mathematics

MDPI

Article

A Quaternion Matrix Equation Model with Application to Color-Image Watermarking

Zhuo-Heng He [1,2], Jin-Xuan Lin [1], Chen-Yang Qi [2] and Shao-Wen Yu [3,*]

[1] Sino-European School of Technology, Shanghai University, Shanghai 200444, China; zhuohenghe@shu.edu.cn (Z.-H.H.); linjinxuan@shu.edu.cn (J.-X.L.)
[2] Department of Mathematics and Newtouch Center for Mathematics, Shanghai University, Shanghai 200444, China; 2802056837@shu.edu.cn
[3] School of Mathematics, East China University of Science and Technology, Shanghai 200237, China
* Correspondence: yushaowen@ecust.edu.cn

Abstract

We study two coupled two-sided quaternion matrix equations sharing an unknown. Because the map $X \mapsto AXB$ is real-linear but not generally quaternion-linear, an order-preserving real representation converts the problem into one real system. It yields an exact augmented-rank solvability criterion, a reduced shared-unknown criterion, and the complete Moore–Penrose solution. We also obtain fourteen necessary quaternion block-rank screening conditions and counterexamples proving that they are not jointly sufficient. For a side-information-assisted dual-watermarking application at embedding strength 0.05, the stored 8-bit image gives a peak signal-to-noise ratio (PSNR) of 35.99 dB and a global structural-similarity estimate of 0.9987; the no-further-processing normalized correlations are 0.3921 and 0.7179, respectively. The ideal path recovers to floating-point roundoff, whereas the stored-image results quantify sensitivity to clipping, quantization, and subsequent processing.

Keywords: quaternion matrix equation; real representation; rank criterion; coupled system; generalized inverse; color image watermarking

MSC: 15A03; 15A24; 15A09

1. Introduction

Matrix decomposition is a central tool in matrix theory. The singular value decomposition (SVD), for example, is fundamental in data analysis, inverse problems, image processing, and control. Generalized decompositions such as the generalized singular value decomposition (GSVD) and restricted SVD extend the object of decomposition from one matrix to several mutually related matrices [1–4]. This multi-matrix viewpoint is especially relevant when several equations share coefficients or unknowns.

Over a division ring, simultaneous decomposition is more delicate because multiplication is not commutative. Wang et al. [5] studied the equivalence canonical form of a matrix triplet with a common number of columns. A simultaneous decomposition of a matrix quaternity was subsequently obtained in [6]. More recently, equivalence canonical forms for two five-matrix arrays and their applications were presented in [7]. These results provide useful structural information about coefficient arrays, but a decomposition-based consistency proof must still preserve every ordered left–right coefficient pair associated with a shared unknown.

Check for updates

Academic Editor: Xiangmin Jiao

Received: 4 August 2026
Revised: 2 September 2026
Accepted: 3 September 2026
Published: 8 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Sylvester-type matrix equations play an important role in control theory [8]. Their consistency, simultaneous equivalence, and solution methods have been studied from several viewpoints [9–12]. For quaternion coefficients, representative work includes Cramer's rules for Sylvester equations and determinantal representations for systems of two-sided equations [13,14]. Using the product singular value decomposition for multiple quaternion matrices, He et al. [15] derived solvability conditions and general solutions for a general chain of coupled Sylvester-type quaternion matrix equations and applied the framework to color-video encryption. A related system with three coupled equations was applied to color-image encryption in [16].

Quaternion matrix decompositions and equations also arise naturally in color-image processing. The quaternion algebra encodes the three red–green–blue (RGB) channels as the imaginary components of one matrix and thereby retains inter-channel relations [17,18]. Structure-preserving quaternion algorithms have consequently been developed for singular value decomposition, Krylov iteration, factorization, recognition, and transmission [4,19–23]. The discrete wavelet transform (DWT) provides a complementary mechanism for embedding payloads into selected frequency sub-bands [24].

Despite these developments, three issues remain insufficiently resolved for coupled two-sided quaternion equations with a shared unknown. First, separate decompositions of left and right coefficient arrays may lose the ordered pairing in each product AXB. Second, compact quaternion block-rank equalities may provide useful inconsistency screens without characterizing the complete real-linear compatibility of the shared variable. Third, applications often report an ideal algebraic recovery without separating it from the clipping and quantization errors introduced by an actually stored image. The present work addresses these gaps for the system below and states explicitly which conclusions are exact, which are only necessary screens, and which belong to the numerical case study.

We consider the coupled system

$$\begin{cases} A_1X_1A_2 + B_1X_2B_2 = \Omega_1, \\ C_1X_2C_2 + D_1X_3D_2 + E_1X_4E_2 = \Omega_2, \end{cases} \tag{1}$$

where all coefficients and right-hand sides are given quaternion matrices and $X_1, \ldots, X_4$ are unknown. The mathematical difficulty is not merely the number of summands. The same X_2 occurs with two different ordered pairs of left and right factors, and quaternion multiplication is noncommutative. Consequently, concatenating left factors and right factors independently can lose the pairing between $B_1X_2B_2$ and $C_1X_2C_2$.

In the general theory the coefficient matrices in (1) are prescribed model data; no probabilistic construction or invertibility assumption is imposed unless stated. The terms describe ordered two-sided actions: a left coefficient acts on row coordinates, a right coefficient acts on column coordinates, and the two occurrences of X_2 couple the equations. Their application-specific interpretation and construction are given in Section 8.

The contributions of this paper are as follows.

(i) We introduce an explicit order-preserving real representation of every map $X \mapsto AXB$ and thereby obtain an exact augmented-rank criterion for (1), without commutativity or artificial equal-size assumptions.

(ii) We eliminate X_1, X_3, X_4 at the operator level and derive a smaller exact criterion for the shared variable X_2.

(iii) We prove fourteen quaternion block-rank equalities useful as necessary screening tests, characterize when they are jointly sufficient for fixed coefficients, and give checkable full-rank and zero-pattern regimes in which sufficiency follows. All subset

combinations of the summands are included, together with two genuinely coupled conditions. A counterexample proves that these tests are not jointly sufficient.

(iv) We give the complete Moore–Penrose solution in real coordinates and clarify the uniqueness, noninvertible, and error-propagation statements used in a reproducible dual-watermarking application that separates ideal floating-point recovery from 8-bit storage.

The rest of the paper is organized as follows. Section 2 fixes dimensions and notation. Section 3 explains how the coefficient arrays are related to equivalence canonical forms and clarifies the role and limitation of such reductions. Section 4 constructs the order-preserving real representation. Exact and screening rank conditions are proved in Section 5, followed by computational remarks and examples in Sections 6 and 7. Section 8 presents the complete watermarking model and experiments. Section 9 discusses the principal findings, limitations, and future directions, and Section 10 concludes the paper.

2. Problem Formulation and Notation

Let $\mathbb{H} = \{q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} : q_s \in \mathbb{R}\}$, where $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$. Throughout the paper the sizes in (1) are

$$\begin{aligned} &A_1 \in \mathbb{H}^{m_1\times p_1}, \quad B_1 \in \mathbb{H}^{m_1\times p_2}, \quad C_1 \in \mathbb{H}^{m_2\times p_2}, \quad D_1 \in \mathbb{H}^{m_2\times p_3}, \quad E_1 \in \mathbb{H}^{m_2\times p_4},\\ &A_2 \in \mathbb{H}^{q_1\times n_1}, \quad B_2 \in \mathbb{H}^{q_2\times n_1}, \quad C_2 \in \mathbb{H}^{q_2\times n_2}, \quad D_2 \in \mathbb{H}^{q_3\times n_2}, \quad E_2 \in \mathbb{H}^{q_4\times n_2},\\ &X_s \in \mathbb{H}^{p_s\times q_s} \quad (s = 1,2,3,4), \qquad \Omega_1 \in \mathbb{H}^{m_1\times n_1}, \quad \Omega_2 \in \mathbb{H}^{m_2\times n_2}. \end{aligned}$$

No equality between the row or column sizes of Ω_1 and Ω_2 is assumed. For a quaternion matrix, $\mathrm{rank}_{\mathbb{H}}$ denotes its rank over the division ring $\mathbb{H}$. For a real matrix, $\mathrm{rank}_{\mathbb{R}}$ denotes its ordinary real rank. Empty block rows or columns appearing below are omitted in the usual way. The conjugate transpose of M is denoted by $M^* = \overline{M}^{\,T}$; unlike ordinary transpose, it satisfies $(MN)^* = N^*M^*$ over $\mathbb{H}$. For equally sized quaternion matrices, the real Frobenius inner product and its induced norm are

$$\langle X, Y\rangle_{\mathbb{R}} = \mathrm{Re}\,\mathrm{tr}(X^*Y), \qquad \|X\|_F = \sqrt{\langle X, X\rangle_{\mathbb{R}}}.$$

3. Coefficient Arrays and Equivalence Reductions

The coefficient incidence pattern of (1) is described by the two arrays

$$\mathfrak{L} = \begin{pmatrix} A_1 & B_1 & 0 & 0\\ 0 & C_1 & D_1 & E_1 \end{pmatrix}, \qquad \mathfrak{R} = \begin{pmatrix} A_2 & 0\\ B_2 & C_2\\ 0 & D_2\\ 0 & E_2 \end{pmatrix}. \tag{2}$$

The zero blocks in these arrays record which unknown occurs in which equation. Equivalence canonical forms for related five-matrix arrays were developed in [7]; the following result summarizes the reduction used on the right coefficient family.

Theorem 1 ([7], Theorem 2). *Let $A \in \mathbb{H}^{p\times m}$, $B \in \mathbb{H}^{r\times m}$, $C \in \mathbb{H}^{r\times n}$, $D \in \mathbb{H}^{s\times n}$, and $E \in \mathbb{H}^{t\times n}$. There exist invertible matrices $\widetilde{P}_1, \ldots, \widetilde{P}_4, \widetilde{Q}_1, \widetilde{Q}_2$ of compatible sizes such that*

$$\widetilde{P}_1 A\widetilde{Q}_1 = S'_a, \quad \widetilde{P}_2 B\widetilde{Q}_1 = S'_b, \quad \widetilde{P}_2 C\widetilde{Q}_2 = S'_c, \quad \widetilde{P}_3 D\widetilde{Q}_2 = S'_d, \quad \widetilde{P}_4 E\widetilde{Q}_2 = S'_e, \tag{3}$$

where the canonical blocks and their dimensions are determined by ranks of subarrays of A, B, C, D, E.

The block-incidence transpose of this array also has an equivalence reduction. Because $\mathbb{H}$ is noncommutative, it must be derived with conjugate transpose rather than ordinary transpose.

Lemma 1 (Adjoint-dual canonical form)**.** *Let $A \in \mathbb{H}^{m\times p}$, $B \in \mathbb{H}^{m\times r}$, $C \in \mathbb{H}^{n\times r}$, $D \in \mathbb{H}^{n\times s}$, and $E \in \mathbb{H}^{n\times t}$. There exist invertible matrices $P_1, P_2, Q_1, Q_2, Q_3, Q_4$ such that*

$$P_1AQ_1 = S_a, \quad P_1BQ_2 = S_b, \quad P_2CQ_2 = S_c, \quad P_2DQ_3 = S_d, \quad P_2EQ_4 = S_e, \tag{4}$$

where $S_a, \ldots, S_e$ are the conjugate transposes of the corresponding canonical matrices in Theorem 1; explicitly,

$$S_a = (S'_a)^*, \quad S_b = (S'_b)^*, \quad S_c = (S'_c)^*, \quad S_d = (S'_d)^*, \quad S_e = (S'_e)^*.$$

Proof. Apply Theorem 1 to A^*, B^*, C^*, D^*, E^*. For example, $\widetilde{P}_1A^*\widetilde{Q}_1 = S'_a$ and $\widetilde{P}_2B^*\widetilde{Q}_1 = S'_b$. Taking conjugate transposes gives

$$\widetilde{Q}_1^*A\widetilde{P}_1^* = (S'_a)^*, \qquad \widetilde{Q}_1^*B\widetilde{P}_2^* = (S'_b)^*.$$

The remaining three relations follow identically. Setting $P_1 = \widetilde{Q}_1^*$, $P_2 = \widetilde{Q}_2^*$, and $Q_j = \widetilde{P}_j^*$ gives (4). In the explicit form of ([7], Formula (13)), the canonical matrices consist of zero and identity blocks, so conjugate transpose reduces to the corresponding block transpose. □

Theorem 1 is valuable for exposing zeros, identity blocks, and common row or column transformations. In the present coupled problem, however, the left and right families must be transformed coherently. The next elementary proposition states precisely what a valid reduction has to preserve. It is formulated independently of the detailed block sizes in a particular canonical form.

Proposition 1 (Coordinated equivalence reduction)**.** *Suppose there are invertible matrices $P_1, P_2, Q_1, Q_2, Q_3, Q_4$ and $\widetilde{P}_1, \widetilde{P}_2, \widetilde{P}_3, \widetilde{P}_4, \widetilde{Q}_1, \widetilde{Q}_2$ such that*

$$P_1A_1Q_1 = S_a, \quad P_1B_1Q_2 = S_b, \quad P_2C_1Q_2 = S_c, \quad P_2D_1Q_3 = S_d, \quad P_2E_1Q_4 = S_e, \tag{5}$$

$$\widetilde{P}_1A_2\widetilde{Q}_1 = S'_a, \quad \widetilde{P}_2B_2\widetilde{Q}_1 = S'_b, \quad \widetilde{P}_2C_2\widetilde{Q}_2 = S'_c, \quad \widetilde{P}_3D_2\widetilde{Q}_2 = S'_d, \quad \widetilde{P}_4E_2\widetilde{Q}_2 = S'_e. \tag{6}$$

Define

$$\widehat{X}_s = Q_s^{-1}X_s\widetilde{P}_s^{-1} \quad (s = 1, 2, 3, 4), \tag{7}$$

$$\widehat{\Omega}_1 = P_1\Omega_1\widetilde{Q}_1, \qquad \widehat{\Omega}_2 = P_2\Omega_2\widetilde{Q}_2. \tag{8}$$

Then (1) is solvable if and only if the transformed system

$$\begin{cases} S_a\widehat{X}_1S'_a + S_b\widehat{X}_2S'_b = \widehat{\Omega}_1, \\ S_c\widehat{X}_2S'_c + S_d\widehat{X}_3S'_d + S_e\widehat{X}_4S'_e = \widehat{\Omega}_2 \end{cases} \tag{9}$$

is solvable. The correspondence between their solution sets is bijective.

Proof. Premultiplying the first equation of (1) by P_1 and postmultiplying it by $\widetilde{Q}_1$, then inserting (5)–(7), gives the first equation of (9). The second equation follows in the same way. All change-of-basis matrices are invertible, so the substitutions can be reversed uniquely. □

Two points are essential. First, the transformations associated with B_1 and C_1 must share the same Q_2, while those associated with B_2 and C_2 must share the same $\widetilde{P}_2$; otherwise they would define two unrelated transformed versions of X_2. Second, even a correct canonical reduction does not permit the pairs (B_1, B_2) and (C_1, C_2) to be separated. The same $\widehat{X}_2$ must satisfy both ordered products in (9). The real representation in the next section retains this pairing automatically and supplies the definitive consistency test. Canonical forms remain useful for structural interpretation and for reducing zero and identity blocks, but they are not used as a substitute for that test. The operator composition law and a proof that the exact rank criterion is invariant under every coordinated equivalence reduction are given in Appendix A.

4. Order-Preserving Real Representation

For $q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$, define $\rho(q) = (q_0, q_1, q_2, q_3)^T$. Left and right multiplication by q are represented by

$$L(q) = \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix}, \qquad R(q) = \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & q_3 & -q_2 \\ q_2 & -q_3 & q_0 & q_1 \\ q_3 & q_2 & -q_1 & q_0 \end{pmatrix}. \tag{10}$$

Thus $\rho(qx) = L(q)\rho(x)$ and $\rho(xq) = R(q)\rho(x)$.

For $X = (x_{\alpha\beta}) \in \mathbb{H}^{p\times q}$, let $\mathrm{rvec}(X)$ be the real vector obtained by stacking $\rho(x_{\alpha\beta})$ column by column. Given $A = (a_{u\alpha}) \in \mathbb{H}^{m\times p}$ and $B = (b_{\beta v}) \in \mathbb{H}^{q\times n}$, define the real block matrix $\mathcal{T}(A, B) \in \mathbb{R}^{4mn\times 4pq}$ by

$$\big[\mathcal{T}(A, B)\big]_{(u,v),(\alpha,\beta)} = L(a_{u\alpha})R(b_{\beta v}), \tag{11}$$

where both pairs of indices are ordered columnwise.

Lemma 2. *For every conformable A, X, B,*

$$\mathrm{rvec}(AXB) = \mathcal{T}(A, B)\,\mathrm{rvec}(X). \tag{12}$$

Proof. The (u, v) entry of AXB is $\sum_{\alpha=1}^{p}\sum_{\beta=1}^{q} a_{u\alpha}x_{\alpha\beta}b_{\beta v}$. By associativity and (10),

$$\rho(a_{u\alpha}x_{\alpha\beta}b_{\beta v}) = L(a_{u\alpha})R(b_{\beta v})\rho(x_{\alpha\beta}).$$

Stacking these identities in the prescribed order gives (12). □

Set

$$\mathcal{A} = \mathcal{T}(A_1, A_2), \quad \mathcal{B} = \mathcal{T}(B_1, B_2), \quad \mathcal{C} = \mathcal{T}(C_1, C_2), \quad \mathcal{D} = \mathcal{T}(D_1, D_2), \quad \mathcal{E} = \mathcal{T}(E_1, E_2), \tag{13}$$

and define

$$\mathcal{M} = \begin{pmatrix} \mathcal{A} & \mathcal{B} & 0 & 0 \\ 0 & \mathcal{C} & \mathcal{D} & \mathcal{E} \end{pmatrix}. \tag{14}$$

In addition, let

$$x = \mathrm{col}\big(\mathrm{rvec}\,X_1, \mathrm{rvec}\,X_2, \mathrm{rvec}\,X_3, \mathrm{rvec}\,X_4\big), \qquad \omega = \mathrm{col}\big(\mathrm{rvec}\,\Omega_1, \mathrm{rvec}\,\Omega_2\big). \tag{15}$$

Lemma 2 shows that (1) is exactly the real linear system $\mathcal{M}x = \omega$.

For reference, Table 1 records the dimensions of all real operators. Besides making the construction directly implementable, this dimension audit prevents the row and column partitions of the two coupled equations from being conflated.

Table 1. Dimensions of the real coefficient operators.

Operator	Dimension
$\mathcal{A}$	$4m_1n_1 \times 4p_1q_1$
$\mathcal{B}$	$4m_1n_1 \times 4p_2q_2$
$\mathcal{C}$	$4m_2n_2 \times 4p_2q_2$
$\mathcal{D}$	$4m_2n_2 \times 4p_3q_3$
$\mathcal{E}$	$4m_2n_2 \times 4p_4q_4$
$\mathcal{M}$	$4(m_1n_1 + m_2n_2) \times 4\sum_{s=1}^{4} p_sq_s$

5. Solvability and Rank Conditions

5.1. Exact Criteria and the General Solution

Theorem 2 (Exact augmented-rank criterion)**.** *System* (1) *is solvable if and only if*

$$\operatorname{rank}_{\mathbb{R}} \mathcal{M} = \operatorname{rank}_{\mathbb{R}} \begin{pmatrix} \mathcal{M} & \omega \end{pmatrix}. \tag{16}$$

Proof. By Lemma 2, quaternion solutions of (1) are in one-to-one correspondence with real solutions of $\mathcal{M}x = \omega$. The real Rouché–Capelli theorem now gives (16). □

Criterion (16) retains every ordered coefficient pair and therefore does not require commutativity or artificial equal-size assumptions. A smaller exact test can be obtained by eliminating the variables that are not shared.

Let $\mathcal{N} = (\mathcal{D}\ \mathcal{E})$. Choose real matrices $\Lambda_{\mathcal{A}}$ and $\Lambda_{\mathcal{N}}$ satisfying

$$\ker \Lambda_{\mathcal{A}} = \operatorname{Im} \mathcal{A}, \qquad \ker \Lambda_{\mathcal{N}} = \operatorname{Im} \mathcal{N}. \tag{17}$$

For instance, their rows may be chosen as bases of the corresponding left nullspaces. Define

$$\mathcal{G} = \begin{pmatrix} \Lambda_{\mathcal{A}}\mathcal{B} \\ \Lambda_{\mathcal{N}}\mathcal{C} \end{pmatrix}, \qquad g = \begin{pmatrix} \Lambda_{\mathcal{A}} \operatorname{rvec}(\Omega_1) \\ \Lambda_{\mathcal{N}} \operatorname{rvec}(\Omega_2) \end{pmatrix}. \tag{18}$$

When the annihilators are chosen with full row rank, let

$$\ell_{\mathcal{A}} = 4m_1n_1 - \operatorname{rank}_{\mathbb{R}} \mathcal{A}, \qquad \ell_{\mathcal{N}} = 4m_2n_2 - \operatorname{rank}_{\mathbb{R}} \mathcal{N}.$$

Then $\Lambda_{\mathcal{A}} \in \mathbb{R}^{\ell_{\mathcal{A}} \times 4m_1n_1}$, $\Lambda_{\mathcal{N}} \in \mathbb{R}^{\ell_{\mathcal{N}} \times 4m_2n_2}$, and $\mathcal{G} \in \mathbb{R}^{(\ell_{\mathcal{A}}+\ell_{\mathcal{N}}) \times 4p_2q_2}$.

Theorem 3 (Reduced shared-variable criterion)**.** *System* (1) *is solvable if and only if*

$$\operatorname{rank}_{\mathbb{R}} \mathcal{G} = \operatorname{rank}_{\mathbb{R}} \begin{pmatrix} \mathcal{G} & g \end{pmatrix}. \tag{19}$$

Proof. For a fixed real vector $x_2 = \operatorname{rvec}(X_2)$, the first equation is solvable in x_1 precisely when $\operatorname{rvec}(\Omega_1) - \mathcal{B}x_2 \in \operatorname{Im} \mathcal{A}$. By (17), this is equivalent to $\Lambda_{\mathcal{A}}\mathcal{B}x_2 = \Lambda_{\mathcal{A}} \operatorname{rvec}(\Omega_1)$. Likewise, the second equation is solvable in (x_3, x_4) precisely when $\Lambda_{\mathcal{N}}\mathcal{C}x_2 = \Lambda_{\mathcal{N}} \operatorname{rvec}(\Omega_2)$. The two requirements form $\mathcal{G}x_2 = g$, whose consistency is equivalent to (19). □

The annihilator test also has a basis-free projector form. Define the orthogonal projectors onto the left nullspaces of $\mathcal{A}$ and $\mathcal{N}$ by

$$P_{\mathcal{A}}^{\perp} = I_{4m_1n_1} - \mathcal{A}\mathcal{A}^{\dagger}, \qquad P_{\mathcal{N}}^{\perp} = I_{4m_2n_2} - \mathcal{N}\mathcal{N}^{\dagger}, \tag{20}$$

and set

$$\mathcal{H} = \begin{pmatrix} P_{\mathcal{A}}^{\perp}\mathcal{B} \\ P_{\mathcal{N}}^{\perp}\mathcal{C} \end{pmatrix}, \qquad h = \begin{pmatrix} P_{\mathcal{A}}^{\perp}\,\mathrm{rvec}(\Omega_1) \\ P_{\mathcal{N}}^{\perp}\,\mathrm{rvec}(\Omega_2) \end{pmatrix}. \tag{21}$$

Corollary 1 (Projector form of the reduced criterion)**.** *System* (1) *is solvable if and only if*

$$\operatorname{rank}_{\mathbb{R}} \mathcal{H} = \operatorname{rank}_{\mathbb{R}} \begin{pmatrix} \mathcal{H} & h \end{pmatrix}. \tag{22}$$

This condition is equivalent to Theorem 3 and does not depend on a particular choice of left-nullspace bases.

Proof. For every real matrix K, the matrix $KK^{\dagger}$ is the orthogonal projector onto $\operatorname{Im} K$. Hence $y \in \operatorname{Im} K$ if and only if $(I - KK^{\dagger})y = 0$. Applying this observation first to $K = \mathcal{A}$ and then to $K = \mathcal{N}$ shows that, for a fixed x_2, the two remaining systems are consistent exactly when $\mathcal{H}x_2 = h$. The real Rouché–Capelli theorem gives (22). Since both this system and $\mathcal{G}x_2 = g$ describe the same admissible set of shared variables, their consistency criteria are equivalent. □

Corollary 2 (Sequential parametrization)**.** *Assume that* (19) *holds, and write*

$$\omega_1 = \mathrm{rvec}(\Omega_1), \qquad \omega_2 = \mathrm{rvec}(\Omega_2), \qquad u = \mathrm{col}\big(\mathrm{rvec}(X_3), \mathrm{rvec}(X_4)\big).$$

Then every solution is obtained from

$$x_2 = \mathcal{G}^{\dagger}g + \big(I - \mathcal{G}^{\dagger}\mathcal{G}\big)y_2, \tag{23}$$

$$x_1 = \mathcal{A}^{\dagger}(\omega_1 - \mathcal{B}x_2) + \big(I - \mathcal{A}^{\dagger}\mathcal{A}\big)y_1, \tag{24}$$

$$u = \mathcal{N}^{\dagger}(\omega_2 - \mathcal{C}x_2) + \big(I - \mathcal{N}^{\dagger}\mathcal{N}\big)y_{34}, \tag{25}$$

where y_2, y_1, y_{34} are arbitrary real vectors of the corresponding dimensions. The solution quadruple is unique if and only if

$$\mathcal{G}, \qquad \mathcal{A}, \qquad \mathcal{N} \quad \textit{all have full column rank.} \tag{26}$$

Proof. Theorem 3 reduces compatibility to $\mathcal{G}x_2 = g$, whose complete solution is (23). For every such x_2, the two defining annihilator relations ensure that $\omega_1 - \mathcal{B}x_2 \in \operatorname{Im}\mathcal{A}$ and $\omega_2 - \mathcal{C}x_2 \in \operatorname{Im}\mathcal{N}$. The standard Moore–Penrose formula therefore gives (24) and (25), and every solution of the original system arises in this way.

If the three matrices in (26) have full column rank, the three nullspace terms vanish and the sequential solution is unique. Conversely, a nonzero vector in $\ker\mathcal{A}$ or $\ker\mathcal{N}$ changes x_1 or u without changing the right-hand sides. If $0 \neq v \in \ker\mathcal{G}$, then $\mathcal{B}v \in \operatorname{Im}\mathcal{A}$ and $\mathcal{C}v \in \operatorname{Im}\mathcal{N}$. Hence x_2 can be replaced by $x_2 + v$ while compensating with suitable changes in x_1 and u, again producing a distinct solution. Thus uniqueness implies all three full-column-rank conditions. □

The projector criterion provides a compatibility test independent of the chosen left-nullspace basis, whereas Corollary 2 gives a constructive blockwise parametrization. For compactness, the same complete solution set can also be written directly in terms of the full operator $\mathcal{M}$, as follows.

Corollary 3 (Complete solution). *If (16) holds, all solutions in real coordinates are*

$$x = \mathcal{M}^{\dagger}\omega + \big(I - \mathcal{M}^{\dagger}\mathcal{M}\big)z, \qquad z \in \mathbb{R}^{4\sum_{s=1}^{4} p_s q_s}, \tag{27}$$

where $\mathcal{M}^{\dagger}$ is the real Moore–Penrose inverse. The quaternion matrices X_s are obtained by applying rvec^{-1} *to the four blocks of x. The solution is unique if and only if $\mathcal{M}$ has full column rank. When the system is consistent, its real affine solution dimension is*

$$4\sum_{s=1}^{4} p_s q_s - \operatorname{rank}_{\mathbb{R}} \mathcal{M}.$$

Proof. Equation (27) is the standard complete solution of a consistent real linear system. Full column rank is equivalent to $\ker \mathcal{M} = \{0\}$. □

5.2. Fourteen Quaternion Block-Rank Screening Conditions

The exact test (16) may be large. The following block-rank equalities can be computationally cheaper necessary tests and are also useful for diagnosing which part of a system fails. In every display below the rank is taken over $\mathbb{H}$.

Proposition 2. *If system (1) is solvable, then all of the following equalities hold:*

$$\operatorname{rank}_{\mathbb{H}}(A_1\ B_1\ \Omega_1) = \operatorname{rank}_{\mathbb{H}}(A_1\ B_1), \tag{28}$$

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} A_2 \\ B_2 \\ \Omega_1 \end{pmatrix} = \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} A_2 \\ B_2 \end{pmatrix}, \tag{29}$$

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & A_2 \\ B_1 & \Omega_1 \end{pmatrix} = \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & A_2 \\ B_1 & 0 \end{pmatrix}, \tag{30}$$

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & B_2 \\ A_1 & \Omega_1 \end{pmatrix} = \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & B_2 \\ A_1 & 0 \end{pmatrix}. \tag{31}$$

For the three summands in the second equation, all eight subset conditions hold:

$$\operatorname{rank}_{\mathbb{H}}(C_1\ D_1\ E_1\ \Omega_2) = \operatorname{rank}_{\mathbb{H}}(C_1\ D_1\ E_1), \tag{32}$$

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} C_2 \\ D_2 \\ E_2 \\ \Omega_2 \end{pmatrix} = \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} C_2 \\ D_2 \\ E_2 \end{pmatrix}, \tag{33}$$

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & 0 & C_2 \\ D_1 & E_1 & \Omega_2 \end{pmatrix} = \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & 0 & C_2 \\ D_1 & E_1 & 0 \end{pmatrix}, \tag{34}$$

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & D_2 \\ 0 & E_2 \\ C_1 & \Omega_2 \end{pmatrix} = \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & D_2 \\ 0 & E_2 \\ C_1 & 0 \end{pmatrix}, \tag{35}$$

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & 0 & D_2 \\ C_1 & E_1 & \Omega_2 \end{pmatrix} = \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & 0 & D_2 \\ C_1 & E_1 & 0 \end{pmatrix}, \tag{36}$$

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & 0 & E_2 \\ C_1 & D_1 & \Omega_2 \end{pmatrix} = \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & 0 & E_2 \\ C_1 & D_1 & 0 \end{pmatrix}, \tag{37}$$

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & C_2 \\ 0 & D_2 \\ E_1 & \Omega_2 \end{pmatrix} = \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & C_2 \\ 0 & D_2 \\ E_1 & 0 \end{pmatrix}, \tag{38}$$

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & C_2 \\ 0 & E_2 \\ D_1 & \Omega_2 \end{pmatrix} = \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & C_2 \\ 0 & E_2 \\ D_1 & 0 \end{pmatrix}. \tag{39}$$

Two additional conditions are obtained by coupled eliminations using the same X_2 and retain some, but not all, cross-equation information:

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & 0 & 0 & 0 & B_2 & C_2 \\ A_1 & B_1 & 0 & 0 & -\Omega_1 & 0 \\ 0 & C_1 & D_1 & E_1 & 0 & \Omega_2 \end{pmatrix} = \operatorname{rank}_{\mathbb{H}}(B_2\ C_2) + \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} A_1 & B_1 & 0 & 0 \\ 0 & C_1 & D_1 & E_1 \end{pmatrix}, \tag{40}$$

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & A_2 & 0 \\ 0 & B_2 & C_2 \\ 0 & 0 & D_2 \\ 0 & 0 & E_2 \\ B_1 & -\Omega_1 & 0 \\ C_1 & 0 & \Omega_2 \end{pmatrix} = \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} A_2 & 0 \\ B_2 & C_2 \\ 0 & D_2 \\ 0 & E_2 \end{pmatrix} + \operatorname{rank}_{\mathbb{H}}\begin{pmatrix} B_1 \\ C_1 \end{pmatrix}. \tag{41}$$

Proof. We first record a general elimination identity. Suppose $T = \sum_{s=1}^{t} L_s Y_s R_s$. For a subset $S \subseteq \{1, \dots, t\}$, put $R_S = \operatorname{col}(R_s : s \in S)$ and $L_{S^c} = (L_s : s \notin S)$. Quaternion elementary block row and column operations give

$$\operatorname{rank}_{\mathbb{H}}\begin{pmatrix} 0 & R_S \\ L_{S^c} & T \end{pmatrix} = \operatorname{rank}_{\mathbb{H}} R_S + \operatorname{rank}_{\mathbb{H}} L_{S^c}. \tag{42}$$

Indeed, for $s \in S$ apply the lower block-row operation

$$\text{lower block row} \ \leftarrow\ \text{lower block row} - (L_s Y_s)(s\text{th upper block row}),$$

whereas for $s \notin S$ apply the last block-column operation

$$\text{last block column} \ \leftarrow\ \text{last block column} - (s\text{th left block column})(Y_s R_s).$$

The order of the factors is unchanged in both operations. Together they change the lower-right block from T to zero. A block permutation then gives the rank sum in (42).

Applying (42) to the two summands in the first equation gives (28)–(31). Applying it to the three summands in the second equation for all 2^3 subsets gives (32)–(39).

It remains to verify the coupled tests. Substitute a solution into the matrix on the left of (40). If its block columns are denoted by $\mathsf{C}_1, \dots, \mathsf{C}_6$ and its block rows by $\mathsf{R}_1, \mathsf{R}_2, \mathsf{R}_3$, perform the invertible shear operations

$$\begin{aligned} &\mathsf{C}_5 \leftarrow \mathsf{C}_5 + \mathsf{C}_1 X_1 A_2 + \mathsf{C}_2 X_2 B_2, \\ &\mathsf{C}_6 \leftarrow \mathsf{C}_6 - \mathsf{C}_2 X_2 C_2 - \mathsf{C}_3 X_3 D_2 - \mathsf{C}_4 X_4 E_2, \\ &\mathsf{R}_2 \leftarrow \mathsf{R}_2 + B_1 X_2 \mathsf{R}_1, \qquad \mathsf{C}_5 \leftarrow \mathsf{C}_5 - \mathsf{C}_2 X_2 B_2. \end{aligned}$$

They transform the matrix to

$$\begin{pmatrix} 0 & 0 & 0 & 0 & B_2 & C_2 \\ A_1 & B_1 & 0 & 0 & 0 & 0 \\ 0 & C_1 & D_1 & E_1 & 0 & 0 \end{pmatrix},$$

whose rank is the right-hand side of (40).

For (41), denote the six block rows by $\mathsf{R}_1, \ldots, \mathsf{R}_6$ and the three block columns by $\mathsf{C}_1, \mathsf{C}_2, \mathsf{C}_3$. The operations

$$\begin{aligned} \mathsf{R}_5 &\leftarrow \mathsf{R}_5 + A_1 X_1 \mathsf{R}_1 + B_1 X_2 \mathsf{R}_2, \\ \mathsf{R}_6 &\leftarrow \mathsf{R}_6 - C_1 X_2 \mathsf{R}_2 - D_1 X_3 \mathsf{R}_3 - E_1 X_4 \mathsf{R}_4, \\ \mathsf{C}_2 &\leftarrow \mathsf{C}_2 + \mathsf{C}_1 X_2 B_2, \qquad \mathsf{R}_5 \leftarrow \mathsf{R}_5 - B_1 X_2 \mathsf{R}_2 \end{aligned}$$

annihilate the two right-hand-side blocks in the last two block rows. After a block permutation, the remaining matrix is

$$\begin{pmatrix} 0 & A_2 & 0 \\ 0 & B_2 & C_2 \\ 0 & 0 & D_2 \\ 0 & 0 & E_2 \\ B_1 & 0 & 0 \\ C_1 & 0 & 0 \end{pmatrix},$$

which is permutation-equivalent to a block diagonal matrix with the two blocks on the right of (41). This proves the result. □

Remark 1. *Conditions* (28)–(41) *are deliberately stated only as necessary conditions. They are not a substitute for* (16)*: quaternion block-rank tests can remove individual summands and some coupled combinations, but they need not preserve the complete real-linear compatibility imposed on the shared* X_2*.*

The following statement characterizes precisely when the screening conditions become sufficient and records two readily verifiable coefficient regimes. For fixed coefficient matrices, let $\mathcal{S}_{14}$ denote the set of pairs (Ω_1, Ω_2) satisfying (28)–(41).

Proposition 3 (Regimes of joint sufficiency)**.** *For fixed coefficients, conditions* (28)–(41) *are jointly sufficient if and only if*

$$h(\Omega_1, \Omega_2) \in \operatorname{Im} \mathcal{H} \quad \text{for every } (\Omega_1, \Omega_2) \in \mathcal{S}_{14}, \tag{43}$$

where $\mathcal{H}$ *and* h *are defined in* (21)*. In particular, joint sufficiency holds in either of the following checkable regimes:*

(a) $\mathcal{G}$ *in* (18)*, constructed from full-row-rank annihilators, has full row rank; in this case every prescribed right-hand-side pair is compatible.*

(b) $B_1 = B_2 = C_1 = C_2 = D_1 = D_2 = 0$*. In this zero-pattern regime the system consists of the independent equations* $A_1 X_1 A_2 = \Omega_1$ *and* $E_1 X_4 E_2 = \Omega_2$*.*

The automatic-consistency condition in Proposition 4 is another sufficient regime, because it makes both blocks of h *identically zero.*

Proof. By Corollary 1, a right-hand-side pair is compatible exactly when $h \in \operatorname{Im} \mathcal{H}$. Restricting this equivalence to $\mathcal{S}_{14}$ proves (43). If $\mathcal{G}$ has full row rank, its image is its entire codomain, so $\mathcal{G}x_2 = g$ is consistent for every g. The equivalence in Theorem 3 then gives the stated conclusion.

In regime (b), the two equations are independent. The equation $AXB = \Omega$ is solvable if and only if $\operatorname{col}(\Omega) \subseteq \operatorname{col}(A)$ and $\operatorname{row}(\Omega) \subseteq \operatorname{row}(B)$; equivalently, $\operatorname{rank}_{\mathbb{H}}(A\ \Omega) = \operatorname{rank}_{\mathbb{H}} A$ and $\operatorname{rank}_{\mathbb{H}} \left(\begin{smallmatrix} B \\ \Omega \end{smallmatrix}\right) = \operatorname{rank}_{\mathbb{H}} B$. Conditions (28), (29), (32), and (33) reduce exactly to these four tests for the two independent equations. Hence all fourteen conditions imply solvability in

this regime. Finally, if $\mathcal{A}$ and $\mathcal{N}$ have full row rank, the projectors in (20) vanish, so $h = 0$ and compatibility is automatic. □

6. Computation and Special Cases

6.1. A Rank-Revealing Computational Procedure

The real formulation gives a direct algorithm that applies to arbitrary rectangular coefficients and does not require constructing a canonical block partition. A practical implementation consists of the following steps.

Step 1. Check the dimensions in Section 2 and assemble $\mathcal{A}, \ldots, \mathcal{E}$ from the 4×4 blocks in (11).

Step 2. Form $\mathcal{M}$ and ω in (14)–(15). Optionally test (28)–(41) first; failure of any one of them certifies inconsistency.

Step 3. Compute a rank-revealing SVD or an orthogonal–triangular (QR) factorization of $\mathcal{M}$ and $(\mathcal{M}\ \omega)$ using the same numerical tolerance. Equality of the two numerical ranks is the consistency decision at the stated tolerance.

Step 4. If the system is consistent, evaluate (27). If only one solution is needed, use $x_0 = \mathcal{M}^{\dagger}\omega$ and recover the quaternion entries by groups of four real coordinates.

For a real matrix $K \in \mathbb{R}^{a\times b}$ with singular values $\sigma_1 \geq \cdots \geq \sigma_{\min(a,b)}$, a common numerical convention is

$$\operatorname{rank}_\tau(K) = \#\{j : \sigma_j(K) > \tau\sigma_1(K)\}, \qquad \tau = \max(a,b)\varepsilon_{\text{mach}}. \tag{44}$$

When data are measured or noisy, τ should instead reflect the estimated noise level. Reporting the tolerance is essential because exact algebraic rank is discontinuous under perturbations.

The reduced test in Theorem 3 can be substantially smaller when p_2q_2 is small relative to the total number of unknown entries. Numerically, $\Lambda_{\mathcal{A}}$ and $\Lambda_{\mathcal{N}}$ may be obtained from the left singular vectors corresponding to zero singular values of $\mathcal{A}$ and $\mathcal{N} = (\mathcal{D}\ \mathcal{E})$. More explicitly, if the columns of $U_{\mathcal{A},0}$ and $U_{\mathcal{N},0}$ are orthonormal bases of the corresponding left nullspaces, one may take

$$\Lambda_{\mathcal{A}} = U_{\mathcal{A},0}^T, \qquad \Lambda_{\mathcal{N}} = U_{\mathcal{N},0}^T.$$

Different orthonormal bases of the same left nullspaces yield row-equivalent reduced systems and hence the same consistency decision.

Let

$$r = 4(m_1n_1 + m_2n_2), \qquad c = 4\sum_{s=1}^{4} p_sq_s.$$

A naive dense realization of $\mathcal{M}$ stores rc real numbers, and a dense SVD has arithmetic cost $O(rc\min\{r,c\})$. In the square $N = 64$ four-unknown image configuration, $r = 32{,}768$ and $c = 65{,}536$; the dense coefficient array would contain 2^{31} double-precision entries, approximately 16 GiB. This storage is unnecessary. The action of $\mathcal{M}$ can be evaluated by forming $A_1X_1A_2$, $B_1X_2B_2$, $C_1X_2C_2$, $D_1X_3D_2$, and $E_1X_4E_2$, followed by addition and rvec. Its transpose action follows from the real Frobenius-adjoint identity

$$\mathcal{T}(A,B)^T \operatorname{rvec}(Y) = \operatorname{rvec}(A^*YB^*). \tag{45}$$

Indeed, $\langle X, Y\rangle_{\mathbb{R}} = \operatorname{rvec}(X)^T \operatorname{rvec}(Y)$ and $\operatorname{Re}\operatorname{tr}(UV) = \operatorname{Re}\operatorname{tr}(VU)$ for conformable quaternion matrices. Therefore

$$\langle AXB, Y\rangle_{\mathbb{R}} = \operatorname{Re}\operatorname{tr}(B^*X^*A^*Y) = \operatorname{Re}\operatorname{tr}(X^*A^*YB^*) = \langle X, A^*YB^*\rangle_{\mathbb{R}},$$

which proves (45). Thus iterative least-squares solvers such as LSQR or LSMR can be implemented without forming $\mathcal{M}$. Such iterations are appropriate for approximate recovery; an exact or tolerance-dependent rank decision still requires a rank-revealing factorization or additional exploitable structure. In the diagonal watermarking specialization below, the operator decomposes entrywise and no dense real coefficient matrix is formed.

6.2. Least-Squares Recovery for Inconsistent Data

In applications, the observed right-hand sides may be contaminated by quantization, channel noise, or an imperfect inverse transform. Even when (16) fails exactly, the real formulation still gives a well-defined approximation. All least-squares solutions are

$$x_{\mathrm{LS}} = \mathcal{M}^{\dagger}\omega + (I - \mathcal{M}^{\dagger}\mathcal{M})z, \qquad z \text{ arbitrary}, \tag{46}$$

and the minimum attainable residual is

$$\min_{x} \|\mathcal{M}x - \omega\|_2 = \|(I - \mathcal{M}\mathcal{M}^{\dagger})\omega\|_2. \tag{47}$$

The vector $\mathcal{M}^{\dagger}\omega$ is the unique minimum-norm least-squares solution. Equations (46) and (47) distinguish exact consistency from approximate reconstruction and provide a natural residual measure for noisy payloads.

6.3. Coefficient Regimes with Automatic Consistency

The exact criterion also identifies simple cases in which no compatibility restriction is needed.

Proposition 4. *Let $\mathcal{N} = (\mathcal{D}\ \mathcal{E})$. If $\mathcal{A}$ and $\mathcal{N}$ both have full row rank, then system (1) is solvable for every pair (Ω_1, Ω_2) of the prescribed sizes.*

Proof. Set $x_2 = 0$. Full row rank of $\mathcal{A}$ makes $\mathcal{A}x_1 = \mathrm{rvec}(\Omega_1)$ solvable. Full row rank of $\mathcal{N}$ makes $\mathcal{N}\,\mathrm{col}(x_3, x_4) = \mathrm{rvec}(\Omega_2)$ solvable independently. The corresponding quaternion matrices solve (1). □

If only one of these operators has full row rank, the corresponding block of Theorem 3 disappears, leaving compatibility constraints solely from the other equation. At the opposite extreme, if $\mathcal{A} = 0$ and $\mathcal{N} = 0$, consistency reduces entirely to the requirement that the two affine conditions $\mathcal{B}x_2 = \mathrm{rvec}(\Omega_1)$ and $\mathcal{C}x_2 = \mathrm{rvec}(\Omega_2)$ admit the same x_2. Example 2 is an explicit 3×3 instance in which the two individual equations are consistent but these requirements conflict.

7. Numerical Examples

Example 1 (A consistent third-order system)**.** *For system (1), let all coefficient and unknown matrices be of order three. Take*

$$A_1 = \begin{pmatrix} \mathbf{i} & 0 & 0 \\ 0 & \mathbf{j} & 0 \\ 0 & 0 & \mathbf{k} \end{pmatrix}, \qquad B_1 = \begin{pmatrix} 1 & \mathbf{i} & 0 \\ \mathbf{j} & 0 & \mathbf{k} \\ 0 & 1 & -\mathbf{i} \end{pmatrix}, \qquad C_1 = \begin{pmatrix} 2\mathbf{j} & -\mathbf{k} & \mathbf{k} \\ 0 & \mathbf{i} & \mathbf{i} \\ -\mathbf{k} & 0 & \mathbf{j} \end{pmatrix},$$

$$D_1 = \begin{pmatrix} -\mathbf{j} & 0 & 0 \\ 0 & \mathbf{k} & 0 \\ 0 & 0 & -\mathbf{k} \end{pmatrix}, \qquad E_1 = \begin{pmatrix} 0 & 0 & \mathbf{i} \\ 0 & \mathbf{i} & 0 \\ \mathbf{j} & 0 & \mathbf{k} \end{pmatrix}, \qquad A_2 = \begin{pmatrix} \mathbf{j} & 0 & 0 \\ 0 & \mathbf{k} & 0 \\ 0 & 0 & \mathbf{i} \end{pmatrix},$$

$$B_2 = \begin{pmatrix} 0 & 1 & 0 \\ \mathbf{k} & -\mathbf{i} & \mathbf{j} \\ \mathbf{i} & 0 & \mathbf{k} \end{pmatrix}, \qquad C_2 = \begin{pmatrix} \mathbf{i} & -\mathbf{j} & 0 \\ 1 & \mathbf{k} & \mathbf{i} \\ 0 & \mathbf{j} & \mathbf{j} \end{pmatrix}, \qquad D_2 = \begin{pmatrix} \mathbf{i} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mathbf{k} \end{pmatrix},$$
$$E_2 = \begin{pmatrix} \mathbf{k} & 0 & 0 \\ 0 & \mathbf{j} & 0 \\ 0 & 0 & \mathbf{i} \end{pmatrix}.$$

Choose the four unknowns as

$$X_1 = \begin{pmatrix} 2\mathbf{j} & 0 & 0 \\ 0 & \mathbf{i} & 0 \\ 0 & 0 & -\mathbf{k} \end{pmatrix}, \qquad X_2 = \begin{pmatrix} \mathbf{i} & -\mathbf{k} & 0 \\ 0 & \mathbf{j} & 1 \\ \mathbf{k} & 0 & -\mathbf{i} \end{pmatrix},$$
$$X_3 = \begin{pmatrix} -2+2\mathbf{i} & 0 & \mathbf{j} \\ 0 & \mathbf{k} & 0 \\ 1 & 0 & -\mathbf{i} \end{pmatrix}, \qquad X_4 = \begin{pmatrix} 0 & \mathbf{i} & 0 \\ \mathbf{j} & 0 & \mathbf{k} \\ 0 & 1 & 0 \end{pmatrix}.$$

Direct quaternion multiplication gives

$$\Omega_1 = A_1X_1A_2 + B_1X_2B_2 = \begin{pmatrix} -1-2\mathbf{i} & \mathbf{i} & -j \\ \mathbf{j}+\mathbf{k} & -1-\mathbf{k} & -\mathbf{i}-k \\ \mathbf{i} & \mathbf{j}+\mathbf{k} & -1+\mathbf{i} \end{pmatrix},$$
$$\Omega_2 = C_1X_2C_2 + D_1X_3D_2 + E_1X_4E_2$$
$$= \begin{pmatrix} -2\mathbf{i}-2\mathbf{k} & 1-\mathbf{i}+2\mathbf{j}+\mathbf{k} & 2+\mathbf{i}+k \\ -1+2\mathbf{k} & -3+\mathbf{j}+\mathbf{k} & 2\mathbf{j}+2k \\ -2-\mathbf{j}+\mathbf{k} & -1-\mathbf{i}-2\mathbf{k} & -\mathbf{i} \end{pmatrix}.$$

Thus the displayed X_1, X_2, X_3, X_4 form a particular solution. The corresponding real operator has 72 rows, and row reduction gives

$$\operatorname{rank}_{\mathbb{R}} \mathcal{M} = \operatorname{rank}_{\mathbb{R}}(\mathcal{M}\,\omega) = 72.$$

Hence the exact criterion (16) independently confirms consistency.

Example 2 (An inconsistent third-order system)**.** *Let*

$$A_1 = A_2 = D_1 = E_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$B_1 = \begin{pmatrix} 1 & \mathbf{i} & \mathbf{j} \\ 0 & \mathbf{j} & \mathbf{k} \\ 0 & 0 & 1 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} \mathbf{i} & 1 & 0 \\ 0 & 1 & \mathbf{k} \\ 0 & 0 & \mathbf{j} \end{pmatrix},$$
$$C_1 = \begin{pmatrix} \mathbf{j} & 1 & \mathbf{i} \\ 0 & \mathbf{k} & \mathbf{j} \\ 0 & 0 & 1 \end{pmatrix}, \qquad C_2 = \begin{pmatrix} \mathbf{k} & \mathbf{i} & 1 \\ 0 & \mathbf{j} & \mathbf{k} \\ 0 & 0 & \mathbf{i} \end{pmatrix},$$
$$D_2 = \begin{pmatrix} 1 & \mathbf{j} & 0 \\ 0 & \mathbf{i} & \mathbf{k} \\ 0 & 0 & \mathbf{k} \end{pmatrix}, \qquad E_1 = \begin{pmatrix} \mathbf{i} & 1 & 0 \\ 0 & \mathbf{k} & \mathbf{j} \\ 0 & 0 & 1 \end{pmatrix},$$
$$\Omega_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \Omega_2 = \begin{pmatrix} \mathbf{i}+\mathbf{k} & -1+\mathbf{i}-2\mathbf{j}+\mathbf{k} & \mathbf{i}-2k \\ -1+\mathbf{j}-\mathbf{k} & -1+\mathbf{j}-\mathbf{k} & \mathbf{j}+k \\ 1+\mathbf{i} & 0 & -2\mathbf{k} \end{pmatrix}.$$

The six nonzero coefficient matrices displayed above are nonsingular. The first equation therefore requires $X_2 = 0$. On the other hand, direct multiplication verifies that

$$\Omega_2 = C_1 Y C_2, \qquad Y = \begin{pmatrix} \mathbf{i} & \mathbf{j} & k \\ 1 & \mathbf{i}+\mathbf{j} & -k \\ \mathbf{j}-\mathbf{k} & 1+\mathbf{i} & 2\mathbf{j} \end{pmatrix} \neq 0.$$

Since $D_1 = 0$ and $E_2 = 0$, the second equation requires $X_2 = Y$, which contradicts the first equation. Equivalently,

$$\operatorname{rank}_{\mathbb{R}} \mathcal{M} = 36 < 37 = \operatorname{rank}_{\mathbb{R}} (\mathcal{M}\ \omega).$$

Each equation separately is solvable, so (28)–(39) hold. Direct quaternion row reduction gives rank $9 = 3 + 6$ on both sides of (40) and rank $9 = 6 + 3$ on both sides of (41). Thus all fourteen screening conditions hold, whereas the exact condition (16) fails.

Example 3 (Why independent aggregate ranks fail)**.** *Let*

$$A_1 = A_2 = D_1 = D_2 = E_1 = E_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$B_1 = C_1 = \begin{pmatrix} 1 & \mathbf{i} & \mathbf{j} \\ 0 & \mathbf{j} & \mathbf{k} \\ 0 & 0 & 1 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} \mathbf{i} & 1 & 0 \\ 0 & 1 & \mathbf{k} \\ 0 & 0 & \mathbf{j} \end{pmatrix},$$

$$C_2 = \begin{pmatrix} 2\mathbf{i} & 2 & 0 \\ 0 & 2 & 2\mathbf{k} \\ 0 & 0 & 2\mathbf{j} \end{pmatrix}, \qquad X_2 = \begin{pmatrix} \mathbf{i} & 1 & \mathbf{j} \\ 0 & \mathbf{j} & \mathbf{k} \\ 0 & 0 & \mathbf{k} \end{pmatrix},$$

$$\Omega_1 = \begin{pmatrix} -1 & 1+\mathbf{i}+\mathbf{k} & -1+2k \\ 0 & -1 & -j \\ 0 & 0 & -\mathbf{i} \end{pmatrix},$$

$$\Omega_2 = \begin{pmatrix} -2 & 2+2\mathbf{i}+2\mathbf{k} & -2+4k \\ 0 & -2 & -2j \\ 0 & 0 & -2\mathbf{i} \end{pmatrix}.$$

A direct calculation gives

$$B_1 X_2 B_2 = \Omega_1, \qquad C_1 X_2 C_2 = \Omega_2,$$

so the displayed X_2 solves both equations. Nevertheless,

$$\operatorname{rank}_{\mathbb{H}} \begin{pmatrix} A_1 & B_1 & 0 & 0 \\ 0 & C_1 & D_1 & E_1 \end{pmatrix} = 3, \qquad \operatorname{rank}_{\mathbb{H}} \begin{pmatrix} A_1 & B_1 & 0 & 0 & \Omega_1 \\ 0 & C_1 & D_1 & E_1 & \Omega_2 \end{pmatrix} = 6.$$

Therefore an independent aggregate rank equality can fail even for a consistent third-order system: it does not retain the ordered pairs (B_1, B_2) and (C_1, C_2) attached to the common X_2.

8. Application to Color-Image Dual-Watermarking

8.1. Quaternion Image Model and Diagonal Keys

Following the standard quaternion color-image representation [17], an $N \times N$ RGB image is represented by the pure quaternion matrix

$$F = R\mathbf{i} + G\mathbf{j} + B\mathbf{k}, \tag{48}$$

where the three real matrices contain normalized red, green, and blue channels. Let X_1 and X_2 be two watermark images and let X_3, X_4 be auxiliary arrays. Choose real diagonal key matrices

$$A_s = \mathrm{diag}(\mathbf{a}_s), \quad B_s = \mathrm{diag}(\mathbf{b}_s), \quad C_s = \mathrm{diag}(\mathbf{c}_s),$$
$$D_s = \mathrm{diag}(\mathbf{d}_s), \quad E_s = \mathrm{diag}(\mathbf{e}_s), \qquad s = 1, 2,$$

with positive diagonal entries. This is the specialization of system (1) used to form the two payloads; the equations are not repeated here. In this interpretation, X_1 and X_2 are the two target watermarks, X_3 and X_4 are independent auxiliary arrays, and Ω_1 and Ω_2 are the algebraic payloads embedded into two wavelet sub-bands. Because the coefficients are real and diagonal, a term such as $A_1 X_1 A_2$ multiplies entry (u, v) by the separable row–column gain $a_{1,u} a_{2,v}$; it does not mix different pixels. The repeated occurrence of X_2 creates the sequential dependence between the two payloads.

The coefficient matrices are not estimated from the host or watermark images. In the reported experiment, their ten diagonal vectors are generated independently by NumPy's MT19937 `RandomState` with seeds $1, \ldots, 10$, in the order listed in Table 2; every entry is sampled uniformly from $[0.5, 3.0)$. The strictly positive interval makes all ten diagonal matrices nonsingular and fixes the coefficient construction completely for reproduction. Other prescribed quaternion coefficients are permitted by the general theory, but would require the exact consistency and recovery formulas of Sections 5 and 6.

Proposition 5 (Pure-quaternion preservation)**.** *If X_1, X_2, X_3, X_4 are pure quaternion matrices and all key matrices are real, then Ω_1 and Ω_2 in (1) are pure quaternion matrices.*

Proof. Real scalars commute with quaternion units and do not create a real component. Every term in (1), and hence each sum, is therefore pure quaternion. □

Remark 2 (Pure-image coordinate reduction)**.** *The four-component real representation in Section 4 is required for arbitrary quaternion coefficients: a general two-sided product need not map the pure-quaternion subspace into itself. Proposition 5 shows that the real diagonal specialization does preserve this subspace. Consequently, the application may be represented by stacking only the three imaginary components, reducing the algebraic coordinate dimension from $4N^2$ to $3N^2$. The reported implementation already evaluates the diagonal quaternion formulas componentwise and never assembles the four-component dense operator; its dominant cost is the image-sized wavelet transform. We therefore retain the uniform four-component notation for the general theory and use direct three-channel arrays in the implementation.*

Once X_3, X_4 are fixed and A_1, A_2, C_1, C_2 are invertible, sequential recovery is unique:

$$X_2 = C_1^{-1}(\Omega_2 - D_1 X_3 D_2 - E_1 X_4 E_2) C_2^{-1}, \tag{49}$$
$$X_1 = A_1^{-1}(\Omega_1 - B_1 X_2 B_2) A_2^{-1}. \tag{50}$$

This is uniqueness of (X_1, X_2) under fixed auxiliary data, not uniqueness of the original four-unknown system. The latter is governed by the full-column-rank condition in Corollary 3.

If any of A_1, A_2, C_1, C_2 is singular, Proposition 5 remains valid, but the inverse Formulas (49) and (50) do not. In that case consistency must be checked with Theorem 3 (or Corollary 1), and all solutions are obtained from Corollary 2 by Moore–Penrose inverses and nullspace parameters. Thus a solution may fail to exist or may be nonunique. For diagonal coefficients, zero diagonal products identify the entries that cannot be recovered uniquely; their corresponding payload entries must satisfy the projector compatibility conditions.

Table 2. Implementation parameters and deterministic reproduction settings.

Item	Version or Setting
Software	Python 3.9.6; NumPy 2.0.2; SciPy 1.13.1; Matplotlib 3.9.4; Pillow 11.3.0; scikit-image 0.24.0; PyWavelets 1.6.0
Image sources	`skimage.data.astronaut()`, `chelsea()`, and `coffee()`; direct square resizing by Pillow LANCZOS
Quaternion encoding	RGB assigned to the $\mathbf{i}, \mathbf{j}, \mathbf{k}$ components; real component set to zero
Coefficient seeds	Independent `RandomState(j)` draws for $j = 1, \ldots, 10$ in the order $\mathbf{a}_1, \mathbf{b}_1, \mathbf{c}_1, \mathbf{d}_1, \mathbf{e}_1, \mathbf{a}_2, \mathbf{b}_2, \mathbf{c}_2, \mathbf{d}_2, \mathbf{e}_2$
Auxiliary arrays	Independent pure quaternion arrays generated by seeds 101 and 102
Wavelet convention	Haar, level 2, symmetric extension; $(cH_2, cV_2, cD_2) = (HL_2, LH_2, HH_2)$ in the PyWavelets tuple order
Embedding strengths	$0.02, 0.05, 0.08, 0.10, 0.15, 0.20, 0.25, 0.30$; selected setting $\alpha = 0.05$
Randomized processing	Global NumPy seed 2024 set once before the ordered processing list
Descriptive scatter plots	A fixed sample of 600 horizontal-neighbor pairs selected with `RandomState(0)`; the correlation analysis uses all valid neighbor pairs
Data and storage path	float64 calculation; $\mathcal{Q}_8$ in (64) for the stored branch; ideal branch evaluated separately

Proposition 6 (Diagonal structure and recovery stability)**.** *Let* $A = \operatorname{diag}(a_1, \ldots, a_N)$ *and* $B = \operatorname{diag}(b_1, \ldots, b_N)$ *be real diagonal matrices. With the columnwise ordering used in Section 4,*

$$\mathcal{T}(A, B) = \operatorname{diag}_{v=1,\ldots,N;\, u=1,\ldots,N}\big((a_u b_v) I_4\big). \tag{51}$$

If all diagonal entries are nonzero, then

$$\kappa_2\big(\mathcal{T}(A, B)\big) = \frac{\max_u |a_u| \, \max_v |b_v|}{\min_u |a_u| \, \min_v |b_v|}. \tag{52}$$

Assume further that A_1, A_2, C_1, C_2 *are nonsingular real diagonal matrices, that the keys and auxiliary arrays are exact, and define* $\Delta\Omega_s = \widehat{\Omega}_s - \Omega_s$*. The sequential recovery errors satisfy*

$$\Delta X_2 = C_1^{-1} \Delta\Omega_2 C_2^{-1}, \tag{53}$$

$$\Delta X_1 = A_1^{-1}\big(\Delta\Omega_1 - B_1 \Delta X_2 B_2\big) A_2^{-1}. \tag{54}$$

Consequently,

$$\|\Delta X_2\|_F \le \gamma_2 \|\Delta\Omega_2\|_F, \tag{55}$$

$$\|\Delta X_1\|_F \le \gamma_1 \|\Delta\Omega_1\|_F + \gamma_1 \beta \gamma_2 \|\Delta\Omega_2\|_F, \tag{56}$$

where

$$\gamma_1 = \|A_1^{-1}\|_2 \|A_2^{-1}\|_2, \qquad \gamma_2 = \|C_1^{-1}\|_2 \|C_2^{-1}\|_2, \qquad \beta = \|B_1\|_2 \|B_2\|_2.$$

Proof. For every entry, $(AXB)_{uv} = a_u x_{uv} b_v$. Since a_u and b_v are real, the corresponding 4×4 real block is $(a_u b_v) I_4$, which proves (51); its largest and smallest singular values give (52). Subtracting (49) and (50) from their perturbed counterparts gives (53) and (54). The bounds then follow from the submultiplicativity of the quaternion Frobenius norm with respect to the spectral norms of the real diagonal factors. □

8.2. Watermark Embedding

Let $F \in \mathbb{H}^{M \times M}$ be the host image, $X_1, X_2 \in \mathbb{H}^{N \times N}$ the watermark images, and $X_3, X_4 \in \mathbb{H}^{N \times N}$ fixed auxiliary arrays that are independent of the watermark images. In the reproducible experiment they are regenerated from two declared seeds rather than stored as

full arrays. We choose $N = M/4$, so the second-level detail sub-bands of a two-level Haar DWT have the same dimensions as the payloads. The embedding procedure is as follows.

Step 1: Algebraic payload generation. Using the diagonal key vectors, compute

$$\Omega_1 = \operatorname{diag}(\mathbf{a}_1) X_1 \operatorname{diag}(\mathbf{a}_2) + \operatorname{diag}(\mathbf{b}_1) X_2 \operatorname{diag}(\mathbf{b}_2), \tag{57}$$

$$\Omega_2 = \operatorname{diag}(\mathbf{c}_1) X_2 \operatorname{diag}(\mathbf{c}_2) + \operatorname{diag}(\mathbf{d}_1) X_3 \operatorname{diag}(\mathbf{d}_2) + \operatorname{diag}(\mathbf{e}_1) X_4 \operatorname{diag}(\mathbf{e}_2). \tag{58}$$

By Proposition 5, the payloads have three imaginary color components and no real component.
Step 2: Frequency decomposition. Apply a two-level Haar DWT to each RGB channel of F. Denote the unmodified second-level horizontal and vertical detail sub-bands by $T_{\mathrm{HL}}^{(0)}$ and $T_{\mathrm{LH}}^{(0)}$. These reference sub-bands are retained as side information, so the present construction is a non-blind watermarking method.
Step 3: Payload insertion. After applying a fixed invertible normalization independently to the three imaginary components, insert the payloads with strength $\alpha \in (0, 1)$:

$$T_{\mathrm{HL}} \leftarrow (1 - \alpha) T_{\mathrm{HL}}^{(0)} + \alpha \ \operatorname{rescale}(\operatorname{Im} \Omega_1), \tag{59}$$

$$T_{\mathrm{LH}} \leftarrow (1 - \alpha) T_{\mathrm{LH}}^{(0)} + \alpha \ \operatorname{rescale}(\operatorname{Im} \Omega_2). \tag{60}$$

Here the operations are performed componentwise on the RGB arrays represented by the pure quaternion payloads. More precisely, for one payload component with range $[o_{\min}, o_{\max}]$ and its target reference sub-band with range $[s_{\min}, s_{\max}]$, the experiment uses the affine map

$$\phi(t) = s_{\min} + \frac{s_{\max} - s_{\min}}{o_{\max} - o_{\min}} (t - o_{\min}), \qquad \phi^{-1}(z) = o_{\min} + \frac{o_{\max} - o_{\min}}{s_{\max} - s_{\min}} (z - s_{\min}). \tag{61}$$

For $r \in \{1, 2\}$ and $c \in \{R, G, B\}$, let

$$\Theta_{r,c} = \left(o_{r,c}^{-}, o_{r,c}^{+}, s_{r,c}^{-}, s_{r,c}^{+}\right) \tag{62}$$

denote the source and target endpoints used by the corresponding map. Hence $\Theta = \{\Theta_{r,c} : r = 1, 2, \ c = R, G, B\}$ contains 24 real scalars. All ranges in the reported experiment are nonzero; a constant payload or reference component must be rejected or handled by a separately specified convention because (61) would otherwise be noninvertible. For a fixed pair (r, c), if the selected sub-band component is perturbed by $\Delta T_{r,c}$ after insertion, the corresponding payload-component perturbation is related exactly through the inverse affine slope:

$$\|\Delta \Omega_{r,c}\|_F = \frac{o_{\max} - o_{\min}}{\alpha (s_{\max} - s_{\min})} \|\Delta T_{r,c}\|_F. \tag{63}$$

Thus decreasing α improves imperceptibility but amplifies extraction errors in proportion to $1/\alpha$.
Step 4: Reconstruction and key storage. Apply the inverse DWT to obtain the unquantized watermarked array F_w^{float}. The stored RGB image and its decoder are defined by

$$\mathcal{Q}_8(z) = \operatorname{uint8}(\operatorname{clip}(255z, 0, 255)), \qquad \mathcal{D}_8(y) = \frac{\operatorname{float64}(y)}{255}. \tag{64}$$

Conversion to `uint8` truncates toward zero after clipping. Thus $F_w^{(8)} = \mathcal{Q}_8(F_w^{\mathrm{float}})$ is distinct from the ideal floating-point branch. Define the public transform record

$$\mathcal{P} = \{M, N, \texttt{haar}, L = 2, \texttt{symmetric}, (cH_2, cV_2, cD_2) = (HL_2, LH_2, HH_2), \mathcal{Q}_8\}. \quad (65)$$

Store

$$\mathcal{K} = \{\mathbf{a}_1, \mathbf{b}_1, \mathbf{c}_1, \mathbf{d}_1, \mathbf{e}_1, \mathbf{a}_2, \mathbf{b}_2, \mathbf{c}_2, \mathbf{d}_2, \mathbf{e}_2, s_3, s_4, T_{\mathrm{HL}}^{(0)}, T_{\mathrm{LH}}^{(0)}, \Theta, \alpha, \mathcal{P}\}, \quad (66)$$

where s_3, s_4 regenerate X_3, X_4 using the fixed auxiliary-array generator specified below. The two reference sub-bands in (66) each contain three color components.

8.3. Watermark Extraction

Given a possibly processed watermarked image F'_w and the key set $\mathcal{K}$, apply the decoder $\mathcal{D}_8$ when the input is an 8-bit RGB image, regenerate $X_3 = \mathrm{Aux}(N, s_3)$ and $X_4 = \mathrm{Aux}(N, s_4)$, and apply the same two-level DWT. Approximate payloads are recovered from

$$\widehat{\Omega}_1 = \mathrm{rescale}^{-1}\left(\frac{\widehat{T}_{\mathrm{HL}} - (1-\alpha)T_{\mathrm{HL}}^{(0)}}{\alpha}\right), \quad (67)$$

$$\widehat{\Omega}_2 = \mathrm{rescale}^{-1}\left(\frac{\widehat{T}_{\mathrm{LH}} - (1-\alpha)T_{\mathrm{LH}}^{(0)}}{\alpha}\right). \quad (68)$$

The recovered watermarks are then computed sequentially:

$$\widehat{X}_2 = C_1^{-1}(\widehat{\Omega}_2 - D_1 X_3 D_2 - E_1 X_4 E_2) C_2^{-1}, \quad (69)$$

$$\widehat{X}_1 = A_1^{-1}(\widehat{\Omega}_1 - B_1 \widehat{X}_2 B_2) A_2^{-1}. \quad (70)$$

There is an important limitation. With fixed X_3, X_4 and invertible A_1, A_2, C_1, C_2, Formula (69) are defined for every extracted pair $(\widehat{\Omega}_1, \widehat{\Omega}_2)$. Algebraic consistency is therefore automatic in this particular implementation and cannot by itself serve as a tamper detector. The screening conditions (28)–(41) also add no authentication in this full-rank regime. Attack effects must be assessed by reconstruction measures, or an independent authentication code must be added if tamper detection is required.

Algorithms 1 and 2 give an implementation-level summary. They explicitly separate the ideal floating-point path from the clipped and quantized storage path and do not introduce a rank condition as an authentication test.

Figure 1 summarizes both data paths and distinguishes stored side information from ordinary data flow.

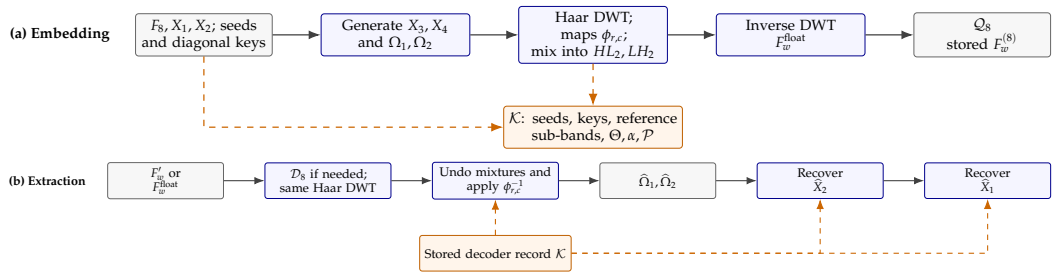


Figure 1. Algorithmic workflow for the reproducible non-blind construction. Solid arrows represent data flow, and dashed orange arrows represent decoder side information. The ideal floating-point branch is distinct from the clipped and quantized 8-bit branch.

Algorithm 1: Quaternion payload generation and non-blind DWT embedding

Input : RGB host $F_8 \in \{0,\dots,255\}^{M\times M\times 3}$; RGB watermarks $W_{1,8}, W_{2,8} \in \{0,\dots,255\}^{N\times N\times 3}$; strength α; coefficient and auxiliary seeds; transform record $\mathcal{P}$

Output: F_w^{float}, stored image $F_w^{(8)}$, and decoder record $\mathcal{K}$

Encode $W_{1,8}, W_{2,8}$ as pure quaternion arrays X_1, X_2 by (48)

Generate the ten real diagonal vectors and independently generate $X_3 = \mathrm{Aux}(N, s_3)$ and $X_4 = \mathrm{Aux}(N, s_4)$

Compute Ω_1, Ω_2 from (57)

foreach $c \in \{R, G, B\}$ **do**

- Compute the two-level Haar DWT of color component $F_8^{(c)}/255$ using symmetric extension
- Retain $S_{1,c} = HL_{2,c}^{(0)}$ and $S_{2,c} = LH_{2,c}^{(0)}$
- Form $\Theta_{1,c}, \Theta_{2,c}$ by (62) and require nonzero source and target ranges
- Replace $HL_{2,c}^{(0)}$ and $LH_{2,c}^{(0)}$ by the two mixtures in (59)
- Apply the inverse DWT, crop to $M \times M$, and assign the result to $F_w^{\mathrm{float},(c)}$

Set $F_w^{(8)} \leftarrow \mathcal{Q}_8(F_w^{\mathrm{float}})$ and assemble $\mathcal{K}$ as in (66)

return $F_w^{\mathrm{float}}, F_w^{(8)}, \mathcal{K}$

Algorithm 2: Side-information-assisted payload extraction and sequential watermark recovery

Input : A stored/processed RGB image F'_w, or the ideal array F_w^{float}; decoder record $\mathcal{K}$

Output: Unclipped floating-point estimates $\widehat{X}_1, \widehat{X}_2$

For the stored path set $Y \leftarrow \mathcal{D}_8(F'_w)$; for the ideal path set $Y \leftarrow F_w^{\mathrm{float}}$

Regenerate X_3, X_4 from s_3, s_4 and initialize $\widehat{\Omega}_1, \widehat{\Omega}_2$ with zero real components

foreach $c \in \{R, G, B\}$ **do**

- Compute the same two-level Haar DWT of $Y^{(c)}$ and read $\widehat{T}_{\mathrm{HL},c}, \widehat{T}_{\mathrm{LH},c}$
- Undo the two mixtures in (67) and apply the inverse affine maps recorded in $\Theta_{1,c}, \Theta_{2,c}$
- Assign the results to the c components of $\widehat{\Omega}_1, \widehat{\Omega}_2$

Recover $\widehat{X}_2$ first and then $\widehat{X}_1$ using (69)

return $\widehat{X}_1, \widehat{X}_2$; clip only copies used for RGB display

8.4. Experimental Setup and Imperceptibility

The numerical case study was implemented in Python using NumPy, PyWavelets, Pillow, SciPy, and scikit-image. For a seed s, $\mathrm{Aux}(N, s)$ denotes the pure quaternion array whose real part is zero and whose three imaginary components are the $N \times N \times 3$ output of an independent NumPy `RandomState(s)` (MT19937) uniform draw on $[0, 1)$. Its principal parameters are:

- Host image: 256×256 "astronaut" from scikit-image;
- Watermark X_1: 64×64 "chelsea", and watermark X_2: 64×64 "coffee";
- Independent auxiliary arrays: $X_3 = \mathrm{Aux}(64, 101)$ and $X_4 = \mathrm{Aux}(64, 102)$;
- Ten real diagonal vectors with entries sampled from $[0.5, 3.0)$ using seeds 1–10;
- Two-level Haar DWT and primary embedding strength $\alpha = 0.05$.

The three standard images were obtained from the `skimage.data` module [25]. According to the scikit-image dataset metadata, `astronaut` is a NASA image with no known copyright restrictions and was released into the public domain; `chelsea` was released under CC0 by the photographer Stefan van der Walt; and `coffee`, photographed by Rachel Michetti and provided courtesy of Pikolo Espresso Bar, was released under CC0. Accordingly, no additional permission is required for their use in the derived experimental figures. The auxiliary arrays contain no samples from either watermark and are regenerated by the decoder from their seeds. All fixed seeds are reproduction metadata, not claims of secret entropy.

Table 2 records the software environment and the remaining deterministic settings used to regenerate every figure and numerical table.

Supplementary Code S1 implements this deterministic pipeline. Running `python3 watermark_experiment_revised.py` from the code directory regenerates all experimental figures and numerical values, up to platform-dependent floating-point rounding. Because the coefficients are diagonal, payload formation requires $O(N^2)$ arithmetic, while each separable DWT or inverse DWT requires $O(M^2)$ operations. The specialized embedding and extraction procedures are therefore $O(M^2 + N^2)$ in time and do not form the dense real operator $\mathcal{M}$.

Watermark recovery is evaluated by normalized cross-correlation (NC), defined explicitly in (75); host-image imperceptibility is evaluated by the two measures below.

Imperceptibility is measured after clipping the inverse-DWT output to $[0, 1]$ and converting it to an 8-bit RGB image. For images F and F_w of size $H \times W$, we use

$$\mathrm{PSNR}(F, F_w) = 10 \log_{10} \frac{255^2 \cdot 3HW}{\sum_{c \in \{R,G,B\}} \|F^{(c)} - F_w^{(c)}\|_F^2} \ \mathrm{dB}, \tag{71}$$

together with the following global, mean-channel structural similarity (SSIM) estimate:

$$S_{\mathrm{glob}}(F, F_w) = \frac{1}{3} \sum_{c \in \{R,G,B\}} \frac{(2\mu_c \widehat{\mu}_c + K_1)(2\sigma_{c\widehat{c}} + K_2)}{(\mu_c^2 + \widehat{\mu}_c^2 + K_1)(\sigma_c^2 + \widehat{\sigma}_c^2 + K_2)}, \tag{72}$$

where $\mu_c, \widehat{\mu}_c$ are whole-channel means, $\sigma_c^2, \widehat{\sigma}_c^2$ are whole-channel population variances, $\sigma_{c\widehat{c}}$ is their population covariance, and $K_1 = (0.01 \cdot 255)^2$, $K_2 = (0.03 \cdot 255)^2$. Equation (72) is stated explicitly because it is a global estimate rather than the standard local-window SSIM implementation; the two should not be compared as if they were identical metrics. Table 3 reports both measures over the tested embedding strengths.

Figure 2 presents the inputs, payloads, stored image, and recovered outputs at the primary setting; the algorithmic data flow itself is shown in Figure 1.

Table 3. PSNR and the global SSIM estimate in (72) after clipping and 8-bit conversion at different embedding strengths.

α	PSNR (dB)	Global SSIM
0.02	43.56	0.9998
0.05	**35.99**	**0.9987**
0.08	31.98	0.9966
0.10	30.07	0.9947
0.15	26.60	0.9882
0.20	24.15	0.9793
0.25	22.26	0.9681
0.30	20.72	0.9547

Note: Bold values indicate the operating point used in the subsequent experiments.

Figure 3 combines the values in Table 3 with the storage-baseline recovery trends as α varies.

Figure 3 represents a three-way tradeoff rather than a universally optimal choice of α. Increasing α raises the payload contribution relative to transform and quantization errors, but also increases host-image distortion and clipping. Conversely, a small α improves imperceptibility while (63) amplifies sub-band perturbations in proportion to $1/\alpha$. At $\alpha = 0.05$, PSNR $= 35.99$ dB and the global structural-similarity estimate is 0.9987. The corresponding storage-baseline NC values are reported below.

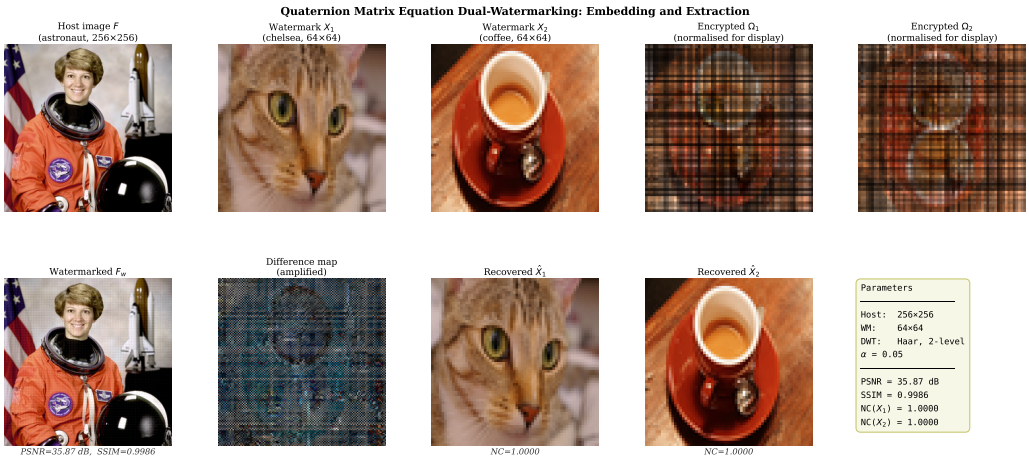


Figure 2. Embedding and extraction overview at $\alpha = 0.05$, regenerated using the clipped and quantized 8-bit image as the no-attack storage baseline.

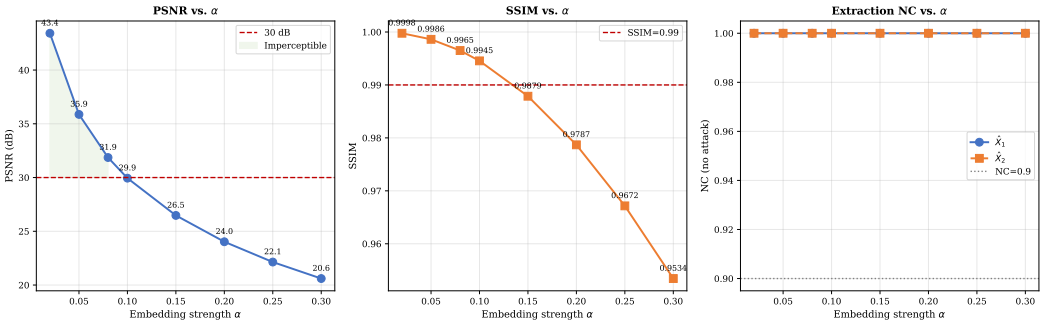


Figure 3. Influence of the embedding strength α on PSNR, the global structural-similarity estimate in (72), and storage-baseline NC. Every point is computed after clipping and 8-bit conversion. The horizontal lines are visual reference levels, and the vertical line identifies the selected setting $\alpha = 0.05$.

8.5. Ideal Floating-Point Recovery and the Storage Baseline

Before clipping and quantization, applying extraction directly to the in-memory floating-point inverse-DWT array gives

$$\|\widehat{\Omega}_1 - \Omega_1\|_\infty \approx 7.73 \times 10^{-14}, \qquad \|\widehat{\Omega}_2 - \Omega_2\|_\infty \approx 1.19 \times 10^{-13}, \tag{73}$$

and NC equal to 1.0000 for both recovered watermarks. This is an ideal transform round trip with the original sub-bands and normalization parameters stored as side information; it is not recovery from a saved 8-bit watermarked image.

In fact, at $\alpha = 0.05$ the unquantized array contains 13,652 channel values outside $[0, 1]$ out of 196,608 values in total. After clipping, 8-bit conversion, and immediate re-extraction without any further attack, the reproduced NC values are approximately

$$\mathrm{NC}(X_1, \widehat{X}_1) = 0.3921, \qquad \mathrm{NC}(X_2, \widehat{X}_2) = 0.7179. \tag{74}$$

These values are produced by the following fixed path: inverse DWT reconstruction, clipping to $[0, 1]$, conversion by $\mathcal{Q}_8$, decoding by $\mathcal{D}_8$, a second two-level DWT, inverse sub-band mixing and affine rescaling, sequential recovery of $\widehat{X}_2$ and then $\widehat{X}_1$, and finally evaluation of (75). NC is the cosine of the angle between the stacked three-component watermark arrays: 1 means perfect directional agreement, whereas values closer to zero indicate weak agreement. Thus 0.3921 for X_1 and 0.7179 for X_2 show substantial storage-induced degradation rather than robust recovery. The lower value for X_1 is consistent with (56), because the error in $\widehat{X}_2$ propagates into the subsequent recovery of $\widehat{X}_1$. Accordingly, the ideal and storage baselines must be reported separately. The discrepancy also motivates a bounded or quantization-aware embedding design.

8.6. Attack-Sensitivity Tests

We next evaluate the storage baseline and 13 post-processing operations, including Joint Photographic Experts Group (JPEG) compression, on the clipped and quantized image. All rows, including the no-further-processing storage baseline, use the same 8-bit data path. Normalized cross-correlation is defined by

$$\mathrm{NC}(X, \widehat{X}) = \frac{\left\langle \mathrm{vec}(\mathrm{Im}\, X), \mathrm{vec}(\mathrm{Im}\, \widehat{X}) \right\rangle}{\| \mathrm{vec}(\mathrm{Im}\, X) \|_2 \, \| \mathrm{vec}(\mathrm{Im}\, \widehat{X}) \|_2}, \tag{75}$$

where $\mathrm{vec}(\mathrm{Im}\, X)$ stacks all three imaginary components. NC is evaluated from the unclipped floating-point estimates; clipping is used only for display. If either norm in (75) is zero, the implementation returns zero by convention; this case does not occur in the reported experiment.

Table 4 specifies every operation. Randomized operations use the single seeded sequence stated in Table 2; deterministic operations use the listed interpolation and boundary conventions.

The operations in Table 4 are applied independently to the same stored baseline. Thus, differences between rows below reflect the stated operation under one common extraction pipeline, not a cryptographic security evaluation.

Figure 4 places the original watermarks beside six representative processed configurations so that visual degradation can be assessed directly. The complete set of numerical comparisons remains in Table 5 and Figure 5. Figure 6 compares the level-two sub-bands before interpreting their differences.

Table 4. Definitions of the post-processing operations applied to the stored 8-bit image.

Operation	Parameter	Implementation
Storage baseline	None	Direct extraction after clipping and $\mathcal{Q}_8$, without additional processing
Gaussian noise	$\sigma = 2, 5$	Independent samples in 8-bit intensity units added to every pixel and channel, followed by clipping and conversion to `uint8`
Salt-and-pepper	Probability 0.01	Each spatial location selected independently with probability 0.01; each selected RGB triplet replaced jointly by black or white with equal probability
JPEG	$Q = 90, 70, 50$	In-memory Pillow encode/decode, 4:4:4 chroma (`subsampling` $= 0$), without optimization or progressive coding

Table 4. *Cont.*

Operation	Parameter	Implementation
Border removal	Width 0.10	Strips of $\lfloor 0.1M \rfloor = 25$ pixels set to zero on all four edges, with canvas size unchanged
Rotation	5°	Counterclockwise about the center; bilinear interpolation, fixed canvas, and black fill
Median filtering	3×3	Applied independently to each channel with reflected boundary handling
Gamma correction	$\gamma = 1.5$	$z \mapsto 255(z/255)^{1.5}$, followed by clipping and 8-bit conversion
Sharpening	Gain 0.5	Per-channel $I + 0.5(I - U_3 I)$, where U_3 is a 3×3 uniform filter with reflected boundaries; followed by clipping and conversion to `uint8`
Histogram equalization	Per channel	Pillow equalization applied independently to RGB
Scaling	Factor 0.5	Resize $256 \to 128 \to 256$ with bilinear interpolation in both steps

Table 5. NC values for the common 8-bit storage baseline and 13 post-processing operations at $\alpha = 0.05$.

Configuration	NC for X_1	NC for X_2
8-bit storage, no further processing	0.3921	0.7179
Gaussian noise ($\sigma = 2$)	0.2229	0.5843
Gaussian noise ($\sigma = 5$)	0.1236	0.3953
Salt-and-pepper noise (1%)	0.0703	0.1249
JPEG compression ($Q = 90$)	0.1617	0.4612
JPEG compression ($Q = 70$)	0.0788	0.3186
JPEG compression ($Q = 50$)	0.0091	0.2930
Zeroing a 10%-wide border	−0.0345	0.1161
Rotation by 5°	−0.0210	0.1062
Median filtering (3×3)	−0.0628	0.3789
Gamma correction ($\gamma = 1.5$)	0.0488	0.1955
Sharpening	0.1972	0.0663
Histogram equalization	0.0912	−0.0359
Scaling by 0.5 and restoring size	−0.0672	0.2988

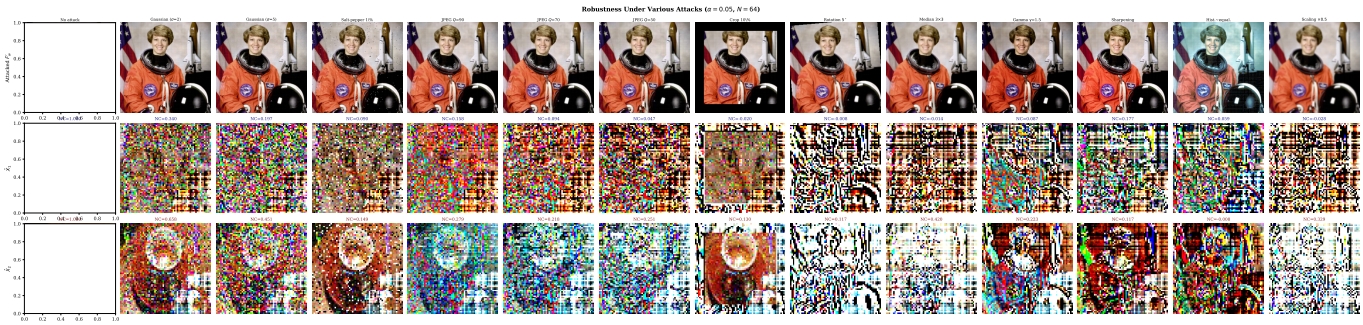


Figure 4. Representative processed images and recovered watermark estimates at $\alpha = 0.05$. The left column gives the original host and watermarks. Solid blue and dashed orange frames distinguish X_1 and X_2, respectively; the NC badges use the common continuous color scale shown below the panels.

Figure 5 combines the data in Table 5 with process-category bands and a direct visual comparison of the two recovered watermarks. The storage baseline already includes clip-

ping and quantization and is distinct from the ideal floating-point result. Each additional operation reduces at least one NC value relative to that baseline, and none reaches the displayed visual reference value of 0.9. Within the noise group, increasing the Gaussian standard deviation from 2 to 5 produces a further decrease for both watermarks; the JPEG group likewise shows an overall loss as quality is reduced from 90 to 50. Geometric changes lead to near-zero or negative NC for X_1, whereas the scaling operation retains an NC of 0.2988 for X_2. The estimates of X_2 are better than those of X_1 in most rows. This agrees with the sequential decoder, because an error in $\widehat{X}_2$ also enters the subsequent formula for $\widehat{X}_1$. These observations characterize the stated experiment and should not be interpreted as evidence of "double protection", authentication, or cryptographic security.

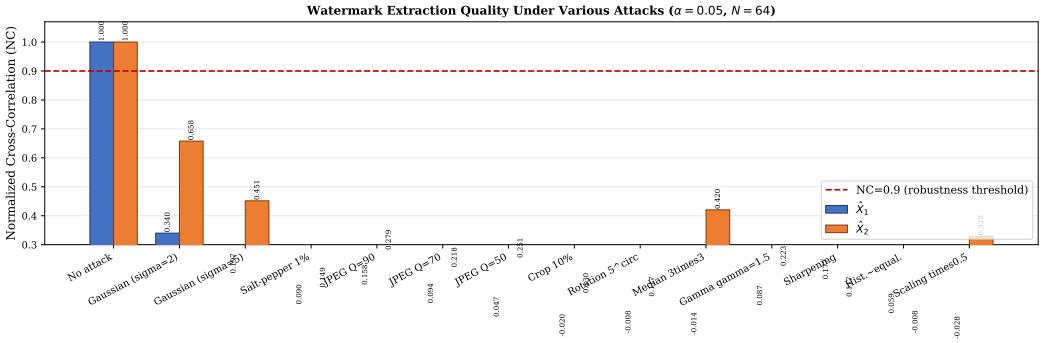


Figure 5. NC values from Table 5, grouped by operation type. Blue diagonally hatched bars represent X_1, and orange dotted bars represent X_2. The dashed line at NC $= 0.9$ is a visual reference only, not a universal robustness criterion; negative values denote anticorrelation under (75).

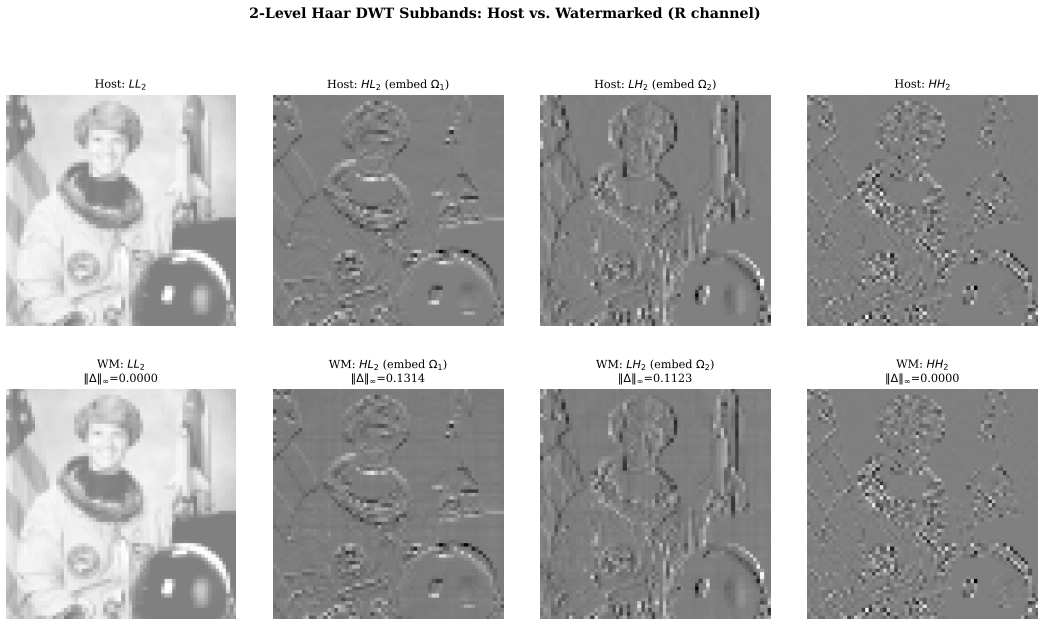


Figure 6. Two-level Haar DWT sub-bands of the host and the clipped, 8-bit watermarked image for the red channel. The largest intended changes occur in HL_2 and LH_2; clipping and quantization introduce smaller changes outside those two sub-bands.

8.7. DWT Sub-Band Inspection

Before clipping and quantization, the construction changes only the selected HL_2 and LH_2 coefficient arrays up to floating-point roundoff. Figure 6 instead uses the stored-image path and therefore reflects the additional perturbation introduced by mapping the inverse-DWT result to a valid 8-bit image. This inspection verifies which frequency bands carry the payload but does not establish resistance to signal-processing attacks.

8.8. Descriptive Payload Statistics

Table 6 reports genuine horizontal, vertical, and diagonal neighbor correlations, calculated using all valid neighboring pixel pairs. The payload values are analyzed before the DWT normalization because an affine rescaling does not change a correlation coefficient.

Table 6. Spatial neighbor correlations of watermark X_1 and payload Ω_1.

Array and Component	Horizontal	Vertical	Diagonal
X_1 (R)	0.7701	0.8695	0.7008
X_1 (G)	0.7879	0.8818	0.7278
X_1 (B)	0.8477	0.9181	0.8080
Ω_1 (**i**)	0.6309	0.4358	0.1383
Ω_1 (**j**)	0.6677	0.4707	0.2214
Ω_1 (**k**)	0.6820	0.5212	0.2919

Figure 7 complements the numerical correlations with the marginal histograms and correctly sampled horizontal-neighbor scatter plots.

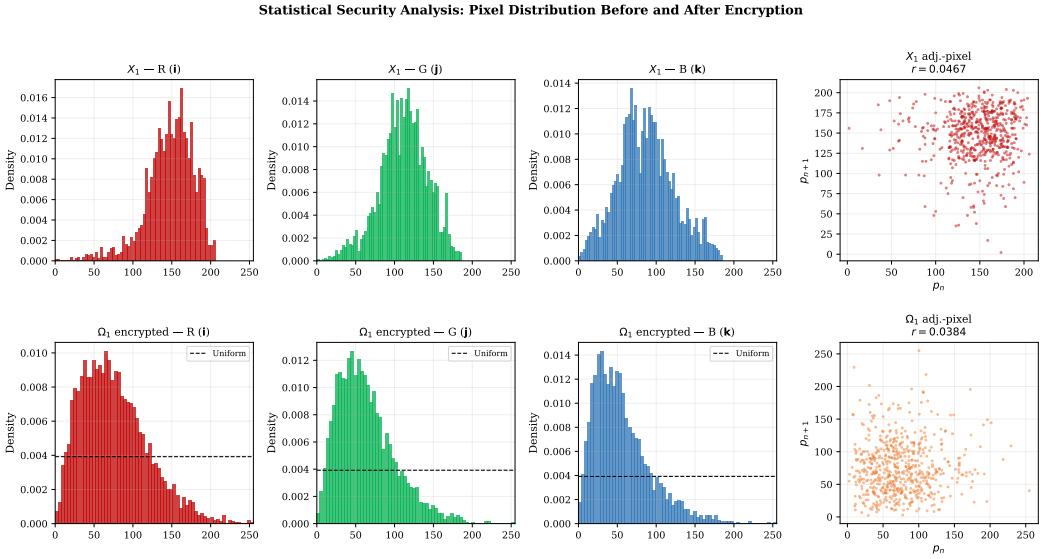


Figure 7. Channel histograms and horizontal-neighbor scatter plots for X_1 and the algebraic payload Ω_1. Each scatter plot uses the same fixed sample of 600 pairs (seed 0), whereas Table 6 uses all valid pairs.

For this example, forming Ω_1 changes the marginal distributions and reduces some spatial correlations, particularly diagonal correlations. Substantial correlations remain, and the payload histograms are not uniform. These descriptive statistics therefore do not establish cryptographic security.

8.9. Key Material, Side Information, and Security Limitations

The decoder stores ten real diagonal vectors, two auxiliary-array seeds, the original HL_2 and LH_2 sub-bands for three color channels, the 24 normalization endpoints, and α. Counting the stored double-precision quantities while excluding the two integer seeds and the small transform record $\mathcal{P}$, this is

$$N_{\text{store}} = 10N + 6N^2 + 25 \tag{76}$$

double-precision values. Here $10N$ counts the ten diagonal vectors, $6N^2$ counts the two three-channel reference sub-bands, and $25 = 24 + 1$ counts the normalization endpoints and

embedding strength. Equation (76) is a storage count, not a measure of secret entropy. IEEE-754 representation bits cannot all be treated as independent key bits; in this reproducible experiment all coefficient and auxiliary generators use disclosed fixed seeds. Crucially, X_3 and X_4 are generated independently of X_1 and X_2; the decoder record therefore no longer contains either target watermark or a deterministic transform of it.

Moreover, the present coefficients are real diagonal matrices. Their action is componentwise real scaling and addition, so quaternion noncommutativity does not supply an additional cryptographic barrier in this implementation. The prototype should be understood as a deterministic, non-blind, side-information-assisted numerical encoding experiment. No claim of semantic security or resistance to known-plaintext, chosen-plaintext, or key-recovery attacks is made.

8.10. Relation to Existing Quaternion Image Models

Table 7 compares only the algebraic structures and application roles of the selected models.

Table 7. Structural comparison with selected quaternion matrix-equation image models. This is not a controlled performance benchmark.

Study	Algebraic Structure	Image Task	Shared-Variable Pattern
He et al. [23]	one Sylvester-type equation	data transmission	none
Qi et al. [16]	three coupled equations	color-image encryption	Y in three equations
Present case study	two coupled equations	dual watermarking	X_2 in two equations

The table compares model structures only. Datasets, image sizes, embedding strengths, side-information assumptions, and metric implementations differ, so PSNR or NC values from separate studies should not be interpreted as a head-to-head ranking. In the present model, the shared variable produces a sequential decoding dependence; it does not provide an independent consistency or authentication check in the diagonal full-rank implementation.

9. Discussion: Key Findings, Limitations, and Future Directions

9.1. Key Findings

The central theoretical finding is that the shared-variable compatibility is an ordinary real linear constraint only after every ordered two-sided product has been represented intact. The augmented-rank test (16) and the reduced criterion (19) are exact, whereas conditions (28)–(41) are computationally useful rejection screens. Proposition 3 separates these roles: it characterizes joint sufficiency through the image of the reduced operator and identifies full-rank and zero-pattern regimes in which the screens are sufficient. The explicit 3×3 counterexamples verify the stated failures of the aggregate and screening tests at third order.

The application illustrates a second distinction. With fixed auxiliary arrays and nonsingular diagonal coefficients, the algebraic payloads are recovered to floating-point roundoff in the ideal in-memory path. The stored-image path is materially different: clipping and 8-bit quantization reduce the baseline NC values to 0.3921 and 0.7179. This gap is consistent with the error-propagation bounds in Proposition 6 and shows why an exact algebraic solution does not by itself imply robust image storage.

9.2. Limitations

The numerical study is a deterministic proof of concept using one host image, two payload images, one principal embedding strength, and disclosed fixed seeds. The reported global mean-channel SSIM estimate is not the standard local-window SSIM. The decoder is non-blind and requires the side information counted in (76); the diagonal real coefficients

provide separable scaling rather than a noncommutative cryptographic barrier. The tested post-processing results therefore characterize sensitivity under the stated pipeline and do not establish security, authentication, or population-level robustness.

The theoretical and experimental scopes must also remain separate. The rank results allow general rectangular quaternion coefficients and do not depend on the image experiment. Conversely, the successful ideal round trip verifies one implementation but is not evidence for the general theorem. The four-component real representation is necessary at that level of generality, although the pure-quaternion diagonal application permits the three-component restriction discussed after Proposition 5.

9.3. Future Directions and Reproducibility

Future work should investigate structured or matrix-free exact solvers, additional coefficient classes in which the screening conditions are sufficient, and regularized recovery for singular or noisy systems. For watermarking, priorities include bounded or quantization-aware sub-band insertion, experiments over multiple host and payload datasets, repeated independent keys, standard local-window SSIM, and controlled comparisons under identical attacks. A deployable authentication method would additionally need an explicit threat model and an independent integrity mechanism.

Supplementary Code S1 reproduces the figures and numerical tables using the software versions, random seeds, image sources, transform conventions, and storage path reported here. The same clipped 8-bit baseline is used for the unprocessed and processed cases, the full NC range is displayed, and spatial correlations use genuine neighboring pairs. OpenAI Codex was used during revision to assist with LaTeX restructuring, language editing, and consistency checks; all formulae, proofs, code outputs, and interpretations were reviewed by the authors.

10. Conclusions

This paper established an order-preserving real representation for a coupled system of two-sided quaternion matrix equations with a shared unknown. It yields an exact augmented-rank criterion, an equivalent reduced criterion, the complete Moore–Penrose solution, and fourteen quaternion block-rank screening conditions. The added sufficiency characterization and the 3×3 counterexamples delineate when those compact screens can and cannot replace the exact real compatibility test.

The dual-watermarking example demonstrates the resulting sequential recovery model while also showing its present limits. At embedding strength 0.05, the stored image attains PSNR 35.99 dB and a global structural-similarity estimate of 0.9987, but clipping and 8-bit quantization reduce the storage-baseline NC values to 0.3921 and 0.7179. The application is therefore an algebraically exact, reproducible non-blind prototype rather than a claim of robust or cryptographically secure watermarking.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/math14183259/s1: Supplementary Code S1, `watermark_experiment_revised.py`, is provided with the revised manuscript and contains the Python source code used to reproduce the dual-watermarking experiment, parameter sweep, post-processing tests, numerical results, and experimental figures. It records the software versions, random seeds, image sources, transform conventions, and storage path used in the reported calculations.

Author Contributions: Conceptualization, Z.-H.H. and S.-W.Y.; methodology, Z.-H.H. and J.-X.L.; software, J.-X.L. and C.-Y.Q.; validation, Z.-H.H., J.-X.L. and C.-Y.Q.; formal analysis, Z.-H.H. and J.-X.L.; writing—original draft preparation, J.-X.L. and C.-Y.Q.; writing—review and editing, Z.-H.H. and S.-W.Y.; supervision, Z.-H.H. and S.-W.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (grant numbers 12271338 and 12371023).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The host image (`astronaut`) and payload images (`chelsea` and `coffee`), together with their source and license metadata, are publicly available in the scikit-image data module (https://scikit-image.org/docs/0.24.x/api/skimage.data.html) (accessed on 13 August 2026).

Acknowledgments: The authors thank the editor and reviewers for their constructive comments. During preparation of this revision, the authors used OpenAI Codex (GPT-5; accessed on 12 August 2026) for LaTeX revision, language editing, and consistency checking. The authors reviewed and edited all outputs and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Real-Operator Identities and Equivalence Invariance

This appendix gives two identities that connect the coordinated canonical reductions in Section 3 with the exact real-rank theory. They also provide independent checks on the order of every noncommutative factor in the construction of $\mathcal{M}$.

Lemma A1 (Composition law)**.** *Let all products below be conformable. Then*

$$\mathcal{T}(A,B)\mathcal{T}(C,D) = \mathcal{T}(AC,DB). \tag{A1}$$

In particular, if P and Q are square and invertible, then

$$\mathcal{T}(P,Q)^{-1} = \mathcal{T}(P^{-1},Q^{-1}). \tag{A2}$$

Proof. For an arbitrary quaternion matrix X, Lemma 2 gives

$$\mathcal{T}(A,B)\mathcal{T}(C,D)\,\mathrm{rvec}(X) = \mathrm{rvec}\big(A(CXD)B\big) = \mathrm{rvec}\big((AC)X(DB)\big).$$

Since this identity holds for every X, it proves (A1). Applying the composition law in both orders to (P,Q) and (P^{-1},Q^{-1}) gives (A2). Notice that the right factors occur as DB, not BD; this reversal is forced by the nesting of the two right multiplications. □

Proposition A1 (Invariance under coordinated equivalence)**.** *Under the transformations in Proposition 1, define*

$$\mathcal{U} = \mathrm{diag}\big(\mathcal{T}(P_1,\widetilde{Q}_1),\mathcal{T}(P_2,\widetilde{Q}_2)\big), \tag{A3}$$
$$\mathcal{V} = \mathrm{diag}\big(\mathcal{T}(Q_1,\widetilde{P}_1),\mathcal{T}(Q_2,\widetilde{P}_2),\mathcal{T}(Q_3,\widetilde{P}_3),\mathcal{T}(Q_4,\widetilde{P}_4)\big). \tag{A4}$$

Let $\widehat{\mathcal{M}}$ and $\widehat{\omega}$ be the real coefficient matrix and right-hand side obtained from the canonical system (9). Then

$$\widehat{\mathcal{M}} = \mathcal{U}\mathcal{M}\mathcal{V}, \qquad \widehat{\omega} = \mathcal{U}\omega, \qquad x = \mathcal{V}\hat{x}. \tag{A5}$$

Consequently,

$$\mathrm{rank}_{\mathbb{R}}\,\widehat{\mathcal{M}} = \mathrm{rank}_{\mathbb{R}}\,\mathcal{M}, \qquad \mathrm{rank}_{\mathbb{R}}\begin{pmatrix}\widehat{\mathcal{M}} & \widehat{\omega}\end{pmatrix} = \mathrm{rank}_{\mathbb{R}}\begin{pmatrix}\mathcal{M} & \omega\end{pmatrix}. \tag{A6}$$

Thus the exact condition (16), the nullity of the coefficient operator, and the affine dimension of every consistent solution set are invariant under coordinated equivalence reduction.

Proof. The variable relation in (7) is equivalent to $X_s = Q_s \widehat{X}_s \widetilde{P}_s$, so Lemma 2 gives $x = \mathcal{V}\widehat{x}$. For a nonzero term in the ith equation involving the sth unknown, denote its ordered left and right coefficients by (L_{is}, R_{is}). Lemma A1 yields

$$\mathcal{T}(P_i, \widetilde{Q}_i)\,\mathcal{T}(L_{is}, R_{is})\,\mathcal{T}(Q_s, \widetilde{P}_s) = \mathcal{T}(P_i L_{is} Q_s, \widetilde{P}_s R_{is} \widetilde{Q}_i).$$

This is precisely the real operator associated with the corresponding canonical term. Applying the same identity blockwise proves the first two relations in (A5). All diagonal blocks in $\mathcal{U}$ and $\mathcal{V}$ are invertible by (A2). Finally,

$$\begin{pmatrix} \widehat{\mathcal{M}} & \widehat{\omega} \end{pmatrix} = \mathcal{U} \begin{pmatrix} \mathcal{M} & \omega \end{pmatrix} \begin{pmatrix} \mathcal{V} & 0 \\ 0 & 1 \end{pmatrix},$$

and left or right multiplication by an invertible real matrix preserves rank. This proves (A6); invariance of nullity and affine dimension follows from rank–nullity. □

References

1. Paige, C.C.; Saunders, M.A. Towards a generalized singular value decomposition. *SIAM J. Numer. Anal.* **1981**, *18*, 398–405. [CrossRef]
2. De Moor, B.; Golub, G.H. The restricted singular value decomposition: Properties and applications. *SIAM J. Matrix Anal. Appl.* **1991**, *12*, 401–425. [CrossRef]
3. De Moor, B.; Zha, H. A tree of generalizations of the ordinary singular value decomposition. *Linear Algebra Its Appl.* **1991**, *147*, 469–500. [CrossRef]
4. Jia, Z.-G.; Ng, M.K.; Song, G.-J. Lanczos method for large-scale quaternion singular value decomposition. *Numer. Algorithms* **2019**, *82*, 699–717. [CrossRef]
5. Wang, Q.-W.; Van der Woude, J.W.; Yu, S.-W. An equivalence canonical form of a matrix triplet over an arbitrary division ring with applications. *Sci. China Math.* **2011**, *54*, 907–924. [CrossRef]
6. Wang, Q.-W.; Zhang, X.; Van der Woude, J.W. A new simultaneous decomposition of a matrix quaternity over an arbitrary division ring with applications. *Commun. Algebra* **2012**, *40*, 2309–2342. [CrossRef]
7. He, Z.-H.; Xu, Y.-Z.; Wang, Q.-W.; Zhang, C.-Q. The equivalence canonical forms of two sets of five quaternion matrices with applications. *Math. Methods Appl. Sci.* **2025**, *48*, 5483–5505. [CrossRef]
8. Bhattacharyya, S.P.; de Souza, E. Pole assignment via Sylvester's equation. *Syst. Control Lett.* **1982**, *1*, 261–263. [CrossRef]
9. Gustafson, W.H. Quivers and matrix equations. *Linear Algebra Appl.* **1995**, *231*, 159–174. [CrossRef]
10. Roth, W.E. The equations $AX - YB = C$ and $AX - XB = C$ in matrices. *Proc. Am. Math. Soc.* **1952**, *3*, 392–396.
11. Wimmer, H.K. Consistency of a pair of generalized Sylvester equations. *IEEE Trans. Autom. Control* **1994**, *39*, 1014–1016. [CrossRef]
12. Lee, S.G.; Vu, Q.P. Simultaneous solutions of matrix equations and simultaneous equivalence of matrices. *Linear Algebra Appl.* **2012**, *437*, 2325–2339. [CrossRef]
13. Kyrchei, I. Cramer's rules for Sylvester quaternion matrix equation and its special cases. *Adv. Appl. Clifford Algebr.* **2018**, *28*, 90. [CrossRef]
14. Kyrchei, I. Determinantal representations of solutions to systems of two-sided quaternion matrix equations. *Linear Multilinear Algebra* **2021**, *69*, 648–672. [CrossRef]
15. He, Z.-H.; Xu, Y.-Z.; Wang, Q.-W.; Wang, X.-X. Solving a system of quaternion matrix equations by using PSVD for multiple matrices with applications. *J. Comput. Appl. Math.* **2026**, *476*, 117072. [CrossRef]
16. Qi, C.-Y.; Liu, C.; He, Z.-H.; Yu, S.-W. A color image encryption model based on a system of quaternion matrix equations. *Mathematics* **2026**, *14*, 319. [CrossRef]
17. Pei, S.-C.; Cheng, C.-M. A novel block truncation coding of color images using a quaternion-moment-preserving principle. *IEEE Trans. Commun.* **1997**, *45*, 583–595. [CrossRef]
18. Le Bihan, N.; Sangwine, S. Quaternion principal component analysis of color images. In *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)*; IEEE: Piscataway, NJ, USA, 2003; Volume 1, pp. I–809–I–812.
19. Zhang, F. Quaternions and matrices of quaternions. *Linear Algebra Appl.* **1997**, *251*, 21–57. [CrossRef]
20. Rodman, L. *Topics in Quaternion Linear Algebra*; Princeton University Press: Princeton, NJ, USA, 2014.
21. Jia, Z.-G.; Ng, M.K. Structure preserving quaternion generalized minimal residual method. *SIAM J. Matrix Anal. Appl.* **2021**, *42*, 616–634. [CrossRef]

22. Ke, Y.-F.; Ma, C.-F.; Jia, Z.-G.; Xie, Y.-J.; Liao, R.-W. Quasi non-negative quaternion matrix factorization with application to color face recognition. *J. Sci. Comput.* **2023**, *95*, 38. [CrossRef]
23. He, Z.-H.; Qin, W.-L.; Tian, J.; Wang, X.-X.; Zhang, Y. A new Sylvester-type quaternion matrix equation model for color image data transmission. *Comput. Appl. Math.* **2024**, *43*, 227. [CrossRef]
24. Lai, C.-C.; Tsai, C.-C. Digital image watermarking using discrete wavelet transform and singular value decomposition. *IEEE Trans. Instrum. Meas.* **2010**, *59*, 3060–3063. [CrossRef]
25. Scikit-Image Developers. `Skimage.Data`—Scikit-Image v0.24.0 Documentation. Available online: https://scikit-image.org/docs/0.24.x/api/skimage.data.html (accessed on 13 August 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.