Calibration and characterization of electromagnetic position and orientation trackers

Robert A. MacLachlan, *Member, IEEE*, Claudia Pelle, Ralph L. Hollis, *Fellow, IEEE*, Elena De Momi, *Senior Member, IEEE*, and Cameron N. Riviere, *Senior Member, IEEE*

***Abstract*—An electromagnetic position and orientation tracker must be calibrated so that magnetic measurements can be converted into spatial data, and the accuracy must also be characterized to know what accuracy is obtained. We describe a magnetic calibration matrix model which is extensible to the common coil configurations. We detail data representations and discuss successful approaches to the calibration optimization. An accurate positioning device is necessary for both calibration and characterization. We discuss the effect of positioning error on calibration accuracy and some design tradeoffs for positioning. We characterize the calibrated accuracy with two different source coil designs and several calibration variations, with the best configuration achieving 208 µm/0.27º uncertainty, an accuracy exceeding published values for commercial EMTs (tested by other methods). We use local measurement nonlinearity to characterize the expected accuracy across small motions, including the cross-coupling between position and orientation. All software described is open source (Apache License 2.0).**

***Index Terms*—Magnetic position and orientation tracking, magnetic model fitting, position measurement.**

I. Introduction

ELECTRO-MAGNETIC trackers (EMTs) have been in use for over 40 years [1], and in that time have seen niche use for short-range measurement of arbitrary motion, especially when the absence of reliable sightlines prevents the use of optical methods. At this point it seems unlikely that EMT will become a mass-market technology, but there is no practical alternative in some uses such as surgical guidance.

EMT hardware is relatively simple and can be constructed almost entirely using standard audio electronics methods. The only specialized components are the source and sensor coils, but Radio Frequency Identification (RFID) technology has made available a variety of coil assemblies that can be used in undemanding EMT applications. What remains a considerable barrier to wider use of EMTs is the need to implement software for signal processing, pose solution, and calibration. EMTs have been developed by numerous manufacturers and many academic research projects, with varying sophistication. Since much of EMT implementation practice is concealed within proprietary products, there is a clear opportunity for open source EMT designs, which can reduce the need to rediscover and reimplement the many hardware and software details.

We have developed the ILEMT (In-Loop ElectroMagnetic Tracker), an open design EMT with adequate speed, resolution and accuracy to be used within a feedback system that stabilizes handheld surgical instruments[2][3]. For this particular realization, the focus is on measurement speed and low noise, which is achieved in the hardware design, and not by calibration, but the calibration methods discussed here are formulated to generalize across a variety of EMT hardware. We demonstrate generality by characterizing two different source designs and two pose solution methods.

We overview methods of generating precise motion for calibration and discuss accuracy testing at length. One important result is a quantification of the impact on accuracy of calibrating with different degrees of freedom than are used during testing. The consistent use of a linear transform pose representation becomes particularly valuable for clarifying the relation between many coordinate systems, such as are present when considering the effect of position error during calibration on the apparent and actual accuracy of the EMT.

This work describes a novel method of EMT position error correction, the Direct Linear Transform (DLT), and introduces a new way of characterizing and visualizing the local error (Differential Nonlinearity) of the EMT output, but we anticipate that a major value of this presentation is combining in one place a clear formulation of the EMT calibration problem together with numerous practical details of implementation. An important aspect of this work is that we have adopted an open science approach. This is not in itself entirely novel [4], but the hardware, software, and data products of our work are a significant contribution to the open resources for EMT development. These are available under Apache License 2.0 on osf.io[3].

Manuscript received May 1, 2022. This work was supported in part by the U.S. National Institutes of Health under Grants R01EB000526, R01EB024564, and R01HL105911. *(Corresponding author: Cameron N. Riviere.)*

Robert A. MacLachlan, Ralph Hollis, and Cameron N. Riviere are with the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: camr@ri.cmu.edu).

Claudia Pelle and Elena De Momi are with the Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy.

II. EMT OVERVIEW

An EMT measures the *pose* (position and orientation) of a *sensor* with respect to a magnetic *source*. This measurement has six degrees of freedom (6 DOF). A single field measurement is not sufficient to determine the pose, so the source incorporates n_{so} distinctly modulated electromagnet source coils, and the sensor measures the local magnetic field along $n_{se} \geq 1$ independent axes. We assume without loss of generality that $n_{so} = n_{se} = 3$, but other combinations are possible. For each pose, each source field experiences a characteristic magnetic coupling into each sensor coil, giving $n_{so}n_{se}$ distinct couplings, which can be represented as the coupling matrix $\boldsymbol{C}$, where $\boldsymbol{C}_{jk}$ is the signed coupling from source coil j into sensor coil k. The sign represents direction of the magnetic field vector with respect to the sensor coil.

Sometimes the source and sensor designs are identical, but the common practice is to make the source much larger (with a correspondingly increased drive current). Then the sensor can be smaller while still maintaining a given signal-to-noise ratio. With this relatively large source, its magnetic aberrations tend to dominate the calibration error.

Frequently the source or sensor coils are wound around a common center (concentric), in a cube or sphere configuration. As well as being more compact than multiple separate coils, when both source and sensor are concentric this also simplifies the pose solution problem[2][3][7]. But even when the coils are fabricated to approach the concentric orthogonal ideal, there is always some deviation, so accuracy can be improved by adding calibration parameters to model the coil position and orientation error. Highly non-concentric coil arrangements can also be useful [8][9][10], in which case we must model the actual coil configuration.

Some EMTs use other sensing principles and source modulations [11], [12], but there is little loss of generality to consider the most common approach, where the source coil has a sinusoidal drive voltage and the sensor uses inductive pickup coils. These methods are applicable to any EMT that operates based on pairwise measurement of coupling between source and sensor axes.

A. The calibration problem

Let $\boldsymbol{c}(\boldsymbol{P})$ be the *measurement function*: the actual coupling $\boldsymbol{C}$ measured in pose $\boldsymbol{P}$. In practice, $\boldsymbol{c}(\boldsymbol{P})$ depends on magnetic interactions that are intractable to model. Instead, we resort to a *measurement model* $\tilde{\boldsymbol{c}}(\boldsymbol{P}, \rho)$. This is a physical model which is parameterized by ρ. Calibration is the process of choosing a suitable $\tilde{\boldsymbol{c}}(\ldots)$ and determining ρ. Our specific measurement model and the structure of ρ is discussed later (§V.)

During calibration we measure $\boldsymbol{c}(\boldsymbol{P}_i)$ for many different known poses $\boldsymbol{P}_{1..n}$ spread across the position and orientation workspace. Then a general nonlinear optimizer is used to determine the ρ that minimizes:

$$\sum_{\boldsymbol{i=1..n}} \left| \frac{\tilde{\boldsymbol{c}}(\boldsymbol{P}_i, \rho) - \boldsymbol{c}(\boldsymbol{P}_i)}{|\boldsymbol{c}(\boldsymbol{P}_i)|_2} \right|_2 \tag{1}$$

Where $|\boldsymbol{A}|_2$ is the matrix 2-norm:

$$\sqrt{\sum_j \sum_k \left(\boldsymbol{A}_{jk}\right)^2}$$

This least-squares formulation permits us to use a least-squares nonlinear solver. These solvers give much more robust convergence than general function minimization, especially when the derivatives of $\tilde{\boldsymbol{c}}(\boldsymbol{P}_i, \rho)$ are unknown and the initial value for ρ is inaccurate. Normalizing by the measured coupling magnitude $|\boldsymbol{c}(\boldsymbol{P}_i)|_2$ gives each point equal weight during optimization, in spite of the large change in the field strength as the source to sensor distance is varied [8][9].

III. RELATED WORK

A. Positioning

Calibration and accuracy testing require placing the sensor in many known poses, in both translation and rotation. This *ground truth* must have significantly higher accuracy than the expected tracker performance. Compounding the difficulty, the sensor positioner must also be non-metallic to avoid interfering with the measurement.

One approach has been to use a nonmetallic manual positioning fixture [9][10]. This procedure is tedious, and rotation effects have often been ignored, perhaps in part because of the large number of poses that must be tested when rotation and translation are varied simultaneously.

An alternative to an accurate positioner is the use of another sufficiently precise position measurement instrument to find the ground truth sensor position, such as an optical tracker [15] or interferometer [12][16]. A coordinate-measuring machine (CMM) mechanically positions a probe in 3D space with very high precision and can be used directly to position the sensor, but is often too expensive to be dedicated to this use, so it is more commonly used to calibrate positioning components [14].

B. Evaluation

There are many performance evaluations of commercial EMTs for medical applications. In addition to the static pose error relevant to calibration, a full evaluation must also measure the noise (precision) and dynamic response.

Studies by EMT vendors use automated positioning systems with traceable calibration to credibly establish the accuracy of the tested EMT [15][12][16]. While sophisticated and costly tests are described, the emphasis is not on documenting procedures such as positioner calibration in enough detail to enable reproducing the results.

There are also many studies by potential users of medical EMTs, both to compare performance of different products, and also to test the suitability of a particular EMT for a specific type of medical procedure [17]. Usual practice is to use a nonmetallic manual fixture (or *phantom*) for sensor positioning [18]. Some tracker models have been evaluated enough times to reveal that the evaluation procedures are not repeatable at the level of accuracy claimed [19]. Aside from the challenge of affordable accurate positioning, inadequate evaluation of the translation error caused by sensor rotation is also a concern. While not comprehensive, these relatively simple procedures do reveal significant variation across EMTs and their use environments. Typical position errors seen in an ideal environment are $0.5\,\text{mm} - 1\,\text{mm}$ RMS at a range of $200\,\text{mm} - 400\,\text{mm}$. In practice EMT pose error is often dominated by interference from metal in or near the workspace, which is highly application dependent, but can exceed 5 mm RMS [19].

C. Output correction

Once tracker output poses have been collected across a range of ground truth poses, it is also common to undertake to correct the tracker error [14][20][21], generally using interpolated lookup tables. This differs from magnetic calibration in that it assumes a pose output and fits a model to the pose error. The tracker calibration is a given, provided by the manufacturer, and embedded into the product's proprietary firmware and calibration data. Output correction can have considerable benefit in the particular calibrated environment, since it compensates for interference from metal objects, but gives exaggerated expectations of compensated performance when the accuracy is evaluated on the same data used for compensation, or without consideration of the positioning uncertainty.

D. Calibration

In comparison to the related work above, there is relatively little detailed discussion of calibration in the literature. While commercial EMT vendors have no interest in revealing their proprietary calibration procedures, calibration **is** usually discussed in descriptions of new EMTs developed for research, but may be specific to the particular hardware used. Nonlinear optimization, such as the approach we describe, is usually used. Refs. [8][9][22] have more specifics on the model parameters, and especially the optimization strategy.

IV. Linear Transforms and Matrices

We use *linear transform* notation for poses and coordinate systems, with the *linear homogeneous* matrix implementation. This is an important matter of practice, and not merely notation. When compared to use of angles and trigonometry, the superiority of transform methods becomes very clear once there are multiple coordinate systems. This practice has become nearly universal in robotics and computer graphics, see [23] for an introduction. There are no angles used in our implementation (but see rotation vectors below).

The measured pose $\boldsymbol{P}$ is the linear transform which maps the source coordinate frame onto the sensor coordinates. A linear transform is the abstraction of a rigid mapping on 6DOF space: translation with rotation. Linear homogeneous coordinates are a particular representation of a transform:

$$\boldsymbol{T} = \begin{bmatrix} \hat{x}_1 & \hat{y}_1 & \hat{z}_1 & x \\ \hat{x}_2 & \hat{y}_2 & \hat{z}_2 & y \\ \hat{x}_3 & \hat{y}_3 & \hat{z}_3 & z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

xyz is the translation part of the transform (position), while $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}$ are the unit vectors defining the axes of the coordinate frame (also known as a rotation matrix or direction cosines). Let $\boldsymbol{T}_1$, $\boldsymbol{T}_2$ be arbitrary linear homogeneous transforms. Successive transforms are composed by matrix multiplication: $\boldsymbol{T}_1\,\boldsymbol{T}_2$. The inverse transformation is the matrix inverse $\boldsymbol{T}^{-1}$. A point is represented by $\boldsymbol{p} = [x\ y\ z\ 1]^T$, while a vector is $\boldsymbol{v} = [x\ y\ z\ 0]^T$. Either is transformed by matrix multiplication:

$$\boldsymbol{p}' = \boldsymbol{T}\,\boldsymbol{p}$$
$$\boldsymbol{v}' = \boldsymbol{T}\,\boldsymbol{v}$$

A vector is only rotated by the transform, while a point is both rotated and translated. Transforms compose on the left when operating on points and vectors: $\boldsymbol{T}_2\,\boldsymbol{T}_1\,\boldsymbol{p}$, etc.

Pose graphs such as Fig. 2, Fig. 3 are a visual notation for systems of transforms which makes it easy to construct needed transform equations by inspection [23]. Each arrow represents the application of a transform that defines the relation between two coordinate systems.

The calibration optimization state includes transforms, but the linear homogeneous representation is unsuitable because it has more than 6 elements. These excess degrees of freedom are suppressed in an orthonormal rotation matrix, but this property is not preserved by the optimizer. Three-angle representations (roll, pitch, yaw) do not have excess degrees of freedom, but suffer from gimbal lock and numerical instability.

We have chosen a *rotation vector* representation during optimization:

$$\boldsymbol{r} = \langle R_x, R_y, R_z \rangle$$

This is closely related to the axis-angle representation $\langle \hat{r}_x, \hat{r}_y, \hat{r}_z, \theta \rangle$, where $\hat{\boldsymbol{r}}$ is the unit vector for the axis of rotation, and θ is the rotation in radians. The corresponding rotation vector is $\boldsymbol{r} = \hat{\boldsymbol{r}}/\theta$. The complete representation for a pose is then:

$$\boldsymbol{P} = \langle x, y, z, R_x, R_y, R_z \rangle \tag{2}$$

A pose or linear transform may be represented as either this vector or as a linear homogeneous matrix, but must be in matrix form when the linear transform is to be applied.

V. The Measurement Model

One way to appreciate the challenge of measurement modeling is that we expect to achieve an accuracy which is much smaller than the size of the source (or sensor). In the tested ILEMT configuration, the source is 40 mm and the sensor 15 mm, while the desired accuracy is $< 400\ \mu\text{m}$. Due to the physics of magnetism, this is not as problematic is it seems. For distance from the source r, and source coil of size d, as r increases to $r \gg d$, the field converges to a dipole.

A. The dipole approximation

The dipole is the simplest physical model for the source magnetic field, yet often works well in magnetic trackers, so we will use it here. For this calibration framework the details of the $\hat{\boldsymbol{c}}(\boldsymbol{P}, \rho)$ model matter little, and higher order models can be used. Where the particular magnetic model does become important is in implementing the pose solution, which is outside our scope.

A dipole possesses only two parameters: location $\boldsymbol{l}$ and vector moment $\boldsymbol{m}$ (Fig. 1). If the sensor is located at $\boldsymbol{p}$, then the field $\boldsymbol{B}(\boldsymbol{p})$ is:

$$\boldsymbol{B}(\boldsymbol{p}) = \mu \frac{1}{|\boldsymbol{r}|^3}[3(\boldsymbol{m} \cdot \hat{\boldsymbol{r}})\hat{\boldsymbol{r}} - \boldsymbol{m}] \tag{3}$$

where $\boldsymbol{r} = \boldsymbol{p} - \boldsymbol{l}$, $\hat{\boldsymbol{r}} = \frac{\boldsymbol{r}}{|\boldsymbol{r}|}$, and μ is the permeability of the medium. $\boldsymbol{B}(\boldsymbol{p})$ is a vector, which each individual sensor coil measures as a scalar voltage $v_i(\boldsymbol{p})$ along some sensor moment $\boldsymbol{m}_i$. The dipole form of the sensor response is the scalar product of the sensor moment with the field:

$$v_i(\boldsymbol{p}) = k_g \boldsymbol{B}(\boldsymbol{p}) \cdot \boldsymbol{m}_i \tag{4}$$

Where k_g is a gain constant that in practice we absorb into the calibrated source and sensor moments, §VI.

B. *Coordinate notation*

These basic magnetic models are coordinate free, but we must introduce multiple coordinate systems to parameterize the full measurement model $\tilde{\boldsymbol{c}}(\boldsymbol{P}, \rho)$ across multiple source and sensor coils. In tracker kinematics, there are two independent parts: the source and the sensor, abbreviated "*so*" and "*se*". Each has its own coordinate system. If an arbitrary point $\boldsymbol{p}$ is represented in the source coordinates, then it is $\boldsymbol{p}^{so}$, while in the sensor coordinates it is $\boldsymbol{p}^{se}$. Our general notation is:

$${}^{part}\boldsymbol{p}_{index}^{coordinates}$$

where $part, coordinates \in \{so, se\}$, and $index$ optionally designates a sub-part. For example, the tracker output $\boldsymbol{P}$ is the sensor pose with respect to the source coordinates, which could be written as ${}^{se}\boldsymbol{P}^{so}$.

We will index the source and sensor coils by $\mathrm{x, y, z}$. Here, "x" is simply a name, not a variable. This is intuitive in the common case the source and sensor coils roughly coincide with the axes of the corresponding coordinate system. We represent the dipole parameters as two 3x3 matrices $\boldsymbol{L}$ and $\boldsymbol{M}$, where the coil locations $\boldsymbol{l}_\mathrm{x}, \boldsymbol{l}_\mathrm{y}, \boldsymbol{l}_\mathrm{z}$ and moments $\boldsymbol{m}_\mathrm{x}, \boldsymbol{m}_\mathrm{y}, \boldsymbol{m}_\mathrm{z}$ are concatenated as column vectors:

$$\boldsymbol{L} = [\boldsymbol{l}_\mathrm{x} \quad \boldsymbol{l}_\mathrm{y} \quad \boldsymbol{l}_\mathrm{z}] \qquad \boldsymbol{M} = [\boldsymbol{m}_\mathrm{x} \quad \boldsymbol{m}_\mathrm{y} \quad \boldsymbol{m}_\mathrm{z}] \tag{5}$$

The source parameters are ${}^{so}\boldsymbol{L}$ and ${}^{so}\boldsymbol{M}$, and the sensor has ${}^{se}\boldsymbol{L}$ and ${}^{se}\boldsymbol{M}$. Since these quantities move with the part, the calibration parameters are determined in the part's coordinate system, eg. ${}^{se}\boldsymbol{L}^{se}$, but when transformed to source coordinates the matrix becomes ${}^{se}\boldsymbol{L}^{so}$.

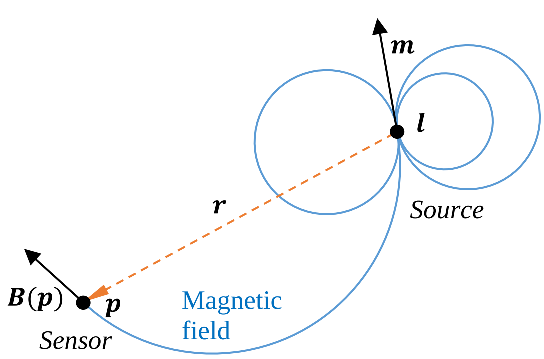

Fig. 1: Dipole model. A source coil is located at $\boldsymbol{l}$ with moment $\boldsymbol{m}$. The sensor located at $\boldsymbol{p}$ measures the source magnetic field vector $\boldsymbol{B}(\boldsymbol{p})$ along several axes.

C. *The multidimensional measurement model*

The dipole model (3), (4) gives a scalar coupling between a single source and sensor coil, while we must find $\tilde{\boldsymbol{c}}(\boldsymbol{P}, \rho)$, the full 3x3 predicted coupling matrix. The source and sensor parameters are defined in separate coordinate systems (Fig. 2). We establish a common frame by transforming the sensor parameters into the source coordinates. This is done by applying $\boldsymbol{P}$ as a linear transform (matrix multiplication):

$${}^{se}\boldsymbol{L}^{so} = \boldsymbol{P}\ {}^{se}\boldsymbol{L}^{se}, \qquad {}^{se}\boldsymbol{M}^{so} = \boldsymbol{P}\ {}^{se}\boldsymbol{M}^{se}$$

We then iterate over all combinations of source and sensor coils to find the pairwise couplings in $\tilde{\boldsymbol{c}}(\boldsymbol{P}, \rho)$.

D. *Fixture transforms*

When calibrating to a high accuracy, one challenge is that the calibration setup itself has unknown kinematic parameters. What exactly is the pose of the source and sensor with respect to the coordinate systems of our calibration fixtures? And where is the source fixture with respect to the sensor fixture?

EMTs are mainly used for measurement of relative motion so the precise mechanical definition of the source and sensor coordinates is relatively unimportant. It is common to perform *registration*, applying a rigid 6DOF transform to the pose measurement to align it with the test positioner or another coordinate system such as a medical image [15][12][18]. We have generalized registration into three *fixture transforms* which absorb the unknowns of the calibration setup. The calibration parameters ρ of the measurement model $\tilde{\boldsymbol{c}}(\boldsymbol{P}, \rho)$ are augmented with the fixture transforms, and the calibration optimization solves for these unknowns also.

Specifically, we need to model the ground truth kinematics so that by independent means we know the correct tracker output ${}^{se}\boldsymbol{P}^{so}$. This is done by a sequential combination of fixture transforms $\boldsymbol{F}$ and known motions $\boldsymbol{J}$. The $\boldsymbol{J}$ can be thought of as the joints in a robot arm, and the $\boldsymbol{F}$ are the fixed links that connect those joints:

$${}^{se}\boldsymbol{P}^{so} = {}^{so}\boldsymbol{F}\ {}^{so}\boldsymbol{J}\ {}^{st}\boldsymbol{F}\ {}^{st}\boldsymbol{J}\ {}^{se}\boldsymbol{F} \tag{6}$$

The $\boldsymbol{J}$ are known, and are varied during the calibration data collection, while the $\boldsymbol{F}$ are constants that are initially only imprecisely known. The "stage motion" ${}^{st}\boldsymbol{J}$ is the combination of the four motorized motions (X Y Z R_z) with the sensor fixture motion ($R_x R_y$). We mechanically align the stage R_z axis with the sensor fixture origin, so that the sensor fixture motion can be combined into the stage motion ${}^{st}\boldsymbol{J}$ without any intervening unknown $\boldsymbol{F}$.

VI. Optimization Constraints

Combining the magnetic and fixture parameters, we have:

$$\rho = \langle {}^{so}\boldsymbol{L}^{so}, {}^{so}\boldsymbol{M}^{so}, {}^{se}\boldsymbol{L}^{se}, {}^{se}\boldsymbol{M}^{se}, {}^{so}\boldsymbol{F}, {}^{st}\boldsymbol{F}, {}^{se}\boldsymbol{F} \rangle$$

We do not optimize all 54 potential variables simultaneously. One reason is that there are redundant degrees of freedom in the model. These variables can "run away" in opposing directions, to no effect, but preventing convergence (see calibration procedures in supplemental material).

If all of the parameters are free, then the source and sensor coordinate frames are underdefined. We chose to tie the coordinate frames to the source and sensor coils by setting:

$$\begin{aligned} {}^{so}\boldsymbol{L}_{\mathrm{z}}^{so} &= {}^{se}\boldsymbol{L}_{\mathrm{z}}^{se} = 0 \\ {}^{so}\boldsymbol{M}_{\mathrm{z}}^{so} &= [0 \quad 0 \quad {}^{so}k]^T \\ {}^{se}\boldsymbol{M}_{\mathrm{z}}^{se} &= [0 \quad 0 \quad {}^{se}k]^T \end{aligned} \tag{7}$$

This places the source and sensor origins at the Z coils, with coil Z axes aligned to the Z coordinate axes. (The coordinates can also be fixed to any other desired location.) To constrain the rotation about Z, we force the X coil Y component to 0 also:

$$\begin{aligned} {}^{so}\boldsymbol{M}_{\mathrm{x}}^{so} &= [{}^{so}\boldsymbol{m}_{\mathrm{xx}} \quad 0 \quad {}^{so}\boldsymbol{m}_{\mathrm{xz}}]^T \\ {}^{se}\boldsymbol{M}_{\mathrm{x}}^{se} &= [{}^{se}\boldsymbol{m}_{\mathrm{xx}} \quad 0 \quad {}^{se}\boldsymbol{m}_{\mathrm{xz}}]^T \end{aligned}$$

These constraints cause the fixture transforms to absorb any pose offset between the Z coil and the fixture.

The source and sensor gains (moment magnitudes) are not independently measurable; we only observe the product of the two. So we must fix either the source or sensor gain: ${}^{so}k$ or ${}^{se}k$ in (7). We initially force ${}^{so}k = 1$, attributing all of the system gain to the sensor. (This pattern of forced values may be clearer in the Table I sample calibration.)

VII. Pose Solution and Error Correction

Magnetic tracking is an *inverse problem*. We can calibrate a measurement model $\tilde{\boldsymbol{c}}(\boldsymbol{P}, \rho)$ to predict the coupling $\boldsymbol{C}$ we expect to see in sensor pose $\boldsymbol{P}$, but this does not tell us how to solve the inverse problem: given a measurement $\boldsymbol{C}$, what is the best estimate of the pose $\widetilde{\boldsymbol{P}}$? Some measurement models permit a closed-form pose solution using linear algebra[5]–[7], but in general it is again necessary to use nonlinear optimization, this time finding $\widetilde{\boldsymbol{P}}$ which minimizes $\left|\tilde{\boldsymbol{c}}(\widetilde{\boldsymbol{P}}, \rho) - \boldsymbol{C}\right|_2$.

In microsurgery the accuracy requirement is high, but the workspace is inherently constrained. When high accuracy is required over a well-defined workspace, it is possible to measure the error at many poses, and then apply a correction to the pose solution (§III.C). As with the calibration of the measurement magnetic model $\tilde{\boldsymbol{c}}(\boldsymbol{P}, \rho)$, it is important that output correction calibration data cover the intended workspace, but because of its ad-hoc non-physical character it tends to generalize even more poorly outside of the calibration volume than does the magnetic model.

General linear transformation is a simple method for transforming spatial data using matrix multiplication. Since this is implemented by a 4x4 transform matrix it is fairly low order, so will not over-fit the calibration data, and has a stable extrapolation beyond edge of the calibrated volume, degrading gradually. A general transform matrix $\boldsymbol{F}$ differs from a linear homogeneous matrix in that all 16 elements are unconstrained.

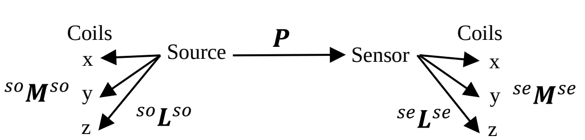

Fig. 2: Pose graph for kinematics of the measurement model. The pose $\boldsymbol{P}$ maps from source coordinates to sensor. Each coil has an independent location $\boldsymbol{l}_i$ and moment $\boldsymbol{m}_i$ within its frame. These are grouped into the $\boldsymbol{L}, \boldsymbol{M}$ matrices.

It is neither orthogonal nor normalized, so is not a rigid body transformation, and can compensate scaling, shearing and trapezoid distortions. An arbitrary point $\boldsymbol{p}_e$ is corrected to $\boldsymbol{p}_c$:

$$\begin{bmatrix} \boldsymbol{p}_f \\ k_n \end{bmatrix} = \boldsymbol{F} \begin{bmatrix} \boldsymbol{p}_e \\ 1 \end{bmatrix}$$

$$\boldsymbol{p}_c = \boldsymbol{p}_f / k_n$$

During calibration we must determine the $\boldsymbol{F}$ which minimizes the mismatch between every calibration point $\boldsymbol{p}_c$ and its ground truth position. General nonlinear optimization can be used, but we have found the simple Direct Linear Transformation (DLT) algorithm to work well here[24]. Only position is corrected using this method. See Table II for examples of the benefit.

VIII. Characterizing Tracker Error

Fig. 3 is a pose diagram of the error kinematics. The pose error of a calibrated EMT is the difference $\boldsymbol{E}(\widetilde{\boldsymbol{P}})$ between the EMT output $\widetilde{\boldsymbol{P}}$ and the true pose $\boldsymbol{P}$. Disregard the stage error and suppose $\boldsymbol{P}$ is known exactly. Then:

$$\begin{aligned} \boldsymbol{P} &= \widetilde{\boldsymbol{P}}\, \boldsymbol{E}(\widetilde{\boldsymbol{P}}) \\ \boldsymbol{E}(\widetilde{\boldsymbol{P}}) &= \widetilde{\boldsymbol{P}}^{-1}\, \boldsymbol{P} \end{aligned} \tag{8}$$

$\boldsymbol{E}(\widetilde{\boldsymbol{P}})$ is itself a 6DOF transform. If the error is computed as a transform matrix by (8) and converted to the eqn. (2) rotation vector format, then the translation and rotation parts represent the error in a somewhat interpretable form. However, it is often preferable to reduce the 6DOF error to translation and rotation magnitudes by taking the vector magnitude of the 3DOF pose parts. This rotation magnitude avoids all difficulties with angle wrapping but does not tell us the error direction.

A. Uncertainty of stage position and tracker measurement

What we wish to determine is the error that can be expected in measurements $\widetilde{\boldsymbol{P}}$ made by the EMT. This typical error is the measurement *standard uncertainty*. Fig. 3 reminds us that we do not know the true $\boldsymbol{P}$, only $\boldsymbol{P}_{\mathrm{stage}}$ constructed using (6), so our actual error computation is:

$$\widetilde{\boldsymbol{E}}(\widetilde{\boldsymbol{P}}) = \widetilde{\boldsymbol{P}}^{-1}\ {}^{so}\boldsymbol{F}\ {}^{so}\boldsymbol{J}\ {}^{st}\boldsymbol{F}\ {}^{st}\boldsymbol{J}\ {}^{se}\boldsymbol{F} \tag{9}$$

${}^{so}\boldsymbol{J}$ and ${}^{st}\boldsymbol{J}$ each have their own standard uncertainties, so $\widetilde{\boldsymbol{E}}(\widetilde{\boldsymbol{P}})$ is an uncertain measurement of the desired $\boldsymbol{E}(\widetilde{\boldsymbol{P}})$. (The fixture transforms $\boldsymbol{F}$ can be regarded as defining the coordinate frames, so have no uncertainty.)

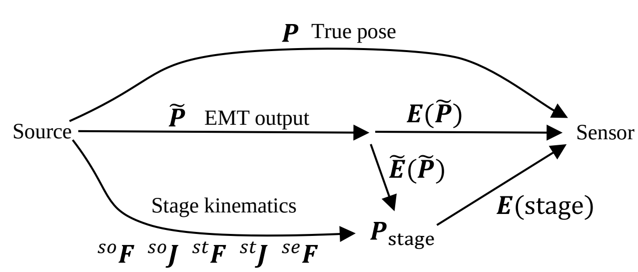

Fig. 3: Pose graph for kinematics of measurement error. $\widetilde{\boldsymbol{P}}$ is the tracker output and $\boldsymbol{E}(\widetilde{\boldsymbol{P}})$ is the error in this output. $\boldsymbol{P}$ is the ideal perfectly accurate pose, which is not directly known. Allowing for stage error, we construct $\boldsymbol{P}_{\mathrm{stage}}$ from the stage kinematics, and compare to $\widetilde{\boldsymbol{P}}$, instead finding the calibration error $\widetilde{\boldsymbol{E}}(\widetilde{\boldsymbol{P}})$. This differs from $\boldsymbol{E}(\widetilde{\boldsymbol{P}})$ by the stage error $\boldsymbol{E}$(stage), so does not represent the full standard uncertainty of the $\widetilde{\boldsymbol{P}}$ measurement.

Let $u(x_i)$ be the standard uncertainty of a measurement x_i, and u_c the *combined standard uncertainty* of a composite measurement, then in standard measurement practice[25]:

$$u_c(x_i + x_j) = u_c(x_i - x_j) = \sqrt{u(x_i)^2 + u(x_j)^2}$$

(For small errors (8) can be regarded as a subtraction.) This Root-Sum-of-Squares (RSS) rule is justifiable supposing that $u(x_i)$ are the standard deviations of independent normally distributed random variables. Because one standard deviation (the combined standard uncertainty) typically captures only about 68% of possible errors, it is common practice to multiply it by a coverage factor (often 2 to 4) to define an expanded uncertainty that encloses the measurement error with higher confidence.

We wish to find the standard uncertainty of $\widetilde{\boldsymbol{P}}$: $u(\widetilde{\boldsymbol{P}})$. Given the true error $\boldsymbol{E}(\widetilde{\boldsymbol{P}})$ we could estimate $u(\widetilde{\boldsymbol{P}})$ as the standard deviation (RMS value) of this error across the test points. Let $u(cal)$ be the RMS standard uncertainty estimate from the available $\widetilde{\boldsymbol{E}}(\widetilde{\boldsymbol{P}})$, eqn. (9). $u(cal)$ alone underestimates $u(\widetilde{\boldsymbol{P}})$, but by including the stage standard uncertainty, the combined uncertainty accounts for this:

$$u(\widetilde{\boldsymbol{P}}) = \sqrt{u(\boldsymbol{P}_{\text{stage}})^2 + u(cal)^2} \tag{10}$$

B. Systematic and random error

Error can be divided into *systematic* (repeatable) error and *random* error. The entire calibration procedure is an attempt to minimize the systematic error in the tracker measurement. The stage has its own systematic and random positioning errors. Since the same stage is used for both calibration and testing, any systematic error in the stage is present in both the calibration and test data. If the calibrated measurement model $\tilde{\boldsymbol{c}}(\boldsymbol{P}, \rho)$ incorporates systematic stage error (such as scale factor), then the true error $\boldsymbol{E}(\widetilde{\boldsymbol{P}})$ becomes correlated with the stage error $\boldsymbol{E}(\text{stage})$. This correlation simultaneously increases the true error $\boldsymbol{E}(\widetilde{\boldsymbol{P}})$ and reduces the measured calibration error $\widetilde{\boldsymbol{E}}(\widetilde{\boldsymbol{P}})$. In the extreme, if random stage error is minimal and $\tilde{\boldsymbol{c}}(\boldsymbol{P}, \rho)$ perfectly models both the magnetics and the systematic stage error then $u(cal)$ from (10) will go to zero. Even so, (10) remains reasonable since then $u(\widetilde{\boldsymbol{P}}) = u(\boldsymbol{P}_{\text{stage}})$.

When choosing the model $\tilde{\boldsymbol{c}}(\boldsymbol{P}, \rho)$ it is important to consider that if ρ has many degrees of freedom (such as with a lookup table) then it will overfit to include the stage error, and this increases the true error even as it reduces the calibration error. If the stage is the only position reference used then this effect is invisible. A lower order model may appear to give worse error, yet actually be more accurate.

If the calibration data has far more DOF than ρ does, then random positioning error will be averaged out in the calibration, while random error during testing will increase $u(cal)$. This is the reverse of the situation with systematic error, since true error $\boldsymbol{E}(\widetilde{\boldsymbol{P}})$ does not increase, but $u(\widetilde{\boldsymbol{P}})$ does.

C. Stage and fixture design

The calibration stage in Fig. 4 provides ± 50mm translation in *x*, *y*, and *z*, and $\pm 100°$ rotation (R_z) and is constructed to minimize magnetic interference with the measurement. The stage position uncertainty is 107 µm, which has little effect on the combined uncertainty (Table II). For details of this uncertainty, the stage construction and characterization see supplement §II. The manual source and sensor fixtures in Fig. 6 create additional 90º rotations of the source and sensor.

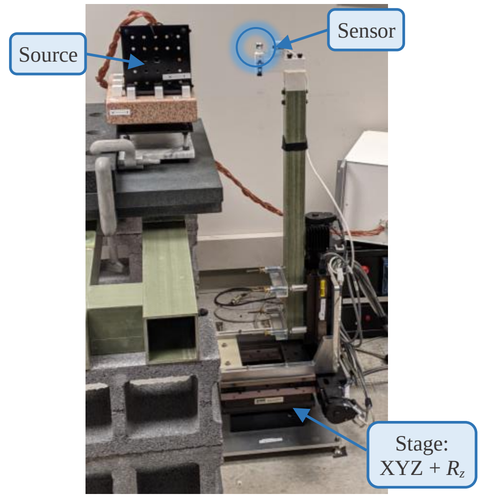

Fig. 4: The stage provides automated motion over a 100mm cube translation workspace, with $\pm 90°$ rotation about Z. The fiberglass pole extends the sensor away from the metal components in the stage so as to avoid interference with the magnetic measurement.

D. Data collection patterns

The calibration and test data can and *should* be different. In any sort of modeling, the model should be evaluated with different data than those used to fit the model. If not, performance may be exaggerated by overfitting. Yet the calibration data must cover the same sort of variation that will be seen in operation (and during testing). Also, in EMT practice, because of the strong constraints of the dipole model, it works well to use smaller data for calibration than testing, which speeds the calibration optimization.

For calibration, the motion pattern we use is a 100 mm cube with a 3x3x3 grid of test points. This cube is offset, with sensor motion $200\text{mm} \leq x \leq 300\text{mm}$. Five Z axis ($R_z$) rotations spaced across $\pm 90^{\circ}$ degrees are also taken at each grid point, giving 135 points in each fixture configuration. The larger test data has a 5x5x5 grid with the same rotations (625 points).

This pattern is collected in each of the 4 rotations of the source fixture and the 3 rotations of the sensor fixture, Fig. 5 and §V.D. In the dipole model (4) the source and sensor response interact linearly. For small deviations from the dipole model the deviation can be decomposed into independent source and sensor effects, so it is not necessary to evaluate the full cross product of source and sensor rotations. Broadly, rotation of the sensor helps to identify the sensor response, while rotations of the source identify the source field. For $\langle {}^{so}\boldsymbol{F}, {}^{se}\boldsymbol{F} \rangle$ identified by Fig. 5 numbering, we test at poses $\langle {}^{so}\boldsymbol{F} \in \{0..3\}, {}^{se}\boldsymbol{F} = 0 \rangle$ and $\langle {}^{so}\boldsymbol{F} = 0, {}^{se}\boldsymbol{F} \in \{1,2\} \rangle$, giving six manual fixture positions with the 625 stage poses repeated at each, for a full test dataset of 3750 points. So even with this coarse grid spacing it is impractical to densely sample across a

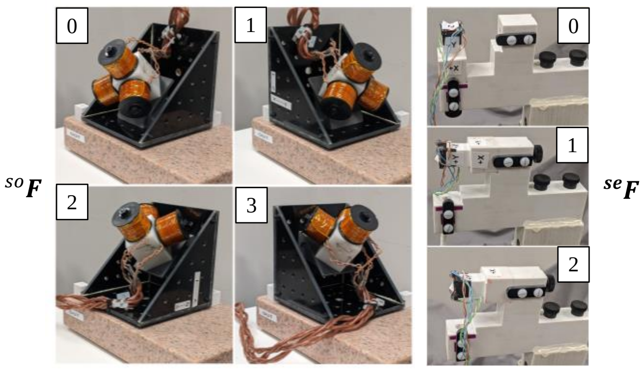

Fig. 5: Source and sensor fixtures are manually positioned during data collection giving test patterns with rotation on all source and sensor axes. Each sensor axis is successively aligned with the stage R_z rotational axis, so is tested at multiple rotations over $\pm 90°$. Source motion is only via the source fixture 90° rotations. Ideally each of the 6 faces of the source would be fixtured toward the sensor, but this is not mechanically possible with this fixture.

6DOF workspace.

E. Tested hardware

Source drive and sensor readout are provided by the ILEMT open source EMT reference design. This hardware is optimized for low noise and high measurement speed rather than cost [2]. While a detailed discussion of ILEMT design and performance is outside the scope of this paper, typical performance with the tested sources-sensor combinations is $< 3\mu$m RMS position noise at 1500 samples/sec (500 Hz bandwidth), at $r = 200$ mm. This a frequency-domain multiplexed AC design, where each source coil continuously emits a sinusoidal field at a distinctive frequency (7.5 kHz, 10.5 kHz and 13.5 kHz).

The coupling recorded for each tested pose is the mean of 1500 measurements, so random variation in the coupling due to broadband noise is negligible. Non-repeatable variation in the coupling is dominated by drift. The collection of calibration and test data takes approximately 8 hours. Drift over this time degrades calibration accuracy, so we powered up the tracker at least 2 hours prior to data collection (supplement §IV.)

The only sensor used is the Premo Magnetics 3DCC10-A-0600J. While this is a fairly large 17x15x14 mm sensor (which improves the signal-to-noise ratio), it is still much smaller than the source, so the sensor response is much closer to a dipole than the source field is. We have characterized performance with a variety of sensors, and the sensor has always had less effect on the accuracy than the source did. Using only one sensor simplifies our presentation.

Results are given for two different source designs: *concentric* and *dipole-approximating* (Fig. 7). The concentric source follows the most common design for EMTs: three orthogonal coils wound around a ferromagnetic cube core. The dipole-approximating source uses three separate coils, each with the geometry from Ref. [26], which gives a close approximation of a dipole field. Because the coils are distinct, this source is non-concentric. Both sources are hand-wound prototypes, and likely have less ideal fields than might be seen with a production source.

For both sources the coil diameter is $d \approx 40$mm, which with a testing radius of $150\text{mm} < r < 250\text{mm}$ gives $3.75 < r/d < 6.25$. Operating at this relatively short range reduces measurement noise, but also increases the difficulty of accurate calibration because the deviation from a dipole model can be significant [27].

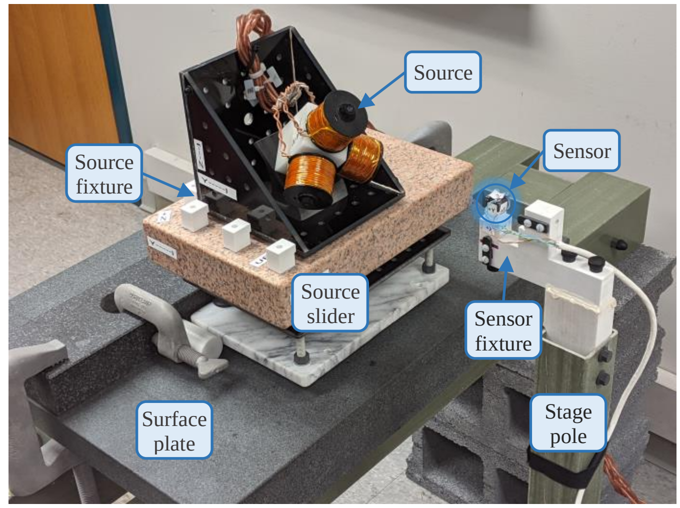

Fig. 6: The source and sensor fixtures permit precise 90-degree rotations of the source and sensor, which enables testing across the entire workspace, supplementing the limited translation and single rotation axis of the stage. The surface plate permits large precise planar and straight-line sliding motion.

These sources and sensor were chosen to meet ILEMT requirements, but from the viewpoint of calibration techniques, many different source and sensor types might be used. The purpose of calibration is to model whatever combination of source and sensor is chosen. To operate at a longer range a proportionally larger source would be used, preserving the signal-to-noise, with the same r/d ratio giving similar accuracy to the dipole approximation.

F. Calibration

These results use three calibration types: the *default* condition uses the general dipole magnetic model, while *concentric* restricts the source and sensor coil positions to their coordinates origin (${}^{so}L^{so} = {}^{se}L^{se} = 0$), and *corrected* applies the DLT output correction to XYZ (§VII). Also, the data may include source fixture motion, or may only use a single source fixture (sensor always on the same side of the source).

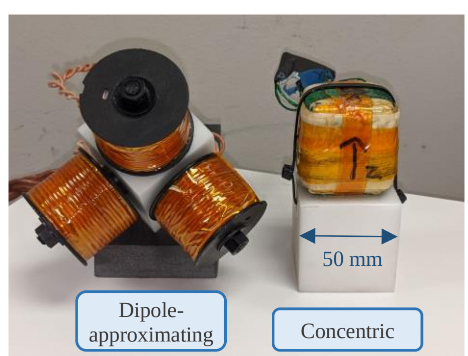

Fig. 7: The concentric source has three orthogonal windings around a single ferrite cube core, giving a high field strength for its size. The dipole-approximating source uses three air-core coils with a geometry that gives a field more closely resembling a magnetic dipole at close ranges.

TABLE I
SAMPLE CALIBRATION (DIPOLE APPROXIMATING SOURCE)

Source position (mm)			Source moment			Sensor position (mm)			Sensor moment		
X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z
45.266	-0.514	0.0	0.950	1.0E-02	0.0	0.144	0.051	0.0	0.158	4.2E-04	0.0
1.249	45.380	0.0	0.0	0.947	0.0	0.110	0.116	0.0	0.0	0.157	0.0
-43.764	-42.830	0.0	-2.6E-02	1.7E-03	1.0	-0.223	-0.185	0.0	2.4E-03	1.9E-03	0.161

TABLE II
ERROR VS. CALIBRATION TYPE AND SOURCE DESIGN

Calibration type	Dipole approximating source						Concentric source					
	XYZ (mm)			RxRyRz (degrees)			XYZ (mm)			RxRyRz (degrees)		
	$u(cal)$	$u(\tilde{P})$	Max	$u(cal)$	$u(\tilde{P})$	Max	$u(cal)$	$u(\tilde{P})$	Max	$u(cal)$	$u(\tilde{P})$	Max
(default)	0.271	0.292	0.747	0.210	0.270	0.529	0.416	0.429	1.022	0.273	0.321	0.637
Corrected	0.178	0.208	0.615	0.210	0.270	0.529	0.263	0.284	0.720	0.273	0.321	0.637
Concentric	28.33	28.326	73.45	20.26	20.26	56.84	0.905	0.911	2.492	0.508	0.536	1.415

$u(cal)$ is the calibration error, while $u(\tilde{\boldsymbol{P}})$is the measurement combined standard uncertainty, including stage standard uncertainty, see eqn (10).

Table I is a sample calibration for the dipole-approximating source. These are the $\boldsymbol{L}$ and $\boldsymbol{M}$ matrices from (5), with coils as columns. Note the values forced to 0.0 and 1.0 to avoid excess degrees of freedom (see §VI). This source is highly non-concentric, with 45mm offset. The sensor is designed to be concentric, but calibration finds position offsets similar in magnitude to the position accuracy we obtain below. Source and sensor non-orthogonality are small, but still significant. The source X Z deviation is 1.6°, which would correspond to a 5.5 mm error at 200 mm. Both the default and concentric calibration methods model axis non-orthogonality.

G. Position and rotation error

The 6DOF pose error is found using (9). Fig. 8 shows translation error vectors across the workspace, with several rotations. This plot is a useful visualization of patterns of error, but if the measurement model $\tilde{\boldsymbol{c}}(\boldsymbol{P},\rho)$ successfully absorbs any such pattern, then a statistical summary is more meaningful.

Table II shows the rotation and translation error of the two calibrated sources as the calibration type is varied. The errors reported are RMS and maximum vector magnitudes of the translation error (XYZ) and the rotation vector of the rotation error ($R_x R_y R_z$). Here both the calibration and test data were taken with three sensor fixtures, but no source fixture rotation. With the general calibration, the dipole-approximating source gives significantly better accuracy than the concentric source (empirical support for [26]). Since this source is not concentric, the accuracy is terrible using a concentric calibration model. For the concentric source, the concentric model has 2x the error of the general dipole model, which for some uses may be an acceptable sacrifice in order to enable a more efficient pose solution. For both sources the section VII output correction significantly improves the XYZ accuracy. The combined measurement uncertainty from (10), which allows for stage error, is only slightly greater than the RMS calibration error $u(cal)$ alone, so system accuracy is not being significantly degraded by stage error.

H. Effect of calibration DOF on accuracy

What happens when the calibration data does not cover the operation workspace? While we understand that such mismatch is not good, it may not be appreciated how severe the accuracy degradation may be. Table III shows the effect of adding dimensions of motion, present in either the calibration data, the test data, or both. For *translation only* there is no sensor rotation at all (not even R_z), then we add all-axis sensor rotation, and finally all-axis source rotation. These results are with the concentric source and *default* calibration. With the dipole-approximating source the pose solution often failed to converge with source fixtures, so is not shown. In Table III we see that calibrating across a higher dimensional workspace increases performance on that workspace, but at the cost of lower accuracy in a more restricted workspace.

Since sensor motion is usually 6DOF, it is at a minimum necessary to calibrate and test all axes of sensor rotation. When test dimensionality is higher than for calibration, then the error is quite high. Comparing the '*' entries in Table III, the worst effect of inadequate calibration is 4.5x. Also, if the EMT is to operate in any hemisphere of the source then calibration and testing must be done in the different source orientations. (This

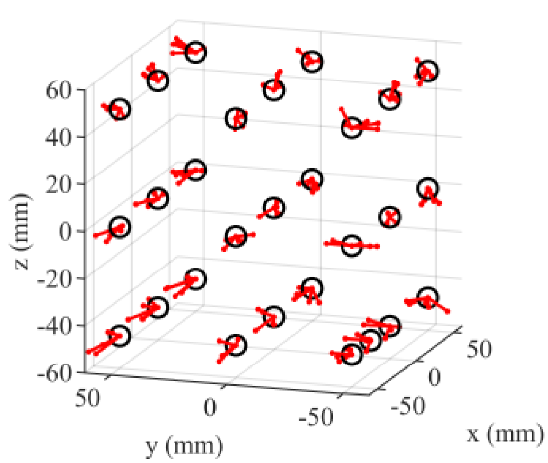

Fig. 8: Translation calibration error across the workspace, under R_z sensor rotation. The error vectors are exaggerated 35x. Each point has 5 error vectors because there are 5 rotations. Note the effect of rotation on translation error. (Data size reduced for illustration, actual test data has many more poses.)

TABLE III
CALIBRATION ERROR VS. SENSOR AND SOURCE ROTATION (CONCENTRIC SOURCE, UNCORRECTED)

	Test data: translation only				Test data: sensor fixtures				Test data: source + sensor fixtures			
	XYZ (mm)		Rxyz (degrees)		XYZ (mm)		Rxyz (degrees)		XYZ (mm)		Rxyz (degrees)	
Calibrate data:	RMS	Max	RMS	Max	RMS	Max	RMS	Max	RMS	Max	RMS	Max
Translation only	**0.231**	**0.613**	**0.109**	**0.308**	1.862*	5.514	1.088	2.809	4.122	11.227	1.793	5.672
Sensor fixture	0.396	0.977	0.276	0.501	**0.416*†**	**1.022**	**0.273**	**0.637**	4.580	12.256	1.826	5.222
Both fixtures	0.844	1.700	0.399	0.888	0.895	2.256	0.540	1.382	**2.306†**	**6.121**	**1.100**	**3.395**

Bold data are conditions where calibration and test workspace dimensionality are the same, giving lowest error for each test condition. Compare the two entries marked '*' for the worst relative effect of calibrating without matched dimensionality, '†' for the reduced accuracy seen in the presence of source fixture rotation even when the calibration and test dimensions are matched. See §VIII.G.

is often unnecessary in surgical applications.) Comparing the '†' entries, adding source fixture rotation has a large 5.5x effect, even with matched calibration. (This may be due to irregularities in the hand-wound source.)

I. Linearity

Measuring position error on the sort of wide-spaced data used in calibration does not directly test what relative accuracy can be expected during fine scale motion. In the grid data used in the above testing, the XYZ increment is 2.5 cm, while the position error may be as much as 100x smaller. It would require an entirely impractical number test points to densely cover the 6DOF workspace. We expect a degree of local *linearity*: that the error will vary smoothly across the workspace, so that a small motion will experience less error than the worst-case full-workspace error, but we have little idea the degree to which this is so. To test this we abandon grid sampling, and instead make dense sweeps, varying only one pose component at a time. This doesn't solve the problem that it is impossible to test all poses, but does densely test *some* part of the workspace, and also enables a useful decomposition of the error.

One shortcoming of the sparse grid error measurement is that it does not characterize the measurement cross-coupling. Any motion of the sensor will result in a change of all 6 measured DOF. We would hope that the largest change represents the true motion, but there are also spurious changes along directions that the sensor did not actually move in (Fig. 9). To simplify presentation, we split these errors into *direct error* and *cross error*, and show the vector magnitude of each. Direct error is the XYZ error caused by XYZ motion (or $R_x R_y R_z$ error caused by $R_x R_y R_z$ motion), while cross error is cross coupling between rotation and translation.

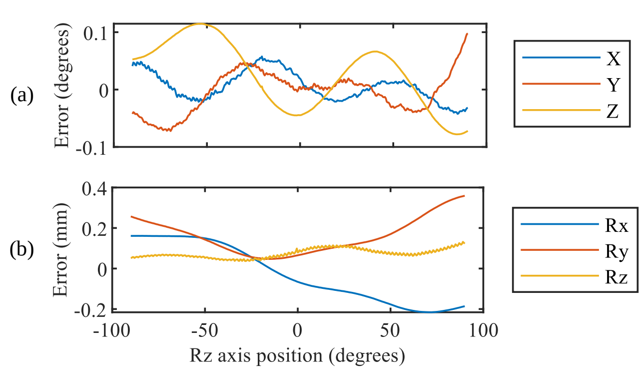

Fig. 9: Actual motion on a single axis (here Rz) creates measurement errors on *all* axes. (a) is the *cross error* (here mm translation error caused by rotation), while (b) is the *direct error* (degrees of rotation error caused by rotation). This is one example from the six possible such plots. See instead Fig. 10 which summarizes this data as rotation and translation error vector magnitudes.

Fig. 10 shows the absolute error and nonlinearity resulting from sensor motion sweeps on all 6 axes. Measurements are at 0.5 mm, 0.5 degree increments over 100 mm, 180 degrees. This uses the dipole approximating coil, without source fixtures (as in Table II corrected condition). We quantify the linearity across the workspace by the derivative of the error (the magnitude of the error gradient). To minimize noise, we differentiate using a Savitzky-Golay filter. The linearity at a point tells us the relative error that can be expected across a small motion. The direct nonlinearity is dimensionless and expressed as a percentage of the motion. While this resembles a scale factor error, it includes off-axis motion as well. Cross nonlinearity is the magnitude of the gradient of translation error with respect to rotational motion, and vice versa, with units mm/° or °/mm.

In Table IV each Fig. 10 curve is summarized by RMS and max, and then these measures are combined by max across the curves (driving axes) in each plot. Absolute error is computed as in Table II, but is now significantly lower. This is mainly because the sweep data only varies one axis, so is on average closer to the center of the workspace, where error is lower. Note how the magnitude of the rotation-to-translation cross error is similar to the direct translation error. This reinforces the point that we cannot characterize EMT translation accuracy by using a translation-only test pattern.

IX. DISCUSSION

As the calibration component of the ILEMT open EMT design we only aimed to match the static accuracy of commercial EMTs. While we have not compared commercial trackers under the same conditions, published results such as [19] suggest we have met and even surpassed this goal. We have achieved this using fairly simple and general methods, which is particularly surprising given our use of handmade sources.

TABLE IV
CROSS COUPLING AND NONLINEARITY DURING AXIS SWEEPS

		Absolute error		Nonlinearity	
		XYZ	RxRyRz	XYZ	RzRyRz
		mm	degrees	Percent	
Direct	RMS	0.093	0.232	0.32%	0.45%
	max	0.152	0.411	0.55%	0.62%
Cross	RMS	0.148	0.099	0.0029	0.0048
	max	0.230	0.182	0.0032	0.0085
		degrees	mm	°/mm	mm/°

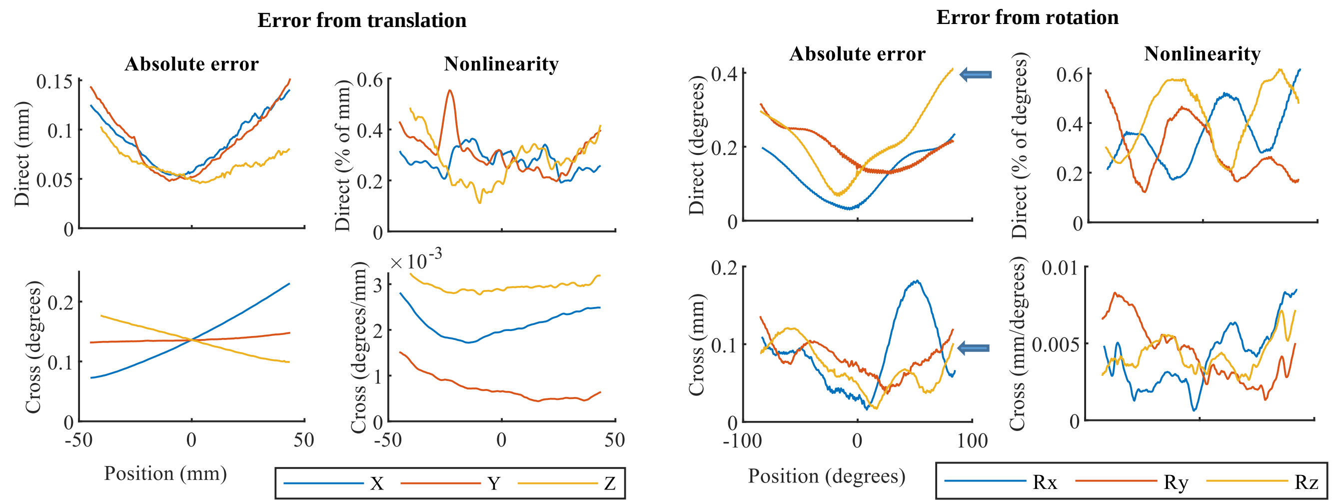

Fig. 10: Error and nonlinearity as each axis is swept across the workspace, one at a time. The absolute error is the vector magnitude of the 3DOF rotation or translation error caused by motion on the axis in the legend. Nonlinearity is the magnitude of the 3DOF derivative (gradient) of the error. As magnitudes, these plots are all nonnegative. The "U" shape of absolute error represents the response passing near an error minimum, not curvature in the measured path. In this figure the data in Fig. 9 has been reduced to two curves (see arrows): the absolute error from rotation (Rz), which is the vector magnitudes of Fig. 9 (b) and Fig. 9 (a).

It is of course unknown whether these methods differ from commercial EMT practice, but the broad principle of fitting a magnetic model using nonlinear optimization is everywhere in the open literature. Aside from the reusable and freely available implementation, our most important additions to EMT practice are the introduction of a general matrix formulation based on linear transforms and linear homogeneous coordinates. We also closely examined characterization methods, including practical details of stage construction and analysis of the effects of stage error on the accuracy of the calibration/test procedure. Characterization of tracker linearity and the rotation/translation cross coupling are particularly important for understanding the performance achievable during small motions.

References

[1] F. Raab, E. Blood, T. Steiner, and H. Jones, "Magnetic Position and Orientation Tracking System," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-15, no. 5, pp. 709–718, Sep. 1979, doi: 10.1109/TAES.1979.308860.

[2] R. A. Maclachlan *et al.*, "Electromagnetic tracker for active handheld robotic systems," in *Proc IEEE Sensors Conf*, Jan. 2016, p. accepted, doi: 10.1109/ICSENS.2016.7808415.

[3] Robert A. MacLachlan, "In-Loop ElectroMagnetic Tracker," *Open Science Framework*, 2024. https://osf.io/6fwjk/ (accessed Feb. 02, 2024).

[4] H. A. Jaeger *et al.*, "Anser EMT: the first open-source electromagnetic tracking platform for image-guided interventions," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 6, pp. 1059–1067, 2017, doi: 10.1007/s11548-017-1568-7.

[5] W. Kim, J. Song, and F. C. Park, "Closed-Form Position and Orientation Estimation for a Three-Axis Electromagnetic Tracking System," *IEEE Trans. Ind. Electron.*, vol. 65, no. 5, pp. 4331–4337, 2018, doi: 10.1109/TIE.2017.2760244.

[6] C. Hu, S. Song, X. Wang, M. Q.-H. H. Meng, and B. Li, "A novel positioning and orientation system based on three-axis magnetic coils," *IEEE Trans. Magn.*, vol. 48, no. 7, pp. 2211–2219, Jul. 2012, doi: 10.1109/TMAG.2012.2188537.

[7] M. Teixeira Silva, E. Santos, L. Batista, and J. Araújo, "Alternative Analytical Solution for Position and Orientation in Electromagnetic Motion Tracking Systems," *WSEAS Trans. Syst.*, vol. 16, pp. 225–233, Aug. 2017.

[8] K. O'Donoghue, "Electromagnetic tracking and steering for catheter navigation," University College Cork, 2014.

[9] T. Bien and G. Rose, "Algorithm for calibration of the electromagnetic tracking system," in *Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics*, 2012, pp. 85–88, doi: 10.1109/BHI.2012.6211512.

[10] A. Plotkin, V. Kucher, Y. Horen, and E. Paperno, "A New Calibration Procedure for Magnetic Tracking Systems," *IEEE Trans. Magn.*, vol. 44, no. 11, pp. 4525–4528, Nov. 2008, doi: 10.1109/TMAG.2008.2003056.

[11] Xin Ge, Yifeng Wang, Ning Ding, Xiaomei Wu, Yuanyuan Wang, and Zuxiang Fang, "An Electromagnetic Tracking Method Using Rotating Orthogonal Coils," *IEEE Trans. Magn.*, vol. 48, no. 12, pp. 4802–4810, Dec. 2012, doi: 10.1109/TMAG.2012.2203917.

[12] M. Schneider and C. Stevens, "Development and testing of a new magnetic-tracking device for image guidance," in *Medical Imaging 2007: Visualization and Image-Guided Procedures*, Mar. 2007, vol. 6505, pp. 65090I-65090I–11, doi: 10.1117/12.713249.

[13] J. Hummel, M. Figl, M. Bax, R. Shahidi, H. Bergmann, and W. Birkfellner, "Evaluation of dynamic electromagnetic tracking deviation," Feb. 2009, p. 72612U, doi: 10.1117/12.813645.

[14] D. Kügler *et al.*, "High-precision evaluation of electromagnetic tracking," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 7, pp. 1127–1135, 2019, doi: 10.1007/s11548-019-01959-5.

[15] D. D. Frantz, A. D. Wiles, S. E. Leis, and S. R. Kirsch, "Accuracy assessment protocols for electromagnetic tracking systems," *Phys. Med. Biol.*, vol. 48, no. 14, pp. 2241–2251, Jul. 2003, doi: 10.1088/0031-9155/48/14/314.

[16] C. Nafis, V. Jensen, L. Beauregard, and P. Anderson, "Method for estimating dynamic EM tracking accuracy of surgical navigation tools," in *Proc.SPIE*, Mar. 2006, vol. 6141, p. 61410K, doi: 10.1117/12.653448.

[17] A. M. Franz, T. Haidegger, W. Birkfellner, K. Cleary, T. M. Peters, and L. Maier-Hein, "Electromagnetic tracking in medicine -A review of technology, validation, and applications," *IEEE Trans. Med. Imaging*, vol. 33, no. 8, pp. 1702–1725, 2014, doi: 10.1109/TMI.2014.2321777.

[18] J. B. Hummel *et al.*, "Design and application of an assessment protocol for electromagnetic tracking systems," *Med. Phys.*, vol. 32, no. 7, p. 2371, 2005, doi: 10.1118/1.1944327.

[19] E. Lugez, H. Sadjadi, D. R. Pichora, R. E. Ellis, S. G. Akl, and G. Fichtinger, "Electromagnetic tracking in surgical and interventional environments: usability study," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, no. 3, pp. 253–262, 2015, doi: 10.1007/s11548-014-1110-0.

[20] V. V. Kindratenko, "A survey of electromagnetic position tracker calibration techniques," *Virtual Real.*, vol. 5, no. 3, pp. 169–182, Sep. 2000, doi: 10.1007/BF01409422.

[21] W. Birkfellner, F. Watzinger, F. Wanschitz, R. Ewers, and H. Bergmann, “Calibration of tracking systems in a surgical environment,” *IEEE Trans. Med. Imaging*, vol. 17, no. 5, pp. 737–742, 1998, doi: 10.1109/42.736028.

[22] F. Attivissimo, A. D. Nisio, A. M. L. Lanzolla, and M. A. Ragolia, “Analysis of Position Estimation Techniques in a Surgical EM Tracking System,” *IEEE Sens. J.*, vol. 21, no. 13, pp. 14389–14396, 2021, doi: 10.1109/JSEN.2020.3042647.

[23] P. Corke, *Robotics, Vision and Control*, 2nd ed., vol. 118. Cham: Springer International Publishing, 2017.

[24] I. E. Sutherland, “Three-dimensional data input by tablet,” *Proc. IEEE*, vol. 62, no. 4, pp. 453–461, 1974.

[25] G. H. White, “Basics of estimating measurement uncertainty,” *Clin. Biochem. Rev.*, vol. 29 Suppl 1, no. Suppl 1, pp. S53–S60, Aug. 2008, [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/18852859.

[26] E. Paperno and A. Plotkin, “Cylindrical induction coil to accurately imitate the ideal magnetic dipole,” *Sensors Actuators, A Phys.*, vol. 112, no. 2–3, pp. 248–252, 2004, doi: 10.1016/j.sna.2004.01.004.

[27] T. Schroeder, “An accurate magnetic field solution for medical electromagnetic tracking coils at close range,” *J. Appl. Phys.*, vol. 117, no. 22, pp. 2–9, Jun. 2015, doi: 10.1063/1.4922667.