Castro-González, N., Dopico, F. M. & Molera, J. M. (2016). Multiplicative perturbation theory of the Moore–Penrose inverse and the least squares problem. *Linear Algebra and Its Applications*, 503, pp. 1–25.

DOI: 10.1016/j.laa.2016.03.027

MULTIPLICATIVE PERTURBATION THEORY OF THE MOORE-PENROSE INVERSE AND THE LEAST SQUARES PROBLEM*

NIEVES CASTRO-GONZÁLEZ†, FROILÁN M. DOPICO‡, AND JUAN M. MOLERA‡

Abstract. Bounds for the variation of the Moore-Penrose inverse of general matrices under multiplicative perturbations are presented. Their advantages with respect to classical bounds under additive perturbations and with respect to other bounds under multiplicative perturbations available in the literature are carefully studied and established. Closely connected to these developments a complete multiplicative perturbation theory for least squares problems, valid for perturbations of any size, is also presented, improving in this way recent multiplicative perturbation bounds which are valid only to first order in the size of the perturbations. The results in this paper are mainly based on exact expressions of the perturbed Moore-Penrose inverse in terms of the unperturbed one, the perturbation matrices, and certain orthogonal projectors. Such feature makes the new results amenable to be generalized in the future to linear operators in infinite dimensional spaces.

Key words. least squares problems, Moore-Penrose inverse, multiplicative perturbation theory

AMS subject classifications. 15A09, 15A12, 65F35

1. Introduction. The main character of this paper is the Moore-Penrose inverse. Therefore, we start by stating its definition and most basic properties. The Moore-Penrose inverse of $A \in \mathbb{C}^{m\times n}$ is defined to be the unique matrix $Z \in \mathbb{C}^{n\times m}$ such that

$$\text{(i) } AZA = A, \quad \text{(ii) } ZAZ = Z, \quad \text{(iii) } (AZ)^* = AZ, \quad \text{(iv) } (ZA)^* = ZA, \tag{1.1}$$

or, equivalently, such that

$$AZ = P_A \quad \text{and} \quad ZA = P_Z, \tag{1.2}$$

where P_A and P_Z stand for the orthogonal projectors onto the column spaces of A and Z, respectively. The equivalence of the four conditions in (1.1) and the two conditions in (1.2) can be easily established and can be found in [4, Theorem 1.1.1]. We will denote by $A^\dagger \in \mathbb{C}^{n\times m}$ the Moore-Penrose inverse of $A \in \mathbb{C}^{m\times n}$. Recall that if $A \in \mathbb{C}^{n\times n}$ is nonsingular, then $A^\dagger = A^{-1}$. It is well known [21, Chapter 3] that the SVD of A allows us to get an expression for $A^\dagger$ and to prove many of its properties. $\mathcal{R}(A)$ will denote the column space of A and $\mathcal{N}(A)$ its null space. It is easy to see that $\mathcal{R}(A^*) = \mathcal{R}(A^\dagger)$, so, according to (1.2), $P_A = AA^\dagger$ and $P_{A^*} = P_{A^\dagger} = A^\dagger A$ are, respectively, the orthogonal projectors onto $\mathcal{R}(A)$ and $\mathcal{R}(A^*)$. One of the most important properties of the Moore-Penrose inverse is that the minimum 2-norm solution of the Least Squares Problem (LSP)

$$\min_{x\in\mathbb{C}^n} \|Ax - b\|_2, \quad A \in \mathbb{C}^{m\times n}, \quad b \in \mathbb{C}^m, \tag{1.3}$$

is $x_0 = A^\dagger b$.

*This work was partially supported by the Ministerio de Economía y Competitividad of Spain through grants MTM-2012-32542, MTM-2015-68805-REDT, MTM-2015-65798-P (Dopico and Molera) and MTM-2010-18057 (Castro-González).

†Facultad de Informática, Universidad Politécnica de Madrid, 28660 Boadilla del Monte, Madrid, Spain (nieves@fi.upm.es).

‡Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. de la Universidad 30, 28911 Leganés, Spain (dopico@math.uc3m.es, molera@math.uc3m.es).

The Moore-Penrose inverse is probably the most important of all other generalized inverses due to its relationship with the LSP, both for full rank and rank deficient matrices. The perturbation theory of the Moore-Penrose inverse is a classical topic in Matrix Analysis and Numerical Linear Algebra. It was studied for the first time by Wedin [22] and nowadays appears in standard references [2, 21].

Classic references only consider additive perturbations of the matrix. Lately there has been an increase of interest in the multiplicative perturbation theory of the Moore-Penrose inverse due, in part, to its application to the error analysis of algorithms that solve structured LSP with High Relative Accuracy (HRA) [3, 5]. In this work we will understand by a multiplicative perturbation of a matrix $A \in \mathbb{C}^{m\times n}$ a matrix of the form $\widetilde{A} = (I+E)A(I+F) \in \mathbb{C}^{m\times n}$, where $(I+E) \in \mathbb{C}^{m\times m}$ and $(I+F) \in \mathbb{C}^{n\times n}$ are nonsingular matrices. Multiplicative perturbation theory [17, 18, 19] has played an important role in the analyses of HRA algorithms for a wide range of Numerical Linear Algebra problems (see [7, 8, 9, 10, 11, 12, 13] and the references therein). In the previous work [5] expressions for $\widetilde{A}^\dagger$ were given and they were used to bound, up to first order, the relative variation for the minimum 2-norm solution of the LSP under multiplicative perturbations, but no bounds for $\|\widetilde{A}^\dagger - A^\dagger\|$ were presented. In this paper we extend those results to give relative perturbation bounds for $\|\widetilde{A}^\dagger - A^\dagger\|$ in terms of the perturbation matrices E and F, and this is done for any normalized unitarily invariant (UI) matrix norm [21, Ch. II, Section 3]. In addition, sharper bounds are presented for the family of UI norms known as Q-norms [1, Def. IV.2.9, p. 95], which includes the spectral or 2-norm, $\|\cdot\|_2$, and the Frobenius norm of matrices, among many others.

Multiplicative perturbation bounds for the Moore-Penrose inverse have also been presented recently in [3]. There are two main differences between the bounds in [3] and the ones presented here: the first is that our bounds are obtained from an explicit expression of $\widetilde{A}^\dagger$ in terms of $A^\dagger$ and the perturbation matrices, which opens the possibility to extend the results to linear operators in infinite dimensional spaces, and the second, and more important, is that we are able to express our relative bounds with respect to either $\|A^\dagger\|_2$ or $\|\widetilde{A}^\dagger\|_2$ (see Theorem 3.5), and not relative to $\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}$. The factor $\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}$ is natural and necessary [21, 22] in the case of additive perturbations if no constraints are imposed on $\widetilde{A} = A + E$. For example it is proved in [21, Ch. III-Theorem 3.8] (see also [22]) that, in any normalized UI norm,

$$\|\widetilde{A}^\dagger - A^\dagger\| \leq \mu \max\{\|A^\dagger\|_2^2, \|\widetilde{A}^\dagger\|_2^2\}\, \|E\|, \tag{1.4}$$

where μ is a moderate constant that depends on the norm used. As G. W. Stewart and J-G. Sun say in [21, p. 145] bounds as (1.4) " ... cannot by themselves insure the convergence of $\widetilde{A}^\dagger$ to $A^\dagger$ as $E \to 0$, since $\widetilde{A}^\dagger$ may grow unboundedly". The reason is that a general additive perturbation can change the rank of A, and it is well known that if the perturbation is not acute then we have the *lower bound* for the size of the perturbation [21, 22]:

$$\|\widetilde{A}^\dagger - A^\dagger\| \geq \frac{1}{\|E\|_2}. \tag{1.5}$$

If, however, the perturbation does not change the rank of the matrix (and that guarantees that the pertubation is acute when its goes to zero [21, p. 145]) we have [21, Ch. III-Theorem 3.9]:

$$\|\widetilde{A}^\dagger - A^\dagger\| \leq \mu \|A^\dagger\|_2 \|\widetilde{A}^\dagger\|_2 \|E\|, \tag{1.6}$$

the factor $\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}$ can be avoided, and $\widetilde{A}^\dagger$ depends continuosly on E. It is reasonable to expect that things will be different for nonsingular multiplicative perturbations because in this case the rank does not change and it should be expected, even in the non acute case, to avoid the factor $\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}$ in the relative bounds for $\|\widetilde{A}^\dagger - A^\dagger\|$. This is what we show in Theorem 3.5, but is missing in [3].

Closely related to the perturbation theory of $A^\dagger$ is the perturbation of the solution of the LSP (1.3). Let $\widetilde{x}_0 = \widetilde{A}^\dagger \widetilde{b}$ be the minimum 2-norm solution of the perturbed LSP

$$\min_{x\in\mathbb{C}^n} \|\widetilde{A}x - \widetilde{b}\|_2, \quad \widetilde{A} \in \mathbb{C}^{m\times n}, \;\; \widetilde{b} \in \mathbb{C}^m. \tag{1.7}$$

The classical bounds [22, Theorem 5.1] for the relative variation of the minimum 2-norm solution and the residual of the LSP under general additive perturbations

$$\widetilde{A} = A + \Delta A, \quad \widetilde{b} = b + \Delta b, \tag{1.8}$$

can be written in the following form [2, Theorem 1.4.6]:

$$\frac{\|\widetilde{x}_0 - x_0\|_2}{\|x_0\|_2} \leq \frac{1}{1-\eta}\left(2\,\kappa_2(A)\,\epsilon_A + \frac{\|A^\dagger\|_2\,\|b\|_2}{\|x_0\|_2}\,\epsilon_b + \kappa_2(A)^2\,\frac{\|r\|_2}{\|A\|_2\,\|x_0\|_2}\,\epsilon_A\right), \tag{1.9}$$

$$\frac{\|\widetilde{r} - r\|_2}{\|b\|_2} \leq \left(\frac{\|A\|_2\|x_0\|_2}{\|b\|_2}\,\epsilon_A + \epsilon_b + \kappa_2(A)\,\frac{\|r\|_2}{\|b\|_2}\,\epsilon_A\right), \tag{1.10}$$

where $r := b - Ax_0$, $\kappa_2(A) = \|A\|_2\,\|A^\dagger\|_2$, $\epsilon_A := \|\Delta A\|_2/\|A\|_2$ and $\epsilon_b := \|\Delta b\|_2/\|b\|_2$, and it is supposed that $x_0 \neq 0$, $\mathrm{rank}\,(A) = \mathrm{rank}\,(\widetilde{A})$, and $\eta := \kappa_2(A)\,\|\Delta A\|_2/\|A\|_2 < 1$. The bounds in (1.9)-(1.10) can be very large if A is ill-conditioned, i.e., $\kappa_2(A) \gg 1$.

A very important application of perturbation theory is the estimation of the forward errors in the computed quantities using backward stable numerical algorithms. In that case the structure of the perturbation is determined by the backward error analysis of the algorithm [14], and the relevant size of the relative perturbations is typically of the order of the unit roundoff of the computer, $\mathtt{u}$. If a numerical algorithm used to solve LSP has *additive* backward errors, as those in (1.8), the relative error in the solution x_0 will have the form (1.9) with $\epsilon_A, \epsilon_b \approx O(\mathtt{u})$. Then the error will be larger than $\mathtt{u}\,\kappa_2(A)$ (in fact, it can be much larger under certain conditions) and, so, (1.9) does not guarantee any digit of accuracy in the computed solution if $\kappa_2(A) \gtrsim 1/\mathtt{u}$, that is, if A is ill-conditioned with respect to the inverse of the unit roundoff. Unfortunately, many types of structured matrices arising in applications (as, for example, Vandermonde matrices, which arise in polynomial data fitting, and Cauchy matrices [14, Chapters 22 and 28]) are extremely ill-conditioned and, so, standard backward stable algorithms for LSP may compute solutions with huge relative errors. However not all algorithms will produce *additive* backward errors. An algorithm to compute the solution of certain structured LSP that has *multiplicative* backward errors has been presented recently [5]. Therefore the study of multiplicative perturbations of the LSP is relevant. We consider in this paper a multiplicatively perturbed LSP as:

$$\min_{x\in\mathbb{C}^n} \|\widetilde{A}x - \widetilde{b}\|_2, \quad \widetilde{A} = (I+E)A(I+F) \in \mathbb{C}^{m\times n}, \;\; \widetilde{b} = b + h \in \mathbb{C}^m, \tag{1.11}$$

where $(I+E) \in \mathbb{C}^{m\times m}$ and $(I+F) \in \mathbb{C}^{n\times n}$ are nonsingular matrices. We show in

Theorem 4.1 that

$$\frac{\|\widetilde{x}_0 - x_0\|_2}{\|x_0\|_2} \leq C_1 + C_2 \frac{\|A^\dagger\|_2 \|b\|_2}{\|x_0\|_2}, \tag{1.12}$$

$$\frac{\|\widetilde{r} - r\|_2}{\|b\|_2} \leq C_3, \tag{1.13}$$

where $\widetilde{x}_0 := \widetilde{A}^\dagger \widetilde{b}$ is the minimum 2-norm solution of (1.11), and C_1, C_2, C_3 are quantities (defined in Theorem 4.1) that depend only on the norms of the perturbations E and F, and on $\|h\|_2/\|b\|_2$. First order asymptotic bounds of the type (1.12) and (1.13) valid only for infinitesimal perturbations E and F have been already proved in [5], while the bounds in this paper are valid for perturbations of any size.

We emphasize that the new multiplicative bounds in (1.12-1.13) are in general completely different from the classical additive bounds in (1.9-1.10). As it has been said, the bound in (1.9) amplifies the perturbations in the data at least by a factor $\kappa_2(A)$ and the amplification can be much larger under certain conditions. However the bound in (1.12) does not depend on $\kappa_2(A)$ and includes only the amplification factor $\|A^\dagger\|_2 \|b\|_2/\|x_0\|_2$. We will show in Subsection 4.2 that $\|A^\dagger\|_2 \|b\|_2/\|x_0\|_2$ is a moderate number, even in the case $\kappa_2(A)$ is very large, except for very particular choices of b. The result (1.12) is therefore an important step to prove that the algorithms in [5] compute with guaranteed HRA solutions of structured LSP that are so ill conditioned that classical algorithms (with additive backward errors) do not guarantee a single digit of accuracy.

The paper is organized as follows. We introduce in Section 2 the basic facts on unitarily invariant norms and Q-norms that will be used throughout the paper. In Section 3 we present explicit expressions for the variation of the Moore-Penrose inverse under multiplicative perturbations and also get bounds for the relative change of $A^\dagger$ in any unitarily invariant norm and in any Q-norm. We get bounds relative to $\|A^\dagger\|_2$ or to $\|\widetilde{A}^\dagger\|_2$, and not to $\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}$. In Section 4 finite multiplicative perturbation bounds for the minimum 2-norm solution of the LSP are presented. We also study the condition number of LSP under multiplicative perturbations and the behavior of the factor $\|A^\dagger\|_2 \|b\|_2/\|x_0\|_2$. Finally in Section 5 some conclusions and lines of future research are discussed.

Note that a few particular cases of the results of this paper have appeared already in [5]. More precisely, Lemma 2.1 in [5] can be obtained from particularizing Lemma 2.2 to the 2-norm. The first equalities in parts (a) and (b) in Lemma 3.1 are Lemma 3.1 in [5], and also some parts of Theorem 3.4 appear in [5] as Theorem 3.2. Theorem 3.4 is a key result in our developments, both here and in [5], though it is used for different purposes in each paper. The proof of Theorem 3.4 is new and also the equations in (3.9) are new results. Finally, observe that the first order bound (4.9) in Section 4, obtained here as a corollary of our main Theorem 4.1 for LSP, appeared in [5] as Theorem 4.1.

2. Preliminaries and basic results. The symbol I_n stands for the $n \times n$ identity matrix, but we will use simply I if the size is clear from the context. A^* denotes the conjugate-transpose of A. We denote the Euclidean vector norm of $x \in \mathbb{C}^n$ by $\|x\|_2$ and the spectral matrix norm, or 2-norm, of A by $\|A\|_2 := \max_{\|x\|_2=1} \|Ax\|_2$. We recall that a norm $\|\cdot\|$ on $\mathbb{C}^{m\times n}$ is unitarily invariant (UI) if $\|U^*AV\| = \|A\|$ for all $A \in \mathbb{C}^{m\times n}$ and any unitary matrices $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$. An UI norm is normalized if $\|A\| = \|A\|_2$ whenever $\mathrm{rank}(A) = 1$. See [1, Section IV.2] or [21, Ch. II, Section 3] for more information on UI norms.

Given $A \in \mathbb{C}^{m\times n}$, with $\mathrm{rank}(A) = r$, its singular values are denoted as $\sigma_1(A) \geq \cdots \geq \sigma_r(A) > \sigma_{r+1}(A) = \cdots = \sigma_p(A) = 0$, with $p = \min\{m, n\}$. Then $\|A\|_2 = \sigma_1(A)$ and the condition number with respect to the 2-norm of A is given by $\kappa_2(A) = \sigma_1(A)/\sigma_r(A)$.

The Ky Fan p-k norms on $\mathbb{C}^{m\times n}$ [1, p. 95], defined by

$$\|A\|_{p,k} = \left(\sigma_1(A)^p + \cdots + \sigma_k(A)^p\right)^{1/p}, \quad p \geq 1,\ k = 1, \ldots, n,$$

are UI norms. If $p = 1$ we obtain the Ky Fan k norms

$$\|A\|_k = \sigma_1(A) + \cdots + \sigma_k(A), \quad k = 1, \ldots, n.$$

These latter norms are important because of the following result [15, Corollary 3.5.9]: If $A, B \in \mathbb{C}^{m\times n}$, then $\|A\| \leq \|B\|$ for every UI norm on $\mathbb{C}^{m\times n}$ if and only if $\|A\|_k \leq \|B\|_k$ for $k = 1, \ldots, n$.

From the previous characterization it follows that $\|A\| = \|A^*\|$ for every UI norm on $\mathbb{C}^{n\times n}$ because the singular values of A are the same as those of A^*. A norm $\|\cdot\|_Q$ is said to be a Q-norm on $\mathbb{C}^{m\times n}$ if there is some unitarily invariant norm $|||\cdot|||$ on $\mathbb{C}^{n\times n}$ such that $\|A\|_Q = |||A^*A|||^{1/2}$ for all $A \in \mathbb{C}^{m\times n}$ [1, Def. IV.2.9]. The norm $\|\cdot\|_Q$ is normalized if the UI norm $|||\cdot|||$ is normalized. Notice that a Q-norm is itself a UI norm. The Ky Fan p-k norms with $p \geq 2$ are Q-norms. Indeed,

$$\|A\|_{p,k} = ((\sigma_1(A^*A))^{p/2} + \cdots + (\sigma_k(A^*A))^{p/2})^{1/p} = \|A^*A\|_{p/2,k}^{1/2}.$$

In particular the spectral and Frobenius matrix norms are Q-norms.

An UI norm $\|\cdot\|$ on $\mathbb{C}^{m\times m}$ induces a UI norm on $\mathbb{C}^{r\times s}$ for any r, s with $\max\{r, s\} \leq m$. For any $A \in \mathbb{C}^{r\times s}$ define $\|A\| = \|\overline{A}\|$, where

$$\overline{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{C}^{m\times m}$$

has been augmented by zero blocks to fill out its size to $m \times m$. These norms will be called a family of UI norms and we will denote by the symbol $\|\cdot\|$ any such family.

Given any family of UI norms on $\mathbb{C}^{m\times m}$, it can be proved that [21, Theorem 3.9, p.80]

$$\|AB\| \leq \|A\|\|B\|_2 \quad \text{and} \quad \|AB\| \leq \|A\|_2\|B\|, \tag{2.1}$$

for all $A \in \mathbb{C}^{r\times s}$, $B \in \mathbb{C}^{s\times t}$ with $\max\{r, s, t\} \leq m$. If we consider a family of UI normalized norms, $\|AB\| \leq \|A\|\|B\|$ also holds. In this paper we will assume that all unitarily invariant norms, included the Q-norms, are normalized, and therefore consistent. We end this section by presenting two Lemmas that will be needed in the rest of the sections[1].

LEMMA 2.1. *Let $A, B \in \mathbb{C}^{m\times n}$ and let $\|\cdot\|$ be a UI norm on $\mathbb{C}^{n\times n}$. Then*
(a) $\|A^*B\|^2 \leq \|A^*A\|\|B^*B\|$.
(b) *If $P \in \mathbb{C}^{m\times m}$ is an orthogonal projector, then $\|A^*PA\| \leq \|A^*A\|$.*
Proof.
(a) See [16, Eq.(3.5.22)].

[1] Lemma 2.2 has appeared as Lemma 2.1 in [5] for the particular case of the 2-norm.

(b) Since $P^* = P = P^2$, the Cauchy-Schwarz type inequality given in (a) implies

$$\|A^*PA\|^2 \leq \|A^*A\| \|A^*PA\|.$$

Hence $\|A^*PA\| \leq \|A^*A\|$.

□

LEMMA 2.2. *Let $B, C \in \mathbb{C}^{m\times n}$, let $\mathcal{S} \subseteq \mathbb{C}^m$ and $\mathcal{W} \subseteq \mathbb{C}^n$ be two vector subspaces, and let $P_\mathcal{S} \in \mathbb{C}^{m\times m}$ and $P_\mathcal{W} \in \mathbb{C}^{n\times n}$ be the orthogonal projectors onto, respectively, $\mathcal{S}$ and $\mathcal{W}$. Let $\|\cdot\|_Q$ be a Q-norm on $\mathbb{C}^{m\times n}$. Then the following statements hold:*

(a) $\|P_\mathcal{S} B + (I - P_\mathcal{S})C\|_Q \leq \sqrt{\|B\|_Q^2 + \|C\|_Q^2}$.

(b) $\|B P_\mathcal{W} + C(I - P_\mathcal{W})\|_Q \leq \sqrt{\|B\|_Q^2 + \|C\|_Q^2}$.

Proof.

(a) Since

$$(P_\mathcal{S} B + (I - P_\mathcal{S})C)^*(P_\mathcal{S} B + (I - P_\mathcal{S})C) = B^* P_\mathcal{S} B + C^*(I - P_\mathcal{S})C,$$

we have

$$\|P_\mathcal{S} B + (I - P_\mathcal{S})C\|_Q^2 = \|\!|B^* P_\mathcal{S} B + C^*(I - P_\mathcal{S})C\|\!| \leq \|\!|B^* P_\mathcal{S} B\|\!| + \|\!|C^*(I - P_\mathcal{S})C\|\!|.$$

Now, from second item in Lemma 2.1 it follows that

$$\|P_\mathcal{S} B + (I - P_\mathcal{S})C\|_Q^2 \leq \|B\|_Q^2 + \|C\|_Q^2.$$

(b) Let $A = B P_\mathcal{W} + C(I - P_\mathcal{W})$. Then $A^* = P_\mathcal{W} B^* + (I - P_\mathcal{W})C^*$. Now, apply part (a) to A^*.

□

3. Multiplicative perturbations for the Moore-Penrose inverse: expressions and bounds. In this section and in Section 4, we consider a multiplicative perturbation of a general matrix $A \in \mathbb{C}^{m\times n}$, that is, a matrix $\widetilde{A} = (I + E)A(I + F)$, where $(I + E) \in \mathbb{C}^{m\times m}$ and $(I + F) \in \mathbb{C}^{n\times n}$ are nonsingular matrices. We are interested in finding expressions for $\widetilde{A}^\dagger$ and, from them, in finding bounds for the relative change of $A^\dagger$ in any UI norm and in any Q-norm. In Theorem 3.5 we get those bounds relative to $\|A^\dagger\|_2$ or to $\|\widetilde{A}^\dagger\|_2$, and not to $\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}$ as it was done in [3], where similar bounds where obtained. The results in [3] will be compared with the new ones in this paper after Theorem 3.5 and we will show that the new ones are superior.

We start with the following technical Lemma[2] that will be needed in the study of the multiplicative perturbation of the Moore-Penrose inverse.

LEMMA 3.1. *Let $A \in \mathbb{C}^{m\times n}$ and $\widetilde{A} = (I + E)A(I + F) \in \mathbb{C}^{m\times n}$, where $(I + E) \in \mathbb{C}^{m\times m}$ and $(I + F) \in \mathbb{C}^{n\times n}$ are nonsingular matrices. Let $\widehat{E} = (I + E)^{-1}E$, $\widehat{F} = (I + F)^{-1}F$, and let $\|\cdot\|$ be a family of unitarily invariant normalized norms on $\mathbb{C}^{q\times q}$, where $q = \max\{m, n\}$, and let us define:*

$$\mathcal{E} := \min\{\|E\|, \|\widehat{E}\|\} \quad \text{and} \quad \mathcal{F} := \min\{\|F\|, \|\widehat{F}\|\}. \tag{3.1}$$

Then the following hold:

[2]The first equation in both items (a) and (b) of Lemma 3.1 have appeared already in Lemma 3.1 in [5].

(a) $P_A(I-P_{\widetilde{A}}) = -P_A E^*(I-P_{\widetilde{A}})$ *and* $P_A = (I+E)^{-1}P_{\widetilde{A}}(I+E)P_A$,
(b) $(I-P_{\widetilde{A}^*})P_{A^*} = -(I-P_{\widetilde{A}^*})F^*P_{A^*}$ *and* $P_{A^*} = (I+F^*)^{-1}P_{\widetilde{A}^*}(I+F^*)P_{A^*}$,
(c) $\|P_{\widetilde{A}}(I-P_A)\| = \|P_A(I-P_{\widetilde{A}})\| \leq \min\{\|EP_A\|, \|(I-P_A)\widehat{E}\|\} \leq \mathcal{E}$,
(d) $\|P_{\widetilde{A}^*}(I-P_{A^*})\| = \|P_{A^*}(I-P_{\widetilde{A}^*})\| \leq \min\{\|P_{A^*}F\|, \|\widehat{F}(I-P_{A^*})\|\} \leq \mathcal{F}$.

Proof.
(a) Since $\mathcal{R}(\widetilde{A}) = \mathcal{R}((I+E)A)$ then $(I-P_{\widetilde{A}})(I+E)A = 0$. Thus, $(I-P_{\widetilde{A}})(I+E)AA^\dagger = (I-P_{\widetilde{A}})(I+E)P_A = 0$, which is equivalent to $P_A(I+E^*)(I-P_{\widetilde{A}}) = 0$. We have also that $P_A(I+E^*) = P_A(I+E^*)P_{\widetilde{A}}$ or $P_A = P_A(I+E^*)P_{\widetilde{A}}(I+E^*)^{-1}$. Hence (a) holds.
(b) Apply (a) to $\widetilde{A}^* = (I+F^*)A^*(I+E^*)$ and conjugate and transpose the first equality.
(c) The subspaces $\mathcal{R}(A)$ and $\mathcal{R}(\widetilde{A})$ have the same dimension. Thus, from [22, Theorem 7.1], the equality $\|P_{\widetilde{A}}(I-P_A)\| = \|P_A(I-P_{\widetilde{A}})\|$ holds for all UI norms. Moreover, by part (a), $\|P_A(I-P_{\widetilde{A}})\| = \|-P_AE^*(I-P_{\widetilde{A}})\| \leq \|P_AE^*\| = \|EP_A\|$. On the other hand, if we write $A = (I-\widehat{E})\widetilde{A}(I-\widehat{F})$, similarly, we get $\|P_{\widetilde{A}}(I-P_A)\| = \|P_{\widetilde{A}}\widehat{E}^*(I-P_A)\| \leq \|\widehat{E}^*(I-P_A)\| = \|(I-P_A)\widehat{E}\|$, which completes the proof of (c).
(d) Part (d) follows from applying part (c) to $\widetilde{A}^* = (I+F^*)A^*(I+E^*)$.

□

REMARK 3.2. *We have used in the proof of Lemma 3.1(c) that $\widetilde{A} = (I+E)A(I+F)$ can be written equivalently as,*

$$A = (I+E)^{-1}\widetilde{A}(I+F)^{-1} = (I-\widehat{E})\widetilde{A}(I-\widehat{F}). \tag{3.2}$$

Therefore, equivalent expressions to those in Lemma 3.1 can be obtained just by the simultaneous interchanges

$$A \longleftrightarrow \widetilde{A}, \quad E \longleftrightarrow -\widehat{E}, \quad \text{and} \quad F \longleftrightarrow -\widehat{F}. \tag{3.3}$$

For example, by doing that to the equations in item (a) in Lemma 3.1 we get the following expressions that will be used later in this paper

$$P_{\widetilde{A}}(I-P_A) = P_{\widetilde{A}}\widehat{E}^*(I-P_A) \quad \text{and} \quad P_{\widetilde{A}} = (I+E)P_A(I-\widehat{E})P_{\widetilde{A}}. \tag{3.4}$$

We will refer to the relationship: Lemma 3.1(a) $\leftrightarrow$ (3.4), as "dual" expressions under the transformations in (3.3). They will be used extensively in this paper to get new expressions from the "original" ones when needed.

Our next step will be to analyze the variation of the orthogonal projectors P_A and P_{A^*} under multiplicative perturbations. The bounds in Theorem 3.3 for the 2-norm will be used in Section 4. The rest of the bounds are included for completeness.

THEOREM 3.3. *Let $A \in \mathbb{C}^{m\times n}$ and $\widetilde{A} = (I+E)A(I+F) \in \mathbb{C}^{m\times n}$, where $(I+E) \in \mathbb{C}^{m\times m}$ and $(I+F) \in \mathbb{C}^{n\times n}$ are nonsingular matrices. Let $\|\cdot\|$ be a family of normalized UI norms and $\|\cdot\|_Q$ be a family of normalized Q-norms. Then*

$$\|P_{\widetilde{A}} - P_A\| \leq 2\mathcal{E}, \quad \|P_{\widetilde{A}} - P_A\|_Q \leq \sqrt{2}\,\mathcal{E}_Q, \quad \text{and} \quad \|P_{\widetilde{A}} - P_A\|_2 \leq \mathcal{E}_2, \tag{3.5}$$

$$\|P_{\widetilde{A}^*} - P_{A^*}\| \leq 2\mathcal{F}, \quad \|P_{\widetilde{A}^*} - P_{A^*}\|_Q \leq \sqrt{2}\,\mathcal{F}_Q, \quad \text{and} \quad \|P_{\widetilde{A}^*} - P_{A^*}\|_2 \leq \mathcal{F}_2, \tag{3.6}$$

$$\|P_{\mathcal{N}(\widetilde{A})} - P_{\mathcal{N}(A)}\| \leq 2\,\mathcal{F}, \quad \|P_{\mathcal{N}(\widetilde{A})} - P_{\mathcal{N}(A)}\|_Q \leq \sqrt{2}\,\mathcal{F}_Q, \quad \text{and} \quad \|P_{\mathcal{N}(\widetilde{A})} - P_{\mathcal{N}(A)}\|_2 \leq \mathcal{F}_2, \tag{3.7}$$

where $\mathcal{E}, \mathcal{E}_Q, \mathcal{E}_2, \mathcal{F}, \mathcal{F}_Q, \mathcal{F}_2$ *are defined as in* (3.1)*, in each case for the corresponding norm.*

Proof. To prove (3.5), we write $P_{\widetilde{A}} - P_A = P_{\widetilde{A}}(I - P_A) - (I - P_{\widetilde{A}})P_A$. Then the first inequality in (3.5) follows from Lemma 3.1(c). Taking Q-norms and applying Lemma 2.2, and again Lemma 3.1(c) we get

$$\begin{aligned} \|P_{\widetilde{A}} - P_A\|_Q^2 &= \|P_{\widetilde{A}}(I - P_A)(I - P_A) - (I - P_{\widetilde{A}})P_A P_A\|_Q^2 \\ &\le \|P_{\widetilde{A}}(I - P_A)\|_Q^2 + \|(I - P_{\widetilde{A}})P_A\|_Q^2 \\ &\le 2\,\mathcal{E}_Q^2. \end{aligned}$$

In addition, it follows from [21, Ch. I, Theorem 5.5] and from [22, Theorem 7.1], that $\|P_{\widetilde{A}} - P_A\|_2 = \|P_{\widetilde{A}}(I - P_A)\|_2 = \|P_A(I - P_{\widetilde{A}})\|_2$; then the third inequality in (3.5) follows immediately from Lemma 3.1(c). The bounds in (3.6) follow from applying the bounds in (3.5) to $\widetilde{A}^* = (I + F^*)A^*(I + E^*)$. Finally, (3.7) follows from (3.6) using $P_{\mathcal{N}(A)} = I - P_{A^*}$, and $P_{\mathcal{N}(\widetilde{A})} = I - P_{\widetilde{A}^*}$. □

Now we present the two main results of this section, that is, Theorems 3.4 and 3.5. First we express $\widetilde{A}^\dagger$ as a projected multiplicative perturbation of $A^\dagger$, and we write the difference between them explicitly in terms of E and F and also of the perturbed projectors $P_{\widetilde{A}^*}$ and $P_{\widetilde{A}}$. Parts of Theorem 3.4 have appeared already in [5, Theorem 3.2]. We present them here again for completeness. Besides, the equations in (3.9) are new results, and also we present a different proof.

Theorem 3.4. *Let* $A \in \mathbb{C}^{m\times n}$ *and* $\widetilde{A} = (I + E)A(I + F) \in \mathbb{C}^{m\times n}$*, where* $(I+E) \in \mathbb{C}^{m\times m}$ *and* $(I+F) \in \mathbb{C}^{n\times n}$ *are nonsingular matrices, and* $I - \widehat{E} := (I+E)^{-1}$, $I - \widehat{F} := (I + F)^{-1}$. *Then*

$$\widetilde{A}^\dagger = P_{\widetilde{A}^*}(I + F)^{-1}A^\dagger(I + E)^{-1}P_{\widetilde{A}}, \tag{3.8}$$

$$\widetilde{A}^\dagger(I + E)P_A = P_{\widetilde{A}^*}(I - \widehat{F})A^\dagger, \qquad A^\dagger(I - \widehat{E})P_{\widetilde{A}} = P_{A^*}(I + F)\widetilde{A}^\dagger, \tag{3.9}$$

and

$$\widetilde{A}^\dagger - A^\dagger = A^\dagger\Theta_E + \Theta_F A^\dagger + \Theta_F A^\dagger\Theta_E, \tag{3.10}$$

where

$$\Theta_E := E^*(I - P_{\widetilde{A}}) - \widehat{E}P_{\widetilde{A}} \quad \text{and} \quad \Theta_F := (I - P_{\widetilde{A}^*})F^* - P_{\widetilde{A}^*}\widehat{F}\,. \tag{3.11}$$

Proof. First, we observe that

$$\widetilde{A}^\dagger = \widetilde{A}^\dagger\widetilde{A}(I + F)^{-1}A^\dagger(I + E)^{-1}\widetilde{A}\widetilde{A}^\dagger = P_{\widetilde{A}^*}(I + F)^{-1}A^\dagger(I + E)^{-1}P_{\widetilde{A}},$$

and (3.8) holds.

To prove (3.9), first let us write, using (3.8):

$$\widetilde{A}^\dagger(I + E)P_A = P_{\widetilde{A}^*}(I + F)^{-1}A^\dagger(I + E)^{-1}P_{\widetilde{A}}(I + E)P_A, \tag{3.12}$$

and noticing that the first equation in Lemma 3.1(a) is equivalent to $(I - P_{\widetilde{A}})(I + E)P_A = 0$ we get

$$\widetilde{A}^\dagger(I + E)P_A = P_{\widetilde{A}^*}(I - \widehat{F})A^\dagger(I + E)^{-1}(I + E)P_A = P_{\widetilde{A}^*}(I - \widehat{F})A^\dagger P_A, \tag{3.13}$$

that is the first equation in (3.9). The second equation in (3.9) is the dual, according to the transformation (3.3), of the first one.

Finally to get (3.10), it follows from (a) and (b) in Lemma 3.1 that

$$A^\dagger\Theta_E = A^\dagger[-(I-P_{\widetilde{A}}) - (I+E)^{-1}EP_{\widetilde{A}}] = A^\dagger((I+E)^{-1}P_{\widetilde{A}} - I), \tag{3.14}$$

$$\Theta_F A^\dagger = [-(I-P_{\widetilde{A}^*}) - P_{\widetilde{A}^*}(I+F)^{-1}F]A^\dagger = (P_{\widetilde{A}^*}(I+F)^{-1} - I)A^\dagger, \tag{3.15}$$

$$\Theta_F A^\dagger\Theta_E = (P_{\widetilde{A}^*}(I+F)^{-1} - I)A^\dagger((I+E)^{-1}P_{\widetilde{A}} - I), \tag{3.16}$$

hence,

$$A^\dagger\Theta_E + \Theta_F A^\dagger + \Theta_F A^\dagger\Theta_E = P_{\widetilde{A}^*}(I+F)^{-1}A^\dagger(I+E)^{-1}P_{\widetilde{A}} - A^\dagger = \widetilde{A}^\dagger - A^\dagger.$$

That is (3.10). □

We emphasize that expression (3.8) can be rewritten in a way (see equation (3.2) in [5, Theorem 3.2]) that ensures that under "small" multiplicative perturbations of A, i.e., small E and F, we obtain "small" multiplicative perturbations of $A^\dagger$. Using Theorem 3.4 we can bound the relative changes for $\|\widetilde{A}^\dagger - A^\dagger\|$ and $\|\widetilde{A}^\dagger - A^\dagger\|_Q$ in terms of the perturbation matrices E and F.

THEOREM 3.5. *Let $A \in \mathbb{C}^{m\times n}$ and $\widetilde{A} = (I+E)A(I+F) \in \mathbb{C}^{m\times n}$, where $(I+E) \in \mathbb{C}^{m\times m}$ and $(I+F) \in \mathbb{C}^{n\times n}$ are nonsingular matrices, and $I-\widehat{E} := (I+E)^{-1}$, $I-\widehat{F} := (I+F)^{-1}$. Let us denote by $\|\cdot\|$ a family of normalized unitarily invariant norms and by $\|\cdot\|_Q$ a family of normalized Q-norms. Then the following bounds hold:*

$$\frac{\|\widetilde{A}^\dagger - A^\dagger\|}{\|A^\dagger\|_2} \le \|E\| + \|\widehat{E}\| + \|F\| + \|\widehat{F}\| + \|\widehat{E}\|\|\widehat{F}\|, \tag{3.17}$$

$$\frac{\|\widetilde{A}^\dagger - A^\dagger\|}{\|\widetilde{A}^\dagger\|_2} \le \|E\| + \|\widehat{E}\| + \|F\| + \|\widehat{F}\| + \|E\|\|F\|, \tag{3.18}$$

$$\frac{\|\widetilde{A}^\dagger - A^\dagger\|_Q}{\|A^\dagger\|_2} \le \sqrt{\|E\|_Q^2 + \|F\|_Q^2 + (\|\widehat{E}\|_Q + \|\widehat{F}\|_Q + \|\widehat{E}\|_Q\,\|\widehat{F}\|_Q)^2}, \tag{3.19}$$

$$\frac{\|\widetilde{A}^\dagger - A^\dagger\|_Q}{\|\widetilde{A}^\dagger\|_2} \le \sqrt{\|\widehat{E}\|_Q^2 + \|\widehat{F}\|_Q^2 + (\|E\|_Q + \|F\|_Q + \|E\|_Q\,\|F\|_Q)^2}. \tag{3.20}$$

Proof. The bound in (3.17) is obtained by first rewriting equation (3.10) in Theorem 3.4 as

$$\widetilde{A}^\dagger - A^\dagger = (I+\Theta_F)A^\dagger\Theta_E + \Theta_F A^\dagger, \tag{3.21}$$

or as

$$\widetilde{A}^\dagger - A^\dagger = \Theta_F A^\dagger(I+\Theta_E) + A^\dagger\Theta_E, \tag{3.22}$$

and from (3.14) and (3.15) we have

$$A^\dagger(I+\Theta_E) = A^\dagger(I+E)^{-1}P_{\widetilde{A}} = A^\dagger(I-\widehat{E})P_{\widetilde{A}}, \tag{3.23}$$

$$(I+\Theta_F)A^\dagger = P_{\widetilde{A}^*}(I+F)^{-1}A^\dagger = P_{\widetilde{A}^*}(I-\widehat{F})A^\dagger. \tag{3.24}$$

From (3.21) and (3.24) we get

$$\widetilde{A}^{\dagger} - A^{\dagger} = P_{\widetilde{A}^*}(I - \widehat{F})A^{\dagger}\Theta_E + \Theta_F A^{\dagger}, \tag{3.25}$$

and from (3.22) and (3.23)

$$\widetilde{A}^{\dagger} - A^{\dagger} = \Theta_F A^{\dagger}(I - \widehat{E})P_{\widetilde{A}} + A^{\dagger}\Theta_E. \tag{3.26}$$

From (3.25), and using (3.23) and (3.11),

$$\begin{aligned}
\|\widetilde{A}^{\dagger} - A^{\dagger}\| &= \|P_{\widetilde{A}^*}(I - \widehat{F})A^{\dagger}\Theta_E + \Theta_F A^{\dagger}\| \\
&= \|P_{\widetilde{A}^*}(I - \widehat{F})A^{\dagger}\Theta_E + (I - P_{\widetilde{A}^*})F^* A^{\dagger} - P_{\widetilde{A}^*}\widehat{F}A^{\dagger}\| \\
&\leq \|P_{\widetilde{A}^*}A^{\dagger}\Theta_E\| + \|P_{\widetilde{A}^*}\widehat{F}A^{\dagger}(I + \Theta_E)\| + \|F^*\|\|A^{\dagger}\|_2 \\
&\leq \|A^{\dagger}\|_2(\|E\| + \|\widehat{E}\| + \|F\| + \|\widehat{F}\|) + \|P_{\widetilde{A}^*}\widehat{F}A^{\dagger}\widehat{E}P_{\widetilde{A}}\| \\
&\leq \|A^{\dagger}\|_2(\|E\| + \|\widehat{E}\| + \|F\| + \|\widehat{F}\| + \|\widehat{E}\|\|\widehat{F}\|),
\end{aligned}$$

which is (3.17).

To get the bound in (3.19) we use (3.21), (3.24), (3.23) and (3.11) to prove that

$$\begin{aligned}
\widetilde{A}^{\dagger} - A^{\dagger} &= (I + \Theta_F)A^{\dagger}\Theta_E + \Theta_F A^{\dagger} \\
&= P_{\widetilde{A}^*}(I - \widehat{F})A^{\dagger}\Theta_E + \Theta_F A^{\dagger} \\
&= P_{\widetilde{A}^*}\left[(I - \widehat{F})A^{\dagger}\Theta_E - \widehat{F}A^{\dagger}\right] + (I - P_{\widetilde{A}^*})F^* A^{\dagger} \\
&= P_{\widetilde{A}^*}\left[A^{\dagger}\Theta_E - \widehat{F}A^{\dagger}(I + \Theta_E)\right] + (I - P_{\widetilde{A}^*})F^* A^{\dagger} \\
&= P_{\widetilde{A}^*}\left[A^{\dagger}E^*(I - P_{\widetilde{A}}) - A^{\dagger}\widehat{E}P_{\widetilde{A}} - \widehat{F}A^{\dagger}(I - \widehat{E})P_{\widetilde{A}}\right] + (I - P_{\widetilde{A}^*})F^* A^{\dagger}.
\end{aligned}$$

It remains to apply Lemma 2.2 to the last equation above and get

$$\begin{aligned}
\|\widetilde{A}^{\dagger} - A^{\dagger}\|_Q^2 &\leq \|A^{\dagger}E^*(I - P_{\widetilde{A}}) - [A^{\dagger}\widehat{E} + \widehat{F}A^{\dagger}(I - \widehat{E})]\, P_{\widetilde{A}}\|_Q^2 + \|F^* A^{\dagger}\|_Q^2 \\
&\leq \|A^{\dagger}E^*\|_Q^2 + \|A^{\dagger}\widehat{E} + \widehat{F}A^{\dagger}(I - \widehat{E})\|_Q^2 + \|F^* A^{\dagger}\|_Q^2 \\
&\leq \|A^{\dagger}\|_2^2 \left(\|E\|_Q^2 + \|F\|_Q^2 + (\|\widehat{E}\|_Q + \|\widehat{F}\|_Q + \|\widehat{E}\|_Q\, \|\widehat{F}\|_Q)^2 \right).
\end{aligned}$$

The last inequality being a consequence of (2.1) and the fact that the Q-norms are normalized UI matrix norms. The inequality above gives the bound for $\|\widetilde{A}^{\dagger} - A^{\dagger}\|_Q / \|A^{\dagger}\|_2$ in (3.19).

The bounds in (3.18) and (3.20) are the dual expressions, according to the transformation (3.3), respectively, of (3.17) and (3.19). □

We highlight now the following points on Theorem 3.5.

REMARK 3.6.

(a) *The bounds in Theorem 3.5 improve significatively the classical bounds for the relative variation of the Moore-Penrose inverse under general additive perturbations* $\widetilde{A} = A + \Delta A$ *in the case* $\mathrm{rank}(\widetilde{A}) = \mathrm{rank}(A)$ *(see [22, Theorem 4.1] or the rearrangement in [21, Ch. III, Corollary 3.10]). The crucial point is that the bounds in Theorem 3.5 do not depend on* $\kappa_2(A)$*, while the classical bounds do.*

(b) *The bounds* (3.17)-(3.18) *in Theorem 3.5 have the advantage that are valid for any normalized unitarily invariant norm, but when they are particularized to* $\|\cdot\|_Q$, *then the bounds in* (3.19)-(3.20) *are always sharper than the ones in* (3.17)-(3.18).

(c) *If A has full row rank, then $\widetilde{A}$ has also full row rank and $P_{\widetilde{A}} = I_m$. Thus, Θ_E in* (3.11) *simplifies to $\Theta_E = -\widehat{E}$ and all the terms containing $\|E\|$ or $\|E\|_Q$ in the bounds of Theorem 3.5 vanish (but one should keep $\|\widehat{E}\|$ and $\|\widehat{E}\|_Q$).*

(d) *If A has full column rank, then $\widetilde{A}$ has also full column rank and $P_{\widetilde{A}^*} = I_n$. Thus, Θ_F in* (3.11) *simplifies to $\Theta_F = -\widehat{F}$ and all the terms containing $\|F\|$ or $\|F\|_Q$ in the bounds of Theorem 3.5 vanish (but one should keep $\|\widehat{F}\|$ and $\|\widehat{F}\|_Q$).*

(e) *If we restrict in Theorem 3.5 the magnitude of the perturbations to be* $\max\{\|E\|_2, \|F\|_2\} < 1$, *then, standard perturbation theory of the inverse [21, Ch. III, Theorem 2.5] and* (2.1) *provide inequalities that can be used in Theorem 3.5 to obtain bounds that are easily computable just in terms of E and F. To this purpose note first that $\|\widehat{E}\| = \|(I+E)^{-1}E\| \leq \|(I+E)^{-1}\|_2\|E\|$, and that an analogous result holds for $\|\widehat{F}\|$, and second that $\|(I+E)^{-1}\|_2 \leq 1/(1-\|E\|_2)$, and the analogue for $\|(I+F)^{-1}\|_2$. As a consequence, we get*

$$\|\widehat{E}\| \leq \frac{\|E\|}{1-\|E\|_2} \quad \text{and} \quad \|\widehat{F}\| \leq \frac{\|F\|}{1-\|F\|_2}. \tag{3.27}$$

(f) *Finally, again with the additional restriction* $\max\{\|E\|_2, \|F\|_2\} < 1$, *Theorem 3.5 can be completed with the following lower and upper bounds for $\|\widetilde{A}^\dagger\|$:*

$$\frac{\|A^\dagger\|}{(1+\|E\|_2)(1+\|F\|_2)} \leq \|\widetilde{A}^\dagger\| \leq \frac{\|A^\dagger\|}{(1-\|E\|_2)(1-\|F\|_2)}.$$

The upper bound follows from (3.8), *the properties used in part (e), and also from $\|P_{\widetilde{A}}\|_2 = \|P_{\widetilde{A}^*}\|_2 = 1$. For getting the lower bound: consider A as a multiplicative perturbation of $\widetilde{A}$, i.e., $A = (I+E)^{-1}\widetilde{A}(I+F)^{-1}$, and apply* (3.8) *with the roles of A and $\widetilde{A}$ exchanged to get $A^\dagger = P_{A^*}(I+F)\widetilde{A}^\dagger(I+E)P_A$.*

As it was said at the beginning of this section and in the Introduction, multiplicative perturbation bounds for the Moore-Penrose inverse have also been presented recently in [3, Section 4]. The bounds in [3] are not based on expressions for $\widetilde{A}^\dagger$ as those in Theorem 3.4, they are obtained following a different approach and therefore they are different. In the rest of this section we are going to compare some of the bounds in [3] with the results of Theorem 3.5.

In order to do that, first we present Theorem 3.7, in which we rewrite in the notation of this paper Theorems 4.1 and 4.2 in [3]. Moreover Theorem 3.7 includes a new result: the bound in (3.30). We will show a new proof of Theorem 4.1 in [3], that is, the bound (3.28) below, using the techniques of this paper and the proof of the new result (3.30), that is similar and comparable to (3.29) [3, Theorem 4.2].

THEOREM 3.7. *[3, Theorems 4.1 and 4.2] With the same hypotheses as in Theorem 3.5:*

$$\frac{\|\widetilde{A}^\dagger - A^\dagger\|}{\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}} \leq \|E\| + \|\widehat{E}\| + \|F\| + \|\widehat{F}\|, \tag{3.28}$$

$$\frac{\|\widetilde{A}^\dagger - A^\dagger\|_Q}{\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}} \leq \sqrt{\frac{3}{2}}\,\sqrt{\|E\|_Q^2 + \|\widehat{E}\|_Q^2 + \|F\|_Q^2 + \|\widehat{F}\|_Q^2}. \tag{3.29}$$

Furthermore, the following bound alternative to (3.29) *also holds:*

$$\frac{\|\widetilde{A}^\dagger - A^\dagger\|_Q}{\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}} \leq \sqrt{\|E\|_Q^2 + \|\widehat{E}\|_Q^2 + \|F\|_Q^2 + \|\widehat{F}\|_Q^2 + 2\,\|E\|_Q\,\|\widehat{F}\|_Q}. \tag{3.30}$$

Proof. To get the bound in (3.28) we use the first equation in (3.9) and (3.25)

$$\|\widetilde{A}^\dagger - A^\dagger\| = \|P_{\widetilde{A}^*}(I-\widehat{F})A^\dagger\Theta_E + \Theta_F A^\dagger\| = \|\widetilde{A}^\dagger(I+E)P_A\Theta_E + \Theta_F A^\dagger\|. \tag{3.31}$$

The term $\widetilde{A}^\dagger(I+E)P_A\Theta_E$ can be written, using (3.14) and the second equation in (3.4), as

$$\begin{aligned} \widetilde{A}^\dagger(I+E)P_A\Theta_E &= \widetilde{A}^\dagger(I+E)P_A((I+E)^{-1}P_{\widetilde{A}} - I) \\ &= \widetilde{A}^\dagger P_{\widetilde{A}} - \widetilde{A}^\dagger(I+E)P_A = \widetilde{A}^\dagger(I-P_A) - \widetilde{A}^\dagger E P_A \\ &= \widetilde{A}^\dagger\widehat{E}^*(I-P_A) - \widetilde{A}^\dagger E P_A, \end{aligned} \tag{3.32}$$

where we have used the first equation in (3.4) for the last equality. Finally, by using (3.15) and the first equation in Lemma 3.1 (b), the second term in (3.31) is written as

$$\Theta_F A^\dagger = (P_{\widetilde{A}^*} - I)A^\dagger - P_{\widetilde{A}^*}\widehat{F}A^\dagger = (P_{\widetilde{A}^*} - I)F^*A^\dagger - P_{\widetilde{A}^*}\widehat{F}A^\dagger. \tag{3.33}$$

Now substituting (3.32) and (3.33) into (3.31) we get

$$\begin{aligned} \|\widetilde{A}^\dagger - A^\dagger\| &\leq \|\widetilde{A}^\dagger(I+E)P_A\Theta_E\| + \|\Theta_F A^\dagger\| \\ &\leq \|\widetilde{A}^\dagger\widehat{E}^*(I-P_A) - \widetilde{A}^\dagger E P_A\| + \|(P_{\widetilde{A}^*} - I)F^*A^\dagger - P_{\widetilde{A}^*}\widehat{F}A^\dagger\| \\ &\leq \|\widetilde{A}^\dagger\|_2(\|\widehat{E}^*\| + \|E\|) + \|A^\dagger\|_2(\|F^*\| + \|\widehat{F}\|) \\ &= \|\widetilde{A}^\dagger\|_2(\|\widehat{E}\| + \|E\|) + \|A^\dagger\|_2(\|F\| + \|\widehat{F}\|), \end{aligned} \tag{3.34}$$

which leads to (3.28).

To get (3.30) we proceed as in (3.31) and (3.33), to obtain

$$\|\widetilde{A}^\dagger - A^\dagger\|_Q^2 = \|\widetilde{A}^\dagger(I+E)P_A\Theta_E - P_{\widetilde{A}^*}\widehat{F}A^\dagger - (I - P_{\widetilde{A}^*})F^*A^\dagger\|_Q^2. \tag{3.35}$$

Now, using Lemma 2.2, (3.32), $\widetilde{A}^\dagger = P_{\widetilde{A}^*}\widetilde{A}^\dagger$, and $A^\dagger = A^\dagger P_A$,

$$\begin{aligned} \|\widetilde{A}^\dagger - A^\dagger\|_Q^2 &\leq \|\widetilde{A}^\dagger(I+E)P_A\Theta_E - \widehat{F}A^\dagger\|_Q^2 + \|A^\dagger\|_2^2\|F\|_Q^2 \\ &\leq \|\widetilde{A}^\dagger\widehat{E}^*(I-P_A) - \widetilde{A}^\dagger E P_A - \widehat{F}A^\dagger\|_Q^2 + \|A^\dagger\|_2^2\|F\|_Q^2 \\ &\leq \|\widetilde{A}^\dagger\widehat{E}^*\|_Q^2 + \|\widetilde{A}^\dagger E - \widehat{F}A^\dagger\|_Q^2 + \|A^\dagger\|_2^2\|F\|_Q^2 \\ &\leq \|\widetilde{A}^\dagger\|_2^2\|\widehat{E}\|_Q^2 + \|A^\dagger\|_2^2\|F\|_Q^2 + (\|\widetilde{A}^\dagger\|_2\|E\|_Q + \|A^\dagger\|_2\|\widehat{F}\|_Q)^2, \end{aligned} \tag{3.36}$$

that can be converted easily into (3.30). □

Let us compare now the results of Theorems 3.5 and 3.7. To begin with note that the presence of $\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}$ in Theorem 3.7 makes it difficult to compare in general the bounds (3.28)-(3.29) with the bounds (3.17)-(3.20), but simultaneously it makes the bounds in Theorem 3.5 more natural and applicable, because the standard situation in perturbation theory is that $A^\dagger$ is known, but $\widetilde{A}^\dagger$ is not. In this context note, for instance, that if $\|A^\dagger\|_2 = \|\widetilde{A}^\dagger\|_2$, then the bound in (3.28) is obviously sharper than (3.17)-(3.18); but if $\|A^\dagger\|_2 \ll \|\widetilde{A}^\dagger\|_2$, then (3.28) does not give any information on $\|\widetilde{A}^\dagger - A^\dagger\|/\|A^\dagger\|_2$, while (3.17) does. *However, as we discuss next, the bounds in Theorem 3.5 are superior to* (3.28)-(3.29)*, both to first order and in terms of wider applicability.*

If we consider tiny perturbations and neglect second order terms, then we can replace both $\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}$ and $\|\widetilde{A}^\dagger\|_2$, simply by $\|A^\dagger\|_2$ (see Remark 3.6(f)), which allows us to make comparisons easier. The bounds (3.17)-(3.18) in Theorem 3.5 and (3.28) in Theorem 3.7 are equal to first order, that is, $\|\widetilde{A}^\dagger - A^\dagger\|/\|A^\dagger\|_2 \leq 2(\|E\| + \|F\|)$. However, to first order, the right-hand side of (3.29) is $\Xi_1 := \sqrt{3}\sqrt{\|E\|_Q^2 + \|F\|_Q^2}$ and the right-hand side of (3.19)-(3.20) is $\Xi_2 := \sqrt{\|E\|_Q^2 + \|F\|_Q^2 + (\|E\|_Q + \|F\|_Q)^2}$. To compare Ξ_1 and Ξ_2, use that $(x+y)^2 \leq 2(x^2+y^2)$, for $x \geq 0$, $y \geq 0$ real numbers. Thus

$$\Xi_2 \leq \sqrt{3\,\|E\|_Q^2 + 3\,\|F\|_Q^2} = \Xi_1,$$

which implies that, to first order, the bounds (3.19)-(3.20) are always sharper than (3.29).

For sufficiently large perturbations, the presence of $\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}$ makes (3.28)-(3.29) unapplicable in certain situations, since, as said before, one of the standard goals of perturbation theory is to bound $\|\widetilde{A}^\dagger - A^\dagger\|$ *without knowing* $\widetilde{A}^\dagger$ *and having only some bounds on the norms of the perturbations* E *and* F.

Let us illustrate this point with an example. Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad E = \begin{bmatrix} -4/5 & 0 & 0 \\ 0 & -4/5 & 0 \\ 0 & 0 & -4/5 \end{bmatrix}, \quad \text{and} \quad F = \begin{bmatrix} -4/5 & 0 \\ 0 & -4/5 \end{bmatrix}. \tag{3.37}$$

An easy computation shows that

$$\frac{\|\widetilde{A}^\dagger - A^\dagger\|_2}{\|A^\dagger\|_2} = 24, \quad \frac{\|\widetilde{A}^\dagger - A^\dagger\|_2}{\|\widetilde{A}^\dagger\|_2} = 0.96, \tag{3.38}$$

$$\|E\|_2 = 0.8, \quad \|F\|_2 = 0.8, \quad \|\widehat{E}\|_2 = 4, \quad \|\widehat{F}\|_2 = 4, \tag{3.39}$$

which give, using $\|.\|_Q = \|.\|_2$, the results in Table 3.1. It can be seen that all three bounds in Theorem 3.7 fail by far in getting a bound for $\|\widetilde{A}^\dagger - A^\dagger\|_2/\|A^\dagger\|_2$, but notice also that (3.30) is the sharpest of the three for $\|\widetilde{A}^\dagger - A^\dagger\|_2/\|\widetilde{A}^\dagger\|_2$, and, finally, notice that the bounds in Theorem 3.5 give the sharpest estimates, in particular those in (3.19) and (3.20).

4. Bounds for the solutions of the Least Squares Problem. In this section we consider the Least Squares Problem (LSP)

$$\min_{x\in\mathbb{C}^n} \|Ax - b\|_2, \quad A \in \mathbb{C}^{m\times n}, \quad b \in \mathbb{C}^m, \tag{4.1}$$

and the multiplicatively perturbed LSP:

$$\min_{x\in\mathbb{C}^n} \|\widetilde{A}x - \widetilde{b}\|_2, \quad \widetilde{A} = (I+E)A(I+F) \in \mathbb{C}^{m\times n}, \quad \widetilde{b} = b + h \in \mathbb{C}^m, \tag{4.2}$$

		Bounds in Theorem 3.5				Bounds in Theorem 3.7		
$\frac{\|\widetilde{A}^\dagger - A^\dagger\|_2}{\|A^\dagger\|_2}$	$\frac{\|\widetilde{A}^\dagger - A^\dagger\|_2}{\|\widetilde{A}^\dagger\|_2}$	(3.17)	(3.18)	(3.19)	(3.20)	(3.28)	(3.29)	(3.30)
24	0.96	25.6	10.24	24.03	6.08	9.6	7.07	6.30

Table 3.1

Example (3.37): Comparing the results of Theorems 3.5 and 3.7. Blue columns should be compared among themselves, and so the black columns.

where $(I+E) \in \mathbb{C}^{m\times m}$ and $(I+F) \in \mathbb{C}^{n\times n}$ are nonsingular matrices. We are interested in finding an upper bound for the relative variation $\|\widetilde{x}_0 - x_0\|_2/\|x_0\|_2$, where $x_0 = A^\dagger b$ and $\widetilde{x}_0 = \widetilde{A}^\dagger \widetilde{b}$ are, respectively, the minimum 2-norm solutions of (4.1) and (4.2). We will also examine the variation of the associated residuals $r = b - Ax_0$ and $\widetilde{r} = \widetilde{b} - \widetilde{A}\widetilde{x}_0$. We will use the multiplicative perturbation theory of $A^\dagger$ developed in Section 3. In addition, the results in this section are compared with the classic additive perturbation bounds of the solution of the LSP [22, Theorem 5.1]. As was said in the Introduction these results are very important for the numerical computation of solutions of structured LSP with high relative accuracy. In particular, one result of this section, the first order bound (4.9) in Corollary 4.2, has already been presented in [5, Theorem 4.1] where it was used in the error analysis of an algorithm that computes with high relative accuracy the minimum 2-norm solutions of structured LSP. Next theorem is the main result in this section.

Theorem 4.1. *Let x_0 and $\widetilde{x}_0$ be the minimum 2-norm solutions of (4.1) and (4.2), respectively, and let $r = b - Ax_0$ and $\widetilde{r} = \widetilde{b} - \widetilde{A}\widetilde{x}_0$ be the corresponding residuals. Let $\widehat{E} := (I+E)^{-1}E$ and $\widehat{F} := (I+F)^{-1}F$, define $\alpha_E := \sqrt{\|E\|_2^2 + \|\widehat{E}\|_2^2}$ and $\alpha_F := \sqrt{\|F\|_2^2 + \|\widehat{F}\|_2^2}$, and assume that $\|h\|_2 \le \epsilon \|b\|_2$. Then,*

$$\|\widetilde{x}_0 - x_0\|_2 \le \alpha_F\, \|x_0\|_2 + \left[\alpha_E\,(1+\alpha_F)\,(1+\epsilon) + \epsilon\,(1+\alpha_F)\right]\, \|A^\dagger\|_2\, \|b\|_2\,, \tag{4.3}$$

$$\|\widetilde{r} - r\|_2 \le \|b\|_2\, \sqrt{(\epsilon + \|E\|_2)^2 + \|E\|_2^2}\,. \tag{4.4}$$

Proof. Let us prove first (4.3). The proof is based on (3.10) in Theorem 3.4 that implies:

$$\begin{aligned}
\widetilde{x}_0 - x_0 &= \widetilde{A}^\dagger (b+h) - A^\dagger b \\
&= \left(\widetilde{A}^\dagger - A^\dagger\right)(b+h) + A^\dagger h \\
&= \left(A^\dagger \Theta_E + \Theta_F A^\dagger + \Theta_F A^\dagger \Theta_E\right)(b+h) + A^\dagger h \\
&= \left(A^\dagger \Theta_E + \Theta_F A^\dagger \Theta_E\right)(b+h) + \Theta_F x_0 + \Theta_F A^\dagger h + A^\dagger h.
\end{aligned}$$

Apply norm inequalities and get

$$\|\widetilde{x}_0 - x_0\|_2 \le \|\Theta_F\|_2 \|x_0\|_2 + \left[\|\Theta_E\|_2\,(1+\|\Theta_F\|_2)\,(1+\epsilon) + \epsilon\,(1+\|\Theta_F\|_2)\right]\, \|A^\dagger\|_2\, \|b\|_2\,.$$

Now, (4.3) follows from Lemma 2.2 that implies $\|\Theta_E\|_2 \leq \alpha_E$ and $\|\Theta_F\|_2 \leq \alpha_F$.

Next, we prove (4.4). First, observe that

$$\begin{aligned}
\widetilde{r} - r &= h - \widetilde{A}\widetilde{x}_0 + Ax_0 \\
&= h - \widetilde{A}\widetilde{A}^\dagger \widetilde{b} + Ax_0 \\
&= (I - \widetilde{A}\widetilde{A}^\dagger)h + Ax_0 - \widetilde{A}\widetilde{A}^\dagger b \\
&= (I - \widetilde{A}\widetilde{A}^\dagger)h + Ax_0 - \widetilde{A}\widetilde{A}^\dagger (r + Ax_0) \\
&= (I - \widetilde{A}\widetilde{A}^\dagger)(h + Ax_0) - \widetilde{A}\widetilde{A}^\dagger r. \qquad (4.5)
\end{aligned}$$

Note that the summands in (4.5) are orthogonal vectors, since $P_{\widetilde{A}} = \widetilde{A}\widetilde{A}^\dagger$, use $Ax_0 = P_A b$ and $r = (I - P_A)b$, recall Lemma 3.1(c), and get (4.4) as follows

$$\begin{aligned}
\|\widetilde{r} - r\|_2^2 &= \|(I - P_{\widetilde{A}})(h + P_A b)\|_2^2 + \|P_{\widetilde{A}}(I - P_A)b\|_2^2 \\
&\leq \left(\|h\|_2 + \|(I - P_{\widetilde{A}})P_A b\|_2\right)^2 + \|P_{\widetilde{A}}(I - P_A)b\|_2^2 \\
&\leq (\epsilon\|b\|_2 + \|E\|_2\|b\|_2)^2 + \|E\|_2^2\|b\|_2^2.
\end{aligned}$$

□

The bound (4.3) simplifies if A has full row or full column rank in the way explained in parts (c) and (d) of Remark 3.6. If A has full row rank, then $\widetilde{r} = r = 0$ and $\|\widetilde{r} - r\|_2 = 0$.

As said in the Introduction the bounds in Theorem 4.1 improve significatively the classical bounds (1.9-1.10) for the relative variation of minimum 2-norm solutions and residuals of LSP under general additive perturbations $\widetilde{A} = A + \Delta A$ [22, Theorem 5.1]. Let us compare them. First, it is convenient to bear in mind that$(\|A\|_2\|x_0\|_2)/\|b\|_2 \leq \kappa_2(A)$ and $\|r\|_2 \leq \|b\|_2$ in (1.10). Next, observe that the bound for $\|\widetilde{r} - r\|_2/\|b\|_2$ in (1.10) includes terms that can be very large even if ϵ_A and ϵ_b are very tiny. This happens if $\kappa_2(A)$ is large and $\|r\|_2 \neq 0$ is not very small. In contrast, if $\|E\|_2$ and ϵ are tiny, then the bound for $\|\widetilde{r} - r\|_2/\|b\|_2$ that follows from (4.4) is always tiny. With respect to the bounds for $\|\widetilde{x}_0 - x_0\|_2/\|x_0\|_2$: the bound in (1.9) amplifies the perturbations in the data at least by a factor $\kappa_2(A)$ and the amplification can be much larger under certain conditions. In addition, (1.9) includes the amplification factor $\|A^\dagger\|_2\, \|b\|_2/\|x_0\|_2$, which is the only potentially large factor in the bound that follows from (4.3). We will show in Subsection 4.2 that $\|A^\dagger\|_2\, \|b\|_2/\|x_0\|_2$ is a moderate number except for very particular choices of b. Therefore, (4.3) always improves (1.9) and, if $\|E\|_2$, $\|F\|_2$, and ϵ are tiny, then (4.3) produces tiny bounds for $\|\widetilde{x}_0 - x_0\|_2/\|x_0\|_2$ for almost all b.

The bounds in Theorem 4.1 depend both on $\widehat{E}$ and $\widehat{F}$. This reduces their applicability in practical situations. Corollary 4.2 overcomes this shortcoming by restricting the magnitude of the perturbations and by using (3.27). Corollary 4.2 follows directly from Theorem 4.1 and is stated in a way that is convenient for its use in error analysis of Least Squares algorithms, as was done in [5].

COROLLARY 4.2. *With the same notation and hypotheses that in Theorem 4.1, assume in addition that* $\|E\|_2 \leq \mu < 1$ *and* $\|F\|_2 \leq \nu < 1$, $x_0 \neq 0$, *and* $b \neq 0$. *Define*

$$\theta_\mu := \mu\sqrt{1 + \frac{1}{(1-\mu)^2}} \quad \textit{and} \quad \theta_\nu := \nu\sqrt{1 + \frac{1}{(1-\nu)^2}}\,. \qquad (4.6)$$

Then the following bounds hold:

$$\frac{\|\widetilde{x}_0 - x_0\|_2}{\|x_0\|_2} \le \theta_\nu + [\theta_\mu(1+\theta_\nu)(1+\epsilon) + \epsilon(1+\theta_\nu)]\,\frac{\|A^\dagger\|_2\|b\|_2}{\|x_0\|_2}\,, \tag{4.7}$$

$$\frac{\|\widetilde{r} - r\|_2}{\|b\|_2} \le \sqrt{(\epsilon+\mu)^2 + \mu^2}\,. \tag{4.8}$$

The bound (4.7) *yields to first order in* ϵ, μ, ν

$$\frac{\|\widetilde{x}_0 - x_0\|_2}{\|x_0\|_2} \le \sqrt{2}\,\nu + \left(\epsilon + \sqrt{2}\,\mu\right)\frac{\|A^\dagger\|_2\|b\|_2}{\|x_0\|_2} + h.o.t\,, \tag{4.9}$$

where $h.o.t$ *stands for "higher order terms" in* ϵ, μ, ν.

To illustrate how the results of Theorem 4.1 and Corollary 4.2 improve classical additive perturbation bounds, we have prepared an simple numerical example run in MATLAB$^{\text{TM}}$. We have generated a 20×10 random matrix A with condition number $\kappa_2(A) = 10^4$, which has been perturbed in two forms: additively $\widetilde{A}_1 := A + \Delta A$ and multiplicatively $\widetilde{A}_2 := (I+E)A(I+F)$ with $\|\Delta A\|_2 = \epsilon\|A\|_2, \|E\|_2 = \|F\|_2 = \epsilon$, and the parameter ϵ has been varied geometrically from 10^{-16} to 10^{-1}. We have also generated a random vector $b \in \mathbb{R}^{20}$, that, for simplicity, has been kept unperturbed. For each value of ϵ we have computed the "exact" minimal length solution x_0 of each Least Squares Problem via the `svd` command of `MATLAB` run in variable precision arithmetic using 50 decimal digits of precision. This has been done for the three LSP:

$$\min_{x\in\mathbb{R}^{10}} \|Ax - b\|_2, \quad \min_{x\in\mathbb{R}^{10}} \|\widetilde{A}_1 x - b\|_2, \quad \text{and} \quad \min_{x\in\mathbb{R}^{10}} \|\widetilde{A}_2 x - b\|_2,$$

and we have called, respectively, the minimal length solutions x_0, x_1 and x_2. In Figure 4.1 we have plotted $\|x_1 - x_0\|_2/\|x_0\|_2$ (continuous line, squares) and $\|x_2 - x_0\|_2/\|x_0\|_2$ (continuous line, circles) against the size of the perturbation ϵ, and we have also represented (dashed line) the bound in the right hand side of (4.9)[3]. It can be seen that for additive perturbations the relative errors increase linearly proportional to $\kappa_2(A)\epsilon$ as predicted in (1.9), however the relative errors for multiplicative perturbations increase linearly as ϵ, independently of the condition number of the matrix $\kappa_2(A)$ as predicted in Theorem 4.1 and Corollary 4.2. Notice also that the factor $\|A^\dagger\|_2\|b\|_2/\|x_0\|_2$ is, in this example, of order 1. It will be argued in Section 4.2 that this is the case most of the times.

Finally, observe that all the results in this section, as well as those in Section 3, are valid for any values of m and n, that is, both if $m \ge n$ or if $m < n$. Thus, they are valid also for multiplicative perturbations of solutions of underdetermined linear systems.

4.1. The condition number under multiplicative perturbations of LSP. In this subsection we prove that $\|A^\dagger\|_2\|b\|_2/\|x_0\|_2$ is essentially, i.e., up to a moderate constant, the condition number under multiplicative perturbations of LSP. After that we will analyze in Subsection 4.2 that condition number, we will argue that it is a

[3] We have omitted, also for simplicity, the results for the residuals that are very similar.

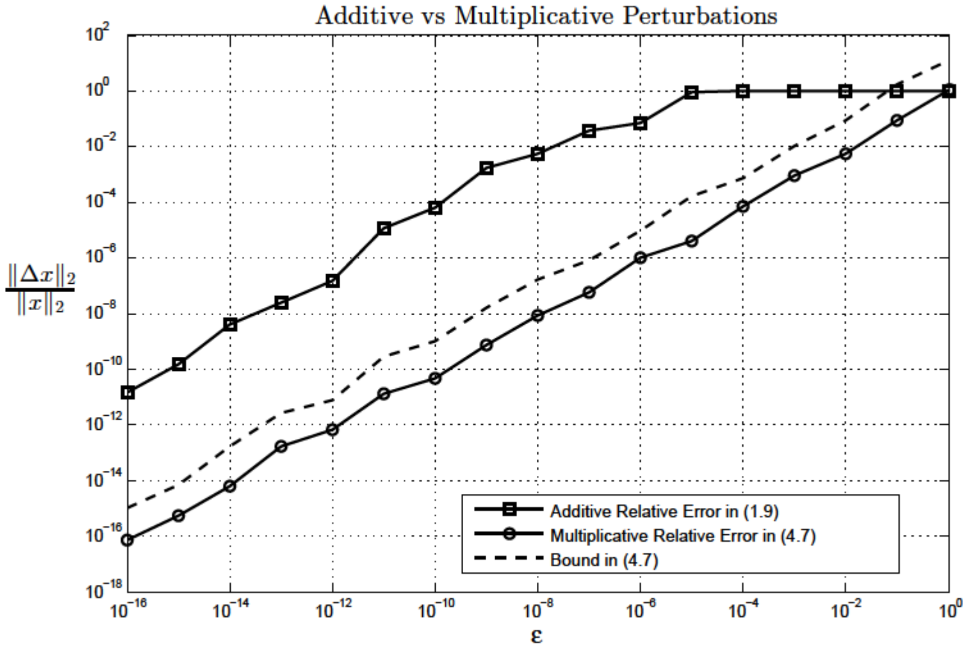

FIG. 4.1. *Additive vs Multiplicative errors in the solution of LSP:* $\min_{x\in\mathbb{R}^{10}} \|Ax-b\|_2, A \in \mathbb{R}^{20\times 10}, \kappa_2(A) = 10^4$ *against the size of the pertubation* ϵ.

moderate number for most vectors b, and we will compare it with the usual[4] condition number for LSP under additive tiny normwise perturbations of A and b, that is,

$$\kappa_{LS}(A,b) := \left(2\,\kappa_2(A) + \frac{\|A^\dagger\|_2\,\|b\|_2}{\|x_0\|_2} + \kappa_2(A)^2\,\frac{\|r\|_2}{\|A\|_2\,\|x_0\|_2} \right). \tag{4.10}$$

First notice that Theorem 4.1 and Corollary 4.2 prove that the sensitivity of the minimum 2-norm solution $x_0 = A^\dagger b$ of a LSP under multiplicative perturbations is governed by $\|A^\dagger\|_2\,\|b\|_2/\|x_0\|_2$. This quantity is well known because it is the condition number for LSP when only the right-hand side b is perturbed (this is easily seem from the second term in the rhs in (1.9)). More precisely, it is easy to prove that if $x_0 = A^\dagger b$, then

$$\frac{\|A^\dagger\|_2\,\|b\|_2}{\|x_0\|_2} = \lim_{\epsilon\to 0}\,\sup \left\{ \frac{\|\widetilde{x}_0 - x_0\|_2}{\epsilon\|x_0\|_2} : \widetilde{x}_0 = A^\dagger\,(b+h),\ \ \|h\|_2 \leq \epsilon\,\|b\|_2 \right\}.$$

Thus Theorem 4.1 essentially proves that multiplicative perturbations have an effect on the minimum 2-norm solution of LSP similar to perturbing only the right-hand side b.

Let us define now the condition number of LSP under multiplicative perturbations and let us determine its magnitude. The reader should notice that, for simplicity, we consider in our definition of the condition number that the left and right mulplicative perturbations, and the relative variation of b, all have the same order.

[4]It is proved in [22, Section 6] that the bound (1.9) is approximately attained to first order in the perturbations.

THEOREM 4.3. *Let us use the same notation and assumptions as in Corollary 4.2 with the parameters μ, ν, and ϵ set equal to η, and let us define the condition number*

$$\kappa_{LS}^{(M)}(A,b) := \lim_{\eta \to 0} \sup \left\{ \frac{\|\widetilde{x}_0 - x_0\|_2}{\eta \, \|x_0\|_2} \; : \; \widetilde{x}_0 = [(I+E)A(I+F)]^{\dagger}\,(b+h), \right. \left. \|E\|_2 \leq \eta, \; \|F\|_2 \leq \eta, \; \|h\|_2 \leq \eta \|b\|_2 \right\} . \tag{4.11}$$

Then

$$\frac{1}{1+2\sqrt{2}}\, \kappa_{LS}^{(M)}(A,b) \leq \frac{\|A^{\dagger}\|_2 \, \|b\|_2}{\|x_0\|_2} \leq \kappa_{LS}^{(M)}(A,b) \, . \tag{4.12}$$

Proof. From (4.9) and $1 \leq \|A^{\dagger}\|_2 \, \|b\|_2 / \|x_0\|_2$, we get

$$\frac{\|\widetilde{x}_0 - x_0\|_2}{\eta \, \|x_0\|_2} \leq \sqrt{2} + \left(1+\sqrt{2}\right) \frac{\|A^{\dagger}\|_2 \|b\|_2}{\|x_0\|_2} + O(\eta) \leq \left(1 + 2\sqrt{2}\right) \frac{\|A^{\dagger}\|_2 \|b\|_2}{\|x_0\|_2} + O(\eta) \, ,$$

which implies the left inequality in (4.12). To prove the right inequality choose a perturbation such that $E = 0$, $F = 0$, and $h = \eta w$, where $\|w\|_2 = \|b\|_2$ and $\|A^{\dagger} w\|_2 = \|A^{\dagger}\|_2 \|w\|_2$. For this perturbation $\|\widetilde{x}_0 - x_0\|_2 / \|x_0\|_2 = \|A^{\dagger} h\|_2 / \|x_0\|_2 = \eta \, \|A^{\dagger}\|_2 \|b\|_2 / \|x_0\|_2$. So, the "sup" appearing in the definition of $\kappa_{LS}^{(M)}(A,b)$ implies $\|A^{\dagger}\|_2 \|b\|_2 / \|x_0\|_2 \leq \kappa_{LS}^{(M)}(A,b)$. ∎

4.2. The factor $\|A^{\dagger}\|_2 \, \|b\|_2 / \|x_0\|_2$ is usually small. In this subsection we are going to show that $\|A^{\dagger}\|_2 \|b\|_2 / \|x_0\|_2$ is usually small compared with $\kappa_{LS}(A,b)$ given in (4.10). In [11, Section 3.2] the same problem was considered for a nonsingular matrix A. The fact that $A \in \mathbb{C}^{m \times n}$ is rectangular forces nontrivial modifications, but we will see that the main conclusions remain the same. First note that for the square nonsingular case [11, Section 3.2] $1 \leq \|A^{\dagger}\|_2 \|b\|_2 / \|x_0\|_2 \leq \kappa_2(A)$. The first inequality also holds in the rectangular case, but, in general, $\|A^{\dagger}\|_2 \|b\|_2 / \|x_0\|_2 \not\leq \kappa_2(A)$. Consider, for example, $A = [1 \; 0; 0 \; 1; 0 \; 0] \in \mathbb{C}^{3 \times 2}$ and $b = [\eta; 0; 1] \in \mathbb{C}^{3 \times 1}$. In this case $x_0 = [\eta; 0] \in \mathbb{C}^{2 \times 1}$ and $\|A^{\dagger}\|_2 \, \|b\|_2 / \|x_0\|_2 = \sqrt{|\eta|^2 + 1} / |\eta|$ which tends to ∞ if $\eta \to 0$ while $\kappa_2(A) = 1$. Nevertheless it is always true that

$$\|A^{\dagger}\|_2 \, \|b\|_2 / \|x_0\|_2 \leq \kappa_{LS}(A,b).$$

The first key point in this section is to show that *if A is fixed*, then $\|A^{\dagger}\|_2 \, \|b\|_2 / \|x_0\|_2$ is a moderate number *for most vectors b*, even if $\kappa_2(A) \gg 1$, and so $\kappa_{LS}(A,b) \gg 1$, which implies that $\|A^{\dagger}\|_2 \, \|b\|_2 / \|x_0\|_2 \ll \kappa_{LS}(A,b)$ for most ill-conditioned LSP whose coefficient matrix is A. However, this is not enough for our purposes, because if $\operatorname{rank}(A) < m$, then for most vectors b the acute angle $\theta(b, \mathcal{R}(A))$ between b and the column space of A is not small, which is equivalent to say that the relative residual $\|Ax_0 - b\|_2 / \|b\|_2 = \sin \theta(b, \mathcal{R}(A))$ is not small. But, very often in practice LSP have small relative residuals, since the problems correspond to inconsistent linear systems $Ax \approx b$ that are close to be consistent. Therefore, the second key point in this section is *if A is fixed* to consider all vectors b such that $\Upsilon = \theta(b, \mathcal{R}(A)) < \pi/2$ *is also fixed*, and then to show that for most *of these vectors b* the factor $\|A^{\dagger}\|_2 \, \|b\|_2 / \|x_0\|_2$ is a moderate number much smaller than $\kappa_{LS}(A,b)$ whenever A is very ill-conditioned.

To explain the properties mentioned above, assume $\operatorname{rank}(A) = r$ and let $A = U \Sigma V^*$ be the SVD of A, where $U \in \mathbb{C}^{m \times r}$ and $V \in \mathbb{C}^{n \times r}$ have orthonormal columns,

$\Sigma = \mathrm{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{C}^{r\times r}$, and $\sigma_1 \geq \cdots \geq \sigma_r > 0$. Observe that $\|x_0\|_2 = \|A^\dagger b\|_2 = \|\Sigma^{-1}U^* b\|_2 \geq |u_r^* b|/\sigma_r$ and

$$\frac{\|A^\dagger\|_2\,\|b\|_2}{\|x_0\|_2} = \frac{\|b\|_2}{\sigma_r\|x_0\|_2} \leq \frac{\|b\|_2}{|u_r^* b|} = \frac{1}{\cos\,\theta(u_r, b)}, \tag{4.13}$$

where u_r is the last column of U and $\theta(u_r, b)$ is the acute angle between u_r and b. Note that the bound on $\|A^\dagger\|_2\,\|b\|_2/\|x_0\|_2$ in (4.13) may be large only if b is "almost" orthogonal to u_r. For example, if A is an extremely ill conditioned fixed matrix (think that $\kappa_2(A) = 10^{1000}$ to be concrete) and b is considered as a random vector whose direction is uniformly distributed in the whole space, then the probability that $0 \leq \theta(u_r, b) \leq \pi/2 - 10^{-6}$ is approximately $1 - 10^{-6}$. Note that the condition $0 \leq \theta(u_r, b) \leq \pi/2 - 10^{-6}$ implies $\|A^\dagger\|_2\,\|b\|_2/\|x_0\|_2 \lesssim 10^6$, which is a moderate number compared to 10^{1000}. In particular, if the perturbation parameters μ, ν, and ϵ in Corollary 4.2 are 10^{-16}, then $\|A^\dagger\|_2\|b\|_2/\|x_0\|_2 \lesssim 10^6$ provides a very good bound for the variation of the minimum 2-norm solution of the LS problem. Even more, it is possible that $\|A^\dagger\|_2\|b\|_2/\|x_0\|_2$ is moderate although $\cos\,\theta(u_r, b) \approx 0$. This can be seen by extending from nonsingular to general matrices the original result by Chan and Foulser in [6, Theorem 1]. We do not present this easy generalization here and refer the reader to the discussion in [11, Section 3.2].

In the argument above, the random vector b may be everywhere in the space. Next, we consider vectors b such that $\Upsilon = \theta(b, \mathcal{R}(A)) < \pi/2$ is kept constant. Let us describe all these vectors as follows: let $y \in \mathbb{C}^r$ be any vector and let $U_\perp \in \mathbb{C}^{m\times(m-r)}$ be such that $[\,U\ U_\perp\,] \in \mathbb{C}^{m\times m}$ is unitary. Then chose any $z \in \mathbb{C}^{m-r}$ such that $\|z\|_2 = \|y\|_2 \tan\Upsilon$, and define $b = Uy + U_\perp z$. It is obvious that $\Upsilon = \theta(b, \mathcal{R}(A))$, because $\mathcal{R}(U) = \mathcal{R}(A)$. In addition, from (4.13), it can be easily proved that these vectors b satisfy

$$\frac{\|A^\dagger\|_2\,\|b\|_2}{\|x_0\|_2} = \frac{\|b\|_2}{\sigma_r\|x_0\|_2} \leq \frac{\|b\|_2}{|u_r^* b|} = \frac{\sqrt{1+\tan^2\Upsilon}}{\cos\,\theta(e_r, y)} = \frac{1}{(\cos\,\Upsilon)\cdot(\cos\,\theta(e_r, y)\,)}, \tag{4.14}$$

where e_r is the rth column of I_r. The bound in (4.14) is a "geometrical" quantity that does not depend on $\kappa_2(A)$ and that, assuming that Υ is not very close to $\pi/2$, is a moderate number for most vectors y, i.e., for most vectors[5] b such that $\Upsilon = \theta(b, \mathcal{R}(A))$.

Finally, we discuss an interesting relationship of the factor $\|A^\dagger\|_2\|b\|_2/\|x_0\|_2$ with the term of $\kappa_{LS}(A, b)$ that depends on $\kappa_2(A)^2$. Note that this term can be upper bounded as follows

$$\Phi := \kappa_2(A)^2\,\frac{\|r\|_2}{\|A\|_2\,\|x_0\|_2} = \kappa_2(A)\,\frac{\|A^\dagger\|_2\,\|r\|_2}{\|x_0\|_2} \leq \kappa_2(A)\,\frac{\|A^\dagger\|_2\,\|b\|_2}{\|x_0\|_2}\,. \tag{4.15}$$

According to our discussion in this subsection $\|A^\dagger\|_2\|b\|_2/\|x_0\|_2$ is a moderate number for most vectors b. Therefore, Φ is upper bounded by a moderate number times $\kappa_2(A)$ for most vectors b and, as a consequence, $\kappa_2(A)^2$ only affects the sensitivity of LSP in very particular situations. In addition, Φ can be written as follows

$$\left(\kappa_2(A)\,\frac{\|A^\dagger\|_2\,\|b\|_2}{\|x_0\|_2}\right)\frac{\|r\|_2}{\|b\|_2} = \Phi\,, \tag{4.16}$$

[5]It seems possible to give a rigorous probabilistic meaning to the loose sentences *"for most vectors b"* that we have used in this section and throughout the paper. A possible strategy would be to consider random vectors whose entries follow uniform distributions in symmetric intervals and to develop results in the spirit of those in [20, Section 3]. This would likely lead to some extra $\sqrt{m}$ factor in the bounds. Due to the length of the paper, we postpone the investigation of these topics.

which implies that for large enough relative residuals (think, for instance, in $\|r\|_2/\|b\|_2 \geq 10^{-3}$) and very ill conditioned matrices A, we have $\|A^\dagger\|_2\|b\|_2/\|x_0\|_2 \ll \Phi \leq \kappa_{LS}(A,b)$, even if $\|A^\dagger\|_2\|b\|_2/\|x_0\|_2$ is large.

4.3. Multiplicative perturbation bounds for other solutions of LSP. Bounds for the variation of solutions different from the minimum 2-norm solution are easily obtained from Theorem 4.1 and (3.7) in Theorem 3.3 and are a minor modification of (4.3). Since the residual of a LS problem is the same for all its solutions, it is not needed to consider again perturbation bounds for the residuals.

THEOREM 4.4. *If $y \in \mathbb{C}^n$ is a solution of the LSP* (4.1)*, then there exists a solution $\widetilde{y} \in \mathbb{C}^n$ of the LSP* (4.2) *such that*

$$\|\widetilde{y} - y\|_2 \leq (\alpha_F + \|F\|_2)\,\|y\|_2 + [\alpha_E\,(1+\alpha_F)\,(1+\epsilon) + \epsilon\,(1+\alpha_F)\,]\ \|A^\dagger\|_2\,\|b\|_2\,,$$

where α_E, α_F, and ϵ are defined as in the statement of Theorem 4.1.

Proof. Given y, there exists a vector $z \in \mathbb{C}^n$ such that $y = x_0 + P_{\mathcal{N}(A)}z$, where x_0 is the minimum 2-norm solution of (4.1). Recall also that $\|y\|_2^2 = \|x_0\|_2^2 + \|P_{\mathcal{N}(A)}z\|_2^2$ and, so, $\|P_{\mathcal{N}(A)}z\|_2 \leq \|y\|_2$. Let us choose the following solution of (4.2), $\widetilde{y} = \widetilde{x}_0 + P_{\mathcal{N}(\widetilde{A})}P_{\mathcal{N}(A)}z$, where $\widetilde{x}_0$ is the minimum 2-norm solution of (4.2). Therefore

$$\|\widetilde{y} - y\|_2 \leq \|\widetilde{x}_0 - x_0\|_2 + \|(P_{\mathcal{N}(\widetilde{A})} - P_{\mathcal{N}(A)})P_{\mathcal{N}(A)}z\|_2 \leq \|\widetilde{x}_0 - x_0\|_2 + \|F\|_2\|y\|_2\,,$$

where we have used (3.7). Now, use (4.3) and $\|x_0\|_2 \leq \|y\|_2$ and get the result. □

Note that the relative variation $\|\widetilde{y}-y\|_2/\|y\|_2$ is governed by $\max\{1, \|A^\dagger\|_2\|b\|_2/\|y\|_2\}$, which is smaller than or equal to $\max\{1, \|A^\dagger\|_2\,\|b\|_2/\|x_0\|_2\}$. Therefore, the minimum 2-norm solution is the most sensitive of the solutions under multiplicative perturbations.

5. Conclusions. We have presented in this paper a thoroughly study of multiplicative perturbations of the Moore-Penrose inverse and the Least Squares Problem, as well as detailed comparisons of the new results with respect to classical additive perturbation results and with respect to other multiplicative perturbation results available in the literature. The multiplicative perturbation bounds for the Moore-Penrose inverse presented here improve previous ones to first order and also in terms of wider applicability. Such improvements are a consequence of the fact that the new bounds are relative to $\|A^\dagger\|_2$ or $\|\widetilde{A}^\dagger\|_2$, instead of relative to $\max\{\|A^\dagger\|_2, \|\widetilde{A}^\dagger\|_2\}$, as are the bounds published so far. On the other hand, the multiplicative perturbation bounds for the Least Squares Problem introduced here are the first ones valid for finite perturbations of arbitrary size, since previously published bounds are asymptotic first order bounds valid only for infinitesimal perturbations. A distinctive feature of the approach followed in this paper is that it is based on the use of adequate orthogonal projectors that allow us to get exact expressions for the perturbed Moore-Penrose inverse which are essential in our developments. A key advantage of this approach is that, under certain conditions, it seems possible to generalize it to linear operators in infinite dimensional spaces, a problem that we plan to study in the future.

REFERENCES

[1] R. Bhatia, *Matrix Analysis*, vol. 169 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1997.

[2] Å. Björck, *Numerical Methods for Least Squares Problems*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996.

[3] L.-X. Cai, W.-W. Xu, and W. Li, *Additive and multiplicative perturbation bounds for the Moore-Penrose inverse*, Linear Algebra Appl., 434 (2011), pp. 480–489.

[4] S. L. Campbell and C. D. Meyer, Jr., *Generalized inverses of linear transformations*, Dover Publications Inc., New York, 1991. Corrected reprint of the 1979 original.

[5] N. Castro-González, J. Ceballos, F. M. Dopico, and J. M. Molera, *Accurate solution of structured least squares problems via rank-revealing decompositions*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1112–1128.

[6] T. F. Chan and D. E. Foulser, *Effectively well-conditioned linear systems*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 963–969.

[7] J. Demmel, *Accurate singular value decompositions of structured matrices*, SIAM J. Matrix Anal. Appl., 21 (1999), pp. 562–580.

[8] J. Demmel, M. Gu, S. Eisenstat, I. Slapničar, K. Veselić, and Z. Drmač, *Computing the singular value decomposition with high relative accuracy*, Linear Algebra Appl., 299 (1999), pp. 21–80.

[9] J. Demmel and K. Veselić, *Jacobi's method is more accurate than QR*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 1204–1246.

[10] F. M. Dopico, P. Koev, and J. M. Molera, *Implicit standard Jacobi gives high relative accuracy*, Numer. Math., 113 (2009), pp. 519–553.

[11] F. M. Dopico and J. M. Molera, *Accurate solution of structured linear systems via rank-revealing decompositions*, IMA J. Numer. Anal., 32 (2012), pp. 1096–1116.

[12] Z. Drmač and K. Veselić, *New fast and accurate Jacobi SVD algorithm. I*, SIAM J. Matrix Anal. Appl., 29 (2008), pp. 1322–1342.

[13] ———, *New fast and accurate Jacobi SVD algorithm. II*, SIAM J. Matrix Anal. Appl., 29 (2008), pp. 1343–1362.

[14] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second ed., 2002.

[15] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[16] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[17] I. C. F. Ipsen, *Relative perturbation results for matrix eigenvalues and singular values*, in Acta numerica, 1998, vol. 7 of Acta Numer., Cambridge Univ. Press, Cambridge, 1998, pp. 151–201.

[18] R.-C. Li, *Relative perturbation theory. I. Eigenvalue and singular value variations*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 956–982.

[19] ———, *Relative perturbation theory. II. Eigenspace and singular subspace variations*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 471–492.

[20] Siegfried M. Rump, *Inversion of extremely ill-conditioned matrices in floating-point*, Japan J. Indust. Appl. Math., 26 (2009), pp. 249–277.

[21] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990.

[22] P.-Å. Wedin, *Perturbation theory for pseudo-inverses*, BIT, 13 (1973), pp. 217–232.