Introduction to Complex-Valued Matrix Differentiation

Are Hjørungnes, *Senior Member, IEEE*, and David Gesbert, *Senior Member, IEEE*.

Abstract

A systematic theory is introduced for finding the derivatives of complex-valued matrix functions with respect to a complex-valued matrix variable and the complex conjugate of this variable. In the framework introduced, the differential of the complex-valued matrix function is used to identify the derivatives of this function. Matrix differentiation results are derived and summarized in tables which can be exploited in a wide range of signal processing related situations. The Hessian matrix of a scalar complex function is also defined, and it is shown how it can be obtained from the second-order differential of the function. The methods given are general such that many other results can be derived using the framework introduced, although only important cases are treated.

Keywords: Complex differentials, non-analytical complex functions, complex matrix derivatives, complex Hessian.

I. Introduction

In many engineering problems, the unknown parameters are complex-valued vectors and matrices and, often, the task of the system designer is to find the values of these complex parameters which optimize a chosen criterion function. Examples of areas where these types of optimization problems might appear, can be in telecommunications, where digital filters can contain complex-valued coefficients [1], electric circuits [2], control theory [3], adaptive filters [4], acoustics, optics, mechanical vibrating systems, heat conduction, fluid flow, and electrostatics [5]. For solving this kind of optimization problems, one approach is to find necessary conditions for optimality. When a scalar real-valued function depends on a complex-valued matrix parameter, the necessary conditions for optimality

Are Hjørungnes is with UniK - University Graduate Center, University of Oslo, Instituttveien 25, P. O. Box 70, N-2027 Kjeller, Norway, email: arehj@unik.no.

David Gesbert is with Mobile Communications Department, Eurécom Institute, 2229 Route des Crêtes, BP 193, F-06904 Sophia Antipolis Cédex, France, email: gesbert@eurecom.fr.

This work was supported by the Research Council of Norway (NFR) and the French Ministry of Foreign Affairs through the Aurora project entitled "Optimization of Broadband Wireless Communications Network" and NFR project 176773/S10.

can be found by either setting the derivative of the function with respect to the complex-valued matrix parameter or its complex conjugate to zero. Differentiation results are well-known for certain classes of functions, e.g., quadratic functions, but can be tricky for others. This paper provides the tools for finding derivatives in a systematic way. In an effort to build adaptive optimization algorithms, it will also be shown that the direction of maximum rate of change of a real-valued scalar function, with respect to the complex-valued matrix parameter, is given by the derivative of the function with respect to the complex conjugate of the complex-valued input matrix parameter. Of course, this is a generalization of a well-known result for scalar functions of vector variables. A general framework is introduced here showing how to find the derivative of complex-valued scalar-, vector-, or matrix functions with respect to the complex-valued input parameter matrix and its complex conjugate. The main contribution of this paper lies in the proposed approach for how to obtain derivatives in a way that is both simple and systematic, based on the so-called differential of the function. In this paper, it is assumed that the functions are differentiable with respect to the complex-valued parameter matrix and its complex conjugate, and it will be seen that these two parameter matrices should be treated as independent when finding the derivatives, as is classical for scalar variables. It is also shown how the Hessian matrix, i.e., second-order derivate, of a complex scalar function can be calculated. The proposed theory is useful when solving numerous problems which involve optimization when the unknown parameter is a complex-valued matrix.

The problem at hand has been treated for *real-valued* matrix variables in [6], [7], [8], [9], [10]. Four additional references that give a brief treatment of the case of real-valued scalar functions which depend complex-valued vectors are Appendix B of [11], Appendix 2.B in [12], Subsection 2.3.10 of [13], and the article [14]. The article [15] serves as an introduction to this area for complex-valued scalar functions with complex-valued argument vectors. Results on complex differentiation theory is given in [16], [17] for differentiation with respect to complex-valued scalars and vectors, however, the more general matrix case is not considered. Examples of problems where the unknown matrix is a complex-valued matrix are wide ranging including precoding of MIMO systems [18], linear equalization design [19], array signal processing [20] to only cite a few.

Some of the most relevant applications to signal and communication problems are presented here, with key results being highlighted and other illustrative examples are listed in tables. The complex Hessian matrix is defined for

TABLE I
CLASSIFICATION OF FUNCTIONS.

Function type	Scalar variables $z, z^* \in \mathbb{C}$	Vector variables $\boldsymbol{z}, \boldsymbol{z}^* \in \mathbb{C}^{N\times 1}$	Matrix variables $\boldsymbol{Z}, \boldsymbol{Z}^* \in \mathbb{C}^{N\times Q}$
Scalar function $f \in \mathbb{C}$	$f(z, z^*)$ $f : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$	$f(\boldsymbol{z}, \boldsymbol{z}^*)$ $f : \mathbb{C}^{N\times 1} \times \mathbb{C}^{N\times 1} \to \mathbb{C}$	$f(\boldsymbol{Z}, \boldsymbol{Z}^*)$ $f : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}$
Vector function $\boldsymbol{f} \in \mathbb{C}^{M\times 1}$	$\boldsymbol{f}(z, z^*)$ $\boldsymbol{f} : \mathbb{C} \times \mathbb{C} \to \mathbb{C}^{M\times 1}$	$\boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*)$ $\boldsymbol{f} : \mathbb{C}^{N\times 1} \times \mathbb{C}^{N\times 1} \to \mathbb{C}^{M\times 1}$	$\boldsymbol{f}(\boldsymbol{Z}, \boldsymbol{Z}^*)$ $\boldsymbol{f} : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}^{M\times 1}$
Matrix function $\boldsymbol{F} \in \mathbb{C}^{M\times P}$	$\boldsymbol{F}(z, z^*)$ $\boldsymbol{F} : \mathbb{C} \times \mathbb{C} \to \mathbb{C}^{M\times P}$	$\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*)$ $\boldsymbol{F} : \mathbb{C}^{N\times 1} \times \mathbb{C}^{N\times 1} \to \mathbb{C}^{M\times P}$	$\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$ $\boldsymbol{F} : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}^{M\times P}$

complex scalar functions, and it is shown how it can be obtained from the second-order differential of the functions. Hessian matrices can be used to check if a stationary point is a minimum, maximum, or a saddle point.

The rest of this paper is organized as follows: Section II contains a discussion on the differences between analytical functions, which are usually studied in mathematical courses for engineers, and non-analytical functions, which are often encountered when dealing with engineering problems of complex variables. In Section III, the complex differential is introduced, and based on this differential, the definition of the derivatives of complex-valued matrix function with respect to the complex-valued matrix argument and its complex conjugate is given in Section IV. The key procedure showing how the derivatives can be found from the differential of a function is also presented in Section IV. Section V contains important results such as the chain rule, equivalent conditions for finding stationary points, and in which direction the function has the maximum rate of change. In Section VI, several key results are placed in tables and some results are derived for various cases with high relevance for signal processing and communication problems. The Hessian matrix of scalar complex-valued functions dependent on complex matrices is defined in Section VII, and it is shown how it can be obtained. Section VIII contains some conclusions. Some of the proofs are given in the appendices.

Notation: The following conventions are always used in this article: Scalar quantities (variables z or functions f) are denoted by lowercase symbols, vector quantities (variables $\boldsymbol{z}$ or functions $\boldsymbol{f}$) are denoted by lowercase boldface symbols, and matrix quantities (variables $\boldsymbol{Z}$ or functions $\boldsymbol{F}$) are denoted by capital boldface symbols. The types of functions used throughout this paper are classified in Table I. From the table, it is seen that all the functions depend on a complex variable and the complex conjugate of the same variable. Let $\jmath = \sqrt{-1}$ be the imaginary unit. Let the symbol $z = x + \jmath y$ denote a complex variable, where the real and imaginary parts are denoted by x and y, respectively. The real and imaginary operators return the real and imaginary parts of the input matrix, respectively. These operators are denoted by $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$. If $\boldsymbol{Z} \in \mathbb{C}^{N\times Q}$ is a complex-valued[1] matrix, then $\boldsymbol{Z} = \mathrm{Re}\{\boldsymbol{Z}\} + \jmath\,\mathrm{Im}\{\boldsymbol{Z}\}$, and $\boldsymbol{Z}^* = \mathrm{Re}\{\boldsymbol{Z}\} - \jmath\,\mathrm{Im}\{\boldsymbol{Z}\}$, where $\mathrm{Re}\{\boldsymbol{Z}\} \in \mathbb{R}^{N\times Q}$, $\mathrm{Im}\{\boldsymbol{Z}\} \in \mathbb{R}^{N\times Q}$, and the operator $(\cdot)^*$ denotes complex conjugate of the matrix it is applied to. The real and imaginary operators can be expressed as $\mathrm{Re}\{\boldsymbol{Z}\} = \frac{1}{2}(\boldsymbol{Z} + \boldsymbol{Z}^*)$ and $\mathrm{Im}\{\boldsymbol{Z}\} = \frac{1}{2\jmath}(\boldsymbol{Z} - \boldsymbol{Z}^*)$.

II. Analytical Versus Non-Analytical Functions

Mathematical courses on complex functions for engineers often only involve the analysis of analytical functions:

Definition 1: *Let $D \subseteq \mathbb{C}$ be the domain of definition of the function $f : D \to \mathbb{C}$. The function f, which depends on a complex scalar variable z, is an* analytical function in the domain D *[5] if* $\lim_{\Delta z \to 0} \frac{f(z+\Delta z) - f(z)}{\Delta z}$ *exists for all $z \in D$.*

If $f(z)$ satisfies the Cauchy-Riemann equations [5], then it is analytical. A function that is complex differentiable is also named *analytic*, *holomorphic*, or *regular*. The Cauchy-Riemann condition can be formulated as one equation as $\frac{\partial}{\partial z^*} f = 0$, where the derivative of f with respect to z^* and z, respectively, is found by treating the variable z and z^* as a constant, i.e., $\frac{\partial}{\partial z^*} z = 0$ and $\frac{\partial}{\partial z} z^* = 0$, see also Definition 2 where this kind of partial derivatives will be defined with more details. From $\frac{\partial}{\partial z^*} f = 0$, it is seen that any analytical function f is *not* dependent on the variable z^*. This can also be seen from Theorem 1, page 804 in [5], which states that any analytical function $f(z)$ can be written as an infinite power series[2] with non-negative exponents of the complex variable z, and this power series is called the Taylor series. The series does *not* contain any terms that depend on z^*. The derivative of a complex-valued function in mathematical courses of complex analysis for engineers is often defined only for analytical functions. However, in many engineering problems, the functions of interest are often *not* analytic, since they are often real-valued functions. Conversely, if a function is only dependent on z, like an analytical function, and not implicitly or explicitly dependent on z^*, then this function cannot *in general* be real-valued, since the function can only be real if the imaginary parts of z can be eliminated, and this is handled by the terms that depend on z^*. An alternative treatment of how to find the derivative of real functions dependent on complex variables than the one used for analytical function is needed. In this article, a systematic theory is provided for finding the derivative of non-analytical functions.

[1] $\mathbb{R}$ and $\mathbb{C}$ are the sets of the real and complex numbers, respectively.

[2] A power series in the variable $z \in \mathbb{C}$ is an infinite sum of the form $\sum_{n=0}^{\infty} a_n (z - z_0)^n$, where $a_n, z_0 \in \mathbb{C}$ [5].

A. Euclidean Distance

In engineering problems, the squared Euclidean distance is often used. Let $f(z) = |z|^2 = zz^*$. If the traditional definition of the derivative given in Definition 1 is used, then the function f is *not* differentiable because:

$$\lim_{\Delta z \to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} = \lim_{\Delta z \to 0} \frac{|z_0 + \Delta z|^2 - |z_0|^2}{\Delta z} = \lim_{\Delta z \to 0} \frac{(\Delta z)\, z_0^* + z_0 (\Delta z)^* + \Delta z (\Delta z)^*}{\Delta z}. \tag{1}$$

This limit does *not* exist because different values are found depending on how Δz is approaching 0. Let $\Delta z = \Delta x + \jmath \Delta y$. Firstly, let Δz approach zero such that $\Delta x = 0$, then the last fraction in (1) is $\frac{\jmath(\Delta y) z_0^* - \jmath z_0 \Delta y + (\Delta y)^2}{\jmath \Delta y} = z_0^* - z_0 - \jmath \Delta y$, which approaches $z_0^* - z_0 = -2\jmath \operatorname{Im}\{z_0\}$ when $\Delta y \to 0$. Secondly, let Δz approach zero such that $\Delta y = 0$, then the last fraction in (1) is $\frac{(\Delta x) z_0^* + z_0 \Delta x + (\Delta x)^2}{\Delta x} = z_0 + z_0^* + \Delta x$, which approaches $z_0 + z_0^* = 2 \operatorname{Re}\{z_0\}$ when $\Delta x \to 0$. For any non-zero complex number z_0 the following holds: $\operatorname{Re}\{z_0\} \neq -\jmath \operatorname{Im}\{z_0\}$, and therefore, the function $f(z) = |z|^2$ is *not* differentiable when using the commonly encountered definition given in Definition 1.

There are two alternative ways [13] to find the derivative of a scalar real-valued function $f \in \mathbb{R}$ with respect to the complex-valued matrix parameter $\boldsymbol{Z} \in \mathbb{C}^{N \times Q}$. The standard way is to rewrite f as a function of the real $\boldsymbol{X}$ and imaginary parts $\boldsymbol{Y}$ of the complex variable $\boldsymbol{Z}$, and then to find the derivatives of the rewritten function with respect to these two independent real variables, $\boldsymbol{X}$ and $\boldsymbol{Y}$, separately. Notice that NQ independent complex parameters in $\boldsymbol{Z}$ correspond to $2NQ$ independent real parameters in $\boldsymbol{X}$ and $\boldsymbol{Y}$. In this paper, we bring the reader's attention onto an alternative approach that can lead to a simpler derivation. In this approach, one treats the differentials of the variables variables $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ as independent, in the way that will be shown by Lemma 1, see also [15]. It will be shown later, that both the derivative of f with respect to $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ can be identified by the differential of f.

B. Formal Partial Derivatives

Definition 2: *The derivatives with respect to z and z^* of $f(z_0)$ are called the* formal partial derivatives *of f at $z_0 \in \mathbb{C}$. Let $z = x + \jmath y$, where $x, y \in \mathbb{R}$, then the formal derivatives, or Wirtinger derivatives [21], are defined as:*

$$\frac{\partial}{\partial z} f(z_0) \triangleq \frac{1}{2} \left(\frac{\partial}{\partial x} f(z_0) - \jmath \frac{\partial}{\partial y} f(z_0) \right) \quad \textit{and} \quad \frac{\partial}{\partial z^*} f(z_0) \triangleq \frac{1}{2} \left(\frac{\partial}{\partial x} f(z_0) + \jmath \frac{\partial}{\partial y} f(z_0) \right). \tag{2}$$

Alternatively, when finding $\frac{\partial}{\partial z} f(z_0)$ and $\frac{\partial}{\partial z^*} f(z_0)$, the parameters z and z^*, respectively, are treated as independent variables, see [15] for more details.

From Definition 2, it follows that $\frac{\partial}{\partial x} f(z_0)$ and $\frac{\partial}{\partial y} f(z_0)$ can be found as:

$$\frac{\partial}{\partial x} f(z_0) = \frac{\partial}{\partial z} f(z_0) + \frac{\partial}{\partial z^*} f(z_0) \quad \text{and} \quad \frac{\partial}{\partial y} f(z_0) = \jmath \left(\frac{\partial}{\partial z} f(z_0) - \frac{\partial}{\partial z^*} f(z_0) \right). \tag{3}$$

If f is dependent on several variables, Definition 2 can be extended. Later in this article, it will be shown how the derivatives, with respect to the complex-valued matrix parameter and its complex conjugate, of all the function types given in Table I can be identified from the differentials of these functions.

Example 1: *Let $f : \mathbb{C} \times \mathbb{C} \to \mathbb{R}$ be $f(z, z^*) = zz^*$. f is differentiable*[3] *with respect to both variables z and z^*, and $\frac{\partial}{\partial z} f(z, z^*) = z^*$ and $\frac{\partial}{\partial z^*} f(z, z^*) = z$. When the complex variable z and its complex conjugate twin z^* are treated as independent variables [15] when finding the derivatives of f, then f is differentiable in both of these variables. However, as shown earlier in this section, the same function is* not *differentiable when studying* analytical *functions defined according to Definition* 1.

III. Complex Differentials

The differential has the same size as the matrix it is applied to. The differential can be found component-wise, that is, $(d\boldsymbol{Z})_{k,l} = d(\boldsymbol{Z})_{k,l}$. A procedure that can often be used for finding the differentials of a complex matrix function[4] $\boldsymbol{F}(\boldsymbol{Z}_0, \boldsymbol{Z}_1)$ is to calculate the difference

$$\boldsymbol{F}(\boldsymbol{Z}_0 + d\boldsymbol{Z}_0, \boldsymbol{Z}_1 + d\boldsymbol{Z}_1) - \boldsymbol{F}(\boldsymbol{Z}_0, \boldsymbol{Z}_1) = \text{First-order}(d\boldsymbol{Z}_0, d\boldsymbol{Z}_1) + \text{Higher-order}(d\boldsymbol{Z}_0, d\boldsymbol{Z}_1), \tag{4}$$

where First-order$(\cdot,\cdot)$ returns the terms that depend on either $d\boldsymbol{Z}_0$ or $d\boldsymbol{Z}_1$ of the first order, and Higher-order$(\cdot,\cdot)$ returns the terms that depend on the higher order terms of $d\boldsymbol{Z}_0$ and $d\boldsymbol{Z}_1$. The differential is then given by First-order$(\cdot,\cdot)$, i.e., the first order term of $\boldsymbol{F}(\boldsymbol{Z}_0 + d\boldsymbol{Z}_0, \boldsymbol{Z}_1 + d\boldsymbol{Z}_1) - \boldsymbol{F}(\boldsymbol{Z}_0, \boldsymbol{Z}_1)$. As an example, let $\boldsymbol{F}(\boldsymbol{Z}_0, \boldsymbol{Z}_1) = \boldsymbol{Z}_0 \boldsymbol{Z}_1$. Then the difference in (4) can be developed and readily expressed as: $\boldsymbol{F}(\boldsymbol{Z}_0 + d\boldsymbol{Z}_0, \boldsymbol{Z}_1 + d\boldsymbol{Z}_1) - \boldsymbol{F}(\boldsymbol{Z}_0, \boldsymbol{Z}_1) = \boldsymbol{Z}_0 d\boldsymbol{Z}_1 + (d\boldsymbol{Z}_0)\boldsymbol{Z}_1 + (d\boldsymbol{Z}_0)(d\boldsymbol{Z}_1)$. The differential of $\boldsymbol{Z}_0\boldsymbol{Z}_1$ can then be identified as all the first-order terms on either $d\boldsymbol{Z}_0$ or $d\boldsymbol{Z}_1$ as $d\boldsymbol{Z}_0\boldsymbol{Z}_1 = \boldsymbol{Z}_0 d\boldsymbol{Z}_1 + (d\boldsymbol{Z}_0)\boldsymbol{Z}_1$.

Definition 3: *The* Moore-Penrose inverse *of $\boldsymbol{Z} \in \mathbb{C}^{N \times Q}$ is denoted by $\boldsymbol{Z}^+ \in \mathbb{C}^{Q \times N}$, and it is defined through the following four relations [22]:*

$$\left(\boldsymbol{Z}\boldsymbol{Z}^+\right)^H = \boldsymbol{Z}\boldsymbol{Z}^+, \tag{5}$$

[3]Here, Definition 2 is used for finding the formal partial derivatives.

[4]The indexes are chosen to start with 0 everywhere in this article.

$$\left(\boldsymbol{Z}^{+}\boldsymbol{Z}\right)^{H} = \boldsymbol{Z}^{+}\boldsymbol{Z}, \tag{6}$$

$$\boldsymbol{Z}\boldsymbol{Z}^{+}\boldsymbol{Z} = \boldsymbol{Z}, \tag{7}$$

$$\boldsymbol{Z}^{+}\boldsymbol{Z}\boldsymbol{Z}^{+} = \boldsymbol{Z}^{+}, \tag{8}$$

where the operator $(\cdot)^H$ *is the Hermitian operator, or the complex conjugate transpose.*

Let $\otimes$ and $\odot$ denote the Kronecker and Hadamard product [23], respectively. Some of the most important rules on complex differentials are listed in Table II, assuming $\boldsymbol{A}$, $\boldsymbol{B}$, and a to be constants, and $\boldsymbol{Z}$, $\boldsymbol{Z}_0$, and $\boldsymbol{Z}_1$ to be complex-valued matrix variables. The vectorization operator $\operatorname{vec}(\cdot)$ stacks the columns vectors of the argument matrix into a long column vector in chronological order [23]. The differentiation rule of the reshaping operator $\operatorname{reshape}(\cdot)$ in Table II is valid for any *linear reshaping*[5] operator $\operatorname{reshape}(\cdot)$ of the matrix, and examples of such operators are the transpose $(\cdot)^T$ or $\operatorname{vec}(\cdot)$. Some of the basic differential results in Table II can be derived by means of (4), and others can be derived by generalizing some of the results found in [7], [9] to the complex differential case.

Differential of the Moore-Penrose Inverse: The differential of the *real-valued* Moore-Penrose inverse can be found in in [7], [8], but the complex-valued version is derived here.

Proposition 1: *Let* $\boldsymbol{Z} \in \mathbb{C}^{N\times Q}$, *then*

$$d\boldsymbol{Z}^{+} = -\boldsymbol{Z}^{+}(d\boldsymbol{Z})\boldsymbol{Z}^{+} + \boldsymbol{Z}^{+}(\boldsymbol{Z}^{+})^{H}(d\boldsymbol{Z}^{H})\left(\boldsymbol{I}_N - \boldsymbol{Z}\boldsymbol{Z}^{+}\right) + \left(\boldsymbol{I}_Q - \boldsymbol{Z}^{+}\boldsymbol{Z}\right)(d\boldsymbol{Z}^{H})(\boldsymbol{Z}^{+})^{H}\boldsymbol{Z}^{+}. \tag{9}$$

The proof of Proposition 1 can be found in Appendix I.

From Table II, the following four equalities follows $d\boldsymbol{Z} = d\operatorname{Re}\{\boldsymbol{Z}\} + \jmath d\operatorname{Im}\{\boldsymbol{Z}\}$, $d\boldsymbol{Z}^* = d\operatorname{Re}\{\boldsymbol{Z}\} - \jmath d\operatorname{Im}\{\boldsymbol{Z}\}$, $d\operatorname{Re}\{\boldsymbol{Z}\} = \frac{1}{2}\left(d\boldsymbol{Z} + d\boldsymbol{Z}^*\right)$, and $d\operatorname{Im}\{\boldsymbol{Z}\} = \frac{1}{2\jmath}\left(d\boldsymbol{Z} - d\boldsymbol{Z}^*\right)$.

The following two lemmas are used to identify the first- and second-order derivatives later in the article. The real variables $\operatorname{Re}\{\boldsymbol{Z}\}$ and $\operatorname{Im}\{\boldsymbol{Z}\}$ are independent of each other and hence are their differentials. Although the complex variables $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ are related, their differentials are linearly independent in the following way:

Lemma 1: *Let* $\boldsymbol{Z} \in \mathbb{C}^{N\times Q}$ *and let* $\boldsymbol{A}_i \in \mathbb{C}^{M\times NQ}$. *If* $\boldsymbol{A}_0 d\operatorname{vec}(\boldsymbol{Z}) + \boldsymbol{A}_1 d\operatorname{vec}(\boldsymbol{Z}^*) = \boldsymbol{0}_{M\times 1}$ *for all* $d\boldsymbol{Z} \in \mathbb{C}^{N\times Q}$, *then* $\boldsymbol{A}_i = \boldsymbol{0}_{M\times NQ}$ *for* $i \in \{0, 1\}$.

[5] The output of the reshape operator has the same number of elements as the input, but the shape of the output might be different, so $\operatorname{reshape}(\cdot)$ performs a reshaping of its input argument.

TABLE II
IMPORTANT RESULTS FOR COMPLEX DIFFERENTIALS.

Function	Differential
$\boldsymbol{A}$	$\boldsymbol{0}$
$a\boldsymbol{Z}$	$a d\boldsymbol{Z}$
$\boldsymbol{A}\boldsymbol{Z}\boldsymbol{B}$	$\boldsymbol{A}(d\boldsymbol{Z})\boldsymbol{B}$
$\boldsymbol{Z}_0 + \boldsymbol{Z}_1$	$d\boldsymbol{Z}_0 + d\boldsymbol{Z}_1$
$\mathrm{Tr}\{\boldsymbol{Z}\}$	$\mathrm{Tr}\{d\boldsymbol{Z}\}$
$\boldsymbol{Z}_0\boldsymbol{Z}_1$	$(d\boldsymbol{Z}_0)\boldsymbol{Z}_1 + \boldsymbol{Z}_0(d\boldsymbol{Z}_1)$
$\boldsymbol{Z}_0 \otimes \boldsymbol{Z}_1$	$(d\boldsymbol{Z}_0) \otimes \boldsymbol{Z}_1 + \boldsymbol{Z}_0 \otimes (d\boldsymbol{Z}_1)$
$\boldsymbol{Z}_0 \odot \boldsymbol{Z}_1$	$(d\boldsymbol{Z}_0) \odot \boldsymbol{Z}_1 + \boldsymbol{Z}_0 \odot (d\boldsymbol{Z}_1)$
$\boldsymbol{Z}^{-1}$	$-\boldsymbol{Z}^{-1}(d\boldsymbol{Z})\boldsymbol{Z}^{-1}$
$\det(\boldsymbol{Z})$	$\det(\boldsymbol{Z})\,\mathrm{Tr}\left\{\boldsymbol{Z}^{-1}d\boldsymbol{Z}\right\}$
$\ln(\det(\boldsymbol{Z}))$	$\mathrm{Tr}\left\{\boldsymbol{Z}^{-1}d\boldsymbol{Z}\right\}$
$\mathrm{reshape}(\boldsymbol{Z})$	$\mathrm{reshape}(d\boldsymbol{Z})$
$\boldsymbol{Z}^*$	$(d\boldsymbol{Z})^*$
$\boldsymbol{Z}^H$	$(d\boldsymbol{Z})^H$
$\boldsymbol{Z}^+$	$-\boldsymbol{Z}^+(d\boldsymbol{Z})\boldsymbol{Z}^+ + \boldsymbol{Z}^+(\boldsymbol{Z}^+)^H(d\boldsymbol{Z}^H)\left(\boldsymbol{I}_N - \boldsymbol{Z}\boldsymbol{Z}^+\right) + \left(\boldsymbol{I}_Q - \boldsymbol{Z}^+\boldsymbol{Z}\right)(d\boldsymbol{Z}^H)(\boldsymbol{Z}^+)^H\boldsymbol{Z}^+$

The proof of Lemma 1 can be found in Appendix II.

Lemma 2: *Let* $\boldsymbol{Z} \in \mathbb{C}^{N\times Q}$ *and let* $\boldsymbol{B}_i \in \mathbb{C}^{NQ\times NQ}$. *If* $\left(d\,\mathrm{vec}^T(\boldsymbol{Z})\right)\boldsymbol{B}_0 d\,\mathrm{vec}(\boldsymbol{Z}) + \left(d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right)\boldsymbol{B}_1 d\,\mathrm{vec}(\boldsymbol{Z}) + \left(d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right)\boldsymbol{B}_2 d\,\mathrm{vec}(\boldsymbol{Z}^*) = 0$ *for all* $d\boldsymbol{Z} \in \mathbb{C}^{N\times Q}$, *then* $\boldsymbol{B}_0 = -\boldsymbol{B}_0^T$, $\boldsymbol{B}_1 = \boldsymbol{0}_{NQ\times NQ}$, *and* $\boldsymbol{B}_2 = -\boldsymbol{B}_2^T$.

The proof of Lemma 2 can be found in Appendix III.

IV. DEFINITION OF THE DERIVATIVE WITH RESPECT TO COMPLEX-VALUED MATRICES

The most general definition of the derivative is given here from which the definitions for less general cases follow and they will later be given in an identification table which shows how the derivatives can be obtained from the differential of the function.

Definition 4: *Let* $\boldsymbol{F} : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}^{M\times P}$. *Then the derivative of the matrix function* $\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) \in \mathbb{C}^{M\times P}$ *with respect to* $\boldsymbol{Z} \in \mathbb{C}^{N\times Q}$ *is denoted* $\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F}$, *and the derivative of the matrix function* $\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) \in \mathbb{C}^{M\times P}$ *with respect to* $\boldsymbol{Z}^* \in \mathbb{C}^{N\times Q}$ *is denoted* $\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F}$ *and the size of both these derivatives is* $MP \times NQ$. *The derivatives* $\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F}$ *and* $\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F}$ *are defined by the following differential expression:*

$$d\,\mathrm{vec}(\boldsymbol{F}) = (\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F})\,d\,\mathrm{vec}(\boldsymbol{Z}) + (\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F})\,d\,\mathrm{vec}(\boldsymbol{Z}^*). \tag{10}$$

$\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$ *and* $\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$ *are called the* Jacobian *matrices of* $\boldsymbol{F}$ *with respect to* $\boldsymbol{Z}$ *and* $\boldsymbol{Z}^*$, *respectively.*

Remark 1: *Definition* 4 *is a generalization of Definition 1, page 173 in [7] to include complex matrices. In [7], several alternative definitions of the derivative of real-valued functions with respect to a matrix are discussed, and*

TABLE III
IDENTIFICATION TABLE.

Function type	Differential	Derivative with respect to z, $\boldsymbol{z}$, or $\boldsymbol{Z}$	Derivative with respect to z^*, $\boldsymbol{z}^*$, or $\boldsymbol{Z}^*$	Size of derivatives
$f(z, z^*)$	$df = a_0 dz + a_1 dz^*$	$\mathcal{D}_z f(z, z^*) = a_0$	$\mathcal{D}_{z^*} f(z, z^*) = a_1$	1×1
$f(\boldsymbol{z}, \boldsymbol{z}^*)$	$df = \boldsymbol{a}_0 d\boldsymbol{z} + \boldsymbol{a}_1 d\boldsymbol{z}^*$	$\mathcal{D}_{\boldsymbol{z}} f(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{a}_0$	$\mathcal{D}_{\boldsymbol{z}^*} f(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{a}_1$	$1 \times N$
$f(\boldsymbol{Z}, \boldsymbol{Z}^*)$	$df = \mathrm{vec}^T(\boldsymbol{A}_0) d\,\mathrm{vec}(\boldsymbol{Z}) + \mathrm{vec}^T(\boldsymbol{A}_1) d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\mathcal{D}_{\boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \mathrm{vec}^T(\boldsymbol{A}_0)$	$\mathcal{D}_{\boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \mathrm{vec}^T(\boldsymbol{A}_1)$	$1 \times NQ$
$f(\boldsymbol{Z}, \boldsymbol{Z}^*)$	$df = \mathrm{Tr}\left\{\boldsymbol{A}_0^T d\boldsymbol{Z} + \boldsymbol{A}_1^T d\boldsymbol{Z}^*\right\}$	$\frac{\partial}{\partial \boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{A}_0$	$\frac{\partial}{\partial \boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{A}_1$	$N \times Q$
$\boldsymbol{f}(z, z^*)$	$d\boldsymbol{f} = \boldsymbol{b}_0 dz + \boldsymbol{b}_1 dz^*$	$\mathcal{D}_z \boldsymbol{f}(z, z^*) = \boldsymbol{b}_0$	$\mathcal{D}_{z^*} \boldsymbol{f}(z, z^*) = \boldsymbol{b}_1$	$M \times 1$
$\boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*)$	$d\boldsymbol{f} = \boldsymbol{B}_0 d\boldsymbol{z} + \boldsymbol{B}_1 d\boldsymbol{z}^*$	$\mathcal{D}_{\boldsymbol{z}} \boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{B}_0$	$\mathcal{D}_{\boldsymbol{z}^*} \boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{B}_1$	$M \times N$
$\boldsymbol{f}(\boldsymbol{Z}, \boldsymbol{Z}^*)$	$d\boldsymbol{f} = \boldsymbol{\beta}_0 d\,\mathrm{vec}(\boldsymbol{Z}) + \boldsymbol{\beta}_1 d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\mathcal{D}_{\boldsymbol{Z}} \boldsymbol{f}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{\beta}_0$	$\mathcal{D}_{\boldsymbol{Z}^*} \boldsymbol{f}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{\beta}_1$	$M \times NQ$
$\boldsymbol{F}(z, z^*)$	$d\,\mathrm{vec}(\boldsymbol{F}) = \boldsymbol{c}_0 dz + \boldsymbol{c}_1 dz^*$	$\mathcal{D}_z \boldsymbol{F}(z, z^*) = \boldsymbol{c}_0$	$\mathcal{D}_{z^*} \boldsymbol{F}(z, z^*) = \boldsymbol{c}_1$	$MP \times 1$
$\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*)$	$d\,\mathrm{vec}(\boldsymbol{F}) = \boldsymbol{C}_0 d\boldsymbol{z} + \boldsymbol{C}_1 d\boldsymbol{z}^*$	$\mathcal{D}_{\boldsymbol{z}} \boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{C}_0$	$\mathcal{D}_{\boldsymbol{z}^*} \boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{C}_1$	$MP \times N$
$\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$	$d\,\mathrm{vec}(\boldsymbol{F}) = \boldsymbol{\zeta}_0 d\,\mathrm{vec}(\boldsymbol{Z}) + \boldsymbol{\zeta}_1 d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\mathcal{D}_{\boldsymbol{Z}} \boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{\zeta}_0$	$\mathcal{D}_{\boldsymbol{Z}^*} \boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{\zeta}_1$	$MP \times NQ$

it is concluded that the definition that matches Definition 4 *is the only reasonable definition. Definition* 4 *is also a generalization of the definition used in [15] for complex-valued* vectors *to the case of complex-valued* matrices.

Table III shows how the derivatives of the different function types in Table I can be identified from the differentials of these functions.[6] To show the uniqueness of the representation in (10), we subtract the differential in (10) from the corresponding differential in Table III to get $(\boldsymbol{\zeta}_0 - \mathcal{D}_{\boldsymbol{Z}} \boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*))\, d\,\mathrm{vec}(\boldsymbol{Z}) + (\boldsymbol{\zeta}_1 - \mathcal{D}_{\boldsymbol{Z}^*} \boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*))\, d\,\mathrm{vec}(\boldsymbol{Z}^*) = \boldsymbol{0}_{MP\times 1}$. The derivatives in the last line of Table III then follow by using Lemma 1 on this equation. Table III is an extension of the corresponding table given in [7], valid in the real variable case. In Table III, $z \in \mathbb{C}$, $\boldsymbol{z} \in \mathbb{C}^{N\times 1}$, $\boldsymbol{Z} \in \mathbb{C}^{N\times Q}$, $f \in \mathbb{C}$, $\boldsymbol{f} \in \mathbb{C}^{M\times 1}$, and $\boldsymbol{F} \in \mathbb{C}^{M\times P}$. Furthermore, $a_i \in \mathbb{C}^{1\times 1}$, $\boldsymbol{a}_i \in \mathbb{C}^{1\times N}$, $\boldsymbol{A}_i \in \mathbb{C}^{N\times Q}$, $\boldsymbol{b}_i \in \mathbb{C}^{M\times 1}$, $\boldsymbol{B}_i \in \mathbb{C}^{M\times N}$, $\boldsymbol{\beta}_i \in \mathbb{C}^{M\times NQ}$, $\boldsymbol{c}_i \in \mathbb{C}^{MP\times 1}$, $\boldsymbol{C}_i \in \mathbb{C}^{MP\times N}$, $\boldsymbol{\zeta}_i \in \mathbb{C}^{MP\times NQ}$, and each of these might be a function of z, $\boldsymbol{z}$, $\boldsymbol{Z}$, z^*, $\boldsymbol{z}^*$, or $\boldsymbol{Z}^*$.

Definition 5: *Let* $\boldsymbol{f} : \mathbb{C}^{N\times 1} \times \mathbb{C}^{N\times 1} \to \mathbb{C}^{M\times 1}$. *The partial derivatives* $\frac{\partial}{\partial \boldsymbol{z}^T} \boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*)$ *and* $\frac{\partial}{\partial \boldsymbol{z}^H} \boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*)$ *of size* $M \times N$ *are defined as*

$$\frac{\partial}{\partial \boldsymbol{z}^T} \boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*) = \begin{bmatrix} \frac{\partial}{\partial z_0} f_0 & \cdots & \frac{\partial}{\partial z_{N-1}} f_0 \\ \vdots & & \vdots \\ \frac{\partial}{\partial z_0} f_{M-1} & \cdots & \frac{\partial}{\partial z_{N-1}} f_{M-1} \end{bmatrix}, \quad \frac{\partial}{\partial \boldsymbol{z}^H} \boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*) = \begin{bmatrix} \frac{\partial}{\partial z_0^*} f_0 & \cdots & \frac{\partial}{\partial z_{N-1}^*} f_0 \\ \vdots & & \vdots \\ \frac{\partial}{\partial z_0^*} f_{M-1} & \cdots & \frac{\partial}{\partial z_{N-1}^*} f_{M-1} \end{bmatrix}, \quad (11)$$

where z_i *and* f_i *is component number* i *of the vectors* $\boldsymbol{z}$ *and* $\boldsymbol{f}$, *respectively.*

[6]For the functions of the type $f(\boldsymbol{Z}, \boldsymbol{Z}^*)$ two alternative definitions for the derivatives are given. The notation $\frac{\partial}{\partial \boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*)$ and $\frac{\partial}{\partial \boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*)$ will be defined in Subsection VI-C.

Notice that $\frac{\partial}{\partial \boldsymbol{z}^T}\boldsymbol{f} = \mathcal{D}_{\boldsymbol{z}}\boldsymbol{f}$ and $\frac{\partial}{\partial \boldsymbol{z}^H}\boldsymbol{f} = \mathcal{D}_{\boldsymbol{z}^*}\boldsymbol{f}$. Using the partial derivative notation in Definition 5, the derivatives of the function $\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$, in Definition 4, are:

$$\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \frac{\partial \operatorname{vec}(\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*))}{\partial \operatorname{vec}^T(\boldsymbol{Z})}, \tag{12}$$

$$\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \frac{\partial \operatorname{vec}(\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*))}{\partial \operatorname{vec}^T(\boldsymbol{Z}^*)}. \tag{13}$$

This is a generalization of the real-valued matrix variable case treated in [7] to the complex-valued matrix variable case. (12) and (13) show how the all the $MPNQ$ partial derivatives of all the components of $\boldsymbol{F}$ with respect to all the components of $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ are arranged when using the notation introduced in Definition 5.

Key result: Finding the derivative of the complex-valued matrix function $\boldsymbol{F}$ with respect to the complex-valued matrices $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ can be achieved using the following three-step procedure:

1) *Compute the differential* $d \operatorname{vec}(\boldsymbol{F})$.
2) *Manipulate the expression into the form given (*10*).*
3) *Read out the result using Table* III.

For less general function types, see Table I, a similar procedure can be used.

V. Fundamental Results

In this section, some useful theorems are presented that exploit the theory introduced earlier. All of these results are important when solving practical optimization problems involving differentiation with respect to a complex-valued matrix. These results include the chain rule, conditions for finding stationary points for a real scalar function dependent on complex-valued matrices, and in which direction the same type of function has the minimum or maximum rate of change. A comment will be made on how this result should be used in the *steepest decent method* [24]. For certain types of functions, the same procedure used for the real-valued matrix case [7] can be used, and this result is stated in a theorem.

1) **Chain Rule:** One big advantage of the way the derivative is defined in Definition 4 compared to other definitions of the derivative of $\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$ is that the chain rule is valid. The chain rule is now formulated.

Theorem 1: *Let* $(S_0, S_1) \subseteq \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q}$*, and let* $\boldsymbol{F} : S_0 \times S_1 \to \mathbb{C}^{M\times P}$ *be differentiable with respect to both its first and second argument at an interior point* $(\boldsymbol{Z}, \boldsymbol{Z}^*)$ *in the set* $S_0 \times S_1$*. Let* $T_0 \times T_1 \subseteq \mathbb{C}^{M\times P} \times \mathbb{C}^{M\times P}$ *be such* *that* $(\boldsymbol{F}(\boldsymbol{Z},\boldsymbol{Z}^*),\boldsymbol{F}^*(\boldsymbol{Z},\boldsymbol{Z}^*)) \in T_0 \times T_1$ *for all* $(\boldsymbol{Z},\boldsymbol{Z}^*) \in S_0 \times S_1$*. Assume that* $\boldsymbol{G} : T_0 \times T_1 \to \mathbb{C}^{R\times S}$ *is differentiable at an interior point* $(\boldsymbol{F}(\boldsymbol{Z},\boldsymbol{Z}^*),\boldsymbol{F}^*(\boldsymbol{Z},\boldsymbol{Z}^*)) \in T_0 \times T_1$*. Define the composite function* $\boldsymbol{H} : S_0 \times S_1 \to \mathbb{C}^{R\times S}$ *by* $\boldsymbol{H}(\boldsymbol{Z},\boldsymbol{Z}^*) \triangleq \boldsymbol{G}(\boldsymbol{F}(\boldsymbol{Z},\boldsymbol{Z}^*),\boldsymbol{F}^*(\boldsymbol{Z},\boldsymbol{Z}^*))$*. The derivatives* $\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{H}$ *and* $\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{H}$ *are obtained through:*

$$\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{H} = (\mathcal{D}_{\boldsymbol{F}}\boldsymbol{G})(\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F}) + (\mathcal{D}_{\boldsymbol{F}^*}\boldsymbol{G})(\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F}^*), \tag{14}$$

$$\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{H} = (\mathcal{D}_{\boldsymbol{F}}\boldsymbol{G})(\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F}) + (\mathcal{D}_{\boldsymbol{F}^*}\boldsymbol{G})(\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F}^*). \tag{15}$$

Proof: *From Definition* 4*, it follows that:*

$$d\,\mathrm{vec}(\boldsymbol{H}) = d\,\mathrm{vec}(\boldsymbol{G}) = (\mathcal{D}_{\boldsymbol{F}}\boldsymbol{G})\,d\,\mathrm{vec}(\boldsymbol{F}) + (\mathcal{D}_{\boldsymbol{F}^*}\boldsymbol{G})\,d\,\mathrm{vec}(\boldsymbol{F}^*). \tag{16}$$

The differentials of $\mathrm{vec}(\boldsymbol{F})$ *and* $\mathrm{vec}(\boldsymbol{F}^*)$ *are given by:*

$$d\,\mathrm{vec}(\boldsymbol{F}) = (\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F})\,d\,\mathrm{vec}(\boldsymbol{Z}) + (\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F})\,d\,\mathrm{vec}(\boldsymbol{Z}^*), \tag{17}$$

$$d\,\mathrm{vec}(\boldsymbol{F}^*) = (\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F}^*)\,d\,\mathrm{vec}(\boldsymbol{Z}) + (\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F}^*)\,d\,\mathrm{vec}(\boldsymbol{Z}^*). \tag{18}$$

*By substituting the results from (*17*) and (*18*), into (*16*), and then using the definition of the derivatives with respect to* $\boldsymbol{Z}$ *and* $\boldsymbol{Z}^*$ *the theorem follows.* ■

2) **Stationary Points:** The next theorem presents conditions for finding stationary points of $f(\boldsymbol{Z},\boldsymbol{Z}^*) \in \mathbb{R}$.

Theorem 2: *Let* $f : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{R}$*. A stationary point*[7] *of the function* $f(\boldsymbol{Z},\boldsymbol{Z}^*) = g(\boldsymbol{X},\boldsymbol{Y})$*, where* $g : \mathbb{R}^{N\times Q} \times \mathbb{R}^{N\times Q} \to \mathbb{R}$ *and* $\boldsymbol{Z} = \boldsymbol{X} + \jmath\boldsymbol{Y}$ *is then found by one of the following three equivalent conditions:*[8]

$$\mathcal{D}_{\boldsymbol{X}}g(\boldsymbol{X},\boldsymbol{Y}) = \mathbf{0}_{1\times NQ} \quad \wedge \quad \mathcal{D}_{\boldsymbol{Y}}g(\boldsymbol{X},\boldsymbol{Y}) = \mathbf{0}_{1\times NQ}, \tag{19}$$

$$\mathcal{D}_{\boldsymbol{Z}}f(\boldsymbol{Z},\boldsymbol{Z}^*) = \mathbf{0}_{1\times NQ}, \tag{20}$$

or

$$\mathcal{D}_{\boldsymbol{Z}^*}f(\boldsymbol{Z},\boldsymbol{Z}^*) = \mathbf{0}_{1\times NQ}. \tag{21}$$

Proof: *In optimization theory [7], a stationary point is a defined as point where the derivatives with respect to all the independent variables vanish. Since* $\mathrm{Re}\{\boldsymbol{Z}\} = \boldsymbol{X}$ *and* $\mathrm{Im}\{\boldsymbol{Z}\} = \boldsymbol{Y}$ *contain only independent variables, (*19*)*

[7]Notice that a stationary point can be a local minimum, a local maximum, or a saddle point.

[8]In (19), the symbol $\wedge$ means that both of the equations stated in (19) must be satisfied at the same time.

gives a stationary point by definition. By using the chain rule in Theorem 1*, on both sides of* $f(\boldsymbol{Z}, \boldsymbol{Z}^*) = g(\boldsymbol{X}, \boldsymbol{Y})$ *and taking the derivative with respect to* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$*, the following two equations are obtained:*

$$(\mathcal{D}_{\boldsymbol{Z}} f)\,(\mathcal{D}_{\boldsymbol{X}} \boldsymbol{Z}) + (\mathcal{D}_{\boldsymbol{Z}^*} f)\,(\mathcal{D}_{\boldsymbol{X}} \boldsymbol{Z}^*) = \mathcal{D}_{\boldsymbol{X}} g, \tag{22}$$

$$(\mathcal{D}_{\boldsymbol{Z}} f)\,(\mathcal{D}_{\boldsymbol{Y}} \boldsymbol{Z}) + (\mathcal{D}_{\boldsymbol{Z}^*} f)\,(\mathcal{D}_{\boldsymbol{Y}} \boldsymbol{Z}^*) = \mathcal{D}_{\boldsymbol{Y}} g. \tag{23}$$

From Table II*, it follows that* $\mathcal{D}_{\boldsymbol{X}} \boldsymbol{Z} = \mathcal{D}_{\boldsymbol{X}} \boldsymbol{Z}^* = \boldsymbol{I}_{NQ}$ *and* $\mathcal{D}_{\boldsymbol{Y}} \boldsymbol{Z} = -\mathcal{D}_{\boldsymbol{Y}} \boldsymbol{Z}^* = \jmath \boldsymbol{I}_{NQ}$*. If these results are inserted into* (22) *and* (23)*, these two equations can be formulated into block matrix form in the following way:*

$$\begin{bmatrix} \mathcal{D}_{\boldsymbol{X}} g \\ \mathcal{D}_{\boldsymbol{Y}} g \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \jmath & -\jmath \end{bmatrix} \begin{bmatrix} \mathcal{D}_{\boldsymbol{Z}} f \\ \mathcal{D}_{\boldsymbol{Z}^*} f \end{bmatrix}. \tag{24}$$

This equation is equivalent to the following matrix equation:

$$\begin{bmatrix} \mathcal{D}_{\boldsymbol{Z}} f \\ \mathcal{D}_{\boldsymbol{Z}^*} f \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\jmath}{2} \\ \frac{1}{2} & \frac{\jmath}{2} \end{bmatrix} \begin{bmatrix} \mathcal{D}_{\boldsymbol{X}} g \\ \mathcal{D}_{\boldsymbol{Y}} g \end{bmatrix}. \tag{25}$$

Since $\mathcal{D}_{\boldsymbol{X}} g \in \mathbb{R}^{1\times NQ}$ *and* $\mathcal{D}_{\boldsymbol{Y}} g \in \mathbb{R}^{1\times NQ}$*, it is seen from* (25)*, that* (19)*,* (20)*, and* (21) *are equivalent.* ■

(24) and (25) are multi-variable generalizations of (2) and (3).

3) **Direction of Extremal Rate of Change:** The next theorem states how to find the maximum and minimum rate of change of $f(\boldsymbol{Z}, \boldsymbol{Z}^*) \in \mathbb{R}$.

Theorem 3: *Let* $f : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{R}$*. The directions where the function* f *have the maximum and minimum rate of change with respect to* $\mathrm{vec}(\boldsymbol{Z})$ *are given by* $[\mathcal{D}_{\boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*)]^T$ *and* $-[\mathcal{D}_{\boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*)]^T$*, respectively.*

Proof: *Since* $f \in \mathbb{R}$*,* df *can be written in the following two ways* $df = (\mathcal{D}_{\boldsymbol{Z}} f)\, d\,\mathrm{vec}(\boldsymbol{Z}) + (\mathcal{D}_{\boldsymbol{Z}^*} f)\, d\,\mathrm{vec}(\boldsymbol{Z}^*)$*, and* $df = df^* = (\mathcal{D}_{\boldsymbol{Z}} f)^*\, d\,\mathrm{vec}(\boldsymbol{Z}^*) + (\mathcal{D}_{\boldsymbol{Z}^*} f)^*\, d\,\mathrm{vec}(\boldsymbol{Z})$*, where* $df = df^*$ *since* $f \in \mathbb{R}$*. Subtracting the two different expressions of* df *from each other and then applying Lemma* 1*, gives that* $\mathcal{D}_{\boldsymbol{Z}^*} f = (\mathcal{D}_{\boldsymbol{Z}} f)^*$ *and* $\mathcal{D}_{\boldsymbol{Z}} f = (\mathcal{D}_{\boldsymbol{Z}^*} f)^*$*. From these results, it follows that:* $df = (\mathcal{D}_{\boldsymbol{Z}} f)\, d\,\mathrm{vec}(\boldsymbol{Z}) + ((\mathcal{D}_{\boldsymbol{Z}} f)\, d\,\mathrm{vec}(\boldsymbol{Z}))^* = 2\,\mathrm{Re}\left\{(\mathcal{D}_{\boldsymbol{Z}} f)\, d\,\mathrm{vec}(\boldsymbol{Z})\right\} = 2\,\mathrm{Re}\left\{(\mathcal{D}_{\boldsymbol{Z}^*} f)^*\, d\,\mathrm{vec}(\boldsymbol{Z})\right\}$*. Let* $\boldsymbol{a}_i \in \mathbb{C}^{K\times 1}$*, where* $i \in \{0, 1\}$*. Then*

$$\mathrm{Re}\left\{\boldsymbol{a}_0^H \boldsymbol{a}_1\right\} = \left\langle \begin{bmatrix} \mathrm{Re}\left\{\boldsymbol{a}_0\right\} \\ \mathrm{Im}\left\{\boldsymbol{a}_0\right\} \end{bmatrix}, \begin{bmatrix} \mathrm{Re}\left\{\boldsymbol{a}_1\right\} \\ \mathrm{Im}\left\{\boldsymbol{a}_1\right\} \end{bmatrix} \right\rangle, \tag{26}$$

where $\langle \cdot, \cdot \rangle$ is the ordinary Euclidean inner product [25] between real vectors in $\mathbb{R}^{2K\times 1}$. Using this

$$df = 2 \left\langle \begin{bmatrix} \mathrm{Re}\left\{ (\mathcal{D}_{\boldsymbol{Z}^*} f)^T \right\} \\ \mathrm{Im}\left\{ (\mathcal{D}_{\boldsymbol{Z}^*} f)^T \right\} \end{bmatrix}, \begin{bmatrix} \mathrm{Re}\left\{ d\,\mathrm{vec}(\boldsymbol{Z}) \right\} \\ \mathrm{Im}\left\{ d\,\mathrm{vec}(\boldsymbol{Z}) \right\} \end{bmatrix} \right\rangle. \tag{27}$$

Cauchy-Schwartz inequality gives that the maximum value of df occurs when $d\,\mathrm{vec}(\boldsymbol{Z}) = \alpha\,(\mathcal{D}_{\boldsymbol{Z}^*} f)^T$ for $\alpha > 0$ and from this, it follows that the minimum rate of change occurs when $d\,\mathrm{vec}(\boldsymbol{Z}) = -\beta\,(\mathcal{D}_{\boldsymbol{Z}^*} f)^T$, for $\beta > 0$. ∎

It is now explained why $[\mathcal{D}_{\boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*)]^T$ it *not* the direction of maximum change of rate of the function f. Let $g : \mathbb{C}^{K\times 1} \times \mathbb{C}^{K\times 1} \to \mathbb{R}$ be $g(\boldsymbol{a}_0, \boldsymbol{a}_1) = 2\,\mathrm{Re}\left\{ \boldsymbol{a}_0^T \boldsymbol{a}_1 \right\}$. If $K = 2$ and $\boldsymbol{a}_0 = [1, \jmath]^T$, then $g(\boldsymbol{a}_0, \boldsymbol{a}_0) = 0$ despite the fact that $[1, \jmath]^T \neq \mathbf{0}_{2\times 1}$. Therefore, the function g is *not* an inner product and Cauchy-Schwartz inequality is not valid for this function. By examining the proof of Theorem 3, it is seen that the reason why $[\mathcal{D}_{\boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*)]^T$ it *not* the direction of maximum change of rate, is that the function g is *not* an inner product.

4) **Steepest Descent Method:** If a real-valued function f is being optimized with respect to the parameter $\boldsymbol{Z}$ by means of the steepest decent method, it follows from Theorem 3 that the updating term must be proportional to $\mathcal{D}_{\boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*)$ and *not* $\mathcal{D}_{\boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*)$. The update equation for optimizing the real-valued function in Theorem 3 by means of the steepest decent method, can be expressed as $\mathrm{vec}^T(\boldsymbol{Z}_{k+1}) = \mathrm{vec}^T(\boldsymbol{Z}_k) + \mu \mathcal{D}_{\boldsymbol{Z}^*} f(\boldsymbol{Z}_k, \boldsymbol{Z}_k^*)$, where μ is a *real positive constant if it is a maximization problem* or a *real negative constant if it is a minimization problem*, and $\boldsymbol{Z}_k \in \mathbb{C}^{N\times Q}$ is the value of the matrix after k iterations. The size of $\mathrm{vec}^T(\boldsymbol{Z}_k)$ and $\mathcal{D}_{\boldsymbol{Z}^*} f(\boldsymbol{Z}_k, \boldsymbol{Z}_k^*)$ is $1 \times NQ$.

Example 2: *Let $f : \mathbb{C} \times \mathbb{C} \to \mathbb{R}$ be $f(z, z^*) = |z|^2 = zz^*$. The partial derivatives are $\mathcal{D}_{z^*} f(z, z^*) = z$ and $\mathcal{D}_z f(z, z^*) = z^*$. This result is in accordance with the intuitive picture for optimizing f, i.e., it is more efficient to maximize f in the direction of $\mathcal{D}_{z^*} f(z, z^*) = z$ than in the direction of $\mathcal{D}_z f(z, z^*) = z^*$ when standing at z.*

5) **Complex Case versus Real Case:** In order to show that the consistence with the real-valued case given in [7], the following theorem shows how the complex-valued case is an extension of [7].

Theorem 4: *Let $\boldsymbol{F} : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}^{M\times P}$ and $\boldsymbol{G} : \mathbb{C}^{N\times Q} \to \mathbb{C}^{M\times P}$. If $\boldsymbol{F}(\boldsymbol{Z}_0, \boldsymbol{Z}_1) = \boldsymbol{G}(\boldsymbol{Z}_0)$, then $\mathcal{D}_{\boldsymbol{Z}} \boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \mathcal{D}_{\boldsymbol{Z}} \boldsymbol{G}(\boldsymbol{Z})$ can be obtained by the procedure given in [7] for finding the derivative of the function $\boldsymbol{G}$ and $\mathcal{D}_{\boldsymbol{Z}^*} \boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \mathbf{0}_{MP\times NQ}$.*

If $\boldsymbol{F}(\boldsymbol{Z}_0, \boldsymbol{Z}_1) = \boldsymbol{G}(\boldsymbol{Z}_1)$, then $\mathcal{D}_{\boldsymbol{Z}} \boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \mathbf{0}_{MP\times NQ}$, and $\mathcal{D}_{\boldsymbol{Z}^*} \boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \mathcal{D}_{\boldsymbol{Z}} \boldsymbol{G}(\boldsymbol{Z})|_{\boldsymbol{Z}=\boldsymbol{Z}^*}$ can be obtained by the procedure given in [7].

TABLE IV
DERIVATIVES OF FUNCTIONS OF THE TYPE $f(\boldsymbol{z}, \boldsymbol{z}^*)$

$f(\boldsymbol{z}, \boldsymbol{z}^*)$	Differential df	$\mathcal{D}_{\boldsymbol{z}} f(\boldsymbol{z}, \boldsymbol{z}^*)$	$\mathcal{D}_{\boldsymbol{z}^*} f(\boldsymbol{z}, \boldsymbol{z}^*)$
$\boldsymbol{a}^T\boldsymbol{z} = \boldsymbol{z}^T\boldsymbol{a}$	$\boldsymbol{a}^T d\boldsymbol{z}$	$\boldsymbol{a}^T$	$\boldsymbol{0}_{1\times N}$
$\boldsymbol{a}^T\boldsymbol{z}^* = \boldsymbol{z}^H\boldsymbol{a}$	$\boldsymbol{a}^T d\boldsymbol{z}^*$	$\boldsymbol{0}_{1\times N}$	$\boldsymbol{a}^T$
$\boldsymbol{z}^T\boldsymbol{A}\boldsymbol{z}$	$\boldsymbol{z}^T\left(\boldsymbol{A}+\boldsymbol{A}^T\right)d\boldsymbol{z}$	$\boldsymbol{z}^T\left(\boldsymbol{A}+\boldsymbol{A}^T\right)$	$\boldsymbol{0}_{1\times N}$
$\boldsymbol{z}^H\boldsymbol{A}\boldsymbol{z}$	$\boldsymbol{z}^H\boldsymbol{A}d\boldsymbol{z} + \boldsymbol{z}^T\boldsymbol{A}^T d\boldsymbol{z}^*$	$\boldsymbol{z}^H\boldsymbol{A}$	$\boldsymbol{z}^T\boldsymbol{A}^T$
$\boldsymbol{z}^H\boldsymbol{A}\boldsymbol{z}^*$	$\boldsymbol{z}^H\left(\boldsymbol{A}+\boldsymbol{A}^T\right)d\boldsymbol{z}^*$	$\boldsymbol{0}_{1\times N}$	$\boldsymbol{z}^H\left(\boldsymbol{A}+\boldsymbol{A}^T\right)$

Proof: *Assume that* $\boldsymbol{F}(\boldsymbol{Z}_0, \boldsymbol{Z}_1) = \boldsymbol{G}(\boldsymbol{Z}_0)$ *where* $\boldsymbol{Z}_0$ *and* $\boldsymbol{Z}_1$ *have independent differentials. Applying the* $\mathrm{vec}(\cdot)$ *and the differential operator to this equation, leads to:* $d\,\mathrm{vec}(\boldsymbol{F}) = (\mathcal{D}_{\boldsymbol{Z}_0}\boldsymbol{F})\, d\,\mathrm{vec}(\boldsymbol{Z}_0) + (\mathcal{D}_{\boldsymbol{Z}_1}\boldsymbol{F})\, d\,\mathrm{vec}(\boldsymbol{Z}_1) = (\mathcal{D}_{\boldsymbol{Z}_0}\boldsymbol{G})\, d\,\mathrm{vec}(\boldsymbol{Z}_0)$. *By setting* $\boldsymbol{Z}_0 = \boldsymbol{Z}$ *and* $\boldsymbol{Z}_1 = \boldsymbol{Z}^*$ *in Lemma* 1, *it is seen that* $\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F} = \boldsymbol{0}_{MP\times NQ}$ *and* $\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \mathcal{D}_{\boldsymbol{Z}}\boldsymbol{G}(\boldsymbol{Z})$. *Since the last equation only has one matrix variable* $\boldsymbol{Z}$, *the same techniques as given in [7] can be used. The first part of the theorem is then proved, and the second part can be shown similarly.* ■

VI. DEVELOPMENT OF DERIVATIVE FORMULAS

A. Derivative of $f(z, z^*)$

The case of scalar function of scalar variables is treated in the literature for one input variable, see for example [5], [26]. If the variables z and z^* are treated as independent variables, the derivatives $\mathcal{D}_z f(z, z^*)$ and $\mathcal{D}_{z^*} f(z, z^*)$ can be found as for scalar functions having two independent variables, see Example 1.

B. Derivative of $f(\boldsymbol{z}, \boldsymbol{z}^*)$

Let $f : \mathbb{C}^{N\times 1} \times \mathbb{C}^{N\times 1} \to \mathbb{C}$ be $f(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{z}^H\boldsymbol{A}\boldsymbol{z}$. This type of function is frequently used in the context of filter optimization and in array signal processing [20]. For instance, a frequently occurring example is that of output power $\boldsymbol{z}^H\boldsymbol{A}\boldsymbol{z}$ optimization where $\boldsymbol{z}$ is the filter coefficients and $\boldsymbol{A}$ is the input covariance matrix. The differential of f is $df = (d\boldsymbol{z}^H)\boldsymbol{A}\boldsymbol{z} + \boldsymbol{z}^H\boldsymbol{A}d\boldsymbol{z} = \boldsymbol{z}^H\boldsymbol{A}d\boldsymbol{z} + \boldsymbol{z}^T\boldsymbol{A}^T d\boldsymbol{z}^*$, where Table II was used. From df, the derivatives of $\boldsymbol{z}^H\boldsymbol{A}\boldsymbol{z}$ with respect to $\boldsymbol{z}$ and $\boldsymbol{z}^*$ follow from Table III, and are $\mathcal{D}_{\boldsymbol{z}} f(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{z}^H\boldsymbol{A}$ and $\mathcal{D}_{\boldsymbol{z}^*} f(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{z}^T\boldsymbol{A}^T$, see Table IV. The other selected examples are given in Table IV and they can be derived similarly. Some of the results included in Table IV can be found in [15].

C. Derivative of $f(\boldsymbol{Z}, \boldsymbol{Z}^*)$

For functions of the type $f(\boldsymbol{Z}, \boldsymbol{Z}^*)$, it is common to arrange the partial derivatives $\frac{\partial}{\partial z_{k,l}} f$ and $\frac{\partial}{\partial z^*_{k,l}} f$ in an alternative way [7] than in the expressions $\mathcal{D}_{\boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*)$ and $\mathcal{D}_{\boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*)$. The notation for the alternative way of organizing all the partial derivatives is $\frac{\partial}{\partial \boldsymbol{Z}} f$ and $\frac{\partial}{\partial \boldsymbol{Z}^*} f$. In this alternative way, the partial derivatives of the elements of the matrix $\boldsymbol{Z} \in \mathbb{C}^{N \times Q}$ are arranged as:

$$\frac{\partial}{\partial \boldsymbol{Z}} f = \begin{bmatrix} \frac{\partial}{\partial z_{0,0}} f & \cdots & \frac{\partial}{\partial z_{0,Q-1}} f \\ \vdots & & \vdots \\ \frac{\partial}{\partial z_{N-1,0}} f & \cdots & \frac{\partial}{\partial z_{N-1,Q-1}} f \end{bmatrix}, \quad \frac{\partial}{\partial \boldsymbol{Z}^*} f = \begin{bmatrix} \frac{\partial}{\partial z^*_{0,0}} f & \cdots & \frac{\partial}{\partial z^*_{0,Q-1}} f \\ \vdots & & \vdots \\ \frac{\partial}{\partial z^*_{N-1,0}} f & \cdots & \frac{\partial}{\partial z^*_{N-1,Q-1}} f \end{bmatrix}. \tag{28}$$

$\frac{\partial}{\partial \boldsymbol{Z}} f$ and $\frac{\partial}{\partial \boldsymbol{Z}^*} f$ are called the *gradient* of f with respect to $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$, respectively. (28) generalizes to the complex case of one of the ways to define the derivative of real scalar functions with respect to real matrices in [7]. The way of arranging the partial derivatives in (28) is different than than the way given in (12) and (13). If $df = \operatorname{vec}^T(\boldsymbol{A}_0) d \operatorname{vec}(\boldsymbol{Z}) + \operatorname{vec}^T(\boldsymbol{A}_1) d \operatorname{vec}(\boldsymbol{Z}^*) = \operatorname{Tr}\left\{\boldsymbol{A}_0^T d\boldsymbol{Z} + \boldsymbol{A}_1^T d\boldsymbol{Z}^*\right\}$, where $\boldsymbol{A}_i, \boldsymbol{Z} \in \mathbb{C}^{N \times Q}$, then it can be shown that $\frac{\partial}{\partial \boldsymbol{Z}} f = \boldsymbol{A}_0$ and $\frac{\partial}{\partial \boldsymbol{Z}^*} f = \boldsymbol{A}_1$, where the matrices $\boldsymbol{A}_0$ and $\boldsymbol{A}_1$ depend on $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ in general. This result is included in Table III. The size of $\frac{\partial}{\partial \boldsymbol{Z}} f$ and $\frac{\partial}{\partial \boldsymbol{Z}^*} f$ is $N \times Q$, while the size of $\mathcal{D}_{\boldsymbol{Z}} f\left(\boldsymbol{Z}, \boldsymbol{Z}^*\right)$ and $\mathcal{D}_{\boldsymbol{Z}^*} f\left(\boldsymbol{Z}, \boldsymbol{Z}^*\right)$ is $1 \times NQ$, so these two ways of organizing the partial derivatives are different. It can be shown, that $\mathcal{D}_{\boldsymbol{Z}} f\left(\boldsymbol{Z}, \boldsymbol{Z}^*\right) = \operatorname{vec}^T\left(\frac{\partial}{\partial \boldsymbol{Z}} f\left(\boldsymbol{Z}, \boldsymbol{Z}^*\right)\right)$, and $\mathcal{D}_{\boldsymbol{Z}^*} f\left(\boldsymbol{Z}, \boldsymbol{Z}^*\right) = \operatorname{vec}^T\left(\frac{\partial}{\partial \boldsymbol{Z}^*} f\left(\boldsymbol{Z}, \boldsymbol{Z}^*\right)\right)$. The steepest decent method can be formulated as $\boldsymbol{Z}_{k+1} = \boldsymbol{Z}_k + \mu \frac{\partial}{\partial \boldsymbol{Z}^*} f(\boldsymbol{Z}_k, \boldsymbol{Z}_k^*)$. Key examples of derivatives are developed in the text below, while others useful results are stated in Table V.

1) **Determinant Related Problems***:* Objective functions that depend on the determinant appear in several parts of signal processing related problems, e.g., in the capacity of wireless multiple-input multiple-output (MIMO) communication systems [19], in an upper bound for the pair-wise error probability [18], [27], and for the exact symbol error rate (SER) for precoded orthogonal space-time block-code (OSTBC) systems [28].

Let $f : \mathbb{C}^{N \times Q} \times \mathbb{C}^{Q \times N} \to \mathbb{C}$ be $f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \det(\boldsymbol{P} + \boldsymbol{Z}\boldsymbol{Z}^H)$, where $\boldsymbol{P} \in \mathbb{C}^{N \times N}$ is independent of $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$. df is found by the rules in Table II as $df = \det(\boldsymbol{P} + \boldsymbol{Z}\boldsymbol{Z}^H) \operatorname{Tr}\left\{(\boldsymbol{P} + \boldsymbol{Z}\boldsymbol{Z}^H)^{-1} d(\boldsymbol{Z}\boldsymbol{Z}^H)\right\} = \det(\boldsymbol{P} + \boldsymbol{Z}\boldsymbol{Z}^H) \operatorname{Tr}\left\{\boldsymbol{Z}^H \left(\boldsymbol{P} + \boldsymbol{Z}\boldsymbol{Z}^H\right)^{-1} d\boldsymbol{Z} + \boldsymbol{Z}^T (\boldsymbol{P} + \boldsymbol{Z}\boldsymbol{Z}^H)^{-T} d\boldsymbol{Z}^*\right\}$. From this, the derivatives with respect to $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ of $\det\left(\boldsymbol{P} + \boldsymbol{Z}\boldsymbol{Z}^H\right)$ can be found, but since these results are relatively long, they are *not* included in Table V.

TABLE V
DERIVATIVES OF FUNCTIONS OF THE TYPE $f(\boldsymbol{Z},\boldsymbol{Z}^*)$

$f(\boldsymbol{Z},\boldsymbol{Z}^*)$	Differential df	$\frac{\partial}{\partial \boldsymbol{Z}} f$	$\frac{\partial}{\partial \boldsymbol{Z}^*} f$
$\mathrm{Tr}\{\boldsymbol{Z}\}$	$\mathrm{Tr}\{\boldsymbol{I}_N d\boldsymbol{Z}\}$	$\boldsymbol{I}_N$	$\boldsymbol{0}_{N\times N}$
$\mathrm{Tr}\{\boldsymbol{Z}^*\}$	$\mathrm{Tr}\{\boldsymbol{I}_N d\boldsymbol{Z}^*\}$	$\boldsymbol{0}_{N\times N}$	$\boldsymbol{I}_N$
$\mathrm{Tr}\{\boldsymbol{A}\boldsymbol{Z}\}$	$\mathrm{Tr}\{\boldsymbol{A}d\boldsymbol{Z}\}$	$\boldsymbol{A}^T$	$\boldsymbol{0}_{N\times Q}$
$\mathrm{Tr}\{\boldsymbol{Z}\boldsymbol{A}_0\boldsymbol{Z}^T\boldsymbol{A}_1\}$	$\mathrm{Tr}\left\{\left(\boldsymbol{A}_0\boldsymbol{Z}^T\boldsymbol{A}_1+\boldsymbol{A}_0^T\boldsymbol{Z}^T\boldsymbol{A}_1^T\right)d\boldsymbol{Z}\right\}$	$\boldsymbol{A}_1^T\boldsymbol{Z}\boldsymbol{A}_0^T+\boldsymbol{A}_1\boldsymbol{Z}\boldsymbol{A}_0$	$\boldsymbol{0}_{N\times Q}$
$\mathrm{Tr}\{\boldsymbol{Z}\boldsymbol{A}_0\boldsymbol{Z}\boldsymbol{A}_1\}$	$\mathrm{Tr}\{(\boldsymbol{A}_0\boldsymbol{Z}\boldsymbol{A}_1+\boldsymbol{A}_1\boldsymbol{Z}\boldsymbol{A}_0)d\boldsymbol{Z}\}$	$\boldsymbol{A}_1^T\boldsymbol{Z}^T\boldsymbol{A}_0^T+\boldsymbol{A}_0^T\boldsymbol{Z}^T\boldsymbol{A}_1^T$	$\boldsymbol{0}_{N\times Q}$
$\mathrm{Tr}\{\boldsymbol{Z}\boldsymbol{A}_0\boldsymbol{Z}^H\boldsymbol{A}_1\}$	$\mathrm{Tr}\left\{\boldsymbol{A}_0\boldsymbol{Z}^H\boldsymbol{A}_1 d\boldsymbol{Z}+\boldsymbol{A}_0^T\boldsymbol{Z}^T\boldsymbol{A}_1^T d\boldsymbol{Z}^*\right\}$	$\boldsymbol{A}_1^T\boldsymbol{Z}^*\boldsymbol{A}_0^T$	$\boldsymbol{A}_1\boldsymbol{Z}\boldsymbol{A}_0$
$\mathrm{Tr}\{\boldsymbol{Z}\boldsymbol{A}_0\boldsymbol{Z}^*\boldsymbol{A}_1\}$	$\mathrm{Tr}\{\boldsymbol{A}_0\boldsymbol{Z}^*\boldsymbol{A}_1 d\boldsymbol{Z}+\boldsymbol{A}_1\boldsymbol{Z}\boldsymbol{A}_0 d\boldsymbol{Z}^*\}$	$\boldsymbol{A}_1^T\boldsymbol{Z}^H\boldsymbol{A}_0^T$	$\boldsymbol{A}_0^T\boldsymbol{Z}^T\boldsymbol{A}_1^T$
$\mathrm{Tr}\{\boldsymbol{A}\boldsymbol{Z}^{-1}\}$	$-\mathrm{Tr}\left\{\boldsymbol{Z}^{-1}\boldsymbol{A}\boldsymbol{Z}^{-1}d\boldsymbol{Z}\right\}$	$-\left(\boldsymbol{Z}^T\right)^{-1}\boldsymbol{A}^T\left(\boldsymbol{Z}^T\right)^{-1}$	$\boldsymbol{0}_{N\times N}$
$\mathrm{Tr}\{\boldsymbol{Z}^p\}$	$p\,\mathrm{Tr}\left\{\boldsymbol{Z}^{p-1}d\boldsymbol{Z}\right\}$	$p\left(\boldsymbol{Z}^T\right)^{p-1}$	$\boldsymbol{0}_{N\times N}$
$\det(\boldsymbol{Z})$	$\det(\boldsymbol{Z})\,\mathrm{Tr}\left\{\boldsymbol{Z}^{-1}d\boldsymbol{Z}\right\}$	$\det(\boldsymbol{Z})\left(\boldsymbol{Z}^T\right)^{-1}$	$\boldsymbol{0}_{N\times N}$
$\det(\boldsymbol{A}_0\boldsymbol{Z}\boldsymbol{A}_1)$	$\det(\boldsymbol{A}_0\boldsymbol{Z}\boldsymbol{A}_1)\,\mathrm{Tr}\left\{\boldsymbol{A}_1\left(\boldsymbol{A}_0\boldsymbol{Z}\boldsymbol{A}_1\right)^{-1}\boldsymbol{A}_0 d\boldsymbol{Z}\right\}$	$\det(\boldsymbol{A}_0\boldsymbol{Z}\boldsymbol{A}_1)\boldsymbol{A}_0^T\left(\boldsymbol{A}_1^T\boldsymbol{Z}^T\boldsymbol{A}_0^T\right)^{-1}\boldsymbol{A}_1^T$	$\boldsymbol{0}_{N\times Q}$
$\det(\boldsymbol{Z}^2)$	$2\det^2(\boldsymbol{Z})\,\mathrm{Tr}\left\{\boldsymbol{Z}^{-1}d\boldsymbol{Z}\right\}$	$2\det^2(\boldsymbol{Z})\left(\boldsymbol{Z}^T\right)^{-1}$	$\boldsymbol{0}_{N\times N}$
$\det(\boldsymbol{Z}\boldsymbol{Z}^T)$	$2\det(\boldsymbol{Z}\boldsymbol{Z}^T)\,\mathrm{Tr}\left\{\boldsymbol{Z}^T\left(\boldsymbol{Z}\boldsymbol{Z}^T\right)^{-1}d\boldsymbol{Z}\right\}$	$2\det(\boldsymbol{Z}\boldsymbol{Z}^T)\left(\boldsymbol{Z}\boldsymbol{Z}^T\right)^{-1}\boldsymbol{Z}$	$\boldsymbol{0}_{N\times Q}$
$\det(\boldsymbol{Z}\boldsymbol{Z}^*)$	$\det(\boldsymbol{Z}\boldsymbol{Z}^*)\,\mathrm{Tr}\left\{\boldsymbol{Z}^*(\boldsymbol{Z}\boldsymbol{Z}^*)^{-1}d\boldsymbol{Z}+(\boldsymbol{Z}\boldsymbol{Z}^*)^{-1}\boldsymbol{Z}d\boldsymbol{Z}^*\right\}$	$\det(\boldsymbol{Z}\boldsymbol{Z}^*)(\boldsymbol{Z}^H\boldsymbol{Z}^T)^{-1}\boldsymbol{Z}^H$	$\det(\boldsymbol{Z}\boldsymbol{Z}^*)\boldsymbol{Z}^T\left(\boldsymbol{Z}^H\boldsymbol{Z}^T\right)^{-1}$
$\det(\boldsymbol{Z}\boldsymbol{Z}^H)$	$\det(\boldsymbol{Z}\boldsymbol{Z}^H)\,\mathrm{Tr}\left\{\boldsymbol{Z}^H(\boldsymbol{Z}\boldsymbol{Z}^H)^{-1}d\boldsymbol{Z}+\boldsymbol{Z}^T\left(\boldsymbol{Z}^*\boldsymbol{Z}^T\right)^{-1}d\boldsymbol{Z}^*\right\}$	$\det(\boldsymbol{Z}\boldsymbol{Z}^H)(\boldsymbol{Z}^*\boldsymbol{Z}^T)^{-1}\boldsymbol{Z}^*$	$\det(\boldsymbol{Z}\boldsymbol{Z}^H)\left(\boldsymbol{Z}\boldsymbol{Z}^H\right)^{-1}\boldsymbol{Z}$
$\det(\boldsymbol{Z}^p)$	$p\det^p(\boldsymbol{Z})\,\mathrm{Tr}\left\{\boldsymbol{Z}^{-1}d\boldsymbol{Z}\right\}$	$p\det^p(\boldsymbol{Z})\left(\boldsymbol{Z}^T\right)^{-1}$	$\boldsymbol{0}_{N\times N}$
$\lambda(\boldsymbol{Z})$	$\frac{\boldsymbol{v}_0^H(d\boldsymbol{Z})\boldsymbol{u}_0}{\boldsymbol{v}_0^H\boldsymbol{u}_0}=\mathrm{Tr}\left\{\frac{\boldsymbol{u}_0\boldsymbol{v}_0^H}{\boldsymbol{v}_0^H\boldsymbol{u}_0}d\boldsymbol{Z}\right\}$	$\frac{\boldsymbol{v}_0^*\boldsymbol{u}_0^T}{\boldsymbol{v}_0^H\boldsymbol{u}_0}$	$\boldsymbol{0}_{N\times N}$
$\lambda^*(\boldsymbol{Z})$	$\frac{\boldsymbol{v}_0^T(d\boldsymbol{Z}^*)\boldsymbol{u}_0^*}{\boldsymbol{v}_0^T\boldsymbol{u}_0^*}=\mathrm{Tr}\left\{\frac{\boldsymbol{u}_0^*\boldsymbol{v}_0^T}{\boldsymbol{v}_0^T\boldsymbol{u}_0^*}d\boldsymbol{Z}^*\right\}$	$\boldsymbol{0}_{N\times N}$	$\frac{\boldsymbol{v}_0\boldsymbol{u}_0^H}{\boldsymbol{v}_0^T\boldsymbol{u}_0^*}$

2) **Eigenvalue Related Problems***:* Let λ_0 be a simple eigenvalue[9] of $\boldsymbol{Z}_0 \in \mathbb{C}^{N\times N}$, and let $\boldsymbol{u}_0 \in \mathbb{C}^{N\times 1}$ be the normalized corresponding eigenvector, such that $\boldsymbol{Z}_0\boldsymbol{u}_0 = \lambda_0\boldsymbol{u}_0$. Let $\lambda : \mathbb{C}^{N\times N} \to \mathbb{C}$ and $\boldsymbol{u} : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times 1}$ be defined for all $\boldsymbol{Z}$ in some neighborhood of $\boldsymbol{Z}_0$ such that:

$$\boldsymbol{Z}\boldsymbol{u}(\boldsymbol{Z}) = \lambda(\boldsymbol{Z})\boldsymbol{u}(\boldsymbol{Z}), \quad \boldsymbol{u}_0^H\boldsymbol{u}(\boldsymbol{Z}) = 1, \quad \lambda(\boldsymbol{Z}_0) = \lambda_0, \quad \boldsymbol{u}(\boldsymbol{Z}_0) = \boldsymbol{u}_0. \tag{29}$$

Let the normalized *left* eigenvector of $\boldsymbol{Z}_0$ corresponding to λ_0 be denoted $\boldsymbol{v}_0 \in \mathbb{C}^{N\times 1}$, i.e., $\boldsymbol{v}_0^H\boldsymbol{Z}_0 = \lambda_0\boldsymbol{v}_0^H$, or equivalently $\boldsymbol{Z}_0^H\boldsymbol{v}_0 = \lambda_0^*\boldsymbol{v}_0$. In order to find the differential of $\lambda(\boldsymbol{Z})$ at $\boldsymbol{Z} = \boldsymbol{Z}_0$, take the differential of both sides of the first expression in (29) evaluated at $\boldsymbol{Z} = \boldsymbol{Z}_0$:

$$(d\boldsymbol{Z})\,\boldsymbol{u}_0 + \boldsymbol{Z}_0 d\boldsymbol{u} = (d\lambda)\,\boldsymbol{u}_0 + \lambda_0 d\boldsymbol{u}. \tag{30}$$

Pre-multiplying (30) by $\boldsymbol{v}_0^H$ gives: $\boldsymbol{v}_0^H\,(d\boldsymbol{Z})\,\boldsymbol{u}_0 = (d\lambda)\,\boldsymbol{v}_0^H\boldsymbol{u}_0$. From Lemma 6.3.10 in [22], $\boldsymbol{v}_0^H\boldsymbol{u}_0 \neq 0$, and hence

$$d\lambda = \frac{\boldsymbol{v}_0^H\,(d\boldsymbol{Z})\,\boldsymbol{u}_0}{\boldsymbol{v}_0^H\boldsymbol{u}_0} = \mathrm{Tr}\left\{\frac{\boldsymbol{u}_0\boldsymbol{v}_0^H}{\boldsymbol{v}_0^H\boldsymbol{u}_0}d\boldsymbol{Z}\right\}. \tag{31}$$

[9]The matrix $\boldsymbol{Z}_0 \in \mathbb{C}^{N\times N}$ has in general N different complex eigenvalues. However, the roots of the characteristic equation, i.e., the eigenvalues need not to be distinct. The number of times an eigenvalue appears is equal to its multiplicity. If one eigenvalue appear only once, it is called a simple eigenvalue [22].

This result is included in Table V, and it will be used later when the derivatives of the eigenvector $\boldsymbol{u}$ and the Hessian of λ are found. The differential of λ^* at $\boldsymbol{Z}_0$ is also included in Table V. These results are derived in [7].

D. Derivative of $\boldsymbol{f}(z, z^*)$

In engineering, the objective function is often a real scalar, and this might be a very complicated function in the parameter matrix. To be able to find the derivative in a simple way, the chain rule in Theorem 1 might be very useful, and then functions of the types $\boldsymbol{f}$ and $\boldsymbol{F}$ might occur.

Let $\boldsymbol{f}(z, z^*) = \boldsymbol{a} f(z, z^*)$, then $d\boldsymbol{f} = \boldsymbol{a} df = \boldsymbol{a}\left(\mathcal{D}_z f(z, z^*)\right) dz + \boldsymbol{a}\left(\mathcal{D}_{z^*} f(z, z^*)\right) dz^*$, where df follows from Table III. From this equation, it follows that $\mathcal{D}_z \boldsymbol{f} = \boldsymbol{a}\mathcal{D}_z f(z, z^*)$ and $\mathcal{D}_{z^*} \boldsymbol{f} = \boldsymbol{a}\mathcal{D}_{z^*} f(z, z^*)$. The derivatives of the vector functions $\boldsymbol{a}z$ and $\boldsymbol{a}zz^*$ follow from these expressions.

E. Derivative of $\boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*)$

First a useful results is introduced. In order to extract the $\operatorname{vec}(\cdot)$ of an inner matrix from the $\operatorname{vec}(\cdot)$ of a multiple matrix product, the following formula [23] is very useful:

$$\operatorname{vec}\left(\boldsymbol{A}\boldsymbol{B}\boldsymbol{C}\right) = \left(\boldsymbol{C}^T \otimes \boldsymbol{A}\right) \operatorname{vec}\left(\boldsymbol{B}\right). \tag{32}$$

Let $\boldsymbol{f} : \mathbb{C}^{N\times 1} \times \mathbb{C}^{N\times 1} \to \mathbb{C}^{M\times 1}$ be given by $\boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*)\boldsymbol{a}$, where $\boldsymbol{F} : \mathbb{C}^{N\times 1} \times \mathbb{C}^{N\times 1} \to \mathbb{C}^{M\times P}$. Then $d\boldsymbol{f} = d\operatorname{vec}(\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*)\boldsymbol{a}) = \left(\boldsymbol{a}^T \otimes \boldsymbol{I}_M\right) d\operatorname{vec}(\boldsymbol{F}) = \left(\boldsymbol{a}^T \otimes \boldsymbol{I}_M\right)\left[\left(\mathcal{D}_{\boldsymbol{z}} \boldsymbol{F}\left(\boldsymbol{z}, \boldsymbol{z}^*\right)\right) d\boldsymbol{z} + \left(\mathcal{D}_{\boldsymbol{z}^*} \boldsymbol{F}\left(\boldsymbol{z}, \boldsymbol{z}^*\right)\right) d\boldsymbol{z}^*\right]$, where (32) and Table III were utilized. From $d\boldsymbol{f}$, the derivatives of $\boldsymbol{f}$ with respect to $\boldsymbol{z}$ and $\boldsymbol{z}^*$ follow.

F. Derivative of $\boldsymbol{f}(\boldsymbol{Z}, \boldsymbol{Z}^*)$

1) **Eigenvector Related Problems***:* First a lemma is stated.

Lemma 3: *Let $\boldsymbol{A} \in \mathbb{C}^{N\times Q}$, then the following equalities holds:*

$$\mathcal{R}\left(\boldsymbol{A}\right) = \mathcal{R}\left(\boldsymbol{A}^+\boldsymbol{A}\right), \;\; \mathcal{C}\left(\boldsymbol{A}\right) = \mathcal{C}\left(\boldsymbol{A}\boldsymbol{A}^+\right), \;\; \operatorname{rank}\left(\boldsymbol{A}\right) = \operatorname{rank}\left(\boldsymbol{A}^+\boldsymbol{A}\right) = \operatorname{rank}\left(\boldsymbol{A}\boldsymbol{A}^+\right) \tag{33}$$

where $\mathcal{C}(\boldsymbol{A})$, $\mathcal{R}(\boldsymbol{A})$, and $\mathcal{N}(\boldsymbol{A})$ are the symbols used for the column, row, and null space of $\boldsymbol{A} \in \mathbb{C}^{N\times Q}$, respectively.[10]

The proof of Lemma 3 can be found in Appendix IV.

[10] $\mathcal{C}(\boldsymbol{A}) = \{\boldsymbol{w} \in \mathbb{C}^{N\times 1} | \boldsymbol{w} = \boldsymbol{A}\boldsymbol{z}, \text{ for some } \boldsymbol{z} \in \mathbb{C}^{Q\times 1}\}$, $\mathcal{R}(\boldsymbol{A}) = \{\boldsymbol{w} \in \mathbb{C}^{1\times Q} | \boldsymbol{w} = \boldsymbol{z}\boldsymbol{A}, \text{ for some } \boldsymbol{z} \in \mathbb{C}^{1\times N}\}$, and $\mathcal{N}(\boldsymbol{A}) = \{\boldsymbol{z} \in \mathbb{C}^{Q\times 1} | \boldsymbol{A}\boldsymbol{z} = \boldsymbol{0}_{N\times 1}\}$.

The differential of the eigenvector $\boldsymbol{u}(\boldsymbol{Z})$ is now found at $\boldsymbol{Z} = \boldsymbol{Z}_0$. The derivation here is similar to the one in [7], where the same result for $d\boldsymbol{u}$ at $\boldsymbol{Z} = \boldsymbol{Z}_0$ was derived, however, more details are included here. Let $\boldsymbol{Y}_0 = \lambda_0 \boldsymbol{I}_N - \boldsymbol{Z}_0$, then it follows from (30), that $\boldsymbol{Y}_0 d\boldsymbol{u} = (d\boldsymbol{Z})\,\boldsymbol{u}_0 - (d\lambda)\,\boldsymbol{u}_0 = (d\boldsymbol{Z})\,\boldsymbol{u}_0 - \frac{\boldsymbol{v}_0^H (d\boldsymbol{Z}) \boldsymbol{u}_0}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \boldsymbol{u}_0 = \left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0 \boldsymbol{v}_0^H}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \right) (d\boldsymbol{Z})\,\boldsymbol{u}_0$, where (31) was utilized. Pre-multiplying $\boldsymbol{Y}_0 d\boldsymbol{u}$ with $\boldsymbol{Y}_0^+$ results in:

$$\boldsymbol{Y}_0^+ \boldsymbol{Y}_0 d\boldsymbol{u} = \boldsymbol{Y}_0^+ \left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0 \boldsymbol{v}_0^H}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \right) (d\boldsymbol{Z})\,\boldsymbol{u}_0. \tag{34}$$

The operator $\dim(\cdot)$ returns the dimension of the vector space. Since λ_0 is a simple eigenvalue, $\dim(\mathcal{N}(\boldsymbol{Y}_0)) = 1$ [5]. Hence, it follows from 0.4.4 (g) on page 13 in [7] that $\text{rank}(\boldsymbol{Y}_0) = N - \dim(\mathcal{N}(\boldsymbol{Y}_0)) = N - 1$. From $\boldsymbol{Y}_0 \boldsymbol{u}_0 = \boldsymbol{0}_{N\times 1}$, it follows from (71), see Appendix V, that $\boldsymbol{u}_0^+ \boldsymbol{Y}_0^+ = \boldsymbol{0}_{1\times N}$. It can be shown by direct insertion in Definition 3 of the Moore-Penrose inverse that $\boldsymbol{u}_0^+ = \boldsymbol{u}_0^H$. From these results it follows that $\boldsymbol{u}_0^H \boldsymbol{Y}_0^+ = \boldsymbol{0}_{1\times N}$. Set $\boldsymbol{C}_0 = \boldsymbol{Y}_0^+ \boldsymbol{Y}_0 + \boldsymbol{u}_0 \boldsymbol{u}_0^H$, then it can be shown from the two facts $\boldsymbol{u}_0^H \boldsymbol{Y}_0^+ = \boldsymbol{0}_{1\times N}$ and $\boldsymbol{Y}_0 \boldsymbol{u}_0 = \boldsymbol{0}_{N\times 1}$ that $\boldsymbol{C}_0^2 = \boldsymbol{C}_0$, i.e., $\boldsymbol{C}_0$ is idempotent. It can be shown by the direct use of Definition 3, that the matrix $\boldsymbol{Y}_0^+ \boldsymbol{Y}_0$ is also idempotent. By the use of Lemma 10.1.1 and Corollary 10.2.2, in [8], it is found that $\text{rank}(\boldsymbol{C}_0) = \text{Tr}\{\boldsymbol{C}_0\} = \text{Tr}\left\{\boldsymbol{Y}_0^+ \boldsymbol{Y}_0\right\} + \text{Tr}\left\{\boldsymbol{u}_0 \boldsymbol{u}_0^H\right\} = \text{rank}\left(\boldsymbol{Y}_0^+ \boldsymbol{Y}_0\right) + 1 = \text{rank}(\boldsymbol{Y}_0) + 1 = N$, where the third equation of (33) was used. From Lemma 10.1.1 and Corollary 10.2.2, in [8], and $\text{rank}(\boldsymbol{C}_0) = N$, it follows that $\boldsymbol{C}_0 = \boldsymbol{I}_N$. Using the differential operator on both sides of the second equation in (29), yields $\boldsymbol{u}_0^H d\boldsymbol{u} = 0$. Using these results, it follows that:

$$\boldsymbol{Y}_0^+ \boldsymbol{Y}_0 d\boldsymbol{u} = \left(\boldsymbol{I}_N - \boldsymbol{u}_0 \boldsymbol{u}_0^H\right) d\boldsymbol{u} = d\boldsymbol{u} - \boldsymbol{u}_0 \boldsymbol{u}_0^H d\boldsymbol{u} = d\boldsymbol{u}. \tag{35}$$

(34) and (35) lead to:

$$d\boldsymbol{u} = (\lambda_0 \boldsymbol{I}_N - \boldsymbol{Z}_0)^+ \left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0 \boldsymbol{v}_0^H}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \right) (d\boldsymbol{Z})\,\boldsymbol{u}_0. \tag{36}$$

From (36), it is possible to find the differential of $\boldsymbol{u}(\boldsymbol{Z})$ evaluated at $\boldsymbol{Z}_0$ with respect to the matrix $\boldsymbol{Z}$:

$$\begin{aligned} d\boldsymbol{u} &= \text{vec}(d\boldsymbol{u}) = \text{vec}\left((\lambda_0 \boldsymbol{I}_N - \boldsymbol{Z}_0)^+ \left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0 \boldsymbol{v}_0^H}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \right) (d\boldsymbol{Z})\,\boldsymbol{u}_0 \right) \\ &= \left(\boldsymbol{u}_0^T \otimes \left[(\lambda_0 \boldsymbol{I}_N - \boldsymbol{Z}_0)^+ \left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0 \boldsymbol{v}_0^H}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \right) \right] \right) d\,\text{vec}(\boldsymbol{Z}), \end{aligned} \tag{37}$$

where (32) was used. From (37), it follows that $\mathcal{D}_{\boldsymbol{Z}} \boldsymbol{u} = \boldsymbol{u}_0^T \otimes \left[(\lambda_0 \boldsymbol{I}_N - \boldsymbol{Z}_0)^+ \left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0 \boldsymbol{v}_0^H}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \right) \right]$. The differential and the derivative of $\boldsymbol{u}^*$ follow by the use of Table II, and the results found here.

The left eigenvector function $\boldsymbol{v} : \mathbb{C}^{N\times N} \rightarrow \mathbb{C}^{N\times 1}$ with the argument $\boldsymbol{Z} \in \mathbb{C}^{N\times N}$, denoted $\boldsymbol{v}(\boldsymbol{Z})$, is defined through the following four relations: $\boldsymbol{v}^H(\boldsymbol{Z})\boldsymbol{Z} = \lambda(\boldsymbol{Z})\boldsymbol{v}^H$, $\boldsymbol{v}_0^H\boldsymbol{v}(\boldsymbol{Z}) = 1$, $\lambda(\boldsymbol{Z}_0) = \lambda_0$, and $\boldsymbol{v}(\boldsymbol{Z}_0) = \boldsymbol{v}_0$. The differential of $\boldsymbol{v}(\boldsymbol{Z})$ at $\boldsymbol{Z} = \boldsymbol{Z}_0$ can be found, using a procedure similar to the one used earlier in this subsection for finding $d\boldsymbol{u}$ at $\boldsymbol{Z} = \boldsymbol{Z}_0$, giving $d\boldsymbol{v}^H = \boldsymbol{v}_0^H\,(d\boldsymbol{Z})\left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0\boldsymbol{v}_0^H}{\boldsymbol{v}_0^H\boldsymbol{u}_0}\right)(\lambda_0\boldsymbol{I}_N - \boldsymbol{Z}_0)^+$.

G. Derivative of $\boldsymbol{F}(z, z^*)$

Examples of functions of the type $\boldsymbol{F}(z, z^*)$ are $\boldsymbol{A}z$, $\boldsymbol{A}zz^*$, and $\boldsymbol{A}f(z, z^*)$, where $\boldsymbol{A} \in \mathbb{C}^{M\times P}$. These functions can be differentiated by finding the differentials of the scalar functions z, zz^*, and $f(z, z^*)$.

H. Derivative of $\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*)$

Let $\boldsymbol{F} : \mathbb{C}^{N\times 1} \times \mathbb{C}^{N\times 1} \rightarrow \mathbb{C}^{N\times N}$ be $\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{z}\boldsymbol{z}^H$. The differential of the $\boldsymbol{F}$ is $d\boldsymbol{F} = (d\boldsymbol{z})\boldsymbol{z}^H + \boldsymbol{z}d\boldsymbol{z}^H$. And from this equation $d\,\mathrm{vec}(\boldsymbol{F}) = [\boldsymbol{z}^* \otimes \boldsymbol{I}_N]\,d\,\mathrm{vec}(\boldsymbol{z}) + [\boldsymbol{I}_N \otimes \boldsymbol{z}]\,d\,\mathrm{vec}(\boldsymbol{z}^H) = [\boldsymbol{z}^* \otimes \boldsymbol{I}_N]\,d\boldsymbol{z} + [\boldsymbol{I}_N \otimes \boldsymbol{z}]\,d\boldsymbol{z}^*$. Hence, the derivatives of $\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{z}\boldsymbol{z}^H$ are given by $\mathcal{D}_{\boldsymbol{z}}\boldsymbol{F} = \boldsymbol{z}^* \otimes \boldsymbol{I}_N$, and $\mathcal{D}_{\boldsymbol{z}^*}\boldsymbol{F} = \boldsymbol{I}_N \otimes \boldsymbol{z}$.

I. Derivative of $\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$

1) **Kronecker Product Related Problems***:* Examples of objective functions which depends on the Kronecker product of the unknown complex-valued matrix is found in [18], [28].

Now, the *commutation matrix* is introduced:

Definition 6: *Let $\boldsymbol{A} \in \mathbb{C}^{N\times Q}$, then there exists a permutation matrix that connects the vectors* $\mathrm{vec}\,(\boldsymbol{A})$ *and* $\mathrm{vec}\left(\boldsymbol{A}^T\right)$*. The* commutation matrix $\boldsymbol{K}_{N,Q}$ *has size $NQ \times NQ$, and it defines the connection between* $\mathrm{vec}\,(\boldsymbol{A})$ *and* $\mathrm{vec}\left(\boldsymbol{A}^T\right)$ *in the following way:* $\boldsymbol{K}_{N,Q}\,\mathrm{vec}(\boldsymbol{A}) = \mathrm{vec}(\boldsymbol{A}^T)$.

The matrix $\boldsymbol{K}_{Q,N}$ is a permutation matrix and it can be shown [7] that $\boldsymbol{K}_{Q,N}^T = \boldsymbol{K}_{Q,N}^{-1} = \boldsymbol{K}_{N,Q}$. The following result is Theorem 3.9 in [7]: Let $\boldsymbol{A}_i \in \mathbb{C}^{N_i\times Q_i}$ where $i \in \{0, 1\}$, then the reason for why the commutation matrix received its name can be seen from the following identity: $\boldsymbol{K}_{N_1,N_0}\,(\boldsymbol{A}_0 \otimes \boldsymbol{A}_1) = (\boldsymbol{A}_1 \otimes \boldsymbol{A}_0)\,\boldsymbol{K}_{Q_1,Q_0}$.

Let $\boldsymbol{F} : \mathbb{C}^{N_0\times Q_0} \times \mathbb{C}^{N_1\times Q_1} \rightarrow \mathbb{C}^{N_0N_1\times Q_0Q_1}$ be given by $\boldsymbol{F}(\boldsymbol{Z}_0, \boldsymbol{Z}_1) = \boldsymbol{Z}_0 \otimes \boldsymbol{Z}_1$, where $\boldsymbol{Z}_i \in \mathbb{C}^{N_i\times Q_i}$. The differential of this function follows from Table II: $d\boldsymbol{F} = (d\boldsymbol{Z}_0) \otimes \boldsymbol{Z}_1 + \boldsymbol{Z}_0 \otimes d\boldsymbol{Z}_1$. Applying the $\mathrm{vec}(\cdot)$ operator to $d\boldsymbol{F}$ yields: $d\,\mathrm{vec}(\boldsymbol{F}) = \mathrm{vec}\,((d\boldsymbol{Z}_0) \otimes \boldsymbol{Z}_1) + \mathrm{vec}\,(\boldsymbol{Z}_0 \otimes d\boldsymbol{Z}_1)$. From $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = \boldsymbol{AC} \otimes \boldsymbol{BD}$ and

Theorem 3.10 in [7], it follows that

$$\begin{aligned}\text{vec}\left((d\boldsymbol{Z}_0)\otimes\boldsymbol{Z}_1\right) &= \left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[(d\,\text{vec}(\boldsymbol{Z}_0))\otimes\text{vec}(\boldsymbol{Z}_1)\right]\\ &= \left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[(\boldsymbol{I}_{N_0Q_0}d\,\text{vec}(\boldsymbol{Z}_0))\otimes(\text{vec}(\boldsymbol{Z}_1)1)\right]\\ &= \left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[(\boldsymbol{I}_{N_0Q_0}\otimes\text{vec}(\boldsymbol{Z}_1))\left(d\,\text{vec}(\boldsymbol{Z}_0)\otimes 1\right)\right]\\ &= \left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\boldsymbol{I}_{N_0Q_0}\otimes\text{vec}(\boldsymbol{Z}_1)\right]d\,\text{vec}(\boldsymbol{Z}_0),\end{aligned}\tag{38}$$

and in a similar way it follows that

$$\begin{aligned}\text{vec}\left(\boldsymbol{Z}_0\otimes d\boldsymbol{Z}_1\right) &= \left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\text{vec}(\boldsymbol{Z}_0)\otimes d\,\text{vec}(\boldsymbol{Z}_1)\right]\\ &= \left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\text{vec}(\boldsymbol{Z}_0)\otimes\boldsymbol{I}_{N_1Q_1}\right]d\,\text{vec}(\boldsymbol{Z}_1).\end{aligned}\tag{39}$$

Inserting the result from (38) and (39) into $d\,\text{vec}(\boldsymbol{F})$ gives:

$$\begin{aligned}d\,\text{vec}(\boldsymbol{F}) &= \left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\boldsymbol{I}_{N_0Q_0}\otimes\text{vec}(\boldsymbol{Z}_1)\right]d\,\text{vec}(\boldsymbol{Z}_0)\\ &+\left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\text{vec}(\boldsymbol{Z}_0)\otimes\boldsymbol{I}_{N_1Q_1}\right]d\,\text{vec}(\boldsymbol{Z}_1).\end{aligned}\tag{40}$$

Define the matrices $\boldsymbol{A}(\boldsymbol{Z}_1)$ and $\boldsymbol{B}(\boldsymbol{Z}_0)$ by $\boldsymbol{A}(\boldsymbol{Z}_1)\triangleq\left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\boldsymbol{I}_{N_0Q_0}\otimes\text{vec}(\boldsymbol{Z}_1)\right]$, and $\boldsymbol{B}(\boldsymbol{Z}_0)=\left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\text{vec}(\boldsymbol{Z}_0)\otimes\boldsymbol{I}_{N_1Q_1}\right]$. It is then possible to rewrite the differential of $\boldsymbol{F}(\boldsymbol{Z}_0,\boldsymbol{Z}_1)=\boldsymbol{Z}_0\otimes\boldsymbol{Z}_1$ as $d\,\text{vec}(\boldsymbol{F})=\boldsymbol{A}(\boldsymbol{Z}_1)d\,\text{vec}(\boldsymbol{Z}_0)+\boldsymbol{B}(\boldsymbol{Z}_0)d\,\text{vec}(\boldsymbol{Z}_1)$. From $d\,\text{vec}(\boldsymbol{F})$, the differentials and derivatives of $\boldsymbol{Z}\otimes\boldsymbol{Z}$, $\boldsymbol{Z}\otimes\boldsymbol{Z}^*$, and $\boldsymbol{Z}^*\otimes\boldsymbol{Z}^*$ can be derived and these results are included in Table VI. In the table, $\text{diag}(\cdot)$ returns the square diagonal matrix with the input column vector elements on the main diagonal [22] and zeros elsewhere.

2) **Moore-Penrose Inverse Related Problems***:* In pseudo-inverse matrix based receiver design, the Moore-Penrose inverse might appear [19]. This is applicable for MIMO, CDMA, and OFDM systems.

Let $\boldsymbol{F}:\mathbb{C}^{N\times Q}\times\mathbb{C}^{N\times Q}\to\mathbb{C}^{Q\times N}$ be given by $\boldsymbol{F}(\boldsymbol{Z},\boldsymbol{Z}^*)=\boldsymbol{Z}^+$, where $\boldsymbol{Z}\in\mathbb{C}^{N\times Q}$. The reason for including both variables $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ in this function definition is that the differential of $\boldsymbol{Z}^+$, see Table II, depends on both $d\boldsymbol{Z}$ and $d\boldsymbol{Z}^*$. Using the $\text{vec}(\cdot)$ operator on the differential of the Moore-Penrose inverse in Table II, in addition to (32) and Definition 6, result in:

$$d\,\text{vec}(\boldsymbol{F})=-\left[\left(\boldsymbol{Z}^+\right)^T\otimes\boldsymbol{Z}^+\right]d\,\text{vec}(\boldsymbol{Z})+\left[\left(\boldsymbol{I}_N-\left(\boldsymbol{Z}^+\right)^T\boldsymbol{Z}^T\right)\otimes\boldsymbol{Z}^+\left(\boldsymbol{Z}^+\right)^H\right]\boldsymbol{K}_{N,Q}d\,\text{vec}(\boldsymbol{Z}^*)$$

TABLE VI
DERIVATIVES OF FUNCTIONS OF THE TYPE $\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$

$\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$	Differential $d\,\mathrm{vec}(\boldsymbol{F})$	$\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$	$\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$
$\boldsymbol{Z}$	$\boldsymbol{I}_{NQ} d\,\mathrm{vec}(\boldsymbol{Z})$	$\boldsymbol{I}_{NQ}$	$\boldsymbol{0}_{NQ \times NQ}$
$\boldsymbol{Z}^T$	$\boldsymbol{K}_{N,Q} d\,\mathrm{vec}(\boldsymbol{Z})$	$\boldsymbol{K}_{N,Q}$	$\boldsymbol{0}_{NQ \times NQ}$
$\boldsymbol{Z}^*$	$\boldsymbol{I}_{NQ} d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\boldsymbol{0}_{NQ \times NQ}$	$\boldsymbol{I}_{NQ}$
$\boldsymbol{Z}^H$	$\boldsymbol{K}_{N,Q} d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\boldsymbol{0}_{NQ \times NQ}$	$\boldsymbol{K}_{N,Q}$
$\boldsymbol{Z}\boldsymbol{Z}^T$	$\left(\boldsymbol{I}_{N^2} + \boldsymbol{K}_{N,N}\right)(\boldsymbol{Z} \otimes \boldsymbol{I}_N)\, d\,\mathrm{vec}(\boldsymbol{Z})$	$\left(\boldsymbol{I}_{N^2} + \boldsymbol{K}_{N,N}\right)(\boldsymbol{Z} \otimes \boldsymbol{I}_N)$	$\boldsymbol{0}_{N^2 \times NQ}$
$\boldsymbol{Z}^T\boldsymbol{Z}$	$\left(\boldsymbol{I}_{Q^2} + \boldsymbol{K}_{Q,Q}\right)\left(\boldsymbol{I}_Q \otimes \boldsymbol{Z}^T\right) d\,\mathrm{vec}(\boldsymbol{Z})$	$\left(\boldsymbol{I}_{Q^2} + \boldsymbol{K}_{Q,Q}\right)\left(\boldsymbol{I}_Q \otimes \boldsymbol{Z}^T\right)$	$\boldsymbol{0}_{Q^2 \times NQ}$
$\boldsymbol{Z}\boldsymbol{Z}^H$	$(\boldsymbol{Z}^* \otimes \boldsymbol{I}_N)\, d\,\mathrm{vec}(\boldsymbol{Z}) + \boldsymbol{K}_{N,N}(\boldsymbol{Z} \otimes \boldsymbol{I}_N)\, d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\boldsymbol{Z}^* \otimes \boldsymbol{I}_N$	$\boldsymbol{K}_{N,N}(\boldsymbol{Z} \otimes \boldsymbol{I}_N)$
$\boldsymbol{Z}^{-1}$	$-\left((\boldsymbol{Z}^T)^{-1} \otimes \boldsymbol{Z}^{-1}\right) d\,\mathrm{vec}(\boldsymbol{Z})$	$-(\boldsymbol{Z}^T)^{-1} \otimes \boldsymbol{Z}^{-1}$	$\boldsymbol{0}_{N^2 \times N^2}$
$\boldsymbol{Z}^p$	$\sum_{i=1}^{p}\left((\boldsymbol{Z}^T)^{p-i} \otimes \boldsymbol{Z}^{i-1}\right) d\,\mathrm{vec}(\boldsymbol{Z})$	$\sum_{i=1}^{p}\left((\boldsymbol{Z}^T)^{p-i} \otimes \boldsymbol{Z}^{i-1}\right)$	$\boldsymbol{0}_{N^2 \times N^2}$
$\boldsymbol{Z} \otimes \boldsymbol{Z}$	$(\boldsymbol{A}(\boldsymbol{Z}) + \boldsymbol{B}(\boldsymbol{Z}))d\,\mathrm{vec}(\boldsymbol{Z})$	$\boldsymbol{A}(\boldsymbol{Z}) + \boldsymbol{B}(\boldsymbol{Z})$	$\boldsymbol{0}_{N^2Q^2 \times NQ}$
$\boldsymbol{Z} \otimes \boldsymbol{Z}^*$	$\boldsymbol{A}(\boldsymbol{Z}^*)d\,\mathrm{vec}(\boldsymbol{Z}) + \boldsymbol{B}(\boldsymbol{Z})d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\boldsymbol{A}(\boldsymbol{Z}^*)$	$\boldsymbol{B}(\boldsymbol{Z})$
$\boldsymbol{Z}^* \otimes \boldsymbol{Z}^*$	$(\boldsymbol{A}(\boldsymbol{Z}^*) + \boldsymbol{B}(\boldsymbol{Z}^*))d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\boldsymbol{0}_{N^2Q^2 \times NQ}$	$\boldsymbol{A}(\boldsymbol{Z}^*) + \boldsymbol{B}(\boldsymbol{Z}^*)$
$\boldsymbol{Z} \odot \boldsymbol{Z}$	$2\,\mathrm{diag}(\mathrm{vec}(\boldsymbol{Z}))d\,\mathrm{vec}(\boldsymbol{Z})$	$2\,\mathrm{diag}(\mathrm{vec}(\boldsymbol{Z}))$	$\boldsymbol{0}_{NQ \times NQ}$
$\boldsymbol{Z} \odot \boldsymbol{Z}^*$	$\mathrm{diag}(\mathrm{vec}(\boldsymbol{Z}^*))d\,\mathrm{vec}(\boldsymbol{Z}) + \mathrm{diag}(\mathrm{vec}(\boldsymbol{Z}))d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\mathrm{diag}(\mathrm{vec}(\boldsymbol{Z}^*))$	$\mathrm{diag}(\mathrm{vec}(\boldsymbol{Z}))$
$\boldsymbol{Z}^* \odot \boldsymbol{Z}^*$	$2\,\mathrm{diag}(\mathrm{vec}(\boldsymbol{Z}^*))d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\boldsymbol{0}_{NQ \times NQ}$	$2\,\mathrm{diag}(\mathrm{vec}(\boldsymbol{Z}^*))$
$\exp(\boldsymbol{Z})$	$\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\sum_{i=0}^{k}\left(\left(\boldsymbol{Z}^T\right)^{k-i} \otimes \boldsymbol{Z}^i\right) d\,\mathrm{vec}(\boldsymbol{Z})$	$\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\sum_{i=0}^{k}\left(\boldsymbol{Z}^T\right)^{k-i} \otimes \boldsymbol{Z}^i$	$\boldsymbol{0}_{N^2 \times N^2}$
$\exp(\boldsymbol{Z}^*)$	$\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\sum_{i=0}^{k}\left(\left(\boldsymbol{Z}^H\right)^{k-i} \otimes (\boldsymbol{Z}^*)^i\right) d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\boldsymbol{0}_{N^2 \times N^2}$	$\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\sum_{i=0}^{k}\left(\left(\boldsymbol{Z}^H\right)^{k-i} \otimes (\boldsymbol{Z}^*)^i\right)$
$\exp\left(\boldsymbol{Z}^H\right)$	$\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\sum_{i=0}^{k}\left((\boldsymbol{Z}^*)^{k-i} \otimes (\boldsymbol{Z}^H)^i\right) \boldsymbol{K}_{N,N} d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\boldsymbol{0}_{N^2 \times N^2}$	$\sum_{k=0}^{\infty}\frac{1}{(k+1)!}\sum_{i=0}^{k}\left((\boldsymbol{Z}^*)^{k-i} \otimes (\boldsymbol{Z}^H)^i\right) \boldsymbol{K}_{N,N}$

$$+\left[\left(\boldsymbol{Z}^{+}\right)^{T}\left(\boldsymbol{Z}^{+}\right)^{*} \otimes \left(\boldsymbol{I}_Q - \boldsymbol{Z}^{+}\boldsymbol{Z}\right)\right] \boldsymbol{K}_{N,Q} d\,\mathrm{vec}(\boldsymbol{Z}^*). \tag{41}$$

This leads to $\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F} = \left\{\left[\left(\boldsymbol{I}_N - \left(\boldsymbol{Z}^{+}\right)^{T}\boldsymbol{Z}^{T}\right) \otimes \boldsymbol{Z}^{+}\left(\boldsymbol{Z}^{+}\right)^{H}\right] + \left[\left(\boldsymbol{Z}^{+}\right)^{T}\left(\boldsymbol{Z}^{+}\right)^{*} \otimes \left(\boldsymbol{I}_Q - \boldsymbol{Z}^{+}\boldsymbol{Z}\right)\right]\right\} \boldsymbol{K}_{N,Q}$ and $\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F} = -\left[\left(\boldsymbol{Z}^{+}\right)^{T} \otimes \boldsymbol{Z}^{+}\right]$. If $\boldsymbol{Z}$ is invertible, then the derivative of $\boldsymbol{Z}^{+} = \boldsymbol{Z}^{-1}$ with respect to $\boldsymbol{Z}^*$ is equal to the zero matrix and the derivative of $\boldsymbol{Z}^{+} = \boldsymbol{Z}^{-1}$ with respect to $\boldsymbol{Z}$ can be found from Table VI. Since the expressions associated with the differential of the Moore-Penrose inverse is so long, it is not included in Table VI.

VII. HESSIAN MATRIX FORMULAS

In the previous sections, we have proposed tools for finding first-order derivatives that can be of help to find stationary points in optimization problems. To identify the nature of these stationary points (minimum, maximum, or saddle point) or to study the stability of iterative algorithms, the second-order derivatives are useful. This can be done by examining the Hessian matrix.

The Hessian matrix of a function is a matrix containing the *second-order derivatives* of the function. In this section, the Hessian matrix is defined and it is shown how it can be obtained. Only the case of scalar functions $f \in \mathbb{C}$ is treated, since this is the case of interest in most practical situations. However, the results can be extended to the vector- and matrix-functions as well. The way the Hessian is defined here follows in the same lines as given in [7], treating the real case. A complex version of Newton's recursion formula is derived in [29], [30], and there the topic of Hessian matrices is briefly treated for real scalar functions. Therefore, the complex Hessian might be used to accelerate the convergence of iterative optimization algorithms.

A. Identification of the Complex Hessian Matrices

When dealing with the Hessian matrix, it is the second-order differential that has to be calculated in order to identify the Hessian matrix. Neither of the matrices $d\boldsymbol{Z}$ nor $d\boldsymbol{Z}^*$ is a function of $\boldsymbol{Z}$ or $\boldsymbol{Z}^*$ and, hence, their differentials are the zero matrix. Mathematically, this can be formulated as:

$$d^2\boldsymbol{Z} = d\left(d\boldsymbol{Z}\right) = \mathbf{0}_{N\times Q} = d\left(d\boldsymbol{Z}^*\right) = d^2\boldsymbol{Z}^*. \tag{42}$$

If $f \in \mathbb{C}$, then $\left(d^2 f\right)^T = d\left(df\right)^T = d^2 f^T = d^2 f$ and if $f \in \mathbb{R}$, then $\left(d^2 f\right)^H = d\left(df\right)^H = d^2 f^H = d^2 f$.

The Hessian depends on *two* variables such that the notation must include which variable the Hessian is calculated with respect to. If the Hessian is calculated with respect to $\boldsymbol{Z}_0$ and $\boldsymbol{Z}_1$, the Hessian is denoted $\mathcal{H}_{\boldsymbol{Z}_0,\boldsymbol{Z}_1} f$. The following definition is used for $\mathcal{H}_{\boldsymbol{Z}_0,\boldsymbol{Z}_1} f$ and it is an extension of the definition in [7], to complex scalar functions.

Definition 7: *Let* $\boldsymbol{Z}_i \in \mathbb{C}^{N_i\times Q_i}$, $i \in \{0,1\}$, *then* $\mathcal{H}_{\boldsymbol{Z}_0,\boldsymbol{Z}_1} f \in \mathbb{C}^{N_1 Q_1 \times N_0 Q_0}$ *and* $\mathcal{H}_{\boldsymbol{Z}_0,\boldsymbol{Z}_1} f \triangleq \mathcal{D}_{\boldsymbol{Z}_0}\left(\mathcal{D}_{\boldsymbol{Z}_1} f\right)^T$.

Later, the following proposition is important for showing the symmetry of the complex Hessian matrix.

Proposition 2: *Let* $f : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}$. *Assumed that the function* $f(\boldsymbol{Z},\boldsymbol{Z}^*)$ *is twice differentiable with respect to all of the variables inside* $\boldsymbol{Z}$ *and* $\boldsymbol{Z}^*$, *when these variables are treated as independent variables. Then [7]* $\frac{\partial^2}{\partial z_{k,l}\partial z_{m,n}} f = \frac{\partial^2}{\partial z_{m,n}\partial z_{k,l}} f$, $\frac{\partial^2}{\partial z^*_{k,l}\partial z^*_{m,n}} f = \frac{\partial^2}{\partial z^*_{m,n}\partial z^*_{k,l}} f$, *and* $\frac{\partial^2}{\partial z^*_{k,l}\partial z_{m,n}} f = \frac{\partial^2}{\partial z_{m,n}\partial z^*_{k,l}} f$, *where* $m,k \in \{0,1,\ldots,N-1\}$ *and* $n,l \in \{0,1,\ldots,Q-1\}$.

Let $p_i = N_i k_i + l_i$ where $i \in \{0,1\}$, $k_i \in \{0,1,\ldots,Q_i-1\}$ and $l_i \in \{0,1,\ldots,N_i-1\}$. As a consequence of Definition 7 and (12), it follows that element number (p_0,p_1) of $\mathcal{H}_{\boldsymbol{Z}_0,\boldsymbol{Z}_1} f$ is given by:

$$\begin{aligned}\left(\mathcal{H}_{\boldsymbol{Z}_0,\boldsymbol{Z}_1} f\right)_{p_0,p_1} &= \left(\mathcal{D}_{\boldsymbol{Z}_0}\left(\mathcal{D}_{\boldsymbol{Z}_1} f\right)^T\right)_{p_0,p_1} = \left[\frac{\partial}{\partial \operatorname{vec}^T\left(\boldsymbol{Z}_0\right)}\left(\frac{\partial}{\partial \operatorname{vec}^T\left(\boldsymbol{Z}_1\right)} f\right)^T\right]_{p_0,p_1} = \frac{\partial}{\partial\left(\operatorname{vec}(\boldsymbol{Z}_0)\right)_{p_1}}\frac{\partial}{\partial\left(\operatorname{vec}(\boldsymbol{Z}_1)\right)_{p_0}} f \\ &= \frac{\partial}{\partial\left(\operatorname{vec}(\boldsymbol{Z}_0)\right)_{N_1 k_1 + l_1}}\frac{\partial}{\partial\left(\operatorname{vec}(\boldsymbol{Z}_1)\right)_{N_0 k_0 + l_0}} f = \frac{\partial^2 f}{\partial\left(\boldsymbol{Z}_0\right)_{l_1,k_1}\partial\left(\boldsymbol{Z}_1\right)_{l_0,k_0}}.\end{aligned} \tag{43}$$

And as an immediate consequence of (43) and Proposition 2, it follows that for twice differentiable functions f:

$$\left(\mathcal{H}_{\boldsymbol{Z},\boldsymbol{Z}} f\right)^T = \mathcal{H}_{\boldsymbol{Z},\boldsymbol{Z}} f, \quad \left(\mathcal{H}_{\boldsymbol{Z}^*,\boldsymbol{Z}^*} f\right)^T = \mathcal{H}_{\boldsymbol{Z}^*,\boldsymbol{Z}^*} f, \quad \left(\mathcal{H}_{\boldsymbol{Z},\boldsymbol{Z}^*} f\right)^T = \mathcal{H}_{\boldsymbol{Z}^*,\boldsymbol{Z}} f. \tag{44}$$

In order to find an identification equation for the Hessians of the function f with respect to all the possible combinations of the variables $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$, an appropriate form of the expression $d^2 f$ is required. This expression is now derived. From (10), the first-order differential of the function f can be written as $df = (\mathcal{D}_{\boldsymbol{Z}} f) d\,\mathrm{vec}(\boldsymbol{Z}) + (\mathcal{D}_{\boldsymbol{Z}^*} f) d\,\mathrm{vec}(\boldsymbol{Z}^*)$, where $\mathcal{D}_{\boldsymbol{Z}} f \in \mathbb{C}^{1\times NQ}$ and $\mathcal{D}_{\boldsymbol{Z}^*} f \in \mathbb{C}^{1\times NQ}$ can be expressed by the use of (10) as:

$$(d\mathcal{D}_{\boldsymbol{Z}} f)^T = \left[\mathcal{D}_{\boldsymbol{Z}} \left(\mathcal{D}_{\boldsymbol{Z}} f\right)^T\right] d\,\mathrm{vec}(\boldsymbol{Z}) + \left[\mathcal{D}_{\boldsymbol{Z}^*} \left(\mathcal{D}_{\boldsymbol{Z}} f\right)^T\right] d\,\mathrm{vec}(\boldsymbol{Z}^*), \tag{45}$$

$$(d\mathcal{D}_{\boldsymbol{Z}^*} f)^T = \left[\mathcal{D}_{\boldsymbol{Z}} \left(\mathcal{D}_{\boldsymbol{Z}^*} f\right)^T\right] d\,\mathrm{vec}(\boldsymbol{Z}) + \left[\mathcal{D}_{\boldsymbol{Z}^*} \left(\mathcal{D}_{\boldsymbol{Z}^*} f\right)^T\right] d\,\mathrm{vec}(\boldsymbol{Z}^*). \tag{46}$$

From (45) and (46) it follows that:

$$d\mathcal{D}_{\boldsymbol{Z}} f = \left[d\,\mathrm{vec}^T(\boldsymbol{Z})\right] \left[\mathcal{D}_{\boldsymbol{Z}} \left(\mathcal{D}_{\boldsymbol{Z}} f\right)^T\right]^T + \left[d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right] \left[\mathcal{D}_{\boldsymbol{Z}^*} \left(\mathcal{D}_{\boldsymbol{Z}} f\right)^T\right]^T, \tag{47}$$

$$d\mathcal{D}_{\boldsymbol{Z}^*} f = \left[d\,\mathrm{vec}^T(\boldsymbol{Z})\right] \left[\mathcal{D}_{\boldsymbol{Z}} \left(\mathcal{D}_{\boldsymbol{Z}^*} f\right)^T\right]^T + \left[d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right] \left[\mathcal{D}_{\boldsymbol{Z}^*} \left(\mathcal{D}_{\boldsymbol{Z}^*} f\right)^T\right]^T. \tag{48}$$

$d^2 f$ is found by applying the differential operator on df and then utilizing the results from (42), (47), and (48):

$$\begin{aligned}
d^2 f &= (d\mathcal{D}_{\boldsymbol{Z}} f)\, d\,\mathrm{vec}(\boldsymbol{Z}) + (d\mathcal{D}_{\boldsymbol{Z}^*} f) d\,\mathrm{vec}(\boldsymbol{Z}^*) \\
&= \left[d\,\mathrm{vec}^T(\boldsymbol{Z})\right] \left[\mathcal{D}_{\boldsymbol{Z}} \left(\mathcal{D}_{\boldsymbol{Z}} f\right)^T\right]^T d\,\mathrm{vec}(\boldsymbol{Z}) + \left[d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right] \left[\mathcal{D}_{\boldsymbol{Z}^*} \left(\mathcal{D}_{\boldsymbol{Z}} f\right)^T\right]^T d\,\mathrm{vec}(\boldsymbol{Z}) \\
&\quad + \left[d\,\mathrm{vec}^T(\boldsymbol{Z})\right] \left[\mathcal{D}_{\boldsymbol{Z}} \left(\mathcal{D}_{\boldsymbol{Z}^*} f\right)^T\right]^T d\,\mathrm{vec}(\boldsymbol{Z}^*) + \left[d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right] \left[\mathcal{D}_{\boldsymbol{Z}^*} \left(\mathcal{D}_{\boldsymbol{Z}^*} f\right)^T\right]^T d\,\mathrm{vec}(\boldsymbol{Z}^*) \\
&= \left[d\,\mathrm{vec}^T(\boldsymbol{Z})\right] \left[\mathcal{D}_{\boldsymbol{Z}} \left(\mathcal{D}_{\boldsymbol{Z}} f\right)^T\right] d\,\mathrm{vec}(\boldsymbol{Z}) + \left[d\,\mathrm{vec}^T(\boldsymbol{Z})\right] \left[\mathcal{D}_{\boldsymbol{Z}^*} \left(\mathcal{D}_{\boldsymbol{Z}} f\right)^T\right] d\,\mathrm{vec}(\boldsymbol{Z}^*) \\
&\quad + \left[d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right] \left[\mathcal{D}_{\boldsymbol{Z}} \left(\mathcal{D}_{\boldsymbol{Z}^*} f\right)^T\right] d\,\mathrm{vec}(\boldsymbol{Z}) + \left[d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right] \left[\mathcal{D}_{\boldsymbol{Z}^*} \left(\mathcal{D}_{\boldsymbol{Z}^*} f\right)^T\right] d\,\mathrm{vec}(\boldsymbol{Z}^*) \\
&= \left[d\,\mathrm{vec}^T(\boldsymbol{Z}^*),\ \ d\,\mathrm{vec}^T(\boldsymbol{Z})\right] \begin{bmatrix} \mathcal{D}_{\boldsymbol{Z}} \left(\mathcal{D}_{\boldsymbol{Z}^*} f\right)^T, & \mathcal{D}_{\boldsymbol{Z}^*} \left(\mathcal{D}_{\boldsymbol{Z}^*} f\right)^T \\ \mathcal{D}_{\boldsymbol{Z}} \left(\mathcal{D}_{\boldsymbol{Z}} f\right)^T, & \mathcal{D}_{\boldsymbol{Z}^*} \left(\mathcal{D}_{\boldsymbol{Z}} f\right)^T \end{bmatrix} \begin{bmatrix} d\,\mathrm{vec}(\boldsymbol{Z}) \\ d\,\mathrm{vec}(\boldsymbol{Z}^*) \end{bmatrix}.
\end{aligned} \tag{49}$$

From Definition 7, it follows that $d^2 f$ can be rewritten as:

$$d^2 f = \left[d\,\mathrm{vec}^T(\boldsymbol{Z}^*),\ \ d\,\mathrm{vec}^T(\boldsymbol{Z})\right] \begin{bmatrix} \mathcal{H}_{\boldsymbol{Z},\boldsymbol{Z}^*} f, & \mathcal{H}_{\boldsymbol{Z}^*,\boldsymbol{Z}^*} f \\ \mathcal{H}_{\boldsymbol{Z},\boldsymbol{Z}} f, & \mathcal{H}_{\boldsymbol{Z}^*,\boldsymbol{Z}} f \end{bmatrix} \begin{bmatrix} d\,\mathrm{vec}(\boldsymbol{Z}) \\ d\,\mathrm{vec}(\boldsymbol{Z}^*) \end{bmatrix}. \tag{50}$$

Assume that it is possible to find an expression of $d^2 f$ in the following form

$$d^2 f = \left[d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right] \boldsymbol{A}_{0,0} d\,\mathrm{vec}(\boldsymbol{Z}) + \left[d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right] \boldsymbol{A}_{0,1} d\,\mathrm{vec}(\boldsymbol{Z}^*)$$

$$+\left[d\,\mathrm{vec}^T(\boldsymbol{Z})\right]\boldsymbol{A}_{1,0}d\,\mathrm{vec}(\boldsymbol{Z})+\left[d\,\mathrm{vec}^T(\boldsymbol{Z})\right]\boldsymbol{A}_{1,1}d\,\mathrm{vec}(\boldsymbol{Z}^*)$$

$$=\left[d\,\mathrm{vec}^T(\boldsymbol{Z}^*),\quad d\,\mathrm{vec}^T(\boldsymbol{Z})\right]\begin{bmatrix}\boldsymbol{A}_{0,0}, & \boldsymbol{A}_{0,1}\\ \boldsymbol{A}_{1,0}, & \boldsymbol{A}_{1,1}\end{bmatrix}\begin{bmatrix}d\,\mathrm{vec}(\boldsymbol{Z})\\ d\,\mathrm{vec}(\boldsymbol{Z}^*)\end{bmatrix}, \tag{51}$$

where $\boldsymbol{A}_{k,l}\in\mathbb{C}^{NQ\times NQ}$, $k,l\in\{0,1\}$, can possibly be dependent on $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$ but *not* on $d\,\mathrm{vec}(\boldsymbol{Z})$ and $d\,\mathrm{vec}(\boldsymbol{Z}^*)$. The four complex Hessian matrices in (50) can now be identified from $\boldsymbol{A}_{k,l}$ in (51) in the following way: Subtract the second-order differentials in (50) from (51), and then use (44) together with Lemma 2 to get

$$\mathcal{H}_{\boldsymbol{Z}^*,\boldsymbol{Z}}f=\frac{\boldsymbol{A}_{0,0}^T+\boldsymbol{A}_{1,1}}{2},\ \ \mathcal{H}_{\boldsymbol{Z},\boldsymbol{Z}^*}f=\frac{\boldsymbol{A}_{0,0}+\boldsymbol{A}_{1,1}^T}{2},\ \ \mathcal{H}_{\boldsymbol{Z},\boldsymbol{Z}}f=\frac{\boldsymbol{A}_{1,0}+\boldsymbol{A}_{1,0}^T}{2},\ \ \mathcal{H}_{\boldsymbol{Z}^*,\boldsymbol{Z}^*}f=\frac{\boldsymbol{A}_{0,1}+\boldsymbol{A}_{0,1}^T}{2}. \tag{52}$$

The complex Hessian matrices of the scalar function f can be computed using a three-step procedure:

1) *Compute* d^2f.
2) *Manipulate* d^2f *into the form given in (*51*).*
3) *Use (*52*) to identify the four Hessian matrices.*

In order to check convexity of f, the middle matrix on the right hand side of (50) must be positive definite.

B. Examples of Calculation of the Complex Hessian Matrices

1) **Complex Hessian of** $f(\boldsymbol{z},\boldsymbol{z}^*)$*:* Let $f:\mathbb{C}^{N\times 1}\times\mathbb{C}^{N\times 1}\to\mathbb{C}$ be defined as: $f(\boldsymbol{z},\boldsymbol{z}^*)=\boldsymbol{z}^H\boldsymbol{\Phi}\boldsymbol{z}$. The second-order differential of f is $d^2f=2\left(d\boldsymbol{z}^H\right)\boldsymbol{\Phi}d\boldsymbol{z}$. From this, $\boldsymbol{A}_{k,l}$ in (51) are $\boldsymbol{A}_{0,0}=2\boldsymbol{\Phi}$, $\boldsymbol{A}_{0,1}=\boldsymbol{A}_{1,0}=\boldsymbol{A}_{1,1}=\boldsymbol{0}_{N\times N}$. (52) gives the four complex Hessian matrices $\mathcal{H}_{\boldsymbol{Z},\boldsymbol{Z}}f=\boldsymbol{0}_{N\times N}$, $\mathcal{H}_{\boldsymbol{Z}^*,\boldsymbol{Z}^*}f=\boldsymbol{0}_{N\times N}$, $\mathcal{H}_{\boldsymbol{Z}^*,\boldsymbol{Z}}f=\boldsymbol{\Phi}^T$, and $\mathcal{H}_{\boldsymbol{Z},\boldsymbol{Z}^*}f=\boldsymbol{\Phi}$.

2) **Complex Hessian of** $f(\boldsymbol{Z},\boldsymbol{Z}^*)$*:* See the discussion around (29) for an introduction to the eigenvalue problem. $d^2\lambda(\boldsymbol{Z})$ is now found at $\boldsymbol{Z}=\boldsymbol{Z}_0$. The derivation here is similar to the one in [7], where the same result for $d^2\lambda$ was found. Applying the differential operator to both sides of (30), results in: $2\left(d\boldsymbol{Z}\right)\left(d\boldsymbol{u}\right)+\boldsymbol{Z}_0d^2\boldsymbol{u}=\left(d^2\lambda\right)\boldsymbol{u}_0+2\left(d\lambda\right)d\boldsymbol{u}+\lambda_0d^2\boldsymbol{u}$. Left-multiplying this equation by $\boldsymbol{v}_0^H$, and solving for $d^2\lambda$ gives

$$d^2\lambda=\frac{2\boldsymbol{v}_0^H\left(d\boldsymbol{Z}-\boldsymbol{I}_Nd\lambda\right)d\boldsymbol{u}}{\boldsymbol{v}_0^H\boldsymbol{u}_0}=\frac{2\boldsymbol{v}_0^H\left(d\boldsymbol{Z}-\boldsymbol{I}_N\frac{\boldsymbol{v}_0^H(d\boldsymbol{Z})\boldsymbol{u}_0}{\boldsymbol{v}_0^H\boldsymbol{u}_0}\right)d\boldsymbol{u}}{\boldsymbol{v}_0^H\boldsymbol{u}_0}=\frac{2\left(\boldsymbol{v}_0^Hd\boldsymbol{Z}-\frac{\boldsymbol{v}_0^H(d\boldsymbol{Z})\boldsymbol{u}_0\boldsymbol{v}_0^H}{\boldsymbol{v}_0^H\boldsymbol{u}_0}\right)d\boldsymbol{u}}{\boldsymbol{v}_0^H\boldsymbol{u}_0}$$
$$=\frac{2\boldsymbol{v}_0^H\left(d\boldsymbol{Z}\right)\left(\boldsymbol{I}_N-\frac{\boldsymbol{u}_0\boldsymbol{v}_0^H}{\boldsymbol{v}_0^H\boldsymbol{u}_0}\right)\left(\lambda_0\boldsymbol{I}_N-\boldsymbol{Z}_0\right)^+\left(\boldsymbol{I}_N-\frac{\boldsymbol{u}_0\boldsymbol{v}_0^H}{\boldsymbol{v}_0^H\boldsymbol{u}_0}\right)\left(d\boldsymbol{Z}\right)\boldsymbol{u}_0}{\boldsymbol{v}_0^H\boldsymbol{u}_0}, \tag{53}$$

where (31) and (36) were utilized. $d^2\lambda$ can be reformulated by means of Theorem 2.3 in [7] in the following way:

$$
\begin{aligned}
d^2\lambda &= \frac{2}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \operatorname{Tr}\left\{ \boldsymbol{u}_0 \boldsymbol{v}_0^H (d\boldsymbol{Z}) \left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0 \boldsymbol{v}_0^H}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \right) (\lambda_0 \boldsymbol{I}_N - \boldsymbol{Z}_0)^+ \left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0 \boldsymbol{v}_0^H}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \right) (d\boldsymbol{Z}) \right\} \\
&= \frac{2}{\boldsymbol{v}_0^H \boldsymbol{u}_0} d\operatorname{vec}^T(\boldsymbol{Z}) \left[\left\{ \left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0 \boldsymbol{v}_0^H}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \right) (\lambda_0 \boldsymbol{I}_N - \boldsymbol{Z}_0)^+ \left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0 \boldsymbol{v}_0^H}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \right) \right\} \otimes \boldsymbol{v}_0^* \boldsymbol{u}_0^T \right] d\operatorname{vec}\left(\boldsymbol{Z}^T\right) \\
&= \frac{2}{\boldsymbol{v}_0^H \boldsymbol{u}_0} d\operatorname{vec}^T(\boldsymbol{Z}) \left[\left\{ \left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0 \boldsymbol{v}_0^H}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \right) (\lambda_0 \boldsymbol{I}_N - \boldsymbol{Z}_0)^+ \left(\boldsymbol{I}_N - \frac{\boldsymbol{u}_0 \boldsymbol{v}_0^H}{\boldsymbol{v}_0^H \boldsymbol{u}_0} \right) \right\} \otimes \boldsymbol{v}_0^* \boldsymbol{u}_0^T \right] \boldsymbol{K}_{N,N} d\operatorname{vec}(\boldsymbol{Z}). \qquad (54)
\end{aligned}
$$

From this expression, it is possible to identify the four complex Hessian matrices by means of (51) and (52).

Let $f : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}$ be given by $f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \operatorname{Tr}\left\{\boldsymbol{Z}\boldsymbol{A}\boldsymbol{Z}^H\right\}$, where $\boldsymbol{Z} \in \mathbb{C}^{N\times Q}$ and $\boldsymbol{A} \in \mathbb{C}^{Q\times Q}$. From Theorem 2.3 in [7], f can written $f = \operatorname{vec}^T(\boldsymbol{Z}^*) \left(\boldsymbol{A}^T \otimes \boldsymbol{I}_N\right) \operatorname{vec}(\boldsymbol{Z})$. By applying the differential operator twice to the latest expression of f, it follows that $d^2 f = 2\left(d\operatorname{vec}^T(\boldsymbol{Z}^*)\right)\left(\boldsymbol{A}^T \otimes \boldsymbol{I}_N\right) d\operatorname{vec}(\boldsymbol{Z})$. From this expression, it is possible to identify the four complex Hessian matrices by means of (51) and (52).

VIII. CONCLUSIONS

An introduction is given to a set of very powerful tools that can be used to systematically find the derivative of complex-valued matrix functions that are dependent on complex-valued matrices. The key idea is to go through the complex differential of the function and to treat the differential of the complex variable and its complex conjugate as independent. This general framework can be used in many optimization problems that depend on complex parameters. Many results are given in tabular form.

APPENDIX I

PROOF OF PROPOSITION 1

Proof: *(8) leads to* $d\boldsymbol{Z}^+ = d\boldsymbol{Z}^+\boldsymbol{Z}\boldsymbol{Z}^+ = (d\boldsymbol{Z}^+\boldsymbol{Z})\boldsymbol{Z}^+ + \boldsymbol{Z}^+\boldsymbol{Z}d\boldsymbol{Z}^+$*. If* $\boldsymbol{Z}d\boldsymbol{Z}^+$ *is found from* $d\boldsymbol{Z}\boldsymbol{Z}^+ = (d\boldsymbol{Z})\boldsymbol{Z}^+ + \boldsymbol{Z}d\boldsymbol{Z}^+$*, and inserted in the expression for* $d\boldsymbol{Z}^+$*, then it is found that:*

$$d\boldsymbol{Z}^+ = (d\boldsymbol{Z}^+\boldsymbol{Z})\boldsymbol{Z}^+ + \boldsymbol{Z}^+(d\boldsymbol{Z}\boldsymbol{Z}^+ - (d\boldsymbol{Z})\boldsymbol{Z}^+) = (d\boldsymbol{Z}^+\boldsymbol{Z})\boldsymbol{Z}^+ + \boldsymbol{Z}^+d\boldsymbol{Z}\boldsymbol{Z}^+ - \boldsymbol{Z}^+(d\boldsymbol{Z})\boldsymbol{Z}^+. \qquad (55)$$

It is seen from (55), that it remains to express $d\boldsymbol{Z}^+\boldsymbol{Z}$ *and* $d\boldsymbol{Z}\boldsymbol{Z}^+$ *in terms of* $d\boldsymbol{Z}$ *and* $d\boldsymbol{Z}^*$*. Firstly,* $d\boldsymbol{Z}^+\boldsymbol{Z}$ *is handled:*

$$d\boldsymbol{Z}^+\boldsymbol{Z} = d\boldsymbol{Z}^+\boldsymbol{Z}\boldsymbol{Z}^+\boldsymbol{Z} = (d\boldsymbol{Z}^+\boldsymbol{Z})\boldsymbol{Z}^+\boldsymbol{Z} + \boldsymbol{Z}^+\boldsymbol{Z}(d\boldsymbol{Z}^+\boldsymbol{Z}) = (\boldsymbol{Z}^+\boldsymbol{Z}(d\boldsymbol{Z}^+\boldsymbol{Z}))^H + \boldsymbol{Z}^+\boldsymbol{Z}(d\boldsymbol{Z}^+\boldsymbol{Z}). \qquad (56)$$

The expression $\boldsymbol{Z}(d\boldsymbol{Z}^+\boldsymbol{Z})$ *can be found from* $d\boldsymbol{Z} = d\boldsymbol{Z}\boldsymbol{Z}^+\boldsymbol{Z} = (d\boldsymbol{Z})\boldsymbol{Z}^+\boldsymbol{Z} + \boldsymbol{Z}(d\boldsymbol{Z}^+\boldsymbol{Z})$*, and it is given by* $\boldsymbol{Z}(d\boldsymbol{Z}^+\boldsymbol{Z}) = d\boldsymbol{Z} - (d\boldsymbol{Z})\boldsymbol{Z}^+\boldsymbol{Z} = (d\boldsymbol{Z})(\boldsymbol{I}_Q - \boldsymbol{Z}^+\boldsymbol{Z})$*. If this expression is inserted into (56), it is found that:*

$d\boldsymbol{Z}^{+}\boldsymbol{Z} = (\boldsymbol{Z}^{+}(d\boldsymbol{Z})(\boldsymbol{I}_Q - \boldsymbol{Z}^{+}\boldsymbol{Z}))^{H} + \boldsymbol{Z}^{+}(d\boldsymbol{Z})(\boldsymbol{I}_Q - \boldsymbol{Z}^{+}\boldsymbol{Z}) = (\boldsymbol{I}_Q - \boldsymbol{Z}^{+}\boldsymbol{Z})\left(d\boldsymbol{Z}^{H}\right)(\boldsymbol{Z}^{+})^{H} + \boldsymbol{Z}^{+}(d\boldsymbol{Z})(\boldsymbol{I}_Q - \boldsymbol{Z}^{+}\boldsymbol{Z})$. *Secondly, it can be shown in a similar manner that:* $d\boldsymbol{Z}\boldsymbol{Z}^{+} = (\boldsymbol{I}_N - \boldsymbol{Z}\boldsymbol{Z}^{+})(d\boldsymbol{Z})\boldsymbol{Z}^{+} + (\boldsymbol{Z}^{+})^{H}\left(d\boldsymbol{Z}^{H}\right)(\boldsymbol{I}_N - \boldsymbol{Z}\boldsymbol{Z}^{+})$. *If the expressions for* $d\boldsymbol{Z}^{+}\boldsymbol{Z}$ *and* $d\boldsymbol{Z}\boldsymbol{Z}^{+}$ *are inserted into* (55)*, (9) is obtained.* ■

APPENDIX II

PROOF OF LEMMA 1

Proof: *Let* $\boldsymbol{A}_i \in \mathbb{C}^{M \times NQ}$ *be an arbitrary complex-valued function of* $\boldsymbol{Z} \in \mathbb{C}^{N \times Q}$ *and* $\boldsymbol{Z}^{*} \in \mathbb{C}^{N \times Q}$*. From Table* II*, it follows that* $d\,\mathrm{vec}(\boldsymbol{Z}) = d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\}) + \jmath d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\})$ *and* $d\,\mathrm{vec}(\boldsymbol{Z}^{*}) = d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\}) - \jmath d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\})$*. If these two expressions are substituted into* $\boldsymbol{A}_0 d\,\mathrm{vec}(\boldsymbol{Z}) + \boldsymbol{A}_1 d\,\mathrm{vec}(\boldsymbol{Z}^{*}) = \boldsymbol{0}_{M\times 1}$*, then it follows that* $\boldsymbol{A}_0(d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\}) + \jmath d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\})) + \boldsymbol{A}_1(d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\}) - \jmath d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\})) = \boldsymbol{0}_{M\times 1}$*. The last expression is equivalent to:* $(\boldsymbol{A}_0 + \boldsymbol{A}_1)d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\}) + \jmath(\boldsymbol{A}_0 - \boldsymbol{A}_1)d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\}) = \boldsymbol{0}_{M\times 1}$*. Since the differentials* $d\,\mathrm{Re}\{\boldsymbol{Z}\}$ *and* $d\,\mathrm{Im}\{\boldsymbol{Z}\}$ *are independent, so are* $d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\})$ *and* $d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\})$*. Therefore,* $\boldsymbol{A}_0 + \boldsymbol{A}_1 = \boldsymbol{0}_{M\times NQ}$ *and* $\boldsymbol{A}_0 - \boldsymbol{A}_1 = \boldsymbol{0}_{M\times NQ}$*. And from this it follows that* $\boldsymbol{A}_0 = \boldsymbol{A}_1 = \boldsymbol{0}_{M\times NQ}$*.* ■

APPENDIX III

PROOF OF LEMMA 2

First three other results are stated and proven:

Lemma 4: *Let* $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{N\times N}$*.* $\boldsymbol{z}^{T}\boldsymbol{A}\boldsymbol{z} = \boldsymbol{z}^{T}\boldsymbol{B}\boldsymbol{z} \;\forall\; \boldsymbol{z} \in \mathbb{C}^{N\times 1}$ *is equivalent to* $\boldsymbol{A} + \boldsymbol{A}^{T} = \boldsymbol{B} + \boldsymbol{B}^{T}$*.*

Proof: *Let* $(\boldsymbol{A})_{k,l} = a_{k,l}$ *and* $(\boldsymbol{B})_{k,l} = b_{k,l}$*. Assume first that* $\boldsymbol{z}^{T}\boldsymbol{A}\boldsymbol{z} = \boldsymbol{z}^{T}\boldsymbol{B}\boldsymbol{z} \;\forall\; \boldsymbol{z} \in \mathbb{C}^{N\times 1}$*, and set* $\boldsymbol{z} = \boldsymbol{e}_k$ *where* $k \in \{0, 1, \ldots, N-1\}$*. Then* $\boldsymbol{e}_k^{T}\boldsymbol{A}\boldsymbol{e}_k = \boldsymbol{e}_k^{T}\boldsymbol{B}\boldsymbol{e}_k$*, gives that* $a_{k,k} = b_{k,k}$ *for all* $k \in \{0, 1, \ldots, N-1\}$*. Setting* $\boldsymbol{z} = \boldsymbol{e}_k + \boldsymbol{e}_l$ *leads to* $(\boldsymbol{e}_k^{T} + \boldsymbol{e}_l^{T})\boldsymbol{A}(\boldsymbol{e}_k + \boldsymbol{e}_l) = (\boldsymbol{e}_k^{T} + \boldsymbol{e}_l^{T})\boldsymbol{B}(\boldsymbol{e}_k + \boldsymbol{e}_l)$*, which results in* $a_{k,k} + a_{l,l} + a_{k,l} + a_{l,k} = b_{k,k} + b_{l,l} + b_{k,l} + b_{l,k}$*. Eliminating equal terms gives* $a_{k,l} + a_{l,k} = b_{k,l} + b_{l,k}$ *which can be written* $\boldsymbol{A} + \boldsymbol{A}^{T} = \boldsymbol{B} + \boldsymbol{B}^{T}$*.*

Assuming that $\boldsymbol{A} + \boldsymbol{A}^{T} = \boldsymbol{B} + \boldsymbol{B}^{T}$*, it follows that* $\boldsymbol{z}^{T}\boldsymbol{A}\boldsymbol{z} = \frac{1}{2}\left(\boldsymbol{z}^{T}\boldsymbol{A}\boldsymbol{z} + \boldsymbol{z}^{T}\boldsymbol{A}^{T}\boldsymbol{z}\right) = \frac{1}{2}\boldsymbol{z}^{T}\left(\boldsymbol{A} + \boldsymbol{A}^{T}\right)\boldsymbol{z} = \frac{1}{2}\boldsymbol{z}^{T}\left(\boldsymbol{B} + \boldsymbol{B}^{T}\right)\boldsymbol{z} = \frac{1}{2}\left(\boldsymbol{z}^{T}\boldsymbol{B}\boldsymbol{z} + \boldsymbol{z}^{T}\boldsymbol{B}^{T}\boldsymbol{z}\right) = \frac{1}{2}\left(\boldsymbol{z}^{T}\boldsymbol{B}\boldsymbol{z} + \boldsymbol{z}^{T}\boldsymbol{B}\boldsymbol{z}\right) = \boldsymbol{z}^{T}\boldsymbol{B}\boldsymbol{z}$*, for all* $\boldsymbol{z} \in \mathbb{C}^{N\times 1}$*.* ■

Corollary 1: *Let* $\boldsymbol{A} \in \mathbb{C}^{N\times N}$*.* $\boldsymbol{z}^{T}\boldsymbol{A}\boldsymbol{z} = 0 \;\forall\; \boldsymbol{z} \in \mathbb{C}^{N\times 1}$ *is equivalent to* $\boldsymbol{A}^{T} = -\boldsymbol{A}$*.*

Proof: *Set* $\boldsymbol{B} = \boldsymbol{0}_{N\times N}$ *in Lemma* 4*, then the corollary follows.* ■

Lemma 5: *Let* $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{N\times N}$*.* $\boldsymbol{z}^{H}\boldsymbol{A}\boldsymbol{z} = \boldsymbol{z}^{H}\boldsymbol{B}\boldsymbol{z} \;\forall\; \boldsymbol{z} \in \mathbb{C}^{N\times 1}$ *is equivalent to* $\boldsymbol{A} = \boldsymbol{B}$*.*

Proof: *Let* $(\boldsymbol{A})_{k,l} = a_{k,l}$ *and* $(\boldsymbol{B})_{k,l} = b_{k,l}$*. Assume first that* $\boldsymbol{z}^{H}\boldsymbol{A}\boldsymbol{z} = \boldsymbol{z}^{H}\boldsymbol{B}\boldsymbol{z} \;\forall\; \boldsymbol{z} \in \mathbb{C}^{N\times 1}$*, and set* $\boldsymbol{z} = \boldsymbol{e}_k$ *where* $k \in \{0, 1, \ldots, N-1\}$*. This gives that* $a_{k,k} = b_{k,k}$ *for all* $k \in \{0, 1, \ldots, N-1\}$*. Also in the same way as in* *the proof of Lemma 4 setting* $\boldsymbol{z} = \boldsymbol{e}_k + \boldsymbol{e}_l$ *leads to* $\boldsymbol{A} + \boldsymbol{A}^T = \boldsymbol{B} + \boldsymbol{B}^T$*. Next, set* $\boldsymbol{z} = \boldsymbol{e}_k + \jmath \boldsymbol{e}_l$*, then manipulations give* $\boldsymbol{A} - \boldsymbol{A}^T = \boldsymbol{B} - \boldsymbol{B}^T$*. The equations* $\boldsymbol{A} + \boldsymbol{A}^T = \boldsymbol{B} + \boldsymbol{B}^T$ *and* $\boldsymbol{A} - \boldsymbol{A}^T = \boldsymbol{B} - \boldsymbol{B}^T$ *imply that* $\boldsymbol{A} = \boldsymbol{B}$*.*

If $\boldsymbol{A} = \boldsymbol{B}$*, then it follows that* $\boldsymbol{z}^H \boldsymbol{A} \boldsymbol{z} = \boldsymbol{z}^H \boldsymbol{B} \boldsymbol{z}$ *for all* $\boldsymbol{z} \in \mathbb{C}^{N\times 1}$*.* ■

Now, the proof of Lemma 2 is given:

Proof: *Substitution of* $d\,\mathrm{vec}(\boldsymbol{Z}) = d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\}) + \jmath d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\})$ *and* $d\,\mathrm{vec}(\boldsymbol{Z}^*) = d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\}) - \jmath d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\})$ *into the second-order differential expression given in the lemma leads to:*

$$\left[d\,\mathrm{vec}^T(\mathrm{Re}\{\boldsymbol{Z}\})\right]\left[\boldsymbol{B}_0 + \boldsymbol{B}_1 + \boldsymbol{B}_2\right] d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\}) + \left[d\,\mathrm{vec}^T(\mathrm{Im}\{\boldsymbol{Z}\})\right]\left[-\boldsymbol{B}_0 + \boldsymbol{B}_1 - \boldsymbol{B}_2\right] d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\})$$

$$+ \left[d\,\mathrm{vec}^T(\mathrm{Re}\{\boldsymbol{Z}\})\right]\left[\jmath\left(\boldsymbol{B}_0 + \boldsymbol{B}_0^T\right) + \jmath\left(\boldsymbol{B}_1 - \boldsymbol{B}_1^T\right) - \jmath\left(\boldsymbol{B}_2 + \boldsymbol{B}_2^T\right)\right] d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\}) = 0. \tag{57}$$

*(*57*) is valid for all* $d\boldsymbol{Z}$ *and the differentials of* $d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\})$ *and* $d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\})$ *are independent. If* $d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\})$ *is set to the zero vector, then it follows from (*57*) and Corollary* 1 *that:*

$$\boldsymbol{B}_0 + \boldsymbol{B}_1 + \boldsymbol{B}_2 = -\boldsymbol{B}_0^T - \boldsymbol{B}_1^T - \boldsymbol{B}_2^T. \tag{58}$$

In the same way, setting $d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\})$ *to the zero vector, then it follows from (*57*) and Corollary* 1 *that:*

$$-\boldsymbol{B}_0 + \boldsymbol{B}_1 - \boldsymbol{B}_2 = \boldsymbol{B}_0^T - \boldsymbol{B}_1^T + \boldsymbol{B}_2^T. \tag{59}$$

*Because of the skew symmetry in (*58*) and (*59*) and the linear independence of* $d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\})$ *and* $d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\})$*, it follows from (*57*) and Corollary* 1 *that:*

$$\left(\boldsymbol{B}_0 + \boldsymbol{B}_0^T\right) + \left(\boldsymbol{B}_1 - \boldsymbol{B}_1^T\right) - \left(\boldsymbol{B}_2 + \boldsymbol{B}_2^T\right) = \boldsymbol{0}_{NQ\times NQ}. \tag{60}$$

*(*58*), (*59*), and (*60*) lead to* $\boldsymbol{B}_0 = -\boldsymbol{B}_0^T$*,* $\boldsymbol{B}_1 = -\boldsymbol{B}_1^T$*, and* $\boldsymbol{B}_2 = -\boldsymbol{B}_2^T$*. Since the matrices* $\boldsymbol{B}_0$ *and* $\boldsymbol{B}_2$ *are skew, Corollary* 1 *reduces* $\left(d\,\mathrm{vec}^T(\boldsymbol{Z})\right)\boldsymbol{B}_0 d\,\mathrm{vec}(\boldsymbol{Z}) + \left(d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right)\boldsymbol{B}_1 d\,\mathrm{vec}(\boldsymbol{Z}) + \left(d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right)\boldsymbol{B}_2 d\,\mathrm{vec}(\boldsymbol{Z}^*) = 0$ *to* $\left(d\,\mathrm{vec}^T(\boldsymbol{Z}^*)\right)\boldsymbol{B}_1 d\,\mathrm{vec}(\boldsymbol{Z}) = 0$*. Then Lemma* 5 *results in* $\boldsymbol{B}_1 = \boldsymbol{0}_{NQ\times NQ}$*.* ■

APPENDIX IV

PROOF OF LEMMA 3

Proof: *Let the symbol* $\subseteq$ *means subset of. From (*7*) and the definition of* $\mathcal{R}(\boldsymbol{A})$*, it follows that* $\mathcal{R}(\boldsymbol{A}) = \left\{\boldsymbol{w} \in \mathbb{C}^{1\times Q} \mid \boldsymbol{w} = \boldsymbol{z}\boldsymbol{A}\left(\boldsymbol{A}^+\boldsymbol{A}\right),\ \text{for some } \boldsymbol{z} \in \mathbb{C}^{1\times N}\right\} \subseteq \mathcal{R}\left(\boldsymbol{A}^+\boldsymbol{A}\right)$*. From the definition of* $\mathcal{R}\left(\boldsymbol{A}^+\boldsymbol{A}\right)$*, it follows that* $\mathcal{R}\left(\boldsymbol{A}^+\boldsymbol{A}\right) = \left\{\boldsymbol{w} \in \mathbb{C}^{1\times Q} \mid \boldsymbol{w} = \boldsymbol{z}\boldsymbol{A}^+\boldsymbol{A},\ \text{for some } \boldsymbol{z} \in \mathbb{C}^{1\times Q}\right\} \subseteq \mathcal{R}(\boldsymbol{A})$*. From* $\mathcal{R}(\boldsymbol{A}) \subseteq \mathcal{R}\left(\boldsymbol{A}^+\boldsymbol{A}\right)$ *and*

$\mathcal{R}\left(\boldsymbol{A}^{+}\boldsymbol{A}\right) \subseteq \mathcal{R}\left(\boldsymbol{A}\right)$, *the first equality of (*33*) follows. From (*7*) and the definition of* $\mathcal{C}\left(\boldsymbol{A}\right)$, *it follows that:* $\mathcal{C}\left(\boldsymbol{A}\right) = \left\{\boldsymbol{w} \in \mathbb{C}^{N\times 1} \mid \boldsymbol{w} = \left(\boldsymbol{A}\boldsymbol{A}^{+}\right)\boldsymbol{A}\boldsymbol{z},\ \textit{for some } \boldsymbol{z} \in \mathbb{C}^{Q\times 1}\right\} \subseteq \mathcal{C}\left(\boldsymbol{A}\boldsymbol{A}^{+}\right)$. *From the definition of* $\mathcal{C}\left(\boldsymbol{A}\boldsymbol{A}^{+}\right)$, *it follows that:* $\mathcal{C}\left(\boldsymbol{A}\boldsymbol{A}^{+}\right) = \left\{\boldsymbol{w} \in \mathbb{C}^{N\times 1} \mid \boldsymbol{w} = \boldsymbol{A}\boldsymbol{A}^{+}\boldsymbol{z},\ \textit{for some } \boldsymbol{z} \in \mathbb{C}^{N\times 1}\right\} \subseteq \mathcal{C}\left(\boldsymbol{A}\right)$. *From* $\mathcal{C}\left(\boldsymbol{A}\right) \subseteq \mathcal{C}\left(\boldsymbol{A}\boldsymbol{A}^{+}\right)$ *and* $\mathcal{C}\left(\boldsymbol{A}\boldsymbol{A}^{+}\right) \subseteq \mathcal{C}\left(\boldsymbol{A}\right)$, *the second equality of (*33*) follows. The third equation of (*33*) is a direct consequence of the two first equations of (*33*).* ■

Appendix V

Lemma Showing Properties of the Moore-Penrose Inverse

Lemma 6: *Let* $\boldsymbol{A} \in \mathbb{C}^{N\times Q}$ *and* $\boldsymbol{B} \in \mathbb{C}^{Q\times R}$, *then the following properties are valid for the Moore-Penrose inverse:*

$$\boldsymbol{A}^{+} = \boldsymbol{A}^{-1} \textit{ for non-singular } \boldsymbol{A}, \tag{61}$$

$$\left(\boldsymbol{A}^{+}\right)^{+} = \boldsymbol{A}, \tag{62}$$

$$\left(\boldsymbol{A}^{H}\right)^{+} = \left(\boldsymbol{A}^{+}\right)^{H}, \tag{63}$$

$$\boldsymbol{A}^{H} = \boldsymbol{A}^{H}\boldsymbol{A}\boldsymbol{A}^{+} = \boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{A}^{H}, \tag{64}$$

$$\boldsymbol{A}^{+} = \boldsymbol{A}^{H}\left(\boldsymbol{A}^{+}\right)^{H}\boldsymbol{A}^{+} = \boldsymbol{A}^{+}\left(\boldsymbol{A}^{+}\right)^{H}\boldsymbol{A}^{H}, \tag{65}$$

$$\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{+} = \boldsymbol{A}^{+}\left(\boldsymbol{A}^{+}\right)^{H}, \tag{66}$$

$$\left(\boldsymbol{A}\boldsymbol{A}^{H}\right)^{+} = \left(\boldsymbol{A}^{+}\right)^{H}\boldsymbol{A}^{+}, \tag{67}$$

$$\boldsymbol{A}^{+} = \left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{+}\boldsymbol{A}^{H} = \boldsymbol{A}^{H}\left(\boldsymbol{A}\boldsymbol{A}^{H}\right)^{+}, \tag{68}$$

$$\boldsymbol{A}^{+} = \left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{H} \textit{ if } \boldsymbol{A} \textit{ has full column rank,} \tag{69}$$

$$\boldsymbol{A}^{+} = \boldsymbol{A}^{H}\left(\boldsymbol{A}\boldsymbol{A}^{H}\right)^{-1} \textit{ if } \boldsymbol{A} \textit{ has full row rank,} \tag{70}$$

$$\boldsymbol{A}\boldsymbol{B} = \boldsymbol{0}_{N\times R} \;\Leftrightarrow\; \boldsymbol{B}^{+}\boldsymbol{A}^{+} = \boldsymbol{0}_{R\times N}. \tag{71}$$

Proof: *(*61*), (*62*), and (*63*) can be proved by direct insertion in Definition* 3 *of the Moore-Penrose inverse.*

The first part of (64) can be proved like this $\boldsymbol{A}^H = \boldsymbol{A}^H \left(\boldsymbol{A}^H\right)^+ \boldsymbol{A}^H = \boldsymbol{A}^H \left(\boldsymbol{A}\boldsymbol{A}^+\right)^H = \boldsymbol{A}^H\boldsymbol{A}\boldsymbol{A}^+$, *where the result from (63) was used. The second part of (64) can be proved in a similar way* $\boldsymbol{A}^H = \boldsymbol{A}^H \left(\boldsymbol{A}^H\right)^+ \boldsymbol{A}^H = \left(\boldsymbol{A}^+\boldsymbol{A}\right)^H \boldsymbol{A}^H = \boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^H$. *The first part of (65) can be shown by* $\boldsymbol{A}^+ = \boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^+ = \left(\boldsymbol{A}^H\left(\boldsymbol{A}^H\right)^+\right)^H \boldsymbol{A}^+ = \boldsymbol{A}^H\left(\boldsymbol{A}^+\right)^H \boldsymbol{A}^+$. *The second part of (65) follows from* $\boldsymbol{A}^+ = \boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{A}^+\left(\left(\boldsymbol{A}^+\right)^H \boldsymbol{A}^H\right)^H = \boldsymbol{A}^+\left(\boldsymbol{A}^H\right)^+ \boldsymbol{A}^H$.

(66) and (67) can be proved by using the results from (64) and (65) in the definition of the Moore-Penrose inverse.

(68) follows from (65), (66), and (67).

(69) and (70) follow from (61), (68), and $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}\left(\boldsymbol{A}^H\boldsymbol{A}\right) = \operatorname{rank}\left(\boldsymbol{A}\boldsymbol{A}^H\right)$, *taken from 0.4.6 in [22].*

Now, (71) will be shown. Firstly, it is shown that $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{0}_{N\times R}$ *implies* $\boldsymbol{B}^+\boldsymbol{A}^+ = \boldsymbol{0}_{R\times N}$. *Assume that* $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{0}_{N\times R}$. *From (68), it follows that:*

$$\boldsymbol{B}^+\boldsymbol{A}^+ = \left(\boldsymbol{B}^H\boldsymbol{B}\right)^+ \boldsymbol{B}^H\boldsymbol{A}^H\left(\boldsymbol{A}\boldsymbol{A}^H\right)^+. \tag{72}$$

$\boldsymbol{A}\boldsymbol{B} = \boldsymbol{0}_{N\times R}$ *leads to* $\boldsymbol{B}^H\boldsymbol{A}^H = \boldsymbol{0}_{R\times N}$, *and then (72) yields* $\boldsymbol{B}^+\boldsymbol{A}^+ = \boldsymbol{0}_{R\times N}$. *Secondly, it will be shown that* $\boldsymbol{B}^+\boldsymbol{A}^+ = \boldsymbol{0}_{R\times N}$ *implies* $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{0}_{N\times R}$. *Assume that* $\boldsymbol{B}^+\boldsymbol{A}^+ = \boldsymbol{0}_{R\times N}$. *Using the implication just proved, i.e., if* $\boldsymbol{C}\boldsymbol{D} = \boldsymbol{0}_{M\times P}$, *then* $\boldsymbol{D}^+\boldsymbol{C}^+ = \boldsymbol{0}_{P\times M}$, *where* M *and* P *are integers given by the size of the matrices* $\boldsymbol{C}$ *and* $\boldsymbol{D}$, *gives* $\left(\boldsymbol{A}^+\right)^+\left(\boldsymbol{B}^+\right)^+ = \boldsymbol{0}_{N\times R}$, *and then the desired result follows from (62).* ■

REFERENCES

[1] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, United Kingdom: Cambridge University Press, May 2003.

[2] F. J. González-Vázquez, "The differentiation of functions of conjugate complex variables: Application to power network analysis," *IEEE Trans. Education*, vol. 31, no. 4, pp. 286–291, Nov. 1988.

[3] S. Alexander, "A derivation of the complex fast Kalman algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 32, no. 6, pp. 1230–1232, Dec. 1984.

[4] A. I. Hanna and D. P. Mandic, "A fully adaptive normalized nonlinear gradient descent algorithm for complex-valued nonlinear adaptive filters," *IEEE Trans. Signal Processing*, vol. 51, no. 10, pp. 2540–2549, Oct. 2003.

[5] E. Kreyszig, *Advanced Engineering Mathematics*, 7th ed. New York, USA: John Wiley & Sons, Inc., 1993.

[6] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits, Syst.*, vol. CAS-25, no. 9, pp. 772–781, Sept. 1978.

[7] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Application in Statistics and Econometrics*. Essex, England: John Wiley & Sons, Inc., 1988.

[8] D. A. Harville, *Matrix Algebra from a Statistician's Perspective*. Springer-Verlag, 1997, corrected second printing, 1999.

[9] T. P. Minka. (December 28, 2000) Old and new matrix algebra useful for statistics. [Online]. Available: http://research.microsoft.com/~minka/papers/matrix/

[10] K. Kreutz-Delgado, "Real vector derivatives and gradients," Dept. of Electrical and Computer Engineering, UC San Diego, Tech. Rep. Course Lecture Supplement No. ECE275A, December 5. 2005, `http://dsp.ucsd.edu/~kreutz/PEI05.html`.

[11] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, New Jersey, USA: Prentice Hall, 1991.

[12] A. H. Sayed, *Fundamentals of Adaptive Filtering*. IEEE Computer Society Press, 2003.

[13] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. John Wiley & Sons, Inc., 1996.

[14] A. van den Bos, "Complex gradient and Hessian," *Proc. IEE Vision, Image and Signal Processing*, vol. 141, no. 6, pp. 380–383, Dec. 1994.

[15] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEE Proc., Parts F and H*, vol. 130, no. 1, pp. 11–16, Feb. 1983.

[16] M. Brookes. Matrix reference manual. [Online]. Available: http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/intro.html

[17] K. Kreutz-Delgado, "The complex gradient operator and the $\mathbb{CR}$-calculus," Dept. of Electrical and Computer Engineering, UC San Diego, Tech. Rep. Course Lecture Supplement No. ECE275A, Sept.-Dec. 2005, `http://dsp.ucsd.edu/~kreutz/PEI05.html`.

[18] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 611–627, Mar. 2002.

[19] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[20] D. H. Jonhson and D. A. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Prentice-Hall, Inc., 1993.

[21] W. Wirtinger, "Zur formalen theorie der funktionen von mehr komplexen veränderlichen," *Mathematische Annalen*, vol. 97, pp. 357–375, 1927.

[22] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press Cambridge, UK, 1985, reprinted 1999.

[23] ——, *Topics in Matrix Analysis*. Cambridge University Press Cambridge, UK, 1991, reprinted 1999.

[24] D. G. Luenberger, *Introduction to Linear and Nonlinear Programming*. Reading, Massachusetts: Addison–Wesley, 1973.

[25] N. Young, *An Introduction to Hilbert Space*. The Pitt Building, Trumpington Street, Cambridge: Press Syndicate of the University of Cambridge, 1990, reprinted edition.

[26] C. H. Edwards and D. E. Penney, *Calculus and Analytic Geometry*, 2nd ed. Englewood Cliffs, New Jersey: Prentice-Hall, Inc, 1986.

[27] H. Sampath and A. Paulraj, "Linear precoding for space-time coded systems with known fading correlations," *IEEE Communications Letters*, vol. 6, no. 6, pp. 239–241, June 2002.

[28] A. Hjørungnes and D. Gesbert, "Precoding of orthogonal space-time block codes in arbitrarily correlated MIMO channels: Iterative and closed-form solutions," *IEEE Trans. Wireless Communications*, 2005, accepted 11.10.2005.

[29] T. J. Abatzoglou, J. M. Mendel, and G. A. Harada, "The constrained total least squares technique and its applications to harmonic superresolution," *IEEE Trans. Signal Processing*, vol. 39, no. 5, pp. 1070–1087, May 1991.

[30] G. Yan and H. Fan, "A Newton-like algorithm for complex variables with applications in blind equalization," *IEEE Trans. Signal Processing*, vol. 48, no. 2, pp. 553–556, Feb. 2000.