Safety Science 158 (2023) 105969

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

Macroscopic and microscopic dynamics of a pedestrian cross-flow: Part II, modelling

Francesco Zanlungo [a,b,*], Claudio Feliciani [c], Zeynep Yücel [b], Katsuhiro Nishinari [c,d], Takayuki Kanda [e]

[a] International Professional University of Technology in Osaka, 3-3-1 Umeda, Kita-ku, 530-0001, Osaka, Japan
[b] Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tsushima-naka Kita-ku, 700-8530, Okayama, Japan
[c] Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, 153-8904, Tokyo, Japan
[d] Department of Aeronautics and Astronautics, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, 153-8656, Tokyo, Japan
[e] Graduate School of Informatics, Kyoto University, Yoshidahonmachi, Sakyo-ku, 606-8317, Kyoto, Japan

ARTICLE INFO

Keywords:
Pedestrian dynamics
Cross-flow
Body orientation
Self-organising patterns

ABSTRACT

In this work, we try to reproduce empirical results concerning the behaviour of a human crowd in a cross-flow using a hierarchy of models, which differ in the details of the body shape (using a disk-shaped body vs a more realistic elliptical shape) and in how collision avoiding is performed (using only information regarding "centre of mass" distance and velocity, or actually introducing body shape information). We verified that the most detailed model (i.e., using body shape information and an elliptical body) outperforms in a significant way the simplest one (using only centre of mass distance and velocity, and disk-shaped bodies). Furthermore, we observed that if elliptical bodies are introduced without introducing such information in collision avoidance, the performance of the model is relatively poor. Nevertheless, the difference between the different models is relevant only in describing the "tails" of the observable distributions, suggesting that the more complex models could be of practical use only in the description of high density settings. Although we did not calibrate our model in order to reproduce "stripe formation" self-organising patterns observed in the crossing area, we verified that they emerge naturally in all models.

1. Introduction

Numerical simulation has become a common practice in the design of pedestrian facilities. Several commercial and non-commercial crowd simulation software are available and used in the construction of train/metro stations, shopping malls, stadia, etc. or to study past accidents (Lovreglio et al., 2020). Crowd simulators are used in a variety of contexts, including urban planning (Arisona et al., 2012), performance assessment of transportation facilities (Hoy et al., 2016) and, more recently, evaluation of policies preventing the spread of infectious diseases (Comai et al., 2020). However, due to their importance in terms of safety, evacuation scenarios are still the ones at which crowd simulations are typically targeted. Due to the widespread use of crowd simulators and the legal implications in case of accidents, a standard for the validation of evacuation codes has been recently proposed (ISO20414 (International Organization for Standardization, 2020)). This shows that simulation models have gradually moved from a purely theoretical topic discussed among researchers to an instrument employed by facility managers, whose validation criteria are defined by international standards. As such, the accuracy of simulation models plays an utmost role now than ever before. In this context, we should note that although the increasing use in realistic scenarios partially proves that simulation software have reached a sufficient degree of quality, they are still far from being perfect. The Social Force Model (SFM) (Helbing and Molnar, 1995), which is commonly used in commercial applications, has still many limitations and in this work we will discuss them and propose an approach which aims at highlighting behavioural features which are most needed in a pedestrian model.

To allow a critical assessment of the model proposed here, we will study the behaviour of a pedestrian crowd in a relatively simple geometry, the crossing area of two orthogonal corridors, each one characterised by a uni-directional flow. This is obviously an extremely idealised condition, since in the real world we do not often have perfectly orthogonal corridors and the ones existing are often created in large areas where people arrive from several directions heading to different

* Corresponding author at: International Professional University of Technology in Osaka, 3-3-1 Umeda, Kita-ku, 530-0001, Osaka, Japan.
E-mail addresses: zanlungo@atr.jp (F. Zanlungo), feliciani@g.ecc.u-tokyo.ac.jp (C. Feliciani), zeynep@okayama-u.ac.jp (Z. Yücel), tknishi@mail.ecc.u-tokyo.ac.jp (K. Nishinari), kanda@i.kyoto-u.ac.jp (T. Kanda).

https://doi.org/10.1016/j.ssci.2022.105969
Received 4 April 2022; Received in revised form 12 September 2022; Accepted 10 October 2022
Available online 29 October 2022

destinations. Nevertheless, the cross-flow scenario represents a good trade-off between a simple experimental setup and realistic observed behaviour. Considering the complex interactions among pedestrians, which require both anticipatory and collision avoidance abilities, we believe the fundamental behaviour observed in crowds in such a setup is not much different than a real scenario, especially when people are heading towards a specific destination such as in evacuations.

However, despite its remarkable properties, the cross-flow has not been explored in detail in the literature and the few studies considering this geometry have mostly analysed only macroscopic quantities, often as part of a comparison with other geometries, without discussing its properties on the microscopic level (Cao et al., 2017, 2018). On the other hand, most studies in pedestrian dynamics focus either on the behaviour of pedestrians in a simpler geometry, a single corridor (either in the uni- or bi-directional (Zhang et al., 2012; Feliciani and Nishinari, 2016; Murakami et al., 2021) case), or on possible variations of the (mostly uni-directional) bottle-neck/evacuation from a room scenario (Seyfried et al., 2009; Adrian et al., 2020; Feliciani et al., 2020b).

All the above cases are obviously very interesting and still present open questions. Even the behaviour in a simple corridor is not completely understood; see for example many works discussing conditions causing the observed variations in fundamental diagrams (density/velocity relations) (Chattaraj et al., 2009; Fujita et al., 2019; Ye et al., 2021). Furthermore, when a second, opposite, flow is allowed, the dynamics becomes considerably more complex, and possibly not completely feasible to be analysed in controlled experiments (due to social norms on walking sides (Zanlungo et al., 2012), stronger influence of social groups (Zanlungo et al., 2020), etc.).

When variables accounting for conditions close to reality are further added, things get even more complex, since route choice (Crociani et al., 2016) or exit selection (Bode et al., 2015), walking impairment or disabilities (Geoerg et al., 2019), collaborative vs competitive behaviours (Von Krüchten and Schadschneider, 2017), level of competitiveness (Feliciani et al., 2020b), social bonds (Ye et al., 2021; Zanlungo et al., 2014, 2015; Zanlungo and Kanda, 2015; Zanlungo et al., 2017a, 2019), available information (Feliciani et al., 2020a), etc. have to be taken into account; all aspects that may even be impossible to realistically reproduce as a whole in a controlled experiment setting.

Nevertheless, all the uni-directional settings share a strongly simplifying feature, namely that all pedestrians have roughly the same goal, and as a result have very similar velocities. While in general a collision avoidance model needs information concerning relative distance *and* velocity, as soon as all the pedestrians have similar velocities, the relative velocity may be (at least in a first, rough zeroth-order approximation) ignored. Although in the bi-directional corridor setting pedestrians are split into two streams with opposite goals and thus roughly opposite velocities, self-organisation (either induced by collision avoidance or, on a much faster scale, by social norm (Zanlungo et al., 2012)) causes the streams to spatially separate. Although the time scale of this separation is a problem that still needs to be fully addressed, the "stationary state" is still (in a first approximation) fundamentally uni-directional (Feliciani et al., 2018).

The above reasons explain why many research works may still rely on the "Circular Specification" (CS) (Helbing and Molnar, 1995) of the Social Force Model, although this model ignores relative velocity and may thus not be able to describe collision scenarios with a more complex dynamics (Johansson et al., 2007; Zanlungo et al., 2011).

On the other hand, the geometry studied in this work naturally induces interactions in which pedestrians in different flows have velocity vectors with relative angles $\approx \pi/2$, and since, differently from the single corridor bi-directional scenario, it does not present a trivial solution in which two continuously flowing streams are physically separated by a single stationary boundary, it does not allow for the above simplification (i.e., ignoring relative velocity).

Previous, although preliminary, studies suggest anyway the emergence of a self-organising pattern that reduces the interactions between the streams while optimising the overall flow, namely the presence of "diagonal stripes" (Naka, 1977; Ando et al., 1988; Helbing et al., 2005). This behaviour, which is well-known to be used in "centrally organised" marching parades (Rokko High School, 2017), deserves to be deeply studied in self-organising pedestrian crowds. Theoretical justification for the emergence of such a pattern can be found in Cividini et al. (2013a), Cividini and Appert-Rolland (2013) (using a discrete lattice model), Cividini et al. (2013b) (using a mean field approach), Hittmeir et al. (2016) (using a partial differential equation model) and Totzeck (2019) (using an an-isotropic agent model). Such a pattern has also been reported in a very recent experimental analysis, which focuses in particular on the relation between stripe orientation and the crossing angle (Mullick et al., 2022). On the other hand, this work focuses on density conditions while keeping the crossing angle fixed at $\pi/2$.

In a previous work (Zanlungo et al., 2022), we analysed six different experimental conditions of the crossing scenario. In each condition, we changed the pedestrian density in the flows (while keeping the density *equal between the flows*), and we repeated the experiment six times for each condition. The analysis was based on the statistical study of nine different macroscopic and microscopic observables. Some of them are quite standard in pedestrian studies (evolution of density in time, speed distribution), while others are studied less often, such as the probability distribution of distances between first neighbours (either in the same or in the crossing flow), and the distribution of the velocity direction. Other features have, at least to our knowledge, been explored even less often, such as the distribution of relative position angles between first neighbours (either in the same or in the crossing flow). Finally, as we believe that, at least in the high density regime, the detailed orientation and movement of the body may play a role in collision avoidance, we studied the body orientation angle, and its deviation with respect to the velocity direction.

In this work, we attempt to explain our findings through the use of computational models. In this respect, we would like to stress that it is not our intention to introduce a new "better" or "best" collision avoidance model, and in particular not a general purpose model, but a model that may allow us to explain the observed data, with particular attention to those data that have not been reported before (orientation and relative angles).

In this respect, we have decided to use as a starting point the "Collision Prediction" (CP) formulation of the Social Force Model (Zanlungo et al., 2011). We use this model for two reasons: (1) it is a conceptually simple (although not necessarily trivial from a computational viewpoint) model that predicts, based on a linear approximation using current *relative position and velocity*, the future possible occurrence of a collision, and generates a force "opposite to the collision", and it is thus a possible way of overcoming the limitations of the CS Social Force Model specification while preserving a similar conceptual framework; (2) it fairly reproduces the emergence of self-organising stripes in a crossing scenario (Zanlungo, 2007b,a; Menge CrowdSim, 2014).

Nevertheless, as the model does not include a specific description of the shape of the human body, nor a description of its orientation dynamics, we modify and extend it to take these aspects in consideration. As stated above, the proposed model is more an "explanation tool" than a "computational tool" to be used for practical purposes. We believe, based on the discussion above and the analysis to come in the rest of this work, that a model of collision avoidance that may properly describe the studied behaviour has to: (1) rely on information concerning relative position, velocity body shape and orientation to understand the possibility of a collision, and (2) determine the variation in velocity and orientation that may avoid such a collision. Our model is based on a conceptually simple implementation of these principles, although not necessarily in the most efficient or practical way.

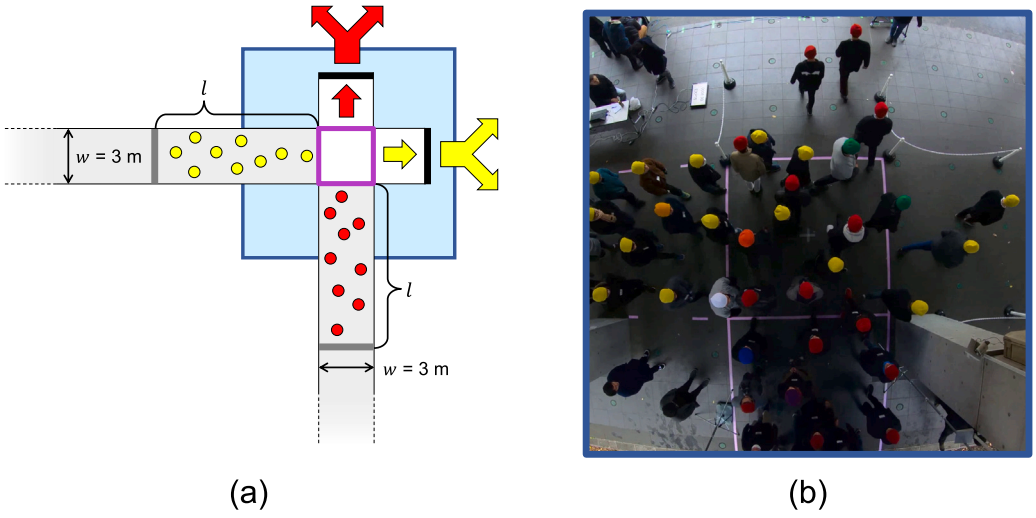

Fig. 1. (a): Schematic illustration of the experimental setup. To simplify the representation, only a limited number of pedestrians is given for reference. Width was set as $w = 3$ m for all experiments, while the size of the starting areas was varied using different values of length l, thus modifying the initial density in each unidirectional flow. (b): frame showing a specific moment (shortly after the start) in the case of an experiment with $\rho_I = 1.5$ ped/m^2 (length of the starting area set at $l = 6.0$ m). The size of the frame corresponds roughly to the blue area shown in the schematic representation in (a). Those participants wearing a cap being neither red nor yellow were wearing the tablet devices used to obtain body orientation (through the in-built gyroscope). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2. Experiments

The model calibration and evaluation performed in this work are based on an experimental campaign aimed to study the behaviour of crowds in a cross-flow scenario, which is shortly described in this section. Details on experimental setup, procedure, participants and data collection are presented by Zanlungo et al. (2022) where a detailed definition of observables is also provided along with a presentation of the empirical results. In this section, only information relevant to the understanding of the data used in this work is presented, and readers interested in subtle empirical aspects are addressed to Zanlungo et al. (2022).

To reproduce the dynamics observed in an orthogonal and symmetric cross-flow geometry, two unidirectional flows were created by having participants line up in "corridor"-like waiting areas. Each waiting area had an equal width $w = 3.0$ m and participants moved by walking straight between the lines delimiting both sides. Both unidirectional flows crossed perpendicularly in a specific area where trajectories of participants were collected. A schematic representation of the experimental setup is presented in Fig. 1(a) and a top-view image extracted from the video of an experiment is also provided in Fig. 1(b).

Density was varied by changing the length of each starting area. For instance, the highest density tested, i.e. $\rho_I = 2.5$ ped/m^2, was created by restricting participants from the entrance to the central area to a distance of 4.5 m. Similarly, in the case of the lowest density, i.e. $\rho_I = 0.25$ ped/m^2, participants were able to line up over a distance of 36 m. Experiments had been designed to have 27 participants in each area (thus, 54 in total), with the aim to create six different density conditions: $\rho_I = 0.25$, $\rho_I = 0.5$, $\rho_I = 1$, $\rho_I = 1.5$, $\rho_I = 2$ and $\rho_I = 2.5$ ped/m^2. However, as no cancellation occurred on the day of the experiments, all 56 present participants took part in the experiment, taking real density slightly above the designed figures. Nonetheless, to keep the presentation of the results clear, original design densities are used to label each experiment.

Participants were instructed on how to align at each density configuration and were simply asked to walk straight and pass the central area after hearing the start signal. Care was taken to ensure that they would not stop after crossing the area to avoid a building up of density obstructing motion. Density conditions were randomly changed to avoid participants to anticipate the condition coming next. In general, given the employed procedure and geometry, it is quite unlikely that participants were able to learn an "optimal strategy" from the first to the last trial. For each density condition 6 trials were performed, with the only exception of $\rho_I = 0.25$ ped/m^2 where the number was increased to 8.

Marking was performed in the starting and central areas, both to direct participants in the right direction and to allow a qualitative inspection of the experiments by checking the videos. In particular, given the importance of the central area a clearly recognisable colour was preferred (pink, as seen in Fig. 1(b)). Also, due to the strong interactions observed in the crossing area and to its importance in the frame of data collection and analysis, this region has been delimited using chain partitions (see Fig. 1(b)). It should be nonetheless remarked that chain partitions were low in height, and thus some participants were able to enter the "forbidden region" outside the crossing area by slightly bending their bodies. This detail was taken into consideration both in data analysis (definition of the density observable) and in simulation settings, as explained below.

Trajectories of participants in the central area were extracted using PeTrack software (Boltes et al., 2010; Boltes and Seyfried, 2013) using caps as markers. In addition, 9 participants were equipped with tablets, fixed to their chest through a bib, and the tablets' built-in gyroscope was used to measure the angular velocity of the upper part of the body and eventually gain body orientation. Data relative to body motion were synchronised with trajectories, thus allowing to know the position and body orientation at any given time for participants using the tablet. More information on data extraction and pre-processing is provided by Zanlungo et al. (2022), with details on the use of tablets to measure body orientation are discussed by Nagao et al. (2018), Feliciani and Nishinari (2022).

The data set is freely available at Feliciani et al. (2022).

3. Observables

In our work we use 9 observables to quantify cross-flow dynamics. In this section, we introduce them providing a qualitative description of their definition and meaning, while a detailed, operative and quantitative definition may be found in Zanlungo et al. (2022). The names and symbols of all observables are summarised in Table 1. Figs. 2 and 3 explain in a graphical way definitions of angles, signs and first neighbours.

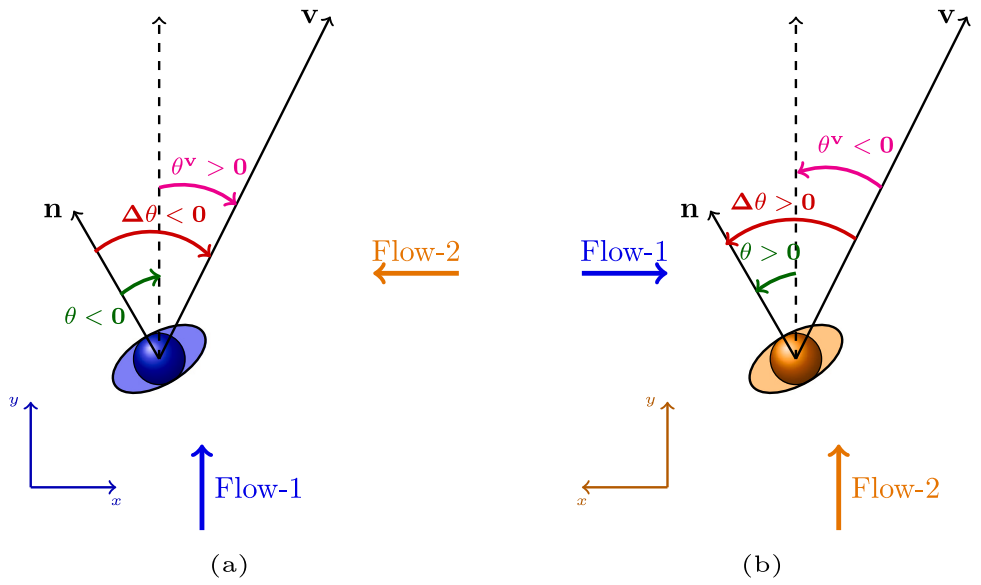

Fig. 2. Graphical definition of observables θ^v, θ and $\Delta\theta$. θ is defined as the angle between the normal to the pedestrian chest $\mathbf{n}$ and the corridor axis y corresponding to the pedestrian's flow. θ^v is defined as the angle between the velocity $\mathbf{v}$ and y, while $\Delta\theta$ is defined as the (modulo 2π) difference between θ and $\Delta\theta$. Angles are defined as positive when spanning in the direction of the x axis, the latter showing the direction from which the crossing pedestrians are coming. To better clarify the definition, (a) shows the angle definitions for pedestrians that see the crossing flow as in-coming from their right, while (b) shows the angle definitions for pedestrians that see the crossing flow as in-coming from their left. Arrows starting from the y axis represent positive angles, arrows ending on the y axis negative ones.

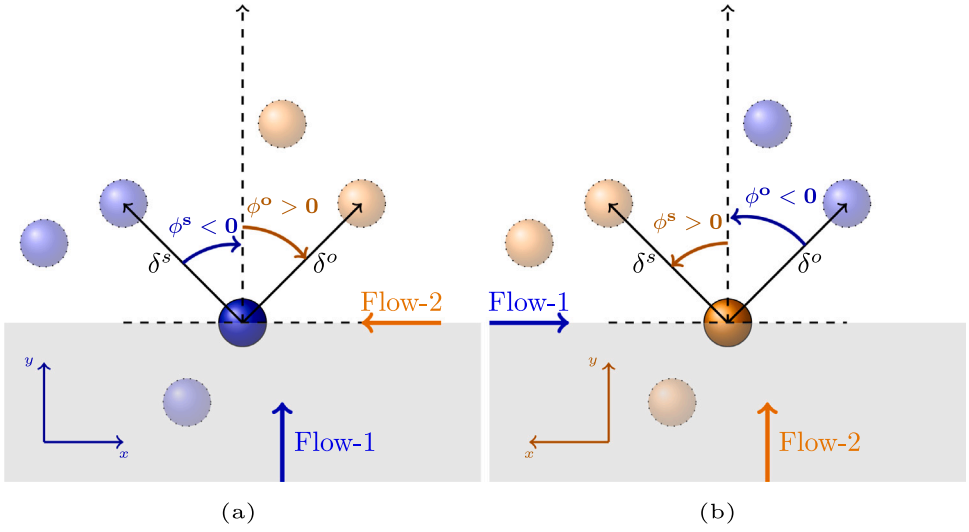

Fig. 3. Graphical definition of observables δ^s, δ^o and ϕ^s and ϕ^o. First neighbours are defined as "the closest pedestrian in the direction of motion", i.e. in the growing y directions (thus the dashed grey area is ignored), and differentiated as belonging to the same or opposite flows. Pedestrians belonging to different flows are shown in different colours. Angle signs are defined as in Fig. 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Summary of observables.

Name	Symbol
Density in the crossing area	ρ
Exit time	E_t
Pedestrian speed	v
Velocity direction angle	θ^v
Body orientation	θ
Difference between θ and θ^v	$\Delta\theta$
Distance to the first neighbour in the same flow	δ^s
Distance to the first neighbour in the crossing flow	δ^o
Same flow first neighbour relative angle	ϕ^s
Crossing flow first neighbour relative angle	ϕ^o

The *density in the crossing area* $\rho(t)$ (1) is the time dependence of the number of pedestrians tracked at each time in the crossing region divided by the area of such region, and it is measured in ped/m^2. This is the only *macroscopic* (i.e., defined at the crowd-level, and not at the individual pedestrian-level) observable that we use. Nevertheless, it is strongly related to a microscopic observable, the *exit time* (from the tracking area) E_t, and we will sometimes refer also to the latter observable because it allows for a more straightforward definition of a probability distribution.

By v we denote the *pedestrian speed* (2), whose probability distribution is denoted with $P(v)$, a notation used also for all the other pedestrian-level (microscopic) observables defined below.

To investigate also the direction of the pedestrian velocity, we study the *velocity direction angle* θ^v (3). This angle assumes values in $[-\pi, \pi)$, $\theta^v = 0$ denoting movement along the corridor axis. The angle is defined

in such a way that $\theta^v > 0$ denotes angles in the direction of the other flow is coming from (Fig. 2).

A subset of ten subjects was carrying a sensor that allowed us to know their *body orientation*. This is denoted with θ (4), assuming values in $[-\pi, \pi)$. Here $\theta = 0$ corresponds to the state in which the normal to the pedestrian chest is aligned with the corridor's axis. Again, $\theta > 0$ denotes angles in the direction the other flow is coming from.

The *difference between* θ *and* θ^v is defined as $\Delta\theta$ (5) in such a way that, again, $-\pi \leq \Delta\theta < \pi$.

For all pedestrians we also measure, as a vector, the distance to their first neighbours, distinguishing between neighbours in the same flow and in the crossing flow (refer to Fig. 3 and, for further details, to Zanlungo et al. (2022)). In order to identify the presence of a self-organising pattern, neighbours are defined to be other pedestrians *on the front* of the pedestrian under consideration, i.e., located in the direction of motion, as identified by the corridor's axis. In such a way, if a stripe pattern emerges, there will be a peak in the relative angle distribution corresponding to the stripe axis (if also neighbours on the back were present such a peak would be duplicated by symmetry).

The *magnitude of the relative distance to the first neighbour in the same flow* is denoted as δ^s (6), while the *magnitude of the relative distance to the first neighbour in the crossing flow* is denoted as δ^o (7). The angles that these distance vectors form with the corridor axis are, respectively, the *same flow first neighbour relative angle* ϕ^s (8) and *crossing flow first neighbour relative angle* ϕ^o (9). These angles are defined to assume values in $[-\pi/2, \pi/2)$, $\phi = 0$ identifying the corridor axis and $\phi > 0$ the direction the other flow is coming from.

4. Computational models

Here we provide a short description of the main features of the models we use to reproduce and hopefully understand cross-flow dynamics. For the operational details of the model implementation, we refer to Appendix C.

The CP model, introduced in its preliminary form by Zanlungo (2007b,a), and developed by Zanlungo et al. (2011), brings the concept of collision avoidance as a way to modify the velocity to avoid a "future predicted collision" (Reynolds, 1987; Van Den Berg et al., 2011; Curtis and Manocha, 2014) in the SFM framework. Using relative velocity information, it improves with respect to the original SFM specification in describing non-trivial situations (Zanlungo et al., 2011) (for a more straightforward modification of SFM to include velocity see Johansson et al. (2007)). It qualitatively reproduces the cross-flow stripe pattern (Zanlungo, 2007b,a), in a way that, at least from a qualitative viewpoint appears to be "more smooth" than in other collision avoidance models (Rokko High School, 2017).

The CP model intentionally does not consider the explicit shape or dimension of the pedestrians' bodies in the prediction of the collision. Actually, the predicted "collision" refers to a point of future maximum approach based on a linear velocity and point-like body approximation. The "repulsive" forces of SFM are then applied using such future "approach distance". As discussed by Zanlungo (2007b,a), Zanlungo et al. (2011), this is done to introduce in the SFM framework the "anticipation" of a collision, and it is intended to reproduce human behaviour in low-medium density range.

Nevertheless, at higher density, when collisions may happen in a short time, it has to be expected that the body shape and orientation have to be taken into consideration. For example, in the data analysis, reported in Zanlungo et al. (2022) we have seen that the deviation between the body orientation and the velocity, which is measured as $\Delta\theta$, increases with density, and we have also seen that the deviation between the body orientation and the "direction to the goal" (corridor axis), which is measured as θ, increases with density, although less strongly. One of the purposes of this work is to investigate the importance of this degree of freedom in crowd dynamics.

There are a few subtle points to be considered when we analyse the importance of body orientation in crowd dynamics. It is clear that, as long as we know that body orientation and velocity orientation are considerably not aligned (as shown in the statistical analysis reported by Zanlungo et al. (2022)), a model according to which the two are identified is not completely satisfying from a scientific viewpoint (Yamamoto et al., 2019). Nevertheless, this does not imply that such a model may not be adequate for most practical purposes. A (non-obese, healthy, normally walking) person, normally occupies a larger "cross-section" orthogonal to motion if body orientation and velocity are aligned. For this reason, a model in which bodies are circular and orientation is simply ignored may, in some geometrical and density settings, perform better than a model with more realistic body dynamics; and may even better fit real data. Nevertheless, this could lead to an over-fitting problem, i.e., the model could fail in describing more challenging conditions in which body orientation may not be ignored anymore. Furthermore, in such conditions, pedestrians may take advantage of their body asymmetry by rotating to decrease their "cross-section" beyond the one of a circular body, creating further differences between the two models (Yamamoto et al., 2019).

The discussion above shows clearly that both factors, namely the deviation between velocity and body orientation, denoted by us with $\Delta\theta$, and the "absolute" (i.e., with respect to the goal, or main walking direction) deviation θ, play an important role. Nevertheless, it is not clear, just by analysing the data reported by Zanlungo et al. (2022), to understand if we are studying conditions in which proper introduction of the body direction degree of freedom and of its dynamics are necessary to reproduce crowd behaviour. For this reason, we are going to introduce a class of models of increasing complexity, to try to understand which elements are necessary to describe the observed behaviour.

To simplify, we may say that we want to compare the CP model (Zanlungo et al., 2011), that did not take into consideration body orientation and its dynamics, with a model that includes such degrees of freedom and their dynamics. Anyway, the distinction is a little bit more subtle, as we may consider whether body shape and the related degree of freedom may be introduced somehow passively, i.e., by introducing the new physical degrees of freedom without modifying the decision dynamics, or more actively, i.e., by performing a detailed "prediction of collisions" between non-circular bodies. As a result, we will be actually comparing four models, summarised below and described in detail in Appendix C.

4.1. Long range (circular)

Except a few substantial differences described in Appendix C.1, this model is fundamentally an updated version of the one introduced by Zanlungo et al. (2011). It is a velocity-dependent implementation of a (Social) Force Model, i.e., a second order differential pedestrian model (Adrian et al., 2019), known to reproduce stripe formation in cross-flows (Zanlungo, 2007b,a; Rokko High School, 2017).

Following the Social Force framework, the "social force", or, more properly, the acceleration due to decisional dynamics, of a pedestrian is given by

$$\ddot{\mathbf{r}} = -k_{v_p}(\dot{\mathbf{r}} - \mathbf{v}_p) + \mathbf{F}^{long}. \tag{1}$$

Here, the first term on the right refers to the tendency of pedestrians to walk with velocity $\mathbf{v}_p$, determined by their preferred speed and wanted walking direction (the parameter k_{v_p} having dimension t^{-1}), while the second term $\mathbf{F}^{long}$ introduces the interaction with the environment related to collision avoidance. The idea behind the specification of this force can be re-conducted to Reynolds (1987). Fundamentally, through first order approximation (i.e. assuming constant velocity), the time and position of maximum approach to an obstacle (possibly a moving one, such as another pedestrian) is computed, and the interaction force (acceleration) is determined based on such "future condition". We

somehow improperly call this model "long range" because the moment of future maximum approach to the obstacle is computed based only on the geometry of trajectories (relative positions and velocity of pedestrian centre of mass), with no regard to actual body shape. Thus, although it describes local interactions with pedestrians and other obstacles, it somehow does it at a "longer range" than the model using detailed body shape and size introduced in Section 4.2.

Although the model may in principle be applied to different body shapes (and indeed we will use it also for elliptical bodies), by not taking into consideration the shape of the body its natural applications are to symmetrical ones (disks).

4.2. Short range (ellipse)

The model of Zanlungo et al. (2011) was designed for moderate densities and did not take into account the shape of the human body. In order to describe the motion of pedestrians at high density, it is necessary to consider at least the fact that the 2D projection of the human body is not rotationally symmetric (Chraibi et al., 2010; Langston et al., 2006; Heliövaara et al., 2012; Hidalgo et al., 2017; Echeverría-Huarte et al., 2020; Yajima et al., 2020) (this asymmetry may be first approximated by using ellipses instead of circles.[1]) When such an asymmetry is introduced, even if we still limit ourselves to the motion of pedestrians on a 2D plane, we need to introduce a new degree of freedom, body orientation angle θ. Assuming body orientation to be equal to velocity orientation would be a too strong limitation, since it would not allow pedestrians to rotate their torso while avoiding a collision without changing considerably their motion direction (Yamamoto et al., 2019). We thus consider the pedestrian to be characterised by 3 degrees of freedom, 2D position $\mathbf{r} = (x, y)$ and angle θ. The latter variable identifies the orientation of an ellipse with axes $A \geq B$.

As we are operating in the Social Force Model paradigm, we deal with second order differential equations[2] for the pedestrian linear and angular acceleration $\dot{\mathbf{v}} \equiv \ddot{\mathbf{r}}$ and $\dot{\omega} \equiv \ddot{\theta}$, as functions of $\mathbf{r}$, $\dot{\mathbf{r}} \equiv \mathbf{v}$, θ and $\omega \equiv \dot{\theta}$. Assuming the $\mathbf{r}$ and θ equations to be decoupled is highly unrealistic, since, while pedestrians are indeed able to walk in a direction different from their body orientation, they prefer moving in the direction of their body orientation (Willems et al., 2020). By ignoring body oscillations due to gait, we propose the following equations

$$\dot{\mathbf{v}} = -k_{v_p}(\mathbf{v} - \mathbf{v}_p) - k_\theta^v(\mathbf{v} - v\mathbf{n}) + (1 - \beta(t^c))\mathbf{F}^{long} + \beta(t^c)\mathbf{F}^{short}, \tag{2}$$

and

$$\dot{\omega} = -k_\omega \omega - k_\theta^\omega \Delta\theta + \beta(t^c)T^{short}. \tag{3}$$

Many terms deserve to be explained. k_ω accounts for the tendency of reducing body oscillations, while k_θ^v and k_θ^ω for the tendency of walking in the direction of body orientation, through the angle difference between velocity and body orientation given either by $\Delta\theta$ (see Fig. 2 and, for further details (Zanlungo et al., 2022)) or the vectorial difference between the pedestrian velocity $\mathbf{v}$ and the velocity they would have if walking with equal speed in the direction given by their chest normal $\mathbf{n}$ (refer again to Fig. 2 and, for further details (Zanlungo et al., 2022)). t^c is the time at which the first collision will happen between the ellipse representing the pedestrian body and an obstacle (such as a wall or another pedestrian), computed using an event-driven algorithm (the algorithm was based on the one proposed by Foulaadvand and Yarifard (2013)) under the assumption of no acceleration. The function β introduces the idea that far away collisions are managed just using the model of Zanlungo et al. (2011), while close ones use forces taking into account body shape ($\mathbf{F}^{short}$, defined below). In detail, β introduces two time scales τ_1, τ_2 such that

$$\beta(t) = \begin{cases} 1 \text{ if } t \leq \tau_1, \\ \frac{\tau_2 - t}{\tau_2 - \tau_1} \text{ if } \tau_1 < t \leq \tau_2. \\ 0 \text{ if } t > \tau_2. \end{cases} \tag{4}$$

As stated above, we use an event-driven algorithm to compute the time of the next collision between ellipses and obstacles (wall polygons or other ellipses). This algorithm can be used to reproduce the dynamics of hard ellipses undergoing elastic collisions, i.e., it provides also the force and torque that the ellipse undergoes at the moment of collision, under the assumption of impenetrability and conservation of energy and momentum.[3]

These forces and momenta can be used as the basis of a collision avoidance method. To understand the logic behind the model, let us first for simplicity consider the case of a 2D disc (i.e., ignoring θ) colliding frontally with a wall at velocity $\mathbf{v}$ in time t^c (t^c is the time from now at which the collision will happen if the pedestrian velocity is not modified). At the moment of collision, the pedestrian will undergo a change in velocity $\boldsymbol{\Delta}\mathbf{v} = -2\mathbf{v}$. If a force $\boldsymbol{\Delta}\mathbf{v}/2t^c$ is applied on the pedestrian, collision will be avoided, with the pedestrian just stopping short of the collision.[4]

In order to generalise this, we rely on an algorithm to detect ellipse overlapping (Foulaadvand and Yarifard, 2013) and compute forces and torques assuming energy and angular momentum conservation in elastic collisions. By this generalisation we may define, with respect to pedestrian i, for all static obstacles (i.e., walls, represented as polygons) and moving obstacles (i.e., pedestrians) j in the environment the predicted time of collision with i as $t_{i,j}^c$, and the predicted change of momentum and angular momentum of i in the collision as, respectively,[5] $\boldsymbol{\Delta}\mathbf{p}_{i,j}$ and $\boldsymbol{\Delta}\mathbf{L}_{i,j}$. Assuming $t_{i,j}^c = +\infty$ if no collision is predicted, and defining

$$t_i^c = \min_j t_{i,j}^c, \tag{5}$$

$$\bar{j} = \operatorname*{argmin}_j t_{i,j}^c, \tag{6}$$

we finally have

$$\mathbf{F}_i^{short} = \frac{\gamma}{t_i^c}\boldsymbol{\Delta}\mathbf{p}_{i\bar{j}}, \tag{7}$$

$$\mathbf{T}_i^{short} = \frac{\gamma}{t_i^c}\boldsymbol{\Delta}\mathbf{L}_{i\bar{j}}, \tag{8}$$

γ being a model parameter.

4.3. Physical dynamics

Eqs. (2) and (3) realise the "social" interaction of pedestrians, i.e., their decisional process, that is, for simplicity's sake, realised at a fixed time step $\Delta t = 0.05$ s (i.e., the decisional dynamics is solved by

[1] Given the theoretical approach of this work, we decided to follow this path since an ellipse is the most natural geometrical generalisation of a circle. Nevertheless, this choice is not to be recommended when the priority is given to issues concerning implementation and computational efficiency (Donev et al., 2005). For a more realistic body shape refer to Alonso-Marroquin et al. (2014).

[2] Or better, with the corresponding difference equation obtained by applying a Euler integrator.

[3] These are actually impulsive forces and torques, i.e., expressed as an instantaneous change in linear and angular momentum.

[4] Obviously, if the equations of motion are exactly integrated in continuous time it can be shown that the pedestrian reaches zero velocity half the way to the obstacle. Nevertheless, the approximation in the text still provides the correct order of magnitude.

[5] Impulsive forces are assumed as orthogonal to the contact surface. As a result, they are central forces for disks, and in the Circular case we have $\boldsymbol{\Delta}\mathbf{L}_{i,j} = \mathbf{0}$, consistent with our choice to ignore the variable θ for disks.

using an Euler integrator.[6]) For the physical dynamics of the system, i.e., in order to deal with *actual* (as opposed to predicted) collisions between pedestrians, we use again an event-driven algorithm (Foulaadvand and Yarifard, 2013) to detect ellipse overlapping and implement physical dynamics assuming collisions to be elastic, i.e., conserving energy and angular momentum. Such an approach has the advantages of forcing absence of overlapping, and thus constraints on space occupation, and providing a quantitative measure for the amount of collisions (kinetic energy, both linear and angular, exchanged in collisions[7] (Zanlungo et al., 2020)). The downside is that elastic collisions are not very realistic between pedestrians, but this problem should not be a serious one for a functional collision avoidance method, i.e. a method in which collisions are extremely reduced.[8]

4.4. Possible model specifications

As stated above, the main purpose of the modelling portion of this work is not to propose the best or most efficient model to reproduce the data, but to *understand which fundamental behaviours or physical features have to be introduced in the model to reproduce the crowd dynamics*. By this, we mean that although most of the details (e.g., the specific form of interaction force function, or even something more fundamental such as using a second order model that specifies accelerations as opposed to using a first order approach that acts directly on velocities) of the proposed model could be replaced by possibly better alternatives obtaining a *quantitative* change in the model's ability to reproduce crowd behaviour, we are here interested in more fundamental and arguably *qualitative* differences.

We believe these differences to be the following:

1. Introducing or not an asymmetry in the shape of the human body, and the corresponding angle dynamics.
2. Introducing or not a *short range* behaviour, in which pedestrians react to close-time collisions by taking explicitly into account the shape of their body.

The asymmetry in the human body is introduced by using an ellipse with semi-axes $A \neq B$, and by using non-zero values of k_θ^v, k_ω, k_θ^ω. We will call such a choice of parameters an *elliptical* specification of the proposed model.[9] In this work, for simplicity's sake, in elliptical specifications we use $A_E = 0.225$ m, $B_E = 0.1$ m, which may be considered to provide a rough approximation of the 2D projection of the body of Japanese young adults that took part in the experiment (Digital Human Laboratory, 1992).

On the other hand, using a circle with $R = A_c = B_c = \sqrt{A_E B_E}$ (so that the same 2D area is occupied by the pedestrian's body) and zero values for k_θ^v,a.24 k_ω, k_θ^ω, corresponds to a *circular* specification of the proposed model.[10]

The short range behaviour is introduced by having non-zero values for parameters γ, τ_1 and τ_2. We refer to such a model specification as *full* (meaning it includes both short and long range behaviours), while a model with $\gamma = \tau_1 = \tau_2 = 0$ is referred to as *no-short* (range behaviour).[11]

We will thus consider, and explicitly calibrate, four possible qualitatively different alternatives of the proposed model: (1) *Full Elliptical* (often just referred to as *Full*) (2) *Full Circular*, (3) *No-short Elliptical* and (4) *No-short Circular* (often referred to as CP). In our analysis of the results, we will need to take into account that different specifications of the model correspond to different dimensions of the parameter space, and thus different optimisation problems. Specifically, 8 parameters need to be calibrated for a *No-short Circular* specification, 11 for the *Full Circular* and *No-short Elliptical* ones, and 14 for a *Full Elliptical* specification (see Appendix C.3).

The source code of the model(s) is freely available at Zanlungo (2022c).

5. Full Elliptical model calibration and evaluation

We used Genetic Algorithms ((GA), 24 runs, each one using 30 generations, 50 genomes and 8 different sets of mutation, crossover and selection hyper-parameters) to calibrate the *Full Elliptical* model (the quantitative comparison to the other models is discussed in Section 6). The fitness function was based on an Earth Mover Distance (EMD, Appendix E) comparison between the observables at initial densities of 0.25 and 1.5 ped/m^2, and the ones produced by the model. Initial conditions (entry points in the crossing area and corresponding velocity) were taken from the data, and each experimental run was reproduced.

After performing an exploratory analysis to decide which observables to include in the fitness function, we noticed that the models were able to reproduce to a good extent the ϕ^s and ϕ^o pdfs (and thus the related qualitative emergence of a stripe pattern) even though such observables were not used for calibration. We thus decided to include in the fitness function only the remaining observables, namely ρ (through the E_t pdf), v, θ^v, θ, $\Delta\theta$, δ^s and δ^o. The results concerning observables ϕ^s and ϕ^0 are anyway shown for evaluation in Section 6.

The fitness function was defined using the EMD (see Appendix E) between the experimental and model observable pdfs. For each model solution and observable, we computed a probability distribution by averaging over repetitions with equal density initial conditions, and compared the result to the experimental one by computing the EMD distance (the used pdfs are normalised to have 1 as the sum over all bin probabilities). Finally, the result (i.e., the EMD distance) was averaged over all the observables and over the two density conditions used in calibration, namely $\rho_I = 0.25$ ped/m^2 and $\rho_I = 1.5$ ped/m^2. The fitness function was finally defined as the negative of such average[12]

Figs. 4–9 show the results for one of the best performing solutions.

[6] As discussed in previous works on the CP model (Zanlungo et al., 2011), it is important that Δt is comparable to the time scale of human decisions, i.e., to the reaction times of humans to external events. In this work, we did not study model dependence on Δt.

[7] This information was not used in this work, since we were missing empirical data on possible collisions between pedestrians in the experiments.

[8] The implemented algorithm presented a convergence problem in extremely packed settings. Namely, events were predicted consistently at time steps close to machine precision, "stalling" the numerical integrator. Since such packed settings are not expected to occur in proper pedestrian behaviour, these convergence problems occur only in the first steps of the optimisation process. We thus introduced in our Genetic Algorithm optimisation procedure an exception to deal with such problems, namely halting the integration and giving a very low fitness to the solution that caused the problem.

[9] This should not be confused with the Elliptical Specification of SFM proposed by Johansson et al. (2007), that corresponds to modifying the SFM force potential creating an asymmetry due to the velocity direction, but unrelated to body shape.

[10] Again, this should not be confused with the original SFM model of Helbing and Molnar (1995), which is referred to as a Circular Specification by Johansson et al. (2007) and Zanlungo et al. (2011) due to the lack of velocity direction dependent asymmetry in its interaction force.

[11] Since we consider the Zanlungo et al. (2011) model as a starting point, due to its ability to reproduce qualitatively the stripe formation in a cross-flow (Zanlungo, 2007b,a; Rokko High School, 2017), all models include the long range module.

[12] This computation is completely equivalent to the one performed in the model evaluation of Section 6, the only difference being that evaluation is performed comparing to the $\rho_I = 2.5$ ped/m^2 initial density condition (and averaged over solutions). Namely, the fitness function for the Full and No-short Elliptical models is equivalent to the "Total evaluation measure" of Table 2, and the fitness function for the Full and No-short Circular models is equivalent to the "No body evaluation measure" of Table 5.

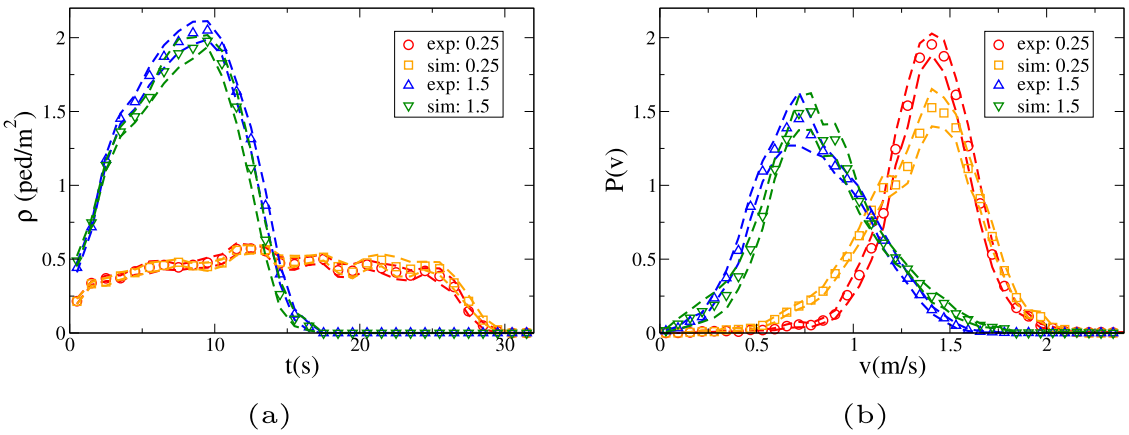

Fig. 4. (a): $\rho(t)$ for the experimental $\rho_I = 0.25$ ped/m^2 initial condition (exp: 0.25, red circles), the calibrated $\rho_I = 0.25$ ped/m^2 initial condition (sim: 0.25, orange squares), the experimental $\rho_I = 1.5$ ped/m^2 initial condition (exp: 1.5, blue up triangles), the calibrated $\rho_I = 1.5$ ped/m^2 initial condition (sim: 1.5, green down triangles). (b): $P(v)$ for the experimental $\rho = 0.25$ ped/m^2 initial condition (exp: 0.25, red circles), the calibrated $\rho = 0.25$ ped/m^2 initial condition (sim: 0.25, orange squares), the experimental $\rho = 1.5$ ped/m^2 initial condition (exp: 1.5, blue up triangles), the calibrated $\rho = 1.5$ ped/m^2 initial condition (sim: 1.5, green down triangles). Dashed lines provide standard error intervals (computed over independent repetitions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

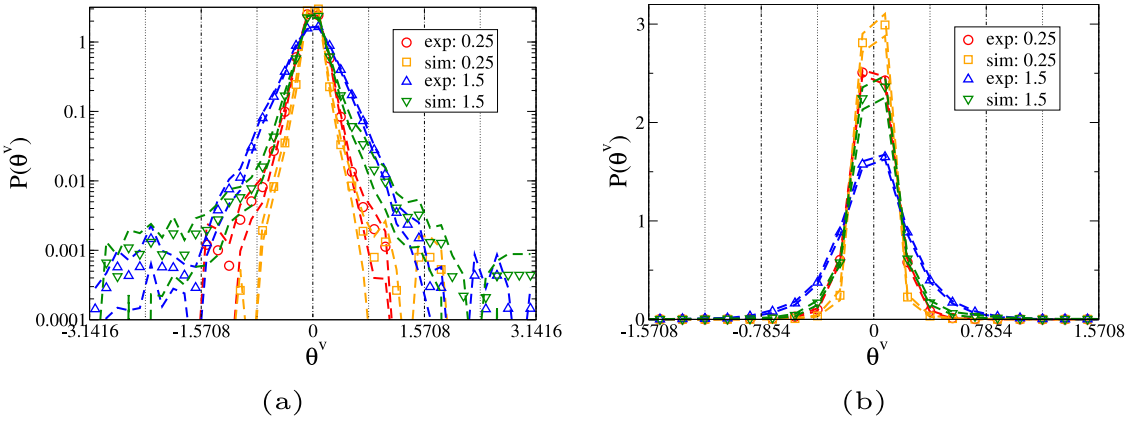

Fig. 5. $P(\theta^v)$ for the experimental $\rho_I = 0.25$ ped/m^2 initial condition (exp: 0.25, red circles), the calibrated $\rho_I = 0.25$ ped/m^2 initial condition (sim: 0.25, orange squares), the experimental $\rho_I = 1.5$ ped/m^2 initial condition (exp: 1.5, blue up triangles), the calibrated $\rho_I = 1.5$ ped/m^2 initial condition (sim: 1.5, green down triangles). (a): logarithmic plot; (b): linear plot. Dashed lines provide standard error intervals (computed over independent repetitions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

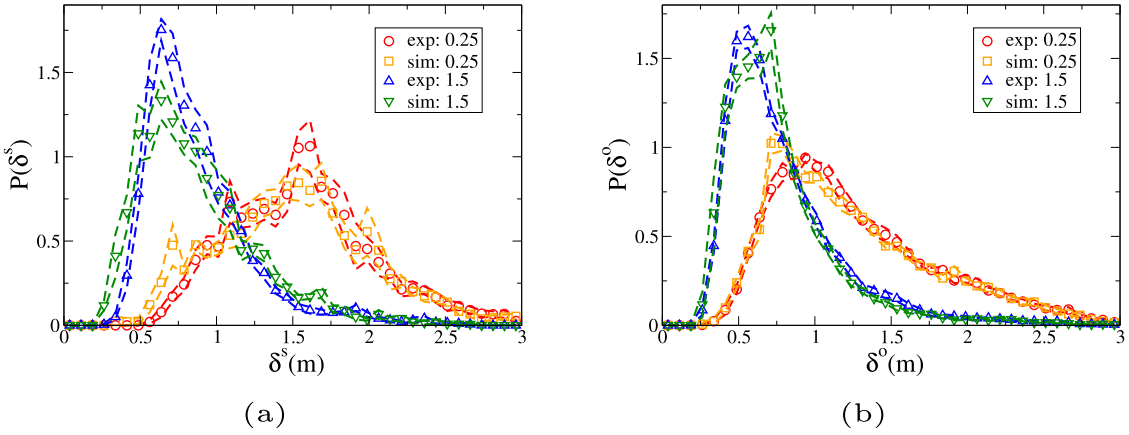

Fig. 6. (a): $P(\delta^s)$ for the experimental $\rho_I = 0.25$ ped/m^2 initial condition (exp: 0.25, red circles), the calibrated $\rho_I = 0.25$ ped/m^2 initial condition (sim: 0.25, orange squares), the experimental $\rho_I = 1.5$ ped/m^2 initial condition (exp: 1.5, blue up triangles), the calibrated $\rho_I = 1.5$ ped/m^2 initial condition (sim: 1.5, green down triangles). (b): $P(\delta^o)$ for the experimental $\rho_I = 0.25$ ped/m^2 initial condition (exp: 0.25, red circles), the calibrated $\rho_I = 0.25$ ped/m^2 initial condition (sim: 0.25, orange squares), the experimental $\rho_I = 1.5$ ped/m^2 initial condition (exp: 1.5, blue up triangles), the calibrated $\rho_I = 1.5$ ped/m^2 initial condition (sim: 1.5, green down triangles). Dashed lines provide standard error intervals (computed over independent repetitions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

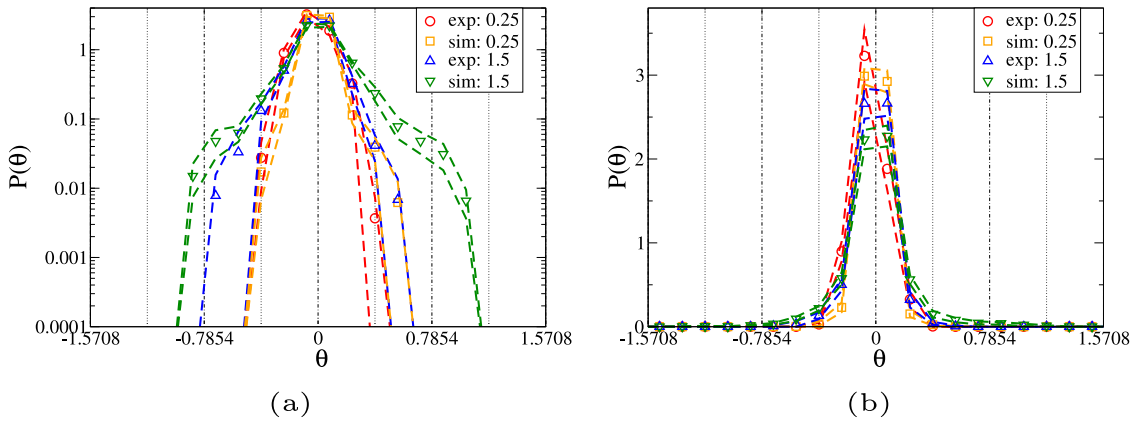

Fig. 7. $P(\theta)$ for the experimental $\rho_I = 0.25$ ped/m^2 initial condition (exp: 0.25, red circles), the calibrated $\rho_I = 0.25$ ped/m^2 initial condition (sim: 0.25, orange squares), the experimental $\rho_I = 1.5$ ped/m^2 initial condition (exp: 1.5, blue up triangles), the calibrated $\rho_I = 1.5$ ped/m^2 initial condition (sim: 1.5, green down triangles). (a): logarithmic plot. (b): linear plot. Dashed lines provide standard error intervals (computed over independent repetitions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

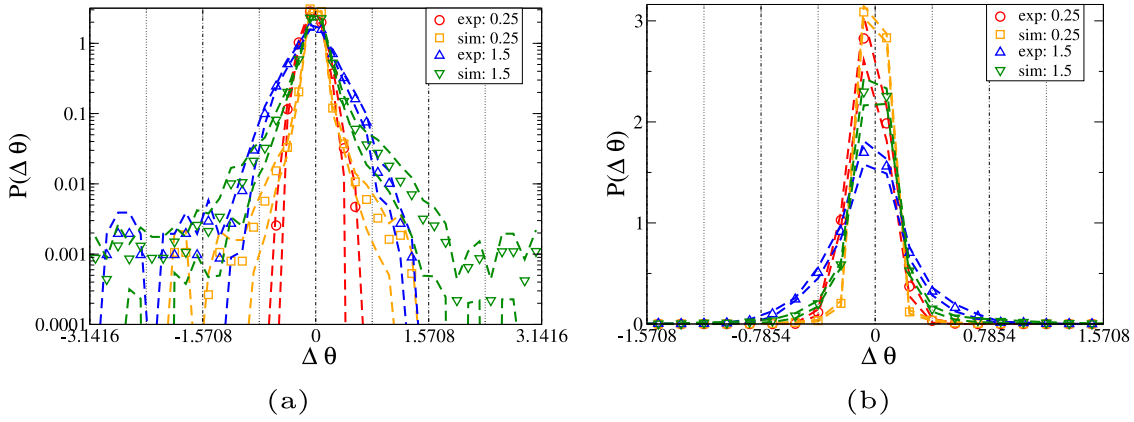

Fig. 8. $P(\Delta\theta)$ for the experimental $\rho_I = 0.25$ ped/m^2 initial condition (exp: 0.25, red circles), the calibrated $\rho_I = 0.25$ ped/m^2 initial condition (sim: 0.25, orange squares), the experimental $\rho_I = 1.5$ ped/m^2 initial condition (exp: 1.5, blue up triangles), the calibrated $\rho_I = 1.5$ ped/m^2 initial condition (sim: 1.5, green down triangles). (a): logarithmic plot. (b): linear plot. Dashed lines provide standard error intervals (computed over independent repetitions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

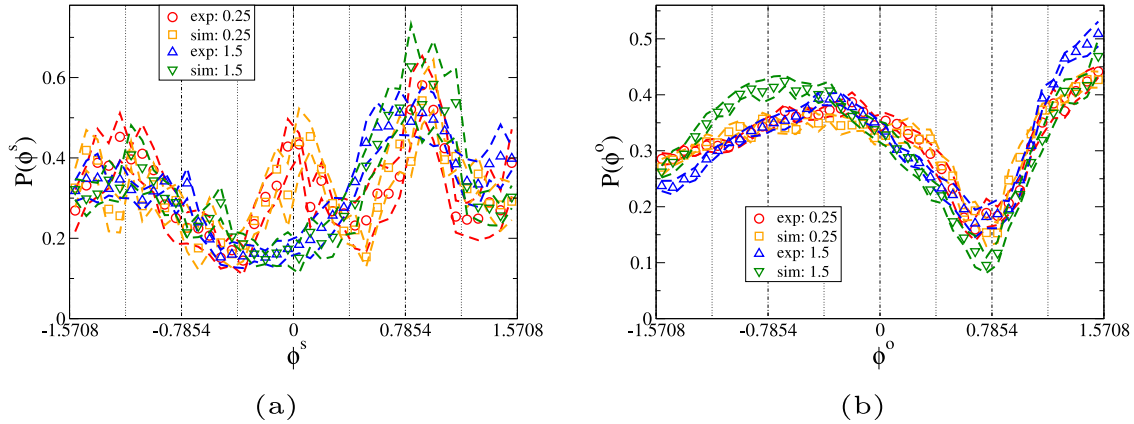

Fig. 9. (a): $P(\phi^s)$ for the experimental $\rho_I = 0.25$ ped/m^2 initial condition (exp: 0.25, red circles), the calibrated $\rho_I = 0.25$ ped/m^2 initial condition (sim: 0.25, orange squares), the experimental $\rho_I = 1.5$ ped/m^2 initial condition (exp: 1.5, blue up triangles), the calibrated $\rho_I = 1.5$ ped/m^2 initial condition (sim: 1.5, green down triangles). (b): $P(\phi^o)$ for the experimental $\rho_I = 0.25$ ped/m^2 initial condition (exp: 0.25, red circles), the calibrated $\rho_I = 0.25$ ped/m^2 initial condition (sim: 0.25, orange squares), the experimental $\rho_I = 1.5$ ped/m^2 initial condition (exp: 1.5, blue up triangles), the calibrated $\rho_I = 1.5$ ped/m^2 initial condition (sim: 1.5, green down triangles). Dashed lines provide standard error intervals (computed over independent repetitions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

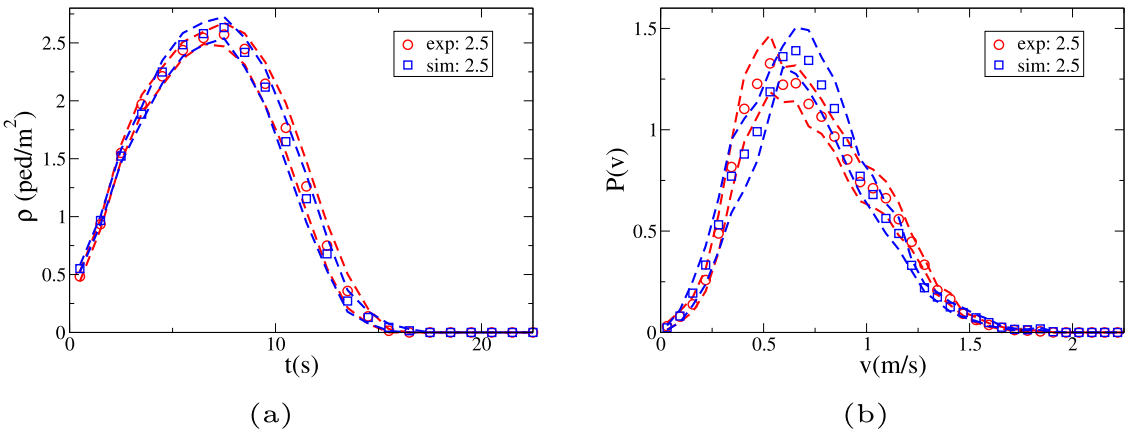

Fig. 10. (a): $\rho(t)$ for the experimental $\rho_I = 2.5$ ped/m^2 initial condition (exp: 2.5, red circles) and the calibrated $\rho_I = 2.5$ ped/m^2 initial condition (sim: 2.5, blue squares). (b): $P(v)$ for the experimental $\rho_I = 2.5$ ped/m^2 initial condition (exp: 2.5, red circles) and the calibrated $\rho_I = 2.5$ ped/m^2 initial condition (sim: 2.5, blue squares). Dashed lines provide standard error intervals (computed over independent repetitions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

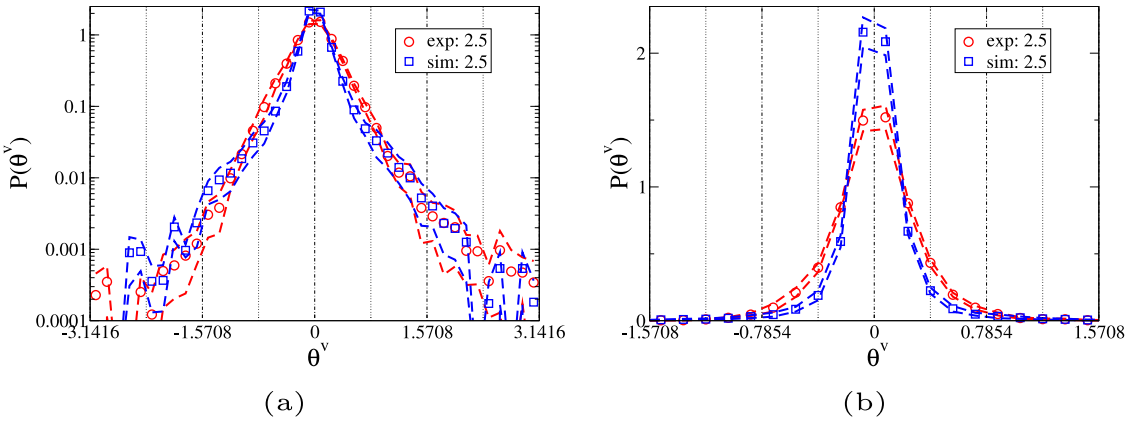

Fig. 11. $P(\theta^v)$ for the experimental $\rho_I = 2.5$ ped/m^2 initial condition (exp: 2.5, red circles) and the calibrated $\rho_I = 2.5$ ped/m^2 initial condition (sim: 2.5, blue squares). (a): logarithmic plot. (b): linear plot. Dashed lines provide standard error intervals (computed over independent repetitions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

We run the model over the $\rho_I = 2.5$ ped/m^2 initial condition and compare it to experimental results in Figs. 10–15 for the same solution as in calibration, while Figs. 16–23, which are discussed in more detail in Section 6, show results averaged over all solutions.

The examined solution appears to reproduce almost perfectly the time evolution of density (Figs. 4 (a) and 10 (a)) for all the three initial conditions, both in calibration and evaluation.[13] This is obviously one of the most important features of pedestrian crowd models. Not all GA solutions are equally good at reproducing the empirical data, and some present a longer tail, due to few pedestrians being stuck for a long time in one corner (at the crossing point of the exiting flows, i.e., on the top-right of Fig. 1(b) or bottom-right of Figs. A.24 (a), (b), (c)). According to a qualitative visual analysis of the simulations, this behaviour is probably mainly caused by the triviality of the path planning mechanism Appendix C.2. Nevertheless, as it happens for high density scenarios, it is clearly accentuated by poor collision avoidance, and it does not appear in good performing models and solutions. The longer tail is better analysed in the logarithmic plots of Fig. 16(b), which includes for all models the average over all solutions, including a few sub-optimal ones.

Concerning the speed distribution (Figs. 4 (b) and 10 (b)), the analysed solution is able to reproduce in a satisfying way calibration and evaluation distributions. On the other hand, some of the solutions reproduced pdfs skewed towards higher velocities at low densities. This tendency could be probably due to the fact that the model assumes a density-independent preferred speed, while the actual preferred speed may be different between free walking and packed settings.

Concerning θ^v, the chosen solution appears to describe better the tails (logarithmic plots in Figs. 5 (a) and 11 (a)) than the bulk (linear plots in Figs. 5 (b) and 11 (b)) of the distributions. This behaviour is not typical, since, as shown in Fig. 18, most solutions have (for any density) a tendency to overestimate the values of $P(\theta^v)$ at $\theta^v \approx 0$ and $|\theta^v| \approx \pi$, while intermediate values are underestimated.

Excluding a slight underestimation of the minimum value in the high-density support (non-zero domain) of $P(\delta^s)$, the reproduction of the δ observables is almost perfect at low and high densities, in calibration and evaluation (Figs. 6 (a), 6 (b), 12 (a) and 12 (b)).

Reproducing the θ observables is probably the main task of the proposed model. The chosen solution is, compared to other solutions

[13] Almost all solutions of all the models reproduce very well the time evolution of density at $\rho_I = 0.25$ ped/m^2. This is due to the fact that at low density pedestrians are (almost) freely-moving, and thus the time evolution of density, including to a large extent also fluctuations, is mostly determined by entry times, which are specified as boundary conditions in the simulations (and thus are independent of the model solution).

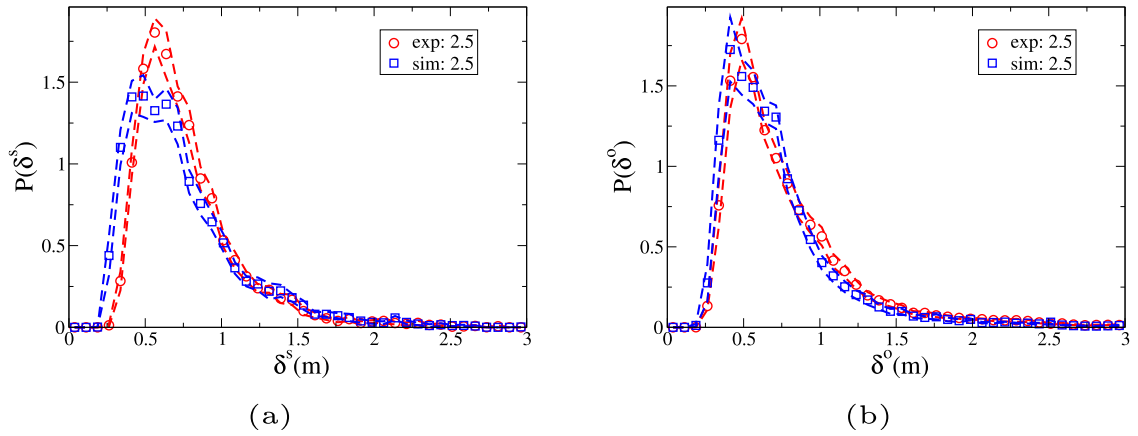

Fig. 12. (a): $P(\delta^s)$ for the experimental $\rho_I = 2.5$ ped/m^2 initial condition (exp: 2.5, red circles) and the calibrated $\rho_I = 2.5$ ped/m^2 initial condition (sim: 2.5, blue squares). (b): $P(\delta^o)$ for the experimental $\rho_I = 2.5$ ped/m^2 initial condition (exp: 2.5, red circles) and the calibrated $\rho_I = 2.5$ ped/m^2 initial condition (sim: 2.5, blue squares). Dashed lines provide standard error intervals (computed over independent repetitions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

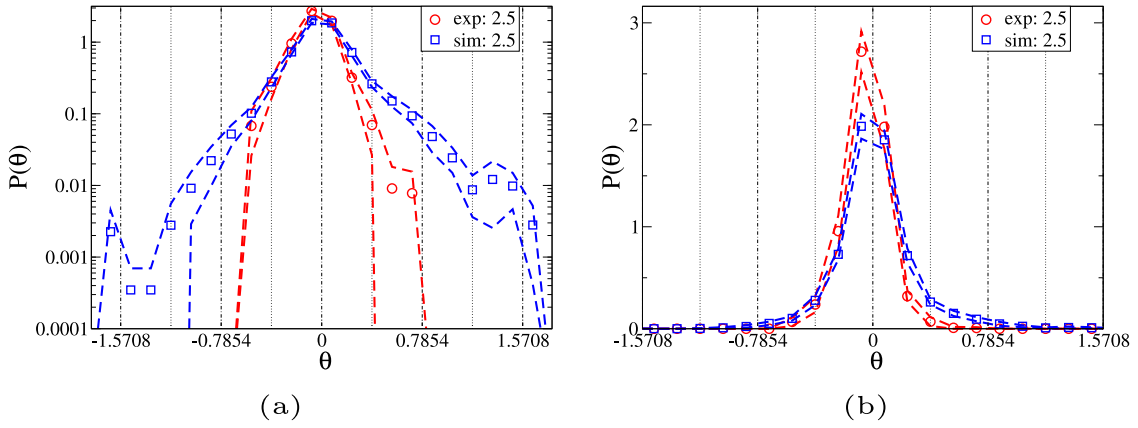

Fig. 13. $P(\theta)$ for the experimental $\rho_I = 2.5$ ped/m^2 initial condition (exp: 2.5, red circles) and the calibrated $\rho_I = 2.5$ ped/m^2 initial condition (sim: 2.5, blue squares). (a): logarithmic plot. (b): linear plot. Dashed lines provide standard error intervals (computed over independent repetitions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

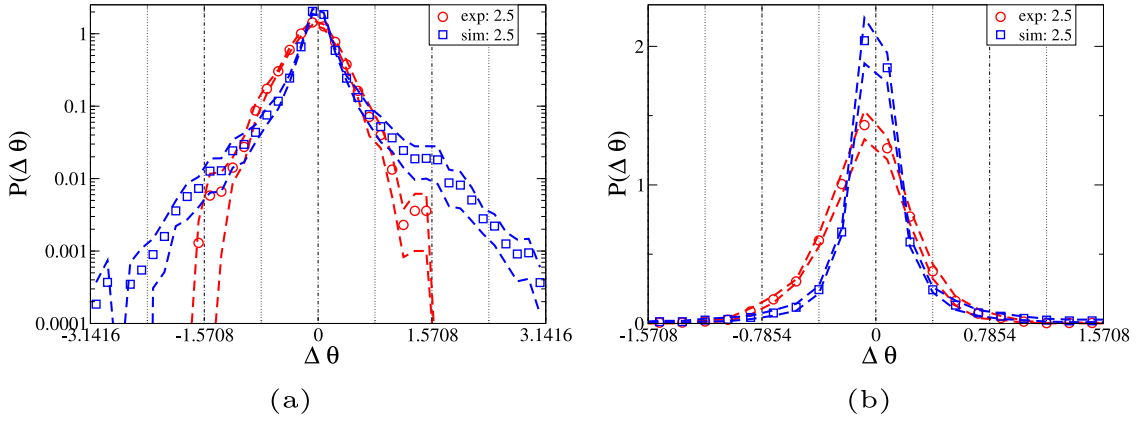

Fig. 14. $P(\Delta\theta)$ for the experimental $\rho_I = 2.5$ ped/m^2 initial condition (exp: 2.5, red circles) and the calibrated $\rho_I = 2.5$ ped/m^2 initial condition (sim: 2.5, blue squares). (a): logarithmic plot. (b): linear plot. Dashed lines provide standard error intervals (computed over independent repetitions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

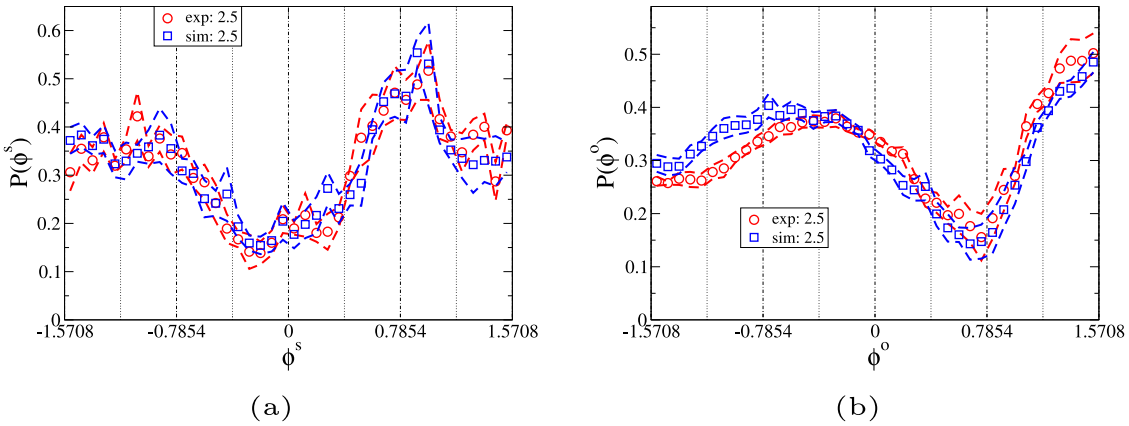

Fig. 15. (a): $P(\phi^s)$ for the experimental $\rho_I = 2.5$ ped/m^2 initial condition (exp: 2.5, red circles) and the calibrated $\rho_I = 2.5$ ped/m^2 initial condition (sim: 2.5, blue squares). (b): $P(\phi^o)$ for the experimental $\rho_I = 2.5$ ped/m^2 initial condition (exp: 2.5, red circles) and the calibrated $\rho_I = 2.5$ ped/m^2 initial condition (sim: 2.5, blue squares). Dashed lines provide standard error intervals (computed over independent repetitions). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(see Figs. 22 and 23), relatively good in reproducing the tails of the distributions (logarithmic plots in Figs. 7 (a), 8 (a), 13 (a) and 14 (a)), while still describing in a satisfactory way the θ and, to a lesser extent, $\Delta\theta$, bulks (linear plots in Figs. 7 (b), 8 (b), 13 (b) and 14 (b)). Nevertheless, even this solution fails in reproducing the fact that in empirical distributions $|\Delta\theta|$ is limited to assume values smaller than $\pi/2$, while $|\theta|$ is limited to assume values smaller than $\pi/4$. This tendency is stronger in less performing GA solutions (Figs. 22 and 23), that usually overestimate the probability of having very large deviations from 0, the latter effect being stronger for θ as compared to $\Delta\theta$.

There are probably two reasons for this phenomenon. One is the choice of the calibration method, since areas with $P \ll 1$ do not give a strong contribution to the EMD metric. The second one is related to the model itself. While our model includes a linear recall force in $\Delta\theta$ (Eq. (3)), Fig. 7(a) and 13 (a) suggest that pedestrians have a strong tendency, almost a constraint, to avoid $|\theta| > \pi/4$. For this reason, non-linearity may be introduced in future models.

Finally, concerning ϕ observables, we may recall that these were not calibrated, so all the results in Figs. 9 (a), 9 (b), 15 (a) and 15 (b) have to be considered as an evaluation of a dynamical property of the system that has not been calibrated. It is our opinion that any model that is able to reproduce fairly the other observables (and maybe even just the time-density relation), should exhibit the stripe pattern described by Zanlungo et al. (2022), and thus reproduce at least qualitatively the maxima and minima in the figures (see Appendix A for a qualitative visualisation of stripe formation in experiments and models).

6. Model comparison

To compare models, we evaluate them on the $\rho_I = 2.5$ ped/m^2 initial condition. More in detail, we obtain for each model 24 independent GA solutions (calibrating on $\rho_I = 0.25$ ped/m^2 and $\rho_I = 1.5$ ped/m^2 initial conditions), and evaluate them on the $\rho_I = 2.5$ ped/m^2 initial condition. For each observable, we compute the EMD, and average it over the different GA solutions. As a way to compare the overall performance of a model, we also provide a "Total evaluation measure" as an average over all the observables on which the model was calibrated. Nevertheless, as observables θ and $\Delta\theta$ are not defined for circular models, all comparisons involving a circular model are performed computing a "No body (angles) evaluation measure" (i.e., not considering θ and $\Delta\theta$). Observables ϕ^s and ϕ^o are not included in the "Total" and "No body" evaluation measures, since relative angles were not used in calibration.

Table 2
EMD metric comparison between Full and No-short elliptical models, evaluation at $\rho_I = 2.5$ ped/m^2.

	Full $\langle\,\rangle \pm \varepsilon(\sigma)$	No-short $\langle\,\rangle \pm \varepsilon(\sigma)$	p	Effect size
Total	0.0147 ± 0.001(0.005)	0.0191 ± 0.0011(0.005)	0.0065	0.84
ρ	0.00967 ± 0.0021(0.01)	0.0142 ± 0.0017(0.0082)	0.1	0.49
v	0.0194 ± 0.0034(0.016)	0.0291 ± 0.003(0.015)	0.037	0.63
θ^v	0.00887 ± 0.00053(0.0025)	0.0109 ± 0.00091(0.0044)	0.063	0.56
δ^s	0.0189 ± 0.0011(0.0054)	0.0257 ± 0.00092(0.0044)	2.5e−05	1.4
δ^o	0.0164 ± 0.0015(0.0074)	0.0198 ± 0.0024(0.011)	0.24	0.35
ϕ^s	0.0207 ± 0.0017(0.0083)	0.0314 ± 0.0021(0.0099)	0.00025	1.2
ϕ^o	0.0424 ± 0.0032(0.015)	0.0384 ± 0.0033(0.016)	0.39	0.26
θ	0.0156 ± 0.0008(0.0039)	0.0164 ± 0.0019(0.0093)	0.69	0.12
$\Delta\theta$	0.0144 ± 0.00066(0.0032)	0.018 ± 0.00099(0.0047)	0.0049	0.87

Nevertheless, the EMD values for these observables are reported in tables.

For each model we performed 24 calibration runs, using 8 different sets of hyper-parameters (i.e., each set is used 3 times). Figs. 16–23 compare the observable pdf obtained by each model on the $\rho_I = 2.5$ ped/m^2 initial condition with the experimental data (experimental data are reported as average over all experiment repetitions, while the model standard errors are computed by averaging over all GA solutions; i.e., for each solution we first compute the average over experiment repetitions, and then use the variation of such averages on GA solutions to compute the standard errors in the figures).

As tools to quantify differences and guide our discussion (without any claim to "prove" or "disprove" anything), we compute for each metric a t test p value and effect size (see Appendix D for definition and interpretation of these statistical indicators).

6.1. Statistical comparison

We first verify what happens if we remove the short range behaviour, i.e., if we compare the *Full* model to a *No-short* Elliptical specification (see Table 2). We then compare the Full Elliptical body model to the Full Circular one in Table 3, the Full Elliptical model to the No-Short Circular (CP) model in Table 4, and finally compare in Table 5 the Full Circular and No-short Circular models.

6.2. Discussion

6.2.1. Long vs. Short range

When comparing the two models that are able to describe the body orientation degree of freedom (the Full and No-short Elliptical models),

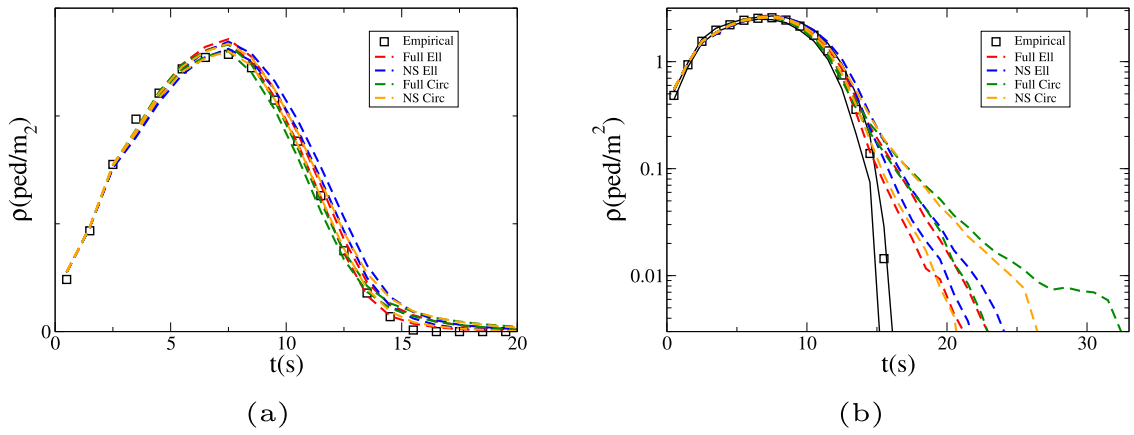

Fig. 16. ρ standard error intervals in different models (dashed lines) compared to the experimental values (squares) for the $\rho_I = 2.5$ ped/m^2 initial condition. (a): linear plot. (b): logarithmic plot. The standard error interval for the experimental values is shown as continuous black lines in the logarithmic plot. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

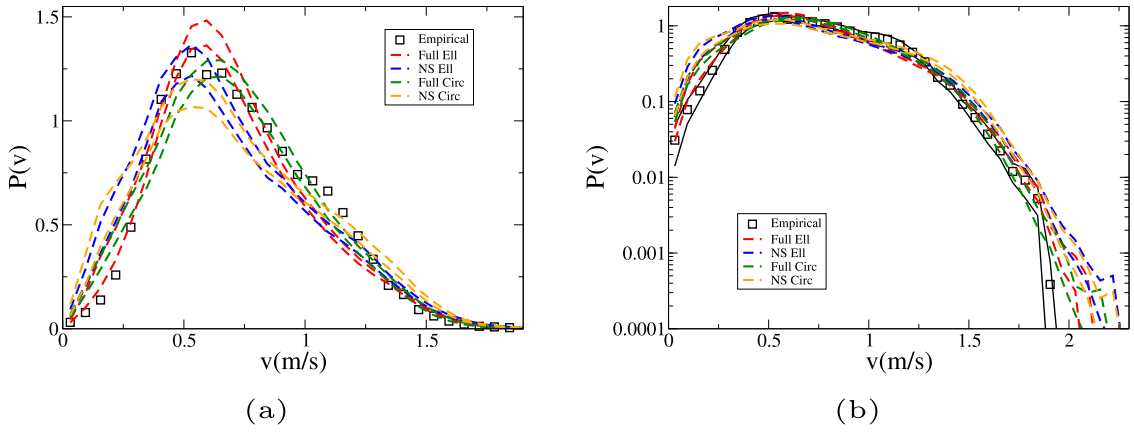

Fig. 17. v standard error intervals in different models (dashed lines) compared to the experimental values (squares) for the $\rho_I = 2.5$ ped/m^2 initial condition. (a): linear plot. (b): logarithmic plot. The standard error interval for the experimental values is shown as continuous black lines in the logarithmic plot. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

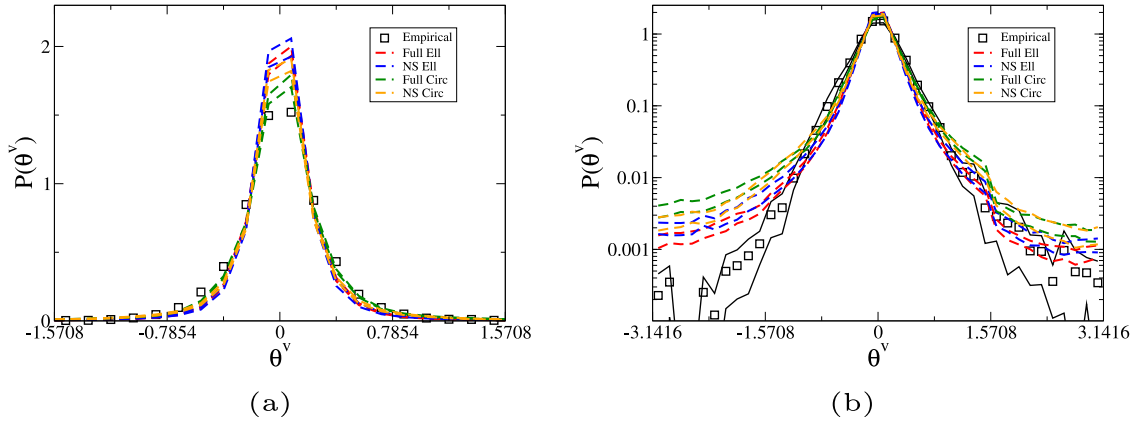

Fig. 18. θ^v standard error intervals in different models (dashed lines) compared to the experimental values (squares) for the $\rho_I = 2.5$ ped/m^2 initial condition. (a): linear plot. (b): logarithmic plot. The standard error interval for the experimental values is shown as continuous black lines in the logarithmic plot. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

we notice a very strong tendency of the Full model to outperform the No-short concerning almost all observables (see Table 2). The only observable on which the No-short model appears to perform better is ϕ^o, on which the models are actually not calibrated. The "Total evaluation measure" on the observables for which the models have been calibrated is different in a significant way, with a p value smaller than 10^{-2} and an effect size close to 1. Quite low p values and large effect sizes are attained concerning almost all the observables, excluding the body orientation θ (although there is a significant difference in $\Delta\theta$, i.e., in the deviation between body and velocity orientation), and the aforementioned ϕ^o. The difference in ρ is less strong than for other observables, not attaining statistical significance, but the No-short

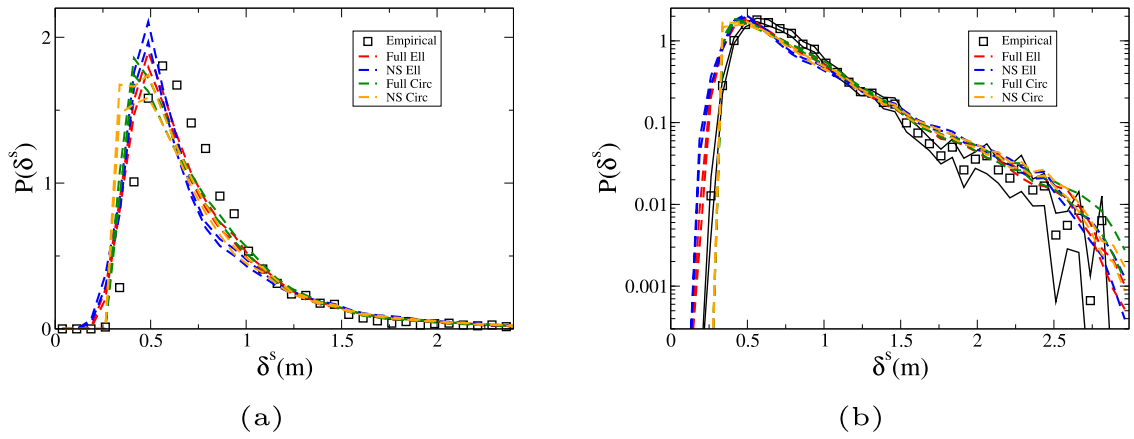

Fig. 19. δ^s standard error intervals in different models (dashed lines) compared to the experimental values (squares) for the $\rho_I = 2.5$ ped/m^2 initial condition. (a): linear plot. (b): logarithmic plot. The standard error interval for the experimental values is shown as continuous black lines in the logarithmic plot. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

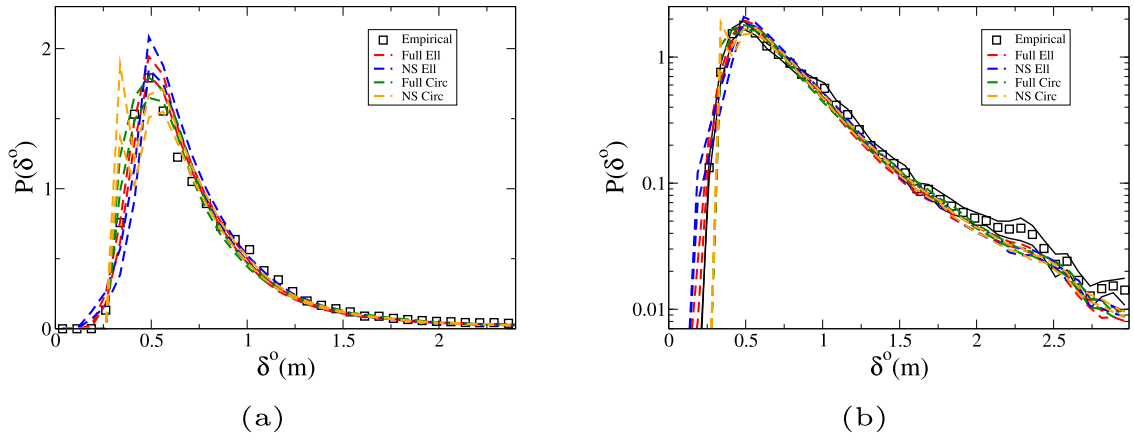

Fig. 20. δ^o standard error intervals in different models (dashed lines) compared to the experimental values (squares) for the $\rho_I = 2.5$ ped/m^2 initial condition. (a): linear plot. (b): logarithmic plot. The standard error interval for the experimental values is shown as continuous black lines in the logarithmic plot. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

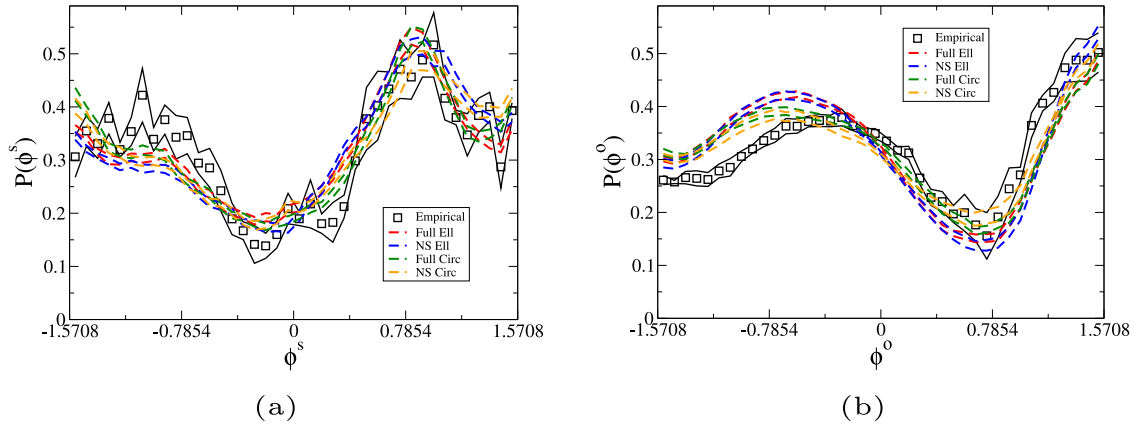

Fig. 21. (a): ϕ^s standard error intervals in different models (dashed lines) compared to the experimental values (squares, standard error interval shown as continuous black lines) for the $\rho_I = 2.5$ ped/m^2 initial condition. (b): ϕ^o standard error intervals in different models (dashed lines) compared to the experimental values (squares, standard error interval shown as continuous black lines). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

model appears to describe poorly the distance between pedestrians, in particular δ^s.

Although the corresponding results are not shown in the manuscript, we verified that there is not a significant change in the difference between the models when calibrated at $\rho_I = 1.5$ ped/m^2 or evaluated at $\rho_I = 2.5$ ped/m^2, the difference being actually a little more pronounced in evaluation, showing that the more complex model does not suffer from over-fitting issues with respect to the simpler one.

It is quite clear, at least from these data, that when an elliptical body is considered, the actual collisions between such elliptical bodies should be predicted by the model.

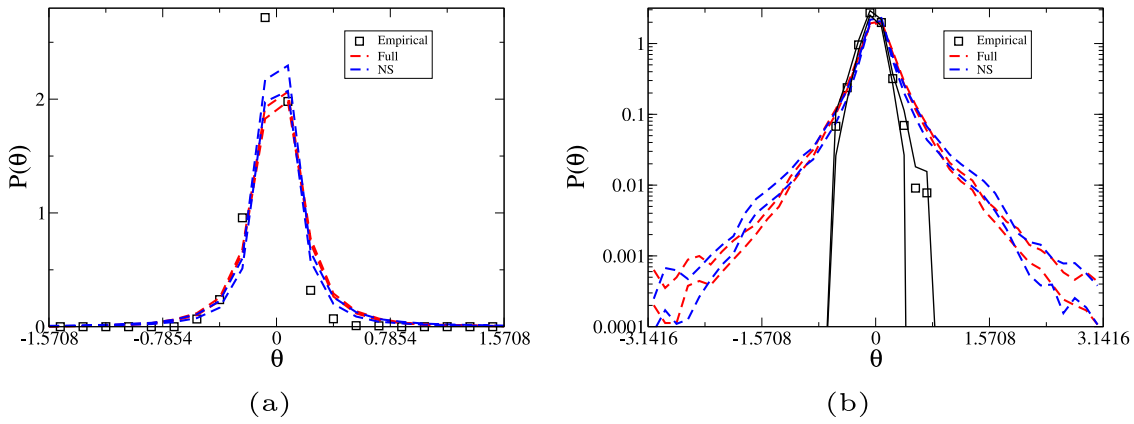

Fig. 22. θ standard error intervals in different models (dashed lines) compared to the experimental values (squares) for the $\rho_I = 2.5$ ped/m^2 initial condition. (a): linear plot. (b): logarithmic plot. The standard error interval for the experimental values is shown as continuous black lines in the logarithmic plot. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

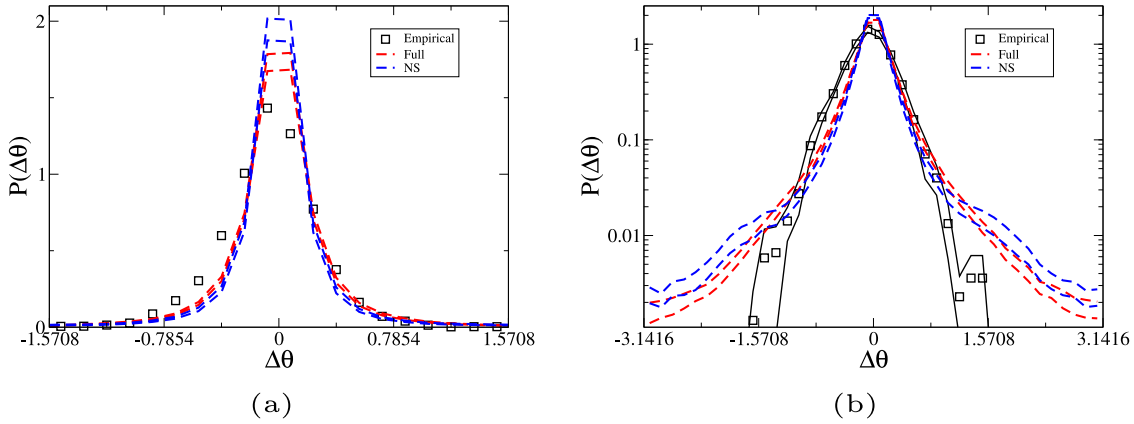

Fig. 23. $\Delta\theta$ standard error intervals in different models (dashed lines) compared to the experimental values (squares) for the $\rho_I = 2.5$ ped/m^2 initial condition. (a): linear plot. (b): logarithmic plot. The standard error interval for the experimental values is shown as continuous black lines in the logarithmic plot. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
EMD metric comparison between Full Elliptical and Full Circular models, evaluation at $\rho_I = 2.5$ ped/m^2.

	Elliptical $\langle\,\rangle \pm \varepsilon(\sigma)$	Circular $\langle\,\rangle \pm \varepsilon(\sigma)$	p	Effect size
No body	$0.0146 \pm 0.0014(0.0067)$	$0.0155 \pm 0.001(0.0049)$	0.61	0.15
ρ	$0.00967 \pm 0.0021(0.01)$	$0.0136 \pm 0.0018(0.0085)$	0.17	0.41
v	$0.0194 \pm 0.0034(0.016)$	$0.0206 \pm 0.0022(0.01)$	0.75	0.095
θ^v	$0.00887 \pm 0.00053(0.0025)$	$0.00818 \pm 0.00082(0.0039)$	0.48	0.21
δ^s	$0.0189 \pm 0.0011(0.0054)$	$0.0192 \pm 0.0012(0.0057)$	0.82	0.069
δ^o	$0.0164 \pm 0.0015(0.0074)$	$0.016 \pm 0.0017(0.0081)$	0.87	0.049
ϕ^s	$0.0207 \pm 0.0017(0.0083)$	$0.0198 \pm 0.0025(0.012)$	0.76	0.09
ϕ^o	$0.0424 \pm 0.0032(0.015)$	$0.0367 \pm 0.0027(0.013)$	0.17	0.41

Table 4
EMD metric comparison between Full Elliptical and No-Short Circular (CP) models, evaluation at $\rho_I = 2.5$ ped/m^2.

	Full $\langle\,\rangle \pm \varepsilon(\sigma)$	CP $\langle\,\rangle \pm \varepsilon(\sigma)$	p	Effect size
No body	$0.0146 \pm 0.0014(0.0067)$	$0.0192 \pm 0.0017(0.0081)$	0.041	0.62
ρ	$0.00967 \pm 0.0021(0.01)$	$0.0139 \pm 0.0025(0.012)$	0.21	0.37
v	$0.0194 \pm 0.0034(0.016)$	$0.0289 \pm 0.004(0.019)$	0.072	0.54
θ^v	$0.00887 \pm 0.00053(0.0025)$	$0.00937 \pm 0.0007(0.0033)$	0.57	0.17
δ^s	$0.0189 \pm 0.0011(0.0054)$	$0.0239 \pm 0.0012(0.0056)$	0.003	0.93
δ^o	$0.0164 \pm 0.0015(0.0074)$	$0.0201 \pm 0.0022(0.011)$	0.17	0.41
ϕ^s	$0.0207 \pm 0.0017(0.0083)$	$0.0283 \pm 0.0021(0.01)$	0.0079	0.82
ϕ^o	$0.0424 \pm 0.0032(0.015)$	$0.0291 \pm 0.0036(0.017)$	0.0081	0.82

Similar considerations can be made when we compare the two models that do not describe body orientation (the Full and No-short Circular models), although the difference in this case is less clear (see Table 5). We indeed have, for the overall evaluation fitness, $p = 0.067$ and a 0.55 effect size for the EMD metric. The differences between the two circular models are actually stronger when the calibration phase is considered (analysis not included in the manuscript). These results suggest that over-fitting issues may be present when comparing the Full and No-short Circular models.

Table 5
EMD metric comparison between Full Circular and No-short Circular models, evaluation at $\rho_I = 2.5$ ped/m^2.

	Full $\langle\,\rangle \pm \varepsilon(\sigma)$	No-short $\langle\,\rangle \pm \varepsilon(\sigma)$	p	Effect size
No body	$0.0155 \pm 0.001(0.0049)$	$0.0192 \pm 0.0017(0.0081)$	0.067	0.55
ρ	$0.0136 \pm 0.0018(0.0085)$	$0.0139 \pm 0.0025(0.012)$	0.93	0.028
v	$0.0206 \pm 0.0022(0.01)$	$0.0289 \pm 0.004(0.019)$	0.073	0.54
θ^v	$0.00818 \pm 0.00082(0.0039)$	$0.00937 \pm 0.0007(0.0033)$	0.27	0.33
δ^s	$0.0192 \pm 0.0012(0.0057)$	$0.0239 \pm 0.0012(0.0056)$	0.0071	0.83
δ^o	$0.016 \pm 0.0017(0.0081)$	$0.0201 \pm 0.0022(0.011)$	0.15	0.44
ϕ^s	$0.0198 \pm 0.0025(0.012)$	$0.0283 \pm 0.0021(0.01)$	0.012	0.77
ϕ^o	$0.0367 \pm 0.0027(0.013)$	$0.0291 \pm 0.0036(0.017)$	0.098	0.5

The, somehow intuitive but nevertheless important, conclusion is that predicting the details of short-time collisions is more important when a non-symmetrical body shape is considered.

6.2.2. *Elliptical vs. Circular*

A straightforward comparison between the newly proposed Full Elliptical model and the CP (Circular No-short) model (see Table 4),

shows that the new model significantly outperforms the old one in overall EMD fitness ($p = 0.041$, effect size 0.62) and in general performs better on most observables, although a comparison with calibration results (not shown in the manuscript), in which the difference between the two models is more pronounced, suggests that the more complex model (i.e., the model with more parameters) may suffer more (as expected) of over-fitting.

Nevertheless, the problem of comparing Circular models to Elliptical ones is more subtle than just referring to their calibration and evaluation performances. Circular models are not able to reproduce orientation dynamics, so for any kind of situation in which such dynamics is relevant, Elliptical ones (or other models including body orientations) should be used. As we have seen in Zanlungo et al. (2022) that the deviation of $\Delta\theta$ from 0 increases with density, this is probably necessary at very high densities. Furthermore, intuitive considerations and previous studies (Yamamoto et al., 2019) suggest this dynamics to be very important in narrow spaces.

A different issue is understanding the regimes in which using body orientation models is necessary. For the studied dynamics, we observe (see Table 3) that when the results are analysed using the EMD metric, that concerns the position of the distribution bulks, very little difference is observed between the Full Elliptical and Full Circular models (it should nevertheless be noted that the Elliptical model has been calibrated to reproduce also body orientation, and thus it is "trying to solve" a more complex dynamics).

Nevertheless, the situation is different when we focus on how models reproduce the tails of the distributions. Fig. 16(b) shows a comparison between the empirical $\rho(t)$ at $\rho_I = 2.5$ ped/m^2 and the evaluation results in all models (for each model the standard error intervals obtained averaging over all solutions are shown). It can be noticed that all models exhibit a tendency to have a fatter tail than the empirical distribution. This is due to some pedestrians getting stuck in an "impasse" near the top-right corner of the crossing area in Fig. 1(b) (bottom-right of Figs. A.24 (a), (b), (c)). This problem could probably be avoided by using a better path-finding behaviour. Nevertheless, the occurrence of such situations are probably the closest the virtual pedestrians ever get to an actual "clogged" situation in our environment. It appears that the Full Elliptical model is better than the Full Circular model in dealing with such situations. The use of the Full Elliptical model could thus be necessary at higher densities/narrower environments than the ones analysed in this work.

Another interesting point can be raised referring to Fig. 18(b), showing a comparison for the pdf of the θ^v observable, namely the only directional observable that may be reproduced by Circular models. We see again that all models present fatter tails than the empirical distribution, but this effect is much stronger in the Circular models. Again, the Full Elliptical model seems better at describing rare events. These results are in agreement with the discussion of empirical results reported by Zanlungo et al. (2022), in which we suggested that in describing rare, high local density events, body orientation could result to be important.

As an overall consideration concerning model comparison, we may say that although the fundamental properties of the dynamics are already well-described by the CP model (refer to the graphs of Figs. 16–23), including the short range interaction leads to a, although reduced, quantitative and qualitative improvement. It is thus natural, also from a purely theoretical viewpoint, to introduce such dynamics by taking explicitly into account the actual shape of the human body. The proposed approximation, using 2D ellipses, indeed leads to an improvement in the description of rare events (tails), for which probably a detailed description of torso movements is needed.

Nevertheless, from a practical viewpoint, at least concerning the description of cross-flows up to 2–3 ped/m^2, the improvement in the description of the dynamics could be too marginal compared to the increase in computational cost (for a quantitative and qualitative improvement in the description of the cross-flow it would probably be more important to improve path-finding or other forms of strategic choices that have not been considered in this work).

Models that take into account, at least in an approximate way, the actual shape of the human body, are known to better describe the behaviour in narrow environments (see for example Yamamoto et al., 2019, Chraibi et al., 2010). Our results suggest that further experiments and studies are necessary to assess their relevance for high density crossing scenarios in medium to large environments as the one studied in this work (i.e., environments whose minimum linear scale is quite larger than the human body; e.g., a corridor as opposed to a door/exit).

6.2.3. Limitations and possible improvements

Arguably, explicit body orientation dynamics will be needed also from a practical viewpoint at higher densities and in more complex (and narrower) environments. To this end, it may be useful to analyse some of the shortcomings of the proposed model. As can be observed in Figs. 22 and 23, the proposed model does not include the strong tendency of human pedestrians to "not turn their back" to their goal. Although this is partially achieved by some solutions (Figs. 7 (a), 8 (a), 13 (a), 14 (a)), this happens to the detriment of the description of the bulk distributions (Figs. 7 (b), 8 (b), 13 (b), 14 (b)). The proposed model is linear in θ angles (i.e., given by a quadratic potential), and thus an improvement could be using a steeper potential for large θ.

Another shortcoming is the description of low δ values. Figs. 12(a) and 19 show that Elliptical models underestimate the minimum value assumed by δ^s, although they describe in a more realistic way the initial slope of the pdf (on the other hand, the Full Elliptical model appears to have a good qualitative description of low δ^o values, see Figs. 12(b) and 20). This is probably due to the limits of the 2D ellipse approximation. Although, as described in Appendix C.1.3, we tried to include in the basic CP model a tendency "not to overlap steps" to account for the fact that our model completely ignores the complex 3D dynamics of leg movement, a more detailed description of limb shape and movement could be needed in future (this effect is intuitively particularly important for elliptical models, since when people are walking in a file and with their torso perpendicular to walking direction, if the movement of the limbs are ignored, they may get closer due to the smaller extension of their body in the walking direction).

Finally, the detailed description of ϕ observables could be improved (see Figs. 9, 15, 21). We nevertheless remind that we decided not to calibrate this behaviour in the current work, and confined ourselves to observing how it was "naturally" emerging in the models.

It may thus be useful to better study models describing body orientation, and possibly develop more simple and efficient ones to decrease issues of calibration and computational cost. Nevertheless, at least for the problem and the range of density studied in this work, it does not appear that body orientation is *crucial* to reproduce the observed dynamics. From a practical viewpoint, for example, it may be that a quantitative and qualitative improvement in the description of the cross-flow may be attained by improving path-finding (or other forms of strategic choices) aspects that have not been considered in this work.

7. Conclusions

In this work, we tried to reproduce the empirical results of a set of controlled experiments concerning the behaviour of a human crowd in a cross-flow scenario by using a hierarchy of models, which differ in the details of the body shape (using a disk-shaped body vs a more realistic elliptical shape) and in how collision avoidance is performed (using only information regarding "centre of mass" distance and velocity, or actually introducing body shape information). We verified that the most detailed model (i.e., using body shape information and an elliptical body) outperforms in a significant way the simplest one (using only centre of mass distance, velocity, and disk-shaped bodies). Furthermore, we observed that if elliptical bodies are introduced without using

such information in collision avoidance, the performance of the model is relatively poor. Nevertheless, the difference between the different models is relevant only in describing the "tails" of the observable distributions, suggesting that the more complex models could be of practical use only in the description of high density settings. Although we did not calibrate our model in order to reproduce the "stripe formation" self-organising pattern, we verified that it naturally emerges, at least at a qualitative level, in all models. A more quantitative study is left for a future work.

CRediT authorship contribution statement

Francesco Zanlungo: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Claudio Feliciani:** Writing – original draft, Data curation, Conceptualization. **Zeynep Yücel:** Writing – review & editing, Writing – original draft, Visualization. **Katsuhiro Nishinari:** Writing – review & editing, Writing – original draft, Funding acquisition. **Takayuki Kanda:** Writing – review & editing, Writing – original draft, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data used in this work has been uploaded to the Julich Pedestrian Dynamics Data Archive, as stated in the manuscript.

Acknowledgements

This research work was in part supported by: JSPS, Japan KAKENHI Grant Number 18H04121, JSPS, Japan KAKENHI Grant Number 20K14992, JST-Mirai Program, Japan Grant Number JPMJMI20D1, JST Moonshot, Japan Grant Number JPMJMS2011. We thank Kazuki Maeda of The University of Fukuchiyama for a valuable suggestion on how to improve the quality of graphical functions in our shared code.

Funding

This research work was in part supported by: JSPS, Japan KAKENHI Grant Number 18H04121, JSPS, Japan KAKENHI Grant Number 20K14992, JST-Mirai Program, Japan Grant Number JPMJMI20D1. The funding sources had no involvement in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Ethics statement

Experimental procedures were approved by the Ethical Commission of The University of Tokyo (approval number 19-274) and conform with the Declaration of Helsinki. Participants received clear information on the nature of the research, methods employed and disclosed data. After a briefing, they gave written permission for participation and data acquisition and received a remuneration in accordance to the university's policies.

Appendix A. Stripe formation in experiments and models

In Fig. A.24 we show some situations in which pedestrians are clearly self-organised in stripes. Fig. A.24(a) shows a frame from the experiments (colours are edited to show clearly which pedestrians belong to which stream, and to suggest the presence of the "clusters"), while Fig. A.24(b) and Fig. A.24(c) show, respectively, frames from simulations using the Full Circular and Full Elliptical models. Although all frames are obtained using the same initial density condition $\rho_I = 1.5$ ped/m^2, they are extracted from different experiment repetitions. The full simulation video corresponding to the results of Fig. A.24(b) may be found at Zanlungo (2022a) while that corresponding to the results of Fig. A.24(c) may be found at Zanlungo (2022b).

Appendix B. Notation

Let us start with some general definitions that will prove to be useful.

The cardinality of a set S is defined as $\#(S)$. Vectors are denoted by boldface, such as $\mathbf{a}$. The standard Euclidean inner (scalar) product between vectors $\mathbf{a}$ and $\mathbf{b}$ is given by

$$(\mathbf{a}, \mathbf{b}), \tag{B.1}$$

and the Euclidean norm by

$$a = (\mathbf{a}, \mathbf{a})^{\frac{1}{2}}. \tag{B.2}$$

As we always deal with 2D vectors, we may consider their vector product as a scalar (given by its projection on the right-handed normal to the plane). Namely, given an arbitrary right handed frame, we define

$$\langle \mathbf{a}, \mathbf{b} \rangle \equiv (\mathbf{a} \times \mathbf{b})_z = a_x b_y - a_y b_x. \tag{B.3}$$

If we are only interested in the absolute value of $\langle \mathbf{a}, \mathbf{b} \rangle$, we need not to worry about the right-handedness of the frame.

In our analysis of empirical observables, we considered only pedestrians that, at a given time t, were located inside a tracking area, which is defined as an L times L square ($L = 3.4$ m), adding a 0.2 m border on the sides of the pink area in Fig. 1, to take into account the absence of hard walls (as discussed by Zanlungo et al. (2022)) and as a consequence has an area $A = L^2$. For each pedestrian i located inside the tracking area (at a given tracking time t) we define the position vector $\mathbf{r}_i$ as the vectorial distance from the origin (e.g., located in the centre of the crossing area). The distance between two pedestrians is defined as

$$\mathbf{r}_{i,j} = \mathbf{r}_j - \mathbf{r}_i. \tag{B.4}$$

At each time t, we also define $f^1(t)$ and $f^2(t)$ as the subset of (tracked) pedestrians belonging, respectively, to flow 1 and 2. Since these sets correspond to tracked pedestrians, they formally depend on tracking time t, although such a dependence is shown only when needed. Their union is the set of tracked pedestrians

$$T(t) = f^1(t) \cup f^2(t). \tag{B.5}$$

Each flow belongs to a corridor, whose axis direction (oriented as the marching direction of the pedestrians, i.e., their goal) is defined through the normalised vector $\mathbf{j}^k$, $k = \{1, 2\}$. We also define the orthogonal normalised vectors as $\mathbf{i}^1 = -\mathbf{j}^2$, $\mathbf{i}^2 = -\mathbf{j}^1$. It should be noted that, through this definition, $\{\mathbf{i}^k, \mathbf{j}^k\}$ are not necessarily right-handed (see Figs. 2, 3).

Furthermore, for each corridor k, we define a reference frame using

$$a_x^k = (\mathbf{a}, \mathbf{i}^k) \quad a_y^k = (\mathbf{a}, \mathbf{j}^k), \tag{B.6}$$

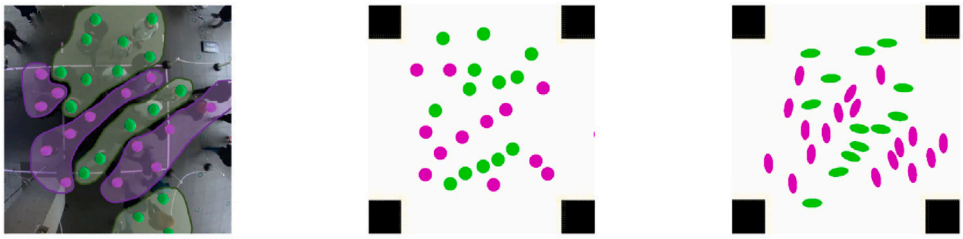

Fig. A.24. Frames showing in a qualitative and visual way stripe formation under the $\rho_I = 1.5$ ped/m^2 initial condition. (a): experiments with subjects, (b) Full Circular model simulation, (c) Full Elliptical model simulation. Pedestrians in green move from top to bottom, those in purple from left to right. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

($a^k_{x,y}$ being the components of an arbitrary vector **a** in frame k) and for each pedestrian i, we define the function

$$F(i) = \begin{cases} 1, & \text{if } i \in f^1 \\ 2, & \text{if } i \in f^2. \end{cases} \tag{B.7}$$

An *empirical probability distribution*[14] of an observable O is computed by defining a bin size

$$\Delta O = \frac{O_{\max} - O_{\min}}{n_b}, \tag{B.8}$$

where n_b is the number of bins, and $O_{\min}$, $O_{\max}$ define the interval in which we study the observable distribution.

The number of observations belonging to a bin j is then defined as

$$n^O_j = \#(\text{ observations } k \text{ with value } O_k \text{ such that } j\Delta O + O_{\min} \le O_k < (j+1)\Delta O + O_{\min}) \tag{B.9}$$

and the empirical probability distribution of O is

$$P(O) = P\left(\left(j + \frac{1}{2}\right)\Delta O + O_{\min}\right) \equiv P_j = \frac{n^O_j}{\sum_l n^O_l}. \tag{B.10}$$

All empirical probability distributions considered in this work are computed by using $n_b = 40$ bins.

For each experimental condition we performed n_e independent repetitions. In graphs, we compare averages and standard errors over all independent repetitions.

Namely, if P^k_j is the value assumed by the jth bin of a given observable with an experimental condition in the kth independent repetition, graphs will show average values and standard errors

$$\langle P_j \rangle \pm \epsilon_j, \tag{B.11}$$

defined according to

$$\langle P_j \rangle = \frac{\sum_{k=1}^{n_e} P^k_j}{n_e}, \tag{B.12}$$

and

$$\langle (P_j)^2 \rangle = \frac{\sum_{k=1}^{n_e} (P^k_j)^2}{n_e}, \qquad \epsilon_j = \sqrt{\frac{\langle (P_j)^2 \rangle - \langle P_j \rangle^2}{n_e - 1}}. \tag{B.13}$$

Appendix C. Detailed description of computational models

C.1. Details in the implementation of the long range module

We start with a detailed description of the long range module, focusing in particular on the differences with the model as introduced Zanlungo et al. (2011). We recall that all the models used in this work are based on the long range module, and thus the properties of the latter are relevant to all models.

The model has been formalised by Zanlungo et al. (2011), although it was introduced in a preliminary version by Zanlungo (2007b,a). Furthermore, while (Zanlungo et al., 2011) dealt with pedestrian motion in a large environment (absence of non-human obstacles), the model was updated to deal with obstacles by Zanlungo et al. (2017b). Moreover, a new term is added in this work to better describe the density/velocity relation in high density crowds. Basically, we may consider the model to consist of 3 terms, the *inter-pedestrian collision prediction term* $\mathbf{F}^p$, the *obstacle collision prediction term* $\mathbf{F}^o$, and the *step overlap term* $\mathbf{F}^s$. The sum of these terms provides an "interaction social force"

$$\mathbf{F}^{long} = \mathbf{F}^p + \mathbf{F}^o + \mathbf{F}^s, \tag{C.1}$$

namely an acceleration that pedestrians perform in order to change their velocity based on their perception of the environment.[15] We recall that according to the Social Force Model paradigm, the pedestrians' decision process determines their acceleration as

$$\ddot{\mathbf{r}} = -k_{v_p}(\dot{\mathbf{r}} - \mathbf{v}_p) + \mathbf{F}^{long}. \tag{C.2}$$

Here $\mathbf{v}_p$ is the preferred velocity of the pedestrians, given by

$$\mathbf{v}_p = v_p \hat{\mathbf{g}}, \tag{C.3}$$

where the preferred speed v_p is assumed to be constant in time for each pedestrian (although it may obviously be different between pedestrians), while the preferred goal direction $\hat{\mathbf{g}}$ may be a function of position and time. k_{v_p} is a parameter of the model and has the dimension of t^{-1}.

C.1.1. Inter-pedestrian collision prediction term

The basic idea of the CP model is the following: for each pedestrian i with position $\mathbf{r}_i$ and velocity $\mathbf{v}_i \equiv \dot{\mathbf{r}}_i$, we define as usual the relative

[14] The term "distribution" is used since observables are theoretically continuous, although from an empirical point of view they are computed over discrete bins. In general, in figures we normalise them in such a way that their integral equals 1, although in EMD computations they are normalised in such a way that the sum over bins equals 1.

[15] In Helbing and Molnar (1995) the Social Force Model was introduced as the sum of terms involving purely physical dynamics (body contact) and terms concerning pedestrian decisional dynamics. Since the action of physical forces obviously depends on pedestrians' masses, the latter were explicitly present in the model. When, as in the current work, the SFM framework is used to describe only decisional dynamics, it is tacitly assumed that pedestrians will apply a force that determines their wanted acceleration, and the term "force" is interchangeable with "acceleration". This could be formalised by multiplying all acceleration terms by the pedestrian's mass, but since we never write equations mixing physical and decisional dynamics, such a formalisation is left implicit. For further details on the relation between "force" and "acceleration" (second order) models, see Adrian et al. (2019).

position and velocity with respect to another pedestrian j respectively as

$$\mathbf{r}_{i,j} = \mathbf{r}_j - \mathbf{r}_i, \tag{C.4}$$

and

$$\mathbf{v}_{i,j} = \mathbf{v}_j - \mathbf{v}_i. \tag{C.5}$$

We define j to be *visible* to i ($j \in V_i^p$) if

$$(\mathbf{v}_i, \mathbf{r}_{i,j}) > 0, \tag{C.6}$$

and *approaching* ($j \in A_i^p$) if

$$(\mathbf{r}_{i,j}, \mathbf{v}_{i,j}) < 0. \tag{C.7}$$

Finally, we define a pedestrian as *colliding* ($j \in C_i^p$) if

$$\left| \left\langle \mathbf{r}_{i,j}, \frac{\mathbf{v}_{i,j}}{v_{i,j}} \right\rangle \right| < d_{ext}. \tag{C.8}$$

This latter condition means that, assuming linear motion, i and j will reach a distance smaller than a threshold d_{ext}, i.e., the maximum distance at which an interaction is expected (not necessarily related to body size). These conditions are slightly different from those used by Zanlungo et al. (2011) and take advantage of the introduction of hard-core potentials of Zanlungo et al. (2017b) (see Eq. (C.13)).

For all pedestrians j satisfying all three conditions we then compute the time $t_{i,j}$ at which the minimum distance between i and j will be reached, assuming both will keep a constant velocity, namely

$$t_{i,j} = -\left(\mathbf{r}_{i,j}, \frac{\mathbf{v}_{i,j}}{v_{i,j}^2} \right). \tag{C.9}$$

We then define the *time of next closest approach to a pedestrian*

$$t_i^p = \min_{j \in (V_i^p \cap A_i^p \cap C_i^p)} t_{i,j}, \tag{C.10}$$

and, again assuming linear motion at constant velocity, we compute, for all *visible* pedestrians j,

$$\mathbf{r}_{i,j}^{min} = \mathbf{r}_j(t) - \mathbf{r}_i(t) + (t_i^p - t)\mathbf{v}_{i,j}, \tag{C.11}$$

i.e., the difference between the (predicted) position of the (moving, i.e., pedestrian) obstacle $\mathbf{r}_j$ and the one of the pedestrian $\mathbf{r}_i$ at t_i^p.

Such "future distance" $\mathbf{r}_{i,j}^{min}$ is then used to define the collision avoidance term of a velocity dependent specification of the social force model as

$$\mathbf{F}_{i,j}^p = -\frac{v_i}{\max(t_i^p, \Delta t)} f(r_{i,j}^{min}) \frac{\mathbf{r}_{i,j}^{min}}{r_{i,j}^{min}}. \tag{C.12}$$

The term $v_i / \max(t_i^p, \Delta t)$ is dimensionally an acceleration and it is introduced assuming that the desired acceleration must be roughly enough for the pedestrian to stop in a time t_i^p, unless such time is smaller than the "reaction time" Δt (the integration step of the behavioural model). As a result, f should be dimensionless, and of order 1.

While in Zanlungo et al. (2011), for historical reasons, the choice of f resembled the one of Helbing and Molnar (1995), just replacing current with future positions, following the analysis of Zanlungo (2007b,a), Zanlungo et al. (2017b) in the current work we use hard core potentials

$$f(r) = \begin{cases} C & \text{if } r \le d_{int}, \\ C \frac{d_{ext} - r}{d_{ext} - d_{int}} & \text{if } d_{int} < r \le d_{ext}, \\ 0 & \text{if } r > d_{ext}. \end{cases} \tag{C.13}$$

C being a dimensionless constant and d_{int}, d_{ext} distance thresholds.

Finally, the interaction force in Eq. (C.1) due to the interaction with the other pedestrians is

$$\mathbf{F}_i^p = \sum_{j \in (V_i^p \cap C_i^p)} \mathbf{F}_{i,j}^p, \tag{C.14}$$

where the sum is performed over all *visible* and *colliding* pedestrians j (i.e., including also non-approaching ones).

C.1.2. Obstacle collision prediction term

Obstacles (walls) are handled using the approach of Zanlungo et al. (2017b). In that work, intended for a robot application, obstacles are expressed as "a cloud of points", and the force generated by such an obstacle is given by an average over all "cloud points". This choice in Zanlungo et al. (2017b) was motivated by the fact that obstacles are perceived by robots as sets of scan points from a sensor (although obviously algorithms to extract higher level information are available). Although here we are dealing with human behaviour, using the same kind of approach appears as a way to express the long range behaviour with respect to obstacles (averaged interaction with the whole obstacle, and not with an explicit collision point).

In detail, we repeat the operations performed above for the interaction of pedestrian j with every other pedestrian in the environment also on a set of "obstacle points".[16] Defining $\mathbf{r}_{i,j}$ as the relative distance between the pedestrian i and the obstacle point j, and assuming the velocity of the obstacle equal to zero, we define j to be *near* ($j \in N_j^o$) if

$$\mathbf{r}_{i,j} < R_{max}, \tag{C.15}$$

on the front ($j \in A_j^o$) if

$$\left(\frac{\mathbf{r}_{i,j}}{r_{i,j}}, \frac{\mathbf{v}_i}{v_i} \right) > \frac{\sqrt{2}}{2}, \tag{C.16}$$

and *colliding* ($j \in C_j^o$) if

$$\left| \left\langle \mathbf{r}_{i,j}, \frac{\mathbf{v}_i}{v_i} \right\rangle \right| < d_{ext}. \tag{C.17}$$

These conditions assure that the collision avoidance process is, although defined as "long range", still of a local nature, i.e., it accounts only for obstacles close in space and collision time. The distance threshold R_{max} was chosen to be equal to 3 m, so that, in general, a pedestrian is interacting with a single obstacle (or very few obstacles) and it is meaningful to use an average of the grid points. The short range model, on the other hand, deals with obstacles as polygons, and computes explicit collision points. The *on the front* condition is stricter than the *visible* one ($j \in V_j^o$), namely Eq. (C.6), to allow pedestrians to be able to walk close to a wall if needed.

For all *near*, *colliding* and *on the front* points, we can compute again

$$t_{i,j} = \left(\mathbf{r}_{i,j}, \frac{\mathbf{v}_i}{v_i^2} \right). \tag{C.18}$$

and define the *time of next closest approach to an obstacle*

$$t_i^o = \min_{j \in (N_j^o \cap A_j^o \cap C_j^o)} t_{i,j}. \tag{C.19}$$

Then, for all *near* and *visible* obstacles (the set of visible obstacles V_i^o is computed in analogy with Eq. (C.6))

$$\mathbf{r}_{i,j}^{min} = \mathbf{r}_j(t) - \mathbf{r}_i(t) - (t_i^p - t)\mathbf{v}_i, \tag{C.20}$$

and

$$\mathbf{F}_{i,j}^o = -\frac{v_i}{\max(t_i^o, \Delta t)} f(r_{i,j}^{min}) \frac{\mathbf{r}_{i,j}^{min}}{r_{i,j}^{min}}. \tag{C.21}$$

Finally, we compute

$$\mathbf{F}_i^o = \frac{1}{\#(N_i^o \cap V_i^o)} \sum_{j \in (N_i^o \cap V_i^o)} \mathbf{F}_{i,j}^o. \tag{C.22}$$

The presence of the threshold R_{max} prevents this average from yielding a very low value due to the sum of many small contributions.

[16] Specifically, we replace walls with a regular grid of points with a 0.05 m lattice step.

C.1.3. Step overlap term

It has been suggested (Curtis and Manocha, 2014) that models based on prediction of collisions are "too efficient", as they typically present too high speeds at high densities. Indeed, the CP model would allow, in the noiseless limit, a crowd of pedestrians with the same goal direction to move with their preferred speed regardless of their densities, or at least up to the maximum density compatible with physical constraints. This is due to the fact that in such a condition, since all pedestrians have the same velocity, no collision is predicted[17] and there is no need of changing velocity.

This is clearly unrealistic, because pedestrians are not 2D rigid bodies. When a uni-directional, single flow motion is considered, the most important correction to the "2D rigid body" assumption amounts to taking into consideration the swing of the legs. As a first order assumption, we proceeded as follows.[18]

To each pedestrian we associate an elliptical[19] "leg swing space", with semi-axes A_l and B_l. A_l is given by half of the pedestrian maximum linear extension (for a circular body with radius R, $A_l = R$; for an elliptical body with semi-axes $A > B$, $A_l = A$), while B_l is a parameter of the model that represents half of the leg swing length. Such an elliptical space is centred at the body centre of mass (elliptical body centre), and oriented in such a way that the B_l axis overlaps with the velocity direction.[20] For each pedestrian pair $\{i, j\}$, and time t, assuming $j \in V_i^p$, the set of visible (i.e., located on i's front) pedestrians, we verify (using again the method based on Foulaadvand and Yarifard, 2013) if their "leg swing space" overlap, and in case they do, we evaluate "how much they overlap" in the following way. We scale both "leg swing spaces" of a linear factor s (i.e., we replace the semi-axes with sA_l and sB_l), and define $0 \le s_{i,j}^f < 1$ as the maximum s such that the two ellipses do not overlap.

Finally, if we add the following extra force term to the interaction term $\mathbf{F}^{long}$ acting on pedestrian i

$$\mathbf{F}_i^s = -\min\left(k_s\left(\sum_{j\in(V_i^p)'}\left(1-s_{i,j}^f\right)\right),(\Delta t)^{-1}\right)\mathbf{v}, \tag{C.23}$$

where $(V_i^p)'$ is the set of visible pedestrians for which the overlapping condition is satisfied.

This term is a velocity-dependent "friction" term (parameter k_s having dimension t^{-1}) which is caused by the overlapping of the pedestrians' "leg swing spaces". The minimum with respect to the inverse of the time step is taken to be sure that, after a decision step, the pedestrian is slowed down, but not pushed into the opposite direction (see Section 4.3 for a discussion of the meaning of the time step Δt in this model).

C.2. "Path-finding" and randomness

All proposed models include the "preferred velocity" term $-k_{v_p}(\dot{\mathbf{r}} - \mathbf{v}_p)$, so it is important to specify how the preferred velocity direction is computed.

The geometry of the crossing corridor is very simple, so that, as long as pedestrians are located inside the borders of their own corridor, their preferred velocity should be directed as their walking direction $\mathbf{j}^{F(i)}$, which in the following discussion we may assume to be given by $(0, 1)$ (i.e., we describe it according to the simulated pedestrian's own frame of reference as defined in Eq. (B.6)). In this model, we include the "unpredictability" of human behaviour by defining the goal vector as

$$\hat{\mathbf{g}} = \frac{\mathbf{g}}{g}, \quad \mathbf{g} = (N^1(\sigma_g), 1 + N^2(\sigma_g)), \tag{C.24}$$

$N^m(\sigma)$ being independent realisations of a Gaussian random variable with zero mean and σ standard deviation.

As a result of collision avoidance, simulated pedestrians may be temporally located in the other stream's corridor, and this situation has to be managed by goal assignment. We verified that using a "minimum path" approach caused overcrowding around the corridors' corners, which did not appear to be qualitatively similar to human behaviour. We thus adopted the following "ad hoc" procedure. Let us identify the positions of the crossing areas' 4 corners in each simulated pedestrian's own frame of reference with coordinates

$$\begin{aligned} c_x &= \pm\frac{L}{2}, \\ c_y &= \pm\frac{L}{2}. \end{aligned} \tag{C.25}$$

Here, L is the width corridor, and the origin of the reference frame is the centre of the crossing area. Since in each pedestrian's frame the walking direction is identified with the y axis, we assume the pedestrians to be located out of their own corridor if the following applies for the pedestrian's position $\mathbf{r} = (x, y)$:

$$|x| \ge \frac{L}{2} - 0.5\text{m}, \; -\frac{L}{2} \le y \le \frac{L}{2}. \tag{C.26}$$

The 0.5 m term concerning the x position is introduced assuming that pedestrians prefer to keep some distance from the corridor's wall.

When such a condition is attained, pedestrians are assumed to re-enter the corridor by walking towards the corner identified by the position

$$\bar{\mathbf{c}} = \left(\text{sign}(x)\left(\frac{L}{2} - 0.5\right), \frac{L}{2}\right) \equiv (\bar{c}_x, \bar{c}_y). \tag{C.27}$$

Nevertheless, to avoid overcrowding around the corner, we assign a stronger weight to the x direction (walking back into the correct corridor) with respect to the y direction (walking towards the goal) defining the goal vector as

$$\hat{\mathbf{g}} = \frac{\mathbf{g}}{g}, \quad \mathbf{g} = (2(\bar{c}_x - x + N^1(\sigma_g)), \bar{c}_y - y + N^2(\sigma_g)). \tag{C.28}$$

As stated above, this is definitely an "ad hoc" solution to the problem, but we are not aware of any quantitative studies regarding pedestrian path choices under such conditions. Nevertheless, we believe that, given the purpose of this work, it was important to adopt a relatively simple approach common to all computational models. For practical applications and future works, it may be interesting to try a machine learning approach aimed also at reproducing pedestrian path choices.

C.3. Parameters

The "long range" model (with the addition of the step overlap term), is characterised by the following parameters (assuming the body size of the pedestrian as given by ellipse semi-axes $A \ge B$, with $A = B$ representing the circular case): C (dimensionless), d_{int} (length), d_{ext} (length) (Eq. (C.13)), the mean μ_v and standard deviation σ_v of the preferred velocity v_p distribution (assumed to be normally distributed), k_{v_p} (Eq. (C.2), inverse of time), the threshold under which walls are considered in collision avoidance R_{max} (length), the size of the leg swing space B_l (length), k_s (Eq. (C.23), inverse of time), and the "goal noise" standard deviation σ_g (Eq. (C.24)). In this work we assumed, following (Zanlungo et al., 2011) $k_{v_p} = 1.52\ \text{s}^{-1}$, and fixed $R_{max} = 2$ m, leaving the 8 parameters C, d_{int}, d_{ext}, μ_v, σ_v, B_l, k_s and σ_g to be calibrated.

A further parameter is the integration step Δt, that was fixed to 0.05 s. The equations governing the behavioural dynamics (collision avoidance) of pedestrians are written as differential equations but

[17] More precisely, $t_i^p \to \infty$.

[18] The proposed term can hardly be regarded as "long range", but we include it in this section because it is considered as part of the "basic" model.

[19] A rectangular shape could be used to decrease computational time.

[20] This means that, for an elliptical body, the body minor axis B is not necessarily aligned with B_l, since we allow body direction and velocity direction to differ.

Table C.6
Summary of parameters.

Parameter	Definition	Calibrated	Model	FE value	FC value
A (major axis)	Section 4.4	No	All	0.225 m	0.15 m
B (minor axis)	Section 4.4	No	All	0.1 m	0.15 m
C	Eq. (C.13)	Yes	All	0.126	7.599
d_{int}	Eq. (C.13)	Yes	All	0.5498 m	0.09154 m
d_{ext}	Eq. (C.13)	Yes	All	1.2084 m	0.11154 m
μ_v	Mean of vel. distr.	Yes	All	1.764 m/s	1.729 m/s
σ_v	Standard deviation of vel. distr.	Yes	All	0.1133 m/s	0.1454 m/s
k_{v_p}	Eq. (C.2)	No	All	1.52 s^{-1}	1.52 s^{-1}
R_{max}	Eq. (C.15)	No	All	2 m	2 m
B_I	Appendix C.1.3	Yes	All	0.5478 m	0.9807 m
k_s	Eq. (C.23)	Yes	All	5.194 s^{-1}	1.98 s^{-1}
σ_g	Eq. (C.24)	Yes	All	0.296	0.
Δt	Sec 4.3	No	All	0.05 s	0.05 s
k_θ^v	Eq. (2)	Yes	Elliptical	0.6329 s^{-1}	
k_ω	Eq. (3)	Yes	Elliptical	8.135 s^{-1}	
k_θ^ω	Eq. (3)	Yes	Elliptical	5.867 s^{-1}	
γ	Eq. (7)	Yes	Full	6.747	10
τ_1	Eq. (4)	Yes	Full	0.7759 s	1 s
τ_2	Eq. (4)	Yes	Full	1.7759 s	1.2681 s

implemented as difference equations, by discretising through an Euler scheme. The integration step is common to all behavioural modules (including the short range ones described below), and assumes that pedestrians take decisions at regular time steps and act based on them. Using a Euler scheme with a time step comparable to human reaction times is thus not an issue, as long as, as described in Section 4.3, physical dynamics is resolved at a finer scale.

The Elliptical models use also the following parameters related to angle dynamics: k_θ^v (inverse of time), k_ω (inverse of time) and k_θ^ω (inverse of squared time). Finally the "short range" module introduces 3 parameters: γ (dimensionless), τ_1 (time) and τ_2 (time) (Eqs. (2)–(8)).

In Table C.6 we summarise all parameters, specifying where they are defined, which models use them, if they are calibrated or not, and providing values assumed in two high performing solutions of the Full Elliptical and Full Circular models, while with the code shared in Zanlungo (2022c) we provide all parameters for each model and GA solution. We would like to stress that these solutions are calibrated to specifically reproduce the pdfs of the observables that we defined to analyse our experiments in a cross-flow setting, and do not necessarily generalise to different and more complex scenarios. Anyway, the different solutions and models provided may hopefully be a starting point for such generalisations.

C.4. Constraints

All terms explained until now are considered to be the results of pedestrian decisional processes, and as such they should provide velocities and accelerations that are compatible with realistic pedestrian behaviour. For this reason, in case after computing the discrete equations, the values of $|\dot{\mathbf{v}}|$ and $\dot{\omega}$ are higher than thresholds $a_{max} = 5$ m s^{-2} and $\dot{\omega}_{max} = 2\pi$ s^{-2}, the final acceleration and angular accelerations are scaled (i.e., preserving their direction and sign) to such maximum values. In an equivalent way, thresholds $v_{max} = 3$ m s^{-1} and $\omega_{max} = 2\pi$ s^{-1} are applied to speed and angular velocity.

Appendix D. Statistical analysis of observables

In this appendix we provide a definition of some recurrent statistical terms used in this work.

We assume that O is an observable whose value is measured over n repetitions, assumed statistically independent, of an experiment. Here, both the term "observable" and "experiment" are used in a very broad sense. "Observable" may refer not only to those defined in Section 3, but also to the values assumed by the evaluation metric defined in Appendix E. By "experiment", we may refer to the actual experiments with human pedestrians, to the simulations using the same initial conditions of a given experiment repetition, or even to an "experiment with pedestrians-simulation" pair in which the simulation uses the same initial conditions as the experiment.

The "n independent repetitions" are assumed to share some common experimental condition (e.g., the initial density ρ_I) or a set of common experimental conditions (e.g., initial density and model specification used in simulation). Given these assumptions, we use some statistical tools to understand what our experiments tell us about the expected behaviour under these conditions and about its statistical variation. All these tools are quite basic and are intended to provide some quantitative tool for comparison, without any claim to "prove" or "disprove" anything (although terms like "statistically significant" are sometimes used in our manuscript when the evidence appears to be compelling, or on the other hand, when we want to stress the lack of such evidence).

Given these premises, writing the value assumed by O in the kth repetition as O_k we define the mean of the observable as

$$\langle O \rangle = \frac{\sum_{k=1}^{n} O_k}{n}, \tag{D.1}$$

its standard deviation as

$$\sigma = \sqrt{\frac{\sum_{k=1}^{n} O_k^2}{n} - \langle O \rangle^2}, \tag{D.2}$$

and its standard error as

$$\varepsilon = \frac{\sigma}{\sqrt{n-1}}. \tag{D.3}$$

When comparing results between two different conditions (e.g., when comparing the evaluation metric values attained by different models), which we may call condition A and B, a possible rule of thumb is to say that O assumes a different value under the two conditions provided that

$$|\langle O \rangle^A - \langle O \rangle^B| \gg \sqrt{2}\max(\varepsilon^A, \varepsilon^B). \tag{D.4}$$

This rule of thumb is obviously related to the t test p values reported in the text. When comparing results between 2 different conditions, corresponding to n^A and n^B repetitions, if we define the difference between their averages as

$$\delta O = \langle O \rangle^A - \langle O \rangle^B, \tag{D.5}$$

and the degree of freedom as

$$\gamma = n^A + n^B - 2, \tag{D.6}$$

and assume, as a null hypothesis, both samples to follow the same Normal distribution, the random variable

$$Z = \frac{\delta O}{\sqrt{\left(\frac{\sigma_A^2}{n^B} + \frac{\sigma_B^2}{n^A}\right)\frac{n^A+n^B}{\gamma}}} \tag{D.7}$$

will follow a t distribution (Ash, 2011), and the probability of having such random variable assuming an absolute value larger than the observed one can be computed numerically (Press et al., 1988) and provides the reported p value.

Let us see how this relates to the rule of thumb for standard errors. Let us assume, as usual in our comparisons

$$n^A = n^B = n \gg 1. \tag{D.8}$$

We have

$$Z \approx \frac{\delta O}{\sqrt{\varepsilon_A^2 + \varepsilon_B^2}}, \tag{D.9}$$

and, according to the central limit theorem (Ash, 2011), the distribution of this random variable can be approximated by a Normal one with average 0 and variance 1 (whose cumulative function we may name $\Phi(x)$), and thus the corresponding p value will be approximated (F being the observed value of Z) by

$$p \approx 2\int_F^{+\infty} \Phi(x)\,\mathrm{d}x, \quad F \gg 1 \Rightarrow p \ll 1, \tag{D.10}$$

justifying the rule of thumb Eq. (D.4).

The value of the ratio in Eq. (D.7) increases with the number of observations, but it may be useful to use also an estimator that does not depend on such number. In this work, we also use effect size, defined as

$$\begin{aligned} \delta &= \frac{\langle O\rangle^A - \langle O\rangle^B}{\overline{\sigma}}, \\ \overline{\sigma} &= \sqrt{\frac{(n^A-1)(\sigma^A)^2 + (n^B-1)(\sigma^B)^2}{n^A+n^B-2}}. \end{aligned} \tag{D.11}$$

While a p-value tells us about the significance of the statistical difference between two distributions, the difference may often be so small that it can be verified only if a large amount of data are collected. But if we have also $\delta \approx 1$, then the two distributions are different enough to be distinguished also using a relatively reduced amount of data.

Appendix E. Earth Mover's Distance

The Earth Mover's Distance (EMD, Cohen and Guibasm, 1999) may be used to compare probability distributions. The EMD process can be visualised as filling holes by moving piles of dirt. Assume that P and Q denote two pdfs, and that a proper metric, named *ground distance* is defined to measure the distance between the bins i and j.[21] Suppose also that a flow $f(i,j)$ is applied to morph P to Q, namely a (signed) quantity is subtracted from $P(i)$ and added to $Q(j)$ in the process of making the P and Q distributions more similar.

In this contest, EMD can be formulated and solved essentially as a transportation problem. Namely, EMD aims at finding the amount of flow f that minimises the overall cost of morphing P to Q. Explicitly, the work required to morph P to Q (or vice versa) *given an explicit flow* f is,

$$\sum_i \sum_j f(i,j)d(i,j), \tag{E.1}$$

where $f(i,j)$ and $d(i,j)$ are respectively, the flow and the ground distance between $P(i)$ and $Q(j)$. By solving for the optimal flow and normalising it with the total flow, EMD is described as,

$$EMD(P,Q) = \frac{\min_f \sum_i \sum_j f(i,j)d(i,j)}{\sum_i \sum_j f(i,j)}. \tag{E.2}$$

The normalisation operation in the above equation yields the average distance travelled by unit weight under the optimal flow.

When computing the EMD between two discrete histograms defined on the same array of bins $i = 0, \ldots, N-1$, the following algorithm may be used

$$\begin{aligned} \mathrm{EMD}_0 &= 0, \\ \mathrm{EMD}_{i+1} &= Q_i - P_i + \mathrm{EMD}_i, \\ \mathrm{EMD} &= \frac{\sum_{i=1}^N |\mathrm{EMD}_i|}{N}. \end{aligned} \tag{E.3}$$

In the last step, we normalise dividing by the number of bins N.

In this work, the comparison is performed between observables averaged over independent repetitions of experiments according to Eq. (B.12).

[21] It is common practice to use the Euclidean distance from i to j as the ground distance. Namely, in the 1D case this reduces to $|i-j|$.

References

Adrian, J., Bode, N., Amos, M., Baratchi, M., Beermann, M., Boltes, M., Corbetta, A., Dezecache, G., Drury, J., Fu, Z., et al., 2019. A glossary for research on human crowd dynamics. Collect. Dyn. 4, 1–13. http://dx.doi.org/10.17815/CD.2019.19.

Adrian, J., Seyfried, A., Sieben, A., 2020. Crowds in front of bottlenecks at entrances from the perspective of physics and social psychology. J. R. Soc. Interface 17 (165), 20190871. http://dx.doi.org/10.1098/rsif.2019.0871.

Alonso-Marroquin, F., Busch, J., Chiew, C., Lozano, C., Ramírez-Gómez, Á., 2014. Simulation of counterflow pedestrian dynamics using spheropolygons. Phys. Rev. E 90 (6), 063305. http://dx.doi.org/10.1103/PhysRevE.90.063305.

Ando, K., Ota, H., Oki, T., 1988. Forecasting the flow of people. Railw. Res. Rev. 45 (8), 8–14, (in Japanese).

Arisona, S.M., Aschwanden, G., Halatsch, J., Wonka, P., 2012. Digital Urban Modeling and Simulation. Vol. 242, Springer, http://dx.doi.org/10.1007/978-3-642-29758-8.

Ash, R.B., 2011. Statistical Inference: A Concise Course. Courier Corporation.

Bode, N.W., Holl, S., Mehner, W., Seyfried, A., 2015. Disentangling the impact of social groups on response times and movement dynamics in evacuations. PLoS One 10 (3), e0121227. http://dx.doi.org/10.1371/journal.pone.0121227.

Boltes, M., Seyfried, A., 2013. Collecting pedestrian trajectories. Neurocomputing 100, 127–133. http://dx.doi.org/10.1016/j.neucom.2012.01.036.

Boltes, M., Seyfried, A., Steffen, B., Schadschneider, A., 2010. Automatic extraction of pedestrian trajectories from video recordings. In: Pedestrian and Evacuation Dynamics 2008. Springer, pp. 43–54. http://dx.doi.org/10.1007/978-3-642-04504-2_3.

Cao, S., Lian, L., Chen, M., Yao, M., Song, W., Fang, Z., 2018. Investigation of difference of fundamental diagrams in pedestrian flow. Physica A 506, 661–670. http://dx.doi.org/10.1016/j.physa.2018.04.084.

Cao, S., Seyfried, A., Zhang, J., Holl, S., Song, W., 2017. Fundamental diagrams for multidirectional pedestrian flows. J. Stat. Mech. Theory Exp. 2017 (3), 033404. http://dx.doi.org/10.1088/1742-5468/aa620d.

Chattaraj, U., Seyfried, A., Chakroborty, P., 2009. Comparison of pedestrian fundamental diagram across cultures. Adv. Complex Syst. 12 (03), 393–405. http://dx.doi.org/10.1142/S0219525909002209.

Chraibi, M., Seyfried, A., Schadschneider, A., 2010. Generalized centrifugal-force model for pedestrian dynamics. Phys. Rev. E 82 (4), 046111. http://dx.doi.org/10.1103/PhysRevE.82.046111.

Cividini, J., Appert-Rolland, C., 2013. Wake-mediated interaction between driven particles crossing a perpendicular flow. J. Stat. Mech. Theory Exp. 2013 (07), P07015. http://dx.doi.org/10.1088/1742-5468/2013/07/P07015.

Cividini, J., Appert-Rolland, C., Hilhorst, H.-J., 2013a. Diagonal patterns and chevron effect in intersecting traffic flows. Europhys. Lett. 102 (2), 20002. http://dx.doi.org/10.1209/0295-5075/102/20002.

Cividini, J., Hilhorst, H., Appert-Rolland, C., 2013b. Crossing pedestrian traffic flows, the diagonal stripe pattern, and the chevron effect. J. Phys. A 46 (34), 345002. http://dx.doi.org/10.1088/1751-8113/46/34/345002.

Cohen, S., Guibasm, L., 1999. The earth mover's distance under transformation sets. In: Proceedings of IEEE International Conference on Computer Vision. Vol. 2, IEEE, pp. 1076–1083. http://dx.doi.org/10.1109/ICCV.1999.790393.

Comai, S., Costa, S., Mastrolembo Ventura, S., Vassena, G., Tagliabue, L., Simeone, D., Bertuzzi, E., Scurati, G., Ferrise, F., Ciribini, A., 2020. Indoor mobile mapping system and crowd simulation to support school reopening because of COVID-19: A case study. In: Proceedings of GeoInformation for Disaster Management Conference. Vol. 44, pp. 29–36. http://dx.doi.org/10.5194/isprs-archives-XLIV-3-W1-2020-29-2020.

Crociani, L., Vizzari, G., Yanagisawa, D., Nishinari, K., Bandini, S., 2016. Route choice in pedestrian simulation: Design and evaluation of a model based on empirical observations. Intell. Artif. 10 (2), 163–182. http://dx.doi.org/10.3233/IA-160102.

Curtis, S., Manocha, D., 2014. Pedestrian simulation using geometric reasoning in velocity space. In: Pedestrian and Evacuation Dynamics 2012. Springer, pp. 875–890. http://dx.doi.org/10.1007/978-3-319-02447-9_73.

Digital Human Laboratory, 1992. Anthropometric dimension database. https://www.airc.aist.go.jp/dhrt/91-92/data/list.html, (in Japanese).

Donev, A., Torquato, S., Stillinger, F.H., 2005. Neighbor list collision-driven molecular dynamics simulation for nonspherical hard particles: II. Applications to ellipses and ellipsoids. J. Comput. Phys. 202 (2), 765–793. http://dx.doi.org/10.1016/j.jcp.2004.08.025.

Echeverría-Huarte, I., Zuriguel, I., Hidalgo, R.C., 2020. Pedestrian evacuation simulation in the presence of an obstacle using self-propelled spherocylinders. Phys. Rev. E 102 (1), 012907. http://dx.doi.org/10.1103/PhysRevE.102.012907.

Feliciani, C., Murakami, H., Nishinari, K., 2018. A universal function for capacity of bidirectional pedestrian streams: Filling the gaps in the literature. PLoS One 13 (12), e0208496. http://dx.doi.org/10.1371/journal.pone.0208496.

Feliciani, C., Murakami, H., Nishinari, K., Perpendicular crossflow experiment with partial information on body orientation, http://dx.doi.org/10.34735/ped.2019.2.

Feliciani, C., Murakami, H., Shimura, K., Nishinari, K., 2020a. Efficiently informing crowds–experiments and simulations on route choice and decision making in pedestrian crowds with wheelchair users. Transp. Res. C 114, 484–503. http://dx.doi.org/10.1016/j.trc.2020.02.019.

Feliciani, C., Nishinari, K., 2016. Empirical analysis of the lane formation process in bidirectional pedestrian flow. Phys. Rev. E 94 (3), 032304. http://dx.doi.org/10.1103/PhysRevE.94.032304.

Feliciani, C., Nishinari, K., Estimation of pedestrian crowds' properties using commercial tablets and smartphones, Transp. B: Transp. Dyn. http://dx.doi.org/10.1080/21680566.2018.1517061.

Feliciani, C., Zuriguel, I., Garcimartín, A., Maza, D., Nishinari, K., 2020b. Systematic experimental investigation of the obstacle effect during non-competitive and extremely competitive evacuations. Sci. Rep. 10 (1), 1–20. http://dx.doi.org/10.1038/s41598-020-72733-w.

Foulaadvand, M.E., Yarifard, M., 2013. Two-dimensional system of hard ellipses: A molecular dynamics study. Phys. Rev. E 88 (5), 052504. http://dx.doi.org/10.1103/PhysRevE.88.052504.

Fujita, A., Feliciani, C., Yanagisawa, D., Nishinari, K., 2019. Traffic flow in a crowd of pedestrians walking at different speeds. Phys. Rev. E 99 (6), 062307. http://dx.doi.org/10.1103/PhysRevE.99.062307.

Geoerg, P., Schumann, J., Holl, S., Hofmann, A., 2019. The influence of wheelchair users on movement in a bottleneck and a corridor. J. Adv. Transp. http://dx.doi.org/10.1155/2019/9717208.

Helbing, D., Buzna, L., Johansson, A., Werner, T., 2005. Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. Transp. Sci. 39 (1), 1–24. http://dx.doi.org/10.1287/trsc.1040.0108.

Helbing, D., Molnar, P., 1995. Social force model for pedestrian dynamics. Phys. Rev. E 51 (5), 4282. http://dx.doi.org/10.1103/PhysRevE.51.4282.

Heliövaara, S., Korhonen, T., Hostikka, S., Ehtamo, H., 2012. Counterflow model for agent-based simulation of crowd dynamics. Build. Environ. 48, 89–100. http://dx.doi.org/10.1016/j.buildenv.2011.08.020.

Hidalgo, R.C., Parisi, D.R., Zuriguel, I., 2017. Simulating competitive egress of noncircular pedestrians. Phys. Rev. E 95 (4), 042319. http://dx.doi.org/10.1016/j.buildenv.2011.08.020.

Hittmeir, S., Ranetbauer, H., Schmeiser, C., Wolfram, M.T., 2016. Derivation and analysis of continuum models for crossing pedestrian traffic. http://dx.doi.org/10.48550/arXiv.1612.07582, arXiv:1612.07582.

Hoy, G., Morrow, E., Shalaby, A., 2016. Use of agent-based crowd simulation to investigate the performance of large-scale intermodal facilities: Case study of Union station in Toronto, Ontario, Canada. Transp. Res. Rec. 2540 (1), 20–29. http://dx.doi.org/10.3141/2540-03.

International Organization for Standardization, 2020. ISO 20414: Fire safety engineering – verification and validation protocol for building fire evacuation models.

Johansson, A., Helbing, D., Shukla, P.K., 2007. Specification of the social force pedestrian model by evolutionary adjustment to video tracking data. Adv. Complex Syst. 10 (supp02), 271–288. http://dx.doi.org/10.1142/S0219525907001355.

Langston, P.A., Masling, R., Asmar, B.N., 2006. Crowd dynamics discrete element multicircle model. Saf. Sci. 44 (5), 395–417. http://dx.doi.org/10.1016/j.ssci.2005.11.007.

Lovreglio, R., Ronchi, E., Kinsey, M.J., 2020. An online survey of pedestrian evacuation model usage and users. Fire Technol. 56 (3), 1133–1153. http://dx.doi.org/10.1007/s10694-019-00923-8.

Menge CrowdSim, 2014. Comparing pedestrian models. https://www.youtube.com/watch?v=LTYFRY-SqRE.

Mullick, P., Fontaine, S., Appert-Rolland, C., Olivier, A.-H., Warren, W.H., Pettré, J., 2022. Analysis of emergent patterns in crossing flows of pedestrians reveals an invariant of 'stripe' formation in human data. PLoS Comput. Biol. 18 (6), e1010210. http://dx.doi.org/10.1371/journal.pcbi.1010210.

Murakami, H., Feliciani, C., Nishiyama, Y., Nishinari, K., 2021. Mutual anticipation can contribute to self-organization in human crowds. Sci. Adv. 7 (12), eabe7758. http://dx.doi.org/10.1126/sciadv.abe7758.

Nagao, K., Yanagisawa, D., Nishinari, K., 2018. Estimation of crowd density applying wavelet transform and machine learning. Physica A 510, 145–163. http://dx.doi.org/10.1016/j.physa.2018.06.078.

Naka, Y., 1977. Mechanism of cross passenger flow - study on complicated passenger flow in railway station (Part I). Trans. Arch. Inst. Jpn 258, 93–102. http://dx.doi.org/10.3130/aijsaxx.258.0_93, (in Japanese).

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1988. Numerical Recipes in C. Cambridge university press Cambridge.

Reynolds, C.W., 1987. Flocks, herds and schools: A distributed behavioral model. In: Proceedings of the Annual Conference on Computer Graphics and Interactive Techniques. pp. 25–34. http://dx.doi.org/10.1145/37401.37406.

Rokko High School, 2017. Rokko high school sport festival 2017: Marching parade. https://youtu.be/Sd5xV5RtF1A?t=610.

Seyfried, A., Passon, O., Steffen, B., Boltes, M., Rupprecht, T., Klingsch, W., 2009. New insights into pedestrian flow through bottlenecks. Transp. Sci. 43 (3), 395–406. http://dx.doi.org/10.1287/trsc.1090.0263.

Totzeck, C., 2019. An anisotropic interaction model with collision avoidance. http://dx.doi.org/10.3934/krm.2020044, arXiv:1912.04234.

Van Den Berg, J., Snape, J., Guy, S.J., Manocha, D., 2011. Reciprocal collision avoidance with acceleration-velocity obstacles. In: Proceedings of IEEE International Conference on Robotics and Automation. IEEE, pp. 3475–3482. http://dx.doi.org/10.1109/ICRA.2011.5980408.

Von Krüchten, C., Schadschneider, A., 2017. Empirical study on social groups in pedestrian evacuation dynamics. Physica A 475, 129–141. http://dx.doi.org/10.1016/j.physa.2017.02.004.

Willems, J., Corbetta, A., Menkovski, V., Toschi, F., 2020. Pedestrian orientation dynamics from high-fidelity measurements. Sci. Rep. 10 (1), 1–10. http://dx.doi.org/10.1038/s41598-020-68287-6.

Yajima, S., Yoshii, K., Sumino, Y., 2020. Aversion of pedestrians to face-to-face situations eases crowding. J. Phys. Soc. Japan 89 (7), 074003. http://dx.doi.org/10.7566/JPSJ.89.074003.

Yamamoto, H., Yanagisawa, D., Feliciani, C., Nishinari, K., 2019. Body-rotation behavior of pedestrians for collision avoidance in passing and cross flow. Transp. Res. B 122, 486–510. http://dx.doi.org/10.1016/j.trb.2019.03.008.

Ye, R., Fang, Z., Lian, L., Wang, Q., Zeng, G., Cao, S., Zhang, J., Song, W., 2021. Traffic dynamics of uni-and bidirectional pedestrian flows including dyad social groups in a ring-shaped corridor. J. Stat. Mech. Theory Exp. 2021 (2), 023406. http://dx.doi.org/10.1088/1742-5468/abdc1a.

Zanlungo, F., 2007a. A collision-avoiding mechanism based on a theory of mind. Adv. Complex Syst. 10 (supp02), 363–371. http://dx.doi.org/10.1142/S0219525907001410.

Zanlungo, F., 2007b. Microscopic dynamics of artificial life systems. (Ph.D. thesis). Università di Bologna.

Zanlungo, F., 2022a. Simulation of pedestrian crossing using full circular model. (Accessed 2022-10-12), https://youtu.be/HevTBCcdmEo.

Zanlungo, F., 2022b. Simulation of pedestrian crossing using full elliptical model. (Accessed 2022-10-12), https://youtu.be/VNLXmvvExNw.

Zanlungo, F., 2022c. Simulation of the model introduced in macroscopic and microscopic dynamics of a pedestrian cross flow: Part-II modelling. (Accessed 2022-10-12), https://github.com/zanlungo/Macroscopic-and-microscopic-dynamics-of-a-pedestrian-cross-flow-Part-II-modelling.

Zanlungo, F., Brščić, D., Kanda, T., 2015. Spatial-size scaling of pedestrian groups under growing density conditions. Phys. Rev. E 91 (6), 062810. http://dx.doi.org/10.1103/PhysRevE.91.062810.

Zanlungo, F., Crociani, L., Yücel, Z., Kanda, T., 2020. The effect of social groups on the dynamics of bi-directional pedestrian flow: a numerical study. In: Traffic and Granular Flow 2019. Springer, pp. 307–313. http://dx.doi.org/10.1007/978-3-030-55973-1_38.

Zanlungo, F., Feliciani, C., Yücel, Z., Nishinari, K., Kanda, T., Macroscopic and microscopic dynamics of a pedestrian cross-flow: Part I, experimental analysis. To be published in Safety Science.

Zanlungo, F., Ikeda, T., Kanda, T., 2011. Social force model with explicit collision prediction. Europhys. Lett. 93 (6), 68005. http://dx.doi.org/10.1209/0295-5075/93/68005.

Zanlungo, F., Ikeda, T., Kanda, T., 2012. A microscopic social norm model to obtain realistic macroscopic velocity and density pedestrian distributions. PLoS One 7 (12), e50720. http://dx.doi.org/10.1371/journal.pone.0050720.

Zanlungo, F., Ikeda, T., Kanda, T., 2014. Potential for the dynamics of pedestrians in a socially interacting group. Phys. Rev. E 89 (1), 012811. http://dx.doi.org/10.1103/PhysRevE.89.012811.

Zanlungo, F., Kanda, T., 2015. A mesoscopic model for the effect of density on pedestrian group dynamics. Europhys. Lett. 111 (3), 38007. http://dx.doi.org/10.1209/0295-5075/111/38007.

Zanlungo, F., Yücel, Z., Brščić, D., Kanda, T., Hagita, N., 2017a. Intrinsic group behaviour: Dependence of pedestrian dyad dynamics on principal social and personal features. PLoS One 12 (11), e0187253. http://dx.doi.org/10.1371/journal.pone.0187253.

Zanlungo, F., Yücel, Z., Ferreri, F., Even, J., Saiki, L.Y.M., Kanda, T., 2017b. Social group motion in robots. In: Proceedings of International Conference on Social Robotics. Springer, pp. 474–484. http://dx.doi.org/10.1007/978-3-319-70022-9_47.

Zanlungo, F., Yücel, Z., Kanda, T., 2019. Intrinsic group behaviour II: On the dependence of triad spatial dynamics on social and personal features; and on the effect of social interaction on small group dynamics. PLoS One 14 (12), e0225704. http://dx.doi.org/10.1371/journal.pone.0225704.

Zhang, J., Klingsch, W., Schadschneider, A., Seyfried, A., 2012. Ordering in bidirectional pedestrian flows and its influence on the fundamental diagram. J. Stat. Mech. Theory Exp. 2012 (02), P02002. http://dx.doi.org/10.1088/1742-5468/2012/02/P02002.