

Next Steps: Learning a Disentangled Gait Representation for Versatile Quadruped Locomotion

Alexander L. Mitchell, Wolfgang Merkt, Mathieu Geisert, Siddhant Gangapurwala, Martin Engelcke, Oiwi Parker Jones, Ioannis Havoutis, and Ingmar Posner

Abstract—Quadruped locomotion is rapidly maturing to a degree where robots now routinely traverse a variety of unstructured terrains. However, while gaits can be varied typically by selecting from a range of pre-computed styles, current planners are unable to vary key gait parameters *continuously* while the robot is in motion. The synthesis, *on-the-fly*, of gaits with unexpected operational characteristics or even the blending of dynamic manoeuvres lies beyond the capabilities of the current state-of-the-art. In this work we address this limitation by learning a latent space capturing the key stance phases of a particular gait, via a generative model trained on a single trot style. This encourages disentanglement such that application of a *drive signal* to a single dimension of the latent state induces holistic plans synthesising a continuous variety of trot styles. In fact properties of this drive signal map directly to gait parameters such as cadence, footstep height and full stance duration. The use of a generative model facilitates the detection and mitigation of disturbances to provide a versatile and robust planning framework. We evaluate our approach on a real ANYmal quadruped robot and demonstrate that our method achieves a continuous blend of dynamic trot styles whilst being robust and reactive to external perturbations.

I. INTRODUCTION

Quadruped locomotion has advanced significantly in recent years, extending its capability towards applications of significant value to industry and the public domain. Driven primarily by advances in optimisation-based [1]–[4] and reinforcement learning-based methods [5]–[7], quadrupeds are now able to traverse over a wide variety of terrains, making them a popular choice for tasks such as inspection, monitoring, search and rescue or goods delivery in difficult, unstructured environments. However, despite recent advances, important limitations remain. Due to the complexity of the system, models used for gait planning and control are often overly simplified and handcrafted for particular gait types (e.g. crawl, trot, gallop, e.g. [1], [8]). In the worst case, this can limit the versatility of the robot as the models deployed are failing to exploit the full capability of the underlying hardware (e.g. [1], [9], [10]). Furthermore, current gait planners are unable to vary key parameters (such as step height) of a gait continuously and online [3], [4]. Instead, contact schedules for discrete gaits are often pre-computed and selected as needed using hierarchical approaches that

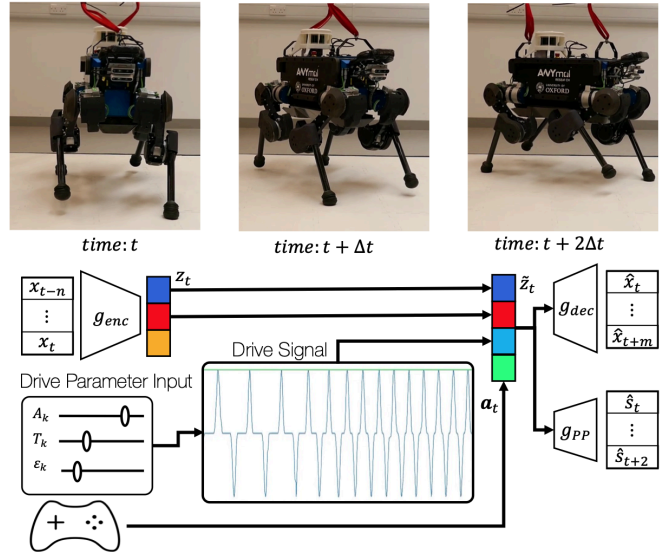


Fig. 1: Using a variational auto-encoder (VAE), our approach learns a structured latent space capturing key stance phases constituting a particular gait. The space is disentangled to a degree such that application of a *drive signal* to a single dimension of the latent variable induces gait styles which can be seamlessly interpolated between. We encode raw sensor information to infer the robot’s gait phase using g_{enc} before applying our drive signal and then decode the augmented latent variable and the base twist action \mathbf{a}_k via g_{dec} and predict the feet in contact using g_{pp} . The drive signal’s amplitude and phase provide continuous control over the robot’s cadence, full-support duration and foot swing height.

rely upon a sequence of optimisations, yielding only a narrow operating window which severely limits their response to external disturbances [3], [4]. The synthesis of gaits which fall outside of this set is typically cumbersome, as approaches that optimise dynamics over gait schedules and footstep lengths or heights are computationally expensive [3], [4], and often require manual intervention [1]. In contrast, the ability to control key gait parameters – such as cadence, swing height, and full-support duration – *on-the-fly* would enable a smooth interpolation between dynamic manoeuvres allowing for swift reaction to external stimuli.

Inspired by recent work on a quadruped that achieves a crawl gait via the traversal of a learned latent space [11], we approach the challenge of continuous contact-schedule variation from the perspective of learning and traversing a structured latent-space. This is enabled by learning a *genera-*

All authors are with the Oxford Robotics Institute, University of Oxford
Email: mitch@robots.ox.ac.uk

This work was supported by a UKRI/EPSCRC Programme Grant [EP/V000748/1], the EPSRC grant ‘Robust Legged Locomotion’ [EP/S002383/1], the EPSRC CDT [EP/L015897/1], the UKRI/EPSCRC RAIN [EP/R026084/1] and ORCA [EP/R026173/1] Hubs and the EU H2020 Project MEMMO (780684). It was conducted as part of ANYmal Research, a community to advance legged robotics.

tive model of locomotion data which, in addition to capturing relevant structure in the space, enables the detection and mitigation of disturbances to provide a versatile and robust planning framework. In particular, we train a variational auto-encoder (VAE) [12], [13] on short sequences of state-space trajectories taken from a single gait type (trot), and predict a set of future states. We show that the resulting latent space has an interpretable structure, lending itself to the generation of a variety of trot styles, depending on how the latent space is traversed. In fact, examining trot trajectories in latent space reveals an oscillatory drive-signal, and we subsequently find that by overwriting this trajectory with a synthetic drive signal, we can *continuously* control the robot’s gait properties whilst the robot is executing the motion. Parameters of this drive signal can be mapped explicitly to gait parameters such as cadence, footstep height, and full-stance support duration.

We illustrate the efficacy of our approach by generating a range of continuously blended trajectories on a real ANYmal quadruped robot – a medium-sized platform (35) standing 0.5 tall. While the latent space is learnt using examples only from a specific gait style, our approach is able to synthesise behaviours significantly beyond this training distribution.

In addition, we leverage our generative approach to both characterise and react to external perturbations. A large impulse applied to the robot’s base triggers a spike in the Evidence Lower Bound (ELBO) which identifies the disturbance as out of the observed training distribution. Inspired by [14], which states that an increase in cadence is a response to slippages in humans, our planner automatically increases the robot’s cadence to aid in counteracting the disturbance and demonstrate a marked improvement in robustness.

To the best of our knowledge, our method is the first which allows continuous variation of the robot’s gait characteristics whilst the robot is walking. It provides a versatile and data-driven approach to quadruped locomotion which additionally allows for disturbance detection and recovery. Beyond these contributions, our results are tantalising as the notion of the drive signal is reminiscent of central pattern generators (CPGs) located in the spinal cord and brain stem of vertebrates [15]. Both CPGs and our synthetic drive signal generate cyclic patterns that effect holistic motor activity and stereotyped rhythmic behaviours such as locomotion.

II. RELATED WORK

Planning and control for quadruped locomotion has advanced in leaps and bounds in recent years. A seminal work in these areas is *Dynamic Gaits* (DG) [1]. DG enables a quadrupedal robot (ANYmal) to execute a wide variety of dynamic gaits (e.g. trot, pace, lateral walk, jump) with real-time motion planning and control. However, to achieve this impressive range of behaviours, DG provides each gait type with its own contact schedule and utilises an environment-specific footstep planner, ultimately limiting its capability.

Latent space approaches for planning and control learn useful and typically low-dimensional representations that can be used to control complex dynamics, without relying on known system models. Classic examples include *Deep*

Variational Bayes Filters (DVBF) [16] and *Embed to Control* (E2C) [17]. *Conditional Neural Movement Primitives* (CNMP) [18] is a more recent latent space approach for robotic arms that generalises between a variety of tasks, such as pick-and-place and obstacle avoidance. Other recent works like UPN and PlaNet [19], [20] show impressive capabilities in simulation but are yet to be applied to real-world systems, including floating-base robots.

In the quadruped domain, *First Steps* [11] learns a structured latent-space based on feasible robot *configurations* and defines a set of performance predictors that are used in an optimisation framework to control the robot. In practice, these performance predictors can be viewed as symbolic inputs (e.g. ‘left front leg up’) but drive the robot in continuous space. Since *First Steps* is trained on static snapshots of robot configurations, it does not learn from observable dynamics and may thus require more explicit structuring of the system than is necessary. Our work addresses this shortcoming and significantly extends this framework to an effective and robust closed-loop planning and control.

The *Motion VAE* (MVAE) [21] learns to represent dynamic trajectories in a structured latent-space for the locomotion of animated humanoids. This is similar to our emphasis of learning representations for dynamic trajectories in the context of locomotion. However, moving from simulated to real physical systems, as is required for robotic applications, necessitates tackling additional complexities like latency, hard real-time requirements, and actuator dynamics. In this work, we tackle these challenges and demonstrate that a single gait style contains sufficient richness to learn a structured latent space that can be exploited to manipulate gait characteristics. Unlike MVAE, our approach does not train on multiple gait styles, despite succeeding in producing them.

III. METHOD

Our aim is to use unsupervised learning to infer a structured latent-space which facilitates real-time and smooth variation of key gait parameters. We conjecture that structure can emerge from the exposure of a suitable generative model to a gait with predetermined and constant characteristics such as cadence, swing height, and full-support duration. We discover a disentangled latent-space where gait parameters are axis-aligned within the structured latent-space. Periodic trajectories in this space can be decoded leading to smooth robot locomotion as depicted in Fig. 1.

VAE Architecture: We train a VAE [12], [13] to create a structured latent-space using observed dynamic data. The input \mathbf{X}_k at time step k consists of N robot states sampled from simulated trot gaits with constant parameters (e.g. cadence and foot-step height). These state-space quantities are values we wish either to control or help to infer the gait phase. These are joint angles; end-effector positions in the base frame; joint torques; contact forces; the base velocity; and the base pose evolution relative to a control frame, which is updated periodically. These quantities are denoted as $\mathbf{x}_k = [\mathbf{q}_k, \mathbf{e}\mathbf{e}_k, \tau_k, \lambda_k, \dot{\mathbf{c}}_k, \Delta\mathbf{c}_k]$. Note that velocities and accelerations do not form part of the VAE’s input, but are

inferred from the input history. This prevents sensitivity to fast-changing quantities such as recorded joint accelerations during inference.

During deployment on the robot, we encode \mathbf{X}_k at the control frequency f_c . However, we construct this input using states \mathbf{x}_k sampled at a frequency f_{enc} lower than f_c :

$$\mathbf{X}_k = [\mathbf{x}_{k-r(N-1)}^\top, \dots, \mathbf{x}_{k-r}^\top, \mathbf{x}_k^\top], \quad (1)$$

where $r = f_c/f_{enc}$ is the ratio between the control and encoder frequencies. Next, the VAE’s decoder output $\hat{\mathbf{X}}_k^+$ predicts the current robot state \mathbf{x}_k as well as M future ones sampled at a frequency of $f_{dec} = f_c$:

$$\hat{\mathbf{X}}_k^+ = [\hat{\mathbf{x}}_k^\top, \hat{\mathbf{x}}_{k+1}^\top, \dots, \hat{\mathbf{x}}_{k+M}^\top] \quad (2)$$

As the *desired-feet-in-contact* \mathbf{s}_k is an input to the tracking controller, we also want to predict this at the current time step, and J steps in the future. Inspired by *First Steps* [11], we therefore utilise a feet-in-contact performance predictor $g_{pp}(\mathbf{z}_k)$. This is attached to the latent space, and estimates the probability of each foot being in contact:

$$\hat{\mathbf{S}}_k = [\hat{\mathbf{s}}_k^\top, \dots, \hat{\mathbf{s}}_{k+J-1}^\top]^\top \quad (3)$$

To command the base twist of the robot, a high-level action \mathbf{a}_k is utilised. This represents longitudinal (x), lateral (y), and yaw (θ) twist in the robot’s base frame. The latent state \mathbf{z}_k and the action \mathbf{a}_k form the input to the decoder.

Training the VAE: We train the VAE and performance predictor together. The VAE’s training loss is the standard ELBO formulation consisting of a reconstruction loss (mean-squared error) plus the Kullback–Leibler (KL) divergence D_{KL} between the inferred posterior $q(\mathbf{z}|\mathbf{X}_k)$ and the prior $p(\mathbf{z})$, multiplied by a hyper-parameter β :

$$\mathcal{L}_{\text{ELBO}} = \text{MSE}(\mathbf{X}_k^+, \hat{\mathbf{X}}_k^+) + \beta D_{\text{KL}}[q(\mathbf{z}|\mathbf{X}_k)||p(\mathbf{z})]. \quad (4)$$

These ELBO terms are then summed with the scaled binary cross-entropy loss between the predicted feet in contact and the recorded ones, such that the overall loss is computed as

$$\mathcal{L} = \mathcal{L}_{\text{ELBO}} + \gamma \text{BCE}(\mathbf{S}_k, \hat{\mathbf{S}}_k) \quad (5)$$

The VAE training loss (Eq. 4), as seen in prior work [22], is responsible for any subsequent disentanglement found in the latent space. The reconstruction error is weighed against the decomposition of the latent space using the hyper-parameter β . This constraint encourages an efficient latent representation, containing only the required information for reconstruction, whilst also reducing the channel capacity. As shown in [22], the D_{KL} term used with an isotropic unit Gaussian ($p(\mathbf{z}) = \mathcal{N}(\mathbf{0}, \mathbf{I})$) encourages conditional independence within \mathbf{z} .

In our approach, as well as that of *First Steps* [11], a structured latent space is encouraged by backpropagating gradients from the performance predictor’s loss through to the encoder input. We hypothesise that useful structuring of this space is inferred from the continuous trajectories used for training, and, in contrast to *First Steps*, no explicit labelling for each stance is required.

A. Investigating the Latent Space

Before solving for locomotion trajectories by planning in latent space, we examine the space to see what structure exists within. Disentangled features are discovered inside.

Latent Space Structure: Fig. 2 shows samples from the latent space colour-coded by their predicted stance. For trot there are four stances: the full-support phase, left front and right hind in contact, another support phase and finally right front plus left hind in contact. Fig. 2 reveals that the latent space has emerged clustered by stance and that, due to the ordering of these stances, a periodic trajectory decodes to a trot gait. This favourable structure is inferred from the continuous trot trajectory input during training.

The Effect of Disentanglement: By examining the latent variables, we discover that oscillations injected into just two dimensions in the latent space decode to continuously varying trot trajectories. This result stems from a latent space where variation in footstep length and cadence is aligned along one dimension, while variation in footstep distance lies along another. The consequence of this disentanglement is that two oscillations (one being $\pi/2$ out of phase with the other) injected into the latent-space is sufficient to produce a broad range of continuously variable trot-trajectories. Given that other work has tried to explicitly build this structure into locomotion systems [23], it is important to emphasise that this disentanglement *emerges* in our study as a result of the training paradigm and data.

B. Control Over the Gait Parameters

The disentangled latent-space is exploited such that the cadence, footstep height, and full-support duration can be controlled online. A specific *drive signal* overwrites a single latent variable, and in this case, we utilise a modified \sin^3 oscillation with commandable amplitude A_k and phase ϕ_k :

$$\mathbf{z}_{k,d_z} = A_k \sin^3(\phi_k). \quad (6)$$

The amplitude A_k controls the foot swing height, and the phase ϕ_k governs cadence and support duration. To control the robot’s swing and stance duration separately, we set the drive signal’s time-period T_k and we employ a stance counter ϵ_k . The time-period T_k is equal to the swing duration, whilst the time that the drive signal is equal to zero is extended by ϵ_k time steps to introduce a full-support duration of $[\text{parse-numbers} = \text{false}, \text{number-math-rm} =](\epsilon_k/f_c)$. Hence, *onceboth* T_k and ϵ_k are used together, the phase dynamics are:

$$\phi_{k+1} = \begin{cases} \phi_k & \text{if } \phi_k \bmod \pi = 0 \text{ and } k_\epsilon < \epsilon_k \\ \phi_k + 2\pi/T_k & \text{otherwise} \end{cases} \quad (7)$$

and, in tandem, the counter ϵ_k is updated as:

$$k_\epsilon \leftarrow \begin{cases} k_\epsilon + 1 & \text{if } \phi_k \bmod \pi = 0 \text{ and } k_\epsilon < \epsilon_k \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

C. Planning for Closed-Loop Control

The VAE is fast enough to act as a planner in a closed-loop controller. Thus, our approach can react to external disturbances and mitigate against real-world effects such as unmodelled dynamics and hardware latency. For closed-loop control, we begin by encoding a history of robot states from the raw sensor measurements to infer the current gait phase. We store a buffer of past robot states and sample from this at f_{enc} to create the encoder’s input.

With an estimate of the current latent variable, we overwrite latent dimension d_z with the drive signal (see Sec. III-B). Next, we employ a second-order Butterworth filter to smooth the latent trajectory and further smooth the locomotion plan. In essence, the drive signal encourages the decoder to output the next open-loop prediction while the other latent variables infer the gait phase from the raw sensor input. This process is repeated at the control frequency (400).

The latent variable \mathbf{z}_k and a desired base twist \mathbf{a}_k are decoded to produce $\hat{\mathbf{X}}_k^+ = g_{dec}(\mathbf{z}_k, \mathbf{a}_k)$. From this, the joint-space trajectory $\hat{\mathbf{Q}}_k$, and local base velocity $\hat{\mathbf{C}}_k$ are extracted and derived or integrated to produce the base and joint positions, velocities and accelerations. These quantities and the predicted contact schedule are sent to the Whole-Body Controller (WBC) [24]. The WBC solves a hierarchical optimisation problem to calculate the joint torques which are sent to the actuators. The series of constraints enforced by the WBC are: contact creation, friction constraints and torque limits. Next, the WBC applies forward kinematics to the VAE’s trajectory to track it in task-space. Note that the WBC does not compensate for infeasible plans.

Disturbance Detection and Response: Our approach is able to both detect and react to disturbances. The VAE is trained using canonical feasible trajectories. Therefore, any disturbances are characterised as out of distribution with respect to the training set. Given the generative nature of our approach, this discrepancy is quantified during operation by the trained model via the Evidence Lower-Bound (ELBO, Eq. 4 where β is set to one). We will show in the evaluation (Sec. IV-E) that even a rudimentary response strategy serves to increase the range of disturbance the system can reject.

IV. EXPERIMENTAL RESULTS

We demonstrate the capabilities of the VAE as a flexible and robust locomotion-planner deployed for closed-loop control on a real robot. In doing so, we investigate (i) the structure induced in the latent space (Sec. IV-B), (ii) the sensitivity of our approach to variations in key hyper parameters (Sec. IV-C), (iii) to what extent the locomotion parameters can be varied online (Sec. IV-D), and finally, (iv) the degree to which disturbance detection, coupled with a rudimentary recovery strategy, further increases the robustness of our approach (Sec. IV-E). Please see the following video for an extended set of experiments along with a brief description of our approach (<https://youtu.be/2XMyAAe7yDs>).

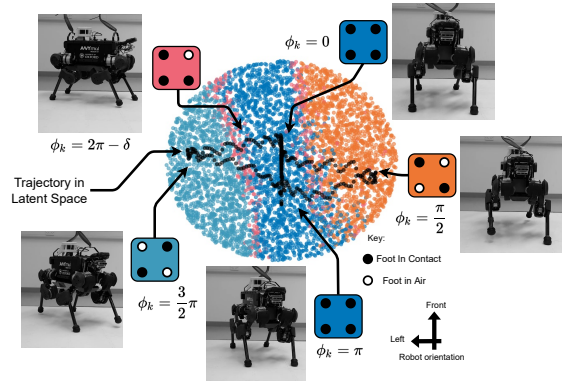


Fig. 2: A slice through the structured latent space, colour coded to illustrate the ordering and clustering of trot gait stances. The horizontal displacement of the latent-space trajectory (black) contributes to the robot’s footstep height, whilst the vertical movement relates to the footstep length. Snapshots of the robot controlled using the VAE-planner illustrate the inter-play between these two latent dimensions.

A. Details on Training and Deployment

Before incorporating our VAE-planner into a closed-loop controller and evaluating the results as deployed on the real ANYmal robot, we discuss the following: firstly, how the dataset is constructed; secondly, how the VAE is trained; and lastly, implementation details for transfer to the real robot.

Dataset Generation: To train the VAE, we require a set of continuous trot trajectories, and use *Dynamic Gaits* (DG) [1] for this. DG is a hierarchical planning and control framework using a *fixed* contact schedule and predefined footstep heights. Specifically, swing, full-stance durations and footstep height are set to 0.5, 75 and 0.10. The components of DG are a footstep planner, a base motion planner and a WBC. The footstep planner computes the next four steps over the gait period using an inverted-pendulum model. Using the footstep positions and schedule, the base motion planner computes the base trajectory over a gait period using a centroidal dynamics model [10] constrained with a Zero-Moment Point (ZMP) [9] criterion. The WBC is described in Sec. III-C and converts the task space trajectories to joint torques, positions and velocities, which are sent to the actuators. The dataset is generated by uniformly sampling desired base twist and executing DG in the *RaiSim* physics simulator [25]. To improve the fidelity of the simulation, the dynamics of the Series-Elastic Actuators (SEA) [26] are modelled using an *actuator network* [7] and accounts for the commanded positions, velocities, feed-forward torques, and low-level PD gains.

Architecture Details: In all our experiments, the VAE’s encoder, decoder and stance performance predictor have two hidden layers and widths of 256 units, using ELU nonlinearities [27]. The encoder input is created using $N = 80$ robot states sampled at 200 – representing a history of 0.4 – from the encoder input, which is of size 5120 units. The input is compressed via a latent space of 125 units which is concatenated with an action of 3 units. Next, the decoder

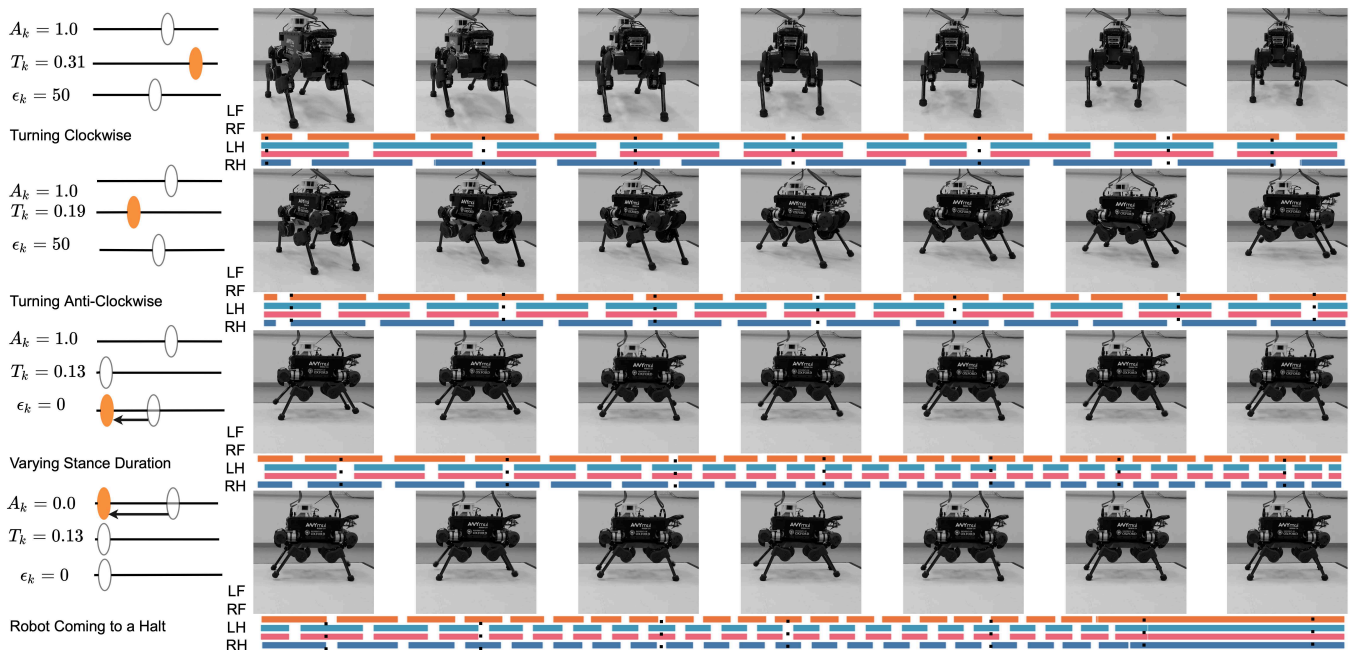


Fig. 3: Closed-loop control of the real ANYmal quadruped showcasing user-controlled variation of gait parameters on the fly. Coloured rectangles represent the full-stance phase, whilst white space denotes the swing duration. The top row shows a trot gait with an introduced quadrupedal stance phase (gait cycle of 0.75 – swing 312.5, stance 62.5; a gait cycle consists of a swing phase for each of the leg pairs). Next, the swing duration T_k is reduced using the time-period slider (gait cycle of 0.5 – swing 188, stance 62.5). The third row illustrates the effect of reducing the stance duration counter ϵ_k to produce trot without full-stance phases (gait cycle of 250 – swing 125, stance 0.0). Finally, transition into standing occurs when the drive signal amplitude A_k is reduced to zero.

outputs the current state and the next $M = 19$ robot states at the control frequency of 400 (preview horizon 47.5, output size: 1216 units). The performance predictor predicts the current feet in contact and two future states. Finally, hyperparameters used for training are $\beta = 1.0$, $\gamma = 0.5$, with a learning rate of $1e-3$ using the Adam optimiser. Training is terminated after $1e+6$ gradient steps.

Domain Transfer: To achieve domain transfer and deployment on the real robot, some modifications are required. The contact forces are artificially set to zero when the robot measures no contact. As mentioned in Sec. III-C, a Butterworth filter with a cutoff frequency of 10 smooths the latent trajectory. Due to the strict computation budget for real-time control, inference times for the VAE are restricted to at most 1, imposing constraints on the capacity of the model (see Sec. IV-C). Our largest model takes approximately 1 for the VAE computation, which is roughly equal to the computation time of the WBC.

B. Structure Induced In The Latent Space

The latent space is inspected to discover what structure exists and if any locomotion properties are disentangled within. Fig. 2 depicts the latent space and is plotted such that the drive signal dimension is along the horizontal, and trot signal dimension is aligned with the vertical. In addition, the stances (coloured areas) are distinctly clustered in sequence of the trot gait, whilst the user-controlled latent-space trajectory (black) moves through the space in a figure

of eight visiting each stance cluster sequentially. See Sec. III-A for detailed discussions.

C. Sensitivity To Hyper Parameters

The size of the latent dimension is reduced from 125 to a minimum of 6 in strides of 32 units. The latent dimension of 6 units deployed in simulation and the real robot produces trajectories which are stable and well tracked by the WBC. Next, the input history is reduced from 0.4 to 0.3 sampled at 200, resulting in poor open-loop performance. Since the gait’s swing duration in training lasts 0.45, an input history long enough to capture this is necessary to infer the gait phase. Finally, the VAE’s width is reduced in 32 unit increments to reduce the channel capacity until the VAE’s trajectories cause the robot to become unstable in simulation. This limit is reached at a width of 128 units, corresponding to a reduction in channel capacity by 52.8. We verified these models on the robot showcasing that a minimal model requires a latent space of 6 units, an input history of 0.4 sampled at 200, and a hidden layer size of 128 neurons.

D. Varying Locomotion Parameters Online

We leverage the disentangled latent-space to smoothly transition between gait parameters whilst the robot is walking. Crucially, cadence, stance duration, and footstep height can be varied during any phase of the gait by modulating the drive signal’s parameters (see Sec. III-B). This results in operating modes which vary from those seen during training.

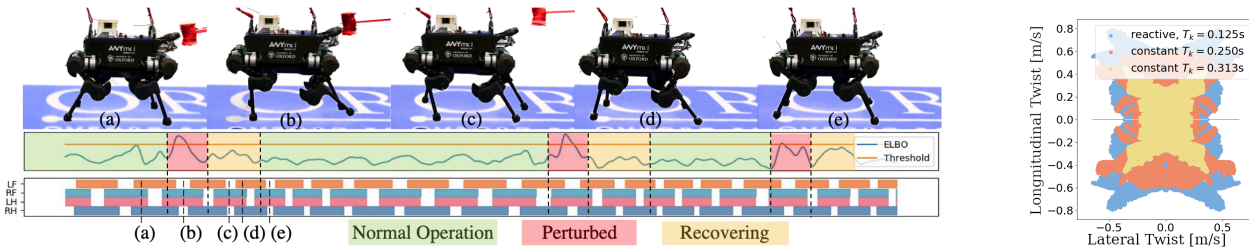


Fig. 4: Push recovery following a shove to the base. A disturbance to the robot causes the ELBO of the form in Eq. 4 to rise above a predetermined threshold (red areas). This identifies a disturbance and triggers an increase in cadence from 250 to 125 as a rudimentary response. This reduction in the swing duration is visible in the contact schedule. The right sub-plot shows how the range of rejectable base-velocity disturbances increases when automatically modulating the cadence. To appreciate the results fully, please see the following video <https://youtu.be/2XMyAAe7yDs>.

Examples of walking motion using the VAE-planner on the real robot are shown in Fig. 3.

Swing Duration: The swing duration is varied over a large operating window on the ANYmal robot. This begins with a swing time-period starting at 312.5, and is smoothly varied until the swing duration reaches 125, effecting a faster step rate. In parallel, we alter the robot’s heading, demonstrating the independence of the action and the latent-space dynamics. Specifically, the top row of Fig. 3 shows the nominal swing duration of 312.5 as the robot turns clockwise, tracking a constant angular velocity command. Next, we demand a slightly faster swing 188 and a constant angular velocity anti-clockwise, before transitioning to the fastest swing (125) in the third row. Here, the coloured contact schedule captures the changes in swing duration as it occurs in real-time.

Stance Duration: Following a successful reduction in cadence, the stance duration is reduced and the robot transitions into a trot with negligible full-stance phase: $\epsilon_k = 0$. Trot gaits with little to no full-support phase are particularly challenging manoeuvres for the system in general, as there is reduced control authority to correct for accrued base pose error. During the swing phase of this gait style, only feet across the diagonal are in contact resulting in a line contact limiting the robot’s ability to steady its base. The transition to a negligible full-stance phase is captured in the third row, where the coloured stance duration reduces in length.

Footstep Height: We vary the amplitude of the drive signal smoothly to zero as is seen in the bottom row (Fig. 3): The footstep height reduces to zero as the white-space in the contact schedule disappears and the robot remains standing. Beyond versatility, e.g. to increase swing heights to overcome irregular ground height, this capability further enables a safe, smooth and natural transition into and out of the VAE control mode (i.e. to start and come to a halt).

E. Disturbance Detection and Recovery

The ELBO is used to detect disturbances as it is a lower bound for the evidence of a sample given a particular distribution. The distribution in question is that learned over the training data from DG. Therefore, any motions which deviate from this due to a perturbation cause a large negative spike in the ELBO value. For simplicity we use ELBO in

the form of Eq. 4 meaning that large positive values are a result of a perturbation (see red areas in Fig. 4).

The VAE-planner is able to reject a wide range of impulses applied to the robot’s base. This operating window is enlarged by increasing the robot’s cadence once a disturbance is detected, see Fig. 4. This is a rudimentary response inspired by [14], which reports that humans increase their cadence to recover from slippage. The cadence is only increased if the ELBO value rises above a pre-selected threshold, below which is considered normal conditions. This value is 11.0 and found by operating the robot for a few minutes and recording the ELBO. The nominal cadence is set to 250 and halves to 125 for 1.50 when the threshold is surpassed.

Fig. 4 depicts the ELBO trace for three push events along with the robot’s contact schedule. The widths of the white spaces in the contact schedule half as the cadence increases to mitigate the disturbance. The robot successfully recovers and this usually requires between three and four steps.

V. CONCLUSION

We present a robust and flexible approach for locomotion planning via traversal of a structured latent-space. We utilise a deep generative model to capture features from locomotion data, and enable detection and mitigation of disturbances. The resulting latent-space is disentangled such that key locomotion features are automatically discovered from a single style of trot gait. This disentanglement is exploited using an oscillatory drive-signal, where the amplitude and phase directly control the gait parameters, namely the cadence, swing height, and full-support duration. Once deployed, the ease with which modulation of the drive signal gives rise to seamless interpolation between gait parameters is demonstrated. Utilising a generative model affords detection of disturbances as out of the distribution seen during training. Although the VAE-planner is able to reject a broad range of impulses, this window is broadened by increasing the cadence once the disturbance is detected – a rudimentary response inspired by human locomotion during slippage [14].

ACKNOWLEDGEMENTS

The authors acknowledge the use of the SCAN facility, and thank Oliver Groth for useful discussions.

REFERENCES

- [1] C. D. Bellicoso, F. Jenelten, C. Gehring, and M. Hutter, "Dynamic locomotion through online nonlinear motion optimization for quadrupedal robots," *IEEE Robot. Automat. Lett. (RA-L)*, vol. 3, no. 3, pp. 2261–2268, 2018.
- [2] C. Mastalli, W. Merkt, J. Marti-Saumell, H. Ferrolho *et al.*, "A direct-indirect hybridization approach to control-limited DDP," *arXiv:2010.00411*, 2021.
- [3] O. Melon, R. Orsolino, D. Surovik, M. Geisert *et al.*, "Receding-horizon perceptive trajectory optimization for dynamic legged locomotion with learned initialization," in *IEEE Int. Conf. Rob. Autom. (ICRA)*, 2021.
- [4] A. W. Winkler, C. D. Bellicoso, M. Hutter, and J. Buchli, "Gait and trajectory optimization for legged systems through phase-based end-effector parameterization," *IEEE Robot. Automat. Lett. (RA-L)*, vol. 3, no. 3, pp. 1560–1567, July 2018.
- [5] J. Hwangbo, J. Lee, A. Dosovitskiy, D. Bellicoso *et al.*, "Learning agile and dynamic motor skills for legged robots," *Science Robotics*, vol. 4, no. 26, 2019.
- [6] S. Gangapurwala, A. Mitchell, and I. Havoutis, "Guided constrained policy optimization for dynamic quadrupedal robot locomotion," *IEEE Robot. Automat. Lett. (RA-L)*, vol. 5, no. 2, pp. 3642–3649, 2020.
- [7] S. Gangapurwala, M. Geisert, R. Orsolino, M. Fallon, and I. Havoutis, "RLOC: Terrain-aware legged locomotion using reinforcement learning and optimal control," *arXiv preprint arXiv:2012.03094*, 2020.
- [8] A. W. Winkler, F. Farshidian, D. Pardo, M. Neunert, and J. Buchli, "Fast trajectory optimization for legged robots using vertex-based ZMP constraints," *IEEE Robot. Automat. Lett. (RA-L)*, vol. 2, no. 4, pp. 2201–2208, Oct 2017.
- [9] M. Vukobratovic and B. Borovac, "Zero-moment point - thirty five years of its life," *Int. J. Hum. Rob. (IJHR)*, vol. 1, pp. 157–173, 2004.
- [10] D. Orin, A. Goswami, and S.-H. Lee, "Centroidal dynamics of a humanoid robot," *Autonomous Robots*, vol. 35, 10 2013.
- [11] A. L. Mitchell, M. Engelcke, O. Parker Jones, D. Surovik *et al.*, "First steps: Latent-space control with semantic constraints for quadruped locomotion," in *IEEE/RSJ Int. Conf. Intell. Rob. Sys. (IROS)*, 2020, pp. 5343–5350.
- [25] J. Hwangbo, J. Lee, and M. Hutter, "Per-contact iteration method for solving contact dynamics," *IEEE Robot. Automat. Lett. (RA-L)*, vol. 3, no. 2, pp. 895–902, 2018.
- [12] D. Kingma and M. Welling, "Auto-encoding variational bayes," in *Int. Conf. on Learn. Repr. (ICLR)*, 2014.
- [13] D. J. Rezende, S. Mohamed, and D. Wierstra, "Stochastic backpropagation and approximate inference in deep generative models," in *Int. Conf. on Mach. Learn. (ICML)*, 2014.
- [14] B. E. Moyer, A. J. Chambers, M. S. Redfern, and R. Cham, "Gait parameters as predictors of slip severity in younger and older adults," *Ergonomics*, vol. 49, pp. 329–343, 2006.
- [15] I. Steuer and P. A. Guertin, "Central pattern generators in the brainstem and spinal cord: an overview of basic principles, similarities and differences," *Brain Research Reviews*, no. 2, pp. 107–164, 2019.
- [16] M. Karl, M. Soelch, J. Bayer, and P. van der Smagt, "Deep variational bayes filters: Unsupervised learning of state space models from raw data," in *Int. Conf. on Learn. Repr. (ICLR)*, 2016.
- [17] M. Watter, J. T. Springenberg, J. Boedecker, and M. A. Riedmiller, "Embed to control: A locally linear latent dynamics model for control from raw images," in *NeurIPS*, 2015.
- [18] M. Y. Seker, M. Imre, J. H. Piater, and E. Ugur, "Conditional neural movement primitives," in *Rob. Sci. Sys. (RSS)*, 2019.
- [19] A. Srinivas, A. Jabri, P. Abbeel, S. Levine, and C. Finn, "Universal planning networks," in *Int. Conf. on Mach. Learn. (ICML)*, 2018.
- [20] D. Hafner, T. Lillicrap, I. Fischer, R. Villegas *et al.*, "Learning latent dynamics for planning from pixels," in *Int. Conf. on Mach. Learn. (ICML)*, vol. 97, 2019, pp. 2555–2565.
- [21] H. Ling, F. Zinno, G. Cheng, and M. Panne, "Character controllers using motion VAEs," *ACM Trans. Graph.*, vol. 39, 07 2020.
- [22] I. Higgins, L. Matthey, A. Pal, C. Burgess *et al.*, " β -VAE: Learning basic visual concepts with a constrained variational framework," in *Int. Conf. on Learn. Repr. (ICLR)*, 2017.
- [23] Y. Yang, T. Zhang, E. Coumans, J. Tan, and B. Boots, "Fast and efficient locomotion via learned gait transitions," in *Conf. on Rob. Learn. (CoRL)*, 2021.
- [24] C. Dario Bellicoso, F. Jenelten, P. Fankhauser, C. Gehring *et al.*, "Dynamic locomotion and whole-body control for quadrupedal robots," in *IEEE/RSJ Int. Conf. Intell. Rob. Sys. (IROS)*, 2017, pp. 3359–3365.
- [26] G. A. Pratt and M. M. Williamson, "Series elastic actuators," in *IEEE/RSJ Int. Conf. Intell. Rob. Sys. (IROS)*, 1995.
- [27] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (elus)," in *Int. Conf. on Learn. Repr. (ICLR)*, 2016.