An Explicit Frame Construction for Normalizing 3D Point Clouds

Justin Baker [*1 2] **Shih-Hsin Wang** [*1 2] **Tommaso de Fernex** [1] **Bao Wang** [1 2]

Abstract

Many real-world datasets are represented as 3D point clouds – yet they often lack a predefined reference frame, posing a challenge for machine learning or general data analysis. Traditional methods for determining reference frames and normalizing 3D point clouds often struggle with specific inputs, lack theoretical guarantees, or require massive data. We introduce a new algorithm that overcomes these limitations and guarantees both universality and compatibility with any learnable framework for 3D point cloud analysis. Our algorithm works with any input point cloud and performs consistently regardless of input complexities, unlike data-driven methods that are susceptible to biases or limited training data. Empirically, our algorithm outperforms existing methods in effectiveness and generalizability across diverse benchmark datasets. Code is available at https://github.com/Utah-Math-Data-Science/alignment.

1. Introduction

Analyzing 3D point clouds – in applications such as molecular modeling (Thomas et al., 2018; Gebauer et al., 2019; Batzner et al., 2022), geometric processing (Mitra et al., 2004; Morell et al., 2014; Wang & Kim, 2019), and robotics (Pomerleau et al., 2015; Mi et al., 2015; Kim et al., 2018) – is challenging. The 3D point clouds can have arbitrary positions and orientations but lack a fixed reference frame. Despite this, certain features of the objects should exhibit equivariance or invariance as the frame undergoes transformations. Understanding and leveraging the equivariance or invariance properties of the point cloud is crucial for efficient, accurate, and robust machine learning (ML), as it significantly boosts the sample efficiency, and promotes the inductive bias of the learning algorithms (Hinton et al., 2011; Cohen & Welling, 2017; Romero et al., 2020; van der Pol et al., 2020; Wang et al., 2024).

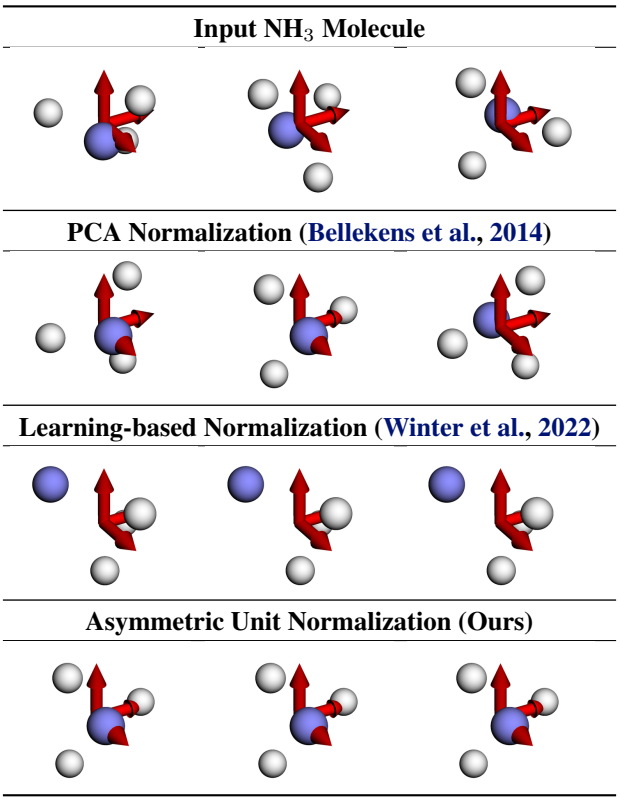


Figure 1. Normalized NH_3 molecules for a set of rotated input (first row). PCA result is inconsistent in both rotation and reflection (second row). Data-driven (learned) normalization is consistent but introduces error in the positional data quantified by a root mean square error (RMSE), with respect to the ground truth positional data, of 0.0924 (third row). Our proposed asymmetric unit normalization is both consistent and error-free (last row).

Shifts in positions of a point cloud can be addressed by centering the point cloud at the origin. Orientation variations, however, present a tougher challenge. Two main categories of approaches have emerged to address orientation variations. The first category approaches incorporate equivariance and invariance directly into ML models through architectural (e.g. neural network) constraints, en-

[*]Equal contribution [1]Department of Mathematics, University of Utah, Salt Lake City, Utah, USA [2]Scientific Computing and Imaging (SCI) Institute, University of Utah, Salt Lake City, Utah, USA. Correspondence to: Bao Wang <wangbaonj@gmail.com>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

suring the prediction results demonstrate equivariance or invariance without capturing the orientation of the reference frame; see e.g. (Anderson et al., 2019; Fuchs et al., 2020; Satorras et al., 2021; Dym & Maron, 2021; Passaro & Zitnick, 2023). The second category approaches focus on determining frames and normalizing the data to its invariant representations (Vranic et al., 2001; Chaouch & Verroust-Blondet, 2008; Winter et al., 2022). More precisely, given a space X of point clouds acting under the transformation group $G = \mathrm{O}(3)$ or $\mathrm{SO}(3)$, we seek a "frame" – a function $\mathcal{F}$ from X to G that maps each point cloud $x \in X$ to a transformation $\mathcal{F}(x) \in G$. This transformation, representing x's orientation, holds a pivotal property: applying its inverse to x yields an invariant representation $\mathcal{F}(x)^{-1}x$, effectively normalizing the point cloud. Further details can be found in Section 2. Importantly, the second category approaches are more flexible and compatible with a wider range of data analysis frameworks; even those without intrinsic equivariance or invariance properties, as highlighted in (Puny et al., 2021; Kaba et al., 2023). In light of this advantage, our focus in this paper is specifically on the second category approaches to eliminate the complicates in data analysis due to orientation variations. In particular, we focus on **constructing frames and utilizing normalization techniques for 3D point clouds with respect to the action of isometries.**

The second category approaches – that determining frames and normalizing data to its invariant representation – can be further classified into two classes: direct and learning-based frame construction. Direct construction methods utilize explicit rules – often relying on principle component analysis (PCA) – to derive a frame and align a given 3D object to a canonical representation; see, e.g. (Vranic et al., 2001; Chaouch & Verroust-Blondet, 2008; Bellekens et al., 2014). However, PCA-derived frames suffer from non-uniqueness and potential ambiguity in sign choices. As demonstrated in Figure 1, PCA is inconsistent in selecting the orientation for both rotation and reflection. This is particularly problematic for molecular structures with strong symmetries like NH_3. While Puny et al. (2021) and Duval et al. (2023) consider a set of frames that include all choices of sign changes, it still cannot handle cases when the eigenvalues of covariance matrices are not distinct. Learning-based frame determination methods leverage neural networks to learn frames that overcome the aforementioned limitations of direct frame construction. Winter et al. (2022) propose a framework that simultaneously learns invariant embeddings and frames; their framework learns an equivariant function from the input embedding space to an intermediate homogeneous space to obtain the reference frame. However, as demonstrated in (Winter et al., 2022), the stabilizer of a given input is isomorphic to that of its corresponding intermediate element in homogeneous space, necessitating distinct homogeneous spaces for inputs with different stabilizers, significantly complicating framework designs. To circumvent such challenges, Kaba et al. (2023) introduces the notion of relaxed equivariance, making learning function $\mathcal{F}$ require only one output space instead of multiple homogeneous spaces.

However, learning-based frames still exhibit some limitations. As highlighted by Kaba et al. (2023), the function $\mathcal{F}$ may lack continuity. In other words, there is no theoretical assurance that a continuous parametrized model can effectively approximate such a non-continuous function. We delve into the inevitability of a non-continuous frame in our setting in Section 3. Additionally, the presence of test data featuring inputs with novel symmetries not present in the training data may compromise the generalizability of learning methods; empirical support for these observations is provided in Figure 1. Using the learning-based technique from (Winter et al., 2022), we observe that the orientation frame is consistent. However, to data is only consistent up to the error of the learning-based method, which makes each of the normalized data in Figure 1 easily distinguishable.

1.1. Our contribution

We tackle the challenge of 3D point cloud analysis by introducing a direct frame construction and normalization algorithm to overcome limitations of existing methods, as discussed before. Our contributions are summarized below:

- We show that in our setting, constructing a continuous frame is impossible. However, even without continuous frames, our proposed frame construction retains the universal approximation ability of any framework. See Section 3 for details.
- We present an explicit training-free frame construction and normalization technique that can seamlessly handle any 3D point cloud. See Section 4 for details.
- We provide various empirical evidence to validate both the theoretical findings and the practical effectiveness of our algorithm. Our method outperforms existing techniques in accuracy and generalizability across various benchmark datasets. See Section 6 for details.

1.2. Organization

We organize this paper as follows: In Section 2, we present necessary background materials and review some existing results. In Section 3, we discuss the bottlenecks of the existing universality of canonicalization and present our new insights on frames and the corresponding universality results. We present our explicit alignment construction in Section 4. We discuss additional related works in Section 5. We compare our explicitly constructed alignment with existing algorithms in Section 6. Technical proofs and missing details are provided in the appendix.

2. Preliminary

Group theory. While our primary focus is on Euclidean groups and permutation groups, it is beneficial to provide a brief overview of fundamental group theory concepts in a broader context. In a G-space, where group G acts on a vector space X, the group action is defined as a group homomorphism mapping each group element to a bijection from X to itself. This action is denoted by $g \cdot x$, where the bijection corresponds to g applied to x. We define the *G-orbit of x* as the set $G \cdot x \coloneqq \{g \cdot x \in X \mid g \in G\} \subseteq X$, and the *stabilizer of* x as the subgroup $G_x \coloneqq \{g \in G \mid g \cdot x = x\} \subseteq G$. Consider two G-spaces, X and Y. A function $f : X \to Y$ is referred to as *G-equivariant* if it satisfies $f(g \cdot x) = g \cdot f(x)$, and as *$G$-invariant* if it satisfies $f(g \cdot x) = f(x)$.

Frame, normalization, and canonicalization. The concept of equivariant moving frames – applicable to any Lie transformation group G – is introduced in (Olver, 2009). In this work, an equivariant moving frame on a G-space X is defined as a G-equivariant smooth function from X to G. To illustrate why frames can be considered as group elements, let us delve into physics. In physics, a reference frame is characterized by a coordinate system established through an orthonormal basis of $\mathbb{R}^3$ and a designated point serving as the origin. Without loss of generality, let origin be $\mathbf{0} = (0, 0, 0)$, then each reference frame can be represented as an orthogonal matrix – a group element $g \in \mathrm{O}(3)$.

However, the existence of equivariant moving frames necessitates a free action, meaning $G_x = \{e\}, \forall x \in X$, where e is the identity in G. This requirement poses a limitation, making it impractical for symmetric inputs x with a non-trivial stabilizer. To overcome this constraint, we relax the condition and define a *frame* as a function $\mathcal{F} : X \to G$, without the necessity of smoothness but ensures the relaxed equivariance condition introduced in (Kaba et al., 2023).

Definition 2.1 (Frame)**.** A frame is defined as a function $\mathcal{F} : X \to G$ satisfying the condition that for any $x \in X$ and $g_1 \in G$, there exists a $g_2 \in g_1 G_x$ such that

$$\mathcal{F}(g_1 \cdot x) = g_2 \cdot \mathcal{F}(x). \tag{1}$$

For simplicity, we may also refer to $\mathcal{F}(x)$ as a frame.

We leave the discussion on the existence of frames in Appendix A. The relaxed equivariance grants a certain flexibility in output transformations, allowing them to differ by elements within stabilizers. Importantly, the relaxed equivariance implies that the function $\mu : X \to X$ defined by $\mu(x) \coloneqq \mathcal{F}(x)^{-1} x \in X$ is group invariant, i.e.,

$$\begin{aligned} \mu(g_1 \cdot x) &= \mathcal{F}(g_1 \cdot x)^{-1}(g_1 \cdot x) = \mathcal{F}(x)^{-1} g_2^{-1}(g_1 \cdot x) \\ &= \mathcal{F}(x)^{-1}(g_2^{-1} g_1 \cdot x) = \mathcal{F}(x)^{-1} x = \mu(x), \end{aligned} \tag{2}$$

since $g_2^{-1} g_1 \in G_x$. Such a mapping – sending each input x to an invariant representation $\mu(x)$ within the same orbit – is commonly referred to as *normalization* or *alignment* (Chaouch & Verroust-Blondet, 2008).

Notice that each input $x \in X$ can be represented as the product of a frame $\mathcal{F}(x)$ and its associated normalization $\mu(x)$. This decomposition effectively breaks down the input data into components that remain invariant and components that exhibit equivariance. Consequently, employing this decomposition allows any function to be rendered G-equivariant through the *canonicalization* technique introduced in (Kaba et al., 2023). In particular, let f be any function between finite-dimensional normed G-spaces X and Y. *The canonicalization of f* through a frame $\mathcal{F} : X \to G$ is a G-equivariant function $\phi : X \to Y$ defined by:

$$\phi(x) \coloneqq \mathcal{F}(x) \cdot f(\mathcal{F}(x)^{-1} \cdot x) \tag{3}$$

We leave the justification of the relaxed G-equivariance of $\phi(x)$ in Appendix A. When the group actions are described by group representations $\rho_X : G \to \mathrm{GL}(X)$ and $\rho_Y : G \to \mathrm{GL}(Y)$, the canonicalization can be written as:

$$\phi(x) \coloneqq \rho_Y(\mathcal{F}(x)) f(\rho_X(\mathcal{F}(x))^{-1} x) \tag{4}$$

In (Kaba et al., 2023), $\rho_X(\mathcal{F}(x))$, $\rho_Y(\mathcal{F}(x))$ are called canonicalization functions and are directly approximated by neural networks in their framework.

Point clouds. Point clouds – collections of points in $\mathbb{R}^n$ – are inherently unordered sets. While often represented as matrices $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m] \in \mathbb{R}^{n \times m}$ where each $\boldsymbol{x}_i \in \mathbb{R}^n$ represents the coordinates of a point and m denote the total number of points, these impose an order that can bias analysis. To address this issue, we consider the action of the symmetric group S_m that permutes matrix columns, ensuring order-independent analysis. This allows us to explore the actions of various transformation groups, such as $\mathrm{E}(n)$, $\mathrm{SE}(n)$, $\mathrm{O}(n)$, or $\mathrm{SO}(n)$, on matrix representations of point clouds without being hindered by order-induced biases. Formally, we investigate the group action of $\mathrm{S}_m \times G$ on the matrix representations $\boldsymbol{X}$, where G denotes any of the aforementioned groups. This action is defined as follows:

$$(\sigma, g) \cdot \boldsymbol{X} \mapsto \left[g \cdot \boldsymbol{x}_{\sigma^{-1}(1)}, \ldots, g \cdot \boldsymbol{x}_{\sigma^{-1}(m)}\right]$$

for any $\sigma \in \mathrm{S}_m$ and $g \in G$.

3. Universality of Canonicalization

We first recall the universality of canonicalization as discussed in (Winter et al., 2022). Let us consider the functions between G-spaces X and Y. We say a parameterized function f is a universal approximator of continuous functions if for any continuous function ψ, any compact set $K \subset X$ and any $\epsilon > 0$, there is a choice of parameters for function f such that $\|\psi(x) - f(x)\| < \epsilon$ for any $x \in K$. Kaba et al. (2023) shows the following theorem:

Theorem 3.1 (Kaba et al. (2023)). *Let $f : X \to Y$ be a parameterized function. Define ϕ as the G-equivariant parameterized function given by,*

$$\phi(x) \coloneqq h'(x) f(h(x)^{-1}x),$$

where $h(x), h'(x)$ are canonicalization functions. Suppose $h(x), h'(x)$ are continuous. Then ϕ is a universal approximator of G-equivariant continuous functions if f is a universal approximator of continuous functions.

Lack of continuity. A crucial condition for the expressiveness guarantee in Theorem 3.1 is the continuity of canonicalization functions. However, as highlighted by Kaba et al. (2023), achieving continuity can be challenging and could affect the guarantee. We establish Proposition 3.2 below, showing the absence of continuity in our particular cases:

Proposition 3.2. *For point clouds in $\mathbb{R}^{n\times m}$ with $m, n \geq 3$, it is impossible to construct a frame $\mathcal{F} : \mathbb{R}^{n\times m} \to G$ that is continuous across the entire domain when $G = \mathrm{E}(n), \mathrm{SE}(n), \mathrm{O}(n)$ or $\mathrm{SO}(n)$.*

Revisiting the universality of canonicalization. Proposition 3.2 highlights a crucial constraint: continuity in frames (canonicalization functions with regular representations) is not achievable for the groups of our interest. This raises concerns about the expressivity guarantee in Theorem 3.1, which relies on continuity. Fortunately, we can circumvent this obstacle by leveraging the orthogonality of representations. Our established Theorem 3.3, presented below, offers an alternative path toward expressiveness guarantees, **even in the absence of continuous canonicalization functions**.

Theorem 3.3. *Let X, Y be two normed G-spaces where the group action are defined by the orthogonal representations*[1] *$\rho_X : G \to \mathrm{GL}(X)$ and $\rho_Y : G \to \mathrm{GL}(Y)$, respectively. Assume that X is finite-dimensional. Suppose f is a parameterized function from X to Y. Define ϕ as the canonicalization of f defined in equation (4) through an arbitrary frame $\mathcal{F} : X \to G$. Then ϕ is a universal approximator of G-equivariant continuous functions as long as f is a universal approximator of continuous functions.*

Leveraging the key insights from Theorem 3.3, we reveal the universality of canonicalization for the group action of $\mathrm{E}(n)$ or $\mathrm{SE}(n)$ on point clouds in Theorem 3.4. This universality guarantees that canonicalization techniques designed for these groups ensure any learning framework operating on point clouds retains its universal approximation capabilities.

The key to this universality lies in the decomposition of $\mathrm{E}(n)$ and $\mathrm{SE}(n)$. Indeed, $\mathrm{E}(n)$ (respectively, $\mathrm{SE}(n)$) can be expressed as a semidirect product $\mathrm{O}(n) \ltimes \mathrm{T}(n)$ (respectively, $\mathrm{SO}(n) \ltimes \mathrm{T}(n)$), where $\mathrm{T}(n)$ is the translational group of $\mathbb{R}^n$. Importantly, all representations of $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ are orthogonal, preserving the desirable properties underlying Theorem 3.3. Moreover, by representing each group element $g \in \mathrm{E}(n)$ as a pair $(\boldsymbol{Q}, \boldsymbol{t})$ with $\boldsymbol{Q} \in \mathrm{O}(n)$ and $\boldsymbol{t} \in \mathrm{T}(n)$, we may focus on frames that relocate the center of the point cloud to the origin. These frames take the following form:

$$\mathcal{F}(\boldsymbol{X}) = \left(\mathcal{F}'\left(\boldsymbol{X} - \frac{1}{m}\boldsymbol{X}\mathbf{1}_m\mathbf{1}_m^\top\right), \frac{1}{m}\boldsymbol{X}\mathbf{1}_m\right) \in \mathrm{E}(n) \tag{5}$$

where $\mathcal{F}' : \mathbb{R}^{n\times m} \to \mathrm{O}(n)$ is a frame and $\mathbf{1}_m = [1, \ldots 1]^\top \in \mathbb{R}^m$, so that $\frac{1}{m}\boldsymbol{X}\mathbf{1}_m$ is the center of the point cloud. A similar construction applies to $\mathrm{SE}(n)$. For this specific type of frame, we have an expressivity guarantee:

Theorem 3.4. *For $G = \mathrm{E}(n)$ or $\mathrm{SE}(n)$, the function $\mathcal{F} : \mathbb{R}^{n\times m} \to G$ constructed via equation (5) is a frame. Moreover, let $f : \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}$ be a parameterized function, and ϕ be its canonicalization through $\mathcal{F}$ defined in equation (3). Then ϕ is a universal approximator of G-equivariant continuous functions if f is a universal approximator of continuous functions.*

4. Explicit Construction of Frames

Motivation. We have demonstrated the absence of a continuous frame in our considered scenarios. To address this limitation, we are intended to manually design a frame rather than relying on learning models to approximate such a frame. Inspired by the specific form of frames studied in Theorem 3.4, we set our sights on explicitly constructing a frame $\mathcal{F} : \mathbb{R}^{3\times m} \to G$ for point clouds centered at the origin for $G = \mathrm{O}(3)$ or $\mathrm{SO}(3)$. Specifically, we aim to find the inverse of $\mathcal{F}$ to effectively normalize the point cloud.

Before delving into the specifics of frame construction, we ask a fundamental question inspired by our intuition: **what intrinsic characteristics of a point cloud – beyond mere coordinates – make choosing a frame or orientation tricky?** One key challenge arises from the presence of inherent symmetries within the point cloud. Imagine a perfectly symmetrical object like a cube. Any rotation that aligns its edges with a standard coordinate system seems equally valid, creating an ambiguity in the "correct" orientation. This occurs because multiple rotations (differing by the cube's symmetry operations) yield visually identical results. To address this inherent ambiguity, we need to understand the intrinsic symmetry of the point cloud. In what follows, we will elucidate this consideration and propose a strategy to construct a frame that handles these symmetries.

Let's briefly revisit the concepts of symmetry operations and point groups, which describe the inherent symmetries of point clouds centered at the origin. For an in-depth discussion, refer to (Miller, 1973; McWeeny, 2002). A *symmetry operation* is a transformation that leaves the intrinsic geo-

[1] Orthogonality implies that the image of group representations consists of orthogonal matrices.

metric configuration of a point cloud unchanged, revealing its underlying symmetries. Formally, for a point cloud represented by matrix $\boldsymbol{X} \in \mathbb{R}^{3\times m}$, a symmetry operation is defined by a pair $(\sigma, g) \in \mathrm{S}_m \times \mathrm{O}(3)$ that preserves the point cloud $\boldsymbol{X}$ unaltered: $(\sigma, g) \cdot \boldsymbol{X} = \boldsymbol{X}$. In other words, it lies within the stabilizer subgroup of $\boldsymbol{X}$. Alternatively, we can consider the action on the unordered (multi-)set representation $\{\!\{\boldsymbol{x}_i\}\!\}_{i=1}^m$. Here, a symmetry operation is simply characterized by a group element $g \in \mathrm{O}(3)$ satisfying $\{\!\{\boldsymbol{x}_i\}\!\} = g \cdot \{\!\{\boldsymbol{x}_i\}\!\} := \{\!\{g \cdot \boldsymbol{x}_i\}\!\}$, meaning the set remains unchanged after applying the transformation g to each point.

The inherent symmetry of a point cloud is then captured by its point group, defined as the subgroup of $\mathrm{O}(3)$ containing all symmetry operations that preserve its structure. It is denoted as $\mathrm{Sym}(\{\!\{\boldsymbol{x}_i\}\!\}) := \{g \in \mathrm{O}(3) \mid g \cdot \{\!\{\boldsymbol{x}_i\}\!\} = \{\!\{\boldsymbol{x}_i\}\!\}\}$.

Objective. While our initial focus was on seeking a frame $\mathcal{F} : \mathbb{R}^{3\times m} \to G$, our ultimate goal is using this frame to normalize point cloud $\{\!\{\boldsymbol{x}_i\}\!\}$. This normalization uses $\mathcal{F}(\boldsymbol{X})^{-1}\{\!\{\boldsymbol{x}_i\}\!\}$ as an invariant representation of $\{\!\{\boldsymbol{x}_i\}\!\}$. However, point clouds, unlike matrices, are unordered sets. To account for this, we introduce a "slightly adjusted" relaxed equivariance condition for $\mathcal{F}$: for any $g_1 \in G$, $\sigma \in \mathrm{S}_m$, there exists $g_2 \in g_1\mathrm{Sym}(\{\!\{\boldsymbol{x}_i\}\!\})$ such that

$$\mathcal{F}((\sigma, g_1) \cdot \boldsymbol{X}) = g_2\mathcal{F}(\boldsymbol{X}). \tag{6}$$

The condition now explicitly mandates invariance under point permutations, ensuring the frame's output is unaffected by the order of $\boldsymbol{X}$. Also, the output transformations are permitted to differ by elements within the point group $\mathrm{Sym}(\{\!\{\boldsymbol{x}_i\}\!\})$[2]. Similarly, one can check $\mathcal{F}(\boldsymbol{X})^{-1}\{\!\{\boldsymbol{x}_i\}\!\}$ is an invariant representation (notice that $g_2^{-1}g_1 \in \mathrm{Sym}(\{\!\{\boldsymbol{x}_i\}\!\})$):

$$\begin{aligned}\mathcal{F}((\sigma, g_1) \cdot \boldsymbol{X})^{-1}\{\!\{(\sigma, g_1) \cdot \boldsymbol{x}_i\}\!\} &= \mathcal{F}(\boldsymbol{X})^{-1}g_2^{-1}g_1\{\!\{\boldsymbol{x}_i\}\!\} \\ &= \mathcal{F}(\boldsymbol{X})^{-1}\{\!\{\boldsymbol{x}_i\}\!\}.\end{aligned} \tag{7}$$

Remark 4.1. While we made an adjustment for relaxed equivariance, the universality guarantee in Theorem 3.4 remains unaffected, as its proof relies on the orthogonality of group representations rather than relaxed equivariance. Moreover, the proof of Proposition 3.2 extends to this adjusted setting, as indicated in the last comment in the proof.

Strategy. Now, we outline a strategy to construct a frame $\mathcal{F}$ satisfying equation (6): First, we extract a sequence of linearly independent vectors within $\boldsymbol{X}$, denoted as $\boldsymbol{v}_1(\boldsymbol{X}), \ldots, \boldsymbol{v}_{\mathrm{rk}(\boldsymbol{X})}(\boldsymbol{X})$ where $\mathrm{rk}(\boldsymbol{X})$ denote the rank of $\boldsymbol{X}$. In particular, we require these points satisfying the condition equation (6): for any $1 \le i \le \mathrm{rk}(\boldsymbol{X})$, $g_1 \in G$, $\sigma \in \mathrm{S}_m$, there exists a $g_2 \in g_1\mathrm{Sym}(\{\!\{\boldsymbol{x}_i\}\!\})$ such that

$$\boldsymbol{v}_i((\sigma, g_1) \cdot \boldsymbol{X}) = g_2\boldsymbol{v}_i(\boldsymbol{X}). \tag{8}$$

[2]This constraint is slightly broader than the original relaxed equivariance in equation (1) since $G_{\boldsymbol{X}} \subseteq \mathrm{Sym}(\{\!\{\boldsymbol{x}_i\}\!\})$

Simplify the notation by denoting each $\boldsymbol{v}_i(\boldsymbol{X})$ as $\boldsymbol{v}_i$. Then we construct $\mathcal{F}$ through the following process:

Step 1: align $\boldsymbol{v}_1$ with the x-axis via a rotation $\boldsymbol{R}_1$.

Step 2: rotate along the x-axis to place $\boldsymbol{v}_2$ in the x-y plane with a positive y-component via $\boldsymbol{R}_2$.

Step 3: reflect along xy plane to ensure a positive z-component for the vector $\boldsymbol{x}_3$ via $\boldsymbol{R}_3$ (skip this step if $G = \mathrm{SO}(3)$)

If the number of steps surpasses $\mathrm{rk}(\boldsymbol{X})$, the process is terminated. Finally, we define the frame as

$$\mathcal{F}(\boldsymbol{X}) := \boldsymbol{R}_1^\top \cdot \cdots \cdot \boldsymbol{R}_{\mathrm{rk}(\boldsymbol{X})}^\top,$$

for $G = \mathrm{O}(3)$ (as $\mathcal{F}(\boldsymbol{X}) := \boldsymbol{R}_1^\top \cdot \cdots \cdot \boldsymbol{R}_{\min\{2,\mathrm{rk}(\boldsymbol{X})\}}^\top$ for $G = \mathrm{SO}(3)$). Proposition B.1, demonstrated in Appendix B, confirms that $\mathcal{F}$ constructed in this manner possesses the desired relaxed equivariance property.

4.1. Asymmetric Unit

To select the sequence of vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{\mathrm{rk}(\boldsymbol{X})}$ – essential for our frame construction – we employ a well-established algorithm with a time complexity of $O(m \log m)$, outlined in (Alt et al., 1988). This algorithm is specifically designed to uncover the inherent symmetries of a point cloud through a process of graph simplification. While we'll leverage its full capabilities, our main focus will be on its graph simplification technique. This technique allows us to extract an important component of point clouds known as the asymmetric unit (Hoffmann, 2020). The reasoning behind this strategic focus will be elucidated later in the paper. We briefly recap the three key steps of this approach below:

Step A: Unit-sphere representation. Firstly, we represent the original point cloud by the set of its corresponding points on the sphere along with their distances to the origin.

1. Project each point $\boldsymbol{x}_i$ onto the unit sphere in $\mathbb{R}^3$ via division by its norm and gather the resulting rescaled points into a set $\mathcal{Z} := \left\{\frac{\boldsymbol{x}_i}{\|\boldsymbol{x}_i\|}\right\}$, removing duplicates.

2. Define $\mathcal{R}_j := \left\{\!\!\left\{\|\boldsymbol{x}_i\| \mid \boldsymbol{z}_j = \frac{\boldsymbol{x}_i}{\|\boldsymbol{x}_i\|}\right\}\!\!\right\}$ for each $\boldsymbol{z}_j \in \mathcal{Z}$, which captures the original radial distances of points mapped to $\boldsymbol{z}_j$.

Remark 4.2. Points on the origin are excluded as their orientation is inherently undefined. Rotations and reflections wouldn't affect them, making them irrelevant when analyzing orientation-dependent frames. Consequently, omitting them does not impact the construction of valid frames.

The *unit-sphere representation* $\{\boldsymbol{z}_j, \mathcal{R}_j\}_{j=1}^{|\mathcal{Z}|}$ holds a remarkable ability: it preserves the core symmetry of a point cloud while offering a more concise and manageable structure. This is because the original point cloud can be reconstructed from this representation using the following formula:

$$\{\!\{\boldsymbol{x}_i\}\!\} = \bigcup_{\boldsymbol{z}_j \in \mathcal{Z}} \{\!\{r \cdot \boldsymbol{z}_j \mid r \in \mathcal{R}_j\}\!\} \tag{9}$$

Therefore, this concise representation retains all the information about symmetries. In other words, $g \in \mathrm{Sym}(\{\!\{\boldsymbol{x}_i\}\!\})$ if and only if $\{\boldsymbol{z}_j, \mathcal{R}_j\} = \{g \cdot \boldsymbol{z}_j, \mathcal{R}_j\}$, i.e. there is a permutation $\sigma \in \mathrm{S}_{|\mathcal{Z}|}$ such that $g\boldsymbol{z}_{\sigma^{-1}(j)} = \boldsymbol{z}_j$ and $\mathcal{R}_{\sigma^{-1}(j)} = \mathcal{R}_j$ for any $1 \leq j \leq |\mathcal{Z}|$.

Step B: Directed labeled graph construction. Next, we convert the unit-sphere representation into a directed labeled graph as follows:

1. Construct the convex hull of $\mathcal{Z}$[3] and represent it as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} := \{j \mid \boldsymbol{z}_j \in \mathcal{Z}\}$ includes all points in $\mathcal{Z}$ as nodes, and $\mathcal{E}$ includes both directions of segments on the convex hull as edges.

2. For each node j, let $j_1, \ldots, j_l$ be the vertices adjacent to j ordered in clockwise fashion. For $1 \leq k \leq l$, compute the distance $d_k = \|\boldsymbol{z}_j - \boldsymbol{z}_{j_k}\|$ and the angle θ_k on the sphere between $\boldsymbol{z}_{j_k}$, $\boldsymbol{z}_j$ and $\boldsymbol{z}_{j_{(k+1) \bmod l}}$.

3. Define $v_j = (\{(\theta_k, d_k) \mid e_{kj} \in \mathcal{E}\}, \mathcal{R}_j)$ and construct the labeled graph $\mathcal{G}$ by assigning v_j to node j.

As stated in (Alt et al., 1988), the directed labeled graph $\mathcal{G}$ mirrors the symmetry of the unit-sphere representation and, consequently, the original point cloud $\{\!\{\boldsymbol{x}_i\}\!\}$. We summarize the results of Alt et al. (1988) in the following lemma.

Lemma 4.3 (Alt et al. (1988))**.** *A permutation $\sigma \in \mathrm{S}_{|\mathcal{Z}|}$ induces a graph automorphism on $\mathcal{G}$, i.e. , $v_{\sigma^{-1}(j)} = v_j$ for any j and $e_{\sigma^{-1}(j)\sigma^{-1}(j')} \in \mathcal{E}$ for any $e_{jj'} \in \mathcal{E}$ if and only if there is a $g \in \mathrm{O}(3)$ s.t. $g\boldsymbol{z}_{\sigma^{-1}(j)} = \boldsymbol{z}_j$ and $\mathcal{R}_{\sigma^{-1}(j)} = \mathcal{R}_j$.*

Step C: Deterministic finite automaton construction and minimization. In the algorithm's final stage, the directed labeled graph $\mathcal{G}$ is converted into a specific structure, called deterministic finite automaton (DFA) (Hopcroft et al., 2001). This transformation allows us to apply Hopcroft's algorithm (Hopcroft, 1971; Aho & Hopcroft, 1974) to identify a unique minimal DFA (Hopcroft et al., 2001). Basically, Hopcroft's algorithm operates by exploiting an equivalence relation between edges derived from the action of $\mathrm{Aut}(\mathcal{G})$ and terminates when no further equivalences exist. The minimal DFA for the NH_3 molecule is shown in Figure 2, while the original DFA before minimization is provided in Figure 7 in Appendix F. For a thorough exposition of Hopcroft's algorithm and its application to this context, see (Hopcroft & Wong, 1974; Wolter et al., 1985; Alt et al., 1988).

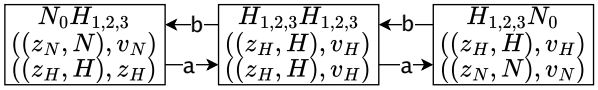


Figure 2. The minimal DFA for the NH_3 molecule. Invariant features are given by $\{H, N\}$, $z_{\{H,N\}}$, and $v_{\{H,N\}}$, node indices are given by $\{0, 1, 2, 3\}$ and state transitions are marked $\boldsymbol{a}, \boldsymbol{b}$.

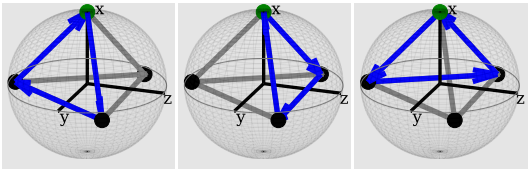


Figure 3. Isomorphic asymmetric units (blue) of NH_3 with N/H colored green/black resp. The asymmetric units reconstruct the polyhedron (grey) generated by the convex hull of the data projected on the unit sphere.

4.2. Constructing a frame from asymmetric units

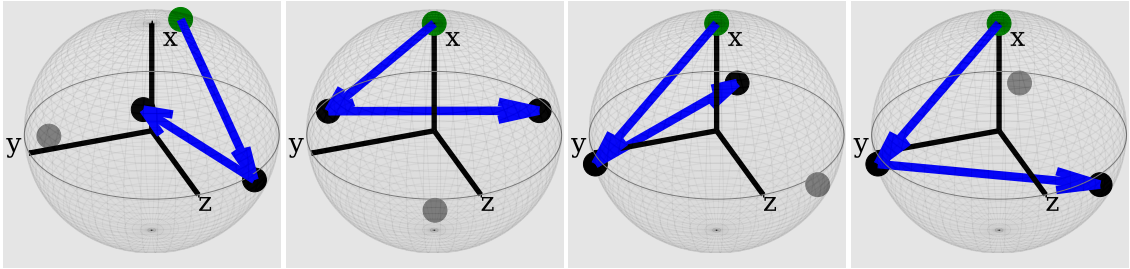
The minimal DFA, while utilized for symmetry detection in (Alt et al., 1988), offers a hidden gem for our frame construction needs. It encodes all "minimal" (directed) subgraphs, called asymmetric unit, that have the remarkable ability to reconstruct the entire graph $\mathcal{G}$, essentially serving as fundamental building blocks. These asymmetric units achieve this reconstruction through the application of graph automorphisms on $\mathcal{G}$, which correspond to symmetry operations as established in Lemma 4.3. This relationship echos the concept of unit cells, where replicating and symmetrically arranging these building blocks generates the complete structure of a crystal. Notably, asymmetric units are smaller than unit cells, lacking internal symmetries, yet still capable of generating the entire graph (Hoffmann, 2020). Figure 3 visualizes three such asymmetric units extracted from the minimal DFA.

To delve deeper, let's formally define an asymmetric unit and unveil its key properties. The group formed by graph automorphisms on $\mathcal{G}$ is denoted by $\mathrm{Aut}(\mathcal{G})$. For $\sigma \in \mathrm{Aut}(\mathcal{G})$ and a subgraph $\mathcal{H} = (\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H}) \subset \mathcal{G}$, the action of σ on $\mathcal{H}$ is defined as

$$\sigma \cdot \mathcal{H} := \left(\{\sigma(k) \mid k \in \mathcal{V}_\mathcal{H}\}, \{e_{\sigma(k)\sigma(k')} \mid e_{kk'} \in \mathcal{E}_\mathcal{H}\}\right). \tag{10}$$

A subgraph $\mathcal{H} \subseteq \mathcal{G}$ is said to *reconstruct* $\mathcal{G}$ if the union of all subgraphs within its orbit $\mathrm{Aut}(\mathcal{G}) \cdot \mathcal{H}$ is the entire graph $\mathcal{G}$, that is, $\mathcal{G} = \bigcup_{\sigma \in \mathrm{Aut}(\mathcal{G})} \sigma \cdot \mathcal{H}$. Moreover, it is said to be minimal if there is no proper subgraph $\mathcal{H}' \subset \mathcal{H}$ that also

[3]The convex hull exhibits the same symmetry as $\mathcal{Z}$.



(a) (b) (c) (d)

Figure 4. Normalization process of NH_3 by aligning N, H, H extracted from the asymmetric unit N→H→H→N (blue): (a) Initial data. (b) N (green) aligns to the x-axis. (c) H (black) aligns to the xy-plane. (d) H (black) aligns to the positive z-direction.

reconstructs $\mathcal{G}$. Interestingly, the minimality of $\mathcal{H}$ is tightly linked to its stabilizer:

Proposition 4.4. *Suppose $\mathcal{H} \subseteq \mathcal{G}$ is a subgraph that reconstructs $\mathcal{G}$. If $\mathcal{H}$ is minimal then for any $\sigma \in \mathrm{Aut}(\mathcal{G})$, we have $\sigma \cdot \mathcal{H} \neq \mathcal{H}$.*

Proposition 4.4 reveals the inherent asymmetry of the minimal building block $\mathcal{H}$, which we can then identify as asymmetric units. This asymmetry is characterized by the absence of automorphism preserving $\mathcal{H}$, implying that there are no equivalent edges, can further simplify its structure while preserving its reconstructive power. This explains why the minimal DFA encodes all asymmetric units. Recall that the minimal DFA is constructed by minimizing a DFA derived from $\mathcal{G}$. During this minimization, equivalent directed edges are identified based on a relationship drawn from the graph automorphisms (corresponding to symmetry operations on the given point cloud). The process terminates when no further equivalences can be found, resulting in a minimal DFA that represents the original DFA. To extract an asymmetric unit, we carefully navigate the equivalence classes of directed edges, choosing one representative from each class, and follow transitions within the DFA. For example, we could start with N_0H_1, followed by H_1H_2, and then H_2N_0. This yields the asymmetric unit depicted in Figure 3. Notably, the extracted asymmetric unit exhibits no inherent symmetry, echoing Proposition 4.4.

S_m-invariance and uniqueness. Notably, any asymmetric unit $\mathcal{H}$ extracted from the minimal DFA is an ordered[4] sequence of distinct directed edges, where consecutive edges are adjacent. It can be represented as $\mathcal{H} = \{e_{k_l k_{l+1}}\}_{l=1}^{|\mathcal{E}_\mathcal{H}|}$ where $e_{k_l k_{l+1}} \neq e_{k_{l'} k_{l'+1}}$ for any $l \neq l'$. The ordering of these edges then guarantees independence from index permutations within the point cloud, ensuring the desired S_m-invariance.

Furthermore, it's essential to note that while there might exist multiple asymmetric units that reconstruct $\mathcal{G}$, the uniqueness of minimal DFA implies that all asymmetric units correspond to equivalent minimal DFAs and hence are isomorphic, i.e., belong to the same orbit under the action of the automorphism group $\mathrm{Aut}(\mathcal{G})$.

We now proceed with the construction of the desired frame $\mathcal{F}$ using the asymmetric unit $\mathcal{H}$, obtained from the algorithm discussed above. Recall that it suffices to select a sequence of linearly independent vectors from the point clouds $\{\!\{\boldsymbol{x}_i\}\!\}$, or equivalently, their unit sphere representation $\{\boldsymbol{z}_j, \mathcal{R}_j\}$, that satisfy the condition equation (6).

We demonstrate our approach for choosing vectors from $\mathcal{H} = \{e_{k_l k_{l+1}}\}_{l=1}^{|\mathcal{E}_\mathcal{H}|}$ and elucidate why it meets the required conditions: $\mathcal{H} = \{e_{k_l k_{l+1}}\}_{l=1}^{|\mathcal{E}_\mathcal{H}|}$ is an S_m-invariant sequence of directed edges. By tracing this directed path, we naturally obtain $\{\boldsymbol{z}_{k_l}\}_{l=1}^{|\mathcal{E}_\mathcal{H}|}$, an S_m-invariant sequence of points in $\mathcal{Z}$. This is a direct consequence of the correspondence between nodes in $\mathcal{G}$ and points in the unit-sphere representation $\{\boldsymbol{z}_j, \mathcal{R}_j\}$. We can then strategically select linearly independent vectors from this sequence as demonstrated in Figure 4(a). Importantly, relaxed equivariance is guaranteed since these vectors represent points in the unit-sphere representation $\{\boldsymbol{z}_j, \mathcal{R}_j\}$ (uniquely determining the original point clouds), and different asymmetric units are distinguished by symmetric operations in $\mathrm{Sym}(\{\!\{\boldsymbol{x}_i\}\!\})$. Figure 4 illustrates how these three vectors define a frame and normalize the NH_3 molecule according to the method in our strategy.

4.3. Point clouds with features

In cases where points may possess additional features, denoted as $\boldsymbol{f}_i$ (e.g., atom features in molecules), our frame determination remains applicable with a slight adjustment. Specifically, we treat $\boldsymbol{f}_i$s as tokens (playing a role similar to atomic numbers) and use a hash function to generate unique indices for these tokens, denoted as a_i. The set of radial distances is redefined as $\mathcal{R}_j = \{\!\{(\|\boldsymbol{x}_i\|, a_i) \mid \boldsymbol{z}_j = \frac{\boldsymbol{x}_i}{\|\boldsymbol{x}_i\|}\}\!\}$ for each j. Clearly, a formula similar to equation (9) holds for this case, ensuring the applicability of our algorithm to cases involving additional features.

[4]This inherent order arises naturally from lexicographical sort when minimizing DFA.

5. Additional Related Works

In this section, we discuss a few representative works revolving around leveraging frames for symmetry-aware ML. Cohen et al. (2019) present a general theory of group equivariant convolutional neural networks on homogeneous spaces. Zhao et al. (2020) develop a roto-translation equivariant 3D capsule module on quaternions that receives a sparse set of local reference frames. The paper (Kofinas et al., 2021) proposes local coordinate frames per node-object to induce roto-translation invariance to the geometric graph of the interacting dynamical system. Puny et al. (2021) develop a systematic framework for adapting known architectures to become invariant or equivariant to new symmetry type leveraging frame averaging. The idea of frame averaging has also been used for equivariant shape space learning (Atzmon et al., 2022). Simeonov et al. (2023) leverage local coordinate frames for SE(3)-equivariant relational rearrangement with neural descriptor fields. Local frames have also been used to build efficient and expressive 3D equivariant graph neural networks (Du et al., 2022; 2023).

6. Experiments

We validate our approach with a comprehensive comparison of our ASymmetric Unit Normalization (**ASUN**) with existing and publicly available methods on benchmark datasets. Specifically, we compare ASUN against PCA (Bellekens et al., 2014), the auto-encoder (AE) as described in (Winter et al., 2022), and frame-averaging (FA) (Puny et al., 2021). To the best of our knowledge, there are no publicly available implementations of the methods discussed in (Kaba et al., 2023). The datasets consist of QM9 (Ramakrishnan et al., 2014) molecular data, ModelNet40 (Wu et al., 2015) CAD point cloud data, and n-body (Satorras et al., 2021) point cloud trajectories[5]. Several state-of-the-art (SOTA) invariant and equivariant neural networks are used for comparison on benchmark tasks. These tasks empirically demonstrate the consistency and generalizability of ASUN.

6.1. QM9 molecular alignment

In this study, we evaluate the consistency and generalizability of ASUN using QM9. For this task, each model constructs a normalized baseline from the positional data of 5000 QM9 molecules. Each molecule is then subjected to a random rotation and translation. The models then normalize the perturbed data. The Wasserstein distance (Villani et al., 2009), i.e. the earth mover distance (EMD), is used to compute the loss between the initial normalization and the normalization after perturbation. Perturbation is applied a total of 100 times per molecular structure. The learning-based AE model is trained on the QM9 positional data following the procedure of Winter et al. (2022). In particular, the RMSE of the predicted positional data is minimized using the Adam optimizer. Once trained, the encoder of the model outputs the learned frame which normalizes the data.

[5]Due to page limit, we leave experiments on the n-body point cloud trajectories to Appendix E.

	Rank 1	**Rank 2**	**Rank 3**
PCA	**0.00014**	0.01793	0.82758
AE	1.15122	0.037539	0.03178
ASUN	**0.00014**	**0.00008**	**0.02826**

Table 1. Results for QM9 molecular normalization by rank. We illustrate the best performance for each rank in bold. The ASUN algorithm falls within the known error range for the QM9 dataset of 0.1Å (Kim et al., 2019) in all ranks.

	C_1	$C_{\infty,v}$	C_s	$D_{6,h}$	T_d
PCA	0.853	0.303	0.679	0.244	0.721
AE	0.041	0.240	0.037	0.023	0.028
ASUN	**0.018**	**0.001**	**0.004**	**0.001**	**0.001**

Table 2. Results for QM9 molecular normalization by point group. We illustrate the best performance for each rank in boldface. ASUN exhibits strongest generalization performance across all different point groups.

Table 1 compares PCA and AE models to ASUN, highlighting the consistency of ASUN. By separating the results by rank, one can easily observe the limitations of PCA and AE. As discussed in Section 1 the PCA model is prone to failure given rank 3 data with a high degree of symmetry. By contrast, the AE model has a limitation for rank 1 and 2 data. The efficacy of our ASUN is supported by the results in Table 1. In particular, PCA performs well for rank 1 data, but significantly underperforms for rank 3 data. We also observe that AE performs poorly for rank 1 data with an error of 1.2Å. ASUN performs the best of all methods across all ranks of the dataset as indicated in boldface.

The advantages of ASUN are illustrated in Fig. 5 for rank 1 and rank 2 data as well as Fig. 1 for rank 3 data. These figures visually illustrate the normalization of perturbed data. ASUN is an error-free explicit frame construction.

The generalizability of AE and PCA is compared to the ASUN method by separating the data based on their point group classification. The results by point group classification are shown in Table 2. In this task, the AE model is trained in the same manner but data is with-held based on its symmetry group. This tests the generalizability of AE to $C_{\infty,v}$, C_s, $D_{6,h}$ and T_d – see Table 4 in the appendix for details of these groups – in comparison to data-free algorithms like PCA and ASUN. AE outperforms PCA, which is generally incapable of handling data with high degrees of symmetry. However, it does not have strong generalizations like the data-free ASUN method.

Invariant/Equivariant	$z/\mathrm{SO}(3)$	$\mathrm{SO}(3)/\mathrm{SO}(3)$	ASUN/ASUN
SchNet (Schütt et al., 2018)	29.74	28.81	28.93
EGNN (Satorras et al., 2021)	22.61	24.63	25.41
MACE (Batatia et al., 2022)	25.36	27.18	27.55
eSCN (Passaro & Zitnick, 2023)	OOM	OOM	OOM
ASUN-Invariant	$z/\mathrm{SO}(3)$	$\mathrm{SO}(3)/\mathrm{SO}(3)$	ASUN/ASUN
GCN (Welling & Kipf, 2016)	13.57	11.91	**18.19**
GCNII (Chen et al., 2020)	21.3	22.97	**30.75**
SchNet-pos	28.40	36.18	**42.79**
EGNN-pos	29.94	47.41	**53.36**

Table 3. Object classification accuracy (%) of ModelNet40 for different train/test augmentations. (OOM – out of memory)

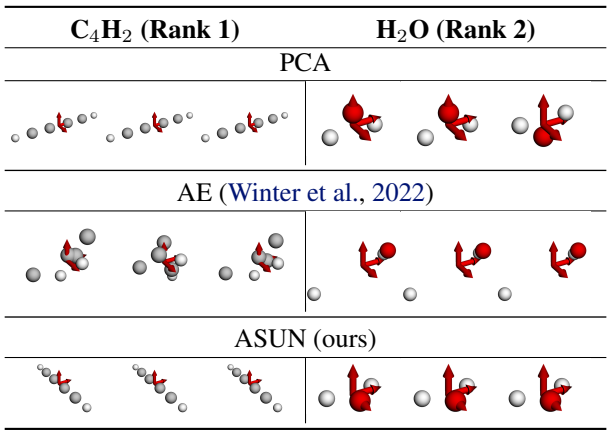


Figure 5. A comparison of aligned C_4H_2 (rank 1) and H_2O (rank 2) molecular structures. We observe inconsistencies at rank 1 for AE and at rank 2 for PCA. ASUN is consistent for all ranks.

6.2. ModelNet40 Shape Classification

Using the ModelNet40 dataset (Wu et al., 2015), we further demonstrate the versatility and effectiveness of ASUN for point cloud alignment. The dataset consists of 40 categorized CAD meshes with 9,843 training and 2,468 testing samples. The nodes and edges of the mesh are used to represent the data as a graph. ModelNet40 is pre-aligned in the xy-plane. The baseline data denoted z is then augmented with random z-axis rotation denoted SO(3) or z-axis alignment denoted ASUN. This benchmark is similar to the existing benchmark (Qi et al., 2017) used in (Kim et al., 2020; Luo et al., 2022; Chen & Cong, 2022). However, the existing benchmark uniformly samples data from the mesh, breaking the data symmetries. Our benchmark uses the full point cloud data which is significantly more challenging.

Several SOTA GNNs, categorized as Invariant/Equivariant and ASUN-Invariant, are trained on each data partition. ASUN-Invariant architectures utilize positional data as input node features and are invariant only when ASUN is applied. We apply this approach to SchNet and EGNN, labeled as SchNet-pos and EGNN-pos, respectively, while excluding MACE and eSCN from this modification. Also, we note that eSCN exceeds the memory capacity of the A100 GPU.

Table 3 reports the classification results for training/testing augmentations $z/\mathrm{SO}(3)$, $\mathrm{SO}(3)/\mathrm{SO}(3)$ and ASUN/ASUN. We note that invariant/equivariant models consistently maintain accuracy across augmentations, as expected, since the output remains invariant under transformations, including the application of ASUN. When positional data is used as a node feature, ASUN significantly improves the accuracy of all models, showing the benefit of ASUN – it allows the positional data to be used directly while maintaining invariance. Additionally, Figure 6 illustrates the application of our ASUN to the ModeNet40 dataset.

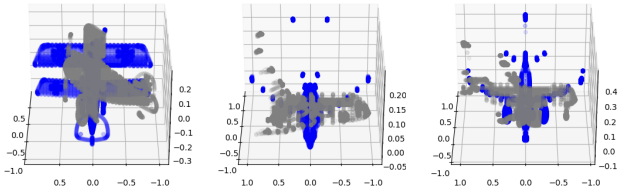


Figure 6. We visualize the application of ASUN on ModeNet40. The original point cloud (grey) is aligned with the blue one using our ASUN method before further analysis.

7. Concluding Remarks

This paper presents a new approach for 3D point cloud normalization, surpassing limitations observed in current approaches. Our direct frame construction and normalization algorithm exhibit robust performance across diverse 3D point cloud datasets, regardless of their complexity or symmetrical properties. This capability to normalize any 3D point cloud extends the applicability of our approach to diverse application domains beyond ML, encompassing fields such as medical imaging, geometric processing, computer vision, and geospatial analysis.

Acknowledgement

This material is based on research sponsored by NSF grants DMS-2152762, DMS-2219956, DMS-2208361, and DMS-2001254 and DOE grant DE-SC0023490.

Impact Statement

This work presents a novel direct frame construction for 3D point cloud normalization, offering performance guarantees within the specific machine learning context. Our proposed normalization outperforms existing algorithms in various settings, demonstrating its potential for broad application in 3D point cloud analysis. We do not see any potential ethical issues in our research.

References

Aho, A. V. and Hopcroft, J. E. *The design and analysis of computer algorithms*. Pearson Education India, 1974.

Alt, H., Mehlhorn, K., Wagener, H., and Welzl, E. Congruence, similarity, and symmetries of geometric objects. *Discrete & Computational Geometry*, 3(3):237–256, 1988.

Anderson, B., Hy, T. S., and Kondor, R. Cormorant: Covariant molecular neural networks. *Advances in neural information processing systems*, 32, 2019.

Atzmon, M., Nagano, K., Fidler, S., Khamis, S., and Lipman, Y. Frame averaging for equivariant shape space learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 631–641, 2022.

Batatia, I., Kovács, D. P., Simm, G. N., Ortner, C., and Csányi, G. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *arXiv preprint arXiv:2206.07697*, 2022.

Batzner, S., Musaelian, A., Sun, L., Geiger, M., Mailoa, J. P., Kornbluth, M., Molinari, N., Smidt, T. E., and Kozinsky, B. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature communications*, 13(1):1–11, 2022.

Bellekens, B., Spruyt, V., Berkvens, R., and Weyn, M. A survey of rigid 3D pointcloud registration algorithms. In *AMBIENT 2014: the Fourth International Conference on Ambient Computing, Applications, Services and Technologies, August 24-28, 2014, Rome, Italy*, pp. 8–13, 2014.

Chaouch, M. and Verroust-Blondet, A. A novel method for alignment of 3D models. In *2008 IEEE International Conference on Shape Modeling and Applications*, pp. 187–195. IEEE, 2008.

Chen, M., Wei, Z., Huang, Z., Ding, B., and Li, Y. Simple and deep graph convolutional networks. In *International Conference on Machine Learning*, pp. 1725–1735. PMLR, 2020.

Chen, R. and Cong, Y. The devil is in the pose: Ambiguity-free 3d rotation-invariant learning via pose-aware convolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7472–7481, 2022.

Cohen, T. and Welling, M. Steerable CNNs. In *International Conference on Learning Representations*, 2017.

Cohen, T. S., Geiger, M., and Weiler, M. A general theory of equivariant cnns on homogeneous spaces. *Advances in Neural Information Processing Systems*, 32, 2019.

Du, W., Zhang, H., Du, Y., Meng, Q., Chen, W., Zheng, N., Shao, B., and Liu, T.-Y. Se (3) equivariant graph neural networks with complete local frames. In *International Conference on Machine Learning*, pp. 5583–5608. PMLR, 2022.

Du, W., Du, Y., Wang, L., Feng, D., Wang, G., Ji, S., Gomes, C., and Ma, Z.-M. A new perspective on building efficient and expressive 3D equivariant graph neural networks. *arXiv preprint arXiv:2304.04757*, 2023.

Duval, A. A., Schmidt, V., Hernandez-Garcia, A., Miret, S., Malliaros, F. D., Bengio, Y., and Rolnick, D. Faenet: Frame averaging equivariant gnn for materials modeling. In *International Conference on Machine Learning*, pp. 9013–9033. PMLR, 2023.

Dym, N. and Maron, H. On the universality of rotation equivariant point cloud networks. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=6NFBvWlRXaG.

Fuchs, F., Worrall, D., Fischer, V., and Welling, M. Se (3)-transformers: 3d roto-translation equivariant attention networks. *Advances in neural information processing systems*, 33:1970–1981, 2020.

Gebauer, N., Gastegger, M., and Schütt, K. Symmetry-adapted generation of 3D point sets for the targeted discovery of molecules. *Advances in Neural Information Processing Systems*, 32, 2019.

Harris, D. and Bertolucci, M. *Symmetry and Spectroscopy: An Introduction to Vibrational and Electronic Spectroscopy*. Dover Books on Chemistry Series. Dover Publications, 1989. ISBN 9780486661445. URL https://books.google.com/books?id=I3W6oSaRlMsC.

Hinton, G. E., Krizhevsky, A., and Wang, S. D. Transforming auto-encoders. In *International conference on artificial neural networks*, pp. 44–51. Springer, 2011.

Hoffmann, F. *Introduction to crystallography*. Springer Nature, 2020.

Hopcroft, J. An n log n algorithm for minimizing states in a finite automaton. In *Theory of Machines and Computations*, pp. 189–196. Elsevier, 1971.

Hopcroft, J. E. and Wong, J.-K. Linear time algorithm for isomorphism of planar graphs (preliminary report). In *Proceedings of the sixth annual ACM symposium on Theory of computing*, pp. 172–184, 1974.

Hopcroft, J. E., Motwani, R., and Ullman, J. D. Introduction to automata theory, languages, and computation. *Acm Sigact News*, 32(1):60–65, 2001.

Kaba, S.-O., Mondal, A. K., Zhang, Y., Bengio, Y., and Ravanbakhsh, S. Equivariance with learned canonicalization functions. In *International Conference on Machine Learning*, pp. 15546–15566. PMLR, 2023.

Kim, H., Park, J. Y., and Choi, S. Energy refinement and analysis of structures in the QM9 database via a highly accurate quantum chemical method. *Scientific Data*, 6 (1):109, Jul 2019. ISSN 2052-4463. doi: 10.1038/s41597-019-0121-7. URL https://doi.org/10.1038/s41597-019-0121-7.

Kim, P., Chen, J., and Cho, Y. K. SLAM-driven robotic mapping and registration of 3D point clouds. *Automation in Construction*, 89:38–48, 2018.

Kim, S., Park, J., and Han, B. Rotation-invariant local-to-global representation learning for 3d point cloud. In *Advances in Neural Information Processing Systems*, 2020.

Kofinas, M., Nagaraja, N., and Gavves, E. Roto-translated local coordinate frames for interacting dynamical systems. *Advances in Neural Information Processing Systems*, 34: 6417–6429, 2021.

Luo, S., Li, J., Guan, J., Su, Y., Cheng, C., Peng, J., and Ma, J. Equivariant point cloud analysis via learning orientations for message passing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18932–18941, 2022.

Martin, G. *Transformation Geometry: An Introduction to Symmetry*. Undergraduate texts in mathematics. Springer, 1982. ISBN 9783540906360. URL https://books.google.com/books?id=ctY0RwAACAAJ.

McWeeny, R. *Symmetry: An introduction to group theory and its applications*. Courier Corporation, 2002.

Mi, C., Shen, Y., Mi, W., and Huang, Y. Ship identification algorithm based on 3d point cloud for automated ship loaders. *Journal of Coastal Research*, (73):28–34, 2015.

Miller, W. *Symmetry groups and their applications*. Academic Press, 1973.

Mitra, N. J., Gelfand, N., Pottmann, H., and Guibas, L. Registration of point cloud data from a geometric optimization perspective. In *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pp. 22–31, 2004.

Morell, V., Orts, S., Cazorla, M., and Garcia-Rodriguez, J. Geometric 3D point cloud compression. *Pattern Recognition Letters*, 50:55–62, 2014.

Olver, P. J. Lectures on moving frames. 2009.

Passaro, S. and Zitnick, C. L. Reducing so(3) convolutions to so(2) for efficient equivariant gnns. In *Proceedings of the 40th International Conference on Machine Learning*, ICML'23. JMLR.org, 2023.

Pomerleau, F., Colas, F., Siegwart, R., et al. A review of point cloud registration algorithms for mobile robotics. *Foundations and Trends® in Robotics*, 4(1):1–104, 2015.

Puny, O., Atzmon, M., Ben-Hamu, H., Misra, I., Grover, A., Smith, E. J., and Lipman, Y. Frame averaging for invariant and equivariant network design. *arXiv preprint arXiv:2110.03336*, 2021.

Qi, C. R., Su, H., Mo, K., and Guibas, L. J. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

Ramakrishnan, R., Dral, P. O., Rupp, M., and von Lilienfeld, O. A. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific Data*, 1(1):140022, August 2014. ISSN 2052-4463. doi: 10.1038/sdata.2014.22. URL https://doi.org/10.1038/sdata.2014.22.

Romero, D. W., Bekkers, E. J., Tomczak, J. M., and Hoogendoorn, M. Wavelet networks: Scale equivariant learning from raw waveforms. *arXiv preprint arXiv:2006.05259*, 2020.

Satorras, V. G., Hoogeboom, E., and Welling, M. E(n) equivariant graph neural networks. In Meila, M. and Zhang, T. (eds.), *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 9323–9332. PMLR, 18–24 Jul 2021. URL https://proceedings.mlr.press/v139/satorras21a.html.

Schütt, K. T., Sauceda, H. E., Kindermans, P.-J., Tkatchenko, A., and Müller, K.-R. SchNet–a deep learning architecture for molecules and materials. *The Journal of Chemical Physics*, 148(24), 2018.

Simeonov, A., Du, Y., Lin, Y.-C., Garcia, A. R., Kaelbling, L. P., Lozano-Pérez, T., and Agrawal, P. Se (3)-equivariant relational rearrangement with neural descriptor fields. In *Conference on Robot Learning*, pp. 835–846. PMLR, 2023.

Thomas, N., Smidt, T., Kearnes, S., Yang, L., Li, L., Kohlhoff, K., and Riley, P. Tensor field networks: Rotation-and translation-equivariant neural networks for 3D point clouds. *arXiv preprint arXiv:1802.08219*, 2018.

van der Pol, E., Worrall, D., van Hoof, H., Oliehoek, F., and Welling, M. MDP homomorphic networks: Group symmetries in reinforcement learning. *Advances in Neural Information Processing Systems*, 33:4199–4210, 2020.

Villani, C. et al. *Optimal transport: old and new*, volume 338. Springer, 2009.

Vranic, D. V., Saupe, D., and Richter, J. Tools for 3D-object retrieval: Karhunen-Loeve transform and spherical harmonics. In *2001 IEEE Fourth Workshop on Multimedia Signal Processing (Cat. No. 01TH8564)*, pp. 293–298. IEEE, 2001.

Wang, Q. and Kim, M.-K. Applications of 3D point cloud data in the construction industry: A fifteen-year review from 2004 to 2018. *Advanced Engineering Informatics*, 39:306–319, 2019.

Wang, S.-H., Hsu, Y.-C., Baker, J., Bertozzi, A. L., Xin, J., and Wang, B. Rethinking the benefits of steerable features in 3d equivariant graph neural networks. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=mGHJAyR8w0.

Welling, M. and Kipf, T. N. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*, 2016.

Winter, R., Bertolini, M., Le, T., Noé, F., and Clevert, D.-A. Unsupervised learning of group invariant and equivariant representations. *Advances in Neural Information Processing Systems*, 35:31942–31956, 2022.

Wolter, J. D., Woo, T. C., and Volz, R. A. Optimal algorithms for symmetry detection in two and three dimensions. *The Visual Computer*, 1:37–48, 1985.

Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., and Xiao, J. 3d shapenets: A deep representation for volumetric shapes. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1912–1920, 2015. doi: 10.1109/CVPR.2015.7298801.

Zhao, Y., Birdal, T., Lenssen, J. E., Menegatti, E., Guibas, L., and Tombari, F. Quaternion equivariant capsule networks for 3d point clouds. In *European Conference on Computer Vision*, pp. 1–19. Springer, 2020.

Appendix for

An Explicit Frame Construction for Normalizing 3D Point Clouds

A. Additional Details for Frame and Normalization

We first show the existence of frames. Notably, there are similarities with certain results presented in (Winter et al., 2022).

Proposition A.1 (Existence of frames)**.** *There is a function $\mathcal{F} : X \to G$ satisfying the G-relaxed equivariance condition that for any $x \in X$ and $g_1 \in G$, there exists a $g_2 \in g_1 G_x$ such that*

$$\mathcal{F}(g_1 \cdot x) = g_2 \cdot \mathcal{F}(x). \tag{11}$$

Proof. Consider the orbits G_x, forming a partition of the set X where each point resides in a unique orbit. Within each G_x, a canonical representative can be chosen following the approach outlined in (Winter et al., 2022). This selection induces a function $\mu : X \to X$ by mapping each $x \in X$ to its canonical representative within its orbit. As the output solely depends on the input's orbits, μ is a G-invariant function.

Moreover, since x and $\mu(x)$ lie in the same orbit, there exists a group element $g_x \in G$ such that $x = g \cdot \mu(x)$. While multiple group elements may satisfy this condition, we choose one (by Axiom of Choice). This defines a well-defined function $\mathcal{F} : X \to G$ by mapping x to g_x. Note that through this construction, we have $\mu(x) = \mathcal{F}(x)^{-1} \cdot x$ for any x. To demonstrate that $\mathcal{F}$ is a frame, let $g_1 \in G$. Since μ is G-invariant, we have

$$\mathcal{F}(x)^{-1} \cdot x = \mu(x) = \mu(g_1 \cdot x) = \mathcal{F}(g_1 \cdot x)^{-1} \cdot g_1 \cdot x.$$

Then we obtain $g_1^{-1}\mathcal{F}(g_1 \cdot x)\mathcal{F}(x)^{-1} \cdot x = x$, which implies that $g_1^{-1}\mathcal{F}(g_1 \cdot x)\mathcal{F}(x)^{-1} \in G_x$. Therefore, we obtain the desired property as follows:

$$g_1^{-1}\mathcal{F}(g_1 \cdot x)\mathcal{F}(x)^{-1} \in G_x \Leftrightarrow \mathcal{F}(g_1 \cdot x)\mathcal{F}(x)^{-1} \in g_1 G_x \Leftrightarrow \mathcal{F}(g_1 \cdot x) = g_2 \cdot \mathcal{F}(x) \text{ for some } g_2 \in g_1 G_x. \tag{12}$$

□

Next, we show that the canonicalization of an arbitrary function f between G-spaces X and Y satisfies relaxed equivariance. Since this is a well-established result in (Kaba et al., 2023), we provide a very concise proof.

Proposition A.2 (G-relaxed equivariance of the canonicalization of functions)**.** *Let X and Y be two G-spaces, and let $\mathcal{F} : X \to G$ be a frame. Then the canonicalization of any function $f : X \to Y$ through $\mathcal{F}$, defined as: $\phi(x) \coloneqq \mathcal{F}(x) \cdot f(\mathcal{F}(x)^{-1} \cdot x)$, is a G-relaxed equivariant function.*

Proof. We observe that $\mathcal{F}(x)^{-1} \cdot x$ is G-invariant. Therefore, $f(\mathcal{F}(x)^{-1} \cdot x)$ must be G-invariant. Then the G-relaxed equivariance of ϕ is inherited from that of $\mathcal{F}$. □

It is evident that if a function is G-equivariant, then its canonicalization remains G-equivariant.

B. Additional Details for Frame Construction

We review and provide a theoretical guarantee of our strategy mentioned in Section 4. First, we extract a sequence of linearly independent vectors $\boldsymbol{v}_1(\boldsymbol{X}), \ldots, \boldsymbol{v}_{\mathrm{rk}(\boldsymbol{X})}(\boldsymbol{X})$ within the point cloud $\{\!\{\boldsymbol{x}_i\}\!\}$ that satisfy the condition: for any $1 \leq i \leq \mathrm{rk}(\boldsymbol{X})$, $g_1 \in G$, $\sigma \in \mathrm{S}_m$, there exists a $g_2 \in g_1 \mathrm{Sym}(\{\!\{\boldsymbol{x}_i\}\!\})$ such that

$$\boldsymbol{v}_i((\sigma, g_1) \cdot \boldsymbol{X}) = g_2 \boldsymbol{v}_i(\boldsymbol{X}). \tag{13}$$

Then we construct $\mathcal{F}$ through the following process:

Step 1: align $\boldsymbol{v}_1$ with the x-axis via a rotation $\boldsymbol{R}_1$.

Step 2: rotate along the x-axis to place $\boldsymbol{v}_2$ in the x-y plane with a positive y-component via $\boldsymbol{R}_2$.

Step 3: reflect along x-y plane to ensure a positive z-component for the vector $\boldsymbol{x}_3$ via $\boldsymbol{R}_3$ (skip this step if $G = \mathrm{SO}(3)$)

If the number of steps surpasses $\mathrm{rk}(\boldsymbol{X})$, the process is terminated. Finally, we define the frame as

$$\mathcal{F}(\boldsymbol{X}) \coloneqq \boldsymbol{R}_1^\top \cdot \cdots \cdot \boldsymbol{R}_{\mathrm{rk}(\boldsymbol{X})}^\top,$$

for $G = \mathrm{O}(3)$ (as $\mathcal{F}(\boldsymbol{X}) \coloneqq \boldsymbol{R}_1^\top \cdot \cdots \cdot \boldsymbol{R}_{\min\{2,\mathrm{rk}(\boldsymbol{X})\}}^\top$ for $G = \mathrm{SO}(3)$).

Proposition B.1. *The function $\mathcal{F} : \mathbb{R}^{3\times m} \to G$ defined above satisfies the relaxed equivariance in equation (6).*

Proof. We observe that it is enough to show that each column vectors in $\mathcal{F}(\boldsymbol{X})$ satisfy the relaxed equivariance. In particular, we will see that its columns are generated by the Gram-Schmidt process.

Define $\boldsymbol{R}_1$ to align $\boldsymbol{v}_1$ with the first standard basis vector $\boldsymbol{e}_1$, ensuring $\|\boldsymbol{v}_1\|\boldsymbol{e}_1 = \boldsymbol{R}_1\boldsymbol{v}_1$. This yields the first column of $\boldsymbol{R}_1^\top$ as $\frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}$, which inherently satisfies relaxed equivariance.

Construct the intermediate vector $\tilde{\boldsymbol{v}}_2 = \boldsymbol{R}_1\boldsymbol{v}_2 - \langle \boldsymbol{R}_1\boldsymbol{v}_2, \boldsymbol{e}_1\rangle \boldsymbol{e}_1$. Define $\boldsymbol{R}_2$ to map $\tilde{\boldsymbol{v}}_2$ to the second standard basis vector while preserving $\boldsymbol{e}_1$. This results in the second column of $\boldsymbol{R}_2^\top\boldsymbol{R}_1^\top$ being $\frac{\boldsymbol{v}_2 - \langle \boldsymbol{v}_2, \boldsymbol{v}_1\rangle \boldsymbol{v}_1}{\|\boldsymbol{v}_2 - \langle \boldsymbol{v}_2, \boldsymbol{v}_1\rangle \boldsymbol{v}_1\|}$, which also satisfies relaxed equivariance.

Define $\boldsymbol{R}_3$ as a reflection along the $x - y$ plane, leaving the first two columns of $\boldsymbol{R}_3^\top\boldsymbol{R}_2^\top\boldsymbol{R}_1^\top$ unchanged. Consider $\tilde{\boldsymbol{v}}_3 = \boldsymbol{R}_2\boldsymbol{R}_1\boldsymbol{v}_3 - \langle \boldsymbol{R}_2\boldsymbol{R}_1\boldsymbol{v}_3, \boldsymbol{e}_1\rangle \boldsymbol{e}_1 - \langle \boldsymbol{R}_2\boldsymbol{R}_1\boldsymbol{v}_3, \boldsymbol{e}_2\rangle \boldsymbol{e}_2$. We see that $\boldsymbol{R}_3\tilde{\boldsymbol{v}}_3 = \|\tilde{\boldsymbol{v}}_3\|\boldsymbol{e}_3$, resulting in the third column being the normalized projection of $\boldsymbol{v}_3$ onto the orthogonal complement of $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, i.e. the normalization of $\boldsymbol{v}_3 - \langle \boldsymbol{v}_3, \boldsymbol{v}_1\rangle \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|} - \langle \boldsymbol{v}_3, \boldsymbol{v}_2 - \langle \boldsymbol{v}_2, \boldsymbol{v}_1\rangle \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}\rangle \frac{\boldsymbol{v}_2 - \langle \boldsymbol{v}_2, \boldsymbol{v}_1\rangle \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}}{\|\boldsymbol{v}_2 - \langle \boldsymbol{v}_2, \boldsymbol{v}_1\rangle \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}\|}$.

Notice that the last two column vectors of $\boldsymbol{R}_1^\top$ when $\mathrm{rk}(\boldsymbol{X}) = 1$ (or the last column vector of $\boldsymbol{R}_2^\top\boldsymbol{R}_1^\top$ when $\mathrm{rk}(\boldsymbol{X}) = 2$) reside outside the space spanned by the vectors $\boldsymbol{x}_i$ within the point cloud. These vectors inherently satisfy the relaxed equivariance property. This is because relaxed equivariance allows for some flexibility in the output transformation, specifically permitting differences by elements within the stabilizer.

□

C. Inherent Symmetry

We reference Martin (1982) as a key source for understanding point set symmetries. Specifically, symmetries in two-dimensional point sets are described by Leonardo's Theorem C.1, and in three-dimensional point sets by Hessel's Theorem C.2.

Point Group	Symmetry Elements	Order of the Group
C_1	E	1
C_i	i	2
C_s	σ	2
C_n	C_n	n
$S_n^\dagger$	S_n	n
C_{nv}	C_n, σ_v	$2n$
C_{nh}	C_n, σ_h	$2n$
D_n	$C_n, \perp C_2$	$2n$
D_{nd}	$C_n, \perp C_2, \sigma_d$	$4n$
D_{nh}	$C_n, \perp C_2, \sigma_h$	$4n$
$C_{\infty v}$	$\mathrm{rk}(\boldsymbol{X}) = 1$ without i	∞
$D_{\infty h}$	$\mathrm{rk}(\boldsymbol{X}) = 1, i$	∞

Table 4. A list of point groups and their symmetry elements. (Harris & Bertolucci, 1989). †: n must be even or $S_n = C_{nh}$

Theorem C.1 (Leonardo's Theorem)**.** *A finite group of isometries for a subset of $\mathbb{R}^2$ is either a cyclic group C_n or a dihedral group D_n.*

Theorem C.2 (Hessel's Theorem)**.** *A finite grup of isometries on a subset of $\mathbb{R}^3$ is one of the following:*

- *The rotation groups $T, C,\ I$ of the platonic solids tetrahedron, cube, and icosahedron*
- *The cyclic groups C_n, $n = 1, 2, \ldots$, and the dihedral groups D_n, $n = 2, 3, \ldots$.*
- *The groups $\bar{T}, \bar{C}, \bar{I}; \bar{C}_1, \bar{C}_2, \ldots; \bar{D}_1, \bar{D}_2, \ldots$; where $\bar{G}$ means the group generated by G and a reflection at some point (inversion).*
- *The groups CT, $C_{2n}C_n$, $D_{2n}D_n$, $n = 2, 3, \ldots$, where GH means $H \cup (G - H) \odot i$ where i is an inversion.*

D. Missing proofs

Proposition D.1. *For point clouds in $\mathbb{R}^{n\times m}$ with $m, n \geq 3$, it is impossible to construct a frame $\mathcal{F} : \mathbb{R}^{n\times m} \to G$ that is continuous across the entire domain when $G = \mathrm{E}(n), \mathrm{SE}(n), \mathrm{O}(n)$ or $\mathrm{SO}(n)$.*

Proof. Without loss of generality, we can narrow our attention to the case of $n = m = 3$. This is achieved by focusing on the first 3×3 minor of any matrix within $\mathbb{R}^{n\times m}$.

Denoting the standard basis vectors of $\mathbb{R}^3$ as $\boldsymbol{e}_1, \boldsymbol{e}_2$, and $\boldsymbol{e}_3$, let $\mathcal{F} : \mathbb{R}^{3\times 3} \to G$ be an abitrary frame. Our goal is to demonstrate that the discontinuity of $\mathcal{F}$ at $\boldsymbol{X}_0 = [\boldsymbol{e}_1, \boldsymbol{0}, \boldsymbol{0}]$. To establish the discontinuity, we seek an $\epsilon > 0$ such that for any $\delta > 0$, there exists a $\boldsymbol{Z} \in \mathbb{R}^{3\times 3}$ satisfying $\|\boldsymbol{X}_0 - \boldsymbol{Z}\|_F < \delta$, yet $\|\mathcal{F}(\boldsymbol{X}_0) - \mathcal{F}(\boldsymbol{Z})\| > \epsilon$.

For $G = \mathrm{O}(3)$ or $\mathrm{SO}(3)$.

Consider $\boldsymbol{X}_{\delta,\theta} = [\boldsymbol{e}_1, \frac{\delta}{2}\boldsymbol{R}_\theta \boldsymbol{e}_2, \frac{\delta}{3}\boldsymbol{R}_\theta \boldsymbol{e}_3] \in \mathbb{R}^{3\times 3}$ where $\boldsymbol{R}_\theta \in \mathrm{SO}(3)$ denote the rotation matrix with axis $\boldsymbol{e}_1$ and angle θ. Note that $\|\boldsymbol{X}_0 - \boldsymbol{X}_{\delta,\theta}\| < \delta$.

Denoting $\mathcal{F}(\boldsymbol{X}_{\delta,0})$ as g_δ, the equivariance of $\mathcal{F}$ and the fact that the stabilizer of $\boldsymbol{X}_{\delta,0}$ is trivial imply $\mathcal{F}(\boldsymbol{X}_{\delta,\theta}) = \boldsymbol{R}_\theta \mathcal{F}(\boldsymbol{X}_{\delta,0}) = \boldsymbol{R}_\theta g_\delta$. Therefore,

$$\|\mathcal{F}(\boldsymbol{X}_0) - \mathcal{F}(\boldsymbol{X}_{\delta,\theta})\| = \|\mathcal{F}(\boldsymbol{X}_0) - \boldsymbol{R}_\theta g_\delta\| = \|\mathcal{F}(\boldsymbol{X}_0) g_\delta^{-1} - \boldsymbol{R}_\theta\|. \tag{14}$$

It suffices to find an $\epsilon > 0$ such that for any $g \in \mathrm{O}(3)$, we have $\|g - \boldsymbol{R}_\theta\| > \epsilon$ for some θ. Fix a θ_0 and choose an $\epsilon > 0$ such that $\|\boldsymbol{R}_{-\theta_0} - \boldsymbol{R}_{\theta_0}\| = 3\epsilon$. By the triangle inequality, $\|g - \boldsymbol{R}_{-\theta_0}\| + \|g - \boldsymbol{R}_{\theta_0}\| \geq \|\boldsymbol{R}_{-\theta_0} - \boldsymbol{R}_{\theta_0}\| = 3\epsilon$, implying $\|g - \boldsymbol{R}_{-\theta_0}\| > \epsilon$ or $\|g - \boldsymbol{R}_{\theta_0}\| > \epsilon$. Thus, $\|g - \boldsymbol{R}_\theta\| > \epsilon$ for $\theta = \theta_0$ or $-\theta_0$, completing the argument for $G = \mathrm{O}(3)$ or $\mathrm{SO}(3)$.

For $G = \mathrm{E}(3)$ or $\mathrm{SE}(3)$.

Remark that $\mathrm{E}(3)$ (resp. $\mathrm{SE}(3)$) has the induced topology from $\mathrm{GL}(4)$ through the inclusion $\mathrm{E}(3) \subset \mathrm{GL}(4)$ (resp. $\mathrm{SE}(3) \subset \mathrm{GL}(4)$), defined as follows:

$$(\boldsymbol{Q}, \boldsymbol{t}) \in \mathrm{E}(3) \mapsto \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{t} \\ 0 & 1 \end{bmatrix}. \tag{15}$$

Consider again the candidate matrix $\boldsymbol{X}_{\delta,\theta}$. Denote $\mathcal{F}(\boldsymbol{X}_0)$ as $(\boldsymbol{Q}_0, \boldsymbol{t}_0)$ and $\mathcal{F}(\boldsymbol{X}_{\delta,0})$ as $(\boldsymbol{Q}_\delta, \boldsymbol{t})$. Similarly, we obtain $\mathcal{F}(\boldsymbol{X}_{\delta,\theta}) = (\boldsymbol{R}_\theta, \boldsymbol{0}) \cdot \mathcal{F}(\boldsymbol{X}_{\delta,0}) = (\boldsymbol{R}_\theta \boldsymbol{Q}_\delta, \boldsymbol{t})$. Since $\|\mathcal{F}(\boldsymbol{X}_0) - \mathcal{F}(\boldsymbol{X}_{\delta,\theta})\| \geq \|\boldsymbol{Q}_0 - \boldsymbol{R}_\theta \boldsymbol{Q}_\delta\|$, an analogous argument establishes the existence of an $\epsilon > 0$ such that for any δ, there exists a θ satisfying $\|\mathcal{F}(\boldsymbol{X}_0) - \mathcal{F}(\boldsymbol{X}_{\delta,\theta})\| > \epsilon$.

Moreover, one can check that the example $\boldsymbol{X}_{\delta,\theta}$ employed in this proof does not exhibit any inherent symmetry.

□

Theorem 3.3. *Let X, Y be two normed G-spaces where the group action are defined by the orthogonal representations*[6] *$\rho_X : G \to \mathrm{GL}(X)$ and $\rho_Y : G \to \mathrm{GL}(Y)$, respectively. Assume that X is finite-dimensional. Suppose f is a parameterized function from X to Y. Define ϕ as the canonicalization of f defined in equation (4) through an arbitrary frame $\mathcal{F} : X \to G$. Then ϕ is a universal approximator of G-equivariant continuous functions as long as f is a universal approximator of continuous functions.*

[6] Orthogonality implies that the image of group representations consists of orthogonal matrices.

Proof. The proof draws inspiration from the methodology introduced by (Kaba et al., 2023) but diverges by avoiding dependence on the continuity of functions. In particular, we utilize the orthogonality of representations to address the absence of continuity.

First of all, let ψ be an arbitrary G-equivariant function. Through the construction of ϕ and the equivariance of ψ, we have

$$\begin{aligned}\|\psi(x)-\phi(x)\| &= \left\|\rho_W(\mathcal{F}(x))\psi\left(\rho_V(\mathcal{F}(x))^{-1}x\right)-\rho_W(\mathcal{F}(x))f\left(\rho_V(\mathcal{F}(x))^{-1}x\right)\right\| \\ &= \left\|\rho_W(\mathcal{F}(x))\left(\psi\left(\rho_V(\mathcal{F}(x))^{-1}x\right)-f\left(\rho_V(\mathcal{F}(x))^{-1}x\right)\right)\right\|.\end{aligned} \tag{16}$$

Given that ρ_W is orthogonal, $\rho_W(\mathcal{F}(x))$ preserves norms. Consequently, we deduce

$$\|\psi(x)-\phi(x)\| = \left\|\psi\left(\rho_V(\mathcal{F}(x))^{-1}x\right)-f\left(\rho_V(\mathcal{F}(x))^{-1}x\right)\right\|, \text{ for any } x\in X. \tag{17}$$

Let $K\subset X$ be a compact set, and $\epsilon>0$ be any positive number. Consider the closure of the set $\{\rho_V(\mathcal{F}(x))^{-1}x \mid x\in K\}$, denoted by $\tilde{K}$. We assert that $\tilde{K}$ is compact. Evidently, it is clear that $\tilde{K}$ is closed. As K is compact and hence bounded, there is an $r>0$ such that $x\in B_r$ for any $x\in K$ with B_r denoting the closed ball centered at the origin with radius r. Using the orthogonality of ρ_V, $\|\rho_V(\mathcal{F}(x))^{-1}x\|=\|x\|\le r$ for any $x\in K$. This implies $\{\rho_V(\mathcal{F}(x))^{-1}x \mid x\in K\}\subset B_r$. As B_r is closed, $\tilde{K}\subset B_r$, which shows that $\tilde{K}$ is bounded and thus is compact.

Now, employing the assumption on the universality of f, there exists a choice of parameters such that:

$$\|\psi(\tilde{x})-f(\tilde{x})\|<\epsilon, \text{ for any } \tilde{x}\in\tilde{K}. \tag{18}$$

Combining this with equation (17), we derive the desired result,

$$\|\psi(x)-\phi(x)\| = \left\|\psi\left(\rho_V(\mathcal{F}(x))^{-1}x\right)-f\left(\rho_V(\mathcal{F}(x))^{-1}x\right)\right\|<\epsilon, \text{ for any } x\in K. \tag{19}$$

□

Theorem 3.4. *For $G=\mathrm{E}(n)$ or $\mathrm{SE}(n)$, the function $\mathcal{F}:\mathbb{R}^{n\times m}\to G$ constructed via equation (5) is a frame. Moreover, let $f:\mathbb{R}^{n\times m}\to\mathbb{R}^{n\times m}$ be a parameterized function, and ϕ be its canonicalization through $\mathcal{F}$ defined in equation (3). Then ϕ is a universal approximator of G-equivariant continuous functions if f is a universal approximator of continuous functions.*

Proof. We first show that $\mathcal{F}:\mathbb{R}^{n\times m}\to\mathrm{E}(n)$ is a frame. Observe that for any $\boldsymbol{X}\in\mathbb{R}^{n\times m}$, the stabilizer $G_{\boldsymbol{X}}$ of $\boldsymbol{X}$ is a subgroup of $\mathrm{O}(n)$ since the action of translations is free. Now consider $g_1\in\mathrm{O}(n)$ and $\boldsymbol{t}\in\mathrm{T}(n)$. We have

$$\begin{aligned}\mathcal{F}\left((g_1,\boldsymbol{t})\cdot\boldsymbol{X}\right)=\mathcal{F}\left(g_1\boldsymbol{X}+\boldsymbol{t}\mathbf{1}_m^\top\right) &= \left(\mathcal{F}'\left(g_1\boldsymbol{X}+\boldsymbol{t}\mathbf{1}_m^\top-\frac{1}{m}\left(g_1\boldsymbol{X}+\boldsymbol{t}\mathbf{1}_m^\top\right)\mathbf{1}_m\mathbf{1}_m^\top\right),\frac{1}{m}\left(g_1\boldsymbol{X}+\boldsymbol{t}\mathbf{1}_m^\top\right)\mathbf{1}_m\right) \\ &= \left(\mathcal{F}'\left(g_1\left(\boldsymbol{X}-\frac{1}{m}\boldsymbol{X}\mathbf{1}_m\mathbf{1}_m^\top\right)\right),\frac{1}{m}\left(g_1\boldsymbol{X}+\boldsymbol{t}\mathbf{1}_m^\top\right)\mathbf{1}_m\right) \\ &= \left(g_2\mathcal{F}'\left(\boldsymbol{X}-\frac{1}{m}\boldsymbol{X}\mathbf{1}_m\mathbf{1}_m^\top\right),g_2\left(\frac{1}{m}\boldsymbol{X}\mathbf{1}_m\right)+\boldsymbol{t}\right) \\ &= (g_2,\boldsymbol{t})\cdot\left(\mathcal{F}'\left(\boldsymbol{X}-\frac{1}{m}\boldsymbol{X}\mathbf{1}_m\mathbf{1}_m^\top\right),\frac{1}{m}\boldsymbol{X}\mathbf{1}_m\right) \\ &= (g_2,\boldsymbol{t})\cdot\mathcal{F}(\boldsymbol{X})\end{aligned} \tag{20}$$

for some $g_2\in g_1G_{\boldsymbol{X}}$. This implies that $\mathcal{F}$ is a frame.

Next, let ψ be an arbitrary G-equivariant continuous function. Through the construction of ϕ and the equivariance of ψ, we have

$$\begin{aligned}\|\psi(\boldsymbol{X})-\phi(\boldsymbol{X})\| &= \left\|\mathcal{F}(\boldsymbol{X})\cdot\psi\left(\mathcal{F}(\boldsymbol{X})^{-1}\cdot\boldsymbol{X}\right)-\mathcal{F}(\boldsymbol{X})\cdot f\left(\mathcal{F}(\boldsymbol{X})^{-1}\cdot\boldsymbol{X}\right)\right\| \\ &= \left\|\mathcal{F}'(\boldsymbol{X}-\frac{1}{m}\boldsymbol{X}\mathbf{1}_m\mathbf{1}_m^\top)\cdot\psi\left(\mathcal{F}(\boldsymbol{X})^{-1}\cdot\boldsymbol{X}\right)-\mathcal{F}'(\boldsymbol{X}-\frac{1}{m}\boldsymbol{X}\mathbf{1}_m\mathbf{1}_m^\top)\cdot f\left(\mathcal{F}(\boldsymbol{X})^{-1}\cdot\boldsymbol{X}\right)\right\|\end{aligned} \tag{21}$$

Since $\mathcal{F}'(\boldsymbol{X})$ preserves norms, we deduce

$$\|\psi(\boldsymbol{X})-\phi(\boldsymbol{X})\| = \left\|\psi\left(\mathcal{F}(\boldsymbol{X})^{-1}\cdot\boldsymbol{X}\right)-f\left(\mathcal{F}(\boldsymbol{X})^{-1}\cdot\boldsymbol{X}\right)\right\|, \text{ for any } \boldsymbol{X}. \tag{22}$$

Similar to the proof of Theorem 3.3, it suffices to show that for any compact set $K \subset \mathbb{R}^{n\times m}$, the set $\{\mathcal{F}(\boldsymbol{X})^{-1}\cdot\boldsymbol{X} \mid \boldsymbol{X} \in K\}$ is bounded. Notice that

$$\begin{aligned}\|\mathcal{F}(\boldsymbol{X})^{-1}\cdot\boldsymbol{X}\| &= \left\|\mathcal{F}'\left(\boldsymbol{X}-\frac{1}{m}\boldsymbol{X}\mathbf{1}_m\mathbf{1}_m^\top\right)^{-1}\left(\boldsymbol{X}-\frac{1}{m}\boldsymbol{X}\mathbf{1}_m\mathbf{1}_m^\top\right)\right\| \\ &= \left\|\boldsymbol{X}-\frac{1}{m}\boldsymbol{X}\mathbf{1}_m\mathbf{1}_m^\top\right\| \\ &= \left\|\boldsymbol{X}\left(\boldsymbol{I}_m-\frac{1}{m}\mathbf{1}_m\mathbf{1}_m^\top\right)\right\|,\end{aligned} \tag{23}$$

where $\boldsymbol{I}_m$ denotes the $m\times m$ identity matrix. The desired result then follows from the fact that $\boldsymbol{I}_m - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m^\top$ is a bounded operator. Therefore, we conclude that ϕ is a universal approximation of G-equivariant continuous function. □

Proposition D.2. *Suppose $\mathcal{H}\subseteq\mathcal{G}$ is a subgraph that reconstructs $\mathcal{G}$. If $\mathcal{H}$ is minimal then for any $\sigma\in\mathrm{Aut}(\mathcal{G})$, we have $\sigma\cdot\mathcal{H}\neq\mathcal{H}$.*

Proof. Suppose $\mathcal{H}$ is minimal and there is a non-trivial σ in $\mathrm{Aut}(\mathcal{G})$ such that $\sigma\cdot\mathcal{H}=\mathcal{H}$. Since σ is non-trivial, it must map at least one edge e_{jk} in $\mathcal{H}$ to a distinct edge $e_{\sigma(j)\sigma(k)}$ in $\mathcal{H}$. Create a proper subgraph $\mathcal{H}'\subset\mathcal{H}$ by removing $e_{\sigma(j)\sigma(k)}$ from $\mathcal{H}$. Notice that $\sigma\mathcal{H}'\cup\mathcal{H}'=\mathcal{H}$, which implies $\mathcal{H}'$ can reconstruct the entire graph since $\mathcal{H}$ can. This contradicts the assumption of $\mathcal{H}$ being minimal, as we have found a proper subgraph $\mathcal{H}'$ capable of reconstructing $\mathcal{H}$.

□

E. n-Body problem

In the following section, we consider the n-body task (Satorras et al., 2021), which uses initial positions and velocities to predict the trajectory of five charged atomic particles. For this task, we use the architectures, training procedures, and hyperparameters as described in (Satorras et al., 2021). Different from (Satorras et al., 2021), we construct the 3000 training trajectories (augmented dataset) by performing six random rotations on 500 trajectories. Then the training is performed on the 3000 trajectories, with 2000 unique trajectories used for validation and 2000 unique trajectories used for testing. In this task, we see a direct comparison between the canonicalization technique of ASUN and architecture-dependent models like TFN, EGNN, and FA. The baseline models also include a linear model and a GNN model adapted from (Satorras et al., 2021).

Method	MSE	Forward time (s)
Linear	0.0714	0.0001
TFN (Thomas et al., 2018)	0.0343	0.0343
GNN	0.0165	0.0032
EGNN (Satorras et al., 2021)	0.0141	0.0062
FA (Puny et al., 2021)	0.0177	0.0041
ASUN-GNN	**0.0139**	0.0032

Table 5. Results for n-body trajectory prediction. We illustrate the best performance in bold.

From Table 5 we observe the benefits of using the data-free and architecture-independent method of ASUN. In particular, we observe that ASUN – used in the baseline GNN setting (ASUN-GNN) – achieves the best results with no additional computational overhead per epoch. This is because ASUN is a data-processing step and does not restrict the architectures. This allows for the use of fast architectures like GNN. ASUN is highly generalizable to real-world datasets that have not been meticulously cleaned or processed.

In Table 6, we empirically demonstrate that message message-passing neural network (MPNN) does not interfere with the universality of canonicalization. This task utilizes the same n-body task as described above. However, this data is not augmented with additional data. We observe that the model performance improves slightly when using ASUN. However, crucially, it does not decrease the model performance or the computational time. Importantly, this shows that ASUN preserves the universality of the methods.

Method	MSE	Forward time (s)
TFN	0.0155	.0343
GNN	0.0107	.0032
EGNN	0.0071	.0062
FA	0.0057	.0041
ASUN-TFN	0.0135	.0343
ASUN-GNN	0.0102	.0034
ASUN-EGNN	0.0068	.0062

Table 6. Results for n-body trajectory prediction.

F. Additional Results

Figure 7 illustrates the full DFA for the NH_3 molecule. In particular, we denote invariant features as $((z_H, H), v_H)$ or $((z_N, N), v_N)$ for hydrogen and nitrogen atoms respectively. The node indices are given by $\{0, 1, 2, 3\}$. The state transitions on the DFA are discussed in (Alt et al., 1988) and are denoted $\boldsymbol{a}$ and $\boldsymbol{b}$.

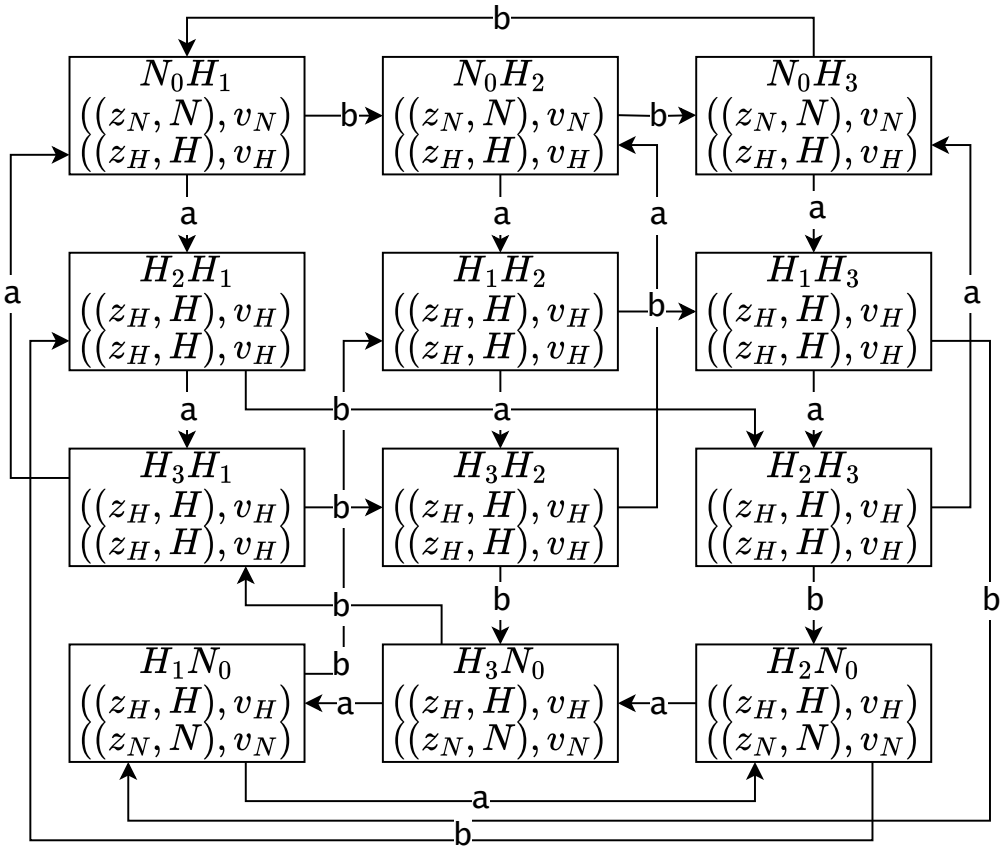


Figure 7. The entire DFA for the NH_3 molecule. Invariant features are given by $\{H, N\}$, $z_{\{H,N\}}$, and $v_{\{H,N\}}$, node indices are given by $\{0, 1, 2, 3\}$ and state transitions are marked $\boldsymbol{a}, \boldsymbol{b}$.