A Generalized Volume-Correlation Subspace Detector and its application in Multiuser Detection

Hailong Shi, Hao Zhang and Xiqin Wang

Abstract

The problem of subspace signal detection, i.e., detecting whether a signal lies within a subspace arises in a wide range of applications in the signal processing community, including communication, radar, array signal processing, hyperspectral image, etc. In this paper, a novel volume-correlation subspace detector is proposed to detect subspace signals contaminated by noises and subspace interferences. It is shown that, the volume of parallelotope, which is spanned by the signal subspace’s basis vectors together with multi-dimensional received signal vectors, can be calculated and used to detect subspace signals with interferences. It is also proved in this paper that, with the knowledge of the signal subspace and the interference subspace, and by accumulating multi-dimensional signal observations, the volume-correlation subspace detector will eliminate the influence of noise as well as the influence of interference asymptotically. And this is an advantage that traditional subspace detectors don’t have. Besides, the application of this volume-correlation subspace detector to Active User Identification in multiuser detection systems with multipath channels is discussed in this paper. Active User Identification is a well known and essential preprocess stage in multiuser detection systems, it can reduce the system complexity and improve the overall decoding performance. It is shown that the proposed volume-correlation subspace detector has its interior advantage to deal with the interference of other users in the system with a multipath channel, and provides asymptotically ideal performance in detection of active users. Numerical simulations is also provided to validate our conclusion.

Index Terms

Subspace Signal Detection, Matched Subspace Detector, Generalized Energy Detector, Volume-Correlation Subspace Detector, Active User Identification, Multiuser Detection

I. INTRODUCTION

The problem of subspace signal detection, i.e., detecting whether a signal lies within a subspace arises in a wide range of applications in the signal processing community, such as communication [1][2][3][4], radar [5][6][7][8][9], array signal processing [10][11], as well as anomaly detection in hyperspectral imagery [12]. The basic problem of subspace detection can be described as a binary hypothesis test, in which the observation signal vectors $r \in \mathbb{R}^P$...
are given, and the problem is to decide whether the signal components in these observations lie within a known linear subspace, it is formulated as

\[
H_0: \quad r = x_I + w, \\
H_1: \quad r = x_s + x_I + w, 
\]

(1)

here \( w \) is the white Gaussian noise component, i.e., \( w \sim N(0, \sigma^2 I_P) \); and it is assumed that the to-be-detected signal, i.e., \( x_s \), obeys the linear subspace model, which means

\[
x_s = X_s \alpha, \quad X_s \in \mathbb{R}^{P \times d}, \quad \alpha \in \mathbb{R}^d,
\]

(2)

where the column vectors of matrix \( X_s \) form the basis of the to-be-detected signal subspace, and \( \alpha \) is the corresponding subspace coefficient vector, and \( d < P \) is the dimension of the to-be-detected signal subspace. For convenience, we call the to-be-detected signal subspace as the **target signal subspace**, and denote it by \( \mathcal{X}_s := \text{span}(X_s) \).

In the same way, we also assume that the interference signal, i.e., \( x_I \), obeys the linear subspace signal model, which means

\[
x_I = X_I \beta, \quad X_I \in \mathbb{R}^{P \times p}, \quad \beta \in \mathbb{R}^p,
\]

(3)

where the column vectors of matrix \( X_I \) form the basis of the interference signal subspace, and \( \beta \) denotes the corresponding subspace coefficient vector, and \( p < P \) is the dimension of the interference signal subspace. In the same way, the **interference signal subspace** is denoted by

\[
\mathcal{X}_I := \text{span}(X_I).
\]

Therefore, our goal is to decide the existence of the target subspace signal \( x_s = X_s \alpha \) from the received signal \( r \in \mathbb{R}^P \) with noise and interference \( x_I = X_I \beta \). Without loss of generality, it is assumed that the target signal subspace \( \mathcal{X}_s \) is linearly independent and disjoint with the interference signal subspace \( \mathcal{X}_I \), i.e., \( \mathcal{X}_s \cap \mathcal{X}_I = \emptyset \); besides, the noise level \( \sigma \) is assumed to be known, while the subspace coefficients \( \alpha \) and \( \beta \) are unknown.

The subspace signal detection described here is a generalization of the conventional signal detection problem, in which the matched filter is used as the common approach, the reason is that subspace signal detection problem concerns about the detection of multi-dimensional subspace signal \( x_s = X_s \alpha \), while the conventional signal detection problem detects signals in the form of individual vectors; in other words, the conventional signal detection problem is a one-dimensional special case of the subspace signal detection. The reason why subspace signal is needed to be dealt with is twofold. First, because of the physical process of signal transmitting and acquisition, for example, the wireless channel’s multi-path effect \([13][2]\), the scattering mechanisms of radar’s target signals and clutters \([6][14]\), as well as the interior connection between image’s neighbor pixels \([12]\), etc, signals we concern are usually multi-dimensional, and needed to be modeled as linear subspaces. Second, due to the randomness in signal modulation and transmitting, it is difficult to obtain the subspace coefficient \( \alpha \) in the receiver, and thus the
vector form of the target signal \( x_s \) is usually unknown. For those reasons, more general approaches are needed to use knowledge of signal subspace \( X_s \) and detect the subspace signal \( x_s = X_s \alpha \).

There have been well-known approaches aiming at detection subspace signals, which is firstly proposed by L. Scharf et.al., it is named as the Matched Subspace Detector [15], and already has tremendous variations and applications [16] [17] [18] [19]. The core technique of the Matched Subspace Detector is to use the principal of Generalized Likelihood Ratio Test, with subspace coefficients \( \alpha \) and \( \beta \) replaced by their maximum likelihood estimation. The test statistics of problem (1) given in [15] is:

\[
t(r) = \frac{1}{\sigma^2} \|(P_{X_I}^\perp - P_{X_{sI}}^\perp) r\|_2^2 = \frac{1}{\sigma^2} \| P_{X_{sI}}^\perp X_s r \|_2^2,
\]

(4)

where \( P_{X_I}^\perp \) is the orthogonal projection operator to subspace \( X_I^\perp \), \( P_{X_{sI}}^\perp \) is the orthogonal projection operator to subspace \( (X_s \oplus X_I) \perp \), and \( P_{X_{sI}}^\perp X_s \) is the orthogonal projection operator to subspace \( P_{X_{sI}}^\perp X_s \). The testing result is obtained by comparing (4) with thresholds derived according to certain conditions (such as Constant False Alarm Rate). (4) indicates that the test statistics is the energy of the part of \( r \) projected orthogonally onto subspace \( P_{X_{sI}}^\perp X_s \), therefore the detector in (4) is also called the Generalized Energy Detector [15].

However, the Generalized Energy Detector cannot eliminate the influence of the interference subspace, because the target signal subspace \( X_s \) may not be orthogonal to the interference signal subspace \( X_I \), which will cause a loss of energy when projecting the target signal, to be more specific, under \( H_1 \) hypothesis, the projection in (4) will yield

\[
P_{P_{X_{sI}}^\perp X_s} r = P_{X_{sI}}^\perp X_s \alpha + P_{P_{X_{sI}}^\perp X_s} w,
\]

(5)

As depicted in figure [1] if the interference subspace \( X_I \) is "more close to" the target subspace \( X_s \), or less orthogonal

---

Fig. 1: Depiction of the Generalized Energy Detector (right), and the influence of interference signal subspace on target signal subspace in Generalized Energy Detector (left)
to $\mathcal{X}_s$, then the signal component $P_{\mathcal{X}_s} P_{\mathcal{X}_I}^\perp \mathbf{x}$ in (5) will be weaker compared to noise, therefore more difficult to detect.

The reason why the detection performance of Generalized Energy Detector is influenced by the interference signal subspace is that the projection in (4) is onto the subspace $P_{\mathcal{X}_s} P_{\mathcal{X}_I}^\perp$, rather than $\mathcal{X}_s$. As a result, the Generalized Energy Detector cannot eliminate the influence of interference, while in this paper, we are going to propose a novel subspace detection technique to asymptotically eliminate the influence of interference. Before we introduce this technique, firstly we begin with a simple demonstration, if we consider the hypothesis test (1) without noise, then under $H_1$ hypothesis, the noiseless received signal is

$$ y := x_s + x_I = X_s \alpha + X_I \beta, \quad (6) $$

then we have

$$ y \in \mathcal{X}_s \bigoplus \mathcal{X}_I. \quad (7) $$

Similarly, under $H_0$ hypothesis, we have

$$ y \in \mathcal{X}_I. \quad (8) $$

If we can get multiple observations of the noiseless received signal $y$, which means the coefficients $\alpha$ (or $\beta$) vary with different observations (without loss of generality, $\alpha$ and $\beta$ can be assumed to be independent random vectors), denote the multiple observations of the received signal by $y^{(1)}, y^{(2)}, \cdots$, then we have

$$ y^{(i)} := x_s^{(i)} + x_I^{(i)} = X_s \alpha^{(i)} + X_I \beta^{(i)}, \quad i = 1, 2, \cdots \quad (9) $$

where $\alpha^{(i)}$ and $\beta^{(i)}$ are the target and interference subspace coefficients corresponding to the $i$'th observation. Then we can have for $m \geq 1$,

$$ H_0 : \text{span}([y^{(1)}, y^{(2)}, \cdots, y^{(m)}]) \subset \mathcal{X}_I, $$

$$ H_1 : \text{span}([y^{(1)}, y^{(2)}, \cdots, y^{(m)}]) \subset \mathcal{X}_s \bigoplus \mathcal{X}_I. \quad (10) $$

Therefore, (10) reminds the use of multi-dimensional information from the multiple observations of the received signal, i.e., the subspace spanned by the columns of the observation matrix $Y^{(m)} := [y^{(1)}, y^{(2)}, \cdots, y^{(m)}]$, to explore the relation between the received signal and the target subspace $\mathcal{X}_s$. It is not difficult to obtain and process multi-dimensional information from the multiple observations of the received signal, and it is well-known that array receivers or sensors, as well as MIMO systems have been widely used in communication, radar and other signal processing systems [10][11][9], and this also makes it possible to exploit the multi-dimensional information hidden in the obtained multiple observations to gain a better performance for the subspace signal detection. Unfortunately, as far as we know, all these applications mentioned above relies on the traditional matched subspace detector approach, and only concerns about a single observation vector $r$ (or in noiseless case $y$). As is said that the traditional matched
subspace detector approach is not interference-resistant, so in this paper, we will make use of the multi-dimensional information from multi-dimensional observations, i.e. the linear subspace spanned by the observations as in (10), and propose a new interference-resistant detection approach to detect subspace signals.

Applying the theory of linear subspace to the domain of signal processing is not a new idea, actually, it has been commonly applied in image processing and computer vision [22][23], machine learning [24][25][26], as well as communication and radar [27][28][29][30]. Briefly speaking, these applications model the multi-dimensional signal as a linear subspace, and deal with the problem of subspace-oriented optimization or estimation using knowledge of an abstract manifold (i.e., the Grassmann Manifold, which is a topological space where each point is a linear subspace); the rich topological structure (such as the Tangent space, Geodesic Distance) as well as various definitions of metrics on Grassmann manifold make it possible to exploit multi-dimensional information and derive better optimization and estimation result. In the following sections, we will utilize the tool of principal angles of subspaces, as well as volumes of matrices to study the subspace signal detection problem (1).

II. PRELIMINARY BACKGROUND

A. Principal Angles between Subspaces

The well-known principal angles [31] can be used to describe the relationship between different subspaces, it is defined as follows:

**Definition 1:** For two linear subspaces \( X_1 \) and \( X_2 \), with dimensions \( \text{dim}(X_1) = d_1, \text{dim}(X_2) = d_2 \). Take \( m = \min(d_1, d_2) \), then the principal angles \( 0 \leq \theta_1 \leq \cdots \leq \theta_m \leq \pi/2 \) between these two subspaces are defined recursively by

\[
\cos \theta_i = \max_{u_i \in X_1, v_i \in X_2} u_i^T v_i, \quad \text{subject to} \quad \|u_i\| = \|v_i\| = 1, \quad u_i^T u_j = 0, \quad v_i^T v_j = 0, \quad (j = 1, \cdots, i - 1, i = 2, \cdots m).
\]

The principal angles are important mathematical tools to describe relations of subspaces. In fact, in the theory of Grassmann manifold [32][33], the Geodesic distance, which is an important metric measure, is defined using the principal angles [32]; besides, according to [33], there are numerous kinds of distance measures that are defined based on the principal angles. In this paper, we will use another distance measure, i.e., the volume, which is closely related to the principal angles, to construct our subspace detector.

B. The Volume of a matrix and the principal angles between subspaces

The \( d \)-dimensional (\( d < P \)) volume of a full-rank matrix \( X \in \mathbb{R}^{P \times d} \) is defined as [34]

\[
\text{vol}_d(X) := \prod_{i=1}^{d} \sigma_i, \quad (11)
\]

where \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_d > 0 \) are the singular values of matrix \( X \). The volume of the matrix \( X \) is also referred to as the \( d \)-dimensional parallelotope spanned by the column vectors of \( X \). For the matrix \( X \) is of full column

June 6, 2014 DRAFT
rank, the volume is equivalently [34][33]

\[ \text{vol}_d(X) = \sqrt{|\text{det}(X^TX)|}. \]  

(12)

Volume provides a measure of separation between two linear subspaces and is closely related to the principal angles between subspaces, according to [31], for two subspaces \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \), with dimensions \( \text{dim}(\mathcal{X}_1) = d_1 \), \( \text{dim}(\mathcal{X}_2) = d_2 \), and basis matrix \( X_1 \) and \( X_2 \), we have

\[ \frac{\text{vol}_{d_1+d_2}([X_1, X_2])}{\text{vol}_{d_1}(X_1) \text{vol}_{d_2}(X_2)} = \prod_{j=1}^{\min(d_1,d_2)} \sin \theta_j(\mathcal{X}_1, \mathcal{X}_2), \]  

(13)

where \( 0 \leq \theta_j(\mathcal{X}_1, \mathcal{X}_2) \leq 2\pi, 1 \leq j \leq \min(d_1, d_2) \) are the principal angles of subspace \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \).

It can be seen intuitively from (13) that, the expression \( \text{vol}_{d_1+d_2}([X_1, X_2])/(\text{vol}_{d_1}(X_1) \text{vol}_{d_2}(X_2)) \) can be used to derive a distance measure of subspace \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \). Because when \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \) intersect, i.e., \( \mathcal{X}_1 \cap \mathcal{X}_2 \neq \emptyset \), we have \( \text{vol}_{d_1+d_2}([X_1, X_2]) = 0 \); when \( \mathcal{X}_1 \) is perpendicular to \( \mathcal{X}_2 \), then \( \text{vol}_{d_1+d_2}([X_1, X_2]) = \text{vol}_{d_1}(X_1) \text{vol}_{d_2}(X_2) \).

The right side of (13) is the product of sines of the principal angles between subspaces \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \). As we know that various metric measures of linear subspaces are defined based on principal angles, typically such as

- Chordal Distance or Projection distance

\[ d_{\text{proj}} := \left( \sum_{j=1}^{\min(d_1,d_2)} \sin^2 \theta_j(\mathcal{X}_1, \mathcal{X}_2) \right)^{1/2}, \]

- Binet-Cauchy Distance

\[ d_{\text{BC}} := \left( 1 - \prod_{j=1}^{\min(d_1,d_2)} \cos^2 \theta_j(\mathcal{X}_1, \mathcal{X}_2) \right)^{1/2}, \]

- Procrustes Distance

\[ d_{\text{proc}} := 2 \left( \sum_{j=1}^{\min(d_1,d_2)} \sin^2 \frac{\theta_j(\mathcal{X}_1, \mathcal{X}_2)}{2} \right)^{1/2}. \]

Additionally, more kinds of distance measures can be defined using functions satisfying specific conditions [33]. Therefore, although we cannot validate whether (13) satisfies the condition of metric functions, i.e., the triangle inequality, following the terminology used in [33], and for simplicity of calculation and convenience, without verifying the triangle inequality, we regard the product of principal sines induced by volume in (13) as a generalized distance measure for linear subspaces in the following analyzes, and use it as the "volume-correlation", which plays a key role in our proposed subspace detector.

### III. The Volume-Correlation Subspace Detector

In this section we will introduce the Volume-Correlation Subspace Detector, as is discussed in the previous section, multi-dimensional observations of the received signal will be used to perform subspace detection, these multiple observations of problem (1) is of the form of:

\[ H_0 : r^{(i)} = X_1\beta^{(i)} + w^{(i)}, \]

\[ H_1 : r^{(i)} = X_s\alpha^{(i)} + X_1\beta^{(i)} + w^{(i)}, \quad i = 1, 2, \ldots, \]  

(14)
where $\alpha^{(i)}$ and $\beta^{(i)}$ are subspace coefficients, $X_s \in \mathbb{R}^{P \times d}$, $X_I \in \mathbb{R}^{P \times p}$, and $w^{(i)}$ is the white Gaussian noise. We need to detect the existence of target subspace signal $X_s = \text{span}(X_s)$ in the observations $r^{(i)} \in \mathbb{R}^P$, $i = 1, 2, \cdots$.

Before the introduction of the subspace detector, we start with another demonstrative observation, as is mentioned previously, in the noiseless situation, and under $H_1$ hypothesis, the multiple observations

$$y^{(i)} = X_s \alpha^{(i)} + X_I \beta^{(i)}, \quad i = 1, 2, \cdots$$

satisfy

$$\text{span}(Y^{(m)}) = \text{span}([y^{(1)}, y^{(2)}, \cdots, y^{(m)}]) \subset X_s \bigoplus X_I.$$

and we have

$$m \leq d + p, \quad \text{rank}(Y^{(m)}) = m, \text{span}(Y^{(m)}) \subset X_s \bigoplus X_I, \quad (15)$$

$$m > d + p, \quad \text{rank}(Y^{(m)}) = d + p, \text{span}(Y^{(m)}) = X_s \bigoplus X_I. \quad (16)$$

Similarly, in the $H_0$ hypothesis, we have

$$m \leq p, \quad \text{rank}(Y^{(m)}) = m, \text{span}(Y^{(m)}) \subset X_I, \quad (17)$$

$$m > p, \quad \text{rank}(Y^{(m)}) = p, \text{span}(Y^{(m)}) = X_I. \quad (18)$$

Then according to the previous analysis, under $H_1$ hypothesis, when $m > p + d$, then the volume of matrix $[Y^{(m)}, X_s]$ will satisfy

$$\text{vol}_{m+d}([Y^{(m)}, X_s]) = 0. \quad (19)$$

Similarly, under $H_0$ hypothesis, when $m > p$, because $X_s$ and $X_I$ are linearly independent, thus the volume of $[Y^{(m)}, X_s]$ will satisfy

$$\text{vol}_{m+d}([Y^{(m)}, X_s]) \neq 0. \quad (20)$$

$(19)$ and $(20)$ implies that, in noiseless situations, when the target signal subspace $X_s$ and the interference subspace $X_I$ are linearly independent, $m \geq d + p$ observations will guarantee the difference of the volume of $[Y^{(m)}, X_s]$ between $H_1$ and $H_0$ hypotheses, and this leads to our Volume-Correlation Subspace Detector.

A. The Volume-Correlation Subspace Detector

In the noiseless situation, the Volume-Correlation Subspace Detector is:

Detector 1’. ($w^{(i)} = 0$, i.e., noiseless scenario)

- Obtain $Y^{(m)} = [y^{(1)}, \cdots, y^{(m)}]$, where

  $$H_0 : \quad y^{(i)} = X_I \beta^{(i)},$$

  $$H_1 : \quad y^{(i)} = X_s \alpha^{(i)} + X_I \beta^{(i)}, \quad i = 1, 2, \cdots, m.$$  

- Determine the rank of $Y^{(m)}$, $r := \text{rank}(Y^{(m)}), m \geq 1$
Remark 2. The estimation of the dimension of the signal subspace (composed of both the target signal and interference signal subspaces) both in noised and noiseless situations is needed. This is because the dimension is a key parameter needed to calculate the volume. For example, in the noiseless situation, and $H_1$ hypothesis, when $m \leq p + d$, the matrix $Y^{(m)}$ is full-rank; but when $m > p + d$, $Y^{(m)} \in \mathbb{R}^{p \times m}$ is not full-rank, then the $m$-dimensional volume $\text{vol}_m(Y^{(m)})$ will reduces to 0, as a result, the Detector 1’ won’t work any more. Therefore, the dimension of the signal subspace $r$ is needed for our volume calculation.
• Remark 3. The detection result is given by comparing the reciprocal of the volume-correlation result $1/t(Y^{(m)})$ (or $1/t(R^{(m)})$) with threshold $1/\tau(\mathcal{X}_s, \mathcal{X}_t)$ (or $c/\tau(\mathcal{X}_s, \mathcal{X}_t)$). The reason for use of reciprocals is to enlarge the difference of $1/t(Y^{(m)})$ (or $1/t(R^{(m)})$) between two hypotheses, which will make the comparison to be easier. Besides, in noised situation, an arbitrary constant $c$ is chosen for the fine tuning the threshold, and to enhance the detection performance. The choosing and analysis of the threshold will be discussed in the following sections. The most important issue which should be noted here is that from (13), it can be seen that $\tau(\mathcal{X}_s, \mathcal{X}_t)$ is a value decided only by the relation between subspaces $\mathcal{X}_s$ and $\mathcal{X}_t$.

• Remark 4. In the noised situation, the Detector 1 uses the well-known eigen-decomposition subspace method as a preprocess for the estimation of signal subspace. As in these classic subspace methods such as MUSIC or PCA, the dimension as well as the basis of the signal subspace is estimated through the eigenvalue decomposition of sampling covariance matrix. As is known that the estimations of dimension and the basis are generally two independent stages, and for the first stage, i.e., the estimation of the dimension of signal subspace, there are already a lot of famous approaches, such as Akaike-Information Criterion (AIC)[35], Minimum Description Length (MDL)[36], Bayesian Information Criterion (BIC)[37], Predictive Description Length (PDL)[38] and the most recent Entropy Estimation of Eigenvalues (EEE)[39]. All those approaches are quite mature and ensure high precision in the dimension estimation. Therefore, we only focus on the second stage, and assume the dimension of signal subspace is precisely estimated, which is also a general assumption in other analysis of the subspace methods. In the following sections, we will analysis the asymptotic error of our Volume-Correlation Detector caused by the signal subspace estimation approach.

• Remark 5. There are two significant factors in the proposed volume-correlation detector: first, the detector makes use of multi-dimensional observations of the received signal, i.e., $r^{(1)}, \ldots, r^{(m)}$ (or $y^{(1)}, \ldots, y^{(m)}$), which will bring much more information than a single observation; second, the test statistics is calculated using volume of the matrix composed of the multi-dimensional observation data $R^{(m)}$ and the basis $X_s$ of target signal subspace $\mathcal{X}_s$. As far as we know, volume is firstly used to construct a subspace detector, and it has advantages in exploiting the multi-dimensional information. Specially, for two 1-dimensional vectors $x_1$ and $x_2$, $\text{vol}_2([x_1, x_2]) = \|x_1\|_2 \|x_2\|_2 \sin \theta(x_1, x_2)$, this means volume is the amplitude of the exterior product, which is in correspondence with the inner product $\langle x_1, x_2 \rangle = \|x_1\|_2 \|x_2\|_2 \cos \theta(x_1, x_2)$ (also known as "correlation", and is used in matched filter detectors). Besides, as is mentioned in the previous section, volume can be used to derive a generalized distance measure for multi-dimensional subspaces. As a result, we regard the volume-based test statistics

$$t(Y^{(m)}) := \text{vol}_{d+r}([X_s, Y^{(m)}]) / (\text{vol}_d(X_s) \text{vol}_r(Y^{(m)})),$$

(21)

and

$$t(R^{(m)}) := \text{vol}_{d+r}([X_s, \hat{U}^{(m)}_{sig}]) / (\text{vol}_d(X_s) \text{vol}_r(\hat{U}^{(m)}_{sig}))$$

(22)

as a generalized volume-correlation, meaning that (21) and (22) imply a kind of measure of correlation between subspaces $\text{span}(X_s)$ and $\text{span}(Y^{(m)})$ (or $\text{span}(\hat{U}^{(m)}_{sig})$). As we know that the exterior product is the product...
defined in Grassmann Algebra, and Grassmann Algebra is a powerful mathematical tool in describing multi-
dimensional geometry; therefore our volume-correlation subspace detector inherits the advantage of Grassmann
Algebra, and is good at dealing with multi-dimensional information by using knowledge of multi-dimensional
geometry.

• Remark 6. Consider the traditional Generalized Energy Detector in (4), denote the orthogonal basis matrix
of subspace \( P_{\tilde{X}_t}X_s \) by \( Q_{P_{\tilde{X}_t}X_s} \in \mathbb{R}^{P \times d} \), then the projection operation can be written in a matrix form
\( Q_{P_{\tilde{X}_t}X_s}^T X_s r \), then we have
\[
\|P_{\tilde{X}_t}X_s r\|_2^2 = \det \left( r^T Q_{P_{\tilde{X}_t}X_s}^T Q_{P_{\tilde{X}_t}X_s} r \right) = \text{vol}_{1}^2 \left( Q_{P_{\tilde{X}_t}X_s}^T X_s r \right).
\]
This implies that the Generalized Energy Detector can be regarded as a special case of this paper’s Volume-
Correlation Subspace Detector. The main difference is that, the Volume-Correlation Subspace Detector makes
use of multi-dimensional observation data \( R^{(m)} = [r^{(1)}, \cdots, r^{(m)}] \), while the Generalized Energy Detector
only uses one-dimensional observation \( r \) and calculates the one-dimensional volume in (23) according to the
GLRT principle. Although currently we cannot give the optimized form of our volume detector under certain
principle, the usage of multi-dimensional information makes our volume-correlation detector indisputably
advantageous, and the study of the optimality will be left for future work.

In the next section, we will give the theoretical analysis of the advantage of the Volume-Correlation Subspace
Detector.

IV. ANALYSIS OF THE VOLUME-CORRELATION SUBSPACE DETECTOR: THE NOISELESS SITUATION

In this section, we will show theoretically that, in the noiseless situation, by using multi-dimensional observa-
tions of the received signal, our Volume-Correlation Subspace Detector will totally eliminate the influence of the
interference signal. The main result is the following theorems:

**Theorem 1:** In noiseless situation, under \( H_1 \) hypothesis, the multiple observations of received signal are of the
form of
\[
y^{(i)} = X_s \alpha^{(i)} + X_I \beta^{(i)}, \quad i = 1, 2, \cdots
\]
where \( X_s \in \mathbb{R}^{P \times d}, \alpha^{(i)} \in \mathbb{R}^d \) denotes the basis and coefficient of the target signal subspace, and \( X_I \in \mathbb{R}^{P \times p}, \beta^{(i)} \in \mathbb{R}^p \) denotes the basis and coefficient of the interference signal subspace, and \( \text{span}(X_s) \) is linearly
independent of \( \text{span}(X_I) \); then for the multi-dimensional observation data matrix
\[
Y^{(m)} = [y^{(1)}, \cdots, y^{(m)}] \in \mathbb{R}^{P \times m},
\]
denote \( r := \text{rank}(Y^{(m)}) \), the following volume-correlation in Detector 1
\[
t(Y^{(m)}) := \text{vol}_{d+r}(X_s, Y^{(m)}) / (\text{vol}_{d}(X_s) \text{vol}_r(Y^{(m)})),
\]
will satisfy:
• When \(1 \leq m < p\),
\[
t(Y^{(m)}) \geq t(Y^{(m+1)});
\] (25)

• When \(m \geq p + 1\),
\[
t(Y^{(m)}) = 0.
\] (26)

Theorem 1 demonstrates the advantage of our volume-correlation subspace detector. It shows that in noiseless situation, and under \(H_1\) hypothesis, as the number of observations \(m\) increases, the volume-correlation \(t(Y^{(m)})\) will generally descend to 0.

Under \(H_0\) hypothesis, here is another theorem:

**Theorem 2:** In noiseless situation, under \(H_0\) hypothesis, the multiple observations of received signal are of the form of
\[
y^{(i)} = X_f \beta^{(i)}, \quad i = 1, 2, \ldots
\]
where \(X_f \in \mathbb{R}^{P \times p}, \beta^{(i)} \in \mathbb{R}^p\) denotes the basis and coefficient of the interference signal subspace; then for the multi-dimensional observation data matrix
\[
Y^{(m)} = [y^{(1)}, \ldots, y^{(m)}] \in \mathbb{R}^{P \times m},
\]
denote \(r := \text{rank}(Y^{(m)})\), the following volume-correlation in Detector 1’
\[
t(Y^{(m)}) := \text{vol}_{d+r}([X_s, Y^{(m)}]) / (\text{vol}_d(X_s) \text{vol}_r(Y^{(m)})),
\] (27)
will satisfy:

• When \(1 \leq m < p\),
\[
t(Y^{(m)}) \geq t(Y^{(m+1)}),
\] (28)

and
\[
t(Y^{(m)}) \geq \tau(X_s, X_f) = \prod_{j=1}^{\min\{d, p\}} \sin \theta_j(X_s, X_f).
\] (29)

• When \(m \geq p\),
\[
t(Y^{(m)}) = \prod_{j=1}^{\min\{d, p\}} \sin \theta_j(X_s, X_f).
\] (30)

where \(\sin \theta_j(X_s, X_f) (1 \leq j \leq \min\{d, p\})\) are the principal angles between subspaces \(X_s\) and \(X_f\).

Theorem 2 shows the property of volume-correlation result under \(H_0\) hypothesis in noiseless situation. It can be seen that as the number of observations \(m\) increases, the volume-correlation \(t(Y^{(m)})\) will generally descend, not to 0, but to a threshold \(\tau(X_s, X_f) = \prod_{j=1}^{\min\{d, p\}} \sin \theta_j(X_s, X_f)\). As is mentioned that \(\prod_{j=1}^{\min\{d, p\}} \sin \theta_j(X_s, X_f)\) is only related to subspaces \(X_s\) and \(X_f\), therefore, as long as \(X_s\) and \(X_f\) are linearly independent, then \(\prod_{j=1}^{\min\{d, p\}} \sin \theta_j(X_s, X_f)\) will be greater than 0, which means under \(H_0\) hypothesis \(t(Y^{(m)})\) will never descend to 0.

Theorem 1 and 2 demonstrates the advantage of the volume-correlation subspace detector: in noiseless situation, as is shown by (25), (26) and (28), (29), the volume correlation \(t(Y^{(m)})\) will descend to 0 under \(H_1\) hypothesis,
and to threshold $\tau(\mathcal{X}_s, \mathcal{X}_I) = \prod_{j=1}^{\min\{d,p\}}$ under $H_0$ hypothesis. Therefore $\tau(\mathcal{X}_s, \mathcal{X}_I) = \prod_{j=1}^{\min\{d,p\}}$ is chosen as the threshold for hypothesis test, and the difference of volume-correlation under these two hypotheses will ensure the validation of test result. The simulated volume-correlation values under both hypotheses are plotted in figures [2] and [3]. In the simulation, we choose $P = 1024$, and $d = 10, p = 40$, the target and interference signal subspaces are chosen arbitrarily (simulation shows that $\tau(\mathcal{X}_s, \mathcal{X}_I) \approx 0.82$), and subspace coefficients $\alpha$ and $\beta$ are randomly generated, it can be seen that when $m \geq p = 40$, $t(Y^{(m)})$ descends to the value predicted by (26) and (29). The simulation implies that in noiseless situation, the influence of interference signal subspace on the volume-correlation $t(Y^{(m)})$ is totally eliminated when the number of observations $m \geq p = 40$. In addition, if the $\tau(\mathcal{X}_s, \mathcal{X}_I)$ is chosen as the threshold for hypothesis test, in the current simulation environment, it is shown that only $m > 3$ observations will ensure the accurate subspace detection result.

However, theorem [1] and [2] is not enough to explain the advantage of our Volume-Correlation Subspace Detector over the Generalized Energy Detector, because currently we have only discussed the noiseless situation. Actually, the projection $P_{P_{\mathcal{X}_I}, \mathcal{X}_s}$ in the Generalized Energy Detector
\[ t(r) = \frac{1}{\sigma^2} \| P_{P_{\mathcal{X}_I}, \mathcal{X}_s} r \|_2^2, \]
comes from the maximum likelihood estimation of $\alpha$ and $\beta$ in the GLRT test. Hence, we need to explore the performance of the Volume-Correlation Subspace Detector in noised situation. As is seen in Detector 1, in the noised situation, the classical eigenvalue decomposition methods are used to estimate the signal subspace $\text{span}(U_{\text{sig}}) = \mathcal{X}_s \oplus \mathcal{X}_I$ (or $\mathcal{X}_I$). And it has been well-known that the estimated signal subspace $\text{span}(\hat{U}_{\text{sig}}^{(m)})$ will approximate the signal subspace $\text{span}(U_{\text{sig}})$, therefore we can intuitively imagine that, in the noised situation, the volume-correlation $t(R^{(m)})$ will also approximate its values in noiseless situation, i.e., $0$ or $\tau(\mathcal{X}_s, \mathcal{X}_I)$. Detailed analysis will be given in the next section.

V. Analysis of the Volume-Correlation Subspace Detector: The noised situation

In this section, we will analyze the asymptotic performance of the volume-correlation subspace detector, and discuss its advantage in noise and interference elimination.

Before we introduce the main theoretical results, we would like to recall the classical eigen-decomposition subspace method used in Detector 1. As is known, the eigen-decomposition subspace method deals with subspace signals contaminated by noise:
\[ r = X\alpha + w, \]
where $X \in \mathbb{R}^{P \times r}$, $\alpha \in \mathbb{R}^r$ are basis and coefficient of the signal subspace, and $w \sim \mathcal{N}(0, \sigma^2 I_P)$ is the white Gaussian noise, then the auto-correlation matrix $C$ of the received signal $r$ can be written as
\[ C = \mathbb{E}\{rr^T\} = X \mathbb{E}\{\alpha\alpha^T\} X^T + \sigma^2 I_P = C_{\text{sig}} + C_{\text{noise}}. \tag{31} \]

The eigenvalues of the auto-correlation matrix are:
\[ \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r \geq \lambda_{r+1} = \cdots = \lambda_P = \sigma^2, \tag{32} \]
and the eigenvectors corresponding to $\lambda_1, \lambda_2, \ldots, \lambda_r$ are

$$u_1, u_2, \ldots, u_r$$

then the matrix $U_{\text{sig}} := [u_1, u_2, \ldots, u_r] \in \mathbb{R}^{p \times r}$ satisfies $\text{span}(U_{\text{sig}}) = \text{span}(X)$, so $\text{span}(U_{\text{sig}})$ is called the signal subspace, and $u_1, u_2, \ldots, u_r$ are called the signal eigenvectors; while the subspace spanned by the remaining noise eigenvectors $u_{r+1}, \ldots, u_p$ is called the noise subspace.

In real systems, the auto-correlation matrix $C^{(m)}$ can only be estimated by the sampled covariance matrix

$$\hat{C}^{(m)} := \frac{1}{m} \sum_{i=1}^{m} r(i) r^T(i),$$

the eigenvectors corresponding to the $r$ largest eigenvalues of $\hat{C}^{(m)}$ is

$$\hat{u}_1^{(m)}, \hat{u}_2^{(m)}, \ldots, \hat{u}_r^{(m)}$$

and is proved to be an asymptotic estimation of the real signal eigenvectors $u_1, u_2, \ldots, u_r$ \cite{40, 41, 42}, which means the estimated signal subspace $\text{span}(\hat{U}_{\text{sig}}^{(m)}) := \text{span}([\hat{u}_1^{(m)}, \hat{u}_2^{(m)}, \ldots, \hat{u}_r^{(m)}])$ is also an asymptotic estimation of the real signal subspace $\text{span}(U_{\text{sig}})$. Combining the above analysis, the main theorems of this section are as follows.
Theorem 3: In noised situation, under $H_1$ hypothesis, the multi-dimensional observations is of the form of

$$H_1 : \quad r^{(i)} = X_s \alpha^{(i)} + X_I \beta^{(i)} + w^{(i)}, \quad i = 1, 2, \ldots, m,$$

where $X_s \in \mathbb{R}^{P \times d}, \alpha^{(i)} \in \mathbb{R}^d$ are the basis and coefficient of the target signal subspace, and $X_I \in \mathbb{R}^{P \times p}, \beta^{(i)} \in \mathbb{R}^p$ are the basis and coefficient of the interference signal subspace; and $w^{(i)} \sim \mathcal{N}(0, \sigma^2 I_P)$ is the white Gaussian noise.

Then the volume-correlation of Detector 1 will satisfy:

Asymptotically (for large $m$), for any $0 < \varepsilon < 1$ and $\delta > 0$, if

$$m \geq \frac{1 + \varepsilon}{(\sqrt{\delta + 1} - 1)^2} \left( \sum_{i,j=1}^{p+d} \frac{\lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} + \sum_{i=1}^{p+d} (P - p - d) \frac{\lambda_i \sigma^2}{(\sigma^2 - \lambda_i)^2} \right),$$

then there exists a constant $C > 0$, such that

$$t^2(R^{(m)}) \leq s_p(Q_{X_s}^T P_{X_s} Q_{X_s}) \delta^d + O(\delta^{d+1}).$$

holds with probability

$$\mathbb{P} \geq 1 - \exp\left\{ - \frac{(p + d) \cdot P \cdot \varepsilon^2}{C} \right\}.$$
Here $\lambda_1, \lambda_2, \cdots, \lambda_r \geq \sigma^2$ are the corresponding eigenvalues of the auto-covariance matrix $C$ shown in (32), and $Q_{X_s,I}$ denotes the orthogonal basis matrix of the subspace $X_s \oplus X_I$; in addition, $s_k(A)$ is defined as the $k$th elementary symmetric function of singular values of matrix $A \in \mathbb{R}^{n \times n}$:

$$s_k(A) := \sum_{1 \leq i_1 \leq \cdots \leq i_k \leq n} \sigma_{i_1} \cdots \sigma_{i_k}, 1 \leq k \leq n. \quad (37)$$

**Theorem 4:** In noised situation, under $H_0$ hypothesis, the multi-dimensional observations is of the form of

$$H_0 : \mathbf{r}^{(i)} = X_I \mathbf{\beta}^{(i)} + \mathbf{w}^{(i)}, \quad (38)$$

where $X_I \in \mathbb{R}^{P \times p}, \mathbf{\beta}^{(i)} \in \mathbb{R}^p$ are the basis and coefficient of the interference signal subspace; and $\mathbf{w}^{(i)} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_P)$ is the white Gaussian noise. Then the volume-correlation of Detector 1 will satisfy:

Asymptotically (for large $m$), for any $0 < \varepsilon < 1$ and $\delta > 0$, if

$$m \geq \frac{1 + \varepsilon}{(\sqrt{\delta} + 1 - 1)^2} \left( \frac{1}{P} \sum_{j \neq i}^{p} \frac{\lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} + \frac{1}{P} \sum_{i=1}^{p} (P - p) \frac{\lambda_i \sigma^2}{(\sigma^2 - \lambda_i)^2} \right), \quad (39)$$

then there exists a constant $C > 0$, such that

$$|t^2(R^{(m)}) - \tau^2(X_s, X_I)| \leq s_{p-1}(Q_{X_s}^T P_{X_s}^\perp Q_{X_s}) \delta + O(\delta^2), \quad (40)$$

holds with probability

$$P \geq 1 - \exp\left\{ - \frac{p \cdot P \cdot \varepsilon^2}{C} \right\}. \quad (41)$$

Also, $\lambda_1, \lambda_2, \cdots, \lambda_r \geq \sigma^2$ are the corresponding eigenvalues of the auto-covariance matrix $C$ shown in (32), and $Q_{X_s}$ denotes the orthogonal basis matrix of subspace $X_s$. In addition, $s_k(A)$ is defined as the $k$th elementary symmetric function of singular values of matrix $A$ in (37).

Theorem 3 and 4 describe the asymptotic performance of our Volume-Correlation Subspace Detector in noised situation. The main result is shown in (35) and (40), together with (34) and (39), it implies that for a large enough number of observations $m$, the volume-correlation subspace detector will converge to its ideal performance without noise asymptotically, which also means the asymptotic elimination of influence of noise and interference. And this is the main advantage of our Volume-Correlation Subspace Detector.

It is noted that the asymptotic results in (35) and (40) are of the probabilistic form, i.e., they hold with a probability (36) and (41). The reason why the result is probabilistic is for the randomness of noise, and the probabilistic result also means that the asymptotic approximation of the volume-correlation is in the statistical sense, i.e., for large number of observations $m$, (35) and (40) will hold statistically.

The condition about the bound of the number of observations $m$ in (34) and (39) is a sufficient condition, it means that for a given parameter $\delta$, when $m$ satisfies (34) and (39), then the volume-correlation will sufficiently satisfy (35) and (40) with a high probability. As we know that sufficient conditions are usually more conservative than reality, in practice the value of volume-correlation would converge much more faster than what is predicted.
in the condition (34) and (39), and this will also be shown in the later numerical simulations as a complementary demonstration.

The bound on the number of observations in (34) and (39) is influenced by three main factors: the probability parameter $\varepsilon$, the error toleration parameter of volume-correlation $\delta$, as well as the eigenvalues of the auto-correlation matrix of the received signal:

$$H_0 : \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq \lambda_{p+1} = \cdots = \lambda_P = \sigma^2,$$

$$H_1 : \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{d+p} \geq \lambda_{d+p+1} = \cdots = \lambda_P = \sigma^2.\quad (42)$$

For smaller error toleration parameter $\delta$, we can see from (34) and (39) that only with more number of observations $m$ can the volume-correlation satisfy (35) and (40). And similar analysis holds for the probability parameter $\varepsilon$. As for the eigenvalues $\lambda_1, \lambda_2, \cdots$, the most noticeable ones among them that may affect the detector’s performance are those that equals to the noise power $\sigma^2$. A larger value of $\sigma^2$ will make the right side of (34) and (39) to be larger, meaning that only with more observations can the detector reach an ideal performance. Besides, the other eigenvalues $\lambda_1, \lambda_2, \cdots$, are determined by the basis $X_s$ (and $X_l$) as well as the statistical property of $\alpha$(and $\beta$), the analysis of their influence on the detector is not clear, and will be left for future work.

The main results (35) and (40) of theorem 4 and 3 are also influenced by two main factors: firstly the deviation parameter $\delta$, it is related to $m$ as is shown in (34) and (39) and analyzed previously. The other important parameters are the elementary symmetry functions of singular values $s_p(Q_{X_s}^T P_{X_s}^+ Q_{X_s})$(and $s_{p-1}(Q_{X_l}^T P_{X_s}^+ Q_{X_l})$). They are the coefficient of the first term of $\delta$ in (35) and (40), and according to their definition in (37), their values are decided only by the relation between subspaces $X_s$ and $X_l$, and are finite-valued, therefore, for a given number of observations $m$, the deviation of the volume-correlation value from its ideal value is decided mainly by the order of $\delta$ in (35) and (40).

The numerical simulation of detector 1 is demonstrated in figure 4 and 5. In the simulation, we we also choose $P = 1024$, and $d = 10, p = 40$, the target and interference signal subspaces are chosen arbitrarily. The average values of 100 monte-carlo simulations of the volume-correlation $t(R^{(m)})$ with respect to different $m$ are plotted in figure 4 and 5 and each values of these 100 monte-carlo simulations are demonstrated by scatter diagram in the small sub-figures. It can be seen from the figures that in the noised situation, as the number of observations $m$ increases, the value of the volume-correlation descends statistically, but slower than the ideal volume-correlation without noise; in addition, the values of the volume-correlation statistically concentrate around their average values, and these average values converge to the ideal values of the volume-correlation without noise, i.e., as $m$ increases, the average of $t(R^{(m)})$ converges to 0 under $H_1$ hypothesis, while under $H_0$ hypothesis, $t(R^{(m)})$ converges to $\tau(X_s, X_l)$. Therefore, as a whole, the simulation validates the result in theorem 3 and 4.

As is mentioned in the previous section, a manual constant $c > 1$ is need to fine tuning the threshold of detector 1, the reason can be obvious as shown in figure 5 if $1/\tau(X_s, X_l)$ is directly used as the threshold for detection, the false alarm rate will be high, therefore, we introduce the constant $c > 1$ to slightly raise the threshold in order to get a favorable false alarm rate. Simulation shows that a value of $c$ slightly greater than 1 (e.g., $c = 1.1$) is enough.
for a satisfactory detection performance.

As a whole, we can conclude that, as the increase and accumulation of observations, the Volume-Correlation Subspace Detector will asymptotically converge to its ideal performance without noise, and eliminate the influence of interference; the accumulation of observations is the price to pay for an ideal detection performance.

![Plot of $t(Y^{(m)})$ and $t(R^{(m)})$ with respect to $m$ (d=10, p=40)](image)

Fig. 4: simulation of the volume-correlation $t(R^{(m)})$ in noised situation

VI. APPLICATION OF VOLUME-CORRELATION SUBSPACE DETECTOR IN MULTI-USER DETECTION

In this section, we will demonstrate the application of our Volume-Correlation Subspace Detector to the process of Active User Identification in multiuser detection systems with multipath channels. And we will demonstrate that our Volume-Correlation Subspace Detector has its advantage in dealing with other users’ subspace interference in the multipath channel.

A. Active User Identification in multipath channels

Consider a typical multiuser detection system with $N$ users, and each user transmitting the Direct sequence CDMA signal (DS-CDMA)[43], then the base band signature waveform of the $n$’th user is:

$$x_n(t) = A_n \sum_{l=1}^{L} c_{n,l} p(t - lT_c), \quad t \in [0, T_s),$$

(44)
where
\[
\mathbf{c}_n = [c_{n,1}, c_{n,2}, \ldots, c_{n,L}]^T \in \mathbb{R}^L, \quad n = 1, \ldots, N,
\]  
(45)
is the \(n\)’th user’s signature sequence with length \(L\), typically the length \(L\) can be smaller than the number of users \(N\). The \(p(t)\) is the chip waveform, which satisfies \(\int_0^{T_c} |p(t)|^2 dt = 1\), \(\int_0^{T_c} p(t-lT_c)p(t-kT_c)dt = 0, l \neq k\); and \(T_c\) is the chip duration, \(T_s\) denotes the symbol duration, \(T_c\) and \(T_s\) satisfy \(LT_c = T_s\); \(A_n\) is the transmit amplitude of the \(n\)’th user.

In wireless communication, the multipath effect of wireless channels is a common physical phenomenon that affects the communication quality, the impulse response of the \(n\)’th user’s multipath channel is generally formulated as a number of delayed impulses:
\[
h_n(t) = \sum_{s=1}^{k_n} h_{n,s} \delta(t - T_{n,s}), \quad 1 \leq n \leq N,
\]  
(46)
where \(k_n\) is the maximum number of paths for the \(n\)’th user, \(h_{n,s}\) is the amplitude of the \(s\)’th path’s channel response for the \(n\)’th user, and \(T_{n,s}\) denotes the channel delay of the \(s\)’th path for the \(n\)’th user. The channel response \(h_n(t)\) will gradually vary with time due to different reasons, e.g., the mobility of each user and channel fading. Here we focus on the scenario of slow fading, i.e., the channel response \(h_n(t)\) remains roughly constant in
one symbol duration $T_s$, but will evolve through several symbol durations.

In order to eliminate the inter-symbol interference, the signature sequence for each user is usually designed to be a pseudo-random sequence (denoted by $s_n$) with a cyclic prefix. Thus, the cyclic prefix technique is also used in this paper, if we denote the discrete path delay by $\tau_{n,s} \triangleq [T_{n,s}/T_c] \in \mathbb{Z}_+$, and the maximum discrete delay by $\tau \triangleq \max_{n,s}\{\tau_{n,s}\}$, with the assumption $1 < \tau < L$, then the transmitted signature sequence for the $n$'th user is designed by appending the last $\tau$ element of the original signature sequence $s_n$ to the head of $s_n$, i.e.,

$$
c_{n,l} = \begin{cases} 
s_{n,P-\tau+l}, & 1 \leq l \leq \tau, \\
s_{n,l}, & \tau + 1 \leq l \leq L 
\end{cases}$$

(47)

where $s_n := [s_{n,1}, \ldots, s_{n,L}]$ is the original pseudo-random sequence assigned for the $n$th user. As a result, any length $P = L - \tau$ sub-sequence of $c_n$ will be a cyclic shift version of $s_n$. Then the signature waveform of the $n$th user in (44) is not affected by ISI in the time interval $t \in (\tau T_c, T)$. For convenience, we denote the $j$th cyclic shift operator by $T_j$, i.e.,

$$
T_j s_n = [s_{n,P-j+1}, \ldots, s_{n,P}, s_{n,1}, \ldots, s_{n,P-j}], \quad 1 \leq j \leq \tau
$$

(48)

The active users transmit signals by modulating their signature waveform (44) with symbols $b_n$, for simplicity, we assume the modulation method is BPSK, i.e., $b_n \in \{-1, 1\}$. The set of active users is denoted by $\mathcal{I} \subset \{1, 2, \ldots, N\}$, with $|\mathcal{I}| \leq N$, then the received signal will be:

$$
y(t) = \sum_{n \in \mathcal{I}} b_n \sum_{s=1}^{k} h_{n,s} x_n(t - T_{n,s}) + w(t), \quad t \in [0, T_s),
$$

(49)

where $h_{n,s}$ and $T_{n,s}$ are the channel response amplitude and path delay of the $s$th path for the $n$th user, respectively, and $w(t)$ is the white Gaussian noise. Before the demodulation and decoding of each user’s symbol $b_n$ from the received signal $y(t)$, a preprocess is usually essential to identify the set of active users, i.e., identify the set $\mathcal{I}$. The reason for this preprocess is that generally there are far less active users than the total $N$ users, identification of the active users will reduce the system complexity, and improve the overall decoding performance [44][45]. Therefore in this section, the goal is to identify the set $\mathcal{I}$ from the received signal $y(t)$.

Fig. 6: Diagram of chip rate sampling front-end

At the receiver, the chip rate sampling front-end in figure 6 will yield a discrete sequence of samples of the receive signal, i.e., $y_1, y_2, \ldots, y_L$, and

$$
y_l = \int_0^{T_c} p(t - lT_c) y(t) dt, \quad l = 1, \ldots, L,
$$

(50)
as is shown in figure 7, in order to eliminate ISI, the last $P = L - \tau$ samples of $y_1, y_2, \cdots, y_L$, is taken as the received signal sample, then the received signal vector is

$$ r = \sum_{n \in I} A_n b_n x_n + n \quad (51) $$

where $r := [y_{\tau+1}, \cdots, y_L]$, and $x_n = S_n a_n$, with

$$ S_n = [s_n, T_1 s_n, \cdots, T_{\tau} s_n] \in \mathbb{C}^{P \times (\tau+1)}, \quad (52) $$

$$ a_n = [0, \cdots, h_{n,1}, 0, \cdots, h_{n,k}, \cdots, 0]^T \in \mathbb{C}^{\tau+1}. \quad (53) $$

The coefficient vector $a_n$ implies information of the $n$th user’s channel state, i.e.,

$$ a_{n,j} = h_{n,s}, \quad j = \tau_{n,s}, \quad (54) $$

$$ a_{n,j} = 0, \quad j \neq \tau_{n,s}. $$

![Fig. 7: Eliminate the Intersymbol interference using cyclic prefix](image)

As a whole, the received signal from each user becomes a linear subspace signal due to the effect of multipath channel, i.e., $x_n = S_n a_n$. Each user’s signature subspace $S_n := \text{span}(S_n)$ is spanned by the cyclic shift version of the signature sequence $s_n$, and the subspace coefficient $a_n$ is determined by each user’s channel response. As is mentioned previously, the channel fading effect makes $a_n$ to be unknown at the receiver, therefore our goal here is to detect the target subspace signal $x_n = S_n a_n$ only with the knowledge of $S_n$.

It is obvious that our Volume-Correlation Subspace Detector has its advantage in dealing with this Active User Identification problem. For the $n$th user, its signature subspace $S_n = \text{span}(S_n)$ will be the target signal subspace $\mathcal{X}_c$, with dimension $d := \tau + 1$; and the other users’ signature subspace $\bigoplus_{m \neq n} S_m$ will become the interference signal subspace $\mathcal{X}_I$, with dimension $p := (u-1)(\tau+1)$. As the Volume-Correlation Subspace Detector will asymptotically eliminate noise and interference using multi-dimensional observations $R^{(m)} = [r^{(1)}, \cdots, r^{(m)}]$, these observations will be chosen from different received signal vectors in different channel coherence intervals, i.e.,

$$ r^{(i)} = \sum_{n \in I} A_n^{(i)} b_n^{(i)} S_n a_n^{(i)} + n^{(i)}, \quad i = 1, 2, \cdots \quad (55) $$

as is shown in figure 8 where the coherence interval $T_{coh}$ is the period that the channel response, i.e., the coefficient $a_n$, remains invariant.
Fig. 8: Multi-dimensional observations of the received signal from different channel coherence intervals

The detection performance of the proposed volume-correlation subspace detector in this application of Active User Identification is demonstrated in the next section.

B. Numerical Simulations

In this section, the performance of our Volume-Correlation Subspace Detector in the application of Active User Identification is demonstrated by monte-carlo simulation. In the simulation, the maximum number of active users is assumed to be 5, i.e., $|\mathcal{Z}| \leq 5$; and the user signature sequence $s_n$ is generated randomly, with values taking $\pm 1$ and $P = 1024$; the number of propagation paths for each user is assumed to be 10, which means the subspace basis and coefficient in (52) and (53) satisfies $S_n \in \mathbb{R}^{1024 \times 10}$ and $a_n \in \mathbb{R}^{10}$; the channel is assumed to obey the slow fading Rayleigh model, therefore each element in $a_n$ is generated randomly from the Rayleigh distribution; and the amplitude of each user’s transmit signal is assumed to be equal. For simplicity, in the simulation we mainly test the performance of our detector on one specific user, and therefore we take one arbitrary user’s signature subspace as the target subspace, and the other 4 active users’ signature subspaces as the interference subspace. We simulate two scenarios when the target user is active or inactive, while in both scenarios the other 4 interference users are assumed active. The detection and false alarm probability are calculated from 1000 monte-carlo simulations for different numbers of observations $m$ and Signal-to-Noise Ratio (SNR). The SNR here is the ratio of signal power to noise power for each user, i.e., $SNR = 10 \log \left( \frac{\| s_n \|^2}{\sigma^2} \right)$. The threshold parameter $c$ is chosen to be 1.1, which ensures an ideal false alarm probability.

1) **Given the number of observations $m$, the performance of the Volume-Correlation Subspace Detector at different level of SNR:** The monte-carlo simulation of the detection probability and false alarm probability when SNR ranges from 0 to 30dB with $m = 5, 10, 40, 50$ is demonstrated in figure 9.

It can be seen in figure 9 that, even when there are only several number of observations, such as $m = 5$ shown in figure 9 our volume-correlation subspace detector still have a good performance for high SNR; and as the number of observations increases, the performance of our volume-correlation subspace detector increases significantly.
Fig. 9: The Mote-Carlo simulation of Detection probability ($P_d$) and the False Alarm probability ($P_f$) for different SNR with $m = 5, 10, 40, 50$

2) Given the SNR, the performance of the Volume-Correlation Subspace Detector for increase number of observations $m$: The monte-carlo simulation of the detection probability and false alarm probability when $m$ ranges from 1 to 50 with SNR = 2.1, 4.3, 8.3, 12.2 is demonstrated in figure 10.

Figure 10 clearly demonstrated the increase of detection probability with increase of number of observations $m$ at different SNR levels. It can be seen that even in a low SNR scenario, such as SNR=2.1dB, by accumulating as many observations, such as $m \geq 50$, our detector will still ensure an ideal detection probability.

C. Further Discussion & Future Work

1) Flexibility about the dimensions of subspaces: As is discussed in the previous sections, in the calculation of volume-correlation in (21) and (22), it is assumed that the target signal subspace $X_s$ and interference signal subspace $X_I$ are known at the receiver, including their dimensions $\dim(X_s) = d$ and $\dim(X_I) = p$. However, in real systems, there is also possibility that the signal and interference subspaces cannot be fully understood at the
Fig. 10: The Monte-Carlo simulation of Detection probability ($P_d$) and the False Alarm probability ($P_f$) with increase of $m$ from 1 to 50, with $SNR = 2.1, 4.3, 8.3, 12.2$

receiver, for example, the only knowledge could be some “larger” subspaces that the real subspaces $X_s$ and $X_I$ lie within. In other words, the real target and signal subspaces, i.e., $X_s$ and $X_I$ satisfy

$$X_s \subset \text{span}(X_s), \quad X_s \in \mathbb{R}^{P \times d}, \quad \text{dim}(X_s) = d_s \leq d;$$

(56)

$$X_I \subset \text{span}(X_I), \quad X_I \in \mathbb{R}^{P \times p}, \quad \text{dim}(X_I) = p_I \leq p.$$  

(57)

The $X_s$ and $X_I$ here are the priori knowledge of the target and interference subspace, and they are used in (21) and (22) to calculate the volume-correlation and perform detection; while the subspaces $X_s$ and $X_I$ are the real target subspaces and interference subspace in the received signal.

The example can be found in the previously mentioned Multiuser Detection system. (51) indicates that each user’s signature subspace is spanned by its signature sequence $s_n$ and its cyclic shift versions, and the subspace coefficient $a_n$ is determined by the multipath channel response $h_{n,s}$ as in (54). From (54), it is obvious that generally the channel can not have paths with all possible delays, which means some elements in $a_n$ will remains zero, and thus
each user’s real signature subspace in the received signal is contained in (but does not equal to) span($S_n$), then we have for the nth user

$$\mathcal{X}_s \subset \text{span}(S_n), \quad \text{with dim}(\mathcal{X}_s) = k_n \leq \text{dim}(S_n) = \tau + 1;$$

$$\mathcal{X}_I \subset \bigoplus_{m \neq n} S_m = \bigoplus_{m \neq n} \text{span}(S_m), \quad \text{with dim}(\mathcal{X}_I) = \sum_{m \in I, m \neq n} k_m \leq \text{dim}(\bigoplus_{m \neq n} S_m) = (u - 1)(\tau + 1),$$

where $\mathcal{X}_s$ is the real subspace of user $n$, and $\mathcal{X}_I$ is the real subspace of the other active users.

Surprisingly, the volume-correlation subspace detector can still be reliable. And the reason is also quite obvious, and can be proved by simply redo the proof procedures of Theorem 1 and 2. In summary, for this scenario, the volume-correlation for the $n$th user in (21) will become

$$t(Y^{(n)}) := \frac{\text{vol}_{\tau+1}([S_n, Y^{(m)}])}{\text{vol}_{\tau+1}(S_n) \text{vol}_r(\text{Y}^{(m)}))},$$

It can be proved that when $m \leq \text{dim}(\mathcal{X}_I) = \sum_{m \neq n} k_m$, the volume correlation $t(Y^{(m)})$ descends with the increase of $m$; and when $m \geq \sum_{m \neq n} k_m$, the volume-correlation descends to 0 under $H_1$ hypothesis, and also satisfies inequality (29) under $H_0$ hypothesis, except that the equality may be unreachable. Besides, similar results to Theorem 3 and 4 can also be derived. This advantage of the volume-correlation subspace detector is demonstrated by simulation in figure 11.

![Plot of $t(Y^{(m)})$ with respect to $m$ (d_s=5, p_I=20)](image1)

![Plot of $t(R^{(m)})$ with respect to $m$ (d_s=5, p_I=20)](image2)

Fig. 11: Output of the volume-correlation without and with noise in Active User Identification, when each user’s real subspace is "smaller" than the known signature subspace span($S_n$), and assume $k_n = 5$

In the simulation, the same settings are used as the previous simulations of active user identification, except that actually each user only have $k_n = 5$ channel paths, and the volume-correlation in (21) and (22) is calculated using the priori-known signature subspace in (52) which assumes channel paths with all delays. The simulation clearly shows that in noiseless situation the volume-correlation $t(Y^{(m)})$ descends to 0 under $H_1$ hypothesis when $m \geq 20,$
but satisfies (29) except the equality under $H_0$ hypothesis. Similar asymptotic results can also be seen for noised situations.

The above analysis and simulation indicates that, the volume-correlation subspace detector allows for significant flexibility about the priori knowledge of the target and interference signal subspace, and this flexibility is also an interior advantage of our detector, and this will be left for our future work.

2) The Phase Transition Phenomenon of random matrices and Future works: When the noise level is significantly high, i.e., the SNR is extremely low, analyzes have shown that there will be significant errors in the step of estimation of signal subspaces using the eigendecomposition subspace methods. The reason is for the so-called Phase Transition Phenomenon about eigenvalues and eigenvectors of random matrices [46][47][48]. Analysis has shown that as $P \to \infty$, and $P/m \to \gamma \in (0,1)$, then we have the following results:

(a) If $\lambda_i / \sigma^2 > 1 + \sqrt{\gamma}$, then
\[
\hat{\lambda}_i^{(m)} \overset{a.s.}{\to} \lambda_i (1 + \frac{\gamma \sigma^2}{\lambda_i - \sigma^2}), \quad i = 1, \ldots, d, \text{or } d + p
\] (58)
\[
|\langle \hat{u}_i^{(m)}(\mathbf{m}), u_i \rangle| \overset{a.s.}{\to} \sqrt{\frac{1}{\lambda_i - \sigma^2}} \left(1 + \frac{\gamma \sigma^4}{(\lambda_i - \sigma^2)^2}\right), \quad i = 1, \ldots, d, \text{or } d + p.
\] (59)

(b) If $\lambda_i / \sigma^2 \leq 1 + \sqrt{\gamma}$, then
\[
\hat{\lambda}_i^{(m)} \overset{a.s.}{\to} \sigma^2 (1 + \sqrt{\gamma})^2, \quad i = 1, \ldots, d, \text{or } d + p
\] (60)
\[
|\langle \hat{u}_i^{(m)}(\mathbf{m}), u_i \rangle| \overset{a.s.}{\to} 0, \quad i = 1, \ldots, d, \text{or } d + p
\] (61)

where $\lambda_i$ and $u_i$ are eigenvalues and eigenvectors of the auto-correlation matrix of the received signal, and $\hat{\lambda}_i^{(m)}$ and $\hat{u}_i^{(m)}$ are the corresponding estimation of $\lambda_i$ and $u_i$ calculated from eigenvalue decomposition of the sampled covariance matrix. This phase transition phenomenon means that when $\lambda_i / \sigma^2$ is lower than a threshold, the estimated eigenvalues $\hat{\lambda}_i^{(m)}$ and eigenvectors $\hat{u}_i^{(m)}$ will resembles the eigenvalues and eigenvectors of the noise subspace almost surely. On the other hand, if the noise variance $\sigma^2$ is significantly high, such that $\lambda_i / \sigma^2$ is lower than $1 + \sqrt{\gamma}$, then the estimated signal subspace $\text{span}(\hat{U}_s^{(m)})$ will not asymptotically converge to the real signal subspace $\text{span}(U_s)$, therefore the volume-correlation will not converge to 0 under $H_1$ hypothesis, or converge to $\tau(X_s, X_I)$ under $H_0$ hypothesis. As a whole, the signal subspace estimation method in Detector 1 has its shortage in low SNR scenarios, so the estimation of signal subspace in a low SNR scenario remains an important future work for our volume-correlation subspace detector.

VII. CONCLUSION

In this paper, we proposed a novel volume-correlation subspace detector, and use it to deal with the problem of subspace signal detection with noise and interference. Subspace signal detection is a common problem in the signal processing community, it arises in a wide range of applications such as communication, radar, array signal processing and hyperspectral imagery. Our proposed volume-correlation subspace detector detects subspace signals by calculating the volume of parallelepiped spanned by the basis of target signal subspace together with the multi-dimensional observations of the received signal. It is proved that by using the multi-dimensional knowledge of the
target and interference signal subspace, as well as multi-dimensional observations, our detector can asymptotically eliminate the influence of noise and interference. Besides, we also discussed the application of the volume-correlation subspace detector to the problem of Active User Identification in multipath channels commonly seen in the multiuser detection systems. Numerical simulations were given to demonstrate the advantage and performance of our volume-correlation subspace detector. As a whole, our volume-correlation subspace detector has its interior advantage in dealing with multi-dimensional subspace detection, and can converge asymptotically to its ideal performance with the accumulation of observations.

APPENDIX A

APPENDIX A. PROOF OF THEOREM 1

Proof:

According to the system model, under $H_1$ hypothesis, if $m \leq d+p$, then $\text{rank}(Y^{(m)}) = m$, the volume-correlation in noiseless situation is:

$$t(Y^{(m)}) = \frac{\text{vol}_{d+m}([X_s, Y^{(m)}])}{\text{vol}_d(X_s) \cdot \text{vol}_m(Y^{(m)})},$$

According to the analysis in [31], we have:

$$\frac{\text{vol}_{d+m}([X_s, Y^{(m)}])}{\text{vol}_d(X_s) \cdot \text{vol}_m(Y^{(m)})} = \prod_{j=1}^{\min(d,m)} \sin \theta_j(X_s, \text{span}(Y^{(m)})), \quad (62)$$

where $0 \leq \theta_j(X_s, \text{span}(Y^{(m)})) \leq 2\pi, 1 \leq j \leq \min(d,m)$ are the principal angles of subspaces $X_s$ and $\text{span}(Y^{(m)})$.

Firstly we are going to prove (25), as is known, for $1 \leq m \leq p$, we have

$$t(Y^{(m)}) = \prod_{j=1}^{\min(d,m)} \sin \theta_j(X_s, \text{span}(Y^{(m)})). \quad (63)$$

$t(Y^{(m)})$ is the product of the sines of principal angles between subspaces $X_s$ and $\text{span}(Y^{(m)})$, thus we can replace the matrices $X_s$ and $Y^{(m)}$ by their orthogonal basis matrices in the volume-correlation expression (62), then equivalently, we have

$$t(Y^{(m)}) = \text{vol}_{d+m}([Q_{X_s}, Q_{Y^{(m)}}]), \quad (64)$$

where $Q_{X_s} \in \mathbb{R}^{P \times d}$ and $Q_{Y^{(m)}} \in \mathbb{R}^{P \times m}$ are the orthogonal bases of subspaces $X_s$ and $\text{span}(Y^{(m)})$, i.e.,

$$\text{span}(Q_{X_s}) = X_s = \text{span}(X_s), \quad Q_{X_s}^T Q_{X_s} = I, \quad \text{span}(Q_{Y^{(m)}}) = \text{span}(Y^{(m)}), \quad Q_{Y^{(m)}}^T Q_{Y^{(m)}} = I. \quad (65)$$

In the same way, for $m + 1$, we have

$$t(Y^{(m+1)}) = \text{vol}_{d+m+1}([Q_{X_s}, Q_{Y^{(m+1)}}]), \quad (67)$$

$Q_{Y^{(m+1)}}$ can be related to $Q_{Y^{(m)}}$, indeed, $Q_{Y^{(m+1)}}$ can be written as

$$Q_{Y^{(m+1)}} = [Q_{Y^{(m)}}, q_{Y^{(m+1)}}], \quad (68)$$
where
\[ q_{y^{(m+1)}} := P_{Y^{(m)}} y^{(m+1)} / \|P_{Y^{(m)}} y^{(m+1)}\|_2, \] 
(69)
and \( P_{Y^{(m)}} \) denotes the orthogonal projection operator onto subspace \( \text{span}(Y^{(m)}) \), therefore \( q_{y^{(m+1)}} \) is the component of \( y^{(m+1)} \) that is perpendicular to the subspace \( \text{span}(Y^{(m)}) \).

Using knowledge of the determinant of block matrices, we have:
\[
t(Y^{(m+1)}) = \text{vol}_{d+m+1}([Q_{X_s}, Q_{Y^{(m+1)}}])
= \det \left( \begin{bmatrix} Q_{X_s}^T & Q_{Y^{(m+1)}_{s}} \end{bmatrix} [Q_{X_s}, Q_{Y^{(m+1)}}] \right)^{1/2}
= \det \left( \begin{bmatrix} Q_{X_s}^T & Q_{X_s}^T & Q_{Y^{(m+1)}_{s}} & Q_{Y^{(m+1)}_{Y}} \end{bmatrix} \right)^{1/2}
= \det \left( \begin{bmatrix} Q_{X_s}^T & Q_{X_s}^T & Q_{Y^{(m+1)}_{s}} & Q_{Y^{(m+1)}_{Y}} \\ Q_{Y^{(m+1)}_{s}}^T & Q_{X_s}^T & Q_{Y^{(m+1)}_{Y}} & Q_{Y^{(m+1)}_{Y}} \end{bmatrix} \right)^{1/2}
\cdot \det \left( q_{y^{(m+1)}}^T q_{y^{(m+1)}} - q_{y^{(m+1)}}^T [Q_{X_s}, Q_{Y^{(m)}}] Q_{X_s}^T Q_{X_s}^T Q_{Y^{(m)}} \right)^{-1}
\cdot [Q_{X_s}^T, Q_{Y^{(m)}}, q_{y^{(m+1)}}]^T
= \det \left( \begin{bmatrix} Q_{X_s}^T & Q_{X_s}^T & Q_{Y^{(m)}_{s}} & Q_{Y^{(m)}_{Y}} \end{bmatrix} \right)^{1/2}
\cdot \|P_{[X_s, Y^{(m)}]} q_{y^{(m+1)}}\|_2
= t(Y^{(m)}) \cdot \|P_{[X_s, Y^{(m)}]} q_{y^{(m+1)}}\|_2, \] 
(71)
where \( P_{[X_s, Y^{(m)}]} \) is the orthogonal projection operator onto the subspace \( \text{span}([X_s, Y^{(m)})] \). Because
\[
\|P_{[X_s, Y^{(m)}]} q_{y^{(m+1)}}\|_2 \leq \|q_{y^{(m+1)}}\|_2 = 1,
\] 
(72)
therefore
\[
t(Y^{(m)}) \geq t(Y^{(m+1)})
\] holds, and then (25) is proved.

Next we prove (26), as is known, the \( i \)th observation of the received signal is
\[ y^{(i)} = X_s \alpha^{(i)} + X_I \beta^{(i)}, \quad i = 1, 2, \ldots, m. \]
then from the subspace perspective, it is known that
\[
\text{span}(Y^{(m)}) = \text{span}([y^{(1)}, \ldots, y^{(m)}]) \subset \text{span}(X_s) \bigoplus \text{span}(X_I),
\] 
(73)
When \( m = p + 1 \), we have
\[
y^{(p+1)} = X_s \alpha^{(p+1)} + X_i \beta^{(p+1)},
\] (74)
it can be seen that the second term in (74) satisfies \( X_i \beta^{(p+1)} \in \text{span}(X_i) \), and all the previous observations, i.e.,
\[
y^{(i)} = X_s \alpha^{(i)} + X_i \beta^{(i)}, \quad i = 1, \cdots, p,
\] (75)
will satisfy
\[
\text{span}(X_i \beta^{(1)}, \cdots, X_i \beta^{(p)}) = \text{span}(X_i).
\] (76)
Therefore, there must exist a sequence of coefficients \( b_j, j = 1, \cdots, p \) that are not all zero, such that
\[
X_i \beta^{(p+1)} = \sum_{j=1}^{p} b_j X_i \beta^{(j)},
\] (77)
then \( y^{(p+1)} \) can be written as
\[
y^{(p+1)} = X_i \beta^{(p+1)} + X_s \alpha^{(p+1)}
\]
\[
= \sum_{j=1}^{p} b_j X_i \beta^{(j)} + X_s \alpha^{(p+1)}
\]
\[
= \sum_{j=1}^{p} b_j (X_s \alpha^{(j)} + X_i \beta^{(j)}) + X_s \alpha^{(p+1)} - \sum_{j=1}^{p} b_j X_s \alpha^{(j)}
\]
\[
= \sum_{j=1}^{p} b_j y^{(j)} + X_s \alpha^{(p+1)} - \sum_{j=1}^{p} b_j \alpha^{(j)}.
\] (78)
So it is obvious that
\[
y^{(p+1)} \in \text{span}([Y^{(p)}; X_s]),
\] (79)
where \( Y^{(p)} = [y^{(1)}, \cdots, y^{(p)}] \). As a result,
\[
t(Y^{(p+1)}) = \frac{\text{vol}_{p+d+1}([X_s, Y^{(p+1)}])}{\text{vol}_d(X_s) \text{vol}_{p+1}(Y^{(p+1)})},
\]
\[
= \frac{\text{vol}_{p+d+1}([X_s, Y^{(p)}, y^{(p+1)}])}{\text{vol}_d(X_s) \text{vol}_{p+1}(Y^{(p+1)})},
\]
\[
= 0.
\] (80)
The same result can be derived for all \( m > p + 1 \), therefore, (26) is proved.

APPENDIX B

APPENDIX B. PROOF OF THEOREM 3

Proof:

For simplicity, we assume \( d \leq p \) in the following proof. Firstly, similar to the proof of theorem 1, it is easy to prove.

Therefore, for \( 1 \leq m \leq p \), we have
\[
t(Y^{(1)}) \geq \cdots \geq t(Y^{(d)}) \geq \cdots t(Y^{(p)}),
\]
and for \( d \leq m \leq p \), we have

\[
\ell(Y^{(m)}) = \prod_{j=1}^{d} \sin \theta_j(\mathcal{X}_s, \text{span}(Y^{(m)})),
\]

so our focus is the relation between \( \prod_{j=1}^{d} \sin \theta_j(\mathcal{X}_s, \text{span}(Y^{(m)})) \) and \( m \). Define the orthogonal bases matrices \( Q_{\mathcal{X}_s} \) and \( Q_{Y^{(m)}} \) which are the same as (65) and (66), then according to the definition of principal angles in [31], the cosines of the principal angles between subspaces satisfy

\[
\cos \theta_j(\mathcal{X}_s, \text{span}(Y^{(m)})) = \sigma_j(Q_{\mathcal{X}_s}^T Q_{Y^{(m)}}, 1 \leq j \leq d,
\]

where \( \sigma_j(Q_{\mathcal{X}_s}^T Q_{Y^{(m)}}) \) are the singular values of the matrix \( Q_{\mathcal{X}_s}^T Q_{Y^{(m)}} \), and is assumed to satisfy \( \sigma_1 \geq \cdots \geq \sigma_d \). Then we have for \( m \leq p \),

\[
\text{span}(Y^{(m)}) \subset \text{span}(X_I),
\]

Denote the orthogonal basis matrix of \( X_I \) by \( Q_{X_I} \), which satisfies

\[
\text{span}(Q_{X_I}) = X_I = \text{span}(X_I),
\]

then \( \text{span}(Q_{Y^{(m)}}) \subset \text{span}(Q_{X_I}) \), and there exists a column-orthogonal matrix \( U_{X_I}^{(m)} \in \mathbb{R}^{p \times m} \), with \( (U_{X_I}^{(m)})^T U_{X_I}^{(m)} = I \), such that

\[
Q_{Y^{(m)}} = Q_{X_I} U_{X_I}^{(m)}, \quad d \leq m \leq p.
\]

In addition, the subspace \( \text{span}(Q_{Y^{(m)}}) \) must have an orthogonal complement in subspace \( X_I \), i.e., there exists another column-orthogonal matrix \( V_{X_I}^{(m)} \in \mathbb{R}^{p \times (p-m)} \), such that

\[
\text{span}(Q_{Y^{(m)})} \perp \text{span}(Q_{X_I} V_{X_I}^{(m)}), \quad d \leq m \leq p.
\]

and

\[
U_{X_I} = [U_{X_I}^{(m)}, V_{X_I}^{(m)}] \in \mathbb{R}^{p \times p},
\]

satisfies \( U_{X_I}^T U_{X_I} = I_p \), in other words, we have

\[
U_{X_I}^{(m)} = U_{X_I} \cdot \begin{bmatrix} I_m \\ O \end{bmatrix},
\]

which means \( U_{X_I}^{(m)} \) is the sub-matrix of \( U_{X_I} \), and

\[
Q_{Y^{(m)}} = Q_{X_I} U_{X_I} \cdot \begin{bmatrix} I_m \\ O \end{bmatrix},
\]

So

\[
Q_{X_I}^T Q_{Y^{(m)}} = Q_{X_I}^T Q_{X_I} U_{X_I} \cdot \begin{bmatrix} I_m \\ O \end{bmatrix}.
\]
According to the relation between singular values and eigenvalues, we have
\[
\sigma_j^2(Q_{X_i}^T Q_{Y(m)}^T) = \lambda_j(Q_{Y(m)}^T Q_{X_i}^T Q_{X_i}^T Q_{Y(m)}^T)
\]
\[
= \lambda_j([I_m, 0] U_{X_i}^T \cdot Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T \cdot U_{X_i} \left[ I_m \ O \right] \cdot 1 \leq j \leq d,
\]
where \(\lambda_j(Q_{Y(m)}^T Q_{X_i}^T Q_{X_i}^T Q_{Y(m)}^T)\) are eigenvalues of the matrix \(Q_{Y(m)}^T Q_{X_i}^T Q_{X_i}^T Q_{Y(m)}^T\). Then \(Q_{Y(m)}^T Q_{X_i}^T Q_{X_i}^T Q_{Y(m)}^T\) is the principal sub-matrix of the matrix \(U_{X_i}^T \cdot Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T \cdot U_{X_i}\).

According to the Interlacing Theorem in [49], we have

**Lemma 1:** For a Hermitian matrix \(X \in \mathbb{R}^{N \times N}\), with its \(m \times m\) principal matrix \(X_m\), the eigenvalues of \(X\) and \(X_m\) satisfies:
\[
\lambda_{N-k+1}(X) \leq \lambda_{m-k+1}(X_m) \leq \lambda_{m-k+1}(X), \quad k = 1, \cdots, m,
\]
where \(\lambda_{N}(X) \leq \cdots \leq \lambda_{1}(X)\) and \(\lambda_{m}(X_m) \leq \cdots \leq \lambda_{1}(X_m)\) are eigenvalues of \(X\) and \(X_m\), respectively.

It has been mentioned that from (90) that \(Q_{Y(m)}^T Q_{X_i}^T Q_{X_i}^T Q_{Y(m)}^T\) is the principal sub-matrix of \(U_{X_i}^T \cdot Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T \cdot U_{X_i}\), besides, it is known that for orthogonal matrix \(U_{X_i}\),
\[
\lambda_j(U_{X_i}^T \cdot Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T U_{X_i}) = \lambda_j(Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T U_{X_i}) = \lambda_j(Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T U_{X_i}) = \sigma_j^2(Q_{X_i}^T Q_{X_i}^T U_{X_i}),
\]
therefore, combining (91) and (92), we have
\[
\lambda_{p-k+1}(Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T U_{X_i}) \leq \lambda_{m-k+1}(Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T U_{X_i}) \leq \lambda_{m-k+1}(Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T U_{X_i})
\]
holds for \(k = 1, \cdots, m\), with \(d \leq m \leq p\), then the right inequality implies that
\[
\lambda_k(Q_{Y(m)}^T Q_{X_i}^T Q_{X_i}^T Q_{Y(m)}^T) \leq \lambda_k(Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T Q_{X_i}^T U_{X_i}).
\]
then
\[
\cos \theta_j(X_s, \text{span}(Y(m))) \leq \cos \theta_j(X_s, X_i), \quad 1 \leq j \leq d.
\]
As a result,
\[
\prod_{j=1}^d \sin \theta_j(X_s, \text{span}(Y(m))) \geq \prod_{j=1}^d \sin \theta_j(X_s, X_i).
\]
So (25) is proved. And it is obvious that when \(m \geq p\), \(Q_{Y(m)} = Q_{X_i} U_{X_i}\), then the equality in (25) holds.

**APPENDIX C**

**APPENDIX C. PROOF OF THEOREM 3 AND THEOREM 4**

**Proof:**

Before the proof of Theorem 3 and Theorem 4 several lemmas are required as intermediate results. As is mentioned, the signal eigenvectors \(\hat{u}_1^{(m)}, \hat{u}_2^{(m)}, \cdots, \hat{u}_r^{(m)}\) calculated from eigenvalue decomposition of the sampled covariance matrix \(\hat{C}^{(m)}\) are estimations of the real signal eigenvectors \(u_1, u_2, \cdots, u_r\), a well-known result is:
Lemma 2: Consider the matrix whose columns are the $r$ largest eigenvectors estimated from of $\hat{C}^{(m)}$, i.e.,

$$\hat{U}_{\text{sig}}^{(m)} = [\hat{u}_1^{(m)}, \hat{u}_2^{(m)}, \ldots, \hat{u}_r^{(m)}],$$

its asymptotic distribution (for large $m$) is jointly Gaussian with mean

$$U_{\text{sig}} = [u_1, u_2, \ldots, u_r]$$

and covariance $\Sigma_1^{(m)}, \ldots, \Sigma_r^{(m)}$, where

$$\Sigma_i^{(m)} := \frac{\lambda_i}{m} \left[ \sum_{j=1}^{r} \frac{\lambda_j}{(\lambda_i - \lambda_j)^2} u_i u_i^T + \sum_{j=r+1}^{P} \frac{\sigma^2}{(\sigma^2 - \lambda_i)^2} u_i u_i^T \right], \quad i = 1, \ldots, r \tag{96}$$

and

$$\mathbb{E}(\hat{u}_i^{(m)} - u_i)(\hat{u}_k^{(m)} - u_k)^T = \Sigma_i^{(m)} \cdot \delta_{i,k}, \quad i, k = 1, \ldots, r \tag{97}$$

where $r$ denotes the dimension of the signal subspace, and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r \geq \lambda_{r+1} = \cdots = \lambda_P = \sigma^2$ are eigenvalues of the auto-correlation matrix in (32), with $u_1, \ldots, u_P$ the corresponding eigenvectors.

Another lemma derived from the well known concentration inequality is need in our proof:

Lemma 3: For the random matrix

$$E = [e_1, \ldots, e_r] \in \mathbb{R}^{P \times r}, \tag{98}$$

where $e_i \sim \mathcal{N}(0, \Sigma_i), 1 \leq i \leq d$, and $\mathbb{E}(e_i e_k^T) = \Sigma_i \cdot \delta_{i,k}$, then for any $0 < \varepsilon < 1$, there exists a constant $C > 0$ that depends on $\Sigma_i$, such that

$$\|E\|_F^2 \leq (1 + \varepsilon) \sum_{i=1}^{r} \text{Tr}(\Sigma_i), \tag{99}$$

holds with probability

$$\mathbb{P} \geq 1 - \exp\left\{-\frac{r \cdot \varepsilon^2}{C}\right\}. \tag{100}$$

Proof of Lemma 3:

From the definition of F-norm, we know that

$$\|E\|_F^2 = \sum_{i=1}^{r} \|e_i\|_2^2, \tag{101}$$

and for any $1 \leq i \leq r$, $e_i \sim \mathcal{N}(0, \Sigma_i)$, denote the eigenvalue decomposition of $\Sigma_i \in \mathbb{R}^{P \times P}$ by

$$\Sigma_i = V_i \Lambda_i V_i^T, \tag{102}$$

where the diagonal matrix $\Lambda_i := \text{diag}(\sigma_{i,1}^2, \ldots, \sigma_{i,P}^2)$ and $\sigma_{i,1}^2 \geq \sigma_{i,2}^2 \geq \cdots \geq \sigma_{i,P}^2 \geq 0$ are eigenvalues of $\Sigma_i$. If we let

$$\tilde{e}_i = V_i^T e_i, \tag{103}$$

then

$$\tilde{e}_i \sim \mathcal{N}(0, \Lambda_i), \quad \|\tilde{e}_i\|_2^2 = \|e_i\|_2^2. \tag{104}$$
Denote the elements of vector $\tilde{e}_i$ by

$$
\tilde{e}_i = [\tilde{e}_{i,1}, \cdots, \tilde{e}_{i,P}]^T,
$$

(105)

then different $\tilde{e}_{i,j}$ are independent and satisfy

$$
\tilde{e}_{i,j} \sim \mathcal{N}(0, \sigma_{i,j}^2), \quad 1 \leq j \leq P.
$$

(106)

Now we stack all these vectors $\tilde{e}_i, 1 \leq i \leq d$ into a single vector, i.e., we let

$$
\tilde{e} := [\tilde{e}_1^T, \tilde{e}_2^T, \cdots, \tilde{e}_P^T]^T \in \mathbb{R}^{r \cdot P},
$$

(107)

then we have

$$
\tilde{e} \sim \mathcal{N}(0, \Lambda) = \text{diag}(\Lambda_1, \cdots, \Lambda_r) = \text{diag}(\sigma_{1,1}^2, \cdots, \sigma_{1,P}^2, \cdots, \sigma_{r,1}^2, \cdots, \sigma_{r,P}^2).
$$

(108)

Therefore, (101) is equivalent to

$$
\|E\|_F^2 = \sum_{i=1}^r \|\tilde{e}_i\|_2^2 = \|\tilde{e}\|_2^2.
$$

(109)

As is known, the norm of a Gaussian random vector will concentrate around its expectation, which is referred to as the concentration of measure phenomenon [50]. It has been proved that the norm of an i.i.d. Gaussian random vector will concentrate around its own expectation (Chapter 4, [51]). The problem of the concentration of $\|\tilde{e}\|_2^2$ here only has a slight difference with the analysis from [51], i.e., the elements of $\tilde{e}$ have different variances, therefore, the proof will be quite similar with the proof of Theorem 4.2 in [51], and in the following proof we only give the part that are different from Theorem 4.2 in [51].

Firstly, we have

$$
\mathbb{E}\{\|\tilde{e}\|_2^2\} = \sum_{i=1}^r \sum_{j=1}^P \sigma_{i,j}^2 = \sum_{i=1}^r \text{Tr}(\Sigma_i),
$$

(110)

Then follow the same approach as Theorem 4.2 in [51], according to the Markov’s Inequality, for any parameter $\beta > 0$ and $\lambda > 0$, we have

$$
\mathbb{P}\{\|\tilde{e}\|_2^2 \geq \beta \sum_{i=1}^r \text{Tr}(\Sigma_i)\} = \mathbb{P}\{\exp(\lambda\|\tilde{e}\|_2^2) \geq \exp(\lambda\beta \sum_{i=1}^r \text{Tr}(\Sigma_i))\}
$$

$$
= \prod_{i=1}^r \mathbb{P}\{\exp(\lambda\|\tilde{e}_i\|_2^2) \geq \exp(\lambda\beta \text{Tr}(\Sigma_i))\}
$$

$$
\leq \prod_{i=1}^r \frac{\mathbb{E}\{\exp(\lambda\|\tilde{e}_i\|_2^2)\}}{\exp(\lambda\beta \text{Tr}(\Sigma_i))}
$$

$$
= \prod_{i=1}^r \prod_{j=1}^P \frac{\mathbb{E}\{\exp(\lambda\tilde{e}_{i,j}^2)\}}{\exp(\lambda\beta \sigma_{i,j}^2)},
$$

(111)

according to (106), the moment generating function of the Gaussian random variable $\tilde{e}_{i,j}$ is:

$$
\mathbb{E}\{\exp(\lambda\tilde{e}_{i,j}^2)\} = \frac{1}{\sqrt{1 - 2\lambda\sigma_{i,j}^2}},
$$

(112)
If we let

\[ \sigma_{\text{max}} := \max_{i,j} \sigma_{i,j}, \quad \sigma_{\text{min}} := \min_{i,j} \sigma_{i,j}, \]  

(113)

then we have

\[ \mathbb{P}\{\|\tilde{e}\|_2^2 \geq \beta \sum_{i=1}^{r} \text{Tr}(\Sigma_i)\} \leq \left( \frac{\exp(-2\lambda\beta \sigma_{\text{min}}^2)}{1 - 2\lambda \sigma_{\text{max}}^2} \right)^{-r/2} \]  

(114)

holds for any \( \lambda > 0 \). The next steps of proof is the same as Theorem 4.2, by replacing \( \lambda \) with its optimal value such that \( \left( \exp(-2\lambda\beta \sigma_{\text{min}}^2)/1 - 2\lambda \sigma_{\text{max}}^2 \right)^{-r/2} \) is minimized, and substituting some complicated formulas involving \( \sigma_{\text{max}} \) and \( \sigma_{\text{min}} \) for a constant \( C \), we can derive the result of this lemma (which is also the result of Corollary 4.1 in [51]):

\[ \mathbb{P}\{\|\tilde{e}\|_2^2 \geq (1 + \varepsilon) \sum_{i=1}^{r} \text{Tr}(\Sigma_i)\} \leq \exp(-\frac{r \cdot P \varepsilon^2}{C}), \]  

(115)

holds for any \( 0 \leq \varepsilon \leq 1 \), where \( C > 0 \) is a constant depending on \( \sigma_{\text{max}} \) and \( \sigma_{\text{min}} \). Therefore, (99) and (100) is proved.

Next, the theory of matrix perturbation will be used to help our analysis about the influence of the estimation error of \( \hat{U}_{\text{sig}}^{(m)} \) on the volume-correlation. Another lemma is need:

**Lemma 4:** (Corollary 2.7 in [52]) For the matrix \( A \in \mathbb{R}^{n \times n} \), and the perturbation matrix \( E \in \mathbb{R}^{n \times n} \), we have

- If \( A \) is full-rank, then
  \[ |\det(A + E) - \det(A)| \leq \sum_{i=1}^{n} s_{n-i}(A)\|E\|_2^i, \]
  (116)

- If \( \text{rank}(A) = k \) for some \( 1 \leq k \leq n - 1 \), then
  \[ |\det(A + E)| \leq \|E\|_2^{n-k} \sum_{i=0}^{k} s_{k-i}(A)\|E\|_2^i. \]
  (117)

where \( s_k(A) \) is the \( k \)th elementary symmetric function of matrix \( A \) defined in [37].

**Proof of Theorem 3**

As theorem 3 describes the situation where \( m \) is sufficiently large, therefore we naturally assume \( m \geq d + p \), the volume-correlation is of the form of:

\[ t(R^{(m)}) = \text{vol}_{d+p+d}([X_s, \hat{U}_{\text{sig}}^{(m)}])/(\text{vol}_d(X_s) \text{vol}_r(\hat{U}_{\text{sig}}^{(m)})), \]

(118)

Denote the orthogonal basis matrix of \( X_s \) by \( Q_{X_s} \), i.e., \( Q_{X_s}^T Q_{X_s} = I_d \) and \( \text{span}(Q_{X_s}) = X_s \), then we have

\[ t(R^{(m)}) = \text{vol}_{d+p+d}([Q_{X_s}, \hat{U}_{\text{sig}}^{(m)}]) \]

\[ = \det \left( \begin{bmatrix} Q_{X_s}^T & Q_{X_s}^{\hat{U}_{\text{sig}}^{(m)}} \\ \hat{U}_{\text{sig}}^{(m)} & \hat{U}_{\text{sig}}^{(m)} \end{bmatrix} \right)^{1/2} \]

\[ = \det(Q_{X_s}^T Q_{X_s})^{1/2} \cdot \det \left( I_d - (\hat{U}_{\text{sig}}^{(m)})^T Q_{X_s} (Q_{X_s}^T Q_{X_s})^{-1} Q_{X_s}^T \right)^{1/2} \]

\[ = \det \left( (\hat{U}_{\text{sig}}^{(m)})^T P_{X_s}^\perp \hat{U}_{\text{sig}}^{(m)} \right)^{1/2}. \]  

(119)
where $P_{X_s}^\perp$ is the orthogonal projection operator onto the subspace $X_s$. According to Lemma 2, the matrix constructed from estimated signal eigenvectors $\hat{U}_{\text{sig}}^{(m)} \in \mathbb{R}^{P \times (p+d)}$ can be regarded as the matrix constructed from real signal eigenvectors plus a perturbation matrix, i.e.,

$$\hat{U}_{\text{sig}}^{(m)} = U_{\text{sig}} + E^{(m)}, \quad \text{(120)}$$

where (since it is under $H_1$ hypothesis)

$$\text{span}(U_{\text{sig}}) = X_s \oplus X_I,$$

and

$$E^{(m)} = [e^{(m)}_1, \ldots, e^{(m)}_{p+d}], \quad \text{(121)}$$

with $e^{(m)}_i \sim N(0, \Sigma^{(m)}_i)$, and different $e^{(m)}_i$ are independent. Then according to (97), we have

$$\Sigma^{(m)}_i = \lambda_i \frac{1}{m} \left( \sum_{j=1, j \neq i}^{p+d} \frac{\lambda_j}{\lambda_i - \lambda_j} u_i u_i^T + \sum_{j=p+d+1}^{p} \frac{\sigma_j^2}{\sigma_i^2 - \lambda_j} u_i u_i^T \right). \quad \text{(122)}$$

Therefore (119) becomes

$$t(R^{(m)}) = \det \left( (\hat{U}_{\text{sig}}^{(m)})^T P_{X_s}^\perp \hat{U}_{\text{sig}}^{(m)} \right)^{1/2} = \det \left( (P_{X_s}^\perp U_{\text{sig}} + P_{X_s}^\perp E^{(m)})^T (P_{X_s}^\perp U_{\text{sig}} + P_{X_s}^\perp E^{(m)}) \right)^{1/2}, \quad \text{(123)}$$

for convenience, let

$$Q := P_{X_s}^\perp U_{\text{sig}}, \quad W := P_{X_s}^\perp E^{(m)},$$

then (123) becomes

$$t^2(R^{(m)}) = \det \left[ (Q + W)^T (Q + W) \right], \quad \text{(124)}$$

now the result in Lemma 4 is ready for use, let

$$A := Q^T Q, \quad E = Q^T W + W^T Q + W^T W,$$

then

$$t^2(R^{(m)}) = \det(A + E),$$

Therefore

$$A = U_{\text{sig}}^T P_{X_s}^\perp U_{\text{sig}}, \quad E = U_{\text{sig}}^T P_{X_s}^\perp E^{(m)} + (E^{(m)})^T P_{X_s}^\perp U_{\text{sig}} + (E^{(m)})^T P_{X_s}^\perp E^{(m)}, \quad \text{(125)}$$

As is known that, under $H_1$ hypothesis, $U_{\text{sig}} \in \mathbb{R}^{P \times (p+d)}$ and $\text{span}(U_{\text{sig}}) = X_s \oplus X_I$, because $X_s$ and $X_I$ are linearly independent, therefore $\text{rank}(A) = p$; then according to (117) in Lemma 4 we have

$$t^2(R^{(m)}) = \det(A + E) \leq \|E\|_{p+d-p} \sum_{i=0}^{p} s_{p-i}(A) \|E\|_2^i, \quad \text{(126)}$$

June 6, 2014 DRAFT
where
\[
\|E\|_2 = \|U_{sig}^T P_{X_i}^\perp E^{(m)} + (E^{(m)})^T P_{X_i}^\perp U_{sig} + (E^{(m)})^T P_{X_i}^\perp E^{(m)}\|_2
\]
\[
\leq 2\|U_{sig}^T P_{X_i}^\perp E^{(m)}\|_2 + \|P_{X_i}^\perp E^{(m)}\|_2
\]
\[
\leq 2\|U_{sig}\|_2\|P_{X_i}^\perp E^{(m)}\|_2 + \|P_{X_i}^\perp E^{(m)}\|_2
\]
\[
= 2\|P_{X_i}^\perp E^{(m)}\|_2 + \|P_{X_i}^\perp E^{(m)}\|_2
\]
\[
\leq 2\|P_{X_i}^\perp E^{(m)}\|_F + \|P_{X_i}^\perp E^{(m)}\|_F.
\]
(127)

Then, according to Lemma 2 and Lemma 3 we have
\[
P_{X_i}^\perp E^{(m)} \sim \mathcal{N}(0, P_{X_i}^\perp \Sigma_i^{(m)}(P_{X_i}^\perp)^T),
\]
(128)

and for any \( \varepsilon > 0 \)
\[
\|P_{X_i}^\perp E^{(m)}\|_F^2 \leq (1 + \varepsilon) \sum_{i=1}^{p+d} \text{Tr}(P_{X_i}^\perp \Sigma_i(P_{X_i}^\perp)^T),
\]
(129)

holds with probability
\[
P \geq 1 - \exp\left\{-\frac{r \cdot P \cdot \varepsilon^2}{C}\right\}.
\]

Consider the right side of (129), we have
\[
\sum_{i=1}^{p+d} \text{Tr}(P_{X_i}^\perp \Sigma_i P_{X_i})
\]
\[
= \sum_{i=1}^{p+d} \left( \frac{1}{m} \sum_{j=1}^{p+d} \frac{\lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} \text{Tr}(P_{X_i}^\perp u_j u_j^T P_{X_i}^\perp)^T) + \sum_{j=p+d+1}^{P} \frac{\lambda_i \sigma^2}{(\sigma^2 - \lambda_i)^2} \text{Tr}(P_{X_i}^\perp u_j u_j^T P_{X_i}^\perp)^T) \right),
\]
(130)

because
\[
\text{Tr}(P_{X_i}^\perp u_j u_j^T P_{X_i}^\perp)^T) = \text{Tr}(u_j^T P_{X_i}^\perp u_j) \leq 1,
\]
(131)

so
\[
\sum_{i=1}^{p+d} \text{Tr}(P_{X_i}^\perp \Sigma_i P_{X_i}) \leq \frac{1}{m} \left( \sum_{i=1}^{p+d} \sum_{j=1}^{p+d} \frac{\lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} + \sum_{i=1}^{p+d} (P - p - d) \frac{\lambda_i \sigma^2}{(\sigma^2 - \lambda_i)^2} \right).
\]
(132)

Combine (132) and (129), because what we need is a sufficient condition, then for any \( \varepsilon > 0 \) and \( 0 \leq \delta < 1 \), we let
\[
\frac{1}{m} \left( \sum_{i=1}^{p+d} \sum_{j=1}^{p+d} \frac{\lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} + \sum_{i=1}^{p+d} (P - p - d) \frac{\lambda_i \sigma^2}{(\sigma^2 - \lambda_i)^2} \right) \leq \frac{\sqrt{\delta + 1} - 1)^2}{1 + \varepsilon},
\]
(133)

then equivalently, for any \( \varepsilon > 0 \) and \( 0 \leq \delta < 1 \), if
\[
m \geq \frac{1 + \varepsilon}{(\sqrt{\delta + 1} - 1)^2} \left( \sum_{i=1}^{p+d} \sum_{j=1}^{p+d} \frac{\lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} + \sum_{i=1}^{p+d} (P - p - d) \frac{\lambda_i \sigma^2}{(\sigma^2 - \lambda_i)^2} \right),
\]
(134)
then
\[ \| P_{\mathcal{X}_i} E^{(m)} \|_F^2 \leq (\sqrt{\delta + 1} - 1)^2, \] (135)
holds with probability
\[ \mathbb{P} \geq 1 - \exp\{- \frac{r \cdot P \cdot \varepsilon^2}{C}\}. \]

Then combining (135) with (126), we get
\[ \| E \|_2 \leq \delta, \] (136)
thus we have
\[ t^2(R^{(m)}) \leq s_p(U_{sig}^T P_{\mathcal{X}_s} U_{sig}) \delta^d + O(\delta^{d+1}), \] (137)
holds with probability \( \mathbb{P} \geq 1 - 2 \exp\{- \frac{r \cdot P \cdot \varepsilon^2}{C}\}. \) According to the definition of elementary symmetric function of singular values in (37), \( s_k(A) \) is unitary-invariant, and \( \text{span}(U_{sig}^{(m)}) = \mathcal{X}_s \oplus \mathcal{X}_l = \text{span}(Q_{\mathcal{X}_l}) \), therefore we can write \( s_p(U_{sig}^T P_{\mathcal{X}_s} U_{sig}) = s_p(Q_{\mathcal{X}_l}^T P_{\mathcal{X}_s} Q_{\mathcal{X}_l}) \).

Then (35) of Theorem 3 is proved.

**Proof of Theorem 4**

The result (40) of Theorem 4 can be similarly derived, as is known that under \( H_0 \) hypothesis, \( U_{sig} \in \mathbb{R}^{P \times p} \), \( \text{span}(U_{sig}) = \mathcal{X}_l \), similarly we can let
\[ A = U_{sig}^T P_{\mathcal{X}_s} U_{sig}, \quad E = U_{sig}^T P_{\mathcal{X}_s} E^{(m)} + (E^{(m)})^T P_{\mathcal{X}_s} U_{sig} + (E^{(m)})^T P_{\mathcal{X}_s} E^{(m)}, \]
and \( A \) is full rank, therefore, according to (116)
\[ |\text{det}(A + E) - \text{det}(A)| \leq s_{n-1}(A) \| E \|_2 + O(\| E \|_2^2), \] (138)
and according to (119), we have
\[ \text{det}(A) = \text{det}(U_{sig}^T P_{\mathcal{X}_s} U_{sig}) = \text{det}(Q_{\mathcal{X}_l}^T P_{\mathcal{X}_s} Q_{\mathcal{X}_l}) = \text{vol}_{d+p}(Q_{\mathcal{X}_s}, Q_{\mathcal{X}_l}) = \tau^2(\mathcal{X}_s, \mathcal{X}_l), \] (139)
Then combined with (127), and according to Lemma 2 and 3 for any \( 0 \leq \delta < 1 \) and \( \varepsilon > 0 \), if
\[ m \geq \frac{1 + \varepsilon}{(\sqrt{\delta + 1} - 1)^2} \left( \sum_{i=1}^p \sum_{j=1}^p \frac{\lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} + \sum_{i=1}^p (P - p) \frac{\lambda_i \sigma^2}{(\sigma^2 - \lambda_i)^2} \right), \] (140)
then
\[ \| P_{\mathcal{X}_s} E^{(m)} \|_F^2 \leq (\sqrt{\delta + 1} - 1)^2, \] (141)
hold with probability
\[ \mathbb{P} \geq 1 - 2 \exp\{- \frac{r \cdot P \cdot \varepsilon^2}{C}\}. \]
therefore
\[ |t^2(R^{(m)}) - \tau^2(\mathcal{X}_s, \mathcal{X}_l)| \leq s_{p-1}(U_{sig}^T P_{\mathcal{X}_s} U_{sig}) \delta + O(\delta^2), \] (142)
where in the same way we can write \( s_{p-1}(U_{sig}^T P_{\mathcal{X}_s} U_{sig}) = s_{p-1}(Q_{\mathcal{X}_l}^T P_{\mathcal{X}_s} Q_{\mathcal{X}_l}), \) (40) is then proved.

June 6, 2014 DRAFT
REFERENCES

[1] W. Xu, M. L. Honig, J. Zeidler, and L. B. Milstein, “Subspace adaptive filtering techniques for multi-sensor ds-cdma interference suppression in the presence of a frequency-selective fading channel,” in *1998 Proc. Asilomar Conference on Signals, Systems, and Computers*, 1998.

[2] M. L. McCloud and L. L. Scharf, “Interference estimation with applications to blind multiple-access communication over fading channels,” *Information Theory, IEEE Transactions on*, vol. 46, no. 3, pp. 947–961, 2000.

[3] M. L. McCloud, L. L. Scharf, and M. K. Varanasi, “Beamforming, diversity, and interference rejection for multuser communication over fading channels with a receive antenna array,” *Communications, IEEE Transactions on*, vol. 51, no. 1, pp. 116–124, 2003.

[4] V. Tuzukov, W.-S. Yoon, and Y. D. Kim, “Wireless sensor networks based on the generalized approach to signal processing with fading channels and receive antenna array,” *WSEAS Transactions on Circuits and Systems*, vol. 3, no. 10, pp. 2149–2155, 2004.

[5] E. Conte, M. Lops, and G. Ricci, “Asymptotically optimum radar detection in compound-gaussian clutter,” *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 31, no. 2, pp. 617–625, 1995.

[6] F. Gini and A. Farina, “Vector subspace detection in compound-gaussian clutter. part i: survey and new results,” *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 38, no. 4, pp. 1295–1311, 2002.

[7] F. Gini, A. Farina, and M. Montanari, “Vector subspace detection in compound-gaussian clutter. part ii: Performance analysis,” *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 38, no. 4, pp. 1312–1323, 2002.

[8] M. Rangaswamy, F. C. Lin, and K. R. Gerlach, “Robust adaptive signal processing methods for heterogeneous radar clutter scenarios,” *Signal Processing, vol. 84, no. 9, pp. 1653–1665, 2004.*

[9] P. Wang, H. Li, and B. Himed, “Moving target detection using distributed mimo radar in clutter with nonhomogeneous power,” *Signal Processing, IEEE Transactions on*, vol. 59, no. 10, pp. 4809–4820, 2011.

[10] O. Besson, L. L. Scharf, and F. Vincent, “Matched direction detectors and estimators for array processing with subspace steering vector uncertainties,” *Signal Processing, IEEE Transactions on*, vol. 53, no. 12, pp. 4453–4463, 2005.

[11] Y. Jin and B. Friedlander, “Detection of distributed sources using sensor arrays,” *Signal Processing, IEEE Transactions on*, vol. 52, no. 6, pp. 1537–1548, 2004.

[12] D. W. Stein, S. G. Beaver, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, “Anomaly detection from hyperspectral imagery,” *Signal Processing Magazine, IEEE*, vol. 19, no. 1, pp. 58–69, 2002.

[13] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, “Subspace methods for the blind identification of multichannel fir filters,” *Signal Processing, IEEE Transactions on*, vol. 43, no. 2, pp. 516–525, 1995.

[14] F. Gini and A. Farina, “Matched subspace cfar detection of hovering helicopters,” *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 35, no. 4, pp. 1293–1305, 1999.

[15] L. L. Scharf and B. Friedlander, “Matched subspace detectors,” *Signal Processing, IEEE Transactions on*, vol. 42, no. 8, pp. 2146–2157, 1994.

[16] L. L. Scharf and L. T. McWhorter, “Adaptive matched subspace detectors and adaptive coherence estimators,” in *Signals, Systems and Computers, 1996. Conference Record of the Thirtieth Asilomar Conference on*. IEEE, 1996, pp. 1114–1117.

[17] S. Kraut and L. L. Scharf, “The cfar adaptive subspace detector is a scale-invariant glrt,” *Signal Processing, IEEE Transactions on*, vol. 47, no. 9, pp. 2538–2541, 1999.

[18] S. Kraut, L. L. Scharf, and L. T. McWhorter, “Adaptive subspace detectors,” *Signal Processing, IEEE Transactions on*, vol. 49, no. 1, pp. 1–16, 2001.

[19] M. N. Desai and R. S. Mangoubi, “Robust gaussian and non-gaussian matched subspace detection,” *Signal Processing, IEEE Transactions on*, vol. 51, no. 12, pp. 3115–3127, 2003.

[20] H. Kwon and N. M. Nasrabadi, “Kernel matched subspace detectors for hyperspectral target detection,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 2, pp. 178–194, 2006.

[21] ——, “Kernel rx-algorithm: a nonlinear anomaly detector for hyperspectral imagery,” *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 2, pp. 388–397, 2005.

[22] X. Wang, Z. Li, and D. Tao, “Subspaces indexing model on grassmann manifold for image search,” *Image Processing, IEEE Transactions on*, vol. 20, no. 9, pp. 2627–2635, 2011.

[23] P. Turaga, A. Veeraraghavan, and R. Chellappa, “Statistical analysis on stiefel and grassmann manifolds with applications in computer vision,” in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.
[50] M. Ledoux, *The Concentration of Measure Phenomenon*, ser. Mathematical Surveys and Monographs. American Mathematical Society, 2001.

[51] M. A. Davenport, “Random observations on random observations: Sparse signal acquisition and processing,” thesis, Rice University, 2010.

[52] I. C. Ipsen and R. Rehman, “Perturbation bounds for determinants and characteristic polynomials,” *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 2, pp. 762–776, 2008.
Plot of $P_d$ and $P_f$ with respect to $m$ (d=10, p=40, SNR=14.9)