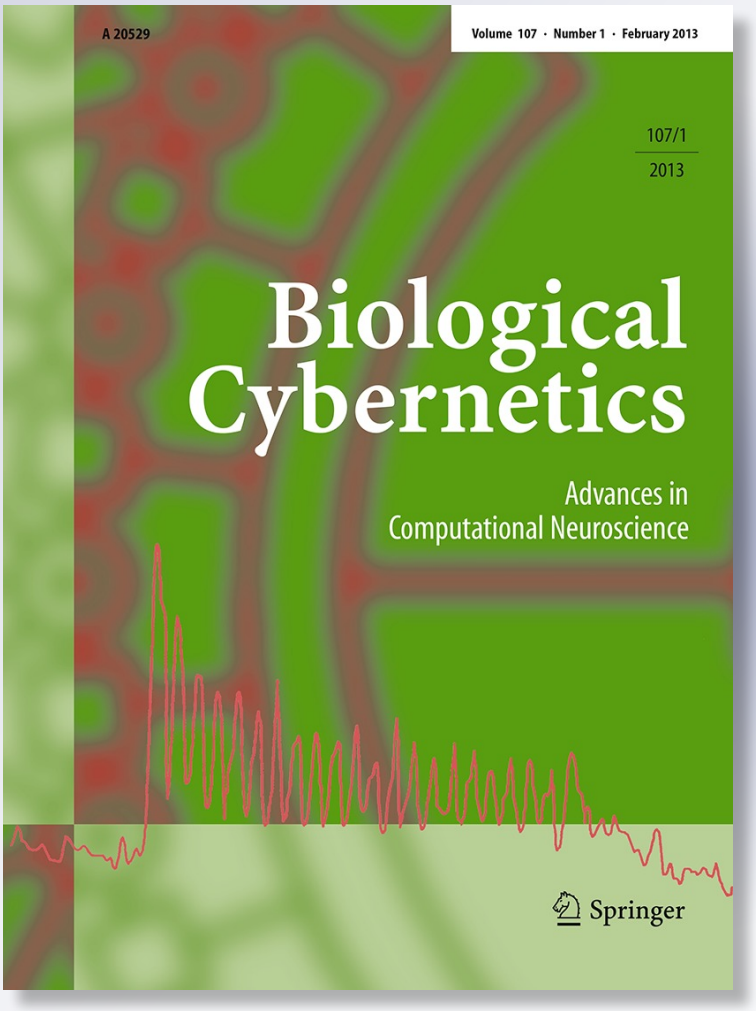
The modular modality frame model: continuous body state estimation and plausibility-weighted information fusion

Stephan Ehrenfeld & Martin V. Butz

Biological Cybernetics
Advances in Computational Neuroscience

ISSN 0340-1200
Volume 107
Number 1

Biol Cybern (2013) 107:61-82
DOI 10.1007/s00422-012-0526-2



Springer

Biol Cybern (2013) 107:61–82
DOI 10.1007/s00422-012-0526-2

ORIGINAL PAPER

Biological Cybernetics

The modular modality frame model: continuous body state estimation and plausibility-weighted information fusion

Stephan Ehrenfeld · Martin V. Butz

Received: 9 March 2012 / Accepted: 25 September 2012 / Published online: 23 October 2012

Abstract Humans show admirable capabilities in movement planning and execution. They can perform complex tasks in various contexts, using the available sensory information very effectively. Body models and continuous body state estimations appear necessary to realize such capabilities. We introduce the Modular Modality Frame (MMF) model, which maintains a highly distributed, modularized body model continuously updating, modularized probabilistic body state estimations over time. Modularization is realized with respect to modality frames, that is, sensory modalities in particular frames of reference and with respect to particular body parts. We evaluate MMF performance on a simulated, nine degree of freedom arm in 3D space. The results show that MMF is able to maintain accurate body state estimations despite high sensor and motor noise. Moreover, by comparing the sensory information available in different modality frames, MMF can identify faulty sensory measurements on the fly. In the near future, applications to lightweight robot control should be pursued. Moreover, MMF may be enhanced with neural encodings by introducing neural population codes and learning techniques. Finally, more dexterous goal-directed behavior should be realized by exploiting the available redundant state representations.

Keywords Modularity · Multimodality · Plausibility estimation · Body schema · State estimation · Robot arm control

S. Ehrenfeld · M. V. Butz (✉)
Department of Computer Science and Department of Psychology, Cognitive Modeling, University of Tübingen, Tübingen Germany
e-mail: martin.butz@uni-tuebingen.de

S. Ehrenfeld
e-mail: stephan.ehrenfeld@uni-tuebingen.de

1 Introduction

Although humans face many challenges when interacting with the environment, they master them with great skill. This *dexterous* control capability (Latash and Turvey 1996) can in part be accounted for solely by the body and hardwired reactive control architectures. Examples are embodied robots presented by Braitenberg, Brooks, or Schmitz, which can produce astonishingly versatile behavior (Braitenberg 1986; Brooks 1990; Schmitz et al. 2008). However, these embodied systems are limited to stereotypic behavior. They cannot generate the full dexterity humans demonstrate.

Instead, a *body model* is necessary for the realization of highly flexible behavior. The notion that humans develop various internal body models is commonly accepted. Various researchers have suggested that these body models are based on forward and inverse representations of sensorimotor correlations and also include estimations of the own body metric (Bernier et al. 2007; Denève et al. 2007; Hoffmann et al. 2010; Longo and Haggard 2010; Streri et al. 1993; Wolpert and Kawato 1998). Humans appear to selectively exploit these representations to predict, plan, and control movements. Thus, suitably modularized bodily representations are crucial for the dexterity apparent in human motor planning and control (Bernier et al. 2007; Latash and Turvey 1996; Latash et al. 2007; Rosenbaum 2010; Wolpert and Kawato 1998).

Nikolai A. Bernstein pointed out that the dexterous control capability of humans must be due to a highly flexible and modular neural representation and control system, which effectively deals with many redundant degrees of freedom of the human body (Latash and Turvey 1996). A key to human dexterity thus lies in the architecture of neurally encoded body models. In the following, we further elaborate on this conjecture. In particular, we propose that a suitable body model

should represent its internal state estimations probabilistically, it should be capable of flexibly integrating multisensory information, and it should maintain redundant state estimations in multiple body-part and sensory-motor respective modules.

1.1 Probabilistic state estimations

Uncertainty arises not only during motor execution and due to sensor noise, but also due to uncertainty about the internal body model, the body state, the task, and the state of the environment. Examples are unpredictable changes in the environment caused by other agents and physical influences. Also the body schema, bodily kinematics, and dynamics change during growth (Wells et al. 2002), injury, bodily exercise, or food intake (Shadmehr and Wise 2005). All these sources of uncertainty give rise to a complicated challenge, as the representations of body, goals, and constraints depend—either directly or indirectly—on the sensory system, which cannot generate noiseless, absolutely precise information. Humans address this uncertainty by means of probabilistic body state estimations (Körding and Wolpert 2004; Doya et al. 2007).

1.2 Multisensory integration

Humans gather state information with a wide variety of sensors in different modalities (like vision and proprioception) and integrate it in their perception of their own body (e.g., Beauchamp 2005; Makin et al. 2008; Maravita et al. 2003). Thus, redundant sources of information *interact* in the human brain in various ways (Calvert et al. 2004; Tononi et al. 1998). These interactions can lead to peculiar effects, such as the rubber hand illusion (Botvinick et al. 1998; Makin et al. 2008), where a representation in one modality frame is strongly modified by information stemming from other modalities and frames of reference. Sensory information from different modalities is thus flexibly combined selectively and probabilistically, dependent on current accuracy and plausibility estimates.

1.3 Modularized representations

Seeing that different sources of information are available in different brain areas, body and environment are selectively and redundantly represented in many state estimation modules. In addition, the brain modularizes its state estimations body-part respectively. For example, bodily self-perception is clearly separated into body segments (de Vignemont et al. 2009). Also, not only primary motor cortex and somatosensory cortex show bodily homunculi, but also many parietal cortical areas and premotor areas exhibit body part distinctions and selective, multimodal, sensory, and reafferent information source integration (Andersen et al. 1997; Shadmehr and Krakauer 2008). With respect to behavior, different parts of the body and bodily components can be moved by selectively activating motor synergies (Gentner and Classen 2006; Latash et al. 2007). With respect to particular movement tasks, a modularized redundant representation allows immediate selective access to the most relevant information amongst the redundant alternatives. Thus, highly modularized, redundant state estimations of body and environment are relevant both, for a robust and flexible representation of state estimations and for dexterous behavioral decision making and control.

Inspired by these characteristics, we propose the Modular Modality Frame (MMF) model. MMF is a highly modularized architecture, which maintains a probabilistic body state estimation over time in redundant modules. Forward and inverse kinematic mappings link these modules and allow for a flexible flow of information, resulting in sensor fusion and local interaction of multiple state estimations. The benefits of our model include high noise robustness, low dimensional representations, and the ability to estimate sensory reliability on the fly.

In Sect. 2, we describe the general architecture of the proposed model: its modality frames and the cooperation of different steps in maintaining the body state representation. These steps are described in more detail in Sect. 3. They include movement prediction, sensor fusion using plausibilities, sensor integration, and crosstalk between modules. The calculation of sensor plausibilities is defined in Sect. 4. The necessary local forward and inverse kinematic mappings are detailed in Sect. 5. In Sect. 6 the experimental setup is described, and results are presented and discussed. The work is concluded with a short summary and a discussion of interesting potential system extensions and applications.

2 MMF model overview

MMF is inspired by the knowledge about the architecture of the human brain and human behavioral control capabilities, as sketched-out above. In particular, MMF represents an internal body model with modularized, locally interactive representations. Each module maintains a particular, probabilistic body state estimation over time by integrating the available multisensory information by means of Bayesian principles.

The body model comprises a full forward kinematics model (used for prediction) and a full inverse kinematics model (used for planning and control), providing high accuracy in simple goal-directed movements. For extensive overviews over other body models in robotics cf. Hoffmann et al. (2010), Nguyen-Tuong and Peters (2011), Sigaud et al. (2011). MMF's modules differ from each other with respect to (I) the arm limb, (II) the sensory modality, which can be a perceived location, direction, or joint angle, and (III) the frame of reference, which can be "global" (i.e., head-

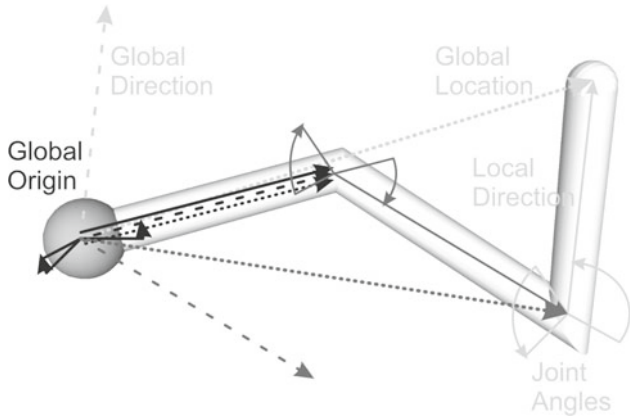


Fig. 1 Schematic of the modules with different modality frames for Limb 1 (*black*), Limb 2 (*dark gray*), and Limb 3 (*light gray*). Joint locations (*dotted arrows*) and limb orientations (*dashed arrows* at origin) are represented shoulder-centered. Moreover, limb orientations are also represented relative to the *upper limb* (*solid arrows* on arm). Finally, joint angles are encoded (*curved arrows*)

centered) or "local" (i.e., respective the previous limb). We term the combination of (II) and (III) a *modality frame*. In particular, the utilized modality frames are termed *global location*[1] (GL), *global orientation*[2] (GO), *local orientation* (LO), and *local (joint) angles* (LA). Other modalities could be used, such as a local location, or also other combinations of modalities and frames of reference. The implemented modules as well as their interactions are illustrated in Figs. 1 and 3. Together, the modules form a redundant estimation of the overall body state.

In each of MMF's modules, the arm state is approximated. In addition, a sensor may exist in each module. Both the approximate arm state and the sensory measurement (if applicable) are assumed to be Gaussian and are each represented with a mean-vector and a Covariance-matrix. The approximate arm state is maintained over time in three steps per iteration (A–C, cf. Fig. 2). We refer to this maintenance over time as *filtering*. For a comparison to similar approaches cf. Sect. 7.1.

The prediction step (A) derives the a priori arm state from the motor command and the forward model. The update-step (B) derives the a posteriori arm state by integrating the combined information from all sensors. This combined information is built by fusing all accessible sensors, weighted by the sensors' covariance-matrices and their *plausibilities* (22). A plausibility measures the trust the system puts in a specific sensor (cf. Sect. 4). MMF assigns these plausibilities autonomously to each sensor, by comparing its readings with other redundant sensory measurements. The third step (C) introduces *crosstalk* between all filters, ensuring that a consistent

[1] The location is a vector from the global origin to the end-point of a limb.

[2] The orientation consists of one unit vector parallel (cf. Fig. 2, "Direction") to a limb and a second unit vector perpendicular to it.

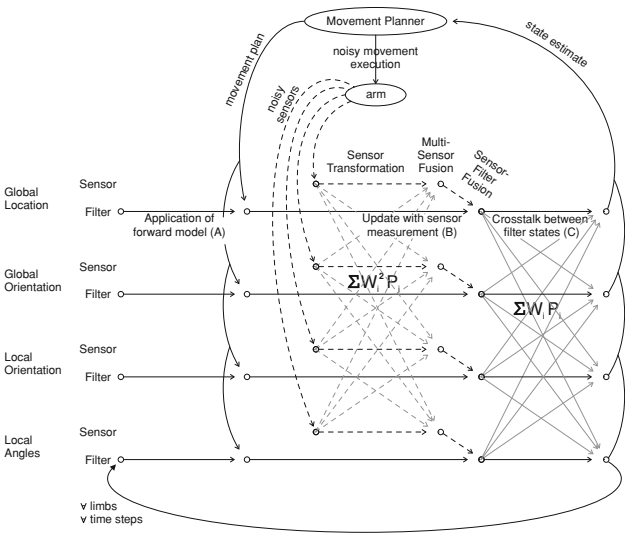


Fig. 2 Data flow: First, the forward model predicts the state estimate after the movement (*A*). Second, the measurements are transformed from all modality frames to all other frames. Comparing the redundant measurements, their respective plausibilities are calculated (cf. Sect. 4) and they are fused, weighted with their plausibility and variance (*dashed lines*). These fused measurements are then integrated in their respective modality frame (*B*). Third, the crosstalk shifts all filter states towards all other filter states, synchronizing them (*C*)

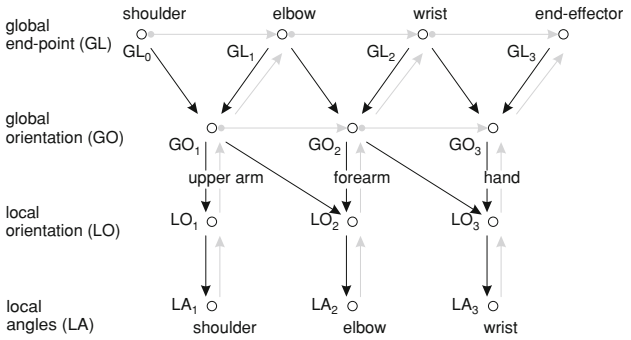


Fig. 3 Dependencies for transforming between different spaces. The *gray arrows* display the transformation from angles to location (forward kinematics), the *black* ones vice versa (inverse kinematics). *Vertical axis*: frames of reference, *horizontal axis*: *limbs, circles*: modules

body state estimation is maintained over time. In this way, the individual state estimates synchronize, resulting in a body state estimate that is highly accurate and consistent across all modality frames. All three steps (A)–(C) are applied successively in the kinematic arm chain starting with the most upper arm limb.

In summary, the redundant and interactive nature of MMF, together with rigorous Bayesian information processing principles allows to fully exploit the available sensory information. MMF explicitly provides state estimations as well as accuracy estimations. Goal-directed motion control is possible despite high sensor and motor noise. Finally, even sensor failures are detectable on the fly, enabling MMF to handle sensory information with randomly changing information

contents. In the following, the three steps and the module interactions within each step are detailed.

3 Body state maintenance

Before we define the prediction (A), update (B), and crosstalk (C) steps in detail, a short overview of transformations between different modality frames is provided below. All three steps A–C (cf. Fig. 2) require mappings that transform states between different modality frames. In classical robotics, the forward (instantaneous) kinematics that map angles $\mathbf{q}$ to a global position ξ is

$$\xi = \mathbf{f}(\mathbf{q}), \tag{1}$$

$$\dot{\xi} = \mathbf{J}(\mathbf{q})\,\dot{\mathbf{q}}, \tag{2}$$

where $\mathbf{f}$ describes the forward-mapping function and $\mathbf{J}$ its Jacobian. Equation (1) maps mean-values, while (2) maps the movement of this mean, and also a standard-deviation-vector (i.e., the square roots of the main diagonal entries of a covariance-matrix). If full covariance-matrices are considered, (2) becomes

$$\mathbf{Cov}(\xi) = \mathbf{J}(\mathbf{q})\,\mathbf{Cov}(\mathbf{q})\,\mathbf{J}^T(\mathbf{q}), \tag{3}$$

with $\mathbf{Cov}(\mathbf{x})$ denoting the covariance matrix of vector $\mathbf{x}$.

As we split the kinematics representation into *four* modality frames, we introduce a more general notation denoting the mean and covariance matrix of a particular modality frame by $\mathbf{x}^i$ and $\mathbf{P}^i$, respectively, where the superscript $i \in \{\text{GL, GO, LO, LA}\}$ specifies the modality frame. Thus, $\mathbf{x}^{\text{GL}} := \xi$, $\mathbf{P}^{\text{GL}} := \mathbf{Cov}(\xi)$, $\mathbf{x}^{\text{LA}} := \mathbf{q}$, and $\mathbf{P}^{\text{LA}} := \mathbf{Cov}(\mathbf{q})$. Additionally, we split the arm representation limb-respectively, denoting a modality frame for a specific limb j by an additional subscript, where necessary. Together, both mean and covariance specify the *state* $\mathbf{s}^i_j$ estimation in a particular modality frame i and for a particular limb j. For example, the state of the global orientation of the lower arm limb is referred to by $\mathbf{s}^{\text{GO}}_2$. These modularized modality frames and the local mappings between the modules are displayed in Fig. 3.

The mappings show that any local mapping $\mathbf{f}^{j\to i}$ from a source modality frame j to a target modality frame i depends on the state $\mathbf{s}^j$ (of the same limb as the target state) and possibly also on the state of the next upper arm limb $\mathbf{s}^k_u$, where the subscript u is used to refer to the corresponding next upper arm limb. Thus, any local mapping can be denoted by

$$\mathbf{x}^i = \mathbf{f}^{j\to i}_{\mathbf{x}}\left(\mathbf{x}^j, \mathbf{x}^k_u\right) \tag{4}$$

$$\mathbf{P}^i = \mathbf{f}^{j\to i}_{\mathbf{P}}\left(\mathbf{P}^j, \mathbf{P}^k_u\right) = \left(\frac{\partial \mathbf{x}^i}{\partial \mathbf{x}^j}\right)\mathbf{P}^j\left(\frac{\partial \mathbf{x}^i}{\partial \mathbf{x}^j}\right)^{\mathrm{T}} + \left(\frac{\partial \mathbf{x}^i}{\partial \mathbf{x}^k_u}\right)\mathbf{P}^k_u\left(\frac{\partial \mathbf{x}^i}{\partial \mathbf{x}^k_u}\right)^{\mathrm{T}}, \tag{5}$$

where $i, j, k \in \{\text{GL, GO, LO, LA}\}$. Note that technically the actual values of the Jacobians $\left(\frac{\partial \mathbf{x}^i}{\partial \mathbf{x}^j}\right)$ and $\left(\frac{\partial \mathbf{x}^i}{\partial \mathbf{x}^k_u}\right)$ still depend on the current state estimates $\mathbf{x}^j$ and $\mathbf{x}^k_u$ of the involved modality frames. For reading convenience, this dependency is not denoted.

3.1 Prediction step

In the prediction step (cf. Fig. 2 step A), the motor command is used to predict the resulting change in the modality state estimations. As a motor command is typically available only in the angular modality frames LA, the reactions of all other frames i need to be calculated using (4–5). Seeing that (4–5) depend on both the current limb and the upper limb, the movements of *both* limbs need to be incorporated in the forward model.

The consequences of the *current* limb movement on a state $\mathbf{s}^i$ are incorporated by a transformation $\mathbf{f}^{i\to\text{LA}}$ of $\mathbf{s}^i$ to the angular frame LA, an addition of the movement in the frame LA, and a back-transformation $\mathbf{f}^{\text{LA}\to i}$ of the moved state to the initial frame i. The consequences of the *upper* limb are integrated by basing the transformation $\mathbf{f}^{i\to\text{LA}}$ on the upper limb state at the *previous time step* $\mathbf{x}_{u,t-1|t-1}$ and basing the back-transformation $\mathbf{f}^{\text{LA}\to i}$ on the upper limb state of the *current time step* $\mathbf{x}_{u,t|t}$. Together, this yields

$$\mathbf{x}^i_{t|t-1} = \mathbf{f}^{\text{LA}\to i}_{\mathbf{x}}(\mathbf{f}^{i\to\text{LA}}_{\mathbf{x}}(\mathbf{x}^i_{t-1|t-1}, \mathbf{x}_{u,t-1|t-1}) + \Delta\mathbf{x}^{\text{LA}}_t, \mathbf{x}_{u,t|t}) \tag{6}$$

$$\mathbf{P}^i_{t|t-1} = \mathbf{f}^{\text{LA}\to i}_{\mathbf{P}}(\mathbf{f}^{i\to\text{LA}}_{\mathbf{P}}(\mathbf{P}^i_{t-1|t-1}, \mathbf{P}_{u,t-1|t-1}) + \Delta\mathbf{P}^{\text{LA}}_t, \mathbf{P}_{u,t|t}). \tag{7}$$

On the left hand side, $\mathbf{x}_{t|t-1}$ specifies the state at time t, given the sensory information at times $0, \ldots, t-1$, i.e., the a priori state estimates after the movement. On the right hand side, $\mathbf{x}_{t-1|t-1}$ is the a posteriori estimate from the *last time step* (after the crosstalk step C) and $\Delta\mathbf{x}^{\text{LA}}_t$ the motor command in the angular frame. We showed in Ehrenfeld and Butz (2011) that the back-forth transformation yields more accurate state estimates than commonly used linear approximation methods via Jacobians. Note that the state estimate of the upper arm limb after Steps (A)–(C), that is, $\mathbf{x}_{u,t|t}$, is available because the steps are computed successively for each limb in the kinematic chain.

3.2 Update step

In the update step (cf. Fig. 2 step B), the sensory information is integrated into the modular state estimates of the MMF. This integration can compensate for the loss of knowledge, which occurs in the prediction step. In order to improve the state estimates, it is desirable to integrate *all* sensor informa-

tion into each individual filter. To achieve this, all sensors are first fused.

3.2.1 Multi-sensor fusion

This fusion (cf. Fig. 2, the dots where four dashed arrows end) is done with a weighted mean. The fused means and variances are

$$\mathbf{x}_{\text{new}} = \sum_i \mathbf{W}_i \, \mathbf{x}_i \tag{8}$$

$$\mathbf{P}_{\text{new}} = \sum_i \mathbf{W}_i^{\mathrm{T}} \, \mathbf{P}_i \mathbf{W}_i \, . \tag{9}$$

Because the sensor fusion should integrate information optimally (with Bayesian principles), the weights must be equal to the (normalized) information content $\mathbf{P}_i^{-1}$.[3] Thus, the weights yield

$$\mathbf{W}_i = \mathbf{P}_i^{-1} \left(\sum_j \mathbf{P}_j^{-1} \right)^{-1} . \tag{10}$$

3.2.2 Sensor-filter fusion

These fused sensor measurements are then integrated (cf. Fig. 2) into the modular body state estimations. Mathematically, this is again a fusion of two noisy sources of information: the fused measurement state $\mathbf{s}_{\text{new}}$ and the a priori state estimate from the prediction step $\mathbf{s}_{t|t-1}$ are combined yielding $\mathbf{s}_{t|t}$. Thus, to accomplish this second information fusion, the same Eqs. (8–10) are used.

3.3 Crosstalk step

In order to provide a coherent estimate over all modality frames, i.e., to ensure that they do not diverge, the filter estimates are synchronized. This is done with crosstalk (cf. Fig. 2 step C), which shifts the state estimates in the different modality frames towards each other. The crosstalk is realized similarly to the sensor fusion described in 3.2.1, i.e., the filter states are transformed to other modality frames and fused (cf. also Fig. 2). All sensor information has already been integrated in 3.2.2, however, the variance must not be further decreased. Instead, a simple *shift* of both means and variances in the direction of the other modules' states is required. As a consequence, (8) and (10) remain the same, but as with the means, the variances are also weighted with only the first power of the weights. This replaces (9) with

$$\mathbf{P}_{\text{new}} = \sum_i \mathbf{W}_i \, \mathbf{P}_i , \tag{11}$$

[3] In the rare case when a matrix is non-invertible, off-diagonal elements may be set to zero.

which models a weighted mean without any information gain (cf. Fig. 2). In sum, both sensor fusion and sensor integration use (8–10), while crosstalk uses (8, 10 and 11).

4 Plausibility of sensor measurements

If a sensor failure occurs, MMF unsupervisedly mistrusts this specific sensor. To do so, MMF computes a sensor plausibility measure in each module by comparing the redundant sensory information available. The plausibility measure is derived from probability theory, as detailed below.

We characterize the failure of a sensor $\mathbf{S}^i$ in a particular modality frame i by a systematic sensor offset error $\mathbf{dz}^i$ or an increase $\mathbf{dR}^i$ in the sensor's variance. The underlying probability density p of the measurement random variable $\mathbf{Z}^i$ at value $\mathbf{z}^i$ given that the real arm state random variable $\mathbf{Y}^i$ takes on the value $\mathbf{y}^i$, is

$$p\left(\mathbf{Z}^i{=}\mathbf{z}^i \,|\, \mathbf{Y}^i{=}\mathbf{y}^i\right) = \mathbf{N}\left(\mathbf{y}^i + \mathbf{dz}^i, \mathbf{R}^i + \mathbf{dR}^i\right)\left(\mathbf{z}^i\right), \tag{12}$$

assuming a Gaussian distribution $\mathbf{N}$ with specified mean and variance. The measurement variance $\mathbf{R}^i$ is assumed to be known to the system. All other components except for the actually observed measurement $\mathbf{z}^i$ are unknown to MMF. Thus, (12) is simply the underlying 3D-Gaussian, unknown to the system, around the shifted arm state and with a shape defined by the increased covariance-matrix.

This probability distribution takes on its maximum when $\mathbf{Z}^i = \mathbf{y}^i$, given the sensor is not currently failing (i.e., $\mathbf{dz}^i = 0$). By normalizing the probability distribution by this maximum, it describes the *plausibility* pl (i.e., the likelihood) that the measurement value $\mathbf{z}^i$ is in agreement with the real arm value $\mathbf{y}^i$. We thus define the plausibility as

$$pl\left(\mathbf{Z}^i{=}\mathbf{z}^i, \mathbf{Y}^i{=}\mathbf{y}^i\right) := \frac{p\left(\mathbf{Z}^i{=}\mathbf{z}^i \,|\, \mathbf{Y}^i{=}\mathbf{y}^i\right)}{p\left(\mathbf{Z}^i{=}\mathbf{y}^i \,|\, \mathbf{Y}^i{=}\mathbf{y}^i\right)} . \tag{13}$$

Thus, the plausibility is equal to 1, if the measurement returns exactly the real arm state, and otherwise a value < 1 but ≥ 0. The plausibility measure is commutative, i.e., $pl\left(\mathbf{Z}^i{=}\mathbf{z}^i, \mathbf{Y}^i{=}\mathbf{y}^i\right) = pl\left(\mathbf{Y}^i{=}\mathbf{y}^i, \mathbf{Z}^i{=}\mathbf{z}^i\right)$ for Gaussian type probability densities. For reading convenience, we drop the random variable notations $\mathbf{Z}$ and $\mathbf{Y}$ in the following.

Since the real arm state $\mathbf{y}^i$ is not known to MMF, $\mathbf{y}^i$ has to be substituted with an estimate. A good choice for a substitute is the reading from a different sensor $\mathbf{S}_2^j\left(\mathbf{z}_2^j, \mathbf{P}_2^j\right)$ in a different modality frame $j \neq i$, which is used for comparison with the sensor in question $\mathbf{S}_1^i\left(\mathbf{z}_1^i, \mathbf{P}_1^i\right)$. After the transformation from modality frame j to i, the plausibility is

$$pl\left(\mathbf{z}_1^i, \mathbf{z}_2^{j\to i}\right) = \frac{p\left(\mathbf{z}_1^i \,|\, \mathbf{z}_2^{j\to i}\right)}{p\left(\mathbf{z}_2^{j\to i} \,|\, \mathbf{z}_2^{j\to i}\right)} . \tag{14}$$

Considering the indirect dependence of $\mathbf{z}_1$ on $\mathbf{z}_2$ via the (unknown) real arm state $\mathbf{y}$, the numerator can be expressed as

$$p\left(\mathbf{z}_1^i|\mathbf{z}_2^{j\to i}\right) = \int \mathbf{dy}^i\; p\left(\mathbf{z}_1^i|\mathbf{y}^i\right) p\left(\mathbf{y}^i|\mathbf{z}_2^{j\to i}\right). \tag{15}$$

Both probability distributions on the right hand side are known. This yields

$$p\left(\mathbf{z}_1^i|\mathbf{z}_2^{j\to i}\right) = \int \mathbf{dy}^i\; N\left(\mathbf{y}^i, \mathbf{P}_1^i\right)\left(\mathbf{z}_1^i\right) \cdot N\left(\mathbf{z}_2^{j\to i}, \mathbf{P}_2^{j\to i}\right)\left(\mathbf{y}^i\right). \tag{16}$$

With $N(\mathbf{y}, \mathbf{P})(\mathbf{z}) = N(\mathbf{0}, \mathbf{P})(\mathbf{z} - \mathbf{y})$, a substitution $\mathbf{y}^i \mapsto \tilde{\mathbf{y}}^i = \mathbf{z}_1^i - \mathbf{y}^i$ and the definition $\Delta\mathbf{z}^{j,i} = \mathbf{z}_1^i - \mathbf{z}_2^{j\to i}$ this yields

$$p\left(\mathbf{z}_1^i|\mathbf{z}_2^{j\to i}\right) = \int \mathbf{d}\tilde{\mathbf{y}}^i\; N\left(\mathbf{0}, \mathbf{P}_1^i\right)\left(\tilde{\mathbf{y}}^i\right) \cdot N\left(\mathbf{0}, \mathbf{P}_2^i\right)\left(\Delta\mathbf{z}^{j,i} - \tilde{\mathbf{y}}^i\right), \tag{17}$$

which is the convolution $*$

$$p\left(\mathbf{z}_1^i|\mathbf{z}_2^{j\to i}\right) = \left(N\left(\mathbf{0}, \mathbf{P}_1^i\right) * N\left(\mathbf{0}, \mathbf{P}_2^i\right)\right)\left(\Delta\mathbf{z}^{j,i}\right) \tag{18}$$

$$= N\left(\mathbf{0}, \mathbf{P}_1^i + \mathbf{P}_2^i\right)\left(\Delta\mathbf{z}^{j,i}\right). \tag{19}$$

With (14) the plausibility yields

$$pl\left(\mathbf{z}_1^i|\mathbf{z}_2^{j\to i}\right) = \exp\left(-\tfrac{1}{2}\left(\Delta\mathbf{z}^{j,i}\right)^{\mathrm{T}}\left(\mathbf{P}_1^i + \mathbf{P}_2^{j\to i}\right)^{-1}\Delta\mathbf{z}^{j,i}\right) \tag{20}$$

This plausibility measures to which extent the measurements $\mathbf{z}_1^i$ and $\mathbf{z}_2^j$ of two sensors $\mathbf{S}_1^i$ and $\mathbf{S}_2^j$ (which both measure the same arm) are in agreement, given the transformations (4–5) and their respective variances.

As mentioned above, this is still only a substitute for the plausibility defined in (13). To improve this substitute, a sensor $\mathbf{S}_k$ is compared not only to one sensor but to all available redundant sensors $\mathbf{S}_l$ ($l \neq k$). For each comparison, a plausibility is calculated and the arithmetic mean over all plausibilities yields the total plausibility of sensor $\mathbf{S}_k$

$$pl_k = \sum_{l=1, l\neq k}^{n} \frac{1}{n-1}\, pl\left(\mathbf{z}_k^{i(k)}|\mathbf{z}_l^{j(l)\to i(k)}\right), \tag{21}$$

where $i(k)$, $j(l)$ are the modality frames, in which the sensors $\mathbf{S}_k$, $\mathbf{S}_l$ return their measurements.

A total of $n \geq 2$ sensors is required to detect that a sensor error occurred, and a total of $n \geq 3$ sensors is required to also assign the error to the faulty sensor. The proposed system uses $n = 4$, which are the simulated sensors in the four modality frames (GL, GO, LO, LA).

After a plausibility pl_k has been assigned to each sensor $\mathbf{S}_k$, an extension of 3.2 incorporates the plausibilities in the sensor fusion: Instead of choosing the weighting $\mathbf{W}_i$ proportional to the inverse variance as was done in (10), the plausibility now serves as another factor, i.e., $\mathbf{W}_k \propto pl_k\, \mathbf{P}_k^{-1}$. To satisfy the constraint $\sum_k \mathbf{W}_k = \mathbf{1}$, the weightings have to be normalized anew

$$\mathbf{W}_k = pl_k\, \mathbf{P}_k^{-1}\left(\sum_{l=1}^{n} pl_l\, \mathbf{P}_l^{-1}\right)^{-1}. \tag{22}$$

Thus, apart from the exact derivations of the local kinematic mappings (4–5), the model has been specified. The mappings are derived in the following section.

5 Transformations

To execute the transformations between modality frames, the mapping functions in (4–5) have to be specified. In this section, the derivation of these local forward and inverse kinematics mapping functions is presented. Before doing so, however, we specify the forward and inverse kinematic chains that need to be defined for the four modality frames in MMF and when which particular forward and inverse kinematic mapping is applied.

5.1 Local kinematic mappings

The local kinematic mappings for one limb and the respective joint are illustrated in Fig. 4. The forward mapping is the *chain* of *transformation steps* LA → LO → GO → GL and the inverse mapping is the inverse chain GL → GO → LO → LA. With those chains, an input state $\mathbf{s}^j$ can be transformed to any output state $\mathbf{s}^i$.

In the above Sects. 3.1, 3.2.1, 3.3 and 4, such chains had to be calculated for four input and up to four output spaces. The chains are implemented in two different ways. In the first approach, all input spaces are transformed separately into each output space. For the inverse kinematics, these transformations can be written as

$$f^{\mathrm{GL}\to\mathrm{LA}} = f^{\mathrm{LO}\to\mathrm{LA}}\left(f^{\mathrm{GO}\to\mathrm{LO}}\left(f^{\mathrm{GL}\to\mathrm{GO}}\left(\mathbf{s}^{\mathrm{GL}}\right)\right)\right) \tag{23}$$

$$f^{\mathrm{GO}\to\mathrm{LA}} = f^{\mathrm{LO}\to\mathrm{LA}}\left(f^{\mathrm{GO}\to\mathrm{LO}}\left(\mathbf{s}^{\mathrm{GO}}\right)\right) \tag{24}$$

$$f^{\mathrm{LO}\to\mathrm{LA}} = f^{\mathrm{LO}\to\mathrm{LA}}\left(\mathbf{s}^{\mathrm{LO}}\right) \tag{25}$$

$$f^{\mathrm{GL}\to\mathrm{LO}} = f^{\mathrm{GO}\to\mathrm{LO}}\left(f^{\mathrm{GL}\to\mathrm{GO}}\left(\mathbf{s}^{\mathrm{GL}}\right)\right) \tag{26}$$

$$f^{\mathrm{GO}\to\mathrm{LO}} = f^{\mathrm{GO}\to\mathrm{LO}}\left(\mathbf{s}^{\mathrm{GO}}\right) \tag{27}$$

$$f^{\mathrm{GL}\to\mathrm{GO}} = f^{\mathrm{GL}\to\mathrm{GO}}\left(\mathbf{s}^{\mathrm{GL}}\right) \tag{28}$$

(cf. Fig. 5, left, black).

In the second approach, any step of the forward and inverse chains is done only once for all inputs at the same time. For

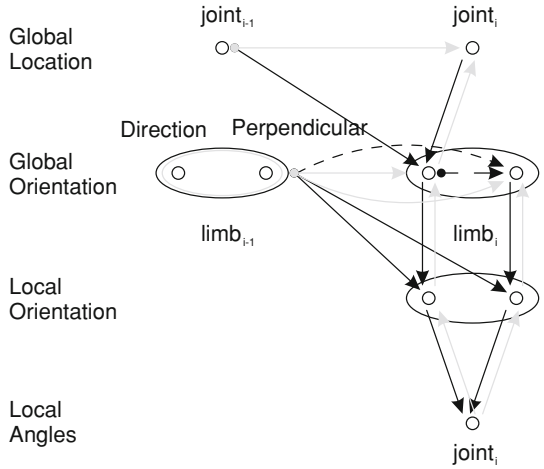


Fig. 4 Local mappings between modality frames for one limb: forward kinematic mappings (*gray*) and inverse kinematic mappings (*black*). The orientation spaces consist of two vectors, the first one parallel to the respective arm limb, the second one perpendicular to it. As perpendicular vector, the third rotation axis is chosen, as defined in Sect. A.3, especially Fig. 19. The *dashed arrow* mapping is defined only for some setups (cf. Sect. A.1)

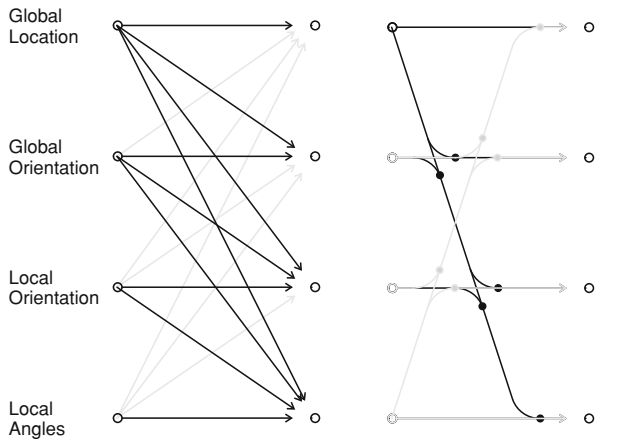


Fig. 5 The two types of transformations used. *Black*: inverse chain, *gray*: forward chain. **a** Each input space is transformed to each output space, where each transformation may consist of several steps. **b** When transforming stepwise, in the inverse chain, for example, the first state ($\mathbf{s}^{\mathrm{GL}}$) is transformed one step to frame GO where the result is fused with the second state ($\mathbf{s}^{\mathrm{GO}}$) and the fused state is transformed one step further, etc. *Dots* mark fusion points

example, for the mapping GL → LA, this can be written as follows:

$$f^{\mathrm{LO}\to\mathrm{LA}}\left(\mathrm{fuse}\left(\mathbf{x}^{\mathrm{LO}},\, f^{\mathrm{GO}\to\mathrm{LO}}\left(\mathrm{fuse}\left(\mathbf{x}^{\mathrm{GO}},\, f^{\mathrm{GL}\to\mathrm{GO}}\left(\mathbf{s}^{\mathrm{GL}}\right)\right)\right)\right)\right) \tag{29}$$

(cf. Fig. 5, right, black), where the fusion is accomplished by a weighted mean as described in (8–10).

Table 1 Transformation type chosen for the respective model steps

Separate	Stepwise
Prediction step (Sect. 3.1)	Sensor fusion (Sect. 3.2.1)
Plausibility calculation (Sect. 4)	Crosstalk step (Sect. 3.3)

The stepwise transformation with fusion approach has two advantages: first, fewer transformations need to be computed, and second, only parts of a state can be included (this becomes important in Sect. 5.2.1). However, the transformation is intermixed with fusion steps. When no intermediate fusion is desired, the first approach has to be chosen. Table 1 specifies which approach is applied in which stage of the MMF body state maintenance process.

5.2 Modality frame transformation steps

We now detail how each modality frame is represented in our implementation of the MMF model. We then detail the local forward and inverse kinematic mappings between neighboring modality frames (cf. Fig. 4), including their derivatives. While some of the mappings are rather standard, others had to be adjusted or newly derived due to the locality and the chosen modality frame encodings. For the sake of replicability and completeness, we detail all mappings.

The two orientation spaces (GO and LO) each consist of two 3D-spaces. The first encodes the direction (GD or LD, which are parallel to the arm limb) and the second encodes a perpendicular (GP or LP, which are perpendicular to the arm limb). Thus, six state vectors and covariance matrices $\mathbf{s}$ are maintained over time for each limb/joint: $\mathbf{s}^{\mathrm{LA}}$, $\mathbf{s}^{\mathrm{LD}}$, $\mathbf{s}^{\mathrm{LP}}$, $\mathbf{s}^{\mathrm{GD}}$, $\mathbf{s}^{\mathrm{GP}}$ and $\mathbf{s}^{\mathrm{GL}}$. In our simulation, each of these states is represented by a 3D-vector and a 3×3 covariance-matrix. Note that because the MMF model continuously updates the respective state and variance estimates stochastically dependent on the current noisy measurements, and also because of the interactions in the crosstalk step, the unit vector constraint cannot be guaranteed. Thus, our applied mappings need to be applicable when the orientation vectors are also only approximate unit vectors, which makes the derivations and resulting mappings noticeably more complex.

To derive the transformation equations, we denote the unit vector by $\mathbf{u}(\mathbf{V})$, the gradient by $\mathrm{grad}(U) = \nabla U$ of a scalar field U, the vector gradient (i.e., Jacobi-matrix) by $\mathrm{grad}(\mathbf{V}) = \nabla\mathbf{V}$ of a vector field $\mathbf{V}$, the divergence by $\mathrm{div}(\mathbf{V}) = \nabla \cdot \mathbf{V}$ and the rotation by $\mathrm{rot}(\mathbf{V}) = \nabla \times \mathbf{V}$. Moreover, we base our derivations on the derivation facts and equations (cf. Bronstein et al (2001))

$$\begin{aligned}\mathrm{grad}(\mathbf{V_1} \cdot \mathbf{V_2}) = (\mathrm{grad}\mathbf{V_1}) \cdot \mathbf{V2} + (\mathrm{grad}\mathbf{V_2}) \cdot \mathbf{V1} \\ + \mathbf{V_1} \times \mathrm{rot}\mathbf{V_2} + \mathbf{V_2} \times \mathrm{rot}\mathbf{V_1}\end{aligned} \tag{30}$$

$$\nabla \frac{1}{|\mathbf{V}|} = -\frac{1}{2\,|\mathbf{V}|^3} \nabla \left(\mathbf{V}^2\right) = -\frac{(\nabla \mathbf{V}) \cdot \mathbf{V} + \mathbf{V} \times \mathrm{rot}\mathbf{V}}{|\mathbf{V}|^3} \tag{31}$$

$$\nabla \mathbf{u}\,(\mathbf{V}) = \frac{\nabla \mathbf{V}}{|\mathbf{V}|} + \mathbf{V} \otimes \nabla \frac{1}{|\mathbf{V}|} \tag{32}$$

$$= \frac{\nabla \mathbf{V}}{|\mathbf{V}|} - \mathbf{eV} \otimes \frac{(\nabla \mathbf{V}) \cdot \mathbf{eV} + \mathbf{eV} \times \mathrm{rot}\mathbf{V}}{|\mathbf{V}|} \tag{33}$$

where $\otimes$ is the dyadic product. Of particular interest is the vector-gradient of the unit-vector $\mathbf{u}\,(\mathbf{V})$ with respect to $\mathbf{V}$:

$$\nabla_{\mathbf{V}} \mathbf{u}\,(\mathbf{V}) = \frac{\mathbf{I} - \mathbf{eV} \otimes \mathbf{eV}}{|\mathbf{V}|} \tag{34}$$

For the three steps in the transformation chain, i.e., GL $\leftrightarrow$ GO, GO $\leftrightarrow$ LO and LO $\leftrightarrow$ LA, we first derive the **inverse** kinematic transformations of the means and the Jacobians, followed by the **forward** kinematic transformations of the means and Jacobians. Together, these transformations form the mappings (4–5).

5.2.1 *Global location $\leftrightarrow$ global orientation*

We start with the first mapping in the **inverse** chain. The mapping specifies, for example, how the elbow location respective the shoulder can be used to determine the direction of the upper arm limb in a global frame of reference. This global direction state $\mathbf{x}^{\mathrm{GD}}$ is simply the unit vector of the difference of two succeeding limb end points (cf. Fig. 1)

$$\mathbf{x}^{\mathrm{GD}} = \mathbf{u}\left(\mathbf{x}^{\mathrm{GL}} - \mathbf{x}_u^{\mathrm{GL}}\right) \tag{35}$$

where u specifies the upper arm limb.

The **Jacobians** of $\mathbf{x}^{\mathrm{GD}}$ are given by

$$\frac{\partial \mathbf{x}^{\mathrm{GD}}}{\partial \mathbf{x}^{\mathrm{GL}}} = \frac{\mathbf{I} - \mathbf{u}\left(\mathbf{x}^{\mathrm{GL}} - \mathbf{x}_u^{\mathrm{GL}}\right) \otimes \mathbf{u}\left(\mathbf{x}^{\mathrm{GL}} - \mathbf{x}_u^{\mathrm{GL}}\right)}{\left|\mathbf{x}^{\mathrm{GL}} - \mathbf{x}_u^{\mathrm{GL}}\right|} \tag{36}$$

$$\frac{\partial \mathbf{x}^{\mathrm{GD}}}{\partial \mathbf{x}_u^{\mathrm{GL}}} = -\frac{\mathbf{I} - \mathbf{u}\left(\mathbf{x}^{\mathrm{GL}} - \mathbf{x}_u^{\mathrm{GL}}\right) \otimes \mathbf{u}\left(\mathbf{x}^{\mathrm{GL}} - \mathbf{x}_u^{\mathrm{GL}}\right)}{\left|\mathbf{x}^{\mathrm{GL}} - \mathbf{x}_u^{\mathrm{GL}}\right|} \tag{37}$$

(cf. 34).

The Jacobian (36) can serve as an example for the aforementioned necessity to include the unit-vector constraint in the transformation. Choosing e.g., $\mathbf{x}^{\mathrm{GL}} - \mathbf{x}_u^{\mathrm{GL}} \equiv (1, 0, 0)^T$ leads to

$$\frac{\partial \mathbf{x}^{\mathrm{GD}}}{\partial \mathbf{x}^{\mathrm{GL}}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{38}$$

$$\Delta \mathbf{x}^{\mathrm{GD}} = \left(0, \Delta \mathbf{x}_2^{\mathrm{GL}}, \Delta \mathbf{x}_3^{\mathrm{GL}}\right)^T \tag{39}$$

with $\Delta \mathbf{x}_1^{\mathrm{GD}} = 0$, correctly as the direction vector can change only perpendicular to its current direction. [4]

The analogous part of the **forward** kinematics mapping is

$$\mathbf{x}^{\mathrm{GL}} = \mathbf{x}_u^{\mathrm{GL}} + l\,\mathbf{u}\left(\mathbf{x}^{\mathrm{GD}}\right). \tag{40}$$

where l is the length of the specified arm limb. The **Jacobian** with respect to $\mathbf{x}_u^{\mathrm{GL}}$ is

$$\frac{\partial \mathbf{x}^{\mathrm{GL}}}{\partial \mathbf{x}_u^{\mathrm{GL}}} = \mathbf{I} \tag{41}$$

and with respect to $\mathbf{x}^{\mathrm{GD}}$:

$$\frac{\partial \mathbf{x}^{\mathrm{GL}}}{\partial \mathbf{x}^{\mathrm{GD}}} = \frac{l_1}{|\mathbf{x}^{\mathrm{GD}}|}\left(\mathbf{I} - \mathbf{u}\left(\mathbf{x}^{\mathrm{GD}}\right) \otimes \mathbf{u}\left(\mathbf{x}^{\mathrm{GD}}\right)\right) \tag{42}$$

(analogous to (36) with $\mathbf{x}_u^{\mathrm{GL}} = \mathbf{0}$).

Above, for the inverse direction, only the direction part $\mathbf{x}^{\mathrm{GL}} \rightarrow \mathbf{x}^{\mathrm{GD}}$ is defined. In general, the perpendicular part $\mathbf{x}^{\mathrm{GL}} \rightarrow \mathbf{x}^{\mathrm{GP}}$ is undefined. If stepwise transformations are used (cf. Fig. 5, right), this does not pose a problem: in the subsequent step GO $\rightarrow$ LO the perpendicular part is simply defined by the GO-source alone. However, if separate transformations are used, the perpendicular part has to be *copied* from the global orientation state $\mathbf{x}^{\mathrm{GO}}$, and it is thus included twice: first in the chain GL $\rightarrow$ LA (23) and second in the chain GO $\rightarrow$ LA (24). This may introduce small but undesired interactions in the prediction step and may also slightly distort the plausibility calculation, which are computed based on the separate transformation approach (cf. Fig. 5 and Table 1).

5.2.2 *Global orientation $\leftrightarrow$ local orientation*

The second step in the **inverse** chain is to transform a global orientation into a body-relative orientation. This is a change of coordinate systems, where the new coordinate system is spanned by the unit vectors $\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right)$, $\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}} \times \mathbf{x}_u^{\mathrm{GD}}\right)$ and $\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}}\right)$. We start with the direction part, i.e. $\mathbf{x}^{\mathrm{GD}} \rightarrow \mathbf{x}^{\mathrm{LD}}\left(\mathbf{x}^{\mathrm{GD}}\right)$

$$\mathbf{x}^{\mathrm{LD}}\left(\mathbf{x}^{\mathrm{GD}}\right) = \begin{pmatrix} \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right)^T \\ \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}} \times \mathbf{x}_u^{\mathrm{GD}}\right)^T \\ \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}}\right)^T \end{pmatrix} \cdot \mathbf{x}^{\mathrm{GD}} \tag{43}$$

(cf. Fig. 6)

As both the specified limb's state $\mathbf{x}^{\mathrm{GD}}$ and the vectors spanning the local coordinate system are noisy, the **Jacobians** with respect to $\mathbf{x}^{\mathrm{GD}}$, $\mathbf{x}_u^{\mathrm{GD}}$ and $\mathbf{x}_u^{\mathrm{GP}}$ are required. Again,

[4] Note that, if the unit vector constraint was not included in the transformation (i.e., if (35) was replaced with $\mathbf{x}^{\mathrm{GD}} = \mathbf{x}^{\mathrm{GL}} - \mathbf{x}_u^{\mathrm{GL}}$), this would lead to $\frac{\partial \mathbf{x}^{\mathrm{GD}}}{\partial \mathbf{x}^{\mathrm{GL}}} = \mathbf{I}$ and thus generally $\Delta \mathbf{x}_1^{\mathrm{GD}} \neq 0$, resulting in wrong variance estimates (ultimately also for e.g., angles).

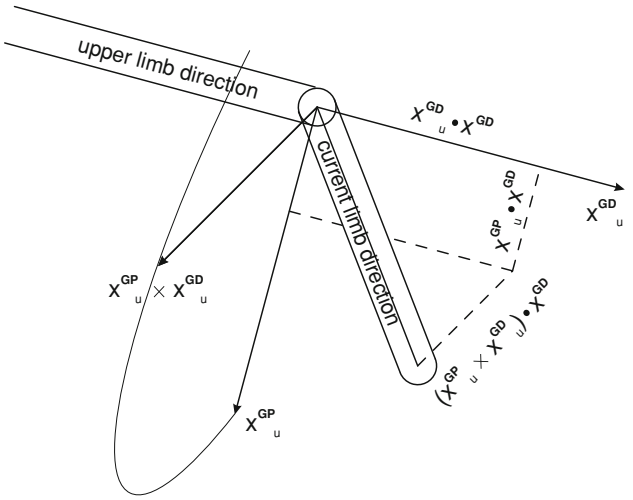


Fig. 6 Transformation from a global to a body relative coordinate system. The new coordinate system is spanned by the unit vectors $\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right)$, $\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}} \times \mathbf{x}_u^{\mathrm{GD}}\right)$ and $\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}}\right)$

as discussed above, derivatives with respect to unit vectors do not suffice. The Jacobian with respect to $\mathbf{x}^{\mathrm{GD}}$ yields

$$\frac{\partial \mathbf{x}^{\mathrm{LD}}}{\partial \mathbf{x}^{\mathrm{GD}}} = \begin{pmatrix} \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right)^T \\ \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}} \times \mathbf{x}_u^{\mathrm{GD}}\right)^T \\ \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}}\right)^T \end{pmatrix}. \tag{44}$$

The Jacobian with respect to $\mathbf{x}_u^{\mathrm{GD}}$ is derived starting from (A.4) to (A.19) and yields

$$\frac{\partial \mathbf{x}^{\mathrm{LD}}}{\partial \mathbf{x}_u^{\mathrm{GD}}} = \begin{pmatrix} \frac{1}{\left|\mathbf{x}_u^{\mathrm{GD}}\right|}\left(\mathbf{x}^{\mathrm{GD}} - \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right)\left(\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right)\cdot\mathbf{x}^{\mathrm{GD}}\right)\right)^T \\ \frac{-1}{|\mathbf{a}|}\left(\mathbf{x}_u^{\mathrm{GP}} \times \left(\mathbf{x}^{\mathrm{GD}} - \mathbf{u}\left(\mathbf{a}\right)\left(\mathbf{u}\left(\mathbf{a}\right)\cdot\mathbf{x}^{\mathrm{GD}}\right)\right)\right)^T \\ 0 \end{pmatrix}. \tag{45}$$

For reasons of symmetry, the Jacobian with respect to $\mathbf{x}_u^{\mathrm{GP}}$ can be derived from (45) and yields

$$\frac{\partial \mathbf{x}^{\mathrm{LD}}}{\partial \mathbf{x}_u^{\mathrm{GP}}} = \begin{pmatrix} 0 \\ \frac{1}{|\mathbf{a}|}\left(\mathbf{x}_u^{\mathrm{GD}} \times \left(\mathbf{x}^{\mathrm{GD}} - \mathbf{u}\left(\mathbf{a}\right)\left(\mathbf{u}\left(\mathbf{a}\right)\cdot\mathbf{x}^{\mathrm{GD}}\right)\right)\right)^T \\ \frac{1}{\left|\mathbf{x}_u^{\mathrm{GP}}\right|}\left(\mathbf{x}^{\mathrm{GD}} - \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}}\right)\left(\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}}\right)\cdot\mathbf{x}^{\mathrm{GD}}\right)\right)^T \end{pmatrix} \tag{46}$$

Note that this transformation can be applied to any vector, so that the same transformation also holds for global perpendicular $\mathbf{x}^{\mathrm{GP}}$ to local perpendicular $\mathbf{x}^{\mathrm{LP}}$.

The **forward** mapping is given by

$$\mathbf{x}^{\mathrm{GD}} = \left(\mathbf{x}^{\mathrm{LD}}\right)_1 \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right) + \left(\mathbf{x}^{\mathrm{LD}}\right)_2 \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}} \times \mathbf{x}_u^{\mathrm{GD}}\right) + \left(\mathbf{x}^{\mathrm{LD}}\right)_3 \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}}\right), \tag{47}$$

the **Jacobian** with respect to $\mathbf{x}^{\mathrm{LD}}$ yields

$$\frac{\partial \mathbf{x}^{\mathrm{GD}}}{\partial \mathbf{x}^{\mathrm{LD}}} = \left(\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right)\ \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}} \times \mathbf{x}_u^{\mathrm{GD}}\right)\ \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}}\right)\right), \tag{48}$$

the Jacobian with respect to $\mathbf{x}_u^{\mathrm{GD}}$

$$\frac{\partial \mathbf{x}^{\mathrm{GD}}}{\partial \mathbf{x}_u^{\mathrm{GD}}} = \frac{\mathbf{I} - \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right) \otimes \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right)}{\left|\mathbf{x}_u^{\mathrm{GD}}\right|}\left(\mathbf{x}^{\mathrm{LD}}\right)_1 + \left(\frac{-\varepsilon \cdot \mathbf{x}_u^{\mathrm{GP}}}{\left|\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right|} + \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right) \otimes \frac{\mathbf{x}_u^{\mathrm{GP}} \times \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right)}{\left|\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right|}\right)\left(\mathbf{x}^{\mathrm{LD}}\right)_2 \tag{49}$$

(cf. (A.21)) and the Jacobian with respect to $\mathbf{x}_u^{\mathrm{GP}}$

$$\frac{\partial \mathbf{x}^{\mathrm{GD}}}{\partial \mathbf{x}_u^{\mathrm{GP}}} = \frac{\mathbf{I} - \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}}\right) \otimes \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}}\right)}{\left|\mathbf{x}_u^{\mathrm{GP}}\right|}\left(\mathbf{x}^{\mathrm{LD}}\right)_3 - \left(\frac{-\varepsilon \cdot \mathbf{x}_u^{\mathrm{GD}}}{\left|\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right|} + \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}} \times \mathbf{x}_u^{\mathrm{GD}}\right) \otimes \frac{\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}} \times \mathbf{x}_u^{\mathrm{GD}}\right)}{\left|\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right|}\right)\left(\mathbf{x}^{\mathrm{LD}}\right)_2 \tag{50}$$

(cf. (A.22)). Again, the transformation from a local orientation state to a global orientation state is the same.

5.2.3 *Local orientation ↔ local angles*

The mapping from a local orientation state to a local angle state and vice versa is a one-to-one mapping: it does not include another limb. It is, however, not trivial, especially as Jacobians and broken constraints of unit vectors are considered. Appendix A.3 describes our implementation of this one-to-one mapping.

5.3 Computational cost

In order to estimate the computational cost c of our model, we consider the number of state transformations necessary during the MMF processing stages of each iteration. Each transformation consists of a varying number of steps (cf. (23–28)). We assume, that the longest inverse transformation GL → LA and forward transformation have approximately the same cost. They thus provide an upper bound for all transformations, and we derive the following in units of the cost c(GL → LA).[5]

[5] Note that the order of the cost c(GL → LA) is generally the same for one through three DOFs per limb. For example, if a planar arm moves in a plane, which is not perpendicular to any of the coordinate axes, then full 3D-vectors would still be required for $\mathbf{x}^{\mathrm{GL}}$, $\mathbf{x}^{\mathrm{GO}}$, $\mathbf{x}^{\mathrm{GP}}$, $\mathbf{x}^{\mathrm{LD}}$ and $\mathbf{x}^{\mathrm{LP}}$. Only the transformation step LO → LA would need fewer computations, but not its inverse LA → LO or the other transformation steps.

With n the number of modality frames and k the number of limbs, the costs for the different steps are

$$c\,(\text{Sensor Fusion}) = 2k$$
$$c\,(\text{Filter Crosstalk}) = 2k$$
$$c\,(\text{Prediction Step}) = 2nk$$
$$c\,(\text{Plausibility Calculation}) = n^2k$$

(cf. Fig. 5; Table 1). In the Prediction Step, the transformations need only be done to the output space (LA) and back. In the Plausibility Calculation, separate forward and inverse transformations must be done for all modality frames separately. Thus, this is the most costly computation stage in MMF. However, the quadratic dependence on the modality frames, which are typically bounded to a small number should not result in a significant processing bottleneck.

When pursuing a goal-directed control task, the state estimate in a task space (typically a location space) will need to be transformed once per time step to the angular space to infer the currently desired motor command. For one limb, the cost is 1; for k limbs, it is approximately

$$c(\text{Non modular}) = k$$

Comparing both, we get the computational overhead, which is due to the usage of multiple modality frames and due to interactions. The effects of the sensor fusion and filter crosstalk are small, the prediction step increases the order of complexity with respect to the number of modality frames by one, and the plausibility calculation again by one. Thus, the addition of further redundant frames of reference has a quadratic scaling effect, as long as full plausibility computations are done. With respect to the number of limbs, the whole model scales linearly so that, for example, the addition of a neck or torso increases the complexity only linearly.

5.3.1 Parallelization

We did not use parallel computation but it can be done with respect to modality frames and limbs. With respect to modality frames, each arrow (Fig. 5, left) of the costly separate transformations can be computed in parallel (cf. also Fig. 2), thus reducing the costs above for all steps to k per parallel unit. In addition, the update step (Sect. 3.2) and the crosstalk step (Sect. 3.3) can be computed in parallel for all limbs if the transformations are based on the previous time step.

6 Experiments

The following experiments investigate whether the sensor information is integrated correctly and how well sensor failure and its effects on the state estimation are captured by our plausibility measure. All results are obtained from a simulation with a kinematic arm that has three limbs and three degrees of freedom in each joint. Thus, an arm with a total of nine degrees of freedom is tested in 3D space. The MMF model thus maintains three dimensional state estimation vectors in each modality frame. In the presented results, only the diagonal components of the modality frame-respective covariance matrices were considered.

Table 2 *GENERAL* settings for sensor standard deviations in all modality frames

LA	LO	GO	GL
0.05rad	0.05	0.05	$0.05 \cdot l$

The sensor standard deviations are equal and independent on different dimensions (i.e., the covariance-matrices are diagonal). The diagonal elements are chosen to be approximately equal in their magnitude (accounting for the transformations (4–5)).

The lengths of all arm limbs are set to 0.5. An ensemble of 150 runs from random start configurations to random angular goal constellations, both given in the angular modality frame LA for all limbs, was generated for each setting. The nine dimensional desired movement is the vector from the estimated arm posture (state estimates in $\mathbf{x}^{\mathrm{LA}}$) to the given angular goal

$$\Delta\mathbf{x}^{\mathrm{LA}}_{\mathrm{des}} = \mathbf{goal}^{\mathrm{LA}} - \mathbf{x}^{\mathrm{LA}}. \tag{51}$$

For each limb m, the respective 3D-part $\Delta\mathbf{x}^{\mathrm{LA}}_{\mathrm{des,m}}$ of the desired movement is capped by a maximum velocity, such that $\left|\Delta\mathbf{x}^{\mathrm{LA}}_{\mathrm{des,m}}\right| \leq \sqrt{3} \cdot 0.1\mathrm{rad}$. Onto this scaled desired movement-vector, a velocity dependent motor noise equal to $0.5 \cdot \left|\Delta\mathbf{x}^{\mathrm{LA}}_{des,m}\right| \frac{1}{\sqrt{3}}$ per dimension and an absolute motor noise equal to 0.005rad per dimension is added. Thus, for big distances between the arm and the goal—where the arm is moving quickly—the velocity-dependent part dominates.

Sensory noise is added to each modular modality frame. We study a setup with approximately equal sensory information available to each modality frame. Thus, we account for the modality frame transformations starting with angular noise of 0.05 rad in the angular modality frame. Table 2 specifies the exact standard deviations applied to the sensors of the respective modality frames. Appendix B details the derivation of these approximately corresponding noise magnitudes.

All results are shown in the global location modality frame (GL) for the second joint.

6.1 Sensor failure detection

In Sect. 4, a convolution of a target and a reference sensor was introduced as the plausibility that both states are in accordance, and an arithmetic mean of plausibilities over multiple references was proposed to be used as the final plausibility of the target modality frame, resulting in (21). Here, we show

Table 3 *Modified* sensor noise settings for the GL modality frame

Setup [a]	Sensor distribution in units of limb length	
	Mean	SD
1a	$y_i^{GL} + 0.5\text{rect}(t)$	$0.005 + 0$
1b	$y_i^{GL} + 0.5\text{sine}(t)$	$0.005 + 0$
1c	$y_i^{GL} + 0$	$0.005 + 0.5\text{rect}(t)$
1d	$y_i^{GL} + 0$	$0.005 + 0.5\text{sine}(t)$

The *known* sensor noise in GL is lowered to 0.005. In the second dimension of the GL-modality frame, an additional noise component or an offset component is added. This additional disruptive component is *unknown* to MMF. As specified in (12), the sensor distribution for the second dimension is a normal distribution

[a]Setup 1a: rectangular offset, 1b: sinusoidal offset, 1c: rectangular increase in variance, 1d: sinusoidal increase in variance. The time-dependent functions rect (t) and sine (t) are displayed in Fig. 7.

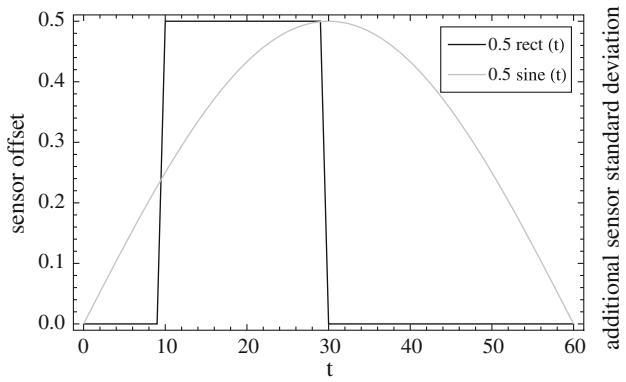


Fig. 7 Time dependency: A rectangular or sinusoidal function is used as either an offset on the Global Location Sensor or as an increase in the Global Location Sensor's standard deviation. In all four settings (cf. also Table 3), this change was unknown to the system, that is, MMF expected an unbiased sensor in GL with a standard deviation of 0.005

that this plausibility can be used to reliably detect sudden sensor offsets or increases in variance. Of particular interest is the scenario where a highly accurate sensor (i.e., with low noise) suddenly returns erroneous measurements. In this case, the system has to switch from strongly relying on this sensor to strongly disregarding it. To model this, a setup is presented in which the Global Location-Sensor has errors but at the same time a lower variance than all other sensors (cf. Table 3).

In Setup 1a, a rectangular offset to this low-noise sensor is introduced at time step 10 and switched off again at time step 30 (cf. Fig. 7, black, left ordinate axis).

Figure 8 shows how the plausibility of the Global Location-Sensor decreases during the application of the offset.

Although the variance is very low, which has an enhancing effect on the weight (cf. $\mathbf{P}_k^{-1}$ in (22)), the low variance has a reducing effect on the plausibility (cf. (20) and Fig. 8), which in turn has a reducing effect on the weight (22). Overall, the

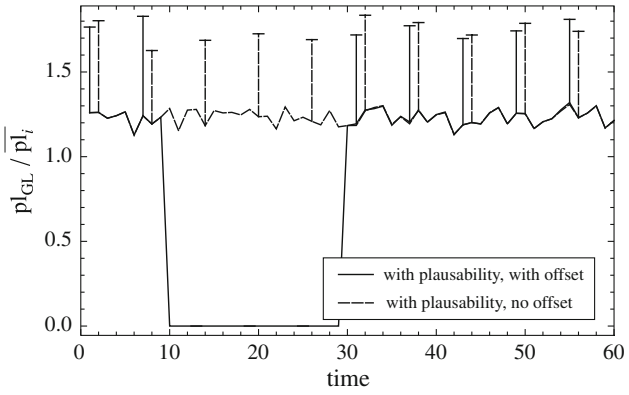


Fig. 8 Setup 1a: The plausibility of the Global Location-Sensor is displayed over time. For the *solid line*, an offset is introduced at time step 10 and is switched off again at time step 30. In this interval, the plausibility reduces strongly as opposed to a system with a non-failing sensor (*dashed line*)

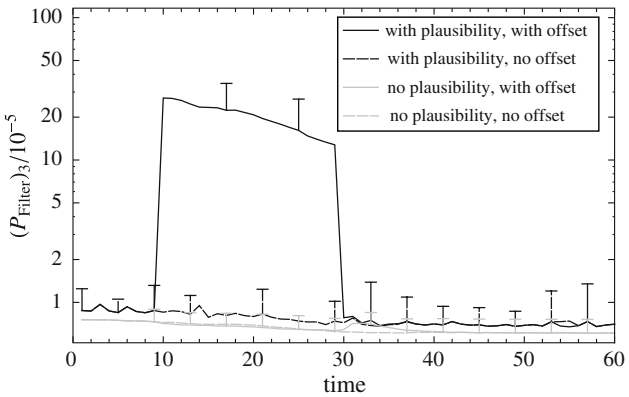


Fig. 9 Setup 1a: The inner filter variance is displayed. An increase is apparent during the application of the offset, which shows an overall loss of information. This is due to a loss of information in the Global Location modality frame, whose sensor is discarded

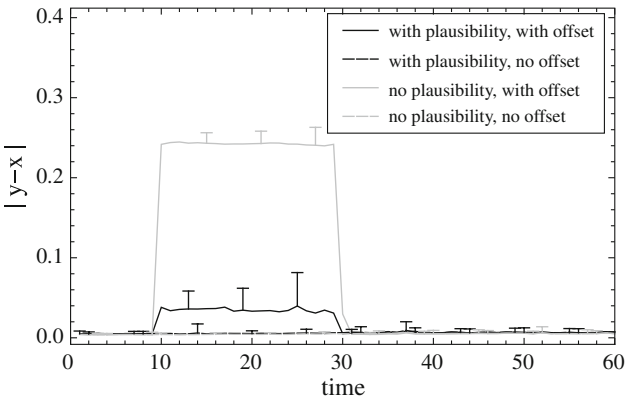


Fig. 10 Setup 1a: The error in the state estimate (i.e., the Euclidean difference of the estimate **x** from the real arm state **y**) is displayed. The introduction of the plausibility helps to minimize this estimation-error: The *solid black graph* is close to the *dashed ones*, while there is a large difference between the *solid gray graph* and the *dashed ones*

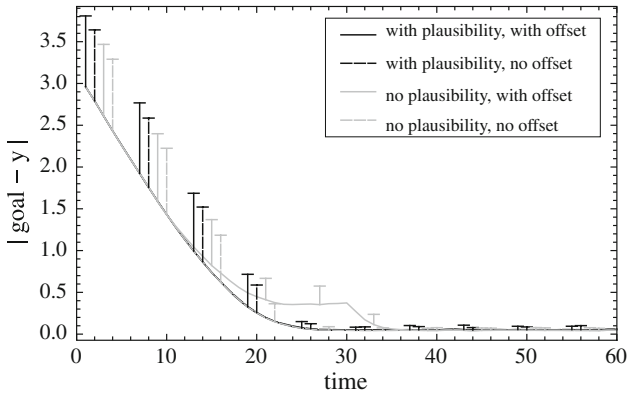


Fig. 11 Setup 1a: The ability to reach goals (i.e., the Euclidean distance of the real arm state from the goal) is displayed. The introduction of plausibility enables the system to reach goals in a setup with a sensor offset nearly as good as in a setup without a sensor offset. If no plausibility is used, the system fails to reach the goal while an offset is present

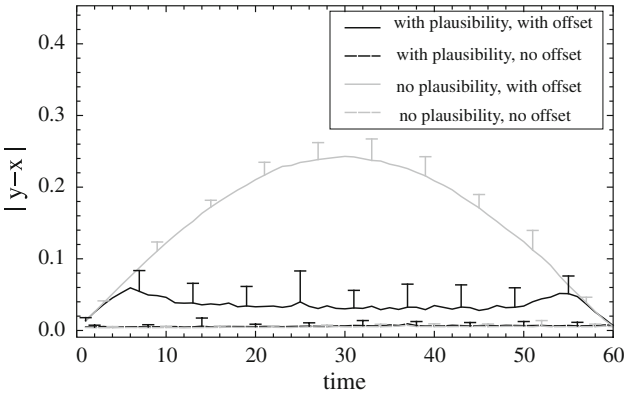


Fig. 12 Setup 1b (sinusoidal sensor offset): The error in the state estimate (i.e., the Euclidean difference of the estimate **x** from the real arm state **y**) is displayed. The introduction of the plausibility helps to minimize this estimation-error: The *solid black graph* is close to the *dashed ones* while there is a large difference between the *solid gray graph* and the *dashed ones*. For low offsets, the system cannot differentiate between a sensor failure or deviations which occur due to known Gaussian noise

weight $\mathbf{W}_{\mathrm{GL}}$ for merging the Global Location-Sensor in (8–10) decreases, and the Global Location-Sensor is discarded while the offset is present. The disregard of the Global Location-Sensor leads to a higher filter variance. This variance serves as a measure of how good the state estimate of the complete system is. Thus, the system "knows" that its estimate is less accurate during this phase, as apparent in Fig. 9. As the influence of the faulty Global Location-Sensor is suppressed and the influence of all other sensors is increased, errors in the state estimate are minimized (c.f. Fig. 10). Finally, Fig. 11 shows how the minimization of the errors in the state estimate enables the system to reach the goal effectively, even under the influence of a strong offset.

To study continuous increases of the sensor offset as well as increases in the sensor variance, setups 1b–1d (cf. Table 3) were simulated. The estimation error $|\mathbf{y} - \mathbf{x}|$ is a good indicator of the system's estimation capabilities, thus only $|\mathbf{y} - \mathbf{x}|$ is shown for these setups. The estimation error for the sinusoidal offset (setup 1b, cf. Fig. 7, gray, left ordinate axis) is shown in Fig. 12. The comparison to Fig. 10 shows that the estimation error is approximately equal for equal amplitudes (at the peak of the sinus) and no "memory effects" are apparent.

Our plausibility measure (20) suggests that offsets which are lower than the standard deviation can barely be detected. The standard deviation of non-location-frames transformed to the location frame is $\sigma^{\neq \mathrm{GL} \to \mathrm{GL}} = 0.025$ (cf. Table 3 with $l = 0.5$). This is approximately the range where the sensor-failure-detection does not work, i.e., where there is no difference between the black solid and gray solid lines (Fig. 12).

Comparing sensor offsets (setups 1a and 1b, Fig. 10 and 12) to increases in the sensor standard deviation (setups 1c and 1d, Fig. 13 and 14) yields the same conclusions: The system can detect and avoid both sudden and continuous sensor failure as long as the magnitude of the error exceeds the "normal" standard deviation. MMF is highly noise robust, especially when multiple sensors are integrated. This has been confirmed in an earlier study (Ehrenfeld and Butz 2011).

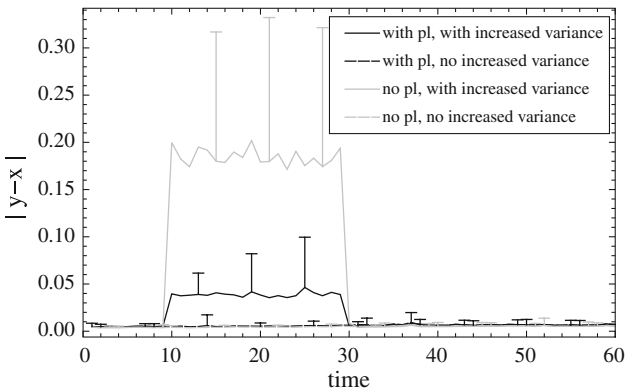


Fig. 13 Setup 1c (rectangular additional noise): The error in the state estimate (i.e., the Euclidean difference of the estimate **x** from the real arm state **y**) is displayed. The introduction of the plausibility helps to minimize this estimation-error: The *solid black graph* is close to the *dashed ones* while there is a large difference between the *solid gray graph* and the *dashed ones*

6.2 Importance of independence

The MMF system evaluated so far may be described as handling *interactive filters* (Fig. 15, solid lines). This interactivity is realized by first fusing all sensors (cf. Sect. 3.2.1) and then integrating them in the body state estimates (cf. Sect. 3.2.2). Thus, information is gained by fusing *independent* informa-

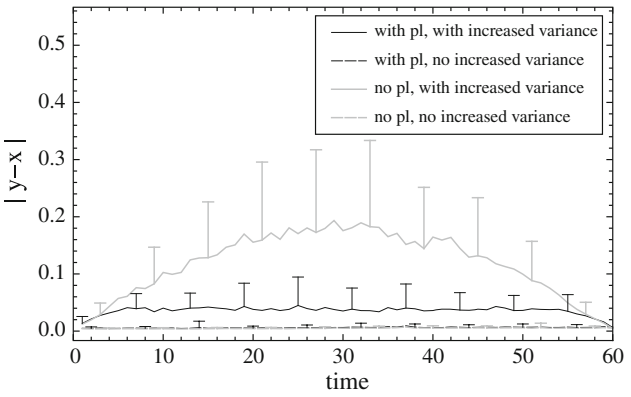


Fig. 14 Setup 1d (sinusoidal additional noise): The error in the state estimate (i.e., the Euclidean difference of the estimate **x** from the real arm state **y**) is displayed. The introduction of the plausibility helps to minimize this estimation-error: The *solid black graph* is close to the dashed ones while there is a large difference between the *solid gray graph* and the *dashed ones*. Small exceedance of the known Gaussian noise is not detected by the system

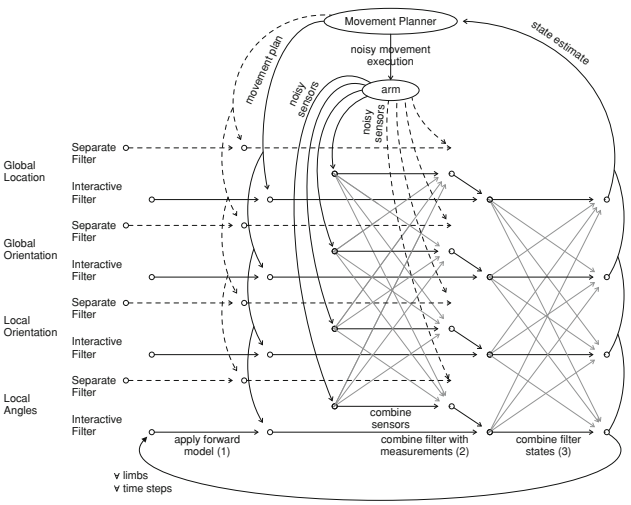


Fig. 15 Data flow in the filter system. The interactive filters (*solid lines*) integrate a weighted mean of sensor measurements and interact in the crosstalk step. The separate filters (*dashed lines*) are each independent on all other filters and integrate only one sensor

tion sources, while in 3.3, a modified weighted mean using (8, 10 and 11) "shifts" the filter states towards each other.

To show the importance of accounting for the fusion of *independent* sources, the proposed setup is compared to a second setup. In the second setup, a set of *separate filters* is used as shown in Fig. 15 with dashed lines. Each sensor is integrated independently into its respective modality frame module but in the crosstalk step the weighted mean is used for the variance estimates as well (8–10). In this way, the information from the other sensors is integrated *indirectly* via the other filter states.

To compare the performance of these two setups, the estimated variances are displayed in Fig. 16 for all four

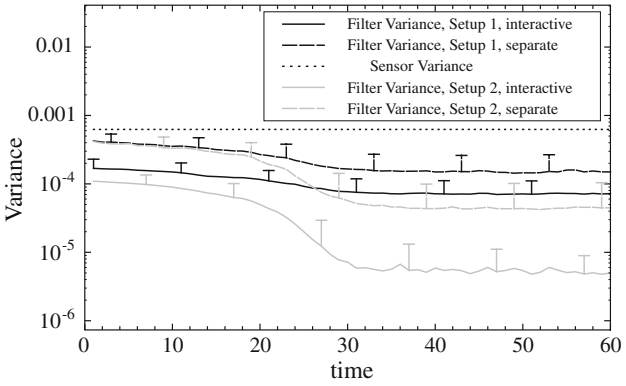


Fig. 16 The variance estimates are lower when dependent sources are fused (*gray*) as opposed to independent sources (*black*). Also the ratio $\frac{\mathbf{P}_{\text{separate}}}{\mathbf{P}_{\text{interactive}}}$ is too large for dependent sources. This ratio, which describes the gain of information, is overestimated in the gray setup, i.e., the gray interactive filter variance is underestimated

experiments (with no sensor failure and the parameters from Table 2). The separate systems serve as a reference and use the same trajectories as the respective interactive system. The ratio of information $\frac{\mathbf{P}^{-1}_{\text{interactive}}}{\mathbf{P}^{-1}_{\text{separate}}} = \frac{\mathbf{P}_{\text{separate}}}{\mathbf{P}_{\text{interactive}}}$ is a factor of how much information is gained through interaction. Even for the most extreme case of equal variances in all modality frames, it should be $\frac{\mathbf{P}_{\text{separate}}}{\mathbf{P}_{\text{interactive}}} \leq 4$. However, for most of the runtime, the gray factor is $\frac{\mathbf{P}_{\text{separate}}}{\mathbf{P}_{\text{interactive}}} \approx 10$, as shown in Fig. 16. Thus, Setup 2 (gray) overestimates the sensory information. That is, it underestimates the remaining variance.

This overestimation has severe consequences: first, because the system believes too strong in its estimate, this estimate converges too early and disregards sensor measurements, which could otherwise improve the estimate. Thus, the accuracy of the estimate decreases dramatically, which is shown as the Euclidean distance of the state estimate **x** from the real arm state **y** half-logarithmic in Fig. 17. Because the movement plan is based on the mean-estimate **x**, the arm stops moving as soon as **x** reaches and stays at the goal. The low variance **P** (c.f. Fig. 16) prevents **x** from being corrected by the measurement and the mainly velocity-dependent motor noise prevents **P** from increasing again. Thus, **x** stays at the goal and the arm is locked in place, even if the real arm state **y** has not yet reached the goal. Figure 18 shows this difference of the real arm state **y** from the goal.

These results show that Setup 2 performs worse than the actually proposed MMF setup. This is due to the fact that the "shift" of the filter states in our proposed setup accepts *dependent* sources, while the "normal" weighted mean of filter states in Setup 2 requires *independent* sources. However, because the filter states depend on each other due to the coupling in the crosstalk step, Setup 2 copies information and re-integrates it in successive time steps. Thus, the

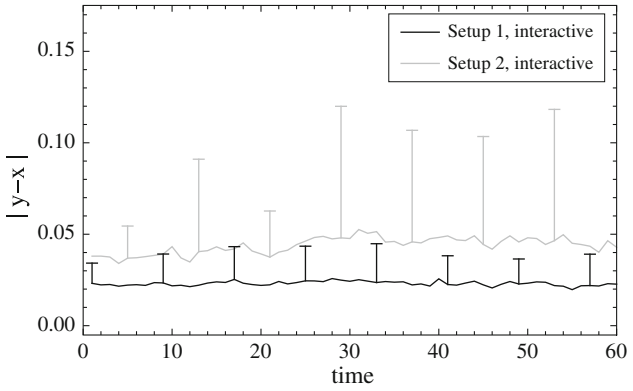


Fig. 17 The Euclidean distance of the real arm state from the estimated state shows that a worse performance is apparent when dependent sources are merged (*gray*)

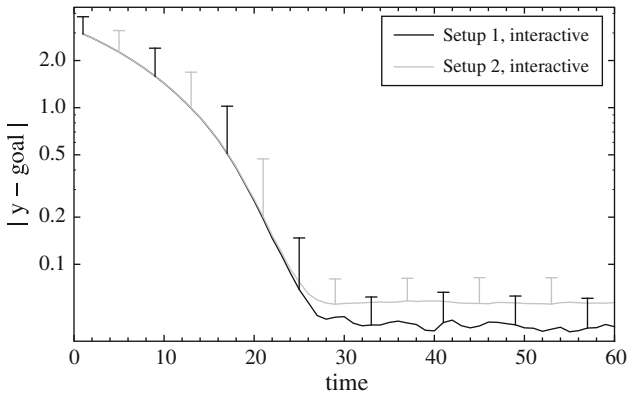


Fig. 18 The Euclidean distance of the goal from the real arm state is larger for the gray setup. Thus, even the control performance is affected negatively

state variance is underestimated. By providing quantitative results for these errors, we showed that it is imperative to not increase information by fusing dependent sources (such as state estimates from different Kalman Filters as done in Setup 2), but rather to fuse *independent* sources only (e.g., different redundant sensors).

7 Related work

The MMF model combines ideas from several related robotic and neurally inspired architectures. To position MMF in the robotics context, we first classify MMF in a framework of standard approaches. Next, we address several related modular and sensor-related robotics architectures.

7.1 MMF classifications

The MMF model can be classified in a framework of standard approaches. MMF essentially encodes a serial chain (Scheinmann and McCarthy 2008) of a generally arbitrary number of limbs, where each joint can have up to 3DOFs (rotations). Translations could be added by additional "translation" limbs. For realizing an extendable limb, the length l in (40) could be changed to an additional modality frame, which would estimate limb length and which would also be updated and maintained over time. The body state is represented probabilistically with means and covariances, as Gaussian noise is assumed. To cope with noise, multiple sensors are fused and integrated into the body state estimate. According to Durrant-Whyte and Henderson (2008), the data fusion can be classified as follows: We use two types of probabilistic data fusion, hierarchical and heterarchical, while the interactions between the components are local and modular for both types.

The fused measurements are integrated over discretized time with Bayesian filtering similar to Kalman filters (Christensen and Hager 2008; Durrant-Whyte and Henderson 2008). The difference between our model and Kalman Filters is with respect to the prediction step. We do not use a linearized movement prediction but instead a prediction based on transformations of body states, which exploits the modularity of the MMF architecture. These transformations are part of an explicit, full body schema (as termed in Hoffmann et al. (2010)). This body schema covers a postural schema and is also close to covering a surface schema. Both short-term and long-term dynamics are covered, and as far as the body image is concerned, a structural or topological representation (Hoffmann et al. 2010) is covered.

Apart from Kalman Filters, other Bayesian approaches such as Particle Filters have been used. In e.g., Gadeyne et al. (2005), the limits of Kalman Filters and advantages of Particle Filters with respect to a state and contact-estimation task are shown. Particle filters may very well be used to approximate the potential distributions in the respective modules of MMF.

7.2 Related models

Many related modular and multi-sensor fusion models have been developed. One example is a recurrent neural network (RNN) based on the MMC principle as applied to forward and inverse kinematics by Cruse and Steinkühler (1993) and Schilling (2011). The MMC network is a self-organizing map where geometrical constraints (which define accessible states) make these accessible states become attractors for the dynamic network. As states in different modalities and frames of reference can be represented in the same network, these dynamics can ensure consistency between them. This can be used, for example, if the state in one modality and frame of reference is externally forced to take on a new (desired) value. The attractor landscape then ensures that (after a transition) all other states match the new value. Thus,

the RNN can be used as an inverse model, a forward model, and for sensory integration.

An elegant example of a modular integration of two mappings applied to a robot is provided in Chinellato et al. (2011). A robot is able to execute eye and arm movements. For both movement types, a separate implicit map is learned unsupervisedly via coordinated looking and reaching movements. Both learned maps can be combined to "form a unique, shared visuomotor map of the peripersonal space." Thus, the three underlying representations or modules (i.e., visual, oculomotor, and arm joint space) are connected.

Another example where two learned implicit maps are connected is the SURE_REACH model (Herbort et al. 2010). SURE_REACH unsupervisedly learns the forward and inverse kinematics mapping (bidirectionally) and encodes the full state and potential motion trajectory redundantly. With excitatory and inhibitory activations continuously propagating through the network, SURE_REACH can adaptively execute goal-directed movements and avoid obstacles. The representations underlying SURE_REACH (target space and joint space) are close to the representations proposed by our model, but they are less modular and thus require a much higher computational cost.

More sensor-related, multimodal cues of the same stimuli have been integrated in a robot to improve the estimation capabilities (Chinellato et al. 2012). Visual measurements were also compared to an internal visual model to improve robot arm state estimation and thus visual servoing (Gratal et al. 2011).

Thus computational approaches exist that are modular with respect to different modalities or different limbs, models that exchange information *adaptively*, that fuse and integrate sensors with rigorous Bayesian integration principles, or models that are able to autonomously detect sensor failure. However, we are not aware of any model that comprises all these characteristics *at once* in a biologically plausible way.

8 Discussion

Inspired by the current knowledge about the human brain and human behavioral control, this paper has proposed a modular modality frame (MMF) architecture. MMF encodes an internal body model in a highly distributed, probabilistic fashion. It maintains an internal state estimate of the sensed and controlled body in the form of local Gaussian distribution estimations. Locality is enforced by separating the standard robotics state representation of end-effector location space ("task space") and joint angle space into four modality frames (shoulder-centered locations, shoulder-respective orientations, orientations respective the next upper limb, and joint angles), each with separate state estimation modules for each arm limb. Local forward and inverse kinematic mappings realize the exchange of sensory and predictive state information sources between the correlated modules via Kalman filtering and probabilistic plausibility estimations. Essentially, each module maintains redundant, complementary body state representations over time (cf. Sect. 3.1), and fuses and integrates multisensory information (cf. Sect. 3.2). Moreover, due to the continuous forward and inverse state estimate exchanges (cf. Sect. 3.3), the modules strive for consistency. Plausibilities of sensor measurements are autonomously calculated and used to adaptively integrate only trustworthy information.

We showed that MMF is able to process the available sensory information highly effectively. Comparisons with non-interactive Kalman-Filters showed that the noisy sensors are well-integrated, and comparisons with a previous version of MMF with a different local information exchange approach showed that the reliabilities of the unfolding state estimations are not overestimated. Scenarios in which the reliability of sensory information changes unknowingly to the MMF have shown that the system is able to identify sensor failures, accounting for their expected (decreased) information contents accordingly. Thus, the system forms and maintains an accurate and consistent body state representation over time, including mean and covariance estimates. The currently still simple control tasks showed that available sensory information can be easily integrated and thus guide the controlled arm to a desired goal state reliably.

Due to these features, we propose that this model should be considered for applications in soft robotics, where sensory information is prone to be noisy and motor commands yield noisy behavior effects. Moreover, similar architectures may be applied for the maintenance of other distributed body state estimations in other bodily simulations, potentially also facing other types of sensors with non-Gaussian noise distributions. Due to the modular structure of our model, the chain of limbs can easily be extended to more limbs, such as a torso or head. The model's computational cost is of the first order with respect to the number of limbs (cf. Sect. 5.3). If the model is applied to a whole humanoid—with two arms, two limbs, and a head attached to the torso, these could be modeled as different chains. As these chains inevitably have at least one limb in common, a crosstalk, such as the one in Sect. 3.3, could ensure synchronization between these chains.

In addition to such straight-forward application potentials, we also believe there is much room for enhancing this work in several respects. The MMF model exchanges information effectively, but further probabilistic information processing may be incorporated in the future. The similarities of MMF with several models and principles of human motor control and the modular neural encodings found in the human brain should be explored further. Moreover, no learning is currently applied in the MMF, which could further enhance the versatility and also the flexible applicability of the architecture.

Finally, MMF offers its modular architecture to distributed planning techniques, in which various goals and constraints may be combined continuously online. In the following, we discuss MMF's relations and potentials in respect to these aspects.

8.1 Further information processing

At the moment, the local forward and inverse kinematics in MMF, which are calculated based on (4) and (5), are assumed to be fully accurate. However, bodies change over time and thus these mappings can never be fully reliable. In future work, we plan to also consider noise or systematic errors in the local forward and inverse kinematics transformations. This will give rise to a challenge, which we intend to meet by adapting the transformations over time.

Currently, MMF only includes influences between different modality frames and between limbs down the kinematic chain. For example, the upper arm state influences the lower arm and hand states. The inverse direction is equally important—if, for example, the hand is perceived at a position incompatible with the upper and lower arm, then all limb states should be adjusted. This seemingly difficult problem may be elegantly addressed with principles from the mentioned MMC system (Cruse and Steinkühler 1993; Schilling 2011). Alternatively, neural population codes may be used, as detailed below.

In Sect. 6.2 and Fig. 15 (dashed lines), a set consisting of filters that act separately in each module was introduced. It served only to generate reference data, but an integration of this second set into MMF could allow MMF to profit from the redundancy even more. Each module would have two filters, one individual sensor-based filter, and one interactive state-estimation filter, where only the interactive filter uses the mappings (4) and (5). Both types of filters may deviate noticeably from each other. Such deviations may yield good signals for detecting kinematic mapping errors, that is, errors in the forward and inverse mappings (4) and (5) across the modules. Adaptation processes of the body model may apply consequently to error detection. The estimation of the actual reliability of the respective mappings may also be inferred. Furthermore, to pre-filter the incoming sensory information, sensory plausibilities may be based on the estimates in the individual sensor-based filters in addition to the actual sensory information.

Finally, at the moment the plausibility determination compares only the redundant sensor readings. These readings could also be compared with the internally predicted state estimates. The current method should be contrasted with one that also considers the internal filtered states in the plausibility determination.

8.2 Neural encodings and human motor control

Clearly, the neo-cortex of the brain processes sensory information from different modalities initially separately. Soon, however, this sensory information is strongly influenced by top-down processes and it is combined in "multisensory" brain regions. The most prominent multisensory region is the parietal cortex. In this respect, Shadmehr and Krakauer have suggested that

> "A function of the parietal cortex is state estimation: to integrate the predicted proprioceptive and visual outcomes with sensory feedback to form a belief about how the commands affected the states of the body and the environment." (Shadmehr and Krakauer 2008, p. 359)

The MMF is essentially a body state estimation system that integrates "proprioceptive" (angles and local limb orientations) and "visual" (location and global, head- or in our simulation shoulder-centered orientations) sensory feedback with the respective predicted outcomes. State estimations in the parietal cortex also exhibit different frames of reference, distinguishing, for example, near- from far-space, or peripersonal spaces that surround particular limbs or the face (Andersen et al. 1997). Thus, limb-respective spatial differentiations have been identified. In the anterior part, the somatosensory cortex clearly distinguishes body parts, processing the respectively available sensory information in a body homunculus-like distribution. These neuro-scientific evidences support the neural validity of our highly modularized architecture, although the brain is likely to process the information more synergistically dependent on motor synergies, which are determined by the body morphology (Latash 2008).

At the moment, though, the MMF model does not implement any neural codes but approximates the likely existence of probabilistic state estimations with Gaussian distributions. The brain appears to approximate probabilistic state estimates with neural population codes (Doya et al. 2007; Pouget et al. 2003). A previous model of human arm reaching has used population codes in the form of rather simple grids to approximate joint and task space (Butz et al. 2007; Herbort et al. 2010). This so-called SURE_REACH model showed many similarities to human arm reaching behavior and yielded particularly highly dexterous arm control capabilities. However, the employed neural encoding does not scale to a humanoid arm with nine degrees of freedom. The modularizations in the MMF architecture circumvent the scalability problem by separating the state representations of the individual limbs and joints. Each modality frame in MMF encodes maximally three degrees of freedom. Thus, neural

population codes can be embedded in these spaces with a reasonable fine granularity.

Once neural population codes are realized and bidirectional, mappings between the population codes of neighboring modality frames are available, information exchange in the inverse direction (from the hand to lower and upper arm limb) will be possible. As discussed above, since the inverse direction is usually not unique, it is very hard to implement it rigorously mathematically. Interestingly, the information exchange up the kinematic chain will allow to account for bodily illusions, such as the rubber hand illusion. Seeing that the MMF model strives for consistency while incorporating sensors with varying sensory reliabilities and respective plausibility estimates, the MMF actually predicts that the illusion will not only affect the stimulated body parts but also those other body parts that need to change their orientation, etc., to maintain an overall consistent body model. To the best of our knowledge, this hypothesis still has to be verified.

8.3 Planning and motor control

Another feature of the SURE_REACH model was its flexible capability of applying a general planning mechanism via model-based reinforcement learning, while flexibly inducing goals and various constraints (including obstacles and movement priorities). Consequently, highly dexterous, goal-oriented behavior was generated on the fly. The recent neuro-scientific literature points out that reach planning is likely to occur concurrently in multiple frames of reference (McGuire and Sabes 2009). Once MMF is implemented with population encoded state representations in each modality frame and neural mappings between the modality frames, we believe that distributed planning similar to the one applied in SURE_REACH will be possible. It will be possible to flexibly take task constraints and behavioral priorities into account. Due to the interactions between the modules, goal-directed behavior should unfold even when goals and constraints are only partially defined and are distributed over various modules. Comparisons with human behavioral data as well as neural behavior data, such as the comparison presented in McGuire and Sabes (2009), will then be imminent.

Due to the expected behavioral dexterity achievable with the resulting neural MMF, the application of the system to control soft robots will then need to be investigated even further. If a dynamic representation is desired that provides stable states, local reinforcement, or inhibition between nodes, the population codes may also be implemented by dynamic neural fields (Erlhagen and Schöner 2002). Recently, goal-directed, redundancy-resolving control has been realized with these techniques (Reimann et al. 2010, 2011). The MMF architecture is even more modularized.

8.4 Learning

The proposed neural population codes may be pre-structured by hand or they may be learned. Due to the potential versatility in the resulting system when the structures are learned from scratch, we are currently implementing neural population codes and neural regression techniques for learning the MMF spatial modules and their respective mappings. Various types of self-organizing maps (Butz et al. 2010; Kohonen 2001; Ritter et al 1992; Toussaint 2006) and neural regression techniques (Schaal and Atkeson 1998; Stalph and Butz 2011; Vijayakumar et al. 2005) may be used for this purpose. As mentioned before, the redundancy of the system can provide measures for the accuracies of the respective mappings. The plausibilities can then be used to adapt the learning rates or to weigh the interactions between the modules during sensory fusion and crosstalk online.

Beyond the learning of the current MMF architecture, additional environmental aspects may also be learned and represented in MMF. The parietal cortex was shown to also represent relative locations and orientations of things in the surrounding environment. These codes are usually particularly relevant for determining expected action consequences in consideration of the environmental context (Shadmehr and Krakauer 2008). As proposed already in the reafference principle (von Holst and Mittelstaedt 1950), a comparison of the predicted effector's reaction (6–7) to the actual sensory consequences of the movement can provide further information about the accuracy of the movement and about other environmental influences (such as external forces). It may even be used to estimate sensor variance online. Thus, as a next step, the system may detect external forces that act upon the system by contrasting the predicted sensory states with the actual experienced ones.

8.5 Conclusion

In conclusion, we believe that the MMF architecture constitutes only the beginning of modeling the brain's approach to human arm control. At the moment, the flexible incorporation of sensory information based on internally generated probabilistic estimates of their plausibilities models the apparent flexible Bayesian integration of redundant sources of information in the human brain. Furthermore, the Bayesian-based information exchange offers itself for the modeling of bodily illusions, such as the rubber hand illusion. Moreover, MMF is applicable for controlling and estimating the state of robots—particularly of robots with potentially partially and temporarily unreliable sensors and actuators. For potential future work, further information processing should be incor-

porated; the comparison with human behavioral data and neural data should be pursued; the modularity of MMF should be exploited to generate even more flexible, goal-directed planning and motor control mechanisms; and finally, the challenge of learning neural representations of the state estimation modules and the mappings between them will enable the even more general applicability of the MMF model—even to body morphologies for which the actual architecture is unknown beforehand.

Acknowledgments The authors thank the Cognitive Modeling team. Funding from the Emmy Noether program (German Research Foundation, DFG, BU1335/3-1) is acknowledged.

A Kinematic transformation derivation details

A.1 Location ↔ global orientation

To the global orientation state $\mathbf{x}^{\mathrm{GP}}$, the Inverse mapping can only provide complete information if the joint has maximally two DOFs and and the first rotation axis is parallel to the arm limb. This case will be considered here exemplarily. The local perpendicular $\mathbf{x}^{\mathrm{LP}}$ can be chosen along the rotation axis of the second DOF, e.g., $\mathbf{Z}$ if angles α and γ are flexible. For $\beta \equiv 0$, Fig. 19 illustrates the following dependency: The local perpendicular $\mathbf{x}^{\mathrm{LP}} = \mathbf{p}$ is perpendicular to the local direction $\mathbf{x}^{\mathrm{LD}}$ and to the upper limb's local direction $\mathbf{x}_u^{\mathrm{LD}} \equiv \mathbf{x}_{\mathrm{old}}^{\mathrm{LD}}$. The same holds true for the global states yielding:

$$\mathbf{x}^{\mathrm{GP}} = \pm\mathbf{u}\left(\mathbf{x}^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GD}}\right) \tag{A.1}$$

With (A.20), the Jacobians with respect to $\mathbf{x}_u^{\mathrm{GD}}$ and $\mathbf{x}^{\mathrm{GD}}$ are:

$$\frac{\partial \mathbf{x}^{\mathrm{GP}}}{\partial \mathbf{x}^{\mathrm{GD}}} = \frac{\pm\,\varepsilon \cdot \mathbf{x}_u^{\mathrm{GD}}}{\left|\mathbf{x}^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GD}}\right|} \pm\mathbf{u}\left(\mathbf{x}^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GD}}\right) \otimes \frac{\mathbf{u}\left(\mathbf{x}^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GD}}\right) \times \mathbf{x}_u^{\mathrm{GD}}}{\left|\mathbf{x}^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GD}}\right|} \tag{A.2}$$

$$\frac{\partial \mathbf{x}^{\mathrm{GP}}}{\partial \mathbf{x}_u^{\mathrm{GD}}} = \frac{\mp\,\varepsilon \cdot \mathbf{x}^{\mathrm{GD}}}{\left|\mathbf{x}^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GD}}\right|} \mp\mathbf{u}\left(\mathbf{x}^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GD}}\right) \otimes \frac{\mathbf{u}\left(\mathbf{x}^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GD}}\right) \times \mathbf{x}^{\mathrm{GD}}}{\left|\mathbf{x}^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GD}}\right|} \tag{A.3}$$

This transformation is displayed as the dashed arrows in Fig. 4.

A.2 Global orientation ↔ local orientation

The first row of the Jacobian of $\mathbf{x}^{\mathrm{LD}}$ with respect to $\mathbf{x}_u^{\mathrm{GD}}$ is given by

$$\frac{\partial\left(\left(\mathbf{x}^{\mathrm{LD}}\right)_1\right)}{\partial \mathbf{x}_u^{\mathrm{GD}}} = \left(\nabla_{\mathbf{x}_u^{\mathrm{GD}}}\left(\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right) \cdot \mathbf{x}^{\mathrm{GD}}\right)\right)^T. \tag{A.4}$$

Equations (30), (34) and

$$\mathrm{rot}(U\mathbf{V}) = U\,\mathrm{rot}\mathbf{V} + \mathrm{grad}U \times \mathbf{V} \tag{A.5}$$

lead to

$$\begin{aligned}\nabla_{\mathbf{V}_1}\left(\mathbf{u}\left(\mathbf{V}_1\right) \cdot \mathbf{V}_2\right) &= \left(\mathrm{grad}\,\mathbf{u}\left(\mathbf{V}_1\right)\right) \cdot \mathbf{V}_2 + \mathbf{V}_2 \times \mathrm{rot}\,\mathbf{u}\left(\mathbf{V}_1\right) \\ &= \frac{1}{|\mathbf{V}_1|}\left(\mathbf{I} - \mathbf{u}\left(\mathbf{V}_1\right) \otimes \mathbf{u}\left(\mathbf{V}_1\right)\right) \cdot \mathbf{V}_2 \\ &= \frac{\mathbf{V}_2 - \left(\mathbf{u}\left(\mathbf{V}_1\right) \cdot \mathbf{V}_2\right) \cdot \mathbf{u}\left(\mathbf{V}_1\right)}{|\mathbf{V}_1|}\end{aligned} \tag{A.6}$$

and this yields:

$$\frac{\partial\left(\left(\mathbf{x}^{\mathrm{LD}}\right)_1\right)}{\partial \mathbf{x}_u^{\mathrm{GD}}} = \left(\frac{\mathbf{x}^{\mathrm{GD}} - \left(\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right) \cdot \mathbf{x}^{\mathrm{GD}}\right) \cdot \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right)}{\left|\mathbf{x}_u^{\mathrm{GD}}\right|}\right)^T \tag{A.7}$$

The second row is

$$\frac{\partial\left(\left(\mathbf{x}^{\mathrm{LD}}\right)_2\right)}{\partial \mathbf{x}_u^{\mathrm{GD}}} = 0 \tag{A.8}$$

To further readability, we introduce $\mathbf{a} = \mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}$ and with Eq. (30) the third row of the Jacobian is

$$\frac{\partial\left(\left(\mathbf{x}^{\mathrm{LD}}\right)_3\right)}{\partial \mathbf{x}_u^{\mathrm{GD}}} = \left(\left(\nabla\mathbf{u}\left(\mathbf{a}\right)\right) \cdot \mathbf{x}^{\mathrm{GD}} + \mathbf{x}^{\mathrm{GD}} \times \mathrm{rot}\,\mathbf{u}\left(\mathbf{a}\right)\right)^T \tag{A.9}$$

The fist term can be rearranged with (32) and the second term with (A.5)

$$\begin{aligned}\frac{\partial\left(\left(\mathbf{x}^{\mathrm{LD}}\right)_3\right)}{\partial \mathbf{x}_u^{\mathrm{GD}}} &= \left(\frac{\nabla\mathbf{a}}{|\mathbf{a}|} \cdot \mathbf{x}^{\mathrm{GD}} + \mathbf{a}\left(\nabla\frac{1}{|\mathbf{a}|} \cdot \mathbf{x}^{\mathrm{GD}}\right)\right. \\ &\quad \left. + \mathbf{x}^{\mathrm{GD}} \times \left(\frac{\mathrm{rot}\,\mathbf{a}}{|\mathbf{a}|} + \nabla\frac{1}{|\mathbf{a}|} \times \mathbf{a}\right)\right)^T\end{aligned} \tag{A.10}$$

With Lagrange's formula for the vector triple product

$$\mathbf{x}^{\mathrm{GD}} \times \left(\nabla\frac{1}{|\mathbf{a}|} \times (\mathbf{a})\right) = \left(\nabla\frac{1}{|\mathbf{a}|}\right)\left(\mathbf{x}^{\mathrm{GD}} \cdot \mathbf{a}\right) - \mathbf{a}(\mathbf{x}^{\mathrm{GD}} \cdot \nabla\tfrac{1}{|\mathbf{a}|}) \tag{A.11}$$

the third row of the Jacobian yields:

$$\begin{aligned}\frac{\partial\left(\left(\mathbf{x}^{\mathrm{LD}}\right)_3\right)}{\partial \mathbf{x}_u^{\mathrm{GD}}} &= \left(\frac{\nabla\mathbf{a}}{|\mathbf{a}|} \cdot \mathbf{x}^{\mathrm{GD}} + \mathbf{x}^{\mathrm{GD}} \times \frac{\mathrm{rot}\,\mathbf{a}}{|\mathbf{a}|}\right. \\ &\quad \left. + \left(\nabla\frac{1}{|\mathbf{a}|}\right)\left(\mathbf{x}^{\mathrm{GD}} \cdot \mathbf{a}\right)\right)^T\end{aligned} \tag{A.12}$$

We introduce ε as the Levi-Civita-Tensor with rank 3, i.e. $(\varepsilon \cdot \mathbf{V})_{i,j} = \varepsilon_{ijk}V_k$ with ε_{ijk} as the Levi-Civita-Symbol.

Given

$$\begin{aligned}\mathrm{rot}\left(\mathbf{V}_1 \times \mathbf{V}_2\right) &= \left(\nabla \mathbf{V}_1\right) \cdot \mathbf{V}_2 - \left(\nabla \mathbf{V}_2\right) \cdot \mathbf{V}_1 \\ &\quad + \mathbf{V}_1 \,\mathrm{div}\mathbf{V}_2 - \mathbf{V}_2 \,\mathrm{div}\mathbf{V}_1 \end{aligned} \tag{A.13}$$

we get

$$\mathrm{rot}_{\mathbf{V}_1}\left(\mathbf{V}_1 \times \mathbf{V}_2\right) = \mathbf{I} \cdot \mathbf{V}_2 - 3\mathbf{V}_2 = -2\mathbf{V}_2 \tag{A.14}$$

and with the identities

$$\nabla_{\mathbf{V}_1}\left(\mathbf{V}_1 \times \mathbf{V}_2\right) = \varepsilon \cdot \mathbf{V}_2 \tag{A.15}$$

$$\left(\varepsilon \cdot \mathbf{V}\right) \cdot \mathbf{V}_3 = \mathbf{V}_3 \times \mathbf{V} \tag{A.16}$$

and (31) and (32), the following simplification holds true:

$$\nabla_{\mathbf{V}_1} \frac{1}{\left|\mathbf{V}_1 \times \mathbf{V}_2\right|} = \frac{\left(\mathbf{V}_1 \times \mathbf{V}_2\right) \times \mathbf{V}_2}{\left|\left(\mathbf{V}_1 \times \mathbf{V}_2\right)\right|^3} \tag{A.17}$$

and with (A.14)–(A.17), the third row of the Jacobian in (A.12) can be further simplified:

$$\begin{aligned}\frac{\partial\left(\left(\mathbf{x}^{\mathrm{LD}}\right)_3\right)}{\partial \mathbf{x}_u^{\mathrm{GD}}} &= \left(\frac{\mathbf{x}^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}}{|\mathbf{a}|}\right. \\ &\quad \left. + \mathbf{x}^{\mathrm{GD}} \times \frac{-2\mathbf{x}_u^{\mathrm{GP}}}{|\mathbf{a}|} + \frac{\mathbf{a} \times \mathbf{x}_u^{\mathrm{GP}}}{|\mathbf{a}|^3}\left(\mathbf{x}^{\mathrm{GD}} \cdot \mathbf{a}\right)\right)^T \\ &= \left(\frac{\mathbf{x}_u^{\mathrm{GP}}}{|\mathbf{a}|} \times \left(\mathbf{x}^{\mathrm{GD}} - \mathbf{u}\left(\mathbf{a}\right)\left(\mathbf{u}\left(\mathbf{a}\right) \cdot \mathbf{x}^{\mathrm{GD}}\right)\right)\right)^T \end{aligned} \tag{A.18}$$

Combining (A.7), (A.8) and (A.18), the complete Jacobian with respect to $\mathbf{x}_u^{\mathrm{GD}}$ is

$$\frac{\partial\left(\mathbf{x}^{\mathrm{LD}}\right)}{\partial \mathbf{x}_u^{\mathrm{GD}}} = \begin{pmatrix} \frac{1}{\left|\mathbf{x}_u^{\mathrm{GD}}\right|}\left(\mathbf{x}^{\mathrm{GD}} - \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right)\left(\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right) \cdot \mathbf{x}^{\mathrm{GD}}\right)\right)^T \\ 0 \\ \frac{1}{|\mathbf{a}|}\left(\mathbf{x}_u^{\mathrm{GP}} \times \left(\mathbf{x}^{\mathrm{GD}} - \mathbf{u}\left(\mathbf{a}\right)\left(\mathbf{u}\left(\mathbf{a}\right) \cdot \mathbf{x}^{\mathrm{GD}}\right)\right)\right)^T \end{pmatrix} \tag{A.19}$$

To calculate the Jacobians for the forward mapping, we first derive the vector gradient of the unit vector of a cross-product. With (32), (A.15) and (A.17) we get

$$\begin{aligned}\nabla_{\mathbf{V}_1}\mathbf{u}\left(\mathbf{V}_1 \times \mathbf{V}_2\right) &= \frac{\varepsilon \cdot \mathbf{V}_2}{\left|\mathbf{V}_1 \times \mathbf{V}_2\right|} + \mathbf{u}\left(\mathbf{V}_1 \times \mathbf{V}_2\right) \\ &\quad \otimes \frac{\mathbf{u}\left(\mathbf{V}_1 \times \mathbf{V}_2\right) \times \mathbf{V}_2}{\left|\mathbf{V}_1 \times \mathbf{V}_2\right|} \end{aligned} \tag{A.20}$$

With (34) and (A.20), the Jacobian with respect to $\mathbf{x}_u^{\mathrm{GD}}$ yields

$$\begin{aligned}\frac{\partial \mathbf{x}^{\mathrm{GD}}}{\partial \mathbf{x}_u^{\mathrm{GD}}} &= \frac{\mathbf{I} - \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right) \otimes \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}}\right)}{\left|\mathbf{x}_u^{\mathrm{GD}}\right|}\left(\mathbf{x}^{\mathrm{LD}}\right)_1 \\ &\quad + \left(\frac{\varepsilon \cdot \mathbf{x}_u^{\mathrm{GP}}}{\left|\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right|} + \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right)\right. \\ &\quad \left. \otimes \frac{\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right) \times \mathbf{x}_u^{\mathrm{GP}}}{\left|\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right|}\right)\left(\mathbf{x}^{\mathrm{LD}}\right)_3 \end{aligned} \tag{A.21}$$

and the Jacobian with respect to $\mathbf{x}_u^{\mathrm{GP}}$:

$$\begin{aligned}\frac{\partial \mathbf{x}^{\mathrm{GD}}}{\partial \mathbf{x}_u^{\mathrm{GP}}} &= \frac{\mathbf{I} - \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}}\right) \otimes \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GP}}\right)}{\left|\mathbf{x}_u^{\mathrm{GP}}\right|}\left(\mathbf{x}^{\mathrm{LD}}\right)_1 - \left(\frac{\varepsilon \cdot \mathbf{x}_u^{\mathrm{GD}}}{\left|\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right|}\right. \\ &\quad - \mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right) \\ &\quad \left. \otimes \frac{\mathbf{u}\left(\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right) \times \mathbf{x}_u^{\mathrm{GD}}}{\left|\mathbf{x}_u^{\mathrm{GD}} \times \mathbf{x}_u^{\mathrm{GP}}\right|}\right)\left(\mathbf{x}^{\mathrm{LD}}\right)_3 \end{aligned} \tag{A.22}$$

A.3 Local orientation ↔ joint angles

We demonstrate the transformation between angles and a body-relative Cartesian coordinate system exemplarily for a 3D angular space. XYZ is introduced as the coordinate system, related to the local orientation of the *upper* arm: $\mathbf{X} \equiv \mathbf{x}_u^{\mathrm{LD}}$ and $\mathbf{Z} \equiv \mathbf{x}_u^{\mathrm{LP}}$ (cf. Fig. 19).

We define the angles $\mathbf{x}^{\mathrm{LA}} = \left(\alpha\ \beta\ \gamma\right)^T$ as Tait-Bryan angles using intrinsic $XY'Z''$ notation. This means, the first rotation is by the angle α around the x-axis $\mathbf{X}$. As the rotation axes are intrinsic (not static), they are rotated: the one (twice, thrice) rotated coordinate system is denoted by one (two, three) dashes. Thus the second rotation is by the angle β around $\mathbf{Y}'$ and the third rotation is by the angle γ around $\mathbf{Z}''$.

Thus, any arbitrary vector $\mathbf{v}$ is rotated by a product of rotation matrices

$$\mathbf{v}''' = R_{\mathbf{Z}''}\left(\gamma\right) R_{\mathbf{Y}'}\left(\beta\right) R_{\mathbf{X}}\left(\alpha\right) \mathbf{v} \tag{A.23}$$

This series of rotations around intrinsic axes is equivalent to the reversed order series of rotations around static axes

$$\mathbf{v}''' = R_{\mathbf{X}}\left(\alpha\right) R_{\mathbf{Y}}\left(\beta\right) R_{\mathbf{Z}}\left(\gamma\right) \mathbf{v} \tag{A.24}$$

$$\begin{aligned}\mathbf{v}''' &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & \cos\alpha & -\sin\alpha \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \\ &\quad \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}\mathbf{v} \end{aligned} \tag{A.25}$$

Rotating the $\left(1\ 0\ 0\right)^T$-vector by this series of rotations yields the **local direction** of the current arm $\mathbf{x}^{\mathrm{LD}}$ (cf. Fig. 19):

$$\mathbf{x}^{\mathrm{LD}}\left(\mathbf{x}^{\mathrm{LA}}\right) = \begin{pmatrix} \cos\beta\cos\gamma \\ \cos\gamma\sin\alpha\sin\beta + \cos\alpha\sin\gamma \\ -\cos\alpha\cos\gamma\sin\beta + \sin\alpha\sin\gamma \end{pmatrix} \tag{A.26}$$

Similarly, rotating $\left(001\right)^T$ yields the **local perpendicular**

$$\mathbf{x}^{\mathrm{LP}}\left(\mathbf{x}^{\mathrm{LA}}\right) = \begin{pmatrix} \sin\beta \\ -\cos\beta\sin\alpha \\ \cos\alpha\cos\beta \end{pmatrix} \tag{A.27}$$

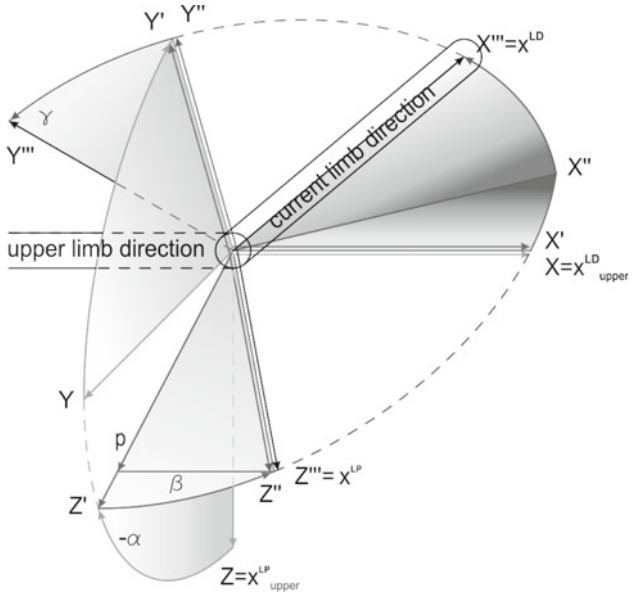


Fig. 19 Transformation from local orientation to local angles and vice versa. Euler angles are used, i.e. rotation axes are not fixed. The *upper limb* (direction and perpendicular) defines $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$, i.e. the coordinate system before any rotation. $\mathbf{X}$, $\mathbf{Y}'$, $\mathbf{Z}''$ serve as rotation axes and the coordinate system after all three rotations is $\mathbf{X}'''$, $\mathbf{Y}'''$, $\mathbf{Z}'''$. $\mathbf{p}$ is the result of the projection of $\mathbf{x}^{\mathrm{LP}}$ on the *YZ*-plane

The **Jacobian** for the local direction is

$$\frac{\partial \mathbf{x}^{\mathrm{LD}}}{\partial \alpha} = \begin{pmatrix} 0 \\ -\sin(\gamma)\sin(\alpha) + \cos(\alpha)\sin(\beta)\cos(\gamma) \\ \cos(\alpha)\sin(\gamma) + \sin(\alpha)\sin(\beta)\cos(\gamma) \end{pmatrix}$$

$$\frac{\partial \mathbf{x}^{\mathrm{LD}}}{\partial \beta} = \begin{pmatrix} -\sin(\beta)\cos(\gamma) \\ \cos(\beta)\sin(\alpha)\cos(\gamma) \\ -\cos(\alpha)\cos(\beta)\cos(\gamma) \end{pmatrix}$$

$$\frac{\partial \mathbf{x}^{\mathrm{LD}}}{\partial \gamma} = \begin{pmatrix} -\cos(\beta)\sin(\gamma) \\ -\sin(\gamma)\sin(\alpha)\sin(\beta) + \cos(\alpha)\cos(\gamma) \\ \cos(\alpha)\sin(\gamma)\sin(\beta) + \sin(\alpha)\cos(\gamma) \end{pmatrix}$$

$$\frac{\partial \mathbf{x}^{\mathrm{LD}}}{\partial \mathbf{x}^{\mathrm{LA}}} = \begin{pmatrix} \frac{\partial \mathbf{x}^{\mathrm{LD}}}{\partial \alpha} & \frac{\partial \mathbf{x}^{\mathrm{LD}}}{\partial \beta} & \frac{\partial \mathbf{x}^{\mathrm{LD}}}{\partial \gamma} \end{pmatrix} \tag{A.28}$$

and the **Jacobian** for the local perpendicular

$$\frac{\partial^{\mathrm{LP}}_{\mathbf{x}}}{\partial^{\mathrm{LA}}_{\mathbf{x}}} = \begin{pmatrix} 0 & \cos(\beta) & 0 \\ -\cos(\beta)\cos(\alpha) & \sin(\beta)\sin(\alpha) & 0 \\ -\cos(\beta)\sin(\alpha) & -\sin(\beta)\cos(\alpha) & 0 \end{pmatrix}. \tag{A.29}$$

The following transition

$$(\alpha, \beta, \gamma) \mapsto (\pi + \alpha, \pi - \beta, \pi + \gamma). \tag{A.30}$$

changes neither the resulting local direction, nor the local perpendicular, even if applied n times in succession. This fact can also be inferred from Fig. 19. As a consequence, the inverse mapping $\mathbf{x}^{\mathrm{LO}} \rightarrow \mathbf{x}^{\mathrm{LA}}$ has multiple solutions—if no multiples of 2π are allowed, two solutions remain, i.e. $n \in \{0, 1\}$.

For convenience, we introduce the result $\mathbf{p}$ of the projection of the local perpendicular $\mathbf{x}^{\mathrm{LP}}$ on the *YZ*-plane (cf. Fig. 19)

$$\mathbf{p} = |\cos\beta| \begin{pmatrix} 0 \\ -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}. \tag{A.31}$$

Next, the **inverse mapping** is derived. To identify the **first angle** α, the two-argument arctangent function as apparent in Fig. 19 is used

$$\alpha = n\pi - \operatorname{atan2}(\mathbf{u}_1(\mathbf{p}), \mathbf{u}_2(\mathbf{p})), \tag{A.32}$$

where $\mathbf{u}_i(\mathbf{p})$ is the i-th dimensional element of the unit vector of $\mathbf{p}$ and n denotes how often (A.30) is applied. To solve for **the second angle** β, we multiply (A.27) by $\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^T$, yielding

$$\beta = n\pi + (-1)^n \arcsin\left(\mathbf{u}_1\left(\mathbf{x}^{\mathrm{LP}}\right)\right). \tag{A.33}$$

The summand $n\pi$ and factor $(-1)^n$ extend the range of arcsin to the whole domain of sin and are in accordance with (A.30).

To solve for the **third angle** γ, we multiply (A.27) with $\mathbf{X}$

$$\cos\gamma = \frac{\mathbf{u}_1\left(\mathbf{x}^{\mathrm{LD}}\right)}{\cos\beta}$$

and with

$$\cos\beta = (-1)^n |\cos\beta| \tag{A.34}$$

(cf. Fig. 19), the inverse mapping for γ yields

$$\gamma = n\pi + \operatorname{sign}\left(\mathbf{X}''' \cdot \mathbf{Y}''\right) \arccos \frac{\mathbf{u}_1\left(\mathbf{x}^{\mathrm{LD}}\right)}{|\cos\beta|}$$

$$\gamma = n\pi + \operatorname{sign}\left(\mathbf{X}''' \cdot \mathbf{Y}''\right) \arccos \frac{\mathbf{u}_1\left(\mathbf{x}^{\mathrm{LD}}\right)}{\sqrt{1 - \left(\mathbf{u}_1\left(\mathbf{x}^{\mathrm{LP}}\right)\right)^2}} \tag{A.35}$$

(cf. Fig. 19). Note that $\mathbf{Y}''$ changes it's sign for $n \mapsto n + 1$.

To obtain the **Jacobian** of the first angle α with respect to the local direction $\mathbf{x}^{\mathrm{LD}}$, (A.27) is multiplied by $\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T$ which yields an expression *alternate* to (A.32)

$$\alpha = \pm \arccos \frac{\mathbf{u}_3\left(\mathbf{x}^{\mathrm{LP}}\right)}{\cos\beta} \tag{A.36}$$

The sign extends the range of arccos and can be specified (cf. Fig. 19) yielding

$$\alpha = (-1)^{n+1} \operatorname{sign}(\mathbf{u}_2(\mathbf{p})) \arccos \frac{\mathbf{u}_3\left(\mathbf{x}^{\mathrm{LP}}\right)}{\cos\beta} \tag{A.37}$$

and with $\mathbf{u}_3\left(\mathbf{x}^{\mathrm{LP}}\right) = \mathbf{p}_3 = \mathbf{u}_3(\mathbf{p})|\cos\beta|$ (cf. Fig. 19) and (A.34), α can be expressed as

$$\alpha = n\pi - \operatorname{sign}(\mathbf{u}_2(\mathbf{p})) \arccos(\mathbf{u}_3(\mathbf{p})), \tag{A.38}$$

which is in accordance with (A.30). With (34), the Jacobian with respect to $\mathbf{p}$ can be derived, and because of $\nabla_{\mathbf{x}}^{\mathrm{LP}}\alpha = \nabla_{\mathbf{p}}\alpha$, the resulting Jacobian with respect to $\mathbf{x}^{\mathrm{LP}}$ is the same

$$\nabla_{\mathbf{x}^{\mathrm{LP}}}\alpha = n\pi + \frac{\operatorname{sign}\left(\mathbf{u}_2\left(\mathbf{p}\right)\right)\cdot\left(\mathbf{Z}-\mathbf{u}_3\left(\mathbf{p}\right)\mathbf{u}\left(\mathbf{p}\right)\right)}{|\mathbf{p}|\sqrt{1-\left(\mathbf{u}_3\left(\mathbf{p}\right)\right)^2}}. \tag{A.39}$$

The equation for the second angle β (A.33) is similar to (A.38). Thus, analogous to (A.39), the **Jacobian** of the second angle β follows as:

$$\nabla_{\mathbf{x}^{\mathrm{LP}}}\beta = \frac{(-1)^n\left(\mathbf{X}-\mathbf{u}_1\left(\mathbf{x}^{\mathrm{LD}}\right)\mathbf{u}\left(\mathbf{x}^{\mathrm{LD}}\right)\right)}{\left|\mathbf{x}^{\mathrm{LP}}\right|\sqrt{1-\left(\mathbf{u}_1\left(\mathbf{x}^{\mathrm{LD}}\right)\right)^2}}, \tag{A.40}$$

The third angle can not be rearranged to gain an expression which solely depends on either the local direction *or* the local perpendicular. Thus, the **Jacobians** of the third angle with respect to both vectors have to be derived. The first is given by

$$\frac{\partial\gamma}{\partial\mathbf{x}^{\mathrm{LD}}} = -\operatorname{sign}\left(\mathbf{X}'''\cdot\mathbf{Y}''\right)\left(1-\left(\frac{\mathbf{u}(\mathbf{x}^{\mathrm{LD}})\cdot\mathbf{X}}{\sqrt{1-(\mathbf{u}(\mathbf{x}^{\mathrm{LP}})\cdot\mathbf{X})^2}}\right)^2\right)^{-\frac{1}{2}} \cdot\left(1-\left(\mathbf{u}\left(\mathbf{x}^{\mathrm{LP}}\right)\cdot\mathbf{X}\right)^2\right)^{-\frac{1}{2}}\cdot\left(\nabla_{\mathbf{x}^{\mathrm{LD}}}\left(\mathbf{u}\left(\mathbf{x}^{\mathrm{LD}}\right)\cdot\mathbf{X}\right)\right)^T$$

and with (A.6) this yields

$$\frac{\partial\gamma}{\partial\mathbf{x}^{\mathrm{LD}}} = \frac{-\operatorname{sign}\left(\mathbf{X}'''\cdot\mathbf{Y}''\right)\left(\mathbf{X}-\left(\mathbf{u}\left(\mathbf{x}^{\mathrm{LD}}\right)\cdot\mathbf{X}\right)\cdot\mathbf{u}\left(\mathbf{x}^{\mathrm{LD}}\right)\right)^T}{\left|\mathbf{x}^{\mathrm{LD}}\right|\sqrt{1-\left(\mathbf{u}\left(\mathbf{x}^{\mathrm{LP}}\right)\cdot\mathbf{X}\right)^2-\left(\mathbf{u}\left(\mathbf{x}^{\mathrm{LD}}\right)\cdot\mathbf{X}\right)^2}} \tag{A.41}$$

The second Jacobian is given by

$$\frac{\partial\gamma}{\partial\mathbf{x}^{\mathrm{LP}}} = -\operatorname{sign}\left(\mathbf{X}'''\cdot\mathbf{Y}''\right)\left(1-\left(\frac{\mathbf{u}(\mathbf{x}^{\mathrm{LD}})\cdot\mathbf{X}}{\sqrt{1-(\mathbf{u}(\mathbf{x}^{\mathrm{LP}})\cdot\mathbf{X})^2}}\right)^2\right)^{-\frac{1}{2}} \left(\mathbf{u}\left(\mathbf{x}^{\mathrm{LD}}\right)\cdot\mathbf{X}\right)\cdot\left(-\tfrac{1}{2}\right)\left(1-\left(\mathbf{u}\left(\mathbf{x}^{\mathrm{LP}}\right)\cdot\mathbf{X}\right)^2\right)^{-\frac{3}{2}} \cdot\left(\nabla_{\mathbf{x}^{\mathrm{LP}}}\left(-\left(\mathbf{u}\left(\mathbf{x}^{\mathrm{LP}}\right)\cdot\mathbf{X}\right)^2\right)\right)^T$$

and with (A.6) this yields

$$\frac{\partial\gamma}{\partial\mathbf{x}^{\mathrm{LP}}} = \frac{-\operatorname{sign}(\mathbf{X}'''\cdot\mathbf{Y}'')\cdot(\mathbf{X}-(\mathbf{u}(\mathbf{x}^{\mathrm{LP}})\cdot\mathbf{X})\cdot\mathbf{u}(\mathbf{x}^{\mathrm{LP}}))^T}{|\mathbf{x}^{\mathrm{LP}}|\sqrt{1-(\mathbf{u}(\mathbf{x}^{\mathrm{LP}})\cdot\mathbf{X})^2-(\mathbf{u}(\mathbf{x}^{\mathrm{LD}})\cdot\mathbf{X})^2}} \cdot\frac{(\mathbf{u}(\mathbf{x}^{\mathrm{LD}})\cdot\mathbf{X})\cdot(\mathbf{u}(\mathbf{x}^{\mathrm{LP}})\cdot\mathbf{X})}{1-(\mathbf{u}(\mathbf{x}^{\mathrm{LP}})\cdot\mathbf{X})^2} \tag{A.42}$$

B Noise parameter derivation

In order to study a setup where all sensors carry approximately equal information, the sensors' Gaussian noise standard deviations $\sigma_{\mathrm{Sensor}}^{j}$ were chosen to reflect the consequences of the transformations (4–5) for all modality frames j. For example, a standard deviation of the angular frame LA, i.e., $\sigma_{\mathrm{Sensor}}^{\mathrm{LA}}$, results in the following corresponding standard deviation in modality frame LO

$$\sigma_{\mathrm{Sensor}}^{\mathrm{LO}} = \sqrt[*]{\left(\frac{\partial\mathbf{x}^{\mathrm{LO}}}{\partial\mathbf{x}^{\mathrm{LA}}}\right)^{*2}\left(\sigma_{\mathrm{Sensor}}^{\mathrm{LA}}\right)^{*2}}, \tag{A.43}$$

(cf. (5)), where ^{*2} is the element-wise square, and $\sqrt[*]{}$ the element-wise square root.

As the sensor noise has to be independent of the arm position, an average of $\sigma_{\mathrm{Sensor}}^{\mathrm{LO}}$ over all arm positions $\mathbf{x}^{\mathrm{LA}}$ has to be estimated. As, however, not all arm positions are reached uniformly distributed, an accurate expression of the average cannot be derived. Instead, we simply set the "average" slightly lower than the maximum. The maximum over all dimensions i and arm states $\mathbf{x}^{\mathrm{LA}}$ is

$$\max_{i,\mathbf{x}^{\mathrm{LA}}}\left(\left(\sigma_{\mathrm{Sensor}}^{\mathrm{LO}}\right)_i\right) \leq \sqrt{2}\cdot\sigma_{\mathrm{Sensor}}^{\mathrm{LA}}/\mathrm{rad}, \tag{A.44}$$

(obtained from (A.29), (A.28) and (A.43)). Using this maximum as a guideline, we approximated the estimated average with

$$\left(\overline{\sigma_{\mathrm{Sensor}}^{\mathrm{LO}}}\right)_k :\approx \left(\sigma_{\mathrm{Sensor}}^{\mathrm{LA}}\right)_k/\mathrm{rad}. \tag{A.45}$$

Here, in all dimensions k, the dimensional elements $(\sigma_{\mathrm{Sensor}}^{j})_k$ were assigned equal values.

Due to different trajectories from random start positions to random goals, the approximation sometimes overestimates and sometimes underestimates the respective standard deviation. It is simply not possible to create even a fictional setup with fixed sensor noise, where each modality frame always contributes the same amount of information. Thus, the modality frames can have stronger or weaker impact on the overall state estimate. We chose the dependencies between all different modality frames j displayed in Table 2 to provide a good compromise.

References

Andersen RA, Snyder LH, Bradley DC, Xing J (1997) Multimodal representation of space in the posterior parietal cortex and its use in planning movements. Ann Rev Neurosci 20:303–330

Beauchamp MS (2005) See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. Curr Opin Neurobiol 15:145–153

Bernier PM, Gauthier GM, Blouin J (2007) Evidence for distinct, differentially adaptable sensorimotor transformations for reaches to visual and proprioceptive targets. J Neurophysiol 98:1815–1819

Botvinick M, Cohen J et al (1998) Rubber hands 'feel' touch that eyes see. Nature 391:756

Braitenberg V (1986) Vehicles: experiments in synthetic psychology. The MIT Press, Cambridge

Bronstein IN, Semendjajew KA, Musiol G, Mühlig H (eds) (2001) Taschenbuch der Mathematik, 5th edn. Harri Deutsch, Frankfurt

Brooks RA (1990) Elephants don't play chess. Robot Auton Syst 6: 3–15

Butz MV, Herbort O, Hoffmann J (2007) Exploiting redundancy for flexible behavior: unsupervised learning in a modular sensorimotor control architecture. Psychol Rev 114:1015–1046

Butz MV, Shirinov E, Reif KL (2010) Self-organizing sensorimotor maps plus internal motivations yield animal-like behavior. Adapt Behav 18:315–337

Calvert GA, Spence C, Stein BE (eds) (2004) The handbook of multisensory processes. The MIT Press, Cambridge

Chinellato E, Antonelli M, Grzyb BJ, del Pobil AP (2011) Implicit sensorimotor mapping of the peripersonal space by gazing and reaching. IEEE Trans Auton Mental Dev 3:43–53

Chinellato E, Grzyb BJ, del Pobil AP (2012) Pose estimation through cue integration: a neuroscience-inspired approach. IEEE Trans Syst Man Cybern Part B 42:530–538

Christensen HI, Hager GD (2008) Sensing and estimation. In: Siciliano B, Khatib O (eds) Springer handbook of robotics. Springer, Berlin, chap 4, pp 87–107

Cruse H, Steinkühler U (1993) Solution of the direct and inverse kinematic problems by a common algorithm based on the mean of multiple computations. Biol Cybern 69:341–351

de Vignemont F, Majid A, Jola C, Haggard P (2009) Segmenting the body into parts: evidence from biases in tactile perception. Q J Exp Psychol 62:500–512

Denève S, Duhamel JR, Pouget A (2007) Optimal sensorimotor integration in recurrent cortical networks: a neural implementation of Kalman filters. J Neurosci 27:5744–5756

Doya K, Ishii S, Pouget A, Rao RPN (2007) Bayesian brain: probabilistic approaches to neural coding. The MIT Press, Cambridge

Durrant-Whyte H, Henderson TC (2008) Multisensor data fusion. In: Siciliano B, Khatib O (eds) Springer handbook of robotics. Springer, Berlin, chap 25, pp 585–610

Ehrenfeld S, Butz MV (2011) A modular, redundant, multi-frame of reference representation for kinematic chains. In: IEEE International Conference on Robotics and Automation, pp 141–147

Erlhagen W, Schöner G (2002) Dynamic field theory of movement preparation. Psychol Rev 109:545–572

Gadeyne K, Lefebvre T, Bruyninckx H (2005) Bayesian hybrid model-state estimation applied to simultaneous contact formation recognition and geometrical parameter estimation. Int J Robot Res 24:615–630

Gentner R, Classen J (2006) Modular organization of finger movements by the human central nervous system. Neuron 52:731–742

Gratal X, Romero J, Kragic D (2011) Virtual visual servoing for real-time robot pose estimation. In: Bittanti S, Cenedese A, Zampieri S (eds) World congress, vol 18, pp 9017–9022

Herbort O, Butz MV, Pedersen G (2010) The SURE_REACH model for motor learning and control of a redundant arm: from modeling human behavior to applications in robotics. In: Sigaud O, Peters J (eds) From motor learning to interaction learning in robots. Springer, Berlin, pp 85–106

Hoffmann M, Marques H, Arieta A, Sumioka H, Lungarella M, Pfeifer R (2010) Body schema in robotics: a review. IEEE Trans Auton Mental Dev 2:304–324

Holst E , von Mittelstaedt H (1950) Das Reafferenzprinzip (Wechselwirkungen zwischen Zentralnervensystem und Peripherie). Naturwissenschaften 37:464–476

Kohonen T (2001) Self-organizing maps. 3rd ed. Springer, Berlin

Körding KP, Wolpert DM (2004) Bayesian integration in sensorimotor learning. Nature 427:244–247

Latash ML (2008) Synergy. Oxford University Press, New York

Latash ML, Turvey MT (eds) (1996) Dexterity and its development. Lawrence Erlbaum Assoc Inc, Mahwah

Latash ML, Scholz JP, Schöner G (2007) Toward a new theory of motor synergies. Motor Control 11:276–308

Longo MR, Haggard P (2010) An implicit body representation underlying human position sense. PNAS 107:11727–11732

Makin TR, Holmes NP, Ehrsson HH (2008) On the other hand: dummy hands and peripersonal space. Behav Brain Res 191:1–10

Maravita A, Spence C, Driver J (2003) Multisensory integration and the body schema: close to hand and within reach. Curr Biol 13:531–539

McGuire LM, Sabes PN (2009) Sensory transformations and the use of multiple reference frames for reach planning. Nat Neurosci 12:1056–1061

Nguyen-Tuong D, Peters J (2011) Model learning for robot control: a survey. Cogn Process 12:319–340

Pouget A, Dayan P, Zemel R (2003) Inference and computation with population codes. Ann Rev Neurosci 26:381–410

Reimann H, Iossifidis I, Schöner G (2010) Integrating orientation constraints into the attractor dynamics approach for autonomous manipulation. In: International conference on humanoid robots, pp 294–301

Reimann H, Iossifidis I, Schoner G (2011) Generating collision free reaching movements for redundant manipulators using dynamical systems. In: 2010 IEEE/RSJ international conference on intelligent robots and systems (IROS 2011), pp 5372–5379

Ritter H, Martinetz T, Schulten K (1992) Neural computation and self-organizing maps—an introduction. Addison-Wesley, New York

Rosenbaum DA (2010) Human motor control. 2nd ed. Academic Press, San Diego

Schaal S, Atkeson CG (1998) Constructive incremental learning from only local information. Neural Comput 10:2047–2084

Scheinmann V, McCarthy JM (2008) Mechanisms and actuation. In: Siciliano B, Khatib O (eds) Springer handbook of robotics. Springer, Berlin, chap 3, pp 67–86

Schilling M (2011) Universally manipulable body models—dual quaternion representations in layered and dynamic MMCs. Auton Robots 30:399–425

Schmitz J, Schneider A, Schilling M, Cruse H (2008) No need for a body model: positive velocity feedback for the control of an 18-DOF robot walker. Appl Bionics Biomech 5:135–147

Shadmehr R, Krakauer JW (2008) A computational neuroanatomy for motor control. Exp Brain Res 185:359–381

Shadmehr R, Wise SP (2005) The computational neurobiology of reaching and pointing: a foundation for motor learning. MIT Press, Cambridge

Sigaud O, Salaun C, Padois V (2011) On-line regression algorithms for learning mechanical models of robots: a survey. Robotics Auton Syst 59:1115–1129

Stalph PO, Butz MV (2011) Learning local linear Jacobians for flexible and adaptive robot arm control. Genetic Programm Evolvable Mach 13:137–157

Streri A, Pownall TT, Kingerlee ST (1993) Seeing, reaching, touching: the relations between vision and touch in infancy. The MIT Press, Cambridge

Tononi G, Edelman GM, Sporns O (1998) Complexity and coherency: integrating information in the brain. Trends Cogn Sci 2:474–484

Toussaint M (2006) A sensorimotor map: modulating lateral interactions for anticipation and planning. Neural Comput 18:1132–1155

Vijayakumar S, D'Souza A, Schaal S (2005) Incremental online learning in high dimensions. Neural Comput 17:2602–2634

Wells JP, Hyler-Both DL, Danley TD, Wallace GH (2002) Biomechanics of growth and development in the healthy human infant: a pilot study. J Am Osteopath Assoc 102:313–319

Wolpert DM, Kawato M (1998) Multiple paired forward and inverse models for motor control. Neural Netw 11:1317–1329