Complex-Valued Matrix Differentiation: Techniques and Key Results

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.

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

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. 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 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

Corresponding author: 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.

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 $\boldsymbol{F} \in \mathbb{C}^{M\times P}$	$\boldsymbol{F}(z, z^*)$ $\boldsymbol{F} : \mathbb{C} \times \mathbb{C} \to \mathbb{C}^{M\times P}$	$\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*)$ $\boldsymbol{F} : \mathbb{C}^{N\times 1} \times \mathbb{C}^{N\times 1} \to \mathbb{C}^{M\times P}$	$\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$ $\boldsymbol{F} : \mathbb{C}^{N\times Q} \times \mathbb{C}^{N\times Q} \to \mathbb{C}^{M\times P}$

contribution of this paper is to generalize the real-valued derivatives given in [1] to the complex-valued case. This is done by finding the derivatives by the so-called complex differentials of the functions. 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. 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 [2], [1], [3], [4], [5]. 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 [6], Appendix 2.B in [7], Subsection 2.3.10 of [8], and the article [9]. The article [10] 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 [11], [12] for differentiation with respect to complex-valued scalars and vectors, however, the more general *matrix* case is not considered. In [13], they find derivatives of scalar functions with respect to complex-valued matrices, however, that paper could have been simplified a lot if the proposed theory was utilized. Examples of problems where the unknown matrix is a complex-valued matrix are wide ranging including precoding of MIMO systems [14], linear equalization design [15], array signal processing [16] 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. For an extended version, see [17].

The rest of this paper is organized as follows: In Section II, 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 III. The key procedure showing how the derivatives can be found from the differential of a function is also presented in Section III. Section IV contains the important results of equivalent conditions for finding stationary points and in which direction the function has the maximum rate of change. In Section V, several key results are placed in tables and some results are derived for various cases with high relevance for signal processing and communication problems. Section VI contains some conclusions. Some of the proofs are given in the appendices.

Notation: 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 $\jmath = \sqrt{-1}$, and and let the real $\mathrm{Re}\{\cdot\}$ and imaginary $\mathrm{Im}\{\cdot\}$ operators return the real and imaginary parts of the input matrix, respectively. 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. 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-valued matrix function[2] $\boldsymbol{F}(\boldsymbol{Z}_0, \boldsymbol{Z}_1)$ is to calculate the difference

$$\boldsymbol{F}(\boldsymbol{Z}_0 + d\boldsymbol{Z}_0, \boldsymbol{Z}_1 + d\boldsymbol{Z}_1) - \boldsymbol{F}(\boldsymbol{Z}_0, \boldsymbol{Z}_1) = \text{First-order}(d\boldsymbol{Z}_0, d\boldsymbol{Z}_1) + \text{Higher-order}(d\boldsymbol{Z}_0, d\boldsymbol{Z}_1), \tag{1}$$

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 (1) 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 $d\boldsymbol{Z}_0\boldsymbol{Z}_1 = \boldsymbol{Z}_0 d\boldsymbol{Z}_1 + (d\boldsymbol{Z}_0)\boldsymbol{Z}_1$.

Let $\otimes$ and $\odot$ denote the Kronecker and Hadamard product [18], 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 $\mathrm{vec}(\cdot)$ stacks the columns vectors of the argument matrix into a long column vector in chronological order [18]. The differentiation rule of the reshaping operator $\mathrm{reshape}(\cdot)$ in Table II is valid for any *linear reshaping*[3] operator $\mathrm{reshape}(\cdot)$ of the matrix, and examples of such operators are the transpose $(\cdot)^T$ or $\mathrm{vec}(\cdot)$. Some of the basic differential results in Table II can be derived by means of (1), and others can be derived by generalizing some of the results found in [1], [4] to the complex differential case.

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

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

[3] 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 $\mathrm{reshape}(\cdot)$ performs a reshaping of its input argument.

TABLE II

IMPORTANT RESULTS FOR COMPLEX DIFFERENTIALS.

Function	$\boldsymbol{A}$	$a\boldsymbol{Z}$	$\boldsymbol{AZB}$	$\boldsymbol{Z}_0+\boldsymbol{Z}_1$	$\operatorname{Tr}\{\boldsymbol{Z}\}$	$\boldsymbol{Z}_0\boldsymbol{Z}_1$	$\boldsymbol{Z}_0\otimes\boldsymbol{Z}_1$
Differential	$\boldsymbol{0}$	$ad\boldsymbol{Z}$	$\boldsymbol{A}(d\boldsymbol{Z})\boldsymbol{B}$	$d\boldsymbol{Z}_0+d\boldsymbol{Z}_1$	$\operatorname{Tr}\{d\boldsymbol{Z}\}$	$(d\boldsymbol{Z}_0)\boldsymbol{Z}_1+\boldsymbol{Z}_0(d\boldsymbol{Z}_1)$	$(d\boldsymbol{Z}_0)\otimes\boldsymbol{Z}_1+\boldsymbol{Z}_0\otimes(d\boldsymbol{Z}_1)$
Function	$\boldsymbol{Z}^*$	$\boldsymbol{Z}^H$	$\det(\boldsymbol{Z})$	$\ln(\det(\boldsymbol{Z}))$	$\operatorname{reshape}(\boldsymbol{Z})$	$\boldsymbol{Z}_0\odot\boldsymbol{Z}_1$	$\boldsymbol{Z}^{-1}$
Differential	$(d\boldsymbol{Z})^*$	$(d\boldsymbol{Z})^H$	$\det(\boldsymbol{Z})\operatorname{Tr}\left\{\boldsymbol{Z}^{-1}d\boldsymbol{Z}\right\}$	$\operatorname{Tr}\left\{\boldsymbol{Z}^{-1}d\boldsymbol{Z}\right\}$	$\operatorname{reshape}(d\boldsymbol{Z})$	$(d\boldsymbol{Z}_0)\odot\boldsymbol{Z}_1+\boldsymbol{Z}_0\odot(d\boldsymbol{Z}_1)$	$-\boldsymbol{Z}^{-1}(d\boldsymbol{Z})\boldsymbol{Z}^{-1}$

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

Differential of the Moore-Penrose Inverse: The differential of the *real-valued* Moore-Penrose inverse can be found in in [1], [3], but the fundamental result of the complex-valued version is derived here.

Definition 1: *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 [19]:*

$$\left(\boldsymbol{Z}\boldsymbol{Z}^+\right)^H = \boldsymbol{Z}\boldsymbol{Z}^+, \qquad \left(\boldsymbol{Z}^+\boldsymbol{Z}\right)^H = \boldsymbol{Z}^+\boldsymbol{Z}, \qquad \boldsymbol{Z}\boldsymbol{Z}^+\boldsymbol{Z} = \boldsymbol{Z}, \qquad \boldsymbol{Z}^+\boldsymbol{Z}\boldsymbol{Z}^+ = \boldsymbol{Z}^+, \tag{2}$$

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

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

$$d\boldsymbol{Z}^+ = -\boldsymbol{Z}^+(d\boldsymbol{Z})\boldsymbol{Z}^+ + \boldsymbol{Z}^+(\boldsymbol{Z}^+)^H(d\boldsymbol{Z}^H)\left(\boldsymbol{I}_N - \boldsymbol{Z}\boldsymbol{Z}^+\right) + \left(\boldsymbol{I}_Q - \boldsymbol{Z}^+\boldsymbol{Z}\right)(d\boldsymbol{Z}^H)(\boldsymbol{Z}^+)^H\boldsymbol{Z}^+. \tag{3}$$

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

The following lemma is used to identify the first-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 $i \in \{0,1\}$.*

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

III. COMPUTATION 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 2: *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:*

$$d\operatorname{vec}(\boldsymbol{F}) = (\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F})\, d\operatorname{vec}(\boldsymbol{Z}) + (\mathcal{D}_{\boldsymbol{Z}^*}\boldsymbol{F})\, d\operatorname{vec}(\boldsymbol{Z}^*). \tag{4}$$

TABLE III

IDENTIFICATION TABLE FOR COMPLEX-VALUED DERIVATIVES.

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
$f(\boldsymbol{z}, \boldsymbol{z}^*)$	$df = \boldsymbol{a}_0 d\boldsymbol{z} + \boldsymbol{a}_1 d\boldsymbol{z}^*$	$\mathcal{D}_{\boldsymbol{z}} f(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{a}_0$	$\mathcal{D}_{\boldsymbol{z}^*} f(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{a}_1$	$1 \times N$
$f(\boldsymbol{Z}, \boldsymbol{Z}^*)$	$df = \mathrm{vec}^T(\boldsymbol{A}_0) d\,\mathrm{vec}(\boldsymbol{Z}) + \mathrm{vec}^T(\boldsymbol{A}_1) d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\mathcal{D}_{\boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \mathrm{vec}^T(\boldsymbol{A}_0)$	$\mathcal{D}_{\boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \mathrm{vec}^T(\boldsymbol{A}_1)$	$1 \times NQ$
$f(\boldsymbol{Z}, \boldsymbol{Z}^*)$	$df = \mathrm{Tr}\left\{\boldsymbol{A}_0^T d\boldsymbol{Z} + \boldsymbol{A}_1^T d\boldsymbol{Z}^*\right\}$	$\frac{\partial}{\partial \boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{A}_0$	$\frac{\partial}{\partial \boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{A}_1$	$N \times Q$
$\boldsymbol{f}(z, z^*)$	$d\boldsymbol{f} = \boldsymbol{b}_0 dz + \boldsymbol{b}_1 dz^*$	$\mathcal{D}_z \boldsymbol{f}(z, z^*) = \boldsymbol{b}_0$	$\mathcal{D}_{z^*} \boldsymbol{f}(z, z^*) = \boldsymbol{b}_1$	$M \times 1$
$\boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*)$	$d\boldsymbol{f} = \boldsymbol{B}_0 d\boldsymbol{z} + \boldsymbol{B}_1 d\boldsymbol{z}^*$	$\mathcal{D}_{\boldsymbol{z}} \boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{B}_0$	$\mathcal{D}_{\boldsymbol{z}^*} \boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{B}_1$	$M \times N$
$\boldsymbol{f}(\boldsymbol{Z}, \boldsymbol{Z}^*)$	$d\boldsymbol{f} = \boldsymbol{\beta}_0 d\,\mathrm{vec}(\boldsymbol{Z}) + \boldsymbol{\beta}_1 d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\mathcal{D}_{\boldsymbol{Z}} \boldsymbol{f}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{\beta}_0$	$\mathcal{D}_{\boldsymbol{Z}^*} \boldsymbol{f}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{\beta}_1$	$M \times NQ$
$\boldsymbol{F}(z, z^*)$	$d\,\mathrm{vec}(\boldsymbol{F}) = \boldsymbol{c}_0 dz + \boldsymbol{c}_1 dz^*$	$\mathcal{D}_z \boldsymbol{F}(z, z^*) = \boldsymbol{c}_0$	$\mathcal{D}_{z^*} \boldsymbol{F}(z, z^*) = \boldsymbol{c}_1$	$MP \times 1$
$\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*)$	$d\,\mathrm{vec}(\boldsymbol{F}) = \boldsymbol{C}_0 d\boldsymbol{z} + \boldsymbol{C}_1 d\boldsymbol{z}^*$	$\mathcal{D}_{\boldsymbol{z}} \boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{C}_0$	$\mathcal{D}_{\boldsymbol{z}^*} \boldsymbol{F}(\boldsymbol{z}, \boldsymbol{z}^*) = \boldsymbol{C}_1$	$MP \times N$
$\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$	$d\,\mathrm{vec}(\boldsymbol{F}) = \boldsymbol{\zeta}_0 d\,\mathrm{vec}(\boldsymbol{Z}) + \boldsymbol{\zeta}_1 d\,\mathrm{vec}(\boldsymbol{Z}^*)$	$\mathcal{D}_{\boldsymbol{Z}} \boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{\zeta}_0$	$\mathcal{D}_{\boldsymbol{Z}^*} \boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{\zeta}_1$	$MP \times NQ$

$\mathcal{D}_{\boldsymbol{Z}}\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$ *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* 2 *is a generalization of Definition 1, page 173 in [1] to include complex-valued matrices. In [1], several alternative definitions of the derivative of real-valued functions with respect to a matrix are discussed, and it is concluded that the definition that matches Definition* 2 *is the only reasonable definition. Definition* 2 *is also a generalization of the definition used in [10] 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.[4] To show the uniqueness of the representation in (4), we subtract the differential in (4) 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 [1], 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 3: *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*

$$\frac{\partial}{\partial \boldsymbol{z}^T}\boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*) = \begin{bmatrix} \frac{\partial}{\partial z_0} f_0 & \cdots & \frac{\partial}{\partial z_{N-1}} f_0 \\ \vdots & & \vdots \\ \frac{\partial}{\partial z_0} f_{M-1} & \cdots & \frac{\partial}{\partial z_{N-1}} f_{M-1} \end{bmatrix}, \quad \frac{\partial}{\partial \boldsymbol{z}^H}\boldsymbol{f}(\boldsymbol{z}, \boldsymbol{z}^*) = \begin{bmatrix} \frac{\partial}{\partial z_0^*} f_0 & \cdots & \frac{\partial}{\partial z_{N-1}^*} f_0 \\ \vdots & & \vdots \\ \frac{\partial}{\partial z_0^*} f_{M-1} & \cdots & \frac{\partial}{\partial z_{N-1}^*} f_{M-1} \end{bmatrix}, \tag{5}$$

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

[4]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 V-A.

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 3, the derivatives of the function $\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*)$, in Definition 2, are:

$$\mathcal{D}_{\boldsymbol{Z}} \boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \frac{\partial \operatorname{vec}(\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*))}{\partial \operatorname{vec}^T(\boldsymbol{Z})}, \tag{6}$$

$$\mathcal{D}_{\boldsymbol{Z}^*} \boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*) = \frac{\partial \operatorname{vec}(\boldsymbol{F}(\boldsymbol{Z}, \boldsymbol{Z}^*))}{\partial \operatorname{vec}^T(\boldsymbol{Z}^*)}. \tag{7}$$

This is a generalization of the real-valued matrix variable case treated in [1] to the complex-valued matrix variable case. (6) and (7) 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 3.

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* (4).
3) *Read out the result using Table* III.

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

Chain Rule: One big advantage of the way the derivative is defined in Definition 2 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, and it might be very useful for finding complicated derivatives.

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:*

$$\mathcal{D}_{\boldsymbol{Z}} \boldsymbol{H} = (\mathcal{D}_{\boldsymbol{F}} \boldsymbol{G}) (\mathcal{D}_{\boldsymbol{Z}} \boldsymbol{F}) + (\mathcal{D}_{\boldsymbol{F}^*} \boldsymbol{G}) (\mathcal{D}_{\boldsymbol{Z}} \boldsymbol{F}^*), \tag{8}$$

$$\mathcal{D}_{\boldsymbol{Z}^*} \boldsymbol{H} = (\mathcal{D}_{\boldsymbol{F}} \boldsymbol{G}) (\mathcal{D}_{\boldsymbol{Z}^*} \boldsymbol{F}) + (\mathcal{D}_{\boldsymbol{F}^*} \boldsymbol{G}) (\mathcal{D}_{\boldsymbol{Z}^*} \boldsymbol{F}^*). \tag{9}$$

Proof: *From Definition* 2, *it follows that* $d \operatorname{vec}(\boldsymbol{H}) = d \operatorname{vec}(\boldsymbol{G}) = (\mathcal{D}_{\boldsymbol{F}} \boldsymbol{G}) \, d \operatorname{vec}(\boldsymbol{F}) + (\mathcal{D}_{\boldsymbol{F}^*} \boldsymbol{G}) \, d \operatorname{vec}(\boldsymbol{F}^*)$. *The differentials of* $\operatorname{vec}(\boldsymbol{F})$ *and* $\operatorname{vec}(\boldsymbol{F}^*)$ *are given by:*

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

$$d \operatorname{vec}(\boldsymbol{F}^*) = (\mathcal{D}_{\boldsymbol{Z}} \boldsymbol{F}^*) \, d \operatorname{vec}(\boldsymbol{Z}) + (\mathcal{D}_{\boldsymbol{Z}^*} \boldsymbol{F}^*) \, d \operatorname{vec}(\boldsymbol{Z}^*). \tag{11}$$

By substituting the results from (10) *and* (11), *into the expression for* $d \operatorname{vec}(\boldsymbol{H})$, *using the definition of the derivatives with respect to* $\boldsymbol{Z}$ *and* $\boldsymbol{Z}^*$ *the theorem follows.* ■

IV. Complex Derivatives in Optimization Theory

In this section, two useful theorems are presented that exploit the theory introduced earlier. Both theorems are important when solving practical optimization problems involving differentiation with respect to a complex-valued matrix. These results include conditions for finding stationary points for a real-valued scalar function dependent on complex-valued matrices and in which direction the same type of function has the minimum or maximum rate of change, which might be used in the *steepest decent method*.

1) **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*[5] *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:*[6]

$$\mathcal{D}_{\boldsymbol{X}} g(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{0}_{1\times NQ} \ \ \wedge \ \ \mathcal{D}_{\boldsymbol{Y}} g(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{0}_{1\times NQ}, \tag{12}$$

$$\mathcal{D}_{\boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{0}_{1\times NQ}, \tag{13}$$

or

$$\mathcal{D}_{\boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{0}_{1\times NQ}. \tag{14}$$

Proof: *In optimization theory [1], 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, (*12*) 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:*

$$\left(\mathcal{D}_{\boldsymbol{Z}} f\right)\left(\mathcal{D}_{\boldsymbol{X}} \boldsymbol{Z}\right) + \left(\mathcal{D}_{\boldsymbol{Z}^*} f\right)\left(\mathcal{D}_{\boldsymbol{X}} \boldsymbol{Z}^*\right) = \mathcal{D}_{\boldsymbol{X}} g, \tag{15}$$

$$\left(\mathcal{D}_{\boldsymbol{Z}} f\right)\left(\mathcal{D}_{\boldsymbol{Y}} \boldsymbol{Z}\right) + \left(\mathcal{D}_{\boldsymbol{Z}^*} f\right)\left(\mathcal{D}_{\boldsymbol{Y}} \boldsymbol{Z}^*\right) = \mathcal{D}_{\boldsymbol{Y}} g. \tag{16}$$

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 (*15*) and (*16*), these two equations can be formulated into block matrix form in the following way:*

$$\begin{bmatrix} \mathcal{D}_{\boldsymbol{X}} g \\ \mathcal{D}_{\boldsymbol{Y}} g \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \jmath & -\jmath \end{bmatrix} \begin{bmatrix} \mathcal{D}_{\boldsymbol{Z}} f \\ \mathcal{D}_{\boldsymbol{Z}^*} f \end{bmatrix}. \tag{17}$$

This equation is equivalent to the following matrix equation:

$$\begin{bmatrix} \mathcal{D}_{\boldsymbol{Z}} f \\ \mathcal{D}_{\boldsymbol{Z}^*} f \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\jmath}{2} \\ \frac{1}{2} & \frac{\jmath}{2} \end{bmatrix} \begin{bmatrix} \mathcal{D}_{\boldsymbol{X}} g \\ \mathcal{D}_{\boldsymbol{Y}} g \end{bmatrix}. \tag{18}$$

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 (*18*), that (*12*), (*13*), and (*14*) are equivalent.* ■

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

[6] In (12), the symbol $\wedge$ means that both of the equations stated in (12) must be satisfied at the same time.

2) **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* $\text{vec}(\boldsymbol{Z})$ *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* $df = (\mathcal{D}_{\boldsymbol{Z}} f)\, d\,\text{vec}(\boldsymbol{Z}) + (\mathcal{D}_{\boldsymbol{Z}^*} f)\, d\,\text{vec}(\boldsymbol{Z}^*)$*, and* $df = df^* = (\mathcal{D}_{\boldsymbol{Z}} f)^*\, d\,\text{vec}(\boldsymbol{Z}^*) + (\mathcal{D}_{\boldsymbol{Z}^*} f)^*\, d\,\text{vec}(\boldsymbol{Z})$*, 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:* $df = (\mathcal{D}_{\boldsymbol{Z}} f)\, d\,\text{vec}(\boldsymbol{Z}) + ((\mathcal{D}_{\boldsymbol{Z}} f)\, d\,\text{vec}(\boldsymbol{Z}))^* = 2\,\text{Re}\{(\mathcal{D}_{\boldsymbol{Z}} f)\, d\,\text{vec}(\boldsymbol{Z})\} = 2\,\text{Re}\{(\mathcal{D}_{\boldsymbol{Z}^*} f)^*\, d\,\text{vec}(\boldsymbol{Z})\}$*. Let* $\boldsymbol{a}_i \in \mathbb{C}^{K\times 1}$*, where* $i \in \{0, 1\}$*. Then*

$$\text{Re}\left\{\boldsymbol{a}_0^H \boldsymbol{a}_1\right\} = \left\langle \begin{bmatrix} \text{Re}\{\boldsymbol{a}_0\} \\ \text{Im}\{\boldsymbol{a}_0\} \end{bmatrix}, \begin{bmatrix} \text{Re}\{\boldsymbol{a}_1\} \\ \text{Im}\{\boldsymbol{a}_1\} \end{bmatrix} \right\rangle, \tag{19}$$

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

$$df = 2 \left\langle \begin{bmatrix} \text{Re}\left\{(\mathcal{D}_{\boldsymbol{Z}^*} f)^T\right\} \\ \text{Im}\left\{(\mathcal{D}_{\boldsymbol{Z}^*} f)^T\right\} \end{bmatrix}, \begin{bmatrix} \text{Re}\{d\,\text{vec}(\boldsymbol{Z})\} \\ \text{Im}\{d\,\text{vec}(\boldsymbol{Z})\} \end{bmatrix} \right\rangle. \tag{20}$$

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

V. Development of Derivative Formulas

A. 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 [1] 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:

$$\frac{\partial}{\partial \boldsymbol{Z}} f = \begin{bmatrix} \frac{\partial}{\partial z_{0,0}} f & \cdots & \frac{\partial}{\partial z_{0,Q-1}} f \\ \vdots & & \vdots \\ \frac{\partial}{\partial z_{N-1,0}} f & \cdots & \frac{\partial}{\partial z_{N-1,Q-1}} f \end{bmatrix}, \quad \frac{\partial}{\partial \boldsymbol{Z}^*} f = \begin{bmatrix} \frac{\partial}{\partial z_{0,0}^*} f & \cdots & \frac{\partial}{\partial z_{0,Q-1}^*} f \\ \vdots & & \vdots \\ \frac{\partial}{\partial z_{N-1,0}^*} f & \cdots & \frac{\partial}{\partial z_{N-1,Q-1}^*} f \end{bmatrix}. \tag{21}$$

$\frac{\partial}{\partial \boldsymbol{Z}} f$ and $\frac{\partial}{\partial \boldsymbol{Z}^*} f$ are called the *gradient*[7] of f with respect to $\boldsymbol{Z}$ and $\boldsymbol{Z}^*$, respectively. (21) generalizes to the complex case of one of the ways to define the derivative of real-valued scalar functions with respect to real matrices in [1]. The way of arranging the partial derivatives in (21) is different than than the way given in (6) and (7). If $df = \text{vec}^T(\boldsymbol{A}_0) d\,\text{vec}(\boldsymbol{Z}) + \text{vec}^T(\boldsymbol{A}_1) d\,\text{vec}(\boldsymbol{Z}^*) = \text{Tr}\left\{\boldsymbol{A}_0^T d\boldsymbol{Z} + \boldsymbol{A}_1^T d\boldsymbol{Z}^*\right\}$, 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.

[7] The following notation also exists [6], [13] for the gradient $\nabla_{\boldsymbol{Z}} f \triangleq \frac{\partial}{\partial \boldsymbol{Z}^*} f$.

TABLE IV

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

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

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 $\mathcal{D}_{\boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*)$ and $\mathcal{D}_{\boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*)$ is $1 \times NQ$, so these two ways of organizing the partial derivatives are different. It can be shown, that $\mathcal{D}_{\boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \mathrm{vec}^T\left(\frac{\partial}{\partial \boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*)\right)$, and $\mathcal{D}_{\boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \mathrm{vec}^T\left(\frac{\partial}{\partial \boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*)\right)$. The steepest decent method can be formulated as $\boldsymbol{Z}_{k+1} = \boldsymbol{Z}_k + \mu \frac{\partial}{\partial \boldsymbol{Z}^*} f(\boldsymbol{Z}_k, \boldsymbol{Z}_k^*)$. The differentials of the simple eigenvalue λ and its complex conjugate λ^* at $\boldsymbol{Z}_0$ are derived in [1] and they are included in Table IV. One key example of derivatives is developed in the text below, while others useful results are stated in Table IV.

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 [15], and in an upper bound for the pair-wise error probability (PEP) [14].

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 $df = \det(\boldsymbol{P} + \boldsymbol{Z}\boldsymbol{Z}^H)\,\mathrm{Tr}\left\{(\boldsymbol{P} + \boldsymbol{Z}\boldsymbol{Z}^H)^{-1} d(\boldsymbol{Z}\boldsymbol{Z}^H)\right\} = \det(\boldsymbol{P} + \boldsymbol{Z}\boldsymbol{Z}^H)\,\mathrm{Tr}\left\{\boldsymbol{Z}^H\left(\boldsymbol{P} + \boldsymbol{Z}\boldsymbol{Z}^H\right)^{-1} d\boldsymbol{Z} + \boldsymbol{Z}^T(\boldsymbol{P} + \boldsymbol{Z}\boldsymbol{Z}^H)^{-T} d\boldsymbol{Z}^*\right\}$. 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 IV.

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

1) **Kronecker Product Related Problems***:* An objective functions which depends on the Kronecker product of the unknown complex-valued matrix is the PEP found in [14]. Let $\boldsymbol{K}_{N,Q}$ denote the *commutation matrix* [1]. Let $\boldsymbol{F} : \mathbb{C}^{N_0\times Q_0} \times \mathbb{C}^{N_1\times Q_1} \to \mathbb{C}^{N_0 N_1\times Q_0 Q_1}$ 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:

TABLE V

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

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

$d\,\mathrm{vec}(\boldsymbol{F})=\mathrm{vec}\left((d\boldsymbol{Z}_0)\otimes\boldsymbol{Z}_1\right)+\mathrm{vec}\left(\boldsymbol{Z}_0\otimes d\boldsymbol{Z}_1\right)$. From Theorem 3.10 in [1], it follows that

$$\begin{aligned}\mathrm{vec}\left((d\boldsymbol{Z}_0)\otimes\boldsymbol{Z}_1\right)&=\left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[(d\,\mathrm{vec}(\boldsymbol{Z}_0))\otimes\mathrm{vec}(\boldsymbol{Z}_1)\right]\\&=\left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\boldsymbol{I}_{N_0Q_0}\otimes\mathrm{vec}(\boldsymbol{Z}_1)\right]d\,\mathrm{vec}(\boldsymbol{Z}_0),\end{aligned}\tag{22}$$

and in a similar way it follows that: $\mathrm{vec}\left(\boldsymbol{Z}_0\otimes d\boldsymbol{Z}_1\right)=\left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\mathrm{vec}(\boldsymbol{Z}_0)\otimes\boldsymbol{I}_{N_1Q_1}\right]d\,\mathrm{vec}(\boldsymbol{Z}_1)$. Inserting the last two results into $d\,\mathrm{vec}(\boldsymbol{F})$ gives:

$$\begin{aligned}d\,\mathrm{vec}(\boldsymbol{F})=&\left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\boldsymbol{I}_{N_0Q_0}\otimes\mathrm{vec}(\boldsymbol{Z}_1)\right]d\,\mathrm{vec}(\boldsymbol{Z}_0)\\&+\left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\mathrm{vec}(\boldsymbol{Z}_0)\otimes\boldsymbol{I}_{N_1Q_1}\right]d\,\mathrm{vec}(\boldsymbol{Z}_1).\end{aligned}\tag{23}$$

Define the matrices $\boldsymbol{A}(\boldsymbol{Z}_1)$ and $\boldsymbol{B}(\boldsymbol{Z}_0)$ by $\boldsymbol{A}(\boldsymbol{Z}_1)\triangleq\left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\boldsymbol{I}_{N_0Q_0}\otimes\mathrm{vec}(\boldsymbol{Z}_1)\right]$, and $\boldsymbol{B}(\boldsymbol{Z}_0)=\left(\boldsymbol{I}_{Q_0}\otimes\boldsymbol{K}_{Q_1,N_0}\otimes\boldsymbol{I}_{N_1}\right)\left[\mathrm{vec}(\boldsymbol{Z}_0)\otimes\boldsymbol{I}_{N_1Q_1}\right]$. 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 $d\,\mathrm{vec}(\boldsymbol{F})=\boldsymbol{A}(\boldsymbol{Z}_1)d\,\mathrm{vec}(\boldsymbol{Z}_0)+\boldsymbol{B}(\boldsymbol{Z}_0)d\,\mathrm{vec}(\boldsymbol{Z}_1)$. From $d\,\mathrm{vec}(\boldsymbol{F})$, 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 V. In the table, $\mathrm{diag}(\cdot)$ returns the square diagonal matrix with the input column vector elements on the main diagonal [19] and zeros elsewhere.

2) **Moore-Penrose Inverse Related Problems***:* In pseudo-inverse matrix based receiver design, the Moore-Penrose inverse might appear [15]. 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 Proposition 1, depends on both $d\boldsymbol{Z}$ and $d\boldsymbol{Z}^*$. Using the $\mathrm{vec}(\cdot)$ operator on the differential of the Moore-Penrose inverse in Table II, in addition to Lemma 4.3.1 in [18] and the definition of the commutation matrix, result in:

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

$$+\left[\left(\boldsymbol{Z}^{+}\right)^{T}\left(\boldsymbol{Z}^{+}\right)^{*}\otimes\left(\boldsymbol{I}_{Q}-\boldsymbol{Z}^{+}\boldsymbol{Z}\right)\right]\boldsymbol{K}_{N,Q}d\,\mathrm{vec}(\boldsymbol{Z}^{*}). \tag{24}$$

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 V.

VI. 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: *(2) 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:*

$$d\boldsymbol{Z}^{+}=(d\boldsymbol{Z}^{+}\boldsymbol{Z})\boldsymbol{Z}^{+}+\boldsymbol{Z}^{+}(d\boldsymbol{Z}\boldsymbol{Z}^{+}-(d\boldsymbol{Z})\boldsymbol{Z}^{+})=(d\boldsymbol{Z}^{+}\boldsymbol{Z})\boldsymbol{Z}^{+}+\boldsymbol{Z}^{+}d\boldsymbol{Z}\boldsymbol{Z}^{+}-\boldsymbol{Z}^{+}(d\boldsymbol{Z})\boldsymbol{Z}^{+}. \tag{25}$$

It is seen from (25), 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:*

$$d\boldsymbol{Z}^{+}\boldsymbol{Z}=d\boldsymbol{Z}^{+}\boldsymbol{Z}\boldsymbol{Z}^{+}\boldsymbol{Z}=(d\boldsymbol{Z}^{+}\boldsymbol{Z})\boldsymbol{Z}^{+}\boldsymbol{Z}+\boldsymbol{Z}^{+}\boldsymbol{Z}(d\boldsymbol{Z}^{+}\boldsymbol{Z})=(\boldsymbol{Z}^{+}\boldsymbol{Z}(d\boldsymbol{Z}^{+}\boldsymbol{Z}))^{H}+\boldsymbol{Z}^{+}\boldsymbol{Z}(d\boldsymbol{Z}^{+}\boldsymbol{Z}). \tag{26}$$

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 (26), 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 (25), (3) 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* $\boldsymbol{A}_0 d\,\mathrm{vec}(\boldsymbol{Z})+\boldsymbol{A}_1 d\,\mathrm{vec}(\boldsymbol{Z}^{*})=\boldsymbol{0}_{M\times1}$*, then it follows that* $\boldsymbol{A}_0(d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\})+\jmath d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\}))+\boldsymbol{A}_1(d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\})-\jmath d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\}))=\boldsymbol{0}_{M\times1}$*. The last expression is equivalent to:* $(\boldsymbol{A}_0+\boldsymbol{A}_1)d\,\mathrm{vec}(\mathrm{Re}\{\boldsymbol{Z}\})+\jmath(\boldsymbol{A}_0-\boldsymbol{A}_1)d\,\mathrm{vec}(\mathrm{Im}\{\boldsymbol{Z}\})=\boldsymbol{0}_{M\times1}$*. 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* $\boldsymbol{A}_0 = \boldsymbol{A}_1 = \boldsymbol{0}_{M\times NQ}$*.* ■

References

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

[2] 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.

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

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

[5] 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`.

[6] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, New Jersey, USA: Prentice Hall, 2002.

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

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

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

[10] 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.

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

[12] 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`.

[13] D. P. Palomar and S. Verdú, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 141–154, Jan. 2006.

[14] 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.

[15] 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.

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

[17] A. Hjørungnes and D. Gesbert. Introduction to complex-valued matrix differentiation. [Online]. Available: http://www.unik.no/personer/arehj/publications/matrixdiff.pdf

[18] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press Cambridge, UK, 1991, reprinted 1999.

[19] ——, *Matrix Analysis*. Cambridge University Press Cambridge, UK, 1985, reprinted 1999.