Transient dynamics of pulse-coupled oscillators with nonlinear charging curves

Kevin P. O'Keeffe

Phys. Rev. E **93**, 032203 — Published 4 March 2016

DOI: 10.1103/PhysRevE.93.032203

Transient dynamics of pulse-coupled oscillators with nonlinear charging curves

Kevin P. O'Keeffe
Department of Physics, Cornell University, Ithaca, NY 14853, USA
(Dated: February 16, 2016)

We consider the transient behavior of globally coupled systems of identical pulse coupled oscillators. Synchrony develops through an aggregation phenomenon, with clusters of synchronized oscillators forming and growing larger in time. Previous work derived expressions for these time dependent clusters, when each oscillator obeyed a linear charging curve. We generalize these results to cases where the charging curves have nonlinearities.

I. INTRODUCTION

During each heartbeat, thousands of pacemaker cells discharge in concert. This collective firing causes the contraction of cardiac muscles, which pump blood around the body. Should these firing fall out of step, heartbeats can become erratic, which inhibits blood flow. In order to maintain healthy heart function, the pacemaker cells must maintain their synchronous firing.

In 1975, Peskin gave the first mathematical analysis of the pacemaker as a self-synchronizing system [1]. He modeled the pacemaker cells as leaky 'integrate-and-fire' oscillators that communicate with each other by firing sudden impulses. He then conjectured that a population of identical leaky oscillators with all-to-all pulsatile coupling would self-organize into synchrony for all $N \geq 2$ and for almost all initial conditions. Mirollo and Strogatz [2] later proved this conjecture.

Since then, pulse-coupled oscillators have been used as models in many other contexts, for example, sensor networks [3], low-powered radio transmission [4], firing neurons [5, 6], earthquakes [7], and economic booms and busts [8]. For greater realism, the associated theoretical work relaxes Peskin's original assumptions, by allowing for example local coupling in lattices [9–11] or networks [12–15]. These effects lead to new phenomena, such as traveling waves, self-organized criticality, partial synchrony, and coexistence. The inclusion of interactions with delays and different sign [16–18] have also been considered, which give rise to multi-stable clustering, transient clustering, phase-lagged synchronization.

Yet even within the simplified context of Peskin's all-to-all model, unanswered theoretical questions remain. In particular, little is known about transient dynamics: in a self-synchronizing system, what does the buildup to synchrony look like? A first step in this direction was presented in [19]. It was shown that synchrony developed through clustering; oscillators start to synchronize in small groups, which grow steadily larger over time. Using tools from aggregation theory [20], this clustering was described quantitatively. In the analysis, it was assumed that each oscillator had a linear charging curve. This idealization is appropriate for electronic oscillators such as those in sensor networks, but not for biological oscillators, like the aforementioned cardiac pacemaker cells or firing neurons. We here extend the analysis in [19] to explore the manner in which these more complicated oscillators achieve synchrony.

II. THE MODEL

We consider $N \gg 1$ identical oscillators coupled all-to-all. Each oscillator is characterized by a voltage-like state variable x_i, which increases from a baseline value of 0 to a threshold set to 1, according to

$$\dot{x_i} = S_0 - \gamma x_i. \tag{1}$$

When an oscillator reaches threshold it does two things: (i) It fires a pulse of size $1/N$. This pulse is received by all other oscillators instantaneously, causing them to discontinuously raise their voltage from x_j to $\min(x_j + 1/N, 1)$. This way, oscillators never exceed the threshold value of 1. To avoid complications with chain reactions of firing oscillators, we assume any oscillators which reach threshold by receiving a pulse, do not themselves fire until the *next* time they reach threshold. (ii) The firing oscillator then resets its voltage to 0, along with any secondary oscillators that were brought to threshold. These oscillators then begin their next cycle synchronized.

If $j > 1$ oscillators reach threshold together, each one fires, so that the pulse has total size j/N (although we later consider other types of pulse).

We note that there is some parameter redundancy, since by rescaling time we could set $S_0 = 1$ without loss of generality. For reasons that will become clear later, a different choice of S_0 is more convenient, so we leave it as a free parameter for now. We remark however that S_0 must be chosen so that $\dot{x_i} > 0$ for $0 \leq x_i \leq 1$.

III. RESULTS

Assume the initial voltages of the oscillators are drawn uniformly at random. How will the dynamics unfold? At the beginning, the oscillators simply increase their voltage according to $\dot{x_i} = S_0 - \gamma x_i$. Then the first oscillator reaches threshold, fires a pulse, and perhaps brings some other oscillators to threshold. As described,

these oscillators begin their next cycle in step. The primary, firing oscillator, and the secondary oscillators it incited to threshold, form a synchronous *cluster*.

As time goes on, other oscillators start firing pulses and start absorbing oscillators which are close enough to threshold. More clusters of synchronized oscillators emerge. In turn, these clusters start reaching threshold and absorbing *other* clusters, growing progressively larger. We note that clusters can only ever increase in size. They can never break apart because (a) the oscillators are identical, and therefore sync'd oscillators have the same speed, and (b) all oscillators receive the same number of pulses (thanks to the global coupling).

The picture is now clear; the system synchronizes through an aggregation phenomenon. Clusters of sync'd oscillators form and get steadily bigger by coalescing with each other. At any time t therefore, there are clusters of various sizes. Let $N_j(t)$ denote the number of clusters of size j at time t: N_1 is the number of singletons, N_2 is the numbers doublets, and so on. These N_j are correlated random quantities. They are correlated because oscillators belonging to clusters of one size are unavailable to clusters of another size, and they are random because of the initial conditions.

To analyze the system's dynamics, we imagine $N_j \gg 1$ so that fluctuations from different realizations of the system are small. Of course, this condition cannot be satisfied for every j, at all t. For example, at the final stages of the process, there will be a small number of very large clusters. We therefore restrict our attention to the portion of the process where $N_j \gg 1$ is approximately true – the opening and middle game, as opposed to the end game.

But how does the end game play out? That is, how does the process terminate? Strogatz and Mirollo [2] showed that for $\gamma > 0$ and pulse size $> 1/N$, then full sync is guaranteed for all IC except for a set of measure zero; the clustering continues until there is one giant cluster of size N. For other values of γ and other pulse sizes, full sync is possible, but not certain to occur.

In this work, we focus only on the transient dynamics, the evolution *to* synchrony. So from now on we implicitly assume we in the early and middle stages of the process, where $N_j \gg 1$ is a valid approximation. We then use ensemble averages to define the *individual cluster densities*,

$$c_j := \langle N_j \rangle / N. \tag{2}$$

We then make the following strong assumptions: (i) fluctuations about the ensemble averages are small, $N^{-1}N_j = c_j + O(N^{-1/2})$, and that (ii) different cluster densities are asymptotically uncorrelated, $N^{-2}N_iN_j = c_ic_j + O(N^{-1/2})$.

We can use these c_j to define a disorder parameter for our system. This is the *total cluster density*,

$$c = \sum_{j=1}^{N} c_j \tag{3}$$

where the index runs over all cluster *sizes*, which range from 1 to N. This density is a measure of the total fragmentation of the system, which we interpret as a kind of disorder. To see this, consider that at $t = 0$, there are N singletons, so $c_1 = 1$, and $c_j = 0$, $\forall j \neq 1$. This means that $c(0) = 1$, correctly identifying that the system begins maximally disordered. At the other extreme as $t \to \infty$, we know there is one giant cluster of size N, so $c = 1/N \approx 0$ for large N. Hence c decreases from 1 to 0 as the system evolves from complete disorder to full synchrony.

A. Total Cluster Density

We first analyse c. It obeys the following rate equation, where R_i is the rate at which clusters of size i fire, and L_i is the number of clusters absorbed during such a firing, for $i = 1, \ldots, N$ (i.e. over all cluster *sizes*):

$$\dot{c} = -\sum_i R_i(t)L_i(t) \tag{4}$$

To find $L_i(t)$, we first define the 'voltage-density' $\rho_j(x,t)dx$ to be the number of j-clusters with voltage between x and $x + dx$ at time t. This has the normalization condition $\int_0^1 \rho_j(x,t)dx = N_j$. Now, when an i-cluster fires, all clusters on the interval $[1 - i/N, 1)$ will be absorbed. This means,

$$L_i(t) = \sum_j \int_{1-i/N}^{1} \rho_j(x,t)dx. \tag{5}$$

We digress briefly to comment on difficulties imposed by voltage density $\rho_j(x,t)$. The nonlinearities in the oscillators' charging curves make $\rho_j(x,t)$ behave non-trivially. It is this key fact which makes nonlinear charging curves much harder to analyze than linear ones. In the linear case, (i.e. when $\gamma = 0$), as considered in [19], $\rho_j(x,t)$ is well approximated by a uniform density. This substantially simplified the calculation of L_i and R_i, and in turn the remainder of the analysis. For instance, (5) reduces to simply $L_i = \sum_j (i/N)N_j = ic$. But when $\gamma \neq 0$, $\rho_j(x,t)$ is an unknown quantity which obeys a complicated PDE. Approximately solving this PDE is a key result of the paper.

We now return to our calculation of L_i. To proceed, we make an approximation. As stated earlier, we are only interested in transient time scales – the opening and middle game. In this regime, most clusters will be small relative to the system size: $j \ll N$. This lets us approximate

the integral above, $\int_{1-i/N}^{1} \rho_j(x,t)dx \approx (i/N)\rho_j(x=1,t)$. Of course, this approximation will get worse as time goes on. We discuss this further in Section IV. Our expression for L_i is then

$$L_i(t) = \frac{i}{N}\sum_j \rho_j(x=1,t). \tag{6}$$

To continue the analysis, we need to find $\rho_j(x,t)$. As mentioned, its behavior is complicated so we defer its calculation, and instead find the firing rate R_i. Naively, one might think that this is simply the flux of i-clusters at threshold: $N^{-1}(\rho_i v)|_{x=1}$ (where N^{-1} is required, since R_i measures the rate of firing of c_i, not N_i). However not every cluster that reaches threshold gets the chance to fire, since some will be absorbed. To account for this effect, we decompose the rate into

$$R_i = R_i^0 - R_i^a. \tag{7}$$

The term R_i^0 is a 'background' firing rate, where we pretend all oscillators get to fire even if they are absorbed. R_i^a is the rate at which i-clusters are *being* absorbed by other clusters of various sizes, and hence deprived of their chance to fire.

We start with R_i^0. To be clear, by background firing rate, we mean the rate i-clusters would fire at, if *every* oscillator fired a pulse when it reached threshold. That is, imagine relaxing our imposition that any secondary oscillators that reach threshold by virtue of a pulse do not fire. In that case,

$$R_i^0 = N^{-1}(\rho_i v)\big|_{x=1}. \tag{8}$$

The speed v of each cluster is non-trivial. This is because in addition to its natural speed $v_0 = \dot{x} = S_0 - \gamma x$, each oscillator receives a steady stream of pulses from firing clusters which increase its voltage:

$$v(x,t) = v_0(x) + v_{pulse}(t). \tag{9}$$

This "pulse velocity" due to the firing of just j-clusters will be (absolute number of pulses per sec) $\times$ (distance per pulse). Since R_j is the firing rate of c_j, $R_j N$ is the absolute number of pulses, while the distance per pulse is j/N. To find the total pulse speed we then sum over all j-clusters: $\sum_j (NR_j)(j/N)$, giving

$$v_{pulse}(t) = \sum_j jR_j(t). \tag{10}$$

Our next target is the absorption rate R_i^a. The calculation is similar to finding L_i, and is given by $R_i^a = \sum_j R_j \int_{1-j/N}^{1} \rho_i(x,t)dx$, which after approximating the integral as before gives,

$$R_i^a = \sum_j R_j(j/N)\rho_i(x=1,t). \tag{11}$$

Substituting R_i^0 and R_i^a into (7) finally gives

$$R_i = \frac{S_0 - \gamma}{N}\rho_i(x=1,t). \tag{12}$$

We now analyze $\rho_j(x,t)$. In principle, it satisfies the the continuity equation with appropriate terms for the absorption of j-clusters at threshold, and the formation of j-clusters from smaller clusters:

$$\dot{\rho_j} + \partial_x(v\rho_j) + Absorption + Gain = 0 \tag{13}$$
$$\partial_x(v\rho_j)|_{x=0} = \partial_x(v\rho_j)|_{x=1} \tag{14}$$

Solving this PDE is the hardest part of the analysis. The absorption and gain terms are the main problem, because they *couple* the voltage densities; through them, $\rho_j(x,t)$ depends all the other $\rho_k(x,t)$. This is because a j-cluster can be created or absorbed by the action of various combinations of other clusters. Enumerating these combinations is by itself difficult, not to mention understanding how they affect the PDE. On top of all that, there is also the non-smoothness of the oscillators' velocity at threshold (which discontinuously jumps from $\dot{x} = S_0 - \gamma$ at $x = 1$, to $\dot{x} = S_0$ at $x = 0$) to deal with.

We can however make progress by observing that the evolution of the system naturally divides into periods $\{T_n\}$. We define a period to be the time take for the full population of oscillators to complete a voltage cycle. More carefully, T_n is earliest time when every oscillator has completed n cycles.

We then solve the continuity equation during a given period, not worrying about what happens before or after. This lets us avoid the complication of the aforementioned non-smoothness of the oscillators' behavior at the boundaries. We also neglect the absorption term. As previously discussed, when an i-cluster fires, only oscillators on $(1-i/N, 1]$ get absorbed. This is a small interval for the 'opening' and 'middle' game we are considering. Hence the absorption term is 0 on most of $[0,1]$ and so we omit it.

But we still have to compute the gain term. As previously discussed, this is combinatorially intensive (we explicitly compute this term later, when calculating the individual cluster densities). However, we can neglect this cumbersome term entirely, by making the following key observation.

Looking at equations (6), (12), we see our desired quantities R_j and L_j depend only on the density of clusters at threshold: $\rho_j(x=1,t)$. Therefore, *during each period*, R_i and L_i are only affected by j-clusters which existed *at the start of that period*, which we call 'original' j-clusters. This is because any 'new' j-clusters won't reach threshold until the *next* period. By 'new', we mean (a)

j-clusters that fired during a period, didn't absorb any other clusters, and so returned to threshold, and (b) any j-clusters that were created by the firing and absorption of other smaller clusters.

So for the purposes of calculating $\rho_j(x=1,t)$ *during a given period*, there is a 'lightcone' between original and new j-clusters. We therefore need to solve the continuity equation for the original j-clusters only, for which the gain term is zero. The problem is then given by (15) below, where $v(x,t)$ is given by (9), $f_0(x)$ is the initial distribution of $\rho_j^{original}$, and the Heaviside functions $H(x), H(1-x)$ are included to confine the I.C. to the interval $[0,1]$.

$$\begin{aligned} &\dot{\rho_j}^{original} + \partial_x(v\rho_j^{original}) = 0 \\ &\rho_j^{original}(x,0) = f_0(x)H(x)H(1-x) \end{aligned} \tag{15}$$

We don't yet know the speed $v(x,t)$. However its structure, $v(x,t) = v_0(x) + v_{pulse}(t)$, lets us derive an approximate solution for $\rho_j^{original}(x,t)$ given by (16) below. The derivation of this key result and the definition of $\Gamma(x,t)$ are shown in the Appendix.

$$\rho_j^{original}(x,t) = e^{\gamma t} f_0\Big(\Gamma(x,t)\Big) H\Big(\Gamma(x,t)\Big) H\Big(1-\Gamma(x,t)\Big). \tag{16}$$

What all this means is, if we known $\rho_j^{original}(x,t)$ at the start of a period, then we know how it will evolve until that period ends. For later convenience, we introduce the following notation. Let $\tilde{x}$ denote that during a period, x is held fixed at its value at the start of that period: $\tilde{x} = x(t = T_n)$ for $T_n < t < T_{n+1}$.

We next make the strong assumption that clusters of all sizes are distributed uniformly in voltage on $[0,1)$ at the start of each period: $\rho_j(x, t = T_n) = \tilde{N}_j$. Then in (16), $f_0(x) = \tilde{N}_j$. We discuss the legitimacy of making this assumption in Section IV.

The Heaviside functions make (16) look complicated. But really, they only enforce that the $\rho_j^{original}$ is zero behind the final j-cluster (cluster with smallest voltage), and ahead of the first j-cluster (cluster with largest voltage). We remark that as it stands, the solution (16) propagates into the unphysical $x \geq 1$ regime. But we of course restrict our attention to just $x \in [0,1]$.

The behavior of $\rho_j^{original}(x,t)$ during each period is therefore simple. The density at each point x simply grows at rate $e^{\gamma t}$ until it drops discontinuously to 0, as the final 'original' j-cluster passes by. This behavior is shown in Figure 1.

Now that we have an expression for $\rho_j(x=1,t)$, which we have argued is $\rho_j^{original}(x=1,t)$, we can complete our expressions for L_i and R_i given by (6) and (12). We then plug the results into (4) to obtain our sought after rate equation for the disorder parameter $c(t)$,

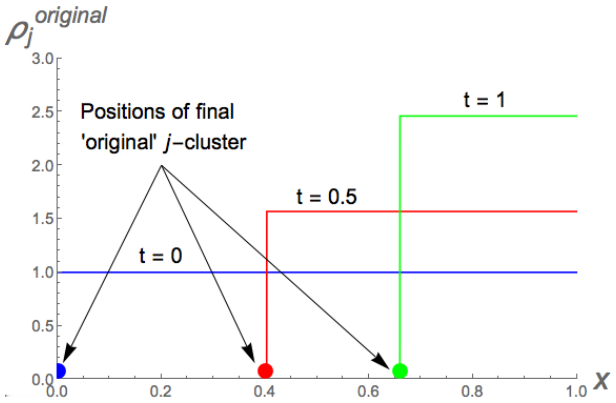

FIG. 1: Evolution of voltage density of original j-clusters during a period, with initial condition $\rho_j^{original}(x,0) = 1$.

$$\dot{c} = -(S_0 - \gamma)e^{2\gamma t}\tilde{c}. \tag{17}$$

which has solution,

$$c(t) = \frac{\tilde{c}}{2\gamma}\Big(S_0 + \gamma + e^{2\gamma t}(\gamma - S_0)\Big). \tag{18}$$

We restate that equations (17) and (18) are only valid during a given period. We can however use (18) to find $c(t)$ for all t, by stitching solutions during successive periods together.

But we still don't know the periods $\{T_n\}$ themselves. To find them, we need the speed v as per (9). Recalling $v_0 = S_0 - \gamma x$, and substituting R_i from (12), gives

$$v(x,t) = (S_0 - \gamma x) + (S_0 - \gamma)e^{\gamma t}. \tag{19}$$

We see that v is the same during each period (i.e. there are no 'tilde' quantities, we denote different values during different periods.). This means that the length of each period is the same: $T_n = nT_0$. We can find T_0 from $T_0(S_0,\gamma) = \int_0^1 v(x)dx$. To compare the effects of different amounts of concavity on equal footing, we want $T_0 = 1$ for every γ. We can achieve this by selecting an appropriate value for S_0, which we have strategically left as a free parameter until now. Doing the integral, this value for S_0 is

$$S_0 = \frac{\left(e^{2\gamma} + 2e^{\gamma} - 1\right)\gamma}{\left(e^{\gamma} - 1\right)\left(e^{\gamma} + 3\right)}. \tag{20}$$

We must be careful when using (20). This is because for sufficiently negative γ, S_0 can become negative. While this choice of S_0 ensures the *total* speed $v = v_0 + v_{pulse}$ is positive, the natural speed $v_0 = \dot{x} = S_0 - \gamma x$ can become negative if S_0 is too negative. This means that the

oscillators *decrease* in voltage in the absence of coupling. We avoid this unphysical regime by requiring $v_0 > 0$ for $0 \le x \le 1$, which leads to $\gamma_{min} \approx -0.881$.

Figure 2 shows the agreement between theory and simulation for $c(t)$ when $\gamma < 0$ and $\gamma > 0$. For comparison, we also show when $\gamma = 0$, which corresponds to the linear charging curve studied in [19]. In the linear case, $c(t)$ is a series of line segments whose slope decreases by a factor of 2 from period to period. But when $\gamma \neq 0$, $c(t)$ has more complicated behavior; it no longer decays linearly during each period.

As can be seen, $c(t)$ declines faster and slower when $\gamma > 0$ and $\gamma < 0$ respectively. This makes physical sense. When $\gamma > 0$, oscillators slow down as they increase in voltage, which makes them clump closer together near $x = 1$. When $\gamma < 0$, the opposite happens; clusters spread further apart closer to threshold. Now suppose a j-cluster fires. When $\gamma > 0$ the interval $[1 - j/N, 1)$ is more likely to contain oscillators than when $\gamma < 0$, thanks to the 'clumping' and 'spreading out', which in turn makes an absorption more likely. The case of zero concavity then interpolates between these two regimes, as evidenced by Figure 2.

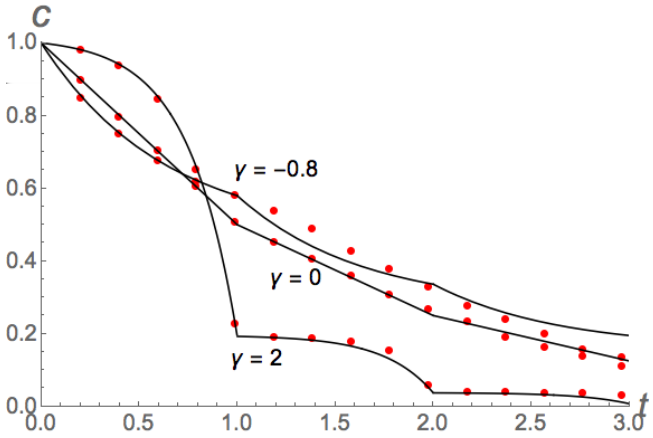

FIG. 2: (Color online) Theoretical and simulated $c(t)$ for $\gamma = 2$, $\gamma = 0$, and $\gamma = -0.8$. Solid lines show theoretical prediction (18), while data points show simulated results for $N = 5 \times 10^4$ oscillators.

We can use (18) for c to estimate the timescale of the transient dynamics. We say transience ends when a cluster of size $\sim N$ has formed, so that $c \sim 1/N$. Our assumptions will likely break down before this, so this is best interpreted as an upper bound. Looking at (18), we see that after one period, c decreases by a factor of,

$$B := \frac{1}{2\gamma}\Big(S_0 + \gamma + e^{2\gamma}(\gamma - S_0)\Big) = \frac{2}{e^{\gamma} + 3}. \tag{21}$$

After n periods, it decreases by B^n. After some algebra, and rounding T_{trans} to the nearest period, we get,

$$T_{trans} \sim \log N / \log B^{-1} \tag{22}$$

B. Individual Cluster Densities

How do the individual densities c_i evolve? We begin with the 1-clusters, whose density is c_1. They are the easiest density to analyze, since they can only decrease. There are two ways this can happen: (i) the loss of a *firing* singleton, when it absorbs other clusters of any size, and (ii) the loss of *absorbed* singletons, due to the firing of another cluster:

$$\dot{c_1} = \mathcal{L}_1^{firing} + \mathcal{L}_1^{absorbed}. \tag{23}$$

We begin with $\mathcal{L}_1^{firing}$. From (12) we know singletons fire at rate $R_1 = (S_0 - \gamma)e^{\gamma t}\tilde{c_1}$. During such a firing, an absorption will take place if there is at least one cluster on $[1 - 1/N, 1)$. This interval contains on average $Nc \times 1/N = c(t) = \tilde{c}e^{\gamma t}$ clusters. Further, the probability that it contains n clusters is given by the Poisson distribution: $\Pi_n = \frac{(\tilde{c}e^{\gamma t})^n}{n!}e^{-\tilde{c}e^{\gamma t}}$. This is the mathematical statement that the clusters are distributed randomly without correlations. The probability that $[1 - 1/N, 1)$ is occupied by at least one cluster is therefore $1 - e^{-\tilde{c}e^{\gamma t}}$. If an absorption takes place, N_1 decreases by 1, since we're only considering the loss of the *firing* oscillator here. The expected loss rate is then $(S_0 - \gamma)e^{\gamma t}\tilde{c_1}[1 \times (1 - e^{-\tilde{c}e^{\gamma t}}) + 0 \times e^{-\tilde{c}e^{\gamma t}}]$, leading to,

$$\mathcal{L}_1^{firing} = (S_0 - \gamma)e^{\gamma t}\tilde{c_1}(1 - e^{-\tilde{c}e^{\gamma t}}). \tag{24}$$

To calculate $\mathcal{L}_1^{absorbed}$, imagine a j-cluster fires and absorbs all the singletons on the interval $[1 - j/N, 1)$. As before, this interval will have on average $Nc_1(t) \times j/N = j\tilde{c_1}e^{\gamma t}$ such singletons. Multiplying this by R_j and summing over j then gives $\sum_j (1 - \gamma)\tilde{c_j}e^{\gamma t} \times j\tilde{c_1}e^{\gamma t}$, which leads to

$$\mathcal{L}_1^{absorbed} = (1 - \gamma)e^{2\gamma t}\tilde{c_i} \tag{25}$$

Substituting $\mathcal{L}_1^{firing}$ and $\mathcal{L}_1^{absorbed}$ into (23) gives,

$$\dot{c_1} = -(S_0 - \gamma)\tilde{c_1}\Big[(1 + e^{\gamma t}) - e^{-\tilde{c}e^{\gamma t}}\Big]. \tag{26}$$

This looks intimidating, but since the quantities $\tilde{c_i}$ are held constant over each period, the R.H.S. is a function of only t. It therefore has an analytic solution, which we show plotted in Figure 3.

Will larger clusters behave similarly? They differ from the singletons in that they can be created as well as absorbed, which makes them harder to calculate. Their absorption rate is easily generalized from that of the singletons:

$$\mathcal{L}_i^{firing} + \mathcal{L}_i^{absorbed} = (S_0 - \gamma t)\tilde{c_i}\Big[(1 + e^{\gamma t}) - e^{-i\tilde{c}e^{\gamma t}}\Big]. \tag{27}$$

Their gain rate is calculated as follows. An i-cluster is created when a cluster of size $k < i$ fires, and absorbs the right combination of other clusters. Suppose there are a_1 1-clusters, a_2 2-clusters, ..., on the interval $[1 - k/N, 1)$. If $a_1 + 2a_2 + \cdots + k = i$, then an i-cluster will be created. Such a combination occurs with probability $\frac{(kc_1)1_1^a}{a_1!}e^{-kc_1} \times \frac{(kc_2)1_2^a}{a_2!}e^{-kc_2} \times \ldots$. Summing first over all such combinations, and then over all k, gives an expected rate gain of

$$\sum_{k=1}^{i-1}(S_0-\gamma)\tilde{c_k}e^{\gamma t}e^{-k\tilde{c}e^{\gamma t}} \sum_{a_1+2a_2+\cdots=i-k}\left(\prod_{p\geq 1}\frac{(k\tilde{c_p}e^{\gamma t})^{a_p}}{a_p!}\right) \tag{28}$$

After combining the loss and gain terms, and some algebraic manipulation, we finally obtain the desired rate equation for i-clusters,

$$\dot{c_i} = -(S_0-\gamma)e^{\gamma t}(1+e^{\gamma t})\tilde{c_i} + \sum_{k=1}^{i}(S_0-\gamma)\tilde{c_k}e^{\gamma t}e^{-k\tilde{c}e^{\gamma t}} \sum_{\sum pa_p=i-k}\left(\prod_{p\geq 1}\frac{(k\tilde{c_p}e^{\gamma t})^{a_p}}{a_p!}\right) \tag{29}$$

This is a set of recursive equations, and so we can solve them successively. As with c_1, the R.H.S. is a pure function of t, so analytic solutions are findable. Figure 3 shows theoretical predictions versus simulation results for c_1 through c_4 when $\gamma = 0.9$. We remark that the effect of a nonlinear versus linear charging curve on the c_i is the same as that for the disorder parameter c: it causes them to no longer decay/grow linearly during each period (note we do not show c_i for the linear charging curve, $\gamma = 0$, for illustrative purpose. See Fig. 6 in the supplemental materials of [19])

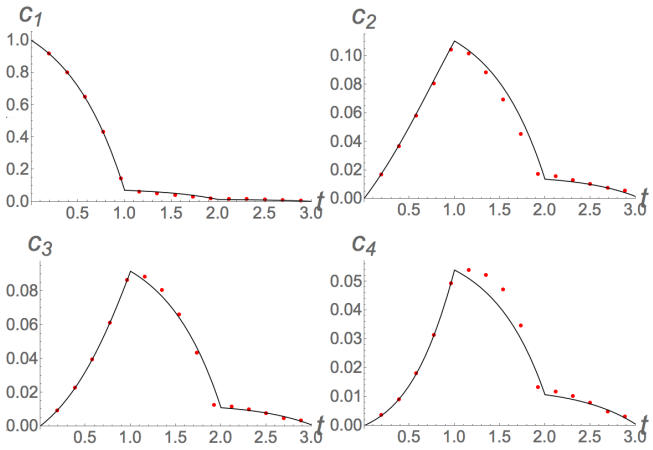

FIG. 3: (Color online) Theoretical and simulated cluster densities c_1 though c_4 for $\gamma = 0.9$. Solid black lines show analytic solutions to (29). Red data points show simulation results for 5×10^4 oscillators.

C. Alternate Coupling Rules

We thus far assumed an i-cluster fired a pulse of size i/N. We now consider two alternatives. The first is simply the original pulse strength with a tunable strength K: $(Ki)/N$. The second is a fixed pulse strength of K/N regardless of the size of the firing cluster. These alterations only modestly change the analysis, so we simply list the results for L_i, R_i, v_{pulse} and c in the table below, where S_0 is given by (20). For illustrative purposes we do not include a formula for c_i, but its calculation is straightforward.

	Variable Pulse: Kj/N	Fixed Pulse: K/N
L_i	$Ki\tilde{c}e^{\gamma t}$	$K\tilde{c}e^{\gamma t}$
R_i	$(S_0-\gamma)\tilde{c}_i e^{\gamma t}$	$(S_0-\gamma)\tilde{c}_i e^{\gamma t}$
v_{pulse}	$K(S_0-\gamma)e^{\gamma t}$	$K(S_0-\gamma)e^{\gamma t}\tilde{c}$
$\dot{c}$	$-K(S_0-\gamma)e^{2\gamma t}\tilde{c}$	$-K(S_0-\gamma)e^{2\gamma t}\tilde{c}^2$
$c(t)$	$\frac{\tilde{c}\left(K(\gamma-S_0)\left(e^{2\gamma t}-1\right)+2\gamma\right)}{2\gamma}$	$\frac{\tilde{c}\left(\tilde{c}K(\gamma-S_0)\left(e^{2\gamma t}-1\right)+2\gamma\right)}{2\gamma}$

As can be seen, there are mostly only minor differences between the two cases. The first thing to note is that $c(t)$ decays more slowly with a fixed pulse strength K/N. Intuitively, this is because large and small clusters now fire with the same strength, which means they absorb all clusters on the fixed interval $[1 - K/N, 1)$. In contrast, for a pulse strength Kj/N, bigger clusters fire bigger pulses, and therefore absorb clusters on an interval proportional to their size: $[1 - Kj/N, 1)$. This is mathematically manifested as $L_i^{variable\ pulse}$ being bigger than $L_i^{fixed\ pulse}$ by a factor of i, which leads to $\dot{c}^{variable\ pulse}$ being bigger than $\dot{c}^{fixed\ pulse}$ by a factor of $\tilde{c}^{-1}$ (remember, $0 \leq c \leq 1$). In turn, this means $c(t)^{fixed\ pulse}$ decays more slowly than $c(t)^{variable\ pulse}$

Also note that v_{pulse} depends on $\tilde{c}$ for the fixed pulse case. The mechanism discussed above is also at play here: since there are fewer clusters in successive periods, and the pulse per cluster is constant, the total 'current' per period will get smaller. This is in contrast to the pulse $= Kj/N$ case, where there are fewer clusters per period also, but larger clusters fire larger pulses, keeping the total 'current' per period constant. A consequence of this decrease in v_{pulse} is that the periods won't be constant for pulse $= K/N$, as there are for Kj/N. They will get longer as v_{pulse} decreases from period to period.

IV. BREAKDOWN OF APPROXIMATIONS

A. Uniformity Assumption

We now discuss the approximations and assumptions we made in our analysis. The first of these was that each cluster density was distributed uniformly in voltage at the start of each period, $\rho_j(x,t) = \tilde{N}_j$. From this, we

derived equations (6) and (12) for L_i and R_i, which in turn led us to our disorder parameter c.

This uniformity assumption clearly cannot be satisfied for each i, at every t. For instance, consider the end of the first period. Perhaps mostly clusters of size < 5 were formed, with only a few larger clusters of size > 10. Then, $\rho_{j<5}(x, t = 1)$ will be approximately uniform, but $\rho_j(x, t = 1)$ will be more sharply peaked. So the uniformity assumption is inaccurate for large clusters, which are few in number. This explains why (18) approximates $c(t)$ well. Since $c(t) = \sum_j c_j$, we see that the sum will be dominated by those c_i which are large, for which the uniformity assumption is accurate.

The fact that the uniformity assumption worsens for larger clusters also means that our results for c_i should get worse for larger i. Figure 4 below shows that this is indeed the case.

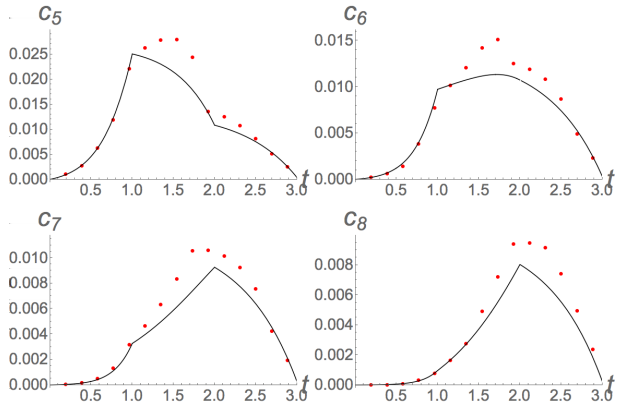

FIG. 4: (Color online) Theoretical and simulated cluster densities c_5 though c_8 for $\gamma = 0.9$. Solid black lines show analytic solutions to (29). Red data points show simulations results for 5×10^4 oscillators. As can be seen, theory and simulation start to disagree

B. Final stages of process

Throughout our analysis, we assumed $N_j \gg 1$. As discussed, this cannot be true $\forall j$, at every t. This assumption is most blatantly incorrect at the end of the process, where there are a small number of macroscopic clusters. Our results should thus substantially disagree with simulation for large t, as is evidenced by Figure 5.

V. CONCLUSION

We have studied the transient dynamics of pulse coupled oscillators with nonlinear charging curves. We derived approximations for the total cluster density $c(t)$ and individual cluster densities $c_i(t)$. These approximations

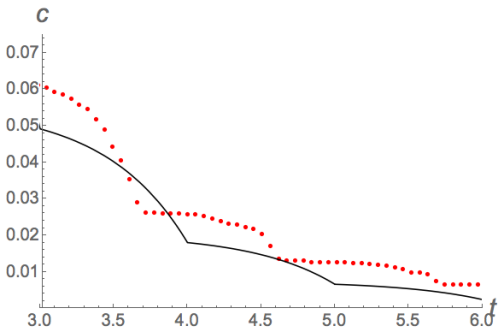

FIG. 5: (Color online) Theoretical and simulated total cluster density $c(t)$ for $\gamma = 0.9$ and $t > 3$. Solid black lines show analytic solution (2). Red data points show simulation results for $N = 10^4$ oscillators. There is a significant disagreement between theory and simulation for later times, when the approximations we made in the analysis breakdown.

were good up to the final stages of the process, where the assumptions made in the analysis breakdown.

Our work could be used to understand clustering in other systems of pulse-coupled oscillators. For instance, Ernst et al [17] reported multi-stable clusters for all-to-all, inhibitory coupling with delays. They found the average number N_c of clusters obeyed $N_c \sim \tau^{-1}$, where τ is the delay. Perhaps adjustments to our analysis could analytically recover this result; the pulse velocity (10) could be made negative to account for the inhibitory coupling, and 'delayed' versions of equations (5), (7) for the firing and loss rates R_i, L_i could be derived.

Furthermore, Mauroya and Sepulcher [15] studied the long term behavior of the system (1) with $\gamma > 0$: the complement to our 'opening' and 'middle' game. They analytically determined the final number of synchronized clusters formed (we remind the reader that when $\gamma > 0$, full synchrony is not *guaranteed* to occur, and so multiple, stable clusters are possible). Perhaps our transient analysis could be united with their steady state results to characterize the full evolution of the Peskin model.

Another possible application of our results is in network detection. Gomez-Gardenes et al [21] showed that transient clustering in the Kuramoto model can be used to approximate the underlying network structure. Could our results could be used to the same effect in networks of pulse-coupled oscillators? Local coupling would however mean that clusters could break apart as well as coalesce. One could account for this by including additional loss terms in our rate equations for c and c_i, (4), (29).

Our model has several idealizations that could be relaxed in future work. For example, local coupling, delayed coupling, and heterogeneity in oscillator speeds and pulse size could be studied. Another modification would be to allow chain reactions, by permitting any clusters that are brought to threshold by another firing cluster, to fire themselves.

We thank Steven Strogatz and Paul Krapivsky for helpful discussions.

VI. ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation through Grant No. DMS-1513179.

[1] C. S. Peskin, *Mathematical aspects of heart physiology* (Courant Institute of Mathematical Sciences, New York University, 1975).

[2] R. E. Mirollo and S. H. Strogatz, SIAM Journal on Applied Mathematics **50**, 1645 (1990).

[3] Y.-W. Hong and A. Scaglione, in *Ultra Wideband Systems and Technologies, 2003 IEEE Conference on* (IEEE, 2003), pp. 190–194.

[4] X. Y. Wang and A. B. Apsel, in *Circuits and Systems, 2007. MWSCAS 2007. 50th Midwest Symposium on* (IEEE, 2007), pp. 1524–1527.

[5] C. Kirst, T. Geisel, and M. Timme, Physical review letters **102**, 068101 (2009).

[6] G. B. Ermentrout and D. H. Terman, *Mathematical foundations of neuroscience*, vol. 35 (Springer Science & Business Media, 2010).

[7] C. H. Scholz, Bulletin of the Seismological Society of America **100**, 901 (2010).

[8] S. Gualdi, J.-P. Bouchaud, G. Cencetti, M. Tarzia, and F. Zamponi, Physical review letters **114**, 088701 (2015).

[9] P. C. Bressloff, S. Coombes, and B. de Souza, Phys. Rev. Lett. **79**, 2791 (1997), URL `http://link.aps.org/doi/10.1103/PhysRevLett.79.2791`.

[10] P. C. Bressloff and S. Coombes, Phys. Rev. Lett. **80**, 4815 (1998), URL `http://link.aps.org/doi/10.1103/PhysRevLett.80.4815`.

[11] A. Corral, C. J. Pérez, A. Díaz-Guilera, and A. Arenas, Phys. Rev. Lett. **74**, 118 (1995), URL `http://link.aps.org/doi/10.1103/PhysRevLett.74.118`.

[12] M. Timme, F. Wolf, and T. Geisel, Phys. Rev. Lett. **89**, 258701 (2002), URL `http://link.aps.org/doi/10.1103/PhysRevLett.89.258701`.

[13] N. Brunel, Journal of computational neuroscience **8**, 183 (2000).

[14] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, Physics Reports **469**, 93 (2008).

[15] A. Mauroya and R. Sepulcher, Chaos: An interdisciplinary journal of nonlinear science **18**, 037122 (2008).

[16] C. Van Vreeswijk, L. Abbott, and G. B. Ermentrout, Journal of computational neuroscience **1**, 313 (1994).

[17] U. Ernst, K. Pawelzik, and T. Geisel, Phys. Rev. Lett. **74**, 1570 (1995), URL `http://link.aps.org/doi/10.1103/PhysRevLett.74.1570`.

[18] W. Gerstner, Physical review letters **76**, 1755 (1996).

[19] K. P. O'Keeffe, P. L. Krapivsky, and S. H. Strogatz, Phys. Rev. Lett. **115**, 064101 (2015), URL `http://link.aps.org/doi/10.1103/PhysRevLett.115.064101`.

[20] P. L. Krapivsky, S. Redner, and E. Ben-Naim, *A kinetic view of statistical physics* (Cambridge University Press, 2010).

[21] J. Gómez-Gardeñes, Y. Moreno, and A. Arenas, Phys. Rev. Lett. **98**, 034101 (2007), URL `http://link.aps.org/doi/10.1103/PhysRevLett.98.034101`.

VII. APPENDIX

We here approximate the density $\rho_j^{original}(x,t)$. For convenience, we drop the superscript 'original'. As shown in the main body of the text, the density solves equation (30) below,

$$\begin{aligned} &\dot{\rho_j} + \partial_x(v\rho_j) = 0 \\ &\rho_j(x,0) = f_0(x)H(x)H(1-x) \end{aligned} \tag{30}$$

where, $f_0(x) = \tilde{N}_j$ (since we are assuming a initial uniform distribution), and

$$v(x,t) = v_0(x) + v_{pulse}(t). \tag{31}$$

While we know $v_0(x) = S_0 - \gamma x$, we don't yet have a complete expression for $v_{pulse}(t)$. In the main text, we derived $v_{pulse} = \sum_j jR_j$, which using (12) for R_i gives

$$v_{pulse}(t) = \sum_j \frac{S_0 - \gamma}{N} j\ \rho_j(x=1,t) \tag{32}$$

This is the source of our difficulty. Our PDE for $\rho_j(x,t)$ depends on v_{pulse}, which in turn depends on the voltage density for *every other* cluster size $\rho_k(x,t)$. To overcome this difficulty, we use a technique similar to the 'leap-frog' or 'split' method used in certain numerical schemes. This involves making a series of recursive approximations for v_{pulse} and ρ_j:

$$v_{pulse} = \left(v_{pulse}^{(0)}, v_{pulse}^{(1)}, v_{pulse}^{(2)}, \dots\right) \tag{33}$$

$$\rho_j = \left(\rho_j^{(0)}, \rho_j^{(1)}, \rho_j^{(2)}, \dots\right) \tag{34}$$

Graphically, our scheme is given by the following, where we have placed the labels of equations used to make the approximations over the arrows.

$$v_{pulse}^{(0)} \xrightarrow{(30)} \rho_j^{(0)} \xrightarrow{(32)} v_{pulse}^{(1)} \xrightarrow{(30)} \rho_j^{(1)} \xrightarrow{(32)} + \dots \tag{35}$$

We begin by setting $v^{(0)}_{pulse} = 0$. The speed is then,

$$v(x,t)^{(0)} = v_0(x) + 0 = S_0 - \gamma x. \tag{36}$$

We plug this into (30) and solve for resulting PDE for $\rho^{(0)}_j(x,t)$. This has solution,

$$\rho^{(0)}_j(x,t) = e^{\gamma t}\tilde{N}_j H\Big(\Gamma_0(x,t)\Big)H\Big(1-\Gamma_0(x,t)\Big). \tag{37}$$

with $\Gamma_0(x,t) = \gamma^{-1}[S_0 - (S_0 - \gamma x)e^{\gamma t}]$. We then use $\rho^{(0)}_j$ to find $v^{(1)}_{pulse}$ using (32), which gives

$$v^{(1)}_{pulse} = (S_0 - \gamma)e^{\gamma t}\tilde{c}. \tag{38}$$

This completes the first step of our scheme. We then repeat the process to find $\rho^{(1)}_j$ and $v^{(2)}_{pulse}$. We use $v^{(1)}_{pulse}$ to update the speed,

$$\begin{aligned} v(x,t)^{(1)} &= v_0(x) + v^{(1)}_{pulse} \\ &= (S_0 - \gamma x) + (S_0 - \gamma)e^{\gamma t}\tilde{c}. \end{aligned} \tag{39}$$

and then plug this into (30) to obtain a PDE for $\rho^{(1)}_j$, which we solve to get,

$$\rho^{(1)}_j(x,t) = e^{\gamma t}\tilde{N}_j H\Big(\Gamma_1(x,t)\Big)H\Big(1-\Gamma_1(x,t)\Big) \tag{40}$$

where $\Gamma_1(x,t) = \Big[\frac{3S_0-\gamma}{2\gamma} + \frac{e^{\gamma t}}{2\gamma}(2\gamma x - 2S_0) + \frac{e^{2\gamma t}}{2\gamma}(\gamma - S_0)\Big]$.

Looking at (37) and (40), we see that $\rho^{(0)}_j$ and $\rho^{(1)}_j$ have the same functional form. They only differ in the arguments of the Heaviside function: $\Gamma_0(x,t) \neq \Gamma_1(x,t)$. This remarkable similarity between $\rho^{(0)}_j$ and $\rho^{(1)}_j$ has an important consequence: it 'closes' our approximation scheme. We see this by substituting $\rho^{(1)}_j$ into (32) to find,

$$v^{(2)}_{pulse} = (S_0 - \gamma)e^{\gamma t}\tilde{c} = v^{(1)}_{pulse}, \tag{41}$$

which implies that $\rho^{(2)}_j = \rho^{(1)}_j$, which in turn implies our scheme terminates at $(v_{pulse}, \rho_j) = (v^{(2)}_{pulse}, \rho^{(1)}_j)$. Our final approximations for v_{pulse} and $\rho_j(x,t)$ are then,

$$v_{pulse}(t) \approx (S_0 - \gamma)e^{\gamma t}\tilde{c} \tag{42}$$

$$\rho_j(x,t) \approx e^{\gamma t}\tilde{N}_j H\Big(\Gamma(x,t)\Big)H\Big(1-\Gamma(x,t)\Big) \tag{43}$$

with $\Gamma(x,t) = \Big[\frac{3S_0-\gamma}{2\gamma} + \frac{e^{\gamma t}}{2\gamma}(2\gamma x - 2S_0) + \frac{e^{2\gamma t}}{2\gamma}(\gamma - S_0)\Big]$.

This concludes our analysis. We state bluntly that our approach is not rigorously justified. Its legitimacy is supported only by the agreement between our analytic results and numerical simulation. We hope future work will elucidate the cause of its efficacy.