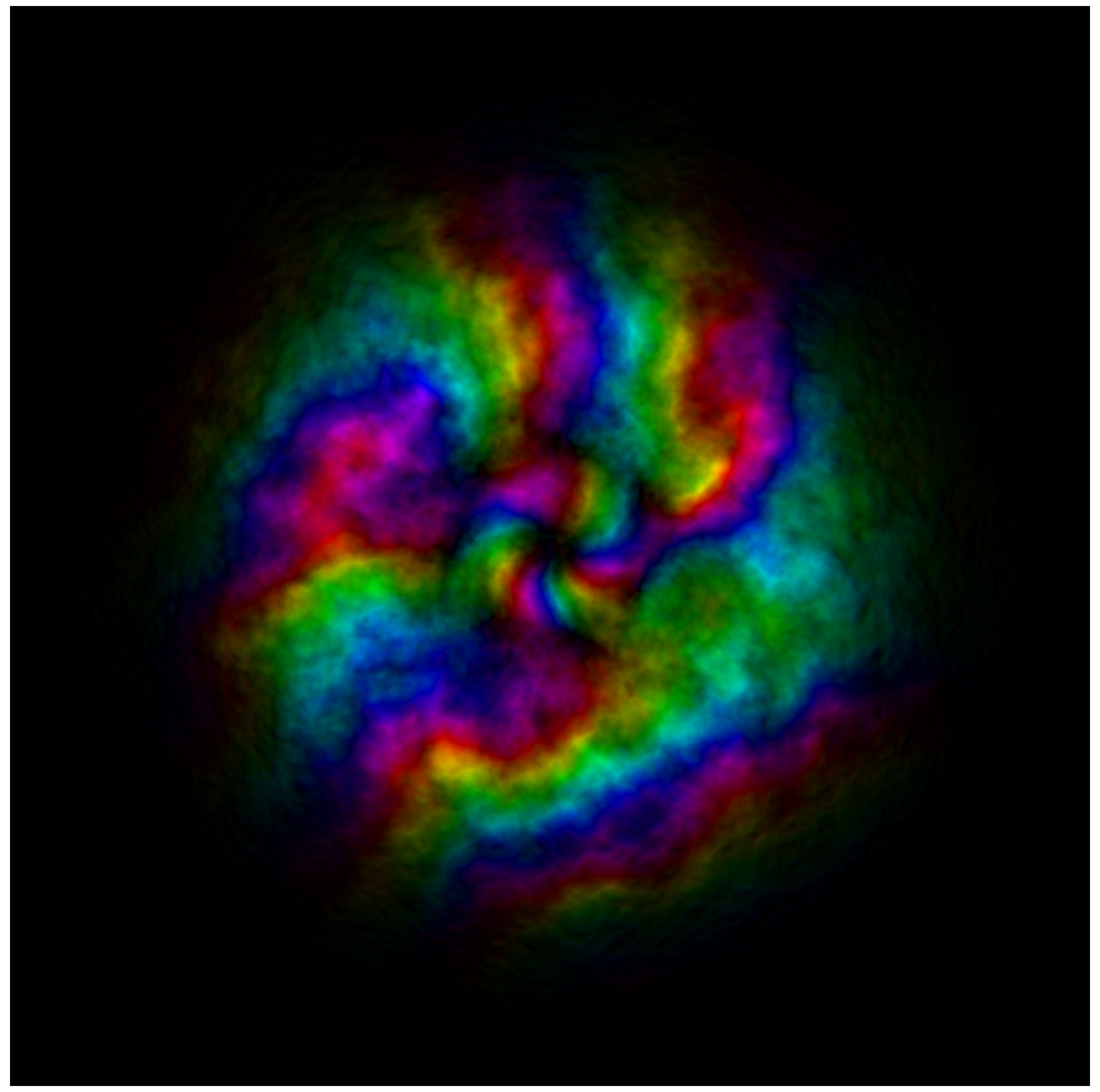
Universidad de los Andes

Numerical classification of structured light beams using optimal transport theory

Jaime Cisternas & Jaime Anguita
Universidad de los Andes - Chile



Numerical Methods for Optimal Transport Problems, Mean Field Games, and Multi-Agent Dynamics (U.T. Federico Santa María, january 2024)

Outline

- Orbital Angular Momenta (OAM) spectra obtained with spatial mode sensors [1d histograms]
- Optimal Transport and Wasserstein distance
- Selection of optimal subset of superpositions
- Classification results
- Partial conclusions
- Extension to Vector Vortex Beams (OAM + polarization) [2d histograms]
- Final conclusions
- **Acknowledgements:** This work is supported by Millennium Institute for Research in Optics (ICN17-012), and ANID-Chile (FONDECYT-1210297).

Twisted light transmission over 143 km

Mario Krenn[a,b,1], Johannes Handsteiner[a,b], Matthias Fink[b], Robert Fickler[a,b,c,d], Rupert Ursin[b], Mehul Malik[a,b], and Anton Zeilinger[a,b,1]

[a]Faculty of Physics, Vienna Center for Quantum Science and Technology, University of Vienna, A-1090 Vienna, Austria; [b]Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, A-1090 Vienna, Austria; [c]Department of Physics, University of Ottawa, Ottawa, ON, Canada K1N 6N5; and [d]Max Planck Centre for Extreme and Quantum Photonics, University of Ottawa, Ottawa, ON, Canada K1N 6N5

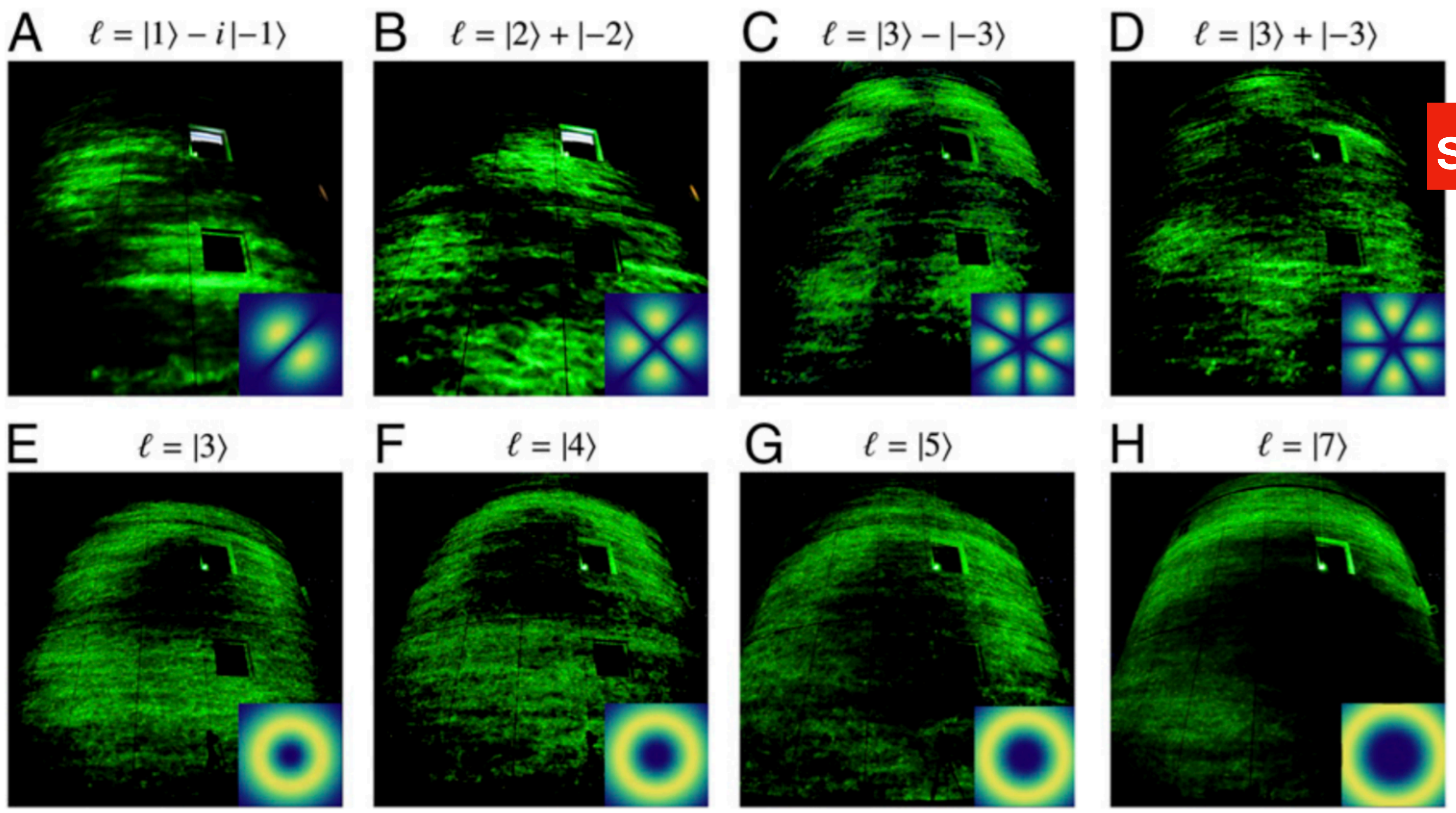
Contributed by Anton Zeilinger, October 13, 2016 (sent for review June 9, 2016; reviewed by Andrew Forbes, Jon

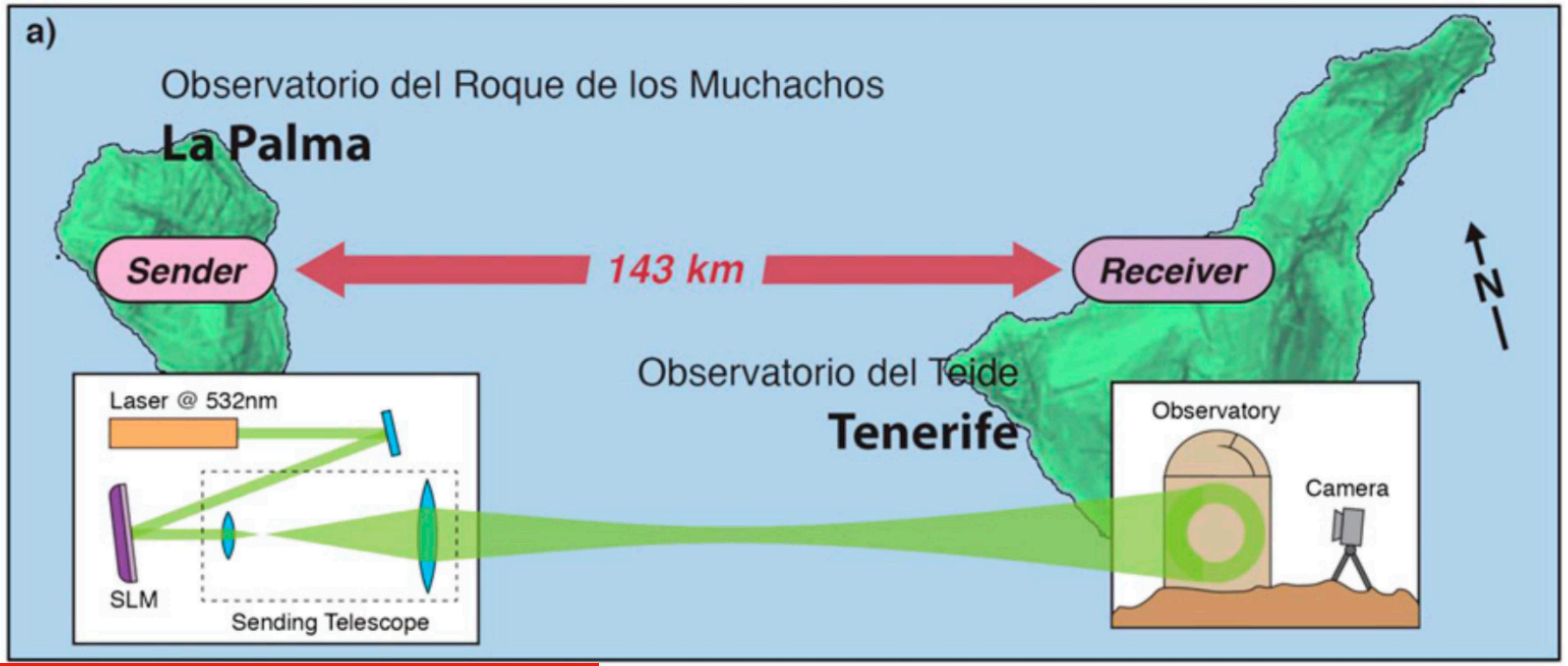
(Anton Zeilinger is an Austrian quantum physicist and Nobel laureate in physics of 2022)

(A few years ago they achieved similar results but sending light from a satellite)

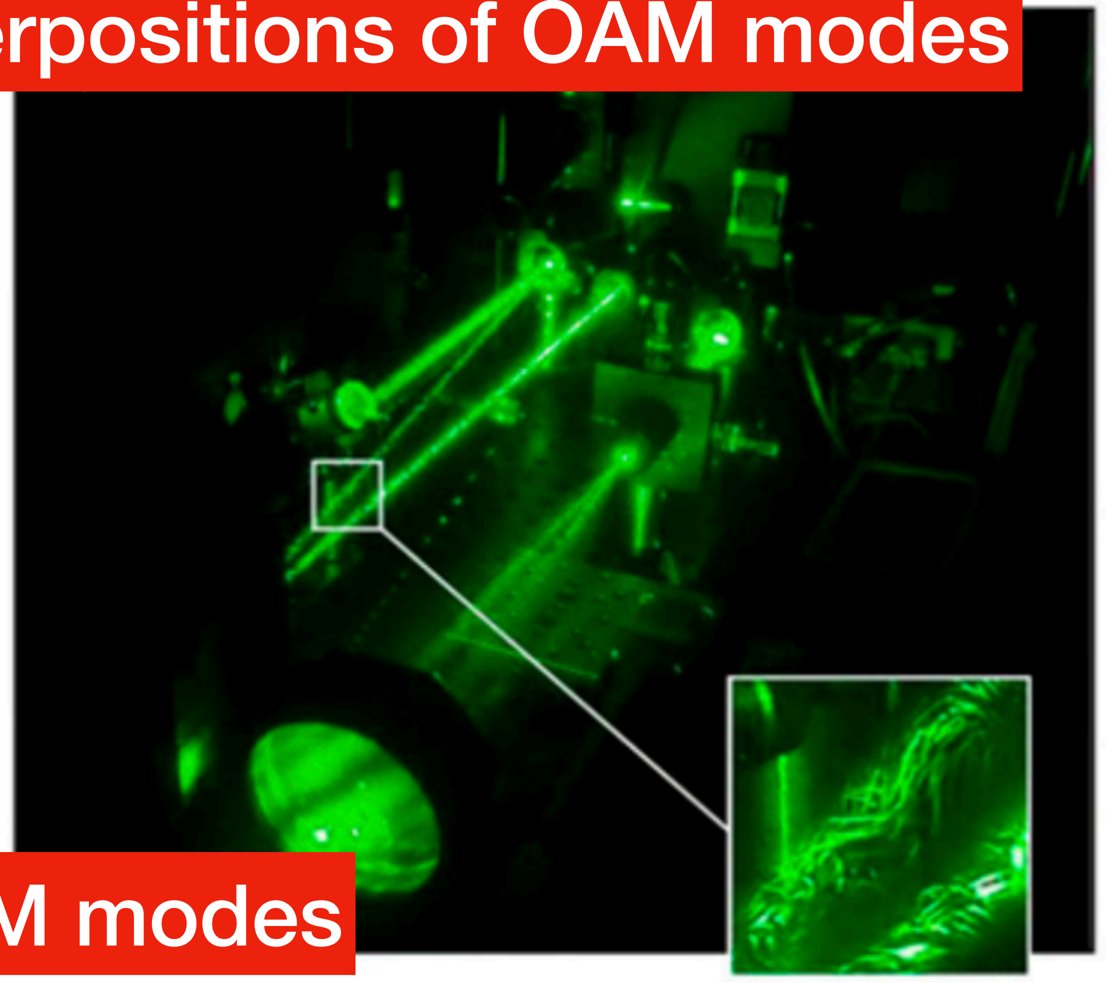
Spatial modes of light can potentially carry a vast amount of information, making them promising candidates for both classical and quantum communication. However, the distribution of such modes over large distances remains difficult. Intermodal coupling complicates their use with common fibers, whereas free-space transmission is thought to be strongly influenced by atmospheric turbulence. Here, we show the transmission of orbital angular momentum modes of light over a distance of 143 km between two Canary Islands, which is 50× greater than the maximum distance achieved previously. As a demonstration of the transmission quality, we use superpositions of these modes to encode a short message. At the receiver, an artificial neural network is used for distinguishing between the different twisted light superpositions. The algorithm is able to identify different mode superpositions with an accuracy of more than 80% up to the third mode order and decode the transmitted message with an error rate of 8.33%. Using our data, we estimate that the distribution of orbital

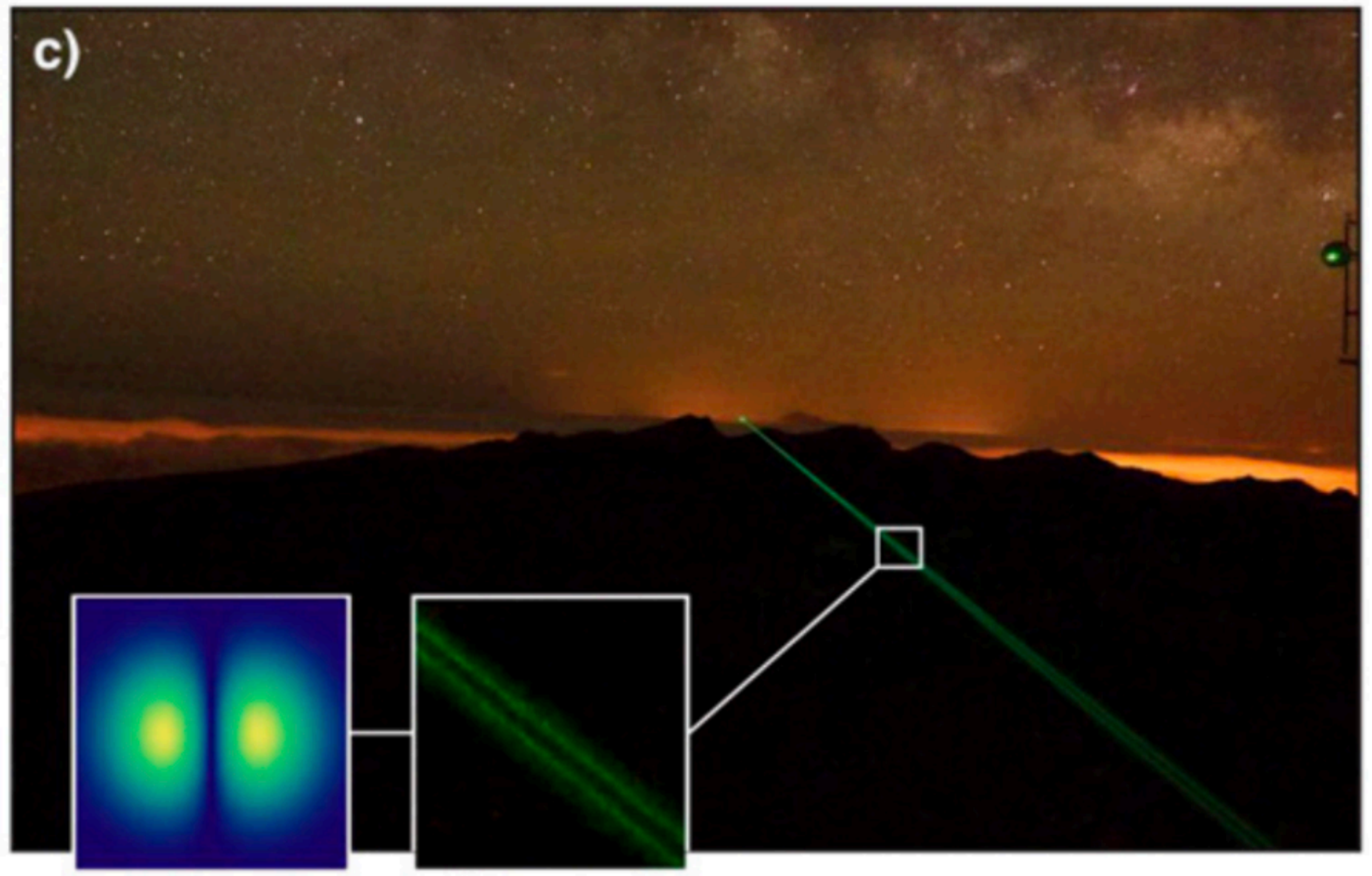
twisted radio waves was perfo (34). Single photons carrying transmitted over ∼210 m in ment in Padua, Italy (35). The large hall, as light in the vis influenced by the turbulenc More recently, 16 different s information for classical com link across Vienna (36). In th glement encoded in the OAM using the first two higher-orde that single-photon spatial coh spatial modes survive in a tu periment in Erlangen, Germa cross-talk of OAM beams ove classical communication exp





superpositions of OAM modes

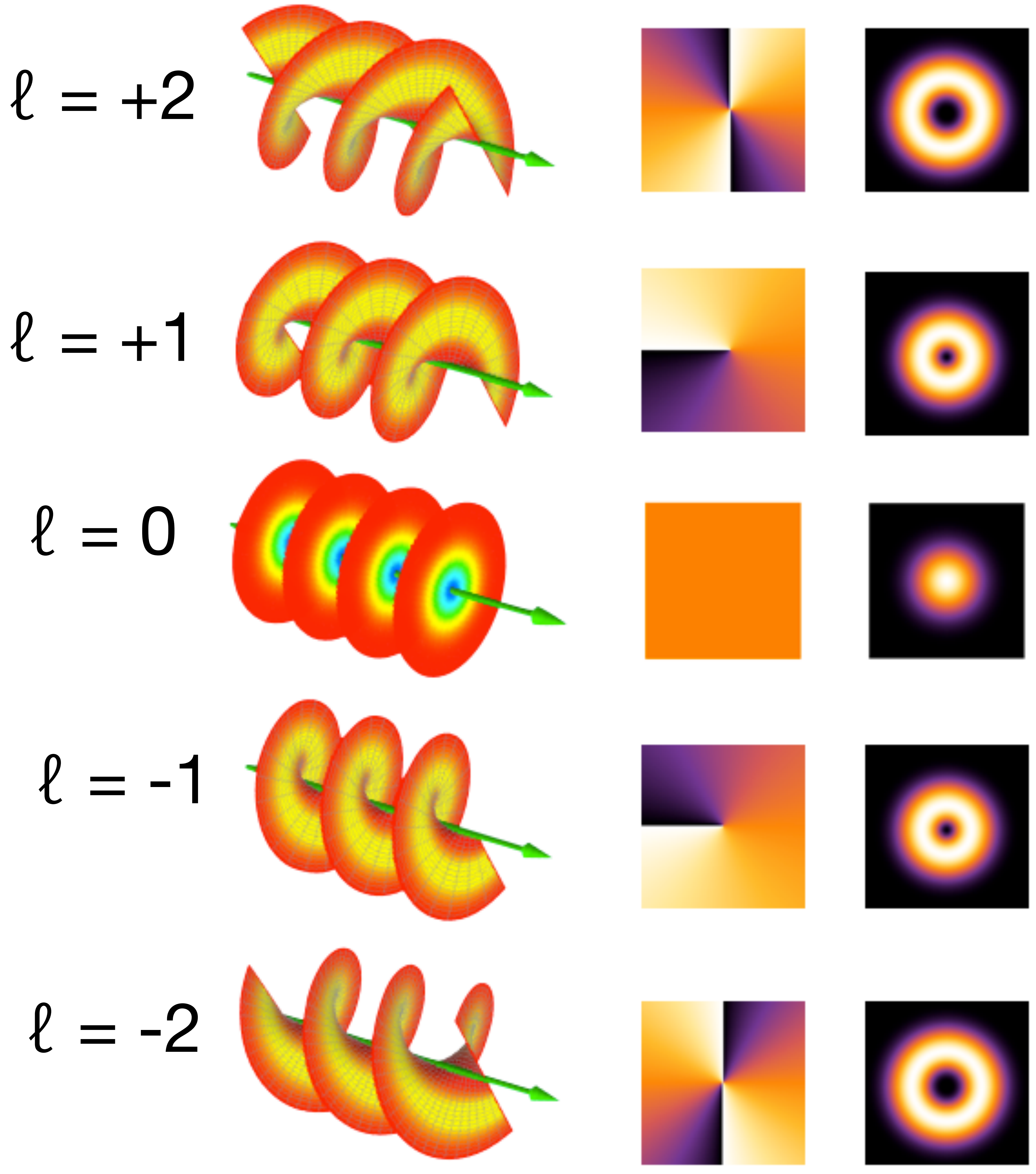




OAM modes

Twisted light / optical vortices / beams with OAM

- Beams carrying orbital angular momentum feature a helical wavefront
- Number of turns around axis (in one wavelength) equals the OAM state ℓ
- Intensity of optical vortices is shaped as a ring (or multiple rings)
- Phase singularity hidden at the core
- We can construct modes that (in the absence of turbulence) propagate without distortions: Laguerre-Gauss, Bessel-Gauss,…
- $$u(r,\varphi,z,t) = A(r,z,t)\exp(i\ell\varphi)$$



(Wikipedia)

Wavefront sensors generate histograms out of OAM superpositions

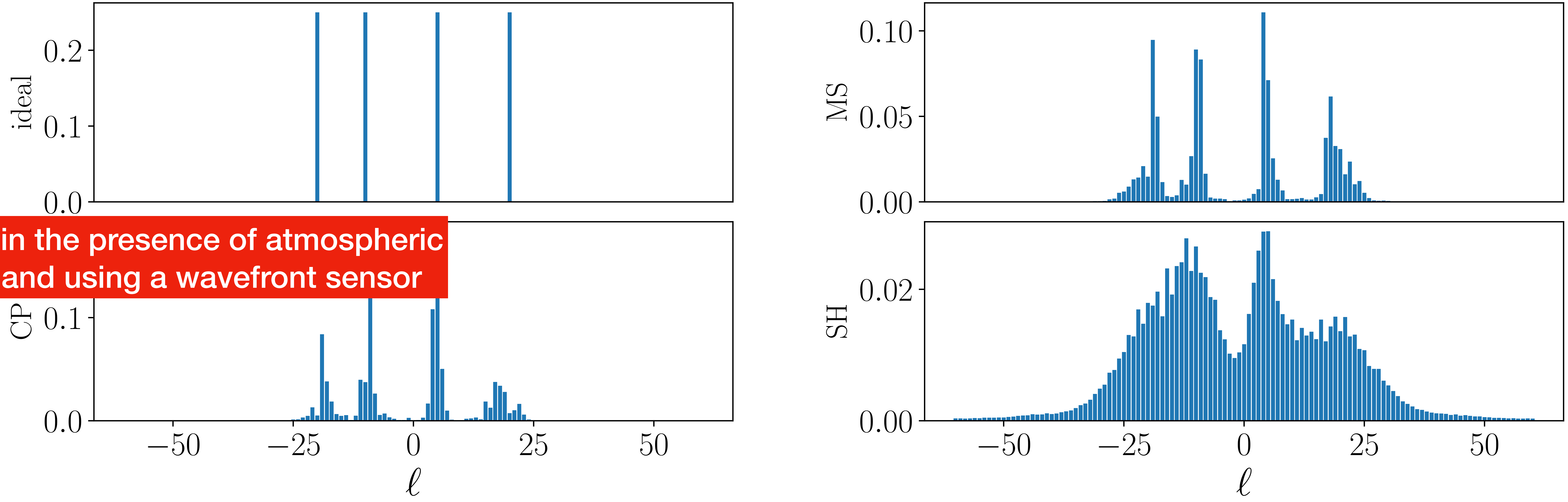
what we expect in the absence of turbulence

$$h_{\text{ideal},s}(\ell) := \frac{1}{|s|} \sum_{\ell=-\ell_{\max}}^{\ell_{\max}} \delta_{\ell \in s}$$

The set of histograms is the probability simplex :

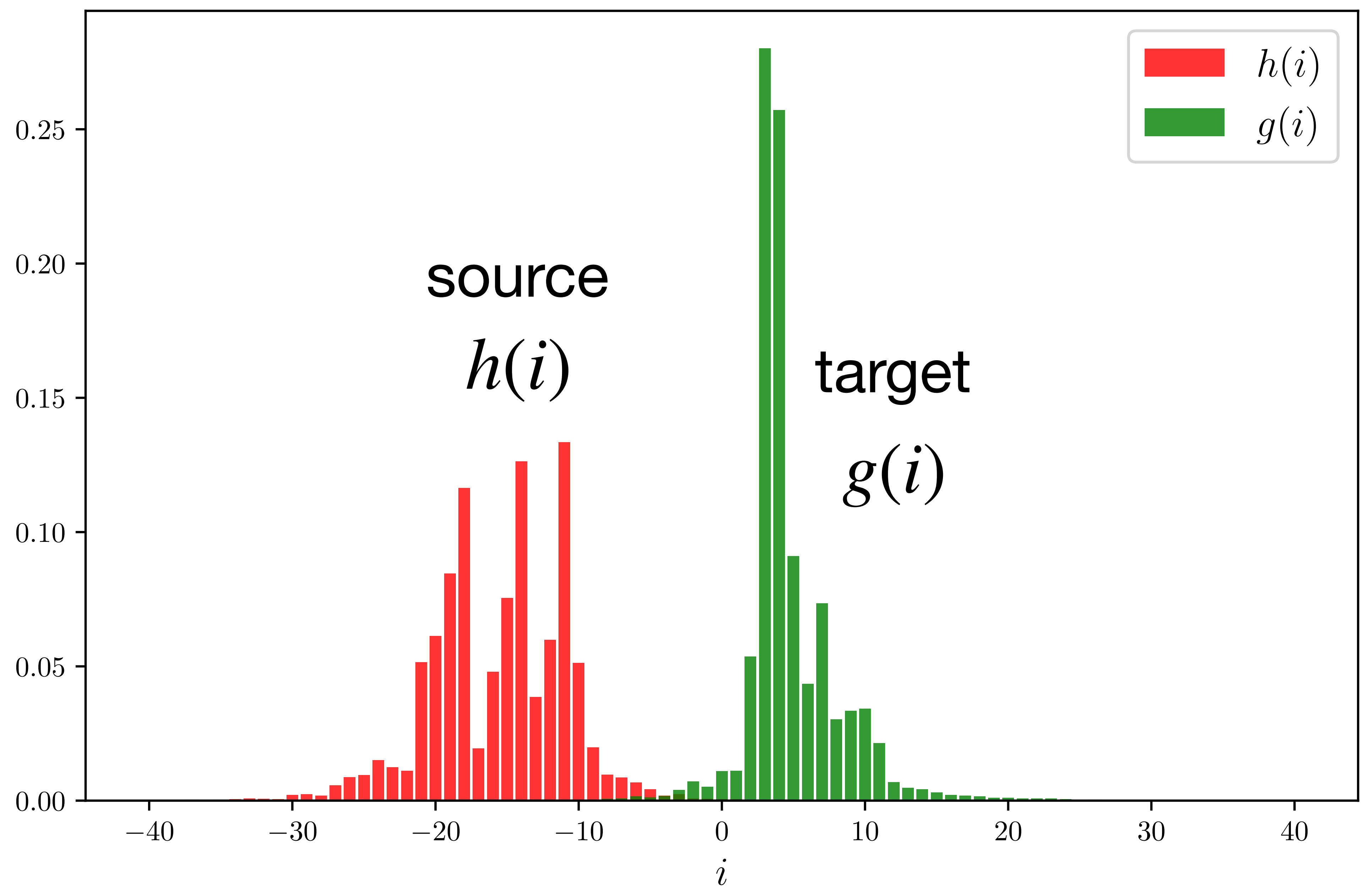
$$\Sigma = \left\{ h(\ell) \geq 0 \quad \text{for} \quad \ell \in \{-\ell_{\max}, \ldots, +\ell_{\max}\} \quad : \quad \sum_{\ell=-\ell_{\max}}^{\ell_{\max}} h(\ell) = 1 \right\}$$

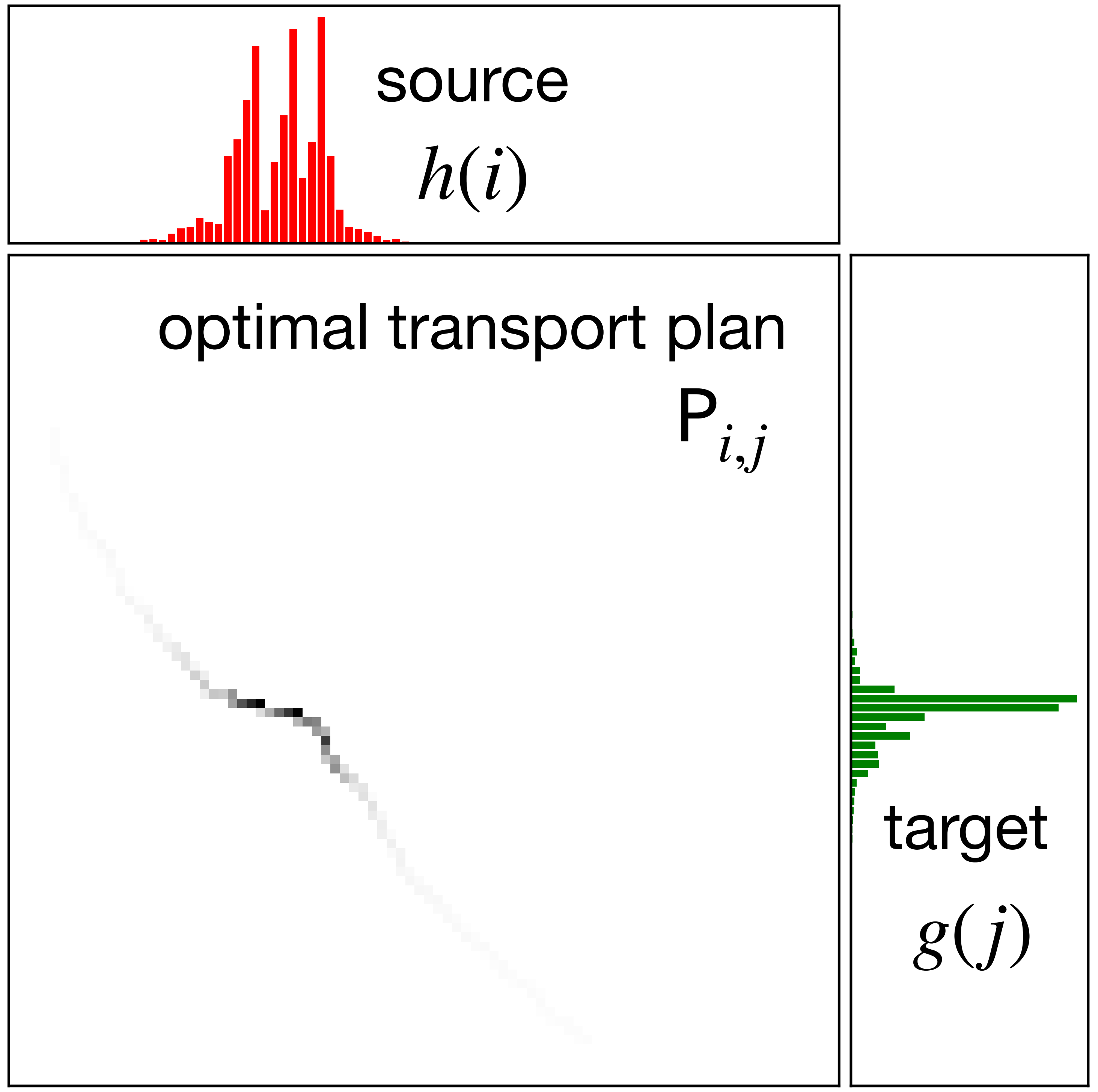
what we get in the presence of atmospheric turbulence and using a wavefront sensor



Here: ℓ in {-20,-10,5,20}, $C_n^2 = 3 \times 10^{-14}\ \mathrm{m}^{-2/3}, L = 1000\ \mathrm{m}$

Optimal Transport (illustration in 1d)





Ground cost,
distance between bins

The optimal transport plan P is the one that minimizes the total cost (interior-point linear optimization)

$$\min_{P \in U(h,g)} \sum_{i,j} C_{i,j} P_{i,j}$$

Here 1d, but can be generalized to histograms of any dimensions

The set of admissible transport plans

$$U(h,g) := \left\{ P \in \mathbb{R}_+^{N \times N} \ : \ \sum_j P_{i,j} = h(i) \ , \ \sum_i P_{i,j} = g(j) \right\}$$

Wasserstein distance

Let’s use for the cost a power…

$$\mathsf{C}_{\ell,\ell'} = d^p(\ell,\ell')$$

…of a well-defined ‘ground’ distance between bins

$$d(\ell,\ell') = d(\ell',\ell)$$

$$d(\ell,\ell') = 0 \iff \ell = \ell'$$

$$d(\ell,\ell'') \leq d(\ell,\ell') + d(\ell',\ell'')$$

…for instance

$$d(\ell,\ell') := |\ell' - \ell|$$

The p-Wasserstein distance:

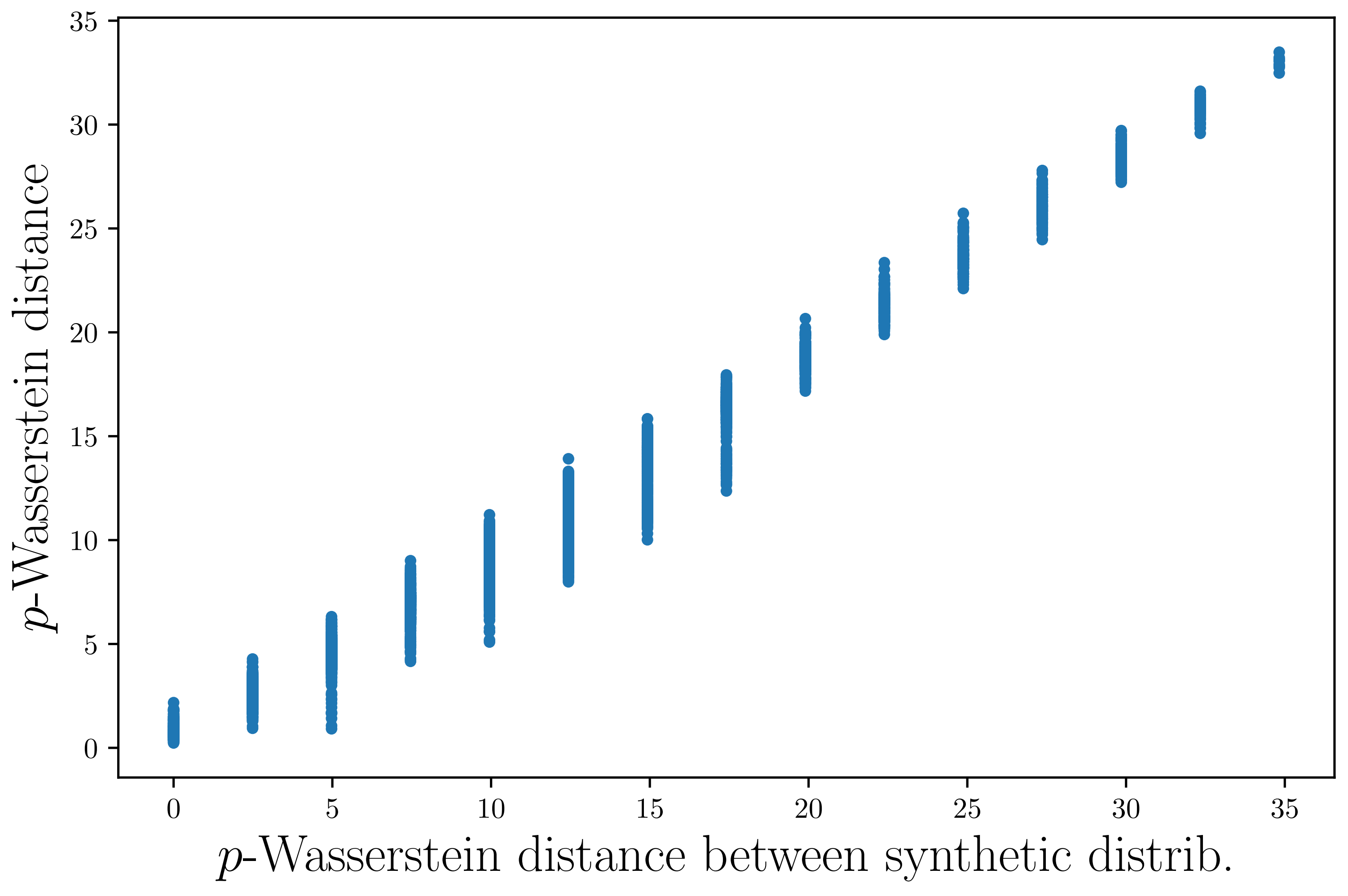
$$W_p(h,g) := \left(\min_{\mathsf{P}\in U(h,g)} \sum_{\ell,\ell'} d(\ell,\ell')^p \; \mathsf{P}_{\ell,\ell'} \right)^{1/p}$$

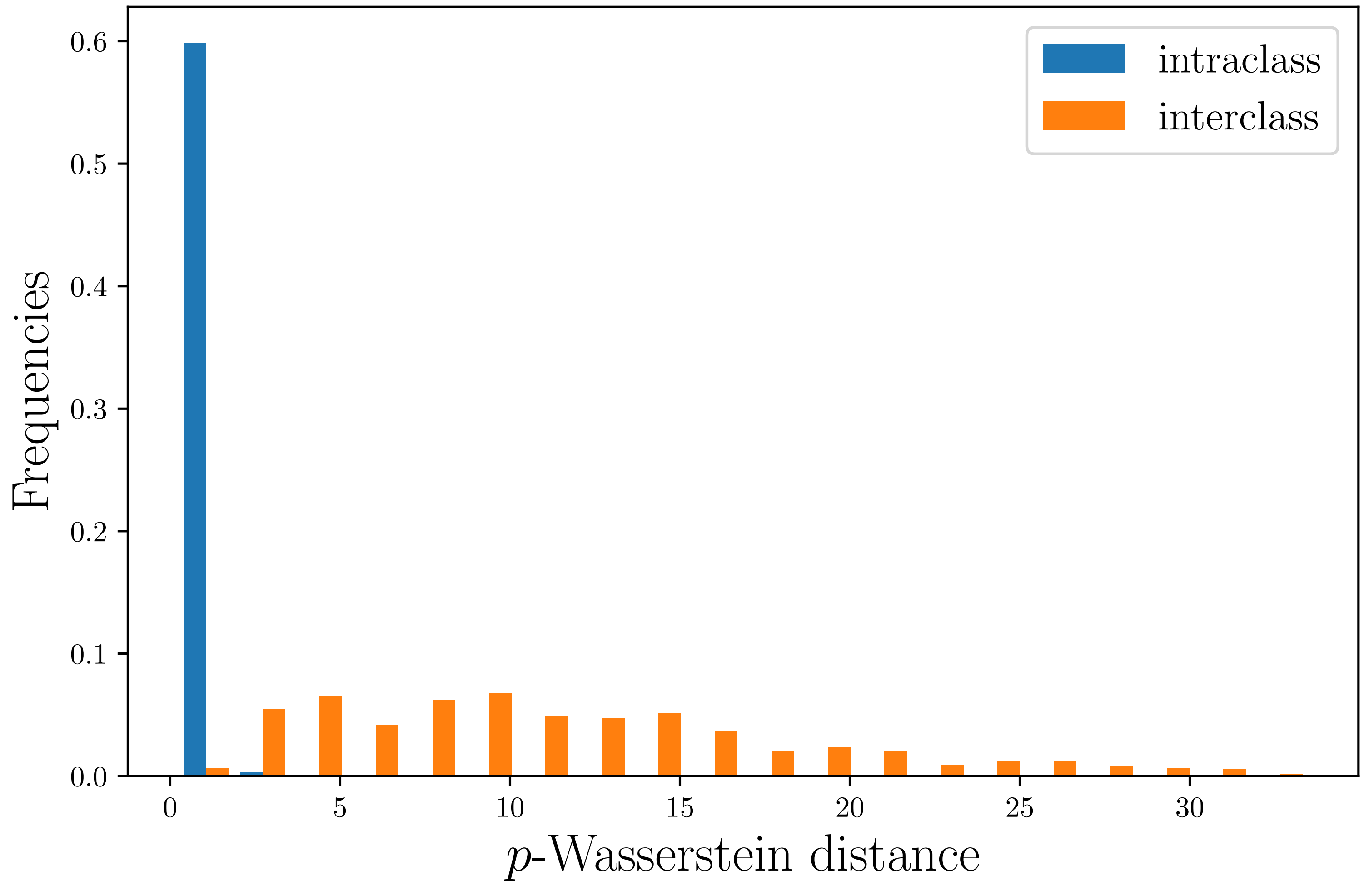
- For $p \geq 1$, this definition of distance verifies the three requisites of a **distance between histograms** *h, g* in Σ.
- It accommodates distortions induced by weak turbulence.
- It can be computed efficiently by several algorithms.
- Can include regularization terms.

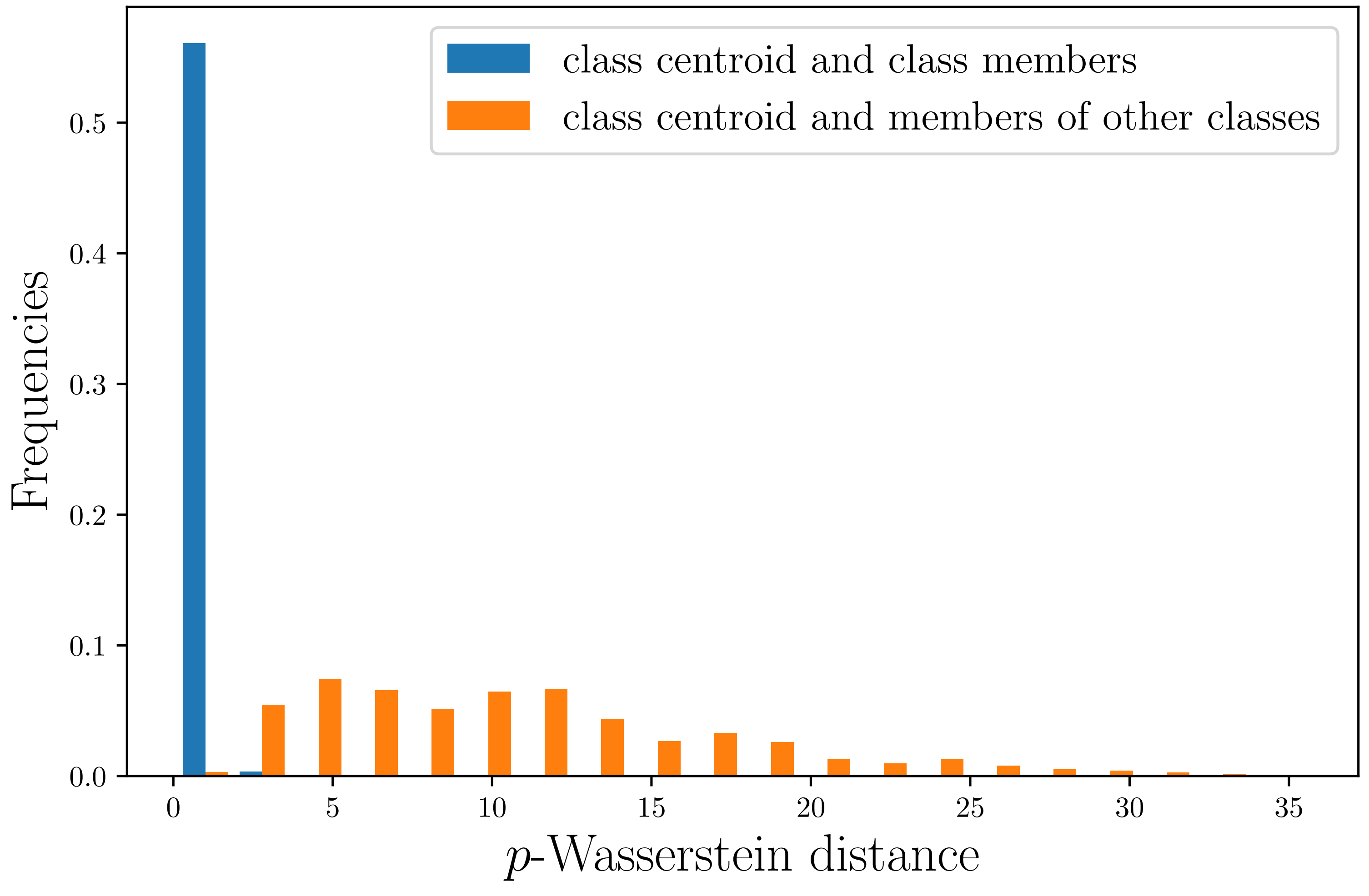
Wasserstein distance

allows the definition of the barycenter of a set of histograms

$$h_{\mathrm{bary},G} := \arg\min_{h\in\Sigma} \sum_{g\in G} \lambda_g W_p^p(h,g)$$







The Wasserstein distance between histograms is preserved during propagation.

Discriminates between histograms corresponding to different superpositions.

Barycenters/centroids provide a simple method of assessing the similarity between an histogram and a group of histograms.

These 3 figures correspond to the Shack-Hartmann histograms from a selection of 28 duets

Here: symbol = class = superposition

Selection of optimal subset of superpositions

A systematic method for selecting a subset of *n* (desirably a power of two) superpositions from a larger set of *N*:

We can pose an integer optimization problem where the entries of the binary vector *x* of length *N* indicate the presence (1) or absence (0) of a given superposition in the optimal subset:

$$\text{minimum} \left(x^T \mathrm{E} x\right), \quad \text{subject to } x \in \{0,1\}^N \text{ and } \sum_{i=1}^{N} x_i = n$$

$$\text{where} \quad \mathrm{E}_{s,t} := 1/W_p^q(h_s, h_t), \quad \text{if } s \neq t, \quad \text{and } 0 \quad \text{otherwise}$$

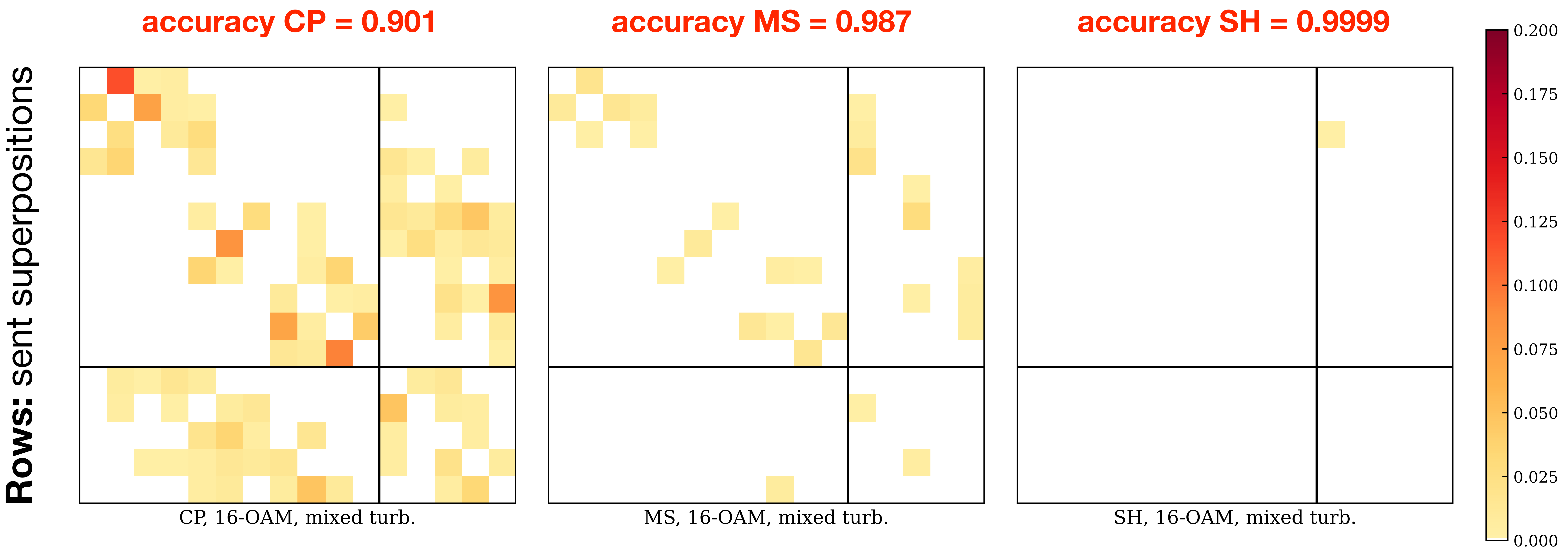
Here h_s could be the ideal spectra, or the barycenter from a training set of labeled spectra associated to the superposition *s*.

We have used a set composed of *N*=98 superpositions (28 duets and 70 quartets),
and the optimal subsets of sizes *n*=64, 32 and 16, always included combinations of duets and quartets.

Nearest Centroid Classification

Requires computation of distances between the new instance and the barycenters of the classes

here we optimize the classification scheme and found the best wavefront sensor



Columns: detected superpositions

A **confusion matrix** contains the result of a classification experiment.
Here diagonal elements are not shown because they are much larger.

$C_{s,d}$ number of observations known to be in class *s* and classified as in class *d*.

$$\text{accuracy} := \frac{\sum_d C_{d,d}}{\sum_{s,d} C_{s,d}}$$

fraction of correct classifications

Partial conclusions

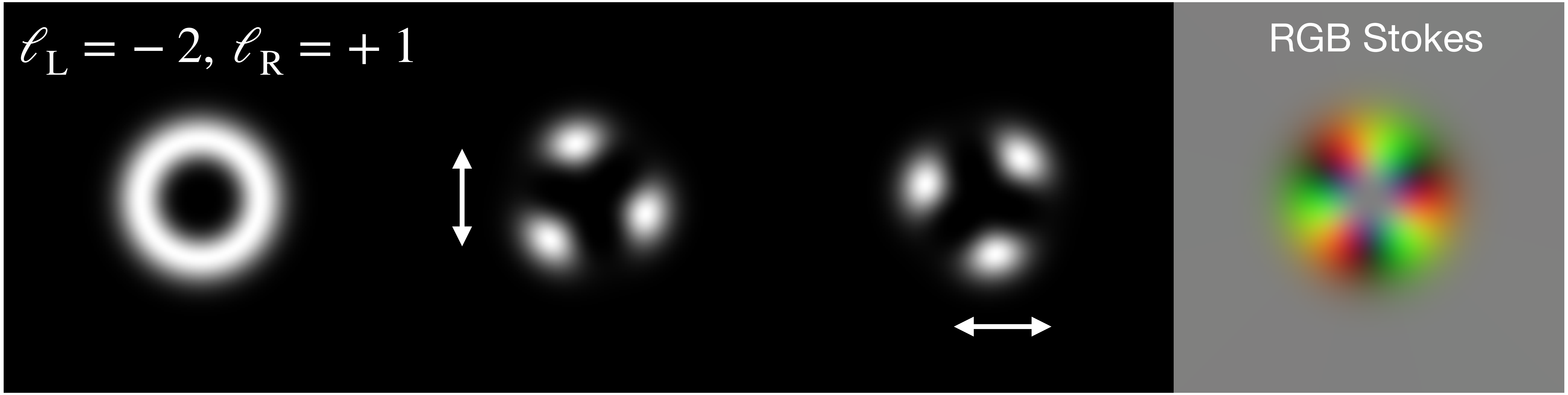
- The **theory of Optimal Transport** provides concepts and algorithms for the appropriate manipulation of empirical OAM spectra distorted by atmospheric turbulence.

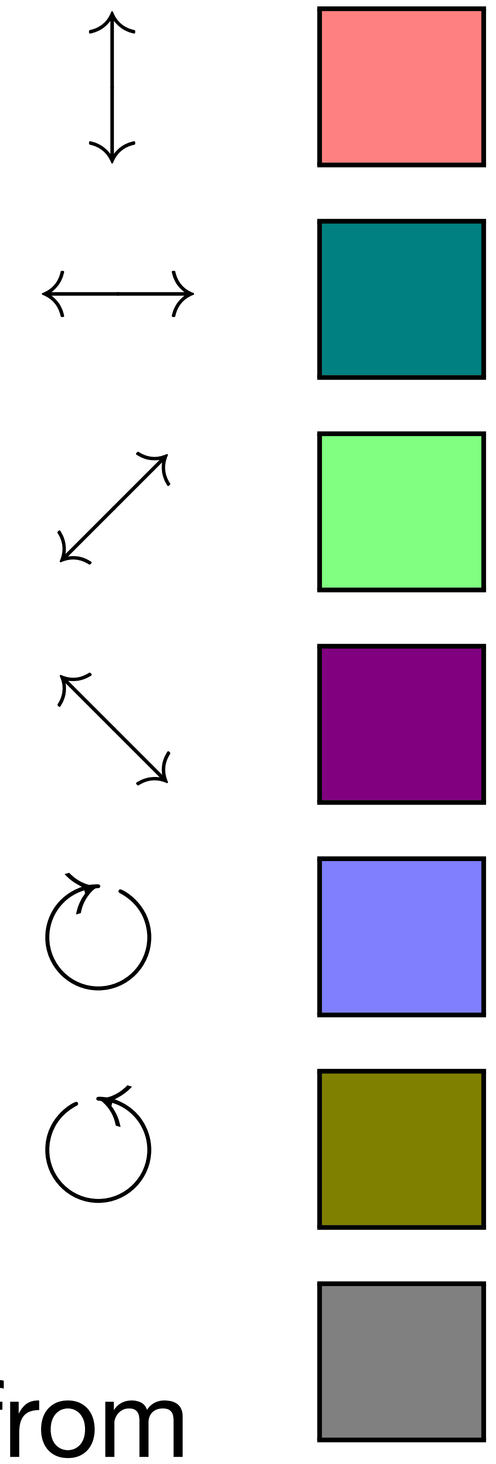
- The definition of a **Wasserstein distance**, and a **barycenter** based on such a distance, can be used to select the **best subsets of superpositions**, and more robust classification schemes based on such limited sets of classes

- The results based on the Wasserstein distance are **consistently better** than the ones obtained with the Kullback-Leibler divergence

Vector Vortex Beams

How can we identify a spatially structured beam with nonuniform polarization?





Does atmospheric turbulence destroy the structure?

Can we use a set/alphabet of VVB modes as data symbols for encoding info and implement a free-space optical link?

RGB color image constructed from the 3 Stokes parameters that characterize polarization

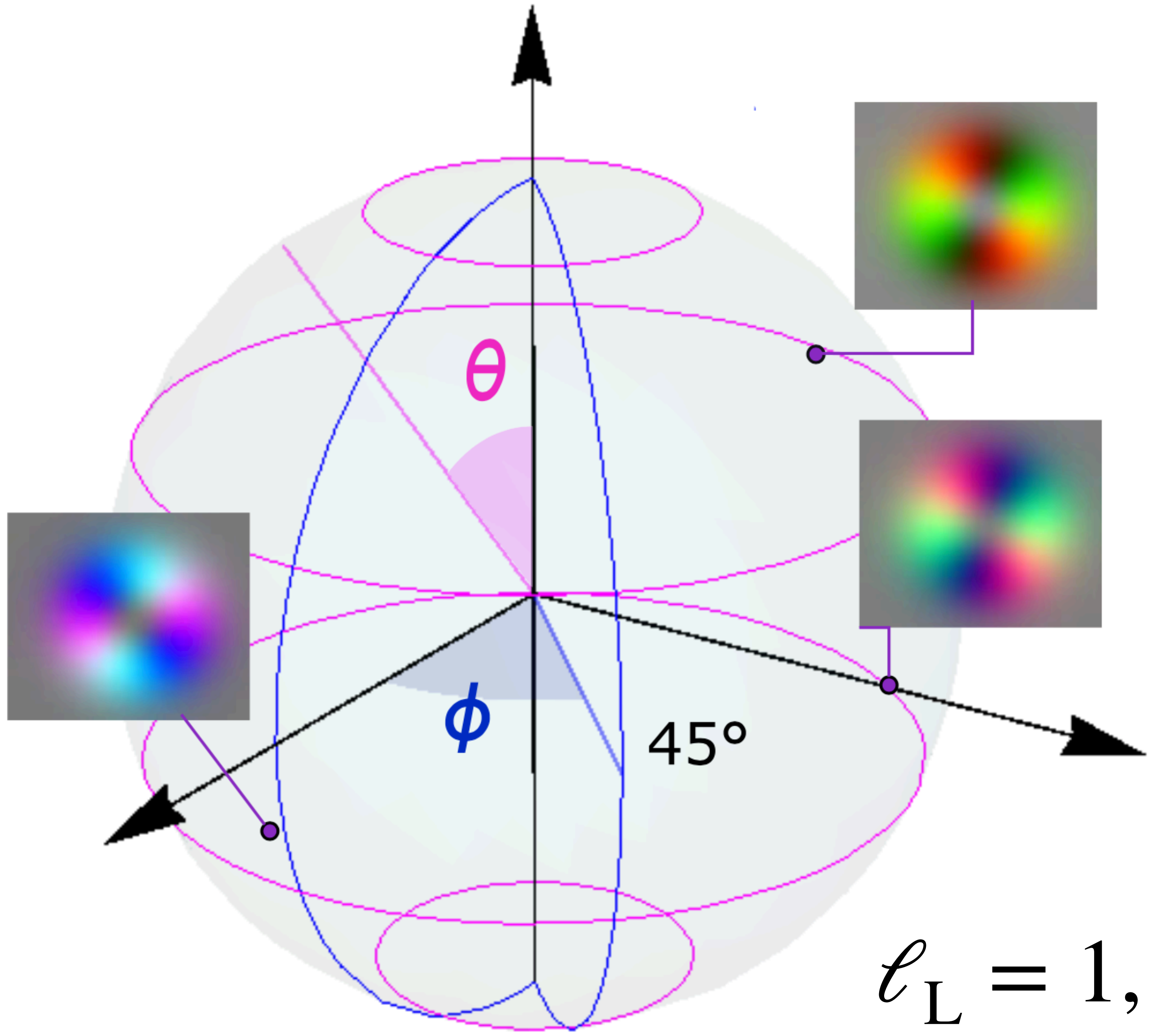
Vector Vortex Beam

Orthogonal polarization states

$$|\mathrm{VVB}\ (p_\mathrm{L},\ell_\mathrm{L},p_\mathrm{R},\ell_\mathrm{R},\theta,\phi)\rangle = \cos\frac{\theta}{2}\ |\mathrm{L}\rangle\ |\mathrm{LG}_{p_\mathrm{L},\ell_\mathrm{L}}\rangle + \sin\frac{\theta}{2}\ e^{i\phi}\ |\mathrm{R}\rangle\ |\mathrm{LG}_{p_\mathrm{R},\ell_\mathrm{R}}\rangle$$

$\ell_\mathrm{L},\ell_\mathrm{R} \in \mathbb{Z}$, $p_\mathrm{L},p_\mathrm{R} \in \mathbb{Z}_{0,+}$, $\theta \in]0,\pi[$, $\phi \in [0,2\pi]$

Orthogonal spatial states with orbital angular momentum (OAM) (Laguerre-Gauss)



$\ell_\mathrm{L} = 1,\ \ell_\mathrm{R} = -1$

For given two OAM indices, the angles *θ,ϕ* define a higher-order Poincaré sphere:

each point represents a nonuniform polarization pattern

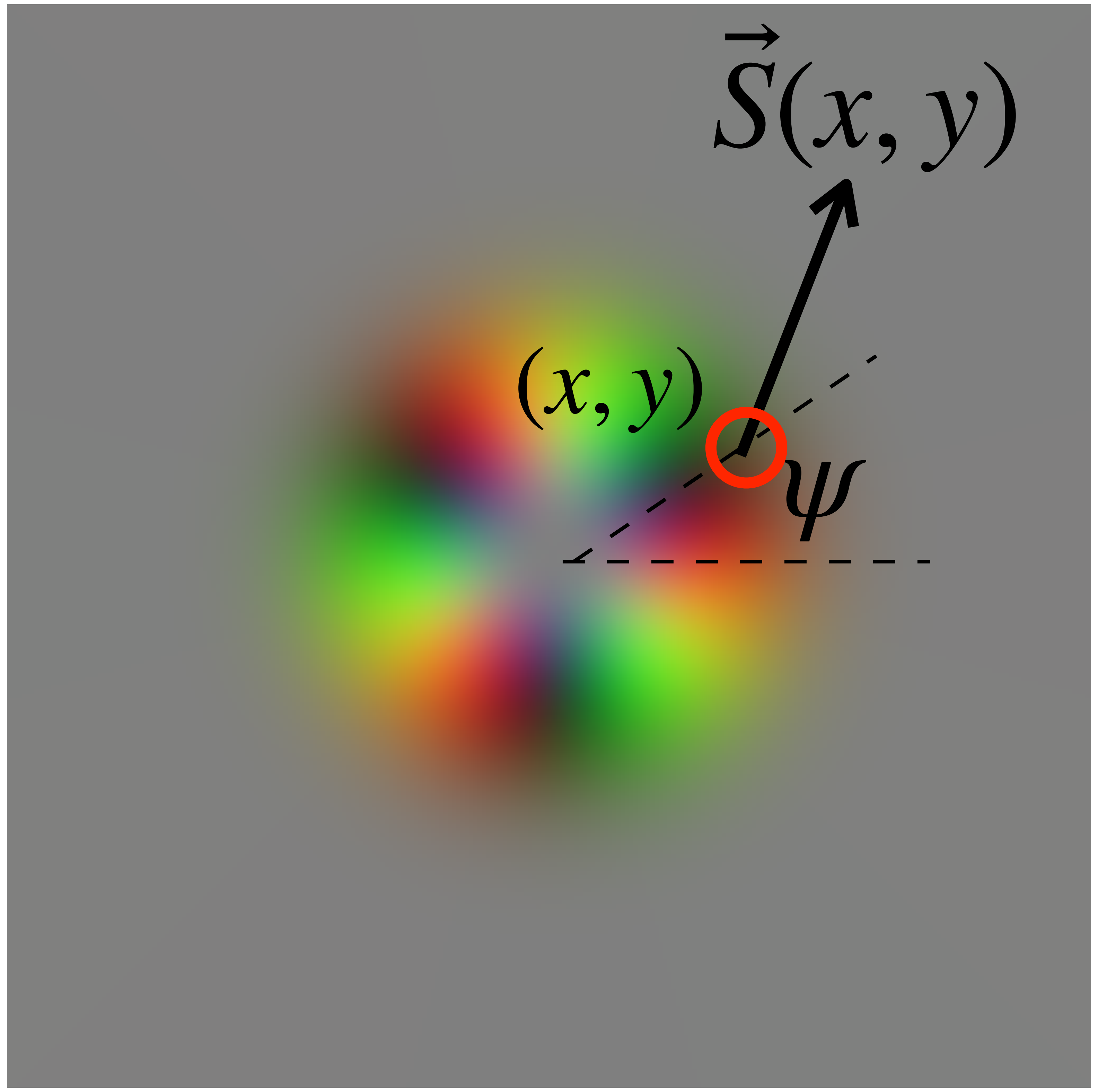
(figure extracted from Giordani 2020)

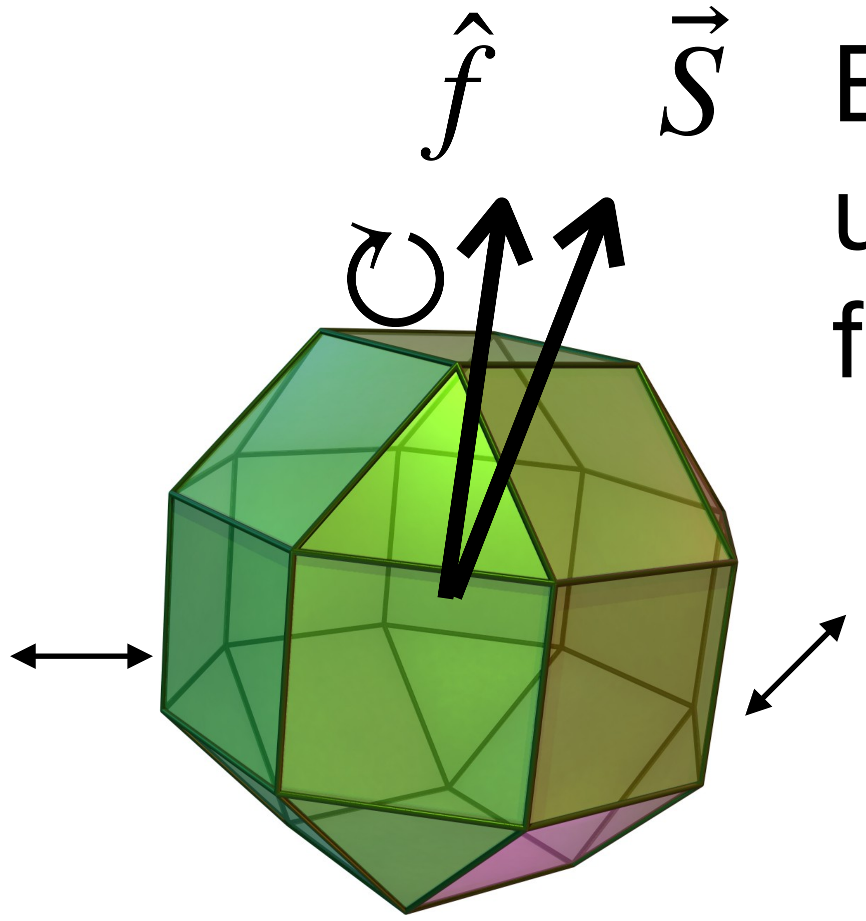
Approach for elliptical polarization

Use VVB defined with several choices of θ,ϕ (selection of points in Poincaré sphere) and then project the measured **nonplanar Stokes vectors** on the faces of a polyhedron (for instance a rhombicuboctahedron with 26 faces and 24 vertices)

The ground distance between bins: $k := (\psi, \hat{f})$

$$d(k,k') := |\psi' - \psi|_{2\pi} + \angle(\hat{f}', \hat{f})$$





Each pixel contributes a unit mass to a particular face of the polyhedron

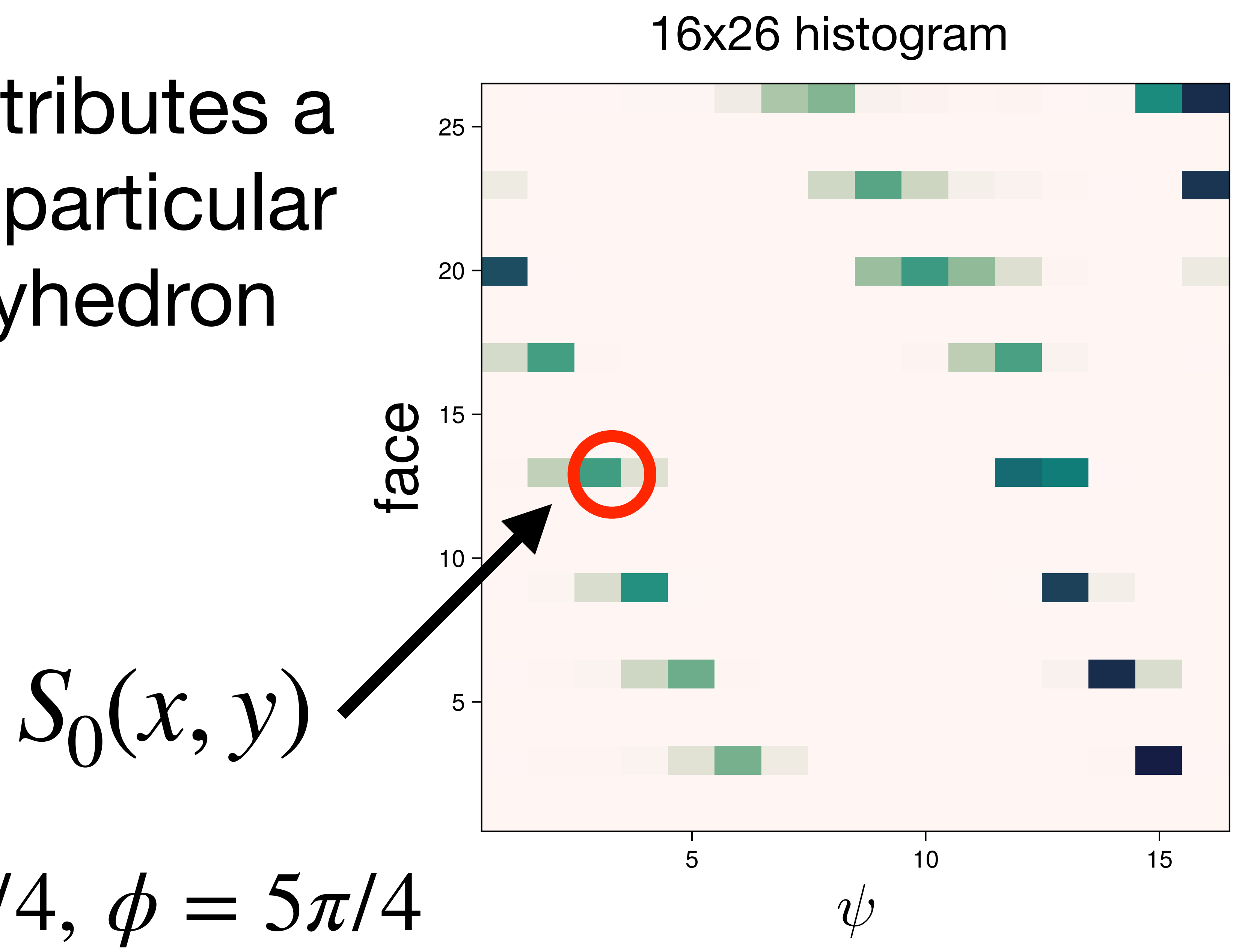


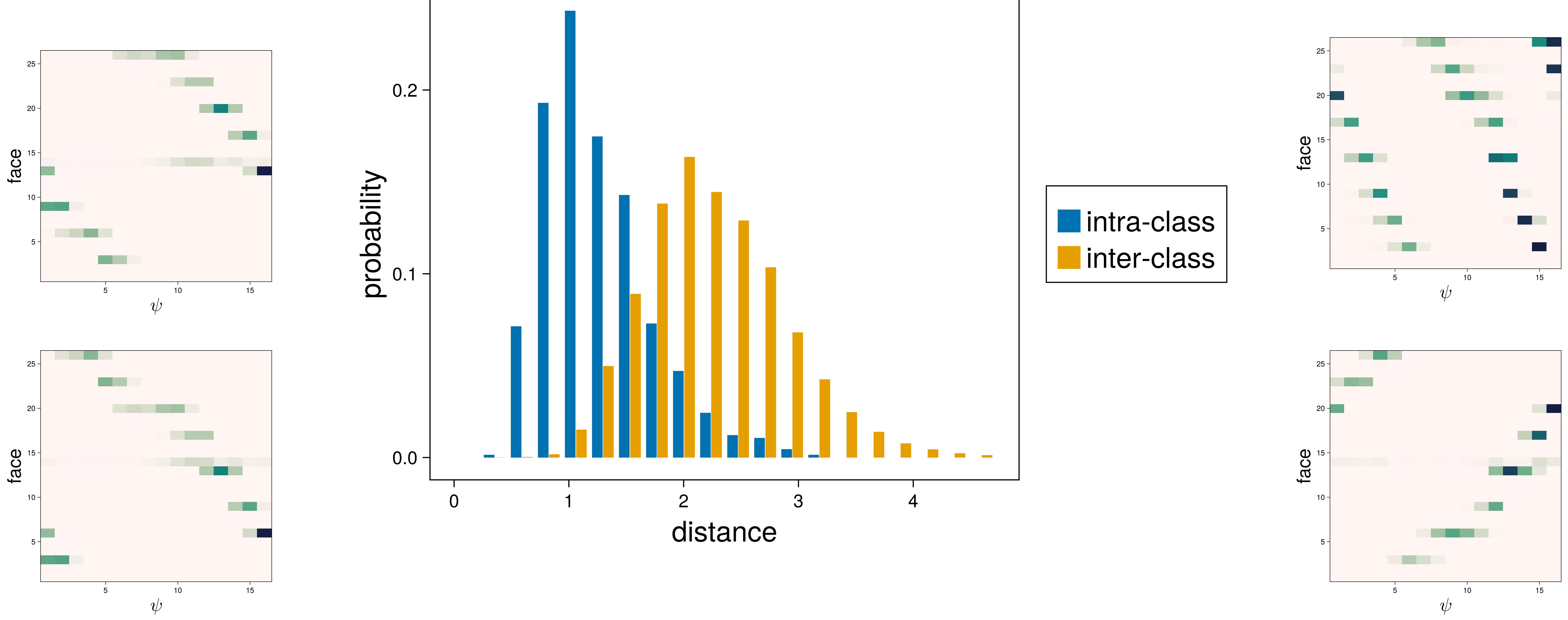
Figure polyhedron: Wikipedia

$\ell_L = +1,\ \ell_R = -1,\ \theta = \pi/4,\ \phi = 5\pi/4$

Statistics of Wasserstein distances

Pairs of instances belonging to the same or different classes (144)

$L = 1 \text{ km}, C_n^2 = 9 \times 10^{-15} \text{ m}^{-2/3}$



Selection of optimal subset of VVBs

A systematic method for selecting a subset of *n* (desirably a power of two) superpositions from a larger set of *NS*:

We can pose an integer optimization problem where the entries of the binary vector *x* of length *NS* indicate the presence (1) or absence (0) of a given superposition in the optimal subset:

$$\text{minimum}\ \left(x^T \mathsf{E} x\right),\ \ \text{subject to}\ x \in \{0,1\}^{N_{\mathcal{S}}}\ \text{and}\ \sum_{i=1}^{N_{\mathcal{S}}} x_i = n$$

$$\text{where}\ \ \mathsf{E}_{s,t} := \left(\frac{\langle W_p(h_s,\cdot\,)\rangle + \langle W_p(\,\cdot\,,h_t)\rangle}{W_p(h_s,h_t)}\right)^q,\ \ \text{if}\ s \neq t,\ \ \text{and}\ 0\ \ \text{otherwise}$$

Here h_s could be the barycenter of a training set of labeled spectra associated to the superposition *s*.

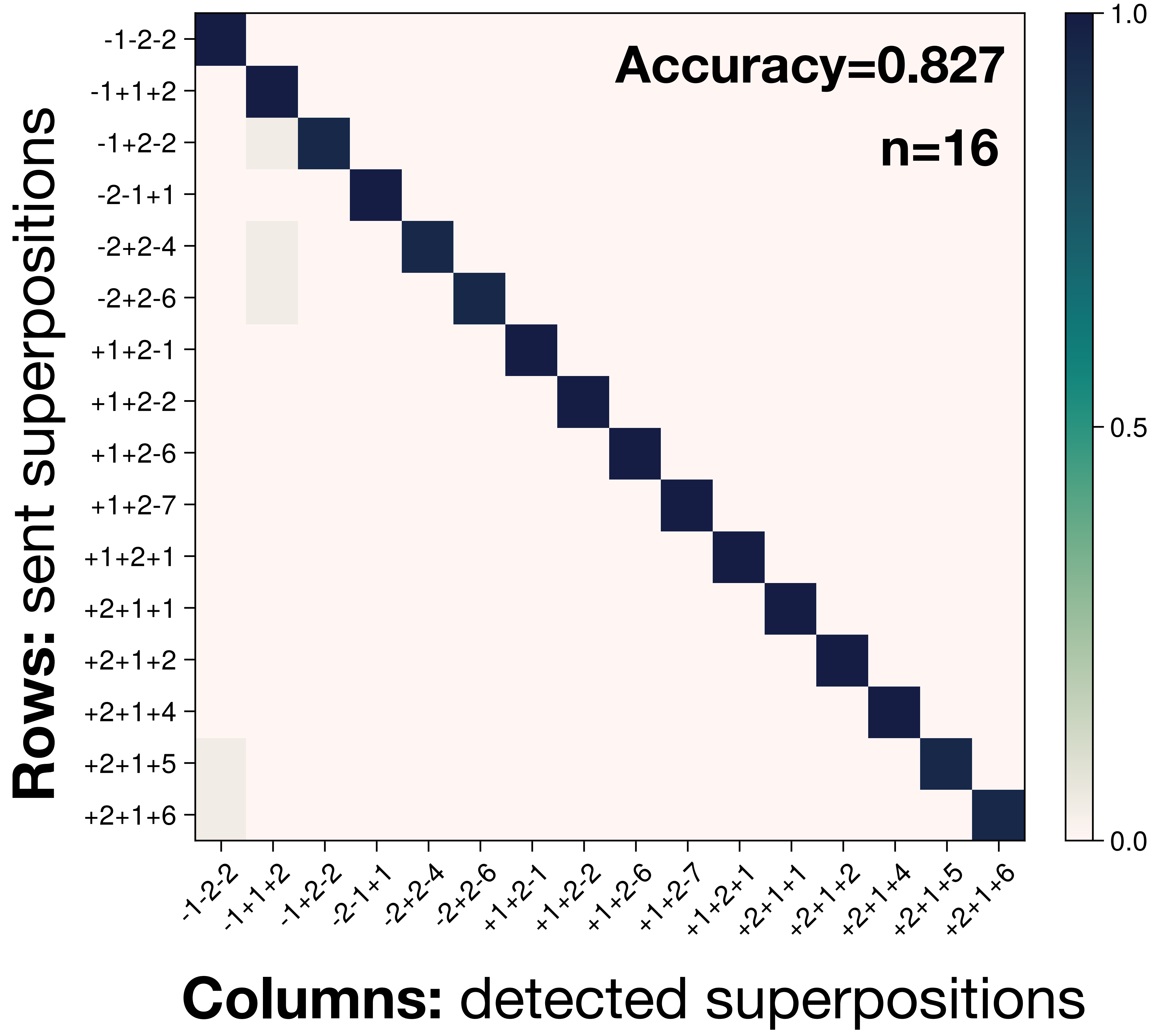
We have used a set composed of *NS*=144 classes of VVBs
(12 selections of *(ψ,ϕ)* times 12 combinations of OAM).

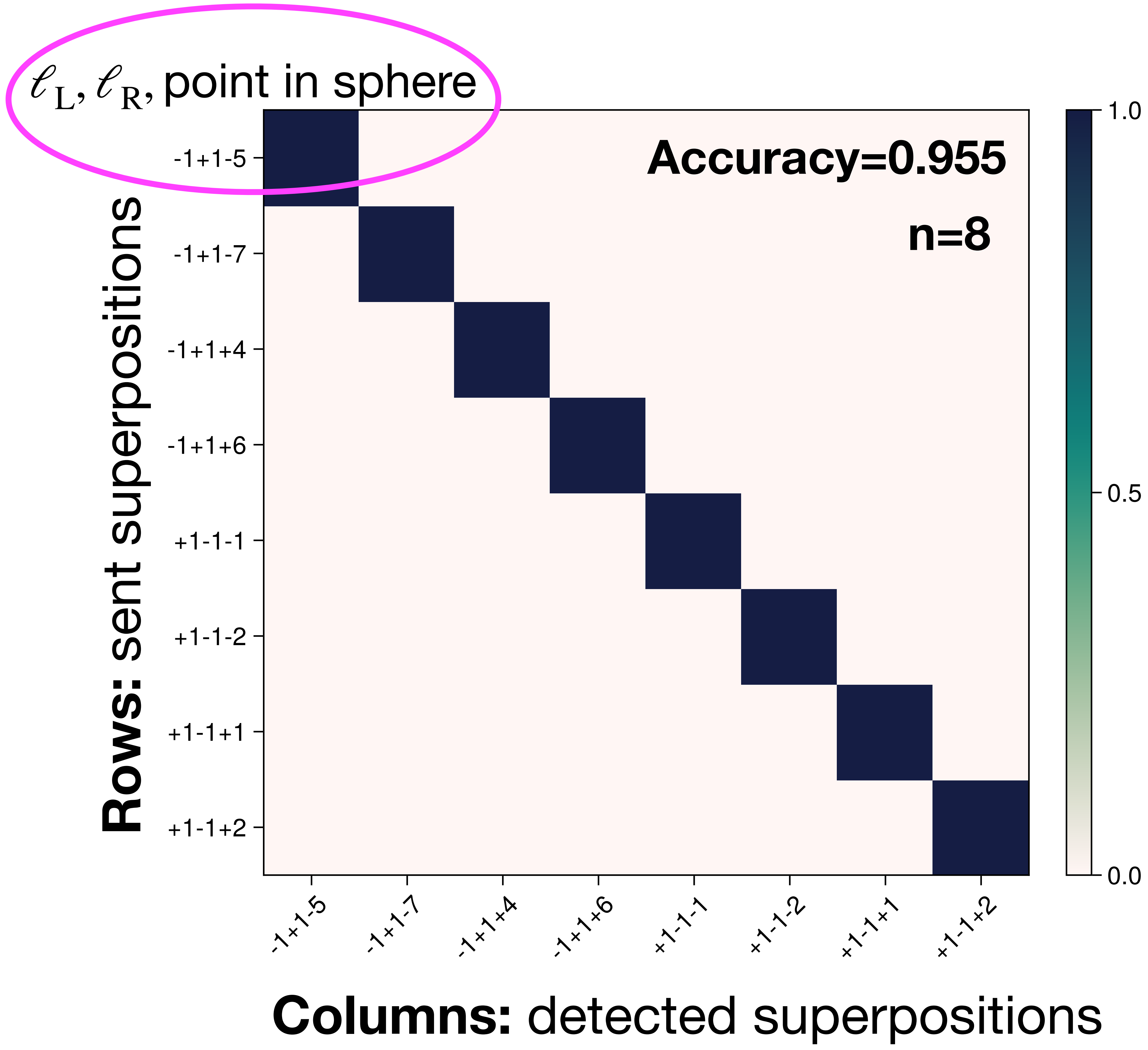
Nearest Centroid Classification

$$\hat{d}(g) := \arg\min_{d} W_p^p(g, h_{\text{bary},d})$$

Requires computation of distances between the new instance and the barycenters of the classes

$$L = 1 \text{ km}, C_n^2 = 2.9 \times 10^{-14} \text{ m}^{-2/3}$$





(For moderate turbulence strengths we obtained perfect classification, zero cross-talk)

Conclusions

- It is possible to construct **2-d histograms** based on empirical **Stokes parameters** obtained from **vector vortex beams** (VVBs) distorted by atmospheric turbulence, for **linear** and **elliptical** polarization.

- The **theory of Optimal Transport** provides concepts and algorithms for the appropriate manipulation of the histograms: definition of a **Wasserstein distance**, and a **barycenter** based on such a distance.

- We can select the **best subsets of superpositions**,

- and more robust **classification** schemes based on such limited sets of classes.

Many thanks for your attention! All questions and inquires are welcome!