Linear reactive control of three-dimensional bipedal walking in the presence of noise and uncertainty

Mark A Locascio, Joseph H Solomon and Mitra JZ Hartmann

Adaptive Behavior published online 31 August 2012
DOI: 10.1177/1059712312453059

Original Paper

Adaptive Behavior

Adaptive Behavior
0(0) 1–18

DOI: 10.1177/1059712312453059
adb.sagepub.com
SAGE

Linear reactive control of three-dimensional bipedal walking in the presence of noise and uncertainty

Mark A Locascio[1], Joseph H Solomon[1] and Mitra JZ Hartmann[2]

Abstract

Walking control of biped robots is a challenging problem, and improving robustness to noise and uncertainty remains difficult. We recently developed a novel control framework for 3D bipedal walking that we call "linear reactive control." It is linear because control torques are computed as simple weighted sums of sensor states. It is reactive because it depends only on the model's current state. The present simulation study shows that this controller performs reliably in the presence of realistic models of joint actuation, sensor noise, and uncertainty in model and contact parameters. The controller is able to maintain a stable gait in the presence of noisy sensor inputs and low-impedance actuation. It also performs reliably on models with high uncertainty (up to 20%) in measurements of their dynamic parameters and widely varying ground contact parameters. The robustness of this controller to realistic conditions validates this method as a promising avenue for bipedal control.

Keywords

Evolutionary robotics, bipedal walking, biped, robot–environment interactions, reactive control

1 Introduction

Despite advances in robotic bipedal locomotion in recent years, state-of-the-art prototypes remain unable to accommodate significant noise and environmental uncertainty. Control schemes for bipedal locomotion broadly include time based kinematic control, particularly zero moment point (ZMP) methods (Jung-Hoon, Jung-Yup, & Jun-Ho, 2010; Kajita, Nagasaki, Kaneko, & Hirukawa, 2007; Sakagami, Watanabe, Aoyama, Matsunaga, Higaki, & Fujimura, 2002), as well as more sophisticated hybrid zero dynamics and virtual constraint methods (Westervelt, Grizzle, & Koditschek, 2003). Each of these control methods has distinct advantages, but there remain limitations in the ability to generate robust 3D walking in the presence of environmental uncertainty.

The emerging field of evolutionary robotics (Nolfi & Floreano, 2000) provides an alternative approach to the development of stable bipedal walking controllers. We recently used this approach to develop a novel, "linear reactive" controller for biped locomotion over rough terrain (Solomon, Locascio, & Hartmann, submitted; Solomon, Wisse, & Hartmann, 2010). In these simulations, there are no mathematical constraints on the control scheme, no *a priori* knowledge of the terrain topology, and the trajectories are not pre-planned in any way. Control torques are simply computed as weighted linear sums of sensor states. Weights are evolved over thousands of learning trials to allow stable, efficient locomotion without any high-level behavioral constraints. Notably, the control is purely reactive, that is, control torques are calculated based only on the current state of the robot, not on any hidden variables or a pre-computed trajectory.

Machine-learning approaches that involve simulations, however, often suffer from the "reality gap" problem, wherein behaviors learned in simulation differ drastically from those observed in hardware when the controller is transferred (Jakobi, Husbands, & Harvey, 1995; Zagal & Ruiz-del-Solar, 2007). In the case of inherently unstable systems such as bipedal walking robots, the likely result is an immediate fall.

The propensity to fall makes investigations of the "reality gap" problem very difficult in hardware.

[1]Department of Mechanical Engineering, Northwestern University, Evanston, IL, USA
[2]Departments of Biomedical and Mechanical Engineering, Northwestern University, Evanston, IL, USA

Corresponding author:
Mitra JZ Hartmann, Department of Biomedical Engineering, 2145 Sheridan Road, Evanston, IL 60208, USA
Email: m-hartmann@northwestern.edu

A good first step is to use simulations, showing that they can work within the "radical envelope of noise" proposed by Jakobi (1997). These types of simulations are one component of an approach to solving the "reality gap" problem. The goal of the present work was to construct simulations and heuristic procedures for tuning locomotion controllers to operate successfully anywhere within the "envelope of noise." In other words, we aim to show that the controller can operate with realistic actuation, and can successfully overcome sensor noise and uncertainty in model parameters and terrain compliance. Our results show stable walking of a simulated 3D biped over slightly rough terrain even in the presence of significant sensor noise and model uncertainty. This type of approach could potentially be useful in solving the "reality gap" problem (Jakobi et al., 1995; Zagal et al., 2007).

2 Background

Stable 3D bipedal walking is a difficult control problem. Several commercially available bipeds can walk under time-based kinematic control, but they tend to have inefficient and unnatural gaits. Honda's ASIMO (Sakagami et al., 2002) is typically cited as the state-of-the-art in bipedal robotics, as it performs well on flat or nearly flat terrain, and on pre-programmed floor plans. ASIMO, like many bipeds (Jung-Hoon et al., 2010; Kajita et al., 2007; Sakagami et al., 2002), uses a zero moment point (ZMP) control scheme. In ZMP-based methods, joint trajectories may be planned such that the ZMP is kept away from the edges of the support polygon of the feet. If a perturbation disturbs the ZMP, moving it toward an edge of the polygon, a rotation about that edge will occur unless a corrective action is taken to stabilize the ZMP (Vukobratovic & Borovac, 2004). The successes of such methods have been impressive, as this framework gives the designer of the trajectories considerable flexibility. However, these controllers are still sensitive to the environment and terrain (or use future-reference-based preview control; Kajita et al., 2003), requiring flat foot contact surfaces. It also requires strong positional control, and trajectories designed to avoid kinematic singularities such as full knee extension (Kuo, 2007). These singularities may be advantageous to walking, and their exclusion (along with the other constraints) results in a characteristic rigid, unnatural, inefficient walking style. Nevertheless, the maneuverability and versatility of ZMP-based methods is largely unparalleled.

Reduced biped models, including the classic compass-gait walking model (Byl & Tedrake, 2009; Goswami, Thuilot, & Espiau, 1998) and 3D bipeds supported by a boom (Iida, Rummel, & Seyfarth, 2008; Manoonpong, Geng, Kulvicius, Porr, & Worgotter, 2007), are also frequently used to develop or investigate control methods. These models often walk well, even on rough terrain (Iida & Tedrake, 2010), and are valuable for developing and gaining insight into the mathematical underpinnings of bipedal gait. Because they lack several degrees of freedom, however, these models generally do not address certain important control issues such as foot actuation and lateral stability.

Dynamic walkers (Collins, Ruina, Tedrake, & Wisse, 2005; Collins, Wisse, & Ruina, 2001; McGeer, 1990) adhere more closely to humanoid design by exploiting the natural dynamics of swinging limbs. Many of these robots also use bio-inspired actuators including elastic "tendons" that introduce compliance in the system (Grizzle, Hurst, Morris, Park, & Sreenath, 2009; Iida et al., 2008; Kuo & Chiou, 2009). They are capable of walking efficiently and stably with a natural-looking gait (Collins et al., 2005; G. A. Pratt, 2002), but their reliance on passive dynamics makes them very sensitive to environmental disturbances, while their compliant actuators make control more difficult by traditional methods.

Previous work has recognized the potential of simple network models for sensorimotor control, but has shown only that such methods are feasible, and has not imposed significant noise or environmental uncertainty. For example, Reil and Husbands developed a feedforward control scheme based on central pattern generators that enables 3D walking on flat terrain in simulation (Reil & Husbands, 2002). Ono et al. developed a simple but effective control scheme in which the applied hip torque is proportional to the swing knee angle and the knee joint is linearly damped, allowing a 2D robotic prototype to traverse a shallow slope (Ono, 2001). Paul used a two-layer feedforward neural network with nonlinear activation functions and an incomplete state vector (no velocities) as the input to achieve simulated 3D walking on flat ground (Paul, 2005). Manoonpong et al. use a hierarchical, feedforward (reactive) neural network control model with an online learning mechanism to allow a 2D robot (supported by a boom) to traverse terrain with variable slopes (Manoonpong et al., 2007). Vaughan's work was similar to ours, but used a multilayer perceptron with nonlinear activation functions for control, and central pattern generators to regulate the gait. His controller was also capable of tolerating some variability in the dynamic parameters (Vaughan, Di Paolo, & Harvey, 2004). One significant issue that was encountered in these studies (except Ono et al.) was the need to tune several parameters to enable stable walking. To this end, techniques from evolutionary robotics were found to be effective.

Evolutionary robotics (Nolfi et al., 2000) is a metaheuristic optimization approach to evolving the controllers (and sometimes morphology) of robotic systems (Bongard, 2011). The approach is based on the process of natural selection, in which a population of candidate solutions (parameter sets) are tested for performance

(fitness), and the best ones are retained and modified in an effort to generate small improvements. The process is repeated until a high-performance solution is found. Two obvious benefits of this approach are that the controller requires no formal analysis to tune, and that it can operate over very large parameter spaces.

We recently used an evolutionary approach to develop a "linear reactive" controller for an actuated dynamic walker that maintains efficient performance on rugged terrain (Solomon et al., submitted; Solomon et al., 2010). The controller demonstrated the ability to stably and efficiently traverse rough terrain in both 2D and 3D, exhibiting a remarkably fluid and human-like gait pattern. However, to implement this controller in hardware, we must show that simulation results can operate within the full range of uncertainties that the robot may plausibly encounter, i.e., within Jakobi's "radical envelope of noise" (Jakobi, 1997). The present work uses evolutionary robotics techniques to develop a linear reactive controller for a 3D biped that is robust to variability in body parameters and various conditions of noise and uncertainty.

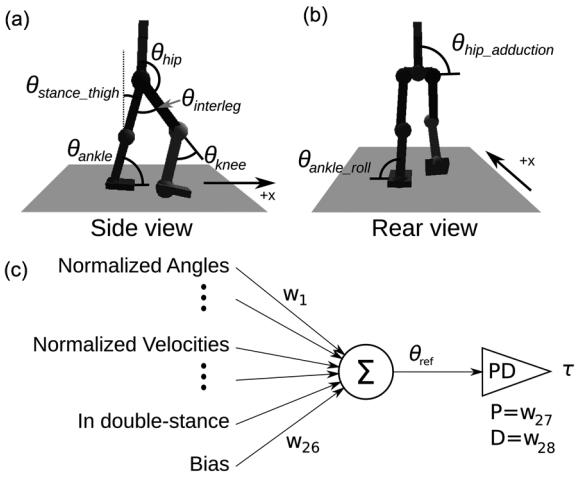

Figure 1. (a) and (b) The dynamic model, showing the hip adduction axis and the actuated and unactuated axes of the ankles. All other joints are indicated by a sphere. Although the terrain looks smooth at the scale shown, there is actually a small amount of (irregular) roughness. (c) A single actuator's control network, in which reference angles are generated as the weighted sum of the 26 inputs (including the bias), and the applied torque is computed as in PD control. There are seven of these networks, one for each actuator.

3 Methods

3.1 Biped and terrain model

The model robot in this study was developed using the Open Dynamics Engine (Smith, 2012). The model uses the dynamic and geometric specifications of the Flame prototype biped developed by Wisse et al. (Hobbelen, de Boer, & Wisse, 2008; Wisse, 2008).

The robot has seven actuators: two ankles, two knees, two hips, and a single actuator controlling the lateral hip adduction angle in the coronal plane. A mechanical linkage ensures that the torso always bisects the hip adduction angle, and that both legs are always equally adducted/abducted. The robot has encoders at each joint, an IMU in the torso, and contact switches on the heel and toe of both feet. There are thus 25 sensory variables that describe the state of the robot: position and velocity of each encoder at each actuated joint (14), position and velocity of the unactuated ankle roll for both feet (4), IMU values for the torso's roll, pitch, and yaw (3) as well as the temporal derivatives of these IMU values (3), and a single binary value indicating if both feet are in contact with the ground (double-stance). The controller receives no visual information, nor any *a priori* knowledge of the terrain. The controller has information about environmental interaction only through the binary double-stance variable, synthesized from the foot contact states. A schematic of the model is shown in Figures 1a and 1b.

The torso pitch and roll (not shown) are relative to gravity. The torso yaw is relative to the x-axis. Note that θ_{hip}, θ_{knee}, and θ_{ankle} are relative to the "parent" link. θ_{hip} is the thigh angle relative to the torso. θ_{knee} is the shank angle relative to the thigh. θ_{ankle} and θ_{ankle_roll} are the foot angles relative to the shank. For each of these, there is an angle for the joint on the stance leg, and one for the angle of the joint on the swing leg. The hip adduction angle is the angle of the hip link relative to the torso in the coronal plane.

Rather than providing the hip joint angles to the controller as angles relative to the torso, the stance and swing hip measurements are cast in the gravity frame. Thus, the controller receives the stance thigh angle relative to gravity and the interleg angle (the angle between the thighs in the sagittal plane) in place of direct measurements of the stance and swing hip angles. These angles are shown in Figure 1a and are computed as:

$$\theta_{stance\_thigh} = \theta_{stance\_hip} - \theta_{pitch}$$
$$\theta_{interleg} = \theta_{swing\_hip} - \theta_{stance\_hip}$$

The velocity of each measurement is computed similarly. The torso-frame hip measurements are replaced with the gravity-frame measurements, keeping the number of sensory variables at 25.

In many studies, perfectly flat, firm ground is used both in hardware and in simulation (Collins et al., 2001; Iida et al., 2008; Morimoto & Atkeson, 2007; Plestan, Grizzle, Westervelt, & Abba, 2003). In this study, we intentionally generated slightly rough terrain, represented as a triangle mesh with 10 cm spacing between the vertices. Terrain roughness prevents the emergence of unrealistic or dangerous behaviors, such as walking with near-zero swing foot height, and it also ensures that all individuals are robust to at least a small amount of unevenness in the terrain.

Table 1. List of sensory connections in the reduced controller

Actuator	Sensory connections to that actuator
Hip adduction	Torso roll, torso yaw, hip adduction (self-connected), stance ankle, torso roll velocity, torso yaw velocity, hip adduction velocity, stance ankle velocity
Upper body	Stance ankle, swing knee velocity, stance ankle velocity
Interleg angle	Stance thigh, swing knee, stance ankle
Stance knee	–
Swing knee	–
Stance ankle	Torso pitch, stance knee, swing ankle, stance thigh velocity, interleg velocity, stance knee velocity, swing ankle velocity
Swing ankle	–

The bias and the P and D gains were always included in the weights of every network.

Each generation of individuals was tested on a unique, randomly generated terrain. Each terrain was generated using a two-dimensional random-walk algorithm, with a step-to-step standard deviation of the resulting terrain height equal to ~0.36% of the leg length. Importantly, note that the terrain unevenness is a property distinct from the contact model of the terrain (see section 3.4.4).

3.2 Controller

There are seven actuators in the biped model (stance/swing ankle, stance/swing knee, stance/swing hip, and hip adduction), each with its own control network. Figure 1c schematizes the control network for a single actuator. We first describe the "fully-connected" controller and then describe the reduced set of weights used in the present work.

In the fully-connected configuration, each actuator's output is computed from 28 values: 25 state variables, plus the P and D gains and one constant bias term. A bias term is standard in machine learning algorithms as a way to provide a constant, non-zero input to a function (e.g., the intercept term of a linear regression). Thus, there are a total of $28 \times 7 = 196$ weights connecting the inputs and actuators. Of these, the 21 values defining the biases and PD gains are treated as weights in the evolution, but are not considered "sensory" connections (of which there are therefore $196 - 21 = 175$). On each time step, the sensor state of the model is measured. Values with fixed limits, such as joint angles, are shifted and scaled to $[-1, +1]$. Values without fixed limits, such as the torso orientation angles and all velocities, are scaled to an approximate range of $[-1, +1]$ based on the maximum and minimum kinematic values recorded during preliminary walking trials. A weighted sum of the ϕ_i values (the normalized sensor inputs) and the bias, 1.0, is computed using the weight vector $\mathbf{w}$ structured as in Figure 1c, and the output is used as a reference angle, θ_{ref}.

$$\theta_{ref} = \sum_{i=1}^{25} \phi_i w_i + 1.0 \cdot w_{26}$$

Note that – in contrast to traditional neural networks – there is no "activation function" applied at the output. The torque is then computed from θ_{ref} using PD control with this value as the reference position, and zero as the reference velocity, in order to apply damping to the closed-loop system. Although there are many inputs to the controller, it is mathematically no more complex than a weighted sum. This stands in contrast to other similar work (Paul, 2005; Vaughan et al., 2004), in which a multilayer perceptron is used, often in conjunction with an internal oscillator.

The hip actuators are re-framed in terms of upper-body and inter-leg control in the same way as the hip sensors. This convention is useful because it effectively isolates control of the upper-body angle from control of the inter-leg angle. The stance hip motor then applies torque as the "upper-body actuator," while the swing hip motor applies torque as the "inter-leg" actuator.

In the present study, not all 175 sensory connections were used. Instead, in all parts of this work, we reduced the controller to 21 sensory connections, which were chosen based on previous work that had determined these connections to be important for 2D walking models (Solomon et al., submitted). Table 1 shows the 21 sensory connections that were used in this work. Note that the bias and P and D gains account for an additional 21 weights, yielding a total of 42 weights that needed to be optimized.

Note that, while this approach is couched in terms of a PD framework, we do not enforce "stiff" control. That is, the reference angles may not be accurately tracked. Instead, an evolutionary algorithm (EA, see section 3.3) tunes the P and D gains to whatever values achieve high-fitness behavior. So, even though the swing knee reference angle is always a constant (as shown in Table 1, it is connected only to the bias input), the swing knee may follow a smooth curve using low gains rather than tracking a fixed reference angle if it is found that such behavior is appropriate. This low-impedance control prevents the high energy use common in stiff positional control, because it does not resist the passive dynamics of the system. This permits more natural complex trajectories with fewer connections.

3.3 Evolution of successful controllers

Because of the high dimensionality of the search space (42 weights), the weights were determined using an evolutionary algorithm. The vector **w** is considered the genetic material of each individual, which is evolved over several hundred generations. Each generation undergoes three steps: diversity-generating breeding and mutation, fitness evaluation, and survivor selection.

In the first step, an initial population of 30 parents is generated and bred to produce 150 children. Two parents are selected at random, and each weight is chosen randomly from either parent ("discrete recombination"; Eiben & Smith, 2003) to produce a child genome. Then, small random mutations are made to each value in the child genome by adding a Gaussian random value with a standard deviation of 0.005. In 5% of mutations, the standard deviation is increased to 0.1 to allow for the possibility of occasional large steps along particular dimensions of the search space. The range of the connection weights are not constrained, so those weights can take any value. The P and D gains, which are treated like weights during the evolution, are limited to positive values.

The fitness function used in the evaluation step is the distance traveled on a fixed amount of energy. That is, the fitness is the distance traveled throughout the simulation, where the end of the simulation is triggered either by a fall (when anything but the feet comes in contact with the terrain), by flight (both feet in the air), or by the exhaustion of the energy supply, whichever comes first. This places emphasis on stability in early generations when individuals fall before using all their energy. In later generations, individuals are stable enough to use their entire energy supply, so "fitter" individuals must be more efficient to travel farther. The fittest individuals must therefore walk stably and efficiently in order to compete for survival. Here, we use "torque cost," denoted C_τ, as our energy quantity (i.e., $C_\tau = \int |\tau| \, dt$ rather than physical energy $E = \int \tau \cdot \omega dt$) (Srinivasan, Westervelt, & Hansen, 2009), because the inclusion of ω, the joint velocity in the cost function, was found to produce very stiff walking behavior that attempted to minimize ω for all joints.

After all 150 child weight vectors have been evaluated, the procedure moves to the final step, survivor selection. Here, we simply rank every child according to its fitness value. The 30 fittest individuals become the parents for the next generation, and the procedure repeats. For simple tasks such as walking without any noise or varying parameters, only 200–300 generations may be needed for fitness to converge to a steady value. For more difficult control tasks such as those described in the next section, upwards of 3000 generations may be required.

We took three precautions to reduce the bootstrapping problem, in which a poorly chosen initial condition or overly difficult behavior prevents evolution from proceeding at all. First, the simulations were started in a configuration compatible with a normal walking gait, in order to prevent unrealistic gait-initiation behaviors. Every simulation began by placing the model in a state similar to the beginning of the "initial swing" phase of gait (Uustal & Baerga, 2004). Second, the initial population of 30 parents was generated by hand-tuning a network consisting only of bias, P, and D values that were sufficient to start the evolution. Uniformly distributed mutations to the weights of this initial set produced a diversified first generation of individuals. These starting behaviors only involved basic motions such as swinging the swing leg forward by increasing the inter-leg angle. No actual walking behavior was encoded in these hand-tuned networks in order to avoid biasing the evolution toward a particular behavior. Third, when training for parameter uncertainty or terrain compliance, we started uncertainty/compliance variability low, and increased variability as the evolution proceeded. For effects that were binary (on or off) such as sensor noise or realistic actuation, the effect was included from the first generation.

3.4 Sources of noise and uncertainty

The goal of the present work was to demonstrate robustness to the following four real-world sources of noise and uncertainty.

3.4.1 Realistic actuation. By default, dynamic simulations of articulated structures apply exactly the requested torque to a joint, essentially modeling a high-impedance, infinitely powerful motor. In contrast to this model, series-elastic actuators (SEAs; J. Pratt, Krupp, & Morse, 2002) are becoming increasingly popular in hardware (Grizzle et al., 2009; Iida et al., 2008; Kuo & Chiou, 2009), as they make the system more compliant and allow direct measurement of applied torque. The tradeoff, however, is that this actuation is a higher-order system than a rigid connection. We modeled actuation in the joints as SEAs by including a virtual spring between the rotor mass and the joint itself. The rotor is free to move independently of the joint, but the two are linked according to the spring equation, $\tau = k(\theta_{motor} - \theta_{joint})$. Values of k are listed in the Appendix, and were chosen based on values approximating the springs of *Flame* (Hobbelen et al., 2008). A damping term was also applied similarly. Additionally, the power output of the simulated motors is limited to 90 W to match the maximum power output of the real motors on *Flame* (Hobbelen et al., 2008). This is implemented simply by determining a maximum torque τ for a given joint velocity ω according to $P = \tau\omega$.

3.4.2 Sensor noise. The controller must be capable of operating with realistic sensing. In this case, we show that the controller is insensitive to errors in the sensor measurements. Simulation software will report simulated joint angles to machine precision, but in reality, a robotic platform would only have access to quantized values from the output of a rotational encoder. We quantized all sensor values in the simulation to 7500 counts/revolution in quadrature, an equivalent of 30000 counts/revolution for each joint. Similarly, the torso orientation of a real robot is usually measured using a set of MEMS gyroscopes and accelerometers, or a prepackaged inertial measurement unit. As a reference, we used the specifications of an Xsens MTi unit with a 100-Hz read rate and $0.05°/s/\sqrt{(Hz)}$ noise. This was modeled in simulation by adding Brownian noise with these same characteristics to the torso orientation values. Given that velocities are typically calculated rather than measured, we ignore ODE's internal access to joint velocities and instead estimate them based on a buffer of quantized angle readings using the adaptive windowing technique of Janabi-Sharifi, Hayward, and Chen (2000).

3.4.3 Uncertainty in model parameters:. Although the CAD drawings of a mechanism allow calculation of the values for the masses, center of mass positions, and inertia tensor components, in practice these will differ from the properties of the machined components for a variety of reasons. For example, the electrical components are not included in the drawings, which therefore exclude parts with significant mass, such as an onboard computer, batteries, and cables. The actual dynamics will therefore not be accurately reflected by the calculated values, and may even vary between seemingly identical copies of the mechanism. In order to develop controllers that are robust to such uncertainty, we require that the controller be insensitive to some variation in these dynamic parameters.

In total, we permitted 86 model parameters to vary simultaneously, gradually increasing from ±5% variation from the nominal CAD values to a maximum of ±20% variation. The first 300 generations of the evolution have the lowest level of variation, ±5%. This comparatively narrow range and the long evolution time (more generations than subsequent variation ranges) bootstrap the evolution for higher ranges in later generations. After these initial 300 generations, the variation range is increased by 5 percentage points every 100 generations. Thus, the first 300 generations have ±5% variation, generations 300–400 have ±10%, 400–500 have ±15%, and generations 500–600 have ±20%. The varied parameters include the parameters of nine parts of the robot (torso, left & right hip, left & right thigh, left & right shank, and left & right foot). Each of the nine parts has 10 parameters that can vary: three components of the center of mass position, six components of the inertia tensor, and the mass. The values of the masses were shared between the left- and right-side parts, removing one possible parameter for each pair of hips, thighs, shanks, and feet.

3.4.4 Uncertainty in terrain compliance:. The ground reaction forces generated during walking are difficult to quantify and represent large sources of uncertainty. In particular, it has been shown that changes in contact stiffness can change the global characteristics of the walking gait of passive dynamic walkers (Qi, Wang, & Li, 2011). We therefore evolved controllers to be robust to variations in terrain compliance, while maintaining a small amount of randomized roughness to prevent fitting to perfectly flat ground. If we consider a firm contact model to be a spring-damper system with proportional and derivative gains $K_p = 100{,}000$ and $K_d = 5{,}000$, then the compliance can be adjusted by scaling these values by some factor between $10^{-1.5}$ and 10^0. Terrain contact was randomly varied by changing this scale factor, and this was implemented by adjusting the constraint force mixing parameter and the error-reduction parameter in ODE in order to achieve the equivalent K_p and K_d.

4 Results

We challenged the controller with realistic actuation via series-elastic actuators (SEAs) driven by finite-power motors, realistic sensor noise, large parameter uncertainties in the dynamic model, and variations in terrain compliance. We first present results for each effect, tested independently. We then present results of a final test in which the controller was trained as all effects were accumulated, in order to demonstrate the operation of the biped controller inside of the envelope of noise (Jakobi, 1997).

4.1 Realistic actuation: robustness to the addition of SEAs and finite motor power

For the given model, the most important deviation from a perfect rigid-body simulation is the addition of SEAs. As described in *Methods*, these actuators are modeled in the simulation by including an additional mass (representing the motor mass) that moves with the joint through the influence of a virtual spring. When a torque is calculated, it is applied only to the motor mass, not directly to the joint itself. This addition radically changes the dynamics of the model, making it more compliant and also harder to control. In addition, we limited motors to an output of 90 W, a feasible motor size for use in an energy-efficient dynamic walker such as *Flame* (Hobbelen et al., 2008).

Initial work showed that adding realistic actuation to all joints simultaneously resulted in a bootstrapping problem that failed to evolve walking behaviors. Recall that the initial "generation 0" parents are biased as described in section 3.3, but are not so heavily biased that they could be considered competent in any way. The bootstrapping done for those is completely insufficient for a walker that starts with all pairs of SEAs included. Therefore, the series-elastic elements were added one at a time, until all joints included SEAs. After adding an SEA-controlled joint, the evolution proceeded for 800 generations, a duration sufficient for complete evolution of the rigidly actuated walkers. In the following tests, SEAs were added to the ankle joints first, then the knees, and finally the hips.

Figure 2a shows the evolution with SEAs added in this manner compared with a rigidly actuated model. The additional compliance and difficulty of control significantly slows down the evolution. In fact, several (approximately seven) attempts were required before we were able to get evolution to proceed at all (data not shown). The rigidly actuated model (labeled "No SEAs" in Figure 2a) only takes ~125 generations to converge to a solution that provides high fitness; the best individuals of that generation can walk about 50 m on 1000 N·m·s of torque-cost.

An additional test of the best individual in the final generation of the rigidly actuated system ("No SEAs" in Figure 2a) was used to compute the specific mechanical cost of transport. This 151.64 N walker does 4511.60 J of positive mechanical work (not torque cost) to walk 81.02 m. These values correspond to a specific mechanical cost of transport of $c_{mt} = 0.37$, which compares favorably with a ZMP-controlled robot such as ASIMO, whose estimated value of $c_{mt} = 1.6$ (Collins et al., 2005).

When ankle SEAs are added, the time to evolve is longer (~600 generations) and the fitness plateaus near 40 m. Addition of SEAs to the knees does not impose a significant additional efficiency penalty. The final pair of SEAs introduced at the hips again substantially reduces efficiency, as the final walker with all SEAs included can walk only a little over half as far on the same amount of energy as a perfectly actuated walker (~30 m on 1000 units of torque-cost).

Figure 2a shows a dramatic drop in fitness after each set of SEAs is added. However, this does not mean that all productive changes to the weights are lost. One might draw an analogy to the replacement of a person's left ankle with a robotic prosthetic. The person's "walking fitness" would initially be near zero, with fitness slowly increasing with time and therapy, but his entire right leg, as well as his left hip and left knee, can all still operate as usual. For five out of six leg joints, walking behaviors learned as a toddler still apply, but additional training is needed to get the entire system's fitness back to its former value.

Similarly, when we add a new pair of SEAs, the simulation is largely bootstrapped, save for one pair of actuators. This is enough for the individuals in the first several generations to fail quickly, because fitness is a measure of how well the entire system performs, not just one set of actuators. Thus, a pair of actuators with altered properties will reduce the entire system's fitness to nearly zero, even if the others function flawlessly.

Figure 2b illustrates the fitness distributions for the best fully evolved individuals. The 10 best walkers from each of the last 50 generations were all evaluated, for a

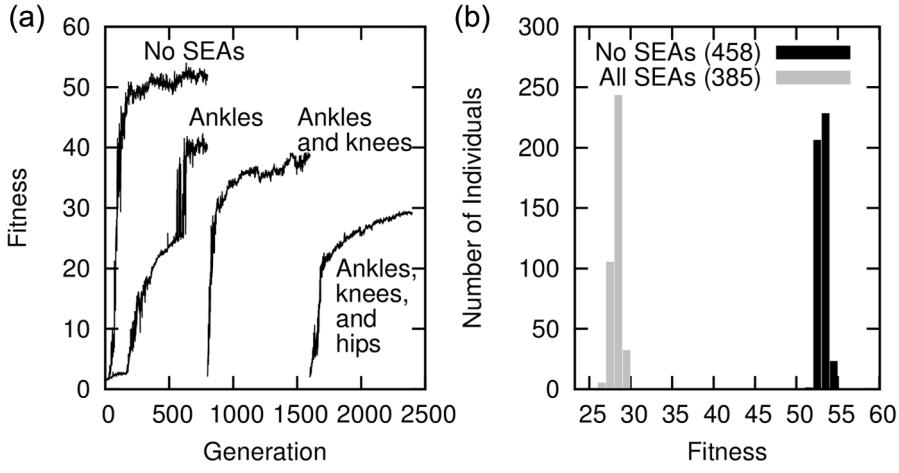

Figure 2. (a) The maximum fitness in each generation with and without series-elastic actuators (SEAs). The addition of SEAs makes evolution more difficult, and the controller less efficient. (b) An evaluation of the 10 best walkers from each of the last 50 generations with and without SEAs and power limits. Only walkers that used all of their energy are shown (458 with perfect actuation, 385 with SEAs and finite power). The linear reactive controller performs very well even for realistically actuated control, though efficiency is lower.

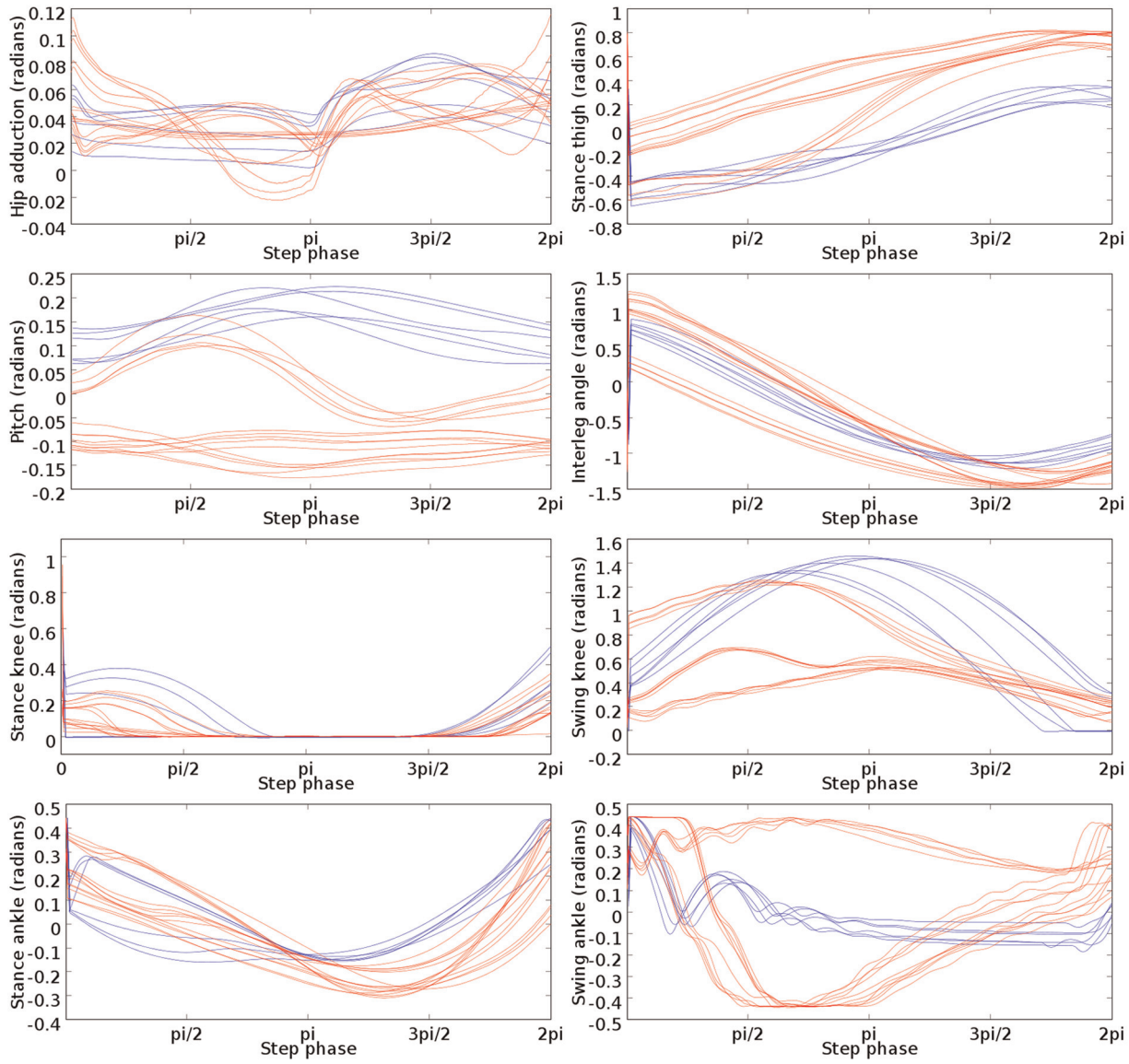

Figure 3. Kinematic trajectories of walkers trained with perfect actuation (blue) compared with those trained with series-elastic actuators (SEAs) (red).

total of 500 individuals. Those that did not fall down (and therefore used all of their energy) are shown; 458/500 walkers with perfect actuation completed their evaluations compared with 385/500 with SEAs and finite power. The distribution of fitnesses for those with perfect actuation is centered around 53 m, whereas those with SEAs and finite power are much lower, around 28 m on 1000 N·m·s of torque-cost. This corresponds roughly to the values in the last 50 generations of each population in Figure 2a. It is clear that including these effects severely limits the efficiency. This result is not surprising, because infinite-power motors and stiff actuation make a simpler control problem. However, the results indicate that linear reactive control is in fact quite capable of producing stable walking even for the greater complexity of a realistically actuated control task. Supplemental videos 1 and 2 demonstrate walking without and with realistic actuation, respectively. Most noticeable is the increased maximum interleg angle (0.77 vs. 0.62 radians), decreased maximum swing knee angle (0.6 vs. 1.3 radians), and reduced walking speed (1.5 vs. 2.6 m/s) when SEAs are present (all significant, with p-values <0.005 using a two-tailed t-test). Note that all videos are recorded at half speed to show detail.

The kinematic trajectories of several important degrees of freedom are shown in Figure 3. To create this figure, three independent evolutions of "trained" walkers and six independent evolutions of "untrained" walkers were run. The five best individuals in the final generation of each trained run and the best individual in the final generation of each untrained run were evaluated, then the trajectories were divided by step and

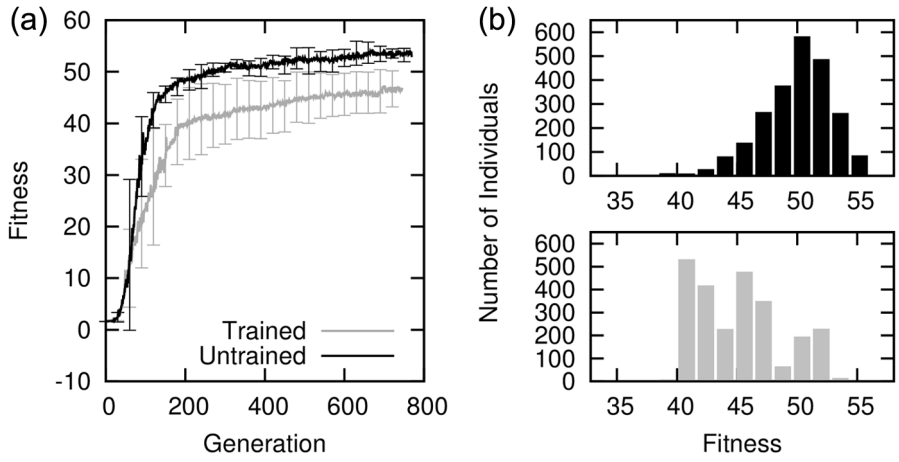

Figure 4. (a) The evolution of controllers with and without sensor noise (curves averaged over six runs). Including sensor noise does not significantly slow down the evolution, though it does result in a slight efficiency penalty. (b) An evaluation of the 10 best walkers in the last 50 generations of six runs. Only walkers that used all of their energy are shown. The sets cope with the presence of sensor noise equally well, as indicated by the very similar number of individuals that walk to exhaustion (2309/3000 untrained vs. 2497/3000 trained).

averaged over the run. Five trained individuals are plotted for each of the three runs to show the tight grouping within a single run. Only a single (untrained) individual is plotted for each of the six untrained runs. To ensure that that we characterized the trajectory without edge effects (slightly different starting behaviors, tripping at the end of an evaluation, etc.), 15 steps were removed at the beginning and five at the end. The length of each step was normalized from 0 to 2π.

Each degree of freedom shows three distinct sets of trajectories for individuals trained with SEAs (red), one for each run. This means that there are several equally fit walking strategies when SEAs are included. The swing knee angles are uniformly lower (implying a straighter swing leg), which in turn requires higher hip adduction angles to prevent toe-scuffing. The upper-body remains more upright, or even leaned slightly backward, to allow the thighs to swing higher. The stance thigh angle starts closer to zero to prevent the swing leg from dragging, and the interleg angle reaches its minimum slightly later, so the swing foot retracts less at the end of the step. Note that although the swing leg is highly variable, it contributes very little to the gait.

4.2 Realistic sensing: robustness to sensor noise

The behavior of any control scheme is highly dependent on the quality of sensor input that it receives. In simulation, the sensor inputs to a controller are known to machine precision, which is clearly unrealistic. To model a realistic amount of sensor noise for all measurements as described above, we used the characteristics of the industry-quality rotational encoders and inertial measurement unit found in *Flame*. These characteristics were used to quantify the amount of noise to add to each sensor, including the quantization of the encoders and the Brownian noise in the IMU.

We evolved six populations of controllers in the absence of sensor noise (i.e., with "perfect" sensing), and six populations in the presence of this noise ("untrained" and "trained" respectively, referring to whether or not they were trained to cope with these sources of noise).

Figure 4a shows the average maximum fitness per generation for each set of populations, illustrating that they evolve equally quickly, though there is a ~13% decrease in maximum fitness between the two populations. Again, the 10 best walkers in the last 50 generations were taken from each "trained" and "untrained" population to make two populations of 3000 high-performance individuals. Both populations were then tested *with* sensor noise. As shown in Figure 4b, the population trained in the presence of sensor noise has only a few more that survived to exhaust their entire energy supply (2309 untrained vs. 2497 trained). The population that was not trained in the presence of sensor noise does have slightly greater efficiency, however, as indicated by the larger mean distance traveled on the same amount of energy.

Overall, the controller performs well in the presence of realistic sensor noise, even when it has not been trained to do so. In other words, the linear reactive controller is inherently robust to noise of the magnitude that would be expected from industry-quality sensing hardware. Supplemental video 3 demonstrates walking in the presence of sensor noise, though the gait does not differ appreciably from video 1.

4.3 Robustness to variations in parameters of the model

Robustness to model parameter variations is among the most important properties of a successful walker. Although CAD drawings can provide highly accurate dynamic parameters of certain components of a bipedal robot, values for other components such as cabling, electronics, and computer hardware are difficult to measure accurately. Similar situations will also arise whenever any computational model differs from its corresponding hardware. We can mitigate the effects of this uncertainty in our walkers by training the controllers to be robust to uncertainty in all 86 model parameters.

Figure 5 shows results from simulations in which the amount of uncertainty in model parameters is slowly increased throughout the evolution. In the first 300 generations, all 86 model parameters are randomly distributed (uniformly) in a range within ±5% of the CAD value in order to avoid a bootstrapping problem early in the evolution. The next 100 generations increase uncertainty to ±10%. Initially, the average fitness (shown as a Bezier-smoothed black line) decreases dramatically due to the higher range of values encountered. Over the course of 100 generations, the controllers are trained to perform better in the presence of this higher range, and fitnesses improve. At generation 400, the uncertainty increases from ±10% to ±15%, with another corresponding drop in fitness, and then an improvement over several generations. Finally, at generation 500, the uncertainty increases one more time, from ±15% to ±20% of the nominal values. Beyond generation 400 (≥10% variation from nominal), the range is large enough to cause unreliable performance, which manifests as high variance in the maximum fitness values.

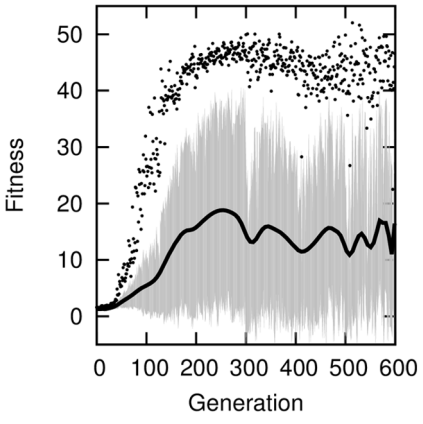

Figure 5. As the amount of uncertainty is slowly increased, the distribution of fitnesses becomes more variable. The maximum fitness in a generation is denoted by a black dot, the Bezier-smoothed average is shown as a black line, and the gray range is ± 1 standard deviation from the average.

We next aimed to compare the performance of walkers trained in the presence of parameter variations (hereafter the "trained" population) with those trained in their absence ("untrained"). However, we needed to ensure that the individual undergoing testing really was the best individual, and not just the best for the one particular set of parameters on which it happened to be tested. Therefore, we conducted a tournament to find the best overall individual in the last 50 generations of the untrained walkers. The single best individual from each of the last 50 generations were each tested using 10 different parameter sets, and the individual with the highest average fitness over those 10 trials was declared the winner. The tournament winner—that is, the "best untrained individual"—was then evaluated using 200 more sets of parameters. Another tournament found and evaluated the "best trained individual" in the same way.

Figure 6a shows the distribution of fitnesses for the best untrained (left) and trained (right) individuals. Comparing the two histograms, it is clear that when untrained walkers are exposed to 20% parameter variation, the results are disastrous. Without training, the controllers are generally unable to accommodate such large parameter variations. The best untrained walker fell before exhausting its energy supply in 182/200 trials, often within the first 5 m. In contrast, the best trained walker performed extremely well, falling in only 51/200 trials. This demonstrates that the linear reactive controller design is capable of performing well even when the controller is trained with an inconsistent model. It simply needs to be trained to be robust to the level of uncertainty present in the measurements of these parameters.

In summary, Figure 6a shows that without training, walkers clearly cannot accommodate uncertainty in dynamic parameters. Additionally, the efficiency of the two populations of walkers was investigated to determine the extent of the tradeoff between robustness and efficiency, and those results are in Figure 6b. The best 10 walkers from each of the last 50 generations of the trained and untrained runs were evaluated *without* parameter variations, so as not to include any gains or losses from lighter or heavier parts. The distributions of fitness are shown in Figure 6b. Similar numbers of walkers complete the run without falling: 458/500 untrained walkers (top), and 443/500 trained walkers (bottom). Only these successful walkers are plotted. The average efficiency (i.e., the distance traveled on a fixed amount of energy) also changes very little between the untrained and trained walkers. However, the standard deviation is significantly wider for those trained in the presence of parameter variations. Supplemental video 4 demonstrates the walking behavior of the best

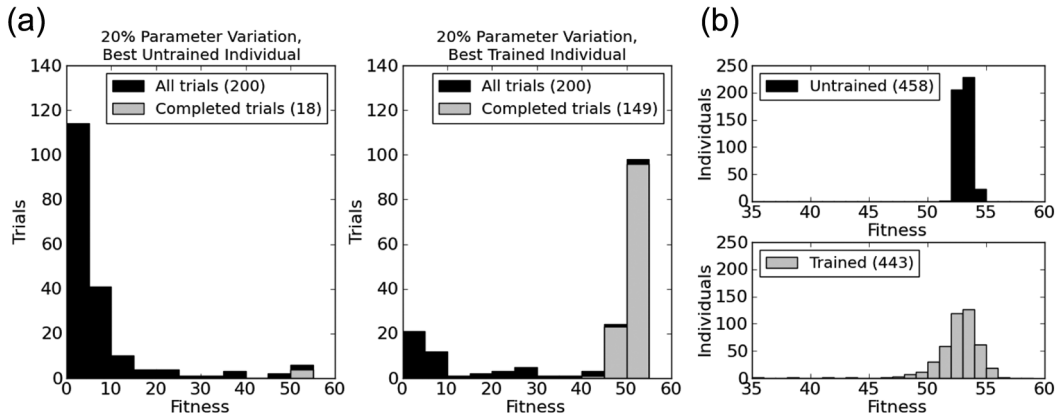

Figure 6. (a) The performance of the best trained and untrained individuals. The untrained individual fails quickly when dynamic parameters differ compared with its training set, whereas the trained individual performs well for a large number of parameter sets. (b) The mean efficiency stays the same after training for robustness to dynamic parameter variations, but the standard deviation increases significantly.

trained controller, which again differs only slightly from the gait of video 1.

The kinematic trajectories of several important degrees of freedom are shown in Figure 7, computed as described in section 4.1. Here there are six independent runs of individuals trained with parameter variations, shown in red. Only the best individual from each run is shown. Unlike the trajectories resulting from the SEA runs, here the six runs found much more similar walking behaviors. For example, notice the very tight grouping of trajectories for the interleg angle, the swing knee angle, and the stance ankle angle. Compared with individuals evolved without variations in their parameters ("untrained" blue trajectories), the swing knee angles of the trained individuals peak earlier and lock before the end of the step. This means that the stance knee is more reliably locked at the beginning of a step. Although there is larger distribution in the upper body and stance thigh angles, the shapes are very uniform. These observations imply that the key to stable walking in the presence of high uncertainty is to develop behaviors that are highly repeatable and, if possible, exploit invariant properties, such as joint angle limits.

4.4 Robustness to variation in terrain compliance

The final variation that we introduce is the stiffness of the ground contact model. Although it has been shown previously that this linear reactive control scheme can be trained to maneuver on rough terrain for a 2D biped model (Solomon et al., submitted), the roughness refers only to the topology, not the compliance of the ground contact. Here, we show that the controller can also be trained to cope with uncertainty in the ground's compliance, so that the trained controllers will not be reliant upon particular behaviors in impacts or ground reaction forces. This also mitigates concerns about the modeling of ground contact, which is notoriously difficult and unreliable.

Collisions in the simulation are essentially modeled as temporary constraints at the contact points, and like other constraints, the spring (K_p) and damper (K_d) constants can be specified by appropriately setting the constraint force mixing and error reduction parameters of ODE. However, moderate changes to K_p and K_d ($<$20%) like in the previous tests do not significantly change the effective compliance of the surface. Instead, the proportional and derivative terms must be scaled by much more extreme factors spanning two orders of magnitude in order to achieve values with truly different contact properties. We conducted a total of six runs with the spring and damper constants scaled as shown in Figure 8, and plotted the average fitness of the 30 best individuals. Even after training, very compliant terrain is difficult to manage. The average fitness falls off quickly when the spring and damper terms are attenuated by more than 1/10. In order to test a challenging yet feasible range, we chose scale factors randomly between $10^{-1.5}$ (0.032) and $10^{-0.5}$ (0.32), in the region of this drop-off.

Similar to the case of dynamic parameter variations, the robustness of the untrained controller is extremely poor (Figure 9a, left). As before, a tournament was conducted to find the overall "best untrained" and "best trained" individual, and then the two were evaluated on 200 different terrain contact models in the range described above. In the model parameter study, 86 parameters were varied, whereas here only a single parameter was varied. Thus, we see a much more bimodal distribution in the untrained individual, which fell immediately in about half of the trials (corresponding to the lower scale factors), and completed only trials with larger scale factors. The trained walker fared much better, walking to exhaustion in 177/200 trials

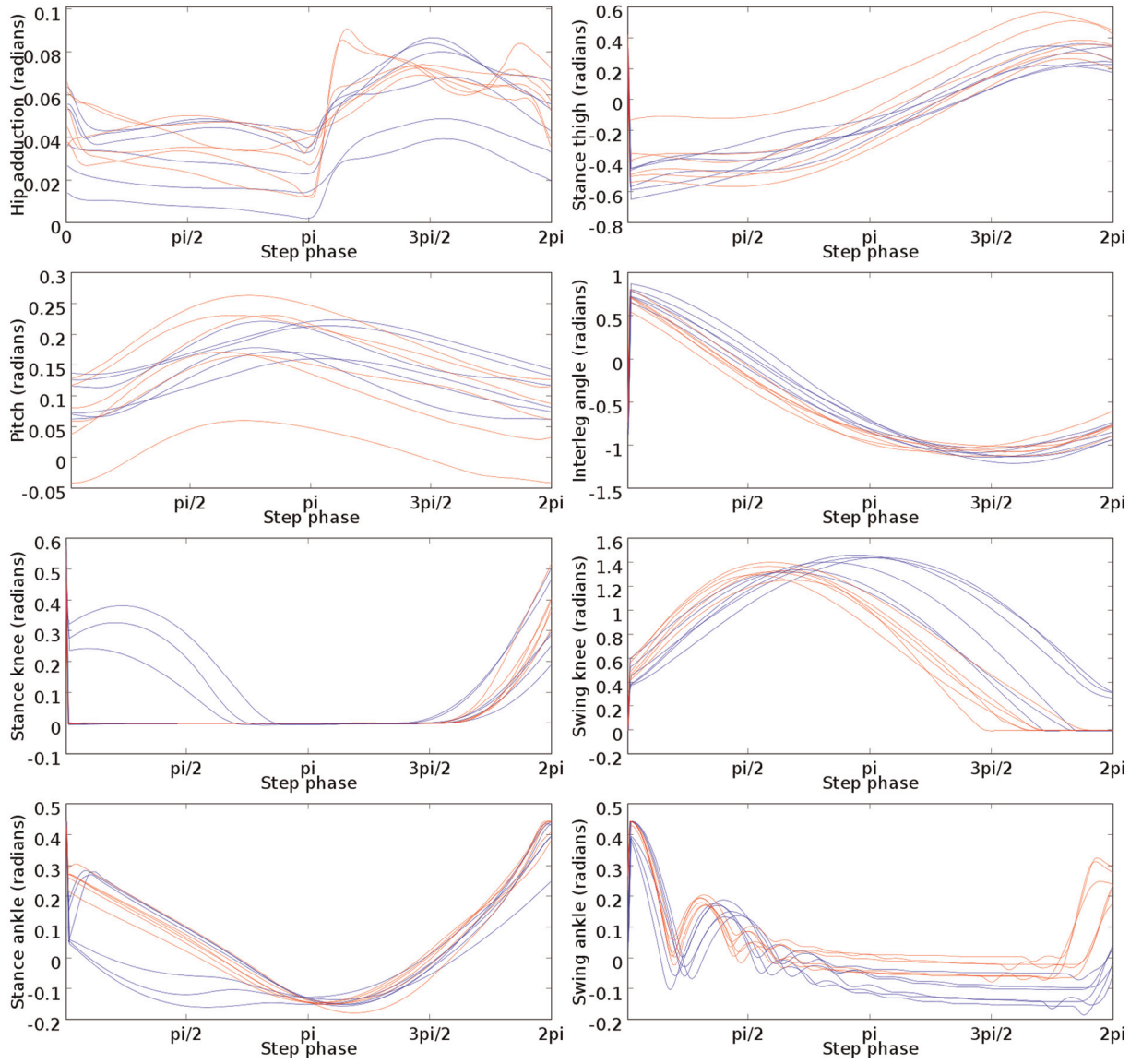

Figure 7. Kinematic trajectories of untrained walkers (blue) compared with those trained in the presence of parameter variations (red).

throughout the tested range of stiffnesses. These results indicate that this controller can be trained to perform well even on many different ground contact models, improving robustness in cases of poorly characterized ground contact.

Efficiency is compared in Figure 9b between a population of 500 walkers trained only on firm ground, and a population of 500 walkers trained on ground of varying properties. Both populations were tested on *firm ground* in order to directly compare the efficiency. Of the individuals that walked until their energy supply was exhausted, the efficiency of trained walkers appears to have increased, while the number that do not fall is much lower than the untrained population. The walkers that fall before using their allotment of energy do so uniformly throughout the range of 0 to 60 m. This indicates that the walkers challenged on unpredictable terrain have developed a strategy that is of greater benefit on unpredictable terrain than on uniformly firm ground. This strategy, although more efficient, is slightly less stable on firm terrain. The effect is exaggerated by the large range of terrain contact models encountered here. In reality, the walker can be trained on a much narrower range, but here we aimed to determine the limit of the controller's ability, to ensure that any realistic surface would fall inside the range of uncertainty for which the controllers are trained. Supplemental video 5 demonstrates the best trained walker.

4.5 All sources of noise and uncertainty

So far, we have shown that the linear reactive control scheme is robust to realistic actuation, imperfect

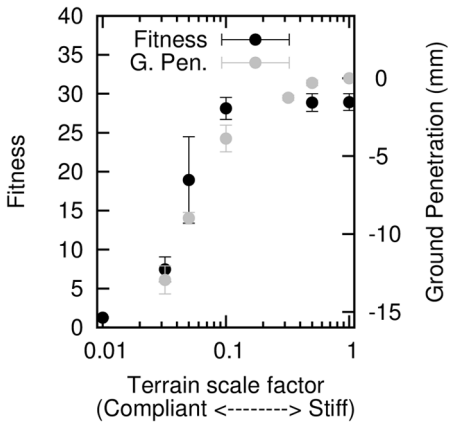

Figure 8. Average fitness decreases rapidly as the terrain becomes less firm. The ground penetration data provide a sense for the real-world compliance of each surface, ranging from firm ground (scale=1.0) to a surface resembling a soft mattress (scale=0.032).

sensing, and uncertain knowledge of the dynamic model and its interactions with the environment. However, controllers were only trained for robustness to a single source of noise or uncertainty at one time. As a final experiment, we now evolve a controller that can operate in the presence of all of the above effects simultaneously.

In these experiments, two populations of walkers were compared. The first population is the same as the population trained with SEAs and finite motor power from section 3.4.1. This population was not trained in the presence of any other effects, and is therefore the "untrained" population. The second population was trained by incrementally adding sources of uncertainty throughout the evolution (SEAs, finite motor power, sensor noise, uncertainty in model parameters, and uncertainty in terrain compliance). Specifically, after the walkers were trained to walk with SEAs and finite motor power, they were additionally challenged with noise in the joint sensor and IMU data, as in section 3.4.2. They were allowed to evolve for 200 generations in the presence of this noise. Then they were trained with model parameter variations. The range of randomly chosen model parameter values was ramped up in the same way as in section 3.4.3, for the same number of generations, but in this case variability peaked at a 10% deviation from the nominal values rather than 20%. Finally, while still randomizing the dynamic parameters on every generation, the ground stiffness was varied between 10% and 100% of the reference stiffness for 300 more generations.

The performance was then tested by conducting a tournament to choose the "best untrained" and the "best trained" individuals from their respective populations. Note that, in order to hold the number of generations constant, "untrained" individuals were allowed to evolve for an additional number of generations equal to the number of generations during which "trained" individuals were exposed to accumulated variations. The individual with the highest fitness in each of the last 50 generations of each population was tested with SEAs, finite power limits, sensor noise, and 10 randomly chosen sets of dynamics and contact parameters in the ranges described above. The winner from each population was determined by the highest average fitness, then each winner was evaluated using 200 more randomly chosen sets of parameters. These evaluations are plotted in Figure 10a.

The best untrained controller is shown on the left side of Figure 10a, and the best trained controller is shown on the right. The best untrained individual fails immediately in nearly all of the 200 trials, completing only 12. By contrast, the best trained individual successfully walks to exhaustion in 117 trials, nearly 10 times as many as the best untrained controller.

Figure 10b compares the efficiency of 500 individuals from the two populations of walkers, evaluated

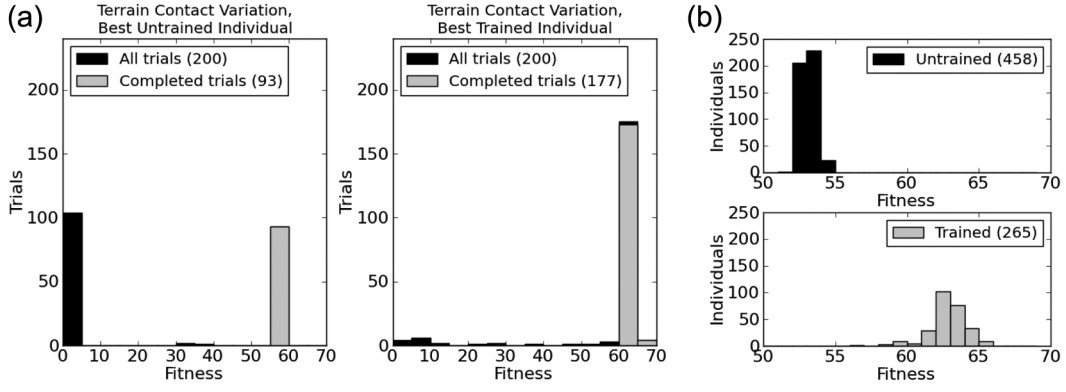

Figure 9. (a) The best untrained individual falls frequently in the first 5 m on unpredictable terrain, whereas the best trained individual is successful on a large range of terrain types. (b) Walkers trained on various terrain types become more efficient, but fewer survive to exhaust their entire energy supply.

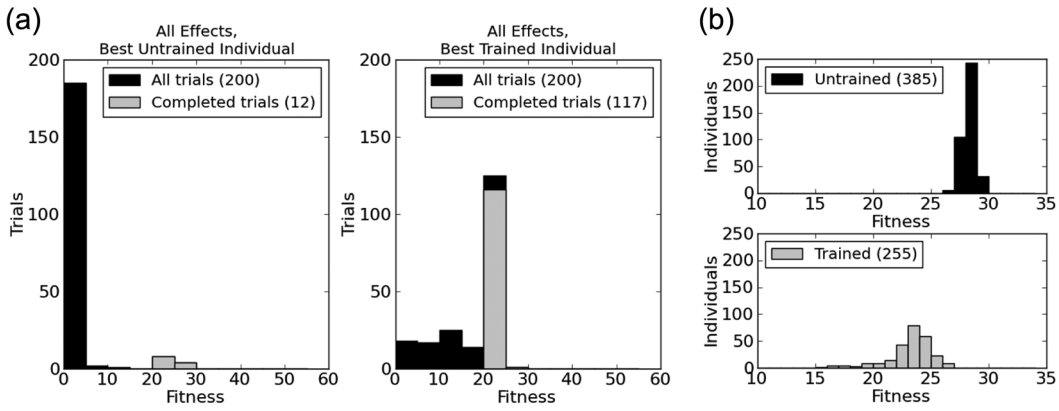

Figure 10. (a) The performance of tournament winners tested in the presence of all effects. (b) Walkers trained in the presence of all sources of noise and uncertainty were slightly less efficient.

using only the nominal parameter values, firm terrain, and noiseless sensors. These groups consisted of the 10 best individuals from the last 50 generations of each evolution. Walkers trained in the presence of all effects were slightly less efficient, and more fall before exhausting their energy supply than in the untrained population. However, as shown in Figure 10a, there is great improvement in robustness under uncertain conditions.

The conditions encountered in this test represent an extraordinarily difficult control task for any biped, given the very large variations in dynamic values and ground interaction. Using our evolutionary method, the best trained walker remained stable in several runs that are well outside the expected range of uncertainty for realistic conditions. This is a remarkable property for a purely reactive linear controller that evolved unsupervised, operating on a fully autonomous, unsupported three-dimensional system. The best trained walker is shown in supplemental video 6. This walker bears the most resemblance to the walker in video 2 (realistic actuation).

5 Discussion

The present work has demonstrated in simulation that a "linear-reactive" controller whose weights are optimized with an evolutionary algorithm can be used to stabilize 3D bipedal walking in a manner robust to multiple sources of uncertainty and noise.

5.1 Evolutionary robotics as an approach to the control of bipedal locomotion

Many studies of bipedal locomotion have the goal of developing control techniques sufficiently general that they can be applied directly to a variety of biped robots, with some tweaking of control parameters (Kuo, 1999; Seyfarth, Geyer, & Heff, 2003; Taga, Yamaguchi, & Shimizu, 1991). The basic idea that motivates these studies is that their generality will facilitate their implementation in hardware.

The present work begins to explore an alternative approach. Given the rate at which computing speed and power are increasing, a possible option for robot control in the future may be simply to evolve controller parameters specific to a particular robot morphology. In other words, a viable, real-world approach to robot control might be to run thousands of simulations to evolve controllers robust to within the "radical envelope of noise," (Jakobi et al., 1995) and then fine-tune these successful controllers in hardware.

In many ways, this approach replicates the process of an animal learning to control its movements. Each individual animal must learn the dynamics of its own limbs, starting at birth. For bipeds such as humans, this learning takes on the order of a year; the controller does not work without extensive learning of synaptic weights.

Although we think it is possible that the evolved weights might be analyzed in a manner that would reveal general control principles, this is not an aim of the present work. Instead, the results presented here demonstrate the possibility of evolving a controller for 3D biped locomotion likely to be robust within the envelope of noise. This in turn suggests that an empirically based solution to the control of 3D biped locomotion might be found using techniques in evolutionary robotics. Ultimately, proof of the viability of this approach will require successful hardware transfer.

5.2 The potential of a linear reactive control scheme to transfer to hardware

Of the many possible choices for an evolutionary approach, we choose a linear reactive control scheme. Our motivation for this choice was that previous work with 5-link and 7-link models (Solomon et al., 2010; Solomon et al., submitted) had suggested the viability

of this approach. Until now, however, linear reactive control had only been shown to be capable of learning walking behaviors under "perfect" conditions, albeit with rough terrain (Solomon et al., submitted; Solomon et al., 2010). In the present work, we introduced significant variability to all components of the simulated model. Rather than operating in a rigid-body simulation environment, we trained the controller with compliant actuators and realistic power limits, similar to those that would be found on a hardware biped. In addition, we were able to train the controller to be robust to realistic values of sensor noise, to uncertainties in the dynamic model, and to variations in the ground contact model. In summary, we have successfully trained the simulations to be robust within certain "envelopes of noise" surrounding the sensing, actuation, ground contact, and dynamics parameters of the simulation (Jakobi, 1997). Robustness to these uncertainties is one of several requisites for successful hardware transfer.

Our evolutionary algorithms cannot technically be said to have performed unbiased searches of the entire space of possible solutions, because we had to bootstrap the initial evolution by hand-tuning the bias weights in the initial set of parents. However, the bootstrapping was minimal. We only tuned one weight per actuator network in order to promote the most basic of walking behaviors, such as a tendency to move the swing leg forward, rather than backward. A search beginning from uniformly random initial conditions would not be feasible, because a strong majority of initial guesses would have been poor starting points for the optimization.

One risk inherent to any dynamic simulation is that the simulator itself may have properties that somehow make the dynamic problem easier to solve. This is true for any choice of physics engine (e.g., Bullet, Magic, ODE). It is possible, therefore, that our simulations of biped walking may inadvertently exploit structure within the simulation environment to enable stable walking. Barring this possibility, however, the present results indicate that this type of linear reactive controller may be appropriate for hardware transfer and viable for real-world applications.

5.3 Efficiency/robustness tradeoffs in linear reactive control

Early hominids may have evolved to exploit an efficiency advantage compared with quadrupeds (Rodman & McHenry, 1980). However, this advantage comes at the cost of stability, making robotic bipedal locomotion a challenging control task. The responses to this challenge have typically been approaches that sacrifice the inherent efficiency of the biped in favor of stability. Such results often lock the motion in known stable, periodic orbits that tend to look unnaturally stiff while walking. The inefficiency of the positional control of these orbits in turn requires large batteries, which increase weight and further decrease efficiency.

The approach used here may help to increase efficiency in two ways. First, linear reactive control is not constrained to behaviors that enforce local or periodic stability. This permits transiently unstable behaviors to emerge and for the system to fully exploit the efficiency inherent to its passive dynamics. Second, training for maximum efficiency is built into the evolutionary selection process. Thus, we are able to find weight sets that enable robustness to parameter variations with no apparent penalty to efficiency at all.

Intriguingly, our results with this controller demonstrate that the traditional efficiency/robustness tradeoff effect does not always manifest itself in a straightforward way. A realistic (series-elastic) actuation model dramatically reduces efficiency (Figure 2), but training for robustness to uncertainty in dynamic parameters does not. A population of walkers trained to accommodate uncertainty in dynamic parameters has the same average efficiency as untrained walkers, but higher variance (Figure 6b). Given that efficiency is an implicit factor in the fitness function, the largest cost of this robustness is the additional evolution time, not efficiency.

5.4 A robust walking controller can be algorithmically simple

The simplicity of linear reactive control is of particular interest, as it illuminates the requirements for robust walking. Even though the controller shown here is a function of 21 state variables (42 weights total), the function itself is trivial: linear reactive control is algorithmically nothing more than a weighted sum. Its implementation as a static function of state variables would indicate that such a minimal controller is sufficient not only for walking, but for very robust bipedal walking.

Within the bounds of this work, we can say that even for particularly challenging walking tasks (e.g., those in which actuator bandwidths are high, sensors are noisy, the model is not well known, and the ground reaction forces vary), linear reactive control would appear to be sufficient for walking. Time-based periodic orbits need not be specified explicitly in order to exhibit stable walking, even when the parameters of the model are not known accurately. In this way, our evolutionary self-tuning procedure is most closely related to direct adaptive control, in which the control gains are tuned online as the plant moves (Åström & Wittenmark, 2008). However, the offline optimization method used here removes training from the feedback loop, resulting in a much simpler control structure.

A more intricate control network was used by Vaughan et al. (2004), and consisted of a multilayer neural network with nonlinear activation functions and

a central pattern generator. These additional components might allow more detailed control maneuvers than our simple linear controller, but much of the interest of this technique is its remarkable robustness in spite of its minimalism. Algorithmically, a weighted sum is simpler than the neural-network approach used in Vaughan's work, yet we have found that our comparatively simple controller demonstrates that the internal rhythmicity of a CPG is not necessary for walking, nor is nonlinearity (introduced by sigmoidal activation functions and multiple layers) required for complex adaptive behaviors. Vaughan showed that changing morphologies could be overcome by his controllers, but all of the parameters that were changed were internal to the robot. In the present work we show not only that the controller is robust to internal variations, but also to more difficult actuation technologies, imperfect sensor information, and widely varying environmental parameters (which are distinctly different from the Gaussian-random external force perturbations applied in Vaughan's work). Finally, we also show that all sources of noise can be accumulated throughout the evolution in order to produce a controller that is robust to many types of noise. This accumulation of robust behaviors is critical, as an attempt to cross the reality gap will necessarily involve robustness to many sources of uncertainty simultaneously.

5.5 Large changes in kinematics do not necessarily imply large changes in weights

The kinematic behavior of both knees changes significantly when SEAs are added (Figure 3), but only one weight in the stance knee actuator network changes beyond one standard deviation from the "untrained" value. Because only the bias and PD gains are available in those networks (Table 1), other joints must be acquiring behaviors that accommodate the new dynamics. There are also noticeable differences in the kinematics of the parameter variation runs. These show subtle but consistent differences in the knee trajectories between "trained" and "untrained" (Figure 7), yet have no significant differences in any of the knee weights. Again, this suggests that the altered knee kinematics could be due to changes in the kinematics of the hip joints, which have networks that have a higher number of sensory inputs. In the case in which sensor noise was added, both the weights and the trajectories of individuals are very similar to the "untrained" runs, differing by more than one standard deviation for only 4/42 weights. Finally, the standard deviation of each weight tended to be larger in trained individuals than in their untrained counterparts.

5.6 Generalizability of linear reactive control

Linear reactive control requires only sensor inputs, with no restrictions on quantity or context. It is completely independent of the type of robot, and can therefore be applied to myriad form factors. It is also easily scalable. New sensors and actuators can be added (or existing sensors and actuators removed) with minimal changes to the architecture of the controller, and no additional design work. The mapping from sensors to actuators simply needs to reflect this change in connections, and then weights must be re-evolved. This makes the controller cheap and easy to deploy.

Beyond this, a linear reactive controller can be trained for robustness under any of the conditions expected for its purpose. The randomization procedure trains the controller not for specific values, but for an entire, unpredictable range of values. This range can be adjusted to suit the amount of uncertainty in the internal hardware and external environment in which the robot will operate. We therefore see this method as a very general control architecture, which could find use in many applications.

Acknowledgments

We thank Martijn Wisse for providing a copy of the *Flame* robot, associated CAD drawings, and other assistance. Some simulations were run using Jet Propulsion Laboratory (JPL) supercomputing resources.

Funding

This work was supported by the Director's Research and Development Fund program of the JPL (California Institute of Technology) to MJZH and JHS, the Graduate Student Research Program of the Jet Propulsion Laboratory (award number 10401 to MAL), and NSF CAREER award IOS-08090000 to MJZH.

References

Åström, K. J., & Wittenmark, B. (2008). *Adaptive control*. New York: Dover Publications.

Bongard, J. (2011). Morphological change in machines accelerates the evolution of robust behavior. *Proceedings of the National Academy of Sciences of the United States of America, 108*(4), 1234–1239.

Byl, K., & Tedrake, R. (2009). Metastable walking machines. *International Journal of Robotics Research, 28*(8), 1040–1064.

Collins, S., Ruina, A., Tedrake, R., & Wisse, M. (2005). Efficient bipedal robots based on passive-dynamic walkers. *Science, 307*(5712), 1082–1085.

Collins, S. H., Wisse, M., & Ruina, A. (2001). A three-dimensional passive-dynamic walking robot with two legs and knees. *International Journal of Robotics Research, 20*(7), 607–615.

Eiben, A. E., & Smith, J. E. (2003). *Introduction to evolutionary computing*. Berlin: Springer.

Goswami, A., Thuilot, B., & Espiau, B. (1998). A study of the passive gait of a compass-like biped robot: Symmetry and chaos. *International Journal of Robotics Research, 17*(12), 1282–1301.

Grizzle, J. W., Hurst, J., Morris, B., Park, H. W., & Sreenath, K. (2009). MABEL, a new robotic bipedal walker and

runner. *2009 American Control Conference, Vols 1–9* (pp. 2030–2036). New York: IEEE.

Hobbelen, D., de Boer, T., & Wisse, M. (2008). *System overview of bipedal robots Flame and TUlip: Tailor-made for limit cycle walking*. New York: IEEE.

Iida, F., Rummel, J., & Seyfarth, A. (2008). Bipedal walking and running with spring-like biarticular muscles. *Journal of Biomechanics, 41*(3), 656–667.

Iida, F., & Tedrake, R. (2010). Minimalistic control of biped walking in rough terrain. *Autonomous Robots, 28*(3), 355–368.

Jakobi, N. (1997). Evolutionary robotics and the radical envelope-of-noise hypothesis. *Adaptive Behavior, 6*(2), 325–368.

Jakobi, N., Husbands, P., & Harvey, I. (1995). Noise and the reality gap: The use of simulation in evolutionary robotics. In F. Moran, A. Moreno, J. J. Merelo & P. Chacon (Eds.), *Advances in artificial life* (Vol. 929, pp. 704–720). Berlin: Springer-Verlag.

Janabi-Sharifi, F., Hayward, V., & Chen, C. S. J. (2000). Discrete-time adaptive windowing for velocity estimation. *IEEE Transactions on Control Systems Technology, 8*(6), 1003–1009.

Jung-Hoon, K., Jung-Yup, K., & Jun-Ho, O. (2010, 6–8 Dec. 2010). *Weight-adaptive walking of the passenger-carrying biped robot, HUBO FX-1*. Paper presented at the 10th IEEE-RAS International Conference on Humanoid Robots (Humanoids).

Kajita, S., Kanehiro, F., Kaneko, K., Fujiwara, K., Harada, K., Yokoi, K., & Hirukawa, H. (2003). Biped walking pattern generation by using preview control of zero-moment point *2003 IEEE International Conference on Robotics and Automation, Vols 1–3, Proceedings* (pp. 1620–1626). New York: IEEE.

Kajita, S., Nagasaki, T., Kaneko, K., & Hirukawa, H. (2007). ZMP-based biped running control—The HRP-2LR humanoid biped robot. *IEEE Robotics & Automation Magazine, 14*(2), 63–72.

Kuo, A. D. (1999). Stabilization of lateral motion in passive dynamic walking. *International Journal of Robotics Research, 18*(9), 917–930.

Kuo, A. D. (2007). Choosing your steps carefully—Trade-offs between economy and versatility in dynamic walking bipedal robots. *IEEE Robotics & Automation Magazine, 14*(2), 18–29.

Kuo, C. H., & Chiou, K. W. (2009). Tendon based full size biped humanoid robot walking platform design. In J. H. Kim, S. S. Ge, P. Vadakkepat, N. Jesse, A. AlManum, S. Puthusserypady, … D. Ahlgren (Eds.), *Advances in Robotics* (Vol. 5744, pp. 135–150). Berlin: Springer-Verlag.

Manoonpong, P., Geng, T., Kulvicius, T., Porr, B., & Worgotter, F. (2007). Adaptive, fast walking in a biped robot under neuronal control and learning. *PLoS Computational Biology, 3*(7), 1305–1320.

McGeer, T. (1990). Passive dynamic walking. *International Journal of Robotics Research, 9*(2), 62–82.

Morimoto, J., & Atkeson, C. G. (2007). Learning biped locomotion—Application of Poincare-map-based reinforcement learning. *IEEE Robotics & Automation Magazine, 14*(2), 41–51.

Nolfi, S., & Floreano, D. (2000). *Evolutionary robotics*. Cambridge, Massachusetts: MIT Press.

Ono, K., Takahashi, R., & Shimada, T. (2001). Self-excited walking of a biped mechanism. *International Journal of Robotics Research*, *20*(12), 953–966.

Paul, C. (2005). Sensorimotor control of biped locomotion. *Adaptive Behavior, 13*(1), 67–80.

Plestan, F., Grizzle, J. W., Westervelt, E. R., & Abba, G. (2003). Stable walking of a 7-DOF biped robot. *IEEE Transactions on Robotics and Automation, 19*(4), 653–668.

Pratt, G. A. (2002). Low impedance walking robots. *Integrative and Comparative Biology, 42*(1), 174–181.

Pratt, J., Krupp, B., & Morse, C. (2002). Series elastic actuators for high fidelity force control. *Industrial Robot, 29*(3), 234–241.

Qi, F., Wang, T. S., & Li, J. F. (2011). The elastic contact influences on passive walking gaits. *Robotica, 29*, 787–796.

Reil, T., & Husbands, P. (2002). Evolution of central pattern generators for bipedal walking in a real-time physics environment. *IEEE Transactions on Evolutionary Computation, 6*(2), 159–168.

Rodman, P. S., & McHenry, H. M. (1980). Bioenergetics and the origin of hominid bipedalism. *American Journal of Physical Anthropology, 52*(1), 103–106.

Sakagami, Y., Watanabe, R., Aoyama, C., Matsunaga, S., Higaki, N., & Fujimura, K. (2002). *The intelligent ASIMO: System overview and integration*. New York: IEEE.

Seyfarth, A., Geyer, H., & Heff, H. (2003). Swing-leg retraction: a simple control model for stable running. *Journal of Experimental Biology, 206*(15), 2547–2555.

Smith, R. (2012). Open Dynamics Engine.

Solomon, J. H., Locascio, M. A., & Hartmann, M. J. Z. (submitted). Linear reactive control for efficient 2D and 3D walking over rugged terrain.

Solomon, J. H., Wisse, M., & Hartmann, M. J. Z. (2010). Fully-interconnected, linear control for limit cycle walking. *Adaptive Behavior, 18*(6), 492–506.

Srinivasan, S., Westervelt, E. R., & Hansen, A. H. (2009). A low-dimensional sagittal-plane forward-dynamic model for asymmetric gait and its application to study the gait of transtibial prosthesis users. *Journal of Biomechanical Engineering-Transactions of the ASME, 131*(3).

Taga, G., Yamaguchi, Y., & Shimizu, H. (1991). Self-organized control of bipedal locomotion by neural oscillators in unpredictable environment. *Biological Cybernetics, 65*, 147–159.

Uustal, H., & Baerga, E. (2004). *Physical medicine and rehabilitation board review*. New York: Demos Medical Publishing.

Vaughan, E. D., Di Paolo, E., & Harvey, I. R. (2004). *The evolution of control and adaptation in a 3D powered passive dynamic walker*. Cambridge: MIT Press.

Vukobratovic, M., & Borovac, B. (2004). Zero-moment point—Thirty five years of its life. *International Journal of Humanoid Robotics, 1*(1), 157–173.

Westervelt, E. R., Grizzle, J. W., & Koditschek, D. E. (2003). Hybrid zero dynamics of planar biped walkers. *IEEE Transactions on Automatic Control, 48*(1), 42–56.

Wisse, M. (2008). *Deliverable 1.2, Hardware Specifications of EU FP6 Project "ESBiRRo"*. TU Delft.

Zagal, J. C., & Ruiz-del-Solar, J. (2007). Combining simulation and reality in evolutionary robotics. *Journal of Intelligent & Robotic Systems, 50*(1), 19–39.

About the Authors

Mark A Locascio received a B.S. degree in engineering from Harvey Mudd College (Claremont, CA) in 2005. He completed an M.S. degree in mechanical engineering from Northwestern University (Evanston, IL) in 2009, and received his Ph.D. degree in mechanical engineering from Northwestern University in 2012. He was a NASA GSRP fellow from 2010 to 2012. His research interests include evolutionary robotics, machine learning, control, and numerical methods.

Joseph H Solomon is currently an adjunct assistant professor in the department of biomedical engineering at Northwestern University, Evanston, IL. He received a B.S. degree in aeronautical and astronautical engineering from the University of Illinois at Urbana-Champaign in 2001, an M.S. degree in mechanical engineering from the University of Illinois at Chicago in 2003, and a Ph.D. degree in mechanical engineering from Northwestern University in 2008. His research interests include evolutionary algorithms, neural networks, artificial general intelligence, biomimetic sensing, and bipedal locomotion.

Mitra JZ Hartmann is an associate professor at Northwestern University with a 50/50 joint appointment between the departments of biomedical engineering and mechanical engineering. She received a B.S. degree in applied and engineering physics from Cornell University, and a Ph.D. degree in integrative neuroscience from the California Institute of Technology. From 2000-2003, she was a postdoctoral scholar at the Jet Propulsion Laboratory in Pasadena, California, where she worked in the Bio-Inspired Technology and Systems group. Dr. Hartmann's research interests involve understanding the mechanics and neuroscience of active sensing behaviors.

Appendix

Series-elastic actuator (SEA) parameters

Runs that included SEAs did not apply computed torques directly to the joints shown in Figure 1. Instead, torque was applied to a motor mass (the "rotor"), which was connected to the joint through a second-order connection. The torque applied to the joint is therefore computed as

$$\tau = k(\theta_{motor} - \theta_{joint}) + c(\dot{\theta}_{motor} - \dot{\theta}_{joint})$$

This is shown visually in the schematic in Figure 11.

The SEAs used in section 3.4.1 use the parameters in Table 2.

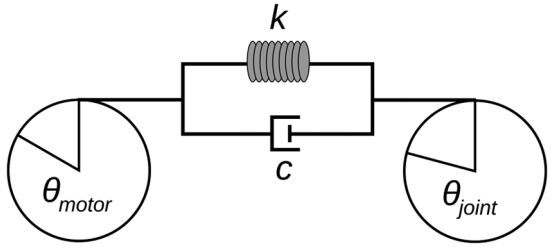

Figure 11. A schematic of the series-elastic actuator (SEA) model. On each timestep, the difference between the positions (and optionally, velocities) of the motor and joint are measured and used to compute the torque applied to the joint through the second-order connection.

Table 2. Series-elastic actuator (SEA) model parameters

	kg Rotor mass	N/m *k*	N/(m/s) *c*
Hip	0.27	413.68	1.70
Knee	0.22	10.18	0.50
Left ankle	0.09	14.80	0.10
Right ankle	0.16	22.83	0.06