Multiple Balance Strategies
From One Optimization Criterion

Christopher G. Atkeson and Benjamin Stephens
Robotics Institute, Carnegie Mellon University
www.cs.cmu.edu/~cga, cga@cmu.edu, bstephens@cmu.edu

Abstract—**Multiple strategies for standing balance have been observed in humans, including using the ankles to apply torque to the ground, using the hips and/or arms to generate horizontal ground forces, and using the knees and hips to squat. This paper shows that multiple strategies can arise from the same optimization criterion. It is likely that humanoid robots will exhibit the same balance strategies as humans.**

I. INTRODUCTION

This paper addresses two questions: 1) Will humanoid robots show the same multiple strategies for standing balance as seen in humans? and 2) If so, do these multiple strategies arise from independent design and control processes, or a single design and control process? This paper demonstrates that it is likely that humanoid robots with backdrivable joints will exhibit the same behavioral strategies seen in humans. A design process using a single optimization criterion results in a controller using multiple strategies. A small perturbation is handled using one strategy, while a large perturbation is handled using another strategy.

Studies of human standing balance have revealed several strategies to compensate for perturbations: the *ankle strategy*, in which torque about the ankle joints is used to balance and the rest of the body is held in a fixed posture, the *hip strategy*, in which torque about the hip joints is used to generate horizontal ground forces moving the center of mass, the *squat strategy*, in which the knees and hips are flexed to lower the center of mass [1], and the *step strategy*, in which a step is taken [2], [3]. These multiple strategies reflect the mechanical constraints faced by humans and humanoid robots. One question this paper addresses is whether each strategy is controlled by a separate controller, as in human eye movement control, or whether a single controller and design process can be used to generate all strategies. In studies of humans, the ankle strategy seems to be used for small and slow perturbations on flat rigid support surfaces, while the hip strategy seems to be used for large or fast perturbations and on narrow or compliant support surfaces [4].

This paper focuses on two balance strategies that do not involve stepping: the ankle and the hip strategies. A future paper will attempt to include stepping. From a humanoid robotics point of view, the ankle strategy turns the body into an inverted pendulum, balanced upright using ankle torque. Hip torque is applied only to keep the hip joint in a fixed position. The hip strategy is that of a two link acrobot [5], where only hip torque is applied and the ankle is unactuated. The acrobot uses hip torque to generate horizontal ground forces, which keeps the center of mass over the foot on average.

Hemami and colleagues analyze the ankle strategy [1]. Several researchers provide examples of humanoid robot standing balance implemented using two hand designed or optimized controllers, one for the ankle strategy and one for the hip strategy, including [6], [7]. Guibard and Gorce present a classifier to select between ankle and hip strategies. Each strategy is separately optimized using different criteria and/or constraints [8]. Kudoh and colleagues choose between an optimized strategy and a predesigned feedback control strategy based on the current state. The optimization finds the best acceleration on the current step (a local or greedy optimization) rather than creating the best response over time [9]. Kuo designs Linear Quadratic Regulators for each perturbation to generate controllers [10]. Different response strategies are generated by changing the optimization criterion based on the size of the perturbation. In Kuo's work the system is linearized. In our work we use a single optimization criterion for all perturbation sizes. Thus, we do not need to "recognize" a perturbation in order to select the appropriate response. We also find that a nonlinear controller outperforms linear controllers and linear controllers with constrained outputs. Abdallah and Goswami explore a balance approach in which two strategies are used in temporal sequence [11].

II. THE ONE LINK MODEL

Figure 1 shows a one link inverted pendulum responding to a perturbation in the sagittal plane (fore/aft motion only). In this simple model all joints except the ankle are held in a fixed position. This model of standing balance has only one strategy available to it, applying torque at the ankles. The amount of ankle torque is limited by the size of the feet. We will use this example to describe our optimization approach.

The model is facing to the right. The one link model is 2 meters high and has a total body weight of 70kg. The ankle angle was bounded by $-0.4 < \theta_a < 0.8$ radians. $\theta_a = 0$ is upright. We assume that in standing the center of pressure is at the center of the foot. We therefore use a symmetric foot 0.2 meters long in our model. This results in a maximum ankle torque of approximately ± 70 Newton-meters.

In this case we model perturbations as impulses applied to the middle of the torso (1.5 meters above the ankle). In this example we present only perturbations that push the model forwards, to simplify our figures. For this model perturbations
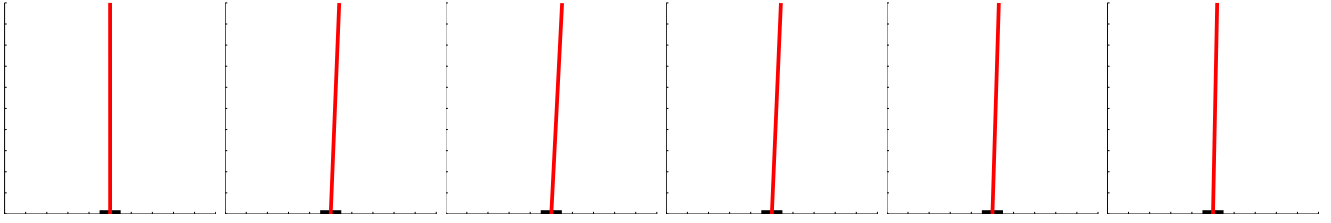
Fig. 1. Configurations every quarter of a second of a simulated one link inverted pendulum response to an impulse of 15 Newton-seconds forward (to the right) The black rectangle indicates the extent of the symmetric foot.
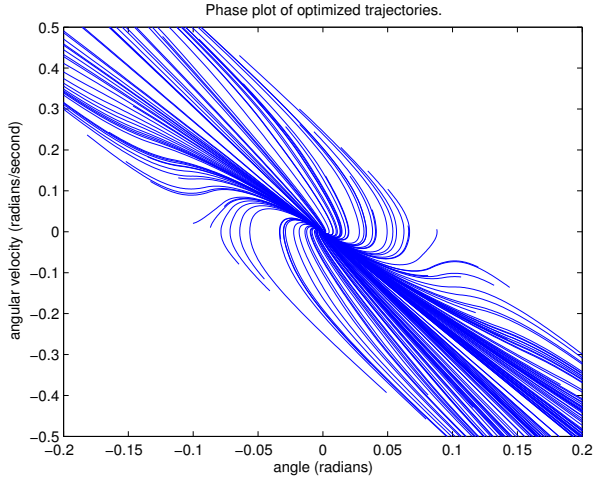

Fig. 2. Optimized trajectories from different starting points.

Fig. 3. Ankle angle trajectories for a range of perturbation sizes (2.5, 5, 7.5, 10, 12.5, and 15 Newton-seconds).

in the other direction lead to symmetric responses. The perturbations instantaneously change the joint velocities from zero to values appropriate for the perturbation:

$$\Delta\dot{\theta} = \mathbf{H}^{-1}\mathbf{J}^{\mathrm{T}} \int force \cdot dt \qquad (1)$$

where elements of the inertia matrix $\mathbf{H}(\theta)$ are the coefficients of angular accelerations in the rigid body dynamics ($\tau = \mathbf{H}\ddot{\theta} + ...$), $\mathbf{J}(\theta)$ is the Jacobean matrix, and $\int force \cdot dt$ is the impulse applied (measured in Newton-seconds). We assume no slipping or other change of contact state during the perturbation.

The one step optimization criterion is a combination of quadratic penalties on the deviations of the joint angles θ_a from zero (straight up) and on the joint torques τ_a: $L(\mathbf{x}, \mathbf{u}) = T * \theta_a^2 + 0.002 * T * \tau_a^2$ where 0.002 weights the torque penalty relative to the position error and T is the time step of the simulation (0.01s). There are no costs associated with joint velocity. The one step criterion is summed over each step of a trajectory to find the total trajectory cost, and the optimal trajectory minimizes the total trajectory cost.

Any trajectory or control law optimization method can be used to find optimal trajectories. To generate control laws for the balancers discussed in this paper, we generate optimal trajectories from either many randomly selected starting points (Figure 2) [12] or a grid of starting points. Each trajectory is locally optimized using Differential Dynamic Programming
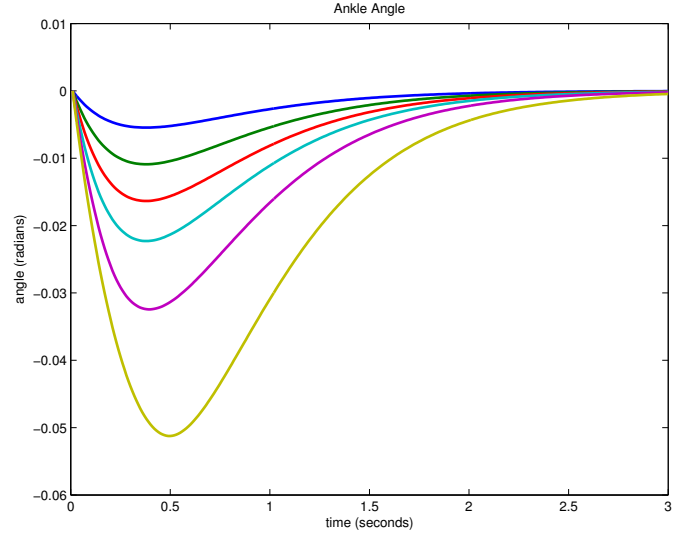
(DDP) [13], [14], [15], [16]. Information is exchanged between trajectories to enable convergence to globally optimal trajectories [12].

Figure 3 shows ankle trajectories for a variety of perturbation sizes, and Figure 4 shows the corresponding torques. Note that the torques saturate for larger perturbations, and that the joint angle motion roughly scales with perturbation size. For this one link system there is only one strategy, so the maximum joint displacements are grow almost linearly with perturbation size (Figure 5).

We have designed a Linear Quadratic Regulator (LQR) that optimizes the same criterion on the one link model. For perturbations of 10 Newton-seconds and higher, this LQR controller saturates the ankle torque as does the controller presented here. For perturbations of 17.5 Newton-seconds and higher, the LQR controller falls down, as does the controller presented here.

III. THE TWO LINK MODEL

In contrast to the one link model, the two link model shows multiple strategies. Figure 6 shows a two link inverted pendulum responding to a perturbation in the sagittal plane. This multi-link pendulum is a useful model for standing balance of a human and a humanoid robot. The bottom link is
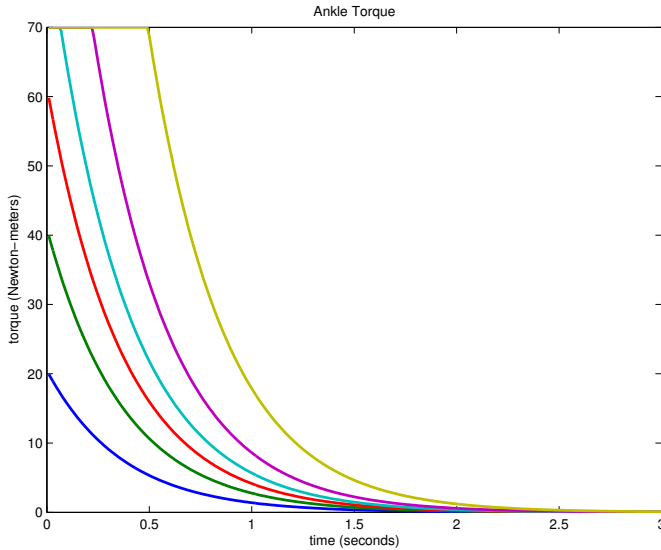
Fig. 4. Ankle torque for the same range of perturbation sizes.
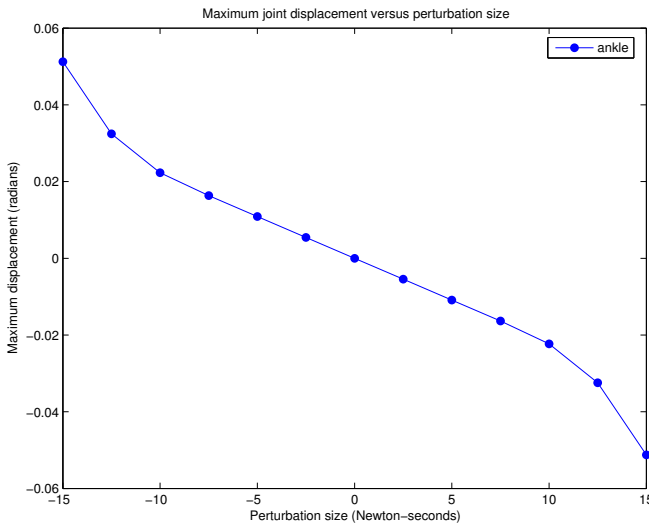

Fig. 5. Maximum joint displacement versus perturbation size.

the leg, and the top link is the torso. Each link is 1 meter long, and is modeled as a thin rod of 35kg for a total body weight of 70kg. We assume that in standing the center of pressure is at the center of the foot. We therefore use a symmetric foot 0.2 meters long in our model.

The model is facing to the right. We model perturbations as impulses applied to the middle of the torso (the top link). In this example we initially present only perturbations that push the model forwards. We summarize results for perturbations in the other direction in a later figure.

Perturbations instantaneously change the joint velocities from zero to values appropriate for the perturbation, as described in Equation 1. The states were bounded by $-0.4 < \theta_a < 0.8$ radians and $-1.5 < \theta_h < 0.1$ radians where θ_a is the ankle angle, θ_h is the hip angle, and $\theta_i = 0$ is upright. The ankle torque was bounded by ± 70Nm to keep the center

of pressure within the foot. The hip torque was bounded by ± 500Nm which approximately matches human capabilities and our humanoid robot. The one step optimization criterion is a straightforward extension of the one link optimization criterion: a combination of quadratic penalties on the deviations of the joint angles from zero (straight up) and on the joint torques: $L(\mathbf{x}, \mathbf{u}) = \mathrm{T} * (\theta_a^2 + \theta_h^2) + 0.002 * \mathrm{T} * (\tau_a^2 + \tau_h^2)$ where 0.002 weights the torque penalty relative to the position error. and T is the time step of the simulation (0.01s). There are no costs associated with joint velocities.

Figures 7 and 9 show the optimized joint angle trajectories for the ankle and knee for a variety of perturbation sizes, and Figures 8 and 10 show the corresponding joint torques. Figures 7 and 9 clearly show the difference in strategy between the responses to the smaller perturbations (2.5, 5, 7.5, 10, 12.5, and 15 Newton-seconds) shown with solid lines and the responses to the larger perturbations (17.5, 20, 22,5, and 25 Newton-seconds) shown with dashed lines. For the small perturbations the ankle angle is entirely negative and the hip movement is small. For the large perturbations the ankle initially moves in a positive direction, and the hip movement is large. In all cases the ankle torque is similar: initially saturating at 70Nm if that torque is reached, and then smoothly decreasing to zero. The hip torques are small and initially positive for small perturbations, but large and initially negative for large perturbations.

Figure 11 shows the maximum ankle and hip displacements for perturbations in both directions (positive perturbations push forwards, negative perturbations push backwards) for a normal size foot and a half sized foot (equivalent to standing on a narrow platform). For smaller perturbations the responses are linearly related to the perturbation size, and are independent of the size of the foot. These responses are associated with the ankle strategy. At some point, the hip strategy is chosen and the ankle and hip perturbations rapidly increase with perturbation size. The location of the strategy transition depends on the size of the foot, which is related to the maximum torque the ankle can apply. The asymmetry of this figure is generated by the angle limits on the hip. The back cannot bend backwards very much (0.1 radian), limiting the use of the hip strategy for backwards (negative) perturbations. Instead the model falls down for these perturbations.

We have designed an LQR controller that optimizes the same criterion on the two link model with the full sized foot. For perturbations of 10 Newton-seconds and higher, the LQR controller saturates the ankle torque, as does the controller presented here. For perturbations of 20 Newton-seconds and higher, the LQR controller falls down, while the controller presented here is able to handle larger perturbations. For the half sized foot, the LQR controller saturates the ankle torque at 5 Newton-seconds, and falls for perturbations of 10 Newton-seconds or more. Our controller for the half sized foot saturates at the same level but can handle perturbations up to 15 Newton-seconds in the forward direction (taking advantage of large hip flexing).
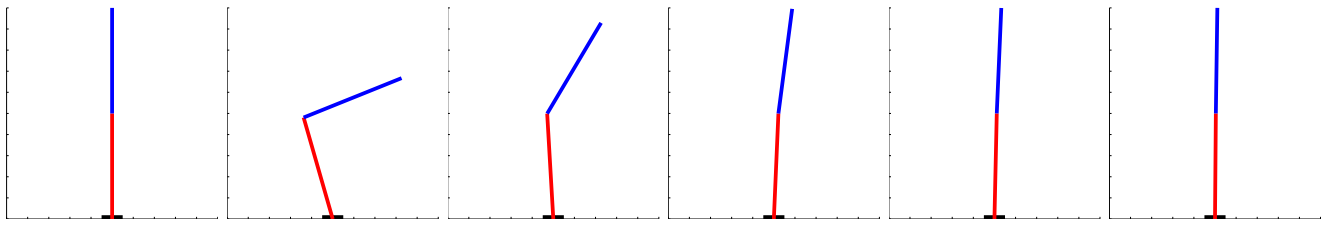
Fig. 6. Configurations every half second of a simulated two link inverted pendulum response to an impulse of 25 Newton-seconds forward (to the right) The black rectangle indicates the extent of the symmetric foot.
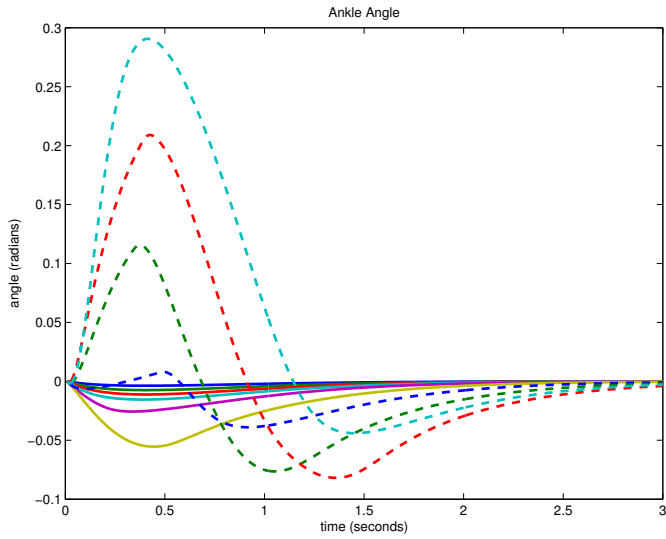

Fig. 7. Ankle angle trajectories for a range of perturbation sizes (2.5, 5, 7.5, 10, 12.5, 15, 17.5, 20, 22,5, and 25 Newton-seconds).
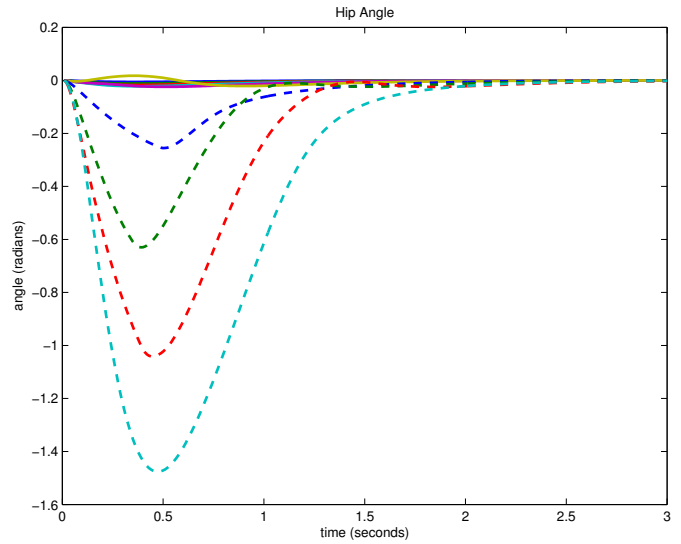

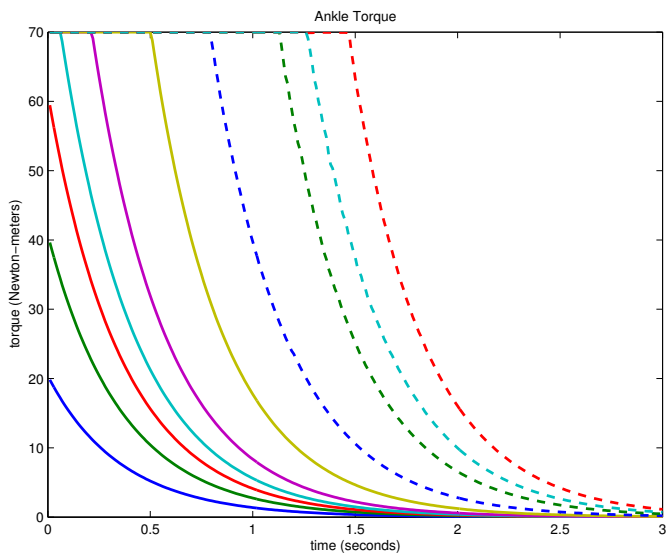Fig. 9. Hip angle trajectories for the same range of perturbation sizes.

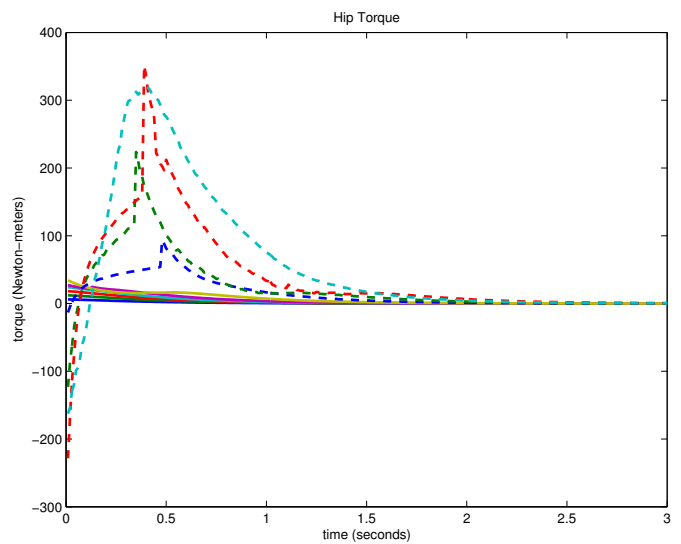Fig. 8. Ankle torque for the same range of perturbation sizes.

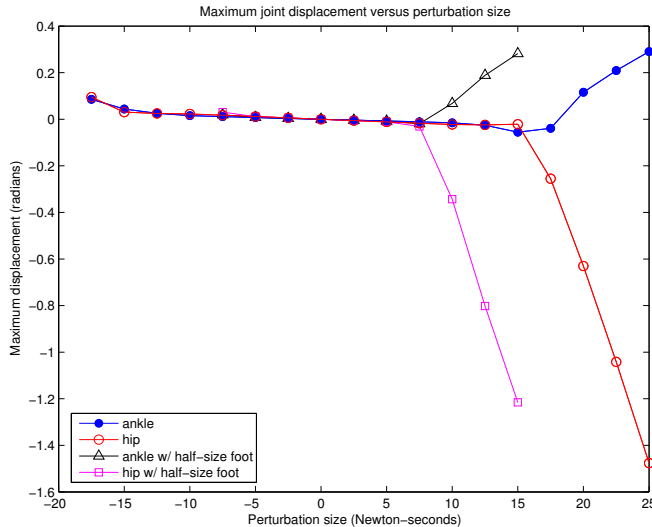Fig. 10. Hip torque for the same range of perturbation sizes.

Fig. 11.   Maximum joint displacement versus perturbation size.

IV. THE FOUR LINK MODEL

To explore more complex and human-like strategies, we created a four link model that included a knee, shoulder, and arm by subdividing the two link model (Figure 12). Each link is modeled as a thin rod, with a calf and thigh length of 0.5 meters, and 17.5kg each. The torso is 1 meter long with a weight of 26.25kg, the arm is 1 meter long with a weight of 8.75kg. The 0.2 meter symmetric foot remained the same. Impulse perturbations were applied horizontally to the middle of the torso as described in Equation 1.

The states were bounded by $-0.4 < \theta_a < 0.8$ radians, $-0.01 < \theta_k < 2.5$ radians, $-1.5 < \theta_h < 0.1$ radians, and $-0.5 < \theta_s < 2.5$ radians where θ_a is the ankle angle, θ_k is the knee angle, θ_h is the hip angle, θ_s is the shoulder angle, and $\theta_i = 0$ is upright with the arms hanging down. The ankle torque was bounded by ± 70Nm to keep the center of pressure within the foot. The knee and hip torques were bounded by ± 500Nm. The shoulder torque was bounded by ± 250Nm. The one step optimization criterion is a combination of quadratic penalties on the deviations of the joint angles from zero (straight up), the joint velocities, and the joint torques: $L(\mathbf{x}, \mathbf{u}) = T * (\theta_a^2 + \theta_k^2 + \theta_h^2 + \theta_s^2) + T * (\dot{\theta}_a^2 + \dot{\theta}_k^2 + \dot{\theta}_h^2 + \dot{\theta}_s^2) + 0.002 * T * (\tau_a^2 + \tau_k^2 + \tau_h^2 + \tau_s^2)$ where 0.002 weights the torque penalty relative to the position and velocity errors. and T is the time step of the simulation (0.01s). In the four link case we introduced a penalty on joint velocities to reduce knee and shoulder oscillations.

Figures 12 and 13 show responses to the largest perturbations that could be handled in each direction. Figure 14 shows the optimized shoulder angle trajectories for a variety of perturbation sizes, Note how the maximum shoulder angle grows suddenly only for the largest perturbation sizes. The shoulder angle is softly constrained to a lower limit of -0.5 radians, which affects the trajectories for the three largest perturbations. The red dashed line in Figure 14 shows a change in arm strategy for the largest backward perturbation (shown

in Figure 13). The arm is initially moved backwards rather than forwards, as was done in the next smallest backwards perturbation (top green line in Figure 14).

Figure 15 shows how the joint maximum deviations initially grow linearly with perturbation size, until an impulse size of approximately 15 Newton-seconds. At this point the strategy changes, and all joints rapidly increase their maximum deviation with perturbation size. This movement is a generalization of the two joint hip strategy, involving the shoulder and knees as well as hips and ankles. Figure 16 shows the effect of reducing the size of the foot (or equivalently standing on a smaller platform). Comparing Figure 16 to Figure 15, we see the hip strategy is used for impulses of 10 Newton-seconds or more for the smaller foot, and that the hip strategy is not effective for backwards perturbations. Interestingly, we see an example of the squat strategy (bent knee and hip) in the response to the largest perturbation in the forward direction (Figure 17).

We have designed an LQR controller that optimizes the same criterion on the four link model with the full sized foot. For perturbations of 12.5 Newton-seconds and higher, the LQR controller saturates the ankle torque. The controller presented here saturates at 10 Newton-seconds. For perturbations of 17.5 Newton-seconds and higher, the LQR controller falls down, while the controller presented here is able to handle larger perturbations. For the half sized foot, the LQR controller saturates the ankle torque at 5 Newton-seconds, and falls for perturbations of 10 Newton-seconds or more. Our controller for the half sized foot saturates at the same level but can handle perturbations up to 15 Newton-seconds in the forward direction (taking advantage of large hip flexing).

V. SUMMARY AND FUTURE WORK

This paper shows that a single (and rather simple) optimization criterion can be used to generate the multiple balance recovery strategies. It appears as if the strategies arise from the mechanical constraints of a jointed structure standing in a gravity field. If possible, the less expensive ankle strategy is used for recovery. If that will not be sufficient, the more expensive hip strategy is used. We expect the same strategies for standing balance will be seen in humanoid robots as are seen in humans, due to the similarity of the mechanical constraints.

This work needs to be extended in several ways. The first is a detailed comparison to human experimental data. In this work simple body models were used to facilitate comparison of models with different numbers of joints. Future work will use a more detailed and accurate model of the human body, including models of the soft foot tissue and floor compliance. Delays will be implemented between sensing and actuation. Imperfect sensing will also be implemented. The location and direction of the perturbation will be varied. Under these conditions, the behavior generated by various optimization criteria will be compared to human experimental data.

The second extension is to actually implement this algorithm on a robot. This will require state estimation based on
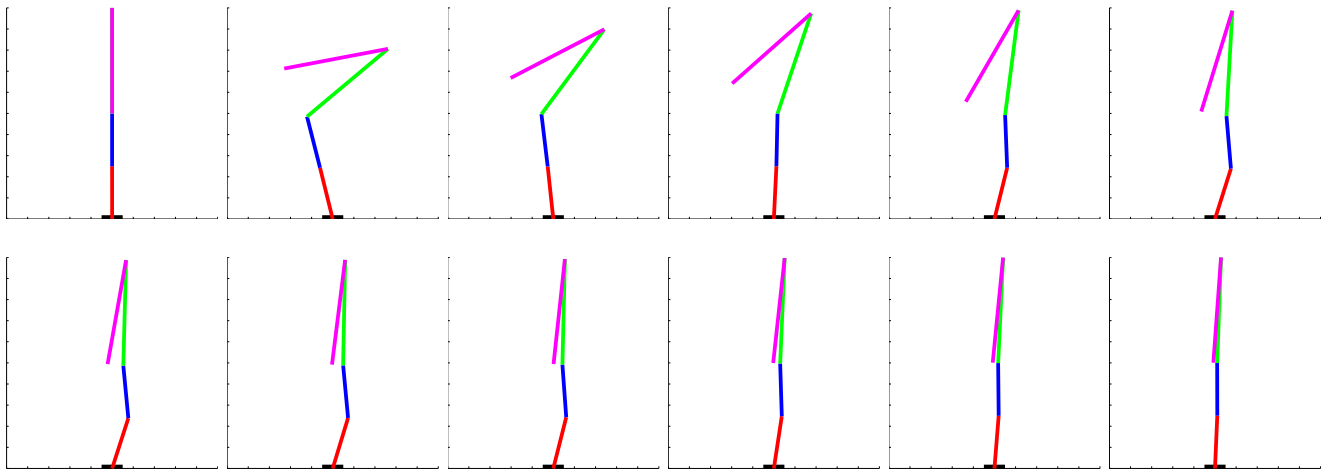
Fig. 12. Configurations every quarter second from a simulated four link inverted pendulum response to a forward impulse (to the right) of 22.5 Newton-seconds. The black rectangle indicates the extent of the symmetric foot.
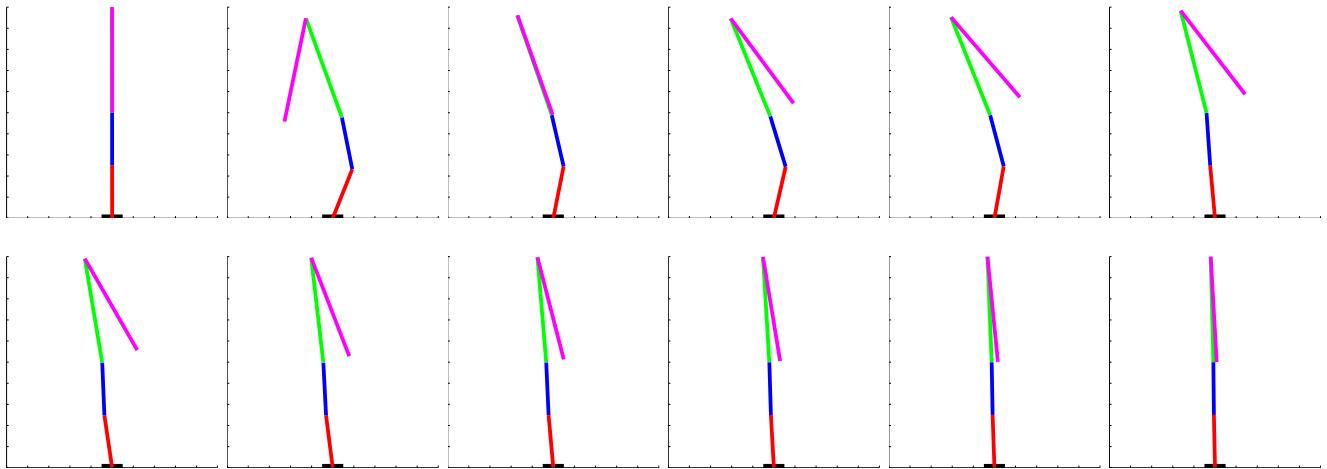

Fig. 13. Configurations every quarter second from a simulated four link inverted pendulum response to a backward impulse (to the left) of 22.5 Newton-seconds

imperfect sensing and dealing with floor compliance. It also requires coordinating the action of both legs and feet [17].

A third extension is to handle perturbations that take a finite amount of time. In this work we avoided the issue of recognizing and predicting the future course of perturbations by applying them instantaneously. However, if a push lasts for a second, it is possible to take that into account in generating a response. If a push will last for an unknown time, the response in the same state might be different depending on the beliefs of the subject as to the future course of the perturbation. It is also possible to recognize perturbations to assist this prediction process. One way to handle this in optimization is to incorporate information about the future course of the perturbation in the state. For complex predictions this might make the optimization much more complex. A simplification is to assume the current perturbation will last indefinitely.

Another extension is to consider stepping as a possible response. We hope to unify the ankle, hip, and step strategies with a single optimization criteria in the future. It may be

the case that when stepping is possible it is chosen instead of using the hip strategy. It may be the case that ankle and hip strategies continue to occur during a step.

REFERENCES

[1] H. Hemami, K. Barin, and Y.-C. Pai, "Quantitative analysis of the ankle strategy under platform disturbance," *EEE Transactions on Neural Systems and Rehabilitation*, vol. 14, pp. 470–80, 2006.
[2] L. Nashner and G. McCollum, "The organization of postural movements: A formal basis and experimental synthesis," *Behavioral Brain Sciences*, vol. 8, pp. 135–72, 1985.
[3] F. Horak and L. Nashner, "Central programming of postural movements: adaptation to altered support-surface configurations," *J Neurophysiology*, vol. 55, no. 6, pp. 1369–81, 1986.
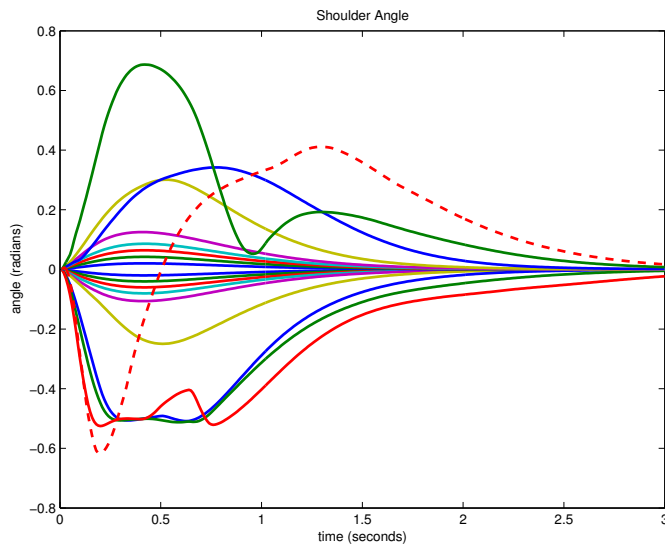
Fig. 14.   Shoulder responses to different size perturbations.

[4] C. Runge, C. Shupert, F. Horak, and F. Zajac, "Ankle and hip postural strategies defined by joint torques," *Gait and Posture*, vol. 10, no. 2, pp. 161–70, 1999.

[5] M. Spong, "The swingup control problem for the acrobot," *IEEE Control Systems Magazine*, vol. 15, no. 1, pp. 49–55, 1995.

[6] A. Hofmann, "Robust execution of bipedal walking tasks from biomechanical principles," PhD Thesis, MIT, EECS, 2006, people.csail.mit.edu/hofma.

[7] B. Stephens, "Integral control of humanoid balance," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007.

[8] M. Guidbard and P. Gorce, "Dynamic control of bipeds using ankle and hip strategies," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2002, pp. 2587–2592.

[9] S. Kudoh, T. Komura, and K. Ikeuchi, "The dynamic postural adjustment with the quadratic programming method," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2002, pp. 2563–2568.

[10] A. Kuo, "An optimal control model for analyzing human postural balance," *IEEE Transactions on Biomedical Engineering*, vol. 42, pp. 87–101, 1995.

[11] M. Abdallah and A. Goswami, "A biomechanically motivated two-phase strategy for biped upright balance control," in *IEEE Int. Conf. on Robotics and Automation*, 2004.

[12] C. G. Atkeson and B. Stephens, "Random sampling of states in dynamic programming," 2007, submitted to NIPS 2007.

[13] P. Dyer and S. McReynolds, *The Computational Theory of Optimal Control*.   Academic, NY, 1970.

[14] D. H. Jacobson and D. Q. Mayne, *Differential Dynamic Programming*. Elsevier, 1970.

[15] C. G. Atkeson, "Using local trajectory optimizers to speed up global optimization in dynamic programming," in *Advances in Neural Information Processing Systems*, J. D. Cowan, G. Tesauro, and J. Alspector, Eds., vol. 6.   Morgan Kaufmann Publishers, Inc., 1994, pp. 663–670.

[16] C. G. Atkeson and J. Morimoto, "Non-parametric representation of a policies and value functions: A trajectory based approach," in *Advances in Neural Information Processing Systems*, vol. 15, 2003.

[17] S. Anderson, C. Atkeson, and J. Hodgins, "Coordinating feet in bipedal balance," in *Proceedings of the 6th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2006)*, 2006.
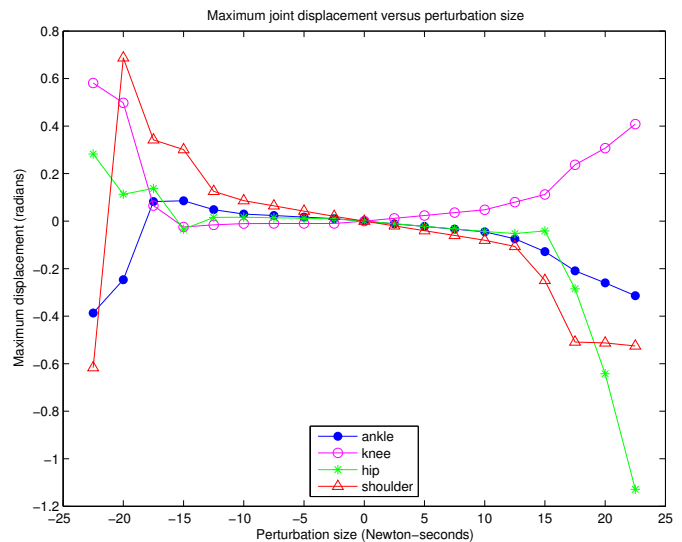
Fig. 15.   Maximum joint displacement versus perturbation size.
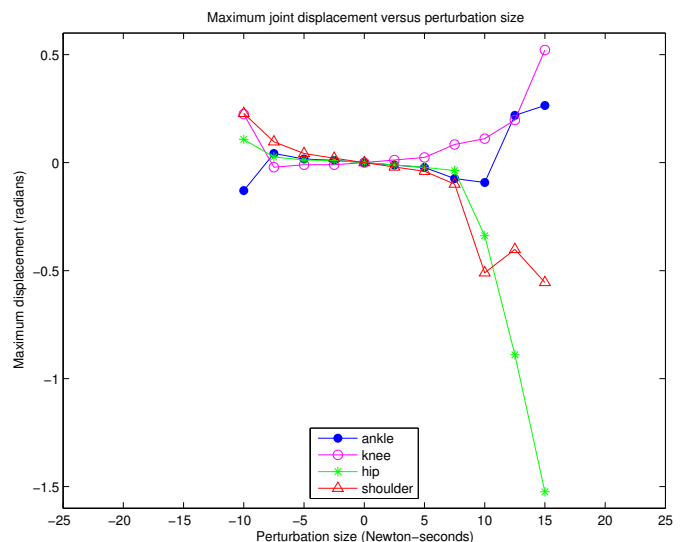

Fig. 16.   Maximum joint displacement versus perturbation size for the half sized foot case.
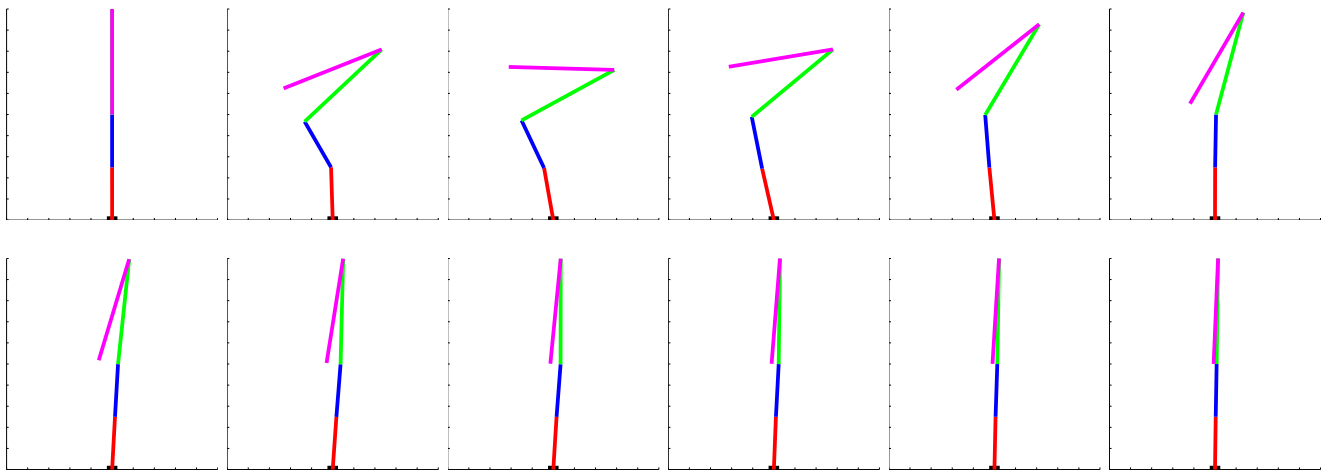
Fig. 17.   Configurations every quarter second from a simulated four link inverted pendulum response to a forward impulse of 15 Newton-seconds. The black rectangle indicates the extent of the half-sized foot.