**Notes and Observations**

**On**

**Simulated Annealing & Cooling Schedules**

*Linear cooling* performs roughly equivalent to gradient descent. As we would expect, a linearly decreasing cooling schedule has a higher probability of accepting worse solutions even until the temperature gets quite low. As we can see, this pushes towards exploration ***over*** exploitation and as a result gives us a very inaccurate approximation to the global optima with as noisy of a surface as the Rastrigin function generates.

*Logarithmic multiplicative cooling* performs poorly, but better than most, with an acceptable distance from the optima of 1 unit away we see a success rate of only ~17%

S.A. exploration is mostly closely related to the time with which we spend at higher temperatures while exploitation is the time spent at lower temps (time = iterations)

G.D. – The fact that the surface itself has fairly uniform gradients across the entire search space leaves us with the conclusion that (without momentum) the probability of finding a global optima is directly proportional to the probability of our random initial point falling somewhere within the peaks surrounding (x,y,z)=(0,0,0)

# iteration control and terminal condition

i = 0

ppts = 0

stagnation = 0

stagnantMax = 5

itersMax = 1000