Learning the eye of the beholder:

Statistical modeling and estimation for personalized color perception

Xuanzhou Chen[1] Austin Xu[2] Jingyan Wang[1] Ashwin Pananjady[1]

[1]Georgia Institute of Technology
[2]Salesforce AI research

Allerton, 2024

Georgia Tech

Motivation + Goal

Motivation:

Help diagnose color recognition difficulties in people with **color blindness**.

Limitation of color-blind test:

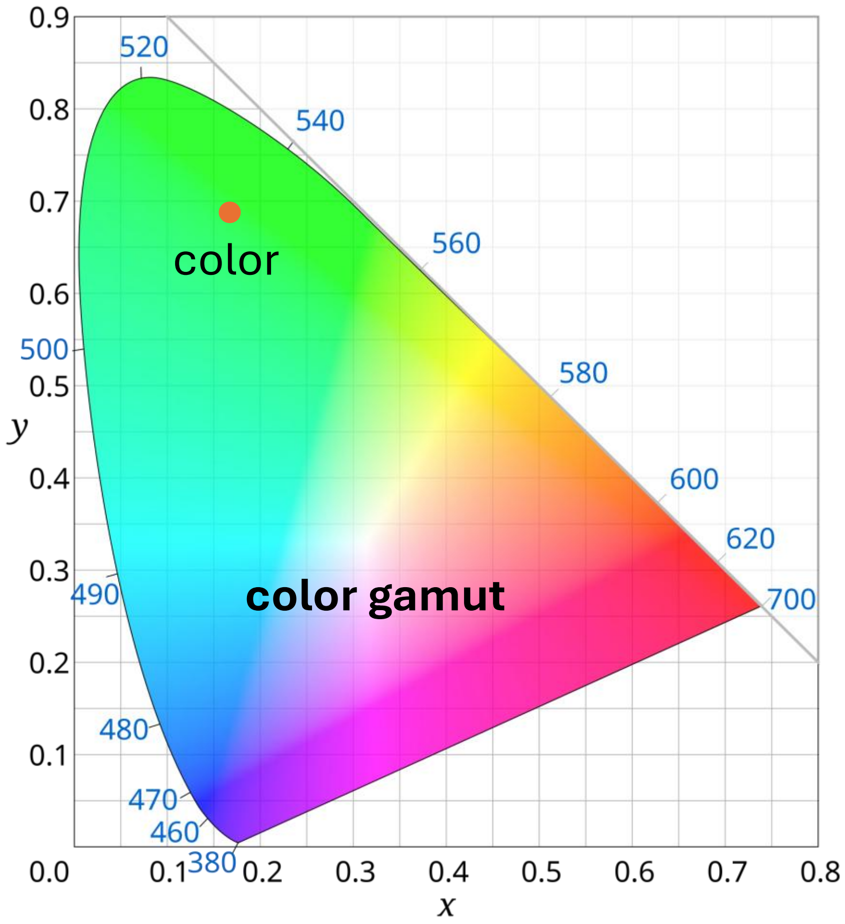
Population level: divide people into two groups--*color blind* or *color normal*.

Fail to capture individual variation in color perception.

Goal:

Understand the individual-level color perception.

Color vision background: color space diagram



CIE xy-chromaticity diagram

color = chromaticity + luminance

hue + colorfulness

Our color space: CIE xy-chromaticity space

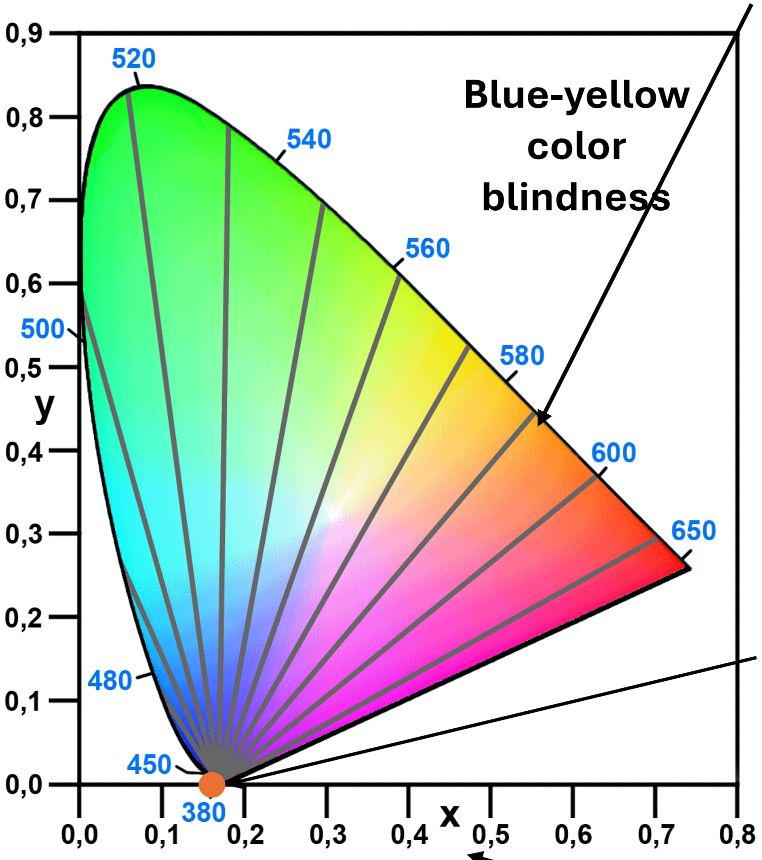
In xy-chromaticity space:

- x-coordinate: the relative proportion of red in the color
- y-coordinate: the relative proportion of green in the color
- Each point in **color gamut** corresponds to a specific color that is visible to human eye.

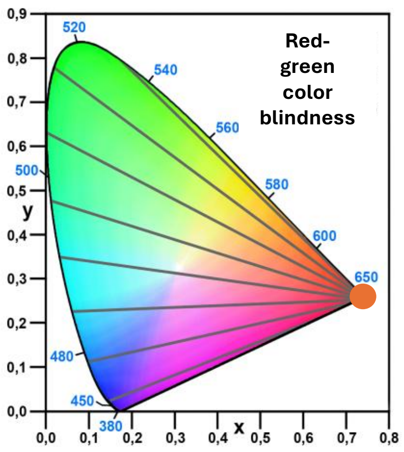
Existing Color Perception Model: color-blind vs. color-normal

Two existing color perception models for color-blind and color-normal

Model for **color-blind**

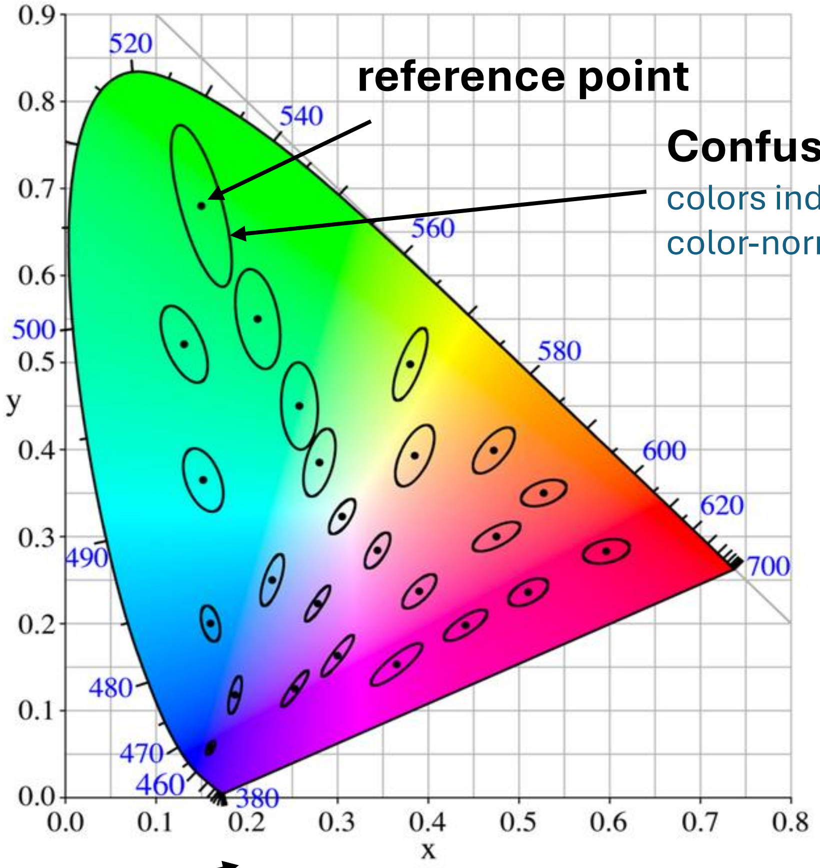


Confusion line:
colors indistinguishable for color-blind



Copunctal point:
Intersection point of all confusion lines.

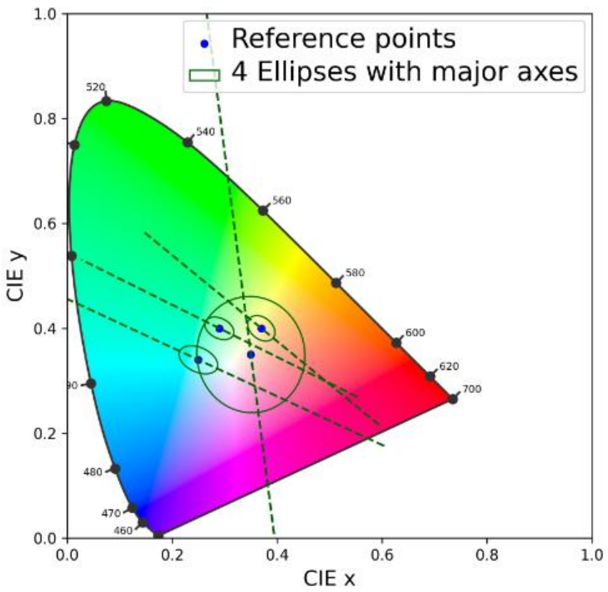
Model for **color-normal**



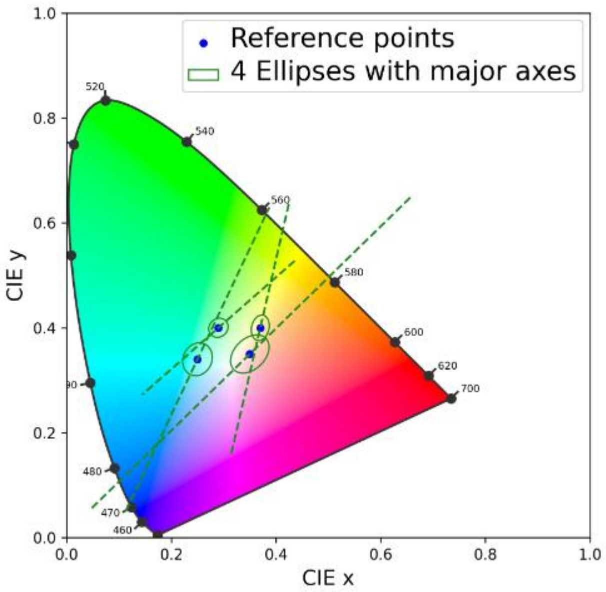
Confusion ellipse:
colors indistinguishable to color-normal

population level

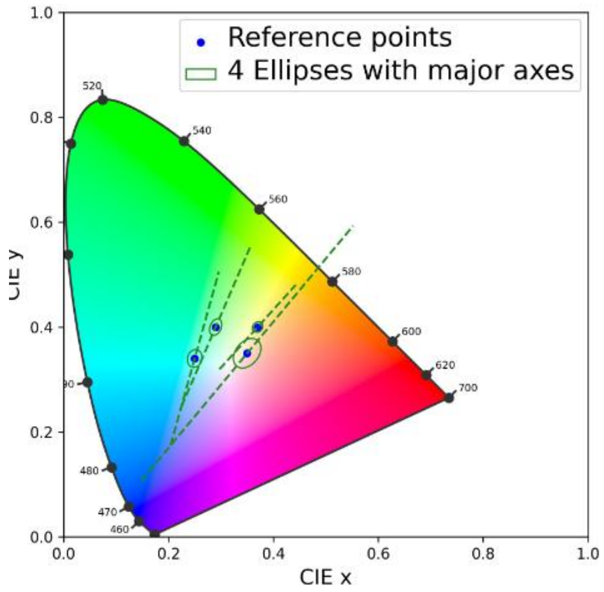
User study: individual color perception variation



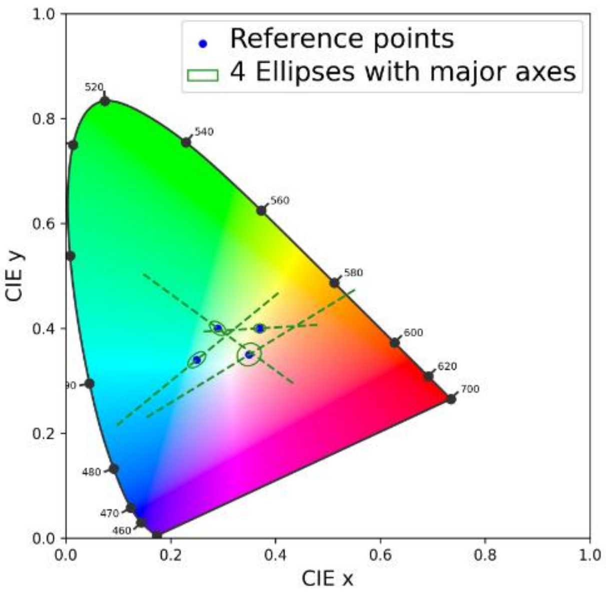
(a) color-blind



(b) color-normal A



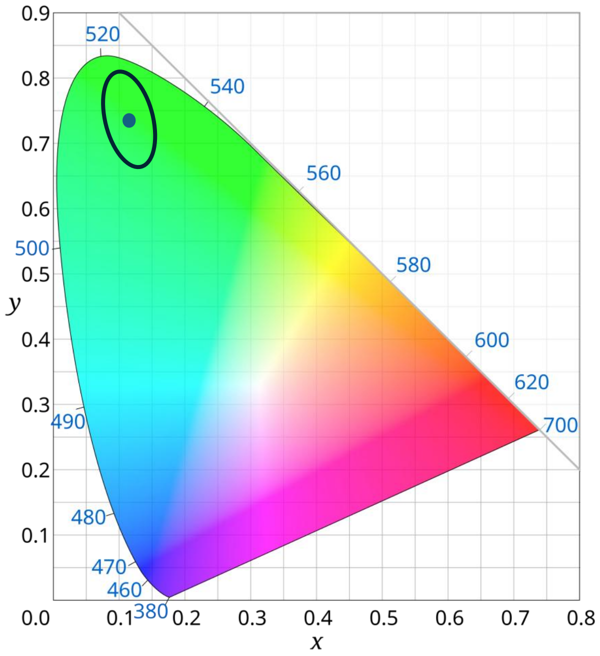
(c) color-normal B



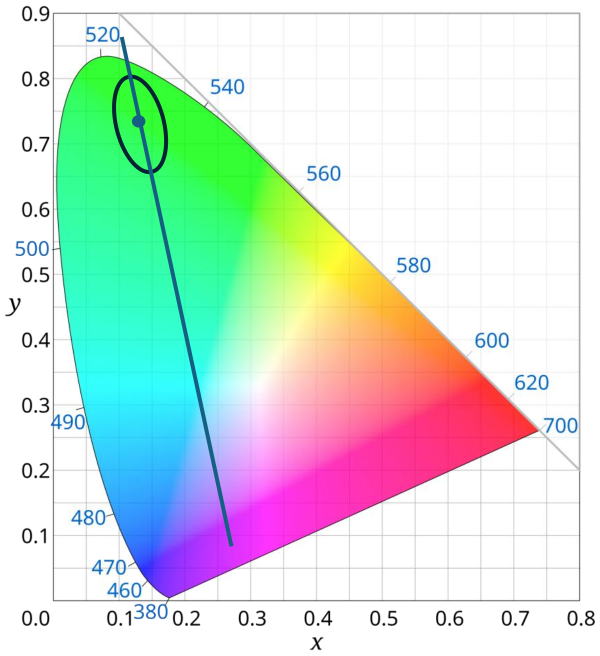
(d) color-normal C

Color Perception Model: our unified model

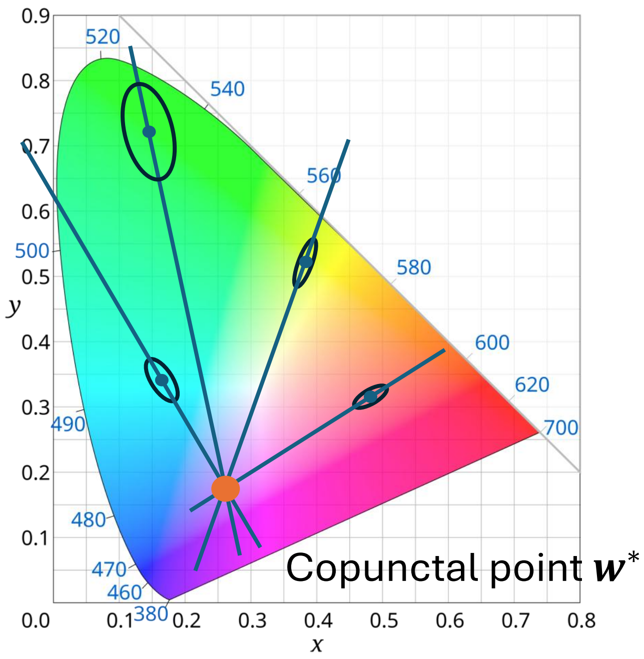
A color perception model based on the “joint structure” of ellipses.



1. **Confusion ellipse** around a color point



2. **Confusion line** is the major axis of the confusion ellipse



3. **Copunctal point** as intersection point of confusion lines

Our color perception model is on individual level.
It is designed for both color-blind and color-normal.

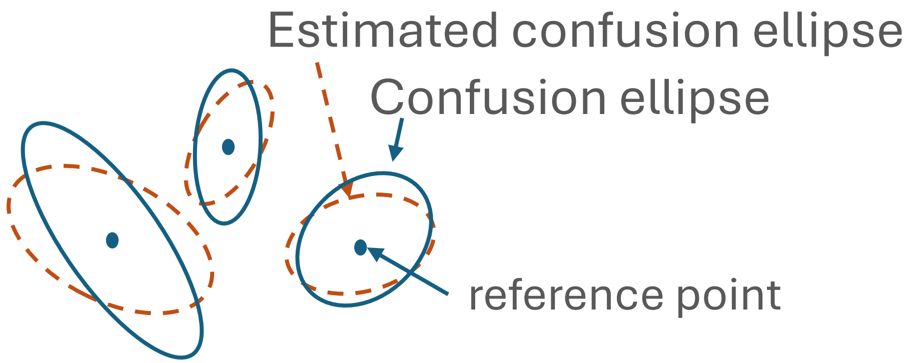
Objective: understand perception by copunctal point estimation

- **Our eventual objective:** Understand individual-level color perception, i.e., learning all ellipses for all users, which is complex.

tackle this.

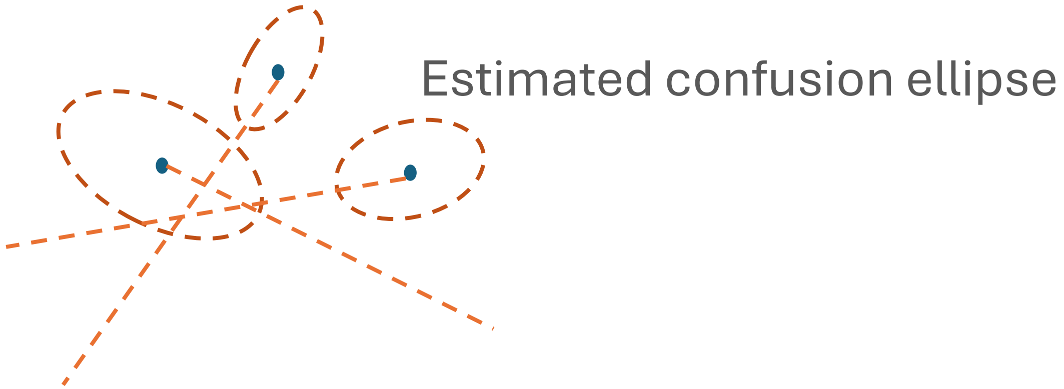
All ellipses share joint structure → Estimate copunctal point for a given user

- **Two challenges when estimate the copunctal point:**



(1) Need to estimate ellipses somehow, how to estimate ellipses?

PART I: Estimate some ellipses



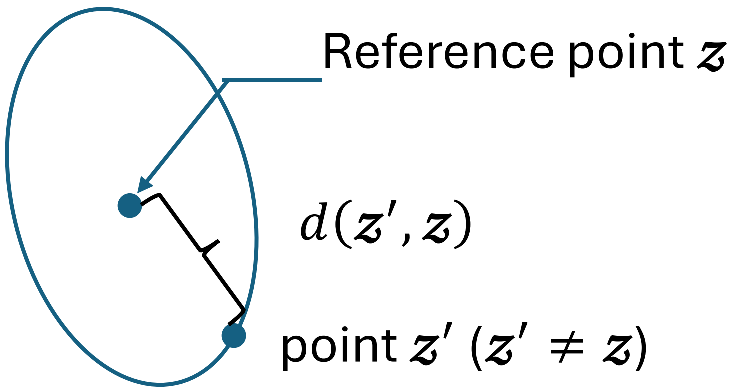
(2) Major axes of estimated ellipses may not intersect. How to estimate the copunctal point?

PART II: Estimate the copunctal point

PART I – Estimate ellipses via metric estimation

How to estimate one ellipse?

$\mathbb{R}^2$ (2D color space)



Q: What is the proper way to define the distance $d(\boldsymbol{z}', \boldsymbol{z})$?

All points on the ellipse centered at $\boldsymbol{z}$ have the same distance from $\boldsymbol{z}$.

Due to scaling invariance, we define the distance to be 1, then for any point $\boldsymbol{z}'$ is on the ellipse, it satisfies

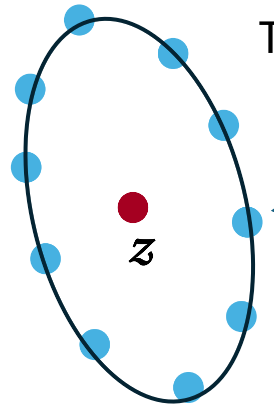
$$\boldsymbol{z}' \in \mathbb{R}^2 \text{ and } \underbrace{(\boldsymbol{z}' - \boldsymbol{z})^T \boldsymbol{\Sigma}_{\boldsymbol{z}}^* (\boldsymbol{z}' - \boldsymbol{z})}_{d(\boldsymbol{z}', \boldsymbol{z})} = 1$$

$\boldsymbol{\Sigma}_{\boldsymbol{z}}^*$: 2 x 2 positive semidefinite matrix.

Note: this problem is called metric learning, $\boldsymbol{\Sigma}_{\boldsymbol{z}}^*$ is called metric and $\boldsymbol{\Sigma}_{\boldsymbol{z}}^*$ determines the ellipse at $\boldsymbol{z}$.

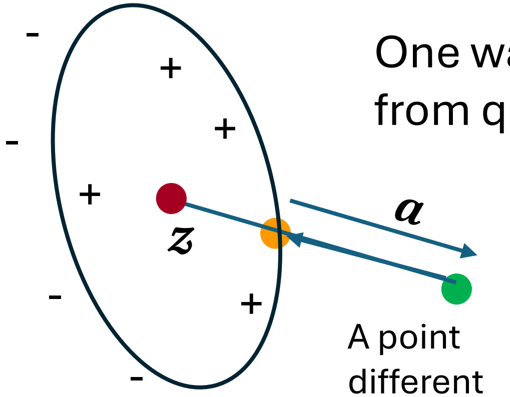
Estimate one ellipse at $\boldsymbol{z}$ ⟷ Estimate metric $\boldsymbol{\Sigma}_{\boldsymbol{z}}^*$

PART I – Collect PAQ data to estimate metric



To determine the ellipse,
we need to collect transition point

(i.e., the farthest point indistinguishable from reference color z)



One way: collect transition point from query from paired comparison

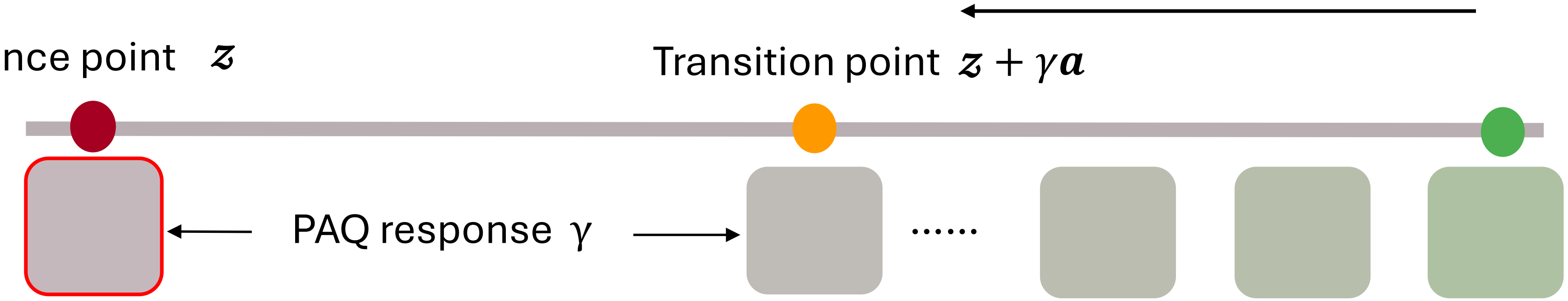
a is the direction vector of the path

PAQ (perceptual adjustment queries)

Slider direction

Reference point z

Transition point $z + \gamma a$

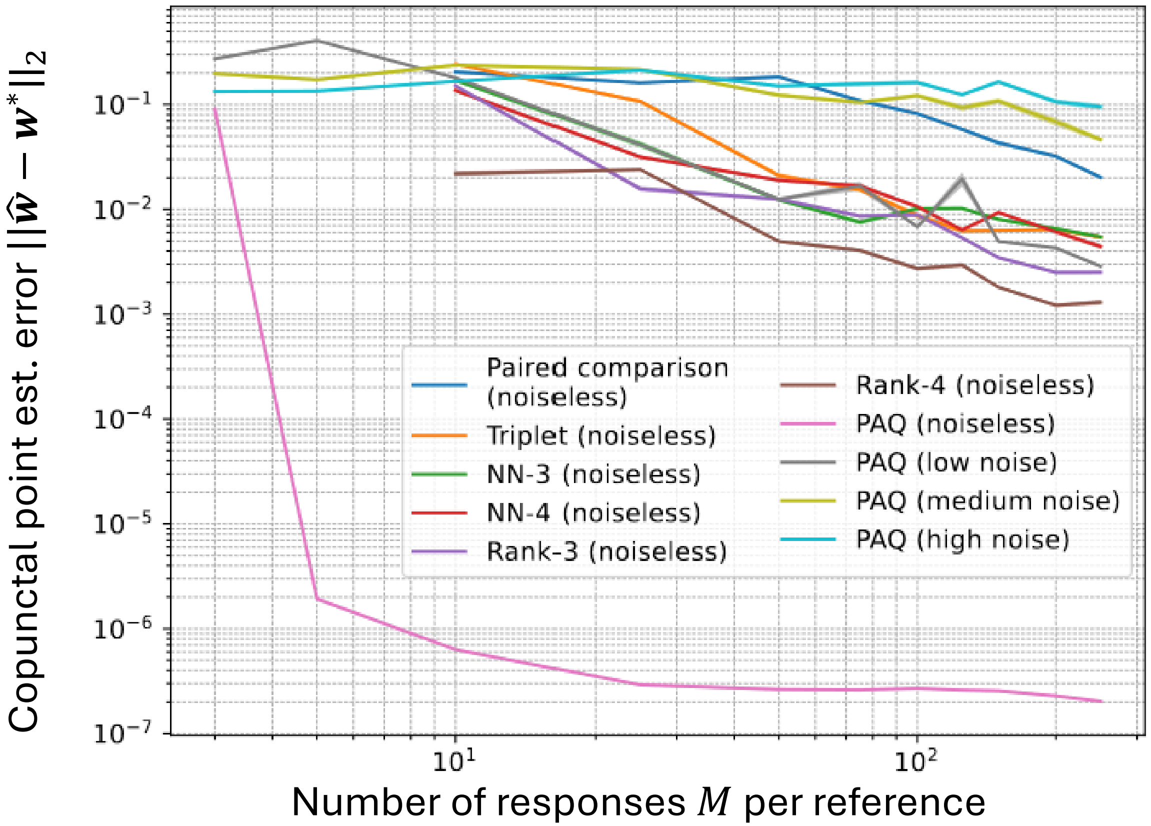


Reference color

PART I – PAQ wins over ordinal queries

Choice of query types

- **Paired comparison**: "Are colors $\boldsymbol{x}$ and $\boldsymbol{z}$ similar?"
- **Triplets**: "which of $\boldsymbol{x}_1$ or $\boldsymbol{x}_1$ is more similar to reference color $\boldsymbol{z}$?"
- **Nearest-neighbor queries**:"which of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ is most similar to reference color $\boldsymbol{z}$?"
- **Ranking queries**:"Rank order $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ in terms of similarity to reference color $\boldsymbol{z}$?"



PAQ is efficient:
Achieve much lower error by requiring a few queries.

PART I – Least squares to estimate metric

Rewrite the ellipse boundary as

$$1 = (\boldsymbol{z}' - \boldsymbol{z})^T \boldsymbol{\Sigma}_{\boldsymbol{z}} (\boldsymbol{z}' - \boldsymbol{z}) \quad = \gamma^2 \, \boldsymbol{a}^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{z}} \boldsymbol{a}$$

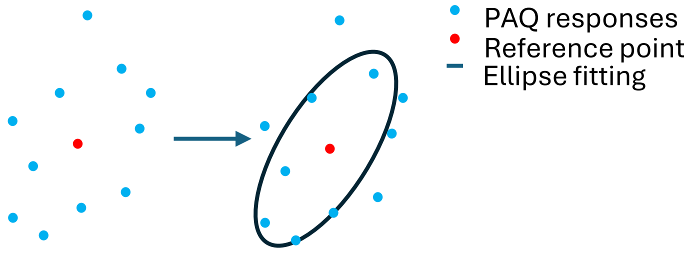
PAQ response (γ^2) — Metric ($\boldsymbol{\Sigma}_{\boldsymbol{z}}$) — Direction vector ($\boldsymbol{a}$)

Least squares estimator to learn one ellipse

Given M PAQ responses $\{\gamma_j\}_{j=1}^{M}$ at each reference point $\boldsymbol{z}$,

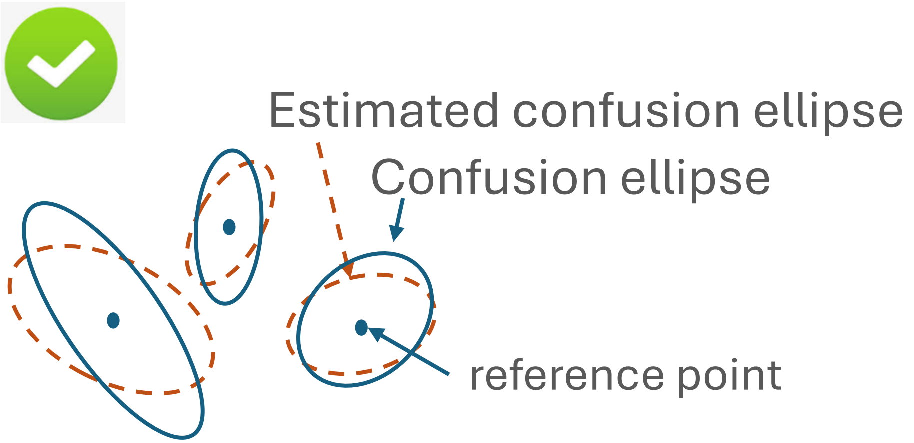
$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} \in \underset{\boldsymbol{\Sigma} \succcurlyeq 0}{\operatorname{argmin}} \frac{1}{M} \sum_{j=1}^{M} \left(\langle \boldsymbol{a}_j \boldsymbol{a}_j{}^T, \boldsymbol{\Sigma} \rangle - \frac{1}{\gamma_j^2} \right)^2 .$$

LS estimation with multiplicative noise

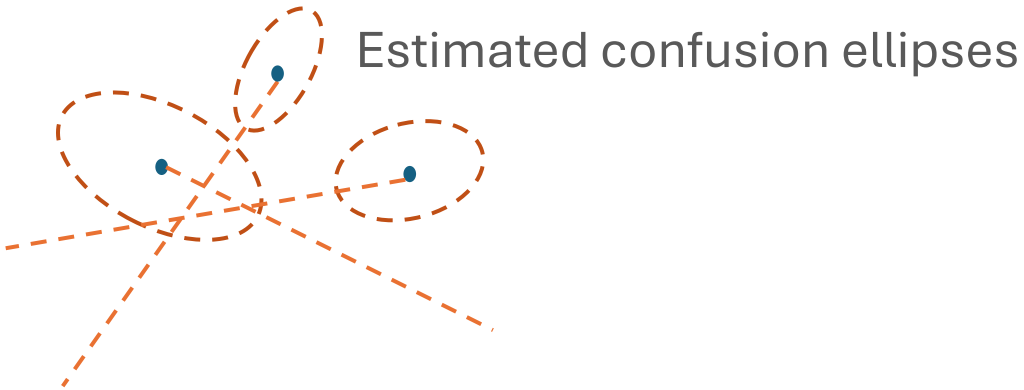


PART I. estimate the ellipse from PAQ data

PART II – Least squares to estimate metric



(1) Need to estimate ellipses somehow, how to estimate ellipses?



(2) Major axes of estimated ellipses may not intersect. How to estimate the copunctal point?

PART I: Estimate some ellipses

PART II: Estimate the copunctal point

Methodology: PART II – Estimate copunctal point

How to estimate copunctal point from noisy ellipses?

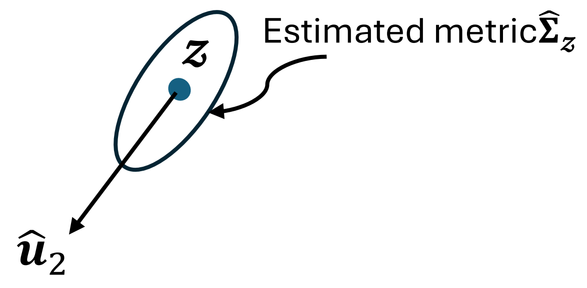
Given estimated ellipse (i.e., $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}}$), $||\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}} - \boldsymbol{\Sigma}^*_{\boldsymbol{z}}||_{\mathrm{op}} \leq \tau_{\boldsymbol{z}}$

- Compute major axis as second eigenvector $\widehat{\boldsymbol{u}}_2$ from SVD of $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}}$.

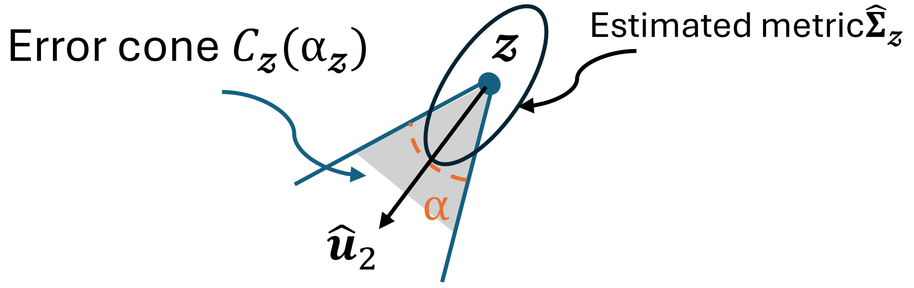
Noisy ellipse → major axes may not intersect

- Construct an error cone $C_{\boldsymbol{z}}(\alpha_{\boldsymbol{z}})$

1. *Vertex at* $\boldsymbol{z}$;
2. $C(\alpha)$ symmetric about $\hat{\mathbf{u}}_2$;
3. *Cone angle* $\alpha_{\boldsymbol{z}} \propto \tau_{\boldsymbol{z}} / |\hat{\lambda}_1 - \hat{\lambda}_2|$, $\tau_{\boldsymbol{z}}$ *is operator norm error bound on metric and* $|\hat{\lambda}_1 - \hat{\lambda}_2|$ *is eigenvalue gap.*



1. compute the major axis as $\widehat{\boldsymbol{u}}_2$



2. construct the error cone $C_{\boldsymbol{z}}(\alpha_{\boldsymbol{z}})$

$\hat{\lambda}_1$: maximum eigenvalue; $\hat{\lambda}_2$: minimum eigenvalue

Methodology: PART II – Estimate copunctal point

Once construct error cones at some reference points, how to find the copunctal point?

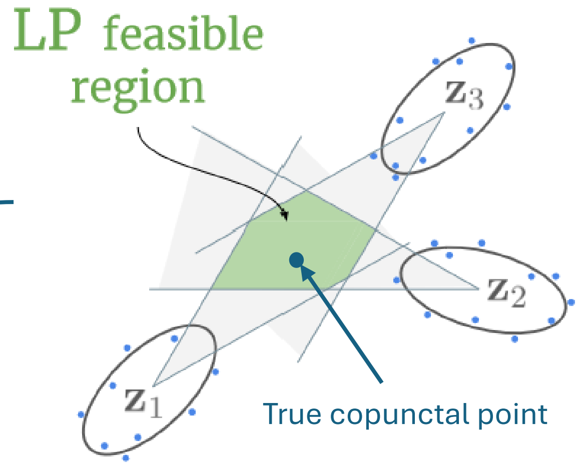
Fact: The true copunctal point is in the intersection of these error cones!

- True major axis is in error cone.
- True copunctal point is the intersection of these true major axes.
- True copunctal point is in the intersection of these error cones.

2. Find copunctal point via linear programming

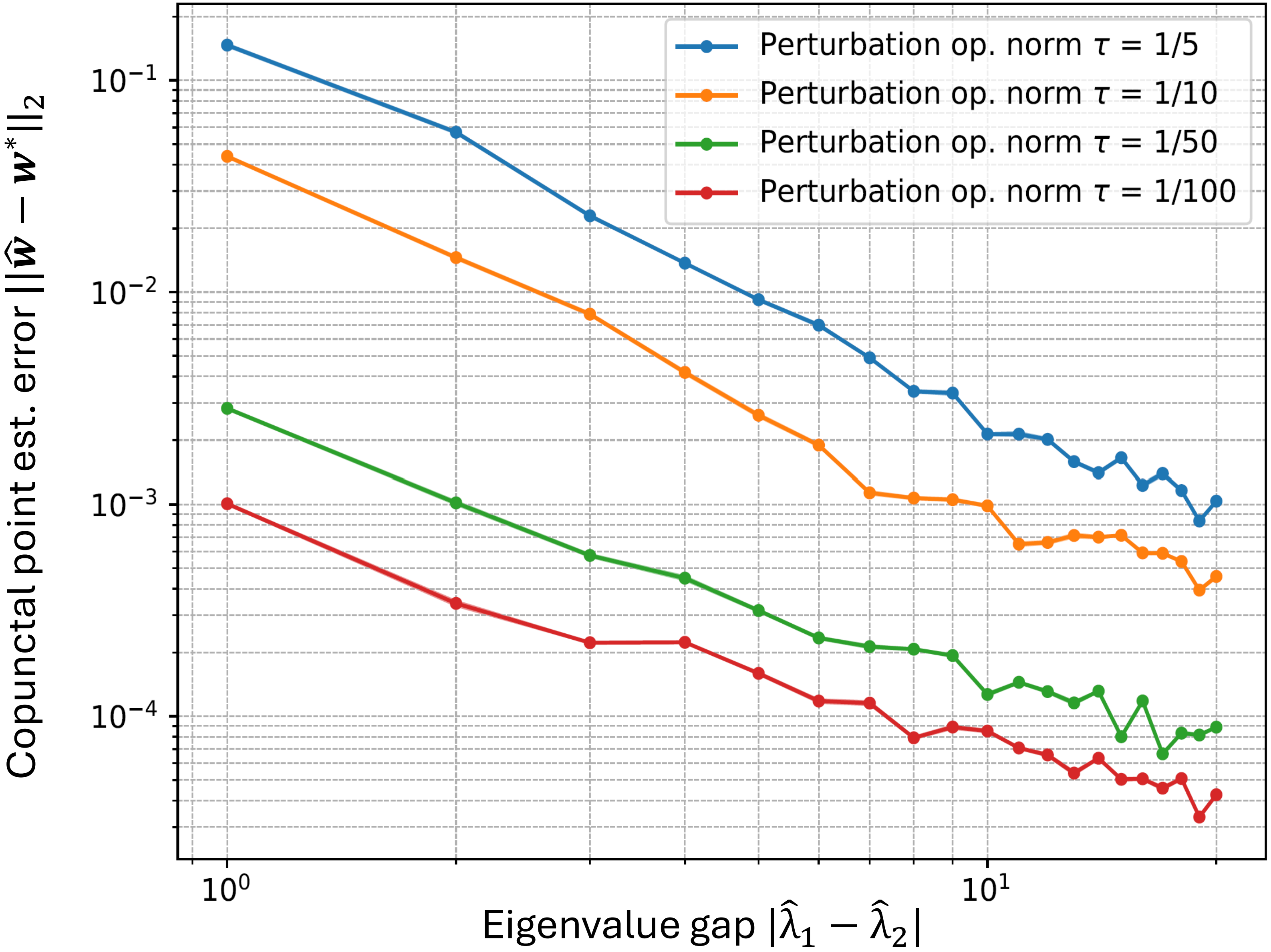
Consider each error cone as two linear constraints,

Estimate of copunctal point ← **LP** (all error cones).
($\hat{\boldsymbol{w}}$)



2. find the copunctal point by LP

Simulation: est. error vs. eigenvalue gap/op. norm bound



Increase operator norm bound on the metric estimation error τ

Theorem 1: eigenvalue gap & op. norm bound

➢ Theorem 1. (Informal)

Given N reference points $\{\boldsymbol{z_i}\}_{i=1}^{N}$, if $||\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z_i}} - \boldsymbol{\Sigma}^*_{\boldsymbol{z_i}}||_{\text{op}} \leq \tau_i$ for any $i \in [N]$, then

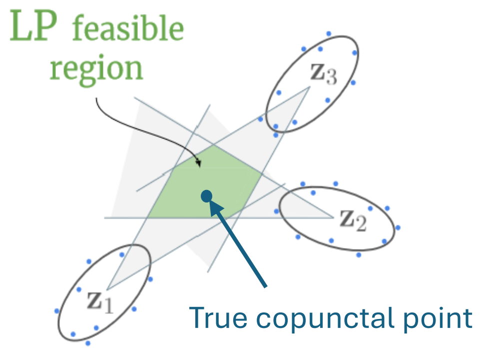
$$||\widehat{\boldsymbol{w}} - \boldsymbol{w}^*||_2 \lesssim \min_{\substack{i,j\in[N],\\ i\neq j}} ||\boldsymbol{z_i} - \boldsymbol{z_j}||_2 \cdot \tan\left(\frac{2\pi\tau_i}{|\widehat{\lambda}_1^{(i)} - \widehat{\lambda}_2^{(i)}|} \vee \frac{2\pi\tau_j}{\left|\widehat{\lambda}_1^{(j)} - \widehat{\lambda}_2^{(j)}\right|}\right),$$

where $\vee$ is Max operator.

Take-away:

1. $||\widehat{\boldsymbol{w}} - \boldsymbol{w}^*||_2 \propto \tau$, $||\widehat{\boldsymbol{w}} - \boldsymbol{w}^*||_2 \propto \frac{1}{|\widehat{\lambda}_1^{(i)} - \widehat{\lambda}_2^{(i)}|}$

If we increase τ or decrease eigenvalue gap, we actually make cone angle larger and therefore, the estimation error of copunctal point is larger.



2. find the copunctal point by LP

Theorem 1: eigenvalue gap & operator norm bound

Denote $\boldsymbol{\Sigma}_{\boldsymbol{z}}^*$ as true metric and $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}}$ as estimated metric at reference point $\boldsymbol{z}$.

➢ Theorem 1. (Informal)

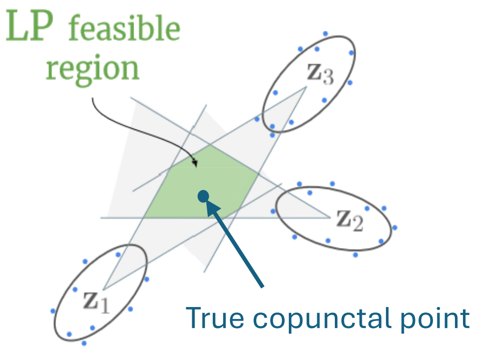
Given N reference points $\{\boldsymbol{z}_i\}_{i=1}^N$, if $||\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_i} - \boldsymbol{\Sigma}_{\boldsymbol{z}_i}^*||_{\mathrm{op}} \leq \tau_i$ for any $i \in [N]$, then

$$||\widehat{\boldsymbol{w}} - \boldsymbol{w}^*||_2 \lesssim \min_{\substack{i,j\in[N],\\ i\neq j}} ||\boldsymbol{z}_i - \boldsymbol{z}_j||_2 \cdot \tan\left(\frac{2\pi\tau_i}{|\hat{\lambda}_1^{(\mathrm{i})} - \hat{\lambda}_2^{(i)}|} \vee \frac{2\pi\tau_j}{\left|\hat{\lambda}_1^{(j)} - \hat{\lambda}_2^{(j)}\right|}\right),$$

where $\vee$ is Max operator.

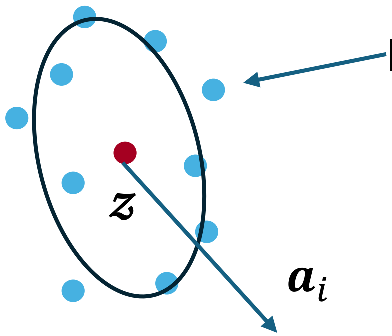
Two steps to get this pairwise upper bound:

- Bound the distance $||\widehat{\boldsymbol{w}} - \boldsymbol{w}^*||_2$ by the diameter of feasible set
- Bound diameter of feasible set by the diameter of arbitrary pairwise cone intersections.



2. find the copunctal point by LP

Theorem 2: # PAQ responses



Assume:

1. the noise η zero-mean and bounded;
2. Each direction vector i.i.d. drawn from the Gaussian distribution, i.e.,$\boldsymbol{a}_i \sim \mathcal{N}(0, \boldsymbol{I}_d)$

Denote $\sigma := Var[1/\eta]$

➢ Theorem 2. (Informal)

Suppose for each reference point $\boldsymbol{z}_i$ for $i \in [N]$, each direction vector $\boldsymbol{a}_i$ is i.i.d. drawn from the Gaussian distribution, i.e.,$\boldsymbol{a}_i \sim \mathcal{N}(0, \boldsymbol{I}_d)$. For any $\delta \in (0,1)$, if the number of PAQ measurements M_i at $\boldsymbol{z}_i$ satisfies $M_i \gtrsim \log^3\left(\frac{M_i}{\delta}\right)$,

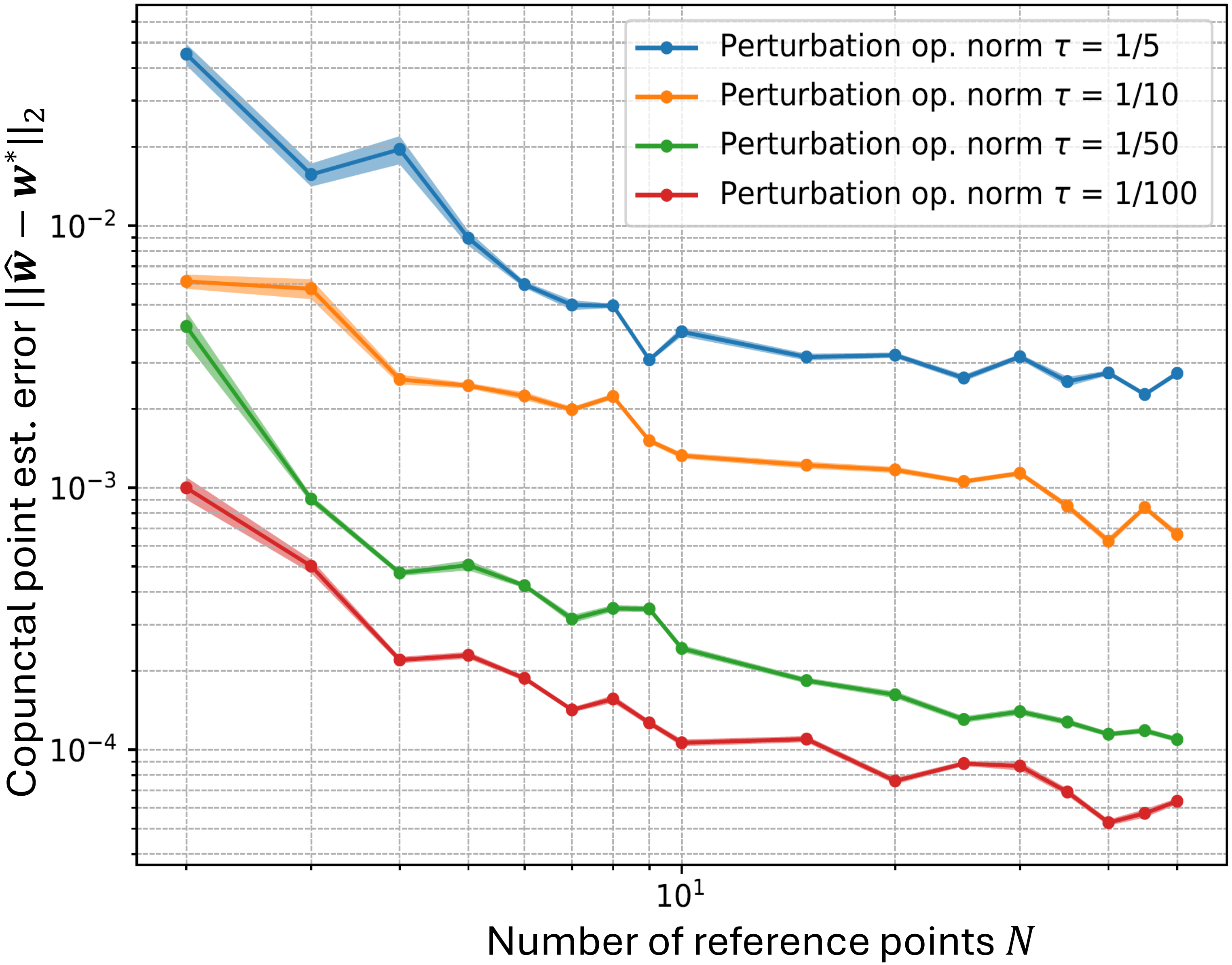
Then with probability greater than $1 - N\delta$,

$$||\hat{\boldsymbol{w}} - \boldsymbol{w}^*||_2 \lesssim \min_{i,j \in [N]} \sigma\, A = \pi r^2 \sqrt{1 + \log^4\left(\frac{1}{\delta}\right)} ||\boldsymbol{z_i} - \boldsymbol{z_j}||_2 \cdot \tan\left(\frac{2\pi}{|\hat{\lambda}_1^{(i)} - \hat{\lambda}_2^{(i)}|\sqrt{M_i}} \vee \frac{2\pi}{|\hat{\lambda}_1^{(j)} - \hat{\lambda}_2^{(j)}|\sqrt{M_j}}\right).$$

Take-away: $||\hat{\boldsymbol{w}} - \boldsymbol{w}^*||_2 \propto \frac{1}{\sqrt{M}}$

If we take more PAQ measurements, then we have a better estimate of metric and therefore we have a better estimate of copunctal point.

Simulation: est. error vs. #reference points



Phenomenon not captured by the theorem:

N ↑, estimation error ↓

Conclusion & Open questions

- **Modeling:** Propose a unified color perception model based on ellipse shared structure.
- **Methodology:** Design algorithm to estimate the copunctal point by PAQ data.
- **Theory:** Provide statistical guarantees on estimation accuracy.
- **Experiment:** Perform simulation and user study.

Open questions

- When we perform **active learning** in collect PAQ responses where direction vectors are not i.i.d. drawn from standard Gaussian, how do we choose the direction vectors and how to deal with the dependencies between the direction vectors?
- How to explain the impact of the number of reference point on estimation error of copunctal point? And how to select the reference point?
- Theorem 1 holds for our model in low dimension. How do we generalize our theoretical results from 2D color space to higher dimension?