IET Signal Processing

Research Article

IET Journals
The Institution of Engineering and Technology

Low-Complexity separable beamformers for massive antenna array systems

ISSN 1751-9675
Received on 1st May 2018
Revised 21st November 2018
Accepted on 18th January 2019
E-First on 19th February 2019
doi: 10.1049/iet-spr.2018.5115
www.ietdl.org

Lucas N. Ribeiro[1] ✉, André L.F. de Almeida[1], Josef A. Nossek[1,2], João César M. Mota[1]

[1]*Wireless Telecommunications Research Group, Federal University of Ceará, Fortaleza, Brazil*
[2]*Department of Electrical and Computer Engineering, Technical University Munich, Munich, Germany*
✉ *E-mail: nogueira@gtel.ufc.br*

Abstract: Future cellular systems will likely employ massive bi-dimensional arrays to improve performance by large array gain and more accurate spatial filtering, motivating the design of low-complexity signal-processing methods. The authors propose optimising a Kronecker-separable beamforming filter that takes advantage of the bi-dimensional array geometry to reduce computational costs. The Kronecker factors are obtained using two strategies: alternating optimisation and sub-array minimum mean square error (MMSE) beamforming with Tikhonov regularisation. According to the simulation results, the proposed methods are computationally efficient but come with source recovery degradation, which becomes negligible when the sources are sufficiently separated in space.

1 Introduction

The number of wireless connected devices has been growing significantly, bringing new challenges to engineers. Future mobile communications systems, for example, are expected to provide very high throughput to several mobile terminals. In order to boost system capacity, new transceiver and network architectures are under investigation. Massive multiple-input-multiple-output (MIMO) technology, which consists of employing a large number of antenna elements at the base station (BS) to serve many multi-antenna users, is expected to yield significant spectral efficiency improvement [1, 2]. Such massive systems should be implemented using planar arrays in order to reduce the array's physical dimensions and to perform elevation and azimuth beamforming. This implementation, known as full-dimension MIMO (FD-MIMO), allows for better interference mitigation and has already been incorporated into 3GPP standards [3]. These technologies pose new engineering challenges concerning computational and energy efficiency [4], calling for research efforts to design computationally efficient signal processing methods for high-dimensional systems.

The high computational complexity of multidimensional filtering systems is not a new problem, though, and the first attempts to tackle this problem can be traced back to some decades ago. For instance, the authors in [5] proposed a multi-stage representation for bi-dimensional filters based on the coefficient matrix eigen decomposition, yielding computer storage, and speed savings. However, computing the eigen decomposition of high-dimensional observations is expensive in general. More recent works have been interested in exploiting the algebraic structure present in some problems to reduce computational costs and to improve system performance. The authors in [6] introduced tensor-based blind source separation methods which reduce the number of parameters to be estimated by exploiting the structure of low-rank signals. Such property implies that signals can be well approximated by a finite sum of low-dimensional Kronecker products. Although this representation simplifies the parameter estimation problem, signal accuracy is degraded. Kronecker separability has also been exploited in [7–9] to increase the convergence rate of adaptive algorithms. Gradient descent-based solutions were presented in [7, 8] to identify second-order Kronecker separable systems, which can be useful to model telephone hybrid-causing electrical echoes. In [9], the authors show that Volterra systems with separable kernels can be expressed in terms of Kronecker products. In [10], we introduce a supervised system identification method to identify third-order Kronecker separable impulse responses based on alternating optimisation. The proposed identification method is applied to identify the telephone hybrid-like impulse response of [7]. Simulation results indicate that the proposed method exhibits better accuracy than the classical Wiener filter solution. In [11], the method of [10] is extended to cope with low-rank Kronecker separable systems, allowing for the identification of more intricate acoustic responses. In [12], fast recursive least squares methods for identifying second-order Kronecker separable (bilinear) systems are presented. Analytical and simulation results confirm the low computational costs and the identification performance of the proposed bilinear methods.

It is well-known that the spatial signature of planar arrays can be decomposed along its two dimensions [13]. Based on this property, beamforming techniques have been proposed. The authors in [14] obtained a low-complexity two-dimensional MIMO precoding scheme by exploiting the Kronecker structure in the steering vectors of rectangular arrays. Therein, the proposed separable zero-forcing (ZF) precoder is presented based on the small angular spread at the elevation domain assumption, which enables algebraic separation of the azimuth and elevation domains by filtering. Results show that when this assumption is satisfied, the proposed separable ZF filter exhibits acceptable performance. However, in more realistic scenarios, this assumption is seldom met, and the performance of the separable ZF filter is severely degraded. In [15], a clever hybrid analogue/digital beamforming method based on the Kronecker product is proposed for multi-cell multi-user MIMO systems. The analogue beamformers are designed exploiting the mixed product property of the Kronecker product to null inter-cell interference and to enhance the desired signal power. In [16], the authors investigate the performance of a tensor global sidelobe canceller (GSC). Simulation results suggest that this tensor-based beamformer requires fewer snapshots than the classical GSC filter to achieve the desired performance. A tensor minimum variance distortionless response beamformer has been introduced in [17] for polarisation sensitive arrays. In [18], we express the received signal vector of a massive MIMO system equipped with a planar array using multi-linear (tensor) algebra. From this model, we derived a two-step low-complexity equaliser which exploits each signal dimension, similar to [14]. It basically consists of sub-array ZF beamforming followed by a low-dimensional minimum mean square error (MMSE) equaliser.

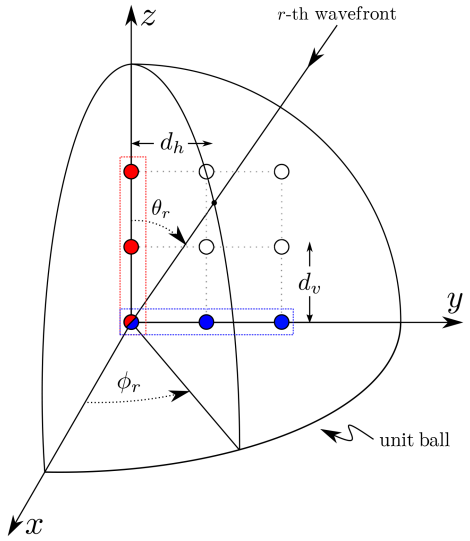

Fig. 1 *Uniform Rectangular Array (URA) in the y – z plane*

In this present work, we propose novel beamforming techniques which exploit the array separability to reduce their computational costs. Our methods are based on the classical MMSE beamformer, also known as Wiener filter [19]. This filter may be computationally impractical due to the inversion of a possibly very large covariance matrix. The matrix inversion lemma [20] can be applied to the MMSE beamformer when the signal statistics are perfectly known, however, in practice, this seldom happens. As alternatives to the classical MMSE solution, we propose methods which aim at optimising a beamforming filter with Kronecker structure, i.e. the coefficients vector admits a Kronecker factorisation. Thus, instead of optimising a large beamforming vector, we propose designing two relatively small beamforming vectors corresponding to the Kronecker factors. We present two strategies to design a Kronecker separable filter. In the first strategy, the mean square error (MSE) function is minimised by means of alternating optimisation. This strategy was first introduced in [21], where the beamforming filter is obtained using sample estimates of the received signal covariance matrix. Here, we derive analytical expressions for the beamformer assuming perfect knowledge of the array manifold matrix. The second strategy consists of a closed-form solution based on the Khatri-Rao factorisation of the separable array manifold matrix. Each sub-beamformer is obtained by performing sub-array MMSE beamforming with Tikhonov regularisation. Simulation results show that the proposed methods can be computationally efficient, however, they come with source recovery degradation, which becomes insignificant when the wavefronts are sufficiently separated in the space.

The following notation is adopted throughout the paper: x denotes a scalar, $\boldsymbol{x}$ a vector, and $\boldsymbol{X}$ a matrix. The (i, j)th entry of $\boldsymbol{X}$ is given by $[\boldsymbol{X}]_{i,j}$. The transposed, conjugated transposed (Hermitian), and pseudo-inverse of $\boldsymbol{X}$ are denoted by $\boldsymbol{X}^{\mathrm{T}}$, $\boldsymbol{X}^{\mathrm{H}}$, and $\boldsymbol{X}^{\dagger}$, respectively. The $(M \times M)$-dimensional identity matrix is represented by $\boldsymbol{I}_M$. The absolute value, the Frobenius and ℓ_2 norms, and the expected value operator are, respectively, denoted by $|\cdot|$, $\|\cdot\|_{\mathrm{F}}$, $\|\cdot\|_2$, and $\mathbb{E}-[\cdot]$. The Kronecker, Khatri-Rao, and n-mode products are represented by $\otimes$, $\diamond$, and $\times_n$, respectively. $O(\cdot)$ represents the Big-O notation.

This work is organised as follows: the system model is introduced in Section 2 and the proposed beamforming methods are presented in Section 3. Therein, we also discuss their computational complexity. Simulation results are shown and discussed in Section 4, and the work is concluded in Section 5.

2 System model

Consider a multi-antenna system equipped with a uniform rectangular array (URA) with N_h antennas in the horizontal axis, and N_v in the vertical axis. This array of $N = N_h N_v$ antennas is distributed along the y–z plane, as illustrated in Fig. 1. Each antenna element has the same beam pattern $g(\phi, \theta)$, where ϕ and θ denote the azimuth and elevation angles, respectively [In practical antenna arrays, the element beampatterns would be different due to phenomena like mutual coupling, among others. To model such scenario, one would need to consider individual antenna beampatterns $g_n(\phi, \theta)$ for all $n \in \{1, \ldots, N\}$.]. The array is illuminated by R independent narrow-band wavefronts in far-field propagation arriving from directions $(\phi_r, \theta_r), r = 1, \ldots, R$ and carrying digitally modulated signals. The wavefronts are assumed to have the same wavelength λ. The modulated signals at discrete-time instant k are denoted by $s_r[k]$ and assumed to be mutually uncorrelated with zero mean and variance σ_s^2.

The received signal at the nth antenna can be modelled as the superposition of the R incoming wavefronts:

$$x_n[k] = \sum_{r=1}^{R} g(\phi_r, \theta_r) a_n(\phi_r, \theta_r) s_r[k] + b_n[k], \tag{1}$$

where $a_n(\phi_r, \theta_r)$ denotes the array response to the r -th wavefront at the n -th antenna, and $b_n[k]$ the complex additive white Gaussian sensor noise (AWGN) with zero mean and variance σ_b^2. The inter-antenna spacing in both the horizontal and vertical axes is $d_h = d_v = \lambda/2$, thus the array response can be written as

$$a_n(\phi_r, \theta_r) = e^{j\pi[(n_h-1)\sin\phi_r \sin\theta_r + (n_v-1)\cos\theta_r]}.$$

with $n = n_h + (n_v - 1)N_h$, $n_h \in \{1, \ldots, N_h\}$, $n_v \in \{1, \ldots, N_v\}$. For notation simplicity, we define direction cosines with respect to the horizontal and vertical axis as $p_r = \sin\phi_r \sin\theta_r$ and $q_r = \cos\theta_r$, respectively. Then, using matrix notation and assuming omni-directional antennas, the received signals vector $\boldsymbol{x}[k] = [x_1[k], \ldots, x_N[k]]^{\mathrm{T}}$ can be represented as

$$\boldsymbol{x}[k] = \sum_{r=1}^{R} \boldsymbol{a}(p_r, q_r) s_r[k] + \boldsymbol{b}[k] = \boldsymbol{A}\boldsymbol{s}[k] + \boldsymbol{b}[k], \tag{2}$$

where $\boldsymbol{a}(p_r, q_r) = [a_1(p_r, q_r), \ldots, a_N(p_r, q_r)]^{\mathrm{T}}$ stands for the array steering vector, $\boldsymbol{s}[k] = [s_1[k], \ldots, s_R[k]]^{\mathrm{T}}$ the symbols vector, and $\boldsymbol{b}[k] = [b_1[k], \ldots, b_N[k]]^{\mathrm{T}}$ the AWGN vector. Note that the model (2) is valid only for a specific angular range where $g(\phi_r, \theta_r) = 1$. Now the array manifold matrix can be written as

$$\boldsymbol{A} = [\boldsymbol{a}(p_1, q_1), \ldots, \boldsymbol{a}(p_R, q_R)] \in \mathbb{C}^{N \times R}. \tag{3}$$

From our assumptions, it follows that the covariance matrix of the received signals is given by

$$\boldsymbol{R}_{xx} = \mathbb{E}[\boldsymbol{x}[k]\boldsymbol{x}[k]^{\mathrm{H}}] = \boldsymbol{A}\boldsymbol{R}_{ss}\boldsymbol{A}^{\mathrm{H}} + \boldsymbol{R}_{bb},$$

where $\boldsymbol{R}_{ss} = \mathbb{E}[\boldsymbol{s}[k]\boldsymbol{s}[k]^{\mathrm{H}}] = \sigma_s^2 \boldsymbol{I}_R$ and $\boldsymbol{R}_{bb} = \mathbb{E}[\boldsymbol{b}[k]\boldsymbol{b}[k]^{\mathrm{H}}] = \sigma_b^2 \boldsymbol{I}_N$. The multi-antenna system employs a beamformer to recover a desired signal among the R incoming signals. We define the signal-to-noise ratio (SNR) as the desired signal power over the AWGN variance, i.e. $\mathrm{SNR} = \sigma_s^2/\sigma_b^2$.

The array response can be *separated* into horizontal and vertical contributions owing to the URA bi-dimensionality [13]. More specifically, the array response with respect to any wavefront can be factorised as

$$a_n(p_r, q_r) = a_{n_h}(p_r) a_{n_v}(q_r), \tag{4}$$

where $a_{n_h}(p_r) = e^{j\pi(n_h-1)p_r}$ and $a_{n_v}(q_r) = e^{j\pi(n_v-1)q_r}$. The sub-array steering vectors are then defined as

$$\boldsymbol{a}_h(p_r) = \left[a_1(p_r), \ldots, a_{N_h}(p_r)\right]^{\mathrm{T}},$$

$$\boldsymbol{a}_v(q_r) = \left[a_1(q_r), \ldots, a_{N_v}(q_r)\right]^{\mathrm{T}}.$$

The horizontal and vertical sub-arrays of a URA are depicted in Fig. 1. The separable representation in (4) leads to the Kronecker factorisation of the array steering vectors:

$$\boldsymbol{a}(p_r, q_r) = \boldsymbol{a}_v(q_r) \otimes \boldsymbol{a}_h(p_r),$$

and, consequently, the array manifold matrix (3) can be written as a Khatri-Rao product

$$\boldsymbol{A} = [\boldsymbol{a}_v(q_1) \otimes \boldsymbol{a}_h(p_1), \ldots, \boldsymbol{a}_v(q_R) \otimes \boldsymbol{a}_h(p_R)] = \boldsymbol{A}_v \diamond \boldsymbol{A}_h, \tag{5}$$

where

$$\boldsymbol{A}_h = [\boldsymbol{a}_h(p_1), \ldots, \boldsymbol{a}_h(p_R)] \in \mathbb{C}^{N_h \times R},$$

$$\boldsymbol{A}_v = [\boldsymbol{a}_v(q_1), \ldots, \boldsymbol{a}_v(q_R)] \in \mathbb{C}^{N_v \times R},$$

stand for the vertical and horizontal sub-array manifold matrices, respectively. Equation (5) emphasises the separable structure of the URA and shall be exploited in beamforming design.

3 Beamforming methods

We are interested in spatially filtering the received signals $\boldsymbol{x}[k]$ to extract $s_d[k]$, the signal of dth (desired) wavefront, while attenuating the interfering signals. To this end, we design the beamforming filter $\boldsymbol{w} \in \mathbb{C}^N$ so that its output $y[k] = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{x}[k]$ approximates the desired signal. We choose to optimise this filter to minimise the mean square error (MSE) function

$$\begin{aligned} J_{\mathrm{MSE}}(\boldsymbol{w}) &= \mathbb{E}\left[\left|s_d[k] - \boldsymbol{w}^{\mathrm{H}}\boldsymbol{x}[k]\right|^2\right] \\ &= \sigma_s^2 - \boldsymbol{p}_{xs}^{\mathrm{H}}\boldsymbol{w} - \boldsymbol{w}^{\mathrm{H}}\boldsymbol{p}_{xs} + \boldsymbol{w}^{\mathrm{H}}\boldsymbol{R}_{xx}\boldsymbol{w}, \end{aligned} \tag{6}$$

where $\boldsymbol{p}_{xs} = \mathbb{E}[\boldsymbol{x}[k]s_d^*[k]] = \boldsymbol{A}\boldsymbol{R}_{ss}\boldsymbol{e}_d \in \mathbb{C}^N$ denotes the cross-covariance vector, and $\boldsymbol{e}_r \in \mathbb{C}^R$ the rth canonical vector in the R-dimensional space. The MMSE beamformer yields the global minimum of $J_{\mathrm{MSE}}(\boldsymbol{w})$ and is given by the Wiener filter $\boldsymbol{w}_{opt} = \boldsymbol{R}_{xx}^{-1}\boldsymbol{p}_{xs}$ [19]. For large array systems, the computation of this filter becomes impractical since it involves the inversion of a very large covariance matrix. Iterative algorithms, such as the gradient descent method, can be used to simplify the calculations, however each of their iterations can still be computationally expensive.

To simplify the calculations of the MMSE beamforming filter, we impose the following Kronecker structure: $\boldsymbol{w} = \boldsymbol{w}_v \otimes \boldsymbol{w}_h$, $\boldsymbol{w}_m \in \mathbb{C}^{N_m}$, $m \in \{v, h\}$. Such a representation is motivated by the computational reduction of the beamformer design, since only $(N_v + N_h)$ parameters need to be optimised, against N_vN_h when separability is not considered. In order to gain more insight into the array separability, let us consider an example with N antennas and $R = 1$ impinging wavefront. The received signal in this case is given by $\boldsymbol{x}[k] = \boldsymbol{a}(p_d, q_d)s[k] + \boldsymbol{b}[k]$. The output signal for the filter $\boldsymbol{w}$ is then written as

$$y[k] = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{x}[k] = \mathrm{AF} \times s[k] + \boldsymbol{w}^{\mathrm{H}}\boldsymbol{b}[k],$$

where $\mathrm{AF} = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{a}(p_d, q_d)$ is the array factor. Note that it can be rewritten as

$$\mathrm{AF} = \left[\boldsymbol{w}_v^{\mathrm{H}}\boldsymbol{a}_v(q_d)\right] \times \left[\boldsymbol{w}_h^{\mathrm{H}}\boldsymbol{a}_h(p_d)\right]. \tag{7}$$

Equation (7) shows that the total array factor is given by the product of the sub-array factors. Note that this property does not depend on the beam pattern of the antenna elements, since it only relies on the factorisation of the array factor. The steering vectors of some array geometries, such as circular arrays, for example, do not permit a Kronecker factorisation. In this case, we cannot directly apply the methods proposed in this work.

We present two novel beamforming strategies based on the MMSE filter that exploits array separability to reduce computational costs. In the first strategy, we recast the MSE function (6) using tensor algebra, and then we devise an iterative beamformer based on alternating minimisation. The reader is referred to [22, 23] for an introduction to tensor algebra. In the second strategy, we obtain a closed-form beamforming filter by employing sub-array MMSE filtering. In the end, we discuss the computational complexity of the proposed methods.

3.1 Tensor MMSE beamformer

Let us first reformulate the received signal model (1) using tensor algebra. Considering array separability (4), the received signal at the nth antenna can be rewritten as

$$x_{n_h,n_v}[k] = \sum_{r=1}^{R} = a_{n_v}^{(v)}(q_r)a_{n_h}^{(h)}(p_r)s_r[k] + b_{n_h,n_v}[k], \tag{8}$$

Now, define the received signals matrix $[\boldsymbol{X}[k]]_{n_h,n_v} = x_{n_h,n_v}[k]$, the array manifold tensor $[\mathscr{A}]_{n_h,n_v,r} = a_{n_h}^{(h)}(p_r)a_{n_v}^{(v)}(q_r)$, and the AWGN matrix $[\boldsymbol{B}[k]]_{n_h,n_v} = b_{n_h,n_v}[k]$. Using tensor modal products [22], the received signals matrix can be expressed as

$$\boldsymbol{X}[k] = \mathscr{A} \times_3 \boldsymbol{s}[k]^{\mathrm{T}} + \boldsymbol{B}[k] \in \mathbb{C}^{N_h \times N_v}. \tag{9}$$

The array manifold tensor $\mathscr{A}$ is a three-dimensional array with dimensions $N_h \times N_v \times R$. The two first array modes refer to the physical array dimensions, whereas the third one represents the transmitted signal dimension, i.e. the number of wavefronts. This tensor can be unfolded into matrices in three different manners [22]:

$$[\mathscr{A}]_{(1)} = \left[\boldsymbol{a}_h(p_1)\boldsymbol{a}_v(q_1)^{\mathrm{T}}, \ldots, \boldsymbol{a}_h(p_R)\boldsymbol{a}_v(q_R)^{\mathrm{T}}\right] \in \mathbb{C}^{N_h \times N_vR},$$

$$[\mathscr{A}]_{(2)} = \left[\boldsymbol{a}_v(q_1)\boldsymbol{a}_h(p_1)^{\mathrm{T}}, \ldots, \boldsymbol{a}_v(q_R)\boldsymbol{a}_h(p_R)^{\mathrm{T}}\right] \in \mathbb{C}^{N_v \times N_hR},$$

$$[\mathscr{A}]_{(3)} = (\boldsymbol{A}_v \diamond \boldsymbol{A}_h)^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} \in \mathbb{C}^{R \times N_vN_h}.$$

Let us now rewrite the beamformer output $y[k] = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{x}[k]$ in terms of $\boldsymbol{w}_h$ and $\boldsymbol{w}_v$ by considering the Kronecker factorisation of $\boldsymbol{w}$ and the bi-dimensional representation of the received signals (8):

$$\begin{aligned} y[k] &= \sum_{n=1}^{N} [\boldsymbol{w}]_n^* x_n[k] \\ &= \sum_{n_h=1}^{N_h}\sum_{n_v=1}^{N_v} [\boldsymbol{w}_h]_{n_h}^*[\boldsymbol{w}_v]_{n_v}^* x_{n_h,n_v}[k]. \end{aligned} \tag{10}$$

Using matrix notation, (10) can be rewritten as

$$y[k] = \boldsymbol{w}_h^{\mathrm{H}}\boldsymbol{X}[k]\boldsymbol{w}_v^* = \boldsymbol{w}_v^{\mathrm{H}}\boldsymbol{X}[k]^{\mathrm{T}}\boldsymbol{w}_h^*.$$

The MSE function (6) can now be reformulated as the following bi-linear function

$$\begin{aligned} &J_{\mathrm{MSE}}(\boldsymbol{w}_h, \boldsymbol{w}_v) \\ &\quad = \mathbb{E}\left[\left|s_d[k] - \boldsymbol{w}_h^{\mathrm{H}}\boldsymbol{X}[k]\boldsymbol{w}_v^*\right|^2\right] = \mathbb{E}\left[\left|s_d[k] - \boldsymbol{w}_v^{\mathrm{H}}\boldsymbol{X}[k]^{\mathrm{T}}\boldsymbol{w}_h^*\right|^2\right]. \end{aligned} \tag{11}$$

Unfortunately, minimising (11) is not straightforward. The gradient of $J_{\mathrm{MSE}}(\boldsymbol{w}_h, \boldsymbol{w}_v)$ with respect to any of its vector variables depends on the other variable. This coupling disables the direct application of methods such as gradient descent, calling for alternating minimisation techniques. To this end, let us define the horizontal and vertical sub-array input signals

$$\boldsymbol{u}_h[k] = \boldsymbol{X}[k]\boldsymbol{w}_v^* \in \mathbb{C}^{N_h}, \tag{12}$$

$$\boldsymbol{u}_v[k] = \boldsymbol{X}[k]^{\mathrm{T}}\boldsymbol{w}_h^* \in \mathbb{C}^{N_v}. \tag{13}$$

and rewrite (11) as

$$J_{\mathrm{MSE}}(\boldsymbol{w}_h, \boldsymbol{w}_v) = \mathbb{E}\left[\left|s_d[k] - \boldsymbol{w}_h^{\mathrm{H}}\boldsymbol{u}_h[k]\right|^2\right] \tag{14}$$

$$= \mathbb{E}\left[\left|s_d[k] - \boldsymbol{w}_v^{\mathrm{H}}\boldsymbol{u}_v[k]\right|^2\right]. \tag{15}$$

It is easy to recognise (14) and (15) as linear functions of $\boldsymbol{w}_h$ and $\boldsymbol{w}_v$, respectively, when the other vector variable is fixed. The proposed beamforming method, referred to as Tensor MMSE (TMMSE), consists of sequentially minimising (14) and (15) using the MMSE solution for each sub-filter until a convergence criterion is satisfied. The sub-beamformers are calculated according to the following theorem:

Theorem 1: The minimisers of (14) and (15) conditioned on $\boldsymbol{w}_v$ *and* $\boldsymbol{w}_h$ *are, respectively, given by*

$$\boldsymbol{w}_h = \boldsymbol{R}_{hh}^{-1}\boldsymbol{p}_{hs},$$

$$\boldsymbol{w}_v = \boldsymbol{R}_{vv}^{-1}\boldsymbol{p}_{vs},$$

where

$$\begin{aligned}\boldsymbol{R}_{hh} &= \mathbb{E}\left[\boldsymbol{u}_h[k]\boldsymbol{u}_h[k]^{\mathrm{H}}\right] \\ &= [\mathscr{A}]_{(1)}\left(\boldsymbol{R}_{ss} \otimes \boldsymbol{w}_v^*\boldsymbol{w}_v^{\mathrm{T}}\right)[\mathscr{A}]_{(1)}^{\mathrm{H}} + \sigma_b^2 \parallel \boldsymbol{w}_v \parallel_2^2 \boldsymbol{I}_{N_h} \in \mathbb{C}^{N_h \times N_h},\end{aligned}$$

$$\begin{aligned}\boldsymbol{R}_{vv} &= \mathbb{E}\left[\boldsymbol{u}_v[k]\boldsymbol{u}_v[k]^{\mathrm{H}}\right] \\ &= [\mathscr{A}]_{(2)}\left(\boldsymbol{R}_{ss} \otimes \boldsymbol{w}_h^*\boldsymbol{w}_h^{\mathrm{T}}\right)[\mathscr{A}]_{(2)}^{\mathrm{H}} + \sigma_b^2 \parallel \boldsymbol{w}_h \parallel_2^2 \boldsymbol{I}_{N_v} \in \mathbb{C}^{N_v \times N_v}\end{aligned}$$

denote the covariance matrices of the sub-array input signals, and

$$\boldsymbol{p}_{hs} = \mathbb{E}[\boldsymbol{u}_h[k]s_d^*[k]] = [\mathscr{A}]_{(1)}(\boldsymbol{R}_{ss}\boldsymbol{e}_d \otimes \boldsymbol{w}_v^*) \in \mathbb{C}^{N_h},$$

$$\boldsymbol{p}_{vs} = \mathbb{E}[\boldsymbol{u}_v[k]s_d^*[k]] = [\mathscr{A}]_{(2)}(\boldsymbol{R}_{ss}\boldsymbol{e}_d \otimes \boldsymbol{w}_h^*) \in \mathbb{C}^{N_v}$$

the cross-covariance vectors between the sub-array input signals and the signal of interest.

Proof: See the Appendix.

Theorem 1: It can be applied when the signals' statistics ($\boldsymbol{R}_{ss}$ and $\boldsymbol{R}_{bb}$), and the array manifold matrix are known. However, such information might not be available in practice, and thus the sub-array covariance matrices and cross-covariance vectors need to be estimated. It can be done by using sample estimates over K time snapshots. In this sense, the covariance matrices $\boldsymbol{R}_{hh}$ and $\boldsymbol{R}_{vv}$ can be estimated as

$$\hat{\boldsymbol{R}}_{hh} = \frac{1}{K}\sum_{k=0}^{K-1}\boldsymbol{u}_h[k]\boldsymbol{u}_h[k]^{\mathrm{H}},$$

$$\hat{\boldsymbol{R}}_{vv} = \frac{1}{K}\sum_{k=0}^{K-1}\boldsymbol{u}_v[k]\boldsymbol{u}_v[k]^{\mathrm{H}},$$

and the cross-covariance vectors $\boldsymbol{p}_{hs}$ and $\boldsymbol{p}_{vs}$ as

$$\hat{\boldsymbol{p}}_{hs} = \frac{1}{K}\sum_{k=0}^{K-1}\boldsymbol{u}_h[k]s_d^*[k],$$

$$\hat{\boldsymbol{p}}_{vs} = \frac{1}{K}\sum_{k=0}^{K-1}\boldsymbol{u}_v[k]s_d^*[k].$$

Note that $\boldsymbol{u}_h[k]$ and $\boldsymbol{u}_v[k]$ can be easily formed by observing $\boldsymbol{x}[k]$, reshaping into $\boldsymbol{X}[k]$, and using (12) and (13), respectively. The steps to compute the TMMSE beamformer are summarised in Algorithm 1.

Algorithm 1: Tensor MMSE algorithm

1: Randomly initialise $\boldsymbol{w}_h$ and $\boldsymbol{w}_v$
2: **repeat**
3: Form $\boldsymbol{R}_{hh}$ and $\boldsymbol{p}_{hs}$
4: $\boldsymbol{w}_h \leftarrow \boldsymbol{R}_{hh}^{-1}\boldsymbol{p}_{hs}$
5: Form $\boldsymbol{R}_{vv}$ and $\boldsymbol{p}_{vs}$
6: $\boldsymbol{w}_v \leftarrow \boldsymbol{R}_{vv}^{-1}\boldsymbol{p}_{vs}$
7 **until** convergence criterion triggers
8: $\boldsymbol{w} \leftarrow \boldsymbol{w}_v \otimes \boldsymbol{w}_h$

3.2 Kronecker MMSE beamformer

Let us consider the following Khatri-Rao product property. Let $\boldsymbol{A} \in \mathbb{C}^{P\times M}$, $\boldsymbol{B} \in \mathbb{C}^{Q\times N}$, $\boldsymbol{C} \in \mathbb{C}^{M\times R}$, and $\boldsymbol{D} \in \mathbb{C}^{N\times R}$. From [24], it follows that

$$(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \diamond \boldsymbol{D}) = (\boldsymbol{A}\boldsymbol{C}) \diamond (\boldsymbol{B}\boldsymbol{D}) \in \mathbb{C}^{PQ\times R}. \tag{16}$$

This result suggests that a Kronecker separable beamformer can be individually applied to the corresponding sub-array manifold matrix in (5). In this case, the filtering operation $y[k] = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{x}[k]$ can be carried out as

$$\begin{aligned}y[k] &= (\boldsymbol{w}_v \otimes \boldsymbol{w}_h)^{\mathrm{H}}(\boldsymbol{A}_v \diamond \boldsymbol{A}_h)\boldsymbol{s}[k] + \boldsymbol{w}^{\mathrm{H}}\boldsymbol{b}[k] \\ &= \left[(\boldsymbol{w}_v^{\mathrm{H}}\boldsymbol{A}_v) \diamond (\boldsymbol{w}_h^{\mathrm{H}}\boldsymbol{A}_h)\right]\boldsymbol{s}[k] + \boldsymbol{w}^{\mathrm{H}}\boldsymbol{b}[k].\end{aligned}$$

Therefore, instead of optimising an N-dimensional beamformer for $\boldsymbol{A}$, we can design two independent low-dimensional beamformers for $\boldsymbol{A}_h$ and $\boldsymbol{A}_v$ individually. According to this approach, each sub-beamformer is fed only with signals from the corresponding antenna sub-array. In this sense, we define the horizontal and vertical observed signals:

$$\boldsymbol{x}_h[k] = \boldsymbol{A}_h\boldsymbol{s}[k] + \boldsymbol{b}_h[k] \in \mathbb{C}^{N_h}$$

$$\boldsymbol{x}_v[k] = \boldsymbol{A}_v\boldsymbol{s}[k] + \boldsymbol{b}_v[k] \in \mathbb{C}^{N_v},$$

where $\boldsymbol{b}_h[k] \in \mathbb{C}^{N_h}$ and $\boldsymbol{b}_v[k] \in \mathbb{C}^{N_v}$ represent the additive Gaussian noise vector observed at the horizontal and vertical sub-arrays, respectively. These vectors are defined as

$$[\boldsymbol{b}_h[k]]_{n_h} = b_{n_h + (n_v - 1)N_h}[k]\big|_{n_v = 1},$$

$$[\boldsymbol{b}_v[k]]_{n_v} = b_{n_h + (n_v - 1)N_h}[k]\big|_{n_h = 1}.$$

We propose to optimise each sub-beamformer according to the MMSE criterion. However, the direct application of the MMSE filter to each sub-beamformer would be prone to numerical problems. Often in many practical scenarios, e.g. mobile communications, different signals are closely separated in an angular domain (azimuth or elevation). In this case, either the vertical or horizontal sub-array manifold matrices become almost rank deficient, turning the MSE minimisation problem ill-posed. To overcome this issue, we resort to Tikhonov regularisation [25], which avoids singular covariance matrices by penalising large-norm solutions. The proposed beamforming method, hereafter referred to as Kronecker MMSE (KMMSE), independently minimises the following cost functions

$$J_{\mathrm{MSE}}^{(h)}(\boldsymbol{w}_h, \rho) = \mathbb{E}\left[\left|s_d[k] - \boldsymbol{w}_h^{\mathrm{H}}\boldsymbol{x}_h[k]\right|^2\right] + \rho \parallel \boldsymbol{w}_h \parallel_2^2, \tag{17}$$

$$J_{\mathrm{MSE}}^{(v)}(\boldsymbol{w}_v, \rho) = \mathbb{E}\left[\left|s_d[k] - \boldsymbol{w}_v^{\mathrm{H}}\boldsymbol{x}_v[k]\right|^2\right] + \rho \parallel \boldsymbol{w}_v \parallel_2^2, \tag{18}$$

where $\rho \geq 0$ denotes the regularisation parameter. Define

$$\boldsymbol{R}_m = \boldsymbol{A}_m \boldsymbol{R}_{ss} \boldsymbol{A}_m^{\mathrm{H}} + \boldsymbol{R}_{bb,m}, \tag{19}$$

$$\boldsymbol{p}_m = \boldsymbol{A}_m \boldsymbol{R}_{ss} \boldsymbol{e}_d, \tag{20}$$

with $\boldsymbol{R}_{bb,m} = \sigma_b^2 \boldsymbol{I}_{N_m}$ for $m \in \{h, v\}$. The minimisers for (17) and (18) are thus given by $\boldsymbol{w}_m = (\boldsymbol{R}_m + \rho \boldsymbol{I}_{N_m})^{-1} \boldsymbol{p}_m$ for $m \in \{h, v\}$. Due to regularisation, the KMMSE output signal is not guaranteed to have the same power as the desired signal. Thus, we employ the following scaling to correct the KMMSE output power: $y_{\mathrm{KMMSE}}[k] = (\sigma_s/\sigma_p) p[k]$, where $p[k] = (\boldsymbol{w}_v \otimes \boldsymbol{w}_h)^{\mathrm{H}} \boldsymbol{x}[k]$ and σ_p denotes the standard deviation of $p[k]$. In a practical implementation, this scaling correction can be performed by the automatic gain control circuit. The computation of the KMMSE filter is summarised in Algorithm 2.

In practice, one might not have a priori knowledge of the sub-array manifold matrices ($\boldsymbol{A}_h$ and $\boldsymbol{A}_v$) and signals' statistics. One can estimate (19) and (20) using the received signals from the horizontal and vertical sub-arrays, represented by

$$\bar{\boldsymbol{x}}_m[k] = \boldsymbol{A}_m \boldsymbol{s}[k] + \boldsymbol{b}_m[k], m \in \{h, v\}\,.$$

For the horizontal sub-array, we define

$$[\bar{\boldsymbol{x}}_h[k]]_{n_h} = x_{n_h + (n_v - 1)N_h}[k]\big|_{n_v = 1} = x_{n_h}[k],$$

with $n_h \in \{1, \ldots, N_h\}$ and $r \in \{1, \ldots, R\}$. Similarly, for the vertical sub-array:

$$[\bar{\boldsymbol{x}}_v[k]]_{n_v} = x_{n_h + (n_v - 1)N_h}[k]\big|_{n_h = 1} = x_{1 + (n_v - 1)N_h}[k],$$

with $n_v \in \{1, \ldots, N_v\}$ and $r \in \{1, \ldots, R\}$. Now, the covariance matrices can be estimated as

$$\hat{\boldsymbol{R}}_h = \left(\frac{1}{K} \sum_{k=0}^{K-1} \bar{\boldsymbol{x}}_h[k] \bar{\boldsymbol{x}}_h[k]^{\mathrm{H}} \right),$$

$$\hat{\boldsymbol{R}}_v = \left(\frac{1}{K} \sum_{k=0}^{K-1} \bar{\boldsymbol{x}}_v[k] \bar{\boldsymbol{x}}_v[k]^{\mathrm{H}} \right),$$

and the cross-covariance vectors as

$$\hat{\boldsymbol{p}}_h = \left(\frac{1}{K} \sum_{k=0}^{K-1} \bar{\boldsymbol{x}}_h[k] s_d^*[k] \right),$$

$$\hat{\boldsymbol{p}}_v = \left(\frac{1}{K} \sum_{k=0}^{K-1} \bar{\boldsymbol{x}}_v[k] s_d^*[k] \right).$$

The proposed closed-form KMMSE beamformer can be seen as a sub-optimal solution which relies on a covariance matrix approximation. According to the mixed product property of the Kronecker product [24], the KMMSE beamformer can be expressed as

$$\boldsymbol{w} = \left[(\boldsymbol{R}_v + \rho \boldsymbol{I}_{N_v}) \otimes (\boldsymbol{R}_h + \rho \boldsymbol{I}_{N_h}) \right]^{-1} (\boldsymbol{p}_v \otimes \boldsymbol{p}_h)\,. \tag{21}$$

The Kronecker product of covariance matrices in (21) can be regarded as an approximation of $\boldsymbol{R}_{xx}$. Also, it is straightforward to see in (21) that the cross-covariance vector $\boldsymbol{p}_{xs}$ can be exactly factorised into $\boldsymbol{p}_v \otimes \boldsymbol{p}_h$. We now conduct an asymptotic analysis of KMMSE to provide insights on its performance.

First, consider the classical MMSE filter

$$\boldsymbol{w}_{opt} = \boldsymbol{R}_{xx}^{-1} \boldsymbol{p}_{xs} = \left(\boldsymbol{A} \boldsymbol{R}_{ss} \boldsymbol{A}^{\mathrm{H}} + \boldsymbol{R}_{bb} \right)^{-1} \boldsymbol{A} \boldsymbol{R}_{ss} \boldsymbol{e}_d\,.$$

Applying the matrix inversion lemma [20], its Hermitian vector can be written as

$$\boldsymbol{w}_{opt}^{\mathrm{H}} = \boldsymbol{e}_d^{\mathrm{T}} \left(\boldsymbol{R}_{ss}^{-1} + \boldsymbol{A}^{\mathrm{H}} \boldsymbol{R}_{bb}^{-1} \boldsymbol{A} \right)^{-1} \boldsymbol{A}^{\mathrm{H}} \boldsymbol{R}_{bb}^{-1}\,.$$

From the signal statistics assumptions in Section 2, we have

$$\boldsymbol{w}_{opt}^{\mathrm{H}} = \boldsymbol{e}_d^{\mathrm{T}} \left(\frac{\sigma_b^2}{\sigma_s^2} \boldsymbol{I}_R + \boldsymbol{A}^{\mathrm{H}} \boldsymbol{A} \right)^{-1} \boldsymbol{A}^{\mathrm{H}}. \tag{22}$$

Now, we rewrite the Kronecker factors of the KMMSE filter using (22) and for $\rho = 0$ to obtain

$$\begin{aligned} \boldsymbol{w}^{\mathrm{H}} = & \left[\boldsymbol{e}_d^{\mathrm{T}} \left(\frac{\sigma_b^2}{\sigma_s^2} \boldsymbol{I}_R + \boldsymbol{A}_v^{\mathrm{H}} \boldsymbol{A}_v \right)^{-1} \boldsymbol{A}_v^{\mathrm{H}} \right] \\ & \otimes \left[\boldsymbol{e}_d^{\mathrm{T}} \left(\frac{\sigma_b^2}{\sigma_s^2} \boldsymbol{I}_R + \boldsymbol{A}_h^{\mathrm{H}} \boldsymbol{A}_h \right)^{-1} \boldsymbol{A}_h^{\mathrm{H}} \right]. \end{aligned} \tag{23}$$

At high SNR, the noise power drops and $\sigma_b^2 \to 0$. If the inverse matrix $(\boldsymbol{A}_m^{\mathrm{H}} \boldsymbol{A}_m)^{-1}$ exists for $m \in \{h, v\}$, then

$$\boldsymbol{w}^{\mathrm{H}} \to \left(\boldsymbol{e}_d^{\mathrm{T}} \boldsymbol{A}_v^{\dagger} \right) \otimes \left(\boldsymbol{e}_d^{\mathrm{T}} \boldsymbol{A}_h^{\dagger} \right). \tag{24}$$

As expected, each sub-array beamformer converges to a ZF filter. Using (16) and (24), we see that the KMMSE output signal at high SNR converges to

$$y[k] \to \left[\left(\boldsymbol{e}_d^{\mathrm{T}} \boldsymbol{A}_v^{\dagger} \boldsymbol{A}_v \right) \diamond \left(\boldsymbol{e}_d^{\mathrm{T}} \boldsymbol{A}_h^{\dagger} \boldsymbol{A}_h \right) \right] \boldsymbol{s}[k] = s_d[k]\,.$$

The inverse $(\boldsymbol{A}_m^{\mathrm{H}} \boldsymbol{A}_m)^{-1}$ exists if and only if $\boldsymbol{A}_m^{\mathrm{H}} \boldsymbol{A}_m$ is not rank deficient, i.e. the wavefronts arrive from different directions. However, when the wavefronts are closely spaced in the angular domain, $\boldsymbol{A}_m^{\mathrm{H}} \boldsymbol{A}_m$ becomes ill-conditioned and the ZF filter performs poorly. Fortunately, when $\rho > 0$, the inverse matrix is defined, allowing for desired signal recovery.

At low SNR, the term $(\sigma_b^2/\sigma_s^2) \boldsymbol{I}_R$ dominates and (23) goes to

$$\begin{aligned} \boldsymbol{w}^{\mathrm{H}} &= \left(\boldsymbol{e}_d^{\mathrm{T}} \frac{\sigma_s^2}{\sigma_b^2} \mathbf{A}_v^{\mathrm{H}} \right) \otimes \left(\boldsymbol{e}_d^{\mathrm{T}} \frac{\sigma_s^2}{\sigma_b^2} \boldsymbol{A}_h^{\mathrm{H}} \right) \\ &= \frac{\sigma_s^4}{\sigma_b^4} \boldsymbol{a}_v(\phi_d, \theta_d)^{\mathrm{H}} \otimes \boldsymbol{a}_h(\phi_d, \theta_d)^{\mathrm{H}}\,. \end{aligned} \tag{25}$$

Equation (25) shows that, as in the classical MMSE filter, the factors of $\boldsymbol{w}$ converge to matched filters which maximise the desired signal power. In this case, the KMMSE output signal can be written as

$$y[k] \to (\sigma_s^4/\sigma_b^4)[(\boldsymbol{a}_v(\phi_d, \theta_d)^{\mathrm{H}} \boldsymbol{A}_v) \diamond (\boldsymbol{a}_h(\phi_d, \theta_d)^{\mathrm{H}} \boldsymbol{A}_h)] \boldsymbol{s}[k] + \boldsymbol{w}^{\mathrm{H}} \boldsymbol{b}[k\,]\,.$$

If the incoming wavefronts are sufficiently separated in the angular domain, i.e. if all $\boldsymbol{a}_v$ and all $\boldsymbol{a}_h$ are mutually orthogonal, then $y[k] \to (\sigma_s^4/\sigma_b^4) s_d[k] + \boldsymbol{w}^{\mathrm{H}} \boldsymbol{b}[k]$. The analysis above shows that the proposed KMMSE filter is able to recover the desired signal from the received signals despite the covariance matrix approximations. Note that it is based on the assumption that the incoming signals are sufficiently separated in the angular domain.

Algorithm 2: Kronecker MMSE filter

1: Select $\rho \geq 0$
2: Form $\boldsymbol{R}_h$ and $\boldsymbol{p}_h$
3: $\boldsymbol{w}_h \leftarrow \left(\boldsymbol{R}_h + \rho \boldsymbol{I}_{N_h} \right)^{-1} \boldsymbol{p}_h$
4: Form $\boldsymbol{R}_v$ and $\boldsymbol{p}_v$

5: $\boldsymbol{w}_v \leftarrow \left(\boldsymbol{R}_v + \rho \boldsymbol{I}_{N_v}\right)^{-1} \boldsymbol{p}_v$

6: $\boldsymbol{w} \leftarrow \boldsymbol{w}_v \otimes \boldsymbol{w}_h$

The proposed beamforming methods work with low-dimensional sub-array manifold matrices to decrease their computational complexity. However, this also reduces their degrees of freedom, which are important for attenuating interfering signals. An MMSE filter designed for N antennas has N degrees of freedom, i.e. it is capable of recovering the desired signal and rejecting $N-1$ undesired sources. Our separable beamforming framework, by contrast, offers N_h and N_v degrees of freedom for the horizontal and vertical sub-arrays. Therefore, the separable filter performance is limited by the least degree of freedom. Hence, the proposed methods are capable of recovering the desired signal and rejecting $\min(N_h, N_v) - 1$ undesired sources. The proposed separable beamforming framework exchanges degrees of freedom for computational complexity reduction.

3.3 Computational complexity

The MMSE filter is known to be computationally complex. However, if the array manifold matrix $\boldsymbol{A}$ and the signal statistics $\boldsymbol{R}_{ss}$ and $\boldsymbol{R}_{bb}$ are known, one can employ the matrix inversion lemma to the MMSE filter and obtain the low-complexity MMSE expression (22), in which a $R \times R$ matrix is inverted. However, this information may not be available, and then sample estimates are needed to compute the MMSE filter. In this case, the inversion lemma cannot be applied, and, thus an $N \times N$ covariance matrix is inverted in order to get the MMSE filter coefficients. Such an operation has complexity $O(N^3)$, which can be overwhelming for massive array systems. In this case, the proposed methods can be used since they are much less expensive in computational terms, as we show in the following.

The TMMSE filter calculates its beamformer coefficients through an iterative process of I iterations, in which N_h – and N_v – dimensional matrices are inverted. Therefore, the TMMSE filter requires $O(I(N_h^3 + N_v^3))$ operations. Therefore, this method is less complex than the classical approach provided that I, N_h, and N_v are not too large. The authors in [26] discussed the convergence of alternating MMSE-based methods and concluded that they are monotonically convergent. Other numerical properties such as convergence rate and stability are not discussed and, to the best of our knowledge, the investigation of these aspects remains a research challenge. The analytical convergence study of the proposed method is beyond the scope of this work.

The KMMSE filter is much simpler than the previous methods since it performs sub-array beamforming using closed-form solutions. To obtain the beamformer coefficients, one needs to invert N_h – and N_v – dimensional matrices only once. Thus, this method carries out $O(N_h^3 + N_v^3)$ operations.

4 Simulation results

We present numerical results from simulations conducted to assess the performance of the proposed methods. At each simulation, signal data are generated as follows: R independent sequences of K QPSK-modulated symbols are generated to form $\boldsymbol{s}[k]$, for all $k \in \{1, \ldots, K\}$. Next, the direction cosines (p_r and q_r) of the R wavefronts are randomly generated according to a uniform distribution in the range $[-0.9, 0.9]$ so the array manifold matrix $\boldsymbol{A}$ is formed. Note that selecting the direction cosines within this range ensures the omnidirectional propagation assumption of Section 2. Finally, the observed signals (2) are formed by contaminating the received symbols with additive noise.

We investigate the computational complexity and source recovery performance of the proposed beamforming methods in terms of floating point operations (flops) and uncoded bit error ratio (BER) of the desired signal. We choose BER as Fig. 2 of merit because it reveals the noise and interference rejection performance. Therefore, if the beamforming operation is correctly carried out, then the interfering wavefronts are attenuated, and the

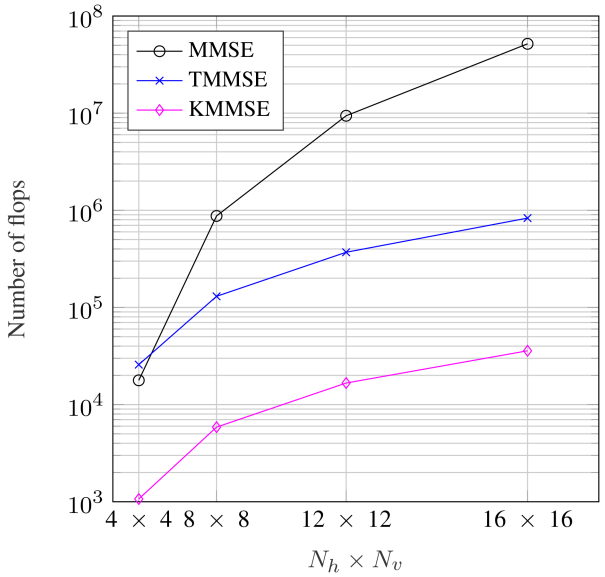

Fig. 2 *Number of flops as function of array size for $R = 4$ wavefronts*

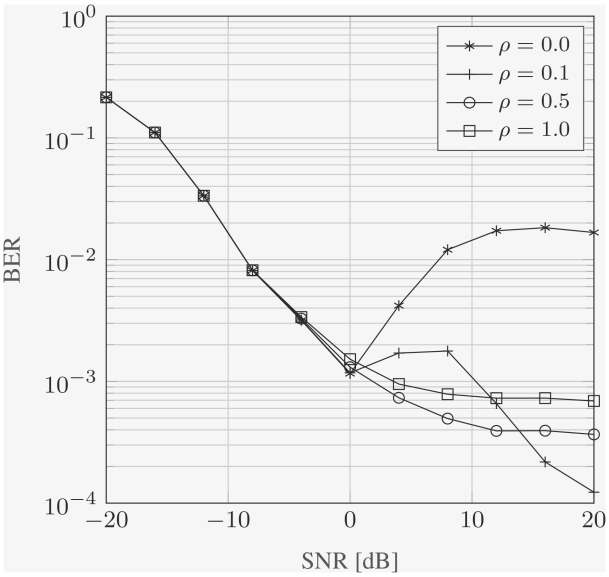

Fig. 3 *KMMSE BER performance for different regularisation parameter ρ. $N_h \times N_v = 8 \times 8$, $R = 4$ wavefronts*

desired signal BER decreases. The graphs in Figs. 3–5 were obtained by averaging the results from 10^5 independent experiments considering $R = 4$ wavefronts, $N_h \times N_v = 8 \times 8$ antennas, and $K = 1000$ symbols. Figs. 6–10, however, were collected from a single experiment with $R = 6$ wavefronts, $N_h \times N_v = 4 \times 4$ antennas, and $K = 1000$ symbols. The parameter selection for the latter figures will be motivated in the following paragraphs. The convergence of the TMMSE method is achieved when the normalised filter residual between consecutive iterations is smaller than a tolerance value $\epsilon > 0$, i.e. $\| (\boldsymbol{w}_i / \| \boldsymbol{w}_i \|_2) - (\boldsymbol{w}_{i-1} / \| \boldsymbol{w}_{i-1} \|_2) \|_2 < \epsilon$, where i denotes the iteration number. In all experiments, we set $\epsilon = 10^{-3}$. Preliminary simulations have shown that the average number of TMMSE iterations is 5.

The computational complexity, measured in flops, is plotted as a function of the array size in Fig. 2. We use the MATLAB Lightspeed toolbox [27] for flops counting since it provides the approximate number of operations required for inverting matrices. This result shows that the proposed method substantially reduces the computational complexity of beamforming design. For an array of 16×16 antennas, the complexity difference between MMSE and KMMSE is around three orders of magnitude. While KMMSE is inexpensive in all scenarios, TMMSE is costly for relatively

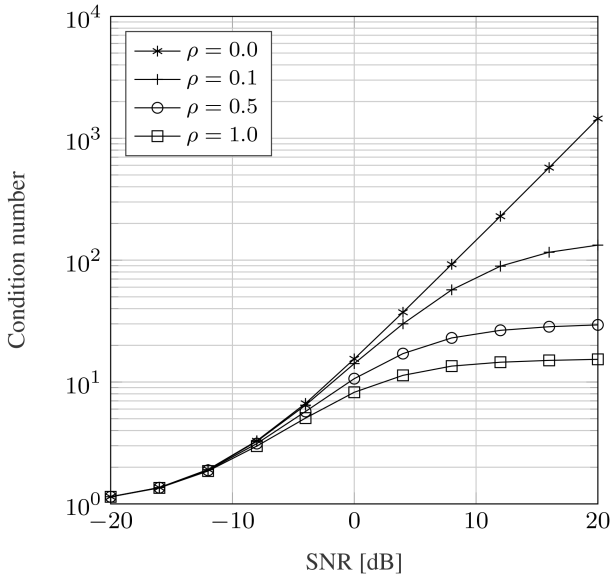

Fig. 4 *Condition number of (19) for different regularisation parameter ρ. $N_h \times N_v = 8 \times 8$, $R = 4$ wavefronts*

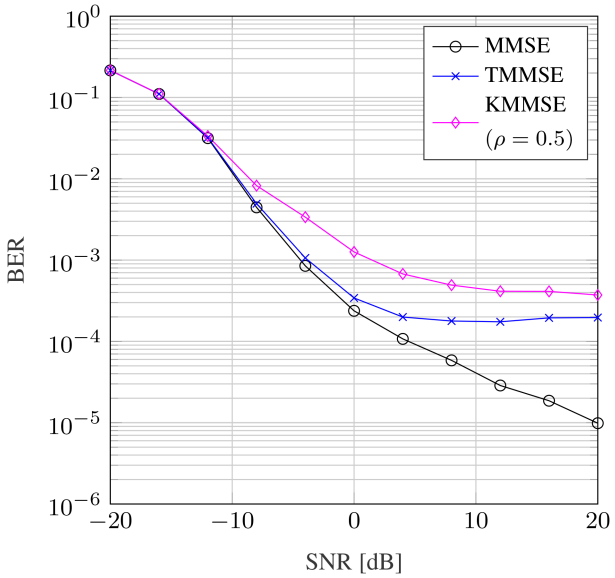

Fig. 5 *$N_h \times N_v = 8 \times 8$, $R = 4$ wavefronts*

small arrays due to the iterative optimisation procedure. For arrays of 8×8 antennas, a set-up expected for 5G systems [3], both separable beamformers are less expensive than MMSE.

We investigate the influence of the regularisation parameter ρ on the KMMSE BER performance in Fig. 3. We observe that regularisation plays a little role in the performance for low SNR (<0 dB). In this case, the noise term on (19) has enough energy to complete the rank of the covariance matrix, decreasing its condition number [28], thus making the horizontal and vertical MMSE beamformers numerically stable. However, for high SNR (≥ 0 dB), regularisation is paramount to achieve a satisfactory performance. This is because the noise term is not strong anymore to fill the covariance matrix rank in this case, and then the horizontal and vertical covariance matrices in (19) become ill-conditioned. To tackle this issue, the regularisation term fills the covariance matrix rank, decreasing the condition number, and thus making the regularised MMSE beamformers stable at high SNR. The link between the covariance matrix condition number and the regularisation parameter ρ is clarified in Fig. 4. We note that the covariance matrix becomes more well-conditioned at high SNR as we increase ρ. More specifically, for $\rho = 0$, the BER worsens with the SNR, owing to the increase in the condition number of the horizontal and/or vertical covariance matrices. For $\rho = 0.1$, KMMSE exhibits a bad performance from 0 dB, to 16 dB SNR. In

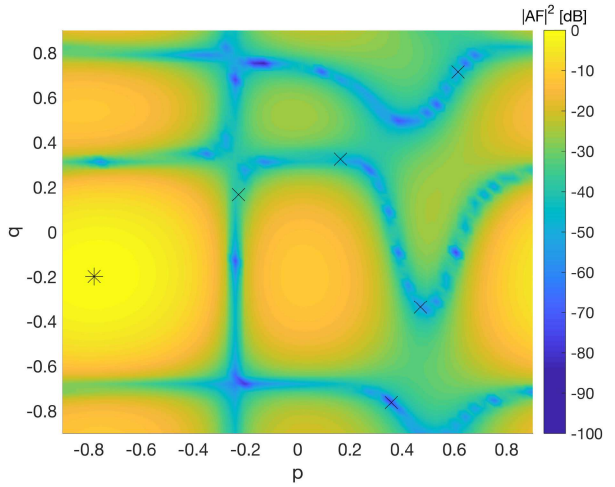

Fig. 6 *MMSE AF squared magnitude. $R = 6$ wavefronts, $N_h \times N_v = 4 \times 4$. Asterisk denotes desired signal, cross interfering signal*

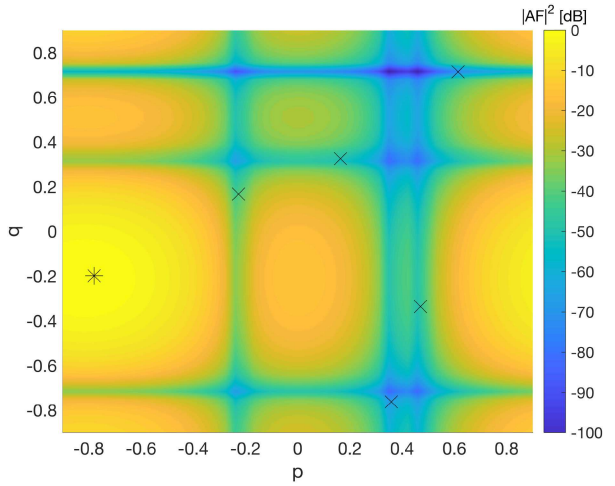

Fig. 7 *TMMSE AF squared magnitude. $R = 6$ wavefronts, $N_h \times N_v = 4 \times 4$. Asterisk denotes desired signal, cross interfering signal*

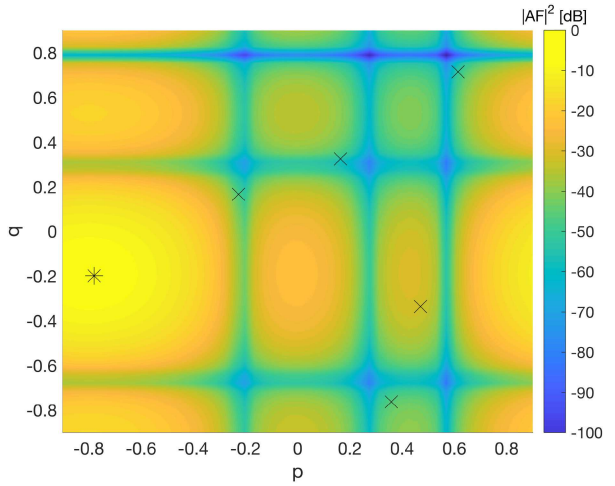

Fig. 8 *KMMSE AF squared magnitude. $R = 6$ wavefronts, $N_h \times N_v = 4 \times 4$, $\rho = 0.5$. Asterisk denotes desired signal, cross interfering signal*

this range, the regularisation factor is not large enough to avoid performance deterioration due to ill-condition of the covariance matrix. Finally, for $\rho > 0.1$, we observe that KMMSE is not much affected by the covariance matrix condition. However, the regularisation term yields a BER bias, which increases with ρ. Fig. 3 shows that $\rho = 0.5$ provides the best stability-performance trade-off, thus we choose this value for the next simulations.

Unfortunately, the important computational complexity reduction observed in Fig. 2 comes with source recovery degradation, as one can see in Fig. 5. This figure shows that TMMSE exhibits good performance from −20 dB to 0 dB, while KMMSE with regularisation parameter $\rho = 0.5$ performs similarly to the benchmark only from −20 dB to −8 dB. At high SNR, the BER performance of the separable beamformers is heavily penalised compared to the benchmark. When one of the direction cosines (p_r or q_r) of the interfering wavefronts is adjacent to those of the desired signal, the separable beamformers fail to recover it, yielding a significant number of bit errors. Since the direction cosines of all wavefronts are randomly selected according to a uniform distribution, it is rather common that the interfering wavefronts are close to the desired signal in either p – or q – domain. As a consequence, beamforming fails due to ill-conditioned covariance matrices, as discussed in the previous paragraph. At high SNR, the poor BER performance is especially accentuated because the noise component does not have enough power to fill the covariance matrix rank. However, whenever the wavefronts are sufficiently separated in space, the separable beamformers exhibit good source recovery performance. The presented results suggest that the proposed methods are appealing alternatives to the standard MMSE beamformer. In Fig. 2, the computational complexity of KMMSE, for example, is two orders of magnitude smaller than that of MMSE for $N_h \times N_v = 8 \times 8$. On the other hand, Fig. 5 indicates that KMMSE is 5 dB apart from MMSE for the uncoded BER of 10^{-3}.

To better understand why the separable beamformers are more sensitive to closely spaced wavefronts, let us investigate their normalised array factor. We consider a scenario in which the proposed beamformers fail due to lack of degrees of freedom. For visualisation easiness, we consider a scenario where a 4×4 array is applied to filter $R = 6$ wavefronts in the following. In this case, each sub-array beamformer has only four degrees of freedom and will fail to filter the $R = 6$ signals. By contrast, the classical MMSE beamformer has 16 degrees of freedom and is able to null the interfering wavefronts and recover the desired signal. Figs. 6–8 show the magnitude of the MMSE, TMMSE, and KMMSE normalised array factors as functions of the direction cosines, respectively. One can see that the MMSE filter accurately places nulls at the interfering wavefronts directions, while a strong beam is pointed towards the desired signal. This is possible because the 16-dimensional filter has sufficient degrees of freedom to separate the wavefronts. In contrast to the benchmark method, KMMSE does not accurately distribute nulls, hindering interference attenuation. In Figs. 9 and 10, one can see that the KMMSE sub-beamformers do not have enough degrees of freedom to attenuate the interfering wavefronts. As a consequence, the undesired signal at $(p, q) = (0.5, -0.3)$ is not properly attenuated, as seen in Fig. 8. We observe that only three nulls are placed to attenuate five interfering wavefronts in Fig. 9. The same is observed in Fig. 10. To solve this issue, one would need to increase the number of antennas to, at least, $N_h \times N_v = 6 \times 6$. Fig. 7 reveals that TMMSE is more accurate than KMMSE at null placement. This is because the null locations are optimised as the alternating algorithm iterates. This accuracy is important especially at high SNR, as one can see in Fig. 5. We conclude that the separable beamformers are more sensitive to the number of impinging signals and closely spaced wavefronts than the classical MMSE beamformer due to the reduced degrees of freedom of the sub-beamformers.

5 Conclusion

We presented two beamforming methods that exploit array separability to reduce the computational complexity of the classical MMSE beamformer. The TMMSE filter is based on tensor algebra and minimises the MMSE by means of alternating minimisation, while the KMMSE filter relies on regularised sub-array MMSE beamforming. Our simulation results show that TMMSE provides moderate computational complexity reduction with small source recovery degradation. By contrast, KMMSE is computationally inexpensive but exhibits poorer BER performance at high SNR. Therefore, TMMSE should be employed when source recovery performance is more important than computation efficiency, and KMMSE on the contrary. This work paves the way to future contributions, including the extension to massive MIMO architectures, e.g. hybrid analogue/digital transceivers. Furthermore, it would be of interest to validate the proposed methods with array responses simulated in an electromagnetic field simulator software.

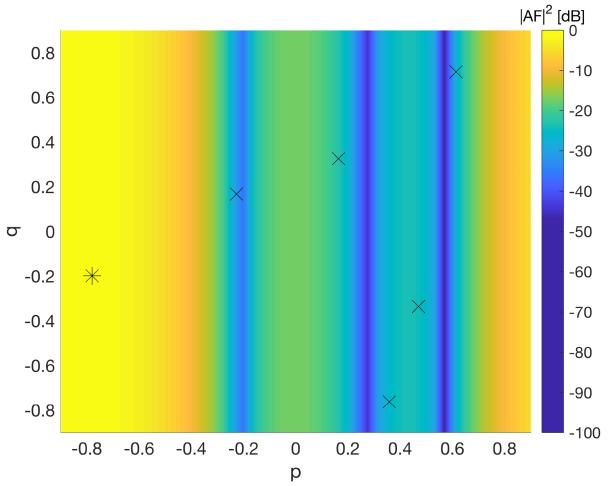

Fig. 9 *Horizontal KMMSE AF squared magnitude. $R = 6$ wavefronts, $N_h = 4$, $\rho = 0.5$. Asterisk denotes desired signal, cross interfering signal*

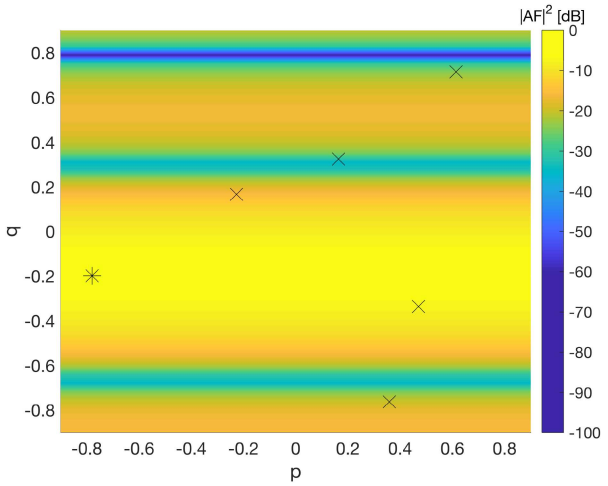

Fig. 10 *Vertical KMMSE AF squared magnitude. $R = 6$ wavefronts, $N_v = 4$, $\rho = 0.5$. Asterisk denotes desired signal, cross interfering signal*

6 Acknowledgment

This work is supported by the National Council for Scientific and Technological Development – CNPq, CAPES/PROBRAL Proc. no. 88887.144009/2017-00, and FUNCAP.

7 References

[1] Larsson, E.G., Edfors, O., Tufvesson, F., *et al.*: 'Massive MIMO for next generation wireless systems', *IEEE Commun. Mag.*, 2014, **52**, (2), pp. 186–195

[2] Schwarz, S., Rupp, M.: 'Society in motion: challenges for LTE and beyond mobile communications', *IEEE Commun. Mag.*, 2016, **54**, (5), pp. 76–83

[3] Ji, H., Kim, Y., Lee, J., *et al.*: 'Overview of full-dimension MIMO in LTE-advanced pro', *IEEE Commun. Mag.*, 2017, **55**, (2), pp. 176–184

[4] Ribeiro, L.N., Schwarz, S., Rupp, M., *et al.*: 'Energy efficiency of mmWave massive MIMO precoding with low-resolution DACs', *IEEE. J. Sel. Top. Signal. Process.*, 2018, **12**, (2), pp. 298–312

[5] Treitel, S., Shanks, J.L.: 'The design of multistage separable planar filters', *IEEE Trans. Geosci. Electron.*, 1971, **9**, (1), pp. 10–27

[6] Boussé, M., Debals, O., De Lathauwer, L.: 'A tensor-based method for large-scale blind source separation using segmentation', *IEEE Trans. Signal Process.*, 2017, **65**, (2), pp. 346–358

[7] Rupp, M., Schwarz, S.: 'A tensor LMS algorithm'. Proc. 2015 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, 2015, pp. 3347–3351

[8] Rupp, M., Schwarz, S.: 'Gradient-based approaches to learn tensor products'. Proc. 2015 23rd European Signal Processing Conf. (EUSIPCO), Nice, 2015, pp. 2486–2490

[9] Pinheiro, F.C., Lopes, C.G.: 'Nonlinear adaptive algorithms on rank-one tensor models'. arXiv preprint arXiv:1610.07520, 2016

[10] Ribeiro, L.N., de Almeida, A.L., Mota, J.C.M.: 'Identification of separable systems using trilinear filtering'. Proc. of the 2015 IEEE 6th Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2015, pp. 189–192

[11] Paleologu, C., Benesty, J., Ciochina, S.: 'Linear system identification based on a kronecker product decomposition', *IEEE/ACM Trans. Audio, Speech, Language Process.*, 2018, **26**, (10), pp. 1793–1808

[12] Elisei-Iliescu, C., Stanciu, C., Paleologu, C., *et al.*: 'Efficient recursive least-squares algorithms for the identification of bilinear forms', *Digit. Signal Process.*, 2018, **83**, pp. 280–296

[13] Van Trees, H.L.: '*Optimum array processing: part IV of detection, estimation and modulation theory*' (John Wiley & Sons, New Jersey, 2002), p. 1

[14] Wang, Z., Liu, W., Qian, C., *et al.*: 'Two-dimensional precoding for 3-D massive MIMO', *IEEE Trans. Veh. Technol.*, 2017, **66**, (6), pp. 5485–5490

[15] Zhu, G., Huang, K., Lau, V.K., *et al.*: 'Hybrid beamforming via the kronecker decomposition for the millimeter-wave massive MIMO systems', *IEEE J. Sel. Areas Commun.*, 2017, **35**, (9), pp. 2097–2114

[16] Miranda, R.K., da Costa, J.P.C., Roemer, F., *et al.*: 'Generalized sidelobe cancellers for multidimensional separable arrays'. 2015 IEEE 6th Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Cancun, 2015, pp. 193–196

[17] Liu, L., Xie, J., Wang, L., *et al.*: 'Robust tensor beamforming for polarization sensitive arrays', *Multidimens. Syst. Signal Process.*, 2018, doi: 10.1007/s11045-018-0580-6

[18] Ribeiro, L.N., Schwarz, S., Rupp, M., *et al.*: 'A low-complexity equalizer for massive MIMO systems based on array separability'. Proc. 2017 25th European Signal Processing Conf. (EUSIPCO), Kos, 2017, pp. 2522–2526

[19] Haykin, S.: '*Adaptive filtering theory*' (Prentice-Hall, Englewood Cliffs, NJ, 1996)

[20] Petersen, K.B., Pedersen, M.S.: '*The matrix cookbook*' (Technical University of Denmark, Copenhagen, 2012)

[21] Ribeiro, L.N., De Almeida, A.L.F., Mota, J.C.M.: 'Tensor beamforming for multilinear translation invariant arrays'. Proc. 2016 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, 2016, pp. 2966–2970

[22] Kolda, T., Bader, B.: 'Tensor decompositions and applications', *SIAM Rev.*, 2009, **51**, (3), pp. 455–500

[23] Comon, P., Luciani, X., De Almeida, A.L.: 'Tensor decompositions, alternating least squares and other tales', *J. Chemometrics: A J. the Chemometrics Soc.*, 2009, **23**, (7–8), pp. 393–405

[24] Liu, S., Trenkler, G.: 'Hadamard, Khatri-Rao, Kronecker and other matrix products', *Int. J. Inf. Syst. Sci.*, 2008, **4**, (1), pp. 160–177

[25] Palomar, D.P., Eldar, Y.C.: '*Convex optimization in signal processing and communications*' (Cambridge University Press, Cambridge, 2010)

[26] Yener, A., Yates, R.D., Ulukus, S.: 'Interference management for CDMA systems through power control, multiuser detection, and beamforming', *IEEE Trans. Commun.*, 2001, **49**, (7), pp. 1227–1239

[27] Minka, T.: 'The lightspeed MATLAB toolbox, version 2.8'. 2017. Available at https://github.com/tminka/lightspeed

[28] Golub, G.H., Van Loan, C.F.: '*Matrix computations*', vol. **3** (JHU Press, Baltimore, MD, 2012)

8 Appendix

It is well-known that the minimisers of (14) and (15) are given by the classical MMSE filters $\boldsymbol{w}_h = \boldsymbol{R}_{hh}^{-1}\boldsymbol{p}_{hs}$ and $\boldsymbol{w}_v = \boldsymbol{R}_{vv}^{-1}\boldsymbol{p}_{vs}$, respectively. In this appendix, we obtain the covariance matrices and cross-covariance vectors necessary to calculate these filters. In our demonstrations, we consider only the horizontal sub-array. The statistics for the vertical sub-array are analogously derived.

First, let us represent (9) in terms of the matrix unfoldings of $\mathscr{A}$. Unfolding this tensor along its first mode gives [22]

$$\boldsymbol{X}[k] = [\mathscr{A}]_{(1)}(\boldsymbol{s}[k] \otimes \boldsymbol{I}_{N_v}) + \boldsymbol{B}[k]\,.$$

Now the horizontal sub-array input (12) can be expressed as

$$\begin{aligned}\boldsymbol{u}_h[k] &\overset{(a)}{=} \big[[\mathscr{A}]_{(1)}(\boldsymbol{s}[k] \otimes \boldsymbol{I}_{N_v}) + \boldsymbol{B}[k]\big](1 \otimes \boldsymbol{w}_v^*) \\ &\overset{(b)}{=} [\mathscr{A}]_{(1)}(\boldsymbol{s}[k] \otimes \boldsymbol{w}_v^*) + \boldsymbol{B}[k]\boldsymbol{w}_v^*,\end{aligned} \tag{26}$$

where (a) follows by considering $\boldsymbol{w}_v^* = 1 \otimes \boldsymbol{w}_v^*$, and (b) is the application of the mixed product property [24]:

$$(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{AC}) \otimes (\boldsymbol{BD})$$

for any matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ with matching dimensions. The covariance matrix of (26) is then given by

$$\boldsymbol{R}_{hh} = [\mathscr{A}]_{(1)}(\boldsymbol{R}_{ss} \otimes \boldsymbol{w}_v^*\boldsymbol{w}_v^{\mathrm{T}})[\mathscr{A}]_{(1)}^{\mathrm{H}} + \boldsymbol{R}_{cc},$$

where $\boldsymbol{R}_{cc} = \mathbb{E}[\boldsymbol{c}[k]\boldsymbol{c}^{\mathrm{H}}[k]]$, and $\boldsymbol{c}[k] = \boldsymbol{B}[k]\boldsymbol{w}_v^* \in \mathbb{C}^{N_h}$. Note we have used the fact that $\mathbb{E}[\boldsymbol{A} \otimes \boldsymbol{B}] = \mathbb{E}[\boldsymbol{A}] \otimes \mathbb{E} - [\boldsymbol{B}]$ for matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ with mutually independent elements, and that the inputs of $\boldsymbol{B}[k]$ and $\boldsymbol{s}[k]$ are uncorrelated. To calculate $\boldsymbol{R}_{cc}$, consider the element-wise representation of $\boldsymbol{c}[k]$:

$$[\boldsymbol{c}[k]]_{n_h} = \sum_{n_v=1}^{N_v} [\boldsymbol{B}[k]]_{n_h,n_v}[\boldsymbol{w}_v]_{n_v}^*, \quad n_h \in \{1,\ldots,N_h\}\,.$$

The elements of $\boldsymbol{R}_{cc}$ are given by:

$$\begin{aligned}[\boldsymbol{R}_{cc}]_{n_h,n'_h} &= \mathbb{E}\Big[[\boldsymbol{c}[k]]_{n_h}[\boldsymbol{c}[k]^{\mathrm{H}}]_{n'_h}\Big] \\ &= \mathbb{E}\left[\sum_{n_v=1}^{N_v}\sum_{n'_v=1}^{N_v}[\boldsymbol{B}[k]]_{n_h,n_v}[\boldsymbol{B}[k]]_{n'_h,n'_v}^*[\boldsymbol{w}_v]_{n_v}[\boldsymbol{w}_v]_{n'_v}^*\right]\end{aligned}$$

for $n_h, n'_h \in \{1,\ldots,N_h\}$. As the AWGN vector has mutually independent elements, it follows that

$$\mathbb{E}\big[[\boldsymbol{B}[k]]_{n_h,n_v}[\boldsymbol{B}[k]]_{n'_h,n'_v}^*\big] = 0$$

for all $n_v \neq n'_v$ and $n_h \neq n'_h$. Therefore, the off-diagonal elements of $\boldsymbol{R}_{cc}$ are zero and those at the main diagonal are given by

$$\begin{aligned}[\boldsymbol{R}_{cc}]_{n_h,n_h} = &\sum_{n_v=1}^{N_v}\mathbb{E}\big[[\boldsymbol{B}[k]]_{n_h,n_v}[\boldsymbol{B}[k]]_{n_h,n_v}^*\big][\boldsymbol{w}_v]_{n_v}[\boldsymbol{w}_v]_{n_v}^* \\ &+\sigma_b^2\sum_{n_v=1}^{N_v}[\boldsymbol{w}_v]_{n_v}[\boldsymbol{w}_v]_{n_v}^* = \sigma_b^2 \parallel \boldsymbol{w}_v \parallel_2^2\end{aligned}$$

and, consequently, we get $\boldsymbol{R}_{cc} = \sigma_b^2 \parallel \boldsymbol{w}_v \parallel_2^2 \boldsymbol{I}_{N_h}$, concluding the derivation of $\boldsymbol{R}_{hh}$. From the definition of $\boldsymbol{p}_{hs}$ and $\boldsymbol{u}_h[k]$, it follows that:

$$\begin{aligned}\boldsymbol{p}_{hs} &= \mathbb{E}\big[([\mathscr{A}]_{(1)}(\boldsymbol{s}[k] \otimes \boldsymbol{w}_v^*) + \boldsymbol{B}[k]\boldsymbol{w}_v^*)(s_d^*[k] \otimes 1)\big] \\ &= [\mathscr{A}]_{(1)}(\boldsymbol{s}[k]s_d^*[k] \otimes \boldsymbol{w}_v^*) = [\mathscr{A}]_{(1)}(\boldsymbol{R}_{ss}\boldsymbol{e}_d \otimes \boldsymbol{w}_v^*),\end{aligned}$$

finalising our proof.

□