

Interpretable Active Inference Gait Control Learning

Rudolf Szadkowski

Jan Faigl

Abstract—Sustaining the gait locomotion in an adversarial environment requires the robot to react to novel experiences adaptively. In Free Energy Principle (FEP), the behavioral reaction is driven by the discrepancy between observation and prediction. Although, for legged robot gait locomotion, the prediction of gait dynamics is challenging as the consequences non-linearly depend on the activity history, the animal gait is robust, adapting to severe motion disruptions seemingly instantly. In biomimetic robotics, the Central Pattern Generator (CPG) relaxes the general dynamics of body-environment interaction to the stable and repetitive dynamics of gait. Based on these observations, we propose self-learning of the gait dynamics model and FEP framework that infers state estimation and gait control. The proposed method is experimentally evaluated on a real hexapod walking robot with 18 controllable degrees of freedom. The robot learns the gait dynamics model indoors and then deploys it in outdoor navigation under various adversarial scenarios. Results show that the developed interpretable gait controller exhibits complex and real-time adaptive behavior when it encounters unknown situations.

I. INTRODUCTION

The navigation performance of legged robots depends on the capability to sustain gait locomotion in a dynamic adversarial environment. In harsh and adversarial conditions, the robot is exposed to novel terrains with different teramechanics and damage, changing the body-terrain dynamic interaction. Consequently, it leads to the difference between the robot’s *World Model* (WM) predictions and sensory observations that turn out to decrease performance. However, under the *Free Energy Principle* (FEP) [1], the prediction error awareness can be used to improve the performance, expanding the WM operational domain from seen to unseen [2].

The FEP is a mathematical cognition processing theory that provides an upper bound on the robot’s *surprisal*, a quantification of the robot’s surprise, which is then minimized. The free energy minimization can be deployed in biomimetic robotics for estimation [3] and has been expanded for control in *Active Inference Control* (AIC) [4], [5]. Since the AIC approach is relatively recent, to the best of the authors’ knowledge, there are no reported deployments of walking robots with challenging body-terrain interaction dynamics.

The gait behavior is a result of past and present interdependent motor activity. Thus, the motor space and temporal space must be considered when describing gait dynamics.

The authors are with the Faculty of Electrical Engineering, Czech Technical University, Technicka 2, 166 27, Prague, Czechia, {szadkrud|faigl.j}@fel.cvut.cz.

The presented work has been supported by the Czech Science Foundation (GAČR) under research project No. 21-33041J. In 2024, the preparation of the manuscript has been supported by the European Union under the project ROBOPROX - Robotics and advanced industrial production (reg. no. CZ.02.01.01/00/22_008/0004590).

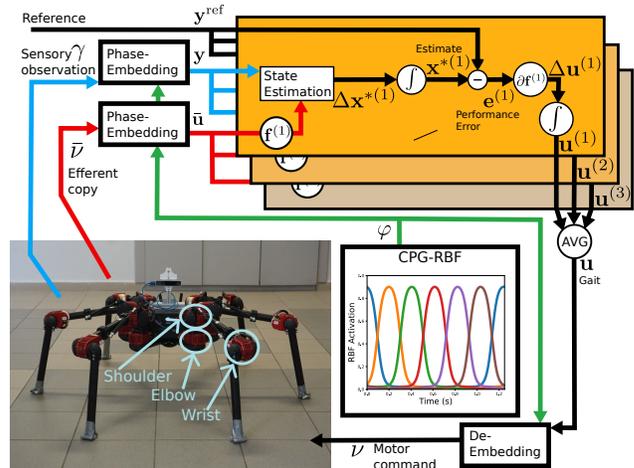


Fig. 1. The proposed gait controller deployed on the hexapod walking robot Daisy with the tracking camera Intel® RealSense™ T265 that provides a relative localization of the robot. The sensory and motor signals are transformed into the phase-embedded space in which the gait dynamics are modeled by linear models $f^{(i)}$. Each particular i th model is utilized both in state estimation $\Delta x^{*(i)}$ and control inference $\Delta u^{(i)}$ that adjusts the gait based on the difference between the reference and estimate. The gaits inferred from the available models are averaged and de-embedded into the command signal sent to the servomotors.

The *Central Pattern Generator* (CPG) is a neural network generating rhythmic signals commonly measured during animal gait [6], [7]. In biomimetic robotics, the CPG models are used for tracking the state evolution in the temporal space for control [8], [9] or estimation [10]. The CPG model can be combined with *Reinforcement Learning* (RL) [11] or self-supervised learning [12], where the CPG reduces the number of parameters needed to control and model the gait dynamics.

In this paper, we propose a robust and adaptive gait for adversarial environments that represents a computationally inexpensive plastic controller. We combine the CPG-based forward model [12] with the AIC framework into a model-based gait controller. The WM is represented as an ensemble of forward models that are incrementally bootstrapped by motor babbling. The trained WM is used by the AIC to provide state estimation and motor control by minimizing the robot’s surprisal; see the system schema depicted in Fig. 1.

The proposed method is deployed on a real hexapod walking robot, which learns to walk in thirty minutes and is capable of adapting to novel experiences. As the forward model is a linear model ensemble and the AIC is applied for estimation and control, the learned parameters and behavior inference are interpretable [13]. Furthermore, the ablation study is performed on a high-fidelity robot model to com-

pare open and closed-loop control with a single model or model ensemble. The results show that the model ensemble outperforms the single model. Besides, the open loop has the capacity to move toward the goal despite the fact that the closed-loop control outperforms the open loop.

Contributions of the paper are considered as follows.

- 1) The biomimetic controller combines the neural and cognitive principles, resulting in robust real-time gait control that self-learns on-site without the need for a dynamic, kinematic, or simulated high-fidelity model.
- 2) The method is implemented using linear models and AIC, which approximates Bayesian inference; therefore, unlike black box solutions, additionally to robot control, the method provides interpretable [13] models and control inference.
- 3) Experimental deployments on a real hexapod walking robot that is subjected to various adversarial scenarios, showing that the controller can adapt to novel experiences and navigate in unknown environments.

In the rest of the paper, an overview of the related work and background on gait locomotion and FEP are provided in Section II. The gait control problem is stated in Section III. The proposed method is described in Section IV. Achieved experimental results are reported in Section V, and the paper is concluded in Section VI.

II. RELATED WORK AND BACKGROUND

The proposed approach combines two biomimetic concepts. The first is a gait control based on the CPG, and the second is control with the FEP. Both concepts are briefly characterized in the following paragraphs to provide an overview of the existing work and background of the approaches used in the proposed method.

A. Gait and Central Pattern Generator (CPG)

The gait is a repetitive motion pattern controlled by the CPG that is a recurrent neural structure providing synchronous rhythmic oscillations that are correlated with the repetitive behavior [14], [15]. The CPG can be mathematically modeled as a self-sustained oscillator or its approximation [16]. While the gait dynamics is history-dependent, heavily non-linear, with codependent variables, the CPG-based controllers drive the gait with relatively significantly less computational effort. In [11], multiple two-layered *Neural Networks* (NN) are driven by CPG, resulting in various hexapod walking robot behaviors. Each network is trained by RL to perform a specific behavior, such as walking on flat, elevated, or curved terrain, and the behaviors are combined by the subsumption architecture.

An opposite architecture is presented in [17], where the NN, learned by RL, drives the CPG that generates rhythmic commands for snake-like soft-robot locomotion. In addition to generating motor commands, the CPG can be used for state estimation [18] to track the gait phase and detect events that are out of the phase. Hence, the CPG can drive and estimate the gait state that is reported in [12] for phase embedding deployed on the soft-robotic arm. The phase

embedding uses the estimated gait phase and sensorimotor signal to transform the control problem into phase-embedded space where the sensorimotor relation can be represented by non-recurrent regression, such as linear model or multilayer perceptron. By incrementally learning the multiple gait dynamics model, the soft arm performs dynamic motions in an open-loop control. The herein proposed approach combines phase embedding with FEP to provide the framework for closed-loop control.

B. Free Energy Principle (FEP)

The FEP is a mathematical theory of the *Predictive Coding* (PC) paradigm under which an agent continually fuses sensory information to represent the environment and predicts the consequences of its actions. The PC is corroborated by experiments in [19], where subjects lose the ability to perform precise movements when predictive capability is inhibited. The FEP has been adapted in robotic manipulators [3] to learn the estimation of the arm joint angles by fusing multiple sensory inputs (tactile, visual, and proprioception) even in the sensory input absence or in perturbed scenarios.

Based on [20], the *Active Inference Control* (AIC) is formulated [21] using a multi-layered prediction architecture that estimates the generalized state (state, its motion velocity, and acceleration) for the robot control to perform reaching, grasping, and tracking tasks. The reported AIC's advantage is robustness to embodiment changes, such as arm damage or weight addition. Furthermore, the authors of [4] show the relation between the PID control and the AIC by correspondences with finding P and I gain parameters. In [5], the AIC outperforms Model Predictive Control in manipulator tasks of reaching and tracking, yet still robust to embodiment changes. Based on the existing results, FEP is considered promising, and we continue on our preliminary work [22] to expand the AIC to the domain of mobile robots in rough environments.

III. PROBLEM STATEMENT

The robot controls its N actuators with commands $v_n(t)$ to track the given M sensory references $y_m^{\text{ref}}(t)$ that consist of planar and turning velocities. The tracking of the given reference results in navigating the robot toward the given goal. The robot measures the M tracked modalities with sensory observations $\gamma_m(t)$. Although the relationship between commands v and sensory observations γ is unknown to the robot, the locomotion is assumed to be the result of the gait behavior that relaxes general dynamics to gait dynamics.

In gait dynamics, we assume that the gait state depends only on events occurring within one *gait period* T . The gait state can be described by the *gait phase* ϕ and the command trajectory within one period $U(t) = (v(\tau))_{\tau=t-T}^{\tau=t}$. A CPG-based forward model predicting sensory state from the gait phase and command history is proposed in [12], where a motor trajectory $U(t)$ is phase-embedded into a vector \mathbf{u} , which relation with sensory observation is modeled as linear.

The present work is to utilize the linear forward model in the AIC model-based controller derived from the FEP.

IV. PROPOSED GAIT AIC BOOTSTRAPPING

The proposed controller incrementally self-learns to track a given reference with continual gait adaptation. We employ the CPG with Radial Basis Function (CPG-RBF) [23] network to transform the sensorimotor signal into phase-embedded vectors and the AIC to track the reference \mathbf{y}^{ref} by updating the gait \mathbf{u} , see Fig. 1. The gait update is performed in real time, where each $\Delta_t = 0.01\text{s}$, the AIC step summarized in Algorithm 1 is evaluated. The AIC utilizes the WM $W = \{\mathbf{f}^{(i)}\}_t^B$, which is incrementally grown by bootstrapping as described in Section IV-B.

Algorithm 1 Active Inference Control Step

- 1: **Input:** Reference \mathbf{y}^{ref} , observation γ , efferent copy $\bar{\mathbf{v}}$.
 - 2: **Output:** Motor command \mathbf{v} .
 - 3: $\phi \leftarrow \phi + 2\pi T^{-1}\Delta_t$ \triangleright Simple CPG model.
 - $\triangleright j$ is the imaginary unit, and $|\cdot|$ denotes the complex norm.
 - 4: $\forall c \in C : \Phi_c \leftarrow |\exp[j\phi] - \exp[j\frac{0.5+c}{C}2\pi]|$
 - 5: $\forall c \in C : \varphi_c \leftarrow \frac{\exp[-\Phi_c^2]}{\sum_d \exp[-\Phi_d^2]}$ \triangleright See Fig. 1 for φ_c evolution.
 - 6: $\forall c \in C : \mathbf{y}_c \leftarrow \mathbf{y}_c + (\gamma - \mathbf{y}_c)\Delta_t \varphi_c$ \triangleright Phase-embedding
 - 7: $\forall c \in C : \bar{\mathbf{u}}_c \leftarrow \bar{\mathbf{u}}_c + (\bar{\mathbf{v}} - \bar{\mathbf{u}}_c)\Delta_t \varphi_c$
 - 8: $c^{\text{current}} \leftarrow \arg \max_c \varphi_c$
 - 9: **if** $c^{\text{current}} \neq c^{\text{previous}}$ **then**
 - 10: $c^{\text{previous}} \leftarrow c^{\text{current}}$
 - 11: **for** $i \in B$ **do**
 - 12: $\mathbf{x}^{*(i)} \leftarrow \mathbf{x}^{*(i)} + \Delta \mathbf{x}^{*(i)}$ update by (6)
 - 13: $\mathbf{u}^{(i)} \leftarrow \mathbf{u}^{(i)} + \Delta \mathbf{u}^{(i)}$ update by (7)
 - 14: **end for**
 - 15: $\mathbf{u} \leftarrow B^{-1} \sum_i^B \mathbf{u}^{(i)}$
 - 16: **end if**
 - 17: **return** $\mathbf{v} \leftarrow \sum_c^C \varphi_c \mathbf{u}_c$ \triangleright De-embedding
-

The employed sensorimotor model [12] utilizes phase-embedding to transform the motor and sensory signals into phase-space, where motor-embedding can be linearly mapped to sensory-embedding. The embedding divides the phase range into C phase segments, $\phi \in [0, 2\pi) = \cup_c^C [2\pi\frac{c-1}{C}, 2\pi\frac{c}{C})$, where for each c th segment, we observe the m th sensory modality y_{mc} and command the n th motor u_{nc} . The phase segment activation is implemented by the CPG-RBF network where the CPG periodically activates C RBF neurons with the period T , see Lines 3–5¹ and Fig. 1. The activation signal φ_c is then used for the phase-embedding of sensory and motor signals, see Lines 6–7. The gait is represented by sequence of motor commands $\mathbf{u} = (\mathbf{u}_c)_c^C$, which is de-embedded at Line 17. The m th sensory modality prediction at the c th phase is

$$f_{mc}(\mathbf{u}) = b_{mc} + \sum_n^N \sum_d^C w_{mc}^{nd} (u_{nd} - u'_{nd}), \quad (1)$$

¹In this work, we use a model of unperturbed CPG phase; however, in principle Lines 3–5 can be replaced by any CPG-RBF implementation.

where the regression parameters $b_{mc}, \mathbf{w}_{mc} = (w_{mc}^{nd})_{nd}^{NC}$, and \mathbf{u}' are bias, weight, and base-gait, respectively. The gait dynamics model $\mathbf{f} = (f_{mc})_{mc}^{MC}$ thus maps the gait to sensory prediction for a single gait period. The model parameters $\{b_{mc}, \mathbf{w}_{mc}\}$ are found by linear regression on the collected data from motor babbling. The dynamics embedded by the CPG-RBF network are utilized in the AIC to infer the gait.

A. Active Inference Control as Free Energy Optimization

In the AIC, the state estimation and control update rules are inferred from the free energy. Its optimization provides an approximation to the posterior state probability

$$p(\mathbf{x}|D, \mathbf{f}) = \frac{p(D|\mathbf{x}, \mathbf{f})p(\mathbf{x}|\mathbf{f})}{p(D|\mathbf{f})} = \frac{p(\mathbf{y}|\bar{\mathbf{u}}, \mathbf{x}, \mathbf{f})p(\bar{\mathbf{u}}|\mathbf{x}, \mathbf{f})p(\mathbf{x}|\mathbf{f})}{p(D|\mathbf{f})}, \quad (2)$$

given the evidence of sensorimotor data $D = (\bar{\mathbf{u}}, \mathbf{y})$, a pair of observed motor commands (efferent copy) and sensory value, and knowledge of sensorimotor model \mathbf{f} . Unlike in [21], we define the optimization in the phase embedded space, where we assume sensorimotor linear relation (1). Hence, a shallow (single layer) architecture is sufficient, and the state \mathbf{x} is a variable of the embedded sensory space.

The state posterior is approximated by the robot's recognition density $q(\mathbf{x}; \mathbf{x}^*, \zeta) = \mathcal{N}(\mathbf{x}; \mathbf{x}^*, \zeta)$ modeled as a Normal distribution with the expected value \mathbf{x}^* and covariance ζ . The expected value is found by minimizing free energy

$$F = -\ln p(\mathbf{y}|\bar{\mathbf{u}}, \mathbf{x}^*, \mathbf{f})p(\bar{\mathbf{u}}|\mathbf{x}^*, \mathbf{f})p(\mathbf{x}^*|\mathbf{f}). \quad (3)$$

Using Laplace approximation [24], normal distributions are used to model the sensory measurement probability $p(\mathbf{y}|\bar{\mathbf{u}}, \mathbf{x}^*, \mathbf{f}) = \mathcal{N}(\mathbf{y}; \mathbf{x}^*, \Sigma^y)$, sensory prediction probability $p(\bar{\mathbf{u}}|\mathbf{x}^*, \mathbf{f}) = \mathcal{N}(\mathbf{f}(\bar{\mathbf{u}}); \mathbf{x}^*, \Sigma^u)$, and sensory prior $p(\mathbf{x}^*|\mathbf{f}) = \mathcal{N}(\mathbf{x}^*; \mathbf{b}, \Sigma^x)$. From that, the free energy can be defined as

$$F(x_{mc}^*, D) = \frac{(y_{mc} - x_{mc}^*)^2}{2\sigma_{mc}^y} + \frac{(f_{mc}(\bar{\mathbf{u}}) - x_{mc}^*)^2}{2\sigma_{mc}^u} + \frac{(x_{mc}^* - b_{mc})^2}{2\sigma_{mc}^x} \quad (5)$$

where we assume variables x_{mc}^* being independent. The state estimate is then updated by gradient descent $\dot{x}_{mc}^* = -\frac{\partial F}{\partial x_{mc}^*}$,

$$\frac{\dot{x}_{mc}^*}{\kappa^x} = \frac{y_{mc} - x_{mc}^*}{\sigma_{mc}^y} + \frac{f_{mc}(\bar{\mathbf{u}}) - x_{mc}^*}{\sigma_{mc}^u} - \frac{x_{mc}^* - b_{mc}}{\sigma_{mc}^x}, \quad (6)$$

where κ^x denotes the learning rate.

Except the estimate \mathbf{x}^* , the free energy is influenced by the motor activity \mathbf{u} [20], assuming the motor activity drives the sensory observations $F(\mathbf{x}^*, \mathbf{y}(\mathbf{u}))$. We adopt the formalized relation between the sensory variable and motor variable gradient [21] as

$$\frac{\dot{u}_{nc}}{\kappa^u} = -\frac{\partial F(\mathbf{x}^*, \mathbf{y}(\mathbf{u}))}{\partial \mathbf{y}(\mathbf{u})} \frac{\partial \mathbf{y}(\mathbf{u})}{\partial u_{nc}} = \sum_{md}^{MC} \frac{y_m^{\text{ref}}(t) - x_{md}^*}{s_{md}^{\text{ref}}} w_{md}^{nc}, \quad (7)$$

where κ^u denotes motor learning rate, with the substitute of the sensory reference \mathbf{y}^{ref} for the sensory variable.

We assume that the model approximates the sensorimotor relation $y_{md}(\mathbf{u}) \approx f_{md}(\mathbf{u})$, thus $\frac{\partial y_{md}}{\partial u_{nc}} \approx w_{md}^{nc}$. Hence, for given

model \mathbf{f} , confidence parameters $\sigma^y, \sigma^u, \sigma^x$, learning rates κ^x, κ^u , and the reference parameters $y^{\text{ref}}, s^{\text{ref}}$, the update rules (6) and (7) provide the estimation and control, respectively. The update is performed for each i th model in the WM, see Lines 12–13, and the controller output is averaged at Line 15.

B. Bootstrapping the World Model

The proposed method calculates the AIC for every model in the WM $W = \{\mathbf{f}^{(i)}\}_i^B$, which is incrementally grown by bootstrapping [12]. Each i th bootstrapping increment produces i th model, where the robot performs two stages: (i) learning by **motor babbling**, and (ii) **reference tracking** using the proposed AIC. At the start of the i th **motor babbling** stage, the current gait is set as the i th base-gait, $\mathbf{u}^{(i)} = \mathbf{u}(t)$, and then, the sensorimotor data $D^{(i)}$ is collected by perturbing the base-gait $\mathbf{u}^{(i)}$ with gaussian noise and measuring the sensory effects \mathbf{y} . After 300 gait periods, the data $D^{(i)} = \{(\bar{\mathbf{u}}, \mathbf{y})\}_j^{300}$ is regressed into model $\mathbf{f}^{(i)}$ that is added to the WM W . In the next stage, the robot **tracks the reference** with Algorithm 1 for 115 gait periods, after which the next bootstrapping increment may follow. The bootstrapping is initialized with the base-gait $\mathbf{u}^{(1)} = \mathbf{0}$. The outcome of B bootstrapping increments results in the WM as of $W = \{\mathbf{f}^{(i)}\}_i^B$.

V. RESULTS

The proposed method has been deployed on real and simulated hexapod walking robots trained separately and only for the real-to-real and sim-to-sim setups. The robustness is examined on the real robot, and a high-fidelity simulation robot model is utilized in the ablation study. The performance evaluation is to assess how the trained controller navigates the robot toward the given goal location.

The **real robot** is the hexapod walking robot HEBI Daisy, which is a six-legged walking robot with $N = 18$ controllable degrees of freedom defined by three servomotors per leg that are controlled by setting their joint angles. The Intel® RealSense™ Tracking Camera T265 provides pose estimates used for control and navigation. Five sensory modalities are utilized from the pose estimates: heading $[y_{\text{head}}]\text{ms}^{-1}$ and side $[y_{\text{side}}]\text{ms}^{-1}$ velocities, roll, pitch, and yaw change denoted as $[y_{\text{roll}}]\text{rads}^{-1}$, $[y_{\text{pitch}}]\text{rads}^{-1}$, and $[y_{\text{yaw}}]\text{rads}^{-1}$, respectively. Together with 18 joint torque servomotor measurements, the robot has $M = 23$ sensory modalities in total. The gait period is set to $T = 1.25\text{s}$, where we track the gait phase with the granularity $C = 6$.

The **simulated robot** is a high-fidelity model of the small six-legged walking robot PhantomX [25] in CoppeliaSim [26]. The robot has a similar morphology to Daisy, and therefore, the same sensorimotor dimensions and symbol semantics are used, but the granularity is $C = 4$.

Two navigation strategies are considered to generate reference \mathbf{y}^{ref} the robots learn to track. Given the robot’s position \mathbf{I} and heading direction α , the strategy steers the robot toward the goal location \mathbf{I}^* . The first navigation strategy is `turngo`, which is used for controller training. It maintains the robot’s

heading toward \mathbf{I}^* and forward heading velocity

$$\mathbf{I}^\Delta = \mathbf{I}^* - \mathbf{I}, \quad (8)$$

$$\alpha^* = \arctan 2(\mathbf{I}_2^\Delta, \mathbf{I}_1^\Delta), \quad (9)$$

$$\alpha^\Delta = \alpha^* - \alpha, \quad (10)$$

$$y_{\text{head}}^{\text{ref}}(t) = \|\mathbf{I}^\Delta\| \cos(\alpha^\Delta), \quad (11)$$

$$y_{\text{yaw}}^{\text{ref}}(t) = \alpha^\Delta \bmod 2\pi, \quad (12)$$

where the reference velocity and heading values are limited to be within the intervals $[0, y_{\text{planar}}^{\text{bound}}]$ and $[-y_{\text{turn}}^{\text{bound}}, y_{\text{turn}}^{\text{bound}}]$, respectively. All other modalities have a reference set to zero.

The second strategy `anygo` is to evaluate trained controller generalization. It uses a combination of side and forward locomotion to reach the goal location

$$y_{\text{head}}^{\text{ref}}(t) = \|\mathbf{I}^\Delta\| \cos(\alpha^\Delta), \quad (13)$$

$$y_{\text{side}}^{\text{ref}}(t) = \|\mathbf{I}^\Delta\| \sin(\alpha^\Delta), \quad (14)$$

with bounded planar velocities to be within $[0, y_{\text{planar}}^{\text{bound}}]$.

The proposed method is implemented in Python using linear regression and runs onboard on the Intel® Core™ i7-10710U processor, which provides sufficient computational power for real-time execution of Algorithm 1. The linear regression in bootstrapping, described in Section IV-B, is implemented with `sklearn` library, which produces a new model in less than a second. Hyperparameter values of $\mathcal{H} = (C, \sigma^y, \sigma^u, \sigma^x, \kappa^x, \kappa^u, s_{\text{head}}^{\text{ref}}, s_{\text{side}}^{\text{ref}}, s_{\text{yaw}}^{\text{ref}}, s_{\text{other}}^{\text{ref}}, y_{\text{planar}}^{\text{bound}}, y_{\text{turn}}^{\text{bound}})$ have been found in preliminary deployments of the real $\mathcal{H}_{\text{real}} = (6, 0.1, 100, 100, 0.02, 0.02, 0.03, 0.03, 0.02, 1, 0.02, 0.03)$ and simulated robot

$$\mathcal{H}_{\text{sim}} = (4, 0.1, 100, 100, 0.02, 0.01, 0.03, 0.03, 0.03, 1, 0.03, 0.12).$$

A. Real Robot Deployment

The real robot Daisy was trained indoors to walk forward on a flat floor and then deployed outdoors on various terrains, see Fig. 3. During the WM bootstrapping on the flat floor, Daisy is tasked to reach 20 m far goal location in the initial heading direction using `turngo` strategy.

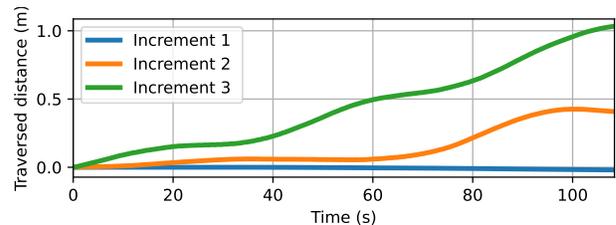


Fig. 2. Traversed distance toward the goal location increases after each bootstrapping increment.

It took three bootstrapping increments to achieve a forward motion. During the first increment, the robot slowly slides on the smooth floor forward and backward. During the second and third increments, the robot improves its forward motion; see Fig. 2. Still, in most experiments, the robot seems to exploit the floor’s smoothness and keeps some legs sliding on the floor, depicted in Figs. 3a and 4a. The WM $W = \{\mathbf{f}^{(1)}, \mathbf{f}^{(2)}, \mathbf{f}^{(3)}\}$ capable of the forward motion is bootstrapped in about 27 min of real world learning.

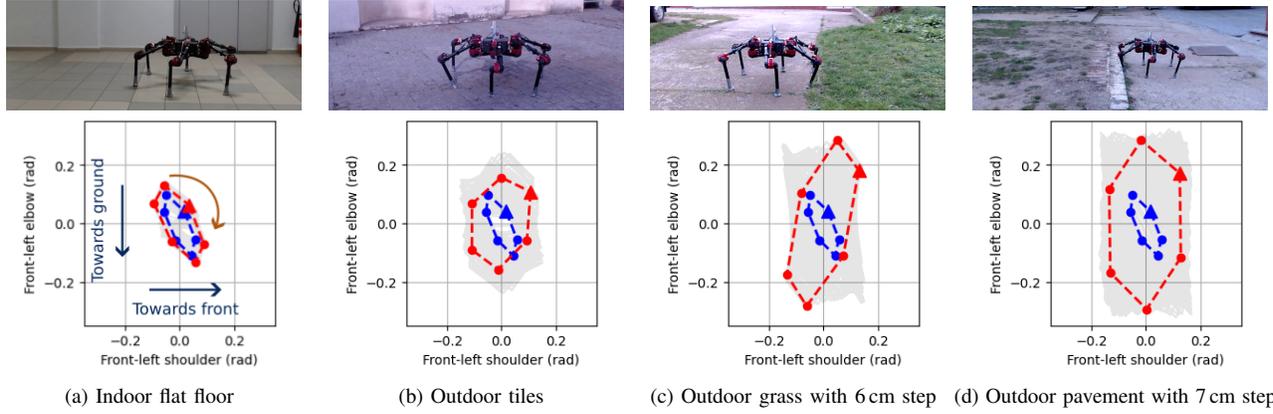
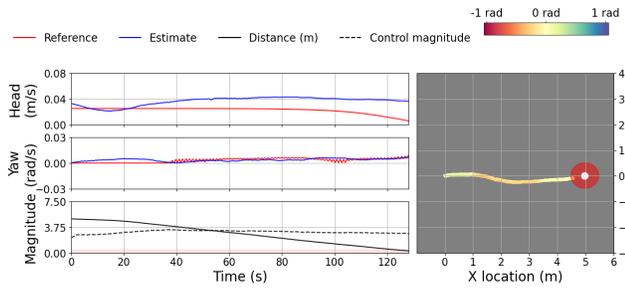
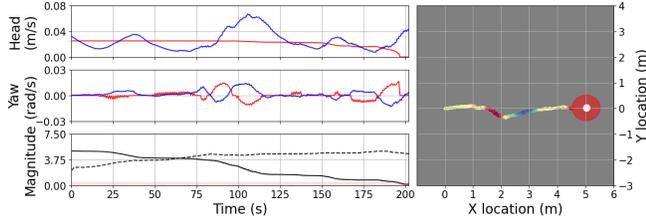


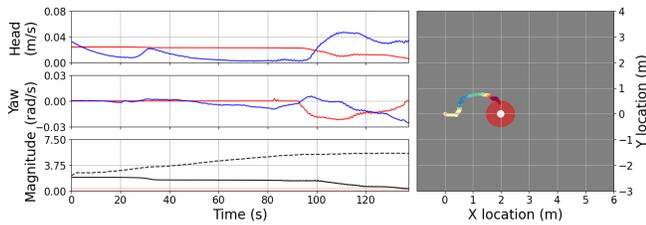
Fig. 3. Learning indoor and deployment outdoor scenarios. In the plots, the first (blue) and last (red) gait of the corresponding experiment, where the angle command \mathbf{u}_j sequence is projected on the middle-right leg shoulder and elbow plane. Each node represents a command at the phase segment executed in a clockwise direction, and the command \mathbf{u}_1 is at the triangular-shaped node. The recordings of the experiments can be found in the supplementary video.



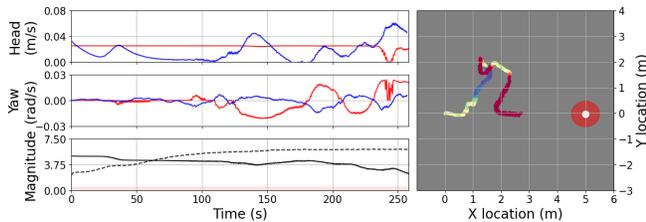
(a) Indoor environment used for WM bootstrapping in Fig. 3a.



(b) Outdoor tiles environment in Fig. 3b.

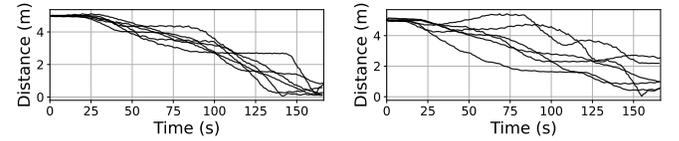


(c) Outdoor grass with a low step as shown in Fig. 3c.



(d) Outdoor pavement with a high step as shown in Fig. 3d.

Fig. 4. Multiple terrain traversal. The left column shows the evolution of tracking the reference, control magnitude $\|\mathbf{u}\|$, and distance from the goal location in meters. The right column shows the robot's navigation with the indicated heading angle by the color corresponding to the color bar on the top right. The translucent red circle indicates the goal region. In (c), the robot traverses a shorter distance of 2 m due to area constraints.



(a) turngo

(b) anygo

Fig. 5. Robot distance to the goal location in the evaluation with manually perturbed initial robot heading. In the first 30 s for each experiment, the largest heading perturbations in radians are (a) $\alpha^\Delta \in \{0.18, -0.08, 0.06, 0.10, 0.86, -0.86\}$, and (b) $\alpha^\Delta \in \{-0.37, 0.71, 1.14, 0.25, 0.50, 1.10\}$.

The WM bootstrapped indoor is used for Daisy deployment outdoors to reach goal location on tiles, concrete, and pavement terrains; see Figs. 3b to 3d. The outdoor terrains are not as smooth as the indoor floor, and the robot initially struggles with locomotion. However, as more reference error causes more updates on the gait, the leg lifting becomes pronounced, and the robot starts walking; see plots in Figs. 4b to 4d. For the final gaits, the legs are lifted higher for terrains with increased elevation differences; see Fig. 3.

The tiled and concrete terrains were used for testing the robot's ability to turn or go sideways with *turngo* and *anygo* navigation strategies, respectively. The robot was tasked to reach goal location 5 m in the initial heading direction and then manually turned, so the robot had to adjust its heading to the goal in the *turngo* navigation, see Fig. 5a. If the turn was slight, the robot adjusted its heading while moving forward. For a sharp turn, the robot slowed down and turned by shortening the stance trajectory of the forward leg on the side of the turn. The *anygo* was tested with the same methodology, and the robot resolved the task by diagonal motion. The *anygo* was slower than *turngo*, see Fig. 5b; however, the robot was of approaching the goal in both cases.

The random elevations on the tiled ground caused the robot to turn its heading from the goal location, yet it was capable of adjusting its heading to reach the goal location, see Fig. 4b. On the pavement, we observed the elevation transition between smooth and rough terrain, which is depicted in Figs. 4c and 4d.

B. Ablation Study using High-fidelity Simulated Robot

The robot model [25] is used in the ablation study with open/closed-loop setups using single/multiple models. The WM for the simulated robot is bootstrapped with three increments using the same methodology as for the real Daisy robot. The simulated robot with the trained WM is tasked to navigate toward the goal location $\mathbf{I}^* = (100\text{m}, 100\text{m})$ using the `turngo` strategy for 30 min. The controller robustness is tested in four *scenarios*, where between 3 min and 21 min of the test, the paralysis is applied to the left front, middle, hind, or none leg.

TABLE I
TRAVELED DISTANCE IN METERS FOR THE ABLATION STUDY.

Control	WM	Front / m	Middle / m	Hind / m	None / m
Open	Single	-0.4(0.0)	-1.8(0.1)	-0.7(0.0)	-0.3(0.1)
Closed	Single	7.0(2.1)	7.1(2.2)	8.1(1.7)	11.5(1.1)
Open	Multi	4.4(0.1)	12.8(0.3)	6.3(0.5)	9.6(0.3)
Closed	Multi	9.9(8.6)	12.3(4.7)	15.5(0.3)	15.9(0.3)

Average values with standard deviation in brackets among five trials.

The proposed closed-loop multi-model control (the very bottom row in Table I) is modified into three *ablations*: open-loop control, single model open-loop, and single model closed-loop. The **open-loop control** is implemented by setting the observation confidence small $\sigma^u = 1000$ and prediction confidence high $\sigma^y = 1$, while the **single model WM** is realized by using only the last model $\mathbf{f}^{(3)}$. Each ablation and scenario is repeated five times, and the average traveled distance (with standard deviation in brackets) is reported in Table I.

The single model open-loop performs the worst, and generally, the closed-loop control performs better than the open-loop control. Interestingly, multi-model open-loop performs better at middle leg paralysis, and overall, the results are close to the single-model closed-loop control, showing that even in the absence of sensory observation, the multi-model WM has the capability to approach the goal location. The multi-model outperforming the single-model is consistent with the previous work [12], where averaging control from multiple models improves the performance.

The ablation experiments for the parameters σ^u and σ^y show that while different values affect locomotion performance, the algorithm consistently sustains locomotion without failure. Moreover, σ^u , σ^y parametrize (6) across all M sensory modalities, indicating that the locomotion capability is robust to parameter variations. Such flexibility of tuning σ^u , σ^y can be used for automated parameter optimization during deployment, which we plan to explore in future work.

C. Model Interpretability

We argue for the interpretability because the models are linear, and the control is the outcome of the AIC, Bayesian inference [13]. In (1), the positive coefficient $w_{m,c}^{n,d}$ can be read as “ n th motor activity at the gait phase d is positively correlated with the m th sensor observation at the consequent

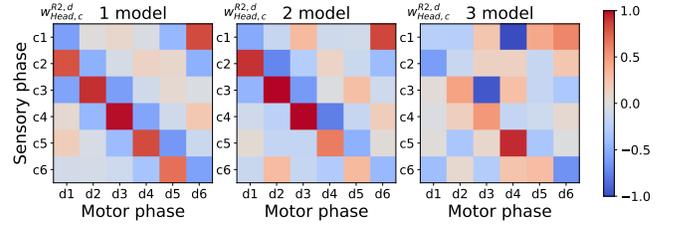


Fig. 6. The phase-phase relation between the right middle leg shoulder joint and heading velocity. The left model is phase-symmetrical, where positive motor activity results in positive heading velocity one phase segment later. The symmetry starts disappearing as the robot lifts its legs in the middle model, where the motor phases 4, 5 do not contribute to the heading velocity as much. The right model is phase-asymmetrical, where only the motor phase 4 contributes to the heading velocity of one phase segment.

phase c .” Following the interpretation, in Fig. 6, we can observe the phase-phase relation between the shoulder joint and forward speed changes $w_{\text{head},d}^{\text{R2},d}$. In the first model, the shoulder joint positive command increases the forward speed in the next phase $w_{\text{head},d+1}^{\text{R2},d} > 0$. However, in the third model, only the motor phase $d = 3$ has a strong sensory response because the leg is at the lowest point during $d = 3$, see Fig. 3, which implies the leg is in the stance. The strong sensory response $w_{\text{head},3}^{\text{R2},3}$ results in control update (7) changing the command at $d = 3$, $u_{\text{R2},3}$, the most.

D. Discussion on Biomimetic Approach

The proposed biomimetic method bootstraps the robot’s WM directly on site, capable of walking in 27 min from scratch. Although the model is linear and trained on a flat floor, the AIC inference is powerful enough to adapt to the rough terrain and adversarial scenarios, shown in Sections V-A and V-B. If the performance is insufficient, the incremental growth of the WM improves the performance [12], as we can see in Fig. 2. Furthermore, the proposed method is interpretable, as demonstrated in Section V-C; thus, besides providing control, the method also provides an analysis of the body-terrain dynamics. It contrasts with black-box controllers, which rely on high-fidelity models and deep machine learning methods [27] that are unsuitable for deployment in uncertain, ever-changing environments without extensive computational infrastructure.

VI. CONCLUSION

We present a novel learnable gait controller that bootstraps its locomotion and solves the Active Inference Control in the phase-embedded sensorimotor space. The embedding allows modeling dynamics by a linear transformation further used by the active inference of the state estimation and motor control. The method is deployed on real and simulated hexapod walking robots and independently tested in real-to-real and sim-to-sim adversarial scenarios. The robot self-learns gait dynamics and gains the capability to walk even under perturbed locomotion. In future work, we plan to focus on the continual scalability of the robot’s locomotion knowledge to address long-term deployments of the mobile platforms.

REFERENCES

- [1] K. J. Friston, J. Kilner, and L. Harrison, “A free energy principle for the brain,” *Journal of Physiology-Paris*, vol. 100, pp. 70–87, 2006. doi: [10.1016/j.jphysparis.2006.10.001](https://doi.org/10.1016/j.jphysparis.2006.10.001)
- [2] T. Taniguchi, S. Murata, M. Suzuki, D. Ognibene, P. Lanillos, E. Ugur, L. Jamone, T. Nakamura, A. Ciria, B. Lara, and G. Pezzulo, “World models and predictive coding for cognitive and developmental robotics: frontiers and challenges,” *Advanced Robotics*, vol. 37, no. 13, pp. 780–806, 2023. doi: [10.1080/01691864.2023.2225232](https://doi.org/10.1080/01691864.2023.2225232)
- [3] P. Lanillos and G. Cheng, “Adaptive robot body learning and estimation through predictive coding,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018. doi: [10.1109/IROS.2018.8593684](https://doi.org/10.1109/IROS.2018.8593684) pp. 4083–4090.
- [4] M. Baioumy, P. Duckworth, B. Lacerda, and N. Hawes, “Active inference for integrated state-estimation, control, and learning,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2021. doi: [10.1109/icra48506.2021.9562009](https://doi.org/10.1109/icra48506.2021.9562009) pp. 4665–4671.
- [5] C. Meo, G. Franzese, C. Pezzato, M. Spahn, and P. Lanillos, “Adaptation through prediction: Multisensory active inference torque control,” *IEEE Transactions on Cognitive and Developmental Systems*, vol. 15, no. 1, pp. 32–41, 2023. doi: [10.1109/tcds.2022.3156664](https://doi.org/10.1109/tcds.2022.3156664)
- [6] K. G. Pearson, “Central programming and reflex control of walking in the cockroach,” *Journal of Experimental Biology*, vol. 56, no. 1, pp. 173–193, 1972. doi: [10.1242/jeb.56.1.173](https://doi.org/10.1242/jeb.56.1.173)
- [7] A. J. Ijspeert, “Central pattern generators for locomotion control in animals and robots: A review,” *Neural Networks*, vol. 21, no. 4, pp. 642–653, 2008. doi: [10.1016/j.neunet.2008.03.014](https://doi.org/10.1016/j.neunet.2008.03.014)
- [8] A. J. Ijspeert, A. Crespi, and J.-M. Cabelguen, *Neuroinformatics*.
- [9] S. Gay, J. Santos-Victor, and A. J. Ijspeert, “Learning robot gait stability using neural networks as sensory feedback function for central pattern generators,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2013. doi: [10.1109/IROS.2013.6696353](https://doi.org/10.1109/IROS.2013.6696353) pp. 194–201.
- [10] R. Héliot and B. Espiau, “Multisensor input for cpg-based sensory—motor coordination,” *IEEE Transactions on Robotics*, vol. 24, no. 1, pp. 191–195, 2008. doi: [10.1109/TRO.2008.915433](https://doi.org/10.1109/TRO.2008.915433)
- [11] M. Thor and P. Manoonpong, “Versatile modular neural locomotion control with fast learning,” *Nature Machine Intelligence*, vol. 4, no. 2, pp. 169–179, 2022. doi: [10.1038/s42256-022-00444-0](https://doi.org/10.1038/s42256-022-00444-0)
- [12] R. Szadkowski, M. S. Nazeer, M. Cianchetti, E. Falotico, and J. Faigl, “Bootstrapping the dynamic gait controller of the soft robot arm,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2023. doi: [10.1109/icra48891.2023.10160579](https://doi.org/10.1109/icra48891.2023.10160579) pp. 2669–2675.
- [13] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, “Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,” *Information Fusion*, vol. 58, pp. 82–115, June 2020. doi: [10.1016/j.inffus.2019.12.012](https://doi.org/10.1016/j.inffus.2019.12.012)
- [14] A. Büschges and J. Schmidt, “Neuronal control of walking: studies on insects,” *e-Neuroforum*, vol. 6, no. 4, pp. 105–112, 2015. doi: [10.1007/s13295-015-0017-8](https://doi.org/10.1007/s13295-015-0017-8)
- [15] S. N. Markin, A. N. Klishko, N. A. Shevtsova, M. A. Lemay, B. I. Prilutsky, and I. A. Rybak, “Afferent control of locomotor cpg: insights from a simple neuromechanical model,” *Annals of the New York Academy of Sciences*, vol. 1198, p. 21, 2010. doi: [10.1111/j.1749-6632.2010.05435.x](https://doi.org/10.1111/j.1749-6632.2010.05435.x)
- [16] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences*, ser. Cambridge Non-linear Science Series. Cambridge University Press, 2001. doi: [10.1017/CBO9780511755743](https://doi.org/10.1017/CBO9780511755743)
- [17] X. Liu, R. Gasoto, Z. Jiang, C. Onal, and J. Fu, “Learning to locomote with artificial neural-network and cpg-based control in a soft snake robot,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020. doi: [10.1109/iros45743.2020.9340763](https://doi.org/10.1109/iros45743.2020.9340763) pp. 7758–7765.
- [18] T. Yan, A. Parri, V. R. Garate, M. Cempini, R. Ronsse, and N. Vitiello, “An oscillator-based smooth real-time estimate of gait phase for wearable robotics,” *Autonomous Robots*, vol. 41, no. 3, pp. 759–774, 2017. doi: [10.1007/s10514-016-9566-0](https://doi.org/10.1007/s10514-016-9566-0)
- [19] M. Desmurget, C. Epstein, R. Turner, C. Prablanc, G. E. Alexander, and S. T. Grafton, “Role of the posterior parietal cortex in updating reaching movements to a visual target,” *Nature Neuroscience*, vol. 2, no. 6, pp. 564–567, 1999. doi: [10.1038/9219](https://doi.org/10.1038/9219)
- [20] C. L. Buckley, C. S. Kim, S. McGregor, and A. K. Seth, “The free energy principle for action and perception: A mathematical review,” *Journal of Mathematical Psychology*, vol. 81, pp. 55–79, 2017. doi: [10.1016/j.jmp.2017.09.004](https://doi.org/10.1016/j.jmp.2017.09.004)
- [21] C. Pezzato, R. Ferrari, and C. H. Corbato, “A novel adaptive controller for robot manipulators based on active inference,” *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2973–2980, 2020. doi: [10.1109/lra.2020.2974451](https://doi.org/10.1109/lra.2020.2974451)
- [22] R. Szadkowski and J. Faigl, “Hexapod gait control through internal model belief update,” in *International Symposium on Adaptive Motion of Animals and Machines*, 2023. doi: [10.18910/92290](https://doi.org/10.18910/92290) pp. 107–108.
- [23] M. Pitchai, X. Xiong, M. Thor, P. Billeschou, P. L. Mailänder, B. Leung, T. Kulvicius, and P. Manoonpong, “Cpg driven rbf network control with reinforcement learning for gait optimization of a dung beetle-like robot,” in *International Conference on Artificial Neural Networks (ICANN)*, 2019. doi: [10.1007/978-3-030-30487-4_53](https://doi.org/10.1007/978-3-030-30487-4_53) pp. 698–710.
- [24] K. Friston, J. Mattout, N. Trujillo-Barreto, J. Ashburner, and W. Penny, “Variational free energy and the laplace approximation,” *NeuroImage*, vol. 34, no. 1, pp. 220–234, 2007. doi: [10.1016/j.neuroimage.2006.08.035](https://doi.org/10.1016/j.neuroimage.2006.08.035)
- [25] M. T. Nguyenová, P. Čížek, and J. Faigl, “Modeling proprioceptive sensing for locomotion control of hexapod crawling robot in robotic simulator,” in *2018 Modelling and Simulation for Autonomous Systems (MESAS)*, 2019. doi: [10.1007/978-3-030-14984-0_17](https://doi.org/10.1007/978-3-030-14984-0_17) pp. 215–225.
- [26] E. Rohmer, S. P. N. Singh, and M. Freese, “Coppeliassim (formerly V-REP): a versatile and scalable robot simulation framework,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2013. doi: [10.1109/IROS.2013.6696520](https://doi.org/10.1109/IROS.2013.6696520) pp. 1321–1326.
- [27] J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, “Learning quadrupedal locomotion over challenging terrain,” *Science Robotics*, vol. 5, no. 47, p. eabc5986, Oct. 2020. doi: [10.1126/scirobotics.abc5986](https://doi.org/10.1126/scirobotics.abc5986)