

$E(n)$ -EQUIVARIANT SPHERICAL DECISION SURFACES

Anonymous authors

Paper under double-blind review

ABSTRACT

We present a constructive derivation of exactly $E(n)$ -equivariant spherical decision surfaces by extending prior $O(n)$ -equivariant hypersphere neurons to include translations. To achieve this, we present a decomposition of the features of the $O(n)$ -equivariant neurons and provide explicit representations for translation and $E(n)$ -transformations to fulfil the respective equivariance constraints. The resulting decision surfaces are exactly $E(n)$ -equivariant without input centring or explicit pairwise differences, and admit explicit closed-form matrix representations. In addition, we numerically verify the correctness of the derivations and perform a downstream check of the resulting geometric primitives.

1 INTRODUCTION

Typically, $O(n)$ -equivariant¹ methods (Esteves et al., 2017; Deng et al., 2021; Assaad et al., 2023; Ruhe et al., 2023) require centring the input to become translation-equivariant, i.e., equivariant under the Euclidean group $E(n)$, or explicitly using the differences between input points, considering input as a collection of points (Thomas et al., 2018; Fuchs et al., 2020; Satorras et al., 2021). Motivated by the fact that spherical decision surfaces (Perwass et al., 2003) produce isometric activations, as shown by Melnyk et al. (2021) (which hints at the equivariance under the full Euclidean group), see Section 2.3, we aim to make $O(n)$ -equivariant hyperspheres (neurons) (Melnyk et al., 2024a) also translation-equivariant, beginning by considering only a single-point input.

Section 3 in its entirety constitutes the contributions of our work. Our goal is to establish and explicitly derive an exactly $E(n)$ -equivariant extension of $O(n)$ -equivariant spherical decision surfaces (Melnyk et al., 2024a) at the neuron level, rather than propose a composable deep architecture. We summarise our contributions as follows:

- we improve the understanding of the features learned by $O(n)$ -equivariant spherical neurons by showing how to decompose them into irreducible representations (irreps) — see Section 3.1;
- we derive the $E(n)$ -equivariant hypersphere neuron based on regular simplices and spherical decision surfaces (see (24) and (26)), with the $E(n)$ -invariant primitives (41) and (43);
- we provide explicit transformation representations fulfilling the equivariance constraints (see (28) and (37)).

In addition, we numerically demonstrate the correctness of our derivations, with the results presented in Section 4. An executable notebook reproducing these experiments will be made publicly available.

2 BACKGROUND

In this section, we present a comprehensive background on the theory of spherical neurons (Perwass et al., 2003; Melnyk et al., 2021) and their rotation/reflection-equivariant version (Melnyk et al., 2022; 2024a), as well as on the general geometric concepts used in our work.

2.1 REGULAR SIMPLICES AND THE SIMPLEX CHANGE OF BASIS

Geometrically, a regular n -simplex represents $n + 1$ equidistant points in nD (Elte, 2006), lying on an nD sphere with unit radius. In the 2D case, the regular simplex is an equilateral triangle; in 3D, a regular tetrahedron, and so on.

¹Here, we also include methods that are strictly speaking only $SO(n)$ -equivariant.

054 Following Cevikalp & Saribas (2023), we compute the Cartesian coordinates of a regular n -simplex
 055 as $n + 1$ vectors $\mathbf{p}_i \in \mathbb{R}^n$:

$$056 \mathbf{p}_i = \begin{cases} n^{-1/2} \mathbf{1}, & i = 1 \\ \kappa \mathbf{1} + \mu \mathbf{e}_{i-1}, & 2 \leq i \leq n + 1 \end{cases}, \quad \kappa = -\frac{1 + \sqrt{n+1}}{n^{3/2}}, \quad \mu = \sqrt{1 + \frac{1}{n}}, \quad (1)$$

059 where $\mathbf{1} \in \mathbb{R}^n$ is a vector with all elements equal to 1 and \mathbf{e}_i is the natural basis vector with the
 060 i -th element equal to 1. For convenience we define \mathbf{P}_n as an $n \times (n + 1)$ matrix holding the regular
 061 n -simplex vertices, \mathbf{P}_i , in its columns:

$$062 \mathbf{P}_n = \begin{bmatrix} | & | & & | \\ \mathbf{p}_0 & \mathbf{p}_1 & \cdots & \mathbf{p}_n \\ | & | & & | \end{bmatrix}. \quad (2)$$

063 Based on the definition of the regular simplices (1), Melnyk et al. (2024a) generalised to n D the
 064 change-of-basis matrix defined in 3D by Melnyk et al. (2022) as an $(n + 1) \times (n + 1)$ orthogonal
 065 matrix \mathbf{M}_n (i.e., representing rotation or reflection) holding in its columns the coordinates of the
 066 regular n -simplex appended with a constant and normalised to unit length:

$$067 \mathbf{M}_n = [\mathbf{m}_i]_{i=1}^{n+1}, \quad \mathbf{m}_i = \frac{1}{p} \begin{bmatrix} \mathbf{p}_i \\ n^{-1/2} \end{bmatrix}, \quad p = \left\| \begin{bmatrix} \mathbf{p}_i \\ n^{-1/2} \end{bmatrix} \right\| = \sqrt{\frac{n+1}{n}}, \quad (3)$$

068 where the norms p are constant, since $\|\mathbf{p}_i\| = \|\mathbf{p}_j\|$ for all i and j by definition of a regular simplex.

069 The following relation between \mathbf{M}_n and \mathbf{P}_n is important for the $O(n)$ -equivariance derivation of the
 070 neurons recapped in Section 2.4:

$$071 \mathbf{M}_n \mathbf{P}_n^\top = p \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0}^\top \end{bmatrix}. \quad (4)$$

072 2.2 EQUI- AND INVARIANCE UNDER $E(n)$ -TRANSFORMATIONS

073 The elements of the orthogonal group $O(n)$ can be represented as $n \times n$ matrices \mathbf{R} with the properties
 074 $\mathbf{R}^\top \mathbf{R} = \mathbf{R} \mathbf{R}^\top = \mathbf{I}_n$, where \mathbf{I}_n is the identity matrix, and $\det \mathbf{R} = \pm 1$, geometrically characterizing
 075 n D rotations and reflections. The special orthogonal group $SO(n)$ is a subgroup of $O(n)$ and includes
 076 only orthogonal matrices with the positive determinant, representing rotations.

077 The Euclidean group $E(n)$ is defined as a semidirect product $E(n) \cong \mathbb{R}^n \rtimes O(n)$; it consists of
 078 all orthogonal transformations together with translations, which can be represented as pairs (\mathbf{R}, \mathbf{t}) :
 079 $E(n) = \{(\mathbf{R}, \mathbf{t}) \mid \mathbf{R} \in O(n), \mathbf{t} \in \mathbb{R}^n\}$ with $(\mathbf{R}_2, \mathbf{t}_2)(\mathbf{R}_1, \mathbf{t}_1) = (\mathbf{R}_2 \mathbf{R}_1, \mathbf{t}_2 + \mathbf{R}_2 \mathbf{t}_1)$.

080 We say that a function $f : \mathcal{X} \rightarrow \mathcal{Y}$ is $E(n)$ -equivariant if for every $\mathbf{R} \in O(n)$ and $\mathbf{t} \in \mathbb{R}^n$, there exists
 081 the transformation representation $\rho(\mathbf{R}, \mathbf{t})$, in the function output space, \mathcal{Y} , such that

$$082 \rho(\mathbf{R}, \mathbf{t}) f(\mathbf{x}) = f(\mathbf{R} \mathbf{x} + \mathbf{t}) \quad \text{for all } (\mathbf{R}, \mathbf{t}) \in E(n), \mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^n, \quad (5)$$

083 where $\rho(\cdot)$ is a group representation, i.e., a homomorphism $\rho : E(n) \rightarrow GL(\mathcal{Y})$, mapping each (\mathbf{R}, \mathbf{t})
 084 to an invertible linear operator on \mathcal{Y} — a matrix operator in some chosen basis.

085 We call a function $f : \mathcal{X} \rightarrow \mathcal{Y}$ $E(n)$ -invariant if for every $\mathbf{R} \in O(n)$ and $\mathbf{t} \in \mathbb{R}^n$, $\rho(\mathbf{R}, \mathbf{t}) = \mathbf{I}_n$. That
 086 is, if

$$087 f(\mathbf{x}) = f(\mathbf{R} \mathbf{x} + \mathbf{t}) \quad \text{for all } (\mathbf{R}, \mathbf{t}) \in E(n), \mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^n. \quad (6)$$

088 Following the convention of prior work (Melnyk et al., 2022; 2024b) hereinafter, we write \mathbf{R} to denote
 089 the same n D rotation/reflection as a $n \times n$ matrix in the Euclidean space \mathbb{R}^n , as a $(n + 1) \times (n + 1)$
 090 matrix in the projective (homogeneous) space $P(\mathbb{R}^n) \subset \mathbb{R}^{n+1}$, and as an $(n + 2) \times (n + 2)$ matrix
 091 in \mathbb{R}^{n+2} . For the latter two cases, we achieve this by appending ones to the diagonal of the original
 092 $n \times n$ matrix without changing the transformation itself (Melnyk et al., 2021).

093 2.3 SPHERICAL NEURONS VIA NON-LINEAR EMBEDDING

094 Spherical neurons (Perwass et al., 2003; Melnyk et al., 2021) are neurons with, as the name suggests,
 095 spherical decision surfaces.

Sphere as a classifier By virtue of conformal geometric algebra (Li et al., 2001), Perwass et al. (2003) proposed to embed the data vector $\mathbf{x} \in \mathbb{R}^n$ and represent the sphere with centre $\mathbf{c} = (c_1, \dots, c_n) \in \mathbb{R}^n$ and radius $r \in \mathbb{R}$ respectively as

$$\mathbf{X} = (x_1, \dots, x_n, -1, -\frac{1}{2}\|\mathbf{x}\|^2) \in \mathbb{R}^{n+2}, \quad \mathbf{S} = (c_1, \dots, c_n, \frac{1}{2}(\|\mathbf{c}\|^2 - r^2), 1) \in \mathbb{R}^{n+2}, \quad (7)$$

and used their scalar product $\mathbf{X}^\top \mathbf{S} = -\frac{1}{2}\|\mathbf{x} - \mathbf{c}\|^2 + \frac{1}{2}r^2$ as a classifier, i.e., the spherical neuron:

$$f_S(\mathbf{X}; \mathbf{S}) = \mathbf{X}^\top \mathbf{S}, \quad (8)$$

with learnable parameters $\mathbf{S} \in \mathbb{R}^{n+2}$. The sign of this scalar product depends on the position of the point \mathbf{x} relative to the sphere (\mathbf{c}, r) : inside the sphere if positive, outside the sphere if negative, and on the sphere if zero (Perwass et al., 2003).

Geometric interpretation and isometric activation Geometrically, the activation of the spherical neuron (8) determines the cathetus length of the right triangle formed by \mathbf{x} , \mathbf{c} , and the corresponding point on the sphere (see Figure 2 in Melnyk et al. (2021)), and, as shown by Melnyk et al. (2021), is *isometric*, i.e., rigid transformations of the spherical neuron (or rather, its spherical decision surface) correspond to the rigid transformations of the vector it acts upon: Using two $(n+2) \times (n+2)$ *motor*, i.e., rigid transformation, operators \mathbf{Q}_S and \mathbf{Q}_X that act on \mathbf{S} and \mathbf{X} (7), respectively,

$$\mathbf{Q}_S = \begin{bmatrix} \mathbf{R} & \mathbf{0} & \mathbf{t} \\ \mathbf{t}^\top \mathbf{R} & 1 & \frac{1}{2}\|\mathbf{t}\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{Q}_X = \begin{bmatrix} \mathbf{R} & -\mathbf{t} & \mathbf{0} \\ \mathbf{0}^\top & 1 & 0 \\ -\mathbf{t}^\top \mathbf{R} & \frac{1}{2}\|\mathbf{t}\|^2 & 1 \end{bmatrix}, \quad (9)$$

such that \mathbf{Q}_X is the adjoint of \mathbf{Q}_S , i.e., $\mathbf{Q}_X^\top \mathbf{Q}_S = \mathbf{I}_{n+2}$, we obtain

$$(\mathbf{Q}_X \mathbf{X})^\top (\mathbf{Q}_S \mathbf{S}) = \mathbf{X}^\top \mathbf{Q}_X^\top \mathbf{Q}_S \mathbf{S} = \mathbf{X}^\top \mathbf{S}. \quad (10)$$

Non-linear activation We note that with respect to the data vector $\mathbf{x} \in \mathbb{R}^n$, a spherical neuron represents a non-linear function $f_S(\cdot; \mathbf{S}) : \mathbb{R}^{n+2} \rightarrow \mathbb{R}$, due to the inherent non-linearity of the embedding (7), and therefore, does not necessarily require an activation function, as observed by Melnyk et al. (2021).

Independent learnable parameters The components of \mathbf{S} in (7) can be treated as *independent* learnable parameters. In this case, a spherical neuron learns a *non-normalised* sphere of the form $\tilde{\mathbf{S}} = (\tilde{s}_1, \dots, \tilde{s}_{n+2}) \in \mathbb{R}^{n+2}$, which represents the same decision surface as its normalised counterpart defined in (7), thanks to the homogeneity of the embedding (Perwass et al., 2003; Li et al., 2001). The last parameter of $\tilde{\mathbf{S}}$ is called the scale parameter $\gamma := \tilde{s}_{n+2}$, so the definition (8) can be written more generally as

$$f_S(\mathbf{X}; \tilde{\mathbf{S}}) = \gamma \mathbf{X}^\top \mathbf{S}. \quad (11)$$

For simplicity and without loss of generality, in the following we suppose $\gamma = 1$.

2.4 $O(n)$ -EQUIVARIANT HYPERSPHERES (NEURONS)

Following the proposal for 3D by Melnyk et al. (2022), which was later utilised in an end-to-end method (Melnyk et al., 2024b), Melnyk et al. (2024a) generalised to n D an $O(n)$ -equivariant hypersphere neuron (also referred to as *equivariant hypersphere*) based on spherical decision surfaces (Perwass et al., 2003; Melnyk et al., 2021) and regular simplices (see Section 2.1).

Definition Particularly, an equivariant hypersphere $\mathbf{F}_n(\mathbf{X}; \mathbf{S})$ is defined as the $(n+1) \times (n+2)$ matrix (\mathbf{S}) for the spherical decision surface $\mathbf{S} \in \mathbb{R}^{n+2}$ with centre $\mathbf{c} \in \mathbb{R}^n$ and an n D input $\mathbf{x} \in \mathbb{R}^n$ embedded as $\mathbf{X} \in \mathbb{R}^{n+2}$ as

$$\mathbf{F}_n(\mathbf{X}; \mathbf{S}) = B(\mathbf{S}) \mathbf{X}, \quad B(\mathbf{S}) = \left[(\mathbf{R}_O^\top \mathbf{R}_{T_i} \mathbf{R}_O \mathbf{S})^\top \right]_{i=1}^{n+1}, \quad (12)$$

where $\{\mathbf{R}_{T_i}\}_{i=1}^{n+1}$ is the \mathbb{R}^{n+2} rotation isomorphism corresponding to the rotation from the first vertex to the i -th vertex of the regular n -simplex (1) (and therefore, $\mathbf{R}_{T_1} = \mathbf{I}_{n+2}$), and $\mathbf{R}_O \in \text{SO}(n)$ is the

geodesic (shortest) rotation² from the sphere centre \mathbf{c} to $\|\mathbf{c}\| \mathbf{p}_1$:

$$\mathbf{R}_O \mathbf{c} = \|\mathbf{c}\| \mathbf{p}_1, \quad (13)$$

where $\mathbf{p}_1 \in \mathbb{R}^n$ is the first vertex of the regular simplex according to (1). Hence, $\mathbf{R}_{T_i} \mathbf{R}_O \mathbf{c} = \|\mathbf{c}\| \mathbf{p}_i$. Note that if the centre \mathbf{c} happens to be $-\mathbf{p}_1$, \mathbf{R}_O is a reflection about the origin. Thus, in principle, we could also write $\mathbf{R}_O \in O(n)$.

The definition of $B(\mathbf{S})$ in (12) can be written by using the definition of a spherical decision surface (7) and a regular simplex (1) and (2), as well as the definition of \mathbf{R}_O (13):

$$\begin{aligned} B(\mathbf{S}) &= \left[(\mathbf{R}_O^\top \mathbf{R}_{T_i} \mathbf{R}_O \mathbf{S})^\top \right]_{i=1}^{n+1} = \left[\mathbf{c}^\top \mathbf{R}_O^\top \mathbf{R}_{T_i}^\top \mathbf{R}_O \quad s_{n+1} \quad 1 \right]_{i=1}^{n+1} \\ &= \left[\|\mathbf{c}\| \mathbf{p}_i^\top \mathbf{R}_O \quad s_{n+1} \quad 1 \right]_{i=1}^{n+1} = \left[\|\mathbf{c}\| \mathbf{P}_n^\top \mathbf{R}_O \quad s_{n+1} \quad \mathbf{1} \quad \mathbf{1} \right], \end{aligned} \quad (14)$$

where $s_{n+1} := \frac{1}{2}(\|\mathbf{c}\|^2 - r^2)$.

$O(n)$ -equivariance The $O(n)$ -equivariance property for the equivariant hypersphere is

$$\mathbf{V}(\mathbf{R}) B(\mathbf{S}) \mathbf{X} = B(\mathbf{S}) \mathbf{R} \mathbf{X}, \quad (15)$$

where \mathbf{R} is the rotation representation in \mathbb{R}^{n+2} obtained by extending the diagonal of the original $n \times n$ matrix with 1s (ones), and the transformation representation in the neuron output space \mathbb{R}^{n+1} is given by the $(n+1) \times (n+1)$ matrix

$$\mathbf{V}(\mathbf{R}) := \mathbf{M}_n^\top \mathbf{R}_O \mathbf{R} \mathbf{R}_O^\top \mathbf{M}_n, \quad (16)$$

where $\mathbf{M}_n \in O(n+1)$ is the change-of-basis matrix defined in (3) and a 1 is appended to the diagonals of \mathbf{R}_O and \mathbf{R} to make them $(n+1) \times (n+1)$. Furthermore, $\mathbf{V}(\mathbf{R})$ belongs to a proper subgroup of $O(n+1)$: $G = O(n) \times \mathcal{S}^n$, i.e., $G < O(n+1)$, where G is formed as $O(n) \times \mathbf{M}_n \left(\frac{\mathbf{c}}{\|\mathbf{c}\|}, 0 \right)$ with $\mathbf{M}_n \left(\frac{\mathbf{c}}{\|\mathbf{c}\|}, 0 \right) \in \mathcal{S}^n$.

The original transformation is found directly as

$$\mathbf{R} = \mathbf{R}_O^\top \mathbf{M}_n \mathbf{V}(\mathbf{R}) \mathbf{M}_n^\top \mathbf{R}_O, \quad (17)$$

followed by the retrieval of the upper-left $n \times n$ sub-matrix.

Non-linearity With respect to the input vector $\mathbf{x} \in \mathbb{R}^n$, the equivariant hypersphere $\mathbf{F}_n(\cdot; \mathbf{S}) : \mathbb{R}^{n+2} \rightarrow \mathbb{R}^{n+1}$ represents a non-linear $O(n)$ -equivariant function. Also, the sum of the output $B(\mathbf{S}) \mathbf{X}$ is an $O(n)$ -invariant scalar due to the regular n -simplex construction.

3 THEORETICAL RESULTS

In this section, we present the main results of our work. Complete derivation and proofs are presented in Appendix A, which, for convenience, follows the same order as this section.

For the derivation of translation equivariance, it is instrumental first to consider the following decomposition of the output \mathbf{y} of the $O(n)$ -equivariant hypersphere neuron.

3.1 IDENTITY SPLIT: EQUIVARIANT HYPERSPHERE OUTPUT DECOMPOSITION INTO IRREPS

Given $\mathbf{X} \in \mathbb{R}^{n+2}$ (7) and $\mathbf{y} := B(\mathbf{S}) \mathbf{X} \in \mathbb{R}^{n+1}$, we decompose it into $O(n)$ -equivariant and invariant parts using (14) as

$$\mathbf{R}_O^\top \mathbf{M}_n \mathbf{y} = \mathbf{R}_O^\top \mathbf{M}_n B(\mathbf{S}) \mathbf{X} = p \|\mathbf{c}\| \left[\alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \right], \quad (18)$$

²In practice, it is computed utilizing the Householder (double-) reflection method, e.g., as described by Golub & Van Loan (2013).

where $\alpha := \frac{\sqrt{n+1}}{2p\|\mathbf{c}\|} = \frac{\sqrt{n}}{2\|\mathbf{c}\|}$, and the invariant part is $\frac{\sqrt{n+1}}{2}(r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2)$, with \mathbf{c} and r being learnable parameters.

Observe that in the case $\|\mathbf{c}\| = 0$, the $O(n)$ -equivariant part, i.e., the original vector \mathbf{x} , zeroes out, and what we are left with is

$$\mathbf{R}_O^\top \mathbf{M}_n \mathbf{y} = \mathbf{R}_O^\top \mathbf{M}_n B(\mathbf{S}) \mathbf{X} = \begin{bmatrix} \mathbf{0} \\ \frac{p\sqrt{n}}{2}(r^2 - \|\mathbf{x}\|^2) \end{bmatrix}. \quad (19)$$

Also, if \mathbf{X} is transformed with $\mathbf{R} \in O(n)$ extended to $(n+2) \times (n+2)$ using the identity matrix, as $\mathbf{R}\mathbf{X}$, defining $\mathbf{y}_R := B(\mathbf{S})\mathbf{R}\mathbf{X}$, we obtain

$$\mathbf{R}_O^\top \mathbf{M}_n \mathbf{y}_R = \mathbf{R}_O^\top \mathbf{M}_n B(\mathbf{S}) \mathbf{R}\mathbf{X} = p\|\mathbf{c}\| \begin{bmatrix} \mathbf{R}\mathbf{x} \\ \alpha(r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \end{bmatrix}. \quad (20)$$

3.2 TRANSLATION EQUIVARIANCE

The desired equivariance property is

$$\mathbf{T}(\mathbf{t}) B(\mathbf{S}) \mathbf{X} = B(\mathbf{S}) \mathbf{T}_{\text{RHS}} \mathbf{X}, \quad (21)$$

where $\mathbf{T}(\mathbf{t})$ is the translation representation we seek and

$$\mathbf{T}_{\text{RHS}} = \begin{bmatrix} \mathbf{I}_n & -\mathbf{t} & \mathbf{0} \\ \mathbf{0}^\top & 1 & 0 \\ -\mathbf{t}^\top & \frac{1}{2}\|\mathbf{t}\|^2 & 1 \end{bmatrix} \quad (22)$$

is the $(n+2) \times (n+2)$ matrix (9) with $\mathbf{R} = \mathbf{I}_n$.

Consider the RHS and denote \mathbf{y}_t to be the output of the equivariant hypersphere given the input \mathbf{x} translated as $\mathbf{x} + \mathbf{t}$ and embedded in \mathbb{R}^{n+2} (see (7)) as \mathbf{X}_t :

$$\mathbf{y}_t := B(\mathbf{S}) \mathbf{X}_t = B(\mathbf{S}) \mathbf{T}_{\text{RHS}}(\mathbf{t}) \mathbf{X} \in \mathbb{R}^{n+1}. \quad (23)$$

As we show in the Appendix, to satisfy (21), we need to make the original $(n+1) \times (n+2)$ equivariant hypersphere $B(\mathbf{S})$ (12) output a *homogeneous* feature. For this, we create the $(n+2) \times (n+2)$ $B_{\text{hom}}(\mathbf{S})$:

$$B_{\text{hom}}(\mathbf{S}) := \begin{bmatrix} B(\mathbf{S}) \\ \mathbf{0}^\top & -p\|\mathbf{c}\| & 0 \end{bmatrix}, \quad (24)$$

so that

$$B_{\text{hom}}(\mathbf{S}) \mathbf{X} = \begin{bmatrix} B(\mathbf{S}) \mathbf{X} \\ p\|\mathbf{c}\| \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ p\|\mathbf{c}\| \end{bmatrix} \in \mathbb{R}^{n+2}, \quad (25)$$

where $\mathbf{c} \in \mathbb{R}^n$ are parameters in \mathbf{S} , and hence no new learnable parameters are added, and the $-$ sign in the $-p\|\mathbf{c}\|$ element in $B_{\text{hom}}(\mathbf{S})$ is due to the second last element of the embedding \mathbf{X} (7) being -1 .

Observe that with respect to the input vector $\mathbf{x} \in \mathbb{R}^n$, the equivariant hypersphere (24) $\mathbf{F}_n(\cdot; \mathbf{S}) : \mathbb{R}^{n+2} \rightarrow \mathbb{R}^{n+2}$ represents a non-linear $E(n)$ -equivariant function. In the general case with the learnable $\gamma \in \mathbb{R}$ parameter (see Section 2.3), we define $B_{\text{hom}}(\tilde{\mathbf{S}})$ as

$$B_{\text{hom}}(\tilde{\mathbf{S}}) := \gamma \begin{bmatrix} B(\mathbf{S}) \\ \mathbf{0}^\top & -p\|\mathbf{c}\| & 0 \end{bmatrix}. \quad (26)$$

We rewrite the desired equivariance property (21) for the new $B_{\text{hom}}(\mathbf{S})$:

$$\mathbf{T}(\mathbf{t}) B_{\text{hom}}(\mathbf{S}) \mathbf{X} = B_{\text{hom}}(\mathbf{S}) \mathbf{T}_{\text{RHS}} \mathbf{X}, \quad (27)$$

where we find that

$$\mathbf{T}(\mathbf{t}) := \mathbf{M}_n^\top \mathbf{R}_O \mathbf{T}_{\text{intern}} \mathbf{R}_O^\top \mathbf{M}_n, \quad (28)$$

where, gently abusing notation, $\mathbf{M}_n \in \text{O}(n+1)$ and $\mathbf{R}_O \in \text{O}(n)$ are extended to $(n+2) \times (n+2)$ using the identity matrix, and $\mathbf{T}_{\text{interm}}$ is the $(n+2) \times (n+2)$ matrix

$$\mathbf{T}_{\text{interm}}(\mathbf{t}) = \begin{bmatrix} \mathbf{I}_n & 0 & \mathbf{t} \\ -2\alpha \mathbf{t}^\top & 1 & -\alpha \|\mathbf{t}\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix}, \quad (29)$$

where $\alpha = \frac{\sqrt{n+1}}{2p\|\mathbf{c}\|} = \frac{\sqrt{n}}{2\|\mathbf{c}\|}$.

We verify that $\mathbf{T}(\mathbf{t})$ is indeed a linear representation as necessary for the equivariance definition (5), by (a) finding its inverse

$$\mathbf{T}(\mathbf{t})^{-1} = \mathbf{M}_n^\top \mathbf{R}_O \mathbf{T}_{\text{interm}}^{-1} \mathbf{R}_O^\top \mathbf{M}_n, \quad \mathbf{T}_{\text{interm}}^{-1} = \begin{bmatrix} \mathbf{I}_n & 0 & -\mathbf{t} \\ 2\alpha \mathbf{t}^\top & 1 & -\alpha \|\mathbf{t}\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix}, \quad (30)$$

which is valid for all $\mathbf{t} \in \mathbb{R}^n$ and α with $\|\mathbf{c}\| \neq 0$ (we will consider this case below), and (b) verifying that the action of $\mathbf{T}(\mathbf{t}_2) \mathbf{T}(\mathbf{t}_1)$ results in the translation of the original $n\text{D}$ \mathbf{x} as $\mathbf{x} + \mathbf{t}_1 + \mathbf{t}_2$:

$$\mathbf{T}(\mathbf{t}_2) \mathbf{T}(\mathbf{t}_1) = \mathbf{M}_n^\top \mathbf{R}_O \mathbf{T}_{\text{interm}}(\mathbf{t}_1 + \mathbf{t}_2) \mathbf{R}_O^\top \mathbf{M}_n. \quad \square \quad (31)$$

Degenerate case $\mathbf{c} = \mathbf{0}$ Geometrically speaking, the case $\mathbf{c} = \mathbf{0}$ means that the entire regular simplex construction in $B(\mathcal{S})$ (14) collapses to a single sphere with three identical copies, all centred at $\mathbf{0}$; additionally, it means that \mathbf{R}_O (13) does not exist. In practice, we can prevent the degenerate case, e.g., by adding some negligible ϵ to $\|\mathbf{c}\|$.

3.3 COMPLETING TO $E(n)$ -EQUIVARIANCE

The desired equivariance property is

$$\mathbf{Q}(\mathbf{R}, \mathbf{t}) B_{\text{hom}}(\mathcal{S}) \mathbf{X} = B_{\text{hom}}(\mathcal{S}) \mathbf{Q}_{\text{RHS}} \mathbf{X}, \quad (32)$$

where $\mathbf{Q}_{\text{RHS}} := \mathbf{Q}_X$ (9).

By using an identity complement and extending \mathbf{V} (16) to $(n+2) \times (n+2)$ on the LHS in (15), we note that $B_{\text{hom}}(\mathcal{S})$ (24) is also $\text{O}(n)$ -equivariant: since the RHS in (15) can be rewritten as

$$B_{\text{hom}}(\mathcal{S}) \mathbf{R} \mathbf{X} = \begin{bmatrix} B(\mathcal{S}) \mathbf{R} \mathbf{X} \\ p\|\mathbf{c}\| \end{bmatrix}, \quad (33)$$

where $p\|\mathbf{c}\| = \sqrt{\frac{n+1}{n}} \|\mathbf{c}\|$ and $\|\mathbf{x}\|^2 = \|\mathbf{R} \mathbf{x}\|^2$ are $\text{O}(n)$ -invariant. The same is valid for the LHS, where the additional $(n+2)$ -th element in the resulting vector is also $p\|\mathbf{c}\|$.

In turn, the RHS of (32) is

$$\mathbf{y}_{\mathbf{R}\mathbf{t}} := B_{\text{hom}}(\mathcal{S}) \mathbf{Q}_{\text{RHS}} \mathbf{X} = \begin{bmatrix} B(\mathcal{S}) \mathbf{Q}_{\text{RHS}} \mathbf{X} \\ p\|\mathbf{c}\| \end{bmatrix}, \quad (34)$$

where using the decomposition of \mathbf{y} (18), we note that

$$\mathbf{R}_O^\top \mathbf{M}_n \mathbf{y}_{\mathbf{R}\mathbf{t}} = p\|\mathbf{c}\| \begin{bmatrix} \mathbf{R} \mathbf{x} + \mathbf{t} \\ \alpha(r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x} + \mathbf{t}\|^2) \\ 1 \end{bmatrix}, \quad (35)$$

where $\|\mathbf{x} + \mathbf{t}\|^2 = \|\mathbf{R} \mathbf{x} + \mathbf{t}\|^2$.

Keeping the extended $(n+2) \times (n+2)$ versions of $\mathbf{M}_n \in \text{O}(n+1)$, $\mathbf{R}_O \in \text{O}(n)$, $\mathbf{V} \in G < \text{O}(n+1)$, using the translation representation \mathbf{T} (28) derived earlier, as well as $\mathbf{y}_{\mathbf{R}\mathbf{t}}$, and following the same order of operations on the LHS as on the RHS of (32) (first an $\text{O}(n)$ action, and then translation), we find that

$$\mathbf{T} \mathbf{V} B_{\text{hom}}(\mathcal{S}) \mathbf{X} = B_{\text{hom}}(\mathcal{S}) \mathbf{Q}_{\text{RHS}} \mathbf{X}. \quad \square \quad (36)$$

Thus, we have proved (32) for

$$\mathbf{Q}(\mathbf{R}, \mathbf{t}) := \mathbf{T}(\mathbf{t}) \mathbf{V}(\mathbf{R}) = \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{R} & 0 & \mathbf{t} \\ -2\alpha \mathbf{t}^\top \mathbf{R} & 1 & -\alpha \|\mathbf{t}\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n, \quad (37)$$

where \mathbf{M}_n , \mathbf{R}_O , and \mathbf{V} (16) are extended to $(n+2) \times (n+2)$ using the identity complement, \mathbf{T} is given by (28), and $\alpha = \frac{\sqrt{n+1}}{2p \|\mathbf{c}\|} = \frac{\sqrt{n}}{2\|\mathbf{c}\|}$.

Similarly to what we did with $\mathbf{T}(\mathbf{t})$ in Section 3.2, we verify that $\mathbf{Q}(\mathbf{R}, \mathbf{t})$ is a linear representation by (a) finding its inverse

$$\mathbf{Q}(\mathbf{R}, \mathbf{t})^{-1} = \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{R}^\top & 0 & -\mathbf{R}^\top \mathbf{t} \\ 2\alpha \mathbf{t}^\top & 1 & -\alpha \|\mathbf{t}\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n, \quad (38)$$

which is valid for all \mathbf{R} in $O(n)$, $\mathbf{t} \in \mathbb{R}^n$ and α with $\|\mathbf{c}\| \neq 0$ (which is the degenerate case considered in Section 3.2), and (b) confirming that the action of $\mathbf{Q}(\mathbf{R}_2, \mathbf{t}_2) \mathbf{Q}(\mathbf{R}_1, \mathbf{t}_1)$ results in the transformation of the original nD \mathbf{x} as $\mathbf{R}_2(\mathbf{R}_1 \mathbf{x} + \mathbf{t}_1) + \mathbf{t}_2 = \mathbf{R}_2 \mathbf{R}_1 \mathbf{x} + \mathbf{R}_2 \mathbf{t}_1 + \mathbf{t}_2$, i.e.,

$$\mathbf{Q}(\mathbf{R}_2, \mathbf{t}_2) \mathbf{Q}(\mathbf{R}_1, \mathbf{t}_1) = \mathbf{Q}(\mathbf{R}_2 \mathbf{R}_1, \mathbf{R}_2 \mathbf{t}_1 + \mathbf{t}_2). \quad \square \quad (39)$$

This completes a constructive existence proof of an explicit $E(n)$ -equivariant representation for spherical decision surfaces, without relying on input centring or pairwise coordinate differences.

3.4 INVARIANT OPERATORS

For the $E(n)$ -equivariant feature in our case (25), we can no longer take the inner product of the input with itself to obtain an invariant feature as it is possible with $O(n)$ -equivariant features (as performed in Melnyk et al. (2024a;b)), and thus require other invariant operators.

Canonical triplets and translation law Within the neuron-specific (indexed with k) canonical frame, i.e., after the change of basis $\mathbf{R}_O^{k\top} \mathbf{M}_n(34) \rightarrow (35)$ and normalisation by $p\|\mathbf{c}_k\|$, the homogeneous feature for an input point $\mathbf{x}_i \in \mathbb{R}^n$ splits into a vector–scalar triple $[\mathbf{x}_i; s_{ik}; 1]$. Under a global translation $\mathbf{x}_i \mapsto \mathbf{x}_i + \mathbf{t}$, the equivariant representation acts as

$$\mathbf{x}_i \mapsto \mathbf{x}_i + \mathbf{t}, \quad s_{ik} \mapsto s_{ik} - 2\alpha_k \mathbf{t}^\top \mathbf{x}_i - \alpha_k \|\mathbf{t}\|^2, \quad (40)$$

where $\alpha_k = \frac{\sqrt{n}}{2\|\mathbf{c}_k\|}$ and $s_{ik} = \alpha_k(r_k^2 - \|\mathbf{c}_k\|^2 - \|\mathbf{x}_i\|^2)$.

Single-point invariant A per-neuron scalar

$$I_{i,k} := s_{ik} + \alpha_k \|\mathbf{x}_i\|^2 = \alpha_k (r_k^2 - \|\mathbf{c}_k\|^2) \quad (41)$$

thus remains invariant under translations and rotations/reflections. It serves as a learned bias/radius term for normalisation or gating across neurons.

Pairwise invariant and distance form For two points i, j under the same neuron k , we define

$$I_{ij,k} := s_{ik} + s_{jk} + 2\alpha_k \mathbf{x}_i^\top \mathbf{x}_j, \quad (42)$$

which is exactly translation-invariant; expanding s_{ik} yields the rotation/reflection-invariant distance form

$$I_{ij,k} = \alpha_k \left(2r_k^2 - 2\|\mathbf{c}_k\|^2 - \|\mathbf{x}_i - \mathbf{x}_j\|^2 \right), \quad (43)$$

so $I_{ij,k}$ depends only on the pairwise distance $\|\mathbf{x}_i - \mathbf{x}_j\|$ and neuron parameters \mathbf{c}_k and r_k .

Table 1: Downstream classification accuracy on the 3D Tetris dataset under unseen $E(n)$ transformations. Chance level is $1/7 \approx 14.3\%$.

Model	Test Accuracy (%)
MLP	14.7
MLP+aug	15.9
MLP+centr+aug	36.5
Ours	100.0

3.5 STRUCTURAL IMPLICATIONS

The results of this work are best interpreted at the level of individual geometric constructions. In particular, enforcing exact $E(n)$ -equivariance for non-linear spherical decision surfaces leads to representations $\mathbf{Q}(\mathbf{R}, \mathbf{t})$ (37) that depend explicitly on the learned geometric parameters $\mathbf{c}_k \in \mathbb{R}^n$ of each neuron. This dependence is intrinsic to the construction and reflects the fact that the geometry of the decision surface itself is learned. As a consequence, the output vector $\begin{bmatrix} \mathbf{y} \\ p \|\mathbf{c}\| \end{bmatrix} \in \mathbb{R}^{n+2}$ (25) is neither in the same space as the input $\mathbf{x} \in \mathbb{R}^n$ nor in the conformal space in which the embedding $\mathbf{X} \in \mathbb{R}^{n+2}$ (7) resides (since the last two parameters of the output are not as they should be in the embedding). For this reason, $B_{\text{hom}}(\tilde{\mathcal{S}})$ cannot be interpreted as a drop-in neural network layer whose outputs can be directly stacked or composed in the usual way. We therefore view $B_{\text{hom}}(\tilde{\mathcal{S}})$ primarily as a theoretically well-defined geometric primitive rather than a drop-in neural network layer. Possible ways of integrating such primitives into larger equivariant systems (e.g., via canonicalisation and invariant extraction) can be envisaged but are not pursued further in this work.

4 DEMONSTRATION

Numerical verification As a sanity check of the derived operators, we use PyTorch (Paszke et al., 2019) and numerically validate the identities from Section 3 by sampling random $\mathbf{x} \in \mathbb{R}^n$ (where required, \mathbf{x}_1 and \mathbf{x}_2), $\mathbf{R} \in O(n)$, $\mathbf{t} \in \mathbb{R}^n$, $\mathbf{c} \in \mathbb{R}^n$, and $r \in \mathbb{R}$, and reporting the distribution of \log_{10} residual norms for each check. Each check is performed 1000 times for each dimension $n = \{2, 3, 4, 5\}$. The resulting histograms presented in Appendix B (Figure 1) show that residuals concentrate at machine precision, supporting the correctness of our derived operators.

3D Tetris data classification We then perform a minimal proof-of-concept experiment verifying that the derived $E(n)$ -invariant primitives (42) behave as predicted in practice. We solve a classification task on a modified version of the 3D Tetris dataset used by Thomas et al. (2018): we remove the mirrored chiral copy to avoid reflection-sensitivity since our demonstration targets $E(n)$ -invariance. Each sample is a 4-point shape in \mathbb{R}^3 , and 7 distinct shapes in total constitute the entire training set. The test set consists of the Tetris shapes subjected to random rotations/reflections and translations, with 7000 samples in total.

We build our model using one $E(3)$ -invariant bank with $K = 10$ neurons (26) (with randomly initialised \mathcal{S}_k and each γ_k initially set to 1), producing closed-form invariants (42), followed by a single linear head classifying from these invariants, in total containing 127 parameters. As a baseline, we construct a 3-layer multilayer perceptron (MLP) with ReLU activations and the same head size, yielding a 191-parameter network. We train three versions of this MLP: (i) as-is (MLP), (ii) with on-the-fly $E(3)$ -augmentation (MLP+aug), and (iii) with centring the input and using $O(3)$ -augmentation (MLP+centr+aug). All the models follow the per-point DeepSet (Zaheer et al., 2017) construction, and are therefore permutation-invariant (we use mean pooling). We train the models for 1000 epochs by minimising the cross-entropy and using the Adam optimiser with the default hyperparameters. The results in Table 1 show that the derived invariant primitives behave as theoretically expected under $E(n)$ transformations, while standard MLPs rely on additional mechanisms to achieve invariance..

5 CONCLUSION

In this work, we presented a constructive derivation of exactly $E(n)$ -equivariant spherical decision surfaces, focusing on explicit representations and geometric correctness rather than architectural design. We hope this work serves as a theoretical reference point for future investigations into geometric decision surfaces and equivariant representations.

432 IMPACT STATEMENT
433

434 This paper presents work aiming to advance the field of machine learning. There are many potential
435 societal consequences of our work, none of which we feel must be specifically highlighted here,
436 perhaps apart from materials science applications, where the development of new materials might
437 have a significant impact on sustainability.
438

439 REFERENCES
440

441 Serge Assaad, C. Downey, Rami Al-Rfou, Nigamaa Nayakanti, and Benjamin Sapp. Vn-transformer:
442 Rotation-equivariant attention for vector neurons. *TMLR*, 2023.
443

444 Hakan Cevikalp and Hasan Saribas. Deep simplex classifier for maximizing the margin in both
445 euclidean and angular spaces. In *Image Analysis: 23rd Scandinavian Conference, SCIA 2023,*
446 *Sirkka, Finland, April 18–21, 2023, Proceedings, Part II*, pp. 91–107. Springer, 2023.
447

448 Congyue Deng, Or Litany, Yueqi Duan, Adrien Poulenard, Andrea Tagliasacchi, and Leonidas J
449 Guibas. Vector neurons: A general framework for SO(3)-equivariant networks. In *Proceedings of*
450 *the IEEE/CVF International Conference on Computer Vision*, pp. 12200–12209, 2021.

451 Emanuel Lodewijk Elte. *The Semiregular Polytopes of the Hyperspaces*. United States: University of
452 Michigan Press, 2006. ISBN 9781418179687. URL [https://books.google.se/books?](https://books.google.se/books?id=9SxtPwAACAAJ)
453 [id=9SxtPwAACAAJ](https://books.google.se/books?id=9SxtPwAACAAJ).

454 Carlos Esteves, Christine Allen-Blanchette, Ameesh Makadia, and Kostas Daniilidis. Learning so(3)
455 equivariant representations with spherical cnns. *CoRR*, 2017. URL [http://arxiv.org/abs/](http://arxiv.org/abs/1711.06721)
456 [1711.06721](http://arxiv.org/abs/1711.06721).
457

458 Fabian Fuchs, Daniel Worrall, Volker Fischer, and Max Welling. Se(3)-transformers: 3d roto-
459 translation equivariant attention networks. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan,
460 and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 1970–1981.
461 Curran Associates, Inc., 2020.
462

463 Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the
464 Mathematical Sciences. Johns Hopkins University Press, 4 edition, 2013. ISBN 978-1-4214-0794-
465 4. doi: 10.56021/9781421407944.

466 Hongbo Li, David Hestenes, and Alyn Rockwood. Generalized homogeneous coordinates for
467 computational geometry. In *Geometric Computing with Clifford Algebras*, pp. 27–59. Springer,
468 2001.
469

470 Pavlo Melnyk, Michael Felsberg, and Mårten Wadenbäck. Embed Me if You Can: A Geometric
471 Perceptron. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*,
472 pp. 1276–1284, October 2021.
473

474 Pavlo Melnyk, Michael Felsberg, and Mårten Wadenbäck. Steerable 3D Spherical Neurons. In
475 *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Pro-*
476 *ceedings of Machine Learning Research*, pp. 15330–15339. PMLR, 17–23 Jul 2022. URL
477 <https://proceedings.mlr.press/v162/melnyk22a.html>.

478 Pavlo Melnyk, Michael Felsberg, Mårten Wadenbäck, Andreas Robinson, and Cuong Le. On Learning
479 Deep O(n)-Equivariant Hyperspheres. In *Proceedings of the 41st International Conference on*
480 *Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 35324–35339.
481 PMLR, 21–27 Jul 2024a. URL [https://proceedings.mlr.press/v235/melnyk24a.](https://proceedings.mlr.press/v235/melnyk24a.html)
482 [html](https://proceedings.mlr.press/v235/melnyk24a.html).
483

484 Pavlo Melnyk, Andreas Robinson, Michael Felsberg, and Mårten Wadenbäck. TetraSphere: A Neural
485 Descriptor for O(3)-Invariant Point Cloud Analysis. In *Proceedings of the IEEE/CVF Conference*
on Computer Vision and Pattern Recognition (CVPR), pp. 5620–5630, June 2024b.

- 486 Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor
487 Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style,
488 high-performance deep learning library. In *Advances in Neural Information Processing Systems*,
489 pp. 8024–8035, 2019.
- 490 Christian Perwass, Vladimir Banarer, and Gerald Sommer. Spherical decision surfaces using confor-
491 mal modelling. In *Joint Pattern Recognition Symposium*, pp. 9–16. Springer, 2003.
- 492
- 493 David Ruhe, Johannes Brandstetter, and Patrick Forré. Clifford Group Equivariant Neural Networks.
494 In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. URL <https://openreview.net/forum?id=n84bzMrGUD>.
- 495
- 496 Victor Garcia Satorras, Emiel Hooeboom, and Max Welling. E(n) equivariant graph neural networks.
497 In *International conference on machine learning*, pp. 9323–9332. PMLR, 2021.
- 498
- 499 Nathaniel Thomas, Tess Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley.
500 Tensor field networks: Rotation-and translation-equivariant neural networks for 3D point clouds.
501 *arXiv preprint arXiv:1802.08219*, 2018.
- 502
- 503 Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabás Póczos, Ruslan Salakhutdinov, and
504 Alexander Smola. Deep sets. In *Advances in Neural Information Processing Systems*, 2017.
- 505
- 506
- 507
- 508
- 509
- 510
- 511
- 512
- 513
- 514
- 515
- 516
- 517
- 518
- 519
- 520
- 521
- 522
- 523
- 524
- 525
- 526
- 527
- 528
- 529
- 530
- 531
- 532
- 533
- 534
- 535
- 536
- 537
- 538
- 539

A COMPLETE DERIVATIONS

In this section, we present complete derivation and proofs, for convenience, following the same organisation as Section 3 and restating expressions.

A.1 IDENTITY SPLIT: EQUIVARIANT HYPERSPHERE OUTPUT DECOMPOSITION INTO IRREPS

Given $\mathbf{X} \in \mathbb{R}^{n+2}$ (7) and $\mathbf{y} := B(\mathbf{S})\mathbf{X} \in \mathbb{R}^{n+1}$, we decompose it into $O(n)$ -equivariant and invariant parts using (14) as

$$\begin{aligned}
\mathbf{R}_O^\top \mathbf{M}_n \mathbf{y} &= \mathbf{R}_O^\top \mathbf{M}_n B(\mathbf{S})\mathbf{X} = \mathbf{R}_O^\top \mathbf{M}_n \left[\begin{bmatrix} \|\mathbf{c}\| \mathbf{P}_n^\top & s_{n+1} \mathbf{1} \end{bmatrix} \mathbf{R}_O \quad \mathbf{1} \right] \mathbf{X} \\
&= \mathbf{R}_O^\top \left[\begin{bmatrix} p \|\mathbf{c}\| \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0}^\top \end{bmatrix} & \mathbf{0} \\ p\sqrt{n} s_{n+1} & \mathbf{R}_O \quad \mathbf{0} \\ & p\sqrt{n} \end{bmatrix} \right] \mathbf{X} \\
&= \left[\begin{bmatrix} p \|\mathbf{c}\| \mathbf{I}_n & \mathbf{0} \\ \mathbf{0}^\top & p\sqrt{n} s_{n+1} \end{bmatrix} \mathbf{R}_O^\top \mathbf{R}_O \quad \mathbf{0} \\ & p\sqrt{n} \end{bmatrix} \mathbf{X} = \left[\begin{bmatrix} p \|\mathbf{c}\| \mathbf{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0}^\top & p\sqrt{n} s_{n+1} & p\sqrt{n} \end{bmatrix} \right] \mathbf{X} \quad (44) \\
&= \left[\begin{bmatrix} p \|\mathbf{c}\| \mathbf{I}_n & \mathbf{0} \\ \mathbf{0}^\top & p\sqrt{n} \frac{1}{2} (\|\mathbf{c}\|^2 - r^2) \end{bmatrix} \mathbf{R}_O^\top \mathbf{R}_O \quad \mathbf{0} \\ & p\sqrt{n} \end{bmatrix} \mathbf{X} = \left[\begin{bmatrix} p \|\mathbf{c}\| \mathbf{x} \\ \frac{p\sqrt{n}}{2} (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \end{bmatrix} \right] \\
&= p \|\mathbf{c}\| \left[\begin{bmatrix} \mathbf{x} \\ \frac{\sqrt{n}}{2\|\mathbf{c}\|} (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \end{bmatrix} \right] = p \|\mathbf{c}\| \left[\begin{bmatrix} \mathbf{x} \\ \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \end{bmatrix} \right],
\end{aligned}$$

where $\alpha := \frac{\sqrt{n+1}}{2p\|\mathbf{c}\|} = \frac{\sqrt{n}}{2\|\mathbf{c}\|}$, and the invariant part is $\frac{\sqrt{n+1}}{2} (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2)$, with \mathbf{c} and r being learnable parameters.

Observe that in the case $\|\mathbf{c}\| = 0$, the $O(n)$ -equivariant part, i.e., the original vector \mathbf{x} , zeroes out, and what we are left with is

$$\mathbf{R}_O^\top \mathbf{M}_n \mathbf{y} = \mathbf{R}_O^\top \mathbf{M}_n B(\mathbf{S})\mathbf{X} = \left[\begin{bmatrix} p \mathbf{0} \mathbf{x} \\ \frac{p\sqrt{n}}{2} (r^2 - 0 - \|\mathbf{x}\|^2) \end{bmatrix} \right] = \left[\begin{bmatrix} \mathbf{0} \\ \frac{p\sqrt{n}}{2} (r^2 - \|\mathbf{x}\|^2) \end{bmatrix} \right]. \quad (45)$$

Also, if \mathbf{X} is transformed with $\mathbf{R} \in O(n)$ extended to $(n+2) \times (n+2)$ using the identity matrix, as $\mathbf{R}\mathbf{X}$, defining $\mathbf{y}_R := B(\mathbf{S})\mathbf{R}\mathbf{X}$, we obtain

$$\mathbf{R}_O^\top \mathbf{M}_n \mathbf{y}_R = \mathbf{R}_O^\top \mathbf{M}_n B(\mathbf{S})\mathbf{R}\mathbf{X} = p \|\mathbf{c}\| \left[\begin{bmatrix} \mathbf{R}\mathbf{x} \\ \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \end{bmatrix} \right]. \quad (46)$$

A.2 TRANSLATION EQUIVARIANCE

The desired equivariance property is

$$\mathbf{T}(\mathbf{t}) B(\mathbf{S})\mathbf{X} = B(\mathbf{S})\mathbf{T}_{\text{RHS}}\mathbf{X}, \quad (47)$$

where $\mathbf{T}(\mathbf{t})$ is the translation representation we seek and

$$\mathbf{T}_{\text{RHS}} = \begin{bmatrix} \mathbf{I}_n & -\mathbf{t} & \mathbf{0} \\ \mathbf{0}^\top & 1 & 0 \\ -\mathbf{t}^\top & \frac{1}{2}\|\mathbf{t}\|^2 & 1 \end{bmatrix} \quad (48)$$

is the $(n+2) \times (n+2)$ matrix (9) with $\mathbf{R} = \mathbf{I}_n$.

Consider the RHS and denote \mathbf{y}_t to be the output of the equivariant hypersphere given the input \mathbf{x} translated as $\mathbf{x} + \mathbf{t}$ and embedded in \mathbb{R}^{n+2} (see (7)) as \mathbf{X}_t :

$$\mathbf{y}_t := B(\mathbf{S})\mathbf{X}_t = B(\mathbf{S})\mathbf{T}_{\text{RHS}}(\mathbf{t})\mathbf{X} \in \mathbb{R}^{n+1}. \quad (49)$$

In the decomposition of \mathbf{y} (44), we can left-multiply both sides with some $(n+1) \times (n+1)$ matrix $\mathbf{T}_{\text{interm}}$:

$$\mathbf{T}_{\text{interm}} \mathbf{R}_O^T \mathbf{M} \mathbf{y} = p \|\mathbf{c}\| \mathbf{T}_{\text{interm}} \begin{bmatrix} \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \\ \mathbf{x} \end{bmatrix}. \quad (50)$$

From this decomposition, it is clear that in order to act as a translation, \mathbf{t} , on the original \mathbf{x} , $\mathbf{T}_{\text{interm}}$ must transform the expression in the following manner:

$$\begin{aligned} p \|\mathbf{c}\| \mathbf{T}_{\text{interm}} \begin{bmatrix} \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \\ \mathbf{x} \end{bmatrix} &= p \|\mathbf{c}\| \begin{bmatrix} \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x} + \mathbf{t}\|^2) \\ \mathbf{x} + \mathbf{t} \end{bmatrix} \\ &= \mathbf{R}_O^T \mathbf{M}_n \mathbf{y}_{\mathbf{t}} = \mathbf{R}_O^T \mathbf{M}_n B(S) \mathbf{T}_{\text{RHS}} \mathbf{X}. \end{aligned} \quad (51)$$

It is achievable with an *affine* transformation, i.e., if instead of left multiplying with $\mathbf{T}_{\text{interm}}$, we added $p \|\mathbf{c}\| \begin{bmatrix} \mathbf{t} \\ -2\alpha \mathbf{t}^T \mathbf{x} - \alpha \|\mathbf{t}\|^2 \end{bmatrix}$ to the decomposition. This would, however, need to depend on the data vector \mathbf{x} . However, this is *impossible* with a linear transformation, that is, in the original $(n+1)$ D space, which means there is not such a translation representation $\mathbf{T}(\mathbf{t}) \in \mathbb{R}^{(n+1) \times (n+1)}$ to satisfy (47).

Homogeneous representation to the rescue If in (51) we instead use a homogeneous representation

$$p \|\mathbf{c}\| \begin{bmatrix} \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \\ \mathbf{x} \\ 1 \end{bmatrix}, \quad (52)$$

resulting in

$$\begin{aligned} p \|\mathbf{c}\| \mathbf{T}_{\text{interm}} \begin{bmatrix} \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \\ \mathbf{x} \\ 1 \end{bmatrix} &= p \|\mathbf{c}\| \begin{bmatrix} \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x} + \mathbf{t}\|^2) \\ \mathbf{x} + \mathbf{t} \\ 1 \end{bmatrix} \\ &= p \|\mathbf{c}\| \begin{bmatrix} \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2 - 2\|\mathbf{x}\| \mathbf{t}^T - \|\mathbf{t}\|^2) \\ \mathbf{x} + \mathbf{t} \\ 1 \end{bmatrix}, \end{aligned} \quad (53)$$

we can find a linear transformation $\mathbf{T}_{\text{interm}}$ that satisfies this equation.

Solving for each row of $\mathbf{T}_{\text{interm}}$ in (53), we obtain the $(n+2) \times (n+2)$

$$\mathbf{T}_{\text{interm}}(\mathbf{t}) = \begin{bmatrix} \mathbf{I}_n & 0 & \mathbf{t} \\ -2\alpha \mathbf{t}^T & 1 & -\alpha \|\mathbf{t}\|^2 \\ \mathbf{0}^T & 0 & 1 \end{bmatrix}, \quad (54)$$

where $\alpha = \frac{\sqrt{n+1}}{2p \|\mathbf{c}\|} = \frac{\sqrt{n}}{2 \|\mathbf{c}\|}$.

Subsequently, we can write the decomposition (50) backwards as

$$\begin{aligned} p \|\mathbf{c}\| \mathbf{T}_{\text{interm}} \begin{bmatrix} \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \\ \mathbf{x} \\ 1 \end{bmatrix} &= p \|\mathbf{c}\| \mathbf{T}_{\text{interm}} \begin{bmatrix} \mathbf{R}_O^T \mathbf{M} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \frac{\mathbf{y}}{p \|\mathbf{c}\|} \\ 1 \end{bmatrix} \\ &= \mathbf{T}_{\text{interm}} \begin{bmatrix} \mathbf{R}_O^T \mathbf{M}_n & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} B(S) \mathbf{X} \\ p \|\mathbf{c}\| \end{bmatrix}. \end{aligned} \quad (55)$$

To obtain the homogeneous representation of \mathbf{y} in the equation above (55), $\begin{bmatrix} \mathbf{y} \\ p \|\mathbf{c}\| \end{bmatrix}$, we need to modify the original $(n+1) \times (n+2)$ equivariant hypersphere $B(\mathbf{S})$ (12), and create the $(n+2) \times (n+2)$ $B_{\text{hom}}(\mathbf{S})$:

$$B_{\text{hom}}(\mathbf{S}) := \begin{bmatrix} \mathbf{0}^\top & B(\mathbf{S}) & \\ & -p \|\mathbf{c}\| & 0 \end{bmatrix}, \quad (56)$$

so that

$$B_{\text{hom}}(\mathbf{S}) \mathbf{X} = \begin{bmatrix} B(\mathbf{S}) \mathbf{X} \\ p \|\mathbf{c}\| \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ p \|\mathbf{c}\| \end{bmatrix} \in \mathbb{R}^{n+2}, \quad (57)$$

where $\mathbf{c} \in \mathbb{R}^n$ are parameters in \mathbf{S} , and hence no new learnable parameters are added, and the $-$ sign in the $-p \|\mathbf{c}\|$ element in $B_{\text{hom}}(\mathbf{S})$ is due to the second last element of the embedding \mathbf{X} (7) being -1 .

Observe that with respect to the input vector $\mathbf{x} \in \mathbb{R}^n$, the equivariant hypersphere (56) $\mathbf{F}_n(\cdot; \mathbf{S}) : \mathbb{R}^{n+2} \rightarrow \mathbb{R}^{n+2}$ represents a non-linear $E(n)$ -equivariant function. In the general case with the learnable $\gamma \in \mathbb{R}$ parameter (see Section 2.3), we define $B_{\text{hom}}(\tilde{\mathbf{S}})$ as

$$B_{\text{hom}}(\tilde{\mathbf{S}}) := \gamma \begin{bmatrix} B(\mathbf{S}) & \\ \mathbf{0}^\top & -p \|\mathbf{c}\| & 0 \end{bmatrix}. \quad (58)$$

We rewrite the desired equivariance property (47) for the new $B_{\text{hom}}(\mathbf{S})$:

$$\mathbf{T}(\mathbf{t}) B_{\text{hom}}(\mathbf{S}) \mathbf{X} = B_{\text{hom}}(\mathbf{S}) \mathbf{T}_{\text{RHS}} \mathbf{X}. \quad (59)$$

We first start with the LHS: To complete the $(n+2) \times (n+2)$ representation $\mathbf{T}(\mathbf{t})$ on the LHS of (59), we complete the last expression of (55), i.e., rotate the $(n+1)$ D part back to the original coordinate system of \mathbf{y} using $\mathbf{M}_n^\top \mathbf{R}_O$:

$$\begin{aligned} \mathbf{T}(\mathbf{t}) B_{\text{hom}}(\mathbf{S}) \mathbf{X} &= \begin{bmatrix} \mathbf{M}_n^\top \mathbf{R}_O & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} \mathbf{T}_{\text{interm}} \begin{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} B_{\text{hom}}(\mathbf{S}) \mathbf{X} \\ &= p \|\mathbf{c}\| \begin{bmatrix} \mathbf{M}_n^\top \mathbf{R}_O & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & 0 & \mathbf{t} \\ -2\alpha \mathbf{t}^\top & 1 & -\alpha \|\mathbf{t}\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \\ \mathbf{x} \\ 1 \end{bmatrix} \\ &= p \|\mathbf{c}\| \begin{bmatrix} \mathbf{M}_n^\top \mathbf{R}_O & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x} + \mathbf{t} \\ \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x} + \mathbf{t}\|^2) \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{M}_n^\top \mathbf{R}_O & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n \mathbf{y}_t \\ p \|\mathbf{c}\| \end{bmatrix} = \begin{bmatrix} \mathbf{M}_n^\top \mathbf{R}_O & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{y}_t \\ p \|\mathbf{c}\| \end{bmatrix} \\ &= \mathbf{I}_{n+2} \begin{bmatrix} B(\mathbf{S}) \mathbf{T}_{\text{RHS}} \mathbf{X} \\ p \|\mathbf{c}\| \end{bmatrix} = B_{\text{hom}}(\mathbf{S}) \mathbf{T}_{\text{RHS}} \mathbf{X}, \quad \square \end{aligned} \quad (60)$$

where $\mathbf{y}_t \in \mathbb{R}^{n+1}$ is defined in (49), effectively proving (59) for $B_{\text{hom}}(\mathbf{S})$ (56) and

$$\mathbf{T}(\mathbf{t}) := \mathbf{M}_n^\top \mathbf{R}_O \mathbf{T}_{\text{interm}} \mathbf{R}_O^\top \mathbf{M}_n, \quad (61)$$

where $\mathbf{T}_{\text{interm}}$ is given by (54) and, gently abusing notation, $\mathbf{M}_n \in O(n+1)$ and $\mathbf{R}_O \in O(n)$ are extended to $(n+2) \times (n+2)$ using the identity matrix.

We verify that $\mathbf{T}(\mathbf{t})$ is indeed a linear representation as necessary for the equivariance definition (5), by (a) finding its inverse

$$\begin{aligned} \mathbf{T}(\mathbf{t})^{-1} &= \mathbf{M}_n^\top \mathbf{R}_O \mathbf{T}_{\text{interm}}^{-1} \mathbf{R}_O^\top \mathbf{M}_n, \\ \mathbf{T}_{\text{interm}}^{-1} &= \begin{bmatrix} \mathbf{I}_n & 0 & -\mathbf{t} \\ 2\alpha \mathbf{t}^\top & 1 & -\alpha \|\mathbf{t}\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix}, \end{aligned} \quad (62)$$

702 which is valid for all $\mathbf{t} \in \mathbb{R}^n$ and α with $\|\mathbf{c}\| \neq 0$, and (b) verifying that the action of $\mathbf{T}(\mathbf{t}_2)\mathbf{T}(\mathbf{t}_1)$
 703 results in the translation of the original nD \mathbf{x} as $\mathbf{x} + \mathbf{t}_1 + \mathbf{t}_2$:
 704

$$\begin{aligned}
 705 & \mathbf{T}(\mathbf{t}_2)\mathbf{T}(\mathbf{t}_1) = \mathbf{M}_n^\top \mathbf{R}_O \mathbf{T}_{\text{interm}}(\mathbf{t}_2) \mathbf{R}_O^\top \mathbf{M}_n \mathbf{M}_n^\top \mathbf{R}_O \mathbf{T}_{\text{interm}}(\mathbf{t}_1) \mathbf{R}_O^\top \mathbf{M}_n \\
 706 & = \mathbf{M}_n^\top \mathbf{R}_O \mathbf{T}_{\text{interm}}(\mathbf{t}_2) \mathbf{T}_{\text{interm}}(\mathbf{t}_1) \mathbf{R}_O^\top \mathbf{M}_n \\
 707 & = \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{I}_n & 0 & \mathbf{t}_2 \\ -2\alpha \mathbf{t}_2^\top & 1 & -\alpha \|\mathbf{t}_2\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & 0 & \mathbf{t}_1 \\ -2\alpha \mathbf{t}_1^\top & 1 & -\alpha \|\mathbf{t}_1\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n \\
 708 & = \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{I}_n & 0 & \mathbf{t}_1 + \mathbf{t}_2 \\ -2\alpha(\mathbf{t}_1 + \mathbf{t}_2)^\top & 1 & -\alpha \|\mathbf{t}_1 + \mathbf{t}_2\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n \\
 709 & = \mathbf{M}_n^\top \mathbf{R}_O \mathbf{T}_{\text{interm}}(\mathbf{t}_1 + \mathbf{t}_2) \mathbf{R}_O^\top \mathbf{M}_n . \quad \square \\
 710 & \tag{63}
 \end{aligned}$$

711 A.3 COMPLETING TO $E(n)$ -EQUIVARIANCE

712 The desired equivariance property is

$$713 \mathbf{Q}(\mathbf{R}, \mathbf{t}) B_{\text{hom}}(\mathbf{S}) \mathbf{X} = B_{\text{hom}}(\mathbf{S}) \mathbf{Q}_{\text{RHS}} \mathbf{X}, \tag{64}$$

714 where $\mathbf{Q}_{\text{RHS}} := \mathbf{Q}_X$ (9).

715 By using an identity complement and extending V (16) to $(n+2) \times (n+2)$ on the LHS in (15), we
 716 note that $B_{\text{hom}}(\mathbf{S})$ (24) is also $O(n)$ -equivariant: since the RHS in (15) can be rewritten as

$$717 B_{\text{hom}}(\mathbf{S}) \mathbf{R} \mathbf{X} = \begin{bmatrix} \mathbf{0}^\top & B(\mathbf{S}) & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} \mathbf{x} \\ -1 \\ -\frac{1}{2} \|\mathbf{x}\|^2 \end{bmatrix} = \begin{bmatrix} B(\mathbf{S}) \mathbf{R} \mathbf{X} \\ p \|\mathbf{c}\| \end{bmatrix}, \tag{65}$$

718 where $p \|\mathbf{c}\| = \sqrt{\frac{n+1}{n}} \|\mathbf{c}\|$ and $\|\mathbf{x}\|^2 = \|\mathbf{R} \mathbf{x}\|^2$ are $O(n)$ -invariant. The same is valid for the LHS,
 719 where the additional $(n+2)$ -th element in the resulting vector is also $p \|\mathbf{c}\|$.

720 In turn, the RHS of (64) is

$$721 \mathbf{y}_{\mathbf{R}\mathbf{t}} := B_{\text{hom}}(\mathbf{S}) \mathbf{Q}_{\text{RHS}} \mathbf{X} = \begin{bmatrix} \mathbf{0}^\top & B(\mathbf{S}) & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} \mathbf{x} + \mathbf{t} \\ -1 \\ -\frac{1}{2} \|\mathbf{x} + \mathbf{t}\|^2 \end{bmatrix} = \begin{bmatrix} B(\mathbf{S}) \mathbf{Q}_{\text{RHS}} \mathbf{X} \\ p \|\mathbf{c}\| \end{bmatrix}, \tag{66}$$

722 where using the decomposition of \mathbf{y} (44), we note that

$$723 \mathbf{R}_O^\top \mathbf{M}_n \mathbf{y}_{\mathbf{R}\mathbf{t}} = \mathbf{R}_O^\top \mathbf{M}_n B(\mathbf{S}) \mathbf{Q}_{\text{RHS}} \mathbf{X} = p \|\mathbf{c}\| \begin{bmatrix} \mathbf{R} \mathbf{x} + \mathbf{t} \\ \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x} + \mathbf{t}\|^2) \end{bmatrix}, \tag{67}$$

724 where $\|\mathbf{x} + \mathbf{t}\|^2 = \|\mathbf{R} \mathbf{x} + \mathbf{t}\|^2$.

725 Keeping the extended $(n+2) \times (n+2)$ versions of $\mathbf{M}_n \in O(n+1)$, $\mathbf{R}_O \in O(n)$, $\mathbf{V} \in G < O(n+1)$,
 using the translation representation \mathbf{T} (61) derived earlier, as well as $\mathbf{y}_{\mathbf{R}\mathbf{t}}$, and following the same
 order of operations on the LHS as on the RHS of (64) (first an $O(n)$ action, and then translation), we

756 rewrite (64) starting from the LHS as
 757
 758
 759

$$\begin{aligned}
 760 \quad T V B_{\text{hom}}(\mathbf{S}) X &= \mathbf{M}_n^\top \mathbf{R}_O T_{\text{interm}} \mathbf{R}_O^\top \mathbf{M} \mathbf{M}_n^\top \mathbf{R}_O \mathbf{R} \mathbf{R}_O^\top \mathbf{M} B_{\text{hom}}(\mathbf{S}) X \\
 761 &= \mathbf{M}_n^\top \mathbf{R}_O T_{\text{interm}} \mathbf{R} \mathbf{R}_O^\top \mathbf{M}_n B_{\text{hom}}(\mathbf{S}) X \\
 762 &= \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{I}_n & 0 & \mathbf{t} \\ -2\alpha \mathbf{t}^\top & 1 & -\alpha \|\mathbf{t}\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R} & 0 & 0 \\ \mathbf{0}^\top & 1 & 0 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n B_{\text{hom}}(\mathbf{S}) X \\
 763 &= \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{R} & 0 & \mathbf{t} \\ -2\alpha \mathbf{t}^\top \mathbf{R} & 1 & -\alpha \|\mathbf{t}\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n B_{\text{hom}}(\mathbf{S}) X \\
 764 &= p \|\mathbf{c}\| \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{R} & 0 & \mathbf{t} \\ -2\alpha \mathbf{t}^\top \mathbf{R} & 1 & -\alpha \|\mathbf{t}\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x}\|^2) \\ 1 \end{bmatrix} \\
 765 &= p \|\mathbf{c}\| \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{R} \mathbf{x} + \mathbf{t} \\ \alpha (r^2 - \|\mathbf{c}\|^2 - \|\mathbf{x} + \mathbf{t}\|^2) \\ 1 \end{bmatrix} = \mathbf{M}_n^\top \mathbf{R}_O \mathbf{R}_O^\top \mathbf{M}_n \begin{bmatrix} \mathbf{y}_{R\mathbf{t}} \\ p \|\mathbf{c}\| \end{bmatrix} \\
 766 &= \mathbf{I}_{n+2} \begin{bmatrix} B(\mathbf{S}) \mathbf{Q}_{\text{RHS}} X \\ p \|\mathbf{c}\| \end{bmatrix} = B_{\text{hom}}(\mathbf{S}) \mathbf{Q}_{\text{RHS}} X. \quad \square \\
 767 & \\
 768 & \tag{68}
 \end{aligned}$$

769 Thus, we have proved (64) for
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779

$$\begin{aligned}
 780 \quad \mathbf{Q}(\mathbf{R}, \mathbf{t}) &:= T(\mathbf{t}) V(\mathbf{R}) = \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{R} & 0 & \mathbf{t} \\ -2\alpha \mathbf{t}^\top \mathbf{R} & 1 & -\alpha \|\mathbf{t}\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n, \\
 781 & \\
 782 & \\
 783 & \\
 784 & \\
 785 & \tag{69} \\
 786 & \\
 787 & \\
 788 & \\
 789 &
 \end{aligned}$$

790 where \mathbf{M}_n , \mathbf{R}_O , and V (16) are extended to $(n+2) \times (n+2)$ using the identity complement, T is
 791 given by (61), and $\alpha = \frac{\sqrt{n+1}}{2p \|\mathbf{c}\|} = \frac{\sqrt{n}}{2\|\mathbf{c}\|}$.
 792

793 Similarly to what we did with $T(\mathbf{t})$ in Section 3.2, we verify that $\mathbf{Q}(\mathbf{R}, \mathbf{t})$ is a linear representation by
 794 (a) finding its inverse
 795
 796
 797
 798
 799

$$\begin{aligned}
 800 \quad \mathbf{Q}(\mathbf{R}, \mathbf{t})^{-1} &= V(\mathbf{R})^{-1} T(\mathbf{t})^{-1} = \mathbf{M}_n^\top \mathbf{R}_O \mathbf{R}^\top T_{\text{interm}}^{-1} \mathbf{R}_O^\top \mathbf{M}_n \\
 801 &= \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{R}^\top & 0 & -\mathbf{R}^\top \mathbf{t} \\ 2\alpha \mathbf{t}^\top & 1 & -\alpha \|\mathbf{t}\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n, \\
 802 & \\
 803 & \tag{70} \\
 804 & \\
 805 & \\
 806 & \\
 807 & \\
 808 & \\
 809 &
 \end{aligned}$$

which is valid for all \mathbf{R} in $\mathbf{O}(n)$, $\mathbf{t} \in \mathbb{R}^n$ and α with $\|\mathbf{c}\| \neq 0$ (which is the degenerate case considered in Section 3.2), and (b) confirming that the action of $\mathbf{Q}(\mathbf{R}_2, \mathbf{t}_2) \mathbf{Q}(\mathbf{R}_1, \mathbf{t}_1)$ results in the transformation

of the original nD \mathbf{x} as $\mathbf{R}_2(\mathbf{R}_1 \mathbf{x} + \mathbf{t}_1) + \mathbf{t}_2 = \mathbf{R}_2 \mathbf{R}_1 \mathbf{x} + \mathbf{R}_2 \mathbf{t}_1 + \mathbf{t}_2$:

$$\begin{aligned}
& \mathcal{Q}(\mathbf{R}_2, \mathbf{t}_2) \mathcal{Q}(\mathbf{R}_1, \mathbf{t}_1) \\
&= \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{R}_2 & 0 & \mathbf{t}_2 \\ -2\alpha \mathbf{t}_2^\top \mathbf{R}_2 & 1 & -\alpha \|\mathbf{t}_2\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{R}_1 & 0 & \mathbf{t}_1 \\ -2\alpha \mathbf{t}_1^\top \mathbf{R}_1 & 1 & -\alpha \|\mathbf{t}_1\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n \\
&= \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{R}_2 \mathbf{R}_1 & 0 & \mathbf{R}_2 \mathbf{t}_1 + \mathbf{t}_2 \\ -2\alpha \mathbf{t}_2^\top \mathbf{R}_2 \mathbf{R}_1 - 2\alpha \mathbf{t}_1^\top \mathbf{R}_1 & 1 & -2\alpha \mathbf{t}_2^\top \mathbf{R}_2 \mathbf{t}_1 - \alpha \|\mathbf{t}_1\|^2 - \alpha \|\mathbf{t}_2\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n \\
&= \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{R}_2 \mathbf{R}_1 & 0 & \mathbf{R}_2 \mathbf{t}_1 + \mathbf{t}_2 \\ -2\alpha (\mathbf{t}_2^\top \mathbf{R}_2 + \mathbf{t}_1^\top) \mathbf{R}_1 & 1 & -\alpha (2 \mathbf{t}_2^\top \mathbf{R}_2 \mathbf{t}_1 + \|\mathbf{R}_2 \mathbf{t}_1\|^2 + \|\mathbf{t}_2\|^2) \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n \\
&= \mathbf{M}_n^\top \mathbf{R}_O \begin{bmatrix} \mathbf{R}_2 \mathbf{R}_1 & 0 & \mathbf{R}_2 \mathbf{t}_1 + \mathbf{t}_2 \\ -2\alpha (\mathbf{t}_2^\top + \mathbf{t}_1^\top \mathbf{R}_2^\top) \mathbf{R}_2 \mathbf{R}_1 & 1 & -\alpha \|\mathbf{R}_2 \mathbf{t}_1 + \mathbf{t}_2\|^2 \\ \mathbf{0}^\top & 0 & 1 \end{bmatrix} \mathbf{R}_O^\top \mathbf{M}_n \\
&= \mathcal{Q}(\mathbf{R}_2 \mathbf{R}_1, \mathbf{R}_2 \mathbf{t}_1 + \mathbf{t}_2). \quad \square
\end{aligned} \tag{71}$$

B DEMONSTRATION

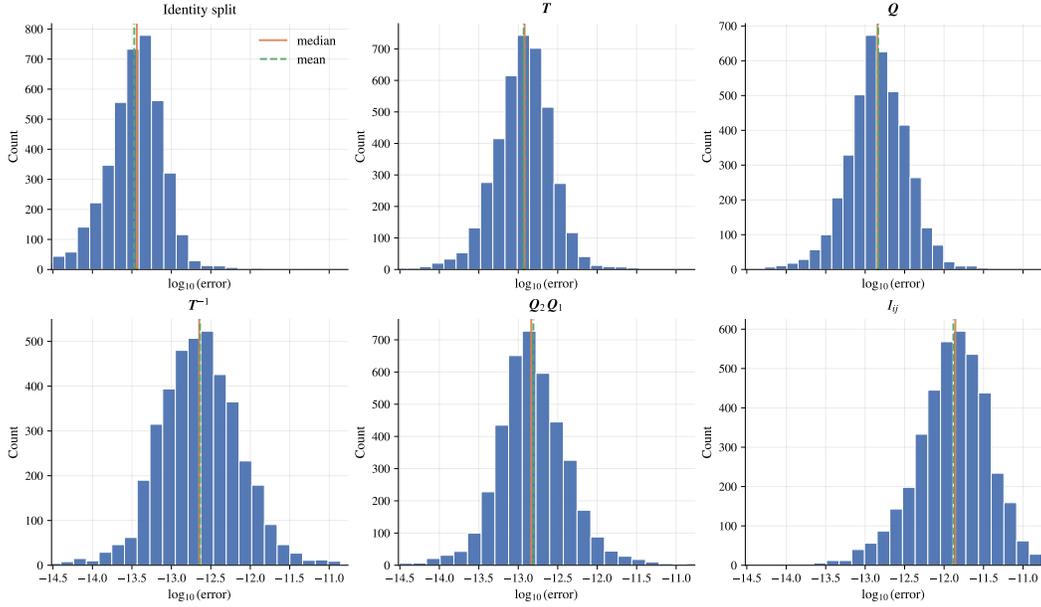


Figure 1: Numerical verification of our construction with histograms of \log_{10} residual norms. Left-to-right and row-by-row, we report errors for: the invariant/equivariant split (identity split) (18), translation equivariance (27) with T , the equivariance constraint (32) with full $E(n)$ actions Q , the inverse of the translation representation T^{-1} (30), group composition $Q_2 Q_1$ (39), and invariants $I_{ij,k}$ (42) and (43). Residuals cluster at machine precision, confirming the derived identities up to numerical roundoff.