# Harnessing the Power of Distributions: Probabilistic Representation Learning on Hypersphere for Multimodal Music Information Retrieval

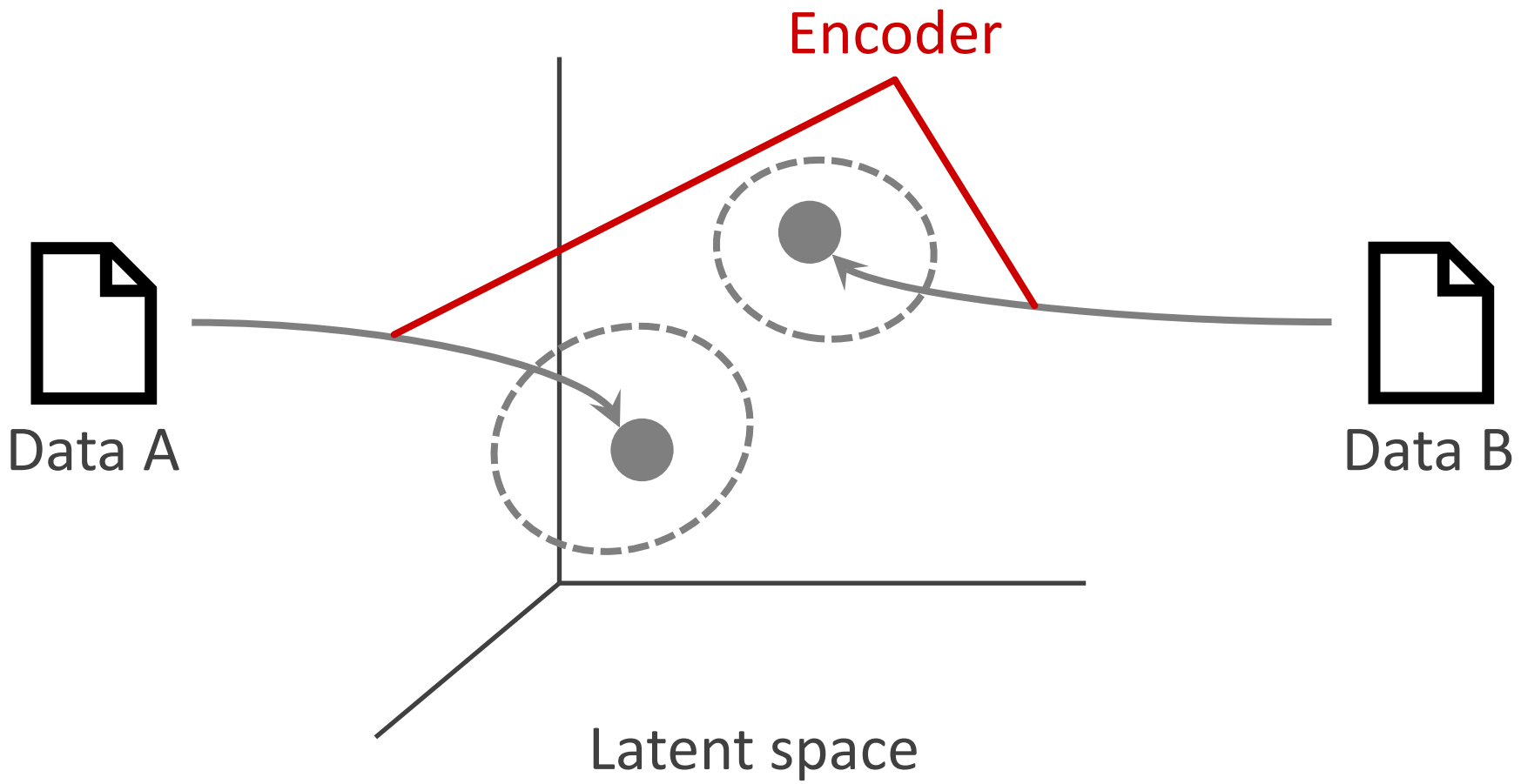
Takayuki Nakatsuka, Masahiro Hamasaki, Masataka Goto

National Institute of Advanced Industrial Science and Technology (AIST)

The 25th International Society for Music Information Retrieval (ISMIR) Conference

# Probabilistic Representation Learning

Provide intricate and diverse representations of data items



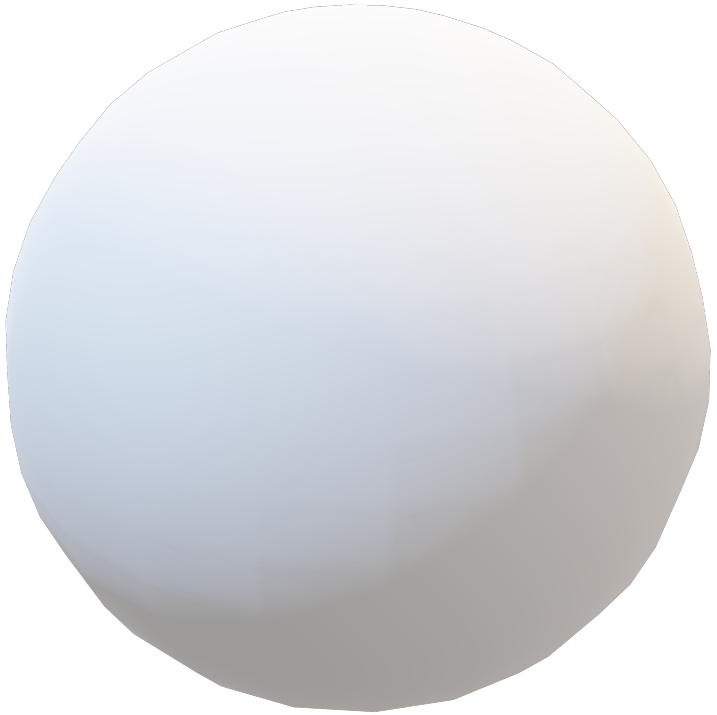
# Probabilistic Representation Learning

Our proposed loss functions are based on

**(1) Contrastive learning**

**(2) Optimal transport**

# Probabilistic Representation Learning on Hypersphere



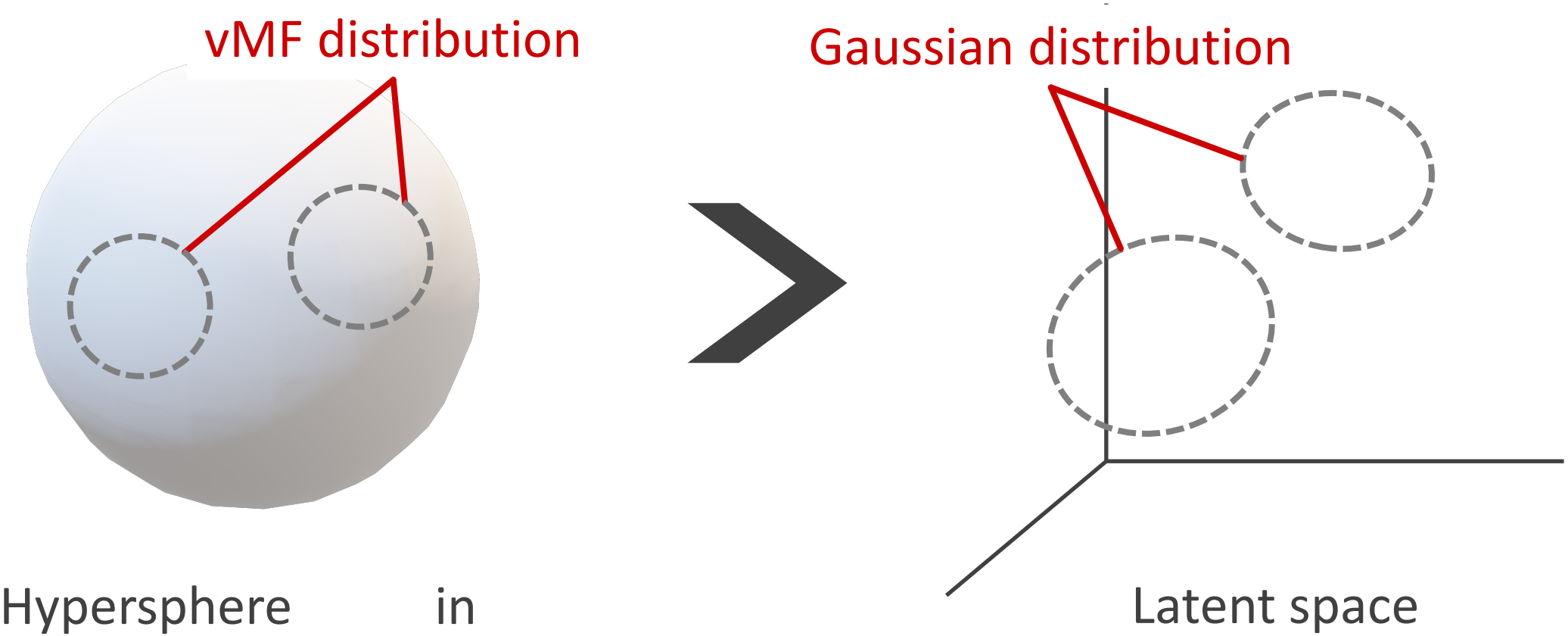
Hypersphere in

# Probabilistic Representation Learning

✓ Methods using the angular distance between distributions [Scott et al., 2021]

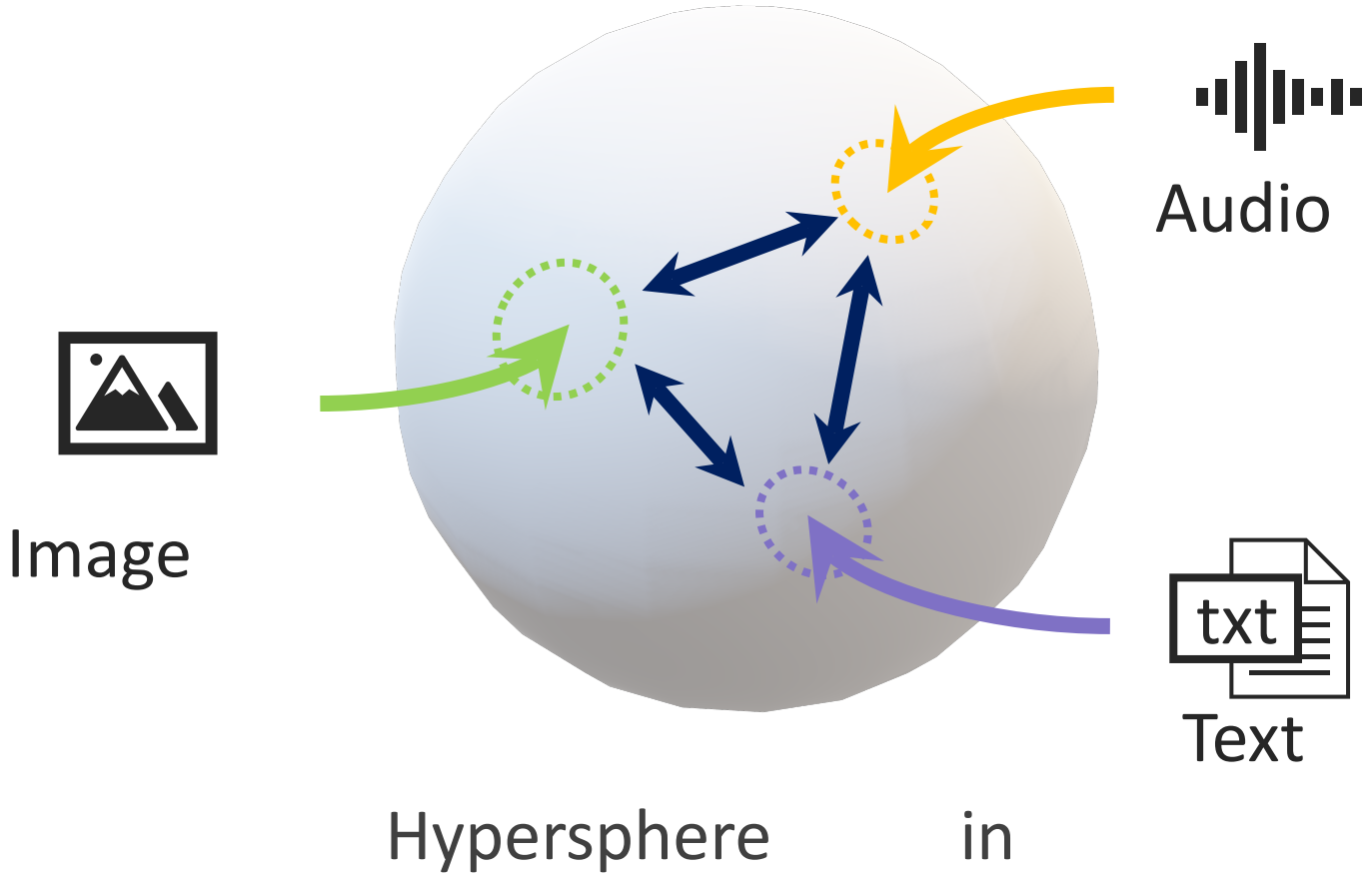
# Probabilistic Representation Learning

✓ Methods using a von Mises-Fisher (vMF) distribution [Li et al., 2021]



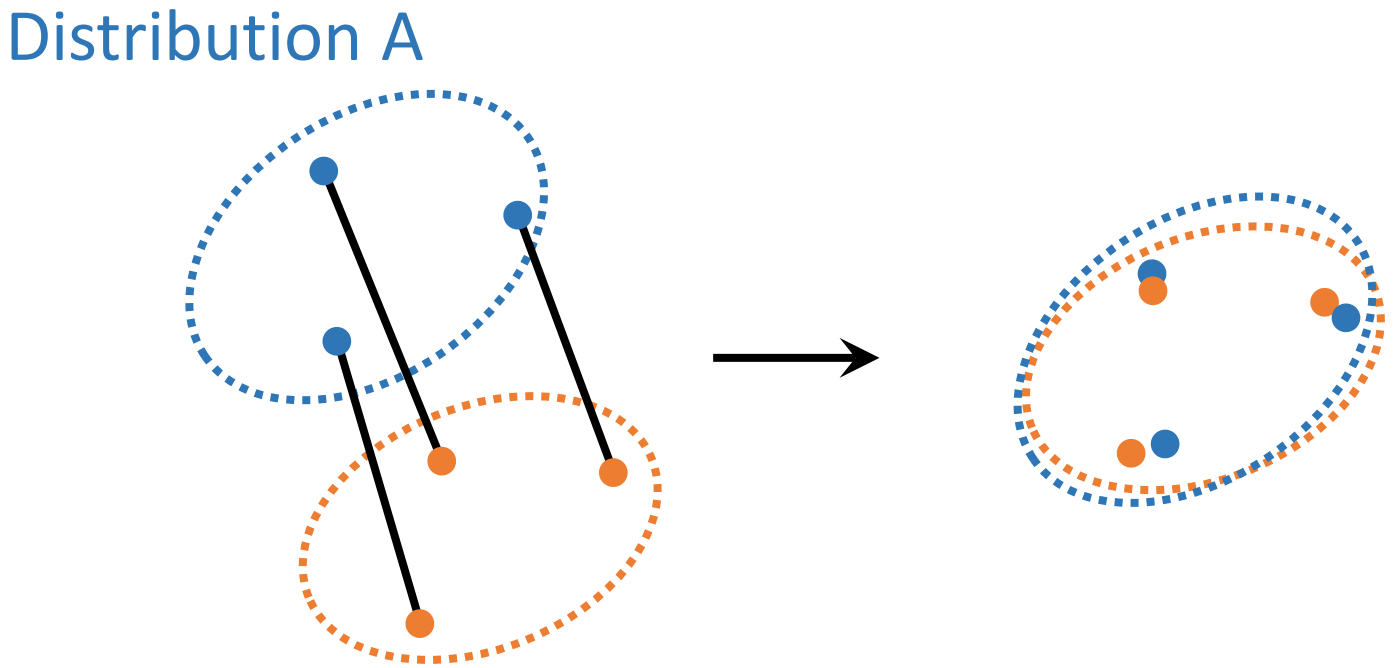
# Probabilistic Representation Learning on Hypersphere

We propose a contrastive loss function on a hypersphere for multiple modalities based on probabilistic contrastive learning [Kirchhof et al., 2023]



# Optimal Transport (OT)

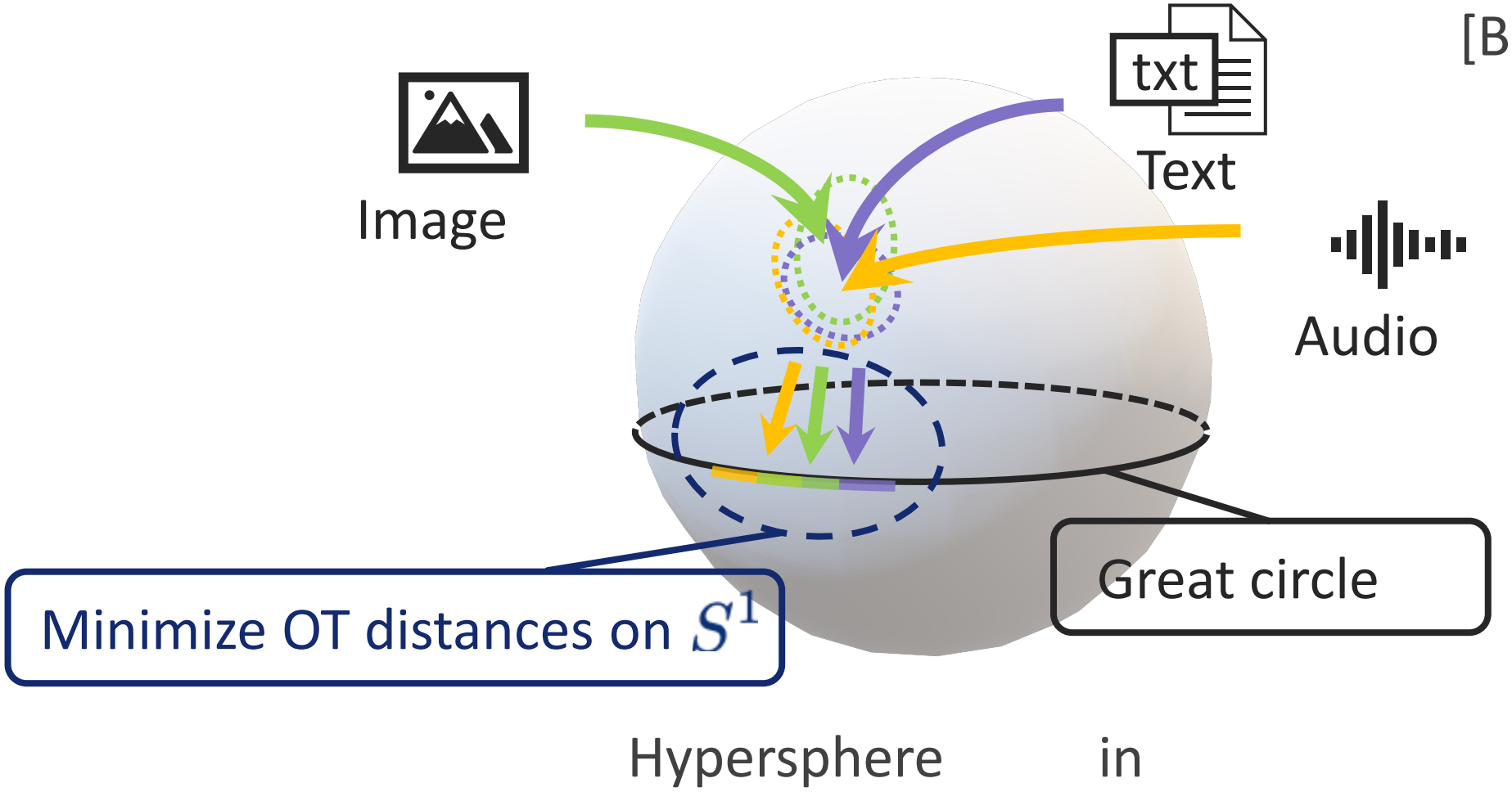
Provide a tool for calculating distances between distributions




● Samples from Distribution B ● Samples from Distribution B

# Probabilistic Representation Learning on Hypersphere

We propose a loss function based on Spherical Sliced-Wasserstein (SSW) distance [Bonet et al., 2023]



# Quantitative Evaluations

Comparison on YT8M-MusicVideo test dataset for Multimodal MIR

Mean Reciprocal Rank

Recall@$k$

Mean Rank

Proposed

Baseline

# Harnessing the Power of Distributions: Probabilistic Representation Learning on Hypersphere for Multimodal Music Information Retrieval

Takayuki Nakatsuka, Masahiro Hamasaki, Masataka Goto

National Institute of Advanced Industrial Science and Technology (AIST)

The 25th International Society for Music Information Retrieval (ISMIR) Conference