Small Worlds and Mega-Minds: Effects of Neighborhood Topology on Particle Swarm Performance

James Kennedy
Bureau of Labor Statistics, USA
Kennedy_Jim@bls.gov

Abstract This study manipulated the neighborhood topologies of particle swarms optimizing four test functions. Several social network structures were tested, with "small-world" randomization of a specified number of links. Sociometric structure and the small-world manipulation interacted with function to produce a significant effect on performance.

1 Introduction

The particle swarm algorithm is a method for optimizing hard numerical functions based on a metaphor of human social interaction (Eberhart, Simpson, and Dobbins, 1996; Kennedy, 1997). Individuals represented as multidimensional points interact with one another and converge on optimal regions of the problem space. While originally developed as a simulation of social behavior, the algorithm has become accepted as a computational intelligence technique related to evolutionary algorithms (Angeline, 1998; Eberhart, Simpson, and Dobbins, 1996).

Human interpersonal interaction, which has been theorized to contribute to cognitive processes (e.g., Brothers, 1997; Levine, Resnick, and Higgins, 1993), occurs in the context of a social structure, often depicted by social scientists as a network of connections between pairs of individuals. Research since the late 1940's (e.g., Bavelas, 1950) has shown that communication within a group, and ultimately the group's performance, is affected by the structure of the social network. Particle swarm research has examined several simple social structures, in particular the interaction of individuals with their immediate adjacent neighbors and fully connected interaction of all individuals in the population, but has not seriously examined effects of various "social structures" on optimization.

2 Small Worlds, Mega-Minds

In the 1960's, Milgram (1967) conducted an experiment to learn how interconnected people in the United States are. Individuals picked at random from the populations of two Midwestern cities were given a folder containing the name of someone in Cambridge, Massachusetts. The individuals were told to send the folder to someone they knew, who they thought might be likely to know the person in Cambridge; the person who received the folder then was to do the same, send it to someone they knew, and so on. The question was, how many links are required to connect two people selected at random? The surprising answer was that a median of only five connections were required to get the folder from a random person in Nebraska to a target in Massachusetts.

Granovetter (1973) theorized that "weak ties" are important in determining the optimality of social structures. According to his much-cited paper, information traveling through distant acquaintances is very important, largely because it provides information that a clique or ingroup might not possess. Strong ties or connections between pairs of individuals are likely to be found within cohesive groups, where interaction is frequent and norms are established and accepted, while weak ties form bridges between strongly-connected clusters. The weak ties then create paths between individuals who are not directly linked, allowing the spread of innovation through a population.

Recently, Watts and Strogetz (1998) examined the effects of changing a small number of randomly-chosen edges in a regular ring lattice. They demonstrated that changing fewer than 1 percent of edges in a network where every individual is connected to its K nearest neighbors (the "lbest" topology used in many particle swarm studies) resulted in a structure that featured high clustering, as found in a regular lattice, but with a greatly reduced average distance between nodes. Clustering is defined by the average number of neighbors that any two connected nodes have in common. In a ring lattice, clustering is high, but the average distance between nodes is also high. Randomizing a small proportion of connections maintains a high level of clustering but greatly reduces average distance. Watts and Strogetz (1998) call this the "small world" effect, in reference to Milgram's research; their paper contains an epidemiological simulation showing that "weak ties," or "small world" shortcuts in an otherwise regular network, have the effect of greatly hastening the spread of a disease through a population.

It is not hard to extrapolate from the disease model to a model of information flowing through a social network, for instance the spread of knowledge of some fact or rumor through a population. According to the small world hypothesis, random connections in an oth-

0-7803-5536-9/99/$10.00 ©1999 IEEE

Authorized licensed use limited to: Politecnico di Milano. Downloaded on January 16,2026 at 14:33:04 UTC from IEEE Xplore. Restrictions apply.

erwise orderly network should propitiate the spread of information throughout a population. For a particle swarm population, a manipulation affecting the topological distance between individuals might affect the rate and degree to which population members are attracted toward a particular solution.

In support of this hypothesis, Hutchins (1995) simulated a group of individuals represented as parallel constraint satisfaction networks. Selected nodes of each network were connected to the corresponding nodes of neighboring networks, and these were allowed to influence one another through the connections. The networks were optimized through the usual technique of asynchronous updating (Hopfield, 1984), allowing however for inputs from other nets. Hutchins found that the groups of networks converged on optima when there were a moderate number of connections among them, but converged on poor solutions when the cognitive structures were highly connected. He concluded that such mega-minds were especially susceptible to confirmation bias, which is a tendency to seek confirmation for hypotheses rather than using the proper logical method of falsifying. Networks that were highly connected in a "mass mental telepathy" (p. 252) quickly succumbed to the attraction of very poor local optima. Thus there appeared to be an optimal level of connectivity among individuals for optimizing the networks.

Latané's social impact theory (Latané, 1981; Nowak, Szamrej, and Latané, 1990) asserts that the probability that someone will adopt an attitude or belief is a function of the Strength, Immediacy, and Number of others who endorse the attitude (immediacy is the reciprocal of distance, while strength is a social attribute like status or persuasiveness). Social impact theory though is presented in terms of univariate social influence, that is, attitudes and beliefs are considered one at a time, independent of one another. Most interesting particle swarm problems are multivariate, for instance, optimizing the pattern of weights in a neural network, as it has long been presumed that cognitive variables should be optimized in relation to one another, i.e., cognitions should be consistent with one another (Abelson et al., 1968). The individual will search the nodes connected to it in the social network – its neighborhood – to determine which single member of the neighborhood has achieved the best performance so far. Thus the sociometric topology of the particle swarm population determines the breadth of influences on the individual, and how many neighbors the individual has in common with its neighbors – how small its world is.

In sum, we have reason to hypothesize that highly connected particle swarm societies might not be as good at finding optima in a problem space, compared to moderately connected social networks. It has been shown (e.g., Kennedy, 1997) that *isolated* particle swarm individuals perform very poorly: the interactions between particles make the algorithm work. What might be the best social structure for particles? The present study begins to investigate this question.

3 The Particle Swarm

The particle swarm algorithm is based on the metaphor of individuals refining their knowledge by interacting with one another. A particle is a moving point in a hyperspace. Besides its position $\vec{x}_i$ and velocity $\vec{v}_i$, each particle stores the best position in the search space it has found thus far in a vector $\vec{p}_i$. The velocity of the particle is adjusted stochastically toward its previous best position, and the best position found by any member of its neighborhood:

$$\vec{v}_i \leftarrow \vec{v}_i + \varphi_1(\vec{p}_i - \vec{x}_i) + \varphi_2(\vec{p}_g - \vec{x}_i)$$

where φ_1 and φ_2 are random numbers defined by their upper limit (usually 2.0). The index g is the index of the particle in the neighborhood with the best performance so far, so that p_g is the best vector found by any member of the neighborhood.

Once $\vec{v}_i$ has been calculated, the particle's position $\vec{x}_i$ is adjusted:

$$\vec{x}_i \leftarrow \vec{x}_i + \vec{v}_i$$

The algorithm is often compared to the family of evolutionary algorithms, as a stochastic population-based search of a problem space. Particle swarm differs from evolutionary methods in an important way, however: it does not implement selection of the fittest. Instead, individuals persist over time, and adapt by changing.

Particles have historically been studied in two general types of overlapping neighborhoods, called *gbest* and *lbest* (Eberhart, Simpson, and Dobbins, 1996). In the *gbest* neighborhood every individual is attracted to the best solution found by any member of the entire population. This structure then is equivalent to a fully connected social network; every individual is able to compare the performances of every other member of the population, imitating the very best. In the *lbest* network each individual is affected by the best performance of its K immediate neighbors in the topological population – a regular ring lattice. In the usual case, K=2, the individual is affected by only its immediately adjacent neighbors.

The choice of social structures used has been thus far a matter of individual artistry, with some lore and little data to help the researcher choose a strategy. The lore suggests that *gbest* populations tend to converge more rapidly than *lbest* populations, when they converge, but are also more susceptible to convergence on local optima.

Authorized licensed use limited to: Politecnico di Milano. Downloaded on January 16,2026 at 14:33:04 UTC from IEEE Xplore. Restrictions apply.

4 Neighborhood Types

In the trials reported below, populations of 20 individuals were configured into these configurations, shown in Figure 1:

- Circles (*lbest*): each individual is connected to its K immediate neighbors only
- Wheels: one individual is connected to all others, and they are connected to only that one
- Stars (*gbest*): every individual is connected to every other individual, and
- Random edges: for N particles, N random symmetrical connections are assigned between pairs of individuals.

In the Circle topology, which is a regular ring lattice as studied by Watts and Strogetz (1998), parts of the population that are distant from one another are also independent of one another. Thus one segment of the population might converge on a local optimum, while another segment converges on a different optimum or keeps searching. Influence spreads from neighbor to neighbor in this topology, until, if an optimum really is the best found by any part of the population, it will eventually pull all the particles in. Small-world reassignments of connections have the effect of shortening the distances between neighborhoods, and one would expect the population to converge faster – perhaps too fast – when the shortcuts are implemented.

The Wheel topology effectively isolates individuals from one another, as all information has to be communicated through the focal individual. This focal individual compares performances of all individuals in the neighborhood, and adjusts its trajectory toward the very best. If adjustments result in improvement in the focal individual's performance, then that improvement is communicated out to the rest of the population. Thus the focal individual serves as a kind of buffer or filter, slowing the speed of transmission of good solutions through the population. (It should be noted that the Wheel is a common configuration for many business and government organizations.)

Small-world shortcuts in the Wheel may have two effects. One is to create mini-neighborhoods, where peripheral individuals are influenced by individuals who are in direct contact with the focal or hub individual. Thus increased communication can result from implementing shortcuts, and we might again expect faster convergence in the collaborating subpopulation. The buffering effect of the focal particle though should prevent overly rapid convergence on local optima. It is also possible to create islands, or disconnected groups of individuals, which may collaborate among themselves to optimize the function. This would introduce a diminishing of communication, as the isolated individuals would not have access to information about the best region found by the population; nor would the rest of the population benefit from their successes.

The Star or Gbest topology links every individual with every other, so that the social source of influence is in fact the best-performing member of the population. Finally, the Random topology simply assigns connections at random between pairs of particles.

Populations were tested on a set of well-studied test functions covering a range of problem types. Circles were defined with K=2, and Circles and Wheels were studied with several degrees of small-world shortcuts; shortcuts are meaningless in either Random or fully connected Star networks.

These trials implemented a modified particle swarm version using a constriction coefficient proposed by Maurice Clerc (Clerc and Kennedy, 1999, under review). This version is simple to implement and has the advantage that the population converges without requiring a V_{max} limit to velocities. A constant coefficient is calculated using the upper limit of $\varphi = \varphi_1 + \varphi_2$:

$$\chi = 1 - \frac{1}{\varphi} + \frac{\sqrt{|\varphi^2 - 4\varphi|}}{2}$$

The particle swarm formula is then modified:

$$\vec{v}_i \leftarrow \chi(\vec{v}_i + \varphi_1(\vec{p}_i - \vec{x}_i) + \varphi_2(\vec{p}_g - \vec{x}_i))$$

This modification, used with values of $\varphi > 4.0$, has been shown to result in excellent optimization of test functions. Since it has desirable convergence properties and removes the necessity of imposing the arbitrary velocity limit, it was used in the present implementations.

5 Method

The particle swarm program was modified to allow control over sociometry. Standard test functions were taken from the literature of evolutionary computation, including De Jong's f1 sphere function, Griewank, Rastrigin, and Rosenbrock functions. All functions were implemented in 30 dimensions. Each function was run with wheel and circle sociometries, with 0, 1, 2, 3, 4, and 5 random small-world shortcuts. Thus the experiment had three independent variables: function (4 levels: Func), basic neighborhood topology (2 levels: Ntype), and shortcuts (6 levels: Nmoved). The dependent variable used was population-best performance on the test function after 1,000 iterations. Each Func×Ntype×Nmoved condition of the experiment was run twenty times.

Two kinds of control groups were run. A randomly connected group was run for each function, as was a *gbest* or fully-connected topology. Because these groups were not orthogonal to the experimental design, they were analyzed separately.

Authorized licensed use limited to: Politecnico di Milano. Downloaded on January 16,2026 at 14:33:04 UTC from IEEE Xplore. Restrictions apply.

All trials used populations of 20 individuals, with φ=4.1. Functions tested are shown in Table 2.

Data were analyzed using analysis of variance (ANOVA). The output of ANOVA is the F statistic. F is the between-group variance (the average difference between cell means) divided by the within-group variance, which is taken to be an estimate of error. A "p-value" indicates the probability of deciding that the null hypothesis of no difference is false when it is actually true. Traditionally a p-value less than 0.05 is considered significant.

6 Results

Because the various functions were scaled differently, resulting in incomparable means and variances, within-function data were standardized to a scale with mean=0 and standard deviation=1, and factorial ANOVA was conducted on these transformed data. As mentioned above, in order to preserve the orthogonality of the experimental design, Random and Gbest conditions were dropped from the analysis. No main effect is reported for Function since scores were standardized to that domain, thus all means are 0.0. Analytic results are shown in Table 1; means and standard deviations are in Appendix 1.

As can be seen, there was no significant main effect for Neighborhood Type, though it interacted significantly with Function and with Number Moved.

The interaction of Neighborhood Type with Function was very strong, as it was seen that populations performed better on three of the functions when they were in the Circle configuration, regardless of the number of edges moved, than in the Wheel configuration. Performance on the Rastrigin function, however, was just the opposite; Rastrigin populations performed better in the Wheel topology.

Table 1. Analysis of variance of 3 factors with interactions. NType=Neighborhood Type, Func=Function, Nmoved=Number of Edges Moved.

Source	DF	F	p
Ntype	1	1.04	0.3090
Func×Ntype	3	13.91	0.0001
Nmoved	5	1.99	0.0780
Func×Nmoved	15	0.43	0.9694
Ntype×Nmoved	5	2.78	0.0168
Func×Ntype×Nmoved	15	0.42	0.9724

The interaction between Neighborhood Type and Number Moved is perhaps anomalous. It was seen that the Sphere-Circle, Sphere-Wheel, Griewank-Wheel, and Rosenbrock-Wheel conditions all had unusually poor performance when four edges were moved. This condition does not interpolate between the results for three and five edges moved; further, the variance was unusually high where Number Moved=4 and topology=Wheel, indicating inconsistent performance. In the Wheel topology, as mentioned above, shortcuts can result in alliances and formation of collaborating subpopulations, or in isolated islands, cut off from the group's progress.

This same effect seems to account for the nearly-significant main effect of Number Moved, $p<0.0780$. The mean performance for Number Moved=4 was worse than that for other values.

A second analysis was performed in order to determine how the Gbest (Star: fully connected topology) and Random (Randomly connected topology) conditions compared to the rest. Scores were standardized per function as before, but including Gbest and Random conditions along with the orthogonal ones. T-tests were conducted, comparing the orthogonal data with Gbest only, and then comparing the orthogonal data with Random only. Interestingly, both Gbest and Random performed better than the other conditions combined, $p<0.05$. Inspection of means showed that Gbest topology was better than the combined mean on every function, while the Random topology outperformed others on all functions except Rastrigin.

7 Discussion

Watts and Strogetz' mathematical model showed the effects of randomizing a very small proportion of connections. In the present particle swarm implementations with populations of 20 individuals it was impossible to move 1% or fewer of the links; smaller population sizes however are usually appropriate with the particle swarm method. Therefore this may not be considered a valid test of the small-world hypothesis, but rather simply an investigation into the effect. On the other hand, these findings do show that the sociometry of the particle swarm has a significant effect on its ability to find optima: the optimal pattern of connectivity among individuals depends on the problem being solved.

This study did not systematically manipulate aspects of test functions, but there are grounds for speculation as to an explanation for the interaction. As seen in Appendix 2, the Sphere and Rosenbrock functions are unimodal, with smooth surfaces. The Griewank function depicted in two dimensions looks like a bumpy bowl that slopes gradually toward the global optimum textured with many slight local optima. The Rastrigin function though features many very steep local optima, with the depth of minima gradually increasing toward the global optimum. Failure occurs when the population converges on a local optimum and is unable to leap from that hill to a better one. One explanation for the current results regarding differences in the effects of sociometric structures on

Authorized licensed use limited to: Politecnico di Milano. Downloaded on January 16,2026 at 14:33:04 UTC from IEEE Xplore. Restrictions apply.

different functions – Wheels performed better than Circles on Rastrigin only – might be that the buffering effect of communicating through a "hub" slows the population's attraction toward the population best, preventing premature convergence on local optima. Thus a hypothesis can be proposed for later research: centralized Wheel topologies may perform better on strongly multimodal landscapes.

Some effects may be due to the convergence enforced by the constriction coefficient. Interestingly, one study with the V_{max}-type particle swarm reported much poorer results on three of these functions – but the "standard" particle swarm performed better on the Rastrigin function (Angeline, 1998). This may suggest that constriction is disadvantageous on problems with many good local optima. On the other hand, that study implemented slightly different forms of these functions, including initialization ranges, and informal testing with the current programs found that the constricted version performed much better than the standard form on all functions.

The second finding, that changing four links on Wheel structures – except on the Rastrigin function – greatly impairs performance, is going to be harder to explain. When three links are randomized, or five links, performance is not especially affected, but four shortcuts result in degraded ability to optimize these test functions. In the Wheel topology, where all individuals are connected to one focal individual, shortcuts can be seen to create subpopulations with individuals that are relatively isolated from the focus. The finding that Random connections performed relatively well – less so on Rastrigin – makes this effect even harder to understand.

In sum, it has been shown that neighborhood topology significantly affects the performance of a particle swarm, and that the effect is dependent on the function.

Bibliography

Abelson, R.P, E. Aronson, W.J. McGuire, T.M. Newcomb, M.J. Rosenberg, & P.H. Tannenbaum (Eds.) (1968). *Theories of Cognitive Consistency: A Sourcebook.* Chicago: Rand-McNally.

Angeline, P. J. (1998). Evolutionary optimization versus particle swarm optimization: Philosophy and performance differences. *Evolutionary Programming VII*, 601-610. Berlin: Springer.

Bavelas, A. (1950). Communication patterns in task-oriented groups. *Journal of the Acoustical Society of America, 22*, 727-730.

Brothers, L. (1997). *Friday's Footprints: How Society Shapes the Human Mind.* Oxford: Oxford University Press.

Clerc, M., and Kennedy, J. (1999, under review). The Particle Swarm: Explosion, Stability, and Convergence in a Multi-Dimensional Complex Space.

Eberhart, R. C., Simpson, P., and Dobbins, R. (1996). *Computational Intelligence PC Tools.* Boston: Academic Press.

Granovetter, M. S. (1973). "The Strength of Weak Ties." *American Journal of Sociology, 78*: 1360-1380.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences, 81*, 3088-3092.

Hutchins, E. (1995). *Cognition in the Wild.* Cambridge, MA: MIT Press.

Kennedy, J. (1997). The particle swarm: Social adaptation of knowledge. *Proceedings of the 1997 International Conference on Evolutionary Computation* (Indianapolis, Indiana), 303-308. IEEE Service Center, Piscataway, NJ.

Latané, B. (1981). The psychology of social impact. *American Psychologist, 36*, 343-356.

Levine, J. M., Resnick, L. B., and Higgins, E. T. (1993). Social foundations of cognition. *Annual Review of Psychology, 44*, 585-612.

Milgram, S. (1967). The small world problem. *Psychology Today, 22*, 61-67.

Nowak, A., Szamrej, J., and Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review, 97*, 362-376.

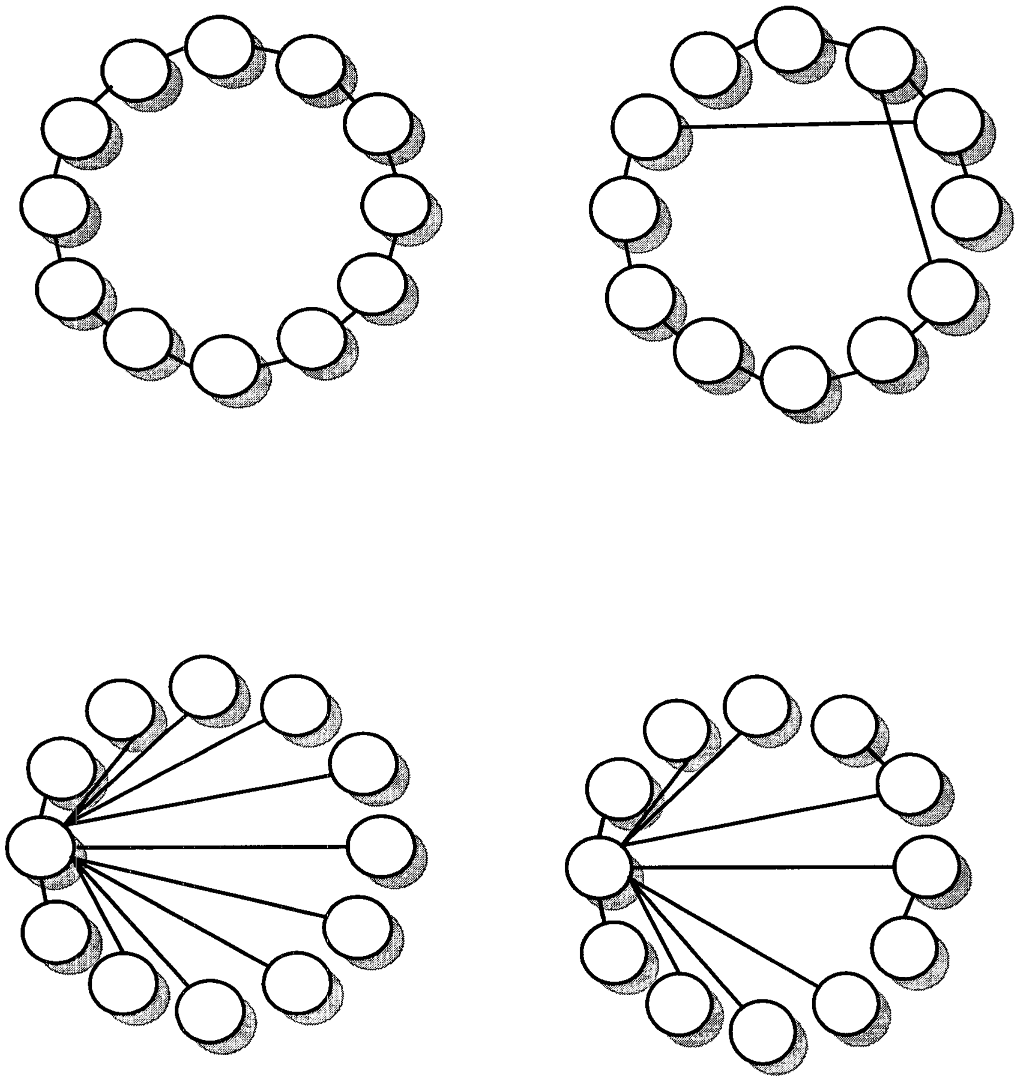
Watts and Strogetz (1998). Collective dynamics of 'small-world' networks. *Nature, 393*, 440-442.

Table 2. Functions tested in particle swarm trials.

Sphere	$f_1(x) = \sum_{i=1}^{n} x_i^2$
Griewank	$f_7(x) = \frac{1}{4000}\sum_{i=1}^{n}(x_i-100)^2 - \prod_{i=1}^{n}\cos\left((x_i-100)/\sqrt{i}\right)+1$
Rosenbrock	$f_9(x) = \sum_{i=1}^{n}(100 + (x_{i+1}-x^2)^2 + (x_i-1)^2)$
Rastrigin	$f_{10}(x) = \sum_{i=1}^{n}\left[x_i^2 - 10\cos(2\pi x_i) + 10\right]$

Authorized licensed use limited to: Politecnico di Milano. Downloaded on January 16,2026 at 14:33:04 UTC from IEEE Xplore. Restrictions apply.

Figure 1. Sociometric topologies of populations of size 12. Top left is circle, K=2, top right is the same neighborhood with two "small-world" shortcuts. On the lower left is a "wheel" configuration, and lower right is the wheel with two small-world changes.



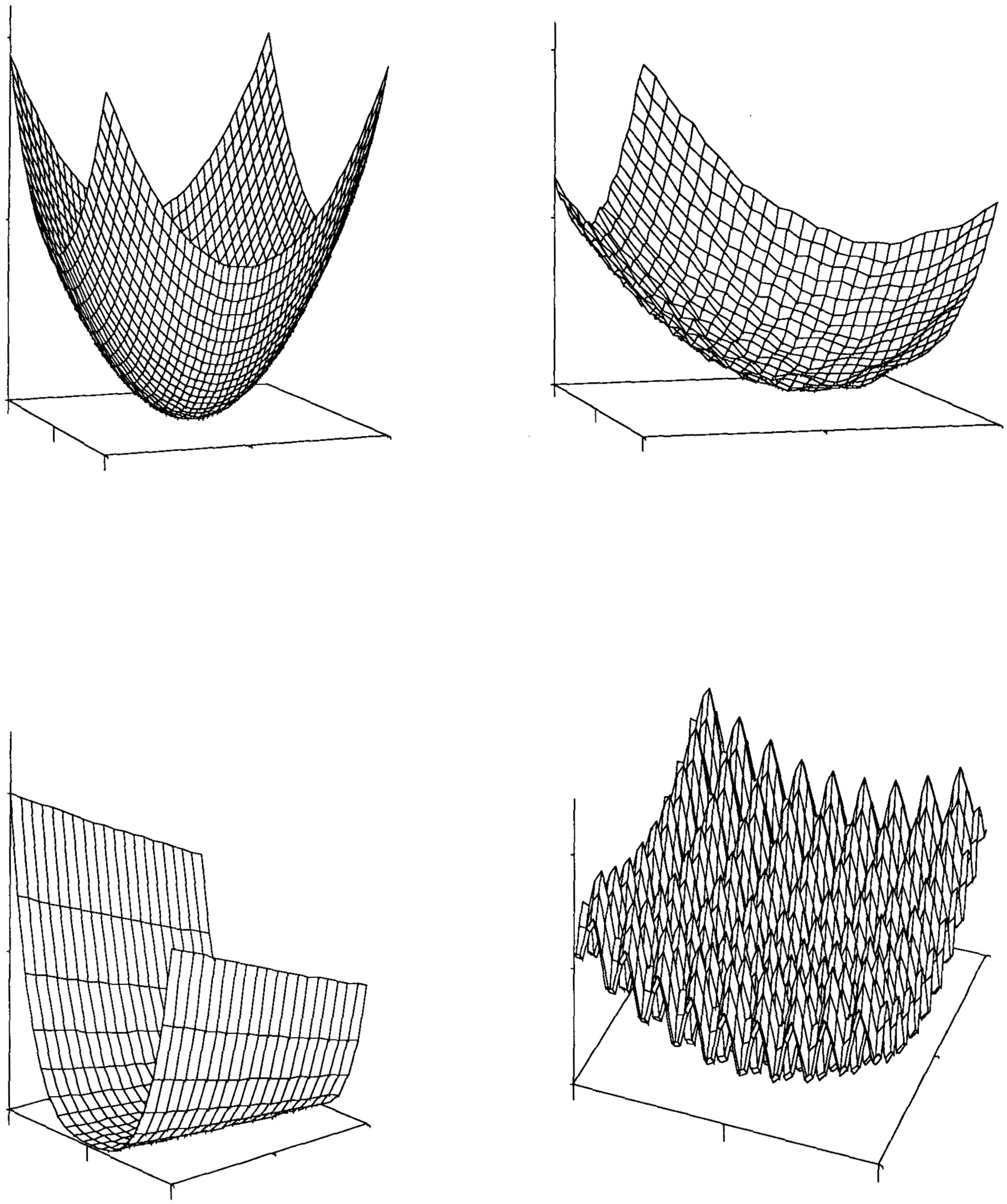
Authorized licensed use limited to: Politecnico di Milano. Downloaded on January 16,2026 at 14:33:04 UTC from IEEE Xplore. Restrictions apply.

Appendix.1. Cell means by condition.

Function	Sociometry	Shortcuts	Mean	S.D.
Sphere	**Circle**	0	0.0002015	0.0001612
		1	0.0003009	0.0001795
		2	0.0002270	0.0001849
		3	0.0002513	0.0002310
		4	0.0003786	0.0004283
		5	0.0002789	0.0004146
	Wheel	0	0.0074977	0.0131162
		1	0.0018735	0.0014357
		2	0.0081846	0.0154843
		3	0.0010437	0.0011029
		4	0.0202361	0.0879435
		5	0.0005530	0.0010135
	Star		0.0000002	0.0000005
	Random		0.0001740	0.0002597
Griewank	**Circle**	0	0.0147821	0.0153143
		1	0.0229308	0.0223021
		2	0.0217059	0.0337333
		3	0.0266354	0.0184261
		4	0.0147307	0.0165112
		5	0.0138850	0.0105729
	Wheel	0	0.1093880	0.1809570
		1	0.0710453	0.1046179
		2	0.0419942	0.0455719
		3	0.0214299	0.0202516
		4	0.6564481	2.6909252
		5	0.0474588	0.0626935
	Star		0.0285562	0.0626092
	Random		0.0162451	0.0224282
Rastrigin	**Circle**	0	82.361307	16.8378058
		1	85.5237802	19.1645739
		2	88.4453155	19.6775784
		3	79.9194073	17.3228016
		4	81.0150061	20.4310513
		5	82.6445681	18.1323865
	Wheel	0	70.2200408	19.9224800
		1	71.4445232	23.2225944
		2	65.1677276	19.7122599
		3	70.0557020	17.1484400
		4	70.9911407	33.4759397
		5	71.6686468	19.4681528
	Star		65.0206740	17.6340843
	Random		76.7315666	19.2301384
Rosenbrock	**Circle**	0	90.3720913	42.6545187
		1	78.8256950	45.3185693
		2	86.3083155	47.1847008
		3	78.1767920	43.9249742
		4	74.6231132	50.2696593
		5	94.6026954	102.520437
	Wheel	0	117.297891	59.6340843
		1	92.6786645	55.7016215
		2	127.445158	74.4664749
		3	96.1995437	42.3662924
		4	174.259688	175.615940
		5	120.005148	91.1367870
	Star		91.6434190	46.7294972
	Random		92.3906794	55.2194654

Authorized licensed use limited to: Politecnico di Milano. Downloaded on January 16,2026 at 14:33:04 UTC from IEEE Xplore. Restrictions apply.

Appendix 2. Functions tested. Top left: sphere function. Top right: Griewank. Bottom left: Rosenbrock. Bottom right: Rastrigin. All functions are plotted for their full range as reported in the text.



Authorized licensed use limited to: Politecnico di Milano. Downloaded on January 16,2026 at 14:33:04 UTC from IEEE Xplore. Restrictions apply.