# The Distribution of Target Registration Error in Rigid-Body Point-Based Registration

J. Michael Fitzpatrick* and Jay B. West

***Abstract*—Guidance systems designed for neurosurgery, hip surgery, spine surgery and for approaches to other anatomy that is relatively rigid can use rigid-body transformations to accomplish image registration. These systems often rely on point-based registration to determine the transformation and many such systems use attached fiducial markers to establish accurate fiducial points for the registration, the points being established by some fiducial localization process. Accuracy is important to these systems, as is knowledge of the level of that accuracy. An advantage of marker-based systems, particularly those in which the markers are bone-implanted, is that registration error depends only on the fiducial localization and is, thus, to a large extent independent of the particular object being registered. Thus, it should be possible to predict the clinical accuracy of marker-based systems on the basis of experimental measurements made with phantoms or previous patients. For most registration tasks, the most important error measure is target registration error (TRE), which is the distance after registration between corresponding points not used in calculating the registration transform. In this paper, we derive an approximation to the distribution of TRE; this is an extension of previous work that gave the expected squared value of TRE. We show the distribution of the squared magnitude of TRE and that of the component of TRE in an arbitrary direction. Using numerical simulations, we show that our theoretical results are a close match to the simulated ones.**

***Index Terms*—Accuracy, error distribution, point-based, target registration error.**

## I. Introduction

THE point-based registration problem is as follows: given a set of homologous points in two spaces, find a transformation that brings the points into approximate alignment. In many cases, the appropriate transformations are rigid, consisting of translations and rotations. Medical applications abound in neurosurgery, for example, where the head can be treated as a rigid body [1]–[7]. The points, which we will call *fiducial* points, may be anatomical landmarks or may be produced artificially by means of attached markers. In the case that we address here, the spaces are three-dimensional (3-D) and may consist, for example, of two magnetic resonance (MR) volumes, a computed tomography volume and an MR volume or PETpositron emission tomography volume, or, in the case of image-guided neurosurgical applications, an image volume and the physical space of the operating room itself. The rigid-body, point-based image registration problem is typically defined to be the problem of finding the translation vector and rotation matrix that produces the least-squares fit of the corresponding fiducial points. The appropriate translation vector is simply the mean displacement between the two point sets. The problem of determining the rotation matrix can be easily reduced to the "Orthogonal Procrustes problem" [8], [9]. Peter Schönemann published the first solution to that problem in 1966 [9]. His solution was rediscovered independently in 1983 by Golub and van Loan [10] and again in 1987 by Arun *et al.* [11]. These latter solutions, unlike the former, employ the method of singular value decomposition (SVD), but they can easily be shown to be equivalent to Schönemann's solution [12].

The solution is unique, but can be expected to yield an imperfect registration in the presence of errors in locating the points. Maurer *et al.* [7], [13] suggested three useful measures of error for analyzing the accuracy of point-based registration methods (see Fig. 1).

1) *Fiducial localization error* (FLE), which is the error in locating the fiducial points.
2) *Fiducial registration error* (FRE), which is the root-mean- square distance between corresponding fiducial points after registration.
3) *Target registration error* (TRE), which is the distance between corresponding points other than the fiducial points after registration.

The term "target" is used to suggest that the points are directly associated with the reason for the registration. In medical applications, they are typically points within, or on the boundary of, lesions to be resected during surgery or regions of functional activity to be examined for diagnostic purposes.

Much work has been done [2], [3], [7], [13]–[17] using numerical simulations to investigate the properties of FRE and TRE. Unknown to many of those performing these simulations, Sibson [18] gave in 1979 an approximation to the distribution of FRE. In 1998, Fitzpatrick *et al.* derived an equation which allows calculation of an approximation to the root-mean-square (rms) value of TRE [19], [20] and agrees with published simulations. In 1999, West and Fitzpatrick presented an expression for an approximation to the *distribution* of TRE, as opposed to merely an rms value [21]. That approximation ignored an anisotropy in the distribution of rotational errors. In what follows, we give an approximation that accounts properly for this anisotropy and we give for the first time expressions for the distribution and rms value of a component of TRE in an arbitrary

Manuscript received June 12, 2000; revised June 16, 2001. This work was supported in part by the National Science Foundation under Grant BES-9802982, and in part by the Engineering and Physical Sciences Research Council of the U.K. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was D. Hawkes. *Asterisk indicates corresponding author.*

*J. M. Fitzpatrick is with the Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37235 USA (e-mail: jmf@vuse.vanderbilt.edu).

J. B. West is with the Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37235 USA

Publisher Item Identifier S 0278-0062(01)08663-3.

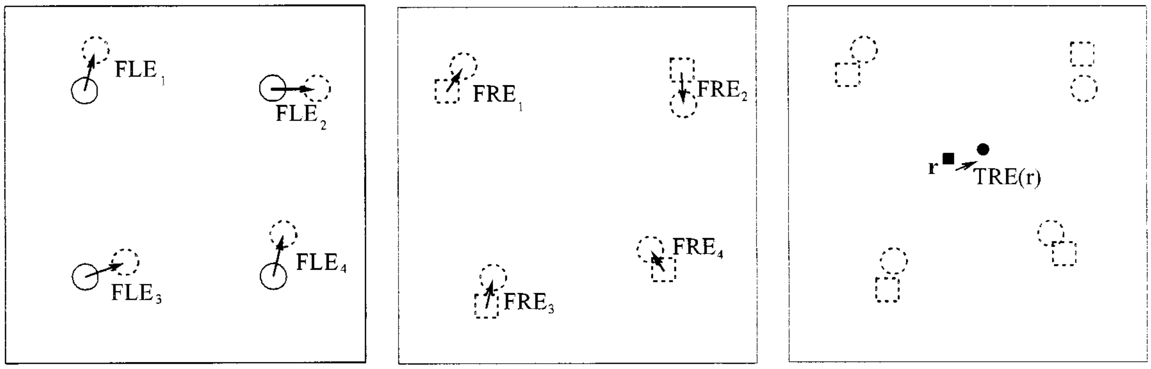


Fig. 1. Various types of registration error. The FLE measured at each fiducial is the distance between the true position (solid circles) and the measured position (dashed circles) of the fiducial. The FRE measured at each fiducial is the distance between the measured position of a fiducial in one space and its counterpart in the other space (dashed circle and dashed square), after registration. The TRE, measured at a point $\mathbf{r}$ relative to some given origin, is the distance after registration between the anatomical location (filled square) represented by $\mathbf{r}$ in one space and the corresponding anatomical point in the other space (filled circle).

direction, expressions that are of crucial importance in applications for which accuracy in a specific direction is critical.

The expected squared value of TRE that we presented previously gives an estimate of the average error, but with the extension of our theory to the calculation of distribution of error, we allow an appreciation of how large the extremal values may be, e.g., we give an answer to the question "*How large might the TRE be, say, once in every 20 times?*".

## II. The Model

We make a simplifying assumption in this work: that the FLE in one space is identically zero. This assumption does not generally hold in real registration problems, but the derivation may easily be extended to the case in which FLE is nonzero in both spaces. Using the result derived in 1985 by Langron and Collins [22], we can solve this two-space FLE problem by simply replacing the variance of FLE in our one-space model with the sum of the variances of FLE in each space. Hence, using $\langle\cdot\rangle$ to denote expected value, we replace $\langle \mathrm{FLE}_1^2\rangle + \langle \mathrm{FLE}_2^2\rangle$ by $\langle \mathrm{FLE}^2\rangle$, where $\mathrm{FLE}_1$ and $\mathrm{FLE}_2$ are the FLEs in the two spaces, respectively.

Here and for the remainder of the paper, we denote the number of fiducial points by $N$ and the dimension of the space containing the points by $K$. The value of $K$ is typically three in medical imaging applications. In general, we may write $X$ as the $N$-by $K$ matrix whose rows correspond to the position vectors of the fiducial points in one space and $Y$ as the $N$-by-$K$ matrix representing the fiducials in the other space. The registration problem is to find a $K$-by-$K$ orthogonal matrix, $R$ and a $K$-by-1 translation vector, $\mathbf{t}$, so that the points $R\mathbf{x}_i + \mathbf{t}$ are in optimal alignment with the corresponding points $\mathbf{y_i}$ in $Y$, where $\mathbf{x}_i$ and $\mathbf{y_i}$ are $K$-by-1 vectors and $i = 1, \ldots, N$. (Note that in this paper we use a nonbold font for scalars and matrices and a bold font for vectors. All vectors are column vectors unless adorned with a superscript $t$ to indicate transposition. Components of matrices and vectors, because they are scalars, are in a nonbold font.) By "optimal alignment," we mean that rms (FRE) is minimized, i.e., $R$ and $\mathbf{t}$ are chosen to minimize

$$G \equiv tr\left((Y - XR - \mathbf{1}_N\mathbf{t}^l)^l(Y - XR - \mathbf{1}_N\mathbf{t}^l)\right). \quad (1)$$

The solution that Schönemann found for the $R$ to minimize $G$ and rms (FRE) is

$$R = BA^t \quad (2)$$

where $ADB^t$ is the SVD of $\hat{Y}^t X$ and $\hat{Y} = Y - \mathbf{1}_N\mathbf{t}^t$. Thus

$$\hat{Y}^t X = ADB^t \quad (3)$$

where $A$, $D$, and $B$ are $K \times K$, $A$ and $B$ are orthogonal, $D$ is diagonal, and the elements of $D$ are nonnegative. This solution, which we will call the "SVD" solution, was an improvement over a solution published in 1952 by Green [23] that was based on the concept of the square root of a symmetric matrix and required that $X^t\hat{Y}$ be nonsingular, a restriction not required for the SVD solution.

In this paper, we assume that $X$ is related to $Y$ by a rigid-body transformation representing a reorientation of the rigid body to which the points are attached and an $N$-by-$K$ matrix $F$ of perturbations representing the FLE. We assume that the elements of $F$ are independent, zero-mean normal variables with equal variance, i.e., that the elements are independent $\mathcal{N}(0, \sigma)$ variables. Thus, FLE has the same distribution at each fiducial point and in each of the coordinate directions at every point. It should be noted that, because it is equal to the sum of the squares of $K$ independent, identically distributed normal variables, $\mathrm{FLE}^2$ is chi-square distributed. Furthermore, the component of localization error along any arbitrary direction is normally distributed. The above assumption about the distribution of elements of $F$ allows the use of a closed-form solution for the registration problem itself and as pointed out by Sibson [18], permits us to neglect the rigid body transformation relating $X$ and $Y$, as FRE and TRE are independent of this reorientation. We note that, under these assumptions, the variance of each element of $F$ is equal to $\langle \mathrm{FLE}^2\rangle/\mathrm{K}$.

We, thus, simplify the problem to that of registering $X$ to $Y = X + F$. As the choice of origin for $X$ is arbitrary, we choose the centroid of $X$ to be the origin.

## III. Derivation of the Distribution

We may adjust the $N$-by-$K$ matrix, $X$ of fiducial positions by choosing the origin of coordinates so that $X$ is centered, meaning that

$$X^t \mathbf{1}_N = 0. \tag{4}$$

We may write the Singular Value Decomposition of $X$ as

$$X = U\Lambda V^t, \tag{5}$$

where

- $U$ $N$-by-$N$ orthogonal matrix;
- $\Lambda$ $N$-by-$K$ diagonal matrix;
- $V$ $K$-by-$K$ orthogonal matrix.

The value of $V$ depends on our choice of orientation of coordinate axes. As with the choice of origin, the choice of axes is also arbitrary. We choose the orientation so that $V = I_K$, the $K$-by-$K$ identity matrix. This choice implies that the coordinate axes coincide with the principal axes of $X$. We now make explicit the relationship of this registration problem to perturbation theory: We write the fiducial set to which $X$ is to be registered as

$$Y = X + \epsilon F \tag{6}$$

where $\epsilon$ is a positive dimensionless constant whose value will be taken to be small enough to allow us to ignore higher order terms in $\epsilon$ as they arise in the derivations below.

### A. Choice of Translation

It is well known that the translation component, $\mathbf{t}$, of the registration that minimizes (1) for any $R$ is simply the translation that aligns the centroids of the two fiducial sets, i.e.,

$$\mathbf{t}^t = \frac{\epsilon \mathbf{1}_N^t F}{N}. \tag{7}$$

We write $\hat{F}$ as the demeaned version of $F$, i.e.,

$$\hat{F} = \frac{F - \mathbf{1}_N \mathbf{1}_N^t F}{N}. \tag{8}$$

Then with (1) and (7) we find that

$$\hat{Y} = Y - \mathbf{1}_N \mathbf{t}^t = X + \epsilon \hat{F}. \tag{9}$$

### B. Choice of Rotation

Using (9) in (1), we see that the rotation component of the optimal registration is the orthogonal matrix $R$ that optimally registers the points in $X$ to the corresponding points in $X + \epsilon \hat{F}$. We do not attempt to derive the exact form of this rotation matrix; instead, we express it as a power series in $\epsilon$

$$R = R^{(0)} + \epsilon R^{(1)} + \mathrm{O}(\epsilon^2) \tag{10}$$

### C. Expression for TRE

Our goal is to find an approximate expression for the distribution of TRE$(\mathbf{r})$. Following Sibson, we note that this distribution, like that of FRE$^2$, depends only on errors in localizing the fiducials, as opposed to gross motion between the two spaces. Also following Sibson, we will continue to treat the case in which the localization error is negligible in the "$X$" space. With these two assumptions we may use (6) The following expression results:

$$\mathbf{TRE}(\mathbf{r}) = R\mathbf{r} + \mathbf{t} - \mathbf{r} \tag{11}$$

where the **TRE** (in bold font) is the displacement vector, as opposed to the magnitude, of TRE. We should note that translation is first order in $\epsilon$, as can be seen from (7). Expanding the rotation matrix in $\epsilon$ and noting that $R = I$ when $\epsilon = 0$, we have $R - I = \epsilon R^{(1)} + \mathrm{O}(\epsilon^2)$. Hence, to first order in $\epsilon$ we may write

$$\mathbf{TRE}(\mathbf{r}) = R^{(1)}\mathbf{r} + \mathbf{t}. \tag{12}$$

We now wish to derive $R^{(1)}$. We begin by imposing the orthogonality requirement on $R$

$$\begin{aligned} R^t R = I &= \left(I + \epsilon R^{(1)t} + \mathrm{O}\left(\epsilon^2\right)\right)\left(I + \epsilon R^{(1)} + \mathrm{O}\left(\epsilon^2\right)\right) \\ &= I + \epsilon\left(R^{(1)t} + R^{(1)}\right) + \mathrm{O}\left(\epsilon^2\right). \end{aligned}$$

Therefore, $R^{(1)}$ is antisymmetric

$$R^{(1)t} = -R^{(1)}. \tag{13}$$

We note from(2) and (3) that for the optimal $R$, $\hat{Y}^t X R = ADA^t$, from which we see that

$$\hat{Y}^t X R = R^t X^t \hat{Y}. \tag{14}$$

(Note that we will make no other use of the SVD solution. In fact, in Schönemann's derivation, this symmetry is established before decomposition is employed. Thus, we do not need to know the complete solution in order to derive the first order approximation.) We use (6), for $Y$, but in order to account for translation, we must use demeaned versions of $X$ and $Y$. As discussed before (4), we have demeaned $X$ by our choice of origin. We demeaned $Y$ in (9), where we found that

$$\hat{Y} = X + \epsilon \hat{F}. \tag{15}$$

Writing (8) in component form gives

$$\hat{F}_{aj} = F_{aj} - \frac{1}{N}\sum_{b=1}^{N} F_{bj} \tag{16}$$

where $1 \leq a \leq N$ is used to label fiducial points and $1 \leq j \leq K$ is used to label coordinate axes. We now use (15), the expansion of $R$ and (13), in (14). The result is a series of equations for each power of $\epsilon$. The linear terms yield

$$X^t X R^{(1)} + R^{(1)} X^t X = X^t \hat{F} - \hat{F}^t X. \tag{17}$$

We wish to solve this equation for $R^{(1)}$. The solution is made difficult by the fact that $R^{(1)}$ occurs multiplied on both the right and left. Following Goodall [24], we perform SVD on $X$ to get $X = U\Lambda V^t$, where $U$ and $V$ are orthogonal and $\Lambda$ is diagonal. Our assumption that the elements of $F$ are identically distributed [see after (6)] assures isotropy in the perturbations. Thus, we can without loss of generality orient our coordinate system in any direction we choose. We pick the orientation to be along the principal axes of the distribution of fiducial points, which means that $V = I$. Thus, we have

$$X = U\Lambda. \tag{18}$$

[Note: Neither this reorientation nor the special positioning of the origin above is necessary to effect a solution to (17), nor for any part of the derivation that follows. However, they do reduce the complexity considerably and they can be easily undone at the end.] Employing (18) enables us to solve (17)

$$R_{ij}^{(1)} = \frac{\Lambda_{ii}Q_{ij} - \Lambda_{jj}Q_{ji}}{\Lambda_{ii}^2 + \Lambda_{jj}^2} \tag{19}$$

which is the result given by Goodall in 1991 [24] and similarly to Goodall we have defined

$$Q = U^t \hat{F}. \tag{20}$$

Because of the antisymmetry of $R^{(1)}$, we may rewrite (12) as

$$\mathbf{TRE}(\mathbf{r}) = \mathbf{\Omega} \times \mathbf{r} + \mathbf{t} \tag{21}$$

where

$$\mathbf{\Omega} = \left[R_{23}^{(1)} - R_{13}^{(1)} R_{12}^{(1)}\right]^T \tag{22}$$

in the case that the spatial dimension $K$ is equal to three. For the rest of this derivation, we will consider only the case $K = 3$. (We have previously derived [20] an expression for the expected value of $\mathrm{TRE}^2(\mathbf{r})$ that holds for general $K$.)

*1) Resolution Into Independent Components:* We wish to resolve $\mathbf{TRE}(\mathbf{r})$ into components along three orthogonal vectors in such a way that each of these three components of TRE is independent of the others. This task is made easier by taking advantage of two facts: 1) A linear combination of normal variables is itself a normal variable and 2) if two $\mathcal{N}(0, \sigma)$ variables are uncorrelated, then they must be independent [25]. With our neglect of higher order terms, any component of **TRE** is necessarily a linear combination of the normally distributed elements of $F$. Because the elements of $F$ have zero mean, the components of **TRE** have zero mean. Thus, our problem is reduced to resolving **TRE** into components that are uncorrelated.

We choose the first of these components to be a vector in the direction $\hat{\mathbf{r}}$, the unit vector in the radial direction and the second to be in the direction of a unit vector $\hat{\mathbf{v}}$, that is perpendicular to $\hat{\mathbf{r}}$. The third vector in this set must, thus, be $\hat{\mathbf{w}} = \hat{\mathbf{r}} \times \hat{\mathbf{v}}$. We denote the components of TRE in these three directions as $\mathrm{TRE}_\mathrm{r}$, $\mathrm{TRE}_\mathrm{v}$, and $\mathrm{TRE}_\mathrm{w}$, respectively. We know that

$$\mathrm{TRE}_\mathrm{r}(\mathbf{r}) = (\mathbf{\Omega} \times \mathbf{r} + \mathbf{t}) \cdot \hat{\mathbf{r}} = \mathbf{t} \cdot \hat{\mathbf{r}}. \tag{23}$$

Similarly, we have that

$$\mathrm{TRE}_\mathrm{v}(\mathbf{r}) = (\mathbf{\Omega} \times \mathbf{r} + \mathbf{t}) \cdot \hat{\mathbf{v}} = (\mathbf{\Omega} \times \mathbf{r}) \cdot \hat{\mathbf{v}} + \mathbf{t} \cdot \hat{\mathbf{v}} \tag{24}$$

and

$$\begin{aligned} \mathrm{TRE}_\mathrm{w}(\mathbf{r}) &= (\mathbf{\Omega} \times \mathbf{r} + \mathbf{t}) \cdot (\hat{\mathbf{r}} \times \hat{\mathbf{v}}) \\ &= ((\mathbf{r} \cdot \hat{\mathbf{v}})\mathbf{r} - (\hat{\mathbf{r}} \cdot \mathbf{r})\hat{\mathbf{v}}) \cdot \mathbf{\Omega} + \mathbf{t} \cdot (\hat{\mathbf{r}} \times \hat{\mathbf{v}}). \end{aligned} \tag{25}$$

Given that $\hat{\mathbf{v}}$ is constrained to be orthogonal to $\hat{\mathbf{r}}$, this simplifies to

$$\mathrm{TRE}_\mathrm{w}(\mathbf{r}) = -r(\hat{\mathbf{v}} \cdot \mathbf{\Omega}) + \mathbf{t} \cdot (\hat{\mathbf{r}} \times \hat{\mathbf{v}}) \tag{26}$$

where $r$ is the magnitude of the vector $\mathbf{r}$. In this way, we may write

$$\mathbf{TRE}(\mathbf{r}) = \mathrm{TRE}_\mathrm{r}\hat{\mathbf{r}} + \mathrm{TRE}_\mathrm{v}\hat{\mathbf{v}} + \mathrm{TRE}_\mathrm{w}\hat{\mathbf{w}} \tag{27}$$

and we wish to choose $\hat{\mathbf{v}}$ so that

$$\hat{\mathbf{r}} \cdot \hat{\mathbf{v}} = 0 \tag{28}$$

and

$$\langle \mathrm{TRE}_\mathrm{r}\mathrm{TRE}_\mathrm{v} \rangle = \langle \mathrm{TRE}_\mathrm{r}\mathrm{TRE}_\mathrm{w} \rangle = \langle \mathrm{TRE}_\mathrm{v}\mathrm{TRE}_\mathrm{w} \rangle = 0. \tag{29}$$

*2) Translational and Rotational Cross Correlations:* In order to derive $\hat{\mathbf{v}}$, we first need expressions for $\langle t_i t_j \rangle$, $\langle \Omega_i t_j \rangle$ and $\langle \Omega_i \Omega_j \rangle$. From (7) we have that

$$t_i = \frac{1}{N} \sum_{a=1}^{N} F_{ai}. \tag{30}$$

We know that

$$\langle t_i t_j \rangle = \frac{1}{N^2} \left\langle \sum_{a=1}^{N} \sum_{b=1}^{N} F_{ai} F_{bj} \right\rangle \tag{31}$$

which may be written as

$$\frac{1}{N^2} \left\langle \sum_{a=1}^{N} F_{ai} F_{aj} \right\rangle \tag{32}$$

because distinct elements of $F$ are independent. Recalling that each element of $F$ is defined to be a zero-mean, normally distributed random variable with variance $\sigma^2$, we have that

$$\langle t_i t_j \rangle = \frac{\sigma^2}{N} \delta_{ij} \tag{33}$$

where $\delta$ is the Kronecker delta function defined as $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$.

From (19) and (22) we have that

$$\Omega_j = \left( \frac{\Lambda_{kk} Q_{kl} - \Lambda_{ll} Q_{lk}}{\Lambda_{kk}^2 + \Lambda_{ll}^2} \right) \tag{34}$$

where $\{j, k, l\} = \{1, 2, 3\}, \{2, 3, 1\}$, or $\{3, 1, 2\}$. Using (20) to expand the elements of $Q$, we have that

$$\langle t_i \Omega_j \rangle = \frac{1}{N} \left\langle \sum_{a=1}^{N} F_{ai} \left( \frac{\sum_{b=1}^{N} (\Lambda_{kk} U_{bk} \hat{F}_{bl} - \Lambda_{ll} U_{bl} \hat{F}_{bk})}{\Lambda_{kk}^2 + \Lambda_{ll}^2} \right) \right\rangle. \tag{35}$$

Using (16) to expand $\hat{F}$ in terms of $F$, we have (36) shown at the bottom of the next page. Using independence of distinct elements of $F$, this may be simplified to (37) shown at the bottom of the page. Clearly

$$\sum_{a=1}^{N}\sum_{b=1}^{N} U_{bk}\delta_{ab} - \frac{1}{N}\sum_{a=1}^{N}\sum_{b=1}^{N}\sum_{c=1}^{N}\delta_{ac}U_{bk} = \sum_{a=1}^{N} U_{ak} - \frac{1}{N}\sum_{a=1}^{N}\sum_{b=1}^{N} U_{bk} = 0. \quad (38)$$

From this we may deduce that the expected product of any $t_i$ and $\Omega_j$ is zero, *i.e.*, they are uncorrelated. Finally, we have products of the form $\langle \Omega_i \Omega_j \rangle$. Any such product may be expressed as

$$\left\langle \left( \frac{\Lambda_{aa}Q_{ab} - \Lambda_{bb}Q_{ba}}{\Lambda_{aa}^2 + \Lambda_{bb}^2} \right) \left( \frac{\Lambda_{cc}Q_{cd} - \Lambda_{cc}Q_{dc}}{\Lambda_{cc}^2 + \Lambda_{dd}^2} \right) \right\rangle . \quad (39)$$

From the definition of $\hat{F}$, we have that

$$\begin{aligned} \left\langle \hat{F}_{ab}\hat{F}_{cd} \right\rangle &= \left\langle \left( F_{ab} - \frac{1}{N}\sum_{i=1}^{N} F_{ib} \right) \left( F_{cd} - \frac{1}{N}\sum_{i=1}^{N} F_{id} \right) \right\rangle \\ &= \sigma^2 \left( \delta_{ac}\delta_{bd} - \frac{1}{N}\delta_{bd} - \frac{1}{N}\delta_{bd} + \frac{1}{N}\delta_{bd} \right) \\ &= \sigma^2 \delta_{bd} \left( \delta_{ac} - \frac{1}{N} \right) \end{aligned}$$

from which we may deduce that

$$\begin{aligned} \langle \Lambda_{aa}Q_{ab}\Lambda_{cc}Q_{cd} \rangle &= \left\langle \sum_{i=1}^{N}\sum_{j=1}^{N} \Lambda_{aa}\Lambda_{cc}U_{ia}U_{jc}\hat{F}_{ib}\hat{F}_{jd} \right\rangle \\ &= \sigma^2 \delta_{bd} \sum_{i=1}^{N}\sum_{j=1}^{N} \Lambda_{aa}\Lambda_{cc}U_{ia}U_{jc} \left( \delta_{ij} - \frac{1}{N} \right). \end{aligned}$$

By orthogonality of $U$, we know that

$$\sum_{i=1}^{N} U_{ia}U_{ic} = U^T U_{ac} = \delta_{ac}. \quad (40)$$

Because $X = U\Lambda$, we also have that

$$\begin{aligned} \sum_{i=1}^{N}\sum_{j=1}^{N} \Lambda_{aa}\Lambda_{cc}U_{ia}U_{jc} &= \sum_{i=1}^{N}\sum_{j=1}^{N} X_{ia}X_{jc} \\ &= \left( \sum_{i=1}^{N} X_{ia} \right) \left( \sum_{i=1}^{N} X_{ic} \right) \\ &= 0. \end{aligned} \quad (41)$$

This gives us the result that

$$\langle \Lambda_{aa}Q_{ab}\Lambda_{cc}Q_{cd} \rangle = \sigma^2 \delta_{ac}\delta_{bd}. \quad (42)$$

Using this result in (39) gives the equation shown at the bottom of the next page. Clearly this is only nonzero when $a = c$ and $b = d$, or when $b = c$ and $a = d$. From the definition of $\mathbf{\Omega}$ in (22), we can see that this happens if and only if $i = j$. Thus, $\langle \Omega_i \Omega_j \rangle = 0$ for $i \neq j$. We denote $\langle \Omega_i^2 \rangle$ as $\omega_i^2$.

*3) Choice of Direction for* $\hat{\mathbf{v}}$*:* Having proved that the expected values of the products of components of $\mathbf{\Omega}$ and $\mathbf{t}$ are all zero, we return to our derivation of $\hat{\mathbf{v}}$. We wish to ensure that the component of TRE along each of our set of vectors is uncorrelated with the others. It happens that the radial component $\mathrm{TRE_r}$ is uncorrelated with $\mathrm{TRE_v}$ and $\mathrm{TRE_w}$ as a consequence of orthogonality. For example

$$\begin{aligned} \langle \mathrm{TRE_r TRE_v} \rangle &= \langle (\mathbf{t}\cdot\hat{\mathbf{r}})\left((\mathbf{\Omega}\times\mathbf{r})\cdot\hat{\mathbf{v}} + \mathbf{t}\cdot\hat{\mathbf{v}}\right) \rangle = \langle (\mathbf{t}\cdot\hat{\mathbf{v}})(\mathbf{t}\cdot\hat{\mathbf{r}}) \rangle \\ &= \left\langle \sum_{i=1}^{3}\sum_{j=1}^{3} t_i t_j \hat{v}_i \hat{r}_j \right\rangle = \sigma^2 \hat{\mathbf{v}}\cdot\hat{\mathbf{r}} = 0 \end{aligned}$$

where we have used uncorrelation of elements of $\mathbf{\Omega}$ and $\mathbf{t}$. Thus, our choice of $\hat{\mathbf{v}}$ is determined by the condition that $\langle \mathrm{TRE_v TRE_w} \rangle$ be equal to zero. We have that

$$\langle \mathrm{TRE_v TRE_w} \rangle = \langle ((\mathbf{\Omega}\times\mathbf{r})\cdot\hat{\mathbf{v}} + \mathbf{t}\cdot\hat{\mathbf{v}})(-\mathrm{r}\mathbf{\Omega}\cdot\hat{\mathbf{v}}) + (\mathbf{t}\cdot(\hat{\mathbf{r}}\times\hat{\mathbf{v}})) \rangle \quad (43)$$

but recalling that the elements of $\mathbf{t}$ and $\mathbf{\Omega}$ are uncorrelated, as are distinct elements of $\mathbf{t}$, this reduces to

$$\langle \mathrm{TRE_v TRE_w} \rangle = \langle -\mathrm{r}((\mathbf{\Omega}\times\mathbf{r})\cdot\hat{\mathbf{v}})(\mathbf{\Omega}\cdot\hat{\mathbf{v}}) \rangle. \quad (44)$$

---

$$\langle t_i \Omega_j \rangle = \frac{1}{N} \left\langle \sum_{a=1}^{N} F_{ai} \left( \frac{\sum_{b=1}^{N} \left( \Lambda_{kk}U_{bk} \left( F_{bl} - \frac{1}{N}\sum_{c=1}^{N} F_{cl} \right) - \Lambda_{ll}U_{bl} \left( F_{bk} - \frac{1}{N}\sum_{c=1}^{N} F_{ck} \right) \right)}{\Lambda_{kk}^2 + \Lambda_{ll}^2} \right) \right\rangle \quad (36)$$

---

$$\langle t_i \Omega_j \rangle = \frac{\sigma^2}{N} \sum_{a=1}^{N} \left( \frac{\sum_{b=1}^{N} \Lambda_{kk}U_{bk} \left( \delta_{ab}\delta_{il} - \frac{1}{N}\sum_{c=1}^{N} \delta_{ac}\delta_{il} \right) - \Lambda_{ll}U_{bl} \left( \delta_{ab}\delta_{ik} - \frac{1}{N}\sum_{c=1}^{N} \delta_{ac}\delta_{ik} \right)}{\Lambda_{kk}^2 + \Lambda_{ll}^2} \right) \quad (37)$$

Because the left-hand side of the above equation is to be set to zero, the equation may be multiplied throughout by a constant. Instead of finding $\hat{\mathbf{v}}$, we first find a vector $\mathbf{v}$ of arbitrary length in the same direction as $\hat{\mathbf{v}}$. We write the three components of $\mathbf{r}$ as $x$, $y$, and $z$, the corresponding components of $\mathbf{v}$ as $\upsilon_x$, $\upsilon_y$, and $\upsilon_z$ and those of $\mathbf{\Omega}$ as $\Omega_x$, $\Omega_y$, and $\Omega_z$. We require that

$$\langle((\mathbf{\Omega} \times \mathbf{r}) \cdot \mathbf{v})(\mathbf{\Omega} \cdot \mathbf{v})\rangle = 0. \tag{45}$$

Expanding this in the components of $\mathbf{v}$, $\mathbf{\Omega}$, and $\mathbf{r}$ gives

$$\begin{aligned}\langle((\Omega_y z - \Omega_z y)\upsilon_x + (\Omega_z x - \Omega_x z)\upsilon_y \\ +(\Omega_x y - \Omega_y x)\upsilon_z)(\Omega_x \upsilon_x + \Omega_y \upsilon_y + \Omega_z \upsilon_z)\rangle = 0.\end{aligned} \tag{46}$$

Recalling that $\langle \Omega_i \Omega_j \rangle = \omega_i^2 \delta_{ij}$, (46) reduces to

$$z\upsilon_x\upsilon_y\left(\omega_y^2 - \omega_x^2\right) + y\upsilon_x\upsilon_z\left(\omega_x^2 - \omega_z^2\right) + x\upsilon_y\upsilon_z\left(\omega_z^2 - \omega_y^2\right) = 0. \tag{47}$$

The other equation that must be satisfied by $\mathbf{v}$ is

$$\mathbf{v} \cdot \mathbf{r} = x\upsilon_x + y\upsilon_y + z\upsilon_z = 0. \tag{48}$$

We first treat the special case in which at least one component of $\mathbf{r}$ is equal to zero. We assume, without loss of generality, that $x = 0$. Then (47) and (48) reduce to

$$z\upsilon_x\upsilon_y\left(\omega_y^2 - \omega_x^2\right) + y\upsilon_x\upsilon_z\left(\omega_x^2 - \omega_z^2\right) = 0 \tag{49}$$

and

$$y\upsilon_y + z\upsilon_z = 0 \tag{50}$$

respectively. There are two ways of satisfying the first equation: by setting $\upsilon_x = 0$ or by setting $\upsilon_y = \upsilon_z = 0$. As this equation is quadratic in $\upsilon_x, \upsilon_y$, and $\upsilon_z$, these must be the only two solutions. If we choose to set $\upsilon_x = 0$, there is a line of possible solutions for $\upsilon_y$ and $\upsilon_z$. We may choose any solution from the line; we then normalize $\mathbf{v}$ to produce $\hat{\mathbf{v}}$. If we choose to set $\upsilon_y = \upsilon_z = 0$, clearly $\hat{\upsilon}_x = 1$. We can see by inspection that these solutions are orthogonal to each other and to $\hat{\mathbf{r}}$. Thus, we have proved that they are the $\hat{\mathbf{v}}$ and $\hat{\mathbf{w}}$ that we seek.

We should also note that if, for any $i \neq j$, $\omega_i^2 = \omega_j^2$, then there are two simple solutions to (47). For example, if $\omega_x^2 = \omega_y^2$, (47) may be satisfied by setting $\upsilon_z = 0$ or by setting $\upsilon_x = \upsilon_y = 0$. As in the case described above, these two solutions may be easily shown to be $\hat{\mathbf{v}}$ and $\hat{\mathbf{w}}$.

We now address the general case in which no component of $\mathbf{r}$ is zero and all the $\omega_i^2$ are distinct. From (48), we have that

$$y\upsilon_y = -z\upsilon_z - x\upsilon_x \tag{51}$$

and multiplying (47) by $y$ and using the above substitution gives

$$\begin{aligned}z\upsilon_x(x\upsilon_x + z\upsilon_z)\left(\omega_x^2 - \omega_y^2\right) + y^2\upsilon_x\upsilon_z\left(\omega_x^2 - \omega_z^2\right) \\ +x\upsilon_z(x\upsilon_x + z\upsilon_z)\left(\omega_y^2 - \omega_z^2\right) = 0.\end{aligned} \tag{52}$$

Rearranging in powers of $\upsilon_x$, this may be rewritten as

$$\begin{aligned}xz\left(\omega_x^2 - \omega_y^2\right)\upsilon_x^2 + \upsilon_x\upsilon_z\left(\left(y^2 + z^2\right)\left(\omega_x^2 - \omega_y^2\right)\right. \\ \left. + \left(x^2 + y^2\right)\left(\omega_y^2 - \omega_z^2\right)\right)\end{aligned} \tag{53}$$

$$xz\left(\omega_y^2 - \omega_z^2\right)\upsilon_z^2 = 0. \tag{54}$$

We may choose to solve for either $\upsilon_x$ or $\upsilon_z$; we choose to solve for $\upsilon_x$ and we wish to check whether (54) always has real roots. Writing the equation as $A\upsilon_x^2 + B\upsilon_x + C = 0$, the condition for real roots is that $B^2 > 4AC$. In this case, we have that

$$B^2 = \upsilon_z^2\left(\left(y^2 + z^2\right)\left(\omega_x^2 - \omega_y^2\right) + \left(x^2 + y^2\right)\left(\omega_y^2 - \omega_z^2\right)\right)^2 \tag{55}$$

and

$$\begin{aligned}4AC &= 4x^2z^2\upsilon_z^2\left(\omega_x^2 - \omega_y^2\right)\left(\omega_y^2 - \omega_z^2\right) \\ &= \upsilon_z^2\left(z^2\left(\omega_x^2 - \omega_y^2\right) + y^2\left(\omega_x^2 - \omega_z^2\right) + x^2\left(\omega_y^2 - \omega_z^2\right)\right)^2.\end{aligned}$$

If $4AC < 0$, clearly $B^2 > 4AC$ and the roots are real. Otherwise, we know that

$$\begin{aligned}4AC &< 4\upsilon_z^2(x^2 + y^2)(y^2 + z^2)\left(\omega_x^2 - \omega_y^2\right)\left(\omega_y^2 - \omega_z^2\right) \\ &= B^2 - \upsilon_z^2\left(\left(y^2 + z^2\right)\left(\omega_x^2 - \omega_y^2\right)\right. \\ &\quad \left. -(x^2 + y^2)\left(\omega_y^2 - \omega_z^2\right)\right)^2\end{aligned} \tag{56}$$

hence

$$\begin{aligned}B^2 - 4AC > \upsilon_z^2\left(\left(y^2 + z^2\right)\left(\omega_x^2 - \omega_y^2\right)\right. \\ \left. -(x^2 + y^2)\left(\omega_y^2 - \omega_z^2\right)\right)^2 > 0\end{aligned} \tag{57}$$

so the roots are real. As $\mathbf{v}$ is of arbitrary length, we set $\upsilon_z = 1$ and solve for $\upsilon_x$. Using the quadratic formula, we have (58)

$$\begin{aligned}\langle \Omega_i \Omega_j \rangle &= \left\langle \left(\frac{\Lambda_{aa}Q_{ab} - \Lambda_{bb}Q_{ba}}{\Lambda_{aa}^2 + \Lambda_{bb}^2}\right)\left(\frac{\Lambda_{cc}Q_{cd} - \Lambda_{dd}Q_{dc}}{\Lambda_{cc}^2 + \Lambda_{dd}^2}\right)\right\rangle \\ &= \frac{\sigma^2(\Lambda_{aa}\Lambda_{cc}\delta_{ac}\delta_{bd} - \Lambda_{aa}\Lambda_{dd}\delta_{ad}\delta_{bc} - \Lambda_{bb}\Lambda_{cc}\delta_{bc}\delta_{ad} + \Lambda_{bb}\Lambda_{dd}\delta_{bd}\delta_{ac})}{(\Lambda_{aa}^2 + \Lambda_{bb}^2)^2} \\ &= \frac{\sigma^2(\delta_{ac}\delta_{bd} - \delta_{bc}\delta_{ad})}{\Lambda_{aa}^2 + \Lambda_{bb}^2}\end{aligned}$$

shown at the bottom of the page. This gives two possible solutions for $\upsilon_x$ and for each solution we may derive a corresponding value for $\upsilon_y$ from (51). We label these as $\upsilon_x^+$, $\upsilon_y^+$, and $\upsilon_x^-$, $\upsilon_y^-$ and the vectors formed from them as $\mathbf{v}^+$ and $\mathbf{v}^-$. We have ensured that these two possible solutions for $\mathbf{v}$ are orthogonal to $\mathbf{r}$ and that their components of TRE are uncorrelated with each other. It remains only to show that the two possible values of $\mathbf{v}$ are orthogonal in order to complete the proof that they are the $\mathbf{v}$ and $\mathbf{w}$ that we seek. We have that

$$\begin{aligned}\mathbf{v}^+\cdot\mathbf{v}^- =& \upsilon_x^+\upsilon_x^- + \upsilon_y^+\upsilon_y^- + \upsilon_z^+\upsilon_z^- \\ =& \upsilon_x^+\upsilon_x^- + \frac{(x\upsilon_x^+ + z\upsilon_z^+)(x\upsilon_x^- + z\upsilon_z^-)}{y^2} \\ &+ \upsilon_z^+\upsilon_z^-.\end{aligned} \tag{59}$$

Using the fact that $\upsilon_z^+ = \upsilon_z^- = 1$ and applying to (54) the formulae for sums and products of roots of a quadratic equation, this may be rewritten as

$$\begin{aligned}\mathbf{v}^+\cdot\mathbf{v}^- =& \frac{(x^2+y^2)\left(\omega_y^2-\omega_z^2\right)}{y^2\left(\omega_x^2-\omega_y^2\right)} \\ &+ \frac{(y^2+z^2)\left(\omega_y^2-\omega_x^2\right)+(x^2+y^2)\left(\omega_z^2-\omega_y^2\right)}{y^2\left(\omega_x^2-\omega_y^2\right)} \\ &+ \frac{(y^2+z^2)\left(\omega_x^2-\omega_y^2\right)}{y^2\left(\omega_x^2-\omega_y^2\right)}\end{aligned} \tag{60}$$

which can be seen to sum to zero. Thus, the two solutions of (54) lead to $\mathbf{v}$ and $\mathbf{w}$.

### D. Expected Values

We now derive the expected value of the square of the magnitude of the total displacement, $\langle \mathrm{TRE}^2\rangle$ and also that of the square of the component of **TRE** in any arbitrary direction. First we derive the variances, $\sigma_r^2$, $\sigma_v^2$, and $\sigma_w^2$, of each of the components of **TRE**. Because their means are zero, the variances of these components are equal to their mean squared values. Thus, from (23) we have that

$$\sigma_r^2 = \left\langle \mathrm{TRE}_\mathrm{r}^2\right\rangle = \left\langle \sum_{i=1}^3\sum_{j=1}^3 \mathrm{t}_j\mathrm{t}_j\right\rangle = \left\langle\sum_{i=1}^3 \mathrm{t}_i\mathrm{t}_i\right\rangle = \frac{\sigma^2}{N}. \tag{61}$$

Similarly, from (24) and (26)

$$\sigma_v^2 = \left\langle \mathrm{TRE}_\mathrm{v}^2\right\rangle = \left\langle \left(\sum_{i=1}^3 \mathrm{w}_i\Omega_i + \mathrm{t}_i\hat{\upsilon}_i\right)^2\right\rangle \tag{62}$$

which may be rewritten as

$$\sigma_v^2 = r^2\omega_w^2 + \frac{\sigma^2}{N} \tag{63}$$

and

$$\sigma_w^2 = \left\langle \mathrm{TRE}_\mathrm{w}^2\right\rangle = \left\langle \left(\sum_{i=1}^3 -\mathrm{r}\hat{\mathrm{v}}_i\Omega_i + \mathrm{t}_i\hat{\mathrm{v}}_i\right)^2\right\rangle = \mathrm{r}^2\omega_v^2 + \frac{\sigma^2}{N}. \tag{64}$$

Using these expressions for the variance of the components, we have that

$$\left\langle \mathrm{TRE}^2(\mathbf{r})\right\rangle = \left\langle \mathrm{TRE}_\mathrm{r}^2 + \mathrm{TRE}_\mathrm{v}^2 + \mathrm{TRE}_\mathrm{w}^2\right\rangle = \frac{3\sigma^2}{N} + \mathbf{r}^2(\omega_v^2+\omega_w^2). \tag{65}$$

Because $\hat{\mathbf{r}}$, $\hat{\mathbf{v}}$ and $\hat{\mathbf{w}}$ form an orthogonal set of axes, we know that

$$\omega^2 = \omega_r^2 + \omega_v^2 + \omega_w^2 \tag{66}$$

and, hence

$$\omega_v^2 + \omega_w^2 = \omega^2 - \omega_r^2. \tag{67}$$

Thus, we may rewrite $\left\langle \mathrm{TRE}^2(\mathbf{r})\right\rangle$ as

$$\begin{aligned}&\frac{3\sigma^2}{N} + \left(x^2+y^2+z^2\right) \\ &\quad\times\left(\omega_1^2+\omega_2^2+\omega_3^2 - \frac{x^2\omega_1^2+y^2\omega_2^2+z^2\omega_3^2}{x^2+y^2+z^2}\right)\end{aligned} \tag{68}$$

which may be further simplified to

$$\frac{3\sigma^2}{N} + (y^2+z^2)\omega_1^2 + (x^2+z^2)\omega_2^2 + (x^2+y^2)\omega_3^2. \tag{69}$$

We write the distance of $\mathbf{r}$ from axis $i$ as $d_i$, (e.g., $d_1 = y^2+z^2$); we then have that

$$\left\langle \mathrm{TRE}^2(\mathbf{r})\right\rangle = \frac{3\sigma^2}{N} + \sum_{i=1}^3 d_j^2\omega_j^2. \tag{70}$$

Finally, we express $\omega_i^2$ in terms of $\sigma$ and $\Lambda$. We have that

$$\begin{aligned}\omega_i^2 = \left\langle \left(R_{jk}^{(1)}\right)^2\right\rangle &= \left\langle\left(\frac{\Lambda_{jj}Q_{jk}-\Lambda_{kk}Q_{kj}}{\Lambda_{jj}^2+\Lambda_{kk}^2}\right)^2\right\rangle \\ &= \frac{\sigma^2}{\Lambda_{jj}^2+\Lambda_{kk}^2}\end{aligned} \tag{71}$$

---

$$\begin{aligned}2xz(\omega_x^2-\omega_y^2)\upsilon_x =& (y^2+z^2)(\omega_y^2-\omega_x^2)+(x^2+y^2)(\omega_z^2-\omega_y^2) \\ &\pm\sqrt{\left((y^2+z^2)(\omega_x^2-\omega_y^2)+(x^2+y^2)(\omega_y^2-\omega_z^2)\right)^2 - 4x^2z^2(\omega_x^2-\omega_y^2)(\omega_y^2-\omega_z^2)}\end{aligned} \tag{58}$$

For $\{i, j, k\} = \{1, 2, 3\}, \{2, 3, 1\}$, or $\{3, 1, 2\}$. But $\Lambda_{jj}^2 + \Lambda_{kk}^2$ is equal to the sum of squared distances of the fiducial points from axis $i$; we will write this quantity as $f_i^2$. We have, then, that

$$\begin{aligned}\langle \mathrm{TRE}^2(\mathbf{r}) \rangle &= \sigma^2 \left( \frac{3}{N} + \sum_{i=1}^{3} \frac{d_i^2}{f_i^2} \right) \\ &= \langle \mathrm{FLE}^2 \rangle \left( \frac{1}{\mathrm{N}} + \frac{1}{3} \sum_{i=1}^{3} \frac{d_i^2}{f_i^2} \right) \end{aligned} \quad (72)$$

which is the 3-D case of the expression we previously derived [20] for the expected squared TRE.

We now derive the expected squared value of the component of **TRE** in an arbitrary direction. We represent this direction by the unit vector $\hat{\mathbf{a}}$. From (27) we can see that the component of **TRE** in direction $\hat{\mathbf{a}}$ is

$$\mathrm{TRE_a} = (\mathrm{TRE_r})\hat{a}_\mathrm{r} + (\mathrm{TRE_v})\hat{a}_\mathrm{v} + (\mathrm{TRE_w})\hat{a}_\mathrm{w} \quad (73)$$

where $\hat{a}_r = \hat{\mathbf{a}} \cdot \hat{\mathbf{r}}, \hat{a}_v = \hat{\mathbf{a}} \cdot \hat{\mathbf{v}}$ and $\hat{a}_w = \hat{\mathbf{a}} \cdot \hat{\mathbf{w}}$. We may then deduce that

$$\langle \mathrm{TRE_a^2} \rangle = \langle \mathrm{TRE_r^2} \rangle \hat{a}_\mathrm{r}^2 + \langle \mathrm{TRE_v^2} \rangle \hat{a}_v^2 + \langle \mathrm{TRE_w^2} \rangle \hat{a}_\mathrm{w}^2 \quad (74)$$

recalling that the fact that the three components of **TRE** are uncorrelated ensures that the expected product of distinct components is zero. From (61), (63), and (64) we can see that we may rewrite (74) as

$$\begin{aligned}\langle \mathrm{TRE_a^2} \rangle &= \frac{\hat{a}_r^2 \sigma^2}{N} + \hat{a}_v^2 \left( \frac{\sigma^2}{N} + r^2 \omega_w^2 \right) + \hat{a}_w^2 \left( \frac{\sigma^2}{N} + r^2 \omega_v^2 \right) \\ &= \frac{\sigma^2}{N} + r^2 \left( \hat{a}_v^2 \omega_w^2 + \hat{a}_w^2 \omega_v^2 \right) \end{aligned} \quad (75)$$

where $\omega_v$ and $\omega_w$ are the components of $\mathbf{\Omega}$ along $\hat{\mathbf{v}}$ and $\hat{\mathbf{w}}$, respectively.

### *E. Distributions*

We have shown how to decompose $\mathbf{TRE}(\mathbf{r})$ into three orthogonal components that are uncorrelated with each other. Because they are linear combinations of $\mathcal{N}(0, \sigma)$ variables, namely the $F_{ij}$, these three components are themselves normal variables with zero mean. Because they are uncorrelated, they are also independent [25].

Since the three orthogonal components of **TRE** are independent, normally distributed variables with zero mean, $\mathrm{TRE}^2$ is distributed as the sum of three chi-square variables

$$\mathrm{TRE}^2 \sim \sigma_\mathrm{r}^2 \chi_1^2 + \sigma_\mathrm{v}^2 \chi_1^2 + \sigma_\mathrm{w}^2 \chi_1^2. \quad (76)$$

From (73), we know that $\mathrm{TRE_a}$ is the sum of three independent, zero mean, normal variables. Hence, we may deduce that $\mathrm{TRE_a}$ itself is a zero mean, normal variable and that its variance $\sigma_a^2$ is equal to $\langle \mathrm{TRE_a^2} \rangle$. We write

$$\mathrm{TRE_a} \sim \mathcal{N}(0, \sigma_\mathrm{a}) \quad (77)$$

where $\sigma_a^2$ is equal to the right-hand side of (75).

## IV. Numerical Simulations

Equation (72) has been shown [20] to be in excellent agreement with numerical simulations. In this paper, we have derived (76) and (77), which are distributions, as opposed to merely expected values. In order to test our approximation and verify our derivations, we have performed some additional simulations. We have chosen (76) for numerical simulation because it includes contributions from all components of TRE.

First, we chose five values of $N$ for which to perform the test: $N = 3$, 4, 10, 20, 50. For each of these values of $N$, we generated the correct number of fiducial positions randomly with uniform distribution within a cube of side 200 mm and one target position randomly with uniform distribution within a cube of side 400 mm. In order to model the FLE of (6), we perturbed independently the $x, y$, and $z$ components of the fiducial positions in one space, using normally distributed independent random variables with zero mean and variance 1/3 mm$^2$. In this way, we produced the same model as that used by Sibson in 1979 [18] and for our previous simulations [20]. We registered the perturbed positions to the original ones, measuring the TRE at the target position according to (11).

One simulation consisted of 100 000 repetitions of the perturbation and registration step, allowing us to estimate the distribution of squared TRE. We performed 100 000 iterations of generating three independent, zero-mean normal variables with variances as derived in Section III-C to represent the three components of TRE; we squared and added these variables to give a squared TRE value. We used the Kolmogorov-Smirnov test [26] to compare the distributions of these simulated and generated values. In the case $N = 10$, there was a significant difference between the two distributions (K-S test, $p < 0.05$), but this difference was not significant to $p < 0.01$. In all other cases, there was no significant difference. In Fig. 2, we show corresponding percentiles of squared TRE for the generated and simulated values. The fact that each plot shows a straight line along $x = y$ demonstrates that the distributions of the two types of values match closely. From the percentile values themselves in these five plots we can also see that, as implied by (72), the magnitude of TRE is decreased as more fiducials are added. The only exception is to this rule is that the $\mathrm{TRE}^2$ percentile values for $N = 20$ are smaller than those for $N = 50$. The explanation for this is that the randomly chosen target for $N = 50$ happens to be more than five times further from the fiducials' centroid than is the case for $N = 20$ and (72) also implies that TRE increases with distance of the target from the fiducials' centroid. In Table I, we compare the predicted variance of each component of TRE from (61), (63), and (64) with the observed variance of these components of the simulated values. For all the fiducial configurations studied, the observed values match the predicted ones within 1%. As pointed out in Section I, our derivations take account of anisotropy in the rotational components of registration error. A convenient measure of the anisotropy is $\omega_v/\omega_w$. In Table I, for the case $N = 3$, $\omega_v/\omega_w = 0.83$ and for the case $N = 4$, $\omega_v/\omega_w = 0.47$ using (63) and (64) to derive this ratio given $\sigma_v^2$ and $\sigma_w^2$; these two cases demonstrate that our theory correctly handles significant amounts of anisotropy in these rotational components.

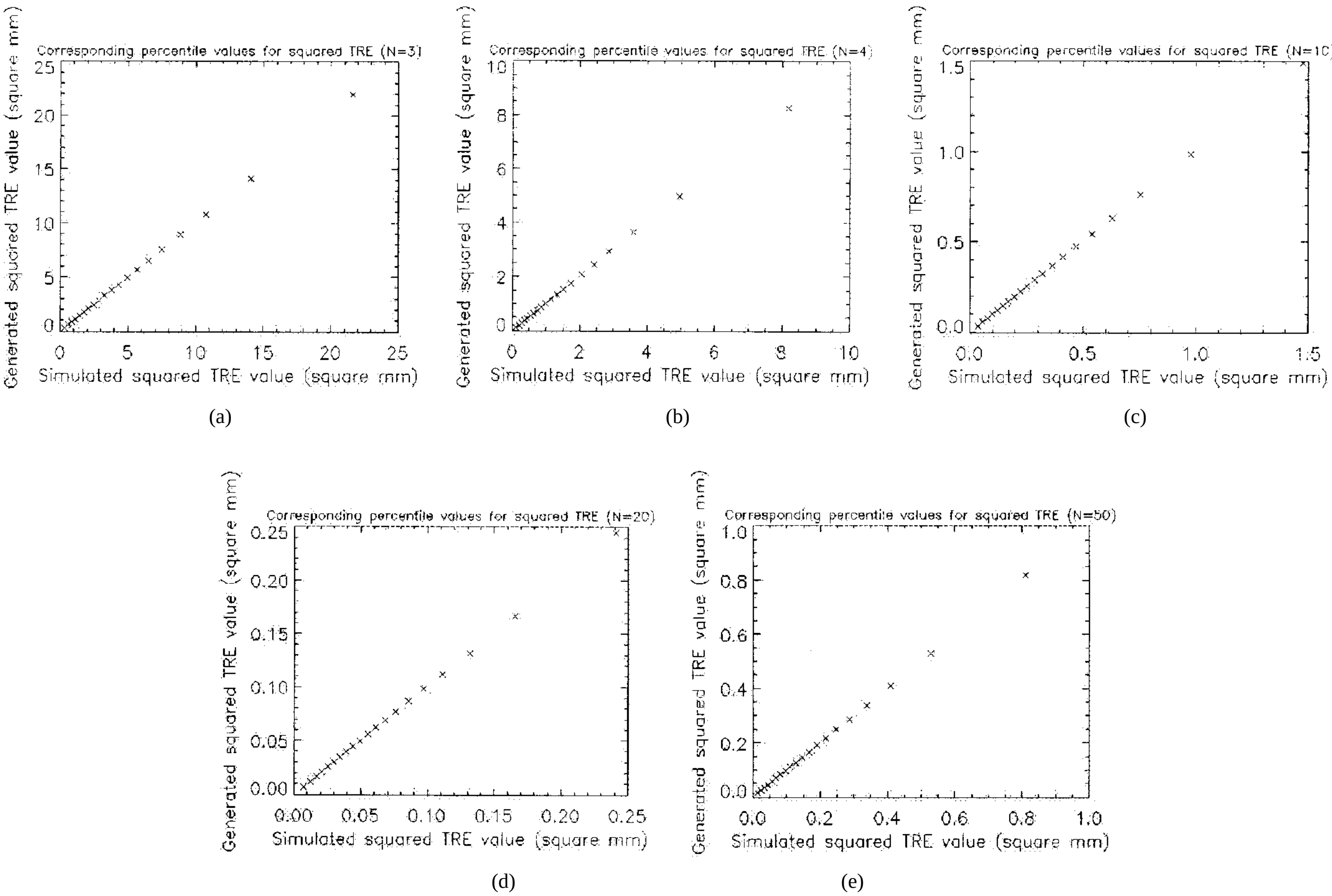


Fig. 2. Plots of corresponding percentile values for squared TRE. On the $x$ axis are percentile values in 5% increments from five to ninety-five and the ninety-ninth percentile value, for the simulated squared TRE. The $y$ axis shows the same percentiles for generated squared TRE values. (a) 3, (b) 4, (c) 10, (d) 20, and (e) 50 fiducial points.

TABLE I

COLUMNS 2, 3, AND 4 GIVE PREDICTED VARIANCES OF $\mathrm{TRE_r}$, $\mathrm{TRE_v}$ AND $\mathrm{TRE_w}$ ACCORDING TO (61), (63), AND (64). COLUMNS 5, 6, AND 7 GIVE THE OBSERVED VARIANCES OF THESE TRE COMPONENTS FROM NUMERICAL SIMULATION. COLUMN 8 GIVES THE MAXIMUM PERCENTAGE DISCREPANCY BETWEEN CORRESPONDING PREDICTED AND OBSERVED VALUES FOR EACH ROW

| $N$ | Predicted values | | | Observed values | | | Max. % error |
|---|---|---|---|---|---|---|---|
| | $\sigma_r^2$ | $\sigma_v^2$ | $\sigma_w^2$ | $\sigma_r^2$ | $\sigma_v^2$ | $\sigma_w^2$ | |
| 3 | 0.1111 | 2.7105 | 1.8970 | 0.1118 | 2.6960 | 1.9070 | 0.63 |
| 4 | 0.0833 | 1.1608 | 0.3192 | 0.0837 | 1.1564 | 0.3194 | 0.50 |
| 10 | 0.0333 | 0.1766 | 0.1342 | 0.0330 | 0.1767 | 0.1352 | 0.91 |
| 20 | 0.0167 | 0.0243 | 0.0221 | 0.0167 | 0.0243 | 0.0220 | 0.45 |
| 50 | 0.0067 | 0.0984 | 0.0755 | 0.0067 | 0.0982 | 0.0753 | 0.27 |

## V. DISCUSSION

We have shown that TRE may be separated into three orthogonal components and that these components have a zero mean, normal distribution with variances given by (61), (63), and (64). We have also demonstrated that the components are independent of each other and, hence, we know that the square of TRE is distributed as in (76). As expected, the radial component of TRE has the smallest variance, as this component contains only translational error; the other two components contain both rotational and translational error. In addition to the square of TRE, we have also derived the distribution of $\mathrm{TRE_a}$, the component of TRE in an arbitrary direction. We find that this distribution is normal. The normal form is inevitable because 1) in the limit of zero perturbation TRE is zero and 2), we have ignored terms of order two and higher in $\epsilon$. $\mathrm{TRE_a}$ is, therefore, a linear combination of the elements of $F$ and is, thus, a zero mean, normally distributed variable. In (75), we have given $\sigma_a^2$, the variance of this variable, which completes the derivation.

The close agreement of the predicted values of $\sigma_r^2, \sigma_v^2$, and $\sigma_w^2$ with those given by the numerical simulations shows that our first-order approximation is a good one for the fiducial configurations we tested. However, when applying this theory to general configurations, care must be taken to ensure that the conditions are such that this approximation remains valid. This will be the case as long as $R^{(2)}$, the second order component of the rotation, remains small compared to $R^{(1)}$. From the orthogonality condition on $R$ we know that

$$((I + R^{(1)} + R^{(2)} + \mathrm{O}(\epsilon^3))(I + R^{(1)t} + R^{(2)t} + \mathrm{O}(\epsilon^3))) = I \tag{78}$$

and equating second order terms in $\epsilon$ gives that

$$R^{(2)} + R^{(2)t} = R^{(1)2}. \tag{79}$$

Thus, we know that the symmetric part of the second order term is much smaller than $R^{(1)}$ in the case that $|R_{ij}^{(1)}| \ll 1$. From the definition of $R^{(1)}$ in (19), we can see that this will be the case

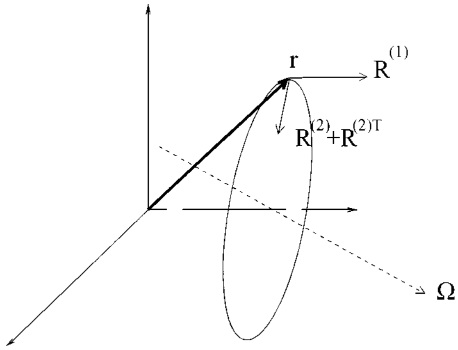


Fig. 3. The rotational component of TRE($\mathbf{r}$). To first order, the error is a motion orthogonal to the axis of rotation $\mathbf{\Omega}$ and to $\mathbf{r}$, represented by $R^{(1)}$. At second order, there is an additional component corresponding to movement toward the axis of rotation, represented by the symmetric part of $R^{(2)}$.

if the magnitude of the perturbations in $F$ and, thus, the magnitude of the elements of $Q$, is much smaller than the moment of the fiducial configuration about each axis. For a near-collinear fiducial configuration, this relationship may not hold. One of the moments $f_i$ will be small; if this moment is of the same order of magnitude as the elements of $F$, representing the FLE, the approximation we use is no longer valid. In the numerical simulations we performed, none of the configurations is collinear by this measure.

We know that we can write $R^{(1)}\mathbf{r}$ as $\mathbf{\Omega} \times \mathbf{r}$; hence, from (79) we can see that we may write $(R^{(2)} + R^{(2)t})\mathbf{r}$ as $\mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{r})$. We illustrate the rotational component of TRE in Fig. 3. This rotational motion may be considered as being analogous to that of the motion of a particle rotating about an axis $\mathbf{\Omega}$; its velocity is equivalent to the first order term and the symmetric part of $R^{(2)}$ represents the second-order "correction" to this velocity, *i.e.,* the particle's acceleration toward the center of the circle. The *antisymmetric* part of $R^{(2)}$ is the second-order "correction" to the rotational axis $\mathbf{\Omega}$ about which the first-order rotation takes place. The determination of this correction is beyond the scope of this present work.

## VI. Conclusion

With (75)–(77), we have derived the approximate expressions for the distribution of TRE that we set out to find. Our approximation agrees closely with simulations, showing that it is an excellent indicator of TRE, as long as the fiducial configuration is sufficiently noncollinear.

We should note that the predictive value is at its greatest when the assumptions under which it was derived remain valid, namely that all the spatial components of localization error for all the fiducials are independent, normally distributed random variables with zero means and equal variances. Because medical images are commonly anisotropic, *i.e.,* they have different spatial resolutions in the coordinate directions, it is not always a good assumption that FLE is isotropically distributed. This problem is ameliorated, however, when the fiducal points are the centroids of markers that are large enough to encompass two or more voxels [27]. Although we have shown elsewhere [28] that, at least in terms of expected value of TRE, our theory is still of value for such cases, it remains an interesting problem to extend our derivations to allow for the anisotropy in FLE that sometimes accompanies anisotropy in the voxel dimensions. This problem will no doubt prove to be a difficult one, because there is no known closed-form solution to the rigid-body registration problem in the case that errors must be weighted differently according to their direction [27]. Another problem is that of unequal weighting among the fiducials in (1). Unequal weighting is appropriate when the localization error varies among the fiducials. This situation arises when the fiducials are derived from landmarks of varying visibility and definition.

While future solutions of the problems of spatial anisotropy and fiducial weighting may extend the application of this work to other medical areas, the assumptions that we have made for these derivations are already realized quite faithfully for bone-attached fiducial markers, used in neurosurgery, orthopedic surgery and radiation oncology. It is hoped that the statistical distributions of the components of TRE that we have derived here will allow surgeons and radiation oncologists to use point-based image registration with more confidence to aid both diagnosis and treatment.

## References


[1] P. Clarysse, D. Gibon, J. Rousseau, S. Blond, C. Vasseur, and X. Marchandise, "A computer-assisted system for 3-D frameless localization in stereotaxic MR," *IEEE Trans. Med. Imag.*, vol. 10, pp. 523–529, Dec. 1991.

[2] D. L. G. Hill, D. J. Hawkes, Z. Hussain, S. E. M. Green, C. F. Ruff, and G. P. Robinson, "Accurate combination of CT and MR data of the head: Validation and applications in surgical and therapy planning," in *3D Advanced Image Processing in Medicine 1992*, C. Roux, G. T. Herman, and R. Collorec, Eds. Piscataway, NJ: IEEE Press, 1992, pp. 79–83.

[3] A. C. Evans, T. M. Peters, D. L. Collins, P. Neelin, and C. Gabe, "Image registration based on discrete anatomic structures," in *Interactive Image-Guided Neurosurgery*, R. J. Maciunas, Ed. Park Ridge, IL: Amer. Assoc. Neurological Surgeons, 1993, pp. 63–80.

[4] C. R. Maurer Jr. and J. M. Fitzpatrick, "A review of medical image registration," in *Interactive Image-Guided Neurosurgery*, R. J. Maciunas, Ed. Park Ridge, IL: Amer. Assoc. Neurological Surgeons, 1993, pp. 17–44.

[5] D. L. G. Hill, D. J. Hawkes, M. J. Gleeson, T. C. S. Cox, A. J. Strong, W.-L. Wong, C. F. Ruff, N. D. Kitchen, D. G. T. Thomas, J. E. Crossman, C. Studholme, A. J. Gandhe, S. E. M. Green, and G. P. Robinson, "Accurate frameless registration of MR and CT images of the head: Applications in surgery and radiotherapy planning," *Radiology*, vol. 191, pp. 447–454, 1994.

[6] J. B. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer Jr., R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. G. Hill, D. J. Hawkes, C. Studholme, J. B. A. Maintz, M. A. Viergever, G. Malandain, X. Pennec, M. E. Noz, G. Q. Maguire Jr., M. Pollack, C. A. Pelizzari, R. A. Robb, D. Hanson, and R. P. Woods, "Comparison and evaluation of retrospective intermodality brain image registration techniques," *J. Comput. Assist. Tomogr.*, vol. 21, pp. 554–566, 1997.

[7] C. R. Maurer Jr., J. M. Fitzpatrick, M. Y. Wang, R. L. Galloway Jr., R. J. Macinnas, and G. S. Allen, "Registration of Head Volume Images Using Implantable Fiducial Markers," Dept. Comput. Sci., Vanderbilt Univ., Nashville, TN, Tech. Rep. CS-96-03, Sept. 1996.

[8] J. R. Hurley and R. B. Cattell, "The procrustes program: Producing direct rotation to test a hypothesized factor structure," *Behav. Sci.*, vol. 7, pp. 258–262, 1962.

[9] P. H. Schönemann, "A generalized solution of the orthogonal procrustes problem," *Psychometrika*, vol. 31, pp. 1–10, 1966.

[10] G. Golub and C. van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1983.

[11] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 698–700, 1987.

[12] V. R. Mandava, "Three-dimensional multimodal image registration using implanted markers," PhD dissertation, Vanderbilt Univ., Nashville, TN, 1991.

[13] C. R. Maurer Jr., J. J. McCrory, and J. M. Fitzpatrick, "Estimation of accuracy in localizing externally attached markers in multimodal volume head images," *Proc. SPIE—Medical Imaging 1993: Image Processing*, vol. 1898, pp. 43–54, 1993.

[14] A. C. Evans, S. Marrett, D. L. Collins, and T. M. Peters, "Anatomical-functional correlative analysis of the human brain using three-dimensional imaging systems," *Proc. SPIE—Medical Imaging III: Image Processing*, vol. 1092, pp. 264–274, 1989.

[15] V. R. Mandava, J. M. Fitzpatrick, C. R. Maurer Jr., R. J. Maciunas, and G. S. Allen, "Registration of multimodal volume head images via attached markers," in *Proc. SPIE—Medical Imaging VI: Image Processing*, vol. 1652, pp. 271–282, 1992.

[16] C. R. Maurer Jr., J. M. Fitzpatrick, M. Y. Wang, R. L. Galloway Jr., R. J. Maciunas, and G. S. Allen, "Registration of head volume images using implantable fiducial markers," *IEEE Trans. Med. Imag.*, vol. 16, pp. 447–462, Aug. 1997.

[17] K. Darabi, P. Grunert, and A. Perneczky, "Accuracy of intraoperative navigation using skin markers," in *Computer Assisted Radiology and Surgery 1997*, H. U. Lemke, M. W. Vannier, and K. Inamura, Eds. Berlin: Springer-Verlag, 1997, pp. 920–924.

[18] R. Sibson, "Studies in the robustness of multidimensional scaling: Perturbational analysis of classical scaling," *J. Roy. Statist. Soc. B*, vol. 41, pp. 217–229, 1979.

[19] J. M. Fitzpatrick, J. B. West, and C. R. Maurer Jr., "Derivation of expected registration error for rigid-body, point- based image registration," *Proc. SPIE—Medical Imaging 1998: Image Processing*, vol. 3338-01, pp. 16–27, 1998.

[20] J. M. Fitzpatrick, J. West, and C. R. Maurer Jr., "Predicting error in rigid-body, point-based registration," *IEEE Trans. Med. Imag.*, vol. 17, pp. 694–702, 1998.

[21] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer Jr., R. M. Kessler, and R. J. Maciunas, "Retrospective intermodality registration techniques: Surface-based versus volume-based," *IEEE Trans. Med. Imag.*, vol. 18, pp. 144–150, Feb. 1999.

[22] S. P. Langron and A. J. Collins, "Perturbation theory for generalized procrustes analysis," *J. Roy. Statist. Soc. B*, vol. 47, pp. 277–284, 1985.

[23] B. F. Green, "The orthogonal approximation of an oblique structure in factor analysis," *Psychometrika*, vol. 17, pp. 429–440, 1952.

[24] C. Goodall, "Procrustes methods in the statistical analysis of shape," *J. Roy. Statist. Soc. B*, vol. 53, pp. 285–339, 1991.

[25] A. C. Rencher, *Linear Models in Statistics*. New York: Wiley, 2000.

[26] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*. New York: Cambridge Univ. Press, 1990.

[27] J. M. Fitzpatrick, D. L. G. Hill, and C. R. Maurer Jr., "Image registration," in *Handbook of Medical Imaging*, M. Sonka and J. M. Fitzpatrick, Eds. Bellingham, WA: SPIE Press, 2000, vol. 2, Medical Image Processing and Analysis.

[28] J. B. West and J. M. Fitzpatrick, *Proc. SPIE—Point-Based Rigid Registration: Clinical Validation of Theory*, vol. 3979, pp. 353–359, 2000.