The Wasserstein distance for CP-violation

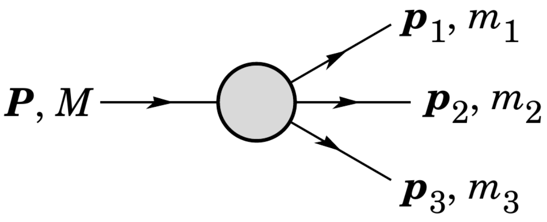
PIKIMO 13

Tony Menzo
PhD Student, University of Cincinnati

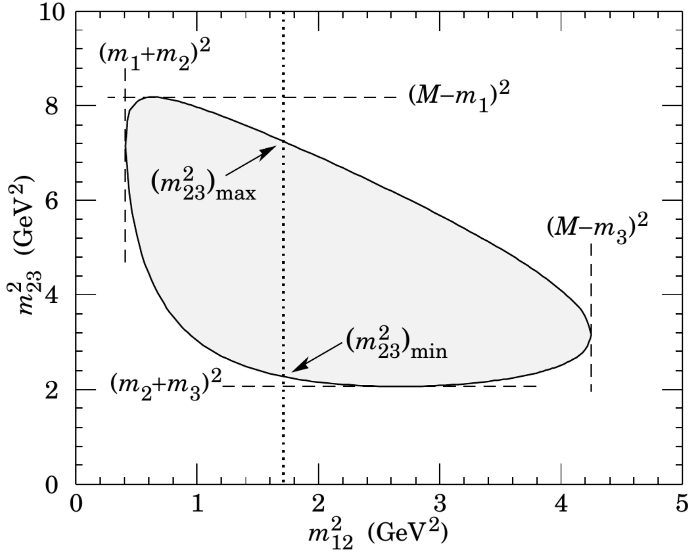
In collaboration with Adam Davis, Ahmed Youssef, and Jure Zupan

Dalitz Plot Analysis – Theory

Consider a 3-body decay:



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

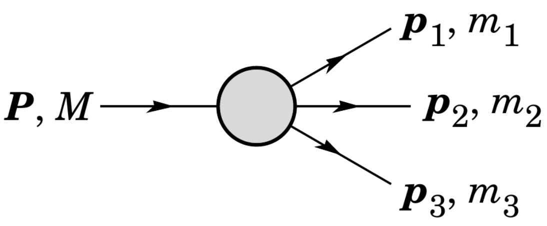


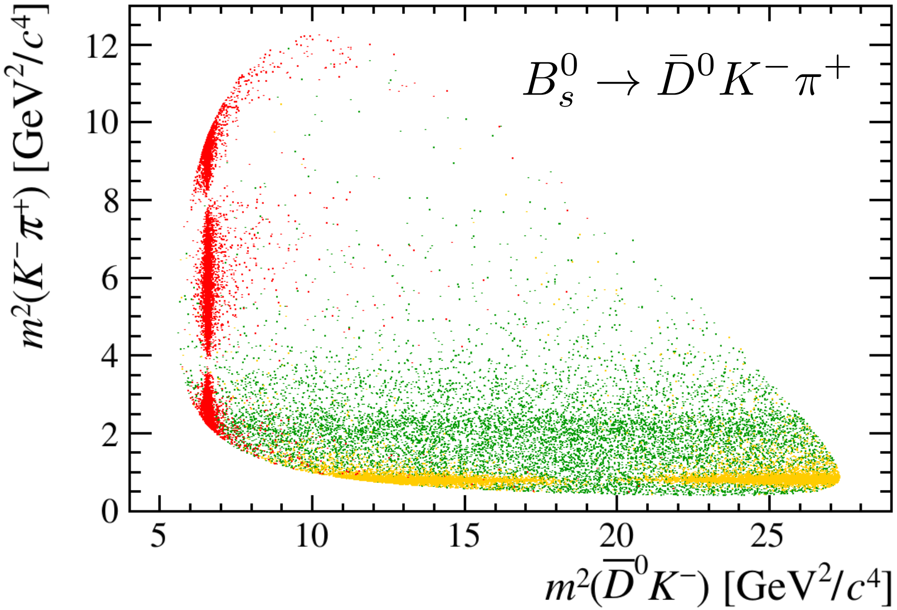
What to remember:

$$d\Gamma \propto |\mathcal{M}|^2 dm_{12}^2 dm_{23}^2$$

Dalitz Plot Analysis – Theory

Consider a 3-body decay:





What to remember:

$$d\Gamma \propto |\mathcal{M}|^2 dm_{12}^2 dm_{23}^2$$

Direct CP violation in the Dalitz Plot

Neccesary condition for direct CPV: $\frac{\mathcal{M}(X \to \cdots)}{\mathcal{M}(\bar{X} \to \overline{\cdots})} \neq 1$

Requires at least two interferring amplitudes:

$$\mathcal{M}(X \to \cdots) = m_1 e^{i\phi_1} e^{i\delta_1} + m_2 e^{i\phi_2} e^{i\delta_2}$$

$$\mathcal{M}(\bar{X} \to \overline{\cdots}) = m_1 e^{-i\phi_1} e^{i\delta_1} + m_2 e^{-i\phi_2} e^{i\delta_2}$$

$$\downarrow$$

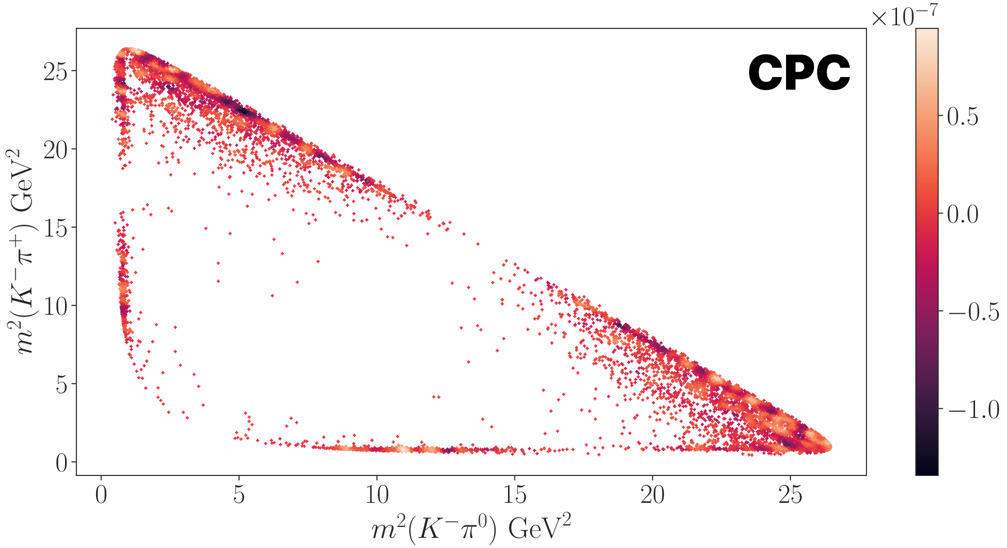
$$\frac{\Gamma(X) - \Gamma(\bar{X})}{\Gamma(X) + \Gamma(\bar{X})} = \frac{-m_1 m_2 \sin(\phi_1 - \phi_2) \sin(\delta_1 - \delta_2)}{|m_1|^2 + |m_2|^2 + 2 m_1 m_2 \cos(\phi_1 - \phi_2) \cos(\delta_1 - \delta_2)}$$

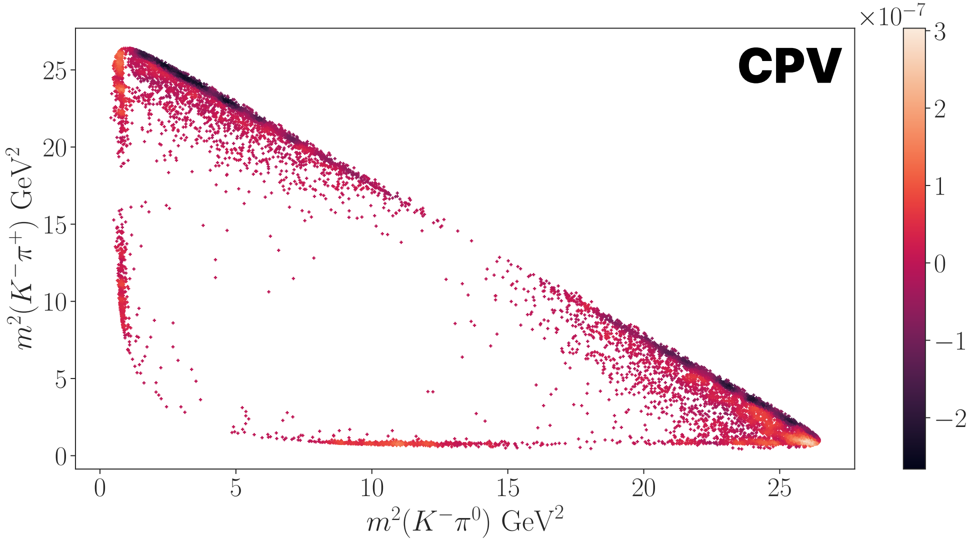
Direct CPV manifests as local **density asymmetries** between conjugate Dalitz plots

Unbinned tests for CPV

Energy Test: $O(n^2)$, tuned on σ

$$T = \sum_{i,j>i}^{N} \frac{\psi_{ij}}{N(N-1)} + \sum_{i,j>i}^{\bar{N}} \frac{\psi_{ij}}{\bar{N}(\bar{N}-1)} - \sum_{i,j}^{N,\bar{N}} \frac{\psi_{ij}}{N\bar{N}}, \qquad \psi_{ij} \equiv \psi(d_{ij};\sigma) = e^{-d_{ij}^2/2\sigma^2}$$





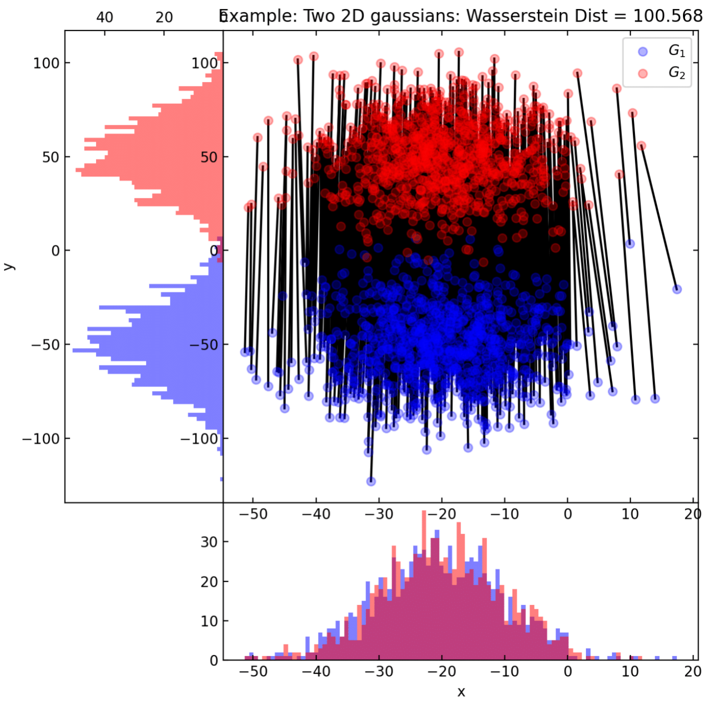
The Optimal Transport Problem

Can you define how similar two distributions are?

One option:

Wasserstein or Earth Mover's distance (EMD)

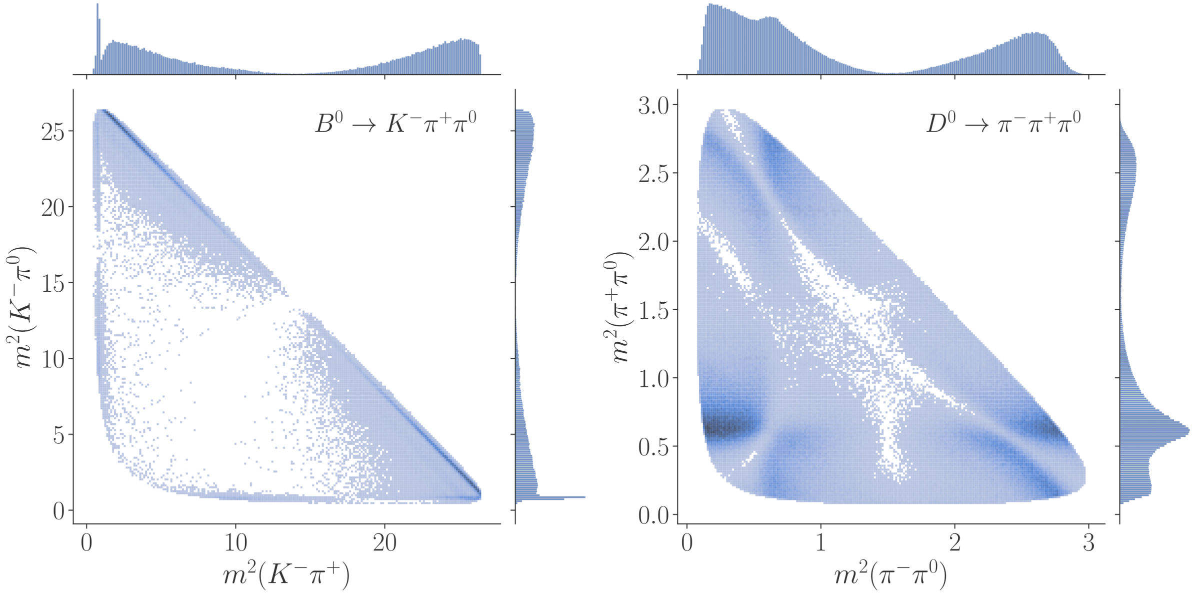
$$W_q(\mathcal{E},\bar{\mathcal{E}}) = \left[\min_{\{f_{ij}\geq 0\}} \sum_{i=1}^{N}\sum_{j=1}^{\bar{N}} f_{ij}\left(\hat{d}_{ij}\right)^q\right]^{1/q}$$



O(n³), tunning on q can be done but not required

Wasserstein-inspired statistic

Implement the EMD on three-body-decay Dalitz plots as a measure of CPV



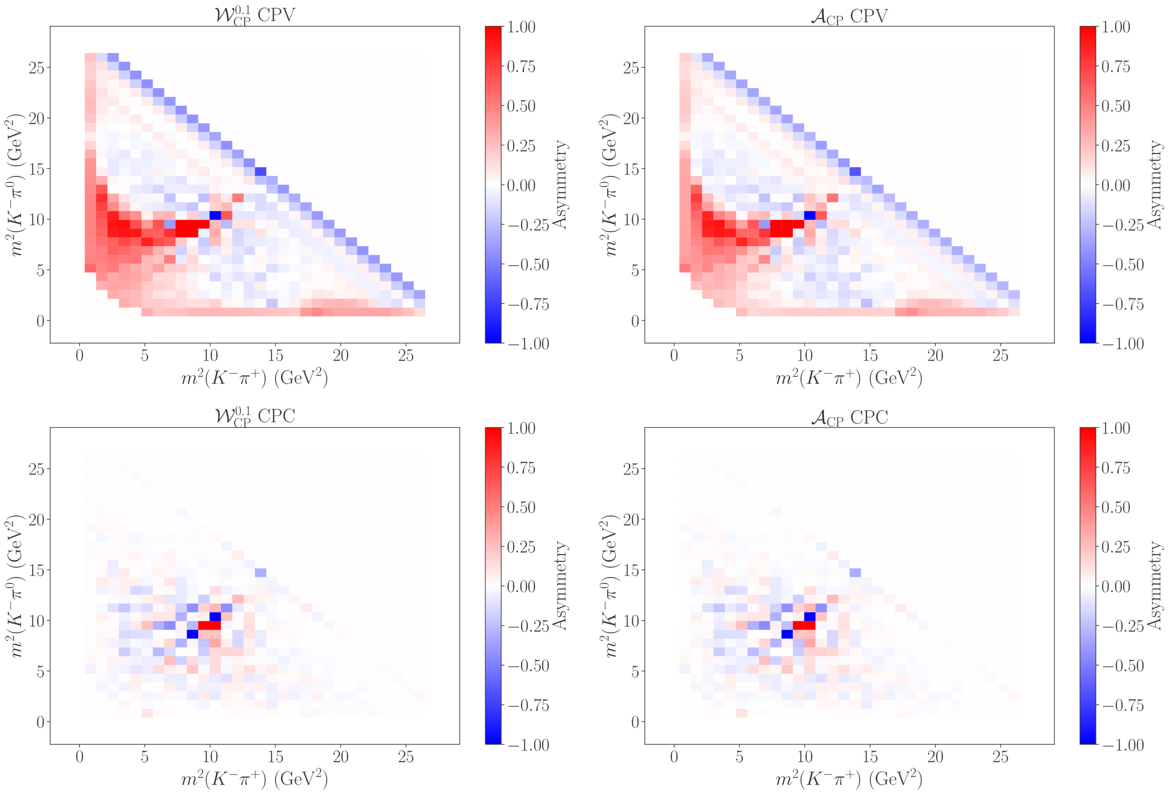
Two analyses: Low statistics (1000 event datasets) and large statistics (10^5 event datasets)

Visualization

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

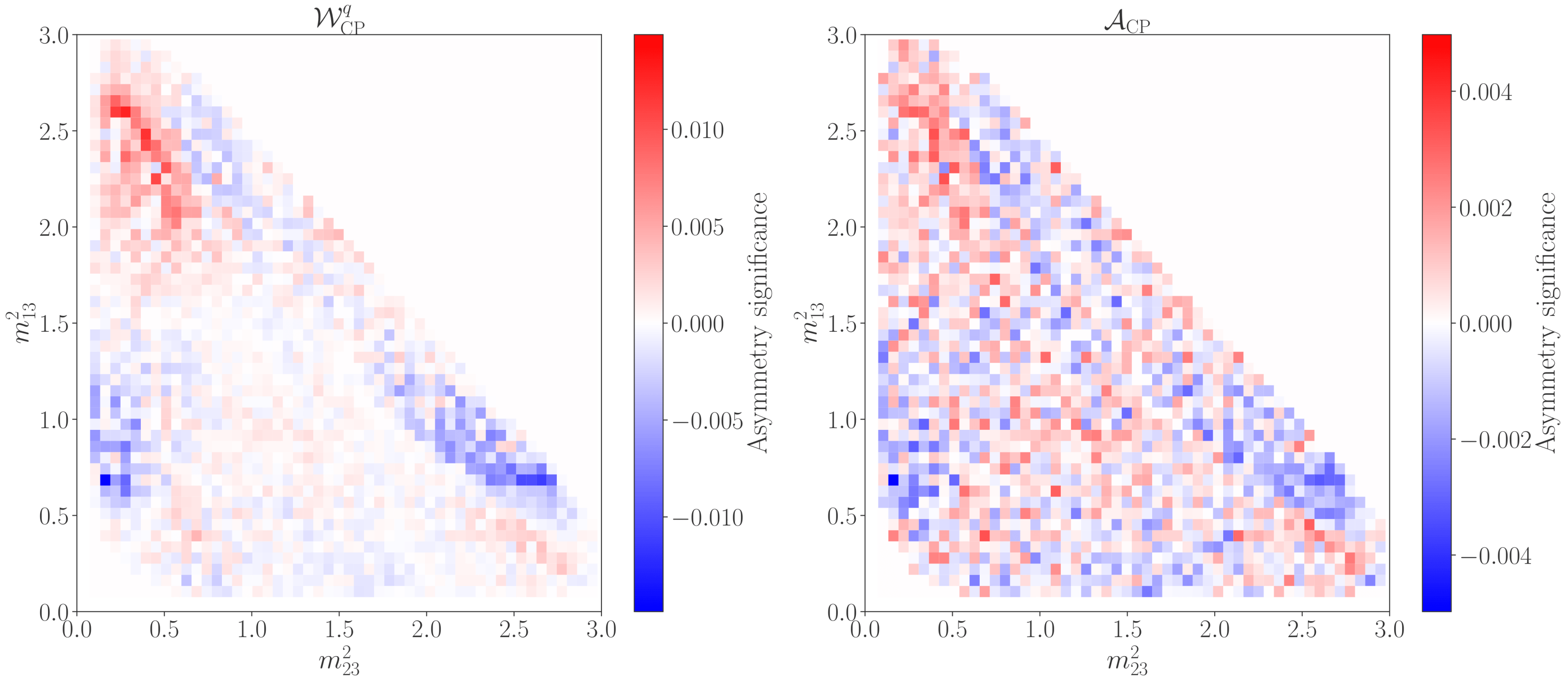
$$W_q = \sum_i \delta W_q(i)$$

$$\mathcal{W}^q_{\text{CP}}(s_{12}, s_{13}) = \frac{\sum_i \delta \bar{W}_i - \sum_i \delta W_i}{\sum_i \delta \bar{W}_i + \sum_i \delta W_i}$$

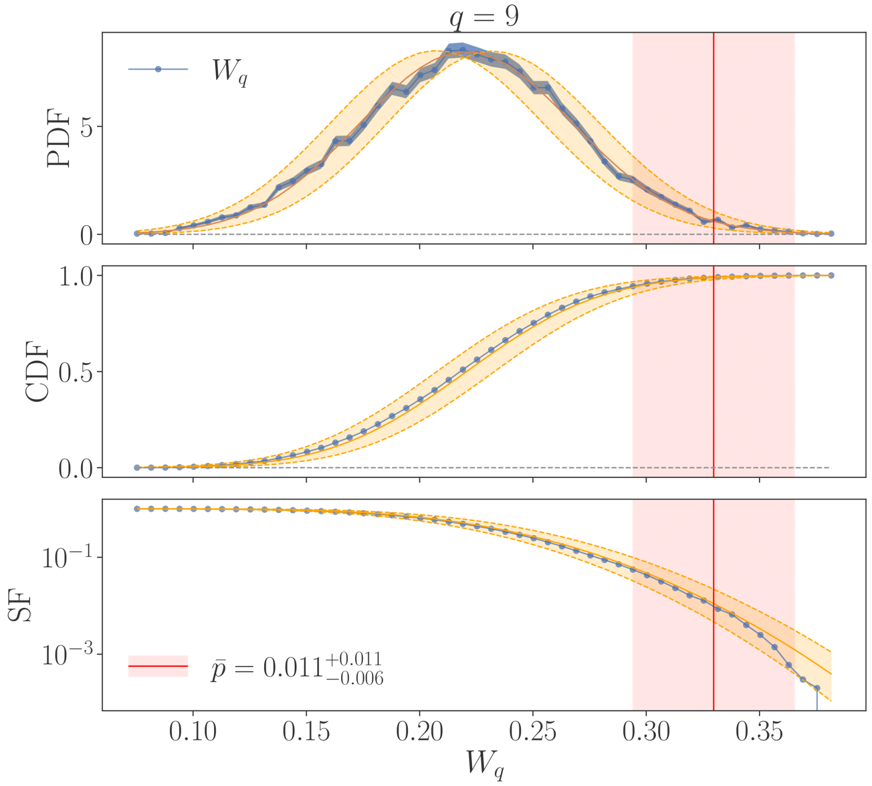


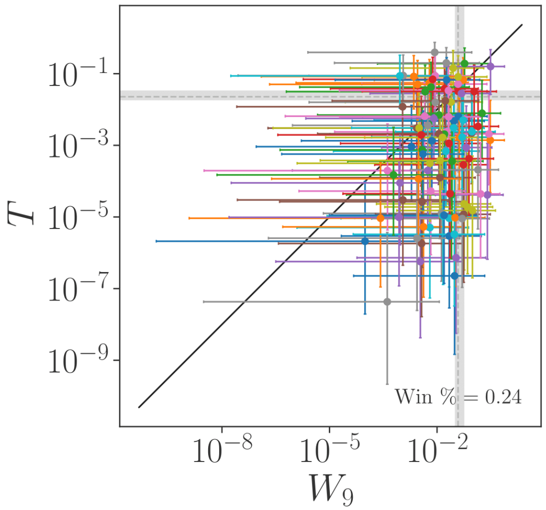
Visualization

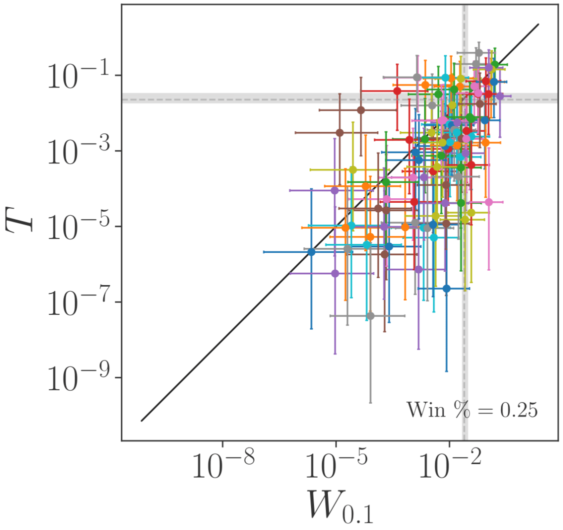
Binned variation for large dataset:



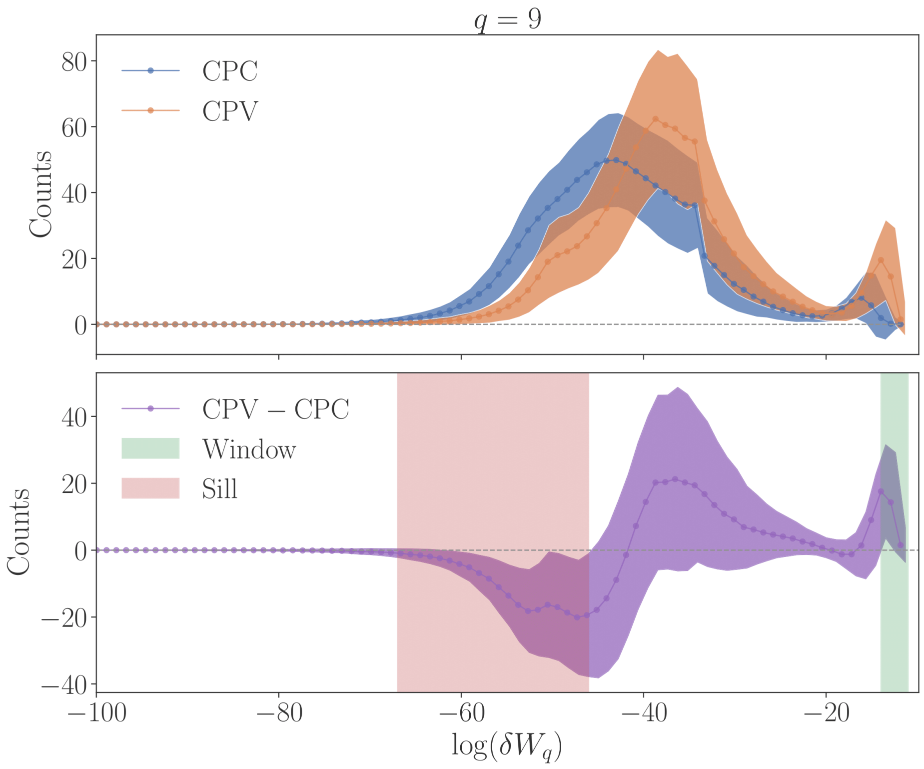
Performance as a statistic



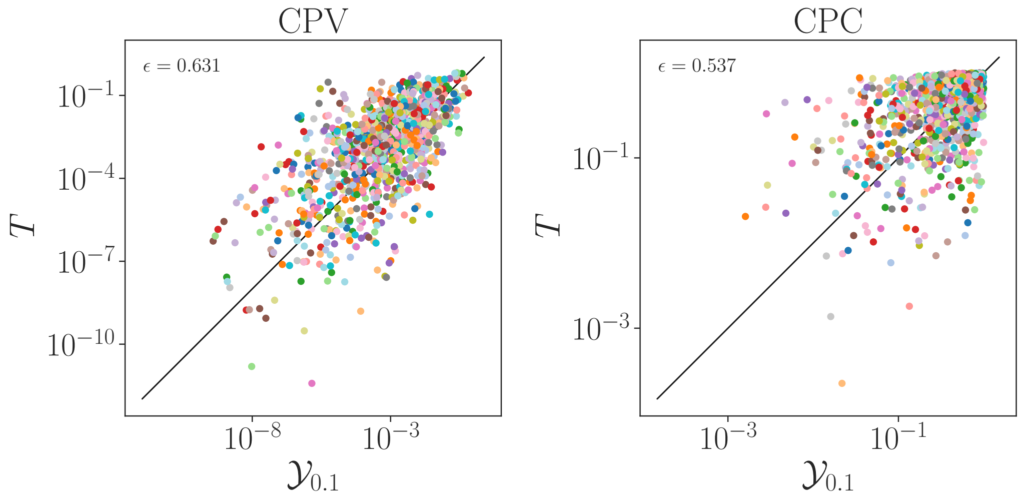




Windowed-Wasserstein statistic



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



Conclusion

The Wasserstein distance can be used as a robust, model independent, unbinned test of CP violation.

- **Future directions:**
  - Time dependent CPV
  - Flavor violation
  - Is there a better test?
  - Model independent way of classifying CPV resonances