Overview

Differentiable Architecture Search (DARTS)

DARTS as Surrogate

Beyond Finite Search Spaces

Overview

Neural Architecture Search (NAS)

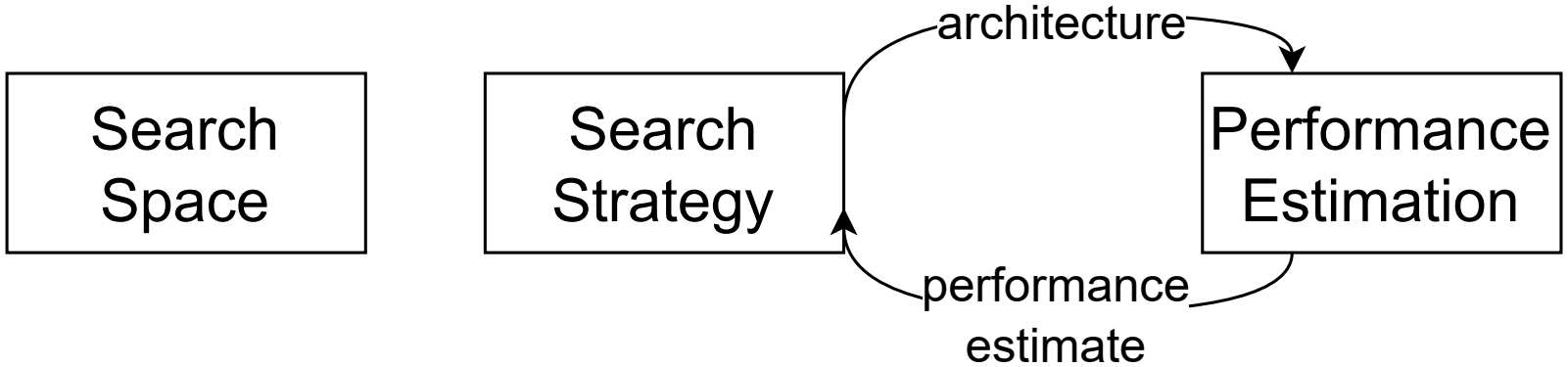
- Automatize choice of neural network architecture

Neural Architecture Search (NAS)

- Automatize choice of neural network architecture
- Discover new architectures

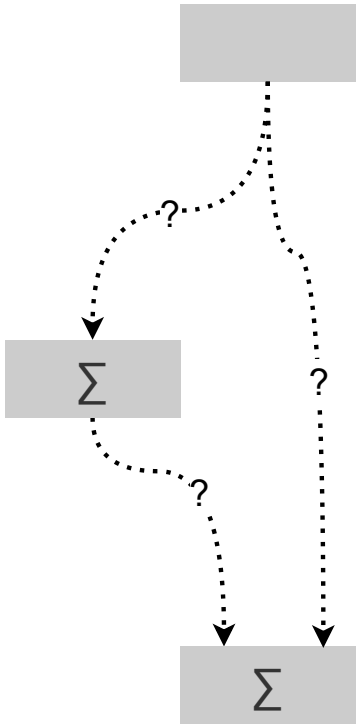
Neural Architecture Search (NAS)

- Automatize choice of neural network architecture
- Discover new architectures



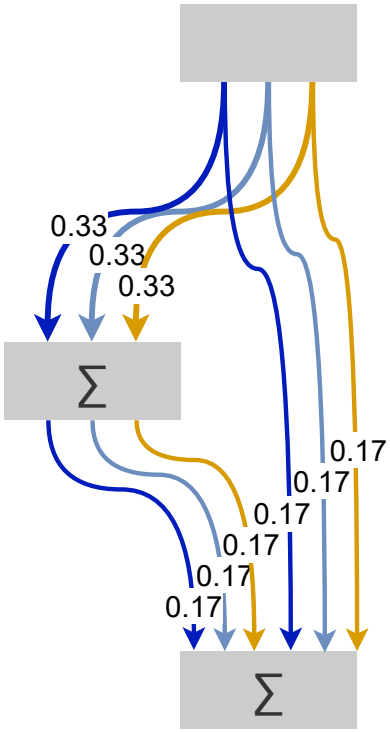
Differentiable Architecture Search

DARTS [Liu et al., 2018] considered as pioneer work



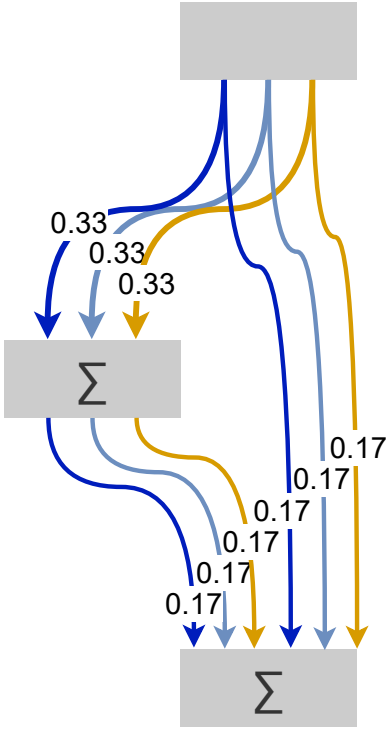
Differentiable Architecture Search

DARTS [Liu et al., 2018] considered as pioneer work

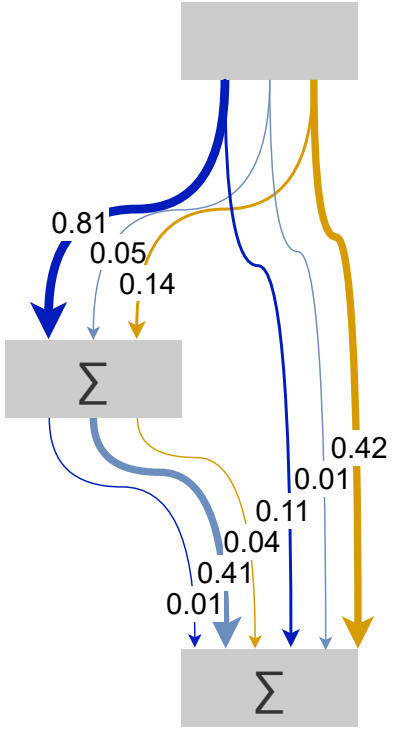


Differentiable Architecture Search

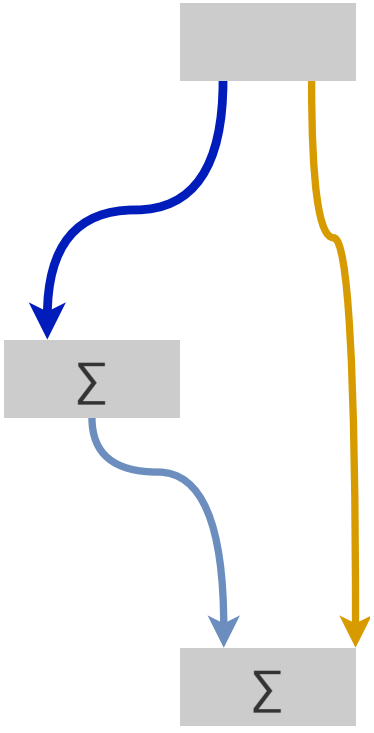
DARTS [Liu et al., 2018] considered as pioneer work



Training start



Training end



Obtain best architecture

Gumbel-Softmax Sampling

We define the *Standard Gumbel* probability density as

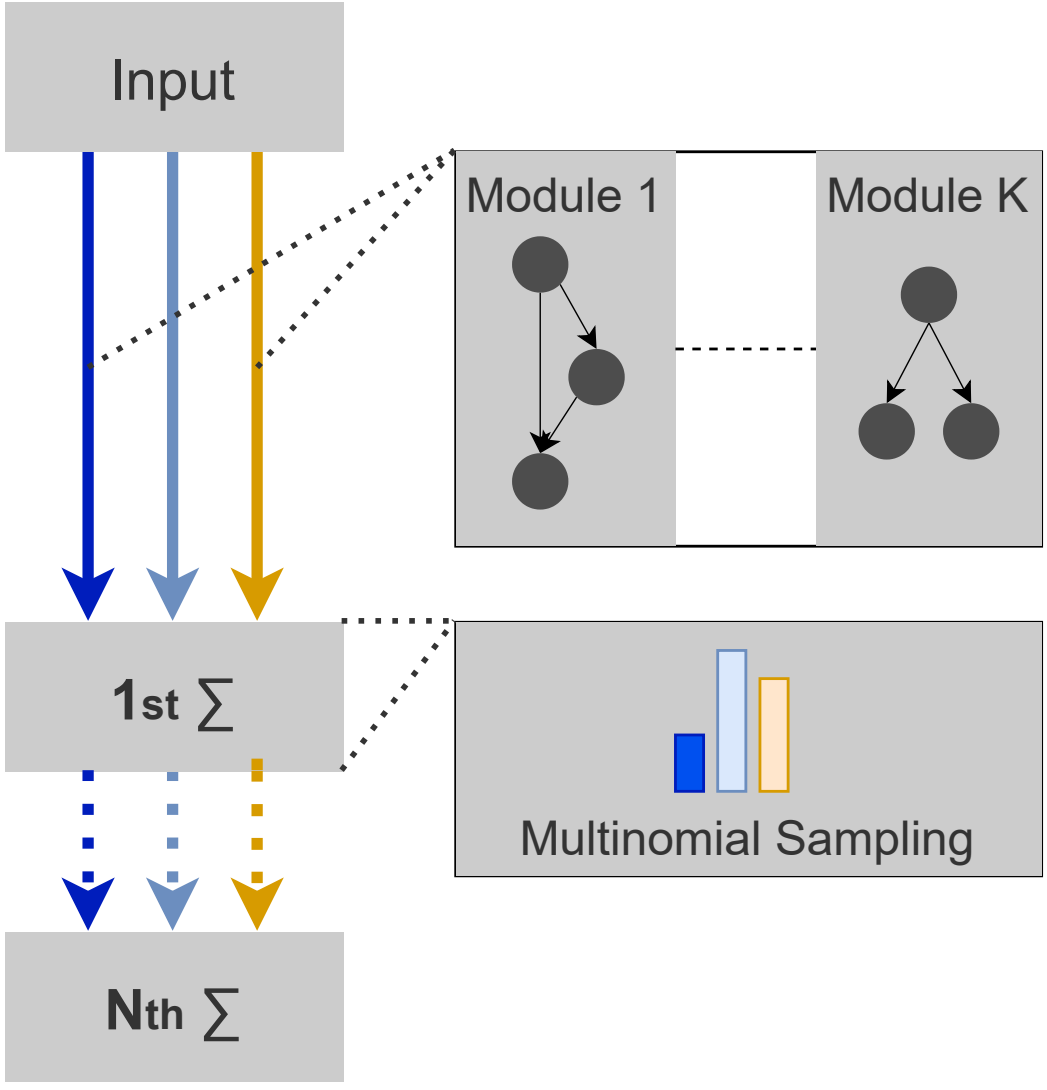
$$g : \mathbb{R} \to [0, 1], x \mapsto \exp^{-(x+\exp^{-x})}$$

For $k \in \mathbb{N}, G \sim P_g^k$ and architecture parameters $a \in \mathbb{R}^k$ it holds:

$$\text{Softmax}(a + G, 0) \sim \text{Multinomial}(1, \text{Softmax}(a))$$

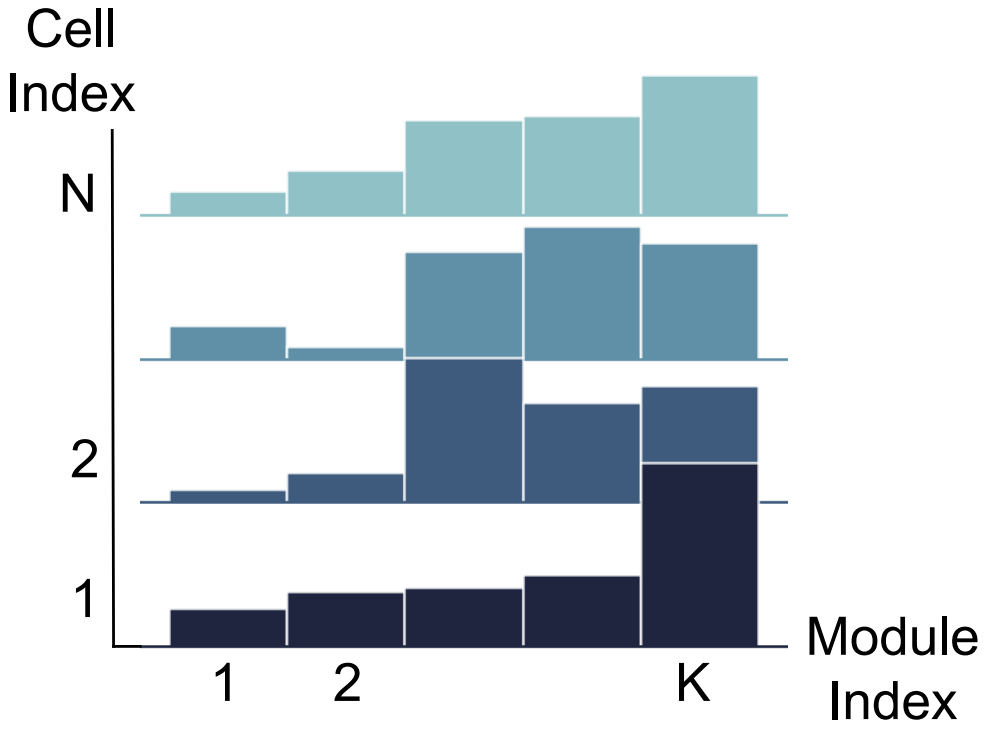
Overview

Search Space



Relative Surrogate

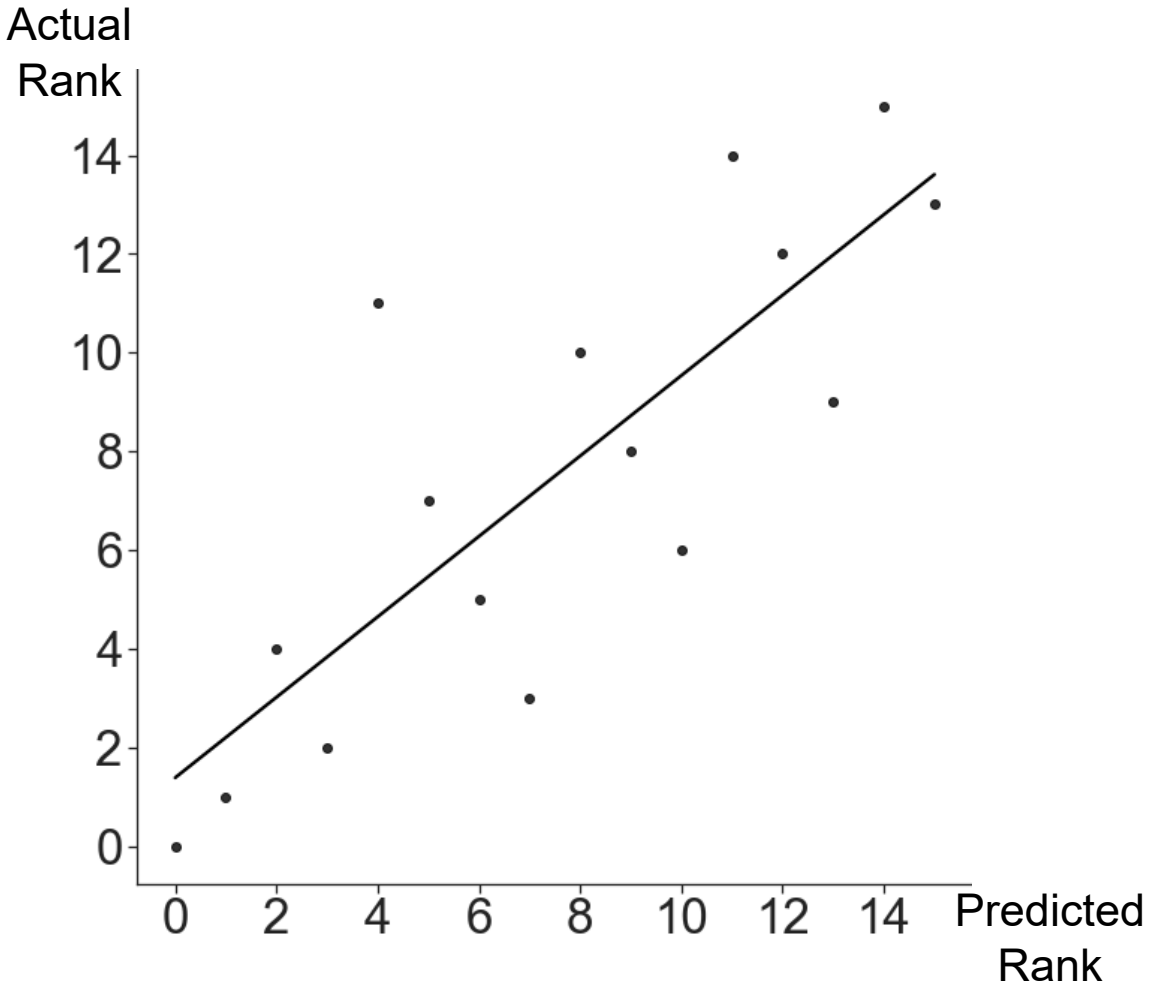
Joint trained multinomials induce ranking on search space



Sampling probability per module per cell

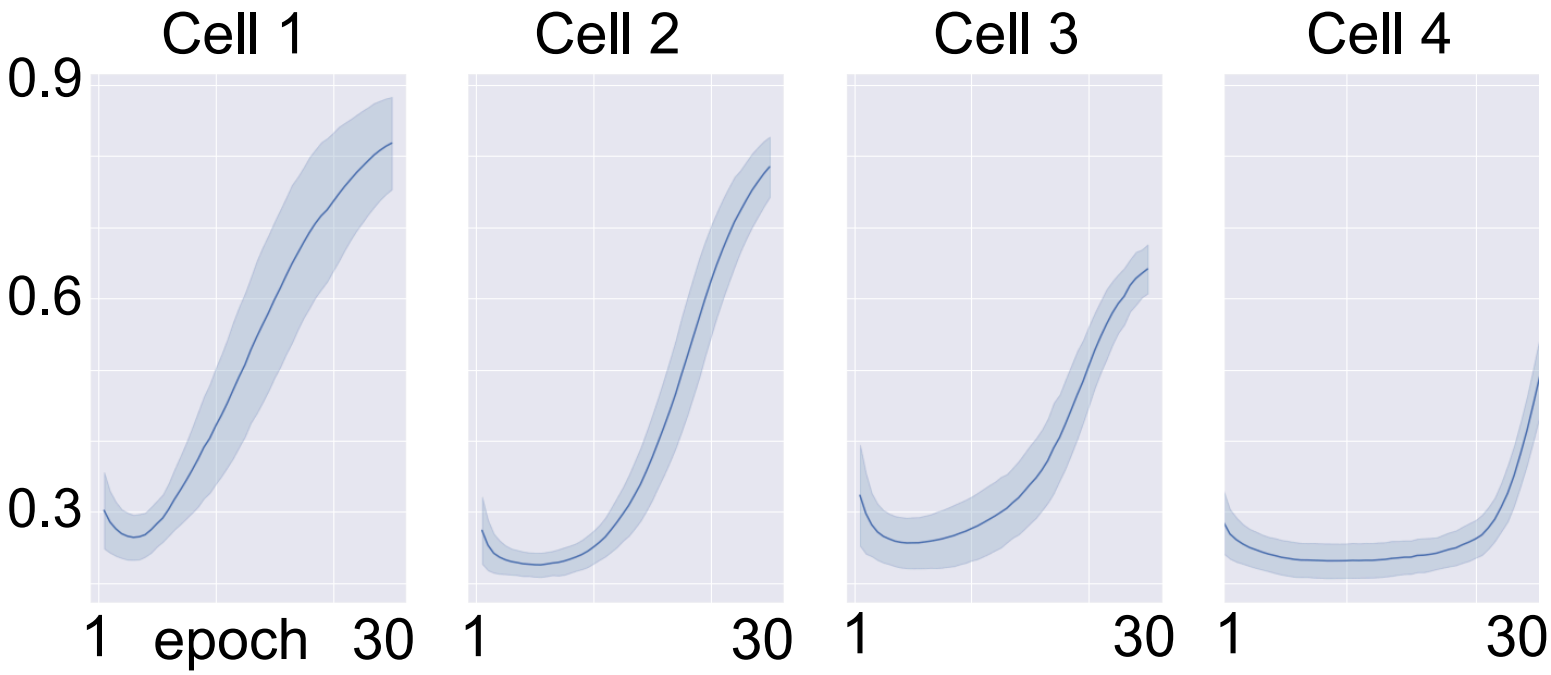
Relative Surrogate

Validate surrogate ranking on actual architecture performances



Architecture Regularization

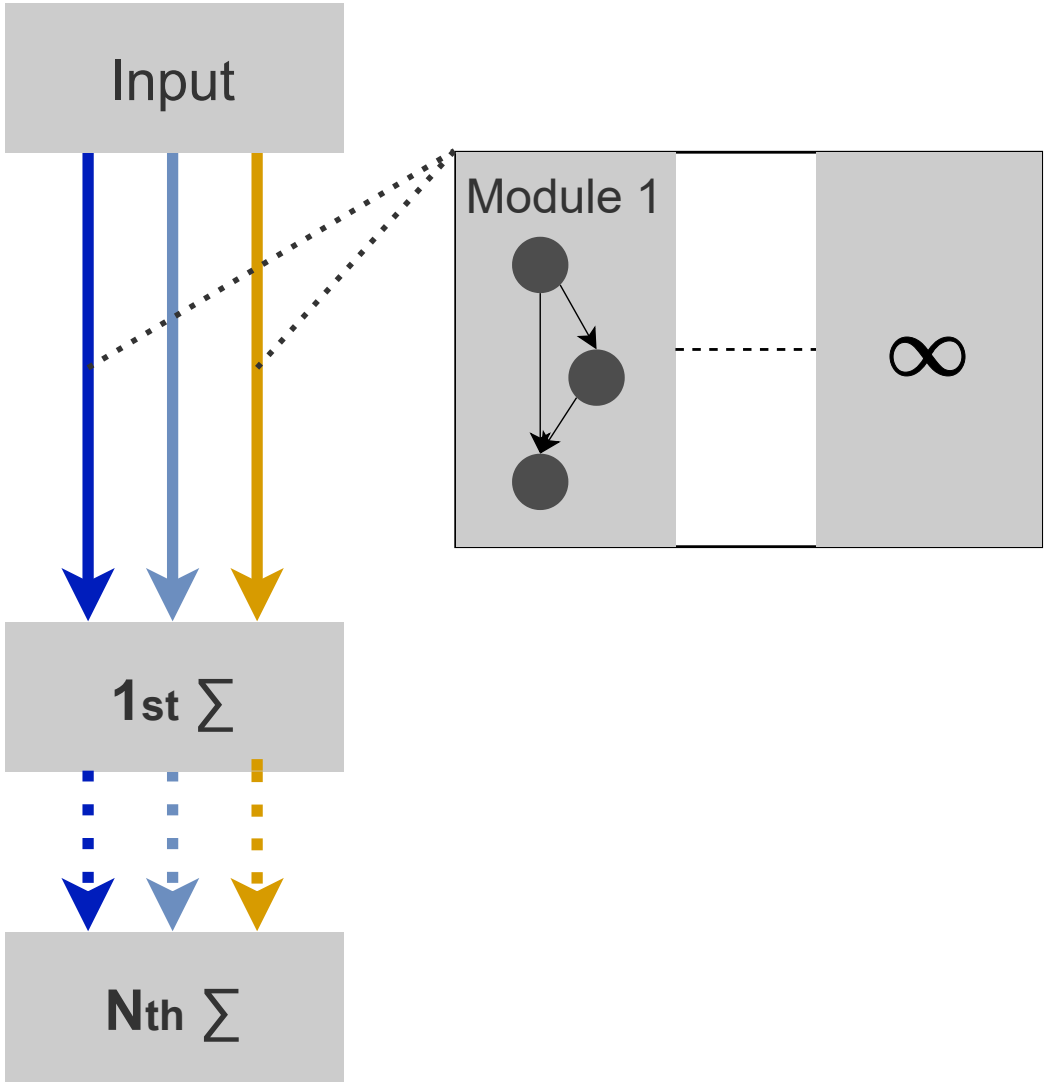
Control speed of convergence dependent on cell index



Maximum norm of architecture parameter vector per epoch per cell

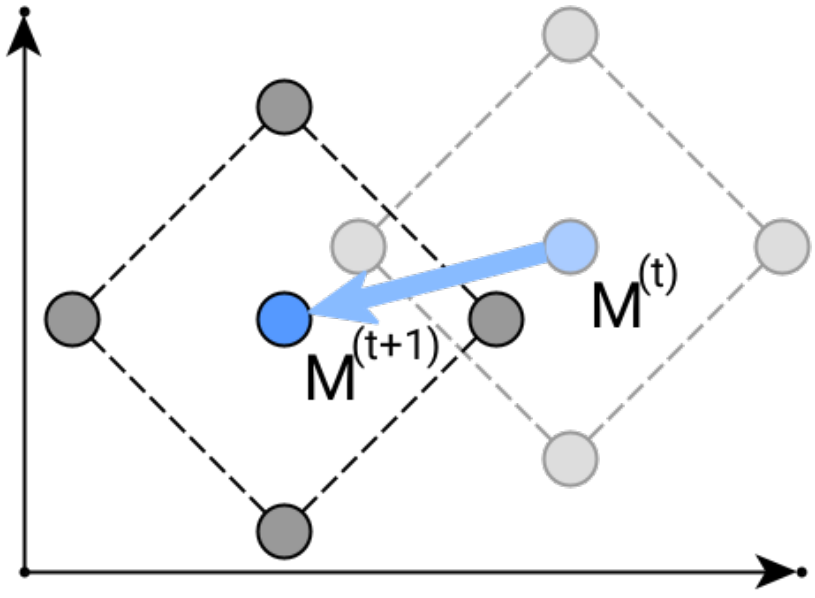
Overview

Search Space Extension



Finite Difference Descent

Finite difference descent on pseudo environment in euclidean search space



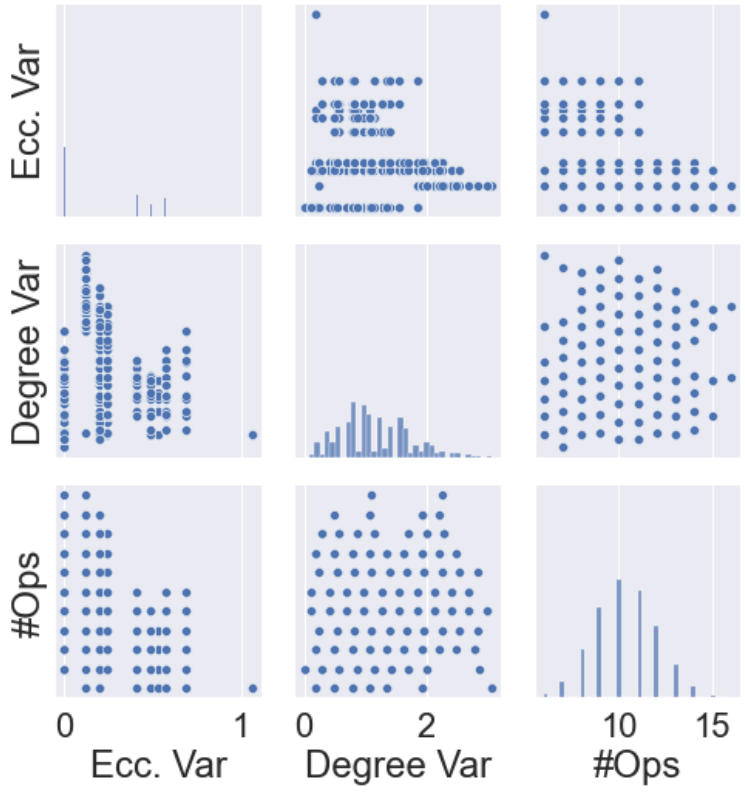
2-dim euclidean search space

Experimental Search Space

We model architectures with directed acyclic graphs (DAG)

Experimental Search Space

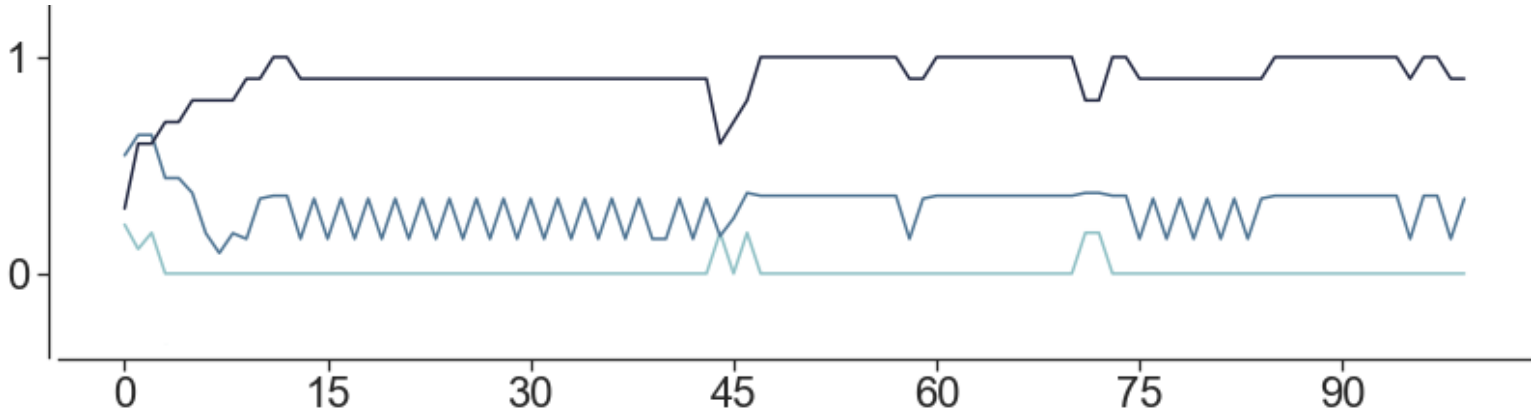
We model architectures with directed acyclic graphs (DAG)



Eccentricity variance, degree variance and # edges for 6-vertice DAGs

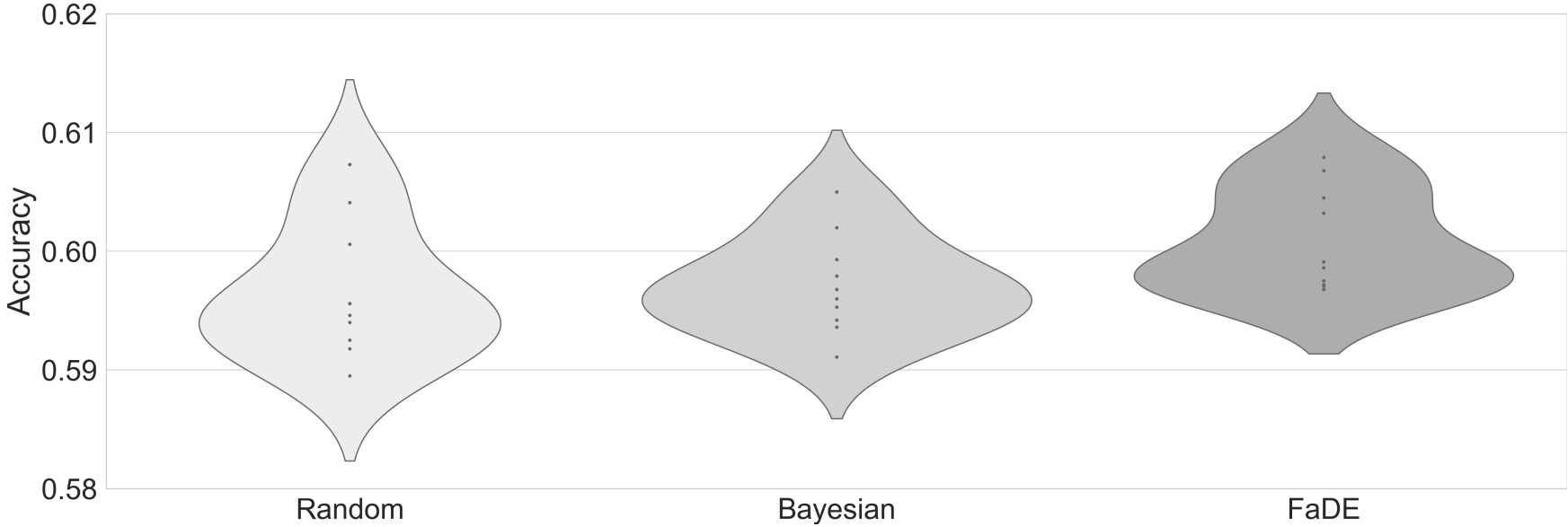
Experiment Results

Search space trajectories (per dimension) for one exemplary cell over 100 epochs of finite difference descent



Experiment Results

Comparing performance of top 10 architectures found by Random Search, Bayesian Search and our approach



References

[Liu et al., 2018] Liu, H., Simonyan, K., and Yang, Y. (2018). Darts: Differentiable architecture search.