Combining Optimal and Neuromuscular Controllers for Agile and Robust Humanoid Behavior

Hartmut Geyer and Chris Atkeson, Carnegie Mellon University

Overview. A central conundrum in robotics is that manually designed controllers are often much more robust than model-based controllers, but model-based controllers are easier to command and coordinate with perceptual systems such as vision. The overall goal of this proposal is to test the hypothesis that a hybrid controller, combining a manually designed neuromuscular model of human control and an optimization-based controller, can achieve the desirable features of both types of control. Exploring hybrid control could lead to agile and robust robot legged locomotion, give us better understanding of the science of locomotion, and potentially lead to insights into human locomotion. Performance goals include human-like walking, achieving human speeds, step length, and cadence; control of locomotion including starting, stopping, speed, and steering; walking on rough terrain including slopes, slippery surfaces, and granular material such as dirt and soil; and whole body locomotion, including the use of hand rails, grab bars, canes, and walkers. The researchers have several legged robots (Sarcos Primus and Atrias), and access to others (Atlas and Coman) for comparative evaluation of current approaches and proposed hybrid approaches. They will evaluate the ability to match human speeds, step lengths, and cadence, control of speed and direction, rough terrain locomotion, whole body locomotion, and the ability to recover from disturbances and errors such as pushes, slipping, and tripping. They will also compare robot and human kinematics and dynamics, including joint velocities and torques, head velocity and acceleration, and ground reaction forces. The project involves theoretical research on nonlinear control, development of new control algorithms, as well as experimental research verifying control theory on humanoid robots. A successful outcome could trigger a paradigm shift in control approaches to robust behavior.

Intellectual Merit. The direct contributions of this research will be: (1) A mathematical framework blending model-based and parametric policy-based control for whole body locomotion in robotics. The framework will be independent of the particular policy, although we focus on a neuromuscular policy as an example in this project. The framework may transfer to other scientific domains where blending optimization with domain knowledge is beneficial. (2) The experimental evaluation in simulation and humanoid robot testbeds of three specific approaches within this framework that focus on the policy as a cost, constraint, or replacement. (3) An understanding of how well these types of blended control generalize across robot platforms of different morphology and sensory and actuation modalities. (4) New hypotheses about human sensory-motor control of locomotion behaviors through the advancement of the neuromuscular policy.

The direct contributions of this research are a deeper understanding of the benefits of combining model-based and parametric policy-based control for locomotion performance in robotics and a comparison of walking controllers for humanoids that robustly navigate complex and uncertain environments. More broadly, the proposed research will contribute to better understanding of robust nonlinear control, and may lead to better understanding of human sensorimotor control.

Broader Impacts. The research will lead to more useful robots. A successful outcome will enable practical controllers for robust locomotion of legged robots in uncertain environments with applications in disaster response and elderly care. The project may generate new insights into human sensorimotor control of behavior, which can lead to better prostheses and exoskeletons and more effective approaches to fall risk management. The proposed project integrates research into education and mentoring by directly training graduate and undergraduate students and by incorporating research outcomes into course work. The project strengthens the infrastructure for research and education by improving the robot testbeds available for research and education. The project supports the broad dissemination of research results by publishing and freely providing simulation and implementation codes. We also expect unplanned broader impacts such as appearing in the news, on TV shows, and in movies. Unplanned broader impacts from previous work include our technologies being demonstrated on entries in the DARPA Robotics Challenge, a graduate student being
a participant on a Discovery Channel TV series, The Big Brain Theory: Pure Genius, a graduate student helping run a group of all female high school students in the robot FIRST competition, and a robot character in a Disney movie being inspired by our work on soft robots (Baymax in Big Hero 6).
PROJECT DESCRIPTION

1 Objectives

Motivation. A central conundrum in robotics is that manually designed parametric policy-based controllers (for example [75, 38, 128, 74, 29, 110, 66]) are often more robust than model-based optimization controllers, but model-based controllers (for example [96, 60, 80, 67, 63, 25, 24]) generalize well to automated behavior generation, which requires direct control of foot or hand placement for visually guided locomotion or the use of handholds, canes, walkers, and other walking aids [17]. We believe that finding a way to combine the advantages of both approaches will lead to more human-like and robust locomotion in humanoid robots, give us a better understanding of the science of locomotion, and potentially lead to new insights into human motor control.

Research Question and Proposed Solution. Our motivation translates into a specific research question: Can these two approaches be combined within one formal framework such that the resulting hybrid approach shares the human-likeness and robustness of policy-based control with the versatility of model-based control? We propose as a solution to this question the use of policy-based control as either a cost, a constraint, or a replacement within the model-based optimization control framework. We hypothesize that each of these variants offers different advantages and shortcomings with respect to the human-likeness, robustness, and computational efficiency of humanoid control. As an example of the model-based controller, we will focus on a hierarchical optimization-based controller developed by Atkeson’s group [96, 25, 24] (Sec. 3). This controller provides direct control of foot and hand placement and steering. It can be combined with a vision system, and can automatically generate a variety of behaviors. For instance, we have demonstrated visually guided rough terrain locomotion, and we have added the use of hands for climbing a ladder with the ATLAS robot [25, 24]. As an example of the policy-based control, we will focus on a neuromuscular model of human control created by Geyer’s group [29, 90, 87, 88] (Sec. 4). The model presents an attractive candidate policy that is inherently human-like, comparably robust to disturbances, and that is less demanding in terms of state estimation (for most joints only position is required and not velocity, for example).

Specific Objectives. Our specific objectives are to formally develop the three variants of a hybrid approach and to evaluate their performance in simulation and hardware experiments with different humanoid robots (Fig. 1) (Sec. 5). The metrics according to which we will evaluate the performance include locomotion speed, cadence, duty factor, as well as kinematics, kinetics and ground reaction force patterns (human-likeness), external push resistance and survival rate on rocky and slippery terrain (robustness to external perturbations), model parameter variations and artificial sensory noise (robustness to internal uncertainty), and cycle time of controller execution (computational costs). For the hardware experiments, our groups are equipped with humanoid robots of different complexity (ATRIAS and SARCOS Primus), and we have access to two more humanoids with different actuation and sensing capabilities (COMAN and ATLAS). This access to a range of hardware platforms will allow us to evaluate how well the observed outcomes generalize across humanoids.

Intellectual Merit. The expected contributions of the proposed research are: (1) A mathematical framework blending model-based and parametric policy-based control for whole body locomotion in robotics. The framework will be independent of the particular policy, although we focus on a neuromuscular policy as an example in this project. (2) The experimental evaluation in simulation and humanoid robot testbeds of three specific approaches within this framework that focus on the policy as a cost, constraint, or replacement. (3) An understanding of how well these types of blended control generalize across robot platforms of different morphology and sensory and actuation modalities. (4) New hypotheses about human sensory-motor control of locomotion behaviors through advancing the neuromuscular policy.
Significance. These contributions can have significant intellectual and societal impact. Although humanoid robots are considered key technological platforms with applications in disaster response and mitigation and elderly care, robust humanoid locomotion remains a research topic in a nascent stage. Successfully developing a mathematical framework that blends model-based and policy-based control, and combines the advantages of both individual control approaches, could trigger a paradigm shift in generating robust locomotion behaviors in robotics with an impact on future robotics technology including humanoids, prosthetics, and exoskeletons. In addition, the mathematical framework may generalize to other scientific domains and problems which benefit from incorporating domain knowledge into optimization.

2 Related Work

Model-based Control in Humanoid Robotics. Force-controlled humanoid robots are often controlled at the behavior level using models of the robot and environment dynamics. These computed torque approaches are generalized to floating base inverse dynamics [96, 60, 80, 67, 63, 25, 24], which enable the robot to track the behavior of an underlying, simplified dynamics model. In legged locomotion, common underlying models are, for instance, the linear inverted pendulum model used in humanoid robots like ASIMO, HRP-4 and HUBO [46, 37, 47, 103, 13], or the zero dynamics manifold for bipedal robots like RABBIT, MABEL, and AMBER, controlled within the hybrid zero dynamics framework [12, 111, 2, 92]. The model-based control paradigm is mathematically elegant, and the corresponding locomotion problem has been solved in theory with optimal control techniques that automatically generate behavior [44, 8, 64, 104, 106, 25]. In practice, however, the paradigm requires to identify the states and dynamic parameters of a robot with high accuracy, and humanoids struggle to demonstrate robust dynamic locomotion.

Parametric Policy-Based Controllers. Parametric policy-based controllers for legged locomotion have been developed in the robotics, computer graphics, and neuromechanics communities [75, 38, 128, 74, 29, 66, 110]. These policy-based controllers capture explicit domain knowledge, acquired either by intuition or by a more formal and time-consuming analysis of individual locomotion behaviors. The Raibert robots in particular demonstrate robust locomotion while navigating rough terrain [74, 66]. In addition, simulation experiments demonstrate that policy controlled models of humans and humanoids can withstand comparably large disturbances including unexpected ground changes and pushes to the body [110, 91, 88, 27].

One motivation for investigating parametric policies is the observation that, in animals and humans, decentralized neural control circuits (central pattern generators and muscle reflex modules) at the level of the spinal cord seem to contribute substantially to the generation and robustness of legged locomotion.
Figure 2. Controller architecture based on hierarchical optimization

[85, 11, 34, 20, 70, 40, 28, 15]. It is believed that imitating these types of parametric policy-based control will help machines to match the agility and robustness of animals [84, 129, 65].

Hybrid Approaches. Very few attempts have been made to combine the benefits of both types of control approaches in humanoid robotics. The control overview reported for the Petman humanoid [66] seems closest in spirit, although the paper does not reveal implementation details. Specifically, during stance the control regulates global goals like the height and orientation of the body similar to other implementations [46, 73, 25]. Desired behaviors also includes target trajectories for the pelvis, back and arm motions, which can be interpreted as costs that improve the human-likeness of behavior. A force distribution algorithm is then used to best achieve the necessary leg forces that track the different behaviors simultaneously.

We are not aware of a formal framework or analysis of a hybrid approach to humanoid control which seamlessly blends model-based optimization control and policy based control.

3 Control Strategy using Model-Based Hierarchical Optimization

Figure 2 shows the architecture of our robot control based on hierarchical optimization. It was successfully tested on our SARCOS and ATLAS platforms (Fig. 1b,c). Our ATLAS robot demonstrated rough terrain walking, ladder climbing, and full body manipulation in the DARPA Robotics Challenge Trials of 2013. This work also received a “Best Oral Paper Award” at Humanoids 2013 [25]. Figure 3 shows a sequence of snapshots of ATLAS walking on rough terrain and plots of measured trajectories of the feet and the center of mass (CoM). Figure 4 shows similar data for ladder climbing.

3.1 Overview

The Behavior Library (Fig. 2) contains learned behaviors and behaviors created from demonstration or motion capture of humans or other robots, and also acts as a cache for planning. The output of the library is used to prime, speed up, and guide the real time optimization process, and can be motion or force trajectories

Figure 3. ATLAS robot practicing for the terrain task of the DRC. Photo snapshots taken every 5 seconds. Measured trajectories in the transversal (a) and sagittal planes (b) of the left (red) and right foot (green) and the center of mass (CoM, blue) are shown. The robot walks in a straight line along \( y = 0 \) in reality. Our state estimator at that time drifted significantly. This has been fixed by adding vision-based SLAM.
Figure 4. ATLAS climbing the top half of the same ladder as used in the DRC. Photo snapshots taken every 13 seconds. The top row shows repositioning of the hook hands, and the bottom row shows stepping up one tread. The two plots show ATLAS climbing the first five treads during the actual run at the DRC Trials. Axis definitions and color codes for foot and CoM trajectories are the same as in Fig. 3. Left and right hand positions are plotted with cyan and magenta dashed lines.

in joint or task coordinates \((x_d(t))\), as well as possibly including value, heuristic cost, or utility functions \((V_c(t))\) that can be used to guide search and optimization. Libraries can be built in advance, or when task specification are known. The rest of the controller can also operate without information from the library if suitable information is not available.

The Multistep Low-Order Dynamics Planner uses receding horizon control to continually optimize trajectories of motion and force of a simplified model of the current task. In walking, this is usually center of mass \((COM_d(t))\), angular momentum \((L_d(t))\), contact force \((\text{contacts}_d(t))\) trajectories, and value, heuristic cost, or utility functions \((V_c(t))\) specifying tradeoffs between these quantities. We optimize with a time horizon of several seconds using Di
d
erential Dynamic Programming [43].

For walking, our planner is similar in spirit to Kajita’s Preview Control for a Linear Inverted Pendulum Model (LIPM) [45]. We are also using a CoM model, reason about Zero Moment Point (ZMP), and use future information to guide the current trajectory. However, our planner can be generalized to nonlinear models and adjust foot steps while optimizing the CoM trajectory. We explicitly add the vertical height in our CoM model to handle height variations on rough terrain. Like capture point methods [72], we take the next few steps into consideration but do not plan to come to rest at the end. Ogura et al. [68] investigated generating human-like walking with heel-strike and toe-off by parameterizing the swing foot trajectory. In contrast, we use very simple rules to guide the low level controller to achieve the same behaviors.

The Full Body QP Optimization uses quadratic programming (QP) with a very short time horizon (1ms on SARCOS and 3ms on ATLAS) to optimize the desired full state of the robot \((\theta_d, \dot{\theta}_d, \tau_d)\), contact forces \((\text{contacts}_d)\), and joint torques \((\tau_d)\). Currently, we pre-specify gain matrices \((K_{\theta}, K_{\dot{\theta}}, K_{\tau}, K_f)\) for the robot controller, but plan to explore how to optimize gains as well as trajectories. This module enforces constraints such as robot kinematics and dynamics and joint, velocity, torque, actuator, friction cone, and center of pressure limits, and resolves redundancies. The module performs inverse kinematics and dynamics to provide us with compliant motions and dynamic behaviors, while compensating for the effects of modeling errors.

For the QP, we use a formulation developed in our group [94, 114, 112]. Unlike [61, 76] who use orthogonal decomposition to project the allowable motions into the null space of the constraint Jacobian, and minimize costs in the contact constraints and the commands, we directly optimize a quadratic cost in terms of state accelerations, torques and contact forces on the full robot model. This design choice allows us to trade off physical quantities of interest. We are also able to directly reason about inequality constraints such as center of pressure within the support polygon, friction, and torque limits. Although it becomes a bigger QP problem, we are still able to solve it in real time. Hutter et al. [41] resolved redundancy in inverse dynamics using a kinematic task prioritization approach that ensures lower priority tasks always exist in the null space of higher priority ones. In contrast to their strictly hierarchical approach, we minimize a sum of
weighted terms. We can directly specify the relative importance of the terms by adjusting the weights.

The **Low-Level Robot Controller** is a blend of our work and controllers provided by the robot manufacturers. For SARCOS and ATLAS, hydraulic valve commands for each joint \(i\) are provided by the manufacturers,

\[
v_i = k_{\theta d}(\theta_i - \theta_{d,i}) + k_{\dot{\theta} d}(\dot{\theta}_i - \dot{\theta}_{d,i}) + k_{\tau d}(\tau_i - \tau_{d,i}) + v_{ff,i},
\]

where \(v_i\) is the valve command, \(\theta_i\) is the joint angle, \(\dot{\theta}_i\) is the joint angular velocity, \(\tau_i\) is the joint torque, and \(v_{ff,i}\) is a feedforward valve command. Subscripts \(d\) indicate desired values, and quantities denoted with \(k\) are scalar gains. This joint level servo runs at 5kHz on SARCOS and at 1kHz on ATLAS.

We augment this independent joint control by coupling joints and providing force control based on force/torque sensors in the wrists, ankles, and skin of the robot,

\[
v = K_\theta(\theta - \theta_d) + K_{\dot{\theta}}(\dot{\theta} - \dot{\theta}_d) + K_{\tau}(\tau - \tau_d) + K_f(f - f_d) + v_{ff}
\]

where \(v\), \(\theta\), \(\dot{\theta}\), \(\tau\), and \(v_{ff}\) are the corresponding vector quantities, and the \(K\) are matrix gains coupling all joints. \(f\) are sensed contact forces. The coupled controller runs at 1kHz on SARCOS and 333Hz on ATLAS.

Figure 2 does not include our state estimator, which provides appropriate state information to all other modules. The details of this work on state estimation are described in [124].

### 3.2 Results from Prior NSF Support Related to this Work

(a) **NSF award**: IIS-0964581 (PI: Hodgins, Atkeson co-PI); **amount**: $699,879; **period**: 7/1/10 - 6/30/14.
(b) **Title**: RI: Medium: Collaborative Research: Trajectory Libraries for Locomotion on Rough Terrain.
(c) **Summary of Results**: This grant has supported work on a variety of approaches to controlling humanoid robots based on trajectory libraries. Section 3.1 focused on the hierarchical optimization approach, which forms one starting point for the proposed work and received a “Best Oral Paper Award” at Humanoids 2013.

**Intellectual Merit.** (1) Identification of a hierarchical approach to online optimal control of behavior, as described above. (2) Algorithms for training robust policies/controllers using multiple models, forming the basis for a new approach to robust learning and for algorithms greatly speeding up reinforcement learning. (3) Simple learning approaches to concisely represent and accurately predict optimal trajectories. (4) Globally optimal control of instantaneously coupled systems (ICS), which is designed by coordinating multiple lower-dimensional optimal controllers. (5) We discovered that many features of optimized walking including costs can be fit using simple global function approximation, such as quadratic function approximation. (6) A paradigm for designing controllers for complex systems, “informed priority control”, which coordinates multiple sub-policies and enables us to prototype complex control system designs faster. (7) State estimation for mobile and humanoid robots with “floating body” dynamics.

**Broader Impacts.** We are working toward more useful robots and knowledge that may help with a significant social problem, why people fall and injure themselves. We have coordinated our outreach activities with the larger outreach efforts of CMU’s NSF Engineering Research Center on Quality of Life Technology to achieve greater reach and effectiveness. Our technologies are being shared by being published, and papers are available electronically. The technologies are also being demonstrated on entries in the DARPA Robotics Challenge. We have created and teach relevant courses, including an undergraduate course on humanoid robotics and a graduate course on optimization of behavior. We put the teaching materials on the web to widely disseminate the results. Our work is being used in Disney Research and through this technology transfer path will eventually be used in entertainment and education applications, and will be available to and inspire the public. A graduate student was a participant on a Discovery Channel TV series, “The Big Brain Theory: Pure Genius”. One purpose of the TV series is getting people excited about engineering. A graduate student helps run a group of all female high school students in the robot FIRST competition. A robot character in a Disney movie is inspired by our work on soft robots (Baymax in Big Hero 6).
Development of Human Resources. The project involved five PhD graduate students (three students have graduated) and four postdocs. We had weekly meetings, weekly lab meetings, and we individually mentored all participants. All participants actively did research, made presentations to our group, gave conference presentations, gave lectures in courses. The students served as teaching assistants.

(d) Publications resulting from this NSF award: [126, 62, 4, 52, 125, 78, 99, 93, 97, 122, 130, 86, 77, 5, 131, 7, 50, 116, 123, 39, 54, 3, 113, 95, 115, 122, 117, 48, 26].

(e) Other research products: [Chris: needs “a brief description of available data, samples, physical collections and other related research products not described elsewhere;” Can also say “none”]

(f) Renewed support. This proposal is not for renewed support.

4 Control Strategy using Neuromuscular Policy

Figure 5 shows the architecture of our human locomotion model controlled by a neuromuscular control policy. We have shown in previous work that this model successfully captures human-like kinematics, kinetics, and ground reaction forces [29, 90, 88], is robust to comparably large external and internal disturbances [29, 89, 90, 87, 91] (Fig. 5b,c), and shows a range of human locomotion behaviors [88] (Fig. 6). We also have applied the neuromuscular policies of this model to the control of prosthetic limbs [21, 105], leading to the first powered ankle prosthesis that is commercially available to transtibial amputees [36, 1].

4.1 Overview

The Musculoskeletal Dynamics are represented in the human model with 8 segments connected by revolute joints (Fig. 5a, i) that are spanned by 18 Hill-type muscles (Fig. 5a, ii). The muscles generate net torques about the joints they span. Each muscle’s force is computed based on a Hill-type model [108] with series and parallel elasticities connected to a contractile element (Fig. 5a, iii). The contractile element’s force $F_{ce} = AF_{max} f_l f_v$ depends on mechanical muscle properties described by the maximum isometric force $F_{max}$ and the nonlinear force-length and force-velocity relationships ($f_l$ and $f_v$), as well as on the muscle activation $A$. Muscle activations are low-pass filtered stimulations $S$ generated by a model of human neural control.

The Neural Control (Figs. 5, iv) of the model almost exclusively represents spinal reflex circuits that are physiologically plausible and embed major functions of legged dynamics and control as local feedback loops [29]. For example, the compliant leg behavior important to legged dynamics in stance [10, 59, 32] is realized with simple positive force reflexes of the anti-gravity muscles, and their stimulation is modeled as...
Figure 6. Examples of steady and transient locomotion behaviors that the neuromuscular policy can generate.

\[ S(t) = S_0 + GF_{ce}(t - \Delta t), \]

where \( S_0 \) is a reference membrane potential at the \( \alpha \)-motoneuron, \( G \) is a synaptic gain, and \( \Delta t \) describes a neural time delay (up to 20ms) [30]. The neural control combines multiple of these functional reflex circuits into control modules. Together they form the neuromuscular control policy.

Our approach to modeling neural control stands in contrast to more common approaches which emphasize central pattern generation [102, 100, 101, 35, 51, 84, 65]. An advantage that our approach provides is that it links physiology to physics-based function, making the contribution of individual control elements clear and transferrable to rehabilitation and humanoid robotics.

The Neuromuscular Control Policy (NMP) represent only one example of several other control policies, which we could consider for a hybrid approach that blends optimization and policy based control (compare Sec. 2). In fact, the main ideas behind the approaches we propose to investigate do not hinge on the particular policy. However, the NMP provides several advantages that make it an attractive candidate policy. It is inherently human-like, does not depend on a high-fidelity model of the full system dynamics, and is comparably robust to disturbances (Fig. 5). In addition, it can express a range of steady and transient locomotion behaviors including slope and stair negotiation, walking and running, accelerating and decelerating, turning and obstacle avoidance (Fig. 6).

4.2 Results from Prior NSF Support Related to this Work

(a) NSF award number: EEC-0540865; amount: $29,560,917; period of support: 6/1/06 - 5/31/15.

(b) Title: NSF Engineering Research Center on Quality of Life Technology (PI: Siewiorek).

(c) Summary of Results: This NSF Engineering Research Center supported during the period 6/2010-5/2013 a research project (lead H. Geyer) on swing leg control in the NMP. We only report on this part of the award.

Intellectual Merit. (1) Identification of a policy-based swing leg controller that robustly places the foot into targets on the ground [18]. (2) Identification of plausible spinal reflex circuits that embed this swing control in an NMP. The NMP generated resulting muscle activations that match human muscle activations in walking and running, providing evidence for a similar control in humans [19].

Broader Impacts. The project has supported our work toward control algorithms in rehabilitation robotics, which may help people who depend on legged assistance to achieve greater mobility and quality of life. The commercially available, powered ankle prosthesis [21, 1] is a testament to the social impact that this work can have. Our results are being shared by being published, and papers are available electronically. The results have been integrated into the curriculum of a graduate course on biomechanics and motor control.

Development of Human Resources. The project trained two graduate students, who presented their results at conferences (ROBIO 2012, ICRA 2013 and EMBC 2013). The graduate students have completed their Master’s degree and won prestigious scholarships (R. Desai: Siebel and Google Anita Borg scholar-
ships) during tenure with the project. They are continuing their training and education within the Robotics Institute’s PhD program.

(d) Publications resulting from this NSF award: [18, 19, 90, 87, 91].
(e) Other research products: None.
(f) Renewed support. This proposal is not for renewed support.

5 Research Plan

Both the model-based optimization approach and the policy-based approach have distinct strengths and weaknesses. The policy-based approach generates quite robust locomotion while being insensitive to kinematic singularities, joint and actuator limits, and model parameters in general, which are all concerns that make it difficult to apply controllers from the model-based optimization approach on humanoid robots. On the other hand, the policy-based approach does not generalize well to automated behavior generation, which requires direct control of foot or hand placement for visually guided locomotion or the use of handholds, canes, walkers, and other walking aids. In this domain, the model-based optimization approach shines. We believe that the fields of humanoid robotics and locomotion science will be pushed forward if the two approaches can be combined to get the best of each. We do not yet know the best combination, and propose to investigate different hybrid approaches.

5.1 Hybrid approaches that we will pursue

Our key concepts for hybrid approaches are to use the neuromuscular policy as either a cost, a constraint, or a replacement (Fig. 7). Each of these ideas can be applied to different stages of the model-based optimization with distinct advantages and shortcomings.

Hybrid Approach 1: Policy as a Cost Term. The concept of the first hybrid approach is to use the neuromuscular control policy (NMP) as a term

\[ J = \sum_i J_i + J_{NMP} \]  

in the cost function at different stages of the hierarchical optimization control. The general advantage of this approach is that it is simple to implement in the state-of-art control frameworks of humanoid controllers [127]. The NMP simply acts as a bias term in existing optimization stages. The main shortcomings

Figure 7. Key concepts for hybrid approaches blending optimal and policy-based control in humanoid locomotion. The neuromuscular policy control acts as a cost (1), constraint (2), or replacement (3) at different levels of the model-based optimization control.
of this approach are that it maintains a high demand on the quality of sensory signals and of the robot-environment dynamics model, $M\ddot{q} + H = S\ddot{r} + J^T\lambda$, and that the computational speed of the optimization remains slow as no dimensionality reduction occurs. We hypothesize that the NMP influence on the cost will improve a humanoid robot’s performance including, for example, more human-like step lengths and timing in locomotion. However, we do not anticipate substantial improvements on locomotion robustness.

**Hybrid Approach 2: Policy as an Optimization Constraint.** Instead of a cost term, the second approach introduces the NMP as a constraint on the hierarchical model-based optimization,

$$\min J \quad s.t. \quad f_i(\bar{x}) = 0, \quad g_j(\bar{x}) > 0,$$

where $\bar{x} = [\bar{q} \quad \bar{\dot{q}}]^T$. Some individual functions of the limbs are common in locomotion. One example is compliant leg behavior. Yet compliant legs tend to buckle [83]. Part of the reflex control in the neuromuscular model counteracts this recurring tendency [29]. Including such common limb functions (domain knowledge) as constraints instead of as costs ensures that the optimization respects known danger zones that threaten locomotion stability. For instance, the knee is particularly prone to buckling, and a torque constraint

$$f_i(\bar{x}) = \tau_k - \tau_{NMP,k},$$

would ensure that the knee torque $\tau_k$ follows the counteraction torque $\tau_{NMP,k}$ produced by the neuromuscular policy. Besides this main advantage, the second approach can also lead to partial dimensionality reduction and a related computational speed improvement if the basis for the optimization is chosen such that it automatically fulfills NMP constraints. The dependence on a high quality dynamics model of the robot remains a shortcoming of this approach. In addition, the implementation of the nonlinear NMP as a constraint will be more challenging. We hypothesize, however, that it improves over the first approach with more stable and robust locomotion behaviors.

**Hybrid Approach 3: Policy as Replacement for Optimization.** The third approach directly replaces optimization by the NMP. The motivation behind this concept is that in contrast to a full optimization using a dynamics model of the entire robot-environment system, the NMP requires much less sensory information and model precision (Fig. 7). For instance, replacing the QP optimization (ii in Fig. 7), the NMP acts as a blackbox that receives a set of task level inputs $\bar{x}_{goal}$ from the dynamics planner (i) and generates desired joint torques for the robot,

$$\bar{\tau}_{cmd} = f(\bar{q}, \hat{\bar{F}}_{cnt}, \hat{\theta}_{float}, \bar{x}_{goals}),$$

based on the current positions $\bar{q}$ of the robot’s joints, a gross estimate $\hat{\bar{F}}_{cnt}$ of the left and right legs’ contact forces, and an estimate of the trunk orientation $\hat{\theta}_{float}$ in the world. The NMP does not require precise measurements of environmental interaction ($\bar{\lambda}$ and $J^T$) or estimates of joint velocities ($\bar{\dot{q}}$). Another advantage of replacement is that it enables very fast computation as optimization is reduced to the low-dimensionality dynamics planner. On the other hand, this approach can only be applied when the topologies match between the NMP and the target robot. Another major shortcoming is that the robots’ behavioral versatility is limited to the behavior versatility of the NMP, which furthermore must be accessible with a few high level inputs $\bar{x}_{goal}$. We hypothesize that the third approach will be the least sensitive to unmodeled robot-environment dynamics for the behaviors that the NMP can address.

5.2 Research activities that we will perform

The starting points for the proposed research are our model-based optimization control framework (detailed in section 3) and neuromuscular control policy (detailed in section 4). The proposed research activities detail the intellectual steps that will take us from these starting points to the evaluation of our hypotheses about hybrid approaches to humanoid control.
Activity 1: Adapt Realtime Optimal Control Framework to ATRIAS and COMAN. Our existing model-based control framework is implemented in realtime on both SARCOS Primus and ATLAS [113, 23]. We will adapt this framework to the ATRIAS and COMAN platforms. The adaptation provides new intellectual challenges. ATRIAS and COMAN have physically compliant, series elastic actuators (SEAs) [71], and it is unclear how these additional states should be treated in the realtime optimization framework. We will investigate two directions, either keeping the low level control of the SEAs separate from optimization or integrating their states into the optimization. From a theoretical standpoint, the second version should provide superior control performance. On the other hand, we have cascaded controllers on the ATRIAS platform for which the SEA control part is an inner loop separated from the behavior control [58, 57]. This separation has provided satisfactory control fidelity in our past work.

Activity 2: Develop Hybrid Approaches with NMP as a Cost or Constraint. Using the NMP as a cost or constraint raises two research questions: Which cost or constraint terms of the NMP maximize control performance? At what stage in the control framework (i-iii, Fig. 7) should the NMP be integrated? We will address the first question by investigating the effect of different cost or constraint terms. In particular, we plan to focus on terms using torques, accelerations, and contact forces. For instance, \( J_{NMP} = \tilde{\tau}^T W_{\tilde{\tau}} \tilde{\tau}_{NMP} + \dot{\theta}^T W_{\dot{\theta}} \dot{\theta}_{NMP} \) would generate a cost biasing torques and accelerations of the optimization toward the NMP. Ideas that we will pursue include the optimization of the cost term weights to find the most robust controller and the ordering of the constraints in a priority list that gets satisfied until degrees of freedom are exhausted. For the second question, we will investigate how the bias with the NMP at different stages affects the control performance. A bias in the planner (i, Fig. 7) may provide better foot placement targets than the linear inverted pendulum currently used in the planner. NMP bias in the full body optimization (ii) can address individual limb functions. Finally, a behavior library populated with the NMP (iii) can provide fast, pre-computed (cached) solutions, and may allow the humanoid to learn behavior over the course of execution.

In general, the NMP constraints are nonlinear constraints. To implement them in the optimization framework, we will use sequential quadratic programming techniques [33].

Activity 3: Develop Hybrid Approach with NMP as a Replacement. Replacing individual stages presents a more radical change than influencing an established optimization framework with additional costs or constraints. To address the feasibility of this approach, we have implemented a preliminary version in a simulation of the 2D ATRIAS testbed (see section 5.3 for details on our testbeds). We removed the foot of the neuromuscular model and adapted it to ATRIAS (Fig. 8a). We then mapped the torques generated by the NMP into desired joint torques (Eq. 6) for a previously validated simulation of the testbed [58, 57]. Finally, we added optimization to plan desired foot placement targets and leg ground clearances in the presence of known obstacles. The resulting control combined robust walking of ATRIAS over terrain with unknown ground level changes up to XXcm (X% leg length) with intentional stepping over obstacles of varying height (tested up to XXcm) and width (XXcm) (Fig. 8b,c) [hg to provide exact results]. The preliminary results suggest the feasibility of the approach. To evaluate it on real hardware, we will embed the NMP replacement.
for the individual robot platforms within the existing realtime optimal control framework (activity 1).

In addition, we will address the research question of whether the NMP can be formulated in a hierarchy that controls versatile behavior with a few high-level task goals $\bar{x}_{\text{goal}}$. Specifically, we will investigate if a hierarchy of reflex control can be identified for the key inputs of target speed, target gait, and target turning rate. Our original NMP produced only walking [29]. An understanding of how the swing leg control of foot placement can be formulated with a hierarchical structure [18, 19] led us to an NMP which can express a rich set of behaviors from walking and running to slope negotiation and turning [88] (compare Fig. 6). Although we can formulate offline optimizations that generate transitions between these behaviors by adapting the reflex control parameters, we do not yet understand how these adaptations can be reinterpreted with hierarchical control structures. A solution to this question not only would enrich the online NMP replacement approach, but also can provide new insights into human sensorimotor control. A human model that adapts its locomotion behaviors based on higher level goals would provide a direct hypothesis about how humans adapt locomotion behaviors.

**Activity 4: Implement, Evaluate and Refine Hybrid Approaches in Simulation.** We will implement the hybrid approaches in simulation to evaluate their performance and refine them before conducting actual hardware tests. The simulations will model the experimental tests that we plan to perform within the robot testbeds. For this activity, we will build on our previous work, in which we have developed similar simulations for the ATRIAS, SARCOS and ATLAS platforms [57, 25, 95, 113, 23].

**Activity 5: Implement, Evaluate and Refine Hybrid Approaches in Robot Experiments.** We will implement the different hybrid approaches on four different robot testbeds. The details of our experimental evaluation plan, including performance metrics and benchmarks, robot testbeds, and data collection, are elaborated in the next section.

### 5.3 How we will evaluate controller performance in experiments

To test our hypotheses about the performance of the three hybrid approaches, we will evaluate the human-likeness, the robustness to external perturbations, the robustness to internal model uncertainty, and the computational cost of the individual humanoid controllers in robot testbeds of increasing complexity.

**Performance Metrics and Benchmarks.** Well-defined standards and benchmarks for assessing locomotion in humanoid robots are rare [107]. Benchmarks that have been used often focus on accomplishing a specific task with discrete pass or fail metrics (climb ladder, open door, negotiate rubble field, [17]). Established performance metrics for a more continuous comparison of competing locomotion controllers include energy-efficiency, the likeness to human locomotion, and the assessment of push recovery [107].

We will quantify human-likeness using established methods and metrics from biomechanical gait analysis focusing on speed, cadence, and duty factor [42, 69], as well as on similarity of kinematic, kinetic, and
ground reaction force patterns [69, 118]. We will evaluate locomotion robustness to external perturbations by quantifying the recovery from pushes in the sagittal and frontal planes, and the survival rate on terrain with unknown ground disturbances of increasing severity. The pushes will be administered with a force-sensor equipped device to a fixed location on the torso near the COM of the robots to quantify disturbance impulse. (We have successfully used this method in previous work assessing standing balance [98].) The ground disturbances will include tracks with pits of rocks (three different granularities ranging from 1cm to 10cm) and of tiny roundball pellets (6mm, emulate behavior of sand), and slippery surfaces of decreasing friction coefficient (steel, oil layer). We will evaluate the robustness to internal model uncertainty by systematically changing the model parameters (mass and inertial properties) and adding artificial sensory noise in the software of the controller during locomotion on level ground. Finally, we will quantify the computational cost by tracking the execution time for each controller.

Benchmarks for some of these metrics have been reported in the literature; however, these benchmarks are highly dependent on the physical capability of the individual robot hardware platform. To objectively assess the benefits and shortcomings of the proposed hybrid approaches to humanoid control, we will establish benchmark values for each of our robot platforms using the state of the art model-based optimization control (detailed in section 3) as a null hypothesis.

**Robot Testbeds.** The groups of the PIs are equipped with humanoid robots of different morphology (ATRIAS and SARCOS Primus) and have access to another humanoid with different actuation and sensing hardware (COMAN) (Fig. 1). This access to a range of hardware platforms will provide us with the unique opportunity to quantify our results across platforms. Such a systematic evaluation of performance, either in simulation or on actual robots, has rarely been done in the field of humanoid robotics, including our previous work.

We will evaluate the control approaches on the simplest robot testbed first (7 degrees of freedom, DOF) and then systematically advance to the testbeds with increasing complexity (up to 34 DOF). First, we will use CMU’s ATRIAS biped constrained to a boom (Fig. 1A[check]). With only 3 external and 4 internal DOF, the testbed substantially simplifies the realtime implementation of the different controllers. Second, we will take ATRIAS off the boom (Fig. ??B[check]). This will increase the level of complexity to 6 external and 6 internal DOF, and will allow us to assess the performance of the controllers including lateral balance. Third, we will apply and test the controllers on CMU’s SARCOS Primus humanoid. With 12 internal DOF used for locomotion, the humanoid matches the full complexity of the neuromuscular model policy control. We will also explore full body control with 34 DOF (includes legs, arms, torso, and neck, but not fingers, eyes, or mouth). Finally, we plan to test the generalizability of our results with the COMAN and ATLAS platforms, which share the complexity of the SARCOS robot.

**Data Collection.** All of the humanoid robots are equipped with position, force and torque sensing. In addition, we will use our motion capture and force plate systems to collect ground truth for joint motions and ground reaction forces during the experiments. The measurements are sufficient to analyze the human-likeness and the robustness to external and internal disturbances with the proposed metrics.

### 6 Management Plan

#### 6.1 Roles of the Investigators

This project will bring together researchers with complementary areas of expertise attacking an interdisciplinary problem. The key personnel will include the two PIs, Chris Atkeson and Hartmut Geyer (CMU Robotics Institute), one postdoctoral researcher, and two graduate students.

**Principal Investigators.** Chris Atkeson has expertise in model-based optimal control of humanoids [95, 113, 23], robot learning, behavior libraries, and humanoid robotics. Hartmut Geyer is an expert in human neuromechanics and robotics. He has broad experience in dynamics and control of legged systems including
conceptual gait models [81, 31, 32, 22, 18, 120], computational models of human neuromuscular control policies [30, 29, 89, 19, 90, 87, 91, 88], and powered prosthetics and legged robots [21, 79, 58, 57]. He also has experience with experimental gait analysis [82, 30, 18, 19].

Neither PI can do the project work alone. A joint effort is necessary to combine our expertise in this interdisciplinary project involving humanoid robotics, model-based optimization, and human neuromechanics. Our collaboration will provide a thrust towards a new generation of students that can take on the challenging questions in this interdisciplinary area of research.

**Postdoctoral Researcher.** We have found that for research involving technologically advanced platforms, the combination of a postdoctoral researcher with several faculty and students provides a very productive mix of expertise. It accelerates the training and learning of the graduate students while maintaining a high level of research productivity. We will recruit a postdoctoral researcher with expertise in humanoid robotics to ensure such a productive environment. She will participate in all activities of the project with a focus on developing a single framework that can seamlessly blend optimization and policy-based control. The attached mentoring plan details our mentoring of the postdoctoral researcher.

**Graduate Students.** The project will involve two graduate students who will be jointly supervised by the PIs and the postdoctoral researcher. For the theoretical parts of the project, one student will focus more on the hybrid approaches using the NMP as a cost or constraint (activity 2), and the other student will focus more on the NMP as a replacement (activity 3). For the experimental validation in simulation and on the robot platforms, the students will work together (activities 1, 4 and 5). The students will be trained by the PIs in legged systems and optimal control theory as well as in model-based optimization and human neuromechanics. They will also acquire expertise in advanced techniques and algorithms for rigid body simulation and realtime hardware control.

**Project Coordination.** The PIs will jointly coordinate the activities supported by this project and will be responsible for all aspects of it. For the day-to-day administration and financial management, they will be supported by the staff of the CMU Robotics Institute. Frequent ad hoc meetings and interactions of all personnel involved are guaranteed as as the PIs’ groups share lab space and participate in a weekly, joint CMU seminar on bipedal locomotion. In addition, all personnel in this project will meet bi-weekly in a more formal setting to review research results and discuss future directions.

### 6.2 Time Line of Activities

**Year 1.** *Theory:* Develop hybrid approaches for planar locomotion (Activities 2 & 3). *Simulation Experiments:* Implement, evaluate and refine theory in simulation experiments of 2D ATRIAS platform (A1 & 4).

**Year 2.** *Theory:* Develop hybrid approaches for 3D locomotion (A2 & 3). *Simulation Experiments:* Implement, evaluate and refine theory in simulation experiments of 3D ATRIAS and SARCOS platforms (A1 & 4). *Robot Experiments:* Evaluate hybrid approaches on 2D (early) and 3D ATRIAS testbed (later) (A5).

**Year 3.** *Theory:* Refine theory based on year 2 robot experiments (A2 & 3). *Simulation Experiments:* Generalize in simulation experiments to ATLAS and COMAN platforms (A1 & 4). *Robots:* Evaluate hybrid approaches on SARCOS (early), ATLAS and COMAN testbeds (later) (A5).

### 6.3 Risk Assessment and Mitigation

**Theory Development Risks (Activities 2 and 3).** Our ideas may not work out as expected, and we may encounter unforeseen barriers to progress. *Mitigation:* The proposed research activities include the exploration of different cost and constraint modalities as well as alternative stages for the integration of the policy based control into the optimization framework (Fig. 7). For activity 2, we have working baseline controllers for
both the SARCOS and ATLAS robots. For activity 3, we have preliminary results that support the viability of the replacement approach (Sec. 5.2). Additionally, the intellectual independence of the three approaches provides a layer of risk mitigation should the theoretical development for one of them be delayed or fail.

**Implementation Risks (Activities 1, 4 and 5).** As for any project including hardware evaluation, we face implementation risks. We may not be able to achieve real time performance of the blended control framework. The humanoid robot platforms may break or have mechanical limitations that hamper implementation (power output limited on ATRIAS due to amplifiers, torque control limited on ATLAS and SARCOS due to stiction in actuators and play in transmission and bearings, joint velocity estimation required for approaches 1 and 2 limited on ATLAS and SARCOS). **Mitigation:** We have extensive prior experience working with three of the four hardware platforms [58, 57, 98, 25, 95, 113, 23] including realtime control implementation and demonstration of locomotion controllers. We are currently updating some of the platforms to improve power output (ATRIAS) and joint velocity estimation (ATLAS and SARCOS). Working with multiple platforms also balances the risk of an implementation delay or failure on a single platform.

7 Broader Impact

**Unplanned Broader Impact.** Often the broader impacts of our work are serendipitous, and not planned in advance. Examples of such ad hoc broader impacts from our recent work include: 1) Our technologies being demonstrated on entries in the DARPA Robotics Challenge. 2) A graduate student was a participant on a Discovery Channel TV series, “The Big Brain Theory: Pure Genius”. One purpose of the TV series is getting people excited about engineering. 3) A graduate student helped run a group of all female high school students in the robot FIRST competition. 4) Our work on soft robotics inspired the soft inflatable robot Baymax in the Disney movie Big Hero 6 [?]. We have participated in extensive publicity as a result. An explicit goal of this movie was to support STEAM. We expect similar unplanned broader impacts to result from this work, especially based on dramatic videos of agile robots.

**Development of Human Resources.** Students working on this project will gain experience and expertise in teaching and mentoring by assisting the course students in class projects. Students working on this project will also have the opportunity to train their communication and inter-personal skills, as they will actively participate in the dissemination of the research results at conferences and in related K-12 outreach programs of the Robotics Institute.

**Participation of Underrepresented Groups.** We will attract both undergraduate and graduate students, especially those from underrepresented groups. We will also make use of existing efforts that are part of ongoing efforts in the Robotics Institute, and CMU-wide efforts. These efforts include supporting minority visits to CMU, recruiting at various conferences and educational institutions, and providing minority fellowships. CMU is fortunate in being successful in attracting an usually high percentage of female undergraduates in Computer Science. Our collaboration with the Rehabilitation Science and Technology Department of the University of Pittsburgh in the area of assistive robotics is a magnet for students with disabilities and students who are attracted by the possibility of working on technology that directly helps people.

**Outreach.** One form of outreach we have pursued is an aggressive program of visiting students and postdocs. This has been most successful internationally, with visitors from the Delft University of Technology (4 students) [6, 56, 119], the HUBO lab at KAIST (1 student and 1 postdoc) [9, 14, 49], and the Chinese Scholarship Council supported 5 students [53, 122]. We welcome new visitors, who are typically paid by their home institutions during the visit. We are currently experimenting with the use of Youtube and lab notebooks on the web to make public preliminary results as well as final papers and videos. We have found this is a useful way to support internal communication as well as potentially create outside interest in our work. We will continue to give lectures to visiting K-12 classes. The Carnegie Mellon Robotics Institute already has an aggressive outreach program at the K-12 level, and we will participate in that program, as well
as other CMU programs such as Andrew’s Leap (a CMU a summer enrichment program for high school students), the SAMS program (Summer Academy for Mathematics + Science: a summer program for diversity aimed at high schoolers), Creative Tech Night for Girls, and other minority outreach programs.

**Curriculum Development.** We will pursue multiple directions for dissemination. First, we will develop course material on robot control and biologically inspired approaches to humanoid and rehabilitation robotics which will directly be influenced by the planned activities of this proposal. The PIs currently teach several courses that will benefit from this material. The first course, *16-642: Manipulation, Mobility & Control*, is part of the Robotics Institute’s recently established professional Master’s degree program that aims at training future leaders in the workforce of robotics and intelligent automation enterprises and agencies. Another course, *16-868: Biomechanics and Motor Control*, directly addresses the research areas in which this proposal is embedded. We teach a course *16-745: Dynamic Optimization*, which addresses optimal control. We also teach a course design to attract undergraduates into the field, *16-264: Humanoids*. All of these courses emphasize learning from interaction with real robots. We will make these course materials freely available on the web.

**Dissemination Plan.** For a more complete description of our plans for dissemination, see our Data Management Plan. We will maintain a public website to freely share our simulations and control code and to document research progress with video material. Finally, we will present our work at conferences and publish it in journals, and use these vehicles to advertise our work to potential collaborators in science and industry.

**Benefits to Society.** The proposed research has the potential to lead to more useful robots. A successful outcome can enable practical controllers for robust locomotion of legged robots in uncertain environments with applications in disaster response and elderly care. The outcomes of this project may also provide new insights into human sensorimotor control, in particular, into how humans adapt locomotion behaviors. Understanding how humans actually control their limbs has the potential to trigger neural and robotic prostheses and exoskeletons which restore legged mobility to people who have damaged or lost sensorimotor functions of their legs due to diabetic neuropathy, stroke, or spinal cord injury as well as improve fall risk management in older adults [109, 16, 55, 121].

**Enhancement of Infrastructure for Research and Education.** The project will upgrade CMU’s ATRIAS biped robot testbed to a 3D locomotion testbed, a powerful tool for research in dynamics and control of legged locomotion and an equally powerful tool for education. This testbed and our Sarcos robot regularly attract the interest of students and inspires them to advanced education and training in robotics.

**Relationship to DRC funding.** We are currently part of one of the funded teams in the DARPA Robotics Challenge [17]. This support will end in the summer of 2015. The DRC focuses on reliability, implementation issues, and logistics. We are not able to develop the proposed ideas in the time frame of the DRC, so longer term NSF funding complements our soon to end DRC funding.
References


