UCSF

UC San Francisco Previously Published Works

Title

Subspace-based interference removal methods for a multichannel biomagnetic sensor array

Permalink

https://escholarship.org/uc/item/8qf346z9

Journal

Journal of Neural Engineering, 14(5)

ISSN

1741-2560

Authors

Sekihara, Kensuke

Nagarajan, Srikantan S

Publication Date

2017-10-01

DOI

10.1088/1741-2552/aa7693

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

J Neural Eng. Author manuscript; available in PMC 2018 December 10.

Published in final edited form as:

J Neural Eng. 2017 October ; 14(5): 051001. doi:10.1088/1741-2552/aa7693.

Subspace-based interference removal methods for a multichannel biomagnetic sensor array

Kensuke Sekihara[1,2] and **Srikantan S Nagarajan**[3]

[1]Signal Analysis Inc., Hachioji, Tokyo, Japan

[2]Department of Advanced Technology in Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519, Japan

[3]Biomagnetic Imaging Laboratory, University of California, San Francisco, 513 Parnassus Avenue, S362, San Francisco, CA 94143, United States of America

Abstract

Objective.—In biomagnetic signal processing, the theory of the signal subspace has been applied to removing interfering magnetic fields, and a representative algorithm is the signal space projection algorithm, in which the signal/interference subspace is defined in the spatial domain as the span of signal/interference-source lead field vectors. This paper extends the notion of this conventional (spatial domain) signal subspace by introducing a new definition of signal subspace in the time domain.

Approach.—It defines the time-domain signal subspace as the span of row vectors that contain the source time course values. This definition leads to symmetric relationships between the time-domain and the conventional (spatial-domain) signal subspaces. As a review, this article shows that the notion of the time-domain signal subspace provides useful insights over existing interference removal methods from a unified perspective.

Main results and significance.—Using the time-domain signal subspace, it is possible to interpret a number of interference removal methods as the time domain signal space projection. Such methods include adaptive noise canceling, sensor noise suppression, the common temporal subspace projection, the spatio-temporal signal space separation, and the recently-proposed dual signal subspace projection. Our analysis using the notion of the time domain signal space projection reveals implicit assumptions these methods rely on, and shows that the difference between these methods results only from the manner of deriving the interference subspace. Numerical examples that illustrate the results of our arguments are provided.

Keywords

interference removal; magnetoencephalography; sensor array processing; signal subspace; biomagnetic imaging; biomagnetism; multi-sensor array

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

1. Introduction

The notion of signal and noise subspaces has been considered useful in the signal processing of data acquired using multichannel sensor arrays [1]. In biomagnetic signal processing, the theory of the signal subspace has been applied to removing interfering magnetic fields, and a representative algorithm in this regard is the signal space projection (SSP) algorithm [2–4]. In this algorithm, the interference subspace is defined as the span of interference-source lead field vectors [2, 5], and it is estimated by utilizing so-called empty-room-noise data. The algorithm projects the measured data onto the subspace orthogonal to the interference subspace to remove the interference.

In biomagnetic signal processing, the signal subspace has been defined to reflect a spatial property of the multichannel data [2, 5]. On the other hand, since biomagnetic signals have rich temporal properties, the incorporation of temporal information can be a natural and fruitful way to extend the notion of the conventional (spatial domain) signal subspace. In this paper, we propose, for the first time (as far as we know), a new definition of signal subspace in the time domain. The time-domain signal subspace is defined as the span of row vectors that contain the source time course values. Such row vectors are referred to as the source time course vectors.

By defining the time-domain signal subspace in this manner, we derive symmetric relationships between the time-domain signal subspace and the conventional (spatial-domain) signal subspace. That is, while a column vector of signal components at a particular time point lies within the spatial-domain signal subspace, a row vector of signal time course from a particular sensor lies within the time-domain signal subspace. While the column space of the signal matrix is equal to the spatial domain signal subspace, the row space of this matrix is equal to the time-domain signal subspace.

As a review article, this paper does not propose a new method, but rather a new way of looking at various existing methods in a unified perspective. Actually, using the time-domain signal subspace, it is possible to interpret various interference removal methods as the time domain signal space projection. It can be shown that these methods differ only in the manner by which they derive the interference subspace. These methods rely on some implicit assumptions that are generally hidden behind their formulations, but they can be revealed by our analysis using the notion of the time domain SSP. Such methods include adaptive noise canceling [6, 7], sensor noise suppression [8], common temporal subspace projection [9], spatio-temporal signal space separation [10] and the recently-proposed dual signal subspace projection [11].

The paper is organized as follows: After reviewing the conventional (spatial-domain) signal subspace, the time-domain signal subspace is defined in section 2. The interference-removal methods utilizing the spatial-domain signal subspace are reviewed in section 3. An important variant of the SSP algorithm, the signal space separation (SSS) method [12–14], is also reviewed in this section. The time-domain SSP algorithm is introduced in section 4. Section 5 presents interpretations of various existing algorithms as the time-domain SSP. Here, we argue how the interference subspace in the time domain can be derived by those algorithms.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

We provide numerical examples that illustrate the results of our arguments in section 6. The Appendix provides several proofs of mathematical arguments, as well as some details of the SSS algorithm needed for readers to follow our arguments.

2. Signal subspaces in the spatial and time domains

2.1. Sensor array measurements

Biomagnetic measurement is usually conducted using a sensor array, which simultaneously measures the biomagnetic signal with multiple sensors. Let us define the measurement of the mth sensor at time t as $y_m(t)$. The measurement from the whole sensor array is expressed as a column vector $\boldsymbol{y}(t)$: $\boldsymbol{y}(t) = [y_1(t), y_2(t), \ldots, y_M(t)]^T$. Here, M is the number of sensors, and the superscript T indicates the matrix transpose. Throughout this paper, plain italics indicate scalars, lower-case boldface italics indicate vectors, and upper-case boldface italics indicate matrices. The location in the three-dimensional space is represented by $\boldsymbol{r}$: $\boldsymbol{r} = (x, y, z)$. The source magnitude at $\boldsymbol{r}$ and time t is denoted as a scalar $s(\boldsymbol{r}, t)$. The source vector is denoted $\boldsymbol{s}(\boldsymbol{r}, t)$, and the source orientation is denoted $\boldsymbol{\eta} = [\eta_x, \eta_y, \eta_z]^T$. We thus have the relationship: $\boldsymbol{s}(\boldsymbol{r}, t) = s(\boldsymbol{r}, t)\boldsymbol{\eta}$.

Let us assume that a unit-magnitude source exists at $\boldsymbol{r}$. When this unit-magnitude source is directed in the x, y, and z directions, the outputs of the mth sensor are respectively denoted by $l_m^x(\boldsymbol{r})$, $l_m^y(\boldsymbol{r})$, and $l_m^z(\boldsymbol{r})$. Let us define an $M \times 3$ matrix $\boldsymbol{L}(\boldsymbol{r})$ whose mth row is equal to $\left[l_m^X(\boldsymbol{r}), l_m^y(\boldsymbol{r}), l_m^z(\boldsymbol{r})\right]$. This matrix $\boldsymbol{L}(\boldsymbol{r})$, referred to as the lead field matrix, represents the sensitivity of the sensor array at $\boldsymbol{r}$. When the unit-magnitude source at $\boldsymbol{r}$ is oriented in the $\boldsymbol{\eta}$ direction, the outputs of the sensor array are expressed as $\boldsymbol{l}(\boldsymbol{r}) = \boldsymbol{L}(\boldsymbol{r})\boldsymbol{\eta}$. This column vector $\boldsymbol{l}(\boldsymbol{r})$, referred to as the lead field vector, represents the sensitivity of the sensor array in the direction of $\boldsymbol{\eta}$ at the location $\boldsymbol{r}$.

The outputs of the sensor array $\boldsymbol{y}(t)$ are expressed as the sum of the signal component $\boldsymbol{y}_S(t)$ and the noise $\boldsymbol{\varepsilon}$:

$$\boldsymbol{y}(t) = \boldsymbol{y}_S(t) + \boldsymbol{\varepsilon}. \quad (1)$$

In equation (1), $\boldsymbol{y}_S(t)$ is called the signal vector, which is expressed as:

$$\boldsymbol{y}_S(t) = \int_{\Omega} \boldsymbol{L}(\boldsymbol{r})\boldsymbol{s}(\boldsymbol{r}, t)\mathrm{d}\boldsymbol{r}, \quad (2)$$

where the integral on the right-hand side is carried out over a three-dimensional volume Ω where signal sources of interest can exist. This Ω is called the source space. In equation (1), an $M \times 1$ random vector $\boldsymbol{\varepsilon}$ represents additive sensor noise, which is assumed to obey the normal distribution:

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$p(\varepsilon) = \mathcal{N}\left(\varepsilon \middle| 0, \varrho^2 \boldsymbol{I}\right), \quad (3)$$

where $\boldsymbol{I}$ is the identity matrix and ϱ^2 is the variance of the sensor noise.

We denote the time series outputs of a sensor array $\boldsymbol{y}(t_1), \ldots, \boldsymbol{y}(t_K)$, where K is the total number of measured time points. It is assumed that $K > M$ in this paper. We define the measured data matrix $\boldsymbol{B}$ as:

$$\boldsymbol{B} = \left[\boldsymbol{y}(t_1), \ldots, \boldsymbol{y}(t_K)\right] = \left[\boldsymbol{y}_1, \ldots, \boldsymbol{y}_K\right], \quad (4)$$

where $\boldsymbol{y}(t_j)$ is denoted $\boldsymbol{y}_j$ for simplicity. We also define a matrix of the signal vector such that

$$\boldsymbol{B}_S = \left[\boldsymbol{y}_S(t_1), \ldots, \boldsymbol{y}_S(t_K)\right] = \left[\boldsymbol{y}_1^S, \ldots, \boldsymbol{y}_K^S\right], \quad (5)$$

where the jth column of $\boldsymbol{B}_S$ is denoted $\boldsymbol{y}_j^S$. This $\boldsymbol{B}_S$ is called the signal matrix in this paper. Then, the data model in equation (1) is expressed in matrix form as:

$$\boldsymbol{B} = \boldsymbol{B}_S + \boldsymbol{B}_{\boldsymbol{\varepsilon}}, \quad (6)$$

where $\boldsymbol{B}_{\boldsymbol{\varepsilon}}$ is the noise matrix whose jth column is equal to the noise vector $\boldsymbol{\varepsilon}$ at time t_j.

2.2. Definition of signal subspace in the spatial domain

Let us assume that a total of Q discrete sources exist. Their locations are denoted by $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_Q$, their orientations by $\boldsymbol{\eta}_1, \cdots, \boldsymbol{\eta}_Q$, and their magnitudes by $s_1(t), \ldots, s_Q(t)$. Then, the source distribution is expressed as:

$$\boldsymbol{s}(\boldsymbol{r}, t) = \sum_{q=1}^{Q} s_q(t) \boldsymbol{\eta}_q \delta\left(\boldsymbol{r} - \boldsymbol{r}_q\right), \quad (7)$$

where $\delta(\boldsymbol{r})$ indicates the delta function. Substituting the equation above into equation (2), the signal vector $\boldsymbol{y}_S(t)$ is expressed as:

$$\boldsymbol{y}_S(t) = \int_{\Omega} \boldsymbol{L}(\boldsymbol{r}) \sum_{q=1}^{Q} s_q(t) \boldsymbol{\eta}_q \delta\left(\boldsymbol{r} - \boldsymbol{r}_q\right) \mathrm{d}\boldsymbol{r} = \sum_{q=1}^{Q} s_q(t) \boldsymbol{l}_q, \quad (8)$$

where $\boldsymbol{l}_q$ represents the lead field vector of the qth source obtained such that $\boldsymbol{l}_q = \boldsymbol{L}(\boldsymbol{r}_q)\boldsymbol{\eta}_q$. We assume that the number of sources Q is smaller than the number of sensors, i.e. $Q < M$. This

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

assumption is referred to as the low-rank signal assumption [5, 15], and we hold this assumption throughout the paper[4].

Equation (8) claims that the signal vector is expressed as a linear sum of the lead field vectors $\boldsymbol{l}_1, \cdots, \boldsymbol{l}_Q$. That is, the signal vector $\boldsymbol{y}_S$ lies within a subspace spanned by $\boldsymbol{l}_1, \cdots, \boldsymbol{l}_Q$. The subspace spanned by the source lead field vectors $\boldsymbol{l}_1, \cdots, \boldsymbol{l}_Q$ is defined as the signal subspace [5], which is denoted by $\mathscr{E}_S$, i.e.

$$\mathscr{E}_S = \mathrm{csp}\left(\left[\boldsymbol{l}_1, \cdots, \boldsymbol{l}_Q\right]\right). \quad (9)$$

Here, the notation csp(·) indicates the column space of the matrix within the parentheses. Equation (8) indicates the relationship,

$$\boldsymbol{y}_S(t) \in \mathscr{E}_S. \quad (10)$$

The signal vector lies within the signal subspace, which is the subspace formed by all possible signal vectors [1].

2.3. Definition of signal subspace in the time domain

This section introduces a novel definition of signal subspace in the time domain. To do so, we define a row vector $\boldsymbol{s}_q$ consisting of the time course of the qth source such that

$$\boldsymbol{s}_q = \left[s_q(t_1), \ldots, s_q(t_K)\right], \quad (11)$$

which we call the time course vector of the qth source. We then prove that a row of the signal matrix $\boldsymbol{B}_S$ is expressed as a linear sum of the time course vectors, $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_Q$. We assume, in this paper, that the source time course vectors $s_1, \cdots, s_Q$ are linearly independent. Substituting equation (8) into equation (5), the following relationship is obtained:

$$\boldsymbol{B}_S = \left[\sum_{q=1}^{Q} s_q(t_1)\boldsymbol{l}_q, \ldots, \sum_{q=1}^{Q} s_q(t_K)\boldsymbol{l}_q\right] = \begin{bmatrix} \sum_{q=1}^{Q}\left[s_q(t_1), \ldots, s_q(t_K)\right]l_q^1 \\ \vdots \\ \sum_{q=1}^{Q}\left[s_q(t_1), \ldots, s_q(t_K)\right]l_q^M \end{bmatrix} = \begin{bmatrix} \sum_{q=1}^{Q} l_q^1 \boldsymbol{s}_q \\ \vdots \\ \sum_{q=1}^{Q} l_q^M \boldsymbol{s}_q \end{bmatrix}, \quad (12)$$

where $l_q^1, \ldots, l_q^M$, are the elements of the lead field vector $\boldsymbol{l}_q$: $\boldsymbol{l}_q = \left[l_q^1, \ldots, l_q^M\right]^T$. Denoting the jth row vector of $\boldsymbol{B}_s$ by $\boldsymbol{\beta}_j^S$, equation (12) shows that

[4]Since we assume that $K > M$, the assumption $K > M > Q$ holds throughout the paper.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\boldsymbol{\beta}_j^S = \sum_{q=1}^{Q} l_q^j \boldsymbol{s}_q . \quad (13)$$

This equation indicates that a row vector of the signal matrix, $\boldsymbol{\beta}_j^S$, is expressed as a linear sum of $\boldsymbol{s}_q$ ($q = 1, \ldots, Q$). That is, we have

$$\boldsymbol{\beta}_j^S \in \mathrm{rsp}\left(\left[\boldsymbol{s}_1^T, \ldots, \boldsymbol{s}_Q^T\right]^T\right), \quad (14)$$

where the notation rsp(·) indicates a row space of the matrix in the parentheses.

Analogous to equations (9) and (10), it is reasonable to define $\mathrm{rsp}\left(\left[\boldsymbol{s}_1^T, \ldots, \boldsymbol{s}_Q^T\right]^T\right)$ as the signal subspace in time domain $\mathscr{K}_S$, i.e.

$$\mathscr{K}_S = \mathrm{rsp}\left(\left[\boldsymbol{s}_1^T, \ldots, \boldsymbol{s}_Q^T\right]^T\right). \quad (15)$$

By defining the time domain signal subspace this way, we can derive symmetric relationships between the time domain signal subspace and the spatial domain signal subspace. That is, we have already shown the relationships:

$$\text{column of } \boldsymbol{B}_S : \boldsymbol{y}_j^S \in \mathscr{E}_S, \quad (16)$$

$$\text{row of } \boldsymbol{B}_S : \boldsymbol{\beta}_j^S \in \mathscr{K}_S . \quad (17)$$

We can show that, with the assumption $K > Q$, the column space of $\boldsymbol{B}_S$ is equal to the spatial domain signal subspace, i.e.

$$\mathscr{E}_S = \mathrm{csp}\left(\boldsymbol{B}_S\right). \quad (18)$$

The proof is presented in appendix A.1. With the assumption $M > Q$, the row space of $\boldsymbol{B}_S$ is equal to the time domain signal subspace, i.e.

$$\mathscr{K}_S = \mathrm{rsp}\left(\boldsymbol{B}_S\right). \quad (19)$$

The proof is presented in appendix A.2.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

3. Interference removal using the spatial-domain signal subspace

3.1. Estimation of spatial-domain signal subspace

Although the signal subspace is defined in equation (9), it is difficult to use this equation to derive the signal subspace, because the source lead field vectors are generally unknown[5]. The signal subspace can be estimated using the time series measurement of the sensor data $\boldsymbol{y}(t)$ as described in the following.

According to equation (18), we can estimate the signal subspace through the estimation of the column space of $\boldsymbol{B}_S$. For this estimation, let us express the singular value decomposition of $\boldsymbol{B}_S$ ($M < K$) as:

$$\boldsymbol{B}_S = \sum_{j=1}^{M} \gamma_j^S \boldsymbol{u}_j^S \left(\boldsymbol{v}_j^S\right)^T .$$

Here, the singular values $\gamma_j^S (j = 1, \ldots, M)$ are numbered in decreasing order, and $\boldsymbol{u}_j^S$ and $\boldsymbol{v}_j^S$ are the spatial and temporal singular vectors, respectively. Since $\boldsymbol{B}_S$ is a matrix whose rank is equal to Q, the singular values of $\boldsymbol{B}_S$ are given as $\gamma_1^S, \ldots, \gamma_Q^S, 0, \ldots, 0$. Namely, $\boldsymbol{B}_S$ has only Q non-zero singular values, and the column space of $\boldsymbol{B}_S$ is equal to the span of singular vectors corresponding to the non-zero singular values, $\boldsymbol{u}_1^S, \ldots, \boldsymbol{u}_Q^S$ That is, we have the relationship

$$\mathscr{E}_S = \mathrm{csp}\left(\left[\boldsymbol{u}_1^S, \ldots, \boldsymbol{u}_Q^S\right]\right). \quad (20)$$

Since the signal matrix $\boldsymbol{B}_S$(and thus its singular vectors) are unknown quantities, we cannot use equation (20) to derive the signal subspace.

The singular vectors $\boldsymbol{u}_1^S, \ldots, \boldsymbol{u}_Q^S$ are estimated using the singular vectors of $\boldsymbol{B}$. Let us denote the spatial singular vectors of the data matrix $\boldsymbol{B}$ that correspond to the Q largest singular values as $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_Q$. Then, using the noise model in equation (3), singular vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_Q$ are asymptotically equal to $\boldsymbol{u}_1^S, \ldots, \boldsymbol{u}_Q^S$. That is, in the limit of infinite number of time samples, we have

$$\mathscr{E}_S = \mathrm{csp}\left(\left[\boldsymbol{u}_1, \ldots, \boldsymbol{u}_Q\right]\right). \quad (21)$$

The proof is presented in appendix A.4. In the case of a finite number of time samples, $\mathrm{csp}([\boldsymbol{u}_1, \ldots, \boldsymbol{u}_Q])$ is equal to the maximum likelihood estimate of the signal subspace $\widehat{\mathscr{E}}_S$:

[5]To derive the source lead field vectors, we must know the locations and orientations of sources. These quantities are generally unknown.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\widehat{\mathscr{E}}_S = \mathrm{csp}\left(\left[\boldsymbol{u}_1, \ldots, \boldsymbol{u}_Q\right]\right). \quad (22)$$

A formal proof of equation (22) can be found in [15] and [16].

3.2. Signal space projection (SSP) algorithm

The signal space projection (SSP) is an algorithm intended to remove the interference overlapped onto the signal [2, 4]. The algorithm is based on the theory of signal subspace and the measurement model is

$$\boldsymbol{y}(t) = \boldsymbol{y}_S(t) + \boldsymbol{y}_I(t) + \varepsilon, \quad (23)$$

where $\boldsymbol{y}_I(t)$ represents the interference overlapped on the signal vector $\boldsymbol{y}_S(t)$. In this paper, this $\boldsymbol{y}_I(t)$ represents interferences originated from outside the source space. The signal vector $\boldsymbol{y}_S(t)$ represents all signals originated from inside the source space[6]. We assume that a total of P sources generate the interference, and $\boldsymbol{y}_I(t)$ is thus expressed as

$$\boldsymbol{y}_I(t) = \sum_{p=1}^{P} \sigma_p(t)\boldsymbol{\xi}_p, \quad (24)$$

where $\boldsymbol{\xi}_p$ is the lead field vector of the pth interference source with its amplitude of $\sigma_p(t)$. Note that we still assume that the row-rank signal assumption $P + Q < M$ holds. According to equation (9), the interference subspace $\mathscr{E}_I$ is defined as

$$\mathscr{E}_I = \mathrm{csp}\left(\left[\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_P\right]\right). \quad (25)$$

The interference subspace can be estimated when the data that contain only interference are available. Such data are called the control data in this paper, and expressed as

$$\boldsymbol{y}_c(t) = \boldsymbol{y}_I(t) + \varepsilon, \quad (26)$$

and the control data matrix is defined as $\boldsymbol{B}_c = [\boldsymbol{y}_c(t_1), \cdots, \boldsymbol{y}_c(t_K)]$ The spatial singular vectors of $\boldsymbol{B}_c$ that correspond to the P largest singular values are denoted by $\boldsymbol{u}_1^c, \ldots, \boldsymbol{u}_P^c$. The projector onto the interference subspace, P_I, is formulated such that

[6]Therefore, $\boldsymbol{y}_S(t)$ includes interferences originated from inside the source space; such interferences include so called brain noise in MEG measurements. In other words, we consider the brain noise as a part of the signal and its removal is not argued in this paper.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\boldsymbol{P}_I = \left[\boldsymbol{u}_1^c, \ldots, \boldsymbol{u}_P^c\right]\left[\boldsymbol{u}_1^c, \ldots, \boldsymbol{u}_P^c\right]^T . \quad (27)$$

Since the interference vector $\boldsymbol{y}_I(t)$ can be expressed as $\boldsymbol{y}_I(t) = \sum_{j=1}^{P} a_j(t)\boldsymbol{u}_j^c$, the relationship $(\boldsymbol{I} - \boldsymbol{P}_I)\boldsymbol{y}_I(t) = 0$ holds. Thus, by projecting the data vector onto the subspace orthogonal to the interference subspace, we have:

$$\hat{\boldsymbol{y}}_S(t) = \left(\boldsymbol{I} - \boldsymbol{P}_I\right)\boldsymbol{y}(t) = \boldsymbol{y}_S(t) - \boldsymbol{P}_I\boldsymbol{y}_S(t) + \left(\boldsymbol{I} - \boldsymbol{P}_I\right)\varepsilon . \quad (28)$$

It is apparent from equation (28) that the projector $\boldsymbol{I} - \boldsymbol{P}_I$ can remove the interference $\boldsymbol{y}_I(t)$ but also affects the signal. The influence on the signal vector is evaluated by the second term on the right-hand side. This term is generally small when the orthogonality between lead field vectors of signal sources $\boldsymbol{l}_j$ ($j = 1, \ldots, Q$) and the basis vectors of the interference subspace $\boldsymbol{u}_i^c (i = 1, \ldots, P)$ is high. The method of interference suppression based on equation (28) is called the signal space projection (SSP) [2, 4]. One problem we have when implementing SSP is the determination of P, the dimension of the interference subspace. This P is usually determined by thresholding the singular values. It is obvious from the arguments above that the underestimation of P results in a stuation that a part of $\boldsymbol{y}_I(t)$ remains in $\hat{\boldsymbol{y}}_S(t)$ and conversely the overestimation may cause a large distortion of the signal vector $\boldsymbol{y}_S(t)$ in $\hat{\boldsymbol{y}}_S(t)$.

3.3. Signal space separation (SSS) algorithm

A method for deriving the signal subspace based not on the span of source lead field vectors but on physical properties of the magnetic field predicted by Maxwell's equations has been proposed [12–14, 17]. In this method, assuming that there are no sources in a region where sensors are located (a region called the sensor region), the sensor measurements can be expressed by an expansion using the vector spherical harmonics, which contains natural separation between the magnetic fields generated from the internal and external regions. Here, the internal region indicates the region closer to the origin than the sensor region, and the external region indicates the region farther from the origin than the sensor region.

Therefore, if the origin is properly set such that the source space Ω is included within the internal region, and the interference sources are located within the external region, the signal can be separated from the interference. Interference removal based on this idea is called signal space separation (SSS). The derivation of the SSS method is presented in the appendix B.

With the proper setting of the origin, we can derive the SSS signal extractor $\boldsymbol{\Gamma}_S$:

$$\boldsymbol{\Gamma}_S = \boldsymbol{C}\boldsymbol{C}^T\left(\boldsymbol{C}\boldsymbol{C}^T + \boldsymbol{D}\boldsymbol{D}^T\right)^{-1} = \boldsymbol{C}\boldsymbol{C}^T\left(\boldsymbol{S}\boldsymbol{S}^T\right)^{-1}, \quad (29)$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

where the matrices $\boldsymbol{C}$ and $\boldsymbol{D}$ are given respectively in equations (B.9) and (B.10), and $\boldsymbol{S} = [\boldsymbol{C},\boldsymbol{D}]$. Note that $\boldsymbol{C}$ and $\boldsymbol{D}$ are $M \times N_C$ and $M \times N_D$ matrices where N_C and N_D are defined in equation (B.14). The derivation of $\boldsymbol{\Gamma}_S$ is presented in appendix B.2. We use the data model in equation (23), and assume that equation (8) holds for $\boldsymbol{y}_S(t)$ and equation (24) for $\boldsymbol{y}_I(t)$. Then, multiplying the signal extractor $\boldsymbol{\Gamma}_S$ to the data vector $\boldsymbol{y}(t)$ gives

$$\boldsymbol{\Gamma}_S \boldsymbol{y}(t) = \boldsymbol{\Gamma}_S \boldsymbol{y}_S(t) + \boldsymbol{\Gamma}_S \boldsymbol{y}_I(t) = \sum_{q=1}^{Q} s_q(t) \boldsymbol{\Gamma}_S \boldsymbol{l}_q + \sum_{p=1}^{P} \sigma_p(t) \boldsymbol{\Gamma}_S \boldsymbol{\xi}_p, \quad (30)$$

where the sensor noise term is dropped for simplicity.

The assumption that the source space Ω is located within the internal region leads to the relationship

$$\mathrm{csp}(\boldsymbol{C}) \supset \mathscr{E}_S .$$

Thus, the lead field vector of a signal source $\boldsymbol{l}_q$ lies within the column space of $\boldsymbol{C}$, and $\boldsymbol{l}_q$ is expressed as a linear sum of the column vectors of $\boldsymbol{C}$, such that

$$\boldsymbol{l}_q = \sum_{j=1}^{N_c} \alpha_j \boldsymbol{c}_j = \boldsymbol{C}\boldsymbol{\alpha}, \quad (31)$$

where $\boldsymbol{c}_j$ is the jth column of $\boldsymbol{C}$, α_j is the jth expansion coefficient, and $\boldsymbol{\alpha}$ is a column vector containing the coefficients, i.e. $\boldsymbol{\alpha} = \left[\alpha_1, \ldots, \alpha_{N_C}\right]^T$. Thus, denoting an $N_D \times 1$ column vector whose elements are all zero by $\mathbf{0}$, and using the derivation in equation (B.20), we can derive the relationship

$$\begin{aligned} \boldsymbol{\Gamma}_S \boldsymbol{l}_q &= \boldsymbol{\Gamma}_S \boldsymbol{C}\boldsymbol{\alpha} = \boldsymbol{\Gamma}_S \boldsymbol{S}\begin{bmatrix}\boldsymbol{\alpha}\\ \mathbf{0}\end{bmatrix} = \boldsymbol{C}\left[\boldsymbol{C}^T\left(\boldsymbol{S}\boldsymbol{S}^T\right)^{-1}\boldsymbol{S}\begin{bmatrix}\boldsymbol{\alpha}\\ \mathbf{0}\end{bmatrix}\right] \\ &= \boldsymbol{C}\left[\boldsymbol{S}^T\left(\boldsymbol{S}\boldsymbol{S}^T\right)^{-1}\boldsymbol{S}\begin{bmatrix}\boldsymbol{\alpha}\\ \mathbf{0}\end{bmatrix}\right]_{[1:N_C]} \approx \boldsymbol{C}\left[\left(\boldsymbol{S}^T\boldsymbol{S}\right)^{-1}\boldsymbol{S}^T\boldsymbol{S}\begin{bmatrix}\boldsymbol{\alpha}\\ \mathbf{0}\end{bmatrix}\right]_{[1:N_C]} \\ &= \boldsymbol{C}\boldsymbol{\alpha} = \boldsymbol{l}_q . \end{aligned} \quad (32)$$

The equation above indicates that the SSS signal extractor $\boldsymbol{\Gamma}_S$ passes the signal-source lead field vector $\boldsymbol{l}_q$ with no distortion.

The assumption that all the interference sources are located within the external region leads to the situation that the column space of $\boldsymbol{D}$ includes the interference subspace, i.e.

$$\mathrm{csp}(\boldsymbol{D}) \supset \mathscr{E}_I .$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Therefore, the lead field vector of an interference source $\boldsymbol{\xi}_p$ lies within the column space of $\boldsymbol{D}$, and $\boldsymbol{\xi}_p$ is expressed as

$$\boldsymbol{\xi}_p = \sum_{j=1}^{N_D} \phi_j \boldsymbol{d}_j = \boldsymbol{D}\boldsymbol{\phi}, \quad (33)$$

where $\boldsymbol{d}_j$ is the jth column of $\boldsymbol{D}$, $\boldsymbol{\phi}_j$ is the jth expansion coefficient, and $\boldsymbol{\phi}$ is a column vector containing these coefficients: $\boldsymbol{\phi} = \left[\phi_1, \ldots, \phi_{N_D}\right]^T$. Again denoting the $N_C \times 1$ column vector whose elements are all zero by $\mathbf{0}$, we have the relationship

$$\boldsymbol{\Gamma}_S \boldsymbol{\xi}_q = \boldsymbol{\Gamma}_S \boldsymbol{D}\boldsymbol{\phi} = \boldsymbol{\Gamma}_S \boldsymbol{S}\begin{bmatrix}\mathbf{0}\\ \boldsymbol{\phi}\end{bmatrix} \approx \boldsymbol{C}\left[\left(\boldsymbol{S}^T\boldsymbol{S}\right)^{-1}\boldsymbol{S}^T\boldsymbol{S}\begin{bmatrix}\mathbf{0}\\ \boldsymbol{\phi}\end{bmatrix}\right]_{\left[1:N_C\right]} = \boldsymbol{C}\mathbf{0} = 0. \quad (34)$$

The equation above indicates that the signal extractor $\boldsymbol{\Gamma}_S$ completely blocks the lead field vector of an interference source $\boldsymbol{\xi}_q$. Consequently, substituting (32) and (34) into (30), we obtain

$$\boldsymbol{\Gamma}_S \boldsymbol{y}(t) = \boldsymbol{\Gamma}_S \boldsymbol{y}_S(t) + \boldsymbol{\Gamma}_s \boldsymbol{y}_I(t) = \sum_{q=1}^{Q} s_q(t)\boldsymbol{\Gamma}_s \boldsymbol{l}_q + \sum_{p=1}^{P} \sigma_p(t)\boldsymbol{\Gamma}_S \boldsymbol{\xi}_p = \sum_{q=1}^{Q} s_q(t)\boldsymbol{l}_q = \boldsymbol{y}_S(t). \quad (35)$$

The equation above shows that by multiplying the signal extractor $\boldsymbol{\Gamma}_S$ with the data vector $\boldsymbol{y}(t)$, the signal vector $\boldsymbol{y}_S(t)$ is selectively extracted with no distortion. This distortionless signal extraction is a major advantage of the SSS algorithm over the SSP algorithm.

4. Time-domain signal space projection

If we obtain the basis vectors of the interference subspace in the time domain, it is possible to remove the interference by projecting the measured data onto the subspace orthogonal to the time-domain interference subspace. We define the interference matrix $\boldsymbol{B}_I$ as

$$\boldsymbol{B}_I = \left[\boldsymbol{y}_I(t_1), \ldots, \boldsymbol{y}_I(t_K)\right]. \quad (36)$$

Then, the data model in equation (23) is expressed as

$$\boldsymbol{B} = \boldsymbol{B}_S + \boldsymbol{B}_I + \boldsymbol{B}_{\boldsymbol{\varepsilon}}. \quad (37)$$

Let us define the time course vector of the pth interference source, $\boldsymbol{\sigma}_p$, as

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\boldsymbol{\sigma}_p = \left[\sigma_p(t_1), \ldots, \sigma_p(t_K)\right]. \quad (38)$$

Then, the interference subspace in the time domain, $\mathscr{K}_I$, is defined as

$$\mathscr{K}_I = \mathrm{rsp}\left(\left[\boldsymbol{\sigma}_1^T, \ldots, \boldsymbol{\sigma}_P^T\right]^T\right). \quad (39)$$

We assume that time course vectors $\boldsymbol{\sigma}_1^T, \ldots, \boldsymbol{\sigma}_P^T$ are linearly independent in this paper.

If basis row vectors of the interference subspace in the time domain are estimated as $\psi_1, \ldots, \psi_P$, defining $\boldsymbol{\Upsilon}$ as $\boldsymbol{\Upsilon} = \left[\psi_1^T, \ldots, \psi_P^T\right]^T$, we can compute the projector onto $\mathscr{K}_I$ such that

$$\boldsymbol{\Pi}_I = \boldsymbol{\Upsilon}^T\left(\boldsymbol{\Upsilon}\boldsymbol{\Upsilon}^T\right)^{-1}\boldsymbol{\Upsilon}. \quad (40)$$

We denote the jth row of $\boldsymbol{B}_I$ by β_j^I. Since β_j^I is expressed as the linear sum of ψ_j: $\beta_j^I = \sum_{j=1}^P c_j \psi_j$, we thus have $\beta_j^I(\boldsymbol{I} - \boldsymbol{\Pi}_I) = 0$. Therefore, projecting the data matrix $\boldsymbol{B}$ onto the subspace orthogonal to $\mathscr{K}_I$, the estimated signal matrix $\widehat{\boldsymbol{B}}_S$ is given by

$$\begin{aligned} \widehat{\boldsymbol{B}}_S &= \boldsymbol{B}(\boldsymbol{I} - \boldsymbol{\Pi}_I) = (\boldsymbol{B}_S + \boldsymbol{B}_I + \boldsymbol{B}_\varepsilon)(\boldsymbol{I} - \boldsymbol{\Pi}_I) \\ &= \boldsymbol{B}_S - \boldsymbol{B}_S\boldsymbol{\Pi}_I + \boldsymbol{B}_\varepsilon(\boldsymbol{I} - \boldsymbol{\Pi}_l). \end{aligned} \quad (41)$$

The method of removing the interference $\boldsymbol{B}_I$ based on equation (41) is referred to as the time-domain signal space projection (time-domain SSP). The influence of the time domain SSP on the signal component is assessed by the second term $\boldsymbol{B}_S\boldsymbol{\Pi}_I$ on the right-hand side of equation (41). This term becomes small when the correlations between the time courses of the signal and interference sources are small. This can be considered an advantage of the time-domain SSP over the spatial-domain SSP. This is because in many real-life applications, the time courses of the signal and interference sources are expected to differ significantly, but the orthogonality of lead field vectors between signal and interference sources may not be so high. In the next section, we show that a number of existing interference removal methods can be interpreted as the time domain SSP, and that these methods differ only in their manner of deriving the basis vectors of the time-domain interference subspace.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

5. Interference removal based on the time-domain SSP

5.1. Adaptive noise canceling(ANC)

Adaptive noise canceling(ANC) is an interference removal method which makes use of data from reference sensors. The reference sensors collect data containing interference but not the signal of interest [6, 7, 18]. It is assumed that the sensor array is equipped with a total of J additional reference sensors, and the outputs of the reference sensors are denoted by a $J \times 1$ column vector $\boldsymbol{y}_R(t)$. Then, the data model is expressed as

$$\boldsymbol{y}(t) = \boldsymbol{y}_S(t) + \boldsymbol{y}_I(t) + \varepsilon, \quad (42)$$

$$\boldsymbol{y}_R(t) = \tilde{\boldsymbol{y}}_I(t) + \tilde{\varepsilon}. \quad (43)$$

Note that $\boldsymbol{y}(t)$, $\boldsymbol{y}_S(t)$, $\boldsymbol{y}_I(t)$ and $\boldsymbol{\varepsilon}$ are $\boldsymbol{M} \times 1$ column vectors, and $\boldsymbol{y}_R(t)$, $\tilde{\boldsymbol{y}}_I(t)$, and $\tilde{\varepsilon}$ are $\boldsymbol{J} \times 1$ column vectors. ANC tries to remove the interference $\boldsymbol{y}_I(t)$ from the data vector $\boldsymbol{y}(t)$ by taking out components maximally correlated with the reference sensor data $\boldsymbol{y}_R(t)$. This removal is carried out by regressing $\boldsymbol{y}(t)$ with $\boldsymbol{y}_R(t)$, i.e.

$$\boldsymbol{y}(t) = \boldsymbol{Z}\boldsymbol{y}_R(t) + \boldsymbol{d}(t), \quad (44)$$

where $\boldsymbol{Z}$ is an $M \times J$ coefficient matrix of this multi-variate regression, and the residual signal $\boldsymbol{d}(t)$ represents the interference removed results. Here, the coefficient matrix $\boldsymbol{Z}$ is obtained by solving the minimization problem:

$$\boldsymbol{Z} = \underset{\boldsymbol{Z}}{\operatorname{argmin}} \left\langle \left\| \boldsymbol{y}(t) - \boldsymbol{Z}\boldsymbol{y}_R(t) \right\|^2 \right\rangle, \quad (45)$$

where $\langle \cdot \rangle$ indicates time average. That is, $\boldsymbol{Z}$ is determined so as to maximize the correlation between $\boldsymbol{y}(t)$ and $\boldsymbol{y}_R(t)$. The interference-removed results are expressed as:

$$\boldsymbol{d}(t) = \boldsymbol{y}(t) - \widehat{\boldsymbol{Z}}\boldsymbol{y}_R(t) = \boldsymbol{y}(t) - \left\langle \boldsymbol{y}(t)\boldsymbol{y}_R^T(t) \right\rangle \left[\left\langle \boldsymbol{y}_R(t)\boldsymbol{y}_R^T(t) \right\rangle \right]^{-1} \boldsymbol{y}_R(t). \quad (46)$$

To understand the relationship between ANC and the time-domain SSP, Let us rewrite equation (46) using a matrix form. To do so, the matrix of the reference sensor data is defined as $\boldsymbol{B}_R$, such that

$$\boldsymbol{B}_R = \left[\boldsymbol{y}_R(t_1), \ldots, \boldsymbol{y}_R(t_K) \right].$$

Using the data matrices of $\boldsymbol{B}_R$ and $\boldsymbol{B}$, we have

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\left\langle \boldsymbol{y}(t)\boldsymbol{y}_R^T(t)\right\rangle = \frac{1}{K}\boldsymbol{B}\boldsymbol{B}_R^T, \quad \text{and} \quad \left\langle \boldsymbol{y}_R(t)\boldsymbol{y}_R^T(t)\right\rangle = \frac{1}{K}\boldsymbol{B}_R\boldsymbol{B}_R^T. \quad (47)$$

Thus, denoting the interference removed results as $\widehat{\boldsymbol{B}}_S$: $\widehat{\boldsymbol{B}}_S = \left[\boldsymbol{d}(t_1), \ldots, \boldsymbol{d}(t_K)\right]$, equation (46) is rewritten as

$$\begin{aligned}\widehat{\boldsymbol{B}}_S &= \boldsymbol{B} - \boldsymbol{B}\boldsymbol{B}_R^T\left(\boldsymbol{B}_R\boldsymbol{B}_R^T\right)^{-1}\boldsymbol{B}_R \\ &= \boldsymbol{B}\left[\boldsymbol{I} - \boldsymbol{B}_R^T\left(\boldsymbol{B}_R\boldsymbol{B}_R^T\right)^{-1}\boldsymbol{B}_R\right] = \boldsymbol{B}\left[\boldsymbol{I} - \boldsymbol{\Pi}_R\right],\end{aligned} \quad (48)$$

where $\boldsymbol{\Pi}_R$ indicates the projector onto the row space of $\boldsymbol{B}_R$ defined such that

$$\boldsymbol{\Pi}_R = \boldsymbol{B}_R^T\left(\boldsymbol{B}_R\boldsymbol{B}_R^T\right)^{-1}\boldsymbol{B}_R. \quad (49)$$

Comparison between equations (41) and (48) makes it clear that ANC is equivalent to the time-domain SSP if the relationship $\boldsymbol{\Pi}_I \approx \boldsymbol{\Pi}_R$ holds where $\boldsymbol{\Pi}_I$ is the projector onto the interference subspace defined in equation (40). Actually, if the sensor noise is small, the row space of $\boldsymbol{B}_R$ can reasonably approximate the row space of $\boldsymbol{B}_I$, which is equal to the interference subspace under the low rank signal assumption $J > Q$. However, when the sensor noise is not negligibly small or when the reference sensor data contains components not contained in the measurement sensor data, $\boldsymbol{\Pi}_R$ and $\boldsymbol{\Pi}_I$ may have some difference, and results from the ANC algorithm may contain errors.

5.2. Common temporal subspace projection (CTSP)

Common temporal subspace projection (CTSP) also removes the interference by making use of the reference sensor data [9]. The difference between CTSP and ANC is that CTSP assumes that the reference sensor data contain components that exist only in the reference sensor data but not in the measurement sensor data. The data model assumed in CTSP is expressed as

$$\boldsymbol{B} = \boldsymbol{B}_S + \boldsymbol{B}_I + \boldsymbol{B}_{\boldsymbol{\varepsilon}}, \quad (50)$$

$$\boldsymbol{B}_R = \widetilde{\boldsymbol{B}}_I + \boldsymbol{B}_w + \widetilde{\boldsymbol{B}}_{\boldsymbol{\varepsilon}}, \quad (51)$$

where a $J \times K$ matrix $\boldsymbol{B}_w$ indicates the components that are contained only in the reference sensor data, and $J \times K$ matrices, $\widetilde{\boldsymbol{B}}_I$ and $\widetilde{\boldsymbol{B}}_{\boldsymbol{\varepsilon}}$, are the interference and noise matrices of the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

reference sensor data. Here, we assume $\boldsymbol{B}_w \boldsymbol{B}_S^T = 0$. According to the arguments in appendix A.5, equation (50) leads to the relationship,

$$\mathrm{rsp}(\boldsymbol{B}) = \mathrm{rsp}\left(\boldsymbol{B}_S + \boldsymbol{B}_I + \boldsymbol{B}_{\boldsymbol{\varepsilon}}\right) \subset \mathrm{rsp}\left(\boldsymbol{B}_S\right) + \mathrm{rsp}\left(\boldsymbol{B}_I\right) + \mathrm{rsp}\left(\boldsymbol{B}_{\boldsymbol{\varepsilon}}\right) = \mathscr{K}_S + \mathscr{K}_I + \mathscr{K}_\varepsilon, \quad (52)$$

where $\mathscr{K}_\varepsilon$ indicates $\mathscr{K}_\varepsilon = rsp\left(\boldsymbol{B}_{\boldsymbol{\varepsilon}}\right)$. Equation (51) leads to

$$\mathrm{rsp}\left(\boldsymbol{B}_R\right) = \mathrm{rsp}\left(\tilde{\boldsymbol{B}}_I + \boldsymbol{B}_w + \tilde{\boldsymbol{B}}_{\boldsymbol{\varepsilon}}\right) \subset \mathrm{rsp}\left(\tilde{\boldsymbol{B}}_I\right) + \mathrm{rsp}\left(\boldsymbol{B}_w\right) + \mathrm{rsp}\left(\tilde{\boldsymbol{B}}_{\boldsymbol{\varepsilon}}\right) = \mathscr{K}_I + \mathscr{K}_w + \widetilde{\mathscr{K}}_\varepsilon, \quad (53)$$

where $\mathscr{K}_w$ and $\widetilde{\mathscr{K}}_\varepsilon$ indicate $\mathscr{K}_w = rsp\left(\boldsymbol{B}_w\right)$ and $\widetilde{\mathscr{K}}_\varepsilon = rsp\left(\tilde{\boldsymbol{B}}_{\boldsymbol{\varepsilon}}\right)$. The relationships $rsp\left(\boldsymbol{B}_S\right) = \mathscr{K}_S$ and $rsp\left(\boldsymbol{B}_I\right) = rsp\left(\tilde{\boldsymbol{B}}_I\right) = \mathscr{K}_I$ are used here. Then, according to the arguments in appendix A.6, equations (52) and (53) lead to the following relationships among sets of basis vectors:

$$\mathscr{S}_B \subset \mathscr{S}_S \cup \mathscr{S}_I \cup \mathscr{S}_\varepsilon \quad (54)$$

$$\mathscr{S}_{B_R} \subset \mathscr{S}_I \cup \mathscr{S}_w \cup \widetilde{\mathscr{S}}_\varepsilon \quad (55)$$

where sets of basis vectors of rsp($\boldsymbol{B}$) and rsp($\boldsymbol{B}_R$) are respectively denoted by $\mathscr{S}_B$, $\mathscr{S}_{B_R}$, i.e. $\mathrm{rsp}\left(\boldsymbol{B}\right) = \mathrm{span}\left(\mathscr{S}_B\right)$ and $\mathrm{rsp}\left(\boldsymbol{B}_R\right) = \mathrm{span}\left(\mathscr{S}_{B_R}\right)$[7]. Also, sets of basis vectors of $\mathscr{K}_S$, $\mathscr{K}_I$, $\mathscr{K}_w$, $\mathscr{K}_\varepsilon$, and $\widetilde{\mathscr{K}}_\varepsilon$ are respectively denoted by $\mathscr{S}_S$, $\mathscr{S}_I$, $\mathscr{S}_w$, $\mathscr{S}_\varepsilon$, and $\widetilde{\mathscr{S}}_\varepsilon$.

We assume that the noise time courses are orthogonal to each other, and they are orthogonal to the signal and interference time courses, resulting in the relationships: $\mathscr{S}_\varepsilon \cap \widetilde{\mathscr{S}}_\varepsilon = \varnothing$, $\mathscr{S}_\varepsilon \cap \mathscr{S}_S = \mathscr{S}_\varepsilon \cap \mathscr{S}_I = \mathscr{S}_\varepsilon \cap \mathscr{S}_w = \varnothing$, and $\widetilde{\mathscr{S}}_\varepsilon \cap \mathscr{S}_S = \widetilde{\mathscr{S}}_\varepsilon \cap \mathscr{S}_I = \widetilde{\mathscr{S}}_\varepsilon \cap \mathscr{S}_w = \varnothing$. Here $\varnothing$ indicates the empty set. We also use $\mathscr{S}_S \cap \mathscr{S}_w = \varnothing$, which results from the orthogonality assumption $\boldsymbol{B}_w \boldsymbol{B}_S^T = 0$. Then, using the proof in appendix A.7 and the distributive property, we can derive

$$\begin{aligned}
&\mathscr{S}_B \cap \mathscr{S}_{B_R} \subset \left(\mathscr{S}_S \cup \mathscr{S}_I \cup \mathscr{S}_\varepsilon\right) \cap \left(\mathscr{S}_I \cup \mathscr{S}_w \cup \widetilde{\mathscr{S}}_\varepsilon\right) \\
&= \left[\mathscr{S}_S \cap \left(\mathscr{S}_I \cup \mathscr{S}_w \cup \widetilde{\mathscr{S}}_\varepsilon\right)\right] \cup \left[\mathscr{S}_I \cap \left(\mathscr{S}_I \cup \mathscr{S}_w \cup \widetilde{\mathscr{S}}_\varepsilon\right)\right] \cup \left[\mathscr{S}_\varepsilon \cap \left(\mathscr{S}_I \cup \mathscr{S}_w \cup \widetilde{\mathscr{S}}_\varepsilon\right)\right] \quad (56) \\
&= \left(\mathscr{S}_S \cap \mathscr{S}_I\right) \cup \mathscr{S}_I \cup \varnothing = \left(\mathscr{S}_S \cap \mathscr{S}_I\right) \cup \mathscr{S}_I = \mathscr{S}_I .
\end{aligned}$$

[7]Here, an expression of $\mathscr{X} = \mathrm{span}\left(\mathscr{S}_X\right)$ indicates that a set of basis vectors $\mathscr{S}_X$ spans the subspace $\mathscr{X}$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Namely,

$$\mathcal{S}_I \supset \mathcal{S}_B \cap \mathcal{S}_{B_R} \quad (57)$$

holds. According to the arguments in appendix A.8, we can obtain the relationship,

$$\begin{aligned}\mathcal{K}_I &= \mathrm{span}(\mathcal{S}_I) \supset \mathrm{span}\left(\mathcal{S}_B \cap \mathcal{S}_{B_R}\right) = \mathrm{span}(\mathcal{S}_B) \cap \mathrm{span}\left(\mathcal{S}_{B_R}\right) \\ &= \mathrm{rsp}(\boldsymbol{B}) \cap \mathrm{rsp}\left(\boldsymbol{B}_R\right).\end{aligned} \quad (58)$$

The equation above claims that the intersection between rsp($\boldsymbol{B}$) and rsp($\boldsymbol{B}_R$) forms a subset of the interference subspace $\mathcal{K}_I$. An algorithm that derives basis vectors of the intersection between two subspaces is described in appendix A.9. Using this algorithm, the orthonormal basis vectors of the intersection rsp($\boldsymbol{B}$) ∩ rsp ($\boldsymbol{B}_R$) can be obtained. Denoting these basis vectors as $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_r$, we can remove the interference and obtain the estimated signal matrix $\widehat{\boldsymbol{B}}_S$ such that

$$\widehat{\boldsymbol{B}}_S = \boldsymbol{B}\left(\boldsymbol{I} - [\psi_1, \ldots, \psi_r][\psi_1, \ldots, \psi_r]^T\right). \quad (59)$$

The time domain SSP in equation (59) cannot perfectly remove the interference because the basis vectors $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_r$ span only a part of the interference subspace. Nonetheless, we can still expect that the method can reduce the interference if rsp($\boldsymbol{B}$) ∩ rsp($\boldsymbol{B}_R$) is a reasonable approximation of $\mathcal{K}_I$. The algorithm that performs the interference removal in a manner described above is called common temporal subspace projection (CTSP) [9].

5.3. Dual signal subspace projection (DSSP)

5.3.1. Pseudo signal subspace projector.—Dual signal subspace projection (DSSP) removes the interference without using either the reference sensor data or control data such as the empty-room-noise data [11]. The algorithm assumes the data model in equation (50) with the assumption that the interference sources are located outside the source space. The DSSP algorithm uses the so-called pseudo signal subspace projector, and to derive it, voxels are defined over the source space Ω, in which the voxel locations are denoted $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$. The pseudo signal subspace $\breve{\mathcal{E}}_S$ is defined such that

$$\breve{\mathcal{E}}_S = \mathrm{csp}\left([\boldsymbol{L}(\boldsymbol{r}_1), \ldots, \boldsymbol{L}(\boldsymbol{r}_N)]\right). \quad (60)$$

If the voxel interval is sufficiently small and voxel discretization errors are negligible, we have the relationship $\breve{\mathcal{E}}_S \supset \mathcal{E}_S$. Therefore, a vector contained in the signal subspace is also contained in the pseudo signal subspace, i.e. if $\boldsymbol{x} \in \mathcal{E}_S$, then $\boldsymbol{x} \in \breve{\mathcal{E}}_S$.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Let us derive the orthonormal basis vectors of the pseudo signal subspace $\breve{\mathscr{E}}_S$. To do so, defining a matrix $\boldsymbol{F}$ as $\boldsymbol{F} = [\boldsymbol{L}(\boldsymbol{r}_1), \boldsymbol{L}(\boldsymbol{r}_N)]$, we compute the singular value decomposition of $\boldsymbol{F}$,

$$\boldsymbol{F} = \sum_{j=1}^{M} \lambda_j \boldsymbol{e}_j \boldsymbol{f}_j^T, \quad (61)$$

where we assume the relationship $M < N$. If the singular values $\lambda_1, \ldots, \lambda_\tau$ are distinctively large, and other singular values $\lambda_{\tau+1}, \ldots, \lambda_M$ are nearly equal to zero, the singular vectors $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_\tau$ forms orthonormal basis vectors of the pseudo signal subspace $\breve{\mathscr{E}}_S$ and the projector onto $\breve{\mathscr{E}}_S$ is obtained using

$$\breve{\boldsymbol{P}}_S = [\boldsymbol{e}_1, \ldots, \boldsymbol{e}_\tau][\boldsymbol{e}_1, \ldots, \boldsymbol{e}_\tau]^T. \quad (62)$$

Note that, since $\breve{\mathscr{E}}_S \supset \mathscr{E}_S$, the orthogonal projector $(\boldsymbol{I} - \breve{\boldsymbol{P}}_S)$ projects out the signal vector, i.e. $(\boldsymbol{I} - \breve{\boldsymbol{P}}_S)\boldsymbol{y}_S(t) = (\boldsymbol{I} - \breve{\boldsymbol{P}}_S)\boldsymbol{B}_S = 0$.

5.3.2. DSSP algorithm.—The DSSP algorithm applyies $\breve{\boldsymbol{P}}_S$ and $\boldsymbol{I} - \breve{\boldsymbol{P}}_S$ to the data matrix $\boldsymbol{B}$ to create two kinds of datasets:

$$\breve{\boldsymbol{P}}_S \boldsymbol{B} = \boldsymbol{B}_S + \breve{\boldsymbol{P}}_S \boldsymbol{B}_I + \breve{\boldsymbol{P}}_S \boldsymbol{B}_\varepsilon, \quad (63)$$

$$(\boldsymbol{I} - \breve{\boldsymbol{P}}_S)\boldsymbol{B} = (\boldsymbol{I} - \breve{\boldsymbol{P}}_S)\boldsymbol{B}_I + (\boldsymbol{I} - \breve{\boldsymbol{P}}_S)\boldsymbol{B}_\varepsilon. \quad (64)$$

In equation (64), the signal of interest is suppressed by multiplying $(\boldsymbol{I} - \breve{\boldsymbol{P}}_S)$ with the data matrix, and a virtual reference time series is estimated[8]. Here, we use $\breve{\boldsymbol{P}}_S \boldsymbol{B}_S = \boldsymbol{B}_S$ and $(\boldsymbol{I} - \breve{\boldsymbol{P}}_S)\boldsymbol{B}_S = 0$. According to the arguments in appendix A.5, the following relationships hold:

$$\mathrm{rsp}(\breve{\boldsymbol{P}}_S \boldsymbol{B}) \subset \mathrm{rsp}(\boldsymbol{B}_S) + \mathrm{rsp}(\breve{\boldsymbol{P}}_S \boldsymbol{B}_I) + \mathrm{rsp}(\breve{\boldsymbol{P}}_S \boldsymbol{B}_\varepsilon), \quad (65)$$

$$\mathrm{rsp}((\boldsymbol{I} - \breve{\boldsymbol{P}}_S)\boldsymbol{B}) \subset \mathrm{rsp}((\boldsymbol{I} - \breve{\boldsymbol{P}}_S)\boldsymbol{B}_I) + \mathrm{rsp}((\boldsymbol{I} - \breve{\boldsymbol{P}}_S)\boldsymbol{B}_\varepsilon). \quad (66)$$

[8]This procedure is called the generalized sidelobe canceller in [19].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The arguments in appendix A.3 prove that the relationships, $rsp\left(\check{\boldsymbol{P}}_S\boldsymbol{B}_I\right) = \mathscr{K}_I$ and $rsp\left(\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}_I\right) = \mathscr{K}_I$ hold. Thus, using $\mathrm{rsp}\left(\boldsymbol{B}_S\right) = \mathscr{K}_S$, equations (65) and (66) lead to

$$\mathrm{rsp}\left(\check{\boldsymbol{P}}_S\boldsymbol{B}\right) \subset \mathscr{K}_S + \mathscr{K}_I + \check{\mathscr{K}}_\varepsilon, \quad (67)$$

$$\mathrm{rsp}\left(\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}\right) \subset \mathscr{K}_I + \check{\mathscr{K}}'_\varepsilon, \quad (68)$$

where we use the notations, $\mathrm{rsp}\left(\check{\boldsymbol{P}}_S\boldsymbol{B}_{\boldsymbol{\varepsilon}}\right) = \check{\mathscr{K}}_\varepsilon$ and $\mathrm{rsp}\left(\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}_{\boldsymbol{\varepsilon}}\right) = \check{\mathscr{K}}'_\varepsilon$. Let us define the sets of basis vectors of $\mathrm{rsp}\left(\check{\boldsymbol{P}}_S\boldsymbol{B}\right)$ and $\mathrm{rsp}\left(\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}\right)$ as $\mathscr{S}_{P_S}$ and $\mathscr{S}_{P_I}$, (i.e. $rsp\left(\check{\boldsymbol{P}}_S\boldsymbol{B}\right) = \mathrm{span}\left(\mathscr{S}_{P_S}\right)$ and $\mathrm{rsp}\left(\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}\right)$. According to the arguments in appendix A.6, equations (67) and (68) can be converted into relationships among the sets of basis vectors, which are written as

$$\mathscr{S}_{P_S} \subset \mathscr{S}_S \cup \mathscr{S}_I \cup \check{\mathscr{S}}_\varepsilon, \quad (69)$$

$$\mathscr{S}_{P_I} \subset \mathscr{S}_I \cup \check{\mathscr{S}}'_\varepsilon, \quad (70)$$

where basis vector sets of $\check{\mathscr{K}}_\varepsilon$ and $\check{\mathscr{K}}'_\varepsilon$ are denoted $\check{\mathscr{S}}_\varepsilon$ and $\check{\mathscr{S}}'_\varepsilon$.

Since the relationship

$$\check{\boldsymbol{P}}_S\boldsymbol{B}_{\boldsymbol{\varepsilon}}\left(\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}_{\boldsymbol{\varepsilon}}\right)^T = \check{\boldsymbol{P}}_S\boldsymbol{B}_{\boldsymbol{\varepsilon}}\boldsymbol{B}_{\boldsymbol{\varepsilon}}^T\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right) = \rho^2\check{\boldsymbol{P}}_S\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right) = 0,$$

holds, $\check{\mathscr{S}}_\varepsilon$ and $\check{\mathscr{S}}'_\varepsilon$ has no intersection, i.e. $\check{\mathscr{S}}_\varepsilon \cap \check{\mathscr{S}}'_\varepsilon = \varnothing$. Also, assuming that $\boldsymbol{B}_S\boldsymbol{B}_{\boldsymbol{\varepsilon}}^T = \boldsymbol{B}_I\boldsymbol{B}_{\boldsymbol{\varepsilon}}^T$, the relationships $\check{\mathscr{S}}_\varepsilon \cap \mathscr{S}_S = \check{\mathscr{S}}_\varepsilon \cap \mathscr{S}_I = \varnothing$ and $\check{\mathscr{S}}'_\varepsilon \cap \mathscr{S}_S = \check{\mathscr{S}}'_\varepsilon \cap \mathscr{S}_I = \varnothing$ hold. Therefore, using the proof in appendix A.7 and the distributive property, we can derive the following relationship:

$$\begin{aligned}
&\mathscr{S}_{P_S} \cap \mathscr{S}_{P_I} \subset \left[\mathscr{S}_S \cup \mathscr{S}_I \cup \check{\mathscr{S}}_\varepsilon\right] \cap \left[\mathscr{S}_I \cup \check{\mathscr{S}}'_\varepsilon\right] \\
&= \left[\mathscr{S}_S \cap \left(\mathscr{S}_I \cup \check{\mathscr{S}}'_\varepsilon\right)\right] \cup \left[\mathscr{S}_I \cap \left(\mathscr{S}_I \cup \check{\mathscr{S}}'_\varepsilon\right)\right] \cup \left[\check{\mathscr{S}}_\varepsilon \cap \left(\mathscr{S}_I \cup \check{\mathscr{S}}'_\varepsilon\right)\right] \quad (71) \\
&= \left(\mathscr{S}_S \cap \mathscr{S}_I\right) \cup \mathscr{S}_I \cup \varnothing = \left(\mathscr{S}_S \cap \mathscr{S}_I\right) \cup \mathscr{S}_I = \mathscr{S}_I.
\end{aligned}$$

Namely, we have

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\mathcal{S}_I \supset \mathcal{S}_{P_S} \cap \mathcal{S}_{P_I}. \quad (72)$$

According to the arguments in appendix A.8, we derive

$$\begin{aligned} \mathcal{K}_I &= \mathrm{span}(\mathcal{S}_I) \supset \mathrm{span}\left(\mathcal{S}_{P_S} \cap \mathcal{S}_{P_l}\right) = \mathrm{span}\left(\mathcal{S}_{P_S}\right) \cap \mathrm{span}\left(\mathcal{S}_{P_I}\right) \\ &= \mathrm{rsp}\left(\check{\boldsymbol{P}}_S \boldsymbol{B}\right) \cap \mathrm{rsp}\left(\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}\right). \end{aligned} \quad (73)$$

The quation above shows that the intersection between $\mathrm{rsp}\left(\check{\boldsymbol{P}}_S \boldsymbol{B}\right)$ and $\mathrm{rsp}\left(\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}\right)$ forms a subset of the interference subspace $\mathcal{K}_I$. The basis vectors of the intersection are derived using the algorithm described in appendix A.9. Once the orthonornal basis vectors of the intersection $\mathrm{rsp}\left(\check{\boldsymbol{P}}_S \boldsymbol{B}\right) \cap \mathrm{rsp}\left(\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}\right)$ are obtained, we can compute the projector onto the intersection, and time-domain SSP can be implemented. If the intersection $\mathrm{rsp}\left(\check{\boldsymbol{P}}_S \boldsymbol{B}\right) \cap \mathrm{rsp}\left(\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}\right)$ is a reasonable approximation of $\mathcal{K}_I$, this time-domain SSP will be able to remove interferences effectively. The method of removing the interference in a manner as described above is called dual signal subspace projection (DSSP) [11].

Note that since the intersection $\mathrm{rsp}\left(\check{\boldsymbol{P}}_S \boldsymbol{B}\right) \cap \mathrm{rsp}\left(\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}\right)$ is only a subset of the interference subspace $\mathcal{K}_I$, the method cannot perfectly remove interferences. However, we can observe that $\mathrm{rsp}\left(\check{\boldsymbol{P}}_S \boldsymbol{B}\right) \cap \mathrm{rsp}\left(\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}\right)$ becomes a better approximation of $\mathcal{K}_I$, if the interference $\boldsymbol{B}_I$ is significantly greater than the signal $\boldsymbol{B}_S$ (and the sensor noise $\boldsymbol{B}_{\boldsymbol{\varepsilon}}$), resulting in a situation that the interference terms are dominated in $\check{\boldsymbol{P}}_S \boldsymbol{B}$ and $\left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)\boldsymbol{B}$. This may explain our empirical findings, which are a little counter-intuitive, that the method works better for larger interferences.

5.4. Spatio-temporal signal space separation (tSSS)

In section 3.3, we argue that one prerequisite of the signal space separation (SSS) algorithm is that all interference sources be located in the external region. However, this requirement is not always fulfilled, as interference sources can also be located fairly close to the source space. In this case, the SSS extractor $\boldsymbol{\Gamma}_S$ cannot adequately remove the interference. Spatio-temporal signal space separation algorithm (tSSS) has been developed for such situations [10].

This algorithm first applies the SSS extractors $\boldsymbol{\Gamma}_S$ and $\boldsymbol{\Gamma}_I$ to the data matrix $\boldsymbol{B}$ to create two kinds of data sets:

$$\boldsymbol{\Gamma}_S \boldsymbol{B} = \boldsymbol{B}_S + \boldsymbol{\Gamma}_S \boldsymbol{B}_I + \boldsymbol{\Gamma}_S \boldsymbol{B}_{\boldsymbol{\varepsilon}}, \quad (74)$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\boldsymbol{\Gamma}_I \boldsymbol{B} = \boldsymbol{\Gamma}_I \boldsymbol{B}_I + \boldsymbol{\Gamma}_I \boldsymbol{B}_{\boldsymbol{\varepsilon}}, \quad (75)$$

where we use the relationships, $\boldsymbol{\Gamma}_S \boldsymbol{B}_S = \boldsymbol{B}_S$, and $\boldsymbol{\Gamma}_I \boldsymbol{B}_S = 0$. We thus obtain,

$$\mathrm{rsp}(\boldsymbol{\Gamma}_S \boldsymbol{B}) \subset \mathscr{K}_S + \mathrm{rsp}(\boldsymbol{\Gamma}_S \boldsymbol{B}_I) + \mathrm{rsp}(\boldsymbol{\Gamma}_S \boldsymbol{B}_{\boldsymbol{\varepsilon}}), \quad (76)$$

$$\mathrm{rsp}(\boldsymbol{\Gamma}_I \boldsymbol{B}) \subset \mathrm{rsp}(\boldsymbol{\Gamma}_I \boldsymbol{B}_I) + \mathrm{rsp}(\boldsymbol{\Gamma}_I \boldsymbol{B}_{\boldsymbol{\varepsilon}}), \quad (77)$$

where we use $\mathrm{rsp}(\boldsymbol{B}_S) = \mathscr{K}_S$.

When an interference source is located close to the source space, the lead field vector of this source, $\boldsymbol{\xi}_p$, may have components expanded by the columns of $\boldsymbol{C}$, as well as components expanded by the columns of $\boldsymbol{D}$, resulting in,

$$\boldsymbol{\xi}_p = \boldsymbol{C}\boldsymbol{\alpha}' + \boldsymbol{D}\phi', \quad (78)$$

where $\boldsymbol{\alpha}'$ and $\boldsymbol{\phi}'$ are vectors containing the expansion coefficients. Therefore, by multiplying the SSS signal extractor to $\boldsymbol{\xi}_p$, we have

$$\boldsymbol{\Gamma}_S \boldsymbol{\xi}_p = \boldsymbol{\Gamma}_S [\boldsymbol{C}\boldsymbol{\alpha}' + \boldsymbol{D}\phi'] = \boldsymbol{C}\boldsymbol{\alpha}' = \boldsymbol{\pi}_p, \quad (79)$$

which shows that $\boldsymbol{\Gamma}_S$ changes the interference lead field from $\boldsymbol{\xi}_p$ to $\boldsymbol{\pi}_p$. Similarly, by applying $\boldsymbol{\Gamma}_I$ to ξ_p, we have

$$\boldsymbol{\Gamma}_I \boldsymbol{\xi}_p = \boldsymbol{\Gamma}_I [\boldsymbol{C}\boldsymbol{\alpha}' + \boldsymbol{D}\phi'] = \boldsymbol{D}\phi' = \tilde{\boldsymbol{\pi}}_p, \quad (80)$$

which shows that the lead field $\boldsymbol{\xi}_p$ is changed to $\boldsymbol{\pi}_p$.

However, these extractors never change the time courses of interference sources. Thus, using the same arguments as in appendix A.3, we can prove that $\mathrm{rsp}(\boldsymbol{\Gamma}_S \boldsymbol{B}_I) = \mathscr{K}_I$ and $\mathrm{rsp}(\boldsymbol{\Gamma}_I \boldsymbol{B}_I) = \mathscr{K}_I$ hold, and equations (76) and (77) become

$$\mathrm{rsp}(\boldsymbol{\Gamma}_S \boldsymbol{B}) \subset \mathscr{K}_S + \mathscr{K}_I + \mathscr{K}'_{\varepsilon}, \quad (81)$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\mathrm{rsp}\left(\boldsymbol{\Gamma}_I \boldsymbol{B}\right) \subset \mathscr{K}_I + \mathscr{K}''_\varepsilon . \quad (82)$$

Sets of basis vectors of rsp($\boldsymbol{\Gamma}_S\boldsymbol{B}$) and rsp($\boldsymbol{\Gamma}_I\boldsymbol{B}$) are respectively denoted by $\mathscr{S}_{\Gamma_S}$ and $\mathscr{S}_{\Gamma_I}$. Then, according to the arguments in appendix A.6, we can obtain

$$\mathscr{S}_{\Gamma_S} \subset \mathscr{S}_S \cup \mathscr{S}_I \cup \mathscr{S}'_\varepsilon, \quad (83)$$

$$S_{\Gamma_l} \subset \mathscr{S}_I \cup \mathscr{S}''_\varepsilon, \quad (84)$$

where sets of basis vectors of $\mathscr{K}_S$, $\mathscr{K}_I$, $\mathscr{K}'_\varepsilon$, and $\mathscr{K}''_\varepsilon$ are respectively denoted by $\mathscr{S}_S$,$\mathscr{S}_I$, $\mathscr{S}'_\varepsilon$, and $\mathscr{S}''_\varepsilon$. Asumuing that relationships, $\mathscr{S}_S \cap \mathscr{S}''_\varepsilon = \mathscr{S}_I \cap \mathscr{S}''_\varepsilon = \varnothing$, and $\mathscr{S}_I \cap \mathscr{S}'_\varepsilon = \mathscr{S}''_\varepsilon \cap \mathscr{S}'_\varepsilon = \varnothing$ hold, we can finally obtain

$$\mathscr{S}_I \supset \mathscr{S}_{\Gamma_S} \cap \mathscr{S}_{\Gamma_I}, \quad (85)$$

and according to the arguments in appendix A.8, we derive the relationship:

$$\mathscr{K}_I \supset \mathrm{rsp}\left(\boldsymbol{\Gamma}_S \boldsymbol{B}\right) \cap \mathrm{rsp}\left(\boldsymbol{\Gamma}_I \boldsymbol{B}\right). \quad (86)$$

The basis vectors of the intersection rsp($\boldsymbol{\Gamma}_S\boldsymbol{B}$) ∩ rsp($\boldsymbol{\Gamma}_I\boldsymbol{B}$) are obtained using the algorithm in appendix A.9, and they can be used for forming the projector onto the interference subspace. Making use of the interference subspace projector obtained in this manner, the tSSS algorithm performs the time-domain SSP for interference removal. Again, although these basis vectors span only a subset of the interference subspace, the method can effectively remove the interference if the intersection rsp($\boldsymbol{\Gamma}_S\boldsymbol{B}$) ∩ rsp($\boldsymbol{\Gamma}_I\boldsymbol{B}$) is a reasonable approximation of $\mathscr{K}_I$. One such case is that the interference terms are dominated in the data sets $\boldsymbol{\Gamma}_S\boldsymbol{B}$ and $\boldsymbol{\Gamma}_I\boldsymbol{B}$.

5.5. Sensor noise suppression (SNS)

The sensor noise suppression (SNS) algorithm has been developed to suppress sensor noise [8]. The algorithm assumes the data model

$$\boldsymbol{B} = \boldsymbol{B}_S + \boldsymbol{B}_{\boldsymbol{\varepsilon}} . \quad (87)$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

A key assumption of the SNS algorithm is that $\boldsymbol{\beta}_j$, the jth row of $\boldsymbol{B}$, lies within a span of all the rows of $\boldsymbol{B}$ except the jth row, i.e.

$$\beta_j \in \mathrm{rsp}\left(\left[\beta_1^T, \ldots, \beta_{j-1}^T, \beta_{j+1}^T, \ldots, \beta_M^T\right]^T\right). \quad (88)$$

Thus, ($\boldsymbol{\beta}_j$ is expressed as the linear sum of the other rows:

$$\beta_j^T = \sum_{i \neq j} \omega_i \beta_i^T = \boldsymbol{\Theta}_j \boldsymbol{\omega}_j, \quad (89)$$

where

$$\boldsymbol{\Theta}_j = \left[\beta_1^T, \ldots, \beta_{j-1}^T, \beta_{j+1}^T, \ldots, \beta_M^T\right], \quad (90)$$

$$\boldsymbol{\omega}_j = \left[\omega_1, \ldots, \omega_{j-1}, \omega_{j+1}, \ldots, \omega_M\right]^T. \quad (91)$$

In the equations above, the notation $\sum_{i \neq j}$ indicates the summation from $i = 1$ to M, except $i = j$, and ω_i ($i = 1, \ldots, j-1, j+1, \ldots, M$) are weights of the linear summation.

The optimum weight, $\hat{\omega}_j$, can be obtained using least squares fitting:

$$\hat{\boldsymbol{\omega}}_j = \underset{\boldsymbol{\omega}_j}{\mathrm{argmin}} \left\| \beta_j^T - \boldsymbol{\Theta}_j \boldsymbol{\omega}_j \right\|^2.$$

The solution is

$$\hat{\boldsymbol{\omega}}_j = \left(\boldsymbol{\Theta}_j^T \boldsymbol{\Theta}_j\right)^{-1} \boldsymbol{\Theta}_j^T \beta_j^T. \quad (92)$$

This $\hat{\omega}_j$ is substituted into equation (89), and the denoised row $\hat{\beta}_j$ is obtained as

$$\hat{\beta}_j = \beta_j \boldsymbol{\Theta}_j \left(\boldsymbol{\Theta}_j^T \boldsymbol{\Theta}_j\right)^{-1} \boldsymbol{\Theta}_j^T. \quad (93)$$

It can be seen in equation (93) that the sensor noise suppression algorithm is the time-domain SSP, assuming that rsp(Θ_j) approximates the time-domain signal subspace $\mathscr{K}_S$ and the projector $\Theta_j\left(\Theta_j^T \Theta_j\right)^{-1} \Theta_j^T$ approximates the projector onto $\mathscr{K}_S$.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

When the number of sensors are large, rsp(Θ_J) should reasonably approximate rsp($\boldsymbol{B}$). When the sensor noise is negligibly small, rsp($\boldsymbol{B}$) then approximates rsp($\boldsymbol{B}_S$), which is equal to $\mathscr{K}_S$, resulting in the projector $\Theta_j\left(\Theta_j^T\Theta_j\right)^{-1}\Theta_j^T$ that approximates the projector onto $\mathscr{K}_S$. However, when the sensor noise is not small, (which is exactly the case where the SNS algorithm is needed,) rsp($\boldsymbol{B}$) differs from rsp($\boldsymbol{B}_S$), and the projector $\Theta_j\left(\Theta_j^T\Theta_j\right)^{-1}\Theta_j^T$ may differ from the signal subspace projector, resulting in the low performance of the SNS algorithm in such cases.

6. Numerical examples

6.1. Data generation

A series of computer simulations were carried out to provide illustrative examples of the results of our arguments in the preceding sections. A sensor alignment of the 275-channel whole-head sensor array from the Omega™ (VMS Medtech, Coquitlam, Canada) neuromagnetometer was used. The configuration of the sensor array arranged on the helmet surface is shown in figure 1, in which the filled circles indicate the locations of sensors. In our experiments, all these sensors were assumed to be magnetometers. We assumed six reference sensors, consisting of two sets of vector magnetometers in addition to the measurement sensors on the helmet, and the reference sensor arrangement is shown also in figure 1[9].

A single dipole source, which generated the signal magnetic field, was assumed to exist at (0,0,10) with the orientation of (1,1,0). The location was 7 cm below the center of the sensor array. The time course assigned to this source is shown in figure 2(a). In our numerical experiments, the time t is expressed with the unit of time points ranging from $t = -1200$ to 1200.

We assumed four interference sources whose coordinates are listed in the top four rows of table 1, which indicates that these interference sources were 500–800 cm away from the signal source. The time courses assigned to these four interference sources are shown in figure 2(b).

Sensor time courses were generated by projecting the source time course in figure 2(a) onto the sensor time courses through the lead field computed using Sarvas' formula [20]. The sensor time courses of the signal magnetic field (after adding sensor noise) are shown in figure 3(a). The magnetic field map across sensors at $t = 520$ is shown in figure 3(b). (The instant at $t = 520$ is shown by the broken vertical line in figure 2(a).) Note that the sensor time courses in figure 3(a) and the field map in figure 3(b) can work as the ground truth in interference removal experiments described below.

Also, the time courses of the interference sources in figure 2(b) were converted to the sensor time courses of the interference using the lead field computed from the Biot-Savart law [18,

[9]Note that this arrangement of the reference sensors is assumed solely for our numerical experiments, and differs from the true arrangement in the Omega neuromagnetometer system.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

20]. The use of the Biot-Savart law here is due to our assumption that these interference sources are non-biological sources. The interference magnetic field was overlapped onto the signal magnetic field with the interference-to-signal ratio (ISR) equal to 6. Here, the ISR is defined as $\|\boldsymbol{B}_I\|/\|\boldsymbol{B}_S\|$ where $\|\boldsymbol{X}\|$ indicates the Frobenius norm of a matrix $\boldsymbol{X}$. The interference-overlapped sensor time courses are shown in figure 3(c). Since the interference is much stronger than the signal magnetic field, the sensor time courses are dominated by the interference. The magnetic field map of the interference-overlapped sensor data is shown in figure 3(d).

6.2. Results of interference removal experiments

6.2.1. Experiments on signal space projection (SSP).—We applied the signal space projecton (SSP) algorithm described in section 3.2 to the interference-overlapped sensor data in figure 3(c). The SSP algorithm requires control data, (which contain only interference) to estimate the interference subspace. Such control data, generated using a set of different interference-source time courses, are shown in figure 4(a). The results of SSP interference removal are shown in figure 4(b), indicating that the interference has been largely removed. The field map of the interference-removed sensor data is shown in figure 4(c). The comparison between the resultant field map in figure 4(c) and the map of the signal-only magnetic field in figure 3(b) indicates that a considerable amount of signal distortion occurs in the SSP interference-removed results.

6.2.2. Experiments on adaptive noise canceling (ANC) and common temporal subspace projection (CTSP).—Adaptive noise canceling (ANC) and common temporal subspace projection (CTSP) were applied to the same interference-overlapped sensor data in figure 3(c). In these experiments, the reference sensor data shown in figure 5(a) were used, and the interference removal results are shown in figure 6. According to the resultant sensor time courses in figures 6(a) and (c), these methods removed most of the interference. The maps of the interference-removed results in figures 6(b) and (d) indicate that almost no signal distortion was caused.

Next, we performed experiments in which the reference sensor data contained, in addition to the interference, a fluctuation that did not exist in the data from the measurement sensors arranged on the helmet. In these experiments, the reference sensor data shown in figure 5(b) were used, and the results of ANC interference removal are shown in figure 7(a). These results show that a fairly large amount of interference remained due to the incorrect reference sensor data. The results of CTSP interference removal are shown in figure 7(b). Here, the intereferncе was nearly completely removed despite the fact that the reference sensor data contained an additional component. This is because the data model of the CTSP algorithm allows components to exist only in the reference data but not in the measured data.

6.2.3. Experiments on signal space separation (SSS).—The signl space separation (SSS) algorithm was tested on the interference-overlapped sensor data in figure 3(c). This algorithm requires neither control noise data nor reference sensor data. It solely relies on the spatial separation between signal and intereference sources. The resultant interference-removed sensor time courses are shown in figure 8(a), and the field map of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

these sensor data at $t = 520$ is shown in figure 8(b). These results show that the SSS algorithm nearly perfectly removed the interference. This is because the prerequisite of the SSS method, the conditions that signal sources be located within the internal region, and interference sources be located within the external region, was nearly perfectly satisfied in these numerical experiments.

We then moved two of the interference sources closer to the signal source by assigning the coordinates in the last two rows of table 1 to these sources. The distances between the new locations of the two interference sources and that of the signal source were approximately 40 cm, and the spatial separation between the signal and intereference sources could be insufficient. The sensor data overlapped by the interference, generated from the nearby interference sources, are shown in figure 9(a), and the field map is shown in figure 9(b). The SSS algorithm was applied to these interference-overlapped data, and the results are shown in figures 9(c) and (d). In these experiments, the SSS method failed to remove the intereference. This is because the two interference sources were located fairly close to the signal source, resulting in the violation of the prerequisite for the SSS method.

6.2.4. Experiments on spatio-temporal signal space separation (tSSS) and dual signal subspace projection (DSSP).—The spatio-temporal signal space separation (tSSS) algorithm was applied to the interference-overlapped sensor data in figure 3(c), and to the nearby interference data in figure 9(a). The interference-removed results are shown in figure 10. The results demonstrate that the tSSS algorithm can effectively remove the interferences not only caused by distant sources but also caused by nearby sources. The field maps in figures 10(b) and (d) demonstrate that almost no signal distortion was caused in the interference removal process.

The dual signal subspace projection (DSSP) algorithm was applied to the sensor data in figure 3(c) and the data in figure 9(a). The interference-removal results are shown in figure 11. The results demonstrate that the DSSP algorithm is also effective in the case of nearby interference sources, as well as in the case of distant interference sources. The field maps in figures 11(b) and (d) demonstrate that the DSSP interference removal results are free of signal distortion.

6.2.5. Experiments on sensor noise suppression (SNS).—The sensor noise suppression (SNS) algorithm was also tested with noisy sensor data. Gaussian noise was added to the signal magnetic data in figure 3(a). Here, the signal to noise ratio (SNR) was set to 10 where the SNR was defined as $\|\boldsymbol{B}_S\|/\|\boldsymbol{B}_\varepsilon\|$. The time courses of resultant noisy sensor data are shown in figure 12(a) and the field map is in (b). The SNS denoised results are shown in figure 12(c). Comparison between the sensor time courses before SNS denoising in figure 12(a) and after denoising in (c) shows a significant reduction in sensor noise through SNS denoising. The denoised magnetic field map in figure 12(d) indicates that no observable distortion is caused.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

7. Summary

This paper reviews subspace-based interference removal methods with an emphasis on the time domain signal subspace. We first provide a concise review on the conventional spatial-domain signal subspace and the signal space projection (SSP) method. We then extend the notion of signal subspace to the time domain by proposing a novel definition of the time-domain signal subspace, and introduce time domain signal space projection. We show that many existing interference removal methods can be interpreted as a form of the time domain SSP. These methods include adaptive noise canceling, sensor noise suppression, common temporal subspace projection, spatio-temporal signal space separation and the recently proposed dual signal subspace projection. We also show that the difference between these methods is primarily due to their manner of deriving the interference subspace. Numerical examples that illustrate the results of our arguments are provided.

Appendix A.: Supplementary mathematical arguments

A.1. Proof of equation (18): $\mathscr{E}_S = \mathrm{csp}(\boldsymbol{B}_S)$

We first prove that $\mathrm{csp}(\boldsymbol{B}_S) = \mathrm{csp}\left(\left[\boldsymbol{y}_1^S, \ldots, \boldsymbol{y}_K^S\right]\right) \supset \mathscr{E}_S$. To do so, we show that if $\boldsymbol{x} \in \mathscr{E}_S$, $\boldsymbol{x} \in \mathrm{csp}\left(\left[\boldsymbol{y}_1^S, \ldots, \boldsymbol{y}_K^S\right]\right) \supset \mathscr{E}_S$ holds. Assuming that $\boldsymbol{x} \in \mathscr{E}_S$, the vector $\boldsymbol{x}$ is expressed as a linear sum of the lead field vectors $\boldsymbol{l}_q$ ($q = 1, \ldots, Q$), such that

$$\boldsymbol{x} = \sum_{q=1}^{Q} \omega_q \boldsymbol{l}_q. \quad \text{(A.1)}$$

In order to show that $\boldsymbol{x} \in \mathrm{csp}\left(\left[\boldsymbol{y}_1^S, \ldots, \boldsymbol{y}_K^S\right]\right)$, we show that $\boldsymbol{x}$ in equation (A.1) is also expressed as a linear sum of $\boldsymbol{y}_j^S$:

$$\boldsymbol{x} = \sum_{j=1}^{K} \alpha_j \boldsymbol{y}_j^S. \quad \text{(A.2)}$$

Actually, substituting equation (8), into (A.2), and using equation (A.1), we get

$$\sum_{j=1}^{K} \alpha_j \sum_{q=1}^{Q} s_q^j \boldsymbol{l}_q = \sum_{q=1}^{Q} \left(\sum_{j=1}^{K} \alpha_j s_q^j \right) \boldsymbol{l}_q = \sum_{q=1}^{Q} \omega_q \boldsymbol{l}_q, \quad \text{(A.3)}$$

where $s_q(t_j)$ is denoted s_q^j. Comparing the coefficients of the vector $\boldsymbol{l}_q$ on the left and right sides gives following set of Q linear equations:

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\begin{aligned} \alpha_1 s_1^1 + \cdots + \alpha_K s_1^K &= \omega_1, \\ \alpha_1 s_2^1 + \cdots + \alpha_K s_2^K &= \omega_2, \\ \vdots \qquad\qquad & \quad \vdots \\ \alpha_1 s_Q^1 + \cdots + \alpha_K s_Q^K &= \omega_Q. \end{aligned}$$

Assuming that $K > Q$, for an arbitrary set of $\omega_1, \ldots, \omega_Q$, a set of $\alpha_1, \ldots, \alpha_K$ exists, and this fact leads to the conclusion that equation (A.2) holds for an arbitrary $\boldsymbol{x}(\boldsymbol{x} \in \mathscr{E}_S)$. Thus, the relationship $\mathrm{csp}\left(\left[\boldsymbol{y}_1^S, \ldots, \boldsymbol{y}_K^S\right]\right) \supset \mathscr{E}_S$ holds. Note that we assume that the source time courses $s_j^1, \ldots, s_j^K (j = 1, \ldots, Q)$ are linearly independent.

We next show that $\mathscr{E}_S \supset \mathrm{csp}\left(\left[\boldsymbol{y}_1^S, \ldots, \boldsymbol{y}_K^S\right]\right)$. To do so, we show that if $\boldsymbol{x} \in \mathrm{csp}\left(\left[\boldsymbol{y}_1^S, \ldots, \boldsymbol{y}_K^S\right]\right)$, $\boldsymbol{x} \in \mathscr{E}_S$ holds. If $\boldsymbol{x} \in \mathrm{csp}\left(\left[\boldsymbol{y}_1^S, \ldots, \boldsymbol{y}_K^S\right]\right)$, equation (A.2) holds. Substituting equation (8) into equation (A.2), we obtain

$$\boldsymbol{x} = \sum_{j=1}^{K} \alpha_j \boldsymbol{y}_j^S = \sum_{j=1}^{K} \alpha_j \sum_{q=1}^{Q} s_q^j \boldsymbol{l}_q = \sum_{q=1}^{Q} \left(\sum_{j=1}^{K} \alpha_j s_q^j \right) l_q, \quad \text{(A.4)}$$

which shows that the vector $\boldsymbol{x}$ is expressed as a linear sum of $\boldsymbol{l}_q (q = 1, \ldots, Q)$, namely, $\boldsymbol{x} \in \mathscr{E}_S$. This leads to the conclusion that $\mathscr{E}_S \supset \mathrm{csp}\left(\left[\boldsymbol{y}_1^S, \ldots, \boldsymbol{y}_K^S\right]\right)$. In summary, since both directions $\mathrm{csp}\left(\left[\boldsymbol{y}_1^S, \ldots, \boldsymbol{y}_K^S\right]\right) \supset \mathscr{E}_S$, and $\mathscr{E}_S \supset \mathrm{csp}\left(\left[\boldsymbol{y}_1^S, \ldots, \boldsymbol{y}_K^S\right]\right)$ hold, the relationship $\mathscr{E}_S = \mathrm{csp}\left(\boldsymbol{B}_S\right)$ holds.

A.2. Proof of equation (19): $\mathscr{K}_S = \mathrm{rsp}\left(\boldsymbol{B}_S\right)$

$\mathscr{K}_S = \mathrm{rsp}\left(\boldsymbol{B}_S\right)$ can be shown just as $\mathscr{E}_S = \mathrm{csp}\left(\boldsymbol{B}_S\right)$ was shown in appendix A.1. We first show that if $\boldsymbol{x} \in \mathscr{K}_S$, the relationship $\boldsymbol{x} \in \mathrm{rsp}\left(\left[\beta_1^S, \ldots, \beta_M^S\right]\right)$ holds. If $\boldsymbol{x} \in \mathscr{K}_S$, the relationship

$$\boldsymbol{x} = \sum_{q=1}^{Q} \omega_q \boldsymbol{s}_q. \quad \text{(A.5)}$$

holds. If this vector $\boldsymbol{x}$ sartisfies $\boldsymbol{x} \in \mathrm{rsp}\left(\left[\beta_1^S, \ldots, \beta_M^S\right]\right)$, $\boldsymbol{x}$ is expressed as

$$\boldsymbol{x} = \sum_{j=1}^{M} \alpha_j \beta_j^S. \quad \text{(A.6)}$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Using equation (13) and the arguments in the preceding section, under the assumption $M > Q$, for an arbitrary set of $\omega_1, \ldots, \omega_Q$, a set of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M$ exists. Therefore, equation (A.6) holds, and $\mathrm{rsp}\left(\left[\beta_1^S, \ldots, \beta_M^S\right]\right) \supset \mathscr{K}_S$ also holds.

We next show that if $\boldsymbol{x} \in \mathrm{rsp}\left(\left[\beta_1^S, \ldots, \beta_M^S\right]\right)$, then $\boldsymbol{x} \in \mathscr{K}_S$ holds. If $\boldsymbol{x} \in \mathrm{rsp}\left(\left[\beta_1^S, \ldots, \beta_M^S\right]\right)$, equation (A.6) holds. Substituting equation (13) into equation (A.6), we get

$$\boldsymbol{x} = \sum_{j=1}^{M} \alpha_j \beta_j^S = \sum_{j=1}^{M} \alpha_j \sum_{q=1}^{Q} l_q^j \boldsymbol{s}_q = \sum_{q=1}^{Q} \left(\sum_{j=1}^{M} \alpha_j l_q^j \right) \boldsymbol{s}_q . \quad \text{(A.7)}$$

The equation above shows that $\boldsymbol{x} \in \mathscr{K}_S$ holds. Threrefore, $\mathscr{K}_S \supset \mathrm{rsp}\left(\left[\beta_1^S, \ldots, \beta_M^S\right]\right)$ holds. In summary, since both directions $\mathrm{rsp}\left(\left[\beta_1^S, \ldots, \beta_M^S\right]\right) \subset \mathscr{K}_S$ and $\mathscr{K}_S \subset \mathrm{rsp}\left(\left[\beta_1^S, \ldots, \beta_M^S\right]\right)$ hold, the relationship $\mathscr{K}_S = \mathrm{rsp}\left(\boldsymbol{B}_S\right)$ holds.

A.3. Proof of $\mathscr{K}_I = \mathrm{rsp}\left(\boldsymbol{X}\boldsymbol{B}_I\right)$ where $\boldsymbol{X} = \check{\boldsymbol{P}}_S$ or $\boldsymbol{X} = \left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)$

Using the same arguments as in the preceding section, we can prove that $\mathscr{K}_I = \mathrm{rsp}\left(\boldsymbol{X}\boldsymbol{B}_I\right)$ holds where $\boldsymbol{X} = \check{\boldsymbol{P}}_S$ or $\boldsymbol{X} = \left(\boldsymbol{I} - \check{\boldsymbol{P}}_S\right)$ and $\check{\boldsymbol{P}}_S$ is the the pseudo signal subspace projector defined in equations (62). Using equation (24), multiplying $\boldsymbol{X}$ with the interference vector $\boldsymbol{y}_I(t)$ gives

$$\boldsymbol{X}\boldsymbol{y}_I(t) = \sum_{p=1}^{P} \sigma_p(t) \boldsymbol{X}\boldsymbol{\xi}_p = \sum_{p=1}^{P} \sigma_p(t) \check{\boldsymbol{\xi}}_p, \quad \text{(A.8)}$$

where we use the notation $\check{\boldsymbol{\xi}}_p = \boldsymbol{X}\boldsymbol{\xi}_p$. That is, the matrix $\boldsymbol{X}$ changes $\boldsymbol{\xi}_p$ to $\check{\boldsymbol{\xi}}_p$, but the multiplication by $\boldsymbol{X}$ never affects the time course $\sigma_p(t)$. Therefore, we get the following relationship:

$$\begin{aligned} \boldsymbol{X}\boldsymbol{B}_I &= \boldsymbol{X}\left[\sum_{p=1}^{P} \sigma_p(t_1)\boldsymbol{\xi}_p, \ldots, \sum_{p=1}^{P} \sigma_p(t_K)\boldsymbol{\xi}_p \right] \\ &= \left[\sum_{p=1}^{P} \sigma_p(t_1)\check{\boldsymbol{\xi}}_p, \ldots, \sum_{p=1}^{P} \sigma_p(t_K)\check{\boldsymbol{\xi}}_p \right] \\ &= \begin{bmatrix} \sum_{p=1}^{P} \left[\sigma_p(t_1), \ldots, \sigma_p(t_K)\right] \check{\boldsymbol{\xi}}_1^p \\ \vdots \\ \sum_{p=1}^{P} \left[\sigma_p(t_1), \ldots, \sigma_p(t_K)\right] \check{\boldsymbol{\xi}}_M^p \end{bmatrix} = \begin{bmatrix} \sum_{p=1}^{P} \check{\xi}_1^p \boldsymbol{\sigma}_p \\ \vdots \\ \sum_{p=1}^{P} \check{\xi}_M^p \boldsymbol{\sigma}_p \end{bmatrix}, \end{aligned} \quad \text{(A.9)}$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

where $\breve{\xi}_1^p, \ldots, \breve{\xi}_M^p$ are the elements of $\breve{\boldsymbol{\xi}}_p$: $\breve{\boldsymbol{\xi}}_p = \left[\breve{\xi}_1^p, \ldots, \breve{\xi}_M^p\right]^T$. Equation (A.9) shows that denoting the jth row of $\boldsymbol{XB}_I$ as $\breve{\beta}_j^I$ we get the relationship $\breve{\beta}_j^I \in \mathscr{K}_I$. Thus, using the arguments in appendix A.2, we can prove that the relationship $\mathscr{K}_I = \mathrm{rsp}\left(\boldsymbol{XB}_I\right)$ holds.

A.4. Asymptotic equivalence between singular vectors of B and B_s

Let us respectively denote sample covariance matrices of $\boldsymbol{B}$, $\boldsymbol{B}_S$, and $B_{\boldsymbol{\varepsilon}}$ by $\boldsymbol{R}$, $\boldsymbol{R}_S$, and $R_{\boldsymbol{\varepsilon}}$, which are obtained such that $\boldsymbol{R} = \frac{1}{K}\boldsymbol{B}\boldsymbol{B}^T$, and $\boldsymbol{R}_S = \frac{1}{K}\boldsymbol{B}_S\boldsymbol{B}_S^T$. Assuming that $\boldsymbol{B}_S$ and $\boldsymbol{B}_{\boldsymbol{\varepsilon}}$ are uncorrelated, from equation (6) we have

$$\boldsymbol{R} = \boldsymbol{R}_S + \boldsymbol{R}_\varepsilon . \quad \text{(A.10)}$$

Assuming the noise model in equation (3) and infinite number of data samples, the noise covariance matrix is given by

$$\boldsymbol{R}_\varepsilon = \varrho^2 \boldsymbol{I} . \quad \text{(A.11)}$$

Sample covariance matrices $\boldsymbol{R}_S$ and $\boldsymbol{R}$ are expressed using the eigen decomposition as:

$$\boldsymbol{R}_S = \sum_{j=1}^{Q} \left(\gamma_j^S\right)^2 \boldsymbol{u}_j^S \left(\boldsymbol{u}_j^S\right)^T, \quad \text{(A.12)}$$

$$\boldsymbol{R} = \sum_{j=1}^{M} \gamma_j^2 \boldsymbol{u}_j \boldsymbol{u}_j^T . \quad \text{(A.13)}$$

Substituting equations (A.12), (A.13) and (A.11) into equation (A.10), we can obtain

$$\sum_{j=1}^{M} \gamma_j^2 \boldsymbol{u}_j \boldsymbol{u}_j^T = \sum_{j=1}^{Q} \left(\gamma_j^S\right)^2 \boldsymbol{u}_j^S \left(\boldsymbol{u}_j^S\right)^T + \varrho^2 \sum_{j=1}^{M} \boldsymbol{u}_j^S \left(\boldsymbol{u}_j^S\right)^T . \quad \text{(A.14)}$$

Therefore, we have the following asymptotic relationships:

$$\gamma_j^2 = \left(\gamma_j^S\right)^2 + \varrho^2 \quad \text{for} \quad j = 1, \ldots, Q, \quad \text{(A.15)}$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\gamma_j^2 = \varrho^2 \quad \text{for} \quad j = Q+1, \ldots, M, \quad \text{(A.16)}$$

$$\boldsymbol{u}_j = \boldsymbol{u}_j^S \quad \text{for} \quad j = 1, \ldots, M. \quad \text{(A.17)}$$

Equation (A.17) indicates the asymptotic equivalence between the singular vectors of $\boldsymbol{B}$ and $\boldsymbol{B}_S$.

A.5. Row space of sum of matrices

Defining two subspaces as $\mathscr{X}$ and $\mathscr{Y}$, the sum of $\mathscr{X}$ and $\mathscr{Y}$ is defined such that [21],

$$\mathscr{X} + \mathscr{Y} = \{x + y | x \in \mathscr{X} \text{ and } y \in \mathscr{Y}\}. \quad \text{(A.18)}$$

That is, the sum of $\mathscr{X}$ and $\mathscr{Y}$ is defined as a set of all possible sums between $\boldsymbol{x}(\boldsymbol{x} \in \mathscr{X})$ and $\boldsymbol{y}(\boldsymbol{y} \in \mathscr{Y})$. Then, we show that, for two arbitrary matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, the relationship

$$\text{rsp}(\boldsymbol{X} + \boldsymbol{Y}) \subset \text{rsp}(\boldsymbol{X}) + \text{rsp}(\boldsymbol{Y}). \quad \text{(A.19)}$$

holds. To show this, let us denote row vectors of $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z} = \boldsymbol{X} + \boldsymbol{Y}$ by $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$, respectively. By definition, for $\boldsymbol{z}$ ($\boldsymbol{z} \in \text{rsp}(\boldsymbol{Z})$), the relationship,

$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$$

holds where $\boldsymbol{x} \in \text{rsp}(\boldsymbol{X})$ and $\boldsymbol{y} \in \text{rsp}(\boldsymbol{Y})$. Thus, $z \in \text{rsp}(\boldsymbol{X}) + \text{rsp}(\boldsymbol{Y})$ holds, and we have shown $\text{rsp}(\boldsymbol{X} + \boldsymbol{Y}) \subset \text{rsp}(\boldsymbol{X}) + \text{rsp}(\boldsymbol{Y})$.

A.6. Basis vectors that span sum of subspaces

Next, we derive the basis vectors of $\mathscr{X} + \mathscr{Y}$. Sets of the basis vectors of $\mathscr{X}$ and $\mathscr{Y}$ are respectively denoted by $\mathscr{S}_X = \left\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\mu\right\}$ and $\mathscr{S}_Y = \left\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_\nu\right\}$. Let us assume that $\boldsymbol{x} \in \mathscr{X}$, $\boldsymbol{y} \in \mathscr{Y}$, and $z \in \mathscr{X} + \mathscr{Y}$. We then have

$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y} = \sum_{j=1}^{\mu} c_j \boldsymbol{x}_j + \sum_{j=1}^{\nu} d_j \boldsymbol{y}_j.$$

Therefore, the relationship

$$\boldsymbol{z} \in \text{span}\left(\left[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\mu, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_\nu\right]\right) \quad \text{(A.20)}$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

holds. Defining a set of the basis vecors of $\mathcal{X}+\mathcal{Y}$ as $\mathcal{S}_{X+Y}$, equation (A.20) indicates that $\mathcal{S}_{X+Y}$ is equal to

$$\begin{aligned}\mathcal{S}_{X+Y} &= \left\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_\mu,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_\nu\right\}\\ &= \left\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_\mu\right\}\cup\left\{\boldsymbol{y}_1,\ldots,\boldsymbol{y}_\nu\right\} = \mathcal{S}_X\cup\mathcal{S}_Y\,.\end{aligned} \tag{A.21}$$

The equation above shows that the basis vectors of the sum of subspaces are obtained as the union of the basis vector sets that span each of the subspaces.

A.7. Proof: If $\mathcal{X}\cap\mathcal{X}'$ and $\mathcal{Y}\subset\mathcal{Y}'$, then $(\mathcal{X}\cap\mathcal{Y})\subset(\mathcal{X}'\cap\mathcal{Y}')$ holds

If $\boldsymbol{x}\in\mathcal{X}\cap\mathcal{Y}$, the relationships, $\boldsymbol{x}\in\mathcal{X}$ and $\boldsymbol{x}\in\mathcal{Y}$, hold. Since $\mathcal{X}\subset\mathcal{X}'$ and $\mathcal{Y}\subset\mathcal{Y}'$, we have $\boldsymbol{x}\in\mathcal{X}'$ and $\boldsymbol{x}\in\mathcal{Y}'$. Therefore, $\boldsymbol{x}\in\mathcal{X}'\cap\mathcal{Y}'$, and we have shown $(\mathcal{X}\cap\mathcal{Y})\subset(\mathcal{X}'\cap\mathcal{Y}')$.

A.8. Proof of $span(\mathcal{S}_X\cap\mathcal{S}_Y)=span(\mathcal{S}_X)\cap span(\mathcal{S}_Y)$

Two subspaces are denoted by $\mathcal{X}$ and $\mathcal{Y}$, and sets of their basis vectors by $\mathcal{S}_X=\left\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_\mu\right\}$ and $\mathcal{S}_Y=\left\{\boldsymbol{y}_1,\ldots,\boldsymbol{y}_\nu\right\}$. Let us also define a set of basis vectors $S_X\cap\mathcal{S}_Y=\left\{z_1,\ldots,z_r\right\}$ where the dimension of the intersection is r. If $\boldsymbol{x}\in\ \mathrm{span}(\mathcal{S}_X\cap\mathcal{S}_Y)$, we can write $\boldsymbol{x}$ as a linear sum: $\boldsymbol{x}=\sum_{j=1}^{r}w_jz_j$. The fact that $z_j\in\mathcal{S}_X$ leads to $\boldsymbol{x}\in\mathcal{X}=\ \mathrm{span}(\mathcal{S}_X)$. Also, the fact that $z_j\in\mathcal{S}_Y$ leads to $\boldsymbol{x}\in\mathcal{Y}=\ \mathrm{span}(\mathcal{S}_Y)$. Therefore, $\boldsymbol{x}\in\mathcal{X}\cap\mathcal{Y}=\ \mathrm{span}(\mathcal{S}_X)\cap\ \mathrm{span}(\mathcal{S}_Y)$ holds, and we can show $\mathrm{span}(\mathcal{X}\cap\mathcal{Y})\subset\ \mathrm{span}(\mathcal{X})\cap\ \mathrm{span}(\mathcal{X})$.

The other direction of the proof is as follows. If $\boldsymbol{x}\in\ \mathrm{span}(\mathcal{S}_X)\cap\ \mathrm{span}(\mathcal{S}_Y)$, we have $\boldsymbol{x}\in\ \mathrm{span}(\mathcal{S}_X)$ and $\boldsymbol{x}\in\ \mathrm{span}(\mathcal{S}_Y)$. Thus, $\boldsymbol{x}$ is expressed as a linear sum of basis vectors that belong to both $\mathcal{S}_X$ and $\mathcal{S}_Y$, namely, a linear sum of $\left\{z_1,\ldots,z_r\right\}$. Therefore, $\boldsymbol{x}\in\ \mathrm{span}(\mathcal{S}_X\cap\mathcal{S}_Y)$, and this indicates $\mathrm{span}(\mathcal{S}_X)\cap\ \mathrm{span}(\mathcal{S}_Y)\subset\ \mathrm{span}(\mathcal{S}_X\cap\mathcal{S}_Y)$. Since we have shown the both directions, we have shown $\mathrm{span}(\mathcal{S}_X\cap\mathcal{S}_Y)=\ \mathrm{span}(\mathcal{S}_X)\cap\ \mathrm{span}(\mathcal{S}_Y)$.

A.9. Derivation of basis vectors that span intersection of two row spaces

Let us assume that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are low-rank data matrices. We define the basis vectors of rsp($\boldsymbol{X}$) as $\mathcal{S}_X=\left\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_\mu\right\}$ where μ is the dimension of rsp($\boldsymbol{X}$), and the basis vectors of rsp($\boldsymbol{Y}$) as $\mathcal{S}_Y=\left\{\boldsymbol{y}_1,\ldots,\boldsymbol{y}_\nu\right\}$ where ν is the dimension of rsp($\boldsymbol{Y}$). The procedure used to find a set of basis vectors of rsp($\boldsymbol{X}$) ∩ rsp($\boldsymbol{Y}$) is described below. The procedure is according to [22].

An orthonormal set of basis vectors of the intersection is obtained as a set of the principal vectors whose principal angles are equal to zero. To find those principal vectors, we define matrices whose columns consist of the basis vectors such that

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\boldsymbol{U} = \left[\boldsymbol{x}_1^T, \ldots, \boldsymbol{x}_\mu^T\right], \quad \text{(A.22)}$$

$$\boldsymbol{V} = \left[\boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_\nu^T\right]. \quad \text{(A.23)}$$

The results of singular-value decomposition of a matrix $\boldsymbol{U}^T\boldsymbol{V}$ are expressed as

$$\boldsymbol{U}^T\boldsymbol{V} = \boldsymbol{Q}\begin{bmatrix} \cos(\theta_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \cos(\theta_\nu) \end{bmatrix}\boldsymbol{T}^T, \quad \text{(A.24)}$$

where $\boldsymbol{Q}$ and $\boldsymbol{T}$ are matrices whose columns consist of singular vectors, and we assume that $\mu > \nu$. In equation (A.24), singular values of the matrix $\boldsymbol{U}^T\boldsymbol{V}$ are equal to the cosines of the principal angles between the two subspaces $\mathrm{csp}\left(\left[\boldsymbol{x}_1^T, \ldots, \boldsymbol{x}_\mu^T\right]\right)(= \mathrm{rsp}(\boldsymbol{X}))$ and $\mathrm{csp}\left(\left[\boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_\nu^T\right]\right)(= \mathrm{rsp}(\boldsymbol{Y}))$. The intersection has the property that the principal angles are equal to zero. Thus, by observing the relation

$$\begin{aligned} \cos(\theta_1) &= \cos(\theta_2) = \cdots = \cos(\theta_r) \\ &\approx 1 > \cos(\theta_{r+1}) \geqslant \cdots \geqslant \cos(\theta_\nu), \end{aligned}$$

the dimension of $\mathrm{csp}(\boldsymbol{U}) \cap \mathrm{csp}(\boldsymbol{V})$, (namely, the dimension of $\mathrm{rsp}(\boldsymbol{X}) \cap \mathrm{rsp}(\boldsymbol{Y})$) is determined to be r. The principal vectors are then obtained either as the first r columns of the matrix $\boldsymbol{UQ}$ or the first r columns of the matrix $\boldsymbol{VT}$. Defining the first r columns of $\boldsymbol{UQ}$ as $\psi_1^T, \ldots, \psi_r^T$, the vectors $\psi_1, \ldots, \psi_r$ form an orthonormal basis set for the intersection $\mathrm{rsp}(\boldsymbol{X}) \cap \mathrm{rsp}(\boldsymbol{Y})$.

Appendix B.: Signal space separation (SSS) method

B.1. Derivation of SSS basis vectors

This section presents an overview of the signal space separation (SSS) method [12–14] with an emphasis on the derivation of SSS signal and interference extractors. One of basic assumptions of the SSS method is that the sensors are installed in a source-free region, which is referred to as the sensor region in this paper. Then, the magnetic field at $\boldsymbol{r}$, $\boldsymbol{B}(\boldsymbol{r})$ is expressed using the spherical polar coordinate $\boldsymbol{r} = (r, \theta, \phi)$ as

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\boldsymbol{B}(\boldsymbol{r}) = -\mu_0 \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \alpha_{\ell,m} \frac{\boldsymbol{\nu}_{\ell,m}(\theta,\phi)}{r^{\ell+2}} - \mu_0 \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \beta_{\ell,m} r^{\ell-1} \omega_{\ell,m}(\theta,\phi), \quad \text{(B.1)}$$

where $\mu 0$ indicates the magnetic permeability of free space. In equation (B.1), $\boldsymbol{v}_{\ell,\mathrm{m}}(\theta,\phi)$ and $\omega_{\ell,m}(\theta,\phi)$ are the modified vector spherical harmonics [12, 23].

In the right-hand-side of equation (B.1), the first term represents the magnetic field generated from sources located closer to the origin than the sensors. The region closer to the origin than the sensors is referred to as the internal region. The second term represents the magnetic field from sources located farther from the origin than the sensors. The region farther from the origin than the sensors is referred to as the external region. If an appropriate choice of the origin can result in the signal sources of interest to be located within the internal region and all interference sources to be located within the external region, equation (B.1) provides a natural separation between the signal and interference. This is the key idea of the SSS method.

Let us derive the SSS basis vectors. To do so, we denote the output of the jth sensor by y_j, its location by $\boldsymbol{r}_j$ and its orientation by ζ_j. Then, we have

$$y_j = \boldsymbol{B}(\boldsymbol{r}_j) \cdot \boldsymbol{\zeta}_j = y_{\mathrm{int}}^j + y_{\mathrm{ext}}^j, \quad \text{(B.2)}$$

where the notation · represents the inner product between two vectors, and y_{int}^j and y^j_{ext} represent the two components of the jth sensor outputs generated from the internal and external regions, respectively. These components are expressed as

$$y_{\mathrm{int}}^j = -\sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \alpha_{\ell,m} \frac{\left[\nu_{\ell,m}(\theta_j,\phi_j) \cdot \boldsymbol{\zeta}_j\right]}{r_j^{\ell+2}}, \quad \text{(B.3)}$$

$$y_{\mathrm{ext}}^j = -\sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \beta_{\ell,m} r_j^{\ell-1} \left[\boldsymbol{\omega}_{\ell,m}(\theta_j,\phi_j) \cdot \boldsymbol{\zeta}_j\right], \quad \text{(B.4)}$$

where we set $\mu_0 = 1$ for simplicity. Let us define the internal and external components of the data vector $\boldsymbol{y}$ as $\boldsymbol{y}_{\mathrm{int}} = \left[y_{\mathrm{int}}^1, \ldots, y_{\mathrm{int}}^M\right]^T$ and $\boldsymbol{y}_{\mathrm{ext}} = \left[y_{\mathrm{ext}}^1, \ldots, y_{\mathrm{ext}}^M\right]^T$ which are expressed such that

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\boldsymbol{y}_{\text{int}} = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \alpha_{\ell,m} \boldsymbol{c}_{\ell,m} \quad \text{(B.5)}$$

$$\boldsymbol{y}_{\text{ext}} = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \beta_{\ell,m} \boldsymbol{d}_{\ell,m}, \quad \text{(B.6)}$$

where column vectors $\boldsymbol{c}_{\ell\text{m}}$ and $\boldsymbol{d}_{\ell\text{m}}$ are given by

$$\boldsymbol{c}_{\ell,m} = \begin{bmatrix} \frac{1}{r_1^{\ell+2}} \left[\boldsymbol{\nu}_{\ell,m}(\theta_1, \phi_1) \cdot \boldsymbol{\zeta}_1 \right] \\ \vdots \\ \frac{1}{r_M^{\ell+2}} \left[\boldsymbol{\nu}_{\ell,m}(\theta_M, \phi_M) \cdot \boldsymbol{\zeta}_M \right] \end{bmatrix} \quad \text{and}$$

$$d_{\ell,m} = \begin{bmatrix} r_1^{\ell-1} \left[\boldsymbol{\omega}_{\ell,m}(\theta_1, \phi_1) \cdot \boldsymbol{\zeta}_1 \right] \\ \vdots \\ r_M^{\ell-1} \left[\boldsymbol{\omega}_{\ell,m}(\theta_M, \phi_M) \cdot \boldsymbol{\zeta}_M \right] \end{bmatrix}. \quad \text{(B.7)}$$

Truncating the summation with respect to the index ℓ to L_C for $\boldsymbol{y}_{\text{int}}$ and L_D for $\boldsymbol{y}_{\text{ext}}$, we finally obtain

$$\boldsymbol{y} = \boldsymbol{y}_{\text{int}} + \boldsymbol{y}_{\text{ext}} = \sum_{\ell=1}^{L_C} \sum_{m=-\ell}^{\ell} \alpha_{\ell,m} \boldsymbol{c}_{\ell,m} + \sum_{\ell=1}^{L_D} \sum_{m=-\ell}^{\ell} \beta_{\ell,m} \boldsymbol{d}_{\ell,m}. \quad \text{(B.8)}$$

Thus, defining

$$\boldsymbol{C} = \left[\boldsymbol{c}_{1,-1}, \boldsymbol{c}_{1,0}, \boldsymbol{c}_{1,1}, \ldots, \boldsymbol{c}_{L_C, L_C} \right], \quad \text{(B.9)}$$

$$\boldsymbol{D} = \left[\boldsymbol{d}_{1,-1}, \boldsymbol{d}_{1,0}, \boldsymbol{d}_{1,1}, \ldots, \boldsymbol{d}_{L_D, L_D} \right], \quad \text{(B.10)}$$

$$\boldsymbol{\alpha} = \left[\alpha_{1,-1}, \alpha_{1,0}, \alpha_{1,1}, \ldots, \alpha_{L_C, L_C} \right]^T, \quad \text{(B.11)}$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\phi = \left[\beta_{1,-1}, \beta_{1,0}, \beta_{1,1}, \ldots, \beta_{L_D, L_D}\right]^T, \quad \text{(B.12)}$$

we obtain

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{\alpha} + \boldsymbol{D}\phi = [\boldsymbol{C}, \boldsymbol{D}]\begin{bmatrix}\boldsymbol{\alpha}\\ \phi\end{bmatrix} = \boldsymbol{S}\boldsymbol{x} \quad \text{(B.13)}$$

where $\boldsymbol{S} = [\boldsymbol{C},\boldsymbol{D}]$ and $\boldsymbol{x} = [\boldsymbol{\alpha}^T, \boldsymbol{\phi}^T]^T$. Here, $\boldsymbol{C}$ is an $M \times N_C$ matrix, and $\boldsymbol{D}$ is an $M \times N_D$ matrix where

$$N_C = L_C^2 + 2L_C, \quad \text{and} \quad N_D = L_D^2 + 2L_D. \quad \text{(B.14)}$$

B.2. SSS signal and interference extractors

When the key assumption that sources of interest are located in the internal region and all interference sources are located in the external region holds, $\boldsymbol{y}_{\text{int}}$ represents the signal of interest and $\boldsymbol{y}_{\text{ext}}$ represents the interference. Assuming that $M > N_C + N_d$, and using equation (B.13), the least squares estimate of $\boldsymbol{x}$ is obtained as:

$$\hat{\boldsymbol{x}} = \left(\boldsymbol{S}^T\boldsymbol{S}\right)^{-1}\boldsymbol{S}^T\boldsymbol{y}. \quad \text{(B.15)}$$

Considering that $\hat{\boldsymbol{x}} = \left[\hat{\boldsymbol{\alpha}}^T, \hat{\phi}^T\right]^T$ and using $\hat{\boldsymbol{\alpha}}$, the signal component $\boldsymbol{y}_{\text{int}}$ is estimated as

$$\hat{\boldsymbol{y}}_{\text{int}} = \boldsymbol{C}\hat{\boldsymbol{\alpha}}. \quad \text{(B.16)}$$

We now derive the SSS signal and interference extractors, and rewrite equation (B.16) using the signal extractor. To do so, let us define an operation to make a new column vector $[a_i, \ldots, a_j]^T$ by using the ith to jth components of $\boldsymbol{a} = [a_1, \ldots, a_M]^T$ as $[\boldsymbol{a}]_{[i:j]}$ (namely, $[\boldsymbol{a}]_{[i:j]} = [a_i, \ldots, a_j]^T$). From equation (B.15), we have

$$\hat{\boldsymbol{\alpha}} = \left[\left(\boldsymbol{S}^T\boldsymbol{S}\right)^{-1}\boldsymbol{S}^T\boldsymbol{y}\right]_{[1:N_C]}. \quad \text{(B.17)}$$

When the condition number of a matrix $\boldsymbol{S}^T\boldsymbol{S}$ is small, i.e. the column spaces of $\boldsymbol{C}$ and $\boldsymbol{D}$ are well separated, the relationship

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$\hat{\boldsymbol{\alpha}} = \left[\left(\boldsymbol{S}^T\boldsymbol{S}\right)^{-1}\boldsymbol{S}^T\boldsymbol{y}\right]_{\left[1:N_C\right]} \approx \left[\left(\boldsymbol{S}^T\boldsymbol{S} + \kappa\boldsymbol{I}\right)^{-1}\boldsymbol{S}^T\boldsymbol{y}\right]_{\left[1:N_C\right]} \quad \text{(B.18)}$$

holds, where $\boldsymbol{\kappa}$ is a small positive constant[10]. Using the matrix inversion formula

$$\left(\boldsymbol{S}^T\boldsymbol{S} + \kappa\boldsymbol{I}\right)^{-1}\boldsymbol{S}^T\boldsymbol{y} = \boldsymbol{S}^T\left(\boldsymbol{S}\boldsymbol{S}^T + \kappa\boldsymbol{I}\right)^{-1}\boldsymbol{y}, \quad \text{(B.19)}$$

we get

$$\begin{aligned}\hat{\boldsymbol{\alpha}} &= \left[\left(\boldsymbol{S}^T\boldsymbol{S} + \kappa\boldsymbol{I}\right)^{-1}\boldsymbol{S}^T\boldsymbol{y}\right]_{\left[1:N_C\right]} = \left[\boldsymbol{S}^T\left(\boldsymbol{S}\boldsymbol{S}^T + \kappa\boldsymbol{I}\right)^{-1}\boldsymbol{y}\right]_{\left[1:N_C\right]} \\ &= \boldsymbol{C}^T\left(\boldsymbol{S}\boldsymbol{S}^T + \kappa\boldsymbol{I}\right)^{-1}\boldsymbol{y} \approx \boldsymbol{C}^T\left(\boldsymbol{S}\boldsymbol{S}^T\right)^{-1}\boldsymbol{y} = \boldsymbol{C}^T\left(\boldsymbol{C}\boldsymbol{C}^T + \boldsymbol{D}\boldsymbol{D}^T\right)^{-1}\boldsymbol{y}\,.\end{aligned} \quad \text{(B.20)}$$

Using equation (B.16), we finally obtain

$$\hat{\boldsymbol{y}}_{\text{int}} \approx \boldsymbol{C}\boldsymbol{C}^T\left(\boldsymbol{C}\boldsymbol{C}^T + \boldsymbol{D}\boldsymbol{D}^T\right)^{-1}\boldsymbol{y}\,. \quad \text{(B.21)}$$

Accordingly, we can conclude that the interference removal by SSS is achieved by multiplying

$$\boldsymbol{\Gamma}_S = \boldsymbol{C}\boldsymbol{C}^T\left(\boldsymbol{C}\boldsymbol{C}^T + \boldsymbol{D}\boldsymbol{D}^T\right)^{-1} \quad \text{(B.22)}$$

with the data vector $\boldsymbol{y}$. That is, the matrix $\boldsymbol{\Gamma}_S$ plays a role as a filter that passes the signal of interest and blocks the interference. In this paper, $\boldsymbol{\Gamma}_S$ is called the SSS signal extractor. Note that, since relationships $(\boldsymbol{\Gamma}_S)^2 = \boldsymbol{\Gamma}_S$ and $(\boldsymbol{\Gamma}_S)^T = \boldsymbol{\Gamma}_S$ do not hold, $\boldsymbol{\Gamma}_S$ is not a projector. In exactly the same manner, we can derive

$$\hat{\boldsymbol{y}}_{\text{ext}} \approx \boldsymbol{D}\hat{\boldsymbol{\phi}} \approx \boldsymbol{D}\boldsymbol{D}^T\left(\boldsymbol{C}\boldsymbol{C}^T + \boldsymbol{D}\boldsymbol{D}^T\right)^{-1}\boldsymbol{y}\,. \quad \text{(B.23)}$$

Therefore, defining $\boldsymbol{\Gamma}_I$ as

$$\boldsymbol{\Gamma}_I = \boldsymbol{D}\boldsymbol{D}^T\left(\boldsymbol{C}\boldsymbol{C}^T + \boldsymbol{D}\boldsymbol{D}^T\right)^{-1}, \quad \text{(B.24)}$$

[10]This equation simply claims that when the condition number of $\boldsymbol{S}^T\boldsymbol{S}$ is small, nearly identical inverse matrices can be obtained either with or without the regularization term $\boldsymbol{\kappa}\boldsymbol{I}$.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$\mathbf{\Gamma}_I$ plays a role as a filter that passes the interference but blocks the signal of interest. This $\mathbf{\Gamma}_I$ is called the SSS interference extractor in this paper.

References

[1]. Paulraj A, Ottersten B, Roy R, Swindlehurst A, Xu G and Kailath T 1993 Subspace methods for directions-of-arrival estimation Handbook of Statistics ed Bose NK and Rao CR (Amsterdam: Elsevier) pp 693–739

[2]. Uusitalo M and Ilmoniemi R 1997 Signal-space projection method for separating MEG or EEG into components Med. Biol. Eng. Comput 35 135–40 [PubMed: 9136207]

[3]. Tesche C, Uusitalo M, Ilmoniemi R, Huotilainen M, Kajola M and Salonen O 1995 Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources Electroencephalogr. Clin. Neurophysiol 95 189–200 [PubMed: 7555909]

[4]. Nolte G and Curio G 1999 The effect of artifact rejection by signal-space projection on source localization accuracy in meg measurements IEEE Trans. Biomed. Eng 46 400–8 [PubMed: 10217878]

[5]. Sekihara K, Poeppel D, Marantz A and Miyashita Y 2000 Neuromagnetic inverse modeling: application of eigenstructure-based approaches to extracting cortical activities from meg data Image, Language, Brain: Papers from the First Mind Articulation Project Symp. (Cambridge, MA: MIT Press) p 197

[6]. Widrow B, Glover JR, McCool JM, Kaunitz J, Williams CS, Hearn RH, Zeidler JR, Dong JE and Goodlin RC 1975 Adaptive noise cancelling: principles and applications Proc. IEEE 63 1692–716

[7]. Adachi Y, Shimogawara M, Higuchi M, Haruta Y and Ochiai M 2001 Reduction of non-periodic environmental magnetic noise in meg measurement by continuously adjusted least squares method IEEE Trans. Appl. Supercond 11 669–72

[8]. De Cheveign A and Simon JZ 2008 Sensor noise suppression J. Neurosci. Methods 168 195–202 [PubMed: 17963844]

[9]. Watanabe T, Kawabata Y, Ukegawa D, Kawabata S, Adachi Y and Sekihara K 2013 Removal of stimulus-induced artifacts in functional spinal cord imaging 35th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (IEEE) pp 3391–4

[10]. Taulu S and Simola J 2006 Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements Phys. Med. Biol 51 1759–68 [PubMed: 16552102]

[11]. Sekihara K, Kawabata Y, Ushio S, Sumiya S, Kawabata S, Adachi Y and Nagarajan SS 2016 Dual signal subspace projection (DSSP): a novel algorithm for removing large interference in biomagnetic measurements J. Neural Eng 13 036007 [PubMed: 27064933]

[12]. Taulu S and Kajola M 2005 Presentation of electromagnetic multichannel data: the signal space separation method J. Appl. Phys 97 124905

[13]. Taulu S, Simola J and Kajola M 2005 Applications of the signal space separation method IEEE Trans. Signal Process. 53 3359–72

[14]. Taulu S, Kajola M and Simola J 2004 Suppression of interference and artifacts by the signal space separation method Brain Topogr. 16 269–75 [PubMed: 15379226]

[15]. Sekihara K and Nagarajan SS 2008 Adaptive Spatial Filters for Electromagnetic Brain Imaging (Berlin: Springer)

[16]. Scharf LL 1991 Statistical Signal Processing: Detection, Estimation and Time Series Analysis (New York: Addison-Wesley)

[17]. Nurminen J, Taulu S and Okada Y 2008 Effects of sensor calibration, balancing and parametrization on the signal space separation method Phys. Med. Biol 53 1975 [PubMed: 18354243]

[18]. Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J and Lounasmaa OV 1993 Magnetoencephalography-theory instrumentation and applications to noninvasive studies of the working human brain Rev. Mod. Phys 65 413–97

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[19]. Mosher JC, Hamalainen MS, Pantazis D, Hui HB, Burgess RC and Leahy RM 2009 Generalized sidelobe canceller for magnetoencephalography arrays IEEE Int. Symp. on Biomedical Imaging: from Nano to Macro (IEEE) pp 149–52

[20]. Sarvas J 1987 Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem Phys. Med. Biol 32 11–22 [PubMed: 3823129]

[21]. Goldberg JL 1991 Matrix Theory with Applications (New York: McGraw-Hill)

[22]. Golub GH and Van Loan CF 2012 Matrix Computations vol 3 (Baltimore, MD: Johns Hopkins University Press)

[23]. Hill E 1954 The theory of vector spherical harmonics Am. J. Phys 22 211–4

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

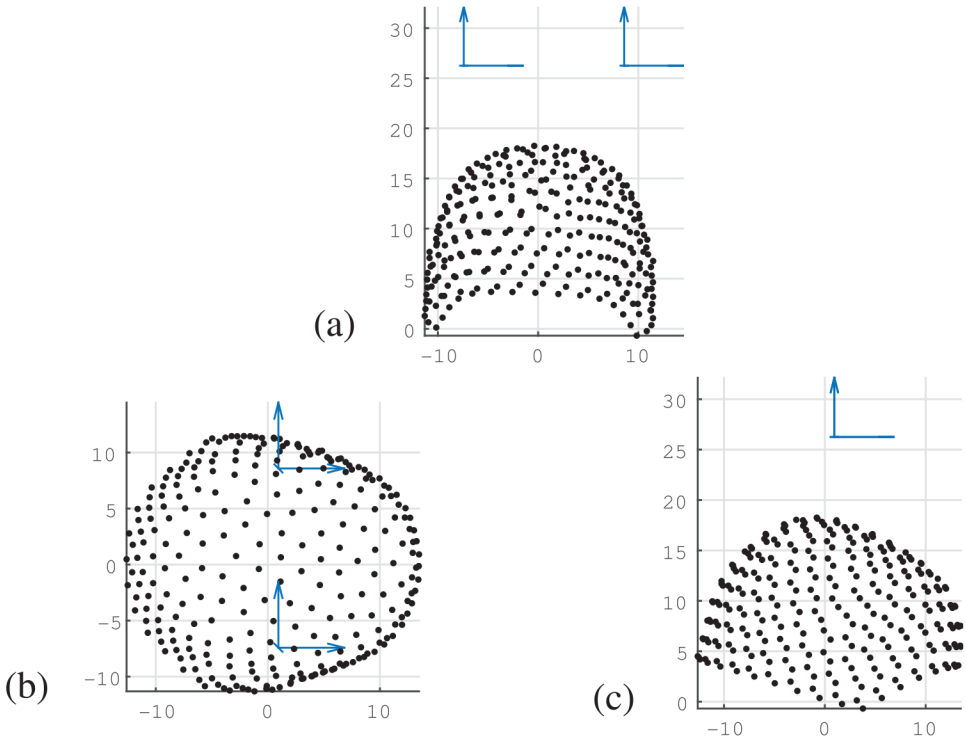

Figure 1.

An arrangement of the 275-channel whole-head sensor array of the Omega™ (VMS Medtech, Coquitlam, Canada) neuromagnetometer used in our numerical experiments. Locations of sensors are indicated by filled circles. Locations and orientations of six reference sensors consisting of two sets of vector magnetometers are shown. Here, (a), (b) and (c), respectively, show coronal, axial, and sagittal views of the sensor arrangement. Note that the arrangement of the reference sensors was assumed solely for our numerical experiments, and differs from the true arrangement in the Omega neuromagnetometer system.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

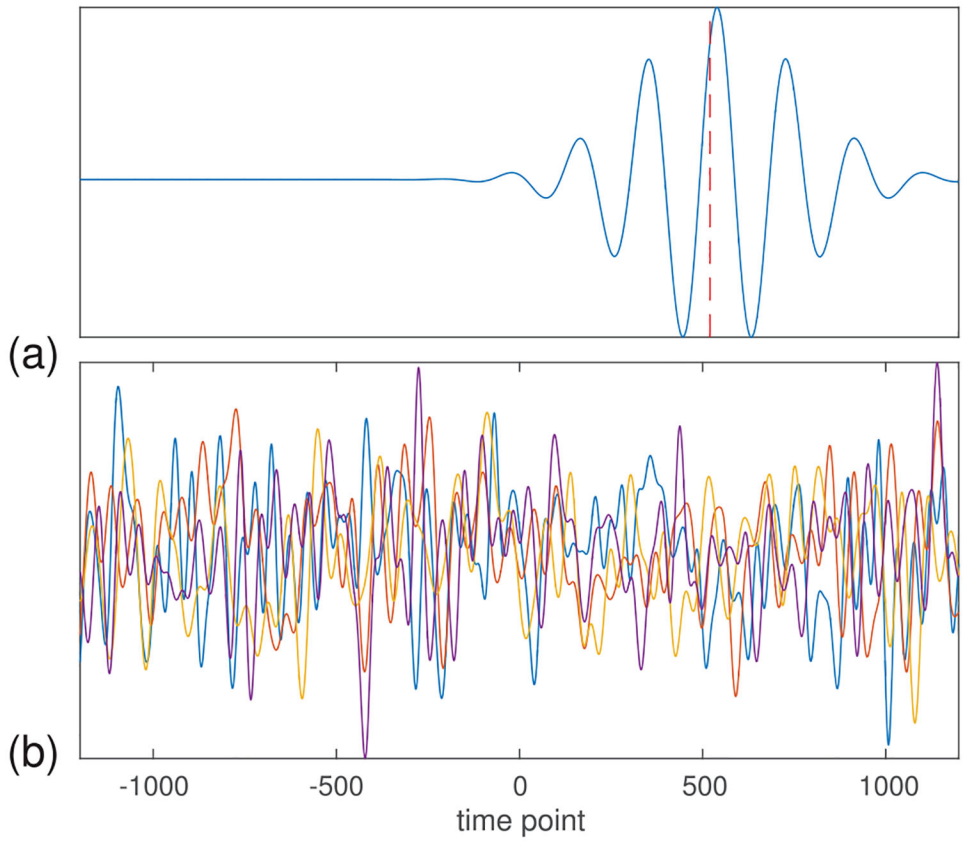

Figure 2.

(a) Time course of the signal source assumed in our numerical experiments. The time t is expressed with the unit of time point ranging from $t = -1200$ to 1200. The broken vertical line indicates the time point at $t = 520$. The field maps at this time point are shown in the following figures. (b) Time courses of the four interference sources assumed in our numerical experiments.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

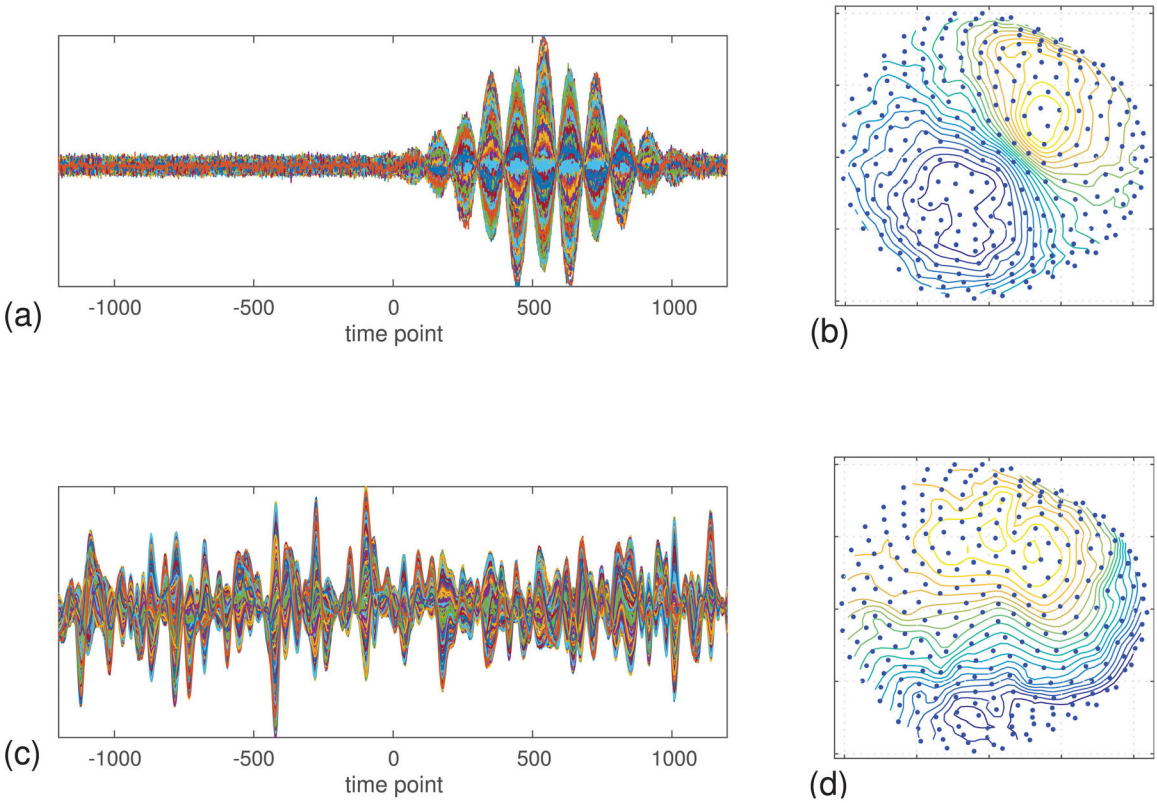

Figure 3.

(a) Time courses of the signal magnetic field with Gaussian noise with the signal-to-noise ratio (SNR) equal to 32. (b) The magnetic field map of the signal magnetic field at $t = 520$. (c) Time courses of interference-overlapped sensor data with the interference-to-signal ratio (ISR) equal to 6. (d) The magnetic field map of the interference overlapped sensor data at $t = 520$.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

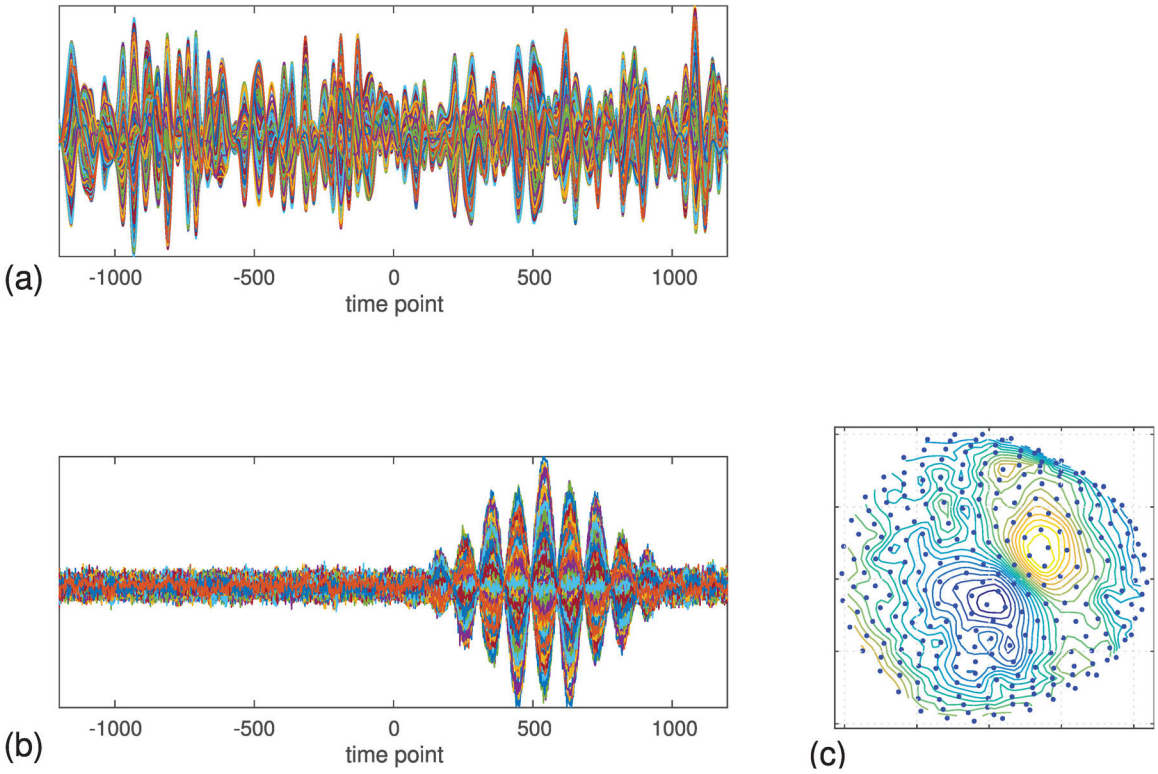

Figure 4.

(a) Simulated empty-room noise data used in the SSP interference removal experiments. (b) Sensor time courses of the SSP interference-removal results. (c) Magnetic field map of the SSP interference removal results. The map at $t = 520$ is shown.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

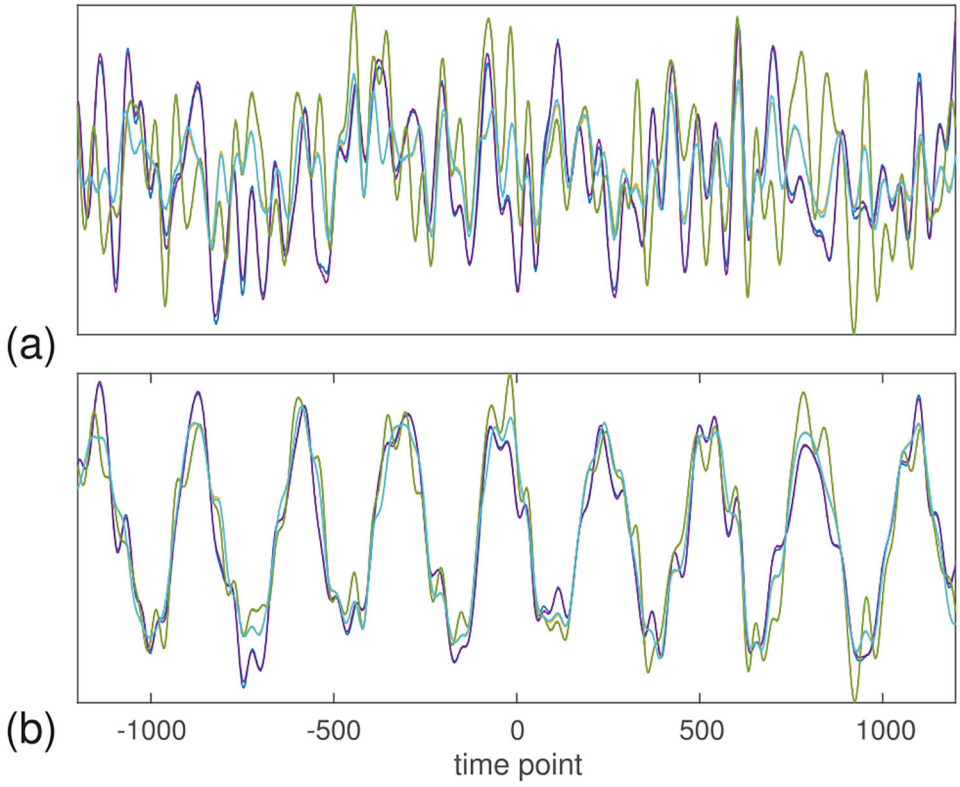

Figure 5.

(a) Time courses of reference sensor data used in ANC and CTSP interference-removal experiments of which results are shown in figure 6. (b) Time courses of reference sensor data containing a large additive fluctuation. The data were used in ANC and CTSP interference removal experiments of which results are shown in figure 7.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

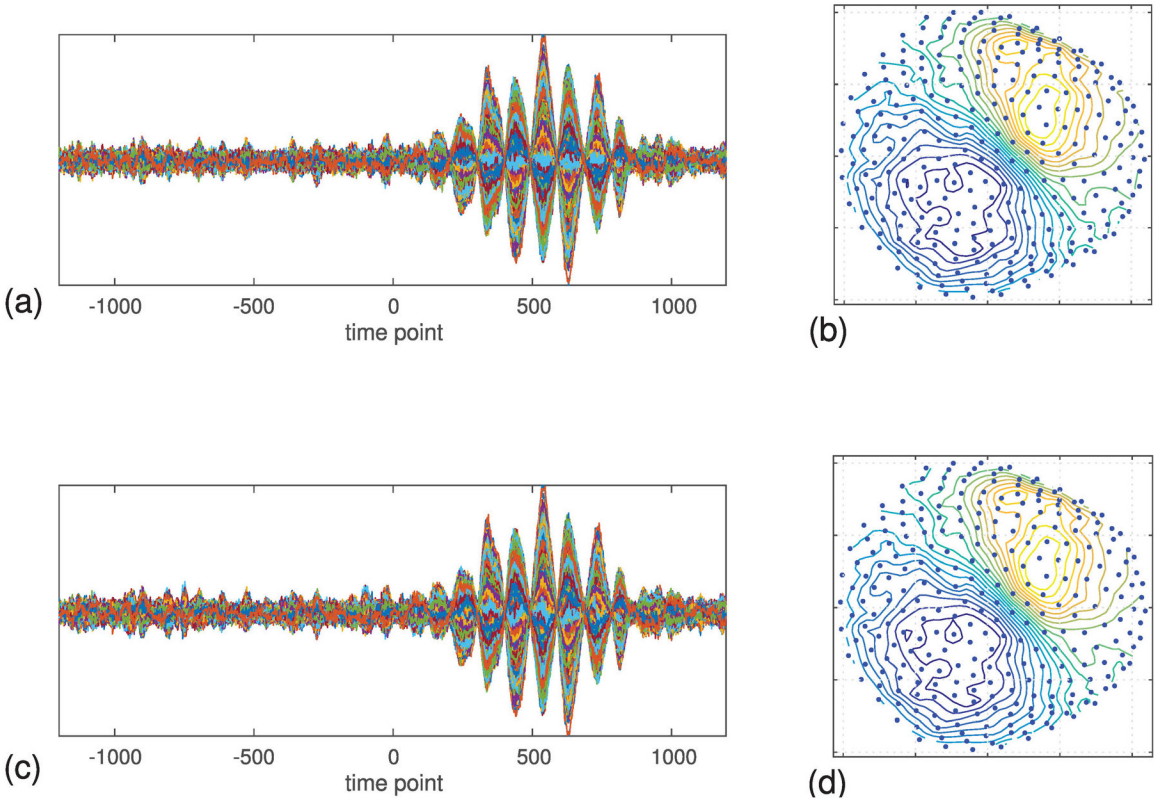

Figure 6.

Results of ANC and CTSP interference removal experiments. The reference sensor data used in the experiments are shown in figure 5(a). (a) Sensor time courses of ANC interference removal results. (b) Magnetic field map of the ANC interference-removal results. (c) Sensor time courses of CTSP interference removal results. (d) Magnetic field map of the CTSP interference-removal results. The field maps at $t = 520$ are shown for (b) and (d).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

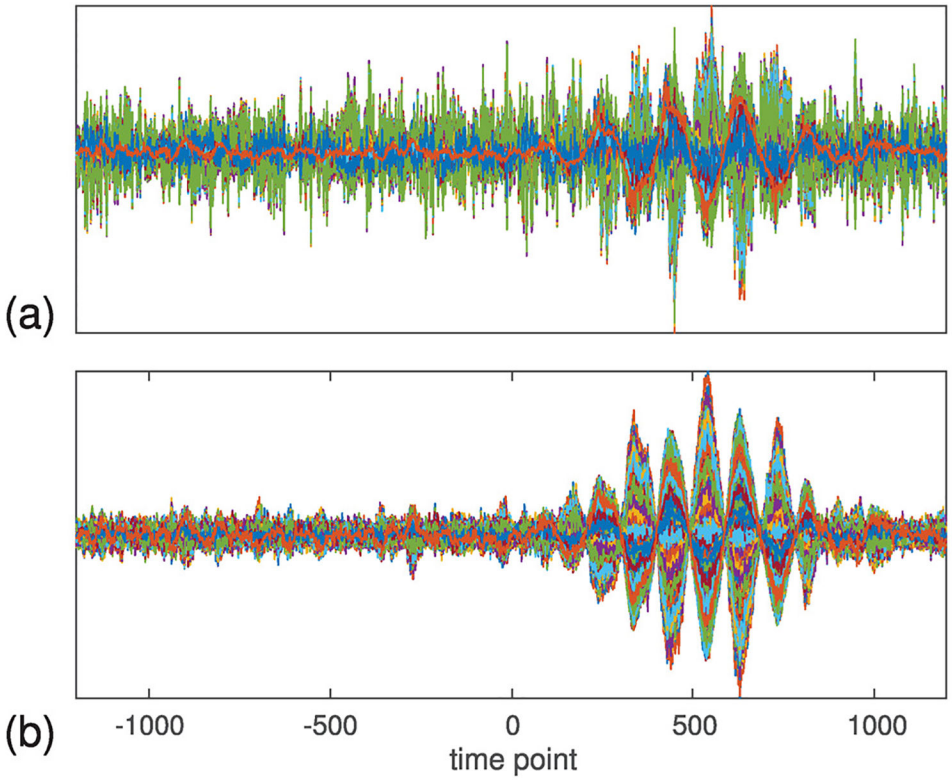

Figure 7.

Results of ANC and CTSP interference removal experiments when the reference sensor data shown in figure 5(b) were used. (a) Sensor time courses of ANC interference-removal results. (b) Sensor time courses of CTSP interference-removal results.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

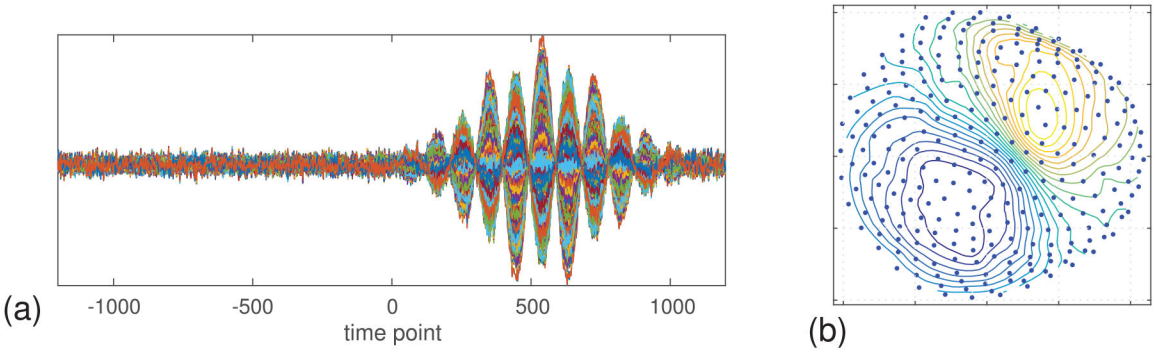

Figure 8.

Results of SSS interference-removal experiments. (a) Sensor time cources of SSS interference removal results. (b) Magnetic field map of the SSS interference removal results. The map at $t = 520$ is shown

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

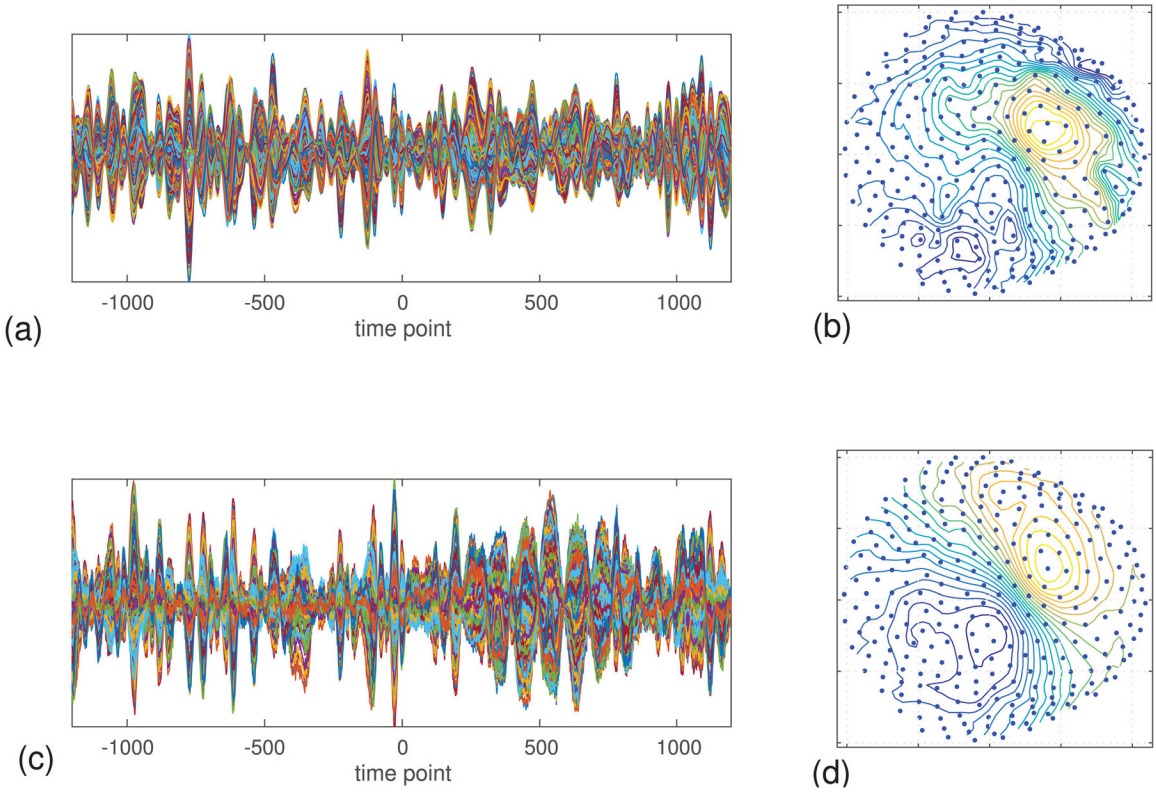

Figure 9.

(a) Sensor time courses of interference-overlapped data generated with nearby interference sources. (b) Magnetic field map of the interference-overlapped data in (a). (c) Sensor time courses of SSS interference removal results obtained from the sensor data in (a). (d) Magnetic field map of the interference removal results obtained from the sensor data in (a). The maps at $t = 520$ are shown for (b) and (d).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

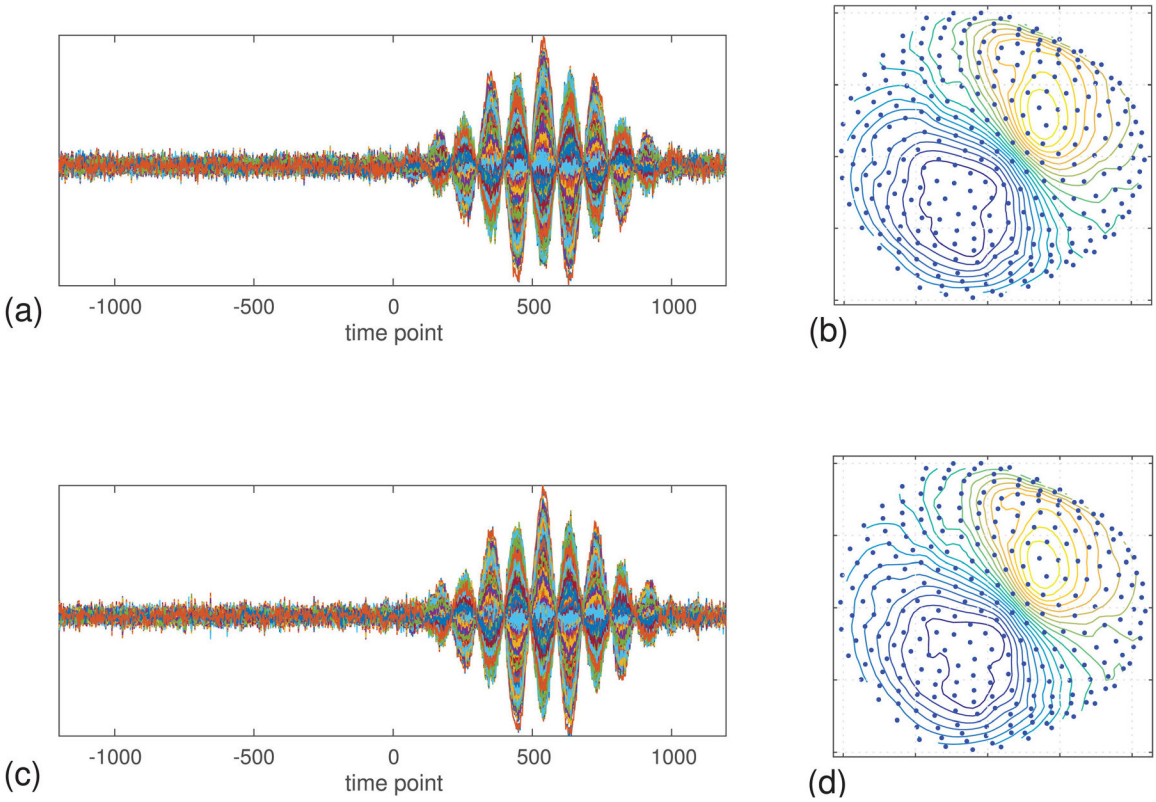

Figure 10.

Results of tSSS interference removal experiments. (a) Sensor time courses of interference removal results obtained using the interference-overlapped data in figure 3(c). (b) Magnetic field map of the interference removal results in (a). (c) Sensor time courses of interference removal results obtained using the nearby-interference data in figure 9(a). (d) Magnetic field map of the interference removal results. The maps at $t = 520$ are shown for (b) and (d).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

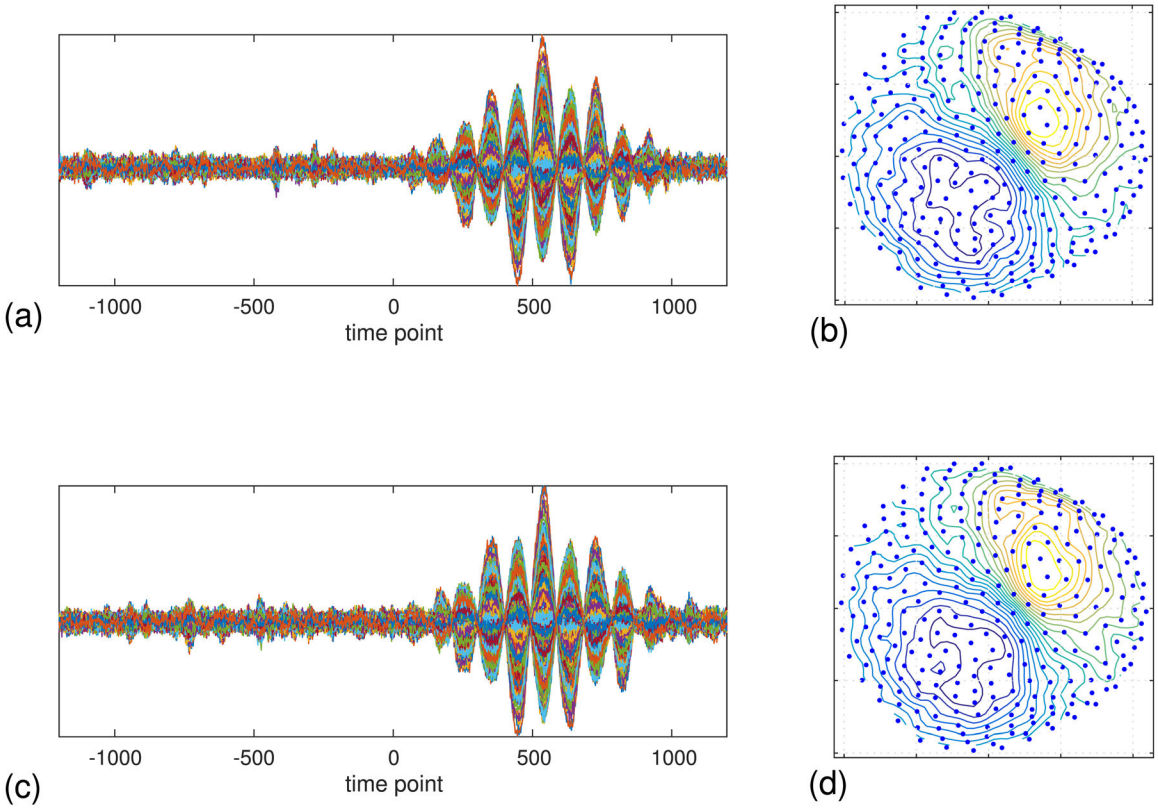

Figure 11.

Results of DSSP interference removal experiments. (a) Sensor time courses of interference removal results obtained using the interference-overlapped data in figure 3(c). (b) Magnetic field map the interference removal results in (a). (c) Sensor time courses of interference removal results obtained using the nearby-interference data in figure 9(a). (d) Magnetic field map of the interference removal results in (c). The maps at $t = 520$ are shown for (b) and (d).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

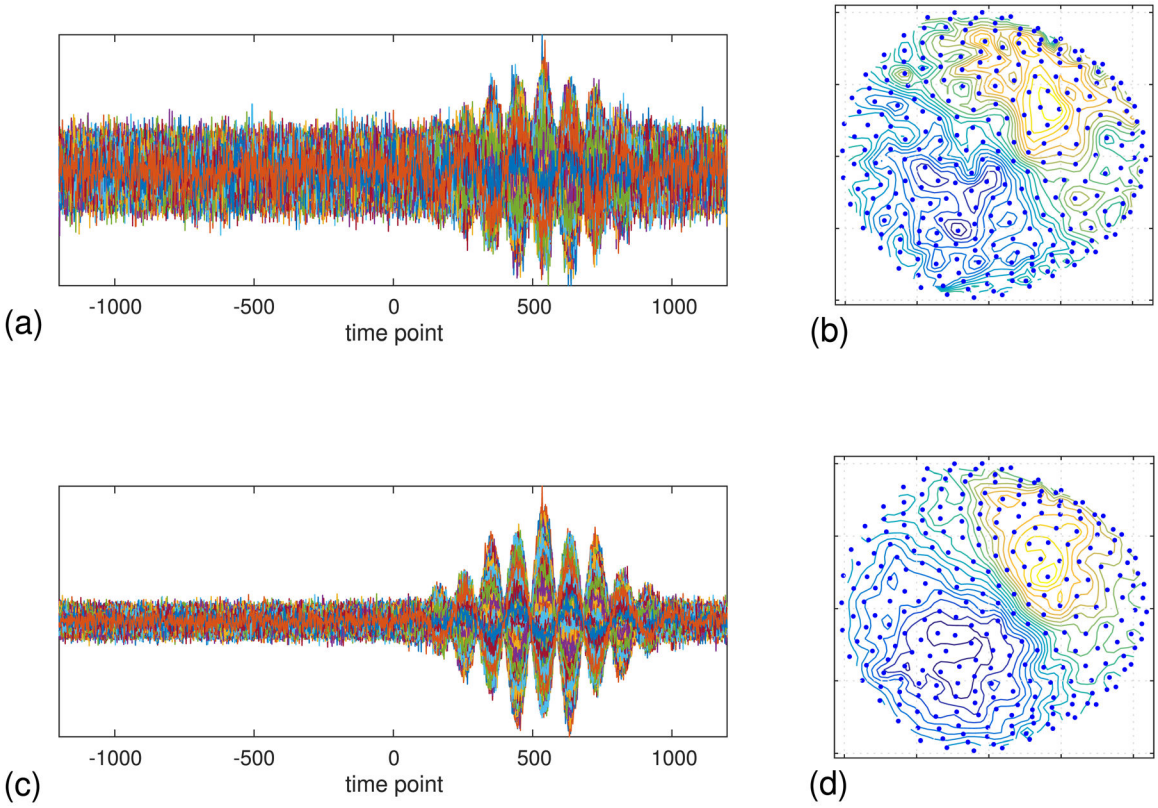

Figure 12.

Results of SNS denoising experiments. (a) Sensor time courses of signal plus Gaussian noise data generated for SNS denoising experiments. The signal-to-noise ratio (SNR) was set at 10. (b) Magnetic field map of the generated sensor data. (c) Sensor time courses of denoised results. (d) Magnetic field map of the denoised results. The maps at $t = 520$ are shown for (b) and (d).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Locations of interference sources assumed in numerical experiments.

Source number	Location (cm)	Distance from signal source (cm)
1	(100, −30, 500)	510
2	(5, −50, 600)	600
3	(−27, −495, 21)	500
4	(755, −342, 37)	830
1	(10, 30, 30)	41
2	(5, −20, −30)	39