ON DISCRETE SYMMETRIES OF ROBOTICS SYSTEMS: A GROUP-THEORETIC AND DATA-DRIVEN ANALYSIS

Anonymous authors
Paper under double-blind review

ABSTRACT

In this work, we study the Morphological Symmetries of dynamical systems with one or more planes of symmetry, a predominant feature in animal biology and robotic systems, characterized by the duplication and balanced distribution of body parts. These morphological symmetries imply that the system's dynamics are symmetric (or approximately symmetric), which in turn imprints symmetries in optimal control policies and in all proprioceptive and exteroceptive measurements related to the evolution of the system's dynamics. For data-driven methods, symmetry represents an inductive bias that justifies data augmentation and the construction of symmetric function approximators. To this end, we use Group Theory to present a theoretical and practical framework allowing for (1) the identification of the system's morphological symmetry Group $\mathcal{G}$, (2) the characterization of how the group acts upon the system state variables and any relevant measurement living in the Euclidean space, and (3) the exploitation of data symmetries through the use of $\mathcal{G}$-equivariant/$\mathcal{G}$-invariant Neural Networks, for which we present experimental results on synthetic and real-world applications, demonstrating how symmetry constraints lead to better sample efficiency and generalization while reducing the number of trainable parameters.

1 INTRODUCTION

Symmetries are a predominant feature in animal biology. The majority of living (and extinct) species are bilaterally or radially symmetric (i.e., having one or more planes of symmetry), a property intuitively recognized by the patterns of balanced distribution and duplication of body parts and shapes (Holló, 2017). Likewise, most robotic systems are symmetric, often featuring more precise symmetries than nature due to accurate duplication of body parts and the tendency to design mechanisms with symmetric volumes and mass distributions. These morphological symmetries of animals and robots imply that the dynamics and control of body motions are also approximately symmetric, resulting in all proprioceptive and exteroceptive measurements, related to the evolution of the system's dynamics (e.g. joint torques, depth images, contact forces), to be also symmetric. This highly relevant inductive bias is frequently left unexploited in most data-driven applications in the fields of robotics, computer graphics, and control.

Recent works in computer graphics (Yeh et al., 2019; Abdolhosseini et al., 2019; Yu et al., 2018) and robotics/dynamical systems (Van der Pol et al., 2020; Ordonez-Apraez et al., 2022; Hamed & Grizzle, 2013) have exploited through different approaches the morphological symmetry group associated with bilateral (or sagittal) symmetry (the reflection group $\mathcal{C}_2$), obtaining improvements in control robustness and in generalization and sample efficiency of function approximators. Notably, Zinkevich & Balch (2001) proved that Markov Decision Processes with state symmetries have symmetric optimal value and policy functions. Despite these encouraging contributions, exploiting the inductive bias of morphological symmetries is not a widespread technique in the research community. We attribute the scarce adoption of these techniques to the lack of a unifying theoretical and practical framework, allowing to identify different morphological symmetries in arbitrary dynamical systems and efficiently and conveniently exploit them in data-driven applications.

This work takes a step towards this unifying framework by studying morphological symmetries through the lens of group theory, the field of mathematics that studies symmetries, which is broadly used in Machine Learning (ML). In summary, our theoretical contributions are:

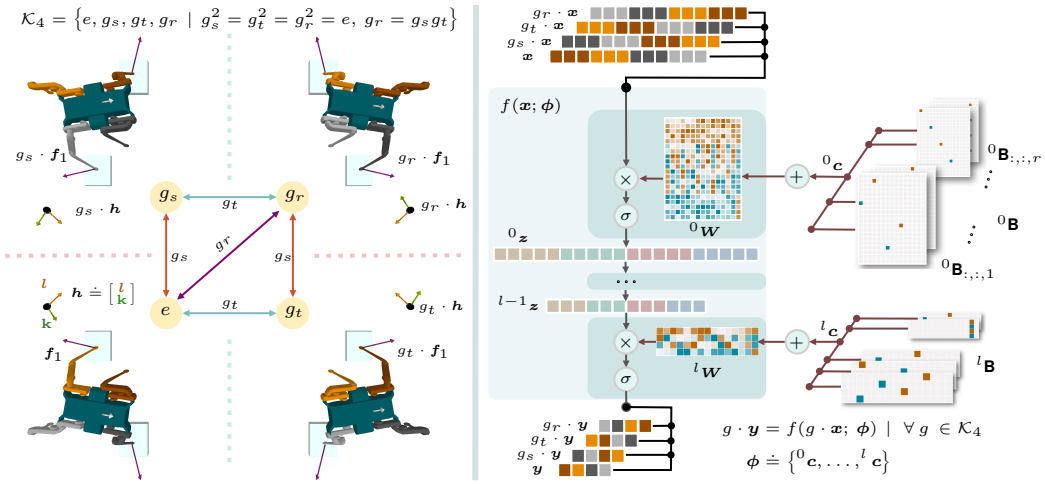


Figure 1: **Left:** Caley diagram and top-view (see 3D animation) of symmetric configurations of the quadruped robot Solo, whose morphological symmetries (described by the Klein four-group $\mathcal{K}_4$) allow it to imitate the effect of reflections (g_s, g_t) and $180°$ rotations of space (g_r). Transformations affect both proprioceptive (state space, CoM linear $\boldsymbol{l}$ and angular $\mathbf{k}$ momentum) and exteroceptive (terrain elevation, external disturbances) quantities. **Right:** Diagram of a $\mathcal{K}_4$-equivariant NN. Each of the layer's linear maps $\boldsymbol{W}$ is constructed as a weighted average of the basis of the space of equivariant linear maps $\mathbf{B}$, computed from the $\mathcal{K}_4$ symmetries of the input-output spaces (see section 5).

- ❊ A mathematical formalization of the concept of discrete morphological symmetries.
- ❊ Characterization of how morphological symmetries affect the system's state variables and any relevant proprioceptive and exteroceptive measurements in Euclidean or joint space.

These contributions open the door for the identification of morphological symmetries in robotic systems and their exploitation through data augmentation or the construction of equivariant function approximators. Additionally, our practical contributions focus on the efficient construction and use of $\mathcal{G}$-equivariant neural networks, for arbitrary morphological symmetry groups $\mathcal{G}$, for which we:

- ✣ Demonstrate that there is a reduction of the architecture's trainable parameters of approximately $1/|\mathcal{G}|$, being $|\mathcal{G}|$ the group order.
- ✣ Derived an optimal initialization for the trainable parameters of an equivariant layer.
- ✣ Provide open-access code[1] for the construction of $\mathcal{G}$-equivariant architectures.

2 Background on Symmetry Groups

In a nutshell, a symmetry group in Group Theory is an abstraction of the concept of symmetries that different geometric objects might have, understanding symmetry as a transformation that when applied to an object conserves a relevant property of its structure. For instance, in fig. 1-left the Klein four-group $\mathcal{K}_4$ describes the symmetries that vectors, pseudo-vectors, rigid bodies, and a quadruped robot have to $180°$ rotations (g_r) and two perpendicular reflections (g_s, g_t). Transformations that preserve vector magnitudes and energy. While on fig. 1-right the same group describes the symmetries of vector spaces, representing the quadruped robot's state $\boldsymbol{x}$ and legs contact state $\boldsymbol{y}$.

Formally, a symmetry group is a set of invertible symmetry transformations (or actions) $\mathcal{G} = \{e, g_1, g_2, \dots\}$, containing the trivial action e (which leaves objects unchanged) and having a binary operator $(\cdot) : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$, that is associative (i.e. $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$), which composes group members into other group members, such as $g_r = g_s \cdot g_t$ for $\mathcal{K}_4$ (see fig. 1). Group representations are characterizations of how each action g transforms a specific geometric object, say $\boldsymbol{x} \in \mathbb{R}^n$. A representation $\rho_{\boldsymbol{x}} : \mathcal{G} \to GL(n)$ (GL : General Linear group) is a group homomorphism associating each g to an invertible linear map $\rho_{\boldsymbol{x}}(g) \in \mathbb{R}^{n \times n}$ specifying how the object $\boldsymbol{x}$ is transformed, that is: $g \cdot \boldsymbol{x} \doteq \rho_{\boldsymbol{x}}(g)\boldsymbol{x}$. Since group actions are abstract, it is common to define different object-dependent representations for each action, as we will see throughout this work.

[1]The link to an anonymous repository is available in the official comments on OpenReview, accessible to reviewers and area chairs. The repository will be made public to the general public upon paper acceptance.

Being this short section undoubtedly an unsatisfactory introduction to Group Theory, we refer the uninitiated and interested reader to Bronstein et al. (2021) for a remarkable introduction to the field.

3 LAGRANGIAN MECHANICS AND SYMMETRIES OF DYNAMICAL SYSTEMS

In this section, we first provide a group-theoretic perspective of symmetries in a dynamical system, which then allows us to introduce the discrete symmetries that are the objective of this work. To this end, let us consider a dynamical system with generalized coordinates $\boldsymbol{q} \in \mathrm{Q} \subset \mathbb{R}^{|\boldsymbol{q}|}$ and velocities $\dot{\boldsymbol{q}} \in \mathrm{TQ} \subset \mathbb{R}^{|\boldsymbol{q}|}$; as well as a Lagrangian scalar function $\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}})$, where Q and TQ represent the constrained configuration space of generalized position and velocity coordinates, respectively.

The symmetries of a dynamical system are defined as transformations in the space of generalized coordinates that keep the energy state of the system unchanged. I.e. if the system's Lagrangian is invariant to the transformations (Ostrowski & Burdick, 1996). In this work, we will focus on time-invariant point-transformations of generalized coordinates $\boldsymbol{q}_g = g(\boldsymbol{q})$, $\dot{\boldsymbol{q}}_g = \dot{g}(\boldsymbol{q}) = \frac{\partial g(\boldsymbol{q})}{\partial \boldsymbol{q}} \dot{\boldsymbol{q}}$. These transformations result in a valid set of new coordinates when the transformation $g(\cdot) : \mathrm{Q} \to \mathrm{Q}$ (and $\dot{g}(\cdot) : \mathrm{TQ} \to \mathrm{TQ}$) is a finite, invertible, continuous, and differentiable function of $\boldsymbol{q}$ (Lanczos, 2020). Formally, we say that a dynamical system has a symmetry group $\mathcal{G}$ when its Lagrangian is $\mathcal{G}$-invariant, that is:

$$\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \mathcal{L}(\boldsymbol{q}_g, \dot{\boldsymbol{q}}_g) = \mathcal{L}\left(g(\boldsymbol{q}), \left(\partial g(\boldsymbol{q}) / \partial \boldsymbol{q}\right) \dot{\boldsymbol{q}}\right) \quad \forall \quad g \in \mathcal{G},\ \boldsymbol{q} \in \mathrm{Q},\ \dot{\boldsymbol{q}} \in \mathrm{TQ}. \tag{1}$$

Because the Lagrangian structure differs among the transformed coordinates, when we derive the Equations of Motion (EoM) of the original and transformed system (say through the Euler-Lagrange equations of second order), these result in equivariant[2] EoMs (Lanczos, 2020). Formally, for a dynamical system with a symmetry group $\mathcal{G}$, we say that its dynamics are $\mathcal{G}$-equivariant, that is:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}})}{\partial \dot{\boldsymbol{q}}}\right) - \frac{\partial \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}})}{\partial \boldsymbol{q}} = \left(\frac{\partial g(\boldsymbol{q})}{\partial \boldsymbol{q}}\right)^{-1} \left[\frac{d}{dt}\left(\frac{\partial \mathcal{L}(\boldsymbol{q}_g, \dot{\boldsymbol{q}}_g)}{\partial \dot{\boldsymbol{q}}_g}\right) - \frac{\partial \mathcal{L}(\boldsymbol{q}_g, \dot{\boldsymbol{q}}_g)}{\partial \boldsymbol{q}_g}\right] \quad \forall \quad g \in \mathcal{G}. \tag{2}$$

Where the original and transformed dynamics are related linearly by the Jacobian of the coordinate transformation $\frac{\partial g(\boldsymbol{q})}{\partial \boldsymbol{q}} \in \mathbb{R}^{|\mathrm{Q}| \times |\mathrm{Q}|}$ (see the proof on Wheeler (2014) and Lanczos (2020)-V.2).

3.1 SYMMETRIES OF FLOATING-BASE ROBOTS

Let us now consider a floating-base robot, whose generalized coordinates $\boldsymbol{q} = [\boldsymbol{X}_B \ \hat{\boldsymbol{q}}]^\intercal$ are composed of the base configuration $\boldsymbol{X}_B \in \mathbb{SE}_3$ ($\mathbb{SE}_3$: Special Euclidean group in 3 dimensions) and the joint-space configuration $\hat{\boldsymbol{q}} \in \mathbb{R}^{n_j}$, where we deliberately abuse notation to keep the homogeneous matrix representation of position and orientation, instead of the usual vector quaternion.

With these coordinates, we can decouple any symmetry action transformation into: $\boldsymbol{q}_g = g(\boldsymbol{q}) = \rho_{\mathrm{Q}}(g)\, \boldsymbol{q} = \begin{bmatrix} \rho_{\mathbb{E}_3}(g) & \mathbf{0} \\ \mathbf{0} & \rho_{\hat{\boldsymbol{q}}}(g) \end{bmatrix} \boldsymbol{q} \quad | \quad \forall\ g \in \mathcal{G}$, where $\rho_{\mathrm{Q}}(g)$ is the representation of the left-action g acting on the configuration space Q. Similarly, $\rho_{\mathbb{E}_3}(g) \in \mathbb{E}_3$—a homogeneous matrix describing the base translation and rotation/reflection associated with g—acts on members of the Euclidean space , and $\rho_{\hat{\boldsymbol{q}}}(g) \in \mathbb{R}^{n_j \times n_j}$ acts on members of the joint-space.

Because we can translate and rotate the base of a floating-base robot without affecting its internal dynamics or modifying its energy state, it is not difficult to see that these systems have as symmetry group the Euclidean group and any of its subgroups $\mathcal{G} \subseteq \mathbb{E}_3$. We can verify this using the defined notation. Let us define the coordinate point-transformation as any rotation/reflection and translation (in the horizontal plane) of space, that leaves the joint-space configuration invariant (i.e. $\rho_{\hat{\boldsymbol{q}}}(g) = \boldsymbol{I}_{n_j}$). Then, reformulating eq. (1) we have:

$$\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \mathcal{L}(\boldsymbol{q}_g, \dot{\boldsymbol{q}}_g) = \mathcal{L}\left(\begin{bmatrix} \rho_{\mathbb{E}_3}(g) \boldsymbol{X}_B \\ \hat{\boldsymbol{q}} \end{bmatrix}, \begin{bmatrix} \rho_{\mathbb{E}_3}(g) \dot{\boldsymbol{X}}_B \\ \dot{\hat{\boldsymbol{q}}} \end{bmatrix}\right) \quad \forall \quad g \in \mathcal{G} \subseteq \mathbb{E}_3. \tag{3}$$

Furthermore, by defining the Lagrangian of the robot as $\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \mathcal{T}(\boldsymbol{q}, \dot{\boldsymbol{q}}) - \mathcal{U}(\boldsymbol{q})$, where $\mathcal{T}(\cdot)$ and $\mathcal{U}(\cdot)$ are the kinetic and potential energy of the system, and reformulating eq. (2) for our floating-base robot with the standard EoM (Wieber et al., 2016) we obtain the $\mathcal{G}$-equivariant dynamics as:

$$\mathbf{M}(\boldsymbol{q})\, \ddot{\boldsymbol{q}} + \mathbf{N}(\boldsymbol{q}, \dot{\boldsymbol{q}})\, \dot{\boldsymbol{q}} - \boldsymbol{\tau} = \rho_{\mathrm{Q}}(g)^{-1} \left[\mathbf{M}(\boldsymbol{q}_g)\, \ddot{\boldsymbol{q}}_g + \mathbf{N}(\boldsymbol{q}_g, \dot{\boldsymbol{q}}_g)\, \dot{\boldsymbol{q}}_g - \boldsymbol{\tau}_g\right] \quad \forall \quad g \in \mathcal{G}, \tag{4}$$

[2]Some authors refer to this property as *covariance* of the EoMs Wheeler (2014); Lanczos (2020).

where for a valid set of generalized coordinates $\boldsymbol{q}$, $\mathbf{M}(\boldsymbol{q})$ is the generalized mass matrix, $\mathbf{N}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ is the Coriolis-centripetal coupling matrix and $\boldsymbol{\tau}$ the generalized forces, encompassing gravitational effects, disturbances, and contacts. What eqs. (3) and (4) state is the notion that for floating-base robots, translations and rotations of its base will have no impact on the system's internal dynamics.

Reflections of Space and discrete symmetries of robotic systems: In this section, we stated that floating-base robots have as symmetry group the Euclidean group $\mathbb{E}_3 = \mathbb{SE}_3 \rtimes \mathcal{C}_2$ and any of its subgroups. However, most robotics literature considers only the effect of proper transformations, that is, actions of $\mathbb{SE}_3$, ignoring the Reflection Group $\mathcal{C}_2$, which involves improper transformations. This deliberative neglect of reflections of space is normally justified by the fact that, in general, no physical machine is capable of effecting a reflection on rigid bodies (Selig, 2005). Think of your own body as a floating-base robot, you can move through space and rotate your base (hip) but you are unable to execute a true reflection (forcing your heart to switch sides and certainly die). Fortunately, your body is symmetric w.r.t. the sagittal plane (we will assume perfect symmetry for now) which allows you to mimic the effects of any reflection of space with feasible changes in your configuration (i.e. changing your body pose and rotating and translating your base). Therefore, you have a discrete morphological symmetry.

4 Discrete morphological symmetries

A Discrete Morphological Symmetry (DMS) is a mathematical formalization of the intuitive property of dynamical systems that can mimic the effect of unfeasible reflections and rotations of space with feasible discrete changes in their configuration (see two examples in fig. 2). Formally, we say that a dynamical system with Lagrangian $\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}})$, has a DMS associated with a finite symmetry group $\mathcal{G}$ if for every state $(\boldsymbol{q}, \dot{\boldsymbol{q}})$, proper or improper action $\overline{g} \in \mathbb{E}_3$—equivalent to a true rotation/reflection and translation of space—, and transformed generalized coordinates $\overline{g}(\boldsymbol{q})$, $\dot{\overline{g}}(\dot{\boldsymbol{q}})$; there exists a proper, non-trivial action $g \in \mathcal{G}$ —involving no reflection of space and a transformation of the joint-space configuration— that mimics the effect of the true transformation of space $\overline{g}$, by reaching a viable state that is equivalent in energy. Using eq. (1) this translates to:

$$\mathcal{L}(\overline{g}(\boldsymbol{q}), \overline{g}(\dot{\boldsymbol{q}})) = \mathcal{L}(g(\boldsymbol{q}), g(\dot{\boldsymbol{q}})) \quad \forall \quad \boldsymbol{q} \in \mathrm{Q}, \dot{\boldsymbol{q}} \in \mathrm{TQ}, (g, \overline{g}) \mid g \in \mathcal{G}, \overline{g} \in \mathbb{E}_3, \tag{5}$$

where $(g(\boldsymbol{q}), g(\dot{\boldsymbol{q}}))$ is a viable state if it can be reached from any other point in the constrained configuration space through a trajectory of states that remain in Q and TQ (Del Prete, 2017). If we reformulate this equation for a specific configuration of a floating base robot, using section 3.1 notation, we would obtain:

$$\mathcal{L}\left(\begin{bmatrix}\rho_{\mathbb{E}_3}(\overline{g})\boldsymbol{X}_B \\ \hat{\boldsymbol{q}}\end{bmatrix}, \begin{bmatrix}\rho_{\mathbb{E}_3}(\overline{g})\dot{\boldsymbol{X}}_B \\ \dot{\hat{\boldsymbol{q}}}\end{bmatrix}\right) = \mathcal{L}\left(\begin{bmatrix}\rho_{\mathbb{E}_3}(g)\boldsymbol{X}_B \\ \rho_{\hat{q}}(g)\ \hat{\boldsymbol{q}}\end{bmatrix}, \begin{bmatrix}\rho_{\mathbb{E}_3}(g)\dot{\boldsymbol{X}}_B \\ \rho_{\hat{q}}(g)\ \dot{\hat{\boldsymbol{q}}}\end{bmatrix}\right), \quad \left| \begin{array}{rcl} \rho_{\mathbb{E}_3}(\overline{g}) & = & \pm 1 \\ \rho_{\mathbb{E}_3}(g) & = & 1 \end{array}\right. \tag{6}$$

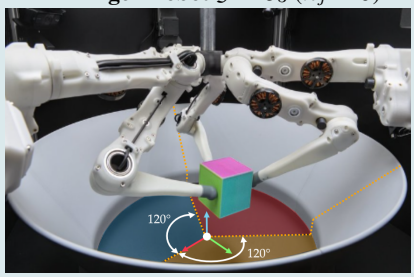
allowing to clearly recognize that the space transformation described by $\overline{g}$ acts only upon members of $\mathbb{E}_3$ (for floating-base robots the base configuration), leaving the joint-space configuration $\hat{\boldsymbol{q}}$ invariant, while the discrete equivalent action g has a proper action representation for members of $\mathbb{SE}_3$ and a non-trivial point-transformation $\rho_{\hat{q}}(g)$ acting on the joint space $\hat{\boldsymbol{q}}$. See fig. 1-left, focus on the robot's base orientation and color of legs. Notice how the robot imitates the reflection of space g_s with a transformation of its base and joints configuration, as with the other group actions.

Implications of morphological symmetries: The value of identifying and exploiting morphological symmetries comes from the $\mathcal{G}$-equivariant nature of the system dynamics (eqs. (2) and (4)). Equivariance implies that any controlled (or uncontrolled) trajectory of the evolution of the system state has up to $|\mathcal{G}|$ equivalent symmetric trajectories (being $|\mathcal{G}|$ the group order). Therefore, equivariance applies also to optimal control policies and any proprioceptive and exteroceptive measurements, relevant to the evolution of the system state (both in the Euclidean space—such as any vector of position, velocity, acceleration, forces, torques, RGB/RGB-Depth images—and in joint space—such as $\hat{\boldsymbol{q}}$, $\dot{\hat{\boldsymbol{q}}}$, $\ddot{\hat{\boldsymbol{q}}}$, torques, noise— (see fig. 1-left). These symmetries motivate the design of algorithms and function approximations that respect or approximate $\mathcal{G}$-equivariance (see section 5).

4.1 Properties of Robotic systems with morphological symmetries

By studying the implications of eq. (1) and eq. (2) for any dynamical system we can conclude several required morphological characteristics. Here, we will focus on floating-base robotic systems

Tri-Finger Robot $\mathcal{G} = \mathcal{C}_3$ ($n_j = 9$)

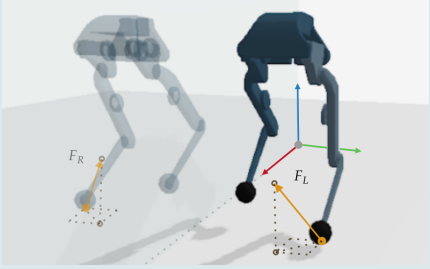


This fixed-based robot is symmetric w.r.t. rotations of space by $\theta = 120°$ in the vertical axis. Therefore, its symmetry group is the cyclic group of order three ($\mathcal{G} = \mathcal{C}_3$). Any action $\overline{g}_\theta$ involving a rotation of the base by an angle θ is *unfeasible*. However, *feasible* actions involving changes in configuration can imitate the effect of the group rotations. Let $\boldsymbol{q} = \hat{\boldsymbol{q}} = [\boldsymbol{q}_{f1}^\intercal, \boldsymbol{q}_{f2}^\intercal, \boldsymbol{q}_{f3}^\intercal]^\intercal$ be composed of each finger's DoF ($\boldsymbol{q}_{fi} \in \mathbb{R}^3$). Then, the feasible actions $g_\theta \in \mathcal{G}$ are given by permutations of the finger DoF:

$$g_{\frac{2\pi}{3}}(\hat{\boldsymbol{q}}) = \rho_{\hat{\boldsymbol{q}}}(g_{\frac{2\pi}{3}})\hat{\boldsymbol{q}}, \qquad g_{\frac{4\pi}{3}}(\hat{\boldsymbol{q}}) = \rho_{\hat{\boldsymbol{q}}}(g_{\frac{4\pi}{3}})\hat{\boldsymbol{q}},$$
$$\rho_{\hat{\boldsymbol{q}}}(g_{\frac{2\pi}{3}}) = \begin{bmatrix} \mathbf{0} & \boldsymbol{I}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I}_3 \\ \boldsymbol{I}_3 & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \rho_{\hat{\boldsymbol{q}}}(g_{\frac{4\pi}{3}}) = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \boldsymbol{I}_3 \\ \boldsymbol{I}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_3 & \mathbf{0} \end{bmatrix}.$$

To ensure dynamics equivariance (eq. (4)), the generalized forces involving fingers contact with world objects should also be transformed $g \cdot \boldsymbol{\tau} = \rho_{\hat{\boldsymbol{q}}}(g)\boldsymbol{\tau}$.

Bolt Bipedal Robot $\mathcal{G} = \mathcal{C}_2$ ($n_j = 6$)



Bolt is a bipedal robot with a sagittal plane reflection symmetry ($\mathcal{G} = \mathcal{C}_2$). This morphological symmetry allows it to imitate the effect of reflections of space ($\overline{g} \in \mathbb{E}_3$) by reconfiguring its base and legs. Defining $\boldsymbol{q} = [\boldsymbol{X}_B, \hat{\boldsymbol{q}}]^\intercal$, and its joint-space $\hat{\boldsymbol{q}} = [\boldsymbol{q}_L^\intercal, \boldsymbol{q}_R^\intercal]^\intercal$ as composed of the left L and right R legs' DoF ($\boldsymbol{q}_{L/R} \in \mathbb{R}^3$). The proper action $g \in \mathcal{G}$ imitating the improper reflection of space $\boldsymbol{R}_{\overline{g}}$ is described by:

$$g(\boldsymbol{q}) = \rho_{\mathrm{Q}}(g)\,\boldsymbol{q} = \begin{bmatrix} \rho_{\mathbb{E}_3}(g) & \mathbf{0} \\ \mathbf{0} & \rho_{\hat{\boldsymbol{q}}}(g) \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_B \\ \hat{\boldsymbol{q}} \end{bmatrix},$$
$$\rho_{\mathbb{E}_3}(g) = \begin{bmatrix} \boldsymbol{R}_g & \mathbf{r}_g \\ \mathbf{0} & 1 \end{bmatrix}, \qquad \rho_{\hat{\boldsymbol{q}}}(g) = \begin{bmatrix} \mathbf{0} & \boldsymbol{I}_3 \\ \boldsymbol{I}_3 & \mathbf{0} \end{bmatrix}.$$

Where $\rho_{\hat{\boldsymbol{q}}}(g)$ represents an exchange in legs configuration, and $\boldsymbol{R}_g$ and $\mathbf{r}_g$ are the rotation matrix and translation required to transform the robot base to coincide with its reflected version. $\boldsymbol{R}_g$ exists due to the **symmetry plane of the base body** (see section 4.1). To ensure dynamics equivariance (eq. (4)), the generalized forces are transformed by $g \cdot \boldsymbol{\tau} = \rho_{\mathrm{Q}}(g)\boldsymbol{\tau}$.

Figure 2: **Example morphological symmetries of the Tri-Finger (Funk et al., 2021) and Bolt robots**, allowing to imitate a rotation of space (left) and a reflection of space (right).

composed of n_b rigid bodies, but a similar analysis can be done for other robot morphologies (e.g. fix-based, soft robots). For clarity of the explanation, it is useful to consider two versions of the robot, an original version and a version rotated/reflected by $\boldsymbol{R}_{\overline{g}}$. The complete derivation of these properties is presented in supplementary A.

Kinematic Three Symmetries: In order for both robots to have the same energy (eq. (1)), the transformation $g(\boldsymbol{q})$ of the original robot state is required to result in a kinematic three indistinguishable from that of the rotated/reflected robot. For systems composed of rigid bodies, these kinematic constraints are summarized by equating the original robot bodies Jacobians with those of the rotated/reflected robot:

$$\mathbf{J}_{T_i}(g(\boldsymbol{q}))\frac{\partial g(\boldsymbol{q})}{\partial \boldsymbol{q}} = \boldsymbol{R}_{\overline{g}} \cdot \mathbf{J}_{T_j}(\boldsymbol{q}), \qquad \mathbf{J}_{R_i}(g(\boldsymbol{q}))\frac{\partial g(\boldsymbol{q})}{\partial \boldsymbol{q}} = |\boldsymbol{R}_{\overline{g}}|\boldsymbol{R}_{\overline{g}} \cdot \mathbf{J}_{R_j}(\boldsymbol{q}), \tag{7}$$

where $\mathbf{J}_{T_i}(\boldsymbol{q}),\ \mathbf{J}_{R_i}(\boldsymbol{q}) \in \mathbb{R}^{3\times|\mathrm{Q}|}$ are the position and orientation Jacobians (describing the instantaneous velocity vectors contributed by each Degree of Freedom (DoF) to body i) of the real robot at a configuration $\boldsymbol{q}$ (Wieber, 2006), and j is an index of bodies of the reflected/rotated robot.

Imitating reflections of space requires symmetric rigid bodies and/or modularity: The robots Solo fig. 1 and Bolt fig. 2 are able to imitate reflections of space due to their base bodies having a symmetric mass distribution (a symmetry plane). This property is fundamental for reflection symmetries because when applying a reflection to a rigid body, the reflected body's inertia $\overline{\mathbf{I}}$ differs from the original body inertia $\mathbf{I}$. Meaning that, in the general case, there does not exist a proper rotation g able to imitate a reflection $\overline{g}$, because even with kinematic symmetries, the difference in inertia results in different angular components of the bodies' kinetic energy, failing to comply with eq. (1). It turns out that, for rigid-body robots, a true reflection of space can only be imitated if: (1) each rigid body has a symmetric mass distribution, or (2) the kinematic three is modular, meaning that each non-symmetric rigid body, there exists a reflected version of the body within the robot's kinematic three (e.g., see how in fig. 2 the left leg's bodies of robot Bolt are reflections of its right leg's

bodies). Both cases ensure the existence of a body-specific reflection $\breve{\boldsymbol{R}}$ $s.t.$ $|\breve{\boldsymbol{R}}| = -1$, aligned with the principle axes of inertia of the body, ensuring the existence of the proper rotation g (see supplementary A.2 and supp.fig 5 for proof).

5 $\mathcal{G}$-EQUIVARIANT AND $\mathcal{G}$-INVARIANT FUNCTION APPROXIMATORS

As discussed in section 4, in data-driven applications involving a dynamical system with a morphological symmetry group $\mathcal{G}$, very often we encounter that the input $\boldsymbol{x}$ and output $\boldsymbol{y}$ spaces of functions we aim to approximate also share the finite symmetry group $\mathcal{G}$ (see section 6 for examples). In such situations, the sample efficiency and generalization capacity of a function approximator $f(\boldsymbol{x};\boldsymbol{\phi})$, parameterized with $\boldsymbol{\phi}$, is enhanced if it complies with the true equivariance or invariance constraints of the input-output spaces (Bronstein et al., 2017; Cohen & Welling, 2016), that is:

$$\underbrace{\rho_{out}(g)\cdot\boldsymbol{y} = f(\rho_{in}(g)\boldsymbol{x};\boldsymbol{\phi}) \quad | \quad \forall \quad g\in\mathcal{G}}_{Equivariance} \quad \text{or} \quad \underbrace{\boldsymbol{y} = f(\rho_{in}(g)\boldsymbol{x};\boldsymbol{\phi}) \quad | \quad \forall \quad g\in\mathcal{G}}_{Invariance}, \tag{8}$$

where equivariance implies that symmetries of the input are mapped to symmetries of the output, while invariance is the special case where symmetries of the input are mapped to an invariant output.

Consider $f(\boldsymbol{x};\boldsymbol{\phi})$ to be composed of multiple perceptron (or convolutional) layers of the form ${}^{l}\boldsymbol{y} := \sigma({}^{l}\boldsymbol{W}{}^{l}\boldsymbol{x} + {}^{l}\boldsymbol{b})$, where ${}^{l}\boldsymbol{x}\in\mathbb{R}^{n}$, ${}^{l}\boldsymbol{y}\in\mathbb{R}^{m}$, ${}^{l}\boldsymbol{W}\in\mathbb{R}^{m\times n}$ and ${}^{l}\boldsymbol{b}$ are the l layer's linear map and bias, respectively; and $\sigma:\mathbb{R}\to\mathbb{R}$ is a strictly monotonic nonlinearity (Ravanbakhsh et al., 2017). With this parametrization, the equivariance constraints of eq. (8) can be reduced to constraints on the linear map $\boldsymbol{W}$ (dropping the layer index l for notation clarity): [3]

$$\rho_{out}(g)\boldsymbol{W} = \boldsymbol{W}\rho_{in}(g) \mid \forall\, g\in\mathcal{G} \quad\Longleftrightarrow\quad (\rho_{\boldsymbol{W}}(g)-\boldsymbol{I})\boldsymbol{w} = \boldsymbol{0} \mid \forall\, g\in\mathcal{G}. \tag{9}$$

The RHS of eq. (9) is a reformulation of the equivariance constraints as a standard set of linear equations, defining $\rho_{\boldsymbol{W}}(g) = \rho_{out}(g)\otimes\rho_{in}(g^{-1})^{\intercal}\in\mathbb{R}^{mn\times mn}$ as the representation of the group action acting on the linear map ($\otimes$ stands for the Kronecker product) and $\boldsymbol{w} = vec(\boldsymbol{W})\in\mathbb{R}^{mn}$ as a vectorized version of $\boldsymbol{W}$ (refer to Finzi et al. (2021) for details). Since the constraint imposed by each g is linear in $\boldsymbol{W}$, we can stack them into a single system of linear equations $\boldsymbol{C}\boldsymbol{w} = \boldsymbol{0}$. The nullspace of this system of equations $\boldsymbol{B}\in\mathbb{R}^{mn\times r}$ describes the r basis vectors spawning the entire space of equivariant linear maps. Allowing to parameterize all $\mathcal{G}$-equivariant $\boldsymbol{W}$ as:

$$\boldsymbol{w} = \textstyle\sum_k^r c_k\boldsymbol{B}_{:,k} \quad\Longleftrightarrow\quad \boldsymbol{W} = \sum_k^r c_k\,\text{unvec}(\boldsymbol{B}_{:,k}) \doteq \sum_k^r c_k\mathbf{B}_{:,:,k}. \tag{10}$$

Where the basis coefficients $\boldsymbol{c}\in\mathbb{R}^{r}$ represent the free variables of the system of equations and the trainable parameters of the equivariant layer (see fig. 1 right).

Dealing with memory complexity of equivariant layers: An equivariant layer needs to store the matrices $\rho_{\boldsymbol{W}}(g)$ and $\boldsymbol{B}$, in addition to the typical memory complexity of a perceptron or convolutional layer. These matrices' memory complexity quickly becomes intractable for moderate input-output dimensions. Fortunately, finite symmetry groups have sparse action matrix representations, resulting in both of the aforementioned matrices being sparse. Our implementation[1] extends the Pytorch API from Finzi et al. (2021) to process finite groups with sparse matrix definitions limiting the additional memory footprint of equivariant layers to a minimum.

Dealing with the computational complexity of determining the equivariant basis $\boldsymbol{B}$: Computing the basis of equivariant linear operators is a computationally demanding process. Finzi et al. (2021) proposes a general approach handling both finite and Lie groups arbitrary action matrix representations, using a Krylov gradient-based method for computing $\boldsymbol{B}$, while Van der Pol et al. (2020) approximates $\boldsymbol{B}$ through SVD of a matrix $\bar{\boldsymbol{W}}\in\mathbb{R}^{z\times mn}$ ($z\geq mn$). Both approaches run in polynomial time $\mathcal{O}(r^2(mn)^2)$ and approximate r numerically.

Fortunately, morphological symmetry groups are finite and have regular representation matrices, allowing to compute $\boldsymbol{B}$ in linear time: Note that the constraints imposed by each $\rho_{\boldsymbol{W}}(g)$ result in parameter sharing constraints (e.g., $w_{10} = -w_2 = \ldots = w_0$). In these cases, every vector of the null-space of $\boldsymbol{C}$ (i.e., $\boldsymbol{B}_i$) simply describes the sharing scheme of a free variable of the system of equations (i.e., the trainable parameter c_i), and this sharing scheme is nothing else but

[3] A similar analysis can be made for the bias vector $\boldsymbol{b}$.

one of the unique r orbits of the dimensions of $\boldsymbol{w}$ when transformed by all group actions, e.g., $\mathcal{G} \cdot w_{10} = \{g \cdot w_{10} \ : \ \forall \ g \in \mathcal{G}\} = \{w_{10}, -w_2, \ldots, w_0\}$ (see the parameter orbits of length 4 in fig. 1-right, for $\mathcal{K}_4$). The orbits of all $w \in \boldsymbol{w}$ are trivially computed with $[\boldsymbol{w}, \rho_{\boldsymbol{W}}(g_1)\boldsymbol{w}, \ldots, \rho_{\boldsymbol{W}}(g_{|\mathcal{G}|})\boldsymbol{w}]$, while the unique r orbits can be identified in $\mathcal{O}(mn)$ time. Our proposed solution can be thought of as a linear-time version of Ravanbakhsh et al. (2017) algorithm for dense layers of finite groups.

Reduction of trainable parameters in equivariant layers: The equivariance constraints of eq. (10) on linear maps of perceptron (or convolutional) layers imply a reduction of trainable parameters from $|\boldsymbol{w}| = mn$ to $|\boldsymbol{c}| = r \leq mn$. For finite groups with regular action representation matrices, r is associated with the number of unique orbits of the elements of $\boldsymbol{w}$. Thus we can compute this value using the orbit-counting theorem (also known as Burnside's Lemma), which states that the number of orbits is the average number of fix-points of $\mathcal{G}$, that is $r = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} |\boldsymbol{w}^g|$, where $\boldsymbol{w}^g \doteq \{w \in \boldsymbol{w} : g \cdot w = w\}$ represents the set of elements of $\boldsymbol{w}$ that are invariant to g (i.e. fix-points). Those fix-points can be identified by the elements on the diagonal of $\rho_{\boldsymbol{W}}(g)$ that are equal to one. Therefore, for a $\mathcal{G}$-equivariant layer, the number of trainable parameters is determined by:

$$r = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \chi^{\mathbf{1}}_{\rho_{\boldsymbol{W}}}(g) = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \chi^{\mathbf{1}}_{\rho_{in}}(g^{-1}) \cdot \chi^{\mathbf{1}}_{\rho_{out}}(g), \tag{11}$$

denoting $\chi^{\mathbf{1}}_{\rho}(g) : \mathcal{G} \to \mathbb{N}$ as the number of fix-points of the action representation $\rho(g)$. Therefore, the number of trainable parameters can range from $|\boldsymbol{w}|/|\mathcal{G}| \leq r \leq |\boldsymbol{w}|$, depending in the fix-points of the layers's input and output spaces.

Optimal trainable parameter initialization for equivariant layers: Proper initialization of the equivariant layer's trainable parameters $\boldsymbol{c}_l$ (eq. (10)) is required to avoid activations/gradients from vanishing or exploding (Klambauer et al., 2017). Following the same derivation of the Kaiming initialization (He et al., 2015) (see supplementary B and supp.fig 7), we can conclude that the parameters should be initially sampled from a distribution with $\mathrm{Var}(\boldsymbol{c}_l) = m/\lambda_{\mathbf{B}}\gamma_\sigma$, to ensure constant variance of activations throughout the network layers. Where $\lambda_{\mathbf{B}} = \sum_i^m \sum_j^n \sum_k^r \mathbf{B}^2_{i:j:k}$ and γ_σ is a nonlinearity dependant scalar (e.g., $\gamma_{\text{ReLu}} = 1/2$, $\gamma_{\text{SeLu}} = 1$ following Klambauer et al. (2017)). This initialization depends only on $\mathbf{B}$. Thus, is applicable to any group for which $\mathbf{B}$ is computable.

6 EXPERIMENTS

We present two experiments of supervised learning, a regression application using synthetic data and a classification application using real-world data. Both experiments aim to illustrate the impact on the model's sample efficiency and generalization capacity when exploiting the robot morphological symmetries, either by using equivariant versions of the models or data augmentation. While we keep the presentation concise, all the technical aspects are detailed in supplementary C and [1].

CoM Momentum Estimation (Regression): In this synthetic experiment we train a function approximator $f(\hat{\boldsymbol{q}}, \hat{\dot{\boldsymbol{q}}}; \boldsymbol{\phi})$ to estimate the robot's center-of-mass momentum $\boldsymbol{h} = [\boldsymbol{l}^\intercal \ \mathbf{k}^\intercal]^\intercal$, given the joint-space position and velocities $\hat{\boldsymbol{q}}$, $\hat{\dot{\boldsymbol{q}}}$, where $\boldsymbol{l}$ and $\mathbf{k}$ are the linear and angular momentum components (referenced to a frame located at the CoM and with the orientation of the base frame, see fig. 1-left). The dataset is generated using $\boldsymbol{h} = \mathbf{A}_G(\hat{\boldsymbol{q}})\hat{\dot{\boldsymbol{q}}}$, where $\mathbf{A}_G$ is the Centroidal Momentum Matrix (CMM) of Orin et al. (2013). This analytical function is highly non-linear and equivariant to the morphological symmetry group $\mathcal{G}$ of a robot: $g \cdot \boldsymbol{h} = \mathbf{A}_G(g \cdot \hat{\boldsymbol{q}})g \cdot \hat{\dot{\boldsymbol{q}}} \mid \forall g \in \mathcal{G}$. Consequently, the function approximator is expected to be equivariant or to approximate equivariance.

We test two robots: (1) Atlas, a $n_j = 30$[DoF] bipedal humanoid robot that has a sagittal plane reflection symmetry. Therefore, its morphological symmetry group is the reflection group $\mathcal{G} = \mathcal{C}_2$. (2) Solo, a $n_j = 12$[DoF] quadruped robot that is symmetric w.r.t. its sagittal and traversal planes. Therefore, its morphological symmetry group is the Klein-4 group $\mathcal{K}_4$ (see fig. 1-left). At the same time, we compare three variants of a function approximation: a standard Multi-Layer Perceptron (MLP), a version of the MLP trained using data-augmentation (MLP-aug), and a version of the MLP with hard-equivariance constraints (E-MLP).

On fig. 3-left-&-middle, we compare the performance of MLP, MLP-aug and E-MLP models with different capacities (number of neurons in hidden layers). In the case of Atlas and Solo, and for all model capacities, the E-MLP and MLP-Aug outperform MLP on sample efficiency (better generalization with less data) and robustness to overfitting when training data is scarce. Comparing

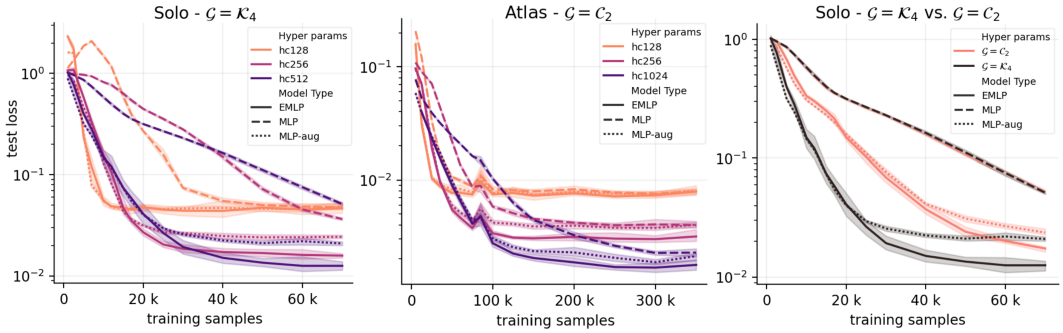


Figure 3: **CoM-estimation results comparing MLP, MLP-aug, and EMLP models**. **Left and Middle:** Test set sample efficiency, for robot Solo and Atlas, of model variants with different capacities (number of hidden layers' neurons *hc*). **Right:** Sample efficiency for robot Solo (fig. 1), of models with $hc = 512$, when exploiting $\mathcal{G} = \mathcal{K}_4$ (sagittal and traversal symmetries) and $\mathcal{G} = \mathcal{C}_2 = \{e, g_s\} \subset \mathcal{K}_4$ (only sagittal symmetry). Reported values represent the average and standard deviation across 10 different seeds.

the E-MLP and MLP-Aug model variants, we see that the lower capacity versions behave similarly, but as capacity increases, E-MLP starts to show better sample efficiency and generalization. Lastly, on (fig. 3-right) we compare, for the robot Solo, the performance of the three model variants when exploiting the robots' entire symmetry group ($\mathcal{K}_4$) and a subgroup of the real symmetry group ($\mathcal{C}_2 \subset \mathcal{K}_4$). The results clearly indicate that sample efficiency and generalization capacity increase with the number of symmetries of the data exploited.

Static-Friction-Regime Contact Detection (Classification): This experiment uses the dataset presented in Lin et al. (2021) for the estimation of static-friction-regime contacts in each of the four legs of the Mini-Cheetah quadruped robot. The dataset samples, collected in the real-world, consist of the history of the past 150 time-frames of proprioceptive data collected from inboard sensors of the robot during locomotion with various gaits and over several terrains, $\boldsymbol{x} \in \mathbb{R}^{54\times150}$, and the ground truth contact state of the robot $\boldsymbol{y} \in \mathbb{R}^{16}$, estimated off-line using a non-causal algorithm (i.e., dependant on the past and future). The goal of the experiment is to train a function approximator $f(\boldsymbol{x};\boldsymbol{\phi})$, that is causal, for estimating the contact state.

Mini-Cheetah has a sagittal plane reflection symmetry, thus, its morphological group is the reflection group $\mathcal{G} = \mathcal{C}_2$. Hence, both the proprioceptive data $\boldsymbol{x}$ and the contact state $\boldsymbol{y}$ are equivariant to their reflected versions. We compare three variants of function approximators: the original NN architecture of Lin et al. (2021) composed of 1D convolutional and perceptron layers (CNN), a version of CNN trained using data-augmentation (CNN-aug), and a version of CNN with hard-equivariance constraints (E-CNN).

The results comparing the sample efficiency of the model variants and the average legs contact state classification are illustrated in fig. 4-left-&-middle. Where the equivariant model E-CNN presents better generalization, performance, and robustness to dataset biases (see supp.fig 6) than the unconstrained models across all training set sizes, followed by CNN-aug. In fig. 4-right we evaluate classification metrics, on the test set, of the model variants when trained with all training data. The E-CNN model outperforms both CNN-aug and CNN on contact state classification and average leg contact detection. When comparing individual leg classification we see that the equivariant model converges to having a similar performance for each symmetric pair of legs, while the unconstrained models converge to an asymmetrical suboptimal state favoring the contact detection of one leg at the expense of reduced performance for the symmetric leg (see LF and RF f1-scores). This asymmetrical performance is attributed to the CNN and CNN-aug models learning to extract temporal features for both symmetric legs separately, increasing the likelihood of converging to asymmetrical local minima. On the contrary, the equivariant model E-CNN can be thought of as learning to extract a single set of *symmetric* temporal features for each symmetric pair of states (a consequence of the model's equivariance and parameter sharing). This implies that the temporal features used for determining the contact state of, say the left frontal leg, would also be used to determine the contact state of the symmetric leg, the right frontal leg, and vice-versa.

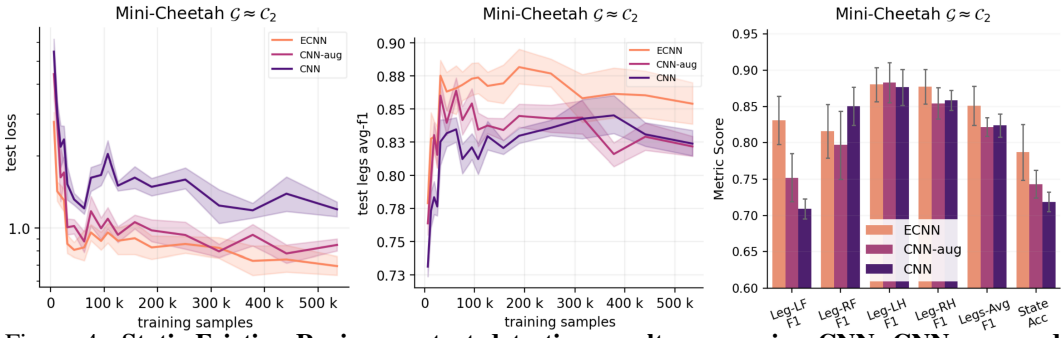


Figure 4: **Static-Friction-Regime contact detection results comparing CNN, CNN-aug, and ECNN. Left:** Sample efficiency in log-log scale. **Middle:** Average legs F1-score vs. training samples. **Right:** Classification metrics on test set performance of models trained with the entire training set. Selected metrics are contact state ($\boldsymbol{y} \in \mathbb{R}^{16}$) accuracy (Acc) and f1-score (F1) for each leg binary contact state. Due to the sagittal symmetry of the robot the left front (LF) and right front (RF) legs are expected to be symmetric, as the left hind (LH) and right hind (RH) legs. F1-score is presented considering the dataset class imbalance (see supplementary C.4 and supp.fig 6). Reported values represent the average and standard deviation across 8 different seeds.

Note that the real-world robot is only approximately symmetric, considering tolerance in the robot's kinematic and dynamic parameters and the external suite of sensors mounted on the robot's back during data collection. Despite the approximate symmetry, the empirical results of the experiment indicate that it is desirable to exploit the morphological symmetry group of the robot either with data-augmentation or equivariance constraints. This fact hints at the reduced capacity of unconstrained NN models to learn symmetries from raw data. Evidence shows that the local minima to which the equivariant models converge have, on average, better performance. Motivating the use of these techniques for data gathered on simulation (where symmetries are exact) and for data gathered in the real world, where the assumption of exact symmetries is broken.

7 Conclusions & Discussion

In this work, we study and characterize discrete morphological symmetries of dynamical systems living in the Euclidean space. We conclude that bilateral and radial symmetries can be associated with the capability of the system to imitate true reflections and rotations of space with discrete changes in the system configuration. By studying these symmetries with the language of Group Theory, we identify a finite symmetry group $\mathcal{G}$, specific to the dynamical system, and show that the system dynamics are $\mathcal{G}$-equivariant and that the system's proprioceptive and exteroceptive data have the same symmetry group $\mathcal{G}$.

Having made the connection between Dynamical Systems and Group Theory, we show why and how these morphological symmetries should be exploited in data-driven applications—to obtain improvements in sample efficiency and generalization capacity—, either by using data-augmentation or $\mathcal{G}$-equivariant Neural Networks. For the latter, we present contributions addressing the memory and computational complexity drawbacks of using this type of architecture. Furthermore, we characterize the parameter-sharing nature of equivariant layers and relate it to the duplication/sharing of dynamics (mass & inertia) and kinematics (Denavit–Hartenberg) parameters of robotic systems with duplicated body parts and morphological symmetries.

Lastly, we present empirical results supporting our claims on two data-driven applications using synthetic and real-world data of three different robots. In both experiments improvements in sample efficiency and generalization are obtained by exploiting the morphological symmetries bias, motivating the use of this technique in applications using data from simulation and/or the real world.

Limitations: Our work makes three main assumptions: (1) Symmetries are identifiable (intuitively or analytically), thus we need not learn them (Weissenbacher et al., 2022). (2) Symmetries are exact, ignoring the possibility of approximate symmetries (Wang et al., 2022). We leave the analysis of this case for future work. (3) Symmetries have regular matrix representations. Although true for finite groups of morphological symmetries, for more general representations the equivariant linear map basis can be computed using Finzi et al. (2021), with an increase in computational cost.

REPRODUCIBILITY STATEMENT

The experimental setup used in the experiments is described in section 6 and supplementary C. Moreover, our implementation[1] will be open-access, where any interested party can find: (1) The original scripts used to run the experiments and generate the results, (2) the parameters of the models used for comparison in the experiments, avoiding the need to retrain the models to test the results, (3) the datasets used in each of our experiments, including the custom partitioning of the dataset of Lin et al. (2021), (4) the scripts used to summarize the results into the figures used in this paper, and (5) 3D interactive environments that allow for the visualization of the morphological symmetries, one of this environments was used for capturing fig. 1-left and its 3D animation.

REFERENCES

Farzad Abdolhosseini, Hung Yu Ling, Zhaoming Xie, Xue Bin Peng, and Michiel Van de Panne. On learning symmetric locomotion. In *Motion, Interaction and Games*, pp. 1–10. 2019.

Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*, 2021.

Justin Carpentier, Guilhem Saurel, Gabriele Buondonno, Joseph Mirabel, Florent Lamiraux, Olivier Stasse, and Nicolas Mansard. The pinocchio c++ library: A fast and flexible implementation of rigid body dynamics algorithms and their analytical derivatives. In *2019 IEEE/SICE International Symposium on System Integration (SII)*, pp. 614–619. IEEE, 2019.

Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International conference on machine learning*, pp. 2990–2999. PMLR, 2016.

Andrea Del Prete. Joint position and velocity bounds in discrete-time acceleration/torque control of robot manipulators. *IEEE Robotics and Automation Letters*, 3(1):281–288, 2017.

Marc Finzi, Max Welling, and Andrew Gordon Wilson. A practical method for constructing equivariant multilayer perceptrons for arbitrary matrix groups. In *International Conference on Machine Learning*, pp. 3318–3328. PMLR, 2021.

Niklas Funk, Charles Schaff, Rishabh Madan, Takuma Yoneda, Julen Urain De Jesus, Joe Watson, Ethan K Gordon, Felix Widmaier, Stefan Bauer, Siddhartha S Srinivasa, et al. Benchmarking structured policies and policy optimization for real-world dexterous object manipulation. *IEEE Robotics and Automation Letters*, 7(1):478–485, 2021.

Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pp. 249–256. JMLR Workshop and Conference Proceedings, 2010.

Kaveh Akbari Hamed and Jessy W Grizzle. Event-based stabilization of periodic orbits for underactuated 3-d bipedal robots with left-right symmetry. *IEEE Transactions on Robotics*, 30(2): 365–381, 2013.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034, 2015.

Gábor Holló. Demystification of animal symmetry: Symmetry is a response to mechanical forces. *Biology Direct*, 12(1):1–18, 2017.

Günter Klambauer, Thomas Unterthiner, Andreas Mayr, and Sepp Hochreiter. Self-normalizing neural networks. *Advances in neural information processing systems*, 30, 2017.

Cornelius Lanczos. *The variational principles of mechanics*. University of Toronto press, 2020.

Tzu-Yuan Lin, Ray Zhang, Justin Yu, and Maani Ghaffari. Legged robot state estimation using invariant kalman filtering and learned contact events. In *5th Annual Conference on Robot Learning*, 2021.

Daniel Ordonez-Apraez, Antonio Agudo, Francesc Moreno-Noguer, and Mario Martin. An adaptable approach to learn realistic legged locomotion without examples. In *2022 International Conference on Robotics and Automation (ICRA)*, pp. 4671–4678. IEEE, 2022.

David E Orin, Ambarish Goswami, and Sung-Hee Lee. Centroidal dynamics of a humanoid robot. *Autonomous robots*, 35(2):161–176, 2013.

Jim Ostrowski and Joel Burdick. Geometric perspectives on the mechanics and control of robotic locomotion. In *Robotics Research*, pp. 536–547. Springer, 1996.

Robert J Quigley. Pseudovectors and reflections. *American Journal of Physics*, 41(3):428–430, 1973.

Siamak Ravanbakhsh, Jeff Schneider, and Barnabas Poczos. Equivariance through parameter-sharing. In *International conference on machine learning*, pp. 2892–2901. PMLR, 2017.

Jon M Selig. *Geometric fundamentals of robotics*, volume 128. Springer, 2005.

Elise Van der Pol, Daniel Worrall, Herke van Hoof, Frans Oliehoek, and Max Welling. Mdp homomorphic networks: Group symmetries in reinforcement learning. *Advances in Neural Information Processing Systems*, 33:4199–4210, 2020.

Rui Wang, Robin Walters, and Rose Yu. Approximately equivariant networks for imperfectly symmetric dynamics. *arXiv preprint arXiv:2201.11969*, 2022.

Matthias Weissenbacher, Samarth Sinha, Animesh Garg, and Kawahara Yoshinobu. Koopman q-learning: Offline reinforcement learning via symmetries of dynamics. In *International Conference on Machine Learning*, pp. 23645–23667. PMLR, 2022.

James T. Wheeler. General coordinate covariance of the euler lagrange equations. Classical Mechanics class notes, 2014. URL `http://www.physics.usu.edu/Wheeler/ClassicalMechanics/CMCoordinateinvarianceofEulerLagrange.pdf`.

P-B Wieber. Holonomy and nonholonomy in the dynamics of articulated motion. In *Fast motions in biomechanics and robotics*, pp. 411–425. Springer, 2006.

Pierre-Brice Wieber, Russ Tedrake, and Scott Kuindersma. Modeling and control of legged robots. In *Springer handbook of robotics*, pp. 1203–1234. Springer, 2016.

Raymond Yeh, Yuan-Ting Hu, and Alexander Schwing. Chirality nets for human pose regression. *Advances in Neural Information Processing Systems*, 32, 2019.

Wenhao Yu, Greg Turk, and C Karen Liu. Learning symmetric and low-energy locomotion. *ACM Transactions on Graphics (TOG)*, 37(4):1–12, 2018.

Martin Zinkevich and Tucker Balch. Symmetry in markov decision processes and its implications for single agent and multi agent learning. In *In Proceedings of the 18th International Conference on Machine Learning*. Citeseer, 2001.

SUPPLEMENTARY
ON DISCRETE SYMMETRIES OF ROBOTICS SYSTEMS: A GROUP-THEORETIC AND DATA-DRIVEN ANALYSIS

A PROPERTIES OF ROBOTIC SYSTEMS WITH MORPHOLOGICAL SYMMETRIES

For clarity of the explanation, let us imagine two different Euclidean spaces and two versions of the robot: the original space (with reference frame o) and robot with coordinates $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$, and the virtual rotated/reflected space (with a reference frame $\overline{o}$, with configuration ${}^{o}\boldsymbol{X}_{\overline{o}} = \begin{bmatrix} \boldsymbol{R}_{\overline{g}} & \mathbf{r}_{\overline{o}} \\ \mathbf{0} & 1 \end{bmatrix}$) and virtual robot with coordinates $\boldsymbol{q}_{\overline{g}} = \overline{g}(\boldsymbol{q})$ and $\dot{\boldsymbol{q}}_{\overline{g}} = \dot{\overline{g}}(\boldsymbol{q})$ referenced to $\overline{o}$. Noting that in the case of a reflection, the virtual robot has reflected versions of each rigid body.

For eqs. (4) and (5) to hold, there must exist an action $g \in \mathcal{G}$ transforming the real robot configuration $\boldsymbol{q}_g = g(\boldsymbol{q})$ and $\boldsymbol{q}_g = \dot{g}(\boldsymbol{q})$ resulting in the same kinetic energy as the virtual robot's kinetic energy:

$$\mathcal{T}(\boldsymbol{q}_g, \dot{\boldsymbol{q}}_g) = \tfrac{1}{2}\sum_{i=1}^{n_b} m_i \dot{\mathbf{r}}_{g,i}^{\intercal}\dot{\mathbf{r}}_{g,i} + \boldsymbol{w}_{g,i}^{\intercal}\mathbf{I}_i\boldsymbol{w}_{g,i} \doteq \tfrac{1}{2}\sum_{j=1}^{n_b} \overline{m}_j \dot{\overline{\mathbf{r}}}_j^{\intercal}\dot{\overline{\mathbf{r}}}_j + \overline{\boldsymbol{w}}_j^{\intercal}\overline{\mathbf{I}}_j\overline{\boldsymbol{w}}_j = \mathcal{T}(\boldsymbol{q}_{\overline{g}}, \dot{\boldsymbol{q}}_{\overline{g}}), \tag{12}$$

where $\dot{\mathbf{r}}_{g,i}$, $\boldsymbol{w}_{g,i}$, m_i and $\mathbf{I}_i$ are the linear and angular velocity, mass and inertia matrix of the transformed body i (referenced to o). Likewise, $\dot{\overline{\mathbf{r}}}_i$, $\overline{\boldsymbol{w}}_i$, $\overline{m}_i$ and $\overline{\mathbf{I}}_i$ are the equivalent quantities for the virtual robot and body i (referenced to $\overline{o}$).

A.1 KINEMATIC SYMMETRIES:

Ignore momentarily the influence of the mass and inertia in terms of the real and virtual bodies. We can assert that for supp.eq 12 to hold, the transformed configuration should result in a kinematic three indistinguishable from the virtual robot's. Thus, for everybody i in the real robot kinematic three there should exist an equivalent virtual body j (as seen in fig. 2, not always $j = i$). By equating the linear and angular velocities of the real and virtual bodies, referenced to o, and expressing the velocities as functions of the generalized coordinates we obtain:

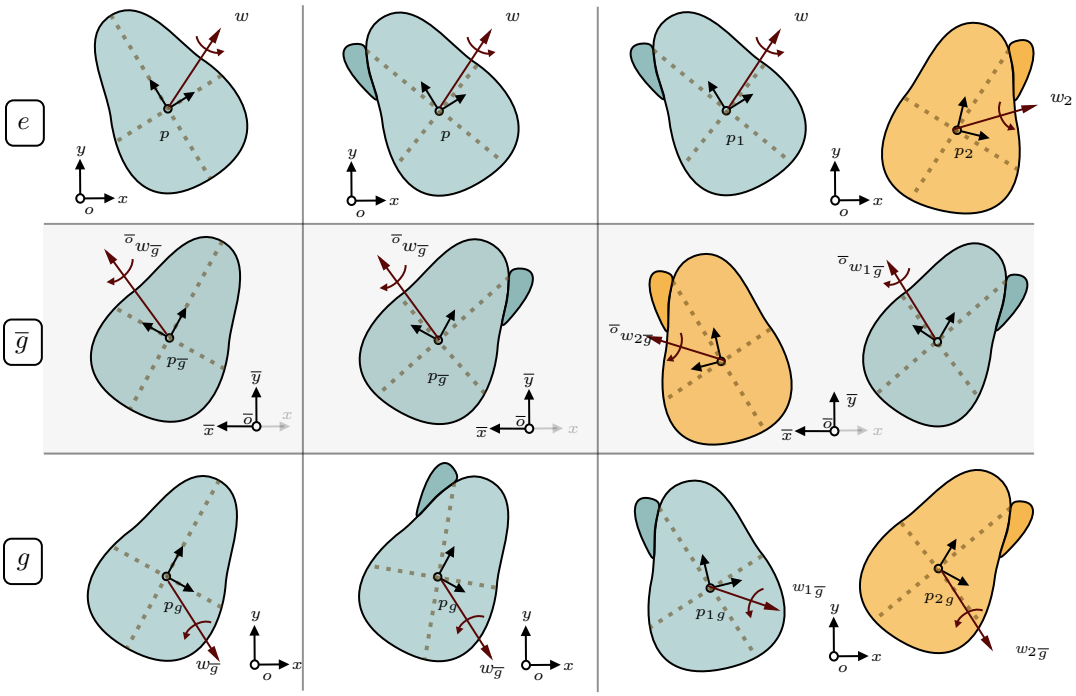
$$\begin{aligned} \dot{\mathbf{r}}_{g,i} &= \dot{\overline{\mathbf{r}}}_j \doteq \boldsymbol{R}_{\overline{g}} \cdot \dot{\mathbf{r}}_j & \boldsymbol{w}_{g,i} &= \overline{\boldsymbol{w}}_j \doteq |\boldsymbol{R}_{\overline{g}}|\boldsymbol{R}_{\overline{g}} \cdot \boldsymbol{w}_j \\ \mathbf{J}_{T_i}(\boldsymbol{q}_g)\dot{\boldsymbol{q}}_g &= \boldsymbol{R}_{\overline{g}} \cdot \mathbf{J}_{T_j}(\boldsymbol{q})\dot{\boldsymbol{q}} & \mathbf{J}_{R_i}(\boldsymbol{q}_g)\dot{\boldsymbol{q}}_g &= |\boldsymbol{R}_{\overline{g}}|\boldsymbol{R}_{\overline{g}} \cdot \mathbf{J}_{R_j}(\boldsymbol{q})\dot{\boldsymbol{q}}, \\ \mathbf{J}_{T_i}(g(\boldsymbol{q}))\tfrac{\partial g(\boldsymbol{q})}{\partial \boldsymbol{q}} &= \boldsymbol{R}_{\overline{g}} \cdot \mathbf{J}_{T_j}(\boldsymbol{q}) & \mathbf{J}_{R_i}(g(\boldsymbol{q}))\tfrac{\partial g(\boldsymbol{q})}{\partial \boldsymbol{q}} &= |\boldsymbol{R}_{\overline{g}}|\boldsymbol{R}_{\overline{g}} \cdot \mathbf{J}_{R_j}(\boldsymbol{q}), \end{aligned} \tag{13}$$

where $\mathbf{J}_{T_i}(\boldsymbol{q}),\ \mathbf{J}_{R_i}(\boldsymbol{q}) \in \mathbb{R}^{3\times|\mathrm{Q}|}$ are the position and orientation analytical Jacobians (describing the instantaneous velocity vectors contributed by each DoF to body i) of the real robot at a configuration $\boldsymbol{q}$ (Wieber, 2006). Formulating supp.eq 13 for each of the n_j bodies of the robot we obtain at best $n_j \times 3 \times |\mathrm{Q}|$ non-linear equations that can be used to assert if g exists. In practice if we define the action as $g(\boldsymbol{q}) = \rho_{\mathrm{Q}}(g)\boldsymbol{q}$ (as in section 3.1), the action representation $\rho_{\mathrm{Q}}(g)$ and specially its component acting on the joint space $\rho_{\dot{\mathbf{q}}}(g)$ can be trivially determined by solving eq. (7) for each body from top to bottom of the kinematic three (i.e., base first, end-effectors last), if g exists.

A.2 REFLECTIONS REQUIRE MODULARITY OR SYMMETRIC RIGID BODIES

Let us assume kinematic symmetry and direct our attention now to the influence of the mass and inertia terms on the kinetic energy of a single rigid body when it is transformed with the action g, which imitates a true reflection of space $\overline{g}$. Focus on the first two columns of supp.fig 5. Because of the kinematic symmetry the CoM of the reflected and transformed bodies coincide, both bodies have equivalent linear components of kinetic energy. However, for arbitrary rigid bodies, the reflected and transform bodies will have different angular components of kinetic energy. Note that in the general case, the transformed and reflected bodies' inertia will differ, thus even if both bodies have the same angular velocities, their kinetic energy will differ.

Let p, $p_{\overline{g}}$ and p_g be frames located at the CoM of the original, reflected and transformed bodies, aligned with the principal axes of inertia of each of the bodies. Similarly, denote ${}^{o}\mathbf{I}$ and ${}^{o}\mathbf{I}_g$ as the original and transformed bodies inertias referenced to o, and ${}^{\overline{o}}\overline{\mathbf{I}}_{\overline{g}}$ as the reflected body inertia



Supplementary Figure 5: Properties of bodies capable of imitating a true reflection $\overline{g}$ of space (w.r.t the yz-plane in this case), with a proper transformation g involving only a rotation and translation. The first row shows the original bodies with their respective angular velocities $\boldsymbol{w}$, subjected to trivial symmetry transformation e (dashed lines represent the principle axes of inertia of the bodies), the second and third rows display the effect of $\overline{g}$ and g on the bodies and angular velocities, respectively. The first column displays a rigid body with symmetric mass distribution, for which g exists, as the reflected and rotated bodies share an equivalent angular kinetic energy. The second column shows a rigid body with asymmetrical mass distribution, for which the rotation g, that produces a kinematic symmetry, results in the reflected and rotated bodies having different angular kinetic energies (eq. (1)). The third column shows two bodies with asymmetrical mass distributions, each a reflected version of the other, in this case, the action g swaps bodies configurations to imitate the configuration and energy state of the reflected bodies transformed with $\overline{g}$. Angular velocity is a pseudo-vector (or axial-vector), for which a reflection transformation is computed as $\boldsymbol{w}_{\overline{g}} = |\boldsymbol{R}_{\overline{g}}|\boldsymbol{R}_{\overline{g}}\boldsymbol{w}$ (see Quigley (1973)).

referenced to the reflected Euclidean space $\overline{o}$. In order to comply with supp.eq 13, we must ensure that:

$$
\begin{aligned}
{}^{o}\boldsymbol{w}_{\overline{g}}^{\intercal}\,{}^{o}\mathbf{I}_{g}\,{}^{o}\boldsymbol{w}_{\overline{g}} &= {}^{\overline{o}}\boldsymbol{w}_{\overline{g}}^{\intercal}\,{}^{\overline{o}}\overline{\mathbf{I}}_{\overline{g}}\,{}^{\overline{o}}\boldsymbol{w}_{\overline{g}},, \\
(\boldsymbol{R}_{\overline{g}}\,{}^{o}\boldsymbol{w})^{\intercal}\,{}^{o}\mathbf{I}_{g}\,(\boldsymbol{R}_{\overline{g}}\,{}^{o}\boldsymbol{w}) &= {}^{\overline{o}}\boldsymbol{w}_{\overline{g}}^{\intercal}\,{}^{\overline{o}}\overline{\mathbf{I}}_{\overline{g}}\,{}^{\overline{o}}\boldsymbol{w}_{\overline{g}} \quad \mid {}^{o}\boldsymbol{w}_{\overline{g}} = |\boldsymbol{R}_{\overline{g}}|\boldsymbol{R}_{\overline{g}}\,{}^{o}\boldsymbol{w}, \\
{}^{o}\mathbf{I}_{g} &= \boldsymbol{R}_{\overline{g}}\,{}^{\overline{o}}\overline{\mathbf{I}}_{\overline{g}}\boldsymbol{R}_{\overline{g}} \quad \mid {}^{o}\boldsymbol{w} \equiv {}^{\overline{o}}\boldsymbol{w}_{\overline{g}}, \quad \boldsymbol{R}_{\overline{g}}\boldsymbol{R}_{\overline{g}} = \boldsymbol{I}, \\
{}^{o}\boldsymbol{R}^{p_g}\,{}^{p_g}\mathbf{I}_{g}\,{}^{o}\boldsymbol{R}^{p_g\,\intercal} &= \boldsymbol{R}_{\overline{g}}\,{}^{\overline{o}}\boldsymbol{R}^{p_{\overline{g}}}\,{}^{p_{\overline{g}}}\overline{\mathbf{I}}_{\overline{g}}\,{}^{\overline{o}}\boldsymbol{R}^{p_{\overline{g}}\,\intercal}\boldsymbol{R}_{\overline{g}} \quad \mid {}^{a}\mathbf{I} = {}^{b}\boldsymbol{R}^{a}\,{}^{b}\mathbf{I}\,{}^{b}\boldsymbol{R}^{a\,\intercal}, \\
{}^{o}\boldsymbol{R}^{p_g}\,{}^{p_g}\mathbf{I}_{g}\,{}^{o}\boldsymbol{R}^{p_g\,\intercal} &= \boldsymbol{R}_{\overline{g}}\,{}^{o}\boldsymbol{R}^{p}\,{}^{p}\boldsymbol{R}^{p_{\overline{g}}}\,{}^{p_{\overline{g}}}\overline{\mathbf{I}}_{\overline{g}}\,{}^{p}\boldsymbol{R}^{p_{\overline{g}}\,\intercal}\,{}^{\overline{o}}\boldsymbol{R}^{p_{\overline{g}}\,\intercal}\boldsymbol{R}_{\overline{g}}, \\
{}^{o}\boldsymbol{R}^{p_g} &\doteq \boldsymbol{R}_{\overline{g}}\,{}^{o}\boldsymbol{R}^{p}\,{}^{p}\boldsymbol{R}^{p_{\overline{g}}} \quad \mid \quad {}^{p}\mathbf{I} \equiv {}^{p_g}\mathbf{I}_{g} \equiv {}^{p_{\overline{g}}}\overline{\mathbf{I}}_{\overline{g}}.
\end{aligned}
\tag{14a}
$$

What eq. (14a) states is that in order for the reflected and transformed bodies to have equivalent angular kinetic energy, both bodies should have co-linear (or aligned) principal axes of inertia. This allow us to describe ${}^{o}\boldsymbol{R}^{p_g}$ as a function of the original body configuration ${}^{o}\boldsymbol{R}^{p}$ and two reflection matrices: the true reflection of space $\boldsymbol{R}_{\overline{g}}$ and a body specific diagonal reflection matrix ${}^{p}\boldsymbol{R}^{p_{\overline{g}}}$, which exists only if the rigid body has a symmetric mass distribution. A visual example for symmetric and asymmetrical rigid bodies is presented in supp.fig 5 left and middle columns.

Note that in this work we use the convention of left actions, implying that $\rho_{\mathbb{E}_3}(g) = \begin{bmatrix} \boldsymbol{R}_g & \mathbf{r}_g \\ \mathbf{0} & 1 \end{bmatrix}$ is a matrix that premultiplies the robot base configuration $\boldsymbol{X}_B$. However, as illustrated in eq. (14a) there is a left and right action affecting $\boldsymbol{X}_B$. This left-right action is only required for rigid bodies with symmetric mass distributions and when using the homogenous matrix representation of the body's position and orientation. If for example, we use the usual vector quaternion representation, a left action can be defined to act on the quaternion and position vectors.

Modularity: As previously described, body i of the real robot should have a reflected equivalent body j with the same CoM position and aligned principle axes of inertia, up to a reflection of any of these axes. If body i is unique in the kinematic three, then $j = i$ and the body must have at least a symmetric mass distribution. When body i is not unique (i.e. there exists another body $j \neq i$ that is its true reflected version), then no constraint of symmetric mass distribution is imposed on i, only the alignment of the principal axes of inertia is required. A didactic example of this scenario is presented in supp.fig 5-right, for Bolt's legs in fig. 2, and for Solo's legs in fig. 1

A.3 Symmetric position and velocity constraint configuration spaces

Although it is implicitly implied on eq. (1) that the constrained position Q and velocity TQ configuration vector spaces should also be symmetric or equivariant, this property might be easily overlooked. As mentioned in section 4 the relevance of morphological symmetries relies in the equivariant nature of the system dynamics (eqs. (2) and (8)), which imprints symmetry constraints on optimal control policies and proprioceptive and exteroceptive measurements. However, with non-symmetric constrained configuration spaces, eq. (1) will not hold for every system state $\boldsymbol{q} \in \mathrm{Q}$, $\dot{\boldsymbol{q}} \in \mathrm{TQ}$, and any uncontrolled or controlled trajectory of the system dynamics shall not have a symmetric equivalent trajectory, as this has the potential to violate the constraints of the configuration space.

B Initialization of equivariant layers

Consider a Equivariant Neural Network architecture composed of multiple layers of Equivariant Linear or Convolutional layers of the form ${}^l\boldsymbol{y} := \sigma({}^l\boldsymbol{W}\,{}^l\boldsymbol{x} + {}^l\boldsymbol{b})$, being l the layer index, ${}^l\boldsymbol{x} \in \mathbb{R}^n$ and ${}^l\boldsymbol{y} \in \mathbb{R}^m$ the layer's input and output vector spaces, ${}^l\boldsymbol{W} \doteq \sum_k^r {}^l c_k \, {}^l\mathbf{B}_{:,:,k} \in \mathbb{R}^{m \times n}$ the layer's linear map, ${}^l\mathbf{B} \in \mathbb{R}^{m \times n \times r}$ the layer's r basis vectors spawning the space of equivariant linear maps, ${}^l\boldsymbol{c} \in \mathbb{R}^r$ the layer's trainable parameters, and ${}^l\boldsymbol{b} \in \mathbb{R}^m$ the layer's bias vector.

For the optimal flow of information throughout the network, it's relevant to initialize the trainable parameters such that the variance of activations (during inference/forward-propagation) and gradients (during back-propagation) is kept constant, avoiding activations/gradients from vanishing or exploiting (Glorot & Bengio, 2010)[4].

The derivation is based on the equivalent process for unconstrained layers presented in He et al. (2015). Let the layer's activations before the non-linearity be denoted by ${}^l\boldsymbol{z} = {}^l\boldsymbol{W}\,{}^l\boldsymbol{x} + {}^l\boldsymbol{b}$, such that ${}^l\boldsymbol{y} = \sigma({}^l\boldsymbol{z})$, and note that ${}^l\boldsymbol{x} = {}^{l-1}\boldsymbol{y}$. Furthermore, we will assume the elements of ${}^l\boldsymbol{c}$ and ${}^l\boldsymbol{x}$ are mutually independent and sampled from two independent distributions, denoting the random variables of the two distributions as ${}^l c$ and ${}^l \mathrm{x}$.

[4]See Pierre Ouannes blog: https://pouannes.github.io/blog/initialization/#xavier-and-kaiming-initialization

The core difference in the initialization of unconstrained and equivariant layers lies in the way the linear map is parameterized. For equivariant layers we have:

$$
\begin{aligned}
\mathrm{Var}({}^l\boldsymbol{W}^l\boldsymbol{x} + {}^l\boldsymbol{b}) &= \sum_i^m \sum_j^n \mathrm{Var}\left({}^lW_{i,j}\,{}^l\mathrm{x}_j\right) && \mid \mathrm{Var}({}^l\boldsymbol{b}) = 0 \\
&= \sum_i^m \sum_j^n \mathrm{Var}\left(\left(\sum_k^r {}^lc_k\,{}^l\mathbf{B}_{m,n,k}\right){}^l\mathrm{x}_j\right) \\
&= \mathrm{Var}\left({}^lc\,{}^l\mathrm{x}\right) \underbrace{\sum_i^m \sum_j^n \sum_k^r \mathbf{B}^2_{m,n,k}}_{\lambda_{{}^l\mathbf{B}}} && \mid \mathrm{Var}\left(\sum_a \underbrace{s_a}_{const}\,\mathbf{p}\right) = \sum_a s_a^2 \mathrm{Var}(\mathbf{p})
\end{aligned}
\tag{15}
$$

In the forward-propagation scenario, we are interested in conserving the variance of the activations throughout layers, that is we must ensure $\mathrm{Var}({}^l\mathbf{z}) = \mathrm{Var}({}^{l-1}\mathbf{z})$. Using supp.eq 15 we obtain:

$$
\begin{aligned}
\mathrm{Var}({}^l\boldsymbol{z}) &= \mathrm{Var}({}^l\boldsymbol{W}^l\boldsymbol{x} + {}^l\boldsymbol{b}) \\
m\,\mathrm{Var}({}^l\mathbf{z}) &= \lambda_{{}^l\mathbf{B}}\mathrm{Var}\left({}^lc\,{}^l\mathrm{x}\right) \\
\mathrm{Var}({}^l\mathbf{z}) &= \frac{\lambda_{{}^l\mathbf{B}}}{m}\left(\underbrace{\mathbb{E}({}^lc^2)}_{\mathrm{Var}({}^lc)}\mathbb{E}({}^l\mathrm{x}^2) - \underbrace{\mathbb{E}({}^lc)^2}_{=0}\mathbb{E}({}^l\mathrm{x})^2\right) \\
\mathrm{Var}({}^l\mathbf{z}) &= \frac{\lambda_{{}^l\mathbf{B}}}{m}\mathrm{Var}({}^lc)\mathbb{E}\left({}^{l-1}\mathrm{y}^2\right) && \mid {}^l\mathrm{x} = {}^{l-1}\mathrm{y} = \sigma({}^{l-1}\mathbf{z}) \\
\mathrm{Var}({}^l\mathbf{z}) &= \frac{\lambda_{{}^l\mathbf{B}}\lambda_\sigma}{m}\mathrm{Var}({}^lc)\mathrm{Var}({}^{l-1}\mathbf{z}) && \mid \mathbb{E}\left({}^{l-1}\mathrm{y}^2\right) = \lambda_\sigma \mathrm{Var}\left({}^{l-1}\mathbf{z}\right) \quad (16) \\
\mathrm{Var}({}^l\mathbf{z}) &\equiv \mathrm{Var}({}^{l-1}\mathbf{z}) && \mid \mathrm{Var}\left({}^{`l}c\right) \doteq \frac{m}{\lambda_{{}^l\mathbf{B}}\lambda_\sigma} \quad (17)
\end{aligned}
$$

where λ_σ in supp.eq 16 is a non-linearity dependent scalar computed analytically or empirically (see He et al. (2015)). In supp.eq 17 we conclude that if we sample the equivariant layer trainable parameters ${}^l\boldsymbol{c}$ from a distribution ensuring $\mathrm{Var}({}^lc) \doteq \frac{m}{\lambda_{{}^l\mathbf{B}}\lambda_\sigma}$, the variance of the activations across equivariant layers remain constant. A similar procedure can be applied to the backward propagation case, concluding that in order to maintain a constant variance of the gradients across the network layers we should sample the trainable parameters ensuring $\mathrm{Var}({}^lc) \doteq \frac{n}{\lambda_{{}^l\mathbf{B}}\lambda_\sigma}$. As remarked in He et al. (2015) both variance values for the forward and backward propagation cases lead to the proper flow of information in the network. On supp.fig 7, it can be appreciated that our method achieves equivalent results for equivariant architectures as He et al. (2015) does for standard linear and convolutional architectures.

C IMPLEMENTATION DETAILS & CODE

Additional to this section, we provide open-access code with the scripts for reproducing the experiments of this work, the parameters of the models used for comparison, along with additional interactive examples visualizing morphological symmetries of both robotic systems and data.

C.1 EFFICIENT DATA AUGMENTATION

Since any input $\boldsymbol{x}$ and output $\boldsymbol{y}$ spaces of equivariant architectures have matrix symmetry action representations, $\rho_{\boldsymbol{x}}(g)$, $\rho_{\boldsymbol{y}}(g)$, it is possible to perform batched data augmentation, reducing the computational complexity of augmenting a batch of N_b samples from N_b matrix-vector multiplications to a single matrix-matrix multiplication, preferably performed after data is loaded to GPU for optimal performance.

Table 1: Robot and Model parameters for CoM and Contact state estimation.

Robot	**Model**	$\mathcal{G}$	**lr**	**Data Samples**
Solo	MLP/EMLP	$\mathcal{C}_2/\mathcal{K}_4$	2.4×10^{-3}	100k
Atlas	MLP/EMLP	$\mathcal{C}_2$	1.5×10^{-3}	500k
MIT-MiniCheetah	CNN	$\mathcal{C}_2$	1.0×10^{-4}	730k
MIT-MiniCheetah	E-CNN	$\mathcal{C}_2$	1.0×10^{-5}	

C.2 Hyperparameter tunning

The only hyper-parameter tunned for each model and model variant was the learning rate. For all model variants presented in this work (except the original Contact-CNN model from Lin et al. (2021), which we retrained using the same hyperparameters reported by the authors) we ran a grid-search in log-scale among 20 different learning rates. In this scenario, we always used the entire training dataset and optimized w.r.t computed loss in the entire validation partition. The learning rate values used for each model are depicted in table 1.

C.3 Experiment: CoM Momentum Estimation

The dataset for the CoM estimation experiment was generated using Pinocchio (Carpentier et al., 2019), which in turn uses the URDF models of the robots Solo and Atlas, to extract the kinematic and dynamic parameters required to compute the Centroidal Momentum Matrix $\mathbf{A}_G(\boldsymbol{q})$ matrix (Orin et al., 2013), with which computing the CoM momentum reduces to:

$$g \cdot \boldsymbol{h} = \mathbf{A}_G(g \cdot \hat{\boldsymbol{q}}) g \cdot \hat{\dot{\boldsymbol{q}}} \qquad | \quad \forall g \in \mathcal{G}. \tag{18}$$

For a robotic system with a morphological symmetry group $\mathcal{G}$, for which every group action g imitates a specific reflection or rotation of space $\overline{g}$, the function to compute the CoM momentum supp.eq 18 becomes equivariant:

$$\begin{bmatrix} \rho_{\mathbb{E}_3}(\overline{g}) & \mathbf{0} \\ \mathbf{0} & \rho_{\mathbb{E}_3}(\overline{g}) \end{bmatrix} \begin{bmatrix} \boldsymbol{l} \\ \mathbf{k} \end{bmatrix} = \mathbf{A}_G\left(\rho_{\mathrm{Q}}(g)\hat{\boldsymbol{q}}\right) \rho_{\mathrm{Q}}(g)\hat{\dot{\boldsymbol{q}}} \qquad | \quad \forall \quad g, \overline{g} \in \mathcal{G}. \tag{19}$$

Note that by definition any improper transformation applied to a pseudo-vector (e.g. angular velocity/momentum, torques) is computed as $|\boldsymbol{R}|\boldsymbol{R} \cdot \mathbf{k}$.

For each of the robots we generate N datasamples (see table 1) composed of $\boldsymbol{x} = [\hat{\boldsymbol{q}}, \hat{\dot{\boldsymbol{q}}}] \in \mathbb{R}^{2|\mathrm{Q}|}$ and $\boldsymbol{y} = \boldsymbol{h}$.

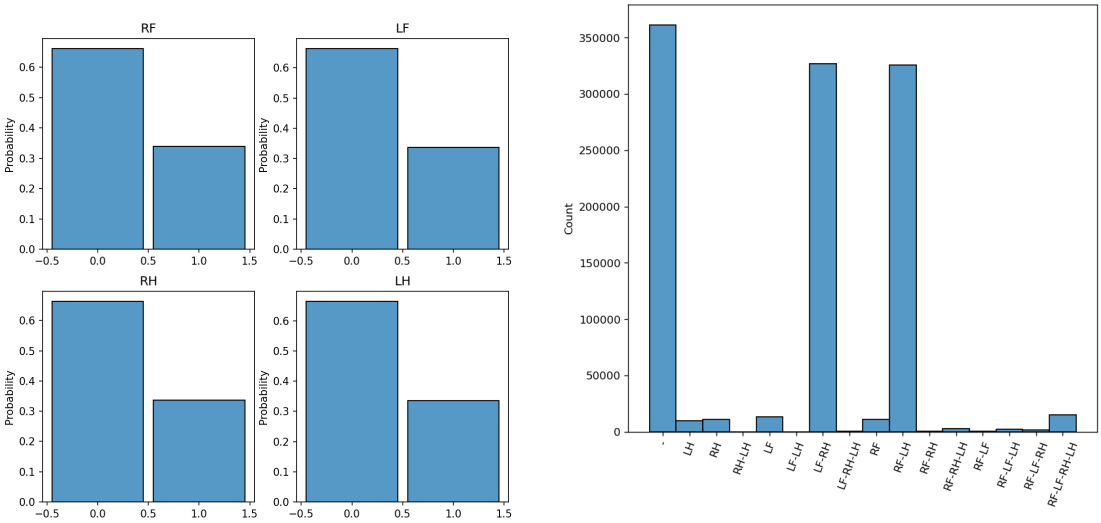
The URDF files of the robots Solo and Atlas are generated using XACRO scripts, which replicate the structure of limbs to their symmetric counterparts, making the dynamics of the robots in simulation exactly $\mathcal{G}$-equivariant. However, the algorithm for computing the CoM momentum from Pinocchio is numerically sensitive, resulting in the orbits of the momentum $\mathcal{G} \cdot \boldsymbol{h}$ deviating slightly from the theoretical orbits. Therefore to reduce numerical errors and ensure the theoretical equivariance of the data, we replace every target variable by the average of its orbit $\boldsymbol{y} = \boldsymbol{h} \doteq \frac{1}{|G|} \sum \mathcal{G} \cdot g^{-}1(\mathbf{A}_G\left(\rho_{\mathrm{Q}}(g)\hat{\boldsymbol{q}}\right) \rho_{\mathrm{Q}}(g)\hat{\dot{\boldsymbol{q}}}) \quad | \quad \forall \quad g \in \mathcal{G}$.

C.4 Experiment: Static-Friction-Regime Contact Detection

The dataset presented in (Lin et al., 2021) is composed of output samples $\boldsymbol{y} \in \mathbb{R}^{16}$, where each dimension of $\boldsymbol{y}$ represents a logit of a specific contact state, among the 16 different combinations of each of the 4 legs possible binary contact states. The input samples $\boldsymbol{y} = \{\boldsymbol{z}_i\}_{i=0}^{150} \in \mathbb{R}^{54\times150}$, are a history of 150 samples $\boldsymbol{z} = [\hat{\boldsymbol{q}}, \hat{\dot{\boldsymbol{q}}}, {}^B\boldsymbol{a}, {}^B\boldsymbol{w}, {}^B\boldsymbol{p}, {}^B\boldsymbol{v}] \in \mathbb{R}^{54}$. Where $\hat{\boldsymbol{q}} \in \mathbb{R}^{n_j}, \hat{\dot{\boldsymbol{q}}} \in \mathbb{R}^{n_j}, {}^B\boldsymbol{a} \in \mathbb{R}^3, {}^B\boldsymbol{w} \in \mathbb{R}^3, {}^B\boldsymbol{p} \in \mathbb{R}^{12}, {}^B\boldsymbol{v} \in \mathbb{R}^{12}$ are the MIT-Mini-Cheetah robot joint-space positions, velocities, base linear acceleration, base angular velocity, and each of the four legs feet's position and velocities, respectively, referenced to the robots base frame B.

The function approximator to learn is expected to be approximately equivariant to the reflection group $\mathcal{C}_2$, considering the sagittal symmetry of the robot morphology. Therefore:

$$\begin{aligned} g \cdot \boldsymbol{y} &= f(g \cdot \boldsymbol{x}; \boldsymbol{\phi}) \quad | \quad g \in \mathcal{G} = \mathcal{C}_2 \\ \rho_{\boldsymbol{y}}(g)\boldsymbol{y} &= f(\rho_{\boldsymbol{x}}(g) \cdot \boldsymbol{x}; \boldsymbol{\phi}) \quad | \quad g \in \mathcal{G} = \mathcal{C}_2, \end{aligned} \tag{20}$$



Supplementary Figure 6: Sample distribution of the training set of the contact dataset presented by Lin et al. (2021). The left figure shows the distribution of the binary contact state for each of the mini-cheetah legs, clearly displaying a similar distribution among legs favoring states of no contact (approximately for each leg 60% of data samples represent states of no contact). The right figure shows the distribution of contact states ($\boldsymbol{y} \in \mathbb{R}^{16}$), clearly showing the dataset imbalance caused by the majority of the data being collected by the robot performing a trot gait.

Table 2: Symmetric contact state for Mini-Cheetah quadruped robot, considering its sagittal symmetry, and morphological symmetry group $\mathcal{C}_2$. Each leg binary contact state (LF: Left Front, RF: Right Front, LH: Left Hind, RH: Right Hind) is displayed with its corresponding robot contact state $\boldsymbol{y}$.

RF	**LF**	**RH**	**LH**	$\boldsymbol{y}$	$g \cdot \boldsymbol{y}$	**LF**	**RF**	**LH**	**RH**
0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	2	0	0	1	0
0	0	1	0	2	1	0	0	0	1
0	0	1	1	3	3	0	0	1	1
0	1	0	0	4	8	1	0	0	0
0	1	0	1	5	10	1	0	1	0
0	1	1	0	6	9	1	0	0	1
0	1	1	1	7	11	1	0	1	1
1	0	0	0	8	4	0	1	0	0
1	0	0	1	9	6	0	1	1	0
1	0	1	0	10	5	0	1	0	1
1	0	1	1	11	7	0	1	1	1
1	1	0	0	12	12	1	1	0	0
1	1	0	1	13	14	1	1	1	0
1	1	1	0	14	13	1	1	0	1
1	1	1	1	15	15	1	1	1	1

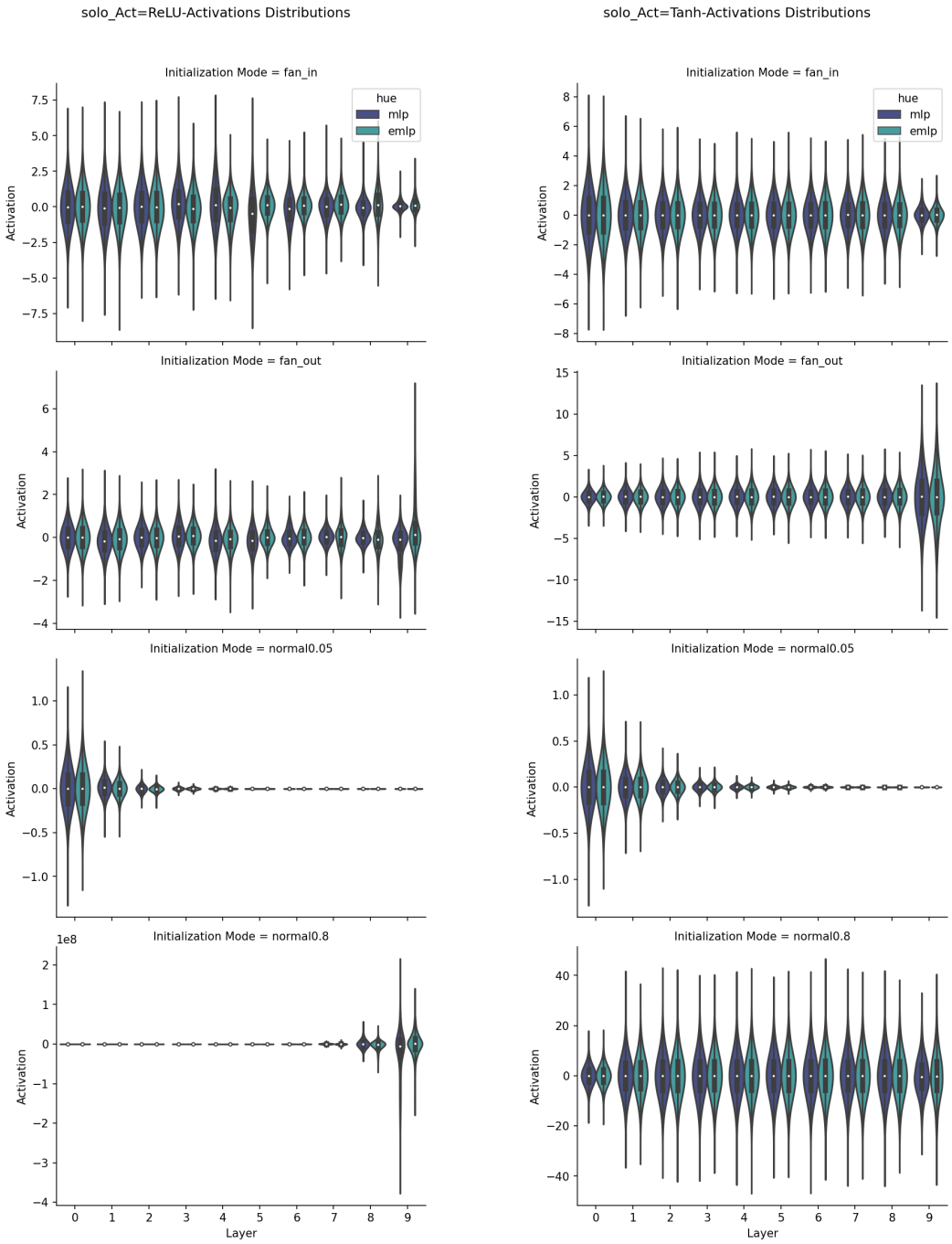
where $\rho_{\boldsymbol{x}}(g) = BlockDiag([\rho_{\hat{\boldsymbol{q}}}(g), \rho_{\hat{\boldsymbol{q}}}(g), [\rho_{\mathbb{E}_3}(g)]) \in \mathbb{R}^{54\times 54}$, and $\rho_{\boldsymbol{y}}(g)$ is given by the permutation matrix relating $\boldsymbol{y}$ and $g \cdot \boldsymbol{y}$ described in table 2.

For details on the construction of the Equivariant 1D Convolutional layers reefer to [1]. Note that the symmetry of a single time-sample $\boldsymbol{z}_i$ is shared across all time-samples $\boldsymbol{y} = \{\boldsymbol{z}_i\}_{i=0}^{150}$.

We modified the original dataset partitioning to properly evaluate the generalization capacity of the models. The original dataset was composed of 15 different recordings varying ground type and gait type used during data collection (most recordings were performed on a trot gait, which heavily biased the dataset to contact states 0, 6, and 9 of table 2. See supp.fig 6). The authors of (Lin et al., 2021)

partitioned all 15 recordings into $(70\%, 15\%, 15\%)$ training, validation and testing. This partition was made such that the first 70% time-samples of each recording were assigned for training, the following 15% to validation, and the rest for testing. Because we are interested in studying the generalization capacity of the models and the out-of-training-distribution performance, we modified this partitioning such that among the 15 different recordings we selected randomly 5 recordings for testing, and the remaining 10 recordings were used for training splitting these recordings into (85%, 15%) training and validation splits as in Lin et al. (2021), that is, for each recording, the first 85% data-samples go for training and the remaining for validation.

The selected training-validation recordings were: air_walking_gait, concrete_difficult_slippery, concrete_left_circle, middle_pebble, rock_road, asphalt_road, concrete_galloping, grass, old_asphalt_road, sidewalk. While the selected testing recordings were: air_jumping_gait, concrete_pronking, concrete_right_circle, forest, small_pebble.



Supplementary Figure 7: Comparison of the initialization method of unconstrained layers of He et al. (2015) with our initialization method for equivariant layers. Left and right columns correspond to MLP and EMLP architectures with $\sigma = ReLu$ and $\sigma = Tanh$ non-linearities, respectively. Each row shows different initialization methods differing in the variance of the initialization distribution of the layer's trainable parameters. First and second rows show the forward and backward propagation cases of He et al. (2015) for MLP and of supplementary B for EMLP, with $\mathrm{Var}({}^l c) \doteq {}^m/_{(\lambda_{l_{\mathbf{B}}} \lambda_\sigma)}$ and $\mathrm{Var}({}^l c) \doteq {}^n/_{(\lambda_{l_{\mathbf{B}}} \lambda_\sigma)}$, respectively. In these cases, the variance of activations through the network depth remains nearly constant, as desired. The last two rows show the initialization of layer parameters with a constant variance of 0.05^2 and 0.8^2, illustrating scenarios of activations vanishings and exploiting. All architectures are composed of 10-layers with 256 neurons on intermediate layers. In the equivariant case, the network is $\mathcal{K}_4$-equivariant.