DD2437
Group name: Vairamuthu_Guan
rvai@kth.se
xguan@kth.se
miloulis@kth.se

Lab assignment 3

Section 3.1 Convergence and attractors

1) All the patterns (x1, x2 x3) converged towards the stored pattern.
2) There were 14 attractors on the network.
3) If the starting pattern is more than half dissimilar, it converges to some other attractor instead of the desired. It also takes longer to converge.

Section 3.2 Sequential Update

1) The 3 patterns are stable. points
2) The network is able to complete a degraded pattern.
    a) The network successfully completed the degrade pattern p10 to p1.
    b) Same is not the case with p2 and p3. In this case the network recollected the image p11 to p2 but not p3. We updated the units sequentially in order.
3) If the units are selected randomly, it takes a longer time to converge and the image slowly converges to one of the attractors

Section 3.3 Energy

1) The energy at the attractors is significantly lower than the other points. Here are the energies for a (order 10^3)
    a) -1.4394
    b) -1.3656
    c) -1.4623
    d) -0.7205
    e) -0.5259
    f) -0.6833
    g) -0.6857
    h) -0.1715
    i) -0.2675
2) The energy decreases every iteration as the point moves closer towards a single attractor.
3) When the generated weight matrix is random, the energy keeps oscillating and does not converge.
4) If the weight matrix is made symmetric, on the other hand, convergence happens. This is because making the weight matrix symmetric creates an unidirectional path to the attractors. Whereas a non symmetric weight matix may result in a loop of switching between different energy states without any proper descent.

Section 3.4 Distortion Resistance

DD2437
Group name: Vairamuthu_Guan
rvai@kth.se
xguan@kth.se
miloulis@kth.se

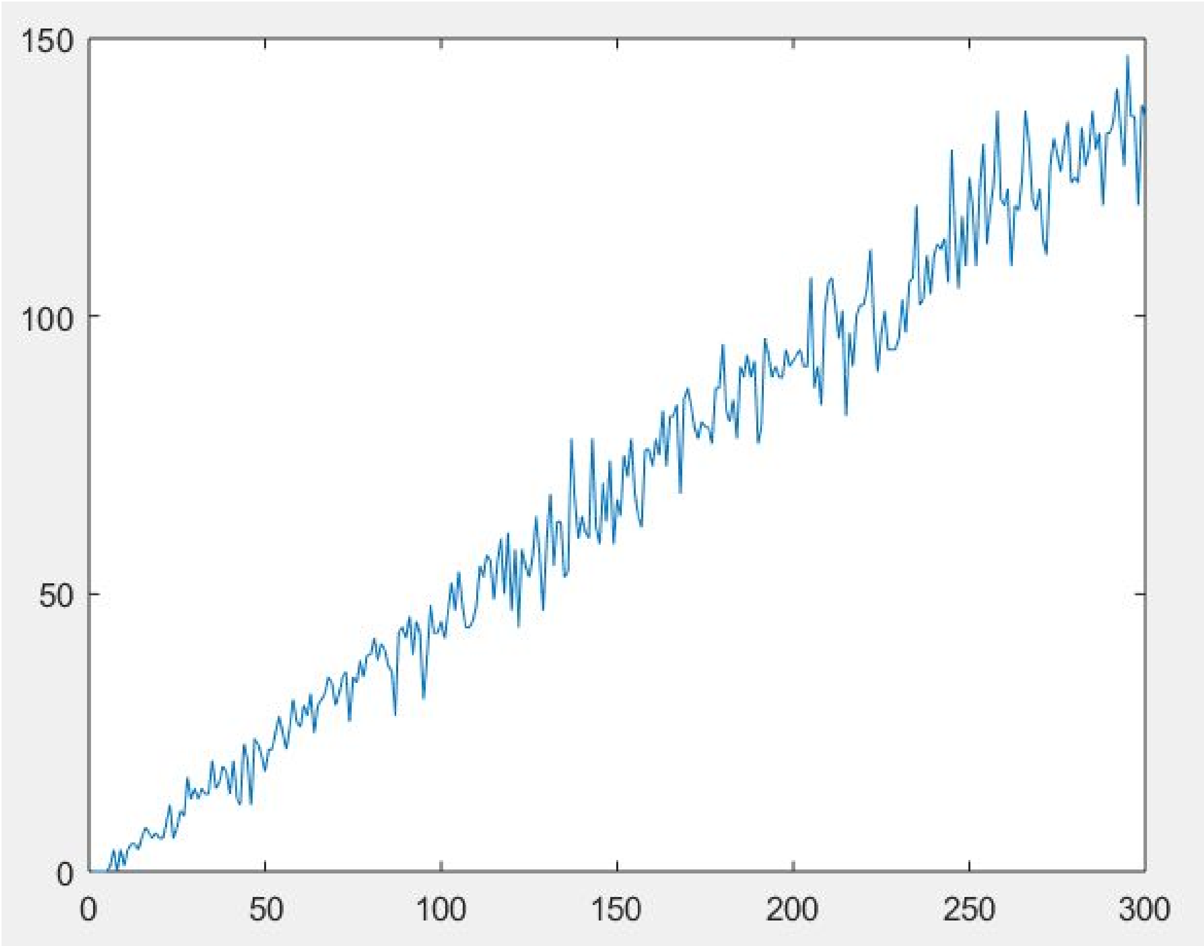
1) Close to around 20-25% noise could be removed, Around 30-40% we can see that there are some incorrect convergence to the wrong patterns. And greater than 50% it fails several times.
2) Yes, p1 and p2 always performed much better than p3. This is probably because there are any other attractors that lie closer to p3 compared to p1 and p2. So it falls to other attractors instead of p3. Also, p3 is a bit biased.
3) Sometimes in the single step recall, if the first iteration isn't successful, the subsequent iteration of the hopfield network helps in converging to the right attractor. But we did not encounter any example where it took more than two iterations at the most.

Section 3.5 Capacity

1) The network is able to safely store 2 patterns safely. There is not much error with the 3rd pattern but with 4 patterns, the network performance drastically reduces. We did our simulations for error ratio = 0.2 and here are the results
    a) incorrectMatchesPercent = [0, 0]
    b) incorrectMatchesPercent = [0, 0, 8.62]
    c) incorrectMatchesPercent = [0, 0, 100,100]
    d) incorrectMatchesPercent = [0, 0, 100, 100, 100]
2) With random patterns the network was able to learn much more
    a) For noise of 0.2, the network was able to recollect all 9 trained images with 100% accuracy
    b) For noise of 0.4, the network was able to recollect the 9 trained images equally with almost 80% accuracy.
3) This behaviour is probably due to the location of these images. If two images are very similar to each other, it means that they lie close to each other in space domain and the tested noisy input may converge to either one of these. Whereas when learning a truly random pattern, it is possible that they are always quite far apart in space, so most of the convergence is unique and accurate. This constitutes a bias in the pictures whereas it is not the case with random patterns.

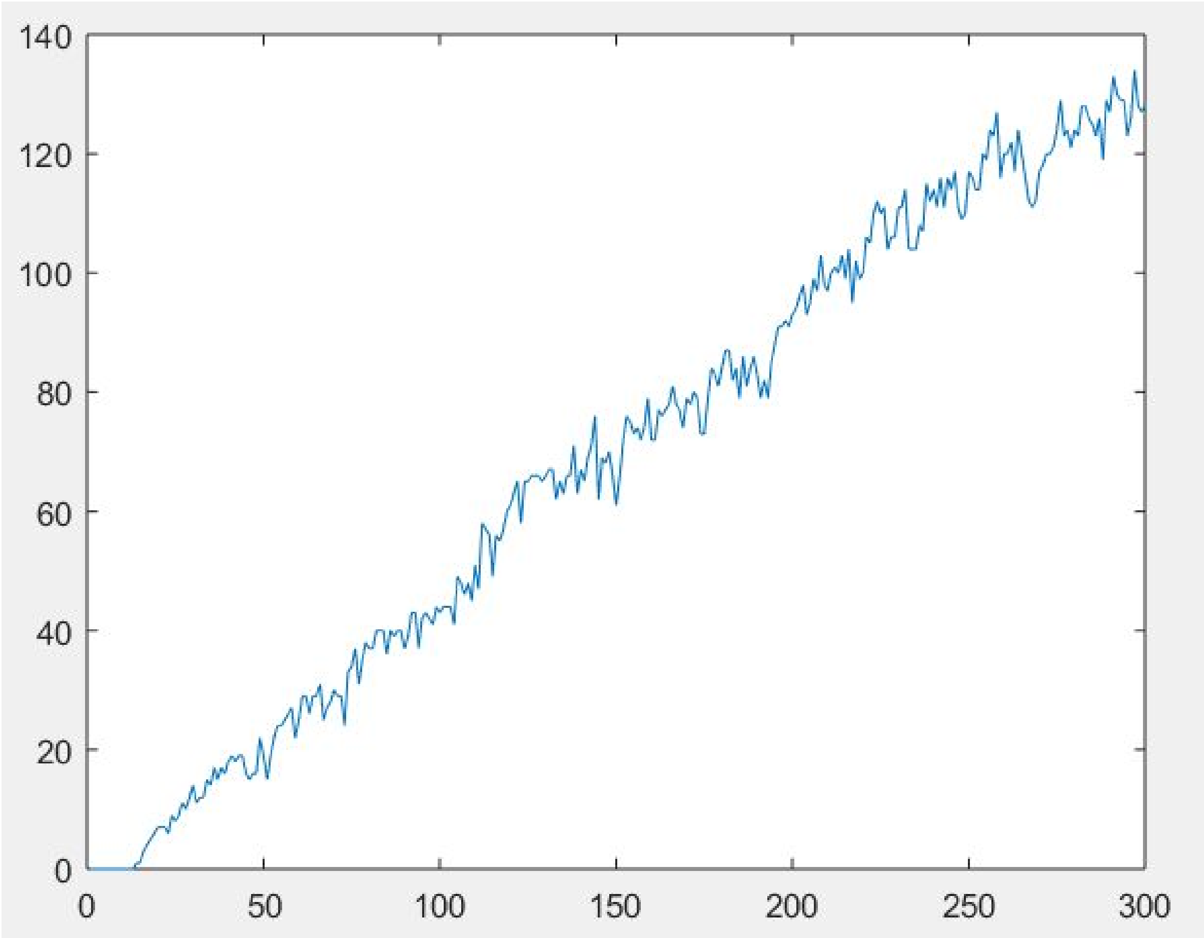
DD2437
Group name: Vairamuthu_Guan
rvai@kth.se
xguan@kth.se
miloulis@kth.se

4) After learning about 12-13 patterns the network struggles to learn more patterns.
5) With the noise added before testing the patterns, the following is observed.

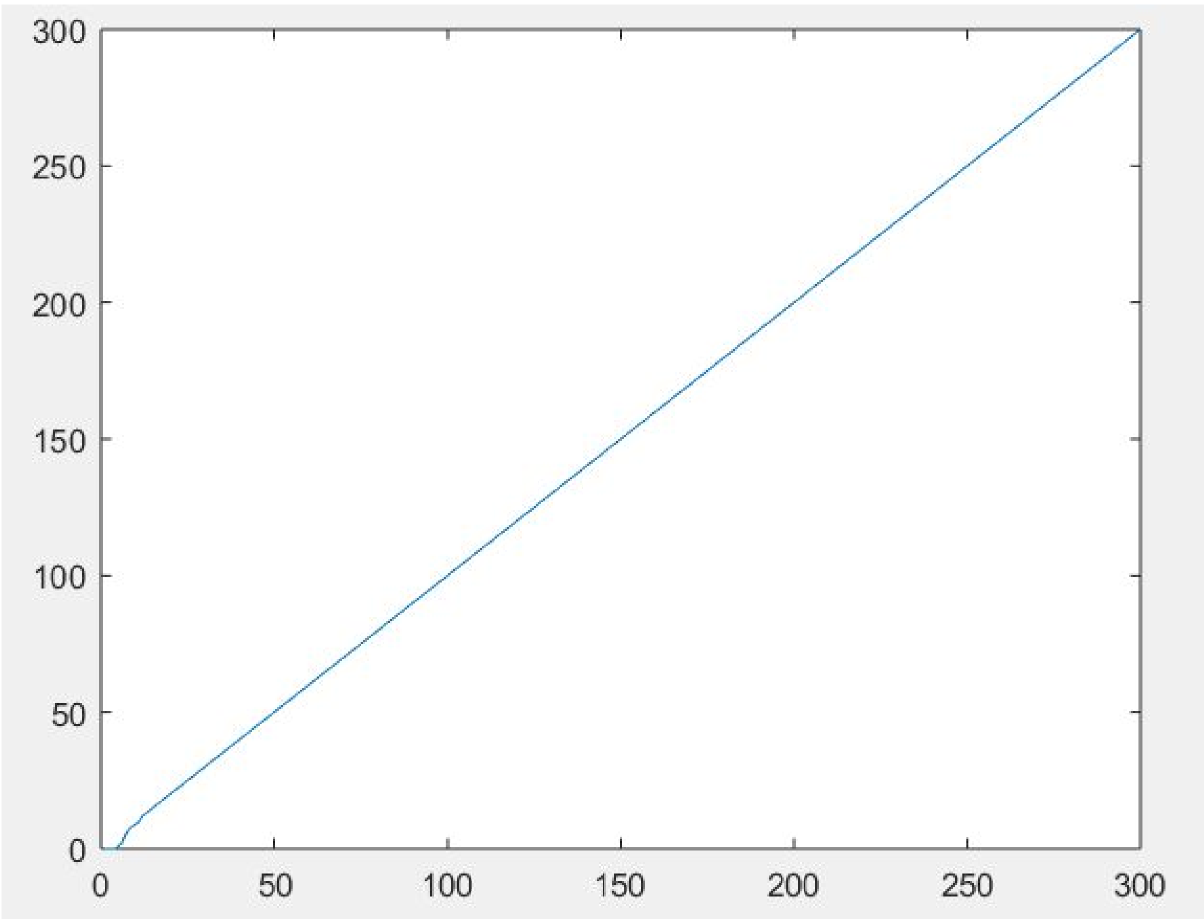


6) It does not change much when the diagonal are set to zero. This helps avoid the spurious patterns.

DD2437
Group name: Vairamuthu_Guan
rvai@kth.se
xguan@kth.se
miloulis@kth.se



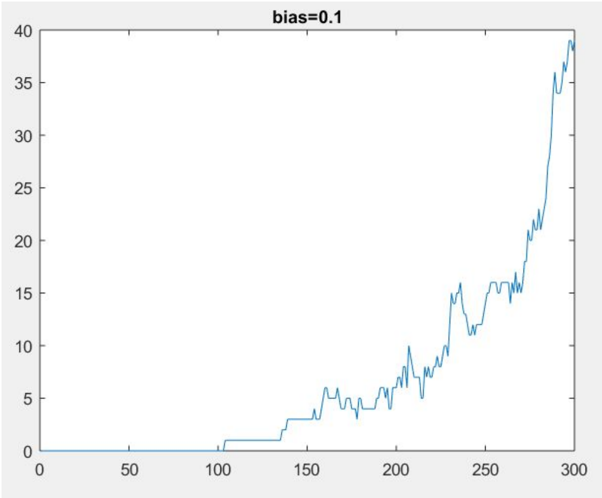
7) The graph becomes very linear in this case. And the number of patterns learnt reduces to just 4.



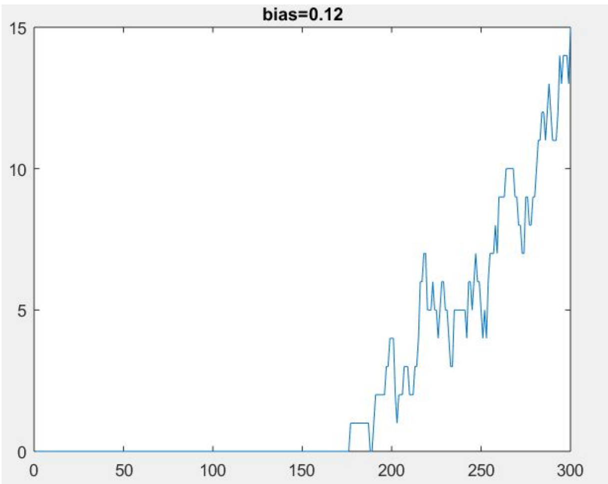
DD2437
Group name: Vairamuthu_Guan
rvai@kth.se
xguan@kth.se
miloulis@kth.se

Section 3.6 Sparse Patterns

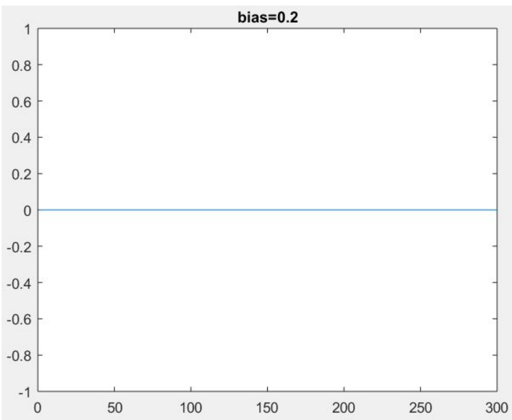
1) Sparsity = 0.1
    a) For bias = 0.1



    b) For bias = 0.12

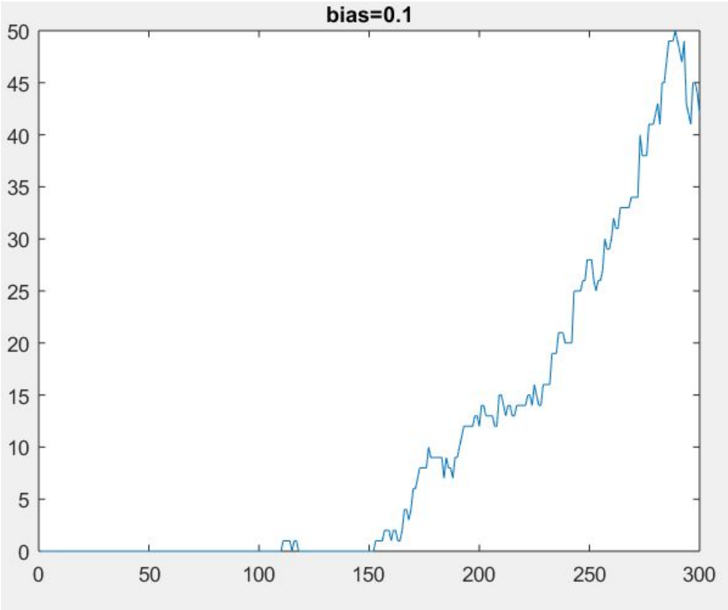


    c) For bias = 0.2

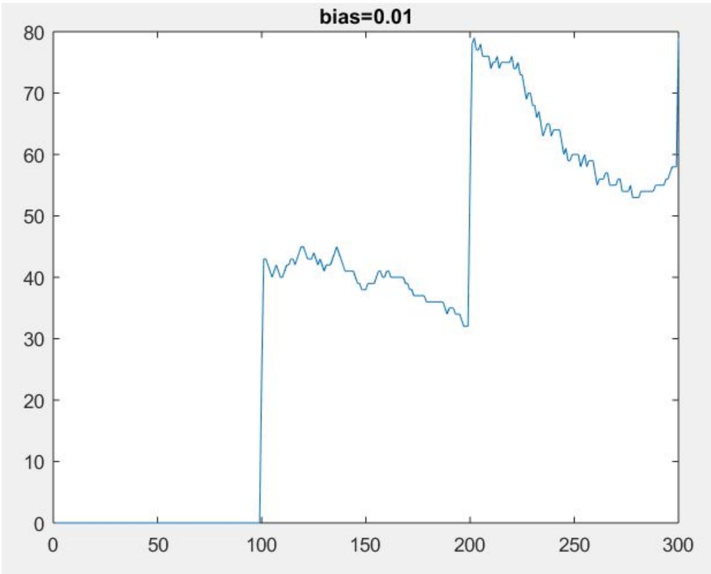


DD2437
Group name: Vairamuthu_Guan
rvai@kth.se
xguan@kth.se
miloulis@kth.se

2) Bias = 0.05
    a) Sparsity = 0.05



    b) Sparsity = 0.01



Using the sparse technique we see that very high biases are handled really well and the number of patterns learnt by the