

## Stochastic fluctuations of diluted pedestrian dynamics along curved paths

Geert G. M. van der Vleuten,<sup>1</sup> Federico Toschi <sup>1,2,3</sup> Wil Schilders,<sup>4</sup> and Alessandro Corbetta <sup>1,2,\*</sup>

<sup>1</sup>Department of Applied Physics and Science Education, Eindhoven University of Technology, Eindhoven 5600MB, The Netherlands

<sup>2</sup>Eindhoven Artificial Intelligence System Institute, Eindhoven University of Technology, Eindhoven 5600MB, The Netherlands

<sup>3</sup>CNR-IAC, Rome I-00185, Italy

<sup>4</sup>Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven 5600MB, The Netherlands



(Received 21 June 2023; accepted 29 November 2023; published 5 January 2024)

As we walk towards our destinations, our trajectories are constantly influenced by the presence of obstacles and infrastructural elements; even in the absence of crowding our paths are often curved. Since the early 2000s pedestrian dynamics have been extensively studied, aiming at quantitative models with both fundamental and technological relevance. Walking kinematics along straight paths have been experimentally investigated and quantitatively modeled in the diluted limit (i.e., in absence of pedestrian-pedestrian interactions). It is natural to expect that models for straight paths may be an accurate approximations of the dynamics even for paths with curvature radii much larger than the size of a single person. Conversely, as paths curvature increase one may expect larger and larger deviations. As no clear experimental consensus has been reached yet in the literature, here we accurately and systematically investigate the effect of paths curvature on diluted pedestrian dynamics. Thanks to a extensive and highly accurate set of real-life measurements campaign, we derive a Langevin-like social-force model quantitatively compatible with both averages and fluctuations of the walking dynamics. Leveraging on the differential geometric notion of covariant derivative, we generalize previous work by some of the authors, effectively casting a Langevin social-force model for the straight walking dynamics in a curved geometric setting. We deem this the necessary first step to understand and model the more general and ubiquitous case of pedestrians following curved paths in the presence of crowd traffic.

DOI: [10.1103/PhysRevE.109.014605](https://doi.org/10.1103/PhysRevE.109.014605)

### I. INTRODUCTION

As we walk towards our destinations, indoor or in open spaces, we typically prefer to follow the most direct (typically straight) path. Yet obstacles, infrastructural elements, or crowd traffic [1,2] make our *preferred paths* unavoidably *curved* (cf. Fig. 1). Additionally, trajectories invariably exhibit *fluctuations* associated with sway and intersubject variability.

Since the early 2000s, pedestrian kinematics has been extensively investigated experimentally [3,4], and the motion of pedestrians walking along straight paths has been thoroughly analyzed and modeled (e.g., Refs. [2,5–7]). Especially in diluted conditions, i.e., in the absence of pedestrian-pedestrian interactions, these analyses were capable of successfully modeling the dynamics, including the stochastic fluctuations around average motions [8,9]. Swaying motion, shoulder oscillations, and intrasubject and intersubject variabilities play a role in the overall fluctuations which, overall, exhibit a Gaussian structure with possible long-tails due to rare events [8]. In the case of paths having curvature radii much larger than the scale of a single person, we expect models for straight dynamics to hold locally. In fact, under these conditions, paths can be reasonably well approximated as being locally straight. One may thus wonder under which conditions and how the

known model for straight paths can be adapted to generic curved paths. Indeed, as paths curvatures increase, one may expect larger and larger deviations from the assumption of a locally straight dynamics. No experimental consensus has yet been reached on how paths curvature affect pedestrians dynamics. Only few, and partially contradictory, studies are available on the topic. These report anticorrelation between velocity and curvature (with linear [1] or power-law trend [10]) or even an apparent absence of curvature effects [11].

The aim of this work is to understand and to quantitatively model the dynamics of pedestrians walking along curved preferred paths, including averages and stochastic fluctuations, considering a broad spectrum of curvature radii even as small as few pedestrian diameters. We opt to address this outstanding issue restricting to crowd scenarios in the diluted limit. Thus, the environment is the only reason pedestrians opt for curved paths. We deem this setting the necessary first step towards the goal of understanding the generic case in which curved paths appear in combination with and as a consequence of the overall crowd traffic.

Understanding the kinematics of pedestrians is part of a challenging and broad multidisciplinary scientific effort with outstanding societal importance due to implications in crowd management [4] and urban design [12] and sharing deep fundamental challenges connected with active flowing matter and statistical physics [3,13]. One of the main obstacles in fully understanding crowd flows is the inherent technical challenge of obtaining measurements with sufficient

\*a.corbetta@tue.nl

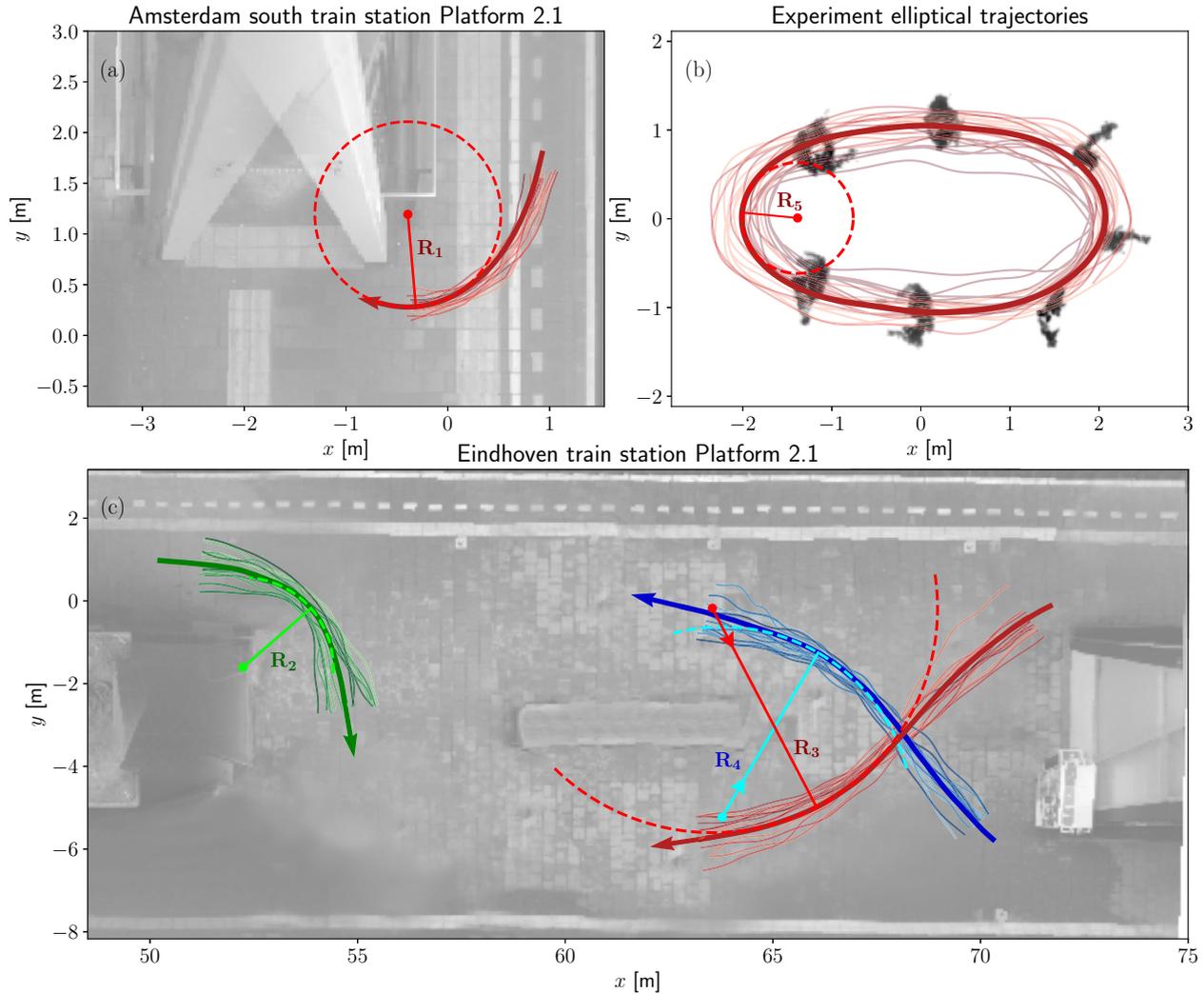


FIG. 1. Few selected trajectories of pedestrians walking along different curved paths in different real-life locations. The three panels provide an overview of the measurement sites employed in this work, respectively in Amsterdam south train station (NL) (a), a laboratory experiment at Eindhoven University of Technology (NL) (where pedestrian were asked to walk along an elliptical path) (b), and at Eindhoven train station (NL) (c). The tangent circles at several points of the preferred-paths (average path across trajectory bundles, cf. Appendix A) are displayed with radii  $R_1 = 0.92$  m,  $R_2 = 2.17$  m,  $R_3 = 5.42$  m,  $R_4 = 4.60$  m, and  $R_5 = 0.63$ , respectively.

spatiotemporal accuracy and statistical resolution, fully capturing the large variability and complexity of pedestrian kinematics. Over the past few years, experimental evidence on pedestrian behavior has been collected mostly in laboratory scenarios, allowing us to probe average behavior, typically studied as a function of the pedestrian density (e.g., Ref. [14]). Average behavior are usually encoded in so-called fundamental diagrams, connecting, e.g., pedestrian density with average velocity or fluxes [15]. Only more recently, accurate, and privacy-respectful large-scale measurements in real-life conditions have become a possibility, either via custom setups developed in research environments [14,16] or via commercial products [17]. Key have been three-dimensional computer vision approaches based on stereoscopic vision or LiDAR-like approaches [18,19]. Data acquisition with a 24/7 schedule in public locations has enabled the collection of highly resolved, high statistics datasets (millions of trajectories), allowing statistical analyses up to rare events and opening new possibilities of model validation [3,20–24].

In this paper, we use high-resolution tracking to collect wide trajectory datasets to investigate the diluted dynamics of pedestrians walking along curved paths. We have performed large-scale data acquisition campaigns in Dutch train stations (Eindhoven, Amsterdam South) and laboratory experiments (in the Eindhoven University of Technology campus, NL). On these bases, we identify the effect of increasing curvature levels on walking velocities, presenting a curvature-velocity fundamental diagram, which we enrich with measurements of the typical fluctuations. This enables us to present a Langevin-like model reproducing quantitatively the complete statistics of position and velocity as curvature changes. Our work generalizes the social forcelike [25] model presented in Ref. [8], which quantitatively reproduces the diluted walking dynamics along straight paths. We effectively cast such a model to a curved geometry: Even in the absence of (social) forces, pedestrians could follow curved trajectories. For this, we employ the language of differential geometry (in particular, through the notion of covariant derivative). On the basis of

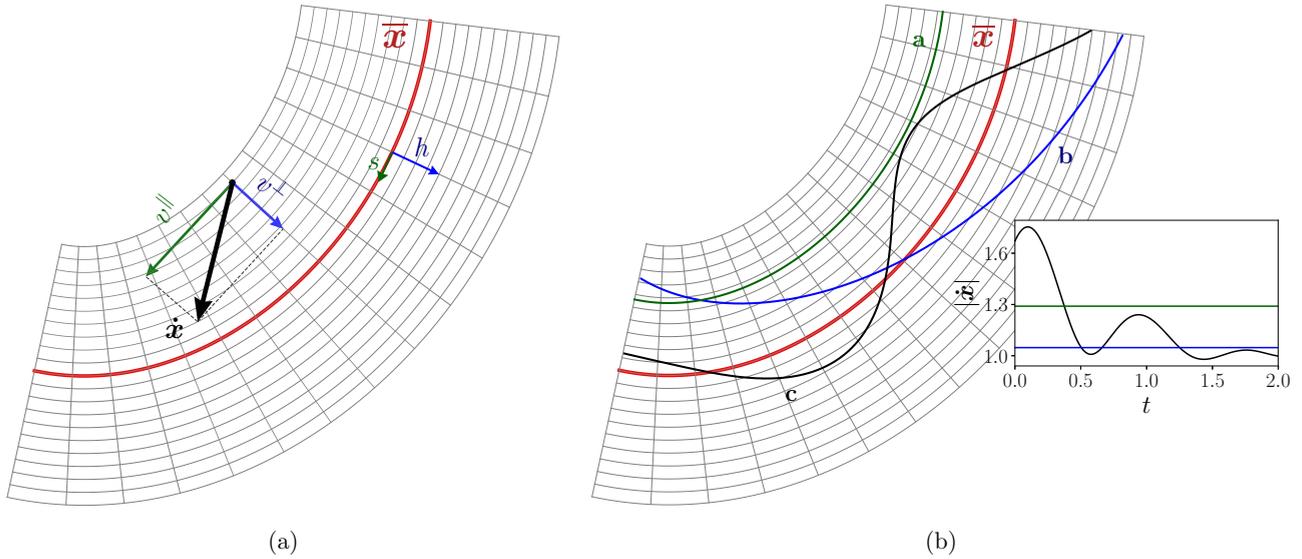


FIG. 2. (a) The average path of the Amsterdam train station measurements, curve  $\bar{x}(s)$ , indicated in red. The tubular neighborhood of  $\bar{x}(s)$  is parameterized by a coordinate frame consisting of coordinate lines equidistant to  $\bar{x}(s)$ , defining the longitudinal direction, and coordinate lines perpendicular to  $\bar{x}(s)$ , defining the transversal direction. Coordinates  $s$  and  $h$  represent the coordinates in the longitudinal and transversal directions, respectively. As indicated, a velocity vector  $V$  can be decomposed in a longitudinal and transversal component ( $v^{\parallel}$  and  $v^{\perp}$ ). (b) Trajectories **a**, in green, and **b**, in blue, are examples of geodesics [solutions of Eq. (12)]. Geodesics with initial velocity parallel to  $\bar{x}(s)$ , such as trajectory **a**, remain parallel to  $\bar{x}(s)$ . The velocity magnitudes of the trajectories are plotted in the inset figure. All geodesics, such as trajectory **a** and **b**, conserve kinetic energy. Trajectory **c**, in black, is a solution of the geodesic equation disturbed by modeling forces [Eq. (17) without noise]. Therefore, the trajectory is forced to oscillate around the path in a damped way. Furthermore, its longitudinal velocity converges to a desired value.

our data analysis, we extend the social-force terms to integrate curvature-dependent effects (with radii down to 0.6 m).

This paper is structured as follows: In Sec. II we introduce the geometric context of tubular neighborhoods of trajectories, central for the forthcoming analyses. In Sec. III, we present the experimental data that we collected for our analyses, together with relevant technical references on data acquisition. Based on the measurements, in Sec. IV, we present a curvature-velocity fundamental diagram, comparing a simple analytic model with measurements. In Sec. V, we present our quantitative Langevin-like model, whose comparison with measurements is reported in the results Sec. VI. A final discussion closes the paper. We opt to postpone most of the technical and formal details connected with differential geometry to the Appendixes.

## II. KINEMATICS OF CURVED WALKING PATHS IN TUBULAR NEIGHBORHOODS

We focus on bundles (i.e., sets containing similarly shaped trajectories) of real-life pedestrian trajectories on the plane  $\mathbf{x} = (x, y)$ :

$$\{t \mapsto \mathbf{x}_\nu(t) = x_\nu(t)\mathbf{e}_x + y_\nu(t)\mathbf{e}_y, \quad \nu = 1, 2, \dots\}, \quad (1)$$

where  $\nu = 1, 2, \dots$ , serves as a trajectory index,  $x_\nu(t)$ ,  $y_\nu(t)$  are the horizontal and vertical components of trajectory  $\nu$  at time  $t$ , and  $(\mathbf{e}_x, \mathbf{e}_y)$  is the (fixed) orthonormal base associated with the  $(x, y)$  coordinates (cf. examples in Fig. 2). These trajectories connect predefined origin and destination, which are separated by, e.g., obstacles or architectural fixtures. The need of bypassing these elements makes typical trajectories,

and thus the whole bundle, nonrectilinear. Due to sway and intersubject variability, trajectories exhibit fluctuations. We analyze such fluctuations in reference with the average path of the bundle,

$$\bar{\mathbf{x}} = \bar{x}(s)\mathbf{e}_x + \bar{y}(s)\mathbf{e}_y,$$

where the variable  $s$  denotes a smooth monotonic parametrization. We identify  $\bar{x}(s)$  with the individual *preferred* path, i.e., the trajectory that each pedestrian aims at following. Examples of such average paths are reported as thick lines in Fig. 1. We postpone the technicalities of the formal definition of the average path,  $\bar{x}(s)$  [Eq. (1)], as a function of the trajectory bundle to Appendix A.

We study fluctuations around  $\bar{x}(s)$  considering its neighborhood. We employ coordinate lines parallel and normal to  $\bar{x}(s)$  [Fig. 2(a)], parameterized by the variables  $s$  and  $h$ , respectively. As mentioned,  $s$  increases as we move along  $\bar{x}(s)$ , whereas  $h$  increases as we move in the orthogonal direction (towards the local curvature center). We name  $(\mathbf{e}_{\parallel}, \mathbf{e}_{\perp})$  the local orthonormal base parallel to these directions. Note that curves defined by  $h = \text{const}$  wrap around  $\bar{x}(s)$  while remaining, in a sense, parallel to it. As such the  $(s, h)$  parametrization of the  $\bar{x}(s)$  neighborhood is usually named *tubular*. For smooth  $\bar{x}(s)$  and limited  $h$ ,  $(s, h)$  uniquely parametrize the tubular neighborhood (e.g., Ref. [26]). We unambiguously decompose velocities,  $\dot{\mathbf{x}} = \dot{x}\mathbf{e}_x + \dot{y}\mathbf{e}_y$ , applied at a point  $\mathbf{x}$  in the neighborhood of  $\bar{x}(s)$ , in a transversal,  $v^{\perp}$ , and a longitudinal component,  $v^{\parallel}$ , respectively perpendicular and parallel to a local coordinate line ( $h = \text{const}$ ). In formulas,

$$\dot{\mathbf{x}} = v^{\parallel}\mathbf{e}_{\parallel} + v^{\perp}\mathbf{e}_{\perp}. \quad (2)$$

Further details on the parametrization of the tubular neighborhood are given in Appendix C.

Our analysis targets kinematic implications on pedestrian trajectories of the curvature of the preferred path. We consider the local curvature of  $\bar{\mathbf{x}}(s)$ ,  $k(s)$ . By definition,  $k(s)$  is the reciprocal of the radius of the circle osculating  $\bar{\mathbf{x}}(s)$  and reads (e.g., Ref. [27])

$$k(s) = \frac{\bar{x}'(s)\bar{y}''(s) - \bar{x}''(s)\bar{y}'(s)}{[(\bar{x}'(s))^2 + (\bar{y}'(s))^2]^{3/2}}, \quad (3)$$

where  $\bar{x}'$  denotes the first derivative of the  $x$  component of  $\bar{\mathbf{x}}(s)$  with respect to  $s$  (the second derivative and operations on the  $y$  component are written accordingly).

### III. MEASUREMENTS

Our study leverages on trajectory datasets acquired via three large-scale pedestrian tracking campaigns all performed in The Netherlands. Our campaigns specifically took place in Amsterdam south train station (AMS), Eindhoven train station (EHV), and on the university campus in Eindhoven (TUE). All our data were acquired in a naturalistic condition (with the exception of the TUE campaign in which pedestrians have been instructed to roughly follow a given path) and in a fully privacy-respectful manner. Commercial or research-grade overhead body-tracking sensors have been employed. All these sensors hinge on similar principles: They measure overhead three-dimensional maps of the scene from which trajectories can be extracted via computer vision approaches (e.g., Refs. [2,17,18]). Since we are interested in the dynamics of undisturbed pedestrians, we consider trajectories in low-density conditions (i.e., in absence of other neighboring pedestrians).

In the following we provide a brief description of the datasets (for technical details about the average paths and the selection procedures, see Appendixes A and B).

#### A. Amsterdam south train station

At this measurement location on platform 2.1, we consider high-resolution data in the vicinity of the staircase [Fig. 1(a)] for the period spanning April 2020 to December 2020 (196 days). Pedestrians arriving by train normally leave the platform via the staircase depicted in the middle. Thus we select some of the many trajectories of pedestrians turning from the platform towards the staircase. The strict selection criteria (Appendix B) result in a selection of 2700 measured trajectories in Amsterdam south train station.

The average path has a gradually increasing curvature and consequently a broad curvature spectrum with a radius of curvature ranging from 5 to 0.9 m. The length of the average path is approximately 2 m.

#### B. Eindhoven train station

At the measurement domain within Eindhoven train station platform 2.1, measurements have been performed between April 2021 and September 2021 with a sample frequency of 10 Hz. We have chosen five winding paths in this train station as preferred paths as these are walked by many pedestrians. Additionally, all paths span wide curvature ranges. A top

view of the platform with three preferred paths is shown in Fig. 1(c).

Totally 2700 measured trajectories are selected in the Eindhoven train station. The average paths in the Eindhoven train station have lengths ranging from 4 to 10 m. The minimum radius of curvature reached by the preferred paths in this station is 2.1 m.

#### C. Eindhoven University of Technology

This measurement campaign is conducted as an experiment at a large public area within the university campus in Eindhoven, the Netherlands, in February 2019. During 1 min, seven participants were asked to walk around two traffic cones, 3 m apart, resulting in elliptical-like trajectories [Fig. 1(b)]. The pedestrians kept their distance to create diluted conditions. The average path has a broad curvature spectrum with a minimal radius of curvature around 0.6 m. The measured trajectories are sampled with a frequency of 30 Hz (further technical information on this experimental setup based on overhead depth sensors are in Ref. [17]).

### IV. CURVATURE-VELOCITY FUNDAMENTAL DIAGRAM AND FLUCTUATIONS

We report here on the effect of the preferred path curvature on the average velocity in the diluted flow limit. We compare a closed-form theoretical model with high statistics measurements. These enable us to derive a fundamental diagram-like relation for average velocity and path curvature.

Consistently with previous research [1,10], we observe that the walking velocity decreases with the curvature of the path.

We assume that body rotation, necessary to adopt a curved trajectory, is the key reason for velocity reduction. Let  $v_{SP}$  denote the velocity pedestrians adopt when walking along straight paths [also straight-path velocity (SPV)]. In each experimental setup, we observed slightly different (average) values of the  $v_{SP}$ , all within the interval [1.10, 1.36] m/s, in agreement with literature velocity measurements in the diluted limit (e.g., Refs. [3,15]). Suppose a pedestrian with body radius  $\delta$  (half body width) walking along a curved path with radius  $R = \frac{1}{k}$  [as in Fig. 3(a)]. We assume the velocity of the body parts following the outer bend to remain equal to the straight-path velocity (i.e., no body part moves faster than the SPV). Considering a rigid body with shoulder line directed toward the curvature center, the body center velocity (BCV),  $v_{BC}$ , satisfies  $v_{BC} < v_{SP}$ , and the following relation among  $v_{SP}$ ,  $v_{BC}$ ,  $\delta$ , and  $R$  holds:

$$\frac{v_{BC}}{R} = \frac{v_{SP}}{R + \delta}. \quad (4)$$

Equation (4) expresses the physical consequence that under our rigid body and shoulder alignment assumptions, and the angular velocity is constant. Linearizing Eq. (4) around  $k = 0$  returns a more familiar fundamental diagram-like expression,

$$v_{BC}(k) = v_{SP}(1 - k\delta). \quad (5)$$

In Fig. 3(b) we compare our model with our experimental measurements. We factor out the context dependency of the velocity by scaling the BCV to the SPV, i.e., we consider the following dimensionless longitudinal velocity at varying

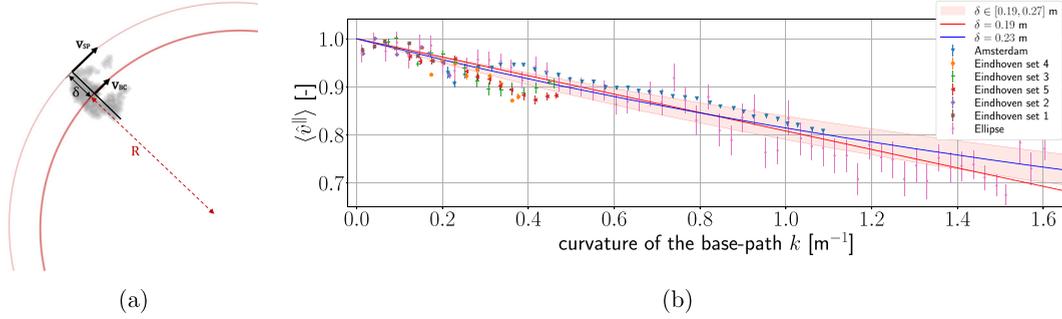


FIG. 3. (a) Sketch of a person following a curved trajectory indicating the body radius,  $\delta$ , radius of curvature,  $R = \frac{1}{k}$ , SPV,  $v_{SP}$ , and the BCV,  $v_{BC}$ . The linearized velocity of the body center behaves according to Eq. (5). (b) The average dimensionless longitudinal velocity  $\langle \hat{v}^{\parallel} \rangle$  as a function of  $k$ , the curvature of the preferred path for the seven different datasets. The error bars indicate the standard deviation. Fits of Eq. (4) of 100 random data partitions are represented by the pink area. The results are compared with the fit of Eq. (4) ( $\delta = 0.23$  m, blue) and the fit of Eq. (5) ( $\delta = 0.19$  m, red).

curvature:

$$\hat{v}^{\parallel}(k) = \left\langle \frac{v_{BC}(k)}{v_{SP}} \right\rangle_k = \frac{\langle v_{BC}(k) \rangle_k}{v_{SP}}, \quad (6)$$

where the average is taken among measurements having the same  $k$  value (where a binning in  $k$  is considered). For each measurement domain, the SPV is determined separately by extrapolating the longitudinal velocity versus curvature relation towards  $k = 0$ . In other terms, we compute  $v_{SP}$  extrapolating the function  $h(k) = \langle v_{BC}(k) \rangle_k$  for  $k \rightarrow 0$ . We report the relation in Eq. (4) with a solid blue line, with body radius  $\delta$  fitted to  $\delta \approx 0.23$  m. The pink area represents a margin of error obtained by fitting Eq. (4) with 100 random partitions of the data, which are compatible with body radii  $\delta \in [19, 27]$  cm consistently with expectations. We report in solid red the linearized relation in Eq. (5) ( $\delta = 0.19$ ). Within the curvature range explored ( $k \in [0, 1.6]$   $\text{m}^{-1}$ ), the complete [Eq. (4)] and linearized relation [Eq. (5)] appear equally compatible with the data. For technical simplicity, in our Langevin-like model proposed in Sec. V we will employ the linearized model.

As curvature change, pedestrians adapt their average velocity almost instantly. In Fig. 4(a), we report as a function of the angular coordinate along the ellipse in the TUE dataset, the

values of average velocity and curvature. Within error bars, the adaptation time of velocity to curvature value appears negligible (approximately 0.3 s). In other terms, it is reasonable to expect that pedestrians move and tune fluctuations in such a way that effectively they manage an instantaneous adaptation of their average velocity.

### A. Velocity fluctuations

We conclude this section reporting on fluctuations beside the curvature-dependent averages [Eq. (6)]. Due to statistics reasons we focus on our richest dataset, AMS. In Fig. 4 we report the probability density function of longitudinal ( $v^{\parallel}$ ) and transversal ( $v^{\perp}$ ) velocity fluctuations. In line with the fundamental diagram [Fig. 3(b)], the means of the longitudinal velocity decrease for higher curvature levels. Compensating for this shift considering

$$v_{\text{SHIFTED}}^{\parallel} := v^{\parallel} - v_{BC}(k) \quad (7)$$

with  $\delta = 19$  cm [cf. Fig. 4(b)], it can be seen that fluctuations in the (shifted) longitudinal velocity are curvature independent and have a Gaussian fluctuation structure with standard deviation  $\sigma_{v^{\parallel}} = 0.19$  m/s. Similarly, fluctuations in transversal velocity do not depend on the curvature and have Gaussian fluctuation,  $\sigma_{v^{\perp}} = 0.15$  m/s. These measurements, after

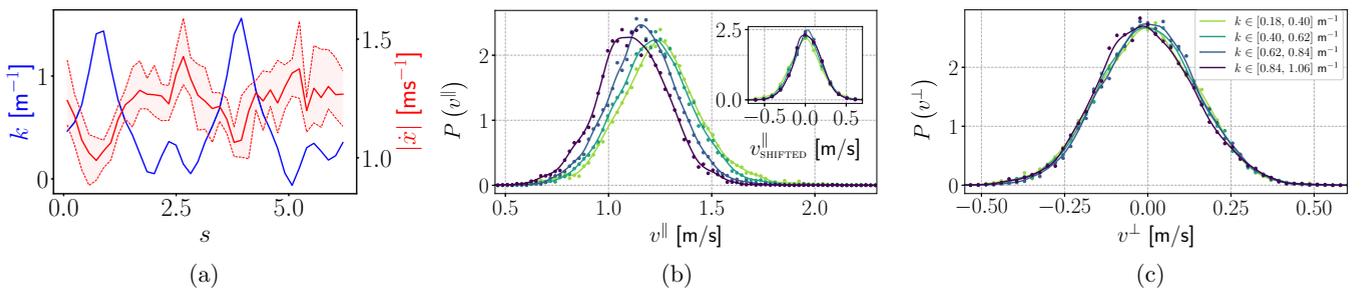


FIG. 4. Curvature ( $k$ ) of the average path and average velocity magnitude ( $|\dot{x}|$ ) for pedestrians walking along an ellipse (TUE dataset). The error bar on the velocity is computed similarly to Fig. 3(b). The  $s$  coordinate parameterizes the phase along the ellipse ( $s \in [0, 2\pi]$ ). High-curvature regions correspond to low-velocity peaks and vice versa. The time offset is approximately 0.3 s and negligible within error bar. (a) Probability distribution functions of the longitudinal velocity,  $v^{\parallel}$ , and of the shifted longitudinal velocity,  $v_{\text{SHIFTED}}^{\parallel}$ , (inset) for four curvature intervals (AMS dataset). The longitudinal velocity fluctuations have different mean. However, the width of the fluctuations is similar for all curvature levels as the inset figure shows. (b) Probability distribution function of the transversal velocity,  $v^{\perp}$ , for the same four curvature intervals. Transversal velocity fluctuations are indistinguishable across different curvature levels.

velocity shifts are compatible with experimental campaigns focusing on straight paths [2,8]. Similarly to Refs. [2,8], the Gaussian behavior of velocity fluctuations will be crucial in modeling perspective, forming the bases to our Langevin-like structure.

## V. LANGEVIN-LIKE MODEL FOR CURVED TUBULAR NEIGHBORHOOD

In this section we show that the walking dynamics around preferred paths can be modeled quantitatively with a Langevin-like model defined on the tubular neighborhood of  $\bar{\mathbf{x}}(s)$ .

### A. Fluctuations around straight paths

The model introduced here extends the Langevin-like model previously proposed by some of the authors and that addresses the case in which  $\bar{\mathbf{x}}(s)$  is a straight trajectory [8]. In Ref. [8], the fluctuating motions of pedestrians have been modeled as a superposition of social forces determining the individual acceleration,  $\ddot{\mathbf{x}}$ . Assuming for simplicity a coordinate system  $(x, y)$  in which  $\bar{\mathbf{x}}(s)$  is the path  $y = 0$ , thus  $x$  identifies the position along  $\bar{\mathbf{x}}(s)$ , and  $y$  is the transversal coordinate (i.e.,  $\bar{\mathbf{x}}(s) = (s, 0)$ ,  $(\mathbf{e}_\parallel, \mathbf{e}_\perp) \equiv (\mathbf{e}_x, \mathbf{e}_y)$ ,  $(\dot{x}, \dot{y}) = (v^\parallel, v^\perp)$ ), individual accelerations read

$$\begin{aligned} \ddot{\mathbf{x}} &= f(\dot{x}, v_{\text{BC}})\mathbf{e}_x + (-2\beta y - 2\mu\dot{y})\mathbf{e}_y + \sigma\dot{\mathbf{W}} \\ &= f(v^\parallel, v_{\text{BC}})\mathbf{e}_\parallel + (-2\beta y - 2\mu v^\perp)\mathbf{e}_\perp + \sigma\dot{\mathbf{W}}. \end{aligned} \quad (8)$$

The previous equation models the following effects:

E1 - self-propulsion along  $\bar{\mathbf{x}}(s)$  driven by  $f(v^\parallel)$ . At first-order Taylor expansion  $f(v^\parallel, v_{\text{BC}})$  is a relaxation term towards a desired walking speed for the body center  $v_{\text{BC}}$ , i.e.,

$$f(v^\parallel, v_{\text{BC}}) = -2\alpha(v^\parallel - v_{\text{BC}}), \quad (9)$$

where  $\alpha$  is inversely proportional to the timescale  $\tau = (2\alpha)^{-1}$  for relaxation towards the desired velocity. Note that this term can be interpreted as an active viscous term with quadratic velocity potential,

$$\Phi_\parallel(v^\parallel, v_{\text{BC}}) = \alpha(v^\parallel - v_{\text{BC}})^2. \quad (10)$$

E2 - transversal confinement in the  $\bar{\mathbf{x}}(s)$  neighborhood, and transversal velocity damping, which is modeled as a damped harmonic oscillator. This is parameterized by a linear stiffness coefficient  $\beta$  and a linear friction coefficient  $\mu$ .

E3 - random noise,  $\dot{\mathbf{W}} := (\dot{W}^\parallel, \dot{W}^\perp)$ , to generate fluctuations and recover randomness in behavior. For simplicity, this is assumed to be  $\delta$  correlated in time, isotropic, with components mutually uncorrelated Gaussian distributed ( $\sigma$  is a scale parameter). This hypothesis quantitatively agrees with the observed fluctuations in terms of correlation structure and probability density of velocities and positions.

Note that in Ref. [2], pairwise interactions to reproduce the statistics of the avoidance behavior have been included in this model.

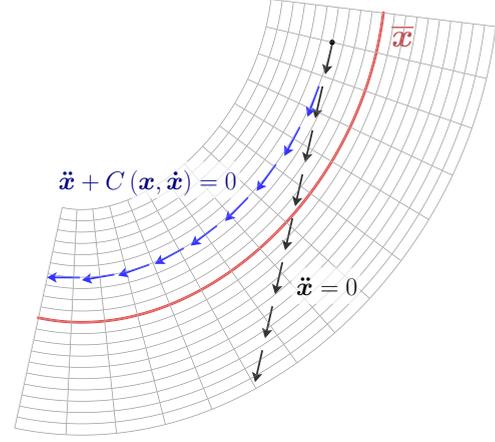


FIG. 5. A tubular coordinate frame around  $\bar{\mathbf{x}}$  with the evolution of a velocity vector respecting  $\ddot{\mathbf{x}} = 0$  and  $\nabla_{\dot{\mathbf{x}}}\dot{\mathbf{x}} = \dot{\mathbf{x}} + C(\mathbf{x}, \dot{\mathbf{x}}) = 0$  in black and blue, respectively.

### B. Parallel dynamics in a tubular neighborhood: Geometric setting

Here we extend model in Eq. (8) to include curvature effect. When pedestrians follow a path with small curvature, we do not expect effects due to curvature: The path appears locally straight. Pedestrians in these conditions would walk following their curved, preferred path. We incorporate this aspect in the left-hand side of the equation of motion (8). Heuristically, we opt to vary the underlying geometry.

First, in absence of forces and noise, Eq. (8) describes a pedestrian conserving their initial momentum:

$$\ddot{\mathbf{x}} = \mathbf{0} \quad \Rightarrow \quad \dot{\mathbf{x}} = \text{const}. \quad (11)$$

This translates into a rectilinear motion (depicted by the black arrows in Fig. 5).

We generalize the left-hand side of Eq. (11), considering broader possibilities of force-free curves (typically addressed as geodesic curves) as solutions of

$$\nabla_{\dot{\mathbf{x}}}\dot{\mathbf{x}} := \ddot{\mathbf{x}} + C(\mathbf{x}, \dot{\mathbf{x}}, \bar{\mathbf{x}}) = \mathbf{0}. \quad (12)$$

Here we adopt the notation  $\nabla_{\dot{\mathbf{x}}}\dot{\mathbf{x}}$  for the covariant derivative of  $\dot{\mathbf{x}}$ . In the field of differential geometry, the covariant derivative is commonly used to express the change of vectors when transporting them in a (curved) geometry [28]. Additionally, the correction term,  $C(\mathbf{x}, \dot{\mathbf{x}}, \bar{\mathbf{x}})$ , is usually expressed by so-called Christoffel symbols of the second kind:  $C(\mathbf{x}, \dot{\mathbf{x}}, \bar{\mathbf{x}}(s)) := \sum_{i,j,k=1,2} \Gamma_{kj}^i \dot{x}^k \dot{x}^j \mathbf{e}_i$  [where the indexed notation satisfies  $(x^1, x^2) := (x, y)$ ,  $(\mathbf{e}_1, \mathbf{e}_2) := (\mathbf{e}_x, \mathbf{e}_y)$ ]. Technical properties of the covariant derivative and Christoffel symbols are postponed to Appendix C.

We aim at a minimal definition of a covariant derivative allowing geodesics that, in heuristic terms, follow the tubular neighborhood [i.e., solutions of Eq. (12), cf. example blue arrows in Fig. 5]. To this purpose, we require geodesics to preserve kinetic energy as well as the velocity in transversal direction ( $v^\perp$ ). The latter implies that geodesics whose initial transversal velocity is zero, i.e., that do not depart from the base path, remain parallel at all times. In formulas, these entail the following properties:

(i) geodesic curves conserve the (Euclidean) kinetic energy, i.e.,

$$\nabla_{\dot{\mathbf{x}}}\dot{\mathbf{x}} = 0 \Rightarrow \frac{d}{dt}\|\dot{\mathbf{x}}\|^2 = \frac{d}{dt}(\dot{x}^2 + \dot{y}^2) = 0. \quad (13)$$

(ii) Geodesic curves preserve the transversal velocity:

$$\nabla_{\dot{\mathbf{x}}}\dot{\mathbf{x}} = 0 \Rightarrow \ddot{h}(t) = 0, \quad \forall t > 0. \quad (14)$$

This implies that geodesic curves initially parallel to  $\bar{\mathbf{x}}(s)$ , i.e., with zero initial orthogonal velocity, remain parallel to  $\bar{\mathbf{x}}(s)$  at all times, i.e.,

$$\begin{cases} v^\perp(t=0) = 0 \\ \nabla_{\dot{\mathbf{x}}}\dot{\mathbf{x}} = 0 \end{cases} \Rightarrow v^\perp(t) = 0, \quad \forall t > 0. \quad (15)$$

This means that if  $\bar{\mathbf{x}}(s)$  is not straight, also geodesics will not be. Two examples of geodesics are shown in Fig. 2(b) (curves **a** and **b**). It can be seen that the properties of remaining parallel and conserving mechanical energy are satisfied, which is ensured by the centripetal-like acceleration  $C(\mathbf{x}, \dot{\mathbf{x}}, \bar{\mathbf{x}}(s))$  (note that this is not a covariant derivative of Levi-Civita type for the Euclidean metric). In our forthcoming simulations we opt to generate trajectories in the physical  $(x, y)$  coordinates. This allows to easily account for forcing terms and possibly generalize our work to include interactions. On the other hand, the correction term remain defined via an implicit system of equations. To prevent this section from becoming needlessly technical, we opt to postpone our derivation of the expression of the correction term following the two hypotheses above as well as their transformation in  $(x, y)$  coordinates in Appendix C.

### C. Pedestrian fluctuations in a tubular neighborhood

To model the fluctuating behavior of pedestrians walking along curved paths, we perturb the force-free dynamics described by Eq. (12), including counterparts of the effects E1–E3. We additionally hypothesize, consistently with the fundamental diagram in Sec. IV, that the body center velocity depends on the instantaneous curvature following Eq. (5). We assume that pedestrians (in absence of stochastic fluctuations) can adjust instantaneously to such velocity as the curvature changes along the path (i.e., when  $\dot{k} \neq 0$ ). This is in agreement with Fig. 4(a) that shows instantaneous velocity adaptations as curvature changes. In principle, the combination of Eq. (9) and Eq. (5) would provide a curvature-dependent propulsion force  $f(v^\parallel, v_{\text{BC}}(k))$ . However, such a propulsion force  $f(v^\parallel, v_{\text{BC}}(k))$  built by bare combination of the two terms would take a time scaling with  $\tau > 0$  to relax to changes in desired velocity due to curvatures changes. This would yield a modeled dynamics that exhibits a systematic nonzero delay to adjust to the body center velocity at a given curvature  $k$ . To model dynamics without such a delay, we correct the propulsion term including a contribution of the curvature time gradient  $\dot{k}$ . This yields our final propulsion term  $\hat{f}(v^\parallel, v_{\text{BC}}(k), \dot{k})$  which reads

$$\begin{aligned} \hat{f}(v^\parallel, v_{\text{BC}}(k), \dot{k}) &= f(v^\parallel, v_{\text{BC}}(k)) - v_{\text{SP}}\delta\dot{k} \\ &= -2\alpha \left[ v^\parallel - v_{\text{SP}}(1 - \delta k) + \frac{v_{\text{SP}}\delta}{2\alpha}\dot{k} \right]. \end{aligned} \quad (16)$$

The component  $-v_{\text{SP}}\delta\dot{k}$  ensures that, in absence of other forces,  $v^\parallel(t) \equiv v_{\text{SP}}(1 - \delta k(t))$  holds at all times and with no time delays. Note that  $\hat{f} \equiv f$  whenever the curvature gradient is zero, e.g., on straight paths or on circular paths with constant radius.

Combining our geodesic flow parallel to the curved preferred path  $\bar{\mathbf{x}}(s)$  [Eq. (12)], the effects E1–E3, and the corrected propulsion term in Eq. (16) yields the following force balance:

$$\nabla_{\dot{\mathbf{x}}}\dot{\mathbf{x}} = \hat{f}(v^\parallel)\mathbf{e}_\parallel + (-2\beta h - 2\mu v^\perp)\mathbf{e}_\perp + \sigma\dot{\mathbf{W}}. \quad (17)$$

An example of a trajectory generated by this model is in Fig. 2(b) (curve **c**), where the modeling forces confine the trajectory around the preferred path  $\bar{\mathbf{x}}(s)$ . In the next section we show that Eq. (17) describes quantitatively the statistics of the fluctuations of pedestrians walking about curved paths. We simulate Eq. (17) via the numerical procedure in Appendix D. This procedure is relatively complex requiring the numerical evaluation of Christoffel symbols. In the idealized case of a circular preferred path, in which center and radius of curvature do not change, Eq. (17) can be written in explicit form in Cartesian coordinates. Although our forthcoming analysis does not employ this simplified setting, we report the related analytic expressions in Appendix E for the readers desiring a simple entry point for implementing our model. We deem this simplified setting a good local approximation whenever the radius and center of curvature change extremely slowly with respect to the fluctuation timescale  $\tau$ .

## VI. RESULTS

In this section we compare the stochastic dynamics modeled by Eq. (17) with experimental data. We focus on trajectories following the curved path at Amsterdam South station (AMS dataset), as it is the richest in amount of trajectories allowing to fully resolve and compare statistical fluctuations. We consider the SPV,  $v_{\text{SP}}$ , and body size radius  $\delta$  determined in Appendix F. We estimate the scale parameters ( $\alpha$ ,  $\beta$ ,  $\mu$ , and  $\sigma$ ) by considering Langevin potentials in longitudinal velocity (shifted as in Sec. IV),

$$\Phi_{v_{\text{SHIFTED}}^\parallel} \sim -\log \mathbb{P}(v_{\text{SHIFTED}}^\parallel), \quad (18)$$

$\mathbb{P}(v_{\text{SHIFTED}}^\parallel)$  here indicates the probability density of  $v_{\text{SHIFTED}}^\parallel$ , and lateral deviation  $\Phi_h$  and transversal velocity  $\Phi_{v^\perp}$ . The fitting procedure follows the approach in Refs. [2,8], and technical details are in Appendix F. We report the values of the model parameters in Table I.

With the estimated parameters from Table I and the simulation procedure explained in Appendix D, we perform simulations of 2700 trajectories with a discretization step size of 0.1 s, comparable to dataset AMS. Figure 6(a) displays a collection of simulated trajectories, qualitatively indistinguishable from the measurements.

Next, we consider stochastic properties by comparing the empiric and simulated probability distributions of the fluctuations in three observables: shifted longitudinal velocity, transversal velocity, and lateral deviation. The empiric probability distribution functions, as well as the ones obtained from the simulations, are shown in Figs. 6(c)–6(e). It can be seen

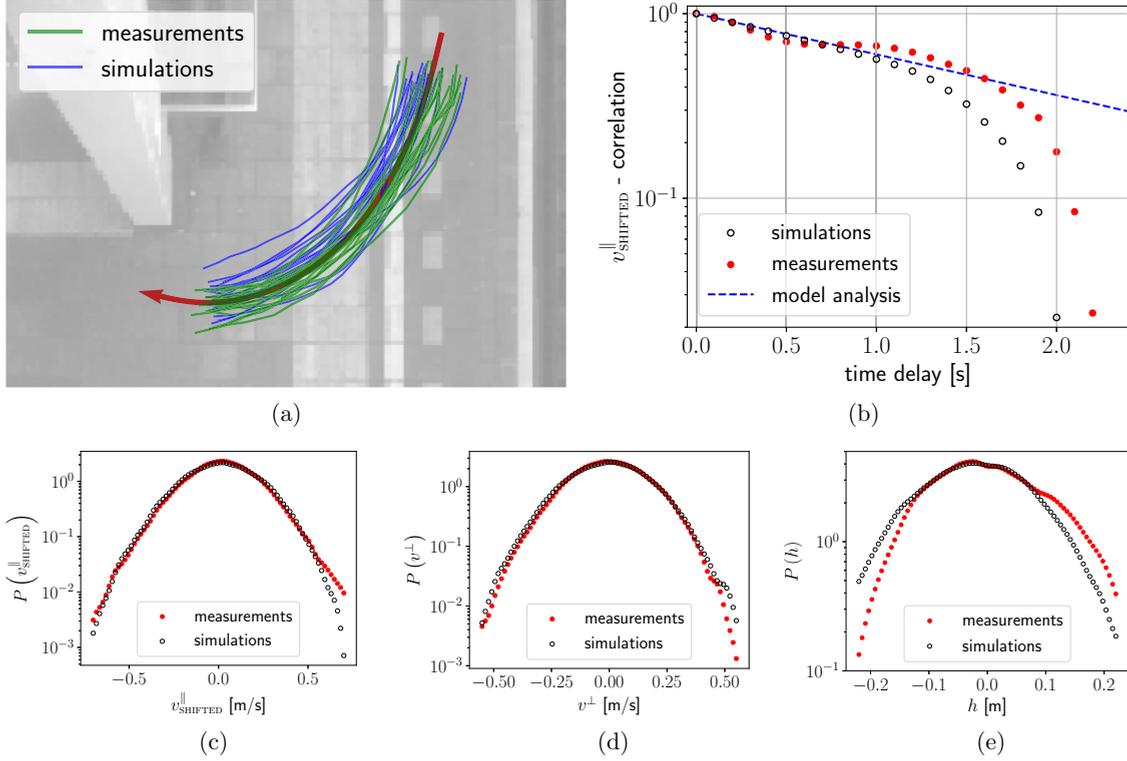


FIG. 6. (a) Top view of Amsterdam south train station platform 2.1 with average path, measured trajectories, and simulated trajectories. (b) Time correlation of the shifted longitudinal velocity,  $v_{\text{SHIFTED}}^{\parallel}$ , of the measured trajectories in the Amsterdam train station (red). The fitted analytic exponential decay [ $\exp(-2\alpha t)$ ] (blue dotted line) is compared with the measurements and simulations. [(c)–(e)] Comparison of the empirical probability distribution functions of the shifted longitudinal velocity (c), transversal velocity (d), and lateral deviation (e) with the distribution functions of the simulation in Amsterdam train station.

that the stochastic properties of the velocity fluctuations are captured by the model. The simulated fluctuations in transversal position are also in good agreement. However, for lateral deviations larger than 10 cm ( $|h| > 0.1$  m), we observe that the empirical fluctuations deviate from the Gaussian behavior. This could potentially be attributed to architectural constraints within the station (e.g., the entrance of the staircase) which could impede inward ( $h < 0$ ) and facilitate outward fluctuations ( $h > 0$ ).

TABLE I. Estimated parameters of the model for the Amsterdam train station dataset.  $\alpha$ , modulating factor of longitudinal propulsion force;  $\beta$ , stiffness coefficient of the transversal linear Langevin dynamics;  $\mu$ , friction coefficient of the transversal linear Langevin dynamics;  $\sigma$ , white noise intensity;  $v_{\text{SP}}$ , straight-path velocity;  $\delta$ , body radius. The parameter estimates and are obtained by fittings of the fundamental diagram, the  $v_{\text{SHIFTED}}^{\parallel}$ -time correlation and Langevin potentials. Further details on the parameter estimation, including error analysis, are provided in Appendix F.

Parameter	Value	
$\alpha$	0.26	$\text{s}^{-1}$
$\beta$	1.17	$\text{s}^{-2}$
$\mu$	0.39	$\text{s}^{-1}$
$\sigma$	0.19	$\text{ms}^{-3/2}$
$v_{\text{SP}}$	1.33	$\text{ms}^{-1}$
$\delta$	0.192	m

Another important statistical property, also used in the model calibration, is the correlation of the shifted longitudinal velocity. In Fig. 6(b), it can be seen that the empiric  $v_{\text{SHIFTED}}^{\parallel}$ -correlation is recovered reasonably well by the model.

## VII. DISCUSSION

We have investigated the fluctuating dynamics of undisturbed pedestrians walking along curved paths with high statistical, space and time accuracy. Our analysis hinged on large trajectory datasets acquired in both real-life conditions and in a experimental setup. The trajectories in the datasets cover a broad range of curvature radii. Thanks to these, we have shown that in the diluted limit a fundamental diagram-like relation between the average longitudinal walking velocity and path curvature exists. Specifically, the average longitudinal velocity decreases for increasing curvature. Notably this reduction is quantitatively compatible with a basic rigid-body-like kinematic model. A first-order expansion of such a model, yield a fundamental diagram-like relation. Based on the large datasets, we have analyzed pedestrian motion beyond averages targeting fluctuations in velocity and lateral deviation. Analogously to the case of straight paths, these fluctuations display Gaussian statistics, as swaying motion, and intersubject/intrasubject variabilities superimpose. Besides, the amplitude of the velocity fluctuations (variance) is independent on the curvature level, at for the range of curvatures observed ( $k \in [0, 1] \text{ m}^{-1}$ ).

Based on these findings, we have extended the quantitative Langevin-like model by Corbetta *et al.* [8] to reproduce, in a statistically quantitative way, the walking dynamics of pedestrians along generic, curved, average paths. In our model, we have considered pedestrians as particle moving according to a custom geodesic flow shaped after the average path. The geodesics we consider are characterized by the conservation of kinetic energy and by the fact that they remain parallel to the average path (when the initial velocity is). We have modeled pedestrian dynamics by perturbing this geodesic flow by (social-like) forces representing (lateral) path adherence, longitudinal propulsion, and random noisy fluctuations. We have validated the model by comparing the probability density functions and the correlation functions generated by repeated model simulations with our measurements at Amsterdam South station. Our model successfully captures the stochastic features of the motion in terms of fluctuations in velocity and position.

We have opted to operate in Cartesian coordinates within a curved geometry, embedding curvature effects in a custom covariant derivative. We believe this choice is instrumental towards further generalization of the model to include, e.g., interactions with other pedestrians and/or different types of forces or noise. All these are typically addressed in Cartesian coordinates. Within the geometric framework we propose, in fact, no coordinate transformations of the forces are required, but only a computation of a correction term (i.e., a Christoffel symbol). Mapping interaction forces in the local coordinate system of each pedestrian would rapidly turn prohibitively complex and computationally expensive.

## ACKNOWLEDGMENTS

This work is part of the HTSM research program “HTCrowd: a high-tech platform for human crowd flows monitoring, modeling and nudging” with Project No. 17962 financed by the Dutch Research Council (NWO). The authors acknowledge ProRail B.V. for providing access to the measurements at Amsterdam south train station and Eindhoven train station.

### APPENDIX A: DEFINITION OF PREFERRED PATH $\bar{\mathbf{x}}(s)$

For the definition of the preferred path, we consider a bundle with  $N$  trajectories,

$$\{\mathbf{x}_\nu(t) \mid \nu = 1, 2, \dots, N\}.$$

Due to the variability in the velocity of pedestrians, we parametrize each trajectory by the relative time,

$$s := \frac{t - t_1}{t_2 - t_1},$$

where  $t_1$  and  $t_2$  are the times that a trajectory enters and leaves the measurement site, respectively. The preferred path,  $\bar{\mathbf{x}}(s)$ , is defined as an ensemble average over the bundle at each relative time instance  $s \in [0, 1]$ :

$$\bar{\mathbf{x}}(s) = \langle \mathbf{x}_\nu(s) \rangle_\nu = \frac{1}{N} \sum_{\nu=1}^N \mathbf{x}_\nu(s). \quad (\text{A1})$$

## APPENDIX B: DATA SELECTION PROCEDURE

### 1. Temporal filters

High-frequency fluctuations (i.e., with frequency greater than 1.2 Hz) do not correspond to typical fluctuations in pedestrian dynamics. These high-frequency fluctuations are presumably connected to experimental errors. To diminish the effect of these high frequencies, we use a Butterworth filter (cut-off frequency 1.2 Hz, order: 5 [29]).

### 2. Trajectory selection AMS

To ensure that the data only contains trajectories under diluted conditions, we restrict to trajectories tracked when no other pedestrian is tracked on the platform. We, furthermore, restrict our study to walking speed by removing trajectories with average velocity outside  $[0.5, 2.5] \text{ ms}^{-1}$ .

We consider the bundle with trajectories starting near the railroad [i.e.,  $(x, y) \in [0.6, 1.3] \times [1.6, 2.3] \text{ m}^2$ ] and finishing at the staircase [i.e.,  $(x, y) \in [-3.0, -0.2] \times [0.7, 3.5] \text{ m}^2$ ] depicted by the two rectangles in Fig. 7(a). We determine an average path,  $\bar{\mathbf{x}}(s)$ , according to Eq. (A1) and parametrize its tubular neighborhood with coordinates  $s$  and  $h$  as in Appendix C. Note that  $h$  represents the normal deviation from  $\bar{\mathbf{x}}(s)$ . For a trajectory  $\mathbf{x}(t)$ , we use the evolution of its  $h$  coordinate,  $h(t)$ , to determine the distance from  $\bar{\mathbf{x}}(s)$ :

$$\|\bar{\mathbf{x}} - \mathbf{x}\| := \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} |h(t)| dt,$$

where the trajectory is defined for time  $t \in [t_1, t_2]$ . We improve the bundle by filtering out the 5% most deviating trajectories. That is  $\|\bar{\mathbf{x}} - \mathbf{x}\| > 24.6 \text{ cm}$ , as depicted in Figs. 7(a) and 7(b).

### 3. Trajectory selection EHV

In contrast with the measurements at Amsterdam train station, nearly always more than one pedestrian is measured at the measurement domain in Eindhoven train station. Therefore we employ a rectangular grid consisting of  $3 \text{ m} \times 3 \text{ m}$  cells. We define the local density as the number of pedestrians in a grid cell. To ensure diluted conditions, we only consider trajectories where the local density does not exceed one during their course. Furthermore, we ensure walking trajectories by applying the same velocity restriction as in AMS trajectory selection.

We group trajectories that originate and terminate in the same areas of the train station into bundles. Five bundles are suited for our analysis as they contain many (curved) paths. Average paths are determined as before. In a similar fashion to AMS trajectory selection, we improve each bundle by discarding the most deviating 10%. The average paths of three bundles are displayed in Fig. 1 (paths in red, blue, and green correspond to bundles 1, 2, and 3, respectively). The average paths of bundles 4 and 5 are displayed in Fig. 7(c).

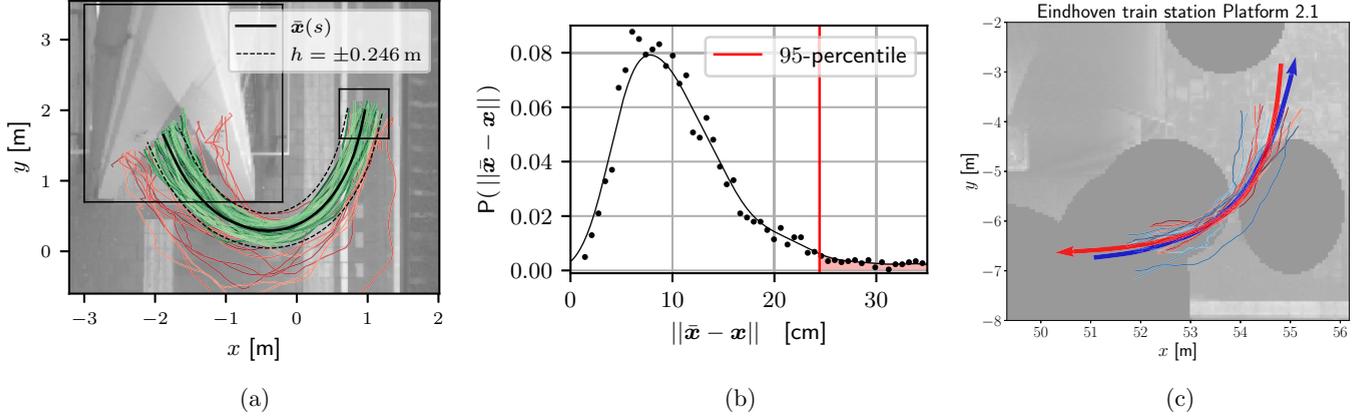


FIG. 7. (a) Collection of trajectories at Amsterdam south train station platform 2.1 including the origination and termination rectangles. Trajectories among the 95% least deviating trajectories are colored green. The 95-percentile of average absolute deviation ( $\|\bar{\mathbf{x}} - \mathbf{x}\| = 24.6$  cm) is indicated with the dotted line. The green trajectories are selected. (b) Probability distribution function of the trajectory deviation from the preferred path ( $\|\bar{\mathbf{x}} - \mathbf{x}\|$ ). The 95-percentile, indicated with the red line, is used for trajectory selection. (c) The average paths of measurement bundle 4 (blue) and bundle 5 (red) with some selected trajectories in Eindhoven train station.

### APPENDIX C: CONSTRUCTION OF TUBULAR NEIGHBORHOOD AND DERIVATION OF THE COVARIANT DERIVATIVE

#### 1. Covariant derivative

A covariant derivative (a.k.a. affine connection) is a mapping that describes how vectors change when transporting them in a smooth collection of tangent spaces. The concept of covariant derivative can be understood as an generalization of the ordinary derivative towards curved surfaces. For  $\mathbf{u}$  and  $\mathbf{v}$  vectors in a tangent space of a curved surface, the covariant derivative of  $\mathbf{u}$  along  $\mathbf{v}$  is denoted as  $\nabla_{\mathbf{v}}\mathbf{u}$  and respects the following properties (e.g., Refs. [28]):

- (i)  $\nabla_{f_1\mathbf{v}_1+f_2\mathbf{v}_2}\mathbf{u} = f_1\nabla_{\mathbf{v}_1}\mathbf{u} + f_2\nabla_{\mathbf{v}_2}\mathbf{u}$ ,
- (ii)  $\nabla_{\mathbf{v}}(\mathbf{u}_1 + \mathbf{u}_2) = \nabla_{\mathbf{v}}\mathbf{u}_1 + \nabla_{\mathbf{v}}\mathbf{u}_2$ ,
- (iii)  $\nabla_{\mathbf{v}}(f\mathbf{u}) = f\nabla_{\mathbf{v}}\mathbf{u} + \mathbf{v}(f) \cdot \mathbf{u}$ ,

for  $\mathbf{u}, \mathbf{u}_1, \mathbf{u}_2, \mathbf{v}, \mathbf{v}_1, \mathbf{v}_2$  in a tangent space and  $f, f_1, f_2$  smooth functions.

We can define the covariant derivative by defining *Christoffel symbols of the second kind*,  $\Gamma_{ij}^k$ . These coefficients determine how basis vectors in different spaces are connected via

$$\Gamma_{ij}^k \mathbf{e}_k := \nabla_{\mathbf{e}_i} \mathbf{e}_j. \quad (\text{C1})$$

Note that from now on, we will use the Einstein summation convention (e.g.,  $\Gamma_{ij}^k \mathbf{e}_k \equiv \sum_k \Gamma_{ij}^k \mathbf{e}_k$ ). Using the properties above, we could write the covariant derivative in terms of Christoffel symbols:

$$\nabla_{\mathbf{v}}\mathbf{u} = \frac{\partial \mathbf{u}}{\partial \mathbf{v}} + u^k u^j \Gamma_{kj}^i \mathbf{e}_i, \quad (\text{C2})$$

where  $\mathbf{u} = u^i \mathbf{e}_i$ .

The covariant derivative can be pushed forward to other coordinate charts using the coordinate transformation  $\phi = \psi_{\beta} \circ \psi_{\alpha}^{-1}$ , which maps from chart  $\psi_{\alpha}$  to chart  $\psi_{\beta}$ . This induces a relation between Christoffel symbols in different coordinate charts:

$$\Gamma_{ij}^k = T_{\ell}^k (S_j^m S_i^n \bar{\Gamma}_{nm}^{\ell} + \partial_j S_i^{\ell}), \quad (\text{C3})$$

with  $T = J_{\phi}$  and  $S = J_{\phi^{-1}} = [J_{\phi}]^{-1}$  the (inverse) Jacobian of  $\phi$  and  $\Gamma_{ij}^k$  and  $\bar{\Gamma}_{ij}^k$  the Christoffel symbols in the coordinate charts  $\psi_{\alpha}$  and  $\psi_{\beta}$ , respectively.

#### 2. Tubular neighborhood

We construct a coordinate chart,  $\psi_{\bar{\mathbf{x}}}$ , that covers the tubular neighborhood of a generic curve  $\bar{\mathbf{x}} : \mathbb{R} \rightarrow \mathbb{R}^2 : s \mapsto (x, y)$  by using the tangent and normal vectors,

$$\mathbf{e}_{\parallel} = \frac{\bar{\mathbf{x}}'(s)}{|\bar{\mathbf{x}}'(s)|} \quad \text{and} \quad \mathbf{e}_{\perp} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \mathbf{e}_{\parallel}, \quad (\text{C4})$$

as basis vectors. The coordinate lines are parallel and normal to  $\bar{\mathbf{x}}(s)$  with coordinates  $s$  and  $h$  representing the parallel and transversal direction, respectively. The coordinate transformation from  $\psi_{\bar{\mathbf{x}}}$  to the Cartesian coordinates is given by

$$\phi(s, h) = \bar{\mathbf{x}}(s) + h \mathbf{e}_{\perp}(s). \quad (\text{C5})$$

#### 3. Energy-conserving connection

Geodesics are generally defined as parallel transport of velocity vectors in their own direction [28],

$$\nabla_{\dot{\mathbf{x}}}\dot{\mathbf{x}} = \ddot{\mathbf{x}} + \Gamma_{kj}^i \dot{x}^k \dot{x}^j \mathbf{e}_i = \mathbf{0}, \quad (\text{C6})$$

analogously to Eq. (12) with correction term  $C(\mathbf{x}, \dot{\mathbf{x}}, \bar{\mathbf{x}}(s)) = \Gamma_{kj}^i \dot{x}^k \dot{x}^j \mathbf{e}_i$ . We derive our affine connection (i.e., derive the Christoffel symbols) such that geodesics respect the physical properties:

- (i) geodesic curves conserve kinetic energy;
- (i) geodesic curves preserve the transversal velocity, as explained in Sec. V. These properties fully describe geodesics in flat space nearby straight paths as  $\ddot{\mathbf{x}} = \mathbf{0}$  ( $\Gamma_{kj}^i = 0 \forall_{i,j,k}$ ). However, this simple connection does not hold for curved paths or curvilinear coordinates.

Energy conservation is ensured by conserving the physical velocity,

$$\|\mathbf{v}\|^2 = \tilde{g}_{ij} \dot{q}^i \dot{q}^j, \quad (\text{C7})$$

where trajectory  $(q^1(t), q^2(t))$  is in generic coordinate chart  $\psi_q$  with metric  $\hat{g}$ . By defining the metric tensor in the Cartesian coordinate chart as  $g_{ij} = \delta_{ij}$ , we define the physical velocity to be the Euclidean velocity ( $\|\mathbf{v}\|^2 = \dot{x}^2 + \dot{y}^2$ ). We define coordinate chart  $\psi_{\bar{x}}$  as in Appendix C 2. Then the metric in  $\psi_{\bar{x}}$  is given by [28]:

$$\hat{g}_{kq} = g_{ij} \frac{\partial \phi^i}{\partial s^k} \frac{\partial \phi^j}{\partial s^q}, \quad (\text{C8})$$

where  $\phi$  denotes the coordinate transformation to the Cartesian coordinates [Eq. (C5)] and  $(s^1, s^2) = (s, h)$ . Note that  $\hat{g}_{sh} = \hat{g}_{hs} = \langle \mathbf{e}_{\parallel}, \mathbf{e}_{\perp} \rangle = 0$  since  $\mathbf{e}_{\parallel} \perp \mathbf{e}_{\perp}$ . Furthermore,  $\hat{g}_{hh} = \|\mathbf{e}_{\perp}\|^2 = 1$  by definition. Therefore the metric in coordinate chart  $\psi_{\bar{x}}$  can be written as

$$\hat{g}_{ij} = \begin{bmatrix} (\partial_s \phi_x)^2 + (\partial_s \phi_y)^2 & 0 \\ 0 & 1 \end{bmatrix}. \quad (\text{C9})$$

Because the metric is diagonal, the physical velocity can be separated into two orthogonal parts,

$$\mathbf{v}^{\parallel} = \sqrt{\hat{g}_{ss}} \dot{s} \quad \text{and} \quad \mathbf{v}^{\perp} = \sqrt{\hat{g}_{hh}} \dot{h}, \quad (\text{C10})$$

which are the longitudinal and transversal velocity components, respectively. To meet the properties, both velocity components need to be conserved, meaning

$$\begin{aligned} \frac{d}{dt} \mathbf{v}^{\parallel} &= \frac{d}{dt} (\sqrt{\hat{g}_{ss}} \dot{s}) = 0, \\ \frac{d}{dt} \mathbf{v}^{\perp} &= \ddot{h} = 0, \end{aligned} \quad (\text{C11})$$

which can be elaborated to

$$\begin{aligned} \ddot{s} + \frac{\partial_s \hat{g}_{ss}}{2\hat{g}_{ss}} \dot{s}^2 + \frac{\partial_h \hat{g}_{ss}}{2\hat{g}_{ss}} \dot{h} \dot{s} &= 0, \\ \ddot{h} &= 0. \end{aligned} \quad (\text{C12})$$

Using Eq. (C2), the Christoffel symbols in  $\psi_{\bar{x}}$  can be determined such that Eq. (C12) is respected:

$$\bar{\Gamma}_{ij}^s = \begin{bmatrix} \frac{\partial_s \hat{g}_{ss}}{2\hat{g}_{ss}} & \frac{\partial_h \hat{g}_{ss}}{2\hat{g}_{ss}} \\ \frac{\partial_h \hat{g}_{ss}}{2\hat{g}_{ss}} & 0 \end{bmatrix}, \quad (\text{C13})$$

and

$$\bar{\Gamma}_{ij}^h = 0. \quad (\text{C14})$$

Hence, the minimum connection that respects the given physical properties is fully described by the Christoffel symbols in Eq. (C13) and Eq. (C14). Note that we can obtain the Christoffel symbols in the Cartesian coordinate chart by applying Eq. (C3).

#### APPENDIX D: NUMERICAL SIMULATIONS

We integrate Eq. (17) by using the Runge-Kutta SRI2 algorithm [30] (via the PyPI library `sdeint` [31]). We choose a discretization step size of 0.1 s, similar to the sampling frequency of our measurements. We initialize our simulations at the beginning of our preferred path  $\bar{\mathbf{x}}(s)$  with  $s(t=0) = 0$  and  $h(0)$ ,  $v^{\perp}(0)$  and  $v^{\parallel}_{\text{SHIFTED}}(0)$  distributed according to the Fokker-Planck equilibrium distributions (see Appendix F).

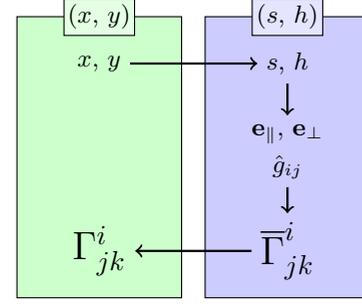


FIG. 8. A schematic overview of the calculation steps for determining the Christoffel symbols. (1) Estimation of tubular coordinates using the Newton-Raphson method. (2) Calculation of tubular neighborhood-dependent variables such as the longitudinal and transversal directions,  $\mathbf{e}_{\parallel}$  and  $\mathbf{e}_{\perp}$ , and the metric  $\hat{g}_{ij}$ . (3) Computation of the Christoffel symbols in the coordinate chart  $\psi_{\bar{x}}$ . (4) Push-forward of the Christoffel symbols to the Cartesian coordinates.

The Christoffel symbols, needed every time step during the integration of Eq. (17), are obtained by the computational steps shown in Fig. 8. For step (1), the computation of the tubular coordinates, we use the two-dimensional Newton-Raphson method [32]. This iterative method solves equations of the form  $\mathbf{f}(\mathbf{s}) = 0$ . If  $\mathbf{s}_0$  is an approximate solution, then the sequence

$$\mathbf{s}_{p+1} = \mathbf{s}_p - J^{-1}(\mathbf{s}_p) \mathbf{f}(\mathbf{s}_p)$$

for  $p = 1, 2, \dots$  and  $J$  Jacobian of  $\mathbf{f}$ , converges to a solution. Given  $\mathbf{x}$ , the tubular coordinates are represented by the roots of function  $\mathbf{f}(\mathbf{s}) = \phi(\mathbf{s}) - \hat{\mathbf{x}}$ . The roots of  $\mathbf{f}$  are estimated with the Newton-Raphson method with the tubular coordinates of the previous time step as an approximated solution. With our typical simulation duration and discretization step size, two iterations of the Newton-Raphson method give a sufficient accurate estimation of coordinates  $s$  and  $h$ .

In step (2), we use Eqs. (C4), (C5), and (C9) to calculate the metric in tubular coordinates,  $\hat{g}_{ss}$ , and the derivatives with respect to  $s$  and  $h$  ( $\partial_s \hat{g}_{ss}$  and  $\partial_h \hat{g}_{ss}$ ). We compute the Christoffel symbols in the tubular coordinate chart in step (3) using Eq. (C13) and (C14). Finally, in step (4), we push the Christoffel symbols to the Cartesian coordinate chart using Eq. (C3).

#### APPENDIX E: MODEL SIMPLIFICATION FOR CIRCULAR PREFERRED PATHS

In the idealized case in which the preferred path has circular shape, the equations for our covariant derivative and our complete Langevin-like model [Eq. (17)] can be written in explicit form. This can serve as a simplified model for conditions in which the preferred path is characterized by long portions of constant curvature (e.g., with respect to the relaxation timescale  $\tau$ ). In this Appendix, we derive the equations for this case. We report this as a practical example that allows to work with our model bypassing the technical challenges. We stress however that in our results Sect. VI, we employ our complete model (via the numerical approach in Appendix D) allowing the curvature center and curvature

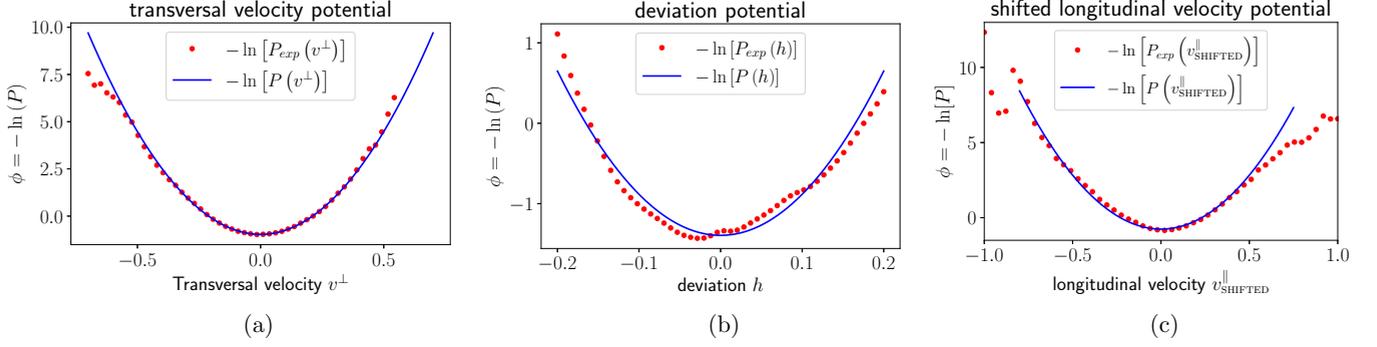


FIG. 9. (a) The empiric potentials,  $\Phi_{\text{exp}} = -\ln(\mathbb{P}_{\text{exp}})$ , of (a) the transversal velocity, (b) the lateral deviation, and (c) the shifted longitudinal velocity obtained from the Amsterdam train station data (red) compared with the fitted model potentials (blue).

radius of the preferred path to change as it happens in our study cases.

As curvature center and curvature radius of the preferred path do not change, certain terms conveniently cancel or simplify. Hence both sides of our model in Eq. (17) can be derived in analytic form and written in Cartesian coordinates. This means that simulations do only require time discretization, while the coordinate transformations in Fig. 8 are not necessary.

Considering a preferred path conveniently centered in the origin and of radius  $R$ , its equations are

$$\bar{\mathbf{x}}(s) = R \cos(s)\mathbf{e}_x + R \sin(s)\mathbf{e}_y, \quad (\text{E1})$$

where the variable  $s$  indicates the phase along the circumference. We can parametrize the tubular neighborhood of this path as

$$\boldsymbol{\phi}(s, h) = (R + h) \cos(s)\mathbf{e}_x + (R + h) \sin(s)\mathbf{e}_y. \quad (\text{E2})$$

The quantity  $\rho = R + h$  is the distance between the origin and the point considered. For the sake of notation, we shall use the symbol  $\rho$  when convenient. In this tubular neighborhood, the longitudinal and transversal velocity component can be written using the polar coordinate local base:

$$\mathbf{e}_\parallel = R_{90} \frac{\mathbf{x}}{\rho} \quad \mathbf{e}_\perp = \frac{\mathbf{x}}{\rho}, \quad (\text{E3})$$

where  $R_{90}$  is a matrix yielding a 90-degree counterclockwise rotation, and which gives

$$v^\parallel = \dot{\mathbf{x}} \cdot R_{90} \frac{\mathbf{x}}{\rho} = \frac{\dot{x}y - \dot{y}x}{\rho} \quad v^\perp = \dot{\mathbf{x}} \cdot \frac{\mathbf{x}}{\rho} = \frac{\dot{x}x + \dot{y}y}{\rho}. \quad (\text{E4})$$

In  $(s, h)$  coordinates, the Euclidean metric reads

$$\hat{g}_{ij} = \begin{pmatrix} \rho^2 & 0 \\ 0 & 1 \end{pmatrix}. \quad (\text{E5})$$

On this basis, we can write our geodesic equation in  $(s, h)$  coordinates using Eq. (C12) that reads

$$\begin{aligned} \ddot{s} + \frac{\dot{h}\dot{s}}{\rho} &= 0, \\ \ddot{h} &= 0. \end{aligned} \quad (\text{E6})$$

The same equations can be mapped in Cartesian coordinates by computing the second time derivative of Eq. (E2) and

replacing Eq. (E6). This yields

$$\begin{aligned} \ddot{x} + x\dot{s}^2 + y\frac{\dot{\rho}\dot{s}}{\rho} &= 0, \\ \ddot{y} + y\dot{s}^2 - x\frac{\dot{\rho}\dot{s}}{\rho} &= 0, \end{aligned} \quad (\text{E7})$$

where

$$\begin{aligned} \rho &= \sqrt{x^2 + y^2}, \\ \dot{s} &= \frac{y\dot{x} - \dot{y}x}{\rho^2} = \frac{v^\parallel}{\rho}. \end{aligned} \quad (\text{E8})$$

Equation (E7) is the left-hand side of Eq. (17). The right-hand side of Eq. (17) can be written immediately considering the components in Eq. (E4), and that  $\dot{k} = 0$  and  $h = \rho - R$ .

## APPENDIX F: MODEL CALIBRATION

The model is calibrated by estimating the model parameters,  $\{\alpha, \beta, \mu, v_{\text{SP}}, \delta, \sigma\}$ . We use the Amsterdam train station measurements to estimate the parameters. The ‘‘straight-path velocity’’  $v_{\text{SP}}$  is estimated by linearly extrapolating the  $v^\parallel - k$  relation towards  $k = 0$ . For the Amsterdam train station this results in  $v_{\text{SP}} = 1.33 \text{ ms}^{-1}$ . The body size radius,  $\delta$ , represent the slope of the fundamental diagram [Fig. 3(b)]. The estimation of  $\delta$  is obtained by a linear fit:  $\delta = 0.19 \text{ m}$ . The remaining four parameters are estimated by applying fits to empirical Langevin potentials and a correlation function. The first fit is applied to the transversal velocity potential. In the stationary regime, the model produces probability distribution of the transversal velocity and lateral deviation from the preferred path,  $\mathbb{P}(h, v^\perp)$ , according the well-known Fokker-Planck equation [33] with solutions

$$\mathbb{P}(h, v^\perp) = \mathcal{N} \exp \left[ -\frac{2\mu}{\sigma^2} (v^\perp)^2 - \frac{4\beta\mu}{\sigma^2} h^2 \right], \quad (\text{F1})$$

where  $\mathcal{N}$  denotes a normalization constant. A Langevin potential can be constructed according to  $\Phi(\cdot) = -\ln(\mathbb{P}(\cdot))$ . The analytical potentials of the transversal dynamics should agree with the empiric potentials such that

$$-\ln(\mathbb{P}_{\text{exp}}(v^\perp)) \approx \frac{2\mu}{\sigma^2} (v^\perp)^2 + K_1 \quad (\text{F2})$$

and

$$-\ln(\mathbb{P}_{\text{exp}}(h)) \approx \frac{4\beta\mu}{\sigma^2} h^2 + K_2. \quad (\text{F3})$$

The constants  $K_1$  and  $K_2$  are normalization constants and  $\mathbb{P}_{\text{exp}}(\cdot)$  denotes the empiric probability distribution function. The fitting can be observed in Figs. 9(a) and 9(b) where the resulting estimated ratios are given by

$$\frac{2\mu}{\sigma^2} \approx 21.77 \quad \text{and} \quad \frac{4\beta\mu}{\sigma^2} \approx 51.08. \quad (\text{F4})$$

The same can be done for the longitudinal dynamics. In the stationary regime, the probability of the shifted longitudinal velocity is distributed according to

$$\mathbb{P}(v_{\text{SHIFTED}}^{\parallel}) = \mathcal{N} \exp\left[\frac{2\alpha}{\sigma^2}(v_{\text{SHIFTED}}^{\parallel})^2\right], \quad (\text{F5})$$

where  $\mathcal{N}$  is a normalization constant. The ratio  $\frac{2\alpha}{\sigma^2}$  is compared to the empirical distribution function of the shifted longitudinal velocity according to

$$-\ln(\mathbb{P}_{\text{exp}}(v_{\text{SHIFTED}}^{\parallel})) \approx \frac{2\alpha}{\sigma^2}(v_{\text{SHIFTED}}^{\parallel})^2 + K_3. \quad (\text{F6})$$

Constant  $K_3$  again represents normalization. The fit [Fig. 9(c)] results in the estimation of the ratio:

$$\frac{2\alpha}{\sigma^2} \approx 14.36. \quad (\text{F7})$$

To complete the parameter estimation, a time correlation function of the shifted longitudinal velocity is used. Using Eq. (9) and the definition of  $v_{\text{SHIFTED}}^{\parallel}$ , the deterministic shifted longitu-

TABLE II. Estimated parameter values with associated uncertainty intervals.

Parameter	Value	Uncertainty interval	
$\alpha$	0.26	[0.22, 0.28]	$\text{s}^{-1}$
$\beta$	1.17	[0.80, 1.67]	$\text{s}^{-2}$
$\mu$	0.39	[0.31, 0.46]	$\text{s}^{-1}$
$\sigma$	0.19	[0.17, 0.20]	$\text{ms}^{-3/2}$
$v_{\text{SP}}$	1.33	[1.29, 1.35]	$\text{ms}^{-1}$
$\delta$	0.192	[0.187, 0.195]	m

dinal dynamics can be described by

$$\frac{d}{dt} v_{\text{SHIFTED}}^{\parallel} = -2\alpha v_{\text{SHIFTED}}^{\parallel}. \quad (\text{F8})$$

Therefore, the time correlation of  $v_{\text{SHIFTED}}^{\parallel}$  should decay as  $\exp(-2\alpha t)$ . An estimated value of  $\alpha$  follows from the fit [Fig. 6(b)]:

$$-2\alpha \approx -0.51. \quad (\text{F9})$$

The estimates obtained by the fitted values result in the parameter values reported in Table II. To determine uncertainty intervals for our estimates, we repeat the fitting procedure five times using randomly selected, equally sized partitions of the data. We then use the fitted values from each of the five partitions to estimate the minimum and maximum values for each parameter. We set these as the lower and upper bounds of the respective intervals.

- [1] D. R. Parisi, P. A. Negri, and L. Bruno, *Phys. Rev. E* **94**, 022318 (2016).
- [2] A. Corbetta, J. A. Meeusen, C.-M. Lee, R. Benzi, and F. Toschi, *Phys. Rev. E* **98**, 062310 (2018).
- [3] A. Corbetta and F. Toschi, *Annu. Rev. Condens. Matter Phys.* **14**, 311 (2023).
- [4] C. Feliciani, K. Shimura, and K. Nishinari, *Introduction to Crowd Management: Managing Crowds in the Digital Era: Theory and Practice* (Springer, Cham, 2021).
- [5] J. Zhang, W. Mehner, S. Holl, M. Boltes, E. Andresen, A. Schadschneider, and A. Seyfried, *Phys. Lett. A* **378**, 3274 (2014).
- [6] J. Zhang, W. Klingsch, A. Schadschneider, and A. Seyfried, *J. Stat. Mech.* (2012) P02002.
- [7] H. Murakami, C. Feliciani, Y. Nishiyama, and K. Nishinari, *Sci. Adv.* **7**, eabe7758 (2021).
- [8] A. Corbetta, C.-M. Lee, R. Benzi, A. Muntean, and F. Toschi, *Phys. Rev. E* **95**, 032316 (2017).
- [9] A. Corbetta and F. Toschi, in *Complexity Science: An Introduction*, edited by M. Peletier, R. Van Santen, and E. Steur (World Scientific, Singapore, 2019), pp. 329–345.
- [10] H. Hicheur, S. Vieilledent, M. J. E. Richardson, T. Flash, and A. Berthoz, *Exp. Brain Res.* **162**, 145 (2005).
- [11] V. Ziemer, A. Seyfried, and A. Schadschneider, in *Traffic and Granular Flow'15*, edited by V. L. Knoop and W. Daamen (Springer, Cham, 2016), pp. 89–96.
- [12] R. Gerike, C. Koszowski, B. Schröter, R. Buehler, P. Schepers, J. Weber, R. Wittwer, and P. Jones, *Sustainability* **13**, 9362 (2021).
- [13] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, *Rev. Mod. Phys.* **85**, 1143 (2013).
- [14] M. Boltes and A. Seyfried, *Neurocomputing* **100**, 127 (2013).
- [15] L. D. Vanumu, K. Ramachandra Rao, and G. Tiwari, *Eur. Transp. Res. Rev.* **9**, 49 (2017).
- [16] S. Seer, N. Brändle, and C. Ratti, *Transp. Res. C Emerg. Technol.* **48**, 212 (2014).
- [17] C. A. S. Pouw, J. Willems, F. van Schadewijk, J. Thureau, F. Toschi, and A. Corbetta, *Coll. Dyn.* **6**, 1 (2022).
- [18] D. Bršćić, T. Kanda, T. Ikeda, and T. Miyashita, *IEEE Trans. Human-Machine Syst.* **43**, 522 (2013).
- [19] A. Gabbana, F. Toschi, P. Ross, A. Haans, and A. Corbetta, *PNAS Nexus* **1**, pgac169 (2022).
- [20] J. Willems, A. Corbetta, V. Menkovski, and F. Toschi, *Sci. Rep.* **10**, 11653 (2020).
- [21] I. Zuriguel, D. R. Parisi, R. C. Hidalgo, C. Lozano, A. Janda, P. A. Gago, J. P. Peralta, L. M. Ferrer, L. A. Pugnaloni, E. Clément *et al.*, *Sci. Rep.* **4**, 7324 (2014).
- [22] D. Bršćić, F. Zanlungo, and T. Kanda, *Trans. Res. Proc.* **2**, 77 (2014).
- [23] F. Zanlungo, T. Ikeda, and T. Kanda, *Phys. Rev. E* **89**, 012811 (2014).

- [24] C. A. S. Pouw, F. Toschi, F. van Schadewijk, and A. Corbetta, *PLoS ONE* **15**, e0240963 (2020).
- [25] D. Helbing and P. Molnár, *Phys. Rev. E* **51**, 4282 (1995).
- [26] A. Gray, *Tubes* (Birkhäuser, Basel, 2003).
- [27] R. Courant and F. John, *Introduction to Calculus and Analysis I* (Springer, Berlin, 1999).
- [28] J. Jost, *Riemannian Geometry and Geometric Analysis* (Springer, Berlin, 2011).
- [29] I. Selesnick and C. Burrus, *IEEE Trans. Sign. Process.* **46**, 1688 (1998).
- [30] A. Röbler, *SIAM J. Numer. Anal.* **48**, 922 (2010).
- [31] M. J. Aburn, Numerical integration of stochastic differential equations (sde) [<https://pypi.org/project/sdeint/>].
- [32] A. Ben-Israel, *J. Math. Anal. Appl.* **15**, 243 (1966).
- [33] J. L. García-Palacios, Introduction to the theory of stochastic processes and Brownian motion problems, [arXiv:cond-mat/0701242](https://arxiv.org/abs/cond-mat/0701242).