



HAL
open science

Cloud of points as discrete measures for Gaussian models and stochastic optimization

Babacar Sow, Rodolphe Le Riche, Sanaa Zannane, Merlin Keller, Julien Pelamatti

► **To cite this version:**

Babacar Sow, Rodolphe Le Riche, Sanaa Zannane, Merlin Keller, Julien Pelamatti. Cloud of points as discrete measures for Gaussian models and stochastic optimization. MASCOT-NUM2023, Apr 2023, LE CROISIC, France. hal-04102275

HAL Id: hal-04102275

<https://hal.science/hal-04102275v1>

Submitted on 6 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Context and problem



- A cloud of points is a sample of n unordered points $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ with $\mathbf{x}_i \in \mathbf{R}^d$.
- Metamodel functions defined over clouds of points using Gaussian process.
- Design stochastic, evolutionary, algorithms with operators based on Wasserstein barycenter.
- Note the set of clouds of points as \mathcal{X} and the set of finite discrete measures defined over \mathbf{R}^d as \mathcal{P}

An evolutionary algorithm

- Clouds can be viewed as measures through the bijective mapping: $\varphi: \mathcal{X} \rightarrow \mathcal{P}$
 $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \mapsto P_X = \frac{1}{n} \sum_{i=1}^n \delta_{\mathbf{x}_i}$
- Given two clouds of points X and X' , we denote $W_2^2(P_X, P_{X'})$ the 2-Wasserstein distance between P_X and $P_{X'}$ (see [3] for details)
- **Crossover as Wasserstein barycenter:** $X'' = \mathbf{W}\text{-barycenter}(X, X')$ where X'' is the discrete measure minimizing $\frac{1}{2}(W_2^2(P_{X''}, P_X) + W_2^2(P_{X''}, P_{X'}))$.
 X'' can have different sizes according to X and X' (see [2] for details).
- **Mutation:** with probability p , do $X''' = \mathbf{W}\text{-barycenter}(X'', (X'' \cup \text{Dom}))$. Dom is a set of points at the domain boundary. For a square domain, Dom can be the corners.

The crossover promotes recombination of shapes while the mutation increases the spread of clouds.

Wind-farm test functions

F_{90} , F_{45} , F_0 and F_{40d} are analytical functions $\mathcal{X} \rightarrow \mathbb{R}$ mimicking the power production of wind farms with wind directions 90° , 45° , 0° and 40 averaged directions, respectively.

References

- [1] Muandet, Krikamol and Fukumizu, Kenji and Sriperumbudur, Bharath and Schölkopf, Bernhard and others: *Kernel mean embedding of distributions: A review and beyond*, Foundations and Trends® in Machine Learning
- [2] Cuturi, Marco and Doucet, Arnaud Fast computation of Wasserstein barycenters. *International conference on machine learning*
- [3] Panaretos, Victor M and Zemel, Yoav, Statistical aspects of Wasserstein distances. *Annual review of statistics and its application*

MMD-based Gaussian process through measure modeling

- Let \mathcal{H} be a Reproducing Kernel Hilbert Space with a characteristic kernel $k_{\mathcal{H}}$ such as Matérn 5/2. The characteristic nature guarantees the injectivity of the embedding map [1]: $P_X \mapsto \mu_X = \int P_X(x)k_{\mathcal{H}}(x, \cdot)dx$. θ_1 and θ_2 are 2 hyper-parameters scaling the dimensions between two points \mathbf{x} and \mathbf{x}' through $|x_1 - x'_1|/\theta_1$ and $|x_2 - x'_2|/\theta_2$.
- The correlation kernel $K(X, X') = \sigma \exp(-\frac{\|\mu_X - \mu_{X'}\|_{\mathcal{H}}^2}{2\theta^2})$ is symmetric and semi-definite positive.
- Learned on 300 clouds of points, we get the following correspondence between the vector θ and the functions. The $Q2$ values obtained by prediction on 1000 unseen (but of same type) clouds of points are

Functions	F_{90}	F_{45}	F_0	F_{40d}
$Q2$	0.899	0.868	0.906	0.906

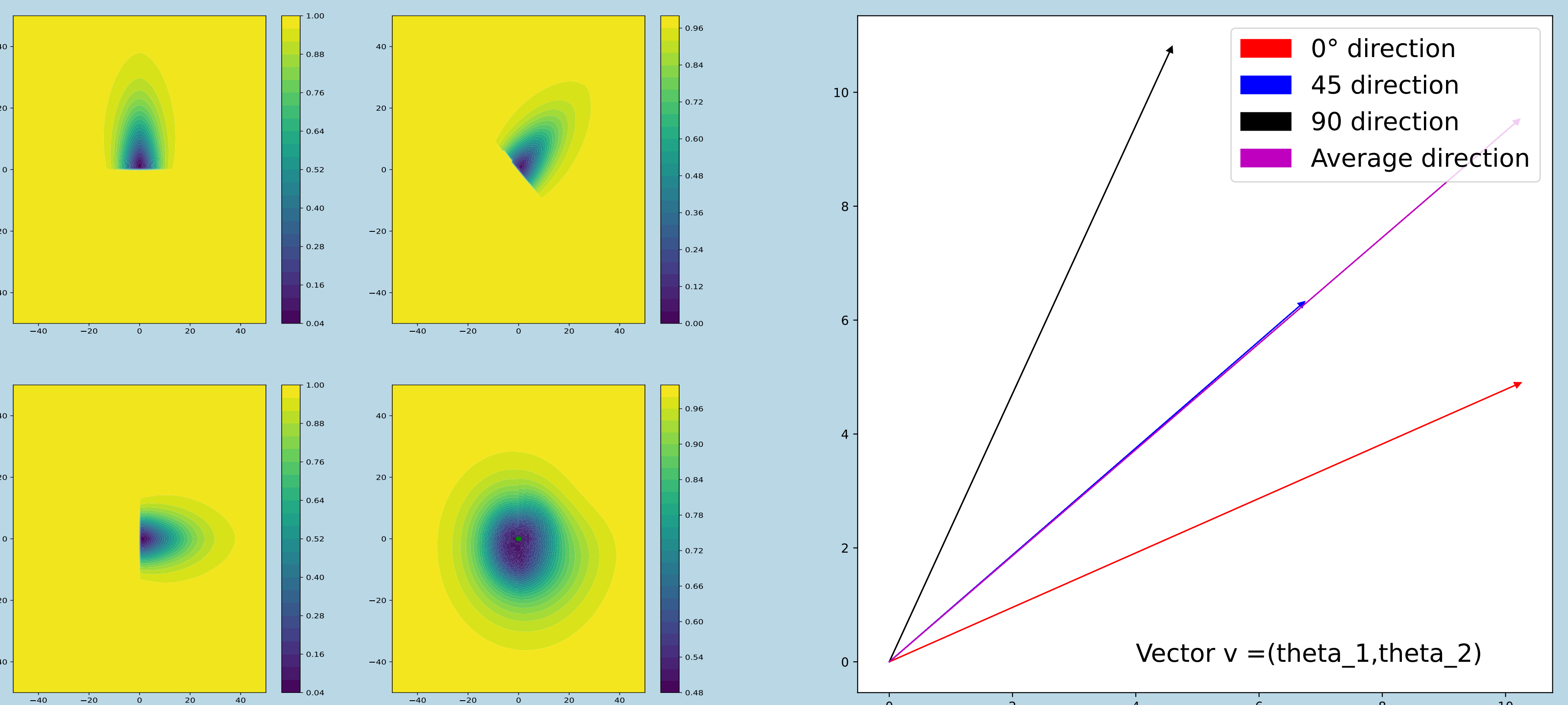


Figure 1: Vectors of length scales of the MMD embedding kernel learned by maximum likelihood on the wind farm proxy for various wind directions. Left: reminder of the turbine contributions for winds at 90° , 45° , 0° and 40 directions (left to right, top to bottom). Right: $(\theta_1, \theta_2)^\top$ vectors of length scales of the embedding kernel.

Best designs found with the evolutionary optimizer

- We add the two following test functions $F_{inert}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) = \sum_{i=1}^n \|\mathbf{x}_i - \bar{X}\|^2$ and $F_{minDist}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) = \min_{i \neq j} \|\mathbf{x}_i - \mathbf{x}_j\|$.
- The following designs correspond to the ones found respectively on F_0 , F_{45} , F_{90} , F_{40d} , F_{inert} and $F_{minDist}$. The inputs are clouds of points (in squares) with varying size (10 to 20).
- We observe an adaptation of the returned designs to the optimized functions. The points are placed optimally according to the wind's direction for wind-farms analytical functions. The algorithm returns also the optimal design for F_{inert} .
- **These results illustrate that modeling clouds of points as finite discrete measures allow their processing without loss of geometrical information.**

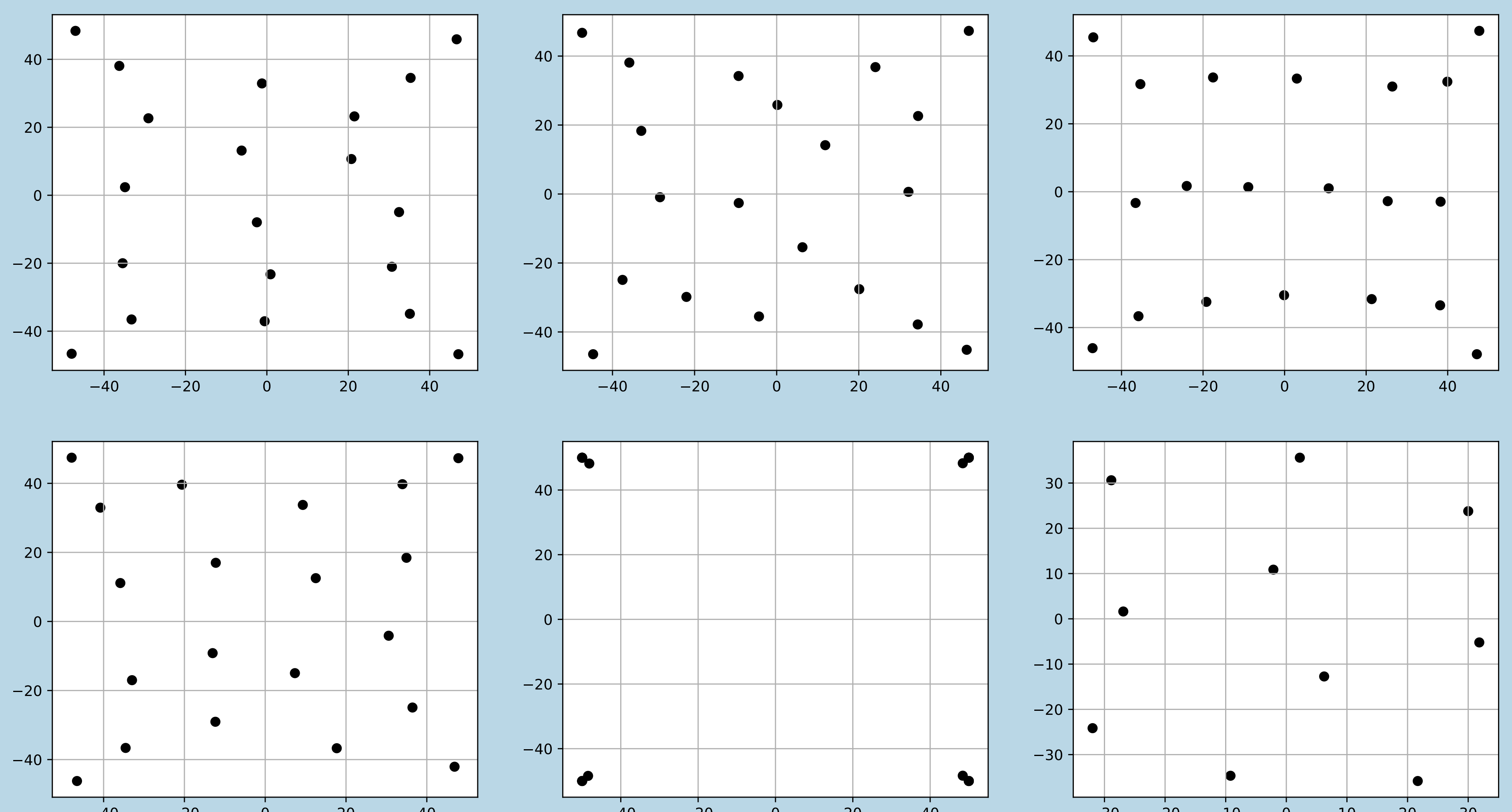


Figure 2: The designs correspond respectively to F_0 , F_{45} , F_{90} , F_{40d} , F_{inert} and $F_{minDist}$ (left to right, top to bottom).