# Universal power law governing pedestrian interactions

Ioannis Karamouzas, <sup>1</sup> Brian Skinner, <sup>2</sup> and Stephen J. Guy <sup>1</sup>

<sup>1</sup>Department of Computer Science and Engineering, University of Minnesota, USA

<sup>2</sup>Materials Science Division, Argonne National Laboratory, USA

(Dated: December 4, 2014)

Human crowds often bear a striking resemblance to interacting particle systems, and this has prompted many researchers to describe pedestrian dynamics in terms of interaction forces and potential energies. The correct quantitative form of this interaction, however, has remained an open question. Here, we introduce a novel statistical-mechanical approach to directly measure the interaction energy between pedestrians. This analysis, when applied to a large collection of human motion data, reveals a simple power law interaction that is based not on the physical separation between pedestrians but on their projected time to a potential future collision, and is therefore fundamentally anticipatory in nature. Remarkably, this simple law is able to describe human interactions across a wide variety of situations, speeds and densities. We further show, through simulations, that the interaction law we identify is sufficient to reproduce many known crowd phenomena.

In terms of its large-scale behaviors, a crowd of pedestrians can look strikingly similar to many other collections of repulsively-interacting particles [1–4]. These similarities have inspired a variety of pedestrian crowd models, including cellular automata and continuum-based approaches [5–8], as well as simple particle or agent-based models [9–15]. Many of these models conform to a long-standing hypothesis that humans in a crowd interact with their neighbors through some form of "social potential" [16], analogous to the repulsive potential energies between physical particles. How to best determine the quantitative form of this interaction potential, however, has remained an open question, with most previous researchers employing a simulation-driven approach.

Previously, direct measurement of the interaction law between pedestrians has been confounded by two primary factors. First, each individual in a crowd experiences a complex environment of competing forces, making it difficult to isolate and robustly quantify a single pairwise interaction. Secondly, a pedestrian's motion is strongly influenced not just by the present position of neighboring pedestrians, but by their anticipated future positions [17– 21], a fact which has influenced recent models [22–25]. Consider, for example, two well-separated pedestrians walking into a head-on collision (Fig. 1a). These pedestrians typically exhibit relatively large acceleration as they move to avoid each other, as would result from a large repulsive force. On the other hand, pedestrians walking in parallel directions exhibit almost no acceleration, even when their mutual separation is small (Fig. 1b).

Here, we address both of the aforementioned factors using a data-driven, statistical mechanics-based analysis that accounts properly for the anticipatory nature of human interactions. This approach allows us to directly and robustly measure the interaction energy between pedestrians. The consistency of our measurements across a variety of settings suggests a simple and universal law governing pedestrian motion.

To perform our analysis, we turn to the large col-

lections of recently published crowd datasets recorded by motion-capture or computer vision-based techniques. These datasets include pedestrian trajectories from several outdoor environments [26, 27] and controlled lab settings [28] (a summary of datasets is given in the Supplemental Material [29]). To reduce statistical noise, datasets with similar densities were combined together, resulting in one Outdoor dataset comprising 1,146 trajectories of pedestrians in sparse-to-moderate outdoor settings, and one Bottleneck dataset with 354 trajectories of pedestrians in dense crowds passing through narrow bottlenecks. In analyzing these datasets, our primary tool for quantifying the strength of interactions between pedestrians is the statistical-mechanical pair distribution function, denoted g.

As in the typical condensed matter setting [30], here we define the pair distribution function g(x) as the observed probability density for two pedestrians to have relative separation x divided by the expected probability density for two non-interacting pedestrians to have the same separation. In general, the probability density for non-interacting pedestrians cannot be known  $a\ priori$ , since it depends on where and how frequently pedestrians enter and exit the environment. However, for large datasets we are able to closely approximate this distribution by sampling the separation between all pairs of pedestrians that are not simultaneously present in the scene (and therefore not interacting). As defined above, small values of the pair distribution function,  $g(x) \ll 1$ , correspond to situations where interactions produce strong avoidance.

If the Cartesian distance r between two pedestrians were a sufficient descriptor of their interaction, we would expect the shape of the pair distribution function g(r) to be independent of all other variables. However, as can be seen in Fig. 1c, g(r) has large, qualitative differences when the data is binned by the rate at which the two pedestrians are approaching each other, v = -dr/dt. In particular, pedestrians with a small rate of approach are more likely to be found close together than those that are

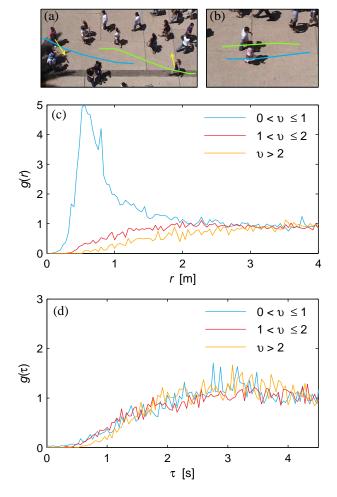


FIG. 1. Analysis of anticipation effects in pedestrian motion. (a) Two pedestrians react strongly to avoid an upcoming collision even though they are far from each other (path segments over an interval of 4s are shown as colored lines, with arrows indicating acceleration). (b) In the same environment, two pedestrians walk close to each other without any relative acceleration. (c) The pair distribution function g as a function of inter-pedestrian separation r shows very different behavior when plotted for pedestrian pairs with different rate of approach v = -dr/dt. Units of v are m/s. (d) In contrast, when g is computed as a function of time-to-collision,  $\tau$ , curves corresponding to different v collapse onto each other.

approaching each other quickly (as evidenced by the separation between the curves at small r). A particularly pronounced difference can be seen for the curve corresponding to small v, where the large peak suggests a tendency for pedestrians with similar velocities to walk closely together.

While the distance r is not a sufficient descriptor of interactions, we find that the pair distribution function can, in fact, be accurately parameterized by a single variable that describes how imminent potentially upcoming collisions are. We refer to this variable as the time-to-collision, denoted  $\tau$ , which we define as the duration of

time for which two pedestrians could continue walking at their current velocities before colliding. As shown in Fig. 1d, when the pair distribution function is plotted as a function of  $\tau$ , curves for different rates of approach collapse onto each other, with no evidence of a separate dependence of the interaction on v. Even when binned by other parameters such as the relative orientation between pedestrians, there is no significant difference between curves (see the Supplemental Material [29]). This consistent collapse of the curves suggests that the single variable  $\tau$  provides an appropriate description of the interaction between pedestrians.

This pair distribution function,  $g(\tau)$ , describes the extent to which different configurations of pedestrians are made unlikely by the mutual interaction between pedestrians. In general, situations with strong interactions (small  $\tau$ ) are suppressed statistically, since the mutual repulsion between two approaching pedestrians makes it very unlikely that the pedestrians will arrive at a situation where a collision is imminent. This suppression can be described in terms of a pedestrian "interaction energy"  $E(\tau)$ . In particular, in situations where the average density of pedestrians does not vary strongly with time, the probability of a pair of pedestrians having time-to-collision  $\tau$  can be assumed to follow a Boltzmann-like relation,  $g(\tau) \propto \exp[-E(\tau)/E_0]$ . Here,  $E_0$  is a characteristic pedestrian energy, whose value is scene-dependent.

This use of a Boltzmann-like relation between  $g(\tau)$  and  $E(\tau)$  amounts to an assumption that the systems being considered are at, or near, statistical equilibrium. In our analysis, this assumption is motivated by the observation that the intensive properties of the system in each of the datasets (e.g., the average pedestrian density and walking speed) are essentially time-independent. If this time-independence is taken as given, a Boltzmann-like relation follows as a consequence of entropy maximization. By rearranging this relation, the interaction energy can be expressed in terms of  $g(\tau)$  as:

$$E(\tau) \propto \ln\left[1/g(\tau)\right].$$
 (1)

A further, self-consistent validation of Eq. (1) is provided below.

Figure 2 plots the interaction law defined by Eq. (1) using the values for  $g(\tau)$  derived from our two aggregated pedestrian datasets. It is worth emphasizing that these two datasets capture very different types of pedestrian motion. The pedestrian trajectories in the *Outdoor* dataset are generally multi-directional paths in sparse-to-moderate densities, with pedestrians often walking in groups or stopping for brief conversations. In contrast, trajectories in the *Bottleneck* dataset are largely unidirectional, with uniformly high density, and with little stopping or grouping between individuals.

Remarkably, despite their large qualitative differences, both datasets reveal the same power-law relationship underlying pedestrian interactions. For both datasets, the

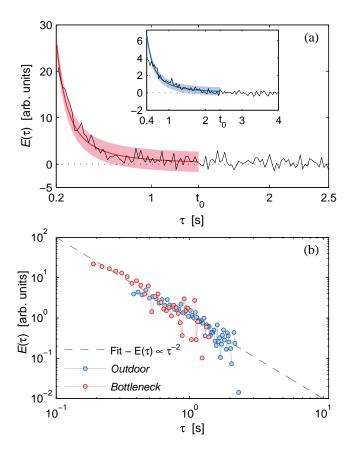


FIG. 2. (a) The interaction energy computed from the dense Bottleneck dataset and from the more sparse Outdoor dataset (inset). The overall constant k is normalized so that E(1) = 1. Both datasets fit well to a power law up to a point marked  $t_0$ , beyond which there is no discernible interaction. Solid lines shows the fit to the data and colored regions show their corresponding 95% confidence interval (Bottleneck,  $R^2 = 0.94$ ; Outdoor,  $R^2 = 0.92$ ). (b) The interaction energy in both datasets is well described by a power law with exponent 2.

interaction energy E shows a quadratic falloff as a function of  $\tau$ , so that  $E(\tau) \propto 1/\tau^2$  over the interval where E is well-defined. For smaller values of  $\tau$  (less than  $\sim 200$  ms), the energy seen in the data saturates to a maximum value, likely as a consequence of finite human reaction times. For sufficiently large values of  $\tau$ , on the other hand, the observed interaction energy quickly vanishes, suggesting a truncation of the interaction when the time-to-collision is large. We denote the maximum observed interaction range as  $t_0$  (Bottleneck:  $t_0 \approx 1.4 \, \mathrm{s}$ ; Outdoor:  $t_0 \approx 2.4 \, \mathrm{s}$ ).

Importantly,  $t_0$  does not, by itself, indicate the intrinsic interaction range between pedestrians, since interactions between distant, non-neighboring pedestrians are screened by the presence of nearest-neighbors, as in other dense, interacting systems [30, 31]. For a crowd with density  $\rho$ , the characteristic "screening time" can be expected to scale as the typical distance between nearest neighbors,  $\rho^{-1/2}$ , divided by the mean walk-

ing speed u. Such scaling is indeed consistent with the trend observed in our data, with the denser Bottleneck dataset ( $\rho = 2.5 \,\mathrm{m}^{-2}$ ,  $u = 0.55 \,\mathrm{m/s}$ ) demonstrating a smaller value of  $t_0$  than the sparser Outdoor dataset ( $\rho = 0.27 \,\mathrm{m}^{-2}$ ,  $u = 0.86 \,\mathrm{m/s}$ ) [32]. While the large- $\tau$  behavior in our datasets is therefore dominated by screening, we can use the largest observed values of  $t_0$  to place a lower bound estimate on the intrinsic range of unscreened interactions (which we denote as  $\tau_0$ ). This estimate suggests that an appropriate value is  $\tau_0 \approx 3 \,\mathrm{s}$ , which is consistent with previous research demonstrating an interaction time horizon of  $2-4 \,\mathrm{s}$  [21].

Since the interaction energy follows a power law with a sharp truncation at large  $\tau$ , we infer from the data the following form of the pedestrian interaction law:

$$E(\tau) = \frac{k}{\tau^2} e^{-\tau/\tau_0}.$$
 (2)

Here, k is a constant that sets the units for energy.

To demonstrate the general nature of the identified interaction law, we performed simulations of pedestrians that adapt their behavior according to Eq. (2) via force-based interactions. In particular, the energy  $E(\tau)$  directly implies a natural definition of the force  ${\bf F}$  experienced by pedestrians when interacting:

$$\mathbf{F} = -\nabla_{\mathbf{r}} \left( \frac{k}{\tau^2} e^{-\tau/\tau_0} \right), \tag{3}$$

where  $\nabla_{\mathbf{r}}$  is the spatial gradient. A full analytical expression for this derivative is given in the Supplemental Material [29].

For the purposes of simulation, each pedestrian is also given a driving force associated with its desired direction of motion, following Ref. 9. The resulting force model is sufficient to reproduce a wide variety of important pedestrian behaviors, including the formation of lanes, arching in narrow passages, slowdowns in congestion, and anticipatory collision avoidance (Fig. 3). Additionally, the simulated pedestrians match the known fundamental diagram [33] of speed-density relationships for real human crowds and qualitatively capture the empirical behavior of g(r) depicted in Fig. 1c [29].

Our simulations also reproduce the anticipatory power law described by Eq. (2), as shown in Fig. 4. In contrast, simulations generated by distance-based interaction forces fail to show a dependence of E on  $\tau$  (Fig. 4). Other, more recent models of pedestrian behavior also cannot consistently capture the empirical power-law relationship (see the Supplemental Material [29]). The ability of our own simulations to reproduce  $E(\tau)$  also provides a self-consistent validation of our use of the Boltzmann relation to infer the interaction energy from data.

Interesting behavior can also be seen when Eq. (3) is applied to walkers propelled forward in the direction of their current velocity without having a specific goal (as

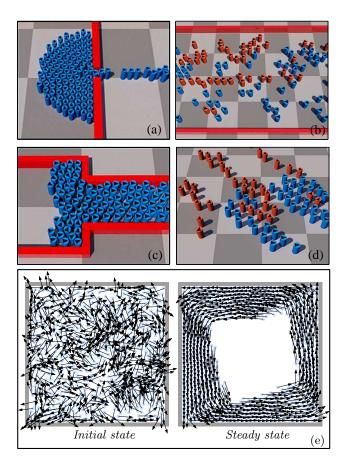


FIG. 3. Stills from simulations of agents following the force law derived from Eq. (2). In figures (a)-(d), agents are represented as cylinders and color-coded according to their goal direction. The simulated agents display many emergent phenomena also seen in human crowds, including arching around narrow passages (a), clogging and "zipping" patterns at bottlenecks (c), and spontaneously self-organized lane formation (b and d). Figure (e) depicts a simulation of agents without a preferred goal direction (arrows represent the agents' current orientations). The agents' interactions lead to large-scale synchronization of their motion. Further simulation details are given in the Supplemental Material [29].

implemented, for example, in Ref. 34). In such cases, complex spatio-temporal patterns emerge, leading eventually to large scale synchronization of motion. An example is illustrated in Fig. 3e, where a collection of pedestrians that is initialized to a high energy state with many imminent collisions settles over time into a low energy state where pedestrians move in unison. This result is qualitatively similar to observed behavior in dense, nongoal-oriented human crowds [35], and is reminiscent of the "flocking" behavior seen in a variety of animal groups [36–40]. A detailed study of such collective behaviors, however, is outside the scope of our present work.

While the model implied by Eq. (3) is widely applicable, it may not be sufficient on its own to capture certain crowd phenomena, such as the shock waves and turbulent

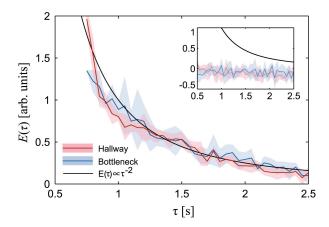


FIG. 4. Inferred interaction energy  $E \propto \ln(1/g)$  as a function of time-to-collision  $\tau$  for different simulations, obtained using the anticipatory force described by Eq. (3), and the distance-dependent force described in Ref. 9 (inset). For simulations with strictly distance-dependent interactions, the inferred interaction energy does not show a dependence on  $\tau$ . In contrast, simulations following our model closely match the observed empirical power law for  $E(\tau)$ . Shaded regions denote average energy values  $\pm$  one standard deviation.

flows that have been reported to occur in extremely high density crowds [41]. In such very dense situations, saturating effects such as finite human reaction time become relevant, and these may alter the quantitative form of the interaction in a way that is not well-captured by our time-to-collision-based analysis. Augmenting our result with an additional close-ranged component of the interaction may give a better description of these extremely dense scenarios, and is a promising avenue for future work.

To conclude, our statistical mechanics-based analysis of a large collection of human data has allowed us to quantify the nature and strength of interactions between pedestrians. This novel type of analysis opens new avenues for studying the behavior of humans using real life data. The data we have analyzed here reveals the existence of a single anticipatory power law governing the motions of humans. The consistency of this law across a variety of scenarios provides a new means to understand how pedestrians behave and suggests new ways to evaluate models of pedestrian interactions. Further, these results suggest a general quantitative law for describing human anticipation that may extend to other studies of human behavior, which may therefore be amenable to a similar type of analysis.

Complete simulation source code, along with videos and links to data used in this study, can be found at our companion webpage: <a href="http://motion.cs.umn.edu/PowerLaw">http://motion.cs.umn.edu/PowerLaw</a>. We would like to thank Anne-Hélène Olivier, Alex Kamenev, Julien Pettré, Igor Aranson, Dinesh Manocha, and Leo Kadanoff for helpful discussions. We also acknowledge support from Intel and from Uni-

versity of Minnesota's MnDRIVE Initiative on Robotics, Sensors, and Advanced Manufacturing. Work at Argonne National Laboratory is supported by the U.S. Department of Energy, Office of Basic Energy Sciences under contract no. DE-AC02-06CH11357.

- D. Helbing, P. Molnár, I. Farkas, and K. Bolay, Env. Plan. B 28, 361 (2001).
- [2] M. Moussaïd, E. G. Guillot, M. Moreau, J. Fehrenbach, O. Chabiron, S. Lemercier, J. Pettré, C. Appert-Rolland, P. Degond, and G. Theraulaz, PLoS Comput Biol. 8, e1002442 (2012).
- [3] W. Yu and A. Johansson, Phys. Rev. E 76, 046105 (2007).
- [4] D. Helbing, Rev. Mod. Phys. **73**, 1067 (2001).
- [5] L. P. Kadanoff, J. Stat. Phys. 39, 267 (1985).
- [6] S. Hoogendoorn and P. H. L. Bovy, Transport. Res. Rec. 1710, 28 (2000).
- [7] R. Hughes, Transp. Res. Part B **36**, 507 (2002).
- [8] A. Kirchner, K. Nishinari, and A. Schadschneider, Phys. Rev. E 67, 056122 (2003).
- [9] D. Helbing, I. Farkas, and T. Vicsek, Nature 407, 487 (2000).
- [10] W. J. Yu, R. Chen, L. Y. Dong, and S. Q. Dai, Phys. Rev. E 72, 026112 (2005).
- [11] M. Chraibi, A. Seyfried, and A. Schadschneider, Phys. Rev. E 82, 046111 (2010).
- [12] S. Hoogendoorn and P. H. L. Bovy, Optim. Control Appl. Methods 24, 153 (2003).
- [13] M. Moussaïd, D. Helbing, S. Garnier, A. Johansson, M. Combe, and G. Theraulaz, Proc. R. Soc. B 276, 2755 (2009).
- [14] B. R. Fajen and W. H. Warren, J. Exp. Psychol. Hum. Percept. Perform. 29, 343 (2003).
- [15] C. W. Reynolds, ACM Comput. Graph. 21, 25 (1987).
- [16] D. Helbing and P. Molnár, Phys. Rev. E 51, 4282 (1995).
- [17] M. Gérin-Lajoie, C. Richards, and B. McFadyen, Motor control 9, 242 (2005).
- [18] T. Ducourant, S. Vieilledent, Y. Kerlirzin, and A. Berthoz, Neurosci. Lett. 389, 6 (2005).
- [19] A. Johansson, Phys. Rev. E 80, 026120 (2009).
- [20] L. Noy, E. Dekel, and U. Alon, Proc. Natl. Acad. Sci. 108, 20947 (2011).
- [21] A.-H. Olivier, A. Marin, A. Crétual, and J. Pettré, Gait Posture 36, 399 (2012).
- [22] J. Ondřej, J. Pettré, A.-H. Olivier, and S. Donikian, ACM Trans. Gr. 29, 1 (2010).
- [23] F. Zanlungo, T. Ikeda, and T. Kanda, EPL (Europhysics Letters) 93, 68005 (2011).
- [24] M. Moussaïd, D. Helbing, and G. Theraulaz, Proc. Natl. Acad. Sci. 108, 6884 (2011).
- [25] S. J. Guy, S. Curtis, M. C. Lin, and D. Manocha, Phys.

- Rev. E 85, 016110 (2012).
- [26] A. Lerner, Y. Chrysanthou, and D. Lischinski, Computer Graphics Forum 26, 655 (2007).
- [27] S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool, in *IEEE International Conference on Computer Vision* (2009) pp. 261–268.
- [28] A. Seyfried, O. Passon, B. Steffen, M. Boltes, T. Rupprecht, and W. Klingsch, Transp. Sci. 43, 395 (2009).
- [29] See the Supplemental Material for experimental datasets, statistical methods, a simulation model derived from our law, and simulation results. The Supplemental Material includes Refs. [42–46].
- [30] G. D. Mahan, Many-Particle Physics (Plenum, New York, 1990).
- [31] M. Babadi, B. Skinner, M. M. Fogler, and E. Demler, Europhys. Lett. 103, 16002 (2013).
- [32] The ratio of  $(\rho^{-1/2})/u$  between the *Bottleneck* and *Outdoor* datasets is  $(1.15 \, \text{s})/(2.25 \, \text{s}) = 0.51$ . For comparison, the measured values of  $t_0$  have a ratio  $(1.4 \, \text{s})/(2.4 \, \text{s}) = 0.58$ .
- [33] U. Weidmann, *Transporttechnik der Fuβgänger*, Literature Research 90 (ETH Zürich, 1993) in German.
- [34] D. Grossman, I. S. Aranson, and E. B. Jacob, New J. Phys. 10, 023036 (2008).
- [35] J. L. Silverberg, M. Bierbaum, J. P. Sethna, and I. Cohen, Phys. Rev. Lett. 110, 228701 (2013).
- [36] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Phys. Rev. Lett. 75, 1226 (1995).
- [37] I. D. Couzin, J. Krause, R. James, G. D. Ruxton, and N. R. Franks, J. Theor. Biol. 218, 1 (2002).
- [38] 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, Proc. Natl. Acad. Sci. 105, 1232 (2008).
- [39] I. L. Bajec and F. H. Heppner, Anim. Behav. 78, 777 (2009).
- [40] T. Vicsek and A. Zafeiris, Phys. Rep. **517**, 71 (2012).
- [41] D. Helbing, A. Johansson, and H. Z. Al-Abideen, Phys. Rev. E 75, 046109 (2007).
- [42] A. Seyfried, M. Boltes, J. Kähler, W. Klingsch, A. Portz, T. Rupprecht, A. Schadschneider, B. Steffen, and A. Winkens, in *Pedestrian and Evacuation Dynamics* 2008, edited by W. Klingsch, C. Rogsch, A. Schadschneider, and M. Schreckenberg (Springer Berlin Heidelberg, 2010) pp. 145–156.
- [43] M. Boltes, A. Seyfried, B. Steffen, and A. Schadschneider, in *Pedestrian and Evacuation Dynamics 2008*, edited by W. Klingsch, C. Rogsch, A. Schadschneider, and M. Schreckenberg (Springer Berlin Heidelberg, 2010) pp. 43–54.
- [44] L. C. Edie, in Proceedings of the Second International Symposium on the Theory of Traffic Flow: London 1963 (OECD, Paris, France, 1965) pp. 139–154.
- [45] S. P. Hoogendoorn and W. Daamen, Transp. Sci. 39, 147 (2005).
- [46] S. Hoogendoorn and W. Daamen, in *Traffic and Granular Flow '03* (Springer Berlin Heidelberg, 2005) pp. 373–382.

# Supplemental material for: Universal power law governing pedestrian interactions

Ioannis Karamouzas, <sup>1</sup> Brian Skinner, <sup>2</sup> and Stephen J. Guy <sup>1</sup>

<sup>1</sup>Department of Computer Science and Engineering, University of Minnesota, USA

<sup>2</sup>Materials Science Division, Argonne National Laboratory, USA

(Dated: December 4, 2014)

### EXPERIMENTAL DATASETS

Our findings draw from datasets published by three different research groups. These datasets are illustrated in Fig. S1 and summarized in Table S1. The experiments labeled b250\_combined and b4\_combined are described in Refs. 1 and 2, and are combined to comprise the Bottleneck dataset. These experimental trials were recorded at 25 fps using multiple cameras, with trajectories automatically tracked and corrected for perspective distortion [1, 3]. The experiments involved participants walking through a 4 m-long corridor that has a bottleneck of width of 2.5 m or 1 m. The remaining four datasets were combined to comprise the Outdoor dataset. The three datasets labeled crowds\_zara01, crowds\_zara02, and students003 were recorded at 25 fps using a single camera and the trajectories of the pedestrians were manually tracked and post-processed to minimize errors and correct the distortion due to pixel noise and the camera perspective [4]. The dataset denoted seq\_eth was recorded at 2.5 fps and tracked in a semi-automatic process [5]. In addition to their different settings, the two datasets types, Bottleneck and Outdoor, are also distinguished from each other by their different pedestrian densities (see Table S1), while within each type the densities are similar.

Matlab was used to process the corresponding 2D positional data of the pedestrians' trajectories after applying a low-pass second order Butterworth filter to remove noise and reduce oscillation effects (zero phase shift; 0.8 normalized cutoff frequency for the four outdoor datasets and 0.24 normalized cutoff frequency for the two bottleneck datasets). In all datasets and for each time instant, we infer the instantaneous velocity of each pedestrian using a discrete derivative. To estimate the time-to-collision  $\tau$ , which indicates when (if ever) a given pair of pedestrians will collide if they continue moving with their current velocities, we assume that each pedestrian can be modeled as a disc with a fixed radius. In our analysis, we used a radius of 0.1 m, which results in less than ten total colliding pairs of pedestrians across all six datasets. The corresponding disc diameter represents the smallest cross-sectional length that a pedestrian can occupy (by rotating or compressing his or her upper body) while resolving a collision. Overall, we collected 119,774 pairwise time-to-collision ( $\tau$ ) samples from the Outdoor datasets and 177,672 samples from the Bottleneck datasets. This number of samples was sufficient to draw statistically significant conclusions about the power law governing pedestrian interactions. While all results presented here correspond to an assumed radius of 0.1 m, a similar power law trend is produced for a wide range of radius values.

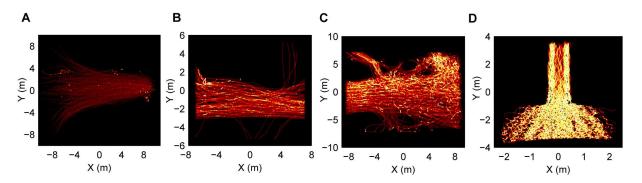


FIG. S1. Pedestrian trajectories for four of the six datasets examined in this paper, colored by time-averaged density from low (dark red) to high (white) values. (A-B) are from sparse, outdoor environments with largely bi-directional flows. (C) is from a moderately dense environment with multi-directional flow. (D) is from controlled experiments where a dense crowd walks through a narrow constriction. Horizontal and vertical axes label the distance from the center of the scene, in meters.

TABLE S1. Characteristics of the datasets analyzed. The first two listed datasets were grouped together to comprise the *Bottleneck* dataset. The remaining four datasets define the combined *Outdoor* dataset. The average density reported in the table is computed using the generalized definition of Edie [6].

Dataset	Type	Description	Flow	Location	No. Ped	Density $(1/m^2)$	Citation	Data files
b250_combined	Lab setting	Participants navigating through a 2.5 m wide bottleneck	Uni- directional	Jülich, Germany	176	2.328	[1, 2]	[7]
l4_ combined	Lab setting	Participants navigating through a 1.5 m wide bottleneck	Uni- directional	Jülich, Germany	178	2.665	[1, 2]	[7]
crowds_ zara01	Outdoor setting	Pedestrians interacting at a shopping street	Bi- directional	Nicosia, Cyprus	148	0.206	[4]	[8]
crowds_ zara02	Outdoor setting	Pedestrians interacting at a shopping street	Bi- directional	Nicosia, Cyprus	204	0.267	[4]	[8]
students003	Outdoor setting	Students interacting at a university campus	Multi- directional	Tel Aviv, Israel	434	0.391	[4]	[8]
seq_eth	Outdoor setting	Students interacting outside the ETH main building	Bi- directional	Zürich, Switzerland	360	0.148	[5]	[9]

## DETAILED DESCRIPTION OF THE PAIR-DISTRIBUTION FUNCTION

For a given variable x that describes the separation between two pedestrians, the pair distribution function g(x) indicates the degree to which a pair separation x is made unlikely by the interactions between pedestrians. Specifically, g(x) is defined by  $g(x) = P(x)/P_{NI}(x)$ , where P(x) is the probability density function for the relative separation x between pedestrians in the dataset, and  $P_{NI}(x)$  is the probability density function for x that would arise, hypothetically, if pedestrians were non-interacting. For large separation values we expect that pedestrians do not influence each other, and therefore  $\lim_{x\to\infty}g(x)=1$ .

While there is no way to remove the interactions between pedestrians in the dataset, we propose the following approach to closely approximate the distribution  $P_{NI}(x)$ . We begin by randomly permuting time information between different pedestrians, so that each moment in a pedestrian's trajectory has its instantaneous position and velocity preserved, but is assigned a randomly permuted time. The resulting "time-scrambled" dataset maintains the same spatially-averaged density at any given instant and the same time-averaged flow rate of pedestrians across any location in the scene as in the original dataset. However, the pedestrian positions in the time-scrambled dataset are uncorrelated with each other at any given value of the scrambled time, since they are drawn from trajectories at different real times. Creating a probability density function of the scrambled data gives  $P_{NI}(x)$ , allowing us to define  $g(x) = P(x)/P_{NI}(x)$ . Figure 1c of the main text shows the result of this process using for the variable x the Cartesian distance r between pedestrians, while Fig. 1d shows the result when the probability density functions are computed for the time-to-collision  $\tau$ .

## STATISTICAL METHODS

## Analysis of similarity between g(r) and $g(\tau)$ curves

We analyzed the effect that the relative velocity between pedestrians has on the distance- and time-to-collision-based pair distribution functions by conducting separate one-way ANOVAs for g(r) and  $g(\tau)$  respectively on the Outdoor

dataset. All tests were performed in STATISTICA (version 8.0) with significance levels set to 5%. To estimate g(r) and  $g(\tau)$ , we used intervals of 0.04 m and 0.04 s, respectively, and clustered pairs of pedestrians into three categories according to their rate of approach v = -dr/dt (Fig. 1c and Fig. 1d) for values of r < 8 m and  $\tau < 8$  s. The analysis reveals that the distance-based pair distribution function g(r) has a significant dependence on the rate of approach, [F(2,576)=27.811, P<0.001]. However, for  $g(\tau)$ , the pair distribution function does not vary for different values of v, [F(2,510)=0.143, P=0.866], indicating that  $\tau$  is a sufficient descriptor of pedestrian interactions.

# Power-law fit for $E(\tau)$

Regarding the power-law fit shown in Fig. 2 of the main text, for both the Outdoor and the Bottleneck datasets, we estimated  $g(\tau)$ , and subsequently the interaction energy  $E(\tau)$ , using intervals of 0.01 s. In both datasets, due to tracking errors and statistical noise, the energy is only well-defined over a finite interval, with the computed energy values fluctuating around a maximum observed energy at very small values of  $\tau$  and becoming indistinguishable from noise at large values of  $\tau$ . To estimate the lower  $\tau$  boundary, we first clustered the data into bins of 0.2 s, and used a series of t-tests between successive bins to determine the first two bins with significantly different interaction energies. In the Outdoor dataset, the analysis revealed a significant difference in  $E(\tau)$  between [0.2 s, 0.4 s) and [0.4 s, 0.6 s) indicating a value of  $\tau = 0.4$  s as an appropriate lower bound [t(38) = 6.664, P < 0.001]. In the Bottleneck dataset, the first two bins already exhibit a statistically significant difference in energy [t(38) = 10.856, P < 0.001], allowing us to select the value of 0.2 s as a lower boundary for  $\tau$ . To estimate the upper boundary value,  $t_0$ , we conducted separate ANOVA tests and determined the first three successive clusters for which the interaction energy does not vary. In the Outdoor dataset, [2.2 s, 2.4 s) denotes the first bin that has the same energy as its two subsequent ones, indicating  $t_0 = 2.4$  s [F(2,57) = 1.883, P = 0.161]. In the Bottleneck dataset, the corresponding bin is [1.2 s, 1.4 s) resulting in the estimate  $t_0 = 1.4$  s [F(2,57) = 1.614, P = 0.208].

Over the interval of well-defined data,  $E(\tau)$  follows a power law. A linear fit of  $\log E$  vs.  $\log \tau$  with bisquare weighting reveals an exponent of  $2.05 \pm 0.123$  for the *Outdoor* dataset and  $2.017 \pm 0.192$  for the *Bottleneck* dataset. As can be seen in Fig. 2b, the interaction energy in both datasets can be well modeled with an exponent of 2[t(174) = 0.809, P = 0.42 for the *Outdoor* and t(106) = 0.171, P = 0.865 for the *Bottleneck*]. We note that, for visual clarity, the data in Fig. 2a and Fig. 2b are down-sampled, showing  $E(\tau)$  samples every  $0.02 \, \mathrm{s}$  and  $0.03 \, \mathrm{s}$  respectively.

## PAIR DISTRIBUTION FUNCTION IN 2D SPATIAL COORDINATES

Figure 1c of the main text demonstrates that the Cartesian distance r between two pedestrians is not a sufficient descriptor of their interaction, as it cannot account for the dependence of the pair distribution function g(r) on the pedestrian rate of approach v. Here we show that even the full two-dimensional displacement vector  $\mathbf{r}$  cannot adequately parameterize pedestrian interactions. In other words, we show that the empirical behavior of the pair distribution function is inconsistent with any form of the interaction that depends only on relative spatial coordinates.

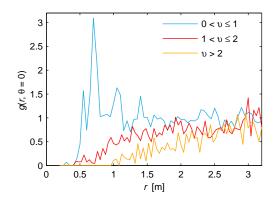


FIG. S2. The pair distribution function g in polar coordinates  $(r, \theta)$  for pedestrians in the *Outdoor* dataset, plotted for the fixed angle  $\theta = 0$ .  $g(r, \theta = 0)$  shows a significant dependence on the rate v at which the pedestrians approach each other, indicating that the displacement vector  $\mathbf{r}$  is not a sufficient descriptor of pedestrian interactions.

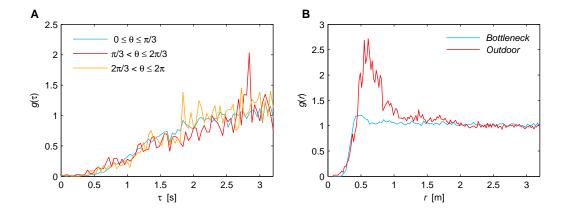


FIG. S3. Appropriateness of the time-to-collision variable  $(\tau)$  as a descriptor of pedestrian-pedestrian interactions. (A) The pair distribution function  $g(\tau)$  as a function of time-to-collision  $\tau$  for different values of the relative orientation  $\theta$  between interacting pedestrians in the *Outdoor* dataset. The different  $g(\tau)$  curves collapse onto each other, suggesting that, in addition to the velocity-independence demonstrated in Fig. 1d of the main text,  $g(\tau)$  is also independent of the pedestrians' relative orientation. (B) The pair distribution function g as a function of the distance r between pedestrian pairs that are not on a collision course (i.e., pairs that have undefined  $\tau$ ). For both the *Bottleneck* and the *Outdoor* datasets, there is no evidence of any repulsion beyond  $r \approx 0.4\,\text{m}$ . This, along with (B) and Fig. 1 of the main text, suggests that  $\tau$  alone provides an appropriate parameterization of pedestrian interactions.

The displacement vector  $\mathbf{r}$  connecting two pedestrians can be parameterized by the vector norm r and the angle  $\theta$  between  $\mathbf{r}$  and a given pedestrian's current heading. We find that, for a given fixed value of  $\theta$ , the pair distribution function  $g(r,\theta)$  shows a significant dependence on the pedestrian rate of approach v. This is shown explicitly for  $\theta = 0$  in Fig. S2. Similar conclusions can also be drawn for different values of  $\theta$ , suggesting that the displacement vector  $\mathbf{r}$  alone cannot accurately quantify pedestrian interactions.

# ORIENTATION-INDEPENDENCE OF $g(\tau)$

In Fig. 1d of the main text, it is shown that  $g(\tau)$  is independent of the rate v at which pedestrians approach other. Here we show that  $g(\tau)$  is also independent of the relative orientation between pedestrians.

Figure S3A plots  $g(\tau)$  for different values of the angle  $\theta$ , defined (as above) as the angle between a pedestrian's velocity vector and the displacement vector connecting the two interacting pedestrians. As in Fig. 1d, the curves for  $g(\tau)$  corresponding to different  $\theta$  collapse onto each other. This suggests that the interaction between pedestrians as captured by  $\tau$  is independent of the pedestrians' relative orientation.

### ABSENCE OF INTERACTION FOR UNDEFINED au

Figure 2 of the main text presents the interaction energy for pedestrian pairs with finite time-to-collision  $\tau$ . Here we demonstrate that for pedestrian pairs that are not on a collision course (i.e., for pairs with undefined  $\tau$ ), there is no evidence of any finite interaction beyond a short-ranged exclusion.

Figure S3B shows the pair distribution function g(r) plotted only for pedestrian pairs with undefined  $\tau$ . For the Bottleneck dataset,  $g(r) \approx 1$  at all  $r \gtrsim 0.4\,\mathrm{m}$ , which suggests that there is no interaction between non-colliding pedestrians once that are separated by more than  $0.4\,\mathrm{m}$ . For the Outdoor dataset, a finite repulsive interaction, corresponding to g(r) < 1, also appears only at  $r \lesssim 0.4\,\mathrm{m}$ . The peak in g(r) at  $r \approx 0.6\,\mathrm{m}$  suggests a positive correlation between non-colliding pedestrians that are not too far away. This can be mainly attributed to the presence of small groups of pedestrians that walk next to each other, either for social reasons or as a strategy for navigating through the crowd.

#### ANALYTICAL EXPRESSION FOR THE SIMULATED INTERACTION FORCE

Equation (2) of the main text defines the interaction energy of a pair of pedestrians with finite time-to-collision  $\tau$ . Within a force-based simulation model, this energy is directly related to the force  $\mathbf{F}_{ij}$  experienced by the pedestrian i due to the interaction with another pedestrian j. In particular,

$$\mathbf{F}_{ij} = -\nabla_{\mathbf{x}_{ij}} E(\tau) = -\nabla_{\mathbf{x}_{ij}} \left( k \tau^{-2} e^{-\tau/\tau_0} \right), \tag{S1}$$

as in Eq. (3) of the main text. Here,  $\tau$  is understood to be a function of the relative displacement  $\mathbf{x}_{ij} = \mathbf{x}_i - \mathbf{x}_j$  between pedestrians and their relative velocity  $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$ .

At any given simulation step, we estimate  $\tau$  by linearly extrapolating the trajectories of the pedestrians i and j based on their current velocities. Specifically, a collision is said to occur at some time  $\tau > 0$  if the corresponding discs of the pedestrians, of radii  $R_i$  and  $R_j$ , respectively, intersect. If no such time exists, the interaction force  $\mathbf{F}_{ij}$  is  $\mathbf{0}$ . Otherwise,  $\tau = \frac{b - \sqrt{d}}{a}$ , where  $a = \|\mathbf{v}_{ij}\|^2$ ,  $b = -\mathbf{x}_{ij} \cdot \mathbf{v}_{ij}$ ,  $c = \|\mathbf{x}_{ij}\|^2 - (R_i + R_j)^2$ , and  $d = b^2 - ac$ . By substituting  $\tau$  into Eq. (S1), the interaction force can be written explicitly as:

$$\mathbf{F}_{ij} = -\left[\frac{ke^{-\tau/\tau_0}}{\|\mathbf{v}_{ij}\|^2 \tau^2} \left(\frac{2}{\tau} + \frac{1}{\tau_0}\right)\right] \left[\mathbf{v}_{ij} - \frac{\|\mathbf{v}_{ij}\|^2 \mathbf{x}_{ij} - (\mathbf{x}_{ij} \cdot \mathbf{v}_{ij}) \mathbf{v}_{ij}}{\sqrt{(\mathbf{x}_{ij} \cdot \mathbf{v}_{ij})^2 - \|\mathbf{v}_{ij}\|^2 (\|\mathbf{x}_{ij}\|^2 - (R_i + R_j)^2)}}\right].$$
 (S2)

A complete simulation is produced by combining this interaction force, and a similar force associated with repulsion from static obstacles, with a driving force. C++ and Python code of our complete force-based simulation model are available at http://motion.cs.umn.edu/PowerLaw/.

## SIMULATION RESULTS

We tested the anticipatory interaction law via computer simulations using the derived force-based model. To approximate the behavior of typical humans, the preferred walking speeds of the pedestrians were normally distributed with an average value of  $1.3 \pm 0.3 \,\text{m/s}$  [10]. In all simulations, we set k = 1.5 and  $\tau_0 = 3 \,\text{s}$  as the default parameter values of the interaction forces. [See Eq. (S2)].

Details of the simulations are listed below:

- Evacuation: 150 pedestrians exit a room (10 m wide × 24 m long) through a narrow doorway. Due to the restricted movement of the pedestrians, arch-like blockings are formed near the exit, leading to clogging phenomena similar to the ones observed in granular media [11, 12]. See Fig. 3a.
- Hallway: 300 pedestrians cross paths while walking from opposite ends of an open hallway that is 20 m wide. The pedestrians dynamically form lanes of uniform walking directions to efficiently resolve collisions. See Fig. 3b.
- Bottleneck: 150 pedestrians start in a 5 m-wide waiting area and have to pass through a 5 m-long bottleneck of variable width (1 m 3 m). In all cases, pedestrians exhibit clogging behaviors in the waiting area (Fig. 3c). With wider bottlenecks, "zipping" patterns emerge inside the constriction. For example, at a width of 2.5 m, pedestrians tend to walk diagonally behind each other, dynamically forming 5-6 overlapping layers that maximize the utility of the bottleneck, as observed in real life [13].
- Crossing: Two groups, of 40 pedestrians each, cross paths perpendicularly. The pedestrians prefer to slow down and let others pass rather than deviate from their planned courses. As such, homogeneous clusters of pedestrians emerge within the two groups, leading to the formation of diagonal line-shaped patterns [14]. See Fig. 3d.
- Collective motion: 750 pedestrians are placed in an enclosed square area of size  $40 \times 40 \,\mathrm{m}$ , and at each time step are propelled forward in the direction of their current velocity without having a specific goal. Pedestrians are initially given random orientations, and after a long enough time they spontaneously form a vortex pattern in which all pedestrians are walking in unison. See Fig. 3e.

Overall, as can be seen in Fig. 3 of the main text, the derived force law described in Eq. (S2) is able to reproduce a wide variety of collective phenomena. We also used the generalized definitions of flow, speed, and density suggested by Edie [6] to measure the density-dependent behavior that the agents exhibit in several of the simulations described

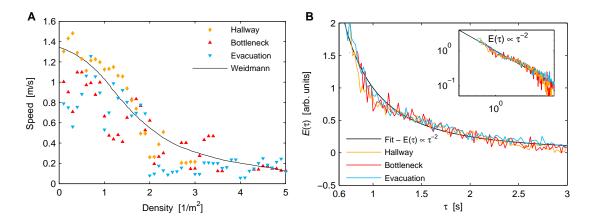


FIG. S4. Analysis of simulations generated using the interaction force described in Eq. (S2). (A) Speed-density relation seen in various simulations. The results are obtained using the generalized definitions of flow, speed and density of Edie [6] by clustering the data into bins of  $0.1 \,\mathrm{ppl/m^2}$ . The corresponding fundamental diagram is compared to measurements of Weidmann [10]. (B) The trajectories of simulated pedestrians reveal the same power-law interaction that we identify from real human crowd data (see Fig. 2 of the main text).

above. Each simulation area was divided into two-dimensional cells measuring  $0.5 \,\mathrm{m} \times 0.5 \,\mathrm{m}$  and for each cell we determined its average density and speed over 4s intervals. Figure S4A shows the corresponding fundamental diagram (here, the bottleneck samples refer to a 1 m wide bottleneck). Our results are in good agreement with the fundamental relation of speed, flow and density of real humans, as reported by Weidmann [10].

Our derived force model is also sufficient to reproduce, qualitatively, the behavior of the spatial pair distribution function g(r). This is shown in Fig. S5A, where simulation results for g(r) are plotted for different values of the pedestrian rate of approach v. The strong dependence of g(r) on v is consistent with what we observe in the Outdoor dataset (Fig. 1c). The peak in g(r) for small v can be mainly attributed to spontaneous lane formation between pedestrians walking in the same direction. Simulation approaches based on distance-dependent forces show somewhat different behavior of g(r), as shown in Fig. S4B, with a weaker dependence on v. Such distance-dependent force simulations also fail to show a strong dependence of g on the time-to-collision (as depicted in Fig. S6A).

Importantly, in addition to reproducing known crowd phenomena and flows, our simulations also reproduce the empirical behavior of  $g(\tau)$  described in this Letter. Figure S4B shows that the inferred pedestrian interaction energy  $E(\tau) \propto \ln(1/g(\tau))$  in the hallway, bottleneck (2.5 m wide), and evacuation simulations closely follows the inverse quadratic power law. In contrast, existing anticipatory models of pedestrian behavior, such as the ones proposed in [15, 16], do not consistently capture this law. This is shown explicitly in Fig. S6 for the hallway and bottleneck scenarios.

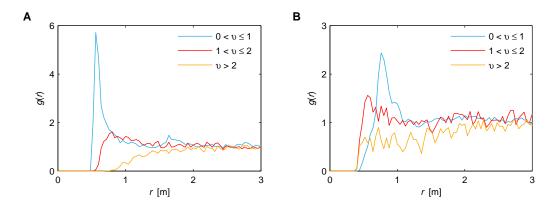


FIG. S5. The pair distribution function g(r) in the hallway simulation generated using (A) the interaction force described in Eq. (S2), and (B) the distance-dependent force described in [11]. The variable v denotes the rate of approach between pedestrians, and is measured in units of m/s. See also Fig. 1c of the main text.

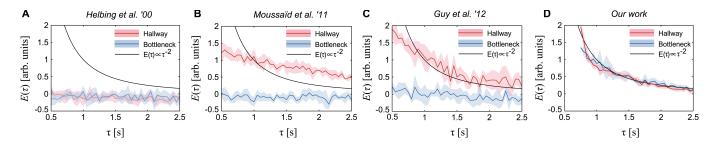


FIG. S6. The inferred interaction energy  $E(\tau) \propto \ln(1/g(\tau))$  as a function of time-to-collision ( $\tau$ ) for simulations obtained using (A) the distance-dependent model described in [11], (B) the behavioral heuristics model proposed in [16], (C) the least-effort model proposed in [15], and (D) our derived anticipatory force model. In all figures, colored lines indicate average energy values every 0.05 s and shaded regions denote  $\pm$  one standard deviation.

### NOTE

C++ and Python implementation of the anticipatory force model is available at <a href="http://motion.cs.umn.edu/PowerLaw/">http://motion.cs.umn.edu/PowerLaw/</a>, along with videos demonstrating simulation results. We also provide links to the data used to derive the power law of human interactions.

- [1] A. Seyfried, O. Passon, B. Steffen, M. Boltes, T. Rupprecht, and W. Klingsch, Transp. Sci. 43, 395 (2009).
- [2] A. Seyfried, M. Boltes, J. Kähler, W. Klingsch, A. Portz, T. Rupprecht, A. Schadschneider, B. Steffen, and A. Winkens, in Pedestrian and Evacuation Dynamics 2008, edited by W. Klingsch, C. Rogsch, A. Schadschneider, and M. Schreckenberg (Springer Berlin Heidelberg, 2010) pp. 145–156.
- [3] M. Boltes, A. Seyfried, B. Steffen, and A. Schadschneider, in *Pedestrian and Evacuation Dynamics 2008*, edited by W. Klingsch, C. Rogsch, A. Schadschneider, and M. Schreckenberg (Springer Berlin Heidelberg, 2010) pp. 43–54.
- [4] A. Lerner, Y. Chrysanthou, and D. Lischinski, Computer Graphics Forum 26, 655 (2007).
- [5] S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool, in *IEEE International Conference on Computer Vision* (2009) pp. 261–268.
- [6] L. C. Edie, in Proceedings of the Second International Symposium on the Theory of Traffic Flow: London 1963 (OECD, Paris, France, 1965) pp. 139–154.
- [7] The b250\_combined and l4\_combined datasets are available at http://www.asim.uni-wuppertal.de/database-new/own-experiments/bottleneck.
- [8] The zara\_01, zara\_02 and students003 datasets are available at http://graphics.cs.ucy.ac.cy/research/downloads/crowd-data.
- [9] The seq\_eth dataset is available at http://www.vision.ee.ethz.ch/datasets/.
- [10] U. Weidmann, Transporttechnik der Fußgänger, Literature Research 90 (ETH Zürich, 1993) in German.
- [11] D. Helbing, I. Farkas, and T. Vicsek, Nature 407, 487 (2000).
- [12] D. Helbing, Rev. Mod. Phys. **73**, 1067 (2001).
- [13] S. P. Hoogendoorn and W. Daamen, Transp. Sci. 39, 147 (2005).
- [14] S. P. Hoogendoorn and W. Daamen, in Traffic and Granular Flow '03 (Springer Berlin Heidelberg, 2005) pp. 373–382.
- [15] S. J. Guy, S. Curtis, M. C. Lin, and D. Manocha, Phys. Rev. E 85, 016110 (2012).
- [16] M. Moussaïd, D. Helbing, and G. Theraulaz, Proc. Natl. Acad. Sci. 108, 6884 (2011).