NEURAL INFORMATION PROCESSING SYSTEMS

TopP&R: Robust Support Estimation Approach for Evaluating Fidelity and Diversity in Generative Models


Pum jun Kim[1]
UNIST

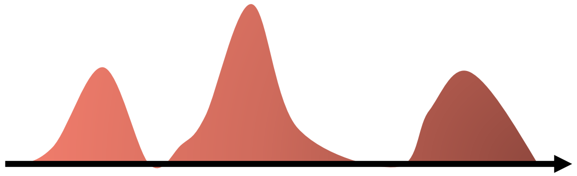

Yoojin Jang
UNIST


Jisu Kim
SNU
Inria


Jaejun Yoo*
UNIST

LAB. OF ADVANCED IMAGING TECHNOLOGY

Image Generative Models and Evaluation Protocol



Real image distribution

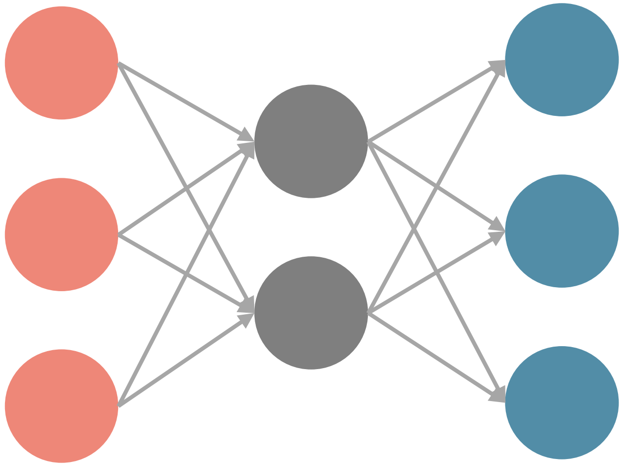
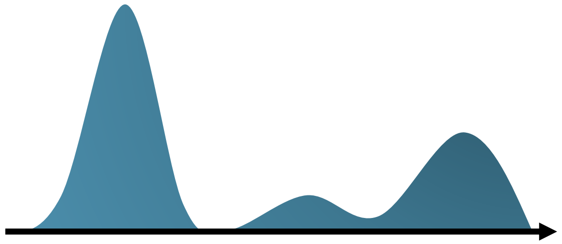
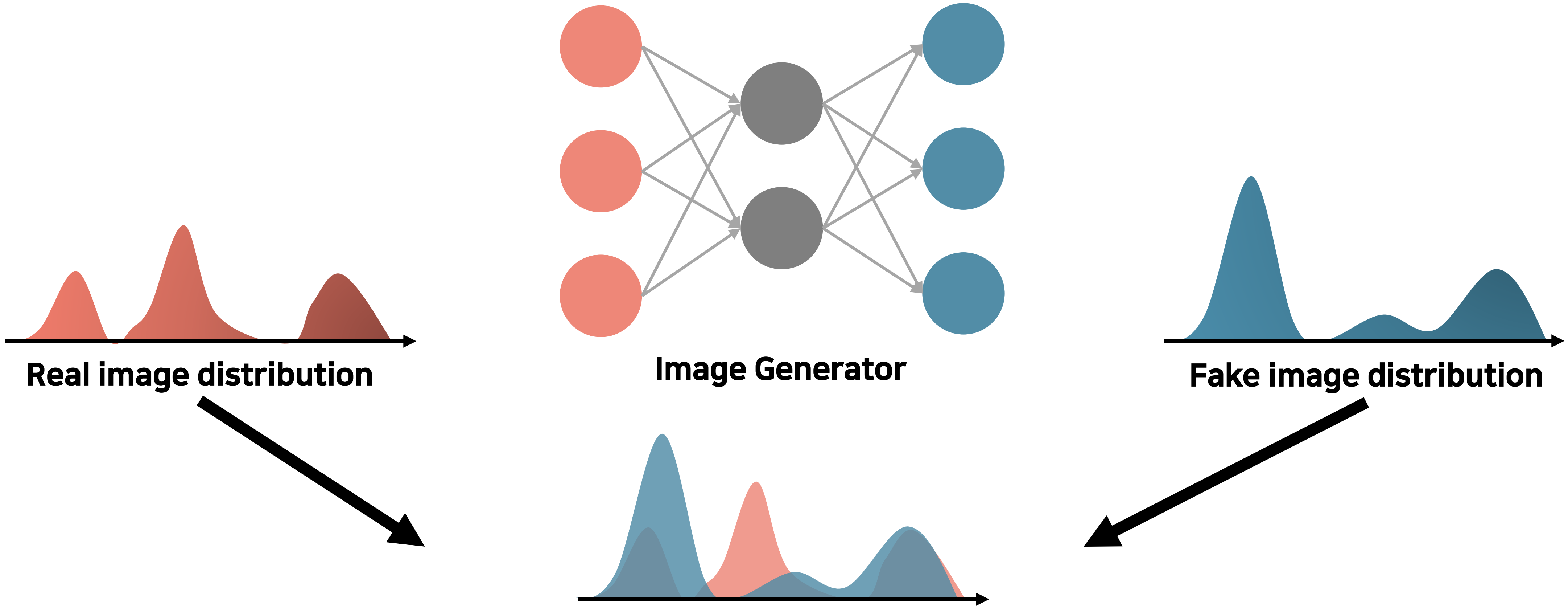


Image Generator



Fake image distribution

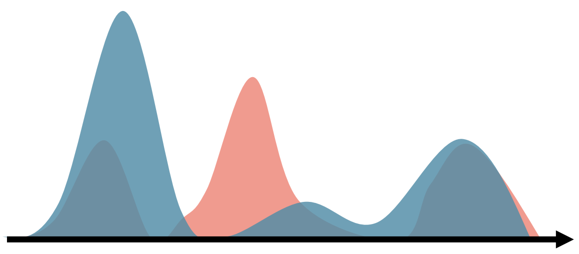
Image Generative Models and Evaluation Protocol



Comparison

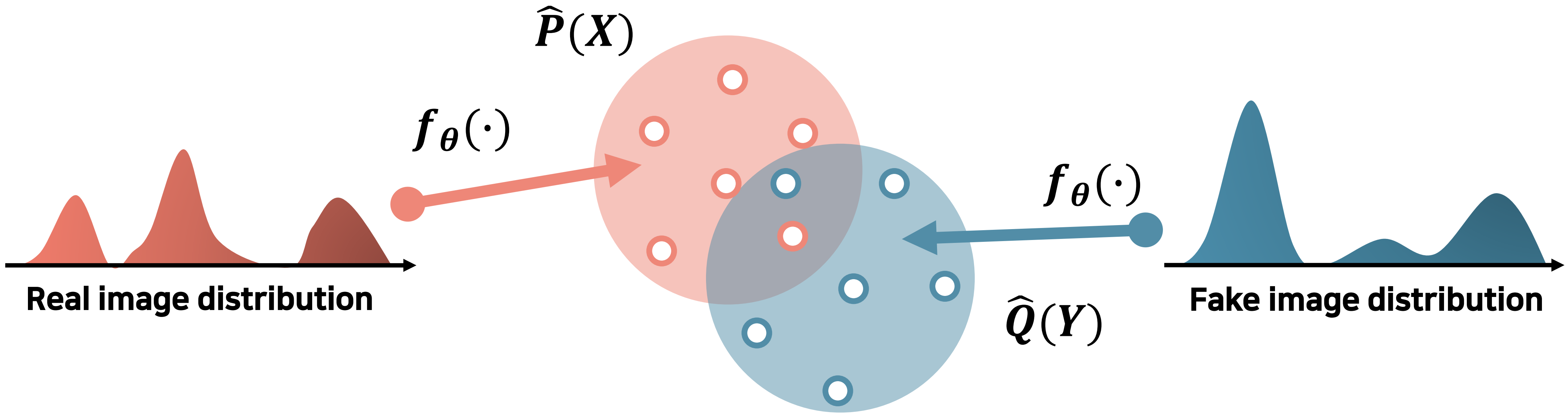
Image Generative Models and Evaluation Protocol

Unknown!



Comparison

Image Generative Models and Evaluation Protocol

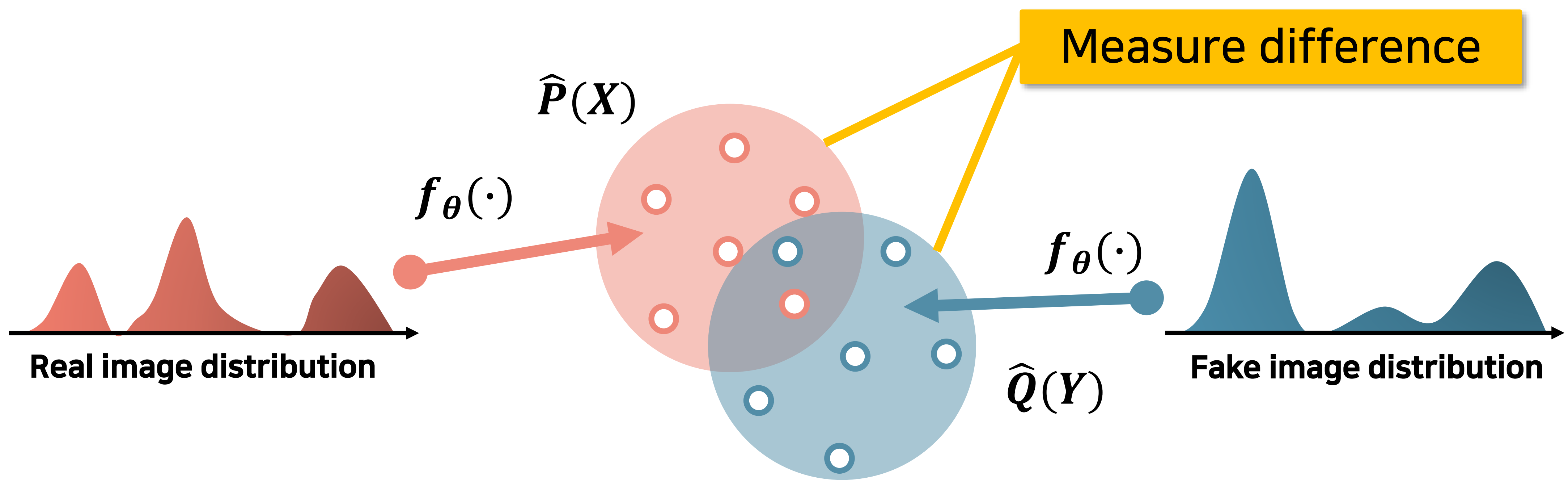


$f_\theta(\cdot)$: Pretrained embedding network


: Extracted real image features


: Extracted fake image features

Image Generative Models and Evaluation Protocol

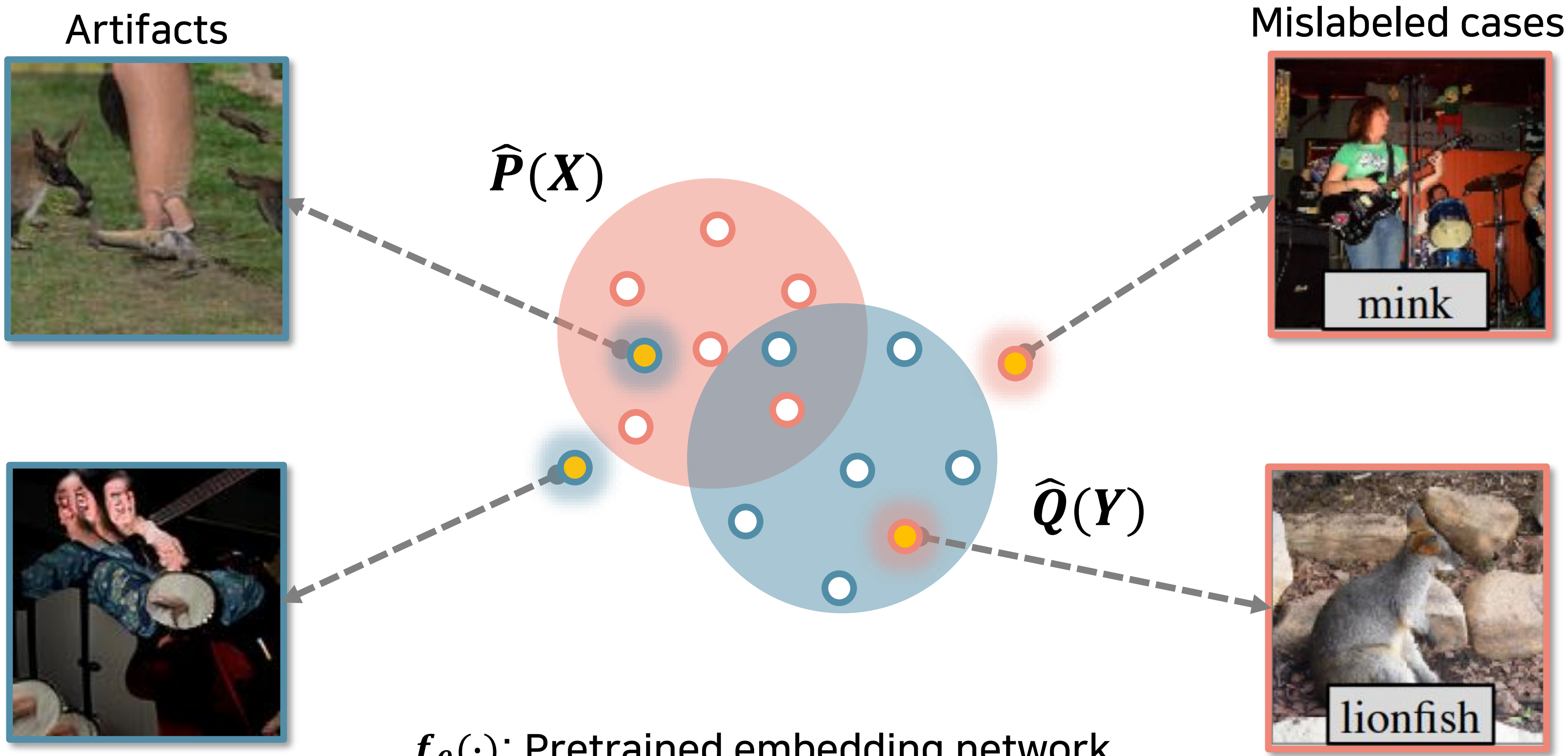


$f_\theta(\cdot)$: Pretrained embedding network

○ : Extracted real image features

○ : Extracted fake image features

Image Generative Models and Evaluation Protocol



$f_\theta(\cdot)$: Pretrained embedding network

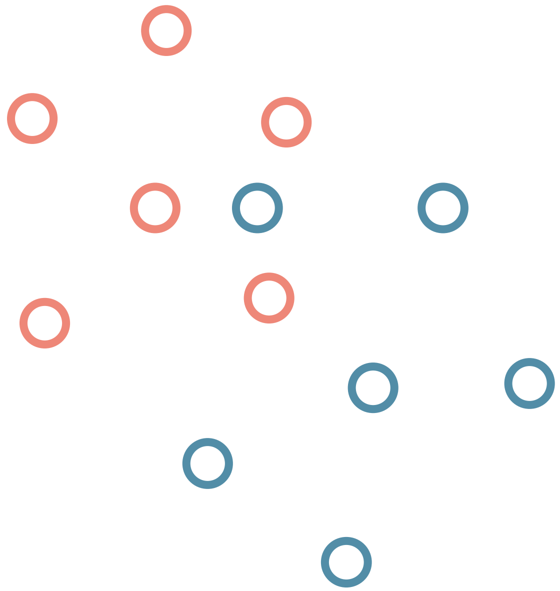
: Extracted real image features

: Extracted fake image features

Pleiss, Geoff, et al. "Identifying mislabeled data using the area under the margin ranking." *Advances in Neural Information Processing Systems* 33 (2020): 17044-17056.

Image Generative Models and Evaluation Protocol

(1) Ideal estimation of distribution



: real image features

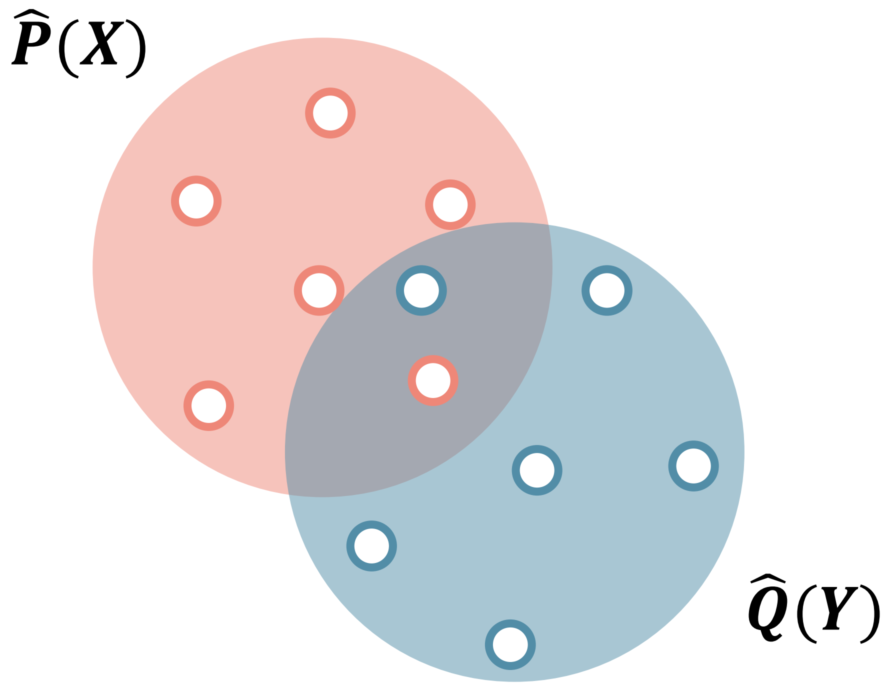
: fake image features

: real noisy features

: fake noisy features

Image Generative Models and Evaluation Protocol

(1) Ideal estimation of distribution



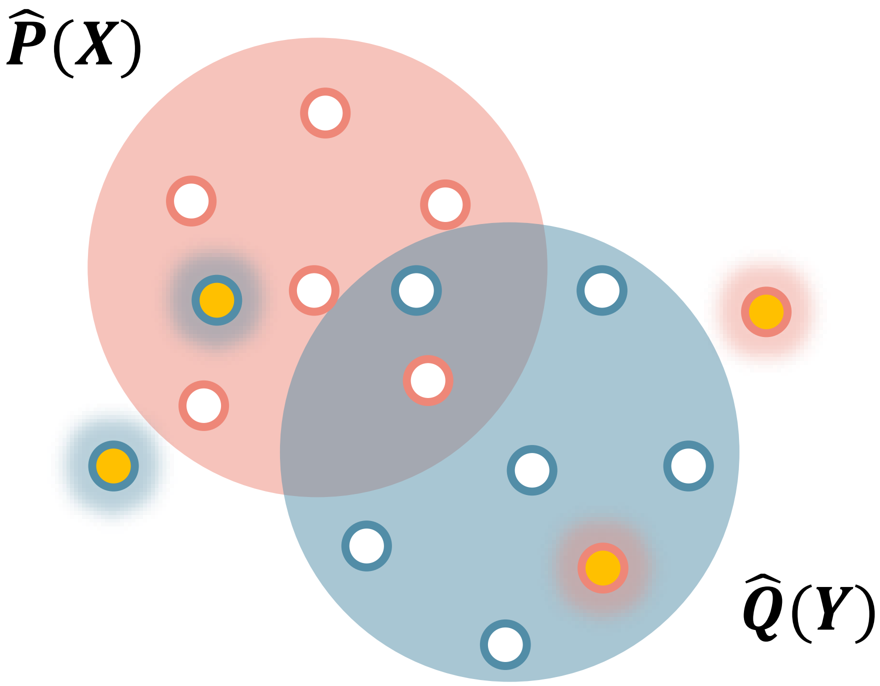
Correctly measured scores

Fidelity $= \boldsymbol{Q(supp(P)) = 0.16}$

Diversity $= \boldsymbol{P(supp(Q)) = 0.16}$

Image Generative Models and Evaluation Protocol

(2) Non-ideal estimation of distribution



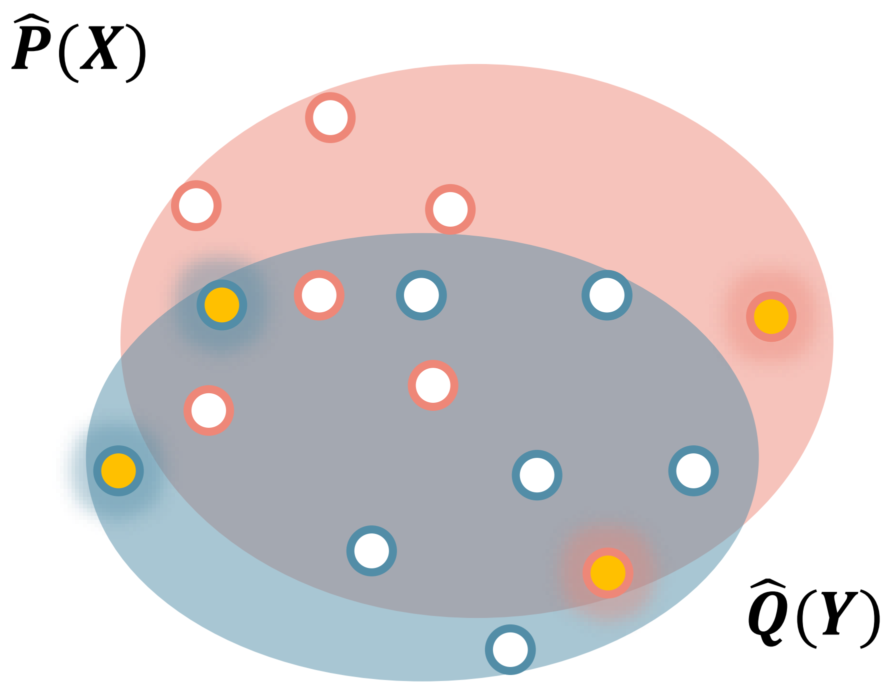
Correctly measured scores

Fidelity $= \boldsymbol{Q(supp(P)) = 0.16}$

Diversity $= \boldsymbol{P(supp(Q)) = 0.16}$

Image Generative Models and Evaluation Protocol

(2) Non-ideal estimation of distribution



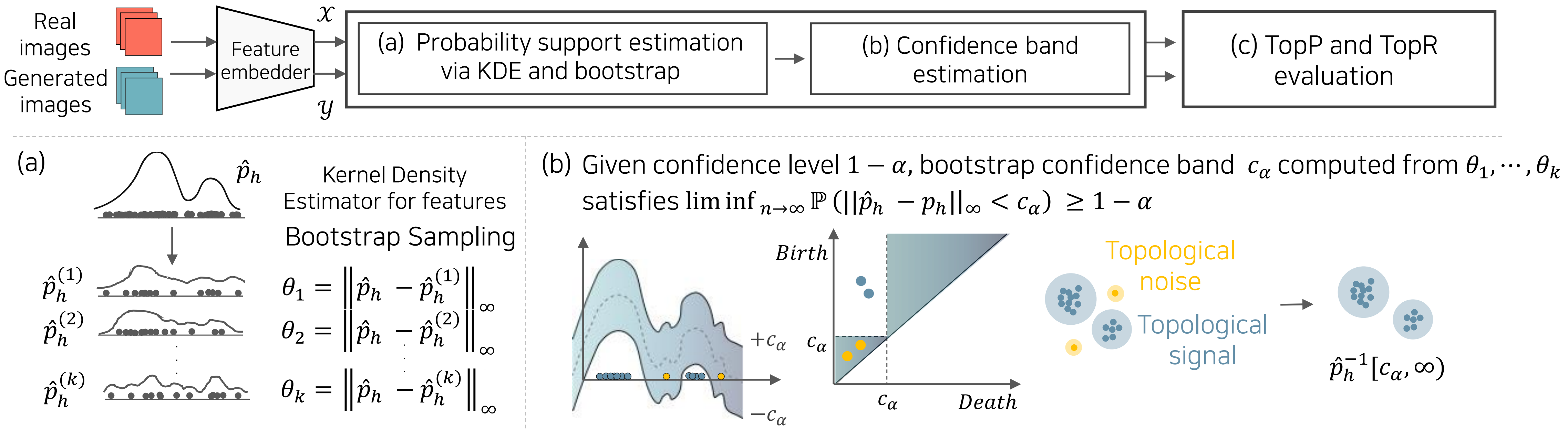
Incorrectly measured scores

Fidelity $= Q(supp(P)) = \mathbf{0.75}$

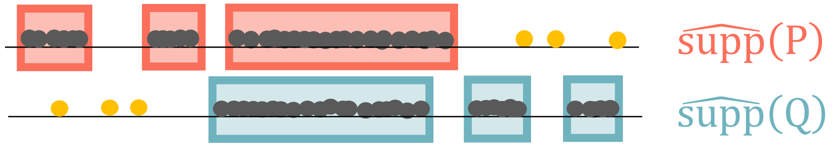
Diversity $= Q(supp(P)) = \mathbf{0.5}$

: real image features
: fake image features
: real noisy features
: fake noisy features

Overview of TopP&R



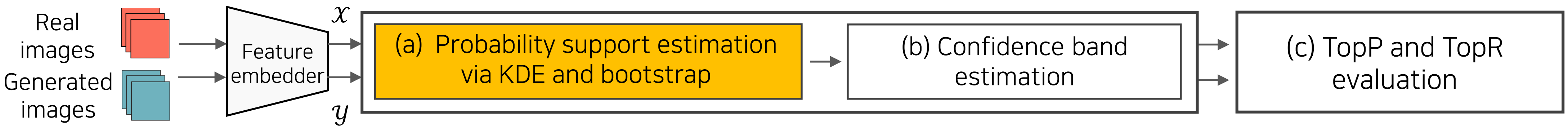
(c) Given $\widehat{supp}(P)$ for real features $\mathcal{X}$ and $\widehat{supp}(Q)$ for generated features $\mathcal{Y}$,

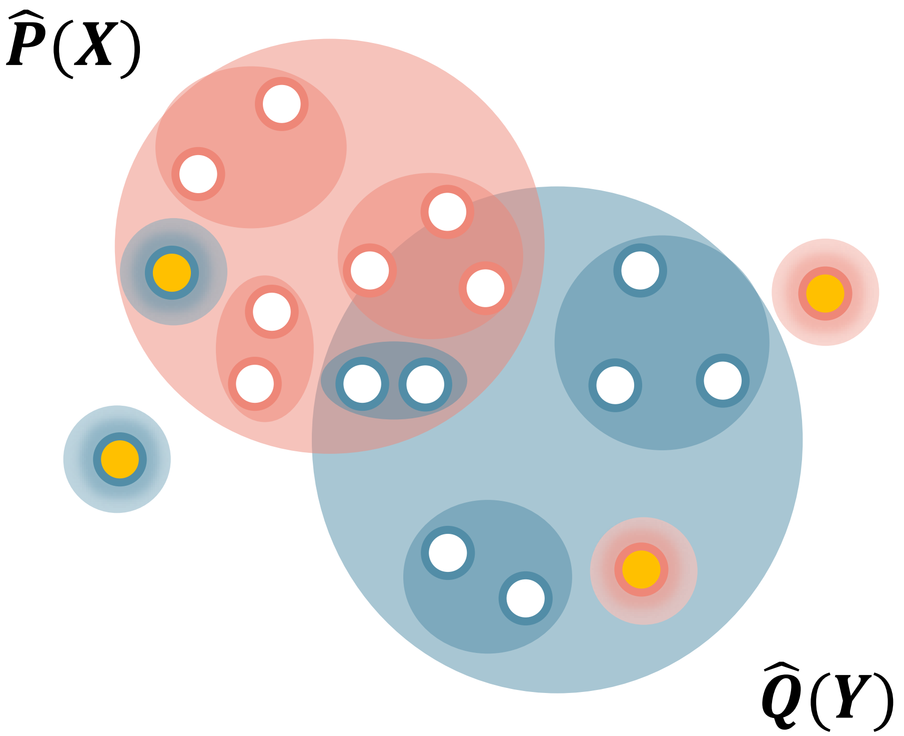


$$TopP_{\mathcal{X}}(\mathcal{Y}) := \frac{\sum_{j=1}^{m} 1\left(Y_j \in \widehat{supp}(P) \cap \widehat{supp}(Q)\right)}{\sum_{j=1}^{m} 1\left(Y_j \in \widehat{supp}(Q)\right)}$$

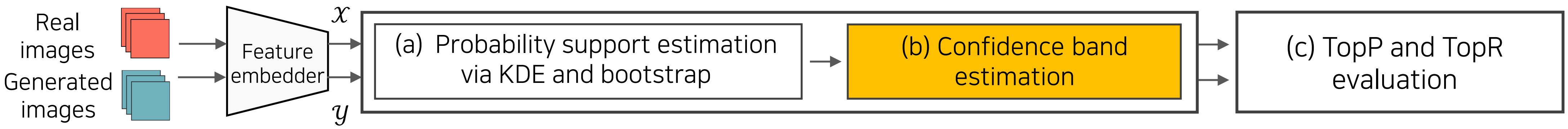
$$TopR_{\mathcal{Y}}(\mathcal{X}) := \frac{\sum_{i=1}^{n} 1\left(X_i \in \widehat{supp}(Q) \cap \widehat{supp}(P)\right)}{\sum_{i=1}^{n} 1\left(X_i \in \widehat{supp}(P)\right)}$$

Overview of TopP&R

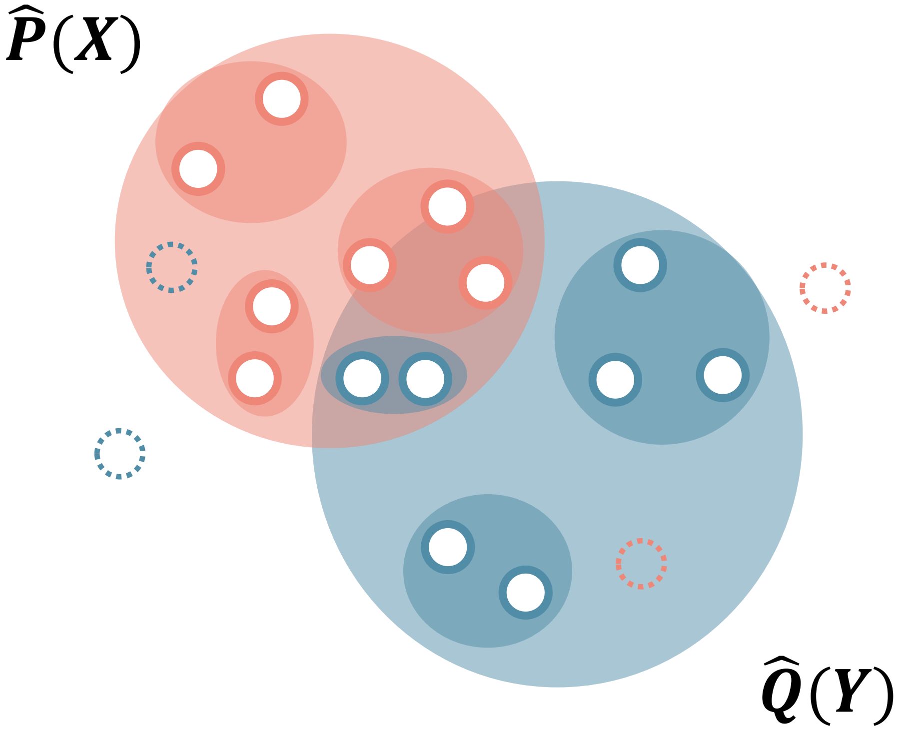




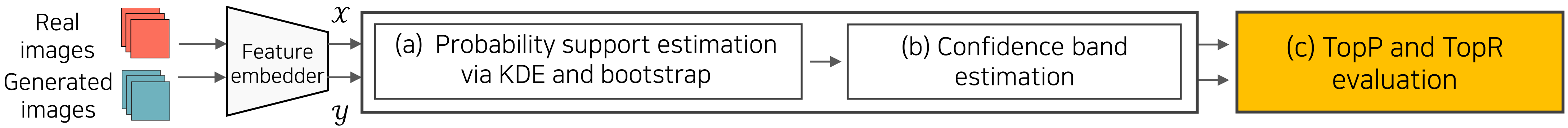
Overview of TopP&R

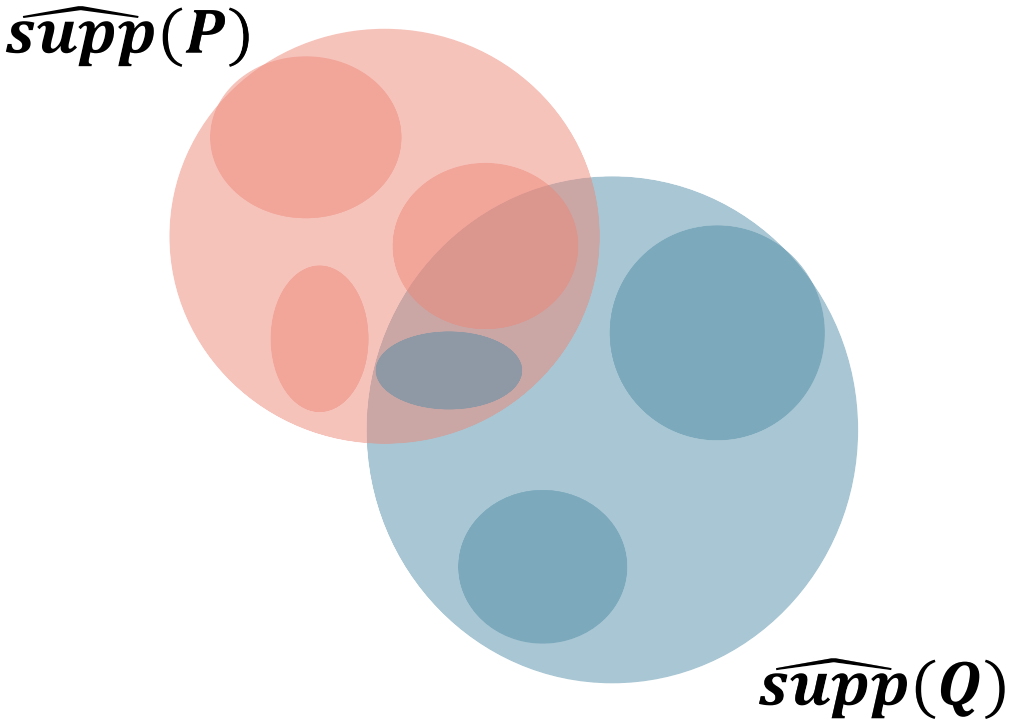


Exclude noisy features under statistical confidence level we set



Overview of TopP&R





$$TopP \xrightarrow[n\to\infty]{} precision$$

$$TopR \xrightarrow[n\to\infty]{} recall$$

See our proposition 4.1 and theorem 4.2

Topological Data Analysis (TDA) as a Solution

Selecting a threshold that effectively removes noisy samples

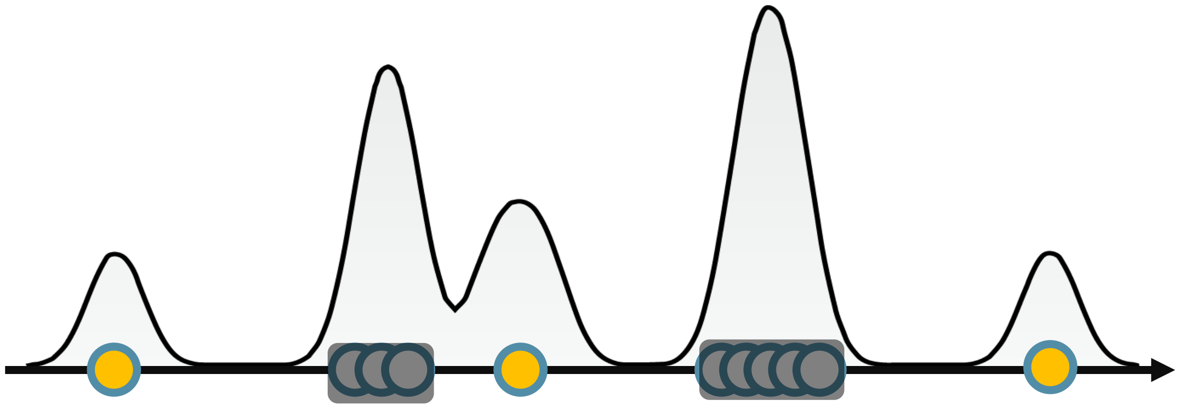
(1) Reference

: significant feature

: noisy features

: support of estimated distribution

: ground truth support of distribution



Topological Data Analysis (TDA) as a Solution

Selecting a threshold that effectively removes noisy samples

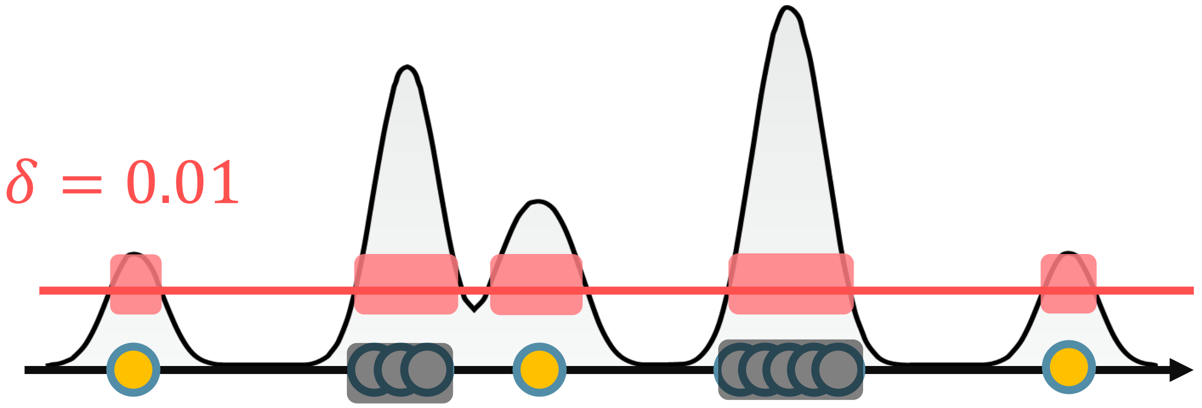
(2) When $\delta = 0.01$ (over-estimated case)

: significant feature

: noisy features

: support of estimated distribution

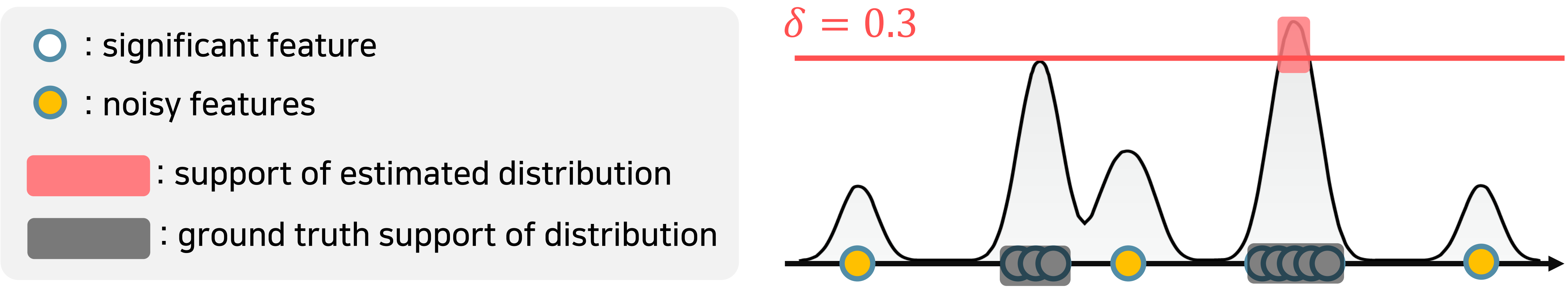
: ground truth support of distribution



Topological Data Analysis (TDA) as a Solution

Selecting a threshold that effectively removes noisy samples

(3) When $\delta = 0.3$ (under-estimated case)



Topological Data Analysis (TDA) as a Solution

Selecting a threshold that effectively removes noisy samples

(4) $\delta \in [0, 1]$

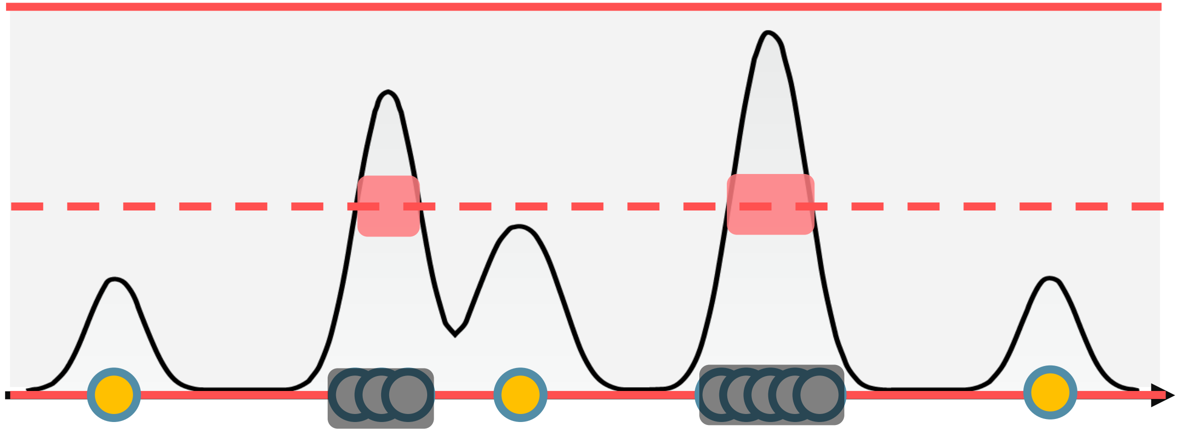
: significant feature

: noisy features

: support of estimated distribution

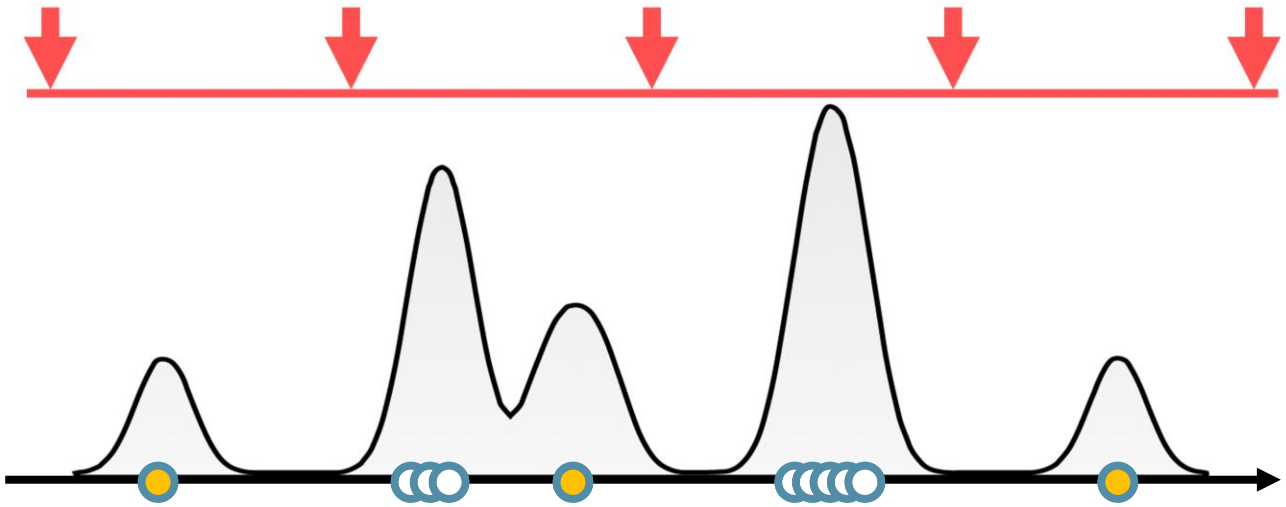
: ground truth support of distribution

$\delta \in [0,1]$ What is the optimal threshold?



Topological Data Analysis (TDA) as a Solution

Tracking when homological feature appear and disappear



: 0-dim homological feature or connected component

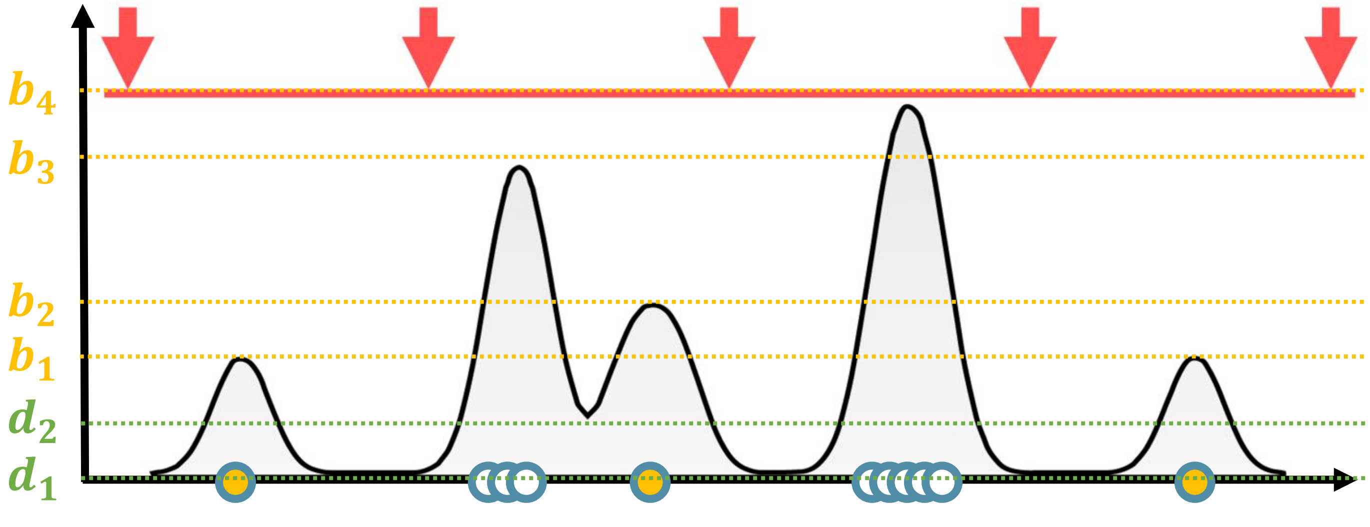
: threshold δ

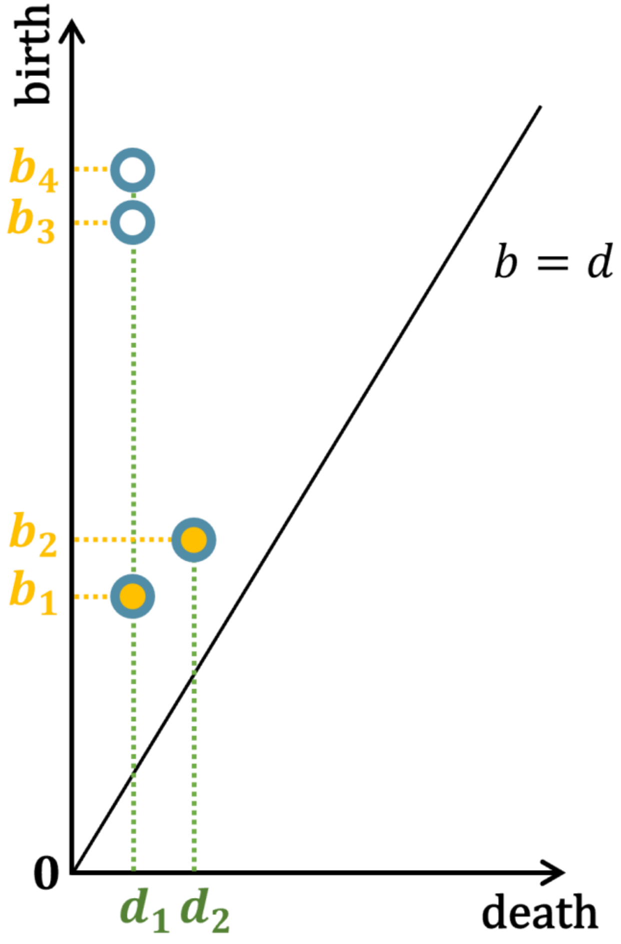
: significant feature

: noisy features

Topological Data Analysis (TDA) as a Solution

Tracking when homological feature appear and disappear





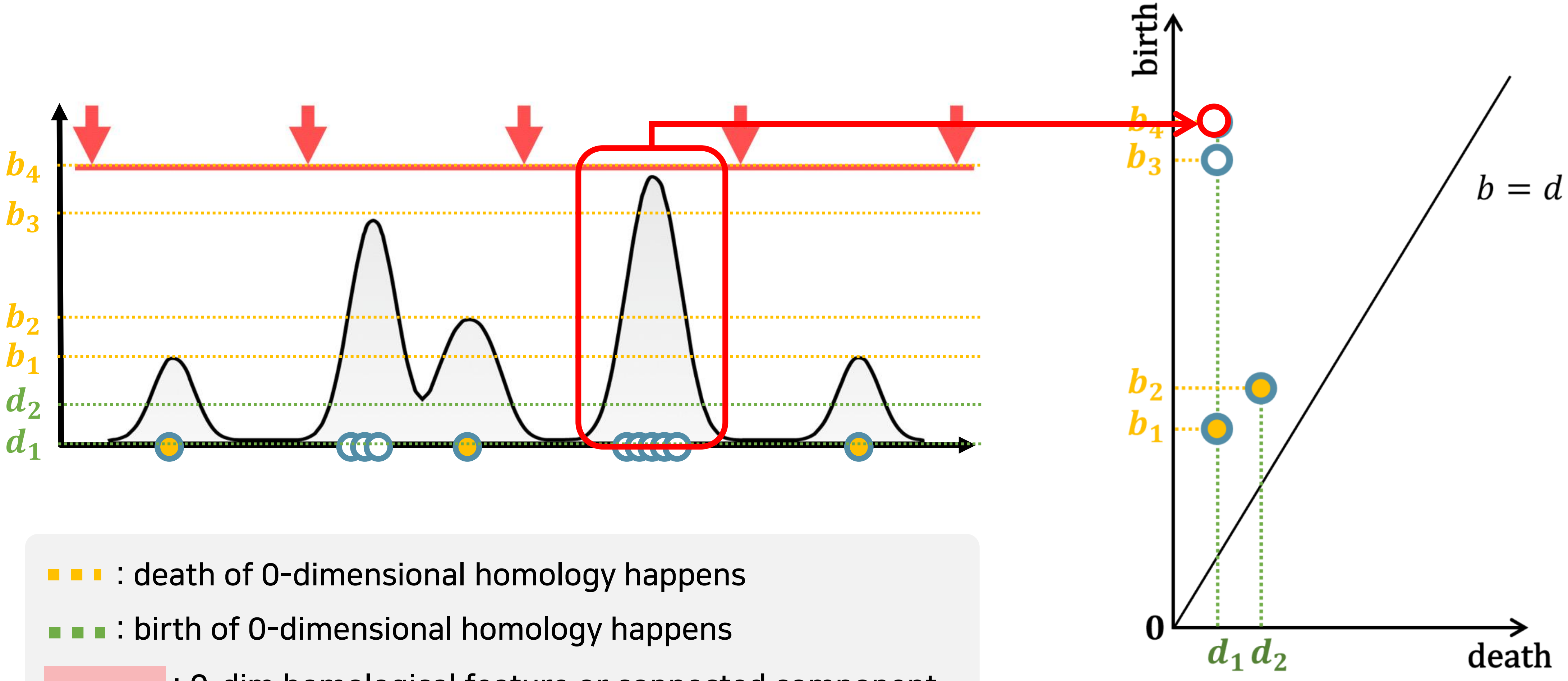
: death of 0-dimensional homology happens

: birth of 0-dimensional homology happens

: 0-dim homological feature or connected component

Topological Data Analysis (TDA) as a Solution

Tracking when homological feature appear and disappear



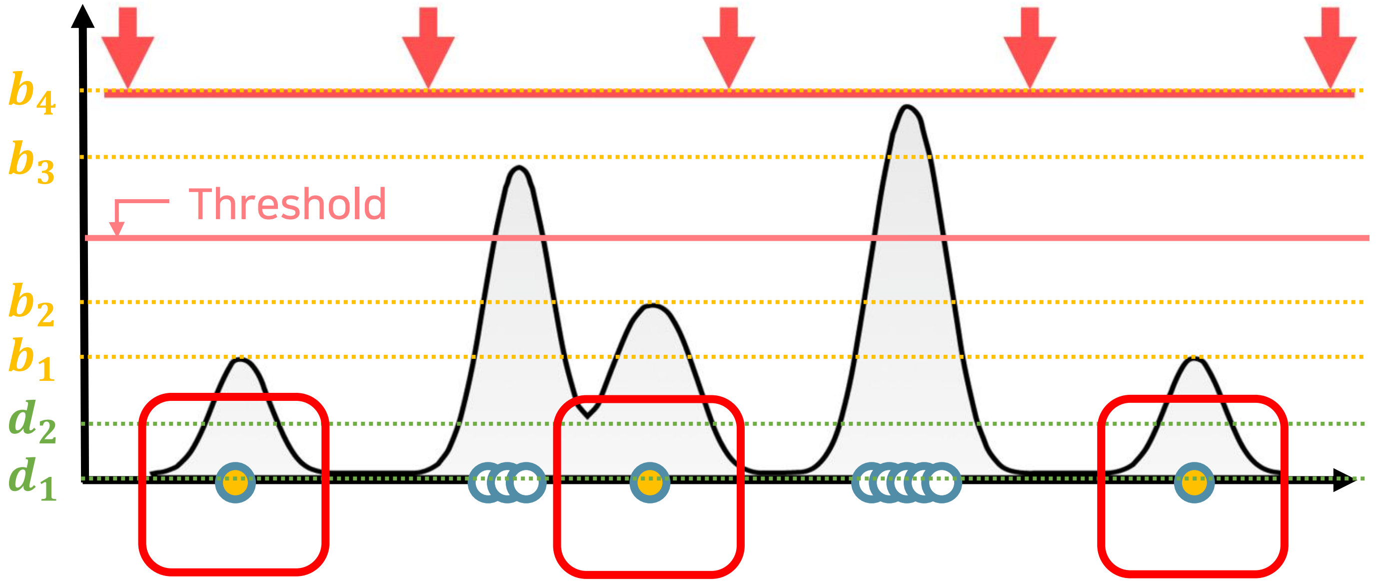
: death of 0-dimensional homology happens

: birth of 0-dimensional homology happens

: 0-dim homological feature or connected component

Topological Data Analysis (TDA) as a Solution

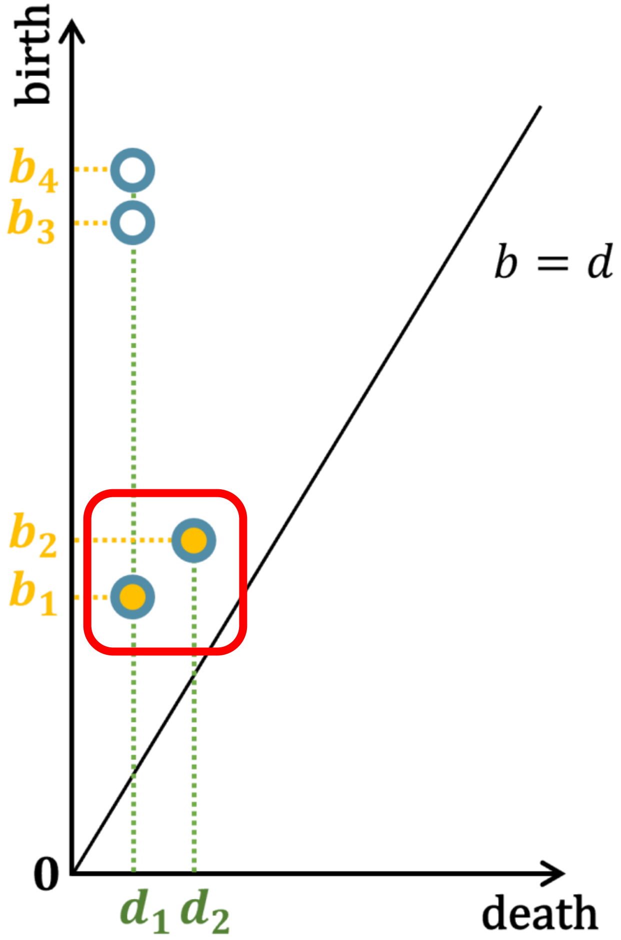
Tracking when homological feature appear and disappear



■ ■ ■ : death of 0-dimensional homology happens

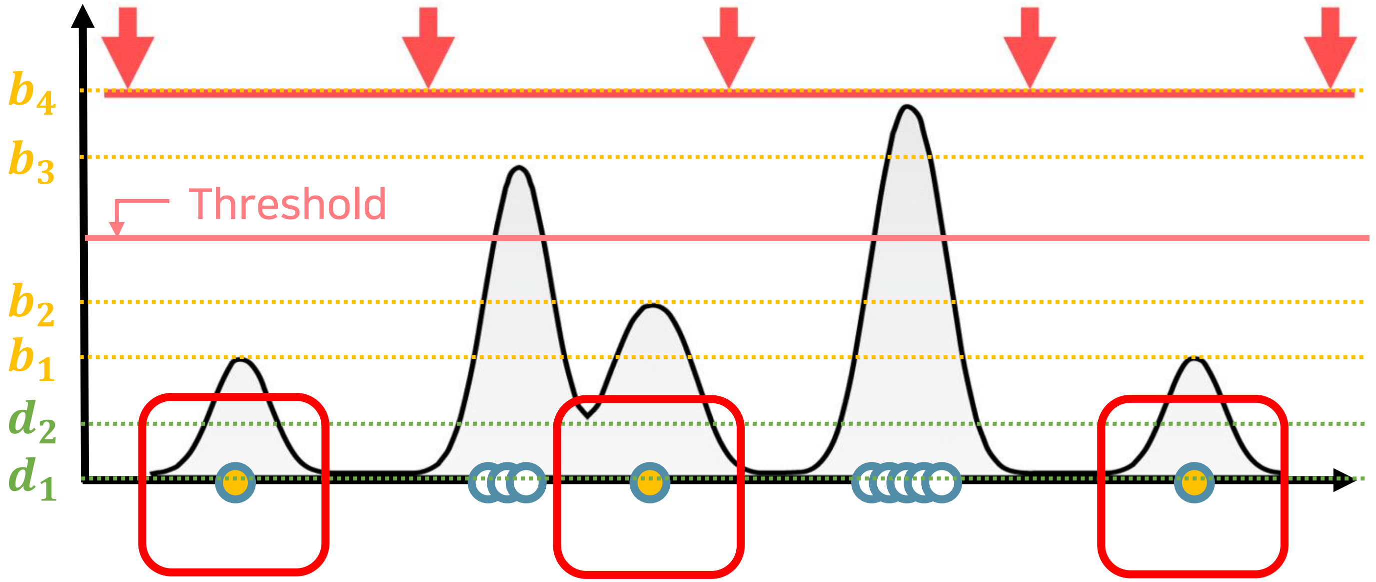
■ ■ ■ : birth of 0-dimensional homology happens

: 0-dim homological feature or connected component



Topological Data Analysis (TDA) as a Solution

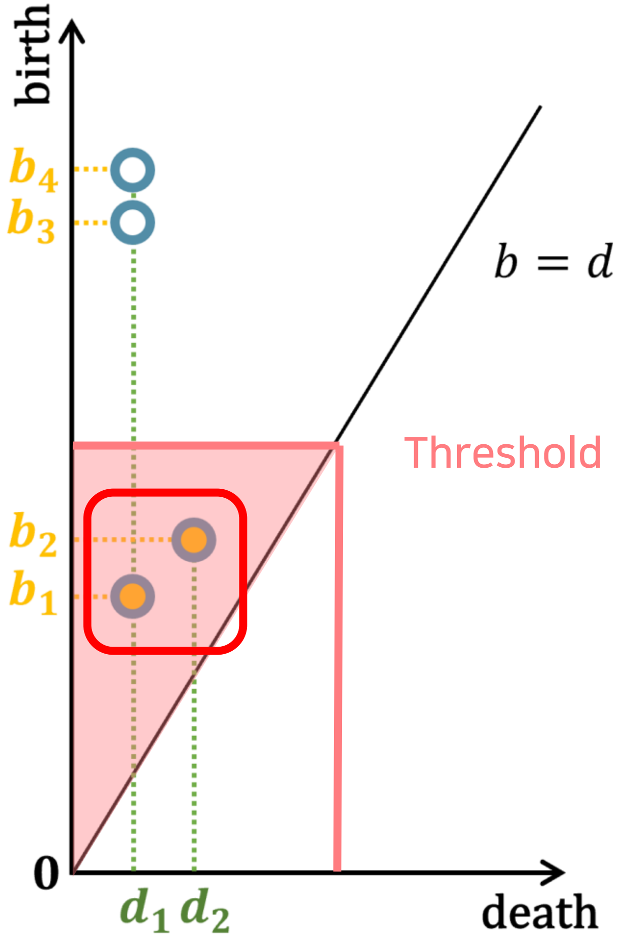
Tracking when homological feature appear and disappear



: death of 0-dimensional homology happens

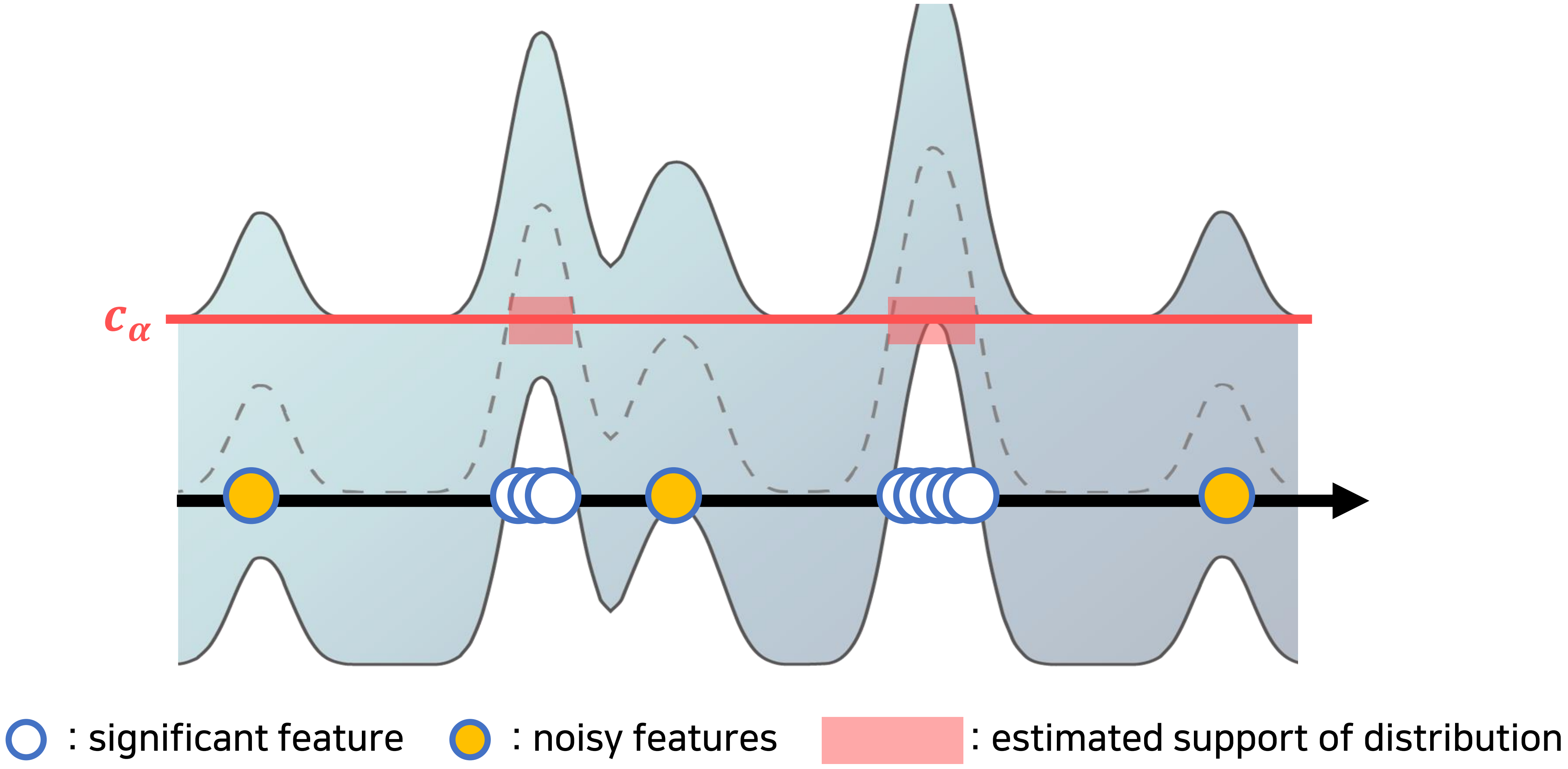
: birth of 0-dimensional homology happens

: 0-dim homological feature or connected component



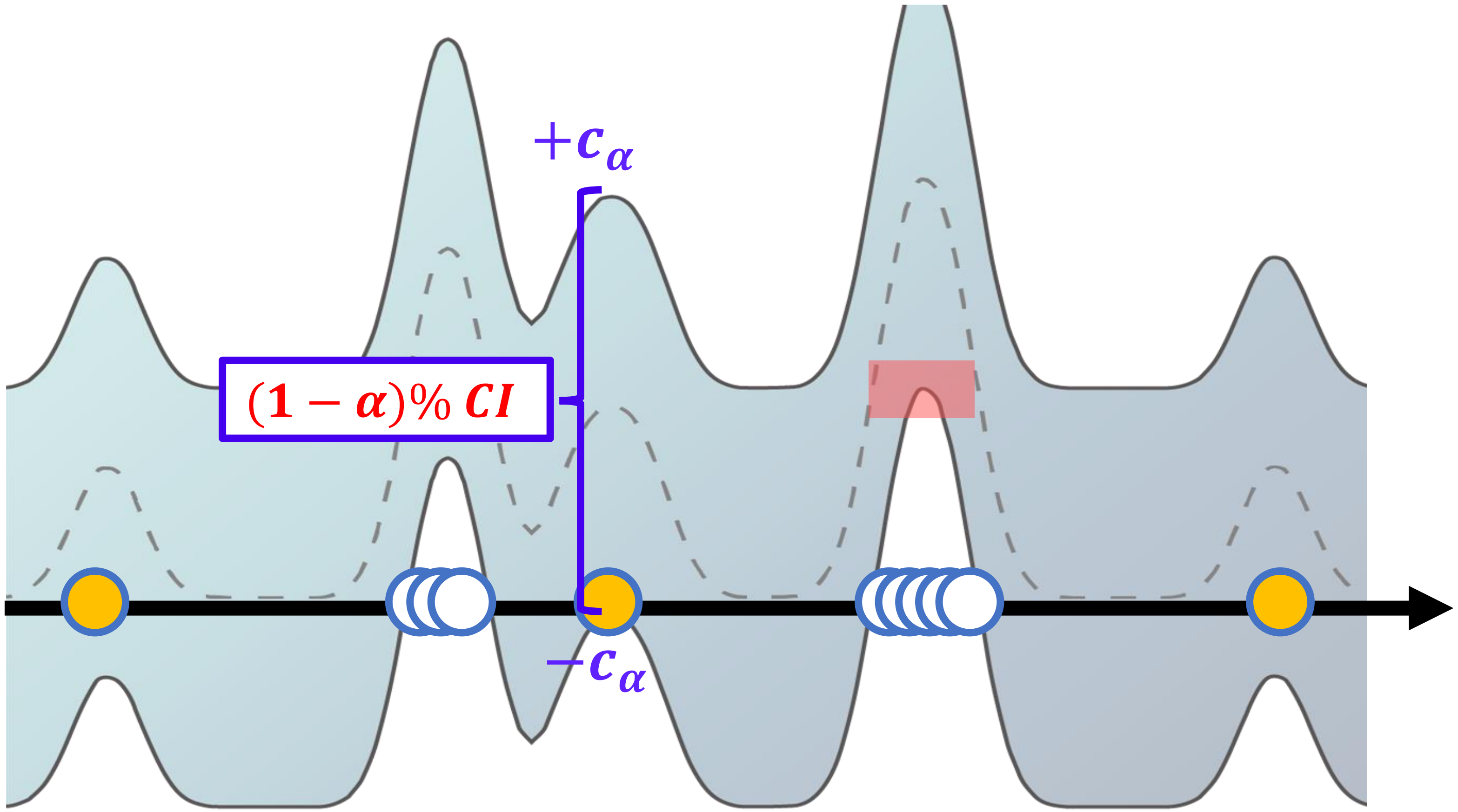
Confidence Band Estimation

Finding the threshold c_α that selects statistically and topologically significant features



Confidence Band Estimation

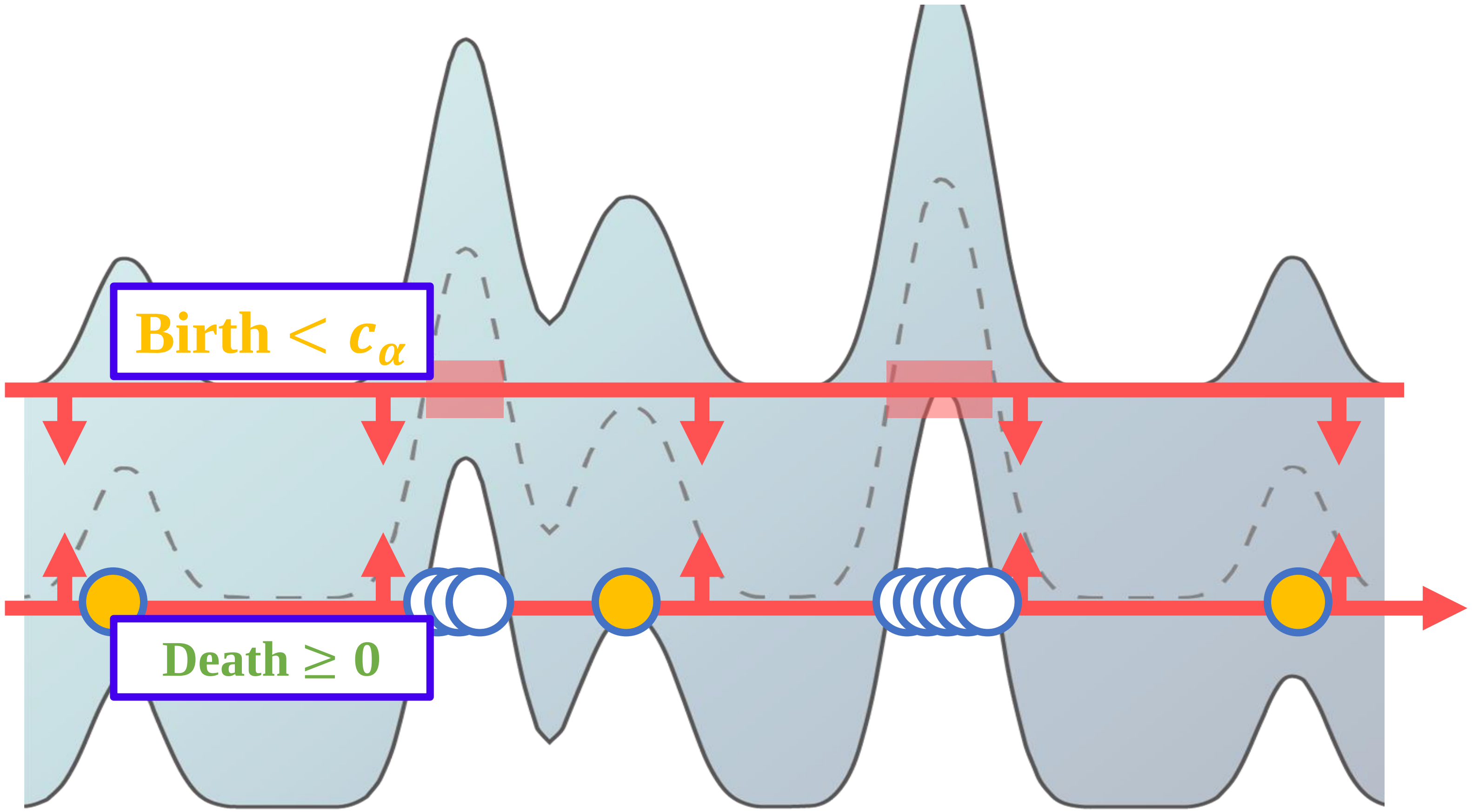
Finding the threshold c_α that selects <u>statistically</u> and <u>topologically</u> <u>significant features</u>



○ : significant feature ● : noisy features ▭ : estimated support of distribution

Confidence Band Estimation

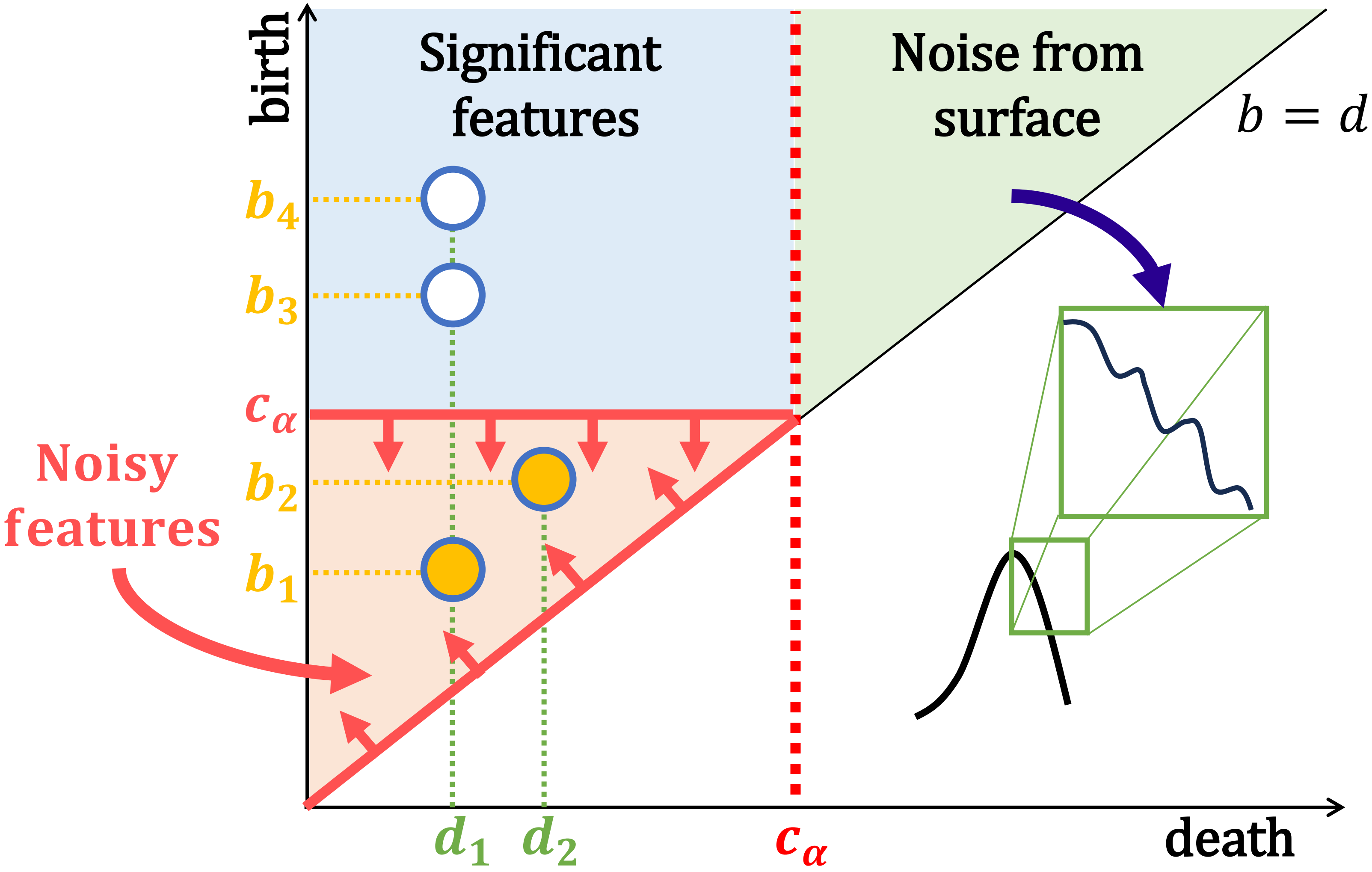
Finding the threshold c_α that selects <u>statistically</u> and <u>topologically</u> <u>significant features</u>



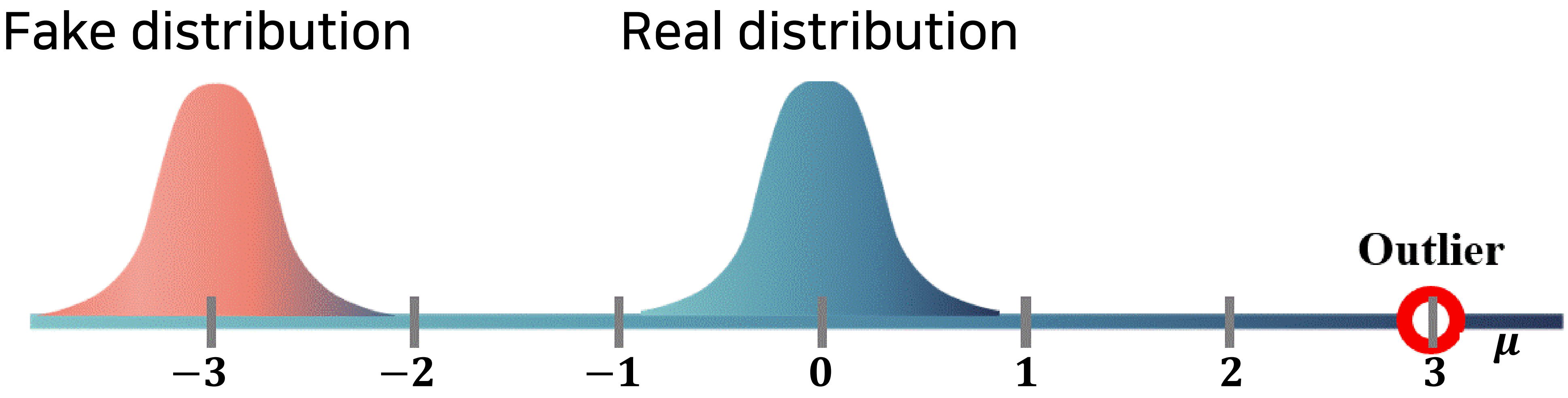
○ : significant feature ● : noisy features ▬ : estimated support of distribution

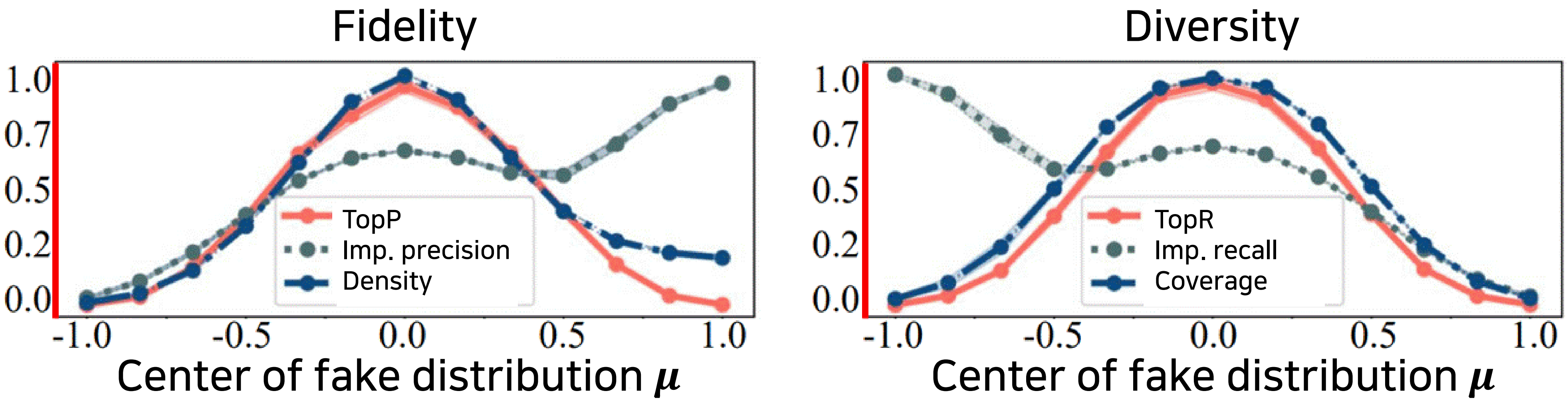
Confidence Band Estimation

Finding the threshold c_α that selects <u>statistically</u> and <u>topologically</u> <u>significant features</u>



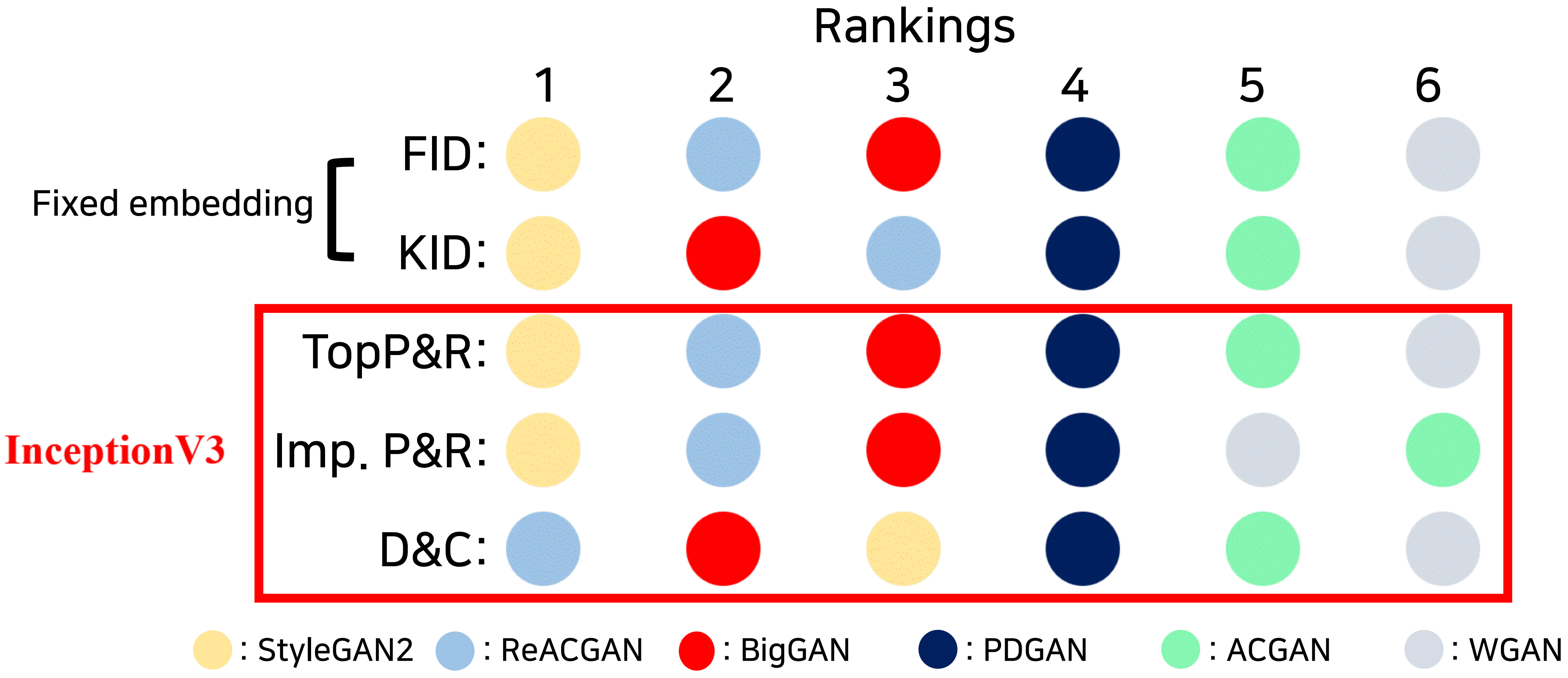
Experiments & Results – perturbation experiment





Experiments & Results – ranking experiment

Metric's consistency "with FID and KID scores" & "with different embeddings"



Thank you

Project Page

Quick Start!

```
pip install top-pr
```

Use our method by only pip command!