Supplemental material for:
Universal Power Law Governing Pedestrian Interactions

Ioannis Karamouzas,[1] Brian Skinner,[2] and Stephen J. Guy[1]

[1] *Department of Computer Science and Engineering, University of Minnesota, USA*
[2] *Materials Science Division, Argonne National Laboratory, USA*

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 τ, 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 (τ) 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.

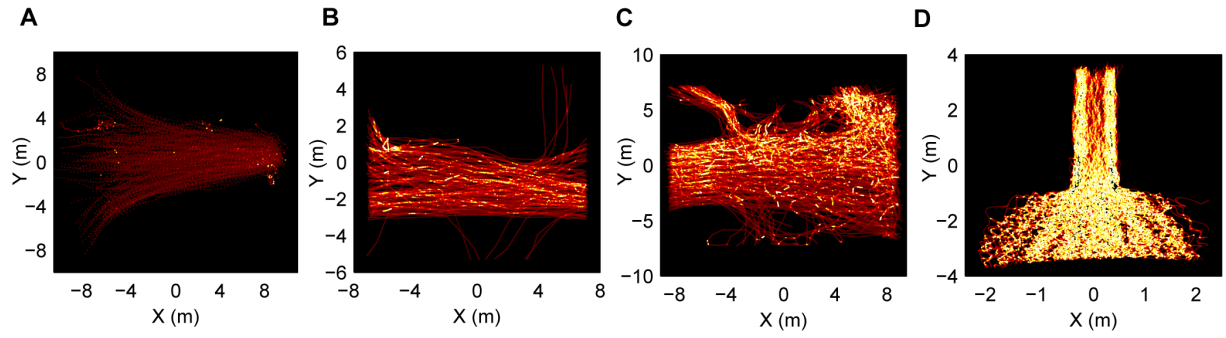

FIG. S1. Pedestrian trajectories for four of the six datasets examined in this paper, colored by time-averaged density from low (dark red) to high (white) values. (A-B) are from sparse, outdoor environments with largely bi-directional flows. (C) is from a moderately dense environment with multi-directional 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 $g(x) = P(x)/P_{\mathrm{NI}}(x)$, where $P(x)$ is the probability density function for the relative separation x between pedestrians in the dataset, and $P_{\mathrm{NI}}(x)$ 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 $P_{\mathrm{NI}}(x)$. 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 $P_{\mathrm{NI}}(x)$, allowing us to define $g(x) = P(x)/P_{\mathrm{NI}}(x)$. 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 τ.

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 τ 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 τ and becoming indistinguishable from noise at large values of τ. To estimate the lower τ 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 τ. 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 ± 0.123 for the *Outdoor* dataset and 2.017 ± 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 s and 0.03 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.

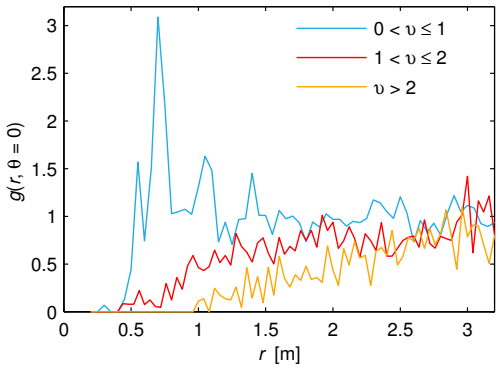

FIG. S2. The pair distribution function g in polar coordinates (r, θ) for pedestrians in the *Outdoor* dataset, plotted for the fixed angle $\theta = 0$. $g(r, \theta = 0)$ shows a significant dependence on the rate v at which the pedestrians approach each other, indicating that the displacement vector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

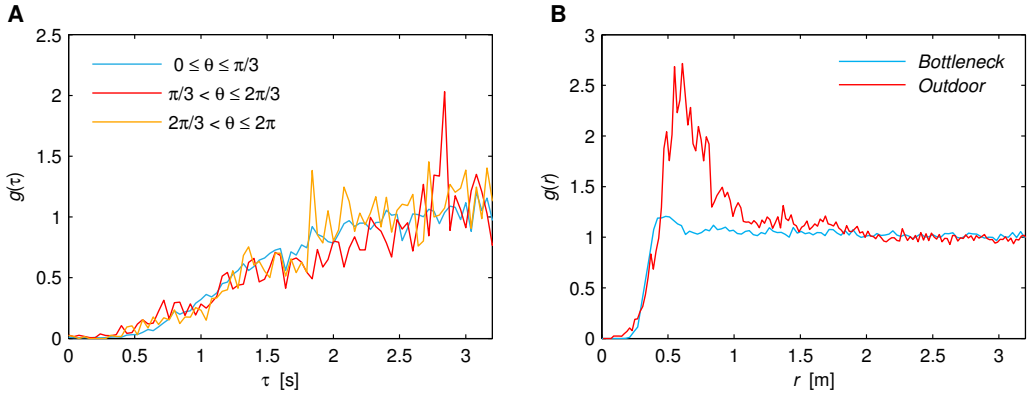

FIG. S3. Appropriateness of the time-to-collision variable (τ) as a descriptor of pedestrian-pedestrian interactions. (A) The pair distribution function $g(\tau)$ as a function of time-to-collision τ for different values of the relative orientation θ between interacting pedestrians in the *Outdoor* dataset. 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 τ). For both the *Bottleneck* and the *Outdoor* datasets, there is no evidence of any repulsion beyond $r \approx 0.4\,\mathrm{m}$. This, along with (B) and Fig. 1 of the main text, suggests that τ 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 θ between $\mathbf{r}$ and a given pedestrian's current heading. We find that, for a given fixed value of θ, 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 θ, 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 θ, 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 θ collapse onto each other. This suggests that the interaction between pedestrians as captured by τ is independent of the pedestrians' relative orientation.

ABSENCE OF INTERACTION FOR UNDEFINED τ

Figure 2 of the main text presents the interaction energy for pedestrian pairs with finite time-to-collision τ. Here we demonstrate that for pedestrian pairs that are not on a collision course (i.e., for pairs with undefined τ), 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 τ. 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 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 τ. 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, τ 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 τ 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 τ into Eq. (S1), the interaction force can be written explicitly as:

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

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 ± 0.3 m/s [10]. In all simulations, we set $k = 1.5$ and $\tau_0 = 3$ 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 × 40 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

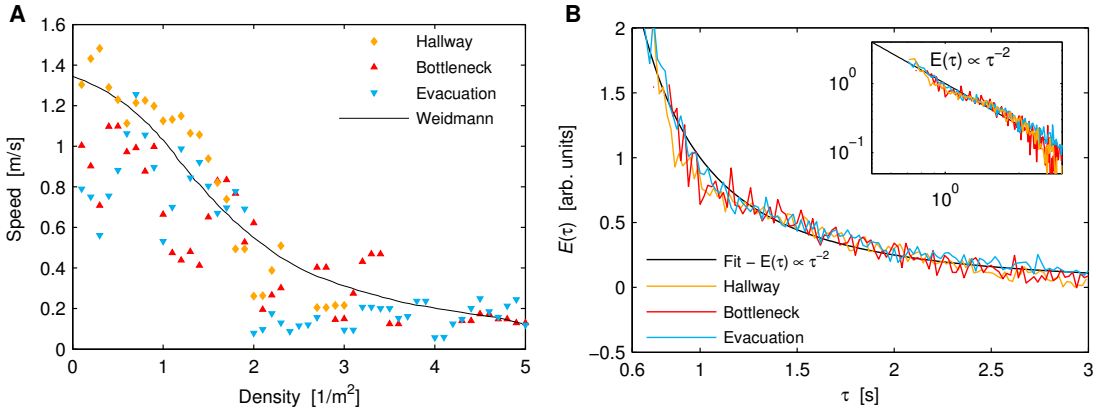

FIG. S4. Analysis of simulations generated using the interaction force described in Eq. (S2). (A) Speed-density relation seen in various simulations. The results are obtained using the generalized definitions of flow, speed and density of Edie [6] by clustering the data into bins of 0.1 ppl/m^2. The 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 m $\times$ 0.5 m and for each cell we determined its average density and speed over 4 s 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.

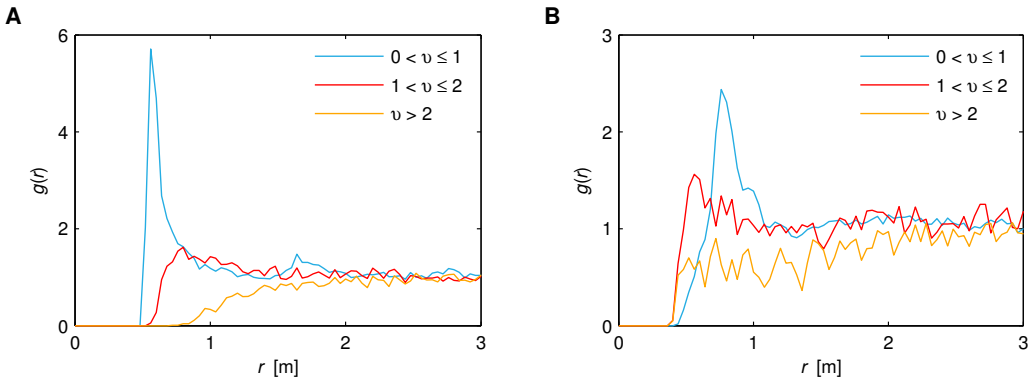

FIG. S5. The pair distribution function $g(r)$ in the hallway simulation generated using (A) the interaction force described in Eq. (S2), and (B) the distance-dependent force described in [11]. The variable v denotes the rate of approach between pedestrians, and is measured in units of m/s. See also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

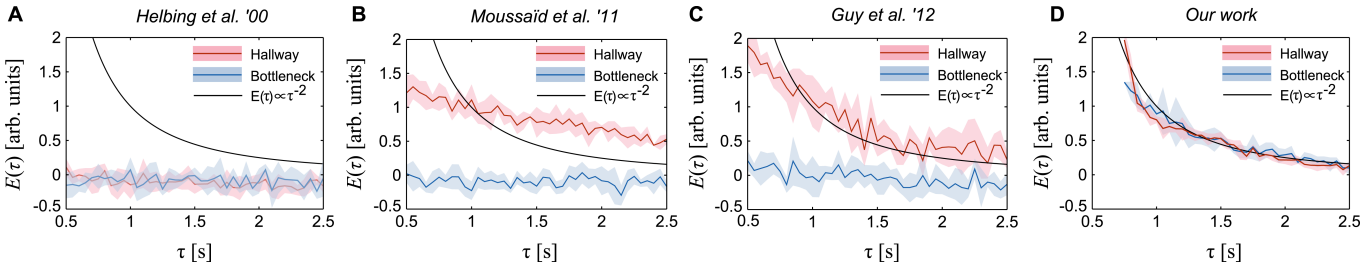

FIG. S6. The inferred interaction energy $E(\tau) \propto \ln(1/g(\tau))$ as a function of time-to-collision (τ) for simulations obtained using (A) the distance-dependent model described in [11], (B) the behavioral heuristics model proposed in [16], (C) the least-effort model proposed in [15], and (D)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

NOTE

C++ and Python implementation of the anticipatory force model is available at `http://motion.cs.umn.edu/PowerLaw/`, 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/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).