Overapproximative Arm Occupancy Prediction for Human-Robot Co-existence Built from Archetypal Movements

Aaron Pereira[1] and Matthias Althoff[1]

Abstract— Human motion is fast and hard to predict. To implement a provably safe collision-avoidance strategy for robots in collaborative spaces with humans, an overapproximative prediction of the occupancy of the human is required, which needs to be calculated faster than real time. We present a method for computing volumes containing the entire possible future occupancy of the human, given its state, faster than real time. The dynamic model of the human is built from analysing a set of archetypal movements performed by test subjects. The occupancy prediction is tested on a publicly available database of motion capture data, and shown to be overapproximative for all movements relating to everyday activities, sport and dance. Our novel algorithm is useful to guarantee safety in human-robot collaboration scenarios.

I. INTRODUCTION

As robots working near humans gain ground in industry, particularly in Small and Medium Enterprises (SMEs), prediction of the movement of nearby humans is paramount to ensuring safe interaction or avoidance. Current safety measures such as speed and separation monitoring, and cages [1], [2] are inflexible or do not guarantee safety. Therefore much research has focussed on predicting human motion to avoid collisions [3]. We present a novel concept to formally bound the spatial occupancy of the human arm based on sensor data in real time, using a model built from analysis of archetypal movements. In contrast to previous work on human motion prediction, this occupancy would cover all potential human occupancies for an uncertain model of human movement given sensor data, at any time up to a prediction horizon. Such real-time prediction may be used in dynamic online obstacle-avoiding trajectory planners e.g. [4].

Humans move fast and have different shapes and sizes; a trajectory planner which plans around all possible human motion is in danger of being overly conservative and prohibiting any robot motion at all. On the other hand, though human motion is often foreseeable, methods which aim to infer the intended motion may not account for an unforeseen movement such as a reflex action, and a collision may occur. A two-stage planner as in Fig. 1 is therefore proposed. A *long-term plan* is generated around the expected human motion. A *continuously updating short-term plan* then accounts for all possible motion and consists of the first section of the long term plan followed by a safe stop. A safety controller as described in [5] generates these plans and guarantees that

[1]The authors are with the Department of Informatics, Technische Universität München, 85748 Garching, Germany. aaron.pereira@tum.de, althoff@in.tum.de

a short-term plan always exists, though the safe stop is only executed if the human performs an unexpected movement and the subsequent section of the long-term plan is unsafe. This approach is tested in simulation in [6].
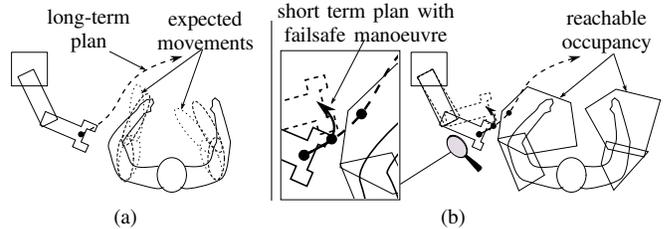
Fig. 1. (a) Long-term planning around expected human movement; (b) short-term failsafe manoeuvre in case of unexpected motion, based on *reachable occupancies* enclosing all possible motion.

Human motion prediction can be classified into 1) predicting a single trajectory and 2) predicting a set or probability distribution of trajectories. In the former case, natural human movements are assumed optimal with respect to an as yet unknown cost function of the dynamics [7]. In point-to-point movements, Morassi et al. [8] observe straight-line spatial hand trajectories and deduce that the human control system operates in Cartesian space. Flash and Hogan [9] show that movements minimise jerk and in [10], the authors predict natural movement of human arms using a muscle-effort minimisation criteria in a Digital Human Model (DHM).

However, we cannot reliably predict the intention of the human, hence several researchers use probabilistic models to predict movement. Koppula and Saxena [11] anticipate intended reaching goals using a temporal conditional random field. Mainprice and Berenson [12] predict future motion using a Gaussian Mixture Model and compute a voxel map of occupancy probabilities, which is used in a human-avoidance cost function to plan trajectories. Ding et al. predict human occupancy with a Hidden Markov Model (HMM) [13]. The latter authors note that unusual or previously unobserved movements would not be accounted for by an HMM and propose that this probabilistic method be supplemented by reachability analysis. Finally, instead of tracking human motion for a prediction model, first principles modeling is used in [14] to obtain a Markov chain to predict human movement around a mobile robot probabilistically.

What happens when humans do not behave as expected? Our short-term prediction must account for *all* movement of humans of *all* shapes and sizes. Reachability analysis can provide formal guarantees in systems where the environment is uncertain, e.g. automated driving [15], though to the
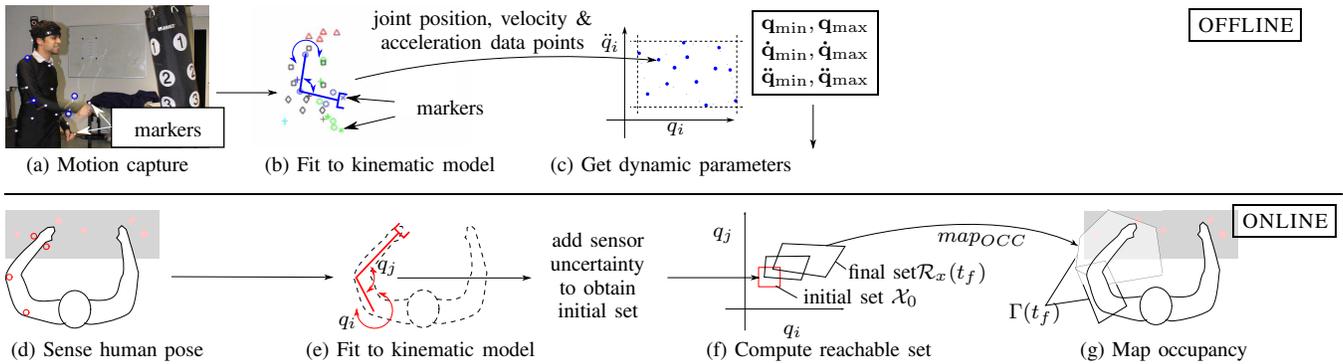
Fig. 2. Above: offline, we capture archetypal movements and fit to a kinematic model; from the resulting joint positions, velocities and accelerations we obtain a dynamic model. Below: online, sensor data (not necessarily motion capture) is fitted to the kinematic model to obtain the state in joint space; we calculate the reachable set at a future time with the dynamics obtained offline, then convert this to a reachable occupancy in Cartesian space

authors' best knowledge, reachability analysis has not yet been used to predict human movement for human-robot co-working. Our overapproximative prediction uses reachability analysis techniques to predict and enclose the entire possible reachable set in space after a given time.

In the following section we define the problem and our approach. In Sec. III we detail the model kinematics, how the archetypal movement data is captured and how this results in the dynamic parameters of the model. In Sec. IV we show how to convert a set of states in joint space to a potential occupancy in Cartesian space. We validate our approach on publicly available data in Sec. V and conclude in Sec. VI.

II. PROBLEM STATEMENT AND APPROACH

Our proposed approach has offline and online elements. *Offline*, we fit motion capture data of archetypal movements (Fig. 2a) to a kinematic model (Fig. 2b) to obtain relevant parameters for a dynamic model (Fig. 2c). *Online in real time*, sensor data is fit to the same kinematic model (Fig. 2d–e), taking into account uncertainty to obtain an initial set of states for the kinematic model. From the dynamics, the reachable set of states after a certain time is calculated (Fig. 2f) and mapped to Cartesian space to determine an overapproximative, reachable occupancy (Fig. 2g). This is then forwarded to a trajectory planner as described in the introduction. Note that the sensors in Fig. 2d need not be infrared motion capture: as we account for sensing error, less accurate pose estimation methods may be used as in [16]. The more uncertain the sensor data, the larger the reachable occupancy, and the more conservative the path planning. We define the reachable set of a system:

Definition 1 (REACHABLE SET). *Given a system with state $\mathbf{x}(t)$, input $\mathbf{u}(t)$ and dynamics $\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{u}(t))$, where t is time. The possible initial states $\mathbf{x}(0)$ and inputs $\mathbf{u}(t)$ are bounded by sets, $\mathbf{x}(0) \in \mathcal{X}_0, \mathbf{u}(t) \in \mathcal{U}(t)$. The reachable set of states $\mathbf{x}$ at $t = t_f$ is (see Fig. 2f):*

$$\mathcal{R}_{\mathbf{x}}(t_f) = \left\{ \mathbf{x}(t_f) = \int_0^{t_f} f(\mathbf{x}(t), \mathbf{u}(t)) dt \;\middle|\; \mathbf{x}(0) \in \mathcal{X}_0, \mathbf{u}(t) \in \mathcal{U}(t) \right\}$$

The human arm is a complex mechanical structure with unknown dynamics and actuation. We obtain knowledge about the position of the arm either directly though sensors or through an observer. As the internal structure of the arm is unknown and complex, we find the state of a simplified kinematic model from the inverse kinematic mapping:

$$map_{IK}(\mathbf{s}) : \mathcal{S} \to \mathcal{P}(\mathcal{X}),$$

which maps an instantaneous sensor reading $\mathbf{s}(t)$ to a set of states $\mathcal{X}(t)$ to account for noisy sensor data, where $\mathcal{P}(\mathcal{X})$ is the power set of $\mathcal{X}$. From the sensor reading at $t = 0$, we obtain the model's initial set of states $\mathcal{R}_{\mathbf{x}}(0) = map_{IK}(\mathbf{s}(0))$. We then define the occupancy mapping, which maps states to a subset of space occupied by the arm, as:

$$map_{OCC}(\mathbf{x}) : \mathcal{X} \to \mathcal{P}(\mathbb{R}^3).$$

Offline, we analyse our motion capture data (Fig. 2a–c) and obtain dynamic parameters for computing the reachable set $\mathcal{R}_{\mathbf{x}}(t)$ online (Fig. 2f) From this, we obtain the *reachable occupancy*, defined as $\Gamma(t) \supseteq \{map_{OCC}(\mathbf{x}) | \mathbf{x} \in \mathcal{R}_{\mathbf{x}}(t)\}$. We require that:

- $\Gamma(t)$ enclose the actual arm for any motion of any human which could possibly occur in a HRI environment;
- the computation time of $\Gamma(t)$ be less than t, so that it can be used in a collision-avoiding path planner;
- $\Gamma(t)$ be as tight as possible, to minimise false-positive collisions detected.

We next describe how the arm model is derived.

III. MODEL

Humans can be modelled as a tree of kinematic chains, e.g. [17]. As an initial demonstration of formally verified occupancy prediction, this work focusses only on the human arm: the arm can achieve high velocities compared to other body parts. Although models of torso and lower body movement exist, whose movement must eventually be integrated into the occupancy prediction (e.g. by adapting approaches from computer vision such as [18]) the purpose of this paper is to demonstrate the validity of the overapproximative prediction, hence we limit our scope to arm movement. In

this section, we describe the motion data collection and show how a well-known 4-degree-of-freedom (DOF) arm model can be reduced to 3-DOF, simplifying the calculation of $\mathcal{R}_x(t)$ and $\Gamma(t)$ by reducing dimensionality. We then detail how collected data is used to find dynamic parameters of the human arm.

A. Motion data

The validity of our prediction is based on accounting for the physical capabilities of a wide demographic of test subjects. Our 38 test subjects are 12 female and 26 male and range in age from 18 to 49 with a median age of 24. 50% do 3 hours or more of sport per week. Since the arm's scope of movement is infinite, 6 distinct archetypal movements were chosen which encompassed lateral, vertical and horizontal movements and ranged from one end of the arm workspace to another, to try and cover the arm's entire workspace. Subjects performed the motions (shown below and in the accompanying video) as fast as possible.
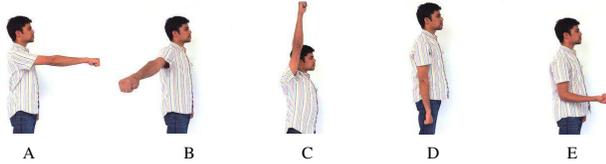
Fig. 3.   From left to right: Positions A–E.

1) Punch[2] to the front then recover to position E.
2) Punch to the front, ending in position A.
3) Position A to position B, elbow allowed to bend.
4) Position A to position B, elbow not allowed to bend.
5) Position C to position D, via position A.
6) Position C to position D, via position B.

A 6-camera Vicon Motion Capture system captures movements at 120Hz. These are then filtered with a 4th order Butterworth filter as in [17] to remove noise. A lower order filter does not sufficiently remove noise, whereas a higher order filter gives artificially attenuated joint accelerations; using this filter, accelerations match well with those observed in [19] from camera footage. Inverse kinematics based on the kinematic model in Sec. III-B are applied to the filtered data to obtain a time series of joint positions. Markers on the right arm are illustrated in Fig. 4. The base coordinate system of the arm is oriented with the clavicle coordinate system recommended in [20] and the origin is $40mm$ below (in negative y-direction from) the shoulder marker (RSHO), as we do not consider movement of the shoulder complex in this low degree of freedom model. The *elbow* was taken at the RELB marker and the *wrist* as the midpoint of RWRA and RWRB (the left arm was similarly defined). The *upper arm* is between shoulder and elbow and the *forearm* between elbow and wrist.

[2]Punching movements were performed with and without a punchbag, in case subjects moved faster against a punchbag due to its familiarity.
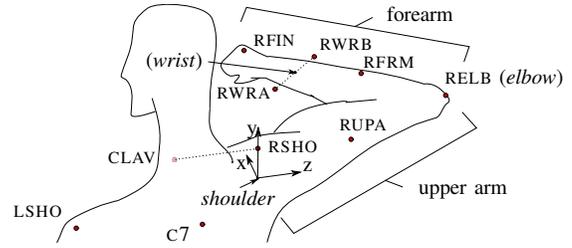
Fig. 4.   Markers and local base coordinate system for the right arm. The base coordinate system has its origin $40mm$ below RSHO and the markers CLAV, C7, T10 and STRN (latter 2 not shown), define its orientation. Note that the z-axis is collinear to RSHO and CLAV. Markers RFIN, RFRM and RUPA are not used for inverse kinematics.
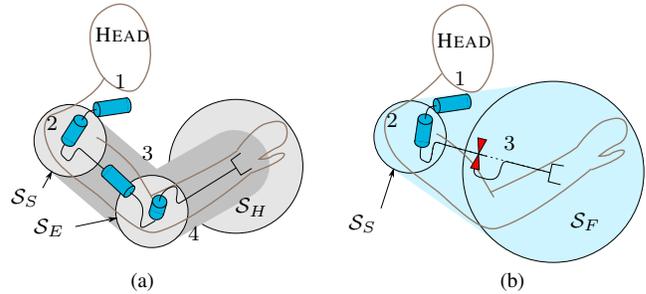
Fig. 5.   (a) A well-known 4-DOF model, (b) a simplified 3-DOF model.

B. Kinematic Model

Arm models with 4-DOF are widely used in modelling arm movement, e.g. for determining arm workspace [21] or predicting position during movements [10]. Wrist movement and forearm rotation are relatively limited and do not greatly affect the occupancy, hence we can account for all movement of the hand in a sphere $\mathcal{S}_H$ around the wrist, see Fig. 5a; the elbow and shoulder are enclosed in spheres $\mathcal{S}_E$ and $\mathcal{S}_S$, and the upper arm and forearm are capsules swept from shoulder to elbow and from elbow to wrist respectively. The radius of $\mathcal{S}_H$ is the 95th percentile hand length from anthropometric studies [22], $0.205m$; that of $\mathcal{S}_S$ and $\mathcal{S}_E$ is $0.08m$, which is larger than the radius of human upper or lower arms but intended to account for clothing. We call the position of the ith joint q_i.

The 4-DOF arm contains a singularity in the elbow when fully extended: as shown in Fig. 6, the inverse kinematics
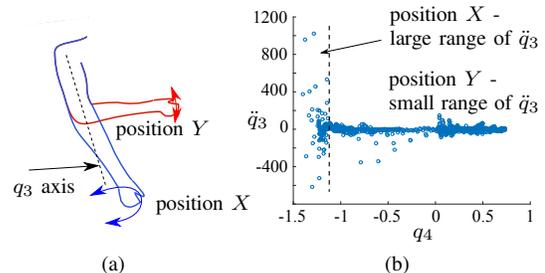
Fig. 6.   (a) positions X near singularity, and Y, away from singularity (b) $\ddot{q}_3$ plotted against q_4 during one movement. At X, extremely high accelerations in are observed, at Y the range is smaller.
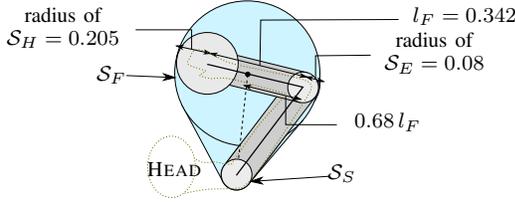
Fig. 7. Determining the radius of $\mathcal{S}_F$ from the maximum length forearm, so as to enclose smaller forearms, and thus determining the relative position of the end effector between elbow and wrist. l_F is forearm length.
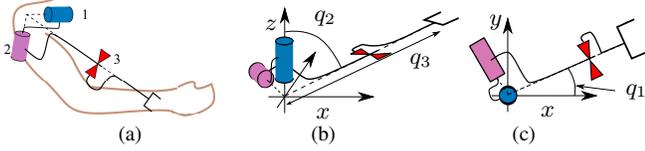
Fig. 8. (a) oblique view (b) lateral view and (c) top view of the 3-DOF model kinematics. q_1 and q_2 are joint angles of the first two rotational joints; q_3 is the extension of the prismatic joint. x, y and z is the base (clavicle) coordinate system.

to determine the 3rd joint are undefined at singularity, and in the vicinity of the singularity q_3 is highly susceptible to noise and small task-space movements. We identify two possible solutions: 1) create a hybrid dynamic model with one dynamic regime near the elbow singularity and another away from it, or 2) make a kinematic parameterisation of the arm which avoids the elbow singularity.

As reachable sets of hybrid systems are time-consuming to calculate and often introduce large overapproximations [23], we opt for the latter option, see Fig. 5b. We enclose the forearm and hand in a sphere $\mathcal{S}_F$ and the shoulder in a sphere $\mathcal{S}_S$ as shown in Fig. 7. The occupancy of the arm is then taken as the convex hull of $\mathcal{S}_F$ and $\mathcal{S}_S$, which encloses the upper arm also by the property of convexity. The diameter of $\mathcal{S}_F$ is $0.627m$, calculated from maximum forearm length observed over our entire dataset, $0.342m$, plus the radius of $\mathcal{S}_H$, plus the radius of $\mathcal{S}_E$. This accounts for parametric uncertainty in the forearm and hand, as smaller length hands and forearms will be enclosed in the largest. In the maximal-length forearm and hand, the centre of $\mathcal{S}_F$, enclosing the forearm, hand and $\mathcal{S}_E$, is a point $\frac{0.627/2 - 0.08}{0.342} = 0.68$ of the distance from elbow to wrist. We call this point the *end effector* and choose its position at 0.68 of the distance from elbow to wrist, for all humans. As both elbow and shoulder are enclosed, by convexity, the upper arm is enclosed regardless of its length.

C. Placement of 1st Joint Axis to Avoid Shoulder Singularity

Consider Fig. 8. If the end effector lies on the axis of joint 1, the inverse kinematics cannot give a value for q_1. This second singularity is only a coordinate singularity: the shoulder is a spherical joint and hence the axis of joint 1 can be freely chosen. We choose it such that it does not intersect the workspace of the end effector (i.e. the end effector can never lie on the joint 1 axis). Fig. 9 shows the sampled workspace over all our data and the chosen joint 1 axis.
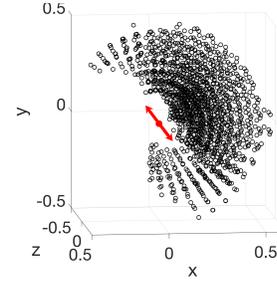
Fig. 9. Sampled workspaces of 3-DOF model end effector, and first joint axis in red. The axis does not intersect the workspace. Coordinate system is that of the shoulder; scale in metres.

The alignment of the joint 1 axis is along the long axis of the clavicle[3].

D. Dynamics of the Arm

Having described the kinematics of the arm model, we now consider its dynamics. The human arm is actuated by antagonistic pairs of muscles exerting torque on one or more joints; this torque is dependent on the moment arm of muscles about the joint, muscle length, and inter-muscular force transmission [24]. Maximum muscle torques are well studied and are used in several digital human models (DHMs), e.g. [25]. These can accurately model human body dynamics so that sensor data may be fit to a DHM to accurately reconstruct body part positions [17]. A model based on maximum torques is impractical as: 1) mass and inertia parameters are uncertain, 2) external torques are unknown, 3) the direction of the gravity force depends on shoulder orientation and 4) the relationship between torque, and joint angles and velocities, is highly nonlinear. Though reachable sets for nonlinear systems with parametric uncertainty can be calculated, such methods (described in e.g. [26]) enlarge the overapproximation and are not real-time compatible. Instead, we use a simple model based on extreme joint positions, speeds and accelerations. To the best knowledge of the authors, a complete study of maximum accelerations of all human arm joints does not exist. Our dynamic model ignores dependencies on joint position or velocity and considers each joint as independent. This is overapproximative: we may include states that are unreachable due to dependencies, but we never exclude any reachable state.

The intersection of three simple models for joint position, velocity and acceleration limits is faster to compute than a complex model which accounts for all limits simultaneously. As in [15, Prop. 1], where the reachable sets $\mathcal{R}_1, ... \mathcal{R}_n$ of n models $\mathcal{M}_1, ... \mathcal{M}_n$ are all overapproximative sets which enclose the exact reachable set $\mathcal{R}_e$, then $\mathcal{R}_e \subseteq \bigcap_{i=1}^{n} \mathcal{R}_i$. We define the Minkowski sum, over two sets A and B, as $A \oplus B = \{a + b \mid a \in A, b \in B\}$.

Letting $\mathcal{Q}(0), \dot{\mathcal{Q}}(0)$ be the sets of initial positions and initial velocities in joint space respectively, and using the

[3]The reader may verify with a little gymnastics that it is impossible to align the humerus with the clavicle, as it is hindered by the Acromial Process.

maximum and minimum accelerations $\ddot{\mathbf{q}}_{\min}, \ddot{\mathbf{q}}_{\max}$, velocities $\dot{\mathbf{q}}_{\min}, \dot{\mathbf{q}}_{\max}$ and positions $\mathbf{q}_{\min}, \mathbf{q}_{\max}$ from the archetypal movements, the reachable sets our models predict are:

1) a 0$^{\text{th}}$ order model: $\mathcal{R}_{\mathbf{q}}^{(0)}(t) = [\mathbf{q}_{\min}, \mathbf{q}_{\max}]$

2) a 1$^{\text{st}}$ order model: $\mathcal{R}_{\mathbf{q}}^{(1)}(t) = \mathcal{Q}(0) \oplus [\dot{\mathbf{q}}_{\min}t \ , \ \dot{\mathbf{q}}_{\max}t]$

3) a 2$^{\text{nd}}$ order model:

$$\mathcal{R}_{\mathbf{q}}^{(2)}(t) = \mathcal{Q}(0) \oplus \dot{\mathcal{Q}}(0)t \oplus \left[\ddot{\mathbf{q}}_{\min}\frac{t^2}{2} \ , \ \ddot{\mathbf{q}}_{\max}\frac{t^2}{2} \right]$$

The intersection $\mathcal{R}_{\mathbf{q}}(t) = \mathcal{R}_{\mathbf{q}}^{(0)}(t) \cap \mathcal{R}_{\mathbf{q}}^{(1)}(t) \cap \mathcal{R}_{\mathbf{q}}^{(2)}(t)$ yields a tighter overapproximation than any one model alone, and joint position, velocity and acceleration limits are not exceeded. $\mathcal{Q}(0), \dot{\mathcal{Q}}(0)$ are box-shaped (we add an independent margin of error based on sensing uncertainty for each joint) hence each of the above models is a product of intervals, whose intersection is fast to compute. As minimising spatial jerk is observed in human trajectory planning [9], adding higher order models would be intuitive; this requires, however, that the model state include acceleration, i.e. $\mathbf{x} = [\mathbf{q}^\top, \dot{\mathbf{q}}^\top, \ddot{\mathbf{q}}^\top]^\top$. Currently, instantaneous velocity is interpolated from consecutive state readings, which is noisy, but accuracy is adequate. Interpolating acceleration is even noisier, which would eradicate any advantages gained with a 3$^{\text{rd}}$ or higher order model.

IV. REPRESENTATION IN SPACE

This section concerns the function map_{OCC} from Sec. II, i.e. translating the joint space reachable set $\mathcal{R}_x(t)$ into the occupancy $\Gamma(t)$. The approach by Täubig et al. [27] is promising, and is applied to human models in [28]. It is overapproximative (as opposed to sampling-based methods e.g. [29]) yet tight, and computes a sphere-swept volume enclosing the occupancy of a link, when the joint angles lie within a product of intervals. A sphere-swept volume is the Minkowski sum of a polytope and a sphere. Our model is the convex hull of two spheres of different radii, which itself is not a sphere-swept volume. As the link must be enclosed in a sphere-swept volume, we create, for this method, a capsule $\mathcal{C}_{FS}$, enclosing both $\mathcal{S}_F$ and $\mathcal{S}_S$; see Fig. 10.
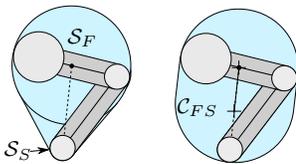
Fig. 10.   3-DOF model enclosed in capsule $\mathcal{C}_{FS}$ for using method in [27].

Polytopes can be represented as a set of intersecting halfspaces (*H-representation*) or a set of vertices (*V-representation*), and the representation of polytopes determines the method of collision checking used. The polytopes obtained in [27] are in V-representation and consequently, the collision check is performed with the GJK algorithm [30]. Collision checking between H-representations, on the other hand, is a simple matter of determining whether a solution exists to a linear system of equations, and can use methods such as Fourier-Motzkin elimination, which allows for parallelisation [30]. Although conversion between representations is possible [31], in order to save computation time, it is advantageous to obtain the polytope in the representation required. We therefore present a method for quickly obtaining the H-representation of the human arm occupancy in the next section.

A. H-Representations

The H-representation for enclosing the occupancy of the 3-DOF model is found as follows. Recall from Fig. 5b that arm occupancy is the convex hull of the spheres $\mathcal{S}_S$ around the origin and $\mathcal{S}_F$ around the end effector. The origin is fixed; the end effector position depends on joint positions $\mathbf{q}$. Let $f(\mathbf{q})$ be the forward kinematic function which gives the Cartesian position of the end effector. We first bound the occupancy of the end effector $f(\mathcal{R}_q) = \{f(\mathbf{q})|\mathbf{q} \in \mathcal{R}_q\}$ with planes (which define a H-representation polytope), then shift these planes outwards by the radius of $\mathcal{S}_F$, so all positions of $\mathcal{S}_F$ are enclosed. If this polytope does not contain $\mathcal{S}_S$, the relevant plane is shifted further out until $\mathcal{S}_S$ is enclosed (see Fig. 13). As the polytope is convex, if $\mathcal{S}_F$ and $\mathcal{S}_S$ are enclosed, so is their convex hull. From the kinematics in Fig. 8, we see that the second rotary joint is orthogonal to the first, the axes of all joints go through the origin, and the prismatic joint is zero at the origin, hence we observe the following:

Observation 1 (Spherical coordinate–joint position correspondance). *The joint positions (q_1, q_2, q_3) correspond to the spherical coordinates of the end effector position (θ, ψ, r) in a coordinate system where the z-axis is aligned with the first joint axis.*

Since $\mathcal{R}_{\mathbf{q}}$ is the product of joint positions intervals $[q_{1,\min}, q_{1,\max}] \times [q_{2,\min}, q_{2,\max}] \times [q_{3,\min}, q_{3,\max}]$, the spherical coordinates of the end effector position $f(\mathcal{R}_{\mathbf{q}})$ satisfy $q_{1,\min} \leq \theta \leq q_{1,\max}$, $q_{2,\min} \leq \psi \leq q_{2,\max}$, $q_{3,\min} \leq r \leq q_{3,\max}$. This is illustrated in Fig. 11, where p_{1-8} are the vertices of $f(\mathcal{R}_{\mathbf{q}})$.
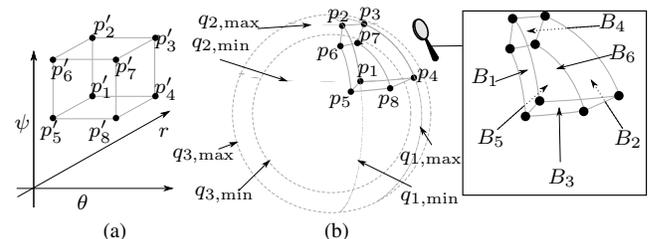
Fig. 11.   The joint space reachable set (a) $\mathcal{R}_{\mathbf{q}}$ and (b) the corresponding end effector occupancy $f(\mathcal{R}_{\mathbf{q}})$. Vertices p_{1-8} of $f(\mathcal{R}_{\mathbf{q}})$ correspond to vertices p'_{1-8} of the reachable set.

The surfaces B_{1-6} bound $f(\mathcal{R}_{\mathbf{q}})$. Letting (θ, ψ, r) be the point defined by spherical coordinates $\theta \in (-\pi, \pi]$, $\psi \in [0, \pi]$, and $r \in \mathbb{R}_{\geq 0}$, each of these surfaces is a subset

of a level set:

$$B_1 \subseteq \{(\theta, \psi, r) | \theta = q_{1,\min}\}; \quad B_2 \subseteq \{(\theta, \psi, r) | \theta = q_{1,\max}\}$$
$$B_3 \subseteq \{(\theta, \psi, r) | \psi = q_{2,\min}\}; \quad B_4 \subseteq \{(\theta, \psi, r) | \psi = q_{2,\max}\}$$
$$B_5 \subseteq \{(\theta, \psi, r) | r = q_{3,\min}\}; \quad B_6 \subseteq \{(\theta, \psi, r) | r = q_{3,\max}\}$$

To build the H-representation, we require a set of bounding planes rather than surfaces of arbitrary shape. We observe that each set of 4 vertices from p_{1-8} defining each surface B_i are coplanar, as $f(\mathcal{R}_\mathbf{q})$ is symmetrical in the plane $\theta = \frac{q_{1,\min}+q_{1,\max}}{2}$. The planes defined by these sets of 4 vertices, however, do not necessarily bound $f(\mathcal{R}_\mathbf{q})$, as explained below.

Observation 2 (Level sets of spherical coordinates). *The locus of points with constant θ is a plane; that with constant ψ is a cone; that with constant r is a sphere. [32]*

Given that $q_{2,\max} - q_{2,\min} < \pi$, each surface B_i is either locally concave, convex or planar. Convex surfaces can be bounded by one or more planes tangent to them. Concave surfaces can be bounded by a plane through extreme points of the surface. From Obs. 2, B_1 and B_2 are planes, hence our first two bounding planes H_1, H_2 are simply these.
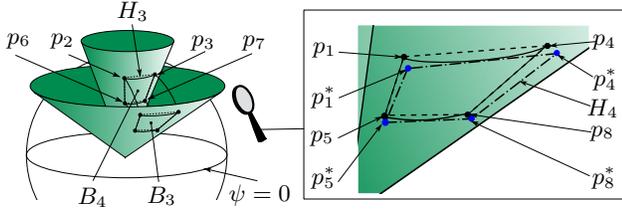
Fig. 12. Shaded areas are level sets of ψ. H_3 is defined by the plane through points $p_{2,3,6,7}$ as the surface is locally concave. However, if $p_{1,4,5,8}$ are in the same hemisphere as $p_{2,3,6,7}$, the surface is locally convex and the plane through $p_{1,4,5,8}$ (right, dashed line) does not bound $f(\mathcal{R}_\mathbf{q})$. Thus we find $p^*_{1,4,5,8}$ such that the plane (dot-dashed line) is tangent to $f(\mathcal{R}_\mathbf{q})$.

Without loss of generality, the surface B_4 lies in the top hemisphere as in Fig. 12. We see that B_4 is locally concave, so the plane passing through the extreme points $p_{2,3,6,7}$ bounds $f(\mathcal{R}_\mathbf{q})$; we take this to be H_4. However, if $q_{2,\max}$ is the same sign as $q_{2,\min}$, then B_3 is convex and the plane passing through the extreme points does not bound $f(\mathcal{R}_\mathbf{q})$ but intersects it. We therefore generate some new points $p^*_4, p^*_4, p^*_5, p^*_8$ (Fig. 12) such that the plane H_3 through these is tangent to B_3 and the θ and r values of p_i and p^*_i are the same (i.e., p_i and p^*_i lie on the same level sets of θ and r). Analytically[4], the ψ values of these points are $\tan^{-1}(\tan(q_{2,\min})\cos(\frac{q_{1,\max}-q_{1,\min}}{2}))$.

Finally, we bound radially, from above and below. Let the surfaces defined by p^*_5, p_6, p_7, p^*_8 and p^*_1, p_2, p_3, p^*_4 be called B'_5 and B'_6, as they are extensions of B_5 and B_6 respectively (they lie on the same level sets, which are boundaries of spheres). B'_5 is concave, so we take H_5 defined by its endpoints. B'_6 is convex so we take $H_{6,7,8,9,10}$ tangent to it at points p^*_1, p_2, p_3, p^*_4 and in the middle. The more tangent planes we take, the tighter the overapproximation, but the more complex the polytope. H_{1-10} are illustrated in Fig. 13.

[4]The proof is omitted for brevity.
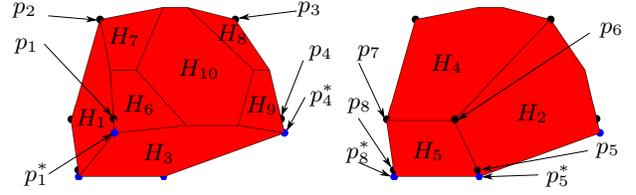
Fig. 13. All planes H_{1-10} bound the occupancy; front and back views.

Thus the reachable occupancy of the end effector is bounded. To enlarge this to $\Gamma(t)$ we shift each halfspace outwards by the radius of the forearm sphere $\mathcal{S}_F$ and, if need be, move the halfspace H_5 outward to enclose the shoulder sphere $\mathcal{S}_S$, see Fig. 14.
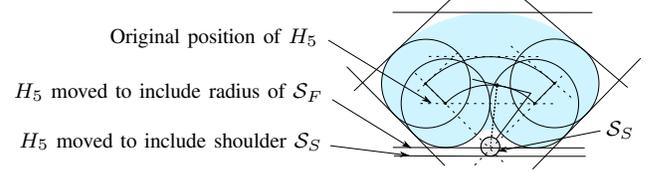
Fig. 14. A 2-D projection showing how to account for $\mathcal{S}_F$ and $\mathcal{S}_S$. Planes bounding $f(\mathcal{R}_\mathbf{q})$ (dashed) are moved outward to account for radius of $\mathcal{S}_F$ (solid). If $\mathcal{S}_S$ is not enclosed, we move H_5 to enclose it.

B. ISO-compliant model

For comparison, we introduce an ISO-based representation $\Gamma_{ISO}(t)$. The Hand Speed Constant $v = 1.6m/s$ is given in [1] as the maximum speed of a non-walking human's body parts. We take the 4-DOF model from Sec. III-B and enlarge it by $v \cdot t$ as shown in Fig. 15.
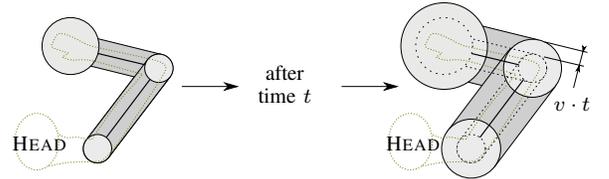
Fig. 15. Prediction $\Gamma_{ISO}(t)$ based on Hand Speed Constant from [1].

V. VALIDATION

In Sec. III-A we choose the archetypal movements such that a wide range of human motion is covered, and they are performed as fast as possible to obtain maximum accelerations and velocities for the model dynamics. To ensure that the parameters obtained from this subset of motion are valid for all motion encountered in an HRI scenario, we validate the dynamic model built from our test data against publicly available motion capture data from Carnegie Mellon University[5]. We group these motions, of lengths from $\frac{1}{2}s$–$30s$, into the following categories:

- *Everyday motions* e.g. construction work, machining work, manipulating objects, stumbling; expected to be

[5]Available at mocap.cs.cmu.edu, accessed on 11.08.15. The authors will happily provide details of which data files were used.

exemplary of the behaviour of humans comfortable working in the vicinity of robots (96 motions)

- *Sports-related motions* e.g throwing, catching and batting balls, boxing. (67 motions)
- *Dance-related motions* e.g Indian dance, modern dance and swing dance. (58 motions)
- *Acrobatic motions* – motions where both feet are simultaneously in the air, e.g. jumps, cartwheels, backflips and swings from a trapeze. (68 motions)

We check, for both arms, whether the arm at time t (in this case, all markers on the arm) is entirely contained in $\Gamma(t)$, for $t = 16.7, 25.0$ and 33.3ms[6] (Tab. I); this is shown in the video attachment. We evaluate the volume of the reachable occupancy (Tab. II) and computation times. As the calculations have no iterative step, hard real-time can be achieved, though for this verification a non-real-time OS running MATLAB R2016a on a 2.8GHz i7 processor with 16GB RAM is used. Polytope visualisation and volume calculation are performed by the MPT toolbox [33]. In case markers are lost or incorrectly tracked by the software, we validate the sensor data using distance checks, i.e. markers on rigid links cannot move relative to each other, and disregard data where markers violate these checks. We account for sensor noise with an uncertainty of $\pm 0.05 \ rad$ and $0.02 \ m$ added to the revolute and prismatic joint positions respectively.

TABLE I

PUBLIC DATABASE MOVEMENTS FULLY CONTAINED IN REACHABLE OCCUPANCY (NO. CONTAINED/TOTAL, FOR $t = 16.7, 25$ AND $33.3ms$)

Category	Everyday	Sport	Dance	Acrobatics
$\Gamma(t)$ (Swept Volume)	96/96	67/67	58/58	60/68
$\Gamma(t)$ (H-representation polytope)	96/96	67/67	58/58	60/68
$\Gamma_{ISO}(t)$ (ISO-based)	93/96	44/67	19/58	17/68

In all cases for the *Everyday*, *Sport* and *Dance* motions, the markers at each prediction horizon were within the respective reachable occupancy, in contrast to the occupancy based on the maximum arm speeds from ISO Standards [1], where only 71% of these motions were within $\Gamma_{ISO}(t)$ at all times. The dynamic model could be extended by including *Acrobatic* motions in the archetypal motions used to determine the model parameters. In this case, the reachable occupancy would be enlarged, making the path planner extra cautious at the expense of performance. As behaviour guidelines in workshops typically prohibit running and acrobatic-like movements, one may argue that such movements would not occur in a factory setting and this range of human motion can be safely ignored in a certifiably safe robot motion planner; injuries resulting from such behaviour on the workshop floor would not be the fault of the robot.

Volume calculation of the ISO-based and swept-volume representations uses the method from [27], which underestimates the true volume. Fig. 16 and Tab. II show the growth

[6]since the data is captured at $120Hz$, these are simply 2, 3 and 4 timesteps in the future respectively.

TABLE II

VOLUME OF OCCUPANCIES (m^3)

Time t	ISO-based	Swept Volume	H-Rep Polytope
16.7ms	0.06	0.28	0.41
25.0ms	0.07	0.38	0.51
33.3ms	0.09	0.54	0.66
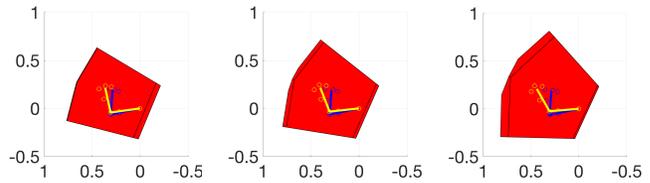
Fig. 16. Left to right: reachable occupancies at 16.7, 25.0 and 33.3ms. Scale in metres; origin at shoulder. Inner (blue) outline is arm at $t = 0$, and outer outline (yellow) is arm at aforementioned times.

of the reachable occupancy over time. Beyond $33ms$, the occupancy grows quite large (e.g. at $50ms$, around $1.5m^3$). If the robot and human are in close proximity, therefore, shorter update cycles for the prediction are required. [21] gives the workspace of the wrist of an average-height person as $0.667\pm0.055m^3$; as the hand is perhaps 25% of the length of the arm, the workspace of the whole arm may be estimated as $\frac{0.667}{0.75^3} \approx 1.6m^3$. For short time intervals, therefore, our prediction is tight.

TABLE III

COMPUTATION TIME (ms)

Process	Inverse kinematics and reachability	Calculating swept volume	Calculating H-representation
Av. time	$0.38ms$	$0.31ms$	$0.08ms$

Computation time is shown in Tab. III – for both representations $\Gamma(t)$ is calculated within $1ms$. Although the H-representation is slightly larger, there is no expensive convex hull computation and it is therefore faster. As time is of the essence, a coarser, faster overapproximation may be an acceptable time-accuracy trade-off, alternatively, a hierarchical or anytime algorithm can be envisaged where collisions detected in the H-representation are checked for false positivity only if computation time remains. To ensure the entire system, i.e. sensing, prediction and path planning/verification, is within real time, fast algorithms must be used. In [6] fast verification and path planning is demonstrated. Infrared motion capture has latencies in the range of $2-4ms$[7], although in industrial scenarios, cameras with larger latencies are more likely.

VI. CONCLUSIONS

We present a method to formally bound the future potential occupancy of a human arm in real time based on sensor

[7]see e.g. `vicon.com/products/software/tracker`

data. Mean computation times of less than $1ms$ demonstrate its suitability for use in an online dynamic path planner. The arm model parameters are generated from archetypal movements of test subjects and are used to predict, in real time and overapproximatively, the subset of space that may be occupied by the arm after a given time. We present two methods for converting joint angles to convex sets: a fast method yielding H-representation polytopes and a slower but tighter method yielding sphere-swept volumes from the literature. The prediction is tested on publicly available motion capture data and found to correctly overapproximate the occupancy of movements in a factory setting. In contrast to probabilistic methods, our technique can be used to certify safety in human-robot interaction and co-working, a goal long unattainable due to insufficient guarantee of safety.

REFERENCES

[1] "Safety of machinery - positioning of safeguards with respect to the approach speeds of parts of the human body," ISO Standard 13855:2010, International Organization for Standardization, 2010.

[2] "Robots and robotic devices - safety requirements for industrial robots - part 1: Robots," ISO Standard 10218-1:2011, International Organization for Standardization, 2011.

[3] A. Zanchettin, N. Ceriani, P. Rocco, H. Ding, and B. Matthias, "Safety in human-robot collaborative manufacturing environments: Metrics and control," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 2, pp. 882–893, 2016.

[4] S. Petti and T. Fraichard, "Safe motion planning in dynamic environments," in *Proc. IEEE-RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 2210 – 2215, 2005.

[5] A. Pereira and M. Althoff, "Safety control of robots under computed torque control using reachable sets," in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, pp. 331 – 338, 2015.

[6] M. J. Zeestraten, A. Pereira, M. Althoff, and S. Calinon, "Online motion synthesis with minimal intervention control and formal safety guarantees," in *Proc. of IEEE Systems, Man and Cybernetics*, 2016. (accepted).

[7] S. Albrecht, M. Leibold, and M. Ulbrich, "A bilevel optimization approach to obtain optimal cost functions for human arm movements," *Numerical Algebra, Control and Optimization*, vol. 2, no. 1, pp. 105–127, 2012.

[8] P. Morasso, "Three dimensional arm trajectories," *Biological Cybernetics*, vol. 48, no. 3, pp. 187–194, 1983.

[9] T. Flash and N. Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model," *The Journal of Neuroscience*, vol. 5, no. 7, pp. 1688–1703, 1985.

[10] V. De Sapio, J. Warren, and O. Khatib, "Predicting reaching postures using a kinematically constrained shoulder model," in *Advances in Robot Kinematics* (J. Lenarčič and B. Roth, eds.), pp. 209–218, Springer Netherlands, 2006.

[11] H. Koppula and A. Saxena, "Anticipating human activities using object affordances for reactive robotic response," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 1, pp. 14–29, 2015.

[12] J. Mainprice and D. Berenson, "Human-robot collaborative manipulation planning using early prediction of human motion," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, pp. 299–306, 2013.

[13] H. Ding, K. Wijaya, G. Reißig, and O. Stursberg, "Online computation of safety-relevant regions for human robot interaction," in *Proc. 43rd Intl. Symp. Robotics (ISR)*, 2012.

[14] F. Rohrmüller, M. Althoff, D. Wollherr, and M. Buss, "Probabilistic mapping of dynamic obstacles using Markov chains for replanning in dynamic environments," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 2504–2510, 2008.

[15] M. Althoff, D. Heß, and F. Gambert, "Road occupancy prediction of traffic participants," in *Proc. of the 16th International IEEE Conference on Intelligent Transportation Systems*, pp. 99–105, 2013.

[16] R. Poppe, "Vision-based human motion analysis: An overview," *Comput. Vis. Image Underst.*, vol. 108, no. 1-2, pp. 4 – 18, 2007.

[17] E. Demircan, T. Besier, S. Menon, and O. Khatib, "Human motion reconstruction and synthesis of human skills," in *Advances in Robot Kinematics: Motion in Man and Machine* (J. Lenarcic and M. M. Stanisic, eds.), pp. 283–292, Springer Netherlands, 2010.

[18] K. Yamane and Y. Nakamura, "Efficient parallel dynamics computation of human figures," in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, vol. 1, pp. 530–537 vol.1, 2002.

[19] A. A. Amis, D. Dowson, and V. Wright, "Analysis of elbow forces due to high-speed forearm movements," *J. Biomech.*, vol. 13, no. 10, pp. 825–831, 1980.

[20] G. Wu, F. C. van der Helm, H. D. Veeger, M. Makhsous, P. V. Roy, C. Anglin, J. Nagels, A. R. Karduna, K. McQuade, X. Wang, F. W. Werner, and B. Buchholz, "ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion-part ii: shoulder, elbow, wrist and hand," *J. Biomech.*, vol. 38, no. 5, pp. 981 – 992, 2005.

[21] N. Klopčar and J. Lenarčič, "Kinematic model for determination of human arm reachable workspace," *Meccanica*, vol. 40, no. 2, pp. 203–219, 2005.

[22] S. Pheasant and C. M. Haslegrave, *Bodyspace: Anthropometry, Ergonomics and the Design of Work*, ch. Hands and Handles, pp. 143–160. Taylor & Francis CRC Press, 2006.

[23] M. Althoff, "Reachability analysis of nonlinear systems using conservative polynomialization and non-convex sets," in *Hybrid Systems: Computation and Control*, pp. 173–182, 2013.

[24] D. Hahn, M. Olvermann, J. Richtberg, W. Seiberl, and A. Schwirtz, "Knee and ankle joint torque-angle relationships of multi-joint leg extension," *J. Biomech.*, vol. 44, pp. 2059–2065, July 2011.

[25] K. Holzbaur, W. Murray, and S. Delp, "A model of the upper extremity for simulating musculoskeletal surgery and analyzing neuromuscular control," *Ann. Biomed. Eng.*, vol. 33, no. 6, pp. 829–840, 2005.

[26] M. Althoff, O. Stursberg, and M. Buss, "Reachability analysis of nonlinear systems with uncertain parameters using conservative linearization," in *Proc. of the 47th IEEE Conference on Decision and Control*, pp. 4042–4048, 2008.

[27] H. Täubig, B. Bäuml, and U. Frese, "Real-time swept volume and distance computation for self collision detection," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 1585–1592, 2011.

[28] M. Ragaglia, A. Zanchettin, and P. Rocco, "Safety-aware trajectory scaling for human-robot collaboration with prediction of human occupancy," in *Proc. Int. Conf. Advanced Robotics*, pp. 85–90, 2015.

[29] T. McMahon, S. Thomas, and N. Amato, "Sampling-based motion planning with reachable volumes: Theoretical foundations," in *Proc. IEEE Int. Conf. on Robotics and Automation*, pp. 6514–6521, 2014.

[30] C. Ericson, *Real-Time Collision Detection*. Morgan Kaufmann, 2004.

[31] V. Kaibel and M. Pfetsch, "Some algorithmic problems in polytope theory," in *Algebra, Geometry and Software Systems* (M. Joswig and N. Takayama, eds.), pp. 23–47, Springer Berlin Heidelberg, 2003.

[32] U. Graf, *Trigonometrie der Ebene, Sphärische Geometrie und Kartenentwürfe*. Quelle & Mener, 1938. (in German).

[33] M. Herceg, M. Kvasnica, C. Jones, and M. Morari, "Multi-Parametric Toolbox 3.0," in *Proc. of the European Control Conference*, (Zürich, Switzerland), pp. 502–510, 2013. http://control.ee.ethz.ch/~mpt.