

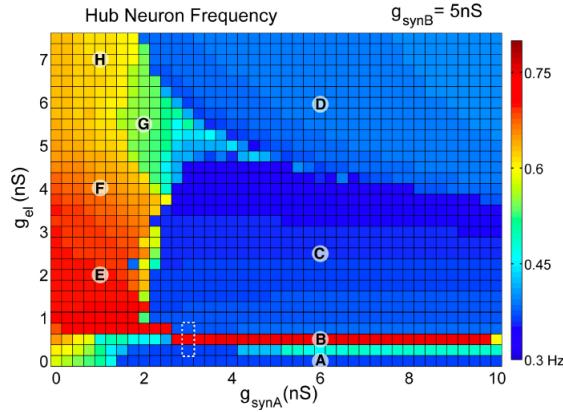
# STG introductory example

Sean Bittner

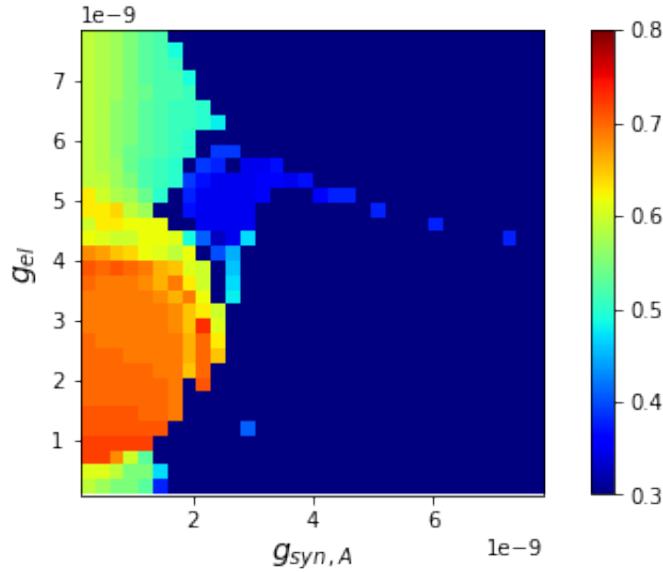
May 22, 2019

## 1 Background

As we've discussed, our introductory example for the DSN is going to be the STG. We want to pick out the regions of conductance space that produce medium levels of hub neuron frequency  $0.525\text{Hz} \pm 0.025\text{Hz}$  (green, Fig. 2, Gutierrez et al. 2013).



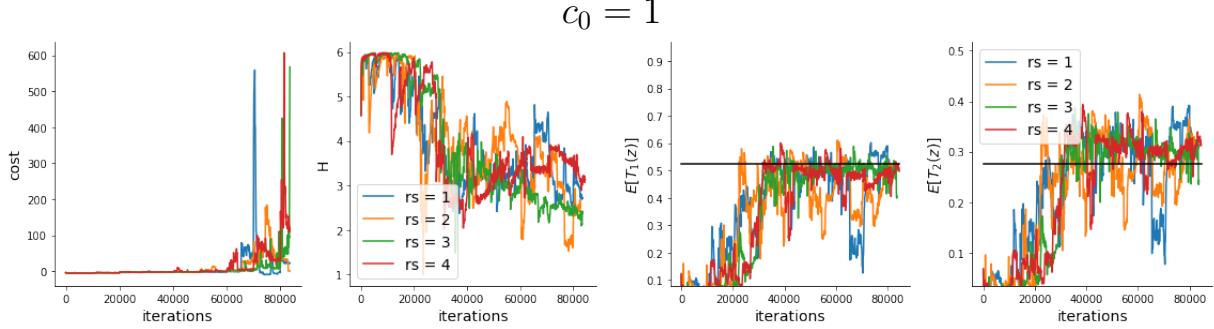
Before training DSNs at the full resolution necessary to exactly reproduce this ( $T=800$ ), I am using  $T=200$  to get an idea of the appropriate augmented lagrangian optimization parameterization that will yield bio-modal (the two green areas below) distributions in this STG setting.



I ran 8 30-hour DSN training processes:  $c_0 = 1, 100$  with 4 random initial seeds each. Each 10 planar flow  $\rightarrow$  scale  $\rightarrow$  shift normalizing flow architecture was initialized to a truncated gaussian

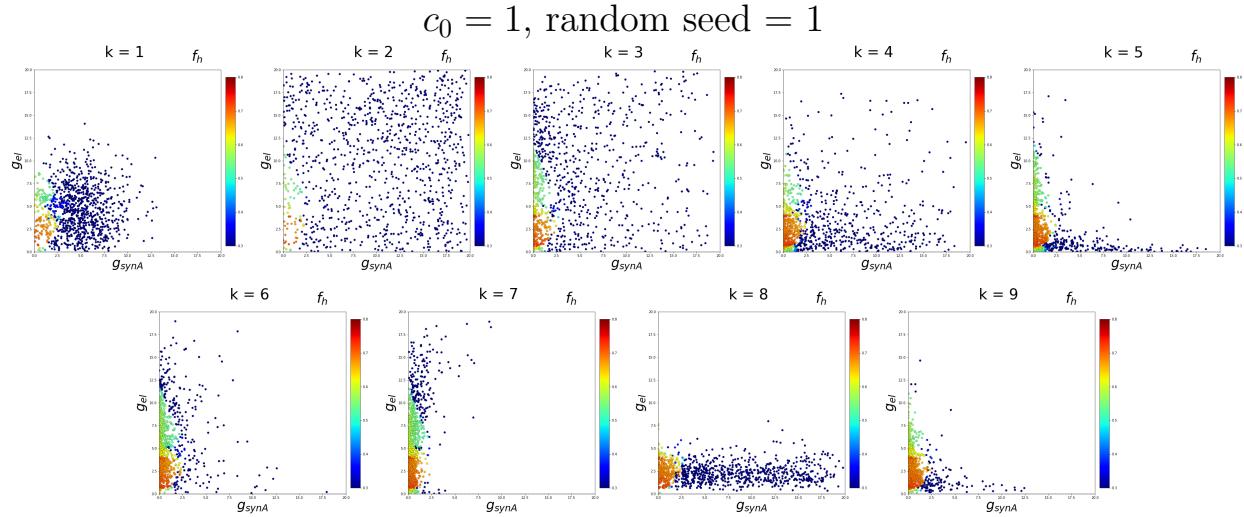
with mean of (4 nS, 4 nS), with a standard deviation of 3 nS. Each augmented lagrangian epoch is 10,000 iterations. This number of iterations for an augmented lagrangian epoch is much higher than I usually use, since we want to get close to convergence for by the end of each epoch. This is desireable, since we want to reason about the dynamics of the entire AL optimization. I am using an AL factor of  $\beta = 4$  as in the MEFN paper, although we can modify this.

## 2 Results

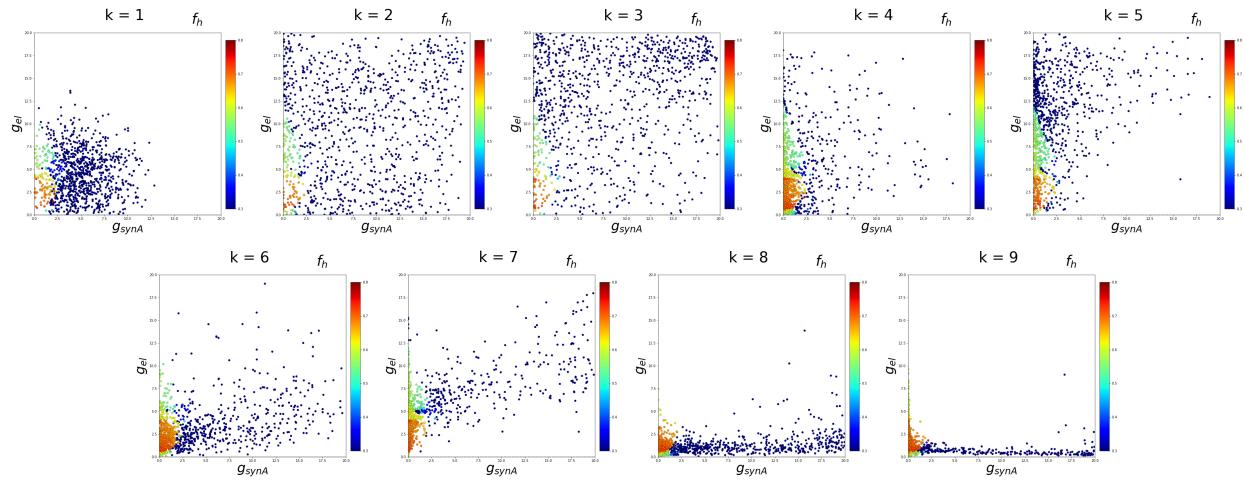


In each of the four optimizations, the entropy ( $H$ ) initially increases, and begins to decrease during the 3-4th AL epoch. At the beginning of the 2nd and 3rd AL epochs ( $k = 2, 3$ ), we should expect a relatively expansive distribution over the interval  $(0,20)$ ,  $(0, 20)$ . This is indeed what we see below, as we visualize the DSN distribution for each random seed at the beginning of each AL epoch.

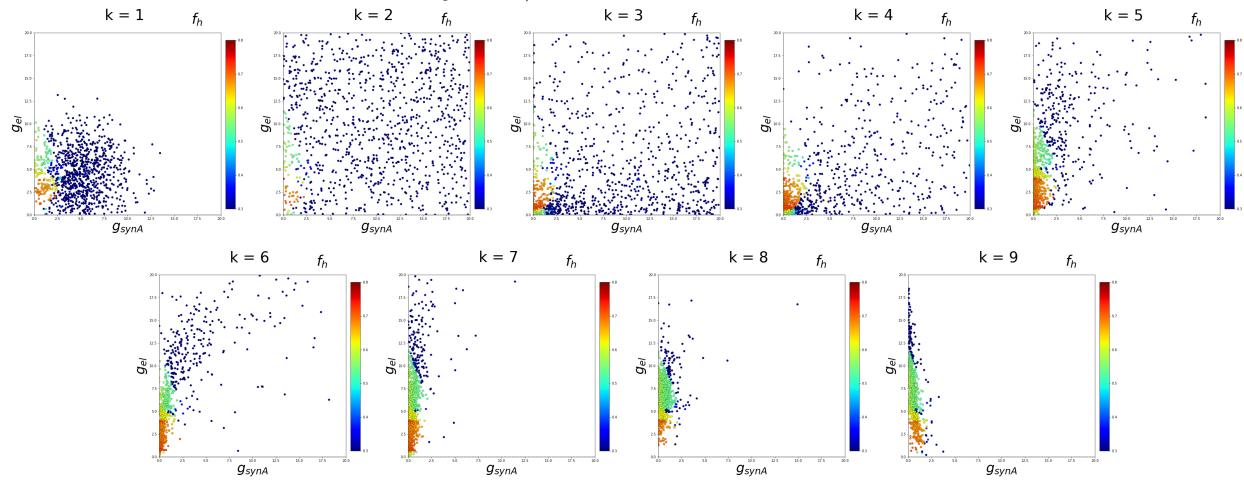
By the end of 8 AL epochs, none of the DSNs have converged via the hypothesis testing criteria, but they are rather close. In general, the means are a bit too low, and the second moments are a bit too high. (If the means are undershot and second moments are overshot, there is a considerably greater amount of variance in these plots then we will have at convergence.) We see a mixture of DSNs which are still emitting samples in both green modes ( $rs = 1, 4$ ), and have focused on the upper, more expansive mode ( $rs = 2, 3$ ).



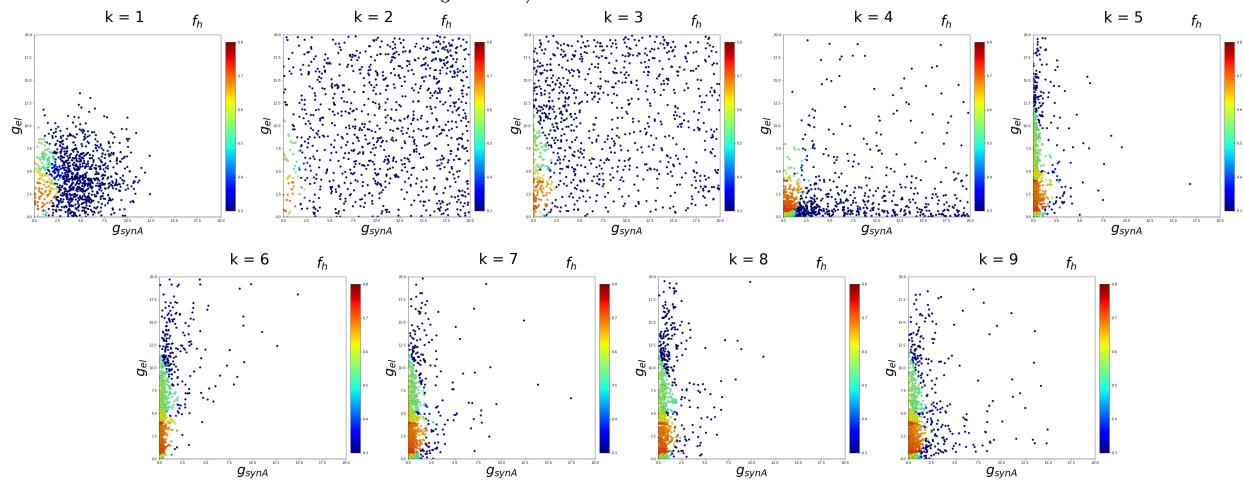
$c_0 = 1$ , random seed = 2



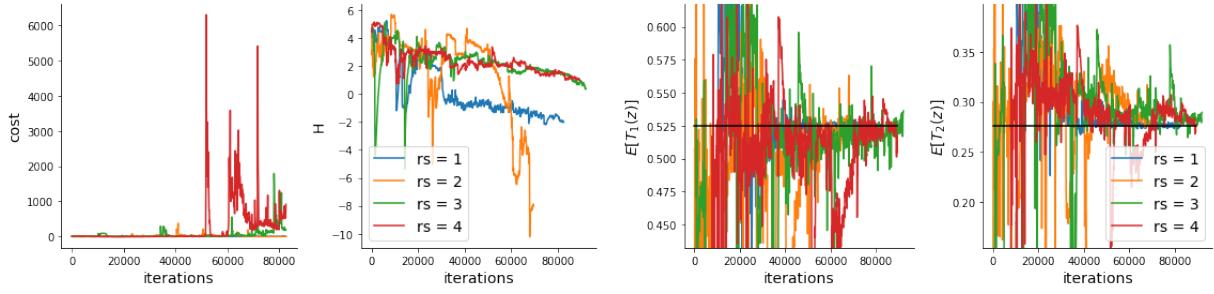
$c_0 = 1$ , random seed = 3



$c_0 = 1$ , random seed = 4

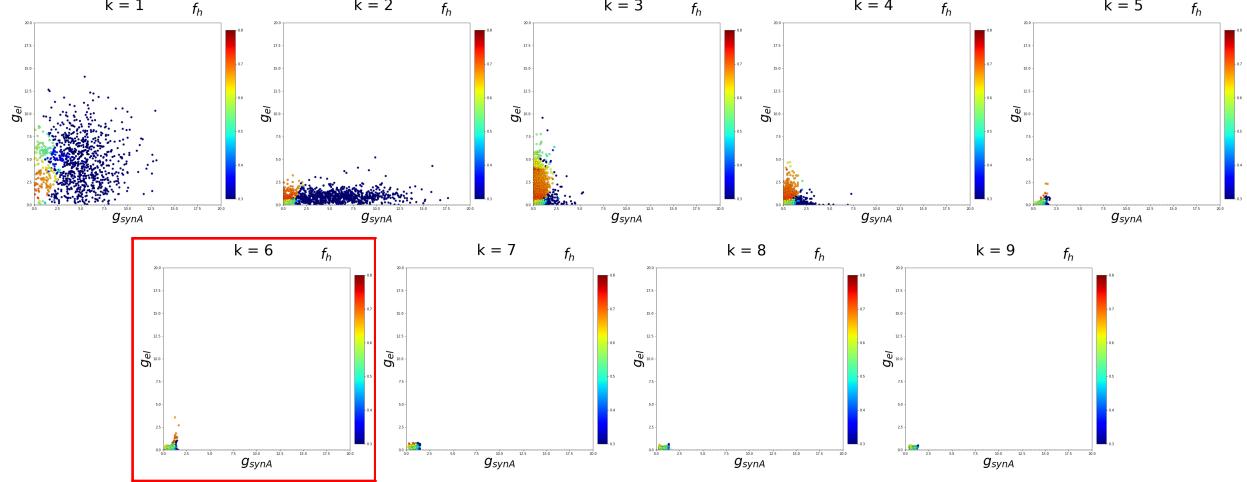


$$c_0 = 100$$

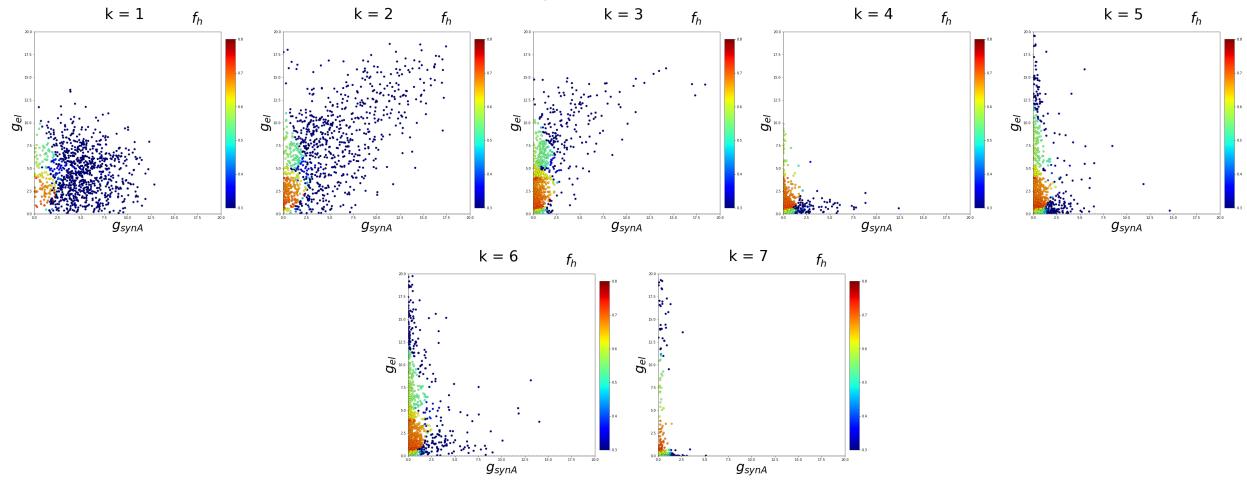


In contrast to  $c_0 = 0$ , we don't see an increase in entropy at the outset. Two of these optimizations actually converge ( $rs=1$ ,  $k=6$  and  $rs=3$ ,  $k=10$ ), but to the individual lower and upper modes, respectively.

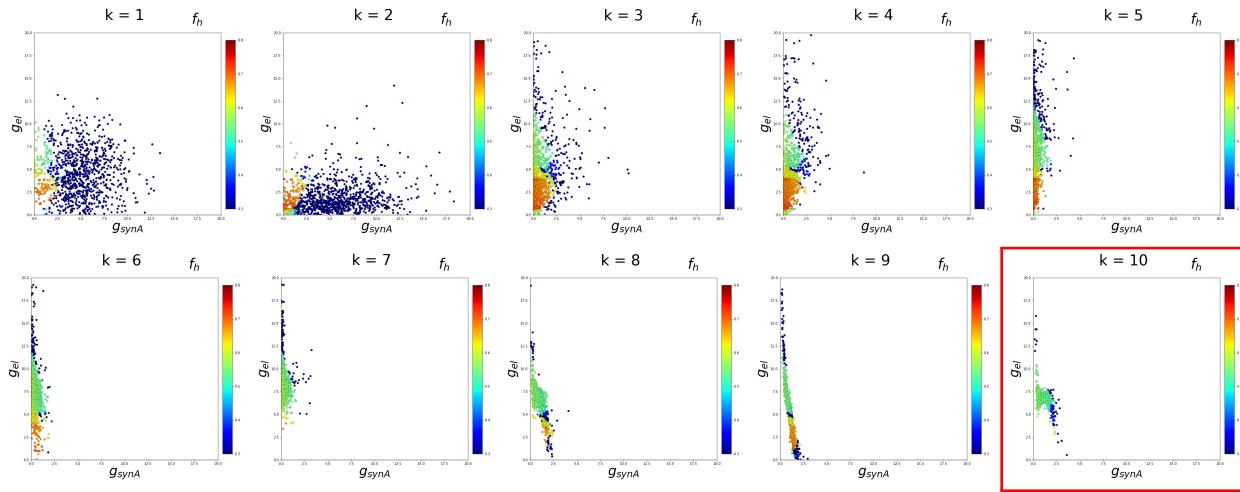
$$c_0 = 100, \text{ random seed} = 1$$



$$c_0 = 100, \text{ random seed} = 2$$



$$c_0 = 100, \text{ random seed} = 3$$



$c_0 = 100$ , random seed = 4

