

## AUTHOR QUERIES

### AUTHOR PLEASE ANSWER ALL QUERIES

**PLEASE NOTE:** We cannot accept new source files as corrections for your article. If possible, please annotate the PDF proof we have sent you with your corrections and upload it via the Author Gateway. Alternatively, you may send us your corrections in list format. You may also upload revised graphics via the Author Gateway.

Carefully check the page proofs (and coordinate with all authors); additional changes or updates WILL NOT be accepted after the article is published online/print in its final form. Please check author names and affiliations, funding, as well as the overall article for any errors prior to sending in your author proof corrections.

AQ:1 = At this time IEEE cannot validate the membership grade for Ralph L. Hollis.

In-order to do this, please provide the IEEE Account ID and/or email address associated with Ralph L. Hollis in the annotated PDF with your corrections.

AQ:2 = Please provide a graphical abstract (GA) for your article. A GA is now required for IEEE Sensors Journal articles. The GA must be different from any image in your accepted article and should represent the article as a whole. If you submit a video as your graphical abstract, please make sure to include an overlay image. Overlay images are usually a screenshot of your video that best represents the video. This is for readers who may not have access to video-viewing software. The overlay image will be used in the pdf and also in front of the video file. Please also provide a caption for the GA. For more information on graphical abstracts, please see the IEEE Author Center (<https://journals.ieeeauthorcenter.ieee.org/create-your-ieee-journal-article/prepare-supplementary-materials/developing-a-graphical-abstract/>).

AQ:3 = Please confirm or add details for any funding or financial support for the research of this article.

AQ:4 = Please confirm whether the edits made in the current affiliation of all the authors are correct.

AQ:5 = Please confirm whether the edits made in the sentence “This is a physical model...” are correct.

AQ:6 = Please check and confirm whether “X, Y, Z, x, y, and z” and superscript “T” present throughout the article should be set in italics or roman.

AQ:7 = The in-text citations of Figs. 1–6 are out of order and are not sequential.

Please update in-text citations so that all figures are cited in sequential order.

AQ:8 = Please check and confirm whether subscripts “x, y, and z” present throughout the article should be set in italics or roman.

AQ:9 = Please provide the report no. for Ref. [8].

AQ:10 = Please provide the publisher name and publisher location for Ref. [12].

AQ:11 = Please confirm the edits made to the biography of all authors.

AQ:12 = Please provide the year of completion when the authors Claudia Pelle received the master’s degree and the Ralph L. Hollis received the Ph.D. degree.

AQ:13 = Please provide a better/higher quality image for the author “Ralph L. Hollis.”

AQ:14 = Please provide the location (city, country) for Thomas J. Watson Research Center.

# Calibration and Characterization of Electromagnetic Position and Orientation Trackers

AQ:1 Robert A. MacLachlan<sup>ID</sup>, Member, IEEE, Claudia Pelle, Ralph L. Hollis, Fellow, IEEE, Elena De Momi<sup>ID</sup>, Senior Member, IEEE, and Cameron N. Riviere<sup>ID</sup>, Senior Member, IEEE

AQ:2 1     **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 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  $208 \mu\text{m}/0.27^\circ$  uncertainty, an accuracy exceeding published values for commercial electromagnetic trackers (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).

AQ:3 18     **Index Terms**—Magnetic model fitting, magnetic position and orientation tracking, position measurement.

PLACE  
GRAPHICAL ABSTRACT  
HERE

## I. INTRODUCTION

AQ:2 ELECTROMAGNETIC 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 for some uses such as surgical guidance.

AQ:3 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,

AQ:4 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 the 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 in-loop electromagnetic tracker (ILEMT), 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

Received 6 August 2024; revised 9 January 2025; accepted 10 January 2025. This work was supported in part by U.S. National Institutes of Health under Grant R01EB000526, Grant R01EB024564, and Grant R01HL105911. The associate editor coordinating the review of this article and approving it for publication was Prof. Pierluigi Salvo Rossi. (Corresponding author: Cameron N. Riviere.)

Robert A. MacLachlan, Ralph L. 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, 20133 Milan, Italy.

This article has supplementary downloadable material available at <https://doi.org/10.1109/JSEN.2025.3533436>, provided by the authors.

Digital Object Identifier 10.1109/JSEN.2025.3533436

of calibrating with different degrees of freedom that 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 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].

## 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  $n_{\text{so}}$  distinctly modulated electromagnet source coils and the sensor measures the local magnetic field along  $n_{\text{se}} \geq 1$  independent axes. We assume without loss of generality that  $n_{\text{so}} = n_{\text{se}} = 3$ , but other combinations are possible. For each pose, each source field experiences a characteristic magnetic coupling into each sensor coil, giving  $n_{\text{so}}n_{\text{se}}$  distinct couplings, which can be represented as the coupling matrix  $\mathbf{C}$ , where  $C_{jk}$  is the signed coupling from source coil  $j$  into sensor coil  $k$ . The sign represents the 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]. However, 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 nonconcentric 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. Calibration Problem

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

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

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

where  $|\mathbf{A}|_2$  is the matrix two-norm

$$\sqrt{\sum_j \sum_k (A_{jk})^2}.$$

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  $\tilde{\mathbf{c}}(\mathbf{P}_i, \rho)$  are unknown and the initial value for  $\rho$  is inaccurate. Normalizing by the measured coupling magnitude  $|\mathbf{c}(\mathbf{P}_i)|_2$  gives each point equal weight during optimization, despite 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 nonmetallic 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 3-D 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 [12], [15], [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 the 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 0.5–1-mm rms at a range of 200–400 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 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. O'Donoghue [8], Bien and Rose [9], and Attivissimo et al. [22] have more specifics on the model parameters, 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 in the following).

The measured pose  $\mathbf{P}$  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 6-DOF space: translation with rotation. Linear homogeneous coordinates are a particular representation of a transform

$$\mathbf{T} = \begin{bmatrix} \hat{x}_1 & \hat{y}_1 & \hat{z}_1 & x \\ \hat{x}_2 & \hat{y}_2 & \hat{z}_2 & y \\ \hat{x}_3 & \hat{y}_3 & \hat{z}_3 & z \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad 229$$

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

$$\begin{aligned} \mathbf{p}' &= \mathbf{T}\mathbf{p} \\ \mathbf{v}' &= \mathbf{T}\mathbf{v}. \end{aligned} \quad 239 \quad 240$$

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:  $\mathbf{T}_2\mathbf{T}_1\mathbf{p}$  and so on.

Pose graphs such as Figs. 2 and 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 six 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, and 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

$$\mathbf{r} = \langle R_x, R_y, R_z \rangle. \quad 258$$

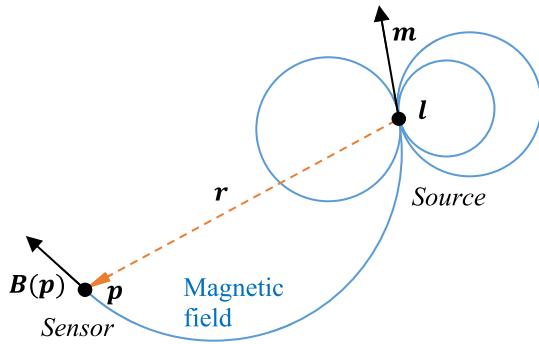
This is closely related to the axis-angle representation  $\langle \hat{r}_x, \hat{r}_y, \hat{r}_z, \theta \rangle$ , where  $\hat{r}$  is the unit vector for the axis of rotation and  $\theta$  is the rotation in radians. The corresponding rotation vector is  $\mathbf{r} = \hat{r}/\theta$ . The complete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

$$\mathbf{P} = \langle x, y, z, R_x, R_y, R_z \rangle. \quad 264$$

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

## V. MEASUREMENT MODEL

One way to appreciate the challenge of measurement modeling is that we expect to achieve an accuracy that is much smaller than the size of the source (or sensor). In the tested ILEMAT configuration, the source is 40 mm and the sensor is 15 mm, while the desired accuracy is  $<400 \mu\text{m}$ . Due to the physics of magnetism, this is not as problematic as 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.



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

### 277 A. Dipole Approximation

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

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

$$288 \quad \mathbf{B}(\mathbf{p}) = \mu \frac{1}{|\mathbf{r}|^3} [3(\mathbf{m} \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}} - \mathbf{m}] \quad (3)$$

289 where  $\mathbf{r} = \mathbf{p} - \mathbf{l}$ ,  $\hat{\mathbf{r}} = (\mathbf{r}/|\mathbf{r}|)$ , and  $\mu$  is the permeability of  
290 the medium.  $\mathbf{B}(\mathbf{p})$  is a vector, which each individual sensor  
291 coil measures as a scalar voltage  $v_i(\mathbf{p})$  along some sensor  
292 moment  $\mathbf{m}_i$ . The dipole form of the sensor response is the  
293 scalar product of the sensor moment with the field

$$294 \quad v_i(\mathbf{p}) = k_g \mathbf{B}(\mathbf{p}) \cdot \mathbf{m}_i \quad (4)$$

295 where  $k_g$  is a gain constant that in practice, we absorb into  
296 the calibrated source and sensor moments, see Section VI.

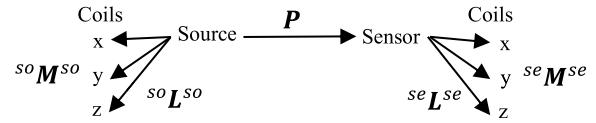
### 297 B. Coordinate Notation

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

$$306 \quad \text{part } \mathbf{p}_{\text{index}}^{\text{coordinates}}$$

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

309 We will index the source and sensor coils by x, y, and z.  
310 Here, “x” is simply a name, not a variable. This is intuitive in  
311 the common case the source and sensor coils roughly coincide with the axes of the corresponding coordinate system.  
312 We represent the dipole parameters as two  $3 \times 3$  matrices



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

**L** and **M**, where the coil locations  $\mathbf{l}_x$ ,  $\mathbf{l}_y$ , and  $\mathbf{l}_z$  and moments  $\mathbf{m}_x$ ,  $\mathbf{m}_y$ , and  $\mathbf{m}_z$  are concatenated as column vectors

$$316 \quad \mathbf{L} = [\mathbf{l}_x \quad \mathbf{l}_y \quad \mathbf{l}_z] \quad \mathbf{M} = [\mathbf{m}_x \quad \mathbf{m}_y \quad \mathbf{m}_z]. \quad 317$$

319 The source parameters are  ${}^{so}\mathbf{L}$  and  ${}^{so}\mathbf{M}$ , and the sensor  
320 has  ${}^{se}\mathbf{L}$  and  ${}^{se}\mathbf{M}$ . Since these quantities move with the  
321 part, the calibration parameters are determined in the part’s  
322 coordinate system, e.g.,  ${}^{se}\mathbf{L}^{se}$ , but when transformed to source  
323 coordinates, the matrix becomes  ${}^{se}\mathbf{L}^{so}$ .

### 324 C. Multidimensional Measurement Model

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

$$332 \quad {}^{se}\mathbf{L}^{so} = \mathbf{P}^{se} \mathbf{L}^{se}, \quad {}^{se}\mathbf{M}^{so} = \mathbf{P}^{se} \mathbf{M}^{se}.$$

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

### 335 D. Fixture Transforms

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

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

352 Specifically, we need to model the ground truth kinematics  
353 so that, by independent means, we know the correct tracker  
354 output  ${}^{se}\mathbf{P}^{so}$ . This is done by a sequential combination of  
355 fixture transforms  $\mathbf{F}$  and known motions  $\mathbf{J}$ .  $\mathbf{J}$  can be thought  
356 of as the joints in a robot arm, and  $\mathbf{F}$  is the fixed link that  
357 connects those joints

$$358 \quad {}^{se}\mathbf{P}^{so} = {}^{so}\mathbf{F}^{so} \mathbf{J}^{st} \mathbf{F}^{st} \mathbf{J}^{se} \mathbf{F}. \quad 359$$

360  $\mathbf{J}$  is known and is varied during the calibration data  
361 collection, while  $\mathbf{F}$  is constants that are initially only

**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

imprecisely known. The “stage motion”  ${}^{\text{st}}\mathbf{J}$  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 with the stage motion  ${}^{\text{st}}\mathbf{J}$  without any intervening unknown  $\mathbf{F}$ .

## VI. OPTIMIZATION CONSTRAINTS

Combining the magnetic and fixture parameters, we have

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

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 the 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

$$\begin{aligned} {}^{\text{so}}\mathbf{L}_z^{\text{so}} &= {}^{\text{se}}\mathbf{L}_z^{\text{se}} = 0 \\ {}^{\text{so}}\mathbf{M}_z^{\text{so}} &= [0 \quad 0 \quad {}^{\text{so}}k]^T \\ {}^{\text{se}}\mathbf{M}_z^{\text{se}} &= [0 \quad 0 \quad {}^{\text{se}}k]^T. \end{aligned} \quad (7)$$

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

$$\begin{aligned} {}^{\text{so}}\mathbf{M}_x^{\text{so}} &= [{}^{\text{so}}\mathbf{m}_{xx} \quad 0 \quad {}^{\text{so}}\mathbf{m}_{xz}]^T \\ {}^{\text{se}}\mathbf{M}_x^{\text{se}} &= [{}^{\text{se}}\mathbf{m}_{xx} \quad 0 \quad {}^{\text{se}}\mathbf{m}_{xz}]^T. \end{aligned}$$

These constraints cause the fixture 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. Therefore, we must fix either the source or sensor gain:  ${}^{\text{so}}k$  or  ${}^{\text{se}}k$  in (7). We initially force  ${}^{\text{so}}k = 1$ , attributing all of the system gain to the sensor. (This pattern of forced values may be clearer in Table I.)

## VII. POSE SOLUTION AND ERROR CORRECTION

Magnetic tracking is an inverse problem. We can calibrate a measurement model  $\tilde{\mathbf{c}}(\mathbf{P}, \rho)$  to predict the coupling  $\mathbf{C}$  we expect to see in sensor pose  $\mathbf{P}$ , but this does not tell us how to solve the inverse problem: given a measurement  $\mathbf{C}$ , what is the best estimate of the pose  $\tilde{\mathbf{P}}$ ? Some measurement models permit a closed-form pose solution using linear algebra [5], [6], [7], but in general, it is again necessary to use nonlinear optimization, this time finding  $\tilde{\mathbf{P}}$ , which minimizes  $|\tilde{\mathbf{c}}(\tilde{\mathbf{P}}, \rho) - \mathbf{C}|_2$ .

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 (Section III-C). As with the calibration of the measurement magnetic model  $\tilde{\mathbf{c}}(\mathbf{P}, \rho)$ , it is important that output correction calibration data cover the intended workspace, but because of its ad hoc nonphysical 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  $4 \times 4$  transform matrix, it is fairly low order, so will not overfit the calibration data, and has a stable extrapolation beyond the edge of the calibrated volume, degrading gradually. A general transform matrix  $\mathbf{F}$  differs from a linear homogeneous matrix in that all 16 elements are unconstrained. It is neither orthogonal nor normalized, so is not a rigid body transformation, and can compensate scaling, shearing, and trapezoid distortions. An arbitrary point  $\mathbf{p}_e$  is corrected to  $\mathbf{p}_c$

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

$$\mathbf{p}_c = \mathbf{p}_f / k_n.$$

During calibration, we must determine  $\mathbf{F}$ , which minimizes the mismatch between every calibration point  $\mathbf{p}_c$  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 shows a pose diagram of the error kinematics. The pose error of a calibrated EMT is the difference  $\mathbf{E}(\tilde{\mathbf{P}})$  between the EMT output  $\tilde{\mathbf{P}}$  and the true pose  $\mathbf{P}$ . Disregard the stage error and suppose that  $\mathbf{P}$  is known exactly. Then,

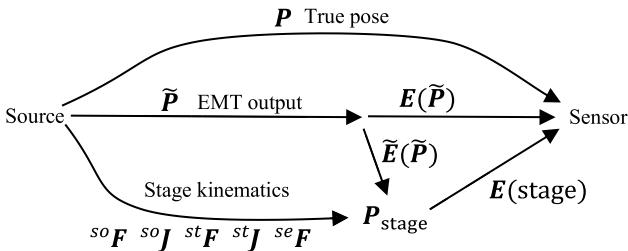
$$\begin{aligned} \mathbf{P} &= \tilde{\mathbf{P}} \mathbf{E}(\tilde{\mathbf{P}}) \\ \mathbf{E}(\tilde{\mathbf{P}}) &= \tilde{\mathbf{P}}^{-1} \mathbf{P}. \end{aligned} \quad (8)$$

$\mathbf{E}(\tilde{\mathbf{P}})$  is itself a 6-DOF transform. If the error is computed as a transform matrix by (8) and converted to (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 6-DOF error to translation and rotation magnitudes by taking the vector magnitude of the 3-DOF pose parts. This rotation magnitude avoids all difficulties with angle wrapping but does not tell us the error direction.

**TABLE II**  
ERROR VERSUS 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

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



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

#### A. Uncertainty of Stage Position and Tracker Measurement

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

$$\tilde{E}(\tilde{P}) = \tilde{P}^{-1} {}^{\text{so}}\mathbf{F} {}^{\text{so}}\mathbf{J} {}^{\text{st}}\mathbf{F} {}^{\text{st}}\mathbf{J} {}^{\text{se}}\mathbf{F}. \quad (9)$$

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

Let  $u(x_i)$  be the standard uncertainty of a measurement  $x_i$  and  $u_c$  be 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-square (RSS) rule is justifiable supposing that  $u(x_i)$  is the standard deviation 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 2–4) to define an expanded uncertainty that encloses the measurement error with higher confidence.

We wish to find the standard uncertainty of  $\tilde{P}$ :  $u(\tilde{P})$ . Given the true error  $E(\tilde{P})$ , we could estimate  $u(\tilde{P})$  as the standard deviation (rms value) of this error across the test points. Let  $u(\text{cal})$  be the rms standard uncertainty estimate from the

available  $\tilde{E}(\tilde{P})$ , (9).  $u(\text{cal})$  alone underestimates  $u(\tilde{P})$ , but by including the stage standard uncertainty, the combined uncertainty accounts for this

$$u(\tilde{P}) = \sqrt{u(P_{\text{stage}})^2 + u(\text{cal})^2}. \quad (10)$$

#### B. Systematic and Random Error

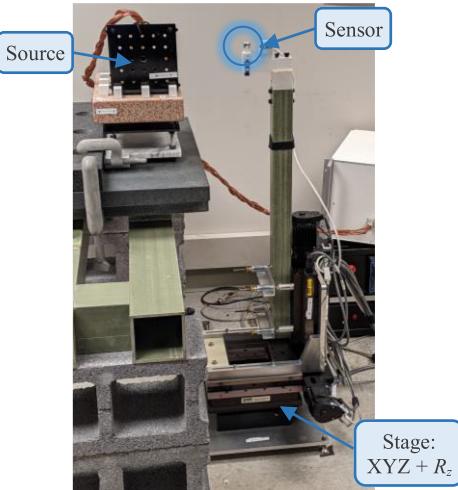
An error can be divided into systematic (repeatable) error and random error. The entire calibration procedure is an attempt to minimize the systematic error in the tracker measurement. The stage has its own systematic and random positioning errors. Since the same stage is used for both calibration and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

When choosing the model  $\tilde{c}(P, \rho)$ , it is important to consider that if  $\rho$  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.

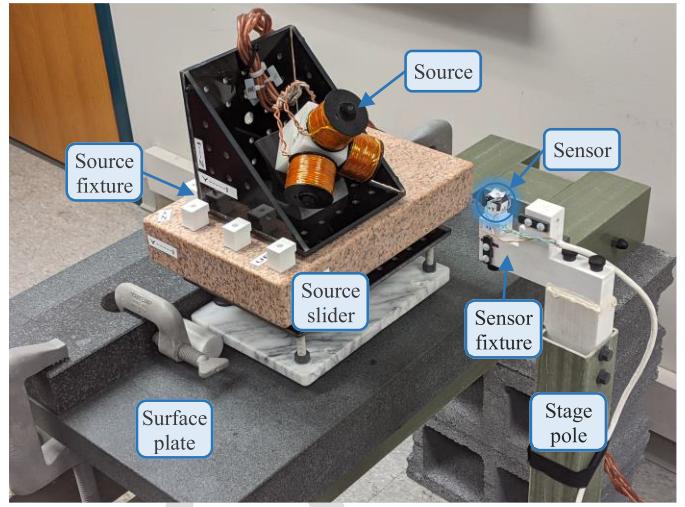
If the calibration data has far more DOF than  $\rho$  does, then random positioning error will be averaged out in the calibration, while random error during testing will increase  $u(\text{cal})$ . This is the reverse of the situation with systematic error, since true error  $E(\tilde{P})$  does not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

#### C. Stage and Fixture Design

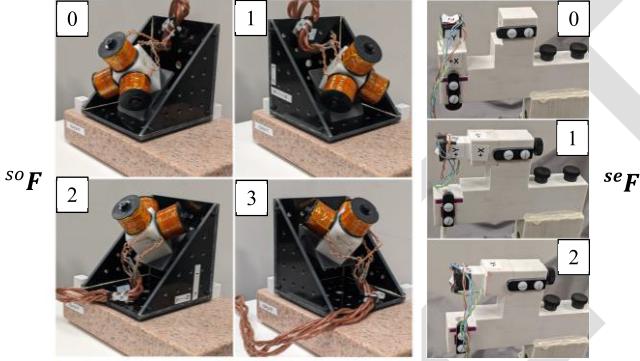
The calibration stage in Fig. 4 provides  $\pm 50$ -mm translation in  $x$ ,  $y$ , and  $z$ , and  $\pm 100^\circ$  rotation ( $R_z$ ) and is constructed to minimize magnetic interference with the measurement. The stage position uncertainty is  $107 \mu\text{m}$ , which has little effect on the combined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



**Fig. 4.** Stage provides automated motion over a 100-mm cube translation workspace, with  $\pm 90^\circ$  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.



**Fig. 6.** Source and sensor fixtures permit precise  $90^\circ$  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.



**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  $\pm 90^\circ$ . Source motion is only via the source fixture  $90^\circ$  rotation. Ideally, each of the six faces of the source would be fixtured toward the sensor, but this is not mechanically possible with this fixture.

#### D. Data Collection Patterns

The calibration and test data can 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  $3 \times 3 \times 3$  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  $\pm 90^\circ$  are also taken at each grid point, giving 135 points in each fixture configuration. The larger test data have a  $5 \times 5 \times 5$  grid with the same rotations (625 points).

This pattern is collected in each of the five rotations of the source fixture and the three rotations of the sensor fixture,

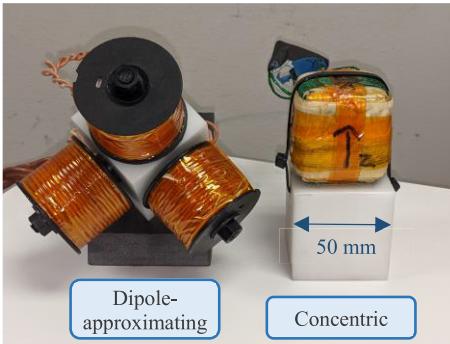
see Fig. 5 and Section V-D. In the dipole model (4), the source and sensor response interact linearly. For small deviations from the dipole model, the deviation can be decomposed into independent source and sensor effects, so it is not necessary to evaluate the full cross product of source and sensor rotations. Broadly, the rotation of the sensor helps to identify the sensor response, while rotations of the source identify the source field. For  $\langle{}^{\text{so}}\mathbf{F}, {}^{\text{se}}\mathbf{F}\rangle$  identified in Fig. 5, we test at poses  $\langle{}^{\text{so}}\mathbf{F} \in \{0.3\}, {}^{\text{se}}\mathbf{F} = 0\rangle$  and  $\langle{}^{\text{so}}\mathbf{F} = 0, {}^{\text{se}}\mathbf{F} \in \{1, 2\}\rangle$ , giving six manual fixture positions with the 625 stage poses repeated at each, for a full test dataset of 3750 points. Therefore, even with this coarse grid spacing, it is impractical to densely sample across a 6-DOF workspace.

#### E. Tested Hardware

Source drive and sensor readout are provided by the ILEMNT 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 ILEMNT design and performance is outside the scope of this article, typical performance with the tested sources and sensor combinations is  $<3\text{-}\mu\text{m}$  rms position noise at 1500 samples/s (500-Hz bandwidth) at  $r = 200$  mm. This is a frequency-domain multiplexed ac design, where each source coil continuously emits a sinusoidal field at a distinctive frequency (7.5, 10.5, 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. Nonrepeatable variation in the coupling is dominated by drift. The collection of calibration and test data takes approximately 8 h. Drift over this time degrades calibration accuracy, so we powered up the tracker at least 2 h prior to data collection (supplement Section IV).

The only sensor used is the Premo Magnetics 3DCC10-A-0600 J. While this is a fairly large  $17 \times 15 \times 14$  mm sensor (which improves the signal-to-noise ratio),



**Fig. 7.** 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.

578 it is still much smaller than the source, so the sensor response  
 579 is much closer to a dipole than the source field is. We have  
 580 characterized performance with a variety of sensors, and the  
 581 sensor has always had less effect on the accuracy than the  
 582 source did. Using only one sensor simplifies our presentation.

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

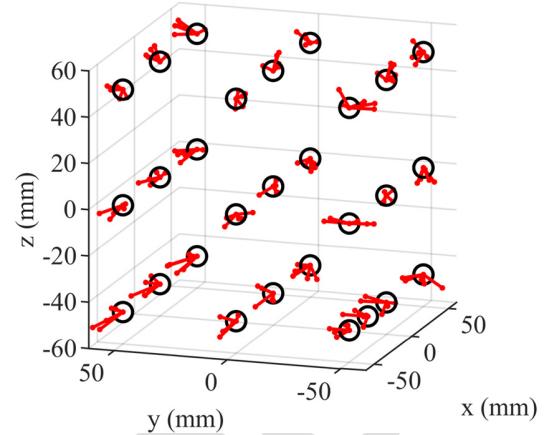
593 For both sources, the coil diameter is  $d \approx 40$  mm, which  
 594 with a testing radius of  $150 \text{ mm} < r < 250 \text{ mm}$  gives  
 595  $3.75 < r/d < 6.25$ . Operating at this relatively short range  
 596 reduces measurement noise but also increases the difficulty  
 597 of accurate calibration because the deviation from a dipole  
 598 model can be significant [27].

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

#### 607 *F. Calibration*

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

615 Table I shows a sample calibration for the dipole approximating  
 616 source. These are the  $L$  and  $M$  matrices from (5), with  
 617 coils as columns. Note the values forced to 0.0 and 1.0 to avoid  
 618 excess degrees of freedom (Section VI). This source is highly  
 619 nonconcentric, with a 45-mm offset. The sensor is designed



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

620 to be concentric, but calibration finds position offsets similar  
 621 in magnitude to the position accuracy we obtain next. Source  
 622 and sensor nonorthogonality are small but still significant.  
 623 The source  $X Z$  deviation is  $1.6^\circ$ , which would correspond  
 624 to a 5.5-mm error at 200 mm. Both the default and concentric  
 625 calibration methods model axis nonorthogonality.

#### 626 *G. Position and Rotation Error*

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

632 Table II shows the rotation and translation error of the two  
 633 calibrated sources as the calibration type is varied. The errors  
 634 reported are rms and maximum vector magnitudes of the  
 635 translation error ( $XYZ$ ) and the rotation vector of the rotation  
 636 error ( $R_x R_y R_z$ ). Here, both the calibration and test data were  
 637 taken with three sensor fixtures, but no source fixture rotation.  
 638 With the general calibration, the dipole approximating source  
 639 gives significantly better accuracy than the concentric source  
 640 (empirical support for [26]). Since this source is not con-  
 641 centric, the accuracy is terrible using a concentric calibration  
 642 model. For the concentric source, the concentric model has  
 643  $2\times$  the error of the general dipole model, which for some  
 644 uses may be an acceptable sacrifice in order to enable a  
 645 more efficient pose solution. For both sources (Section VII),  
 646 output correction significantly improves the  $XYZ$  accuracy.  
 647 The combined measurement uncertainty from (10), which  
 648 allows for stage error, is only slightly greater than the rms  
 649 calibration error  $u(\text{cal})$  alone, so system accuracy is not being  
 650 significantly degraded by stage error.

#### 651 *H. Effect of Calibration DOF on Accuracy*

652 What happens when the calibration data do not cover  
 653 the operation workspace? While we understand that such  
 654 mismatch is not good, it may not be appreciated how  
 655 severe the accuracy degradation may be. Table III shows  
 656 the effect of adding dimensions of motion, present in either

**TABLE III**  
CALIBRATION ERROR VERSUS SENSOR AND SOURCE ROTATION (CONCENTRIC SOURCE, UNCORRECTED)

Calibrate data:	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)	
	RMS	Max	RMS	Max	RMS	Max	RMS	Max	RMS	Max	RMS	Max
Translation only	<b>0.231</b>	<b>0.613</b>	<b>0.109</b>	<b>0.308</b>	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	<b>0.416*†</b>	<b>1.022</b>	<b>0.273</b>	<b>0.637</b>	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	<b>2.306†</b>	<b>6.121</b>	<b>1.100</b>	<b>3.395</b>

**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.

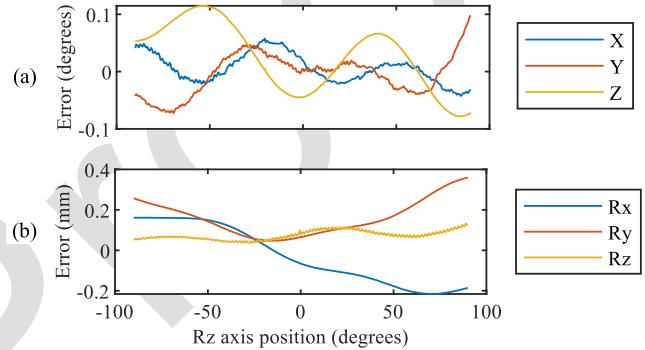
the calibration data, the test data, or both. For translation only, there is no sensor rotation at all (not even  $R_z$ ), then we add all-axis sensor rotation and finally all-axis source rotation. These results are with the concentric source and default calibration. With the dipole approximating source, 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 6-DOF, 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  $4.5\times$ . 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 is often unnecessary in surgical applications.) Comparing the “†” entries, adding source fixture rotation has a large  $5.5\times$  effect, even with matched calibration. (This may be due to irregularities in the hand-wound source.)

### I. 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  $100\times$  smaller. It would require an entirely impractical number test points to densely cover the 6-DOF 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 of 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 does not 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 six measured DOFs. 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 the 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



**Fig. 9.** Actual motion on a single axis (here  $R_z$ ) creates measurement errors on all axes. (a) Cross error (here mm translation error caused by rotation), while (b) 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 these data as rotation and translation error vector magnitudes.

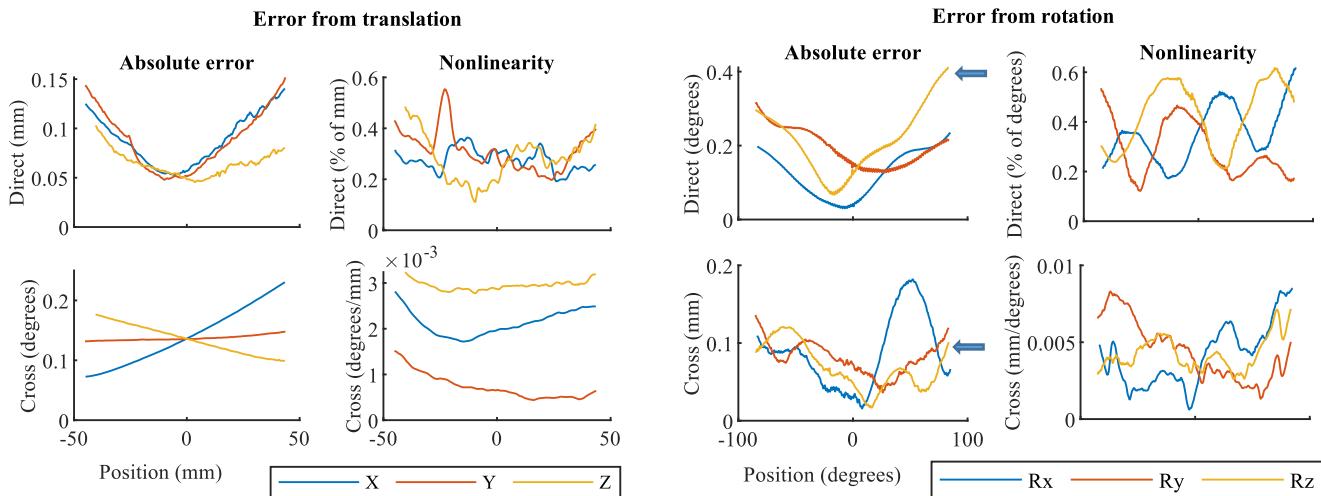
**TABLE IV**  
CROSS-COUPLING AND NONLINEARITY DURING AXIS SWEEPS

	Absolute error		Nonlinearity	
	XYZ	RxRyRz	XYZ	RxRyRz
	mm	degrees	Percent	
Direct	RMS	0.093	0.232	0.32%
	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/°

(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. 10 shows the absolute error and nonlinearity resulting from sensor motion sweeps on all six axes. Measurements are at 0.5 mm, 0.5° increments over 100 mm, 180°. 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 curve in Fig. 10 is summarized by rms and max, and then, these measures are combined by max across



**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 3-DOF rotation or translation error caused by motion on the axis in the legend. Nonlinearity is the magnitude of the 3-DOF 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 have been reduced to two curves (see arrows): the absolute error from rotation ( $R_z$ ), which is the vector magnitudes of Fig. 9(b) and (a).

722 the curves (driving axes) in each plot. An absolute error is  
 723 computed as in Table II but is now significantly lower. This  
 724 is mainly because the sweep data only vary one axis, so is  
 725 on average closer to the center of the workspace, where the  
 726 error is lower. Note how the magnitude of the rotation-to-  
 727 translation cross error is similar to the direct translation error.  
 728 This reinforces the point that we cannot characterize EMT  
 729 translation accuracy by using a translation-only test pattern.

## IX. DISCUSSION

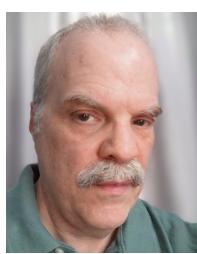
730 As the calibration component of the ILEMTR open EMT  
 731 design, we only aimed to match the static accuracy of com-  
 732 mercial EMTs. While we have not compared commercial  
 733 trackers under the same conditions, published results such  
 734 as [19] suggest that we have met and even surpassed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

738 It is of course unknown whether these methods differ  
 739 from commercial EMT practice, but the broad principle of  
 740 fitting a magnetic model using nonlinear optimization is every-  
 741 where in the open literature. Aside from the reusable and  
 742 freely available implementation, our most important additions  
 743 to EMT practice are the introduction of a general matrix  
 744 formulation based on linear transforms and linear homoge-  
 745 neous coordinates. We also closely examined characterization  
 746 methods, including practical details of stage construction and  
 747 analysis of the effects of stage error on the accuracy of the  
 748 calibration/test procedure. Characterization of tracker linearity  
 749 and the rotation/translation cross-coupling are particularly  
 750 important for understanding the performance achievable dur-  
 751 ing 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](https://doi.org/10.1109/TAES.1979.308860).
- [2] R. A. MacLachlan et al., “Electromagnetic tracker for active handheld robotic systems,” in *Proc. IEEE SENSORS*, Oct. 2016, pp. 1–3, doi: [10.1109/ICSENS.2016.7808415](https://doi.org/10.1109/ICSENS.2016.7808415).
- [3] Robert A. MacLachlan. (2024). *In-Loop ElectroMagnetic Tracker*. Open Science Framework. Accessed: Feb. 2, 2024. [Online]. Available: <https://osf.io/6fwjk/>
- [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, Jun. 2017, doi: [10.1007/s11548-017-1568-7](https://doi.org/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, May 2018, doi: [10.1109/TIE.2017.2760244](https://doi.org/10.1109/TIE.2017.2760244).
- [6] C. Hu, S. Song, X. Wang, M. Q.-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](https://doi.org/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,” Univ. College Cork, Cork, Ireland, Tech. Rep., 2014.
- [9] T. Bien and G. Rose, “Algorithm for calibration of the electromagnetic tracking system,” in *Proc. IEEE-EMBS Int. Conf. Biomed. Health Informat.*, Jan. 2012, pp. 85–88, doi: [10.1109/BHI.2012.6211512](https://doi.org/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](https://doi.org/10.1109/TMAG.2008.2003056).
- [11] X. Ge, Y. Wang, N. Ding, X. Wu, Y. Wang, and Z. 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](https://doi.org/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*, vol. 6505, 2007, Art. no. 650901, doi: [10.1117/12.713249](https://doi.org/10.1117/12.713249).
- [13] J. Hummel, M. Figl, M. Bax, R. Shahidi, H. Bergmann, and W. Birkfellner, “Evaluation of dynamic electromagnetic tracking deviation,” in *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, Feb. 2009, p. 72612U, doi: [10.1117/12.813645](https://doi.org/10.1117/12.813645).

- 804 [14] D. Kügler et al., "High-precision evaluation of electromagnetic tracking," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 7, pp. 1127–1135, Jul. 2019, doi: [10.1007/s11548-019-01959-5](https://doi.org/10.1007/s11548-019-01959-5).
- 805 [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](https://doi.org/10.1088/0031-9155/48/14/314).
- 806 [16] C. Nafis, V. Jensen, L. Beauregard, and P. Anderson, "Method for estimating dynamic EM tracking accuracy of surgical navigation tools," *Proc. SPIE*, vol. 6141, Mar. 2006, Art. no. 61410K, doi: [10.1117/12.653448](https://doi.org/10.1117/12.653448).
- 807 [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. Imag.*, vol. 33, no. 8, pp. 1702–1725, Aug. 2014, doi: [10.1109/TMI.2014.2321777](https://doi.org/10.1109/TMI.2014.2321777).
- 808 [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](https://doi.org/10.1118/1.1944327).
- 809 [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, Mar. 2015, doi: [10.1007/s11548-014-1110-0](https://doi.org/10.1007/s11548-014-1110-0).
- 810 [20] V. V. Kindratenko, "A survey of electromagnetic position tracker calibration techniques," *Virtual Reality*, vol. 5, no. 3, pp. 169–182, Sep. 2000, doi: [10.1007/BF01409422](https://doi.org/10.1007/BF01409422).
- 811 [21] W. Birkfellner, F. Watzinger, F. Wanschitz, R. Ewers, and H. Bergmann, "Calibration of tracking systems in a surgical environment," *IEEE Trans. Med. Imag.*, vol. 17, no. 5, pp. 737–742, Oct. 1998, doi: [10.1109/42.736028](https://doi.org/10.1109/42.736028).
- 812 [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 Sensors J.*, vol. 21, no. 13, pp. 14389–14396, Jul. 2021, doi: [10.1109/JSEN.2020.3042647](https://doi.org/10.1109/JSEN.2020.3042647).
- 813 [23] P. Corke, *Robotics, Vision and Control*, vol. 118, 2nd ed. Cham, Switzerland: Springer, 2017.
- 814 [24] I. E. Sutherland, "Three-dimensional data input by tablet," *Proc. IEEE*, vol. 62, no. 4, pp. 453–461, Apr. 1974.
- 815 [25] G. H. White, "Basics of estimating measurement uncertainty," *Clin. Biochemist Rev.*, vol. 29, pp. S53–S60, Aug. 2008. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/18852859>
- 816 [26] E. Paperno and A. Plotkin, "Cylindrical induction coil to accurately imitate the ideal magnetic dipole," *Sens. Actuators A, Phys.*, vol. 112, nos. 2–3, pp. 248–252, May 2004, doi: [10.1016/j.sna.2004.01.004](https://doi.org/10.1016/j.sna.2004.01.004).
- 817 [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](https://doi.org/10.1063/1.4922667).



**Robert A. MacLachlan** (Member, IEEE) received the B.S. degree in applied mathematics from Carnegie Mellon University, Pittsburgh, PA, USA, in 1987.

He has been a Research Software Developer and an Engineer with Carnegie Mellon University since 1983. At the Robotics Institute, he has developed high-precision measurement systems and intelligent medical instruments, with a particular focus on electromagnetic and optical tracking systems for microsurgical applications. His research interests include medical robotics, precision instrumentation, and real-time control systems, with an emphasis on combining sensitive measurements with signal processing to enable new mechatronic solutions for surgical challenges.



**Claudia Pelle** received the master's degree in biomedical engineering from the Politecnico di Milano, Milan, Italy, with a focus on technologies for electronics. Her master's thesis, titled "Electromagnetic Tracker Implementation for Localization of Precision Manipulators," was developed at The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA.

She is a Software Embedded Engineer with Thales Italia S.p.A, Rome, Italy. Her professional work primarily involves developing and software for Thales' products, contributing to innovative solutions in areas such as aerospace, defense, and security technologies.



**Ralph L. Hollis** (Fellow, IEEE) received the Ph.D. degree in physics from the University of Colorado, Boulder, CO, USA.

He was a Research Staff Member with the Thomas J. Watson Research Center from 1978 to 1993, where worked in magnetism, acoustics, and robotics, and was the Manager of Advanced Robotics from 1986 and 1993. He is currently a Research Professor of robotics and electrical and computer engineering at Carnegie Mellon University, Pittsburgh, PA, USA. He is a Founding Director of the Microdynamic Systems Laboratory, Carnegie Mellon University, where his research centers on haptics, agile manufacturing, and dynamically stable mobile robots.

Dr. Hollis is a member of American Physical Society.



**Elena De Momi** (Senior Member, IEEE) received the M.Sc. degree in biomedical engineering and the Ph.D. degree in bioengineering from the Politecnico di Milano, Milan, Italy, in 2002 and 2006, respectively.

She was the Co-Founder of the Neuroengineering and Medical Robotics Laboratory, in 2008, being responsible for the Medical Robotics Section. She is responsible for the lab course in medical robotics, the courses on Clinical Technology Assessment and Smart Hospital of the M.Sc degree in biomedical engineering, Politecnico di Milano, and the course in healthcare robotics and active aging at MEDTEC (Humanitas University), Rozzano, Italy. She serves on the Board Committee of the Ph.D. course in bioengineering and the National Ph.D. in robotics and intelligent machines. She is currently a Full Professor with the Electronic Information and Bioengineering Department (DEIB), Politecnico di Milano.

Dr. De Momi has been the Publication Chair and the Co-Chair of IEEE ICRA 2019, 2023, and 2024 and is the Program Chair of IROS 2025. She is currently a Senior Editor of *International Journal of Robotics Research*, an Editor of *IEEE Robotics and Automation Magazine*, and an Associate Editor of *IEEE TRANSACTIONS ON ROBOTICS AND ROBOTICS AND AUTOMATION LETTERS* and *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*.



**Cameron N. Riviere** (Senior Member, IEEE) received the Ph.D. degree in mechanical engineering from Johns Hopkins University, Baltimore, MD, USA, in 1995.

In 1995, he joined the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA, where he is currently a Research Professor. His research interests include medical robotics, control systems, signal processing, and human-machine interfaces.

Dr. Riviere has been an Associate Editor on the Conference Editorial Boards of the IEEE Robotics and Automation Society and the IEEE Engineering Medicine and Biology Society, and a Guest Editor of special issues of *IEEE TRANSACTIONS ON MEDICAL ROBOTICS AND BIONICS* and *PROCEEDINGS OF THE IEEE*.