A novel control mechanism for natural flocks

Andrea Cavagna1,2, Antonio Culla2,1, Xiao Feng1,2, Irene Giardina2,1,3, Tomas S. Grigera4,5,6, Willow Kion-Crosby1,2, Stefania Melillo1,2, Giulia Pisegna2,1, Lorena Postiglione1,2, Pablo Villegas1,2
1 Istituto Sistemi Complessi, Consiglio Nazionale delle Ricerche, UOS Sapienza, 00185 Rome, Italy
2 Dipartimento di Fisica, Università Sapienza, 00185 Rome, Italy
3 INFN, Unità di Roma 1, 00185 Rome, Italy
4 Instituto de Física de Líquidos y Sistemas Biológicos CONICET - Universidad Nacional de La Plata, La Plata, Argentina
5 CCT CONICET La Plata, Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina and
6 Departamento de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Argentina

Speed fluctuations of individual birds within natural flocks are moderate, as it is natural to expect given the aerodynamic, energetic and biomechanical constraints of flight. Yet the spatial correlations of such fluctuations have a range as wide as the entire group: the speed change of a bird influences and is influenced by that of birds on the other side of the flock. Long-range correlations and limited speed fluctuations set conflicting constraints on the mechanism controlling the individual speed of each bird, as the factors boosting correlations also tend to amplify speed fluctuations, and vice versa. Here, building on experimental field data on starling flocks with group sizes spanning an unprecedented interval of over two orders of magnitude, we develop a novel type of speed control mechanism that generates scale-free correlations yet allowing for biologically plausible values of the flocks’ speed. The bare-bones hallmark of this new control is a strongly nonlinear speed-restoring force that ignores small deviations of the individual speed from its natural reference value, while ferociously suppressing larger speed fluctuations. Numerical simulations of self-propelled particles flocks in the same size range as the empirical data fully support our theoretical results. The match of field empirical data, theory and simulations over such an extensive range of sizes suggests that the novel speed control we propose is the most economic theoretical mechanism able to reproduce all key experimental traits of real flocks.

Since the early stages of the effort to formulate a mathematical description of organized flocking behavior, as observed for example in starlings (Sturnus vulgaris), dunlins (Calidris alpina), or - on a smaller scale - pigeons (Columba livia), the fundamental dynamical rule common to all models has been that of mutual local imitation: each individual within the group tends to adjust its state of motion to make it as similar as possible to that of its neighbors [1,2]. This type of imitative behavior can be either explicitly prescribed by the model through a direct interaction between the animals’ velocities [1,3–5], or it may be an effective interaction emerging from simpler positional rules, as attraction and repulsion [2,4–8], depending on the coarse-graining level we decide to work at. In either case, imitation of the local neighbors is the cornerstone of organized flocking dynamics. Imitation is also a fundamental player in the context of statistical physics - in the form of ferromagnetic interactions [1,5] - and therefore this simple behavioral trait establishes a strong link between the study of collective behavior in biology and in physics [1,6]. Most of the classic models of flocking assume that all the individuals within the group move with the same constant speed [1,6]. In that case, mutual imitation only requires the individuals to adapt the orientation of their velocity orientation to that of their neighbors, so that imitation is entirely expressed by an effective alignment interaction. Beside the unimpeachable strive for simplicity of the early models, which allowed to make rapid progress in the understanding of the emergence of collective order, the fixed-speed assumption seemed to be quite harmless from a biological point of view, as in all the classic biological examples of collective behavior it seemed that speed fluctuations did not play any significant role. In fact, empirical evidence showed otherwise.

In real flocks the individual speeds of the birds are not all the same and they are not constant in time: data gathered in the field show very clearly that speed fluctuations cannot be neglected [1,7]. Hence, the general mechanism of mutual imitation requires that a bird also adjusts its speed to be as close as possible to that of its neighbors, similarly to the way it regulates its direction of motion. There is an important difference, though. The orientation of the velocity vector is expressed by an angle, hence it cannot drift to unreasonably large (or small) values; therefore, unless there is a bias towards a certain direction - which is typically not the case in cluster flocks during aerial display - it makes perfect sense to leave individual orientations entirely subject to the self-organized imitative interaction among the birds, without setting any preferential direction of motion. After all, the true marvel of collective behavior is that coordinated motion of the group emerges even without any preferential directional bias. In contrast, the control of the individual speed cannot be left just to mutual imitation, as nothing then would prevent the birds to move - all in sync - at the speed of light! Clearly, common sense urges us to devise a control mechanism aimed at keeping the in-

1 Classic reviews of collective behavior in biological systems can be found in [2,3].
Loosely speaking, we can think of $v_0$ as the speed at which a single free bird would cruise in absence of any external perturbations.

\[ \text{FIG. 1: Qualitative sketch of linear vs marginal speed-restoring force.} \] In the linear case the force pulls the speed back to its natural reference value $v_0$ proportionally to the deviation from $v_0$, that is with a constant rate; the slope of the linear force, $g$, is the speed stiffness. Instead, in the marginal case, the force is extremely weak for small deviations from the reference speed, while it increases very sharply for large deviations, harshly suppressing them in a highly nonlinear fashion.

Individual speed of each bird at least in the ballpark of some reasonable biological value. For aerodynamic and biomechanical reasons, birds of a given species can fly only within a certain speed range, not too slow, not too fast. Coming to an example which is relevant for the empirical data we discuss here, starlings fly at a typical speed of about 12 meters-per-second (m/s) \[18\]. We shall call this the reference natural speed of the species under consideration and indicate it with $v_0$. Hence, speed has to respond to an inter-individual interaction based on mutual imitation, like orientation, but, unlike orientation, it must also be subject to a control mechanism that acts on a purely individual basis.

By far the most straightforward way to maintain the individual speed of each animal around the natural reference value, $v_0$, is to have in the equations of motion a linear speed-restoring force: whenever the speed $v_i$ of bird $i$ deviates from the natural reference value $v_0$, it gets ‘pushed back’ proportionally to the deviation from $v_0$ (see Fig.1). This force is fully described by just one parameter, namely its slope – or speed stiffness – $g$. Linear speed control is widely used in the field of collective behavior, not only to describe bird flocks [19, 20], but also fish schools [21], pedestrian collectives [22, 24], and even group dynamics of flying robots \[27\] and vehicle crowd interaction \[26\]. Apart from its simplicity, linear control neatly captures a very important empirical trait of real flocks, namely the fact that the correlations of the speed fluctuations are long-ranged \[17\], endowing these systems with remarkable collective coordination. It has been shown in [19] that the lower is the force stiffness $g$, the larger is the range of the correlation, so that when $g$ is small enough, long-range speed correlations are correctly reproduced by linear speed control. Linear speed control has also been tested in the very large scale numerical simulations of [20], which showed that too strong a stiffness $g$ not only hinders long-range correlations, but it also leads to significant group fragmentation.

Despite its simplicity, success and ubiquity in the literature, linear speed control entails an intrinsic conflict that has never been uncovered until now, and that we will show to be extremely difficult to resolve. Very large natural flocks display very long-range correlations; to describe this situation, a flocking model based on linear speed control requires a very low value of the linear speed stiffness, $g$; but in so doing the speed-restoring force becomes very weak, and as a consequence the animal’s speed becomes too loosely bounded around its reference value, $v_0$. When this happens, the typical speed of a flock may drift, growing unreasonably larger than the reference natural speed, $v_0$. We will present in this work new empirical data showing that in real flocks of birds the conflict between long-range correlation and limited natural speed entailed by linear control is impossible to resolve without fine-tuning in a tricky size-dependent fashion the speed stiffness, which seems rather unrealistic from a biological point of view. We will also perform numerical simulations of self-propelled particle flocks ruled by linear speed control, which demonstrate the aforementioned conflict in a very compelling way.

Here, we propose a different control mechanism, whose fundamental idea is quite simple: small speed fluctuations elicit nearly zero restoring force, while larger speed fluctuations are pushed back extremely sharply (see Fig.1). For mathematical reasons that will be clearer later on, we call this marginal speed control. This type of strongly non-linear control is inspired by some general biological considerations. There are two main factors regulating speed fluctuations: (i) energetic expenditure concerns; (ii) biomechanical constraints connected to the aerodynamics of flight. Small speed fluctuations are not prevented by biomechanical constraints, hence they are only regulated by energetic factors, but how sharply? Birds like starlings have high metabolic rates, so one would expect saving energy to be a critical issue. And yet, starlings during aerial display prove to be very liberal about their energy expenditure habits, as it has been much reflected upon in the literature \[27, 29\]. When starlings return at dusk to their roost, flight time would be expected to be kept at a minimum, because their metabolic rate is ten times higher in flight than while roosting \[27\]. Hence, from an energetics standpoint, the
most economic behavior would be to fly directly to the roost and immediately land (incidentally, this would also be by far the best course of action from an antipredatory point of view [29]). But this is not what happens. Instead, flocks will turn and wheel over the roost for a half hour before landing, a time during which they are expending energy at an outrageous rate. This empirical fact suggests that, at least during roost flocking, energy expenditure is a very mild concern of starlings [30]. It is reasonable, then, that small extra energy expenditures (and therefore small speed fluctuations) are weaker-than-linearly suppressed. On the other hand, large speed fluctuations clash against biomechanical and aerodynamic constraints, which are set very stringently by anatomy, physiology and physics [31-33]; therefore, a stronger-than-linear suppression of large speed fluctuations also seems reasonable. In a nutshell, marginal control simply means that birds exploit all speed fluctuations they are allowed to by the hard biomechanical constraints. The great advantage of this kind of control is that its low stiffness at small fluctuations boosts correlation, while its sharp increase for large fluctuations always grants a plausible speed to the group. Here, we will provide theoretical, numerical and - most importantly - field empirical evidence indicating that only a model of flocking based on marginal speed control can produce long-range correlations and biologically acceptable speed without the need to fine-tune any parameters.

Empirical evidence from field observations

The first large-scale 3D experimental data on flocks of starlings in the field were reported in [34] and [35], and they are summarised in Fig.2, in which we also added the data-acquisition campaigns of [36] and the new data from a campaign conducted in 2019-2020 (see SI for details of the experiments, and Table S1 in the SI for all biological data in each acquisition). Overall, this data-set spans group sizes between \( N = 10 \) and \( N = 3000 \) animals, a wider interval than any previously reported study. The three main experimental results that are of interest for us here are the following:

i) Flocks are highly ordered systems. The polarization, \( \Phi = (1/N) \sum v_i/v_i \) (where \( N \) is the total number of birds in the flock, \( v_i \) is the vector velocity of bird \( i \), and \( v_i \) is its modulus, i.e. speed), is always quite large, typically above 0.9 (Fig.2a). This observation rules out the possibility that flocks are close to a standard ordering transition, as for example that occurring in the Vicsek model [4, 14] in the biological context, or the classical Heisenberg model [15, 37–39] in the physical one. Instead, these are clearly systems deep into their ordered phase. This is nothing new, of course, but it is still conceptually relevant to keep it in mind when trying to find a comprehensive theory. From a statistical physics point of view, strong order implies that flocks cannot be described by a theory close to a standard critical point [15], a fact that is particularly relevant when we come to the next piece of experimental evidence.

ii) Empirical data indicate that speed fluctuations are correlated over long distances [17]. Even though di-
rect inter-individual interactions between birds are very much short-ranged, only involving a handful of neighbors around each focal bird [40, 41], equal-time spatial correlations of the speed fluctuations (see SI for the mathematical definition) have a spatial range that scales with the system’s linear size, $L$ (Fig.2b). This means that, in a statistical sense, the change of orientation and speed of some bird $i$ influences and is influenced by that of another bird $j$ on the other side of the flock, irrespective of the size of the flock. When this state of affairs holds, we say that the system enjoys scale-free correlations; we stress that scale-free correlations in starling flocks are observed both in the fluctuations of the direction of motions of the birds and in the fluctuations of their speed.\footnote{Unlike the case of orientations, scale-free correlations of the speed cannot be explained as the effect of spontaneous breaking of a continuous symmetry - i.e. through the Goldstone mechanism\cite{37,39} - simply because the speed is a scalar degree of freedom and the fact that flocks have a certain non-zero speed does not break any continuous symmetry.} If we consider the previous empirical result, namely large collective order, scale-free correlations of the speed appear to be a very unusual feature in the context of statistical physics, as fluctuations of the modulus of the degrees of freedom are normally heavily suppressed in the ordered phase and very much short-range correlated\cite{37}. Hence, this is the piece of phenomenology that really needs some non-trivial speed control mechanism.

iii) Flock-to-flock speed fluctuations are quite moderate (Fig.2c). The average cruising speed of starlings within a flock is about 12 meters-per-second (m/s), with typical fluctuations of 2 m/s\cite{17}. This is also the typical cruising speed of an entire flock, namely $v = \langle 1/N \sum v_i \rangle$ (where $v_i$ is the speed of bird $i$ and $N$ the total number of birds in the flock), whose distribution is reported in (Fig.2c); even the most extreme deviations of the speed of the group with respect to the typical value are rarely beyond 4 m/s. It is important to stress again that this is true both at the individual and at the collective level: what we report in Fig.2c is the distribution of the mean speed of the flock, across all recorded flocks, while in\cite{17} we reported the distribution of the bird-to-bird speed fluctuations within a flock; neither the individuals, nor the groups ever cruise at a speed much different from the natural reference value. This empirical result is an utterly natural thing to expect, given the dynamics of flight and the biomechanical constraints that this entails; we would be very surprised indeed to find variations of several orders of magnitude in the speed of the flocks! Yet, as we shall see, once we will give a mathematical description of the problem, this seemingly puny requirement may become tremendously difficult to fulfill, and - together with scale-free correlations - it poses very stringent constraints on the speed control mechanism.

Theory and simulations

The reference flocking dynamics we will consider here is the Vicsek model\cite{3, 11, 45}, in which the animals’ velocities interact through an explicit velocity coupling, aimed at describing the effective imitation between neighboring individuals. This kind of dynamics can be written in a compact way as follows,

$$\frac{dv_i}{dt} = -\frac{\partial H}{\partial v_i} + \eta_i$$

(1)

$$\frac{dx_i}{dt} = v_i$$

(2)

where $\eta_i$ is a white noise with strength proportional to $T$, a parameter playing the role of an effective temperature in the statistical physics context, namely $\langle \eta_i(t)\eta_j(t') \rangle = 2dT\delta_{ij}\delta(t-t')$ (we will work in $d = 3$ spatial dimensions in the following); $H$ is a cost function (or effective Hamiltonian), whose derivatives with respect to $v_i$ represents the forces acting on the particles’ velocities.\footnote{The effective friction coefficient in front of $\dot{v}_i$ in\cite{1} can be set to 1 through an appropriate rescaling of time\cite{39}.} At a fairly general level, we can write,

$$H = \frac{1}{2} J \sum_{i,j} n_{ij} (v_i - v_j)^2 + \sum_i V(v_i)$$

(3)

where the first term represents the imitation interaction between particles’ velocities, having strength $J$, and the second term is the speed control term, which affects each particle independently. The adjacency matrix, $n_{ij}$, is 1 for interacting neighbors and 0 otherwise. The self-propulsion part of the dynamics,\cite{4}, implies that the interaction network depends on time, $n_{ij} = n_{ij}(t)$: this fact, plus the velocity-dependent form of Hamiltonian, makes the dynamics an out-of-equilibrium one\cite{45}.

In the original Vicsek model, as in most of the original models of collective behavior, the speed of the particles is kept constant through a hard constraint, $|v_i| = v_0$. Here we want to study speed fluctuations and their correlations, hence we have relaxed the hard Vicsek constraint and introduced the control ‘potential’, $V(v_i)$, whose aim is to keep the speed of each particle around the natural reference value, $v_0$. Since the potential enters into the equation for the velocity through its derivative, this derivative is what we called speed-restoring force in the Introduction. Note that we want the speed control potential to regulate just the modulus of the velocity, not its direction, hence it must be a rotationally symmetric function of the velocity vector.

We notice something important: once the speed is left free to fluctuate, the imitation term in the dynamics effectively encourages neighbors to have both a similar orientation and a similar speed. Therefore, the actual speed
of an individual may be different from \( v_0 \) because of mutual imitation. This highlights an important point: the natural reference speed \( v_0 \), and the tendency to stay close to it, is an \emph{a priori} individual factor, while the actual speed of a bird is the product of the interaction between that individual factor and the inter-individual interactions among the birds. These two levels, mutual interaction vs individual control, can compete with each other, making the dynamics of speed control significantly more complex than a dynamics solely based on alignment.

**Linear speed control**

The simplest control, and indeed the one used in all models to date with fluctuating speed, consists of a Gaussian potential confining the speed,

\[
V(v_i) = g (v_i - v_0)^2
\]

where \( v_i = |v_i| \). This potential generates a linear restoring force acting on the speed in the equation of motion \( i \), so this is the \emph{linear speed control} we discussed in the Introduction. The constant \( g \) is the stiffness of the restoring force, and it can be interpreted as the elastic constant of a spring keeping the speed around its natural reference value, \( v_0 \). Although the bounding potential is Gaussian, it is so in the \emph{speed}, i.e. the modulus of the velocity, \( |v_i| \), which is not a linear function of \( v_i \); hence, the model is in fact not Gaussian, which is the reason why some of the calculations we report in the SI are not as straightforward as one would expect.

The determination of the range of spatial correlations of the speed fluctuations, also known as correlation length \( \xi_{sp} \), in the case of a linear speed control has been worked out in [19], giving,

\[
\xi_{sp} = r_1 \left( \frac{Jn_e}{g} \right)^{1/2}
\]

where \( r_1 \) is the mean inter-particle distance and \( n_e \) is average number of interacting nearest-neighbors. The meaning of this result in terms of statistical physics is quite clear: the theory defined by [3], with the potential [4], has a Gaussian critical point at \( g = 0 \), where the correlation length diverges. Conversely, large values of the speed stiffness \( g \) suppress speed fluctuations and the range of their correlations [19]. This mechanism is shown in Fig[3], where we report the correlation length of the speed fluctuations, \( \xi_{sp} \), vs the system’s size, \( L \), in field empirical observations (black points) and in numerical simulations of self-propelled particles regulated by linear speed control, Eqs. [14] (colored points). As anticipated, the empirical correlation length scales linearly with the system size over the whole range of experimentally captured flocks. In simulations with linear speed control, one observes that when the speed stiffness \( g \) is small enough, \( \xi_{sp} \) correctly scales linearly with \( L \) over the whole range (dark red points). To understand how small is “small enough”, we notice that in order to have scale-free correlation at all observed sizes, one needs \( \xi_{sp} \gg L \) for each \( L \), a condition that, together with [5], leads to,

\[
g \ll \frac{a}{L_{\text{max}}}
\]

where \( L_{\text{max}} \) is the size of the largest flock in the dataset and \( a = r_1^2 Jn_e \) collects all size-independent quantities. When the stiffness is so small that [6] holds, scale-free correlations over the entire range of \( L \) are reproduced (dark red points in Fig[3]). Conversely, if \( g \) is larger than (or too close to) the bound set by [7], the range of the correlation grows linearly with \( L \) only up to a certain maximum size, and then it saturates to its bulk value given by [8] (orange and yellow points). We conclude that, if correlations were our only experimental concern, there would be no need to increase the speed stiffness \( g \) beyond its bound [6], and everything would be fine. When we turn to the next piece of empirical evidence, though, the plot thickens.

The second key observable we have to take into consideration in our match between theory and experiments is the mean speed of the flock,

\[
s = \frac{1}{N} \sum_i v_i
\]

As we have already remarked upon, \( s \) will be influenced by the speed control mechanism, trying to keep the speed of each bird close to the natural reference value \( v_0 \), but also by the imitation social force: due to equations [1]-[3], birds tend to copy the entire velocity of their neighbors, orientation \emph{and} speed, hence the speed control competes with the imitative interaction to determine the collective speed, \( s \), of the entire flock. This interplay and the non-analytic form of the linear control, makes the calculation of the probability distribution of \( s \) rather intricate. However, under some reasonable and experimentally justified approximations (see SI for the mathematical details), one obtains the following result,

\[
P(s) = \frac{1}{Z} s^2 \exp \left[ -\frac{Ng}{T} (s-v_0)^2 \right]
\]

where \( Z \) is a normalization factor. The average value of this distribution can be computed numerically for any value of \( g \) and \( N \), and is reported in Fig[3] (full lines) together with data from simulations and experiments. But to say something analytical we can use the peak of the distribution, that is the typical value of the mean speed, \( s_{\text{typical}} \), which is a very good proxy for the average in the case of [8]; this is given by,

\[
s_{\text{typical}} = \frac{1}{2} v_0 \left( 1 + \sqrt{1 + \frac{4T}{Ng^2v_0^2}} \right)
\]

There is much to learn from this result. First, for \( N \rightarrow \infty \) we get \( s_{\text{typical}} = v_0 \), so that all is good for large flocks,
FIG. 3: Linear vs Marginal speed control. a: Speed correlation length, $\xi_{sp}$, vs system’s size, $L$. Black points represent experimental data on natural flocks of starlings, clearly showing a scale-free scenario, in which $\xi_{sp}$ scales linearly with $L$ (Pearson coefficient $r = 0.97$, p-value $p < 5.0 \times 10^{-3}$; the black solid line is a linear fit to the data). Colored points (from light yellow to dark red) represent numerical simulations of self-propelled flocks ruled by linear speed control, at several values of the speed stiffness $g$. Numerical simulations yield scale-free correlations over the entire range only at the smallest value of the stiffness $g$ (dark red), while for larger values of $g$ the correlation length saturates to its bulk value at some finite value of $L$, thus deviating from experiments. b: Mean speed, $s$, vs number of individuals, $N$. Black points represent experimental data, showing no detectable dependence of mean speed on $N$ (Spearman coefficient $r_s = -0.13$, p-value $p_s = 0.21$; the black horizontal line marks the average value over all the flocks, which is equal to 11.9 m/s). Colored points represent numerical simulations with linear speed control for the same values of the speed stiffness $g$ as in panel a. Linear speed control gives a near constant speed compatible with experiments only at the largest value of the stiffness $g$ (light yellow), while smaller values of $g$ give a growth of the mean speed at small $N$ completely incompatible with experiments. Full lines (from light yellow to dark red) represent the analytical prediction from (8). A comparison between panels a and b shows that linear speed control with a fixed value of the stiffness $g$ is unable to reproduce at the same time experimental data for both correlation length and mean speed. c: Black points and line represent correlation length vs system’s size and their linear fit for experimental data, as in panel a. Blue points represent simulations of flocks ruled by marginal speed control, at one single value of the temperature/noise $T$, low enough to stay in the deeply ordered phase. Scale-free behavior is satisfactorily reproduced. d: Mean speed vs number of individuals from experimental data as in panel b (black points an line). Blue points represent marginal speed control simulations at the same value of the temperature (and of all other parameters) as in panel c, showing very weak size dependence of the speed across the whole biological range of the data. The blue solid line represents the theoretical prediction for marginal speed control, $\xi_{sp}$ (Pearson coefficient $r_s = -0.02$, p-value $p_s = 0.98$). Inset: Same numerical and theoretical data as in main panel, plotted over a smaller range to better appreciate the agreement between theory and simulations. Marginal speed control is able to match theory with experiments for both correlation length and mean speed with a single set of parameters. Numerical simulations have been performed using a discrete version of (1) and (2), as described in detail in the SI. For both kinds of simulations (linear and marginal) the parameters are chosen to ensure large polarization (between 0.9 and 1) as in the experimental data. Colored points in the figure correspond to averages over numerical data, error bars to standard deviations. Black points correspond to the median (over time) of experimental data for each individual flocking event, error bars to median absolute deviations. Numerical correlation lengths are reported on the same scale as experimental ones by matching the $\xi_{sp}$ vs $L$ curves at the fully scale-free value of the parameters. Numerical/theoretical speeds are reported on the same scale as experimental ones by matching the $v$ vs $N$ curves at the largest value of $N$.

as their typical (and also their mean) speed is just the same as the natural reference speed, $v_0$. Troubles emerge for finite systems, though, as the typical speed grows for decreasing $N$, potentially becoming absurdly larger than the natural reference value, $v_0$; and because in the combination $Ng$ appears, this disastrous effect is the more serious the smaller the speed stiffness $g$, so that for very weak control, even relatively large flocks will have
a biologically implausible speed. The physical reason for this is that when the imitation force is strong (as it must be in the ordered phase of a flock) and the speed-restoring force is weak, the system is allowed to gain entropy by increasing in a coordinated fashion all the individual speeds of the particles, thus forming an unreasonably fast flock.

This theoretical prediction is confirmed by numerical simulations. Fig. 3b shows that, once the reference speed \(v_0\) of theory and experiments is matched at the largest sizes, for small values of \(N\) and \(g\) numerical flocks with linear speed control (dark red points) have a mean speed that is completely incompatible with actual experimental data in the same range of \(N\) (black points). To contrast the increase of the typical mean speed in smaller flocks one needs a larger value of the speed stiffness \(g\) (light yellow points), but we know from the previous discussion that this depresses the range of the speed correlations, so that one fails to reproduce scale-free correlations, Fig. 3a. This is the blanket-too-short dilemma of linear speed control: either we use a speed stiffness \(g\) small enough to reproduce scale-free correlations even at the largest observed values of \(N\), but in that case we get implausible large speed at low \(N\) (dark red points), or we increase \(g\) to tame the speed within the experimental fluctuations at low \(N\), but then we lose scale-free correlations at large \(N\) (light yellow points). Linear speed control cannot yield both results at the same time.

We can quantify the conflict within linear control by setting a second bound on the speed stiffness \(g\). From [9] we see that, in order to have a typical flock’s speed reasonably close to the natural reference value, \(v_0\), one must ensure that \(T/(Ngv_0^2) \ll 1\) for all observed sizes; if we use the reasonable approximation \(N \sim L^3/r_1^3\), where \(r_1\) is the mean nearest neighbor distance, we obtain the condition,

\[
g \gg \frac{b}{L_{\text{min}}^3}
\]

where \(L_{\text{min}}\) is the size of the smallest flock in the data-set, and where once again we have grouped into the parameter \(b = r_1^3T/v_0^2\) all size-independent constants. Once the spectrum of observed values of \(L\) is wide enough, the two bounds [9] and [10] cannot be satisfied both with one single value of the speed control stiffness \(g\). The only way to reconcile linear speed control with the empirical observations would be to assume the existence of a tuning mechanism such that the speed stiffness \(g\) depends on the size \(L\) of the flock, in a way to satisfy the following condition,

\[
b \frac{L}{L^3} \ll g(L) \ll a \frac{L}{L^2}
\]

This is a rather narrow strip for \(g(L)\) to live in, so that a biological mechanism fulfilling [11] would require some very tricky size-dependent fine-tuning. But even that could be insufficient: the two asymptotic inequalities in [11] require the stiffness \(g\) to stay well clear of both boundaries, \(b/L^3\) and \(a/L^2\), which for medium-small values of \(L\) may be very hard to achieve. Therefore, although not entirely impossible to hypothesize, this scenario seems far less likely than a theory that yields all pieces of empirical phenomenology at the same time, with no need of size-dependent tuning of the parameters.

**Beyond linear control: general theory**

In statistical physics, the correlation length \(\xi\) of a system is connected to the inverse of the quadratic curvature of the (renormalized) potential, calculated at its minimum [28, 39, 47]: very small curvature implies very large correlation length, so that a divergent \(\xi\) is always due to a zero second derivative (or marginal mode) along some direction of the (renormalized) potential. This is also the case for the Gaussian speed control [4]: the second derivative of the potential along the speed is proportional to \(g\), hence when \(g\) is small, the correlation length is large. The problem with the Gaussian case, however, is that by decreasing \(g\) we suppress the whole potential bounding the speed, not only its curvature at the minimum, hence giving a freeway to excessive fluctuations of the speed, ultimately resulting in an implausible large speed of flocks with small \(N\).

This state of affairs suggests that we must turn to a bounding potential that does not vanish entirely when its curvature does. To find this potential we proceed through general considerations of symmetry and simplicity. First, the potential must keep the speed around the reference natural value \(v_0\) and it must diverge for large values of the speed; secondly, it must be rotationally symmetric; third, it must have the simplest mathematical form compatible with the previous conditions and with the experimental evidence. The most general form of a rotationally symmetric potential that bounds the speed around the natural reference value \(v_0\) is \(V(v_i) = (v_i - v_0)^p\), where the integer power \(p \geq 2\) must be even, in order to produce a minimum of the potential at \(v_0\). For \(p = 2\) we have the classic \(O(n)\) potential of standard vector ferromagnets [37, 38] (see Fig 4 left). This theory is not suitable for our purposes, though, as in the low-temperature symmetry-broken phase it cannot develop zero curvature in the speed direction unless we put an amplitude \(g\) in front of the whole potential and let \(g\) itself go to zero; but this is what we already did for the Gaussian case, and we know it does not work.\(^5\) Indeed, for \(p = 2\) and for \(v_i = |v_i| \sim v_0\), we can write, \(V \sim (v_i - v_0)^2\), which is nothing else than the Gaussian potential analyzed in

\(^5\) Modulus fluctuations in the standard \(O(n)\) theory are different from longitudinal fluctuations; the latter, being coupled to the massless transverse fluctuations, are also scale-free (albeit with a weaker divergence), while the former are truly massive fluctuations, hence they have finite (in fact very small) correlation length and susceptibility [28, 39].
the previous Section, giving all the shortcomings of linear speed control.

Marginal speed control

The next simplest possibility is \( p = 4 \), which gives the following speed-control potential,

\[
V(v_i) = \frac{1}{v_0^4} \lambda (v_i \cdot v_i - v_0^2)^4
\]

(12)

where, thanks to the \( v_0^4 \) normalisation, the amplitude \( \lambda \) has the same physical dimensions as \( J \) and \( g \). This potential was first studied on purely theoretical grounds in \cite{51} and its shape is depicted in Fig.4, right. The crucial feature of the potential in (12) is that its second derivative in the direction of the speed is always zero. This lack of second derivative, which implies a lack of linear term in the speed control force in the equation of motion, is a feature called marginality \cite{47,51,52}, hence we will call this marginal speed control. The important point is that the marginal potential has zero second derivative irrespective of the value of the amplitude \( \lambda \); higher order powers in the expansion of the potential are nonzero, though, and in fact they are very steep, thus bounding the speed much more effectively to its reference value, \( v_0 \), compared to linear control. Correspondingly, the speed-restoring force is very weak for small deviations from \( v_0 \), but very strong for large deviations, as we anticipated at a qualitative level in Fig.1.

Marginal potential gives rise to an infinite correlation length of speed fluctuations under all physical conditions; in fact, it is not so. Speed correlations are regulated by the bounding potential (i.e. the energy), but also by the fluctuations induced by the noise (i.e. the entropy). A theoretical study of (12) shows that at very low noise the marginal potential dominates, so that speed fluctuations are indeed scale-free, while by increasing the noise the correlation is increasingly suppressed by entropic fluctuations \cite{51}. As a consequence, the marginal theory has a zero-noise (or zero-temperature) critical point, where the correlation length of the speed fluctuations diverges. A mean-field analysis \cite{51} shows that the correlation length diverges as,

\[
\xi_{sp} \sim \frac{1}{T^{1/2}}
\]

(13)

where the generalized temperature \( T \) is the strength of the noise in (1). This scenario has an interesting and very positive consequence: in the marginal theory, by simply decreasing the noise strength \( T \), we bring home two out of three empirical traits, that is a large polarization and a large correlation length, a somewhat unusual result within standard statistical systems.

\footnote{In field theory terms what happens is that at finite temperature entropy provides a non-zero second derivative of the renormalized potential, i.e. a non-zero mass of speed fluctuations, and therefore a finite speed correlation length. The mass goes to zero at \( T = 0 \), where the speed correlation length diverges \cite{51}.}

---

FIG. 4: Speed control potentials. **Left.** The classic O(n) potential, \( V(v_i) = (v_i \cdot v_i - v_0^2)^2 \), has flat curvature along the orientation direction (Goldstone mode), but non-zero curvature along the speed direction, hence giving short-range speed correlations in the low temperature ordered phase. The only way to increase the speed correlation length in this case is to decrease the amplitude of the whole potential, but in so doing one boosts speed fluctuations in a way incompatible with experiments. **Right.** The marginal potential, \( V(v_i) = (v_i \cdot v_i - v_0^2)^4 \), is the simplest non-Gaussian form with zero curvature also along the speed direction, giving scale-free speed correlations for low enough temperature, irrespective of the overall amplitude of the potential. The sharp increase of the potential for large values of \( |v_i| \) tames very sharply speed fluctuations, thus giving results perfectly in line with the experiments also for the mean speed of the flock. In both cases the minimum of the potential runs along the circle, \( |v_i| = v_0 \). For the sake of visualisation both potentials are displayed here for the simplest case of \( n = 2 \), that is for a two-components velocity vector, while of course in reality we have \( n = 3 \).
What about the constraint of having a biologically reasonable speed in this marginal case? The calculation of the distribution of the mean speed of flocks under the marginal potential is more complicated than in the linear case (see SI), but under some reasonable approximations one obtains,

\[
P(s) = \frac{1}{Z} s^2 \exp \left[ -\frac{N \lambda}{Tv_0^2} (v_0^2 - s^2)^2 \right]
\]  

(14)

which is formally similar to the linear case, but with very different powers into the exponential, which tame speed fluctuations extremely sharply. The maximum of this distribution, i.e. the typical mean speed of the flock, is given by (see SI for details),

\[
s_{\text{typical}} \simeq \begin{cases} 
  v_0 & \text{for } N \gg \frac{T}{\lambda v_0^2} \\
  v_0 \left( \frac{T}{4N\lambda v_0^2} \right)^{1/8} & \text{for } N \ll \frac{T}{\lambda v_0^2}
\end{cases}
\]  

(15)

As in the linear case, for \( N \to \infty \), the typical speed of the flocks becomes the same as the natural reference speed, \( v_0 \), while it increases for smaller sizes. However, in the marginal case there are two crucial differences with respect to the linear case: first, because the strength \( T \) of the noise is small (to have large correlation length), the crossover size \( N = T/(\lambda v_0^2) \), below which the mean speed increases, is small indeed, thus shielding this regime; second, the exponent of the speed’s increase for small \( N \) is 1/8, significantly smaller than the exponent 1/2 of linear speed control (see equation (9)). For these two reasons marginal speed control is so much more effective than linear control at small \( N \).

These theoretical results are fully confirmed by numerical simulations, and comparison with experimental data indicates that marginal control is indeed the right way to go (Fig.3a and Fig.3b): with just one reasonable set of parameters, the only crucial one in fact being the low noise strength, \( T \), numerical simulation of self-propelled flocks with marginal speed control fit the experimental data very well: the speed fluctuations correlation length \( \xi_s \) scales linearly with size up to the largest \( L \), and the mean speed shows only a very moderate increase at low \( N \), well within the scatter of the empirical data. Incidentally, we also notice that the analytical prediction given by (14) fits the numerical data quite well (Fig.3b-inset). We stress once again that all three pieces of phenomenology – large polarization, large correlation length, moderate speed at all group sizes – are achieved by marginal control by doing just one very sensible thing, namely setting the system in the low-noise phase. On the contrary, in the linear case large polarization must be achieved by tuning the alignment strength, \( J \), while correlation and speed are tuned by a second independent parameter, that is the stiffness \( g \); most worrying, correlation and speed are regulated by \( g \) in a conflicting way by linear control, a conflict completely resolved by marginal control.

Conclusions

Pioneering models and theories of collective behavior in biological systems, assumed constant speed of the microscopic particles, chiefly to simplify the mathematical description of the problem. In real natural systems, though, and in particular in bird flocks, speeds are not constant and they fluctuate in a very specific way, as they are long-range correlated, an unusual feature in standard statistical systems in their ordered phase. A first step forward towards understanding this phenomenon was done in [19] and [20], where it was shown that the simplest linear speed control mechanism is able to give scale-free correlations provided that the speed control stiffness is small enough. Due to the limited span of group sizes previously available to empirical observations in the field, though, these previous investigations were unable to detect that a linear speed control cannot grant both scale-free correlation and a biologically reasonable speed in the whole range of biological sizes. The reason is quite simple: in linear speed control there is an unavoidable conflict between increasing the range of the correlation and keeping moderate the mean speed of the group.

We showed here that among all nonlinear speed control mechanisms the marginal one stands out according to a criterion of mathematical simplicity and agreement with the experimental data. At the modeling level, the advantages of switching from linear to marginal speed control are significant. While with linear control collective order and speed correlation need to be tuned by independent parameters – the strength of imitation to achieve strong correlation – the marginal theory catches these two pigeons with one stone, as both high polarization and strong correlation are achieved by simply lowering the noise, that is by pushing deeper into the ordered phase. Most importantly, marginal control maintains the mean speed of the flocks limited, well within the experimental variability of the speed in real flocks, thanks to the non-linear terms in the speed-restoring force. Finally, the marginal potential is the simplest mathematical function satisfying all the above requirements. At the biological level, marginal control appears to be a rather natural mechanism, according to which small speed fluctuations are essentially free, at least during a flight phase in which birds’ priority is definitely not to save energy, while large speed fluctuations are heavily depressed by biomechanical constraints. This argument is fairly general, suggesting that marginal control may be applicable to a wider class of natural systems than bird flocks.

An interesting question is whether or not a theory with marginal speed control requires any tuning of the parameters, and in particular tuning close to criticality [53]. The success of the marginal theory is based on the fact that the second derivative of the potential is exactly zero at \( v_0 \), a condition that seems to require some tuning. In fact, a small non-zero quadratic term would still be acceptable in the marginal case, as long as its am-
plitude is much smaller than $1/L_{\text{max}}^{2}$, while - thanks to the steep nonlinear rise of the marginal potential - there is no lower bound for it; this tuning is therefore rather lukewarm. On the other hand, to reproduce the correct phenomenology of flocks the marginal theory requires the system to be deep into its ordered phase, namely close to the zero-temperature critical point, so in this sense there is a case for near-criticality. However, in the marginal case there is no need for any finite-size tuning: a zero-temperature critical point (unlike standard critical points) is not shifted by finite-size effects, so that it has always just one physical side (positive temperature); hence, we can just push the system at low temperature, without worrying to cross the critical point in any way; as a result, the control parameter does not need to depend on size to keep the system close to criticality. As we have seen, the situation would be different in the case of linear control, as the speed stiffness must be carefully tuned in a size-dependent way to remain close, but not too close, to the critical point.

It would be interesting to perform a full statistical physics analysis of the marginal theory, although computing analytically the critical exponents is not easy, even in the equilibrium case of a fixed network [51], let alone in the off-equilibrium active case of realistic biological systems. Surely, the equilibrium case should be analyzed first, by using a Renormalization Group (RG) approach. The calculation could be interesting, because of the existence of two critical points: the finite temperature one, where the order-disorder transition occurs, and the zero-temperature critical point, which is the truly relevant one for flocks. The structure of the RG flow between these two fixed points is likely to be very nontrivial. To tackle the off-equilibrium case one should turn to the hydrodynamic theory of flocking of Toner and Tu [54, 55], perhaps in the incompressible flavour of [50], and possibly with the inclusion of inertial mode-coupling terms [57], to account for linear information transfer within flocks [30]. Although extremely complicated, these seem the most compelling theoretical directions to explore in future studies.

Acknowledgements

This work was supported by ERC Advanced Grant RG BIO (785932) to ACa, and ERANET-CRIB grant to ACa and TSG. TSG was also supported by grants from CONICET, ANPCyT and UNLP (Argentina). The authors acknowledge several illuminating discussions with William Bialek regarding speed control in flocks. ACa wishes to thank Victor Martin-Major for careful advice on the numerics. ACa warmly thanks Frank Heppner for reading the original manuscript, and for a year-long conversation on collective avian behavior.

[1] C. W. Reynolds, “Flocks, herds and schools: A distributed behavioral model,” in Proceedings of the 14th annual conference on Computer graphics and interactive techniques, pp. 25–34, 1987.
[2] F. Heppner and U. Grenander, “A stochastic nonlinear model for coordinated bird flocks,” The ubiquity of chaos, vol. 233, p. 238, 1990.
[3] A. Huth and C. Wissel, “The simulation of the movement of fish schools,” Journal of theoretical biology, vol. 156, no. 3, pp. 365–385, 1992.
[4] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, “Novel type of phase transition in a system of self-driven particles,” Phys Rev Lett, vol. 75, pp. 1226–1229, Aug 1995.
[5] I. D. Couzin, J. Krause, R. James, G. D. Ruxton, and N. R. Franks, “Collective memory and spatial sorting in animal groups,” Journal of theoretical biology, vol. 218, no. 1, pp. 1–11, 2002.
[6] H. Chaté, F. Ginelli, G. Grégoire, F. Peruani, and F. Raynaud, “Modeling collective motion: Variations on the vicsek model,” Éur. Phys. J. B, vol. 64, pp. 451–456, 08 2008.
[7] P. Romanczuk and L. Schimansky-Geier, “Swarming and pattern formation due to selective attraction and repulsion,” Interface focus, vol. 2, no. 6, pp. 746–756, 2012.
[8] R. Grossmann, L. Schimansky-Geier, and P. Romanczuk, “Self-propelled particles with selective attraction-repulsion interaction: from microscopic dynamics to coarse-grained theories,” New Journal of Physics, vol. 15, no. 8, p. 085014, 2013.
[9] J. K. Parrish and W. M. Hamner, Animal Groups in Three Dimensions. Cambridge: Cambridge University Press, 1997.
[10] S. Camazine, N. R. Franks, J. Sneyd, E. Bonabeau, J.-L. Deneubourg, and G. Theraulaz, Self-Organization in Biological Systems. Princeton, NJ, USA: Princeton University Press, 2001.
[11] J. Krause and G. D. Ruxton, Living in groups. Oxford University Press, 2002.
[12] I. Couzin and J. Krause, “Self-organization and collective behavior in vertebrates,” Advances in the Study of Behavior, vol. 32, pp. 1–75, 2003.
[13] D. J. Sumpter, Collective animal behavior. Princeton University Press, 2010.
[14] T. Vicsek and A. Zafeiris, “Collective motion,” Physics Reports, vol. 517, no. 3, pp. 71–140, 2012.
[15] S.-K. Ma, Modern theory of critical phenomena. Advanced book classics, Perseus Pub, 2000.
[16] A. Cavagna, I. Giardina, and T. S. Grigera, “The physics of flocking: Correlation as a compass from experiments to theory,” Physics Reports, vol. 728, pp. 1–62, 2018.
[17] A. Cavagna, A. Cimarelli, I. Giardina, G. Parisi, R. Santagati, F. Stefanini, and M. Viale, “Scale-free correlations in starling flocks,” Proc Natl Acad Sci USA, vol. 107, pp. 11865–70, Jun 2010.
[18] J. M. Rayner, P. W. Viscardi, S. Ward, and J. R. Speakman, “Aerodynamics and energetics of intermittent flight in birds,” American Zoologist, vol. 41, no. 2, pp. 188–204,
W. J. Hamilton III, W. M. Gilbert, F. H. Heppner, and D. Yang, U. Ozguner, and K. Redmill, “Social force dynamics,” *European Journal of Social Psychology*, vol. 158, no. 3, pp. 563–578, 2015.

D. Helbing and P. Molnár, “Social force model for pedestrian dynamics,” *Phys. Rev. E*, vol. 51, pp. 4282–4286, May 1995.

D. Helbing, I. Farkas, and T. Vicsek, “Simulating dynamical features of escape panic,” *Nature*, vol. 407, pp. 487–490, September 2000.

D. Helbing, I. Farkas, P. Molnar, and T. Vicsek, “Simulation of pedestrian crowds in normal and evacuation situations,” vol. 21, pp. 21–38, 01 2002.

A. Garrell, L. Garza-Elizondo, M. Villamizar, F. Herrero, and A. Sanfeliz, “Aerial social force model: A new framework to accompany people using autonomous flying robots,” in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 7011–7017, 2017.

D. Yang, U. Ozguner, and K. Redmill, “Social force based microscopic modeling of vehicle-crowd interactions,” in *2018 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1537–1542, 2018.

W. J. Hamilton III, W. M. Gilbert, F. H. Heppner, and R. J. Planck, “Starling roost dispersal and a hypothetical mechanism regulating rhythmic animal movement to and from dispersal centers,” *Ecology*, vol. 48, no. 5, pp. 825–833, 1967.

F. H. Heppner, “Avian flight formations,” *Bird-bandng*, vol. 45, no. 2, pp. 160–169, 1974.

I. L. Bajec and F. H. Heppner, “Organized flight in birds,” *Animal Behaviour*, vol. 78, no. 4, pp. 777–789, 2009.

F. H. Heppner, “Private communication,” 2007.

C. J. Pennycook, “Mechanical constraints on the evolution of flight,” *Memoirs of the California Academy of Sciences*, vol. 8, pp. 83–98, 1986.

J. M. Rayner, “Form and function in avian flight,” in *Current ornithology*, pp. 1–66, Springer, 1988.

J. M. Rayner, “Biomechanical constraints on size in flying vertebrates,” in *Symposia of the Zoological Society of London*, no. 69, pp. 83–110, London: The Society, 1990–1999.

A. Cavagna, I. Giardina, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, “The starflag handbook on collective animal behaviour: 1. empirical methods,” *Anim Behav*, vol. 76, pp. 217–230, Jan 2008.

A. Cavagna, I. Giardina, A. Orlandi, G. Parisi, and A. Procaccini, “The starflag handbook on collective animal behaviour: 2. three-dimensional analysis,” *Anim Behav*, vol. 76, pp. 237–248, Jan 2008.

A. Attanasi, A. Cavagna, L. del Castello, I. Giardina, T. S. Grigera, A. Jelić, S. Melillo, L. Parisi, O. Pohl, E. Shen, et al., “Information transfer and behavioural inertia in starling flocks,” *Nature physics*, vol. 10, no. 9, pp. 691–696, 2014.

A. Z. Patashinskii and V. L. Pokrovskii, *Fluctuation Theory of Phase Transitions*. Pergamon Press, 1979.

M. Le Bellac, *Quantum and Statistical Field Theory*. Clarendon Press Oxford, 1991.

N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group*. Reading, Massachusetts: Perseus Books, 1992.

M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, et al., “Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study,” *Proceedings of the national academy of sciences*, vol. 105, no. 4, pp. 1232–1237, 2008.

W. Bialek, A. Cavagna, I. Giardina, T. Mora, E. Silvestri, M. Viale, and A. M. Walczak, “Statistical mechanics for natural flocks of birds,” *Proc Natl Acad Sci USA*, vol. 109, pp. 4786–91, Mar 2012.

A. Cavagna, L. Del Castello, S. Dey, I. Giardina, S. Melillo, L. Parisi, and M. Viale, “Short-range interactions versus long-range correlations in bird flocks,” *Physical Review E*, vol. 92, no. 1, p. 012705, 2015.

H. Ling, G. E. Melvor, J. Westley, K. Van der Vaart, J. Yin, R. T. Vaughan, A. Thornton, and N. T. Ouellette, “Collective turns in jackdaw flocks: kinematics and information transfer,” *Journal of the Royal Society Interface*, vol. 16, no. 159, p. 20190450, 2019.

H. Ling, G. E. Melvor, J. Westley, K. van der Vaart, R. T. Vaughan, A. Thornton, and N. T. Ouellette, “Behavioural plasticity and the transition to order in jackdaw flocks,” *Nature communications*, vol. 10, no. 1, pp. 1–7, 2019.

F. Ginelli, “The physics of the Vicsek model,” *The European Physical Journal Special Topics*, vol. 225, no. 11-12, pp. 2099–2117, 2016.

R. Zwanzig, *Nonequilibrium statistical mechanics*. Oxford University Press, USA, 2001.

J. J. Binney, N. Dowrick, A. Fisher, and M. Newman, *The theory of critical phenomena: an introduction to the renormalization group*. Oxford University Press, Inc., 1992.

E. Brézin, D. Wallace, and K. G. Wilson, “Feynman-graph expansion for the equation of state near the critical point,” *Physical Review B*, vol. 7, no. 1, p. 232, 1973.

E. Brézin and D. Wallace, “Critical behavior of a classical heisenberg ferromagnet with many degrees of freedom,” *Physical Review B*, vol. 7, no. 5, p. 1967, 1973.

A. Patashinskii and V. Pokrovskii, “Longitudinal susceptibility and correlations in degenerate systems,” *Zh. Eksp. Teor. Fiz.*, vol. 64, p. 1445, 1973.

A. Cavagna, A. Culla, L. Di Carlo, I. Giardina, and T. S. Grigera, “Low-temperature marginal ferromagnetism explains anomalous scale-free correlations in natural flocks,” *Comptes Rendus Physique*, vol. 20, pp. 319–328, May-Jun 2019.

C. Lupo, G. Parisi, and F. Ricci-Tersenghi, “The random field XY model on sparse random graphs shows replica symmetry breaking and marginally stable ferromagnetism,” *Journal of Physics A: Mathematical and Theoretical*, vol. 52, p. 284001, Jun 2019.

T. Mora and W. Bialek, “Are biological systems poised at criticality?,” *J Stat Phys*, vol. 144, pp. 268–302, Jul 2011.
Supplemental Materials: A novel control mechanism for natural flocks

Experiments

Experiments on starling flocks have been performed in Rome, from the terrace of Palazzo Massimo alle Terme, in front of a large roosting site at the Termini Railway Station. The experimental technique used is stereoscopic photography, where multiple synchronized video-sequences of a flocking event are acquired from different observation points with a calibrated multi-camera video-acquisition system [58]. Digital images are then analyzed with a specifically designed tracking software [59], in order to extract from the raw data the three-dimensional trajectories of the individual birds in the flock.

Data have been collected across the years during several experimental campaigns. The very first data were collected in the context of the Starflag project [34, 35], between 2007 and 2010, using Canon D1-Mark II cameras, shooting interlaced at 10 frames-per-second (FPS), with a resolution of 8.2 Megapixels (MP). A total of 24 acquisitions (i.e. flocking events) were reconstructed, with flocks sizes ranging between 100 and 2500 birds. This data-set was used in the analysis of [17, 19], where a linear speed control theory was considered.

A second campaign took place between 2011 and 2012, using faster cameras, namely IDT M5, shooting at 170FPS with a resolution of 5MP. During this campaign acquisitions were taken with group sizes ranging between 50 and 600 birds. The main aim of the experiment was to study orientational information transfer across flocks [36], so speed fluctuations were recorded but not analyzed at the time. The data from this set - included in the present work - fully support the previous scale-free results of speed correlations of [17], and usefully expand the size range for the determination of the mean flock speed in the low $N$ region.

A final campaign took place in the last months of 2019 and in January and February 2020. This campaign uses state-of-the-art IDT OS10-4K cameras, shooting at 155FPS with a resolution of 9.2MP. These new experiments focus on a dynamic protocol (i.e. acquisitions are taken with a panning stereo-system). However, 8 acquisitions have been taken with a static set-up analogous to the previous campaigns, and have been included in the present analysis. The group range of these acquisition is between 10 and 100 birds, further and crucially expanding the range at small $N$ to clearly differentiate between linear and marginal speed control.

All campaigns have been conducted with a three camera system exploiting trifocal geometry [60]. The image analysis - segmentation of individual birds, stereometric matching and dynamical tracking - have been performed using the method of [34, 35] for the first campaign, and the most advanced method of [59] for the second and third campaigns.

We summarize in Table I for each flocking event, all the experimental quantities used in our analysis.
Spatial correlation functions

The speed spatial connected correlation function is defined as \[16\]

\[
C(r) = \frac{\sum_{i,j} \delta v_i \delta v_j \delta(r - r_{ij})}{\sum_{i,j} \delta(r - r_{ij})}
\]  
(S1)

where \(N\) is the number of individuals in the system, \(r_{ij} = |r_i - r_j|\) is the mutual distance between individuals \(i\) and \(j\), and

\[
\delta v_i = v_i - \frac{1}{N} \sum_k v_k
\]  
(S2)

is the fluctuation of the individual speed \(v_i = |v_i|\) with respect to the mean speed of the group \(s = (1/N) \sum_i v_i\), evaluated at a given instant of time. The function \(C(r)\) represents the instantaneous average of mutual correlations among all pairs at distance \(r\): in systems with local, distance-dependent interactions, for large enough system sizes, this quantity is a good proxy of the typical correlation at that distance, as computed with the correct theoretical measure. A full discussion of definition (S1), its asymptotic limit, finite size effects, and behavior in known cases, can be found in \[16\]. Here we notice that the correlation function (S1) is the only possible definition applicable to experimental data, where no a priori information is available on the true nature of the dynamics. This definition has indeed been used in all the previous analysis of speed correlations mentioned in this paper. We display in Fig. S1 the correlation function (S1) computed, respectively, from experimental data (panel a), and from numerical simulations with a linear speed control model (panel b) and a marginal speed control model (panel c).

For each configuration of the system (at a given time) we estimate the correlation length as:

\[
\xi = \frac{\int_{0}^{r_0} \int_{0}^{r} r C(r) \, dr \, dC(r)}{\int_{0}^{r_0} \int_{0}^{r} dC(r)}
\]  
(S3)

and then we perform a time average of this quantity over different configurations. The point \(r_0\) is the first point for which \(C(r) = 0\). Such a point always exists due to the very definition of correlation function given in (S1) (the integral of \(C(r)\) between 0 and the size of the system \(L\) always vanishes). (S3) provides a reliable estimate of the correlation length in every regime, both when the system is scale-free with long-range correlations and when the system is far from criticality with short-range correlations (in the linear speed control model we can see all this phenomenology, simply by changing the parameter \(g\), see Fig. 3a in main text). The reason is that (S3) makes use of the information encoded in the zero-crossing point \(r_0\), together with the shape of the entire function \(C(r)\). When correlations are short-range and the correlation function is nearly exponential \[16\], e.g. \(C(r) \sim e^{-r/\xi}\), given that it almost vanishes after the point \(r_0\), we can think of extending the integrals of (S3) up to \(L\), and we obtain \(\xi \sim \hat{\xi}\). Conversely, when correlations are long range, regardless of the precise shape of the \(C(r)\), the first point of zero-crossing \(r_0\) dominates, hence we obtain \(\xi \sim r_0\). Other possible definitions are legitimate but, either they are reliable near criticality and they fail in the short-range correlation regime, or they capture the correlation range when it is much smaller than the system size, completely failing when the correlation range becomes extensive. For example, if one chooses \(r_0\) as an estimate for the correlation length, it works efficiently when the system is scale-free, and \(r_0\) identifies the size of the correlated domains, while it fails when the system has an intrinsic length-scale \[16\] \[17\]. On the other hand, in the phase of short-range correlations, it is easy to determine the correlation length via an exponential fit, because the \(C(r)\) has an exponential shape \[16\]; however this procedure is unfeasible in the critical regime when the correlation functions do not have an exponential behavior (see Fig. S1).

Distribution of the mean speed: linear speed control

In this section we describe how to derive the approximate mean speed distribution of Eq. (8) in the main text. The starting point is the pseudo-Hamiltonian with the Gaussian potential:

\[
H(\{v_i\}) = \frac{J}{2} \sum_{i,j} \delta v_i \delta v_j^2 + g \sum_i (v_i - v_0)^2
\]  
(S4)
FIG. S1: Some examples of connected correlation functions. We report some examples of speed connected correlation functions, computed by (S1). The first point of zero-crossing \( r_0 \) is visible for each function. All the functions are normalized such that \( C(r = 0) = 1 \).

\( a \): Example of speed connected correlation function in experimental data. 
\( b \): Example of speed connected correlation function in linear speed control model simulations. 
\( c \): Example of speed connected correlation function in marginal speed control model simulations. For \( b \) and \( c \) the distance \( r \) is measured in simulation units.

where all the sums are from 1 to the number of particles in the system \( N \). We are dealing with an active system, hence the matrix \( n_{ij} = n_{ij}(t) \) depends on time. However, it has been shown in [61] that, due to the large polarization of real flocks, the relaxation time scale of \( n_{ij}(t) \) is significantly larger than that of the velocities, so that a quasi-equilibrium approach to the problem is reasonable; from now on we will then consider a time-independent \( n_{ij} \). The validity of this approach is retrospectively confirmed by the remarkable agreement between the predictions of the approximate equilibrium theory derived here below, and the results from self propelled particles simulations, as displayed in Fig.3b and Fig.3d of the main text. In the context of quasi-equilibrium, we can assume a Boltzmann-like distribution for the velocities

\[
P(\{v_i\}) = \frac{1}{Z} \exp (-\beta H(\{v_i\})) .
\]

where \( \beta = 1/T \) is the inverse temperature, and quantifies the degree of noise in the system. Our aim is now to marginalize (S5) to get a probability distribution for the mean speed. It is convenient to rewrite (S4) in terms of the individual speeds \( v_i = |v_i| \) and flight directions \( \sigma_i = v_i/v_i \). In the very ordered phase, one can use the “spin-wave approximation” (SW) [62], as already done in previous analysis of starling flocks [19, 41]. When the polarization is large (enforced in our model by choosing \( J \gg 1 \)), the flight direction of each individual is very close to the polarization vector. Hence:

\[
\begin{align*}
\mathbf{v}_i &= v_i \sigma_i \quad \text{with} \quad |\sigma_i| = 1 \\
\sigma_i &\simeq n \left( 1 - \frac{\pi_i^2}{2} \right) + \pi_i
\end{align*}
\]

where \( n \) is the unit vector along the polarization vector \( \Phi = \frac{1}{N} \sum_i \sigma_i \), and the \( \pi_i \) are the fluctuations orthogonal to \( n \). The constraint \( \sum_i \pi_i = 0 \) holds by construction and, in the high ordered regime, \( \pi_i^2 \ll 1 \) for every \( i \). The Hamiltonian (S4) then becomes, up to order \( \pi_i^2 \):

\[
H(\{v_i\}, \{\pi_i\}) = J \sum_{i,j} \Lambda_{ij} v_i v_j + g \sum_i (v_i - v_0)^2 + J \sum_{i,j} \tilde{\Lambda}_{ij}(\{v_k\}) \pi_i \cdot \pi_j ,
\]

where we defined the matrices:

\[
\begin{align*}
\Lambda_{ij} &= -n_{ij} + \delta_{ij} \sum_k n_{ik} \quad \text{(Discrete Laplacian)} \\
\tilde{\Lambda}_{ij}(\{v_k\}) &= -n_{ij} v_i v_j + \delta_{ij} \sum_k n_{ik} v_i v_k
\end{align*}
\]

(S10)
In terms of the variables, the probability density \( P(\{v_i\}, \{\pi_i\}) \) becomes,

\[
P(\{v_i\}, \{\pi_i\}) = \frac{\delta \left( \sum_k \pi_k \right) \prod_i v_i^{d-1} e^{-\beta H}}{\int Dv' D\pi' \delta \left( \sum_k \pi'_k \right) e^{-\beta H} \prod_i v_i'^{d-1}}
\]  
\[(S11)\]

where \( Dv' = \prod_k dv'_k \), \( D\pi' = \prod_k d\pi'_k \) and \( d \) is the dimension of the velocity vector. We now need to integrate out the fluctuations \( \pi_i \), to obtain the marginalized distribution of the individual speeds \( v_i \). Let us define

\[
\Omega(\{v_i\}) = \prod_j v_j^{d-1} \int D\pi \exp \left[ -\beta J \sum_{i,j} \tilde{\Lambda}_{ij} (\{v_k\}) \pi_i \cdot \pi_j \right] \delta \left( \sum_k \pi_k \right)
\]  
\[(S12)\]

The integral can be easily performed upon a change of integration variables from the \( \{\pi_i\} \) to the eigenvectors \( \{\tilde{\pi}_\alpha\} \) of the matrix \( \tilde{\Lambda} \). Both \( \Lambda \) and \( \tilde{\Lambda} \) inherit the translational invariance of the original Hamiltonian and have a constant eigenvector corresponding to a zero mode, since \( \sum_j \tilde{\Lambda}_{ij} = \sum_j \Lambda_{ij} = 0 \). The constraint on the \( \{\pi_i\} \) becomes a constraint on the zero mode, i.e. \( \delta(\tilde{\pi}_0) \), making the integral finite and leaving out only \( d-1 \) eigenvalues. We get

\[
\Omega(\{v_i\}) = \left[ \prod_j v_j^{d-1} \right] \left[ \prod_{\alpha \neq 0} \tilde{\lambda}_\alpha (\{v_k\}) \right]^{-\frac{d-1}{2}}
\]  
\[(S13)\]

where the \( \{\tilde{\lambda}_\alpha\} \) are the eigenvalues of \( \tilde{\Lambda} \) and depend on the \( \{v_i\} \) in some complicated way. Since we are interested in the distribution of the mean speed \( s = (1/N) \sum_i v_i \), we will now estimate the behavior of \( \Omega \) to leading order in \( s \). Once again, it is convenient to make a change of variables, going from real space to the space of the eigenvectors \( \{\tilde{v}_a\} \) of the discrete Laplacian \( \Lambda \). Each \( v_i \) can be decomposed into its \( \tilde{v}_a \) components using the formula \( v_i = \sum_a w_i^{(a)} \tilde{v}_a \), where \( w_i^{(a)} \) is the change of basis matrix. As mentioned above, the zero-mode has constant coefficients \( w_i^{(0)} = 1/\sqrt{N} \) and the zero-mode eigenvector is therefore proportional to the mean speed, i.e. it is exactly \( \sqrt{N} s = \left( 1/\sqrt{N} \right) \sum_i v_i \). This also implies that for each \( v_i \) we have

\[
v_i = s + \delta v_i = s + \sum_{a \neq 0} w_i^{(a)} \tilde{v}_a.
\]  
\[(S14)\]

We can now express the function \( \Omega \), in terms of this new representation

\[
\Omega \sim \left[ \prod_{\alpha \neq 0} \tilde{\lambda}_\alpha (\{v_k\}) \right]^{-\frac{d-1}{2}} f \left( s + \sum_{a \neq 0} w_k^{(a)} \tilde{v}_a \right)^{d-1} \left[ \prod_j \left[ 1 + \sum_{a \neq 0} \frac{w_j^{(a)} v_j}{s} \right] \right]^{d-1} = \left( 1 + \sum_{a \neq 0} \frac{w_k^{(a)} \tilde{v}_a}{s} \right)^{d-1}.
\]  
\[(S15)\]

Here \( f \) is a generic rational function of its argument. The function \( f \) is a generic polynomial of order \( (N - 1)(d - 1) \) in its argument (from dimensional analysis), hence it is safe to extract a \( s^{(N-1)(d-1)} \), because \( s \) is present in the expansion of every \( v_k \).

The term \( \Omega \) describes the contribution to the measure coming from the integration of the directional fluctuations. Once we integrate the directional fluctuations, we have an Hamiltonian that only depends on the moduli \( \{v_i\} \). Also in this case, we can express everything in terms of \( s \) and the non-zero modes \( \{\tilde{v}_a\} \) of \( \Lambda \). Remembering that \( \sum_i w_i^{(a)} w_i^{(b)} = \delta_{a,b} \), we get:

\[
H = J \sum_{i,j} \Lambda_{ij} v_i v_j + g \sum_i (v_i - v_0)^2 = \sum_{a=1}^N (J \lambda_a + g) \tilde{v}_a^2 + gN(s - v_0)^2
\]  
\[(S16)\]

where, with a slight abuse of notation, we still indicate with \( H \) the marginalised Hamiltonian depending only on the speeds. After these manipulations we get the distribution

\[
P(\{s, \tilde{v}_a\}) = \frac{\Omega(\{s, \tilde{v}_a\}) e^{-\beta H}}{\int ds' Dv' \Omega(\{s', \tilde{v}_b\}) e^{-\beta H}}.
\]  
\[(S17)\]
with \( a \neq 0 \) and \( \text{Do}^\prime = \prod_{b \neq 0} d\hat{v}_b \) We can now derive the distribution of the mean speed \( s = \frac{1}{N} \sum_i v_i \) by marginalizing over all the non-zero modes \( \hat{v}_a \). To this end, we note that since \( |w_i^{(a)}| < 1 \) for every \( i \) and \( a \), we have \( \hat{v}_a = \sum_i w_i^{(a)} v_i < \sum_i v_i = Ns \). The domain of the variables appearing in \( \text{S17} \) is therefore:

\[
0 \leq s < \infty \\
-Ns \leq \hat{v}_a \leq Ns \quad \text{for } a \neq 0
\]

We then get

\[
P(s) = \frac{1}{Z_s} \exp \left[ -N\beta g(s - v_0)^2 \right] J_{N\beta g} \Omega(s, \{ \hat{v}_a \}) \exp \left[ -\beta \sum_{a=1}^{N} (J\lambda_a + g) \hat{v}_a^2 \right]
\]

\[
= \frac{1}{Z_s} s^{d-1} \exp \left[ -N\beta g(s - v_0)^2 \right] \int_{-Ns}^{Ns} \text{Do}^\prime \left( 1 + \sum_{a \neq 0} \frac{w_i^{(a)} \hat{v}_a}{s} \right) \exp \left[ -\beta \sum_{a=1}^{N} (J\lambda_a + g) \hat{v}_a^2 \right]
\]

where \( Z_s \) is the normalization of the distribution and the integral in \( \text{Do}^\prime \) is over all the non-zero modes. We omitted all the irrelevant constants that cancel out through simplification between the distribution and its normalization. In the approximation where the relative fluctuations of the individual speeds are small, we can expand the function \( h \) appearing in the above expression and compute the remaining Gaussian integral for large values of \( N \). We obtain, at leading order

\[
P(s) = \frac{1}{Z_s} s^{d-1} \exp \left[ -\frac{Ng}{T} (s - v_0)^2 \right]
\]

We stress that the above approximation is quite reasonable in the deeply ordered phase. The quantity \( \delta v_i = \sum_{a \neq 0} w_i^{(a)} \hat{v}_a \) indeed represents the fluctuation of the individual speed with respect to the mean speed of the group, \( s \), and it must not be confused with the fluctuations of the mean speed itself. At low noise, when mutual adaptation is strong, individuals efficiently coordinate both their directions and speeds so that we expect individual deviations from the group mean flight direction (the polarization), and the mean speed to be small (as confirmed by simulations, see Fig[2]). On the other hand, if the value of \( g \) is small, i.e. the control on the individual speeds is loose, the \( \{ v_i \} \) can remain coordinated and at the same time wildly fluctuate (e.g. everyone speeds up), giving rise to large fluctuations of \( s \), while keeping the relative deviations \( \delta v_i \) small.

The average value of the mean speed computed from distribution \( \text{S21} \) has been plotted in Fig.3b of the main paper: it predicts very nicely the values measured through numerical simulations of the off-lattice linear control model (see next section for details), confirming the validity of the approximations performed in the calculation (i.e. large directional order, quasi-equilibrium, small relative fluctuations of the speed). To get an analytical estimate of the typical speed, we can compute the maximum of the distribution. By imposing \( \frac{\partial P}{\partial s} = 0 \) for \( d = 3 \), we obtain the following equation:

\[
s^2_{\text{typical}} - s_{\text{typical}} v_0 - \frac{T}{Ng} = 0
\]

that gives us the expression for the maximum:

\[
s_{\text{typical}} = v_0 \left[ \frac{1}{2} + \frac{1}{2} \sqrt{1 + \frac{4T}{Ng v_0^2}} \right]
\]

This result confirms the idea that the mean speed is substantially different from \( v_0 \) for small \( N \), if \( g \) is too small, as clearly shown in Fig.3b of the main paper.

We wish to draw the reader’s attention on the fact that, despite the approximations we used to derive them (in particular the fixed network assumption), the analytical results of this section are in perfect agreement with numerical simulations performed by using an actual self-propelled particle model (see Fig.3b in the main text). This is not surprising, considering that in the deeply ordered flocking phase the time scale to reschedule the interaction network is much larger than the time of local relaxation [61].
Relative fluctuations of the speed. We report in this plot the relative fluctuations of the individual speed $\Delta s/s$ as a function of the mean speed, computed from numerical simulations for different values of $N$, and for $g = 10^{-3}$. The fluctuation is defined as $\Delta s = \left[ \frac{1}{N} \sum_i \delta v_i^2 \right]^{1/2}$. Each point in the plot corresponds to a distinct configuration, and all points of the same color are drawn from the same simulation performed at a given value of $N$. The black line is a fit of the data with a $f(x) = a/x$ function. The vertical dashed line corresponds to $s = v_0$, which is the asymptotic value for the mean speed in the thermodynamic limit. The fluctuations themselves are small and depend on $s$ only very weakly (inset), so that the relative fluctuations decay as $1/s$. Relative fluctuations therefore only increase due to the decrease of the average value of $s$ at large sizes. However, such value is limited by below ($s < v_0$) and the relative fluctuations therefore remain small in the whole range of parameters.

Distribution of the mean speed: marginal speed control

Let us now consider the marginal speed control model, that has the pseudo-Hamiltonian:

$$H(\{v_i\}) = \frac{J}{2} \sum_{i,j} n_{ij}(v_i - v_j)^2 + \frac{\lambda}{v_0^6} \sum_i (v_i^2 - v_0^2)^4$$

We can follow a similar procedure as the one used for linear control, i.e. we apply the SW approximation to deal with directional fluctuations and we decompose in normal modes for the speed fluctuations. We end up with a distribution with the same structure as the one of (S17) with

$$H = \sum_{a=1}^{N} J \lambda_a \hat{v}_a^2 + \frac{\lambda}{v_0^6} \sum_i \left( \sum_{a,b} w_i^{(a)} w_i^{(b)} \hat{v}_a \hat{v}_b - v_0^2 \right)^4$$

Integration over the non-zero modes with this effective Hamiltonian is clearly a hard task, due to the non-Gaussian contributions. However, in the approximation where the relative speed fluctuations are small things simplify: we can easily extract the zero mode contribution $\simeq N \frac{\lambda}{v_0^6} (s^2 - v_0^2)^4$ in the exponent, while at leading order the integration over the remaining modes (which is non Gaussian in this case) will produce a constant integral. The distribution for the average speed $s$ will then be:

$$P(s) = \frac{1}{Z s^{d-1} \exp \left[ -N \frac{\lambda}{v_0^6} (s^2 - v_0^2)^4 \right]}$$

The agreement between theory and simulations is less accurate than in the linear speed control case, but we still have a satisfying match between the predicted average mean speed and the value measured from numerical simulations (Fig.3d of the main paper). Once again we can compute the maximum of the distribution to estimate the typical mean speed. For $d = 3$ we get:

$$1 - \frac{4N\lambda v_0^2}{T} \left( \frac{s_{\text{typical}}}{v_0} \right)^2 \left( \frac{s_{\text{typical}}}{v_0} - 1 \right)^3 = 0$$
Since we are interested in the behavior of $s_{typical}$ in $N$ at fixed $T$ and $\lambda$, we can solve this equation in the two limits of big $N$ and small $N$, obtaining:

$$s_{typical} \simeq \begin{cases} v_0 \left[ 1 + \left( \frac{T}{32 N \lambda v_0^2} \right)^{1/3} \right] & \text{for } N \gg \frac{T}{\lambda v_0^2} \\ v_0 \left( \frac{T}{4 N \lambda v_0^2} \right)^{1/8} & \text{for } N \ll \frac{T}{\lambda v_0^2} \end{cases} \quad (S28)$$

**Numerical Simulations**

To investigate the flocking dynamics described by equations (1) and (2) of the main text, we perform numerical simulations with a system of self-propelled particles. The flock is modeled as a set of particles moving in a three-dimensional space with update rules for positions and velocities, which are a discretized version of Eqs. (1)(2). Following a simple Euler integration scheme [63], we get

$$v_i(t + \Delta t) = v_i(t) + \Delta t F_i + \delta \eta_i \quad (S29)$$

$$r_i(t + \Delta t) = r_i(t) + \Delta t v_i(t) \quad (S30)$$

Here the force $F_i = F_{int} + F_{sc}$ acting on particle $i$ contains both an alignment term $F_{int}$,

$$F_{int} = -J \sum_{j} n_{ij}(t) (v_i(t) - v_j(t)) \quad (S31)$$

and a speed control term $F_{sc}$, which can be either linear (by taking the derivative of the Hamiltonian $[S4]$):

$$F_{sc} = 2g \frac{v_i}{|v_i|} (v_0 - |v_i|) \quad (S32)$$

or marginal (by taking the derivative of the Hamiltonian $[S24]$):

$$F_{sc} = \frac{8\lambda}{v_0} v_i (v_0^2 - v_i^2)^3 \quad (S33)$$

The last term in $[S29]$ is a white gaussian noise with zero mean and variance:

$$\sigma_{\eta}^2 = 2d T \Delta t \quad (S34)$$

where $d = 3$ is the space dimension and $T$ is the effective temperature. The matrix $n_{ij}(t)$ is the adjacency matrix that defines which pairs interact; its entries can assume only the values 0 and 1, according to a rule of interaction that can be metric (i.e. $n_{ij} \neq 0$ if and only if $r_{ij} < r_c$) or topological (i.e. $n_{ij} \neq 0$ if $j$ is one of $i$’s first $n_c$ neighbors) [40] [41].

When working at fixed average density and in the very low temperature region where density fluctuations are small, there is not great difference between metric and topological interaction. Even though natural flocks are known to have analogous to the one of real flocks (to reproduce the behavior of the mean speed as
a function of $N$), and we also need to span a range in linear sizes $L$ of the same order of what found in experiments (to reproduce the scaling of the correlation length). However, natural flocks have non-trivial aspect ratios (being flatter in the direction of gravity [64]), which implies that their spatial extension $L_{\text{bio}}$ (defined as the maximum distance between two individuals) does not scale with the group size $N_{\text{bio}}$ as $L_{\text{bio}} = (N_{\text{bio}})^{1/3}$. For this reason, we need to perform two sets of simulations to satisfy the criteria mentioned above. In the first set, we perform simulations with $N = 8 \div 2744$, that are the minimum and maximum number of individuals in the recorded natural flocks. With these numerical data we compute the mean speed distribution and its average, that are then displayed in Fig.3b,d. Then, we perform a second set of simulations with $L/r_1$ up to the maximum experimental value of $L_{\text{bio}}/r_{1\text{bio}} = 70$, where $r_{1\text{bio}}$ is the average nearest neighbor distance of real flocks. With these data we compute the correlation lengths that are plotted in Fig2.a,c; for this set of simulations we have $N = 125 \div 343 \times 10^3$. An alternative strategy would have been to perform simulations in a box reproducing the aspect ratio of natural flocks, rather than in a cubic box. This is however not convenient. The aspect ratio for real flocks is not a stable quantity, but fluctuates from flock to flock and - for the same flock - from time to time. This would require an extremely painful calibration of the chosen simulation box. On the other hand, analysis of real data show that the aspect ratio does not influence the statistics of speeds and correlations (e.g., correlations of flocks with different aspect ratios rescale very well, safe for boundary effects at very large distances [17]). Choosing a cubic a box is therefore perfectly legitimate in terms of the physics, and it very much simplifies the data analysis.

We choose the value of the reference speed of the particles $v_0$ and of the integration step $\Delta t$ to ensure an average displacement $\Delta r \simeq v_0 \Delta t$ much smaller than the size of the box $L$. In this way there is a weak rewiring of the interaction network during the time of simulation, consistently with the quasi-equilibrium condition of natural flocks [61]. The step of integration is selected as the maximum value granting a robust numerical integration in terms of errors and stationarity of the energy of the system (absence of trends in time or in size). In simulations with marginal speed control this algorithmic stability is achieved with $\Delta t = 0.01$, while linear speed control requires a $\Delta t = 0.001$. Every simulation consists in a run of length $N_{\text{steps}} = 2 \times 10^4$ steps for thermalization and an independent run long $N_{\text{steps}} = 1.2 \times 10^6$ steps. From the latter we extract configurations every 1000 steps in order to compute the quantities needed by our analysis.

In equilibrium ferromagnetic models in their low temperature phase, the polarization $\Phi$ only depends on the ratio between ferromagnetic coupling $J$ and temperature $T$ through the relation [62],

$$\Phi = 1 - \alpha \frac{T}{J} \tag{S35}$$

where $\alpha$ is a constant of order 1 whose value depends on the specific structure of the interaction network. This relation, though, is only valid when the vectorial degrees of freedom $v_i$ have modulus 1, whereas if they have modulus equal to $v_0$ the relation changes to,

$$\Phi = 1 - \alpha \frac{T}{v_0^2 J} \tag{S36}$$

For out-of-equilibrium models, as the present case, the polarization depends in principle on all parameters; however, in the deeply ordered phase that we are considering here the main contribution to $\Phi$ is still given by the ratio between alignment and noise, so that [S36] remains a very useful rule of thumb to fix the parameters of the model such to have a polarization equal to that of natural flocks, namely $\Phi \simeq 0.89 \div 0.99$. In Table II we report the values of the other parameters used in the simulations.
| Acquisition | No. of birds N | Flock's size L, m | Polarization Φ | Mean speed s, m/s | Correlation length ξ, m |
|-------------|----------------|------------------|----------------|------------------|---------------------|
| 16-05       | 1548           | 68.1             | 0.961          | 15.5             | 9.1                 |
| 17-06       | 380            | 40.5             | 0.935          | 10.0             | 5.8                 |
| 21-06       | 530            | 26.6             | 0.973          | 11.0             | 3.8                 |
| 25-08       | 1079           | 52.5             | 0.962          | 12.7             | 7.1                 |
| 25-10       | 696            | 30.1             | 0.991          | 12.6             | 3.3                 |
| 25-11       | 854            | 33.1             | 0.957          | 10.7             | 3.4                 |
| 28-10       | 1122           | 32.3             | 0.982          | 11.2             | 3.2                 |
| 29-03       | 422            | 28.1             | 0.963          | 10.8             | 3.7                 |
| 31-01       | 1565           | 67.3             | 0.921          | 7.5              | 8.5                 |
| 32-06       | 690            | 18.4             | 0.981          | 10.0             | 2.6                 |
| 42-03       | 366            | 27.2             | 0.979          | 10.2             | 3.4                 |
| 48-17       | 709            | 25.7             | 0.886          | 13.5             | 3.0                 |
| 49-05       | 636            | 15.1             | 0.995          | 13.7             | 2.0                 |
| 54-08       | 2548           | 66.6             | 0.971          | 14.2             | 8.9                 |
| 57-03       | 2559           | 76.1             | 0.978          | 14.3             | 10.7                |
| 58-06       | 351            | 19.4             | 0.987          | 10.8             | 2.2                 |
| 58-07       | 445            | 15.3             | 0.977          | 10.9             | 2.4                 |
| 63-05       | 712            | 47.2             | 0.978          | 10.3             | 4.1                 |
| 69-09       | 206            | 13.4             | 0.985          | 11.8             | 1.8                 |
| 69-10       | 994            | 32.4             | 0.987          | 12.0             | 4.1                 |
| 69-13       | 1238           | 39.2             | 0.937          | 10.1             | 5.9                 |
| 69-19       | 617            | 21.0             | 0.975          | 14.3             | 3.6                 |
| 72-02       | 101            | 7.2              | 0.993          | 13.3             | 1.5                 |
| 77-07       | 131            | 6.5              | 0.978          | 9.2              | 1.5                 |

**TABLE I: Experimental data.** This table reports all the data required to perform the analysis presented in this paper. Each line corresponds to a different acquisition (i.e. flocking recording), for all the three experimental campaigns considered (acquisitions labeling system changed from one campaign to another). Acquisitions belonging to different campaigns are separated by a straight line. For each acquisition we have: the median number of individuals $N$, the median flock’s size $L$, the mean polarization $\Phi = 1/N \sum_i v_i/v_i$, the median of the mean speed $s = 1/N \sum_i v_i$ and the median correlation length $\xi$, computed via eq. (S3). Every median (or mean), relative to a particular acquisition, is made over all the frames in that recording. Since the measured polarization value depends on time resolution (higher resolution bringing more noise), acquisition of the second and third campaign (that are acquired at much faster rates) have been re-sampled at the same rate of the first campaign so as to have homogeneous measurements for all the data.
| Speed control | $g$     | $\lambda$ | $T$           | $J$ |
|---------------|---------|------------|---------------|-----|
| Linear        | 0.001   | -          | $2.5 \times 10^{-3}$ | 10  |
|               | 0.03    | -          | $2.5 \times 10^{-3}$ | 10  |
|               | 0.1     | -          | $2.5 \times 10^{-3}$ | 10  |
|               | 1.0     | -          | $2.5 \times 10^{-3}$ | 10  |
| Marginal      | -       | 0.001      | $1.25 \times 10^{-4}$ | 1.0 |

**TABLE II: Parameters of simulations.** In this table we report the values of relevant parameters used in the numerical simulations. The other parameters are $r_c = 1.2$, $v_0 = 0.05$, $\Delta t_{MRG} = 0.01$, $\Delta t_{GAUSS} = 0.001$. 