\left\langle (\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{G}, \operatorname{sign}((\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{B})\right\rangle\right| - M\eta_{\boldsymbol{\mathcal{O}},c} - D.
\end{aligned}$$

Considering the first term, we have

$$\begin{aligned}
&\left|\sum_{j=1}^{N}\left\langle(\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{G},\operatorname{sign}((\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{B})\right\rangle-\sum_{j=1}^{N}\left\langle\widehat{\boldsymbol{x}}_j^\top\boldsymbol{G},\operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{B})\right\rangle\right|\\
&=\left|\sum_{j=1}^{N}\left\langle\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\cos(\boldsymbol{\Theta})-\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\sin(\boldsymbol{\Theta}),\operatorname{sign}\left(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\sin(\boldsymbol{\Theta})+\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\cos(\boldsymbol{\Theta})\right)\right\rangle\right.\\
&\qquad\left.-\sum_{j=1}^{N}\left\langle\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\cos(\boldsymbol{\Theta}),\operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\sin(\boldsymbol{\Theta}))\right\rangle\right|\\
&\le\left|\sum_{j=1}^{N}\left\langle\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\cos(\boldsymbol{\Theta}),\operatorname{sign}\left(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\sin(\boldsymbol{\Theta})+\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\cos(\boldsymbol{\Theta})\right)-\operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\sin(\boldsymbol{\Theta}))\right\rangle\right| \qquad (3.115)\\
&\qquad+\left|\sum_{j=1}^{N}\left\langle\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\sin(\boldsymbol{\Theta}),\operatorname{sign}\left(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\sin(\boldsymbol{\Theta})+\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\cos(\boldsymbol{\Theta})\right)\right\rangle\right|\\
&\le\sum_{j=1}^{N}\left|\left\langle\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\cos(\boldsymbol{\Theta}),\operatorname{sign}\left(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\sin(\boldsymbol{\Theta})+\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\cos(\boldsymbol{\Theta})\right)-\operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\sin(\boldsymbol{\Theta}))\right\rangle\right|\\
&\qquad+\sum_{j=1}^{N}\left\|\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\sin(\boldsymbol{\Theta})\right\|.
\end{aligned}$$

Letting $\boldsymbol{a}_1:=\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\cos(\boldsymbol{\Theta}),\boldsymbol{a}_2:=\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\sin(\boldsymbol{\Theta}),\boldsymbol{e}:=\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\cos(\boldsymbol{\Theta})$, we have

$$\begin{aligned}
&|\langle\boldsymbol{a}_1,\operatorname{sign}(\boldsymbol{a}_2+\boldsymbol{e})-\operatorname{sign}(\boldsymbol{a}_2)\rangle|\\
&=\left|\frac{\langle\boldsymbol{a}_1,\boldsymbol{a}_2+\boldsymbol{e}\rangle}{\|\boldsymbol{a}_2+\boldsymbol{e}\|}-\frac{\langle\boldsymbol{a}_1,\boldsymbol{a}_2\rangle}{\|\boldsymbol{a}_2\|}\right|=\left|\frac{\langle\boldsymbol{a}_1,\boldsymbol{a}_2+\boldsymbol{e}\rangle\|\boldsymbol{a}_2\|-\langle\boldsymbol{a}_1,\boldsymbol{a}_2\rangle\|\boldsymbol{a}_2+\boldsymbol{e}\|}{\|\boldsymbol{a}_2+\boldsymbol{e}\|\|\boldsymbol{a}_2\|}\right|\\
&=\left|\frac{\langle\boldsymbol{a}_1,\boldsymbol{a}_2\rangle(\|\boldsymbol{a}_2+\boldsymbol{e}\|-\|\boldsymbol{a}_2\|)-\langle\boldsymbol{a}_1,\boldsymbol{e}\rangle\|\boldsymbol{a}_2\|}{\|\boldsymbol{a}_2+\boldsymbol{e}\|\|\boldsymbol{a}_2\|}\right|\\
&\le\left|\frac{\langle\boldsymbol{a}_1,\boldsymbol{a}_2\rangle(\|\boldsymbol{a}_2+\boldsymbol{e}\|-\|\boldsymbol{a}_2\|)}{\|\boldsymbol{a}_2+\boldsymbol{e}\|\|\boldsymbol{a}_2\|}\right|+\left|\frac{\langle\boldsymbol{a}_1,\boldsymbol{e}\rangle}{\|\boldsymbol{a}_2+\boldsymbol{e}\|}\right|\\
&\le\frac{\|\boldsymbol{a}_1\|\|\boldsymbol{e}\|}{\|\boldsymbol{a}_2+\boldsymbol{e}\|}+\frac{\|\boldsymbol{a}_1\|\|\boldsymbol{e}\|}{\|\boldsymbol{a}_2+\boldsymbol{e}\|}=\frac{2\|\boldsymbol{a}_1\|\|\boldsymbol{e}\|}{\|\boldsymbol{a}_2+\boldsymbol{e}\|}.
\end{aligned}$$

Plugging this back into (3.115), we have

$$\begin{aligned}
&\left|\sum_{j=1}^{N}\left\langle(\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{G},\operatorname{sign}((\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{B})\right\rangle-\sum_{j=1}^{N}\left\langle\widehat{\boldsymbol{x}}_j^\top\boldsymbol{G},\operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{B})\right\rangle\right|\\
&\leq 2\sum_{j=1}^{N}\frac{\|\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\cos(\boldsymbol{\Theta})\|\|\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\cos(\boldsymbol{\Theta})\|}{\|\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\sin(\boldsymbol{\Theta})+\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\cos(\boldsymbol{\Theta})\|}+\sin(\theta_c)\sum_{j=1}^{N}\left\|\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\right\| \qquad (3.116)\\
&\leq 2\sum_{j=1}^{N}\left(\frac{\|\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\|}{|\widehat{\boldsymbol{x}}_j^\top\boldsymbol{p}_c\sin(\theta_c)+\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{q}_c\cos(\theta_c)|}+\sin(\theta_c)\right)\left\|\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\right\|.
\end{aligned}$$

As in the proof of Lemma 3, we expect that in the noisy case the angle between $\operatorname{Span}(\boldsymbol{B})$ and $\mathcal{S}^\perp$, i.e., θ_c, is either near zero or close to $\pi/2$. On the one hand, from (3.114), we have

$$\begin{aligned}
0 &\geq \left|\sum_{j=1}^{N}\left\langle(\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{G},\operatorname{sign}((\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{B})\right\rangle\right|-M\eta_{\boldsymbol{\mathcal{O}},c}-D\\
&\geq \left|\sum_{j=1}^{N}\left\langle\widehat{\boldsymbol{x}}_j^\top\boldsymbol{G},\operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{B})\right\rangle\right|-M\eta_{\boldsymbol{\mathcal{O}},c}-D\\
&\quad-\left|\sum_{j=1}^{N}\left\langle(\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{G},\operatorname{sign}((\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{B})\right\rangle-\sum_{j=1}^{N}\left\langle\widehat{\boldsymbol{x}}_j^\top\boldsymbol{G},\operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{B})\right\rangle\right| \qquad (3.117)\\
&> \cos(\theta_c)\sum_{j=1}^{N}\left|\widehat{\boldsymbol{x}}_j^\top\boldsymbol{p}_c\right|-\frac{2}{\sin(\theta_c)}\sum_{j=1}^{N}\left\|\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\right\|-M\eta_{\boldsymbol{\mathcal{O}},c}-D\\
&> \cos(\theta_c)Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min}-\frac{2}{\sin(\theta_c)}Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}-M\overline{\eta}_{\boldsymbol{\mathcal{O}},c}
\end{aligned}$$

where the second inequality follows from the reverse triangular inequality, the third inequality follows from (3.116), and the last inequality uses the definitions of $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$, $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}$ and $\overline{\eta}_{\boldsymbol{\mathcal{O}},c}$. Thus, we obtain

$$0 > \cos(\theta_c)Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min}-\frac{2}{\sin(\theta_c)}Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}-M\overline{\eta}_{\boldsymbol{\mathcal{O}},c}, \qquad (3.118)$$

which has the same formulation of (3.26). According to the proof of Lemma 3, the lower valid region for θ_c is $\theta_c \le \sin^{-1}(t_1)$, where t_1 is the smallest nonnegative root of (3.110). On the other hand, for sufficiently large θ_c, (3.116) is bounded by $\frac{2}{\cos(\theta_c)} \sum_{j=1}^N \|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{Q}\|$, and thus similar to (3.118), we have

$$\begin{aligned} & 0 > \cos(\theta_c) N c_{\widehat{\boldsymbol{\mathcal{X}}},\min} - \frac{2}{\cos(\theta_c)} N c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} - M \overline{\eta}_{\boldsymbol{\mathcal{O}},c} \\ \Leftrightarrow \quad & \cos^2(\theta_c) - \cos(\theta_c) R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} - 2 R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c} < 0 \\ \Leftrightarrow \quad & \cos(\theta_c) < \frac{1}{2}\left(R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} + \sqrt{R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} + 8 R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}} \right). \end{aligned} \tag{3.119}$$

By defining

$$t_2 := \sqrt{1 - \frac{1}{4}\left(R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} + \sqrt{R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} + 8 R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}} \right)^2},$$

and combining (3.112) and (3.119), we obtain an upper valid region for θ_c as $\theta_c \ge \sin^{-1}(t_2)$. In summary, any critical point $\boldsymbol{B}$ of problem (3.75) spans a subspace that has an angle θ_c from $\boldsymbol{\mathcal{S}}^\perp$ satisfying

$$\theta_c \le \sin^{-1}(t_1) \quad \text{or} \quad \theta_c \ge \sin^{-1}(t_2).$$

Finally, we show that $t_2 \ge t_1$. Note that it follows from (3.37) and (3.38) that $t_1^2 \le \frac{1}{2}(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c})$, so that it is sufficient to show

$$t_2^2 = 1 - \frac{1}{4}\left(R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} + \sqrt{R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} + 8 R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}} \right)^2 \ge \frac{1}{2}\left(1 - R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}\right),$$

which is equivalent to

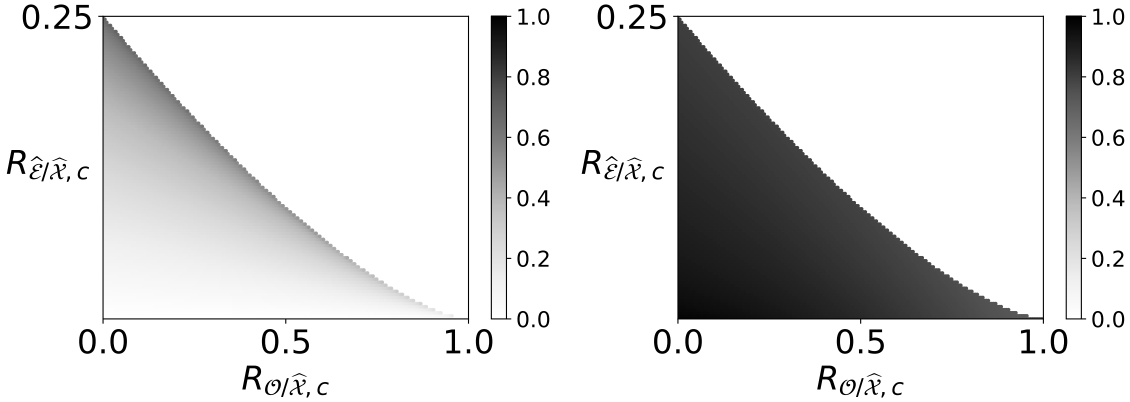
$$\frac{1}{2}R_{\mathcal{O}/\widehat{\mathcal{X}},c}\sqrt{R^2_{\mathcal{O}/\widehat{\mathcal{X}},c} + 8R_{\widehat{\mathcal{E}}/\widehat{\mathcal{X}},c}} + 2R_{\widehat{\mathcal{E}}/\widehat{\mathcal{X}},c} - \frac{1}{2} \leq 0, \tag{3.120}$$

which is guaranteed by condition (3.108), and thus completes the proof. □

Discussion of Lemma 14. The feasible region for $(R_{\mathcal{O}/\widehat{\mathcal{X}},c}, R_{\widehat{\mathcal{E}}/\widehat{\mathcal{X}},c})$ with condition (3.108) satisfied is shown as the area under the curve in Figure 3.4, which implies that the outlier-to-inlier ratio and the noise-to-inlier ratio cannot be very large at the same time. In other words, larger noise levels restrict the number of outliers that the holistic DPCP problem (3.75) can tolerate. Next, (3.109) indicates that any critical point $\boldsymbol{B}$ of the noisy problem (3.75) spans a subspace that is close to either $\mathcal{S}^\perp$ or $\mathcal{S}$. Figure 3.4 provides a better understanding of t_1 and t_2: with smaller outlier-to-inlier ratio and noise-to-inlier ratio, t_1 is closer to 0 (lighter) and t_2 is closer to 1 (darker), making the geometric location of $\boldsymbol{B}$ more restricted. Compared with Lemma 9 for the noiseless case where $\boldsymbol{B}$ is an exact orthonormal basis of $\mathcal{S}^\perp$ if it is sufficiently far from $\mathcal{S}$, here we can only guarantee that it lies in a neighborhood of $\mathcal{S}^\perp$, i.e., $\theta \leq \sin^{-1}(t_1)$, due to the noise. According to Proposition 1, one can further bound t_1 by

$$t_1 \leq \frac{25R_{\widehat{\mathcal{E}}/\widehat{\mathcal{X}},c}}{(1 - R_{\mathcal{O}/\widehat{\mathcal{X}},c})^2}. \tag{3.121}$$

When there is no noise, from (3.121) we have $t_1 = 0$, and from (3.111) we have $t_2 = \sqrt{1 - R^2_{\mathcal{O}/\widehat{\mathcal{X}},c}}$, which is consistent with Lemma 9. Moreover, (3.121) shows that t_1 is small with small outlier-to-inlier ratio and noise-to-inlier ratio, and is proportional to



(a) Value of t_1

(b) Value of t_2

Figure 3.4. Plot of (a) t_1 and (b) t_2 in Lemma 14 given $(R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}, R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c})$ pairs such that condition (3.108) holds true (area below the curve).

the effective noise level. Finally, compared with Lemma 3 that analyzes the geometry of the critical points for the noisy problem (3.1) with $c = 1$, where both t_1 and t_2 are defined by the nonnegative roots of (3.15), in this generalized analysis t_2 is decoupled from (3.110) (see (3.111)) since we have used a different proof technique for problem (3.75) defined over the Grassmannian.

3.2.3.2 Geometry of the global solutions

Using Lemma 14, we now characterize the global solution of the holistic DPCP problem (3.75) in the noisy setting.

Theorem 7. *If* $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} < 1$, (3.108) *holds, and*

$$R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} + \left(\frac{M(c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c})}{Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min}} + 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}\right)^2 + 8R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c} < 1, \tag{3.122}$$

then any global solution $\boldsymbol{B}^*$ *of problem* (3.75) *must span a subspace that has an angle*

θ^* from $\mathcal{S}^\perp$ satisfying

$$\theta^* \leq \sin^{-1}(t_1), \tag{3.123}$$

where $0 \leq t_1 \leq 1$ is the smallest nonnegative root of (3.15)*.*

Proof. Since $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} < 1$ and (3.108) holds, we can apply Lemma 14 to obtain that any critical point $\boldsymbol{B}$ of problem (3.75) must have principal angle θ from $\mathcal{S}^\perp$ satisfy

$$\sin(\theta) \leq t_1 \quad \text{or} \quad \cos(\theta) \leq \frac{1}{2}\left(R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} + \sqrt{R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} + 8R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}}\right),$$

where $0 \leq t_1 \leq 1$ is the smallest nonnegative root of (3.110). Since a global minimizer $\boldsymbol{B}^*$ must be a critical point, to reach a contradiction, let us assume that (3.123) does not hold, so that

$$\cos(\theta^*) \leq \sqrt{R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} + 8R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}}. \tag{3.124}$$

Utilizing the fact that $\boldsymbol{B}^*$ is a global solution, we have

$$\begin{aligned}
\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}^*\|_{1,2} &\leq \min_{\boldsymbol{B}\in\mathbb{O}(D,c),\boldsymbol{B}\perp\mathcal{S}} \|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}\|_{1,2} \\
&= \min_{\boldsymbol{B}\in\mathbb{O}(D,c),\boldsymbol{B}\perp\mathcal{S}} \left\{\|\widehat{\boldsymbol{\mathcal{E}}}^\top \boldsymbol{B}\|_{1,2} + \|\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}\|_{1,2}\right\} \\
&\leq N c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} + M c_{\boldsymbol{\mathcal{O}},\max,c}.
\end{aligned} \tag{3.125}$$

On the other hand, by utilizing a similar decomposition of $\boldsymbol{B}^*$ as in (3.84), we can write $\boldsymbol{B}^* = \boldsymbol{P}\sin(\boldsymbol{\Theta}) + \boldsymbol{Q}\cos(\boldsymbol{\Theta})$, where $\boldsymbol{P} \in \mathbb{R}^{D\times c}$ and $\boldsymbol{Q} \in \mathbb{R}^{D\times c}$ are orthonormal matrices satisfying $\mathrm{Span}(\boldsymbol{P}) \subseteq \mathcal{S}$ and $\mathrm{Span}(\boldsymbol{Q}) \subseteq \mathcal{S}^{\perp}$,and $\boldsymbol{\Theta}$ is the diagonal matrix whose diagonal entries $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_c$ are the principal angles between $\mathrm{Span}(\boldsymbol{B}^*)$ and $\mathcal{S}^{\perp}$ ($\theta_c \equiv \theta^*$ is also the subspace angle between $\mathrm{Span}(\boldsymbol{B}^*)$ and $\mathcal{S}^{\perp}$). Then

$$\begin{aligned}
\|\widetilde{\boldsymbol{\mathcal{X}}}^{\top}\boldsymbol{B}^*\|_{1,2} &= \sum_{j=1}^{N} \|(\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^{\top}\boldsymbol{B}^*\|_2 + \sum_{j=1}^{M} \|\boldsymbol{o}_j^{\top}\boldsymbol{B}^*\|_2 \\
&= \sum_{j=1}^{N} \|\widehat{\boldsymbol{x}}_j^{\top}\boldsymbol{P}\sin(\boldsymbol{\Theta}) + \widehat{\boldsymbol{\epsilon}}_j^{\top}\boldsymbol{Q}\cos(\boldsymbol{\Theta})\|_2 + \sum_{j=1}^{M} \|\boldsymbol{o}_j^{\top}\boldsymbol{B}^*\|_2 \\
&\geq \sum_{j=1}^{N} \|\widehat{\boldsymbol{x}}_j^{\top}\boldsymbol{P}\sin(\boldsymbol{\Theta})\|_2 - \sum_{j=1}^{N} \|\widehat{\boldsymbol{\epsilon}}_j^{\top}\boldsymbol{Q}\cos(\boldsymbol{\Theta})\|_2 + \sum_{j=1}^{M} \|\boldsymbol{o}_j^{\top}\boldsymbol{B}^*\|_2 \\
&\geq \sum_{j=1}^{N} \sin(\theta^*)\left|\widehat{\boldsymbol{x}}_j^{\top}\boldsymbol{p}_c\right| - \sum_{j=1}^{N} \|\widehat{\boldsymbol{\epsilon}}_j^{\top}\boldsymbol{Q}\|_2 + \sum_{j=1}^{M} \|\boldsymbol{o}_j^{\top}\boldsymbol{B}^*\|_2 \\
&\geq \sin(\theta^*) N c_{\widehat{\boldsymbol{\mathcal{X}}},\min} - N c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} + M c_{\boldsymbol{\mathcal{O}},\min,c},
\end{aligned}$$

which together with (3.125) gives

$$\sin(\theta^*) \leq \frac{M(c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c}) + 2N c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}}{N c_{\widehat{\boldsymbol{\mathcal{X}}},\min}}. \tag{3.126}$$

Combining (3.124) and (3.126), we obtain

$$\begin{aligned}
1 &= \sin^2(\theta^*) + \cos^2(\theta^*) \\
&\leq R^2_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} + \left(\frac{M(c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c})}{N c_{\widehat{\boldsymbol{\mathcal{X}}},\min}} + 2R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}\right)^2 + 8R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c},
\end{aligned}$$

which contradicts (3.122), thus completing the proof. □

Discussion of Theorem 7. Condition (3.122) is sufficient to ensure that global solutions of problem (3.75) span a subspace that is close to $\mathcal{S}^{\perp}$. We interpret (3.122) as follows: with fixed M/N, as data points are increasing ($c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c} \to 0$) and well-distributed (large $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$, small $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}$), and the effective noise is mild (small $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}$), (3.122) will be satisfied and global solutions of (3.75) must be close to $\mathcal{S}^{\perp}$. Note that in the noiseless case, condition (3.122) is equivalent to condition (3.89) and $t_1 = 0$, which means Theorem 7 is precisely Theorem 5 in the noiseless setting.

3.2.3.3 Probabilistic analysis

We now provide a probabilistic characterization of global optimality for problem (3.75) in the noisy setting. We have already derived the concentration bounds for $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$ (Lemma 7), $\eta_{\boldsymbol{\mathcal{O}},c}$ (Lemma 12) and $c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c}$ (Lemma 13). For a statistical analysis of the deterministic result in Theorem 7, we are left to derive a concentration inequality for $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}$.

Bounding $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}$. In the following, we present the concentration bound for $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}$ under the random spherical model specified in Definition 1.

Lemma 15. *Consider the random spherical model defined in Definition 1. For a fixed number $t > 0$, we have*

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} \le \left(1 + \frac{2\sqrt{2c}}{\sqrt{N}}\right)\delta(\sigma) + \frac{t}{\sqrt{N}}\right] \ge 1 - 2e^{-\frac{t^2}{2}} \tag{3.127}$$

where $\delta(\sigma)$ is defined in (3.65).

Proof. According to the generative model in Definition 1, let $\overline{\boldsymbol{x}}_1, \cdots, \overline{\boldsymbol{x}}_N \sim \mathcal{N}\left(\mathbf{0}, \frac{1}{d}\mathcal{P}_{\mathcal{S}}\right)$

and $\overline{\boldsymbol{\epsilon}}_1, \cdots, \overline{\boldsymbol{\epsilon}}_N \sim \mathcal{N}\left(\mathbf{0}, \frac{\sigma^2}{D}\mathbf{I}\right)$ be from (3.56) and (3.57), and $\widehat{\boldsymbol{\mathcal{E}}} = \boldsymbol{\mathcal{E}}_{\boldsymbol{n}}$. Then, we have

$$c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} = \max_{\boldsymbol{B}\in\mathbb{O}(D,c)} \frac{1}{N}\sum_{j=1}^{N} \left\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}\right\|_2 = \max_{\boldsymbol{B}\in\mathbb{O}(D,c)} \frac{1}{N}\sum_{j=1}^{N} \left\|\frac{\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}\top}\boldsymbol{B}}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2}\right\|_2 \le \frac{1}{N}\sum_{j=1}^{N} \frac{\|\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}}\|_2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2}$$

where the last inequality follows from $\|\boldsymbol{B}\|_2 = 1$. Defining the random variable

$$R_j := \frac{\|\overline{\boldsymbol{\epsilon}}_j^{\boldsymbol{n}}\|_2}{\|\overline{\boldsymbol{x}}_j + \overline{\boldsymbol{\epsilon}}_j\|_2},$$

we are interested in $\mathbb{E}\left[R_j\right]$. According to the proof of Lemma 8 and (3.66), we have

$$\mathbb{E}\left[R_j\right] \le \sqrt{\sigma} + \sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)},$$

which leads to

$$\mathbb{E}_0 := \mathbb{E}\left[\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}\|_2\right] \le \mathbb{E}[R_j] \le \sqrt{\sigma} + \sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)}. \tag{3.128}$$

We are now ready to bound $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}$. Note that

$$c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} = \sup_{\boldsymbol{B}\in\mathbb{O}(D,c)} \frac{1}{N}\sum_{j=1}^{N} \left\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}\right\|_2 = \sup_{\boldsymbol{B}\in\mathbb{O}(D,c)} \left(\frac{1}{N}\sum_{j=1}^{N} \left\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}\right\|_2 - \mathbb{E}_0\right) + \mathbb{E}_0. \tag{3.129}$$

Since $\mathbb{O}(D,c)$ is compact, there exists $\boldsymbol{B}^+ \in \mathbb{O}(D,c)$ that achieves the supremum in

(3.129). Therefore, for any $\widehat{\boldsymbol{\epsilon}}_1, \widehat{\boldsymbol{\epsilon}}_2, \cdots, \widehat{\boldsymbol{\epsilon}}_N, \widehat{\boldsymbol{\epsilon}}'_k$, we have

$$\begin{aligned}
&\left|\sup\left(\frac{1}{N}\sum_{j=1}^{N}\left\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}\right\|_2 - \mathbb{E}_0\right) - \sup\left(\frac{1}{N}\sum_{j\neq k}\left(\left\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}\right\|_2 + \left\|\widehat{\boldsymbol{\epsilon}}_k'^\top \boldsymbol{B}\right\|_2\right) - \mathbb{E}_0\right)\right| \\
&\leq \left|\frac{1}{N}\sum_{j=1}^{N}\left\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}^+\right\|_2 - \mathbb{E}_0 - \left(\frac{1}{N}\sum_{j\neq k}\left(\left\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}^+\right\|_2 + \left\|\widehat{\boldsymbol{\epsilon}}_k'^\top \boldsymbol{B}^+\right\|_2\right) - \mathbb{E}_0\right)\right| \qquad (3.130) \\
&= \left|\frac{1}{N}\left(\left\|\widehat{\boldsymbol{\epsilon}}_k^\top \boldsymbol{B}^+\right\|_2 - \left\|\widehat{\boldsymbol{\epsilon}}_k'^\top \boldsymbol{B}^+\right\|_2\right)\right| \leq \frac{2}{N}.
\end{aligned}$$

Applying Lemma 4 with $c_k = \frac{2}{N}$, we have

$$\mathbb{P}\left[\left|\sup\left(\frac{1}{N}\sum_{j=1}^{N}\left\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}\right\|_2 - \mathbb{E}_0\right) - \mathbb{E}\left[\sup\left(\frac{1}{N}\sum_{j=1}^{N}\left\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}\right\|_2 - \mathbb{E}_0\right)\right]\right| \geq \epsilon\right] \leq 2e^{-\frac{\epsilon^2 N}{2}}. \qquad (3.131)$$

Moreover, we have

$$\begin{aligned}
&\mathbb{E}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)}\left(\frac{1}{N}\sum_{j=1}^{N}\left\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}\right\|_2 - \mathbb{E}_0\right)\right] \\
&\leq 2\mathbb{E}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)}\frac{1}{N}\sum_{j=1}^{N}\varepsilon_j\left\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}\right\|_2\right] \leq 2\sqrt{2}\mathbb{E}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)}\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{\varepsilon}_j^\top \boldsymbol{B}^\top \widehat{\boldsymbol{\epsilon}}_j\right] \\
&= \frac{2\sqrt{2}}{N}\mathbb{E}\left[\sup_{\boldsymbol{B}\in\mathbb{O}(D,c)}\left\langle \boldsymbol{B}, \sum_{j=1}^{N}\widehat{\boldsymbol{\epsilon}}_j\boldsymbol{\varepsilon}_j^\top\right\rangle\right] \leq \frac{2\sqrt{2c}}{N}\mathbb{E}\left[\left\|\sum_{j=1}^{N}\widehat{\boldsymbol{\epsilon}}_j\boldsymbol{\varepsilon}_j^\top\right\|_F\right] \\
&\leq \frac{2\sqrt{2c}}{N}\sqrt{\mathbb{E}\left[\left\|\sum_{j=1}^{N}\widehat{\boldsymbol{\epsilon}}_j\boldsymbol{\varepsilon}_j^\top\right\|_F^2\right]} = \frac{2\sqrt{2c}}{N}\sqrt{\mathbb{E}\left[\sum_{j=1}^{N}\|\widehat{\boldsymbol{\epsilon}}_j\|_2^2 + \sum_{i\neq j}\boldsymbol{\varepsilon}_i^\top\boldsymbol{\varepsilon}_j\widehat{\boldsymbol{\epsilon}}_i^\top\widehat{\boldsymbol{\epsilon}}_j\right]} \\
&\leq \frac{2\sqrt{2c}}{\sqrt{N}}\sqrt{\sigma + (1-\sigma)F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)} \leq \frac{2\sqrt{2c}}{\sqrt{N}}\left(\sqrt{\sigma} + \sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)}\right),
\end{aligned} \qquad (3.132)$$

where the first inequality follows from Lemma 6, the second inequality follows from Lemma 11 by taking $h_j = \|\cdot\|_2$, the third inequality follows from Cauchy-

Schwartz, the fourth inequality follows from the Jensen's Inequality, and the fifth inequality follows from an upper bound for $\mathbb{E}[\|\widehat{\boldsymbol{\epsilon}}_j\|_2^2] = \mathbb{E}[R_j^2]$ that is similar to (3.66). Applying (3.132) to (3.131), we obtain

$$\mathbb{P}\left[\sup\left(\frac{1}{N}\sum_{j=1}^{N}\left\|\widehat{\boldsymbol{\epsilon}}_j^\top \boldsymbol{B}\right\|_2 - \mathbb{E}_0\right) \geq \frac{2\sqrt{2c}}{\sqrt{N}}\left(\sqrt{\sigma}+\sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)}\right)+\epsilon\right]$$
$$\leq 2e^{-\frac{\epsilon^2 N}{2}}.$$

Therefore, from (3.129), we have

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} \geq \mathbb{E}_0 + \frac{2\sqrt{2c}}{\sqrt{N}}\left(\sqrt{\sigma}+\sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)}\right)+\epsilon\right] \leq 2e^{-\frac{\epsilon^2 N}{2}}.$$

Applying the upper bound for $\mathbb{E}_0$ in (3.128), we obtain

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} \geq \left(1+\frac{2\sqrt{2c}}{\sqrt{N}}\right)\left(\sqrt{\sigma}+\sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)}\right)+\epsilon\right] \leq 2e^{-\frac{\epsilon^2 N}{2}},$$

and by setting $\epsilon = \frac{t}{\sqrt{N}}$ we have

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} \geq \left(1+\frac{2\sqrt{2c}}{\sqrt{N}}\right)\left(\sqrt{\sigma}+\sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\frac{(D-d)\sigma}{D+d\sigma^2}\right)}\right)+\frac{t}{\sqrt{N}}\right] \leq 2e^{-\frac{t^2}{2}}.$$

Noting that $\frac{(D-d)\sigma}{D+d\sigma^2} < \sigma$ and all the CDFs are nondecreasing, we get

$$\mathbb{P}\left[c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} \geq \left(1+\frac{2\sqrt{2c}}{\sqrt{N}}\right)\left(\sqrt{\sigma}+\sqrt{1-\sigma}\sqrt{F_{d,D-d}\left(\sigma\right)}\right)+\frac{t}{\sqrt{N}}\right] \leq 2e^{-\frac{t^2}{2}},$$

which completes the proof. □

Discussion of Lemma 15. It has been shown in Lemma 8 that $\delta(\sigma) = O(\sigma^{\frac{d}{4}} + \sigma^{\frac{1}{2}})$, implying that the concentration for $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}$ in (3.127) implies that $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} = O(\sigma^{\frac{d}{4}} + \sigma^{\frac{1}{2}})$ with high probability. However, as in the discussion after Lemma 8, the concentration bound for $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}$ does not immediately imply $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} = 0$ when $\sigma = 0$ because of the term $\frac{t}{\sqrt{N}}$ (this is usually very small since N is very large compared with t), which appears to be an artifact of the proof technique; improvement is left as future work.

We now give the probabilistic characterization of the globally optimal solutions for the problem (3.75).

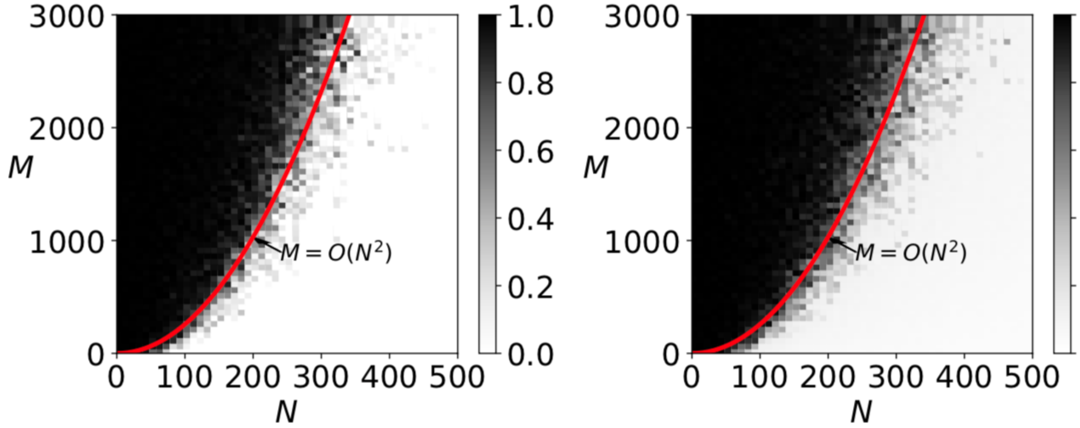
Theorem 8. *Consider the random spherical model defined in Definition 1. Assume $N > c$. Then for any positive $t < 2\Big(\sqrt{\frac{2N}{\pi d}}\rho(\sigma) - 2\Big)$, with probability at least $1 - 8e^{-t^2/2}$, any global solution $\boldsymbol{B}^*$ of problem (3.75) must span a subspace that has an angle θ^* from $\mathcal{S}^\perp$ satisfying*

$$\sin(\theta^*) \le \frac{C_1\delta(\sigma) + \frac{25t}{\sqrt{N}}}{\sqrt{\frac{2}{\pi d}}\rho(\sigma) - C_2\frac{t\sqrt{M} + \sqrt{cDM}\log D}{N} - \frac{4+t}{2\sqrt{N}}} \tag{3.133}$$

as long as

$$\begin{aligned} &M\left((8\sqrt{2c} + 2t)^2 + C_3(\sqrt{cD}\log D + t)^2\right) \\ &\qquad \le N^2\Bigg[\left(\sqrt{\frac{2}{\pi d}}\rho(\sigma) - \frac{4+t}{2\sqrt{N}}\right)^2 - C_4\delta(\sigma) - \frac{16t^2}{N} - \frac{8t}{\sqrt{dN}}\Bigg], \end{aligned} \tag{3.134}$$

where C_1, C_2, C_3, C_4 are universal constants independent of N, M, D, d, c, t and σ.



(a) $\sigma = 0$

(b) $\sigma = 0.1$

Figure 3.5. Plot of the subspace angle between Span($\boldsymbol{B}^*$) and $\mathcal{S}^\perp$ with $\boldsymbol{B}^*$ the computed solution to the noisy holistic DPCP problem (3.75) when varying N and M for noise level (a) $\sigma = 0$ and (b) $\sigma = 0.1$. Here we fix $D = 30$ and $c = 5$.

Proof. Theorem 8 follows directly from Theorem 7 and

$$t_1 \leq \frac{25 R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}}{(1 - R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c})^2} \tag{3.135}$$

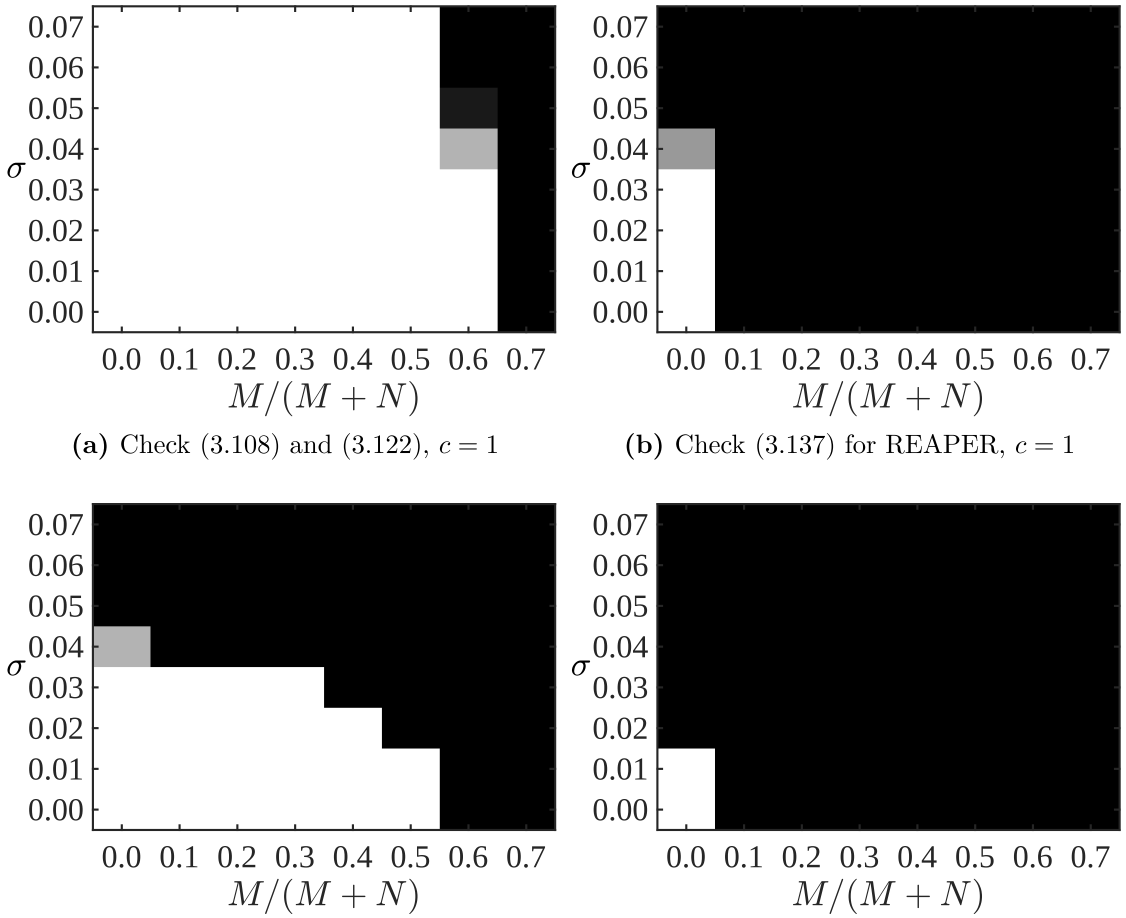
by plugging the concentrations for $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$ from (3.54), $\eta_{\boldsymbol{\mathcal{O}},c}$ from (3.95), $c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c}$ from (3.101), and $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}$ from (3.127) into (3.122) and (3.135). □

Discussion of Theorem 8. Towards interpreting Theorem 8, first recall that $\delta(\sigma) \to 0$ and $\rho(\sigma) \to 1$ as $\sigma \to 0$. Then, (3.133) indicates that the angle θ^* between $\mathcal{S}^\perp$ and the subspace spanned by a global solution $\boldsymbol{B}^*$ of (3.75) becomes close to zero as $\sigma \to 0$, and $\sin(\theta^*) = O(\sigma^{d/4} + \sigma^{\frac{1}{2}})$ which is of the same order as $\delta(\sigma)$. Furthermore, the sufficient condition (3.134) implies that problem (3.75) can also tolerate $O(N^2)$ outliers for learning the entire orthonormal basis for $\mathcal{S}^\perp$ with noisy data, as illustrated in Figure 3.5. Finally, we remark that condition (3.134) does not necessarily have the same form as condition (3.105) when $\sigma = 0$ or the condition in (3.74) when $c = 1$ because the proof used is different; however, they all reveal that the DPCP problems (both (3.1) and (3.75)) can handle $O(N^2)$ outliers, which is an apparent advantage over other RSR methods [62] that can only deal with $O(N)$ outliers in theory.

3.3 Comparison with state-of-the-art

As noted in Section 2.2, DPCP is very closely related to least absolute deviations subspace learning methods. Two important representatives of that class are REAPER [61] and the Geodesic Gradient Descent (GGD) method of [76]. In particular, the GGD problem (2.6) shares a similar formulation as (3.75), which optimizes over $\mathbb{G}(D, d)$ and aims at recovering an orthonormal basis for the underlying subspace $\mathcal{S}$ instead of a basis for the dual space $\mathcal{S}^\perp$ as in DPCP, while the problem of REAPER (2.5) can be viewed as a convex relaxation of it. In this section, we compare the theoretical results of DPCP to those known for REAPER and GGD. We show that the global optimality conditions for DPCP given in the previous sections are much tighter compared to those required for REAPER. In fact, they are even an improvement over conditions that enable a local stability characterization of the function landscape given by [76].

Comparison with REAPER [61]. For the global optimality analysis, [61, Theorem 2.1] asserts that any global minimizer of the REAPER problem (2.5) spans



(a) Check (3.108) and (3.122), $c = 1$

(b) Check (3.137) for REAPER, $c = 1$

(c) Check (3.108) and (3.122), $c = 5$

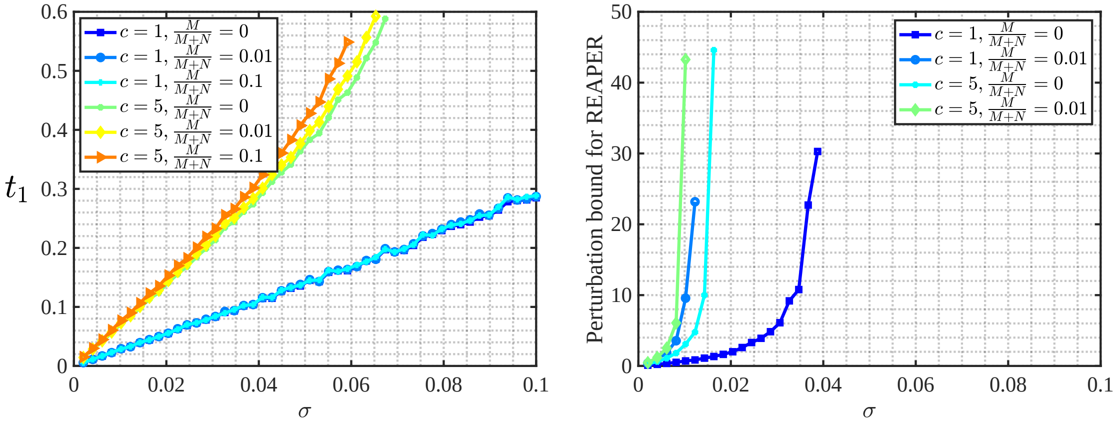
(d) Check (3.137) for REAPER, $c = 5$

Figure 3.6. Check whether (3.108) and (3.122) for DPCP and (3.137) for REAPER [61] are satisfied (white) or not (black) when varying the outlier ratio $M/(M+N)$ and σ. Here we fix $D = 30$ and $N = 1500$.

a subspace that has an angle θ^* from $\mathcal{S}$ satisfying

$$\sin(\theta^*) \le \frac{2N\mathscr{R}(\mathcal{S})}{\left[\frac{N}{4\sqrt{d}}c_{\widehat{\boldsymbol{\mathcal{X}}},\min} - M\mathscr{A}(\mathcal{S}) - N\mathscr{R}(\mathcal{S})\right]_+}, \tag{3.136}$$

where $\mathscr{R}(\mathcal{S}) := \frac{1}{N}\sum_{j=1}^{N}\|\widehat{\boldsymbol{\epsilon}}_j\|_2$ is the total inlier residual (recall that it is an upper bound for $c_{\widehat{\boldsymbol{\mathcal{E}}},\max}$ or $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}$), $\mathscr{A}(\mathcal{S}) := \frac{1}{M}\|\boldsymbol{\mathcal{O}}\|_2\|\overline{\mathcal{P}_{\mathcal{S}^\perp}\boldsymbol{\mathcal{O}}}\|_2 \ge 0$ is an *alignment statistic* that measures the amount of linear structure in the outliers, and $[a]_+ = a$ if $a > 0$ and 0 otherwise. Here $\mathcal{P}_{\mathcal{S}^\perp}$ is the orthoprojection onto $\mathcal{S}^\perp$ and the overline spherization



(a) t_1 in Theorem 7

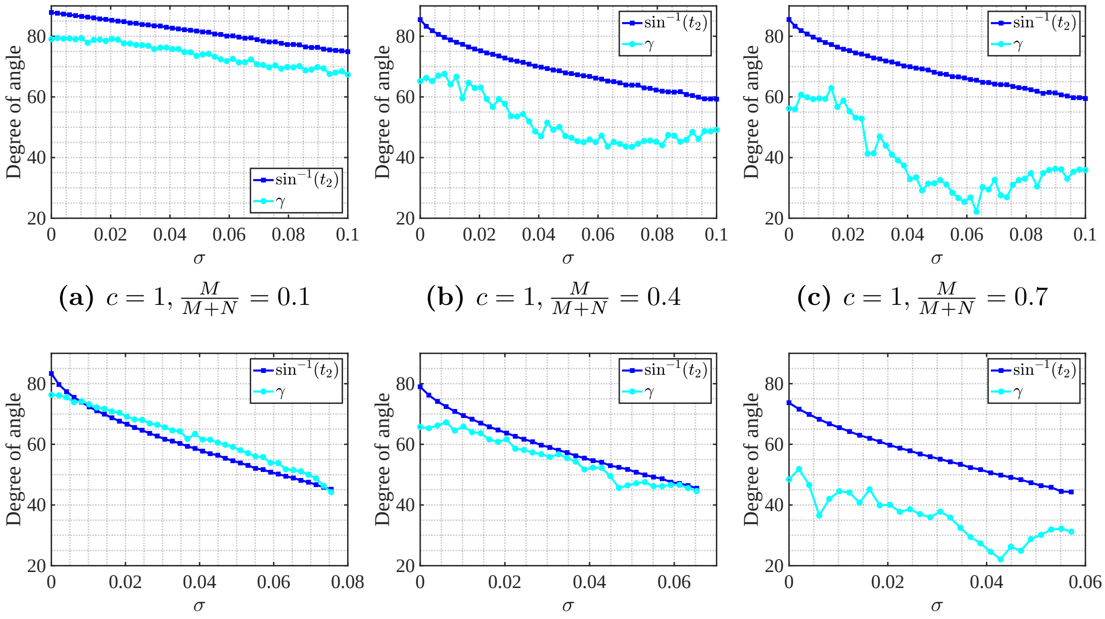
(b) Upper bound for $\sin(\theta^*)$ in (3.136)

Figure 3.7. Evaluation of (a) t_1 in Theorem 7 and (b) upper bound for $\sin(\theta^*)$ in (3.136), with $D = 30$ and $N = 1500$. In (b), we only plot (3.136) for $\frac{M}{M+N} \in \{0, 0.01\}$ since it is only meaningful for a mild size of the outlier ratio.

operator normalizes the columns of a matrix. To make (3.136) meaningful, we require

$$\frac{M\mathscr{A}(\mathcal{S})}{Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min}} < \frac{1}{4\sqrt{d}} - R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}. \tag{3.137}$$

We compare the necessary condition (3.137) for REAPER to (3.108) and (3.122) for the DPCP problem (3.75) (see Theorem 7). In a special case that there are no outliers, i.e., $\mathscr{A}(\mathcal{S}) = 0$, (3.137) requires $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c} < \frac{1}{4\sqrt{d}}$. By contrast, (3.108) only requires $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}}} < \frac{1}{4}$ (see Figure 3.4). More generally, in the presence of outliers, $M\mathscr{A}(\mathcal{S})$ in (3.137) scales as $O(M)$ under the Haystack model [61], whereas the quantity $M(c_{\boldsymbol{\mathcal{O}},\max,c} - c_{\boldsymbol{\mathcal{O}},\min,c})$ in (3.122) scales as $O(\sqrt{M})$ as proved in Lemma 13, indicating that the theoretical analysis for DPCP potentially tolerates more outliers. Numerically, this is captured in Figure 3.6, in which we observe that (3.108) and (3.122) are satisfied for a much larger range of outlier ratio and noise levels. Finally, note that $\mathscr{R}(\mathcal{S})$ appears both in the numerator and denominator in the RHS of (3.136), which makes the entire upper



(a) $c = 1, \frac{M}{M+N} = 0.1$ **(b)** $c = 1, \frac{M}{M+N} = 0.4$ **(c)** $c = 1, \frac{M}{M+N} = 0.7$

(d) $c = 5, \frac{M}{M+N} = 0.1$ **(e)** $c = 5, \frac{M}{M+N} = 0.4$ **(f)** $c = 5, \frac{M}{M+N} = 0.7$

Figure 3.8. Comparison between the quantity γ of [76] and $\sin^{-1}(t_2)$ in the cases of $c = 1$ (top row) and $c = 5$ (bottom row) with outlier ratio $\frac{M}{M+N} \in \{0.1, 0.4, 0.7\}$. Here we fix $D = 30$ and $N = 3000$.

bound blow up quickly when the noise level increases; see Figure 3.7b. In contrast, according to Theorem 7 and (3.121), the upper bound for $\sin(\theta^*)$ in our analysis, i.e., t_1, is roughly proportional to the effective noise level (see Figure 3.7a), and thus provides more insight into the problem.

Comparison with the local optimality conditions of [76]. The GGD paper [76] only provides local optimality analysis for the problem (2.6), which is exactly the dual form of the holistic DPCP problem (3.75) considered in this chapter. [76, Theorem 2] asserts that, given $0 < \eta < \gamma < \pi/2$ such that a certain stability condition holds, any critical point of (2.6) spans a subspace that has an angle θ from $\mathcal{S}$ satisfying

$$\theta < \eta \quad \text{or} \quad \theta > \gamma. \tag{3.138}$$

Note that Lemma 14 has similar statements in characterizing the geometry of the critical points for (3.75). Particularly, in (3.109) we have

$$\theta \leq \sin^{-1}(t_1) \quad \text{or} \quad \theta \geq \sin^{-1}(t_2). \tag{3.139}$$

For both results, a tighter analysis corresponds to a smaller η or $\sin^{-1}(t_1)$ (closer to 0) and a larger γ or $\sin^{-1}(t_2)$ (closer to $\pi/2$) so that the geometric distribution of the critical points are more restricted. As a simulation, we compare (3.138) and (3.139) by manually setting η equal to $\sin^{-1}(t_1)$ and then compare $\sin^{-1}(t_2)$ and γ. Figure 3.8 shows the comparison between γ and $\sin^{-1}(t_2)$ under different codimensions, and percentages of outlier ratio and noise levels. In most of the cases, we can observe that $\sin^{-1}(t_2)$ is larger than γ by a significant amount, under the restriction that η is equal to $\sin^{-1}(t_1)$, thus suggesting that (3.139) is a tighter result compared with (3.138). Moreover, (3.138) is sensitive to the variation of the outliers, while (3.139) is rather stable (compare Figure 3.8a to Figure 3.8c and Figure 3.8d to Figure 3.8f). Finally, we mention that the relationship between η and γ in [76] is captured by the complex inlier-outlier stability statistic, which is not as clear as for our t_1 and t_2, with the latter being explicitly defined by (3.110) and (3.111). In conclusion, we believe that Lemma 14 represents a theoretical and practical improvement over the characterization of the critical points of (3.75) given previously by [76] for a dual formulation of the problem.

Chapter 4

Efficient Algorithms for Learning a Single Subspace with DPCP

We have established the theory of DPCP for learning a single subspace with any codimension under both noiseless and noisy settings in Chapter 3. Nevertheless, the existing algorithms (and their convergence theory) for DPCP are designed the case of codimension equal to 1 and noiseless data. The other scenarios (i.e., codimension larger than 1 and noisy data) call for the design of a unified algorithmic framework that is both scalable and emits a convergence theory for all of the above cases.

In this chapter, we focus on a linearly convergent method for non-smooth non-convex optimization on the Grassmannian, which will cover robust subspace learning via DPCP as a particular application. In Section 4.1, we briefly introduce optimization on the Grassmannian along with the necessary background and notation. In Section 4.2, we present a Projected Riemannian Sub-Gradient Method (PRSGM) with linear convergence guarantees, and show that PRSGM applied to the holistic DPCP problem (2.9) can provably recover a basis (respectively, an approximate basis) for the orthogonal complement of the underlying subspace in the noiseless setting (respectively, noisy setting). Experiments using synthetic and real data in Section 4.3 demonstrate the effectiveness and superiority of PRSGM.

4.1 Introduction

Optimization problems on the Grassmannian $\mathbb{G}(D, c)$ (a.k.a. the Grassmann manifold consisting of the linear c-dimensional subspaces in $\mathbb{R}^D$) appear in a wide variety of applications. A problem of interest in this thesis is a robust subspace recovery problem, namely learning a d-dimensional subspace $\mathcal{S} \subset \mathbb{R}^D$ from corrupted data. As discussed in previous chapters, the original DPCP problem (2.2) involves optimization on the sphere ($\mathbb{G}(D, 1)$), the holistic DPCP problem (2.9) estimates an entire orthonormal basis for $\mathcal{S}^\perp$ by optimizing over $\mathbb{G}(D, c)$ (recall that $c = D - d$ is the codimension of the underlying subspace), and the GGD problem (2.6) learns an orthonormal basis for $\mathcal{S}$ by optimizing on $\mathbb{G}(D, d)$. A key challenge to such problems is that the optimization problems are non-convex since the Grassmannian is a non-convex set.

One approach to solving optimization problems on the Grassmannian is to exploit the fact that the Grassmannian is a Riemannian manifold, and develop generic Riemannian optimization techniques. When the objective function is twice differentiable, [8] shows that Riemannian gradient descent and Riemannian trust-region methods converge to first- and second-order stationary solutions, respectively. Newton algorithms on the Grassmannian have been developed in [34]. When Riemannian gradient descent is randomly initialized, [60] further shows that it converges to a second-order stationary solution almost surely, but without any guarantee on the convergence rate. Non-smooth trust region algorithms [47], gradient sampling methods [22, 23], and proximal gradient methods [18] have also been proposed for non-smooth manifold optimization when the objective function is not continuously differentiable. However, the available theoretical results establish convergence to stationary points from an arbitrary initialization with either no rate of convergence guarantee, or at best a sublinear rate[1].

On the other hand, when the constraint set is convex, [25, 26, 65] show that subgradient methods can handle non-smooth and non-convex objective functions as long as the problem satisfies certain regularity conditions called *sharpness* and *weak convexity*. In such a case, R-linear convergence[1] is guaranteed (e.g., see robust phase retrieval [33] and robust low-rank matrix recovery [65]). Analogous to other regularity conditions for smooth problems, such as the regularity condition of [14] and the error bound condition in [73], sharpness and weak convexity capture regularity properties of non-convex and non-smooth optimization problems. However, these two properties have not yet been exploited for solving problems on the Grassmannian, or other non-convex manifolds.

A related regularity condition, which in this thesis we call the Riemannian Regularity Condition (RRC), has been exploited for orthogonal dictionary learning (ODL)

[1] Suppose the sequence $\{\boldsymbol{x}_k\}$ converges to $\boldsymbol{x}^\star$. We say it converges sublinearly if $\lim_{k\to\infty} \|\boldsymbol{x}_{k+1} - \boldsymbol{x}^\star\|/\|\boldsymbol{x}_k - \boldsymbol{x}^\star\| = 1$, and R-linearly if there exists $C > 0, q \in (0,1)$ such that $\|\boldsymbol{x}_k - \boldsymbol{x}^\star\| \leq Cq^k$, $\forall k \geq 0$.

[3], which solves an ℓ_1 minimization problem on the sphere, a manifold parameterizing $\mathbb{G}(D,1)$. However, under this RRC, Projected Riemannian Sub-Gradient Methods have only been proved to converge at a *sublinear* rate. On the other hand, a Projected Sub-Gradient Method (DPCP-PSGM) [152, 153] has been successfully used and proved to converge at a *piecewise linear* rate for the DPCP problem (2.2). However, i) it is restricted to optimization on the sphere ($\mathbb{G}(D,1)$) even for subspaces of codimension higher than 1, so it may not be applicable to problem (2.9); (ii) it has only be shown to converge to a basis element of $\mathcal{S}^\perp$ with noiseless data; and (iii) the convergence analysis does not reveal the origin of the improved convergence rate.

4.1.1 Background

In this chapter, we consider minimization problems on the Grassmannian $\mathbb{G}(D,c)$. We adopt the same notation as in Section 3.2.1. In particular, we parameterize points on the Grassmannian by representing an element of $\mathbb{G}(D,c)$ by an orthonormal matrix in $\mathbb{O}(D,c) = \{\boldsymbol{B} \in \mathbb{R}^{D\times c} : \boldsymbol{B}^\top \boldsymbol{B} = \mathbf{I}_c\}$, which is also the well-known Stiefel manifold. When $D = c$, we denote $\mathbb{O}(c,c)$ by $\mathbb{O}(c)$, the orthogonal group. This matrix representation is not unique since $\operatorname{Span}(\boldsymbol{BQ}) = \operatorname{Span}(\boldsymbol{B})$ for any $\boldsymbol{Q} \in \mathbb{O}(c)$. Thus, we say that $\{\boldsymbol{A}, \boldsymbol{B}\} \subset \mathbb{G}(D,c)$ are equivalent if $\operatorname{Span}(\boldsymbol{A}) = \operatorname{Span}(\boldsymbol{B})$. With this understanding, we use $\boldsymbol{B}$ to represent the equivalence class $[\boldsymbol{B}] = \{\boldsymbol{BQ} : \boldsymbol{Q} \in \mathbb{O}(c)\}$ and consider the parameterized problem studied in [34, 49] give by

$$\operatorname*{minimize}_{\boldsymbol{B}\in\mathbb{O}(D,c)} \; f(\boldsymbol{B}) \tag{4.1}$$

where $f : \mathbb{R}^{D\times c} \to \mathbb{R}$ is lower semi-continuous, possibly non-convex and non-smooth, and invariant to the action of $\mathbb{O}(c)$, i.e., $f(\boldsymbol{B}) = f(\boldsymbol{BQ})$ for any $\boldsymbol{Q} \in \mathbb{O}(c)$. Again, the global minimum of (4.1) is not unique since if $\boldsymbol{B}^*$ is a global minimum, then any point in $[\boldsymbol{B}^*]$ is also a global minimum.

For any $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{O}(D, c)$, as specified in Definition 2, the principal angles between $\mathrm{Span}(\boldsymbol{A})$ and $\mathrm{Span}(\boldsymbol{B})$ are defined as $\theta_i(\boldsymbol{A}, \boldsymbol{B}) = \arccos(\sigma_i(\boldsymbol{A}^\top \boldsymbol{B}))$ for $i = 1, \ldots, c$, where $\sigma_i(\cdot)$ denotes the i-th largest singular value. As before, the largest principal angle $\theta_c(\boldsymbol{A}, \boldsymbol{B})$ is referred to as the subspace angle between $\mathrm{Span}(\boldsymbol{A})$ and $\mathrm{Span}(\boldsymbol{B})$. We then define the distance between $\boldsymbol{A}$ and $\boldsymbol{B}$ as

$$\mathrm{dist}(\boldsymbol{A}, \boldsymbol{B}) := \sqrt{2\sum_{i=1}^{c}\Big(1 - \cos(\theta_i(\boldsymbol{A}, \boldsymbol{B}))\Big)} = \min_{\boldsymbol{Q}\in\mathbb{O}(c)} \|\boldsymbol{B} - \boldsymbol{AQ}\|_F \tag{4.2}$$

where the last term is also known as the *orthogonal Procrustes problem.* The second equality in (4.2) follows from the result [51] according to which the optimal rotation matrix $\boldsymbol{Q}$ minimizing $\|\boldsymbol{B} - \boldsymbol{AQ}\|_F$ is $\boldsymbol{Q}^* = \boldsymbol{UV}^\top$, where $\boldsymbol{U\Sigma V}^\top$ is the SVD of $\boldsymbol{A}^\top\boldsymbol{B}$. Thus, $\mathrm{dist}(\boldsymbol{A}, \boldsymbol{B}) = 0$ iff $\mathrm{Span}(\boldsymbol{A}) = \mathrm{Span}(\boldsymbol{B})$. We also define the projection of $\boldsymbol{B}$ onto $[\boldsymbol{A}]$ as

$$\mathcal{P}_{\boldsymbol{A}}(\boldsymbol{B}) = \boldsymbol{AQ}^*, \quad \text{where} \quad \boldsymbol{Q}^* = \mathop{\arg\min}_{\boldsymbol{Q}\in\mathbb{O}(c)} \|\boldsymbol{B} - \boldsymbol{AQ}\|_F.$$

Here $\boldsymbol{AQ}^*$ is in $[\boldsymbol{A}]$, with $\boldsymbol{Q}^*$ representing a nonlinear transformation of $\boldsymbol{A}^\top\boldsymbol{B}$, as described above. The following result implies that $\theta_c(\boldsymbol{A}, \boldsymbol{B})$ and $\mathrm{dist}(\boldsymbol{A}, \boldsymbol{B})$ are equivalent in characterizing how close $\boldsymbol{A}$ and $\boldsymbol{B}$ are to each other.

Proposition 2. *The definition* (4.2) *of* $\operatorname{dist}(\boldsymbol{A}, \boldsymbol{B})$ *is equivalent to the subspace angle* $\theta_c(\boldsymbol{A}, \boldsymbol{B})$ *in measuring the similarity between* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *in the following sense:*

$$\sin(\theta_c(\boldsymbol{A}, \boldsymbol{B})) \le \operatorname{dist}(\boldsymbol{A}, \boldsymbol{B}) \le \sqrt{2c} \cdot \sin(\theta_c(\boldsymbol{A}, \boldsymbol{B})).$$

Proof. To prove the first inequality, we have

$$\begin{aligned}
\operatorname{dist}(\boldsymbol{A}, \boldsymbol{B}) &= \sqrt{2\sum_{i=1}^{c}(1 - \cos(\theta_i(\boldsymbol{A}, \boldsymbol{B})))} = \sqrt{\sum_{i=1}^{c} 4\sin^2\left(\theta_i(\boldsymbol{A}, \boldsymbol{B})/2\right)} \\
&\ge 2\sin(\theta_c(\boldsymbol{A}, \boldsymbol{B})/2) \\
&\ge 2\sin(\theta_c(\boldsymbol{A}, \boldsymbol{B})/2)\cos(\theta_c(\boldsymbol{A}, \boldsymbol{B})/2) \\
&= \sin(\theta_c(\boldsymbol{A}, \boldsymbol{B})).
\end{aligned}$$

To prove the second inequality, we have

$$\begin{aligned}
\operatorname{dist}(\boldsymbol{A}, \boldsymbol{B}) &= \sqrt{\sum_{i=1}^{c} 4\sin^2\left(\theta_i(\boldsymbol{A}, \boldsymbol{B})/2\right)} \\
&\le 2\sqrt{c} \cdot \sin(\theta_c(\boldsymbol{A}, \boldsymbol{B})/2) \\
&\le \sqrt{2c} \cdot \sin(\theta_c(\boldsymbol{A}, \boldsymbol{B}))
\end{aligned}$$

where we used the fact that $\sin(a/2) \le \frac{\sin(a)}{\sqrt{2}}$ for $a \in [0, \pi/2]$. □

Since f can be non-smooth and non-convex, we utilize the Clarke subdifferential, which generalizes the gradient for smooth functions and the subdifferential in convex analysis. The Clarke subdifferential [3] of a locally Lipschitz function f at $\boldsymbol{B}$ is

$$\partial f(\boldsymbol{B}) := \operatorname{conv}\left\{\lim_{i\to\infty} \nabla f(\boldsymbol{B}_i) : \boldsymbol{B}_i \to \boldsymbol{B}, f \text{ differentiable at } \boldsymbol{B}_i\right\}$$

where conv denotes the convex hull. When f is differentiable at $\boldsymbol{B}$, its Clarke subdifferential is simply $\{\nabla f(\boldsymbol{B})\}$. When f is not differentiable at $\boldsymbol{B}$, the Clarke subdifferential is the convex hull of the limit of gradients taken at differentiable points. Note that the Clarke subdifferential $\partial f(\boldsymbol{B})$ is a nonempty and convex set since a locally Lipschitz function is differentiable almost everywhere.

Since we consider problems on the Grassmannian, we use tools from Riemannian geometry to state optimality conditions. From [34], the tangent space of the Grassmannian at $[\boldsymbol{B}]$ is defined as $T_{\boldsymbol{B}} := \{\boldsymbol{W} \in \mathbb{R}^{D\times c} : \boldsymbol{W}^\top \boldsymbol{B} = \boldsymbol{0}\}$, and the orthogonal projector onto the tangent space is $\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top$, which is well-defined and does not depend on the class representative since $\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{B}\boldsymbol{B}^\top$ for any $\boldsymbol{A} \in [\boldsymbol{B}]$. We generalize the definition of the Clarke subdifferential and denote by $\widetilde{\partial} f$ the Riemannian subdifferential of f [3]:

$$\widetilde{\partial} f(\boldsymbol{B}) := \operatorname{conv}\left\{\lim_{i\to\infty} (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\nabla f(\boldsymbol{B}_i) : \boldsymbol{B}_i \to \boldsymbol{B}, f \text{ differentiable at } \boldsymbol{B}_i\right\}.$$

We say that $\boldsymbol{B}$ is a critical point of (4.1) if and only if $\boldsymbol{0} \in \widetilde{\partial} f(\boldsymbol{B})$, which is a necessary condition for being a minimizer to (4.1).

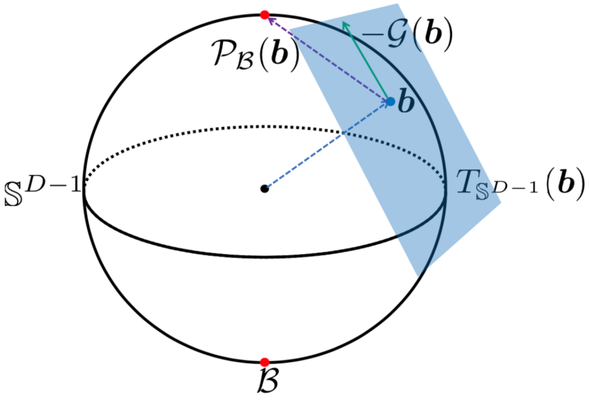

$-\mathcal{G}(\boldsymbol{b})$
$\mathcal{P}_{\mathcal{B}}(\boldsymbol{b})$
$\boldsymbol{b}$
$\mathbb{S}^{D-1}$
$T_{\mathbb{S}^{D-1}}(\boldsymbol{b})$
$\mathcal{B}$

Figure 4.1. Illustration of the Riemannian regularity condition in Definition 3. Red nodes denote $[\boldsymbol{B}^*]$, with the top one closest to $\boldsymbol{B}$. Inequality (4.3) requires the angle between $\mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}) - \boldsymbol{B}$ (purple arrow) and $-\mathcal{G}(\boldsymbol{B})$ (blue arrow) to be sufficiently small.

4.2 Projected Riemannian Sub-Gradient method

In this section, we state our key Riemannian regularitity condition (RRC, Section 4.2.1), propose a Projected Riemannian Sub-Gradient Method (Section 4.2.2) based on RRC, analyze its convergence properties (Section 4.2.3), and show it can be applied to solving the DPCP problem (2.9) under both noiseless and noisy settings (Section 4.2.4).

4.2.1 Riemannian Regularity Condition (RRC)

Definition 3. *Let* $\{\alpha, \epsilon\} > 0$ *and* $\boldsymbol{B}^* \in \mathbb{O}(D, c)$*. We say* $f : \mathbb{R}^{D\times c} \to \mathbb{R}$ *satisfies the* $(\alpha, \epsilon, \boldsymbol{B}^*)$*-Riemannian regularity condition (RRC) if for every* $\boldsymbol{B} \in \mathbb{O}(D, c)$ *satisfying* $\operatorname{dist}(\boldsymbol{B}, \boldsymbol{B}^*) \leq \epsilon$*, there exists a Riemannian subgradient* $\mathcal{G}(\boldsymbol{B}) \in \widetilde{\partial} f(\boldsymbol{B})$ *such that*

$$\langle \mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}) - \boldsymbol{B}, -\mathcal{G}(\boldsymbol{B})\rangle \geq \alpha \operatorname{dist}(\boldsymbol{B}, \boldsymbol{B}^*). \tag{4.3}$$

Strictly speaking, Definition 3 is extrinsic since we view the Grassmannian as embedded in the Euclidean space and (4.3) uses the standard inner product in the Euclidean space. Recently, a particular instance of Definition 3 was shown to hold [3] in the context of ODL. Note that $-\mathcal{G}(\boldsymbol{B})$ is not necessarily a descent direction for all $\mathcal{G}(\boldsymbol{B}) \in \widetilde{\partial} f(\boldsymbol{B})$, and that the set of allowable Riemannian subgradients that satisfy (4.3) need not include the minimum norm element from $\widetilde{\partial} f(\boldsymbol{B})$ even though that one is known to be a descent direction [43]. In Section 4.2.4, we show that a natural choice of Riemannian subgradient satisfies (4.3) for DPCP, where $\boldsymbol{B}^*$ is a target solution. As illustrated in Figure 4.1, condition (4.3) implies that the negative of the chosen Riemannian subgradient $\mathcal{G}(\boldsymbol{B})$ has a positive angle with $\mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}) - \boldsymbol{B}$. To see this, let

$$\xi := \sup \left\{ \|\mathcal{G}(\boldsymbol{B})\|_F : \operatorname{dist}(\boldsymbol{B}, \boldsymbol{B}^*) \leq \epsilon \right\} \tag{4.4}$$

denote an upper bound on the size of the Riemannian subgradients in a neighborhood of $\boldsymbol{B}^*$ (we assume that $\xi < \infty$). From (4.3) we have

$$\frac{\langle \mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}) - \boldsymbol{B}, -\mathcal{G}(\boldsymbol{B}) \rangle}{\|\mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}) - \boldsymbol{B}\|_F \|\mathcal{G}(\boldsymbol{B})\|_F} = \frac{\langle \mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}) - \boldsymbol{B}, -\mathcal{G}(\boldsymbol{B}) \rangle}{\operatorname{dist}(\boldsymbol{B}, \boldsymbol{B}^*) \|\mathcal{G}(\boldsymbol{B})\|_F} \geq \frac{\alpha}{\xi},$$

which gives a bound on the sum of the cosines of the principal angles between $\mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}) - \boldsymbol{B}$ and $-\mathcal{G}(\boldsymbol{B})$ and implies that

$$\xi \geq \alpha. \tag{4.5}$$

In fact, by applying the Cauchy-Schwartz inequality to (4.3), we have

$$\|\mathcal{G}(\boldsymbol{B})\|_F \operatorname{dist}(\boldsymbol{B}, \boldsymbol{B}^*) \geq \langle \boldsymbol{B} - \mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}), \mathcal{G}(\boldsymbol{B}) \rangle \geq \alpha \operatorname{dist}(\boldsymbol{B}, \boldsymbol{B}^*),$$

which leads to

$$\|\mathcal{G}(\boldsymbol{B})\|_F \geq \alpha, \ \forall \boldsymbol{B} \notin [\boldsymbol{B}^*], \operatorname{dist}(\boldsymbol{B}, \boldsymbol{B}^*) \leq \epsilon. \tag{4.6}$$

We will show in Section 4.2.3 that if the $(\alpha, \epsilon, \boldsymbol{B}^*)$-RRC holds, then a Projected Riemannian Sub-Gradient Method will converge to $\boldsymbol{B}^*$ when an appropriate initialization and step size strategy are used.

4.2.1.1 Comparison with regularity conditions for non-smooth functions

Definition 3 is similar in nature to other regularity conditions that characterize geometric properties of the objective function. Perhaps the most closely related ones for non-smooth functions are *sharpness* and *weak convexity*. Consider a function $h : \mathbb{R}^D \to \mathbb{R}$ and assume that the set of global minima

$$\mathcal{X} := \{\boldsymbol{z} \in \mathbb{R}^D : h(\boldsymbol{z}) \leq h(\boldsymbol{x}) \text{ for all } \boldsymbol{x} \in \mathbb{R}^n\} \tag{4.7}$$

is non-empty. Then, h is said to be sharp with parameter $\nu > 0$ (see [12]) if

$$h(\boldsymbol{x}) - \min_{\boldsymbol{z} \in \mathbb{R}^D} h(\boldsymbol{z}) \geq \nu \operatorname{dist}(\boldsymbol{x}, \mathcal{X}) \tag{4.8}$$

holds for all $\boldsymbol{x} \in \mathbb{R}^D$. The function h is said to be weakly convex with parameter $\tau \geq 0$ if $\boldsymbol{x} \mapsto h(\boldsymbol{x}) + \frac{\tau}{2}\|\boldsymbol{x}\|^2$ is convex [122]. If h is both sharp and weakly convex, then [25, 65] show that

$$\langle \mathcal{P}_{\mathcal{X}}(\boldsymbol{x}) - \boldsymbol{x}, -\boldsymbol{d} \rangle \geq \nu \operatorname{dist}(\boldsymbol{x}, \mathcal{X}) - \frac{\tau}{2} \operatorname{dist}^2(\boldsymbol{x}, \mathcal{X}) \tag{4.9}$$

for any $\boldsymbol{x} \in \mathbb{R}^D$ and any $\boldsymbol{d} \in \partial h(\boldsymbol{x})$, where $\mathcal{P}_{\mathcal{X}}$ is the orthogonal projector onto the set $\mathcal{X}$. Note that (4.9) is useful when its RHS is nonnegative, i.e., when $\operatorname{dist}(\boldsymbol{x}, \mathcal{X}) \leq (2\nu)/\tau$. Thus, for any $\epsilon < (2\nu)/\tau$, we have

$$\langle \mathcal{P}_{\mathcal{X}}(\boldsymbol{x}) - \boldsymbol{x}, -\boldsymbol{d} \rangle \geq \left(\nu - \frac{\tau}{2}\epsilon\right) \operatorname{dist}(\boldsymbol{x}, \mathcal{X}) \text{ for all } \boldsymbol{d} \in \partial h(\boldsymbol{x}) \tag{4.10}$$

whenever $\boldsymbol{x}$ satisfies $\operatorname{dist}(\boldsymbol{x}, \mathcal{X}) \leq \epsilon$. Noting the similarity between (4.10) and (4.3) ($\boldsymbol{B}^*$ can be taken as a minimizer of h), the RRC (4.3) can be viewed as a generalization of (4.10) (the consequence of sharpness and weak convexity) to the Riemannian manifold. There are two main differences. First, (4.3) differs from (4.10) in that its LHS involves the Riemannian subgradient due to the Grassmannian constraint. Second, (4.3) is only required to hold for a particular Riemannian subgradient at $\boldsymbol{B}$, while (4.10) holds for all subgradients, thus imposing a slightly stronger regularity condition on the problem.

4.2.1.2 Comparison with regularity conditions for smooth functions

Aside from the weak convexity and sharpness, another regularity condition related to Definition 3 is the one proposed in [14]: we say a continuously differentiable function

$g : \mathbb{R}^D \to \mathbb{R}$ satisfies the $(\alpha, \gamma, \epsilon)$-regularity condition, if for all $\boldsymbol{x} \in \mathbb{R}^D$ such that $\operatorname{dist}(\boldsymbol{x}, \mathcal{X}) \leq \epsilon$, where $\mathcal{X}$ is the set of global minima of g as defined in (4.7), we have

$$\langle \mathcal{P}_{\mathcal{X}}(\boldsymbol{x}) - \boldsymbol{x}, -\nabla g(\boldsymbol{x}) \rangle \geq \alpha \operatorname{dist}^2(\boldsymbol{x}, \mathcal{X}) + \gamma \|\nabla g(\boldsymbol{x})\|^2. \tag{4.11}$$

We now compare (4.3) with (4.11). On the one hand, (4.11) has a form similar to (4.3) as both attempt to provide lower bounds for the inner product between the gradient (or Riemannian subgradient) and the vector $\boldsymbol{x} - \mathcal{P}_{\mathcal{X}}(\boldsymbol{x})$ for any $\boldsymbol{x}$ that is close to $\mathcal{X}$. On the other hand, (4.11) mainly differs from (4.3) in two aspects. First, compared with (4.3), the RHS of (4.11) has an additional term that depends on the magnitude of the gradient, i.e., $\|\nabla g(\boldsymbol{x})\|^2$, while it is impossible to include the Riemannian subgradient term $\|\mathcal{G}(\boldsymbol{B})\|_F$ into the RHS of (4.3) since then as its LHS goes to 0 when $\boldsymbol{B}$ tends to $\boldsymbol{B}^\star$ the term $\|\mathcal{G}(\boldsymbol{B})\|_F$ does not vanish due to (4.6). Moreover, by applying the Cauchy-Schwartz inequality to the LHS of (4.11), we obtain

$$\gamma \|\nabla g(\boldsymbol{x})\|^2 \leq \operatorname{dist}(\boldsymbol{x}, \mathcal{X}) \|\nabla g(\boldsymbol{x})\| - \alpha \operatorname{dist}^2(\boldsymbol{x}, \mathcal{X}),$$

which implies $\|\nabla g(\boldsymbol{x})\| \to 0$ as $\operatorname{dist}(\boldsymbol{x}, \mathcal{X}) \to 0$, hence in sharp contrast to (4.6).

4.2.2 Projected Riemannian Sub-Gradient method on the Grassmannian

We now propose to solve (4.1) using the Projected Riemannian Sub-Gradient Method (PRSGM), which is summarized in Algorithm 1. Given the t-th iterate $\boldsymbol{B}_t$, the next

Algorithm 1 Projected Riemannian Sub-Gradient Method (PRSGM)

```
Initialization: set $\boldsymbol{B}_0$ and $\mu_0$;
for $t = 0, 1, \dots$ do
    Obtain $\mathcal{G}(\boldsymbol{B}_t) \in \widetilde{\partial} f(\boldsymbol{B}_t)$ satisfying (4.3) with $\boldsymbol{B} = \boldsymbol{B}_t$;
    Compute a step size $\mu_t$ according to a certain rule;
    Update the iterate:
        $\widehat{\boldsymbol{B}}_{t+1} \leftarrow \boldsymbol{B}_t - \mu_t \mathcal{G}(\boldsymbol{B}_t)$ and $\boldsymbol{B}_{t+1} \leftarrow \operatorname{orth}(\widehat{\boldsymbol{B}}_{t+1})$;    (4.12)
end for
```

iterate $\boldsymbol{B}_{t+1}$ is obtained by first moving in a direction opposite to a Riemannian subgradient at $\boldsymbol{B}_t$ that satisfies the regularity condition (4.3), and then performing orthonormalization. In Section 4.2.4, we will show that such a projected Riemannian subgradient can be computed for the DPCP problem. We remark that Algorithm 1 is an *extrinsic* method since the iterates are not computed by moving along the Grassmannian; rather, the intermediate point $\widehat{\boldsymbol{B}}_{t+1}$ is projected onto the Grassmannian.

In order to justify (4.12), we show that $\widehat{\boldsymbol{B}}_{t+1}$ in (4.12) always has full column rank given $\boldsymbol{B}_t \in \mathbb{O}(D, c)$. In fact, since $\widehat{\boldsymbol{B}}_{t+1} = \boldsymbol{B}_t - \mu_t \mathcal{G}(\boldsymbol{B}_t)$, we have

$$\widehat{\boldsymbol{B}}_{t+1}^{\top} \widehat{\boldsymbol{B}}_{t+1} = \mathbf{I} + \mu_t^2 \left(\mathcal{G}(\boldsymbol{B}_t)\right)^{\top} \mathcal{G}(\boldsymbol{B}_t), \tag{4.13}$$

where the equality follows because $\boldsymbol{B}_t \in \mathbb{O}(D, c)$ is orthogonal to $\mathcal{G}(\boldsymbol{B}_t)$. Thus, the eigenvalues of $\widehat{\boldsymbol{B}}_{t+1}^{\top} \widehat{\boldsymbol{B}}_{t+1}$ are always greater than or equal to 1. Therefore, all singular values of $\widehat{\boldsymbol{B}}_{t+1}$ are non-vanishing, which means $\widehat{\boldsymbol{B}}_{t+1}$ has full column rank.

Note that there are multiple ways to orthonormalize $\widehat{\boldsymbol{B}}_{t+1}$, although for our purpose they are all equivalent since they all correspond to the same subspace. In (4.12), orth refers to any method that finds an orthonormal basis for $\operatorname{Span}(\widehat{\boldsymbol{B}}_{t+1})$. For example, one can compute $\boldsymbol{B}_{t+1}$ to be the Gram-Schmidt orthonormalization of $\widehat{\boldsymbol{B}}_{t+1}$, or as the first c left singular vectors of $\widehat{\boldsymbol{B}}_{t+1}$. Also, no specific step size rule is provided in Algorithm 1, whereas specific choices are made for the convergence analysis in Section 4.2.3.

4.2.2.1 Connection to the projected subgradient and the geodesic subgradient methods

We now relate the Projected Riemannian Sub-Gradient Method (PRSGM) with the Projected Sub-Gradient Method (PSGM) used in [152, 153] and the Geodesic Gradient Descent (GGD) method used in [76]. In particular, PSGM is developed and analyzed for solving the DPCP problem (2.2) on the sphere, i.e., $\mathbb{O}(D,1)$. Consider $c=1$ in our objective problem (4.1) so that $\{\boldsymbol{B}_t\} \subset \mathbb{O}(D,1)$ in Algorithm 1. First, we claim that PRSGM and PSGM are essentially the same except that the step sizes are scaled differently. For a subgradient $\boldsymbol{d}_t \in \partial f(\boldsymbol{B}_t) \subset \mathbb{R}^D$, the PSGM uses the update

$$\widehat{\boldsymbol{B}}^{\natural}_{t+1} \leftarrow \boldsymbol{B}_t - \mu^{\natural}_t \boldsymbol{d}_t \quad \text{and} \quad \boldsymbol{B}^{\natural}_{t+1} \leftarrow \widehat{\boldsymbol{B}}^{\natural}_{t+1} / \|\widehat{\boldsymbol{B}}^{\natural}_{t+1}\|_2,$$

which is the same as Algorithm 1 except that the Riemannian subgradient $\mathcal{G}(\boldsymbol{B}_t)$ in (4.12) is replaced by the subgradient $\boldsymbol{d}_t$, and $\mu^{\natural}_t$ is the step size for PSGM. To relate $\widehat{\boldsymbol{B}}^{\natural}_{t+1}$ with $\widehat{\boldsymbol{B}}_{t+1}$, we observe that

$$\begin{aligned}\widehat{\boldsymbol{B}}^{\natural}_{t+1} &= \boldsymbol{B}_t - \mu^{\natural}_t \boldsymbol{d}_t = \boldsymbol{B}_t - \mu^{\natural}_t \boldsymbol{B}_t \boldsymbol{B}_t^{\top} \boldsymbol{d}_t - \mu^{\natural}_t \mathcal{G}(\boldsymbol{B}_t) \\ &= (1 - \mu^{\natural}_t \boldsymbol{B}_t^{\top} \boldsymbol{d}_t) \left(\boldsymbol{B}_t - \frac{\mu^{\natural}_t}{1 - \mu^{\natural}_t \boldsymbol{B}_t^{\top} \boldsymbol{d}_t} \mathcal{G}(\boldsymbol{B}_t) \right),\end{aligned}$$

which implies that $\widehat{\boldsymbol{B}}^{\natural}_{t+1}$ is the scaled version of $\widehat{\boldsymbol{B}}_{t+1}$ if we set $\mu_t = \frac{\mu^{\natural}_t}{1-\mu^{\natural}_t \boldsymbol{B}_t^\top \boldsymbol{d}_t}$ in (4.12), or equivalently, $\mu^{\natural}_t = \frac{\mu_t}{1+\mu_t \boldsymbol{B}_t^\top \boldsymbol{d}_t}$. With this choice, $\boldsymbol{B}^{\natural}_{t+1} = \boldsymbol{B}_{t+1}$ if $\mu^{\natural}_t$ is sufficiently small so that $(1 - \mu^{\natural}_t \boldsymbol{B}_t^\top \boldsymbol{d}_t) > 0$. Thus, the convergence guarantee for PRSGM in Section 4.2.3 can be directly applied for PSGM by using the step size $\mu^{\natural}_t = \frac{\mu_t}{1+\mu_t \boldsymbol{B}_t^\top \boldsymbol{d}_t}$, which is close to μ_t as long as μ_t is small.

Both PRSGM and PSGM are extrinsic since the iterates are allowed to move outside the underlying Grassmannian. In contrast, the GGD method proposed in [76] is *intrinsic*, for which the geodesic derivatives are formulated such that the iterates always move along the Grassmannian. With this in mind, consider optimization over $\mathbb{O}(D, 1)$. The geodesic subgradient method uses the update

$$\boldsymbol{B}^{\diamond}_{t+1} \leftarrow \cos(\mu^{\diamond}_t)\boldsymbol{B}_t - \sin(\mu^{\diamond}_t)\tfrac{\mathcal{G}(\boldsymbol{B}_t)}{\|\mathcal{G}(\boldsymbol{B}_t)\|_2}$$

where $\mu^{\diamond}_t$ is the corresponding step size. Note that $\boldsymbol{B}^{\diamond}_{t+1}$ is always on the sphere due to the fact that $\mathcal{G}(\boldsymbol{B}_t)$ is orthogonal to $\boldsymbol{B}_t$. Again, by writing

$$\boldsymbol{B}^{\diamond}_{t+1} = \cos(\mu^{\diamond}_t)\left(\boldsymbol{B}_t - \tan(\mu^{\diamond}_t)\frac{\mathcal{G}(\boldsymbol{B}_t)}{\|\mathcal{G}(\boldsymbol{B}_t)\|_2}\right)$$

and following a similar argument as before, we can see that the convergence analysis for PRSGM in Section 4.2.3 can also be applied to the geodesic subgradient method in this case with the step size $\mu^{\diamond}_t = \arctan(\mu_t \|\mathcal{G}(\boldsymbol{B}_t)\|_2)$.

4.2.3 Convergence analysis

Our convergence analysis for Algorithm 1 relies in the RRC of Definition 3. When this regularity condition holds, we show that the iterates of Algorithm 1 exhibit the following properties: (i) they converge to a neighborhood of the set $\boldsymbol{B}^*$ when a constant step size is used, and (ii) they converge at an R-linear rate to $\boldsymbol{B}^*$ when a geometrically diminishing step size is used.

4.2.3.1 Constant step size

We first consider the convergence of Algorithm 1 when a constant step size is used.

Proposition 3. *Suppose that for some $(\alpha, \epsilon, \boldsymbol{B}^*)$ the function f satisfies the $(\alpha, \epsilon, \boldsymbol{B}^*)$-RRC in Definition 3. Let $\{\boldsymbol{B}_t\}$ be generated by Algorithm 1 with step size*

$$\mu_t \equiv \mu \leq \frac{\alpha\epsilon}{\xi^2}$$

and initial iterate $\boldsymbol{B}_0$ satisfying $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*) \leq \epsilon$, where ξ is defined in (4.4). Then, for all $t \geq 0$, it holds that

$$\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) \leq \max\left\{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*) - \frac{\mu\alpha t}{2}, \frac{\mu\xi^2}{\alpha}\right\}. \tag{4.14}$$

Proof. We have already shown that in Algorithm 1 that $\widehat{\boldsymbol{B}}_{t+1}$ always has full column rank. Let $\widehat{\boldsymbol{B}}_{t+1} = \boldsymbol{P\Omega Q}^\top$ be a reduced SVD of $\widehat{\boldsymbol{B}}_{t+1}$, where $\boldsymbol{\Omega}$ is an $c \times c$ diagonal matrix with singular values $w_1, \ldots, w_c$ along the diagonals. According to (4.13), the eigenvalues of $\widehat{\boldsymbol{B}}_{t+1}^\top \widehat{\boldsymbol{B}}_{t+1}$ are always greater than or equal to 1, which implies that

$w_1, \ldots, w_c \geq 1$. Therefore, for any $\boldsymbol{U} \in \mathbb{O}(D, c)$, it follows that

$$\begin{aligned}
&\|\widehat{\boldsymbol{B}}_{t+1} - \boldsymbol{U}\|_F^2 - \|\boldsymbol{B}_{t+1} - \boldsymbol{U}\|_F^2 \\
&= \|\boldsymbol{P\Omega Q}^\top\|_F^2 - \|\boldsymbol{PQ}^\top\|_F^2 - 2\,\mathrm{trace}((\boldsymbol{\Omega} - \mathbf{I})\boldsymbol{P}^\top\boldsymbol{UQ}) \\
&\geq \sum_{i=1}^c \omega_i^2 - 1 - 2(\omega_i - 1) = \sum_{i=1}^c (\omega_i - 1)^2 \geq 0,
\end{aligned} \tag{4.15}$$

where we have chosen $\boldsymbol{B}_{t+1}$ to be $\boldsymbol{PQ}^\top$, and the last line directly follows Von Neumann's inequality, i.e., $\mathrm{trace}(\boldsymbol{F}^\top\boldsymbol{G}) \leq \sum_i \sigma_i(\boldsymbol{F})\sigma_i(\boldsymbol{G})$ where $\sigma_1(\cdot) \geq \sigma_2(\cdot) \geq \cdots \geq 0$ are the singular values of a matrix.

We prove (4.14) by induction. It is clear that (4.14) holds when $t = 0$. Now assume that (4.14) holds at the t-th iteration, which implies that $\mathrm{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) \leq \epsilon$. Then,

$$\begin{aligned}
&\mathrm{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{B}^*) \\
&\leq \|\boldsymbol{B}_{t+1} - \mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}_t)\|_F^2 \leq \|\widehat{\boldsymbol{B}}_{t+1} - \mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}_t)\|_F^2 \\
&= \|\boldsymbol{B}_t - \mu\mathcal{G}(\boldsymbol{B}_t) - \mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}_t)\|_F^2 \\
&= \|\boldsymbol{B}_t - \mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}_t)\|_F^2 - 2\mu\langle \boldsymbol{B}_t - \mathcal{P}_{\boldsymbol{B}^*}(\boldsymbol{B}_t), \mathcal{G}(\boldsymbol{B}_t)\rangle + \mu^2\|\mathcal{G}(\boldsymbol{B}_t)\|_F^2 \\
&\leq \mathrm{dist}^2(\boldsymbol{B}_t, \boldsymbol{B}^*) - 2\alpha\mu\,\mathrm{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) + \mu^2\xi^2
\end{aligned} \tag{4.16}$$

where the second line uses (4.15), and the last line uses the RRC (4.3).

It is clear from (4.16) that $\mathrm{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{B}^*) \leq \mathrm{dist}^2(\boldsymbol{B}_t, \boldsymbol{B}^*)$ if $\mathrm{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) \geq \frac{\mu\xi^2}{2\alpha}$.

In particular, when $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) \geq \frac{\mu\xi^2}{\alpha}$, we have

$$\begin{aligned}\operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{B}^*) &\leq \operatorname{dist}^2(\boldsymbol{B}_t, \boldsymbol{B}^*) - \alpha\mu \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) + \mu^2\xi^2 - \alpha\mu \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) \\ &\leq \left(\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) - \frac{\mu\alpha}{2}\right)^2,\end{aligned}$$

which implies that

$$\operatorname{dist}(\boldsymbol{B}_{t+1}, \boldsymbol{B}^*) \leq \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) - \frac{\mu\alpha}{2}$$

since $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) \geq \frac{\mu\xi^2}{\alpha} \geq \mu\alpha$.

On the other hand, when $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) \leq \frac{\mu\xi^2}{\alpha}$, it also follows from (4.16) that

$$\begin{aligned}\operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{B}^*) &\leq \max\left\{\left(\frac{\mu\xi^2}{\alpha}\right)^2 - 2\mu\alpha\frac{\mu\xi^2}{\alpha} + \mu^2\xi^2, \mu^2\xi^2\right\} \\ &= \max\left\{\left(\frac{\mu\xi^2}{\alpha}\right)^2 - \mu^2\xi^2, \mu^2\xi^2\right\} \\ &\leq \max\left\{\left(\frac{\mu\xi^2}{\alpha}\right)^2 - \mu^2\xi^2, \mu^2\xi^2\frac{\xi^2}{\alpha^2}\right\} \\ &= \left(\frac{\mu\xi^2}{\alpha}\right)^2\end{aligned}$$

where the first inequality follows from the fact $h(a) := a^2 - 2\alpha\mu a$ is increasing in $[a', \infty]$ for any a' such that $h(a') \geq 0$, and the second inequality utilizes (4.5). Thus by induction, (4.14) holds for all iterations $t \geq 0$. □

Discussion of Proposition 3. Towards interpreting Proposition 3, first consider the case $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*) > \mu\xi^2/\alpha$, in which case (4.14) implies that after at most $T = 2(\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*) - \mu\xi^2/\alpha)/(\mu\alpha)$ iterations, the inequality $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) \leq \mu\xi^2/\alpha$

will hold for all $t \geq T$. In that sense, Proposition 3 essentially says that no further decay of $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*)$ can be guaranteed. This agrees with empirical evidence regarding Algorithm 1 with a constant step size (see Section 4.3). Note that (4.14) also suggests a tradeoff in selecting the step size μ. A larger step size μ leads to a faster decrease on the bound but a larger universal upper bound of $\mu\xi^2/\alpha$, which may even exceed $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)$ if μ is too large.

4.2.3.2 Geometrically diminishing step size

A useful strategy to balance the tradeoff discussed in the case of constant step size is to use a diminishing step size that starts relatively large and decreases to zero as the iterates proceed. As the universal upper bound $\frac{\mu\xi^2}{\alpha}$ in (4.14) is proportional to μ, it is expected that the decay rate of the step size will determine the convergence rate of the iterates. In this section, we consider a geometrically diminishing step size scheme, i.e., we decrease the step size by a fixed fraction between iterations. Our argument is inspired by [25, 65]. Convergence with geometrically diminishing step size is guaranteed by the following result, which shows that if we choose the decay rate and initial step size properly, then the PRSGM converges to $\boldsymbol{B}^*$ at an R-linear rate.

Theorem 9. *Suppose that the function f satisfies the $(\alpha, \epsilon, \boldsymbol{B}^*)$-RRC in Definition 3. Let $\{\boldsymbol{B}_t\}$ be the sequence generated by Algorithm 1 with step size*

$$\mu_t = \mu_0 \beta^t \tag{4.17}$$

and initialization $\boldsymbol{B}_0$ *satisfying* $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*) \leq \epsilon$. *Assume that*

$$\begin{aligned} &\mu_0 \leq \frac{\alpha \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)}{2\xi^2} \quad \textit{and} \\ &\sqrt{1 - 2\frac{\alpha\mu_0}{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)} + \frac{\mu_0^2\xi^2}{\operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{B}^*)}} =: \underline{\beta} \leq \beta < 1, \end{aligned} \tag{4.18}$$

where ξ *is defined in* (4.4). *Then, the sequence* $\{\boldsymbol{B}_t\}$ *satisfies*

$$\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) \leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)\beta^t \quad \textit{for all} \quad t \geq 0. \tag{4.19}$$

Proof. We prove (4.19) by induction. It is clear that (4.19) holds when $t = 0$. Now assume that (4.19) holds at the t-th iteration, which implies that $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) \leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)\beta^t$. Since $\boldsymbol{B}_t$ satisfies the Riemannian regularity condition (4.3), according to the proof of Proposition 3, we know that (4.16) holds:

$$\begin{aligned} \operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{B}^*) &\leq \operatorname{dist}^2(\boldsymbol{B}_t, \boldsymbol{B}^*) - 2\alpha\mu_t \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) + \mu_t^2\xi^2 \\ &= (\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) - \alpha\mu_t)^2 + \mu_t^2(\xi^2 - \alpha^2). \end{aligned} \tag{4.20}$$

From the assumption that $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) \leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)\beta^t$ and

$$\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)\beta^t \geq 2\frac{\mu_0\xi^2}{\alpha}\beta^t \geq 2\alpha\mu_0\beta^t = 2\alpha\mu_t \geq \alpha\mu_t,$$

where the first inequality follows from assumption (4.18) and the second inequality follows from $\xi \geq \alpha$ in (4.5). Therefore, (4.20) achieves its maximum at $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*) =$

$\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)\beta^t$. Plugging this observation into (4.20) gives

$$\begin{aligned}
\operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{B}^*) &\leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)^2\beta^{2t} - 2\alpha\mu_t \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)\beta^t + \mu_t^2\xi^2 \\
&= \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)^2\beta^{2t} - 2\alpha\mu_0 \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)\beta^{2t} + \mu_0^2\beta^{2t}\xi^2 \\
&= \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)^2\beta^{2t}\left(1 - 2\frac{\alpha\mu_0}{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)} + \frac{\mu_0^2\xi^2}{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)^2}\right) \\
&\leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)^2\beta^{2(t+1)}
\end{aligned} \tag{4.21}$$

where the last line holds because $\beta \geq \underline{\beta} = \sqrt{1 - 2\frac{\alpha\mu_0}{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)} + \frac{\mu_0^2\xi^2}{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)^2}}$ in (4.18). Hence, the induction proof is complete. □

Discussion of Theorem 9. The rate at which $\{\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{B}^*)\}_{t\geq 0}$ tends to zero in (4.19) is determined by β, which has to satisfy (4.18). Note that $\underline{\beta}$ is well defined and is strictly less than 1 in (4.18). To see this, on the one hand, $\mu_0 \leq \alpha \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)/2\xi^2$ and $\xi \geq \alpha$ together imply $1 - 2\alpha\mu_0/\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*) \geq 0$. On the other hand, $-2\alpha\mu_0/\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*) + \mu_0^2\xi^2/\operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{B}^*) < 0$ is a decreasing function of μ_0 when $\mu_0 \in (0, \alpha \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)/2\xi^2]$. In particular, when $\mu_0 = \alpha \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)/2\xi^2$, we have $\underline{\beta} = \sqrt{1 - 3\alpha^2/4\xi^2}$, giving the fastest decaying rate by setting $\beta = \underline{\beta}$. Note that if $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{B}^*)$ is not known a priori, then one can replace it by its upper bound ϵ in (4.18) and (4.19) and the results still hold. Finally, we remark that the decaying rate of $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp)$ is determined by the diminishing factor β. A large β may lead to a slow convergence rate while a small β, e.g., smaller than $\underline{\beta}$, may lead to divergence. We will see this tradeoff more clearly with numerical experiments as presented in Section 4.3.

4.2.4 Applications to DPCP

In this section, we show that Algorithm 1 achieves an R-linear convergence rate when applied to the DPCP problem (2.9) for estimating a basis for the orthogonal complement of the underlying subspace under both noiseless (Section 4.2.4.1) and noisy (Section 4.2.4.2) settings.

4.2.4.1 Data corrupted by outliers only

Recall the DPCP problem (2.9) with noiseless data, which was given as

$$\min_{\boldsymbol{B}\in\mathbb{O}(D,c)} f(\boldsymbol{B}) := \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}\right\|_{1,2} \equiv \sum_{j=1}^{L} \left\|\widetilde{\boldsymbol{x}}_j^\top \boldsymbol{B}\right\|_2$$

where $\widetilde{\boldsymbol{\mathcal{X}}} = [\boldsymbol{\mathcal{X}},\ \boldsymbol{\mathcal{O}}]\boldsymbol{\Gamma} \in \mathbb{R}^{D\times L}$ is the dataset with inliers $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{D\times N}$ spanning a d-dimensional subspace $\mathcal{S}$ of $\mathbb{R}^D$, outliers $\boldsymbol{\mathcal{O}} \in \mathbb{R}^{D\times M}$, unknown permutation $\boldsymbol{\Gamma}$, and $c = D - d$ is the codimension of $\mathcal{S}$. We will show that the DPCP problem (2.9) satisfies the RRC, which will then be used to establish convergence rates. Since the objective function f is regular [3], it follows from [140] that $\widetilde{\partial} f(\boldsymbol{B}) = (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\partial f(\boldsymbol{B})$. Also note that the ℓ_2 norm is subdifferentially regular, thus by the chain rule one natural choice for the Riemannian subgradient is

$$\mathcal{G}(\boldsymbol{B}) = (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\sum_{j=1}^{L} \widetilde{\boldsymbol{x}}_j \operatorname{sign}(\widetilde{\boldsymbol{x}}_j^\top \boldsymbol{B}), \tag{4.22}$$

where $\operatorname{sign}(\boldsymbol{a})$ is defined in (3.78). With the geometric quantities $c_{\boldsymbol{\mathcal{X}},\min}$ in (3.4) and $\eta_{\boldsymbol{\mathcal{O}},c}$ in (3.79), we now state a key insight, namely that the DPCP problem (2.9) satisfies the RRC of Definition 3.

Lemma 16. *For any* $\epsilon < \sqrt{2\Big(1 - M\eta_{\boldsymbol{\mathcal{O}},c}/Nc_{\boldsymbol{\mathcal{X}},\min}\Big)}$*, the DPCP problem* (2.9) *satisfies the* $(\alpha, \epsilon, \boldsymbol{S}^\perp)$*-RRC with* $\alpha = ((1 - \epsilon^2/2)Nc_{\boldsymbol{\mathcal{X}},\min} - M\eta_{\boldsymbol{\mathcal{O}},c})/\sqrt{2c}$ *and any orthonormal basis* $\boldsymbol{S}^\perp$ *for* $\mathcal{S}^\perp$*. Also,*

$$\|\mathcal{G}(\boldsymbol{B})\|_F \leq \sqrt{N}\,\|\boldsymbol{\mathcal{X}}\|_2 + M\eta_{\boldsymbol{\mathcal{O}},c},\ \forall \boldsymbol{B} \in \mathbb{O}(D, c) \tag{4.23}$$

where $\|\boldsymbol{A}\|_2$ *denotes the spectral norm of a matrix* $\boldsymbol{A}$*.*

Proof. Let $\boldsymbol{S} \in \mathbb{R}^{D\times d}$ be an orthonormal basis of the subspace $\mathcal{S}$ and let $\boldsymbol{S}^\perp \in \mathbb{R}^{D\times c}$ be an orthonormal basis of the orthogonal complement $\mathcal{S}^\perp$. By utilizing similar decomposition as in (3.84), we have

$$\boldsymbol{B} = \boldsymbol{P}\sin(\boldsymbol{\Theta}) + \boldsymbol{Q}\cos(\boldsymbol{\Theta})$$

where $\boldsymbol{P}$ and $\boldsymbol{Q}$ are orthonormal matrices satisfying $\operatorname{Span}(\boldsymbol{P}) \subseteq \mathcal{S}$ and $\operatorname{Span}(\boldsymbol{Q}) \subseteq \mathcal{S}^\perp$, and $\boldsymbol{\Theta}$ is the diagonal matrix whose diagonal entries $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_c$ are the principal angles between $\operatorname{Span}(\boldsymbol{B})$ and $\mathcal{S}^\perp$. After defining

$$\boldsymbol{G} = \boldsymbol{P}\cos(\boldsymbol{\Theta})\sin(\boldsymbol{\Theta}) - \boldsymbol{Q}\sin^2(\boldsymbol{\Theta}),$$

we have

$$\begin{aligned}
&\langle -\mathcal{G}(\boldsymbol{B}), \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}) - \boldsymbol{B}\rangle = \langle -\mathcal{G}(\boldsymbol{B}), \boldsymbol{Q}\rangle \\
&= -\left\langle (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\left(\sum_{j=1}^{L} \widetilde{\boldsymbol{x}}_j \operatorname{sign}(\widetilde{\boldsymbol{x}}_j^\top \boldsymbol{B})\right), \boldsymbol{Q}\right\rangle \\
&= \left\langle \sum_{j=1}^{N} \boldsymbol{x}_j \operatorname{sign}(\boldsymbol{x}_j^\top \boldsymbol{B}) + \sum_{j=1}^{M} \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B}), \boldsymbol{G}\right\rangle \\
&= \sum_{j=1}^{N} \left\langle \boldsymbol{x}_j^\top \boldsymbol{P} \cos(\boldsymbol{\Theta}) \sin(\boldsymbol{\Theta}), \operatorname{sign}(\boldsymbol{x}_j^\top \boldsymbol{P} \sin(\boldsymbol{\Theta}))\right\rangle \\
&\qquad - \left\langle \boldsymbol{G}, (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top) \sum_{j=1}^{M} \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B})\right\rangle
\end{aligned} \tag{4.24}$$

where the second equality follows from $(\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\boldsymbol{Q} = \boldsymbol{G}$ and the very last line uses the fact that $\boldsymbol{G} \in \operatorname{Span}(\boldsymbol{B}^\perp)$. We bound the first term in the last line of (4.24) by

$$\begin{aligned}
&\sum_{j=1}^{N} \left\langle \boldsymbol{x}_j^\top \boldsymbol{P} \sin(\boldsymbol{\Theta}) \cos(\boldsymbol{\Theta}), \operatorname{sign}(\boldsymbol{x}_j^\top \boldsymbol{P} \sin(\boldsymbol{\Theta}))\right\rangle \\
&\geq \cos(\theta_c) \sum_{j=1}^{N} \left\langle \boldsymbol{x}_j^\top \boldsymbol{P} \sin(\boldsymbol{\Theta}), \operatorname{sign}(\boldsymbol{x}_j^\top \boldsymbol{P} \sin(\boldsymbol{\Theta}))\right\rangle \\
&= \cos(\theta_c) \sum_{j=1}^{N} \left\| \boldsymbol{x}_j^\top \boldsymbol{P} \sin(\boldsymbol{\Theta}) \right\|_2 \geq \cos(\theta_c) \sin(\theta_c) \sum_{j=1}^{N} \left| \boldsymbol{x}_j^\top \boldsymbol{p}_c \right| \\
&\geq \cos(\theta_c) \sin(\theta_c) N c_{\boldsymbol{\mathcal{X}},\min}
\end{aligned} \tag{4.25}$$

where the first inequality follows because $0 \leq \theta_1 \leq \theta_2 \leq \cdots \theta_c \leq \frac{\pi}{2}$, and the last inequality utilizes the definition of $c_{\boldsymbol{\mathcal{X}},\min}$ in (3.4) since $\boldsymbol{p}_c \in \mathcal{S} \cap \mathbb{S}^{D-1}$. On the other hand, the second term in the last line of (4.24) can be bounded by

$$\begin{aligned}
&\left|\left\langle \boldsymbol{G}, (\mathbf{I}-\boldsymbol{B}\boldsymbol{B}^\top)\sum_{j=1}^M \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B})\right\rangle\right| \\
&= \left|\left\langle \boldsymbol{P}\cos(\boldsymbol{\Theta})\sin(\boldsymbol{\Theta}) - \boldsymbol{Q}\sin^2(\boldsymbol{\Theta}), (\mathbf{I}-\boldsymbol{B}\boldsymbol{B}^\top)\sum_{j=1}^M \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B})\right\rangle\right| \\
&\le \sin(\theta_c)\left|\left\langle \boldsymbol{P}\cos(\boldsymbol{\Theta}) - \boldsymbol{Q}\sin(\boldsymbol{\Theta}), (\mathbf{I}-\boldsymbol{B}\boldsymbol{B}^\top)\sum_{j=1}^M \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B})\right\rangle\right| \\
&\le \sin(\theta_c)\left\|(\mathbf{I}-\boldsymbol{B}\boldsymbol{B}^\top)\sum_{j=1}^M \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B})\right\|_F \le \sin(\theta_c) M \eta_{\boldsymbol{\mathcal{O}},c}
\end{aligned} \tag{4.26}$$

where the first inequality follows because $0 \le \theta_1 \le \theta_2 \le \cdots \theta_c \le \frac{\pi}{2}$, the second inequality utilizes the fact that $\boldsymbol{P}\cos(\boldsymbol{\Theta}) - \boldsymbol{Q}\sin(\boldsymbol{\Theta})$ is an orthonormal matrix, and the last inequality follows from the definition of $\eta_{\boldsymbol{\mathcal{O}},c}$ in (3.79).

Plugging the bounds from (4.25) and (4.26) into (4.24), we obtain

$$\langle -\mathcal{G}(\boldsymbol{B}), \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}) - \boldsymbol{B}\rangle \ge \sin(\theta_c)(\cos(\theta_c) N c_{\boldsymbol{\mathcal{X}},\min} - M\eta_{\boldsymbol{\mathcal{O}},c}). \tag{4.27}$$

According to Proposition 2, we know that $\|\mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}) - \boldsymbol{B}\|_F^2 = \operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^\perp)^2 \le 2c\sin^2(\theta_c)$, which together with (4.27) leads to

$$\langle -\mathcal{G}(\boldsymbol{B}), \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}) - \boldsymbol{B}\rangle \ge \frac{\cos(\theta_c) N c_{\boldsymbol{\mathcal{X}},\min} - M\eta_{\boldsymbol{\mathcal{O}},c}}{\sqrt{2c}} \operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^\perp). \tag{4.28}$$

On the other hand, we have

$$\|\boldsymbol{B} - \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B})\|_F^2 = 2\sum_{i=1}^c (1-\cos(\theta_i)) \ge 2(1-\cos(\theta_c)),$$

which implies that $\cos(\theta_c) \geq 1 - \frac{\|\boldsymbol{B}-\mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B})\|_F^2}{2}$. Combining this with (4.28) and $\|\boldsymbol{B} - \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B})\|_F \leq \epsilon$ shows that the DPCP problem (2.9) satisfies the $(\alpha, \epsilon, \boldsymbol{S}^\perp)$-RRC.

Finally, we prove (4.23). Let us denote $\boldsymbol{Z}$ as the matrix with $\left\{\operatorname{sign}(\boldsymbol{x}_j^\top \boldsymbol{B})\right\}_{j=1}^N$ being its rows, so that $\|\boldsymbol{Z}\|_F \leq \sqrt{N}$. Moreover, utilizing the definition of $\eta_{\boldsymbol{\mathcal{O}},c}$ in (3.79), we are able to bound the Riemannian subgradient in (4.22) as

$$\begin{aligned}
\|\mathcal{G}(\boldsymbol{B})\|_F &= \left\|(\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\sum_{j=1}^N \boldsymbol{x}_j \operatorname{sign}(\boldsymbol{x}_j^\top \boldsymbol{B}) + (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\sum_{j=1}^M \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B})\right\|_F \\
&\leq \left\|(\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\boldsymbol{\mathcal{X}}\boldsymbol{Z}\right\|_F + \left\|(\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\sum_{j=1}^M \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B})\right\|_F \\
&\leq \left\|(\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\boldsymbol{\mathcal{X}}\right\|_2 \|\boldsymbol{Z}\|_F + M\eta_{\boldsymbol{\mathcal{O}},c} \\
&\leq \sqrt{N}\,\|\boldsymbol{\mathcal{X}}\|_2 + M\eta_{\boldsymbol{\mathcal{O}},c}
\end{aligned} \tag{4.29}$$

where the second inequality follows from $\|\boldsymbol{A}\boldsymbol{B}\|_F \leq \|\boldsymbol{A}\|_2\|\boldsymbol{B}\|_F$. □

Combining Lemma 16 with Theorem 9 allows us to conclude the linear convergence of Algorithm 1, when applied to problem (2.9) in the noiseless setting, to any orthonormal basis of $\mathcal{S}^\perp$ when a geometrically diminishing step size is used.

Theorem 10. *Suppose that the initialization $\boldsymbol{B}_0$ satisfies*

$$\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) < \sqrt{2\left(1 - M\eta_{\boldsymbol{\mathcal{O}},c}/Nc_{\boldsymbol{\mathcal{X}},\min}\right)},$$

with $\boldsymbol{S}^\perp$ any orthonormal basis for $\mathcal{S}^\perp$. Let $\{\boldsymbol{B}_t\}$ be the sequence generated by Algorithm 1 for solving the DPCP problem (2.9) *with $\mathcal{G}(\boldsymbol{B}_t)$ in* (4.22) *and step size $\mu_t = \mu_0\beta^t$, where μ_0 and β satisfy* (4.18) *with $\epsilon = \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)$, $\alpha = ((1 - \epsilon^2/2)Nc_{\boldsymbol{\mathcal{X}},\min} -$*

$M\eta_{\mathcal{O},c})/\sqrt{2c}$, and $\xi = \sqrt{N}\left\|\boldsymbol{\mathcal{X}}\right\|_2 + M\eta_{\mathcal{O},c}$. Then, it holds that

$$\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t, \ \forall t \geq 0,$$

i.e., $\{\boldsymbol{B}_t\}$ converges to $\boldsymbol{S}^\perp$ at an R-linear rate.

Proof. The proof of Theorem 10 directly follows the RRC for problem (2.9) as stated in Lemma 16 and the convergence result for Algorithm 1 as stated in Theorem 9. □

Discussion of Theorem 10. Theorem 10 implies that the PRSGM applied to problem (2.9) with a good initialization converges to an orthonormal basis of $\mathcal{S}^\perp$ at an R-linear rate, which is a significant improvement over the alternative approach of solving a sequence of c problems of the form (2.2) on $\mathbb{G}(D,1)$. Note that when $c = 1$, PSGM was proved to have a piecewise linear convergence rate in [152, 153]. Nevertheless, in this case, Theorem 10 of PRSGM still improves upon PSGM in three ways: (i) it allows for a simpler strategy for selecting the step size than does the piecewise geometrically diminishing step size, which has two more parameters controlling when and how often to decay the step size; (ii) it provides a more transparent convergence analysis since its proof follows directly from the RRC and Theorem 9; and (iii) it places a slightly weaker requirement on the initialization $\boldsymbol{B}_0$, which in practice is implemented by spectral initialization [112, 152, 153], namely the bottom eigenvectors of $\widetilde{\boldsymbol{\mathcal{X}}}\widetilde{\boldsymbol{\mathcal{X}}}^\top$ are used. The next result provides theoretical guarantees for the spectral initialization in the sense that $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)$ is reasonably small.

Proposition 4. *The spectral initialization $\boldsymbol{B}_0$, which is defined by taking the bottom* *c eigenvectors of $\widetilde{\boldsymbol{\mathcal{X}}}\widetilde{\boldsymbol{\mathcal{X}}}^\top$, satisfies*

$$\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) \leq \sqrt{\frac{\sum_{j=1}^{c} \sigma_j^2(\boldsymbol{\mathcal{O}}) - \sum_{j=D-c+1}^{D} \sigma_j^2(\boldsymbol{\mathcal{O}})}{\sigma_d^2(\boldsymbol{\mathcal{X}})}} \tag{4.30}$$

where $\sigma_\ell(\cdot)$ denotes the ℓ-th largest singular value.

Proof. Note that for any $\boldsymbol{B}$ that is orthogonal to $\mathcal{S}$, we have

$$\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}\|_F^2 = \|\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}\|_F^2 = \operatorname{trace}(\boldsymbol{B}^\top \boldsymbol{\mathcal{O}}\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}) = \sum_{j=1}^{c} \boldsymbol{b}_j^\top \boldsymbol{\mathcal{O}}\boldsymbol{\mathcal{O}}^\top \boldsymbol{b}_j \leq \sum_{j=1}^{c} \sigma_j^2(\boldsymbol{\mathcal{O}}). \tag{4.31}$$

Since $\boldsymbol{B}_0 = \arg\min_{\boldsymbol{B}\in\mathbb{O}(D,c)} \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}\right\|_F^2$, we have

$$\left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}_0\right\|_F^2 \leq \sum_{j=1}^{c} \sigma_j^2(\boldsymbol{\mathcal{O}}). \tag{4.32}$$

On the other hand, let $\boldsymbol{S}$ be an orthonormal basis for $\mathcal{S}$ and let $\boldsymbol{\Phi}$ be the coefficients of $\boldsymbol{\mathcal{X}}$ in $\boldsymbol{S}$, i.e., $\boldsymbol{\mathcal{X}} = \boldsymbol{S}\boldsymbol{\Phi}$. Then, it follows that

$$\begin{aligned}\left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}_0\right\|_F^2 &= \left\|\boldsymbol{\mathcal{X}}^\top \boldsymbol{B}_0\right\|_F^2 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}_0\right\|_F^2 = \left\|\boldsymbol{\mathcal{X}}^\top \boldsymbol{S}\boldsymbol{S}^\top \boldsymbol{B}_0\right\|_F^2 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}_0\right\|_F^2 \\ &= \left\|\boldsymbol{\Phi}^\top \boldsymbol{S}^\top \boldsymbol{B}_0\right\|_F^2 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}_0\right\|_F^2 \geq \sigma_{\min}^2(\boldsymbol{\Phi}) \left\|\boldsymbol{S}^\top \boldsymbol{B}_0\right\|_F^2 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}_0\right\|_F^2 \\ &\geq \sigma_d^2(\boldsymbol{\mathcal{X}}) \left\|\boldsymbol{B}_0 - \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}_0)\right\|_F^2 + \sum_{j=D-c+1}^{D} \sigma_j^2(\boldsymbol{\mathcal{O}})\end{aligned} \tag{4.33}$$

where we first utilize the fact that $\boldsymbol{\mathcal{X}}$ lies in $\mathcal{S}$ so that $\boldsymbol{\mathcal{X}} = \boldsymbol{S}\boldsymbol{S}^\top \boldsymbol{\mathcal{X}}$, the inequality follows because $\|\boldsymbol{A}\boldsymbol{B}\|_F^2 = \operatorname{trace}(\boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{B}\boldsymbol{B}^\top) \geq \sigma_{\min}(\boldsymbol{A}^\top \boldsymbol{A}) \left\|\boldsymbol{B}\boldsymbol{B}^\top\right\|_F$ for any $\boldsymbol{A}, \boldsymbol{B}$, and the last line follows from $\left\|\boldsymbol{S}^\top \boldsymbol{B}_0\right\|_F^2 = \left\|\boldsymbol{S}\boldsymbol{S}^\top \boldsymbol{B}_0\right\|_F^2 = \left\|\boldsymbol{B}_0 - \boldsymbol{S}^\perp(\boldsymbol{S}^\perp)^\top \boldsymbol{B}_0\right\|_F^2 = \|\boldsymbol{B}_0 - \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}_0)\|_F^2$. Combining (4.32), (4.33), and the fact that $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) =$

$\|\boldsymbol{B}_0 - \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}_0)\|_F$, we obtain

$$\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)^2 \leq \frac{\sum_{j=1}^{c} \sigma_j^2(\boldsymbol{\mathcal{O}}) - \sum_{j=D-c+1}^{D} \sigma_j^2(\boldsymbol{\mathcal{O}})}{\sigma_d^2(\boldsymbol{\mathcal{X}})},$$

which completes the proof. □

4.2.4.2 Data corrupted by outliers and noise

It has been shown in Section 4.2.4.1 that the PRSGM applied to the DPCP problem (2.9) with noiseless data converges linearly to an orthonormal basis, say $\boldsymbol{S}^\perp$, of $\mathcal{S}^\perp$. However, the analytical result cannot be immediately generalized to noisy data of the form $\widetilde{\boldsymbol{\mathcal{X}}} = [\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}},\ \boldsymbol{\mathcal{O}}]\boldsymbol{\Gamma}$ with the noise matrix $\boldsymbol{\mathcal{E}} \neq \mathbf{0}$ denoting the additive noise imposed on the inliers $\boldsymbol{\mathcal{X}}$. In this case, one can only expect that PRSGM at best converges to a neighborhood of $\boldsymbol{S}^\perp$ as suggested by the noisy analyses in Section 3.2.3. Note that the convergence analysis of PRSGM is built upon a particular RRC (Definition 3), which is a local geometric property of the problem relative to a point of interest, e.g., $\boldsymbol{S}^\perp$ in our case. In this section, we will show that when data is corrupted by noise, the RRC for (2.9) only holds outside a neighborhood of $\boldsymbol{S}^\perp$ with a radius proportional to the effective noise level, which is then used to show that the PRSGM converges linearly to that neighborhood of $\boldsymbol{S}^\perp$.

In line with the noisy analysis in Section 3.2.3, we reorganize the noisy inliers $\boldsymbol{\mathcal{X}}+\boldsymbol{\mathcal{E}}$ by $\widehat{\boldsymbol{\mathcal{X}}} + \widehat{\boldsymbol{\mathcal{E}}}$ with $\widehat{\boldsymbol{\mathcal{X}}}$ denoting the effective inliers and $\widehat{\boldsymbol{\mathcal{E}}}$ denoting the effective noise, and $\operatorname{Span}(\widehat{\boldsymbol{\mathcal{X}}}) \subseteq \mathcal{S}$ and $\operatorname{Span}(\widehat{\boldsymbol{\mathcal{E}}}) \subseteq \mathcal{S}^\perp$. Also, we will use the geometric quantities introduced in Section 3.1.2 and Section 3.2.3, e.g., $c_{\widehat{\boldsymbol{\mathcal{X}}},\min}$, $c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}$, $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}$ and $R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}$, for the rest of the convergence analysis. The following result gives the RRC for the DPCP problem (2.9) with noisy data.

Lemma 17. *For any $\epsilon > 0$ satisfying*

$$\epsilon\left(1 - R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c} - \epsilon^2/2\right) \geq 4\sqrt{2c} R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}, \tag{4.34}$$

let

$$\alpha := \frac{N c_{\widehat{\boldsymbol{\mathcal{X}}},\min}}{2\sqrt{2c}}\left(\left(1 - \frac{\epsilon^2}{2}\right) - R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}\right).$$

Then for any $\boldsymbol{B} \in \mathbb{O}(D, c)$ satisfying

$$\epsilon \geq \operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^{\perp}) \geq \omega := \frac{2N c_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}}{\alpha} \tag{4.35}$$

and $\mathcal{G}(\boldsymbol{B}) \in \widetilde{\partial} f(\boldsymbol{B})$ defined as in (4.22)*, it holds that*

$$\langle -\mathcal{G}(\boldsymbol{B}), \mathcal{P}_{\boldsymbol{S}^{\perp}}(\boldsymbol{B}) - \boldsymbol{B}\rangle \geq \alpha \operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^{\perp}). \tag{4.36}$$

Also,

$$\|\mathcal{G}(\boldsymbol{B})\|_F \leq \xi := \sqrt{N}\|\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}}\|_2 + M\eta_{\boldsymbol{\mathcal{O}},c}, \ \forall \boldsymbol{B} \in \mathbb{O}(D, c). \tag{4.37}$$

Proof. Let $\boldsymbol{S} \in \mathbb{R}^{D\times d}$ be an orthonormal basis of the subspace $\mathcal{S}$ and let $\boldsymbol{S}^{\perp} \in \mathbb{R}^{D\times c}$ be an orthonormal basis of the orthogonal complement $\mathcal{S}^{\perp}$. By utilizing a similar decomposition as in (3.84), we have

$$\boldsymbol{B} = \boldsymbol{P}\sin(\boldsymbol{\Theta}) + \boldsymbol{Q}\cos(\boldsymbol{\Theta})$$

where $\boldsymbol{P}$ and $\boldsymbol{Q}$ are orthonormal matrices satisfying $\operatorname{Span}(\boldsymbol{P}) \subseteq \mathcal{S}$ and $\operatorname{Span}(\boldsymbol{Q}) \subseteq \mathcal{S}^\perp$, and $\boldsymbol{\Theta}$ is the diagonal matrix whose diagonal entries $\theta_1 \le \theta_2 \le \cdots \le \theta_c$ are the principal angles between $\operatorname{Span}(\boldsymbol{B})$ and $\mathcal{S}^\perp$. Defining

$$\boldsymbol{G} = \boldsymbol{P}\cos(\boldsymbol{\Theta})\sin(\boldsymbol{\Theta}) - \boldsymbol{Q}\sin^2(\boldsymbol{\Theta}),$$

which is orthogonal to $\boldsymbol{B}$, it follows that

$$\begin{aligned}
&\langle -\mathcal{G}(\boldsymbol{B}), \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}) - \boldsymbol{B}\rangle = \langle -\mathcal{G}(\boldsymbol{B}), \boldsymbol{Q}\rangle \\
&= -\left\langle (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\left(\sum_{j=1}^{L} \widetilde{\boldsymbol{x}}_j \operatorname{sign}(\widetilde{\boldsymbol{x}}_j^\top \boldsymbol{B})\right), \boldsymbol{Q}\right\rangle \\
&= -\left\langle \sum_{j=1}^{N}(\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)\operatorname{sign}((\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{B}) + \sum_{j=1}^{M} \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B}), \boldsymbol{G}\right\rangle \\
&= -\left\langle \sum_{j=1}^{N}(\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{G}, \operatorname{sign}((\widehat{\boldsymbol{x}}_j + \widehat{\boldsymbol{\epsilon}}_j)^\top \boldsymbol{B})\right\rangle \\
&\qquad - \left\langle (\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\sum_{j=1}^{M} \boldsymbol{o}_j \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{B}), \boldsymbol{G}\right\rangle
\end{aligned} \tag{4.38}$$

where the second equality follows from $(\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)\boldsymbol{Q} = \boldsymbol{G}$ and the very last line utilizes the fact that $\boldsymbol{G} \in \operatorname{Span}(\boldsymbol{B}^\perp)$.

For the first term in (4.38), according to the proof of Lemma 14, we have

$$\begin{aligned}
&\left|\sum_{j=1}^{N}\left\langle -(\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{G}, \operatorname{sign}((\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{B})\right\rangle - \sum_{j=1}^{N}\left\langle -\widehat{\boldsymbol{x}}_j^\top\boldsymbol{G}, \operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{B})\right\rangle\right| \\
&= \Bigg|\sum_{j=1}^{N}\langle \widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\cos(\boldsymbol{\Theta})\sin(\boldsymbol{\Theta}) - \widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\sin^2(\boldsymbol{\Theta}), \operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\sin(\boldsymbol{\Theta}) + \widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\cos(\boldsymbol{\Theta}))\rangle \\
&\qquad - \sum_{j=1}^{N}\left\langle \widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\cos(\boldsymbol{\Theta})\sin(\boldsymbol{\Theta}), \operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\sin(\boldsymbol{\Theta}))\right\rangle\Bigg| \\
&\le \sum_{j=1}^{N}\left(\frac{2\|\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\|}{|\widehat{\boldsymbol{x}}_j^\top\boldsymbol{p}_c\sin(\theta_c) + \widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{q}_c\cos(\theta_c)|} + \sin(\theta_c)\right)\|\widehat{\boldsymbol{\epsilon}}_j^\top\boldsymbol{Q}\| \le 2Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}.
\end{aligned} \tag{4.39}$$

Moreover, from (4.25), we already know that

$$\begin{aligned}
\sum_{j=1}^{N}\left\langle -\widehat{\boldsymbol{x}}_j^\top\boldsymbol{G}, \operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{B})\right\rangle &= \sum_{j=1}^{N}\left\langle \widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\cos(\boldsymbol{\Theta})\sin(\boldsymbol{\Theta}), \operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{P}\sin(\boldsymbol{\Theta}))\right\rangle \\
&\ge \cos(\theta_c)\sin(\theta_c)Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min}.
\end{aligned} \tag{4.40}$$

Therefore, we obtain

$$\begin{aligned}
&\sum_{j=1}^{N}\left\langle -(\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{G}, \operatorname{sign}((\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{B})\right\rangle \\
&\ge \sum_{j=1}^{N}\left\langle -\widehat{\boldsymbol{x}}_j^\top\boldsymbol{G}, \operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{B})\right\rangle \\
&\qquad - \left|\sum_{j=1}^{N}\left\langle -(\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{G}, \operatorname{sign}((\widehat{\boldsymbol{x}}_j+\widehat{\boldsymbol{\epsilon}}_j)^\top\boldsymbol{B})\right\rangle - \sum_{j=1}^{N}\left\langle -\widehat{\boldsymbol{x}}_j^\top\boldsymbol{G}, \operatorname{sign}(\widehat{\boldsymbol{x}}_j^\top\boldsymbol{B})\right\rangle\right| \\
&\ge \sin(\theta_c)\cos(\theta_c)Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min} - 2Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}.
\end{aligned}$$

On the other hand, the second term in (4.38) is bound by $\sin(\theta_c)M\eta_{\boldsymbol{\mathcal{O}},c}$ as shown in (4.26). Combining these bounds with (4.38), we arrive at

$$\begin{aligned}
&\langle -\mathcal{G}(\boldsymbol{B}), \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}) - \boldsymbol{B}\rangle \\
&\geq \sin(\theta_c)\Big(\cos(\theta_c) N c_{\widehat{\boldsymbol{\mathcal{X}}},\min} - M\eta_{\boldsymbol{\mathcal{O}},c} - D\Big) - 2Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max,c} \\
&\geq \sin(\theta_c) N c_{\widehat{\boldsymbol{\mathcal{X}}},\min}\Big(\cos(\theta_c) - R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}\Big) - 2Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}
\end{aligned}$$

where we used the definition of $R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}$ in (3.107). According to Proposition 2, we know that $\operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^\perp) \leq \sqrt{2c}\sin(\theta_c)$, which allows us to conclude that

$$\begin{aligned}
&\langle -\mathcal{G}(\boldsymbol{B}), \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}) - \boldsymbol{B}\rangle \\
&\geq \frac{Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min}}{\sqrt{2c}}\Big(\cos(\theta_c) - R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}\Big)\operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^\perp) - 2Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}.
\end{aligned}$$

For any $\epsilon > 0$ and $\operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^\perp) \leq \epsilon$, from the definition of $\operatorname{dist}(\cdot,\cdot)$ in (4.2) we have

$$\epsilon \geq \sqrt{2\sum_{i=1}^{c}(1-\cos(\theta_i))} \geq \sqrt{2(1-\cos(\theta_c))} \quad \Rightarrow \quad \cos(\theta_c) \geq 1 - \frac{\epsilon^2}{2}.$$

Hence we obtain

$$\langle -\mathcal{G}(\boldsymbol{B}), \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}) - \boldsymbol{B}\rangle \geq \frac{Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min}}{\sqrt{2c}}\left(\left(1-\frac{\epsilon^2}{2}\right) - R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}\right)\operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^\perp) - 2Nc_{\widehat{\boldsymbol{\mathcal{E}}},\max,c}.$$

We now let $\operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^\perp) \geq \frac{4\sqrt{2c}R_{\widehat{\boldsymbol{\mathcal{E}}}/\widehat{\boldsymbol{\mathcal{X}}},c}}{1-R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}-\epsilon^2/2}$, in which case we have

$$\langle -\mathcal{G}(\boldsymbol{B}), \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}) - \boldsymbol{B}\rangle \geq \frac{Nc_{\widehat{\boldsymbol{\mathcal{X}}},\min}}{2\sqrt{2c}}\left(\left(1-\frac{\epsilon^2}{2}\right) - R_{\boldsymbol{\mathcal{O}}/\widehat{\boldsymbol{\mathcal{X}}},c}\right)\operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^\perp), \tag{4.41}$$

which completes the proof of (4.36).

Finally, we prove (4.37). Similar to the proof of (4.29), we denote $\boldsymbol{Z}$ as the matrix with $\left\{\operatorname{sign}(\boldsymbol{x}_j + \boldsymbol{\epsilon}_j)^\top \boldsymbol{B}\right\}_{j=1}^N$ being its rows; note that $\|\boldsymbol{Z}\|_F \leq \sqrt{N}$. Hence we have

$$\|\mathcal{G}(\boldsymbol{B})\|_F \leq \|(\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^\top)(\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}})\|_2 \|\boldsymbol{Z}\|_F + M\eta_{\boldsymbol{\mathcal{O}},c} \leq \sqrt{N}\|\boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{E}}\|_2 + M\eta_{\boldsymbol{\mathcal{O}},c},$$

which completes the proof. □

Discussion of Lemma 17. First, condition (4.35) specifies both an upper bound and a lower bound that $\operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^\perp)$ needs to satisfy: the upper bound ϵ indicates that the RRC is a local geometric property around $\boldsymbol{S}^\perp$, which is the same as in Lemma 16 when $\boldsymbol{\mathcal{E}} = \mathbf{0}$, while the lower bound ω implies the RRC may not hold within a small radius of $\boldsymbol{S}^\perp$ due to the existence of noise. Note that the lower bound ω for $\operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^\perp)$ leads to a region around $\mathcal{S}^\perp$ inside which the RRC is not guaranteed and its radius ω is proportional to the effective noise level (vanishing as $\boldsymbol{\mathcal{E}} \to \mathbf{0}$), making the entire lemma reduce to the noiseless one as stated in Lemma 16. We remark that (4.34) gives a valid range for ϵ and thus ensures the validity of (4.35). Given $\operatorname{dist}(\boldsymbol{B}, \boldsymbol{S}^\perp) \in [\omega, \epsilon]$, the RRC condition (4.36) states that a negative Riemannian subgradient $-\mathcal{G}(\boldsymbol{B})$ has a small angle with the direction pointing towards $\boldsymbol{S}^\perp$ at $\boldsymbol{B}$.

With the RRC for problem (2.9) for the noisy setting stated in Lemma 17, we provide the convergence analysis for PRSGM (Algorithm 1) to any orthonormal basis of $\mathcal{S}^\perp$ with two different strategies of updating the step size: constant step size and geometrically diminishing step size.

Proposition 5. *Consider α, ϵ, ω and ξ defined in Lemma 17. Suppose the initialization*

$\boldsymbol{B}_0$ *of Algorithm 1 satisfies* $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) \leq \epsilon$*, and let* $\{\boldsymbol{B}_t\}$ *be the iterates generated with constant step size* $\mu_t \equiv \mu$ *satisfying*

$$\mu \leq \frac{\alpha(\epsilon - \omega)}{\xi^2}. \tag{4.42}$$

Then it holds that

$$\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \max\left\{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) - \frac{t\alpha\mu}{2}, \frac{\mu\xi^2}{\alpha} + \omega\right\}. \tag{4.43}$$

Proof. We prove (4.43) by induction. Obviously, it is true when $t = 0$. Next, suppose it holds for the t-th iteration, i.e., that

$$\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \max\left\{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) - \frac{t\alpha\mu}{2}, \frac{\mu\xi^2}{\alpha} + w\right\}. \tag{4.44}$$

From (4.42) and $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) \leq \epsilon$ we know that $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \epsilon$. It now follows for the $(t+1)$-th iteration that

$$\begin{aligned} &\operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp) \leq \|\boldsymbol{B}_{t+1} - \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}_t)\|_F^2 \\ &\leq \|\widehat{\boldsymbol{B}}_{t+1} - \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}_t)\|_F^2 \\ &= \|\boldsymbol{B}_t - \mu\mathcal{G}(\boldsymbol{B}_t) - \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}_t)\|_F^2 \\ &= \|\boldsymbol{B}_t - \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}_t)\|_F^2 - 2\mu\langle \boldsymbol{B}_t - \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}_t), \mathcal{G}(\boldsymbol{B}_t)\rangle + \mu^2\|\mathcal{G}(\boldsymbol{B}_t)\|_F^2 \end{aligned} \tag{4.45}$$

where the second line follows from (4.15).

Case (I): $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \geq \omega$. Utilizing the Riemannian regularity condition (4.36), from the last line in (4.45) we obtain

$$\operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp) \le \operatorname{dist}^2(\boldsymbol{B}_t, \boldsymbol{S}^\perp) - 2\mu\alpha \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) + \mu^2\xi^2. \tag{4.46}$$

It is clear that $\operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp) \le \operatorname{dist}^2(\boldsymbol{B}_t, \boldsymbol{S}^\perp)$ if $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \ge \frac{\mu\xi^2}{2\alpha}$. In particular, when $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \ge \frac{\mu\xi^2}{\alpha}$, we have

$$\begin{aligned}
\operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp) &\le \operatorname{dist}^2(\boldsymbol{B}_t, \boldsymbol{S}^\perp) - \alpha\mu \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) - \alpha\mu \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) + \mu^2\xi^2 \\
&= \left(\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) - \frac{\mu\alpha}{2}\right)^2 - \alpha\mu \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) + \mu^2\xi^2 - \frac{\mu^2\alpha^2}{4} \\
&\le \left(\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) - \frac{\mu\alpha}{2}\right)^2,
\end{aligned}$$

which implies that

$$\operatorname{dist}(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp) \le \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) - \frac{\mu\alpha}{2} \tag{4.47}$$

since $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \ge \frac{\mu\xi^2}{\alpha} \ge \mu\alpha$ due to (4.5).

On the other hand, when $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) < \frac{\mu\xi^2}{\alpha}$, it also follows (4.46) that

$$\begin{aligned}
\operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp) &\le \max\left\{\left(\frac{\mu\xi^2}{\alpha}\right)^2 - 2\mu\alpha\frac{\mu\xi^2}{\alpha} + \mu^2\xi^2, \mu^2\xi^2\right\} \\
&= \max\left\{\left(\frac{\mu\xi^2}{\alpha}\right)^2 - \mu^2\xi^2, \mu^2\xi^2\right\} \\
&\le \max\left\{\left(\frac{\mu\xi^2}{\alpha}\right)^2 - \mu^2\xi^2, \mu^2\xi^2\frac{\xi^2}{\alpha^2}\right\} \\
&\le \left(\frac{\mu\xi^2}{\alpha}\right)^2,
\end{aligned}$$

where the first inequality follows from the fact that $h(a) := a^2 - 2\alpha\mu a + \mu^2\xi^2$ is upper bounded by $\max\{h(\mu\xi^2/\alpha), h(0)\}$ when $\frac{\mu\xi^2}{\alpha} \geq \mu\alpha$, and the second inequality utilizes (4.5). This tells us that

$$\operatorname{dist}(\boldsymbol{B}_{t+1}, \boldsymbol{S}^{\perp}) \leq \frac{\mu\xi^2}{\alpha}. \tag{4.48}$$

Combining (4.44), (4.47) and (4.48), we obtain that

$$\operatorname{dist}(\boldsymbol{B}_{t+1}, \boldsymbol{S}^{\perp}) \leq \max\left\{\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^{\perp}) - \frac{(t+1)\alpha\mu}{2}, \frac{\mu\xi^2}{\alpha} + \omega\right\}. \tag{4.49}$$

Case (II): $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^{\perp}) < \omega$. The assumptions for RRC in Lemma 17 do not hold, but we can bound the last line in (4.45) such that

$$\begin{aligned}\operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{S}^{\perp}) &\leq \operatorname{dist}^2(\boldsymbol{B}_t, \boldsymbol{S}^{\perp}) + 2\mu \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^{\perp})\xi + \mu^2\xi^2 \\ &< \omega^2 + 2\mu\omega\xi + \mu^2\xi^2 = (\mu\xi + \omega)^2 \\ &\leq \left(\mu\xi\frac{\xi}{\alpha} + \omega\right)^2 = \left(\frac{\mu\xi^2}{\alpha} + \omega\right)^2,\end{aligned} \tag{4.50}$$

where the last line utilizes (4.5), which implies that

$$\operatorname{dist}(\boldsymbol{B}_{t+1}, \boldsymbol{S}^{\perp}) < \frac{\mu\xi^2}{\alpha} + \omega. \tag{4.51}$$

Combining (4.49) and (4.51), we have

$$\operatorname{dist}(\boldsymbol{B}_{t+1}, \boldsymbol{S}^{\perp}) \leq \max\left\{\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^{\perp}) - \frac{(t+1)\alpha\mu}{2}, \frac{\mu\xi^2}{\alpha} + \omega\right\},$$

which completes the proof. □

Discussion of Proposition 5. Proposition 5 shows that with a constant step size, Algorithm 1 applied to the noisy DPCP problem (2.9) ensures convergence to a neighborhood of $\boldsymbol{S}^\perp$ if properly initialized. If $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) > \mu\xi^2/\alpha + \omega$, then $\{\boldsymbol{B}_t\}$ will get closer to $\boldsymbol{S}^\perp$ until the iterates enter the region where $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \mu\xi^2/\alpha + \omega$, after which no further decay is guaranteed. Also, a larger step size μ results in faster convergence of $\boldsymbol{B}_t$ to a larger neighborhood of $\boldsymbol{S}^\perp$. Compared with Proposition 3, the valid range for step size μ in (4.42) gets more restricted by an amount $\frac{\alpha}{\xi^2}\omega$ that reflects the influence of the noise. Moreover, the guaranteed neighborhood of convergence in (4.43) with noisy data is enlarged by ω. This makes sense because, according to Lemma 17, the RRC may not hold inside a region around $\boldsymbol{S}^\perp$ with radius ω. Finally, since $\omega \to 0$ as $\boldsymbol{\mathcal{E}} \to \boldsymbol{0}$, the results of Proposition 5 reduce to that of Proposition 3 for problem (2.9) with noiseless data.

We now consider diminishing step sizes.

Theorem 11. *Consider α, ϵ, ω and ξ as defined in Lemma 17. Suppose the initialization $\boldsymbol{B}_0$ of Algorithm 1 satisfies $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) \leq \epsilon$, and let $\{\boldsymbol{B}_t\}$ be the iterates generated with step size*

$$\mu_t = \mu_0\beta^t$$

satisfying

$$\begin{aligned} &\mu_0 \leq \frac{\alpha}{\xi^2} \min\left\{\frac{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)}{2}, \epsilon - \omega\right\} \quad \textit{and} \\ &\sqrt{1 - 2\frac{\alpha\mu_0}{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)} + \frac{\mu_0^2\xi^2}{\operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{S}^\perp)}} =: \underline{\beta} \leq \beta < 1. \end{aligned} \tag{4.52}$$

Then it holds that

$$\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t + \omega, \ \forall t \geq 0. \tag{4.53}$$

Proof. The validity of the definition of $\underline{\beta}$ is given after Theorem 9. Let us prove (4.53) as well as $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \epsilon$ for all t by induction. Obviously, it is true when $t = 0$. Now suppose that it holds for the t-th iteration, i.e., that

$$\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t + \omega \quad \text{and} \quad \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \epsilon. \tag{4.54}$$

Consider the $(t + 1)$-th iteration.

Case (I): $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \geq \omega$. Utilizing the Riemannian regularity condition (4.36), from the last line in (4.45) we obtain

$$\begin{aligned} \operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp) &\leq \operatorname{dist}^2(\boldsymbol{B}_t, \boldsymbol{S}^\perp) - 2\alpha\mu_t \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) + \mu_t^2\xi^2 \\ &= (\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) - \alpha\mu_t)^2 + \mu_t^2(\xi^2 - \alpha^2). \end{aligned} \tag{4.55}$$

From $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t + \omega$ and

$$\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t \geq 2\frac{\mu_0 \xi^2}{\alpha}\beta^t \geq 2\alpha\mu_0\beta^t = 2\alpha\mu_t,$$

where the first inequality follows from (4.52) and the second inequality follows from (4.5), we know that (4.55) achieves its maximum at $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) = \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t + \omega$. Plugging this back into (4.55) gives

$$\begin{aligned}
&\operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp)\\
&\leq \operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^{2t} + 2\omega \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t\\
&\qquad + \omega^2 - 2\alpha\mu_t\omega - 2\alpha\mu_t \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t + \mu_t^2\xi^2\\
&= \operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^{2t} - 2\alpha\mu_0 \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^{2t} + \mu_0^2\beta^{2t}\xi^2\\
&\qquad + \omega^2 + 2\omega \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t - 2\alpha\mu_t\omega\\
&= \operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^{2t}\left(1 - 2\frac{\alpha\mu_0}{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)} + \frac{\mu_0^2\xi^2}{\operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{S}^\perp)}\right)\\
&\qquad + \omega^2 + 2\omega \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t - 2\alpha\mu_t\omega\\
&\leq \operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^{2(t+1)} + \omega^2 + 2\omega \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t - 2\alpha\mu_t\omega\\
&\leq \operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^{2(t+1)} + \omega^2 + 2\omega \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^{t+1}\\
&= \left(\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^{t+1} + \omega\right)^2
\end{aligned}$$

where the second inequality follows from the definition of $\underline{\beta}$ and $\underline{\beta} \leq \beta$ in (4.52), and the last inequality follows from

$$\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t - \alpha\mu_t \leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^{t+1}. \tag{4.56}$$

To see why (4.56) holds, first note that after writing $\mu_t = \mu_0 \beta^t$, (4.56) is equivalent to $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) - \alpha\mu_0 < \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta$. In fact, from (4.5) we have $\frac{\mu_0^2 \xi^2}{\operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{S}^\perp)} \geq \frac{\mu_0^2 \alpha^2}{\operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{S}^\perp)}$, and thus

$$\begin{aligned}
\underline{\beta} &= \sqrt{1 - 2\frac{\alpha\mu_0}{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)} + \frac{\mu_0^2\xi^2}{\operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{S}^\perp)}} \\
&\geq \sqrt{1 - 2\frac{\alpha\mu_0}{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)} + \frac{\mu_0^2\alpha^2}{\operatorname{dist}^2(\boldsymbol{B}_0, \boldsymbol{S}^\perp)}} \\
&= 1 - \frac{\alpha\mu_0}{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)}.
\end{aligned}$$

From $\beta \geq \underline{\beta}$, we have

$$\beta \geq 1 - \frac{\alpha\mu_0}{\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)},$$

which implies $\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) - \alpha\mu_0 < \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta$. Hence we conclude that

$$\operatorname{dist}(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp) \leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^{t+1} + \omega.$$

Similarly, from $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \epsilon$ and $\epsilon \geq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)\beta^t \geq 2\alpha\mu_t$, plugging $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) = \epsilon$ back into (4.55) gives

$$\operatorname{dist}^2(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp) \leq \epsilon^2 - 2\alpha\mu_t\epsilon + \mu_t^2\xi^2 \leq \epsilon^2$$

where the last inequality comes from (4.52) and

$$\frac{\mu_t \xi^2}{\alpha} = \frac{\mu_0 \beta^t \xi^2}{\alpha} \leq \frac{\mu_0 \xi^2}{\alpha} \leq \epsilon - \omega \leq 2\epsilon \quad \Rightarrow \quad -2\alpha \mu_t \epsilon + \mu_t^2 \xi^2 \leq 0.$$

Hence we also have $\operatorname{dist}(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp) \leq \epsilon$ in this case.

Case (II): $\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) < \omega$. The assumptions for RRC in Lemma 17 do not hold, but similar to (4.50), we can bound the last line in (4.45) such that

$$\begin{aligned}
\operatorname{dist}(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp) &\leq \mu_t \xi + \omega = \mu_0 \beta^t \xi + \omega \\
&\leq \frac{\alpha \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)}{2\xi^2} \beta^t \xi + \omega \\
&\leq \frac{1}{2} \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) \beta^t + \omega \\
&\leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) \beta^{t+1} + \omega
\end{aligned}$$

where the second inequality utilizes the upper bound of μ_0 in (4.52), the third inequality follows from (4.5), and the last inequality follows from $\beta \geq \underline{\beta} \geq \sqrt{1 - \frac{3\alpha^2}{4\xi^2}} \geq \frac{1}{2}$. Note that from (4.52) we have

$$\mu_0 \beta^t \xi + \omega \leq \mu_0 \xi + \omega \leq \epsilon - \omega + \omega = \epsilon,$$

and thus $\operatorname{dist}(\boldsymbol{B}_{t+1}, \boldsymbol{S}^\perp) \leq \epsilon$ also holds in this case.

Therefore, from the above discussion, the following holds for all $t \geq 0$:

$$\operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) \beta^t + \omega \;\text{ and }\; \operatorname{dist}(\boldsymbol{B}_t, \boldsymbol{S}^\perp) \leq \epsilon,$$

which completes the proof. □

Discussion of Theorem 11. With a strategy of geometrically diminishing step size in Algorithm 1, Theorem 11 implies that PRSGM applied to the noisy DPCP problem (2.9) with proper initialization converges to a neighborhood of $\boldsymbol{S}^\perp$ at a linear rate, whose radius ω is proportional to the effective noise level. This is in sharp contrast with the convergence analysis with noiseless data in Theorem 10 for which the PRSGM converges linearly to $\boldsymbol{S}^\perp$. Moreover, we note that the requirement for the initial step size μ_0 is more restricted by an amount of $\frac{\alpha}{\xi^2}\omega$ due to the existence of noise. Finally, if no noise is present, we have $\omega = 0$, which implies a linear convergence to $\boldsymbol{S}^\perp$, which is consistent with Theorem 10.

We now provide a result that guarantees that the spectral initialization provides a good enough starting point for solving the noisy problem (2.9).

Proposition 6. *The spectral initialization $\boldsymbol{B}_0$, which is obtained by taking the bottom c eigenvectors of $\widetilde{\boldsymbol{\mathcal{X}}}\widetilde{\boldsymbol{\mathcal{X}}}^\top$, satisfies*

$$\operatorname{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) \leq \sqrt{\frac{\sum_{j=1}^{c} \sigma_j^2(\boldsymbol{\mathcal{O}}) - \sum_{j=D-c+1}^{D} \sigma_j^2(\boldsymbol{\mathcal{O}}) + 2\sum_{j=1}^{c} \sigma_j^2(\widehat{\boldsymbol{\mathcal{E}}})}{\sigma_d^2(\widehat{\boldsymbol{\mathcal{X}}})}} \tag{4.57}$$

where $\sigma_\ell(\cdot)$ denotes the ℓ-th largest singular value.

Proof. Similar to the proof of Proposition 4, we have

$$\begin{aligned}
\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}_0\|_F^2 &= \min_{\boldsymbol{B} \in \mathbb{O}(D,c)} \|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}\|_F^2 \\
&\leq \min_{\boldsymbol{B} \in \mathbb{O}(D,c), \mathrm{Span}(\boldsymbol{B}) = \mathcal{S}^\perp} \|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}\|_F^2 \\
&= \min_{\boldsymbol{B} \in \mathbb{O}(D,c), \mathrm{Span}(\boldsymbol{B}) = \mathcal{S}^\perp} \left\{ \|(\widehat{\boldsymbol{\mathcal{X}}} + \widehat{\boldsymbol{\mathcal{E}}})^\top \boldsymbol{B}\|_F^2 + \|\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}\|_F^2 \right\} \qquad (4.58) \\
&= \min_{\boldsymbol{B} \in \mathbb{O}(D,c), \mathrm{Span}(\boldsymbol{B}) = \mathcal{S}^\perp} \left\{ \|\widehat{\boldsymbol{\mathcal{E}}}^\top \boldsymbol{B}\|_F^2 + \|\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}\|_F^2 \right\} \\
&\leq \sum_{j=1}^{c} \sigma_j^2(\boldsymbol{\mathcal{O}}) + \sum_{j=1}^{c} \sigma_j^2(\widehat{\boldsymbol{\mathcal{E}}})
\end{aligned}$$

where we used the fact that $\mathrm{Span}(\widehat{\boldsymbol{\mathcal{X}}}) \subseteq \mathcal{S}$, $\mathrm{Span}(\widehat{\boldsymbol{\mathcal{E}}}) \subseteq \mathcal{S}^\perp$, and (4.31).

On the other hand, since $\mathrm{Span}(\widehat{\boldsymbol{\mathcal{X}}}) \subseteq \mathcal{S}$, let $\boldsymbol{S}$ be an orthonormal basis for $\mathcal{S}$ and let $\boldsymbol{\Phi}$ be the coefficients when $\widehat{\boldsymbol{\mathcal{X}}}$ is expressed with $\boldsymbol{S}$, i.e., $\widehat{\boldsymbol{\mathcal{X}}} = \boldsymbol{S}\boldsymbol{\Phi}$. Then, we have

$$\begin{aligned}
\left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}_0\right\|_F^2 &= \left\|(\widehat{\boldsymbol{\mathcal{X}}} + \widehat{\boldsymbol{\mathcal{E}}})^\top \boldsymbol{B}_0\right\|_F^2 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}_0\right\|_F^2 \\
&\geq \left\|\widehat{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{B}_0\right\|_F^2 - \left\|\widehat{\boldsymbol{\mathcal{E}}}^\top \boldsymbol{B}_0\right\|_F^2 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{B}_0\right\|_F^2 \qquad (4.59) \\
&\geq \sigma_d^2(\widehat{\boldsymbol{\mathcal{X}}}) \left\|\boldsymbol{B}_0 - \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}_0)\right\|_F^2 - \sum_{j=1}^{c} \sigma_j^2(\widehat{\boldsymbol{\mathcal{E}}}) + \sum_{j=D-c+1}^{D} \sigma_j^2(\boldsymbol{\mathcal{O}})
\end{aligned}$$

where the last inequality follows from (4.33). Combining (4.58), (4.59), and the fact that $\mathrm{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp) = \|\boldsymbol{B}_0 - \mathcal{P}_{\boldsymbol{S}^\perp}(\boldsymbol{B}_0)\|_F$, we obtain

$$\mathrm{dist}(\boldsymbol{B}_0, \boldsymbol{S}^\perp)^2 \leq \frac{\sum_{j=1}^{c} \sigma_j^2(\boldsymbol{\mathcal{O}}) - \sum_{j=D-c+1}^{D} \sigma_j^2(\boldsymbol{\mathcal{O}}) + 2\sum_{j=1}^{c} \sigma_j^2(\widehat{\boldsymbol{\mathcal{E}}})}{\sigma_d^2(\widehat{\boldsymbol{\mathcal{X}}})},$$

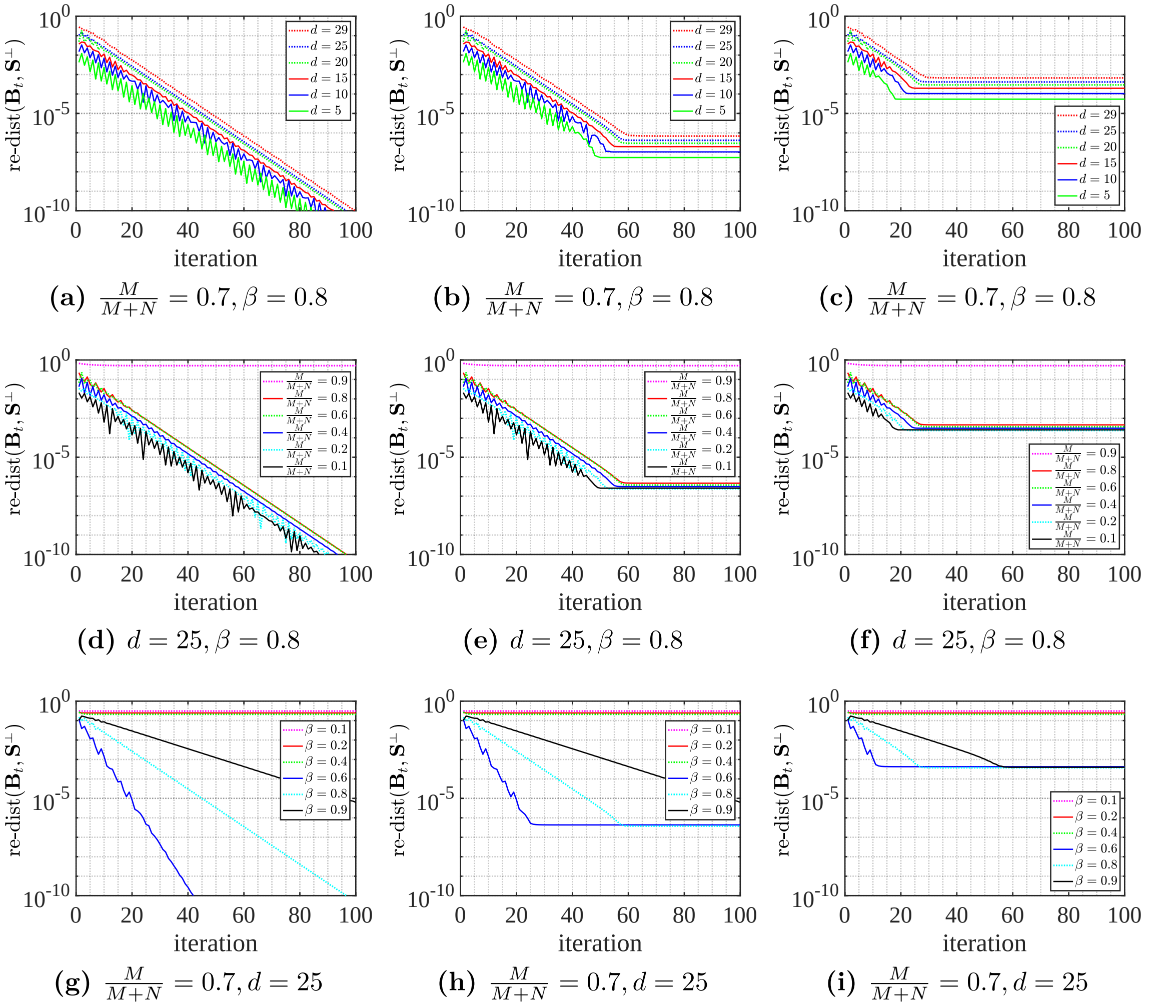
which completes the proof. □

4.3 Experiments

In this section we evaluate PRSGM (Algorithm 1) applied to solve the DPCP problem (2.9) experimentally. In Section 4.3.1 we investigate the convergence properties of PRSGM and its performance of robustly learning a subspace of high relative dimension using synthetic data. We further demonstrate its superiority by experimenting on roadplane detection using real 3D data in Section 4.3.2.

4.3.1 Synthetic data

Convergence of PRSGM. We first conduct experiments under different settings to verify the convergence properties of PRSGM (Algorithm 1) with geometrically diminishing step sizes for solving problem (2.9). The data are generated according to the random spherical model in Definition 1, where we fix $D = 30$ and $N = 500$. We use the spectral initialization as stated before, and compute the initial step size μ_0 by one iteration of a backtracking line search. Figure 4.2 demonstrates the convergence of PRSGM with different subspace dimension d (or codimension $c = D - d$), outlier ratio $\frac{M}{M+N}$, and the geometric decreasing factor β. Each of the three columns, from left to right, corresponds to a noise level with $\sigma = 0, 10^{-6}$ and 10^{-3}, respectively. In particular, Figures 4.2a, 4.2b and 4.2c show the convergence of PRSGM with $\frac{M}{M+N} = 0.7, \beta = 0.8$ under different subspace dimension d and noise level σ. We observe that the PRSGM converges linearly to $\boldsymbol{S}^{\perp}$ with noiseless data, and converges to a neighborhood of $\boldsymbol{S}^{\perp}$ when the noise level is moderate, regardless of the subspace dimension d; hence numerically justifying Theorem 11. In Figures 4.2d, 4.2e and 4.2f,



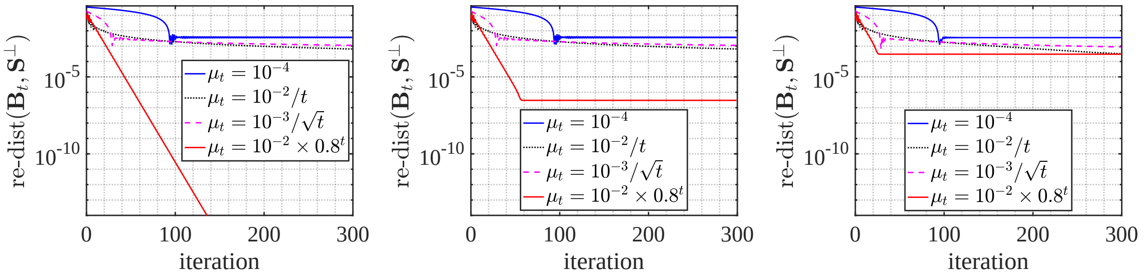
(a) $\frac{M}{M+N} = 0.7, \beta = 0.8$ **(b)** $\frac{M}{M+N} = 0.7, \beta = 0.8$ **(c)** $\frac{M}{M+N} = 0.7, \beta = 0.8$

(d) $d = 25, \beta = 0.8$ **(e)** $d = 25, \beta = 0.8$ **(f)** $d = 25, \beta = 0.8$

(g) $\frac{M}{M+N} = 0.7, d = 25$ **(h)** $\frac{M}{M+N} = 0.7, d = 25$ **(i)** $\frac{M}{M+N} = 0.7, d = 25$

Figure 4.2. Convergence of PRSGM (Algorithm 1) for the noisy DPCP problem (2.9). Each of the three columns, from left to right, corresponds to a noise level with $\sigma = 0, 10^{-6}$ and 10^{-3}, respectively. For all the cases, we fix $D = 30, N = 500$. Moreover, we choose $\boldsymbol{B}_0$ as the bottom c eigenvectors of $\widetilde{\boldsymbol{\mathcal{X}}}\widetilde{\boldsymbol{\mathcal{X}}}^\top$ and compute the initial step size μ_0 by a backtracking line search method. The relative distance re-dist$(\boldsymbol{B}_t, \boldsymbol{S}^\perp)$ is defined by dist$(\boldsymbol{B}_t, \boldsymbol{S}^\perp)/\sqrt{c}$.

we set $d = 25, \beta = 0.8$ while varying the outlier ratio $\frac{M}{M+N}$ and noise level σ. We also observe linear convergence to a neighborhood of $\boldsymbol{S}^\perp$, except for the case $\frac{M}{M+N} = 0.9$ in which case we have many more outliers than inliers. Finally, in Figures 4.2g, 4.2h and 4.2i, we set $\frac{M}{M+N} = 0.7, d = 25$ while vary the factor β that controls the geometrically diminishing step size and noise level σ. In particular, it verifies the role of β as indicated by Theorem 11, namely that β controls the convergence speed.



(a) $\sigma = 0$ **(b)** $\sigma = 10^{-6}$ **(c)** $\sigma = 10^{-3}$

Figure 4.3. Performance of PRSGM (Algorithm 1) for the noisy DPCP problem (2.9) with different step size choices μ_t. For all the cases, we fix $D = 30, d = 20, N = 500$ and $\frac{M}{M+N} = 0.7$. Moreover, we choose $\boldsymbol{B}_0$ as the bottom c eigenvectors of $\widetilde{\boldsymbol{\mathcal{X}}}\widetilde{\boldsymbol{\mathcal{X}}}^\top$. The relative distance re-dist$(\boldsymbol{B}_t, \boldsymbol{S}^\perp)$ is defined by dist$(\boldsymbol{B}_t, \boldsymbol{S}^\perp)/\sqrt{c}$.

When β is too small, e.g., $\beta \in \{0.1, 0.2, 0.4\}$, convergence may not occur, which agrees with (4.52) and (4.53). However, when $\beta \in \{0.6, 0.8, 0.9\}$ the algorithm converges at an R-linear rate, with larger values of β resulting in slower convergence speeds.

We further investigate the performance of PRSGM with different choices of step size μ_t, as illustrated in Figure 4.3. Similar to the patterns in Figure 4.2, we observe linear convergence for the geometrically diminishing step size, which converges much faster than when a constant step size or classical diminishing step size ($O(1/k)$ and $O(1/\sqrt{k})$) is used, under both noiseless and noisy settings.

Robust subspace learning with DPCP solved by PRSGM. After numerically justifying the convergence properties of PRSGM, we now turn our focus onto applying PRSGM to the DPCP problem (2.9) for robustly learning a subspace $\mathcal{S}$ of high relative dimension. As a comparison, we also try the approach of solving (2.2) recursively (see Algorithm 2). Note that in Algorithm 2, the subproblem is slightly different from the original DPCP problem (2.2) in that it has one more constraint on $\boldsymbol{b}$, i.e., $\boldsymbol{b} \perp \operatorname{Span}(\mathcal{B})$. However, the additional constraint can be removed by

Algorithm 2 The Recursive DPCP Approach for Learning a Subspace

```
Input: data $\widetilde{\boldsymbol{\mathcal{X}}}$, codimension $c$;
Set $\mathcal{B} \leftarrow \emptyset$;
for $i = 1, 2, \cdots, c$ do
    Compute $\boldsymbol{b}^{(i)} \leftarrow \arg\min_{\boldsymbol{b}\in\mathbb{S}^{D-1}, \boldsymbol{b}\perp \mathrm{Span}(\mathcal{B})} \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_1$;
    Update $\mathcal{B} \leftarrow \mathcal{B} \cup \left\{\boldsymbol{b}^{(i)}\right\}$;
end for
```

transformation of the optimization variable. Consider the following subproblem:

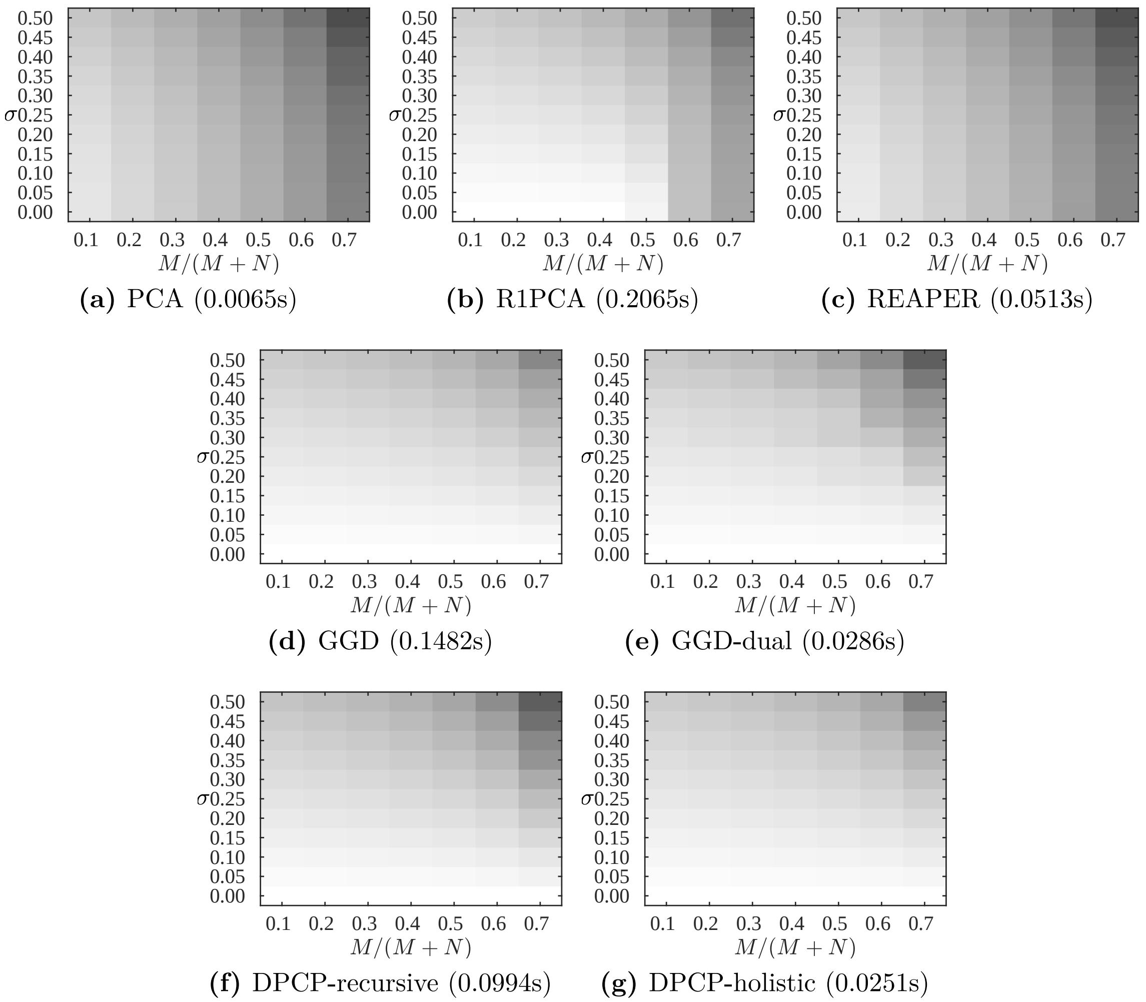
$$\min_{\boldsymbol{b}\in\mathbb{S}^{D-1}, \boldsymbol{b}\perp \mathrm{Span}(\mathcal{B})} \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_1 \tag{4.60}$$

where $\mathcal{B} = \left\{\boldsymbol{b}^{(1)}, \cdots, \boldsymbol{b}^{(p)}\right\}$ is an orthonormal set with $1 \leq p < c$. Let $\boldsymbol{A}^\perp \in \mathbb{R}^{D\times(D-p)}$ be an orthonormal matrix that is orthogonal to Span($\mathcal{B}$), thus allowing the constraint $\boldsymbol{b} \perp \mathrm{Span}(\mathcal{B})$ can be parameterized as $\boldsymbol{b} = \boldsymbol{A}^\perp \boldsymbol{\tau}$, making subproblem (4.60) is equivalent to

$$\min_{\boldsymbol{\tau}\in\mathbb{S}^{D-p-1}} \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{A}^\perp \boldsymbol{\tau}\right\|_1.$$

This is an optimization problem over the sphere that can be solved by PRSGM.

Besides the holistic and recursive approaches of DPCP, we also consider other closely related subspace recovery methods that include PCA, R1PCA [28], REAPER [61], and GGD [76]. Note that R1PCA, REAPER and GGD are primarily designed for learning a low-dimensional subspace. Observing that the objective function of GGD is similar to (2.9) except that it learns a basis for $\mathcal{S}$ instead of $\mathcal{S}^\perp$, we also apply GGD to learn a basis of $\mathcal{S}^\perp$, and call it GGD-dual. For the DPCP approaches implemented with PRSGM, we use the spectral initialization, compute the initial step size μ_0 by one iteration of a backtracking line search, and set the diminishing factor $\beta = 0.6$.



(a) PCA (0.0065s) **(b)** R1PCA (0.2065s) **(c)** REAPER (0.0513s)

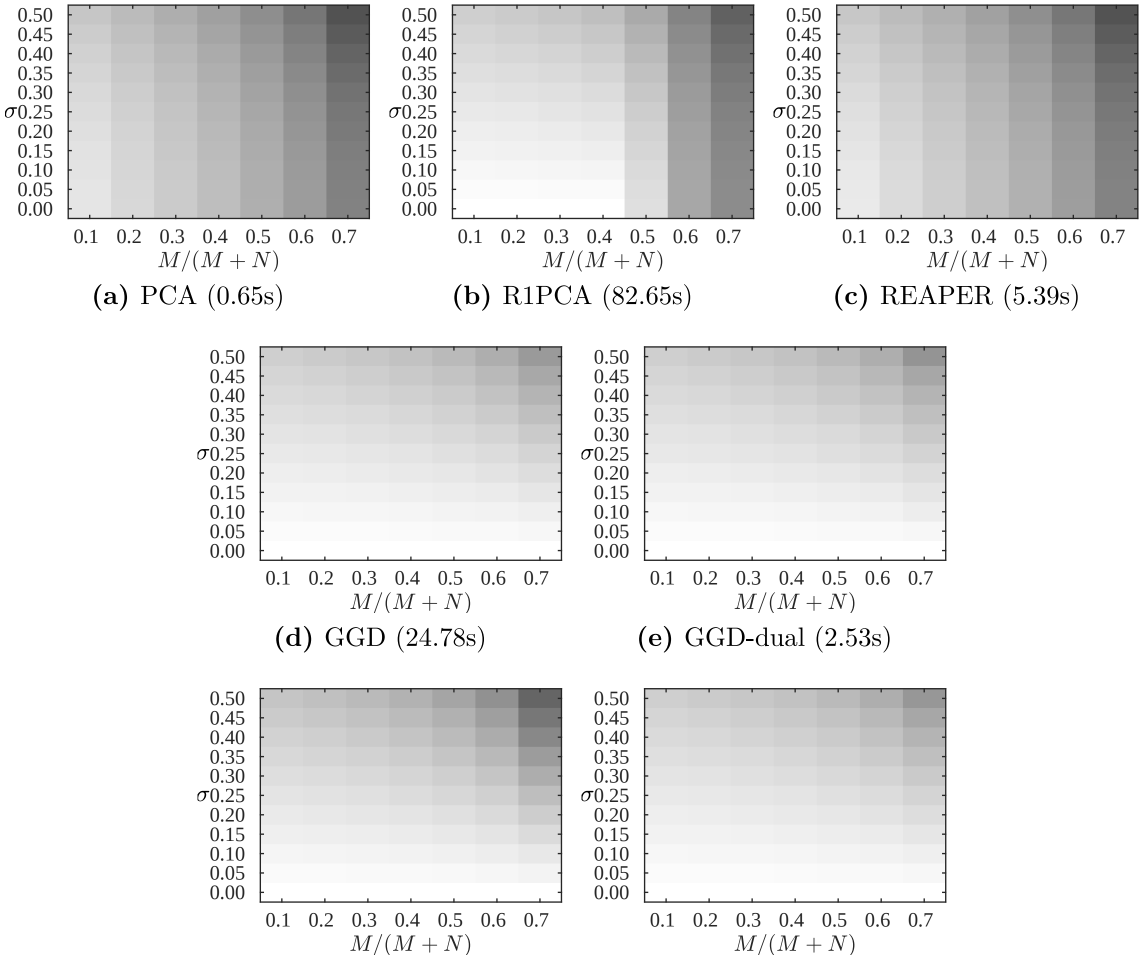
(d) GGD (0.1482s) **(e)** GGD-dual (0.0286s)

(f) DPCP-recursive (0.0994s) **(g)** DPCP-holistic (0.0251s)

Figure 4.4. Phase transition of the distance between the ground-truth basis for the (dual) subspace and the computed basis by different methods when varying the outlier ratio $M/(M+N)$ and σ. The lighter the color, the smaller the distance. The mean running time for each method is also recorded. Here we fix $D = 100, c = \lceil 0.05D \rceil = 5, N = 10D$, and the results are averaged over 100 experiments.

For all the methods, the maximal number of iterations is set to 200, and the relative convergence accuracy, wherever applicable, is set to 10^{-6}. We conduct the experiments with $D \in \{100, 1000\}$, $c = \lceil 0.05D \rceil$ and $N = 10D$ and plot the phase transition of the distance between the ground-truth basis for the (dual) subspace and the basis computed by different methods when varying the outlier ratio $\frac{M}{M+N}$ and noise level σ.

As demonstrated in Figures 4.4 and 4.5, PCA and REAPER are the least com-



(a) PCA (0.65s)
(b) R1PCA (82.65s)
(c) REAPER (5.39s)
(d) GGD (24.78s)
(e) GGD-dual (2.53s)
(f) DPCP-recursive (79.74s)
(g) DPCP-holistic (1.78s)

Figure 4.5. Phase transition of the distance between the ground-truth basis for the (dual) subspace and the computed basis by different methods when varying the outlier ratio $M/(M+N)$ and σ. The lighter the color, the smaller the distance. The mean running time for each method is also recorded. Here we fix $D = 1000, c = \lceil 0.05D \rceil = 50, N = 10D$, and the results are averaged over 100 experiments.

petitive methods in the test. PCA is not robust to outliers although it is the fastest for its simplicity. We conjecture that REAPER does not perform well as a robust subspace recovery method because it needs more inlier points for the underlying convex relaxation to be effective (in contrast to the non-convex approaches used by GGD and DPCP). R1PCA performs well with moderate outliers but is still unable to handle a high outlier ratio. Meanwhile, it is very time-consuming for estimating a subspace of high relative dimension. Next, GGD, GGD-dual and DPCP-holistic perform similarly well in terms of accurately estimating a ground-truth basis even with a high outlier ratio. However, GGD takes significantly longer time since it optimizes over $\mathbb{G}(D, d)$, which is inefficient in the high relative dimension regime. We see that applying GGD to learn the dual subspace in $\mathbb{G}(D, c)$, i.e., GGD-dual, is much faster, although not as fast as our holistic DPCP approach that solves (2.9) with PRSGM. Finally, we note that the recursive DPCP approach (Algorithm 2) is slow due to its recursive nature; moreover, as the outlier ratio and noise level increase, its estimation of the underlying subspace becomes less accurate since the error tends to accumulate during the recursive procedure. We conclude that the proposed holistic DPCP approach performs favorably against the competitors in the high relative dimension regime.

4.3.2 Roadplane detection using real 3D data

In this section, we use the experimental setup of [153] to further compare DPCP and alternative methods in the task of 3D roadplane detection. As introduced in Section 1.1.2.1, given a 3D point cloud of a road scene our goal is to learn an affine plane $\mathcal{A} = \mathcal{H} + \boldsymbol{t} \subset \mathbb{R}^3$ as a model for the road, where $\mathcal{H}$ is a plane through the origin with normal vector $\boldsymbol{b}$ and $\boldsymbol{t}$ is its translation with respect to the origin. We convert it to a linear subspace learning problem by working in homogeneous coordinates, i.e., by adding 1 at the fourth coordinate and embedding $\mathcal{A}$ into the linear hyperplane $\bar{\mathcal{H}} \subset \mathbb{R}^4$ with normal vector $\bar{\boldsymbol{b}} = [\boldsymbol{b}^\top \ \ -\boldsymbol{t}^\top \boldsymbol{b}]^\top$.

We use the 3D point clouds from the KITTI dataset [40]. In addition to the 7 frames annotated in [153], we further annotate 131 frames. Each point cloud contains around 10^5 points with approximately 50% outliers. The data are homogenized and normalized to unit ℓ_2-norm. We compare DPCP-PRSGM (Algorithm 1) to PCA, RANSAC [39], R1PCA [28], REAPER [61] and GGD [76]. Since the task involves optimization over the sphere, we also compare with the previously developed DPCP methods for learning a hyperplane, namely DPCP-PSGM [152, 153], DPCP-IRLS and DPCP-d [112] (see problem (2.8)). Additionally, for DPCP-PRSGM and DPCP-PSGM, we test with both geometrically decaying step size and a modified backtracking line search as described in [153]; the latter is known to perform well in practice but lacks a convergence theory. As a result, we denote these variants as DPCP-PRSGM-decay, DPCP-PRSGM-ls, DPCP-PSGM-decay, and DPCP-PSGM-ls.

Since DPCP-PRSGM-decay, DPCP-PSGM-decay and DPCP-d are among the fastest methods with comparable running times, we let them run to convergence and then set the running time of the slowest as the time budget for the remaining methods. For RANSAC, we also include a version with 10× and 100× that time budget. For all the DPCP approaches, we use the spectral initialization and compute the initial step size by one iteration of a backtracking line search. For DPCP-PRSGM-decay and DPCP-PSGM-decay, we set the diminishing factor to 0.6. We tune the parameters of the other algorithms on a randomly selected training set of 13 frames and use the rest of the frames for evaluation. Each method is tuned to achieve an optimal error and then re-tuned to be as fast as possible without exceeding 5% of that error. The λ of DPCP-d is set to $\frac{2.76}{\sqrt{N+M}}$, the minimum step size allowed for the DPCP approaches is

Table 4.1. 3D road plane estimation using 125 annotated frames of the KITTI dataset.

Methods/metric	ROC	$\hat{\bar{\theta}}$	$\hat{\theta}$	$\hat{\boldsymbol{t}}$	# iterations	time (in msec)
PCA	0.76	4.40	1.73	14%	N/A	1
RANSAC×1	0.78	3.74	4.18	12%	3.8	31
RANSAC×10	0.91	1.58	2.85	5%	18.7	149
RANSAC×100	**0.93**	**1.47**	2.77	**4%**	64.1	515
R1PCA	0.89	2.24	0.93	8%	6.1	25
REAPER	0.88	2.48	1.07	8%	4.1	27
GGD	0.80	3.40	1.59	11%	3.0	26
DPCP-IRLS	0.81	3.67	1.48	12%	3.0	29
DPCP-d	0.92	1.51	0.82	5%	6.5	16
DPCP-PSGM-ls	0.92	1.59	**0.76**	5%	37.3	24
DPCP-PSGM-decay	0.85	2.90	1.15	10%	31.1	14
DPCP-PRSGM-ls	0.92	1.59	**0.76**	5%	35.8	24
DPCP-PRSGM-decay	0.85	2.96	1.17	10%	31.1	14

set to 10^{-9}, and the relative convergence accuracy, wherever applicable, is set to 10^{-6}.

Table 4.1 reports geometric, clustering and algorithmic metrics for the various methods. Once a method has computed an estimated normal vector $\hat{\bar{\boldsymbol{b}}} \in \mathbb{R}^4$, we extract from it estimates $\hat{\boldsymbol{b}}, \hat{\boldsymbol{t}}$. We report the corresponding estimation errors, i.e., the angle $\hat{\bar{\theta}}$ between $\bar{\boldsymbol{b}}^*$ and $\hat{\bar{\boldsymbol{b}}}$, the angle $\hat{\theta}$ between $\boldsymbol{b}^*$ and $\hat{\boldsymbol{b}}$, and $100 \left\|\boldsymbol{t}^* - \hat{\boldsymbol{t}}\right\|_2 / \left\|\boldsymbol{t}^*\right\|_2 \%$, where $\bar{\boldsymbol{b}}^*, \boldsymbol{b}^*, \boldsymbol{t}^*$ are the ground-truth values. By varying a threshold on the distances of all points to the estimated affine plane, the area under the ROC curve is obtained (this is also the internal thresholding parameter for RANSAC), with higher values indicating better performance. Finally, the number of averaged iterations executed by each method and its running time in msec are also reported. Notably, not only does DPCP-PRSGM-ls, DPCP-PSGM-ls and DPCP-d outperform RANSAC×1 and RANSAC×10, but its performance is comparable with that of RANSAC×100, which they still surpass in estimating the orientation of the normal vector $\boldsymbol{b}^*$: RANSAC×100

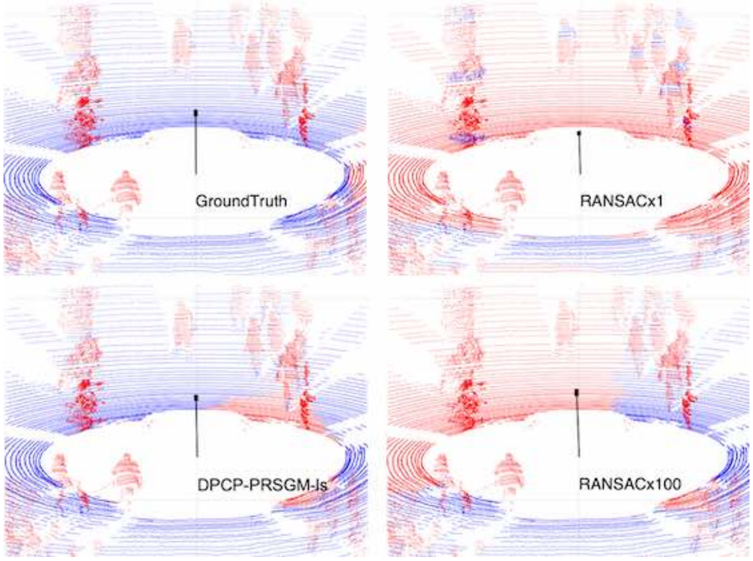


Figure 4.6. 3D point clouds and estimated translations for frame 328 of KITTY-CITY-71, with inliers in blue and outliers in red.

is off by 2.77° on average, while DPCP-PRSGM-ls (DPCP-PSGM-ls) and DPCP-d are only off by 0.76° and 0.82°, respectively; see Figure 4.6. We also note that although DPCP-PRSGM-decay and DPCP-PSGM-decay are among the fastest methods, they are not competitive to their counterparts that use the modified line search for updating the step size. On the other hand, DPCP-IRLS and REAPER make heavy use of the SVD, which makes them slow to run on $\mathcal{O}(10^5)$ points, and eventually inaccurate given the limited time budget. As illustrated in Figures 4.7 and 4.8, we visualize the results of the above methods by projecting the 3D point clouds onto the image.

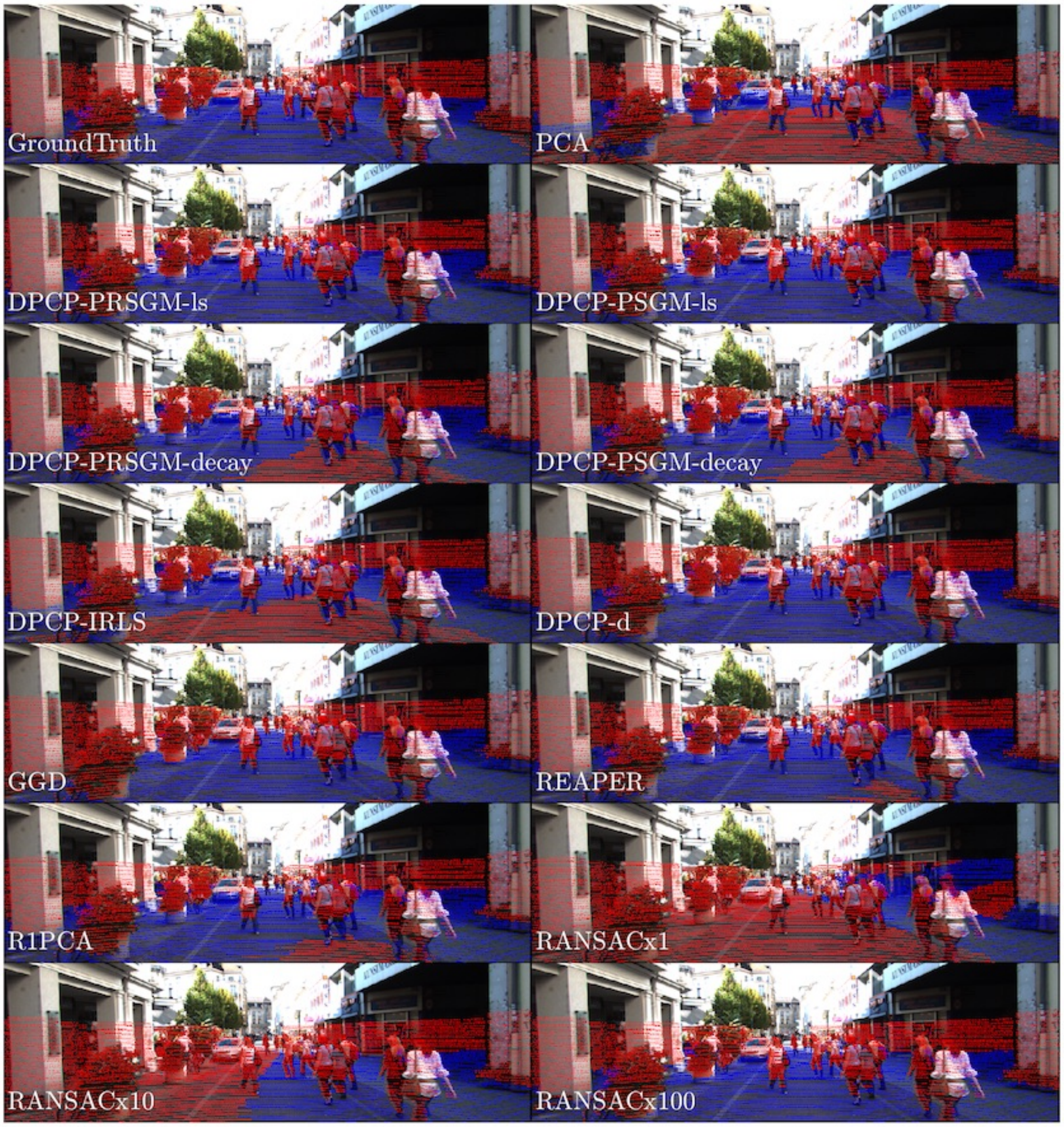


Figure 4.7. Projections of 3D point clouds for frame 328 of KITTY-CITY-71 onto the image, with inliers in blue and outliers in red.

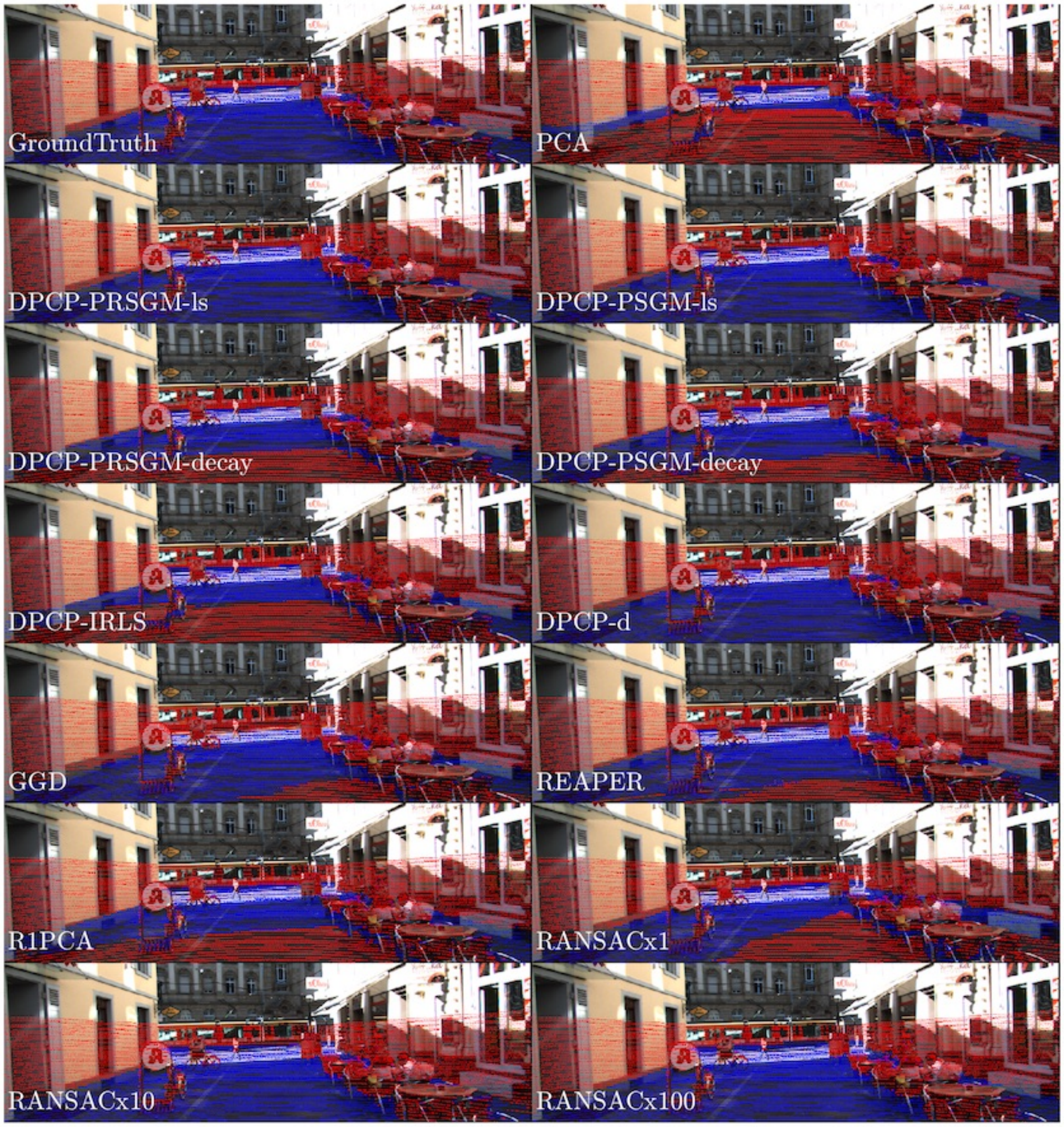


Figure 4.8. Projections of 3D point clouds for frame 881 of KITTY-CITY-71 onto the image, with inliers in blue and outliers in red.

Chapter 5

Learning a Union of Hyperplanes with DPCP

In Chapter 3 and Chapter 4, we established theory and developed algorithms for learning a single subspace of high relative dimension with DPCP. It is known [109, 113] that DPCP is able to handle data points drawn from a union of hyperplanes (UoH), where it is used to learn the normal vector to a dominant hyperplane. Nevertheless, existing analyses of DPCP in the multi-hyperplane case lack a precise characterization of the distribution of the data and are difficult to interpret. Furthermore, the provable algorithm based on solving a recursion of linear programs is inefficient. Thus, it is reasonable to ask whether we can provide a more transparent analysis by leveraging the geometric quantities and analytical techniques from Chapter 3 as well as extend the PRSGM proposed in Chapter 4 to solve the DPCP problem under a UoH model.

We structure this chapter as followings. In Section 5.1, we introduce the problem of learning a hyperplane under a UoH model, and discuss the limitations of closely related work. Next, we provide an improved analysis of the DPCP problem for a UoH in Section 5.2. The extension of the PRSGM applied to DPCP for a UoH is given in Section 5.3. We present how to do hyperplane clustering with DPCP in Section 5.4. Finally, the results of our numerical experiments are presented in Section 5.5.

5.1 Introduction

DPCP has been analyzed for learning a hyperplane from data under a UoH model [106, 109, 113] by optimizing the same problem (2.2) as for learning a single hyperplane:

$$\min_{\boldsymbol{b}\in\mathbb{R}^D} \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_1 \text{ s.t. } \|\boldsymbol{b}\|_2 = 1. \tag{5.1}$$

The distinction is that the dataset now has the form of $\widetilde{\boldsymbol{\mathcal{X}}} = [\boldsymbol{\mathcal{X}}_1, \cdots, \boldsymbol{\mathcal{X}}_K]\boldsymbol{\Gamma} \in \mathbb{R}^{D\times N}$, where $\bigcup_{k=1}^K \boldsymbol{\mathcal{X}}_k = \boldsymbol{\mathcal{X}} \in \mathbb{R}^{D\times N}$ are N inlier points that lie in a union of K hyperplanes $\mathcal{H}_1, \cdots, \mathcal{H}_K$ of $\mathbb{R}^D$ with unit normal vectors $\boldsymbol{n}_1, \cdots, \boldsymbol{n}_K$, respectively, and $\boldsymbol{\mathcal{X}}_k$ are N_k inlier points that belong to $\mathcal{H}_k$ for every $k \in [K] := \{1, \cdots, K\}$.

As discussed in Section 2.1.2, the data modeling for a UoH is fundamentally different than a single hyperplane learning case since when we treat the data points from one specific hyperplane as inliers, the points from other hyperplanes cannot be merely viewed as regular outliers as before since they exhibit additional linear structures, and hence need to be treated differently in the analysis. The problem (5.1) becomes even more challenging if the dataset $\widetilde{\boldsymbol{\mathcal{X}}}$ also contains regular outliers, i.e.,

$\widetilde{\boldsymbol{\mathcal{X}}} = [\boldsymbol{\mathcal{X}}_1, \cdots, \boldsymbol{\mathcal{X}}_K, \boldsymbol{\mathcal{O}}]\boldsymbol{\Gamma} \in \mathbb{R}^{D \times (N+M)}$, with $\boldsymbol{\mathcal{O}} \in \mathbb{R}^{D \times M}$ the M outlier points that do not exhibit any certain structures, which is excluded in the analysis of (5.1) in [106, 109, 113]. It is not known, however, whether DPCP can learn a normal to one of the hyperplanes in the presence of *both* structured and regular outliers. In fact, several related questions remain unanswered. Under what conditions is a global optimum of the DPCP problem (5.1) a normal to one of the hyperplanes? When the global optimum is a normal, which hyperplane is it a normal to? Can the convergence of some optimization algorithm to a global solution to the non-convex DPCP problem under the UoH[1] data model be guaranteed? This chapter addresses all of these challenges.

Before moving on, we discuss the limitations of the most closely related work. Note that [113] has partially addressed the previous challenges of DPCP for a UoH without outliers, while [64] analyzed ℓ_p recovery of a single subspace from a union of subspaces with problem (5.1) as a special case (i.e. $p = 1$ and subspaces are of dimension $d = D - 1$). Three key aspects of their limitations should be emphasized (see Table 5.1 for a summary). First, in the analysis of which hyperplane is recovered, [113] and [64] introduce different notions of a "dominant" or "most significant" hyperplane, which depend only on the (expected) number of points in each group. In particular, the hyperplane (say $\mathcal{H}_1$) with the most number of points is defined as the *dominant hyperplane* in [113], i.e.,

$$N_1 > \max_{k \geq 2} N_k. \tag{5.2}$$

It is proved in [113] that a global solution of (5.1) is a normal vector of $\mathcal{H}_1$ under

[1] For the rest of the chapter, when we say "a UoH model", we assume it contains regular outliers.

certain conditions that implicitly make use of the distribution of the data, but are deterministic in nature and difficult to interpret. On the other hand, [64] considers a random model where inliers are sampled from $(\cup_{k=1}^{K}\mathcal{H}_k)\cap\mathbb{S}^{D-1}$ with weights $\{\psi_k\}_{k=1}^{K}$ (ψ_k is the weight of sampling inliers in $\mathcal{H}_k$) and outliers are sampled from $\mathbb{S}^{D-1}$ with weight ψ_0, and $\sum_{k=0}^{K}\psi_k = 1$. Then $\mathcal{H}_1$ is defined as the *most significant hyperplane* if

$$\psi_1 > \sum_{k=2}^{K}\psi_k. \tag{5.3}$$

The number of sampled points, in expectation, is equivalent to $N_1 > \sum_{k\geq 2} N_k$. We argue that the global optimum depends not only on the (expected) number of data points in each group, but also on geometric quantities related to their distribution. Currently there is no notion of *geometric dominance* that captures these aspects. Second, [113] provides geometric conditions under which the global minimum is a normal to the "dominant" hyperplane, and [64] provide probabilistic conditions. However, neither have *both* types of analyses, nor do the analyses make connections to geometric dominance. Third, the provably convergent algorithm in [113], which is based on a recursion of linear programs (LPs), is not scalable, while the recommended *Iteratively Reweighted Least Squares* (IRLS) [61, 63] approach does not have a guarantee for the DPCP problem. Meanwhile, [64] does not provide concrete algorithms for solving the problem. In other words, there lacks an algorithm that is scalable and enjoys a convergence guarantee for learning a single hyperplane under a UoH model. It is desirable that the PRSGM developed in Chapter 4 can be provably extended to solve (5.1) for a specific hyperplane.

Table 5.1. The theory and algorithms for learning a hyperplane under a UoH model for the most closely related work.

	Theory			Algorithms		
	Which hyperplane does it recover?	Handle outliers	Analytical approach		Convergence guarantee	Scale well?
[64]	most significant plane (see (5.3))	✓	probabilistic	–	–	–
[113]	dominant plane (see (5.2))	✗	geometric	LPs IRLS	✓ ✗	✗ ✓
This work	geometrically dominant plane (see Definition 4)	✓	probabilistic + geometric	PRSGM	✓	✓

5.2 Analysis of DPCP for a union of hyperplanes

In this section, we introduce a new notion of geometric dominance for determining the hyperplane that is learned by the DPCP problem (5.1) under a UoH model (Section 5.2.1), present deterministic geometric analyses of its critical points (Section 5.2.2) and global solutions (Section 5.2.3), and also provide an interpretable probabilistic analysis (Section 5.2.4).

5.2.1 Geometrically dominant hyperplane

Building upon problem (5.1), we consider a dataset $\widetilde{\boldsymbol{\mathcal{X}}}$ that also contains the regular outlier term $\boldsymbol{\mathcal{O}}$. If $\boldsymbol{b}$ is a normal vector to a hyperplane, it is orthogonal to all the data points within this hyperplane. Thus, we attempt to find a normal vector to one specific hyperplane by solving

$$\min_{\boldsymbol{b}\in\mathbb{S}^{D-1}} f(\boldsymbol{b}) := \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\right\|_1 = \sum_{k=1}^{K} \left\|\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b}\right\|_1 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{b}\right\|_1 .$$

Note that for learning a *single* hyperplane, say $\mathcal{H}_1$, when the inliers are uniformly distributed in $\mathcal{H}_1 \cap \mathbb{S}^{D-1}$ and the outliers are uniformly distributed in $\mathbb{S}^{D-1}$, according to the theory established in Chapter 3, we know that the DPCP problem (5.1) can provably recover the true normal vector to $\mathcal{H}_1$ provided that the number of outliers is big-O of the square of the number of inliers. When $\boldsymbol{\mathcal{X}}$ consists of inliers from a union of K hyperplanes, as is the case considered in this chapter, the analysis of a single hyperplane *cannot* be applied here by treating the data points from one hyperplane as inliers and the rest as outliers since the data distribution in other planes is far from uniform and thus violates the prior assumptions.

Geometric quantities. Since the outlier term $\boldsymbol{\mathcal{O}}$ is the same as before, we adopt the quantities $c_{\boldsymbol{\mathcal{O}},\max}$ and $c_{\boldsymbol{\mathcal{O}},\min}$ defined in (3.5) and $\eta_{\boldsymbol{\mathcal{O}}}$ defined in (3.6) to characterize the distribution of the outliers. Next, for the inlier subset $\boldsymbol{\mathcal{X}}_k$ in hyperplane $\mathcal{H}_k$, similar to the definition of $c_{\boldsymbol{\mathcal{X}},\min}$ in (3.4), we define

$$\begin{aligned} c_{\boldsymbol{\mathcal{X}}_k,\min} &:= \frac{1}{N_k} \min_{\boldsymbol{b} \in \mathcal{H}_k \cap \mathbb{S}^{D-1}} \left\| \boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b} \right\|_1 \quad \text{and} \\ c_{\boldsymbol{\mathcal{X}}_k,\max} &:= \frac{1}{N_k} \max_{\boldsymbol{b} \in \mathcal{H}_k \cap \mathbb{S}^{D-1}} \left\| \boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b} \right\|_1 . \end{aligned} \tag{5.4}$$

A well-distributed $\boldsymbol{\mathcal{X}}_k$ leads to a large value of $c_{\boldsymbol{\mathcal{X}}_k,\min}$ and small value of $c_{\boldsymbol{\mathcal{X}}_k,\max}$ since it is difficult to find a single direction $\boldsymbol{b}$ that is orthogonal to (or in line with) many points in $\boldsymbol{\mathcal{X}}_k$. Finally, parallel to the definition of $\eta_{\boldsymbol{\mathcal{O}}}$ in (3.6), we define the following quantity that further characterizes the distribution of inliers:

$$\eta_{\boldsymbol{\mathcal{X}}_k} := \frac{1}{N_k} \max_{\boldsymbol{b} \in \mathcal{H}_k \cap \mathbb{S}^{D-1}} \left\| (\mathcal{P}_{\mathcal{H}_k} - \boldsymbol{b}\boldsymbol{b}^\top) \boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b}) \right\|_2$$

where $\mathcal{P}_{\mathcal{H}_k}$ is the orthonormal projection onto $\mathcal{H}_k$ and $\operatorname{sign}(\boldsymbol{a})$ denotes that we apply $\operatorname{sign}(\cdot)$ as defined in (3.2) element-wise to a vector $\boldsymbol{a}$. Note that how $\eta_{\boldsymbol{\mathcal{X}}_k}$ is different from $\eta_{\boldsymbol{\mathcal{O}}}$ in (3.6): the unit vector $\boldsymbol{b}$ in the definition of $\eta_{\boldsymbol{\mathcal{X}}_k}$ is also restricted to be inside $\mathcal{H}_k$ for the sake of characterizing the distribution of inliers $\boldsymbol{\mathcal{X}}_k$. We will see shortly that the optimality analysis for (5.1) based on these geometric quantities is easier to interpret and facilitates a probabilistic analysis.

Geometrically dominant hyperplane. For the objective in (5.1), the outlier term $\left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{b}\right\|_1$ should be nearly constant for well distributed outliers, so that the minimizer of (5.1) is determined by the relative importance of the inlier terms $\left\|\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b}\right\|_1$. We also expect the relative orientation of the underlying hyperplanes to play an important role in determining the solution to (5.1). For example, in the case that data are uniformly sampled and each plane has the same point weights, the solution of (5.1) has a bias towards the normals of the planes that are close to each other. Noting that the geometric relationships among $\{\mathcal{H}_k\}$ are determined by the principal angles between $\{\boldsymbol{n}_k\}$, we define $\theta_{k\ell} \in [0, \pi/2]$ to be the principal angle between $\boldsymbol{n}_k$ and $\boldsymbol{n}_\ell$. By analyzing the first-order necessary condition for problem (5.1), we define ζ_k as the following measure of the relative dominance for $\boldsymbol{\mathcal{X}}_k$ that considers the point weights, data distribution, and relative orientation of the hyperplanes:

$$\zeta_k := \frac{N_k c_{\boldsymbol{\mathcal{X}}_k,\min}}{\sqrt{\boldsymbol{1}^\top \boldsymbol{W}_{(k,k)}^{\max} \boldsymbol{1}} + \sum_{\ell \neq k} N_\ell \eta_{\boldsymbol{\mathcal{X}}_\ell} + M \overline{\eta}_{\boldsymbol{\mathcal{O}}}} \tag{5.5}$$

where

$$\boldsymbol{W}^{\max} := \left(N_k c_{\boldsymbol{\mathcal{X}}_k,\max} N_\ell c_{\boldsymbol{\mathcal{X}}_\ell,\max} \cos(\theta_{k\ell})\right)_{1\le k,\ell\le K} \in \mathbb{R}^{K\times K} \tag{5.6}$$

whose (k,ℓ)th entry represents the joint importance of $\boldsymbol{\mathcal{X}}_k$ and $\boldsymbol{\mathcal{X}}_\ell$ weighted by $\cos(\theta_{k\ell})$, $\boldsymbol{W}^{\max}_{(k,k)}$ is the principal submatrix obtained by deleting the kth row and kth column of $\boldsymbol{W}^{\max}$, $\mathbf{1}$ is the vector of all ones, and $\overline{\eta}_{\boldsymbol{\mathcal{O}}} := \eta_{\boldsymbol{\mathcal{O}}} + D/M$ is defined in (3.7). Note the following: (i) the numerator $N_k c_{\boldsymbol{\mathcal{X}}_k,\min}$ of (5.5) gives the contribution from $\boldsymbol{\mathcal{X}}_k$; (ii) the term $\mathbf{1}^\top \boldsymbol{W}^{\max}_{(k,k)}\mathbf{1}$ in the denominator counts the sum of the entries in $\boldsymbol{W}^{\max}_{(k,k)}$, capturing the total contributions from $\{\boldsymbol{\mathcal{X}}_\ell\}_{\ell\neq k}$; and (iii) the last term $\sum_{\ell\neq k} N_\ell \eta_{\boldsymbol{\mathcal{X}}_\ell} + M\overline{\eta}_{\boldsymbol{\mathcal{O}}}$ is typically small[2] compared with the former two terms. Overall, ζ_k measures the relative dominance of $\boldsymbol{\mathcal{X}}_k$ over $\{\boldsymbol{\mathcal{X}}_\ell\}_{\ell\neq k}$. We see that a larger relative dominance of $\boldsymbol{\mathcal{X}}_k$ (i.e. larger ζ_k) results from better distributed data points, larger N_k relative to M and N_ℓ for $\ell \neq k$, and better separation of the other hyperplanes (large $\theta_{ij}, i, j \neq k, i \neq j$).

Definition 4. *With ζ_k in* (5.5)*, we say that $\mathcal{H}_k$ is a* geometrically dominant hyperplane *if and only if $\zeta_k \geq \zeta_\ell, \forall \ell \in [K]$.*

The notion of geometric dominance makes the deterministic analysis more interpretable (Sections 5.2.2 and 5.2.3) and allows a probabilistic analysis (Section 5.2.4) that is easier to be satisfied with only a mild number of sampled points.

Proposition 7. *There is at most one $k \in [K]$ such that $\zeta_k > 1$, in which case it also holds that $\zeta_\ell < 1$ for all $\ell \in [K]\backslash k$.*

[2] Assuming points in $\boldsymbol{\mathcal{X}}_k$ and $\boldsymbol{\mathcal{O}}$ are uniformly sampled from $\mathbb{S}^{D-1}\cap\mathcal{H}_k$ and $\mathbb{S}^{D-1}$, respectively, according to Lemma 2, both $N_k c_{\boldsymbol{\mathcal{X}}_k,\max}$ and $N_k c_{\boldsymbol{\mathcal{X}}_k,\min}$ scale as $O(N_k)$, while $N_k\eta_{\boldsymbol{\mathcal{X}}_k}$ scales as $O(\sqrt{N_k})$ and $M\eta_{\boldsymbol{\mathcal{O}}}$ scales as $O(\sqrt{M})$.

Proof. Without loss of generality, assume $\zeta_1 > 1$ and $\zeta_2 \geq 1$. From (5.5) we have

$$1 < \zeta_1 = \frac{N_1 c_{\boldsymbol{\mathcal{X}}_1,\min}}{\sqrt{\mathbf{1}^\top \boldsymbol{W}_{(1,1)}^{\max}\mathbf{1} + \sum_{\ell\neq 1} N_\ell \eta_{\boldsymbol{\mathcal{X}}_\ell} + M\overline{\eta}_{\boldsymbol{\mathcal{O}}}}} < \frac{N_1 c_{\boldsymbol{\mathcal{X}}_1,\min}}{\sqrt{\mathbf{1}^\top \boldsymbol{W}_{(1,1)}^{\max}\mathbf{1}}} < \frac{N_1 c_{\boldsymbol{\mathcal{X}}_1,\min}}{N_2 c_{\boldsymbol{\mathcal{X}}_2,\max}}$$

where we used the fact that $\mathbf{1}^\top \boldsymbol{W}_{(1,1)}^{\max}\mathbf{1} > N_2^2 c_{\boldsymbol{\mathcal{X}}_2,\max}^2$, from which we obtain that $N_1 c_{\boldsymbol{\mathcal{X}}_1,\min} > N_2 c_{\boldsymbol{\mathcal{X}}_2,\max}$. Similarly, we have

$$1 \leq \zeta_2 = \frac{N_2 c_{\boldsymbol{\mathcal{X}}_2,\min}}{\sqrt{\mathbf{1}^\top \boldsymbol{W}_{(2,2)}^{\max}\mathbf{1} + \sum_{\ell\neq 2} N_\ell \eta_{\boldsymbol{\mathcal{X}}_\ell} + M\overline{\eta}_{\boldsymbol{\mathcal{O}}}}} < \frac{N_2 c_{\boldsymbol{\mathcal{X}}_2,\min}}{\sqrt{\mathbf{1}^\top \boldsymbol{W}_{(2,2)}^{\max}\mathbf{1}}} < \frac{N_2 c_{\boldsymbol{\mathcal{X}}_2,\min}}{N_1 c_{\boldsymbol{\mathcal{X}}_1,\max}}$$

so that $N_2 c_{\boldsymbol{\mathcal{X}}_2,\min} \geq N_1 c_{\boldsymbol{\mathcal{X}}_1,\max}$. Combining these results with $c_{\boldsymbol{\mathcal{X}}_2,\max} \geq c_{\boldsymbol{\mathcal{X}}_2,\min}$ gives

$$N_1 c_{\boldsymbol{\mathcal{X}}_1,\min} > N_2 c_{\boldsymbol{\mathcal{X}}_2,\max} \geq N_2 c_{\boldsymbol{\mathcal{X}}_2,\min} \geq N_1 c_{\boldsymbol{\mathcal{X}}_1,\max},$$

which contradicts the fact that $c_{\boldsymbol{\mathcal{X}}_1,\min} \leq c_{\boldsymbol{\mathcal{X}}_1,\max}$, hence completing the proof. □

Discussion of Proposition 7. It follows from Proposition 7 that if $\zeta_k > 1$ then $\mathcal{H}_k$ is the *unique* geometrically dominant hyperplane. For the rest of the analysis, we assume that there always exists $k \in [K]$ such that $\zeta_k > 1$; the scenario that such a geometrically dominant hyperplane does not exist is left for future work. We note that this assumption ensures a simple landscape of the non-convex DPCP problem (5.1) that allows us to show that under certain conditions the global minimizers of (5.1) are guaranteed to be normal vectors of the geometrically dominant hyperplane (Theorem 12). The assumption may be stronger than needed in theory[3]

[3] In fact, [113, Proposition 5] shows that for three equi-angular hyperplanes, global minimizers of (5.1) can be normal vectors of any of the planes when they are well-separated and the data points are well-distributed.

since it excludes the possibility that normals of the other hyperplanes are global solutions to (5.1), which are also of interest. Furthermore, we remark that other works make similar assumptions—[113] requires (5.2) and [64] requires (5.3). We will see that, when data is sampled from a specific random spherical model (Theorem 13), the geometric dominance not only implies that both (5.2) and (5.3) hold, but also that it has the advantage of explicitly characterizing the data distribution. Finally, this assumption is likely to be satisfied in the subspace estimation step of K-subspaces (KSS) [1, 10] where most of the points in the estimated cluster are expected to be sampled from one dominant hyperplane with the remaining points belonging to the other hyperplanes; this works well in practice as we will see in Section 5.5.

5.2.2 Geometry of the critical points

Without loss of generality, we assume $\zeta_1 > 1$, i.e., that $\mathcal{H}_1$ is the geometrically dominant hyperplane. In the next result, we characterize critical points of (5.1) with respect to the geometrically dominant hyperplane $\mathcal{H}_1$.

Lemma 18. *Any critical point $\boldsymbol{b}$ of problem* (5.1) *must belong to* $\{\pm\boldsymbol{n}_1\}$ *or have its principal angle θ from $\boldsymbol{n}_1$ satisfy* $\theta \geq \arcsin\left(\sqrt{1-(1/\zeta_1)^2}\right)$.

Proof. Our goal is to characterize the geometry of a critical point $\boldsymbol{b}$ of (5.1) with respect to the geometrically dominant hyperplane $\mathcal{H}_1$. We observe that a Riemannian subgradient for the inlier term $\left\|\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b}\right\|_1$ is of the form

$$(\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b}) = (\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{N_k} \operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top \boldsymbol{b})\boldsymbol{x}_j^{(k)},$$

where $\boldsymbol{x}_j^{(k)}$ as the jth point in $\boldsymbol{\mathcal{X}}_k$ and $\operatorname{sign}(\cdot)$ is defined in (3.2). Let θ_k be the principal angle between $\boldsymbol{b}$ and $\boldsymbol{n}_k$. We first show that for any $k \in [K]$ and $\boldsymbol{b}$ that is not orthogonal to $\mathcal{H}_k$ i.e., $\boldsymbol{b} \notin \{\pm \boldsymbol{n}_k\}$, it holds that

$$\begin{aligned}\cos(\theta_k) N_k c_{\boldsymbol{\mathcal{X}}_k,\min} \leq & \left\| (\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top) \sum_{j=1}^{N_k} \operatorname{sign}(\boldsymbol{x}_j^{(k)\top} \boldsymbol{b}) \boldsymbol{x}_j^{(k)} \right\|_2 \\ & \leq \cos(\theta_k) N_k c_{\boldsymbol{\mathcal{X}}_k,\max} + N_k \eta_{\boldsymbol{\mathcal{X}}_k}.\end{aligned} \tag{5.7}$$

By decomposing $\boldsymbol{b} = \sin(\theta_k)\bar{\boldsymbol{s}}_k + \cos(\theta_k)\bar{\boldsymbol{n}}_k$, where $\bar{\boldsymbol{s}}_k \in \mathcal{H}_k, \bar{\boldsymbol{n}}_k \in \mathcal{H}_k^\perp$, and $\|\bar{\boldsymbol{s}}_k\|_2 = \|\bar{\boldsymbol{n}}_k\|_2 = 1$, it follows that

$$\begin{aligned}& \left\| (\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top) \sum_{j=1}^{N_k} \operatorname{sign}(\boldsymbol{x}_j^{(k)\top} \boldsymbol{b}) \boldsymbol{x}_j^{(k)} \right\|_2^2 \\ = & \left\| \sum_{j=1}^{N_k} \operatorname{sign}(\boldsymbol{x}_j^{(k)\top} \boldsymbol{b}) \left(\boldsymbol{x}_j^{(k)} - \sin(\theta_k)(\boldsymbol{x}_j^{(k)\top} \bar{\boldsymbol{s}}_k)\boldsymbol{b} \right) \right\|_2^2 \\ = & \left\| \sum_{j=1}^{N_k} \operatorname{sign}(\boldsymbol{x}_j^{(k)\top} \boldsymbol{b}) \boldsymbol{x}_j^{(k)} - \sin(\theta_k) \sum_{j=1}^{N_k} \operatorname{sign}(\boldsymbol{x}_j^{(k)\top} \boldsymbol{b})(\boldsymbol{x}_j^{(k)\top} \bar{\boldsymbol{s}}_k)\boldsymbol{b} \right\|_2^2 \\ = & \left\| \boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b}) - \sin(\theta_k) \left\| \boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k \right\|_1 \boldsymbol{b} \right\|_2^2 \\ = & \left\| \boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b}) \right\|_2^2 + \sin^2(\theta_k) \left\| \boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k \right\|_1^2 - 2\sin(\theta_k) \left\| \boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k \right\|_1 \boldsymbol{b}^\top \boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b}) \\ = & \left\| \boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b}) \right\|_2^2 + \sin^2(\theta_k) \left\| \boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k \right\|_1^2 - 2\sin^2(\theta_k) \left\| \boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k \right\|_1^2 \\ = & \left\| \boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b}) \right\|_2^2 - \sin^2(\theta_k) \left\| \boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k \right\|_1^2\end{aligned} \tag{5.8}$$

where the first equality follows from $\boldsymbol{b}^\top \boldsymbol{x}_j^{(k)} = \sin(\theta_k)\bar{\boldsymbol{s}}_k^\top \boldsymbol{x}_j^{(k)}$ and the third equality follows from $\operatorname{sign}(\boldsymbol{x}_j^{(k)\top} \boldsymbol{b}) = \operatorname{sign}(\sin(\theta_k)\bar{\boldsymbol{s}}_k^\top \boldsymbol{x}_j^{(k)}) = \operatorname{sign}(\bar{\boldsymbol{s}}_k^\top \boldsymbol{x}_j^{(k)})$. To bound the last line in (5.8), we note that

$$\begin{aligned}\left\|\boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b})\right\|_2^2 &= \left\|\boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b})\right\|_2^2 \|\bar{\boldsymbol{s}}_k\|_2^2 \\ &\geq \left(\sum_{j=1}^{N_k} \operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top \boldsymbol{b}) {\boldsymbol{x}_j^{(k)}}^\top \bar{\boldsymbol{s}}_k\right)^2 = \left\|\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k\right\|_1^2\end{aligned} \tag{5.9}$$

where we used the Cauchy-Schwartz inequality and $\operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top \boldsymbol{b}) = \operatorname{sign}(\sin(\theta_k)\bar{\boldsymbol{s}}_k^\top \boldsymbol{x}_j^{(k)}) = \operatorname{sign}(\bar{\boldsymbol{s}}_k^\top \boldsymbol{x}_j^{(k)})$. Using the bound from (5.9) in (5.8), we obtain

$$\begin{aligned}\left\|(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{N_k} \operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top \boldsymbol{b})\boldsymbol{x}_j^{(k)}\right\|_2^2 &\geq \left\|\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k\right\|_1^2 - \sin^2(\theta_k)\left\|\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k\right\|_1^2 \\ &= \cos^2(\theta_k)\left\|\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k\right\|_1^2 \geq \cos^2(\theta_k) N_k^2 c_{\boldsymbol{\mathcal{X}}_k,\min}^2,\end{aligned}$$

which proves the lower bound in (5.7). To show the upper bound in (5.7), we have

$$\begin{aligned}&\left\|\boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b})\right\|_2^2 = \left\|\boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k)\right\|_2^2 \\ &= \left\|\bar{\boldsymbol{s}}_k\bar{\boldsymbol{s}}_k^\top \boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k) + \left(\mathcal{P}_{\mathcal{H}_k} - \bar{\boldsymbol{s}}_k\bar{\boldsymbol{s}}_k^\top\right)\boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k)\right\|_2^2 \\ &= \left\|\bar{\boldsymbol{s}}_k\bar{\boldsymbol{s}}_k^\top \boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k)\right\|_2^2 + \left\|\left(\mathcal{P}_{\mathcal{H}_k} - \bar{\boldsymbol{s}}_k\bar{\boldsymbol{s}}_k^\top\right)\boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k)\right\|_2^2 \\ &= \left\|\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k\right\|_1^2 + \left\|\left(\mathcal{P}_{\mathcal{H}_k} - \bar{\boldsymbol{s}}_k\bar{\boldsymbol{s}}_k^\top\right)\boldsymbol{\mathcal{X}}_k \operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k)\right\|_2^2\end{aligned} \tag{5.10}$$

where the last line utilizes $\|\bar{\boldsymbol{s}}_k\|_2 = 1$. Plugging (5.10) into (5.8), we obtain

$$\begin{aligned}
&\left\|(\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{N_k}\operatorname{sign}(\boldsymbol{x}_j^{(k)\top}\boldsymbol{b})\boldsymbol{x}_j^{(k)}\right\|_2^2 \\
&= \left\|\boldsymbol{\mathcal{X}}_k^\top\bar{\boldsymbol{s}}_k\right\|_1^2 + \left\|\left(\mathcal{P}_{\mathcal{H}_k} - \bar{\boldsymbol{s}}_k\bar{\boldsymbol{s}}_k^\top\right)\boldsymbol{\mathcal{X}}_k\operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top\bar{\boldsymbol{s}}_k)\right\|_2^2 - \sin^2(\theta_k)\left\|\boldsymbol{\mathcal{X}}_k^\top\bar{\boldsymbol{s}}_k\right\|_1^2 \\
&= \cos^2(\theta_k)\left\|\boldsymbol{\mathcal{X}}_k^\top\bar{\boldsymbol{s}}_k\right\|_1^2 + \left\|\left(\mathcal{P}_{\mathcal{H}_k} - \bar{\boldsymbol{s}}_k\bar{\boldsymbol{s}}_k^\top\right)\boldsymbol{\mathcal{X}}_k\operatorname{sign}(\boldsymbol{\mathcal{X}}_k^\top\bar{\boldsymbol{s}}_k)\right\|_2^2 \\
&\leq \left(\cos(\theta_k)N_k c_{\boldsymbol{\mathcal{X}}_k,\max} + N_k\eta_{\boldsymbol{\mathcal{X}}_k}\right)^2,
\end{aligned} \tag{5.11}$$

which completes the proof of (5.7).

Let $f(\boldsymbol{b}) := \left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top\boldsymbol{b}\right\|_1$. For any critical point $\boldsymbol{b}$ of problem (5.1), there exists $\boldsymbol{v}\in\partial f(\boldsymbol{b})$ such that $(\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\boldsymbol{v} = \mathbf{0}$. Due to the general position [152, 153] assumption of the data and $\boldsymbol{b}\notin\{\pm\boldsymbol{n}_1\}$, $\boldsymbol{b}$ can be orthogonal to at most $R < D$ data points, meaning that we can write

$$\mathbf{0} = (\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\left(\sum_{k=1}^{K}\sum_{j=1}^{N_k}\operatorname{sign}(\boldsymbol{x}_j^{(k)\top}\boldsymbol{b})\boldsymbol{x}_j^{(k)} + \sum_{j=1}^{M}\operatorname{sign}(\boldsymbol{o}_j{}^\top\boldsymbol{b})\boldsymbol{o}_j + \boldsymbol{\xi}\right) \tag{5.12}$$

where $\boldsymbol{\xi} = \sum_{r=1}^R \tau_{j_r}\widetilde{\boldsymbol{x}}_{j_r}$ with $\widetilde{\boldsymbol{x}}_{j_1},\cdots,\widetilde{\boldsymbol{x}}_{j_R}$ the columns of $\widetilde{\boldsymbol{\mathcal{X}}}$ orthogonal to $\boldsymbol{b}$, and $\{\tau_{j_1},\cdots,\tau_{j_R}\}\subset[-1,1]$. Therefore, we can write

$$\begin{aligned}
0 \geq &\left\|(\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\sum_{k=1}^{K}\sum_{j=1}^{N_k}\operatorname{sign}(\boldsymbol{x}_j^{(k)\top}\boldsymbol{b})\boldsymbol{x}_j^{(k)}\right\|_2 - \left\|(\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{M}\operatorname{sign}(\boldsymbol{o}_j{}^\top\boldsymbol{b})\boldsymbol{o}_j^{(1)}\right\|_2 \\
&\qquad - \left\|(\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\sum_{r=1}^{R}\tau_{j_r}\widetilde{\boldsymbol{x}}_{j_r}\right\|_2 \\
\geq &\left\|(\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{N_1}\operatorname{sign}(\boldsymbol{x}_j^{(1)\top}\boldsymbol{b})\boldsymbol{x}_j^{(1)}\right\|_2 - \left\|(\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\sum_{k=2}^{K}\sum_{j=1}^{N_k}\operatorname{sign}(\boldsymbol{x}_j^{(k)\top}\boldsymbol{b})\boldsymbol{x}_j^{(k)}\right\|_2 - M\overline{\eta}_{\boldsymbol{\mathcal{O}}}.
\end{aligned}$$

Together with (5.7), we have

$$\begin{aligned}\cos(\theta_1)N_1c_{\boldsymbol{\mathcal{X}}_1,\min} &\le \left\|(\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{N_1}\operatorname{sign}({\boldsymbol{x}_j^{(1)}}^\top\boldsymbol{b})\boldsymbol{x}_j^{(1)}\right\|_2\\ &\le \left\|(\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\sum_{k=2}^{K}\sum_{j=1}^{N_k}\operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top\boldsymbol{b})\boldsymbol{x}_j^{(k)}\right\|_2+M\overline{\eta}_{\boldsymbol{\mathcal{O}}}\\ &\le \sum_{k=2}^{K}\left\|(\mathbf{I}-\boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{N_k}\operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top\boldsymbol{b})\boldsymbol{x}_j^{(k)}\right\|_2+M\overline{\eta}_{\boldsymbol{\mathcal{O}}}\\ &\le \sum_{k=2}^{K}\cos(\theta_k)N_kc_{\boldsymbol{\mathcal{X}}_k,\max}+\sum_{k=2}^{K}N_k\eta_{\boldsymbol{\mathcal{X}}_k}+M\overline{\eta}_{\boldsymbol{\mathcal{O}}}.\end{aligned} \tag{5.13}$$

Moreover, from [109, Lemma 12], we know that

$$\sum_{k=2}^{K}\cos(\theta_k)N_kc_{\boldsymbol{\mathcal{X}}_k,\max}\le\left[\sum_{k=2}^{K}N_k^2c_{\boldsymbol{\mathcal{X}}_k,\max}^2+2\sum_{i\neq j,i,j\neq 1}N_iN_jc_{\boldsymbol{\mathcal{X}}_i,\max}c_{\boldsymbol{\mathcal{X}}_j,\max}\cos(\theta_{ij})\right]^{\frac{1}{2}} \tag{5.14}$$

where $\theta_{ij}\in(0,\pi/2]$ is the principal angle between $\boldsymbol{n}_i$ and $\boldsymbol{n}_j$. By the definition of $\boldsymbol{W}^{\max}$ in (5.6) and $\boldsymbol{W}^{\max}_{(1,1)}$, the RHS of (5.14) can be simplified as $\sqrt{\mathbf{1}^\top\boldsymbol{W}^{\max}_{(1,1)}\mathbf{1}}$. Plugging it into (5.13), we have

$$\cos(\theta_1)\le\frac{\sqrt{\mathbf{1}^\top\boldsymbol{W}^{\max}_{(1,1)}\mathbf{1}}+\sum_{k=2}^{K}N_k\eta_{\boldsymbol{\mathcal{X}}_k}+M\overline{\eta}_{\boldsymbol{\mathcal{O}}}}{N_1c_{\boldsymbol{\mathcal{X}}_1,\min}}=\frac{1}{\zeta_1}, \tag{5.15}$$

which leads to $\sin(\theta_1)\ge\sqrt{1-1/\zeta_1^2}$ with ζ_k defined in (5.5). Since among $\{\zeta_k\}_{k=1}^K$, ζ_1 is the only one greater than 1, (5.15) is informative and well-defined, and we conclude that any critical point $\boldsymbol{b}$ of (5.1) must satisfy either

$$\boldsymbol{b}\in\{\pm\boldsymbol{n}_1\}\quad\text{or}\quad\theta_1\ge\arcsin\left(\sqrt{1-(1/\zeta_1^2)}\right),$$

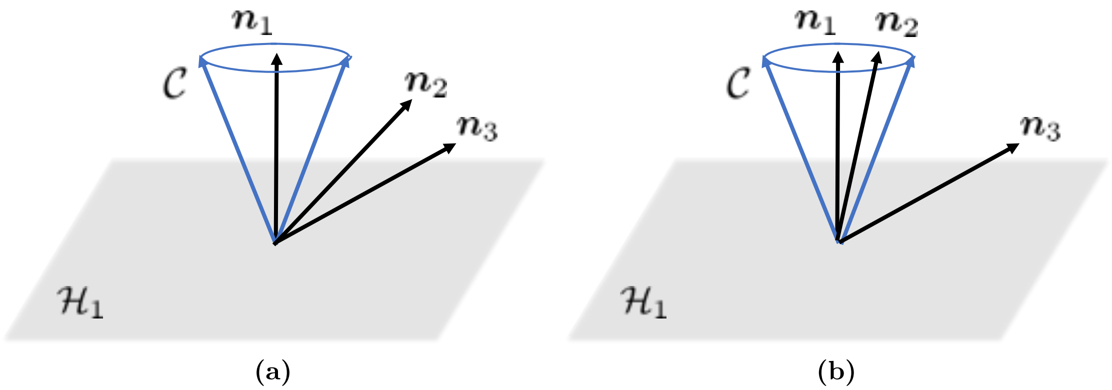


Figure 5.1. Illustration on the distribution of the critical points of problem (5.1). (a) Since $\boldsymbol{n}_2, \boldsymbol{n}_3 \notin \mathcal{C}$, they could be critical points; (b) Since $\boldsymbol{n}_2 \in \mathcal{C}$ it cannot be a critical point, but $\boldsymbol{n}_3$ could be because $\boldsymbol{n}_3 \notin \mathcal{C}$.

which completes the proof. □

Discussion of Lemma 18. Intuitively, Lemma 18 suggests that any critical point of (5.1) is either a normal vector of $\mathcal{H}_1$, or very close to $\mathcal{H}_1$ (i.e., within a region defined by the geometric dominance level of $\boldsymbol{\mathcal{X}}_1$). As the relative dominance of $\boldsymbol{\mathcal{X}}_1$ increases (larger ζ_1), the location of a critical point $\boldsymbol{b}$ becomes more restricted. In particular, Lemma 18 allows us to conclude that $\boldsymbol{n}_1$ is the single (up to direction) critical point inside of the cone $\mathcal{C} := \{\boldsymbol{y} \in \mathbb{R}^D : |\boldsymbol{y}^\top \boldsymbol{n}_1| > 1/\zeta_1, \|\boldsymbol{y}\|_2 = 1\}$ centered around $\pm\boldsymbol{n}_1$. The above observation ensures that every normal in the set $\{\pm\boldsymbol{n}_2, \cdots, \pm\boldsymbol{n}_K\}$ that lies inside of $\mathcal{C}$ is not a critical point (see Figure 5.1). We will later see in Section 5.3 how this facilitates the convergence of PRSGM to $\{\pm\boldsymbol{n}_1\}$ when it is initialized inside $\mathcal{C}$ because $\boldsymbol{n}_1$ (up to direction) is the only possible solution within the region.

5.2.3 Geometry of the global solutions

Lemma 18 is useful in helpful us understand the geometry of the global solutions of (5.1). To show that any global minimizer $\boldsymbol{b}^*$ is a normal vector to the geometrically dominant hyperplane $\mathcal{H}_1$, i.e., $\boldsymbol{b}^* \in \{\pm \boldsymbol{n}_1\}$, we need to ensure that every critical point close to $\mathcal{H}_1$ is not a global solution. Inspired by the analysis in [113], we define

$$\gamma_k := \frac{N_k c_{\boldsymbol{\mathcal{X}}_k,\min}}{\sum_{\ell \neq k} N_\ell c_{\boldsymbol{\mathcal{X}}_\ell,\max} \sin(\theta_{k\ell}) - \sqrt{\sum_{i=2}^{K-1} \lambda_i\left(\boldsymbol{W}^{\min}_{(k,k)}\right)} + M(c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min})}. \tag{5.16}$$

Here, $\boldsymbol{W}^{\min}$ is the same as $\boldsymbol{W}^{\max}$ in (5.6) by replacing $c_{\boldsymbol{\mathcal{X}}_k,\max} c_{\boldsymbol{\mathcal{X}}_\ell,\max}$ with $c_{\boldsymbol{\mathcal{X}}_k,\min} c_{\boldsymbol{\mathcal{X}}_\ell,\min}$, and $\lambda_1(\boldsymbol{A}) \geq \cdots \geq \lambda_n(\boldsymbol{A})$ are the eigenvalues of an n-by-n matrix $\boldsymbol{A}$. In fact, we can show that every global solution of (5.1) is not far from $\{\pm \boldsymbol{n}_1\}$ in the sense that its principal angle θ from $\boldsymbol{n}_1$ satisfies $\theta \leq \arcsin(1/\gamma_1)$. Combining this fact with Lemma 18 establishes our main theoretical result as follows.

Theorem 12. *Any global solution $\boldsymbol{b}^*$ of problem* (5.1) *is a normal vector to the geometrically dominant hyperplane $\mathcal{H}_1$ if*

$$\frac{1}{\zeta_1^2} + \frac{1}{\gamma_1^2} < 1. \tag{5.17}$$

Proof. To show that any global minimizer $\boldsymbol{b}^*$ is a normal vector to $\mathcal{H}_1$, we first prove that every critical point close to $\mathcal{H}_1$ is not a global solution. In other words, any global solution of (5.1) is not far from $\{\pm \boldsymbol{n}_1\}$.

For $k \in [K] \backslash \{1\}$, recall that $\theta_{k\ell} \in [0, \pi/2]$ is defined as the principal angle between

$\boldsymbol{n}_k$ and $\boldsymbol{n}_\ell$. We rewrite $\boldsymbol{n}_1 = \sin(\theta_{1k})\bar{\boldsymbol{s}}_k + \cos(\theta_{1k})\bar{\boldsymbol{n}}_k$, where $\bar{\boldsymbol{s}}_k \in \mathcal{H}_k, \bar{\boldsymbol{n}}_k \in \mathcal{H}_k^\perp$, and $\|\bar{\boldsymbol{s}}_k\|_2 = \|\bar{\boldsymbol{n}}_k\|_2 = 1$. Since $\boldsymbol{b}^*$ is a global minimizer, we have

$$\begin{aligned}
\left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}^*\right\|_1 &= \sum_{k=1}^{K} \left\|\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b}^*\right\|_1 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{b}^*\right\|_1 \\
&\leq \sum_{k=2}^{K} \left\|\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{n}_1\right\|_1 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{n}_1\right\|_1 \\
&= \sum_{k=2}^{K} \left\|\boldsymbol{\mathcal{X}}_k^\top \left(\sin(\theta_{1k})\bar{\boldsymbol{s}}_k + \cos(\theta_{1k})\bar{\boldsymbol{n}}_k\right)\right\|_1 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{n}_1\right\|_1 \qquad (5.18) \\
&= \sum_{k=2}^{K} \sin(\theta_{1k}) \left\|\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}_k\right\|_1 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{n}_1\right\|_1 \\
&\leq \sum_{k=2}^{K} \sin(\theta_{1k}) N_k c_{\boldsymbol{\mathcal{X}}_k,\max} + M c_{\boldsymbol{\mathcal{O}},\max}
\end{aligned}$$

where we used the fact that $\bar{\boldsymbol{n}}_k$ is orthogonal to $\boldsymbol{\mathcal{X}}_k$. On the other hand, for all $k \in [K]$, we decompose $\boldsymbol{b}^* = \sin(\theta_k)\bar{\boldsymbol{s}}'_k + \cos(\theta_k)\bar{\boldsymbol{n}}'_k$ where θ_k is the principal angle between $\boldsymbol{b}^*$ and $\boldsymbol{n}_k$, $\bar{\boldsymbol{s}}'_k \in \mathcal{H}_k$, $\bar{\boldsymbol{n}}'_k \in \mathcal{H}_k^\perp$, and $\|\bar{\boldsymbol{s}}'_k\|_2 = \|\bar{\boldsymbol{n}}'_k\|_2 = 1$. Then we have

$$\begin{aligned}
\left\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}^*\right\|_1 &= \sum_{k=1}^{K} \left\|\boldsymbol{\mathcal{X}}_k^\top \boldsymbol{b}^*\right\|_1 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{b}^*\right\|_1 \\
&= \sum_{k=1}^{K} \left\|\boldsymbol{\mathcal{X}}_k^\top \left(\sin(\theta_k)\bar{\boldsymbol{s}}'_k + \cos(\theta_k)\bar{\boldsymbol{n}}'_k\right)\right\|_1 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{b}^*\right\|_1 \\
&= \sin(\theta_1) \left\|\boldsymbol{\mathcal{X}}_1^\top \bar{\boldsymbol{s}}'_1\right\|_1 + \sum_{k=2}^{K} \sin(\theta_k) \left\|\boldsymbol{\mathcal{X}}_k^\top \bar{\boldsymbol{s}}'_k\right\|_1 + \left\|\boldsymbol{\mathcal{O}}^\top \boldsymbol{b}^*\right\|_1 \qquad (5.19) \\
&\geq \sin(\theta_1) N_1 c_{\boldsymbol{\mathcal{X}}_1,\min} + \sum_{k=2}^{K} \sin(\theta_k) N_k c_{\boldsymbol{\mathcal{X}}_k,\min} + M c_{\boldsymbol{\mathcal{O}},\min} \\
&\geq \sin(\theta_1) N_1 c_{\boldsymbol{\mathcal{X}}_1,\min} + \left[\sum_{k=2}^{K} N_k^2 c_{\boldsymbol{\mathcal{X}}_k,\min}^2 - \lambda_1(\boldsymbol{W}_{(1,1)}^{\min})\right]^{\frac{1}{2}} + M c_{\boldsymbol{\mathcal{O}},\min}
\end{aligned}$$

where we used the fact that $\boldsymbol{n}'_k$ is orthogonal to $\boldsymbol{\mathcal{X}}_k$, and the last inequality follows from [109, Lemma 13]. Note that since

$$\sum_{k=2}^{K} N_k^2 c_{\boldsymbol{\mathcal{X}}_k,\min}^2 = \operatorname{trace}(\boldsymbol{W}_{(1,1)}^{\min}) = \sum_{i=1}^{K-1} \lambda_i(\boldsymbol{W}_{(1,1)}^{\min}),$$

we can simplify (5.19) and write it as

$$\left\| \widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}^* \right\|_1 \geq \sin(\theta_1) N_1 c_{\boldsymbol{\mathcal{X}}_1,\min} + \sqrt{\sum_{i=2}^{K-1} \lambda_i(\boldsymbol{W}_{(1,1)}^{\min})} + M c_{\boldsymbol{\mathcal{O}},\min}. \tag{5.20}$$

Combining (5.18) and (5.20), we obtain

$$\begin{aligned} &\sin(\theta_1) \\ \leq\ & \frac{\sum_{k=2}^{K} \sin(\theta_{1k}) N_k c_{\boldsymbol{\mathcal{X}}_k,\max} - \sqrt{\sum_{i=2}^{K-1} \lambda_i(\boldsymbol{W}_{(1,1)}^{\min})} + M(c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min})}{N_1 c_{\boldsymbol{\mathcal{X}}_1,\min}} = \frac{1}{\gamma_1}, \end{aligned} \tag{5.21}$$

which indicates that any global minimizer $\boldsymbol{b}^*$ must be close enough to $\{\pm \boldsymbol{n}_1\}$ such that its principal angle θ_1 from $\boldsymbol{n}_1$ satisfies $\theta_1 \leq \arcsin\left(1/\gamma_1\right)$.

We now prove Theorem 12 by contradiction. Suppose there exists a global minimizer $\boldsymbol{b}^*$ that satisfies $\boldsymbol{b}^* \notin \{\pm \boldsymbol{n}_1\}$, then by Lemma 18, we have $\cos(\theta_1) \leq 1/\zeta_1$. Moreover, (5.21) tells us that $\sin(\theta_1) \leq 1/\gamma_1$, which when combined together yields

$$1 = \sin^2(\theta_1) + \cos^2(\theta_1) \leq \frac{1}{\zeta_1^2} + \frac{1}{\gamma_1^2},$$

which contradicts (5.17), and thus completes the proof. □

We first give additional interpretations of γ_k. Note that γ_k is similar to ζ_k, which characterizes the relative dominance of $\boldsymbol{\mathcal{X}}_k$ from a different perspective. First, the term $M(c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min})$ in the denominator of (5.16) represents the impact of outliers: uniformly distributed outliers with $M \to \infty$ cause the difference $c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min}$ to vanish, making the term small (see Lemma 2). Next, to better understand the square root part in (5.16), for simplicity we consider the equi-angular case for $\{\mathcal{H}_\ell\}_{\ell \neq k}$ such that $\theta_{ij} \equiv \theta'$ for all $i, j \neq k, i \neq j$. Then, one can obtain $\sum_{i=2}^{K-1} \lambda_i(\boldsymbol{W}_{(k,k)}^{\min}) = (1 - \cos(\theta')) \sum_{\ell \neq k,r} N_\ell^2 c_{\boldsymbol{\mathcal{X}}_\ell,\min}^2$, where $r = \arg\max_{\ell \neq k} N_\ell c_{\boldsymbol{\mathcal{X}}_\ell,\min}$. For a global solution to be a normal of $\mathcal{H}_k$, one may expect: (i) a large relative disparity in significance between $\boldsymbol{\mathcal{X}}_k$ and $\boldsymbol{\mathcal{X}}_\ell$ for all $\ell \neq k$ so that $\frac{N_k c_{\boldsymbol{\mathcal{X}}_k,\min}}{N_\ell c_{\boldsymbol{\mathcal{X}}_\ell,\max}}$ is large; (ii) $\mathcal{H}_k$ to be relatively close to the other planes so that the energy concentrated around $\mathcal{H}_k$ is relatively large, i.e., $\theta_{k\ell}$ is relatively small; and (iii) the other planes $\{\mathcal{H}_\ell\}_{\ell \neq k}$ are relatively well separated so that the energy concentrated around any of them is relatively small, i.e., θ' is relatively large. All these conditions lead to γ_k being large.

Discussion of Theorem 12. An interpretation of Theorem 12 follows from the above discussion about ζ_k and γ_k: for a fixed number of inliers $\{N_k\}$ and outliers M, if data points are well-distributed (large $c_{\boldsymbol{\mathcal{X}}_k,\min}$, small $c_{\boldsymbol{\mathcal{X}}_k,\max}$, small $\eta_{\boldsymbol{\mathcal{X}}_k}$, small $\eta_{\boldsymbol{\mathcal{O}}}$, small $c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min}$) and $\mathcal{H}_1$ is closer to the other planes (relatively small $\theta_{1\ell}, \ell \neq 1$) than the other planes are to each other (relatively large $\theta_{ij}, i, j \neq 1, i \neq j$), then both ζ_1 and γ_1 tend to be large, (5.17) is more likely to be satisfied, and any global minimizer is a normal vector of $\mathcal{H}_1$. In contrast to the discrete result in [113], which is based on a continuous variant of (5.1) without outliers and uses quantities such as the *spherical cap discrepancy* or *circumradii of polytopes* that are difficult to interpret, the global geometric analysis here focuses on the discrete problem (5.1) and leverages geometric quantities to explicitly characterize the underlying distribution of both inliers and outliers. Finally, when the dataset is further contaminated with noise, one may expect that the error between the global minimizer and the true normal vector to $\mathcal{H}_1$ to be proportional to the noise level, as analyzed for a single subspace case in Section 3.1.2. The extension to noisy data is by no means trivial, and we leave it as future work.

5.2.4 Probabilistic analysis

In this section, we present a probabilistic characterization of the global optimal solutions of problem (5.1) under a UoH model. We first explicitly state the random spherical model for a UoH that we will consider.

Definition 5 (Random spherical model for a UoH). *Consider a random spherical model where the M columns of $\boldsymbol{\mathcal{O}}$ are drawn uniformly from the sphere $\mathbb{S}^{D-1}$, and the N_k columns of $\boldsymbol{\mathcal{X}}_k$ are drawn uniformly from $\mathbb{S}^{D-1} \cap \mathcal{H}_k$ for $k \in [K]$, where $\mathcal{H}_k$ is a given hyperplane in $\mathbb{R}^D$ with $\dim(\mathcal{H}_k) = D - 1$.*

Compared with the random spherical model for a single subspace given in Definition 1, Definition 5 specifies a similar generative model for data drawn from a UoH. Note that we already have the concentration properties for the outlier-related quantities, namely $c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min}$ and $\eta_{\boldsymbol{\mathcal{O}}}$, under such models as stated in Lemma 2. On the other hand, the concentration bounds for the inlier-related geometric quantities, namely $c_{\boldsymbol{\mathcal{X}}_k,\min}, c_{\boldsymbol{\mathcal{X}}_k,\max}$ and $\eta_{\boldsymbol{\mathcal{X}}_k}$, follow from [152, Lemma 4], which leads to

$$\begin{aligned}
&\mathbb{P}\left[c_{\boldsymbol{\mathcal{X}}_k,\min} \geq \bar{C}_0 - \left(2+\frac{t}{2}\right)/\sqrt{N_k}\right] \geq 1-2e^{-\frac{t^2}{2}}, \\
&\mathbb{P}\left[c_{\boldsymbol{\mathcal{X}}_k,\max} \leq \bar{C}_0 + \left(2+\frac{t}{2}\right)/\sqrt{N_k}\right] \geq 1-2e^{-\frac{t^2}{2}}, \quad \text{and} \\
&\mathbb{P}\left[\eta_{\boldsymbol{\mathcal{X}}_k} \leq C_1\left(\sqrt{D}\log D + t\right)/\sqrt{N_k}\right] \geq 1-2e^{-\frac{t^2}{2}}
\end{aligned} \tag{5.22}$$

for any number $t > 0$, where

$$\begin{aligned}
\bar{C}_0 &:= \frac{(D-3)!!}{(D-2)!!} \cdot \begin{cases} \frac{2}{\pi}, & \text{for even } D, \\ 1, & \text{for odd } D, \end{cases} \\
n!! &:= \begin{cases} n(n-2)(n-4)\cdots 4\cdot 2, & \text{if } n \text{ is even}, \\ n(n-2)(n-4)\cdots 3\cdot 1, & \text{if } n \text{ is odd}, \end{cases}
\end{aligned} \tag{5.23}$$

and C_1 is a universal constant that is independent of K, $\{N_k\}$, M, D and t. We are now able to state our probabilistic result.

Theorem 13. *For the random spherical model in Definition 5, the probability that any global solution of* (5.1) *is a normal vector of* $\mathcal{H}_1$ *is at least* $1-2(K+1)e^{-t^2/2}$*, where* $t > 0$ *satisfies, with* $\bar{C}_0$ *defined in* (5.23)*, the inequality*

$$\begin{aligned}
\bar{C}_0 \sum_{k\neq 1} N_k + \left(C_1\sqrt{D}\log(D) + \frac{3t}{2}\right)\sum_{k\neq 1}\sqrt{N_k} + \left(C_2\sqrt{D}\log D + t\right)\sqrt{M} \\
< \bar{C}_0 N_1 - \left(\sqrt{2} + \frac{t}{2\sqrt{2}}\right)\sqrt{N_1}
\end{aligned} \tag{5.24}$$

where C_1, C_2 *are universal constants that are independent of* K*,* $\{N_k\}$*,* M*,* D *and* t*.*

Proof. We first note that

$$\begin{aligned}\widetilde{\zeta}_1 &:= \frac{N_1 c_{\boldsymbol{\mathcal{X}}_1,\min}}{\sum_{k\neq 1} N_k c_{\boldsymbol{\mathcal{X}}_k,\max} + \sum_{k\neq 1} N_k \eta_{\boldsymbol{\mathcal{X}}_k} + M\overline{\eta}_{\boldsymbol{\mathcal{O}}}} < \zeta_1 \quad \text{and} \\ \widetilde{\gamma}_1 &:= \frac{N_1 c_{\boldsymbol{\mathcal{X}}_1,\min}}{\sum_{k\neq 1} N_k c_{\boldsymbol{\mathcal{X}}_k,\max} + M(c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min})} < \gamma_1,\end{aligned}$$

so that

$$\frac{1}{\widetilde{\zeta}_1^2} + \frac{1}{\widetilde{\gamma}_1^2} < 1, \tag{5.25}$$

meaning that (5.17) holds. Therefore, any global solution of (5.1) is a normal vector of $\mathcal{H}_1$. Therefore, Theorem 13 follows directly from Theorem 12 by plugging the concentrations for $c_{\boldsymbol{\mathcal{O}},\max} - c_{\boldsymbol{\mathcal{O}},\min}, \eta_{\boldsymbol{\mathcal{O}}}$ from (3.9) and $c_{\boldsymbol{\mathcal{X}}_k,\min}, c_{\boldsymbol{\mathcal{X}}_k,\max}, \eta_{\boldsymbol{\mathcal{X}}_k}$ from (5.22) into (5.25). □

Discussion of Theorem 13. Note that $\bar{C}_0 \in \left[\sqrt{\frac{2}{\pi(D-1)}}, \sqrt{\frac{1}{D-1}}\right]$ (see [152, footnote 9]) is a constant for fixed D. As the number of inliers from the hyperplanes goes to infinity and the other parameters are fixed, (5.24) roughly requires $\sum_{k\neq 1} N_k < N_1$, which coincides with (5.3) of [64] (in expectation). Also, as the number of inliers goes to infinity, (5.24) implies that the DPCP approach for a UoH can tolerate $M = O((N_1 - \sum_{k\neq 1} N_k)/D)^2)$ outliers, which generalizes the result in [152, 153] for a single subspace. Finally, since (5.24) is linear in t, it gives an upper bound for t, which is roughly $O((N_1 - \sum_{k\geq 2} N_k - \sqrt{M})/(\sum_k \sqrt{N} + \sqrt{M}))$.

A similar probabilistic result is provided in [64, Theorem 1.1] but for a different generative model where the number of points sampled in each hyperplane is not fixed in advance, as opposed to M and $\{N_k\}$ here, but is controlled by the sampling weights $\{\psi_k\}_{k=0}^K$ (see (5.3)). With this difference in mind, we now compare [64, Theorem 1.1] with (5.24). Towards that goal, dividing both sides of (5.24) by the total number of data points $N+M$, and viewing $\frac{M}{N+M}$ as ψ_0 and $\frac{N_k}{N+M}$ as ψ_k, gives

$$\psi_1 > \sum_{k=2}^K \psi_k + \frac{3\sqrt{D}\cdot t + \rho(D)}{\sqrt{N+M}} \sum_{k=0}^K \sqrt{\psi_k}, \tag{5.26}$$

where $\rho(D) := \sqrt{2}D\log D \max(C_1, C_2)$. Our result and [64, Theorem 1.1] require a similar condition on ψ_k to guarantee that any global solution of (5.1) is a normal vector of $\mathcal{H}_1$ with certain probability. On one hand, (5.26) requires ψ_1 to be larger than $\sum_{k=2}^K \psi_k$ by a positive amount (which goes to 0 if the total number of points goes to infinity), which is slightly stronger than (5.3) in [64]. On the other hand, [64, Theorem 1.1] only ensures a probability of $1 - C_3 \exp\left(-\frac{N+M}{C_4}\right)$, where $C_3 = O\left(D^{D(D-1)/2} + D^{8(D-1)}\right)$ and $C_4 = O\left(D^{16}\right)$ (assuming the other parameters such as K are fixed), thus needing to sample $\Omega(D^{18}\log D)$ points to make the probability overwhelming (e.g., probability of $1 - O(\exp(-D))$ if $N + M = \Omega(D^{19}\log D)$). For comparison, by taking $t = \sqrt{\frac{N+M}{D^3}}$, Theorem 13 now requires ψ_1 to be larger than $\sum_{k=2}^K \psi_k$ by a small amount of $\left(\frac{3}{D} + \frac{\rho(D)}{\sqrt{N+M}}\right)\sum\sqrt{\psi_k}$ and guarantees with probability $1-2(K+1)\exp\left(-\frac{N+M}{2D^3}\right)$, which only requires a total sampling of $\Omega(D^3)$ points to make the probability overwhelming (e.g., probability of $1-O(\exp(-D))$ if $N+M = \Omega(D^4)$), which is much smaller than the $\Omega(D^{18}\log D)$ needed in [64].

Algorithm 3 Projected Riemannian Sub-Gradient Method (PRSGM) for solving (5.1)

```
1: Initialization: $\boldsymbol{b}_0 \in \mathbb{S}^{D-1}$, step size $\mu_0$, and $\beta \in (0,1)$.
2: for $t = 0, 1, 2, \cdots$ do
3:     Compute a Riemannian subgradient: $\mathcal{G}(\boldsymbol{b}_t) \leftarrow (\mathbf{I} - \boldsymbol{b}_t\boldsymbol{b}_t^\top)\widetilde{\boldsymbol{\mathcal{X}}}\operatorname{sign}(\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}_t)$;
4:     Update the step size in a geometrically diminishing fashion: $\mu_t \leftarrow \mu_0\beta^t$;
5:     Update the iterate:
           $\widehat{\boldsymbol{b}}_{t+1} \leftarrow \boldsymbol{b}_t - \mu_t\mathcal{G}(\boldsymbol{b}_t)$ and $\boldsymbol{b}_{t+1} \leftarrow \widehat{\boldsymbol{b}}_{t+1}/\|\widehat{\boldsymbol{b}}_{t+1}\|_2$;
6: end for
```

5.3 Projected Riemannian Sub-Gradient method for learning a union of hyperplanes

In Section 5.2, we have shown that the non-convex DPCP problem (5.1) is effective in robustly recovering a specific hyperplane for a UoH. The work of [113] proposed to solve (5.1) by either an LP-based algorithm that involves a sequence of convex optimization problems thus is computationally expensive, or an IRLS algorithm that requires doing an SVD in each iteration and lacks a convergence guarantee. In this work, motivated by the Projected Riemannian Sub-Gradient Method (PRSGM) analyzed in Chapter 4 for solving optimization problems over the Grassmannian (Section 4.2.3) and its successful application to the DPCP problem (2.9) for learning a single subspace (Section 4.2.4), we now extend it to solve problem (5.1) with data drawn from a UoH.

As summarized in Algorithm 3, we apply the general PRSGM framework (see Algorithm 1) for solving (5.1) and focus on its convergence to the geometrically dominant hyperplane $\mathcal{H}_1$. In particular, each iterate of the PRSGM computes a natural Riemannian subgradient $(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\widetilde{\boldsymbol{\mathcal{X}}}\operatorname{sign}(\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b})$, which only involves matrix-vector multiplications, hence is computationally efficient compared with solving an LP. Moreover, since

PRSGM has been proved (see Theorem 10 and Theorem 11) to converge to a global solution at a linear rate with appropriate initialization and geometrically diminishing step size in the single subspace case, we extend this analysis to the UoH model and prove a linear convergence rate. Towards that goal, we measure the distance between any vector $\boldsymbol{b} \in \mathbb{S}^{D-1}$ and our target solution set $\{\pm \boldsymbol{n}_1\}$ by

$$\operatorname{dist}(\boldsymbol{b}, \{\pm \boldsymbol{n}_1\}) = \min(\|\boldsymbol{b} - \boldsymbol{n}_1\|_2, \|\boldsymbol{b} + \boldsymbol{n}_1\|_2),$$

which is a special case of (4.2). Also, it is clear that

$$\mathcal{P}_{\{\pm \boldsymbol{n}_1\}}(\boldsymbol{b}) = \operatorname{sign}(\boldsymbol{b}^\top \boldsymbol{n}_1)\boldsymbol{n}_1.$$

The next result establishes the Riemannian regularity condition (RRC) (see Definition 3) for problem (5.1), which we use to obtain a linear convergence rate.

Lemma 19. *For any* $\epsilon \in \left(0, \sqrt{2(1 - 1/\zeta_1)}\right)$ *and* $\alpha = \frac{\sqrt{2}}{2} N_1 c_{\boldsymbol{\mathcal{X}}_1,\min} \left((1 - \epsilon^2/2) - 1/\zeta_1\right)$ *with* ζ_1 *defined in* (5.5)*, the DPCP problem* (5.1) *satisfies the following* $(\alpha, \epsilon, \boldsymbol{n}_1)$*-RRC: for every* $\boldsymbol{b} \in \mathbb{S}^{D-1}$ *satisfying* $\operatorname{dist}(\boldsymbol{b}, \{\pm \boldsymbol{n}_1\}) \leq \epsilon$*, we have*

$$\langle \operatorname{sign}(\boldsymbol{b}^\top \boldsymbol{n}_1)\boldsymbol{n}_1 - \boldsymbol{b}, -(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\widetilde{\boldsymbol{\mathcal{X}}} \operatorname{sign}(\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b})\rangle \geq \alpha \operatorname{dist}(\boldsymbol{b}, \{\pm \boldsymbol{n}_1\}). \tag{5.27}$$

Proof. First note that for any $\epsilon \in \left(0, \sqrt{2(1 - 1/\zeta_1)}\right)$ and $\operatorname{dist}(\boldsymbol{b}, \{\pm \boldsymbol{n}_1\}) \leq \epsilon$, we must have $\boldsymbol{b} \notin \mathcal{H}_1$ and thus $\operatorname{sign}(\boldsymbol{b}^\top \boldsymbol{n}_1) \neq 0$. In fact, if $\boldsymbol{b} \in \mathcal{H}_1$, then

$$\operatorname{dist}(\boldsymbol{b}, \{\pm \boldsymbol{n}_1\}) = \min(\|\boldsymbol{b} - \boldsymbol{n}_1\|_2, \|\boldsymbol{b} + \boldsymbol{n}_1\|_2) = \sqrt{2} > \epsilon.$$

Without loss of generality, let us assume $\operatorname{sign}(\boldsymbol{b}^\top \boldsymbol{n}_1) > 0$ since the analysis for the case of $\operatorname{sign}(\boldsymbol{b}^\top \boldsymbol{n}_1) < 0$ is similar. For any $\boldsymbol{b} \in \mathbb{S}^{D-1}$, the projection of $\boldsymbol{b}$ onto $\{\pm \boldsymbol{n}_1\}$ is

$$\mathcal{P}_{\{\pm \boldsymbol{n}_1\}}(\boldsymbol{b}) = \underset{\boldsymbol{z} \in \{\pm \boldsymbol{n}_1\}}{\arg\min} \|\boldsymbol{z} - \boldsymbol{b}\|_2 = \operatorname{sign}(\boldsymbol{b}^\top \boldsymbol{n}_1)\boldsymbol{n}_1 = \boldsymbol{n}_1.$$

Letting $\theta_1 \in [0, \pi/2)$ be the angle between $\boldsymbol{b}$ and $\boldsymbol{n}_1$, we can write

$$\boldsymbol{b} = \sin(\theta_1)\boldsymbol{s}_1 + \cos(\theta_1)\boldsymbol{n}_1$$

where $\boldsymbol{s}_1 \in \mathcal{H}_1 \cap \mathbb{S}^{D-1}$. Next, we define

$$\begin{aligned}\boldsymbol{g} &:= (\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\boldsymbol{n}_1 = \boldsymbol{n}_1 - \boldsymbol{b}(\boldsymbol{b}^\top \boldsymbol{n}_1) \\ &= \boldsymbol{n}_1 - (\sin(\theta_1)\boldsymbol{s}_1 + \cos(\theta_1)\boldsymbol{n}_1)\cos(\theta_1) \\ &= \sin(\theta_1)(-\cos(\theta_1)\boldsymbol{s}_1 + \sin(\theta_1)\boldsymbol{n}_1) = \sin(\theta_1)\widehat{\boldsymbol{g}}\end{aligned} \tag{5.28}$$

where $\widehat{\boldsymbol{g}} = -\cos(\theta_1)\boldsymbol{s}_1 + \sin(\theta_1)\boldsymbol{n}_1$ is orthogonal to $\boldsymbol{b}$ and $\|\widehat{\boldsymbol{g}}\|_2 = 1$. We have

$$\begin{aligned}
&\left\langle \operatorname{sign}(\boldsymbol{b}^\top \boldsymbol{n}_1)\boldsymbol{n}_1 - \boldsymbol{b}, -(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\widetilde{\boldsymbol{\mathcal{X}}}\operatorname{sign}(\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}) \right\rangle \\
&= \left\langle \boldsymbol{n}_1, -(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\widetilde{\boldsymbol{\mathcal{X}}}\operatorname{sign}(\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}) \right\rangle \\
&= -\left\langle \widetilde{\boldsymbol{\mathcal{X}}}\operatorname{sign}(\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}), (\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\boldsymbol{n}_1 \right\rangle = -\left\langle \widetilde{\boldsymbol{\mathcal{X}}}\operatorname{sign}(\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}), \boldsymbol{g} \right\rangle \\
&= -\left\langle \sum_{j=1}^{N_1} \operatorname{sign}({\boldsymbol{x}_j^{(1)}}^\top \boldsymbol{b})\boldsymbol{x}_j^{(1)} + \sum_{k=2}^{K}\sum_{j=1}^{N_k} \operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top \boldsymbol{b})\boldsymbol{x}_j^{(k)} + \sum_{j=1}^{M} \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{b})\boldsymbol{o}_j, \boldsymbol{g} \right\rangle \\
&= \sin(\theta_1)\cos(\theta_1)\left\|\boldsymbol{\mathcal{X}}_1^\top \boldsymbol{s}_1\right\|_1 - \left\langle \sum_{k=2}^{K}\sum_{j=1}^{N_k} \operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top \boldsymbol{b})\boldsymbol{x}_j^{(k)} + \sum_{j=1}^{M} \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{b})\boldsymbol{o}_j, \boldsymbol{g} \right\rangle
\end{aligned} \tag{5.29}$$

where we used the result in (5.28). Now consider the second term in (5.29):

$$\begin{aligned}
&\left\langle \sum_{k=2}^{K}\sum_{j=1}^{N_k} \operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top \boldsymbol{b})\boldsymbol{x}_j^{(k)} + \sum_{j=1}^{M} \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{b})\boldsymbol{o}_j, \boldsymbol{g} \right\rangle \\
&= \sin(\theta_1)\left\langle \sum_{k=2}^{K}\sum_{j=1}^{N_k} \operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top \boldsymbol{b})\boldsymbol{x}_j^{(k)} + \sum_{j=1}^{M} \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{b})\boldsymbol{o}_j, \widehat{\boldsymbol{g}} \right\rangle \\
&= \sin(\theta_1)\left\langle (\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\sum_{k=2}^{K}\sum_{j=1}^{N_k} \operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top \boldsymbol{b})\boldsymbol{x}_j^{(k)}, \widehat{\boldsymbol{g}} \right\rangle \\
&\qquad + \sin(\theta_1)\left\langle (\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{M} \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{b})\boldsymbol{o}_j, \widehat{\boldsymbol{g}} \right\rangle \\
&\le \sin(\theta_1)\sum_{k=2}^{K}\left\| (\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{N_k} \operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top \boldsymbol{b})\boldsymbol{x}_j^{(k)} \right\|_2 \\
&\qquad + \sin(\theta_1)\left\| (\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{M} \operatorname{sign}(\boldsymbol{o}_j^\top \boldsymbol{b})\boldsymbol{o}_j \right\|_2 \\
&\le \sin(\theta_1)\left(\sum_{k=2}^{K}\cos(\theta_k)N_k c_{\boldsymbol{\mathcal{X}}_k,\max} + \sum_{k=2}^{K} N_k \eta_{\boldsymbol{\mathcal{X}}_k} + M\eta_{\boldsymbol{\mathcal{O}}} \right) \\
&\le \sin(\theta_1)\frac{N_1 c_{\boldsymbol{\mathcal{X}}_1,\min}}{\zeta_1}
\end{aligned} \tag{5.30}$$

where the first equality follows from $\boldsymbol{g} = \sin(\theta_1)\widehat{\boldsymbol{g}}$ in (5.28), the second equality follows from $(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\widehat{\boldsymbol{g}} = \widehat{\boldsymbol{g}}$, the first inequality follows from the Cauchy-Schwartz inequality and $\|\widehat{\boldsymbol{g}}\|_2 = 1$, the second inequality follows from (5.7), and the last inequality follows from the definition of ζ_1 in (5.5). Plugging (5.30) back into (5.29), we obtain

$$\begin{aligned}&\left\langle \operatorname{sign}(\boldsymbol{b}^\top \boldsymbol{n}_1)\boldsymbol{n}_1 - \boldsymbol{b}, -(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\widetilde{\boldsymbol{\mathcal{X}}} \operatorname{sign}(\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b})\right\rangle \\ &\geq \sin(\theta_1) N_1 c_{\boldsymbol{\mathcal{X}}_1,\min}(\cos(\theta_1) - 1/\zeta_1).\end{aligned} \tag{5.31}$$

Since $\operatorname{dist}(\boldsymbol{b}, \{\pm\boldsymbol{n}_1\}) = \min(\|\boldsymbol{b}-\boldsymbol{n}_1\|_2, \|\boldsymbol{b}+\boldsymbol{n}_1\|_2)$, and θ_1 is the principal angle between $\boldsymbol{b}$ and $\boldsymbol{n}_1$, we obtain

$$\operatorname{dist}^2(\boldsymbol{b}, \{\pm\boldsymbol{n}_1\}) = \|\boldsymbol{b}\|_2^2 + \|\boldsymbol{n}_1\|_2^2 - 2\boldsymbol{b}^\top \boldsymbol{n}_1 = 2 - 2\cos(\theta_1). \tag{5.32}$$

Moreover, from Proposition 2, we have

$$\operatorname{dist}(\boldsymbol{b}, \{\pm\boldsymbol{n}_1\}) \leq \sqrt{2}\sin(\theta_1). \tag{5.33}$$

For any $\epsilon > 0$ such that $\operatorname{dist}(\boldsymbol{b}, \{\pm\boldsymbol{n}_1\}) \leq \epsilon$, from (5.32) we have

$$\cos(\theta_1) = \frac{2 - \operatorname{dist}^2(\boldsymbol{b}, \{\pm\boldsymbol{n}_1\})}{2} \geq \frac{2-\epsilon^2}{2}. \tag{5.34}$$

According to (5.31), and making use of (5.33) and (5.34), we have

$$\begin{aligned}&\left\langle \operatorname{sign}(\boldsymbol{b}^\top \boldsymbol{n}_1)\boldsymbol{n}_1 - \boldsymbol{b}, -(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\widetilde{\boldsymbol{\mathcal{X}}} \operatorname{sign}(\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b})\right\rangle \\ &\geq \sin(\theta_1) N_1 c_{\boldsymbol{\mathcal{X}}_1,\min}(\cos(\theta_1) - 1/\zeta_1). \\ &\geq \frac{\sqrt{2}}{2} N_1 c_{\boldsymbol{\mathcal{X}}_1,\min}\left(\left(1 - \frac{\epsilon^2}{2}\right) - \frac{1}{\zeta_1}\right)\operatorname{dist}(\boldsymbol{b}, \{\pm\boldsymbol{n}_1\}).\end{aligned} \tag{5.35}$$

To ensure that the RHS of (5.35) is nonnegative, we require $\epsilon < \sqrt{2(1-1/\zeta_1)}$, which completes the proof. □

Discussion of Lemma 19. In words, (5.27) guarantees that when $\boldsymbol{b}$ is close to a target solution $\pm\boldsymbol{n}_1$ (a normal vector of the geometrically dominant hyperplane $\mathcal{H}_1$), the negative Riemannian subgradient points toward the target solution. The choice of ϵ and α in Lemma 19 depends on the geometric dominance level of $\boldsymbol{\mathcal{X}}_1$. A larger dominance level for $\boldsymbol{\mathcal{X}}_1$ (larger ζ_1) leads to a larger ϵ (i.e., a larger initialization region) and larger α (i.e., the negative Riemannian subgradient points closer to $\pm\boldsymbol{n}_1$). Using the RRC in (5.27), we are now able to apply Theorem 9 to obtain a convergence result for Algorithm 3.

Theorem 14. *Let $\{\boldsymbol{b}_t\}$ be the sequence generated by Algorithm 3 for solving problem (5.1) with initialization $\boldsymbol{b}_0$ satisfying $\widehat{\theta}_0 = \arccos(|\boldsymbol{n}_1^\top \boldsymbol{b}_0|) < \arccos(1/\zeta_1)$ and step size $\mu_t = \mu_0\beta^t$ such that*

$$0 < \mu_0 \le \frac{\alpha\epsilon}{2\xi^2} \quad \textit{and} \quad 1 > \beta \ge \sqrt{1 - 2\frac{\alpha\mu_0}{\epsilon} + \frac{\mu_0^2\xi^2}{\epsilon^2}},$$

where $\epsilon = \sqrt{2(1-\cos(\widehat{\theta}_0))}$, $\alpha = \frac{\sqrt{2}}{2} N_1 c_{\boldsymbol{\mathcal{X}}_1,\min}\Big(\cos(\widehat{\theta}_0) - 1/\zeta_1\Big)$, and

$$\xi = \sqrt{\mathbf{1}^\top \boldsymbol{W}^{\max}\mathbf{1}} + \sum_{k=1}^{K} N_k \eta_{\boldsymbol{\mathcal{X}}_k} + M\eta_{\boldsymbol{\mathcal{O}}} + D. \tag{5.36}$$

Then, the principal angle $\widehat{\theta}_t$ between $\boldsymbol{b}_t$ and $\boldsymbol{n}_1$ decays at a linear rate:

$$\sin(\widehat{\theta}_t) \le \epsilon \cdot \beta^t \textit{ for all } t \ge 0.$$

Proof. First note that

$$\epsilon = \sqrt{2(1-\cos(\widehat{\theta}_0))} = \operatorname{dist}(\boldsymbol{b}_0, \{\pm\boldsymbol{n}_1\}) \tag{5.37}$$

where the last equality follows from (5.32). To satisfy the RRC in (5.27), we require

$$\epsilon < \sqrt{2(1-1/\zeta_1)}.$$

Combining this inequality with (5.37) gives the requirement on the initialization, i.e.,

$$\widehat{\theta}_0 < \arccos(1/\zeta_1).$$

In other words, by choosing the initialization $\boldsymbol{b}_0$ satisfying $\widehat{\theta}_0 < \arccos(1/\zeta_1)$, and $\epsilon = \sqrt{2(1-\cos(\widehat{\theta}_0))}$, $\alpha = \frac{\sqrt{2}}{2} N_1 c_{\boldsymbol{\mathcal{X}}_1,\min}\Big(\cos(\widehat{\theta}_0) - 1/\zeta_1\Big)$, from Lemma 19 the $(\alpha, \epsilon, \{\pm\boldsymbol{n}_1\})$-RRC in (5.27) is satisfied. Moreover, for the Riemannian subgradient $\mathcal{G}(\boldsymbol{b})$ used in Algorithm 3, we have

$$\begin{aligned}
&\|\mathcal{G}(\boldsymbol{b})\|_2 = \left\|(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\widetilde{\boldsymbol{\mathcal{X}}}\operatorname{sign}(\widetilde{\boldsymbol{\mathcal{X}}}^\top\boldsymbol{b})\right\|_2 \\
&= \left\|(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\sum_{k=1}^{K}\sum_{j=1}^{N_k}\operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top\boldsymbol{b})\boldsymbol{x}_j^{(k)} + (\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{M}\operatorname{sign}(\boldsymbol{o}_j^\top\boldsymbol{b})\boldsymbol{o}_j\right\|_2 \\
&\le \sum_{k=1}^{K}\left\|(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{N_k}\operatorname{sign}({\boldsymbol{x}_j^{(k)}}^\top\boldsymbol{b})\boldsymbol{x}_j^{(k)}\right\|_2 + \left\|(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\sum_{j=1}^{M}\operatorname{sign}(\boldsymbol{o}_j^\top\boldsymbol{b})\boldsymbol{o}_j\right\|_2 \\
&\le \sum_{k=1}^{K}\cos(\theta_k)N_k c_{\boldsymbol{\mathcal{X}}_k,\max} + \sum_{k=1}^{K}N_k\eta_{\boldsymbol{\mathcal{X}}_k} + M\eta_{\boldsymbol{\mathcal{O}}} \\
&\le \sqrt{\mathbf{1}^\top\boldsymbol{W}^{\max}\mathbf{1}} + \sum_{k=1}^{K}N_k\eta_{\boldsymbol{\mathcal{X}}_k} + M\eta_{\boldsymbol{\mathcal{O}}}
\end{aligned}$$

where the second inequality follows from (5.7), and the last inequality from (5.14). We now apply Theorem 9 by specifying ξ as in (5.36) to obtain that $\{\boldsymbol{b}_t\}$ satisfies

$$\operatorname{dist}(\boldsymbol{b}_t, \{\pm \boldsymbol{n}_1\}) \leq \operatorname{dist}(\boldsymbol{b}_0, \{\pm \boldsymbol{n}_1\})\beta^t, \forall k \geq 0. \tag{5.38}$$

From Proposition 2, we have $\operatorname{dist}(\boldsymbol{b}_t, \{\pm \boldsymbol{n}_1\}) \geq \sin(\widehat{\theta}_t)$, and thus from (5.38) we obtain

$$\sin(\widehat{\theta}_t) \leq \operatorname{dist}(\boldsymbol{b}_0, \{\pm \boldsymbol{n}_1\})\beta^t = \epsilon \cdot \beta^t$$

where the last equality follows (5.37). □

Discussion of Theorem 14. Theorem 14 ensures that a properly initialized Algorithm 3 converges linearly to a normal vector of the geometrically dominant hyperplane $\mathcal{H}_1$, i.e., $\pm \boldsymbol{n}_1$, provided a certain geometrically diminishing step size is used. Note that Theorem 12 implies that $\pm \boldsymbol{n}_1$ are global solutions to (5.1) when condition (5.17) is satisfied. The initialization requirement coincides with Lemma 18, which states that any critical point inside the cone $\mathcal{C} = \{\boldsymbol{y} \in \mathbb{R}^D : |\boldsymbol{y}^\top \boldsymbol{n}_1| > 1/\zeta_1, \|\boldsymbol{y}\|_2 = 1\}$ must be a normal vector of $\mathcal{H}_1$ (see Figure 5.1). Moreover, as discussed after Theorem 9, the diminishing factor β is crucial to the convergence properties of the PRSGM in Algorithm 3: convergence may fail if β is too small, and convergence may be slow when β is too large, which will be further illustrated in Section 5.5. Finally, a result similar to Proposition 4 that ensures a spectral initialization $\boldsymbol{b}_0 = \arg\min_{\boldsymbol{b} \in \mathbb{S}^{D-1}} \left\| \widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b} \right\|_2^2$ for Algorithm 3 is close enough to $\{\pm \boldsymbol{n}_1\}$ can be stated as follows.

Proposition 8. *The spectral initialization* $\boldsymbol{b}_0$ *computed as the bottom eigenvector of* *the matrix $\widetilde{\boldsymbol{\mathcal{X}}}\widetilde{\boldsymbol{\mathcal{X}}}^\top$ satisfies*

$$\operatorname{dist}(\boldsymbol{b}_0, \{\pm \boldsymbol{n}_1\}) \leq \sqrt{\frac{\sum_{k=2}^K \left(\sigma_1^2(\boldsymbol{\mathcal{X}}_k) - \sigma_{D-1}^2(\boldsymbol{\mathcal{X}}_k)\right) + \sigma_1^2(\boldsymbol{\mathcal{O}}) - \sigma_D^2(\boldsymbol{\mathcal{O}})}{\sigma_{D-1}^2(\boldsymbol{\mathcal{X}}_1)}} \tag{5.39}$$

where $\sigma_\ell(\cdot)$ denotes the ℓ-th largest singular value.

5.4 Hyperplane clustering with DPCP

Recall that K-subspaces (KSS) [1, 10] is a simple iterative framework for subspace clustering that alternates between assigning data points to clusters and fitting a subspace to each cluster. The previous sections concentrated on the theory and algorithms for solving the DPCP problem (5.1) for a UoH, showing it recovers the geometrically dominant hyperplane. Inspired by the fact that condition (5.24) in Theorem 13 is likely to hold in the subspace estimation step of KSS (since we expect most of the points in the estimated cluster to belong to a *single* hyperplane), we use a family of KSS variants for hyperplane clustering. Note that the better performance of the iterative KSS approach over the sequential approach, which fits one hyperplane at a time and removes the points belonging to the previously selected subspace, was observed in [113] where the DPCP problem was solved by IRLS.

Aside from the standard KSS, we also consider the following two improved variants.

Ensemble KSS (EKSS). The performance of KSS is sensitive to its initialization because the problem is non-convex. A practical approach is to repeat the process for multiple random initializations and then pick the best one, or combine the results together in a certain way. Based on the fact that partially-correct clustering information from each random initialization of KSS can be combined to obtain a better clustering result, the Ensemble KSS (EKSS) [68] constructs an affinity matrix whose (i, j)th entry is the number of times the ith and jth points are clustered together, and then applies spectral clustering to obtain the final clustering results.

Cooperative Re-initialization (CoRe) KSS. The Cooperative Re-initialization (CoRe) [58] framework optimizes a group of clustering results (replicas) by greedily swapping clusters between them to improve the overall quality. Both EKSS and CoRe expect the clustering in each replica to be partially correct, and that the same pattern of errors will not be made by all replicas. CoRe is capable of identifying bad clusters in a replica and swapping them with better alternatives by monitoring the change in the objective value, and hence it is observed to be more efficient than EKSS.

Since the above variants of KSS use PCA as the standard way to fit a hyperplane to a cluster, we denote them as PCA-KSS, PCA-EKSS, and PCA-CoRe-KSS. To improve their performance, we replace PCA by our DPCP approach with the PRSGM (Algorithm 3) and denote these KSS variants by DPCP-KSS, DPCP-EKSS, and DPCP-CoRe-KSS. We also use the CoP [88] to fit the hyperplane for each cluster, resulting in the three KSS variants CoP-KSS [42], CoP-EKSS [68], and CoP-CoRe-KSS. Experimental results using both synthetic and real data for all of these algorithmic variants are presented in the next section.

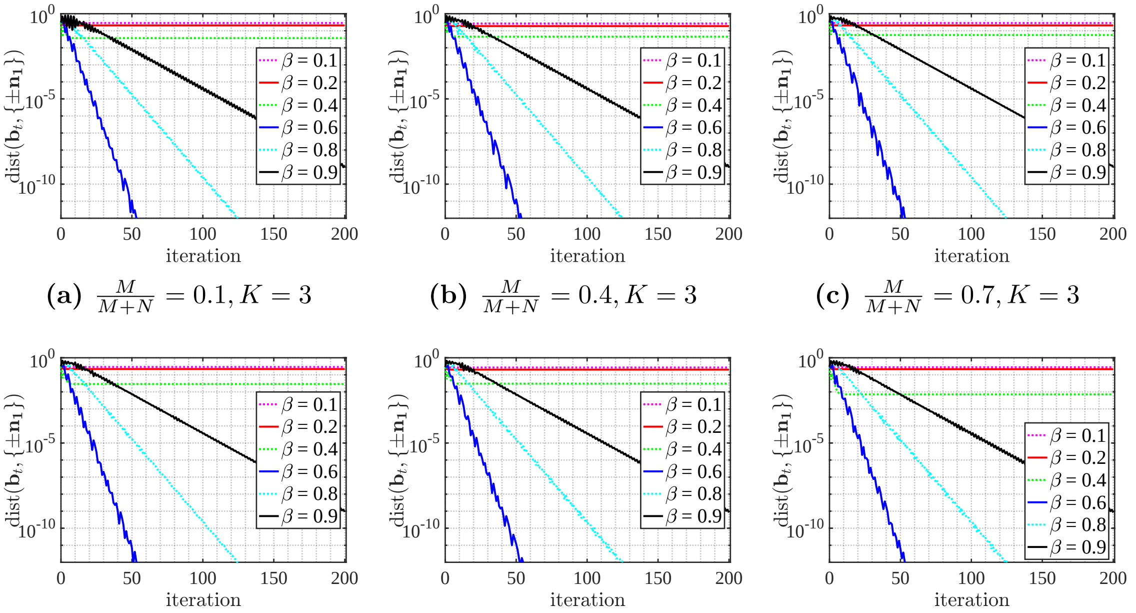
5.5 Experiments

In this section, we evaluate the PRSGM (Algorithm 3) for solving the DPCP problem (5.1) under a UoH model. In Section 5.5.1 we investigate the performance of integrating DPCP into various KSS variants (see Section 5.4) using synthetic data. We further demonstrate its performance and superiority by experimenting on plane clustering using real 3D data in Section 5.5.2.

5.5.1 Synthetic data

Convergence of PRSGM. We first numerically justify the convergence properties of PRSGM (Algorithm 3) for solving the DPCP problem (5.1) under a UoH model. The data are generated based on the random spherical model in Definition 5, where we fix $D = 9, N = 2000$ with $N_{k+1} = 0.5N_k$. Figure 5.2 shows the convergence of PRSGM for various values of the geometric diminishing factor β, different outlier ratios $\frac{M}{M+N}$, and different numbers of underlying hyperplanes K, where we use the spectral initialization and set the initial step size μ_0 to 0.01. One can observe linear convergence of PRSGM to $\mathcal{H}_1$ for all the cases if the diminishing factor β is tuned properly. In particular, it verifies the role of β in Theorem 14, which is similar to that of Theorem 9, namely that it controls the convergence speed. When β is too small, e.g., $\beta \in \{0.1, 0.2, 0.4\}$, convergence may fail, which agrees with (4.18) and (4.19). However, when $\beta \in \{0.6, 0.8, 0.9\}$, the algorithm converges linearly, with larger values of β resulting in slower convergence speeds.

Hyperplane clustering. Next, we compare the performance of the methods



(a) $\frac{M}{M+N} = 0.1, K = 3$ **(b)** $\frac{M}{M+N} = 0.4, K = 3$ **(c)** $\frac{M}{M+N} = 0.7, K = 3$

(d) $K = 4, \frac{M}{M+N} = 0.7$ **(e)** $K = 5, \frac{M}{M+N} = 0.7$ **(f)** $K = 6, \frac{M}{M+N} = 0.7$

Figure 5.2. Convergence of PRSGM (Algorithm 3) to $\mathcal{H}_1$ for the DPCP problem (5.1) under a UoH model. In the experiments, we fix $D = 9, N = 2000$ with $N_{k+1} = 0.5N_k$. We choose $\boldsymbol{b}_0$ as the bottom eigenvector of $\widetilde{\boldsymbol{\mathcal{X}}}\widetilde{\boldsymbol{\mathcal{X}}}^\top$ and set the initial step size μ_0 to 0.01.

discussed in Section 5.4. Following the setup in [113], we test with ambient dimensions $D = 4, 9$ for the synthetic experiments. And we test with $K = 2, 3, 4, 5$, $N = 50KD$ (each plane has the same number of points so that $N_k = 50D$), and $\frac{M}{M+N} = 0.3$. Since the KSS-style methods (without ensemble) are sensitive to initialization, we run them 10 times with random initializations until convergence (tolerance of 0.001) or 100 iterations is reached, and then select the best (i.e., the one with the lowest objective value). The CoRe methods operate directly on these 10 replicas to return an improved clustering result by aggregating the knowledge. For the EKSS-like methods, in each replica we run the KSS-style methods for only 10 iterations but build the affinity matrix based on 1000 such replicas, which is suggested in [68]. For the KSS variants involve DPCP, we make the conservative choice of fixing $\beta = 0.9$ in PRSGM, which

Table 5.2. Mean hyperplane clustering accuracy (runtime in seconds) over 50 independent experiments when $D = 4$.

	$D = 4$			
	$K = 2$	$K = 3$	$K = 4$	$K = 5$
MKF	0.7937 (0.13)	0.6263 (0.19)	0.5548 (0.23)	0.4643 (0.29)
SCC	0.9445 (0.21)	0.9209 (0.46)	0.9093 (0.77)	0.8784 (1.48)
EnSC	0.7011 (0.14)	0.4912 (0.23)	0.3913 (0.30)	0.3254 (0.41)
SSC-ADMM	0.6801 (0.86)	0.4810 (2.30)	0.3795 (4.32)	0.3175 (9.91)
SSC-OMP	0.5707 (0.07)	0.4134 (0.09)	0.3291 (0.12)	0.2747 (0.38)
DPCP-KSS	**0.9834** (0.11)	**0.9463** (0.46)	0.8985 (0.77)	0.8103 (1.05)
CoP-KSS	0.9614 (0.11)	0.8747 (0.42)	0.8300 (0.81)	0.7630 (1.24)
PCA-KSS	0.9601 (0.01)	0.8623 (0.05)	0.8142 (0.12)	0.7461 (0.19)
DPCP-EKSS	**0.9889** (5.85)	0.8807 (8.19)	**0.9778** (9.45)	**0.9489** (12.67)
CoP-EKSS	0.8278 (10.90)	0.8393 (16.69)	0.8772 (20.46)	0.7938 (29.40)
PCA-EKSS	0.8278 (4.10)	0.8274 (6.03)	0.8517 (7.46)	0.7542 (10.58)
DPCP-CoRe-KSS	0.9832 (0.20)	**0.9715** (0.55)	**0.9561** (1.23)	**0.9599** (1.93)
CoP-CoRe-KSS	0.9612 (0.10)	0.8992 (0.48)	0.9065 (0.96)	0.8907 (1.74)
PCA-CoRe-KSS	0.9603 (0.02)	0.8981 (0.12)	0.8769 (0.32)	0.8586 (0.80)

empirically works well but additional tuning for β is still possible. Besides those KSS variants, we also test the performance of other state-of-the-art subspace clustering algorithms that include MKF [146], SCC [17], SSC-ADMM [35], EnSC [141], and SSC-OMP [143]. For MKF, we set the step size for gradient boosting to be 0.001, the maximal allowed iterations to be 10000; and for SCC, we use the linear spectral curvature clustering implementation; for EnSC, we set $\lambda = 0.95$ and $\alpha = 200$; for SSC-ADMM, we set $\rho = 1$ and $\alpha = 20$; for SSC-OMP, we set $k_{\max} = 5$ and $\epsilon = 10^{-8}$. Table 5.2 and Table 5.3 report the mean clustering accuracy (runtime in seconds) of the methods on 50 independent instances with the highest two clustering accuracies in each column given in bold when $D = 4$ and $D = 9$, respectively.

One can see that the SC methods EnSC, SSC-ADMM, and SSC-OMP, which are designed primarily for the low relative dimension setting, are among the least competitive for clustering *hyperplanes*. Notably, SSC-ADMM is significantly slower than other competitors, especially when the ambient dimension and the number of underlying hyperplanes become large. Also, MKF and SCC do not perform well. Among the other methods, we observe that within each scheme, algorithms that involve DPCP (implemented by PRSGM in Algorithm 3) almost always perform the best. As a result, in each column the best method is the one that uses DPCP as the internal solver for identifying the dominant hyperplane in a cluster. We find that with as little as 10 replicas, the methods built on the CoRe framework perform very well. We believe this result is because CoRe is more aggressive in dealing with bad clusters, i.e., swapping them with other estimates, while for EKSS even bad clusters still have a good chance of influencing the final clustering results. Finally, the EKSS-like methods take significantly more time compared with other variants because its success relies on a large number replicas to build an affinity matrix of high quality.

Table 5.3. Mean hyperplane clustering accuracy (runtime in sec) over 50 independent experiments when $D = 9$.

	$D = 9$			
	$K = 2$	$K = 3$	$K = 4$	$K = 5$
MKF	0.5840 (0.19)	0.3973 (0.22)	0.2949 (0.22)	0.2470 (0.29)
SCC	0.9126 (1.46)	0.5940 (3.53)	0.3138 (5.83)	0.2519 (12.33)
EnSC	0.6223 (0.49)	0.3996 (1.63)	0.3125 (2.27)	0.2540 (2.76)
SSC-ADMM	0.6683 (10.09)	0.4010 (46.49)	0.2999 (112.77)	0.2548 (296.68)
SSC-OMP	0.5267 (0.13)	0.3573 (0.54)	0.2732 (0.71)	0.2232 (0.93)
DPCP-KSS	0.9927 (0.66)	**0.9807** (1.24)	0.8051 (1.29)	0.5004 (1.80)
CoP-KSS	0.9706 (0.75)	0.9358 (2.78)	0.8380 (5.28)	0.5110 (8.35)
PCA-KSS	0.9619 (0.05)	0.9243 (0.22)	0.8074 (0.51)	0.5130 (0.93)
DPCP-EKSS	**0.9938** (11.53)	0.9517 (16.15)	0.4908 (33.10)	0.2920 (44.41)
CoP-EKSS	0.8271 (43.96)	0.7900 (60.34)	0.3706 (107.24)	0.2867 (133.72)
PCA-EKSS	0.8221 (7.48)	0.7539 (14.00)	0.3660 (28.56)	0.2868 (39.68)
DPCP-CoRe-KSS	**0.9928** (0.96)	**0.9857** (3.98)	**0.9784** (7.83)	**0.9628** (11.12)
CoP-CoRe-KSS	0.9706 (0.78)	0.9415 (2.89)	0.9258 (5.64)	0.9089 (10.22)
PCA-CoRe-KSS	0.9619 (0.07)	0.9370 (0.38)	**0.9278** (1.23)	**0.9083** (4.11)

5.5.2 Plane clustering using real 3D data

We explore the performance of DPCP in hyperplane clustering using the real dataset NYUdepthV2 [80], as introduced in Section 1.1.2.2, which contains indoor RGB images of size 480 × 640 × 3 together with depth information for each pixel. We use the experimental setup of [113], where the hyperplane annotation is done manually on 92 indoor RGBd images taken by Microsoft Kinect, but only the 89 of them that contain more than one hyperplane are preserved. Thus, each RGBd image consists of 480 × 640 depth values and K_i planes with $K_i > 1, i \in \{1, 2, ..., 89\}$. After the camera calibration, 307,200 3D points are obtained from each image, which has dominant hyperplanes such as floors, walls and so on. Ground-truth labels indicate that each point either belongs to plane of index from $\{1, 2, \cdots, K_i\}$, or is an outlier (index 0).

Table 5.4. Mean clustering error (running time in seconds) for KSS variants with different "backbones" on 89 annotated images of NYUdepthV2.

	KSS	CoRe-KSS	EKSS
DPCP	**10.2%** (0.09)	**9.3%** (1.11)	**8.0%** (15.81)
PCA	12.4% (0.04)	11.7% (1.17)	10.8% (12.72)
CoP	11.0% (0.04)	10.8% (0.83)	13.8% (18.12)

Pre-processing. For computational reasons, we perform superpixel representation where each image is segmented into about 1000 superpixels and the set of pixels corresponding to each superpixel is substituted by their median depth. Since the planes associated with an indoor scene are affine in $\mathbb{R}^3$, we use homogeneous coordinates by appending 1 at the fourth coordinate and normalize it to unit length in $\mathbb{R}^4$. Finally, since different superpixels represent different numbers of underlying pixels, we adopt the practice in [113] that each homogenized superpixel is further weighted according to its size so that points representing larger numbers of superpixels have more influence.

Evaluation. Given the estimated clusters for the superpixels, we assign the original pixels to the same cluster as their representatives. Note that none of the algorithms considered are explicitly configured to detect outliers, instead they assign every point to some plane. We only evaluate the clustering error of the inlier subset in the estimated clusters as was the practice in [113]. The clustering error is defined to be the sum of mismatches from each cluster divided by the total number of inliers. Since the labels of the ground-truth clusters could be mismatched with the estimations, we report the minimum clustering error after performing a linear assignment.

Results. We now compare the KSS variants with different "backbones" as introduced in Section 5.4, namely PCA, CoP and DPCP, in clustering hyperplanes on the 89 annotated images of NYUdepthV2. The parametric setting for each method is the same as for the synthetic experiments. Note that here we exclude the other general subspace clustering algorithms discussed in the synthetic experiments since they have been shown to be less competitive for the hyperplane clustering task (see Section 5.5.1). We first show in Table 5.4 the average clustering error for the KSS variants applied to the real data. One can see that a similar phenomenon appears as in the synthetic experiments, namely that the algorithm achieving the lowest mean clustering error is one using DPCP as the internal subproblem (solved by PRSGM in Algorithm 3) for estimating the dominant hyperplane within each KSS framework. On the other hand, although the KSS method runs very fast, it is generally not comparable with CoRe-KSS or EKSS in this test. Note that EKSS takes significantly longer time for its construction of the affinity matrix, which is based on 1000 KSS replicas. Finally, in Figures 5.3, 5.4 and 5.5, we give visual comparisons of various approaches on clustering hyperplanes from 3D point clouds of image 55, image 5 and image 60 in NYUdepthV2, respectively.

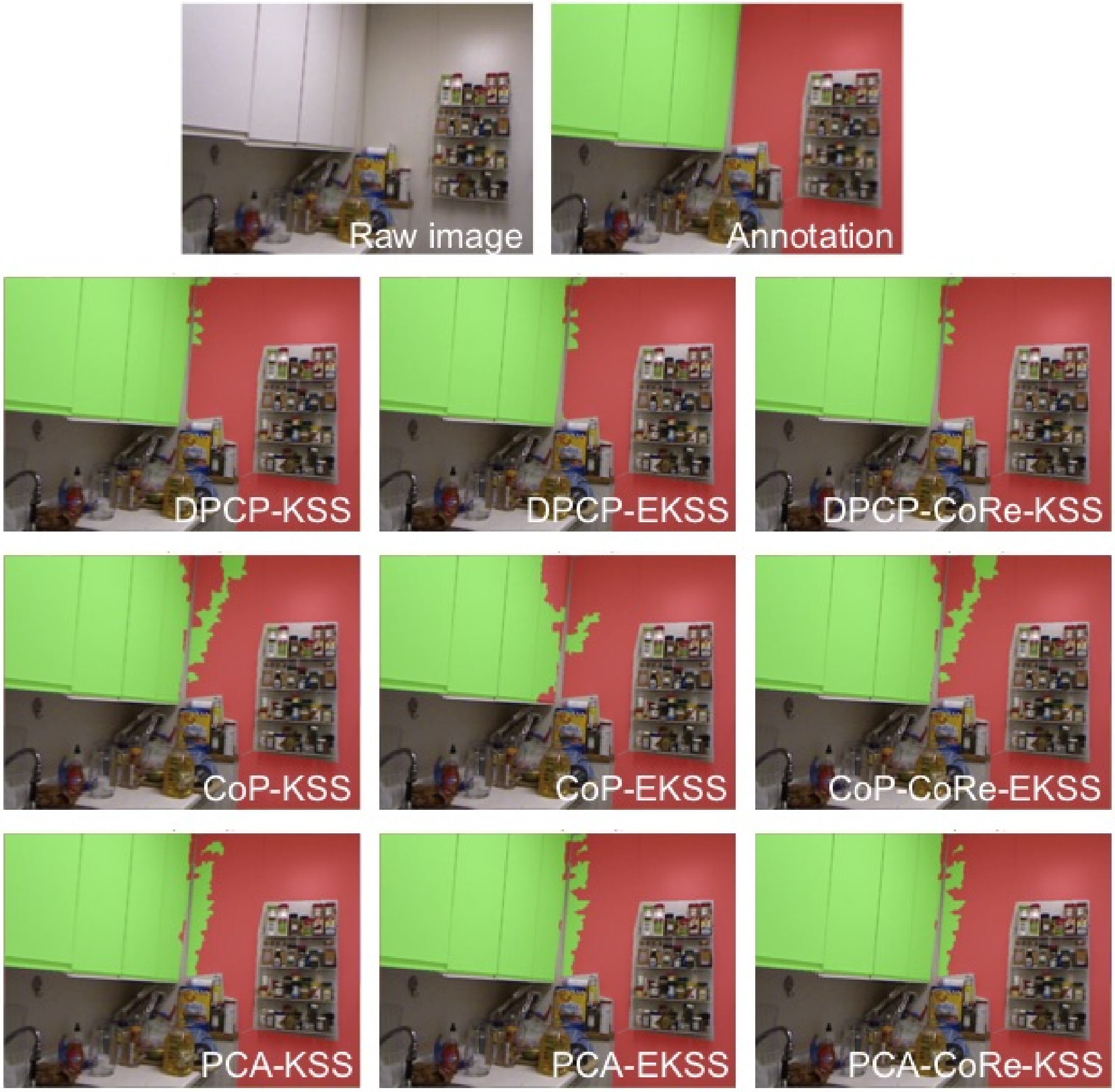


Figure 5.3. Visualization of various approaches in clustering two hyperplanes from a 3D point cloud of image 55 in NYUdepthV2.

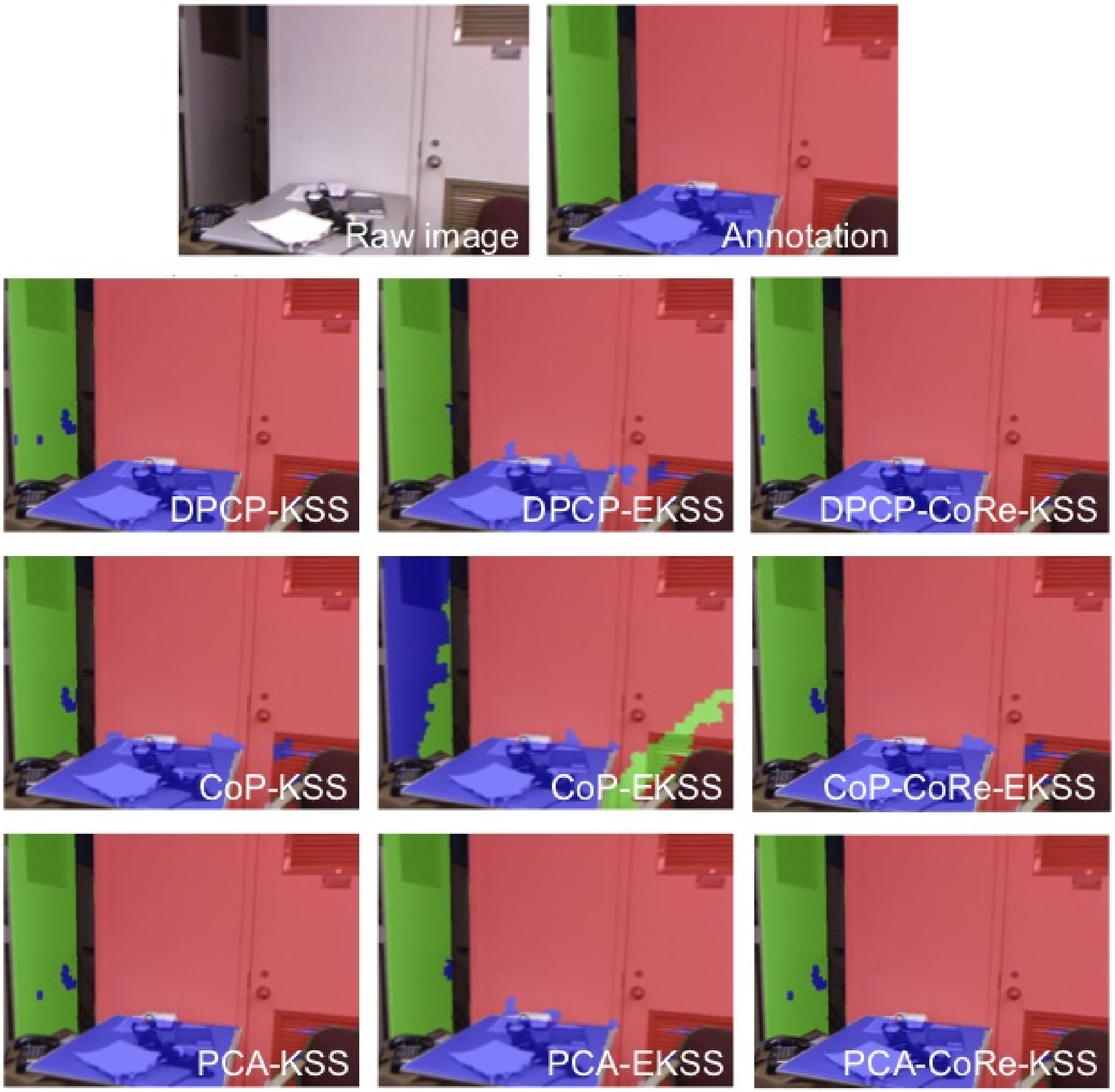


Figure 5.4. Visualization of various approaches in clustering three hyperplanes from a 3D point cloud of image 5 in NYUdepthV2.

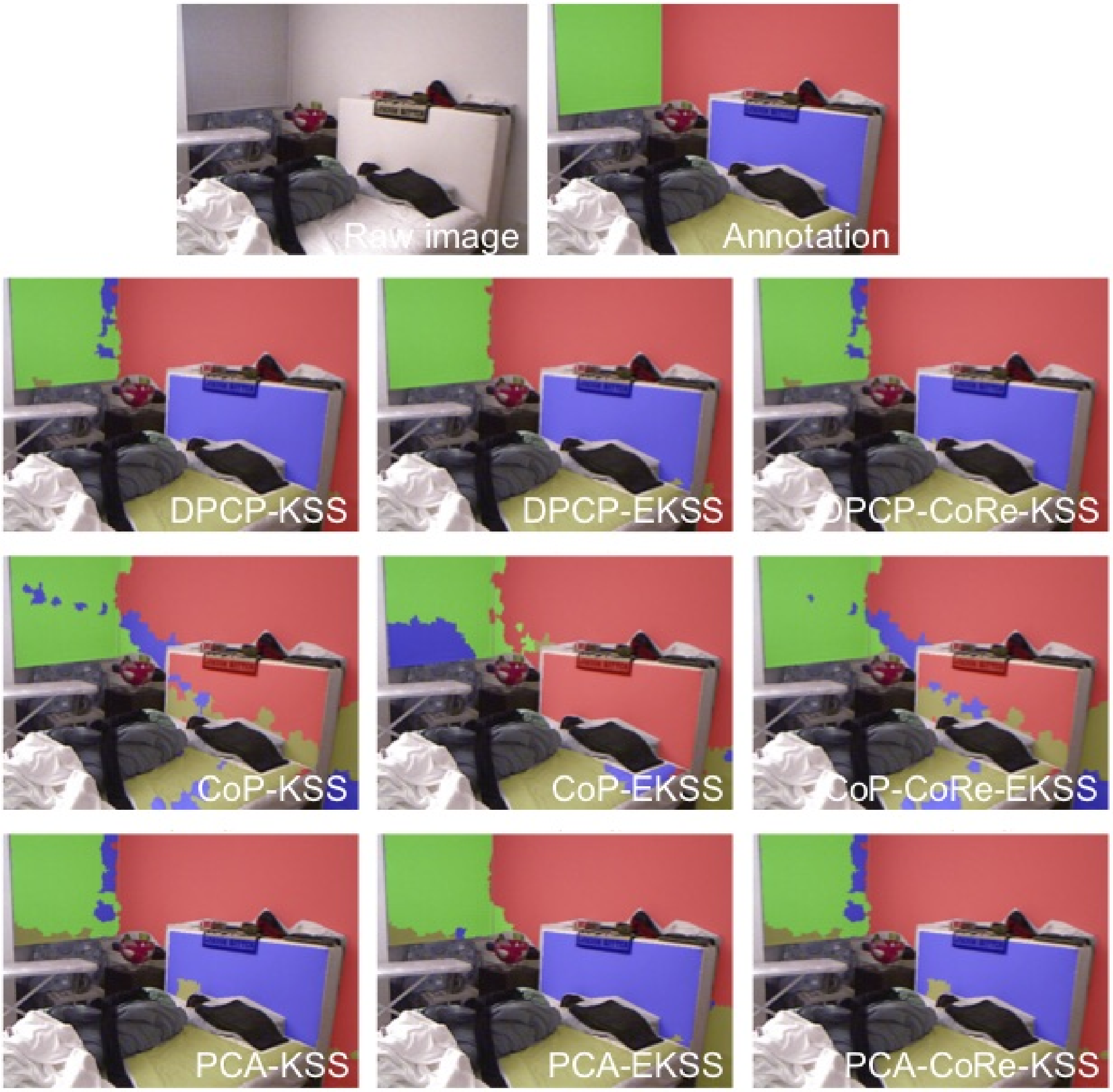


Figure 5.5. Visualization of various approaches in clustering four hyperplanes from a 3D point cloud of image 60 in NYUdepthV2.

Chapter 6

Conclusions

This thesis developed extensive theory and algorithms for subspace learning for data arising from subspaces of high relative dimension.

In Chapter 3 we extended the global optimality analysis of the Dual Principal Component Pursuit (DPCP) method from learning a hyperplane with noiseless data to a subspace of any dimension in the high relative dimension regime with noisy data. We established a geometric analysis that revealed that the subspace angle between the global solution to the non-convex DPCP problem and the orthogonal complement of the subspace is upper bounded by an amount that is proportional to the noise level. We also derived a probabilistic analysis that shows that the DPCP problem for learning a subspace of high relative dimension can handle $O((\#\text{inliers})^2)$ outliers even in the noisy setting, which is superior to other existing robust subspace recovery methods that can tolerate at best $O(N)$ outliers in theory.

In Chapter 4 we presented a Projected Riemannian Sub-Gradient Method (PRSGM) and showed that with proper initialization and step size, it converges linearly to some points at which the objective function satisfies a certain Riemannian regularity condition (RRC). We then applied PRSGM to the DPCP problem for learning a single subspace and proved that it converges linearly to a neighborhood of the orthogonal complement subspace, whose region is proportional to the noise level. Experiments on synthetic data and 3D roadplane detection demonstrated the effectiveness of using PRSGM as the subproblem solver for DPCP in robust single subspace learning.

In Chapter 5 we improved the existing global optimality theory of DPCP for a union of hyperplanes (UoH) by deriving a more transparent geometric analysis and a new probabilistic analysis. Our analysis shows that under certain conditions any global solution to DPCP for a UoH is a normal vector to a geometrically dominant hyperplane. Also, we proved a convergence result for PRSGM when used for DPCP under a UoH model. Finally, by integrating DPCP into KSS (DPCP-KSS) and utilizing ensembles of DPCP-KSS, experiments on synthetic data and 3D plane clustering showed that we achieve state-of-the-art performance in hyperplane clustering.

There are many interesting directions for future work. Theoretically, one can extend the analysis of DPCP for a UoH to a union of high dimensional subspaces. Algorithmically, one can design an efficient intrinsic optimization method for solving DPCP wherein the iterates move along geodesic directions, which contrasts the extrinsic PRSGM. Finally, one can explore additional applications of DPCP such as clustering deep features extracted from images of a single object category in ImageNet.

Bibliography

[1] Pankaj K Agarwal and Nabil H Mustafa. “K-means projective clustering.” In: *Proceedings of the twenty-third ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. 2004, pp. 155–165.

[2] Alex M Andrew. “Multiple view geometry in computer vision.” In: *Kybernetes* (2001).

[3] Yu Bai, Qijia Jiang, and Ju Sun. “Subgradient descent learns orthogonal dictionaries.” In: *7th International Conference on Learning Representations, ICLR 2019*. 2019.

[4] Laurent Bako. “Identification of switched linear systems via sparse optimization.” In: *Automatica* 47.4 (2011), pp. 668–677.

[5] Laura Balzano, Robert Nowak, and Benjamin Recht. “Online identification and tracking of subspaces from highly incomplete information.” In: *2010 48th Annual allerton conference on communication, control, and computing (Allerton)*. IEEE. 2010, pp. 704–711.

[6] Daniel Barath, Jiri Matas, and Jana Noskova. “MAGSAC: marginalizing sample consensus.” In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019, pp. 10197–10205.

[7] Ronen Basri and David W Jacobs. “Lambertian reflectance and linear subspaces.” In: *IEEE transactions on pattern analysis and machine intelligence* 25.2 (2003), pp. 218–233.

[8] Nicolas Boumal, Pierre-Antoine Absil, and Coralia Cartis. “Global rates of convergence for nonconvex optimization on manifolds.” In: *IMA Journal of Numerical Analysis* 39.1 (2019), pp. 1–33.

[9] Thierry Bouwmans and El Hadi Zahzah. “Robust PCA via principal component pursuit: A review for a comparative evaluation in video surveillance.” In: *Computer Vision and Image Understanding* 122 (2014), pp. 22–34.

[10] Paul S Bradley and Olvi L Mangasarian. “K-plane clustering.” In: *Journal of Global Optimization* 16.1 (2000), pp. 23–32.

[11] J Paul Brooks, José H Dulá, and Edward L Boone. “A pure L1-norm principal component analysis.” In: *Computational statistics & data analysis* 61 (2013), pp. 83–98.

[12] James V Burke and Michael C Ferris. “Weak sharp minima in mathematical programming.” In: *SIAM Journal on Control and Optimization* 31.5 (1993), pp. 1340–1359.

[13] Emmanuel J Candes, Xiaodong Li, Yi Ma, and John Wright. “Robust principal component analysis?” In: *Journal of the ACM (JACM)* 58.3 (2011), pp. 1–37.

[14] Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi. “Phase retrieval via Wirtinger flow: Theory and algorithms.” In: *IEEE Transactions on Information Theory* 61.4 (2015), pp. 1985–2007.

[15] Emmanuel J Candes, Michael B Wakin, and Stephen P Boyd. "Enhancing sparsity by reweighted ℓ_1 minimization." In: *Journal of Fourier analysis and applications* 14.5-6 (2008), pp. 877–905.

[16] Rick Chartrand and Wotao Yin. "Iteratively reweighted algorithms for compressive sensing." In: *2008 IEEE international conference on acoustics, speech and signal processing*. IEEE. 2008, pp. 3869–3872.

[17] Guangliang Chen and Gilad Lerman. "Spectral curvature clustering (SCC)." In: *International Journal of Computer Vision* 81.3 (2009), pp. 317–330.

[18] Shixiang Chen, Shiqian Ma, Anthony Man-Cho So, and Tong Zhang. "Proximal gradient method for manifold optimization." In: *arXiv preprint arXiv:1811.00980* 5.6 (2018), p. 8.

[19] Xinlei Chen, Abhinav Shrivastava, and Abhinav Gupta. "Neil: Extracting visual knowledge from web data." In: *Proceedings of the IEEE international conference on computer vision*. 2013, pp. 1409–1416.

[20] Ondrej Chum, Jiri Matas, and Josef Kittler. "Locally optimized RANSAC." In: *Joint Pattern Recognition Symposium*. Springer. 2003, pp. 236–243.

[21] Joao Paulo Costeira and Takeo Kanade. "A multibody factorization method for independently moving objects." In: *International Journal of Computer Vision* 29.3 (1998), pp. 159–179.

[22] Frank E Curtis, Tim Mitchell, and Michael L Overton. "A BFGS-SQP method for nonsmooth, nonconvex, constrained optimization and its evaluation using relative minimization profiles." In: *Optimization Methods and Software* 32.1 (2017), pp. 148–181.

[23] Frank E Curtis and Xiaocun Que. “A quasi-Newton algorithm for nonconvex, nonsmooth optimization with global convergence guarantees.” In: *Mathematical Programming Computation* 7.4 (2015), pp. 399–428.

[24] Ingrid Daubechies, Ronald DeVore, Massimo Fornasier, and C Sinan Güntürk. “Iteratively reweighted least squares minimization for sparse recovery.” In: *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences* 63.1 (2010), pp. 1–38.

[25] Damek Davis, Dmitriy Drusvyatskiy, Kellie J MacPhee, and Courtney Paquette. “Subgradient methods for sharp weakly convex functions.” In: *Journal of Optimization Theory and Applications* 179.3 (2018), pp. 962–982.

[26] Damek Davis, Dmitriy Drusvyatskiy, and Courtney Paquette. “The nonsmooth landscape of phase retrieval.” In: *IMA Journal of Numerical Analysis* 40.4 (2020), pp. 2652–2695.

[27] Allison Del Giorno, J Andrew Bagnell, and Martial Hebert. “A discriminative framework for anomaly detection in large videos.” In: *European Conference on Computer Vision*. Springer. 2016, pp. 334–349.

[28] Chris Ding, Ding Zhou, Xiaofeng He, and Hongyuan Zha. “R 1-pca: rotational invariant l 1-norm principal component analysis for robust subspace factorization.” In: *Proceedings of the 23rd international conference on Machine learning*. 2006, pp. 281–288.

[29] Tianjiao Ding, Yunchen Yang, Zhihui Zhu, Daniel P Robinson, René Vidal, Laurent Kneip, and Manolis C Tsakiris. “Robust Homography Estimation via Dual Principal Component Pursuit.” In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 6080–6089.

[30] Tianyu Ding, Zhihui Zhu, Tianjiao Ding, Yunchen Yang, Rene Vidal, Manolis Tsakiris, and Daniel Robinson. "Noisy dual principal component pursuit." In: *Proceedings of the International Conference on Machine learning.* 2019, pp. 1617–1625.

[31] Tianyu Ding, Zhihui Zhu, Manolis Tsakiris, Rene Vidal, and Daniel Robinson. "Dual Principal Component Pursuit for Learning a Union of Hyperplanes: Theory and Algorithms." In: *International Conference on Artificial Intelligence and Statistics*. PMLR. 2021, pp. 2944–2952.

[32] Tianyu Ding, Zhihui Zhu, Rene Vidal, and Daniel Robinson. "Dual Principal Component Pursuit for Robust Subspace Learning: Theory and Algorithms for a Holistic Approach." In: *Proceedings of the International Conference on Machine learning*. 2021.

[33] John C Duchi and Feng Ruan. "Solving (most) of a set of quadratic equalities: Composite optimization for robust phase retrieval." In: *Information and Inference: A Journal of the IMA* 8.3 (2019), pp. 471–529.

[34] Alan Edelman, Tomás A Arias, and Steven T Smith. "The geometry of algorithms with orthogonality constraints." In: *SIAM journal on Matrix Analysis and Applications* 20.2 (1998), pp. 303–353.

[35] Ehsan Elhamifar and Rene Vidal. "Sparse subspace clustering: Algorithm, theory, and applications." In: *IEEE transactions on pattern analysis and machine intelligence* 35.11 (2013), pp. 2765–2781.

[36] Ehsan Elhamifar and René Vidal. "Clustering disjoint subspaces via sparse representation." In: *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE. 2010, pp. 1926–1929.

[37] Ehsan Elhamifar and René Vidal. “Sparse subspace clustering.” In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE. 2009, pp. 2790–2797.

[38] Paolo Favaro, René Vidal, and Avinash Ravichandran. “A closed form solution to robust subspace estimation and clustering.” In: *CVPR 2011*. IEEE. 2011, pp. 1801–1807.

[39] Martin A Fischler and Robert C Bolles. “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography.” In: *Communications of the ACM* 24.6 (1981), pp. 381–395.

[40] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. “Vision meets robotics: The kitti dataset.” In: *The International Journal of Robotics Research* 32.11 (2013), pp. 1231–1237.

[41] Pando Georgiev, Fabian Theis, and Andrzej Cichocki. “Sparse component analysis and blind source separation of underdetermined mixtures.” In: *IEEE transactions on neural networks* 16.4 (2005), pp. 992–996.

[42] Andrew Gitlin, Biaoshuai Tao, Laura Balzano, and John Lipor. “Improving K-Subspaces via Coherence Pursuit.” In: *IEEE Journal of Selected Topics in Signal Processing* 12.6 (2018), pp. 1575–1588.

[43] Jean-Louis Goffin. *Subgradient optimization in nonsmooth optimization (including the Soviet revolution)*. Groupe d’études et de recherche en analyse des décisions, 2012.

[44] Alvina Goh and René Vidal. “Segmenting motions of different types by unsupervised manifold clustering.” In: *2007 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE. 2007, pp. 1–6.

[45] Gene H Golub and Charles F Van Loan. *Matrix computations*. Vol. 3. JHU press, 2013.

[46] Venu Madhav Govindu. “A tensor decomposition for geometric grouping and segmentation.” In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*. Vol. 1. IEEE. 2005, pp. 1150–1157.

[47] Philipp Grohs and Seyedehsomayeh Hosseini. “Nonsmooth trust region algorithms for locally Lipschitz functions on Riemannian manifolds.” In: *IMA Journal of Numerical Analysis* 36.3 (2016), pp. 1167–1192.

[48] Amit Gruber and Yair Weiss. “Multibody factorization with uncertainty and missing data using the EM algorithm.” In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.* Vol. 1. IEEE. 2004, pp. I–I.

[49] Jihun Hamm and Daniel D Lee. “Grassmann discriminant analysis: a unifying view on subspace-based learning.” In: *Proceedings of the 25th international conference on Machine learning*. 2008, pp. 376–383.

[50] Zhaoshui He and Andrzej Cichocki. “An efficient K-hyperplane clustering algorithm and its application to sparse component analysis.” In: *International Symposium on Neural Networks*. Springer. 2007, pp. 1032–1041.

[51] Nick Higham and Pythagoras Papadimitriou. “Matrix procrustes problems.” In: *Rapport technique, University of Manchester* (1995).

[52] Harold Hotelling. “Analysis of a complex of statistical variables into principal components.” In: *Journal of educational psychology* 24.6 (1933), p. 417.

[53] Xudong Jiang. “Linear subspace learning-based dimensionality reduction.” In: *IEEE Signal Processing Magazine* 28.2 (2011), pp. 16–26.

[54] Bangti Jin, Dirk A Lorenz, and Stefan Schiffler. “Elastic-net regularization: error estimates and active set methods.” In: *Inverse Problems* 25.11 (2009), p. 115022.

[55] Ian T Jolliffe. “Principal components in regression analysis.” In: *Principal component analysis*. Springer, 1986, pp. 129–155.

[56] Sham Kakade. *Symmetrization and rademacher averages*. Lecture Notes on Statistical Learning Theory, (Lecture 11), 2011.

[57] Andrew V Knyazev and Peizhen Zhu. “Principal angles between subspaces and their tangents.” In: *arXiv preprint arXiv:1209.0523* (2012).

[58] Connor Lane, Benjamin Haeffele, and René Vidal. “Adaptive online k-subspaces with cooperative re-initialization.” In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2019, pp. –.

[59] Michel Ledoux and Michel Talagrand. *Probability in Banach Spaces: isoperimetry and processes*. Springer Science & Business Media, 2013.

[60] Jason D Lee, Ioannis Panageas, Georgios Piliouras, Max Simchowitz, Michael I Jordan, and Benjamin Recht. “First-order methods almost always avoid saddle points.” In: *arXiv preprint arXiv:1710.07406* (2017).

[61] G. Lerman, M. B. McCoy, J. A. Tropp, and T. Zhang. “Robust computation of linear models by convex relaxation.” In: *Foundations of Computational Mathematics* 15.2 (2015), pp. 363–410.

[62] Gilad Lerman and Tyler Maunu. “An overview of robust subspace recovery.” In: *Proceedings of the IEEE* 106.8 (2018), pp. 1380–1410.

[63] Gilad Lerman and Tyler Maunu. "Fast, robust and non-convex subspace recovery." In: *Information and Inference: A Journal of the IMA* 7.2 (2018), pp. 277–336.

[64] Gilad Lerman and Teng Zhang. "l_p-Recovery of the Most Significant Subspace Among Multiple Subspaces with Outliers." In: *Constructive Approximation* 40.3 (2014), pp. 329–385.

[65] Xiao Li, Zhihui Zhu, Anthony Man-Cho So, and Rene Vidal. "Nonconvex robust low-rank matrix recovery." In: *SIAM Journal on Optimization* 30.1 (2020), pp. 660–686.

[66] David Liebowitz and Andrew Zisserman. "Metric rectification for perspective images of planes." In: *Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No. 98CB36231)*. IEEE. 1998, pp. 482–488.

[67] John Lipor, David Hong, Yan Shuo Tan, and Laura Balzano. "Subspace clustering using ensembles of k-subspaces." In: *Information and Inference: A Journal of the IMA* 10.1 (2021), pp. 73–107.

[68] John Lipor, David Hong, Dejiao Zhang, and Laura Balzano. "Subspace Clustering using Ensembles of K-Subspaces." In: *ArXiv* abs/1709.04744v2 (2018).

[69] Guangcan Liu, Zhouchen Lin, Shuicheng Yan, Ju Sun, Yong Yu, and Yi Ma. "Robust recovery of subspace structures by low-rank representation." In: *IEEE transactions on pattern analysis and machine intelligence* 35.1 (2012), pp. 171–184.

[70] Guangcan Liu, Zhouchen Lin, and Yong Yu. “Robust subspace segmentation by low-rank representation.” In: *Proceedings of the 27th international conference on machine learning (ICML-10)*. 2010, pp. 663–670.

[71] Wei Liu, Gang Hua, and John R Smith. “Unsupervised one-class learning for automatic outlier removal.” In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014, pp. 3826–3833.

[72] Can-Yi Lu, Hai Min, Zhong-Qiu Zhao, Lin Zhu, De-Shuang Huang, and Shuicheng Yan. “Robust and efficient subspace segmentation via least squares regression.” In: *European conference on computer vision*. Springer. 2012, pp. 347–360.

[73] Zhi-Quan Luo and Paul Tseng. “Error bounds and convergence analysis of feasible descent methods: a general approach.” In: *Annals of Operations Research* 46.1 (1993), pp. 157–178.

[74] Yi Ma, Allen Y Yang, Harm Derksen, and Robert Fossum. “Estimation of subspace arrangements with applications in modeling and segmenting mixed data.” In: *SIAM review* 50.3 (2008), pp. 413–458.

[75] Panos P Markopoulos, Sandipan Kundu, Shubham Chamadia, Nicholas Tsagkarakis, and Dimitris A Pados. “Outlier-resistant data processing with L1-norm principal component analysis.” In: *Advances in Principal Component Analysis*. Springer, 2018, pp. 121–135.

[76] Tyler Maunu, Teng Zhang, and Gilad Lerman. “A well-tempered landscape for non-convex robust subspace recovery.” In: *Journal of Machine Learning Research* 20.37 (2019), pp. 1–59.

[77] Andreas Maurer. “A vector-contraction inequality for rademacher complexities.” In: *International Conference on Algorithmic Learning Theory*. Springer. 2016, pp. 3–17.

[78] Colin McDiarmid. “On the method of bounded differences.” In: *Surveys in combinatorics* 141.1 (1989), pp. 148–188.

[79] Brian McWilliams and Giovanni Montana. “Subspace clustering of high-dimensional data: a predictive approach.” In: *Data Mining and Knowledge Discovery* 28.3 (2014), pp. 736–772.

[80] Pushmeet Kohli Nathan Silberman Derek Hoiem and Rob Fergus. “Indoor Segmentation and Support Inference from RGBD Images.” In: *ECCV*. 2012.

[81] David Nistér. “Preemptive RANSAC for live structure and motion estimation.” In: *Machine Vision and Applications* 16.5 (2005), pp. 321–329.

[82] Necmiye Ozay, Mario Sznaier, Constantino Lagoa, and Octavia Camps. “GPCA with denoising: A moments-based convex approach.” In: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE. 2010, pp. 3209–3216.

[83] Yannis Panagakis and Constantine Kotropoulos. “Elastic net subspace clustering applied to pop/rock music structure analysis.” In: *Pattern Recognition Letters* 38 (2014), pp. 46–53.

[84] Karl Pearson. “LIII. On lines and planes of closest fit to systems of points in space.” In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2.11 (1901), pp. 559–572.

[85] Q. Qu, J. Sun, and J. Wright. “Finding a sparse vector in a subspace: Linear sparsity using alternating directions.” In: *Advances in Neural Information Processing Systems* (2014), pp. 3401–3409.

[86] Qing Qu, Ju Sun, and John Wright. “Finding a sparse vector in a subspace: Linear sparsity using alternating directions.” In: *IEEE Transactions on Information Theory* 62.10 (2016), pp. 5855–5880.

[87] Rahul Raguram, Ondrej Chum, Marc Pollefeys, Jiri Matas, and Jan-Michael Frahm. “USAC: a universal framework for random sample consensus.” In: *IEEE transactions on pattern analysis and machine intelligence* 35.8 (2012), pp. 2022–2038.

[88] Mostafa Rahmani and George K Atia. “Coherence pursuit: Fast, simple, and robust principal component analysis.” In: *IEEE Transactions on Signal Processing* 65.23 (2017), pp. 6260–6275.

[89] Shankar R Rao, Roberto Tron, René Vidal, and Yi Ma. “Motion segmentation via robust subspace separation in the presence of outlying, incomplete, or corrupted trajectories.” In: *2008 IEEE conference on computer vision and pattern recognition*. IEEE. 2008, pp. 1–8.

[90] Aparajithan Sampath and Jie Shan. “Segmentation and reconstruction of polyhedral building roofs from aerial lidar point clouds.” In: *IEEE Transactions on geoscience and remote sensing* 48.3 (2009), pp. 1554–1567.

[91] Florian Schroff, Antonio Criminisi, and Andrew Zisserman. “Harvesting image databases from the web.” In: *IEEE transactions on pattern analysis and machine intelligence* 33.4 (2010), pp. 754–766.

[92] Nathan Silberman, Derek Hoiem, Pushmeet Kohli, and Rob Fergus. “Indoor segmentation and support inference from rgbd images.” In: *European conference on computer vision*. Springer. 2012, pp. 746–760.

[93] Rim Slama, Hazem Wannous, Mohamed Daoudi, and Anuj Srivastava. “Accurate 3D action recognition using learning on the Grassmann manifold.” In: *Pattern Recognition* 48.2 (2015), pp. 556–567.

[94] Mahdi Soltanolkotabi, Emmanuel J Candes, et al. “A geometric analysis of subspace clustering with outliers.” In: *The Annals of Statistics* 40.4 (2012), pp. 2195–2238.

[95] H. Spath and G. A. Watson. “On orthogonal linear ℓ_1 approximation.” In: *Numerische Mathematik* 51.5 (1987), pp. 531–543.

[96] D. A. Spielman, H. Wang, and J. Wright. “Exact recovery of sparsely-used dictionaries.” In: *Proceedings of the 23d international joint conference on Artificial Intelligence*. AAAI Press, 2013, pp. 3087–3090.

[97] Michael Steinbach, George Karypis, and Vipin Kumar. “A comparison of document clustering techniques.” In: (2000).

[98] Waqas Sultani, Chen Chen, and Mubarak Shah. “Real-world anomaly detection in surveillance videos.” In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 6479–6488.

[99] J. Sun, Q. Qu, and J. Wright. “Complete dictionary recovery using nonconvex optimization.” In: *International Conference on Machine Learning*. 2015, pp. 2351–2360.

[100] Ju Sun, Qing Qu, and John Wright. “Complete dictionary recovery over the sphere.” In: *2015 International Conference on Sampling Theory and Applications (SampTA)*. IEEE. 2015, pp. 407–410.

[101] Ju Sun, Qing Qu, and John Wright. “Complete dictionary recovery over the sphere I: Overview and the geometric picture.” In: *IEEE Transactions on Information Theory* 63.2 (2016), pp. 853–884.

[102] Ju Sun, Qing Qu, and John Wright. “Complete dictionary recovery over the sphere ii: Recovery by riemannian trust-region method.” In: *IEEE Transactions on Information Theory* 63.2 (2016), pp. 885–914.

[103] Michael E Tipping and Christopher M Bishop. “Mixtures of probabilistic principal component analyzers.” In: *Neural computation* 11.2 (1999), pp. 443–482.

[104] Carlo Tomasi and Takeo Kanade. “Shape and motion from image streams under orthography: a factorization method.” In: *International journal of computer vision* 9.2 (1992), pp. 137–154.

[105] Roberto Tron and René Vidal. “A benchmark for the comparison of 3-d motion segmentation algorithms.” In: *2007 IEEE conference on computer vision and pattern recognition*. IEEE. 2007, pp. 1–8.

[106] Manolis C Tsakiris. “Dual Principal Component Pursuit and Filtrated Algebraic Subspace Clustering.” PhD thesis. Johns Hopkins University, 2017.

[107] Manolis C Tsakiris and Rene Vidal. “Filtrated algebraic subspace clustering.” In: *SIAM Journal on Imaging Sciences* 10.1 (2017), pp. 372–415.

[108] Manolis C Tsakiris and Rene Vidal. "Filtrated spectral algebraic subspace clustering." In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2015, pp. 28–36.

[109] Manolis C Tsakiris and Rene Vidal. "Hyperplane Clustering Via Dual Principal Component Pursuit." In: *arXiv preprint arXiv:1706.01604* (2017).

[110] Manolis C Tsakiris and René Vidal. "Algebraic clustering of affine subspaces." In: *IEEE transactions on pattern analysis and machine intelligence* 40.2 (2017), pp. 482–489.

[111] Manolis C Tsakiris and René Vidal. "Dual principal component pursuit." In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2015, pp. 10–18.

[112] Manolis C Tsakiris and René Vidal. "Dual principal component pursuit." In: *The Journal of Machine Learning Research* 19.1 (2018), pp. 684–732.

[113] Manolis C Tsakiris and René Vidal. "Hyperplane clustering via dual principal component pursuit." In: *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org. 2017, pp. 3472–3481.

[114] Paul Tseng. "Nearest q-flat to m points." In: *Journal of Optimization Theory and Applications* 105.1 (2000), pp. 249–252.

[115] Duygu Ucar, Qingyang Hu, and Kai Tan. "Combinatorial chromatin modification patterns in the human genome revealed by subspace clustering." In: *Nucleic acids research* 39.10 (2011), pp. 4063–4075.

[116] Konstantin Usevich and Ivan Markovsky. "Optimization on a Grassmann manifold with application to system identification." In: *Automatica* 50.6 (2014), pp. 1656–1662.

[117] Laurens Van Der Maaten, Eric Postma, and Jaap Van den Herik. "Dimensionality reduction: a comparative." In: *J Mach Learn Res* 10.66-71 (2009), p. 13.

[118] Aad W Van der Vaart. *Asymptotic statistics*. Vol. 3. Cambridge university press, 2000.

[119] Namrata Vaswani, Thierry Bouwmans, Sajid Javed, and Praneeth Narayanamurthy. "Robust subspace learning: Robust PCA, robust subspace tracking, and robust subspace recovery." In: *IEEE signal processing magazine* 35.4 (2018), pp. 32–55.

[120] Namrata Vaswani and Praneeth Narayanamurthy. "Static and dynamic robust PCA and matrix completion: A review." In: *Proceedings of the IEEE* 106.8 (2018), pp. 1359–1379.

[121] Roman Vershynin. "Introduction to the non-asymptotic analysis of random matrices." In: *arXiv preprint arXiv:1011.3027* (2010).

[122] Jean-Philippe Vial. "Strong and weak convexity of sets and functions." In: *Mathematics of Operations Research* 8.2 (1983), pp. 231–259.

[123] Rene Vidal. "Subspace clustering." In: *IEEE Signal Processing Magazine* 28.2 (2011), pp. 52–68.

[124] René Vidal and Paolo Favaro. "Low rank subspace clustering (LRSC)." In: *Pattern Recognition Letters* 43 (2014), pp. 47–61.

[125] René Vidal and Richard Hartley. "Motion segmentation with missing data using powerfactorization and gpca." In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.* Vol. 2. IEEE. 2004, pp. II–II.

[126] René Vidal, Yi Ma, and Jacopo Piazzi. "A new GPCA algorithm for clustering subspaces by fitting, differentiating and dividing polynomials." In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.* Vol. 1. IEEE. 2004, pp. I–I.

[127] Rene Vidal, Yi Ma, and Shankar Sastry. "Generalized principal component analysis (GPCA)." In: *IEEE transactions on pattern analysis and machine intelligence* 27.12 (2005), pp. 1945–1959.

[128] René Vidal, Yi Ma, Stefano Soatto, and Shankar Sastry. "Two-view multibody structure from motion." In: *International Journal of Computer Vision* 68.1 (2006), pp. 7–25.

[129] René Vidal, Stefano Soatto, Yi Ma, and Shankar Sastry. "An algebraic geometric approach to the identification of a class of linear hybrid systems." In: *42nd IEEE International Conference on Decision and Control (IEEE Cat. No. 03CH37475).* Vol. 1. IEEE. 2003, pp. 167–172.

[130] René Vidal, Roberto Tron, and Richard Hartley. "Multiframe motion segmentation with missing data using Power Factorization and GPCA." In: *International Journal of Computer Vision* 79.1 (2008), pp. 85–105.

[131] Ulrike Von Luxburg. "A tutorial on spectral clustering." In: *Statistics and computing* 17.4 (2007), pp. 395–416.

[132] Richard Y Wang, Veda C Storey, and Christopher P Firth. "A framework for analysis of data quality research." In: *IEEE transactions on knowledge and data engineering* 7.4 (1995), pp. 623–640.

[133] Yu-Xiang Wang, Huan Xu, and Chenlei Leng. "Provable Subspace Clustering: When LRR meets SSC." In: *NIPS.* Vol. 1. 2. 2013, p. 5.

[134] Tomas Werner and Andrew Zisserman. “New techniques for automated architectural reconstruction from photographs.” In: *European conference on computer vision*. Springer. 2002, pp. 541–555.

[135] Yan Xia, Xudong Cao, Fang Wen, Gang Hua, and Jian Sun. “Learning discriminative reconstructions for unsupervised outlier removal.” In: *Proceedings of the IEEE International Conference on Computer Vision*. 2015, pp. 1511–1519.

[136] Huan Xu, Constantine Caramanis, and Sujay Sanghavi. “Robust PCA via outlier pursuit.” In: *IEEE transactions on information theory* 58.5 (2012), pp. 3047–3064.

[137] Xu Xu, Mingjun Zhong, and Chonghui Guo. “A hyperplane clustering algorithm for estimating the mixing matrix in sparse component analysis.” In: *Neural Processing Letters* 47.2 (2018), pp. 475–490.

[138] Jingyu Yan and Marc Pollefeys. “A general framework for motion segmentation: Independent, articulated, rigid, non-rigid, degenerate and non-degenerate.” In: *European conference on computer vision*. Springer. 2006, pp. 94–106.

[139] Allen Y Yang, Shankar R Rao, and Yi Ma. “Robust statistical estimation and segmentation of multiple subspaces.” In: *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW’06)*. IEEE. 2006, pp. 99–99.

[140] Wei Hong Yang, Lei-Hong Zhang, and Ruyi Song. “Optimality conditions for the nonlinear programming problems on Riemannian manifolds.” In: *Pacific Journal of Optimization* 10.2 (2014), pp. 415–434.

[141] Chong You, Chun-Guang Li, Daniel P Robinson, and René Vidal. “Oracle based active set algorithm for scalable elastic net subspace clustering.” In:

Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, pp. 3928–3937.

[142] Chong You, Daniel P Robinson, and René Vidal. "Provable self-representation based outlier detection in a union of subspaces." In: *Proceedings of the ieee conference on computer vision and pattern recognition.* 2017, pp. 3395–3404.

[143] Chong You, Daniel Robinson, and René Vidal. "Scalable sparse subspace clustering by orthogonal matching pursuit." In: *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2016, pp. 3918–3927.

[144] Teng Zhang. "Robust subspace recovery by Tyler's M-estimator." In: *Information and Inference: A Journal of the IMA* 5.1 (2016), pp. 1–21.

[145] Teng Zhang and Gilad Lerman. "A novel m-estimator for robust pca." In: *The Journal of Machine Learning Research* 15.1 (2014), pp. 749–808.

[146] Teng Zhang, Arthur Szlam, and Gilad Lerman. "Median k-flats for hybrid linear modeling with many outliers." In: *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops.* IEEE. 2009, pp. 234–241.

[147] Teng Zhang, Arthur Szlam, Yi Wang, and Gilad Lerman. "Hybrid linear modeling via local best-fit flats." In: *International journal of computer vision* 100.3 (2012), pp. 217–240.

[148] Yuqian Zhang, Han-wen Kuo, and John Wright. "Structured local minima in sparse blind deconvolution." In: *Proceedings of the 32nd International Conference on Neural Information Processing Systems.* 2018, pp. 2328–2337.

[149] Zhengyou Zhang. "A flexible new technique for camera calibration." In: *IEEE Transactions on pattern analysis and machine intelligence* 22.11 (2000), pp. 1330–1334.

[150] Xingquan Zhu and Xindong Wu. “Class noise vs. attribute noise: A quantitative study.” In: *Artificial intelligence review* 22.3 (2004), pp. 177–210.

[151] Zhihui Zhu, Tianyu Ding, Daniel Robinson, Manolis Tsakiris, and René Vidal. “A linearly convergent method for non-smooth non-convex optimization on the grassmannian with applications to robust subspace and dictionary learning.” In: *Advances in Neural Information Processing Systems*. 2019, pp. 9437–9447.

[152] Zhihui Zhu, Yifan Wang, Daniel P Robinson, Daniel Q Naiman, Rene Vidal, and Manolis C Tsakiris. “Dual principal component pursuit: probability analysis and efficient algorithms.” In: *arXiv preprint arXiv:1812.09924* (2018).

[153] Zhihui Zhu, Yifan Wang, Daniel Robinson, Daniel Naiman, Rene Vidal, and Manolis Tsakiris. “Dual principal component pursuit: Improved analysis and efficient algorithms.” In: *Advances in Neural Information Processing Systems*. 2018, pp. 2171–2181.

Vita

Tianyu Ding was born in Anhui, China in 1992. He received his Bachelor of Science degree in Mathematics from Sun Yat-Sen University in 2014 and won the title of Outstanding Graduate. He received his Master of Science and Engineering degree in Financial Mathematics from the Department of Applied Mathematics and Statistics at Johns Hopkins University in 2016. From 2016 to 2021, he enrolled in the Ph.D. program in the same department, during which time he received another Master of Science and Engineering degree in Computer Science from Johns Hopkins University in 2020. His research interest lies in the intersection of numerical optimization, machine learning, deep learning, and computer vision. In the summer of 2020, he was a Research Intern at Microsoft and worked for designing deep neural architectures for video frame interpolation.