LieGG: Studying learned Lie Group Generators

Artem Moskalev, Anna Sepliarskaia, Ivan Sosnovik, Arnold Smeulders

$$g = \exp(\mathfrak{h} \cdot t)$$

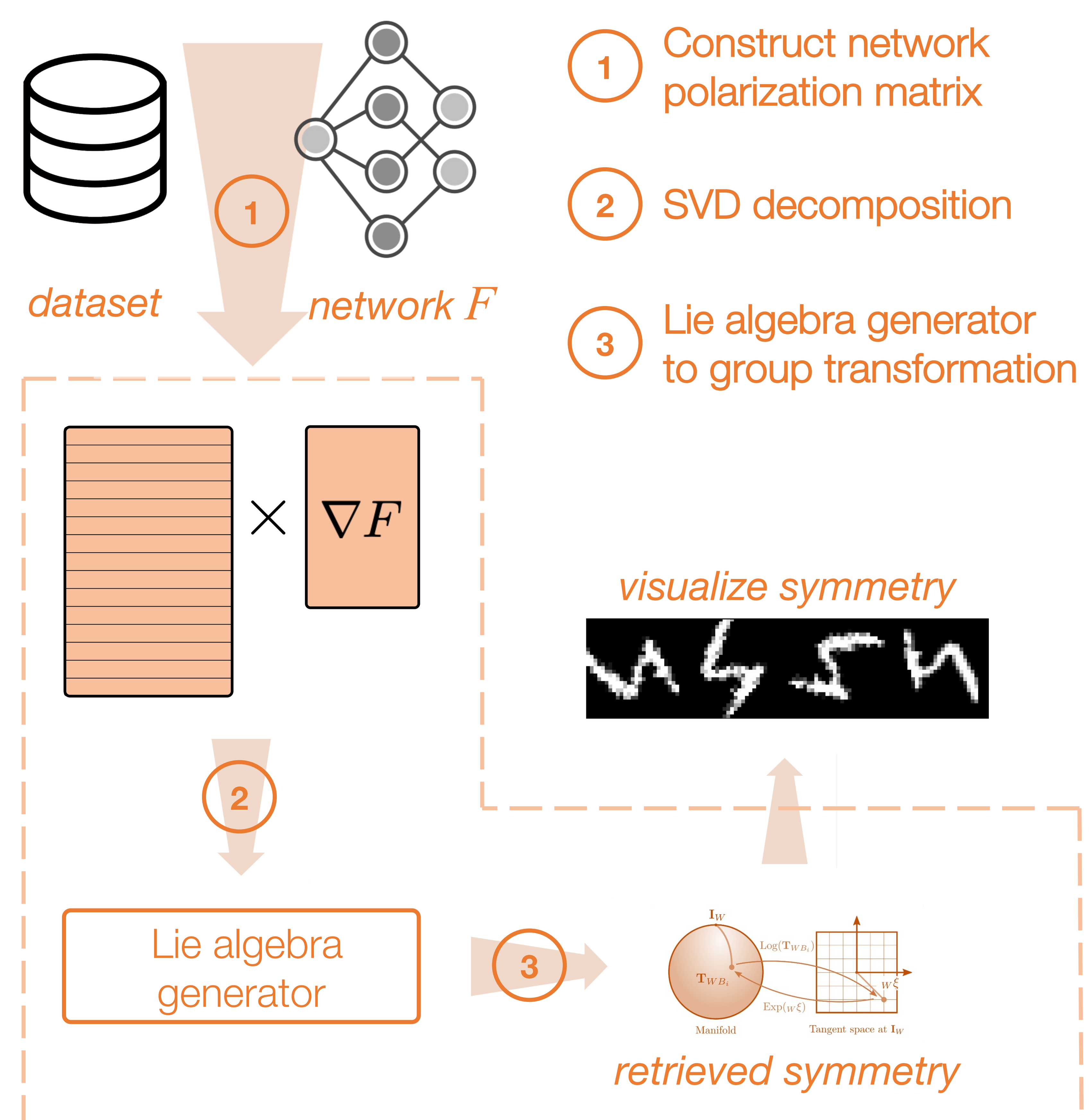
Summary:

We present a method to **extract symmetries** learned by a neural network and to **evaluate** the degree to which a network is invariant to them.

We utilize **Lie group – Lie algebra correspondence** to explicitly retrieve learned invariances in a form of the generators of corresponding Lie-groups **without prior knowledge** of symmetries in the data.

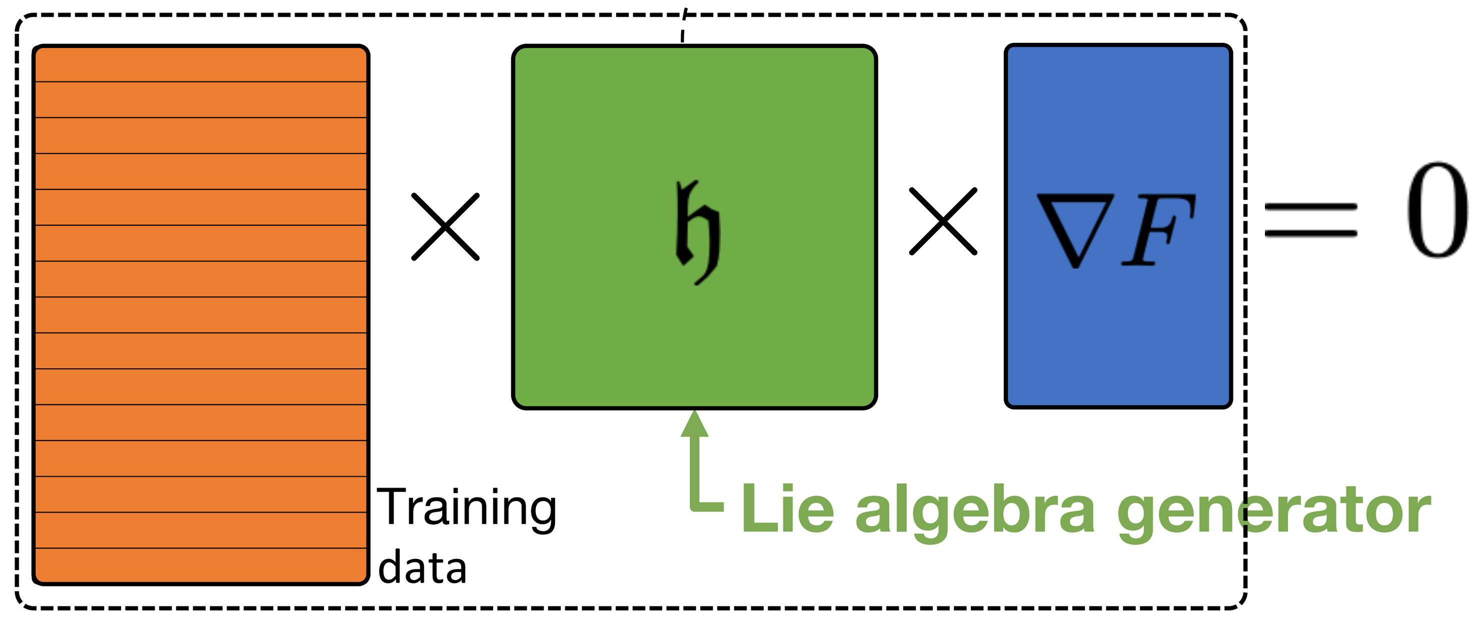
We study how the symmetrical properties of neural networks depend on parameterization and configuration of a model.

LieGG overview:



1. Construct network polarization matrix
2. SVD decomposition
3. Lie algebra generator to group transformation

Lie algebra:



Lie group:

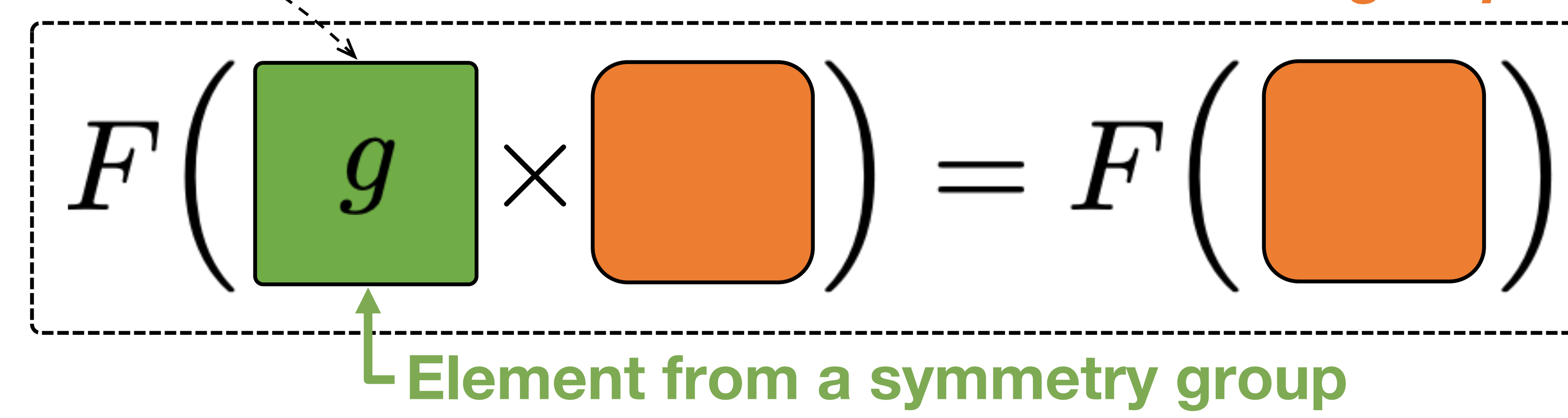


Figure 1:
We use invariance criterion for Lie algebras to turn symmetry extraction into a matrix nullspace problem.

Symmetry generator as the nullspace:

$$\sum_i^n \sum_j^n \frac{\partial F}{\partial x_i} \cdot \mathfrak{h}_{ij} \cdot x_j = 0$$

$$\mathcal{E} \overset{\text{SVD}}{=} U\Sigma V^T$$
$$\mathfrak{h} \in Nullspace(\mathcal{E})$$

*network polarization matrix

Singular vectors (w/ near-zero singular values) form a **nullspace**

They are the **generators** of Lie algebra

Symmetry variance & Symmetry bias:

*Symmetry **variance**:*

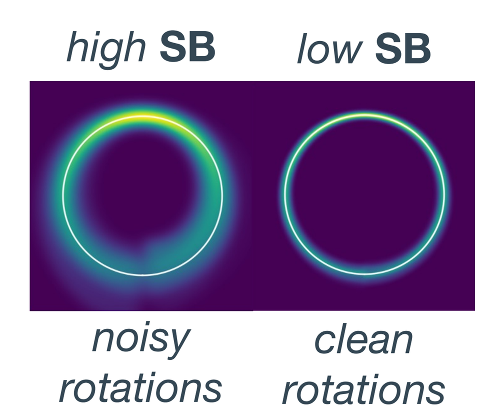
$$\mathcal{V}(\mathcal{E}) = \sigma_{min}(\mathcal{E})^2/|D| \sim \mathbb{E}_{x\sim\mathcal{D}}[(F(gx) - F(x))^2]$$

*how invariant a network is to the learned symmetry

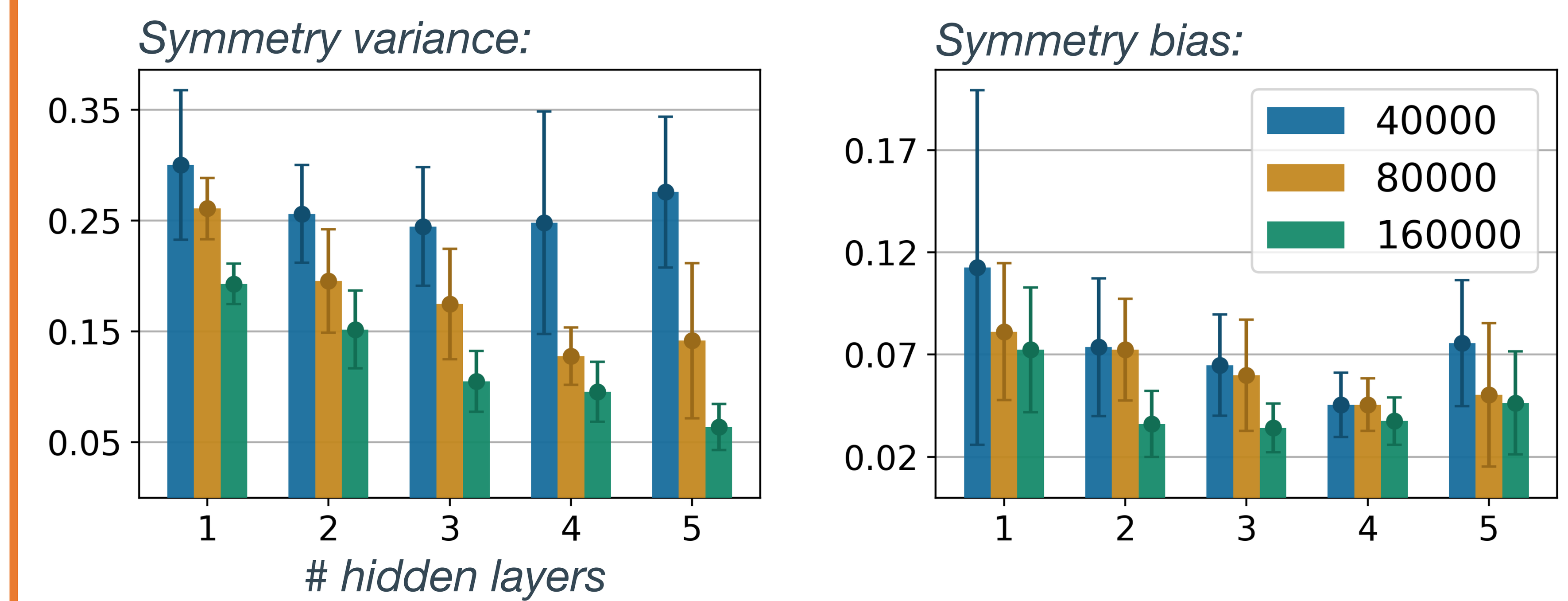
*Symmetry **bias**:*

$$\mathcal{B}_i = ||(V^T)_i - vec(\mathfrak{h}_{true})||_2$$

*how close is the learned symmetry to the true symmetry

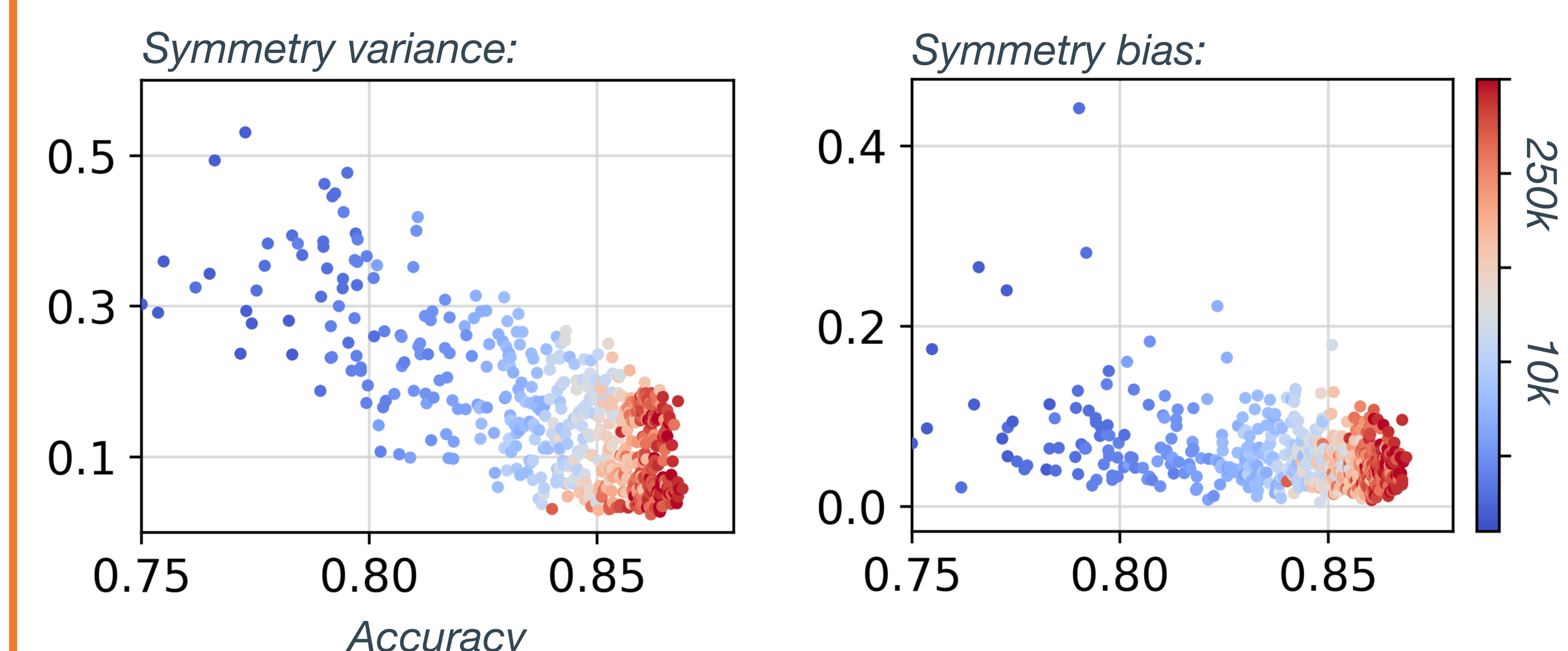


Symmetrical properties vs Configuration:



deeper networks are better symmetry learners than wider ones

Symmetry learning vs Accuracy:



#parameters matters