NEUROSCI C151 Spring 2024 Final Presentation

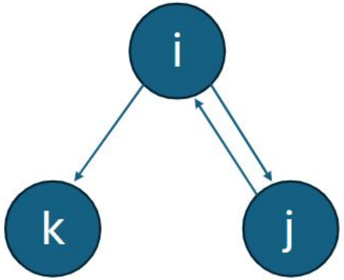
Analyzing different initial conditions in motif extraction to identify patterns in distribution of multiplicity

MOTIF GANG
Pranav Daggubati, Lyla Liu

Instructor: Sharmila Venugopal
TA: Ryan Kochis

Introduction & Background

- Motif: the representation of a possible pattern of connections within a group of cells, with connections depicted as edges and the cells as nodes
- Motif classes are groups of motifs that show congruent connections. In total, there are 64 motifs that fall into one of these classes.
- <u>Bidirectional Connections</u>: Often associated with stabilizing network activity and enhancing communication between neurons.
- Motif Model: Network connections represented by pre and post synaptic matrix coordinates
- Motifs represented by vectors



Post

$$\text{Pre}\quad\begin{array}{c|c} & \begin{matrix} i & j & k \end{matrix} \\ \begin{matrix} i \\ j \\ k \end{matrix} & \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$[3 \quad 1 \quad 0]$$

$$1 * i,j + 2 * j,i$$
$$1 * i,k + 2 * k,i$$
$$1 * j,k + 2 * k,j$$

1. Biemann, C., Krumov, L., Roos, S., Weihe, K. (2016). Network Motifs Are a Powerful Tool for Semantic Distinction. In: Mehler, A., Lücking, A., Banisch, S., Blanchard, P., Job, B. (eds) Towards a Theoretical Framework for Analyzing Complex Linguistic Networks. Understanding Complex Systems. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-47238-5_4
2. Elisabeth Wong, Brittany Baur, Saad Quader, Chun-Hsi Huang, Biological network motif detection: principles and practice, Briefings in Bioinformatics, Volume 13, Issue 2, March 2012, Pages 202–215, https://doi.org/10.1093/bib/bbr033
3. R. Milo et al. ,Network Motifs: Simple Building Blocks of Complex Networks. Science 298, 824-827 (2002). https://doi.org/10.1126/science.298.5594.824

Problems

- Certain motifs do get favored in these randomized modeling studies.
- Motifs incorporating complex/fully connected connections have much higher ratios of prevalence than less complex connected networks.
- Certain motifs show up more significantly than others when given certain starting conditions.

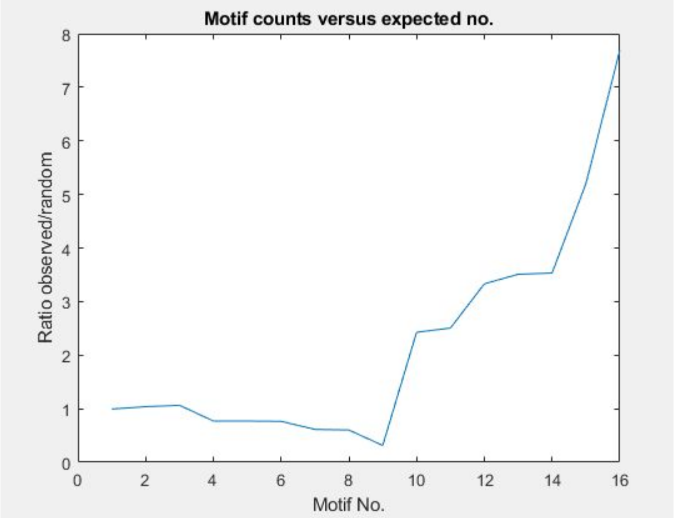


Figure: Motif counts versus expected number. The graph illustrates the ratio of observed motif counts to the expected counts in a randomized model for different motifs numbered 1 through 16. The x-axis represents the motif numbers, while the y-axis shows the ratio of observed counts to expected counts in a random model.

Hypothesis

There are starting conditions which will show a significant preference for the rise of complex motifs over non-complex motifs.

Approach Overview

1. Determining what sets of initial conditions to use to check a broad range of possibilities
2. Conducting motif extraction on these different initial conditions
3. Conducting analysis of motif prevalence in each group of conditions
4. Calculate statistics on whether some motif modalities are more prevalent than others in certain situations.

Krauss, P., Zankl, A., Schilling, A., Schulze, H., & Metzner, C. (2019). Analysis of structure and dynamics in three-neuron motifs. Frontiers in Computational Neuroscience, 13, 5. https://doi.org/10.3389/fncom.2019.000050

Initial Conditions:

All probabilities are mapped with each other

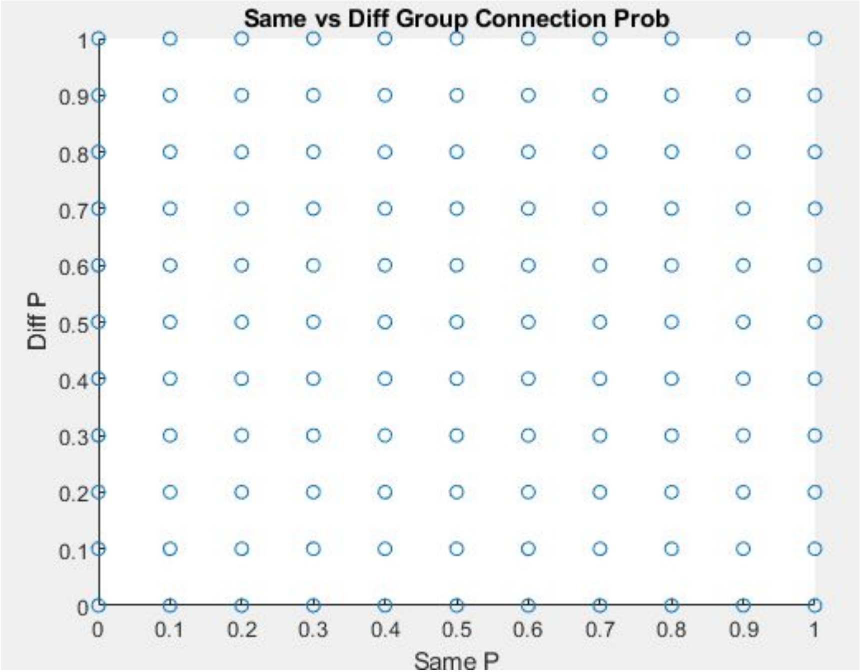


Figure: Same vs Different Group Connection Probability. This figure shows the mapping of connection probabilities within the same group (Same P) and between different groups (Diff P). Each point represents a combination of these probabilities, covering the entire range from 0 to 1 for both parameters. This mapping is used to explore various initial conditions for motif extraction and analysis.

Workflow

STEP 1

Creating Initial Variables

- Need a meshgrid of p_same_group and p_diff_group values
- These values determine the initial probability of a cell connecting to its own group or to a different group

STEP 2

Make Connectivity Matrices

- For each pair of probability values, generate a connectivity matrix. This involves assigning group IDs to cells and calculating connection probabilities based on the initial conditions.

STEP 3

Get ratios of motifs and motif classes

- Extract and calculate the observed-to-expected ratios of motifs from each connectivity matrix. This helps in identifying prevalent motifs under different initial conditions.

STEP 4

Analysis

- Conduct statistical analysis on the motif ratios to identify patterns and correlations. This step involves examining the distribution and significance of various motif classes under different initial conditions.

Functions

- generateConnectivityMatrix.m
  - Creates a connectivity matrix and a correlation matrix (unused)
- tripletMotifs.m
  - Analyzes the connectivity matrix and returns the expected probabilities and experimental probabilities of motif expression
- motifRatios.m
  - Runs generateConnectivityMatrix and tripletMotifs.m as a function to return a 2-value cell
    - Motif Ratios and Class Ratios
  - This is fed into a loop to repeat experiment multiple times

```
num_iters = 15;
motif_cell = cell(1,num_iters);
class_cell = cell(1,num_iters);

for i = 1:num_iters
    ratio_cell = motifRatios(Ngroups, Ncells, p_same_group_values, p_diff_group_values);
    motif_cell{i} = ratio_cell{1};
    class_cell{i} = ratio_cell{2};

end
```

Parameters Used

Number of trials: 15

Number of cells/nodes: 200

Probabilities: 0 -> 1 with interval 0.1

Total number of combinations: 121

Final Data Structures: 121 x 16 x 15 matrix and 121 x 4 x 15 matrix

Motifs

Classes

Results: Motif and class prevalence in 1 trial of 15

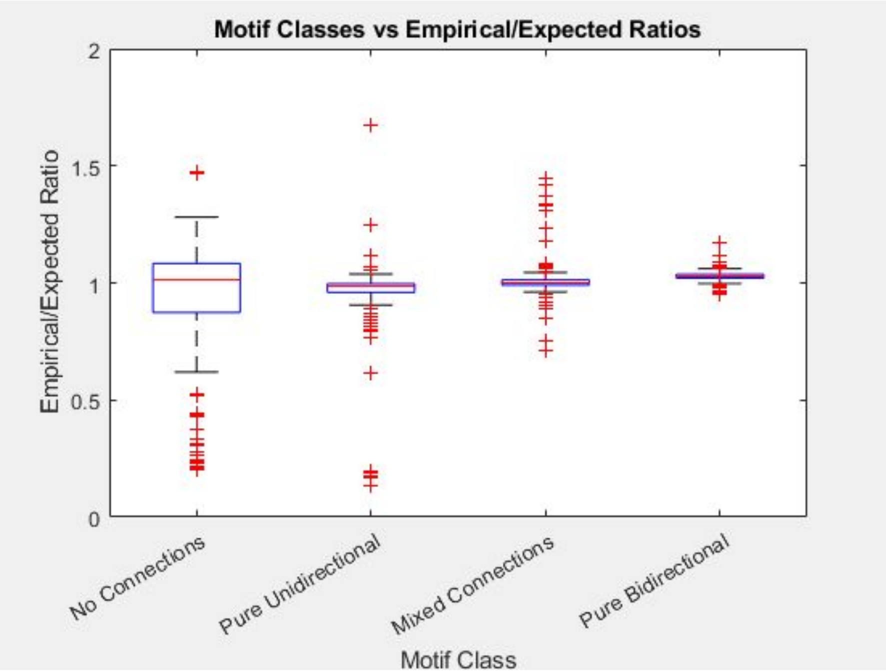


Figure: Motif Classes vs Empirical/Expected Ratios. The boxplot shows the empirical to expected ratios for different motif classes, which include: No Connections, Pure Unidirectional, Mixed Connections, Pure Bidirectional.

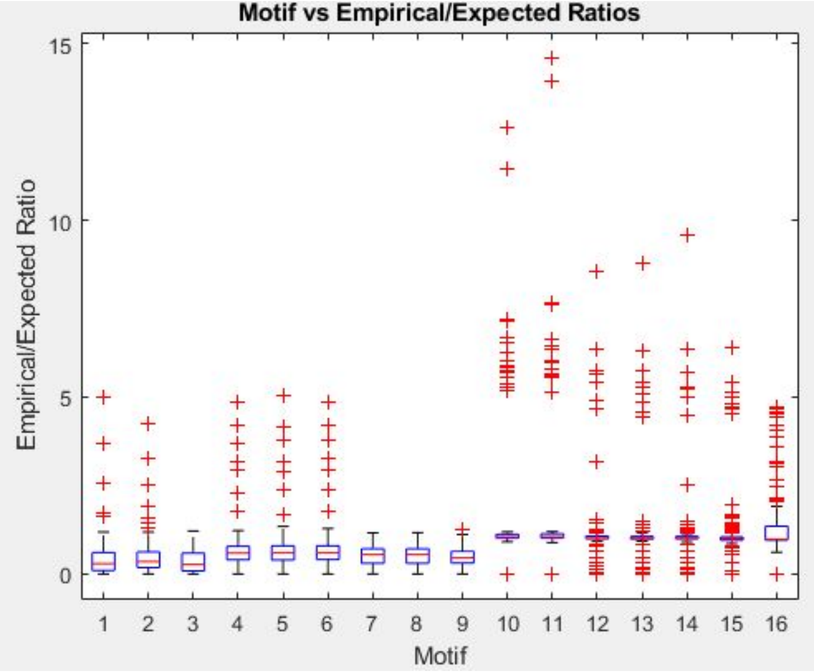


Figure: Motif vs Empirical/Expected Ratios. The boxplot illustrates the empirical to expected ratios for individual motifs numbered 1 through 16.

Results: Statistics of Class Distributions

```
%% Class Analysis

class_p_values = zeros(1,4);

% Prepare data for ANOVA
for classcol = 1:4
    % Extract data for the current column across all cells
    data = [];
    group = [];
    for cellIdx = 1:15
        columnData = class_cell{cellIdx}(:, classcol);
        data = [data; columnData];
        group = [group; repmat(cellIdx, size(columnData))];
    end

    % Perform ANOVA for the current column
    [p, tbl, stats] = anova1(data, group, 'off');

    class_p_values(classcol) = p;

end

% Display all p-values
disp('P-values for each class:');
disp(class_p_values);
```

```
P-values for each class:
    1.0000    1.0000    1.0000    1.0000
```

ANOVA shows that ratio of seen to expected distributions in motif classes is not stable. It is random

Results: Statistics of Motif Distributions

```
%% Motif Analysis

motif_p_values = zeros(1,16);

% Prepare data for ANOVA
for motifcol = 1:16
    % Extract data for the current column across all cells
    data = [];
    group = [];
    for cellIdx = 1:15
        columnData = motif_cell{cellIdx}(:, motifcol);
        data = [data; columnData];
        group = [group; repmat(cellIdx, size(columnData))];
    end

    % Perform ANOVA for the current column
    [p, tbl, stats] = anova1(data, group, 'off');

    motif_p_values(motifcol) = p;
end

% Display all p-values
disp('P-values for each motif:');
disp(motif_p_values);
```

```
P-values for each motif:
  Columns 1 through 8

    1.0000    1.0000    1.0000    1.0000    1.0000    1.0000    1.0000    1.0000

  Columns 9 through 16

    1.0000    1.0000    1.0000    1.0000    1.0000    1.0000    1.0000    0.5834
```

ANOVA shows that ratio of seen to expected distributions in motifs is not stable. It is random

Conclusions

- The study identifies patterns in motif prevalence and provides a statistical framework for analysis.
- We see that more complexly connected networks might be more likely, but no statistically significant results are found for this.
- Findings show that initial conditions do not elicit the same or similar result every time, as seen by the non-significance seen with ANOVA.
- As such, the hypothesis that initial conditions cause complex structures to be preferred is not supported.

Future Work:

- Shrinking scope to either large number of iterations or large numbers of initial probability pairs. Having both be large makes it difficult to identify false conclusions

Contributions

Pranav	Lyla
created the meshgrid	came up with a way to hold all of the connectivity matrices motif_ratio_matrix
generateConnectivityMatrix.m function	looped tripletMotifs.m to get all of the ratio_arrays
optimized the conditionals in the tripletMotifs.m code	graphed the ratios of motifs

Acknowledgements

Thank you Professor Venugopal and Ryan for providing feedback on how to improve our work.

Thank you to the classmates for challenging us with important questions and making us reflect on optimal analysis strategies.

Thank you!