Natural variability of walking behavior in *Drosophila melanogaster*

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen

Fakultät der Universität zu Köln

vorgelegt von

Vincent Godesberg

aus Köln

2024

Berichterstatter:

Prof. Dr. Ansgar Büschges

Prof. Dr. Martin Nawrot

Table of Contents

Abstract

Walking behavior in insects serves as a model for studying motor control, offering insights into general locomotor principles. Although much is known about how the nervous system regulates insect locomotion, variability in walking behavior within and across individuals of the same species has been understudied. This thesis aims to close the gap of knowledge by investigating the natural variability of walking behavior in *Drosophila melanogaster.* High-speed videography and automated, deep learning-based tracking were used to collect extensive data of leg-tip kinematics in 2D from a large cohort of wildtype flies.

Principal component analysis (PCA) was applied to unbiasedly explore patterns of covariance in the data set. Five principal components (PCs) were found to explain the majority of the observed variability. A set of two PCs primarily represented general interleg coordination patterns, the former reflecting the characteristic tripod gait that is widely observed in insect locomotion and the latter reflecting the systematic deviation from canonical tripod coordination. In contrast, the remaining three PCs captured idiosyncrasies in posture and leg movements, which remained stable across different speeds, suggesting that each fly maintains a unique, consistent walking signature. Although displaying the same idiosyncrasies across walking speeds, the relative fraction of variability described by these three PCs decreased for higher speeds, while the PC describing tripod coordination markedly gained significance. This finding indicates, that interindividual variability of leg-tip kinematics decreases with increasing walking speeds, which is further investigated by analyzing the correlation of spatial and temporal variability with walking speed and measures of interleg coordination. For this purpose, tripod coordination strength (TCS) was used as a proxy for interleg coordination and a novel coherence measure was designed to quantify how well timed intersegmental influences between adjacent legs are. I show that the coherence measure in direct comparison outcompetes walking speed and TCS at predicting both spatial and temporal variability. This result suggests that the stronger entrainment of weakly coupled oscillators controlling single leg movements is causally related to the observed decrease in variability as walking speed increases. Finally, I developed an advanced experimental setup to enable continuous high-

resolution recordings of freely walking flies over long durations, capturing kinematic data at up to 200 Hz. This setup, combined with optimized data processing tools, enabled a robust collection of leg-tip kinematics and efficient filtering for behavior of interest, which was demonstrated by a rudimentary analysis of curved walking behavior.

Overall, this thesis provides a comprehensive quantitative analysis of the natural variability in *Drosophila* walking behavior. By separating general motor control features from idiosyncrasies and systematically analyzing how key parameters of walking behavior interact with variability, this work offers novel insights into the underlying neural mechanisms enabling consistent and adaptive motor control. The presented findings might serve as a foundation for future studies on how neural circuits generate flexible and yet robust behaviors in both insects and vertebrates.

List of Abbreviations

2D	two-dimensional
3D	three-dimensional
AEPs	anterior extreme positions
BDSC	Bloomington *Drosophila* stock center
BL	body length
COM	center of mass
CPG	central pattern generators
DLC	DeepLabCut
fCO	femoral chordotonal organ
GB	gigabyte
GUI	graphic user interface
GtACR1	*Guillardia theta* anion channelrhodopsin 1
IR	infrared
L1 – L3	left front, middle, and hind leg
LED	light emitting diode
MB	megabyte
PC	principal component
PCA	principal component analysis
PEPs	posterior extreme positions
R1 – R3	right front, middle, and hind leg
RAM	random access memory
RMSE	root mean square error
SC	step cycle
SD	standard deviation
TCS	tripod coordination strength
t-SNE	t-distributed stochastic neighbor embedding
UAS	upstream activating sequence
UMAP	uniform manifold approximation and projection

1 Chapter 1: General Introduction

Legged locomotion, particularly walking, is a common and fundamental behavior in terrestrial animals, playing a crucial role in activities like foraging, exploration, predator avoidance, and migration. Locomotion, as a result, has been a key factor in animal evolution, leading to a vast diversity of body morphologies optimized for various modes of movement (Dickinson, 2000; Zenkevich, 1945). These modes of movement include walking, crawling, swimming, and flying, as well as more specialized forms like brachiation, burrowing, or hopping (Ijspeert, 2003). Despite this diversity, the fundamental principle of locomotion remains the same: animals generate force against their environment to propel themselves in a desired direction (Dickinson, 2000; Holmes et al., 2006). This force is produced by rhythmic muscle activity, coordinated by the nervous system, which is continuously adjusted by sensory feedback to cope with perturbations and changing environmental conditions. In addition, higher areas of the nervous system modulate locomotor circuits to achieve goal-oriented behaviors. In summary, locomotion results from a complex interplay of neural processes, muscle activity, biomechanics, body morphology, and environmental influences (Dickinson, 2000; Holmes et al., 2006; Seipel et al., 2017).

Walking, as the most common type of legged locomotion, is the primary mode of movement for terrestrial animals (Biewener and Patek, 2018). It involves the coordinated movement of legs in a cyclic pattern, typically divided into two phases: the stance (power stroke), where the leg contacts the ground to propel the body, and the swing (return stroke), where the leg is lifted and repositioned for the next stance phase. During walking, the nervous system must coordinate both the movements of the individual segments within each leg (intraleg coordination) and the interactions between different legs (interleg coordination). The number of legs which need to be coordinated ranges from bipeds (e.g., humans) to hexapods (e.g., insects) and even millipedes, which can have hundreds of legs. In many walking species, interleg coordination can be categorized into distinct gaits that are optimized for different speeds (Alexander, 2003). For example, horses walk at low speeds, trot at intermediate speeds, and gallop at high speeds, transitioning between gaits to optimize their energetic cost

function (Hoyt and Taylor, 1981). Other animals, like insects, usually do not exhibit distinct gaits, but rather use a continuum of interleg coordination patterns (Nirody, 2021; Szczecinski et al., 2018; Wosnitza et al., 2013).

While interleg coordination necessarily relies on information exchange between the neural circuits controlling individual legs, intraleg coordination has been shown to depend on central pattern generators (CPGs), which are crucial for generating the rhythmic motor patterns underlying locomotion in all animals (Grillner, 2006; MacKay-Lyons, 2002; Mantziaris et al., 2020). CPGs are neural networks that can produce rhythmic motor outputs necessary for timing and coordinating motor neurons without relying on rhythmic sensory input or descending signals. However, for locomotion to be adaptive and flexible, activity of CPGs is shaped by sensory feedback and modulated by descending signals from higher neural centers. While significant progress has been made in understanding the function and role of CPGs, there is still much to learn about how CPGs interact in the context of legged locomotion (Grillner and Kozlov, 2021).

Motor control of walking is frequently studied in insects, as they share many general locomotion principles with vertebrates (Büschges, 2005; Duysens et al., 2000; Pearson, 1993) while their simpler nervous systems facilitate the examination of underlying neural mechanisms (Bidaye et al., 2018). In contrast to vertebrates, which possess a maximum of four legs, insects have six legs, which are attached pairwise to one of three thoracic segments. While the coordination of six walking legs requires specialized control strategies, the basis for coordinating the legs still consists of state-dependent information exchange between CPGs, mechanical coupling, and sensory feedback. Sensory feedback in the legs of insects results from various sensors, such as campaniform sensilla sensing strain and stress in the cuticle (Dickerson et al., 2021; Zill et al., 2004; Tuthill and Wilson, 2024), or chordotonal organs measuring flexion and extension of leg segments as well as sounds and vibrations (Büschges, 1994; Field and Matheson, 1998; Mamiya et al., 2018; Zill, 1985). The feedback provided by these sensors in the context of interleg coordination can trigger phase transitions during walking and is used to adjust and correct parameters of movement such as force or velocity in ongoing motor actions

(Akay et al., 2001; Cruse et al., 1990). A set of coordination rules, often referred to as Cruse rules after Holk Cruse, relies, at least partially, on sensory feedback. The Cruse rules have been described and experimentally confirmed in the stick insect *Carausius morosus* (Cruse, 1990; Cruse et al., 2007) and served as the basis of the distributed neural network controller Walknet (Dürr et al., 2004; Schilling et al., 2013). Walknet successfully demonstrated, both in a model and in robots, that functional hexapodal walking behavior can be generated solely with decentralized rules. Whether the Cruse rules are actually implemented as such in the neural control algorithms of insects and how this decentralized concept relates to CPGs is still not fully understood.

Among the established model organisms, *Drosophila melanogaster* provides an particularly well-suited model to explore walking behavior due to its relatively simple nervous system, the ever-extending capabilities of the Gal4-UAS-system for precise genetic interventions (Duffy, 2002; Hales et al., 2015; Venken et al., 2011), and the capacity for high-throughput behavioral analysis (Bidaye et al., 2018). Walking behavior of *Drosophila* has been extensively studied with a focus on how the nervous system initiates forward and backward walking (Bidaye et al., 2020, 2014), how different parameters change systematically with walking speed (DeAngelis et al., 2019; Strauß and Heisenberg, 1990; Wosnitza et al., 2013), how inhibition of sensory feedback affects leg kinematics (Chockley et al., 2022; Mendes et al., 2013), how different modes of interleg coordination interact with static stability (Szczecinski et al., 2018), or the 3D leg kinematics and the degrees of freedom of every single joint (Haustein et al., 2024). Despite this rich body of knowledge, one aspect which has been mostly neglected is the dimension of variability in walking behavior.

Variability can be observed in any type of animal behavior. The causes of this *behavioral variability* are manifold and operate on several levels, from the sub-cellular level (ion channels) to ontogeny (including genetics and learned adaptations) and, finally, to evolutionary processes. Investigating these causes can teach us a lot about how animals have evolved to cope with the demands of their environment and how nervous systems ultimately produce

behavior. To uncover the underlying principles, it is essential to disentangle the different types, aspects, and dimensions of variability.

A basic distinction must be made between inter- and intra-individual behavioral variability. The latter deals with instance-to-instance variability observed in a particular individual, while the former tells us about the whole range of different forms of expression in a complete species. In the context of evolution, inter-individual variability serves as the basis for natural selection to work on. In other words: without different versions of genes, phenotypes, strategies, and behaviors in a species, adaptation to changing environmental conditions is fairly difficult, if not impossible (Roseman, 2020). Another fundamental differentiation can be seen between systematic and unsystematic, i.e. random, behavioral variability. Systematic variability means that changing parameters are correlated with other internal or external parameters. Differentiation between systematic and unsystematic variability can be difficult, because not all potentially influential parameters can be measured in a single experiment, thus correlation with an unknown parameter cannot be excluded.

Unsystematic behavioral variability, to some extent, results from noise in the nervous system, challenging animals to develop not only efficient but also robust and stable solutions for the expression of behavior (Faisal et al., 2008). This especially needs to be considered when we try to model behaviors, because computers work deterministically and tend to find highly effective solutions which may break apart if only a single parameter changes slightly. Trial-to-trial variability can also result from fatigue (Jones, 2016) or changing inner states of the animal (Pflüger, 1999). In general, the many different causes of variability make it difficult to determine the exact contribution of each factor.

Inter-individual variability emerges from genetic and epigenetic differences (Yamamoto et al., 2024), diet (Kraus et al., 2022), or social and developmental imprinting (Hayden et al., 2020). Previous studies have highlighted that behavioral idiosyncrasies, including aspects like left-right preferences in a Y-maze or locomotor handedness, can be partly hereditary (Ayroles et al., 2015; Buchanan et al., 2015). Furthermore, even isogenic individuals reared in identical

conditions will vary to some extent. In addition to differences in the mean form of expression, individuals can also vary in their degree of intra-individual variability or the probability of expressing the behavior at all. However, variability affects all interactions with the environment and thus plays a fundamental part in both evolution and individual development. Therefore, it should be taken seriously as an important aspect of biological phenomena in general and behavior in particular.

In this thesis, I investigated the different dimensions of natural variability in *Drosophila* walking behavior by analyzing detailed kinematic data from a large cohort of wild-type flies. The basis for this study were 2D leg-tip kinematics of freely walking flies which were obtained by employing high-speed videography and a deep learning-based automated tracking method (Mathis et al., 2018). I applied principal component analysis (PCA, introduced in chapter 2) to unbiasedly explore the covariant patterns in the data set, identifying key components that describe both shared and individual-specific aspects of movement and variability. I show that the individual-specific aspects remain largely constant across different walking speeds, while the contribution of general coordination patterns increased at higher speeds, supporting the notion that faster walking is accompanied by stricter coordination (chapter 2). This reduction of variability in leg-tip kinematics with increasing walking speeds is then focused by systematically analyzing the relationship between walking speed, spatial and temporal variability, and interleg coordination. The results suggest a causal connection between the synchronization of leg movements and variability, while walking speed seems to be more indirectly correlated (chapter 3). Finally, I introduce an advanced setup for recording freely-walking fruit flies enabling continuous high-resolution recordings at 200 Hz for multiple hours. The capabilities of this new setup and the data processing software specifically designed for it are demonstrated in a curved-walking analysis (chapter 4).

2 Chapter 2: Natural variability and individuality of walking behavior in *Drosophila*

The following chapter consists of a manuscript which is in the third round of revision in the *Journal of Experimental Biology*, last submission on September 6th, 2024

Authors

Vincent Godesberg, Till Bockemühl, and Ansgar Büschges

Author contributions

V.G. carried out experiments. V.G. analyzed data and created figures. V.G., T.B., and A.B. conceptualized the study. V.G., T.B., and A.B. wrote the manuscript. A.B. provided funding.

2.1 Abstract

Insects use walking behavior in a large number of contexts, such as exploration, foraging, escape and pursuit, or migration. A lot is known on how nervous systems produce this behavior in general and also how certain parameters vary with regard to walking direction or speed, for instance. An aspect that has not received much attention is if and how walking behavior varies across individuals of a particular species. To address this, we created a large corpus of kinematic walking data of many individuals of the fruit fly *Drosophila*. We only selected instances of straight walking in a narrow range of walking speeds to minimize the influence of high-level parameters, like turning and walking speed, aiming to uncover more subtle aspects of variability. Using high-speed videography and automated annotation we captured the positions of the six leg tips for thousands of steps and used principal components analysis to characterize the postural space individuals used during walking. Our analysis shows that the largest part of walking kinematics can be described by five principal components (PCs). Separation of these five PCs into a 2-dimensional and a 3-dimensional subspace divided the description of walking behavior into invariant features shared across individuals and features that relate to the specifics of individuals; the latter features can be regarded as idiosyncrasies. We also demonstrate that this approach can detect the effects of experimental interventions in an unbiased manner and that general aspects of individuality, such as the individual walking posture, can be described.

2.2 Introduction

Legged locomotion, commonly subsumed under the term walking, is found in most terrestrial animals. Walking is used in a diverse set of behavioral contexts, such as exploration and foraging, escape, pursuit, mating, and migration, making it a central component of an animal's behavioral repertoire. The diversity of these contexts requires walking to be highly adaptable and flexible. Consequently, the task- and situation specific neuronal control of walking behavior is important role for its proper execution.

The neuronal control and kinematics of walking have been extensively studied in a large variety of arthropods (Nirody, 2021; Nirody, 2023), particularly in insects, from small insects, like fruit flies (*Drosophila melanogaster,* DeAngelis et al., 2019; Mendes et al., 2013; Mendes et al., 2014; Strauß and Heisenberg, 1990; Wosnitza et al., 2012) or desert ants (*Cataglyphis,* Pfeffer et al., 2019; Wahl et al., 2015; Zollikofer, 1994), to large ones, like cockroaches (*Periplaneta americana,* Couzin-Fuchs et al., 2015; Delcomyn, 1971; Delcomyn, 1989), locusts (*Schistocerca gregaria,* Burns, 1973; Niven et al., 2010; Pearson and Franklin, 1984), or stick insects (*Carausius morosus,* Cruse, 1976; Dallmann et al., 2016; Dürr and Ebeling, 2005; Gruhn et al., 2009). There is a large body of knowledge in these groups, ranging from the sensorimotor control of individual legs, on how interleg coordination of the six legs is achieved, to high-level descending and central neuronal control (Bidaye et al., 2018; Cruse, 1990; Dürr et al., 2004). A number of common features governing the neuronal and kinematic aspects of insect walking have been identified. Walking speed, as a major aspect, seems to be generally controlled by changes in stance duration, while stance amplitude and swing duration are largely kept constant (DeAngelis et al., 2019; Wosnitza et al., 2012). These changes in stance duration are accompanied by systematic changes in interleg coordination. Unlike larger vertebrates, however, which use distinct gaits in a speed-dependent manner (Diedrich and Warren, 1995; Hoyt and Taylor, 1981), insects exhibit a continuum of interleg coordination patterns (Szczecinski et al., 2018; Wosnitza et al., 2012). These systematic effects and other invariant features are generally present in walking insects as an evolved phenotypic trait; however, the neuronal control and kinematics of walking must have exhibited hereditary

interindividual and inter-species variability during evolution, thereby adapting to co-evolving traits like body morphology or ecological demands. Indeed, a previous study on the evolution of walking behavior in a large set of drosophilids showed that there exist systematic differences in walking behavior across species and strains. These differences evolve and diverge rapidly in closely-related species, but re-converge to shared features in more distantly-related ones (York et al., 2022). However, while this previous study establishes an evolutionary approach to investigate the variability of walking in insects, it was mainly based on the characteristics of the velocities at which different species walk and did not focus more specifically on detailed leg kinematics. Thus, how variable walking is within and between individuals of a given species or strain on a level closer to the actual motor output is still largely unexplored.

Walking behavior differs in multiple parameters between individual flies, such as average posture, preferred coordination pattern, walking speeds, or degree of intraindividual variability, as anecdotal observations indicated. However, a quantitative description of these differences is still lacking. To explore this aspect in greater detail, in the present study we wanted to explicitly acknowledge variability as an important aspect of walking behavior on the inter- and intra-individual level and to find an unbiased and more comprehensive way of characterizing and interpreting the observed variability. Behavioral variability on the individual level is a topic that has started to receive more attention in recent years. For instance, two studies on handedness in *Drosophila* show that individual flies can have left/right preferences in a Y-maze decision paradigm and that these preferences are partly hereditary (Ayroles et al., 2015; Buchanan et al., 2015). A study on several *Drosophila* wildtype strains established systematic differences in odor preference that can be traced back to phenotypic differences related to these strains (Ruebenbauer et al., 2008). Further studies in *Drosophila* showed that grooming behavior shows variability between different wildtype lab strains, but also between isogenic individuals of the same strain (Mueller et al., 2022), and that the kinematics of trajectories during object orientation are highly individualized and that neuronal asymmetries drive this behavioral individuality (Linneweber et al., 2020).

Here, we investigated the natural intra- and interindividual variability of low-level kinematic parameters of walking in the fruit fly *Drosophila melanogaster*. To control for and exclude known influences of walking speed and curve walking on kinematics, we initially focused on straight walking at intermediate speeds in a large set of male flies, each of which spontaneously produced a large corpus of walking behavior in an unrestrained free-walking paradigm. Using high-speed video recording and automated annotation based on deep learning methods (DeepLabCut, DLC Mathis et al., 2018), we extracted the positions of two body markers, as well as the tarsal tips from all video frames in these straight sequences, automatically determined positions and times of lift-off and touch-down events of the legs, and calculated walking speed and interleg coordination for all step cycles. Across individuals, the data set we created in this way contained more than 36,000 steps for each leg. We used principal components analysis (PCA) to find a compact description of this large data set and systematically explored correlations and the variability between leg kinematics an individual basis as well as across individuals. In the field of motor control, biomechanics, and kinematics PCA has been used to extract motor synergies in the context of cortical control of hand movements in monkeys (Mollazadeh et al., 2014) and joint-angle correlations of targeted catching movements in humans (Bockemühl et al., 2010), to detect altered kinematic profiles in stroke patients (Milovanović and Popović, 2012), to characterize recovery of locomotor function after spinal injury in mice (Takeoka et al., 2014; Takeoka and Arber, 2019), or to evaluate the complexity of wing kinematics in bats (Riskin et al., 2008). In *Drosophila* locomotion, PCA has been applied, for instance, to evaluate the effect of neurotoxins on a large number of kinematic parameters during walking (Cabrita et al., 2022). Furthermore, an approach mathematically related to PCA has been used to extract a set of wing movement patterns that allow for flight control in *Drosophila* (Chakraborty et al., 2015).

Here, PCA revealed that most of the kinematic variability in our data set (approx. 80%) is contained in the first five PCs. We show that two subsets of these five PCs describe interindividually applicable dynamics of interleg coordination, on the one hand, and individual characteristics of walking behavior, what might be called idiosyncrasies, on the other. The first

subset contains two PCs which mainly capture interleg coordination-specific aspects of walking – how tripod-like a particular movement pattern is or the general repetitive sequence of alternating swing and stance movements of individual legs. In contrast, the contribution of a second subset of three PCs relates to how individuals differ from each other in the way they walk; data for different individuals occupy different regions within this PC subspace, highlighting interindividual differences. Simultaneously, the same flies are indistinguishable in the subspace related to interleg coordination, supporting the notion that these two PCs describe universal aspects of intra-individual variability. The importance and applicability of these two subsets of PCs is further substantiated by (1) relating them to a quantitative measure of tripod coordination strength (TCS), (2) showing that individual-specific contributions to walking remain constant with regard to walking speed, (3) a use case for characterizing changes in walking behavior induced by optogenetic inhibition of sensory structures in the legs, and, finally, (4) deriving a two-dimensional measure for postural adaptations. Our results suggest that the variability observed in walking flies is systematic and that PCA is a suitable approach for the quantification of and decomposition into idiosyncrasies, inter-leg coordination patterns, and effects of experimental interventions.

2.3 Materials and Methods

2.3.1 Fly strains and husbandry

Male flies of the wild-type strain Berlin-K (Bloomington Drosophila Stock Center (BDSC, #8522) were used for those experiments which formed the basis for PCA (see below). Inhibition experiments (see below) were performed with F1 flies resulting from crosses between *iav*-Gal4 (O'Dell and Burnet, 1988) (BDSC #52273) and UAS-GtACR1 (Govorunova et al., 2015; Mohammad et al., 2017) (BDSC #92983). Both, the use of the wildtype strain Berlin-K as well as the *iav*-Gal4 line, are owed to the fact that we wanted to have a direct connection with the data and results from a previous study (Chockley et al., 2022).

Wild-type flies were raised on 12h/12h light/dark cycle, while transgenic flies were raised in the dark to prevent premature activation of GtACR1 channels and potential adaptation prior to experiments. All flies were kept at 25°C and approximately 60% humidity on a standard food medium (Backhaus et al., 1984). To improve the function of GtACR1, transgenic flies had 60 µg of all-trans-retinal in their food for at least three days prior to the experiments. All animals used in this study were five days old. To increase walking activity, flies were isolated and starved for approximately 24 hours before experiments, but had access to moist tissue paper during this period.

2.3.2 Experimental setup

The experimental setup described here is largely identical to the one used in a previous study (Chockley et al., 2022). However, for clarity we describe it in detail here again. The recording arena (Fig. 1A) consisted of an inverted glass petri dish (diameter: 60 mm) as the walking substrate and a watch glass (diameter: 100 mm) as the lid. This arrangement formed a closed chamber with a curved dome tapered towards the edge of the petri dish, similar to an inverted FlyBowl (Simon and Dickinson, 2010). The inner side of the watch glass was coated with SigmaCote (SL2, Sigma-Aldrich, St. Louis, Missouri, USA). This resulted in a hydrophobic surface on which flies found less grip; walking on the ceiling was thus reduced. The petri dish

and the watch glass were placed in a custom-made plastic holder with a cutout that allowed for video recordings from below (holder not shown in Fig. 1A).

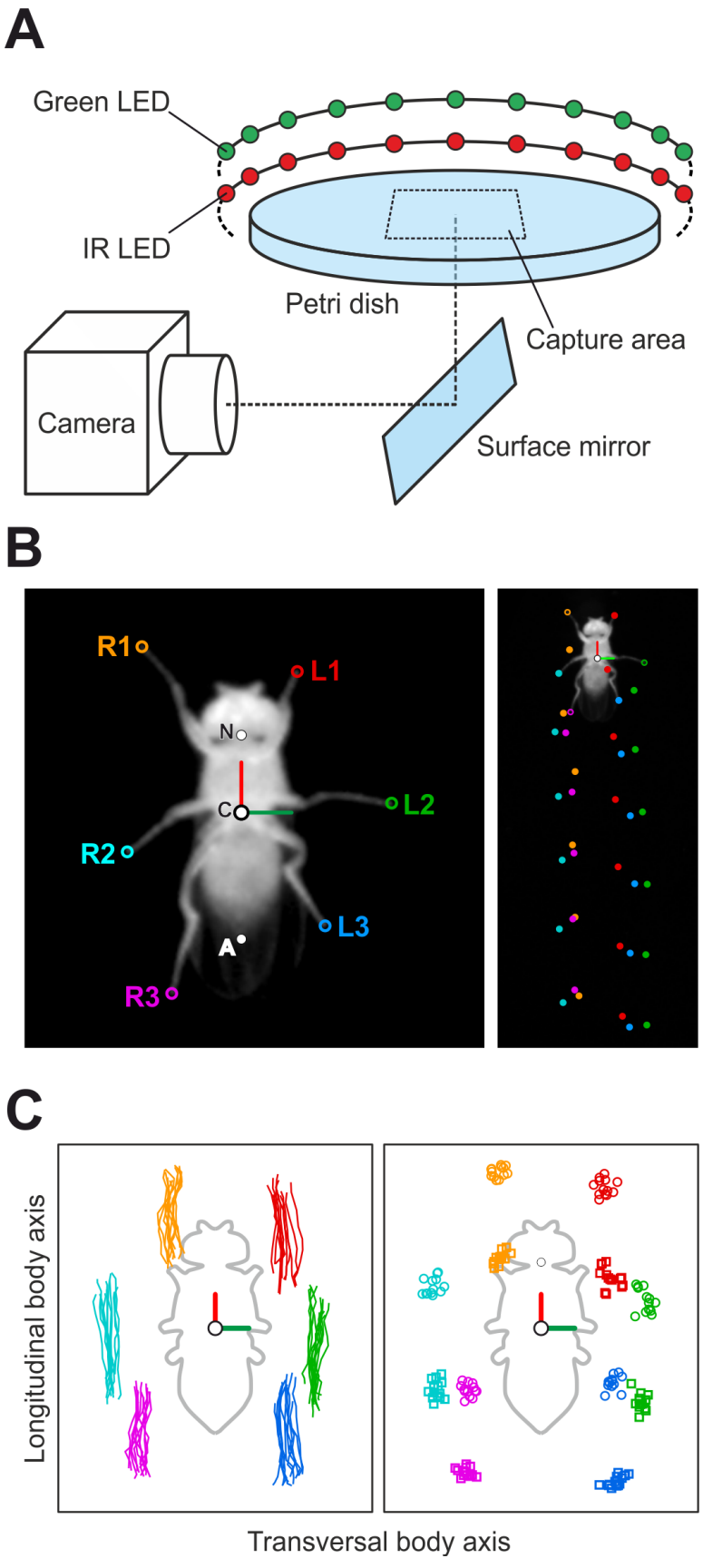


Figure 1: Setup and data acquisition. (A) Schematic of the experimental setup. The watch glass covering the petri dish is not depicted. (B) Examples of the body-centered bottom view (left) and the camera view (right) with annotated leg tips (right legs R1 - R3, and left legs L1 - L3), abdomen (AT), center of the body (C) and neck (N). (C) Examples of stance trajectories (left) and tarsus extreme positions (right) of one walking bout after detection of lift-off and touch-down positions. Round markers: AEPs, square markers: PEPs. Fly silhouette (gray outline) for reference.

To record flies walking on the glass substrate we used a camera (model: VC-2MC-M340, Vieworks, Anyang, Republic of South Korea) equipped with an object-space telecentric lens (focal length 55 mm, model: Computar TEC-55, CBC America, Cary, North Carolina, USA). The telecentric lens provided an orthographic projection, reducing image position-dependent changes in apparent fly posture. The camera was located on the side of the setup and its view was directed at the experiment chamber from below via a surface mirror tilted at an angle of 45°. To increase video resolution the camera view was focused on a square area in the center of the walking substrate (30 mm side length, resolution 1000 by 1000 pixels, 33.3 pixels mm^{-1}). The typical body length (BL, defined as the distance between the anterior end of the head and the tip of the abdomen) of a fly of approximately 2 mm corresponded therefore to 65 - 70 pixels; this was sufficient to image all body parts annotated during analysis with sufficient accuracy, particularly the tarsal tips (Fig. 1B, left). The scene was illuminated with 60 infrared (IR) LEDs (wavelength: 890 nm, opening angle: 20°) arranged in a concentric ring around the chamber. Light from these LEDs was emitted mainly parallel to the arena surface. Thus, only the fly reflected any appreciable amount of light, resulting in a strong contrast between it and the background (see Fig. 1B). Since ambient light during experiments was generally low, flies walked largely in the dark. However, contrast was further enhanced by adding a visible-light filter (upper cut-off frequency 790 nm) to the camera's lens, eliminating any remaining ambient light from the room or optogenetic inhibition (see below). Video data was acquired at 200 Hz and a shutter time of 400 µs; this low shutter time prevented motion blur and ensured detectability of the leg tips even during very fast movements. Acquisition of individual video frames during this shutter time and IR illumination were synchronized with a pulse generator. For inhibition experiments (see below) we added a second LED ring around the chamber. This

ring consisted of 60 green LEDs (wavelength: 525 nm) whose light was directed at the recording chamber. These LEDs could be switched on and off programmatically during an ongoing experiment via a multi-function I/O device (USB-6001, National Instruments Corporation, Austin, TX). The experimental setup was custom-designed and built (Electronics workshop, Department of Animal Physiology, University of Cologne).

2.3.3 Behavioral experiments

Prior to an experiment, single flies were aspirated into a tube and then transferred into the arena; no CO_2 was used for this step. Flies were then allowed to acclimate to the setup for 5 minutes prior to the start of video acquisition. The total duration of an experiment was up to 3 hours. During this time, flies walked spontaneously in the chamber and frequently crossed the capture area. For the complete duration of an experiment, video data was acquired continuously and the last 1000 frames (equivalent to 5 seconds) were stored in a ring buffer. Custom-written software functions evaluated the recorded frames in real-time and determined if the fly was present in a particular video frame and if it had produced a continuous walking track with a minimum length of 3 BL and a minimum walking speed of 2 BL s^{-1}. Once the fly had produced such a track and then stopped or left the capture area the contents of the frame buffer were committed to storage as a valid trial. After this, acquisition automatically started anew. Note, that at this time no additional selection criteria (such as curvature or walking speeds, for instance) were applied to determine validity of a trial. The set of videos acquired in this way merely served as a large and relatively unconstrained initial set of walking behavior (see next section on further criteria for data inclusion).

The inhibition experiments with *iav*-Gal4>UAS-GtACR1 used for evaluation of the descriptive power of PCA (see below) were performed in the same setup. Technically, these experiments are a replication of those conducted in one of our previous studies (Chockley et al., 2022, see their Fig. 3). In this previous study, we confirmed that the wildtype strain Berlin-K used in the present study did not show strong changes in leg kinematics simply due to exposure to green light (Fig. S2 in Chockley et al., 2022). We therefore restricted our inhibition experiments in the

present study to *iav*-Gal4>UAS-GtACR1 and extended the experiments as follows. Flies in the inhibition experiments either walked in the dark (wild-type control condition) or under green-light illumination (inhibition condition). Trials in the dark and in the light were alternated; once the fly had produced a trial in the dark, the green-light illumination was switched on and the system was primed to record the next trial. If the fly did not produce a valid trial within the first 60 seconds of the green-light condition, the light was switched off and data acquisition was suspended for 30 seconds to prevent adaptation to the inhibition and potentially harmful effects of prolonged exposure to the green light. After this cooldown period, the green light was switched on again, and data acquisition was resumed. Once a trial had eventually been recorded in the inhibition condition, the light was switched off again and the process was repeated. Data from these inhibition experiments were sorted into control trials (recorded in the dark) and inhibition trials (recorded during green-light illumination). Video acquisition, online data evaluation during experiments, and high-level hardware control were implemented with custom-written software in MATLAB (2018b, The Mathworks, Natick, Massachusetts, USA).

2.3.4 Processing of video data

Because curve walking has a strong influence on leg kinematics in insects (Dürr and Ebeling, 2005; Gruhn et al., 2008; Jander, 1985) we initially restricted our analyses to straight walking. Instances of this in the complete set of walking trials (see previous section) were detected with custom-written algorithms, whose parameters were determined empirically (see Supplementary Material). Only segments of video data associated with straight walking were extracted from valid trials. In each video frame of these segments eight different body parts (the tarsal tips of all legs, the neck, and the posterior tip of the abdomen) were detected with DeepLabCut (DLC, Mathis et al., 2018, Fig. 1B). To improve robustness of DLC we first detected the general position of the fly within a video frame using a threshold operation, conversion to a binary image, and, finally, calculation of the centroid of the largest contiguous region of white pixels (which corresponded to the fly). Using this position, we then cropped the

fly from the video. With these cropped and fly-centered views we used three different instances of DLC in a two-step analysis. The first step was to detect the neck and abdomen of the fly. These positions were used to define a fly-centric coordinate system (Fig. 1C); all data presented in this study (apart from the identification of swing and stance phases) are based on these body-centric coordinates. The positions of all body parts were then normalized to each fly's body length to allow for body size-independent comparisons between flies. We also used the neck and abdomen positions to rotate the cropped views and align the fly's longitudinal axis vertically. These cropped and rotated data were then used in the second DLC analysis, in which the tarsal tips of each body side were detected by two independently-trained instances of DLC, one for the left and one for the right legs. In general, DLC performance was very good; to ensure highest accuracy, however, we also visually inspected all automatically generated annotations for errors and corrected these manually, where necessary.

Swing and stance phases of all legs were determined automatically based on the respective speeds at which tarsal tips moved in an arena-centric coordinate system: whenever a tarsus is stationary in this coordinate system, i.e. it co-moves with the ground, we assumed the leg to be in stance phase (Fig. 1B, right). Conversely, movements of more than 1.5 pixels per frame were empirically defined as swing phase activity (see also Supplementary Material). A transition between stance and swing phase was defined as lift-off event. The last position of the tarsal tip on the ground before lift-off was defined as the posterior extreme position (PEP) for that step (Fig. 1C, right). Conversely, a transition between swing and stance phase was defined as touch-down event; the first tarsal position with ground contact associated with this event was identified as the anterior extreme position (AEP) (Fig. 1C, right). The time of onset of a particular step was defined as its lift-off, and a complete step of a leg was defined as its movement between two consecutive lift-off events, i.e. a swing phase followed by a stance phase. In contrast to steps of individual legs, step cycles (SCs) were defined as follows: start and end of an SC were determined by the respective step of the right middle leg, which was selected arbitrarily for this purpose. All six tarsal tip positions for this interval comprised the data of one SC. Since the stepping period of the six legs in straight walking flies is almost

identical and constant for small time windows, each leg completes its own cycle during an SC, although they all start and end at individual positions and phases, respectively. In other words, all six leg tip positions at the beginning and the end of an SC are usually highly similar, not just for the reference leg.

The walking speed associated with a step or an SC was defined as the average walking speed of the animal between on- and offset and was used to allow for the selection of steps and SCs within a certain range of walking speeds for analysis. The main body of data used in the present study was based on steps and SCs whose associated walking speed was between 5 and 7 BL s^{-1} (see Fig. S1 for all speed ranges of all individuals). Initially, we restricted the range of walking speeds in this way to facilitate comparability between individuals. Furthermore, previous studies have shown that walking speed has a strong and systematic influence on many of the kinematic parameters investigated here (Mendes et al., 2013; Strauß and Heisenberg, 1990; Wosnitza et al., 2012); this general influence might have an unwanted effect on the initial analysis if individuals walk at different preferred speeds. However, for a later analysis we expanded the range of walking speeds to lower and higher speeds (see below).

2.3.5 Principal component analysis (PCA)

PCA is a tool for dimensionality reduction and can be used to find linear correlations of multiple parameters in high-dimensional data sets. Mathematically, calculating PCA is identical to finding the eigenvectors and eigenvalues of a data set's covariance matrix (Manly and Alberto, 2019). The principal components (PCs) form a new coordinate system that is aligned with the directions of highest variability in the data set. Importantly, PCs are ordered according to their relevance, i.e. the relative amount of variance they describe.

To evaluate whether a given PC describes a fraction of variability that is meaningful, i.e. larger than what would be expected by chance, a reference data matrix can be constructed that contains the original data but in which the columns (corresponding to the individual variables) are randomly and independently permuted. This removes any correlations between rows (individual observations). Repeating the PCA for this randomized data set will give eigenvalues

for each PC that can serve as reference levels whose values need to be exceeded for a PC of the original data to be identified as meaningful.

We applied PCA to the positions of all six leg tips during SCs for 88 individual flies. 30 SCs between 5 and 7 BL s^{-1} for each of these 88 flies were used as the basis for PCA. Resampling and interpolation were used to acquire exactly 100 postures for each SC in the analysis, resulting in matrices of 3000-by-12 data points per fly, with each row representing x and y-coordinates of the six leg tips and a total of 30 (SCs) times 100 (normalized number of data points per cycle) rows. The final data matrix for PCA comprising data from all flies contained 2,640 SCs, represented by 264,000 data points with 12 parameters each. Prior to the analysis, each column (equivalent to one parameter) was standardized to a mean of 0 and unit variance (equivalent to z-scores). PCA was carried out on this standardized matrix in MATLAB 2018b (function *pca.m*). To establish reference levels for meaningful contributions of PCs, we used the same data in a randomized second PCA as outlined above.

Here, PCs describe spatial covariations of the positions of all six tarsal tips and can readily be interpreted as movements of the tarsal positions when multiplied by a non-zero factor. Using this fact, individual PCs were visualized by varying their values systematically from minus 2 to plus 2 of their respective standard deviations before transferring the data back into the original parameter space. The resulting positional changes are depicted as arrows to indicate the direction of covariation. In mathematical terms, these arrows correspond to the *loadings* of each PC. For subsequent data analysis, the original data was transformed either in its entirety (264,000-by-12 matrix) or on a per-fly basis (3000-by-12) into the new coordinate system established by the PCs. In the context of PCA, these transformed data are typically also referred to as *scores*.

2.3.6 PCs and tripod coordination strength (TCS)

The fraction of variability described by a PC for one complete cycle of both tripod groups was compared to the respective coordination pattern. Tripod coordination strength (TCS), as used in previous studies (e.g. Ramdya et al., 2017; Wahl et al., 2015; Wosnitza et al., 2012), was

used to quantify the synchronicity of the swing activity in the two tripod groups. TCS was calculated as follows: for each tripod group (a set of ipsilateral front and hind legs and the contralateral middle leg), the time in which all three legs were simultaneously in swing phase was divided by the time from the earliest swing onset to the latest swing termination in any of these three legs. Hence, a perfect overlap of all three swing phases resulted in a maximal TCS value of 1 and would correspond to canonical tripod coordination. The minimal TCS value of 0 was assigned in cases of no overlap between swing phases in a tripod group. The TCS values of the tripod groups were averaged and the fractions of variability described by each principal component for all positions during the movements of these two tripod groups were calculated. This was achieved by transferring all leg tip positions in this respective time frame into the principal component space and measuring the fraction of variability described by each PC.

2.3.7 Analysis of different walking speeds and time of recording

To test how the results of our PCA analysis change for different walking speeds we selected a lower (2 to 4 BL s^{-1}) and a higher (8 to 10 BL s^{-1}) speed range and transformed data from these speed ranges into the PC space we established for 5 to 7 BL s^{-1}. We only allowed flies with 30 or more steps in a given speed range to be part of the analyses. Both additional data sets contained flies which had not been part of the data set of 5 to 7 BL s^{-1} and vice versa. For all three datasets we calculated the fraction of relative variability described by each principal component. In addition, we compared how the walking behavior of all individual flies changed with walking speeds in the dimensions of PCs 1 to 5 regarding their mean positions and standard deviations. A similar approach was taken with regard to the time a particular trial took place during an experiment (early vs. late trials, for more detailed methods and results see Supplementary Information).

2.3.8 Evaluation experiments

To test the suitability of the PCA based approach for the description and analysis of idiosyncrasies in *Drosophila* walking behavior we used an optogenetic approach. For this, the Gal4-UAS-system was used to express GtACR1, an anion-selective channelrhodopsin, in a

group of mechanosensory neurons in the legs. When activated optogenetically with green light, GtACR1 inhibits neurons expressing it (Govorunova et al., 2015; Mohammad et al., 2017). We used the transgenic *iav*-Gal4 line (O'Dell and Burnet, 1988) to target all chordotonal organs, including the femoral chordotonal organ (fCO), the largest sensory organ in the fly's legs. The resulting transgenic flies have previously been shown to exhibit a systematic and noticeable phenotype in walking behavior in an inhibition paradigm (Chockley et al., 2022). In contrast, wildtype control flies in that study (wildtype Berlin-K crossed to UAS-GtACR1) did not show a strong phenotype during phases of illumination with green light (see Supplementary Figure S2 in Chockley et al., 2022). Replicating this previous study, flies walked spontaneously in the arena, initially in the dark (control condition, identical to the way in which the main data set in the present study was collected). Once a fly had produced a valid trial, the green-light illumination was switched on (inhibition condition). After a valid trial was produced in the inhibition condition, the green-light illumination was switched off again; conditions were alternated in this way for the duration of the experiment.

We used these data to evaluate whether we can detect the known and strong fCO inhibition-specific effects in the PCA approach explored here and whether we could distinguish them reliably from putatively smaller stochastic effects based on general variability in walking behavior (these exist even between two randomly chosen samples from a larger data set). For this, a minimum of 30 SCs each for dark (control) and light condition (inhibition) for individual flies was compared regarding the mean tarsal trajectories and the shift observed in the respective mean positions in PCs 2, 4, and 5. We tested the observed effect size for significance by comparing it to a bootstrap analysis performed on our original wildtype data set we established earlier for the speed range of 5 to 7 BL s^{-1}: for 14 individual flies we randomly selected two unique sets of 30 SCs each and compared these two sets with each other in the same way, we compared the two conditions (control and inhibition) for the transgenic flies.

2.3.9 Symmetry axis in subspace of PCs 2, 4, and 5

The evidence that PCs 2, 4, and 5 describe interindividual differences in leg kinematics and posture (see Results section) suggests that these PCs describe more general differences in the mean posture of individual flies. Among other aspects, this might refer to how sprawled the posture of an animal is or individual-specific distances between AEPs and PEPs, for instance. To give a comprehensible example, we systematically searched for an axis in the subspace of PCs 2, 4, and 5 along which the posture changes symmetrically with respect to the left and right body side. This axis was supposed to go through the center of the three-dimensional subspace and respective postures were calculated from -5 to +5 times the standard deviation. Symmetry was measured by calculating the RMSE between contralateral leg-pairs after projecting the positions of one body side to the other. The axis with the highest symmetry scores and the respective postures are shown in Figure 7.

2.4 Results

We recorded 103 male flies of the wild-type strain Berlin-K with a total of 36,942 straight walking step cycles (SCs) with a median of 242 SCs per fly with 10% and 90% percentiles of 93.8 and 841 SCs, respectively (min: 28; max: 1717). 88 flies yielded 30 or more SCs within the speed range of 5 to 7 BL s^{-1} targeted in our analysis and thus were included in the PCA (for details see Supp. Fig. S1). Restriction to this speed range was important, because many parameters of walking, like duty cycle, stepping frequency, step amplitude, or phase relationships, change systematically with walking speed. During evaluation experiments, we recorded 28 flies resulting from crosses between *iav*-Gal4 and UAS-GtACR1. They performed a total of 9,694 SCs in the control and inhibition condition with a median of 279.5 SCs per fly with 10% and 90% percentiles of 108.9 and 726.6 SCs, respectively (min: 43; max: 918). 14 of these flies produced 30 or more SCs within the speed range of 5 to 7 BL s^{-1} for both conditions and were included in the analysis.

2.4.1 Flies walk in an idiosyncratic manner

Despite being all male, of identical age, from the same highly interbred wildtype strain, and tested under identical conditions, the flies analyzed here showed idiosyncrasies which are already apparent in qualitative inspection. To first illustrate these qualitative aspects and give a first exemplary impression of these idiosyncrasies that we would like to investigate further, Figure 2 shows AEPs and PEPs (Fig. 2Ai - Aiii) as well as leg tip trajectories (Fig. 2Bi - Biii) of three exemplary flies over the course of three individual walking bouts. Several distinguishing kinematic characteristics can be identified, such as (1) the overall width of the complete posture or leg pairs (compare Fig. 2Bi and Biii, for instance), (2) the straightness and shape of tarsal trajectories, or (3) left-right symmetry. Qualitatively, it is also evident that intra-individual step-to-step (e.g. distance between AEPs and PEPs in one leg) variability is lower as compared to interindividual variability (i.e. difference in posture or asymmetries between Fig. 2Bi and Biii). These observations show that flies indeed walked idiosyncratically.

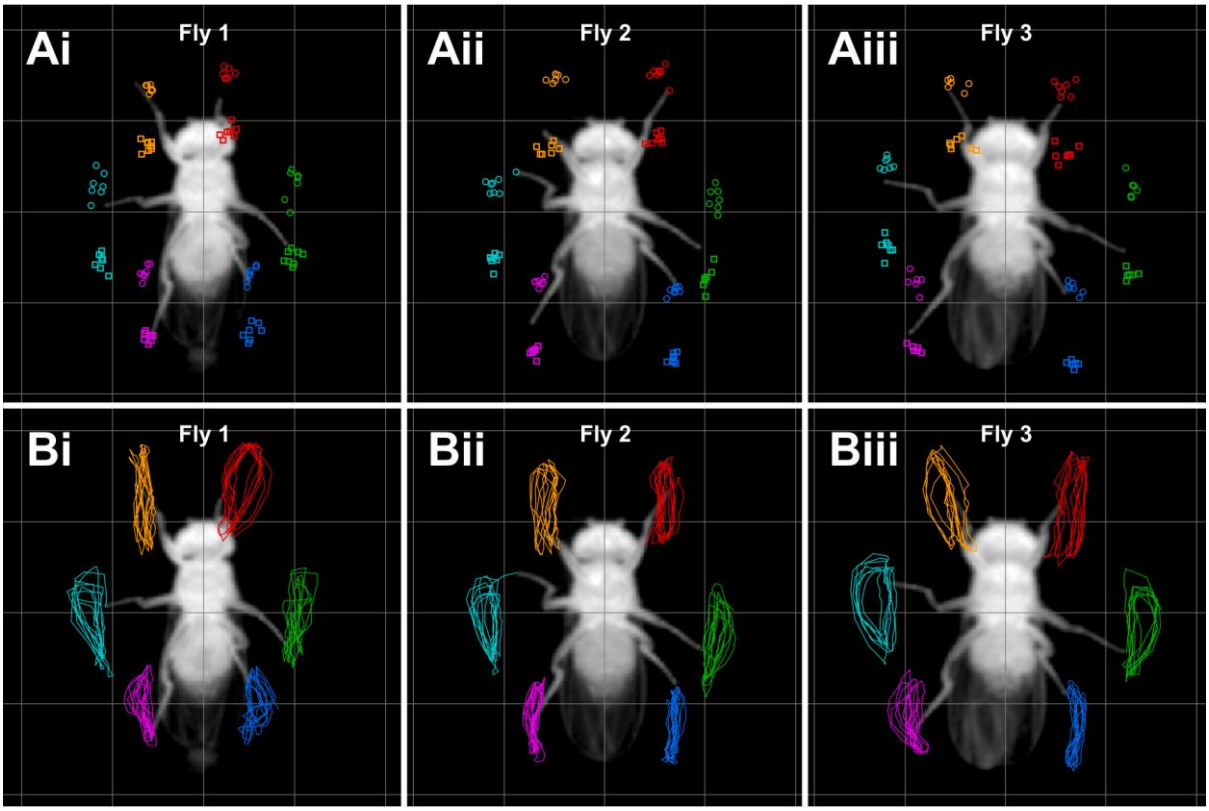


Figure 2: Qualitative inspection shows flies walk in an idiosyncratic manner. Here, interindividual variability of leg kinematics in three exemplary individuals is shown Sub-indices i to iii refer to three different animals, respectively; each column is a different fly. (A) Anterior and posterior extreme positions (AEPs and PEPs) of single walking bouts for three exemplary flies. Round markers: AEPs, square markers: PEPs. (B) Complete tarsal tip trajectories of the same walking bouts and flies as in A. Leg color coding: red = L1; green = L2; blue = L3; orange = R1; teal = R2; magenta = R3. These three flies are also referenced as exemplary flies in Fig. 5 and Fig. 7.

2.4.2 PCA detects significant correlations of leg tip positions

We used PCA to characterize linear covariations of leg tip movements. Such covariations have either a systematic biological basis or are random fluctuations. Here, the fraction of variability described by each PC for the randomized data set approximated the inverse of the number of dimensions (~8.33%). In contrast, fractions of variability described by PCs 1 to 5 for the original and unpermuted data set exceeded this reference value, accounting for more variability than expected for randomized data (Fig. 3A) and described a cumulative ~78% of the variability. For further analyses we therefore focused on these first five PCs.

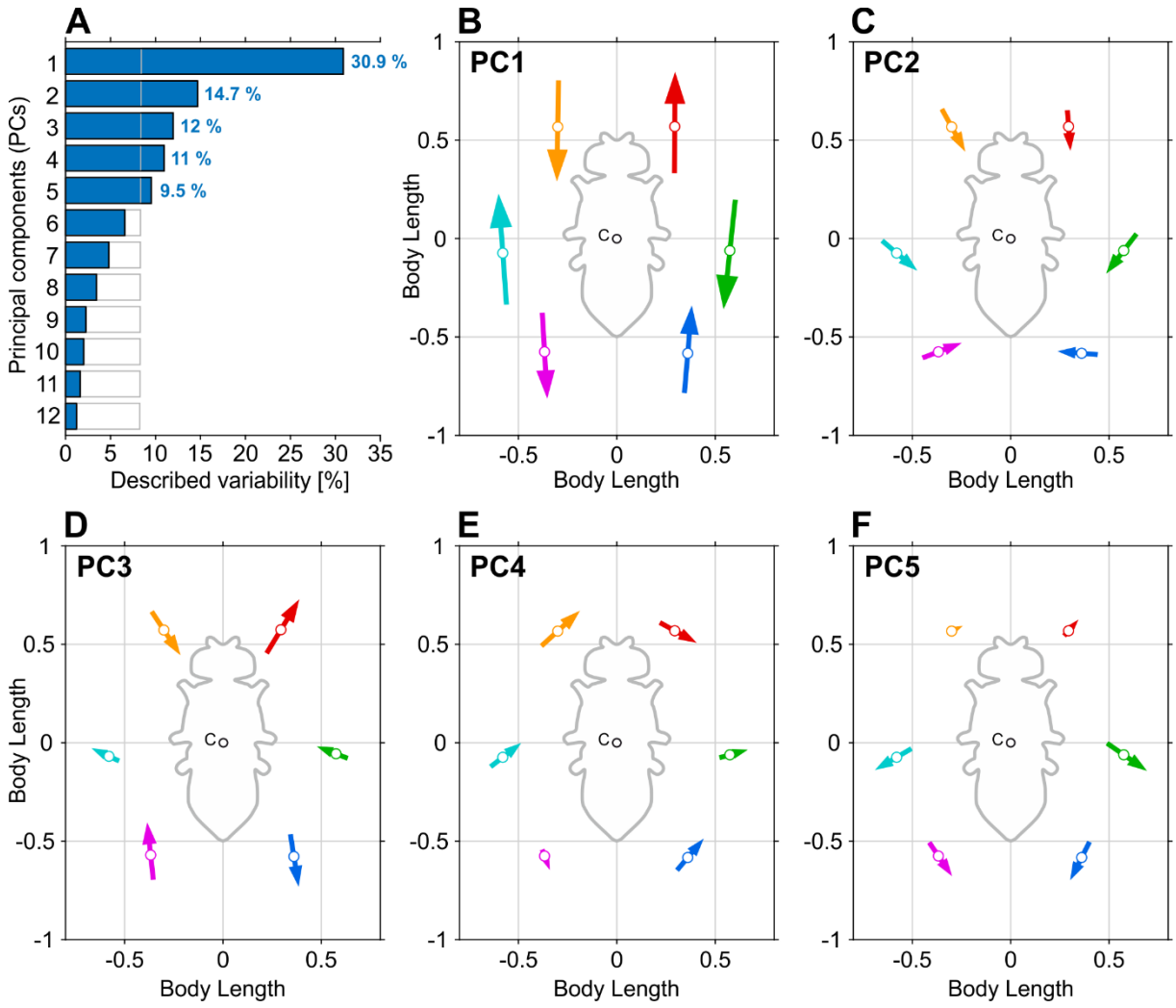


Figure 3: PCA finds significant correlations of leg tip positions. (A) Fractions of variability in leg tip kinematics described by each principal component. Filled blue bars: results based on PCA on the data set acquired in this study (number of flies: 88, number of steps per fly: 30). Gray open bars: results for the same, but randomly permuted data set (see results section). PCs 1 to 5 exceed the described variability expected based on the permuted data set (respective percentages shown at the bars). (B) to (F) Visualization of covariation of tarsal tip kinematics captured by PCs 1 to 5, respectively. White dots on arrows depict the mean positions of all data points in the analysis. Arrow directions indicate the directions for a particular PC. Arrow lengths (magnitude) equal twice the standard deviation. As such, the contribution of each PC to the positioning of individual leg tips can be directly compared. C: center of the body. Leg color coding: red = L1; green = L2; blue = L3; orange = R1; teal = R2; magenta = R3.

Here, each PC's loadings intuitively describe how 2D positions of the tarsal tips covary. This covariation can be a result of active leg movements during walking, i.e. protractions and retractions. Covariation in the two front legs, for instance, can be expected, because during

anterograde movement of one front leg the other one generally moves posterograde. Further covariation is based on general postural difference between individuals; how sprawled an individual's posture is, for instance, will affect its six legs fairly evenly and will be detected as positional covariation that describes the distance of the tarsi from the body. Individual PCs (i.e. their loadings) can be visualized as arrows to facilitate the comparison of the direction and magnitude (according to the amount of described variability) of covariations (Fig. 3B - F). PC 1 (30.9% described variability, Fig. 3B) captured the counter-directed anterograde and posterograde movements of neighboring legs, either ipsilaterally or contralaterally, and probably comprised the major component of forward locomotion. This PC suggestively groups the six legs into two groups (R1, R3, and L2, as well as L1, L3, and R2, respectively). The positions of legs in each of these two groups covary in the same direction and with the same magnitude, while the two groups have movement directions that are exactly in opposite direction. We hypothesize that this covariation corresponds to an idealized tripod movement. However, the first PC described less than a third of the dynamics occurring in the leg tips. PCs 2 (14.7%) and 5 (9.5%) show shifts that are mirror symmetric along the longitudinal body axis for all three leg pairs. The directions of these covariations indicate that PCs 2 and 5 might describe interindividual differences in overall postural width (Fig. 2 Ai and Aiii, for instance). PC 3 (12%) showed anterograde and posterograde directions of covariation for the front and hind legs (Fig. 3D). Here, however, the sign of covariation is negative on both body sides as compared to PC 1, and the middle legs do not display a comparable covariation in the direction of movement. Interestingly, in the context of interleg coordination, PC 3 would allow for a dissociation from the hypothesized strict tripod coordination as defined by PC 1 (but see next section for further elaboration). PC 4 (11%) only exhibits asymmetric covariations, mainly shifting the overall positions of legs left and right; this is most pronounced in the front legs. While particularly this left and right covariation of the front legs in PC 4 (and to some extent of the other legs in PCs 2 and 5) intuitively gives the impression of curve walking, we found it to be an actual asymmetry between individuals which mainly resulted from differences in mean posture between flies and not from curve walking (compare front leg positions of fly 1 and 3 in

Fig. 2 Ai and Aiii). Supplementary Fig. S2 shows quantitatively that flies that scored high or low on PCs 2, 4, or 5 did not produce reliable curve walking in either direction. Furthermore, we could not detect a strong relationship between the scores of PCs 2, 4, and 5 in a particular SC and the curvatures that were associated with that SC (Suppl. Fig. S3A to C). The same holds for the relationship between curve walking and PCs 1 and 3, respectively (Suppl. Fig. S3D and E). Taken together, these interpretations of the loadings of PCs 1 to 5 tentatively suggest that these might describe aspects of interleg coordination (PCs 1 and 3), and posture (PCs 2, 4, and 5). In the following analyses, we present further evidence for this hypothesis.

2.4.3 TCS reveals the relevance of PCs 1 and 3 for interleg coordination

Using TCS as a measure for how tripod-like a stepping sequence is, we tested for correlations between inter-leg coordination patterns and scores of individual PCs. Insects use a speed-dependent continuum of interleg coordination patterns (Szczecinski et al., 2018; Wosnitza et al., 2012). High walking speeds are associated with the so-called tripod coordination, slower walking speeds effect coordination that deviates from tripod coordination (Szczecinski et al., 2018). Canonical tripod coordination (rarely observed in its ideal form) thereby corresponds to anti-phasic activity of two tripod leg groups (sets of ipsilateral front and hind legs and the contralateral middle leg). In this context, TCS is a single value that describes the similarity between a particular coordination pattern and ideal tripod coordination (see Wahl et al., 2015 and Wosnitza et al., 2012). We used TCS to test whether covariations described by PCs were indeed correlated with certain coordination patterns. We first calculated TCS for all instances in which complete leg cycles of a particular tripod group were available. Then, where possible, we averaged the TCS values of two subsequent tripod cycles, effectively extending the TCS definition from single tripod groups to all six legs. The values calculated this way contained steps from the complete data set of 88 flies (all walking speeds) and yielded 30,216 instances of two complete and subsequent tripod cycles (Fig. 4A). The distribution of TCS values generally reflected the distribution of walking speeds; TCS values were predominantly found

in the range between 0.5 and 0.75 and are associated with walking speeds around 6 BL s^{-1} (see Fig. S1A and B).

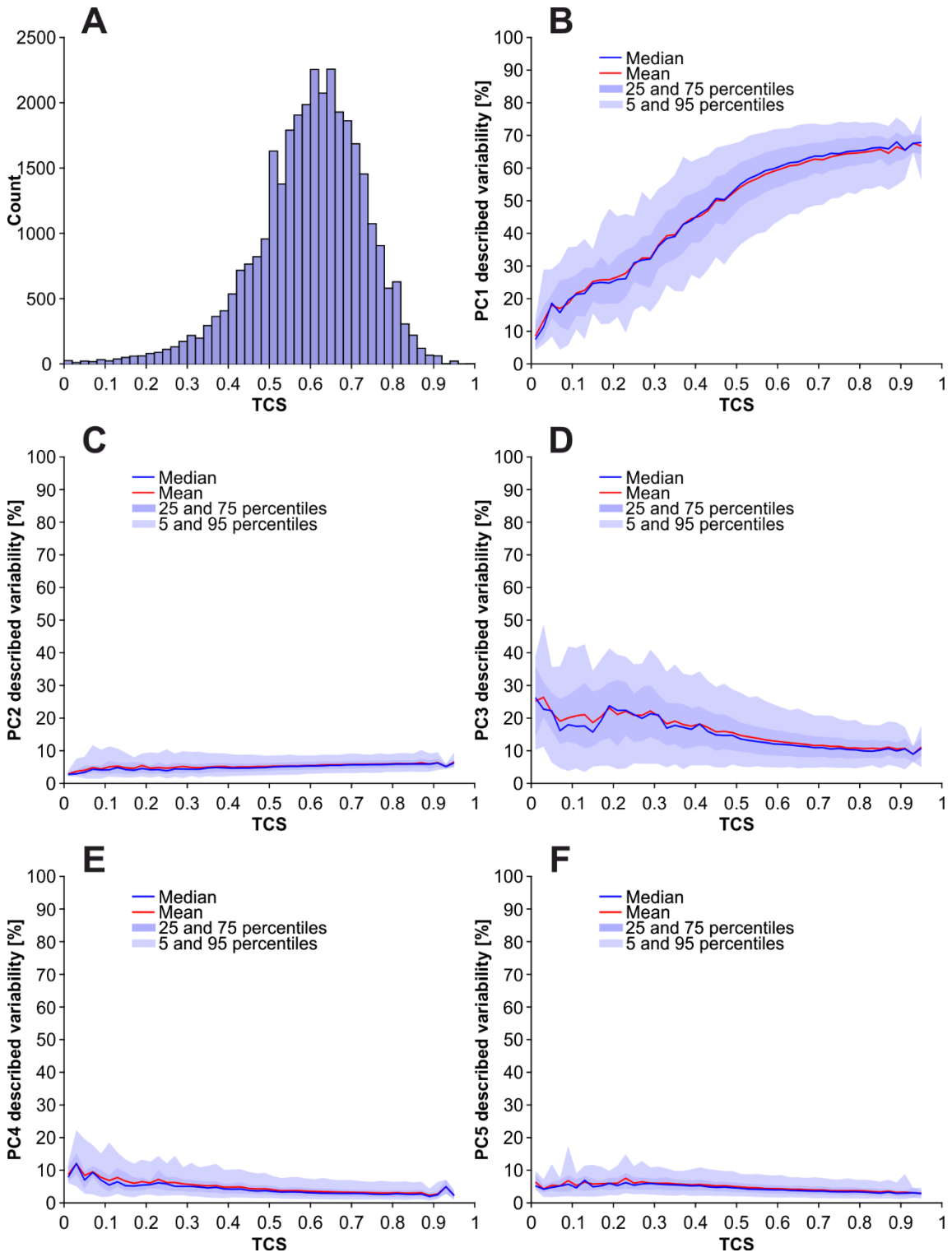


Figure 4: TCS reveals the relevance of PCs 1 (B) and 3 (D) for interleg coordination. (A) Histogram of TCS values for all steps of all flies in the analysis pool (n = 30,216). Note, that the majority of steps have TCS values in the range of 0.5 to 0.75; this is largely due to the fact

that these TCS values are associated with walking speeds of around 6 BL s^{-1}. A majority of flies tend to produce walking behavior at this speed (Suppl. Fig. 1A and B). (B) to (F): TCS values plotted against described variability of PCs 1 to 5. TCS was calculated for consecutive cycles of both tripod groups. Scores for PCs 1 to 5 were determined by transferring the tarsal tip positions of respective steps into the PC space and calculating the relative described variability for each PC and each cycle. The higher the TCS the more tripod-like a particular cycle is.

Leg tip positions within the respective interval were mapped into PC space and relative fractions of variability described by each PC for this subset of data were calculated. The fraction of variability described by PC 1 was found to be strongly positively correlated with TCS (Fig. 4B). Conversely, the described variance of PCs 2, 4, and 5 did not show a clear correlation with the respective TCS values (Fig. 4C, E, and F). PC 3 was negatively correlated with the TCS values, but this was less pronounced as compared to PC 1 (Fig. 4D). Taken together, these results suggest that PCs 1 and 3 describe a substantial part of the dynamic and coordination-related aspects of walking behavior. In contrast, PCs 2, 4, and 5 seem to describe features that are independent of these dynamic aspects; we explore their relevance in greater detail in the following sections.

2.4.4 PCs 2, 4, and 5 describe interindividual kinematic differences

Individual flies can be compared in subspaces spanned by the first five PCs to check whether interindividual differences of fly walking behavior are described by these PCs. To do this, and as a first step, we plotted data from five exemplary flies in 2-dimensional PC subspaces (Fig. 5A - D), either spanned by pairs of PCs 2, 4, and 5 (Fig. 5A - C) or PCs 1 and 3 (Fig. 5D).

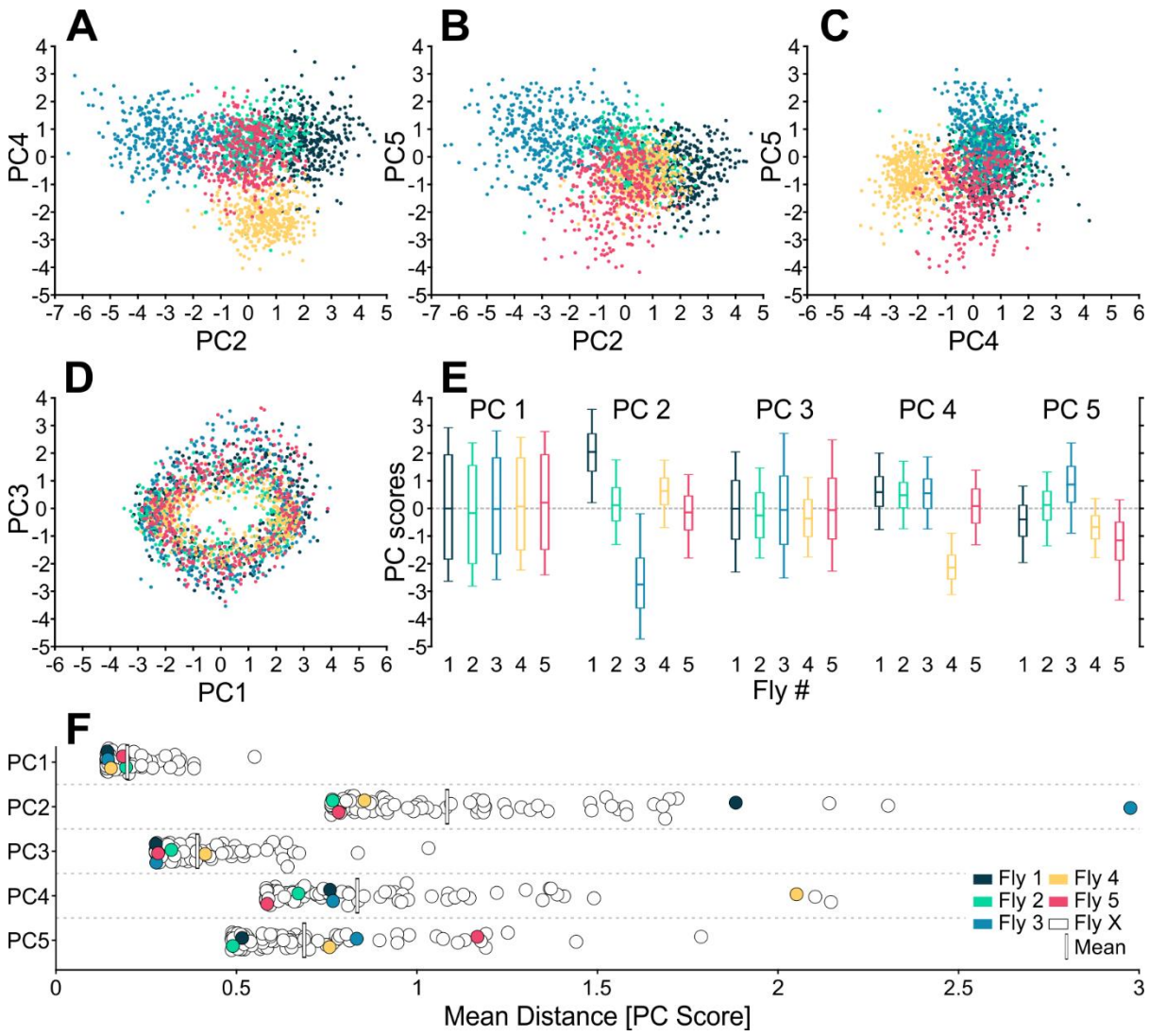


Figure 5: PCs 2, 4, and 5 describe interindividual differences of leg kinematics. Subspaces spanned by (A) PCs 2 and 4, (B) PCs 2 and 5, (C) PCs 4 and 5, (D) PCs 1 and 3. Individual colors correspond to individual flies. Each dot represents a fly's posture at a given time, selected randomly from the data which contributed to the analysis (n = 400 per fly). Note the difference between the distributions in A to C as compared to D. Also note the systematic elliptic shape of the distribution (see Movies M2 and M3). (E) Boxplots indicating the distributions of scores of the five flies in A to D (same color code) for the first five PCs. Flies 1, 2, and 3 refer to the data of the exemplary flies presented in Fig. 2. The data for flies 4 and 5 have been added to further populate and exemplify the subspaces in A to D. See Supplementary Figure S6 panels ii for summarized data of all flies (n = 88) and Movie M1 for an animation of a plot of all data, analogous to panels A to D in this figure. (F) Mean distance of each fly to all other flies along the dimensions of the first five PCs. Flies 1 to 5 are color coded in the same way as for panels A to E. Mean values for each PC are depicted as vertical lines. Note that flies cluster more in PCs 1 and 3, indicating higher similarity in these PCs (see also Fig. S9D).

The subspaces spanned by PCs 2, 4, and 5 show relatively clear separations between flies, while in PCs 1 and 3 (Fig. 5D) the data largely overlap (for the complete data set of 88 flies see Suppl. Movie M1). Interestingly, data plotted in Figure 5D forms an elliptic ring around the origin, indicating that they describe fundamentally different aspects of walking. We found that over the course of a SC the postural representation in this 2D subspace spanned by PCs 1 and 3 cycles the origin once (Movies M2 and M3). The dimensions of this ring were largely invariant for individuals, but depended on interleg coordination. Figure 5E confirms that all five flies had individual signatures regarding their scores for PCs 2, 4, and 5, while scores for PC 1 and 3 were comparatively similar (median values for PC 1 and 3 were approx. zero for these flies). Extending this analysis to the complete data set (Movie M1), we find that the average distances between flies in the dimensions defined by the five PCs is larger in PCs 2, 4, and 5 than in PCs 1 and 3 (Fig. 5F). These results demonstrate that PCs 2, 4, and 5 describe some of the inter-individual differences, while the absence of separation in the subspace spanned by PCs 1 and 3 shows that these describes features that are largely invariant between individuals (such as interleg coordination or the basic movement of legs).

2.4.5 Influence of walking speed and recording time

To investigate how walking speed affects aspects of interleg coordination and idiosyncrasies of walking we analyzed additional subsets of data from different speed ranges in the original PC space. We selected the range between 2 and 4 BL s^{-1} and 8 and 10 BL s^{-1}, to extend the analysis to slower and faster walking speeds, respectively. Because fewer steps were recorded for these speeds (see also Fig. S1A and B), fewer flies were included in these analyses; several of these flies had not been part of the original PCA. Nevertheless, for the lower speed range we found 23 flies and for higher range we found 43 flies with 30 or more steps. Supplementary Figure S5 shows how described variability changed for the lower (Fig. 6A) and higher (Fig. 6C) ranges of walking speeds compared to the original speed range (Fig. 6B, also compare Fig. 3A). The fraction of described variability of PC 1 strongly increases by approx. 21% from the low to the high speed range, while PC 3 increases by approx. 3.4%. In

contrast, fractions of described variability for all other PCs decreased, resulting in the accumulation of described variability in PC 1 and PC 3. This accumulation of PCs 1 and 3 and the reduction in PCs 2, 4, and 5 suggests that the overall variability of the flies' walking behavior decreased with increasing walking speeds.

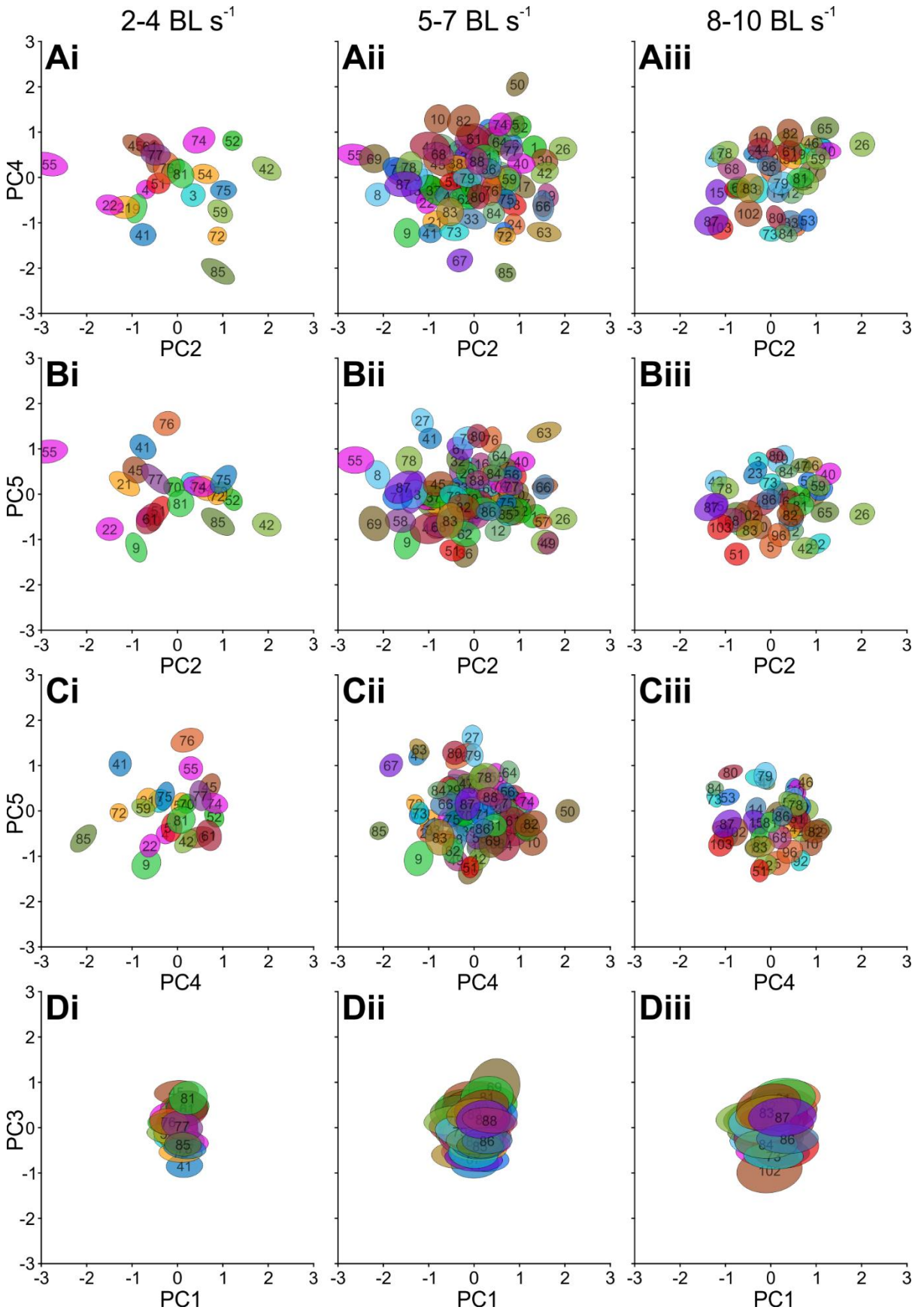


Figure 6: Data distribution of flies for three different speed ranges in the subspaces defined by PCs 1 to 5. Instead of individual data points, like in Fig. 5A-D, all data points are depicted

here in the form of ellipses that describe the bivariate distribution of the underlying data set for each fly (0.3 times the standard deviation). First column: 2 to 4 BL s^{-1}, n = 23 flies. Second column: 5 to 7 BL s^{-1}, n = 88 flies. Third column: 8 to 10 BL s^{-1}, n = 43 flies. (A) Subspace of PCs 2 and 4, corresponding to Fig. 5A. (B) Subspace of PCs 2 and 5, corresponding Fig. 5B. (C) Subspace of PCs 4 and 5, corresponding to Fig. 5C. (D) Subspace of PCs 1 and 3, corresponding Fig. 5D. Note, that the spread of the data in D (PCs 1 and 3) is much smaller than in A to C and that most flies occupy the same position in this subspace (also compare Fig. 5D and S5). The changes in mean positions of flies for different walking speeds are depicted in Suppl. Figure S5.

If idiosyncrasies are constant and can be described by PCs 2, 4, and 5, as we hypothesize here, flies should vary strongly with respect to each other in this space but should keep their individual positions when walking speed changes. To investigate this, we characterized the average positions and distributions of all individual flies in PC space separately for all three speed ranges and how their positions changed between speeds (Fig. 6). Analogously to the exemplary flies in Figure 5, the complete population of flies varies strongly in the space spanned by PCs 2, 4, and 5; this is true for the original speed range (Fig. 6Aii to Cii), as well as for the lower (Fig. 6Aii to Cii) and higher range (Fig. 6 Aiii to Ciii), supporting the notion that idiosyncratic aspects of walking are mainly described by these PCs, even when considering different speeds. In combination, Figure 6A to C show that most flies can be distinguished from others in their position with regard to at least one of these three PCs. In contrast, the data for individual flies overlap strongly in the space spanned by PCs 1 and 3 (Fig. 6D); at the same time, each individual fly's data distribution (ellipses in Fig. 6) increases in the direction of PC 1 with walking speed, suggesting an increase in importance of this PC hypothesized to be reflective of tripod coordination. With increasing walking speed, the average distance from any fly to all others in the subspace of PCs 2, 4, and 5 decreases from 1.87 to 1.75 to 1.54 (also compare Fig. 5F). In the subspace of PCs 1 and 3 we measured average distances of 0.51, 0.47, and 0.49, respectively. Not only are the average distances between flies generally much larger for PCs 2, 4, and 5 as compared to PCs 1 and 3, they are also more strongly affected by increasing walking speeds. Most importantly here, however, is the relative constancy of positions in the space of PCs 2, 4, and 5 (see Fig. 6 for absolute positions and Suppl. Fig. S6

for shifts). This indicates that the kinematics described by these PCs largely remain constant for an individual, even when that individual changes its walking speed. Together with the clear identifiability of individuals in this space, this suggests that these PCs indeed describe individual aspects of walking that are independent of walking speed.

2.4.6 Optogenetic inhibition demonstrates the descriptive potential of PCs

An optogenetic inhibition experiment was performed to evaluate the potential of PCs 2, 4, and 5 to describe systematic differences in posture and leg positioning. We crossed the *iav*-Gal4 driver line with UAS-GtACR1. The resulting F1 generation expressed the inhibitory channelrhodopsin GtACR1 in the neurons of chordotonal organs, including the femoral chordotonal organ (fCO). Previous experiments have shown that inhibiting these sensory structures causes systematic changes in walking kinematics (Chockley et al., 2022). Here, we used these expected changes to test if they can be detected within the subspace spanned by PCs 2, 4, and 5. We focused on this particular PC subspace because the changes due to inhibition of the fCOs during walking were found to be largely postural (Chockley et al., 2022) and, as we have shown in the preceding analyses (Figs. 4, 5, and 6), this subspace seems to capture postural changes well.

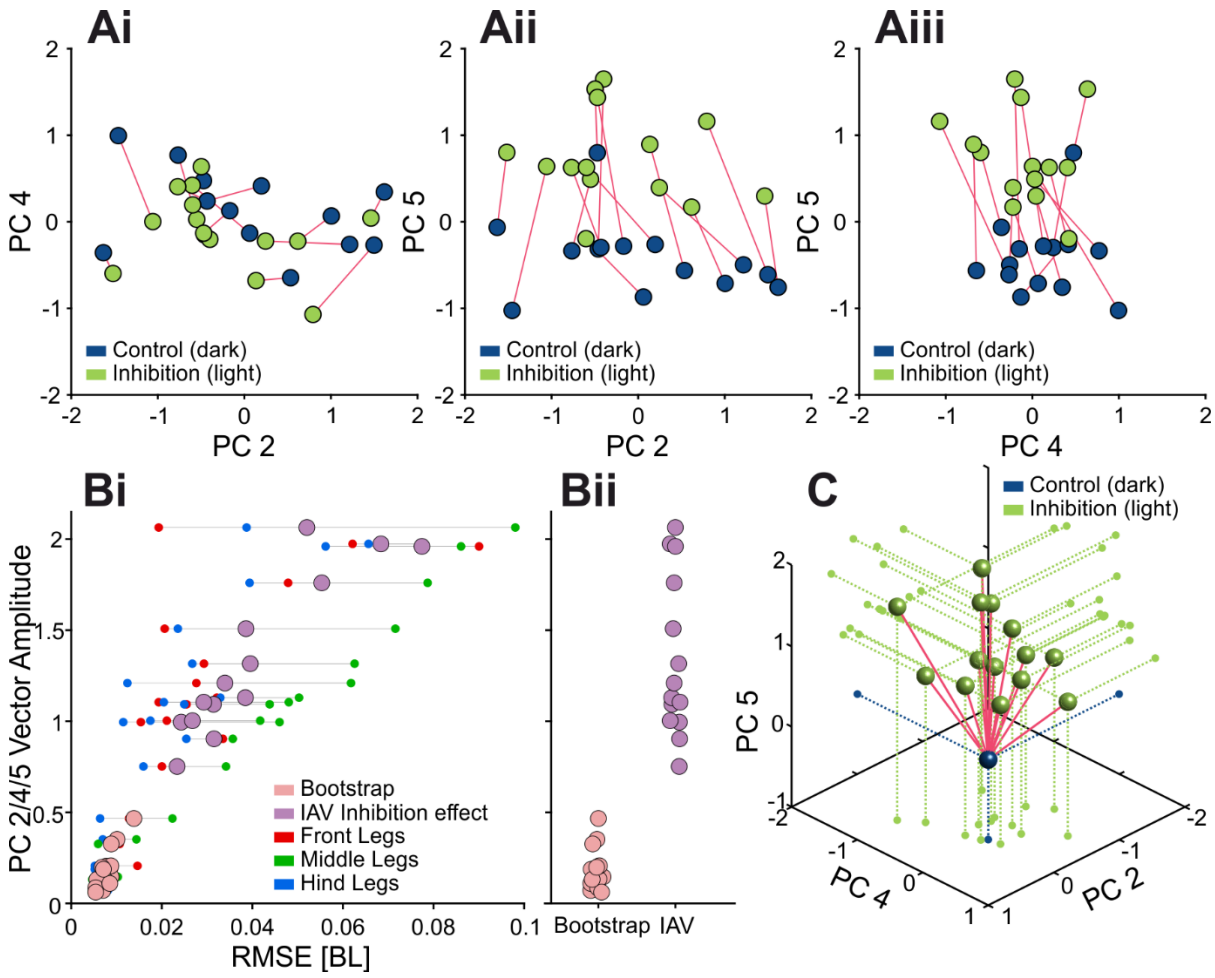


Figure 7: An optogenetic inhibition experiment demonstrates the descriptive potential of PCs. (A) 2-D representations of the inhibition effect in subspaces spanned by combinations of PCs 2, 4, and 5 (steps in dark: blue; steps in green light (fCO inhibited): green). Red lines connect mean positions of individual flies in the two conditions. (B) Vector amplitudes (i.e. effect size) of the inhibition experiments and the bootstrap analysis plotted against the RMSE of the leg tip trajectories. Teal dots depict results for the bootstrap analysis, orange dots for inhibition experiments. Red, green, and blue dots indicate the RMSE for front-, middle-, and hind legs. (C) 3-D representation of the inhibition effect (green) relative to the control condition (blue). For clarity, control conditions were set to the origin.

Due to the alternating darkness-light paradigm (see Methods), each individual fly produced trials in the dark (corresponding to wildtype condition) as well as during green-light illumination (corresponding to inhibition); each fly therefore served as its own control. For each fly with more than 30 steps in the control (dark) and the inhibition condition (light), the mean positions for both conditions were analyzed in the subspace spanned by PCs 2, 4, and 5 (Fig. 7A). The length of the vector describing the difference between dark and light condition was compared

to a bootstrap analysis. For this, two sets of 30 steps each were randomly drawn from single flies of the original data set and plotted in the same way as the data from the inhibition experiments (Fig. 7B). The results for the inhibition experiments show no strong preference for shift directions for PCs 2 and 4, but a clear and consistent shift towards more positive values for PC 5 (Fig. 7Ai - Aiii). In contrast, the bootstrap analysis expectedly resulted in more random shift directions whose magnitude was also smaller. The mean effect sizes measured for PCs 2, 4, and 5 (i.e. vector lengths) were plotted against the observed differences in the mean leg tip trajectories between control and inhibition, expressed as root mean squared error (RMSE), showing a positive correlation (Fig. 7Bi); the stronger the difference between control and inhibition on the level of leg tip kinematics, the larger the shift in PC space was. The difference in RMSE is mainly driven by the middle legs (green dots), matching the effect described by (Chockley et al., 2022). Figure 7Bii depicts the vector lengths for the inhibition experiments and the bootstrap analysis, showing that effect size was much larger on average for the inhibition experiments. PC 2 shows shifts to more negative values for 10 individuals, but also four in the opposite direction, although with much smaller amplitudes. PC 4 shows even smaller and less consistent effects, likely because the inhibition of fCOs affected both body sides equally while PC 4 describes only asymmetric covariations (Fig. 3E). However, the clear and consistent shifts from dark to light condition in PC 5 demonstrate that PCs in general can be used to compare and quantify effects and their magnitude in a reduced number of dimensions.

2.4.7 A symmetry axis demonstrates how individual postures are encoded in PCs 2, 4, and 5

To further explore which interindividual differences might be encoded by PCs 2, 4, and 5, we systematically searched for an axis in this PC subspace that reflected symmetric postural changes with regard to the longitudinal axis of the body. We found an axis which described the mean distance of the leg tips to the fly body (Fig. 8Aii).

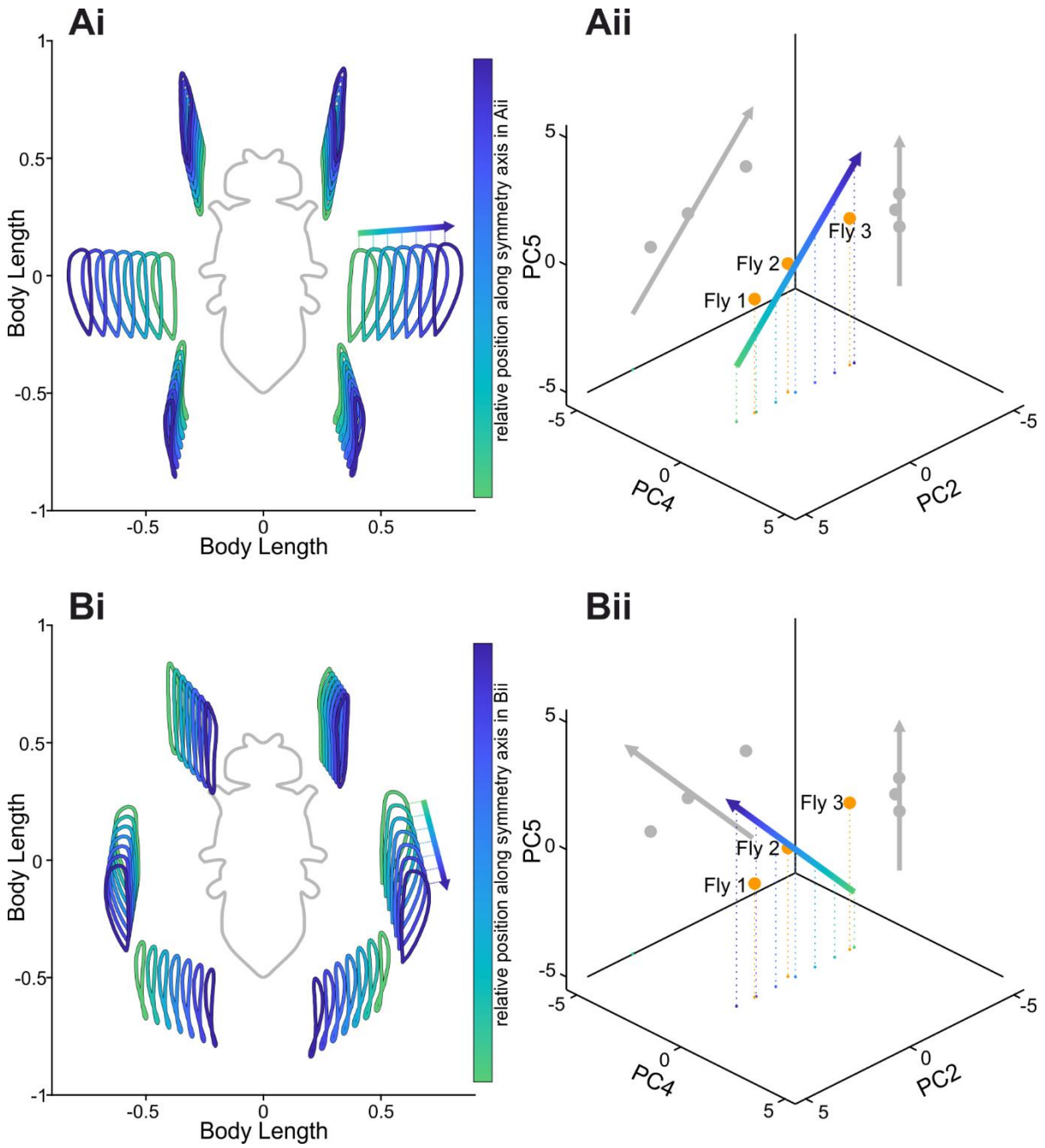


Figure 8: A symmetry axis demonstrates how individual postures are encoded in PCs 2, 4, and 5. (A) Leg tip trajectories for different points on the (B) symmetry axis in the subspace of PCs 2, 4, and 5 describing systematic differences in postural width shown in (A). Color code in A corresponds to color shading along the axis in B and is additionally shown in the form of a gradient bar on the right side of panel A. Orange dots in B depict the mean positions of flies 1, 2, and 3 introduced in Fig. 2 and also referenced in Fig. 5. Gray: projections of the three-dimensional sequence of explored combinations into two-dimensional subspaces spanned by PCs 2 and 5 as well as PCs 4 and 5, respectively.

The contribution of PC 4 was very weak, which is not surprising since we searched for highest symmetry and PC 4 describes asymmetric covariations. Interestingly, flies 1 to 3, (referred to in Figs. 2 and 5), are close to the symmetry axis and the qualitative differences in their postures, which mainly involved how sprawled these flies walked, are represented well in the artificial postures shown in Figure 8Ai. PC 4, on the other hand, seems to describe less symmetric differences between individuals as presented for fly 4 in Figure 5. To additionally demonstrate how PCs 2, 4, and 5 might describe differences in postures, we constructed an axis that was perpendicular to that shown in Figure 8Aii. The resulting axis was still mainly co-planar with the subspace of PC 2 and 5 (Fig. 8Bii). The respective changes of fly posture displayed in Figure 8Bi are almost perpendicular to those in Ai, mainly shifting the average leg positions anterograde and posterograde. Together, these two PCs compactly describe a large range of different, but symmetrically arranged fly postures. Interestingly, PC 4 did not show up in our unbiased approach to find these symmetry axes. This is unsurprising, when we consider that this PC largely describes left and right shifts in leg positions (Fig. 3E). Together with PCs 2 and 5, it is then possible to more precisely describe the left-right asymmetry and, due to the linear nature of PCA, PC 4 would simply shift the posture accordingly.

2.5 Discussion

Here, we used a data set of a large number of flies and many instances of walking sequences to investigate the naturally occurring intra- and interindividual variability of walking on the level of leg tip kinematics. Initial qualitative observations readily replicated anecdotal impressions from previous studies that flies walk in an idiosyncratic manner (Fig. 2). A quantitative analysis using PCA revealed that much of the observed kinematic variability can be captured by a lower-dimensional representation of five PCs (Fig. 3A). Two principal components (PCs 1 and 3) captured general variability aspects across individuals related to invariant movement patterns, such as the general positional changes during a SC and interleg coordination (Fig. 3B and D, Fig. 4B and D); at the same time, three other PCs (2, 4, and 5) described individual-specific aspects of walking (Fig. 3C, E, and F, Figs. 5 and 8). PCA was able to separate these two different sources of variability into sub-domains (Fig. 5, Fig. 6 panels ii) and the contributions of the individual-specific PCs can be regarded as a fingerprint of a fly's idiosyncrasies during walking (Fig. 5E, Fig 6Aii to Cii) which are distinct from more general features shared by all flies. Investigation of different ranges of walking speeds, as well as the comparison of trials recorded at different times during experiments, showed that idiosyncratic aspects of walking are largely constant for an individual, while general invariant dynamics of walking, as captured by PCs 1 and 3, increased in significance with an increase in walking speed (Fig. 6 and Suppl. Figs. S4 to S6). Exploration of the individual-specific PCs 2, 4, and 5 revealed that general high-level features, like kinematic changes after experimental interventions (Fig. 7) or overall posture of tarsal tips (Fig. 8) can be compactly detected, described, and quantified.

Postural variability in the present data occurs on the intra-step level, the step-to-step level, and the interindividual level. The intra-step level thereby refers to positional variations of the tarsal tips over time, essentially the periodic back and forth movement of legs during protraction and retraction. The specifics of these movements are mainly captured by PCs 1 and 3, with PC 1 reflecting the canonical tripod coordination (Figure 3B) and PC 3 reflecting deviations from this tripod coordination. Because this dynamical component of variability is large as compared to more subtle and static postural aspects, it is not surprising that PC 1 is the most prominent PC

and generally captures a large part of the variability across individuals. Together with PC 3 it therefore is suited to compactly represent the overall coordination an animal uses. Previously, there have been only few attempts to compactly characterize coordination of the six legs of walking insects; TCS is one measure (Wahl et al., 2015; Wosnitza et al., 2012), but it has limitations, especially if the walking pattern deviates strongly from tripod coordination. A characterization based on these two PCs, whose contributions also seem to be negatively correlated with each other (see Fig. 4B and D), could serve as an alternative approach to efficiently describe coordination in future studies. An aspect that was not yet analyzed in detail here is the temporal variation of the contribution of these two PCs, their score time courses. Since PC 1 and 3 captured the actual movement of the legs, the scores of these two PCs will generally be modulated periodically over the time course of the stepping movements of individual legs (Movies 2 and 3). Further examination in this regard could reveal more subtle relationships between how PC 1 and 3 are modulated in the context of interleg coordination and how individual flies combine these two patterns to establish tripod coordination (PC 1) and deviations from it (PC 3).

In contrast, PCs 2, 4, and 5 capture more static and interindividual postural differences (Figs. 4, 5, and 8). The postures of individual flies occupy different parts of this PC subspace. Consequently, this subspace is helpful to describe these individual differences in the first place, but this can be readily expanded. We explored several exemplary expansions here (discussion for walking speed-related results further below). The first was an experimental intervention that introduced known changes in leg kinematics which, in turn, were picked up clearly in the space spanned by PCs 2, 4, and 5. The second expansion was a top-down search in this space that describes overall changes in postural width. This approach can be useful in other novel interventions, in which putative effects are to be detected but for which a more unbiased analysis is desirable or in which several effects combine more subtly. Previous studies focused on kinematic parameters like stance amplitude, durations of swing or stance phases, or AEPs and PEPs (Mendes et al., 2013, Wosnitza et al., 2012, Strauss & Heisenberg, 1990, Ramdya et al., 2017). While these singular measures are informative, a more unbiased way of looking

at putative effects of interventions might reveal other effects that are less intuitive, more complicated, or interdependent.

PCA has the general limitation of only capturing linear correlations between analyzed variables. It is conceivable that more complicated approaches for dimensionality reduction will yield more compact or clearer description and separation of the different levels of behavioral variability. A previous study, for instance, used Uniform Manifold Approximation and Projection (UMAP) to explore high-dimensional kinematics of walking flies and found similar reductions of these data (DeAngelis et al., 2019). However, the analysis we use here already captures many interesting and helpful aspects of variability; the strength of PCA in the present context is its intuitive interpretability with regard to what individual PCs mean for kinematics.

Considering the putative roles of the five PCs, with 1 and 3 responsible for aspects related to interleg coordination and 2, 4, and 5 responsible for postural aspects, we can speculate that a particular combination of all of these in an individual has implications for how walking is controlled, especially with regard to static stability (Szczecinski et al., 2018). Stability is dependent on the duration of ground contact of the legs as well as their positioning in relation to the center of mass. This might directly tie back to the individual PCs. Based on the findings presented here, we assume that the ratio of PCs 1 and 3 mainly determines interleg coordination (Movies M2 and M3) and that the remaining PCs determine overall posture (including AEPs and PEPs). Thus, from a motor control perspective we speculate that PCs 1 and 3 might be dynamically mixed together, based on the intended interleg coordination and walking speed, but also on aspects of posture that have become fixed in ontogeny, like its symmetry. High walking speeds, require exact alternation of the swing and stance phases of neighboring legs, in the sense of stability. More asymmetric overall postures (e.g. in flies that score high on PC 4, for instance) might therefore necessitate an earlier switch to more tripod-like and therefore more stable coordination patterns (as reflected in strong contribution of PC 1) to compensate for deviations from perfect symmetry.

Importantly, the first analysis in the present study focused on straight walking in a narrow speed range and on male flies of the same age, reared in identical conditions, and from the same highly interbred wildtype strain. These restrictions were intentional, as we wanted to first establish the general approach and its usefulness for a more controlled subset of all possible behavioral data. Therefore, we wanted to exclude additional sources of potential variability, based on parameters like walking speed, age, or sex. Even in this relatively controlled data set we nevertheless found a diversity of idiosyncratic ways of walking. But it is true that this first analysis necessarily only reflected kinematic details which were contained in the selected data and the conclusions, particularly with respect to the constancy of idiosyncrasies, could be limited.

We therefore extended the analysis to different walking speeds. Walking speed has strong effects on many kinematic parameters and interleg coordination in walking insects (DeAngelis et al., 2019; Strauß and Heisenberg, 1990; Szczecinski et al., 2018; Wosnitza et al., 2012). This analysis showed that idiosyncrasies remain largely constant when individuals change walking speed, but that aspects of interleg coordination change to some extent. Especially PC 1 becomes more important, a finding that is in line with previous established results that coordination becomes more tripod-like the faster a fly walks (Szczecinski et al., 2018; Wosnitza et al., 2012). Another interesting observation (until now anecdotal), relates to how strictly flies walk in dependence of walking speed. These observations suggest that AEP and PEP positioning or the phase relationships between legs, for instance, become less variable at higher walking speeds. Our results now support this observation in a quantitative manner: flies tend to become relatively more similar to each other with increasing walking speed, i.e. the described variability of PCs 2, 4, and 5 (interindividual variability) decreases and PCs 1 and 3 (aspects of interleg coordination) increase in importance. While the exact reason for this is unknown at this point, we hypothesize that it could be related to the stability requirements of fast walking; a previous study suggests, that the region in phase space associated with stable walking at high speeds is smaller than that for low speeds (Szczecinski et al., 2018) and that fast walking necessitates tripod coordination. Flies should therefore walk (1) in a less variable

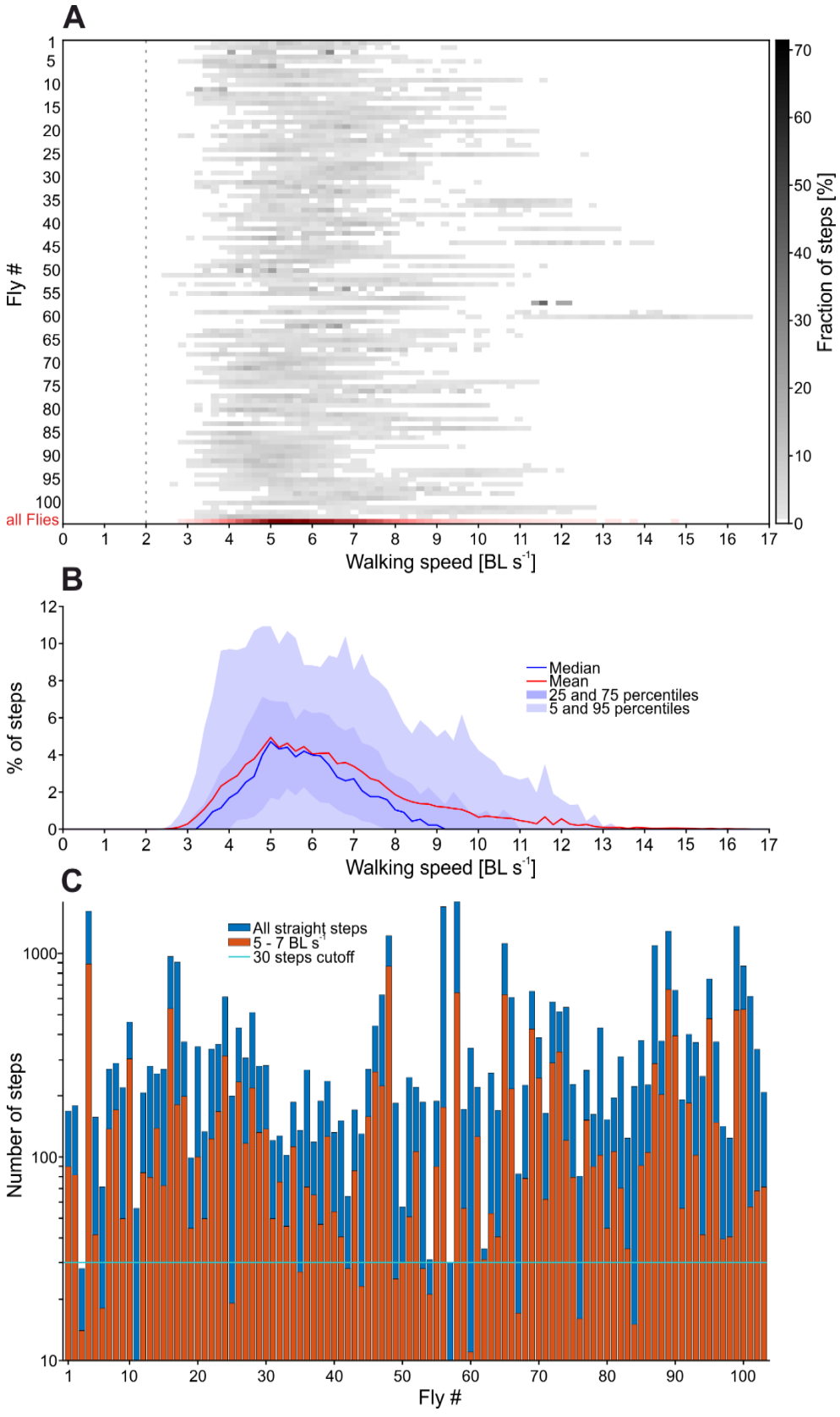
way at these speeds and (2) should also become more similar to each other. Our results support this notion and future studies will investigate this aspect more closely.

Other natural extensions of this analysis might focus on the influence of curve walking or age, among others. Expanding the number of these factors a controlled way should successively elucidate the full spectrum of natural variability in fly walking behavior, thereby enabling more precise investigation of the fundamental principles underlying the motor control of insect walking. Even stronger changes in the specific makeup of PCs and their contributions are expected when not only straight but also curve walking enters the data set. Curve walking entails pronounced changes in the kinematics of all legs as indicated by previous studies, mainly on stick insects (Dürr and Ebeling, 2005; Gruhn et al., 2008). Legs contribute in a specific kinematic manner to curve walking. These changes will be reflected in the detected PCs, adding another layer of variability that reflects invariant aspects of curve walking, on the one hand, and potential interindividual differences in the way single flies implement this change.

Here, we provide the first extensive study of variability on the kinematic level in *Drosophila* walking behavior in genetically very similar and highly interbred individuals. The combination of dimensionality reduction using PCA, the interleg coordination measure TCS, and inhibition experiments shed some light on the natural variability present in walking behavior and suggests an enhanced way of characterizing data of freely-walking flies. We thereby hope to build the fundament for further systematic investigation of the principles underlying the natural variability of walking.

2.6 Supplementary material chapter 2

2.6.1 Speed distribution across all recorded flies



Supplementary Figure 1: Distribution of walking speeds across all recorded flies and step cycles (SCs). (A) Each row shows one fly. Bin size for walking speed was set to 0.2 BL s^{-1}. Walking bouts below 2 BL s^{-1} were not recorded. Last row (red) shows the total walking speed distribution for all flies and 36,942 recorded SCs. (B) Distribution of walking speeds for all flies, same data as in A (last row). (C) Logarithmic histogram of the number of SCs for all recorded flies. Blue = total number of straight SCs; orange = number of SCs between 5 and 7 BL s^{-1}; teal = cutoff line for flies which were taken into the analysis pool at 30 SCs.

Figure S1 displays the distribution of walking speeds for all flies and all SCs pooled together (see line labeled “all flies”). Some flies tend to walk faster than others; most of them, however, preferred walking speeds between 5 and 7 BL s^{-1}, which is why this speed range was selected for the study. Line total indicates that walking speed across steps tends to be log normal distributed. No steps associated with walking speeds close to the cutoff point at 2 BL/s were recorded. Of the 103 recorded flies, 88 produced 30 or more steps cycles in the target speed range of 5 to 7 BL s^{-1}.

2.6.2 Automated detection and annotation of flies

While detection and annotation of the flies was fully automated, each video was manually checked for annotation errors and corrected if necessary. The detection and annotation process was structured as follows. For each image frame in a video, we created a binary image with a simple threshold operation, i.e. bright pixels above the threshold became white, dark pixels below the threshold became black. To remove small regions of white pixels we used a morphological opening operation (equivalent to an erosion operation, followed by a dilation, MATLAB function *imopen.m*). After this operation, only the region of white pixels corresponding to the fly was left. The position of the fly was then detected as the mean position of this remaining blob. This position was used to crop a square region containing the fly from the image. The size of this cropped region was adjusted to the fly’s length. Using DLC, we then detected the neck and abdomen in this cropped subregion and used this information to rotate the subregion such that the fly’s longitudinal axis was aligned in parallel to the y-axis of the image. In two subsequent runs of DLC we then detected the leg tips on the right body side, as well as those on the left side, respectively (see also Fig. 1B).

2.6.3 Discrimination of swing and stance phases

Transition points between swing and stance phases, i.e. lift-off and touchdown, were automatically determined by analyzing the movement of the markers (i.e. their absolute speed) in the arena-centric coordinate system of the original video. For every marker and every frame, the distance between the current marker position to the position in the subsequent frame was calculated. Due to noise the exact location of a marker slightly varies from frame to frame independently of actual leg movement. Therefore, a threshold had to be defined to discriminate frames with legs moving relatively to the ground from frames with legs having ground contact. An empirically determined threshold of 1.5 pixels per frame produced precise and reliable results and was used for the complete analysis in this study. Additional conditions were defined to further increase the robustness of the algorithm. For example, a swing phase had to last at least 3 frames (equal to 15 ms), otherwise it was not detected as such. Additionally, the correct automatic determination of lift-offs and touch-downs was reviewed manually during the process of screening for errors in the DLC annotations.

2.6.4 Training of DeepLabCut

Automated annotation of all six leg tips, the neck and the tip of the abdomen was carried out with DeepLabCut (DLC), a deep neural network approach (Mathis et al., 2018). DLC was trained on exemplary frames of freely walking fruit flies along with manually generated information about the locations of the eight body parts of interest (neck and abdominal tip, as well as the tips of all six legs). Respective training data was produced by manually annotating video frames in a custom-written MATLAB program.

Coordinates of the body parts of interest were stored together with the cropped video in a single MAT-file. For the initial training data set, 500 frames were collected from ten flies in various positions. Training data sets were digitally augmented by randomly applying rotations, translocations, and resizing. This prevented adaptation of the network to a certain orientation,

size, or position in the video frame. Rotations were not applied for the training of the networks which were used for leg tip detection, since the fly body axis was aligned vertically after the detection of neck and abdomen. Here, complexity of possible inputs was reduced to potentially increase accuracy and reliability of the automated annotation.

Even though the arena surface was cleaned regularly, small pieces of dust or debris accumulated on it to some extent. These pieces typically can have sizes and shapes that make them similar in appearance to the tips of the fly's legs and were sometimes detected as such. To make the trained networks more robust to these pieces, randomly generated, virtual specks of dust and debris were added to each training frame. Modifications of training frames were carried out in a custom-written MATLAB program. The same program was used to create training directories, which contained all files necessary for training of DLC. Training was run for 400 thousand to 1 million iterations, taking 8 to 20 hours.

2.6.5 Selection of straight segments

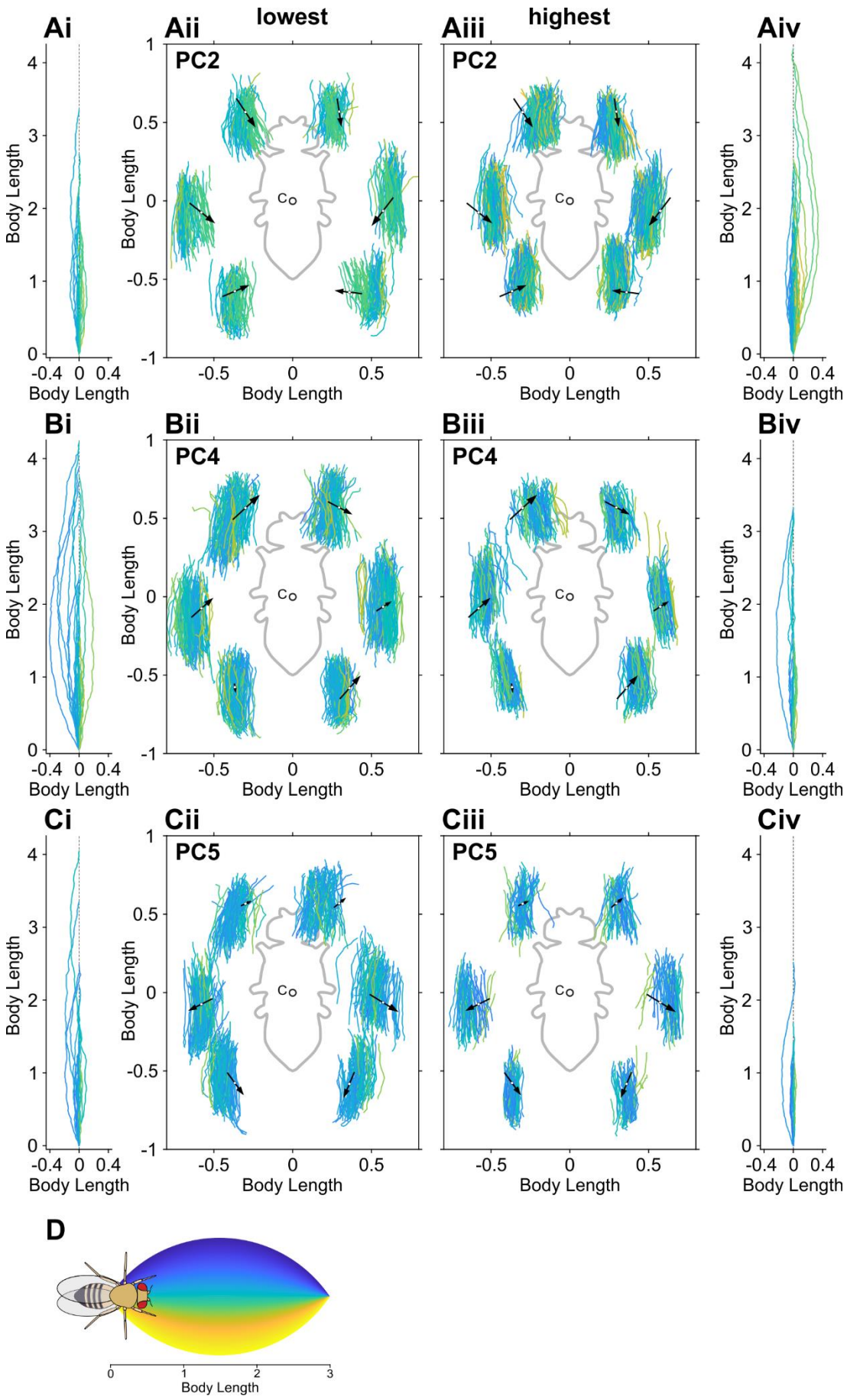
The recording software triggered data acquisition dependent on the length of the trajectory rather than its shape. Hence, the acquired video data needed to be scanned for segments of straight walking. For thousands of videos it is fairly labor-intensive to search for such segments manually, isolate them from the remaining video, and store them in another directory. Moreover, to what extent a trajectory is considered to be straight is subjective and the criteria may change unconsciously over time. Therefore, an algorithm judging each trajectory by objective criteria was considered to be the best option. Since all further analyzed trajectories should be of certain straightness while losing as few frames as possible, two approaches for detecting straight segments were combined to compensate for their individual disadvantages. For both approaches, the original trajectory was first divided into equidistant pieces of 0.5 mm length. One approach extracted the angles between neighboring pieces as an estimator for straightness, in a way that two or more pieces of trajectory with an angle of 5° or less to each other were considered to be straight. The other approach used the quintic spline function in MATLAB to fit a spline to equidistant points. Each point with its two direct neighbors can be

used to define a circle with a radius correlating to the curvature at this point of the trajectory. An infinite radius would mean that all three points are on a perfect straight line. The smaller the radius becomes, the higher the curvature is. Constellations of points with an inverse radius of 0.1 or less were assessed as straight.

2.6.6 Curve walking and PCs 2, 4, and 5

Flies with very low or very high mean values in the directions of PCs 2, 4, and 5 were analyzed with respect to the curvature of their walking trajectories and corresponding leg tip trajectories. Figure S2 shows exemplary walking trajectories of these flies with low and high means, together with the respective leg tip trajectories, which were color-coded according to the same code (Fig. S2D) as the walking trajectories. This color code differentiates between left and right turns, with blueish colors for right and yellowish colors for left turns. If remaining curve walking behavior is the cause of the differences between the extreme flies of any PC, we should see a clear tendency of yellowish leg-tip trajectories being on one side of the continuum and blueish ones on the other side. Instead, we see mostly equally distributed colors of leg-trajectories, while sometimes even fully yellow and blue ones overlap each other (e.g. Fig. S2Bii, right hind leg). PC 2 in Figure S2A features fly 1 (Aiii) and fly 3 (Aii) which are also shown in Figure 2 Bi and Biii, Figure 5, and Figure 8. Here we see how PC 2 describes the difference in posture width between both flies. Figure S3B shows low and high scoring examples along the dimension of PC 4, which gives an impression of potential curve walking behavior from the arrow depiction in Figure 3E, mainly because of the side shift in both front legs. Leg-tip trajectories of the extreme flies along the dimension of PC 4 shown in panels Bii and Biii in Figure S2 demonstrate that the individual posture preference is much more clearly correlated with PC 4 scores than the curvature of the walking trajectory. Hence, PC 4 describes first and foremost idiosyncrasies in walking behavior and is not directly correlated with remaining aspects of curve walking. Extreme cases of PC 5 in Figure S2C again show no correlation with trajectory curvature, but illustrate how PC 5 describes symmetric differences in posture which

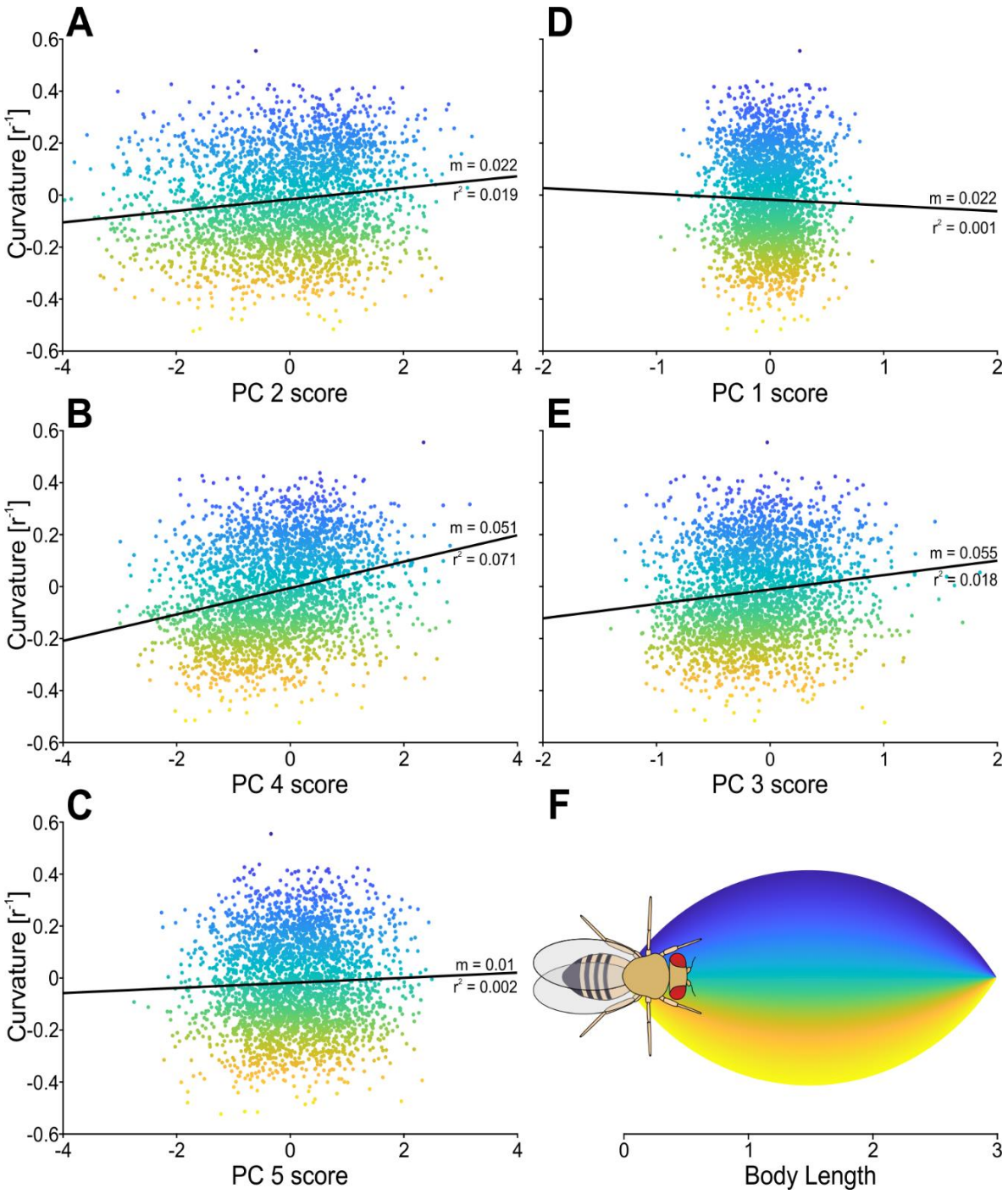
are mostly perpendicular to those described by PC 2. Panel D provides a legend for the color code of curvature values relative to the flies' body length.


Ai
Aii
lowest
Aiii
highest
Aiv
PC2
PC2
Bi
Bii
PC4
Biii
PC4
Biv
Ci
Cii
PC5
Ciii
PC5
Civ
Co
Body Length
D
0
1
2
3
Body Length

Supplementary Figure 2: Walking trajectories and leg-tip stance trajectories of exemplary flies with low (i and ii) and high (iii and iv) scores for PCs 2, 4, and 5. (A) PC 2. (B) PC 4. (C) PC 5. Leg tip trajectories as well as trajectories are color-coded according to the curvature of the trajectory of their associated trial. (D) Color code for curvature of walking trajectories. Yellow hues indicate trajectories that are curved to the right, blue hues indicate trajectories that are curved to the left. (A) PC 2 mainly scales the overall posture width, with the fly scoring high (Aii) having a more sprawled posture than the one that scores low (Aii). (B) PC 4 most strongly affects the front legs (Bii and Biii, for high and low scores in PC 4) and due to its asymmetric nature suggests a curve walking context most directly. However, the associated trajectories (Bi and Biv, respectively) do not show a clear curve walking phenotype. (C) PC 5 mainly affects middle and hind legs. The curve walking implications for a fly that scores high (Cii) or low (Ciii) are very weak when the respective trajectories are evaluated (Ci and Civ, respectively). (D) Color code for curvature of walking trajectories. Yellow hues indicate trajectories that are curved to the right (positive values), blue hues indicate trajectories that are curved to the left (negative).

2.6.7 Curvature vs PC scores

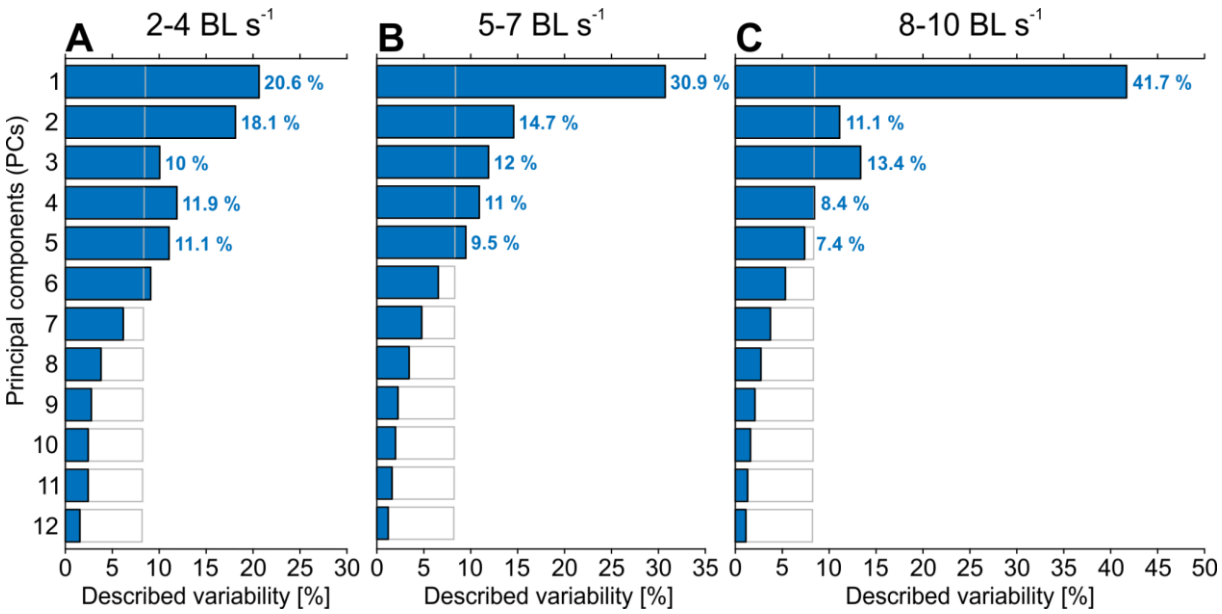
To more systematically test for correlations between slightly curved walking trajectories and certain PC scores we calculated the curvature of all recorded straight walking bouts in the data set and color coded it as in Figure S3D. We then calculated the fraction of variability described by all PCs for every individual walking bout and plotted the results in Figure S3. We also plotted regression lines, their slopes (m) and their r^2-values. Figure S3A shows a weak correlation for PC 2, as it explains only 1.9 % of the variance. The results for PC 4 in panel B show that it explains about 7.1 % of the curvature variance. This correlation is substantially weaker than the influence of individual preferences in average leg tip positions as demonstrated in Figure S2B. PC 5 in panel C does not show any clear correlation with only 0.2 % explained variance. The two PCs we identified as mostly related to aspects of interleg coordination (PCs 1 and 3, panels D and E) have similarly low coefficients of determination with regard to curvature and describe less than 1% and 1.8% of variance due to curvature, respectively.



Supplementary Figure 3: Curvature of individual walking trajectories plotted against their respective scores in all PCs. (A to C) PCs 2, 4, and 5, respectively. (D and E) PCs 1 and 3. Black: regression line with slope (m) and coefficient of determination (r^2). (F) Color code for curvature of walking trajectories. Yellow hues indicate trajectories that are curved to the right (positive values), blue hues indicate trajectories that are curved to the left (negative). Coefficients of determination are generally very low, indicating no strong correlation between the curvature of walking trajectories and the described variance of the respective PC.

2.6.8 Effect of walking speed on PC contributions

Steps between 2 and 4 as well as between 8 and 10 BL s^{-1} were transformed into the PC space resulting from the initial analysis that was based on a range of 5 to 7 BL s^{-1} (see Fig. 6 in the main text). The results of this analysis results are shown in Figs. S4A and S4C, respectively. Fractions of described variability were calculated and plotted in the same way as in Figure 3A (for comparison, Fig. 3A was replicated here as Fig. S4B). Figure S4A shows the results for 2 to 4 BL s^{-1}, with a more than 10% smaller fraction of described variability for PC 1 and 1% smaller for PC3, while all other PCs described more variance compared to 5 to 7 BL s^{-1}. Panel C shows a continuation of the trend from A to B with even higher fractions of variability being described by PCs 1 and 3 while the described variability decreases for all other PCs. This result is a hint for an overall decrease of inter-individual variability with increasing walking speed, since relatively more variability is described by fewer PCs, particularly by PCs 1 and 3, which we identified as the ones that describe more general and invariant aspects of walking that all flies have.

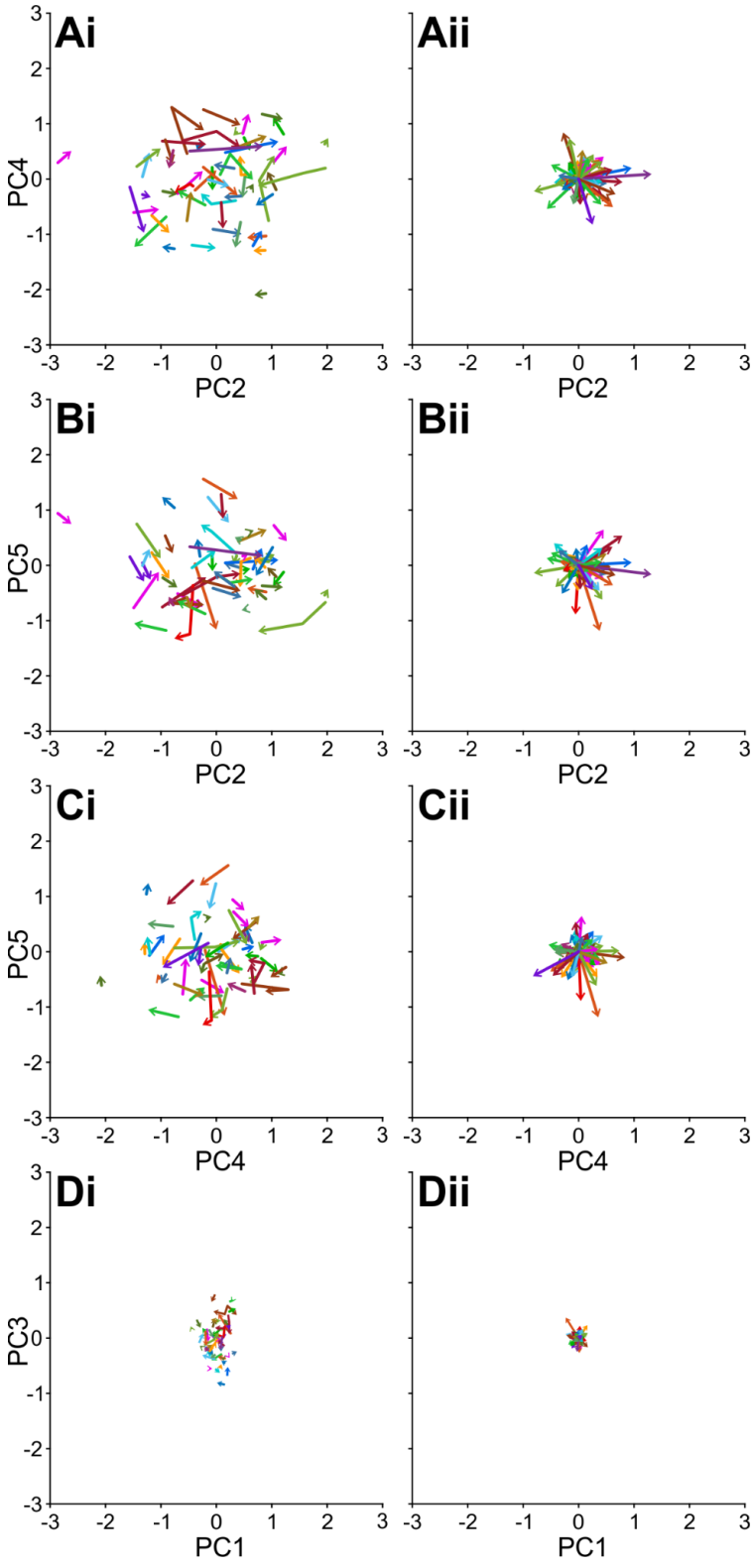


Supplementary Figure 4: Fractions of variability in leg tip kinematics described by each principal component for three different ranges of walking speed. Filled blue bars: Fractions of variability described by each principal component for the three subsets of data. Gray open bars: results for the same, but randomly permuted subsets of data (see results section). Percentages of PCs 1 to 5 are shown next to their respective bars. The minimum number of

steps per fly was set to 30 for all conditions. (A) Subset of steps between 2 and 4 BL s^{-1}, n = 23 flies. (B) Original subset of steps between 5 and 7 BL s^{-1} which was used for the PCA analysis (same data in Fig. 3A), n = 88 flies. (C) Subset of steps between 8 and 10 BL s^{-1}, n = 43 flies.

2.6.9 Walking speed-associated shifts in PC subspaces

Figure S5 shows the walking speed-associated changes of individual fly positions in the space spanned by PCs 2, 4, and 5 and PCs 1 and 3, respectively, (for absolute positions see Fig. 6 in the main text). Shifts are visualized as arrows. The left column (panels i) shows the shifts in position as they are located in the subspace. Flies with 30 or more steps in all three walking speed ranges are displayed with a single arrow with a bend at the position for 5 to 7 BL s^{-1}. All arrows point to the position of faster walking, while some start with the position at 5 to 7 BL s^{-1} (if fewer than 30 steps were recorded between 2 and 4 BL s^{-1}) and others end with position at 5 to 7 BL s^{-1} (if fewer than 30 steps were recorded between 8 and 10 BL s^{-1}). Flies which yielded the minimum of 30 or more steps for only one of the three conditions are not represented in this figure. The right column shows a centered version of the arrows shown on the left side for better comparison of their direction and length. Shifts are generally small and have no clear preference (also see panels ii for a centered version); this indicates that flies walked very similarly in the three respective walking speed ranges. Particularly the shifts in the space spanned by PCs 1 and 3 are almost non-existent. This supports the notion that these two PCs capture aspects of walking that are largely invariant and describe interleg coordination and overall movement of the legs. For more details see the Results section in the main text.



Supplementary Figure 5: Arrows showing the difference of average positions of individual flies for different walking speeds in the subspaces of PCs 1 to 5 shown in Fig. 6. The left column shows the absolute positions and their shifts as they are located in the subspace (also compare columns in Fig. 6). Flies with 30 or more steps in all three walking speed ranges are displayed with a single arrow with a bend at the position for 5 to 7 BL s^{-1}. All arrows point to

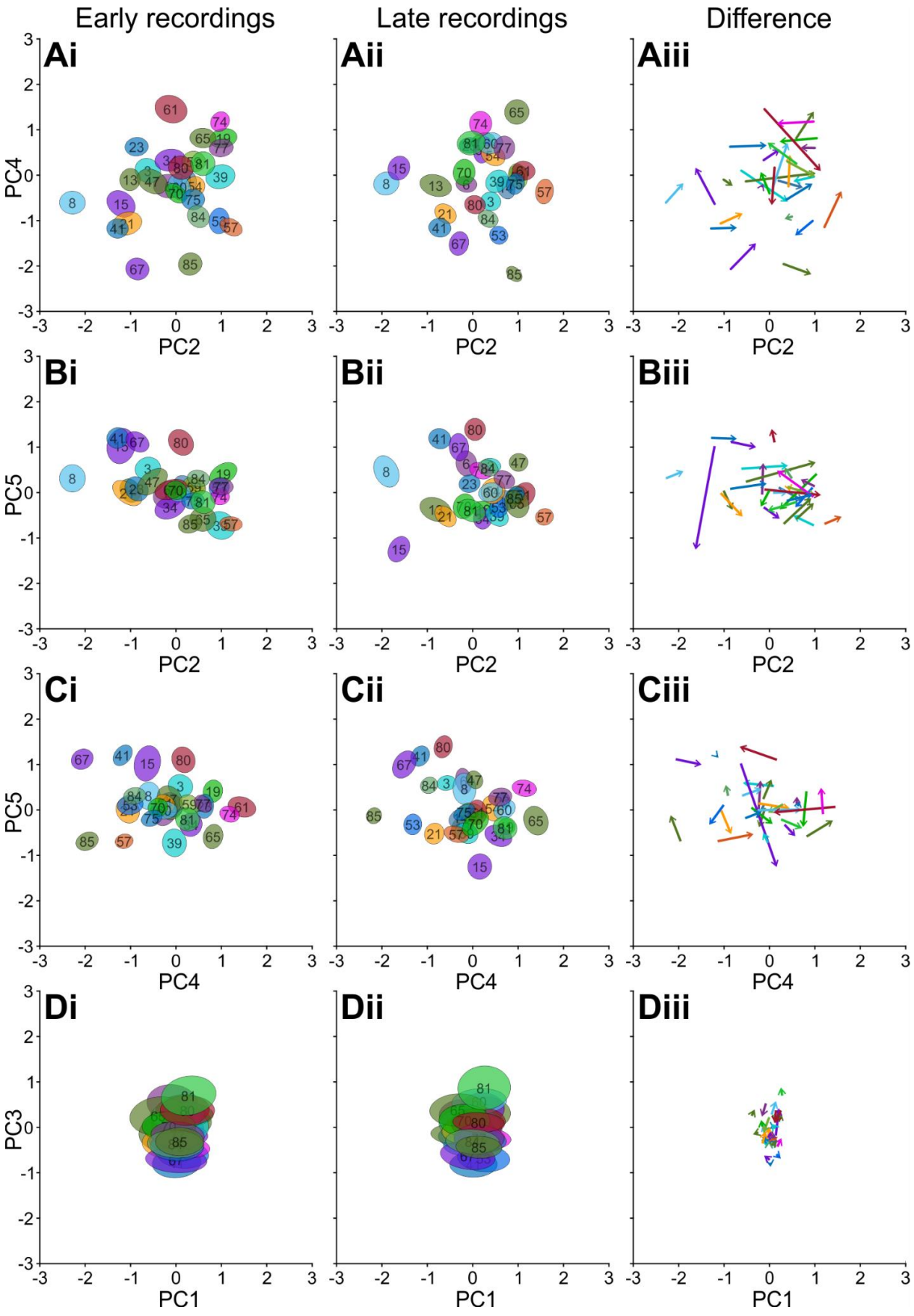
the position of faster walking, while some start with the position at 5 - 7 BL s^{-1} (if less than 30 steps were recorded between 2 and 4 BL s^{-1}) and others end with position at 5 - 7 BL s^{-1} (if less than 30 steps were recorded between 8 and 10 BL s^{-1}). Flies which yielded the set minimum of 30 or more steps for only one of the three conditions are not represented in this figure. The right column shows a centered version of the arrows shown on the left side for better comparison of their direction and length. The color code is identical for both columns and matches the coloring in Figure 6. (A) Subspace of PCs 2 and 4. (B) Subspace of PCs 2 and 5. (C) Subspace of PCs 4 and 5. (D) Subspace of PCs 1 and 3.

2.6.10 Recording time-associated shifts in PC subspaces

Flies regularly spent 2 hours in the experimental setup and produced walking behavior throughout that time. Valid signatures of individuality (i.e. the positions of flies in the subspace of PCs 2, 4, and 5 we proposed as related to individual aspects of walking) should remain largely unchanged over the course of an experiment. To verify this, we analyzed how steps recorded during early and late phases of the session relate to each other in that subspace. For this purpose, we grouped trials of individual flies into an early and a late set. These two sets contained trials that were recorded at least 60 minutes apart from each other, but were produced at the same speed. Each set needed to contain at least 30 steps; 29 flies in total met these criterium. Analogously to the approach in the analysis of walking speed-dependent effects (see Fig. 6) we then transferred these subsets of data into the PC space and examined how different their average positions were in PC space (Fig. S6). Generally, the flies' average positions remained largely constant in the subspace spanned by PCs 2, 4, and 5 (Fig. S6, panels i to iii). While there were some shifts, many were small and in the order of the depicted SD ellipses (0.3 times the standard deviation) or smaller (compare vector lengths in panels Aiii to Ciii). At the same time, flies in the two conditions were still clearly distinguishable. Furthermore, there were no general directional trends in the shifts, making systematic changes in walking behavior due to exhaustion during a prolonged experiment, for instance, unlikely. Taken together, this supports the notion that PCs 2, 4, and 5 describe inter-individual differences in walking behavior that persist at least over the course of the experiments that

were done here. The shifts in the subspace spanned by PCs 1 and 3 are even smaller; most shifts do not have an appreciable magnitude (Fig. S6Diii). This finding makes sense if we assume, that these two PCs describe variability in the walking behavior mostly related to coordination and speed and supports this notion even further.

There are a few singular exceptions. A particularly large shift was detected for fly 15 (purple color, Fig. S6Aiii to Ciii). Closer visual inspection of the underlying data revealed that, for unknown reasons, this particular fly clearly increased the spread of its hind legs towards the end of the recording session. While this individual is an exception of the rule and indeed did change its idiosyncratic way of walking, it became also evident that this produced a large signal in our analysis; however, combined with the fact that most flies did not produce these large signals this makes us more confident, that this approach can truly detect idiosyncrasies in walking behavior.



Supplementary Figure 6: Recording time-associated shifts in PC subspaces. First column: Average positions plus 0.3 times the standard deviations of flies for steps which were recorded

early during the recording session. Second column: Average positions of flies for steps which were recorded at least 60 minutes after the last step of the early recordings group (first column). The minimum number of steps for early and late recording sets was 30 each. The resulting number of flies that produced enough steps in both sets was n = 29. Third column: Arrows showing the difference between average positions (center of ellipses) of early (first column) and late (second column) recordings of individual flies in the subspaces spanned by PCs 1 to 5. Arrows point from average positions in early recordings to the average position during later recordings. (A) Subspace of PCs 2 and 4. (B) Subspace of PCs 2 and 5. (C) Subspace of PCs 4 and 5. (D) Subspace of PCs 1 and 3.

Movie 1: All data points (i.e. individual postures) of all flies in the analysis pool plotted, one after the other, on top of each other in 2D subspaces spanned by PCs 1 to 5. Arrangement as presented in Figure 5. Individual colors correspond to individual flies (n = 88) but are not unique throughout the complete visualization. Note the relatively invariant elliptic shape of the data associated with individual flies in the subspace spanned by PC 1 and 3.

Movie 2: Example of a high-TCS walking bout with the respective footfall pattern and visualization of PCs 1 and 3. Footfall pattern: black indicates swing phase, white indicates stance phase. Video view: tracked leg tips (round markers) with parts of their prospective movements for clarity. Open markers indicate swing phase, filled markers indicate stance phase. Note that the walking speed is similar to that in Movie M3. Movie M2 shows relatively high TCS values, resulting in a relatively flat shape of the time courses of PCs 1 and 3, with major amplitudes of oscillation in the direction of PC 1. The resulting elliptic shape can be cut in half horizontally to differentiate between the swing phases of the left (lower half) and right (upper half) tripod group, respectively.

Movie 3: Example of a low-TCS walking bout with the respective footfall pattern and visualization of PCs 1 and 3. Footfall pattern: black indicates swing phase, white indicates stance phase. Video view: tracked leg tips (round markers) with parts of their prospective movements for clarity. Open markers indicate swing phase, filled markers indicate stance phase. Note that the walking speed is similar to that in Movie M2. Movie M3 shows relatively low TCS values, associated with increased amplitudes in direction of PC 3 and simultaneously decreased PC 1 activity. Unlike to the example in Movie M2, the differentiation of swing and stance activity of individual legs or tripod groups in the subspace spanned by PC 1 and 3 is not directly possible for low TCS examples, as the activity of individual legs does not align with the score time courses as during more tripod-like coordination.

3 Chapter 3: Speed-related changes in kinematic variability in walking *Drosophila*

Preface

The following part of the present thesis was initially arranged as a paper draft, which is the reason for using the first person plural. It was then extended and adapted to serve as a monographic chapter 3 in this thesis which builds on the first part. Therefore, the introduction skips universal aspects of walking behavior, in general, and in *Drosophila*, in particular, which have already been introduced in the first part of this thesis, and starts with a more detailed view on walking speed as a central aspect of walking behavior in *Drosophila*.

3.1 Introduction

One of the most critical factors defining walking behavior is speed, which requires animals to maintain stability and coordination while increasing frequencies and velocities of single leg movements. In fruit flies (*Drosophila melanogaster*), walking speed has been shown to correlate mainly with interleg coordination and the duration of stance phases, while leaving other kinematic aspects, such as swing duration and stance amplitude, relatively invariant (DeAngelis et al., 2019; Mendes et al., 2013; Wosnitza et al., 2013). These speed-dependent modulations in walking behavior reflect the underlying neural mechanisms that allow flies to smoothly transition between different locomotor patterns without discrete gait shifts (Hughes, 1952; Nirody, 2021; Owaki et al., 2017; Tóth and Daun, 2019), as observed in many vertebrates (Hoyt and Taylor, 1981). Instead, a continuous modulation of interleg coordination patterns in response to changes in speed is well documented (DeAngelis et al., 2019; Wosnitza et al., 2013). At higher walking speeds, flies tend to exhibit more regular, tripod-like coordination patterns, where the two tripod groups (the ipsilateral front and hind legs, and the

contralateral middle leg) move in a synchronized anti-phasic pattern. In contrast, at slower speeds, the coordination not only shifts gradually to more complex patterns, but also becomes more variable (DeAngelis et al., 2019; Mendes et al., 2013). Which aspects of fly walking behavior are affected by this increase of variability at lower speeds and to what extent has not yet received a lot of attention. The present study aims to bridge this gap by analyzing in detail how walking speed correlates with both spatial and temporal variability in single leg movements and interleg coordination in *Drosophila melanogaster*.

The question arising naturally from the investigation of speed-dependent changes in stepping variability is why variability changes at all and whether it is to some extent advantageous for the fly. The common hypothesis is that the increase in variability at slow walking results from a more sensory feedback-driven control of leg movements, which allows for step-to-step adjustments in response to environmental perturbations (Schilling and Cruse, 2020). One obvious concept the sensory feedback at slow walking could be used on is static stability. Static stability quantifies how well the center of mass is covered by the support polygon, which consists of all legs with ground contact (Gonzalez De Santos et al., 1998; Szczecinski et al., 2018). The shortest distance of the projection of the center of mass to the support polygon serves as a measure for static stability. However, if the projection of the center of mass is outside of the support polygon, the respective posture is statically unstable. Due to their small size, flies are much more affected by elastic forces of their muscles and joints and viscous forces of the air surrounding them than by inertia (Hooper, 2012; Hooper et al., 2009), which is why static stability plays a much larger role for walking in flies compared to mammals. A previous study demonstrated that average phase relations between adjacent legs at both slow and fast walking follow the theoretical optimum for static stability (Szczecinski et al., 2018). However, Szczecinski et al. used averaged stance trajectories and modulated phase relations without noise for their analysis, leaving the question open how step-to-step variability and static stability relate to each other. Here, we address this issue by systematically analyzing whether spatial and temporal variability of leg movements have a positive or negative effect on static stability.

Another candidate for explaining the differences in variability with changing walking speeds is the higher complexity of interleg coordination patterns which is observed in slow walking. While strict tripod coordination can appear robot-like, with the three legs of each tripod group moving in synchrony, we observe much more versatile and asynchronous patterns of coordination below the upper end of the continuum. It seems reasonable that the ubiquitous anti-phasic relationship of directly adjacent legs in tripod coordination is easier to maintain by the nervous system than a set of more individual phase relations. Indeed, it is conceivable that the sensory information at slow walking is predominantly used to ensure that directly adjacent legs are not simultaneously in swing phase. Coordination rules (sometimes called Cruse rules, after Holk Cruse) prohibiting such unstable configurations have been described and experimentally demonstrated to be active between directly adjacent legs of the stick insect *Carausius morosus* (Cruse, 1990; Cruse et al., 2007). The first Cruse rule (1) says that the lift-off of an ipsilaterally anterior leg is prohibited if a middle or hind leg is in swing phase. The second (2) Cruse rule says that protraction of a leg is facilitated when the adjacent ipsilaterally posterior leg or contralaterally adjacent leg just finished its swing phase. The third (3) Cruse rule says that protraction is enforced in an adjacent ipsilaterally posterior or contralaterally adjacent leg with an increasing influence over the course of the stance phase. While similar experiments could not yet be successfully performed on *Drosophila* due to its much smaller size, it is highly plausible that similar rules are active here. However, it is yet unclear, how these rules relate to central pattern generators (CPGs), which have been shown to be crucial for generating the rhythmic motor patterns underlying locomotion in all animals (Grillner, 2006; MacKay-Lyons, 2002; Mantziaris et al., 2020). CPGs are neural networks that can produce rhythmic motor outputs necessary for timing and coordinating motor neurons without relying on rhythmic sensory input or descending signals. CPGs govern the intraleg coordination of individual legs and can be understood as oscillators which are coupled by signals coming from sensory feedback and CPGs of contra- and ipsilaterally adjacent legs (Berendes et al., 2016; Cruse, 1990). This interaction of the leg oscillators leads to synchronization, which is called entrainment, and entrainment of any oscillator is the strongest when all incoming influences

are coherent and do not contradict each other (Pikovsky et al., 2001). In this study, we follow the idea that less coherent and more contradicting influences exchanged between adjacent legs at low walking speeds could negatively affect the entrainment (Berendes et al., 2016; Jiménez et al., 2022) of those legs, which again might lead to increased variability. For this purpose, we have designed a new measure which is based on the first three Cruse rules and serves for quantifying the coherence of these rules as they are exerted on single legs. This coherence measure was then correlated with variability and tested against walking speed and tripod coordination strength (Wosnitza et al., 2013) with regard to their capabilities for predicting spatial and temporal variability.

Here, we provide a comprehensive view of how walking speed interacts with step-to-step variability in *Drosophila*. By separating our analysis into spatial kinematic variability (i.e. changes in AEP and PEP clustering) and temporal coordination metrics, we aim to disentangle the interplay between central coordinating influences between legs and the final motor output of individual legs. We show that static stability is higher for averaged steps compared to individual steps, indicating that step-to-step adjustments for improving static stability can be neglected as an explanation for the observed variability. Additional analysis of the static stability for different walking speeds revealed that faster walking is correlated with less stable configurations, thus tolerating less variability, which might have caused the nervous system to evolve in a way which automatically decreases variability with increasing speeds. Our coherence measure is designed to detect how coherent potential intersegmental influences are with each other. We use this coherence measure to substantiate our hypothesis that the observed decrease in variability with increasing walking speeds is caused by a stronger entrainment of the weakly coupled oscillators which control the leg movements. The results suggest that variability in walking is closely related to interleg coordination, with coherence serving as the most effective predictor among the parameters studied.

This study of how individual variability interacts with speed-dependent changes in walking behavior has general implications for understanding the neural control of movement in both invertebrates and vertebrates, as it highlights the dynamic interplay between speed,

coordination, and variability in motor output. It builds on previous work by examining how walking speed modulates both the coordination and variability of leg movements in *Drosophila*. Our findings suggest that while walking speed imposes constraints on kinematic variability, it also plays a crucial role in shaping individual walking patterns. By quantifying these relationships, we aim to contribute to a deeper understanding of the neural and biomechanical principles that enable insects to maintain locomotor stability across a wide range of speeds.

3.2 Materials and Methods

3.2.1 Fly strains and husbandry

Male flies of the wild-type strain Berlin-K (Bloomington Drosophila Stock Center (BDSC, #8522) were used for this study. Flies were raised on 12h/12h light/dark cycle at 25°C and approximately 60% humidity on a standard food medium (Backhaus et al., 1984). To control for age-related effects all animals used in this study were five days old. Prior to an experiment flies were isolated and starved for approximately 24 hours to increase their walking activity. To prevent desiccation, flies had access to moist tissue paper during this period of isolation.

3.2.2 Experimental setup

We used the same data set as for the first part of this thesis and hence the same experimental setup. For the sake of completeness we describe the most important features here again. For a more detailed description, please see **2.3.2**.

The recording arena consisted of an inverted glass petri dish (diameter: 60 mm) as the walking substrate and a watch glass (diameter: 100 mm) as the lid. We used a camera (model: VC-2MC-M340, Vieworks, Anyang, Republic of South Korea) equipped with an object-space telecentric lens (focal length 55 mm, model: Computar TEC-55, CBC America, Cary, North Carolina, USA). The camera was directed at the experiment chamber from below via a surface mirror. The camera view was focused on a square area in the center of the arena (30 mm side length, resolution 1000 by 1000 pixels, 33.3 pixels mm^{-1} or 30 µm $pixel^{-1}$). The scene was

illuminated with 60 infrared (IR) LEDs (wavelength: 890 nm, opening angle: 20°) arranged in a concentric ring around the chamber. Video data was acquired at 200 Hz and a shutter time of 400 µs.

3.2.3 Behavioral experiments

As stated above, we used the same data set as for the first part of this thesis. The most important aspects are described here again, please see **2.3.3** for a more detailed description.

Single flies were recorded for a total duration of up to 3 hours. During this time, flies walked spontaneously in the chamber and frequently crossed the capture area.

Video acquisition, online data evaluation during experiments, and general high-level hardware control were implemented with custom-written software in MATLAB (2018b, The Mathworks, Natick, Massachusetts, USA).

3.2.4 Processing of video data

Processing of video data was identical to the first part of this thesis. Here we describe the most important aspects again, please see the Methods section in **2.3.4** for a more detailed description.

Instances of straight walking in the complete set of walking trials were detected with custom-written algorithms. In each video frame of these straight walking segments the tarsal tips of all six legs and two additional body parts (the neck and the posterior tip of the abdomen) were detected with DeepLabCut (Mathis et al., 2018).

Swing and stance phases of all legs were determined automatically based on the respective speeds at which tarsal tips moved in an arena-centric coordinate system. The last position of the tarsal tip on the ground before lift-off was defined as the posterior extreme position (PEP) for that step. The first tarsal position with ground contact after a swing phase was identified as the anterior extreme position (AEP). A complete step of a leg was defined as its movement between two consecutive lift-off events, i.e. a swing phase followed by a stance phase. In

contrast to steps of individual legs, step cycles (SCs) were defined as follows: start and end of an SC were determined by the respective step of the right middle leg. All six tarsal tip positions for this interval comprised the data of one SC. The definition of a SC in this study was exclusively used for the static stability analysis.

3.2.5 Tripod coordination strength (TCS)

Tripod coordination strength (Ramdya et al., 2017; Wahl et al., 2015; Wosnitza et al., 2013) was used to quantify the synchronicity of the swing activity in the two tripod groups. Briefly, the TCS was calculated as follows: for each tripod group (a set of ipsilateral front and hind legs and the contralateral middle leg), the time in which all three legs were simultaneously in swing phase was divided by the time from the earliest swing onset to the latest swing termination in any of these three legs. Hence, a perfect overlap of all three swing phases resulted in a maximal TCS value of 1 and would correspond to canonical tripod coordination. The minimal TCS value of 0 was assigned in cases of no overlap between swing phases in a tripod group. The TCS values of the tripod groups were averaged.

3.2.6 Definition of the tripod cycle

To define a unit of walking behavior which allowed for determination and comparison of different measurements regarding interleg coordination we came up with the definition of the tripod cycle. For our purposes it was important to include the complete swing and stance phases of all six legs. The tripod cycle therefore starts with the first lift-off of one tripod group and ends with the last touch-down of the other. This way a tripod cycle always contains a whole cycle (a complete swing and complete stance phase) of all six legs and hence enables calculation of TCS values for both tripod groups, which were averaged to obtain a single TCS value for each tripod cycle. Measurements with more than a single value over the course of a tripod cycle, such as walking speed, were averaged over the time of the tripod cycle. Tripod cycles were not allowed to overlap, ensuring that every recorded time point only was part of a single tripod cycle.

3.2.7 Spatial variability analysis

To evaluate spatial variability, we analyzed the clustering of anterior and posterior extreme positions (AEPs and PEPs, i.e. touch down and lift-off positions). We effectively determined the 2D-variance of AEP and PEP distributions by taking the average of the two eigenvalues. For this purpose, for one fly at a time, x- and y-components of AEPs or PEPs within a certain range of walking speeds were collected and the eigenvalues of the respective covariance matrix were calculated. This calculation was performed for 30 or more steps per fly and walking speed range, as we qualitatively assessed that 30 steps in a narrow range of walking speeds provide a representative sample of the underlying distribution. Because the spatial variability between flies differed largely we normalized the results by dividing all values by the variability between 5 and 6 BL s^{-1}. This way variability values of different flies became directly comparable. Average variability was plotted for speed ranges with data of 5 or more flies.

3.2.8 Temporal variability analysis

Temporal variability was assessed by calculating the circular standard deviation of phase relations between adjacent legs. Phase relations describe whether and how much the activity of adjacent legs is shifted against each other. For the calculation of phase relations a reference and a dependent leg need to be defined. The phase value tells us about the relative shift between the two legs. Perfect in-phase activity has a phase value of 0°, while alternating anti-phasic activity has a phase value of 180°. Delayed activity of a dependent anti-phasic leg would result in values above 180°, while premature activity would result in values below 180°. The phase of the reference leg was defined to start and end with consecutive liftoff events. At the time of the lift-off of the dependent leg the current phase value of the reference leg was taken as the phase value of that pair. For ipsilateral leg pairs we always selected the more posterior leg and for contralateral leg pairs we selected the left leg as the reference leg. Analogously to the procedure for spatial variability, we only considered temporal variability values for flies and speed ranges with 30 or more measurements for each pair of adjacent legs. Variability values were averaged separately for contra- and ipsilateral pairs and normalization was achieved by

dividing all results by the variability between 5 and 6 BL s^{-1}. Average variability was plotted for speed bins with data of 5 or more flies (Figure 5).

3.2.9 Static stability analysis

Static stability in this context describes how stable configuration of legs in stance phase is (Gonzalez De Santos et al., 1998). For this purpose, the leg tip positions of all legs in stance phase form a support polygon, which is defined by the area between these points without any concave outlines. In other words, a leg does not contribute to the support polygon if its position is already covered by the other legs. A configuration is considered unstable if the projection of the center of mass (COM) is outside of the support polygon. If the projection of the COM is inside the support polygon, the shortest distance from the COM to an edge of the support polygon equals the static stability for that configuration. Calculation of static stability for single SCs was performed as described by Szczecinski et al.. Namely we used the same procedure to estimate the COM of the flies and defined all stance trajectories relative to that COM. In contrast to the model approach from Szczecinski et al. we used data from individual steps for our analysis. We normalized the length of each SC to 100 data points by interpolating between the recorded leg tip positions. For each of these 100 time points we constructed a stability polygon between the leg tip positions of all legs which were in stance phase at the time and defined the minimum distance between the COM and an edge of the polygon as static stability. We averaged the static stability values over the course of SCs to be able to compare between SCs.

To evaluate the influence of variability on static stability, we calculated average SCs for individual flies and compared these averaged SCs to the individual SCs with regard to their static stability. Because every SC has slightly different phase relations between legs and varying step and swing durations, we could not simply average the leg tip positions as they were, because this would have meant averaging stance positions from one SC with swing positions from another and vice versa. Instead, we extracted swing and stance positions separately from all SCs before we normalized and averaged them. We then calculated average

phases and average swing and stance durations for all legs and used all these components to construct an artificial SC which represented the average of each parameter as closely as possible.

To differentiate between the influence of spatial and temporal variability on the static stability we created partially averaged SCs which consisted either of the original stance trajectories and the averaged phase relations, or of the averaged stance durations and the original phases of the individual SCs. In other words, we either applied the averaged phase relations and swing and stance durations on the individual stance trajectories, or we applied the individual phase relations and swing and stance durations on the averaged stance trajectories. In the first case, we used the kinematics as performed by the animal, but with the timing of the average step of that same animal. In the second case, we used the averaged stance trajectories, but with the timing of the individual SCs, as they were performed by the animal. This way we could assess which effect the spatial and the temporal component each have on the static stability.

For all three versions of fully or partially averaged SCs we used the exact same procedure as for the original SCs to calculate the static stability. For each fly and speed range we used exactly 30 steps to perform this analysis and averaged the resulting average stability values per fly to compactly plot and compare them in a single figure.

3.2.10 Coherence

To substantiate the idea that the relative timing of intersegmental influences, as described by the Cruse rules (Cruse, 1990; Cruse et al., 2007), might have a direct effect on the degree of observed kinematic variability, we designed a new measure for quantifying the temporal alignment of these influences. We hypothesize that contradicting influences from different neighboring legs might result in increased variability, while well-aligned and temporally coherent influences would create a strong entrainment between the oscillators, which again should lead to more robust, repetitive, and, thus, less variable patterns of motor outputs.

The coherence of intersegmental sensory influences was estimated by assuming the presence of the so-called Cruse rules (Cruse, 1990) in *Drosophila*. For our approach, we took only the first three rules. All of these have temporal effects and are exclusively active between directly adjacent legs, either ipsilaterally or contralaterally. These rules say that (1) lift-off is suppressed in an anterior leg while an adjacent and posterior middle or hind leg is in swing phase; (2) early protraction is facilitated in anterior legs as well as contralateral legs at the end of the swing phase; and (3) late protraction in a posterior or contralateral leg is enforced by the swing phase in an anterior or contralateral leg with an increasing influence over the course of the stance phase.

None of these influences have a decisive effect, they merely increase or decrease the likelihood of lift-off in the affected leg at a given time. In effect, their combined influence determines when the PEP is reached in a given step of a leg. Rule 1 is a suppressing influence (i.e. lift-off is delayed to later times), while rules 2 and 3 have an activating effect (i.e. lift-off occurs at earlier times). We wanted our coherence measure to be low if the effect of rule 1 overlapped often with effects of rules 2 and 3, because that would mean a contradictory influence and weaken the entraining effect of these rules, while perfect alignment of rules 2 and 3 would strengthen such effects in our theory.

For each individual leg, we combined these negative (suppressing) and positive (facilitating) effects as follows: Rule (1) was represented with a value of -1 over the course of the swing phase. Rule (2) incremented to +1 over the next three frames and fell then back to zero. Rule (3) incremented its effect from 0 to +1 over the course of the stance phase. All effects were summed up and the standard deviation over the course of the selected time frame was taken as our coherence measure. The more contradicting and unaligned influences occurred, the lower the resulting standard deviation was, and the less contradicting and better-aligned the influences were, the higher the resulting standard deviation was. For more details and a visualized description please see Results section (Figure 3).

To evaluate our coherence measure, we used our extensive data set to extract how phase relations, as well as swing, stance, and step durations changed with walking speed and used these values to create artificial foot fall patterns, which we then used our coherence algorithm on. The resulting footfall patterns and their coherence values are plotted in Figure X C and D, as well as the coherence values for all straight walking bouts in our data set.

3.2.11 Assessing spatial and temporal variability through parameter range comparison

To analyze how closely the three parameters – walking speed, TCS, and coherence – are related to spatial and temporal variability, we kept the range of one parameter constant while allowing for wide range variation of another. In other words, we had always one narrowed and one free parameter and tested all data points in the narrow range along the dimension of the free parameter. Broken down into single steps we have proceeded as follows:

For every fly with 60 or more steps within the narrowed range we divided the data into an upper 50% and a lower 50% half along the dimension of the free parameter. For both halves we calculated the spatial and temporal variability. We normalized the results by dividing the upper half by the lower half to directly show, whether the variability was increased (values above 1) or decreased (values below 1) for higher values of the free parameter. Then we swapped the role of both parameters (former free one was now narrowed and former narrowed one was now free) and performed the same steps again. We tested for statistical significance both between the parameters (Wilcoxon rank sum test) and between upper and lower halves within each parameter (Wilcoxon signed rank test).

For all parameters, the width of the narrowed ranges was set to values which contained exactly 17 flies with 60 or more steps in that range. We found that this number was a good compromise between the contradicting goals of keeping the ranges as narrow as possible and having representative sample sizes. This way we could directly test parameters against each other with regard to their capacity for predicting variability.

3.2.12 A proxy for spatial and temporal variability

Variability is mathematically undefined for single measurements. Until this point, we therefore only calculated variability for 30 or more measurements. This approach, however, makes it necessary to group measurements before assigning a variability value to all of them. To directly test different parameters for their predictive power regarding the variability, we aimed to find a proxy for variability which can be assigned to every tripod cycle. For spatial variability, we decided to use the deviation of single AEPs and PEPs from the average position of the respective group. We first calculated the average positions of AEPs and PEPs for every fly individually (Figure 1 A). Then we calculated the distance of each AEP and PEP to these mid-points (Figure 1 B). Finally, we divided the distances by the standard deviation of that group to obtain Z-Scores (Figure 1 C). We averaged the results for all six legs, resulting in one AEP and one PEP variability value in form of a Z-Score.

For temporal variability we used a similar approach. We calculated average phases for all adjacent leg pairs and took the absolute difference between each single phase values and the respective average value. Finally, we again calculated Z-Scores by dividing the resulting values by the standard deviation. We averaged the Z-Scores for all contra- and ipsilateral leg pairs, resulting in two values for each tripod cycle.

For depiction of the results we calculated probability density functions for each parameter and the corresponding Z-scores and displayed it in form of heat maps with contour lines. To quantitatively assess the correlation of our spatial and temporal variability proxies (Z-Scores) with walking speed, TCS, and coherence, we performed a simple linear regression analysis and stated the slopes and r-squared values in the figures.

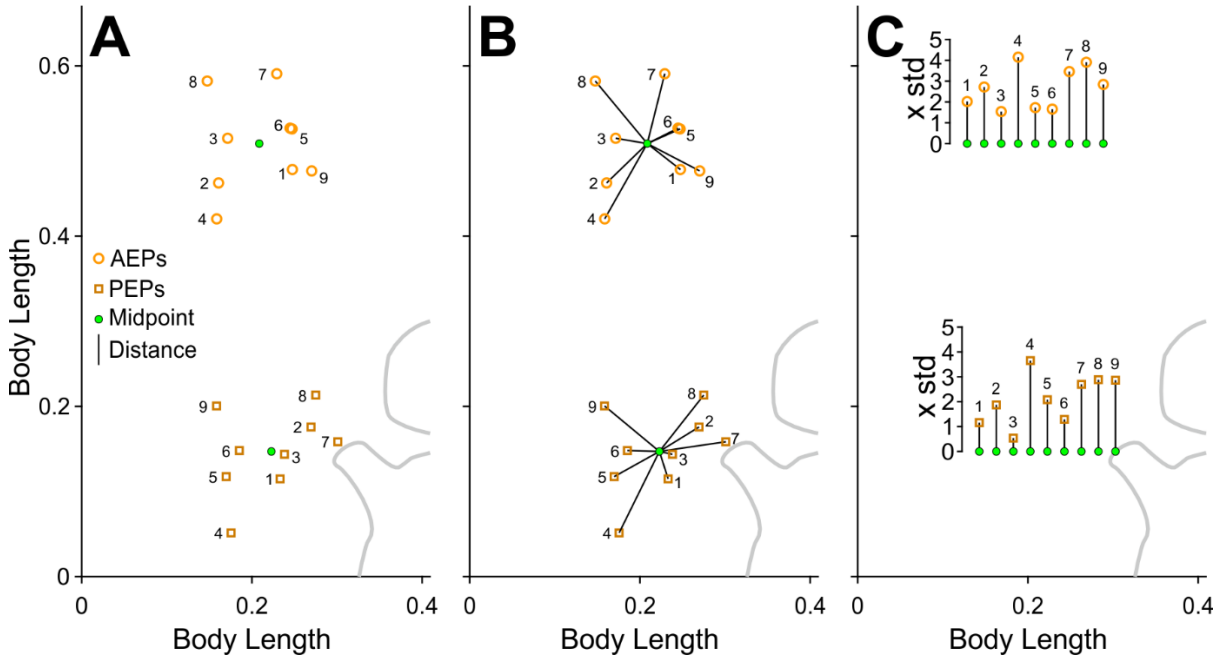


Figure 1: Deduction of spatial proxies for variability from AEPs and PEPs. For clarity, only AEPs and PEPs of the left front leg are shown. The grey lines on the lower right are part of the fly silhouette indicating the relative position of the fly body. (A) A group of AEPs (circles) and PEPs (squares) of the right front leg with their arithmetic mid-point (green dot). (B) Straight lines from each AEP and PEP to their respective mid-point indicate the distance to the average. (C) AEPs and PEPs are aligned and their distance to the mid-point is on display. The scale to the left demonstrates how the distances are measured in terms of the standard deviation of the whole group of either AEPs or PEPs. The whole procedure was performed for every fly on all its AEPs and PEPs to enable a more proportional assessment of changing variability. Results for the six legs were averaged to obtain a single value for AEP and PEP variability for every SC.

3.3 Results

3.3.1 Low-level parameters change with walking speeds

Walking speed is a central parameter of walking behavior and significantly correlates with most other parameters. Here, we measured basic kinematic parameters of walking behavior and how they relate to walking speed for our extensive data set of straight walking *Drosophila*. Panels A to F in Figure 2 combine data from all six legs. Panels G and H combine data from the left and right body side. Figure 2A shows how overall step duration is drastically reduced

with increasing walking speeds, while most of this reduction results from the decrease in stance duration and only to a small fraction from decreasing swing duration (Figure 2 B and C). Generally, all aspects of step duration vary more at low walking speeds.

The stronger reduction of step duration compared to swing duration results in a reduction of the duty factor from 0.75 at low walking to around 0.55 at high walking speeds (Figure 2 D). All effects mentioned above together result in a highly linear and more than two and a half fold increase of step frequency from 6.5 Hz to 16.5 Hz (Figure 2E). Aside from the before mentioned temporal aspects, flies on average also almost double the amplitude of their strides (distance between AEP and consecutive PEP in the fly coordinate system) with increasing walking speeds from 0.26 BL to 0.51 BL (Figure 2F).

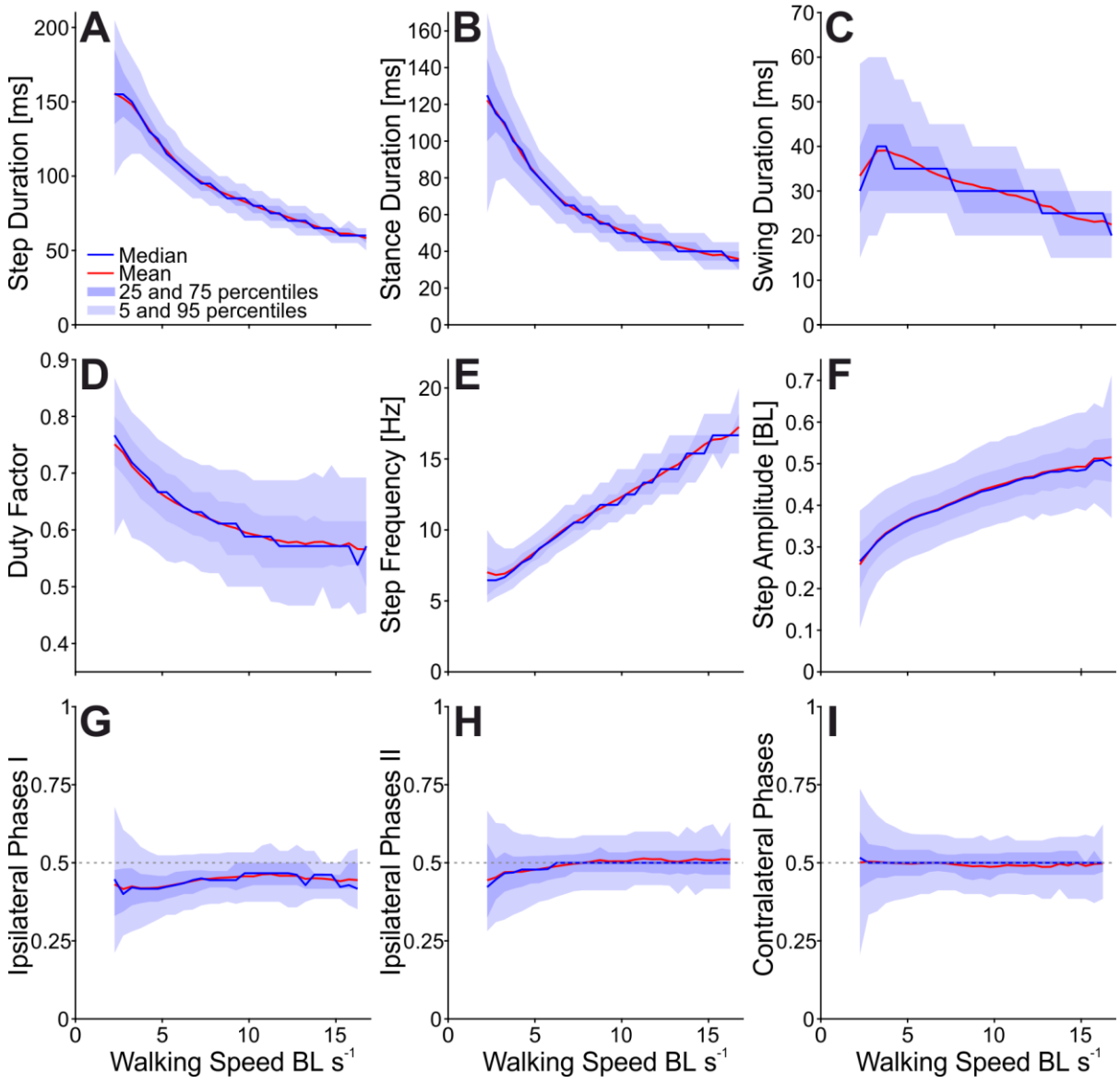


Figure 2: Basic walking parameters plotted against walking speed. (A) Step duration. (B) Stance duration. (C) Swing duration. (D) Duty factor. (E) Step frequency. (F) Step amplitude (G) Ipsilateral phase relations between middle legs and front legs. (H) Ipsilateral phase relations between hind legs and middle legs. (I) Contralateral phase relations. Panels A to F combine data from all six legs. Panels G and H combine data from the left and right body side. For all plots, means, medians, 25 and 75, as well as 5 and 95 are shown. The bin width was set to 0.5 BL s^{-1}. The minimum number of measurements per bin was set to 30.

Ipsilateral phase relations change slightly with increasing walking speed, approaching a value of 0.5. Phase relations between middle and front legs (Figure 2H) reach an average value of 0.5 around 7 BL s^{-1} and stay constant from there on. Contralateral phase relations show a stable average of 0.5 over the whole range of walking speeds (Figure 2I). All phase relations show more variation at low walking speeds compared to faster walking.

3.3.2 Coherence analysis yields similar results for artificial data and recorded behavior

Our coherence measure aims to quantify how orderly (explanation following below) the coordinating influences coming in from directly adjacent legs are. For our measure, we assumed the presence of the so-called Cruse rules (Figure 3A) in *Drosophila*. Figure 3 Bi to Biii show exemplary footfall patterns with the coordinating influences from neighboring legs plotted below. The shape (horizontal bar, vertical bar, or ramp) indicates which rule is exerted (see bottom of Figure 3A for legend) and the color gives information about the leg which caused the influence. In Figure 3Bi we see multiple examples of rule 1 (horizontal bar, suppressing) and rules 2 and 3 (vertical bar and ramp, facilitating) overlapping. While Bii and Biii still show some examples of overlapping suppressing and facilitating influences it becomes clear, that the orderliness of intersegmental influences increases with higher walking speeds and TCS values. Increasing orderliness in this context means that rules 2 and 3 overlap more often with them self and/or each other and especially that rule one shows less overlap with the contradicting influences of rules 2 and 3. Some examples for these increasing orderliness with increasing walking speeds can be seen in the first row of panels Ci to Ciii in Figure 3. The highest displayed speed and TCS in panel Biii shows a high degree of order with suppressing influences being mostly separated from facilitating ones, which indicates that the signals between adjacent legs are less contradicting. At the same time, rules 2 and 3 are better aligned with each other, resulting in a summation of facilitating influences, which might again lead to a more robust entrainment of neighboring legs. Figure 3C presents examples of panels Bi to Biii to illustrate how the coherence value is calculated. First the individual rules are translated into signals; Rule (1) was represented with a value of -1 over the course of the swing phase. Rule (2) incremented to +1 over three frames and fell then back to zero. Rule (3) incremented its effect from 0 to +1 over the course of the stance phase. All effects were summed up and the standard deviation over the course of the selected time frame was taken as our coherence measure. We always averaged coherence values of all six legs and for the whole time period

analyzed, which is why the coherence values resulting from the examples in C do not match the values plotted in E.

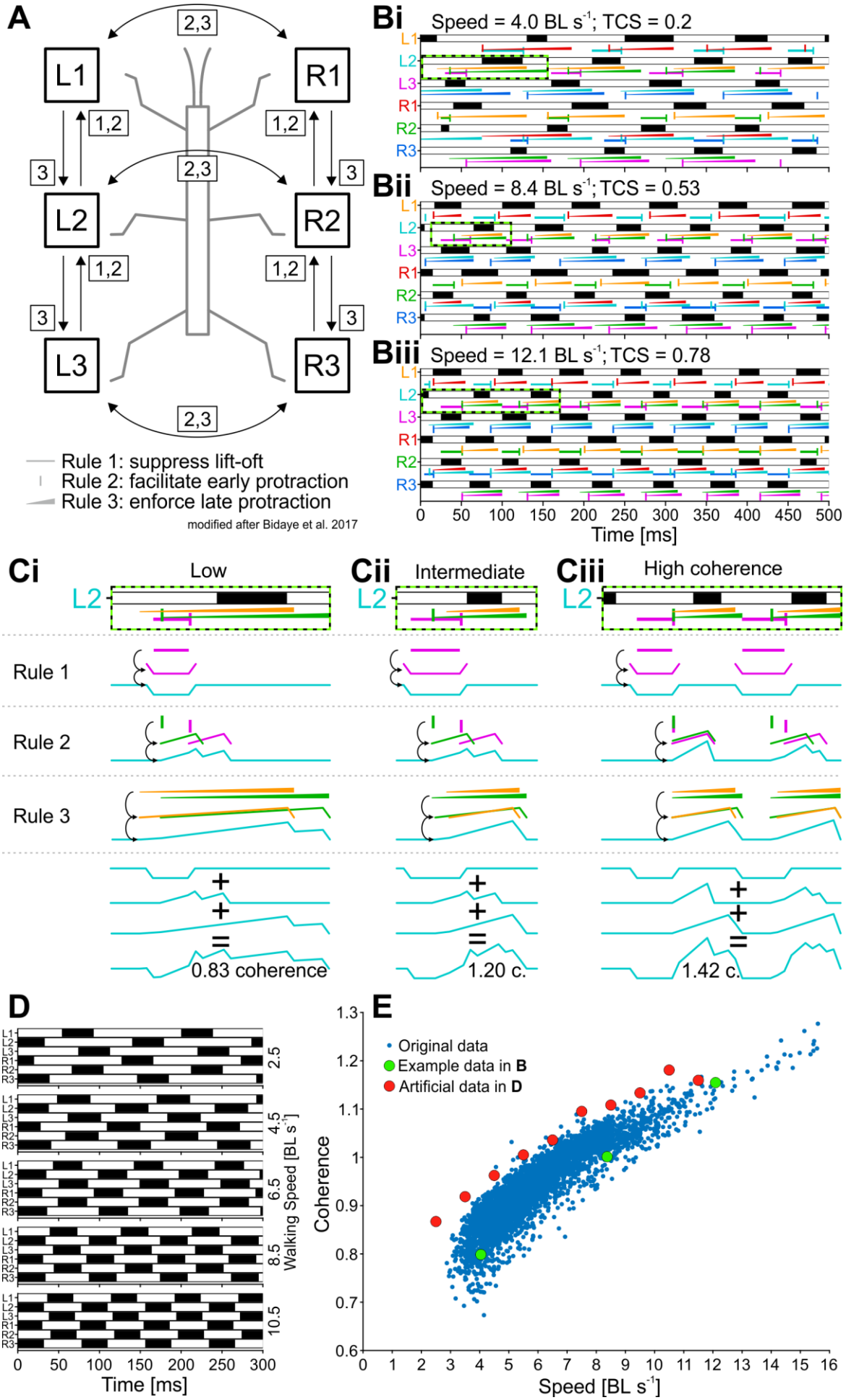

A
L1
R1
L2
R2
L3
R3
2,3
1,2
3
Rule 1: suppress lift-oft
Rule 2: facilitate early protraction
Rule 3: enforce late protraction
modified after Bidaye et al. 2017
Bi
Speed = 4.0 BL s-1; TCS = 0.2
Bii
Speed = 8.4 BL s-1; TCS = 0.53
Biii
Speed = 12.1 BL s-1; TCS = 0.78
Time [ms]
Ci
Low
Cii
Intermediate
Ciii
High coherence
Rule 1
Rule 2
Rule 3
0.83 coherence
1.20 c.
1.42 c.
D
Walking Speed [BL s-1]
Time [ms]
E
Original data
Example data in B
Artificial data in D
Coherence
Speed [BL s-1]

Figure 3: Overview and rationale of our coherence measure. (A) Schematic of a stick insect with the first three so called Cruse rules below. The numbers next to the arrows indicate from which leg onto which other leg a rule is exerting its effect. Arrows with two heads indicate, that the respective rules are active in both directions. (B) Exemplary footfall patterns for walking bouts with low (4.0 BL s^{-1}, 0.2 TCS), intermediate (8.4 BL s^{-1}, 0.53 TCS) and high (12.1 BL s^{-1}, 0.78 TCS) walking speeds and TCS values. Colored bars and triangles below footfall rows indicate when the respective leg is affected by a rule and the color indicates, which other leg it is affected by. Horizontal bars indicate the action of rule 1, vertical bars of rule 2, and triangles of rule 3. The effect of rule 1 is lift-off (=start of protraction) suppressing, while effects of rules 2 and 3 are protraction facilitating/enforcing. Overlap of rule 1 with rules 2 and/or 3 is interpreted as incoherent, while perfect overlap of rules 2 and 3 results in increased coherence values. (C) Examples for the determination of the coherence value. Top row shows exemplary fractions for the left middle leg (L2) from Bi to Biii. Following rows visualize how rules are translated into curves and summed up. Our coherence measure equals the standard deviation of the summed curves. For a complete footfall pattern we averaged the results for all individual legs. (D) Artificial footfall patterns generated after the mean values for step, swing, and stance durations as well as contra- and ipsilateral phase relations shown in Fig. 1 A to F. Selected mean walking speeds at which the values were extracted are displayed to the right of each footfall pattern. Coherence values of the depicted footfall patterns are shown in orange in (E). (E) Coherence values of all straight walking bouts in the data set (blue), the example trials shown in (B), and of the artificially generated footfall patterns shown in (D) (orange).

Figure 3D displays examples of artificially created footfall patterns, which were produced with the average values for stance duration, duty factor, and phase relations shown in Figure 2. The corresponding coherence values for the patterns in Figure 3D are also plotted in panel E, which depicts the average coherence values of all straight walking bouts in our data set. For both recorded and artificially produced data, our coherence measure is positively correlated with walking speed, while the results for the artificial data (orange dots) trace the upper limit of the distribution of recorded trials (blue dots). However, these results show that our coherence measure indeed reflects mainly the speed dependent changes in step durations and phase relations as shown in Figure 2, rather than being a product of the reduced variability at higher walking speeds. In contrast to TCS, which also reflects the speed dependent changes of step durations and phase relations, our coherence measure is based on signals between directly

adjacent legs which have been demonstrated experimentally in the stick insect (Cruse, 1990; Cruse et al., 2007) and might hence be more physiological than the definition of TCS.

3.3.3 Variability decreases with increasing walking speeds

To analyze the connection between walking speed and motor variability in detail we separately measured and compared spatial and temporal variability for different speed ranges. Spatial variability was measured in form of 2-dimensional (x- and y-components) clustering of anterior and posterior extreme positions (AEPs and PEPs, i.e. touch down and lift off positions, Figure 4 Ai and Aii). For each leg and for both AEPs and PEPs the eigenvalues were calculated and averaged to provide an estimate of the variability in the bivariate distributions. Temporal variability was measured in form of the circular standard deviation of phase relations between ipsi- and contralateral leg pairs (Figure 5 Ai and Aii).

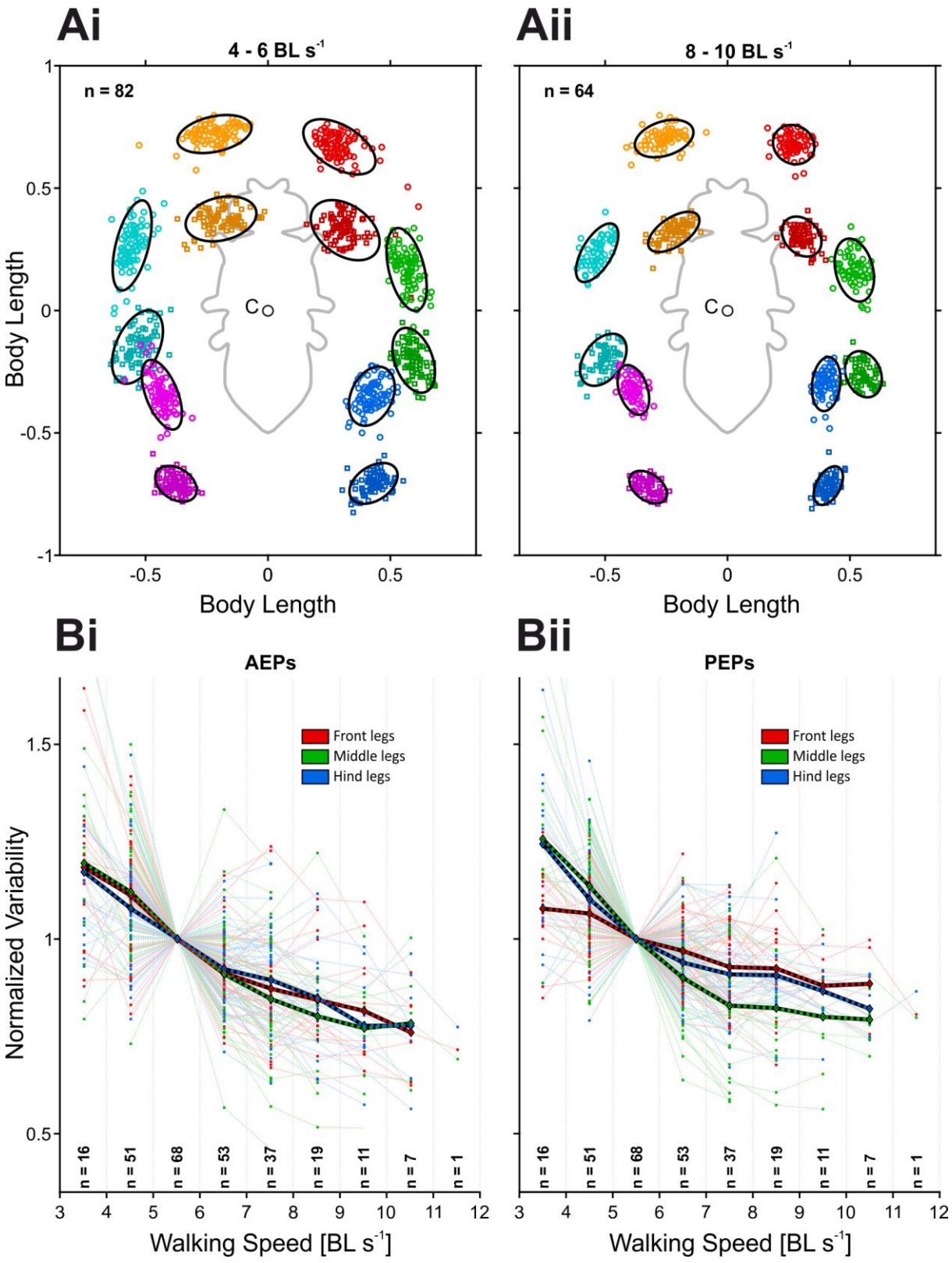


Figure 4: Spatial variability decreases with increasing walking speeds. (A) Example of the relationship between extreme position clustering and walking speed. Ellipses indicate twice the standard deviations of the bivariate distributions. AEPs are depicted as circles and PEPs as squares. Leg color coding: red = R1; green = R2; blue = R3; orange = L1; teal = L2; magenta = L3. (B) Variability distribution of AEPs and PEPs for different speed ranges. Each dot

represents the normalized variability of 30+ steps of one fly. Only flies with 30 or more measurements between 5 and 6 BL s^{-1} were added to the analysis pool, because this speed bin was selected for normalization purposes. The variability distributions of the two body sides were averaged (30+ steps on each side). Faint dashed lines connect dots of individual flies, thick black lines connect the average values of all bins. Average values were calculated for all bins with 5 or more flies. ***n***-values display the number of flies in each speed bin.

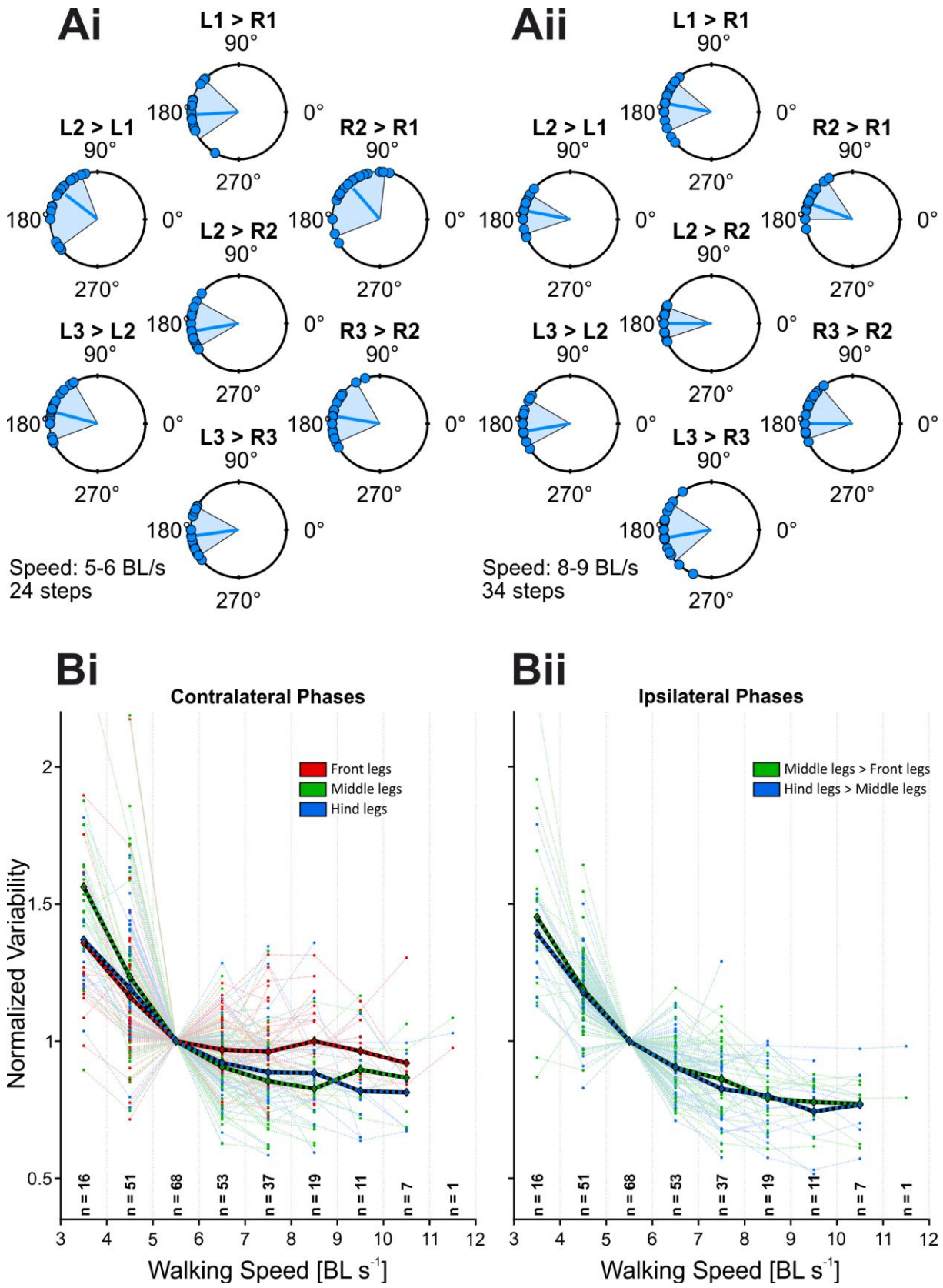


Figure 5: Temporal variability decreases with increasing walking speeds. (A) Phase diagrams of a representative fly for different speed ranges. Each dot represents the current phase of the reference leg at the time of the dependent leg lift-off. The reference leg is for contralateral pairs the left and for ipsilateral pairs always the posterior leg. The vector lines (blue) point to the

average phases and have average resultant length. Light blue wedges indicate 5 and 95 percentiles of the distribution. (B) Variability distribution of ipsi- and contralateral phase relations for different speed ranges. Each dot represents the normalized variability of 30+ steps of one fly. Only flies with 30 or more measurements between 5 and 6 BL s^{-1} were added to the analysis pool, because this speed bin was selected for normalization purposes. The variability distributions of ipsilateral pairs of both body sides have been averaged (30+ steps for each pair). Faint dashed lines connect dots of individual flies, thick black lines connect the average values of all bins. Average values were calculated for all bins with 5 or more flies. ***n***-values display the number of flies in each speed bin.

We measured spatial and temporal variability for individual flies and different speed ranges and normalized the result by dividing all values by the variability between 5 and 6 BL s^{-1} to achieve comparability. The average of the relative variability values for speed ranges was taken to illustrate the trend. Normalized and averaged variability values for all flies reveal a distinct decrease of both spatial (Figure 4 Bi and Bii) and temporal (Figure 5 Bi and Bii) variability with increasing walking speeds. The average decrease for AEP clustering is almost linear and similar for all legs. The decrease for PEPs is more prominent for walking speeds between 3 and 8 BL s^{-1} and is substantially weaker for front legs.

The decrease in temporal variability described by contralateral phase relations is again weaker for front legs and reached a negative plateau for values above 6 BL s^{-1}. Ipsilateral phase relations are almost identical between middle and front legs and hind and middle legs and show a very smooth transition from strong decreases at low walking speeds towards a negative plateau at higher walking speeds.

Overall, variability decreased more strongly for the temporal aspects. Spatial variability decreased linearly and temporal variability seemed to reach a negative plateau at high walking speeds. However, all measured aspects on average showed a clear and consistent decrease of variability in walking behavior with increasing walking speeds.

3.3.4 Static stability necessitates increased accuracy at higher walking speeds

The static stability of original step cycles (SCs) was analyzed and compared to standardized steps to test how sensitive static stability is to temporal and spatial variability.

Static stability is a likely candidate for what flies want to control for when generating their motor output, because stumbling or falling due to instable configurations is inefficient and can, for example, obstruct reproductive success in male flies, because they have to successfully chase a female to reproduce (Bateman, 1948). Unsystematic variability can be understood as random deviations from the intended movement and should therefore result in less stable stepping for most cases. Systematic variability on the other hand can result from neural or biomechanical mechanisms which have the purpose to optimize on a step-by-step basis for a certain parameter of walking behavior, such as static stability.

The hypothesis for our analysis was as follows: if the higher level of variability at low walking speeds results from spontaneous adjustments that increase the acute static stability, then individual steps should be more stable, on average, compared to averaged steps, because averaging largely eliminates the exact matching of timing and positioning between the legs for a particular step or SC.

However, the stability of the averaged step was higher than the average stability of the individual steps for all flies and speed ranges. This shows that flies – at least from the static stability view point – should aim to perform an idealized step which they deviate from due to their unsystematic variability.

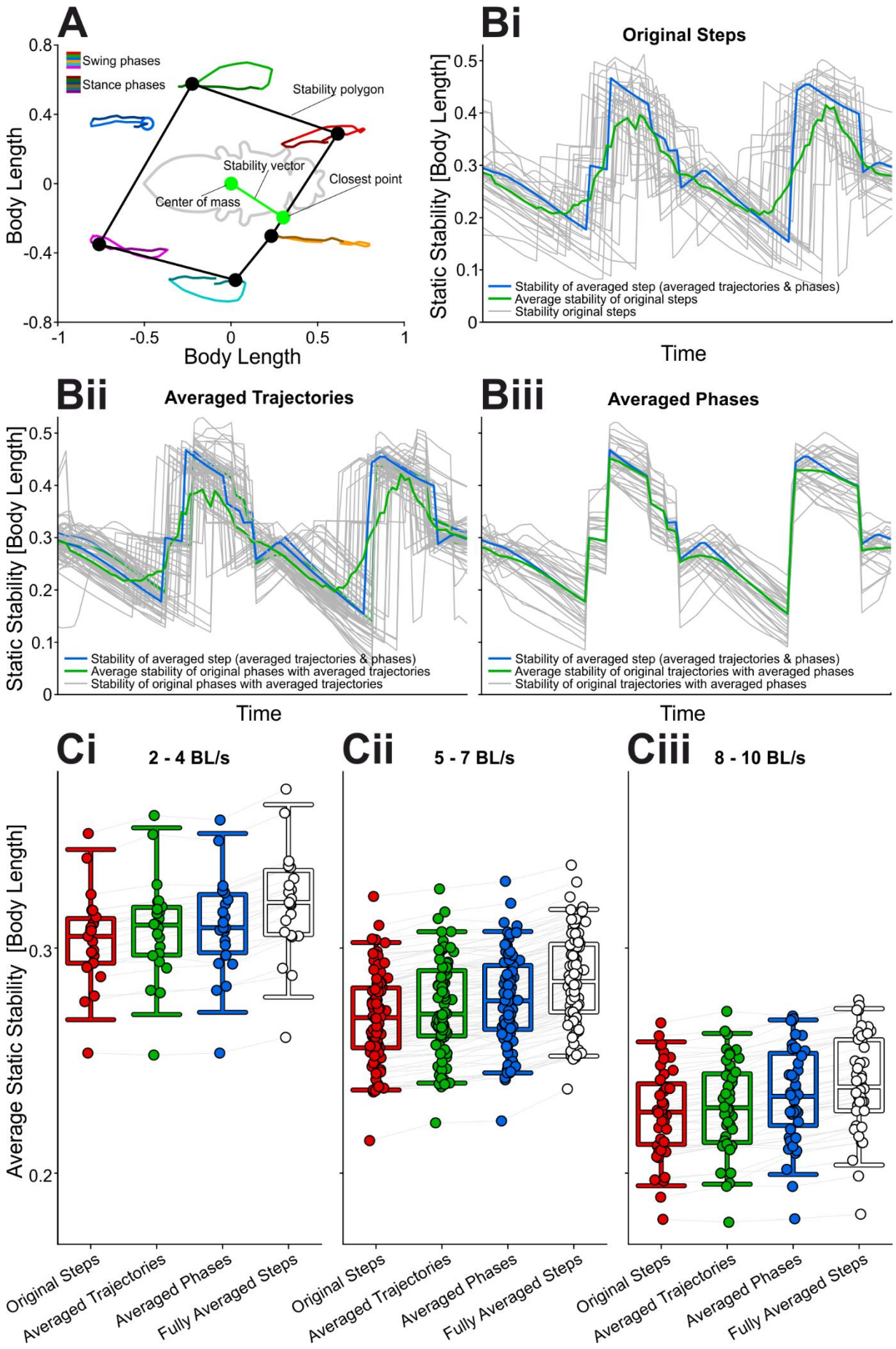

A
Swing phases
Stance phases
Stability polygon
Stability vector
Closest point
Center of mass
Body Length
Body Length
Bi
Original Steps
Static Stability [Body Length]
Stability of averaged step (averaged trajectories & phases)
Average stability of original steps
Stability original steps
Time
Bii
Averaged Trajectories
Static Stability [Body Length]
Stability of averaged step (averaged trajectories & phases)
Average stability of original phases with averaged trajectories
Stability of original phases with averaged trajectories
Time
Biii
Averaged Phases
Stability of averaged step (averaged trajectories & phases)
Average stability of original trajectories with averaged phases
Stability of original trajectories with averaged phases
Time
Ci
2 - 4 BL/s
Cii
5 - 7 BL/s
Ciii
8 - 10 BL/s
Average Static Stability [Body Length]
Original Steps
Averaged Trajectories
Averaged Phases
Fully Averaged Steps

Figure 6: Static stability demands increased accuracy at higher walking speeds. (A) Example of a static stability polygon (black connected dots) formed by all legs in stance phase at the presented instant in time. Green connected dots indicate the estimated center of mass and the closest point on the stability polygon. (B) Static stability graphs over the course of individual and partially or completely averaged SCs. Grey lines in (Bi) display the stability values for the steps as they were originally performed by the animal. In (Bii) and (Biii) show stability values for the same steps as on (Bi), but with averaged trajectories or phases, respectively. Blue lines show the stability of the corresponding average step. Green lines indicate the average stability of all individual steps (grey lines, partially averaged or original). (C) Summarized results for all flies with at least 30 steps in one of the three speed ranges. Each dot represents the average stability of 30 steps of one fly. Boxplots show the 5, 25, 50, 75, and 95 percentiles of the distributions. Red dots correspond to steps as they were originally performed by the animal, green dots to original phases with averaged trajectories, blue dots to averagde phases with original trajectories, and white dots to averaged phases with averaged trajectories. (Ci) 2 to 4 BL s^{-1}, (Cii) 5 to 7 BL s^{-1}, (Ciii) 8 to 10 BL s^{-1}.

Original steps and averaged steps both became generally less stable the faster flies walk. This can readily be explained by the decreasing number of legs which are simultaneously in stance phase with increasing walking speed. The decreased stability at high walking speeds means that disturbances of or imperfections in the motor output more quickly result in stumbling or falling. As shown in the previous section, flies compensate for that by decreasing the variability of stepping with increasing speeds. The question arising therefore is, whether flies walk more variable at low speeds just because the longer stance phases allow for it, or due to some yet unknown factor.

3.3.5 Walking speed, TCS, and coherence are all correlated with variability

To analyze and compare how walking speed, TCS, and coherence relate to spatial and temporal variability, we performed a similar analysis as presented in Figures 4 and 5. Here we used tripod cycles (see Materials and Methods) instead of SCs to be able to calculate TCS and coherence, which is why Figures 7A and 8A differ slightly from Figures 4B and 5B. Apart from this difference we performed exactly the same analysis: we defined bins for all three parameters, namely walking speed, TCS, and coherence. We took data of each fly with 30 or more steps in a parameter bin and calculated the spatial and temporal variability. Finally we

normalized the results by dividing all variability values by the variability of the bin which contained data of the largest number of flies. The resulting normalized variability values are plotted on Figures 7 and 8.

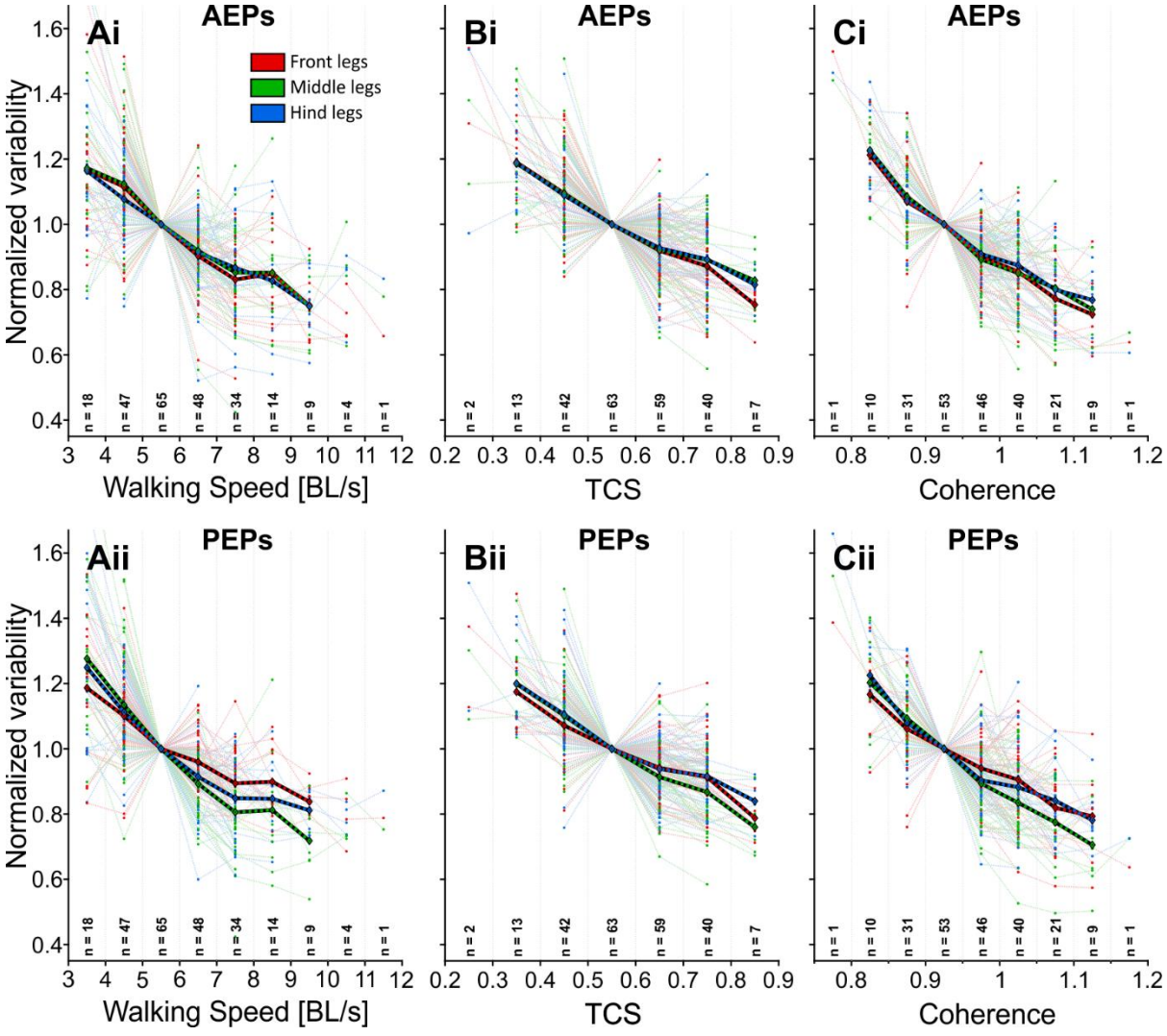


Figure 7: Variability distribution of AEPs and PEPs for different walking speeds, TCS, and coherence values. Each dot represents the normalized variability of 30+ steps of one fly. The variability distributions of the two body sides were averaged (30+ steps on each side). Faint dashed lines connect dots of individual flies, thick black lines connect the average values of all bins. n-values display the number of flies in each speed bin. Average values were calculated for all bins with 5 or more flies. (A) Variability of AEPs and PEPs plotted against walking speed, (B) against TCS, and (C) against coherence.

Figure 7 shows that walking speed, TCS, and coherence are all negatively correlated with spatial variability measured as the 2D variance of AEP and PEP distributions. Overall, the correlations seem to be linear and highly similar for all three parameters, as well as for AEPs and PEPs.

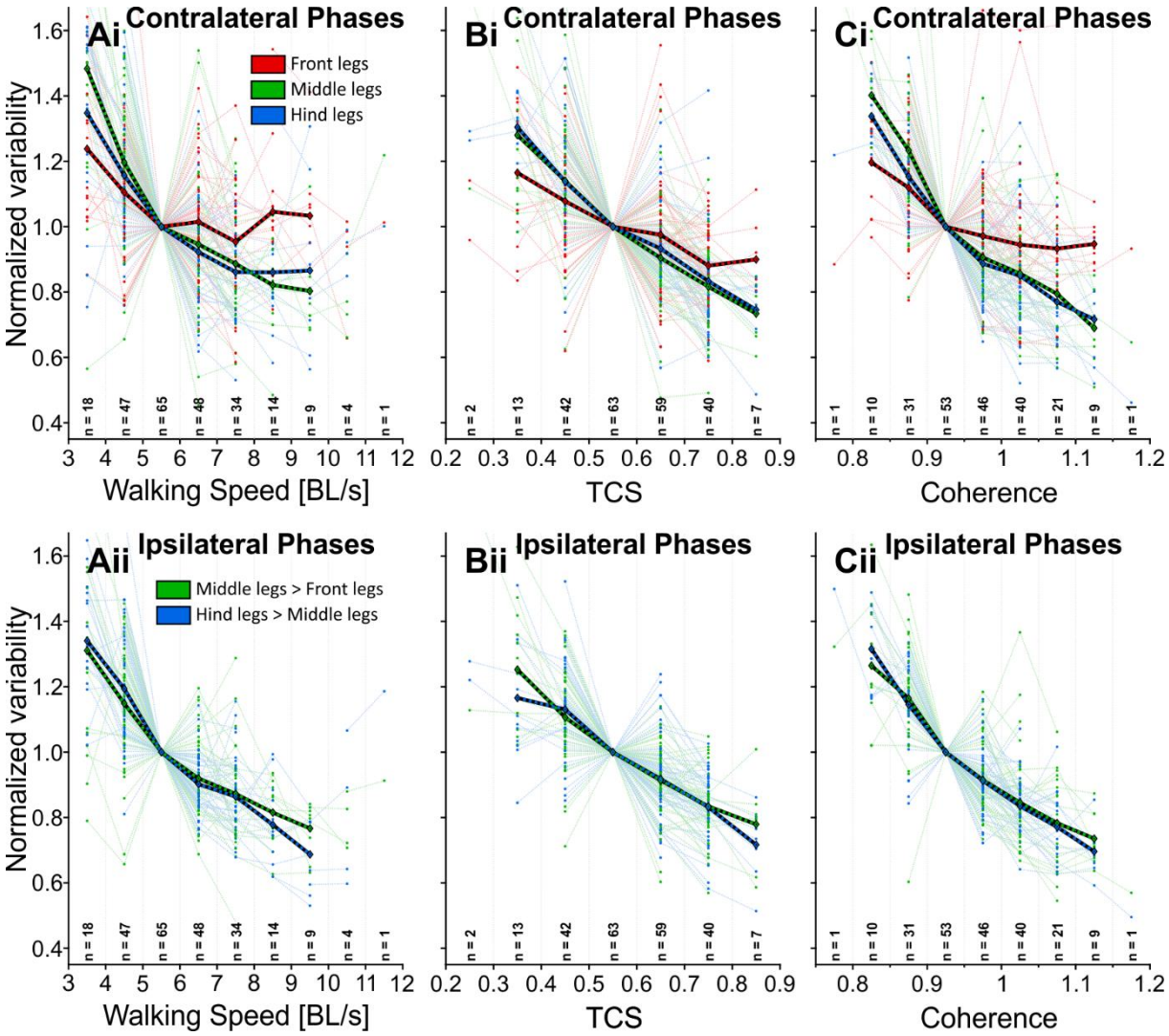


Figure 8: Variability distribution of contra- and ipsilateral phases for different walking speeds, TCS, and coherence values. Each dot represents the normalized variability of 30+ steps of one fly. The variability distributions of ipsilateral pairs of both body sides have been averaged (30+ steps for each pair). Faint dashed lines connect dots of individual flies, thick black lines connect the average values of all bins. ***n***-values display the number of flies in each speed bin. Average values were calculated for all bins with 5 or more flies. (A) Variability of contra- and ipsilateral phases plotted against walking speed, (B) against TCS, and (C) against coherence.

The variability of contralateral and ipsilateral phases is also negatively correlated with all three parameters. Contralateral phases of front legs show the weakest correlation in all cases. TCS

shows an almost linear relationship with the variability of ipsilateral and contralateral phases. Walking speed and coherence on the other hand show a stronger correlation at lower and a slightly weaker correlation at the higher part of the parameter ranges. Apart from these small differences, correlations with temporal variability seem very similar for all three parameters, as we already observed for the spatial variability depicted in Figure 7. This high similarity between all three parameters demonstrates, that they all, despite their different definitions, capture important aspects of walking behavior, which are connected to motor variability.

3.3.6 Proxies for variability yield similar results for TCS and coherence

We used Z-scores describing the deviation from the average of the whole distribution as proxies for variability, which allowed us to directly assign a variability value to every step in the data set. We plotted these Z-scores separately for AEPs and PEPs (Figure 9) as well as for contralateral and ipsilateral phases (Figure 10) against walking speed, TCS, and coherence. Because the data was too dense to observe individual data points, we calculated probability density functions, which we visualized in form of heat maps in Figures 9 and 10, together with contour lines for visual guidance and regression lines for quantification purposes.

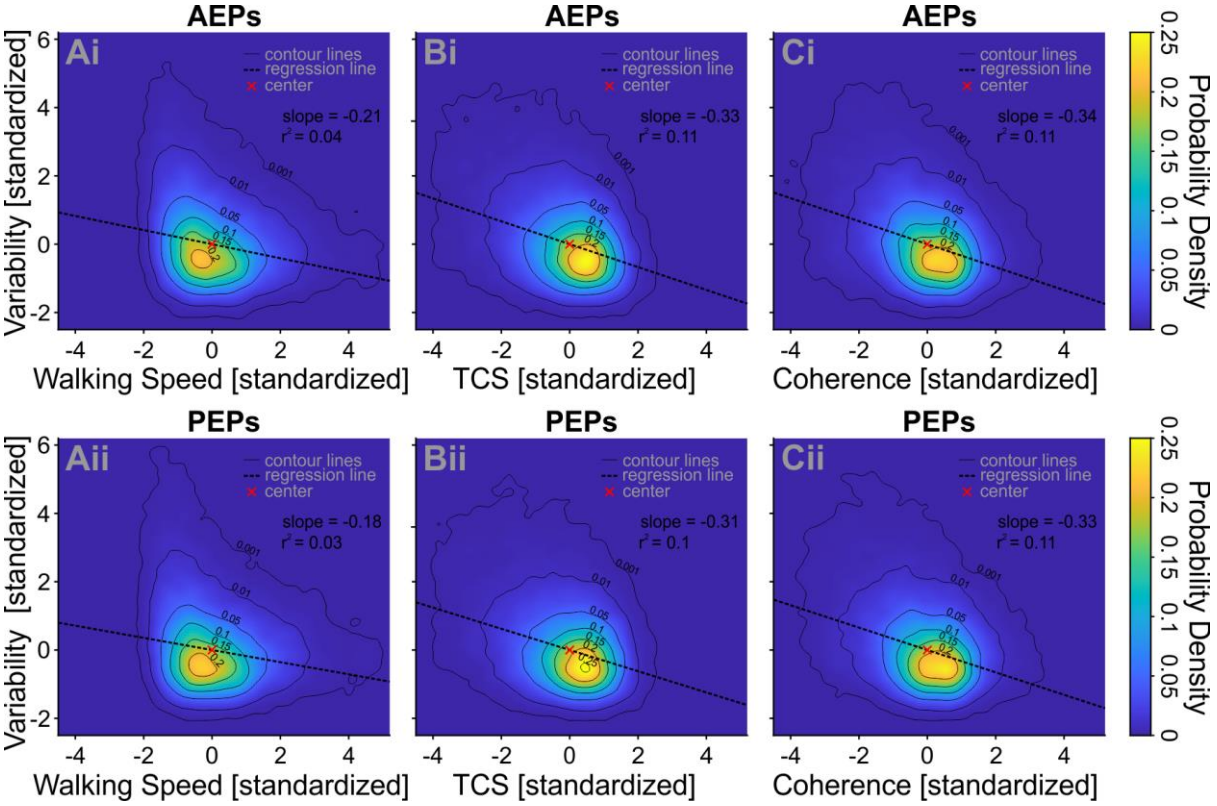


Figure 9: Heat maps displaying the probability density functions of Z-Scores against walking speed, TCS, and coherence. Z-Scores here describe the standardized deviation from average values of AEP and PEP distributions (see Figure 1). Standardized means that the average value was subtracted and everything was divided by the standard deviation. The color bar to the right indicates the probability density and the contour lines help to identify the regions for certain density values. Slope and r^2 values belong to the dashed regression line. The center of all data points is marked with a red x. (A) Z-Scores of AEPs and PEPs plotted against walking speed, (B) against TCS, and (C) against coherence.

Figure 9 shows a different shape of the distribution of spatial variability Z-scores and walking speed compared to TCS and coherence. While the former approximates a triangle shape with the highest data concentration on the lower left side of the center, TCS and coherence both show a more elliptical distribution with the highest data concentration shifted towards the lower right side of the center. Both the slope and r^2-values indicate a stronger correlation of TCS and coherence compared to walking speed.

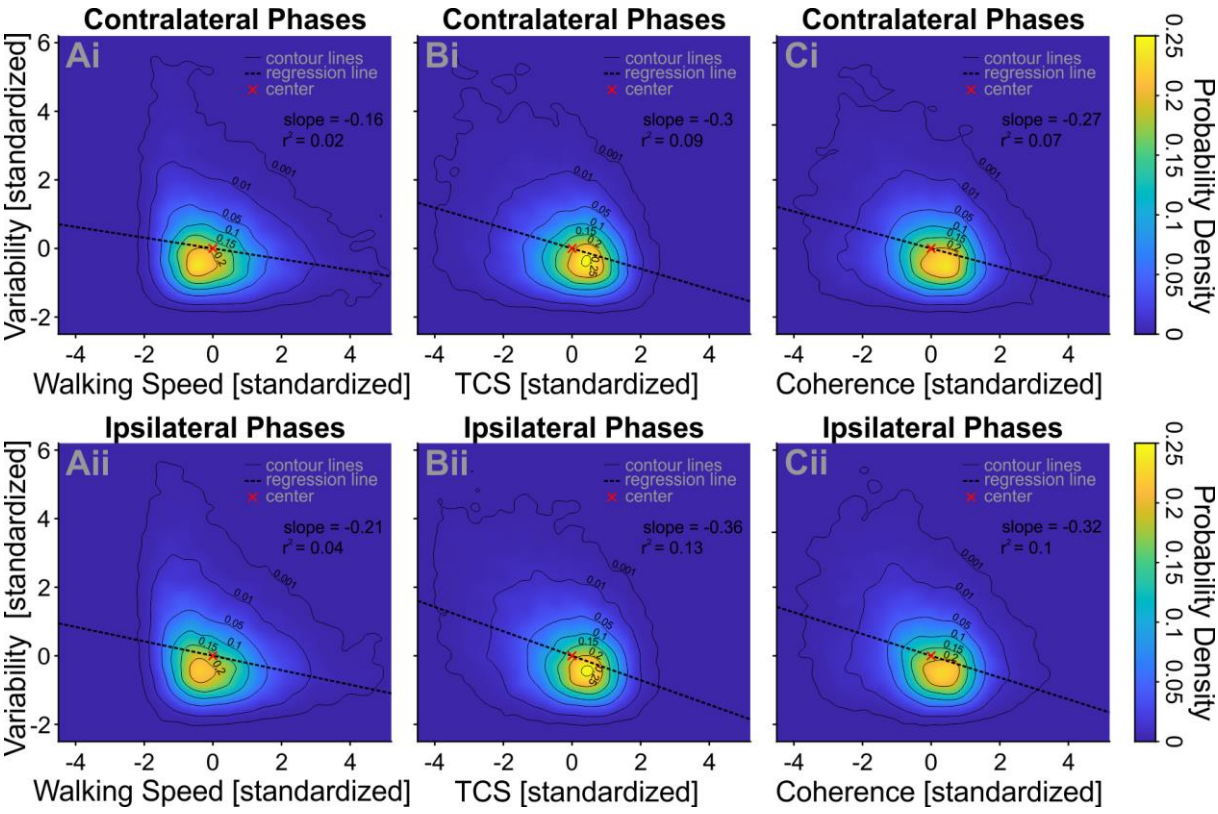


Figure 10: Heat maps displaying the probability density functions of Z-Scores against walking speed, TCS, and coherence. Z-Scores here describe the standardized deviation from contra- and ipsilateral average phases. Standardized means that the average value was subtracted and everything was divided by the standard deviation. The color bar to the right indicates the probability density and the contour lines help to identify the regions for certain density values. Slope and r^2 values belong to the dashed regression line. The center of all data points is marked with a red x. (A) Z-Scores of contra- and ipsilateral phases plotted against walking speed, (B) against TCS, and (C) against coherence.

Results for temporal variability Z-scores in Figure 10 are highly similar to those for the spatial ones in Figure 9. Figure 10B shows an even higher concentration of data for TCS than for coherence in panel C, which might contribute to the slightly larger r^2-values for TCS. Additionally, the r^2-values are larger for all parameters for ipsilateral than for contralateral phases variability.

Overall the analysis using variability proxies revealed a weak correlation with all three parameters. It was the weakest for walking speed, which also showed a different shape of correlation, as the contour lines for both spatial and temporal variability aspects demonstrate.

The high similarity for TCS and coherence on the other hand indicates that they measure similar aspects of walking behavior.

3.3.7 TCS is a better predictor for variability than walking speed

We used tripod coordination strength (TCS) to quantify the tripodness of stepping sequences and compared this proxy for interleg coordination against walking speed regarding the relative spatial and temporal variability. TCS is a simple measure which indicates how close a sequence of walking behavior is to an ideal tripod coordination. It is positively correlated with walking speed, while TCS values for certain speed values can vary substantially, especially for intermediate speeds (Figure 11A). This variability of TCS values at intermediate speeds allowed us to separate the two factors in the following way: first we restricted the speed range to 5.05 to 5.8 BL s^{-1} (Figure 11A), calculated the variability of steps with high TCS values and divided it by the variability of the low TCS steps. In this analysis, values below 1 are associated with flies with reduced variability for high TCS values and above 1 for those with increased variability. We then restricted the TCS range to 0.64 to 0.70 TCS, keeping the largest part of the walking speed range (Figure 11B) and proceeded with the same steps as for the restricted walking speed range.

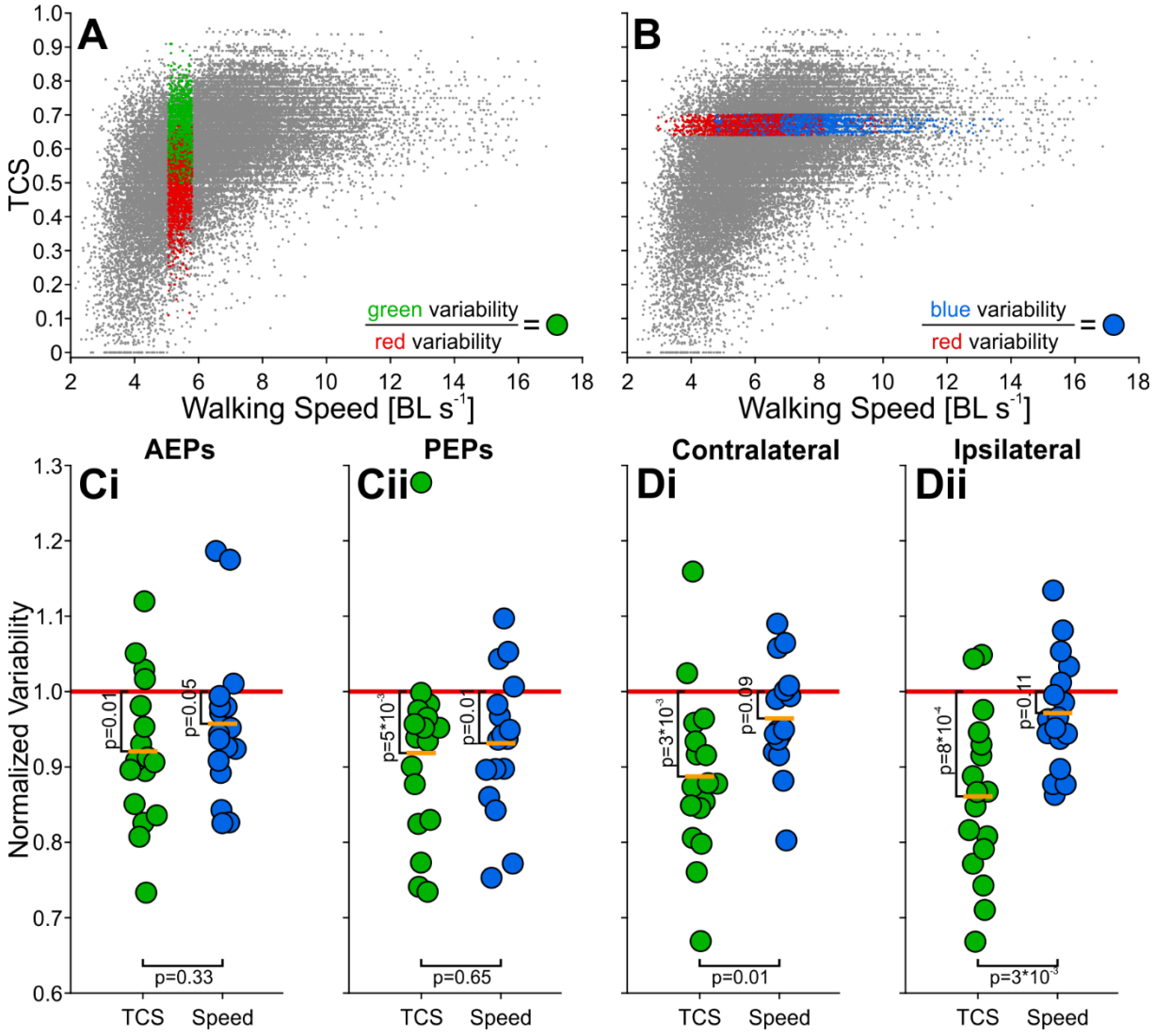


Figure 11: TCS is a better predictor for variability than walking speed. (A) Walking speed range limited to 5.05 to 5.8 BL s^{-1}. Steps belonging to the lower 50 % of TCS values are shown in red, upper 50 % in green. Every fly has its own median TCS value, which is why the border between lower and upper halves of TCS values are not separated at one identical TCS value for all flies. (B) TCS range limited to 0.64 to 0.7 TCS. Steps with low walking speeds are shown in red, high walking speeds in blue. As for TCS, separation of lower and upper halves is individual for each fly. (C) Relative spatial and temporal variability for the subsets of data shown in (A) and (B). For each fly, the variability of high TCS (green) and high walking speed (blue) steps was calculated and divided by the variability of the respective low TCS and low walking speed group. This way, normalized variability values below one demonstrate, that the fly walked less variable with higher TCS values or walking speeds, respectively. (Ci) Normalized variability of AEPs. (Cii) Normalized variability of PEPs. (Di) Normalized variability of contralateral phase relations. (Dii) Normalized variability of ipsilateral phase relations. Each dot represents results for one fly, mean results are indicated with an orange bar. (D) Variability of high TCS or walking speed contralateral and ipsilateral phase relations divided by their low

scoring counterparts. Each dot represents results for one fly, mean results are indicated with an orange bar.

As the results in Figure 11 C and D show, high-to-low TCS comparison yielded significantly reduced variability values for high TCS steps both spatially and temporally, while the high-to-low walking speed comparison did only result in a significant difference for PEP variability. The effect size along the TCS dimension is much larger for temporal (phase relations in Fig. 11D) than for spatial variability (APEs and PEPs in Fig. 11C), while it was the other way around for the coherence dimension. Overall, this analysis demonstrates that interleg coordination (here, approximated with TCS) is negatively correlated with variability, while walking speed alone only shows significant effects for spatial variability measures. This result indicates that walking speed does not have a strong direct impact on motor variability, but affects it due to its correlation with interleg coordination.

3.3.8 Coherence is a better predictor for variability than walking speed and TCS

To directly compare walking speed, TCS, and coherence with regard to their capacity for the prediction of spatial and temporal variability we extended the analysis presented in Figure 11 by comparing walking speed and TCS to our coherence measure (Figure 12 and 13). To keep all three analyses comparable, we used the same limits for the narrowed ranges of the three parameters.

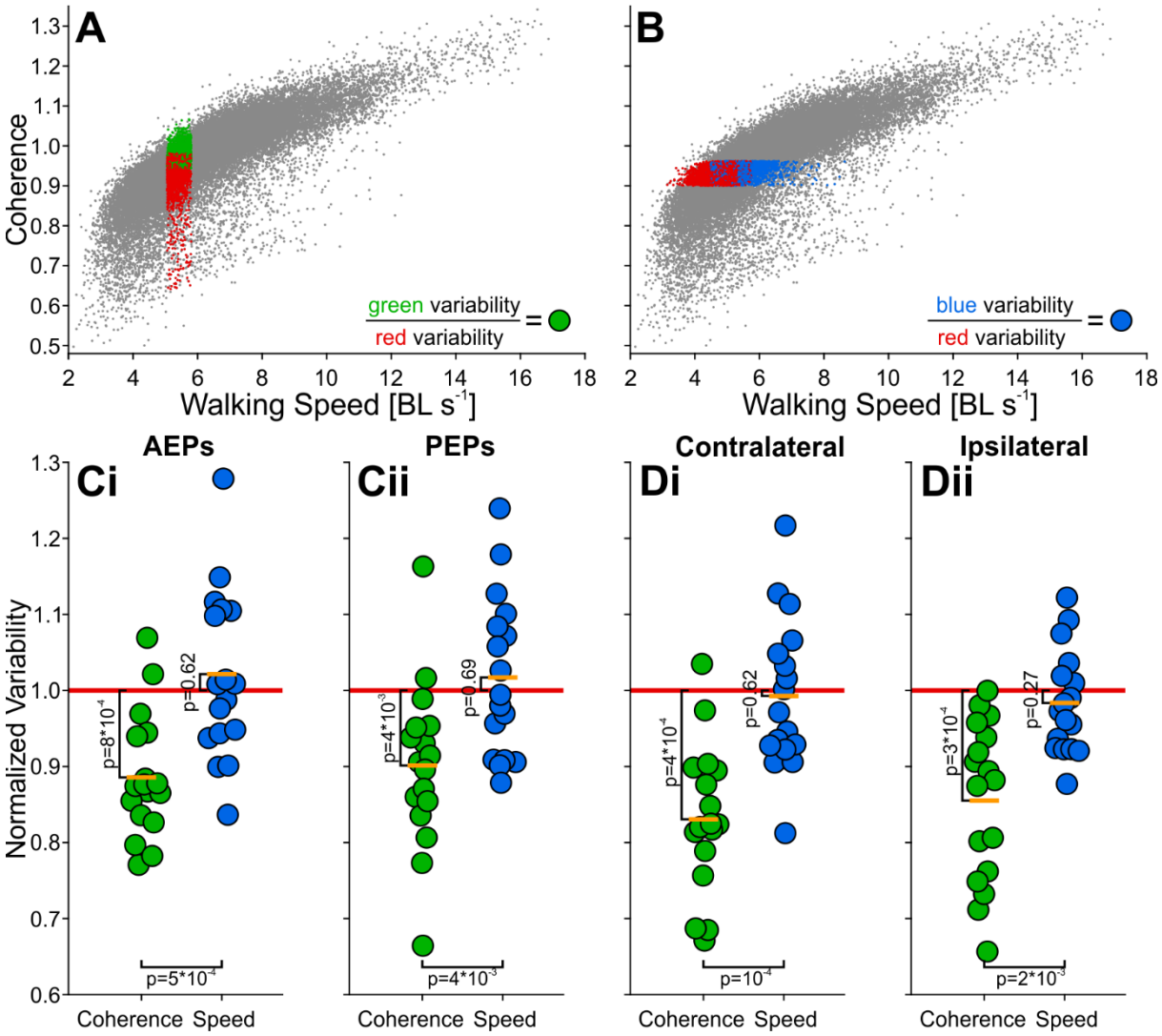


Figure 12: Coherence is a better predictor for variability than walking speed. (A) Walking speed range limited to 5.05 to 5.8 BL s^{-1}. Steps belonging to the lower 50 % of Coherence values are shown in red, upper 50 % in green. Every fly has its own median Coherence value, which is why the border between lower and upper halves of TCS values are not separated at one identical Coherence value for all flies. (B) Coherence range limited to 0.9 to 0.962 Coherence. Steps with low walking speeds are shown in red, high walking speeds in blue. As for Coherence, separation of lower and upper halves is individual for each fly. (C) Relative spatial and temporal variability for the subsets of data shown in (A) and (B). For each fly, the variability of high Coherence (green) and high walking speed (blue) steps was calculated and divided by the variability of the respective low Coherence and low walking speed group. This way, normalized variability values below one demonstrate, that the fly walked less variable with higher Coherence values or walking speeds, respectively. (Ci) Normalized variability of AEPs. (Cii) Normalized variability of PEPs. (Di) Normalized variability of contralateral phase relations. (Dii) Normalized variability of ipsilateral phase relations. Each dot represents results for one fly, mean results are indicated with an orange bar. (D) Variability of high Coherence or

walking speed contralateral and ipsilateral phase relations divided by their low scoring counterparts. Each dot represents results for one fly, mean results are indicated with an orange bar.

Figure 12 C and D show significant effects along the coherence dimension with the narrowed range of walking speeds for AEPs and PEPs as well as contralateral and ipsilateral phases. In contrast, we could not find any significant difference along the walking speed dimension with the narrowed range of coherence values. This finding suggests that our coherence measure can predict variability with even higher accuracy than TCS, which, in turn, was already shown to be more predictive than walking speed in Figure 7.

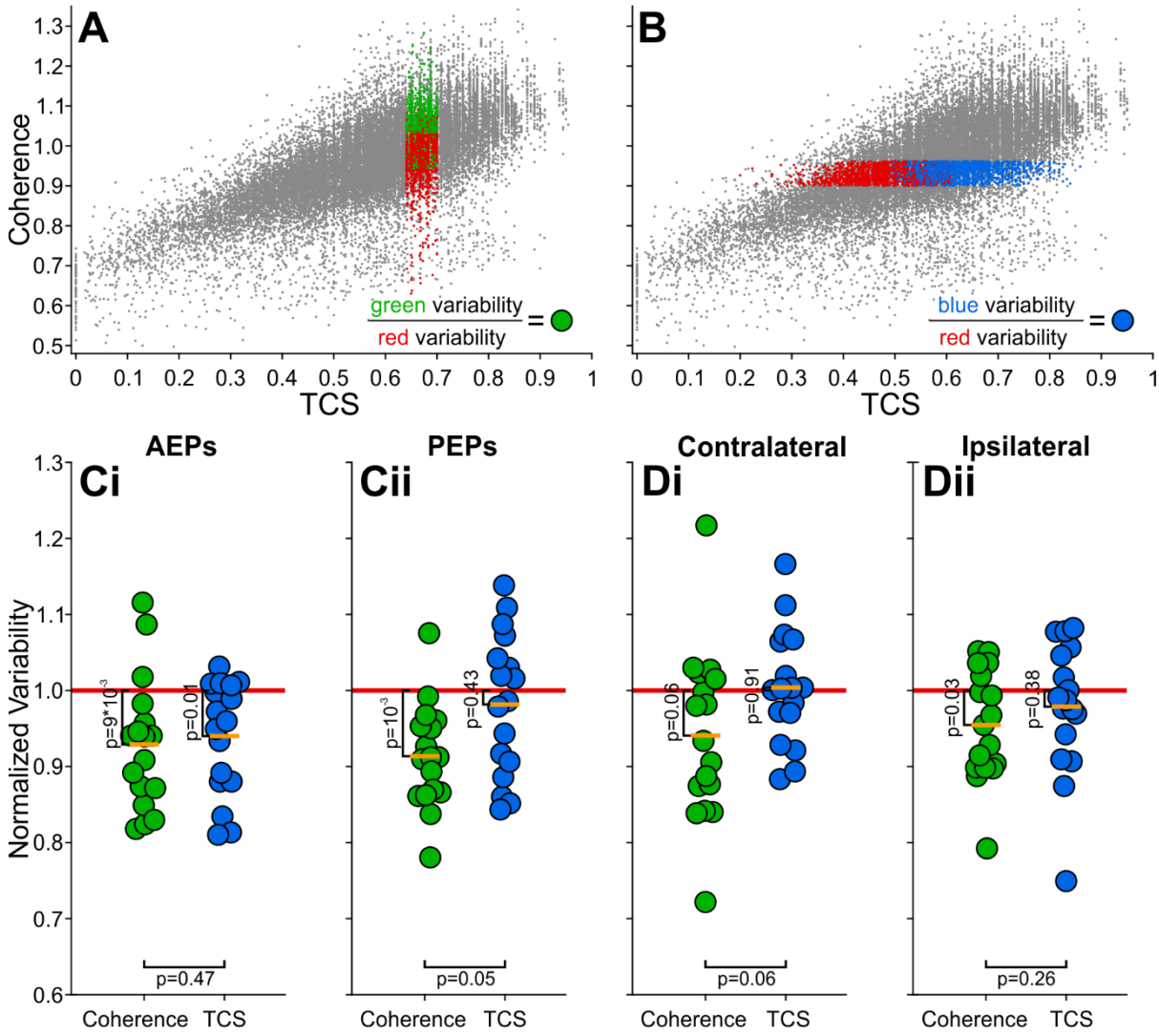


Figure 13: Coherence is a better predictor for variability than TCS. (A) TCS range limited to 0.64 to 0.7 BL s^{-1}. Steps belonging to the lower 50 % of Coherence values are shown in red, upper 50 % in green. Every fly has its own median Coherence value, which is why the border

between lower and upper halves of TCS values are not separated at one identical Coherence value for all flies. (B) Coherence range limited to 0.9 to 0.962 Coherence. Steps with low TCS are shown in red, high TCS in blue. As for Coherence, separation of lower and upper halves is individual for each fly. (C) Relative spatial and temporal variability for the subsets of data shown in (A) and (B). For each fly, the variability of high Coherence (green) and high TCS (blue) steps was calculated and divided by the variability of the respective low Coherence and low TCS group. This way, normalized variability values below one demonstrate, that the fly walked less variable with higher Coherence values or TCS, respectively. (Ci) Normalized variability of AEPs. (Cii) Normalized variability of PEPs. (Di) Normalized variability of contralateral phase relations. (Dii) Normalized variability of ipsilateral phase relations. Each dot represents results for one fly, mean results are indicated with an orange bar. (D) Variability of high Coherence or TCS contralateral and ipsilateral phase relations divided by their low scoring counterparts. Each dot represents results for one fly, mean results are indicated with an orange bar.

The overall difference between the TCS and coherence dimension depicted in Figure 13 is small. Especially for AEPs and ipsilateral phases the relative variability values for TCS and coherence are highly similar. For PEPs and contralateral phases on the other hand we observe a reasonably large decrease along the coherence dimension, while values along the TCS dimension are neutral on average. However, the small effect size together with the mixed results for both spatial and temporal aspects of variability are first and foremost another indication of the general similarity of TCS and coherence. Still, taken together with the indirect comparison over walking speed in Figures 7 and 12, our findings indicate that the coherence measure is an even better predictor of variability than TCS.

3.4 Discussion

In this study, we analyzed an extensive data set of straight walking fruit flies along the dimension of walking speed and focused on systematic changes in stepping variability and how intersegmental influences relate to them. We used this data set to analyze in detail how common parameters of walking behavior, such as step durations, duty cycles, or phase relations change systematically with walking speed (Figure 2). To quantify how potential intersegmental influences affect the timing and coordination of leg movements we established a measure of coherence of these influences. This measure was used on artificially created footfall patterns (Figure 3 D and E) and proofed to be proportional to the systematic changes of interleg coordination presented in Figure 2. For both spatial and temporal variability we found negative correlations with walking speeds (Figure 4 B and Figure 5 B). Systematic analysis of static stability revealed that mean steps are more stable than individual steps, even if only the temporal or the spatial aspect were averaged. Additionally, static stability markedly decreases with increasing walking speeds, possibly suggesting the necessity for increased accuracy of individual leg stepping and interleg coordination (Figure 6 C). Correlating spatial and temporal variability with walking speed, TCS, and coherence in Figures 7 and 8 showed, that all three parameters are negatively and mostly linearly correlated with all measured aspects of variability. Usage of a single step-based proxy for variability in Figures 9 and 10 revealed that walking speed differs from TCS and coherence with regard to its data concentration and overall shape in the heat map depiction. Limiting the range of walking speed and then TCS while allowing almost full range variance of the other parameter (Figure 11 A and B) revealed that TCS, in direct comparison, can more accurately predict the stepping variability than walking speed (Figure 11 C and D). Directly testing the predictive capabilities of coherence against walking speed in Figure 12 showed that coherence can predict both spatial and temporal variability more accurately than walking speed, even more consistent and with a larger effect size than TCS in the homologous analysis in Figure 11. Direct comparison of TCS and coherence in Figure 13 again indicates that coherence is the best predictor for variability analyzed in this study.

Walking speed is arguably the most important high level parameter affecting all over aspects of walking behavior. We calculated how nine basic kinematic parameters change with walking speeds in our extensive data set and plotted the results in Figure 2. In doing so, we were able to reproduce most findings of previous studies (DeAngelis et al., 2019; Mendes et al., 2013; Strauß and Heisenberg, 1990; Szczecinski et al., 2018; Wosnitza et al., 2013). One difference was that swing duration for our data set decreased from ~40 ms to ~25 ms, while DeAngelis et al. show an increase from ~30 to ~40 ms for a similar range of walking speeds. Mendes et al. and Sczcecinski et al. on the other hand report a small decrease in swing duration which was in a similar range as what we observed. However, we used the mean results from Figure 2 to construct artificial footfall patterns (Fig. 3D) which did not contain any noise, but strictly followed the statistics of recorded walking behavior. These artificial footfall patterns were used to evaluate our coherence measure, which we established to quantify how well attuned the incoming coordinating signals from different adjacent legs were. For the estimation of these intersegmental signals we assumed that similar rules as described by Cruse in 1990 for the stick insect (Fig. 3A) are also valid in the fruit fly. We found a consistent positive correlation between our coherence measure and increasing walking speeds, both for recorded behavior and artificial data (Fig. 3E). Testing the coherence measure on artificial data without step-to-step variability thereby proved that we did not just reproduce the inverse of the decrease in variability which can be observed with increasing walking speeds.

Many aspects of changes in interleg coordination related to different walking speeds are widely studied (DeAngelis et al., 2019; Mendes et al., 2013; Szczecinski et al., 2018; Wosnitza et al., 2013); however, one aspect which has not yet received a lot of attention is the step-to-step variability and how it changes with walking speed. We were able to successfully replicate the finding that the temporal variability of leg movements decreases with increasing walking speeds (Figure 5) (DeAngelis et al., 2019; Mendes et al., 2013). Additionally, we could for the first time quantify the decrease of spatial variability with increasing walking speeds for a large data set (Figure 4). The finding of decreasing variability for higher walking speeds seems counterintuitive at first, since larger activation amplitudes of muscles are usually associated

with proportionally increased motor noise (Harris and Wolpert, 1998; Schmidt et al., 1979; Todorov, 2004), or in other words, variability. A common hypothesis to explain this discrepancy in insects is that sensory feedback plays a diminishing role with increasing walking speeds (Berendes et al., 2016; Isakov et al., 2016), or from a different perspective, that inter-limb coupling becomes stronger at high speeds (Aminzare et al., 2018; Aminzare and Holmes, 2019). The first interpretation is based on the idea that sensory feedback produces variability due to situation-dependent adaptions of leg movements, while the second interpretation congruently says that fast walking is mostly centrally driven and open-loop aspects dominate. Following these interpretations that sensory feedback is the cause of high variability at slow walking, it should be expected that this variability is to some degree systematic and serves a purpose.

One parameter that could be optimized through sensory feedback is static stability. Small animals like flies are more affected by elastic and viscous forces than by inertia; hence static stability is crucial, because it directly impacts the fly's ability to maintain balance and avoid falling while walking (Szczecinski et al., 2018). We therefore analyzed the static stability of individual steps and compared it to averaged steps calculated for the same animal. This way we could test whether the variability of individual steps increases the static stability compared to averaged steps, in which the situation dependent adaptions of leg movements are smoothed out. We found that averaged steps in all cases were more stable compared to the stability of individual steps, even if only the temporal or the spatial aspect were averaged (Figure 6C). Individual steps being less stable than averaged ones indicates that either sensory feedback is not mainly used to improve static stability, or other factors are producing the observed variability. We also observed that static stability decreases significantly with increasing walking speeds, reducing the margin for variability, as deviations from the optimal configuration more rapidly lead to instability when static stability is already diminished. Conversely, at lower walking speeds, longer stance phases allow for greater variability in stepping. (Schilling et al., 2013). It is possible that flies produce more variable motor output during slow walking simply

because precise control is not required to maintain static stability. However, it seems more likely that the changes in variability are driven by the neural control system itself.

A fly's nervous system produces a continuum of interleg coordination patterns converging to the so-called tripod coordination at high walking speeds (DeAngelis et al., 2019; Wosnitza et al., 2013). TCS is a simple measure of how tripod a certain pattern is. It produces reliably high values for tripod-like patterns and gradually lower values for patterns further down the continuum (Wosnitza et al., 2013). It is impossible to properly assess the temporal interaction of six legs with a single number, but we found TCS to be a well-suited proxy to analyze how variability changes with interleg coordination and whether it is significantly different from walking speed. In addition to walking speed and TCS we tested our coherence measure against variability. Similar to TCS, the coherence measure is also based on the timing of lift-off and touchdown events, but it is based on assumed coordinative signaling between directly adjacent legs. In contrast, TCS is based on the overlap of swing phases in tripod groups and thus combines legs which are not directly adjacent. However, all three parameters show a clear negative and mostly linear correlation with both spatial and temporal variability, as depicted in Figures 7 and 8. This high similarity shows that walking speed, TCS, and coherence describe important aspects of walking behavior which are connected to each other and to the causes of spatial and temporal variability. Direct comparison of all three parameters with regard to the strength of correlation is difficult, because it was impossible to use the same data subsets for their calculation. However, to allow fair comparison, we used the standardized deviation from a fly's mean AEP, PEP, contra-, or ipsilateral phase relation as a proxy for variability. This procedure enabled us to assign spatial and temporal variability values to each tripod cycle, which were then correlated with walking speed, TCS, and the coherence of the same tripod cycle. The results in Figures 9 and 10 replicate the finding of negative correlations for all three parameters, but additionally reveal that walking speed is distinct from TCS and coherence, as it shows a different shape of contour lines and a shift of the highest data concentration to the left of the mean. TCS and coherence on the other hand show larger r^2-values than walking speed and are again highly similar to each other. This similarity might

come from the fact that both measures are derived from timings of lift off and touch down events, although the coherence measure only considers directly adjacent legs and TCS in contrast is based on swing phases within tripod groups, which do not contain directly adjacent legs. However, walking speed, TCS, and coherence on average all show shape wise similar correlations with variability (Figures 7 and 8), while direct comparison using a proxy for variability indicates that TCS and coherence are stronger correlated with variability than walking speed (Figures 9 and 10). An even more direct comparison of all three parameters is needed to make more definitive statements.

The central problem of comparing walking speed, TCS, and coherence with regard to variability is that all three parameters are correlated with each other. Our idea to solve this issue was to separate the factors from each other as much as possible. We made use of the fact that we observe a range of different TCS values for all walking speeds, particularly for intermediate speeds (Figure 11A). This allowed us to separate the two factors by restricting the variability of one at a time while allowing almost the full range of the other. Comparison of the relative variability of TCS and walking speed in Figure 11 C and D revealed that TCS, or the underlying change in coordination patterns, is better correlated with variability than walking speed. TCS captures highly significant differences in relative variability for both contra- and ipsilateral phase relations while walking speed shows a much smaller effect (Fig. 11D); hence walking speed seems to more indirectly affect temporal variability due to its effect on interleg coordination. For spatial variability the effect is much smaller, but also shows a slight advantage of TCS over walking speed (Fig. 11C). However, these results suggest that the decrease in variability with increasing walking speeds arises from systematic changes in phase relationships between adjacent legs along the continuum of interleg coordination patterns.

The coordination rules proposed by Cruse et al. provide a well-established framework to explain how insect legs are controlled in a decentralized manner (Cruse, 1990; Cruse et al., 2007). While the existence and importance of central pattern generators (CPGs) for insect locomotion is extensively documented (Bässler and Wegner, 1983; Bidaye et al., 2018; Pearson, 1972; Wilson, 1961; MacKay-Lyons, 2002; Mantziaris et al., 2020), it remains largely

unclear how their activity is aligned with each other. Sensory feedback and/or state-dependent signaling are the most likely candidates to fill the gap. The individual leg CPGs can be understood as oscillators and the signals coming from sensory feedback and CPGs of contra- and ipsilaterally adjacent legs provide coupling (Berendes et al., 2016; Cruse, 1990). Entrainment of any oscillator is the strongest when all incoming influences are coherent and do not contradict each other (Pikovsky et al., 2001). We applied all temporal influences as described by Holk Cruse in 1990 on exemplary footfall patterns and observed, that these assumed signals indeed become more coherent and less contradicting the faster a fly walks (Fig. 3B). This observation is well captured in our coherence measure, as the examples in Figure 3C and the coherence values of all recorded walking bouts in Figure 3E show. Additionally, the coherence measure is shown to yield similar results for artificial footfall patterns (Fig. 3D) and recorded data (Fig. 3E). Coherence values of artificial footfall patterns are systematically higher than most coherence values of recorded walking bouts, probably because the unsystematic noise in the recorded data negatively affects the coherence of intersegmental influences, while the artificial footfall patterns are based on averaged values and therefore do not contain noise. Direct comparison of our coherence measure and walking speed in Figure 12 revealed highly significant differences between them. The coherence measure on average shows a clear effect for all spatial and temporal parameters, while walking speed completely fails to detect a significant correlation for any of the four parameters. This clear result suggests that our coherence measure, and even more than TCS, captures most of the systematic variability, while walking speed seems to only indirectly affect variability due to its correlation with the continuum of interleg coordination patterns. Finally, we addressed the arising question if our coherence measure can more accurately predict variability than TCS (Figure 13). TCS only showed a clear effect on the variability on AEPs, while our coherence measure additionally detected a variability decrease for PEPs and showed slightly larger effect sizes for contralateral and ipsilateral phase relations. Figures 7 and 8 showed that all three parameters are linearly correlated with variability, hence similar differences should be expected for the ranges of high and low scoring groups shown in panels A and B of Figures

11 to 13. We currently have no hypothesis for the different effects of TCS on AEP and PEP variability in Figure 13C, or an explanation why the difference between TCS and walking speed in Figure 11 C and D are much more pronounced for temporal aspects than for spatial once. However, the overall consistency of the findings that TCS and coherence in direct comparison were superior to walking speed in predicting variability and the fact that coherence more reliably detected variability than TCS clearly hint towards a connection of intersegmental influences and variability.

4 Chapter 4: An advanced setup for studying freely-walking *Drosophila*

4.1 Introduction

Investigating the walking behavior of freely-moving *Drosophila* has the potential to provide key insights into the neural and biomechanical mechanisms underlying legged locomotion. However, capturing this behavior with high spatial and temporal resolution presents significant technical challenges, especially when considering the rapid leg movements and fine-scale kinematic changes characteristic of walking behavior in flies. For instance, the typical swing phase duration of *Drosophila melanogaster* is approximately 30 ms (Chapter 3), necessitating the use of frame rates around 200 Hz to capture a minimum of 5 to 6 frames per swing phase. This is critical for accurately detecting the transition points between the swing and stance phases, as well as the precise lift-off and touch-down positions. Additionally, the small body size of *Drosophila*, approximately 2 mm in length, presents a further challenge. High-resolution recordings of fly behavior require either close proximity of the camera to the subject or the application of significant optical zoom. Achieving a large depth of field is advantageous for obtaining sharp images, yet it poses a tradeoff with the need for short exposure times to minimize motion blur, because optimizing both parameters reduces the amount of light reaching the camera sensor during image acquisition. Previous approaches for studying fly locomotion, such as multi-animal tracking systems (TRex: Walter and Couzin, 2021) or ball-tracking setups (Berendes et al., 2016; Seelig et al., 2010), have their own limitations. TRex for example struggles to balance the need for high frame rates and high resolution over extended periods of live recordings. Ball tracking setups, on the other hand, involve tethering of the fly and letting it walk on an air-suspended ball, which might have unforeseeable effects on the expressed behavior, such as the influence of the ball's curvature on walking kinematics or the absence of visual flow.

Here, I describe an advanced setup that is designed to address these challenges by continuously recording walking behavior in freely moving *Drosophila* at frame rates of up to 200 Hz with high video resolution. The setup allows for long-term behavioral studies without sacrificing detail in the captured data, which is crucial for analyzing the fine but systematic nuances in both leg tip kinematics and interleg coordination patterns (see chapters 2 and 3). The higher resolution of the new camera enabled the capture of the entire arena, while maintaining the relative size of the fly in pixels consistent with the previous setup. This improvement ensures that the entire body of the fly is captured in every video frame, whereas data from the previous setup were sometimes unsuitable for analysis when parts of the fly were outside the frame. Furthermore, this new system incorporates robust data pre-selection methods, enabling researchers to focus on specific parameters of interest, such as walking speed or curvature of walking trajectories, while filtering out irrelevant or unwanted behaviors.

A use case for the new setup presented here is the analysis of trajectory curvature based on three different analytical approaches. The first approach is an integral component of the data pre-selection algorithm applied on the raw data which comes out of the setup. It is used for detection of segments of walking behavior within a defined range of curvature values and is based on smoothing spline functions. The second approach approximates segments of walking behavior with a circle and thereby provides a simple and easy to interpret way for quantifying the curvatures of trajectories. The third approach measures how flies adjust their body orientation in relation to their path, providing an interesting extension to the other two methods which are entirely based on the curvature of walking trajectories. The advantages and limitations of each approach are discussed and compared, resulting in the suggestion to combine the spline function and the body orientation approach for further analysis of curved walking behavior. However, the example of curvature analysis demonstrates how the enhanced capabilities of the setup presented here can be leveraged to analyze specific aspects of *Drosophila* walking behavior.

In summary, this advanced setup offers a comprehensive solution for studying *Drosophila* walking behavior with high spatial and temporal resolution, addressing key limitations of previous tracking systems.

4.2 Materials and Methods

4.2.1 Experimental setup

Except for the camera the exact same hardware components as for parts one and two of this thesis were used. Please go section **2.3.2** for a detailed description. However, the most important aspects of the setup are described here again.

The recording arena consisted of an inverted glass petri dish (diameter: 60 mm) as the walking substrate and a watch glass (diameter: 100 mm) as the lid. This arrangement formed a closed chamber with a curved dome tapered towards the edge of the petri dish, similar to an inverted FlyBowl (Simon and Dickinson, 2010). To record flies walking on the glass substrate the Atlas ATX051S-MC camera (LUCID Vision Labs, Richmond, BC, Canada) was used for this new version of the setup. The major advantages of this new camera are the more than two times higher resolution (as compared with the previous setup) and the 10 Gigabit Ethernet connection which allowed us to continuously grab frames with a resolution of 2000-by-2000 pixels at a recording rate of 200 Hz. The camera was equipped with an object-space telecentric lens (focal length 55 mm, model: Computar TEC-55, CBC America, Cary, North Carolina, USA). The camera was located on the side of the setup and its view was directed at the experiment chamber from below via a surface mirror tilted at an angle of 45° (Figure 1).

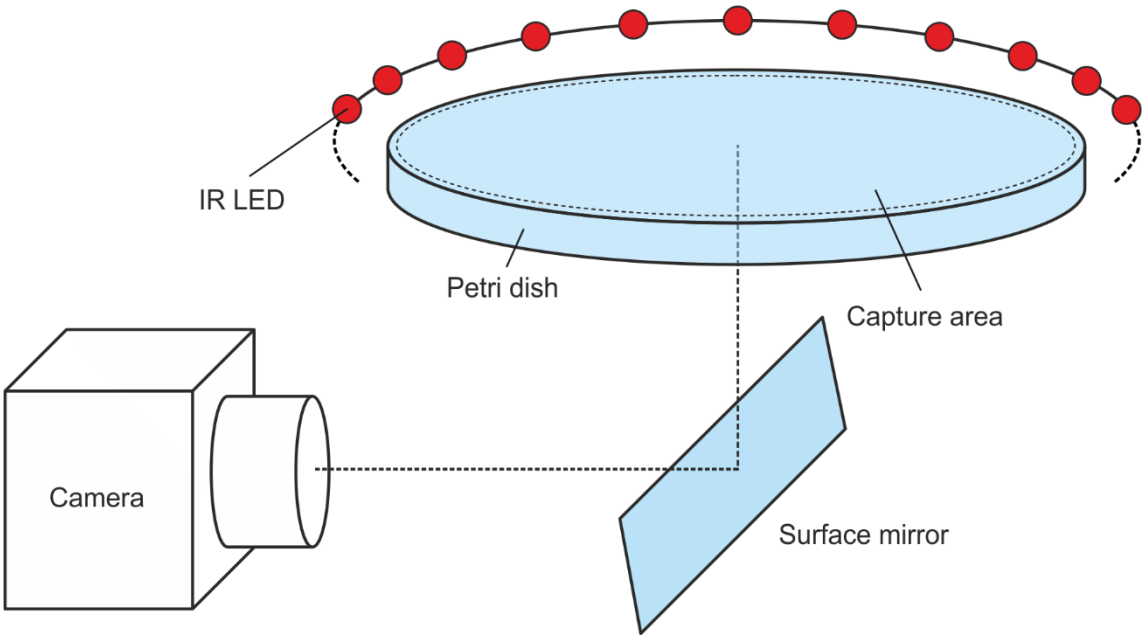


Figure 1: Schematic of the experimental setup. The watch glass covering the petri dish is not depicted.

The scene was illuminated with 60 infrared (IR) LEDs (wavelength: 830 nm, opening angle: 20°) arranged in a concentric ring around the chamber. Video data was acquired at 200 Hz and a shutter time of 400 µs.

4.2.2 Software changes

The new camera enabled capturing of the whole arena, but for continuous recording at 200 Hz a new data processing solution was needed, because full resolution frames could not be stored fast enough. The first step was to implement the software in Python instead of Matlab, which allowed for usage of free source software packages. The first adaption was the use of multiprocessing to more efficiently distribute the load of tasks required for data acquisition on multiple processor kernels. However, the central, yet unsolved, problem was the amount of data which needed to be handled in a very short period of time: a single grayscale frame at 2000-by-2000 pixels resolution is 4 megabytes (MB) of disc space. At a capture rate of 200 Hz this accumulates to 800 MB in a single second. This amount of data per second cannot be written on a conventional hard drive and even if it was possible, the data load would be inefficient to handle.

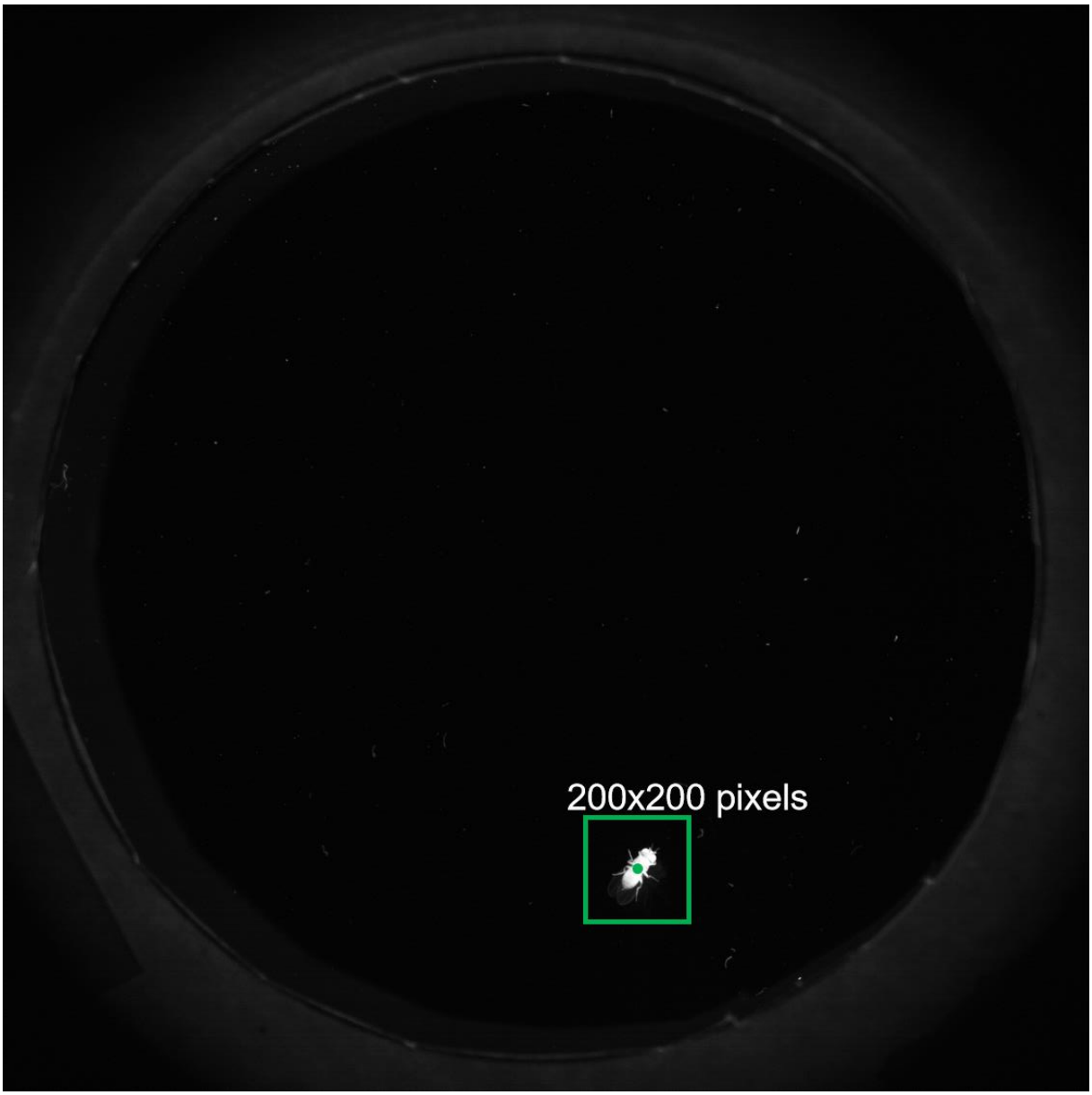


Figure 2: Example shot from the setup at full resolution of 2000-by-2000 pixels (4 MB disc space). The green outlines around the fly measures exactly 200-by-200 pixels (0.04 MB disc space).

The simple solution was to only keep the part of each image which is relevant, namely the part in which the fly was located (Figure 2, green outlines). To reliably isolate the fly from the original video frames, a fast and robust detection algorithm was required. With a frame capture rate of 200 Hz, the interval between consecutive frames is only 5 ms, necessitating the completion of all operations within this time frame to prevent data accumulation in RAM. In the previous setup, a threshold operation followed by a blob detection algorithm was used to determine the fly's position. However, for full-scale frames, this process exceeded the 5 ms limit and was

therefore performed after the recording had concluded. To enable faster, yet reliable real-time tracking of the fly's position during live recordings, a new detection algorithm was developed. This approach involved (1) down sampling the video frames by a factor of 4, (2) converting them into binary images using a brightness threshold, and (3) applying the median function to the white pixel values along the x- and y-axes. This way, the center of the largest object with brightness values above threshold (the fly) could reliably be detected in less than 5 ms and with an accuracy of 4 pixels. Around the position of the fly a 200-by-200 pixels large square was cut out of the video frame. In this snippet, the position of the fly was again determined with the same approach as before but without down sampling to have more exact coordinates and a smoother walking trajectory. The resulting 200-by-200 pixels snippet was written to the hard drive. This way. each frame needed only one percent of the disc space the full frame would have needed, which was even further reduced by low-loss compression algorithms applied with FFmpeg (codec: libx264, compression factor: 14). One hour of recordings using these settings results in a video file of approximately 2 gigabytes (GB), whereas the uncompressed, full-resolution version (2000-by-2000 pixels) of the same video would require approximately 5760 GB of disk space. In addition to the frame snippet containing the fly the exact position of that snippet in the original video frame was stored to keep track of the walking trajectory and to enable reconstruction of the original view. Exact time points of recording for each frame were saved to enable the post hoc detection of frame drops. Every 20 frames the full resolution frame showing the whole arena and the cropped version showing just the fly were displayed in the graphic user interface (GUI) to enable live supervision during experiments (Figure 3).

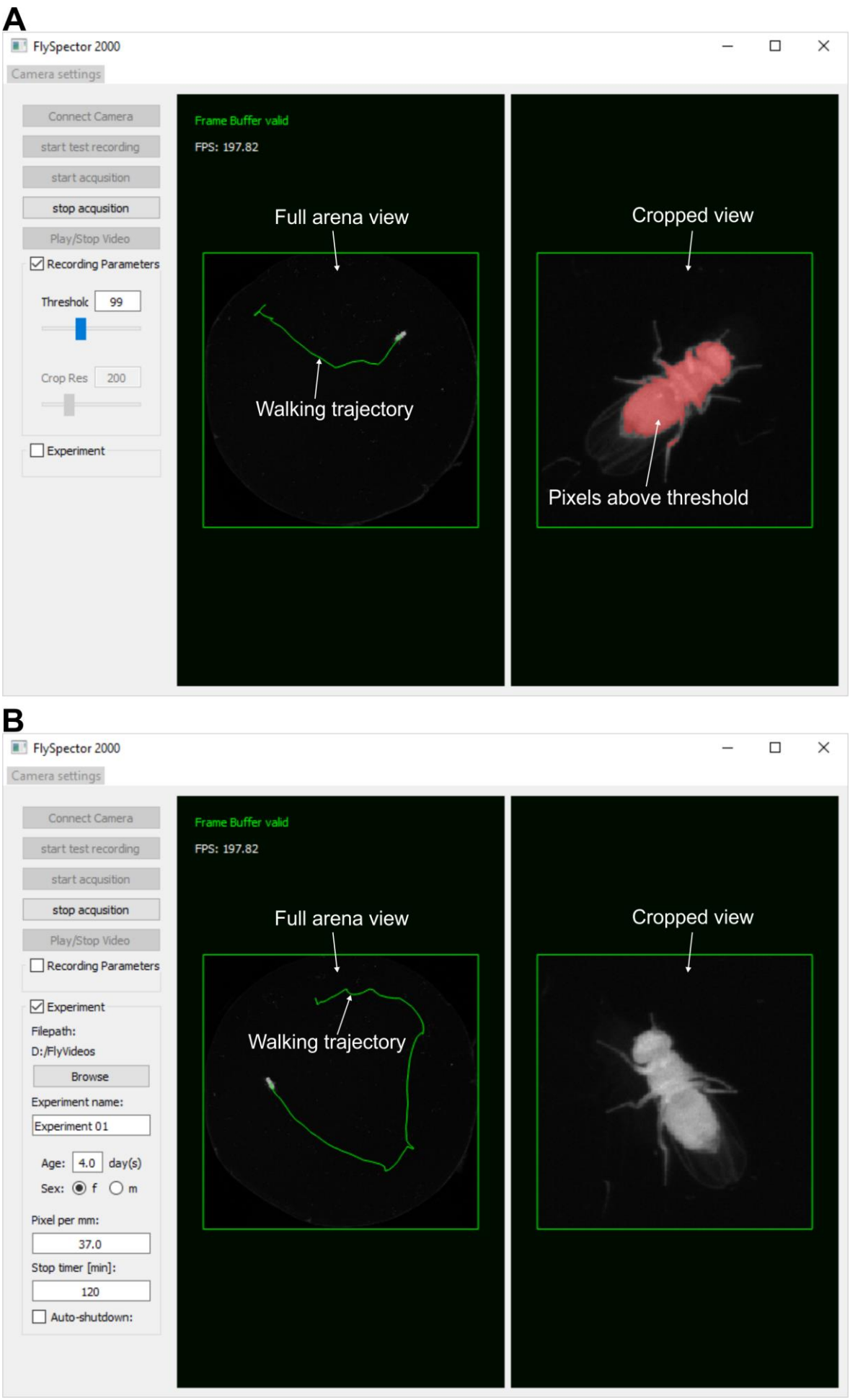

A
FlySpector 2000
Camera settings
Connect Camera
start test recording
start acqusition
stop acqusition
Play/Stop Video
Recording Parameters
Threshold
99
Crop Res
200
Experiment
Frame Buffer valid
FPS: 197.82
Full arena view
Walking trajectory
Cropped view
Pixels above threshold
B
FlySpector 2000
Camera settings
Connect Camera
start test recording
start acqusition
stop acqusition
Play/Stop Video
Recording Parameters
Experiment
Filepath:
D:/FlyVideos
Browse
Experiment name:
Experiment 01
Age:
4.0
day(s)
Sex:
f
m
Pixel per mm:
37.0
Stop timer [min]:
120
Auto-shutdown:
Frame Buffer valid
FPS: 197.82
Full arena view
Walking trajectory
Cropped view

Figure 3: Screenshots of the graphic user interface (GUI) used for controlling the setup. (A) Test recording view. The left side shows the control panel for camera and file handling. The lower part of the control panel shows the selected recording parameters section, which provides a slider for setting the brightness threshold. The pixels with brightness values above the current threshold are displayed in real time with a red overlay in the cropped view on the right. The whole arena and the most recent walking trajectory of the fly are shown in the middle. (B) Experimentation view. The lower part of the control panel on the left now shows the Experiment parameters which are used to select where files are saved and what additional information is stored. The cropped view on the right shows the video frames as they are written to the hard drive, but with a 20-fold lower frame rate.

The positions of the fly for the last 5 seconds were depicted with a green line on the whole arena view (mid sections of panels A and B in Figure 3). A test-acquisition mode without data storage was implemented to adjust lighting, focus, camera settings, and recording parameters, such as the brightness threshold, prior to an experiment (Figure 3A). The GUI also provided the opportunity to add information like the age and sex of a fly (Figure 3B, control panels on the lower left side). The pixels-to-mm ratio could also be saved with the other information, but needed to be manually determined by placing a piece of grid paper on top of the walking substrate and measuring the number of pixels equaling 1 mm.

4.2.3 Behavioral experiments

Experiments were performed in the same way as for the old setup (see **2.3.3**), but the whole arena was captured continuously due to the advanced hardware (camera and the 10 Gbit s^{-1} ethernet connection) and optimized detection algorithm described above.

4.2.4 Processing of recorded behavior

The adjustments in hard- and software allowed for continuous video recordings for multiple hours. The resulting data sets contained the complete set of behavior a fly can produce in the setup, such as walking, grooming, inspecting the outer rim of the arena, and standing still for minutes. To algorithmically filter these data sets, a set of three parameters was defined. These parameters were: (1) walking speed, (2) curvature of the walking trajectory, and (3) distance to the center. All three parameters were calculated and assigned to each individual video

frame. In addition, a minimum trajectory length was established to filter out segments that, while valid, were too short for analysis purposes.

Walking speed was calculated by measuring the distances between the fly positions in every 21th frame (corresponding) to smooth out speed fluctuations which happen when legs transition between swing and stance phase. An additional benefit of smoothing was the reduction of noise caused by both the discretization inherent in the detection algorithm and the brightness fluctuations of the moving fly body. These fluctuations can affect which pixels surpass the brightness threshold, thereby influencing the determination of the fly position.

Trajectory curvature was approximated by fitting a spline function to an equidistantly resampled version of the trajectory, followed by evaluating this spline to compute a curvature value at each point along the original trajectory. The application of the spline function effectively filtered out minor bumps, wiggling of the fly body, and artifacts from the detection algorithm, while preserving genuine changes in the direction of movement. To acquire a basis for the spline function to work on, the original walking trajectories were rearranged into equidistant pieces. This means that the original trajectory was resampled and after every 0.25 mm a new anchor point was taken. These anchor points then formed the equidistant trajectory and were used to calculate separate spline functions for the x- and y-components. These separate spline functions could be combined afterwards, resulting in a smoothed version of the original trajectory, which left out smaller bumps and spikes in the trajectory but reflected changes in the overall direction of movement. Because this smoothed trajectory was fully based on the spline functions, these functions could directly be used to calculate the curvature at any point of the trajectory.

For the calculation of the distance to the center the position of the arena in the original camera frame needed to be estimated. For this purpose, the smallest circle containing all recorded fly positions was calculated and the center of that circle was the approximation of the center of the arena. The distance of each point on the walking trajectory to this approximated arena

center was calculated to enable subtraction of every behavior which happened close to the rim of the arena.

All consecutive frames meeting the set criteria for walking speed, curvature, and distance to the center formed one potential walking bout. If this bout met the set minimum trajectory length, it was committed to the analysis pool.

4.2.5 Additional approaches for curvature analysis

In addition to the curvature analysis based on the spline function described above two additional approaches for assessing the curvature of walking trajectories were employed to analyze and compare advantages and limitations of each approach. The second approach was angular velocity, or more precisely the average change in body orientation for each mm of walking trajectory. The calculation of body orientation was based on determining the main body axis, defined by neck and abdomen markers, which were automatically annotated using DeepLabCut (Mathis et al., 2018). The difference in body orientation between the first and the last frame of each trajectory was measured in degrees and divided by the total length of the trajectory in mm. Because each trajectory already followed a single curve in either right or left direction, this approach reliably provided the average turning rate of the fly for that specific trajectory.

For the third approach a circle was fitted to the walking trajectory and the inverse of the circle's radius was used as the curvature value. To achieve this, the trajectory was rotated so that the first and last positions were aligned vertically. The first and last positions, along with the most extreme x-position of the rotated trajectory, were used to construct a circle that passed through all three points. The inverse of the radius of this circle was then taken as the curvature value for the trajectory.

To compare all three approaches, a set of curved trajectories was extracted from a 2-hour recording session with a single fly. The settings for this extraction were as follows: walking speed range: 1.5 to 20 mm s^{-1} (corresponds to approx. 1 to 10 body lengths (BL) per second);

relative distance to center: up to 90 % of the diameter of the arena; curvature: 0.05 to 0.3 radius^{-1} (radius in mm); minimum trajectory length: 6 mm (corresponds to approx. 3 BL). The resulting trajectories were categorized into left and right turns and ordered based on their curvature values as determined by the spline function approach. Each trajectory was color-coded according to the curvature values calculated with all three methods – spline function, body orientation, and circle-fitting. This color-coding allowed for a visual comparison of how each approach represented the curvature of each trajectory and highlighted the similarities and differences between the methods.

4.3 Results

4.3.1 Selection of valid segments

The relevance of specific parts of the recorded behavior largely depends on the scientific question being addressed. The approach presented here enables the flexible definition of ranges for two key parameters of walking behavior: walking speed and the curvature of walking trajectories. Setting lower and upper limits for both parameters typically results in segmentation of the original trajectories, as demonstrated in Figure 4B and Figure 5B.

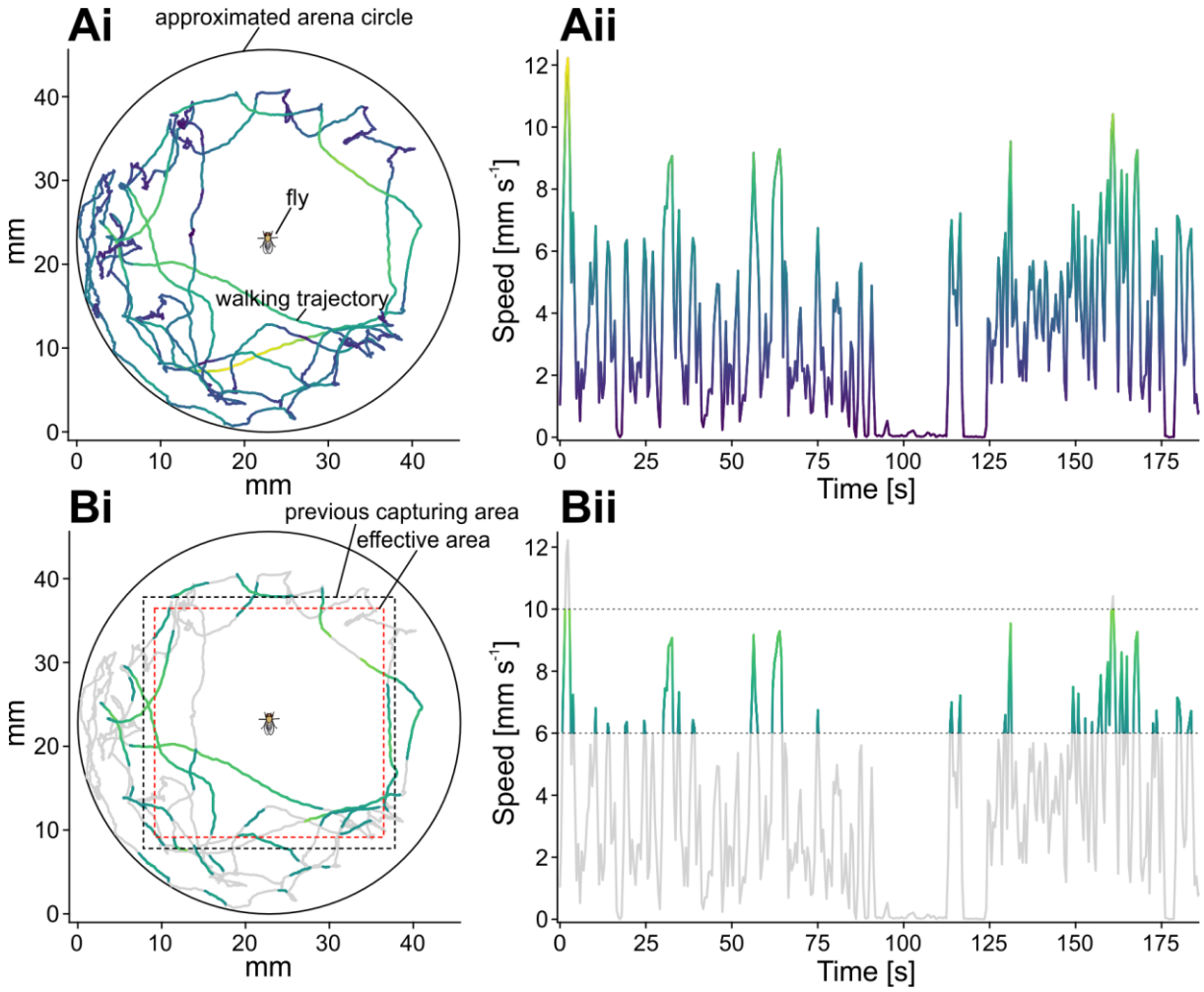


Figure 4: Example of the speed analysis. (A) Left: Walking trajectory of a fly colored according to its walking speed displayed on the right. Walking speed is color-coded with a color map ranging from dark blue (standing/slow walking) to yellow (fast walking). A schematic figure of a fly is displayed in the arena center for size comparison. (B) Left: Same walking trajectory as in A, but only segments between 6 and 10 mm s^{-1} are selected and colored according to the walking speed displayed on the right. The black dashed square indicates the capture area of

the previous version of the setup. The red dashed square highlights the region that could be effectively used for data analysis, as parts of the fly were usually not out of the frame when its center was located outside of the effective area.

Figure 4A illustrates exemplarily how walking speed varies over time during a relatively short part of a recording session. Despite the applied smoothing (see Methods), it is evident that the fly tended to accelerate rapidly for short periods before decelerating just as quickly. This behavior results in long sequences of walking within a narrow range of speeds being relatively rare, as seen in the example for a range of walking speeds between 6 and 10 BL s^{-1} in Figure 4B. The advantage of continuously capturing the entire arena is also evident in this case, as a large fraction of the behavior of interest was observed outside of the previous capturing area (Figure 4B, black dashed square, compare **2.3.2**). Note that only frames displaying the entire fly body could be used for analysis, which is why the effective area (Figure 4B, red dashed square) is even smaller than the total previous capture area.

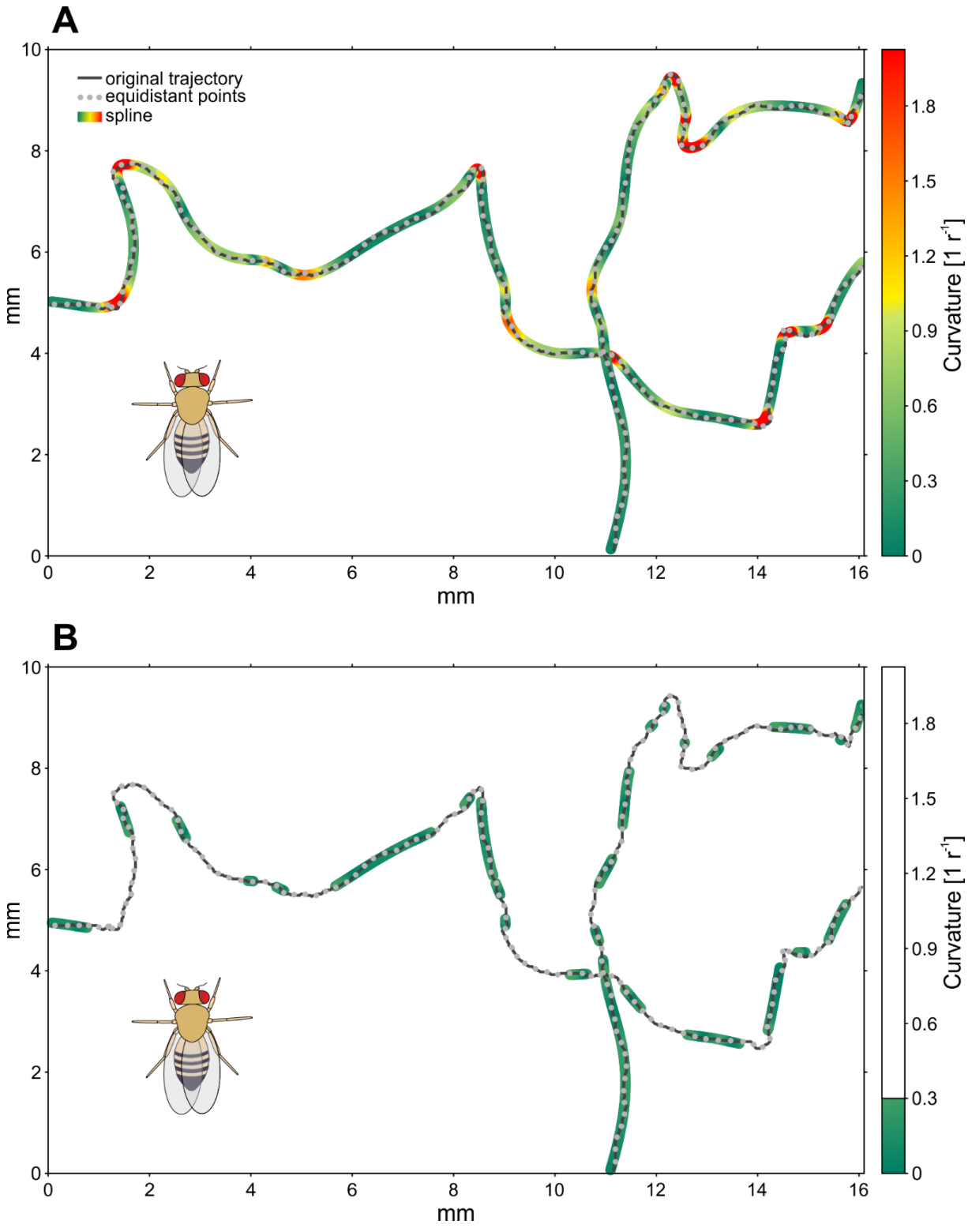


Figure 5: Example of the curvature analysis. (A) Walking trajectory of a fly (dark grey line), equidistant points along the trajectory (light grey dots), and the spline function which was fitted to the equidistant points. The spline function is colored according to its curvature. A schematic figure of a fly is displayed in the lower left corner for size comparison. (B) Only regions of the spline function below 0.3 curvature are shown; this is associated with straight walking.

Figure 5A demonstrates that the curvature of the fly's walking behavior also fluctuates regularly, similar to what was observed for walking speed. Restricting the curvature to values below 0.3, i.e. straight walking, as shown in Figure 5B, leads to segmentation of the original trajectory. Many of these segments within the defined curvature range are very short, making them unsuitable for further analysis.

Figure 6A shows the walking trajectory recorded over a 2-hour session with a single fly. While not all flies exhibit this level of spontaneous activity, this example illustrates the substantial amount of data that can be generated with the presented setup. However, it is evident that the fly primarily walked along the outer rim of the arena, a common observation in these kinds of setups. While this behavior may also be of interest, most types of analyses will likely focus on free-walking behavior without direct interaction with the arena's borders. To address this, a third criterion based on the fly's distance from the center was implemented to efficiently filter out behavior occurring too close to the rim. Figure 6B displays all behavior recorded within 90 percent of the arena’s diameter.

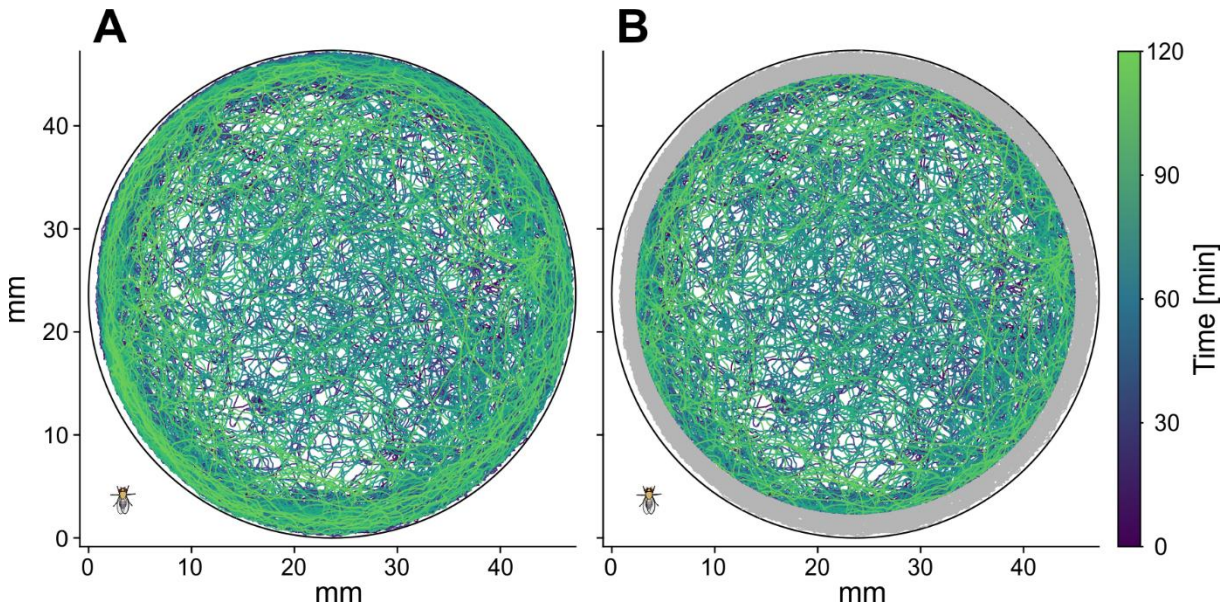


Figure 6: Example of the distance-to-center analysis. (A) The whole trajectory recorded for one fly over 2 hours. The trajectory is colored according to the time at which it was recorded during the session (see color bar on the right). A schematic figure of a fly is displayed in the lower left corner for size comparison. (B) Trajectory segments outside 90 percent of the arena's diameter are shown in grey, indicating that they are excluded from the analysis in this example.

For most experiments, all three parameters presented here – walking speed, curvature, and distance to the center – need to be controlled. The fact that each frame is assigned a value for these parameters enables to systematically and efficiently search for segments of walking behavior which meet all criteria simultaneously (Figure 7).

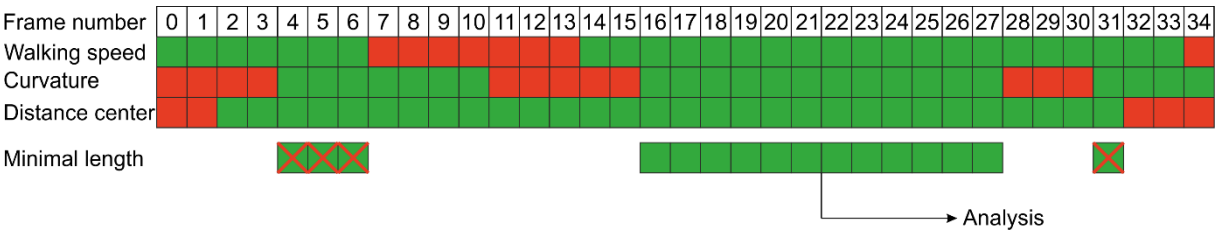


Figure 7: Example of the selection process for valid segments. Red filled squares indicate that the value of the parameter assigned to a frame was not in the defined range, while green squares indicate frames within parameter range. Only segments of walking behavior which are in range of all three parameters and are additionally of at least minimal length are chosen for analysis.

After the selection process schematically illustrated in Figure 7, the video segments selected for analysis, along with the associated meta data, are saved separately to a designated directory. The data format of these extracted segments was designed to be compatible with the analysis pipeline used before, enabling the automated annotation of body parts with DeepLabCut (DLC, Mathis et al. 2018) and application of all analysis tools which were developed for data from the previous setup. The whole process of data pre-selection described here takes only a short time (in the order of seconds) and all used parameters are readily adjustable. Overall, the advanced version of the experimental setup together with the optimized analysis pipeline allows to record more data and analyze it faster and more conveniently while retaining the same spatial and temporal resolution as for the previous setup.

4.3.2 Use case: Curvature analysis

To showcase the new analysis pipeline and to investigate the advantages and limitations of the curvature assessment with a spline function (see Methods), two alternative approaches were employed. Figure 8A shows the whole trajectory over the course of a two-hour experiment as well as the selected segments of curved walking behavior. The settings (see Methods) resulted in 67 selected trajectories, which were grouped into 27 turns in left and 40 in right direction (Figure 8B). The results of the circle fitting approach depicted in Figure 8Cii are similar to those of the spline function approach with an average difference of 0.016 radius^{-1}. The observed differences between the two approaches probably result from the fact that both of them are derived from application of fitting algorithms, which necessarily result in simplification of the underlying data. In other words: the spline-function and the circle-fitting approach are both biased, but in different directions. The clear advantage of the spline function is that it can be applied to any shape of trajectory, while the circle fitting only works well for preselected, relatively even arcs. A characteristic of the spline function approach is that it is based on multiple parameters, from the distance between equidistant points on the original trajectory to tolerance and the weighting of different points on the trajectory for the smoothing splines. These number of parameters, which can all have a strong effect on the outcome, can be an advantage, as they allow for configuration of the approach for different purposes, but they also make the approach less objective, because there is no perfect set of parameters and each variation yields different results, to some extent. The curve fitting on the other hand is a simple algorithm and provides an easily understandable interpretation of trajectory curvature, as all smooth curves eventually approximate a circle.

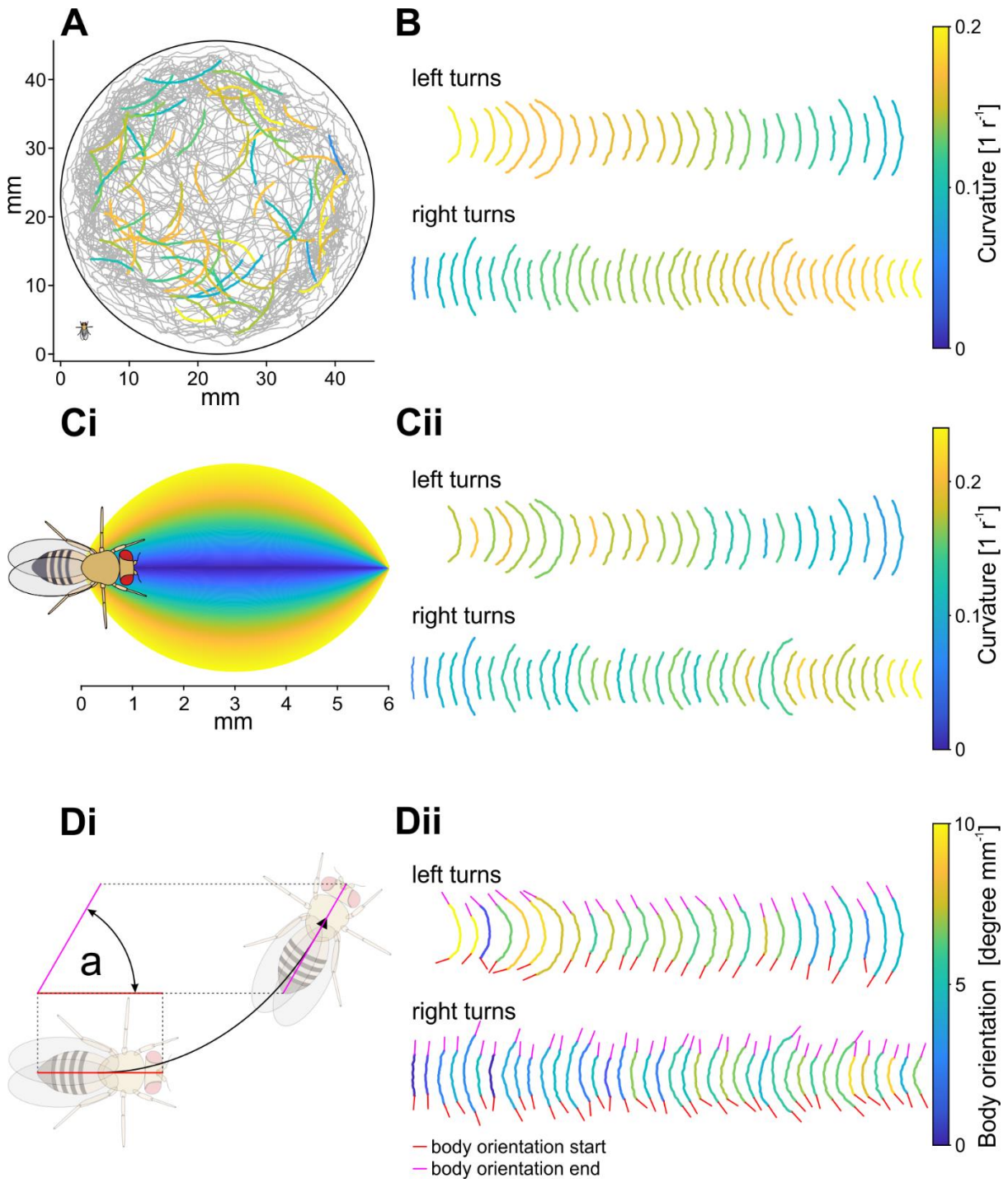


Figure 8: Comparison of different approaches for assessing the curvature of walking trajectories. (A) The whole trajectory recorded for one fly over two hours in grey and all segments selected for analysis colored according to their mean curvature value as determined with the spline functions approach (see Methods section). A schematic figure of a fly is displayed in the lower left corner for size comparison. (B) Trajectories from (A) grouped in left and right turns and ordered according to their mean curvature value as determined with the spline-function approach. (C) Left: Schematic illustrating segments of circles with different radii colored according to the corresponding curvature values in radius^{-1} (see color bar on the right).

A schematic figure of a fly is displayed for size comparison. Right: The same trajectories as in B in the same order, but colored according to the curvature value as determined with the circle-fitting approach. (D) Left: Schematic illustrating how the difference in body orientation between the first and the last frame of a trajectory were calculated. The angle 'a' was divided by the total length of the walking trajectory to determine the curvature value. Right: The same trajectories as in B and C in the same order, but colored according to results of the body orientation approach. Red and magenta lines at the start and end point of the trajectories indicate the body positions in the first and last frame.

Although displaying the same trend as in B and C, color coding of trajectories according to the change of body orientation in Figure 8Dii shows clear differences to the spline function approach. One particularly clear example for this is the third trajectory of left turns, which had a relatively high curvature value according to the spline function approach and simultaneously shows very low change in body orientation. Observation of the raw video data revealed that the fly was standing prior to this walking segment and turned into the walking trajectory with the initiation of walking. This issue could be reduced by using a higher minimum walking speed to increase the distance between the walking behavior and standing. Another idea would be to cut of the first part of each selected segment to reduce the influence of what ever happened before the parameters entered the set range, but this approach would also reduce the amount of data which is left for analysis. Other examples, like the fourth right turn in Figure 8Dii, show a final body orientation (magenta line) which is not in line with the last piece of the trajectory. Here, the orientation of the fly's body changed more slowly than the direction of walking, resulting in slight side walking behavior (data not shown). While the change in body orientation is a good measure for detecting and quantifying these idiosyncrasies, it does not always correlate well with the observed curvature on the trajectory level.

The different methods for quantifying curve walking behavior presented here all have their own advantages and limitations. The spline function approach is the only one applicable on the raw data sets, as circle-fitting only works for relatively short and smooth trajectories and the precise determination of the body angle depends on previous annotation of the body axis. However, for further analysis a combination of either spline-function or circle-fitting with body orientation

is probably the best choice to differentiate between the shape of the trajectory and how the fly oriented itself.

4.4 Discussion

The changes applied to the setup, consisting of the new camera, the optimized fly detection, and the new approach for data pre-selection, provide multiple advantages. Now, the entire arena can be continuously captured for practically unlimited periods of time, allowing for a more comprehensive view of a fly's behavior and much richer data sets. The fast and reliable filtering of behavior based on walking speed and curvature as two of the most central parameters defining walking behavior represents a critical advancement. The new analysis pipeline is well-suited to quickly extract subsets of data for addressing specific scientific questions.

However, a fundamental limitation of the approach presented here is that it can only be used for studying a single fly at a time, while other state of the art approaches, like TRex (Walter and Couzin, 2021), can track hundreds of individuals at once. TRex reliably tracks the position, speed, and orientation of multiple individuals and can be used on various different species (Walter and Couzin, 2021). In contrast, the approach presented here is highly specialized and allows for the detailed analysis of single flies at frame rates of up to 200 Hz. TGrabs, the real-time tracking version of TRex, only works with frame rates up to 60 Hz (Walter and Couzin, 2021), which is too low for the analysis of *Drosophila* leg kinematics, because *Drosophila* can exhibit stepping frequencies of up to 20 Hz (approx. 50 ms per step) and swing durations below 30 ms (see **3.3.1**). At a frame rate of 60 Hz, this would on average result in less than 3 frames per swing phase, which is insufficient to accurately capture the transitions between swing and stance phases, as well as precise lift-off and touch-down positions. A frame rate of 60 Hz would also lead to significant discretization effects, impairing the differentiation between inter leg coordination patterns and resulting in an overall loss of detail. Live detection of the fly’s position is crucial though, because the amount of data accumulating at 200 Hz and a resolution of 2000-by-2000 pixels cannot be stored to a hard drive continuously. The high spatial resolution

again is needed to track freely moving flies in a relatively large area to obtain sufficiently long segments for the characterization of walking behavior.

One approach to solve this trade-off situation between recording frequency and video resolution is the FlyMAD system, which uses a mirror attached to a two-axis galvanometer to keep the camera view directed to the fly position (Bath et al., 2014). This way, the FlyMAD system simultaneously achieves high frequency and high resolution recordings of freely walking *Drosophila* while keeping the resolution low enough to allow for continuous capturing. In a sense, FlyMAD works very similarly to what the setup described here does, with the difference that FlyMAD achieves this largely via additions to hardware. However, the FlyMAD system is a highly developed and complex system which is difficult to implement and maintain, hence it might be less robust and prone for malfunctions. These putative difficulties in the handling of the FlyMAD system might be a reason for it not being used for any study since it has been described for the first time in 2014.

Another alternative is the ball tracking setup, in which flies are tethered to a fine wire and placed on an air-suspended ball that acts as a spherical treadmill (Berendes et al., 2016; Seelig et al., 2010). This approach can be used for high resolution recordings of walking flies from multiple angles with a high resolution of the fly's body, enabling high detail 3-D reconstruction of leg kinematics (Günel et al., 2019; Haustein et al., 2024; Lobato-Rios et al., 2022). However, continuous capturing using multiple cameras is not feasible due to the rapid accumulation of large amounts of data. Simpler versions with a single camera still necessitate tethering flies and placing them in the right position over the ball prior to experiments. This procedure makes all walking behavior recorded in the ball tracking setup somewhat artificial, because flies have to walk on a curved surface, do not experience visual flow, get different feedback from pushing against the ground, and have to deal with increased inertia, because the ball is multiple times heavier than the fly.

Overall, the advantages of the setup presented here lie in its relatively simple and robust design, as well as its ability to continuously record freely moving flies with high spatial and

temporal resolution over extended periods of time. This combination of precision and durability makes it well-suited for detailed long-term behavioral studies.

The presented use-case of comparison between different curvature assessment methods highlights key advantages and limitations within each approach. The spline function approach, while flexible and capable of capturing complex trajectory shapes, relies on various adjustable parameters, such as the distance between equidistant points, tolerance, and the option to apply different weightings to every point on the trajectory. This flexibility is beneficial for adapting the method to different experimental needs, but it also introduces some degree of subjectivity, because the settings strongly depend on the researcher's assessment of the respective fit. In contrast, the circle-fitting method, though more constrained and primarily suited for relatively smooth and continuous arcs, provides a straightforward and easily interpretable measure of curvature. The differences observed between assessment of the trajectory curvature and measuring changes in body orientation show that both aspects need to be considered when analyzing curved walking behavior in *Drosophila*, because leg kinematics probably depend significantly on both.

In conclusion, the combination of advanced hardware, optimized detection algorithms, and flexible analytical approaches provides a powerful framework for studying fly locomotion. The use-case of curvature analysis showed how specific segments of walking behavior could be extracted and further analyzed from a large data set, giving an example for the new opportunities provided by the improved setup. Future work could build upon these findings by incorporating additional parameters at the level of raw recordings, such as body orientation or leg kinematics, to enable even more precise and efficient data selection. This would allow for a deeper analysis of *Drosophila* locomotion and provide a more comprehensive and versatile approach for addressing specific scientific questions.

5 Chapter 5: General Discussion

In this dissertation, I created an extensive data set of freely-walking *Drosophila* and applied PCA to investigate systematic components of inter- and intraindividual variability (chapter 2). The same data set was used to analyze the relationship between walking speed and variability of leg tip kinematics (chapter 3). Finally, I developed an advanced version of the setup which was initially used for generating the data set (chapter 4).

A key contribution of this work is the finding that *Drosophila* walking behavior is not uniform across individuals, even within highly controlled experimental conditions. Despite being genetically similar and raised in identical environments, the flies exhibited distinct walking styles, particularly in terms of interleg coordination and posture. This interindividual variability, captured by PCA, reveals that specific principal components (PCs 2, 4, and 5) are associated with idiosyncratic aspects of walking behavior, while other components (PCs 1 and 3) are linked to more universal features of locomotion, such as the continuum of interleg coordination patterns which is observed in flies. The ability to decompose these distinct aspects of walking behavior into different PCs provides a powerful framework for characterizing and quantifying data of freely-walking flies. One example for this is the presented use of PCs 2, 4, and 5 for identification of systematic postural changes caused by optogenetic inhibition of sensory neurons (chapter 2). This demonstrates the potential of this approach to serve as a sensitive and efficient method for quantifying the effects of genetic or environmental perturbations on walking behavior.

The findings presented for the PCA analysis also raise important questions about the relationship between walking speed and behavioral variability. As walking speed increases, the contribution of interindividual differences (captured by PCs 2, 4, and 5) decreases, while the contribution of interleg coordination patterns (PCs 1 and 3) increases. This suggests that higher walking speeds impose stricter constraints on motor coordination, potentially due to the stability requirements of fast locomotion. This preliminary finding from the PCA analysis that variability decreases with increasing walking speeds is further elaborated on by specifically

examining the relationship between walking speed, spatial and temporal variability, and interleg coordination. I show that both spatial and temporal variability of leg movements decrease as walking speed increases and that TCS, as a measure of interleg coordination, is a better predictor for variability than walking speed. I additionally introduce the concept of coherence as a measure of how well-timed putative intersegmental influences between adjacent legs are. The coherence measure is based on a set of coordination rules, often referred to as Cruse rules after Holk Cruse, which have been demonstrated experimentally in the stick insect *Carausius morosus*. In direct comparison, this coherence measure outperforms both walking speed and TCS in predicting motor variability, suggesting that the synchronization of leg movements plays a crucial role in reducing variability at higher speeds (chapter 3).

Another feature of this thesis is the development and application of an advanced experimental setup designed to address several challenges in studying *Drosophila* walking behavior. Previous methods, such as the multi-animal tracking systems TRex (discussed in chapter 4), presented limitations in terms of spatial and temporal resolution, which were insufficient for capturing the rapid leg movements and subtle kinematic details that characterize fly locomotion. The new system introduced in this work resolves these issues by offering continuous, high-resolution recordings at frame rates of up to 200 Hz. This setup not only enables long-term behavioral studies of freely moving flies but also ensures that fine-scale details of leg tip kinematics and interleg coordination are captured. Another strength of the new system is its ability to isolate specific parameters of interest through a robust data pre-selection process. This process allows for the usage of key parameters such as walking speed and trajectory curvature to efficiently and reliably select segments of interest or exclude unwanted behavior. As for the previous version of the setup, the selected segments of behavior can be automatically annotated with DLC and the resulting data of leg-tip and body marker positions can be used to address specific scientific questions. One example for this is the presented analysis of trajectory curvatures with different approaches (chapter 4), in which the annotation of the fly's body axis was used to quantify the change in body orientation relative to the trajectory length. Here, the opportunities of the new setup were used to conveniently filter for

curved trajectories, which was combined with DLC annotations to demonstrate that both the trajectory shape and the relative orientation of the fly need to be considered when studying curve walking.

In summary, this thesis explores the variability and individuality of *Drosophila* walking behavior, revealing that despite genetic and environmental similarities, flies exhibit distinct locomotor patterns. Principal components analysis (PCA) effectively decomposes walking behavior into both universal coordination features and individual-specific adaptations. The identification of these components facilitates the quantification of behavioral responses to genetic and environmental interventions, as demonstrated through optogenetic manipulation. Furthermore, the analysis of walking speed and variability highlights that higher speeds reduce individual differences while enhancing coordination, suggesting greater motor constraints at high speeds. The application of coherence as a measure of leg synchronization suggests that systematic changes in interleg coordination towards more tripod-like patterns play a larger role in the decrease of variability than the increase of walking speed itself. The development of an advanced experimental setup, capable of capturing high-resolution behavioral data at 200 Hz, addresses limitations of previous methods, enabling detailed and long-term analyses of walking dynamics. The system's capacity for efficient data pre-selection and annotation allows for precise examination of specific aspects of walking behavior, as shown in the trajectory curvature analysis. This work not only advances insights into insect locomotion but also provides a foundation for future research into motor control mechanisms.

Looking forward, future research could build on this work by expanding the analysis to include a wider range of behaviors, such as curve walking or obstacle negotiation, as well as investigating the effects of aging, sex, or genetic diversity on the idiosyncrasies and variability of walking behavior. The application of more advanced dimensionality reduction techniques, such as UMAP or t-SNE, could also provide deeper insights into potential non-linear relationships within walking kinematics. The Cruse rules or a different set of coordination rules could be proposed and confirmed for *Drosophila*, enabling a more detailed and precise analysis of coherence and its relationship with kinematic variability. The same is true for a

potential model based approach reflecting the weakly coupled oscillators controlling single leg movements in insects, which could yield novel insights into whether and how the entrainment dynamics of multiple CPGs affect variability. For the experimental setup an obvious expansion would be the option to record two or more individuals and their interactions simultaneously.

In conclusion, this thesis represents a significant step towards a better understanding of the interplay between general locomotor patterns and individual-specific motor output. The methods and findings discussed here have far-reaching implications for the study of motor control across species, offering a framework that could be applied to other insects or even vertebrates to explore the genetic and environmental factors contributing to behavioral variability.

List of Figures

Chapter 2

Chapter 3

Chapter 4

References

Akay, T., Bässler, U., Gerharz, P., Büschges, A., 2001. The Role of Sensory Signals From the Insect Coxa-Trochanteral Joint in Controlling Motor Activity of the Femur-Tibia Joint. Journal of Neurophysiology 85, 594–604. https://doi.org/10.1152/jn.2001.85.2.594

Alexander, R.M., 2003. Principles of animal locomotion. Princeton university press.

Aminzare, Z., Holmes, P., 2019. Heterogeneous Inputs to Central Pattern Generators Can Shape Insect Gaits. SIAM J. Appl. Dyn. Syst. 18, 1037–1059. https://doi.org/10.1137/18M120021X

Aminzare, Z., Srivastava, V., Holmes, P., 2018. Gait Transitions in a Phase Oscillator Model of an Insect Central Pattern Generator. SIAM J. Appl. Dyn. Syst. 17, 626–671. https://doi.org/10.1137/17M1125571

Ayroles, J.F., Buchanan, S.M., O'Leary, C., Skutt-Kakaria, K., Grenier, J.K., Clark, A.G., Hartl, D.L., De Bivort, B.L., 2015. Behavioral idiosyncrasy reveals genetic control of phenotypic variability. Proc. Natl. Acad. Sci. U.S.A. 112, 6706–6711. https://doi.org/10.1073/pnas.1503830112

Bässler, U., Wegner, U., 1983. Motor Output of the Denervated Thoracic Ventral Nerve Cord in the Stick Insect *Carausius Morosus*. Journal of Experimental Biology 105, 127–145. https://doi.org/10.1242/jeb.105.1.127

Bateman, A.J., 1948. Intra-sexual selection in Drosophila. Heredity 2.3, 349–368.

Bath, D.E., Stowers, J.R., Hörmann, D., Poehlmann, A., Dickson, B.J., Straw, A.D., 2014. FlyMAD: rapid thermogenetic control of neuronal activity in freely walking Drosophila. Nat Methods 11, 756–762. https://doi.org/10.1038/nmeth.2973

Berendes, V., Zill, S.N., Büschges, A., Bockemühl, T., 2016. Speed-dependent interplay between local pattern-generating activity and sensory signals during walking in *Drosophila*. J Exp Biol 219, 3781–3793. https://doi.org/10.1242/jeb.146720

Bidaye, S.S., Bockemühl, T., Büschges, A., 2018. Six-legged walking in insects: how CPGs, peripheral feedback, and descending signals generate coordinated and adaptive motor rhythms. Journal of Neurophysiology 119, 459–475. https://doi.org/10.1152/jn.00658.2017

Bidaye, S.S., Laturney, M., Chang, A.K., Liu, Y., Bockemühl, T., Büschges, A., Scott, K., 2020. Two Brain Pathways Initiate Distinct Forward Walking Programs in Drosophila. Neuron 108, 469-485.e8. https://doi.org/10.1016/j.neuron.2020.07.032

Bidaye, S.S., Machacek, C., Wu, Y., Dickson, B.J., 2014. Neuronal Control of *Drosophila* Walking Direction. Science 344, 97–101. https://doi.org/10.1126/science.1249964

Biewener, A., Patek, S., 2018. Movement on Land, in: Animal Locomotion. Oxford University Press.

Bockemühl, T., Troje, N. F. and Dürr, V. (2010). Inter-joint coupling and joint angle synergies of human catching movements. *Hum Movement Sci* 29, 73–93.

Buchanan, S.M., Kain, J.S., de Bivort, B.L., 2015. Neuronal control of locomotor handedness in *Drosophila*. Proc Natl Acad Sci USA 112, 6700–6705. https://doi.org/10.1073/pnas.1500804112

Burns, M. D. (1973). The Control of Walking in Orthoptera. I. Leg Movements in Normal Walking. *J Exp Biol* 58, 45–58.

Büschges, A., 2005. Sensory Control and Organization of Neural Networks Mediating Coordination of Multisegmental Organs for Locomotion. Journal of Neurophysiology 93, 1127–1135. https://doi.org/10.1152/jn.00615.2004

Büschges, A., 1994. The physiology of sensory cells in the ventral scoloparium of the stick insect femoral chordotonal organ. Journal of Experimental Biology 189, 285–292. https://doi.org/10.1242/jeb.189.1.285

Cabrita, A., Medeiros, A. M., Pereira, T., Rodrigues, A. S., Kranendonk, M. and Mendes, C. S. (2022). Motor dysfunction in Drosophila melanogaster as a biomarker for developmental neurotoxicity. *iScience* 25, 104541.

Chakraborty, S., Bartussek, J., Fry, S. N. and Zapotocky, M. (2015). Independently Controlled Wing Stroke Patterns in the Fruit Fly Drosophila melanogaster. *PLoS ONE* 10, e0116813.

Chockley, A.S., Dinges, G.F., Di Cristina, G., Ratican, S., Bockemühl, T., Büschges, A., 2022. Subsets of leg proprioceptors influence leg kinematics but not interleg coordination in *Drosophila melanogaster* walking. Journal of Experimental Biology 225, jeb244245. https://doi.org/10.1242/jeb.244245

Couzin-Fuchs, E., Gal, O., Holmes, P. and Ayali, A. (2015). Differential control of temporal and spatial aspects of cockroach leg coordination. *J Insect Physiol* 79, 96–104.

Cruse, H. (1976). The function of the legs in the free walking stick insect, *Carausius morosus*. *J Comp Physiol* 112, 235–262.

Cruse, H., 1990. What mechanisms coordinate leg movement in walking arthropods? Trends in Neurosciences 13, 15–21. https://doi.org/10.1016/0166-2236(90)90057-H

Cruse, H., Dean, J., Heuer, H., Schmidt, R.A., 1990. Utilization of Sensory Information for Motor Control, in: Neumann, O., Prinz, W. (Eds.), Relationships Between Perception and Action. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 43–79. https://doi.org/10.1007/978-3-642-75348-0_4

Cruse, H., Dürr, V., Schmitz, J., 2007. Insect walking is based on a decentralized architecture revealing a simple and robust controller. Phil. Trans. R. Soc. A. 365, 221–250. https://doi.org/10.1098/rsta.2006.1913

DeAngelis, B.D., Zavatone-Veth, J.A., Clark, D.A., 2019. The manifold structure of limb coordination in walking Drosophila. eLife 8, e46409. https://doi.org/10.7554/eLife.46409

Dallmann, C. J., Dürr, V. and Schmitz, J. (2016). Joint torques in a freely walking insect reveal distinct functions of leg joints in propulsion and posture control. *Proc Royal Soc B Biological Sci* 283, 20151708.

Delcomyn, F. (1971). The Locomotion of the Cockroach *Periplaneta Americana*. *J Exp Biol* 54, 443–452.

Delcomyn, F. (1989). Walking in the American cockroach: the timing of motor activity in the legs during straight walking. *Biol Cybern* 60, 373–384.

Dickerson, B.H., Fox, J.L., Sponberg, S., 2021. Functional diversity from generic encoding in insect campaniform sensilla. Current Opinion in Physiology 19, 194–203. https://doi.org/10.1016/j.cophys.2020.11.004

Dickinson, M.H., 2000. How Animals Move: An Integrative View. Science 288, 100–106. https://doi.org/10.1126/science.288.5463.100

Diedrich, F. J. and Warren, W. H. (1995). Why Change Gaits? Dynamics of the Walk–Run Transition. *J. Exp. Psychol.: Hum. Percept. Perform.* 21, 183–202.

Duffy, J.B., 2002. GAL4 system in *drosophila* : A fly geneticist's swiss army knife. Genesis 34, 1–15. https://doi.org/10.1002/gene.10150

Dürr, V., Schmitz, J., Cruse, H., 2004. Behaviour-based modelling of hexapod locomotion: linking biology and technical application. Arthropod Structure & Development 33, 237–250. https://doi.org/10.1016/j.asd.2004.05.004

Dürr, V. and Ebeling, W. (2005). The behavioural transition from straight to curve walking: kinetics of leg movement parameters and the initiation of turning. *J Exp Biol* 208, 2237-2252.

Duysens, J., Clarac, F., Cruse, H., 2000. Load-Regulating Mechanisms in Gait and Posture: Comparative Aspects. Physiological Reviews 80, 83–133. https://doi.org/10.1152/physrev.2000.80.1.83

Faisal, A.A., Selen, L.P.J., Wolpert, D.M., 2008. Noise in the nervous system. Nat Rev Neurosci 9, 292–303. https://doi.org/10.1038/nrn2258

Field, L.H., Matheson, T., 1998. Chordotonal Organs of Insects, in: Advances in Insect Physiology. Elsevier, pp. 1–228. https://doi.org/10.1016/S0065-2806(08)60013-2

Gonzalez De Santos, P., Jimenez, M.A., Armada, M.A., 1998. Dynamic Effects in Statically Stable Walking Machines. Journal of Intelligent and Robotic Systems 23, 71–85. https://doi.org/10.1023/A:1007993923530

Govorunova, E. G., Sineshchekov, O. A., Janz, R., Liu, X. and Spudich, J. L. (2015). Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics. *Science* 349, 647–650.

Grillner, S., 2006. Biological Pattern Generation: The Cellular and Computational Logic of Networks in Motion. Neuron 52, 751–766. https://doi.org/10.1016/j.neuron.2006.11.008

Grillner, S., Kozlov, A., 2021. The CPGs for Limbed Locomotion–Facts and Fiction. IJMS 22, 5882. https://doi.org/10.3390/ijms22115882

Gruhn, M., Zehl, L. and Büschges, A. (2008). Straight walking and turning on a slippery surface. *J Exp Biol* 212, 194–209.

Gruhn, M., Uckermann, G. von, Westmark, S., Wosnitza, A., Büschges, A. and Borgmann, A. (2009). Control of Stepping Velocity in the Stick *Insect Carausius morosus*. *J Neurophysiol* 102, 1180–1192.

Günel, S., Rhodin, H., Morales, D., Campagnolo, J., Ramdya, P., Fua, P., 2019. DeepFly3D: A deep learning-based approach for 3D limb and appendage tracking in tethered, adult *Drosophila* (preprint). Animal Behavior and Cognition. https://doi.org/10.1101/640375

Hales, K.G., Korey, C.A., Larracuente, A.M., Roberts, D.M., 2015. Genetics on the Fly: A Primer on the *Drosophila* Model System. Genetics 201, 815–842. https://doi.org/10.1534/genetics.115.183392

Harris, C.M., Wolpert, D.M., 1998. Signal-dependent noise determines motor planning. Nature 394, 780–784. https://doi.org/10.1038/29528

Haustein, M., Blanke, A., Bockemühl, T., Büschges, A., 2024. A leg model based on anatomical landmarks to study 3D joint kinematics of walking in Drosophila

melanogaster. Front. Bioeng. Biotechnol. 12, 1357598. https://doi.org/10.3389/fbioe.2024.1357598

Hayden, L., Lochovska, K., Sémon, M., Renaud, S., Delignette-Muller, M.-L., Vilcot, M., Peterkova, R., Hovorakova, M., Pantalacci, S., 2020. Developmental variability channels mouse molar evolution. eLife 9, e50103. https://doi.org/10.7554/eLife.50103

Holmes, P., Full, R.J., Koditschek, D., Guckenheimer, J., 2006. The Dynamics of Legged Locomotion: Models, Analyses, and Challenges. SIAM Rev. 48, 207–304. https://doi.org/10.1137/S0036144504445133

Hooper, S.L., 2012. Body size and the neural control of movement. Current Biology 22, R318–R322. https://doi.org/10.1016/j.cub.2012.02.048

Hooper, S.L., Guschlbauer, C., Blumel, M., Rosenbaum, P., Gruhn, M., Akay, T., Buschges, A., 2009. Neural Control of Unloaded Leg Posture and of Leg Swing in Stick Insect, Cockroach, and Mouse Differs from That in Larger Animals. Journal of Neuroscience 29, 4109–4119. https://doi.org/10.1523/JNEUROSCI.5510-08.2009

Hoyt, D.F., Taylor, C.R., 1981. Gait and the energetics of locomotion in horses. Nature 292, 239–240. https://doi.org/10.1038/292239a0

Hughes, G.M., 1952. The Co-Ordination of Insect Movements. Journal of Experimental Biology 29, 267–285. https://doi.org/10.1242/jeb.29.2.267

Ijspeert, A.J., 2003. Vertebrate locomotion, in: The Handbook of Brain Theory and Neural Networks. pp. 649–654.

Isakov, A., Buchanan, S.M., Sullivan, B., Ramachandran, A., Chapman, J.K.S., Lu, E.S., Mahadevan, L., De Bivort, B., 2016. Recovery of locomotion after injury in *Drosophila* depends on proprioception. Journal of Experimental Biology jeb.133652. https://doi.org/10.1242/jeb.133652

Jander, J. P. (1985). Mechanical stability in stick insects when walking straight and around curves. p. Berlin [W. Ger.]: Parey, 1985.

Jiménez, A., Lu, Y., Jambhekar, A., Lahav, G., 2022. Principles, mechanisms and functions of entrainment in biological oscillators. Interface Focus. 12, 20210088. https://doi.org/10.1098/rsfs.2021.0088

Jones, J.H., 2016. Comparative Physiology of Fatigue. Medicine & Science in Sports & Exercise 48, 2257–2269. https://doi.org/10.1249/MSS.0000000000000985

Kraus, S., Monchanin, C., Gòmez-Maracho, Lihoreau, M., 2022. Insect Diet, in: Encyclopedia of Animal Cognition and Behavior. Springer International Publishing.

Linneweber, G. A., Andriatsilavo, M., Dutta, S. B., Bengochea, M., Hellbruegge, L., Liu, G., Ejsmont, R. K., Straw, A. D., Wernet, M., Hiesinger, P. R., Hassan, A. H. (2020). A neurodevelopmental origin of behavioral individuality in the *Drosophila* visual system. *Science* 367, 1112–1119.

Lobato-Rios, V., Ramalingasetty, S.T., Özdil, P.G., Arreguit, J., Ijspeert, A.J., Ramdya, P., 2022. NeuroMechFly, a neuromechanical model of adult Drosophila melanogaster. Nat Methods 19, 620–627. https://doi.org/10.1038/s41592-022-01466-7

MacKay-Lyons, M., 2002. Central Pattern Generation of Locomotion: A Review of the Evidence. Physical Therapy 82, 69–83. https://doi.org/10.1093/ptj/82.1.69

Mamiya, A., Gurung, P., Tuthill, J.C., 2018. Neural Coding of Leg Proprioception in Drosophila. Neuron 100, 636-650.e6. https://doi.org/10.1016/j.neuron.2018.09.009

Manly, B. F. J. and Alberto, J. A. N. (2019). Multivariate Statistical Methods, A Primer, Fourth Edition.

Mantziaris, C., Bockemühl, T., Büschges, A., 2020. Central pattern generating networks in insect locomotion. Developmental Neurobiology 80, 16–30. https://doi.org/10.1002/dneu.22738

Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., Bethge, M., 2018. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci 21, 1281–1289. https://doi.org/10.1038/s41593-018-0209-y

Mendes, C.S., Bartos, I., Akay, T., Márka, S., Mann, R.S., 2013. Quantification of gait parameters in freely walking wild type and sensory deprived Drosophila melanogaster. eLife 2, e00231. https://doi.org/10.7554/eLife.00231

Mendes, C. S., Rajendren, S. V., Bartos, I., Márka, S. and Mann, R. S. (2014). Kinematic Responses to Changes in Walking Orientation and Gravitational Load in *Drosophila melanogaster*. *PLoS ONE* 9, e109204.

Milovanović, I. and Popović, D. B. (2012). Principal Component Analysis of Gait Kinematics Data in Acute and Chronic Stroke Patients. *Comput. Math. Methods Med.* 2012, 649743.

Mohammad, F., Stewart, J. C., Ott, S., Chlebikova, K., Chua, J. Y., Koh, T.-W., Ho, J. and Claridge-Chang, A. (2017). Optogenetic inhibition of behavior with anion channelrhodopsins. *Nat Methods* 14, 271–274.

Mollazadeh, M., Aggarwal, V., Thakor, N. V. and Schieber, M. H. (2014). Principal components of hand kinematics and neurophysiological signals in motor cortex during reach to grasp movements. *J. Neurophysiol.* 112, 1857–1870.

Mueller, J. M., Zhang, N., Carlson, J. M. and Simpson, J. H. (2022). Variation and Variability in Drosophila Grooming Behavior. *Front. Behav. Neurosci.* 15, 769372.

Nirody, J.A., 2021. Universal Features in Panarthropod Inter-Limb Coordination during Forward Walking. Integrative and Comparative Biology 61, 710–722. https://doi.org/10.1093/icb/icab097

Nirody, J. A. (2023). Flexible locomotion in complex environments: the influence of species, speed and sensory feedback on panarthropod inter-leg coordination. *J. Exp. Biol.* 226.

Niven, J. E., Buckingham, C. J., Lumley, S., Cuttle, M. F. and Laughlin, S. B. (2010). Visual Targeting of Forelimbs in Ladder-Walking Locusts. *Curr Biol* 20, 86–91.

O'Dell, K. and Burnet, B. (1988). The effects on locomotor activity and reactivity of the hypoactive and inactive mutations of *Drosophila melanogaster*. *Heredity* 61, 199–207.

Owaki, D., Goda, M., Miyazawa, S., Ishiguro, A., 2017. A Minimal Model Describing Hexapedal Interlimb Coordination: The Tegotae-Based Approach. Front. Neurorobot. 11, 29. https://doi.org/10.3389/fnbot.2017.00029

Pearson, K. G. and Franklin, R. (1984). Characteristics of Leg Movements and Patterns of Coordination in Locusts Walking on Rough Terrain. *Int. J. Robot. Res.* 3, 101–112.

Pearson, K.G., 1993. Common Principles of Motor Control in Vertebrates and Invertebrates. Annual Reviews.

Pearson, K.G., 1972. Central Programming and Reflex Control of Walking in the Cockroach. Journal of Experimental Biology 56, 173–193. https://doi.org/10.1242/jeb.56.1.173

Pfeffer, S. E., Wahl, V. L., Wittlinger, M. and Wolf, H. (2019). High-speed locomotion in the Saharan silver ant, *Cataglyphis bombycina*. *J Exp Biol* 222, jeb198705.

Pflüger, H.-J., 1999. Neuromodulation during motor development and behavior. Current Opinion in Neurobiology 9, 683–689. https://doi.org/10.1016/S0959-4388(99)00026-4

Pikovsky, A., Rosenblum, M., Kurths, J., 2001. Synchronization: A Universal Concept in Nonlinear Sciences, 1st ed. Cambridge University Press. https://doi.org/10.1017/CBO9780511755743

Ramdya, P., Thandiackal, R., Cherney, R., Asselborn, T., Benton, R., Ijspeert, A.J., Floreano, D., 2017. Climbing favours the tripod gait over alternative faster insect gaits. Nat Commun 8, 14494. https://doi.org/10.1038/ncomms14494

Riskin, D. K., Willis, D. J., Iriarte-Díaz, J., Hedrick, T. L., Kostandov, M., Chen, J., Laidlaw, D. H., Breuer, K. S. and Swartz, S. M. (2008). Quantifying the complexity of bat wing kinematics. *J. Theor. Biol.* 254, 604–615.

Roseman, C.C., 2020. Exerting an influence on evolution. eLife 9, e55952. https://doi.org/10.7554/eLife.55952

Schilling, M., Cruse, H., 2020. Decentralized control of insect walking: A simple neural network explains a wide range of behavioral and neurophysiological results. PLoS Comput Biol 16, e1007804. https://doi.org/10.1371/journal.pcbi.1007804

Schilling, M., Hoinville, T., Schmitz, J., Cruse, H., 2013. Walknet, a bio-inspired controller for hexapod walking. Biol Cybern 107, 397–419. https://doi.org/10.1007/s00422-013-0563-5

Schmidt, R.A., Zelaznik, H., Hawkins, B., Frank, J.S., Quinn, J.T., 1979. Motor-output variability: A theory for the accuracy of rapid motor acts. Psychological Review 86, 415–451. https://doi.org/10.1037/0033-295X.86.5.415

Seelig, J.D., Chiappe, M.E., Lott, G.K., Dutta, A., Osborne, J.E., Reiser, M.B., Jayaraman, V., 2010. Two-photon calcium imaging from head-fixed Drosophila during optomotor walking behavior. Nat Methods 7, 535–540. https://doi.org/10.1038/nmeth.1468

Seipel, J., Kvalheim, M., Revzen, S., A. Sharbafi, M., Seyfarth, A., 2017. Conceptual Models of Legged Locomotion, in: Bioinspired Legged Locomotion. Elsevier, pp. 55–131. https://doi.org/10.1016/B978-0-12-803766-9.00004-X

Simon, J.C., Dickinson, M.H., 2010. A New Chamber for Studying the Behavior of Drosophila. PLoS ONE 5, e8793. https://doi.org/10.1371/journal.pone.0008793

Strauß, R., Heisenberg, M., 1990. Coordination of legs during straight walking and turning in Drosophila melanogaster. Journal of Comparative Physiology A 167, 403–412. https://doi.org/10.1007/BF00192575

Szczecinski, N.S., Bockemühl, T., Chockley, A.S., Büschges, A., 2018. Static stability predicts the continuum of interleg coordination patterns in *Drosophila*. J Exp Biol 221, jeb189142. https://doi.org/10.1242/jeb.189142

Takeoka, A. and Arber, S. (2019). Functional Local Proprioceptive Feedback Circuits Initiate and Maintain Locomotor Recovery after Spinal Cord Injury. *Cell Rep.* 27, 71-85.e3.

Takeoka, A., Vollenweider, I., Courtine, G. and Arber, S. (2014). Muscle Spindle Feedback Directs Locomotor Recovery and Circuit Reorganization after Spinal Cord Injury. *Cell* 159, 1626–1639.

Todorov, E., 2004. Optimality principles in sensorimotor control. Nat Neurosci 7, 907–915. https://doi.org/10.1038/nn1309

Tóth, T.I., Daun, S., 2019. A kinematic model of stick-insect walking. Physiol Rep 7, e14080. https://doi.org/10.14814/phy2.14080

Tuthill, J.C., Wilson, R.I., 2016. Mechanosensation and Adaptive Motor Control in Insects. Current Biology 26, R1022–R1038. https://doi.org/10.1016/j.cub.2016.06.070

Venken, K.J.T., Simpson, J.H., Bellen, H.J., 2011. Genetic Manipulation of Genes and Cells in the Nervous System of the Fruit Fly. Neuron 72, 202–230. https://doi.org/10.1016/j.neuron.2011.09.021

Wahl, V., Pfeffer, S.E., Wittlinger, M., 2015. Walking and running in the desert ant Cataglyphis fortis. J Comp Physiol A 201, 645–656. https://doi.org/10.1007/s00359-015-0999-2

Walter, T., Couzin, I.D., 2021. TRex, a fast multi-animal tracking system with markerless identification, and 2D estimation of posture and visual fields. eLife 10, e64000. https://doi.org/10.7554/eLife.64000

Wilson, D.M., 1961. The Central Nervous Control of Flight in A Locust. Journal of Experimental Biology 38, 471–490. https://doi.org/10.1242/jeb.38.2.471

Wosnitza, A., Bockemuhl, T., Dubbert, M., Scholz, H., Buschges, A., 2013. Inter-leg coordination in the control of walking speed in Drosophila. Journal of Experimental Biology 216, 480–491. https://doi.org/10.1242/jeb.078139

Yamamoto, A., Huang, W., Anholt, R.R.H., Mackay, T.F.C., 2024. The genetic basis of variation in Drosophila melanogaster mating behavior. iScience 27, 109837. https://doi.org/10.1016/j.isci.2024.109837

York, R. A., Brezovec, L. E., Coughlan, J., Herbst, S., Krieger, A., Lee, S.-Y., Pratt, B., Smart, A. D., Song, E., Suvorov, A., et al. (2022). The evolutionary trajectory of drosophilid walking. *Curr. Biol.* 32, 3005-3015.e6.

Zenkevich, L.A., 1945. The evolution of animal locomotion. Journal of Morphology 77, 1–52. https://doi.org/10.1002/jmor.1050770102

Zill, S., Schmitz, J., Büschges, A., 2004. Load sensing and control of posture and locomotion. Arthropod Structure & Development 33, 273–286. https://doi.org/10.1016/j.asd.2004.05.005

Zill, S.N., 1985. Plasticity And Proprioception In Insects I. Responses and Cellular Properties of Individual Receptors of the Locust Metathoracic Femoral Chordotonal Organ. Journal of Experimental Biology 116, 435–461. https://doi.org/10.1242/jeb.116.1.435

Zollikofer, C. (1994). Stepping patterns in Ants – Influence of Speed and Curvature. *J Exp Biol* 192, 95–106.

Danksagung

Ich möchte folgenden Personen danken:

Prof. Dr. Ansgar Büschges für die Möglichkeit, die Zeit meiner Promotion in seiner Arbeitsgruppe zu verbringen, für seine ansteckende Begeisterung für die Wissenschaft und vor allem für seine tatkräftige und ideenreiche Unterstützung in schwierigen Phasen.

Prof. Dr. Martin Nawrot für die Übernahme der Rolle des Zweitgutachters und seine stehts fundierten und hilfreichen Anregungen in den TAC-Meetings.

PD Dr. Tom Weihmann für seinen Einsatz als Mentor und seine inspirierenden Fragen während der TAC-Meetings.

Dr. Till Bockemühl, der mich bereits bei der Durchführung meiner Bachelor- und Masterarbeit angeleitet und mir im Lauf der Jahre, in unzähligen Stunden der gemeinsamen Arbeit und des Austauschs, sehr viel beigebracht hat, sowohl im Bezug auf die wissenschaftliche Denk- und Vorgehensweise, als auch darüber hinaus.

Dr. Moritz Haustein, der stets Zeit gefunden hat, über meine Daten und Auswertungen zu diskutieren und mir mit seinem Rat häufig weitergeholfen hat. Außerdem für die Hilfe bei der Erstellung der programmatischen Grundstruktur des neuen Setups sowie für die Bereitstellung der von ihm erstellten Schemaabbildung einer Fruchtfliege, welche in den Abbildungen zu Kapitel 4 häufig zum Einsatz kam.

Der gesamten Arbeitsgruppe Büschges für die freundliche Aufnahme, vielseitige Hilfestellung, das angenehme Arbeitsklima und die interessanten Gespräche über die Wissenschaft und das Leben im Allgemeinen.

Meiner Familie und meinen Freunden dafür, dass sie immer hinter mir stehen und mir dadurch erst die nötige Sicherheit gegeben haben, auch die besonders herausfordernden Phasen durchzustehen.

Vielen Dank!

Data availability statement

This dissertation project involved the acquisition of primary data and subsequent analyses. The primary data, intermediate data, and the source code of all analysis routines were archived on the online storage file system for institutions of the University of Cologne (https://rrzk.uni-koeln.de/en/data-storage-and-share/online-storage-sofs). Data integrity is ensured by a routinely backup procedure. Data and analysis routines are available from Vincent Godesberg or Prof. Dr. Ansgar Büschges upon reasonable request.

Erklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Vincent Godesberg