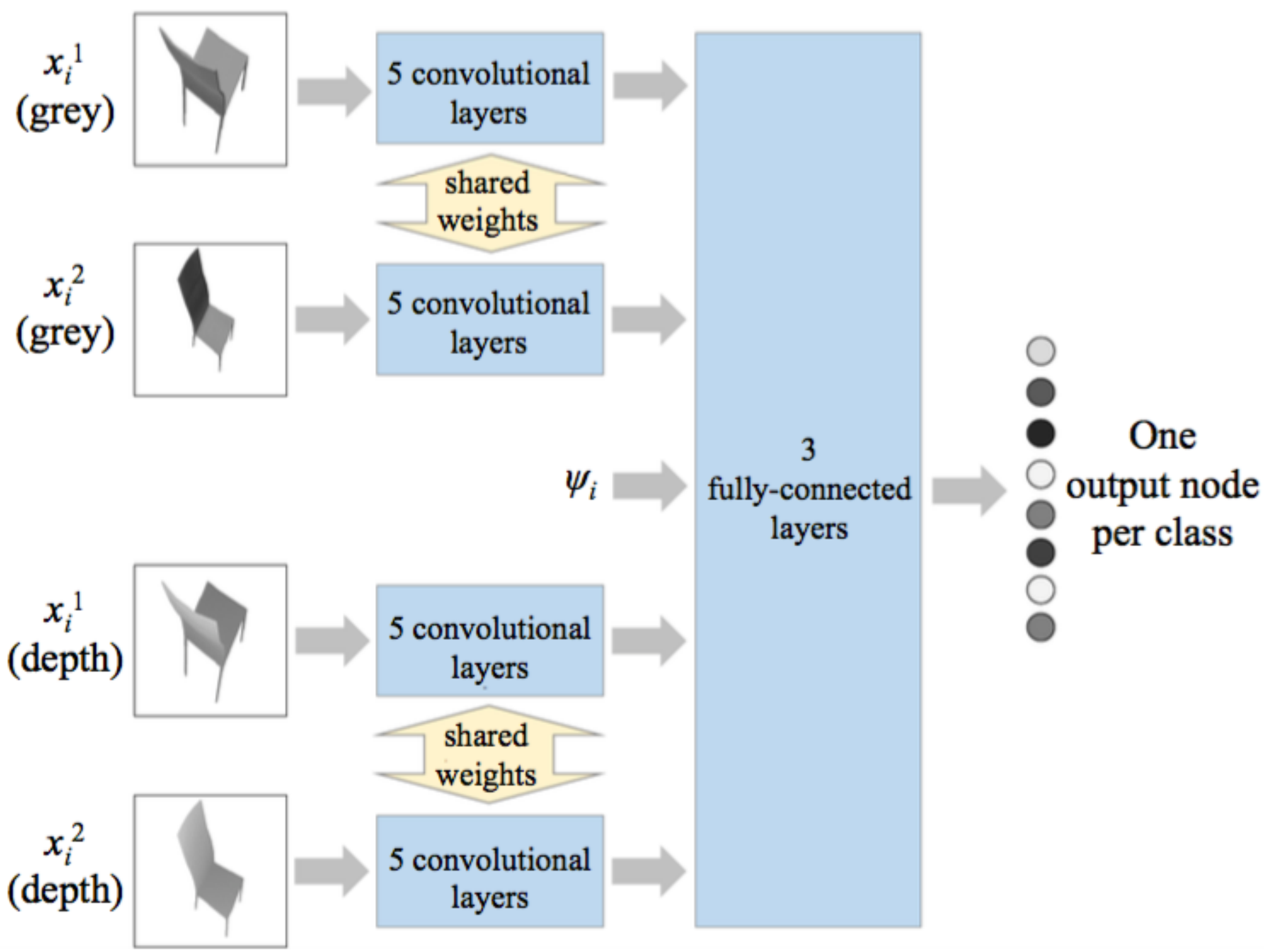
PAIRWISE DECOMPOSITION OF IMAGE SEQUENCES FOR ACTIVE MULTI-VIEW RECOGNITION(EXPERIMENT)

Dongguang You

RECAP

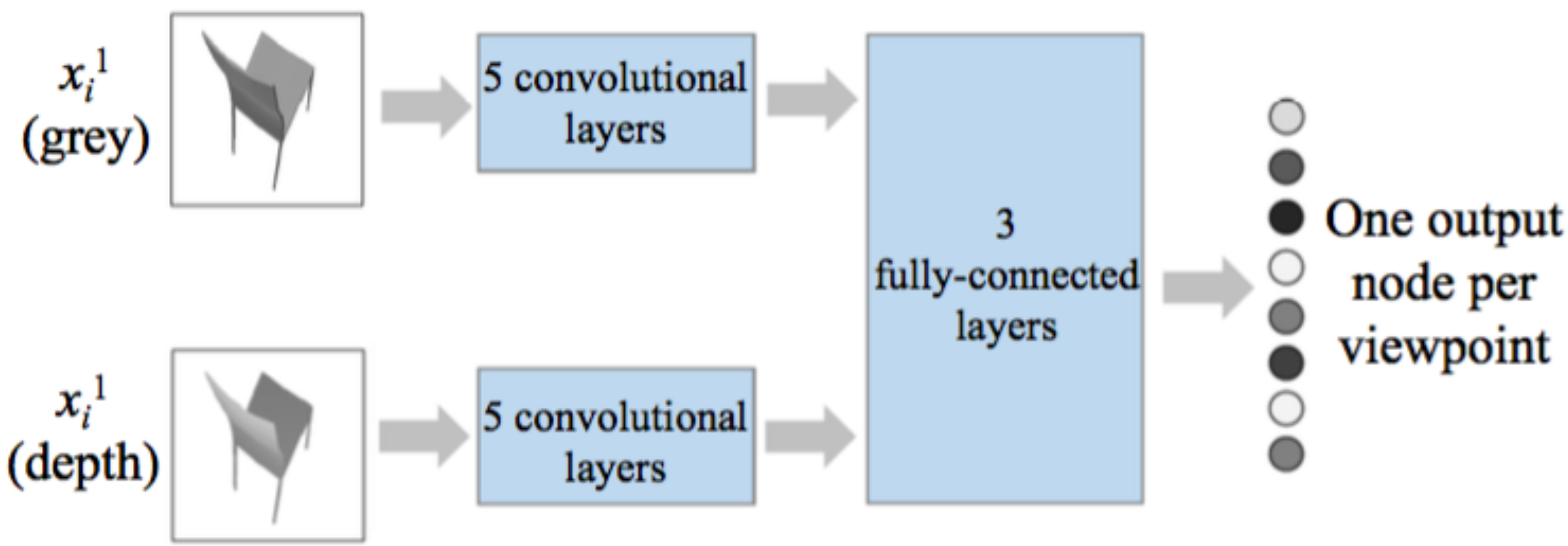
- Pairwise Classification



$$f(\mathbf{y}|w_1 ... w_N) = \sum_{i=1}^{i=N} \lambda_i \, p(\mathbf{y}|w_i) \, . \tag{1}$$

RECAP

- Pairwise Classification
- Next Best View selection/Trajectory Optimisation



TRAJECTORY OPTIMISATION

- Goal: maximize

$$\sum_{i,j \in Sequence} predictedCrossEntropy(i,j)$$

- At each step: find a trajectory that maximizes

$$\sum_{i \in Observed, j \in unobserved} predictedCrossEntropy(i,j)$$

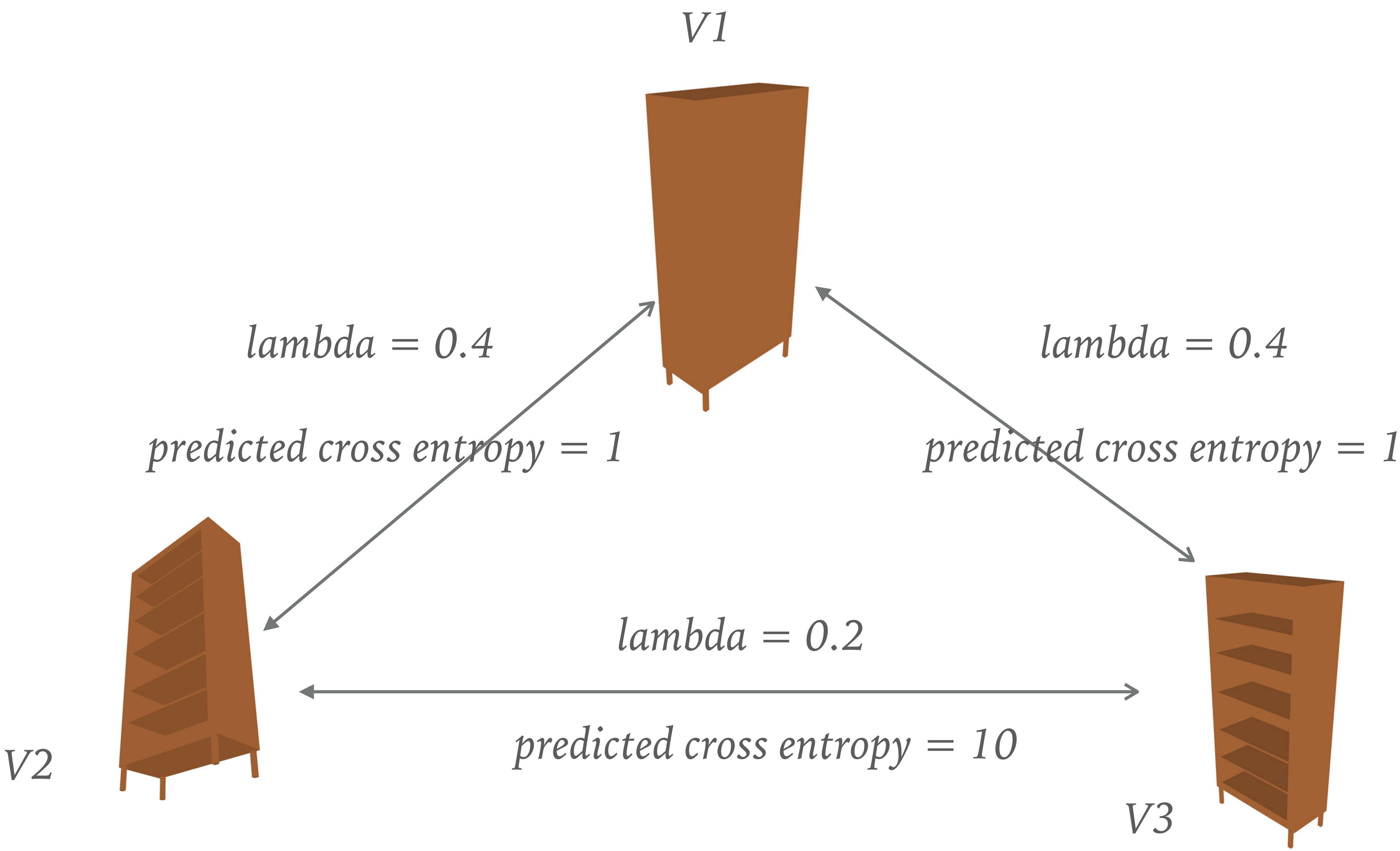
MOTIVATION

- Recall lambda in $$f(\mathbf{y}|w_1 \dots w_N) = \sum_{i=1}^{i=N} \lambda_i \, p(\mathbf{y}|w_i) \,. \qquad (1)$$
- lambda only depends on the relative pose

Failure case:

- Predicted cross entropy of pairs in two trajectories: [1, 10, 1] and [3, 3, 3]
- Choose [1, 10, 1] over [3, 3, 3]
- Lambda for the three pairs in [1, 10, 1]: 0.4, 0.2, 0.4
- Sadly a small weight is assigned to the critical pair during classification

Failure case



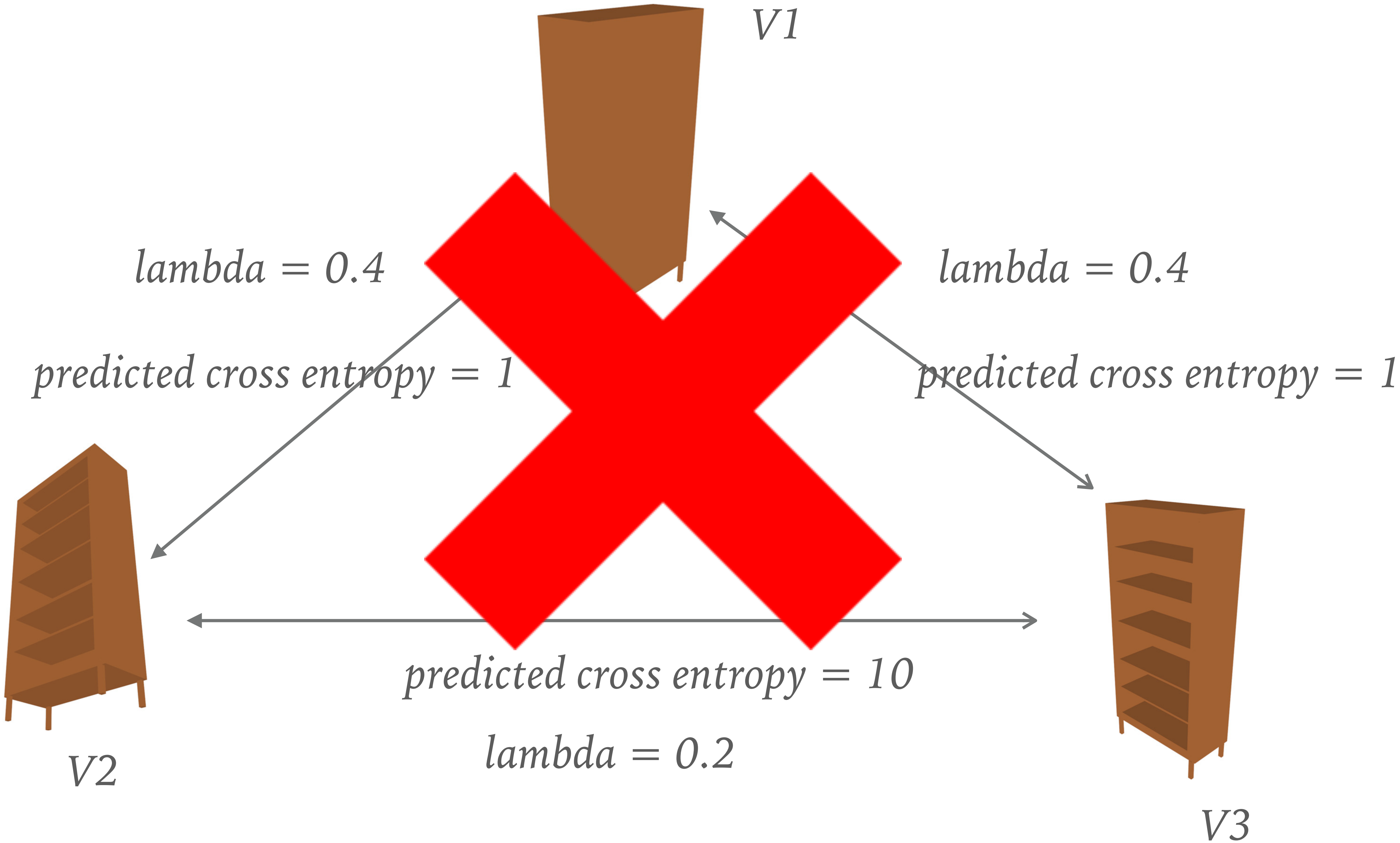
MOTIVATION

- Problem: lambda and predicted cross entropy may conflict

- Solution1: incorporate lambda into trajectory optimisation

$$\sum_{i \in Observed, j \in unobserved} \lambda(i,j) * predictedCrossEntropy(i,j)$$

  - choose [3,3,3] over [1,10,1] given lambda = [0.4,0.2,0.4]



$$\sum_{i \in Observed, j \in unobserved} \lambda(i,j) * predictedCrossEntropy(i,j)$$

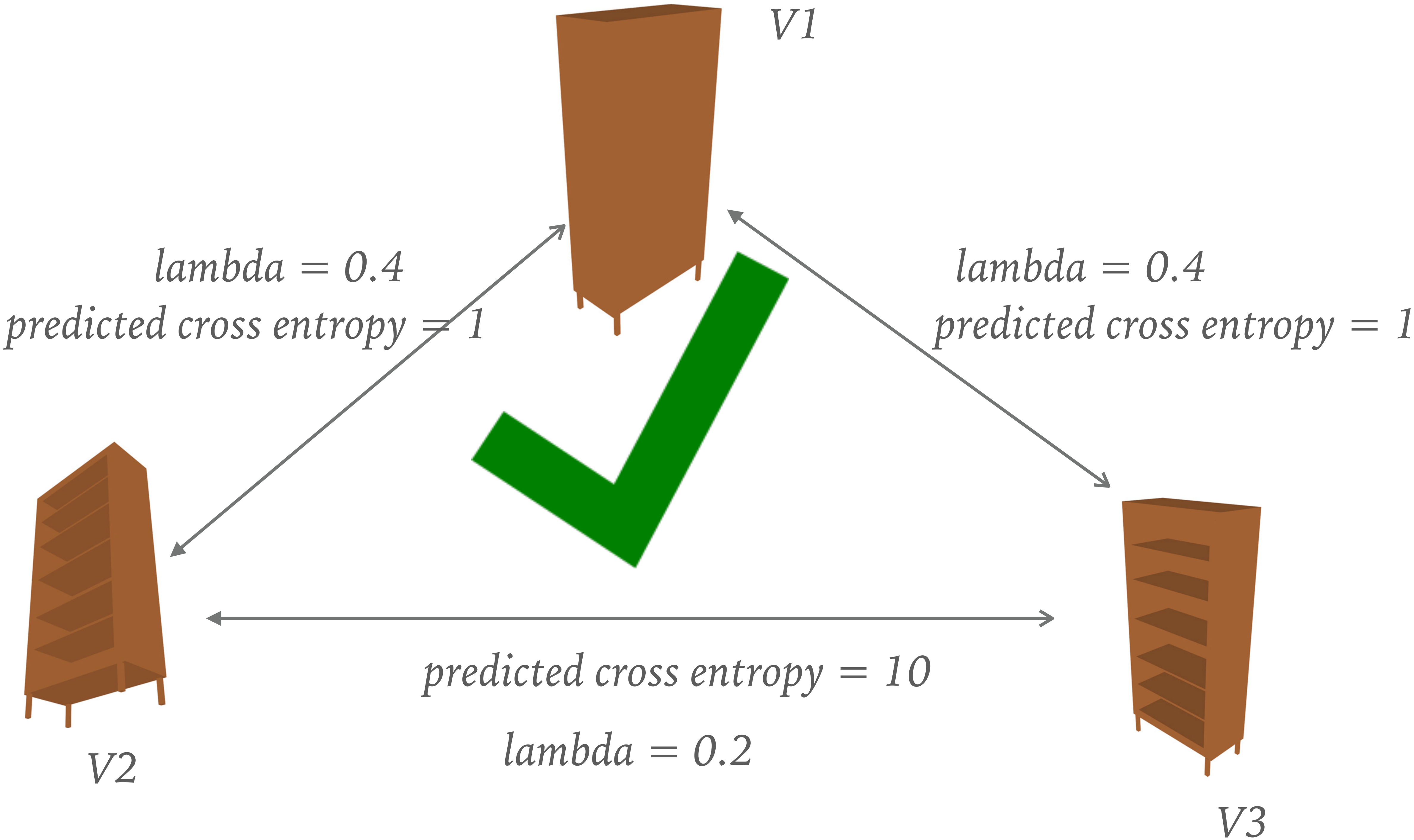
MOTIVATION

- Problem: lambda and predicted cross entropy conflict

- Solution2: replace lambda with predicted cross entropy

$$f(y|w_1 ... w_N) = \sum_{i=1}^{i=N} predictedCE(w_i) * p(y|w_i)$$

  - choose [1,10,1] over [3,3,3], and assign a weight = [1,10,1]/12 to the 3 pairs



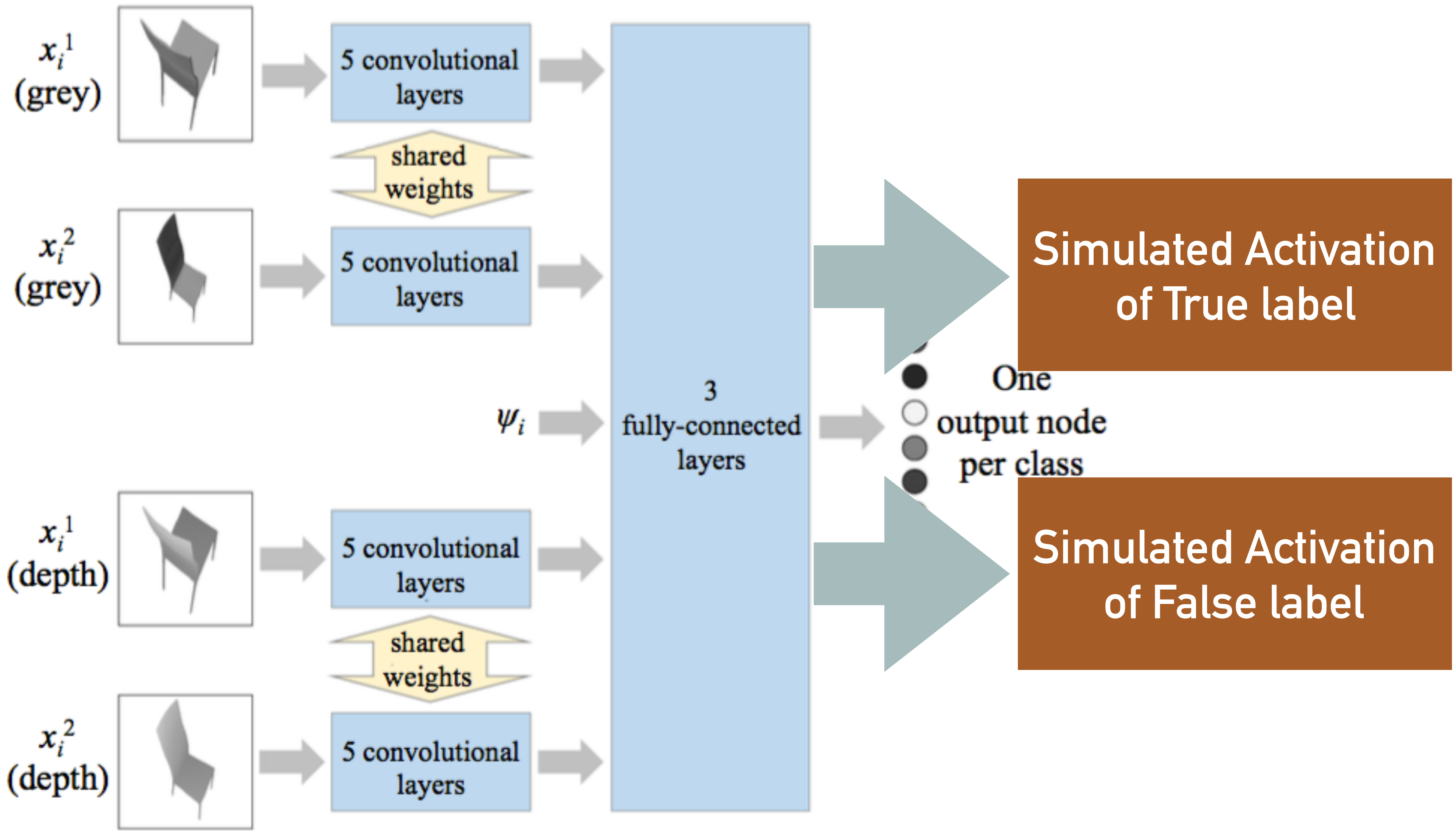
$$f(y|w_1...w_N) = \sum_{i=1}^{i=N} predictedCE(w_i) * p(y|w_i)$$

EXPERIMENT SETUP

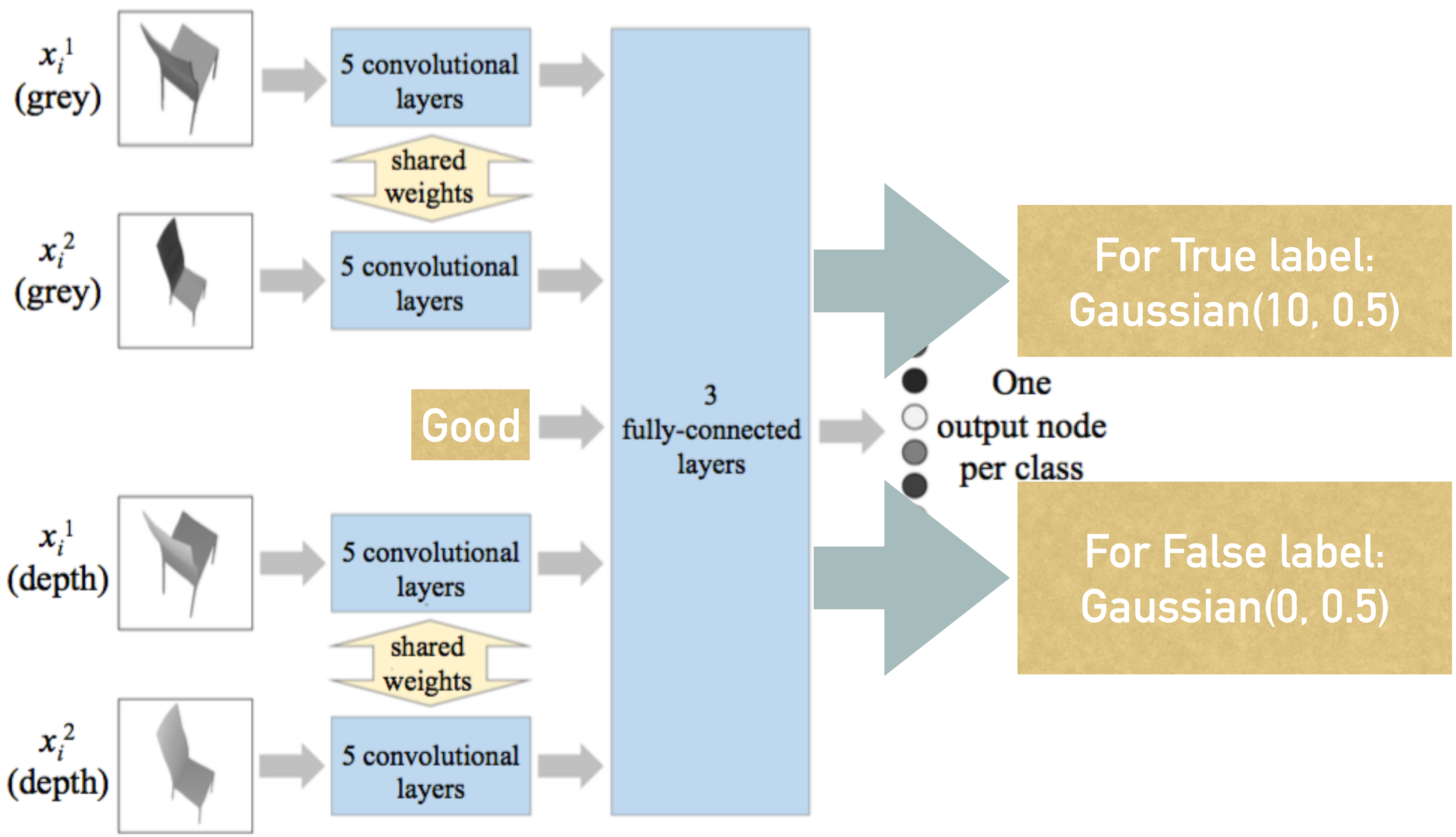
- Simplified setting
  - binary classification
  - relative poses are either good or bad
  - consider testing data of one label

- Simulate the activation of the pairwise classification net
  - assuming the activation follows Gaussian distribution

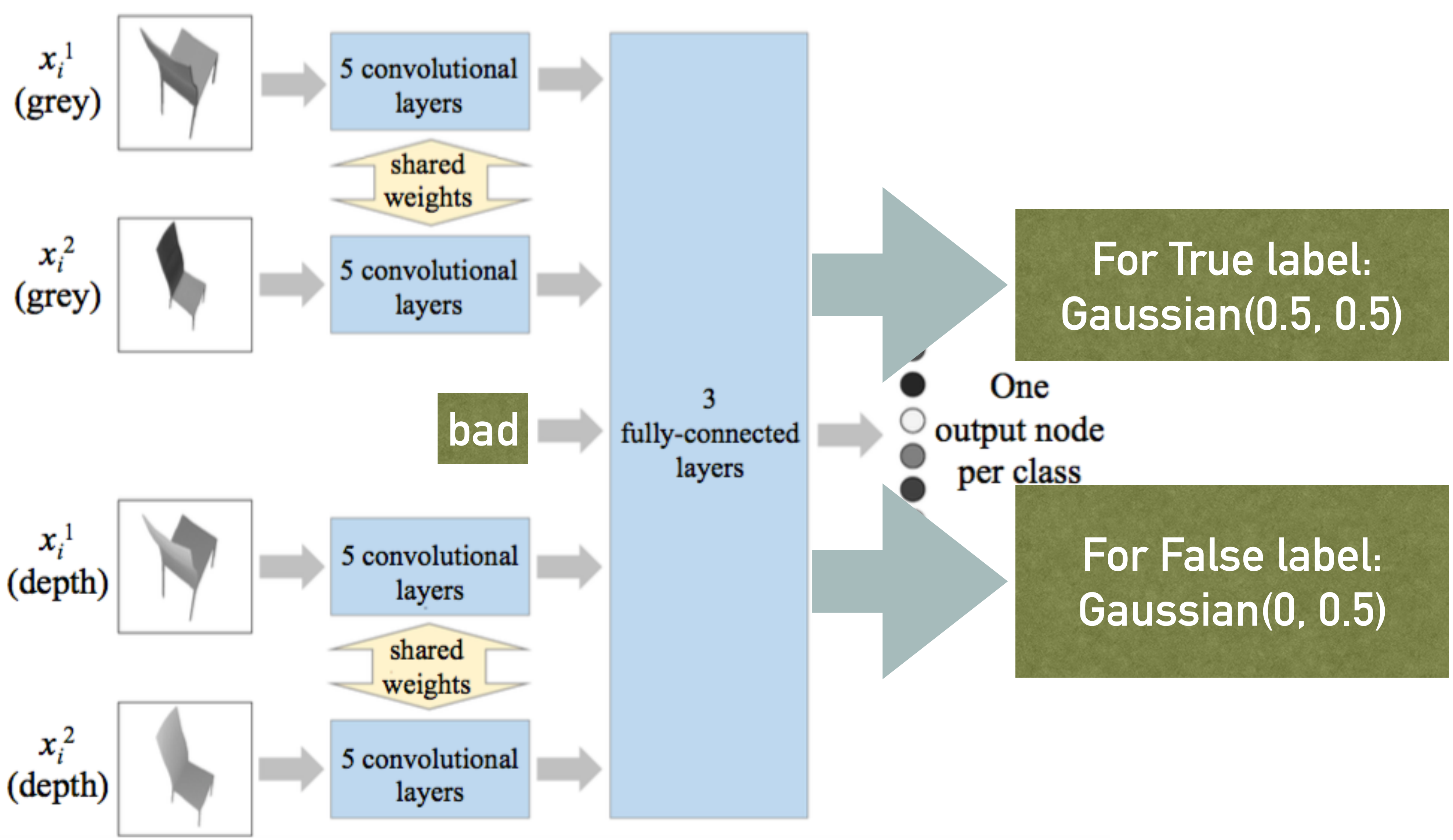
ACTIVATION SIMULATION



Good relative pose



Bad relative pose



RELATIVE POSE SIMULATION

For each test sample

- 4*4 grids of viewpoints
- 120 pairs
- 60 pairs in good relative pose, 60 pairs in bad relative pose

CROSS ENTROPY PREDICTION SIMULATION

- Compute ground-truth cross entropy for each pair

- Predicted cross entropy $\sim$ Gaussian(truth cross entropy, 0.5)

CONVERTING LAMBDA AND CROSS ENTROPY

- lambda and cross entropy are negative

The author didn't make this clear. He pick the pairs that are good by maximising the cross-entropy, so I assume he is using sum(p(x) * log(p'(x))), which is nonpositive

- converted lambda = lambda - min(lambda) - max(lambda)
  - [-1.5, -1] -> [1, 1.5]
  - [-2, -1.2 , -0.6] -> [0.6, 1.4, 2]

- Same for cross entropy

EXPERIMENT 1

- Proposed: incorporate lambda into trajectory optimisation

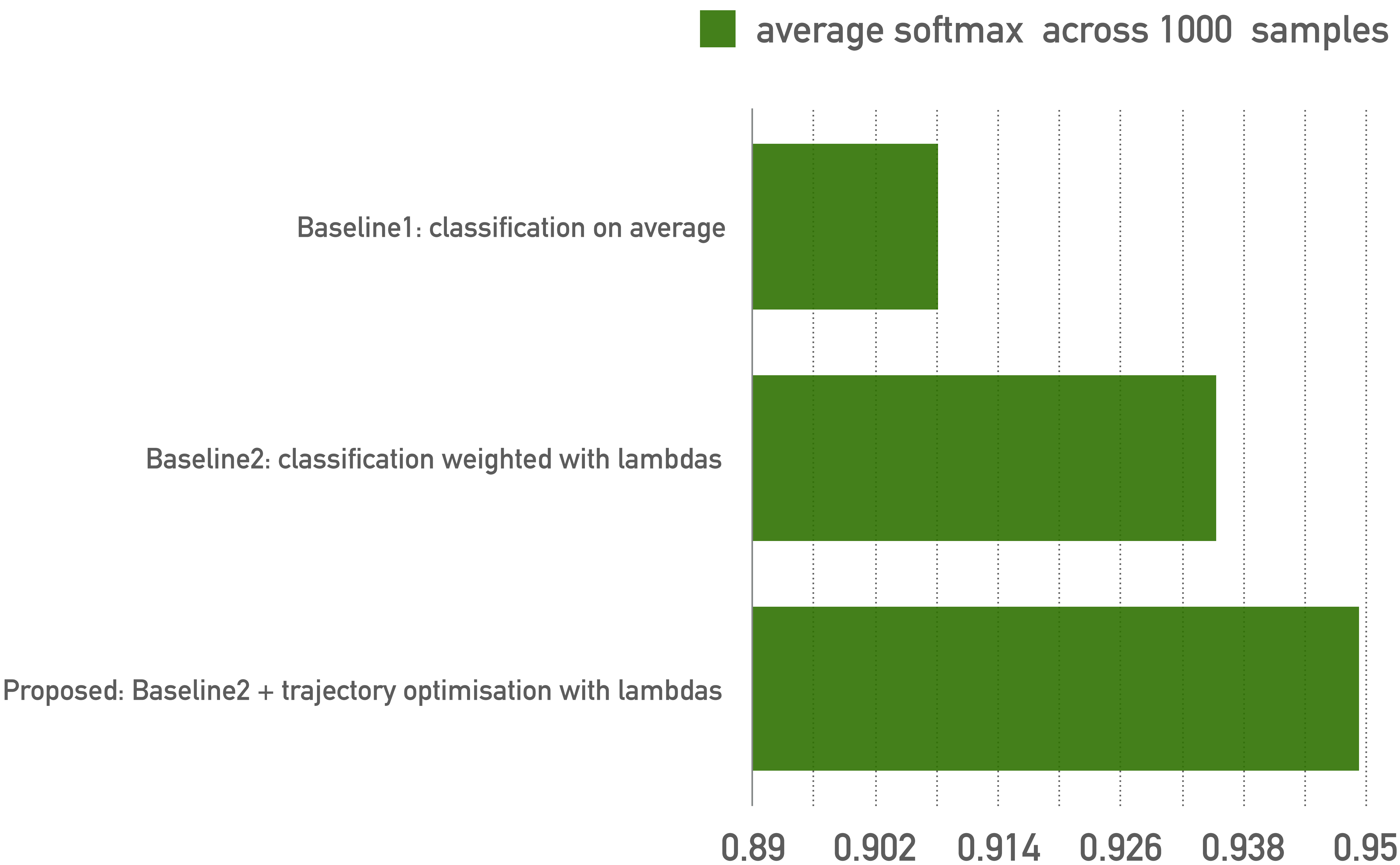
$$\sum_{i \in Observed, j \in unobserved} \lambda(i,j) * predictedCrossEntropy(i,j)$$

- Baselines:

$$\sum_{i \in Observed, j \in unobserved} predictedCrossEntropy(i,j)$$

    - Baseline 1: averaged classification
    - Baseline 2: classification weighted with lambda

RESULT1



average softmax across 1000 samples
Baseline1: classification on average
Baseline2: classification weighted with lambdas
Proposed: Baseline2 + trajectory optimisation with lambdas
0.89
0.902
0.914
0.926
0.938
0.95

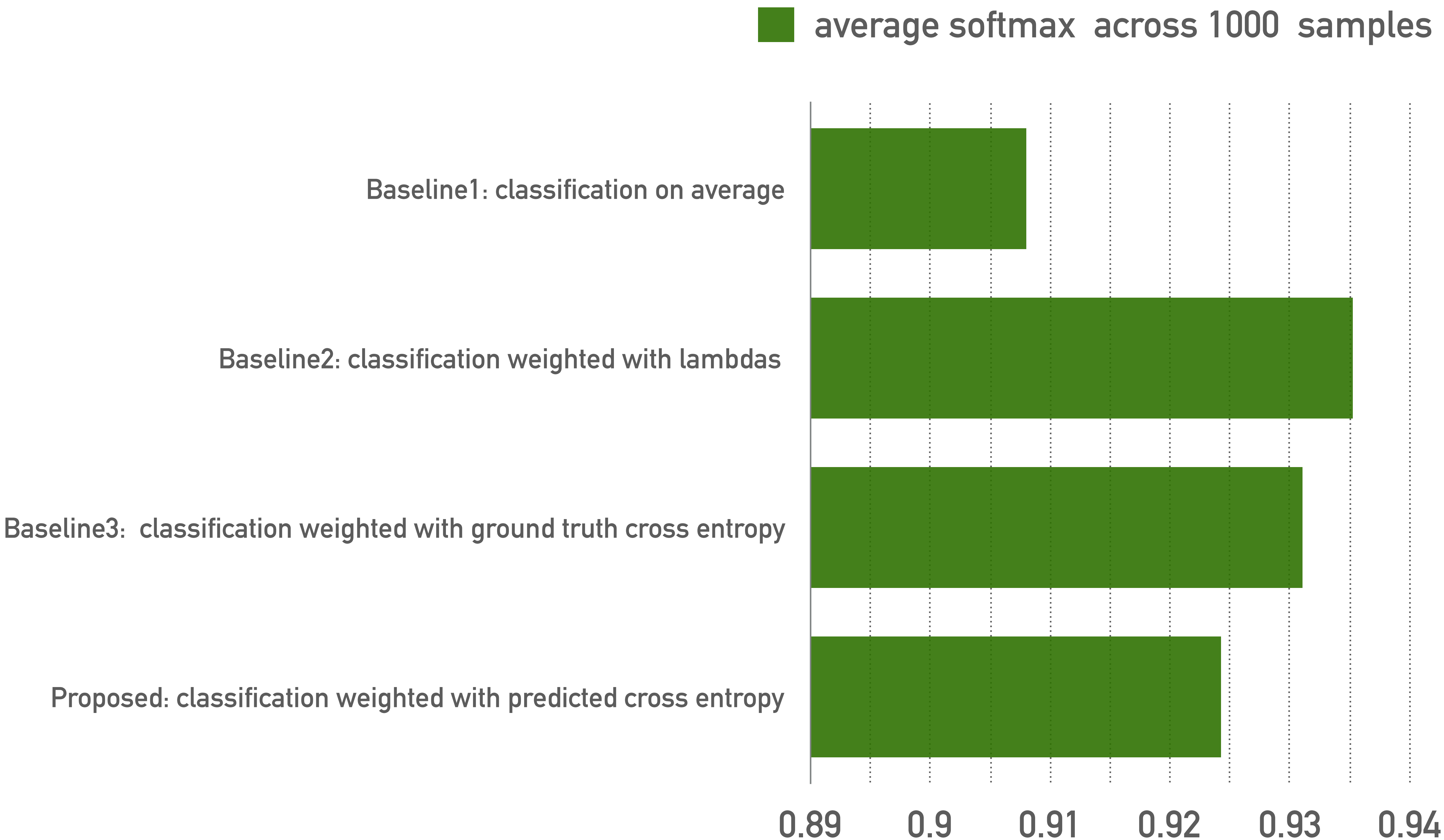
EXPERIMENT2

- Proposed: use the predicted cross entropy as the weight, instead of lambda

$$f(y|w_1...w_N) = \sum_{i=1}^{i=N} predictedCE(w_i) * p(y|w_i)$$

- Baseline 1: averaged classification result
- Baseline 2: classification result weighted with lambda
- Baseline 3: classification result weighted with ground truth cross entropy

RESULT2


average softmax across 1000 samples
Baseline1: classification on average
Baseline2: classification weighted with lambdas
Baseline3: classification weighted with ground truth cross entropy
Proposed: classification weighted with predicted cross entropy
0.89
0.9
0.91
0.92
0.93
0.94

EXPERIMENT2*

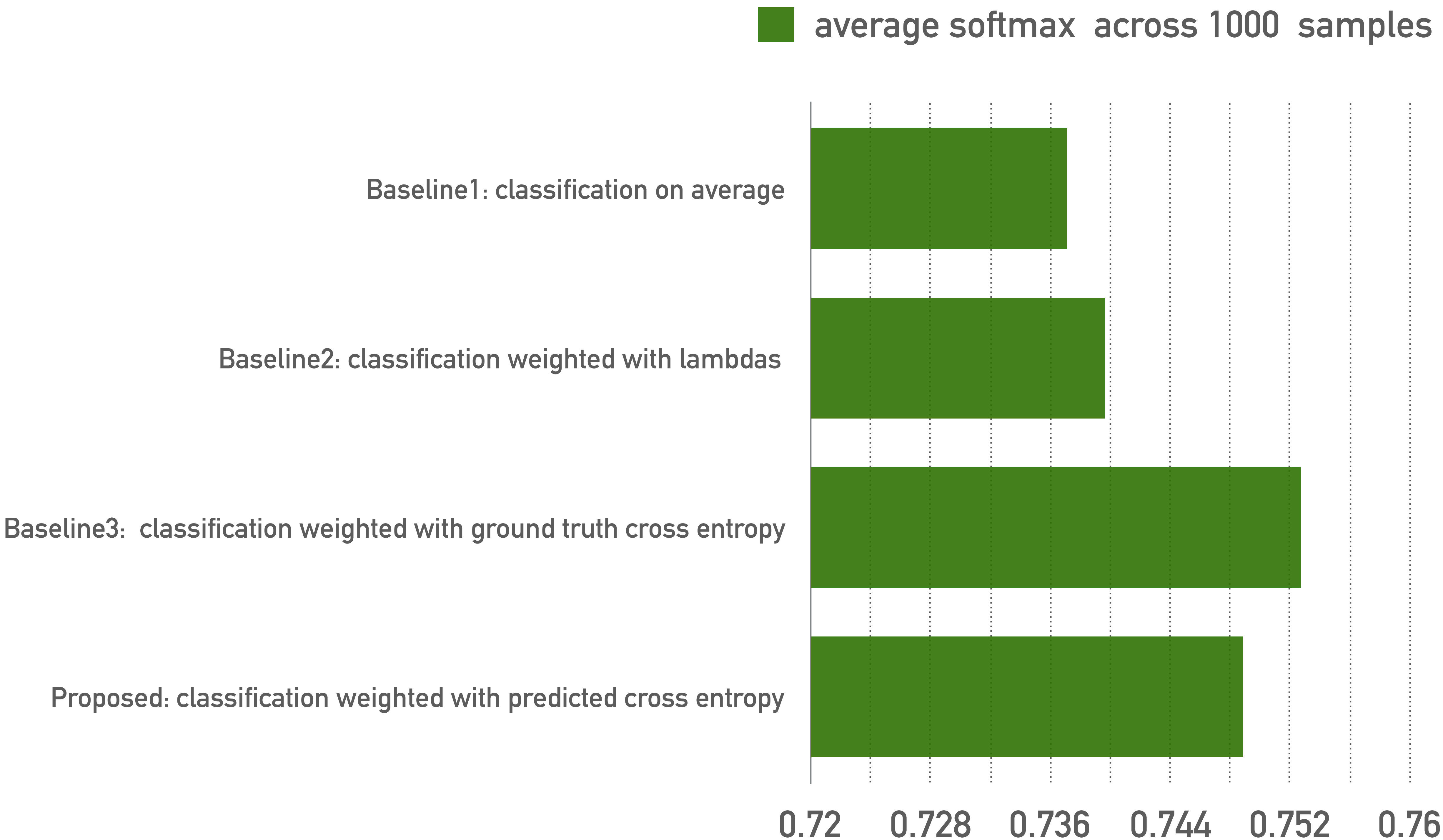
- What if the effect of relative pose is weaker?

The activation of correct label is modified:

- Good relative pose ~ Gaussian(1, 0.5) instead of Gaussian(10, 0.5)
- Bad relative pose ~ Gaussian(0.5,0.5), same as before

- What would the comparisons look like?

RESULT2*


average softmax across 1000 samples
Baseline1: classification on average
Baseline2: classification weighted with lambdas
Baseline3: classification weighted with ground truth cross entropy
Proposed: classification weighted with predicted cross entropy
0.72
0.728
0.736
0.744
0.752
0.76

LIMITATION OF THE PAIRWISE METHOD

- do not have a global view(as compared to “Look ahead before you leap”)

- range of entropy is (-inf, 0), hard to guarantee the accuracy of regression

CONCLUSION

- When the effect of relative pose is strong
  - incorporating lambda into trajectory optimisation might improve the prediction
- When the effect of relative pose is weak
  - predicted cross entropy could be a better choice for weight than lambda