

# Measuring synchronization of stochastic oscillators in biology

Z Deng<sup>1</sup>, S Arsenault<sup>2</sup>, L Mao<sup>2</sup>, and J. Arnold<sup>3</sup>

<sup>1</sup>College of Engineering, University of Georgia

<sup>2</sup>Institute of Bioinformatics and Entomology Department, University of Georgia

<sup>3</sup>Genetics Department, University of Georgia

arnold@uga.edu

**Abstract.** A fundamental problem in physics is measuring and modeling the synchronization of coupled stochastic oscillators. The problem is relatively recent in biology, where it has become possible to measure stochastic oscillators in single cells. A variety of synchronization measures have been proposed to describe a field of coupled stochastic oscillators. We introduce a synchronization measure new to this problem (but old to Genetics) called the intraclass correlation (ICC). The ICC is simple to interpret and has a statistical framework for inference. We illustrate ICC behaviour in the Kuramoto phase-locking model and on a field of over 25,000 oscillators in single cells measured every half-hour over a ten day interval.

## 1. Introduction

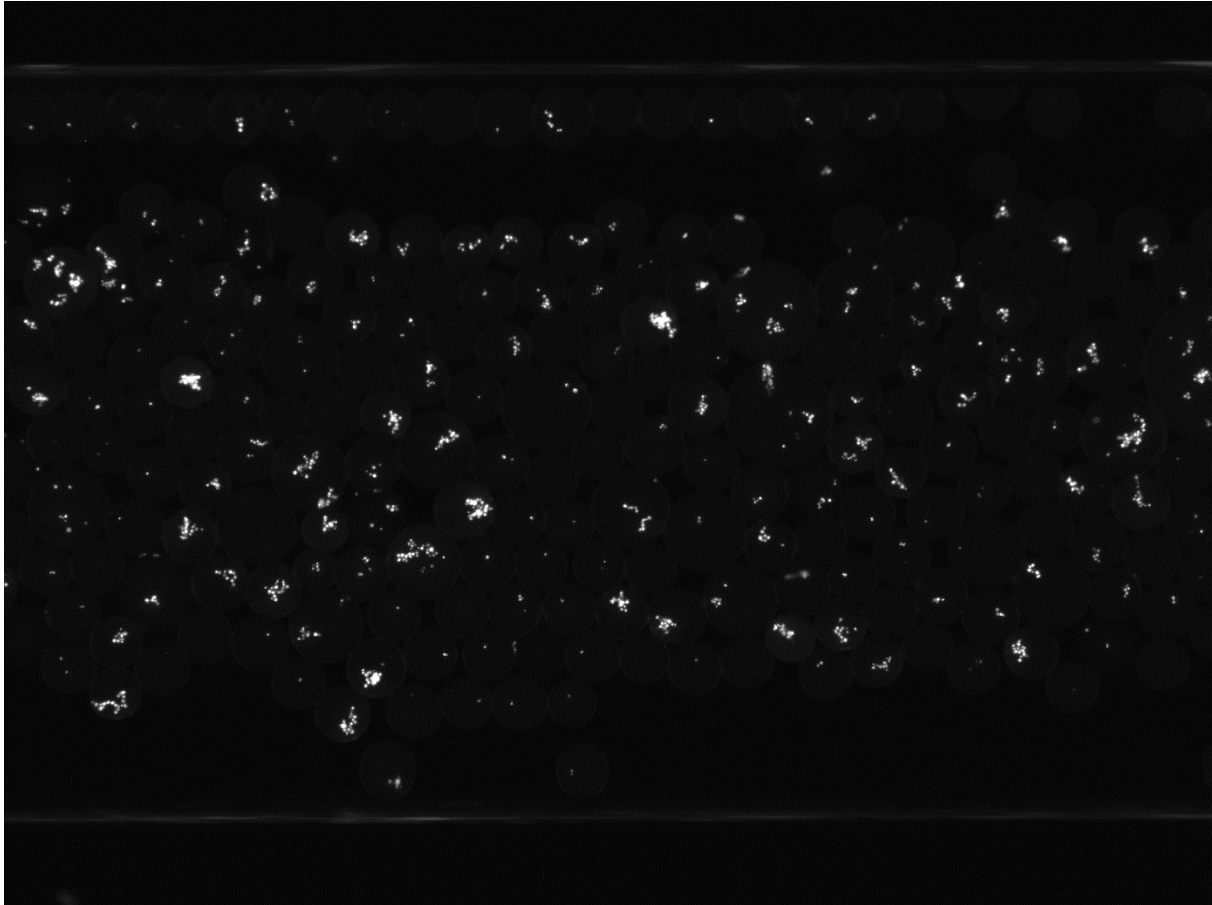
The following paper is dedicated to H.-Bernd Schuttler on his 60<sup>th</sup> birthday in honor of his contributions in systems biology drawn from statistical physics(1-4). His introduction of ensemble methods into systems biology constituted a sea change for the field, one now that many of us take for granted. The focus today is on a new problem on which Bernd is working that is central to physics and more recently in biology, namely measuring and modeling the synchronization of coupled stochastic oscillators.

The problem arises in physics in the synchronization of lasers, Josephson-Junction arrays, relativistic magnetrons, and the coupled signals of two interferometers measuring gravitational waves (5, 6). Only recently has the problem come to the attention of biologists, particularly in the context of circadian rhythms(7). With the advent of new technologies for measuring the clocks in single cells(8), it is now possible to make progress on how the clocks of individual cells become synchronized. Being able to measure the clocks of individual cells fundamentally changes the scale of measurement usually performed on mass cultures of  $10^7$  cells. Because of the change in scale, new behaviors are evident that demand new models for the stochastic oscillations of single cells(9). Many of the new stochastic models describing the behavior of individual cells have been drawn from physics(10).

The problem to be described here is simple. Each cell has its own oscillator, and the average behavior of these oscillators has been successfully described with ensemble methods applied to deterministic models at the level of  $10^7$  cells (2, 3). The challenge today is to measure individually



each of the oscillators of >25,000 cells over a ten day period and determine how the cells synchronize over time. In contrast to earlier experiments, these new measurements are done at the single cell level. The data are the trajectories of each cell's oscillator over a ten day window measured every half hour (h). The data for one time point are shown in Figure 1 for one field of view and resemble a star field observed at one of successive time points. All of the data at one time point is the union of many fields of view (~30). We will introduce measures of synchronization between the oscillators (stars) drawn from physics and genetics. We will illustrate their behavior in a physical model. Then we will examine these synchronization measures and their behavior on >25,000 individual cells over a 10 day window.



**Figure 1.** One field of view on oscillators in single cells at a single time point.

## 2. Synchronization measures

A good synchronization measure should have a number of properties. One, it should capture the similarity between oscillator trajectories as summarized in their phase, amplitude, and period. Second, any measurement of similarity should be done relative to the variability in the measurements between replicates. This will allow an assessment of significance of trajectory similarities found. Three, the measure should have a simple interpretation. Finally, the measure should have well known statistical properties to facilitate inference about synchronization of stochastic oscillators.

A variety of synchronization measures have been introduced to describe the coupled behavior of stochastic oscillators in physics. These include the cross-correlation between cellular trajectories (with respect to frequency), mutual information shared between trajectories, the phase synchronization from either the Hilbert or Wavelet transform of individual cellular trajectories, and an index based on the circular variance known as the phase-locking value(11).

Of these methods, we chose to utilize the phase synchronization statistic (or phase locking index) extracted using the Hilbert Transform. By taking the Hilbert Transform of the data

$$\tilde{x}(t) = \frac{1}{\pi} * p.v. \int_{-\infty}^{+\infty} \frac{x(t')}{t - t'} dt'$$

It is possible to define the analytic signal

$$Z_x(t) = x(t) + i\tilde{x}(t) = A_x^H(t)e^{i\varphi_x^H(t)}$$

and consequently extract the phase with respect to time.

$$\varphi_x^H(t) = \arctan \frac{\tilde{x}(t)}{x(t)}$$

An index of phase locking can then be calculated to compare the synchronization of two signals.

$$\gamma = \left| \frac{1}{N} \sum_{j=1}^N e^{i[\varphi_x(t_j) - \varphi_y(t_j)]} \right|$$

One additional measure has been recently introduced for the study of biological oscillators(12). Garcio-Ojalvo et al.(12) have suggested one measure of synchronization ( $R_G$ ), the variance in the measured mean of the oscillator  $X_{ij}$  along a trajectory (in time) divided the variance in  $X_{ij}$  of an individual trajectory:

$$R_G = \frac{<[(n^{-1}\sum \bullet X_{ij})^2] - <(n^{-1}\sum \bullet X_{ij})>^2>}{(<[X_{ij}]^2> - <[X_{ij}]>^2)}, \quad (1)$$

where  $n$  is the number of cell trajectories observed. If the cells are out of phase, then the numerator of  $R_G$  will be small relative to the denominator due to phase cancellation. On the other hand, a high degree of synchronization will essentially reduce observations on  $n$  cells to the same as looking at one cell, and the measure  $R$  will then be near 1. This measure does have then a simple interpretation.

There is also one time-tested measure of similarity that has been used in genetics for over 150 years, the intraclass correlation(13). Let us suppose the measurements on the state of oscillator  $i$  within the  $j$ th neighborhood at one time point are denoted by  $X_{ij}$ . A neighborhood is collection of cells that are neighbors. Only oscillators within a neighborhood can be coupled. We observe each cell within a neighborhood every half hour over a ten day period. We can think of neighborhoods of the same size as replicate experiments.

The measurements on the cells are assumed to have the following covariance structure

$$\begin{aligned} \text{COV}(X_{ij}, X_{kl}) &= 0, \quad i \neq k \text{ (different neighborhoods)} \\ &= \rho\sigma^2 \quad i = k, j \neq l \text{ (different cells within a neighborhood)} \end{aligned}$$

(2)

$$= \sigma^2 \quad i = k, j = l \text{ (same cell within a neighborhood)}$$

Stochastic oscillators within a neighborhood under this model can communicate and are coupled. The parameter  $\rho$  is the intraclass correlation and is a direct measure of the coupling of oscillators.

The intraclass correlation has a distinguished history in genetics(13) as a heritability measure (in which neighborhoods are families) and two simple interpretations. Firstly, the intraclass correlation is the correlation between measurements on different cells within the same neighborhood. Second, the intraclass correlation can also be interpreted as the fraction of variation in the oscillator explained within neighborhoods relative to that between neighborhoods. In this way the intraclass correlation has a built-in internal control to compare the communication between oscillators within neighborhoods relative to the variance in oscillators across replicate neighborhoods of the same size (a). We will denote the sample estimate of the intraclass correlation by  $R$ . This estimate will depend not only on the size of the neighborhoods (a), but will also vary with time. So, the intraclass correlation should be thought of as a synchronization surface varying both in time and neighborhood size. In other words, there will be trios, quartets, and quintets all singing different pieces simultaneously and over time, and these ensembles of musicians will be replicated.

There is a substantial body of theory for inference about the intraclass correlation(13). The total variation in oscillator behavior ( $X_{ij}$ ) at any particular time point can be partitioned into a between

neighborhood sum of squares ( $SS_B$ ) and a within neighborhood sum of squares ( $SS_W$ ) in Table 1 such that the total sum of squares  $SS_T = SS_B + SS_W$ . The ratio of Estimated Mean Squares (EMS) provide a test of the significance of the intraclass correlation  $\rho$  known as an F-test.

**Table 1.** Analysis of Variance (ANOVA) of oscillators between and within neighborhoods is used to estimate the intraclass correlation  $\rho$ . The model in (2) is a balanced variance components model with cells grouped by neighborhoods (13). The number of cells in a neighborhood is denoted by  $a$ ; the number of neighborhoods is denoted by  $n$ . The variance in a single oscillator measurement is  $\sigma^2$  and the intraclass correlation is  $\rho$ . The intraclass correlation is a measure of the fraction of total variation ( $SS_T$ ) within neighborhoods as well as the correlation between oscillator measurements on different cells within the same neighborhood. The quantities,  $\bar{X}_{B_i}$  and  $\bar{X}_{BAR}$ , are sample means over neighborhood  $i$  and over cells and neighborhoods, respectively. The estimated mean square (EMS) is a sum of squares (SS) divided by its degrees of freedom. The EMS estimates its expectation in the last column. Setting the EMS equal to their expectations allows the solution for  $\rho$ .

Source	degrees of freedom (df)	Sums of Squares (SS)	Estimated Mean Square (EMS)	Expectation of EMS
Between neighborhoods	$n-1$	$SS_B = a \sum_i (\bar{X}_{B_i} - \bar{X}_{BAR})^2$	$SS_B/(n-1)$	$\sigma^2 + (a-1)\rho\sigma^2$
Within neighborhoods	$n(a-1)$	$SS_W = \sum_i \sum_j (X_{ij} - \bar{X}_{B_i})^2$	$SS_W/n(a-1)$	$(1-\rho)\sigma^2$
Corrected total	$na-1$	$SS_T = \sum_i \sum_j (X_{ij} - \bar{X}_{BAR})^2$		

From the Analysis of Variance (ANOVA) Table 1, an estimator of the intraclass correlation can be constructed by setting the Estimated Mean Squares (EMS) equal to their expectations. The result is two equations in two unknowns,  $\rho$  and  $\sigma^2$ . Solving these two equations yields the estimator  $R$ :

$$R = (EMS_B - EMS_W) / (EMS_B + (a-1)EMS_W) \quad (3)$$

The estimator  $R$  defines a surface describing the synchronization of the oscillators in the system as a function of time ( $t$ ) and neighborhood size ( $a$ ). This synchronization measure has the four properties desired of a synchronization measure outlined earlier.

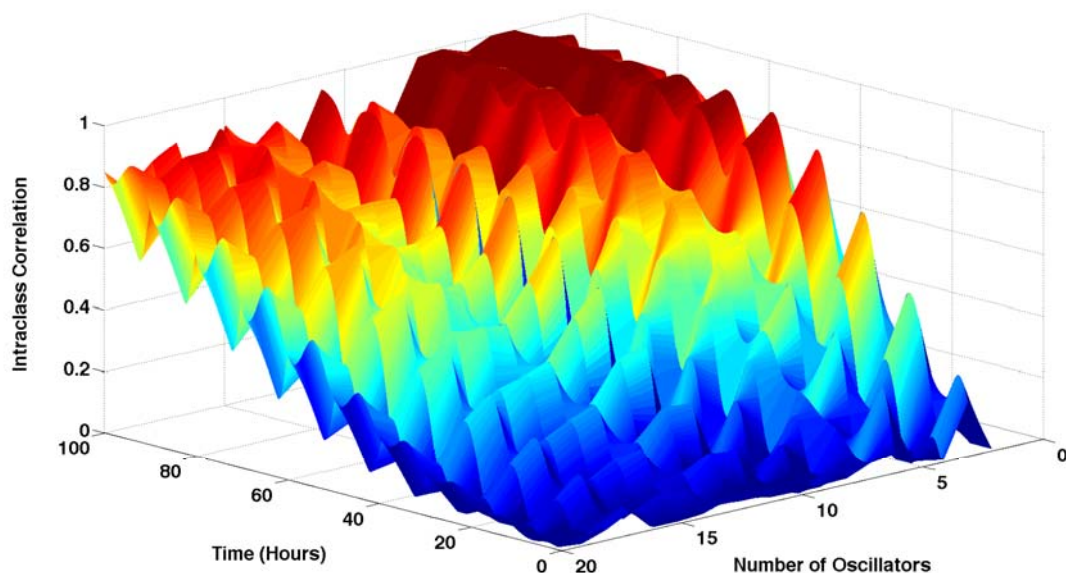
### 3. Kuramoto phase locking model

Kreuz et al. (11) concluded from a comparative analysis of varied synchronization measures in different physical systems that no one synchronization measure was to be preferred in different physical systems, but the modeling context determined the appropriate synchronization measure to be used in a particular physical system. One of the classic physical models of phase locking is the Kuramoto model (14). If the oscillators within cells are constrained to a ring indexed by  $r$ , then neighboring oscillators communicate. The model is defined by the following equations. For  $s \in [1, N]$ ,

$$\dot{\phi}_s = \omega_s + \frac{K}{N} * \sum_{r=1}^N \sin(\phi_r - \phi_s)$$

The parameters used here are  $\omega_s = \frac{2*\pi}{20}$ ,  $K = 0.05$ . The mapping onto the fluorescent measurements is  $X = \sin \phi$ .

In Figure 2 is displayed the synchronization surface for the Kuramoto phase-locking model using the intraclass correlation  $R$  in Eqn. (3). The synchronization surface (Figure 2) has substantial structure. The measure increases with time for different neighborhood sizes. There is also an alternating structure to the surface, where even numbers of neighbors have ridges, while odd numbers of neighbors have valleys. There is more synchronization with an even number of singers, such as a quartet, than for example, a trio.



**Figure 2.** Synchronization surface for the Kuramoto phase-locking model has substantial structure as a function of time ( $t$ ) and neighborhood size ( $a$ ).

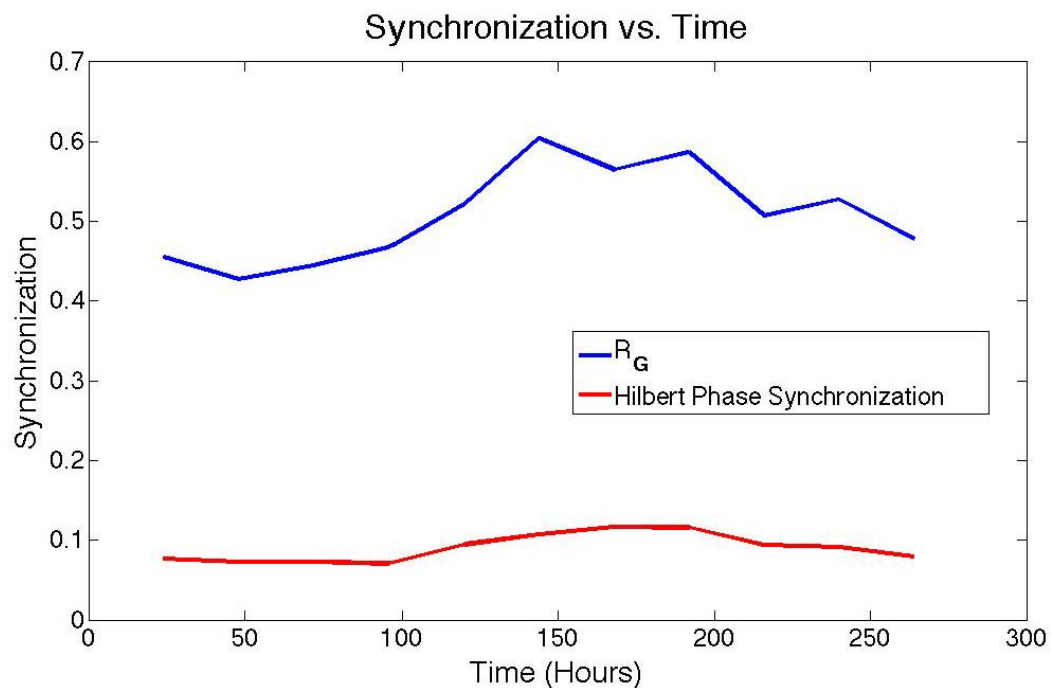
#### 4. Measuring synchronization

Synchronization was measured on a “star field” of over 25,000 cells in neighborhoods varying in size from 2 to 47 cells (Table 2).

Table 2. Number of Replicate Neighborhoods with Neighborhood size.

Number of Neighborhoods (n)	Number of cells per neighborhood (a)
670	2
436	3
422	4
336	5
310	6
281	7
244	8
224	9
186	10
162	11
120	12
88	13
95	14
66	15
53	16
44	17
29	18
29	19
16	20
13	21
13	22
6	23
14	24
7	25
4	26
2	27
2	29
3	30
4	31
2	32
2	46
2	47

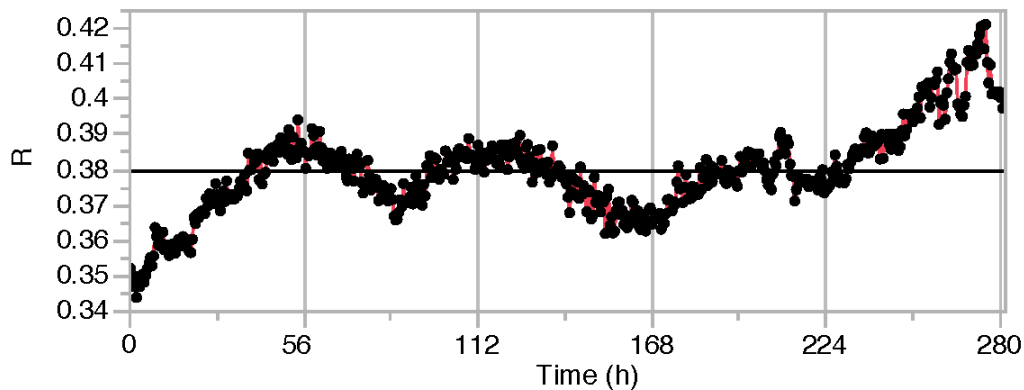
We computed the two older measures of synchronization, the phase of the Hilbert Transform and the  $R_G$  measure in (1) for neighborhoods of size 4 (Figure 3) using a one day non-overlapping window. We can see that both measures are roughly increasing with time. These measures only capture one aspect of the similarities of oscillator trajectories, their phase, and do not factor in the between replicate variability across neighborhoods.



**Figure 3.** Synchronization measures increase with time for a neighborhood of size 4. Measures graphed are the phase extracted from the Hilbert transform of each trajectory and the synchronization measure  $R_G$  in (2).

One of the advantages of the intraclass correlation as a measure of similarity between trajectories is that it also measures the fraction of variation explained by communication within neighborhoods. For example, the intraclass correlation (Figure 5A below) asymptotes at ~50% for most neighborhoods, implying that ~50% of the oscillator variation can be attributed to that within neighborhoods and only 50% to variation between neighborhoods. This begs the question of why oscillator behavior is so similar within neighborhoods.

In particular with the inference tools available about the intraclass correlation(13), we can ask questions about the significance of the synchronization. For example, is there any evidence of synchronization when there are only 2 oscillators in a neighborhood (Figure 4)? We constructed the ANOVA tables (from Table 1) for time point 0 and time point 282 hours (Tables 3 and 4). The synchronization was highly significant with  $R = 0.35$  at time 0 and  $R = 0.40$  at time = 282 h. All synchronization values (Figure 4) are significant at all time points even after a multiple test correction.



**Figure 4.** The intraclass correlation  $R$  is significant for all time points for neighborhoods with 2 oscillators even after a multiple test correction. This is a slice through Figure 5A.

**Table 3.** Analysis of Variance (ANOVA) of oscillators between and within neighborhoods is used to test the significance of the intraclass correlation  $\rho$  for neighborhoods of size 2 at time point 0. The method of computation of the ANOVA is given in Table 1. The value is significant at least the 0.001 level (\*\*\*).

Source	degrees of freedom (df)	Sums of Squares (SS)	Estimated Mean Square (EMS)	Expectation of EMS	F
Between neighborhoods	670-1	84639.67452	126.5167033	$\sigma^2 + (a-1)\rho\sigma^2$	
Within neighborhoods	670(2-1)	40642.82316	60.66093009	$(1-\rho)\sigma^2$	0.48***
Corrected total	(670)(2)-1	125282.4977			

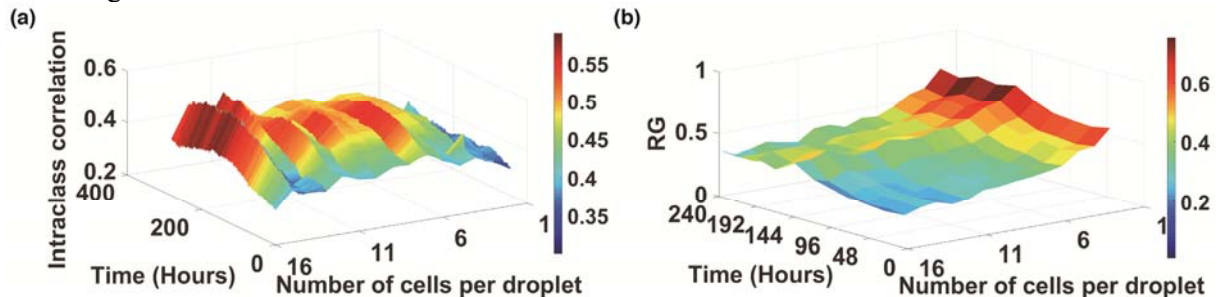
**Table 4.** Analysis of Variance (ANOVA) of oscillators between and within neighborhoods is used to test the significance of the intraclass correlation  $\rho$  for neighborhoods of size 2 at time point 282 h. The method of computation of the ANOVA is given in Table 1.

Source	degrees of freedom (df)	Sums of Squares (SS)	Estimated Mean Square (EMS)	Expectation of EMS	F
Between neighborhoods	670-1	26295.44231	39.30559389	$\sigma^2 + (a-1)\rho\sigma^2$	
Within neighborhoods	670(2-1)	11364.75646	16.96232307	$(1-\rho)\sigma^2$	0.43***
Corrected total	(670)(2)-1	37660.19877			

As an alternative, we constructed the synchronization surface as described in Table 1 and (3) for varying time and neighborhood size (Figure 5A). We can see that the synchronization surface highly resembles that produced by the Kuramoto phase-locking model (Figure 2). The ridges and



valleys are evident. The ridges in synchronization become higher as time advances. As a comparison we also constructed  $R_G$  (Figure 5B). There appears to be an increase of synchronization with time and drop off with neighborhood size.



**Figure 5.** The measured synchronization surface of >25,000 oscillators resembles that of the Kuramoto phase-locking model. (A) Synchronization surface for  $R$  in (3). (B) Synchronization surface for  $R_G$  in (1) using a one day window.

## 5. Discussion

The degree of synchronization not only varies with time, but also with the number of neighbors. There is a social environment that affects the behaviour of oscillators. There is structure to the synchronization surface as well (Figure 5). Whether there are an odd or even number of neighbors has an impact on the extent of synchronization. The synchronization will be hard to measure unless the observer controls for neighborhood size (Figure 5).

The intraclass correlation in (3) is readily interpreted in two ways. One, the synchronization measure  $R$  asymptoting in time at ~50% is the correlation in individual trajectories within a neighborhood. The final correlation between trajectories is quite high relative to the variation between neighborhoods. The ~50% is also the fraction of variation in oscillator trajectories captured within neighborhoods measured relative to the variation between replicate neighborhoods. The measure also allowed us to test how synchronization arises (Figure 4).

As synchronization takes place with time, there is not a sharp transition as in many physical and chemical systems(15). The synchronization is gradual as captured in the intraclass correlation. There is not an abrupt transition to synchronization as seen in some physical systems. It is possible that there is another state of the system as determined by cell density and the degree of coupling of cell oscillators that a sharp phase transition would occur(15). The neighborhoods, such as  $a=2$  or the large neighborhoods ( $a=42$ ), display non-monotone synchronization because of noise or the small number of replicate neighborhoods to average over to produce the surface value. For the other neighborhoods, the synchronization surface is increasing in time.

The synchronization surface (Figure 5) is also a rich source for testing various physical and biological models of communication. The Kuramoto phase locking model reproduces the structure in the synchronization surface, implying that there is a local phase-locking mechanism in operation between oscillators in cells within neighborhoods. Further information could be obtained by ensemble fitting of the Kuramoto model to the trajectories of individual oscillators(2, 16).

The problem of measuring synchronization in physics and biology is a general one. The collective behavior of individual actors and their synchronization is observed in a variety of biological and physical systems. One recent example is the highly synchronized marching of locusts or the behavior of human crowds(16, 17). The study of synchronization in physical and biological problems will continue to be challenging experimentally and discerning its underlying mechanism(s), a fundamental problem in both physics and biology.

## References

1. D. Battogtokh, D. K. Asch, M. E. Case, J. Arnold, H. B. Schuttler, An ensemble method for identifying regulatory circuits with special reference to the *qa* gene cluster of *Neurospora crassa*. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 16904-16909 (2002).
2. Y. Yu *et al.*, A genetic network for the clock of *Neurospora crassa*. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 2809-2814 (2007).
3. W. Dong *et al.*, Systems biology of the clock in *Neurospora crassa*. *PloS one* **3**, e3105 (2008).
4. A. Al-Omari *et al.*, Discovering Regulatory Network Topologies Using Ensemble Methods on GPGPUs With Special Reference to the Biological Clock of *Neurospora crassa*. *IEEE Access* **3**, 27-42 (2015).
5. S. H. Strogatz, Exploring complex networks. *Nature* **410**, 268-276 (2001).
6. B. Abbott *et al.*, Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters* **116**, 061102 (2016).
7. A. T. Winfree, On emerging coherence. *Science* **298**, 2336-2337 (2002).
8. O. Mondragon-Palomino, T. Danino, J. Selimkhanov, L. Tsimring, J. Hasty, Entrainment of a population of synthetic genetic oscillators. *Science* **333**, 1315-1319 (2011).
9. Z. Deng *et al.*, Single cell measurements on the biological clock by microfluidics. *Proc. of the 18th International Conference on Miniaturized Systems for Chemistry and Life Sciences (MicroTAS)*. San Antonio, TX **18**, 881-883 (2014).
10. J. D. Challenger, A. J. McKane, Synchronization of stochastic oscillators in biochemical systems. *Physical Review E* **88**, 012107 (2013).
11. T. Kreuz *et al.*, Measuring synchronization in coupled model systems: A comparison of different approaches. *Physica D: Nonlinear Phenomena* **225**, 29-42 (2007).
12. J. Garcia-Ojalvo, M. B. Elowitz, S. H. Strogatz, Modeling a synthetic multicellular clock: repressilators coupled by quorum sensing. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 10955-10960 (2004).
13. S. R. Searle, G. Casella, C. E. McCulloch, *Variance Components*. (John Wiley & Sons, 2009), vol. 391.
14. Y. Kuramoto, *Chemical oscillations, waves, and turbulence*. (Springer Science & Business Media, 2012), vol. 19.
15. A. F. Taylor, M. R. Tinsley, F. Wang, Z. Huang, K. Showalter, Dynamical quorum sensing and synchronization in large populations of chemical oscillators. *Science* **323**, 614-617 (2009).
16. D. J. Sumpter, *Collective animal behavior*. (Princeton University Press, 2010).
17. J. Buhl *et al.*, From disorder to order in marching locusts. *Science* **312**, 1402-1406 (2006).