# Representation of Cooperative Firing Activity Among Simultaneously Recorded Neurons

# G. L. GERSTEIN AND A. M. H. J. AERTSEN

Department of Physiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

# SUMMARY AND CONCLUSIONS

1. Simultaneous and separable extracellular recording of substantial populations of neurons under chronic and behavioral conditions is becoming experimentally feasible. We have recently described a conceptual transformation of such multiple spike train data that allows the experimenter to analyze the entire network of observed neurons as an entity rather than as a summation of neuron pairs. The basic transformation represents each of N neurons as a particle in an N-space. Each particle is given a "charge" that is related to the spike train of the corresponding neuron. The resulting forces on the N particles cause aggregation of those particles that represent neurons with time-related firing. The present paper extends the visualization and possibilities of this way of analyzing properties of neuronal assemblies. Data are taken from computer-simulated neuronal networks in order to provide known properties.

2. We demonstrate projection of particle positions from the N space to a plane. Under the right conditions the spatial arrangement of the particles forms a Venn diagram of functional relationships in the entire neural network.

3. We introduce revised force rules in the transformation that allow detection and study of inhibitory connections among the observed neurons. Sensitivity is lower than for excitatory connections.

4. We introduce revised "charge" rules that improve "signal-to-noise" properties and in addition allow inference of directed connectivity. The original transformation only allows identification of neurons with time-related firing. The two-charge transformation allows explicit identification of presynaptic and postsynaptic neurons.

5. Finally we examine sensitivity of the transformation to individual and near-coincident firing rates. Some criteria are presented for choice of charge normalization rules in the transformation.

# INTRODUCTION

A number of neurophysiological laboratories have recently turned their attention to the direct observation of neuronal assembly activity (a very partial list includes Refs. 3, 4, 9, 11, 16, 18). Such multineuron experiments involve the simultaneous recording of many spike trains and subsequently require an interpretation of the data in terms of neuronal assembly properties. The principal approach to the interpretation phase of such work until now has been based on computation of crosscorrelograms among pairs or triplets of spike trains. This well-tried approach allows the inference of connectivity between and to the corresponding pair of neurons, as well as isolation of direct stimulus effects on the neurons. A tutorial with many references to the original literature appears in Glaser and Ruchkin (10).

The use of pairs forces a combinatorial proliferation of calculations (10 neurons means 45 pairs,  $\times$  2 for 2 time resolutions,  $\times$  2 for stimulus shift controls = 180 cross-correlograms for a *single* stimulus or behavioral condition). If the experiment compares 10 different stimulus conditions, the overloaded experimenter must assimilate some 1,800 crosscorrelograms, not to mention another 100 poststimulus time histograms, in order to specify how the assembly of 10 neurons behaves in these particular circumstances. The combinatorial situation is much worse if we calculate with triplets of spike trains. It is therefore entirely appropriate to seek a conceptual tool that allows a more direct interpretation of multineuron experiments in terms of neural assemblies.

In a recent paper (8) we have described a new representation for activity of the spike activities recorded in a multineuron experiment. This "gravitational" representation transcends the usual and problematic approach of addressing a group of neurons as a collection of all possible pairs or triplets and stresses the properties of the entire group of neurons as an entity. The basic representation associates each of N observed neurons with a particle in N-space, the *i*th particle located at  $\mathbf{x}_i$ . Each such particle carries a time-varying charge  $(q_i)$ for the *i*th particle), which is a filtered and normalized version of the spike train from the corresponding neuron. The resulting forces between charges will cause relative movement of the particles. Those particles corresponding to neurons that tend to fire in close time relationship will tend to aggregate. Different clusters of particles will signify different functional groups of neurons.

Explicit equations describing the charges, propulsive fields, forces, and displacements are given in the previous paper (8). At the location of a particular particle, we calculate the propulsive field due to the charges of all other particles (this is a vector summation). The resulting force on that particle is given by the product of the propulsive field and the local particle's charge. In formal notation total force on the *i*th particle is

$$\mathbf{F}_i = q_i \sum_{j \neq i} q_j \mathbf{r}_{ij}$$

where  $\mathbf{r}_{ij}$  is the unit vector pointing from particle *i* to particle *j*. We have used a high-viscosity model, so that there are no accelerations; velocity is proportional to the force. Simple integration leads to the displacement of the particle. This process is carried out for all particles, resulting in a set of trajectories in *N*space.

In the previous paper (8) it was shown that spike trains from simulated nets of neurons containing excitatory connections indeed led to appropriate clustering. This was inferred from the time development of all interparticle pair distances. In the present paper we show

more explicit projection methods to visualize the actual clustering in N-space and demonstrate that the aggregation actually produces a kind of Venn diagram for the neural network. Such projections are best seen on color displays and, for their dynamic aspects, in an animation movie. Some of the flavor is apparent in the present monochrome illustrations. In subsequent parts of the paper we examine the analogous investigation of inhibitory connections and then present a variation of the original representation, which addresses the problem of direction of connections rather than just association of firing times. This modification also has favorable consequences for sensitivity and signal-to-noise properties. Finally, we examine the consequences of rate normalization in calculating the charge on each particle.

### RESULTS

#### Projection

The basic gravitational representation of multineuronal activity, which we have presented earlier (8), maps the activity of the Nobserved neurons into an N-dimensional space. The N points that respectively represent each of the neurons move in an N-1 dimensional subspace under the influence of forces produced by a fictional "charge" on each point. The restriction to N-1 dimensions is because all forces are *internal* to the system of N particles. Furthermore, since any two particles exert equal forces on each other, the center of mass remains stationary. Various rules and convenient parameters in this representation have been presented in the previous paper; in the basic version, the timevarying "charge" of each point is related to the times of occurrence of action potentials from the neuron represented by that point. In due course, points in this representation will aggregate if the corresponding neurons fire in a time-correlated way. The arrangement of points in this aggregation and its time course can be used to infer the functional (and perhaps dynamic) connectivity of the entire neuronal assembly under examination. Note that in examining the aggregation of points in Nspace we are able to go beyond the usual statements about neuron *pair* relations as measured by cross-correlation; the entire assembly can be studied as an entity rather than as a combination of neuron pairs.

Mathematically there is no serious difficulty in dealing with the full *N*-dimensional space. It is only if we want to visualize the aggregation process that we are forced into some sort of projected representation. For convenience we restrict this discussion to a two-dimensional projection. Thus we are looking for a projection plane that adequately represents the movements of points in the *N*-space.

Any mapping that involves reduction of dimensionality (like projection) necessarily involves loss of information. Different types of mapping may lose different types and amounts of information; there may be an optimal choice for retaining particular features of the original N-dimensional distribution. Let us first examine a straightforward geometric projection onto a plane within the N-space; the plane will be chosen on the basis of preknowledge about the neuronal network. This is of course inappropriate to analysis of real neuronal data; however by using data produced by a neuron-simulator program we may better examine the possibilities of the projection rules.

In the following we use the same computersimulated neuronal circuit that we have examined in the previous paper (8); 10 neurons are arranged so that 1) there are two independently firing presynaptic neurons, 2) there is a set of four postsynaptic neurons for each presynaptic one, two neurons in each of these sets being driven by both presynaptic neurons, and 3) there are two independently firing control neurons. All connections are taken to be excitatory; the corresponding cross-correlograms are shown in Fig. 6 of Ref. 8. The strength of these interactions falls into the weak-to-moderate category encountered in physiological recordings. A Venn diagram of the neuronal circuit is shown in an inset in Fig. 1A. A plot of various pairwise distances during the aggregation process was shown in Fig. 7 of Ref. 8 and is replicated here in Fig. 1A.

We choose a projection plane that is determined by the three points corresponding to 1) the particles 1 and 2 (representing the 2 drivers in the circuit), and 2) the point at the center of mass of particles 9 and 10 (representing the independent neurons). The origin is chosen at the center of mass of particles 1 and 2 (i.e., halfway between them). Note that this definition results in a dynamic projection plane; as the several points move in the N-space so will the plane. In the following results we have determined this projection plane and the projections of all particles at each step of the calculation.

The projected trajectories of all points throughout the entire time span (0-16 s) of the calculation are presented in Fig. 1D. In effect this is a "time exposure" of the entire calculation. Identification of the traces that were originally in color obviously is lost in this black-and-white reproduction. Some of the traces and their directions can be identified with the aid of Fig. 1, E and F, which correspond to early (0-4 s) and middle (4-8.8 s)time spans. The independent particles (9 and 10) spend most of their time in an arc that evolves from the bottom of the display window and passes through the upper left. Particles 1 and 2 (the drivers) start at the extreme right and left of the screen and move inward in straight lines, as expected from the definition of the projection plane. All other particles (the driven) start at a point near the middle of the screen and are rapidly dragged to a small region near the origin of the projection plane (indicated by the dashed frame in Fig. 1D) where they congregate with particles 1 and 2.

Two typical enlarged "snapshots" of this region are shown in Fig. 1, B (at 4.84 s) and C (at 8.8 s). Even before 4.84 s this projection shows a configuration of particles that is reminiscent of the Venn diagram that describes the original network. Each driver particle has drawn in its own exclusively driven particles; the jointly driven particles (7 and 8) form a "bridge" between these two clusters, as shown in Fig. 1*B*. As time moves on, the configuration "boils" but continues to aggregate. Ultimately a single dense cluster is produced but still maintains the basic character of the Venn diagram (Fig. 1*C*).

A very good impression of the whole aggregation process can be obtained by combining "snapshots" into an animation movie. Such a movie clearly demonstrates that after the initial condensation of each driver with its directly driven particles, the bridge (7 and 8) takes over the role of attractor and pulls together the two clusters that had formed around its shared drivers (1 and 2). Note that condensation to the Venn diagram state has occurred before 5 s. This is an extraordinary sensitivity, since by then each neuron has given only some





50 spikes. (We have chosen the particular times shown in Fig. 1 for maximum clarity in separating overlapping trajectories and particle positions. The Venn diagram state persists for the entire period from 4.5 to 9 s in the calculation.)

We have investigated a number of other choices of projection plane within the same general framework. Variations we tried included other choices of the three basis points defining the plane and freezing the plane at some appropriate point during the computation. Many of these projections were about as informative as the example of Fig. 1.

The projections described here all depend fundamentally on a priori knowledge about the structure of the network for choosing the projection plane. This obviously is not satisfactory for an arbitrary set of data resulting from a physiological experiment. Other methods of visualization, not depending on preknowledge, will have to be considered; we return to this in the DISCUSSION.

# Detection of inhibition

The previous paper (8) was mostly concerned with excitatory neuronal connections. In order to deal with inhibition in terms of our representation we wish to produce aggregations as a consequence of favored configurations like presynaptic spike followed by postsynaptic silence (rather than presynaptic spike followed by postsynaptic spike). A simple way to attain this result is to reverse the sign of the force. This, in reversal of the original representation, will produce particle aggregation for spike-silence combinations and particle repulsion for spike-spike and silence-silence combinations.

The sign-reversed version of the gravity representation was calculated for a set of simulated spike trains from 10 neurons in which unit 1 was inhibiting unit 2 and the remaining 8 units were firing independently (the simulation algorithm is described in Ref. 1). The time development of a number of interpair distances from this set are shown in Fig. 2.4. Note that the distance between the two particles corresponding to the inhibitory pair (1, 2) decreases considerably faster and further than do the distances between unrelated pairs; the sign-reversed gravity representation indeed has the potential to detect inhibition among neurons. The rate of coalescence for the inhibitory pair in Fig. 2*A* compared with the results for excitation (cf. Fig. 1*A* and the results in Ref. 8), however, indicates a substantial reduction in sensitivity for inhibitory interaction. A similar asymmetry in sensitivity for the detection of excitation and inhibition was recently observed in cross-correlation analysis of spike trains (1) and can be shown to be formally related.

For the present spike-train simulation the strength of the inhibitory connection was set to be quite large (0.75 on a possible range of 0 to 1). Nevertheless the inherent noise in the spike trains causes some of the unrelated pairs to "diffuse" towards each other at rates that are uncomfortably close to that of the inhibitory pair (cf. Fig. 2.4). In realistic situations this may well lead to serious problems with respect to signal-to-noise ratio. Evidently some sort of criterion for significance and an appropriate measure for the "thermal noise" will be necessary; we leave this point for future theoretical analysis.

A surprising phenomenon in Fig. 2A is that there is a clear time variation in the velocity of approach of the particles corresponding to the inhibitory pair. The distance function shows an initial fast decrease (0-1 s) followed by an interval of slow or even absent decrease (1-4 s); then the rate of decrease rises to a higher and approximately constant value (4-14 s), and finally there is virtually no decrease at all (14-32 s). The phenomenon is real and data dependent and not due to a saturation inherent to the representation or to its numeric implementation. This is proven by the observation that time reversal of the spike trains essentially resulted in a time reversal of the distance curve.

Deviations from the calculated average firing rate produce variations in force that in turn could produce effects as in Fig. 2.4. Such spikerate fluctuations (of  $\sim 1$  s) can be ruled out here because in this particular example the charge history of each particle was normalized with a 20-interval moving average of the corresponding neuron's firing rate. This leaves only the possibility of a real time dependence (due in this case to statistical fluctuation) in the degree of synchrony of firing and silence of the units involved. This was investigated by evaluation of the cross-correlation function of the two spike trains. Cross-correlograms were calculated over two separate observation in-



FIG. 2. Gravitational clustering of inhibitory pair (neuron 1 inhibits neuron 2) among a group of 10 otherwise unrelated, simulated neurons. All 10 spike trains were Poisson-like; mean rates were 10/s; the inhibitory synapse had a strength of -0.75, a delay of 5 ms, and silent intervals in the postsynaptic train with duration of 80-100 ms (1). A shows the time development of a number of interpair distances; the distance between the pair of particles corresponding to the inhibitory neuron pair is indicated by the label 1,2; all other curves correspond to unrelated pairs. Arrows along the 1,2 distance curve mark the time intervals with different rates of aggregation of the 1,2 pair. The cross-correlation functions of the spike trains from neurons 1 and 2 over these two intervals are shown in B (4-14 s) and C (14-24 s). Note the relative similarity in the inhibitory trough in the correlogram (positive time shift, between 5 and 100 ms) and the distinct excess in correlated firing prior to the inhibition in C compared with B (negative time shift between -32and 0 ms). Note that this excess correlation is due to statistical fluctuations. Further explanation in text.

tervals corresponding to the case of fast (4-14 s) and slow (14-24 s) coalescence. Results are shown in Fig. 2, *B* (fast) and *C* (slow). Comparison of these correlograms shows that in this particular example it is not the synchrony between events in train 1 and silences in train 2 that is changing; the troughs in both correlograms are virtually identical. However, the two periods differ in the numbers of nearly coincident firings of the two units; there is an increased probability of firing of unit 2 pre-

ceding an event from unit 1 in the interval of 14-24 s (Fig. 2C) compared with the earlier interval of 4-14 s (Fig. 2B). Note that this excess in correlated firing prior to the inhibition is purely due to statistical fluctuations in the simulation; the model did not contain any excitatory connection. During the interval corresponding to Fig. 2C the excess correlation causes a repulsive force, which in this case manages to balance the attractive force due to inhibition. Consequently the rate of change of

the interpair distance, which can be shown to be related to the area under the time-dependent correlogram, drops to approximately zero.

The conclusion from the observed time dependence of the interparticle distance is twofold. The gravity representation apparently is sensitive to time dependence of the degree of connectivity as manifested in the correlation of firing and/or silences. This obviously should be considered an advantageous property for detection of possible dynamic changes in strength of functional neuronal connections associated with stimulation, context, behavior, learning, and the like. At the same time, however, the example of Fig. 2 shows that the performance of the original gravity procedure is adversely affected by the symmetry of the force rules: event from 1 precedes event from 2 and event from from 2 precedes event from 1 are treated equally. With a unidirectional neural interaction this symmetry causes roughly half of the elementary contributions to the force history to consist of noise. The sensitivity and selectivity will drop accordingly. This is the situation during the period between 4 and 14 s. During the later period between 14 and 24 s the situation becomes even worse; the irrelevant part of the correlation (left side of the correlogram) not only adds noise but also adds a force component which on average counteracts and actually balances the attractive component.

This adverse consequence of symmetry in the force rules becomes manifest in the case of inhibition, whereas it was not evident for excitatory interaction. The reason for this is that the gravity representation is considerably less sensitive for inhibition to begin with. The foregoing discussion clearly points to the desirability of having asymmetric force rules. Not only can this possibly address the question of direction of the connection (causality), it should also improve the sensitivity and selectivity of the gravity representation by a considerable amount.

# Direction of connections: the two-charge representation

In the original model (8) the force on an individual particle is the product of its own charge and the propulsive field at that particular position. The field in turn is calculated as the vector sum of the individual fields generated by the *other* particles, i.e., a function of their individual charges and positions. Thus "charge" is used for two purposes: 1) for generation of the field, and 2) as a coefficient that evaluates the force on the individual particle. Now let us consider a model in which *separate* classes of charge are associated with these two separate functions. One class of charge will be used to calculate the propulsive field; we will call this the *effector charge*  $q_e$ . Thus the propulsive field at particle *i* is now given by

$$\mathbf{f}_i = \sum_{j \neq i} q_{\mathbf{e}j} \mathbf{r}_{ij} \tag{1}$$

where  $\mathbf{r}_{ii}$  represents the unit vector given by

$$\mathbf{r}_{ij} = (\mathbf{x}_j - \mathbf{x}_i) / S_{ij} \tag{2}$$

Here  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are vectors representing the positions of particles *i* and *j*, while  $S_{ij}$  is the Euclidian distance between them.

The second class of charge will be used to calculate the net force on that particular particle; we will call this the *acceptor charge*  $q_a$ . Thus the force on particle *i* is

$$\mathbf{F}_i = q_{\mathbf{a}i}\mathbf{f}_i \tag{3}$$

In this modification each particle is associated with two distinct charge histories, each serving a different purpose in the calculation. In the original model these charges and purposes were combined.

In the original model, each charge was normalized in two ways: for average spike rate and to obtain zero mean charge. Such normalizations are included in our definitions of q; we have made a slight modification by using a sliding rather than global spike-rate average. As in the original model we will continue to have high viscosity. Therefore the velocity of each particle will be proportional to the force upon it

$$d\mathbf{x}_i/dt = \sigma q_{ai}(t)\mathbf{f}_i(t) \tag{4}$$

with  $1/\sigma$  representing the viscosity. Finally, the displacement is calculated by simple integration

$$\Delta \mathbf{x}_i = h \sigma q_{\mathbf{a}i}(t) \mathbf{f}_i(t) \tag{5}$$

where h denotes the time step in the computation.

This modification of the charge rules together with appropriate operation rules leads to a model from which it is possible to infer causality, i.e., to produce *directed* graphs (wiring diagrams). The original model, due to its inherent symmetry, only detects temporal association of firing regardless of pre- and postsynaptic identification.

The most important requirement for the new charge rules is that they differentiate between two favored patterns: a spike from neuron 1 precedes a spike from neuron 2 versus the reverse situation. Such favored patterns are the prototype signatures of the two possible directions of neuronal interaction: 1 drives 2 versus 2 drives 1.

For the two required charges let us introduce the functions shown in Fig. 3; one charge consists of a decaying exponential starting at and following the spike, whereas the other charge consists of a rising exponential that terminates at the spike. Note that we have not yet assigned these charges to "effector" or "acceptor" roles; in fact, it will be necessary to consider both possible assignments, as we will see below.

Let us examine the forces and resulting position changes of two particles when the fa-



FIG. 3. Conversion of a train of action potentials (1 spike schematically shown in *top trace*) to a charge function used in the gravitational calculations may be based on different elementary charge functions. Two examples are shown here: a decaying exponential starting at and following the spike (*middle trace*) and a rising exponential that terminates at the spike (*bottom trace*).

vored spike configuration is that a spike from neuron 1 precedes that from neuron 2 (i.e., neuron 1 drives neuron 2). For the three possible assignments of charges (original and 2 new) Fig. 4 depicts schematically the forces on each particle resulting from the overlap of the appropriate charge histories as well as the final net position change. Original positions are indicated by the dashed circles. In Fig. 4 column 1 refers to the original model; the two alternative charge assignments for the two-charge model are given in columns 2 and 3. The charge profiles are given in the top row. The interesting observation to be made from this figure is the difference in the three sets of position changes indicated in the bottom row. In the two-charge model only one particle will move while the other stays stationary; one particle is acting as a stationary attractor of the other. In the conditions of column 2, the particle representing the presynaptic neuron moves towards the particle representing the postsynaptic neuron. In the conditions of column 3 the point representing the presynaptic neuron acts as the attractor. In the original model both points move towards their center of mass; no inference about pre- and postsynaptic identification is possible. Examples of such movements in the original and two charge models are shown in Fig. 6. The circuit here was the same inhibitory connection used for Fig. 2. See DISCUSSION for additional aspects of Fig. 6.

If the favored spike pattern had been reversed, i.e., neuron 2 drives neuron 1, Fig.'4 still holds, provided that 1 and 2 are everywhere interchanged. For this simple synaptically related neuron pair, it is not necessary to consider both versions of the two-charge model; the causal implications of columns 2 and 3 are identical. However even in slightly more complicated neuronal configurations both calculations are essential, as shown below.

Figure 5 shows a small catalog of three-neuron circuits and the movements and final positions associated with each of the three sets of charge rules. The circuits considered here have at most one direct connection between any two neurons. Since only three neurons are involved, the movements of the points are restricted to a two-dimensional plane and can be fully depicted without the complications of

0 <b>—</b> @			
			1
a = 1			
force			
on 1			
a = 2			
e = 1			
force on 2	1 -2	0 0	1 -2
net displacement	୦୦ ୭୦	0 0	<b>୦</b> ହ

FIG. 4. Forces and resulting position changes of 2 particles for 3 possible assignments of charge functions when the favored spike configuration is that a spike from neuron 1 precedes a spike from neuron 2 (1 drives 2). The three assignment rules of charges are given in the *top row*: original gravity representation (forward acceptor and forward effector, *left column*) and the two alternative assignments for the present two-charge representation (forward acceptor and backward effector, *middle column*; backward acceptor and forward effector, *right column*). In rows 2 and 3 the forces on each of the two particles involved are graphically derived from the overlap in the relevant charge histories. *Bottom row* shows the resulting position changes of the two-particle configuration for each of the three charge assignments (original positions are indicated by the *dashed circles*). Note that both versions of the corresponding particles, unlike the original representation, which only shows that both neurons are associated, without allowing inference regarding the direction.

projection. As before, each of the position columns corresponds to a particular set of charge rules (as in Fig. 4). Overall calibration is provided by row one, corresponding to a set of three independent neurons; there is no systematic position change in any of the three representations. The simple connection of Fig. 4 is repeated in the second row. Divergence and convergence of simple excitatory connections are contrasted in rows 3 and 4. Row 5 considers an open cascade, whereas a closed loop is shown in row 6. The remaining possible three-neuron circuit is given in row 7.

Using the original version of the gravity

representation (position column 1) all of the networks with more than one connection (rows 3–7) produce identical trajectories to a single cluster at the center of mass. This means that the different circuits *cannot* be distinguished in this way. However if we have available the results of both two-charge calculations, all circuits are clearly distinguished. Note that for the charge assignment in position column 2, the points representing postsynaptic elements act as attractors. For the assignments of position column 3, the points representing presynaptic elements act as attractors. For some networks, it is necessary to know both



FIG. 5. Catalog of 3-neuron circuits (*left column*) and the movements and final positions of the corresponding particles in the gravity representation (original positions are indicated by the *dashed circles*). As in Fig. 4, each of the three position columns corresponds to a particular set of charge rules, given in the *top row*. Circuits considered here have at most 1 direct connection between any 2 neurons. The different cases considered are: 1) three independent neurons, 2) one excitatory connection (cf. Fig. 4), 3) divergence of excitation, 4) convergence of excitation, 5) cascade of three neurons, 6) closed loop of three neurons, and 7) cascade, shunted by direct connection (or combination of divergence and convergence). Note that the original gravity representation (1st position column) does not distinguish between neuron configurations 3-7; combination of final positions and trajectories from the two alternative 2-charge representations leads to a clear distinction in all cases. Further explanation in text.

final particle positions and the trajectories. An example is the distinction between the open cascade (row 5) and the mixed cascade (row 7). Note that in the latter the signature of the open cascade dominates the convergence and divergence that are also present. Finally, the sense of the connections in the closed loop (row 6) can be determined from the sense of the trajectories. The analysis of similar circuits involving more neurons will generally follow similar principles.

We stress that unambiguous and robust conclusions require the calculation of both possible charge assignments, i.e., both righthand columns in Fig. 5. All the above assumes that all synapses have equal strength and that there is negligible contribution from random fluctuations. In particular, the final configurations where a particle arrives at the center of mass of the others (cf. Fig. 5 row 3, position column 2) are unstable equilibria; if there is a small difference between the two synaptic strengths, the moving particle will rapidly reach the vicinity of the center of mass and will subsequently slowly drift towards the particle representing the neuron with the stronger connection. Random fluctuations of the interparticle forces resulting from the stochastic nature of the spike trains set a lower limit to the synaptic strengths or differences of strengths to be discerned.

If we allow more connections between any two neurons, the situation becomes more complicated, but the analysis proceeds along similar lines. For instance, a similar arrangement as in row 2, now however with two connections in obverse directions (mutually excitatory pair), will result in both corresponding particles being attracted to their center of mass for each of the three sets of charge rules; all three position columns in that case will show the particle arrangement given here in Fig. 5 (row 2, position column 1).

#### Aggregation and the normalizations

We have observed in a number of instances that two data sets with apparently quite similar cross-correlograms between spike trains of the interacting neurons show different rates of aggregation in the gravity representation. This discrepancy turns out to be partly a function of visual interpretation of relatively sparse cross-correlograms (i.e., containing many bins with no counts). More important, however, the discrepancy depends on the type of spikerate normalization that is installed in the gravity representation. Note that the discrepancy is only a matter of aggregation *rate*; the correct aggregations always occur.

In all variants of the gravity representation that we have so far presented, each particle charge was normalized in two ways. One normalization adjusted for zero mean charge on each particle. This served to keep particles from aggregating unless their charge histories were temporally correlated. An improvement we have used throughout this paper is to use a sliding rather than a global average for such normalization.

The second normalization adjusted for average firing rate. Again the average is best taken in sliding form. The original rationale for rate normalization was to keep a fast-firing neuron from dominating movements of the particles. It turns out, however, that there are tradeoffs in overall performance of the gravity representation when different rules for rate normalization are invoked.

We have arbitrarily chosen the elemental charge to be a decaying exponential of time constant, which is incremented by q at the time of an action potential. Within this context there seem to be several logical choices for firing rate normalization. These are calculated as follows:

The area of an elemental charge described by  $q = q_0 e^{-t/\tau}$  is just  $q_0 \tau$ . If there are *M* such events in time *T*, the (global) mean charge is

$$\bar{q} = (Mq_0\tau)/T = q_0\tau/\mu \tag{6}$$

where  $\mu$  is the mean interval between spikes of that train. In order to attain rate normalization,  $\bar{q}$  must be made independent of  $\mu$  by appropriate definition of  $q_0$  and/or  $\tau$ . Some possibilities are as follows:

MODEL 1. Charge-increment normalization. Let  $q_0 = \mu/\mu_0$  where  $\mu_0 = 1$  ms. We may choose any reasonable  $\tau$  (which will be used for *all* particles). Thus for *all* particles

$$\bar{q} = \tau/\mu_0.$$

MODEL 2. Time-constant normalization. Let  $\tau = \mu/c$  where c is a constant (c should be ~10 to preserve the temporal pattern details of the spike train in the charge history; c determines the degree of smoothing). We may choose any reasonable  $q_0$  (which will be used for *all* particles). Thus for *all* particles

$$\bar{q} = q_0/c.$$

MODEL 3. No normalization. Here we are free to choose both  $q_0$  and  $\tau$  (which will be used for all particles). Thus

$$\bar{q} = q_0 \tau / \mu$$

but this is *different* for each particle because the value of  $\mu$  is different.

In all three models,  $\bar{q}$  is subtracted from each instantaneous charge value to produce an effective charge of zero mean as required by the first (mean charge) normalization. Note that the value to be subtracted is universal for all particles in models 1 and 2, but different for each particle in model 3.

Model 1, charge-increment normalization, has been used throughout this and the previous paper (8). Model 2, time-constant normalization, was suggested to us by Prof. H. Jasper as a way of attaining better sensitivity to detailed firing patterns.

It turns out that the three choices of normalization produce different velocity of particle aggregation for a given near-coincidence rate. In addition, performance with the three normalizations differs in its sensitivity to individual spike rates. To make this explicit, let us characterize two spike trains by each of their rates  $\rho_1$  and  $\rho_2$  as well as the rate of near coincidences  $\rho_{12}$ . The number of counts in the cross-correlogram peak (calculated for data duration T) is therefore  $\rho_{12}T$ . The correlogram "background" will be  $T\rho_1\rho_2\Delta t$  where  $\Delta t$  is the bin width.

For each of the rate normalization models defined above we may use the gravity calculation to obtain the reduction of distance between the two particles whose charges represent these two spike trains. The results are

MODEL 1. 
$$\Delta S_{12} = K_1 T \rho_{12} / \rho_1 \rho_2$$
  
MODEL 2.  $\Delta S_{12} = K_2 T \rho_{12} / (\rho_1 + \rho_2)$   
MODEL 3.  $\Delta S_{12} = K_3 T \rho_{12}$  (10)

(The three constants  $K_1$ ,  $K_2$ , and  $K_3$  of course have different units.)

Thus, if we contrast two spike train *pairs* that have the same near-coincidence rate (i.e., identical peaks in the cross-correlation) but with different spike rates, we find that the higher rate pair will show slower aggregation in model 1 and model 2. Only model 3 will show aggregation independent of individual spike rates and hence will behave more in keeping with our intuitive interpretation of cross-correlogram peaks. The conclusion of these calculations is that model 3 (no rate normalization) is probably the best choice for unbiased analysis of unknown data.

### DISCUSSION

We have demonstrated that the gravity representation is very capable of analyzing the interrelationships within a group of observed neurons without having to decompose the assembly into pairs or triplets. The sensitivity is extraordinary, as shown in Fig. 1, where a mere 50 spikes from each contributing neuron were enough to produce aggregation and to show the essential Venn diagram of the complete network. This sensitivity transcends at least by an order of magnitude that of the joint PST scatter diagram (7) or the three-neuron "snowflake" (17), which produce comparable information, but only for two- and three-neuron networks, and require hundreds to thousands of spikes from each neuron. The basic improvement in sensitivity is related to the time integration implicit in the gravity representation.

In the present paper we have dealt with four issues: 1) visualization of the clustering process in *N*-space by projection to a plane, 2) detection of inhibition by tailoring the force rules of the gravity representation, 3) inferring causality (i.e., direction of connectivity) by introducing new charge rules, and 4) some consequences of the normalization rules.

#### Projection

Our experience with this type of visualization involved a color display and showed the utility of movies for representing dynamic aspects of the clustering process. Neither of these tools could be demonstrated here. Nevertheless, some of the flavor has been shown by the "time exposures" and "snapshots" of Fig. 1. For an actual investigation such tools remain essential.

The projections described all involved choice of an actual plane in the *N*-space, and depended critically on a priori knowledge about the connectivity of the network. In a real laboratory situation this preknowledge is absent, and more general methods from pattern recognition and cluster analysis must be employed (5). One promising possibility would be the so-called "structure preserving" nonlinear mapping introduced by Sammon (19) or modifications thereof (2, 20). These methods best preserve the shortest distances in the original distribution; therefore we would expect to see the same clusters as in the direct geometric projections presented here. It remains to be seen, however, if the relationships between clusters, as in the Venn diagrams shown, will be preserved in the nonlinear mappings. There are a number of unsolved problems involved in this type of data-dependent and iterative mapping, such as the dependence of the mapping on starting configuration in the projection plane and the distortion of trajectories. Moreover, these methods generally involve very heavy computation.

The utility of projection schemes is to provide visual feedback to the experimenter in a readily interpretable form involving information about the entire assembly of neurons. It may be useful to avoid this procedure, and to develop methods that describe directly the evolution of clusters and their shapes in the *N*-space. For example an undistorted description in terms of hyperspheres and hypertubes in the *N*-space might be more accurate than projections. Visualization could be obtained by (mentally) projecting such gestalts (rather than individual points) to a 2- or 3-space (21).

#### Inhibition

We analyzed inhibition by reversing the direction of force between any two particles, thus causing aggregation when the corresponding neurons show association of spike and silence. We noted that the aggregation process was slower than for comparable excitatory connections.

The underlying reason for this difference in sensitivity is the relationship of forces and charges in the gravitational representation. Force is produced by the vector sum of products of suitably normalized charges. Among other things the normalization subtracts average charge. Thus the normalized charge of a particle is large and positive immediately after a spike of the corresponding neuron and subsequently decays to a small negative value during the interval between spikes. A spikespike configuration between two neurons implies a product of two relatively large and positive charge values (little decay has occurred). The spike-silence configuration has one charge large and positive, and the other small and negative; the product therefore is smaller and negative. Finally the silence-silence configuration involves two small charges, both negative; the product therefore is small and positive.

When we dealt with excitatory connections the force in the first and third configurations was defined as attractive; the force in the second configuration was correspondingly repulsive. The modification of force rules that we used for detection of inhibition involved the reversal of these relationships. Keeping in mind the relative magnitude of the several components, it is obvious that the aggregation will be slower for the rules used in the inhibitory case. In this form, the gravitational representation is less sensitive for inhibition than for excitation.

Other possible rules may not suffer from this limitation. For example, as proposed in the previous paper (8), we may define a positive charge to favor silences of some particular duration. This would be a two-charge representation of a different sort than discussed in the present paper.

#### Two-charge representation

The forward and backward charge histories described in the present paper allow us to infer the direction of connections between neurons. It is essential to evaluate both the "final" positions of the particles and the paths by which they get there. Furthermore the catalog in Fig. 5 shows that it is necessary to calculate both possible sets of charge assignments in order to attain robust conclusions.

We have shown that it is necessary to evaluate excitatory and inhibitory connections by using different rules in the charges and resulting forces. Charge rules are selected to favor some particular configuration of spikes and/or silences. If we assume only unidirectional connections between any two neurons, the original gravity representation, due its inherent symmetry, is vulnerable to statistical interference from the "noncausal" half of the correlogram. (Presumably a unidirectional connection, whether single or multisynaptic, produces a feature either to the left or the right of the correlogram origin.) Such interference was demonstrated for an inhibitory connection in Fig. 2.

A consequence of introducing the twocharge representation is an increase in immunity to interference from the "noncausal" half of the correlogram. In effect, sensitivity is increased. This effect is demonstrated in Fig. 6*A*, which shows the development of interparticle distance for the same data and pairs



as used in Fig. 2. Computation parameters were similar, except that Fig. 6A was obtained from the two-charge representations. The projected positions of the particles in the single charge (as in Fig. 2) and both two-charge representations are shown in Fig. 6, B-D.

The effect of statistical perturbations may be less marked in the particle trajectories for the two-charge representations, and the inference of causality is trivial. The two-charge representation does not eliminate the utility of the original symmetric representation. There is a doubling of calculation time because both possible sets of charge assignments must be used. In addition the interpretation is much more complicated than for the simple representation (cf. Fig. 5). These time costs suggest that initial screening with the original representation is more appropriate for a quasi-online experimental situation.

In the original representation the center of mass of the total configuration of particles remained stationary at all times. In the twocharge model Newton's "action equals reaction" has been discarded, and as a consequence the center of mass is no longer a conserved quantity. Since we are dealing here with an entirely imaginary construct, this violation of physical law should be no cause for alarm.

The previous paper (8) introduced a different type of two-charge representation than we have discussed here. The earlier formulation involved complications such as two "alter egos" for each particle and the arbitrary spatial fixing of the locations of the effector "egos." The present formulation allows the same types of functional inference with considerably more conceptual clarity and without such arbitrary restrictions.

# Normalization

We have shown that velocity of aggregation is sensitive to the individual neuronal firing rates as well as to the near-coincidence firing rates. Several different sets of rules for normalization were examined. Each of these produced different sensitivities to individual firing rates in relation to the sensitivity for near-coincidence firing. The choice among different normalization rules must take into account additional factors. For example, the several normalizations also have different selectivity for firing pattern or burst structure. Thus normalization rules involve choices and tradeoffs; there is no universal optimal solution. At the same time, we emphasize that appropriate aggregation will take place no matter which normalization is used. Only the velocity and variance of aggregation are affected.

#### Implementation

It unfortunately remains necessary to run these computations on a minicomputer rather than a laboratory microcomputer. The work shown in this paper was done in FORTRAN running on a Data General MV10000 computer (comparable with DEC VAX 785). With this computer power (but no special array processors) typical runs for 16 or 32 s of spike data from 10 neurons took 3 min. Although some improvement is undoubtedly possible, this speed will suffice for quasi-on-line guidance of a multineuron experiment.

# Unresolved issues

A number of questions require additional work: 1) What is the relation between the gravity representation and the generalized cross-correlation of spike trains? The computations are obviously different and are differently subject to statistical difficulty. We have already remarked that in order to arrive at significant conclusions the gravity representation requires far less data than do the joint PST scatter diagram and snowflake (forms of 2- and 3-way correlation functions for spike trains).

2) For unrelated neurons, the corresponding particles execute motions that are vaguely reminiscent of Brownian motion (6, 14). However, the situation is fundamentally different; Brownian motion assumes that a par-

FIG. 6. Performance of the two-charge representation: development of interparticle distances and projected particle positions for neuron 1 inhibits neuron 2, the same spike trains as used in Fig. 2. Charge rules are indicated by *inset* in each panel. A: distances vs. time. Note the steady decrease of the 1,2-distance and the improved "signal-to-noise ratio" compared with the results from the original representation (Fig. 2). B: corresponding projected positions. Note that particle 1 (presynaptic) moves towards an essentially stationary particle 2 (postsynaptic). C: projected positions in the single charge representation. Note that particles 1 and 2 move symmetrically towards each other. D: projected positions in the dual of B. Note that particle 2 (postsynaptic) moves towards an essentially stationary particle 1 (presynaptic).

ticle is acted upon by a very large number of forces, and that the whole system is defined in 3-space. In the gravitational representation, only N forces are involved, where N is a relatively small number. However, the system is in N-space, and N is considerably larger than 3. At what time in the development of the system will these diffusion-like movements effectively swamp those indicative of real neuronal interactions? Can something like a diffusion constant be defined? Such considerations determine the limits of sensitivity and signal-tonoise ratios that can be obtained.

3) Recent papers (12, 13) about multistate systems suggest that an efficient way to reach the global minimum "energy" state is to use a decreasing temperature schedule. The whole process becomes analogous to annealing (15). Can a similar approach be used in the gravity representation, for instance by using a particle configuration-dependent value of the viscosity parameter (see Eq. 4)?

#### REFERENCES

- AERTSEN, A. M. H. J. AND GERSTEIN, G. L. Evaluation of neuronal connectivity: sensitivity of crosscorrelation. *Brain Res.* 340: 341–354, 1985.
- BISWAS, G., JAIN, A. K., AND DUBES, R. C. Evaluation of projection algorithms. *IEEE Trans. Pattern Anal. Mach. Intelligence* PAMI3: 701–708, 1981.
- EGGERMONT, J. J., EPPING, W. J. M., AND AERTSEN, A. M. H. J. Stimulus dependent neural correlations in the auditory midbrain in the grassfrog (*Rana temporaria L.*) *Biol. Cybern.* 47: 103–117, 1983.
- 4. FROSTIG, R. D., GOTTLIEB, Y., VAADIA, E., AND ABELES, M. The effects of stimuli on the activity and functional connectivity of local neuronal groups in the cat auditory cortex. *Brain Res.* 272: 211–222, 1983.
- 5. FUKUNAGA, K. Introduction to Statistical Pattern Recognition. New York: Academic, 1972.
- GARDINER, C. W. Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences. Berlin: Springer-Verlag, 1983.
- GERSTEIN, G. L. AND PERKEL, D. H. Mutual temporal relationships among neuronal spike trains. *Biophys.* J. 12: 453–473, 1972.
- GERSTEIN, G. L., PERKEL, D. H., AND DAYHOFF, J. E. Cooperative firing activity in simultaneously recorded populations of neurons: detection and measurement. J. Neurosci. 5: 881–889, 1985.
- GERSTEIN, G. L., BLOOM, M. J., ESPINOSA, I. E., EVANCZUK, S., AND TURNER, M. R. Design of a laboratory for multi-neuron studies. *IEEE Trans. Syst. Man Cybern.* SMC13: 668-675, 1983.
- GLASER, E. M. AND RUCHKIN, D. S. Principles of Neurobiological Signal Analysis. New York: Academic, 1976.
- 11. GRINVALD, A., COHEN, L. B., LESHER, S., AND BOYLE, M. B. Simultaneous monitoring of optical ac-

4) Finally, it is time to apply the gravity representation to real data from real neurons. This will necessitate generalization to time-varying firing rates induced by stimulation. In the case of repeated stimulation, various schemes of shuffling of the data, comparable with the procedures used in cross-correlation, will allow the separate assessment of purely stimulus related effects.

#### ACKNOWLEDGMENTS

This work was supported through National Institutes of Health Grant NS-05606 to G. L. Gerstein, a Fulbright Program Travel Grant to A. Aertsen, and System Development Foundation Grant SDF-0013 to G. L. Gerstein.

Received 13 December 1984; accepted in final form 11 July 1985.

tivity of many neurons in invertebrate ganglia using a 124 element photodiode array. J. Neurophysiol. 45: 829-840, 1981.

- HINTON, G. E., SEJNOWSKI, T. J., AND ACKLEY, D. H. Boltzmann machines: constraint satisfaction networks that learn. Pittsburgh, PA: Carnegie-Mellon Univ., 1984. (Tech. report CMU-CS-84-119).
- 13. HOPFIELD, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA.* 79: 2554–2558, 1982.
- KAMPEN, N. G. VAN. Stochastic Processes in Physics and Chemistry. Amsterdam: Elsevier/North-Holland, 1981.
- KIRKPATRICK, S., GELATT, C. D., JR., AND VECCHI, M. P. Optimization by simulated annealing. *Science* 220: 671–680, 1983.
- KRUGER, J. A 12-fold microelectrode for recording from vertically aligned cortical neurons. J. Neurosci. Methods 6: 347–350, 1982.
- PERKEL, D. H., GERSTEIN, G. L., SMITH, M. S., AND TATTON, W. G. Nerve-impulse patterns: a quantitative display technique for three neurons. *Brain Res.* 100: 271–296, 1975.
- REITBOECK, H. J. A 19-channel matrix drive with individually controllable fiber microelectrodes for neurophysiological applications. *IEEE Trans. Syst. Man Cybern.* SMC13: 676–682, 1983.
- SAMMON, J. W., JR. A nonlinear mapping for data structure analysis. *IEEE Trans. Comput.* C18: 401– 409, 1969.
- TEREKHINA, A. YU. Methods of multi-dimensional data scaling and visualization (survey). Avtom. Telemekh. 7: 80-94, 1973.
- ZAHN, C. T. Graph-theoretical methods for detecting and describing gestalt clusters. *IEEE Trans. Comput.* C20: 68–86, 1971.

Present address of A. M. H. J. Aertsen: Max-Planck-Institut für Biologische Kybernetik, 74 Tübingen, West Germany.