Novelty Search In Infinite Dimensional Spaces

Pete Grindrod CBE

Mathematical Institute

University of Oxford

Oxford Mathematics

Abandoning Objective Functions

Prizing Novelty

- Novelty Search*: what is it*?

- Meta objectives: (i) range of exploration and (ii) leaving “no holes” or no systematic holes (iii) a feasible (tolerable) set

- Searching within distinct types of parameter space: for non-unique novel solutions (random element of Markov serach chains)

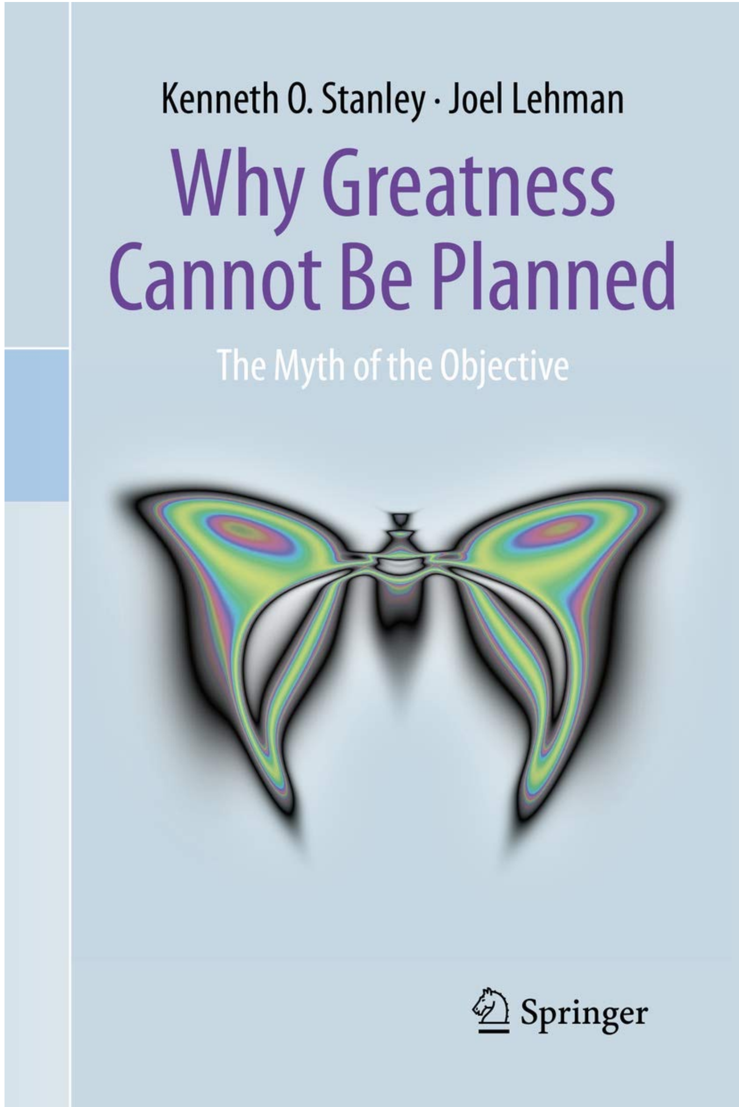
- Applications to search and generating novel solutions to problems

- General/Strategic applications to innovation and risk-taking

- Aim is to give *Advice*

** NOT to be confused with “novelty detection” : single class (normal) supervised discrimination*

Oxford Mathematics



ELSEVIER
Communications in Nonlinear Science and Numerical Simulation
Volume 125, October 2023, 107339

Research paper

Desperately searching for something

Clive E. Bowman, Peter Grindrod

Show more

+ Add to Mendeley   Share   Cite

https://doi.org/10.1016/j.cnsns.2023.107339   Get rights and content

Highlights

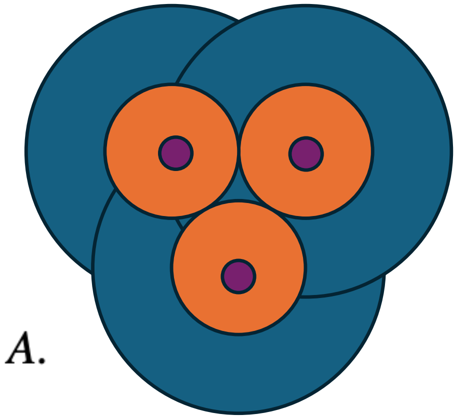
- Novelty search is a form of parameter sampling without a pre-defined objective function, with the meta objectives of coverage (growing reach) and density (having no unexplored regions).
- Methods deployed must depend critically on the nature of the parameter space to be sampled.
- As the dimension of the space increases the desire for coverage outweigh that for density.

Oxford Mathematics

Parameter space – any metric space

For any subset $A \subset U$, the ε-neighbourhood of A is the union of the ε-neighbourhoods of all points in A:

$$A_\varepsilon = \cup_{a \in A} N_\varepsilon(a) = \{x \in X | \min_{a \in A} \delta(x,a) < \varepsilon\}.$$



Definition

i An ε-cover is a subset $A \subset U$ such that $X = A_\varepsilon$.

ii An ε-packing is a subset $A \subset X$ such that $\delta(a,b) > 2\varepsilon$ for all $a, b \in A$.

iii An ε-net is a subset $A \subset X$ that is an ε-cover of X and an $\varepsilon/2$-packing.

Intuitively an ε-net is a set of points that is well spread out, yet is an ε-cover for X.

An ε-net is not a unique entity: there may be many such nets over the same space. Any ε-net is often merely an aspiration, possessing both coverage (reach) and packing (density) attributes.

Build an archive of trial points

- Add successive points to the archive, A, a subset of parameter metric space X

- Reward/prefer novelty

- Add new points to A provided they are far enough (> ε) away from A

Oxford Mathematics

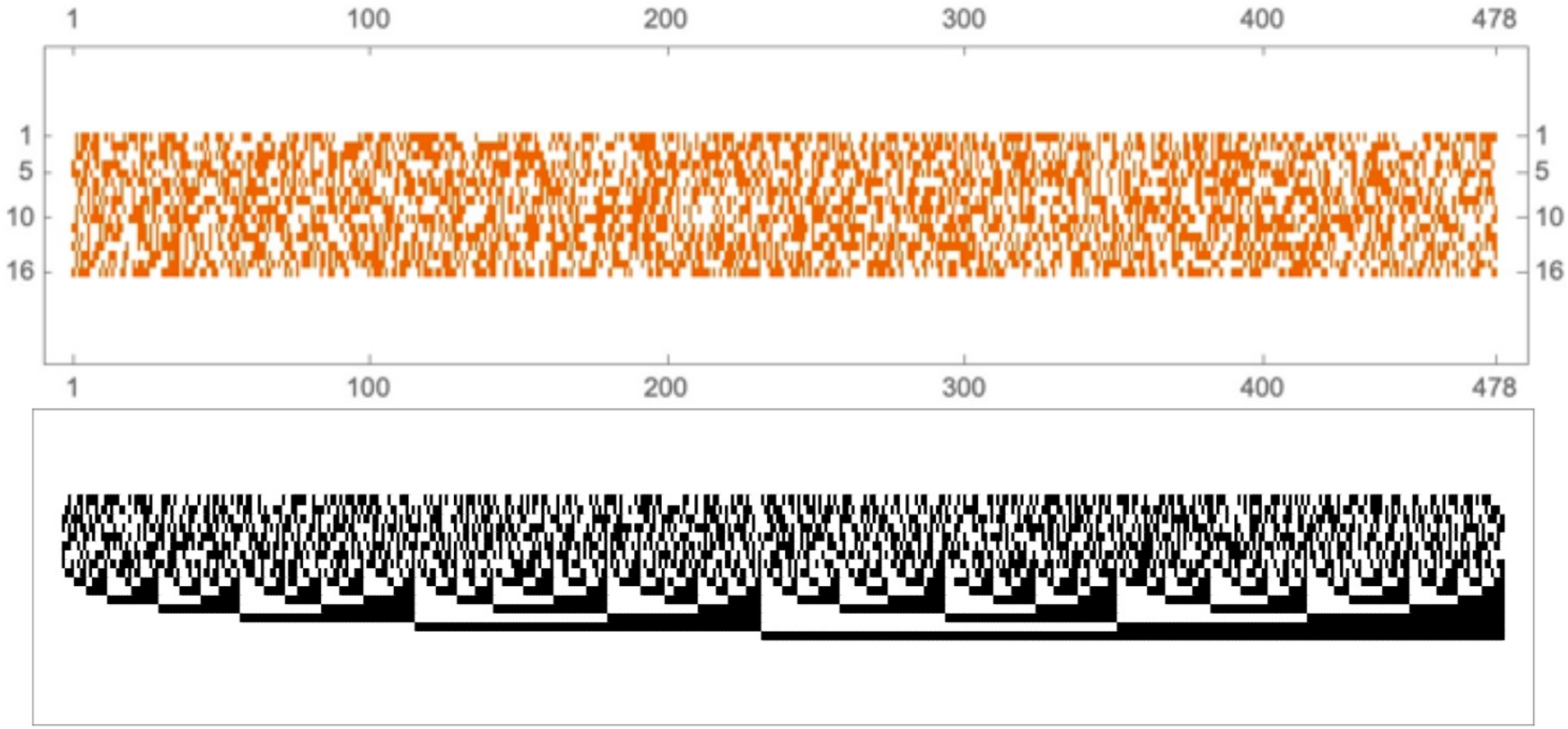


Figure 1: An ε-net for the $n = 16$ Hamming space with $\varepsilon = 4$, and 478 elements: each lies with a minimum pairwise distance of 4 from all others. We show (above, orange) each element vertically as the columns and the arrange the elements horizontally in the order in which they were added to the net. We show (below, black) the same elements this time ordered horizontally by their binary value.

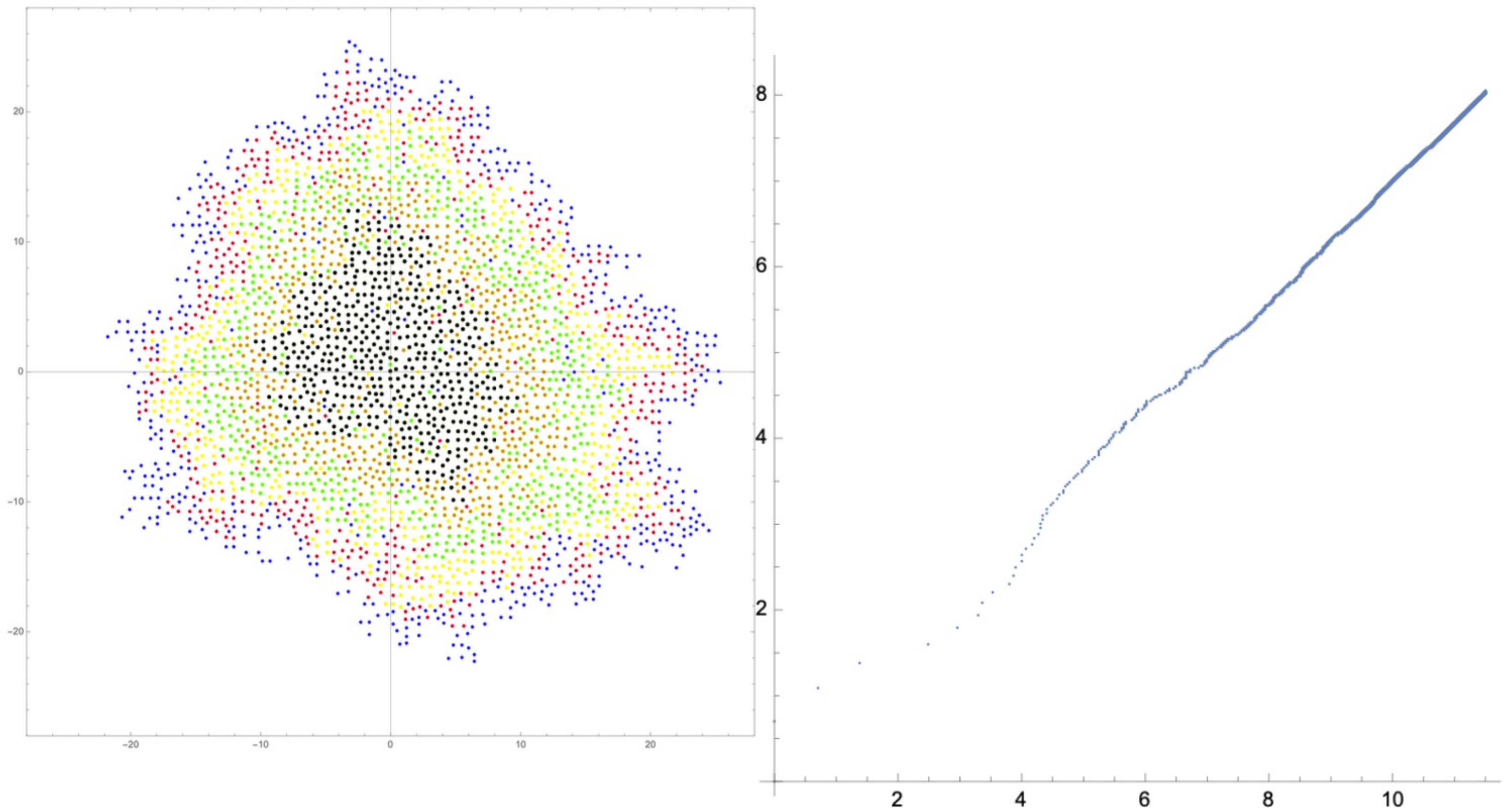


Figure 2: Left: a growing ε-net for $\mathbb{R}^2$ and $\varepsilon = 1/2$, with 500, 1000, 1500,...,3000 successive points. Right:$\log_e$ of the points added versus $\log_e$ of the candidates tested.

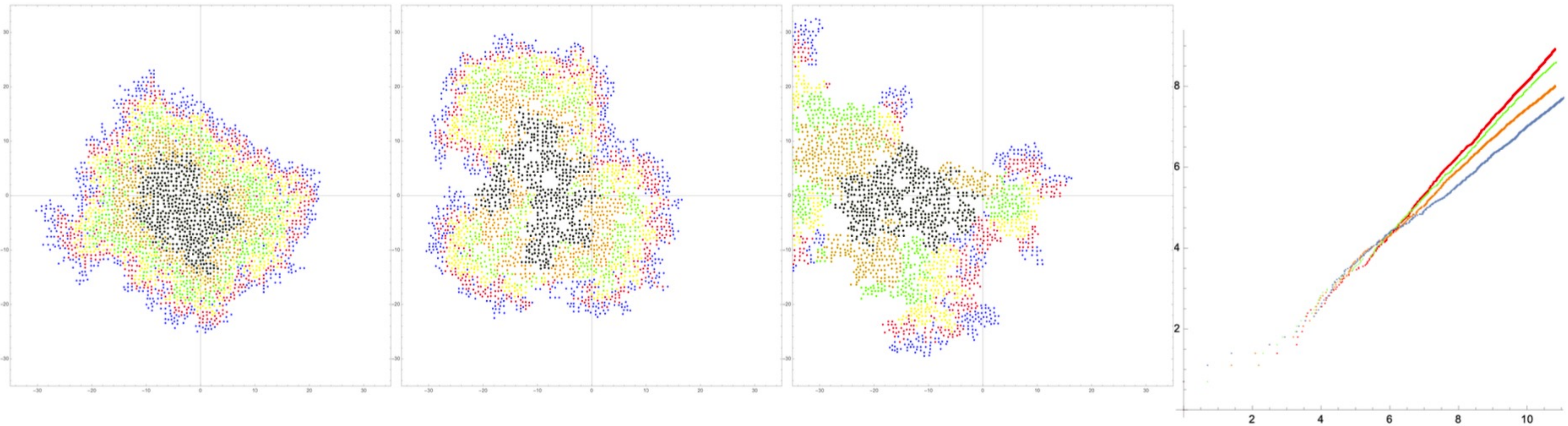


Figure 3: Growing ε-nets for $\mathbb{R}^2$ and $\varepsilon = 1/2$, with 500, 1000, 1500,...,3000 successive points. From the left: selecting parents from the most recent 50%, 20% and 10% of the archive. Right: $\log_e$ of the points added versus $\log_e$ of the candidates tested, selecting parents from the most recent 100% (blue), 50% (orange), 20% (green), and 10% (red) of the archive.

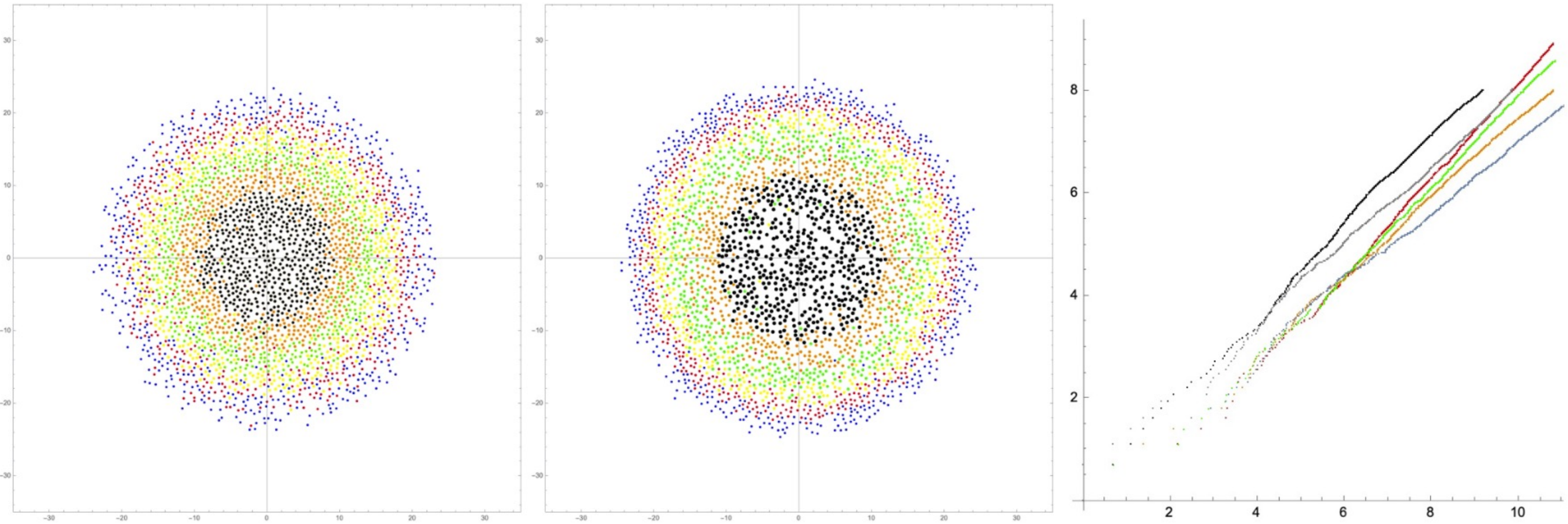


Figure 4: From the left: Growing ε-nets for $\mathbb{R}^2$ and $\varepsilon = 1/2$, with 500, 1000, 1500,...,3000 successive coloured points, by selecting and perturbing parents drawn from the most recent 50% and the most recent 20% of the archive together with an additional random rotations. Right: $\log_e$ of the points added versus $\log_e$ of the candidates tested: for sampling the most recent 50% (gray) and 20% (black), compared to the results from Figure 3 (selecting parents from the most recent 100% (blue), 50% (orange), 20% (green), and 10% (red) of the archive, with no additional rotations).

The volume a ball in $\mathbb{R}^n$ of radius R is is

$$V(n,R) = \frac{\pi^{n/2} R^n}{\Gamma(\frac{n}{2}+1)}.$$

Optimal packing densities, ρ_n, for hyper-spheres in n dimensions come into play here. These volume fractions are known [9] for $n = 1, .., 8$.

n	ρ_n
2	$\frac{\pi}{2\sqrt{3}} = 0.9069$
3	$\frac{\pi}{3\sqrt{2}} = 0.7405$
4	$\frac{\pi^2}{16} = 0.6169$
5	$\frac{\pi^2}{15\sqrt{2}} = 0.4653$
6	$\frac{\pi^3}{48\sqrt{3}} = 0.3729$
7	$\frac{\pi^3}{105} = 0.2953$
8	$\frac{\pi^4}{384} = 0.2537$

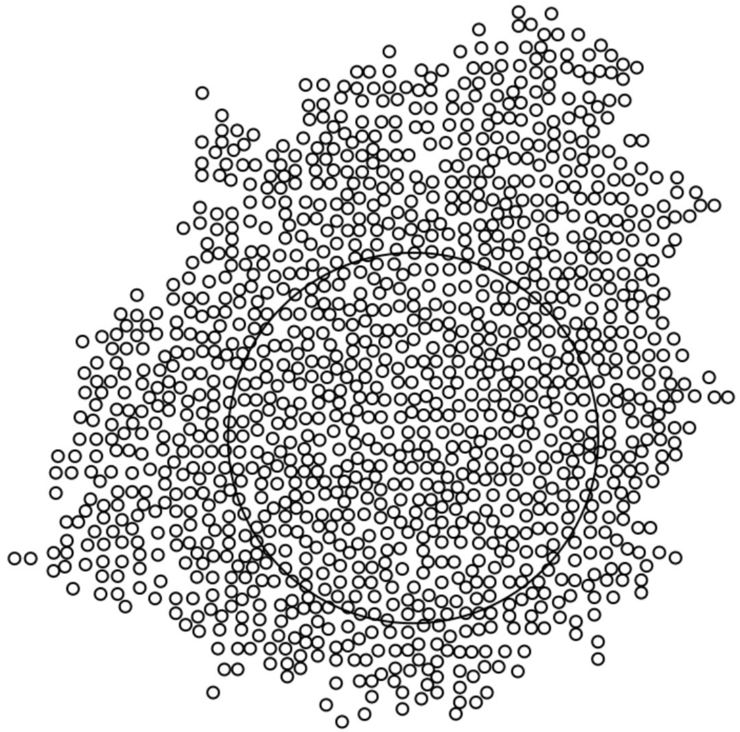


Figure 5: In $n = 2$ dimensions, 437 spheres of radius $\varepsilon/2 = 1/4$ are located inside a larger circular domain of radius 8. We estimate a packing density of ρ_a =0.43, compared to the optimal density of 0.74.

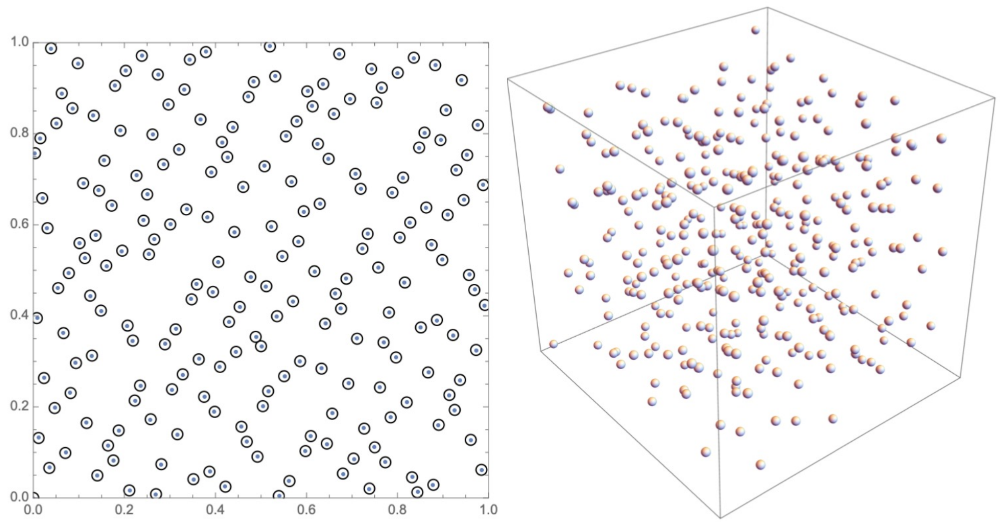


Figure 6: Left: A set of 200 Halton points in $\mathbb{R}^2$, generated with $p_1 = 2$ and $p_2 = 3$, with the largest balls, of radii= 0.01184, centred at each point, such that no two balls intersect. Right: A set of $7^3 = 343$ Halton points $\mathbb{R}^3$ generated with $p_1 = 2$, $p_2 = 3$ and $p_3 = 5$.

In many applications Lévy flight method are developed with controlled truncations of the Cauchy distribution. Note that these archives are formed from the superposition of many walks/flights, with Markov chains formed from successive (parent, successful candidate) pairs.

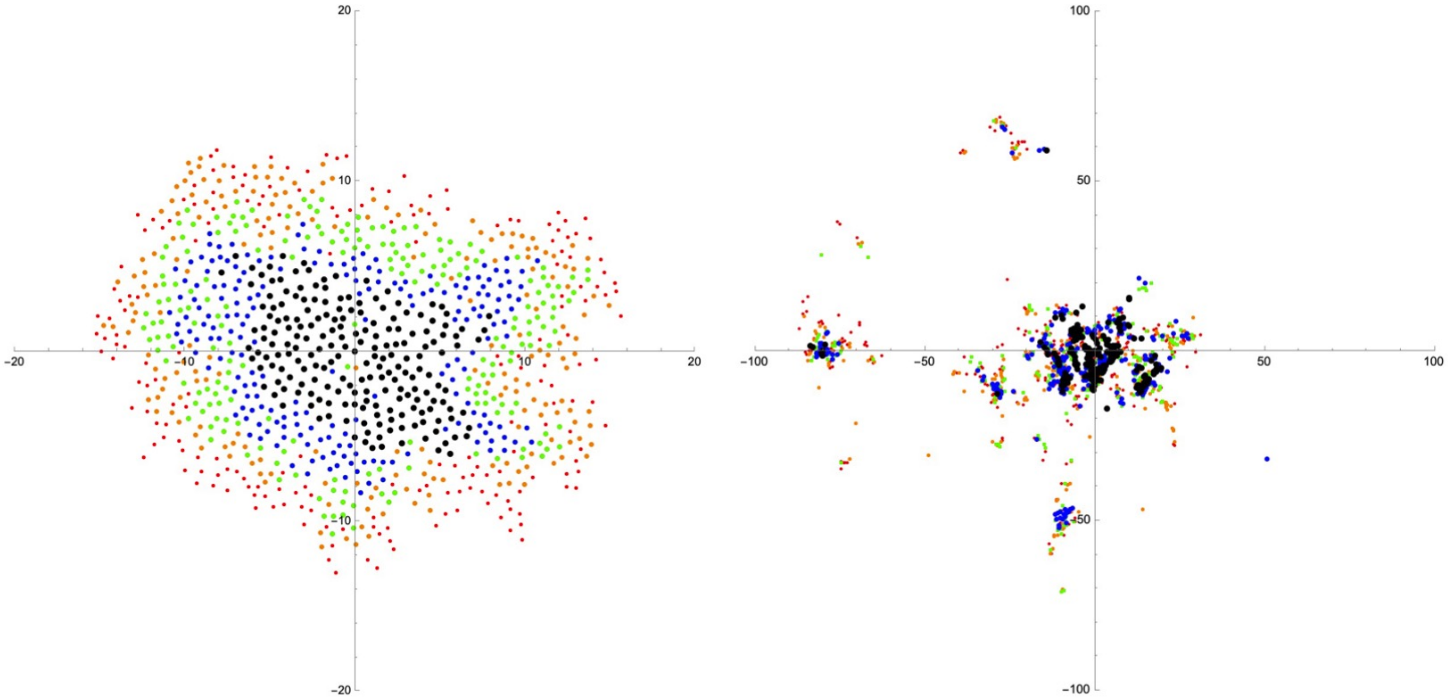


Figure 7: 1000 point archives developed with a Brownian process (left) and a Lévy flight process (right), with a minimum distance of $\varepsilon = 0.5$ from any existing archive member, and considering only the most recent 50% of the archive as parents. Points are coloured by their ordering: black points $\{1,200\}$, blue $\{201,400\}$, green $\{401,600\}$, orange $\{601,800\}$, red $\{801,1000\}$. Note the scale of the right hand plot (where a few far away points are not shown).

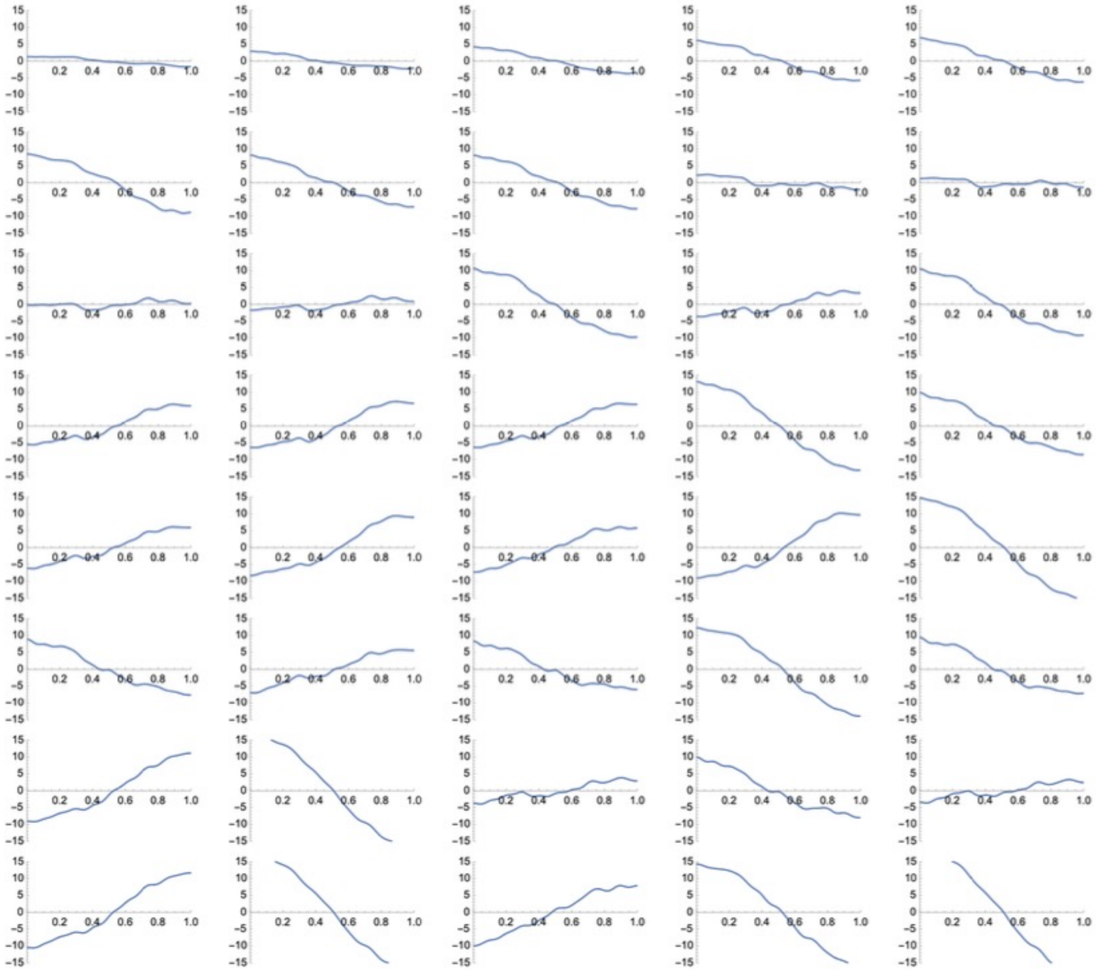


Figure 8: An archive of 160 successive functions (we show every 4th element added), defined on $\Omega = [0,1]$ satisfying zero Neumann boundary conditions, with each being a minimal separation of $\varepsilon = 0.5$ in the $L^2([0,1])$ norm from all previous elements.

Advice for Desperate Searchers

- If the parameter space is finite or else is compact within n-dim Euclidean space
  - The random archive packing is possible, yet may take computational effort to squeeze it out.
  - We recommend pseudo-lattice methods, which avoid rectilinear alignments yet also limit clustering, such as the use of Halton points for the archive generation.

- If the parameter space, is unbounded within n-dim Euclidean space where packing (the existence of unexplored voids) is very important, then any exhaustive is search impossible.
  - Archive methods using mutations equivalent to Brownian motion
  - Recency-bias in parent selection so the archive are grown efficiently
  - As the dimension becomes large though, diffusion may be far too slow.

- If the parameter space is unbounded in within n-dim Euclidean space where packing is far less important than achieving a wide reach (a large sampling variation),
  - We recommend Levy flights (occasional, very large, jumps).
  - This is explorative (and known to be effective within *foraging*).

- If the parameter space is an infinite dimensional (usually a function space)
  - packing is impossible (even for the unit ball: one may always add new elements equidistant from all existing members of an archive).
  - Choosing a suitable basis (compactly supported wavelets?)
  - The only relevant aim can be reach/exploration

Oxford
Mathematics

This project

- Devise ways in which to generate archives via random walks in a functional space (Levy flights, diffusion,…)
  - Take $L^2[\Omega]$ where Ω has dimension 1 (say [0,1]) or higher, say 2.
- Devise suitable performance measures to evaluate and exploit novelty

Oxford Mathematics