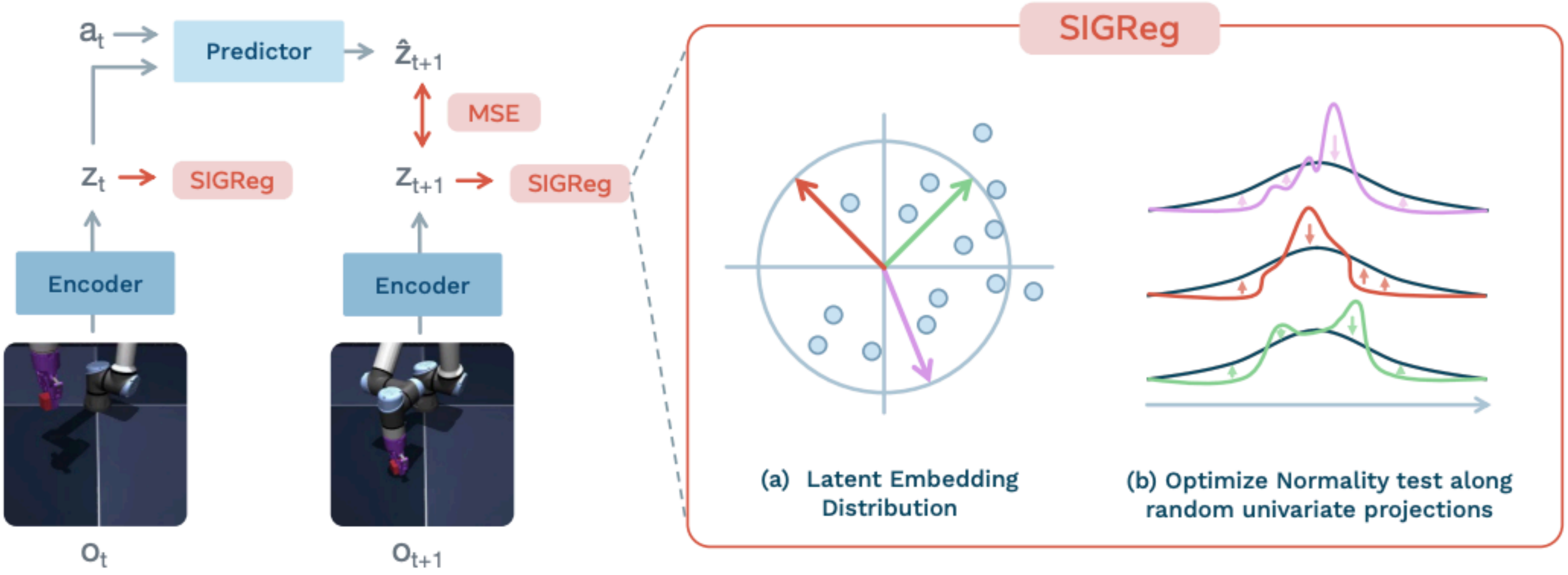
CSCI1470

Deep Learning

Randall Balestriero

Recap

LeWorldModel

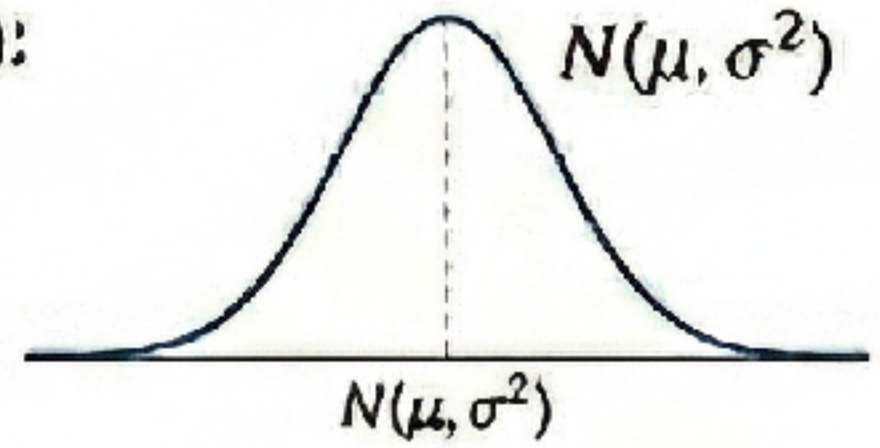


Any idea?

THE GAUSSIAN DISTRIBUTION

1. Probability Density Function (PDF):

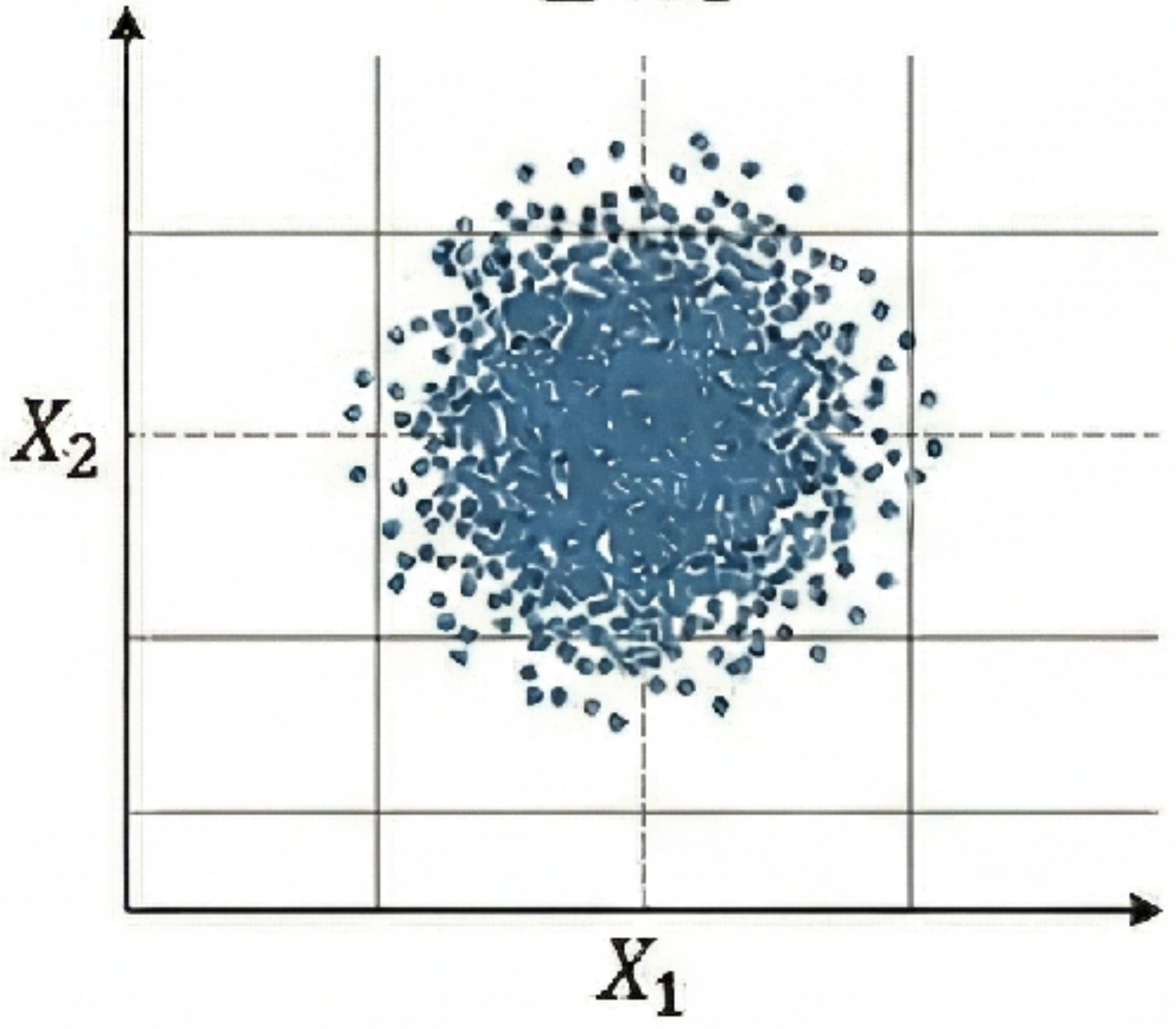
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



2. 2D Gaussian Samples:

Isotropic Covariance

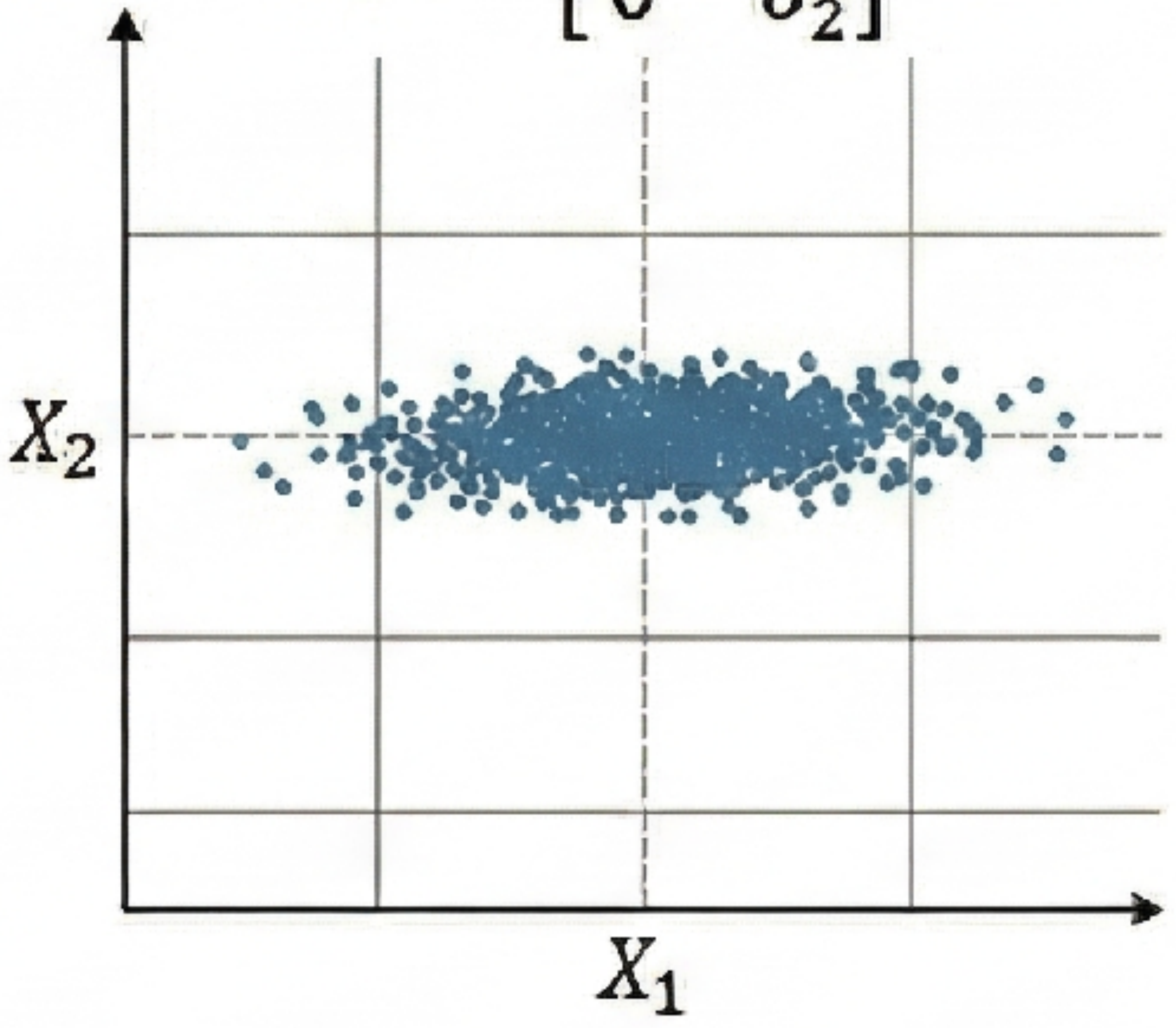
$\mathbf{\Sigma} \propto \mathbf{I}$



Isotropic Covariance ($\mathbf{\Sigma} = \sigma^2 \mathbf{I}$) - Circular Symmetry

Diagonal Covariance (UNCORRELATED)

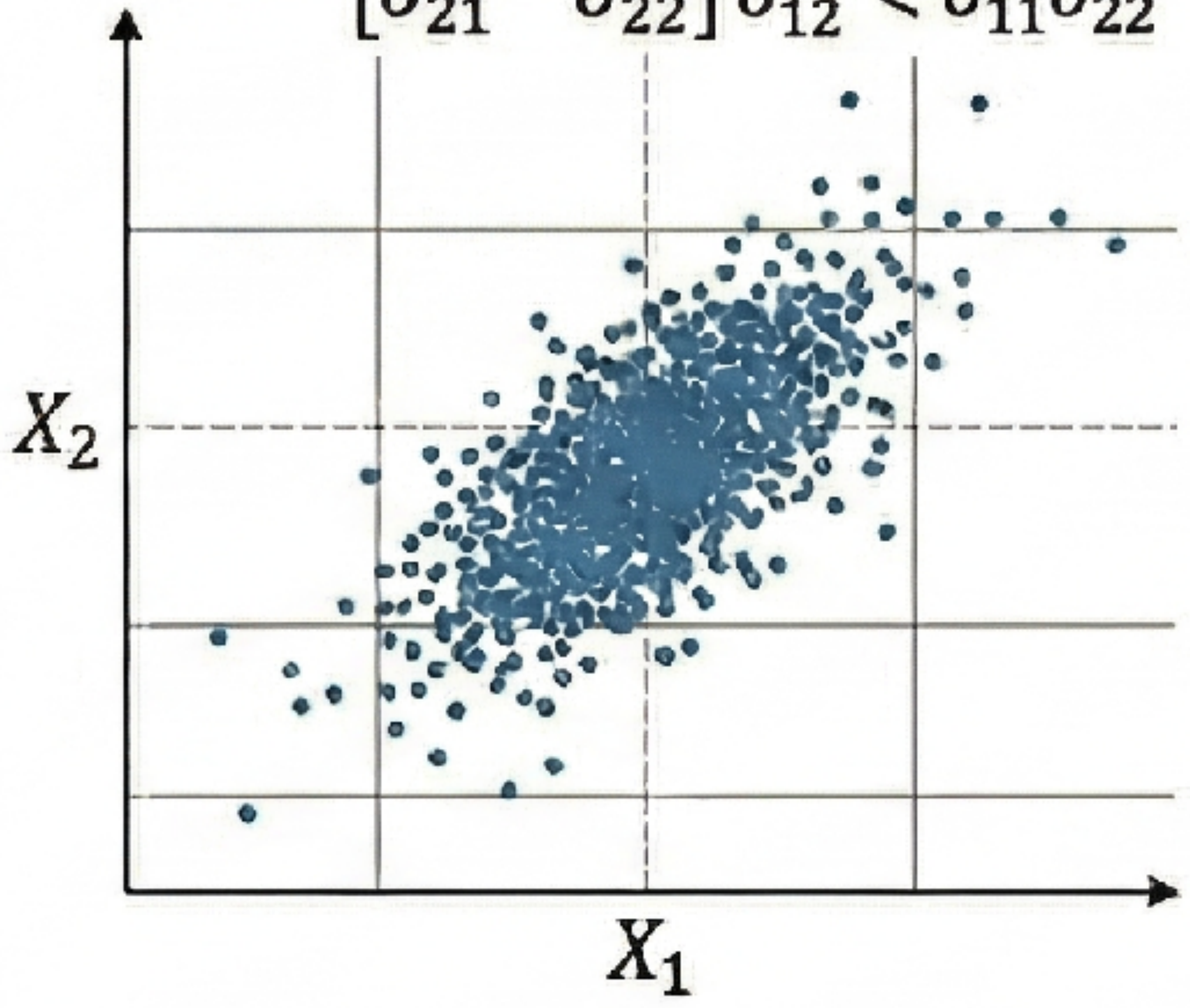
$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$



Diagonal Covariance ($\mathbf{\Sigma} = diag(\sigma_1^2, \sigma_2^2)$) - Axis-Aligned

Full Covariance (GENERAL $\mathbf{\Sigma}$)

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \quad \sigma_{12} = \sigma_{21}, \quad \sigma_{12}^2 < \sigma_{11}\sigma_{22}$$



Full Covariance (General $\mathbf{\Sigma}$) - Rotated Ellipse

The shape of the scatter changes based on the covariance matrix.

THE CHARACTERISTIC FUNCTION

1. The Math:

1. The Math:

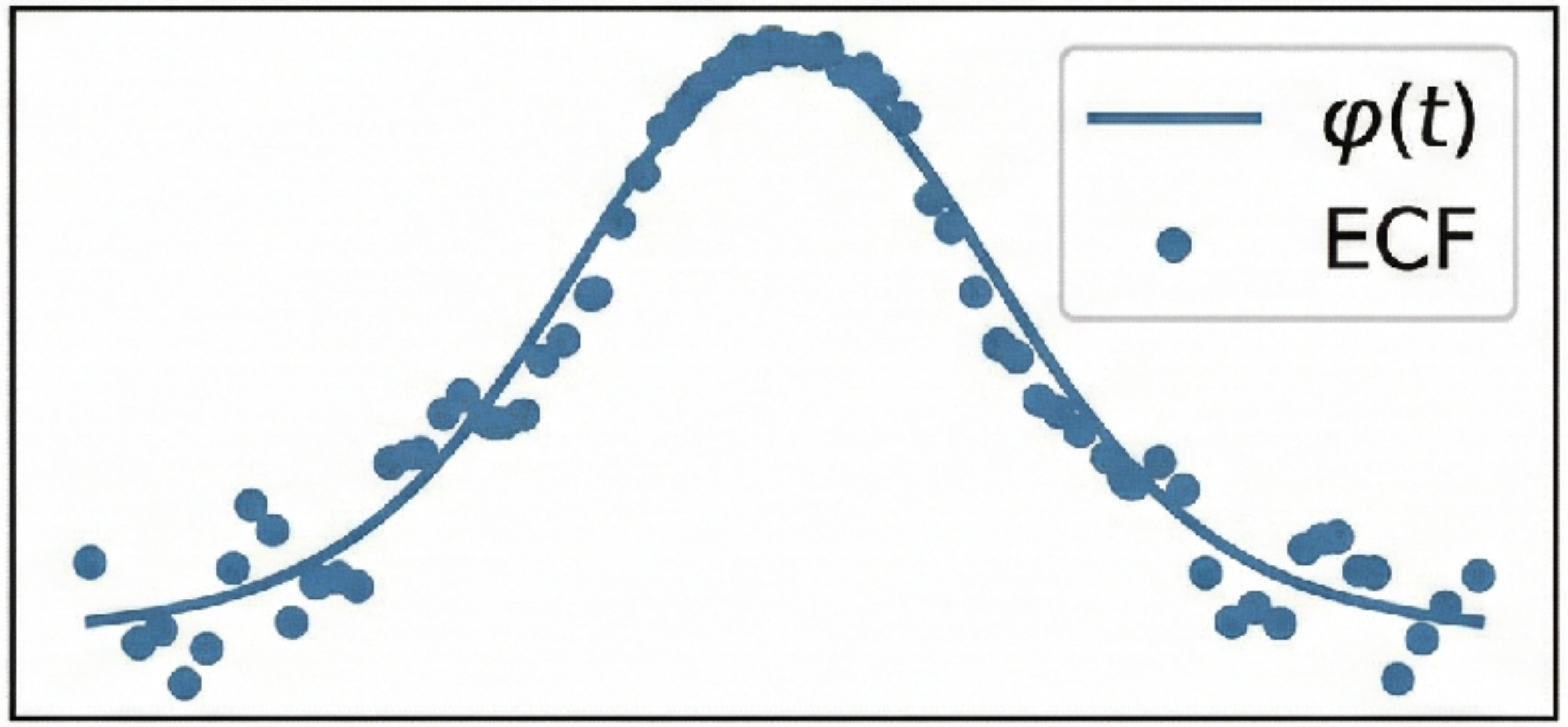
$\int^{\infty} e^{itx} f(x)dx$

$$\varphi_X(t) = E[e^{itX}] = \int_{-\infty}^{\infty} e^{itx} f(x)dx$$

- $\varphi_X(t)$ is the characteristic function of random variable X
- t is a real number argument
- $f(x)$ is the Probability Density Function (PDF) of X

2. Empirical Characteristic Function (ECF):

$$\hat{\varphi}_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{itX_j}$$



The Empirical Characteristic Function estimates the theoretical function from sampled data.

PROPERTIES OF CHARACTERISTIC FUNCTIONS

1. Key Mathematical Properties:

   a) **Uniqueness**: The characteristic function determines the distribution of **X** uniquely.

   b) **Symmetry**: If **X** is a real-valued random variable, $\varphi_X(-t) = \overline{\varphi_X(t)}$.

   c) **Boundedness**: $|\varphi_X(t)| \leq 1$ and $\varphi_X(0) = 1$.

   d) **Linearity**: For constant constants $a, b \in \mathbb{R}$, $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$.

2. Linear Combination Property (Convolution):

$$\varphi_{\sum X_j}(t) = \prod \varphi_{X_j}(t)$$

Characteristic functions simplify the analysis of linear combinations of random variables.

DERIVATION OF THE GAUSSIAN CHARACTERISTIC FUNCTION

1. Standard Normal Setup:
   - Assume $X \sim \mathcal{N}(0,1)$.
   - PDF is: $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.
   - CF Definition: $\varphi(t) = E[e^{itX}]$.
2. Integrating over the PDF:
   - $\varphi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-x^2/2} dx$.
     - $\varphi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2-2itx)} dx$.
3. Completing the Square:
   - In the exponent: $x^2 - 2itx = (t)^2 - (it)^2$.

$$x^2 - 2itx = (x - it)^2 + t^2.$$

4. Substituting Back and Evaluating:
   - $\varphi(t) = e^{-t^2/2} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-it)^2} dx \right]$.
     - The term in [brackets] is the integral of a *shifted* Gaussian PDF, so it evaluates to 1.

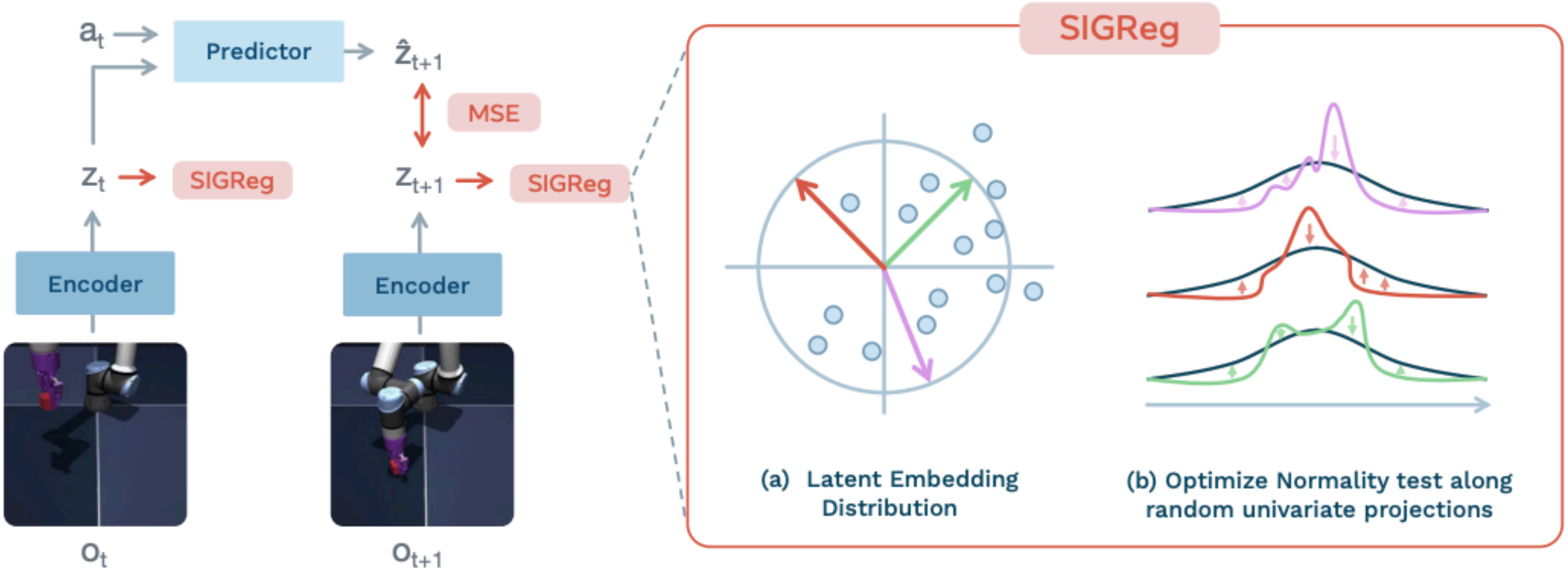
$$\boldsymbol{\varphi(t) = e^{-t^2/2}}.$$

The Gaussian characteristic function has an exponential quadratic form.

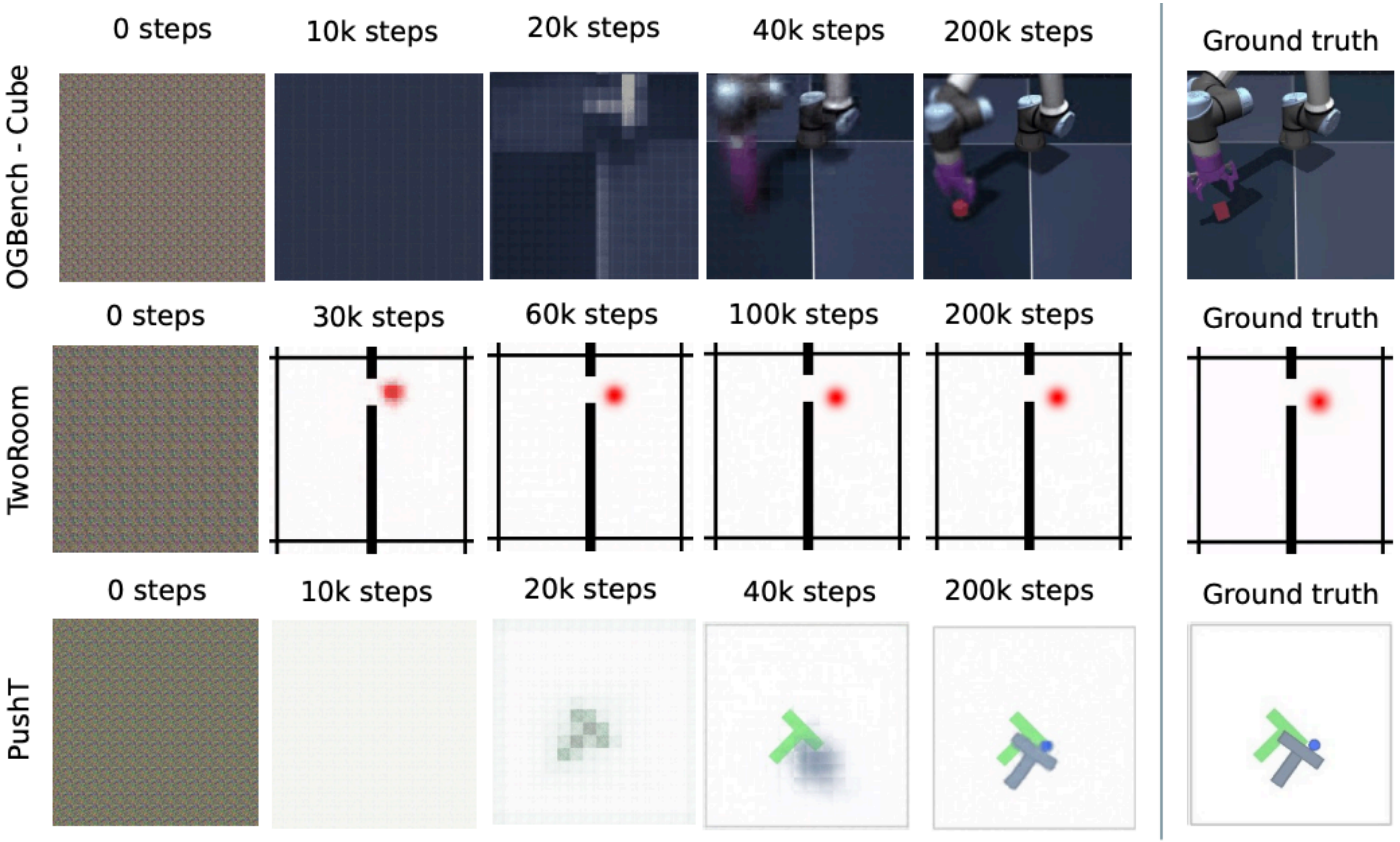
Any idea?

```python
def SIGReg(x, global_step, num_slices=256):
    # slice sampling -- synced across devices --
    dev = dict(device=x.device)
    g = torch.Generator(**dev)
    g.manual_seed(global_step)
    proj_shape = (x.size(1), num_slices)
    A = torch.randn(proj_shape, generator=g, **dev)
    A /= A.norm(p=2, dim=0)
    # -- Epps-Pulley stat. see Sec. 4.3 for alt. --
    # integration points
    t = torch.linspace(-5, 5, 17, **dev)
    # theoretical CF for N(0, 1) and Gauss. window
    exp_f = torch.exp(-0.5 * t**2)
    # empirical CF -- gathered across devices --
    x_t = (x @ A).unsqueeze(2) * t  # (N, M, T)
    ecf = (1j * x_t).exp().mean(0)
    ecf = all_reduce(ecf, op="AVG")
    # weighted L2 distance
    err = (ecf - exp_f).abs().square().mul(exp_f)
    N = x.size(0) * world_size
    T = torch.trapz(err, t, dim=1) * N
    return T
```

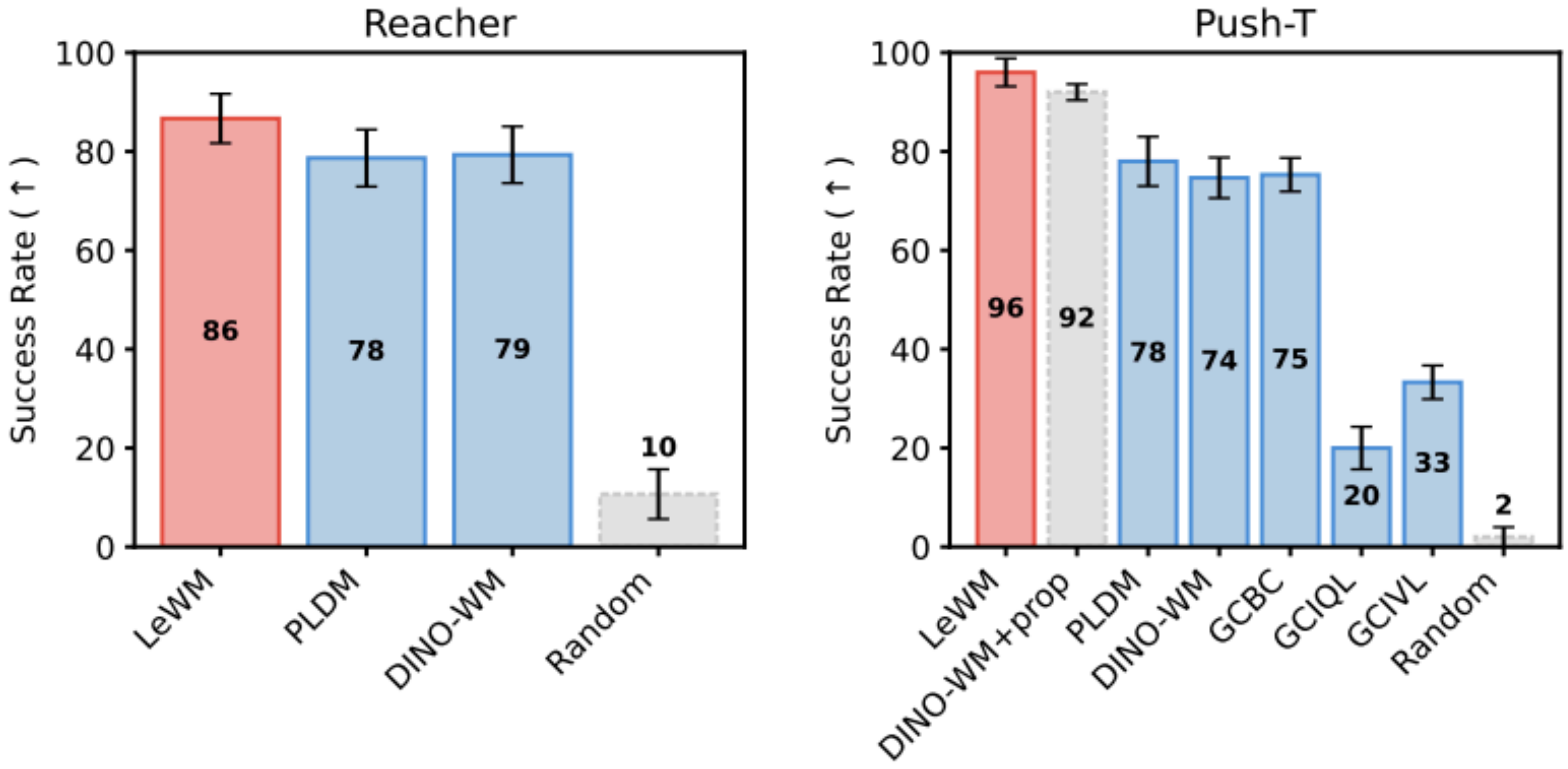
LeWorldModel



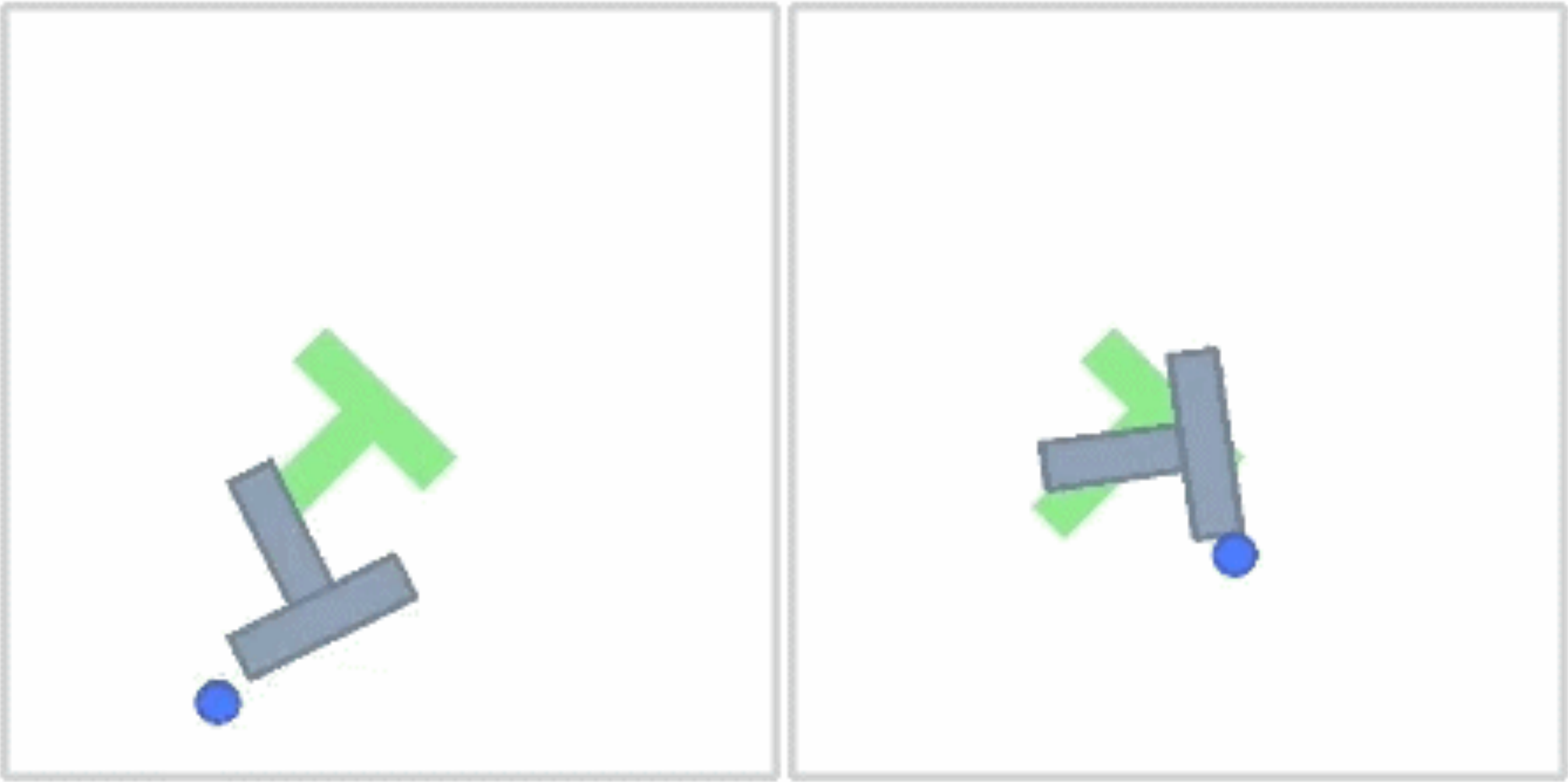
LeWorldModel

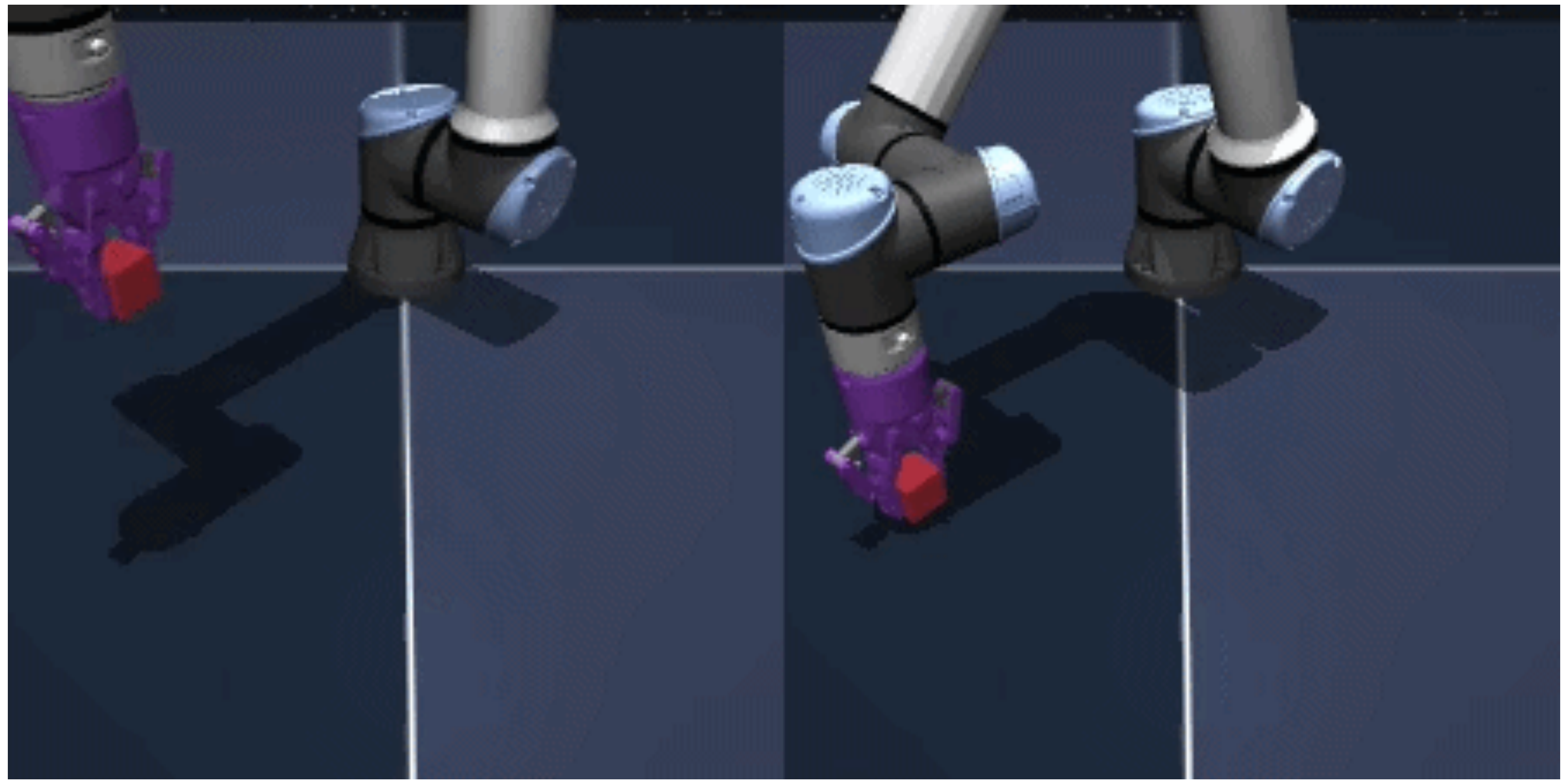


LeWorldModel

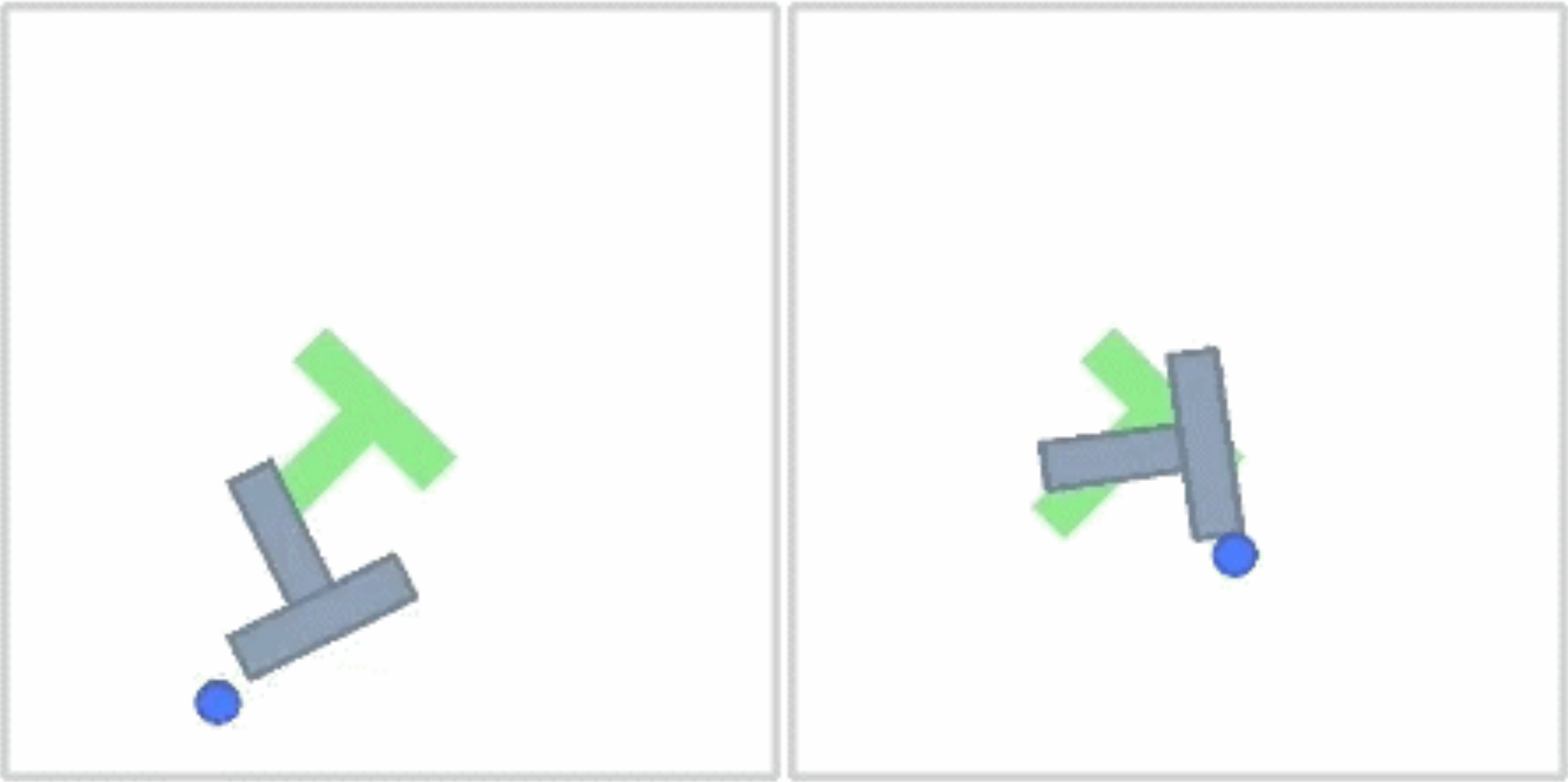


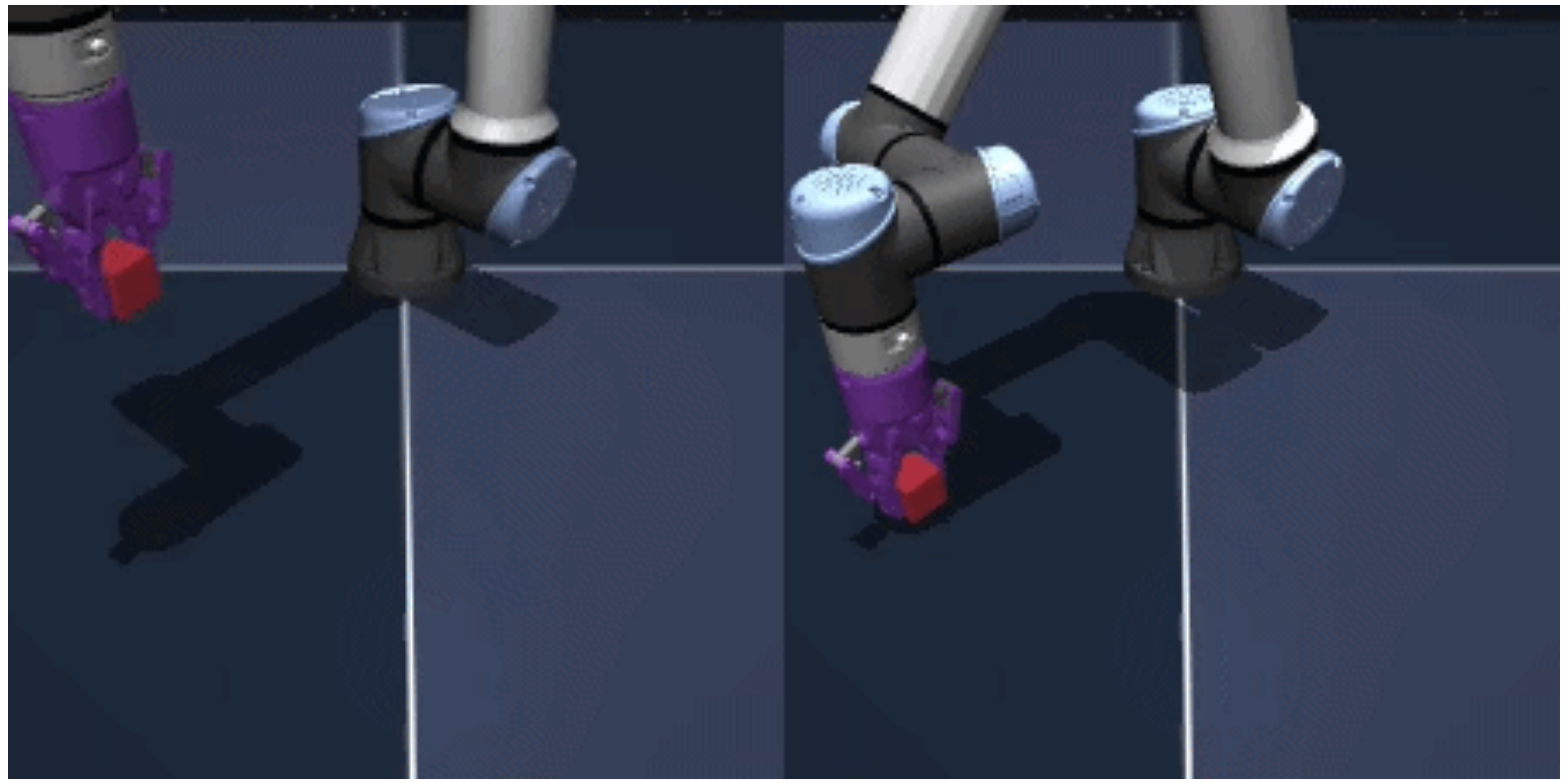
LeWorldModel





LeWorldModel





stable-worldmodel

World model research made simple. From data collection to training and evaluation.

Docs | Tests passing | pypi v0.0.6 | PyTorch | Ruff

Quick Example | Environments | Installation | Documentation | Contributing | Citation

Thank you!
See you Wednesday!