Gaussian Process Kernels for Rotations and 6D Rigid Body Motions

Muriel Lang, Oliver Dunkley and Sandra Hirche

Abstract— **Gaussian Processes (GPs) are gaining increasing popularity due to their expressive power for learning the dynamics of non-linear time series data, e.g. for human motion prediction. However, so far they are restricted to Euclidean space: input data such as position and velocity need to be Euclidean. In this paper, we examine GPs over time series of 6D rigid body motions including large rotations. As the use of Euler angles with large rotations results in inaccurate predictions, we present an extension of the valid input data to quaternions $\mathbb{H}$ and dual quaternions $\mathbb{H}_D$. The quality of a GP prediction over unit quaternions is compared with GP prediction over Euler angles. The results are evaluated based on experimental data from a quadrotor and in a learning task of a collision free 6D motion trajectory incorporating large rotations based on artificial data from a motion planner.**

I. INTRODUCTION

In human-robot interaction the robot needs to adapt to a human partner to maximize assistive performance. The goal-directed interaction of robots with humans is one of the most challenging issues in unstructured environments such as for instance in elderly care or other general assistive scenarios. Precise predictions of behavioral goals of humans are essential for human-robot interaction, in terms of intention estimation from gestures and motions. Due to the absence of physical models that can accurately describe human movements, stochastic learning models are gaining popularity owing to its robustness in predicting dynamic human movements from mere observations.

In order for the robot to learn and adapt to nonlinear dynamics of human behavior in human-robot interaction, Gaussian Processes (GPs) [1], Gaussian mixture models (GMMs) [2] and Hidden Markov models (HMMs) [3], [4] have been used. While HMMs are one of the most popular methods for describing motion primitives supporting human-robot interaction, these methods are limited to a discrete state space and less suitable for modeling continuous motion. Therefore, in this work we focus on GPs for estimating highly nonlinear dynamics of motion such as human behavior. Furthermore, the GPs requires minimum model assumptions [5] and the process is more generalizable. However, GPs (as well as GMMs) have a drawback so far; they are only defined in the Euclidean space because, for the matrix inversion, orthogonal coordinate axes are required. Hence, the input data for the models are limited to position, force data, translational velocities or accelerations. When orientation is required in addition, Euler

angles are typically used as a rotation representation (see for GPs [6], [7] and for GMMs [2]). When the rotation angles are small, this approach is sufficient as the space of Euler orientations close to zero is approximately Euclidean. However, another rotation parametrization is required when the rate of orientation change becomes significantly large at high angular speed, low sampling frequency or if loss of input data occurs. One solution for such problem is a use of the axis-angle orientation representation. However, it requires more extensive probability distributions than the normal distribution such as the wrapped Gaussian or the von Mises distribution [8]. Furthermore, no efficient composition algorithm necessary for the GP prediction is available in the space of axis-angles as the vector sum is not defined. In contrast, the representation by unit quaternions defined on the unit sphere S_3 is appropriate to solve GP regression problems regarding large rotations. In addition, the unit quaternions is extended to dual quaternions so that it can singularly represent 6D rigid body motions in the space $S_3 \times \mathbb{R}^3$.

In this paper, we therefore demonstrate how rotations and translations represented by dual-quaternions can suitably be incorporated to GPs using continuous data. The significant contribution of the current paper is that we introduce an extension of the valid GP input space to the non-Euclidean space of quaternions $\mathbb{H}_D$ and dual quaternions $\mathbb{H}$ to perform GP predictions on 6D rigid body motions. We define three squared exponential kernels each based on a metric on the unit quaternions. We then evaluate the performance of a GP prediction over unit quaternions and Euler angles on real rotation data and learn a 6D motion trajectory over dual quaternions.

The present paper is organized as follows. In Section II the GP kernels are introduced: we define the GP and explain briefly the 6D pose representation by dual quaternions. Secondly, we introduce the GP over pure rotations and 6D rigid body motions. Finally, in Section III, we compare the different metrics using real data and present a solution from the GP over dual quaternions for an example data set which includes large rotations and highlight it's superiority over regular GPs.

II. GP FOR 6D RIGID MOTIONS

In this section the GP over 6D rigid motions, parametrized by dual quaternions is introduced. This GP can be used for modeling and predicting system dynamics including significant rotations and translations. Firstly, some background information on GP regression is provided to biefly introduce a 6D pose representation by dual quaternions. For better

M. Lang, O. Dunkley and S. Hirche are with the Institute for Information-Oriented Control, Faculty of Electrical Engineering and Information Technology, Technische Universität München, D-80290 Munich
`muriel.lang at tum.de`

understanding the GP is first examined only on rotations before it is extended to full 6D rigid motions.

A. GP Definition

A GP is a collection of random variables, which are jointly Gaussian distributed. It can be understood as a generalization of a Gaussian distribution to an infinitely number of variables. Instead of having a defining mean vector $\mu \in \mathbb{R}^n$ and covariance matrix $\Sigma \in M^{n \times n}$, $n \in \mathbb{N}$, the GP is fully specified by a mean function $m(\mathbf{x})$ and a covariance function $k(\mathbf{x}, \mathbf{x}')$. The covariance function describes the dependencies between function values for any input pair $(\mathbf{x}, \mathbf{x}')$ and defines the elements of the kernel matrix $\mathbf{K}$ by $(\mathbf{K})_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ [9]. Hence, a GP is defined through

$$f(x) = \mathrm{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')). \tag{1}$$

In practice, the mean function is often set to zero, as this simplifies calculation without limiting the expressive power of the process. The essential part in GP model learning is the kernel function $k(\mathbf{x}, \mathbf{x}')$ and the estimation of its free hyperparameters, such as length-scale λ, or signal variance σ_f. To learn the optimal hyperparameters of the GP, a gradient descent algorithm is used. The squared exponential covariance function

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \, \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\lambda^2}\right), \tag{2}$$

with hyperparameters σ_f, λ, is probably the most widely used kernel function in machine learning [5].

B. 6D Pose Representation by Dual Quaternions

A unit quaternion $\mathbf{q}_r = q_w + q_x i + q_y j + q_z k \in \mathbb{H}$, with parameter entries $q_w, q_x, q_y, q_z \in \mathbb{R}$ such that $\|\mathbf{q}_r\| = 1$ is used to represent a 3D orientation. An imaginary quaternion $\mathbf{q}_t = q'_x i + q'_y j + q'_z k$, with $q'_x, q'_y, q'_z \in \mathbb{R}$ is used to represent a 3D position. Together, the quaternions define a 6D pose and a rigid motion $(\mathbf{q}_r, \mathbf{q}_t)$. The representations

$$SE(3) \simeq (\mathbf{q}_r, \mathbf{q}_t) \simeq S_3 \times \mathbb{R}^3 \tag{3}$$

are isomorphic for rigid motions.

The ring of the dual quaternions is defined as

$$\mathbb{H}_D = \{\, \mathbf{dq} \mid \mathbf{dq} = \mathbf{q}_{\mathrm{re}} + \varepsilon\, \mathbf{q}_{\mathrm{du}} \,\&\, \mathbf{q}_{\mathrm{re}}, \mathbf{q}_{\mathrm{du}} \in \mathbb{H}\}, \tag{4}$$

where ε is a dual unit, which holds $\varepsilon^2 = 0$ [10]. The dual quaternion representing the pose $(\mathbf{q}_r, \mathbf{q}_t)$, is calculated by

$$\mathbf{dq} := \mathbf{q}_r + \frac{\varepsilon}{2}\, \mathbf{q}_t\, \mathbf{q}_r. \tag{5}$$

To transform a point $\mathbf{p} = (p_x, p_y, p_z)^\top \in \mathbb{R}^3$ by a rigid motion parameterized by a dual quaternion $\mathbf{dq}$, the point needs to be represented by an imaginary quaternion $\mathbf{q}_{\mathbf{p}} = p_x i + p_y j + p_z k$. Then the transformed point $\mathbf{p}'$ is represented by the imaginary quaternion

$$\mathbf{q}_{\mathbf{p}'} = \mathbf{dq} * \mathbf{q}_{\mathbf{p}} * \overline{\mathbf{dq}}, \tag{6}$$

where $*$ is the dual quaternion multiplication and $\overline{\mathbf{dq}} = \overline{\mathbf{q}}_{\mathrm{re}} + \varepsilon \overline{\mathbf{q}}_{\mathrm{du}}$ is the conjugate of $\mathbf{dq}$ and $\overline{\mathbf{q}}$ is the quaternion conjugate of $\mathbf{q}$ [11].

C. Gaussian Process over Pure Rotations

A GP over pure rotations parametrized by unit quaternions is introduced in this section. We subdivide into explaining the GP mapping from input to output space, briefly introducing the mean function of the GP over the unit sphere S_3, before we specify the metrics over unit quaternions, which we use for the kernel function. Finally we compare the kernels to highlight their differences.

1) Input Space to Output Space Mapping: The training data we introduce for learning the GP mapping from input to output space, consists of unit quaternion samples $\mathbf{q}$ on the sphere S_3 with corresponding velocities. Each rotational velocity is represented by a vector $\mathbf{v}_{\mathrm{TS}}$ in the tangent space $\mathrm{TS}_{\mathbf{q}}$ of the corresponding rotation quaternion $\mathbf{q}$, i.e. each 3D velocity vector is defined in its tangent space to the unit sphere S_3 with respective basis. Hence, the Gaussian distribution over velocities is in a Euclidean tangent space and we avoid the need to define a Gaussian on the hypersphere S_3. The velocity in the tangent space $\mathrm{TS}_{\mathbf{q}}$ can be equivalently represented by a unit quaternion $\dot{\mathbf{q}}$, using the central projection of $\mathbf{v}_{\mathrm{TS}}$ to the sphere, $\mathbf{q}' = \mathbf{q}\dot{\mathbf{q}}$. In the GP prediction the velocity vector is projected to the sphere to obtain the predicted orientation as a unit quaternion. The central projection from the tangent space $\mathrm{TS}_{\mathbf{q}}$ of the quaternion $\mathbf{q}$, to the hypersphere is defined as

$$\Pi_{\mathbf{q}} : \mathrm{TS}_{\mathbf{q}} \to S_3$$
$$\Pi_{\mathbf{q}}(\mathbf{v}_{\mathrm{TS}}) = \left\{ -\frac{\mathbf{v}}{\|\mathbf{v}\|}, \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\}, \tag{7}$$

where $\mathbf{v} = \mathbf{q} + \mathbf{B}\mathbf{v}_{\mathrm{TS}}$. The basis $\mathbf{B}$ is the canonical representation of the 3D tangent space $\mathrm{TS}_{\mathbf{q}}$ in the 4D space $\mathbb{R}^4$ [12]. With this representation we avoid cumbersome learning restrictions in the GP, to assure the output to be a unit quaternion.

2) Mean Functions: As well as ordinary GP, the GP over quaternions is fully specified by a mean function $m(\mathbf{x})$ and a covariance function $k(\mathbf{x}, \mathbf{x}')$. The mean function is defined by orientations and is therefore restricted to the unit hypersphere, $m : S_3 \to S_3$. As the mean function in practice is often set to zero rotation, i.e. $m(\mathbf{q}) = (1, 0, 0, 0)^\top$, we focus on distance functions.

3) Unit Quaternion Metrics: We introduce three metrics over points on the unit hypersphere S_3, which are visualized in Fig. 1. The first two metrics depend on the scalar product over quaternions, which is defined as

$$\langle \mathbf{q}, \mathbf{q}' \rangle = q_w q'_w + q_x q'_x + q_y q'_y + q_z q'_z = \mathbf{q}\overline{\mathbf{q}'}. \tag{8}$$

The unit quaternions are a double coverage of the rotations, so opposite quaternions represent the same orientation. To avoid ambiguity, both quaternions $\mathbf{q}, \mathbf{q}'$ are chosen to lay on the same hemisphere, such that the scalar product is nonnegative $\langle \mathbf{q}, \mathbf{q}' \rangle \geq 0$. Firstly, the quaternion norm $\|\mathbf{q} - \mathbf{q}'\|$ is a good approximation of the distance between similar rotations. It yields the same result as the regular vector norm of points $\forall \mathbf{q}, \mathbf{q}' \in S_3$

$$d_{\mathrm{norm}}(\mathbf{q}, \mathbf{q}') = \sqrt{\mathbf{q}\overline{\mathbf{q}'}} = \|\mathbf{q} - \mathbf{q}'\|. \tag{9}$$
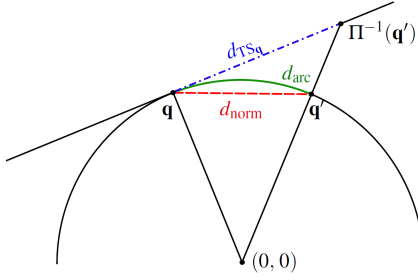
Fig. 1. The distance measures consisting of quaternion norm, arc length and projection to the tangent space are depicted with a red dashed, green solid and blue dashed-dotted line.

This distance measure is calculated in the Euclidean space $\mathbb{R}^4$ and is not part of the sphere S_3. Secondly, an alternative distance measure between unit quaternions $\mathbf{q}$ and $\mathbf{q}'$ is the arc length between the quaternions. It is obtained by

$$d_{\mathrm{arc}}(\mathbf{q}, \mathbf{q}') = \arccos(\langle \mathbf{q}, \mathbf{q}' \rangle) = \arccos(\mathbf{q}\overline{\mathbf{q}'}). \quad (10)$$

The advantage of distance d_{arc} over d_{norm} is, that the arc is a subset of S_3. This means, the arc length is a distance measure on the unit sphere and it precisely measures how similar orientations are instead of approximating it. Thirdly, we determine the distance between unit quaternions $\mathbf{q}, \mathbf{q}'$ in the tangent space. One quaternion, for instance $\mathbf{q}'$ is projected to the other's tangent space $\mathrm{TS}_\mathbf{q}$ by central projection (7). The quaternion $\mathbf{q}$ is represented by the origin of $\mathrm{TS}_\mathbf{q}$. Then the tangent space distance $d_{\mathrm{TS}_\mathbf{q}}(\mathbf{q}, \mathbf{q}')$ between the projected quaternions is the norm of $\Pi_\mathbf{q}^{-1}(\mathbf{q}')$ in the Euclidean space $\mathrm{TS}_\mathbf{q} \simeq \mathbb{R}^3$,

$$d_{\mathrm{TS}_\mathbf{q}}(\mathbf{q}, \mathbf{q}') = \|\Pi_\mathbf{q}^{-1}(\mathbf{q}')\| = \|\mathbf{q}'_{\mathrm{TS}_\mathbf{q}}\|. \quad (11)$$

As this distance measure is symmetric, it holds $d_{\mathrm{TS}_\mathbf{q}}(\mathbf{q}, \mathbf{q}') = d_{\mathrm{TS}_{\mathbf{q}'}}(\mathbf{q}', \mathbf{q}) \forall \mathbf{q}, \mathbf{q}' \in S_3$. If the quaternions represent opposite orientations, i.e. the quaternions are perpendicular, the projected quaternion hits the tangent space at the point ∞. Therefore, we include the compact boundary $\overline{\mathbb{R}}^3 = \mathbb{R}^3 \cup \{\infty\}$ in the tangent space, to obtain a continuous completion of the distance $d_{\mathrm{TS}_\mathbf{q}}$ on $\mathbb{R} \cup \{\infty\}$. Choosing any point $\mathbf{p} \neq \mathbf{q}, \mathbf{q}'$ between the quaternions $\mathbf{q}$ and $\mathbf{q}'$ as the tangent point would yield a bounded distance measures $d_{\mathrm{TS}_\mathbf{p}}$, like $d_{\mathrm{norm}}, d_{\mathrm{arc}}$ are bounded. The maximal distance occurring for 2π rotations is $\sqrt{2}$ in the quaternion norm measure d_{norm} and $\pi/2$ in the arc length measure d_{arc}. The distance measures $d_{\mathrm{norm}}, d_{\mathrm{arc}}$ and $d_{\mathrm{TS}_\mathbf{q}}$ define metrics on the unit sphere S_3, as they all are non-negative, symmetric and fulfill the triangle inequality.

4) Covariance Functions: Over the metrics from the previous subsection, covariance functions on S_3 are defined. For simplicity we focus linear kernel and the family of squared exponential kernels in the present study. The definition of a linear kernel over quaternions follows straight forward from the scalar product over quaternions (8),

$$k_{\mathrm{lin}}(\mathbf{q}, \mathbf{q}') := \langle \mathbf{q}, \mathbf{q}' \rangle. \quad (12)$$

The kernel functions over unit quaternions corresponding to the squared exponential kernel (2) provide high correlations

between close unit quaternions and, smoothly descending, smaller correlations for the far ones. From the d_{norm} (9) we obtain the regular squared exponential kernel

$$k_{\mathrm{norm}}(\mathbf{q}, \mathbf{q}') = \sigma_f^2 \exp\left(-\frac{\|\mathbf{q} - \mathbf{q}'\|^2}{2\lambda^2}\right), \quad (13)$$

where the quaternions $\mathbf{q}, \mathbf{q}'$ can be regarded as vectors in the $\mathbb{R}^4$. From the square arc length (10) and the linear kernel (12) we construct a new kernel function

$$k_{\mathrm{arc}}(\mathbf{q}, \mathbf{q}') = \sigma_f^2 \exp\left(-\frac{(\arccos^2(\langle \mathbf{q}, \mathbf{q}' \rangle))}{2\lambda^2}\right). \quad (14)$$

That k_{arc} defines a kernel function which can be proven by induction. With the metric over projected points in the tangent space (11), a squared exponential kernel function is defined through

$$k_{\mathrm{TS}}(\mathbf{q}, \mathbf{q}') = \sigma_f^2 \exp\left(-\frac{\|\mathbf{q}'_{\mathrm{TS}_\mathbf{q}}\|^2}{2\lambda^2}\right). \quad (15)$$

The proof that (15) defines a covariance function is analogous to the one for the regular squared exponential kernel (2) in the Euclidean space.

5) Comparison of Covariance Functions: In this paragraph, we investigate the behavior of the covariance functions over unit quaternions, as introduced above. To compare them, the covariances at each time step are calculated between a rotated point and the starting configuration for a one dimensional full turn. In Fig. 2, the kernel func-
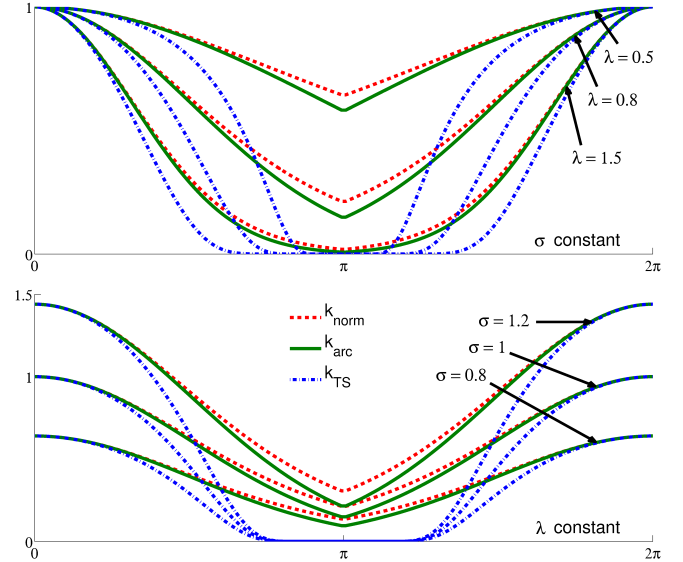

Fig. 2. The full turn about the z-axis is parametrized by quaternions $\mathbf{q}$. Then the function value of the covariance functions k_{norm}, k_{arc} and k_{TS} are calculated for $(0, \mathbf{q})$ and are visualized by the red dashed, blue dashed-dotted, green solid line, respectively.

tions $k_{\mathrm{norm}}(\mathbf{0}, \mathbf{q})$, $k_{\mathrm{arc}}(\mathbf{0}, \mathbf{q})$ and $k_{\mathrm{TS}}(\mathbf{0}, \mathbf{q})$ are visualized for $\mathbf{q} \in \mathbb{H}$ representing the rotation from 0 to 2π about the z-axis. The function values correspond to the first row, respectively first column of the corresponding covariance matrix $\mathbf{K}$. In all cases the covariance decreases, as expected, on rotating to the farthest orientation π and symmetrically

increases as it departs from π. In the upper part, the covariances are depicted three times for fixed signal noise σ_f and varying length scale λ. In the lower part, the length scale is fixed, while the signal variance varies. The steep decrease of kernel k_{TS}, visualized by a blue dashed-dotted line, compared to the kernels k_{norm} and k_{arc}, depicted by the red dashed and green solid line, encodes the fast increase of the underlying metric d_{TS_q}. Hence, high correlations are obtained only for highly similar orientations.

The unit sphere S_3 can be combined with any Euclidean space $\mathbb{R}^n$ using the Cartesian set product. This offers a straight forward method to perform GP regression over spaces that consist of rotation and translation, translational velocities, forces etc. However, the Cartesian set product neglects all correlations between the subsets and a case specific manual weighting between rotation and translation is required. Instead, we propose the GP over dual quaternions for 6D rigid motions, as it nicely captures this correlation.

D. Gaussian Process over 6D Rigid Body Motions

In this section, we introduce a Gaussian Process over $S_3 \times \mathbb{R}^3$. We extend our input data from pure orientations to whole 6D poses, parametrized by dual quaternions $\mathbb{H}_D$. On the special Euclidean group we do not address the mean function $m : SE(3) \to SE(3)$, and it is set to zero.

1) Input Space to Output Space Mapping: The GP is trained on a dataset consisting of input poses and output velocities. The training input is parametrized by dual quaternions, whereas the training output is a concatenation of rotational velocities in the tangent space and translational velocities parametrized by a simple 3D vector. The velocities in the tangent space are projected to the hypersphere S_3 by (7) and with the imaginary quaternion representation for the translational velocities, they can be combined to a dual quaternion velocity using (5). Hence, the dual quaternion velocity represents the rigid motion velocity at a certain 6D pose.

2) Dual Quaternion Distance Measures: We introduce three magnitude measures for dual quaternion differences. Analogous to how we consider the quaternion distance $d_{\text{norm}}(\mathbf{q}, \mathbf{q}') = \Delta \mathbf{q}$ as a 4D vector in the squared quaternion norm, we now consider the dual quaternion $\mathbf{dq} = \mathbf{q}_{\text{re}} + \varepsilon \cdot \mathbf{q}_{\text{du}}$ as a vector tuple $(\mathbf{q}_{\text{re}}, \mathbf{q}_{\text{du}})$. Then, the vector tuple distance $d_{\text{tuple}}(\mathbf{dq}, \mathbf{dq}') = \Delta \mathbf{dq}$ of the dual quaternions $\mathbf{dq}, \mathbf{dq}'$ can be defined as

$$d_{\text{tuple}}(\mathbf{dq}, \mathbf{dq}') = \|\mathbf{q}_{\text{re}} - \mathbf{q}'_{\text{re}}\| + \|\mathbf{q}_{\text{du}} - \mathbf{q}'_{\text{du}}\|. \quad (16)$$

In this distance measure it is disregarded the fact, that the second part of a dual quaternion lies in a dual space. However, the correlation between rotation and translation is captured in the second summand of the distance d_{tuple}. As a second distance measure, we propose the squared dual quaternion norm, since this measure is well-known and captures the duality of the representation

$$d_{\text{dqn}}^2(\mathbf{dq}, \mathbf{dq}') = \|\mathbf{dq} - \mathbf{dq}'\|^2 = \|\Delta \mathbf{dq}\|^2$$
$$= \underbrace{\Delta \mathbf{q}_{\text{re}} \overline{\Delta \mathbf{q}_{\text{re}}}}_{x_{\text{re}}} + \varepsilon (\underbrace{\Delta \mathbf{q}_{\text{re}} \overline{\Delta \mathbf{q}_{\text{du}}} + \Delta \mathbf{q}_{\text{du}} \overline{\Delta \mathbf{q}_{\text{re}}}}_{x_{\text{du}}}) \quad (17)$$

The dual quaternion difference $\Delta \mathbf{dq}$ between two 6D poses $\mathbf{dq}$ and $\mathbf{dq}'$ yields a dual quaternion, which is not representing a rigid motion. Therefore, we obtain a dual number $x_{\text{re}} + \varepsilon x_{\text{du}}$ for d_{dqn}^2. As the dual part x_{du} can be negative, there is no canonical method to extract a real number from the squared dual quaternion norm necessary for assuring the covariance to be positive semi-definite. More information about dual numbers can be found in [13]. The last magnitude measure we introduce, is calculated from the transformation rigid motion $\vec{\mathbf{dq}}$ between dual quaternions. It denotes the rigid motion which has to be applied to the pose $\mathbf{dq}$, to arrive in $\mathbf{dq}'$ and is obtained by

$$\vec{\mathbf{dq}} = \overline{\mathbf{dq}} * \mathbf{dq}'. \quad (18)$$

From the transformation dual quaternion $\vec{\mathbf{dq}} = \vec{\mathbf{q}}_{\text{re}} + \varepsilon \vec{\mathbf{q}}_{\text{du}}$, rotation and translation are obtained by dual quaternion decomposition. We define the squared transformation magnitude measure as

$$d_{\text{mag}}^2(\mathbf{dq}, \mathbf{dq}') = \|\vec{\mathbf{q}}_r\|^2 + \|\vec{\mathbf{q}}_t\|^2 + c\|\vec{\mathbf{q}}_r \vec{\mathbf{q}}_t\|, \quad (19)$$

where $\vec{\mathbf{q}}_r$ is the unit quaternion for rotation, $\vec{\mathbf{q}}_t$ the imaginary quaternion for translation and $c \geq 0$ a nonnegative constant. The quaternion norm (9) provides real positive numbers, which are combined to capture the correlation between rotation and translation. In the transformation dual quaternion the correlation of rotation and translation is already captured, such that there is no need to scale $\vec{\mathbf{q}}_r$ against $\vec{\mathbf{q}}_t$. The factor c defines how strong the composed part is weighted in comparison to pure rotation and translation.

3) Covariance Functions: The vector tuple distance (16) allows us to define a linear kernel on dual quaternions $k_{\text{lin}}(\mathbf{dq}, \mathbf{dq}') = d_{\text{tuple}}^2(\mathbf{dq}, \mathbf{dq}')$. Over dual quaternions no scalar product is defined, but over quaternion tuples $\mathbf{dq} = (\mathbf{q}_{\text{re}}, \mathbf{q}_{\text{du}})$ and $\mathbf{dq}' = (\mathbf{q}'_{\text{re}}, \mathbf{q}'_{\text{du}})$, the analogon to a scalar product is

$$d_{\text{tuple}}^2(\mathbf{dq}, \mathbf{dq}') = \langle \mathbf{q}_{\text{re}}, \mathbf{q}'_{\text{re}} \rangle + \langle \mathbf{q}_{\text{du}}, \mathbf{q}'_{\text{du}} \rangle$$
$$+ \|\mathbf{q}_{\text{re}} - \mathbf{q}'_{\text{re}}\| \|\mathbf{q}_{\text{du}} - \mathbf{q}'_{\text{du}}\|. \quad (20)$$

We obtain a squared exponential kernel function, by inserting the squared transformation magnitude measure (19) in the exponential function

$$k_{\text{mag}}(\mathbf{dq}, \mathbf{dq}') = \sigma_f^2 \exp\left(-\frac{d_{\text{mag}}^2(\mathbf{dq}, \mathbf{dq}')}{2\lambda^2}\right). \quad (21)$$

Figure 3 illustrates the behavior of the squared exponential kernel k_{mag} on dual quaternions, representing simultaneously constant rotation and translation. In x-direction constant rotation about a single axis from 0 to 2π is applied, while in y-direction constant translation in one direction is applied. The value of the covariance function $k_{\text{mag}}(\mathbf{dq}', \mathbf{dq})$ is color coded for $\mathbf{dq}'$ representing the origin $(0,0)$, i.e. $\mathbf{dq}' = (1,0,0,0)^\top + \varepsilon(0,0,0,0)^\top$ and $\mathbf{dq}$ representing the rotations and translations in the grid. The highest correlations are obtained for points, where both, rotation and translation are small, i.e. just a small rigid motion has to be applied to move from $\mathbf{dq}'$ to $\mathbf{dq}$. The color
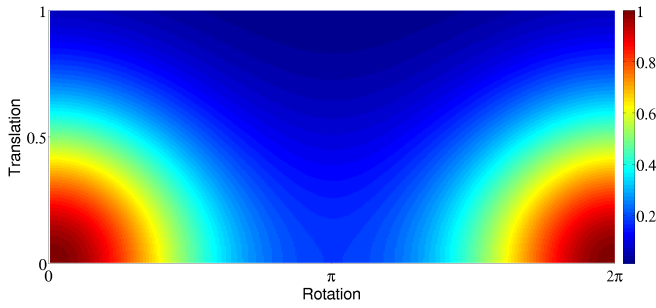
Fig. 3. In x-direction a full turn about one axis is shown and in y-direction a constant translation from 0 to 1. Each point in the plane is converted to a dual quaternion **dq**, representing the rotation x and translation y. The hyperparameters $(\lambda, \sigma_f) = (1, 1)$ and $c = 0$. The value of the covariance function $k_{\mathrm{mag}}(0, \mathbf{dq})$ is visualized by the color code of the bar on the right side.

values are symmetric, as the distance on the hypersphere S_3 decreases on both sides of π.

III. EXPERIMENTAL EVALUATION

In this section, we evaluate the GP over 6D rigid body motions. Firstly, the performance of the proposed squared exponential kernels for pure orientation is compared with a GP over Euler angles in a real data experiment. Subsequently, we apply the GP over dual quaternions to an example setup including large rotations, translations and variances, which is poorly handled by an ordinary GP. We use the MATLAB toolbox GPML from C.E. Rasmussen and H. Nickisch, provided on http://gaussianprocess.org/ in our implementation to perform GP regression over dual quaternions. All code is written in MATLAB.

A. Real Data Experiment: 3D Rotation

We estimate the rotation in a flight trajectory of a quadrocopter using real data as ground truth. The 6D pose trajectory of a Crazyflie Nano Quadrotor, visualized by Fig. 4 is measured in a passive marker based motion capture system (Qualisys, Sweden) at 10Hz. The GP is trained on 7%
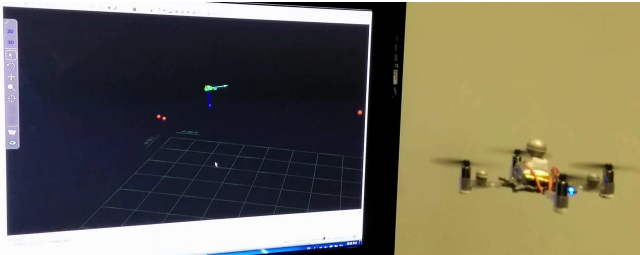

Fig. 4. The flight trajectory of the Crazyflie Nano Quatrotor captured in a motion capture system is used to compare GPs using Quaternions with GPs using Euler angles.

random samples of 150 poses, that are artificially pulled samples from the motion data. The tracking system noise is modeled by white Gaussian noise with standard deviation 0.03. We use exact inference and a Gaussian likelihood and set the hyperparameters $(\sigma_f, \lambda) = (0.08, 0.4)$. Each flight trajectory we was recorded for 15 seconds and this data set contains rotations of more than 2π. The GP regression is

performed over pure 3D rotations of the flight trajectory. We compare the GP over unit quaternions using the three different squared exponential kernels $k_{\mathrm{norm}}, k_{\mathrm{arc}}$ and k_{TS} with a GP over Euler angles (ϕ, ψ, θ) treating the angles like a vector in $\mathbb{R}^3$. From the estimated orientation trajectory exemplary one signal output is visualized in Fig. 5. We arbitrarily select the fourth component q_z of the rotation quaternion $(q_w, q_x, q_y, q_z)^\top \in S_3$ and visualize it over time using the different underlying metrics $d_{\mathrm{norm}}, d_{\mathrm{arc}}$ and $d_{\mathrm{TS_q}}$. For comparison the corresponding signal yaw θ of the Euler angles is also shown. The predictions using k_{norm} and k_{TS}
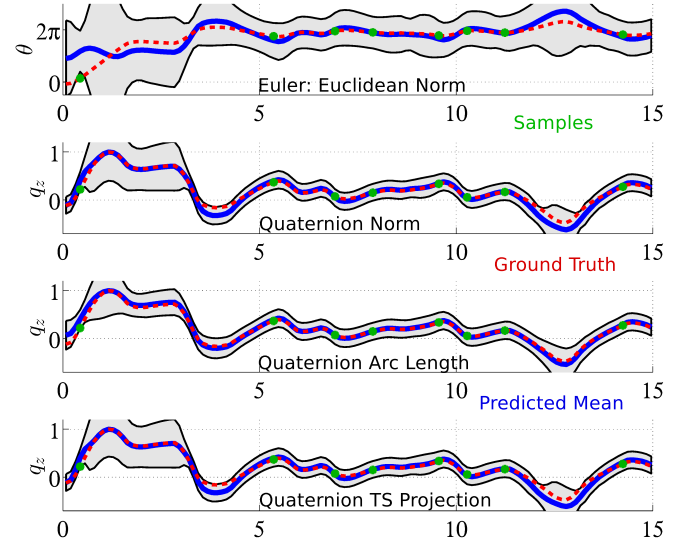

Fig. 5. All signals visualize one DoF of the performed rotation over time. The red dashed line depicts the ground truth rotation, the green stars the samples used for learning and the blue solid line the mean prediction. The first signal shows the yaw rotation of the flight trajectory estimated by a GP over Euler angles using the Euclidean norm in the kernel. The other signals each depict the fourth entry of the GP prediction over unit quaternions using the kernels $k_{\mathrm{norm}}, k_{\mathrm{arc}}$ and k_{TS}. The gray shaded area shows an upper bound for the sextuple standard deviation of the prediction.

resemble each other, as the underlying metrics are similar, those using k_{arc} are more accurate for large rotations. As the underlying metric of k_{TS} precisely measures the orientation distance, mean prediction is less erroneous for instance in second 13. All GPs over unit quaternions on S_3 outperform the GP over Euler angles. Even though the yaw angle θ is unwrapped to avoid the jump at 2π, the predictions close to large rotation fall behind any prediction using quaternions. Further, the covariance encoding the prediction uncertainty is significantly bigger.

B. Simulated Data Experiment: 6D Rigid Motion

In this section we apply the GP over dual quaternions to an example setup incorporating large rotations and translations. The aim of this experiment is to learn a collision free 6D motion trajectory of a stylus from random starting configurations to a fixed goal configuration inside a hole. A set of 500 trajectories to sufficiently cover the 6D space is generated by the Open Motion Planning Library (OMPL) [14] using the path planner RRTconnect with path simplification and smoothing. Each trajectory consists of about 200 to 300

sequential poses visualized in Fig. 6. The GP over 6D rigid motions is trained with 20% randomly chosen sample poses from 108406 poses in total. The prediction is performed
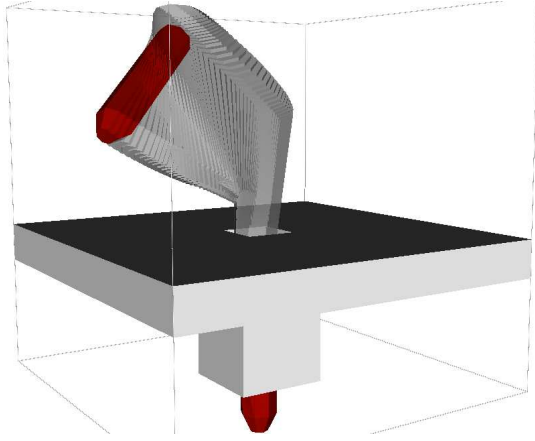

Fig. 6. Artificially generated 6D rigid motion trajectory of a longish red object from a random starting configuration to the goal configuration inside the hole.

stepwise and we iteratively apply the predicted velocity to the current pose. To save computational effort, the prediction is based on the set of $n \in \mathbb{N}$ closest poses according to the transformation magnitude measure (19). The starting configurations are restricted to the 6D space covered by the training trajectories. This assures positive correlation with training poses required by the GP prediction. The experiment is considered successful, when the object maneuvers into the hole without collision. In Fig. 7, an initial random configuration is visualized by a large coordinate system. The smaller coordinate systems represent a subset of training trajectory starting poses. The predicted motion trajectory correctly follows the training data and arrives at the goal pose inside the hole without collision.
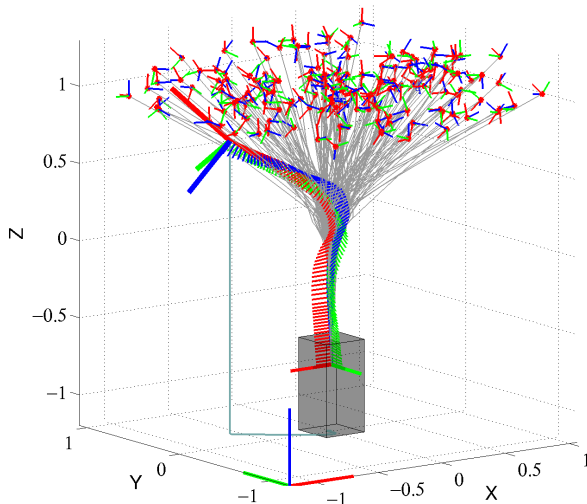

Fig. 7. A successful 6D rigid body motion prediction is visualized. The object moves from a random starting configuration to a goal configuration. The regression is performed, using GP over dual quaternions with the kernel k_{mag}.

IV. CONCLUSION

In this paper we introduce the GP over dual quaternions. The essential contribution is the extension of the input data of the GP to 6D rigid body motions using the squared exponential kernel k_{mag} defined over dual quaternions. We introduce and compare three metrics on the hypersphere that induce the kernels k_{norm}, k_{arc} and k_{TS} for pure rotations. GP regression is performed for a task involving large rotations, which is poorly handled by a regular GP using Euler angles, while the GP using dual quaternions can successfully overcome the problem and arrives in the goal configuration without collision. For future work we plan to predict human behavior in human-robot interaction using the GP over dual quaternions, as for instance human gaze motion trajectories include large rotations.

REFERENCES

[1] J. M. Wang, D. J. Fleet, S. Member, and A. Hertzmann, "Gaussian process dynamical models for human motion," *IEEE Trans. Pattern Anal. Machine Intell*, 2007.

[2] S. M. Khansari Zadeh and A. Billard, "Learning and control of uav maneuvers based on demonstrations." Presented at Robotics Science and Systems, Seattle, June 28 July 1, 2009., 2009.

[3] D. Kulić, W. Takano, and Y. Nakamura, "Incremental learning of full body motions via adaptive factorial hidden markov models," in *in 7th International Conference on Epigenetic Robotics*, 2007.

[4] J. M. Hernndez, M. Lawitzky, A. Mrtl, D. Lee, and S. Hirche, "An experience-driven robotic assistant acquiring human knowledge to improve haptic cooperation," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011.

[5] C. Rasmussen and C. Williams, *Gaussian processes for machine learning*. Adaptative computation and machine learning series, University Press Group Limited, 2006.

[6] J. Ko and D. Fox, "GP-BayesFilters: Bayesian filtering using Gaussian process prediction and observation models," in *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pp. 3471–3476, 2008.

[7] S. Kim and A. Billard, "Estimating the non-linear dynamics of free-flying objects.," *Robotics and Autonomous Systems*, vol. 60, no. 9, pp. 1108–1122, 2012.

[8] M. Jammalamadaka, J and A. Sengupta, *Topics in Circular Statistics*. Series on Multivariate Analysis (Book 5), Singapore: World Scientific Pub Co Inc; Har/Dskt edition (March 15, 2001), March 2001.

[9] C. E. Rasmussen, "Gaussian processes for machine learning," MIT Press, 2006.

[10] E. Ata and Y. Yayli, "Dual unitary matrices and unit dual quaternions," *Differential Geometry - Dynamical Systems*, vol. 10, pp. 1–12, 2008.

[11] W. Feiten, M. Lang, and S. Hirche, "Rigid Motion Estimation using Mixtures of Projected Gaussians," 16th International Conference on Information Fusion, 2013.

[12] M. Lang and W. Feiten, "MPG - Fast Forward Reasoning on 6 DOF Pose Uncertainty," in *7th German Conference on Robotics (ROBOTIK 2012)*, (Munich, Germany), May 2012.

[13] F. W. B. Vasantha Kandasamy, *Dual Numbers*. American Research Press.

[14] I. A. Şucan, M. Moll, and L. E. Kavraki, "The Open Motion Planning Library," *IEEE Robotics & Automation Magazine*, vol. 19, pp. 72–82, December 2012.