

A Generalized Matrix Inverse that is Consistent with Respect to Diagonal Transformations

Jeffrey Uhlmann

University of Missouri-Columbia
 201 EBW, Columbia, MO 65211
 Email: uhlmannj@missouri.edu

Abstract

A new generalized matrix inverse is derived which is consistent with respect to arbitrary nonsingular diagonal transformations, e.g., it preserves units associated with variables under state space transformations, thus providing a general solution to a longstanding open problem relevant to a wide variety of applications in robotics, tracking, and control systems. The new inverse complements the Drazin inverse (which is consistent with respect to similarity transformations) and the Moore-Penrose inverse (which is consistent with respect to unitary/orthonormal transformations) to complete a trilogy of generalized matrix inverses that exhausts the standard family of analytically-important linear system transformations. Results are generalized to obtain unit-consistent and unit-invariant matrix decompositions and examples of their use are described.

Keywords: Drazin Inverse, Generalized Matrix Inverse, Inverse Problems, Linear Estimation, Linear Systems, Machine Learning, Matrix Analysis, Moore-Penrose Pseudoinverse, Scale Invariance, Singular Value Decomposition, SVD, System Design, Unit Consistency.

I. INTRODUCTION

For a nonsingular $n \times n$ matrix¹ \mathbf{A} there exists a unique matrix inverse, \mathbf{A}^{-1} , for which certain properties of scalar inverses are preserved, e.g., commutativity:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \quad (1)$$

while others have direct analogs, e.g., matrix inversion distributes over nonsingular multiplicands as:

$$(\mathbf{X}\mathbf{A}\mathbf{Y})^{-1} = \mathbf{Y}^{-1}\mathbf{A}^{-1}\mathbf{X}^{-1} \quad (2)$$

When attempting to generalize the notion of a matrix inverse for singular \mathbf{A} it is only possible to define an approximate inverse, $\tilde{\mathbf{A}}^{-1}$, which retains a subset of the algebraic properties of a true matrix inverse. For example, a generalized inverse definition might simply require the product $\mathbf{A}\tilde{\mathbf{A}}^{-1}$ to be idempotent in analogy to the identity matrix. Alternative definitions might further require:

$$\mathbf{A}\tilde{\mathbf{A}}^{-1}\mathbf{A} = \mathbf{A} \quad (3)$$

or

$$\tilde{\mathbf{A}}^{-1}\mathbf{A} = \tilde{\mathbf{A}}^{-1}\mathbf{A}\tilde{\mathbf{A}}^{-1} \quad (4)$$

and/or other properties that may be of analytic or application-specific utility.

The vast literature² on generalized inverse theory spans more than a century and can inform the decision about which of the many possible generalized inverses is best suited to the needs of a particular

¹It is assumed throughout that matrices are defined over an associative normed division algebra (real, complex, and quaternion). The inverse of a unitary matrix \mathbf{U} can therefore be expressed using the conjugate-transpose operator as \mathbf{U}^* .

²Much of the literature on generalized inverses is covered in the comprehensive book by Ben-Israel and Greville [3].

application. For example, the *Drazin* inverse, \mathbf{A}^D , satisfies the following for any square matrix \mathbf{A} and nonsingular matrix \mathbf{X} [8], [5], [3]:

$$\mathbf{A}^D \mathbf{A} \mathbf{A}^D = \mathbf{A}^D \quad (5)$$

$$\mathbf{A} \mathbf{A}^D = \mathbf{A}^D \mathbf{A} \quad (6)$$

$$(\mathbf{X} \mathbf{A} \mathbf{X}^{-1})^D = \mathbf{X} \mathbf{A}^D \mathbf{X}^{-1} \quad (7)$$

Thus it is applicable when there is need for commutativity (Eq.(6)) and/or consistency with respect to similarity transformations (Eq.(7)). On the other hand, the Drazin inverse is only defined for square matrices and does not guarantee the rank of \mathbf{A}^D to be the same as \mathbf{A} . Because the rank of \mathbf{A}^D may be less than that of \mathbf{A} (and in fact is zero for all nilpotent matrices), it is not appropriate for recursive control and estimation problems (and many other applications) that cannot accommodate progressive rank reduction.

The Moore-Penrose *pseudoinverse*, \mathbf{A}^P , is defined for any $m \times n$ matrix \mathbf{A} and satisfies conditions which include the following for any conformant unitary matrices \mathbf{U} and \mathbf{V} :

$$\text{rank}[\mathbf{A}^P] = \text{rank}[\mathbf{A}] \quad (8)$$

$$\mathbf{A} \mathbf{A}^P \mathbf{A} = \mathbf{A} \quad (9)$$

$$\mathbf{A}^P \mathbf{A} \mathbf{A}^P = \mathbf{A}^P \quad (10)$$

$$(\mathbf{U} \mathbf{A} \mathbf{V})^P = \mathbf{V}^* \mathbf{A}^P \mathbf{U}^* \quad (11)$$

Its use is therefore appropriate when there is need for unitary consistency, i.e., as guaranteed by Eq.(11). Despite its near-universal use throughout many areas of science and engineering ranging from tomography [4] to genomics analysis [2], the Moore-Penrose inverse is not appropriate for many problems to which it is commonly applied, e.g., state-space applications that require consistency with respect to the choice of units for state variables. In the case of a square singular transformation matrix \mathbf{A} , for example, a simple change of units applied to a set of state variables may require an inverse $\tilde{\mathbf{A}}^{-1}$ to be preserved under diagonal similarity

$$(\mathbf{D} \mathbf{A} \mathbf{D}^{-1})^{-1} = \mathbf{D} \tilde{\mathbf{A}}^{-1} \mathbf{D}^{-1} \quad (12)$$

where the nonsingular diagonal matrix \mathbf{D} defines an arbitrary change of units. The Moore-Penrose inverse does not satisfy this requirement because $(\mathbf{D} \mathbf{A} \mathbf{D}^{-1})^P$ does not generally equal $\mathbf{D} \mathbf{A}^P \mathbf{D}^{-1}$. As a concrete example, given

$$\mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & -1/2 \end{bmatrix} \quad (13)$$

it can be verified that

$$\mathbf{A}^P = \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & -1/2 \end{bmatrix} \quad (14)$$

and that

$$\mathbf{D} \mathbf{A}^P \mathbf{D}^{-1} = \begin{bmatrix} 1/2 & 1/4 \\ -1 & -1/2 \end{bmatrix} \quad (15)$$

which does not equal

$$(\mathbf{D} \mathbf{A} \mathbf{D}^{-1})^P = \begin{bmatrix} 0.32 & 0.64 \\ -0.16 & -0.32 \end{bmatrix}. \quad (16)$$

To appreciate the significance of unit consistency, consider the standard linear model

$$\hat{\mathbf{y}} = \mathbf{A} \cdot \hat{\boldsymbol{\theta}} \quad (17)$$

where the objective is to identify a vector $\hat{\boldsymbol{\theta}}$ of parameter values satisfying the above equation for a data

matrix \mathbf{A} and a known/desired state vector $\hat{\mathbf{y}}$. If \mathbf{A} is nonsingular then there exists a unique \mathbf{A}^{-1} which gives the solution

$$\hat{\boldsymbol{\theta}} = \mathbf{A}^{-1} \cdot \hat{\mathbf{y}} \quad (18)$$

If, however, \mathbf{A} is singular then the Moore-Penrose inverse could be applied as

$$\hat{\boldsymbol{\theta}} = \mathbf{A}^{\text{P}} \cdot \hat{\mathbf{y}} \quad (19)$$

to obtain a solution. Now suppose that $\hat{\mathbf{y}}$ and $\hat{\boldsymbol{\theta}}$ are expressed in different units as

$$\hat{\mathbf{y}}' = \mathbf{D}\hat{\mathbf{y}} \quad (20)$$

$$\hat{\boldsymbol{\theta}}' = \mathbf{E}\hat{\boldsymbol{\theta}} \quad (21)$$

where the diagonal matrices \mathbf{D} and \mathbf{E} represent changes of units, e.g., from imperial to metric, or rate-of-increase in liters-per-hour to a rate-of-decrease in liters-per-minute, or any other multiplicative change of units. Then Eq.(17) can be rewritten in the new units as

$$\hat{\mathbf{y}}' = \mathbf{D}\hat{\mathbf{y}} = (\mathbf{DAE}^{-1}) \cdot \mathbf{E}\hat{\boldsymbol{\theta}} = (\mathbf{DAE}^{-1}) \cdot \hat{\boldsymbol{\theta}}' \quad (22)$$

but for which

$$\mathbf{E}\hat{\boldsymbol{\theta}} \neq (\mathbf{DAE}^{-1})^{\text{P}} \cdot \hat{\mathbf{y}}' \quad (23)$$

In other words, the change of units applied to the input does not generally produce the same output in the new units. This is because the Moore-Penrose inverse only guarantees consistency with respect to unitary transformations (e.g., rotations) and not with respect to nonsingular diagonal transformations. To ensure unit consistency in this example a generalized matrix inverse $\tilde{\mathbf{A}}^{-1}$ would have to satisfy

$$(\mathbf{DAE}^{-1})^{\tilde{-1}} = \mathbf{E}\tilde{\mathbf{A}}^{-1}\mathbf{D}^{-1} \quad (24)$$

Stated more generally, if \mathbf{A} represents a mapping $V \rightarrow W$ from a vector space V to a vector space W then the inverse transformation $\tilde{\mathbf{A}}^{-1}$ must preserve consistency with respect to the application of arbitrary changes of units to the coordinates (state variables) associated with V and W .

Unit consistency (UC) has been suggested in the past as a critical consideration in specific applications (e.g., robotics [9], [7] and data fusion [22]), but the means for enforcing it have been limited because the most commonly applied tools in linear systems analysis, the eigen and singular-value decompositions, are inherently not unit consistent and therefore require UC alternatives. This may explain why in practice it is so common – *almost reflexive* – for an arbitrary criterion such as “least-squares” minimization (which is implicit when the Moore-Penrose inverse is used) to be applied without consideration for whether it is appropriate for the application at hand. In this paper the necessary analytical and practical tools to support unit consistency are developed.

The structure of the paper is as follows: Section II describes a simple and commonly-used (at least implicitly) mechanism for obtaining one-sided unit consistency. Section III develops a unit-consistent generalized inverse for elemental nonzero matrices, and Section IV develops the fully general unit-consistent generalized matrix inverse. Section V applies the techniques used to achieve unit consistency for the generalized inverse problem to develop unit-consistent and unit-invariant alternatives to the singular value decomposition (SVD) and other tools from linear algebra. Finally, Section VI summarizes and discusses the contributions of the paper.

II. LEFT AND RIGHT UC GENERALIZED INVERSES

Inverse consistency with respect to a nonsingular left diagonal transformation, $(\mathbf{DA})^{\text{L}} = \mathbf{A}^{\text{L}}\mathbf{D}^{-1}$, or a right nonsingular diagonal transformation, $(\mathbf{AD})^{\text{R}} = \mathbf{D}^{-1}\mathbf{A}^{\text{R}}$, is straightforward to obtain. The solution has been exploited implicitly in one form or another in many applications over the years; however, its

formal derivation and analysis is a useful exercise to establish concepts and notation that will be used later to derive the fully-general UC solution.

Definition II.1. Given an $m \times n$ matrix \mathbf{A} , a *left diagonal scale function*, $\mathcal{D}_L[\mathbf{A}] \in \mathbb{R}_+^{m \times m}$, is defined as giving a positive diagonal matrix satisfying the following for all conformant positive diagonal matrices \mathbf{D}_+ , unitary diagonals \mathbf{D}_u , permutations \mathbf{P} , and unitaries \mathbf{U} :

$$\mathcal{D}_L[\mathbf{D}_+\mathbf{A}] \cdot (\mathbf{D}_+\mathbf{A}) = \mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A}, \quad (25)$$

$$\mathcal{D}_L[\mathbf{D}_u\mathbf{A}] = \mathcal{D}_L[\mathbf{A}], \quad (26)$$

$$\mathcal{D}_L[\mathbf{P}\mathbf{A}] = \mathbf{P} \cdot \mathcal{D}_L[\mathbf{A}] \cdot \mathbf{P}^T, \quad (27)$$

$$\mathcal{D}_L[\mathbf{A}\mathbf{U}] = \mathcal{D}_L[\mathbf{A}] \quad (28)$$

In other words, the product $\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A}$ is invariant with respect to any positive left-diagonal scaling of \mathbf{A} , and $\mathcal{D}_L[\mathbf{A}]$ is consistent with respect to any left-permutation³ of \mathbf{A} and is invariant with respect to left-multiplication by any diagonal unitary and/or right-multiplication by any general unitary.

Lemma II.2. Existence of a left-diagonal scale function according to Definition II.1 is established by instantiating $\mathcal{D}_L[\mathbf{A}] = \mathbf{D}$ with

$$\mathbf{D}(i, i) \doteq \begin{cases} 1/\|\mathbf{A}(i, :)\| & \|\mathbf{A}(i, :)\| > 0 \\ 1 & \text{otherwise} \end{cases} \quad (29)$$

where $\mathbf{A}(i, :)$ is row i of \mathbf{A} and $\|\cdot\|$ is a fixed unitary-invariant vector norm⁴.

Proof. $\mathcal{D}_L[\cdot]$ as defined by Lemma II.2 is a strictly positive diagonal as required, and the left scale-invariance condition of Eq.(25) holds trivially for any row of \mathbf{A} with all elements equal to zero and holds for every nonzero row i by homogeneity for any choice of vector norm as

$$\mathbf{D}_+(i, i)\mathbf{A}(i, :) / \|\mathbf{D}_+(i, i)\mathbf{A}(i, :)\| = \mathbf{D}_+(i, i)\mathbf{A}(i, :) / (\mathbf{D}_+(i, i) \cdot \|\mathbf{A}(i, :)\|) \quad (30)$$

$$= \mathbf{A}(i, :) / \|\mathbf{A}(i, :)\|. \quad (31)$$

The left diagonal-unitary-invariance condition of Eq.(26) is satisfied as $|\mathbf{D}_u(i, i)| = 1$ implies

$$|(\mathbf{D}_u\mathbf{A})(i, j)| = |\mathbf{A}(i, j)| \quad (32)$$

for every element j of row i of $\mathbf{D}_u\mathbf{A}$. The left permutation-invariance of Eq.(27) holds as element $\mathbf{D}(i, i)$ is indexed with respect to the rows of \mathbf{A} , and the right unitary-invariance condition of Eq.(28) is satisfied by the assumed unitary invariance of the vector norm applied to the rows of \mathbf{A} . \square

If \mathbf{A} has *full support*, i.e., no row or column with all elements equal to zero, then $\mathcal{D}_L[\mathbf{D}_+\mathbf{A}] = \mathbf{D}_+^{-1} \cdot \mathcal{D}_L[\mathbf{A}]$. If, however, there exists a row i of \mathbf{A} with all elements equal to zero then the i th diagonal element of $\mathcal{D}_L[\mathbf{D}_+\mathbf{A}]$ is 1 according to Lemma II.2, so the corresponding element of $\mathbf{D}_+^{-1} \cdot \mathcal{D}_L[\mathbf{A}]$ will be different unless $\mathbf{D}_+^{-1}(i, i) = 1$. Eq.(25) holds because such elements are only applied to scale rows of \mathbf{A} with all elements equal to zero. The following similarly holds in general

$$\mathcal{D}_L[\mathbf{D}_+\mathbf{A}] \cdot \mathbf{A} = \mathbf{D}_+^{-1} \cdot \mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A}, \quad (33)$$

and because any row i of zeros in \mathbf{A} implies that column i of \mathbf{A}^P will be zeros, the following also holds in general

$$\mathbf{A}^P \cdot \mathcal{D}_L[\mathbf{D}_+\mathbf{A}] = \mathbf{A}^P \cdot \mathbf{D}_+^{-1} \cdot \mathcal{D}_L[\mathbf{A}]. \quad (34)$$

³Note: Eq.(27) here includes a \mathbf{P}^T term errantly missing in the final journal version.

⁴The unitary-invariant norm used here is necessary only because of the imposed right-invariant condition of Eq.(28).

At this point it is possible to derive a left generalized inverse of an arbitrary $m \times n$ matrix \mathbf{A} , denoted \mathbf{A}^{\perp} , that is consistent with respect to multiplication on the left by an arbitrary nonsingular diagonal matrix.

Theorem II.3. For $m \times n$ matrix \mathbf{A} , the operator

$$\mathbf{A}^{\perp} \doteq (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^{\text{p}} \cdot \mathcal{D}_L[\mathbf{A}] \quad (35)$$

satisfies for any nonsingular diagonal matrix \mathbf{D} :

$$\mathbf{A}\mathbf{A}^{\perp}\mathbf{A} = \mathbf{A}, \quad (36)$$

$$\mathbf{A}^{\perp}\mathbf{A}\mathbf{A}^{\perp} = \mathbf{A}^{\perp}, \quad (37)$$

$$(\mathbf{D}\mathbf{A})^{\perp} = \mathbf{A}^{\perp}\mathbf{D}^{-1}, \quad (38)$$

$$\text{rank}[\mathbf{A}^{\perp}] = \text{rank}[\mathbf{A}] \quad (39)$$

and is therefore a left unit-consistent generalized inverse.

Proof. The first two generalized inverse properties can be established from the corresponding properties of the Moore-Penrose inverse as:

$$\mathbf{A}\mathbf{A}^{\perp}\mathbf{A} = \mathbf{A} \cdot \{(\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^{\text{p}} \cdot \mathcal{D}_L[\mathbf{A}]\} \cdot \mathbf{A} \quad (40)$$

$$= \mathbf{A} \cdot \{(\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^{\text{p}} \cdot (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})\} \quad (41)$$

$$= (\mathcal{D}_L[\mathbf{A}]^{-1} \cdot \mathcal{D}_L[\mathbf{A}]) \cdot \mathbf{A} \cdot \{(\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^{\text{p}} \cdot (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})\} \quad (42)$$

$$= \mathcal{D}_L[\mathbf{A}]^{-1} \cdot \left\{ \underline{(\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A}) \cdot (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^{\text{p}} \cdot (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})} \right\} \quad (43)$$

$$= \mathcal{D}_L[\mathbf{A}]^{-1} \cdot (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A}) \quad (44)$$

$$= \mathbf{A} \quad (45)$$

and

$$\mathbf{A}^{\perp}\mathbf{A}\mathbf{A}^{\perp} = \{(\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^{\text{p}} \cdot \mathcal{D}_L[\mathbf{A}]\} \cdot \mathbf{A} \cdot \{(\mathcal{D}_L[\mathbf{A}]^{-1}\mathbf{A})^{\text{p}} \cdot \mathcal{D}_L[\mathbf{A}]\} \quad (46)$$

$$= \left\{ \underline{(\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^{\text{p}} \cdot (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A}) \cdot (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^{\text{p}}} \right\} \cdot \mathcal{D}_L[\mathbf{A}] \quad (47)$$

$$= (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^{\text{p}} \cdot \mathcal{D}_L[\mathbf{A}] \quad (48)$$

$$= \mathbf{A}^{\perp} \quad (49)$$

The left unit-consistency condition $(\mathbf{D}\mathbf{A})^{\perp} = \mathbf{A}^{\perp}\mathbf{D}^{-1}$, for any nonsingular diagonal matrix \mathbf{D} , can be established using a polar decomposition $\mathbf{D} = \mathbf{D}_+ \mathbf{D}_u$:

$$\mathbf{D}_+ = \text{Abs}[\mathbf{D}] \quad (50)$$

$$\mathbf{D}_u = \mathbf{D}\mathbf{D}_+^{-1} \quad (51)$$

and exploiting unitary-consistency of the Moore-Penrose inverse, i.e., $(\mathbf{U}\mathbf{A})^{\text{p}} = \mathbf{A}^{\text{p}}\mathbf{U}^*$, and commutativity of $\mathcal{D}_L[\cdot]$ with other diagonal matrices:

$$(\mathbf{DA})^\perp = (\mathcal{D}_L[\mathbf{DA}] \cdot \mathbf{DA})^p \cdot \mathcal{D}_L[\mathbf{DA}] \quad (52)$$

$$= (\mathcal{D}_L[\mathbf{D}_+ \mathbf{D}_u \mathbf{A}] \cdot \mathbf{D}_+ \mathbf{D}_u \mathbf{A})^p \cdot \mathcal{D}_L[\mathbf{D}_+ \mathbf{D}_u \mathbf{A}] \quad (53)$$

$$= (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{D}_+^{-1} \cdot \mathbf{D}_+ \mathbf{D}_u \mathbf{A})^p \cdot \mathbf{D}_+^{-1} \cdot \mathcal{D}_L[\mathbf{A}] \quad (54)$$

$$= (\mathcal{D}_L[\mathbf{A}] \cdot (\mathbf{D}_+^{-1} \mathbf{D}_+) \cdot \mathbf{D}_u \mathbf{A})^p \cdot \mathbf{D}_+^{-1} \cdot \mathcal{D}_L[\mathbf{A}] \quad (55)$$

$$= (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{D}_u \mathbf{A})^p \cdot \mathbf{D}_+^{-1} \cdot \mathcal{D}_L[\mathbf{A}] \quad (56)$$

$$= (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^p \cdot \mathbf{D}_u^* \cdot \mathbf{D}_+^{-1} \cdot \mathcal{D}_L[\mathbf{A}] \quad (57)$$

$$= (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^p \cdot (\mathbf{D}_u^* \mathbf{D}_+^{-1}) \cdot \mathcal{D}_L[\mathbf{A}] \quad (58)$$

$$= (\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^p \cdot \mathbf{D}^{-1} \cdot \mathcal{D}_L[\mathbf{A}] \quad (59)$$

$$= \{(\mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A})^p \cdot \mathcal{D}_L[\mathbf{A}]\} \cdot \mathbf{D}^{-1} \quad (60)$$

$$= \mathbf{A}^\perp \mathbf{D}^{-1}. \quad (61)$$

Lastly, the rank-consistency condition, $\text{rank}[\mathbf{A}^\perp] = \text{rank}[\mathbf{A}]$, is satisfied as every operation performed according to Lemma II.2 preserves the rank of the original matrix. In particular, the rank consistency of \mathbf{A}^\perp derives from the fact that $\text{rank}[\mathbf{A}^p] = \text{rank}[\mathbf{A}]$. \square

A right unit-consistent generalized inverse clearly can be derived analogously or in terms of the already-defined left operator as

$$\mathbf{A}^R \doteq ((\mathbf{A}^T)^\perp)^T. \quad (62)$$

In terms of the linear model of Eq.(17) for determining values for parameters $\hat{\theta}$,

$$\begin{aligned} \hat{\mathbf{y}} &= \mathbf{A} \cdot \hat{\theta} \\ &\Downarrow \\ \hat{\theta} &= \mathbf{A}^{\sim} \cdot \hat{\mathbf{y}} \end{aligned}$$

the inverse \mathbf{A}^{\sim} could be instantiated with *either* \mathbf{A}^\perp or \mathbf{A}^R to provide, respectively, consistency with respect to the application of a change of units to $\hat{\mathbf{y}}$ or a change of units to $\hat{\theta}$ – *but not both*.

III. UC GENERALIZED INVERSE FOR ELEMENTAL-NONZERO MATRICES

The derivations of separate left and right UC inverses from the previous section cannot be applied to achieve general unit consistency, i.e., to obtain a UC generalized inverse \mathbf{A}^U which satisfies

$$(\mathbf{DAE})^U = \mathbf{E}^{-1} \mathbf{A}^U \mathbf{D}^{-1} \quad (63)$$

for arbitrary nonsingular diagonals \mathbf{D} and \mathbf{E} . However, a *joint* characterization of the left and right diagonal transformations can provide a basis for doing so.

Lemma III.1. The transformation of an $m \times n$ matrix \mathbf{A} as \mathbf{DAE} , with $m \times m$ diagonal \mathbf{D} and $n \times n$ diagonal \mathbf{E} , is equivalent to a Hadamard (elementwise) matrix product $\mathbf{X} \circ \mathbf{A}$ for some rank-1 matrix \mathbf{X} .

Proof. Letting $\mathbf{d}_m = \text{Diag}[\mathbf{D}]$ and $\mathbf{e}_n = \text{Diag}[\mathbf{E}]$, the matrix product \mathbf{DAE} can be expressed as

$$\mathbf{DAE} = (\mathbf{d}_m \mathbf{1}_n^T) \circ \mathbf{A} \circ (\mathbf{1}_m \mathbf{e}_n^T) \quad (64)$$

$$= \{(\mathbf{d}_m \mathbf{1}_n^T) \circ (\mathbf{1}_m \mathbf{e}_n^T)\} \circ \mathbf{A} \quad (65)$$

$$= (\mathbf{d}_m \mathbf{e}_n^T) \circ \mathbf{A} \quad (66)$$

where $\mathbf{1}_n^T$ is a row vector of n ones and $\mathbf{1}_m$ is a column vector of m ones. Letting $\mathbf{X} = \mathbf{d}_m \mathbf{e}_n^T$ completes the proof. \square

Definition III.2. For an $m \times n$ matrix \mathbf{A} , left and right *general-diagonal scale functions* $\mathcal{D}_U^L[\mathbf{A}] \in \mathbb{R}_+^{m \times m}$ and $\mathcal{D}_U^R[\mathbf{A}] \in \mathbb{R}_+^{n \times n}$ are defined as jointly satisfying the following for all conformant positive diagonal matrices \mathbf{D}_+ and \mathbf{E}_+ , unitary diagonals \mathbf{D}_u and \mathbf{D}_v , and permutations \mathbf{P} and \mathbf{Q} :

$$\mathcal{D}_U^L[\mathbf{D}_+ \mathbf{A} \mathbf{E}_+] \cdot (\mathbf{D}_+ \mathbf{A} \mathbf{E}_+) \cdot \mathcal{D}_U^R[\mathbf{D}_+ \mathbf{A} \mathbf{E}_+] = \mathcal{D}_U^L[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_U^R[\mathbf{A}] \quad (67)$$

$$\mathcal{D}_U^L[\mathbf{P} \mathbf{A} \mathbf{Q}] \cdot (\mathbf{P} \mathbf{A} \mathbf{Q}) \cdot \mathcal{D}_U^R[\mathbf{P} \mathbf{A} \mathbf{Q}] = \mathbf{P} \cdot \{\mathcal{D}_U^L[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_U^R[\mathbf{A}]\} \cdot \mathbf{Q} \quad (68)$$

$$\mathcal{D}_U^L[\mathbf{D}_u \mathbf{A} \mathbf{D}_v] = \mathcal{D}_U^L[\mathbf{A}], \quad (69)$$

$$\mathcal{D}_U^R[\mathbf{D}_u \mathbf{A} \mathbf{D}_v] = \mathcal{D}_U^R[\mathbf{A}] \quad (70)$$

The function $\mathcal{S}_U[\mathbf{A}]$ is defined to be the rank-1 matrix guaranteed by Lemma III.1

$$\mathcal{S}_U[\mathbf{A}] \circ \mathbf{A} \equiv \mathcal{D}_U^L[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_U^R[\mathbf{A}] \quad (71)$$

i.e.,

$$\mathcal{S}_U[\mathbf{A}] = \text{Diag}[\mathcal{D}_U^L[\mathbf{A}]] \cdot \text{Diag}[\mathcal{D}_U^R[\mathbf{A}]]^T \quad (72)$$

Definition III.3. A matrix \mathbf{A} is defined to be an *elemental-nonzero matrix* if and only if it does not have any element equal to zero.

The following lemma uses the elementwise matrix functions $\text{LogAbs}[\cdot]$ and $\text{Exp}[\cdot]$, where $\text{LogAbs}[\mathbf{A}]$ represents the result of taking the logarithm of the magnitude of each element of \mathbf{A} and $\text{Exp}[\mathbf{A}]$ represents the taking of the exponential of every element of \mathbf{A} .

Lemma III.4. Existence of a general diagonal scale function according to Definition III.2 for arguments without zero elements is established by instantiating $\mathcal{D}_U^L[\mathbf{A}]$ and $\mathcal{D}_U^R[\mathbf{A}]$ as

$$\mathcal{D}_U^L[\mathbf{A}] = \text{Diag}[\mathbf{x}_m] \quad (73)$$

$$\mathcal{D}_U^R[\mathbf{A}] = \text{Diag}[\mathbf{y}_n] \quad (74)$$

for

$$\mathbf{x}_m \cdot \mathbf{y}_n^T = \mathcal{S}_U[\mathbf{A}] = \text{Exp}[\mathbf{J}_m \mathbf{L} \mathbf{J}_n - \mathbf{L} \mathbf{J}_n - \mathbf{J}_m \mathbf{L}] \quad (75)$$

where $\mathbf{L} = \text{LogAbs}[\mathbf{A}]$ and \mathbf{J}_m has all elements equal to $1/m$ and \mathbf{J}_n has all elements equal to $1/n$.

Proof. First it must be shown that $\text{Exp}[\mathbf{J}_m \mathbf{L} \mathbf{J}_n - \mathbf{L} \mathbf{J}_n - \mathbf{J}_m \mathbf{L}]$ is a rank-1 matrix. This can be achieved by expanding as

$$\mathbf{J}_m \mathbf{L} \mathbf{J}_n - \mathbf{L} \mathbf{J}_n - \mathbf{J}_m \mathbf{L} = \left(\frac{1}{2} \mathbf{J}_m \mathbf{L} \mathbf{J}_n - \mathbf{L} \mathbf{J}_n\right) + \left(\frac{1}{2} \mathbf{J}_m \mathbf{L} \mathbf{J}_n - \mathbf{J}_m \mathbf{L}\right) \quad (76)$$

$$= \left(\frac{1}{2} \mathbf{J}_m \mathbf{L} - \mathbf{L}\right) \mathbf{J}_n + \mathbf{J}_m \left(\frac{1}{2} \mathbf{L} \mathbf{J}_n - \mathbf{L}\right) \quad (77)$$

$$= \mathbf{u}_m \mathbf{1}_n^T + \mathbf{1}_m \mathbf{v}_n^T \quad (78)$$

where

$$\mathbf{u}_m = \frac{1}{n} \left(\frac{1}{2} \mathbf{J}_m - \mathbf{I}_m\right) \mathbf{L} \cdot \mathbf{1}_n \quad (79)$$

$$\mathbf{v}_n^T = \mathbf{1}_m^T \cdot \mathbf{L} \left(\frac{1}{2} \mathbf{J}_n - \mathbf{I}_n\right) / m \quad (80)$$

and then noting that the elementwise exponential of $\mathbf{u}_m \mathbf{1}_n^T + \mathbf{1}_m \mathbf{v}_n^T$ is the strictly positive rank-1 matrix $\text{Exp}[\mathbf{u}_m] \cdot \text{Exp}[\mathbf{v}_n^T]$, i.e., $\mathcal{D}_U^L[\mathbf{A}] = \text{Diag}[\text{Exp}[\mathbf{u}_m]]$ and $\mathcal{D}_U^R[\mathbf{A}] = \text{Diag}[\text{Exp}[\mathbf{v}_m]]$, which confirms existence and strict positivity as required. Eq.(67) can be established by observing that

$$\mathcal{D}_U^L[\mathbf{D}_+ \mathbf{A} \mathbf{E}_+] \cdot (\mathbf{D}_+ \mathbf{A} \mathbf{E}_+) \cdot \mathcal{D}_U^R[\mathbf{D}_+ \mathbf{A} \mathbf{E}_+] \equiv \mathcal{S}_U[\mathbf{D}_+ \mathbf{A} \mathbf{E}_+] \circ (\mathbf{D}_+ \mathbf{A} \mathbf{E}_+) \quad (81)$$

and letting $\mathbf{d}_m = \text{Diag}[\mathbf{D}_+]$, $\mathbf{e}_n = \text{Diag}[\mathbf{E}_+]$, $\mathbf{u}_m = \text{LogAbs}[\mathbf{d}_m]$, $\mathbf{v}_n = \text{LogAbs}[\mathbf{e}_m]$, and $\mathbf{L} = \text{LogAbs}[\mathbf{A}]$:

$$\mathcal{S}_U[\mathbf{D}_+ \mathbf{A} \mathbf{E}_+] = \mathcal{S}_U[(\mathbf{d}_m \mathbf{1}_n^T) \circ \mathbf{A} \circ (\mathbf{1}_m \mathbf{e}_n^T)] \quad (82)$$

$$= \mathcal{S}_U[\text{Exp}[\mathbf{u}_m \mathbf{1}_n^T] \circ \mathbf{A} \circ \text{Exp}[\mathbf{1}_m \mathbf{v}_n^T]] \quad (83)$$

$$= \text{Exp}[\mathbf{J}_m (\mathbf{u}_m \mathbf{1}_n^T + \mathbf{L} + \mathbf{1}_m \mathbf{v}_n^T) \mathbf{J}_n \quad (84)$$

$$- (\mathbf{u}_m \mathbf{1}_n^T + \mathbf{L} + \mathbf{1}_m \mathbf{v}_n^T) \mathbf{J}_n \quad (85)$$

$$- \mathbf{J}_m (\mathbf{u}_m \mathbf{1}_n^T + \mathbf{L} + \mathbf{1}_m \mathbf{v}_n^T)] \quad (86)$$

$$= \text{Exp}[(\mathbf{J}_m \cdot \mathbf{u}_m \mathbf{1}_n^T + \mathbf{J}_m \mathbf{L} \mathbf{J}_n + \mathbf{1}_m \mathbf{v}_n^T \cdot \mathbf{J}_n) \quad (87)$$

$$- (\mathbf{u}_m \mathbf{1}_n^T + \mathbf{L} \mathbf{J}_n + \mathbf{1}_m \mathbf{v}_n^T \cdot \mathbf{J}_n) \quad (88)$$

$$- (\mathbf{J}_m \cdot \mathbf{u}_m \mathbf{1}_n^T + \mathbf{J}_m \mathbf{L} + \mathbf{1}_m \mathbf{v}_n^T)] \quad (89)$$

$$= \text{Exp}[(-\mathbf{u}_m \mathbf{1}_n^T) + (\mathbf{J}_m \mathbf{L} \mathbf{J}_n - \mathbf{L} \mathbf{J}_n - \mathbf{J}_m \mathbf{L}) + (-\mathbf{1}_m \mathbf{v}_n^T)] \quad (90)$$

$$= \text{Exp}[-\mathbf{u}_m \mathbf{1}_n^T] \circ \mathcal{S}_U[\mathbf{A}] \circ \text{Exp}[-\mathbf{1}_m \mathbf{v}_n^T] \quad (91)$$

$$= \mathbf{D}_+^{-1} \cdot \mathcal{S}_U[\mathbf{A}] \cdot \mathbf{E}_+^{-1} \quad (92)$$

where the last step recognizes that $-\mathbf{u}_m = \text{LogAbs}[\text{Diag}[\mathbf{D}_+^{-1}]]$ and $-\mathbf{v}_n = \text{LogAbs}[\text{Diag}[\mathbf{E}_+^{-1}]]$. The identity of Eq.(67) can then be shown as:

$$\mathcal{D}_U^L[\mathbf{D}_+ \mathbf{A} \mathbf{E}_+] \cdot (\mathbf{D}_+ \mathbf{A} \mathbf{E}_+) \cdot \mathcal{D}_U^R[\mathbf{D}_+ \mathbf{A} \mathbf{E}_+] = \mathcal{S}_U[\mathbf{D}_+ \mathbf{A} \mathbf{E}_+] \circ (\mathbf{D}_+ \mathbf{A} \mathbf{E}_+) \quad (93)$$

$$= (\mathbf{D}_+^{-1} \cdot \mathcal{S}_U[\mathbf{A}] \cdot \mathbf{E}_+^{-1}) \circ (\mathbf{D}_+ \mathbf{A} \mathbf{E}_+) \quad (94)$$

$$= \mathcal{S}_U[\mathbf{A}] \circ \mathbf{A} \quad (95)$$

$$= \mathcal{D}_U^L[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_U^R[\mathbf{A}] \quad (96)$$

Eq.(68) holds as the indexing of the rows and columns of $\mathcal{D}_U^L[\mathbf{A}]$ and $\mathcal{D}_U^R[\mathbf{A}]$ (and $\mathcal{S}_U[\mathbf{A}]$) is the same as that of \mathbf{A} . Eqs.(69) and (70) hold directly because Lemma III.4 only involves functions of the absolute values of the elements of the argument matrix \mathbf{A} . \square

Theorem III.5. For an elemental-nonnzero $m \times n$ matrix \mathbf{A} , the operator

$$\mathbf{A}^U \doteq \mathcal{D}_U^R[\mathbf{A}] \cdot (\mathcal{D}_U^L[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_U^R[\mathbf{A}])^{-P} \cdot \mathcal{D}_U^L[\mathbf{A}] \quad (97)$$

satisfies for any nonsingular diagonal matrices \mathbf{D} and \mathbf{E} :

$$\mathbf{A} \mathbf{A}^U \mathbf{A} = \mathbf{A}, \quad (98)$$

$$\mathbf{A}^U \mathbf{A} \mathbf{A}^U = \mathbf{A}^U, \quad (99)$$

$$(\mathbf{D} \mathbf{A} \mathbf{E})^U = \mathbf{E}^{-1} \mathbf{A}^U \mathbf{D}^{-1}, \quad (100)$$

$$\text{rank}[\mathbf{A}^U] = \text{rank}[\mathbf{A}] \quad (101)$$

and is therefore a general unit-consistent generalized inverse.

Proof. The first two generalized inverse properties can be established from the corresponding properties

of the MP-inverse as:

$$\mathbf{A}\mathbf{A}^{\mathbf{U}}\mathbf{A} = \mathbf{A} \cdot \{\mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}] \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}]\} \cdot \mathbf{A} \quad (102)$$

$$= (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}]^{-1} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}]) \cdot \quad (103)$$

$$\mathbf{A} \cdot \{\mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}] \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}]\} \cdot \mathbf{A} \quad (104)$$

$$\cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}] \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}]^{-1}) \quad (105)$$

$$= \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}]^{-1} \cdot \quad (106)$$

$$\frac{(\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}]) \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}} \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])}{\cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}]^{-1}} \quad (107)$$

$$\cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}]^{-1} \quad (108)$$

$$= \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}]^{-1} \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}]) \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}]^{-1} \quad (109)$$

$$= \mathbf{A} \quad (110)$$

and

$$\mathbf{A}^{\mathbf{U}}\mathbf{A}\mathbf{A}^{\mathbf{U}} = \{\mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}] \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}]\} \cdot \mathbf{A} \quad (111)$$

$$\cdot \{\mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}] \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}]\} \quad (112)$$

$$= \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}] \quad (113)$$

$$\cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}} \cdot \frac{(\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}]) \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}}}{\cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}]} \quad (114)$$

$$\cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \quad (115)$$

$$= \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}] \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \quad (116)$$

$$= \mathbf{A}^{\mathbf{U}} \quad (117)$$

The general UC condition $(\mathbf{DAE})^{\mathbf{U}} = \mathbf{E}^{-1}\mathbf{A}^{\mathbf{U}}\mathbf{D}^{-1}$, for any nonsingular diagonal matrix \mathbf{D} , can be established using a polar decompositions $\mathbf{D} = \mathbf{D}_+\mathbf{D}_u$ and $\mathbf{E} = \mathbf{E}_+\mathbf{E}_u$:

$$(\mathbf{DAE})^{\mathbf{U}} = \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{DAE}] \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{DAE}] \cdot (\mathbf{DAE}) \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{DAE}])^{\mathbf{P}} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{DAE}] \quad (118)$$

$$= \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{D}_+\mathbf{A}\mathbf{E}_+] \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{D}_+\mathbf{A}\mathbf{E}_+] \cdot (\mathbf{DAE}) \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{D}_+\mathbf{A}\mathbf{E}_+])^{\mathbf{P}} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{D}_+\mathbf{A}\mathbf{E}_+] \quad (119)$$

$$= \mathbf{E}_+^{-1} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}] \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{D}_+^{-1} \cdot (\mathbf{DAE}) \cdot \mathbf{E}_+^{-1} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{D}_+^{-1} \quad (120)$$

$$= \mathbf{E}_+^{-1} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}] \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{D}_u \cdot \mathbf{A} \cdot \mathbf{E}_u \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{D}_+^{-1} \quad (121)$$

$$= \mathbf{E}_+^{-1} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}] \cdot \mathbf{E}_u^* \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}} \cdot \mathbf{D}_u^* \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{D}_+^{-1} \quad (122)$$

$$= (\mathbf{E}_+^{-1} \cdot \mathbf{E}_u^*) \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}] \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot (\mathbf{D}_u^* \cdot \mathbf{D}_+^{-1}) \quad (123)$$

$$= \mathbf{E}^{-1} \cdot \{\mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}] \cdot (\mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{R}}[\mathbf{A}])^{\mathbf{P}} \cdot \mathcal{D}_{\mathbf{U}}^{\mathbf{L}}[\mathbf{A}]\} \cdot \mathbf{D}^{-1} \quad (124)$$

$$= \mathbf{E}^{-1} \cdot \mathbf{A}^{\mathbf{U}} \cdot \mathbf{D}^{-1}. \quad (125)$$

The rank-consistency condition of the theorem holds exactly as for the proof of Theorem II.3. \square

The elemental-nonzero condition of Lemma III.4 is required to ensure the existence of the elemental logarithms for $\mathbf{L} = \text{LogAbs}[\mathbf{A}]$, so the closed-form solution for the general unit-consistent matrix inverse of Theorem III.5 is applicable only to matrices without zero elements. In many contexts involving general matrices there is no reason to expect any elements to be identically zero, but in some applications, e.g., compressive sensing, zeros are structurally enforced. Unfortunately, Lemma III.4 cannot be extended to accommodate zeros by a simple limiting strategy; however, results from matrix scaling theory can be applied to derive an unrestricted solution.

IV. THE FULLY GENERAL UNIT-CONSISTENT GENERALIZED INVERSE

Given a nonnegative matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ with full support, m positive numbers $S_1 \dots S_m$, and n positive numbers $T_1 \dots T_n$, Rothblum & Zenios [17] investigated the problem of identifying positive diagonal matrices $\mathbf{U} \in \mathbb{R}^{m \times m}$ and $\mathbf{V} \in \mathbb{R}^{n \times n}$ such that the product of the nonzero elements of each row i of $\mathbf{A}' = \mathbf{U}\mathbf{A}\mathbf{V}$ is S_i and the product of the nonzero elements of each column j of \mathbf{A}' is T_j . They provided an efficient solution, referred to in their paper as *Program II*, and analyzed its properties. Specifically, for vectors $\mu \in \mathbb{R}^m$ and $\eta \in \mathbb{R}^n$, defined in their paper⁵, they proved the following:

Theorem IV.1. (Rothblum & Zenios⁶) The following are equivalent:

- 1) Program II is feasible.
- 2) Program II has an optimal solution.
- 3) $\prod_{i=1}^m (S_i)^{\mu_i} = \prod_{j=1}^n (T_j)^{\eta_j}$.

If a solution exists then the matrix $\mathbf{A}' = \mathbf{U}\mathbf{A}\mathbf{V}$ is the unique positive diagonal scaling of \mathbf{A} for which the product of the nonzero elements of each row i is S_i and the product of the nonzero elements of each column j is T_j .

Although \mathbf{A}' is unique in Theorem IV.1, the diagonal scaling matrices \mathbf{U} and \mathbf{V} may not be. The implications of this, and the question of existence, are addressed by the following theorem.

Theorem IV.2. For any nonnegative matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ with full support there exist positive diagonal matrices $\mathbf{U} \in \mathbb{R}^{m \times m}$ and $\mathbf{V} \in \mathbb{R}^{n \times n}$ such that the product of the nonzero elements of each row i and column j of $\mathbf{X} = \mathbf{U}\mathbf{A}\mathbf{V}$ is 1, and $\mathbf{X} = \mathbf{U}\mathbf{A}\mathbf{V}$ is the unique positive diagonal scaling of \mathbf{A} which has this property. Furthermore, if there do exist distinct positive diagonal matrices \mathbf{U}_1 , \mathbf{V}_1 , \mathbf{U}_2 , and \mathbf{V}_2 such that

$$\mathbf{X} = \mathbf{U}_1\mathbf{A}\mathbf{V}_1 = \mathbf{U}_2\mathbf{A}\mathbf{V}_2 \quad (126)$$

then $\mathbf{V}_1^{-1}\mathbf{A}^p\mathbf{U}_1^{-1} = \mathbf{V}_2^{-1}\mathbf{A}^p\mathbf{U}_2^{-1}$.

Proof. The existence (and uniqueness) of a solution for Program II according to Theorem IV.1 is equivalent to

$$\prod_{i=1}^m (S_i)^{\mu_i} = \prod_{j=1}^n (T_j)^{\eta_j} \quad (127)$$

which holds unconditionally, i.e., independent of μ and η , for the case in which every S_i and T_j is 1. Proof of the *Furthermore* statement is given in Appendix A. \square

Lemma IV.3. Given an $m \times n$ matrix \mathbf{A} , let \mathbf{X} be the matrix formed by removing every row and column of $\text{Abs}[\mathbf{A}]$ for which all elements are equal to zero, and define $r[i]$ to be the row of \mathbf{X} corresponding to row i of \mathbf{A} and $c[j]$ to be the column of \mathbf{X} corresponding to column j of \mathbf{A} . Let \mathbf{U} and \mathbf{V} be the diagonal matrices guaranteed to exist from the application of Program II to \mathbf{X} according to Theorem IV.2. Existence of a general-diagonal scale function according to Definition III.2 for \mathbf{A} is provided by instantiating

⁵As will be seen, the precise definitions of μ and η will prove irrelevant for purposes of this paper.

⁶This theorem combines results from theorems 4.2 and 4.3 of [17]. The variables \mathbf{U} and \mathbf{V} are used here for positive diagonal matrices purely for consistency with that paper despite their exclusive use elsewhere in this paper to refer to unitary matrices. Although not stated explicitly by Rothblum and Zenios, Program II is easily verified to be permutation consistent.

$\mathcal{D}_U^L[\mathbf{A}] = \mathbf{D}$ and $\mathcal{D}_U^R[\mathbf{A}] = \mathbf{E}$ where

$$\mathbf{D}(i, i) = \begin{cases} \mathbf{U}(r[i], r[i]) & \text{row } i \text{ of } \mathbf{A} \text{ is not zero} \\ 1 & \text{otherwise} \end{cases} \quad (128)$$

$$\mathbf{E}(j, j) = \begin{cases} \mathbf{V}(c[j], c[j]) & \text{column } j \text{ of } \mathbf{A} \text{ is not zero} \\ 1 & \text{otherwise} \end{cases} \quad (129)$$

Proof. In the case that \mathbf{A} has full support so that $\mathbf{X} = \text{Abs}[\mathbf{A}]$ then Theorem IV.1 guarantees that $\mathcal{D}_U^L[\mathbf{X}] \cdot \mathbf{X} \cdot \mathcal{D}_U^R[\mathbf{X}]$ is the unique diagonal scaling of \mathbf{X} such that the product of the nonzero elements of each row and column is 1. Therefore, the scale-invariance condition of Eq.(67):

$$\mathcal{D}_U^L[\mathbf{D}_+ \mathbf{X} \mathbf{E}_+] \cdot (\mathbf{D}_+ \mathbf{X} \mathbf{E}_+) \cdot \mathcal{D}_U^R[\mathbf{D}_+ \mathbf{X} \mathbf{E}_+] = \mathcal{D}_U^L[\mathbf{X}] \cdot \mathbf{X} \cdot \mathcal{D}_U^R[\mathbf{X}] \quad (130)$$

holds for any positive diagonals \mathbf{D}_+ and \mathbf{E}_+ as required. For the case of general \mathbf{A} the construction defined by Lemma IV.3 preserves uniqueness with respect to nonzero rows and columns of $\mathcal{D}_U^L[\mathbf{A}] \cdot \text{Abs}[\mathbf{A}] \cdot \mathcal{D}_U^R[\mathbf{A}]$, i.e., those which correspond to the rows and columns of $\mathbf{U}\mathbf{X}\mathbf{V}$, by the guarantee of Theorem IV.1, and any row or column with all elements equal to zero is inherently scale-invariant, so Eq.(67) holds unconditionally for the construction defined by Lemma IV.3. The remaining conditions (permutation consistency and invariance with respect to unitary diagonals) hold equivalently to the proof of Lemma III.4. \square

At this point it is possible to establish the existence of a fully-general, unit-consistent, generalized matrix inverse.

Theorem IV.4. For an $m \times n$ matrix \mathbf{A} there exists an operator

$$\mathbf{A}^U \doteq \mathcal{D}_U^R[\mathbf{A}] \cdot (\mathcal{D}_U^L[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_U^R[\mathbf{A}])^{-P} \cdot \mathcal{D}_U^L[\mathbf{A}] \quad (131)$$

which satisfies for any nonsingular diagonal matrices \mathbf{D} and \mathbf{E} :

$$\mathbf{A}\mathbf{A}^U\mathbf{A} = \mathbf{A}, \quad (132)$$

$$\mathbf{A}^U\mathbf{A}\mathbf{A}^U = \mathbf{A}^U, \quad (133)$$

$$(\mathbf{D}\mathbf{A}\mathbf{E})^U = \mathbf{E}^{-1}\mathbf{A}^U\mathbf{D}^{-1}, \quad (134)$$

$$\text{rank}[\mathbf{A}^U] = \text{rank}[\mathbf{A}]. \quad (135)$$

Proof. The proof of Theorem III.5 applies unchanged to Theorem IV.4 except that the elemental-nonzero condition imposed by Lemma III.4 is removed by use of Lemma IV.3. \square

For completeness, the example of Eqs.(13-16) with

$$\mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & -1/2 \end{bmatrix} \quad (136)$$

can be revisited to verify that

$$(\mathbf{D}\mathbf{A}\mathbf{D}^{-1})^U = \mathbf{D}\mathbf{A}^U\mathbf{D}^{-1} = \begin{bmatrix} 1/2 & 1/4 \\ -1 & -1/2 \end{bmatrix} \quad (137)$$

where $(\mathbf{D}\mathbf{A}\mathbf{D}^{-1})^U = \mathbf{D}\mathbf{A}^U\mathbf{D}^{-1}$ as expected. Extending the example with

$$\mathbf{E} = \begin{bmatrix} 5 & 0 \\ 0 & -3 \end{bmatrix} \quad (138)$$

it can be verified that

$$(\mathbf{DAE})^{\text{U}} = \mathbf{E}^{-1} \mathbf{A}^{\text{U}} \mathbf{D}^{-1} = \begin{bmatrix} 1/10 & 1/20 \\ 1/6 & 1/12 \end{bmatrix} \quad (139)$$

with equality as expected.

In practice the state space of interest may comprise subsets of variables having different assumed relationships. For example, assume that m state variables have incommensurate units while the remaining n state variables are defined in a common Euclidean space, i.e., their relationship should be preserved under orthonormal transformations. This assumption requires that a linear transformation \mathbf{A} should be consistent with respect to state-space transformations of the form

$$\mathcal{T} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \quad (140)$$

where \mathbf{D} is a nonsingular $m \times m$ diagonal matrix and \mathbf{R} is an $n \times n$ orthonormal matrix. Thus the inverse of \mathbf{A} cannot be obtained by applying either the UC inverse or the Moore-Penrose inverse, and the two inverses cannot be applied separately to distinct subsets of the state variables because all of the variables mix under the transformation. This can be seen from a block-partition:

$$\mathbf{A} = \begin{bmatrix} \mathbf{W} & \mathbf{X} \\ \mathbf{Y} & \mathbf{Z} \end{bmatrix} \begin{matrix} \} m \\ \} n \end{matrix} \quad (141)$$

$\underbrace{\hspace{1.5cm}}_m \quad \underbrace{\hspace{1.5cm}}_n$

and noting that consistency in this case requires a generalized inverse that satisfies:

$$(\mathcal{T}_1 \cdot \mathbf{A} \cdot \mathcal{T}_2)^{\tilde{-1}} = \begin{bmatrix} \mathbf{D}_1 \mathbf{W} \mathbf{D}_2 & \mathbf{D}_1 \mathbf{X} \mathbf{R}_2 \\ \mathbf{R}_1 \mathbf{Y} \mathbf{D}_2 & \mathbf{R}_1 \mathbf{Z} \mathbf{R}_2 \end{bmatrix}^{\tilde{-1}} = \mathcal{T}_2^{-1} \cdot \mathbf{A}^{\tilde{-1}} \cdot \mathcal{T}_1^{-1}. \quad (142)$$

In the case of nonsingular \mathbf{A} the partitioned inverse is unique:

$$\mathbf{A}^{-1} = \begin{bmatrix} (\mathbf{W} - \mathbf{XZ}^{-1}\mathbf{Y})^{-1} & -\mathbf{W}^{-1}\mathbf{X}(\mathbf{Z} - \mathbf{YW}^{-1}\mathbf{X})^{-1} \\ -\mathbf{Z}^{-1}\mathbf{Y}(\mathbf{W} - \mathbf{XZ}^{-1}\mathbf{Y})^{-1} & (\mathbf{Z} - \mathbf{YW}^{-1}\mathbf{X})^{-1} \end{bmatrix} \quad (143)$$

and is unconditionally consistent. Respecting the block constraints implicit from Eq.(142), the desired generalized inverse for singular \mathbf{A} under the present assumptions can be verified as:

$$\mathbf{A}^{\tilde{-1}} = \begin{bmatrix} (\mathbf{W} - \mathbf{XZ}^{\text{P}}\mathbf{Y})^{\text{U}} & -\mathbf{W}^{\text{U}}\mathbf{X}(\mathbf{Z} - \mathbf{YW}^{\text{U}}\mathbf{X})^{\text{P}} \\ -\mathbf{Z}^{\text{P}}\mathbf{Y}(\mathbf{W} - \mathbf{XZ}^{\text{P}}\mathbf{Y})^{\text{U}} & (\mathbf{Z} - \mathbf{YW}^{\text{U}}\mathbf{X})^{\text{P}} \end{bmatrix}. \quad (144)$$

The general case involving different assumptions for more than two subsets of state variables (possibly different for the left and right spaces of the transformation) can be solved analogously with appropriate partitioning.

The generalized inverse of Theorem IV.4 is unique when instantiated using the construction defined by Lemma IV.3 by virtue of the uniqueness of both the Moore-Penrose inverse and the scaling of Theorem IV.1 (alternative scalings are discussed in Appendix B), and it completes a trilogy of generalized matrix inverses that exhausts the standard family of transformation invariants. Specifically, the Drazin inverse is consistent with respect to similarity transformations, the Moore-Penrose inverse is consistent with respect to unitary/orthonormal transformations, and the new generalized inverse is consistent with respect to diagonal transformations.

In the next section it is demonstrated that the general approach for obtaining the UC inverse can be efficiently applied to a wide variety of other matrix decompositions and operators (including other generalized matrix inverses) to impose unit consistency.

V. UNIT-CONSISTENT/INVARIANT MATRIX DECOMPOSITIONS

Unit consistency has been suggested in the past as a critical consideration in specific applications (e.g., robotics [9], [7] and data fusion [22]), but the means for enforcing it have been limited because the most commonly applied tools in linear systems analysis, the eigen and singular-value decompositions, are inherently not unit consistent and therefore require UC alternatives. This motivates the need to extend unit-consistency to other areas of matrix analysis. This clearly includes transformations $T[\mathbf{A}]$ which can be redefined in UC form as

$$\mathcal{D}_U^L[\mathbf{A}]^{-1} \cdot T[\mathcal{D}_U^L[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_U^R[\mathbf{A}]] \cdot \mathcal{D}_U^R[\mathbf{A}]^{-1} \quad (145)$$

and functions $f[\mathbf{A}]$ which can be redefined in unit scale-invariant form as $f[\mathcal{D}_U^L[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_U^R[\mathbf{A}]]$, but it also extends to matrix decompositions.

The Singular Value Decomposition (SVD) is among the most powerful and versatile tools in linear algebra and data analytics [23], [16], [13], [1]. The Moore-Penrose generalized inverse of \mathbf{A} can be obtained from the SVD of \mathbf{A}

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^* \quad (146)$$

as

$$\mathbf{A}^p = \mathbf{V}\mathbf{S}^{\sim}\mathbf{U}^* \quad (147)$$

where \mathbf{U} and \mathbf{V} are unitary, \mathbf{S} is the diagonal matrix of singular values of \mathbf{A} , and \mathbf{S}^{\sim} is the matrix obtained from inverting the nonzero elements of \mathbf{S} . This motivates the following definition.

Definition V.1. The Unit-Invariant Singular-Value Decomposition (UI-SVD) is defined as

$$\mathbf{A} = \mathbf{D} \cdot \mathbf{U}\mathbf{S}\mathbf{V}^* \cdot \mathbf{E} \quad (148)$$

with $\mathbf{D} = \mathcal{D}_U^L[\mathbf{A}]^{-1}$, $\mathbf{E} = \mathcal{D}_U^R[\mathbf{A}]^{-1}$, and $\mathbf{U}\mathbf{S}\mathbf{V}^*$ is the SVD of $\mathbf{X} = \mathcal{D}_U^L[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_U^R[\mathbf{A}]$. The diagonal elements of \mathbf{S} are referred to as the *unit-invariant* (UI) *singular values* of \mathbf{A} .

Given the UI-SVD of a matrix \mathbf{A}

$$\mathbf{A} = \mathbf{D} \cdot \mathbf{U}\mathbf{S}\mathbf{V}^* \cdot \mathbf{E} \quad (149)$$

the UC generalized inverse of \mathbf{A} can be expressed as

$$\mathbf{A}^u = \mathbf{E}^{-1} \cdot \mathbf{V}\mathbf{S}^{\sim}\mathbf{U}^* \cdot \mathbf{D}^{-1}. \quad (150)$$

Unlike the singular values of \mathbf{A} , which are invariant with respect to arbitrary left and right unitary transformations of \mathbf{A} , the UI singular values are invariant with respect to arbitrary left and right nonsingular diagonal transformations⁷. Thus, functions of the unit-invariant singular values are unit-invariant with respect to \mathbf{A} .

The largest k singular values of a matrix (e.g., representing a photograph, video sequence, or other object of interest) can be used to define a unitary-invariant signature [6], [15], [10], [14], [12], [11] which supports computationally efficient similarity testing. However, many sources of error in practical applications are not unitary. As a concrete example, consider a system in which a passport or driving license is scanned to produce a rectilinearly-aligned and scaled image that is to be used as a key to search an existing image database. The signature formed from the largest k unit-invariant singular values can

⁷A *left* UI-SVD can be similarly defined as $\mathbf{A} = \mathbf{D} \cdot \mathbf{U}\mathbf{S}\mathbf{V}^*$ with $\mathbf{D} = \mathcal{D}_L[\mathbf{A}]^{-1}$ and $\mathbf{U}\mathbf{S}\mathbf{V}^*$ being the SVD of $\mathbf{X} = \mathcal{D}_L[\mathbf{A}] \cdot \mathbf{A}$. The resulting *left unit-invariant singular values* are invariant with respect to nonsingular left diagonal transformations and right unitary transformations of \mathbf{A} (the latter property is what motivated the right unitary-invariance requirement of Definition II.1 and the consequent use of a unitary-invariant norm in the construction of Lemma II.2.). A *right* UI-SVD can be defined analogously to interchange the left and right invariants.

be used for this purpose to provide robustness to amplitude variations among the rows and/or columns of the image due to the scanning process.

The UI-SVD may also offer advantages as an alternative to the conventional SVD, or truncated SVD, used by existing methods for image and signal processing, cryptography, digital watermarking, tomography, and other applications in order to provide state-space or coordinate-aligned robustness to noise⁸.

More generally, the approach used to define unit scale-invariant singular values can be applied to other matrix decompositions, though the invariance properties may be different. In the case of scale-invariant eigenvalues for square \mathbf{A} , i.e., $\text{eig}[\mathcal{D}_U^L[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_U^R[\mathbf{A}]]$, the invariance is limited to diagonal transformations \mathbf{DAE} such that \mathbf{DE} is nonnegative real, e.g., $\mathbf{D}_+ \mathbf{A} \mathbf{E}_+$, \mathbf{DAD} (or \mathbf{DAD} for complex \mathbf{D}), and \mathbf{DAD}^{-1} . In applications in which changes of units can be assumed to take the form of positive diagonal transformations, the scale-invariant (SI) eigenvalues can therefore be taken as a complementary or alternative signature to that provided by the UI singular values.

VI. DISCUSSION

The principal contribution of this paper is the derivation of a unit-consistent generalized matrix inverse. Its consistency with respect to diagonal transformations provides an alternative to the Drazin inverse, which is only consistent with respect to similarity transformations, and the Moore-Penrose inverse, which is only consistent with respect to unitary/orthonormal transformations. This inverse has broad potential applications as a replacement to Moore-Penrose, e.g., for unit-consistent or unit-invariant gradient descent for optimization and deep learning. Another contribution of the paper is a demonstration that a partitioned inverse can be constructed to respect different consistency assumptions associated with different subsets of variables of a defined state space. This is essential to permit a large, complex linear system to be mathematically expressed and analyzed in a unified manner. More generally, different inverses can also be applied in parallel to obtain sets of distinct solutions, e.g., in order to extract information under different invariance and/or consistency assumptions in pattern recognition and machine learning applications.

It has been emphasized that the appropriate choice of inverse depends critically on the assumed properties of the system of interest. It should also be emphasized that while the Drazin, Moore-Penrose, and UC inverses may be *analytically* special because of their respective algebraic properties, there are many practical situations in which a non-algebraic “heuristic” inverse may represent the more appropriate choice. For example, consider the following singular matrix:

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (151)$$

From a practical perspective, if the matrix is assumed to be nonnegative then its “true” inverse is

$$\mathbf{A}^{-1} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \infty \end{bmatrix} \quad (152)$$

and could be approximated with a large number in place of infinity as

$$\tilde{\mathbf{A}}^{-1} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 10^{15} \end{bmatrix} \quad (153)$$

⁸Such applications are examined in greater detail in a longer preprint from which this paper was derived [21]. In applications in which the signature provided by the set of UI singular values may be too concise [20], e.g., because it is permutation invariant, a vectorization of the matrix $\mathbf{A} \circ (\mathbf{A}^U)^T$ can be used as a more discriminating UI signature.

so as to capture the blow-up of zero when inverted. This blow-up behavior is lost completely if an algebraic inverse is applied:

$$\mathbf{A}^{\text{D}} = \mathbf{A}^{\text{P}} = \mathbf{A}^{\text{U}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (154)$$

This counter-intuitive (and in many practical situations *wrong*) result derives from the fact that no finite value can be chosen in a way that preserves any form of pure *algebraic* consistency⁹. This is just one more example of the danger of blindly using the Moore-Penrose pseudoinverse (or any other generalized inverse) to deal generically with the problem of inverting a singular matrix. In summary, it is essential to make the choice of generalized matrix inverse based on the application-specific properties that are necessary to preserve, and this paper has provided the appropriate choice for unit consistency.

APPENDIX A UNIQUENESS OF THE UC INVERSE

By virtue of the uniqueness of the Moore-Penrose inverse, the UC inverse from Theorem IV.4 is uniquely determined given a scaling $\mathbf{A} = \mathcal{D}_v^L \mathbf{X} \mathcal{D}_v^R$ produced according to Theorem IV.2. However, the positive diagonal matrices $\mathcal{D}_v^L[\mathbf{A}]$ and $\mathcal{D}_v^R[\mathbf{A}]$ are not necessarily unique, so there may exist distinct positive diagonal matrices \mathbf{D}_1 and \mathbf{D}_2 and \mathbf{E}_1 and \mathbf{E}_2 such that

$$\mathbf{A} = \mathbf{D}_1 \mathbf{X} \mathbf{E}_1 = \mathbf{D}_2 \mathbf{X} \mathbf{E}_2. \quad (155)$$

What remains is to establish the uniqueness of \mathbf{A}^{U} in this case, i.e., that

$$\mathbf{D}_1 \mathbf{X} \mathbf{E}_1 = \mathbf{D}_2 \mathbf{X} \mathbf{E}_2 \implies \mathbf{E}_1^{-1} \mathbf{X}^{\text{P}} \mathbf{D}_1^{-1} = \mathbf{E}_2^{-1} \mathbf{X}^{\text{P}} \mathbf{D}_2^{-1}. \quad (156)$$

We begin by noting that if an arbitrary $m \times n$ matrix \mathbf{A} has rank r then it can be factored [3] as the product of an $m \times r$ matrix \mathbf{F} and an $r \times n$ matrix \mathbf{G} as

$$\mathbf{A} = \mathbf{F} \mathbf{G} \quad (157)$$

The Moore-Penrose inverse can then be expressed in terms of this rank factorization as

$$\mathbf{A}^{\text{P}} = \mathbf{G}^* \cdot (\mathbf{F}^* \cdot \mathbf{A} \cdot \mathbf{G}^*)^{-1} \cdot \mathbf{F}^* \quad (158)$$

where \mathbf{G}^* and \mathbf{F}^* are the conjugate transposes of \mathbf{G} and \mathbf{F} .

Because $\mathbf{D}_1 \mathbf{X} \mathbf{E}_1 = \mathbf{D}_2 \mathbf{X} \mathbf{E}_2$ implies

$$\mathbf{X} = \mathbf{D}_2^{-1} \mathbf{D}_1 \mathbf{X} \mathbf{E}_1 \mathbf{E}_2^{-1} \quad (159)$$

then from the rank factorization $\mathbf{X} = \mathbf{F} \mathbf{G}$ we can obtain an alternative factorization

$$\mathbf{X} = \mathbf{F}' \mathbf{G}' = (\mathbf{D}_2^{-1} \mathbf{D}_1 \mathbf{F}) (\mathbf{G} \mathbf{E}_1 \mathbf{E}_2^{-1}) \quad (160)$$

from the fact that the ranks of \mathbf{F} and \mathbf{G} are unaffected by nonsingular diagonal scalings. Applying the

⁹It might not be inappropriate to formally define the result of Eq.(153) as a *blow-up consistent* generalized matrix inverse.

rank factorization identity for the Moore-Penrose inverse then yields

$$\mathbf{X}^P = (\mathbf{F}'\mathbf{G}')^{-P} \quad (161)$$

$$= (\mathbf{G}\mathbf{E}_1\mathbf{E}_2^{-1})^* \cdot ((\mathbf{D}_2^{-1}\mathbf{D}_1\mathbf{F})^* \cdot \mathbf{X} \cdot (\mathbf{G}\mathbf{E}_1\mathbf{E}_2^{-1})^*)^{-1} \cdot (\mathbf{D}_2^{-1}\mathbf{D}_1\mathbf{F})^* \quad (162)$$

$$= \mathbf{E}_1\mathbf{E}_2^{-1}\mathbf{G}^* \cdot ((\mathbf{F}^*\mathbf{D}_2^{-1}\mathbf{D}_1)\mathbf{X}(\mathbf{E}_1\mathbf{E}_2^{-1}\mathbf{G}^*))^{-1} \cdot \mathbf{F}^*\mathbf{D}_2^{-1}\mathbf{D}_1 \quad (163)$$

$$= (\mathbf{E}_1\mathbf{E}_2^{-1}) \cdot \mathbf{G}^* \cdot \left(\mathbf{F}^* \cdot (\mathbf{D}_2^{-1}\mathbf{D}_1\mathbf{X}\mathbf{E}_1\mathbf{E}_2^{-1}) \cdot \mathbf{G}^* \right)^{-1} \cdot \mathbf{F}^* \cdot (\mathbf{D}_2^{-1}\mathbf{D}_1) \quad (164)$$

$$= (\mathbf{E}_1\mathbf{E}_2^{-1}) \cdot (\mathbf{G}^* \cdot (\mathbf{F}^*\mathbf{X}\mathbf{G}^*)^{-1} \cdot \mathbf{F}^*) \cdot (\mathbf{D}_2^{-1}\mathbf{D}_1) \quad (165)$$

$$= \mathbf{E}_1\mathbf{E}_2^{-1} \left(\mathbf{G}^* \cdot (\mathbf{F}^*\mathbf{X}\mathbf{G}^*)^{-1} \cdot \mathbf{F}^* \right) \mathbf{D}_2^{-1}\mathbf{D}_1 \quad (166)$$

$$= \mathbf{E}_1\mathbf{E}_2^{-1}\mathbf{X}^P\mathbf{D}_2^{-1}\mathbf{D}_1 \quad (167)$$

which implies¹⁰

$$\mathbf{E}_1^{-1}\mathbf{X}^P\mathbf{D}_1^{-1} = \mathbf{E}_1^{-1} \cdot (\mathbf{E}_1\mathbf{E}_2^{-1}\mathbf{X}^P\mathbf{D}_2^{-1}\mathbf{D}_1) \cdot \mathbf{D}_1^{-1} \quad (168)$$

$$= (\mathbf{E}_1^{-1}\mathbf{E}_1) \cdot \mathbf{E}_2^{-1}\mathbf{X}^P\mathbf{D}_2^{-1} \cdot (\mathbf{D}_1\mathbf{D}_1^{-1}) \quad (169)$$

$$= \mathbf{E}_2^{-1}\mathbf{X}^P\mathbf{D}_2^{-1} \quad (170)$$

and thus establishes that $\mathbf{E}_1^{-1}\mathbf{X}^P\mathbf{D}_1^{-1} = \mathbf{E}_2^{-1}\mathbf{X}^P\mathbf{D}_2^{-1}$ and therefore that the UC generalized inverse \mathbf{A}^U is unique.

Using a similar but more involved application of rank factorization it can be shown that the UC generalized matrix inverse satisfies

$$\mathbf{A}^U \cdot (\mathbf{A}^U)^{-U} \cdot \mathbf{A}^U = \mathbf{A}^U \quad (171)$$

which is weaker than the uniquely-special property of the Moore-Penrose inverse:

$$(\mathbf{A}^P)^{-P} = \mathbf{A}. \quad (172)$$

APPENDIX B ALTERNATIVE CONSTRUCTIONS

The proofs of Theorems II.3 and III.5 (and consequently Theorem IV.4) do not actually require the general unitary consistency property of the Moore-Penrose inverse and instead only require diagonal unitary consistency, e.g., in Eqs.(56)-(57) as

$$(\mathbf{D}_u\mathbf{A})^{-P} = \mathbf{A}^P\mathbf{D}_u^* \quad (173)$$

and in Eqs.(121)-(122) as

$$(\mathbf{D}_u\mathbf{A}\mathbf{E}_u)^{-P} = \mathbf{E}_u^* \mathbf{A}^P \mathbf{D}_u^* \quad (174)$$

for unitary diagonal matrices \mathbf{D}_u and \mathbf{E}_u . Thus, the Moore-Penrose inverse could be replaced with an alternative which maintains the other required properties but satisfies this weaker condition in place of general unitary consistency.

Similarly, the scalings defined by Lemmas III.4 and IV.3 are not necessarily the only ones that may be used to satisfy the conditions of Definition III.2. More specifically, Lemmas III.4 and IV.3 define left and right nonnegative diagonal scaling functions $\mathcal{D}_U^L[\mathbf{A}]$ and $\mathcal{D}_U^R[\mathbf{A}]$ satisfying

$$\mathcal{D}_U^L[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_U^R[\mathbf{A}] = \mathcal{D}_U^L[\mathbf{D}_+\mathbf{A}\mathbf{E}_+] \cdot \mathbf{D}_+\mathbf{A}\mathbf{E}_+ \cdot \mathcal{D}_U^R[\mathbf{D}_+\mathbf{A}\mathbf{E}_+] \quad (175)$$

for all positive diagonals \mathbf{D}_+ and \mathbf{E}_+ . Because the unitary factors of the elements of \mathbf{A} are unaffected by the nonnegative scaling, the scalings can be constructed without loss of generality from $\text{Abs}[\mathbf{A}]$.

¹⁰Note that the diagonal matrices commute and are real so, e.g., $\mathbf{D}^* = \mathbf{D}$.

If nonnegative \mathbf{A} is square, irreducible, and has full support then such a scaling can be obtained by alternately normalizing the rows and columns to have unit sum using the Sinkhorn iteration [18], [19]. The requirement for irreducibility stems from the fact that the process cannot always converge to a finite left and right scaling. For example, the matrix

$$\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \quad (176)$$

cannot be scaled so that the rows and columns sum to unity unless the off-diagonal element b is driven to zero, which is not possible for any finite scaling. In other words, the Sinkhorn unit-sum condition cannot be jointly satisfied with respect to both the set of row vectors and the set of column vectors. What is needed, therefore, is a measure of vector “size” that can be applied within a Sinkhorn-type iteration but is guaranteed to converge to a finite scaling¹¹.

Definition B.1. For all vectors \mathbf{u} with elements from a normed division algebra, a nonnegative composable size function $s[\mathbf{u}]$ is defined as satisfying the following conditions for all α :

$$s[\mathbf{u}] = 0 \Leftrightarrow \mathbf{u} = \mathbf{0} \quad (177)$$

$$s[\alpha\mathbf{u}] = |\alpha| \cdot s[\mathbf{u}] \quad (178)$$

$$s[\mathbf{b}] = 1 \quad \forall \mathbf{b} \in \{0, 1\}^n - \mathbf{0}_n \quad (179)$$

$$s[\mathbf{u}] = s[\mathbf{u} \otimes \mathbf{b}] = s[\mathbf{b} \otimes \mathbf{u}] \quad \forall \mathbf{b} \in \{0, 1\}^n \quad (180)$$

The defined size function provides a measure of scale that is homogeneous, permutation-invariant, and invariant with respect to tensor expansions involving identity and zero elements. More intuitively, however, $s[\mathbf{u}]$ can be thought of as a “mean-like” measure taken over the magnitudes of the nonzero elements of \mathbf{u} . With the imposed condition $s[\mathbf{0}] \doteq 0$ the following instantiations can also be verified to satisfy the definition:

$$s_{\times}[\mathbf{u}] \doteq \left(\prod_{k \in S} |\mathbf{u}_k| \right)^{1/|S|} \quad j \in S \text{ iff } \mathbf{u}(j) \neq 0 \quad (181)$$

$$s_p[\mathbf{u}] \doteq \|\mathbf{u}\|_p / |S|^{1/p} \quad j \in S \text{ iff } \mathbf{u}(j) \neq 0 \quad (182)$$

$$s_{a,b}[\mathbf{u}] \doteq \left(\frac{\sum_i |\mathbf{u}_i|^{a+b}}{\sum_i |\mathbf{u}_i|^a} \right)^{1/b} \quad a > 0, b > 0 \quad (183)$$

The first case, $s_{\times}(\mathbf{u})$, is more easily interpreted as the geometric mean of the nonzero elements of \mathbf{u} . Its application in a Sinkhorn-type iteration converges to a unique scaling in which the *product* of the nonzero elements in each row and column has unit magnitude. If a , b , and c are positive for the matrix of Eq.(176) then the scaled result using $s_p[\mathbf{u}]$ is

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad (184)$$

where the product of the nonzero elements in each row and column is unity and the particular left and right diagonal scalings are determined by the values of a , b , and c . It can be shown that for all elemental nonzero matrices that the scaling produced using $s_{\times}(\mathbf{u})$ is equivalent to that produced by the constructions defined by Lemmas III.4 and IV.3 and that the iteration is fast-converging.

The row/column conditions imposed by $s_p[\mathbf{u}]$ can most easily be understood in the case of $p = 1$, for which it is equivalent to the mean of the absolute values of the nonzero elements of \mathbf{u} . In the case of

¹¹This definition and the subsequently-defined instance, $s_{a,b}[\mathbf{u}]$, may be of independent interest for analyzing properties of low-rank subspace embeddings in high-dimensional vector spaces, e.g., infinite-dimensional spaces.

$p = 2$, if a vector \mathbf{v} is formed from the m nonzero elements of \mathbf{u} then

$$s_2[\mathbf{u}] = \|\mathbf{v}\|_2 / m^{1/2} \quad (185)$$

In the example of the 2×2 matrix of Eq.(176) the scaled result produced using $s_p[\mathbf{u}]$ for any $p > 0$ happens to be the same as that produced using $s_\times[\mathbf{u}]$. For nontrivial matrices, however, the results for different p are not generally (nor typically) equivalent to each other or to that produced by $s_\times(\mathbf{u})$.

The third size function, $s_{a,b}[\mathbf{u}]$, satisfies the required conditions without imposing special treatment of zero elements. In other words, it is a continuous function of the elements of \mathbf{u} and would therefore appear to be a more natural choice for instantiating $\mathcal{D}_U^L[\mathbf{A}]$ and $\mathcal{D}_U^R[\mathbf{A}]$ for analysis purposes, e.g., in the limit as a and b go to zero where $s_{a,b}[\mathbf{u}] \equiv s_\times[\mathbf{u}]$. (It should be noted that the homogeneity properties of $s_{a,b}[\mathbf{u}]$ hold generally for any a and b from a normed division algebra with $0^0 \doteq 1$, and it subsumes $s_p[\mathbf{u}]$ in the limiting cases $a \rightarrow 0$ and/or $b \rightarrow 0$.)

APPENDIX C IMPLEMENTATIONS

Below are basic Octave/Matlab implementations of some of the methods developed in the paper. Although not coded for maximum efficiency or numerical robustness, they should be sufficient for experimental corroboration of theoretically-established properties.

The following function computes \mathbf{A}^U for $m \times n$ real or complex matrix \mathbf{A} . It has complexity dominated by the Moore-Penrose inverse calculation, which is $O(mn \cdot \min(m, n))$.

```
function Ai = uinv(A)
    [S, dl, dr] = dscale(A);
    Ai = pinv(S) .* (dl * dr)';
end
```

The following function evaluates the UC/UI singular values of the real or complex matrix \mathbf{A} .

```
function s = usvd(A)
    s = svd(dscale(A));
end
```

The following function evaluates the UC/UI singular-value decomposition of the $m \times n$ real or complex matrix \mathbf{A} .

```
function [D, U, S, V, E] = usv_decomp(A)
    [S, dl, dr] = dscale(A);
    D = diag(1./dl); E = diag(1./dr);
    [U, S, V] = svd(S);
end
```

The following function computes the unique positively-scaled matrix $\mathbf{S} = \mathcal{D}_U^L[\mathbf{A}] \cdot \mathbf{A} \cdot \mathcal{D}_U^R[\mathbf{A}]$ with diagonal left and right scaling matrices $\mathcal{D}_U^L[\mathbf{A}] = \text{diag}[dl]$ and $\mathcal{D}_U^R[\mathbf{A}] = \text{diag}[dr]$. It has $O(mn)$ complexity for $m \times n$ real or complex matrix \mathbf{A} .

```
function [S, dl, dr] = dscale(A)
    tol = 1e-15;
    [m, n] = size(A);
    L = zeros(m, n); M = ones(m, n);
    S = sign(A); A = abs(A);
```

```

idx = find(A > 0.0); L(idx) = log(A(idx));
idx = setdiff(1 : numel(A), idx);
L(idx) = 0; M(idx) = 0;
r = sum(M, 2); c = sum(M, 1);
u = zeros(m, 1); v = zeros(1, n);
dx = 2*tol;
while (dx > tol)
    idx = c > 0;
    p = sum(L(:, idx), 1) ./ c(idx);
    L(:, idx) = L(:, idx) - repmat(p, m, 1) .* M(:, idx);
    v(idx) = v(idx) - p; dx = mean(abs(p));
    idx = r > 0;
    p = sum(L(idx, :), 2) ./ r(idx);
    L(idx, :) = L(idx, :) - repmat(p, 1, n) .* M(idx, :);
    u(idx) = u(idx) - p; dx = dx + mean(abs(p));
end
dl = exp(u); dr = exp(v);
S = S.* exp(L);
end

```

REFERENCES

- [1] O. Alter, P.O. Brown, D. Botstein, "Singular Value Decomposition for Genome-Wide Expression Data Processing and Modeling," *Proc Natl Acad Sci*, 97(18):10101-6, 2000.
- [2] O. Alter O, G.H. Golub, "Integrative Analysis of Genome-Scale Data by Using Pseudoinverse Projection Predicts Novel Correlation Between DNA Replication and RNA Transcription," *Proc Natl Acad Sci*, 101(47):16577-16582, 2004.
- [3] A. Ben-Israel and N.E. Greville, *Generalized Inverses: Theory and Applications*, 2nd Edition, Springer-Verlag, 2003.
- [4] J. G. Berryman, "Analysis of Approximate Inverses in Tomography. I. Resolution analysis," *Optimization and Engineering*, 1, 87-117, 2000.
- [5] S.L. Campbell, C.D. Meyer, and N.J. Rose, "Applications of the Drazin Inverse to Linear Systems of Differential Equations with Singular Constant Coefficients," *SIAM Journal of Applied Mathematics*, Vol. 31, No. 3, 1976.
- [6] B. Cui, Z. Zhao, W.H. Tok, "A Framework for Similarity Search of Time Series Cliques with Natural Relations," *IEEE Transaction on Data and Knowledge Engineering*, 2012.
- [7] K. L. Doty, C. Melchiorri, and C. Bonivento, "A Theory of Generalized Inverses Applied to Robotics," *International Journal of Robotics Research*, vol. 12, no. 1, pp. 1-19, 1995.
- [8] M. Drazin, "Pseudo-Inverses in Associative Rings and Semigroups," *The American Mathematical Monthly*, 65:7, 1958.
- [9] J. Duffy, "The Fallacy of Modern Hybrid Control Theory that is Based on 'Orthogonal Complements' of Twists and Wrenches Spaces", *Int. J. of Robotic Systems*, 7(2), 1990.
- [10] Z-Q Hong, "Algebraic feature extraction of image for recognition," *Pattern Recognition*, 24(3), 211-219, 1991.
- [11] KM Jeong and J-J Lee, "Video Sequence Matching Using Normalized Dominant Singular Values," *Journal of the Korea Multimedia Society*, Vol.12:12, Page 785-793, 2009.
- [12] KM Jeong, J-J Lee, Y-H Ha, "Video sequence matching using singular value decomposition," *Proc. 3rd Int. Conf. Image Analysis and Recognition (ICIAR)*, pp 426-435, 2006.
- [13] F. Leblond, K.M. Tichauer, B.W. Pogue, "Singular Value Decomposition Metrics Show Limitations of Detector Design in Diffuse Fluorescence Tomography," *Biomedical Optics Express.*, 1(5):1514-1531, 2010.
- [14] J. H. Luo and C. C. Chen, "Singular Value Decomposition for Texture Analysis," *Applications of Digital Image Processing XVII*, SPIE Proceedings, vol. 2298, pp.407-418, 1994.
- [15] G.J. Meyer, *Classification of Radar Targets using Invariant Features*, Dissertation, Air Force Institute of Technology, AFIT/DS/ENG/03-04, 2003.
- [16] R.J. Michelena, "Singular Value Decomposition for Cross-Well Tomography," *Geophysics*, 58(11):1655-1661, 1993.
- [17] U.G. Rothblum and S.A. Zenios, "Scalings of Matrices Satisfying Line-Product Constraints and Generalizations," *Linear Algebra and Its Applications*, 175: 159-175, 1992.
- [18] R. Sinkhorn, "A relationship between arbitrary positive matrices and doubly stochastic matrices," *Ann. Math. Statist.*, 35, 876-879, 1964.
- [19] R. Sinkhorn and P. Knopp, "Concerning nonnegative matrices and doubly stochastic matrices," *Pacific J. Math.*, 21, 343-348, 1967.

- [20] Y. Tian, T. Tan, and Y. Wang, "Do singular values contain adequate information for face recognition?," *Pattern Recognition*, 36:649-655, 2003.
- [21] J.K. Uhlmann, "Unit Consistency, Generalized Inverses, and Effective System Design Methods," arXiv:1604.08476v2 [cs.NA] 11 Jul 2017.
- [22] J.K. Uhlmann, *Dynamic Map Building and Localization: New Theoretical Foundations*, pp. 86-87, Doctoral Dissertation, University of Oxford, 1995.
- [23] H. Yanai, K. Takeuchi, Y. Takane, *Projection Matrices, Generalized Inverse Matrices, and Singular Value Decomposition*," Springer, ISBN-10:1441998861, 2011.