On Closed-Form Formulas for the 3D Nearest Rotation Matrix Problem

Soheil Sarabandi, Arya Shabani, Josep M. Porta and Federico Thomas

Abstract—The problem of restoring the orthonormality of a noisy rotation matrix by finding its *nearest* correct rotation matrix arises in many areas of robotics, computer graphics, and computer vision. When the Frobenius norm is taken as the measure of closeness, the solution is usually computed using the singular value decomposition (SVD). A closed-form formula exists but, as it involves the roots of a polynomial of third degree, it is assumed to be too complicated and numerically ill-conditioned. In this paper, we show how, by carefully using some algebraic recipes scattered in the literature, it is possible to derive a simple and yet numerically stable formula for most practical applications. Moreover, by relying on a result that permits obtaining the quaternion corresponding to the sought optimal rotation matrix, we present another closed-form formula that provides a good approximation to the optimal one using only the elementary algebraic operations of addition, subtraction, multiplication and division. These two closed-form formulas are compared with respect to the SVD in terms of accuracy and computational cost.

Index Terms—Rotation matrices, quaternions, singular value decomposition, third degree polynomials.

I. INTRODUCTION

Rotations in 3D are commonly represented using 3×3 proper orthogonal matrices (also known as *rotation matrices*). A matrix, say $\mathbf{R}$, is said to be orthogonal if $\mathbf{R}\mathbf{R}^T$ is equal to the identity and to be proper if, in addition, $\det(\mathbf{R}) = 1$. In other words, the three row and column vectors of $\mathbf{R}$ represent a right-handed orthonormal reference frame. There are some applications in robotics, computer vision, and computer graphics in which *noisy* rotation matrices are generated. That is, rotation matrices that satisfy the two aforementioned conditions approximately. Due to floating point precision errors, we have that, for example, the result of cumulatively multiplying rotation matrices is a noisy rotation matrix [1]. Likewise, the integration of angular velocity differential equations leads to a rotation matrix that progressively departs from orthogonality as time increases [2]. A way to alleviate these problems consists in representing rotations using unit quaternions, which are only converted to rotation matrices when required. Floating point precision errors in these cases lead to non-unit quaternions, but normalizing them is a trivial task. Unfortunately, using quaternions is not always desirable or even possible. For example, rotating a vector by a quaternion using the sandwich formula [3] is computationally much more expensive than rotating it using the standard multiplication by the corresponding rotation matrix [4]. Noisy rotation matrices not only arise as the result of floating point operations. For example, one simple solution to the problem of determining the rotation matrix that is the best fit to a given set of measured rotations consists in averaging all measurements (see [5] and the references therein). However, the result is not, in general, a proper orthogonal matrix. Then, the need arise for orthonormalizing noisy rotation matrices, a process that essentially
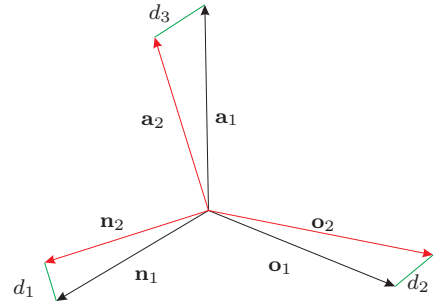
Fig. 1. Following the standard robotics notation, the column vectors of the rotation matrix $\mathbf{R}_i$ can be made explicit as $\mathbf{R}_i = (\mathbf{n}_i \ \mathbf{o}_i \ \mathbf{a}_i)$ [6]. These vectors determine a reference frame. Then, in this figure, the Frobenius norm of the difference between $\mathbf{R}_1$ and $\mathbf{R}_2$ is equal to $\sqrt{d_1^2 + d_2^2 + d_3^2}$.

consist in finding the *nearest* rotation matrix that satisfies the two aforementioned conditions exactly.

A naive way to restore the orthonormality of a noisy rotation matrix consists in applying the Gram-Schmidt process to its rows or columns. Despite its popularity due to its simple geometric interpretation, the result is rather arbitrary as it depends on the order in which the rows or the columns of the matrix are taken. This is why, whenever possible, the problem is solved by relying on the SVD of the noisy rotation matrix which computes its nearest rotation matrix according to the Frobenius norm. This paper focuses on the derivation of new closed-form formulas to obtain this result.

This paper is organized as follows. Section II briefly reviews some basic facts on the nearest exact rotation matrix to a given noisy rotation matrix in terms of the Frobenius norm. An exact closed-form formula for the computation of this rotation matrix is presented in Section III. This formula combines Smith's method for the computation of the required eigenvalues and Franca's method to avoid the explicit computation of the corresponding eigenvectors. The goal in Section IV is to introduce an alternative closed-form formula, involving only the four elementary arithmetic operations (addition, subtraction, multiplication, and division). This derivation is based on a heuristic argument for averaging quaternions which leads to an approximation to the optimum. Section V presents an exhaustive performance analysis of the the two closed-form formulas. Finally, Section VI summarizes the main contributions.

II. BACKGROUND

The first problem faced when trying to find the *nearest* proper orthogonal matrix, say $\hat{\mathbf{R}}$, to a noisy rotation matrix, say $\mathbf{R}$, is how to define a measure of closeness endowed with physical meaning. The Euclidean or Frobenius norm (denoted as $\|\cdot\|_F$) of $\mathbf{R}-\hat{\mathbf{R}}$ (*i.e.*, the square root of the sum of squares of the elements of $\mathbf{R}-\hat{\mathbf{R}}$) is commonly used to this end. The reasons for choosing this norm are multiple:

- Although it does not seem to be a reasonable measure of angular difference, it has a simple geometric interpretation (see Fig. 1).
- It leads to a closed-form formula. If we would use the 2-norm the closeness measure would be given by the largest singular value of $\mathbf{R}-\hat{\mathbf{R}}$ which is not easy to deal with [7].
- Using it, the solution is unique. If we would use the 2-norm, the solution is not necessarily unique [7].

Thus, using the Frobenius norm, our problem can be stated in algebraic terms as follows: given the noisy rotation matrix $\mathbf{R}$, the problem consists in finding the matrix $\hat{\mathbf{R}}$ that minimizes

$$\left\| \hat{\mathbf{R}}-\mathbf{R} \right\|_F^2 = \operatorname{trace}\left((\hat{\mathbf{R}}-\mathbf{R})(\hat{\mathbf{R}}-\mathbf{R})^T \right) = \rho_1^2 + \rho_2^2 + \rho_3^2, \quad (1)$$
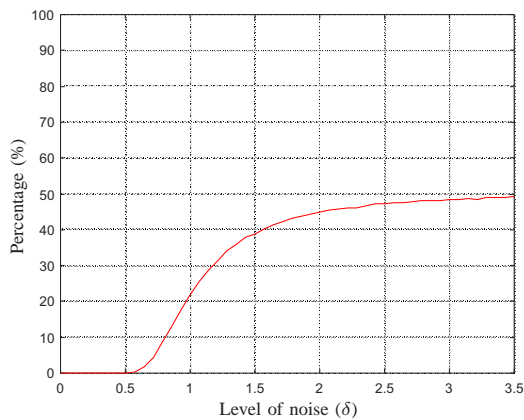
Fig. 2. Percentage of matrices with negative determinant as a function of the level of noise added to randomly generated rotation matrices.

subject to $\hat{\mathbf{R}}^T\hat{\mathbf{R}} = \mathbf{I}$, where ρ_i, $i = 1, 2, 3$, are the eigenvalues of $\hat{\mathbf{R}} - \mathbf{R}$. Using Lagrange multipliers, it can be proved that the optimal solution to this constrained optimization problem, in the case that $\mathbf{R}$ is not singular, is given by [8], [9]:

$$\hat{\mathbf{R}} = \mathbf{R}\left(\mathbf{R}^T\mathbf{R}\right)^{-\frac{1}{2}} = \mathbf{R}\left(\mathbf{I} + \mathbf{E}\right)^{-\frac{1}{2}}, \tag{2}$$

where

$$\mathbf{E} = \mathbf{R}^T\mathbf{R} - \mathbf{I} \tag{3}$$

can be seen as an error matrix.

Now, isolating $\mathbf{R}$ from (2), we have that

$$\mathbf{R} = \hat{\mathbf{R}}\mathbf{H}, \tag{4}$$

where $\mathbf{H} = (\mathbf{R}^T\mathbf{R})^{\frac{1}{2}}$. Equation (4) is called the polar decomposition of $\mathbf{R}$. The polar decomposition can be obtained from the SVD (Singular Value Decomposition) and vice versa. Indeed if $\mathbf{R} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ is an SVD then $\mathbf{R} = (\mathbf{U}\mathbf{V}^T)(\mathbf{V}\boldsymbol{\Sigma}\mathbf{V}^T) = \hat{\mathbf{R}}\mathbf{H}$ is a polar decomposition. Therefore, $\hat{\mathbf{R}} = \mathbf{U}\mathbf{V}^T$ [10].

It is easy to verify that $\hat{\mathbf{R}}$ thus obtained is orthonormal, *i.e.* $\hat{\mathbf{R}}^T\hat{\mathbf{R}} = \mathbf{I}$. However, there is no guarantee that $\det(\hat{\mathbf{R}}) = +1$. To represent a proper rotation, the orthonormal matrix $\hat{\mathbf{R}}$ has to satisfy this condition as well, otherwise it represents a reflection. There is no easy way to enforce this condition, and with highly noisy rotation matrices, we may have that $\det(\hat{\mathbf{R}}) = -1$.

If $\det(\mathbf{R}) < 0$, the nearest rotation matrix can be obtained as [11, p. 421]:

$$\mathbf{R} = \mathbf{U}\operatorname{diag}(1, 1, -1)\mathbf{V}^T. \tag{5}$$

This is a useful refinement that cannot be introduced in (2). As a consequence, this is an advantage of using SVD.

To assess the prevalence of the cases in which $\det(\mathbf{R}) < 0$, let us randomly generate 10^6 rotation matrices whose elements are contaminated with additive uncorrelated uniformly distributed noise in the interval $[-\delta, \delta]$. Then, if we plot the percentage of the resulting matrices whose determinant is negative as we vary δ from 0 to 3.5, we obtain the plot in Fig. 2. Inversions of signs arise for $\delta > 0.5$ which is an extremely high level of noise. Thus, for most practical applications, equation (2) can be directly used provided that we have an efficient way to compute $\left(\mathbf{R}^T\mathbf{R}\right)^{-\frac{1}{2}}$. Next, we show how to derive a closed-form formula for this computation that avoids the numerical computation of the matrix square root.

III. AN EXACT CLOSED-FORM FORMULA

If $\det(\mathbf{R}) \neq 0$, $\mathbf{A} = \mathbf{R}^T\mathbf{R}$ is symmetric and positive definitive. Then, it can be diagonalized as follows:

$$\mathbf{A} = \mathbf{V}^T \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \mathbf{V}, \tag{6}$$

where $\{\lambda_i\}$ is the set of non-negative real eigenvalues of $\mathbf{A}$. Then, it can be proved that [12]

$$\mathbf{A}^{-\frac{1}{2}} = \mathbf{V}^T \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{\lambda_3}} \end{pmatrix} \mathbf{V}. \tag{7}$$

The problem is thus essentially reduced to diagonalize $\mathbf{A}$. Although conceptually simple, the derivation of a closed-form formula based on this idea is full of unexpected difficulties. Let us start with the computation of $\{\lambda_i\}$.

A simple method to compute $\{\lambda_i\}$ can be found in [13]. Let us define $3m = \operatorname{trace}(\mathbf{A})$, $2q = \det(\mathbf{A} - m\mathbf{I})$, and let $6p$ equal to the sum of squares of the elements of $(\mathbf{A} - m\mathbf{I})$. Then, we have that

$$\lambda_1 = m + 2\sqrt{p}\cos\theta, \tag{8}$$
$$\lambda_2 = m + 2\sqrt{p}(\cos\theta + \sqrt{3}\sin\theta), \tag{9}$$
$$\lambda_3 = m + 2\sqrt{p}(\cos\theta - \sqrt{3}\sin\theta), \tag{10}$$

where

$$\theta = \frac{1}{3}\operatorname{atan2}\left(\sqrt{p^3 - q^2}, q\right). \tag{11}$$

Since the characteristic polynomial of $\mathbf{A}$ is a third-order polynomial, it is not surprising to recognize Cardano's formulas in (8)-(11). The term $p^3 - q^2$ corresponds to the discriminant of the characteristic polynomial of $\mathbf{A}$ [14]. When $\mathbf{A}$ has two equal eigenvalues, round-off errors might lead to a small negative value for this discriminant. Any implementation should consider this possibility. We here use atan2, the 2-argument arctangent, instead of the standard arctangent function, because it directly resolves the quadrant ambiguity by taking into account the sign of q (for $q > 0$, the solution must lie in the first quadrant, for $q < 0$ it must be located in the second). Moreover, $\operatorname{atan2}(x, 0)$ is well-defined contrarily to what happens with its single argument counterpart. Furthermore, observe that, whatever is the value returned by $\operatorname{atan2}(0, 0)$, we have that $\lambda_1 = \lambda_2 = \lambda_3 = m$ because $p = 0$.

Alternative formulations to the above one can be found, for example, in [15] and [16]. In [15], the case in which $p^3 - q^2$ is below a certain tolerance is treated separately. In [16], this term is expanded and simplified to improve numerical accuracy. Nevertheless, the implementation of these alternative formulations shows that they perform worse because they explicitly compute the coefficients of the characteristic polynomial of $\mathbf{A}$ thus introducing unnecessary operations.

Now, we should compute the eigenvectors of $\mathbf{A}$ (*i.e.*, the columns of $\mathbf{V}$). Nevertheless, for low levels of noise, the three eigenvalues are arbitrarily close to 1, but finding the eigenvectors of a multiple eigenvalue is numerically ill-conditioned [17]. Thus, it is better to avoid the explicit computation of these eigenvectors. To this end, two strategies can be found in the literature.

First, observe that $\mathbf{V}$ in (6) can also be seen as a rotation matrix which can be factorized using, for instance, XYZ Euler angles. Then, equation (6) can be expressed as

$$\mathbf{A} = \mathbf{R}_z(-\phi_3)\mathbf{R}_y(-\phi_2)\mathbf{R}_x(-\phi_1)$$
$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \mathbf{R}_x(\phi_1)\mathbf{R}_y(\phi_2)\mathbf{R}_z(\phi_3), \qquad (12)$$

which can be solved for ϕ_i. This approach is followed in [18] and [19]. Unfortunately, the resulting formulas become quite involved.

Alternatively, the approach presented in [15], which has received little attention, provides a simpler solution to the problem. Indeed, if we apply the Cayley-Hamilton theorem —which states that every square real matrix satisfies its own characteristic polynomial— to $\mathbf{A}^{\frac{1}{2}}$, we have that

$$\left(\mathbf{A}^{\frac{1}{2}} - \sqrt{\lambda_1}\mathbf{I}\right)\left(\mathbf{A}^{\frac{1}{2}} - \sqrt{\lambda_2}\mathbf{I}\right)\left(\mathbf{A}^{\frac{1}{2}} - \sqrt{\lambda_3}\mathbf{I}\right)$$
$$= \mathbf{A}^{\frac{3}{2}} - a_2\mathbf{A} + a_1\mathbf{A}^{\frac{1}{2}} - a_0\mathbf{I} = \mathbf{0}, \qquad (13)$$

where

$$a_2 = \sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3},$$
$$a_1 = \sqrt{\lambda_1\lambda_2} + \sqrt{\lambda_1\lambda_3} + \sqrt{\lambda_2\lambda_3},$$
$$a_0 = \sqrt{\lambda_1\lambda_2\lambda_3}.$$

Now, if we multiply (13) by $\mathbf{A}^{\frac{1}{2}}$ and we substitute $\mathbf{A}^{\frac{3}{2}}$ from (13) in the result, we obtain

$$\mathbf{A}^{\frac{1}{2}} = \frac{1}{a_2a_1 - a_0}\left(a_2a_0\mathbf{I} + (a_2^2 - a_1)\mathbf{A} - \mathbf{A}^2\right). \qquad (14)$$

Moreover, if we multiply (13) by $\mathbf{A}^{-\frac{1}{2}}$ we obtain

$$\mathbf{A}^{-\frac{1}{2}} = \frac{1}{a_0}\left(a_1\mathbf{I} - a_2\mathbf{A}^{\frac{1}{2}} + \mathbf{A}\right). \qquad (15)$$

Finally, substituting (14) in (15), rearranging terms, and substituting the result in (2), we have that

$$\hat{\mathbf{R}} = \mathbf{R}\mathbf{A}^{-\frac{1}{2}} = \mathbf{R}(b_2\mathbf{A}^2 - b_1\mathbf{A} + b_0\mathbf{I}), \qquad (16)$$

where

$$b_2 = \frac{a_2}{a_0(a_2a_1 - a_0)}, \qquad (17)$$
$$b_1 = \frac{a_0 + a_2(a_2^2 - 2a_1)}{a_0(a_2a_1 - a_0)}, \qquad (18)$$
$$b_0 = \frac{a_2a_1^2 - a_0(a_2^2 + a_1)}{a_0(a_2a_1 - a_0)}. \qquad (19)$$

Now, it it worth observing what happens for low levels of noise. In this case, $\lambda_i \approx 1$. Then, $a_2 \approx 3$, $a_1 \approx 3$, and $a_0 \approx 1$. As a consequence,

$$\hat{\mathbf{R}} \approx \frac{1}{8}\mathbf{R}(3\mathbf{R}^T\mathbf{R}\mathbf{R}^T\mathbf{R} - 10\mathbf{R}^T\mathbf{R} + 15\mathbf{I}). \qquad (20)$$

It is interesting to note that this formula coincides with the one obtained by computing the Taylor series expansion of $\mathbf{E}$ up to the third term and substituting the result in (2). This formula is actually used in [20] in an iterative algorithm intended to converge to $\hat{\mathbf{R}}$.

Equation (16) clearly fails if one of the eigenvalues is zero. Moreover, since the sign of $\det(\hat{\mathbf{R}})$ is the same as that of $\det(\mathbf{R})$, it actually provides a closed-form formula for the nearest orthogonal matrix, not the nearest rotation matrix. Thus, it is only valid if $\det(\mathbf{R}) > 0$, that is, if $\delta < 0.5$. Unfortunately, this bound is actually optimistic because the above formulation leads to a loss of accuracy when the eigenvalues differ significantly in magnitude [21], [22].
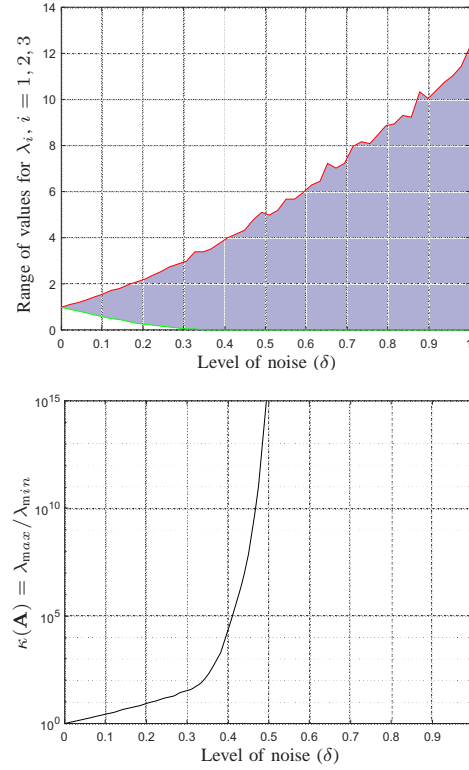
Fig. 3. Range of values for the eigenvalues of $\mathbf{A}$ (top), and its condition number (bottom) as a function of the added noise. When this noise is over 0.5, the lower bound for the eigenvalues is zero. As a consequence, the condition number goes to infinity.

For exact rotation matrices, the three eigenvalues of $\mathbf{A}$ are equal to 1. To see what happens with noisy matrices, we can randomly generate 10^6 rotation matrices whose elements, as in the experiment performed in Section II, are contaminated with additive uncorrelated uniformly distributed noise in the interval $[-\delta, \delta]$. Then, we compute the range of the eigenvalues of these 10^6 matrices. If we repeat this experiment for different values of δ, we obtain the region depicted in Fig. 3 (top). The ratio between the upper and lower bound gives us a measure of dispersion between the eigenvalues. Observe that, since $\mathbf{A}$ is symmetric, this ratio gives us an upper bound for its condition number. For $\delta > 0.5$, the condition number is clearly unbounded [Fig. 3 (bottom)]. As a rule of thumb, if the condition number $\kappa(\mathbf{A}) = 10^k$, then one may lose up to k digits of accuracy. Therefore, since the machine epsilon using single precision arithmetics is approximately 10^{-7}, the level of noise should not exceed, according to Fig. 3(bottom), 0.4 for accurate results. This bound is experimentally verified in Section V.

IV. AN APPROXIMATE CLOSED-FORM FORMULA

Euler's theorem of rigid-body rotations states that the orientation of a body after having undergone any sequence of rotations is equivalent to a single rotation of that body through an angle θ about an axis that we will represent by the unit vector $\hat{\mathbf{n}} = (n_x\ n_y\ n_z)^T$. A rotation matrix, expressed in terms of $\hat{\mathbf{n}}$ and θ, has the following form (see [6, p. 30] for an elementary deduction):

$$\mathbf{R}(\hat{\mathbf{n}}, \theta) =$$
$$\begin{pmatrix} c + n_x^2(1-c) & n_xn_y(1-c) - n_zs & n_xn_z(1-c) + n_ys \\ n_yn_x(1-c) + n_zs & c + n_y^2(1-c) & n_yn_z(1-c) - n_xs \\ n_zn_x(1-c) - n_ys & n_zn_y(1-c) + n_xs & c + n_z^2(1-c) \end{pmatrix}, \quad (21)$$

TABLE I
THE FOUR CONSISTENT SETS OF SIGNS FOR THE COMPONENTS OF $\hat{\mathbf{q}}$.

σ_1	$+$	$\text{sign}(r_{32} - r_{23})$	$\text{sign}(r_{13} - r_{31})$	$\text{sign}(r_{21} - r_{12})$
σ_2	$\text{sign}(r_{32} - r_{23})$	$+$	$\text{sign}(r_{21} + r_{12})$	$\text{sign}(r_{13} + r_{31})$
σ_3	$\text{sign}(r_{13} - r_{31})$	$\text{sign}(r_{21} + r_{12})$	$+$	$\text{sign}(r_{32} + r_{23})$
σ_4	$\text{sign}(r_{21} - r_{12})$	$\text{sign}(r_{13} + r_{31})$	$\text{sign}(r_{32} + r_{23})$	$+$

where $s = \sin\theta$ and $c = \cos\theta$. Now, if we introduce the following change of variables

$$q_0 = k\cos(\theta/2), \tag{22}$$

$$q_1 = k\,n_x\sin(\theta/2), \tag{23}$$

$$q_2 = k\,n_y\sin(\theta/2), \tag{24}$$

$$q_3 = k\,n_z\sin(\theta/2), \tag{25}$$

where k is an arbitrary non-null real number, then (21) can be rewritten as

$$\mathbf{R}(\mathbf{q}) = \frac{1}{q_0^2+q_1^2+q_2^2+q_3^2}\begin{pmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2(q_1q_2-q_0q_3) & 2(q_1q_3+q_0q_2) \\ 2(q_1q_2+q_0q_3) & q_0^2-q_1^2-q_2^2-q_3^2 & 2(q_2q_3-q_0q_1) \\ 2(q_1q_3-q_0q_2) & 2(q_2q_3+q_0q_1) & q_0^2-q_1^2-q_2^2+q_3^2 \end{pmatrix}. \tag{26}$$

where $\mathbf{q} = (q_0, q_1, q_2, q_3)^T$ is a vector whose elements can be seen as the components of a non-unit quaternion representing the same rotation as $\mathbf{R}$ [23]. Since equation (26) is independent from the value of k, this vector can actually be treated as a vector of homogeneous coordinates [24].

Now, given $\mathbf{R} = (r_{ij})_{1\leq i,j\leq 3}$, let us define the associated 4×4 symmetric matrix

$$\mathbf{U} = \frac{1}{4}\begin{pmatrix} r_{11}+r_{22}+r_{33}+1 & r_{32}-r_{23} & r_{13}-r_{31} & r_{21}-r_{12} \\ r_{32}-r_{23} & r_{11}-r_{22}-r_{33}+1 & r_{21}+r_{12} & r_{31}+r_{13} \\ r_{13}-r_{31} & r_{21}+r_{12} & r_{22}-r_{11}-r_{33}+1 & r_{32}+r_{23} \\ r_{21}-r_{12} & r_{31}+r_{13} & r_{32}+r_{23} & r_{33}-r_{11}-r_{22}+1 \end{pmatrix}. \tag{27}$$

It was proved in [25], and later independently rediscovered in [26] and [27], that the dominant eigenvector of $\mathbf{U}$ is the quaternion whose corresponding rotation matrix is equal to $\mathbf{R}(\mathbf{R}^T\mathbf{R})^{\frac{1}{2}}$ (the dominant eigenvector is the eigenvector associated with the eigenvalue whose absolute value is maximal). A closed-form formula is derived in [25] to obtain this dominant eigenvector, but it is far from trivial as it involves the computation of the roots of a fourth order polynomial. Alternatively, a numerical method is proposed in [27].

When $\mathbf{R}$ is not contaminated by noise, $\mathbf{U}$ is identically equal to (see, for example, [3] or [28]):

$$\mathbf{q}\mathbf{q}^T = \begin{pmatrix} q_0q_0 & q_0q_1 & q_0q_2 & q_0q_3 \\ q_1q_0 & q_1q_1 & q_1q_2 & q_1q_3 \\ q_2q_0 & q_2q_1 & q_2q_2 & q_2q_3 \\ q_3q_0 & q_3q_1 & q_3q_2 & q_3q_3 \end{pmatrix}. \tag{28}$$

Thus, for exact rotation matrices, all the columns of $\mathbf{U}$ are equal up to a scalar factor. The same applies to its rows as it is symmetric. For noisy rotation matrices, this is not longer true. Then, assuming that the elements of the rotation matrix are contaminated by uncorrelated noise, it is reasonable to average the column vectors of (27) in some way to get an estimation of $\mathbf{q}$.

The problem of averaging quaternions arises in many different applications [29]. The simple average (excluding constant factors as $\mathbf{q}$ is treated as a vector in homogeneous coordinates)

$$\hat{\mathbf{q}} = \sum_{i=1}^{4}\mathbf{u}_i, \tag{29}$$

where $\mathbf{u}_i$ denotes the i column of $\mathbf{U}$, has a subtle flaw. Since $\mathbf{u}_i$ and $-\mathbf{u}_i$ represent the same rotation (quaternions provide a double covering of the rotation group), changing the sign of any $\mathbf{u}_i$ should not change the average. Nevertheless, it is clear that (29) does not have this property.

Therefore, one possibility is to *homogenize* the signs of $\mathbf{u}_i$ prior to averaging them. A simple way to implement this idea reads as follows:

$$\hat{\mathbf{q}} = \sum_{i=1}^{4}\text{sign}(\mathbf{u}_j\cdot\mathbf{u}_i)\mathbf{u}_i, \tag{30}$$

where the dot operator stands for the scalar product, and $\mathbf{u}_j$ is chosen so that $\|\mathbf{u}_j\|^2 \geq \|\mathbf{u}_i\|^2$, $i = 1,\ldots,4$.

Summing up, the method we propose can be simply summarized as follows: given the noisy rotation matrix $\mathbf{R}$, we compute $\mathbf{U}$ using (27), then $\hat{\mathbf{q}}$ using (30), and finally $\mathbf{R}(\hat{\mathbf{q}})$ using (26). Observe that this method always returns a rotation matrix independently of the sign of $\det(\mathbf{R})$ by using only the four basic arithmetic operations.

The idea of orthonormalizing a noisy rotation matrix by converting it to quaternion form and then obtaining back the corresponding proper rotation matrix is not new. It is commonly implemented using the Shepperd-Markley method [30]. Very recently, the method presented in [28], based on Cayley's factorization [31], is shown to provide closer results to those obtained using the SVD than Shepperd-Markley method. A unified view of the method we have just proposed and the one resulting from Cayley's factorization is obtained by substituting the arithmetic mean given in (30) by the squared mean root. In this case, the result reads as follows:

$$\hat{\mathbf{q}} = (\sigma_1\|\mathbf{u}_1\|, \sigma_2\|\mathbf{u}_2\|, \sigma_3\|\mathbf{u}_3\|, \sigma_4\|\mathbf{u}_4\|)^T, \tag{31}$$

where the signs of the components of $\hat{\mathbf{q}}$, σ_i, have to be calculated separately. Since we have a global sign ambiguity, we can assume that the largest component of $\hat{\mathbf{q}}$ is positive and assign the remaining signs according to Table I.

In the next section, it is shown that the use of the arithmetic mean in (30) provides better results in the worst case than the use of the squared mean root in (31).

V. PERFORMANCE ANALYSIS

This paper has supplementary downloadable multimedia material which includes a C++ library implementing the methods described in this paper, and a main program needed to reproduce the analysis described below. The methods that have been analyzed are the following:

1) The new exact method resulting from applying the closed-form formula given by equation (16).
2) The new approximate method resulting from applying the closed-form formula given by equation (30).
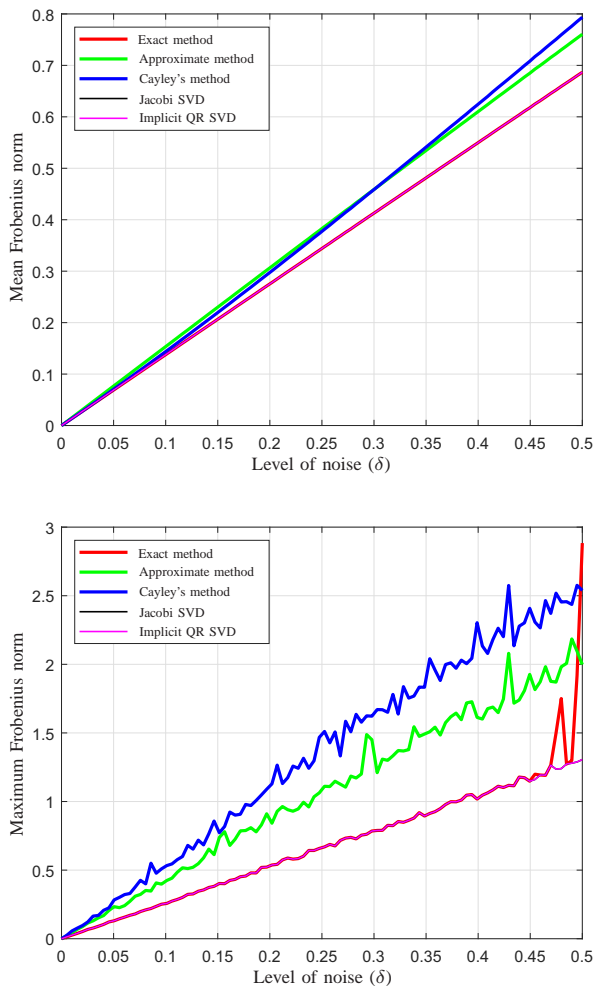
Fig. 4. Mean (top) and maximum (bottom) Frobenius norm between randomly generated noisy rotation matrices and the corresponding nearest rotation matrices obtained with the five compared methods implemented in single precision arithmetics. 10^6 random matrices are generated for each value of δ.
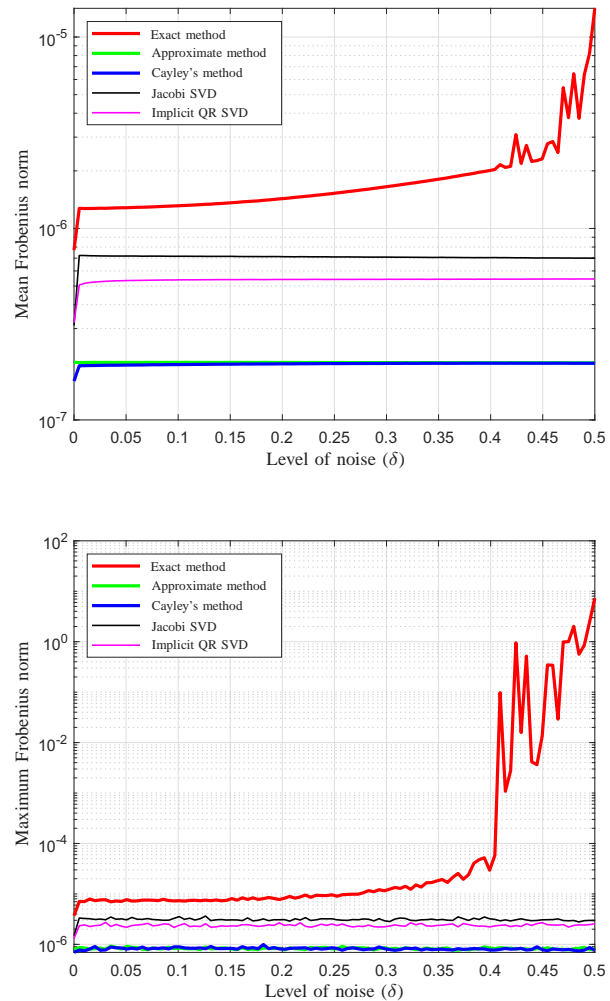

Fig. 5. Mean (top) and maximum (bottom) orthogonality error of the rotation matrices obtained with the five compared methods implemented in single precision arithmetics. 10^6 random matrices are generated for each value of δ.

3) The method given by equation (31) which results from using Cayley's factorization [28].
4) The Jacobi SVD implemented in the Eigen3 C++ library [32].
5) The Implicit QR SVD algorithm recently presented in [33]. This algorithm is specialized for $3{\times}3$ matrices and hence faster than the general Jacobi SVD.

To assess the performance of these methods, we have implemented the following procedure using single-precision arithmetic on a PC with a 4,2 GHz Intel Core i7 processor:

1) Generate 10^6 random quaternions using the algorithm detailed in [34], which permits to generate sets of points uniformly distributed on $\mathbb{S}^3$.
2) Convert these quaternions to rotation matrices whose elements are then contaminated with additive uncorrelated uniformly distributed noise in the interval $[-\delta, \delta]$.
3) Compute the nearest rotation matrices for these 10^6 noisy rotation matrices using each of the above methods.
4) Compute the maximum and the average Frobenius norm between the noisy matrices and the obtained matrices for each method.

If this procedure is repeated for values of δ ranging from 0 to 0.5, the plots in Fig. 4 are obtained. The curves for the mean Frobenius norm error obtained using the exact close-form formula, Jacobi SVD,

and Implicit QR SVD overlap as they all provide the optimal result. The curves for the maximum Frobenius norm error also overlap for $\delta < 0.45$. For larger values, as already predicted in Section III, the exact method exhibits numerical instabilities. After linear regression, the mean Frobenius norm error for the optimal results is found to be equal to 1.375δ. Despite the approximate method is based on an arithmetic average resulting from a heuristic argument, it performs better than Cayley's method. Its superiority is clear when observing the maximum error curves. After linear regression, the mean Frobenius norm error for this approximate method results is 1.526δ.

If the elements of the rotation matrix are contaminated with zero-mean Gaussian noise with variance equal to $\delta/2$, instead of uniformly distributed noise in the interval $[-\delta, \delta]$, the results do not differ essentially from the ones given above. The possibility of generating the corresponding curves is also included as an option in the downloadable multimedia material.

To asses the orthogonality errors of the obtained results, we have also computed the mean and the maximum of the Frobenius norm of $\hat{\mathbf{R}}\hat{\mathbf{R}}^T{-}\mathbf{I}$ for δ ranging from 0 to 0.5. The results are plotted in Fig. 5. The mean and the maximum orthogonality error curves for the approximate and Cayley's method overlap. They both provide the minimum orthogonality error because the last step in both methods

consists in applying equation (26) which returns an exact orthogonal matrix for any set of values for q_i, $i = 0, \ldots, 4$, provided that not all them are zero. The maximum orthogonality error curve for the exact method reveals that numerical instabilities arise for $\delta > 0.4$, as already observed. The orthogonality error is also the highest one for this method because it involves square roots and trigonometric functions, but it is still neglectable for most practical applications.

The average execution time of the five methods compared above is 0.13, 0.07, 0.03, 1.32, and 0.22 microseconds, respectively[1]. Thus, the most commonly used method –the Jacobi SVD method– is the computationally most expensive alternative. The Implicit QR SVD method provides an important improvement thanks to its particularization to 3×3 matrices. An extra improvement is obtained by using the exact closed-form formula presented in this paper. Finally, if we accept an approximation to the optimal solution (with is actually better than the one obtained using Cayley's method), a further improvement can be obtained using the closed-form formula introduced in this paper that involves neither trigonometric functions nor square roots.

VI. Conclusion

The problem of obtaining the nearest rotation matrix, in Frobenius norm, to a given noisy rotation matrix can be straightforwardly solved by computing the SVD of the noisy rotation matrix. Nevertheless, this approach requires higher computational resources than simpler closed-form solutions.

We have presented a compact exact closed-form formula which is numerically stable, for most practical problems, despite it involves the computation of the roots a third-order polynomial. Moreover, a new approximate method has also been presented. It was designed, based on a simple heuristic argument, to rely only on the four basic arithmetic operations. Nevertheless, to our surprise, it performs better than Cayley's method which, in turn, performs better than Shepperd-Markley method, the standard method used to orthonormalize a noisy rotation matrix by converting it to quaternion form and then obtaining back the corresponding proper rotation matrix. This approximate method is particularly useful for implementation in embedded micro-controllers with limited computational resources because it requires neither square root nor trigonometric computations.

In the light of the presented results, we can say that the use of SVD for solving the nearest rotation matrix problem can be advantageously substituted by closed-form formulations in most practical applications.

References

[1] D. Herbison-Evans and D.S. Richardson, "Control of round-off propagation in articulating the human figure," *Computer Graphics and Image Processing*, Vol. 17, No. 4, pp. 386-393, 1981.

[2] C. Rucker, "Integrating rotations using nonunit quaternions," *IEEE Robotics and Automation Letters*, Vol. 3, No. 4, pp. 2979-2986, 2018.

[3] F. Thomas, "Approaching dual quaternions from matrix algebra," *IEEE Transactions on Robotics*, Vol. 30, No. 5, pp. 1037-1048, 2014.

[4] D. Eberly, "Rotation representations and performance issues," web document, 2008. Available online at https://www.geometrictools.com/Documentation/RotationIssues.pdf, last accessed October 2019.

[5] H. Zhuang, Z.S. Roth, and R. Sudhakar, "Practical fusion algorithms for rotation matrices: A comparative study," *Journal of Robotic Systems*, Vol. 9, No. 7, pp. 915-931, 1992.

[6] R.P. Paul, *Robot Manipulators: Mathematics, Programming, and Control*, MIT Press Cambridge, MA, 1982.

[7] W. Kahan, "The nearest orthogonal or unitary matrix," University of California at Berkeley, 2011. Available online at ⟨https://people.eecs.berkeley.edu/~wkahan/Math128/NearestQ.pdf⟩, last accessed October 2019.

[8] C.R. Giardina, R. Bronson, and L. Wallen, "An optimal normalization scheme," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-11, No. 4, pp. 443-446, 1975.

[9] I. Bar-Itzhack, "Iterative optimal orthogonalization of the strapdown matrix," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-11, No. 1, pp. 30-37, 1975.

[10] J. Mao, "Optimal orthonormalization of the strapdown matrix by using singular value decomposition," *Computers and Mathematics with Applications*, Vol. 12, No. 3, pp. 353-362, 1986.

[11] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, *Robotics: Modelling, Planning and Control*, Springer Verlag, 2009.

[12] H.C. Schweinler and E.P. Wigner, "Orthogonalization methods," *Journal of Mathematical Physics*, Vol. 11, No. 5, pp. 1693-1694, 1970.

[13] O.K. Smith, "Eigenvalues of a symmetric 3×3 matrix," *Communications of the ACM*, Vol. 4, No. 4, p. 168, 1961.

[14] J.F. Blinn, "How to solve a cubic equation, part 1: The shape of the discriminant," *IEEE Computer Graphics and Applications*, Vol. 26, No. 3, pp. 84-93, 2006.

[15] L.P. Franca, "An algorithm to compute the square root of a 3×3 positive definite matrix," *Computers and Mathematics with Applications*, Vol. 18, No. 5, pp. 459-466, 1989.

[16] J. Kopp, "Efficient numerical diagonalization of Hermitian 3×3 matrices," *International Journal of Modern Physics C*, Vol. 19, No. 03, pp. 523-548, 2008.

[17] A. Ruhe, "Properties of a matrix with a very ill-conditioned eigenproblem," *Numerische Mathematik*, Vol. 15, No. 1, pp. 57-60, 1970.

[18] A.W. Bojanczyk and A. Lutoborski, "Computation of the Euler angles of a symmetric 3×3 matrix," *SIAM Journal of Matrix Analysis and Application*, Vol. 12, no.1, pp. 41-48, 1991.

[19] M.J. Kronenburg, "A method for fast diagonalization of a 2×2 or 3×3 real symmetric matrix," arXiv:1306.6291v4, 2015.

[20] M.A. Hasan, "Square-root free orthogonalization algorithms," *IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, pp. 3173-3176, 2009.

[21] W. Kahan, "To solve a real cubic equation," Lecture Notes for a Numerical Analysis Course, University of California, Berkeley, CA, 1986. Available online at ⟨https://people.eecs.berkeley.edu/~wkahan/Math128/Cubic.pdf⟩, last accessed October 2019.

[22] J.F. Blinn, "How to solve a cubic equation, part 5: Back to numerics," *IEEE Computer Graphics and Applications*, Vol. 27, No. 3, pp. 78-89, 2007.

[23] A.S. Hathaway, "Quaternions as numbers in four-dimensional space," *Bulletin of the American Mathematical Society*, Vol. 4, pp. 54-57, 1897.

[24] A. Purwar and Q.J. Ge, "On the effect of dual weights in Computer Aided Design of rational motions," *ASME Journal of Mechanical Design*, Vol. 127, No. 5, pp. 967-972, 2005.

[25] B.K.P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society A*, Vol. 4, No. 4, pp. 629-642, 1987.

[26] I.Y. Bar-Itzhack, "New method for extracting the quaternion from a rotation matrix," *Journal of Guidance Control and Dynamics*, Vol. 23, No. 6, pp. 1085-1087, 2000.

[27] N.J. Higham and V. Noferini, "An algorithm to compute the polar decomposition of a 3×3 matrix," *Numerical Algorithms*, Vol. 73, No. 2, pp. 349–369, 2016.

[28] S. Sarabandi, A. Pérez-Gracia, and F. Thomas, "On Cayley's factorization with an application to the orthonormalization of noisy rotation matrices," *Advances in Applied Clifford Algebras*, pp. 29-49, July 2019.

[29] F.L. Markley, Y. Cheng, J.L. Crassidis, and Y. Oshman, "Averaging quaternions," *Journal of Guidance, Control, and Dynamics*, Vol. 30, No. 4, pp. 1193-1196, 2007.

[30] F.L. Markley, "Unit quaternion from rotation matrix," *Journal of Guidance, Control and Dynamics*, Vol. 31, No. 2, pp. 440-442, 2008.

[31] A. Pérez-Gracia and F. Thomas, "On Cayley's factorization of 4D rotations and applications," *Advances in Applied Clifford Algebras*, Vol. 27, No. 1, pp. 523–538, 2017.

[32] G. Guennebaud, "Eigen: A C++ linear algebra library," *Eurographics/CGLibs*. Available online at http://downloads.tuxfamily.org/eigen/eigen_CGLibs_Giugno_Pisa_2013.pdf, last accessed October 2019.

[33] T. Gast, C. Fu, C. Jiang, and J. Teran, "Implicit-shifted symmetric QR singular value decomposition of 3×3 matrices," DTIC Report, University of California, Los Angeles, 2018.

[34] G. Marsaglia, "Choosing a point from the surface of a sphere," *Annals of Mathematical Statistics*, Vol. 43, pp. 645-646, 1972.

---

[1]The code for all methods were compiled with the '-O3' option to optimize the performance and preserve the IEEE compliance.