

# Modeling and Cohesiveness Analysis of Midge Swarms

Kevin M. Passino<sup>1</sup>

Dept. Electrical and Computer Engineering,  
Dept. Evolution, Ecology, and Organismal Biology, and  
Mathematical Biosciences Institute,  
The Ohio State University,  
2015 Neil Avenue, Columbus, OH 43210-1272

Received \_\_\_\_\_; accepted \_\_\_\_\_

## Abstract

Midges (*Anarete pritchardi*) coordinate their flight motions to form a cohesive group during swarming. In this paper, individual midge motion dynamics, sensing abilities, and flight rules are represented with a midge swarm model. The sensing accuracy and flight rule are adjusted so that the model produces trajectory behavior, and velocity, speed, and acceleration distributions, that are remarkably similar to those found in midge swarm experiments. Mathematical analysis of the validated swarm model shows that the distances between the midges' positions and the swarm position centroid, and the midges' velocities and the swarm velocity centroid, are ultimately bounded (i.e., eventually satisfy a bound expressed in terms of individual midge parameters). Likewise, the swarm position and velocity centroids are shown to be ultimately bounded. These analytical results provide insights into why the identified individual midge sensing characteristics and flight rule lead to cohesive swarm behavior.

Keywords: *Anarete pritchardi* Kim (Diptera: Cecidomyiidae), midge, swarm, boundedness analysis.

---

<sup>1</sup>Bibliographic information: Passino, Kevin M., “Modeling and Cohesiveness Analysis of Midge Swarms,” Int. Journal of Swarm Intelligence Research, Vol. 4, No. 4, pp. 1–22, Oct.-Dec. 2013. The author may be contacted at [k.passino@osu.edu](mailto:k.passino@osu.edu), <http://www.ece.osu.edu/~passino/>

## 1. Introduction

Coordinated group motion has been studied extensively for a wide range of species (e.g., bacteria, insects, fish, and birds) via experiments, simulations, and mathematical analyses (Parrish and Hamner, 1997; Gueron et al., 1996; Mogilner and Edelstein-Keshet, 1999; Okubo et al., 2001; Couzin et al., 2002; Mogilner et al., 2003; Ballerini et al., 2008; Schultz et al., 2008). In a particularly early and influential study, Okubo and Chiang (1974) conducted a series of experiments where they filmed midge swarms flying over a white-colored “swarm marker,” an object over which midges are attracted to swarm for the purpose of mating (Chiang, 1968; Chiang et al., 1978, 1980). By using shadows from the sun on the white marker, they were able to distinguish between individual midge trajectories for all members of a swarm. This gave them position trajectories in the  $(x, y)$  plane, from which they computed position variances and velocities for the midges at each sampling time. They computed the swarm position centroid as it varied over time. They showed that the speed distribution of the midges in a swarm obeyed a 2-dimensional Maxwell-Boltzmann distribution, and that the distributions of the velocities relative to the velocity centroid, in the  $x$  and  $y$  dimensions, were Gaussian. They showed that the velocity autocorrelation coefficient varied with lag, and that the density functions for midge position relative to the swarm position centroid are not Gaussian, but are peaked near zero. Moreover, they showed the mean midge velocity depends on the distance from the middle of the swarm, with a tendency for higher velocities near the edge of the swarm. Using the data sets from (Okubo and Chiang, 1974), Okubo et al. (1976) computed the acceleration field, which shows that accelerations are generally higher near the edges of the swarm. Mathematical analysis of midge swarms was initiated in (Okubo and Chiang, 1974), and later expanded upon in (Okubo, 1986). In both these works, the focus was on comparisons between diffusion processes and swarm dynamics in terms of autocorrelation functions.

In first part of this paper, a nonlinear stochastic discrete-time (“individual-based”) model is introduced and it is shown how to tune the parameters of the midge’s individual rule for flight and sensing accuracy so that simulated midge flight characteristics and distributions (e.g., velocity and acceleration) closely match those in (Okubo and Chiang, 1974; Okubo et al., 1976) (since the experimental data sets are not available, only a visual comparison is possible). Apparently, this is the first time that an individual-based model has been tuned to represent the midge swarm data. In the second part of this paper, swarm cohesiveness is studied via mathematical analysis of its boundedness properties. Boundedness of discrete-time swarms was first studied in (Liu and Passino, 2004b), with corresponding continuous-time results given in (Liu and Passino, 2004a). In both these studies, however, the work was entirely mathematical, with no connection to a biological swarm. Here, the work in (Liu and Passino, 2004a,b) is extended so it applies to midge swarms. This involves including a representation of the swarm marker and specific rules to represent how individual midges decide to fly relative to the swarm marker and other members of the swarm. Also, it requires the development of ultimate boundedness results for both the distances between individual positions/velocities and swarm position/velocity centroids, and the swarm positions/velocity centroids themselves. Our analysis is unique with respect to past analysis on midge swarms (Okubo and Chiang, 1974; Okubo, 1986) and more generally the literature on analysis of biological swarms (e.g., the work in (Gueron et al., 1996; Mogilner and Edelstein-Keshet, 1999; Okubo et al., 2001; Mogilner et al., 2003)) both in the form of model used, and in the properties studied.

There is an important philosophical issue that needs to be raised about the contribution of this paper. First, the primary focus here is *not* to produce a perfectly accurate model of midge swarm behavior, analogous to how other studies of groups of animals have approached modeling (e.g., honey bees). Clearly, a good model would have to rely on better data (rather than the relatively old data used here, and for a relatively small number of

midges) and possibly additional aspects of the mathematical representation (e.g., other terms). Second, the primary focus here is not on the mathematical analysis of a model, for example, to obtain the most general conditions (and least conservative results) possible for cohesive behavior of groups. What is the focus? It is a “wholistic” focus. The focus here is to come up with a reasonably good model that is still analytically tractable, and show a useful property in an analytical way, which has significant value beyond just computational studies (e.g., in providing insights into mechanisms that lead to cohesive behavior). The paper succeeds at this approach, and hence is a significant contribution over what has been found in the literature along these lines. This philosophy is not new. It is essentially the philosophy that is used in control system design in engineering. You come up with a plant model that is “good enough,” develop a controller, and then prove mathematically that the close-loop system will, for instance, be stable. You do not want a perfect plant model as you will not be able to use any standard controller synthesis technique, and you will not be able to mathematically prove stability. The key is to take a wholistic view, taking into consideration model complexity in terms of the whole control system design process and goals and what you can achieve for them.

## 2. Midge Swarm Model

A nonlinear stochastic mathematical model is introduced and shown to be able to represent the data taken from real midge swarms in (Okubo and Chiang, 1974; Okubo et al., 1976). The model used for individual midge dynamics and sensing is kept particularly simple (e.g., detailed insect flight dynamics are ignored) to ensure that when multiple midges are composed into a set of interacting midges in a swarm, analytical tractability is maintained for the study in the next section.

## 2.1. Mathematical Model

Consider a swarm of  $N$  midges where the  $i^{\text{th}}$  midge has point mass dynamics

$$x^i(k+1) = x^i(k) + v^i(k)T \quad (1)$$

$$v^i(k+1) = v^i(k) + \frac{1}{M_i}u^i(k)T \quad (2)$$

Here,  $x^i \in \mathfrak{R}^n$  is the position (in meters),  $v^i \in \mathfrak{R}^n$  is the velocity (m/sec),  $M_i$  is the mass, and  $u^i \in \mathfrak{R}^n$  is the force input (Newtons) for the  $i^{\text{th}}$  midge. The sampling period is  $T$  seconds and to simplify notation “ $(kT)$ ” was replaced with “ $(k)$ ” where  $k$  is the time index. Basically, Equation (2) is an Euler approximation of “force equals mass times acceleration.” For simulations  $n = 3$ , but in analysis  $n$  is arbitrary.

Let  $\bar{x}(k) = \frac{1}{N} \sum_{i=1}^N x^i(k)$  and  $\bar{v}(k) = \frac{1}{N} \sum_{i=1}^N v^i(k)$  be the position centroid and velocity centroid of the swarm at the  $k^{\text{th}}$  time step, respectively. Let  $e_p^i(k) = x^i(k) - \bar{x}(k)$  and  $e_v^i(k) = v^i(k) - \bar{v}(k)$ . Let  $x_m \in \mathfrak{R}^n$  be a fixed position relative to a “swarm marker,” and where the midges seek to swarm (for convenience, we will refer to  $x_m$  as the swarm marker, when in actuality it is a fixed distance from the marker). Let  $e_m^i(k) = x^i(k) - x_m$ .

Assume each midge can sense its position relative to both  $\bar{x}$  and  $x_m$ , but with some errors. Let  $d_p^i \in \mathfrak{R}^n$  and  $d_m^i \in \mathfrak{R}^n$  be these sensing errors (e.g., noise) for midge  $i$ , respectively. Hence, each midge senses

$$\begin{aligned} \hat{e}_p^i(k) &= e_p^i(k) - d_p^i(k) \\ \hat{e}_m^i(k) &= e_m^i(k) - d_m^i(k) \end{aligned}$$

and uses these values in the flight rule below to decide how to coordinate inter-midge motion and their position relative to the swarm marker. Assume that the sensing errors have magnitudes bounded by

$$\|d_p^i(k)\| \leq D_p$$

$$\|d_m^i(k)\| \leq D_m \quad (3)$$

for any  $i$ , where  $D_p$  and  $D_m$  are known non-negative constants and  $\|\cdot\|$  is the 2-norm. Since no constraints are placed on the size of  $D_p$  and  $D_m$ , *arbitrarily poor sensing* of the errors  $e_p^i(k)$  and  $e_m^i(k)$  can be represented. For instance, a particular midge could be near the position centroid (center) of the swarm, and far from the swarm marker, but sense that it is near the edge of the swarm and close to the swarm marker. Essentially, the assumption here is that some subset of the swarm members appear as spots on the retina of the midge and an averaging process is completed to determine some type of estimate as to where the body of the swarm is. There is no additional limit placed on the rate of variation of the sensing errors. In particular,  $d_p^i(k)$  and  $d_m^i(k)$  are allowed to be bounded noise so that the misperceptions of relative positions can change significantly from one sampling instant to the next for each midge in the swarm.

Suppose the force input (“flight rule”) for the  $i^{\text{th}}$  midge at the  $k^{\text{th}}$  step is

$$\begin{aligned} u^i(k) = & -M_i k_p \hat{e}_p^i(k) - M_i k_m \hat{e}_m^i(k) - M_i k_d v^i(k) \\ & + M_i k_r \sum_{j=1, j \neq i}^N \exp\left(\frac{-\frac{1}{2}\|x^i - x^j\|^2}{r_s^2}\right) (x^i - x^j) \end{aligned} \quad (4)$$

This equation is from (Liu and Passino, 2004a). Here,  $-M_i k_p \hat{e}_p^i(k)$  is an “attraction term” that represents that individual  $i$  tries to move toward its noisy estimate of the swarm position centroid. The term  $-M_i k_m \hat{e}_m^i(k)$  represents a midge’s attraction to the swarm marker. Since the noisy signals  $\hat{e}_p^i(k)$  and  $\hat{e}_m^i(k)$  are used in these terms, aggregation around the swarm centroid and swarm marker are only loosely sought (e.g., one particular midge may at some time be attracted to aggregate at a different swarm marker location than another midge due to the sensing noise). The scalars  $k_p > 0$  and  $k_m > 0$  are “attraction gains” which indicate how aggressive each midge is in aggregating (coordinating its motion to maintain a tightly cohesive group) and staying close to the swarm marker. The gain

$k_d > 0$  is a “velocity damping gain.” The gain  $k_r > 0$  is a “repulsion gain” which sets how much the midges want to be kept at a distance from other midges, and  $r_s$  represents its “repulsion range.” A relatively small value, near the dimensions of a midge, will be used for  $r_s$ . The fast roll-off of the Gaussian nonlinearity in the repulsion term is used to represent that the  $i^{\text{th}}$  midge has little influence on the  $j^{\text{th}}$  midge if they are far apart, and hence, two such midges do not need to be able to sense each other’s positions. On the other hand, if midges are close together, the approach to repulsion indicates that midges can accurately sense other midge positions and thereby repel each other, or that if they collide they will push each other away. Simulations show that this repulsion term is effective in avoiding any two midges being at nearly the same position at the same time.

## 2.2. Representing Midge Swarm Data

The best midge swarm representation that could be derived in this paper was obtained by tuning the model just presented in Section 2.1. This model is presented first. Afterwards, the alternatives that were considered are discussed in detail.

Although typically a swarm has  $N \leq 20$ , choose  $N = 25$ , the number of midges in the “series 5 data” (Okubo and Chiang, 1974). Assume each midge weighs 0.12 mg which sets the value of  $M_i$ . Let  $T = 0.01$  sec. Let  $x_m = [0, 0, 0]^T$ . Let  $d_p^i(k)$  and  $d_m^i(k)$ , for each  $k$  and  $i$ , have all their components uniformly distributed on  $k_{dp}[-1, 1]$  and  $k_{dm}[-1, 1]$ , respectively, with  $k_{dp} = k_{dm} = 0.1$  so that midges’ sensing of  $e_p^i(k)$  and  $e_m^i(k)$  can be off by  $\pm 10$  cm in each dimension at each  $k$  (changing these values has impact on simulated swarm behavior, e.g., if the parameters are made larger, the swarm is “noisier” and not as compact). Since a swarm’s diameter is 10–12 cm (Okubo and Chiang, 1974), this is a relatively high magnitude of sensing noise (e.g., a midge at the edge of a swarm can perceive itself to be on the other side of the swarm and a midge that perceives itself to be exactly at  $x_m$ , could be 10 cm



from it). The choice of the noise sequences gives  $D_p = D_m = \sqrt{0.03} = 0.1732$ . Tuning the model led to  $k_p = 100$ ,  $k_m = 10$ ,  $k_d = 5$ , and  $k_r = 1000$ . A value  $r_s = 0.001$  was used since a midge is about 2 mm long (Okubo, 1986). Initial conditions for the midge positions were set at random values such that they were  $\pm 5$  cm from  $x_m$  in each dimension. Initial conditions for the midge velocities were set at random directions with speeds that were uniformly distributed on  $[0, 80]$  cm/sec. Other distributions for initial conditions produced similar results. Simulations represent 20 sec. of real time; this is a significantly longer data collection time period than in (Okubo and Chiang, 1974).

Plots of 9 of the  $N = 25$  midges for 1 sec. at the end of the 20 sec. simulation are shown in Figure 1 with trajectories only shown in the  $(x, y)$  plane. Interestingly, these trajectories have characteristic shapes similar to the individual midge trajectories shown in Figs. 13–22 in (Okubo and Chiang, 1974) (notice the occasional abrupt turns and erratic behavior that interrupts smooth arcs). Figure 2 shows that swings of 4–5 cm can occur in the components of  $\bar{x}(k)$  and that the standard deviations on components of  $x^i(k)$  are in the range of 2 cm as in Figs. 7–10 in (Okubo and Chiang, 1974).

The distributions for the normalized components of  $e_p$  and  $e_v$  are shown in Figure 3. The top three plots show three normalized distributions for components of  $e_p$ , which roughly match the distribution shown in Fig. 42 (series 5 data) in (Okubo and Chiang, 1974). The bottom three plots of Figure 3 show a good fit with the Gaussian  $0.3 \exp(-\psi^2/800)$  (where  $\psi$  is a normalized value in the  $x$ ,  $y$ , or  $z$  dimension) that is taken from Fig. 35 (series 5 data) in (Okubo and Chiang, 1974).

Figure 4 shows the speed distribution for the 2-dimensional case along with a Maxwell-Boltzmann distribution fit taken from Fig. 32 (series 5 data) in (Okubo and Chiang, 1974). This shows that the speed distribution matches the one found from the experimental data very well. The mean 2-d velocity was found to be 31.3 cm/sec. for series

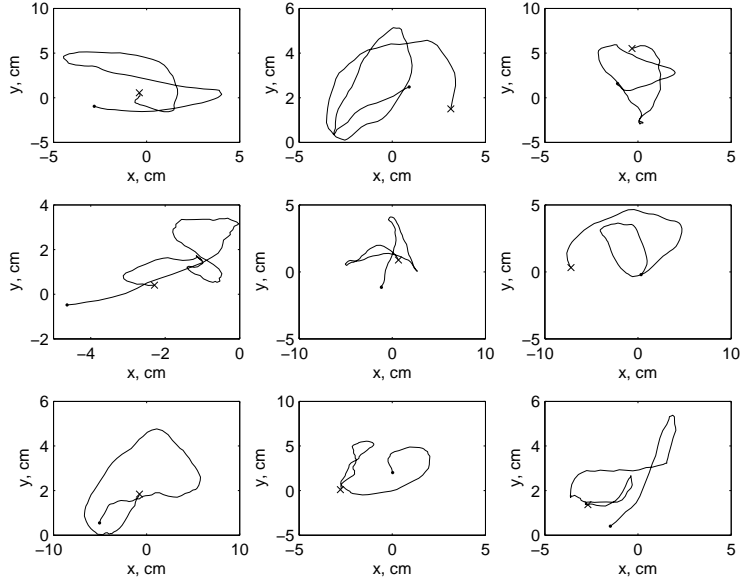


Fig. 1.— Simulation of midge position trajectories (shown in  $(x, y)$ -plane). The “ $\times$ ” marks starting points of the trajectories, and the dots represent the midges at the end of the simulation.

4 data and 25 cm/sec. for series 5 data (Okubo and Chiang, 1974), whereas here the mean 2-d velocity for the model is 26.19 cm/sec., a good match.

The normalized velocity distributions for the  $x$  direction are shown in Figure 5. Similar plots are found for the  $y$  and  $z$  cases. The shape of the plot is similar to the corresponding one for the  $x$  direction as seen in Fig. 56 in (Okubo and Chiang, 1974). The normalized acceleration distributions for the  $x$  direction are shown in Figure 6. Similar plots are found for the  $y$  and  $z$  cases. The shape of the plot is quite similar to the corresponding one for the  $x$  direction as seen in Figs. 5–6 in (Okubo et al., 1976) for the series 5 data (e.g., the peaks in the median accelerations are at about  $\pm 500$  cm/sec<sup>2</sup> in both cases).

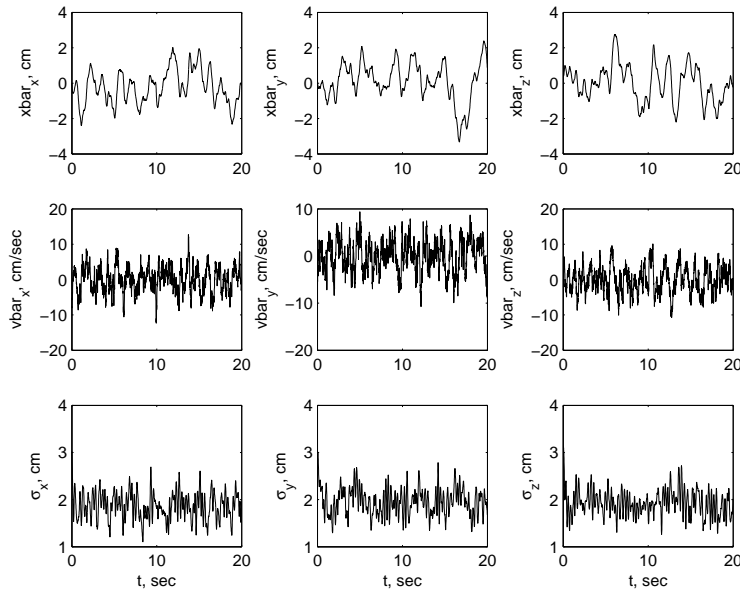


Fig. 2.— Midge swarm simulation, plots of  $\bar{x}(k)$ ,  $\bar{v}(k)$ , and standard deviations of position trajectories in each dimension.

### 2.3. Alternative Flight Rules for Midge Swarm Representation

The simulations of the last subsection show that flight rules used by the midges in a swarm *could* be specified via Equation (4) with the gains given above (e.g.,  $k_p$ ,  $k_m$ ,  $k_d$ ,  $k_r$ , and  $r_s$ ). In this subsection, an evaluation of alternative flight rules for Equation (4) is given.

#### 2.3.1. Adding and Deleting Terms from the Midge Flight Rule

First, evaluations were made of whether additional terms in Equation (4) result in an even more accurate representation of the midge swarm experimental data. A common term used in swarm simulations and analysis (e.g., see (Couzin et al., 2002)) is an “alignment” term which has also been called a “velocity attraction” term. The purpose of such a term is for each midge to align its direction of flight to other members of the swarm (or, for

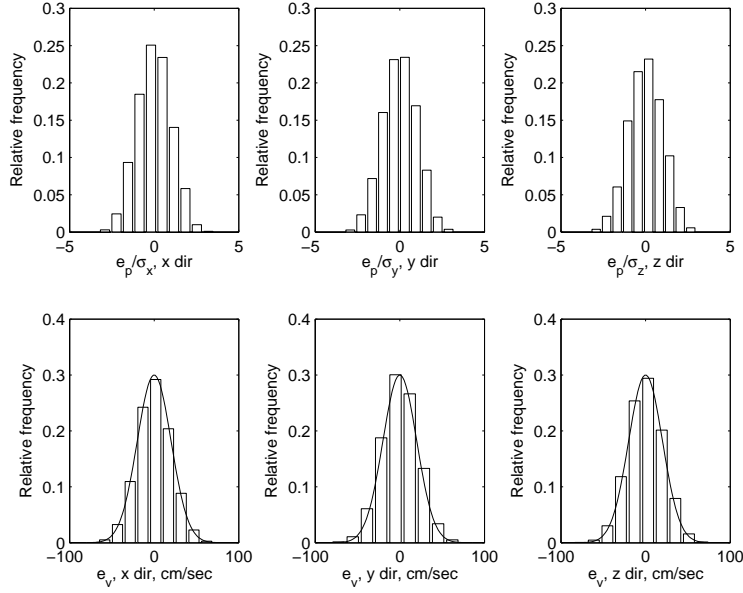


Fig. 3.— Distributions of normalized components of  $e_p$  and  $e_v$ . The normalization factors are  $\sigma_x$ ,  $\sigma_y$ , and  $\sigma_z$ , the standard deviations of the appropriate component of  $x^i$  over  $i$  at each  $k$ .

example, the noisily sensed swarm velocity centroid). When such a term (i.e., a term  $-M_i k_v \hat{e}_v^i(k)$ ) is added to Equation (4) and the gains tuned, it was not possible to match the experimental data in (Okubo and Chiang, 1974; Okubo et al., 1976) better than is done in the previous subsection. In fact, when reasonably close matches were found to the position, velocity, and acceleration distributions, it was found that the swarm wandered as a cohesive group with variations of almost  $\pm 10$  cm in the components of  $\bar{x}(k)$ . This is not surprising as the velocity attraction term causes the midge flight directions to align with each other thereby causing the swarm to stay cohesive but move purposefully in a single direction. Other alignment rules result in similar behavior.

The principle of parsimony would support the validity of flight rules other than the one given by Equation (4), if for example, one of the terms from Equation (4) could be

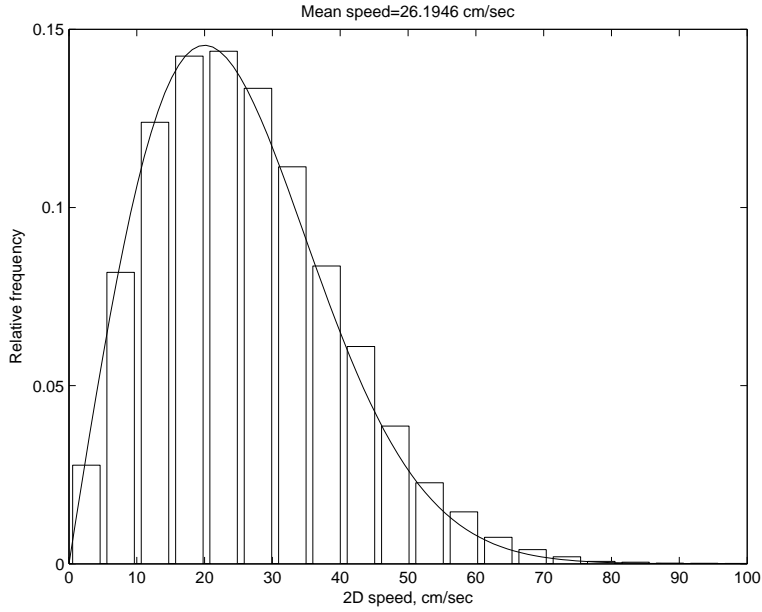


Fig. 4.— Speed distribution for the 2-dimensional case (bar plot) along with a 2-dimensional Maxwell-Boltzmann speed distribution  $0.012v \exp(-v^2/800)$  (line plot) where  $v$  is the speed  $v = \sqrt{v_x^2 + v_y^2}$ .

eliminated but the swarm data is still adequately represented. Consider each term of Equation (4). The repulsion term must be present in order to ensure that two midges are not in the same position at the same time. The velocity damping term  $-M_i k_d v^i(k)$  in Equation (4) is used to represent that midges cannot fly arbitrarily fast; it represents physical constraints and hence must be present. The swarm marker attraction term  $-M_i k_m \hat{e}_m^i(k)$  must be present to represent that in nature swarms do indeed hover over swarm markers (letting  $k_m = 0$  results in a swarm position centroid that will eventually wander long distances). It is not known, however, if the aggregation term  $-M_i k_p \hat{e}_p^i(k)$  (or one like it) is used by the midges. It was included in the last section since experiments show that relatively small swarms are maintained over larger markers (e.g., via Fig. 1 in (Okubo and Chiang, 1974) for  $N = 20$ , a swarm marker area is about  $750 \text{ cm}^2$ , and the

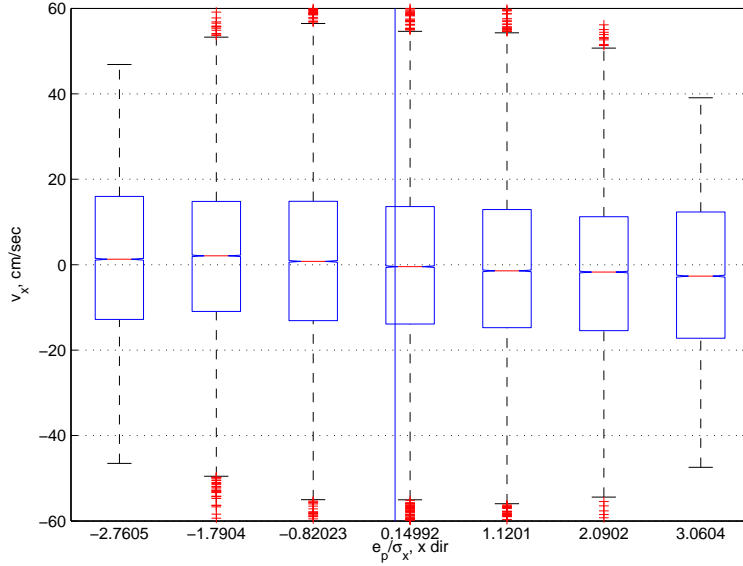


Fig. 5.— Velocity distributions for the  $x$  direction as a function of normalized  $x$  component of  $e_p$ . The vertical line represents zero on the horizontal axis. For the boxplots, the middle line in each box is the median value, boxes with notches that do not overlap represent that the medians of the two groups differ at the 5% significance level, the edges of the boxes are the 25<sup>th</sup> and 75<sup>th</sup> percentiles, whiskers (dashed lines) represent 1.5 times the interquartile range, and outliers are designated with a “+”.

swarm diameter is only 10-12 cm). Interestingly, if the  $-M_i k_p \hat{e}_p^i(k)$  term is removed from Equation (4), and the value of  $k_m$  is changed from  $k_m = 10$  to  $k_m = 100$ , then simulations show that all the distributions (e.g., velocity and acceleration) match those from (Okubo and Chiang, 1974) reasonably well (the reason for this is that on average  $\bar{x}(k)$  is reasonably close to  $x_m$ ). However, the plot for  $k_m = 100$  corresponding to Figure 2 (not included here in the interest of space) shows that  $\bar{x}(k)$  only has swings of 2 cm, which is inconsistent with (Okubo and Chiang, 1974). Basically,  $k_m$  had to be increased to  $k_m = 100$  to get the mean velocity of the midges in the swarm to match the one found in experiments;

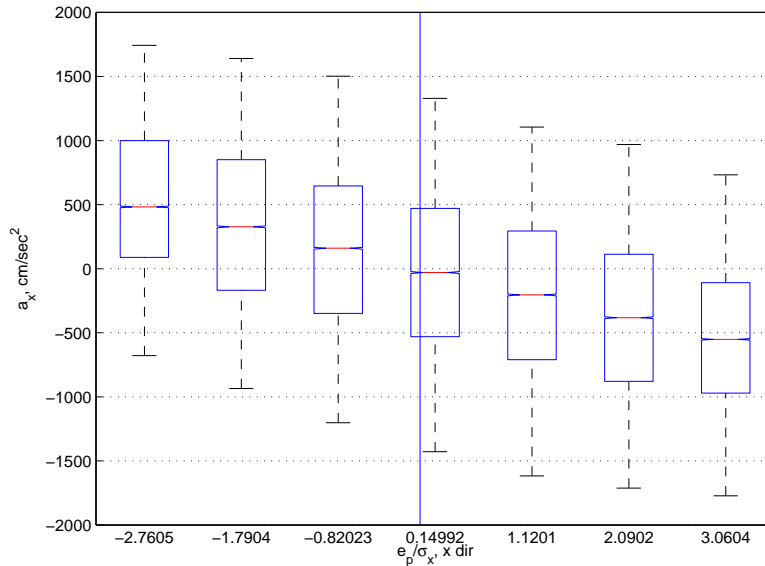


Fig. 6.— Acceleration distributions for the  $x$  direction as a function of normalized  $x$  component of  $e_p$ . The vertical line represents zero on the horizontal axis. The boxplot is explained in the caption of Figure 5.

however, such a high gain results in very close regulation of the swarm position centroid to the swarm marker. Above, by having both the  $k_p$  and  $k_m$  gains it was possible to match the acceleration and velocity distributions, and only loosely regulate the swarm around the swarm marker. Hence, it seems that terms of the type shown in Equation (4) are *all* needed for a valid representation of the experimental midge swarm data.

### 2.3.2. Rules Based on Metric Distance

While it has been shown that terms of the *type* seen in Equation (4) are all needed to represent the experimental data, we have not established that terms of the *specific* form in Equation (4) are needed. In this and the following subsection we examine whether there

are other attraction and repulsion rules that use less information and result in matching the swarm data as well as was done in the last subsection.

Common rules for attraction and repulsion in swarms (e.g., see (Couzin et al., 2002; Parrish and Hamner, 1997))) say that a midge should move away from other midges that are within a “repulsion range”  $r_r$ , but toward other midges that are outside the radius  $r_r$ , but inside the “attraction range”  $r_a > r_r$ . Let

$$N_r(i, k) = \{j : \|x^i(k) - x^j(k)\| \leq r_r, j = 1, \dots, N\}$$

denote the set of indices of the neighbors of midge  $i$  that are within the repulsion range of  $r_r$  at time  $k$ . Here, as with the rule in Equation (4), the value of  $r_r = 1$  mm since a midge is 2 mm in length. Let

$$N_a(i, k) = \{j : r_r < \|x^i(k) - x^j(k)\| \leq r_a, j = 1, \dots, N\}$$

denote the set of indices of the neighbors of midge  $i$  that are in the “attraction range” of midge  $i$  at time  $k$ . Let

$$\bar{x}_r^i(k) = \frac{1}{|N_r(i, k)|} \sum_{j \in N_r(i, k)} x^j(k)$$

denote the position centroid of the neighbors of midge  $i$  in the repulsion range  $r_r$  at time  $k$ . Note that for all  $k$ ,  $|N_r(i, k)| \neq 0$  since  $i \in N_r(i, k)$ . Let

$$\bar{x}_a^i(k) = \frac{1}{|N_a(i, k)|} \sum_{j \in N_a(i, k)} x^j(k)$$

denote the position centroid of the neighbors of midge  $i$  in the attraction range at time  $k$ . Let  $e_{pa}^i(k) = x^i(k) - \bar{x}_a^i(k)$  and  $\hat{e}_{pa}^i(k) = e_{pa}^i(k) - d_p^i(k)$  and use this in place of  $\hat{e}_p^i(k)$  in Equation (4). Let  $e_{pr}^i(k) = x^i(k) - \bar{x}_r^i(k)$  and use  $+M_i k_r e_{pr}^i(k)$  in place of the repulsion term in Equation (4). Note that if  $r_a$  is too small, then midges will frequently become “disconnected” from the swarm. Here, however, drastic effects like a midge wandering off



and never returning (a permanent disconnection) are not seen due to the presence of the attraction term to the swarm marker in Equation (4).

After running many simulations, it was determined that (i) the noise needed to be at a high magnitude in order to match the swarm data and in particular the speed distribution (hence,  $k_{dp} = k_{dm} = 0.1$  was used in all cases); and (ii) it was logical to begin with a low value of  $r_a$  and increase its value while keeping the other midge flight rule parameters the same as in Section 2.2 (i.e., in no cases was it found that further tuning of the parameters improved the matching to the swarm data). Using  $r_a = 1$  cm it was not possible to tune the gains so that the swarm data is even close to being matched. Other small values of  $r_a$  produced the same poor results. Intermediate values of  $r_a$  (e.g.,  $r_a = 5$  cm) result in matching a few features of the swarm data, but with midges occasionally swinging far outside the swarm body. Only when  $r_a$  is increased to  $r_a = 9$  cm did the matching of the swarm data improve. There were, however, still a few relatively wide swings of a few midges away from the body of the swarm. When  $r_a = 10$  cm this effect was eliminated and the model matched the swarm data as well as in Section 2.2. Of course, values of  $r_a > 10$  cm also resulted in good matching of the data. Considering the diameter of the swarm (10–12 cm) these high values of  $r_a$  correspond to the case where each midge has almost all the other midges within  $r_a$  so that this attraction rule becomes indistinguishable from the one used in Section 2.2.

### 2.3.3. Rules Based on Topological Distance

Next, in (Ballerini et al., 2008) it was shown that birds interacted with each other using a “topological distance” (a fixed number of nearest neighbors) rather than the metric distance like we just considered. Let  $N_n(i, k)$  denote the set of indices of the  $N_c$ ,  $1 \leq N_c \leq N$ , nearest neighbors to midge  $i$  at time  $k$  (for consistency with the last rule, this

includes midge  $i$  as its own nearest neighbor). Let

$$\bar{x}_{nn}^i(k) = \frac{1}{N_c} \sum_{j \in N_n(i,k)} x^j(k)$$

denote the position centroid of the  $N_c$  nearest neighbors of midge  $i$  at time  $k$ . Let  $\hat{e}_{pnn}^i(k) = x^i(k) - \bar{x}_{nn}^i(k) - d_p^i(k)$  and use this in place of  $\hat{e}_p^i(k)$  in Equation (4). Since the repulsion rule has little effect, we simply keep the one in Equation (4). Note that with  $N_c = N = 25$  this attraction rule is the same as the one used in Section 2.2 to match the swarm data. For the birds in (Ballerini et al., 2008) a value of  $N_c$  of 6 or 7 was found when  $N$  was in the range of hundreds to tens of thousands. This implies that  $N_c = 2$  is a natural starting point for trying to match the swarm data (although in this case groups of size 2 could become disconnected from the swarm, the attraction to the swarm marker alleviates this problem as it did above). Simulations show that if a value of  $N_c = 2$  is used it is not possible to tune the gains of the rule to even get close to matching the experimental swarm data. Similar poor results are obtained for other small values of  $N_c$ . For  $10 \leq N_c \leq 20$ , however, it was possible to match much of the swarm data except for periodic swings of the trajectories far outside the main body of the swarm. For  $20 < N_c \leq 25$  the frequency and magnitude of the swings outside the main body of the swarm decrease so the model matched the swarm data as well as in Section 2.2.

In summary, while both the metric and topological approaches to defining the attraction rule can be tuned to represent the swarm data, good representations were only obtained when the rules closely approximated the attraction rule in Section 2.2 (i.e., with high values of  $r$  and  $N_c$ ). Hence, it seems that if the midges use less information to define an attraction rule, the swarm data cannot be matched. It is for this reason that cohesiveness analysis is only considered for the case when the flight rule in Equation (4) is used with the gains defined in Section 2.2.

### 3. Swarm Cohesiveness Analysis

In this section, the dynamics of the error system for the  $i^{\text{th}}$  midge are defined as the differences between the midge position and velocity and the swarm position and velocity centroids, respectively. After some preliminary technical results, the first main result is given. It states that the error system is ultimately bounded, which means that all midge position and velocity trajectories are eventually bounded relative to the swarm position and velocity centroid trajectories. Next, the dynamics of the velocity centroid and swarm position centroid relative to the swarm marker are defined. Then, in the second main result it is shown that the distance between the swarm position centroid and the swarm marker is ultimately bounded. Moreover, the velocity centroid is ultimately bounded.

#### 3.1. Error Dynamics

Consider an error system with  $e_p^i(k) = x^i(k) - \bar{x}(k)$  and  $e_v^i(k) = v^i(k) - \bar{v}(k)$ . Using simple algebra, the error dynamics are

$$\begin{aligned} e_p^i(k+1) &= e_p^i(k) + e_v^i(k)T \\ e_v^i(k+1) &= e_v^i(k) + \frac{1}{M_i}u^i(k)T - \frac{1}{N}\sum_{j=1}^N \frac{1}{M_j}u^j(k)T \end{aligned} \quad (5)$$

Substitute the  $u^i$  in Equation (4) into the error dynamics in Equation (5). First, notice that

$$\begin{aligned} \frac{1}{M_i}u^i(k) &= -k_p e_p^i(k) + k_p d_p^i(k) - k_m e_m^i(k) + k_m d_m^i(k) - k_d v^i(k) \\ &+ k_r \sum_{j=1, j \neq i}^N \exp\left(\frac{-\frac{1}{2}\|x^i - x^j\|^2}{r_s^2}\right) (x^i - x^j) \end{aligned} \quad (6)$$

Also,

$$\frac{1}{N}\sum_{j=1}^N \frac{1}{M_j}u^j(k) = \frac{1}{N}\sum_{j=1}^N k_p d_p^j(k) - k_m \bar{x}(k) + k_m x_m$$

$$+\frac{1}{N} \sum_{j=1}^N k_m d_m^j(k) - k_d \bar{v}(k) \quad (7)$$

where we used the facts that  $\frac{1}{N} \sum_{j=1}^N e_p^j = 0$  and

$$\frac{1}{N} \sum_{j=1}^N k_r \sum_{\ell=1, \ell \neq j}^N \exp\left(\frac{-\frac{1}{2}\|x^j - x^\ell\|^2}{r_s^2}\right) (x^j - x^\ell) = 0$$

From Equations (5), (6), and (7) we have

$$e_v^i(k+1) = -(k_p + k_m)T e_p^i(k) + (1 - k_d T) e_v^i(k) + U^i(k)$$

where

$$\begin{aligned} U^i(k) &= k_p T d_p^i(k) + k_m T d_m^i(k) \\ &\quad - \frac{1}{N} \sum_{j=1}^N k_p T d_p^j(k) - \frac{1}{N} \sum_{j=1}^N k_m T d_m^j(k) \\ &\quad + k_r T \sum_{j=1, j \neq i}^N \exp\left(\frac{-\frac{1}{2}\|x^i - x^j\|^2}{r_s^2}\right) (x^i - x^j) \end{aligned}$$

which is a nonlinear non-autonomous system.

Define  $E^i = [e_p^{i\top}, e_v^{i\top}]^\top$  and  $E = [E^1\top, E^2\top, \dots, E^N\top]^\top$ . With  $I$  an  $n \times n$  identity matrix, the error dynamics of the  $i^{\text{th}}$  midge may be written as

$$E^i(k+1) = \overbrace{\begin{bmatrix} I & TI \\ -(k_p + k_m)TI & (1 - k_d T)I \end{bmatrix}}^A E^i(k) + \overbrace{\begin{bmatrix} 0 \\ I \end{bmatrix}}^B U^i(k) \quad (8)$$

If we view the swarm as an interconnected system with each midge being a subsystem, then the matrix  $A$  in Equation (8) specifies the internal system dynamics for each midge subsystem in the error system, and  $U^i(k)$  gives the external input for each midge  $i$  at time step  $k$ .

### 3.2. Preliminaries

The following result will be used in both the theorems below.

**Lemma 1** *The matrix  $A$  in Equation (8) is convergent if*

$$T < \begin{cases} \frac{4}{k_d + \sqrt{k_d^2 - 4(k_p + k_m)}} & \text{if } k_d^2 - 4(k_p + k_m) \geq 0 \\ \frac{k_d}{k_p + k_m} & \text{if } k_d^2 - 4(k_p + k_m) < 0 \end{cases} \quad (9)$$

**Proof:** It can be proven that matrix  $A$  has  $n$  repeated values of the eigenvalues of matrix  $\tilde{A} = \begin{bmatrix} 1 & T \\ -(k_p + k_m)T & 1 - k_d T \end{bmatrix}$ . From  $|zI - \tilde{A}|$  the characteristic equation  $z^2 + [k_d T - 2]z + 1 - k_d T + (k_p + k_m)T^2 = 0$  which has roots

$$z_{1,2} = \frac{1}{2} \left[ 2 - k_d T \pm T \sqrt{k_d^2 - 4(k_p + k_m)} \right]$$

To have a convergent  $A$ :

1. If  $k_d^2 - 4(k_p + k_m) \geq 0$ , then we need

$$-1 < \frac{1}{2} \left[ 2 - k_d T \pm T \sqrt{k_d^2 - 4(k_p + k_m)} \right] < 1$$

Notice that  $z_{1,2} < 1$  always holds with  $(k_p + k_m) > 0$ . To have  $-1 < z_{1,2}$ , we need

$$-2 < 2 - k_d T - T \sqrt{k_d^2 - 4(k_p + k_m)}$$

that is

$$T < \frac{4}{k_d + \sqrt{k_d^2 - 4(k_p + k_m)}}$$

2. If  $k_d^2 - 4(k_p + k_m) < 0$ , we need  $\|z_{1,2}\| < 1$ . So

$$(2 - k_d T)^2 + T^2 [4(k_p + k_m) - k_d^2] < 4$$

$$4 - 4k_d T + 4(k_p + k_m)T^2 < 4$$

That is  $0 < T < \frac{k_d}{(k_p + k_m)}$ .

■

The proofs in the remainder of the paper rely on a few facts. First, when matrix  $A$  is convergent, for any given matrix  $\tilde{Q} = \tilde{Q}^\top > 0$  (positive definite), there exists a unique matrix  $\tilde{P} = \tilde{P}^\top > 0$  which is the solution of the discrete Lyapunov equation  $A^\top \tilde{P} A - \tilde{P} = -\tilde{Q}$ . Given  $\tilde{P}$  and  $\tilde{Q}$  that satisfy the discrete Lyapunov equation above, define  $\beta_{M,m}$  respectively as twice the values of the maximum and minimum eigenvalues of  $\tilde{P}$  given  $\tilde{Q} = I$ , i.e.,  $\beta_M = 2\lambda_{\max}(\tilde{P}|_{\tilde{Q}=I})$  and  $\beta_m = 2\lambda_{\min}(\tilde{P}|_{\tilde{Q}=I})$ . Next, with  $\tilde{P}$ ,  $\tilde{Q}$  and  $\beta_M$  defined above, the minimum of function  $f(\tilde{P}, \tilde{Q}) = \frac{2\lambda_{\max}(\tilde{P})}{\lambda_{\min}(\tilde{Q})}$  is  $\beta_M$ , that is,

$$\beta_M = \left. \frac{2\lambda_{\max}(\tilde{P})}{\lambda_{\min}(\tilde{Q})} \right|_{\tilde{Q}=I} = \min_{\tilde{Q}} f(\tilde{P}, \tilde{Q}) \quad (10)$$

### 3.3. Ultimate Boundedness of Midge Swarm Error Dynamics

Our first main result, proven next, states that there exists a known ultimate bound (in terms of the midge parameters) on the distances between the midge positions and velocities and the swarm centroid position and centroid velocity, and the error system trajectories will eventually satisfy this bound.

**Theorem 1** *Consider the error system described by the model in Equation (5). Assume the noise satisfies Equation (3). Assume  $T$  is sufficiently small such that the condition in Lemma 1 holds. Then, the trajectories of the swarm error system are uniformly ultimately bounded, and  $E^i$  for all  $i$  will converge to the set  $\Omega_b$ , where*

$$\Omega_b = \left\{ E : \|E^i\| \leq \frac{\gamma}{2} \left( \beta_M \|A\| + \sqrt{\|A\|^2 \beta_M^2 + 2\beta_M} \right), i = 1, 2, \dots, N \right\} \quad (11)$$

*is attractive and compact, with  $\beta_M$  defined in Equation (10) and*

$$\gamma = \left( 2k_p D_p + 2k_m D_m + (N-1)k_r \exp\left(-\frac{1}{2}\right) r_s \right) T \quad (12)$$

**Proof:** To study the boundedness of the error dynamics, it is convenient to choose a Lyapunov function for each midge as

$$V_i(k) = E^i(k)^\top P E^i(k) \quad (13)$$

with  $P = P^\top > 0$  a  $2n \times 2n$  positive definite matrix. Then we have

$$\begin{aligned} V_i(k+1) &= E^i(k+1)^\top P E^i(k+1) \\ &= E^i(k)^\top A^\top P A E^i(k) + 2U^i(k)^\top B^\top P A E^i(k) \\ &\quad + U^i(k)^\top B^\top P B U^i(k) \end{aligned}$$

and

$$\begin{aligned} \Delta V_i(k) &= V_i(k+1) - V_i(k) \\ &= E^i(k)^\top \underbrace{(A^\top P A - P)}_{-Q} E^i(k) + 2U^i(k)^\top B^\top P A E^i(k) \\ &\quad + U^i(k)^\top B^\top P B U^i(k) \end{aligned} \quad (14)$$

Note that given any  $Q = Q^\top > 0$ , the existence of a desired  $P$  is guaranteed. Choose for the composite system

$$V(k) = \sum_{i=1}^N V_i(k)$$

where  $V_i(k)$  is given in Equation (13). In what follows, we seek to find conditions where  $\Delta V_i(k) < 0$  for all  $i$  so that  $\Delta V(k) < 0$ .

First, since for any matrix  $M = M^\top > 0$  and vector  $X$

$$\lambda_{\min}(M) X^\top X \leq X^\top M X \leq \lambda_{\max}(M) X^\top X$$

where  $\lambda_{\min}(M)$  and  $\lambda_{\max}(M)$  denote the minimum and maximum eigenvalue of  $M$ , respectively, then we have

$$\sum_{i=1}^N \left( \lambda_{\min}(P) \|E^i(k)\|^2 \right) \leq V(k) \leq \sum_{i=1}^N \left( \lambda_{\max}(P) \|E^i(k)\|^2 \right) \quad (15)$$

Next, we need to bound the  $U^i(k)$  contribution in Equation (8). Notice that any function  $F(\psi) = \exp\left(\frac{-\frac{1}{2}\|\psi\|^2}{r_s^2}\right) \|\psi\|$ , with  $\psi$  any real vector, has a unique maximum value of  $\exp\left(-\frac{1}{2}\right) r_s$  which is achieved when  $\|\psi\| = r_s$  (Gazi and Passino, 2003). So

$$\left\| \sum_{j=1, j \neq i}^N \exp\left(\frac{-\frac{1}{2}\|x^i - x^j\|^2}{r_s^2}\right) (x^i - x^j) \right\| \leq (N-1) \exp\left(-\frac{1}{2}\right) r_s$$

Thus, from Equation (3) we have

$$\begin{aligned} \|U^i(k)\| &\leq k_p T D_p + k_m T D_m + k_p T \frac{1}{N} \sum_{j=1}^N D_p + k_m T \frac{1}{N} \sum_{j=1}^N D_m \\ &\quad + (N-1) k_r T \exp\left(-\frac{1}{2}\right) r_s \end{aligned}$$

Using Equations (14) and (15), the fact that  $\|B\| = 1$ , and the fact that since  $P > 0$ ,  $\|P\| = \lambda_{max}(P)$ , we have

$$\begin{aligned} \Delta V(k) &= \sum_{i=1}^N \Delta V_i(k) \\ &\leq \sum_{i=1}^N \left[ -\lambda_{min}(Q) \|E^i(k)\|^2 + 2\lambda_{max}(P) \|U^i(k)\| \|A\| \|E^i(k)\| \right. \\ &\quad \left. + \lambda_{max}(P) \|U^i(k)\|^2 \right] \\ &\leq \sum_{i=1}^N \left[ -\|E^i(k)\|^2 + \beta_M \gamma \|A\| \|E^i(k)\| + \frac{\beta_M \gamma^2}{2} \right] \end{aligned}$$

where  $\beta_M = 2\lambda_{max}(P)$ . For the last step, by inspecting the above inequality it can be seen that minimizing  $\beta_M$  is desirable for achieving boundedness with a small bound so  $Q = I$  was chosen.

To find the set  $\Omega_b$  note that

$$\Delta V_i(k) \leq -\|E^i(k)\|^2 + \beta_M \gamma \|A\| \|E^i(k)\| + \frac{\beta_M \gamma^2}{2} \quad (16)$$

Solving the equation on the right-hand-side gives that  $\Delta V_i(k) < 0$  for all  $i$  (and hence,  $\Delta V(k) < 0$ ) when

$$\|E^i(k)\| > \frac{\gamma}{2} \left( \beta_M \|A\| + \sqrt{\|A\|^2 \beta_M^2 + 2\beta_M} \right)$$



So the set  $\Omega_b$  given in the statement of the theorem is attractive and compact. ■

**Remark 1** *The fact that  $\Omega_b$  is “attractive” means that if  $E^i(k)$  trajectories start outside  $\Omega_b$ , they will eventually enter that set, and stay in that set for all time. Hence, if a midge starts at a large distance from the main body of the swarm, and at zero velocity, it will eventually join the group and stay within a known bound from the swarm position and velocity centroid.*

**Remark 2** *If there is a higher magnitude of sensing noise (quantified by the noise bounds  $D_p$  and  $D_m$ ), or higher values of the repulsion gain  $k_r$ , repulsion range  $r_s$ , or  $N$ , then intuitively it is expected that the size of the swarm will grow. The above result is consistent with this intuition. In particular, if  $D_p$ ,  $D_m$ ,  $k_r$ ,  $r_s$ , or  $N$  increases, the size of  $\Omega_b$ , denoted by  $|\Omega_b|$ , increases as can be seen by inspecting Equations (11) and (12).*

### 3.4. Ultimate Boundedness of Midge Swarm Position and Velocity Centroids

To study the boundedness of the midge swarm position and velocity centroids their dynamics are first defined. Note that

$$\begin{aligned}\bar{x}(k+1) - x_m &= \bar{x}(k) - x_m + \bar{v}(k)T \\ \bar{v}(k+1) &= \bar{v}(k) + \frac{1}{N} \sum_{j=1}^N \frac{1}{M_j} u^j(k)T \\ &= -k_m T (\bar{x}(k) - x_m) + (1 - k_d T) \bar{v}(k) + (k_p \bar{d}_p + k_m \bar{d}_m) T\end{aligned}\tag{17}$$

with  $\bar{d}_p(k) = \frac{1}{N} \sum_{i=1}^N d_p^i(k)$  and  $\bar{d}_m(k) = \frac{1}{N} \sum_{i=1}^N d_m^i(k)$ . Let  $C(k) = [(\bar{x}(k) - x_m)^\top \bar{v}(k)^\top]^\top$ .

With  $I$  an  $n \times n$  identity matrix, the dynamics are

$$C(k+1) = \overbrace{\begin{bmatrix} I & TI \\ -k_m TI & (1 - k_d T) I \end{bmatrix}}^{\bar{A}} C(k) + \overbrace{\begin{bmatrix} 0 \\ I \end{bmatrix}}^{\bar{B}} \bar{U}(k)\tag{18}$$

where  $\bar{U}(k) = (k_p \bar{d}_p(k) + k_m \bar{d}_m(k))T$ .

Our second main result, proven next, states that there exists a known ultimate bound (in terms of the midge parameters) on the difference between the swarm position centroid and the swarm marker that will eventually be satisfied. Also, it shows that the velocity centroid is ultimately bounded. The pattern of the formulation and proof follow that of the last subsection closely; hence, the presentation is brief.

**Theorem 2** *Consider the dynamics described by the model in Equation (18). Assume the noise satisfies Equations (3). Assume  $T$  satisfies Equation (21). Then, the swarm position and velocity centroids are uniformly ultimately bounded, and  $C(k)$  will converge to the set  $\bar{\Omega}_b$ , where*

$$\bar{\Omega}_b = \left\{ C : \|C\| \leq \frac{\bar{\gamma}}{2} \left( \bar{\beta}_M \|\bar{A}\| + \sqrt{\|\bar{A}\|^2 \bar{\beta}_M^2 + 2\bar{\beta}_M} \right), i = 1, 2, \dots, N \right\} \quad (19)$$

is attractive and compact, with  $\bar{\beta}_M = 2\lambda_{\max}(\bar{P})$  and

$$\bar{\gamma} = (k_p D_p + k_m D_m) T \quad (20)$$

**Proof:** Using Lemma 1,  $\bar{A}$  is convergent if

$$T < \begin{cases} \frac{4}{k_d + \sqrt{k_d^2 - 4k_m}} & \text{if } k_d^2 - 4k_m \geq 0 \\ \frac{k_d}{k_m} & \text{if } k_d^2 - 4k_m < 0 \end{cases} \quad (21)$$

Choose a Lyapunov function  $\bar{V}(k) = C(k)^\top \bar{P} C(k)$  where  $\bar{P} = \bar{P}^\top > 0$  is a  $2n \times 2n$  positive definite matrix. Then, using the approach in the last proof,  $-\bar{Q} = \bar{A}^\top \bar{P} \bar{A} - \bar{P}$ , let  $\bar{Q} = I$ , and  $\bar{\beta}_M = 2\lambda_{\max}(\bar{P})$ . ■

**Remark 3** *Since  $\bar{\Omega}_b$  is attractive, if the swarm position centroid starts out large relative to the swarm marker, it will eventually become within a known bound of the swarm marker.*

Also, if the swarm velocity centroid starts out very large, it will eventually become within a known size.

**Remark 4** *If there is a higher magnitude of sensing noise (quantified by the noise bounds  $D_p$  and  $D_m$ ), then intuitively it is expected that the error between the swarm position centroid and swarm marker, along with the velocity centroid, will grow. The above result is consistent with this intuition. In particular, if  $D_p$  or  $D_m$  increase, the size of  $\bar{\Omega}_b$ , denoted by  $|\bar{\Omega}_b|$ , increases as can be seen by inspecting Equations (19) and (20). Note, however, that unlike in Theorem 1 neither the repulsion gain  $k_r$ , size of the repulsion range  $r_s$ , nor  $N$  affect the size of the bound.*

**Remark 5** *Note that if  $N$  is large, and  $d_p^i(k)$  and  $d_m^i(k)$  are zero mean random variables for each  $k$ , then  $D_p$  and  $D_m$  in the statement of the theorem can be replaced by  $\frac{1}{2}D_p$  and  $\frac{1}{2}D_m$ , respectively, since these overbound  $\bar{d}_p(k)$  and  $\bar{d}_m(k)$ . Hence, in some cases  $|\bar{\Omega}_b|$  can be reduced in size.*

**Remark 6** *Each of the two theorems require that  $T$  be sufficiently small in order to obtain the bounds. These conditions, in Equations (9) and (21), are both satisfied for the values of  $T$ ,  $k_p$ ,  $k_m$ , and  $k_d$  found for the validated midge swarm model in Section 2. The bounds determined in each of the two theorems can be computed from the parameters for the validated midge swarm model in Section 2. Generally, however, the bounds found in this paper are conservative since multiple overbounding arguments are used in the proofs of the theorems. Nevertheless, the analysis process itself provides significant insights into the swarm dynamics that emerge from local flight rules of individual midges.*

#### 4. Conclusions

A mathematical model was introduced and shown to produce trajectory behavior, and velocity, speed, and acceleration distributions, that are quite similar to those found in the midge swarm experiments in (Okubo and Chiang, 1974; Okubo et al., 1976). The first main mathematical result showed that all midge position and velocity trajectories are eventually bounded relative to the swarm position and velocity centroid trajectories, respectively. The second main result showed that the distance between the swarm position centroid and the swarm marker is ultimately bounded, as is the velocity centroid. Explicit bounds were provided in terms of midge parameters for both results, and several insights were gained on how the midge flight rule results in swarm cohesion.

In the modeling study, it is noteworthy that the only attraction term for the flight rule that led to a good representation of the experimental data was  $-M_i k_p \hat{e}_p^i(k)$  with sensing noise of a significant magnitude. Other attraction rules that used less information (e.g., the ones based on metric and topological distance in Sections 2.3.2 and 2.3.3) could not match the experimental data unless they approximated  $-M_i k_p \hat{e}_p^i(k)$ . This provides evidence that individual swarm members try to move toward a noisy estimate of the swarm centroid in order for the swarm to maintain cohesion. Superficially, though, it appears to be impossible for any midge to implement the  $-M_i k_p \hat{e}_p^i(k)$  attraction term since it is unlikely that any midge will know the positions of all other midges at any point in time so it can compute the centroid. It is not, however, claimed here that any midge explicitly *computes* the centroid. All that is needed is for the midge to be able to aggregate midge position information (e.g., via gross patterns detected in its visual system); its aggregation of the information can be a very noisy estimate of the underlying centroid of positions of the other midges. Recall that when a midge is in the middle of the swarm, it will undoubtedly perceive the centroid to be some distance away (e.g., along the edge of the swarm), and hence will move away from

the actual swarm position centroid. There is only one instance when the  $-M_i k_p \hat{e}_p^i(k)$  term contributes a force input that definitely corresponds to a midge making progress toward moving to the position centroid: If the midge is far away from the position centroid (e.g., outside the main body of the swarm), the term will result in movement in the *general direction* of the swarm position centroid since the noise is bounded about the position centroid.

Next, note that the matching of the midge’s experimental speed distribution to the Maxwell-Boltzmann speed distribution for molecules in a gas was first recognized in (Okubo and Chiang, 1974). Clearly, this matching means that there is significant randomness in speeds of the midges in a swarm. Indeed, the modeling study here shows that without noise of a significant magnitude driving the velocity (i.e., the second equation in Equation (2)) it was not possible to obtain the Maxwell-Boltzmann speed distribution so that the swarm data could be properly represented. Here, noise was injected in two places. First, it played a significant role in defining the sensing noise on the attraction term as was just discussed. Second, there was noise in sensing the swarm marker. Having these two noise sources allowed for independent adjustment of how close the swarm tracked the swarm marker so that this feature could also be represented with the model.

While the noisy attraction term plays a key role in representing the swarm, the repulsion term *does not* since it acts on such short distances, essentially defining what happens when midges collide. Indeed, the experimental data could have been represented just as well with either the repulsion term in Equation (4) or the  $+M_i k_r e_{pr}^i(k)$  repulsion term defined in Section 2.3.2. Moreover, since  $e_{pr}^i(k) \leq (N - 1)r_r$ , the cohesiveness results of the last section will also hold if  $+M_i k_r e_{pr}^i(k)$  is used as the repulsion term in Equation (4) (but, different bounds will be obtained).

In the analysis, it is noteworthy that two key derivations made the cohesiveness

analysis possible. First, both the average error between any midge and the swarm position centroid and the average repulsion force for the swarm were zero (see equations just below Equation (7)). This arises since if there were no noise, every pair of midges in the swarm attract each other or repel each other by equal amounts, but in opposite directions. Since the average repulsion force is zero, one less term is included in  $U^i(k)$  that needs to be over-bounded. The fact that  $\frac{1}{N} \sum_{j=1}^N e_p^j = 0$  simplifies the proof; if it were not zero then there is additional cross-coupling in the swarm between the  $E^i(k)$  and  $E^j(k)$ ,  $j \neq i$ , that would have to be accounted for in the analysis. Second, it was possible to load all the terms that could be bounded from above into  $U^i(k)$ . These two facts allowed for the derivation of the error dynamics in Equation (8) that had the special form of a linear term (specified via  $A$ ) and a nonlinear stochastic term quantified via  $U^i(k)$ . Lemma 1 showed that Equation (8) is a stable linear system driven by a bounded nonlinear input term. The remainder of the proof of ultimate boundedness used a sequence of standard ideas from algebra and analysis. Next, the swarm position centroid minus the swarm marker, and the velocity centroid dynamics, were derived using similar ideas to the error dynamics in Equation (8) (e.g, the fact that  $\frac{1}{N} \sum_{j=1}^N e_p^j = 0$ ). Due to the similarity of the two error systems it was possible to use the same proof pattern to show that the dynamics were uniformly ultimately bounded. For both cases, the matrices  $A$  and  $\bar{A}$  were shown to have eigenvalues in the unit disk for the swarm parameters found in Section 2. The ultimate bounds were also computed from the midge parameters, but *for the simulations conducted* these bounds were quite conservative (several orders of magnitude larger than the achieved diameter of the swarm by the end of a 20 sec. simulation run).

For future work, it would first be of interest to perform experiments to determine if the midge flight rule is indeed encapsulated by Equation (4). Second, it would be useful to show that the mathematical results still hold if there are different gains used in the midge flight rule in Equation (4) for each midge (e.g., different  $k_p$  gains on each midge) since it is

unlikely that individual midges use *exactly* the same gains. Third, it would be useful to find tighter ultimate bounds on the error and centroid trajectories.

**Acknowledgements:** The author would like to thank Kevin M. Schultz and Benjamin E. Keller for inputs, along with Yanfei Liu’s earlier collaboration on our past work that impacted this work.

## REFERENCES

- M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *PNAS*, 105(4):1232–1237, 2008.
- H. C. Chiang. Ecology of insect swarms V. Movement of individuals *Anarete Pritchardi* within a swarm. *Ann. Entomological Soc. of America*, 61(3):584–587, 1968.
- H. C. Chiang, B. J. Mettler, A. Okubo, and A. S. Robbins. Coupling of midge individuals in a swarm *Anarete Pritchardi* Kim (Diptera: Cecidomyiidae). *Ann. Entomological Soc. of America*, 71:859–861, 1978.
- H. C. Chiang, A. Goldsmith, and A. Okubo. Interaction of male and female midges *Anarete Pritchardi* Kim leading to coupling. *Ann. Entomological Soc. of America*, 73(5): 504–513, 1980.
- I. D. Couzin, J. Krause, R. James, G. D. Ruxton, and N. R. Franks. Collective memory and spatial sorting in animal groups. *J. Theor. Biol.*, 218:1–11, 2002.
- V. Gazi and K. M. Passino. Stability analysis of swarms. *IEEE Trans. on Automatic Control*, 48(4):692–697, April 2003.
- S. Gueron, S. A. Levin, and D. I. Rubenstein. The dynamics of herds: From individuals to aggregations. *J. Theor. Biol.*, 182:85–98, 1996.
- Y. Liu and K. Passino. Stable social foraging swarms in a noisy environment. *IEEE Trans. on Automatic Control*, 49(1):30–44, 2004a.
- Y. Liu and K. M. Passino. Cohesive behaviors of multiple cooperative mobile discrete-time agents in a noisy environment. In D. Grundel, R. Murphey, and P. Pardalos, editors,



- Theory and Algorithms for Cooperative Systems*, pages 213–238. World Scientific Pub., NJ, 2004b.
- A. Mogilner and L. Edelstein-Keshet. A non-local model for a swarm. *J. Math. Biol.*, 38: 534–570, 1999.
- A. Mogilner, L. Edelstein-Keshet, L. Bent, and A. Spiros. Mutual interactions, potentials, and individual distance in a social aggregation. *J. Math Biol.*, 47:353–389, 2003.
- A. Okubo. Dynamical aspects of animal grouping: Swarms, schools, flocks, and herds. *Adv. Biophys.*, 22:1–94, 1986.
- A. Okubo and H. C. Chiang. An analysis of the kinematics of swarming of *Anarete Pritchardi Kim* (Diptera: Cecidomyiidae). *Res. Popul. Ecol.*, 16:1–42, 1974.
- A. Okubo, H. C. Chiang, and C. C. Ebbesmeyer. Acceleration field of individual midges *Anarete Pritchardi Kim* within a swarm. *Can. Entomol.*, 109:149–156, 1976.
- A. Okubo, D. Grunbaum, and L. Edelstein-Keshet. The dynamics of animal grouping. In A. Okubo and S. A. Levin, editors, *Diffusion and Ecological Problems: Modern Perspectives*, pages 197–237. Springer Verlag, NY, 2001.
- J.K. Parrish and W.M. Hamner, editors. *Animal Groups in Three Dimensions*. Cambridge Univ. Press, Cambridge, England, 1997.
- K. M. Schultz, K. M. Passino, and T. D. Seeley. The mechanism of flight guidance in honeybee swarms: Subtle guides or streaker bees? *J. Exp. Biol.*, 211:3287–3295, 2008.