5
A distance constrained synaptic plasticity model of *C. elegans* neuronal network

The quest for understanding broad structural organization, functional building blocks and mechanisms of control of nervous systems has been central to neuroscience [E.R. Kandel et al., 2000]. Vast knowledge of cellular and molecular mechanisms garnered through reductionist studies over decades, while enriching our understanding of brain mechanisms, have highlighted the need for holistic perspective of neural architecture [Rubinov and Sporns, 2010]. This urge to delve into systems properties has propelled efforts into connectome projects that attempt to map and model neural wirings to the finest detail possible [Chiang et al., 2011; Sporns, 2013b; White et al., 1986; Zingg et al., 2014]. *C. elegans* connectome is the only complete neuronal wiring diagram available till date [Beth L Chen et al., 2006; Towlson et al., 2013; White et al., 1986]. Along with the rich understanding available on the biology of this model organism [Altun, Z.F., Herndon, L.A., Crocker, C., Wolkow, C.A., Lints, R. and Hall, 2016; Howe et al., 2016], its connectome presents an opportunity to learn basic governing principles that drive structure and function of neuronal architecture.

Despite its apparently simple nervous system, *C. elegans* is known to possess complex functions associated to sensation, movement, conditioning and memory [Ardiel and Rankin, 2010; Chatterjee and Sinha, 2008]. This multi-cellular nematode has been extensively investigated to understand neural mechanisms involved in response to chemicals, temperature, mechanical stimulation as well as mating and egg laying behaviors [Chatterjee and Sinha, 2008; Hobert, 2003]. These biological functions have neuronal basis and are a reflection of emergent properties of signal dynamics over the network. Its nervous system has evolved to confer evolutionary benefits under constant tinkering and is known to undergo synaptic rewiring during the course of its life [Eric R Kandel et al., 2014]. Beyond the broad evolutionary architecture, synaptic plasticity offers additional adaptive advantage to respond to the environment and perhaps to achieve better functional efficiency. The key role of distance constraint in shaping the architecture of complex networks has been well studied and highlighted [Amaral, Scala, Barthelemy, and Stanley, 2000; Avena-Koenigsberger, Goñi, Solé, and Sporns, 2015; Barthélemy, 2011].

The *C. elegans* neuronal system could be modelled as a network and studied for structural properties of its neuronal architecture as well as for network dynamics (Figure 5.1). The *C. elegans* neuronal network (CeNN) has been mapped to a high resolution with details of its neurons, their locations and synaptic connectivity [Choe et al., 2004]. The network, comprising of 277 neurons that are interlinked with 2105 synapses, has been studied for its broad structural features as well as towards identification of motifs that potentially contribute to the dynamics over the network. Using graph theoretical measures, CeNN has been observed to have a small world architecture with small path length and high clustering [Watts and Strogatz, 1998]. This has been proposed to be due to processes that leave the network critically poised between absolute order and extreme randomness. The small world nature may render this neural network (as well as other neuronal systems) efficient for information dynamics. When probing for network sub-structures that could form the building block of the CeNN, Milo *et al.* identified feed forward motifs (FFMs) to be significantly over-represented [R Milo et al., 2002]. Such structural building blocks have been suggested to be of functional relevance to biological systems (in addition to other networked systems). How exactly such building blocks may offer functional advantage to networked

systems and whether these entities have evolved to optimize the building blocks is not clearly understood yet.

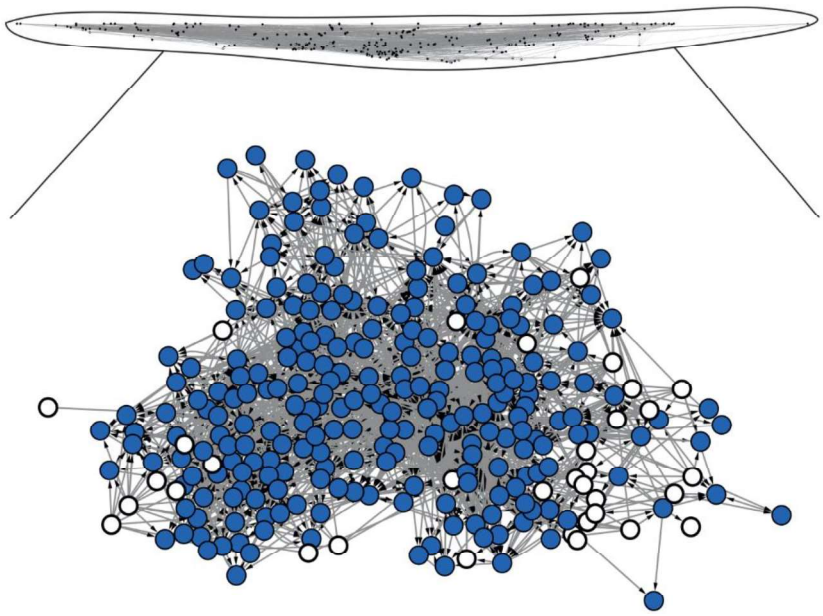

Figure 5.1 : Network structure of *C. elegans* nervous system. Functionally relevant driver neurons (34 nodes highlighted in white) were identified with maximum matching criterion. Beyond explaining the small world nature, saturation of feed forward motifs and observed number driver neurons, the distance constrained synaptic plasticity model accurately identifies specific driver neurons.

Control systems approach to complex networks provides a better perspective of dynamics over the network and ability to steer its 'state' [Y.-Y. Liu et al., 2011]. Neuronal architecture of CeNN forms an important underlying framework which specifies phenotypic features of *C. elegans*. Important behavioral traits as well as cognitive processes (such as movement, sensation, egg laying, mechanoception, chemosensation and memory) are known to have neuronal basis. A network is said to be controllable if it can be reached to a desired state from any initial state by providing inputs to certain nodes [Lin, 1974; Y.-Y. Liu et al., 2011]. The set of nodes that facilitate such a control are named driver nodes [Y.-Y. Liu et al., 2011].

By studying genotypic and phenotypic aspects of CeNN, in our earlier study (Chapter 4) we have shown that 'driver neurons' are associated with important biological functions such as reproduction, signaling processes and anatomical structural development [Badhwar and Bagler, 2015]. Interestingly, randomized controls have no driver neurons as compared to CeNN which presents a sizeable number of driver neurons that are crucial for its control. While earlier studies have shown connectivity of neurons in CeNN partially explains the observed number of driver neurons [Y.-Y. Liu et al., 2011], no model has so far been developed that accounts for its small world architecture, over-representation of FFMs as well as controllability.

In this study, we create one-dimensional (1D) and two-dimensional (2D) network models of *C. elegans* neuronal system to investigate the role of FFMs as building blocks in conferring controllability and small world nature. With the help of a simple 1D ring model we show such a network is critically poised for the number of FFMs, neuronal clustering and characteristic path-length in response to synaptic rewiring, indicating optimal rewiring. We found that synaptic connections between neurons are characterized with a strong distance constraint in CeNN. Using this as a guiding principle, we created a distance constrained synaptic plasticity model that simultaneously explains small world nature, FFM saturation and controllability of the network. This model accounted for the observed number of driver neurons and also accurately identified

specific driver neurons. Thus this model presents realistic process of distance constrained synaptic plasticity as a plausible basis of nature of functional sub-structures and controllability observed in CeNN.

5.1 TOPOLOGICAL PROPERTIES OF *C. elegans* NEURONAL NETWORK

The *C. elegans* neuronal network was modeled as interconnections among 277 somatic neurons via 2105 electrical and chemical synapses (For more details, please see Chapter 3). A typical neuron in CeNN had on an average 7.59 synaptic connections. Starting with CeNN, we studied its topological properties that provide insights into its structure and function [Albert and Barabasi, 2002; Dorogovtsev, 2014]. Consistent with previous reports, we observed that *C. elegans* neuronal network is a small world network by virtue of high clustering coefficient $\bar{C} = 0.172$ and comparable characteristic path length ($L = 4.018$), with respect to its randomized counterpart ($\bar{C}_{ER} = 0.028$ and $L_{ER} = 2.97$)(Table 3.1) [Watts and Strogatz, 1998]. Beyond these global topological features, CeNN is known to be over represented with feed forward motifs [R Milo et al., 2002] that are functionally associated with mechanisms of memory [Mozzachiodi and Byrne, 2010]. We observed that, FFMs were significantly overrepresented in CeNN ($Zscore = 151.12$) as compared to those in corresponding random graphs. In CeNN, feed forward motifs are most prevalent among all unidirectional motifs as shown in Figure 5.2.

Table 5.1 : Topological properties of CeNN and its controls.

	CeNN	**ER**	**DD**
$\bar{C}$	0.172	0.028±0.001	0.067±0.003
L	4.018	2.97±0.01	2.981±0.018
n_D	34	0.28±0.514	22.38±1.153
n_{FFM}	3776	438.3±22.1	1699.6±57.5

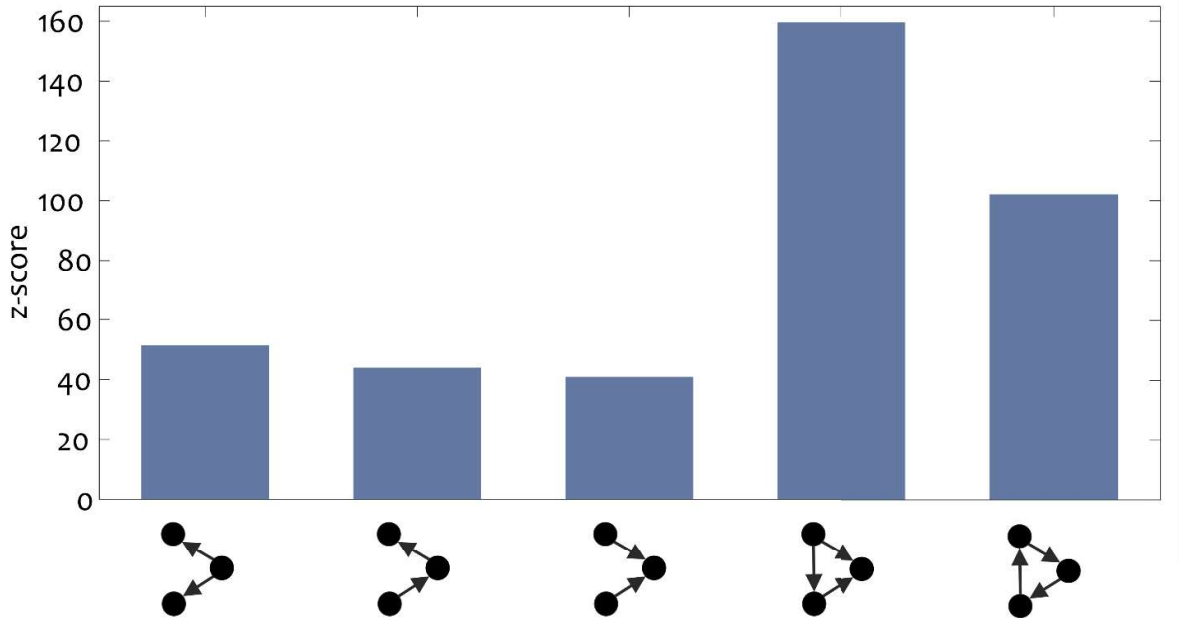

Figure 5.2 : Statistics for unidirectional three node motifs depicting over-representation of feed forward motifs. The Z-Score was computed in comparison to 100 instances of random controls (ER) of CeNN.

From control systems perspective CeNN can be controlled through a small set of driver neurons (34) to any desired state in finite time [Y.-Y. Liu et al., 2011]. The number of driver neurons in CeNN is significantly higher in comparison to its random counterpart. As described in Chapter 4, driver neurons in CeNN are genotypically and phenotypically associated with biological functions such as reproduction and maintenance of cellular processes [Badhwar and Bagler, 2015]. This alludes to the fact that driver neurons serve a critical role in the neuronal architecture of *C. elegans* and the number of driver neurons therefore has functional bearing on its control.

Table 5.1 depicts topological features of CeNN that are potentially critical for specifying its function. Other than the small world nature, evident from high clustering among neurons, the CeNN is characterized with significantly higher number of driver nodes as well as number of feed forward motifs. While connectivity/degree of neurons (DD) partially explains the increase in FFMs as well as that in n_D, at the same time it cannot account for observed clustering. No comprehensive model that can explain all of these functionally relevant features is hitherto known.

5.2 1D RING MODEL OF CeNN

We constructed a ring graph model of CeNN so as to maximise the number of FFMs while preserving the number of neurons n as well as average neuronal connectivity k of CeNN (Figure 5.3). While the core idea and strategy implemented in this model is analogous to that of Watts and Strogatz's [Watts and Strogatz, 1998], it is extended to represent directed edges (synapses), and hence naturally accommodates network motifs and controllability analysis. Starting with $n(=277)$ nodes arranged in a circular manner, every node was connected (in anti-clockwise sense) to its next nearest neighbour with a directed edge. The procedure was repeated to connect every node with its nearest neighbour and the next nearest neighbour until the out-degree of every node matched with that of average out-degree of CeNN ($k = 7.59 \approx 8$). This strategy maximises the number of FFMs to $nk\,(k-1)/2$ in the regular graph model of CeNN and represents an asymptotic version saturated with FFMs.

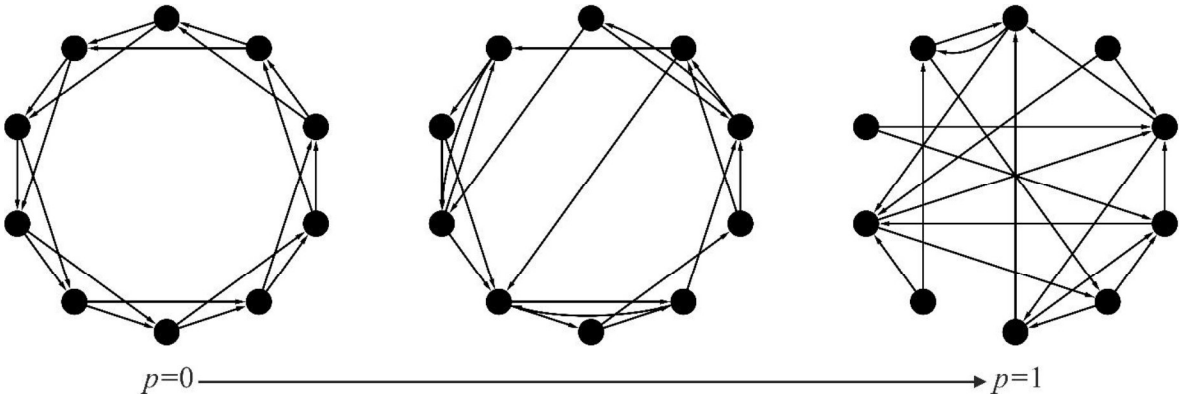

Figure 5.3 : The 1D ring model, with neurons linked for maximizing number of feed forward motifs, was rewired with increasing probability of synaptic rewiring. Starting with an asymptotic model (with 277 nodes and 8 out-going edges) saturated with FFMs, synaptic rewiring was emulated with probability p. The model exhibits a spectrum of topological variations between extreme regularity and randomness. The figure shows an illustration for 10 nodes and 2 outgoing edges. See Figure C.1 of Annexure C for another illustration with larger network size.

Starting from such a regular ring graph maximally saturated with 7756 FFMs, we simulated random synaptic rewiring to observe its effect on topological features. We rewired every edge in this network with a certain 'probability of rewiring (p)'. Every out-going edge connecting a node to its nearest neighbor was chosen and rewired randomly with probability p by ensuring that there were no duplicate edges or self-edges and that the network is always

connected. In the second lap, the process was repeated for the edges made with next-nearest neighbors and so on. All edges are thus exhaustively considered for rewiring in k laps. For every probability of rewiring 1000 instances of graphs were created for a range of $p = 10^{-4}$ to $p = 1$. Topological properties ($\bar{C}, L, n_{FFM}$ and n_D) were computed for every instance of graph thus generated.

In addition to FFM saturation, the regular graph had very high average clustering coefficient $\overline{C_{reg}} = 0.35$ as well as characteristic path-length $L_{reg} = 17.69$. From an analogous undirected Watts and Strogatz model it was anticipated that with increase in synaptic rewiring the clustering as well as path-length would decrease to approach that of random graph asymptotically [Watts and Strogatz, 1998]. This simulation of synaptic rewiring was also expected to provide insights into its impact on number of FFMs and driver neurons. As shown in the Figure 5.4, with increasing probability of synaptic rewiring the number of FFMs is unaffected up to $p \approx 0.01$ before falling sharply. While this result points at a critical threshold for number of FFMs in response to probability of synaptic rewiring, no driver neurons were presented by the model across the simulation ($n_D \approx 0 \; \forall \; 0 \leq p \leq 1$). For 1D ring graph, these results highlight a critical threshold of rewiring for which the network has optimum saturation of FFMs. Such a simple model can only provide topological insights devoid of biological basis. Search for a more realistic model prompted us to look for biological constraints that may dictate synaptic rewiring as well as to build a 2D model that could possibly reveal mechanisms that render observed controllability in CeNN.

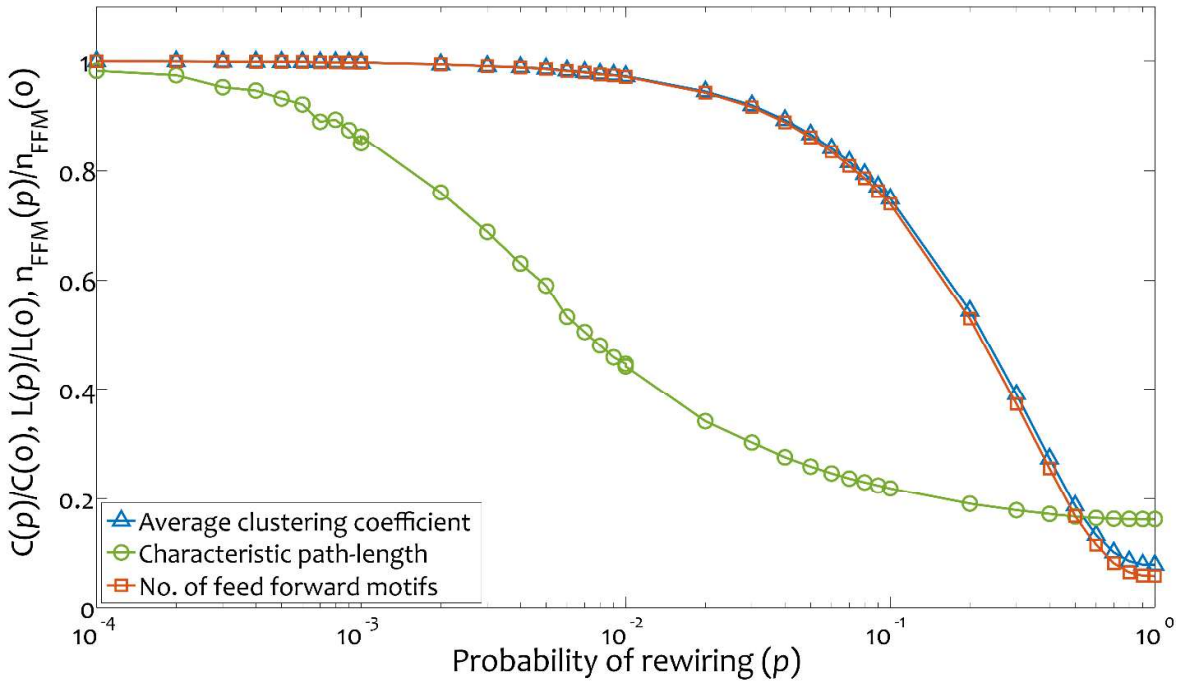

Figure 5.4 : Response of 1D ring model with changing probability of synaptic rewiring was measured in terms of average clustering coefficient ($\bar{C}$), characteristic path-length (L), number of FFMs (n_{FFM}) and number of driver nodes (n_D). For intermediate values of p, the model exhibits small world phenomenon as well as FFM saturation, but cannot account for controllability ($n_D = 0 \; \forall \; p$). All parameters were normalized with respect to the initial ring graph ($p = 0$). Error bars represent standard deviation over 100 instances. Please see Figure C.2, Figure C.3 and Figure C.4 in Annexure C for non-normalized data.

5.3 CeNN FOLLOWS DISTANCE CONSTRAINED SYNAPTIC CONNECTIVITY PATTERN

We measured the connectivity pattern in CeNN by enumerating number of neuron pairs that are synaptically connected and Cartesian distance between them. We observed that the synaptic connections were constrained by distance as evident from the power law observed from neuronal connectivity data Figure 5.5. The probability of two neurons being connected scales as a power law $p(d) \propto d^{-\alpha}$, with presence of a few exceptional long distance connections with an exponent $\alpha = 2.02$ ($p - value = 0.92$). The power law nature of data was established following the strategy prescribed by Clauset et al. [Clauset et al., 2009].

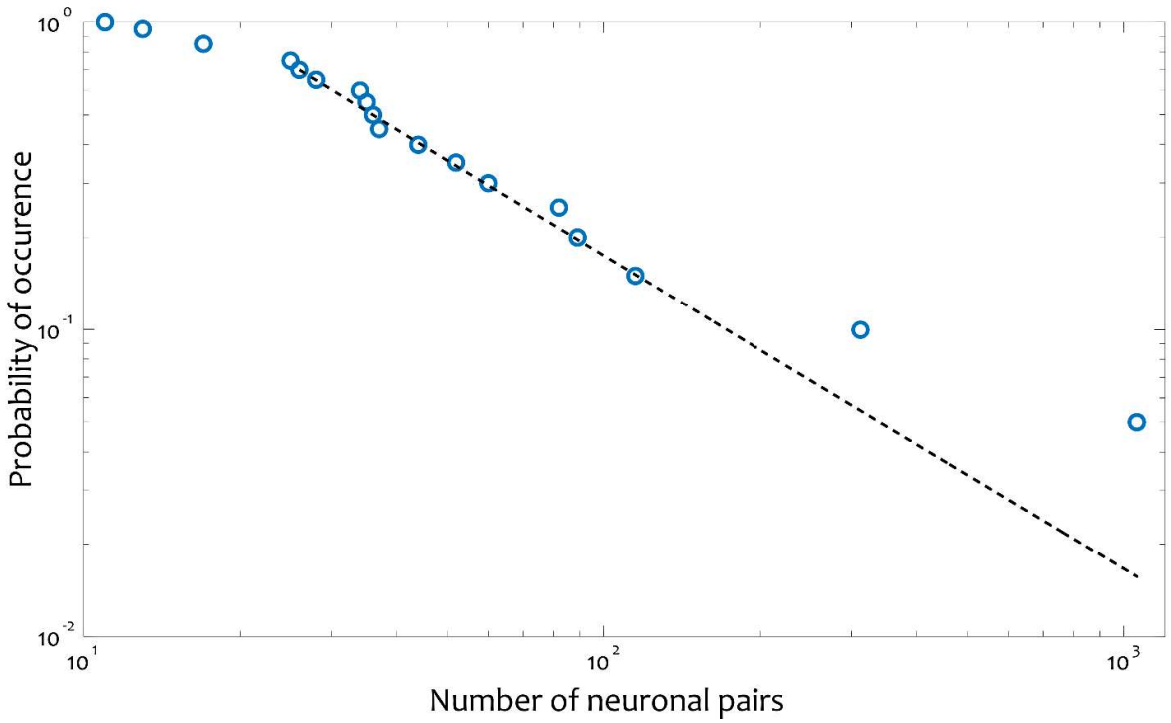

Figure 5.5 : Empirical distance constrained synaptic connectivity pattern observed in *C. elegans* neuronal wiring. The number of synapses that connect neurons at distance d follows a power law pattern with an exponent of $\alpha = 2.02$ ($p - value = 0.92$) [Clauset et al., 2009].

To incorporate this empirical distance constrained connectivity pattern, we created more realistic 2D models: (1) Distance constrained random (DCR) model that adds distance constraint starting from ER control, and (2) Distance constrained plasticity (DCP) model that overlays the distance constraint starting from the DD control. Along with the ER and DD random controls these models allow us to segregate the contribution of degree (connectivity) vis-à-vis distance constrained synaptic wiring towards conferring observed topological features upon CeNN.

5.4 DISTANCE CONSTRAINED MODELS OF CeNN

We created 2D distance constrained models that, similar to distance constraint observed in CeNN, follow a restraint on synaptic connectivity based on distance between two neurons. These 2D models are based on positional data of *C. elegans* neurons, that have been mapped to a high resolution [Choe et al., 2004]. In these models, the probability $P(d)$ that two neurons at a distance d are connected with a synapse approximately follows a power law pattern observed from empirical data (Eq.(5.1)).

$$P \propto d^{-\alpha} \tag{5.1}$$

The distance constraint is modulated by the exponent $0 \leq \alpha \leq \infty$. Here, the distance between neurons i and j, $d(i,j)$, was calculated as the Euclidean distance (Eq.(5.2)).

$$d(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{5.2}$$

The power law nature of neuronal connectivity was established following the recipe suggested by Clauset *et al.* [Clauset et al., 2009].

We created two models of CeNN based on the distance constraint: Distance constrained random (DCR) and Distance constrained synaptic plasticity (DCP).

5.4.1 Distance constrained random (DCR) model

The underlying framework for DCR model is that of ER control. Starting with ER (random) control, we rewired every edge to impose distance constraint for specific exponent α. Statistics of topological parameters were computed over 100 instances. Response of DCR model was observed by varying the value of exponent between $0 \leq \alpha \leq 3$.

5.4.2 Distance constrained synaptic plasticity (DCP) model

In contrast to DCR model, the underlying framework for DCP model is that of DD control which preserves the synaptic connectivity of each neuron. Starting with DD control, every edge was rewired to impose distance constraint for specific exponent α ($0 \leq \alpha \leq 3$) and statistics of topological parameters were computed over 100 instances.

5.4.3 Cartesian graph model of CeNN

The deterministic Cartesian graph model of CeNN was created by ensuring that every neuron is connected to its spatially nearest neurons. Beginning with (n =)277 neurons placed at Cartesian coordinates matching their observed position in the nervous system of *C. elegans* [Choe et al., 2004], every neuron was connected to (k =)8 spatially nearest neurons. This model reflects preferential deterministic connections made by a neuron based on its distance from another neuron.

5.5 RESULTS

5.5.1 Distance constrained random model

The distance constrained random model is a 2D model in which the number of neurons, number of synapses and neuronal locations were preserved. Starting from initial random connectivity (ER) every synapse was probabilistically rewired to follow distance constraint with a certain α Figure 5.6. The lower asymptotic limit of this model converges to ER model for $\alpha = 0$. With increasing α the probability of long distance synaptic connections decreases. For extremely large values of α this model converges to the Cartesian model in which every neuron is deterministically connected to its spatially nearest neighbors. We varied the value of α between 0 and 3 to assess its impact on the topology of neural network. We found that the average clustering coefficient, characteristic path-length and number of FFMs monotonously increase with increasing α. For these parameters the DCR model was closest to actual neuronal network of *C. elegans* for $\alpha = 0.6$. While this model took us closer to CeNN, it did not reflect controllability measured in terms of n_D. The driver neurons vanish for asymptotic limits of α with maximum $n_D = 15.7$ for $\alpha = 0.4$. The fact that DD control, in which number of synapses of every neuron is preserved, matches with CeNN better in controllability (Table 5.1) prompted us to create a more refined 'distance constrained synaptic plasticity model'.

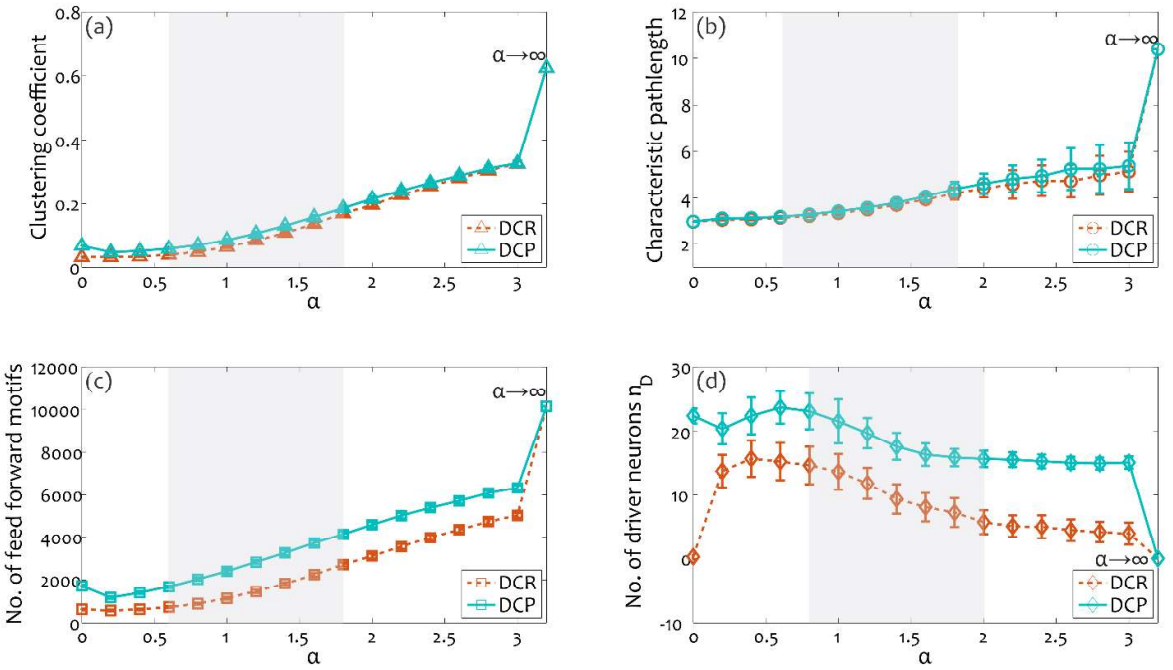

Figure 5.6 : Response of distance constrained models of CeNN (DCR and DCP) with increasing constraint (α) measured in terms of (a) average clustering coefficient ($\bar{C}$), (b) characteristic path-length (L), (c) number of FFMs (n_{FFM}), and (d) number of driver nodes (n_D). The lower the α more heterogenous are the synaptic lengths (larger proportion of long range synapses). For $\alpha = 0$, DCR and DCP models converge to ER and DD controls, respectively. For asymptotic limits of $\alpha < \infty$ both the models converge to the Cartesian model, a regular model with saturation of FFMs coupled with high clustering but devoid of driver nodes. While the small world nature (reflected in high clustering and low path-length) and FFM saturation is realized by both DCR and DCP models, DCP model stands out in reproducing all key features of CeNN for $0.6 < \alpha < 1.8$ (highlighted with gray background). Please see Figure C.5, Table C.6 and Table C.7 of Annexure C for data associated with this figure.

5.5.2 Distance constrained synaptic plasticity model

The distance constrained synaptic plasticity (DCP) model preserves the number of synapses of every neuron over and above the number of neurons and their locations. While following the distance constraint, this model mimics synaptic rewiring that is known to take place in CeNN [Eric R Kandel et al., 2014; Shen and Bargmann, 2003]. We observed the response of topological features for varying extent of distance constraint $0 \leq \alpha \leq 3$ (Figure 5.6). In addition to the clustering and characteristic path-length, interestingly, this model successfully realised number of FFMs as well as number of driver neurons. We found that for an intermediate distance constraint of $\alpha = 0.6$ this model is closest to CeNN in reproducing number of driver neurons that are critical for control of the network (Figure 5.6(d)). This model presents a range of distance constraint parameter $0.6 \leq \alpha \leq 1.8$ for which the small world nature as well as functionally relevant features of regulatory motifs and controllability were realistically exhibited. The number of driver nodes in DCP model was always higher than those returned by DCR model. For $\alpha < 1.8$ i.e. in the presence of strong distance constraint, DCP model returned significantly high number of driver nodes higher than maximally displayed by DCR model. This points at the role of long distance synaptic connections in conferring observed nature of control in CeNN.

5.5.3 Identification of specific driver neurons

We obtained the set of specific driver neurons using maximum matching algorithm [Pothen and Fan, 1990]. The success of DCR, DCP models in accurate identification of driver neurons was measured with the help of $F1\ score$. Using the driver neurons set identified from the CeNN (34) as the basis (Details of driver neurons is provided in Table C.8 of Annexure C), we identified true positives (TP) and true negatives (TN) (neurons that are correctly classified) as well as false positives (FP) and false negatives (FN) (neurons that were incorrectly marked as

driver neurons, and neurons that were incorrectly marked as non-driver neurons, respectively) for DCR and DCP models across 100 instances. The $F1\ score$, which is used for computing the quality of binary classification is defined as Eq.(5.3).

$$F1 = \frac{2TP}{2TP + FP + FN} \tag{5.3}$$

While the DCP model successfully reproduces key topological features important for function and control of CeNN (Figure 5.6), the question is whether it can also capture specific neurons implicated in control of the network and not just the number of driver neurons (Figure 3.5(d)). We compared specific neurons identified by the minimum driver neurons set obtained from real-world CeNN with that obtained from distance constraint models, for varying extent of distance constraint (Figure 5.7). Neurons that were consistently identified as driver neurons over 100 random instances of DCP and DCR models were distilled (True Positives). The classification accuracy of these models was assessed using $F1\ score$. Interestingly, we found that the performance of DCP model was significantly better compared to DCR model (the accuracy of which was indistinguishable from that obtained from random sampling) and superior to DD control in the presence of dominant long distance connections ($\alpha \leq 1$). Thus DCP model is not only closer to real-world network in terms of number of driver nodes but also accurately identified specific neurons that can drive the network dynamics. In summary, the DCP model, that embeds empirically observed phenomenon of neuronal rewiring in addition to fixed neuronal connectivity, successfully recreates topological features of functional relevance to *C. elegans*.

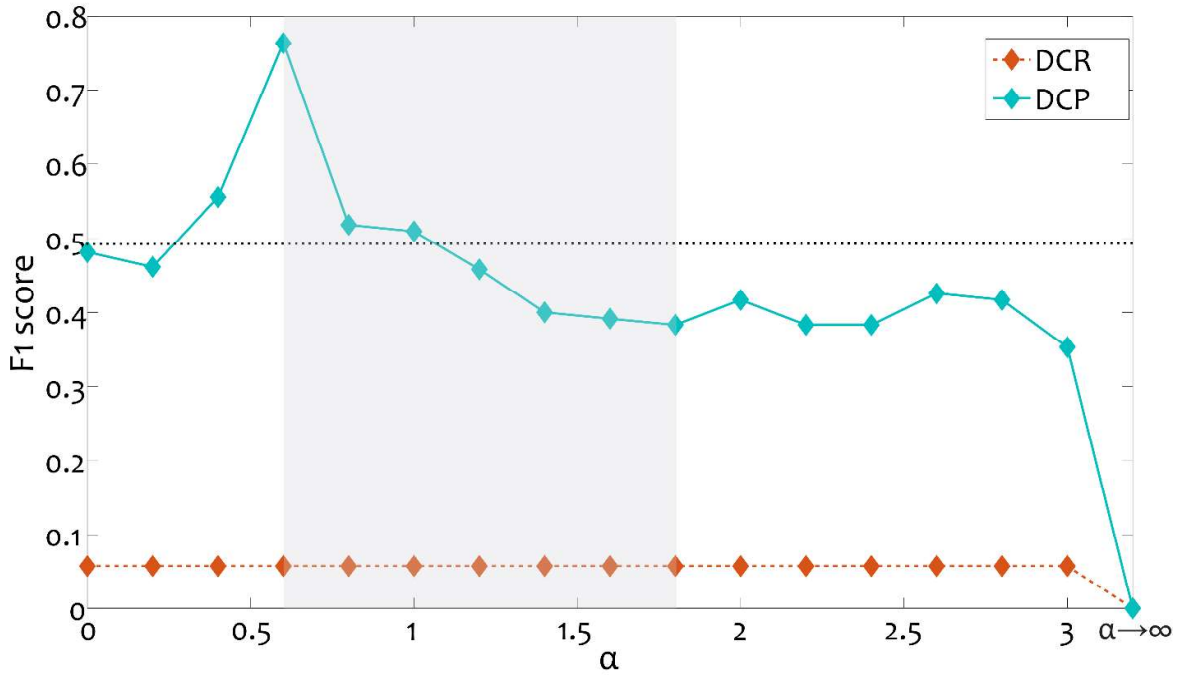

Figure 5.7 : Accuracy of identification of specific driver neurons with changing distance constraint exponent. The DCR model, with random synaptic connectivity pattern, fared poorly. The performance of DCP model was consistently better than that of the DCR model indicating the critical role played by the distance constraint in specifying the control of the neuronal network. For optimum distance constraint ($\alpha = 0.6$) DCP model provides the best match with the reality ($F1\ score = 0.77$), better than what could be accounted for by only neuronal connectivity (DD control; indicated with a dashed line). The spectrum of distance constraint regime for which DCP model is closest to CeNN ($0.6 \leq \alpha \leq 1.8$) is highlighted with gray background.

5.6 DISCUSSION

Brain research has been driven by inquiry for principles of brain structure organization and its control mechanisms. The neuronal wiring map of *C. elegans*, the only complete connectome available till date, presents an incredible opportunity to learn basic governing principles that drive structure and function of its neuronal architecture. Despite its apparently simple nervous system, *C. elegans* is known to possess complex functions. The neuronal architecture forms an important underlying framework which specifies phenotypic features associated to sensation, movement, conditioning and memory [Ardiel and Rankin, 2010]. CeNN seemingly has evolved as a small world network with high clustering and low characteristic path-length for functional benefits [Watts and Strogatz, 1998]. Other than the small world global architecture CeNN is reported to be enriched with number of feed forward motifs among all possible three and four node motifs [R Milo et al., 2002]. Our results suggest that the heterogeneous composition of motifs dictated by FFMs contributes to increased clustering as well as control of the network.

Analysis of neuronal architecture of CeNN has revealed that the network is optimally wired [Beth L Chen et al., 2006; Pérez-Escudero and de Polavieja, 2007] and is dictated by constraints [Itzhack and Louzoun, 2010; Pan, Chatterjee, and Sinha, 2010; Towlson et al., 2013]. Till date a few simple null models of CeNN have been implemented with network feature constraints [Erdös, P., 1984; Maslov and Sneppen, 2002]. These studies suggest that neuronal connectivity plays a key role in rendering clustering as well as presentation of as many driver neurons as observed in CeNN [Y.-Y. Liu et al., 2011; Maslov and Sneppen, 2002]. None of these models has been able to explain all network features, especially clustering and number of driver neurons, claimed to be of biological relevance [Badhwar and Bagler, 2015; Y.-Y. Liu et al., 2011].

Here, we present a distance constrained synaptic plasticity model that accounts for high clustering, FFMs saturation and large number of driver nodes. With a 1D ring model maximized for feed forward motifs, we show that such a model exhibits critical phenomenon in response to increased probability of synaptic rewiring. While this simple model lends interesting insights into the mechanisms of CeNN architecture, it cannot capture the aspect of controllability. Rooted in empirical observation of distance constraint followed in neuronal connections, we built more realistic 2D distance constrained models with random connectivity (DCR) and degree preserved connectivity (DCP). The latter model, that mimics real-world *C. elegans* neuronal wiring and follows a distance constrained synaptic plasticity mechanism, comes closest to the CeNN in presenting small world architecture, dominance of FFMs and nature of controllability within a range of free variable α. The DCP model also successfully captures specific driver neurons with impressive accuracy. Our results suggest that the extent of synaptic plasticity in CeNN is optimized so as to acquire key structural and dynamical network features.

Clearly, while the DCP model highlights the role of synaptic plasticity and distance constrained neuronal connectivity in specifying structural features and control architecture of CeNN, it is limited in many ways. The present study overlooks functional differences of synaptic links such as chemical synapses and gap junctions. Also, the strength of synaptic connections (edge weight) were ignored in these unweighted network models. Models that factor in such biologically relevant aspects, which are ignored in this study in favor of simplicity, may yield more enriched representations of CeNN.

…