Theory of the Snowflake Plot and Its Relations to Higher-Order Analysis Methods

Gabriela Czanner
gabriela@neurostat.mgh.harvard.edu
Neuroscience Statistics Research Laboratory,
Massachusetts General Hospital and Harvard Medical School,
Boston, MA 02114, U.S.A.

Sonja Grün
gruen@neurobiologie.fu-berlin.de
Institute for Biology–Neurobiology, Free University,
Berlin, 14195, Germany

Satish Iyengar
si@stat.pitt.edu
Department of Statistics, University of Pittsburgh
Pittsburgh, PA 15260, U.S.A.

The snowflake plot is a scatter plot that displays relative timings of three neurons. It has had rather limited use since its introduction by Perkel, Gerstein, Smith, and Tatton (1975), in part because its triangular coordinates are unfamiliar and its theoretical properties are not well studied. In this letter, we study certain quantitative properties of this plot: we use projections to relate the snowflake plot to the cross-correlation histogram and the spike-triggered joint histogram, study the sampling properties of the plot for the null case of independent spike trains, study a simulation of a coincidence detector, and describe the extension of this plot to more than three neurons.

1 Introduction

The snowflake plot (or joint configuration scatter diagram) of Perkel, Gerstein, Smith and Tatton (1975) is a scatter plot that displays the relative timings of firings from three simultaneously recorded neurons. It assumes that the relationships between the neurons can be fully described by the differences between their firing times. The plot uses a triangular coordinate system that treats the neurons symmetrically, and it is bounded by a hexagonal box—hence, its name. A point in the plot corresponds to the three differences between the firing times of the three neurons.

Perkel et al. (1975) studied the geometry of the plot and did many simulations to show the patterns in the snowflake plot under different dependence structures; however, there has been no rigorous theoretical study of the properties of this plot yet. One possible reason for the limited interest in the snowflake plot is that it is rather computer intensive; only recently have computers become powerful enough to allow quick construction of this plot and to do simulation studies to understand its properties better.

There are also concerns about whether the snowflake plot can be helpful in analyzing data. One reason is that it can contain artifacts that complicate interpretations and can make it more difficult to understand and use. Next, the triangular system used to plot the snowflake is not as familiar as the Cartesian system, so what patterns one should expect given a certain circuit of three neurons are not immediately clear. This is especially hard for the more complex circuits.

Nevertheless, the snowflake plot can be helpful (Perkel et al., 1975) as a potential screening device for higher-order correlations, which are hard to detect by other displays. An important example is coincidence detection (Perkel et al., 1975), in which a neuron fires in response to receiving (nearly) coincident spike inputs (Abeles, 1982b; see section 5 below). The simplest circuit that resembles a coincidence detector is a neuron that receives inputs from two neurons, and the threshold of the receiving neuron is such that two coincident input spikes suffice to reach threshold. Such a relation of two synchronous spikes followed (with some small delay) by an output spike cannot be detected by mere pairwise analysis, for example, using a cross-correlation histogram (CCH). Instead, methods are required that allow the study of more than two neurons at a time as provided in the snowflake plot.

How the snowflake plot works for excitatory or inhibitory connections has been studied using simulations (Perkel et al., 1975). If neuron A excites neuron B and if a third neuron C is independent of the others, then there will be an excess of cases when a spike from neuron A is followed by a spike from neuron B. This will be reflected in the snowflake plot as a narrow band of increased density of the points. Note, however, that for each pair of spikes from neurons A and B, we need a spike (although unrelated) from neuron C to draw the point in the snowflake. Thus, we need neuron C to fire persistently to bring out the relationship between A and B in the snowflake plot (Perkel et al., 1975). For this reason, the authors suggest the use of the snowflake plot in addition to the CCH and autocorrelation histogram.

A related development is the joint peristimulus time scatter diagram of Gerstein and Perkel (1969, 1972) and joint peristimulus time histogram (JPSTH) of Aertsen, Gerstein, Habib, and Palm (1989) for two neurons that are subject to repeated stimulation. The same concept is used by Prut et al. (1998) for detecting synchrony among three neurons. Here the trigger event is a spike emitted by a neuron instead of a stimulus. Prut et al. (1998) call the resulting matrix of raw counts a threefold correlation matrix, or the counts matrix instead of JPSTH. Since the term *counts matrix* is a rather

generic name and since it is a histogram, we will call it the spike-triggered joint histogram (STJH) instead. Finally, in their work, the three neurons are treated asymmetrically, with one neuron playing the role of a trigger and the other two neurons being the reference units. In that case, if two neurons fire after the third neuron fires and if we do not have prior knowledge about the connections between neurons, we may need three separate STJHs to detect the dependence structure. While this is not difficult for just three neurons, it can become cumbersome when studying many more neurons in groups of three.

Many of the ideas that we explore in this article about projection to explain the relationship between different plots and extensions to more than three neurons were qualitatively stated in earlier work (Perkel et al., 1975; Gerstein & Perkel, 1972; Kristan & Gerstein, 1970; Abeles & Gerstein, 1988). This article provides quantitative results that put the early qualitative work on a rigorous mathematical basis. After reviewing some basics in section 2, we show in section 3 that the CCH and the snowflake plot are results of orthogonal projections—the CCH from two- to one-dimensional space and the snowflake plot from three- to two-dimensional space. In section 4, we show that the snowflake plot is a generalization of the STJH in the sense that it treats neurons symmetrically because the STJH can be viewed as a result of a nonorthogonal projection from three- to two-dimensional space; finally, we show that the snowflake is the union of all three projections (STJHs of three neurons) so it gives us the same information as the three STJHs at once. Hence, the snowflake plot can bring simplification in computations. In section 5, we will study in detail the example of a simulated coincidence detector convergent circuit using the snowflake. In section 6, we derive analytically the distribution of the points on the snowflake plot under the null assumption that the neurons fire independently from three homogeneous Poisson processes, and we discuss its properties. In section 7 we discuss the geometry of a higher-dimensional version of the snowflake plot for four or more neurons and end with a discussion in section 8.

2 Snowflake Plot Basics

We begin with a brief account of the geometry and construction of the snowflake plot; this material is extracted from Perkel et al. (1975). Let A, B, and C denote the times of spikes from three neurons. Since $(B - A) + (C - B) + (A - C) = 0$, we can plot the point (A, B, C) on a plane using a triangular coordinate system (see point P in Figure 1A). The Cartesian coordinates of the plotted point P are

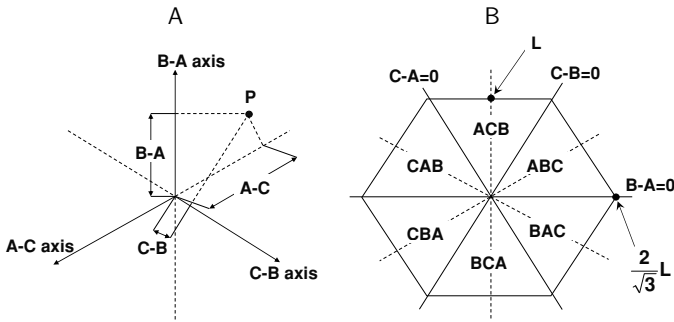$$\left(\frac{2C - (A + B)}{\sqrt{3}}, \ B - A \right). \tag{2.1}$$

Figure 1: Geometry of the snowflake plot. (A) The triangular coordinate system for the snowflake plot. (B) The coincidence lines (solid) divide the hexagon into six triangles. Each triangle corresponds to a particular set of timing sequences: the top triangle contains sequences *ACB*, then to the right *ABC*, *BAC*, *BCA*, *CBA*, and *CAB*.

Figure 1B shows again the axis of the triangular system. The lines that are perpendicular to them (solid) are called coincidence lines; for example, at the $B - A = 0$ line, neurons A and B fire at the same time.

The solid lines (see Figure 1B) divide the hexagon into six sectors. Each sector of the snowflake is characterized by the order of firings; for example, *ABC* refers to the sector representing triples in which A precedes B, which in turn precedes C. Furthermore, each point on the snowflake plot corresponds to a precise sequence of the three neurons A, B, and C (see Figure 1B) defined by the order of the three neurons and two time differences between them.

Following Perkel et al. (1975), we often consider only those triples (A, B, C) that satisfy

$$|B - A| < L \,, \ |C - B| < L \,, \ \text{and} \ |A - C| < L, \tag{2.2}$$

for some time interval, or *span L*, that is shorter than the recording period τ. The plot of such points will again form a snowflake; that is, they will all lie inside a smaller hexagon centered at the origin. Such a restriction is useful when the dependence among the neurons that is of interest is over small time intervals.

3 Snowflake Plot Generalizes the CCH

Aertsen et al. (1989) show the correspondence between the CCH and JPSTH. The CCH can be viewed as an orthogonal projection from the
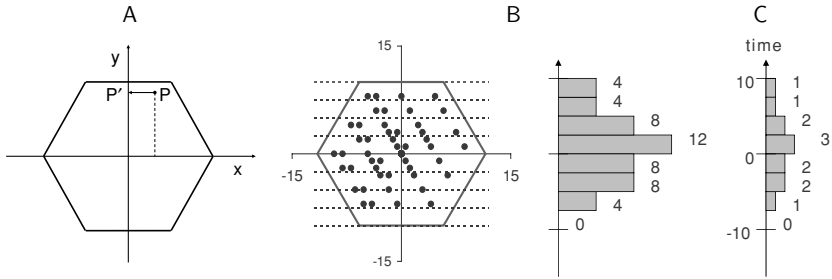
Figure 2: Orthogonal projection of points of the snowflake plot onto the vertical axis. (A) $P = (\frac{1}{\sqrt{3}}(2C - B - A), B - A)$, $P' = B - A$. (B) The snowflake plot and the histogram of all projected points. Neuron A fires at times 1, 5, 9; B fires at 2, 4, 7, 9; C fires at 1, 2, 5, 9. The total time is $\tau = 10$. (C) The raw CCH of A and B. The histogram from B and the raw CCH have the same shape; they are equal up to a constant multiple, $n_C = 4$, which is the number of times neuron C fired.

two-dimensional space spanned by times when neurons A and B fired onto the one-dimensional space spanned by the vector $(-1, 1)$—that is, onto the hyperplane $x + y = 0$—up to a scale constant $\frac{1}{\sqrt{2}}$. The projection is simply

$$B - A = (-1, 1) \begin{pmatrix} A \\ B \end{pmatrix}.$$

Similarly, the snowflake plot can be considered an orthogonal projection from the cube $[0, \tau]^3$ onto the plane $x + y + z = 0$:

$$\begin{pmatrix} \frac{1}{\sqrt{3}}(2C - B - A) \\ B - A \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{3}} \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix}.$$

Thus, the snowflake plot is a generalization of the CCH, from two to three neurons. This approach suggests that if we can find the joint spike density function in the cube, we can—at least in principle—then find the induced density of the projection on any hyperplane, for example, in the snowflake plot; in section 6, we use this approach to study the null case of three independent Poisson spike trains.

Moreover, we can project the snowflake plot onto, for example, the vertical axis (see Figure 2A) and create a histogram of the points in the projection. Then the number of points occurring in the interval $(t_1, t_2]$ is equal to the number of points in the cross-correlation histogram multiplied by n_C, the number of times neuron C fired (which is independent on t_1, t_2). The reason is that each possible difference $(B - A)$ is represented by n_C points in the
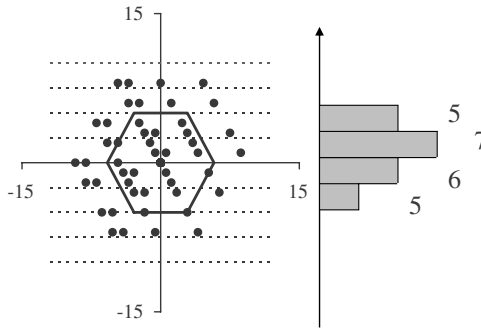
Figure 3: The projection when $L = 5 < \tau = 10$. Here we are interested only in the points within the span $L = 5$, which is illustrated by the hexagon. All points within the hexagon define the snowflake plot. All other points should not be plotted; however, for illustrative purposes, we draw them here. Neuron A fires at times 1, 5, 9; B fires at 2, 4, 7, 9; C fires at 1, 2, 5, 9; the total time is $\tau = 10$. This histogram is not equal to the raw CCH in Figure 2C.

snowflake plot. For illustration, see Figure 2B, where neuron C fired four times: the histogram of the projection is hence equal to raw counts CCH (of A and B) multiplied by 4.

However, the details are different when we work with the plot of points within the span L, with $L < \tau$. Consider again the projection of the points onto the vertical axes and a histogram of the points in that projection. Then the projection of the snowflake plot is not equal to the CCH (up to a constant multiple) (see Figure 3). The reason is that each difference $(B - A)$ is plotted in this smaller snowflake only if there is a close spike (within span L) from the third neuron C.

4 Three STJHs Embedded in the Snowflake Plot

Consider now the case of three neurons, where one of the three plays the role of a stimulus. Prut et al. (1998) used the idea of JPSTH for detecting the synchrony among three neurons. Here we will show that the snowflake plot can be divided into three parts, each corresponding to one STJH. This suggests that the plot can be considered as a generalization of the STJH by treating the three neurons symmetrically. This can be beneficial if we do not have prior information about which neuron is a trigger, for example, when many neurons are recorded and examined in groups of three.

The analysis for spatiotemporal patterns by Prut et al. (1998) was motivated by predictions of the synfire chain model (Abeles, 1982a, 1991): "In such a model, every time the chain is activated, the participating cells will fire in a sequential manner. The temporal spacing between the spikes will
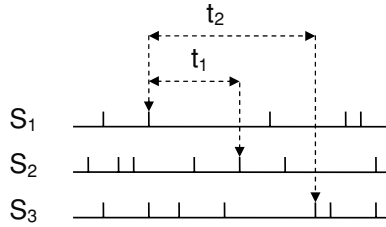
Figure 4: Precise firing sequence (PFS). Here, S_1, S_2, and S_3 are three different neurons.
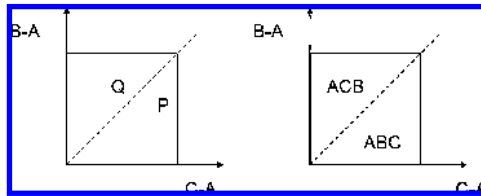

Figure 5: The geometry in STJH where neuron A is a trigger. For P, we have $C - A > B - A$; hence, $C > B$; hence, the PFS is ABC. For Q, we have $C - A < B - A$; hence, $C < B$; hence, the PFS is ACB.

correspond to the relative location of the parent cells along the chain. Such a sequence of spikes is termed a precise firing sequence (PFS)" Prut et al., 1998, pp. 2857–2858), Studies have detected PFSs in cortical activity of the frontal lobe (Abeles, Bergman, Margalit, & Vaadia, 1993; Villa & Fuster, 1992).

We will use Prut et al.'s notation. A precise firing sequence is defined by its unit composition and time delays between spikes. The three units are denoted by S_1, S_2, S_3. Unit S_1 fires first, unit S_2 fires t_1 time units after S_1 fires, and unit S_3 fires t_2 time units after S_1 fires, where $t_1 < t_2$ (i.e. $t_1 = S_2 - S_1$, $t_2 = S_3 - S_1$) (see Figure 4). So the PFS can be written as $(S_1, S_2, S_3; t_1, t_2)$. Call the first unit the trigger and the other two units reference units. Units S_1, S_2, S_3 are either three different neurons, or they can represent one neuron.

For three neurons $A, B,$ and C, assume that A is a trigger (i.e., $S_1 = A$). We can plot $C - A$ on the x-axis and $B - A$ on the y-axes (see Figure 5). All such points create a spike-triggered type of diagram (scatter plot). If we create a grid above it and count the number of points in each square, we get the STJH. Moreover, the triangle to the right of the main diagonal represents the PFSs: $(A, B, C; t_1 = B - A, t_2 = C - A)$ (since $C - A > B - A$ so $B < C$; see point P in Figure 5); analogously, the other triangle represents the PFSs: $(A, C, B; t_1 = C - A, t_2 = B - A)$ (see point Q in Figure 5). The PFSs also
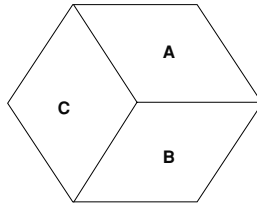
Figure 6: Solid lines divide the snowflake plot into three parts, each defined by a trigger unit (e.g., *B* means that *B* is a trigger).

divide the snowflake plot into triangles (see the six triangles, *ABC*, *ACB*, *BAC*, *BCA*, *CAB*, *CBA*, in Figure 1B). In this respect, the snowflake plot contains the same information as the three STJHs—where each STJH is defined by a different trigger neuron—at once.

We now show that there is a nice transformation from the snowflake plot onto the STJHs. Let us divide the snowflake plot into three parts (see parts A, B, and C in Figure 6). Each part is determined by the trigger neuron, the one that fired first in the sequence.

From the definition of the snowflake plot, there is a one-to-one correspondence between part A of the snowflake plot and the STJH from Figure 5. Note that each point in our STJH has coordinates $[x_{STJH}, y_{STJH}] = [C - A, B - A]$. Then

$$x_{SNOW} = \frac{1}{\sqrt{3}}[2(C - A) - (B - A)] = \frac{2}{\sqrt{3}}x_{STJH,A} - \frac{1}{\sqrt{3}}y_{STJH,A} \qquad (4.1)$$

and

$$y_{SNOW} = B - A = y_{STJH,A} \qquad (4.2)$$

since

$$\frac{1}{\sqrt{3}}(2C - B - A) = \frac{1}{\sqrt{3}}[(C - B) + (C - A)]$$

$$= \frac{1}{\sqrt{3}}[(C - A) - (B - A) + (C - A)]$$

$$= \frac{1}{\sqrt{3}}[2(C - A) - (B - A)],$$

where $(x_{STJH,A}, y_{STJH,A})$ are coordinates of the point in the STJH with *A* being a trigger, and (x_{SNOW}, y_{SNOW}) are coordinates of the same point in the
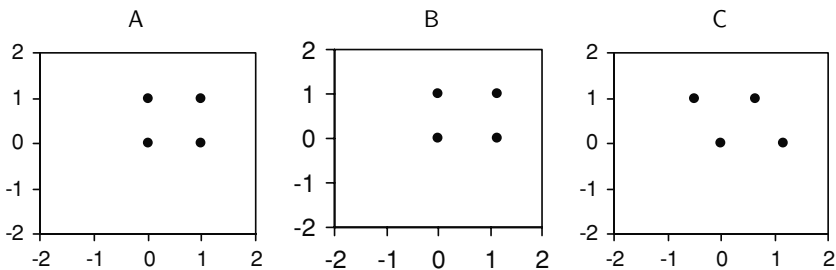
Figure 7: Transformation (see equation 4.3) of the points in the STJH to the snowflake. (A) Points $(0, 0)$, $(1, 0)$, $(0, 1)$ and $(1, 1)$. (B) The four points after scaling. (C) The four points after scaling and shearing.

snowflake plot. In the same way we can get a transformation between parts B, C, and corresponding STJHs.

Note that our transformation matrix (in equations 4.1 and 4.2) is

$$\begin{bmatrix} \frac{2}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{3}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix}, \tag{4.3}$$

the composition (product) of two affine transformations: first, a shear matrix that preserves the horizontal axis with shear factor $-1/2$, followed by a scaling (in this case, rescaling the horizontal axis by the factor $2/\sqrt{3}$). Hence, it is also an affine transformation.

Figure 7 provides an example. Figure 7A shows four points $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$. Figure 7B shows these points after applying the scaling, where the point $(1, 0)$ transforms into $(1.155, 0)$, because

$$\begin{bmatrix} \frac{2}{\sqrt{3}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{pmatrix} 1.155 \\ 0 \end{pmatrix}.$$

Similarly, the points $(0, 0)$, $(0, 1)$, $(1, 1)$ transform into $(0, 0)$, $(0, 1)$ and $(1.155, 1)$, respectively. Figure 7C shows the four points after applying the second transformation: it is a shearing, which transforms into $(1.155, 0)$. Also, the points $(0, 0)$, $(0, 1)$, $(1.155, 1)$ transform into $(0, 0)$, $(-0.5, 1)$, and $(0.655, 1)$, respectively.

In summary, each of the three STJHs is represented by one-third of a snowflake plot; hence, the three STJHs give exactly the same information as the snowflake plot.
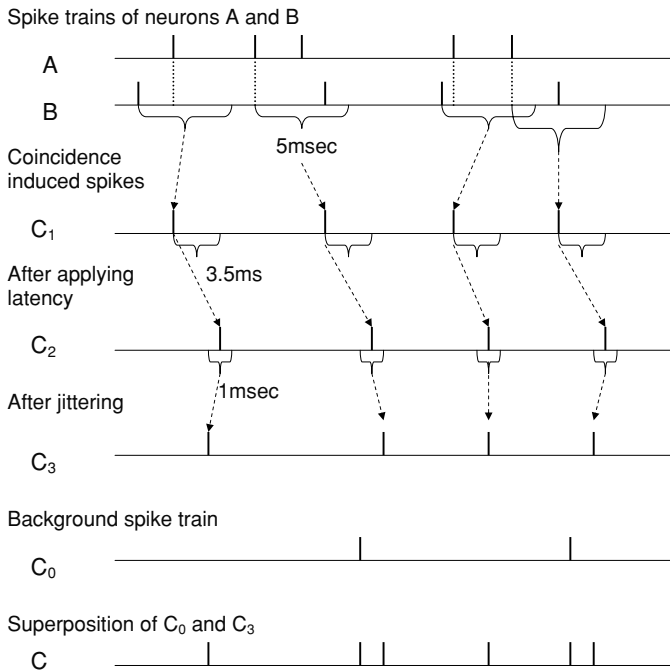
Spike trains of neurons A and B

Figure 8: Simulation of the circuit with the coincidence detector neuron.

5 Example

In this section, we show the use of the snowflake with an example: a simulated circuit with a coincidence detector neuron. Such a circuit has already been discussed by Perkel et al. (1975). However they used an integrate-and-fire (IF) model to simulate the data. We use a different approach: we inject the coincidence-induced spikes into the spike train of the coincidence detector, as described below and depicted in Figure 8.

We define a coincidence detector as a neuron (neuron C) that fires (with a latency 3 to 4 ms) if it sees an A spike followed by a B spike or vice versa within 5 ms. Also, we assume that the coincidence detector is not perfect; that is, it can fire (with some low probability) even when there is no coincidence.

In our simulations, we generated A and B independently as Poisson processes with the same intensity—15 spikes per second. Neuron C was made to fire (with a latency 3 to 4 ms) if it saw a near coincidence (within 5 ms) between A and B spikes. Finally, we injected (see Grün, Diesmann, & Aertsen, 2002; Kuhn, Aertsen, & Rotter, 2003) low-intensity background

spikes into the C spike train. (For more details about the simulation, see appendix A.)

The simulated data are shown in Figure 9. The coincidence detector is an example of a PFS in the following sense. This circuit can be described by a set of two kinds of PFSs: $(A, B, C; t_1, t_2)$, and $(B, A, C; t_3, t_4)$. In $(A, B, C; t_1, t_2)$, neuron A fires first; in $(B, A, C; t_3, t_4)$, neuron B fires first.

Figure 9 shows the raw CCHs, raw STJHs, and the corresponding snowflake plot. The first STJH (see Figure 9E) shows that the PFSs $(A, B, C; t_1, t_2)$ take place where $0 < t_1 < 5$ and $t_2 \in (t_1 + 3, t_1 + 4)$ ms. The second STJH (Figure 9F) says that $(B, A, C; t_3, t_4)$ also takes place with $t_3 < 5$ and $t_4 \in (t_3 + 3, t_3 + 4)$. The first STJH corresponds to part A (see Figure 6) of the snowflake plot (contains the upper part of the chevron), the second corresponds to part B (contains the lower part of the chevron), the third to part C. This example illustrates the correspondence among the three JSTHs and the snowflake plot.

6 The Null Distribution

In this section, we analyze the null case of the three neurons firing independently according to homogeneous Poisson processes, with possibly different rates. In practice, the null distribution can then be subtracted from data to detect deviations from it in the snowflake plot. In appendix C, we calculated the mean and variance of counts in bins of the cube $[0, \tau]^3$, which allows us to study the variability of the snowflake.

In appendix B, we show that the joint spike intensity function over the snowflake plot is $\lambda_A \lambda_B \lambda_C f_{X,Y}(x, y)$, where λ_A, λ_B, and λ_C are the intensities of neurons A, B, and C, respectively:

$$f_{X,Y}(x, y) = \begin{cases} \frac{\sqrt{3}}{2\tau^3}(\tau - y) & \text{in sector ACB} \\ \frac{\sqrt{3}}{2\tau^3}(\tau - \frac{\sqrt{3}}{2}x - \frac{1}{2}y) & \text{in sector ABC} \\ \frac{\sqrt{3}}{2\tau^3}(\tau - \frac{\sqrt{3}}{2}x + \frac{1}{2}y) & \text{in sector BAC} \\ \frac{\sqrt{3}}{2\tau^3}(\tau + y) & \text{in sector BCA} \\ \frac{\sqrt{3}}{2\tau^3}(\tau + \frac{\sqrt{3}}{2}x + \frac{1}{2}y) & \text{in sector CBA} \\ \frac{\sqrt{3}}{2\tau^3}(\tau + \frac{\sqrt{3}}{2}x - \frac{1}{2}y) & \text{in sector CAB,} \end{cases}$$

τ is the recording period, and (x, y) are the Cartesian coordinates used in equation 2.1.

The density $f_{X,Y}$ has a tent-like shape (see Figure 10A). To illustrate this result, we simulated three independent neurons with individual intensities of 10 spikes per second and recording period 15 seconds. The neurons fired
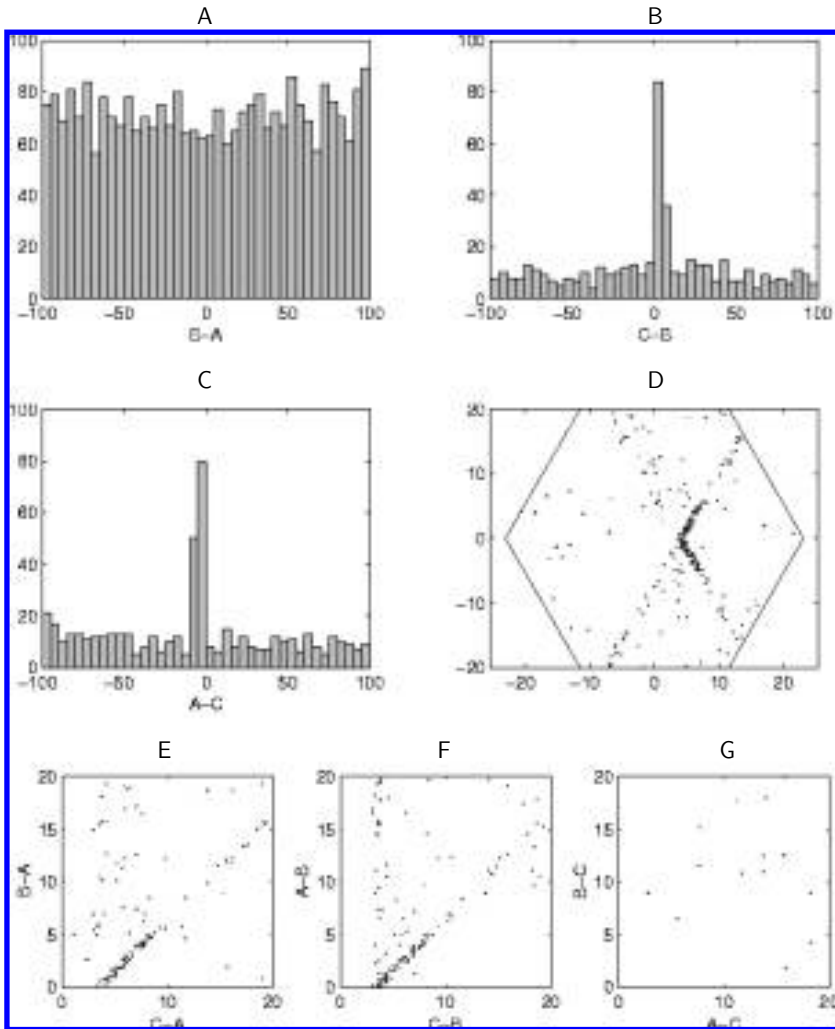
Figure 9: Simulated coincidence detector. Neurons A and B are from two independent Poisson processes with intensities 15 spikes per sec. If neuron C sees an A spike followed by a B spike within 5 ms, then C fires within 3 to 4 ms after the B spike. C also fires if it sees a B spike followed by an A spike (see appendix A). The total time of simulation is 1 minute; neuron A fired 897 times and neuron B fired 932 times. Neuron C fired 125 times, with 12 background and 113 coincidence-induced spikes. (A) Raw CCH of A and B. (B) Raw CCH of B and C. (C) Raw CCH of A and C. (D) Snowflake plot of the simulated data. (E) Spike-triggered scatter diagram of B and C (A is a trigger). STJH can be obtained as a histogram above the plot. (F) Spike-triggered scatter diagram of A and B (C is a trigger). (G) Spike-triggered scatter diagram of A, and B (C is a trigger).
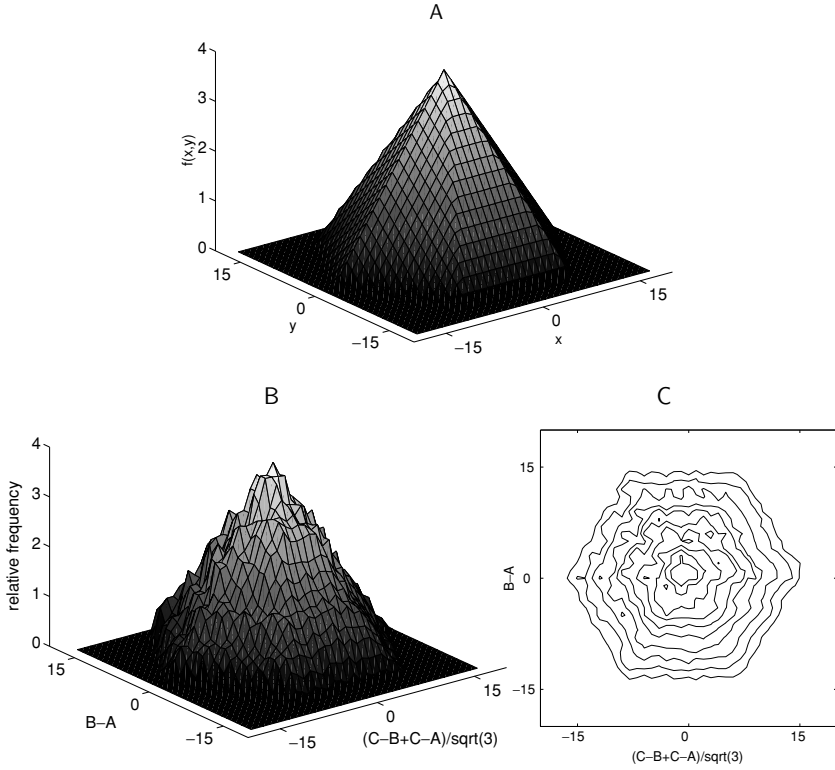
Figure 10: (A) The density $f_{X,Y}$ for the independence case with $\tau = 15$ sec. The value at the center is 0.003849. (B) The 3D histogram for the independence case with $\tau = 15$ sec. Data are obtained from a simulation of three independent Poisson processes with intensities of 10 spikes per second. The neurons A, B, and C fired 146, 147, and 152 times, respectively. Bin size is 1 ms^2. The height of each bin is equal to number of points in the snowflake plot divided by the total number of points, $146 \times 147 \times 152$. The value at the center is 0.003774. (C) Contour plot of the 3D histogram in B.

146, 147, and 152 times. We created bins over the snowflake plot and plotted the relative counts (See the histogram and contour plot in Figures 10B and 10C). We also derived the joint spike intensity function over the span L,

$$f_{X,Y/L}(x, y) = \lambda_A \lambda_B \lambda_C \, \frac{3L^3}{\tau^3} \left(\frac{\tau}{L} - \frac{2}{3} \right) f_{X,Y}(x, y) \, I\{(x, y) \in \text{span L}\},$$

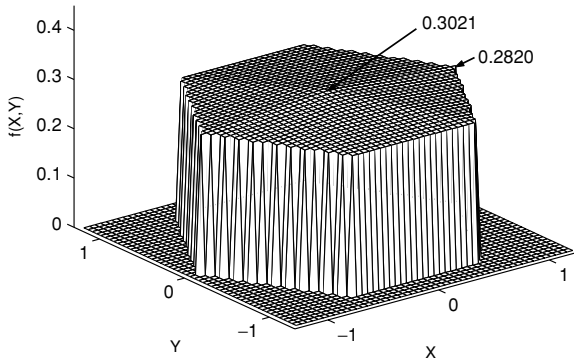where I is the indicator function. Figure 11 shows the intensity function for

Figure 11: The density function $f_{X,Y/L}$ for the independence case; $\tau = 15$ sec, $L = 1$ sec.

the case of L = 1 sec and total recording time 15 sec. It is not constant, and it also has a maximum at origin. However, the shorter L is or the larger τ is, the more nearly constant the density appears near origin. This explains the observation in Perkel et al. (1975) that the distribution of the points in the snowflake plot should have approximately constant density over the hexagon.

To assess the applicability of this null distribution result to trains other than the Poisson, we also simulated three independent spike trains, this time with gamma (α, β) distributed interspike intervals, of orders $\alpha = 2$ and 3 (the Poisson train has intervals that are gamma with order 1), and scale parameter $\beta = 100/\alpha$ so that the mean interspike interval for all three cases is 100 ms. The three-dimensional histograms and the corresponding contour plots are given in Figure 12. Note that both resemble those from the Poisson process. Note that as the order increases, the contours appear smoother; this may reflect the fact that the coefficient of variation of the gamma, $1/\sqrt{\alpha}$, decreases as α increases.

Furthermore, our theoretical calculations in appendix C show that for Poisson spike trains, the expected number of points in any small cube contained in $[0, \tau]^3$ depends on intensities only through their product $\lambda_A \lambda_B \lambda_C$, which is an overall measure of the density of points in the cube, and that the variance depends on the intensities separately. We also show in appendix C that for a fixed value of this product, the variance of the points in any small cube is minimized if the three intensities are equal. Because the snowflake plot is a linear projection of the points in the cube onto a plane, these qualitative properties are inherited by the plot. This supports the observation in Perkel et al. (1975) that the points in the snowflake appear more random over the hexagon if the neurons' spiking intensities are equal. It also suggests that large differences between intensities cause larger variability across samples.
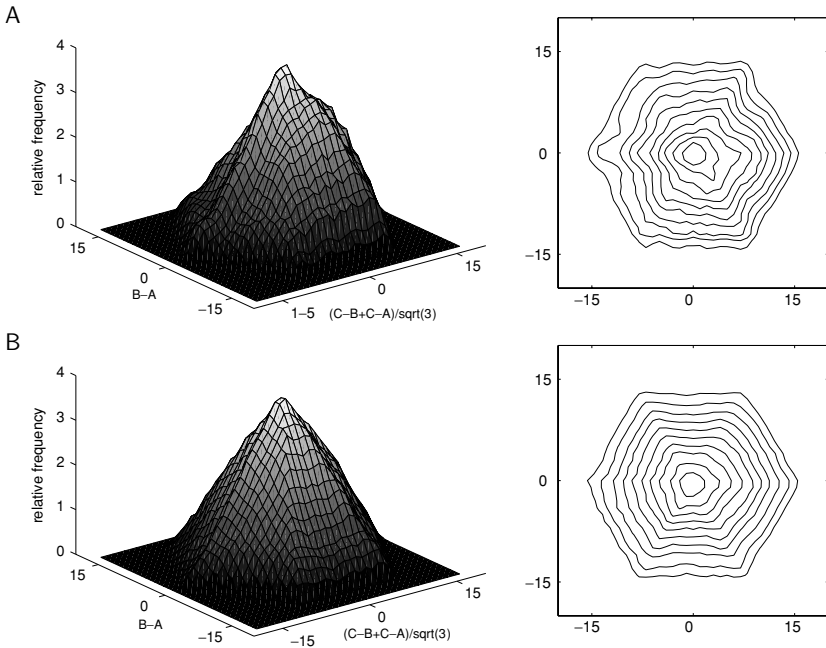
A

B

Figure 12: Two simulations of three independent gamma trains and the corresponding 3D histograms and contour plots on the snowflake plot. Total simulation time is $\tau = 15$ sec. The two simulations have different parameters of scale and order; however, they give the same mean interspike interval, 100 ms (as was also the case in Figure 10). (A) Trains simulated from gamma (2,100/2). The neurons fired 159, 148, and 136 times. (B) Trains simulated from gamma (3,100/3). The neurons fired 152, 148, and 138 times.

7 Geometry for More Than Three Neurons

Simultaneous recordings of more than three neurons are increasingly common; see, for example Beggs and Plenz (2004), Buzsaki (2004), Hoffman and McNaughton (2002), and Warren, Fernandez, and Normann, (2001). Hence, we explored the idea of generalizing the snowflake for more neurons. In this section, we show how the data from k neurons can be projected onto a $(k-1)$-dimensional space, where $k \geq 4$, assuming that relative spike timings contain all information about the dependencies among neurons. Although we cannot visualize the projection (except for four neurons), an analysis that is similar to the three-neurons case is possible. In particular, one can derive the joint density in the projection and subtract it from the projected data to search for patterns. This can be especially useful for exploratory analysis.

In the case of three neurons (i.e., in the snowflake plot), the angle between the axes of triangular system is $\arccos(-1/2) = 120°$. From section 3, we have that the snowflake plot is the result of projecting (up to a constant, $1/\sqrt{2}$) the cube onto the plane $x + y + z = 0$ using the projection matrix given there. For the case of k neurons, the projection matrix has vectors

$$
\begin{array}{cccccccc}
\sqrt{2}(1, & 1, & 1, & \ldots, & 1, & & 1, & -(k-1))/\sqrt{k(k-1)} \\
\sqrt{2}(1, & 1, & 1, & \ldots, & 1, & -(k-2), & & 0)/\sqrt{(k-2)(k-1)} \\
& & & \cdots & & & & \\
(1, & 1, & 1, & -3, & 0, & & \ldots, & 0)/\sqrt{6} \\
(1, & 1, & -2, & 0, & 0, & & \ldots, & 0)/\sqrt{3} \\
(1, & -1, & 0, & 0, & 0, & & \ldots, & 0).
\end{array}
$$

Once again, they are perpendicular, and their norm is again $\sqrt{2}$. These vectors, together with vector $\sqrt{2}(1, 1, 1, \ldots, 1)/\sqrt{k}$, form a Helmert matrix (see Mardia, Kent, & Bibby, 1979) up to a constant $\sqrt{2}$.

Note that the point $(1, 1, \ldots, 1, 1, 0)$ projects onto

$$
\sqrt{2}\left(\frac{(k-1)}{\sqrt{k(k-1)}}, 0, \ldots, 0\right),
$$

and $(1, 1, \ldots, 1, 0, 1)$ projects onto

$$
\sqrt{2}\left(\frac{-1}{\sqrt{k(k-1)}}, \frac{(k-2)}{\sqrt{(k-1)(k-2)}}, 0, \ldots, 0\right).
$$

In general, the points $(1, 1, \ldots, 1, 1, 0)$ and $(1, 1, \ldots, 1, 0, 1)$ represent the $(k-1)$-way coincidences; that is, some $(k-1)$ of the neurons fire simultaneously, and the remaining one fire at some different time. The projections of these points give us information about the coincidence lines in the projection. For example, the angle between these two coincidence lines is

$$
\arccos\left(\frac{-\frac{k-1}{k(k-1)/2}}{\frac{k-1}{\sqrt{k(k-1)/2}}\frac{\sqrt{1+k(k-2)}}{\sqrt{k(k-2)/2}}}\right) = \arccos\left(-\frac{1}{k-1}\right).
$$

So for $k = 3$ neurons, the angle between coincidence lines is $\arccos(-1/2) = 120°$; for $k = 4$, $\arccos(-1/3) = 109.47°$; and for large k,

$$
\lim_{k \to \infty} \arccos(-1/(k-1)) = \frac{\pi}{2};
$$

thus, the angle between coincidence lines becomes perpendicular as $k \to \infty$.

Finally, we show how to derive the axes for the case of four neurons. Note that the coincidence line ABC must be perpendicular to the plane where the values $B - A$ are varying; also, the coincidence line ABD must be perpendicular to the plane where values $B - A$ are varying. This gives us enough information to find the axes $B - A$ as being perpendicular to the two coincidence lines:

$$
\begin{aligned}
v_{B-A} &= (\qquad 0, \qquad 0, \quad 1.414) \\
v_{A-C} &= (\qquad 0, -1.225, -0.707) \\
v_{D-A} &= (\ 1.155, \quad 0.408, \quad 0.707) \\
v_{C-B} &= (\qquad 0, \quad 1.225, -0.707) \\
v_{B-D} &= (-1.155, -0.408, \quad 0.707) \\
v_{D-C} &= (\ 1.155, -0.815, \qquad 0).
\end{aligned}
$$

Notice that $v_{A-C} = -(v_{B-A} + v_{C-B})$, so the vectors $v_{B-A}, v_{C-B}, v_{A-C}$ lie in a common plane ($x = 0$) and have an equal angle, $120°$. Hence the projection onto the plane $x = 0$ is again the (snowflake) hexagon. Furthermore, it can be shown that angles among $v_{B-A}, v_{A-D}, v_{D-B}$ are $120°$; among v_{A-C}, v_{C-D}, and v_{D-A}, they are $120°$. Finally, the angle between v_{B-A} and v_{D-C} is $90°$; the same holds for v_{A-C} and v_{B-D} and for v_{D-A} and v_{C-B}.

8 Discussion

In this article, we have begun a quantitative study of the snowflake plot of Perkel et al. (1975). Our findings can be summarized as follows.

First, we used the unifying idea of projections of spike trains to show how the snowflake plot relates to the cross-correlogram (CCH) and the joint peristimulus time histogram (JPSTH). We treated the CCH and the snowflake plot as results of orthogonal projections—the CCH from two-onto one-dimensional space, the snowflake plot from three- onto two-dimensional space. We also showed how the snowflake plot generalizes the spike-triggered joint histogram (STJH) symmetrically: The snowflake is the union of all three STJHs of the three neurons.

Second, we analytically derived the distribution of points on the snowflake plot under the assumption that the neurons fire independently from homogeneous Poisson processes. The joint spike density function over the snowflake plot has a tent-like shape. However, if we are interested in the firings that lie within some span, then the analytical density of the points in the snowflake converges to a constant density as total observation time increases and span decreases.

Third, we studied the geometry of the snowflake for more than three neurons. Although one cannot easily visualize the projections, except for four neurons, an analysis similar to that for three neurons can be done. One

can derive the joint density in the projection in the null case and subtract it from the projected data to search for patterns. This can be useful for exploratory analysis to analyze several spike trains simultaneously.

How does the snowflake help us in analyzing data? It can be used for screening purposes to find potential PFSs and as a graphical device to visualize the PFSs. The idea is to create a grid above the snowflake plot and construct a histogram. The advantage is that all three STJHs (i.e., the raw counts matrices in Prut et al., 1998) are in one histogram, but the disadvantage is that the display is triangular rather than orthogonal.

However, to find the raw counts is only the beginning of the process of detecting spatiotemporal patterns. One then needs to calculate the corrected counts (corrected for the expected counts) and derive a test to identify them as significant events. Prut et al. (1998) developed a method that allows the identification of significant events, which is based on counting. However, there is still a long way to go in order to evaluate significance, since higher-order correlations need to be evaluated, and therefore one needs to correct for by chance coincidence of lower-order correlations (see Schneider & Grün, 2003; Nakahara & Amari 2002; Gütig, Aertsen, & Rotter, 2003). If the interest is only to identify deviations from full independence, one can apply the same procedure as in the unitary events analysis (see the joint surprise in Grün et al., 2002). Analytical work will be of limited value because significance calculations for such models are quite complicated; it is more likely that simulation approaches that use shuffled or jittered trains (see Brown, Kass, & Mitra, 2004) will be needed, but these must be executed with care (Gerstein, 2004).

There are other avenues of research besides of the evaluation of significance in the snowflake plot. A thorough study about the artifact of the snowflake plot—the woven pattern (Perkel et al., 1975)—is needed. It is caused by multiple entries, where one firing of a neuron is represented by multiple points in the plot. The artifact is always present, but it is less visible if the span L is small. Also of interest would be to generalize the independence assumption and find the distribution of the points in the snowflake plot. Another generalization can be made in modifying the Poisson assumption. Furthermore, the derivations of the joint spike density can be generalized for more than three neurons.

Appendix A: Simulation of a Coincidence Detector

To simulate the homogeneous Poisson process with intensity λ, one can use a direct method. Consider a spike at time t, $t \in [0, \tau]$. Then we can draw a random number, E, from exponential distribution with mean time interval $1/\lambda$. This defines the next spike at time $t + E$. This method can give the times of spikes with any precision in principle.

An equivalent way of simulating a Poisson process is to first draw the total number, n, of spikes from Poisson distribution with parameter $\lambda\tau$.

Then we can draw a random sample of size n from uniform distribution over the interval $[0, \tau]$. Finally, we sort the sample in increasing order to get the Poisson process.

We took the following steps to simulate a coincidence detector based on Poisson trains (see also Figure 8):

1. We simulated neurons A and B as two independent realizations of Poisson processes with intensities $\lambda_A = \lambda_B = 15$ spikes per second. In order to simulate the Poisson process, we used the second algorithm above.

2. We created the spike train C_1 by looking for coincidences between A and B. If an A spike is followed by a B spike (within 5 ms), we create a C_1 spike at the time of the B spike. Analogously, if a B spike is followed by an A spike (within 5 ms), we create a C_1 spike at the time of the A spike.

3. We applied a latency of 3.5 ms to each C_1 spike to get the C_2 spike train.

4. The train C_3 is a noisy version of C_2. It is obtained by jittering each spike of C_2 independently and uniformly over a time window of ± 0.5 ms around its original position.

5. We simulated the background process C_0 as Poisson process with intensity 0.2 spike per second, independent of the spike trains A and B. This will be the background process for neuron C.

6. We injected the spikes of C_0 into the train C_1 to create the spike train C. In the theory of point processes, this is also called the superposition of point processes.

Appendix B: Derivation of Null Distribution of Snowflake Plot

We now derive the joint density function $f_{X,Y}(x, y)$ (see section 6) of the points on the snowflake plot when the three neurons are firing independently according to homogeneous Poisson processes with possibly different rates.

We showed in section 3 that the snowflake plot is an orthogonal projection of the observations from the cube $[0, \tau]^3$ onto the plane $x + y + z = 0$. Hence, we first compute the joint spike density function in the cube; then we derive the induced probability distribution of the projection.

We first recall the following fact about a homogeneous Poisson process with n points in the interval $[0, \tau]$ (Karlin & Taylor, 1975): conditional on n, the points have the same distribution as uniform order statistics of size n in the interval $[0, \tau]$. Hence, conditional on the number of spikes in each of the three spike trains, the distribution of the points in the cube is the

same as the distribution of a random vector (U_1, U_2, U_3), where U_1, U_2, U_3 are independent and identically distributed with uniform distribution on $[0, \tau]$. We next project this distribution onto the plane $x + y + z = 0$.

Now construct the random vector

$$(X, Y, Z) = \left(\frac{2}{\sqrt{3}}\left[U_3 - \frac{U_1 + U_2}{2}\right], U_2 - U_1, U_3\right). \tag{B.1}$$

To find the distribution of the snowflake plot (X, Y), we first find the distribution of (X, Y, Z) and then integrate out Z. The inverse transformation of equation B.1 is

$$(U_1, U_2, U_3) = (-X\sqrt{3}/2 - Y/2 + Z, -X\sqrt{3}/2 + Y/2 + Z, Z),$$

which has Jacobian determinant $= -\sqrt{3}/2$; hence, the joint probability distribution function of (X, Y, Z) is

$$f_{XYZ}(x, y, z) = \frac{\sqrt{3}}{2\tau^3}I\{0 \le -x\sqrt{3}/2 - y/2 + z \le \tau, 0 \le -x\sqrt{3}/2 + y/2$$
$$+ z \le \tau, 0 \le z \le \tau\},$$

where I is the indicator function: $I\{0 \le x \le \tau\} = 1$ if $0 \le x \le \tau$ and zero otherwise. Next, we integrate out z from $f_{XYZ}(x, y, z)$. This is equivalent to finding the intersection of the three intervals $[x\sqrt{3}/2 + y/2, \tau + x\sqrt{3}/2 + y/2]$, $[x\sqrt{3}/2 - y/2, \tau + x\sqrt{3}/2 - y/2]$, and $[0, \tau]$ for each combination of (x, y). Now fix a sector: for example, if $(x, y) \in$ sector ACB, then $0 \le y \le \tau$, $x\sqrt{3} \le y$, $-x\sqrt{3} \le y$. Since $0 \le y$, we have

$$\left[x\sqrt{3}/2 + y/2, \tau + x\sqrt{3}/2 + y/2\right] \bigcap \left[x\sqrt{3}/2 - y/2, \tau + x\sqrt{3}/2 - y/2\right]$$
$$= \left[x\sqrt{3}/2 + y/2, \tau + x\sqrt{3}/2 - y/2\right].$$

Next, since $-x\sqrt{3} \le y$ and $x\sqrt{3} \le y$, we have $x\sqrt{3}/2 + y/2 \ge 0$ and $\tau + x\sqrt{3}/2 - y/2 \le \tau$. Hence,

$$\left[x\sqrt{3}/2 + y/2, \tau + x\sqrt{3}/2 - y/2\right] \bigcap [0, \tau]$$
$$= \left[x\sqrt{3}/2 + y/2, \tau + x\sqrt{3}/2 - y/2\right].$$

Finally, the length of the last interval is $\tau - y$, so $f_{XY}(x, y) = \frac{\sqrt{3}}{2\tau^3}(\tau - y)$ in the sector ACB. The proof for the other sectors is similar.

Appendix C: Null Mean and Variance of Counts in Cubes

Suppose that the three neurons are firing according to independent Poisson processes with intensities λ_A, λ_B, and λ_C. Denote the firing times of the three neurons by (A_i, B_j, C_k), with $1 \leq i \leq n_A, 1 \leq j \leq n_B$, and $1 \leq k \leq n_C$. Next, define the statistic

$$N\{(a, b] \times (c, d] \times (e, f]\}$$

$$= \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} \sum_{k=1}^{n_C} I\{(A_i, B_j, C_k) \in (a, b] \times (c, d] \times (e, f])\}$$

$$= \sum_{i=1}^{n_A} I\{A_i \in (a, b]\} \sum_{j=1}^{n_B} I\{B_j \in (c, d]\} \sum_{k=1}^{n_C} I\{C_k \in (e, f]\},$$

which counts the number of points in the cube $(a, b] \times (c, d] \times (e, f]$. Thus, conditional on the number of spikes, this count is the product of three independent binomial variates, say, $X \sim Bin(n_A, \frac{b-a}{\tau})$, $Y \sim Bin(n_B, \frac{d-c}{\tau})$, and $Z \sim Bin(n_C, \frac{f-e}{\tau})$. Thus, the conditional mean of $N = N\{(a, b] \times (c, d] \times (e, f]\}$ is

$$E(N|n_A, n_B, n_C) = \frac{1}{\tau^3} n_A n_B n_C (b - a)(d - c)(f - e),$$

and since n_i are independent Poisson variates with means τn_i for $i = A, B, C$, the unconditional mean is $E(N) = \lambda_A \lambda_B \lambda_C (b - a)(d - c)(f - e)$. Thus, the conditional mean depends on the number of spikes n_A, n_B, n_C through their product only, and the unconditional mean depends on intensities $\lambda_A, \lambda_B, \lambda_C$ through their product only. In both cases, they are proportional to the volume of the box $(a, b] \times (c, d] \times (e, f]$.

Next, the computation and study of the conditional variance of N is rather lengthy, so we sketch the details here and refer to Czanner (2004) for the technical details. The conditional variance is

$$\text{var}(N|n_A, n_B, n_C) = \text{var}(XYZ) = E\left[(XYZ)^2\right] - [E(XYZ)]^2.$$

To simplify notation, suppose that $b - a = d - c = f - e = h$, so the box is a cube. Using the independence of X, Y, and Z and binomial moments, we

get $\mathrm{var}(N|n_A, n_B, n_C)$ is

$$\frac{n_A n_B n_C h^3 (1-h)^3}{\tau^6} \left[1 + \frac{h}{1-h}(n_A + n_B + n_C) \right.$$
$$\left. + \left(\frac{h}{1-h}\right)^2 (n_A n_B + n_B n_C + n_C n_A) \right],$$

which depends on n_A, n_B, n_C separately, not just on their product.

This conditional variance is smallest if $n_A = n_B = n_C$ (assuming $(b-a) = (d-c) = (f-e) = h$ and $n_A n_B n_C = K_0$ where K_0 is a given positive constant that measures the overall density of points in the cube). First, notice that the variance has the form

$$H(n_A, n_B, n_C) = n_A n_B n_C \left[K_1 + K_2(n_A n_B + n_B n_C + n_C n_A) \right.$$
$$\left. + K_3(n_A + n_B + n_C) \right],$$

for some positive constants K_1, K_2, K_3 when the counts are all nonzero and h is not 0 or 1. To minimize H subject to the constraint $n_A n_B n_C = K_0$, we treat the counts as continuous variables and use Lagrange multipliers to show that the minimum occurs when $n_A = n_B = n_C$.

Finally, to get the unconditional variance, we use the fact that for any two random variables or vectors U and V, $\mathrm{var}[U] = E(\mathrm{Var}[U|V]) + \mathrm{var}(E[U|V])$. Now if we let $U = N$ and $V = (n_A, n_B, n_C)$, and use the moments of the Poisson distribution, we get that the unconditional variance of N is

$$\mathrm{var}(N) = \lambda_A \lambda_B \lambda_C [K_4 + K_5(\lambda_A + \lambda_B + \lambda_C) + K_6(\lambda_A \lambda_B + \lambda_A \lambda_C + \lambda_B \lambda_C)]$$

where K_4, K_5, and K_6 are all positive constants depending on τ and h only. Note that the unconditional variance has the same form as the conditional variance, so that for a given value of the product $\lambda_A \lambda_B \lambda_C$, the unconditional variance is minimized when $\lambda_A = \lambda_B = \lambda_C$.

Acknowledgments

References

Abeles, M. (1982a). *Local cortical circuits: An electrophysiological study*, New York: Springer.

Abeles, M. (1982b). Role of cortical neuron: Integrator or coincidence detector? *Israel J. Medical Science, 18*, 83–92.

Abeles, M. (1991). *Corticonics: Neural circuits of the cerebral cortex*. Cambridge: Cambridge University Press.

Abeles, M., Bergman, H., Margalit, E., & Vaadia, E. (1993). Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *J. Neurophysiology, 70*, 1629–1638.

Abeles, M., & Gerstein, G. L. (1988). Detecting spatiotemporal firing patterns among simultaneously recorded single neurons. *J. Neuroscience, 60*, 909–924.

Aertsen, A. M. H. J., Gerstein, G. L., Habib, M. K., & Palm, G. (1989). Dynamics of neuronal firing correlation: Modulation of "effective connectivity." *J. Neurophysiology, 61*, 900–917.

Beggs, J. M., & Plenz D. (2004). Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures. *J. Neuroscience, 24*, 5216–5229.

Brown, E. N., Kass, R. E., & Mitra, P. P. (2004). Multiple neural spike train data analysis: State-of-the-art and future challenges. *Nature Neuroscience, 7*, 456–461.

Buzsaki, G. (2004). Large-scale recording of neuronal ensembles. *Nature Neuroscience, 7*, 446–451.

Czanner, G. (2004). *Applications of statistics in neuroscience*. Unpublished doctoral dissertation, University of Pittsburgh.

Gerstein, G. (2004). Searching for significance in spatiotemporal firing patterns. *Acta Neurobiologiiae Experimentalis, 64*, 203–207.

Gerstein, G. L., & Perkel, D. H. (1969). Simultaneously recorded trains of action potentials: Analysis and functional interpretations. *Science, 164*, 828–830.

Gerstein, G. L., & Perkel, D. H. (1972). Mutual temporal relationships among neuronal spike trains: Statistical techniques for display and analysis. *Biophysical Journal, 12*, 453–473.

Grün, S., Diesmann, M., & Aertsen, A. (2002). Unitary events in multiple single-neuron spiking activity: I. Detection and significance. *Neural Computation, 14*, 43–80.

Gütig, R., Aertsen, A., & Rotter, S. (2003). Analysis of higher-order neuronal interactions based on conditional inference. *Biological Cybernetics, 88*, 352–359.

Hoffman, K. L., & McNaughton, B. L. (2002). Coordinated reactivation of distributed memory traces in primate neocortex. *Science, 297*, 2070–2073.

Karlin, S., & Taylor, H. M. (1975). *A first course in stochastic processes*. New York: Academic Press.

Kristan W. B., & Gerstein G. L. (1970). Plasticity of synchronous activity in a small neural net. *Science 169*, 1336–1339.

Kuhn, A., Aertsen, A., & Rotter, S. (2003). Higher-order statistics of input ensembles and the response of simple model neurons. *Neural Computation, 15*, 67–101.

Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis*. New York: Academic Press.

Nakahara, H., & Amari, S. (2002). Information-geometric measure for neural spikes. *Neural Computation, 14*, 2269–2316.

Perkel, D. H., Gerstein, G. L., Smith, M. S., & Tatton, W. G. (1975). Nerve-impulse patterns: A quantitative display technique for three neurons. *Brain Research, 100*, 271–296.

Prut, Y., Vaadia, E., Bergman, H., Haalman, I., Slovin, H., & Abeles, M. (1998). Spatiotemporal structure of cortical activity: Properties and behavioral relevance. *J. Neurophysiology, 79*, 2857–2874.

Schneider, G., & Grün, S. (2003). Analysis of higher-order correlations in multiple parallel processes. *Neurocomputing, 52–54*, 771–777.

Villa, A. E., & Fuster, J. M. (1992). Temporal correlates of information processing during visual short-term memory. *Neuroreport, 3*, 113–116.

Warren, D. J., Fernandez, E., & Normann, R. A. (2001). High-resolution two-dimensional spatial mapping of cat striate cortex using a 100-microelectrode array. *Neuroscience, 105*, 19–31.