

# DUAL QUATERNION ROTATIONAL AND TRANSLATIONAL EQUIVARIANCE IN 3D RIGID MOTION MODELLING

Guilherme Vieira\*, Eleonora Grassucci†, Marcos Eduardo Valle\*, Danilo Comminiello†

\*Dept. of Applied Mathematics, Universidade Estadual de Campinas (Unicamp), Brazil

†Dept. of Information Engineering, Electronics, and Telecomm., Sapienza University of Rome, Italy

## ABSTRACT

Objects' rigid motions in 3D space are described by rotations and translations of a highly-correlated set of points, each with associated  $x, y, z$  coordinates that real-valued networks consider as separate entities, losing information. Previous works exploit quaternion algebra and their ability to model rotations in 3D space. However, these algebras do not properly encode translations, leading to sub-optimal performance in 3D learning tasks. To overcome these limitations, we employ a dual quaternion representation of rigid motions in the 3D space that jointly describes rotations and translations of point sets, processing each of the points as a single entity. Our approach is translation and rotation equivariant, so it does not suffer from shifts in the data and better learns object trajectories, as we validate in the experimental evaluations. Models endowed with this formulation outperform previous approaches in a human pose forecasting application, attesting to the effectiveness of the proposed dual quaternion formulation for rigid motions in 3D space.

**Index Terms**— Dual Quaternions, Rigid Motions, Translation and Rotation Equivariance, Human Pose Forecasting.

## 1. INTRODUCTION

Human pose forecasting (HPF) aims to predict the position of human keypoints – such as head, limbs, torso, etc – in the near future. Accurately estimating the position of human joints and their orientations allows for a better understanding of individuals moving in 3D space. By leveraging this understanding, it is possible to predict the motions of individuals in the immediate future, which is invaluable for applications in self-driving cars [1, 2], home healthcare [3], and security surveillance [4], for example. The estimated keypoints can also be used to faithfully reconstruct 3D human body models, having therefore notable applications in augmented and virtual reality [5].

HPF data usually comprise a set of joint data points representing the poses in the space. Such kind of data is inherently a multichannel signal since positions are represented by their coordinates in 3D space. The sequence of positions of a joint

in several instants is therefore a time series with highly correlated components. Moreover, joints are part of the same body so they are also correlated among themselves, especially those located in the same limb. Higher dimensional algebras can be used to represent HPF data preserving such correlations.

Learning representations of multichannel data using higher dimension algebras has been proved to outperform real-valued approaches in several problems [6, 7, 8, 9]. Quaternion-valued networks already boast works on a significant array of diverse tasks such as audio signal and image processing, computer vision, and dynamic system modeling [10, 11, 12, 13]. Human pose estimation using quaternion-valued networks has been done in a few different manners [14, 15, 16, 17]. Quaternions are often used since they are able to encapsulate multiple spatial coordinates in one entity and also efficiently represent rotations in 3D space [18]. However, the quaternion product is limited to describing rotations around axes containing the origin. Such operations are not robust to translations, which are extremely common transformations in applications to augmented reality, 3D space modeling, and computer vision.

This work considers an elegant representation for jointly modeling rotations and translations of objects in the 3D space using dual quaternions. Thanks to such representation, we can enclose a full rigid motion, i.e., translation and rotation, in a single entity [19], considering object movements in the space as a combination of highly-correlated elements. Differently from previous attempts that only focus on simulated tests [20, 21], we model human skeleton motions in real-world scenarios, showing the crucial role of the dual quaternion representation in learning body translations in space. Moreover, we provide practical proof that our dual quaternion formulation is both translation and rotation equivariant, which are highly desirable properties for all applications involving 3D movement modeling. Thus, neural models described with our approach significantly mitigate losses from translations and rotations in the data. In the experimental evaluation, we show that dual quaternion models outperform their real and quaternion-valued counterparts according to every metric considered, and visually display the translation and rotation equivariance properties of our approach.

This paper is structured as follows: Section 2 briefly reviews quaternions and dual numbers. Section 3 encompasses the definition of dual quaternions, a description of rigid motions, and the translation and rotation equivariance properties. Section 4 presents the experimental validation, also showcasing the equivariance of dual quaternions. Finally, Section 5 contains some concluding remarks.

## 2. OVERVIEW OF QUATERNIONS, DUAL NUMBERS, AND DUAL QUATERNIONS

Quaternions, denoted by  $\mathbb{H}$ , are one of the most well-known hypercomplex algebras. Commonly seen as an extension to complex numbers  $\mathbb{C}$ , this algebra was introduced in 1843 by W. R. Hamilton while searching for a domain that could encapsulate operations in three-dimensional space. A quaternion is a number of the form

$$q = q_W + q_X\hat{i} + q_Y\hat{j} + q_Z\hat{k} \in \mathbb{H}, \quad (1)$$

where  $q_W, q_X, q_Y, q_Z$  are real numbers. The symbols  $\hat{i}, \hat{j}, \hat{k}$  denote the hyperimaginary units and follow the multiplication rules  $\hat{i}^2 = \hat{j}^2 = \hat{k}^2 = \hat{i}\hat{j}\hat{k} = -1$ , known as the Hamilton rules.

A quaternion given by (1) efficiently describes a rotation in 3D space. Formally,  $q_W$  is called the real part and  $\bar{\mathbf{q}} = q_X\hat{i} + q_Y\hat{j} + q_Z\hat{k}$  is referred to as the vector part of  $q$ . Hence, a quaternion can be written as  $q = q_W + \bar{\mathbf{q}}$ . We say that  $q$  is a pure quaternion if  $q_W = 0$ . The conjugate of  $q = q_W + \bar{\mathbf{q}}$  is easily expressed in this notation as  $q^* = q_W - \bar{\mathbf{q}}$ . Moreover, we can write a quaternion in polar form

$$q_\theta = \|q\| \left( \cos\left(\frac{\theta}{2}\right) + \sin\left(\frac{\theta}{2}\right) \bar{\mathbf{u}} \right), \quad (2)$$

for  $\theta \in [0, 2\pi)$  and a pure quaternion  $\bar{\mathbf{u}} = u_X\hat{i} + u_Y\hat{j} + u_Z\hat{k}$ , where  $\|q\| = \sqrt{q_W^2 + q_X^2 + q_Y^2 + q_Z^2}$  is the absolute value of  $q$ . The rotation of a 3D vector  $(p_X, p_Y, p_Z)$  by an angle  $\theta$  around the axis determined by  $(u_X, u_Y, u_Z)$  can be efficiently determined by

$$p_{rot} = qpq^*, \quad (3)$$

where  $p = p_X\hat{i} + p_Y\hat{j} + p_Z\hat{k}$  and  $u = u_X\hat{i} + u_Y\hat{j} + u_Z\hat{k}$  are pure quaternions encoding the position and axis vectors, and  $q$  is the quaternion given by (2).

We would like to remark that, if  $q$  is not a unitary quaternion, then  $p_{rot}$  given by (3) is scaled by  $\|q\|^2$ . Therefore, it is common to normalize  $q$ , that is,  $q \leftarrow \frac{q}{\|q\|}$ . Moreover, the representation in (3) is unique due to the domain of  $\theta$ . Thus, a quaternion uniquely defines a rotation-dilation and vice-versa.

Dual numbers yield a hypercomplex algebra of dimension 2 over  $\mathbb{R}$ . A dual number has the form  $\hat{a} = a_0 + \varepsilon a_\varepsilon$ , where  $a_0, a_\varepsilon \in \mathbb{R}$  and  $\varepsilon$ , called the dual unit, satisfies  $\varepsilon^2 = 0$ . Using

distributivity, we conclude that the product of two dual numbers satisfies  $(a_0 + \varepsilon a_\varepsilon)(b_0 + \varepsilon b_\varepsilon) = a_0b_0 + \varepsilon(a_0b_\varepsilon + a_\varepsilon b_0)$ . In particular,  $(\varepsilon a_\varepsilon)(\varepsilon b_\varepsilon) = 0$  even if  $a_\varepsilon, b_\varepsilon \neq 0$ .

The so-called dual quaternions, denoted by  $\mathbb{D}$ , are quaternions whose components are dual numbers. In mathematical terms, a dual quaternion is given by  $\hat{\mathbf{q}} = \hat{q}_W + \hat{q}_X\hat{i} + \hat{q}_Y\hat{j} + \hat{q}_Z\hat{k}$ , where  $\hat{q}_W, \hat{q}_X, \hat{q}_Y, \hat{q}_Z$  are dual numbers. Interestingly, the dual unit  $\varepsilon$  commutes with the quaternion hyperimaginary units  $\hat{i}, \hat{j}, \hat{k}$ . Thus, the set of dual quaternions  $\mathbb{D}$  can be defined equivalently as dual numbers in which each part is a quaternion, i.e., a dual quaternion can be represented by  $\hat{\mathbf{q}} = (q_0 + \varepsilon q_\varepsilon)$ , where  $q_0, q_\varepsilon \in \mathbb{H}$ . The norm of a dual quaternion is defined by  $\|\hat{\mathbf{q}}\| = \sqrt{\|q_0\|^2 + \|q_\varepsilon\|^2}$  and  $\hat{\mathbf{q}}$  is a unit dual quaternion if  $\|\hat{\mathbf{q}}\| = 1$ .

In this work, we will represent dual quaternions as bold-face letters with a hat  $\hat{\mathbf{q}}$ , while dual numbers will be regular letters with hat  $\hat{q}$ .

## 3. DUAL QUATERNION REPRESENTATION OF RIGID MOTIONS: EQUIVARIANCE PROPERTIES

Any rigid motion in 3D space can be reduced to a rotation-translation with respect to a screw axis  $\vec{h}$ . Using a proper dual quaternion representation, we can characterize a full rigid motion as a unique entity and leverage algebra properties to build a more robust model [19].

Let us take a unit dual quaternion  $\hat{\mathbf{q}} = \hat{q}_W + \hat{q}_X\hat{i} + \hat{q}_Y\hat{j} + \hat{q}_Z\hat{k}$ . On the one hand, the real part  $\hat{q}_W$  of  $\hat{\mathbf{q}}$  yields a dual number  $\hat{\theta} = \arccos(\hat{q}_W) = \frac{\theta}{2} + \varepsilon \frac{s}{2}$ , with  $\theta \in [0, 2\pi)$ . Here,  $\frac{\theta}{2}$  is the rotation angle around the screw axis  $\vec{h}$  and  $\frac{s}{2}$  is the translation distance along that axis [19]. On the other hand, the imaginary part of  $\hat{\mathbf{q}}$  contains information regarding the direction of  $\vec{h}$ , which consists of a vector through the origin  $\vec{u}$  translated by a vector  $\vec{d}$ . Formally, we have

$$\hat{\mathbf{q}} = \hat{q}_W + \hat{q}_X\hat{i} + \hat{q}_Y\hat{j} + \hat{q}_Z\hat{k},$$

where

$$\begin{aligned} \hat{q}_W &= \cos\frac{\theta}{2} - \frac{\varepsilon}{2} \vec{u} \cdot \vec{d} \sin\frac{\theta}{2} \\ \hat{q}_X &= u_x \sin\frac{\theta}{2} + \frac{\varepsilon}{2} \left[ d_x \cos\frac{\theta}{2} - \sin\frac{\theta}{2} (u_y d_z - u_z d_y) \right] \\ \hat{q}_Y &= u_y \sin\frac{\theta}{2} + \frac{\varepsilon}{2} \left[ d_y \cos\frac{\theta}{2} - \sin\frac{\theta}{2} (u_z d_x - u_x d_z) \right] \\ \hat{q}_Z &= u_z \sin\frac{\theta}{2} + \frac{\varepsilon}{2} \left[ d_z \cos\frac{\theta}{2} - \sin\frac{\theta}{2} (u_x d_y - u_y d_x) \right] \end{aligned} \quad (4)$$

Here,  $u = u_x\hat{i} + u_y\hat{j} + u_z\hat{k}$  and  $d = d_x\hat{i} + d_y\hat{j} + d_z\hat{k}$  are purely imaginary quaternions, and  $\vec{u} \cdot \vec{d} = u_x d_x + u_y d_y + u_z d_z$  on the first equation denotes the dot product.

Equation (4) highlights an interesting property: the dual part is responsible for the translation of the rigid motion. Indeed, if there is no translation the vector representing the

translation of the screw axis is  $\vec{d} = 0\hat{i} + 0\hat{j} + 0\hat{k}$ . By setting  $d_x = d_y = d_z = 0$  in the above equations we find that the dual part becomes null, and the resulting equation is equivalent to (2). This is a straightforward fact: a rigid motion with no translation is simply a rotation around an axis that contains the origin, a motion fully described by a quaternion.

On the other hand, if we take a rigid motion without rotation, i.e.,  $\theta = 0$  then (4) yields  $\widehat{\mathbf{q}}_{\mathbf{t}} = 1 + \frac{\varepsilon}{2}\vec{d}$ . This further reinforces that the dual part is responsible for the translation and shows that the imaginary part of the non-dual part is associated to the rotation operation. This elegant formulation puts in evidence the role played by the dual part and formally shows why dual quaternions can encapsulate both operations while quaternions are restricted to representing rotations.

Lastly, a unit dual quaternion  $\widehat{\mathbf{q}}$  can be written as

$$\widehat{\mathbf{q}} = \cos\left(\frac{\widehat{\theta}}{2}\right) + \widehat{\mathbf{h}} \sin\left(\frac{\widehat{\theta}}{2}\right), \quad (5)$$

where  $\widehat{\mathbf{h}}$  is a unit dual quaternion with zero scalar part. Note that this equation is similar to the quaternion polar form (2), except that  $\widehat{\theta} = \theta_0 + \varepsilon\theta_\varepsilon$  is a dual number and  $\widehat{\mathbf{h}} = h_0 + \varepsilon h_\varepsilon$  is a dual quaternion. Indeed, this representation yields the screw motion parameters directly:  $\theta_0/2$  is the rotation angle around the axis defined by  $h_0$  and  $\theta_\varepsilon/2$  is the translation along that same axis.  $h_\varepsilon$  is the so-called moment of the axis and provides an unambiguous representation of the axis in space. It is defined as  $h_\varepsilon = \vec{p} \times h_0$  where  $\vec{p}$  is a vector from the origin pointing to any point of the axis  $h_0$ . We note that any choice of vector yields the same moment, since  $(\vec{p} + \alpha h_0) \times h_0 = \vec{p} \times h_0 = h_\varepsilon, \forall \alpha \in \mathbb{R}$ . This illustrates one very desirable property: dual quaternions express rigid motions based on an unambiguous representation of the screw axis [22] and thus are independent of the coordinate system, therefore equivariant, as opposed to quaternions whose rotations around axes containing the origin are deformed by translations.

## 4. EXPERIMENTAL EVALUATION

### 4.1. Lorenz System: Equivariance Properties

The Lorenz system is an ordinary differential equations (ODE) system that describes the movement of a free particle in atmospheric domain effects. This movement is characterized by rigid motions since the particle is under effects of translation at all times with its direction constantly rotating. The equations system is

$$\begin{cases} \frac{dx}{dt} = \sigma(y - x), \\ \frac{dy}{dt} = x(\rho - z) - y, \\ \frac{dz}{dt} = xy - \beta z, \end{cases} \quad (6)$$

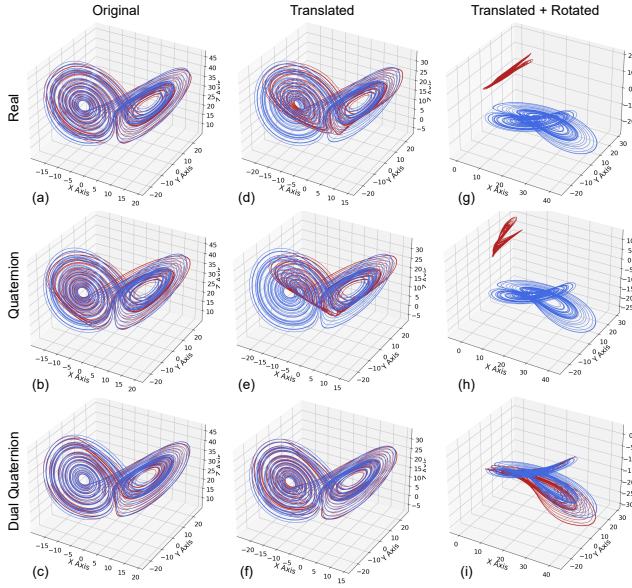
where  $\sigma, \beta, \rho$  are constants. This system exhibits chaotic behavior for certain values of these constants, such as  $\sigma = 10$ ,  $\beta = 8/3$ , and  $\rho = 28$ , meaning that a slight deviation in the initial position results in a large deviation in the particle trajectory. Using these constant values we generate a time series of  $10k$  consecutive positions, 10% of which are used for training, and the remaining 90% are used for testing. By strongly limiting the size of the training set we ensure that the performance of the network is more closely related to how well it learns the rigid motions rather than overfitting.

We employ single-hidden-layer MLPs with ReLU activation in the hidden layer and identity in the output layer in this example. The networks were trained for a one-step-ahead prediction task using a 2-step sliding window, i.e., the network received as inputs two consecutive positions and the desired output is the immediate third one. The inputs and outputs for the real-, quaternion-, and dual quaternion-valued models are formatted as follows:

- **Real-Valued Model:** 6 input values containing the 2 positions as  $(x_{t-1}, y_{t-1}, z_{t-1}, x_t, y_t, z_t)$ , 3 output values as  $(x_{t+1}, y_{t+1}, z_{t+1})$ .
- **Quaternion-Valued Model:** 2 quaternion inputs, one for each position, as  $(0 + x_{t-1}\hat{i} + y_{t-1}\hat{j} + z_{t-1}\hat{k}, 0 + x_t\hat{i} + y_t\hat{j} + z_t\hat{k})$ , 1 output quaternion as  $(0 + x_{t+1}\hat{i} + y_{t+1}\hat{j} + z_{t+1}\hat{k})$ .
- **Dual Quaternion-Valued Model:** 1 dual quaternion input containing the 2 positions as  $(0 + x_{t-1}\hat{i} + y_{t-1}\hat{j} + z_{t-1}\hat{k}) + \varepsilon(0 + x_t\hat{i} + y_t\hat{j} + z_t\hat{k})$ , one dual quaternion output  $(0 + x_{t+1}\hat{i} + y_{t+1}\hat{j} + z_{t+1}\hat{k}) + \varepsilon(0 + 0\hat{i} + 0\hat{j} + 0\hat{k})$ .

The real network is an MLP with a 6-128-3 architecture, i.e., 6 inputs, 128 neurons in the hidden layer, 3 units in the output layer. Analogously, the quaternion model is an MLP with a 2-80-1 architecture, and the dual quaternion model is an MLP with a 1-53-1 architecture. Networks architectures were chosen to feature an equal number of total trainable parameters, namely 1280. Each model was trained for 250 epochs with the SGD optimizer, a learning rate of 0.0003 for the real model and 0.0009 for the other two models, and a single batch consisting of the entire training set.

We use the prediction step to showcase the translation and rotation equivariance properties. We generate a random translation and a random rotation. After training on the original training set, the networks are tested on 4 different test sets: i) the original test set; ii) the translated test set; iii) the rotated test set; and iv) the translated+rotated test set. Sets ii-iv) are obtained by applying the aforementioned randomly generated operations. We reiterate that there is no retraining between the tests. Hence only the original training set was presented to the networks during training. Figure 1 shows the predicted trajectories for the original, translated, and translated+rotated test sets. From the left column, it is clear the models are able to solve the prediction task, albeit with visible flaws from the



**Fig. 1.** Predicted trajectories in red, expected in blue. Rows contain real-, quaternion- and dual quaternion-valued outputs, respectively. Columns show outputs for the original, translated, and translated+rotated test sets, respectively.

real and quaternion models (Fig. 1 (a)(b)). In the middle column, it can be seen that the translation barely affects the dual quaternion model, while the other two fail to adapt. Finally, the right column shows the real and quaternion models utterly fail to predict the translated+rotated test set, outputting a trajectory not even close to the expected (Fig. 1 (g)(h)). On the other hand, the dual quaternion model (Fig. 1(i)) shows an accurate prediction, correctly outputting a rotated and translated trajectory.

Table 1 shows the mean squared error (MSE) and the prediction gain per model for each of the 4 test sets. The prediction gain expresses the ratio between a reference level ( $\sigma_s$ ) and the signal error ( $\sigma_e$ ). Formally, the prediction gain is defined by  $R = 10 \log_{10} \frac{\sigma_s^2}{\sigma_e^2}$ , where  $\sigma_s^2$  is the estimated variance of the input signal and  $\sigma_e^2$  denotes the estimated variance of the prediction error. It is clear from the table that the dual quaternion model wins every scenario, adapting to the transformations in the test set seamlessly. Since there is no retraining at all, this constitutes strong evidence that the model learns the rigid motions of the Lorenz system rather than mimicking a local behavior. In fact, since a dual quaternion number uniquely describes a rigid motion, the dual quaternion MLP learns the correct rigid motions of the Lorenz system by fine-tuning its dual quaternion parameters. Moreover, this shows that the equivariance properties are a natural consequence of endowing the model with the dual quaternion algebra, which in turn is simply the replacement of sum and product by the respective operations.

The possibility to train a network on a dataset and achieve roughly the same performance on translated and/or rotated objects is invaluable and a direct consequence of the rotation and translation equivariance properties of the underlying algebra.

## 4.2. Human Pose Forecasting

We present a variational autoencoder endowed with dual quaternion numbers (DQVAE) that leverages both global and local information by means of dual quaternion representation of rigid motions to more accurately predict the pose of individuals [23]. Global information is contained in the center of mass coordinates since a human moving is essentially described by a translation applied to each point of the body. Local information, instead, is more accurately described as joint positions relative to the center of mass and more closely related to the action being executed. Indeed, the global trajectory information is contained in the coordinates  $x_c, y_c, z_c$  of the center of mass, while the local information is given by the relative position of joints with respect to the center of mass coordinates as  $(x - x_c), (y - y_c), (z - z_c)$ .

We consider as baseline the decoupled representations for pose forecasting (DeRPoF) [23] that involves two different models for trajectory and local pose estimation separately, with an LSTM autoencoder and an LSTM variational autoencoder, respectively. Leveraging the higher-dimensional nature of hypercomplex numbers, we propose the quaternion and dual quaternion coupled representation for pose forecasting models, referred to as Quaternion CoRPoF and Dual Quaternion CoRPoF, respectively. Thanks to the high-dimensional nature of the dual quaternions, the latter model is able to enclose both global and local information in a single entity as

$$\hat{\mathbf{q}} = x_c \hat{i} + y_c \hat{j} + z_c \hat{k} + \varepsilon ((x - x_c) \hat{i} + (y - y_c) \hat{j} + (z - z_c) \hat{k}). \quad (7)$$

We carry out a validation experiment in the 3D Poses in the Wild (3DPW) dataset. The task is predicting short-term immediate future poses for human bodies based on joint information. The 3DPW dataset contains over 51k registered frames from 60 short videos portraying humans performing basic activities such as hugging, arguing, playing basketball, and dancing, among others. Data is provided in 2 forms: RGB images and structured data. We focus on the latter, which comprises the 3D coordinates of the center of mass and of 13 key joints of individuals. Hence, this task consists of predicting rigid motions of sets of points, similar to the Lorenz system task above.

The proposed Dual Quaternion CoRPoF has an encoder composed of single-layer dual quaternion LSTM (DQLSTM) blocks of hidden dimension 64, along with two fully connected (DQFC) layers with latent dimension 32, which encode the mean and variance of the latent vector. The decoder of the proposed dual quaternion variational autoencoder (DQVAE) comprises two DQLSTM layers with the same hidden

**Table 1.** MSE and Prediction Gain for the Lorenz system prediction task with the original, translated, rotated and translated+rotated test sets, respectively.

Model	Test MSE↓				Test Prediction Gain↑			
	Original	Translated	Rotated	T+R	Original	Translated	Rotated	T+R
Real-valued	0.433	4.648	118.719	192.766	57.187	31.050	14.024	10.360
Quaternion-valued	0.756	7.952	178.102	263.051	51.918	24.841	6.847	3.729
Dual Quaternion-valued	<b>0.272</b>	<b>0.183</b>	<b>2.140</b>	<b>3.617</b>	<b>63.606</b>	<b>52.722</b>	<b>43.797</b>	<b>40.108</b>

**Table 2.** Results for 3D human pose estimation on the 3DPW dataset. VIM scores are in centimeters, as suggested in [23].

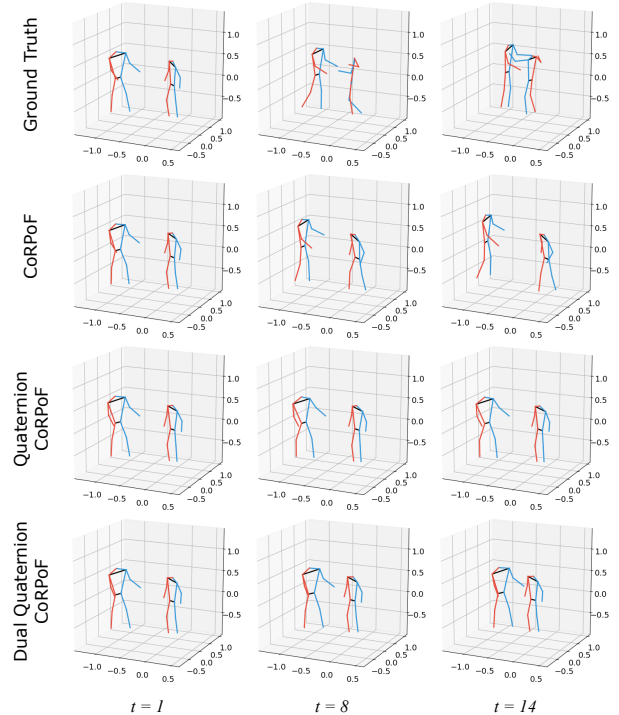
Model	VIM↓	FDE↓	Val Loss↓
SC-MPF [24]	46.28	-	-
Nearest Neighbour [25]	27.34	-	-
Zero Velocity [26]	29.35	-	-
DeRPoF [23]	19.07 ± .005	0.360 ± .007	-
CoRPoF [23]	16.76 ± .003	0.317 ± .001	0.118 ± .004
Quaternion CoRPoF	16.35 ± .009	0.271 ± .002	0.105 ± .010
Dual Quaternion CoRPoF	<b>15.23 ± .002</b>	<b>0.266 ± .001</b>	<b>0.103 ± .006</b>

dimension and the final DQFC layers to output the predicted pose. This decoder does not merely reconstruct the input but instead aims at generating future estimations while optimizing the following variational bound:

$$\mathcal{L} = \sum_{t=t_{obs}+1}^{t_{obs}+t_{fut}} \|p_t - \hat{p}_t\|^2 + \beta \text{KL}(p(z) \| \mathcal{N}(\mathbf{0}, \mathbf{I})), \quad (8)$$

where  $p_t$  is the future ground truth pose,  $\hat{p}_t$  is the predicted one,  $p(z)$  is the latent distribution, and  $\beta$  is the hyperparameter that balances the two elements. The first term of (8) cares about the trajectory estimation, while the second element is crucial for local poses. The model is trained for 1000 epochs by the Adam optimizer with a learning rate of 0.01 and an adaptive scheduler, as suggested in [23]. It is worth mentioning that due to the properties of the Hamilton product, the quaternion and dual quaternion models reduce the number of parameters of the network with respect to the real-valued baseline CoRPoF by 75% and 88%, respectively.

To assess the performance of our model and to be consistent with the previous literature we compute the visibility ignored metric (VIM), which is the average of the distances between each predicted joint and the corresponding point in the ground truth, and the final displacement error (FDE), which is instead an L2 distance. Table 2 reports these objective metrics as average scores and standard deviations over multiple runs. As expected, there is notable improvement operated by our dual quaternion formulation due to the effectiveness of higher-dimensional representations for movements in the 3D space. The comparisons with SC-MPF, Nearest Neighbours, and Zero velocity are performed with the scores reported in their papers, as their code is not available. Figure 2 dis-



**Fig. 2.** Estimated poses at different time steps. The dual quaternion CoRPoF models the translations, while the real- and quaternion-valued networks fail to do so.

plays a ground truth sample and the corresponding predicted skeletons by the Dual Quaternion CoRPoF and its quaternion- and real-valued counterparts. Although real CoRPoF tries to model local poses, it fails to learn the right movement trajectory and skeletons are stuck in the same coordinates. The quaternion model on the other hand shows little translation of the skeletons but fails to model their fine movements. In the last row, the dual quaternion CoRPoF learns the correct trajectory of both skeletons while also displaying some of the expected fine movements. This showcases the natural ability of the proposed dual representation to model translations in 3D space.

## 5. CONCLUDING REMARKS

In this paper, we present a formulation of dual quaternions that puts in evidence its connection to modeling rigid motions in 3D. We formally explain how a dual quaternion jointly encases information regarding translations and rotations, and provide a practical example of the translation and rotation equivariance properties using the Lorenz system. We proceed to show how models endowed with this formulation outperform other approaches in an application to human pose forecasting, proving our theoretical claims that the dual quaternion-valued models are robust to data shifting and more adequate to such tasks.

## 6. REFERENCES

- [1] K. Mangalam, E. Adeli, K.-H. Lee, A. Gaidon, and J. C. Niebles, "Disentangling human dynamics for pedestrian locomotion forecasting with noisy supervision," in *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2020, pp. 2784–2793.
- [2] H. Razali, T. Mordan, and A. Alahi, "Pedestrian intention prediction: A convolutional bottom-up multi-task approach," *Transportation research part C: emerging technologies*, vol. 130, pp. 103259, 2021.
- [3] L. Kidziński, B. Yang, J. L. Hicks, A. Rajagopal, S. L. Delp, and M. H. Schwartz, "Deep neural networks enable quantitative movement analysis using single-camera videos," *Nature communications*, vol. 11, no. 1, pp. 1–10, 2020.
- [4] M. Cormier, A. Clepe, A. Specker, and J. Beyerer, "Where are we with human pose estimation in real-world surveillance?," in *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2022, pp. 591–601.
- [5] Z. Li, J. Liu, Z. Zhang, S. Xu, and Y. Yan, "CLIFF: Carrying location information in full frames into human pose and shape estimation," *arXiv preprint arXiv:2208.00571*, 2022.
- [6] S. P. Talebi and D. P. Mandic, "A quaternion frequency estimator for three-phase power systems," in *IEEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, 2015, pp. 3956–3960.
- [7] E. Grassucci, D. Comminiello, and A. Uncini, "A quaternion-valued variational autoencoder," in *EEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, 2021, pp. 3310–3314.
- [8] G. Vieira and M. E. Valle, "Acute lymphoblastic leukemia detection using hypercomplex-valued convolutional neural networks," *arXiv preprint arXiv:2205.13273*, 2022.
- [9] T. Parcollet, M. Morchid, and G. Linarès, "A survey of quaternion neural networks," *Artificial Intelligence Review*, vol. 53, no. 4, pp. 2957–2982, 4 2020.
- [10] Y. Zhou, L. Jin, G. Ma, and X. Xu, "Quaternion capsule neural network with region attention for facial expression recognition in color images," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, no. 4, pp. 893–912, 2022.
- [11] Z. Jia, Q. Jin, M. K. Ng, and X.-L. Zhao, "Non-local robust quaternion matrix completion for large-scale color image and video inpainting," *IEEE Trans. on Image Process.*, vol. 31, pp. 3868–3883, 2022.
- [12] E. Grassucci, E. Cicero, and D. Comminiello, "Quaternion generative adversarial networks," in *Generative Adversarial Learning: Architectures and Applications*, R. Razavi-Far, A. Ruiz-Garcia, V. Palade, and J. Schmidhuber, Eds., pp. 57–86. Springer International Publishing, Cham, 2022.
- [13] C. Brignone, G. Mancini, E. Grassucci, A. Uncini, and D. Comminiello, "Efficient sound event localization and detection in the quaternion domain," *IEEE Trans. on Circuits and Systems II: Express Briefs*, vol. 69, no. 5, pp. 2453–2457, 2022.
- [14] D. Pavlo, D. Grangier, and M. Auli, "Quaternion: A quaternion-based recurrent model for human motion," *arXiv preprint arXiv:1805.06485*, 2018.
- [15] D. Pavlo, C. Feichtenhofer, M. Auli, and D. Grangier, "Modeling human motion with quaternion-based neural networks," *Int. Journal of Computer Vision*, vol. 128, no. 4, pp. 855–872, 2020.
- [16] H.-W. Hsu, T.-Y. Wu, S. Wan, W. H. Wong, and C.-Y. Lee, "QuatNet: Quaternion-based head pose estimation with multi-regression loss," *IEEE Trans. on Multimedia*, vol. 21, no. 4, pp. 1035–1046, 2019.
- [17] C. Papaioannidis and I. Pitas, "3D object pose estimation using multi-objective quaternion learning," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 30, no. 8, pp. 2683–2693, 2020.
- [18] X. Zhang, S. Qin, Y. Xu, and H. Xu, "Quaternion product units for deep learning on 3d rotation groups," *IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 7302–7311, 2020.
- [19] E. Pennestrì and P. P. Valentini, "Dual quaternions as a tool for rigid body motion analysis: A tutorial with an application to biomechanics," *Archive of Mechanical Engineering*, vol. 57, no. 2, pp. 187–205, 2010.
- [20] J. Pöppelbaum and A. Schwung, "Predicting rigid body dynamics using dual quaternion recurrent neural networks with quaternion attention," *IEEE Access*, vol. 10, pp. 82923–82943, 2022.
- [21] A. Schwung, J. Pöppelbaum, and P. C. Nutakki, "Rigid body movement prediction using dual quaternion recurrent neural networks," in *IEEE Int. Conf. on Industrial Technology (ICIT)*, 2021, vol. 1, pp. 756–761.
- [22] L. Kavan, S. Collins, C. O'Sullivan, and J. Zara, "Dual quaternions for rigid transformation blending," *Trinity College Dublin, Tech. Rep.*, vol. 46, 2006.
- [23] B. Parsaeifard, S. Saadatnejad, Y. Liu, T. Mordan, and A. Alahi, "Learning decoupled representations for human pose forecasting," *IEEE/CVF Int. Conf. on Computer Vision Workshops (ICCVW)*, pp. 2294–2303, 2021.
- [24] V. Adeli, E. Adeli, I. D. Reid, J. C. Niebles, and H. Rezatofighi, "Socially and contextually aware human motion and pose forecasting," *IEEE Robotics and Automation Letters*, vol. 5, pp. 6033–6040, 2020.
- [25] J. Y. Zhang, P. Felsen, A. Kanazawa, and J. Malik, "Predicting 3d human dynamics from video," *IEEE/CVF Int. Conf. on Computer Vision (ICCV)*, pp. 7113–7122, 2019.
- [26] J. Martinez, M. J. Black, and J. Romero, "On human motion prediction using recurrent neural networks," *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 4674–4683, 2017.