# Face Recognition Using Random Walks on Graphs: Real-Time Learning without Explicit Feedback

Michal Valko (Computer Science Department)

Branislav Kveton, Ling Huang, Daniel Ting

Grad Expo 2010, March 22nd, 2010

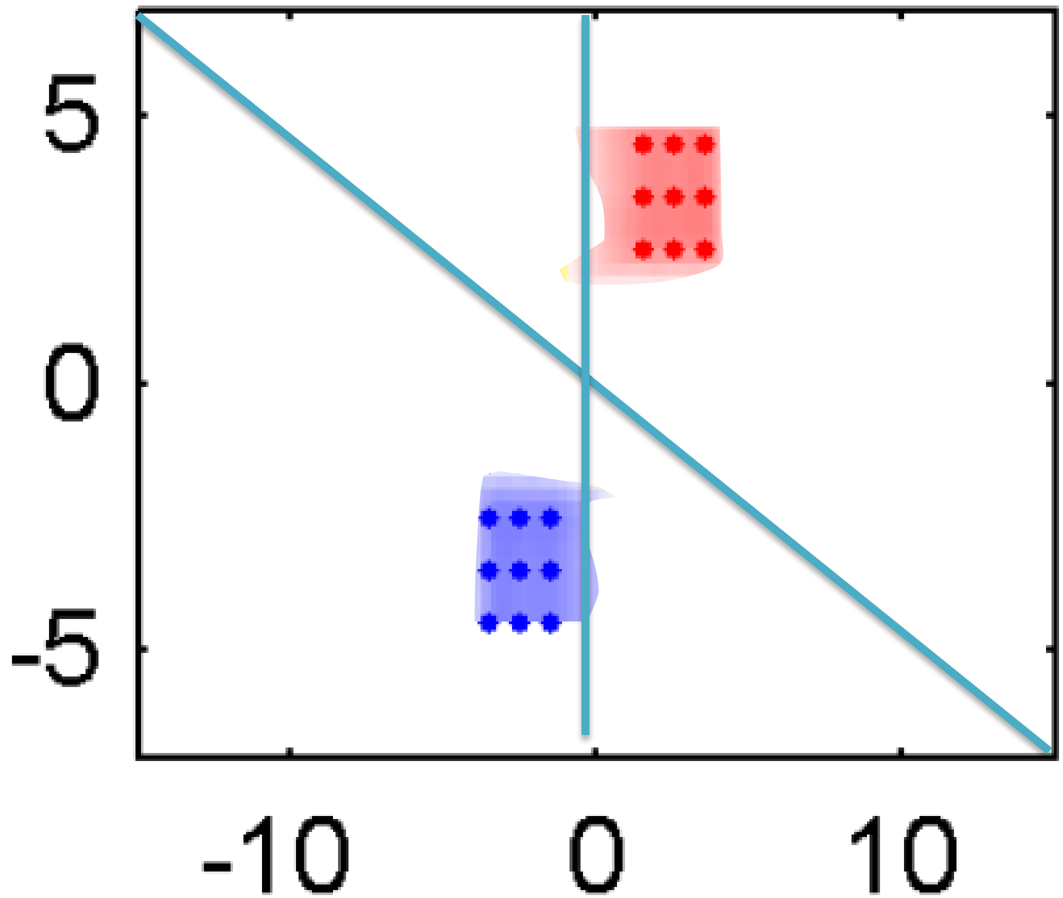
# Main Ideas

- **Goal**:  Adaptation to patterns with minimal human feedback (labels)
  - Most of data around is unlabeled
  - Labeling is expensive

- **Solution**:  Semi – Supervised learning (Machine Learning)
  - Labeled examples are provided in the beginning
    - Provide initial bias
  - Unlabeled examples come as available

- **Approach**:  Regularized graph–based inference + quantization

# Semi-supervised learning

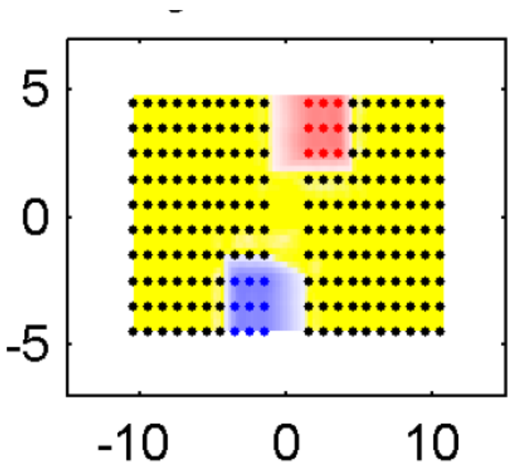


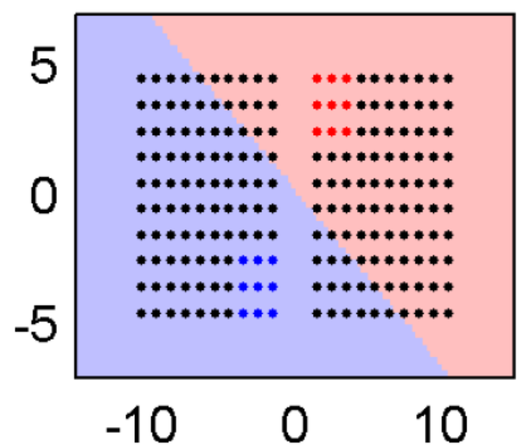
# Semi-supervised learning

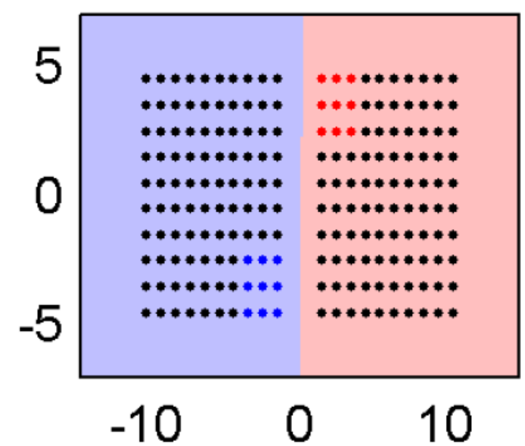
Data

Supervised

Semi-Supervised



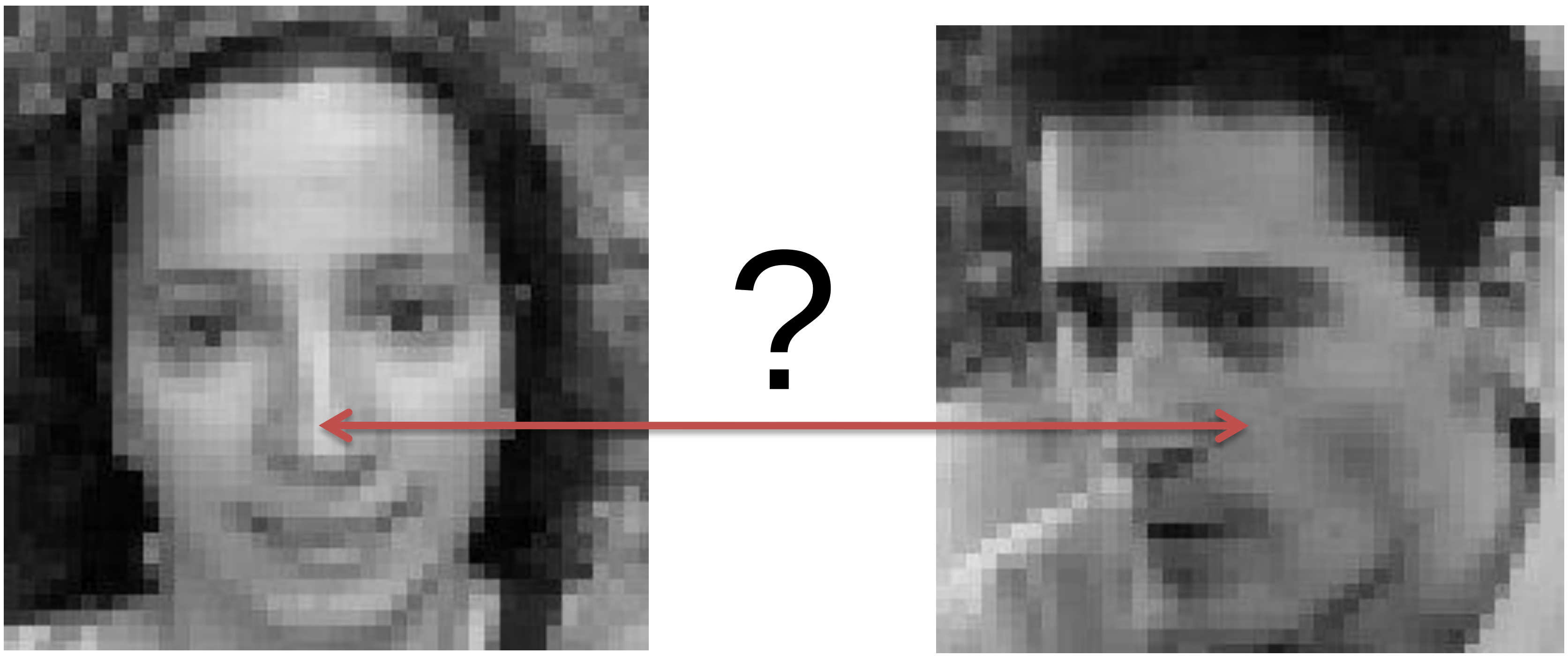
# Face Similarities

?

# Face Similarities

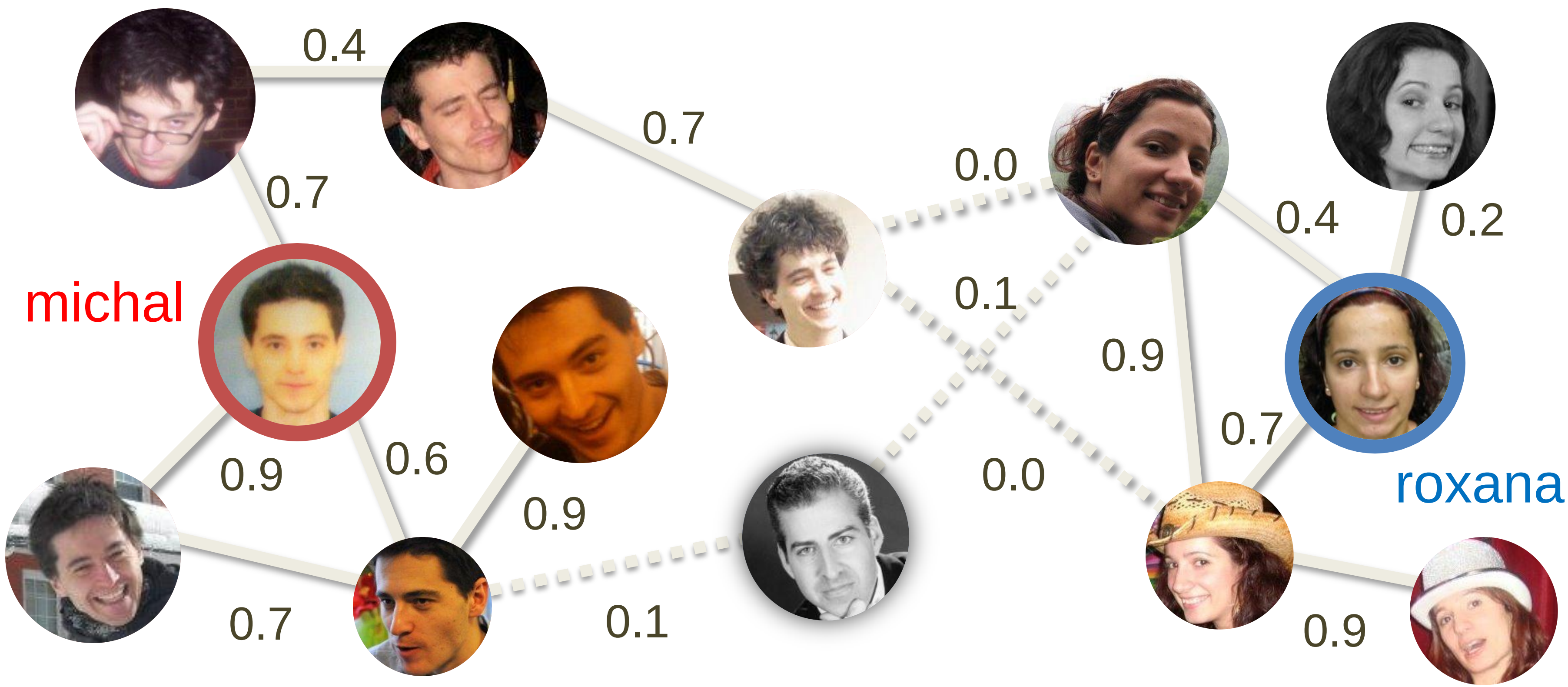
?

# Face Similarities



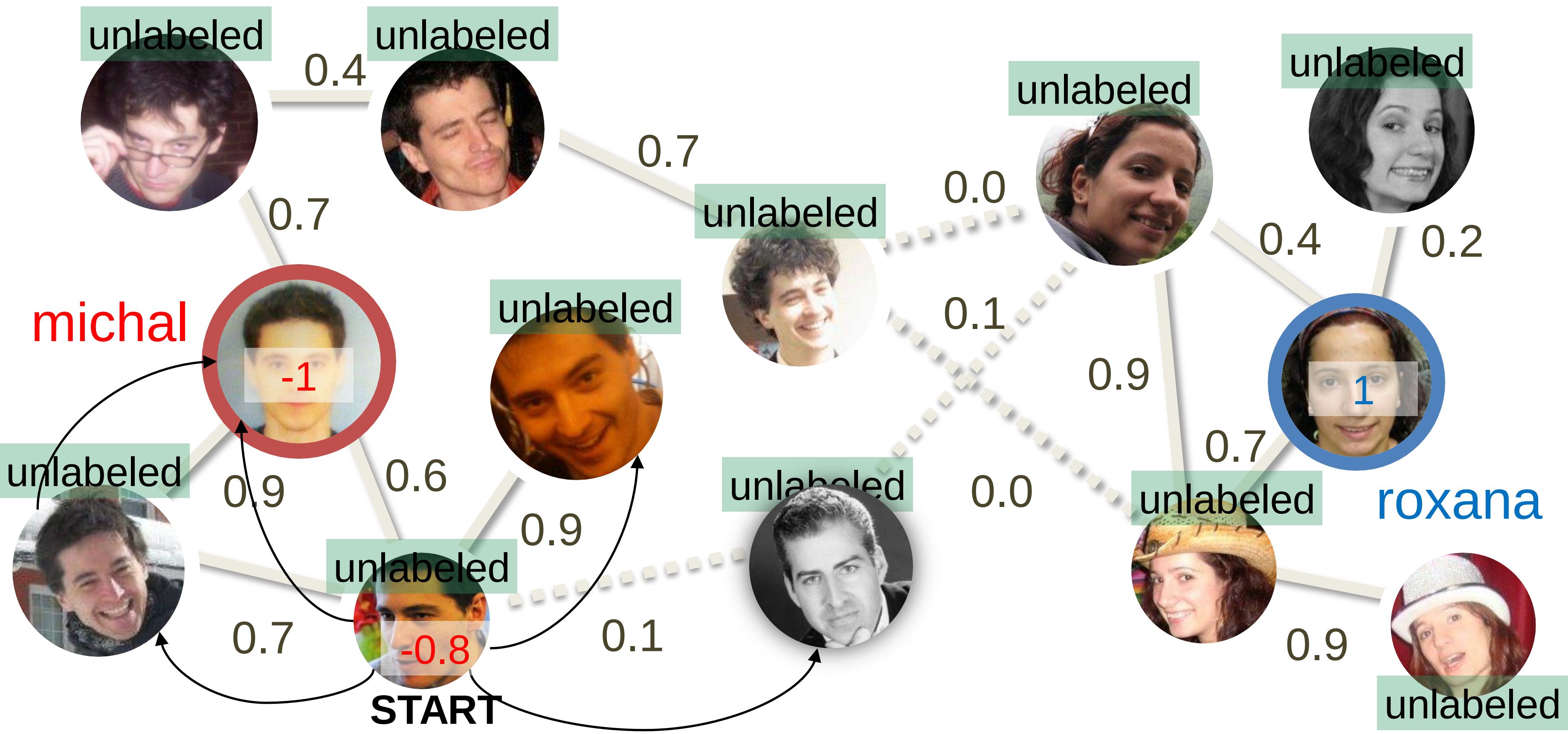

$$d(\mathbf{x}_i, \mathbf{x}_j) = \min \begin{cases} \|\mathbf{x}_i - \mathbf{x}_j\|_{2,\psi}\,, \\ \|(\mathbf{x}_i - \bar{\mathbf{x}}_i) - (\mathbf{x}_j - \bar{\mathbf{x}}_j)\|_{2,\psi} \\ \|\mathbf{x}_i/\bar{\mathbf{x}}_i - \mathbf{x}_j/\bar{\mathbf{x}}_j\|_{2,\psi} \end{cases}$$

# Graph from faces



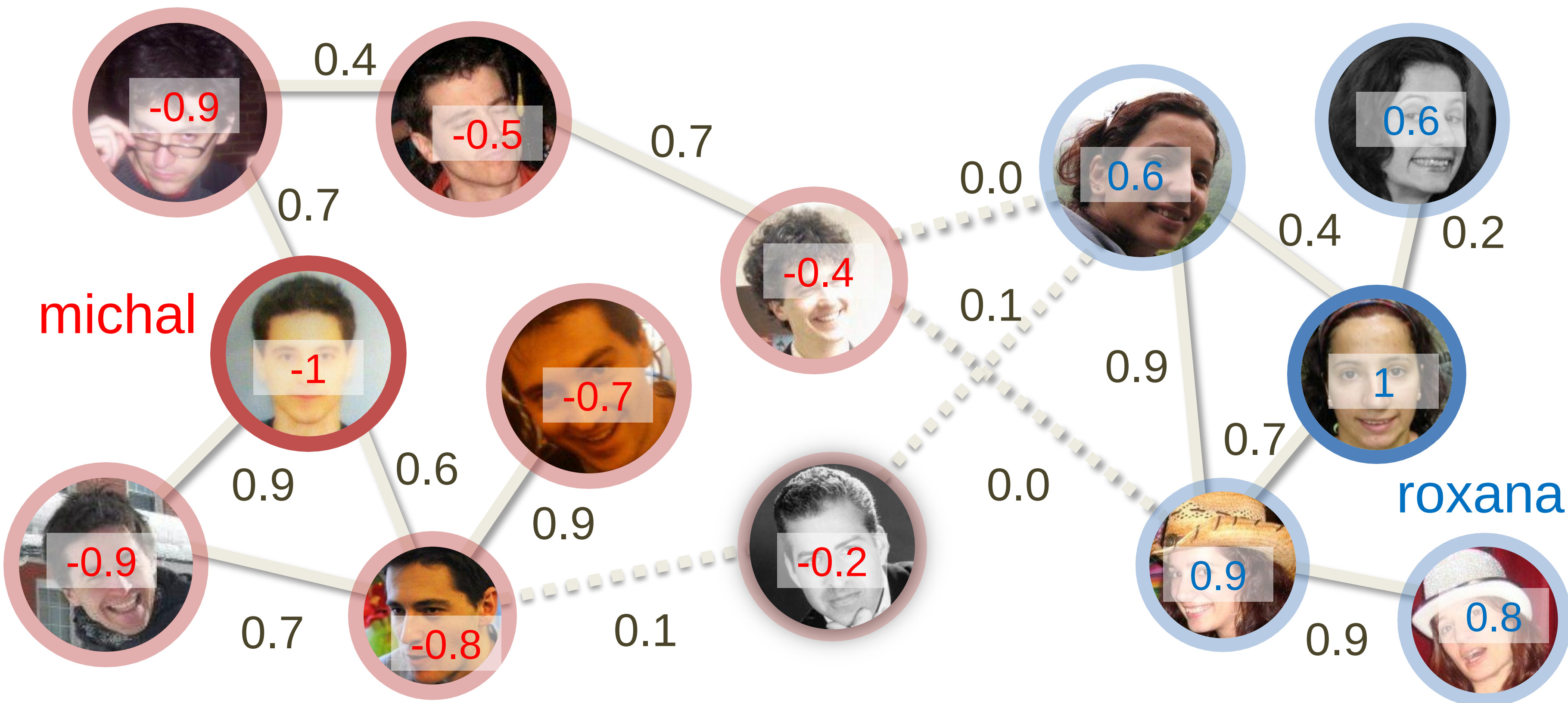
# Graph-based Semi-Supervised Learning

$$\ell^* = \arg\min_{\ell} \; \ell^\top L \ell$$

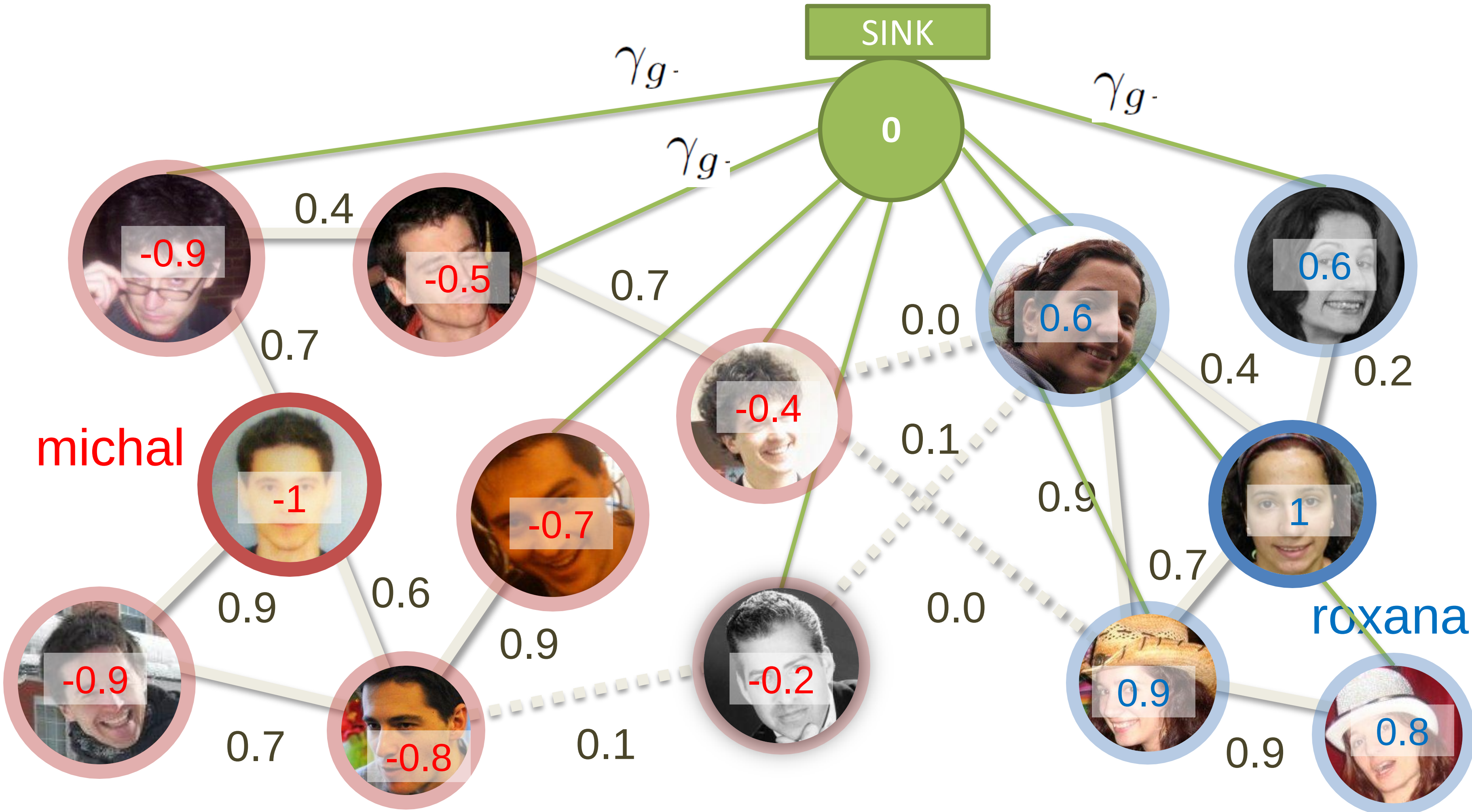


# Harmonic Function Solution (HFS)

- Labels of unlabeled vertices are inferred using the harmonic function solution



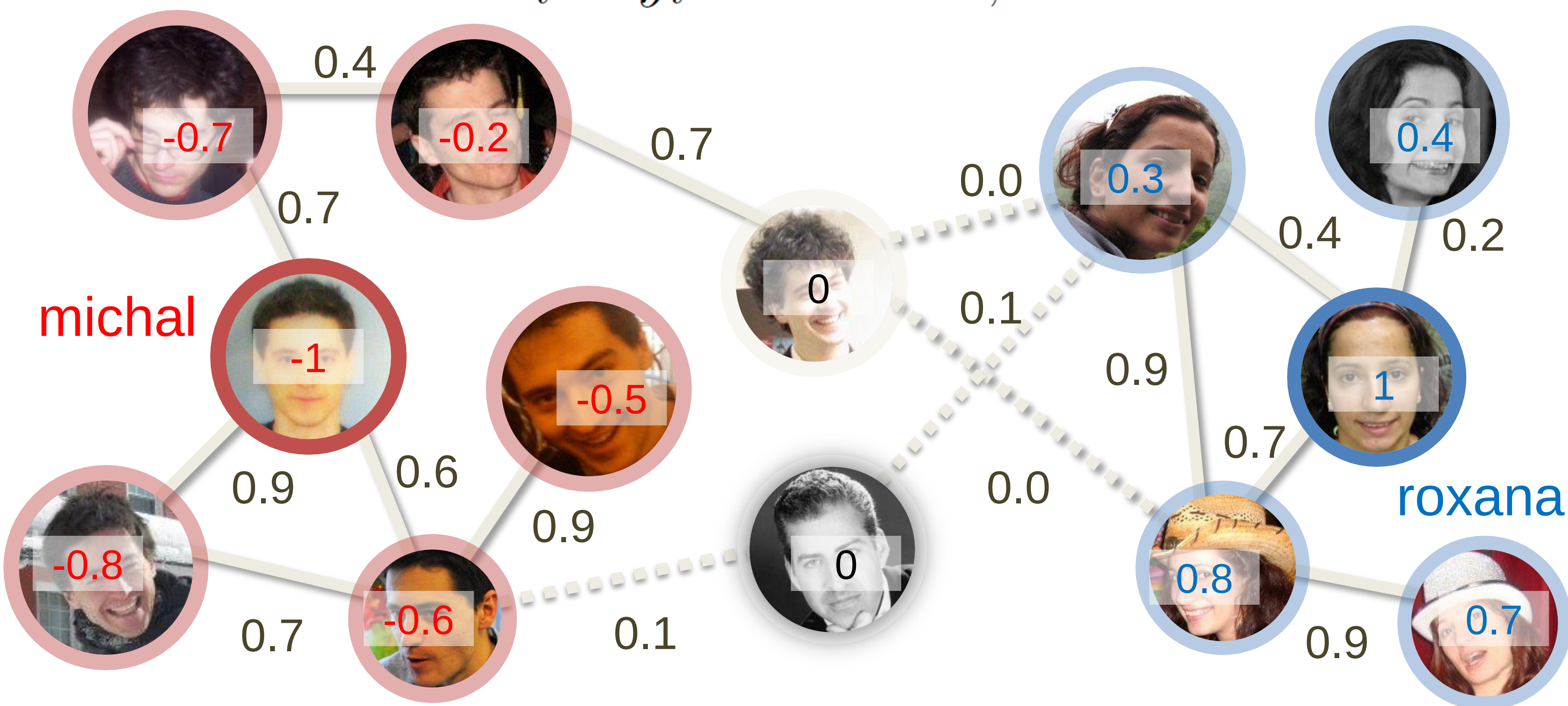
# Dealing with Outliers



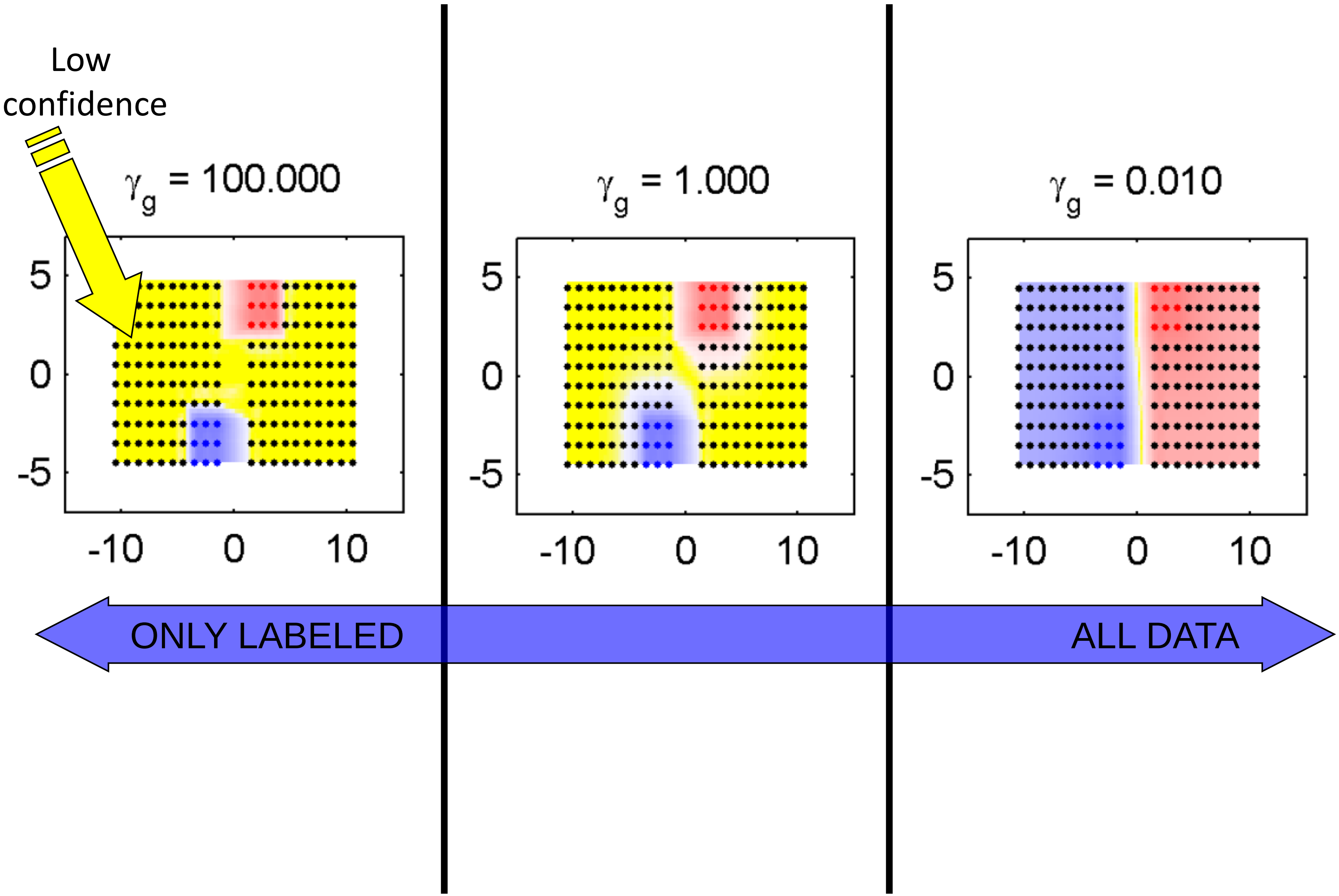
# Regularized HFS

$$\boldsymbol{\ell}^* = \arg\min_{\boldsymbol{\ell}} \boldsymbol{\ell}^{\mathsf{T}}(\boxed{\gamma_g} I + L)\boldsymbol{\ell}$$

$$\text{s.t. } \ell_i = y_i \text{ for all } i \in l;$$



# Regularization



# Online HFS

**Inputs:** an example $x_t$, a data adjacency graph W

**Algorithm:**

What is wrong with this algorithm?

Add $x_t$ to the graph W and compute the Laplacian L

O(t)

Infer labels on the graph:

$$\min_{\lambda \in \Re^N} \lambda^{\mathrm{T}} (L + \gamma_g I)\lambda \quad \text{s.t. } \lambda_i = y_i \text{ for all } i \in l$$

O($t^3$)

Predict $\hat{y}_t = \lambda_t$

**Outputs:** a prediction $\hat{y}_t$, an updated data adjacency graph W

# Online HFS

**Inputs:** an example $x_t$, a data adjacency graph W

**Algorithm:**

If the graph W has more than M vertices, quantize it

Add $x_t$ to the graph W and compute the Laplacian L

Infer labels on the graph:

$$\min_{\lambda \in \Re^N} \lambda^{\mathrm{T}} (L + \gamma_g I)\lambda \quad \text{s.t. } \lambda_i = y_i \text{ for all } i \in l$$

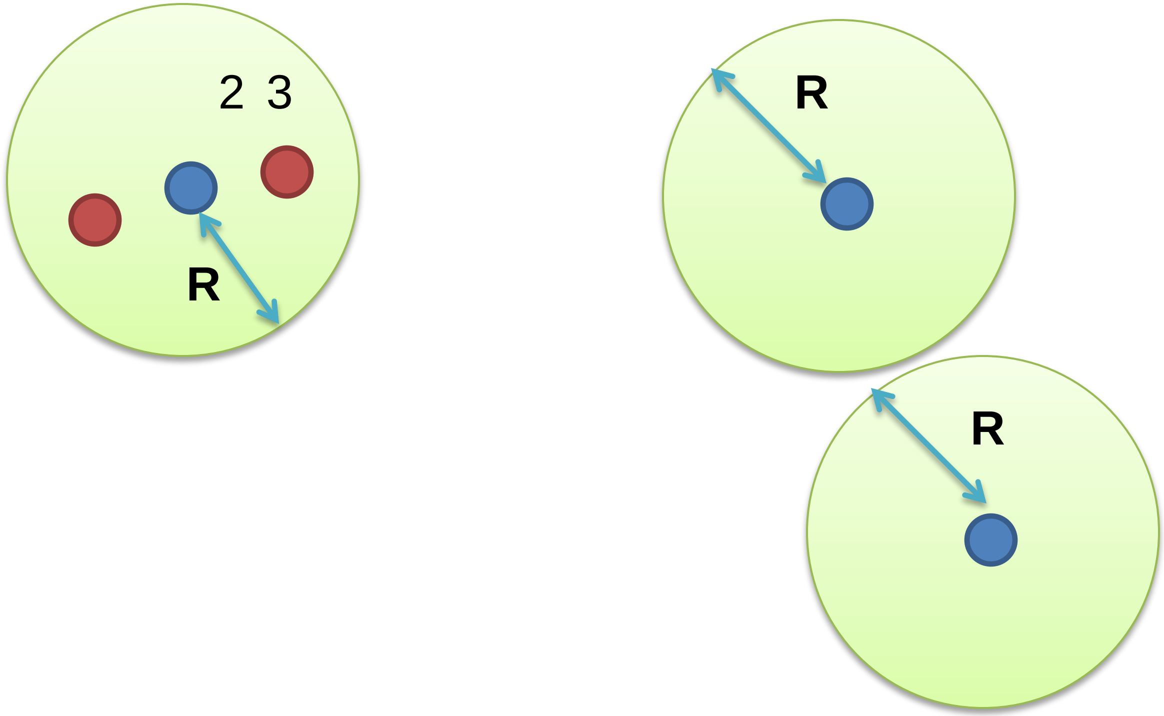
Predict $\hat{y}_t = \lambda_t$

O(M)

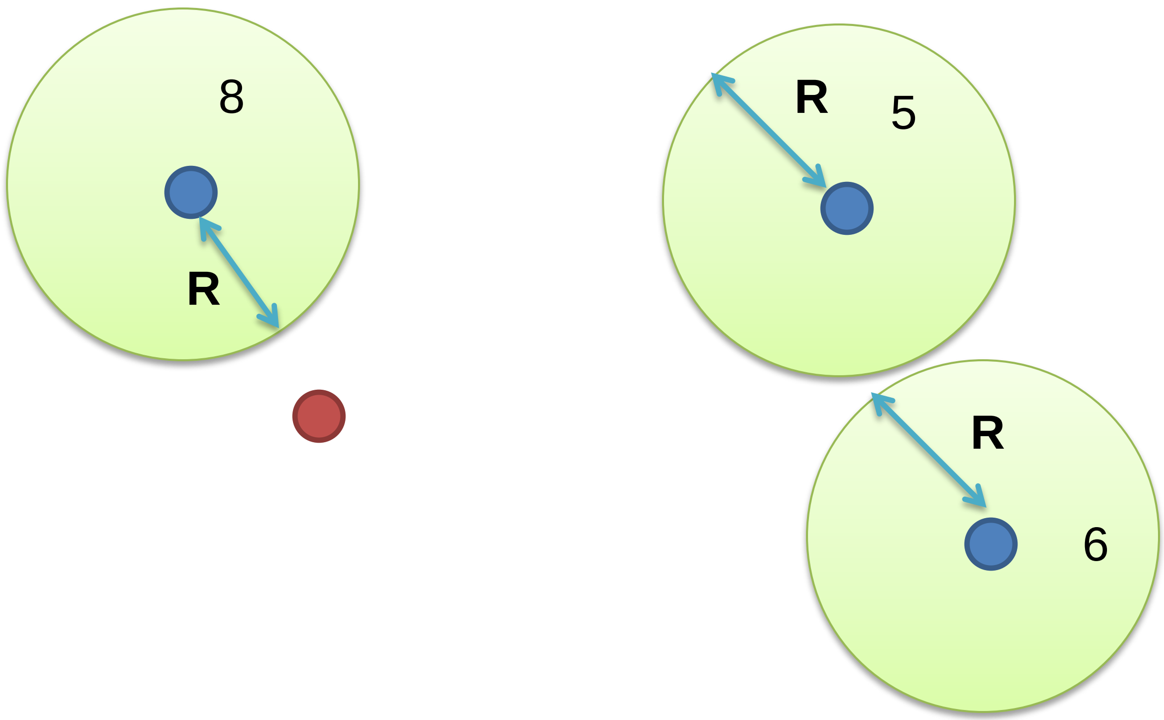
O(M³)

**Outputs:** a prediction $\hat{y}_t$, an updated data adjacency graph W

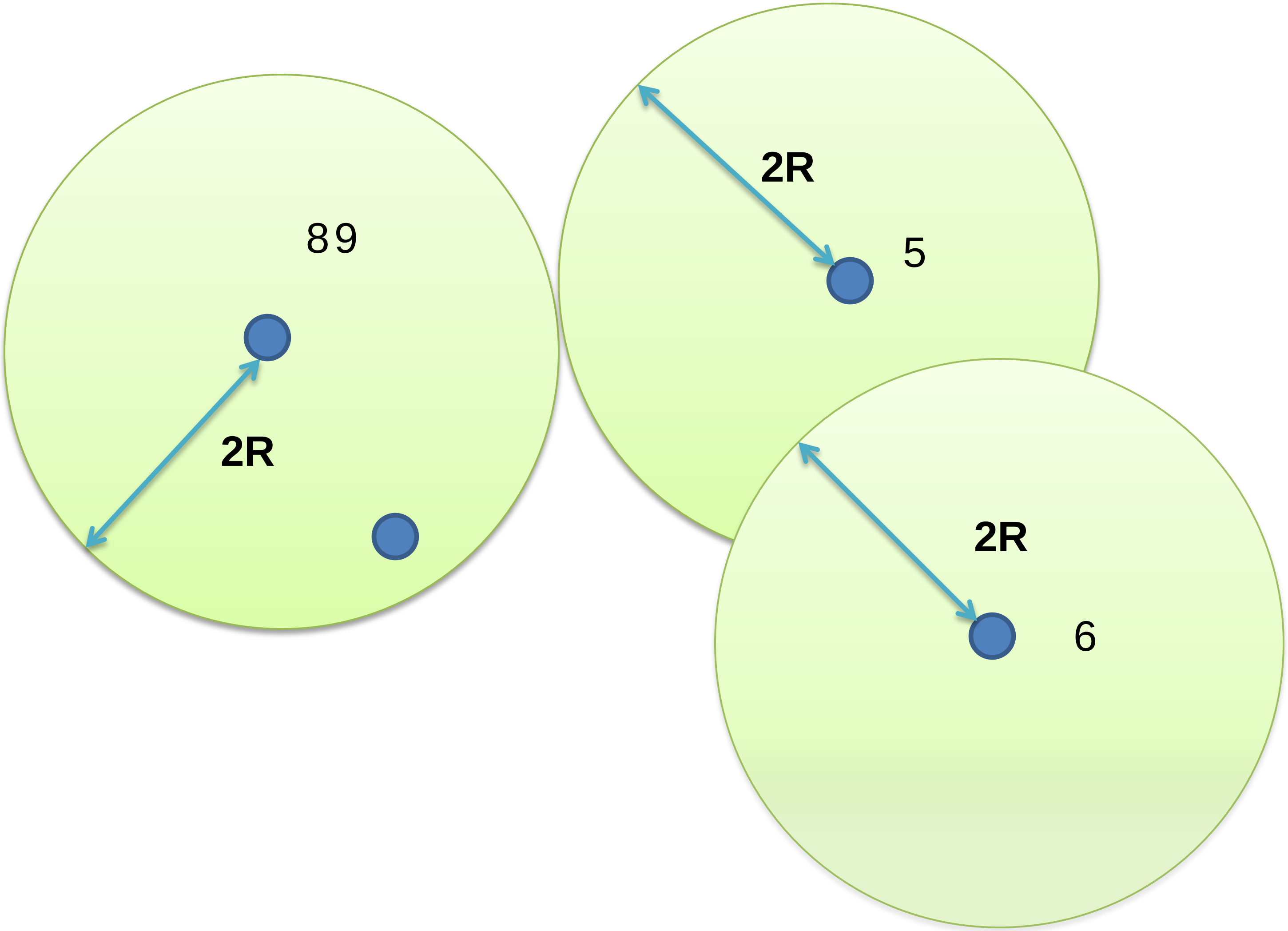
# Incremental k-centers



# Incremental k-centers



# Incremental k-centers



# Demostration

# Theoretical Guarantees

- We seek a regret bound of the form:

$$
\begin{aligned}
\frac{1}{N}\sum_t (\hat{y}_t - y_t)^2 \quad \le \quad & \frac{1}{N}\sum_t (y_t^* - y_t)^2 \\
& + \\
& \frac{1}{N}\sum_t (y_t' - y_t^*)^2 \\
& + \\
& \frac{1}{N}\sum_t (\hat{y}_t - y_t')^2
\end{aligned}
$$

Online learning risk

Offline learning error

Online learning error

Quantization error

- The errors should be bounded on the order of $O\left(\sqrt{N}\right)$

# OfficeSpace Dataset

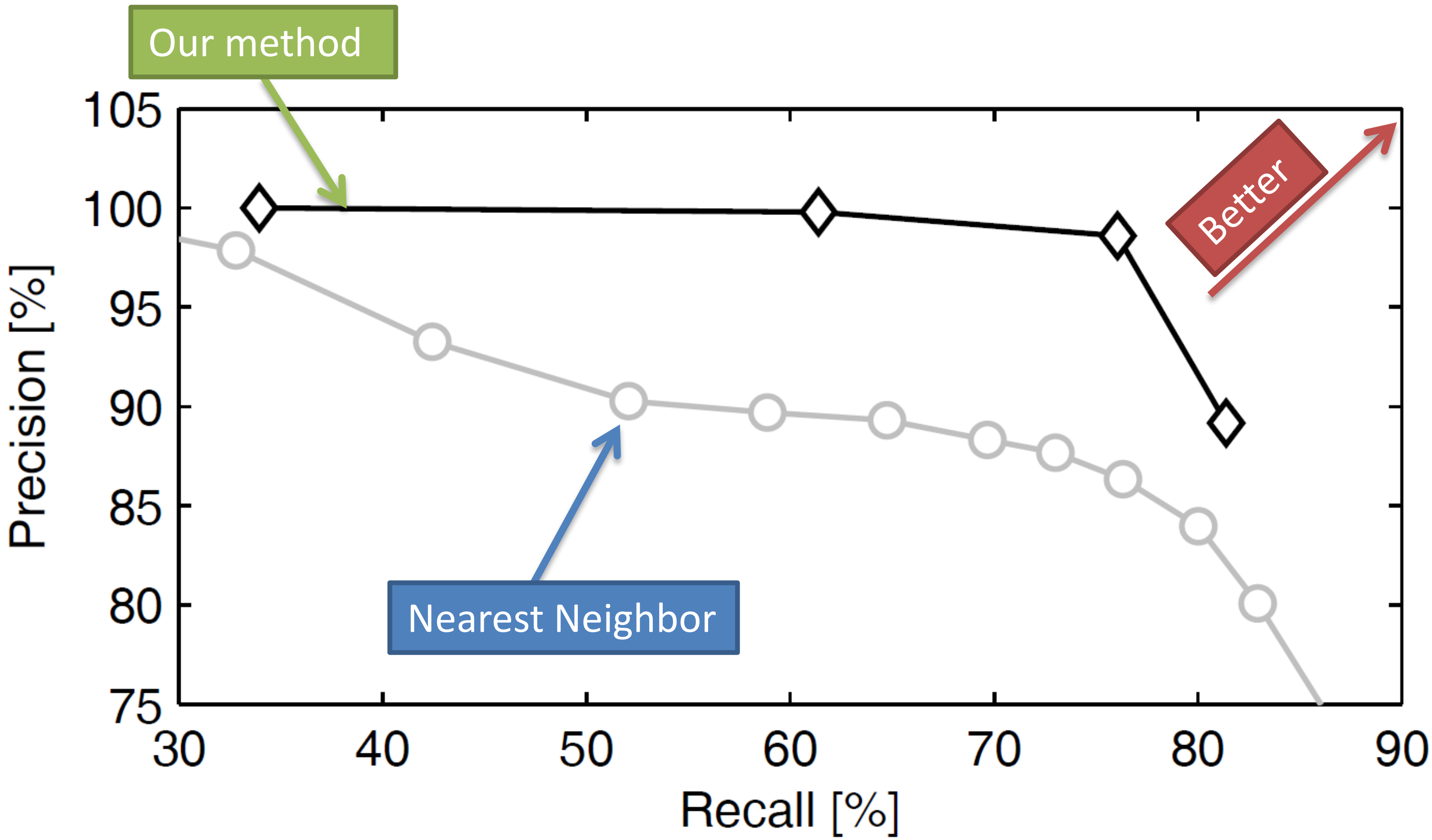
Snapshots

V1

V2

- 8 people
- Only 4 faces are labeled

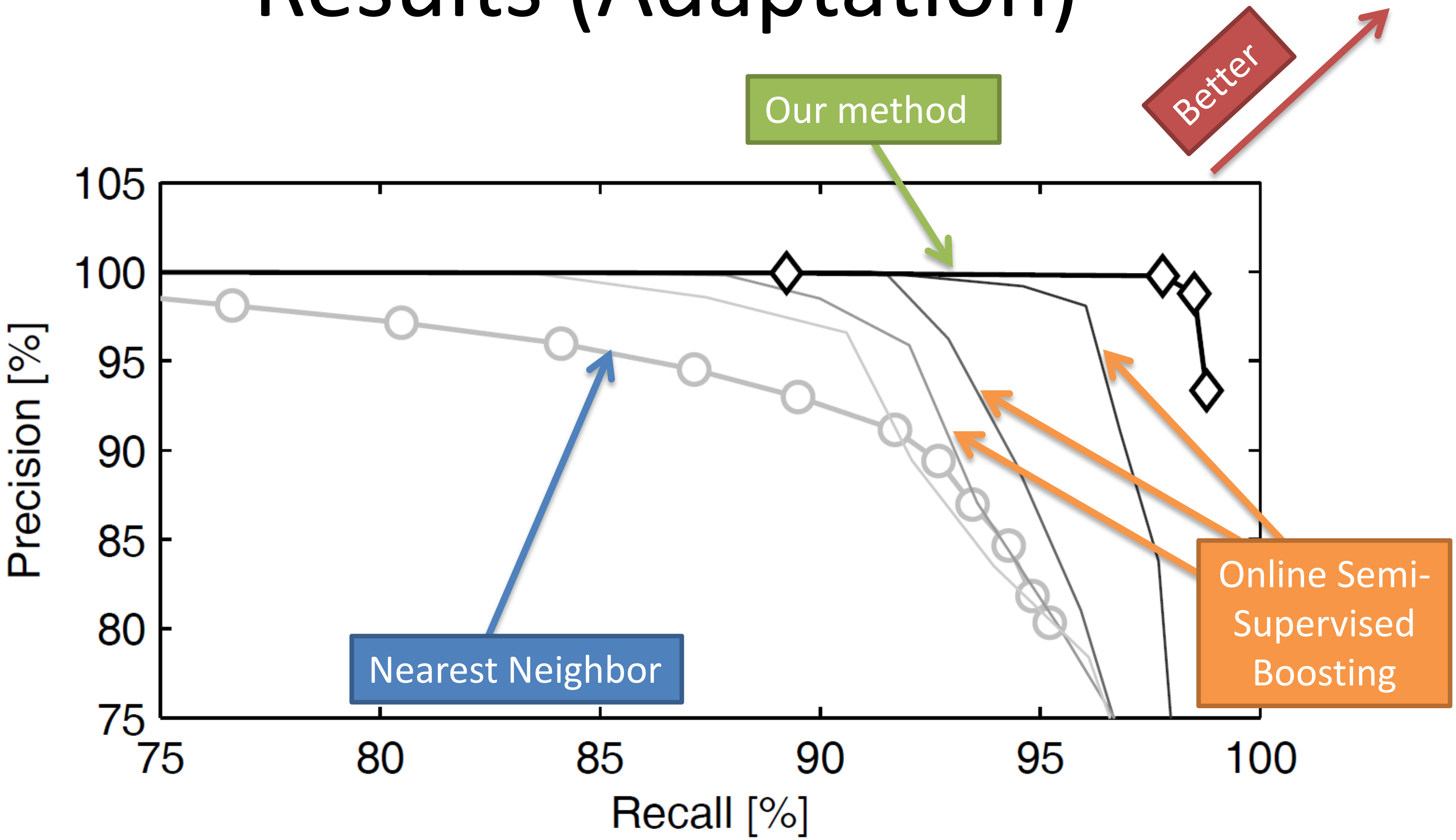
# Results (OfficeSpace)



# Adaptation Dataset

- 3 locations, different light conditions
- 8 camera positions

# Results (Adaptation)



# Conclusions

- Algorithm for semi-supervised learning
  - Takes advantage of the manifold structure in the data
- Requires minimal feedback
  - Only 1 or few labeled examples
- Works online and requires constant storage
- Theoretical guarantees on success rates of our methods
- Future work:
  - other data reduction methods
  - other domains: object recognition, augmented reality