frontiers in Neurorobotics

ORIGINAL RESEARCH
published: 25 September 2015
doi: 10.3389/fnbot.2015.00010

Distributed recurrent neural forward models with synaptic adaptation and CPG-based control for complex behaviors of walking robots

Sakyasingha Dasgupta[1,2,3]*, *Dennis Goldschmidt*[2], *Florentin Wörgötter*[1,2] and *Poramate Manoonpong*[2,4]

[1] Institute for Physics - Biophysics, George-August-University, Göttingen, Germany, [2] Bernstein Center for Computational Neuroscience, George-August-University, Göttingen, Germany, [3] Laboratory for Neural Computation and Adaptation, Riken Brain Science Institute, Saitama, Japan, [4] CBR, Embodied AI and Neurorobotics Lab, The Mærsk Mc-Kinney Møller Institute, University of Southern Denmark, Odense, Denmark

Walking animals, like stick insects, cockroaches or ants, demonstrate a fascinating range of locomotive abilities and complex behaviors. The locomotive behaviors can consist of a variety of walking patterns along with adaptation that allow the animals to deal with changes in environmental conditions, like uneven terrains, gaps, obstacles etc. Biological study has revealed that such complex behaviors are a result of a combination of biomechanics and neural mechanism thus representing the true nature of embodied interactions. While the biomechanics helps maintain flexibility and sustain a variety of movements, the neural mechanisms generate movements while making appropriate predictions crucial for achieving adaptation. Such predictions or planning ahead can be achieved by way of internal models that are grounded in the overall behavior of the animal. Inspired by these findings, we present here, an artificial bio-inspired walking system which effectively combines biomechanics (in terms of the body and leg structures) with the underlying neural mechanisms. The neural mechanisms consist of (1) central pattern generator based control for generating basic rhythmic patterns and coordinated movements, (2) distributed (at each leg) recurrent neural network based adaptive forward models with efference copies as internal models for sensory predictions and instantaneous state estimations, and (3) searching and elevation control for adapting the movement of an individual leg to deal with different environmental conditions. Using simulations we show that this bio-inspired approach with adaptive internal models allows the walking robot to perform complex locomotive behaviors as observed in insects, including walking on undulated terrains, crossing large gaps, leg damage adaptations, as well as climbing over high obstacles. Furthermore, we demonstrate that the newly developed recurrent network based approach to online forward models outperforms the adaptive neuron forward models, which have hitherto been the state of the art, to model a subset of similar walking behaviors in walking robots.

Keywords: neural control, forward models, recurrent networks, locomotion, adaptive behavior, walking robots, synaptic adaptation

OPEN ACCESS

Edited by:
Andrea Soltoggio, Loughborough University, UK

Reviewed by:
Sebastian Risi, IT University of Copenhagen, Denmark
Auke Ijspeert, École Polytechnique Fédérale de Lausanne, Switzerland
Qinggang Meng, Loughborough University, UK

***Correspondence:**
Sakyasingha Dasgupta, Riken Brain Science Institute, 2-1 Hirosawa, Wako, 351-0198 Saitama, Japan
sakyasingha.dasgupta@riken.jp

Received: *02 April 2015*
Accepted: *31 August 2015*
Published: *25 September 2015*

Citation:
Dasgupta S, Goldschmidt D, Wörgötter F and Manoonpong P (2015) Distributed recurrent neural forward models with synaptic adaptation and CPG-based control for complex behaviors of walking robots. Front. Neurorobot. 9:10. doi: 10.3389/fnbot.2015.00010

1. Introduction

Walking animals show diverse locomotor skills to deal with a wide range of terrains and environments. These involve intricate motor control mechanisms with internal prediction systems and learning (Huston and Jayaraman, 2011), allowing them to effectively cross gaps (Blaesing and Cruse, 2004), climb over obstacles (Watson et al., 2002), and even walk on uneven terrain (Cruse, 1976; Pearson and Franklin, 1984). These capabilities are realized by a combination of biomechanics of their body and neural mechanisms. The main components of these neural mechanisms include central pattern generators (CPGs), internal forward models, and limb-reflex control systems. The CPGs generate basic rhythmic motor patterns for locomotion, while the reflex control employs direct sensory feedback (Pearson and Franklin, 1984). However, it is argued that biological systems need to be able to predict the sensory consequences of their actions in order to be capable of rapid, robust, and adaptive behavior. As a result, similar to the observations in vertebrate brains (Kawato, 1999), insects can also employ internal forward models as a mechanism to predict their future state (predictive feedbacks) given the current state or sensory context (sensory feedback) and the control signals (efference copies), in order to shape the motor patterns for adaptation (Webb, 2004; Mischiati et al., 2015). Essentially, such a forward model acts as an internal feedback loop, that uses a copy of the motor command, in order to predict the expected sensory input. Comparing this to the actual input, appropriate modulations of this signal or adaptive behaviors can be carried out.

In order to make such accurate predictions of future actions to satisfy changing environmental demands, the internal forward models require some degree of memory of the previous sensory-motor information. However, given that, such motor control happens on a very fast timescale, keeping track of temporal information is integral to such very short-term memory processes. Reservoir-based recurrent neural networks (RNNs) (Maass et al., 2002; Jaeger and Haas, 2004; Sussillo and Abbott, 2009), with their inherent ability to deal with temporal information and fading memory of sensory stimuli, thus provide a suitable platform to model such internal predictive mechanisms. Taking this perspective, here, we utilize a newly developed model of self-adaptive reservoir networks (SARN) (Dasgupta et al., 2013; Dasgupta, 2015), to act as the forward models for sensorimotor prediction. This works in conjunction with other neural mechanisms for motor control and generates complex adaptive locomotion in an artificial walking robotic system. Specifically, by exploiting the adaptive recurrent layer of our model it is possible to achieve complex motor transformations at different walking gaits, which is significantly difficult to achieve by currently existing adaptive forward models employed with walking robots (Dearden and Demiris, 2005; Schröder-Schetelig et al., 2010; Manoonpong et al., 2013).

We present for the first time a distributed forward model architecture using six SARN-based forward models on a hexapod robot, each of which is for sensory prediction and state estimation of an individual robot leg. The outputs of the models are compared with foot contact sensory signals (actual sensory feedback) and the differences between them are used for motor adaptation, in an online manner. This is integrated as part of the neural mechanism framework consisting of (1) single central pattern generator-based control for generating basic rhythmic patterns and coordinated movements, (2) distributed reservoir forward models and (3) searching and elevation action control for adapting the movement of an individual leg based on the forward model predictions, in order to deal with changing environmental conditions. The distributed nature of the SARN-based forward models allows each leg to act independently with its own feedback and adapt to various environmental situations. This has hitherto, been a difficult problem with centralized motor prediction architectures (Dearden and Demiris, 2005; Pfeifer et al., 2007). Although, there have been some influential distributed architectures for locomotion control of insect inspired robots (Beer et al., 1992; Cruse et al., 1998), they are largely reactive without any prediction (forward model) ability at each leg. In this work, our distributed approach to motor prediction can not only significantly decrease computational demands but also enable each leg with inherent memory in order to make predictions based on its history of sensorimotor signals. This naturally lends to flexibility and robustness of the overall locomotive behavior. Furthermore, each SARN forward model can learn to make predictions for multiple different walking gaits, which was also hitherto not possible in the current state of the art adaptive neuron forward model architecture (Manoonpong et al., 2013). Additionally, the ability to deal with sensorimotor noise or missing information (corrupt signals) of motor commands can be crucial under real environmental conditions. In this work, we will show that the long internal memory of recurrent neural networks naturally allow our forward models to be noise robust and deal with such abnormal conditions to produce truly adaptive locomotion. Overall the neural mechanisms framework presented in this paper makes primary contributions toward making better controllers for insect inspired legged robots (Ijspeert, 2014). While at the same time the developed adaptation mechanism could also suggest a possible role in animal motor control, given the biological plausibility of a distributed neural architecture (Beer et al., 1992) and local leg control (Berg et al., 2015).

In the following section we describe the architectural setup of the neural mechanisms used for the design of adaptive locomotion control in a walking robot, along with a description of the simulated hexapod robot AMOS II and the modular robot control environment used as the development platform for our proposed control system. In Section 3, we present the materials and methods used in this study. Specifically, we introduce the setup and implementation of the distributed reservoir-based adaptive forward model, with details of the learning procedure. Section 4 presents experimental results of the learning mechanism and the resulting behaviors of the simulated hexapod AMOS II on different complex locomotion scenarios likes crossing a large gap, walking on uneven (rough) terrains, overcoming obstacles and dealing with leg damage scenarios. The results obtained from the reservoir based forward models are juxtaposed with the previous state of the art adaptive neuron forward models setup. Finally, in Section 5,

we discuss our results and provide an outlook of further future directions.

2. Neural Mechanisms for Complex Locomotion

The neural mechanisms (**Figure 1A**) for locomotion control, are designed based on a modular architecture, such that, they comprise of, (i) central pattern generator (CPG)-based control, (ii) reservoir-based adaptive forward models, and (iii) searching and elevation action control. The CPG-based control and the searching and elevation control have been previously discussed in detail in Manoonpong et al. (2013), thus here we will only provide a brief overview of these mechanisms, while the reservoir-based adaptive forward models, which forms the main topic of this work, will be presented in detail in the following section.

The CPG-based control primarily generates a variety of rhythmic patterns and coordinates all leg joints of a simulated hexapod robot AMOSII (**Figure 1B**), thereby, leading to a multitude of different behavioral patterns and insect-like leg movements. The patterns include omnidirectional walking and insect-like gaits (Manoonpong et al., 2013). All these patterns can be set manually, or autonomously driven by exteroceptive sensors, like a camera (Zenker et al., 2013), a laser scanner (Kesper et al., 2013), or range sensors. While the CPG-based control provides versatile autonomous behaviors, the searching and elevation control at each leg uses the accumulated error signals provided by the reservoir-based adaptive forward models in order to adapt the movement of an individual leg of the robot and deal with changes in environmental conditions.

The CPG-based control (see Supplementary Figure 1 for detailed description) itself is designed as a modular neural network that consists mainly of the following four elements:

1. CPG mechanism with neuromodulation for generating different rhythmic signals. Inspired by biological findings, here the CPG circuit is designed as a two-neuron fully connected recurrent network (Pasemann et al., 2003) (Supplementary Figure 1, top left), such that using different external neuromodulatory inputs different walking gaits can be achieved.
2. CPG post-processing units (PCPG) for shaping CPG output signals.

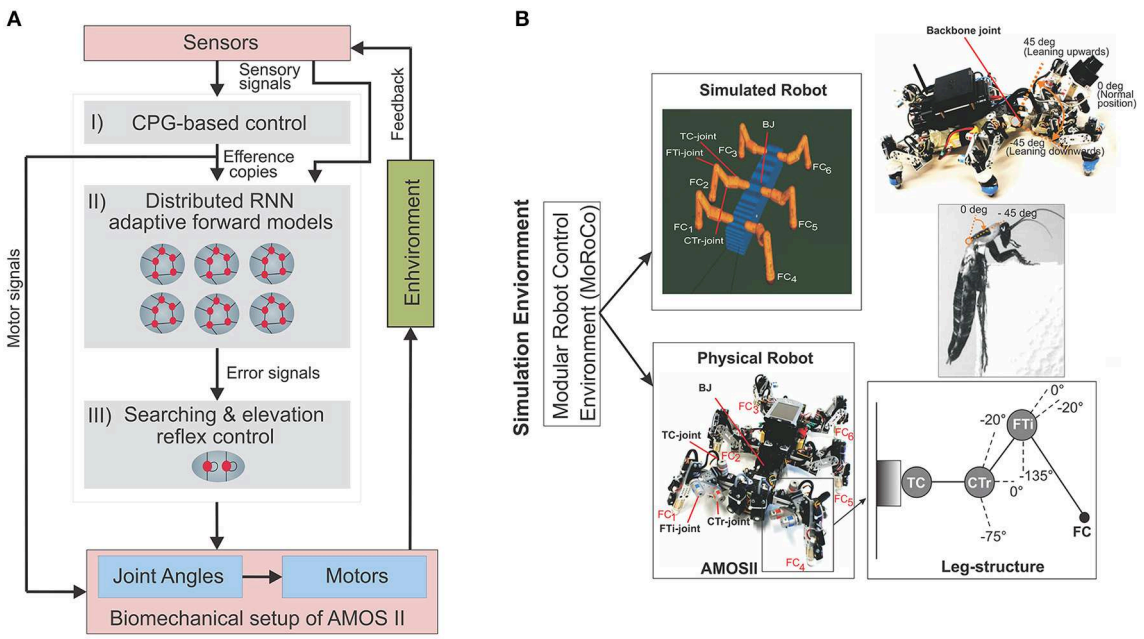

FIGURE 1 | (A) The closed-loop architectural diagram of an artificial bio-inspired walking system consisting of the sensors (i.e., proprioceptive and exteroceptive sensors) that receive environmental inputs and feedback, the neural mechanisms (i, ii, iii) for adaptive locmotion control, and the biomechanical setup of the hexapod robot AMOSII [i.e., six 3-jointed legs, a segmented body structure with one active backbone joint (BJ), actuators, and passive compliant components Manoonpong et al., 2013]. **(B)** Modular Robot Control Environment embedded in the LPZRobots simulation toolkit (Der and Martius, 2012; Hesse et al., 2012). (Top left) The simulation environment provides the main testbed for developing the controller, testing it on the simulated hexapod robot, and finally transferring it to the physical agent. Here we evaluate our model and results primarily on the simulated robot (bottom left), which accurately embodies the characteristics of its physical equivalent, AMOS II robot (bottom left). Here, FC_1, FC_2, FC_3, FC_4, FC_5, and FC_6 are foot contact sensors installed in the robot legs, which are used as the main sensory stimuli compared against the predicted signal from the RNN-based (reservoir) forward models. Each leg (bottom right inset) consists of three joints: the innermost thoraco-coxal (TC-) joint enables forward and backward movements, the middle coxa-trochanteral (CTr-) joint enables elevation and depression of the leg, and the outermost femur-tibia (FTi-) joint enables extension and flexion of the tibia. The morphology of these multi-jointed legs were designed based on a cockroach leg (Zill et al., 2004). (Top right) The front and back parts of the body are connected with a backbone joint (BJ) which primarily allows upwards and downwards tilting of the front body segment (along the horizontal axis). Thus, this is used for climbing and gap crossing purposes. This is also based on a similar joint structure found in the cockroach morphology, allowing it to climb large obstacles. More details on BJ control for climbing can be found in Goldschmidt et al. (2014).

3. Phase switching network (PSN) and velocity regulating networks (VRNs) for walking directional control.
4. Motor neurons with embedded fixed delay lines for transmitting motor commands to all leg joints of AMOS II. These delay lines are utilized to realize the inter-limb coordination, in which they introduce phase differences between the transmitted signals to all leg joints. As a result, the desired walking gait can be achieved.

All neurons of the control network are modeled as discrete-time rate-coded neurons. They are updated with a frequency of approximately 27 Hz (1 time step ≈37 ms). The activity ϕ_i of each neuron in the control network develops according to:

$$\phi_i(t) = \sum_{j=1}^{n} W_{ij}\theta_j(t-1) + \epsilon_i, i = 1, \ldots, n. \quad (1)$$

where, n denotes the number of units, ϵ_i is an internal bias signal or stationary input to each unit i, W_{ij} are the synaptic strength of the connections from neuron j to neuron i. The output θ_i of all neurons of the control network are calculated by using the hyperbolic tangent (tanh) transfer function, i.e., $\theta_i = tanh(\phi_i)$, $\in [-1, 1]$, except for the CPG postprocessing neurons use a step function, the motor output neurons use a piecewise linear transfer function.

The discrete time dynamics of activity states ϕ_i of the two-neuron ($i \in 1, 2$) fully connected CPG circuit, and its output states θ_i follows Equation (1) and a tanh transfer function, respectively. The initial states of the CPG neurons are set to a small positive value, e.g., 0.1. An external excitatory modulatory input MI is introduced to the synaptic connections of the neurons (Supplementary Figure 1, above), in order to modulate the outputs of the CPG. Here different values of MI generates different walking gait patterns (wave, tetrapod, catterpillar, tripod etc.). Although this can be set automatically using sensory inputs (Manoonpong et al., 2013), here we set their values by hand using empirical evaluations. As such, the synaptic weights of the CPG circuit follows:

$$W_{11,22} = d_0, \quad (2)$$
$$W_{12m} = W_{d1} + MI, \quad (3)$$
$$W_{21m} = -(W_{d1} + MI). \quad (4)$$

where, $W_{11,22}$ are fixed synapses with value $d_0 = 1.4$ and $W_{12m,21m}$ are modulated or plastic synapses. Here, W_{d0} and W_{d1} are default synaptic weights selected such that basic periodic signals can be generated. They need to be selected in accordance with the dynamics of the system that generates periodic or quasi-periodic attractors (Pasemann et al., 2003).

The searching and elevation control at each leg, consist of single recurrent neurons that receive the difference (instantaneous error) between the predicted forward model signal and the actual sensory feedback. Due to the recurrent self-connection, this error is accumulated over time. The accumulated error can then be used to either extend specific leg joints in order to get better foothold (searching action) during the stance phase, or elevate further to overcome obstacles during the swing phase (see **Figure 6E** in Section 4.1). Similar to the CPG-based control, all neurons in the searching and elevation control are modeled as discrete-time rate-coded neurons with piece-wise linear activation functions (see Manoonpong et al., 2013, for details), respectively.

3. Materials and Methods

3.1. Reservoir-based Distributed Adaptive Forward Models

We design, six identical adaptive RNN-based forward models ($RF_{1,2,3,\ldots,6}$), one for each leg of the walking robot (**Figure 2A**). These serve the purpose of online sensorimotor prediction as well as state estimation. Specifically, each forward model learns to correctly transform the efference copy of the actual motor signal for each leg joint (i.e., here the CTr-joint motor signal)[1], into an expected or predicted sensory signal. This predicted signal is then compared with the actual incoming sensory feedback signals (i.e., here the foot contact signal—**Figure 2B**, of each leg) and, based on the error accumulated over time, it triggers the appropriate action (searching or elevation) and modulate the locomotive behavior of the robot. Each forward model is based on a random RNN architecture of the self-adaptive reservoir network type (Dasgupta et al., 2013; Dasgupta, 2015). Due to the presence of rich recurrent feedback connections, the dynamic reservoir and intrinsic homeostatic adaptations, the network exhibits a wide repertoire of non-linear activity and long fading memory. This can be primarily exploited for the purpose of specific leg joint-motor signal transformation, act as motor memory and for the prediction of sensorimotor patterns arising in the current context.

3.2. Network Setup

The basic setup of each reservoir forward model can be divided into three layers: input, hidden (or internal), and readout layers (**Figure 2B**). The internal layer consists of a large recurrent neural network driven by time-varying stimuli (CPG motor signals). These driving signals are projected via the input layer. The internal layer is constructed as a random RNN with fixed randomly initialized synaptic connectivity (in this setup we only modify the reservoir-to-readout neuron weights). Using a discrete time version of SARN, with a step size of Δt, the discrete time state dynamics of each reservoir neuron is given by the following equations:

$$x_i(t+1) = \left(1 - \frac{\Delta t}{\tau_i}\right) x_i(t) + \frac{\Delta t}{\tau_i}\left(g\sum_{j=1}^{N} W^{rec}_{i,j} r_j(t) + W^{in}_{i,1} u(t) + B_i\right), \quad i = 1, \ldots, N. \quad (5)$$

$$r_i(t) = \tanh(a_i x_i(t) + b_i), \quad (6)$$

$$\mathbf{z}(t) = \left[\mathbf{W}^{out}\right]^T \mathbf{r}(t). \quad (7)$$

[1] We use the CTr-joint motor signal instead of the TC- and FTi-motor signals since this shows clear swing (off the ground) and stance (on the ground) phases which can be qualitatively matched to the actual foot contact signal.

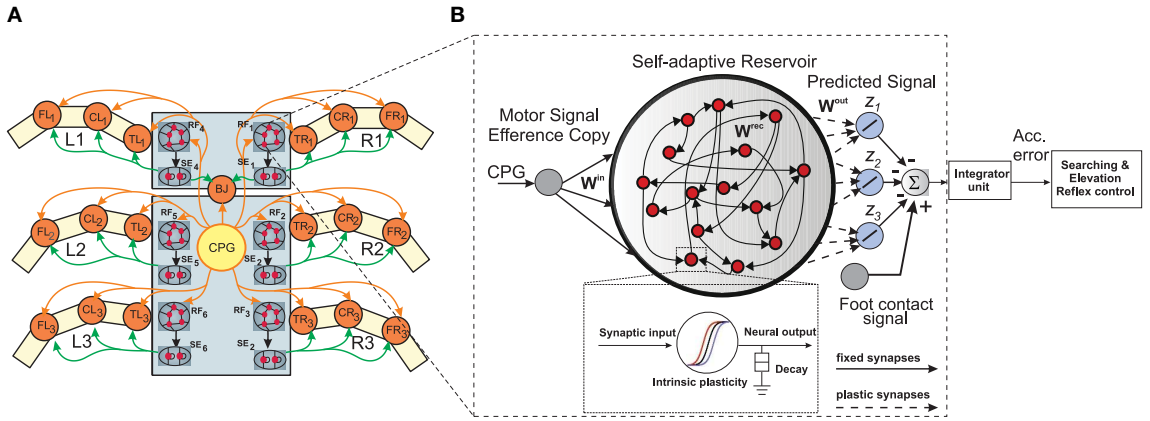

FIGURE 2 | (A) Neural mechanisms implemented on the bio-inspired hexapod robot AMOSII. The yellow circle (*CPG*) represents the neural locomotion control mechanism (see Supplementary Figure 1). The gray circles ($RF_{1,2,3,...,6}$) represent the reservoir-based adaptive forward models. The green circles ($SE_{1,2,3,...,6}$) represent searching and elevation control modules. The orange circles represent leg joints where TR_i, CR_i, FR_i are TC-, CTr- and FTi-joints of the right front leg ($i = 1$), right middle leg ($i = 2$), right hind leg ($i = 3$) and TL_i, CL_i, FL_i are left front leg ($i = 1$), left middle leg ($i = 2$), left hind leg ($i = 3$), respectively. *BJ* is a backbone joint. The orange arrow lines indicate the motor signals which are converted to joint angles for controlling motor positions. The black arrow lines indicate error signals. The green arrow lines indicate signals for adapting joint movements to deal with different circumstances. **(B)** An example of the reservoir-based adaptive forward model. The dashed frame shows a zoomed in view of a single reservoir neuron. In this setup, the input to each of the reservoir network comes from the CTr-joint of the respective leg. The reservoir learns to produce the expected foot contact signal for three different walking gaits (z_1, z_2, z_3). The signals of the output neurons are combined and compared to the actual foot contact sensory signal. The error from the comparison is transmitted to an integrator unit. The unit accumulates the error over time. The accumulated error is finally used to adapt joint movements through searching and elevation control.

The RNN model consists of N neurons, such that the membrane potential at the soma (at time t) of the reservoir neurons, resulting from the incoming excitatory and inhibitory synaptic inputs, is given by a N dimensional vector of neuron state activations. $x(t) = x_1(t), x_2(t), ..., x_N(t)$. The RNN here, does not explicitly model action potentials, but describes neuronal firing rates. Where in, the variable $r_i(t)$ describes the instantaneous firing rate (N dimensional) of the reservoir neurons and is calculated as a non-linear function of the state activation $x_i(t)$ (Equation 5). Each reservoir neuron i, receives inputs from other neurons in the network with firing rates $r_j(t)$ via synaptic connections of strength W_{ij}^{rec} along with incoming stimuli from the input layer via synapses of strength W_{ij}^{in}. Each reservoir neuron is also provided with an auxiliary bias B_i. The parameter g (Sompolinsky et al., 1988; van Vreeswijk and Sompolinsky, 1996) acts as the scaling factor for the recurrent connection weights allowing different dynamic regimes from stable ($g < 1$) to highly irregular chaotic ($g > 1$) (Sussillo and Abbott, 2009), being present in the network.

The input to the reservoir $u(t)$, consists of a single CTr-joint motor signal. This acts as an efference copy of the post-processed CPG motor output. The readout layer consists of three neurons, with their activity being represented by the three-dimensional vector $\mathbf{z}(t)$. Although typically $M < N$ readout neurons can be connected to the reservoir, here we restricted it to three neurons, as each readout here learns the predictive signal for one of the following different walking gaits: wave (z_1), tetrapod (z_2), and caterpillar (z_3) gaits. The wave, tetrapod, and caterpillar gaits are used for climbing over an obstacle, walking on uneven terrain, and crossing a large gap, respectively[2]. Subsequent to the supervised training of the reservoir-to-readout connections $\mathbf{W}^{out}$, each readout neuron basically learns to predict the expected foot contact signal associated with each of these gaits. The decay rate for each reservoir neuron is given by $\frac{1}{\tau_i}$, where τ_i is the individual membrane time constant. The input-to-reservoir connections weights $\mathbf{W}^{in}$ and internal recurrent weights $\mathbf{W}^{rec}$ were drawn randomly from the uniform distribution $[-0.1, 0.1]$ and a Gaussian distribution of zero mean and variance $\frac{g^2}{\sqrt{p_c N}}$, respectively. Where, the parameter p_c controls the probability of connections inside the recurrent layer and is set to be 20%. In order to select the appropriate reservoir size, empirical evaluations were carried out (**Figures 3A,B**), such that we achieved a moderate network size of $N = 30$, for which the minimum prediction error was obtained at the readout layer, irrespective of the walking gait. The recurrent weights were subsequently scaled by the factor of $g = 0.95$ (see **Figure 3**), such that the spontaneous network dynamics is in a stable regime and achieves the best performance of the chosen network size. In accordance with the SARN model, unsupervised intrinsic plasticity (Triesch, 2005) and neuron timescale adaptation (Dasgupta, 2015) were carried out in order to learn the transfer function parameters (a_i and b_i) and the

[2]These three gaits were empirically selected among 19 other possibilities. Previous studies have demonstrated that the wave and tetrapod gaits are the most effective for climbing and walking on uneven terrains, respectively. While in this particular study we observed that the caterpillar gait was the most effective for crossing a gap. However, without any loss of performance, additional walking gaits can be applied easily by adding further readout neurons.

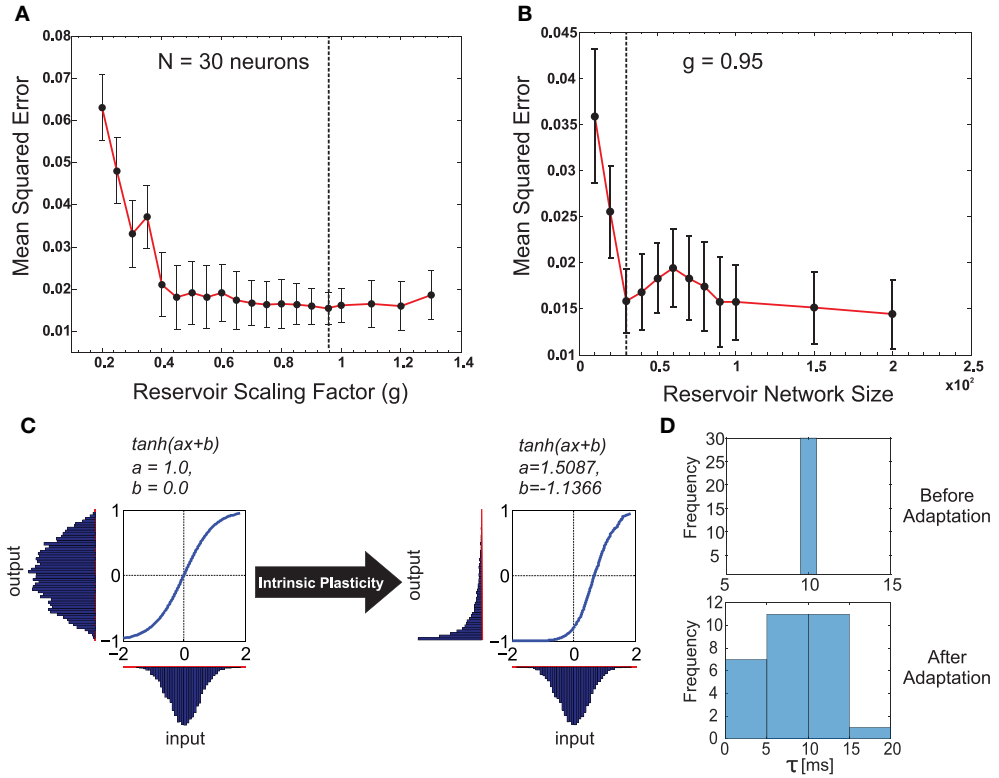

FIGURE 3 | (A) Plot of the change in the mean squared error for the forward model task for one of the front legs (R_1) of the walking robot with respect to the scaling of the recurrent layer synaptic weights W^{rec} with different g-values. As observed, very small values in g have a negative impact on performance compared with values closer to one being better. Interestingly, the performance did not change significantly for $g > 1.0$ (chaotic domain). This is mainly due to homeostasis introduced by intrinsic plasticity in the network. The optimal value of $g = 0.95$ selected for our experiments is indicated with a dashed line. **(B)** Plot of the change in mean squared error with respect to different reservoir sizes (N). g was fixed at the optimal value. Although increasing the reservoir size in general tends to increase performance, a smaller size of $N = 30$ gave the same level of performance as $N = 100$. Accordingly keeping in mind the trade off between network size and learning performance, we set the forward model reservoir size to 30 neurons. Results were averaged over 10 trials with different parameter initializations on the forward model task for a single leg and a fixed walking gait. **(C)** Example of the intrinsic plasticity to adjust the reservoir neuron non-linearity parameters a and b. Initially the the reservoir neuron fires with an output distribution of Gaussian shape matching that of the input distribution. However, after adjustment using intrinsic plasticity mechanism (Dasgupta et al., 2013) the reservoir neuron adapts the parameters a and b, such that, now for the same Gaussian input distribution the output distribution follow a maximal entropy Exponential-like distribution. **(D)** Distribution of the reservoir forward model individual neuron time constants before and after adaptation.

reservoir time constant parameters τ_i for each individual neuron (**Figures 3C,D**).

3.3. Readout Weight Adaptation

Here we used a modified version of the original recursive least squares (RLS) algorithm (Simon, 2002; Jaeger and Haas, 2004) based on the FORCE learning formulation (Sussillo and Abbott, 2009), in order to learn the reservoir-to-readout connection weights $\mathbf{W}^{out}$ at each time step, while the CPG input $u(t)$ is being fed into the reservoir. The readout weights $\mathbf{W}^{out}$ are calculated such that the overall error at the readout neurons is minimized; thereby the network can learn to accurately transform the CTr-motor signal to the expected foot contact signal, for each walking gait. The instantaneous error signal ($e(t)$) at the readout layer, can be calculated as the difference between the reservoir predicted output ($z(t)$) and the desired output, $d(t)$ (i.e., here the expected foot contact signal). Based on Equation (7), this can be formulated as:

$$e(t) = \sum_{j=1}^{3} W_j^{out}(t-1)r_j(t) - d(t). \tag{8}$$

Using the RLS algorithm, and minimizing this error, the readout weights W_j^{out} update can be defined by,

$$W_i^{out} = W_i^{out}(t-1) - e(t)\sum_j P_{ij}(t)r_j(t). \tag{9}$$

Where, $\mathbf{P}$ is a $N \times N$ square matrix proportional to the inverse of the correlation matrix of the reservoir neuron firing rate vector $\mathbf{r}$. $\mathbf{P}$ is initialized using the identity matrix $\mathbf{I}$ and a small constant parameter δ_c as, $\mathbf{P}(0) = \frac{\mathbf{I}}{\delta_c}$. $\mathbf{P}$, here, acts as the adaptive learning rate for updating the readout weights with weight modifications automatically slowing down as $\mathbf{P}$ decreases with time. This allows the learning to occur stably and eventually converge to a solution. $\mathbf{P}$ is updated as each time point as,

$$\mathbf{P}(t) = \mathbf{P}(t-1) - \left(\frac{\mathbf{P}(t-1)\mathbf{r}(t)\mathbf{r}^T(t)\mathbf{P}(t-1)}{1 + \mathbf{r}^T(t)\mathbf{P}(t-1)\mathbf{r}(t)} \right). \quad (10)$$

The reservoir-to-readout neuron weights were initialized to zero at start. Details of all the fixed parameters and initial settings for the reservoir based forward model networks are summarized in Supplementary Table 1.

4. Results

4.1. Learning the Reservoir Forward Model (Motor Prediction)

The entire learning and testing procedure of the SARN-based forward models can be divided into three stages, namely:

1. Pre-training: This stage is used primarily for gathering preliminary sensorimotor data in order to adapt the SARN individual neuron parameters. Here, no reservoir-to-readout weight adaptations occur. This stage can be further divided into,
   - Sequential learning: The robot walks under normal conditions, while sequentially transitioning from one walking gait to another (fixed duration of time). The process is stopped after all the gaits are completed.
   - Offline SARN adaptation: The sensorimotor data collected from the above process is used to adapt the reservoir neuron non-linearity and time constant parameters (Dasgupta et al., 2013).
2. Online training: The same procedure of sequential learning is carried out, however now with ongoing adaptations of the reservoir-readout neuron connection weights based on the RLS algorithm (Equation 9).
3. Testing: After only a single online training cycle the learned forward models are tested on the different experimental conditions.

We now provide a more in depth explanation of these different learning stages.

4.1.1. Pre-training (Without Weight Adaptation)

In order to train the six forward models ($RF_1 toRF_6$) in an online manner, one for each leg, we let the simulated robot AMOSII walk under normal conditions (i.e., walking on a flat terrain with the three different gaits). Initially, we let the robot walk with a certain walking pattern, and then every 2500 time steps (here one time step is equivalent to 37 ms, therefore 2500 time steps is equal to 92.5 s), the gait pattern was sequentially altered (this occurs by changing the modulatory input to the CPG—see Supplementary Figure 1). As a result, the robot sequentially transitions from wave gait, to tetrapod gait, to caterpillar gait repeatedly (here these gaits were empirically selected as the most efficient for the different tasks, however multiple different such gaits can be learned by a single forward model. For an example with commonly used tripod gait, see Supplementary Figure 2). Using this procedure, we let the robot walk for three complete cycles (22,500 time steps) and collected the corresponding CTr-motor signal and foot contact sensor readings for all legs. Intrinsic plasticity and neuron time constant adaptations (Dasgupta et al., 2013; Dasgupta, 2015), were then carried out using 20 epochs of 1000 time steps overlapping time windows. After this pre-training phase, all the reservoir neuron non-linearity parameters and individual time constants (τ_i) were fixed (see **Figure 3D** for the distribution of neuronal time constants before and after training).

4.1.2. Online Training (With Weight Adaptation)

Subsequent to the pre-training phase, normal training of the reservoir-to-readout weights $\mathbf{W}^{out}$ was carried out using the online RLS learning algorithm with the same process of making the robot walk on a flat, regular terrain and sequential switching between the three gait patterns every 2500 time steps. As such, at any given point in time only one of the readout neurons (specific to the walking gait) are active. In this manner, synaptic weights projecting from reservoir to the first readout neuron (z_1) corresponding to the foot contact signal prediction for the wave gait, and synaptic weights projecting to the second (z_2) and third (z_3) readout neurons corresponding to the foot contact signal prediction of the tetrapod and caterpillar gaits, are learned, respectively. Within this experimental setup, as observed from **Figures 4A–C** the readout weights corresponding to each gait converges very quickly, in less than the trial period of 2500 time steps[3]. As a result, every time the CTr-motor signal changes due to walking gait transformations, the RF associated with each leg learns to predict the expected foot contact signal robustly. The training process was carried out only once under normal walking conditions. This was subsequently used as the baseline in order to compare with the actual foot contact signals (sensory feedback) while walking under the situations of crossing a gap, climbing, and negotiating uneven terrains.

Figure 5 shows an example of the forward model prediction (training) during the three different walking gaits, for the right front leg of AMOSII (R_1). Visual inspection clearly demonstrates that according to the corresponding efference copy of CTr-motor signal at a particular gait, the expected foot contact (FC) signal is precisely predicted at each time point. Similarly, the foot contact signals for the other legs are also predicted online, given the current context of CTr-signal (not shown). Note that the FC signals of the other legs normally show slightly different periodic patterns. Furthermore, there exists considerable lag between the expected stance phase according to the motor signal and that observed from the FC signal (difference between dotted green lines in **Figure 5**). Due to the internal memory of the incoming motor signal in the reservoir, we see that the output neurons can adapt to these time lags efficiently, even when the frequency of the signal increases with a change in walking gaits. Furthermore, the reservoir-based forward models enable the robust generation of the predicted FC signal, even in the presence of high noise corruption or missing information in the incoming CTr-joint motor signal (**Figures 5J,K**). Due to the fact that the CTr-motor signals are obtained after appropriate

[3] Due to intrinsic noise and nature of the reservoir-to-readout synaptic adaptation, the weights still show minute fluctuations after successful learning; therefore here convergence applies that the norm of the readout weights $|W^{out}|$ remains constant with a small finite value (Sussillo and Abbott, 2009).

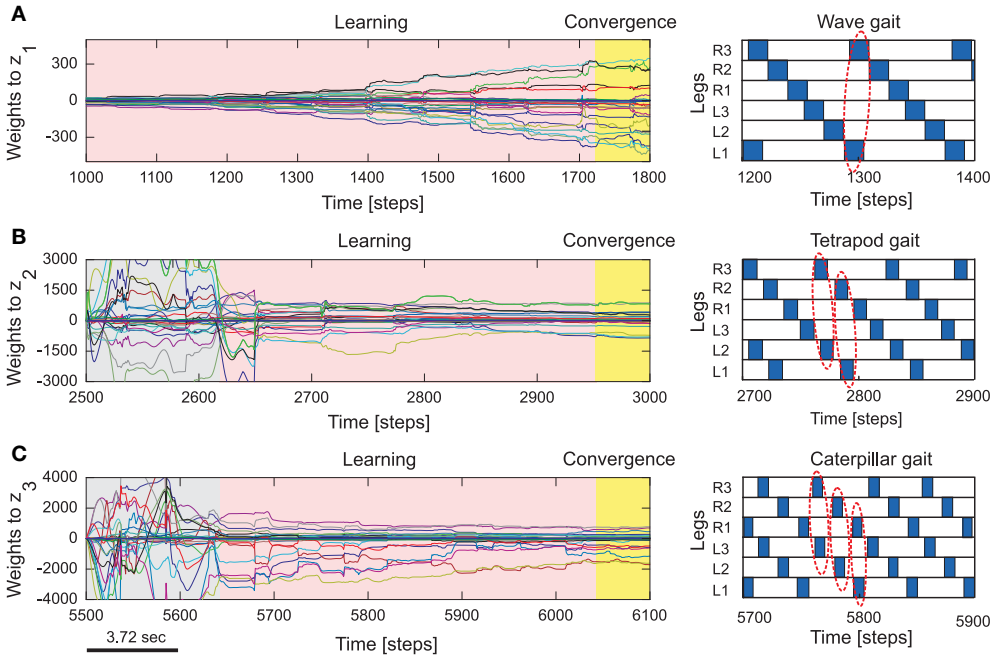

FIGURE 4 | Reservoir-to-readout weight adaptation during online learning. (A) Changes of 30 weights projecting to the first readout neuron (z_1) of the forward model of the right front leg (R_1) while walking with a wave gait. During this period, weights projecting to the second (z_2) and third (z_3) output neurons remain unchanged (i.e., they are zero). **(B)** Changes of the weights to z_2 while walking with a tetrapod gait. During this period, the weights to z_3 still remain unchanged and the weights to z_1 converge to around zero. **(C)** Changes of the weights to z_3 while walking with a caterpillar gait. During this period, the weights to z_1 and z_2 converge to around zero. At the end of each gait, all weights are stored such that they will be used for locomotion in different environments. The gray areas represent transition phases from one gait to another gait and the yellow areas represent convergence. The gait diagrams are shown on the right. They are observed from the motor signals of the CTr-joints (**Figure 5**). White areas indicate ground contact or stance phase and blue areas refer to no ground contact during swing phase. As frequency increases, some legs step in pairs (dashed enclosures). Here convergence implies no significant change in the vector norm of the readout weights.

post-processing of original CPG signals and passage through the motor neurons coupled with different time delays. Such signal corruption can occur at various levels. Therefore, the ability of the forward model to deal with such abrupt noise in the motor signals in a robust manner is crucial to the adaptive mechanisms. Furthermore, such signal corruptions can also occur, due to entrainment mechanisms applied for the automatic tuning or adaptation of CPG outputs (Nachstedt et al., 2013). Such online adaptation for sudden motor signal variations, was not possible in the previous state of the art adaptive neuron forward models (Manoonpong et al., 2013). This model inherently lacked the ability to deal with variations in the temporal properties of the signal. As such, a simple square wave matching the timing of the motor signal efference copy was used, providing a limited range of behavior, as well as being biologically implausible. However, here our reservoir-based model can accurately estimate the spatiotemporal properties of the signal and robustly learn the exact shape, as well as the timing of the actual FC signals.

4.2. Simulated Complex Environments

In order to assess the ability of the reservoir-based forward models to generate adaptive complex locomotive behaviors in a neural closed-loop control system (see **Figure 1**), we conducted simulation experiments under different situations including crossing a gap, walking on uneven terrain and climbing over high obstacles (similar to the behaviors observed in real insects). In all cases, we used the same training procedure for the forward models by allowing the robot to walk under normal conditions on a flat even terrain.

During testing of the learned behavior, while AMOSII walks under different environmental conditions and a specific gait, the output of each trained forward model (i.e., the predicted FC signal, **Figure 6A**) is used to compare it to the actual incoming FC signal of the leg (**Figure 6B**). The difference (instantaneous error signal Δ) between them determines the walking state where a positive value ($+\Delta$) indicates losing ground contact during the stance phase and a negative value ($-\Delta$) indicates stepping on or hitting obstacles during the swing phase.

$$\Delta_i(t) = RF_i(t) - FC_i(t). \tag{11}$$

where $i \in \{1, 2, ..., 6\}$ represents each leg of the robot.

Thus, we use the positive value for searching control (**Figure 6D**, above). This is then accumulated through a single recurrent neuron S with a linear transfer function and is always reset to 0.0 at the beginning of swing phase. Similarly, the negative value is used for elevation control (**Figure 6D**, below). The value is also accumulated through a recurrent neuron E with a linear transfer function. These accumulated errors (**Figure 6C**) thus allow the robot leg to be either elevated (on hitting an

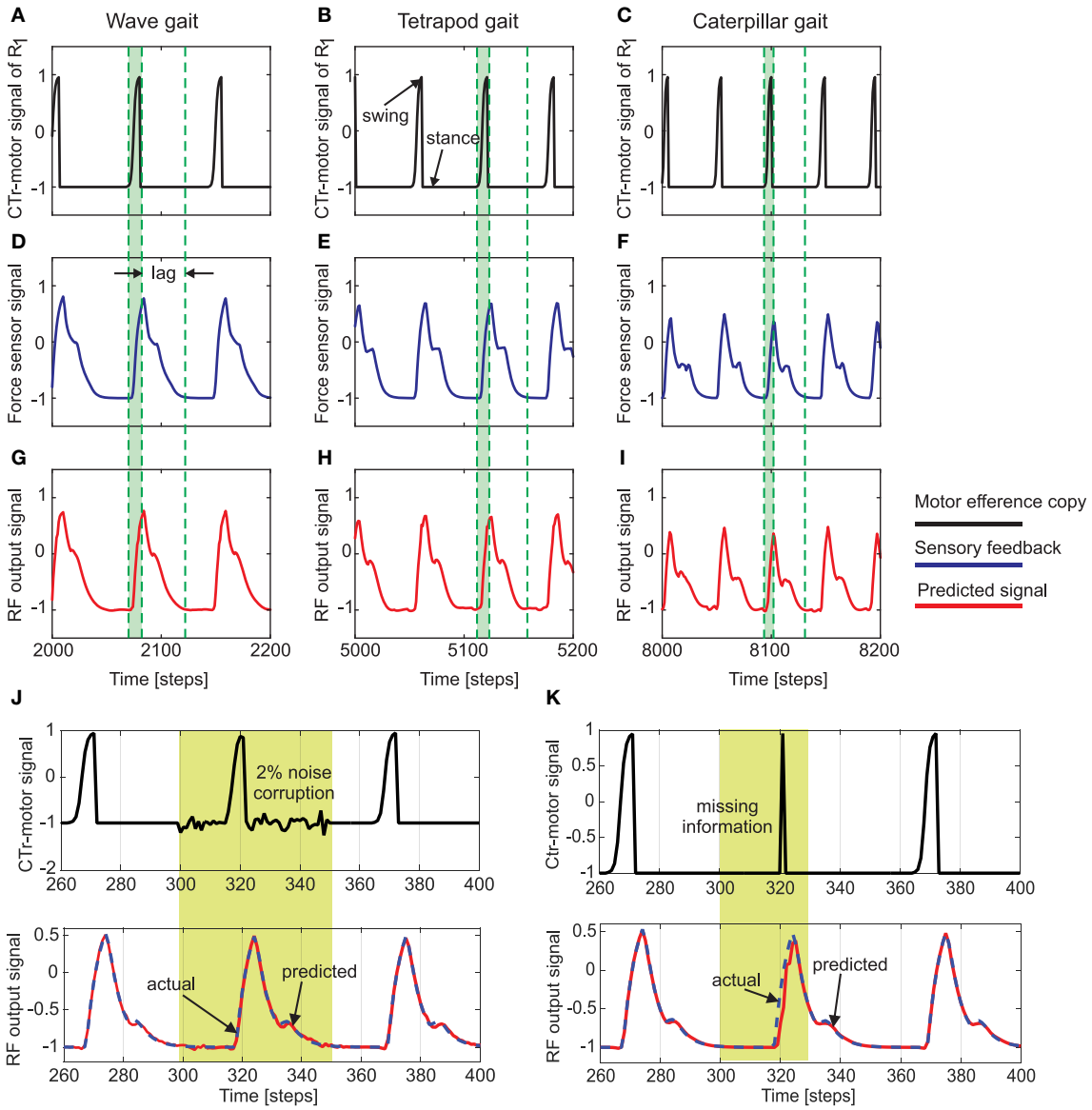

FIGURE 5 | (A–C) The CTr-joint motor signal of the right front leg (R_1) for wave, tetrapod, and caterpillar gaits, respectively. This motor signal provides the efference copy or the input to the reservoir forward models. **(D–F)** The actual foot contact signal (force sensor signal under normal walking conditions) used as the target signal of the reservoir models. **(G–I)** The predicted foot contact signal or the final learned output of the forward model for each walking gait (*RF* output signal). The green shaded region indicates the time interval between swing and stance phase for the CTr motor signal at the three walking gaits. As observed the actual foot contact signal is considerably lagged in time compared to the motor signal. Effectively, this lag decreases with an increase in the gait frequency. The single RF adaptively accounts for these different delay times in order to accurately predict the expected foot contact signal. **(J)** above—CTr-joint motor signal demonstrated for a single leg, with 2% Gaussian noise injected between 300 and 350 time steps (yellow shaded region), below—Despite the noise corruption of the motor signal, the reservoir forward model is able to generate the correct predicted FC signal (blue dotted—target FC signal, red solid—predicted signal). **(K)** above—The CTr-joint motor signal corrupted with missing information between 280 and 320 time steps. As a result, the motor signal shows a narrow spike between 310 and 330 time steps (yellow shaded region), below—Reservoir forward model predicted signal (red) as compared to the desired FC signal (dotted blue). Although the CTr motor signal was transiently missing, the reservoir is able to generate the desired FC signal considerably well, while at the same time maintaining the correct temporal sequence of the signals.

obstacle) or searching for a foothold during the swing and stance phases, respectively (see Manoonpong et al., 2013, for more details of the searching and elevation control). As depicted in **Figures 6A,B**, while walking on a rough terrain (in this case with tetrapod walking gait), the currently recorded sensory feedback or foot contact sensor reading differs considerably from the reservoir predicted signal. As a result, there is a high accumulation of error between each swing or stance phase (**Figure 6C**). It should be noted that the initial (≈50 time steps) abruptly high amplitude signal observed in the reservoir forward

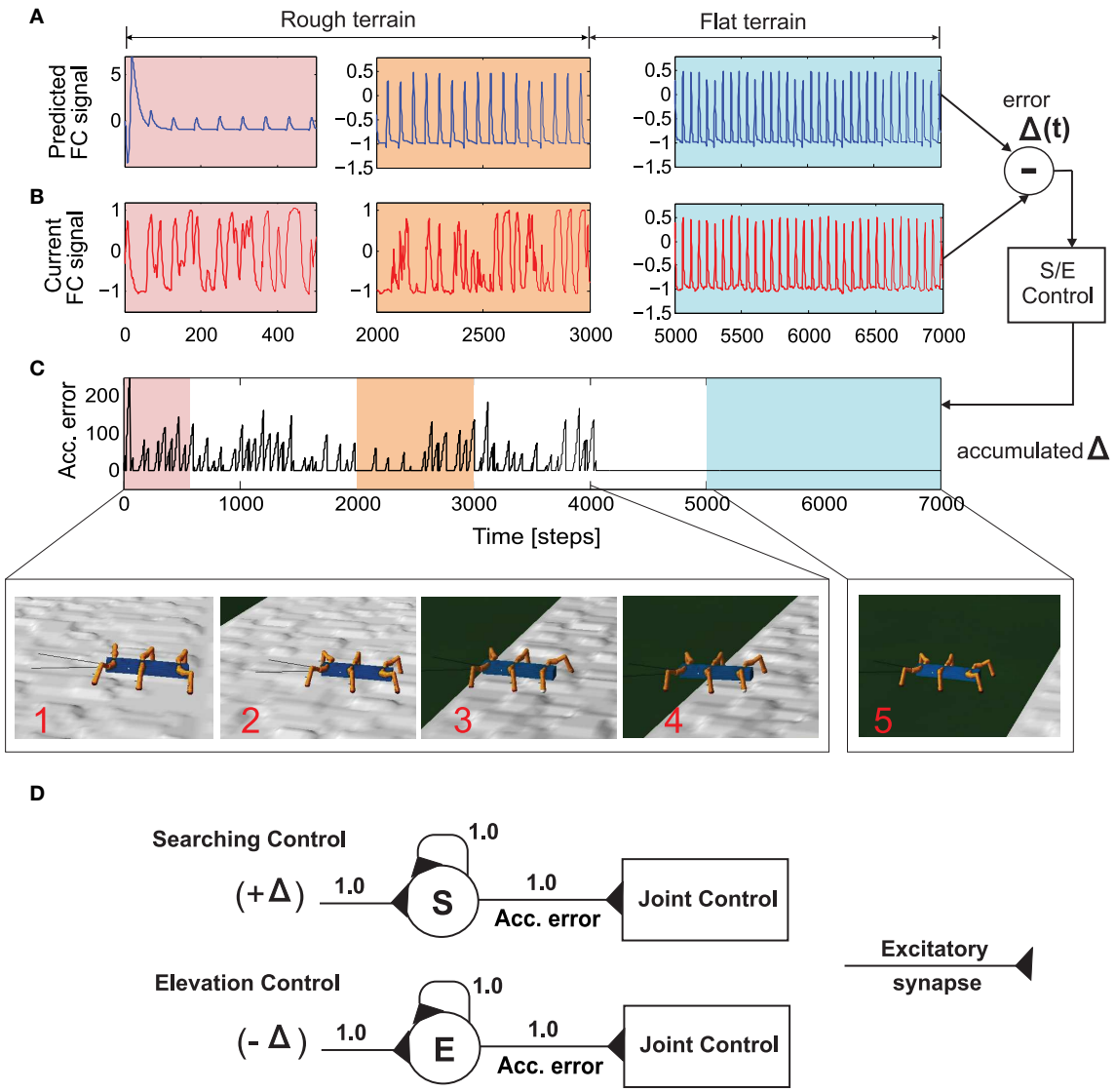

FIGURE 6 | Successfully navigating rough terrain with reservoir forward model. (A) The reservoir forward model predicted, expected foot contact signal. After a small initial transient the reservoir output quickly converges to the expect signal for normal walking condition. **(B)** The actual sensory feedback (foot contact signal) while walking on the rough surface **(C)** Accumulated error calculated from the instantaneous error ($\Delta(t)$) after passing through the recurrent neuron in the searching and elevation control. **(D)** The searching and elevation action control system consisting of individual recurrent neurons as signal accumulators. After 4000 time steps, the robot successfully overcomes the rough terrain and continuous walking on a flat surface. As a result, there is zero accumulated error since the predicted foot contact signal almost exactly matches the actual signal. See the experiment Supplementary Video 3.

model prediction, is caused due to the transient recovery time needed by reservoir readout neurons to settle to the exact learned patterns. This is overcome within the next few time steps and RF predicted FC signal continues to occur in a robust manner. The accumulated error causes the corresponding leg action control mechanism to kick in and the robot successfully navigates out of the rough terrain (after ≈4000 time steps). Once the robot moves into the flat terrain, the reservoir predicted foot contact signal matches almost perfectly with the actual sensory feedback. As a result, the accumulated error becomes zero and normal walking without any additional searching or elevation control mechanisms, can continue. In essence based on the reservoir forward models, while traversing from the uneven terrain (**Figure 6**, inset 1–4) to the flat terrain (**Figure 6**, inset 5), the robot can adapt its legs individually to deal with the change of terrain. That is, it depressed its leg and extended its tibia to search for a foothold when loosing ground contact during the stance phase. Losing ground contact information is detected by a significant change of the accumulated errors (**Figure 6C**). In case of both walking on uneven terrain and climbing, this accumulated error causes shifting of the CTr- and FTi-joints causing the respective leg to search for a foothold. However, in the

specific case of crossing a gap (**Figure 7**), we use the accumulated error in order to control tilting of the backbone joint (BJ) and shifting of the TC- and FTi-joints such that the front legs can be extended forward continuously till the robot can find a foothold. In addition to this leg joint control, reactive backbone joint control using the additional ultrasonic sensors in front of the robot can also be used to learn to lean up the BJ for climbing over obstacles (this has been previously successfully applied using classical conditioning based learning in Goldschmidt et al. (2014) and as such not discussed here).

We now take the example of the more complex, multiple gap crossing experiment in order to look in detail at the learning outcome of the forward models. This experiment was divided into two components, consisting of one larger gap (15cm length) and another relatively shorter gap of 11 cm length. The two gaps were separated by considerable distance where the robot was allowed to walk on a regular flat terrain. In order to learn to cross a gap, we let AMOS II walk with a caterpillar gait (see **Figure 4C**, right), such that each left and right pair of legs moves simultaneously. Empirically this is observed to be the most suited gait for overcoming large gaps, as well as supported by experimental observations in stick insects (Blaesing and Cruse, 2004). As shown in **Figure 7**(1), at the beginning AMOS II walked forward straight toward the initial gap. In this period, as it walks on the flat surface of the platform, it performed regular movements similar to the training period under normal walking conditions (training on a flat regular surface). Eventually, it encounters a 15 cm wide gap (≈44% of body length—the maximum cross-able distance). In this situation, during the subsequent stance phase the front legs of the robot loose ground contact (**Figures 7D,E**). As a result, the foot contact sensors from the front legs do not record any value. However, the reservoir forward model still predicts the expected foot contact signal, causing a positive instantaneous error (Equation 11). This leads to a gradual ramping of the accumulated error signal between each stance phase and swing phase, for the front legs (**Figure 7A**).

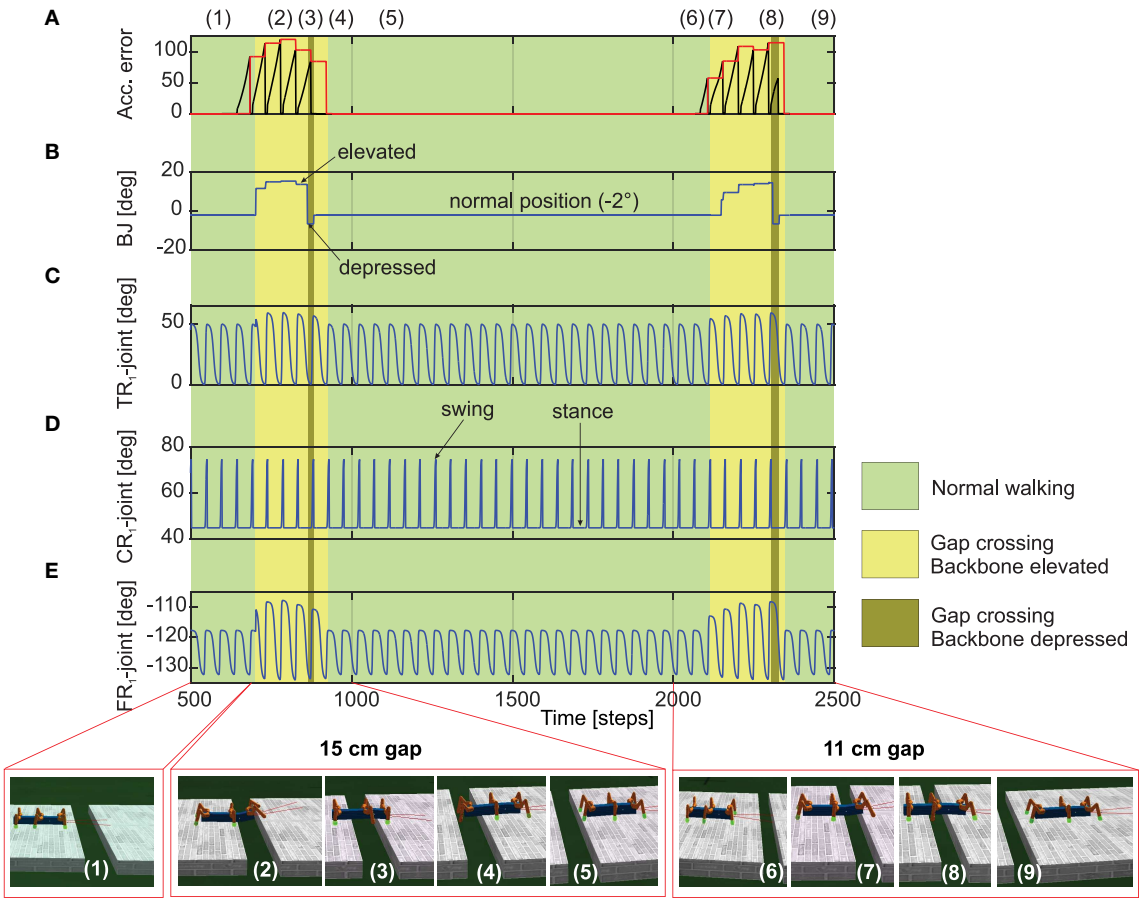

FIGURE 7 | Real-time data of walking and crossing multiple gaps using the forward model predictions. (A) The accumulated error (black line) and the maximum accumulated error value at the end of each stance phase (red line) of the right front leg (R_1). The accumulated error is reset to zero every swing phase. **(B)** The backbone joint (BJ) angle during walking and gap crossing. The BJ stays at the normal position (−2°) during normal walking. On encountering a gap (15cm), it leans upwards in a step like fashion and then finally bent downwards in order to cross the gap. This procedure is repeated for the second gap (11cm), however with different degree of elevations. **(C–E)** The TC-, CTr-, and FTi-joint angles of right front leg R_1 during normal walking and gap crossing. The joint adaptation was controlled by the maximum accumulated error value of the previous step (red line). Below pictures show snap shots of the locomotion of AMOS II during the experiment. Note that one time step is ≈0.037 s. For further details interested readers are recommended to see the experiment Supplementary Videos 1, 2.

Please note that here the slope of the accumulated error signal was empirically adjusted. Too small or too large values for the slope of the ramp may cause inadequate or large extensions of the leg.

In order to activate the BJ and adapt the leg movements due to the difference between the reservoir predicted FC signal and the actual sensory feedback of the FC sensors (error signals), we used the maximum accumulated error value of the previous step (**Figure 7A**, red line) and control the BJ and leg movements in the subsequent step. In this manner, the BJ started to lean upwards incrementally (step like manner) at around 680–850 time steps [**Figure 7**(2)]. Simultaneously, the TC- and FTi-joint movements of the left and right front legs were also adapted accordingly in order to carry out elevation action (this is reflected in the higher amplitude of these two signals in this time period). Due to a predefined time-out period for tilting upwards, at around 850 time steps [**Figure 7**(3)], the backbone joint automatically moved downwards recording a negative value. Consequently, the front legs touch the ground of the second platform at the middle of the stance phase; thereby, causing the accumulated error signals to decrease. Due to another time-out period for tilting downwards at around 900 time steps [**Figure 7**(4)], the BJ automatically moved to the normal position (-2°). Since now the situation is similar to walking on flat terrain, the RF predicted foot contact signal matches the one recorded by the foot sensors, with accumulated error dropping to zero. Thereafter, the TC- and FTi-joints perform regular movements. Subsequently left and right hind legs loose the ground contact, and AMOSII continues to walk forward. Here the movements of the TC- and FTi-joints were slightly adapted allowing AMOS II to successfully cross the gap and continue walking on the second platform [**Figure 7**(5)]. As the terrain now resembles a regular flat surface (similar to the original training terrain) AMOSII two continues to walk forward in normal manner with no accumulated errors being present. However, the same procedure is repeated once again, when AMOSII re-encounters the second gap at around 2100 time steps. However, in this case, since the gap length is much smaller, the elevation in the BJ occurs with an initial increment of smaller amplitude [**Figure 7**(2)] as compared to the previous case. Thereafter, a similar process is followed and AMOSII can once again successfully overcome this gap and continue walking on the other end of the platform [**Figure 7**(9)]. This clearly demonstrates the adaptive yet robust performance of the forward model based predictions in order to successively cross gaps of different length.

Figure 8 shows that the reservoir forward model in combination with the neural locomotion control mechanisms, not only successfully generates gap crossing behavior of AMOS II and learns to walk on uneven terrain, but also allows it to climb over single and multiple obstacles (e.g., up a fleet of stairs). In all these cases, we directly used the accumulated errors for movement adaptation via the searching and elevation control mechanisms. For climbing, the reactive backbone joint control was also applied to the system (see Goldschmidt et al., 2014, for more details) and a slow wave gait walking pattern (see **Figure 4A**, right) was used.

Experimentally the wave gait was found to be the most effective for climbing, which allows AMOSII to overcome the highest climbable obstacle (i.e., 15 cm height which equals ≈86% of its leg length) and to surmount a fleet of stairs. For walking on uneven terrain, a tetrapod gait (see **Figure 4B**, right) was used without the backbone joint control. This is the most effective gait for walking on uneven terrain (see also Manoonpong et al., 2013). Recall that in all experiments the forward models basically generate the expected foot contact signals (i.e., sensory prediction), which are compared to the actual incoming ones. Errors between the expected and actual signals during locomotion serve as state estimation and are used to adapt the joint movements accordingly. It is important to note that, the best gait for each specific scenario was experimentally determined and fixed. However, this could be easily extended with learning mechanisms (see Steingrube et al., 2010) to switch to the desired gait when the respective behavioral

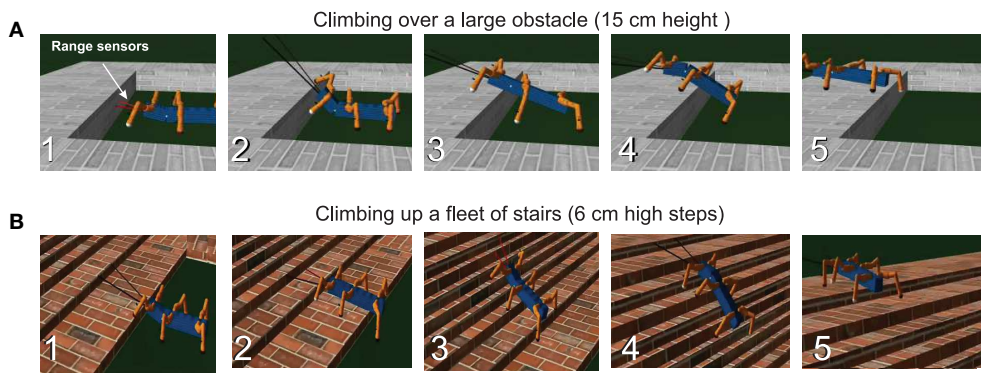

FIGURE 8 | Snapshots showing the learned behavior during climbing over a high obstacle and climbing up a fleet of stairs. (A) AMOSII walked with the wave gait and approached a 15 cm high obstacle (1). It detected the obstacle using its range sensors installed at its front part. The low-pass filtered range sensory signals control the BJ to tilt upwards (2) and then back to its normal position (3). Due to the missing foot contact of the front legs, the BJ moved downwards to ensure stability (4). During climbing, middle and hind legs lowered downwards due to the occurrence of the accumulated errors, showing leg extension, to support the body. Finally, it successfully surmounted the high obstacle (5). For further details see the Supplementary Video 4 **(B)** AMOSII climbed up a fleet of stairs (1–5) using the wave gait as well as the reactive BJ control. The climbing behavior is also similar to the one described in the case **(A)**. For further details see Supplementary Video 5.

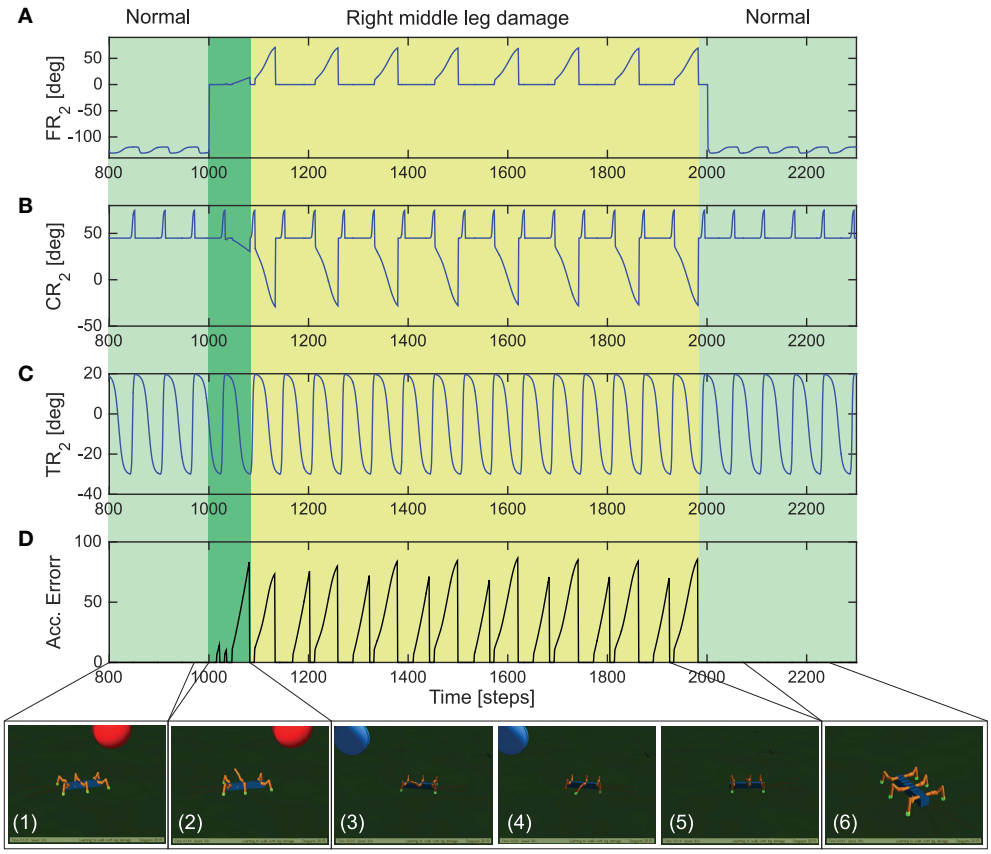

FIGURE 9 | Real-time data for adaptive locomotion to overcome leg damage. (A) The FT-i joint angles of the right middle leg R_2. **(B)** The CT-i joint angles of the right middle leg R_2. **(C)** The TC-i joint angles of the right middle leg. **(D)** Accumulated error signal at the end of each stand phase. It is reset to zero at every swing phase. Below pictures show the locomotion of AMOSII during the experiment (temporal spacing of the panels are not exact). Please see the Supplementary Video 6 for closer look at the exact adaptive behavior.

scenarios are encountered, without any additional influence on the performance of the reservoir forward models.

Adaptations in both biological and robotic systems, not only requires the ability to deal with different environmental conditions for complex locomotion (as demonstrated with the gap crossing, climbing and uneven terrain navigation examples) but can also require the ability to adapt to sudden or abrupt changes in body properties, like growth or lesions (e.g., damage to robot joint motors or connections being disengaged) (Cully et al., 2015). Therefore, here, we demonstrate that the distributed reservoir-based forward models allows the robot to adapt the movements of a damaged leg and its walking gait, in order to deal with sudden leg damage situations. In this scenario, post learning of the forward models under the three different walking gaits, we initially let the robot walk with a tetrapod gait (**Figure 4B**, right). After 1000 steps (≈37 s) we constrained (deactivated) the FT-i joint (outermost) of the right middle leg such that the leg remains suspended in air and cannot achieve ground contact in this configuration. Thus, simulating leg damage scenario. AMOSII was then allowed to continue walking on the flat terrain under this damaged condition.

As observed in **Figure 9**, initially AMOSII walks under normal conditions (photo panel 1) with the right middle leg FT-i joint functioning normally. The FT-i joint was then constrained to 0° maximum and minimum angle of clearance (**Figure 9A**) thereby causing the right middle leg to be suspended in the air (photo panel 2). As a result the reservoir forward model prediction mismatches the current footcontact signal on the damaged leg, causing the accumulated error to gradually ramp up (**Figure 9D**). After a short transient period of AMOSII trying to walk in this configuration (dark green section in **Figure 9**), this results in adaptations in the FT-i and CT-i joints (yellow highlighted section in **Figures 9A,B**) thereby, allowing the robot to extend the damaged leg further down and support the locomotion (photo panels 3, 4, and 5). As a result, AMOSII was able to successfully keep walking straight with a slightly modified tetrapod gait despite the damaged right middle leg. Finally, after 2000 time steps (≈74 s), the FT-i joint was once again allowed to function normally, causing the accumulated error to become zero (the forward model prediction matches the actual footcontact signal). The robot then continues to walk as in the undamaged condition with a tetrapod gait. For further details, we encourage the readers to see the Supplementary Video 6 of the entire experiment. These results, thus clearly demonstrate that the distributed reservoir forward models not only allow complex locomotive behaviors, but also enable the

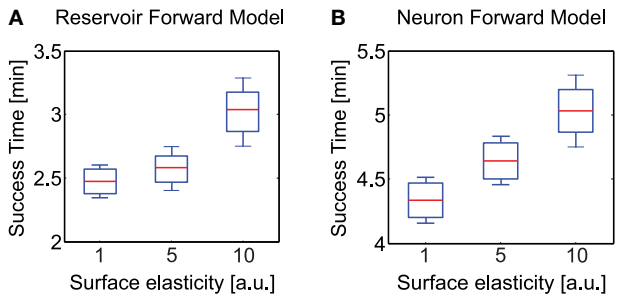

FIGURE 10 | Average time to successfully overcome uneven terrains of different elasticity (hard, moderate, highly elastic). (A) Average success time for reservoir-based forward model. **(B)** Average success time for adaptive neuron forward model from Manoonpong et al. (2013). Here the whiskers indicate one standard deviation above and below the mean value. Note the difference in scale of the y-axis in both plots. The experimental surface here consisted of the rough terrain as presented in **Figure 6** consisting of irregular undulations, however with varying degree of elasticity for the three cases.

robot to deal with unwanted changes in body properties in a robust manner.

In order to evaluate the performance of our adaptive reservoir forward model in comparison to the state of the art model recently presented in Manoonpong et al. (2013) (single recurrent neural with low-pass filter), we carried out simulation experiments with AMOSII walking on different types of surfaces. Specifically, after training on a flat surface (under normal conditions) we carried out 10 trials each with the robot walking on uneven terrains (laid with multiple obstacles of height 8 cm), having three different elastic properties[4]. The surfaces were divided into hard (1.0), moderately elastic (5.0) and highly elastic (10.0). A tetrapod walking gait was used in all three cases. Starting from a fixed position, we noted the total time taken by the robot to successfully cross the uneven terrain region and move into a flat surface region. As observed in **Figures 10A,B**, the reservoir forward model enables the robot to traverse the uneven region considerably faster as compared to the adaptive neuron forward model, in all three scenarios. Both the models can be seen to overcome the hard surface much better as compared to the elastic ones. This was expected due to the changes in surface stiffness resulting in additional forces on the robot legs. However, the reservoir model performance was considerably more robust with a mean difference in success time of 1.86 min for the hardest surface and approximately 2 min for the most elastic surface, cases. Given that the walking gait was fixed, here the success time can be thought as an indicator of the robot's energy efficiency. In the absence of additional body mechanisms to deal with changing surface stiffness, the reservoir based model outperforms the previous implementations of adaptive forward models by ≈25% on average. In the climbing and gap crossing scenarios, the performance of the two forward models are comparable (not shown here explicitly) unless there are significant changes in the ground reaction forces (e.g., climbing or crossing gaps on different types of terrain). As such the reservoir forward model offers a more generalized architecture for adaptive locomotion. Furthermore, as demonstrated previously, this model is also capable of robustly coping with missing motor information and a high degree of sensory noise; making use of the SARN internal memory and multiple timescales (Dasgupta, 2015). This was very difficult to achieve with the previous simple single recurrent neuron forward models. Moreover, the previous study also required that a separate forward model be learned for every different walking gait. Thus, creating a scalability issue for real robot implementations. Here, however, a single SARN can be trained online to predict the foot contact signals for multiple different walking gaits (here we show three gaits, but it can be easily extended to many more patterns—see Supplementary Figure 2, for tripod gait example).

[4]Here the elasticity coefficients do not strictly represent Young's modulus values. These were local parameter setting defined in the simulation, with increasing values causing greater elasticity.

5. Discussion

In this study, we presented adaptive forward models using the self-adaptive reservoir network for locomotion control. The model is implemented on each leg of a simulated bio-inspired hexapod robot. It is trained online during walking on a flat terrain in order to transform an efference copy (motor signal) into an expected foot contact signal (i.e., sensory prediction). Afterwards, the learned model of each leg is used to estimate walking states by comparing the expected foot contact signal with the actual incoming one. The difference between the expected and actual foot contact signals is used to adapt the robot's leg through elevation and searching control. Each leg is adapted independently. This enables the robot to successfully walk on uneven terrains. Moreover, using a backbone joint, the robot can also successfully cross a large gap and climb over a high obstacle as well as up a fleet of stairs. In this approach, basic walking patterns are generated by CPG-based control along with local leg control mechanisms that make use of the reservoir prediction to adapt the robot's behavior. The key neural mechanisms presented in this work, namely, CPG -based neural control, internal forward models and local leg control, are essential for robust, adaptive locomotion control. However, only individual instances of them has been successfully realized on artificial and bio-mimetic robotic systems (Bläsing, 2004; Pfeifer et al., 2007; Lewinger and Quinn, 2011; Ren et al., 2012; Schilling et al., 2012; Christensen et al., 2014; Cully et al., 2015); thereby achieving partial solutions. Furthermore, although a few studies have focused on a combination of these neural mechanisms, they have largely been tailored for adaptive locomotion in quadruped robots (Lewis and Bekey, 2002; Silva et al., 2012), without the ability to climb obstacles or cross large gaps, as observed in real animals and insects. Thus, this work demonstrates how the combination of these essential components, coupled with the power of the adaptive recurrent neural forward models can achieve very rich behavioral repertoire in bio-inspired hexapod robots. Thus, supporting the idea that such embodied neural control (Floreano et al., 2014) is indeed a potential powerful future alternative of more conventional control methods.

It is important to note that the usage of reservoir networks, as forward models here, provides the crucial benefit of an

inherent representation of time and fading memory (due to the internal feedback loops and input dependent adaptations). Such memory of the time-varying motor or sensory stimuli is required to overcome intrinsic time lags between expected sensory signals and motor outputs (Wolpert et al., 1998), as well as in behavioral scenarios with considerable dependence on the history of motor output (Lonini et al., 2009). This is very difficult in most of the previous implementations of forward internal models using either simple single recurrent neuron implementations (Manoonpong et al., 2013), feed-forward multi-layered neural networks (Schröder-Schetelig et al., 2010), or Bayesian network models (Dearden and Demiris, 2005; Sturm et al., 2008). Furthermore, in this case, online adaptation of only the reservoir-to-readout weights (readout) makes such networks beneficial for simple and online learning. The pre-training phase of the current setup was carried out only to gather sufficient statistics of the CTr-motor signals and foot-contact signals while walking under the different gaits, in order to learn the optimal reservoir neuron non-linearity and time constant parameters (Dasgupta et al., 2013). Subsequent to this, reservoir-to-readout weight learning occurs continuously without the need of any offline batch mode phase. Moreover, only a single learning trial under normal walking conditions was enough to learn the forward model for leg adaptations under different environmental situations. As a result making the reservoir based forward models very suitable for fast learning under real robot implementations.

The concept of forward models with efference copies in conjunction with neural control has been suggested since the mid-twentieth century (Holst and Mittelstaedt, 1950; Held, 1961) and increasingly employed for biological investigations (Webb, 2004). This is because it can explain mechanisms which biological systems use to predict the consequence of their action based on sensory information, resulting in adaptive and robust behaviors in a closed-loop scenario. This concept also forms a major motivation for robots inspired by biological systems. Within this context, the work presented here, verifies that a combination of CPG-based neural control, adaptive reservoir forward models with efference copies, and searching and elevation control can be used for robustly generating complex locomotion and adaptive behaviors in an artificial walking system. Additionally, although in this study we specifically focused on locomotive behaviors for walking robots, (such) SARN based motor prediction systems can be easily generalized to a number of other applications. Specifically for neuro-prosthetics (Ganguly and Carmena, 2009), sensor-driven orthotic control (Lee and Lee, 2005; Braun et al., 2014) or brain-machine interface devices (Golub et al., 2012), that require the learning of such predictive models using highly non-stationary, temporal signals, applying SARN models can provide high performance gains with embedded memory, as compared to the current static feed-forward neural network solutions.

In the future, we will transfer the reservoir-based adaptive forward models to the physical hexapod robot AMOS-II (Manoonpong et al., 2013) in order to test the adaptive behaviors in a real environment. Typically, the transfer of learning from simulation studies to physical hardware involves additional sensory and motor noise. As demonstrated in **Figures 5J,K**, the SARN based forward models are robust to significant levels of sensory noise as well as capable of dealing with corruption of motor signals. As such, although the currently presented results are in simulation, we envision that a transfer to a noisy real robot platform can be easily achieved. Furthermore, while the work presented here uses only a single CPG, the control mechanism and the distributed nature of the forward models allow for easy extension to multiple CPGs (Barikhan et al., 2014; Ren et al., 2015). For multiple CPGs, synchronization can emerge from continuous interactions of distributed CPGs, body dynamics, and the environment through local sensory feedback of each leg as shown in Barikhan et al. (2014); or can be also achieved by using a master-client mechanism with learning as demonstrated in our previous work (Ren et al., 2015).

Author Contributions

SD, FW, and PM designed the research. SD and PM implemented the model, analyzed data and carried out simulations. DG carried out the climbing experiments. SD and PM wrote the manuscript.

Acknowledgments

This research was supported by the Emmy Noether Program (DFG, MA4464/3-1), the Federal Ministry of Education and Research (BMBF) by a grant to the Bernstein Center for Computational Neuroscience II Göttingen (01GQ1005A, project D1) and the International Max Planck Research School for Physics of Biological and Complex Systems scholarship.

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fnbot.2015.00010

References

Barikhan, S. S., Wörgötter, F., and Manoonpong, P. (2014). "Multiple decoupled cpgs with local sensory feedback for adaptive locomotion behaviors of bio-inspired walking robots," in *From Animals to Animats 13* (Castellón: Springer), 65–75.

Beer, R. D., Chiel, H. J., Quinn, R. D., Espenschied, K. S., and Larsson, P. (1992). A distributed neural network architecture for hexapod robot locomotion. *Neural Comput.* 4, 356–365. doi: 10.1162/neco.1992.4.3.356

Berg, E. M., Hooper, S. L., Schmidt, J., and Büschges, A. (2015). A leg-local neural mechanism mediates the decision to search in stick insects. *Curr. Biol.* 25, 2012–2017. doi: 10.1016/j.cub.2015.06.017

Blaesing, B., and Cruse, H. (2004). Stick insect locomotion in a complex environment: climbing over large gaps. *J. Exp. Biol.* 207, 1273–1286. doi: 10.1242/jeb.00888

Bläsing, B. (2004). "Adaptive locomotion in a complex environment: simulation of stick insect gap crossing behaviour," in *From Animals to Animats 8* (Los Angeles, CA), 173–182.

Braun, J. M., Wörgötter, F., and Manoonpong, P. (2014). "Internal models support specific gaits in orthotic devices," in *Mobile Service Robotics, Number 17 in Proceedings of the International Conference on Climbing and Walking Robots* (Poznań), 539–546.

Christensen, D. J., Larsen, J. C., and Stoy, K. (2014). Fault-tolerant gait learning and morphology optimization of a polymorphic walking robot. *Evol. Syst.* 5, 21–32. doi: 10.1007/s12530-013-9088-3

Cruse, H. (1976). The control of body position in the stick insect (carausius morosus), when walking over uneven surfaces. *Biol. Cybern.* 24, 25–33. doi: 10.1007/BF00365591

Cruse, H., Kindermann, T., Schumm, M., Dean, J., and Schmitz, J. (1998). Walkneta biologically inspired network to control six-legged walking. *Neural Netw.* 11, 1435–1447. doi: 10.1016/S0893-6080(98)00067-7

Cully, A., Clune, J., Tarapore, D., and Mouret, J.-B. (2015). Robots that can adapt like animals. *Nature* 521, 503–507. doi: 10.1038/nature14422

Dasgupta, S. (2015). *Temporal Information Processing and Memory Guided Behaviors with Recurrent Neural Networks*. Ph.D. thesis, Georg-August University, Göttingen.

Dasgupta, S., Wörgötter, F., and Manoonpong, P. (2013). Information dynamics based self-adaptive reservoir for delay temporal memory tasks. *Evol. Syst.* 4, 235–249. doi: 10.1007/s12530-013-9080-y

Dearden, A., and Demiris, Y. (2005). "Learning forward models for robots," in *International Joint Conference on Artificial Intelligence*, Vol. 5 (San Francisco, CA), 1440.

Der, R., and Martius, G. (2012). "The LPZRobots simulator," in *The Playful Machine* (Berlin; Heidelberg: Springer), 293–308. doi: 10.1007/978-3-642-20253-7_16

Floreano, D., Ijspeert, A. J., and Schaal, S. (2014). Robotics and neuroscience. *Curr. Biol.* 24, 910–920. doi: 10.1016/j.cub.2014.07.058

Ganguly, K., and Carmena, J. M. (2009). Emergence of a stable cortical map for neuroprosthetic control. *PLoS Biol.* 7:e1000153. doi: 10.1371/journal.pbio.1000153

Goldschmidt, D., Wörgötter, F., and Manoonpong, P. (2014). Biologically-inspired adaptive obstacle negotiation behavior of hexapod robots. *Front. Neurorobot.* 8:3. doi: 10.3389/fnbot.2014.00003

Golub, M. D., Yu, B., and Chase, S. M. (2012). "Internal models engaged by brain-computer interface control," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE* (San Diego, CA: IEEE), 1327–1330.

Held, R. (1961). Exposure-history as a factor in maintaining stability of perception and coordination. *J. Nerv. Ment. Dis.* 132, 26–32. doi: 10.1097/00005053-196101000-00005

Hesse, F., Martius, G., Manoonpong, P., Biehl, M., and Wörgötter, F. (2012). "Modular robot control environment testing neural control on simulated and real robots," in *Frontiers in Computational Neuroscience, Conference Abstract: Bernstein Conference* (Munich), 1416–1420.

Holst, E., and Mittelstaedt, H. (1950). Das reafferenzprinzip. *Naturwissenschaften* 37, 464–476. doi: 10.1007/BF00622503

Huston, S. J., and Jayaraman, V. (2011). Studying sensorimotor integration in insects. *Curr. Opin. Neurobiol.* 21, 527–534. doi: 10.1016/j.conb.2011.05.030

Ijspeert, A. J. (2014). Biorobotics: using robots to emulate and investigate agile locomotion. *Science* 346, 196–203. doi: 10.1126/science.1254486

Jaeger, H., and Haas, H. (2004). Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science* 304, 78–80. doi: 10.1126/science.1091277

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Curr. Opin. Neurobiol.* 9, 718–727. doi: 10.1016/S0959-4388(99)00028-8

Kesper, P., Grinke, E., Hesse, F., Wörgötter, F., and Manoonpong, P. (2013). Obstacle/gap detection and terrain classification of walking robots based on a 2d laser range finder. *Chapter* 53, 419–426. doi: 10.1142/9789814525534/0053

Lee, J.-W., and Lee, G.-K. (2005). Gait angle prediction for lower limb orthotics and prostheses using an emg signal and neural networks. *Int. J. Control Autom. Syst.* 3, 152–158.

Lewinger, W. A., and Quinn, R. D. (2011). Neurobiologically-based control system for an adaptively walking hexapod. *Ind. Robot. Int. J.* 38, 258–263. doi: 10.1108/01439911111122752

Lewis, M. A., and Bekey, G. A. (2002). Gait adaptation in a quadruped robot. *Auton. Robots* 12, 301–312. doi: 10.1023/A:1015221832567

Lonini, L., Dipietro, L., Zollo, L., Guglielmelli, E., and Krebs, H. I. (2009). An internal model for acquisition and retention of motor learning during arm reaching. *Neural Comput.* 21, 2009–2027. doi: 10.1162/neco.2009.03-08-721

Maass, W., Natschläeger, T., and Markram, H. (2002). Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Comput.* 14, 2531–2560. doi: 10.1162/089976602760407955

Manoonpong, P., Parlitz, U., and Wörgötter, F. (2013). Neural control and adaptive neural forward models for insect-like, energy-efficient, and adaptable locomotion of walking machines. *Front. Neural Circuits* 7:12. doi: 10.3389/fncir.2013.00012

Mischiati, M., Lin, H.-T., Herold, P., Imler, E., Olberg, R., and Leonardo, A. (2015). Internal models direct dragonfly interception steering. *Nature* 517, 333–338. doi: 10.1038/nature14045

Nachstedt, T., Wörgötter, F., Manoonpong, P., Ariizumi, R., Ambe, Y., and Matsuno, F. (2013). "Adaptive neural oscillators with synaptic plasticity for locomotion control of a snake-like robot with screw-drive mechanism," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on* (Karlsruhe: IEEE), 3389–3395.

Pasemann, F., Hild, M., and Zahedi, K. (2003). "So(2)-networks as neural oscillators," in *Computational Methods in Neural Modeling*, eds J. Mira and J. R. Álvarez (Berlin: Springer), 144–151.

Pearson, K., 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. doi: 10.1177/027836498400300209

Pfeifer, R., Lungarella, M., and Iida, F. (2007). Self-organization, embodiment, and biologically inspired robotics. *Science* 318, 1088–1093. doi: 10.1126/science.1145803

Ren, G., Chen, W., Dasgupta, S., Kolodziejski, C., Wörgötter, F., and Manoonpong, P. (2015). Multiple chaotic central pattern generators with learning for legged locomotion and malfunction compensation. *Inf. Sci.* 294, 666–682. doi: 10.1016/j.ins.2014.05.001

Ren, G., Chen, W., Kolodziejski, C., Wörgötter, F., Dasgupta, S., and Manoonpong, P. (2012). "Multiple chaotic central pattern generators for locomotion generation and leg damage compensation in a hexapod robot," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on* (Vilamoura: IEEE), 2756–2761.

Schilling, M., Paskarbeit, J., Schmitz, J., Schneider, A., and Cruse, H. (2012). "Grounding an internal body model of a hexapod walker control of curve walking in a biologically inspired robot," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on* (IEEE), 2762–2768.

Schröder-Schetelig, J., Manoonpong, P., and Wörgötter, F. (2010). Using efference copy and a forward internal model for adaptive biped walking. *Auton. Robots* 29, 357–366. doi: 10.1007/s10514-010-9199-7

Silva, P., Matos, V., and Santos, C. P. (2012). "Adaptive quadruped locomotion: learning to detect and avoid an obstacle," in *From Animals to Animats 12* (Odense: Springer), 361–370.

Simon, H. (2002). *Adaptive Filter Theory,* Vol. 2. Englewood Cliffs, NJ: Prentice Hall.

Sompolinsky, H., Crisanti, A., and Sommers, H. (1988). Chaos in random neural networks. *Phys. Rev. Lett.* 61:259. doi: 10.1103/PhysRevLett.61.259

Steingrube, S., Timme, M., Wörgötter, F., and Manoonpong, P. (2010). Self-organized adaptation of a simple neural circuit enables complex robot behaviour. *Nat. Phys.* 6, 224–230. doi: 10.1038/nphys1508

Sturm, J., Plagemann, C., and Burgard, W. (2008). "Adaptive body scheme models for robust robotic manipulation," in *Robotics: Science and Systems* (Zurich).

Sussillo, D., and Abbott, L. F. (2009). Generating coherent patterns of activity from chaotic neural networks. *Neuron* 63, 544–557. doi: 10.1016/j.neuron.2009.07.018

Triesch, J. (2005). "A gradient rule for the plasticity of a neurons intrinsic excitability," in *Artificial Neural Networks: Biological Inspirations–ICANN 2005* (Warsaw: Springer), 65–70.

van Vreeswijk, C., and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274, 1724–1726. doi: 10.1126/science.274.5293.1724

Watson, J. T., Ritzmann, R. E., Zill, S. N., and Pollack, A. J. (2002). Control of obstacle climbing in the cockroach, blaberus discoidalis. i. kinematics. *J. Compa. Physiol. A* 188, 39–53. doi: 10.1007/s00359-002-0277-y

Webb, B. (2004). Neural mechanisms for prediction: do insects have forward models? *Trends Neurosci.* 27, 278–282. doi: 10.1016/j.tins.2004.03.004

Wolpert, D. M., Miall, R. C., and Kawato, M. (1998). Internal models in the cerebellum. *Trends Cogn. Sci.* 2, 338–347. doi: 10.1016/S1364-6613(98)01221-2

Zenker, S., Aksoy, E. E., Goldschmidt, D., Wörgötter, F., and Manoonpong, P. (2013). "Visual terrain classification for selecting energy efficient gaits of a hexapod robot," in *Advanced Intelligent Mechatronics (AIM), 2013 IEEE/ASME International Conference on* (IEEE), 577–584.

Zill, S., Schmitz, J., and Büschges, A. (2004). Load sensing and control of posture and locomotion. *Arthropod Struct. Dev.* 33, 273–286. doi: 10.1016/j.asd.2004.05.005

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.