University of CINCINNATI

Earth Mover‘s Distance as a meassure of CP violation

CKM 2023, Santiago de Compostela, Spain

Based on [J. High Energ. Phys. 2023, 98 (2023)]

Ahmed Youssef
Ph.D. Candidate, University of Cincinnati
youssead@ucmail.uc.edu

Sept 19th, 2023

In collaboration with:
Adam Davis, Tony Menzo, and Jure Zupan

Outline

Current state of the art: Energy Test

Earth Mover' Distance (EMD) as test statistic
→ B decay

Modified EMD for large samples
→ D decay

Conclusion and Outlook

youssead@ucmail.uc.edu

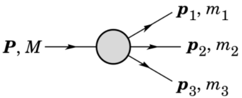
Direct CPV

How do we quantify direct CP violation in 3 body decay?

youssead@ucmail.uc.edu

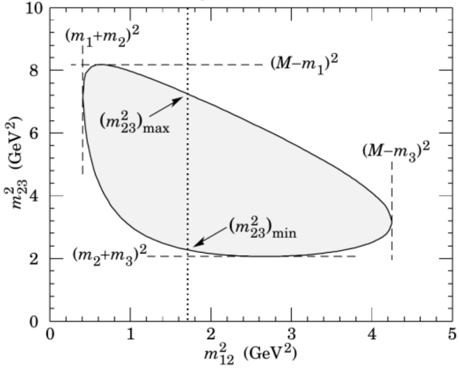
Direct CPV

How do we quantify direct CP violation in 3 body decay?



$$\frac{d\Gamma(B \to f)}{d\,\Omega} \neq \frac{d\Gamma(\bar{B} \to \bar{f})}{d\,\Omega}$$ ➡ **Visualize using Dalitz plots!**

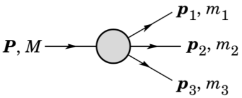
$$m_{ij}^2 = (p_i + p_j)^2$$



m_{ij} - invariant mas of a final state particle

Visualizes the differential decay rate across the phase space of the three-body decay

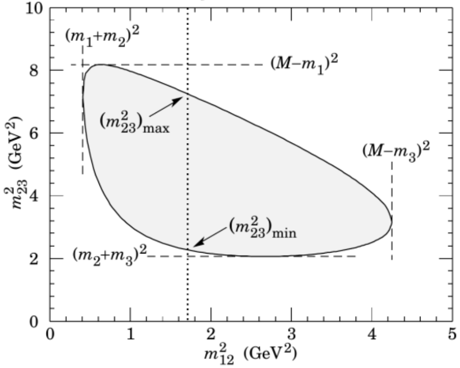
How do we quantify direct CP violation in 3 body decay?



$$\frac{d\Gamma(B\to f)}{d\,\Omega} \neq \frac{d\Gamma(\bar{B}\to\bar{f})}{d\,\Omega}$$

→ **Visualize using Dalitz plots!**

$$m_{ij}^2 = (p_i + p_j)^2$$



m_{ij} - **invariant mas of a final state particle**

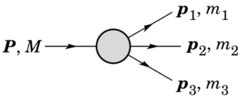
Visualizes the differential decay rate across the phase space of the three-body decay

Compare particle and its antiparticle distribution

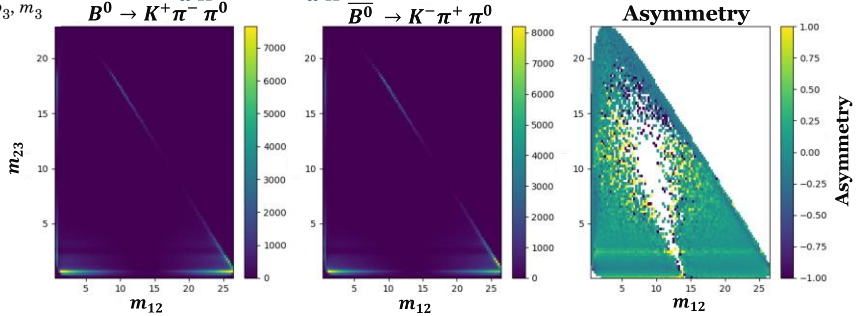
→ **Hints to CP violation**

Direct CPV

How do we quantify direct CP violation in 3 body decay?



$$\frac{d\Gamma(B \to f)}{d\,\Omega} \neq \frac{d\Gamma(\bar{B} \to \bar{f})}{d\,\Omega}$$ → **Visualize using Dalitz plots!**



→ **Direct CPV manifests as local density asymmetries between conjugate Dalitz plots!**

Current State of the art

Model the amplitude → **Very Complicated!**

Use a Test Statistic

- **Binned Test Statistic**
  - → **e.g. Miranda Test** → **Washes out Asymmetry!**
- **Unbinned Test Statistic**
  - → **e.g. Energy Test** → **Hard to interpret!**

Current State of the art

Model the amplitude → Very Complicated!

Use a Test Statistic

- **Binned Test S**
  - e … a Test → Washes out Asymmetry!
- **Unbinned Test Statistic**
  - e.g. Energy Test → Hard to interpret!

We need an alternative approach!

Motivation

What requirements do we need?

Motivation

What requirements do we need?

Is it highly sensitive to CP violation?

Can we interpret it?

Motivation

What requirements do we need?

Is it highly sensitive to CP violation?

Can we interpret it?

Earth Mover‘s Distance (EMD) as test statistic

Motivation

What requirements do we need?

Is it highly sensitive to CP violation?

Can we interpret it?

Earth Mover's Distance (EMD) as test statistic

→ **Comparable sensitivity to established method! (Comparison with the Energy Test)**

→ **Tells us which part of the Dalitz plot the CPV originated from!**

Energy Test

The energy test has already been successfully applied to search for CPV in multibody decays LHCB Collaboration, Phys. Lett. B 740 (2015) 158

Energy Test

The energy test has already been successfully applied to search for CPV in multibody decays

LHCB Collaboration, [Phys. Lett. B 740 (2015) 158]

Unbinned two-sample test utilizing a test statistic:

$$T = \sum_{i,j>i}^{N} \frac{\psi_{ij}}{N(N-1)} + \sum_{i,j>i}^{\bar{N}} \frac{\psi_{ij}}{\bar{N}(\bar{N}-1)} - \sum_{i,j}^{N,\bar{N}} \frac{\psi_{ij}}{N\bar{N}},$$

Sum over index i — **Sum over index j** — **Sum over indices i,j**

Weighting distance function:

$$\psi_{ij} \equiv \psi(d_{ij};\sigma) = e^{-d_{ij}^2/2\sigma^2}$$

$$B^0(\bar{B}^0) \to f\ (\bar{f})$$

i – B^0 sample
j – $\bar{B}^0$ sample

Energy Test

The energy test has already been successfully applied to search for CPV in multibody decays LHCB Collaboration, [Phys. Lett. B 740 (2015) 158]

Unbinned two-sample test utilizing a test statistic:

$$T = \underbrace{\sum_{i,j>i}^{N} \frac{\psi_{ij}}{N(N-1)}}_{\text{Sum over index i}} + \underbrace{\sum_{i,j>i}^{\bar{N}} \frac{\psi_{ij}}{\bar{N}(\bar{N}-1)}}_{\text{Sum over index j}} - \underbrace{\sum_{i,j}^{N,\bar{N}} \frac{\psi_{ij}}{N\bar{N}}}_{\text{Sum over indices i,j}},$$

Weighting distance function:

$$\psi_{ij} \equiv \psi(d_{ij};\sigma) = e^{-d_{ij}^2/2\sigma^2}$$

$$B^0(\bar{B}^0) \rightarrow f\ (\bar{f})$$

i – B^0 sample
j – $\bar{B}^0$ sample

Events from two identical distribution ➡ T close to zero

Events from two dissimilar distribution ➡ T is non zero

Energy Test

The energy test has already been successfully applied to search for CPV in multibody decays

LHCB Collaboration, [Phys. Lett. B 740 (2015) 158]

Unbinned two-sample test utilizing a test statistic:

$$T = \sum_{i,j>i}^{N} \frac{\psi_{ij}}{N(N-1)} + \sum_{i,j>i}^{\bar{N}} \frac{\psi_{ij}}{\bar{N}(\bar{N}-1)} - \sum_{i,j}^{N,\bar{N}} \frac{\psi_{ij}}{N\bar{N}},$$

Sum over index i — **Sum over index j** — **Sum over indices i,j**

Weighting distance function:

$$\psi_{ij} \equiv \psi(d_{ij};\sigma) = e^{-d_{ij}^2/2\sigma^2}$$

$B^0(\bar{B}^0) \to f\ (\bar{f})$

i – B^0 sample
j – $\bar{B}^0$ sample

Events from two identical distribution → **T close to zero**

Events from two dissimilar distribution → **T is non zero**

Perform a hypothesis test to obtain a p-value!

Energy Test

The energy test has already been successfully applied to search for CPV in multibody decays

LHCB Collaboration, Phys. Lett. B 740 (2015) 158

Unbinned two-sample test utilizing a test statistic:

$$T = \sum_{i,j>i}^{N} \frac{\psi_{ij}}{N(N-1)} + \sum_{i,j>i}^{\bar{N}} \frac{\psi_{ij}}{\bar{N}(\bar{N}-1)} - \sum^{N,\bar{N}} \frac{\psi_{ij}}{N\bar{N}}$$

Sum over index i

j

Sum over indices i,j

Energy Test is not interpretable!

Weighting distance function:

$$\psi_{ij} \equiv \psi(d_{ij};\sigma) = e^{-d_{ij}^2/2\sigma^2}$$

$B^0(\bar{B}^0) \rightarrow f\ (\bar{f})$

i – B^0 sample
j – $\bar{B}^0$ sample

Events from two identical distribution → T close to zero

Events from two dissimilar distribution → T is non zero

Perform a hypothesis test to obtain a p-value!

Energy Test

The energy test has already been successfully applied to search for CPV in multibody decays LHCB Collaboration, [Phys. Lett. B 740 (2015) 158]

Unbinned two-sample test utilizing a test statistic:

$$T = \sum_{i,j>i}^{N} \frac{\psi_{ij}}{N(N-1)} + \ldots$$

Sum over index i

Wei... ting distance ...nction

$$\ldots_{ij};\sigma) = e^{-d_{ij}^2/2\sigma^2}$$

$(\bar{f})$

i – B^0 sample
j – $\bar{B}^0$ sample

...indices i,j

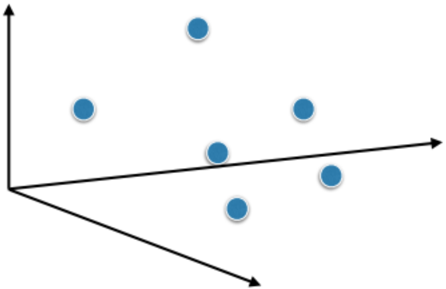
Can we come up with a more interpretable test statistic?

Events from two identical distribution → T close to zero

Events from two dissimilar distribution → T is non zero

Perform a hypothesis test to obtain a p-value!

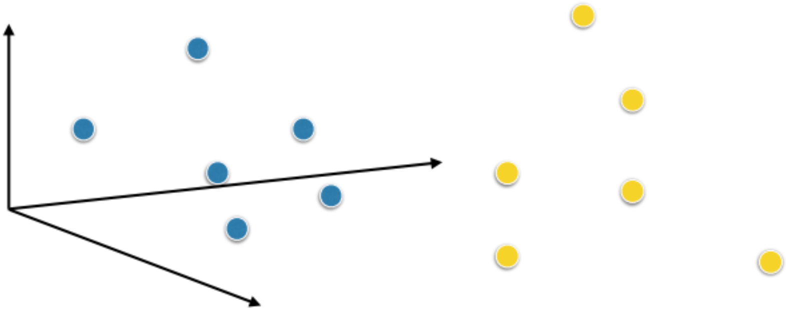
Optimal Transport (OT)



Example from: Marco Cuturi, MLSS summer school presentaation

youssead@ucmail.uc.edu

Optimal Transport (OT)

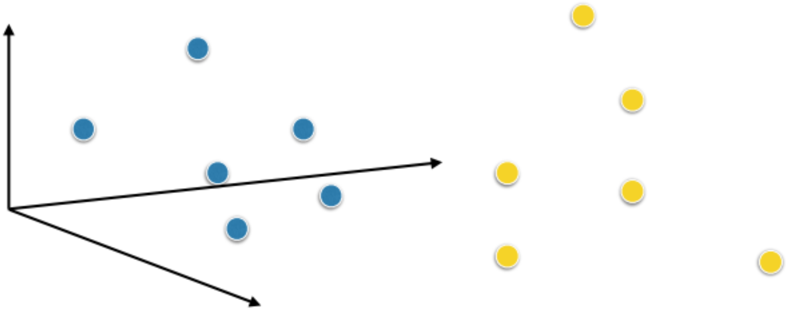


Example from: Marco Cuturi, MLSS summer school presentaation

youssead@ucmail.uc.edu

Optimal Transport (OT)

What is the most optimal way to move one sample to another?

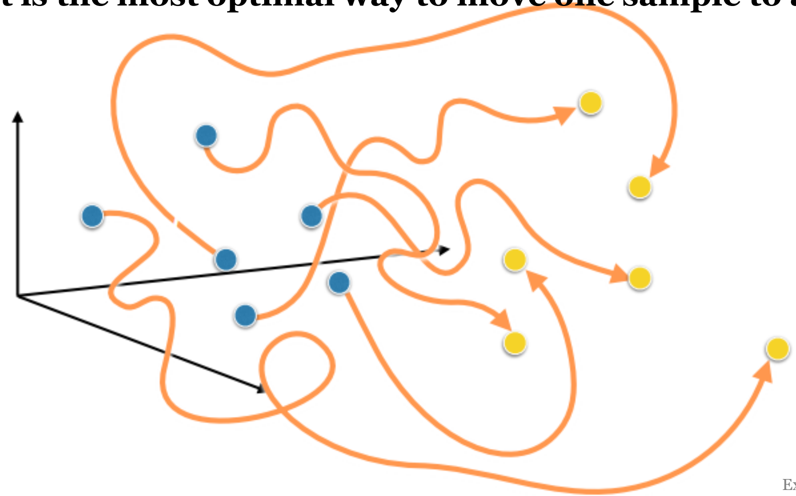


Example from: Marco Cuturi, MLSS summer school presentaation

youssead@ucmail.uc.edu

Optimal Transport (OT)

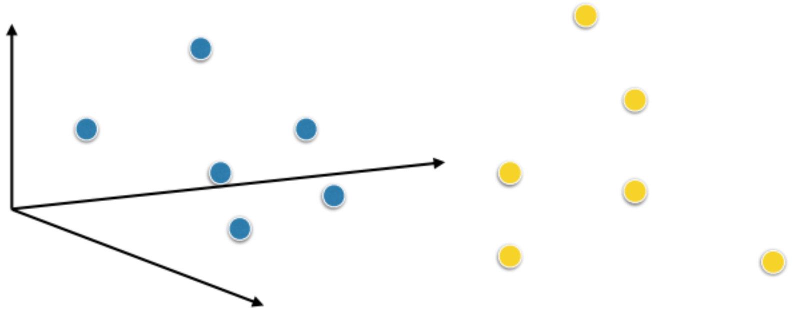
What is the most optimal way to move one sample to another?



Example from: Marco Cuturi, MLSS summer school presentaation

Optimal Transport (OT)

What is the most optimal way to move one sample to another?

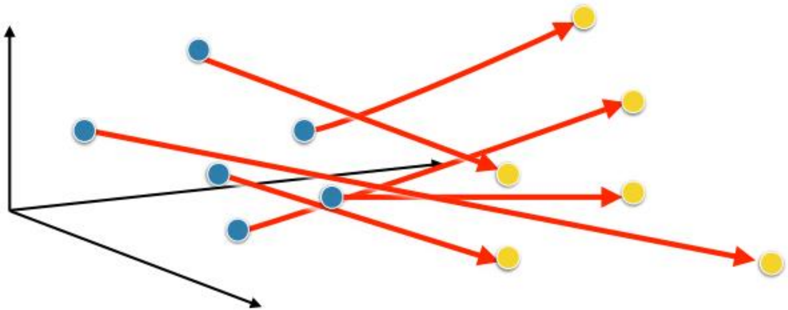


Goal of OT: Find the most "natural" way to move points

Example from: Marco Cuturi, MLSS summer school presentaation

Optimal Transport (OT)

What is the most optimal way to move one sample to another?

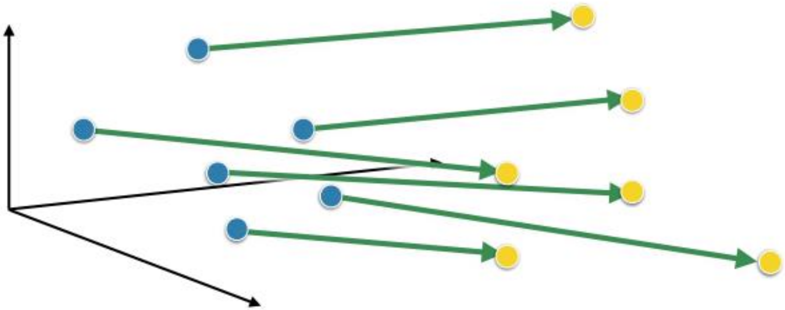


Goal of OT: Find the most "natural" way to move points

Example from: Marco Cuturi, MLSS summer school presentaation

Optimal Transport (OT)

What is the most optimal way to move one sample to another?

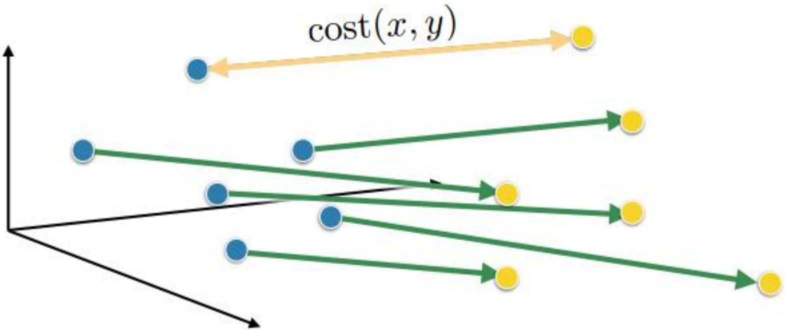


Goal of OT: Find the most “natural” way to move points

Example from: Marco Cuturi, MLSS summer school presentaation

Optimal Transport (OT)

What is the most optimal way to move one sample to another?

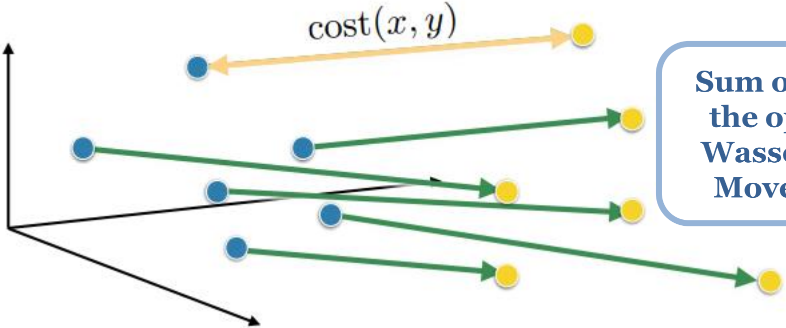


Goal of OT: Find the most "natural" way to move points

Example from: Marco Cuturi, MLSS summer school presentaation

Optimal Transport (OT)

What is the most optimal way to move one sample to another?



Sum over the cost of the optimal path = Wasserstein (Earth Mover's) distance

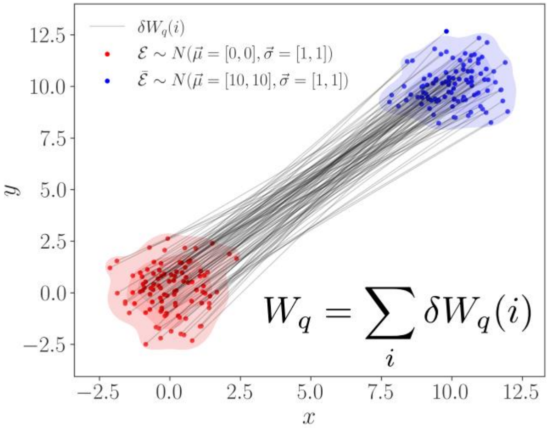
Goal of OT: Find the most "natural" way to move points

Example from: Marco Cuturi, MLSS summer school presentaation

Wasserstein Distance

Wasserstein distance (WD)

$$W_q(\mathcal{E},\bar{\mathcal{E}}) = \left[\min_{\{f_{ij}\geq 0\}} \sum_{i=1}^{N}\sum_{j=1}^{\bar{N}} f_{ij}\left(\hat{d}_{ij}\right)^q\right]^{1/q}$$



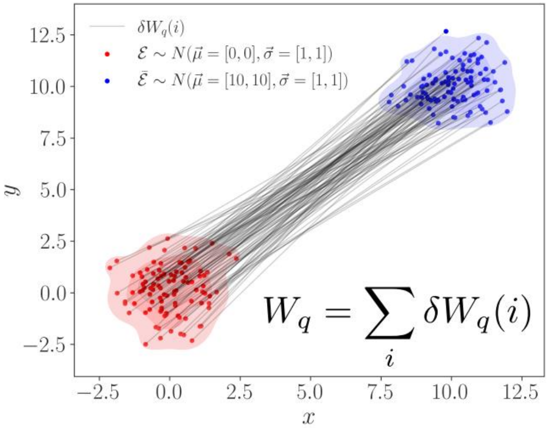
Wasserstein Distance

Wasserstein distance (WD)

$$W_q(\mathcal{E},\bar{\mathcal{E}}) = \left[\min_{\{f_{ij}\geq 0\}} \sum_{i=1}^{N}\sum_{j=1}^{\bar{N}} f_{ij}\left(\hat{d}_{ij}\right)^q\right]^{1/q}$$

Events from two identical distribution → **Small Wq**

Events from two dissimilar distribution → **Larger Wq**



Wasserstein Distance

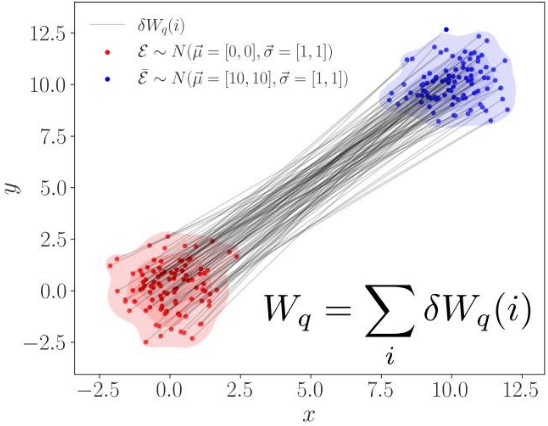
Wasserstein distance (WD)

$$W_q(\mathcal{E},\bar{\mathcal{E}}) = \left[\min_{\{f_{ij}\geq 0\}} \sum_{i=1}^{N}\sum_{j=1}^{\bar{N}} f_{ij}\left(\hat{d}_{ij}\right)^q\right]^{1/q}$$

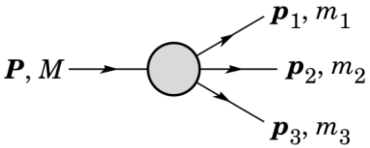
Events from two identical distribution → **Small Wq**

Events from two dissimilar distribution → **Larger Wq**

Perform a hypothesis test to obtain a p-value!

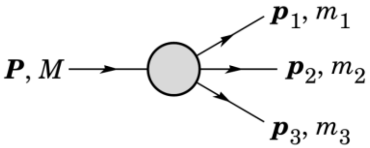


Application to 3 Body decay



youssead@ucmail.uc.edu

Application to 3 Body decay



Sample Size = $\sim 10^3$

$$B^0 \to K^+\pi^-\pi^0$$
$$\bar{B}^0 \to K^-\pi^+\pi^0$$

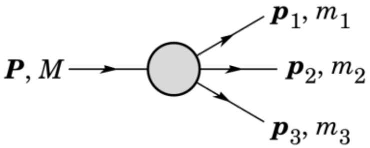
EMD as a test statistic

Sample Size = $\sim 10^5 - 10^6$

$$D^0 \to \pi^+\pi^-\pi^0$$
$$\bar{D}^0 \to \pi^-\pi^+\pi^0$$

„Modified“ EMD as a test statistic

Application to 3 Body decay



Sample Size = $\sim 10^3$

$$B^0 \to K^+\pi^-\pi^0$$
$$\bar{B}^0 \to K^-\pi^+\pi^0$$

EMD as a test statistic

Sample Size = $\sim 10^5 - 10^6$

$$D^0 \to \pi^+\pi^-\pi^0$$
$$\bar{D}^0 \to \pi^-\pi^+\pi^0$$

„Modified“ EMD as a test statistic

youssead@ucmail.uc.edu

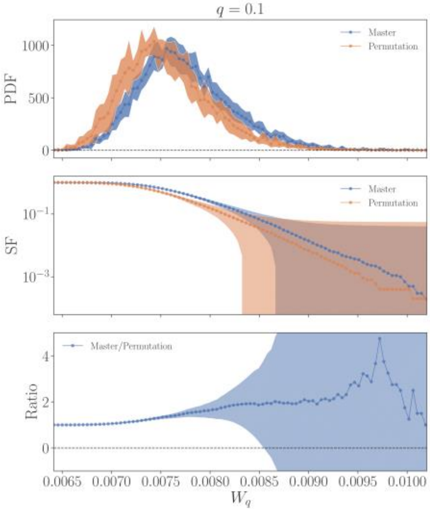
Hypothese Test

Obtain the null hypotheses pdf from your test statistic by calculating it n times

youssead@ucmail.uc.edu

Hypothese Test

Obtain the null hypotheses pdf from your test statistic by calculating it n times



Permutation Method

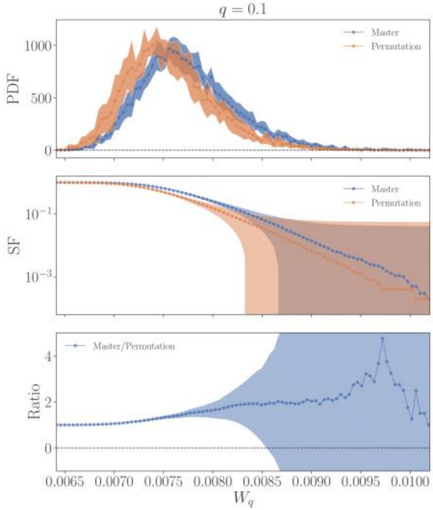
- Permuting the original B^0 and $\bar{B}^0$ samples
- Calculate the test statistic for each permutation

Master Method

- Generate an ensemble of B^0 and $\bar{B}^0$ decay event samples, using the B^0 decay model for both
- Calculate the test statistic for each sample pair

Hypothese Test

Obtain the null hypotheses pdf from your test statistic by calculating it n times



Permutation Method

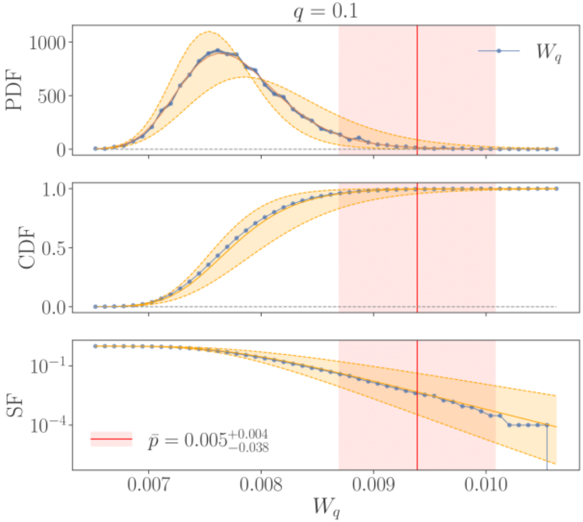
- Permuting the original B^0 and $\bar{B}^0$ samples
- Calculate the test statistic for each permutation

Master Method

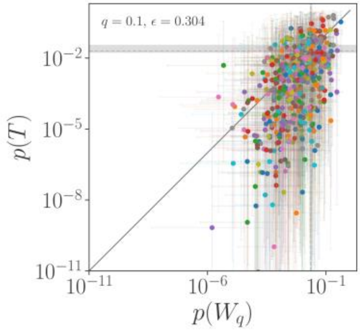
- Generate an ensemble of B^0 and $\bar{B}^0$ decay event samples, using the B^0 decay model for both
- Calculate the test statistic for each sample pair

Compare the sensitivity of W_q and Energy test using the master method

Results for B decay

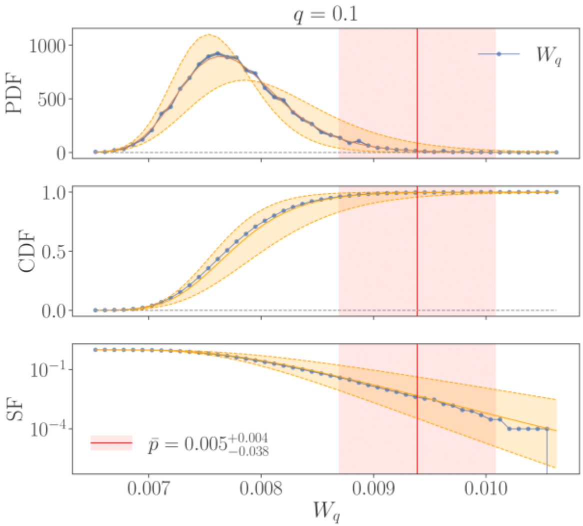


$$\epsilon \equiv \frac{1}{N_e} \sum_{i=1}^{N_e} \begin{cases} +1 & p_i(W_q) < p_i(T), \\ 0 & \text{otherwise,} \end{cases}$$

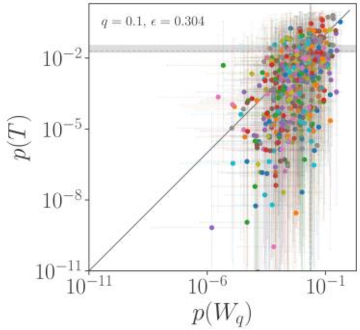


youssead@ucmail.uc.edu

Results for B decay



$$\epsilon \equiv \frac{1}{N_e}\sum_{i=1}^{N_e}\begin{cases}+1 & p_i(W_q) < p_i(T),\\ 0 & \text{otherwise},\end{cases}$$



We have a similar sensitivity as the ET

youssead@ucmail.uc.edu

Can we locate the CP violation?

Can we locate the CP violation?

EMD traces the variation of the CP asymmetry across the Dalitz plot!

Can we locate the CP violation?

EMD traces the variation of the CP asymmetry across the Dalitz plot!

CP asymmetry:

$B^0(\bar{B}^0) \to K^+\pi^-\pi^0\ (K^-\pi^-\pi^0)$

$$\mathcal{A}_{\mathrm{CP}}(s_{12}, s_{13}) = \frac{d\bar{\Gamma}(\bar{s}_{12}, \bar{s}_{13}) - d\Gamma(s_{12}, s_{13})}{d\bar{\Gamma}(\bar{s}_{12}, \bar{s}_{13}) + d\Gamma(s_{12}, s_{13})}$$

BaBar amplitude model

BarBar Collaboration, Phys. Rev. D 83 (2011) 112010

youssead@ucmail.uc.edu

Can we locate the CP violation?

EMD traces the variation of the CP asymmetry across the Dalitz plot!

CP asymmetry:

$B^0(\bar{B}^0) \to K^+\pi^-\pi^0\ (K^-\pi^-\pi^0)$

$$\mathcal{A}_{\mathrm{CP}}(s_{12}, s_{13}) = \frac{d\bar{\Gamma}(\bar{s}_{12}, \bar{s}_{13}) - d\Gamma(s_{12}, s_{13})}{d\bar{\Gamma}(\bar{s}_{12}, \bar{s}_{13}) + d\Gamma(s_{12}, s_{13})}$$

BaBar amplitude model

BarBar Collaboration, Phys. Rev. D 83 (2011) 112010

$$W_q^q = \sum_i \delta W_q(i)$$

EMD asymmetry:

$$\mathcal{W}_{\mathrm{CP}}^q(s_{12}, s_{13}) = \frac{\sum_{\bar{i}} \delta\bar{W}_q(\bar{i}) - \sum_i \delta W_q(i)}{\sum_{\bar{i}} \delta\bar{W}_q(\bar{i}) + \sum_i \delta W_q(i)}$$

Can we locate the CP violation?

EMD traces the variation of the CP asymmetry across the Dalitz plot!

$B^0(\bar{B}^0) \to K^+\pi^-\pi^0\ (K^-\pi^+\pi^0)$

CP asymmetry:

$$\mathcal{A}_{CP}(s_{12}, s_{13}) = \frac{d\bar{\Gamma}(\bar{s}_{12}, \bar{s}_{13}) - d\Gamma(s_{12}, s_{13})}{d\bar{\Gamma}(\bar{s}_{12}, \bar{s}_{13}) + d\Gamma(s_{12}, s_{13})}$$

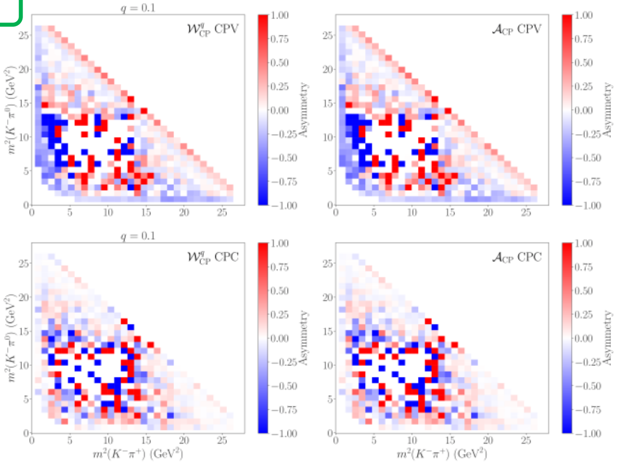
BaBar amplitude model

BarBar Collaboration, Phys. Rev. D 83 (2011) 112010

$$W_q^q = \sum_i \delta W_q(i)$$

EMD asymmetry:

$$\mathcal{W}_{CP}^q(s_{12}, s_{13}) = \frac{\sum_{\bar{i}} \delta \bar{W}_q(\bar{i}) - \sum_i \delta W_q(i)}{\sum_{\bar{i}} \delta \bar{W}_q(\bar{i}) + \sum_i \delta W_q(i)}$$



Can we locate the CP violation?

EMD traces the variation of the CP asymmetry across the Dalitz plot!

$B^0(\bar{B}^0) \rightarrow K^+\pi^-\pi^0\ (K^-\pi^+\pi^0)$

CP asymmetry:

$$\mathcal{A}_{CP}(s_{12}, s_{13}) = \frac{d\bar{\Gamma}(\bar{s}_{12}, \bar{s}_{13}) - d\Gamma(s_{12}, s_{13})}{d\bar{\Gamma}(\bar{s}_{12}, \bar{s}_{13}) + d\Gamma(s_{12}, s_{13})}$$

BaBar amplitude model

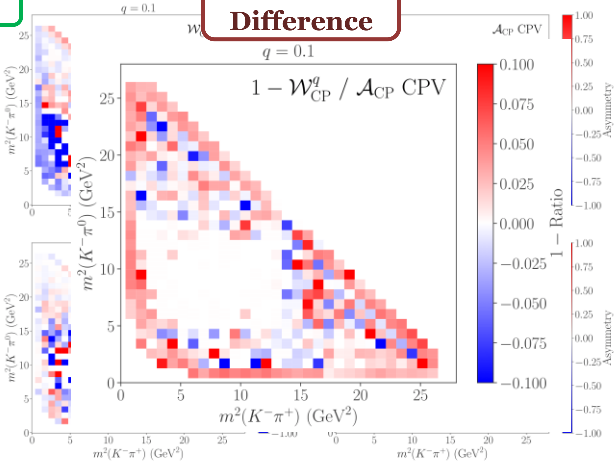
BarBar Collaboration, Phys. Rev. D 83 (2011) 112010

$$W_q^q = \sum_i \delta W_q(i)$$

EMD asymmetry:

$$\mathcal{W}_{CP}^q(s_{12}, s_{13}) = \frac{\sum_{\bar{i}} \delta\bar{W}_q(\bar{i}) - \sum_i \delta W_q(i)}{\sum_{\bar{i}} \delta\bar{W}_q(\bar{i}) + \sum_i \delta W_q(i)}$$

Relative Difference



youssead@ucmail.uc.edu

What about larger Data sets?

What about larger Data sets?

$D^0 \to \pi^+\pi^-\pi^0$
$\overline{D}^0 \to \pi^-\pi^+\pi^0$

→ **Very small non-zero CP violation**

→ **Studied at the LHCb using the ET**

What about larger Data sets?

$$D^0 \to \pi^+\pi^-\pi^0$$
$$\overline{D}^0 \to \pi^-\pi^+\pi^0$$

→ **Very small non-zero CP violation**

→ **Studied at the LHCb using the ET**

Can we still use the EMD?

What about larger Data sets?

$$D^0 \to \pi^+\pi^-\pi^0$$
$$\overline{D}^0 \to \pi^-\pi^+\pi^0$$

→ **Very small non-zero CP violation**

→ **Studied at the LHCb using the ET**

Can we still use the EMD?

Yes, but …

youssead@ucmail.uc.edu

What about larger Data sets?

$$D^0 \to \pi^+\pi^-\pi^0$$
$$\overline{D}^0 \to \pi^-\pi^+\pi^0$$

→ Very small non-zero CP violation

→ Studied at the LHCb using the ET

Can we still use the EMD?

Yes, but ...

→ Computationally expensive

→ Very memory intensive

What about larger Data sets?

$D^0 \to \pi^+\pi^-\pi^0$
$\bar{D}^0 \to \pi^-\pi^+\pi^0$

➡ Very small no… o CP violation

➡ Stud… using the ET

We need a more efficient solution for larger data sets!

Ca… the EMD?

Yes, but …

➡ Computationally expensive

➡ Very memory intensive

Alternative Solution

We propose two solutions

Binned Wasserstein distance

Not part of this talk

Sliced Wasserstein distance

youssead@ucmail.uc.edu

Sliced Wasserstein distance

Use Sliced Wasserstein Distance as test statistic!

youssead@ucmail.uc.edu

Sliced Wasserstein distance

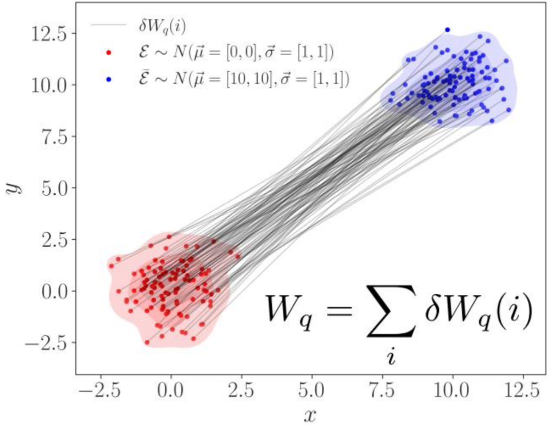
Use Sliced Wasserstein Distance as test statistic!

Wasserstein distance (WD)

$$W_q(\mathcal{E},\bar{\mathcal{E}}) = \left[\min_{\{f_{ij}\geq 0\}} \sum_{i=1}^{N}\sum_{j=1}^{\bar{N}} f_{ij}\left(\hat{d}_{ij}\right)^q\right]^{1/q}$$

Sliced Wasserstein distance

- Projects high dimensional data into one dimensional “slices”
- WD in 1D has a closed form solution
  - Sorted Difference of the two samples



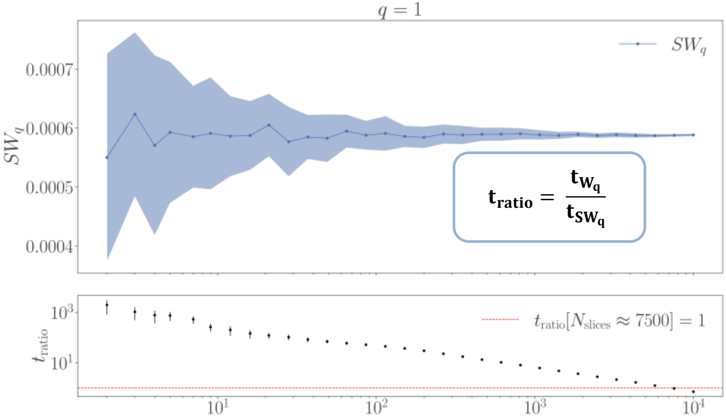
Sliced Wasserstein distance

How many slices do we need to converge?

youssead@ucmail.uc.edu

Sliced Wasserstein distance

How many slices do we need to converge?

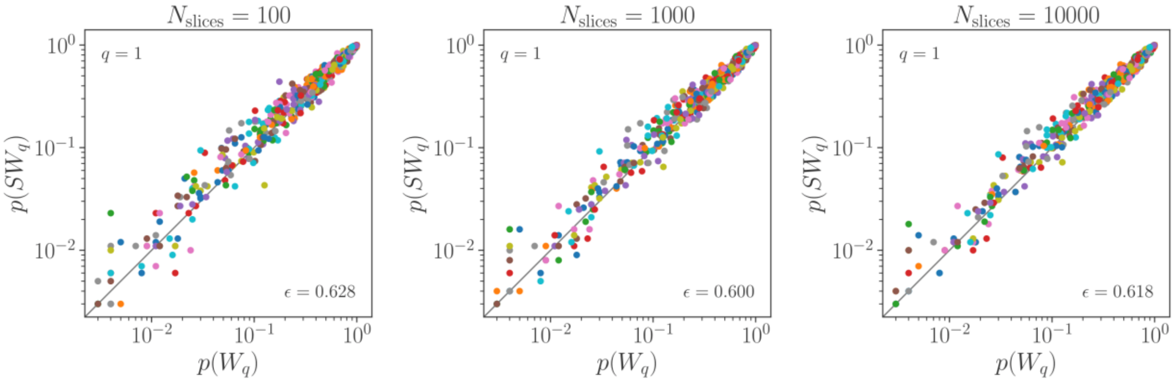


$$t_{ratio} = \frac{t_{W_q}}{t_{SW_q}}$$

- **Starts converging at 1000 slices**
- **1000 slices: speed up of a factor ~7 over W_q**
- **Comp of N_{Slices} can be done in parallel!**
- **Does not require large memory resources!**

Sliced Wasserstein distance

Comparison with W_q



Conclusion and Outlook

EMD is a robust, model independent, and unbinned test statistic for CPV!

highly sensitive to CPV

Interpretable

Future work

- Improving the test further
- Time-dependent CPV
- Flavor Violation

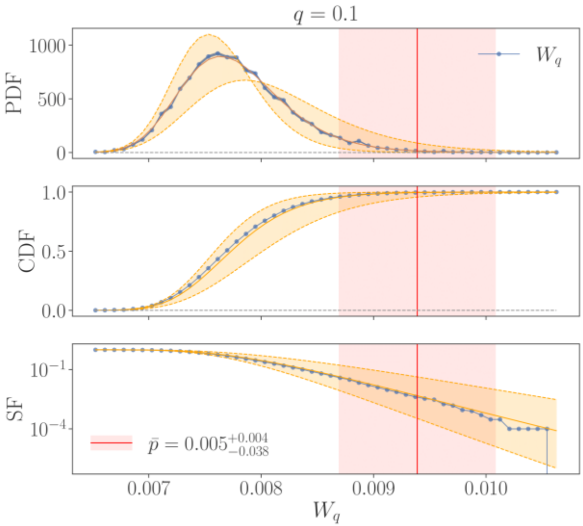
Public code:

https://github.com/adamdddave/EMD4CPV

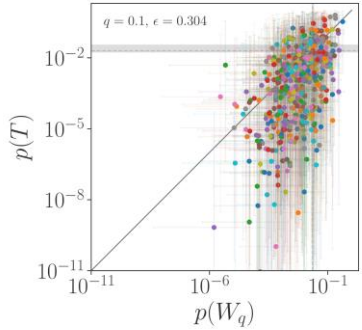
youssead@ucmail.uc.edu

Back up

Results for B decay



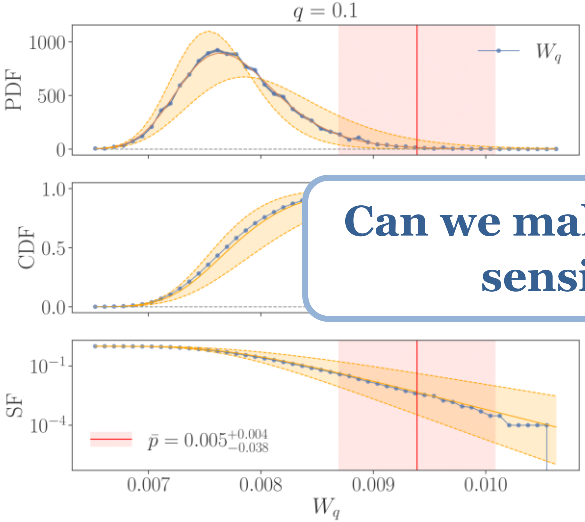
$$\epsilon \equiv \frac{1}{N_e} \sum_{i=1}^{N_e} \begin{cases} +1 & p_i(W_q) < p_i(T), \\ 0 & \text{otherwise,} \end{cases}$$



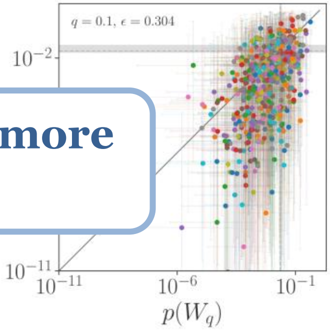
$\delta W_q > 0$: W_q receives contributions from non CPV areas

→ **Slightly smaller sensitivity than ET**

Results for B decay



$$\epsilon \equiv \frac{1}{N_e} \sum_{i=1}^{N_e} \begin{cases} +1 & p_i(W_q) < p_i(T), \\ 0 & \text{otherwise}, \end{cases}$$

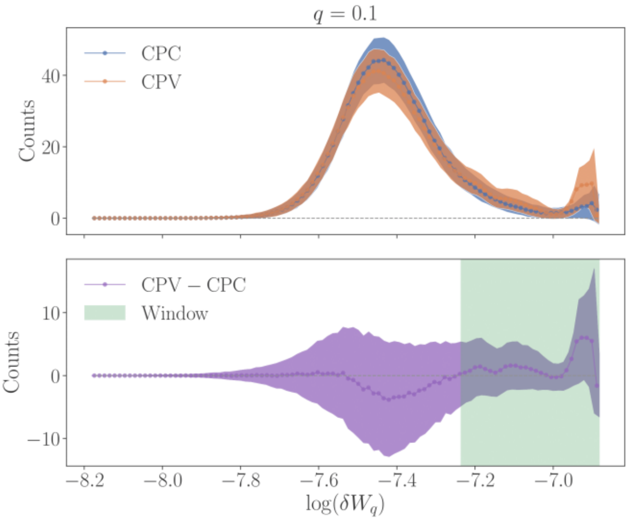


Can we make W_q more sensitive?

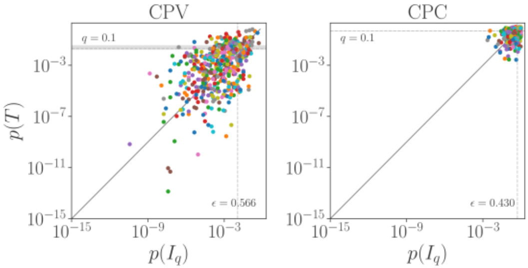
$\delta W_q > 0$: W_q receives contributions from non CPV areas

→ Slightly smaller sensitivity than ET

Windowed Wasserstein distance

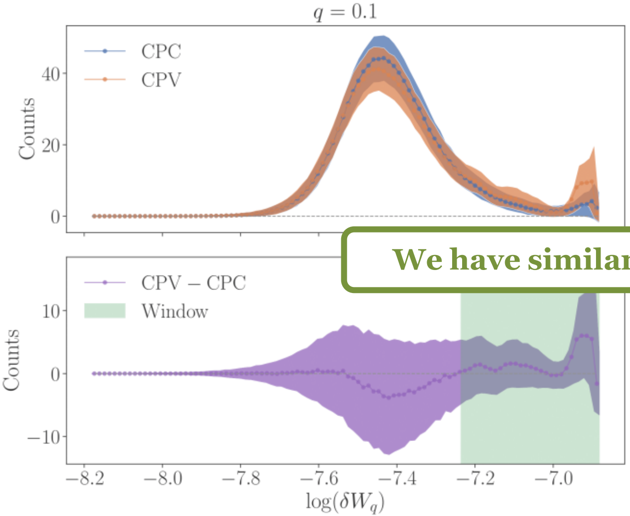


$$w(x) = \begin{cases} +1 & x \in [\delta W^{\text{win}}_{\text{min}}, \delta W^{\text{win}}_{\text{max}}], \\ 0 & \text{otherwise.} \end{cases}$$



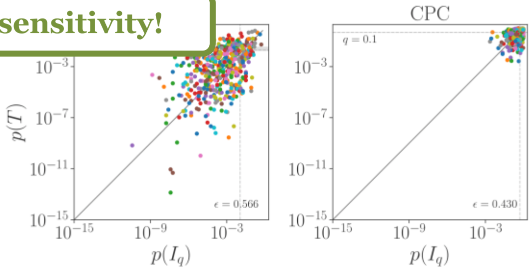
youssead@ucmail.uc.edu

Windowed Wasserstein distance



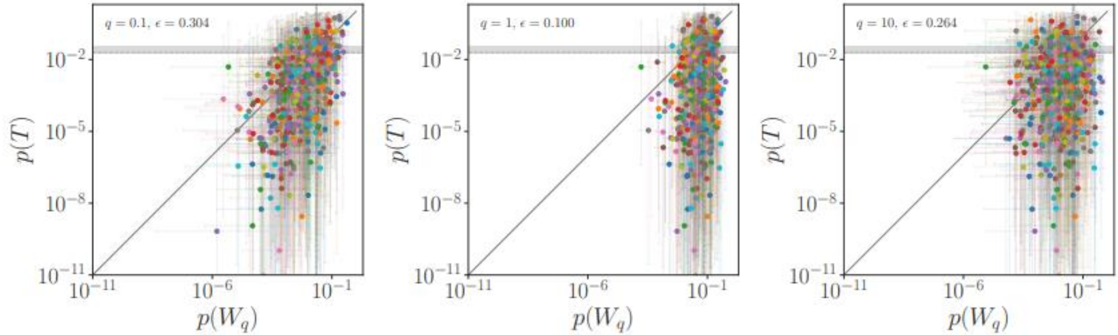
$$w(x) = \begin{cases} +1 & x \in [\delta W^{\text{win}}_{\text{min}}, \delta W^{\text{win}}_{\text{max}}], \\ 0 & \text{otherwise.} \end{cases}$$

We have similar sensitivity!

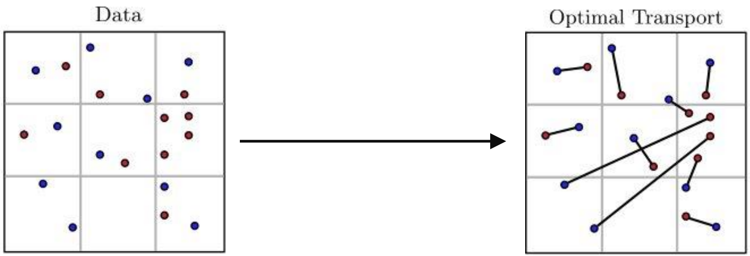


youssead@ucmail.uc.edu

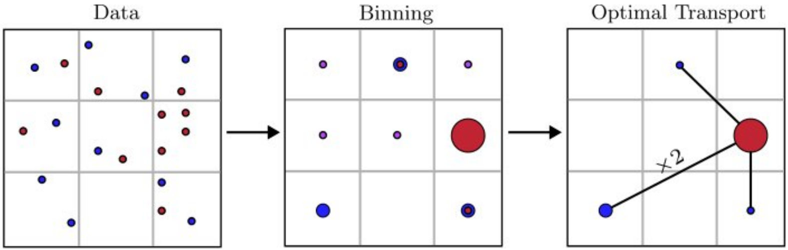
Results for B decay



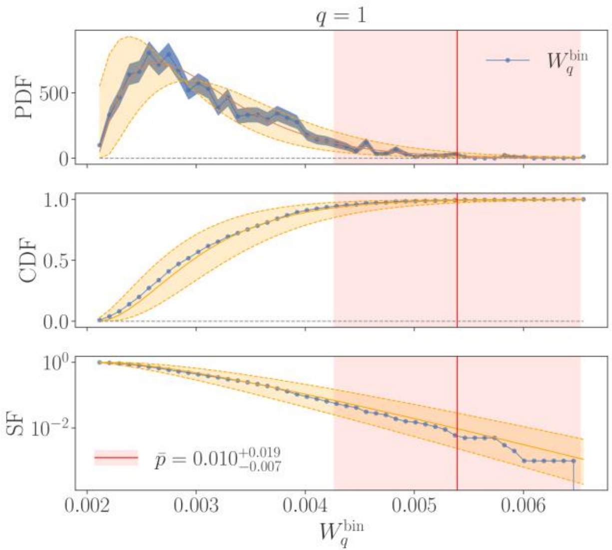
“Normal” Wasserstein distance

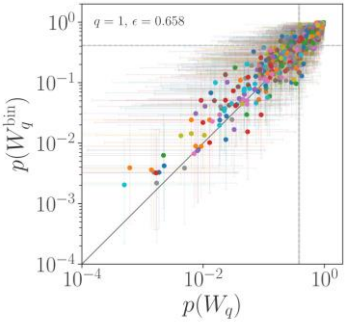


Binned Wasserstein distance

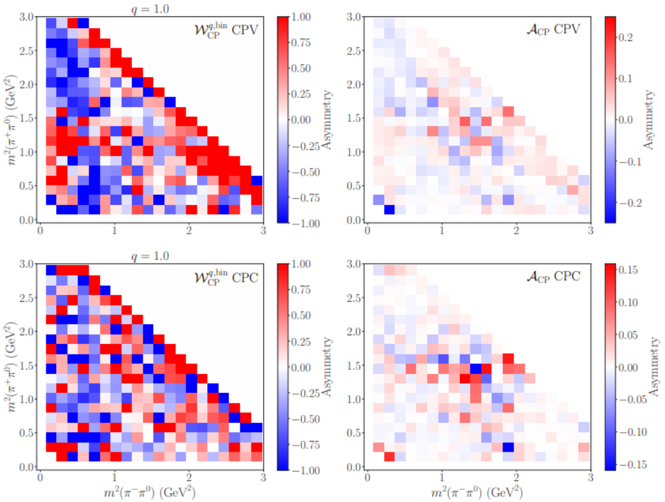


Binned Wasserstein distance





Binned EMD



youssead@ucmail.uc.edu