GEVD Based Speech and Noise Correlation Matrix Estimation for Multichannel Wiener Filter Based Noise Reduction

Robbe Van Rompaey
Dept. of Electical Engineering-ESAT, STADIUS
KU Leuven
Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
robbe.vanrompaey@esat.kuleuven.be

Marc Moonen
Dept. of Electical Engineering-ESAT, STADIUS
KU Leuven
Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
marc.moonen@esat.kuleuven.be

***Abstract*—In a single speech source noise reduction scenario, the frequency domain correlation matrix of the speech signal is often assumed to be a rank-1 matrix. In multichannel Wiener filter (MWF) based noise reduction, this assumption may be used to define an optimization criterion to estimate the positive definite speech correlation matrix together with the noise correlation matrix, from sample 'speech+noise' and 'noise-only' correlation matrices. The estimated correlation matrices then define the MWF. In generalized eigenvalue decomposition (GEVD) based MWF, this optimization criterion involves a prewhitening with the sample 'noise-only' correlation matrix, which in particular leads to a compact expression for the MWF. However, a more accurate form would include a prewhitening with the estimated noise correlation matrix instead of with the sample 'noise-only' correlation matrix. Unfortunately this leads to a more difficult optimization problem, where the prewhitening indeed involves one of the optimization variables. In this paper, it is demonstrated that the modified optimization criterion, remarkably, leads to only minor modifications in the estimated correlation matrices and eventually the same MWF, which justifies the use of the original optimization criterion as a simpler substitute.**

***Index Terms*—Noise reduction, speech enhancement, Wiener filter, multichannel Wiener filter (MWF), generalized eigenvalue decomposition (GEVD).**

I. Introduction

Multichannel noise reduction is an important speech processing task in cell phones, hearing instruments and speech recognition systems. To suppress the environmental noise while minimizing speech distortion, a crucial operation is the estimation of a valid frequency domain speech correlation matrix and noise correlation matrix. The estimated correlation matrices then define the multichannel Wiener filter (MWF).

The speech and noise correlation matrices are assumed to agree with the considered scenario, i.e. in a scenario with S speech sources, the speech correlation matrix is assumed to be a positive definite rank S matrix. In this paper, we consider a single speech source scenario ($S = 1$). In multichannel Wiener filter (MWF) based noise reduction, this assumption may then be used to define an optimization criterion to estimate the rank-1 speech correlation matrix together with the noise correlation matrix, from sample 'speech+noise' and 'noise-only' correlation matrices [1]. In generalized eigenvalue decomposition (GEVD) based MWF, this optimization criterion involves a prewhitening with the sample 'noise-only' correlation matrix, which in particular leads to a compact expression for the MWF. However, a more accurate form would include a prewhitening with the estimated noise correlation matrix instead of with the sample 'noise-only' correlation matrix. Unfortunately this leads to a more difficult optimization problem, where the prewhitening indeed involves one of the optimization variables.

In this paper, it is demonstrated that the modified optimization criterion leads to only minor modifications in the estimated correlation matrices and eventually the same MWF, which justifies the use of the original optimization criterion as a simpler substitute. The conclusions from the experiments in [1] are by consequence also valid when the modified optimization criterion is adopted.

The remainder of this paper is organized as follows. In Section II the MWF is reviewed and the need for a proper estimation of the rank-1 speech and noise correlation matrix is discussed. Also some common optimization criteria to estimate the speech correlation matrix together with the noise correlation matrices, from sample 'speech+noise' and 'noise-only' correlation matrices, are presented. In Section III the modified optimization criterion is presented together with a derivation of its optimal solution.

II. Problem Statement

A. MWF Based Noise Reduction

Let N denote the number of observed microphone signals. Frequency domain processing is considered where the N complex microphone signals of a frequency bin (bin index omitted for brevity) are stacked in a vector $\boldsymbol{x}$, and consist

Acknowledgements: The work of R. Van Rompaey was supported by a doctoral Fellowship of the Research Foundation Flanders (FWO-Vlaanderen). This work was carried out at the ESAT Laboratory of KU Leuven in the frame of KU Leuven internal funding C2-16-00449 'Distributed Digital Signal Processing for Ad-hoc Wireless Local Area Audio Networking'. The scientific responsibility is assumed by its authors.

ISBN 978-90-827970-1-5 © EURASIP 2018

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 13,2021 at 14:10:40 UTC from IEEE Xplore. Restrictions apply.

of a single speech source component $\boldsymbol{x}_s$ and additive noise component $\boldsymbol{x}_n$:

$$\boldsymbol{x} = \boldsymbol{x}_s + \boldsymbol{x}_n. \tag{1}$$

If the speech and noise signals are assumed to be uncorrelated then the correlation matrices

$$R_{x_{r_1}} = E\{\boldsymbol{x}\boldsymbol{x}^H\} \tag{2}$$
$$R_{s_{r_1}} = E\{\boldsymbol{x}_s\boldsymbol{x}_s^H\} \tag{3}$$
$$R_{n_{r_1}} = E\{\boldsymbol{x}_n\boldsymbol{x}_n^H\} \tag{4}$$

can be related by

$$R_{x_{r_1}} = R_{s_{r_1}} + R_{n_{r_1}} \tag{5}$$

where $E\{.\}$ is the expected value operator. Here $R_{s_{r_1}}$ is a rank-1 matrix for a single speech source scenario.

The MWF is a linear filter estimating a specific desired signal based on the observed signals $\boldsymbol{x}$. The desired signal can be arbitrarily chosen to be the (unknown) speech component of the first microphone signal $\boldsymbol{e}_1^H\boldsymbol{x_s}$, where $\boldsymbol{e}_1$ denotes the first unity vector. The MWF minimizes the Mean Squared Error (MSE) criterion:

$$J_{MWF}(\boldsymbol{w}) = E\{|\boldsymbol{w}^H\boldsymbol{x} - \boldsymbol{e}_1^H\boldsymbol{x_s}|^2\}. \tag{6}$$

where ^{H} denotes the Hermitian transpose.

The optimal solution is given by

$$\boldsymbol{w}_{MWF} = (R_{s_{r_1}} + R_{n_{r_1}})^{-1}R_{s_{r_1}}\boldsymbol{e}_1. \tag{7}$$

The correlation matrices $R_{x_{r_1}}$ and $R_{n_{r_1}}$ are first estimated by (recursive) time-averaging during 'noise+speech' periods and 'noise only' periods respectively, where the distinction is based on a speech activity detection, assuming that the noise and speech are (spatially) stationary. This results in the sample 'speech+noise' matrix R_x and the sample 'noise-only' correlation matrix R_n. $R_{s_{r_1}}$ may then be estimated (based on (5)) as the difference between the sample 'speech+noise' correlation matrix R_x and the sample 'noise-only' correlation matrix R_n, i.e.

$$R_s = R_x - R_n. \tag{8}$$

However R_s has mostly a rank larger than one, especially in low SNR scenarios, such that better correlation matrix estimation methods are necessary. In [1], the estimation of a positive definite rank-1 speech correlation matrix $R_{s_{r_1}}$ and a corresponding noise correlation matrix $R_{n_{r_1}}$ is presented based on two different optimization criteria depending on the sample 'speech+noise' correlation matrix R_x and the sample 'noise-only' correlation matrices R_n. These are explained in the next sections.

B. EVD Based Speech and Noise Correlation Matrix Estimation

A first optimization criterion defined in [1] is used to estimate the correlation matrices as

$$\min_{R_{n_{r_1}}, R_{s_{r_1}} = \text{ rank-1}} \alpha \parallel R_x - (R_{n_{r_1}} + R_{s_{r_1}}) \parallel_F^2 + (1-\alpha) \parallel R_n - R_{n_{r_1}} \parallel_F^2 \tag{9}$$

where α is a constant ($0 \leq \alpha \leq 1$) assigning a weight to the different approximations and $\parallel . \parallel_F$ denotes the Frobenius norm. The solution to this problem is based on the symmetric eigenvalue decomposition [2] of $R_x - R_n = KDK^H$ with the eigenvalues sorted from high to low in D and is given by

$$R_{s_{r_1}} = K \cdot \text{diag}\{D_1, 0, ..., 0\} \cdot K^H \tag{10}$$
$$R_{n_{r_1}} = \alpha(R_x - R_{s_{r_1}}) + (1-\alpha)R_n. \tag{11}$$

C. GEVD Based Speech and Noise Correlation Matrix Estimation

In the optimization criterion in (9), an unweighted Frobenius norm with absolute (squared) approximation errors is used. As suggested in [1], it can be more appropriate to consider relative approximation errors depending on the noise. To this end a noise prewhitening operation is included depending on the GEVD of the matrix pencil $\{R_x, R_n\}$ [3], [4]:

$$\begin{aligned} R_n &= Q\Sigma_n Q^H \\ R_x &= Q\Sigma_x Q^H \\ \Rightarrow R_x R_n^{-1} &= Q\Sigma_x\Sigma_n^{-1}Q^{-1} = Q\Sigma Q^{-1} \end{aligned} \tag{12}$$

where Q is an invertible matrix, the columns of which are unique up to a scalar and define the generalized eigenvectors. Σ_x, Σ_n and Σ are real-valued diagonal matrices where $\Sigma_x = \text{diag}\{\sigma_{x_1}, .., \sigma_{x_N}\}, \Sigma_n = \text{diag}\{\sigma_{n_1}, .., \sigma_{n_N}\}$, and $\Sigma = \text{diag}\{\frac{\sigma_{x_1}}{\sigma_{n_1}}, .., \frac{\sigma_{x_N}}{\sigma_{n_N}}\}$ define the generalized eigenvalues sorted from high to low.

From the GEVD of the matrix pencil $\{R_x, R_n\}$, the prewhitening matrix is defined as $V_n = \Sigma_n^{-1/2}Q^{-1}$ such that $R_n = (V_n^H V_n)^{-1}$ and $V_n R_n V_n^H = I$ and (9) is reformulated as:

$$\min_{R_{n_{r_1}}, R_{s_{r_1}} = \text{ rank-1}} \alpha \parallel V_n \left(R_x - (R_{n_{r_1}} + R_{s_{r_1}})\right) V_n^H \parallel_F^2 + (1-\alpha) \parallel V_n \left(R_n - R_{n_{r_1}}\right) V_n^H \parallel_F^2 \tag{13}$$

The solution to this problem is based on the GEVD and is given by

$$R_{s_{r_1}} = Q \cdot \text{diag}\{\sigma_{x_1} - \sigma_{n_1}, 0, ..., 0\} \cdot Q^H \tag{14}$$
$$R_{n_{r_1}} = Q \cdot \text{diag}\{\sigma_{n_1}, \alpha\sigma_{x_2} + (1-\alpha)\sigma_{n_2}, ... , ..., \alpha\sigma_{x_N} + (1-\alpha)\sigma_{n_N}\} \cdot Q^H. \tag{15}$$

ISBN 978-90-827970-1-5 © EURASIP 2018

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 13,2021 at 14:10:40 UTC from IEEE Xplore. Restrictions apply.

As it was stated in [1], the GEVD effectively selects the mode with the highest SNR and it allows a more reliable estimation of the rank-1 speech correlation matrix. In the considered MWF context, the GEVD based estimation outperforms (in terms of SNR improvement achieved by the MWF) other correlation matrix estimation methods, like the first column decomposition or EVD based estimation, at the cost of a more sophisticated matrix decomposition. Also note that unlike the EVD-based MWF, the GEVD-based MWF is completely immune to scaling and linear combining of the input signals [5], i.e. the output signal and output SNR is independent of such scaling and combining, which is a desirable property.

III. Modified GEVD Based Speech and Noise Correlation Matrix Estimation

In stead of using the sample 'noise-only' correlation matrix R_n for the prewhitening as in (13), it may be more accurate to use the estimated noise correlation matrix, i.e. reformulate (13) as

$$\min_{\substack{R_{n_{r_1}} = (V^H V)^{-1} \\ R_{s_{r_1}} = \text{rank-1}}} \alpha \parallel V\left(R_x - (R_{n_{r_1}} + R_{s_{r_1}})\right) V^H \parallel_F^2 + (1-\alpha) \parallel V\left(R_n - R_{n_{r_1}}\right) V^H \parallel_F^2 \quad (16)$$

where the matrix V defines the prewhitening but is now part of the optimization problem and connected with the noise correlation matrix $R_{n_{r_1}}$ by the constraint

$$R_{n_{r_1}} = R_{n_{r_1}}^{1/2} R_{n_{r_1}}^{H/2} = V^{-1}V^{-H} = (V^H V)^{-1}. \quad (17)$$

Unfortunately the technique used to solve (13) in [1] can not be used to solve (16). However, as the next theorem states, the GEVD again forms the solution to this modified criterion.

Theorem: The matrix $R_{n_{r_1}}$ and positive definite rank-1 matrix $R_{s_{r_1}}$ that form the only stationary point of the modified optimization problem (16) are given by

$$R_{s_{r_1}} = Q \cdot \text{diag}\{\sigma_{x_1} - \sigma_{n_1}, 0, ..., 0\} \cdot Q^H \quad (18)$$

$$R_{n_{r_1}} = Q \cdot \text{diag}\{\sigma_{n_1}, \frac{\alpha\sigma_{x_2}^2 + (1-\alpha)\sigma_{n_2}^2}{\alpha\sigma_{x_2} + (1-\alpha)\sigma_{n_2}}, ... , ..., \frac{\alpha\sigma_{x_N}^2 + (1-\alpha)\sigma_{n_N}^2}{\alpha\sigma_{x_N} + (1-\alpha)\sigma_{n_N}}\} \cdot Q^H \quad (19)$$

if $\sigma_{x_1} - \sigma_{n_1} \geq 0$, else given by

$$R_{s_{r_1}} = 0 \quad (20)$$

$$R_{n_{r_1}} = Q(\alpha\Sigma_x + (1-\alpha)\Sigma_n)^{-1}(\alpha\Sigma_x^2 + (1-\alpha)\Sigma_n^2)Q^H \quad (21)$$

where Q, Σ_n, Σ_x are defined in (12).

Proof: The modified optimization criterion in (16) can be simplified using (17) leading to

$$J(V, R_{s_{r_1}}) = \alpha \parallel VR_xV^H - I - VR_{s_{r_1}}V^H \parallel_F^2 + (1-\alpha) \parallel VR_nV^H - I \parallel_F^2 . \quad (22)$$

Replacing the (unknown) Hermitian matrix VR_xV^H by its EVD defining a unitary matrix P and real diagonal matrix Λ gives

$$VR_xV^H = P\Lambda P^H \quad (23)$$

and

$$VR_xV^H - I = P(\Lambda - I)P^H. \quad (24)$$

Then it is known from low rank matrix approximation theory that the optimal positive definite rank-1 matrix $R_{s_{r_1}}$ is given by

$$VR_{s_{r_1}}V^H = \lambda_{max}\boldsymbol{u}_{max}\boldsymbol{u}_{max}^H \quad (25)$$

with $\lambda_{max} = \max(\{\Lambda_{i,i} - 1\}_{i=1..N}, 0)$, to make $R_{s_{r_1}}$ positive definite. Also $\boldsymbol{u}_{max}$ is a column of P and unitary ($\boldsymbol{u}_{max}^H\boldsymbol{u}_{max} = 1$). Here it is assumed that the largest eigenvalue λ_{max} is unique [1], so that $\boldsymbol{u}_{max}$ is uniquely determined (up to a sign ambiguity). Hence the first part in the r.h.s of (22) can be rewritten as

$$\parallel VR_xV^H - I - VR_{s_{r_1}}V^H \parallel_F^2 = \parallel VR_xV^H - I \parallel_F^2 - \lambda_{max}^2. \quad (26)$$

Assume from now on that $\lambda_{max} > 0$ and denote the corresponding largest eigenvalue of VR_xV^H with Λ_{max}. The case where $\lambda_{max} = 0$ will be discussed at the end of the proof.

The optimization criterion (22) is currently given by

$$J(V) = \alpha \parallel VR_xV^H - I \parallel_F^2 - \alpha\lambda_{max}(V)^2 + (1-\alpha) \parallel VR_nV^H - I \parallel_F^2 \quad (27)$$

$$= \alpha tr(VR_xV^HVR_xV^H - 2VR_xV^H + I) + (1-\alpha)tr(VR_nV^HVR_nV^H - 2VR_nV^H + I) - \alpha\lambda_{max}(V)^2 \quad (28)$$

where $tr()$ denotes the trace operation. To find possible stationary points of this non-constrained optimization problem (only constraint is that V is invertible), the following differential are defined:

$$\begin{aligned} df_1(V) &= dtr\left(VRV^H\right) \\ &= tr\left(d(VRV^H)\right) \\ &= tr\left(dVRV^H + VRdV^H\right) \\ &= tr\left(RV^HdV + dV^HVR\right) \end{aligned} \quad (29)$$

[1]The case where λ_{max} has a multiplicity larger than one is not considered here for the sake of a compact exposition.

ISBN 978-90-827970-1-5 © EURASIP 2018

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 13,2021 at 14:10:40 UTC from IEEE Xplore. Restrictions apply.

$$\begin{aligned} df_2(V) &= dtr\left(VRV^HVR^HV^H\right) \\ &= tr\left(d(VRV^HVR^HV^H)\right) \\ &= tr\left(dVRV^HVR^HV^H\right) + tr\left(VRdV^HVR^HV^H\right) \\ &\quad + tr\left(VRV^HdVR^HV^H\right) + tr\left(VRV^HVR^HdV^H\right) \\ &= tr\left((RV^HVR^HV^H + R^HV^HVRV^H)dV\right. \\ &\quad \left. + dV^H(VR^HV^HVR + VRV^HVR^H)\right) \end{aligned} \tag{30}$$

$$\begin{aligned} df_3(V) &= d\lambda_{max}^2(V) \\ &= 2\lambda_{max}d\lambda_{max}(V) \\ &= 2\lambda_{max}\boldsymbol{u}_{max}^H d(VR_xV^H - I)\boldsymbol{u}_{max} \\ &= 2\lambda_{max}\left(\boldsymbol{u}_{max}^H dVR_xV^H\boldsymbol{u}_{max}\right. \\ &\quad \left. + \boldsymbol{u}_{max}^H VR_xdV^H\boldsymbol{u}_{max}\right) \\ &= 2\lambda_{max}tr\left(R_xV^H\boldsymbol{u}_{max}\boldsymbol{u}_{max}^H dV\right. \\ &\quad \left. + dV^H\boldsymbol{u}_{max}\boldsymbol{u}_{max}^H VR_x\right) \end{aligned} \tag{31}$$

where the linear and cyclic properties of the trace operation are used in (29) - (31) and the derivative of an eigenvalue with multiplicity of one as discussed in [6] is used in (31).

Now using a result from [7] that if the differential $df(V)$ can be written as

$$df(V) = tr(A_0^T dV + dV^H A_1) \tag{32}$$

where A_0 and A_1 may depend on V and ^{T} denotes the transpose operation, then the partial derivatives of $f(V)$ with respect to the complex-valued matrix V and to the complex conjugate of V (denoted with V^*) is given by

$$\frac{\partial f(V)}{\partial V} = A_0 \tag{33}$$

$$\frac{\partial f(V)}{\partial V^*} = A_1. \tag{34}$$

Combining (29), (30) and (31) with (34), an equation for stationary points of $J(V)$ is obtained as

$$\begin{aligned} \frac{\partial}{\partial V^*}J &= 2\alpha VR_xV^HVR_x - 2\alpha VR_x \\ &\quad + 2(1-\alpha)VR_nV^HVR_n - 2(1-\alpha)VR_n \\ &\quad - 2\alpha\lambda_{max}\boldsymbol{u}_{max}\boldsymbol{u}_{max}^H VR_x \\ &= 0. \end{aligned} \tag{35}$$

Dividing both sides by 2 and multiplying with $V^H\boldsymbol{u}_{max}$ gives

$$\begin{aligned} 0 &= \alpha VR_xV^HVR_xV^H\boldsymbol{u}_{max} - \alpha VR_xV^H\boldsymbol{u}_{max} \\ &\quad + (1-\alpha)VR_nV^HVR_nV^H\boldsymbol{u}_{max} \\ &\quad - (1-\alpha)VR_nV^H\boldsymbol{u}_{max} \\ &\quad - \alpha\lambda_{max}\boldsymbol{u}_{max}\boldsymbol{u}_{max}^H VR_xV^H\boldsymbol{u}_{max} \\ &= \alpha\Lambda_{max}^2\boldsymbol{u}_{max} - \alpha\Lambda_{max}\boldsymbol{u}_{max} \\ &\quad + (1-\alpha)(VR_nV^HVR_nV^H - VR_nV^H)\boldsymbol{u}_{max} \\ &\quad - \alpha(\Lambda_{max}-1)\Lambda_{max}\boldsymbol{u}_{max}. \end{aligned} \tag{36}$$

This simplifies to

$$(VR_nV^HVR_nV^H - VR_nV^H)\boldsymbol{u}_{max} = 0. \tag{37}$$

Hence $\boldsymbol{u}_{max}$ is an eigenvector of $VR_nV^HVR_nV^H - VR_nV^H = VR_nV^H(VR_nV^H - I) = (VR_nV^H - I)VR_nV^H$ with eigenvalue zero, so also an eigenvector of VR_nV^H with eigenvalue one or zero. Since R_n and V are assumed to be invertible, VR_nV^H must be nonsingular and hence $\boldsymbol{u}_{max}$ must be an eigenvector of VR_nV^H with eigenvalue one. If $\boldsymbol{u}_{max}$ is an eigenvector of VR_xV^H and of VR_nV^H, it is also an eigenvector with eigenvalue Λ_{max} of

$$(VR_xV^H)(VR_nV^H)^{-1} = VR_xR_n^{-1}V^{-1}. \tag{38}$$

By substituting the GEVD (12) of $\{R_x, R_n\}$ in (38), the eigenvalue decomposition of (38) is obtained as

$$VR_xR_n^{-1}V^{-1} = \tilde{V}(\Sigma_x\Sigma_n^{-1})\tilde{V}^{-1} \tag{39}$$

with $\tilde{V} = VQ$.

Hence $\boldsymbol{u}_{max}$ is the normalized version the column of $\tilde{V}$ corresponding to the eigenvalue Λ_{max}. Assume that $\frac{\sigma_{x_1}}{\sigma_{n_1}} = \Lambda_{max}$. Then $\boldsymbol{u}_{max}$ can be written as

$$\boldsymbol{u}_{max} = \frac{\tilde{V}e_1}{\|\tilde{V}e_1\|} = \frac{1}{\sigma_{v_1}^{1/2}}\tilde{V}e_1. \tag{40}$$

Plugging (40) and (12) into the stationary point equation (35), results in

$$\begin{aligned} 0 &= \alpha\tilde{V}\Sigma_x\tilde{V}^H\tilde{V}\Sigma_xQ^H - \alpha\tilde{V}\Sigma_xQ^H \\ &\quad + (1-\alpha)\tilde{V}\Sigma_n\tilde{V}^H\tilde{V}\Sigma_nQ^H - (1-\alpha)\tilde{V}\Sigma_nQ^H \\ &\quad - \alpha\frac{\lambda_{max}}{\sigma_{v_1}}\tilde{V}e_1e_1^H\tilde{V}^H\tilde{V}\Sigma_xQ^H \end{aligned} \tag{41}$$

Since Q^H and $\tilde{V}$ are invertible, right-multiplying with $\tilde{V}^{-1}$ and left-multiplying with Q^{-H} gives

$$\begin{aligned} 0 &= \alpha\Sigma_x\Sigma_V\Sigma_x + (1-\alpha)\Sigma_n\Sigma_V\Sigma_n - \alpha\frac{\lambda_{max}}{\sigma_{v_1}^2}e_1e_1^H\Sigma_V\Sigma_x \\ &\quad - \alpha\Sigma_x - (1-\alpha)\Sigma_n \end{aligned} \tag{42}$$

where $\Sigma_V = \tilde{V}^H\tilde{V}$ and $\sigma_{v_1} = (\Sigma_V)_{1,1}$.

It can be verified that in general Σ_V has to be diagonal ($\Sigma_V = \text{diag}\{\sigma_{v_1}, .., \sigma_{v_N}\}$) to solve (42). Using the fact that $\lambda_{max} = \Lambda_{max} - 1 = \frac{\sigma_{x_1}}{\sigma_{n_1}} - 1$, finally Σ_V is obtained as given in (43) at the top of the next page.

The optimal solution is thus shown to be

$$V = \Sigma_V^{1/2}Q^{-1} \tag{44}$$

$$R_{n_{r_1}} = (V^HV)^{-1} = Q\Sigma_V^{-1}Q^H \tag{45}$$

$$\begin{aligned} R_{s_{r_1}} &= V^{-1}u_{max}\lambda_{max}u_{max}^HV^{-H} \\ &= \frac{1}{\sigma_{v_1}}Q\Sigma_V^{-1/2}\Sigma_V^{1/2}e_1(\frac{\sigma_{x_1}}{\sigma_{n_1}} - 1)e_1^H\Sigma_V^{1/2}\Sigma_V^{-1/2}Q^H \\ &= Q\cdot\text{diag}\{\sigma_{x_1} - \sigma_{n_1}, 0, ..., 0\}\cdot Q^H \end{aligned} \tag{46}$$

ISBN 978-90-827970-1-5 © EURASIP 2018

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 13,2021 at 14:10:40 UTC from IEEE Xplore. Restrictions apply.

$$\Sigma_V = \text{diag}\{\alpha\sigma_{x_1}^2 + (1-\alpha)\sigma_{n_1}^2, \alpha\sigma_{x_2}^2 + (1-\alpha)\sigma_{n_2}^2, ..., \alpha\sigma_{x_N}^2 + (1-\alpha)\sigma_{n_N}^2\}^{-1}$$
$$\left(\begin{bmatrix} \alpha\sigma_{x_1} + (1-\alpha)\sigma_{n_1} & 0 & \cdots & 0 \\ 0 & \alpha\sigma_{x_2} + (1-\alpha)\sigma_{n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha\sigma_{x_N} + (1-\alpha)\sigma_{n_N} \end{bmatrix} - \begin{bmatrix} \alpha\lambda_{max}\sigma_{x_1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \right)$$
$$= \text{diag}\{\sigma_{n_1}^{-1}, \frac{\alpha\sigma_{x_2} + (1-\alpha)\sigma_{n_2}}{\alpha\sigma_{x_2}^2 + (1-\alpha)\sigma_{n_2}^2}, ..., \frac{\alpha\sigma_{x_N} + (1-\alpha)\sigma_{n_N}}{\alpha\sigma_{x_N}^2 + (1-\alpha)\sigma_{n_N}^2}\}. \quad (43)$$

It remains to show that Λ_{max} is equal to the largest generalized eigenvalue ratio $\frac{\sigma_{x_1}}{\sigma_{n_1}}$. This can be seen from the fact that Λ_{max} is the largest positive eigenvalue of $VR_xV^H = P\Lambda P^H$. Indeed the eigenvalues of VR_xV^H can be determined using (12) and (44):

$$\Lambda = \text{diag}\{\frac{\sigma_{x_1}}{\sigma_{n_1}}, \{\frac{\alpha\frac{\sigma_{x_i}^2}{\sigma_{n_i}^2} + (1-\alpha)\frac{\sigma_{x_i}}{\sigma_{n_i}}}{\alpha\frac{\sigma_{x_i}^2}{\sigma_{n_i}^2} + (1-\alpha)}\}_{i=2..N}\}. \quad (47)$$

It suffices that $\frac{\sigma_{x_1}}{\sigma_{n_1}} (\geq 1) \geq \max(\{\frac{\sigma_{x_i}}{\sigma_{n_i}}\}_{i=2..N})$ since then

$$\alpha(\frac{\sigma_{x_1}}{\sigma_{n_1}} - 1)\frac{\sigma_{x_i}^2}{\sigma_{n_i}^2} + (1-\alpha)(\frac{\sigma_{x_1}}{\sigma_{n_1}} - \frac{\sigma_{x_i}}{\sigma_{n_i}}) \geq 0 \quad \forall i \quad (48)$$

or equivalently

$$\frac{\sigma_{x_1}}{\sigma_{n_1}} \geq \frac{\alpha\frac{\sigma_{x_i}^2}{\sigma_{n_i}^2} + (1-\alpha)\frac{\sigma_{x_i}}{\sigma_{n_i}}}{\alpha\frac{\sigma_{x_i}^2}{\sigma_{n_i}^2} + (1-\alpha)} \quad \forall i. \quad (49)$$

From this it is seen that $\frac{\sigma_{x_1}}{\sigma_{n_1}}$ is the largest eigenvalue of VR_xV^H.

The previous derivation up to (43) is also valid for the trivial case where $\lambda_{max} = 0$. For this case the solution is obtained as

$$R_{s_{r_1}} = 0 \quad (50)$$
$$R_{n_{r_1}} = Q \cdot \text{diag}\{\frac{\alpha\sigma_{x_1}^2 + (1-\alpha)\sigma_{n_1}^2}{\alpha\sigma_{x_1} + (1-\alpha)\sigma_{n_1}}, \frac{\alpha\sigma_{x_2}^2 + (1-\alpha)\sigma_{n_2}^2}{\alpha\sigma_{x_2} + (1-\alpha)\sigma_{n_2}}, ...$$
$$..., \frac{\alpha\sigma_{x_N}^2 + (1-\alpha)\sigma_{n_N}^2}{\alpha\sigma_{x_N} + (1-\alpha)\sigma_{n_N}}\} \cdot Q^H$$
$$= Q(\alpha\Sigma_x + (1-\alpha)\Sigma_n)^{-1}(\alpha\Sigma_x^2 + (1-\alpha)\Sigma_n^2)Q^H. \quad (51)$$

where the optimal $R_{n_{r_1}}$ is simply a non-linear interpolation depending on α between R_x and R_n. □

Remark: An extension of the proof to a scenario with S speech sources is possible in a similar way where the optimal rank-S speech correlation matrix $R_{s_{r_S}}$ is then given by:

$$R_{s_{r_S}} = Q \cdot \text{diag}\{\sigma_{x_1} - \sigma_{n_1}, ..., \sigma_{x_S} - \sigma_{n_S}, 0, ..., 0\} \cdot Q^H \quad (52)$$

where $\sigma_{x_1} - \sigma_{n_1}, ..., \sigma_{x_S} - \sigma_{n_S}$ are the S largest generalized eigenvalue differences.

While the optimal $R_{s_{r_1}}$ in (18) is the same as in (14), the $R_{n_{r_1}}$ in (19) is different from (15). The resulting α-dependent interpolation between the generalized eigenvalues is more involved in (19) than the simpler linear interpolation in (15). However the resulting MWF (7) is straightforwardly shown to be the same in both cases and is given by

$$\boldsymbol{w}_{MWF} = Q^{-H} \cdot \text{diag}\{\frac{\sigma_{x_1} - \sigma_{n_1}}{\sigma_{x_1}}, 0, ..., 0\} \cdot Q^H \boldsymbol{e}_1. \quad (53)$$

This justifies the use of the original optimization criterion (13) as a simpler substitute. Formula (53) demonstrates also that the GEVD based $\boldsymbol{w}_{MWF}$ is independent of the value of α in the optimization criterion, which is a desirable property.

REFERENCES

[1] R. Serizel, M. Moonen, B. V. Dijk, and J. Wouters, "Low-rank approximation based multichannel wiener filter algorithms for noise reduction with application in cochlear implants," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 4, pp. 785–799, April 2014.

[2] ——, "Rank-1 approximation based multichannel wiener filtering algorithms for noise reduction in cochlear implants," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 8634–8638.

[3] M. Dendrinos, S. Bakamidis, and G. Carayannis, "Speech enhancement from noise: A regenerative approach," *Speech Communication*, vol. 10, no. 1, pp. 45 – 57, 1991. [Online]. Available: http://www.sciencedirect.com/science/article/pii/016763939190027Q

[4] S. Doclo and M. Moonen, "Gsvd-based optimal filtering for single and multimicrophone speech enhancement," *IEEE Transactions on Signal Processing*, vol. 50, no. 9, pp. 2230–2244, Sep 2002.

[5] A. Hassani, A. Bertrand, and M. Moonen, "Cooperative integrated noise reduction and node-specific direction-of-arrival estimation in a fully connected wireless acoustic sensor network," *Signal Processing*, vol. 107, pp. 68 – 81, 2015, special Issue on ad hoc microphone arrays and wireless acoustic sensor networks Special Issue on Fractional Signal Processing and Applications. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168414004071

[6] J. R. Magnus, "On differentiating eigenvalues and eigenvectors," *Econometric Theory*, vol. 1, no. 2, pp. 179–191, 1985. [Online]. Available: http://www.jstor.org/stable/3532409

[7] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2740–2746, June 2007.

ISBN 978-90-827970-1-5 © EURASIP 2018

Authorized licensed use limited to: KU Leuven Libraries. Downloaded on September 13,2021 at 14:10:40 UTC from IEEE Xplore. Restrictions apply.