

Bayesian Optimization of a Wearable Assistive Device Using an Estimator Stopping Process

A thesis presented

by

Charles Liu

to

Institute of Applied Computational Science

in partial fulfillment of the requirements

for the degree of

Master of Engineering

in the subject of

Computational Science & Engineering

Harvard University

Cambridge, Massachusetts

May 2018

© 2018 Charles Liu

All rights reserved.

Thesis Advisor:

Professor Scott Kuindersma

Author:

Charles Liu

Bayesian Optimization of a Wearable Assistive Device Using an Estimator Stopping Process

Abstract

Recent *human-in-the-loop* (HIL) optimization studies using wearable devices have shown an improved average metabolic reduction by optimizing control parameters online during short-duration experiments. However, the coupling of slow metabolic dynamics, high measurement noise, and hard experimental time constraints creates significant practical challenges for scaling optimization methods to more expressive control strategies. Prior work on applying gradient descent and Bayesian optimization methods to perform HIL optimization have decoupled the estimation and parameter selection problems, which leads to fixed estimation intervals and imposes a hard limit on the number of parameter evaluations possible in a given time budget. In this work, a different approach is taken that couples estimation and parameter selection, allowing the algorithm to spend less time refining the metabolic estimates for parameters that are unlikely to improve performance over the best values observed so far. We represent this early stopping mechanism with two different models that are incorporated in a standard Bayesian optimization scheme using an unscented Kalman filter for estimating metabolic rate. Performance is analyzed in several numerical examples and in pilot experiments with two human subjects optimizing 6 free control parameters of a hip soft exosuit.

Contents

Abstract	iii
Acknowledgments	x
Introduction	1
1 Bayesian Optimization	5
1.1 Gaussian Processes	5
1.2 GP Regression and Acquisition Functions	6
1.3 Hyperparameter Optimization and Student-t Noise Model	8
1.4 Implementation	9
2 Metabolic Cost Estimation	12
2.1 Instantaneous Energetic Cost	12
2.2 Parameter Estimation	13
2.3 UKF Covariance Parameters	14
2.4 Initial Trials and Future Work	16
3 Measurement Stopping Algorithms	19
3.1 Stopping Problem	19
3.2 σ -scaled Offset	20
3.3 Gittins Index	21
3.4 Adaptive Thresholds	23
3.5 Simulations	23
4 Subject Trials	27
4.1 Experimental Protocol	27
4.2 Experimental Subjects	29
4.3 Results	30
References	34
Appendix A Bayesian Optimization	38

Appendix B Unscented Kalman Filter	40
Appendix C Optimal Stopping Time Derivation	42
Appendix D Simulation Results	44

List of Tables

2.1	Tuning Parameters for metabolic estimator	15
3.1	Setup of finite horizon stopping problem	19
3.2	Optimization functions for simulations	25
4.1	6 parameters for providing a force profile	27
4.2	Subject Trial Summary	31

List of Figures

1	Flow diagram of optimization problem	3
1.1	Sample GP Regression of a function using squared exponential and Matérn kernels.	7
1.2	Gaussian MAP estimate vs. Student-t Gibbs Sampler. Data points added with Gaussian noise including some outliers with very high variance.	10
1.3	Demonstration of the effect of fixing measurement noise. Using the same data from Figure 1.2, the noise associated with each data point is fixed and inversely proportional to the true accuracy - in effect telling the GP that the minority outliers are the more accurate values. This may arise in our stopping problem where a minimum is found (measured for the full duration, resulting variance is relatively low) and many nearby points are subsequently sampled and stopped early (resulting variances are very high).	11
2.1	Sample estimation process for instantaneous energetic cost.	13
2.2	With a true instantaneous energetic cost at 350 and prior at 300, keeping all other parameters constant, the effect of using various P_w values on measurement variance.	16
2.3	Tuned UKF Estimator over previously collected data. Accuracy assessed over partial data.	17
2.4	Tuned UKF Estimator shows different levels of volatility during optimization studies on different subjects.	18
4.1	Experimental setup. Assistance parameters for the soft exosuit were optimized on-line. Respiratory device measured respiratory rate to estimate metabolic cost for a given condition with a covariance. In the optimization computer, the Gittins process determined when to stop sampling and Bayesian optimization was used to select a next parameter set. The parameters were sent to the real-time computer (speedgoat) to generate a force trajectory. Then, the actuator provided a force to a subject through a bowden cables.	30
4.2	Diagram of Optimization Process	31

4.3 Force profiles of the two subjects	31
A.1 Four Iterations of Bayesian Optimization	39
D.1 Branin Function	45
D.2 Ackley Function	46
D.3 Levy Function	47
D.4 Hartmann-6 Function	48

List of Algorithms

1	Bayesian Optimization Outline	8
2	Determining threshold levels for either model based on subject's exploration data. With the σ -offset, the higher threshold is more risk averse, while with the Gittins model the lower is.	24

Acknowledgments

I am very grateful for having Scott Kuindersma as my advisor. It has been a privilege to be his student and I appreciate his guidance, patience, and immense knowledge in helping me throughout the course of my study. It has also been a pleasure to work with Myunghee Kim, who was integral to the work in this thesis. Her wealth of experience in the field as well as overseeing the various levels of development was instrumental. I'd also like to thank Professor Conor Walsh, his lab and the Wyss staff, in particular Nikos Karavas, Sangjun Lee, Jinsoo Kim, Chih-Kang Chang, Asa Eckert-Erdheim, Maria Athanassiu, Brice Mikala Iwangou, Nicolas Menard, and Sarah Sullivan. This work would not have been possible without the hard work of everyone involved from the design and fabrication of the exosuit to the development of the controllers. Finally, I thank my parents, my friends, and the rest of my family for their amazing support.

Introduction

Wearable robotic devices are intended as a means to augment human economy, strength, and endurance. Over the past decade, a number of devices have been developed for reducing the metabolic cost of walking for able-bodied individuals (Malcolm *et al.*, 2013; Mooney and Herr., 2016; Caputo and Collins, 2014; Panizzolo *et al.*, 2016a). With tethered and portable hardware platforms having advanced considerably (Malcolm *et al.*, 2013; Lee *et al.*; Panizzolo *et al.*, 2016b), it is now possible to accurately control settings such as the timing and magnitude of the delivered torque (Malcolm *et al.*, 2013; Collins *et al.*, 2015; Ding *et al.*, 2017; Kim and Collins, 2015, 2017; Lee *et al.*, 2016). Increased attention has been directed towards how control strategy and control parameters influence overall system performance (Caputo and Collins, 2014; Kim and Collins, 2015, 2017; Quinlivan *et al.*, 2017). Traditionally, control parameters have been tuned manually by researchers with expert knowledge on both the device and biomechanics of human walking or through exploring the parameter space in a systematic sweep (Caputo *et al.*, 2015; Quinlivan *et al.*, 2017). However, despite these efforts, several studies have shown significant inter-subject variability in the observed metabolic benefit for a fixed parameter setting (Quesada *et al.*, 2016). Being able to automatically recover the optimal parameter setting on an individual basis is therefore a fundamental component in designing effective wearable robotic devices.

Recently, several groups have explored different formulations of *human-in-the-loop* (HIL) optimization that offer the possibility to avoid conventional exhaustive search (Koller *et al.*, 2016; Zhang *et al.*, 2017; Ding *et al.*, 2018; Kim *et al.*, 2017). These studies use instantaneous energetic cost (Selinger and Donelan., 2014), an estimation of steady-state metabolic cost, as

an objective measurement for automatically optimizing parameter settings (Koller *et al.*, 2016; Felt *et al.*, 2015; Zhang *et al.*, 2017; Ding *et al.*, 2018). Previous research has applied this HIL optimization approach to find optimal onset actuation timing of a bilateral pneumatic ankle exoskeleton by automatically adjusting a single control parameter (Koller *et al.*, 2016). A recent study with an ankle exoskeleton (Zhang *et al.*, 2017) and hip exosuit (Ding *et al.*, 2018) showed the potential to achieve larger metabolic benefits by simultaneously optimizing multiple parameters using *Bayesian optimization*.

Motivation

Bayesian optimization is a sequential optimization strategy that has been used in a variety of applications from hyperparameter tuning for machine learning algorithms (Snoek *et al.*, 2012; Mahendran *et al.*, 2012) to portfolio allocation (Brochu *et al.*, 2010) and experimental design (Brochu and De Freitas, 2010). As it is a general framework to optimize noisy black box functions, many variations have been introduced to account for different use cases. In particular, an area of interest has been in tackling the noisiness of data samples (Rue *et al.*, 2009; Nogueira *et al.*, 2016; Assael *et al.*, 2014) and the myopia of the algorithm (Lam *et al.*, 2016; González *et al.*, 2015).

The algorithm is typically framed in the context of an iteration budget, abstracting away the actual data sampling process. However, in the HIL problem there is a limited time constraint rather than an iteration constraint, and data acquisition exhibits a direct tradeoff between noisiness and elapsed time. In previous studies, a fixed observation window was used for every parameter evaluation, effectively reducing the time constraint to an iteration constraint. Unfortunately, this means spending valuable time measuring parameter settings that are unlikely to improve upon the best value observed so far.

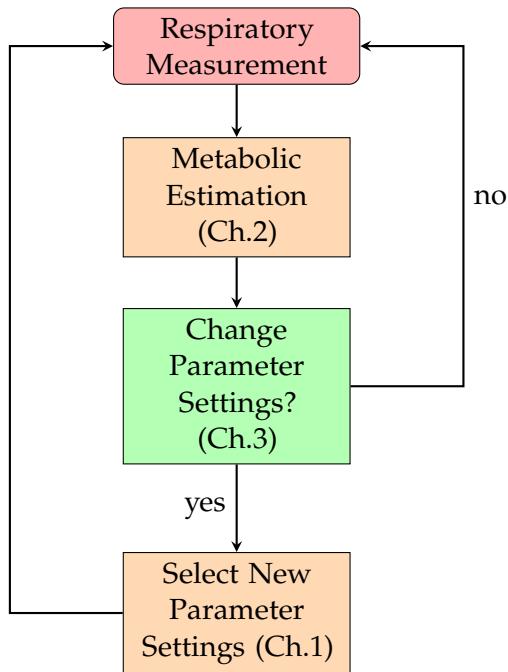


Figure 1: Flow diagram of optimization problem

Summary of contributions

The main contribution of this work is to introduce a stopping problem within a given evaluation of Bayesian optimization. Rather than having a fixed observation window, a stopping rule is determined after every respiratory measurement to maximize the time spent evaluating promising parameter settings. To determine the optimal stopping time, a metabolic cost estimator¹ is also developed to provide a probability distribution of the instantaneous energetic cost after every measurement. Depending on the paramterization of the estimator, two formulations of the stopping problem are introduced and evaluated under different circumstances.

¹Primary development by Myunghee Kim

Outline

Chapter 1 gives background information on Bayesian optimization and Gaussian processes. Chapter 2 presents the metabolic estimator and its robustness under various levels of noise. Initial experimental data is provided and future work on the estimator is discussed. Chapter 3 introduces two algorithms for determining the optimal stopping point within an iteration of Bayesian Optimization. Simulations of various optimization functions are provided using the two algorithms under different scenarios. Chapter 4 provides results from an experimental protocol using one of the algorithms. Finally, we conclude this thesis with a discussion of future work.

Chapter 1

Bayesian Optimization

1.1 Gaussian Processes

Given a parameter space $\mathbb{X} \subset \mathbb{R}^m$ and an unknown, noisy cost function f that is expensive to evaluate, we aim to find the solution to the minimization problem

$$x^* = \arg \min_{x \in \mathbb{X}} f(x), \quad (1.1)$$

by iteratively choosing points to evaluate until a stopping criteria is met. Bayesian optimization, a sequential design strategy, computes a posterior distribution on possible objective functions given all previous evaluations and optimizes an acquisition function on this distribution to globally select the next value of x to evaluate, typically in a way that carefully balances exploration and exploitation. The distribution over f is modeled as a Gaussian process (Rasmussen, 2006) \mathbb{G} ,

$$f \sim \mathbb{G}(\mu, \kappa), \quad (1.2)$$

where $\mu : \mathbb{X} \rightarrow \mathbb{R}$ is a mean function typically set to zero, and $\kappa : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}$ is a covariance kernel that characterizes the correlation between different points in the domain. There are a number of options for covariance kernels, both stationary and non-stationary, but the most

widely known is the squared exponential

$$\kappa(x_i, x_j | \theta) = \sigma_\theta^2 \exp\left(-\frac{1}{2} d^2\left(\frac{x_i}{l}, \frac{x_j}{l}\right)\right), \quad (1.3)$$

with parameter $\theta = [\sigma_\theta^2, l]$ where σ_θ is a scalar, l is either a scalar or vector of dimension m , and $d(x_i, x_j)$ is the Euclidean distance. The Matérn covariance kernel is a more generalized form of the squared exponential, defined as

$$\kappa(x_i, x_j | \theta) = \sigma_\theta^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} d\left(\frac{x_i}{l}, \frac{x_j}{l}\right)\right)^\nu K_\nu\left(\sqrt{2\nu} d\left(\frac{x_i}{l}, \frac{x_j}{l}\right)\right) \quad (1.4)$$

with an additional scalar ν and K_ν is a modified Bessel function.

1.2 GP Regression and Acquisition Functions

By definition of a GP, any collection of points forms a multivariate normal distribution defined by μ and κ . Under this assumption, a posterior distribution given a set of training samples can be solved analytically. Formally, given some training samples $S = \{(x_i, y_i)\}_{i=1}^n$, and assuming each sample follows a Gaussian noise model $y_i \sim f(x_i) + \mathcal{N}(0, \sigma_n^2)$, the posterior distribution at x is Gaussian with mean $\bar{\mu}(x)$ and variance $\bar{\sigma}^2(x)$ evaluated as

$$\bar{\mu}(x) = K(X, x)^T [K(X, X) + \sigma_n^2 I]^{-1} Y \quad (1.5)$$

$$\bar{\sigma}^2(x) = \kappa(x, x | \theta) - K(X, x)^T [K(X, X) + \sigma_n^2 I]^{-1} K(X, x) \quad (1.6)$$

$$K(X, x)_i = \kappa(x_i, x | \theta)$$

$$K(X, X)_{ij} = \kappa(x_i, x_j | \theta),$$

where $X = \{x_i\}_{i=1}^n$ and $Y = \{y_i\}_{i=1}^n$.

Given a Gaussian distribution at point x , an acquisition function captures how attractive that point is to sample next. A common choice of acquisition function is Expected Improvement (EI), which returns the expected reduction in cost over the best parameters observed

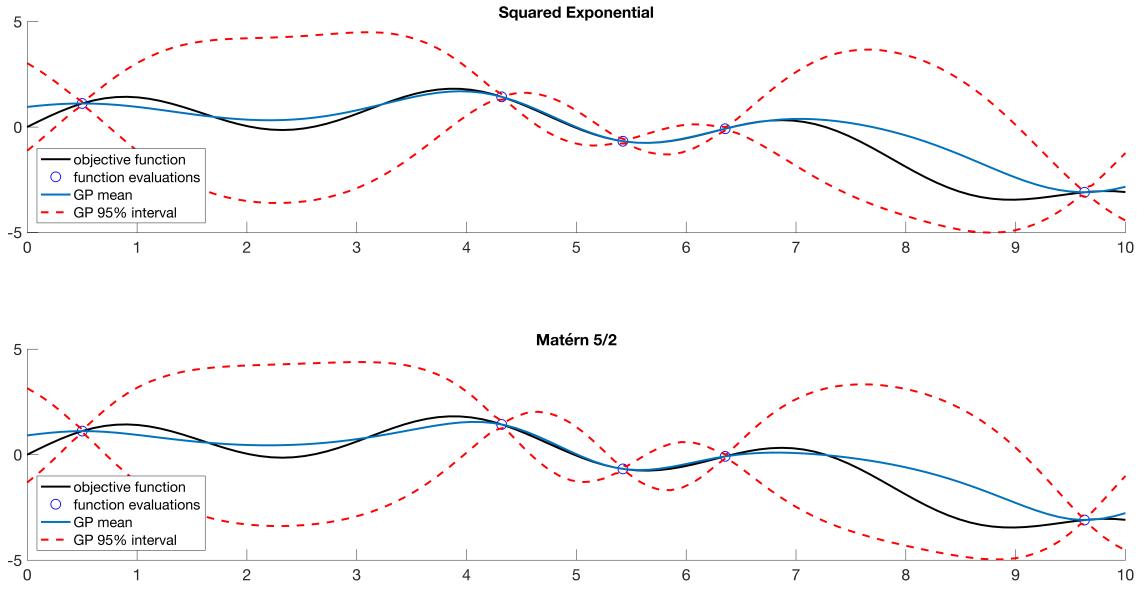


Figure 1.1: Sample GP Regression of a function using squared exponential and Matérn kernels.

so far. For the given GP model, EI can be computed in closed form:

$$\begin{aligned}
 EI(x|\mathcal{S}) &= \int_{\infty}^{\infty} \max(0, y^* - y) p(y|x) dy \\
 &= z\bar{\sigma}(x)\Phi(z) + \bar{\sigma}(x)\phi(z) \\
 z &= \frac{y^* - \bar{\mu}(x) + \xi}{\bar{\sigma}(x)},
 \end{aligned} \tag{1.7}$$

where y^* is the best value observed so far, $\Phi(z)$ and $\phi(z)$ are the standard normal CDF and PDF functions, and ξ is a scaling parameter to adjust the tradeoff between exploration-exploitation (Lizotte, 2008). The next point to evaluate is then chosen as $x^* = \arg \max_x EI(x|\mathcal{S})$. Figure A.1 shows a few iterations of Bayesian Optimization on a sample function.

The choice of y^* can be defined as the minimum observed value so far, the minimum $\bar{\mu}(x)$ under the current GP model over all previous observations, or some biased version thereof (Lizotte, 2008).

```

Objective Function  $F(x)$ 
Acquisition Function  $g(\mu, \sigma^2)$ 
Specify Exploration Points  $\mathbb{E} = \{e_1, e_2, \dots, e_n\}$ 
Training Samples  $S$ 
for  $i = 1$  to  $n$ 
     $S = S \cup \{e_i, F(e_i)\}$ 
end
while  $t < T$ 
    Update GP Hyperparameters  $\theta$ 
    Given  $f(x|\theta, S) \sim \mathcal{N}(\mu_x, \sigma_x^2)$ ,
     $x^* = \arg \max_{x \in \mathbb{X}} g(\mu_x, \sigma_x^2)$ 
     $S = S \cup \{x^*, F(x^*)\}$ 
end

```

Algorithm 1: Bayesian Optimization Outline

1.3 Hyperparameter Optimization and Student-t Noise Model

The values of the hyperparameters in the kernel function are critical to determining the posterior distribution. Typically, these hyperparameters are tuned to sampled points with a maximum a posteriori (MAP) point estimate (Snoek *et al.*, 2012; Jarno Vanhatalo and Vehtari, 2013). Specifically, assuming our Gaussian noise model the MAP estimate is defined as

$$\begin{aligned}
& \arg \max_{\theta, \sigma_n^2} \log P(S|\theta, \sigma_n^2) + \log P(\theta, \sigma_n^2) \\
&= \log \int p(Y|f, \sigma_n^2) p(f|X, \theta) df + \log P(\theta, \sigma_n^2) \\
&= -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log(K(X, X) + \sigma_n^2 I) - \frac{1}{2} Y^T (K(X, X) + \sigma_n^2 I)^{-1} Y + \log P(\theta, \sigma_n^2)
\end{aligned} \tag{1.8}$$

Instead of using a single point estimate, a full Bayesian approach would involve marginalizing out the hyperparameters through an MCMC approximation. This would be particularly beneficial in the case where the posterior is multi-modal or sensitive to small changes in hyperparameter values. Many sampling approaches have been proposed, and we refer the reader to (Barber *et al.*, 2011, Chapter 14) for a summary.

The MCMC approach is also used in more complex noise models whose posterior distribution and marginal likelihood are no longer analytically tractable. Vanhatalo *et al.* (2009); Jyläniemi *et al.* (2011) developed a more robust noise model based on the Student-

t distribution to handle occurrences of large outliers in the collected data. The noise distribution follows

$$p(y|f, \sigma_n^2, \nu) = \left(\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2}) \sqrt{\nu \pi} \sigma_n} \right) \left(1 + \frac{(y-f)^2}{\nu \sigma_n^2} \right)^{-\frac{(\nu+1)}{2}} \quad (1.9)$$

where ν is a degrees of freedom hyperparameter. A Gibbs sampler can be performed through a hierarchical model where each training sample's output noise follows an independent inverse chi-squared distribution

$$\begin{aligned} y_i|f &\sim \mathcal{N}(f, V) \\ V &\sim \text{Inv-}\chi^2(\nu, \sigma_i^2), \end{aligned} \quad (1.10)$$

with θ alternately sampled according to the Gaussian Process. Figure 1.2 shows a comparison of the robust Student-t noise with the standard Gaussian noise. The majority of data points exhibit low variance, while some extreme outliers are also produced. While the Gaussian noise model treats each data point equally, the Student-t model is able to effectively pick out the outliers and recreate the underlying function.

1.4 Implementation

There are a number of implementation choices relevant to HIL optimization. While the MCMC approach has shown to be more robust, GP regression is already an $O(n^3)$ operation where n represents the number of function evaluations. With MCMC, the GP regression becomes $O(mn^3)$ for m samples. Optimizing the acquisition function using a sufficient sample size is then increasingly expensive, taking on the order of minutes to evaluate. However, once a stopping condition is met the next parameter setting must immediately be communicated to the wearable device. In our implementation, we start an MCMC sampler concurrently with a new parameter evaluation. When the stopping condition is met, if the sampler has completed optimizing EI we use the returned parameter setting. Otherwise, we default to the MAP estimate to find the next parameter selection. There is then a tradeoff between the standard MAP estimate with a more robust MCMC sampler that uses one less

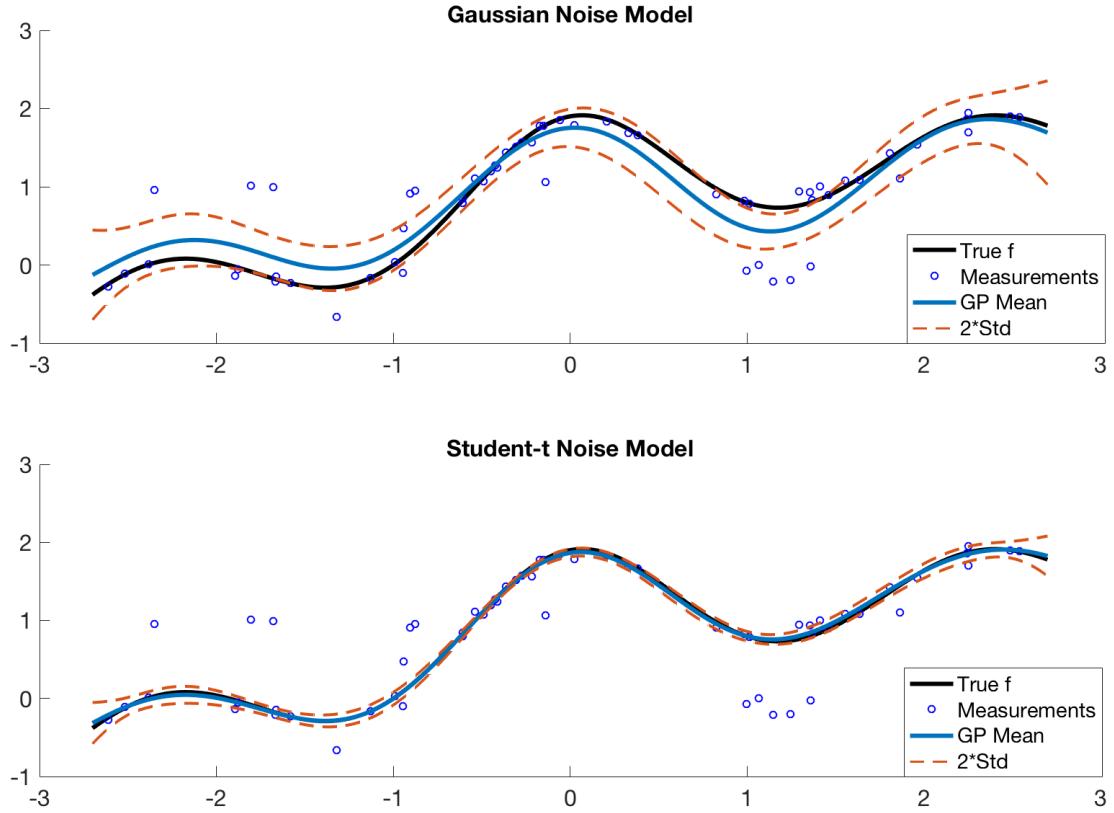


Figure 1.2: Gaussian MAP estimate vs. Student-t Gibbs Sampler. Data points added with Gaussian noise including some outliers with very high variance.

training sample.

As Chapter 2 will discuss, the returned metabolic estimates have an associated variance; earlier stopping points exhibit a higher variance representing greater uncertainty for the estimated cost. The GP estimates may incorrectly inflate the best observed value if a number of parameter settings nearby are stopped early (see MAP estimate in Figure 1.2 for $x \in (-2, -1)$). To appropriately weight the training samples based on their estimator variances, we fix separate noise hyperparameters corresponding to each sample. In the Gaussian noise model, the regression equation then becomes

$$\begin{aligned}\bar{\mu}(x) &= K(X, x)^T [K(X, X) + \Sigma]^{-1} Y \\ \bar{\sigma}^2(x) &= \kappa(x, x | \theta) - K(X, x)^T [K(X, X) + \Sigma]^{-1} K(X, x),\end{aligned}\tag{1.11}$$

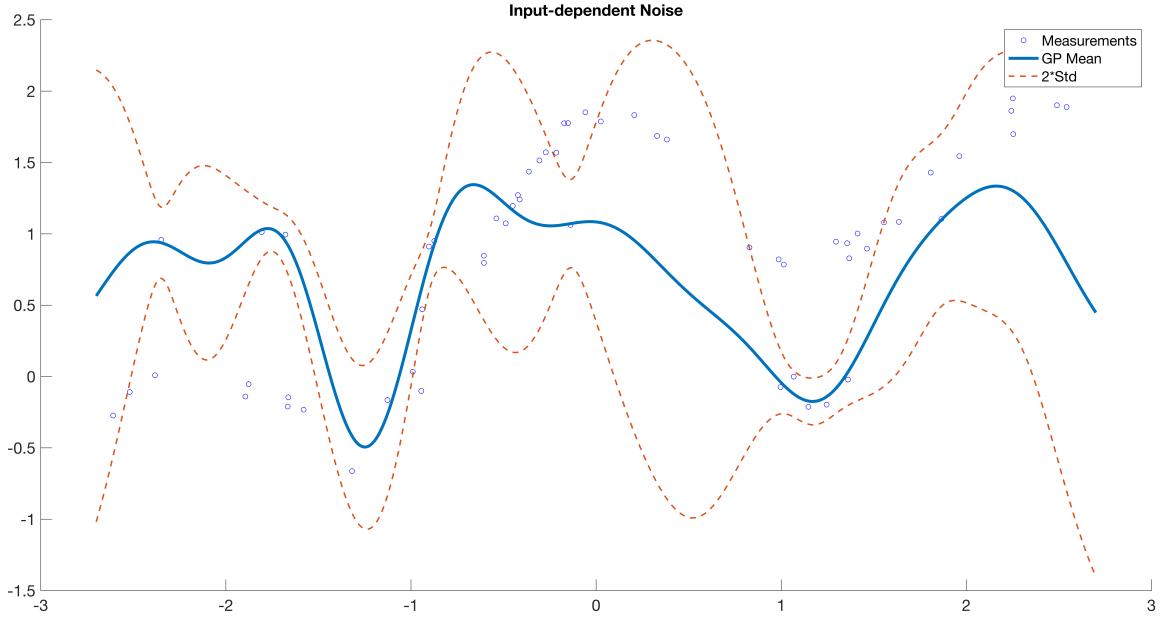


Figure 1.3: Demonstration of the effect of fixing measurement noise. Using the same data from Figure 1.2, the noise associated with each data point is fixed and inversely proportional to the true accuracy - in effect telling the GP that the minority outliers are the more accurate values. This may arise in our stopping problem where a minimum is found (measured for the full duration, resulting variance is relatively low) and many nearby points are subsequently sampled and stopped early (resulting variances are very high).

where Σ is a diagonal matrix of entries $\{\sigma_i^2\}$ corresponding to scaled variance estimates of each data point (Kersting *et al.*, 2007). In the case of the MCMC model, each y_i is defined independently with a noise parameter already. Figure 1.3 gives an example of the effect of fixed noise hyperparameters on the GP. Additionally, we define y^* in Expected Improvement as the best mean estimate rather than the observed value. The best observed value is equally likely to be an outlier among a set of nearby points that were all measured for the full duration, so the Gaussian process estimates are used instead.

Chapter 2

Metabolic Cost Estimation

2.1 Instantaneous Energetic Cost

The instantaneous energetic cost, our objective function, is estimated through continuous breath measurements over a lengthy duration of time. Brockway (1987) developed a formula for converting CO_2 and O_2 measurements into a metabolic cost, which is then fit into a first-order dynamical model. Given the noisiness of respiratory measurements, previous studies (Felt *et al.*, 2015; Selinger and Donelan., 2014) allowed three minutes while subjects reached a steady state before measuring for an additional three minutes to estimate an instantaneous energetic cost. In a previous Bayesian optimization study (Ding *et al.*, 2018), each iteration used a fixed two minutes of respiratory measurements.

Rather than requiring a fixed time interval for each evaluation, we aim to ascertain an accurate estimate only for promising parameter settings while stopping early for those that are unlikely to improve upon the best settings found so far. To make that determination, an online estimator for instantaneous energetic cost was developed using the metabolic cost model

$$m_t(c_0, c, \tau_0, \tau) = c(1 - e^{\frac{-t}{\tau}}) + c_0 e^{\frac{-t}{\tau_0}}, \quad (2.1)$$

where t is the total measurement time in seconds and τ_0, τ are time constants characterizing

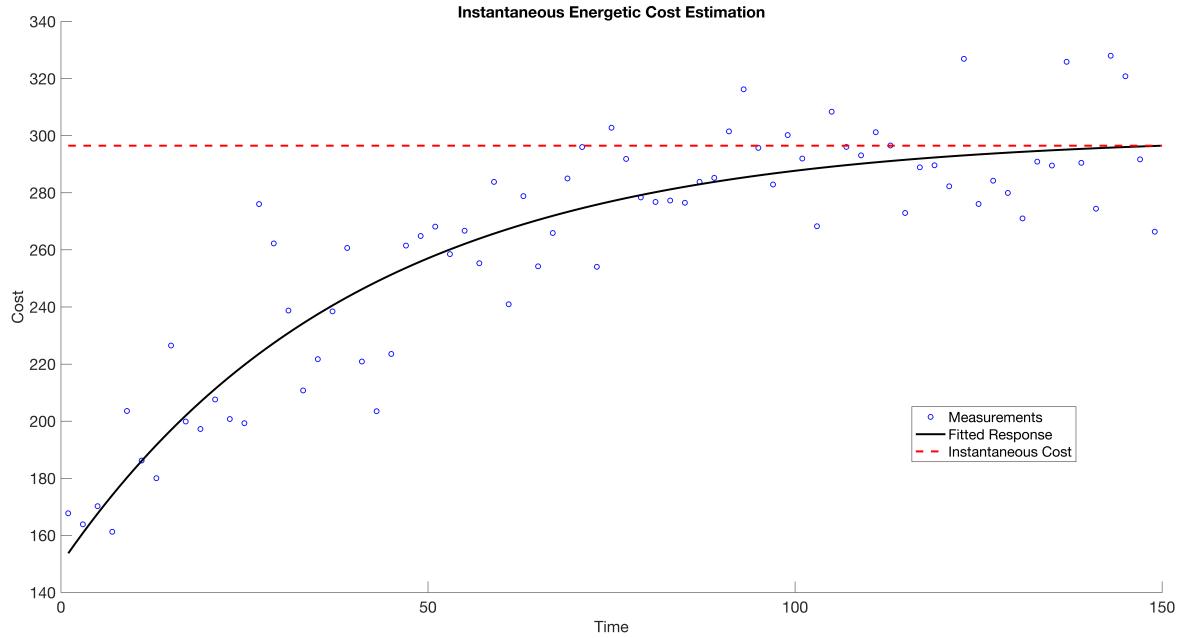


Figure 2.1: Sample estimation process for instantaneous energetic cost.

the rate of change for the initial cost c_0 and the instantaneous energetic cost c respectively.

2.2 Parameter Estimation

The time and cost parameters are estimated through an Unscented Kalman Filter (Julier, 1997). Kalman filters are applied to discrete time systems of the form

$$\begin{aligned} x(t+1) &= F(x(t), v(t), t) \\ z(t) &= H(x(t), w(t), t), \end{aligned} \tag{2.2}$$

where $x(t)$ represents the unobserved state of the system, $z(t)$ the observed measurement, and zero-mean Gaussian vectors $v(t)$ for process disturbances or modeling errors and $w(t)$ for measurement noise.

First, a prediction is made for the state of the system after one timestep. Let $\hat{x}(t|t-1)$ represent the estimate for $x(t)$ using measurements $Z(t-1) = [z(1), z(2), \dots, z(t-1)]$ and

$P_x(t|t-1)$ the covariance for the estimate. The prediction step aims to calculate

$$\begin{aligned}\hat{x}(t|t-1) &= \mathbb{E}[F(x(t-1), v(t-1), t-1)|\mathbb{Z}(t-1)] \\ P_x(t|t-1) &= \mathbb{E}[\{\hat{x}(t|t-1) - x(t)\}\{\hat{x}(t|t-1) - x(t)\}^T|\mathbb{Z}(t-1)].\end{aligned}\tag{2.3}$$

The prediction is then reconciled with the observed measurement through a minimum mean-squared estimate $\hat{x}(t)$ and covariance $P_x(t)$ according to the following equations

$$\begin{aligned}\hat{x}(t) &= \hat{x}(t|t-1) + Ky \\ P_x(t) &= P_x(t|t-1) - KP_y(t|t-1)K^T \\ K &= P_{xy}(t|t-1)P_y^{-1}(t|t-1) \\ y &= z(t) - H(\hat{x}(t|t-1), w(t), t).\end{aligned}\tag{2.4}$$

The underlying requirement for these equations is an accurate propagation of the first two moments of $x(t)$ and $z(t)$. As such, Gaussian estimates are used as they assume the least information. When F or H is nonlinear, rather than linearizing the functions with a first order approximation the UKF approximates its probability distribution through an unscented transformation. The complete implementation details are outlined in Appendix B.

As we consider τ_0, τ, c, c_0 to be constant parameters, our UKF setup is

$$\begin{aligned}x &= [c_0 \ c \ \tau_0 \ \tau] \\ F(x(t), v(t), t) &= x(t) + v(t) \\ H(x(t), w(t), t) &= m_t(x(t)) + w(t).\end{aligned}\tag{2.5}$$

2.3 UKF Covariance Parameters

The UKF algorithm has covariance parameters P_x for the state estimate, P_w for measurement noise, and P_v for process noise that determine the sensitivity of the state estimate to new measurements. Table 2.1 summarizes the effect of increasing one of these covariance parameters while holding all others constant.

The volatility of the estimator is dependent on the calibration of the covariance parame-

Increase	Effect On State Estimate
P_x	A higher initial covariance on the state estimate implies higher uncertainty with the state prior. Early observations will have a greater effect on the state update. However, P_x decreases after every state update; observations therefore will have a smaller effect on the state estimate over time
P_w	A higher measurement noise will reduce an observation's effect on the state update. The state will be slow to move away from the prior, particularly in early observations. P_w also affects P_x on future timesteps, as it is bounded by P_w
P_v	While F is essentially the identity function in the metabolics model as the cost and time parameters are assumed to be constant, P_v encapsulates the error in that model as well as outside disturbances that can't be measured such as fatigue. Increasing P_v can be used to enforce a lower bound on the effect of an observation to the state update

Table 2.1: Tuning Parameters for metabolic estimator

ters with respect to the noise of the measurements. Figure 2.2 shows three estimators: one that underestimates the measurement noise by a factor of 10, one that is exact, and one that overestimates by a factor of 10. Underestimating leads to a higher weight given to the measurement on subsequent state updates, with the volatility initially quite high. As more measurements are taken and P_x decreases, the estimator eventually reaches a steady state. On the other hand, assuming too high a measurement noise will cause the estimator to slowly update towards the true value.

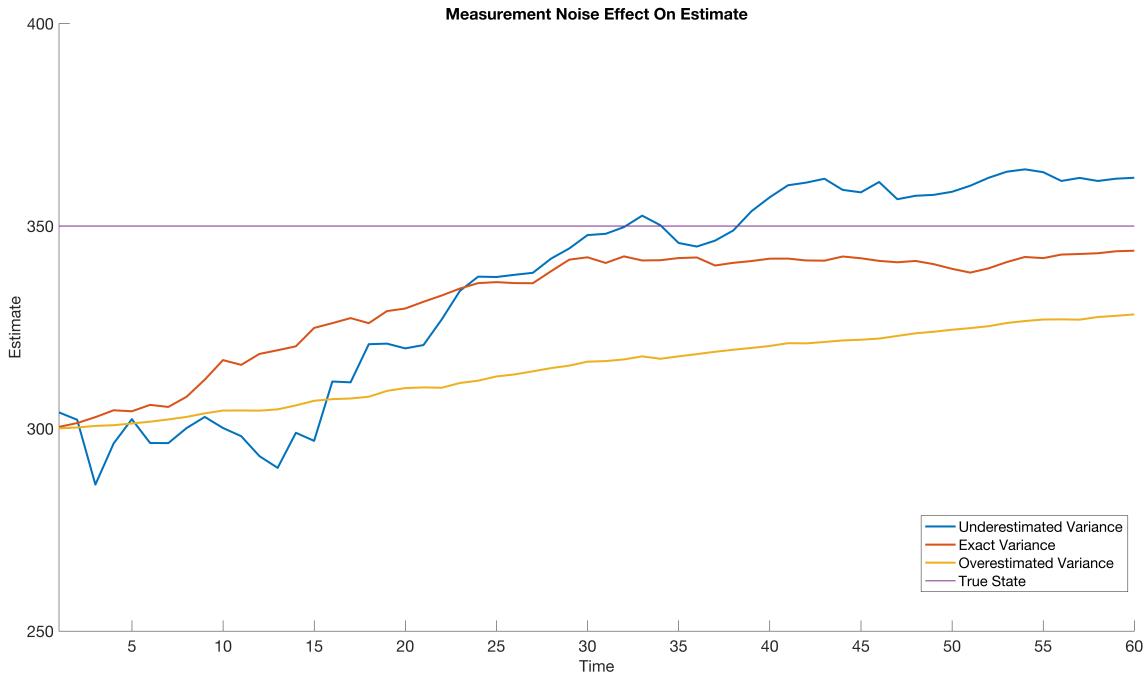


Figure 2.2: With a true instantaneous energetic cost at 350 and prior at 300, keeping all other parameters constant, the effect of using various P_w values on measurement variance.

2.4 Initial Trials and Future Work

Using previously attained metabolic data, Figure 2.3 shows the raw and noisy measurements taken and the metabolic cost prediction that the model developed in response to a representative participant¹. The bottom plot displays the covariance associated with the cost, beginning with a covariance of 1 and converging at different rates depending upon the raw data. When tuned properly, both trials see a relatively smooth rate of convergence of the estimate. With varying amounts of data provided, the accuracy in percent error was calculated by comparing with a "ground truth" value at each termination condition (5 min, 2 min, 1.5 min, 20 breaths, and 30 breaths). This ground truth value is the average of the last two minutes of data, when the subject has reached steady-state.

During trial testing for our optimization methods, the same noise covariances used on the UKF estimator showed various levels of volatility across subjects (Figure 2.4). This

¹Testing carried out by Myunghee Kim

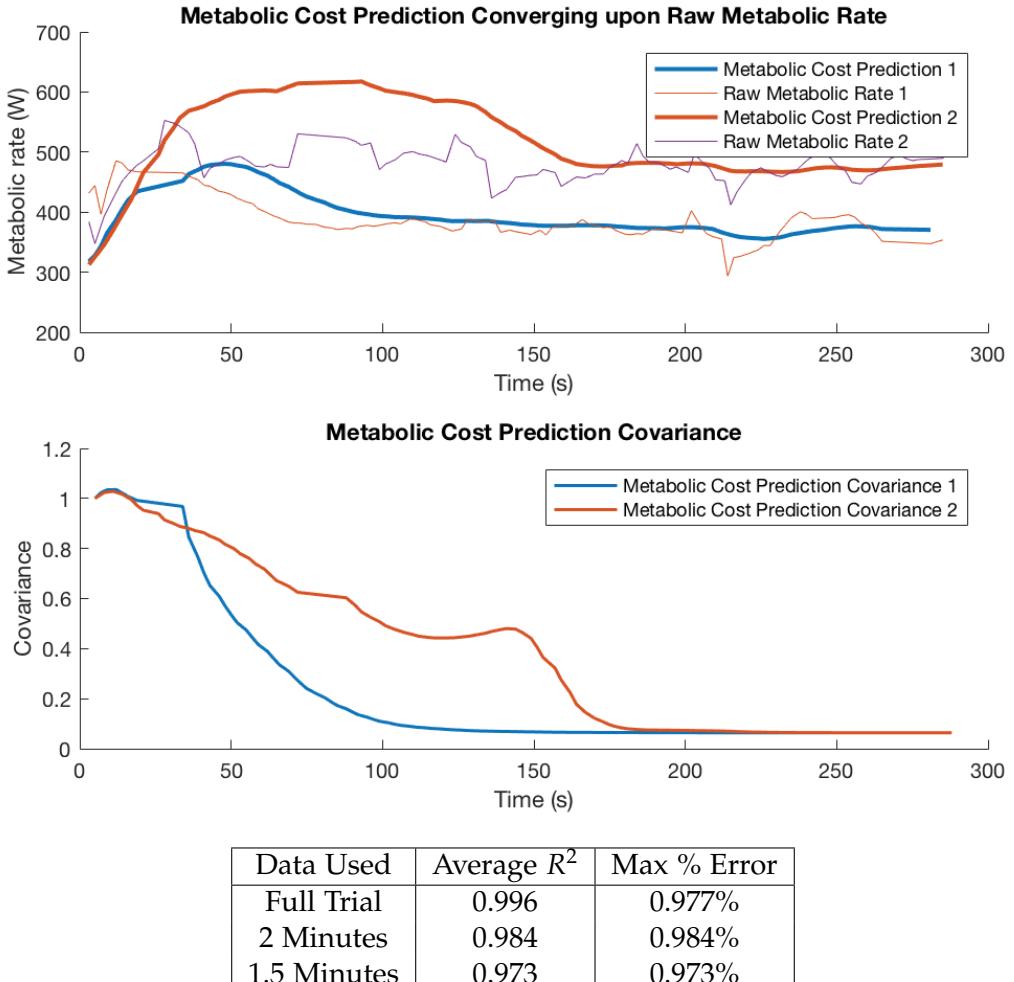


Figure 2.3: Tuned UKF Estimator over previously collected data. Accuracy assessed over partial data.

suggests the estimator should also have some subject-to-subject individualization based on the volatility of respiratory measurements. Future work would involve creating a cost function to optimize for this variation, balancing between the smoothness of the estimation curve and eventual convergence to the "true" instantaneous energetic cost. As Bayesian optimization starts with an exploration phase, the estimator can then be tuned a posteriori. Depending on the speed of the optimization, the estimator could potentially be tuned after every evaluation that takes the full duration, making it adaptable over time to external

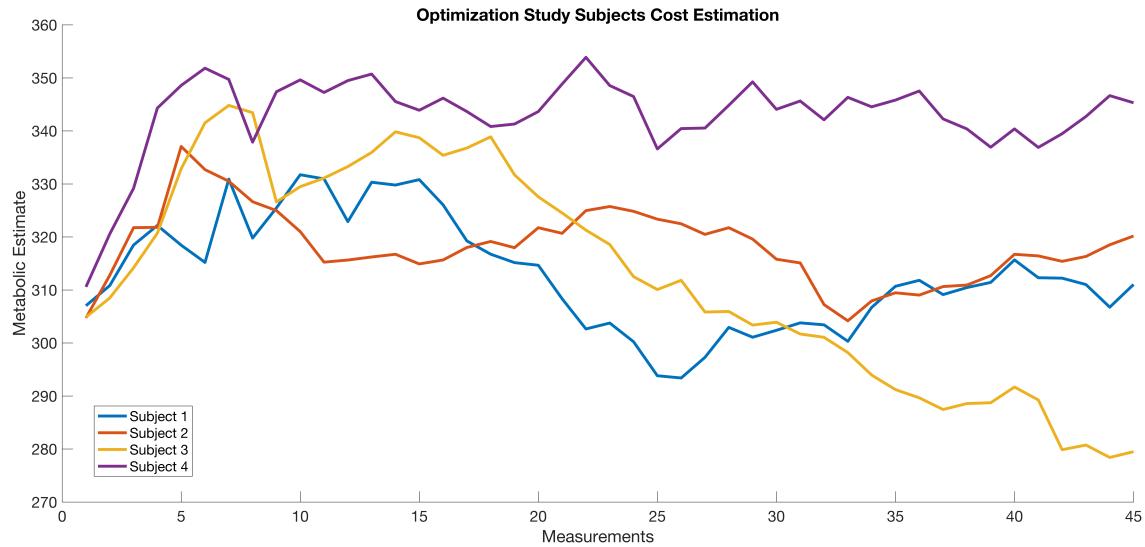


Figure 2.4: Tuned UKF Estimator shows different levels of volatility during optimization studies on different subjects.

factors such as fatigue. Improvements in tuning the estimator would have residual effects in the information provided to both the Gaussian process and the stopping problem. Without it, Chapter 3 will introduce the stopping problem and different models that take into account the accuracy of the estimator.

Chapter 3

Measurement Stopping Algorithms

3.1 Stopping Problem

With an estimate and associated variance for the instantaneous energetic cost, the determination for when to stop measuring is known as a finite-horizon stopping problem. Formally, the stopping problem is defined with

N	Finite horizon
X_t	State at time t
$P(X_t X_{t-1})$	State transitions, typically Markovian
λ	Discount Factor $\in (0, 1]$
$r(X)$	Bounded reward function for continuing at state X
$g(X)$	Bounded reward function for stopping at state X

Table 3.1: Setup of finite horizon stopping problem

The objective is to find τ that maximizes

$$V(x) = \max_{\tau: \tau \leq N} \mathbb{E}[\lambda^\tau g(X_\tau) + \sum_{i=0}^{\tau-1} \lambda^i r(X_i) | X_0 = x]. \quad (3.1)$$

The value function can be computed via backward induction by defining

$$\begin{aligned} J_N(x) &= g(x) \\ J_n(x) &= \max\{g(x), r(x) + \lambda \mathbb{E}_{P(y|x)}[J_{n+1}(y)]\}, \end{aligned} \tag{3.2}$$

then $V(x) = J_0(x)$ and the optimal stopping time is

$$\tau = \min_{t \geq 0} J_t(X_t) = g(X_t). \tag{3.3}$$

For a derivation of the optimal stopping time, see Appendix C. We aim to formulate the stopping problem in terms of the metabolic estimator. Let N be the maximum measurements allowed at a parameter evaluation, $\hat{x}_t \sim \mathcal{N}(\mu_{x_t}, \sigma_{x_t}^2)$ the cost estimate and associated variance for the current evaluation, and $\hat{x}^* \sim \mathcal{N}(\mu_{x^*}, \sigma_{x^*}^2)$ the cost estimate and variance for the best parameter setting found thus far as defined by the Gaussian process. As Bayesian optimization starts with an exploration phase, \hat{x}^* will always be defined (Algorithm 1).

3.2 σ -scaled Offset

The σ -scaled offset model directly embeds \hat{x}_t and \hat{x}^* into the stopping problem. The best estimate of the difference in energetic costs is given by the following Gaussian distribution

$$\begin{aligned} \hat{x}_t - \hat{x}^* &\sim \mathcal{N}(\mu, \sigma^2) \\ \mu &= \mu_{x_t} - \mu_{x^*} \\ \sigma^2 &= \sigma_{x_t}^2 + \sigma_{x^*}^2. \end{aligned} \tag{3.4}$$

The stopping problem state is then the estimate difference, with a non-Markovian transition model of independent draws from the same Gaussian distribution. The lower x_t seems relative to x^* , the higher the reward function incentivizes continuing. In summary,

$$\begin{aligned}
X_t &= \mu \\
P(X_t|X_{t-1}) &= P(X_t) \sim \mathcal{N}(\mu, \sigma^2) \\
r(x) &= K\sigma - x \\
g(x) &= 0,
\end{aligned} \tag{3.5}$$

where K is a σ -scaled offset to provide some risk averseness to the estimate. The optimal stopping point for this formulation can be characterized as the first time $X_t > K\sigma$. As more measurements are taken and P_x decreases, σ^2 decreases the offset accordingly.

As shown in Section 2.3, underestimating P_w can cause the estimator to exhibit high volatility. Additionally, since a subsequent state covariance update $P_x(t+1)$ is bounded by P_w , a lower P_w will also cause σ^2 to be underestimated. With the σ -scaled offset model, underestimating the noisiness of observations could potentially cause promising parameter settings to stop measuring early as there is no fault tolerance for crossing the threshold.

3.3 Gittins Index

A more fault-tolerant model is drawn from literature related to the multi-armed bandit problem (MAB), which represents a sequential decision making problem where at any given time a player must select from different options (or arms) that then responds with a variable reward. Historically this has been used in a clinical context, where a physician must select from one of several different treatment options when presented a patient and receives binary rewards of either successes or failures (Sofia S. Villar and Wason, 2015). A stopping problem has also been described as a one-armed bandit problem (Ferguson, 2006).

Successes and failures can be based on μ, σ^2 using the Probability of Improvement (PI) metric

$$PI(\mu, \sigma^2) = \Phi\left(\frac{\mu}{\sigma}\right), \tag{3.6}$$

where Φ is the CDF for a standard normal distribution. A success can then be defined

as $PI(\mu, \sigma^2) > T$ for some defined risk tolerance measure T . As σ^2 decreases over more measurements, PI becomes less influenced by σ^2 and more by μ .

The stopping problem state follows a Beta distribution where parameters α, β correspond to the count of successes and failures, and transitions follow the posterior of the Beta distribution. In summary,

$$X_t = (\alpha_t + \alpha_0, \beta_t + \beta_0)$$

$$P(X_{t+1}) = \begin{cases} (\alpha_t + \alpha_0 + 1, \beta_t + \beta_0) \\ \text{w.p. } \frac{\alpha_t + \alpha_0}{\alpha_t + \alpha_0 + \beta_t + \beta_0} \\ (\alpha_t + \alpha_0, \beta_t + \beta_0 + 1) \\ \text{w.p. } \frac{\beta_t + \beta_0}{\alpha_t + \alpha_0 + \beta_t + \beta_0} \end{cases} \quad (3.7)$$

$$r(X_t) = \frac{\alpha_t + \alpha_0}{\alpha_t + \alpha_0 + \beta_t + \beta_0},$$

where α_0, β_0 are smoothing priors and α_t, β_t are successes and failures up to time t .

Gittins and Jones showed that an optimal policy could be constructed by calculating a value called the Gittins Index (Chakravorty and Mahajan, 2014). Let $g(x) = K$ for some constant K , then the Gittins Index is defined as

$$\nu(X_t) = (1 - \lambda) \min_K \{J_t(X_t) = K\}. \quad (3.8)$$

The Gittins Index is generally used for infinite horizon stopping problems, with $g(x) = \frac{\nu(x)}{1-\lambda}$ representing an infinite geometric sum to compensate for all future rewards. By setting $\lambda = 1 - \frac{1}{N}$, $\forall x \ 0 \leq \nu(x) \leq 1$.

In the case of a multi-armed bandit problem, the optimal policy is to play the arm with the highest Gittins Index. Ferguson (2006) shows that the stopping problem, or the one-armed bandit problem, is equivalent to having a second arm that provides a constant reward. The optimal policy then corresponds to playing the arm as long as its Gittins Index is above a specified threshold V . As T and V are correlated, one of the parameters can be fixed (such as $T = 0.5$) while the other is tuned to a preferred risk tolerance for expected

success rate.

This approach essentially acts as a low-pass filter for a potentially noisy metabolic estimator. The choice for V can similarly be thought of in terms of a Gittins Index

$$\begin{aligned} \nu((\alpha, \beta)) &= V \\ \alpha + \beta &= N, \end{aligned} \tag{3.9}$$

where α, β are successes and failures over the maximum duration. This is in essence asking the question: "If I were to continue measuring this parameter setting, what is the minimum number of successes that I would need to warrant the cost of experimental time?".

3.4 Adaptive Thresholds

Both the σ -offset and Gittins model have a risk tolerance threshold that may be difficult to tune across subjects. Another approach is to use the exploration phase of Bayesian optimization to create a schedule for the thresholds adapted to the subject's measurements. A simple approach would be to gauge the trajectory of thresholds if the stopping problem were run on the exploration phase, and accordingly use the most risk averse during the optimization. Let M and Σ be $|\mathbb{E}| \times N$ matrices providing traces where M_{ij}, Σ_{ij} correspond to $\mu_{x_j}, \sigma_{x_j}^2$ for exploration point i and measurement j . For each exploration point, the \hat{x}^* distribution can be the final distribution, (M_{iN}, Σ_{iN}) , to estimate the magnitude of thresholds when comparing similar cost estimates. Algorithm 2 details the calculation of a threshold vector A .

3.5 Simulations

Simulations were run on four optimization functions (Table 3.2) with Gaussian noise added proportional to the standard deviation of the respective function. The Hartmann-6 and Ackley functions have larger domain spaces with lower standard deviations, generally making them more difficult functions to optimize using Bayesian optimization. While the

```

Init  $|\mathbb{E}| \times N$  Matrix  $X$ 
Init  $N$ -Vector  $A$ 
Define Window  $W$ 
for  $i = 1$  to  $|\mathbb{E}|$ 
     $\alpha = \alpha_0, \beta = \beta_0$ 
    for  $j = 1$  to  $N$ 
         $\mu = M_{ij} - M_{iN}$ 
         $\sigma^2 = \Sigma_{ij} + \Sigma_{iN}$ 
        if  $Offset$ 
             $X_{ij} = \frac{\mu}{\sigma}$ 
        end
        if  $Gittins$ 
             $\alpha = \alpha + \mathbb{1}_{PI(\mu, \sigma^2) \geq T}$ 
             $\beta = \beta + \mathbb{1}_{PI(\mu, \sigma^2) < T}$ 
             $X_{ij} = v((\alpha, \beta))$ 
        end
    end
    end
for  $i = 1$  to  $N$ 
    if  $Offset$ 
         $A_i = MAX\{X_{rc} | 1 \leq r \leq |\mathbb{E}|, i - W \leq c \leq i + W\}$ 
    end
    if  $Gittins$ 
         $A_i = MIN\{X_{rc} | 1 \leq r \leq |\mathbb{E}|, i - W \leq c \leq i + W\}$ 
    end
end
Return  $A$ 

```

Algorithm 2: Determining threshold levels for either model based on subject's exploration data. With the σ -offset, the higher threshold is more risk averse, while with the Gittins model the lower is.

Levy and Branin functions have large regions of local minima, and in Branin's case three global minima, much higher noise is added due to the standard deviations of the respective functions.

Both the σ -offset and Gittins models were tested each with risk thresholds $T = [0, 0.25, 0.5, AD]$ where AD is the adaptive threshold method. A Gaussian state estimate was continually reconciled with new observations, equivalent to a Kalman filter with the identity function plus additive noise for both F and H , and measurement noise was tuned to different scenarios $P_w = [\frac{\sigma^2}{10}, \sigma^2, 10\sigma^2]$. P_x was initialized to μ, σ^2 of the function, and $P_v = \frac{\sigma^2}{100}$. The same observations from exploration points were provided for each function

Function	Domain	Minimum	Mean	Std
Hartmann-6	$[0, 1]^6$	-3.322	-0.259	0.383
Ackley	$[-32.768, 32.768]^4$	0	20.882	1.027
Levy	$[-10, 10]^4$	0	42.544	27.939
Branin	$([-5, 10], [0, 15])$	0.398	54.452	51.129

Table 3.2: Optimization functions for simulations

under all scenarios, and the maximum duration was set to 45 measurements.

For time t , let x_t be the GP's estimated minimum point. The figures chart

$$y_t = \frac{f(x_t) - f(x^*)}{f(x_0) - f(x^*)}, \quad (3.10)$$

where f is the true, noiseless function and $f(x^*)$ is the function minimum.

Results from simulations averaged over 100 trials are provided in Appendix D. The exploration phase is omitted from the figures as the same observations were used for all scenarios, though final estimates may vary across different P_w values. The case of $T = 0$ for the Gittins model is equivalent to a fixed measurement window and can be used as the point of comparison. Stopping early shows a clear benefit in Hartmann and Ackley, the less noisy functions. The σ -offset method generally outperforms the Gittins in these situations, and is comparable only to the most aggressive $T = 0.5$ case. As there is minimal noise, different levels of P_w show only slight differences.

In the case of higher noise with Branin and Levy, only Levy shows a clear advantage to stopping early. With the more fault-tolerant Gittins approach, different levels of P_w still show relatively consistent performance, while having a constant threshold in the σ -offset model quickly stagnates when underestimating measurement noise. Note these functions are not strictly decreasing - this implies a training sample has been added that causes the

Gaussian process to incorrectly assume a new minimum point. As P_w increases this becomes less apparent; a higher P_w causes σ^2 of the model to also increase, and as both models are σ -scaled the thresholds are rarely met until a more refined estimate is made.

The adaptive thresholds show fairly consistent performance across all scenarios, including with underestimating P_w in the σ -offset model. Generally a higher P_w estimate allows for a more predictable trajectory, which makes the adaptive thresholds more effective. In these experiments, eight exploration points were used, making the exploration phase one third of the entire trial time. Additional tests would include decreasing the number of exploration points, especially in the case of underestimating P_w with either the Branin or Levy functions, and altering the maximum duration window.

To date, wearable devices have shown to provide a metabolic reduction on the order of twenty percent. The standard deviation in metabolic rate across parameter settings, while heavily subject dependent, has ranged from ten to thirty percent across our trials so far, making the Branin and Levy functions a more appropriate comparison. Though difficult to measure the noisiness of metabolics as there is no true instantaneous energetic costs to compare against, the importance of the relative magnitudes between P_w and P_x, P_v are clear in both affecting the stopping process as well as Gaussian process estimates. The bandit model aims to reduce the effect of incorrectly assigning values to these covariances, while the σ -offset model generally performed the best when tuned properly.

Chapter 4

Subject Trials

4.1 Experimental Protocol

Optimization Parameters

An initial pilot was run varying 6 parameters

Peak Force	Maximum force in Newtons provided
Onset Timing	Start of assistance during the gait cycle
Peak Timing	Point of peak force provided during the gait cycle
Offset Timing	End of assistance during the gait cycle
Curvature 1	Force provided at midpoint of onset and peak
Curvature 2	Force provided at midpoint of peak and offset

Table 4.1: 6 parameters for providing a force profile

Hip-only Soft Exosuit

The textile components of the hip-only soft exosuit consist of a short spandex base layer, a waist belt, and two thigh braces as described in (Lee *et al.*). Two IMUs (MTi-3 AHRS, Xsens Technologies B.V., Enschede, Netherlands) mounted on the anterior part of the thigh measure the orientation and angular velocity of thigh segments which are used for gait cycle estimation in real-time (Ding *et al.*, 2016). The anchoring points of the Bowden cable are at the bottom left and right of the waist belt as well as at the middle center of the thigh braces from the posterior view. Two load cells (LSB200, Futek Advanced Sensor Technology Inc., CA, USA) are integrated in the aluminum attachment piece of the anchoring point to measure the cable force. When the actuator pulls the inner cable, the soft exosuit generates a hip extension torque by shortening the distance between two anchoring points.

Actuation System

The two DOF actuators in the off-board system are used to deliver hip extension assistance to both legs. Each actuator consists of a customized brushless motor (Allied Motion, Colorado, USA), a 9.25:1 spiroid gear set (ITW Heartland, MN, USA) and a 45 mm radius pulley. The pulley groove is designed to be able to wrap the inner Bowden cable 1.5 times around the pulley, allowing 420 mm of cable travel (Lee *et al.*). A 360 counts/rev incremental encoder (AS5134-ZSST, ams AG, Premstaetten, Austria) is mounted on the customized motor to measure its position and velocity. The Bowden cable sheath of the actuator side connects to the pulley cover and inner cable attaches to the pulley. When the motor rotates, the actuator can either pull in the inner cable to generate a force or push the cable out so it can go slack.

Control System Architecture

The electronics hardware has a three-layer configuration for the control system architecture: real-time target machine (top layer), microprocessor (middle layer), and servomotor driver (bottom layer). In the top layer, the Speedgoat real-time target machine runs a Simulink model in a host laptop. A PCI interface card (CAN-AC PCI, Softing AG, Haar, Germany) for

CAN communication is installed on the target machine to acquire sensor signals and send the command to the servomotor driver through the microprocessor. The microprocessor (Due, Arudino, Ivrea, Italy) in the middle layer acts as a signal hub between target machine and servomotor driver while protecting the system from undesired behavior of the actuator due to anomalous errors from the target machine. In the bottom layer, Gold Twitter (Elmo Motion Control Ltd, Israel) with 55V maximum supply voltage and 60 A peak current is selected as the current servomotor driver. It tracks the current command from the microprocessor to drive the actuator. The CAN communication between layers closes the control loop at 1kHz loop rate (Lee *et al.*).

Another laptop collects the metabolic data, runs an optimization algorithm, and sends the assistance profile parameters that will be explored in the next condition to the host laptop. Ethernet communication is used between a real-time target machine, a host laptop, and the second laptop. Based on the profile parameters calculated by the algorithm, the host laptop sends the current command to the servomotor drive to deliver assistance to the subject.

4.2 Experimental Subjects

Two male subjects (S1: 27 yrs, 74 kg, 177 cm; S2: 48 yrs, 85 kg, 178 cm) participated in this preliminary study. The subjects walked on the treadmill at 1.25 m/s under four different conditions: Normal walking without Exosuit (N), Exosuit-off (OFF), Fixed Exosuit Assistance (FIX) and Optimal Exosuit Assistance (OPT). During FIX condition, the force profiles assisting hip extension were based on the previous study of (Ding *et al.*, 2016), whereas force profiles in OPT condition were determined through a two-part procedure (8 exploration points, 5 minute break, then 40 minute continuous optimization) which was performed before walking test conditions. Metabolic cost of walking across different conditions was measured using a portable gas analysis system (K4b2, Cosmed, Roma, Italy).

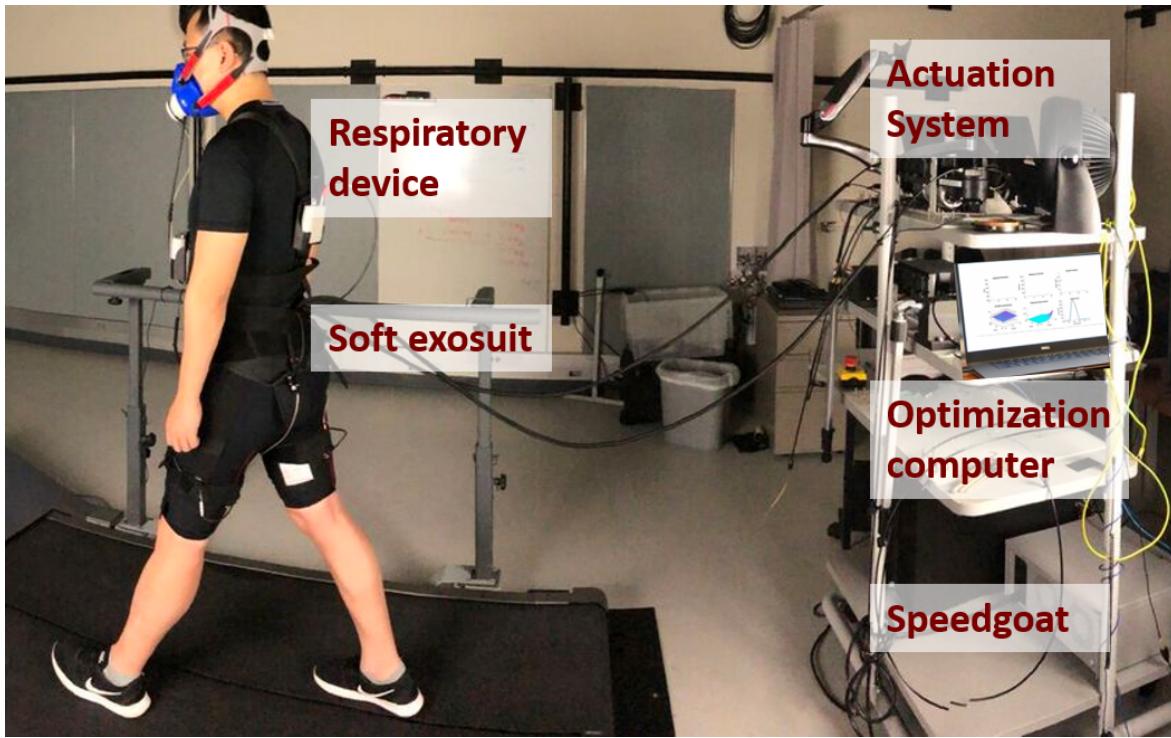


Figure 4.1: Experimental setup. Assistance parameters for the soft exosuit were optimized on-line. Respiratory device measured respiratory rate to estimate metabolic cost for a given condition with a covariance. In the optimization computer, the Gittins process determined when to stop sampling and Bayesian optimization was used to select a next parameter set. The parameters were sent to the real-time computer (speedgoat) to generate a force trajectory. Then, the actuator provided a force to a subject through a bowden cables.

4.3 Results

Two human trials with the same subjects from a previous study were conducted as a comparison. The Gittins method was used with a relatively risk averse $T = 0.35$ to account for the possibility of a very low signal-to-noise ratio, as well as the practical consideration of parameters requiring time to fully propagate to the suit. A summary of the metabolic reduction can be found in Table 4.2 with figures for the optimized force assistance in Figure 4.3. The soft exosuit followed desired trajectory within 5% error, relative to the peak force.

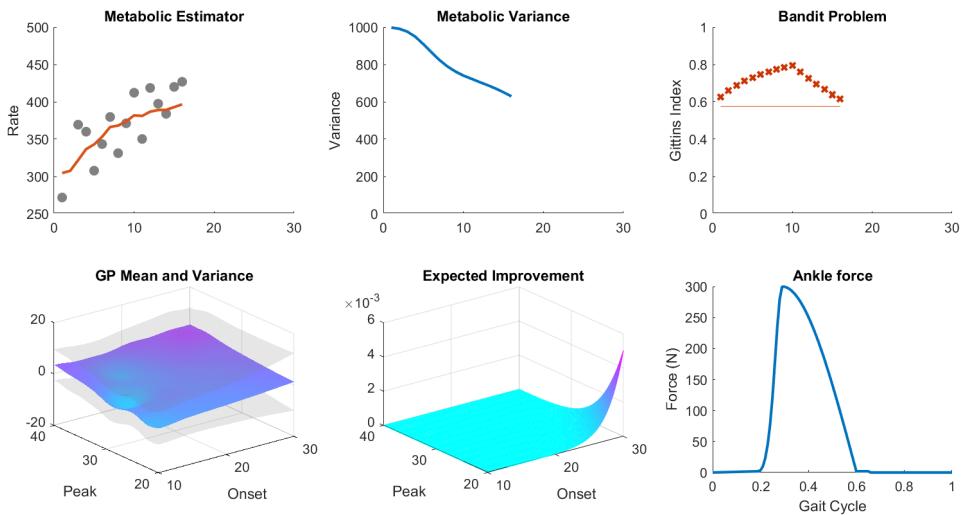


Figure 4.2: Diagram of Optimization Process

Subj	OPT	FIX	OPT Peak F	FIX Peak F
1	35%	29%	223N	247N
2	7%	-4%	174N	230N

Table 4.2: Subject Trial Summary

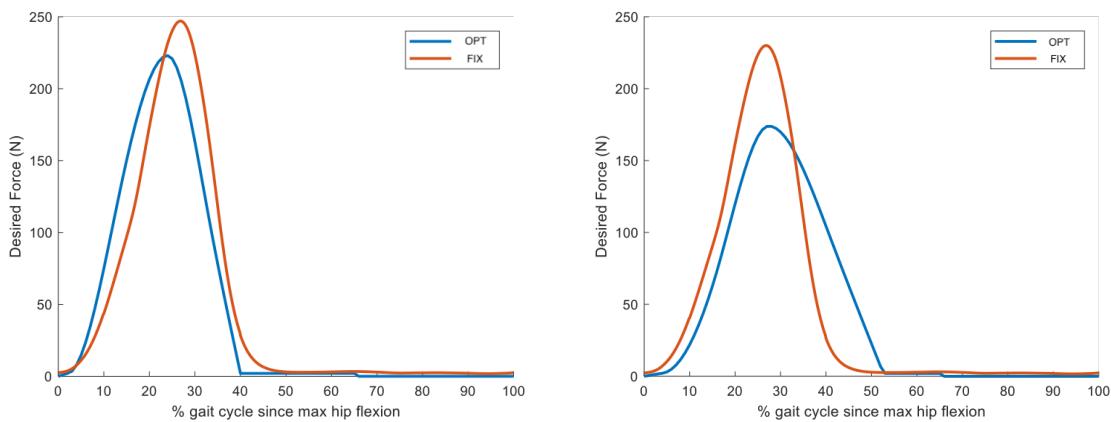


Figure 4.3: Force profiles of the two subjects

Conclusion

Summary

This thesis aims to introduce a stopping criteria for measurement time within HIL optimization. A metabolic estimator based off the Unscented Kalman Filter algorithm was designed and has effectively shown the ability to track the instantaneous energetic cost online with varying amounts of data. Two methods were developed for early stopping: a simple σ -offset threshold strategy and a bandit process strategy based off the Gittins Index. Depending on the noisiness of the estimator, the Gittins approach can provide a higher fault tolerance for noisy observations. Two trials were then conducted using the Gittins approach to optimize over six hip parameters and similar metabolic reduction was achieved compared to a previous study with the same subjects. While the previous study had a shorter duration time, it was only a two-parameter study and ample time was spent prior to the subject tests isolating the parameters individually to find the two with highest variability (peak and offset timing). In our case, we were able to find similar results without the need for an extensive pre-trial exploration phase. Interestingly we also found a lower peak force to provide similar metabolic reduction in each of the two subjects. Having a lower peak force provides two benefits: lower battery requirements on the suit and reduced strain on the harness and subsequent discomfort to the user.

Challenges and Future Work

There were a number of implementation choices that can be improved upon in future work. While collecting ample metabolic data is a time-constrained challenge, in the future a prior metabolic landscape could potentially be developed as a mean function for the Gaussian process. With respect to the Bayesian optimization, accounting for the early stopping of parameter settings led to fixed noise hyperparameters based off the metabolic estimator. Alternatively, a fully heteroscedastic approach could be experimented with that involves fitting output noise to a secondary Gaussian process. For the metabolic estimator, the high subject-to-subject variability in measurements requires some optimization for the covariance matrices. Given Bayesian optimization has an exploration phase, tuning the estimator based on this data should improve reliability. Similarly, different methods for tuning the risk tolerances of the stopping algorithms could be analyzed through additional subject tests. Two subject trials is certainly not enough, and many more subject tests should be carried out for a better analysis of the stopping algorithms in practice. Finally, an area of interest for future studies is multi-joint optimization using both ankle and hip actuation (Lee *et al.*, 2018). This would be a natural use case for the stopping algorithm as the number of optimization parameters could potentially double. Initially, a pilot would involve optimizing four parameters: peak timing in both the hip and ankle, onset timing in the ankle, and offset timing in the hip.

References

- ASSAEL, J.-A. M., WANG, Z., SHAHRIARI, B. and DE FREITAS, N. (2014). Heteroscedastic treed bayesian optimisation.
- BARBER, D., CEMGIL, A. and CHIAPPA, S. (2011). *Bayesian Time Series Models*. Bayesian Time Series Models, Cambridge University Press.
- BROCHU, E., HOFFMAN, M. W. and DE FREITAS, N. (2010). Portfolio allocation for bayesian optimization.
- BROCHU, V. M. C., ERIC and DE FREITAS, N. (2010). A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *Technical Report*, pp. TR-2009-23.
- BROCKWAY, J. M. (1987). Derivation of formulae used to calculate energy expenditure in man. **41C**, 463–471.
- CAPUTO, J. M., ADAMCZYK, P. G. and COLLINS, S. H. (2015). Informing ankle-foot prosthesis prescription through haptic emulation of candidate devices. pp. 6445 – 6450.
- and COLLINS, S. H. (2014). A universal ankle-foot prosthesis emulator for experiments during human locomotion. *JBME*, **136**, 035002.
- CHAKRAVORTY, J. and MAHAJAN, A. (2014). *Multi-Armed Bandits, Gittins Index, and its Calculation*, Wiley-Blackwell, chap. 24, pp. 416–435.
- COLLINS, S. H., WIGGIN, M. B. and SAWICKI, G. S. (2015). Reducing the energy cost of human walking using an unpowered exoskeleton. *Nature*, **522**, 212–215.
- DING, Y., GALIANA, I., ASBECK, A. T., ROSSI, S. M. M. D., BAE, J., SANTOS, T. R. T., DE ARAUJO, V. L., LEE, S., HOLT, K. G. and WALSH, C. (2017). Biomechanical and physiological evaluation of multi-joint assistance with soft exosuits. **25** (2), 119–130.
- , KIM, M., KUINDERSMA, S. and WALSH, C. J. (2018). Human-in-the-loop optimization of hip assistance with a soft exosuit during walking. *Science Robotics*, **3** (15).
- , PANIZZOLO, F. A., SIVIY, C., MALCOLM, P., GALIANA, I., HOLT, K. G. and WALSH, C. J. (2016). Effect of timing of hip extension assistance during loaded walking with a soft exosuit. *JNER*, **13** (1), 87.

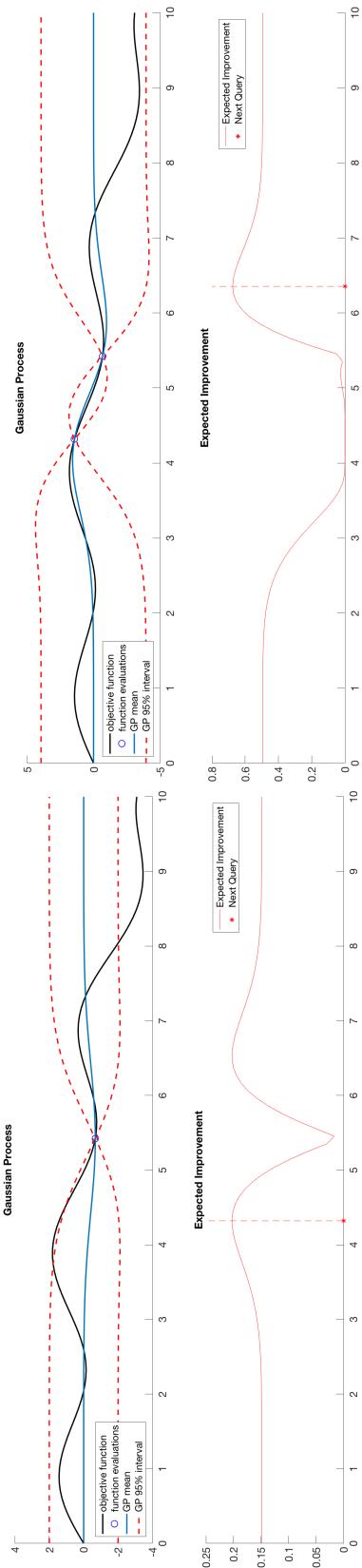
- FELT, W., SELINGER, J. C., DONELAN, J. M. and REMY, C. D. (2015). "Body-In-The-Loop": Optimizing Device Parameters Using Measures of Instantaneous Energetic Cost. *PloS one*, **10** (8), e0135342.
- FERGUSON, T. (2006). *Optimal Stopping and Applications*.
- GONZÁLEZ, J., OSBORNE, M. and LAWRENCE, N. D. (2015). Glasses: Relieving the myopia of bayesian optimisation.
- JARNO VANHATALO, J. H. P. J. V. T., JAAKKO RIIHIMÄKI and VEHTARI, A. (2013). Gpstuff: Bayesian modeling with gaussian processes. *Journal of Machine Learning Research*, **14**, 1175–1179.
- JULIER, S. J. (1997). A new extension of the kalman filter to nonlinear systems. *Int. Symp. Aerospace/Defense Sensing, Simul. and Controls*, **3**.
- JYLÄNKI, P., VANHATALO, J. and VEHTARI, A. (2011). Robust gaussian process regression with a student-t likelihood. *J. Mach. Learn. Res.*, **12**, 3227–3257.
- KERSTING, K., PLAGEMANN, C., PFAFF, P. and BURGARD, W. (2007). Most likely heteroscedastic gaussian process regression. In *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, New York, NY, USA: ACM, pp. 393–400.
- KIM, M. and COLLINS, S. H. (2015). Once-per-step control of ankle-foot prosthesis push-off work reduces effort associated with balance during walking. **12** (1), 43.
- and — (2017). Step-to-step ankle inversion/eversion torque modulation can reduce effort associated with balance. *Frontiers in neurorobotics*, **11**, 62.
- , DING, Y., MALCOLM, P., SPEECKAERT, J., SIVIY, C. J., WALSH, C. J. and KUINDERSMA, S. (2017). Human-in-the-loop bayesian optimization of wearable device parameters. *PloS one*, **12** (9), e0184054.
- KOLLER, J. R., GATES, D. H., FERRIS, D. P. and REMY, C. D. (2016). 'body-in-the-loop' optimization of assistive robotic devices: A validation study. In *Robotics: Science and Systems*.
- LAM, R., WILLCOX, K. and WOLPERT, D. H. (2016). Bayesian optimization with a finite budget: An approximate dynamic programming approach. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon and R. Garnett (eds.), *Advances in Neural Information Processing Systems* 29, Curran Associates, Inc., pp. 883–891.
- LEE, G., DING, Y., BUJANDA, I. G., KARAVAS, N., ZHOU, Y. M. and WALSH, C. J. (). Improved assistive profile tracking of soft exosuits for walking and jogging with off-board actuation.
- LEE, S., CREA, S., MALCOLM, P., GALIANA, I., ASBECK, A. and WALSH, C. (2016). Controlling negative and positive power at the ankle with a soft exosuit. pp. 3509–3515.
- , KARAVAS, N., QUINLIVAN, B., RYAN, D., PERRY, D., ECKERT-ERDHEIM, A., MURPHY, P., GREENBERG GOLDY, T., MENARD, N., ATHANASSIU, M., KIM, J., LEE, G., GALIANA, I. and WALSH, C. (2018). Autonomous multi-joint soft exosuit for assistance with walking overground. In *IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, Australia, May 21-25.

- LIZOTTE, D. J. (2008). *Practical Bayesian Optimization*. Ph.D. thesis, Edmonton, Alta., Canada, aAINR46365.
- MAHENDRAN, N., WANG, Z., HAMZE, F and FREITAS, N. D. (2012). Adaptive mcmc with bayesian optimization. In N. D. Lawrence and M. Girolami (eds.), *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*, La Palma, Canary Islands: PMLR, *Proceedings of Machine Learning Research*, vol. 22, pp. 751–760.
- MALCOLM, P., DERAVE, W., GALLE, S. and DE CLERCQ, D. (2013). A simple exoskeleton that assists plantarflexion can reduce the metabolic cost of human walking. *PLoS one*, **8** (2), e56137.
- MOONEY, L. M. and HERR., H. M. (2016). Biomechanical walking mechanisms underlying the metabolic reduction caused by an autonomous exoskeleton. **13** (1), 4.
- NOGUEIRA, J., MARTINEZ-CANTIN, R., BERNARDINO, A. and JAMONE, L. (2016). Unscented bayesian optimization for safe robot grasping.
- PANIZZOLO, F. A., GALIANA, I., ASBECK, A. T., SIVIY, C., SCHMIDT, K., HOLT, K. G. and WALSH, C. J. (2016a). A biologically-inspired multi-joint soft exosuit that can reduce the energy cost of loaded walking. **13** (1), 43.
- , —, —, —, — and WALSH., C. J. (2016b). A biologically-inspired multi-joint soft exosuit that can reduce the energy cost of loaded walking. *JNER*, **13**, 1–13.
- QUESADA, R. E., CAPUTO, J. M. and COLLINS, S. H. (2016). Increasing ankle push-off work with a powered prosthesis does not necessarily reduce metabolic rate for transtibial amputees. **49** (14), 3452–3459.
- QUINLIVAN, B., LEE, S., MALCOLM, P., ROSSI, D., GRIMMER, M., SIVIY, C., KARAVAS, N., WAGNER, D., ASBECK, A., GALIANA, I. and WALSH, C. (2017). Assistance magnitude versus metabolic cost reductions for a tethered multiarticular soft exosuit. *Science Robotics*, **2** (2), eaah4416.
- RASMUSSEN, C. E. (2006). Gaussian processes for machine learning.
- RUE, H., MARTINO, S. and CHOPIN, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **71** (2), 319–392.
- SELINGER, J. C. and DONELAN., J. M. (2014). Estimating instantaneous energetic cost during non-steady-state gait. **117** (11), 1406–1415.
- SNOEK, J., LAROCHELLE, H. and ADAMS, R. P. (2012). Practical bayesian optimization of machine learning algorithms. In F. Pereira, C. J. C. Burges, L. Bottou and K. Q. Weinberger (eds.), *Advances in Neural Information Processing Systems 25*, Curran Associates, Inc., pp. 2951–2959.
- SOFIA S. VILLAR, J. B. and WASON, J. (2015). Multi-armed bandit models for the optimal design of clinical trials: Benefits and challenges. *Statistical Science*, **30** (2), 199–215.

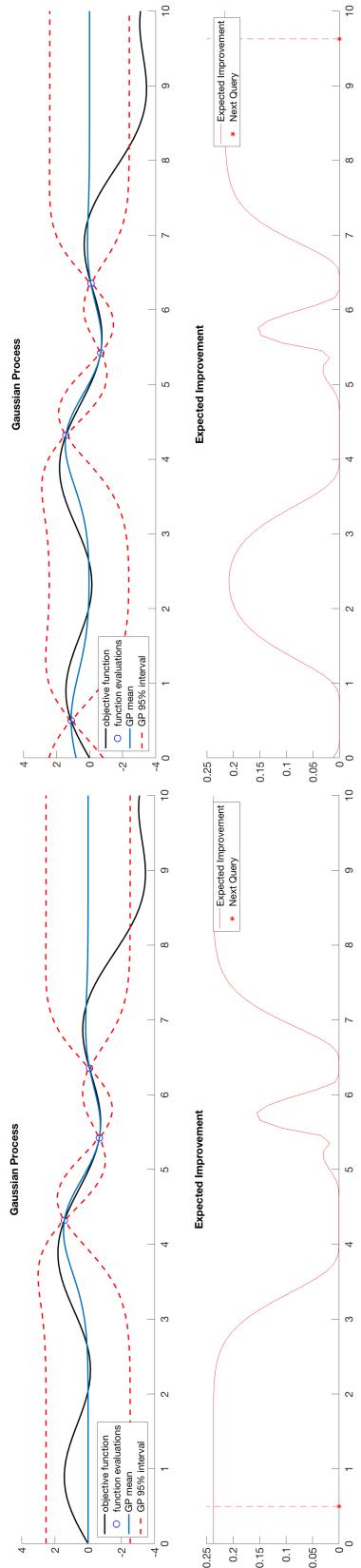
- VANHATALO, J., JYLÄNKI, P. and VEHTARI, A. (2009). Gaussian process regression with student-t likelihood. In Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams and A. Culotta (eds.), *Advances in Neural Information Processing Systems 22*, Curran Associates, Inc., pp. 1910–1918.
- WAN, E. A. and MERWE, R. V. D. (2000). The unscented kalman filter for nonlinear estimation. pp. 153–158.
- ZHANG, J., FIERS, P., WITTE, K. A., JACKSON, R. W., POGGENSEE, K. L., ATKESON, C. G. and COLLINS, S. H. (2017). Human-in-the-loop optimization of exoskeleton assistance during walking. *Science*, **356** (6344), 1280–1284.

Appendix A

Bayesian Optimization



(a) Iterations 1 and 2



(b) Iterations 3 and 4

Figure A.1: Four Iterations of Bayesian Optimization

Appendix B

Unscented Kalman Filter¹

The unscented transformation selects a representative set of points (sigma points) with associated weights to propagate through a nonlinear function. More formally, given an N -dimensional random variable x with mean \bar{x} and covariance P_x we calculate the following sigma vectors

$$\begin{aligned}\chi_0 &= \bar{x} \\ \chi_i &= \bar{x} + (\sqrt{(N + \lambda)P_x})_i & i = 1, \dots, N \\ \chi_i &= \bar{x} - (\sqrt{(N + \lambda)P_x})_{i-N} & i = N + 1, \dots, 2N \\ W_0^{(m)} &= \lambda / (N + \lambda) \\ W_0^{(c)} &= \lambda / (N + \lambda) + (1 - \alpha^2 + \beta) \\ W_i^m = W_i^c &= 1 / \{(2(N + \lambda))\} & i = 1, \dots, 2N,\end{aligned}\tag{B.1}$$

where $(\sqrt{(N + \lambda)P_x})_i$ represents the i th row of the matrix square root, $\lambda = \alpha^2(N + \kappa) - N$, and α , κ , and β are scaling parameters. For a given nonlinear function $y = g(x)$, an estimate for the mean and covariance are then

$$\begin{aligned}\bar{y} &\approx \sum_{i=0}^{2N} W_i^{(m)} g(\chi_i) \\ P_y &\approx \sum_{i=0}^{2N} W_i^{(c)} (g(\chi_i) - \bar{y})(g(\chi_i) - \bar{y})^T\end{aligned}\tag{B.2}$$

¹Wan and Merwe (2000)

Accurate to the second order, the Unscented Kalman Filter incorporates this transformation in its F and H functions as follows:

Initial State Estimate and Covariance $\hat{x}(0), P_x(0)$

Define Augmented State $x^a = [x^T \quad v^T \quad w^T]^T$

Define Augmented Sigma Points $\chi^a = [(\chi^x)^T \quad (\chi^v)^T \quad (\chi^w)^T]^T$

Initial Augmented State Estimate $\hat{x}^a(0) = [\hat{x}(0)^T \quad 0 \quad 0]^T$

Initial Augmented Covariance $P^a(0) = \begin{bmatrix} P_x(0) & 0 & 0 \\ 0 & P_v & 0 \\ 0 & 0 & P_w \end{bmatrix}$

for $t = 1$ **to** T

Using B.1 Calculate Sigma Points $\chi^a(t)$ And Weights $W^{(m)}, W^{(c)}$

Time Update:

$$\chi_i^x(t|t-1) = F(\chi_i^x(t-1), \chi_i^v(t-1), t-1)$$

$$\hat{x}(t|t-1) = \sum_{i=0}^{2N} W_i^{(m)} \chi_i^x(t|t-1)$$

$$P_x(t|t-1) = \sum_{i=1}^{2N} W_i^{(c)} \{ \chi_i^x(t|t-1) - \hat{x}(t|t-1) \} \{ \chi_i^x(t|t-1) - \hat{x}(t|t-1) \}^T$$

$$\mathbb{Z}_i(t|t-1) = H(\chi_i^x(t|t-1), \chi_i^v(t|t-1), t-1)$$

$$\hat{z}(t|t-1) = \sum_{i=1}^{2N} W_i^{(m)} \mathbb{Z}_i(t|t-1)$$

Measurement Update:

$$P_y(t|t-1) = \sum_{i=1}^{2N} W_i^{(c)} \{ \mathbb{Z}_i(t|t-1) - \hat{z}(t|t-1) \} \{ \mathbb{Z}_i(t|t-1) - \hat{z}(t|t-1) \}^T$$

$$P_{xy}(t|t-1) = \sum_{i=1}^{2N} W_i^{(c)} \{ \chi_i^x(t|t-1) - \hat{x}(t|t-1) \} \{ \mathbb{Z}_i(t|t-1) - \hat{z}(t|t-1) \}^T$$

$$K = P_{xy}(t|t-1) P_y^{-1}(t|t-1)$$

$$\hat{x}(t) = \hat{x}(t|t-1) + K(z(t) - \hat{z}(t|t-1))$$

$$P_x(t) = P_x(t|t-1) - K P_y(t|t-1) K^T$$

end

Appendix C

Optimal Stopping Time Derivation

Define the process

$$Z_n = \sum_{i=0}^{n-1} \lambda^i r(X_i) + \lambda^n J_n(X_n)$$

We have

$$\begin{aligned} \mathbb{E}[Z_{n+1}|X_n, X_{n-1}, \dots, X_0] &= \sum_{i=0}^n \lambda^i r(X_i) + \lambda^{n+1} \mathbb{E}[J_{n+1}(X_{n+1})|X_n] \\ &= \sum_{i=0}^{n-1} \lambda^i r(X_i) + \lambda^n (r(X_n) + \lambda \mathbb{E}[J_{n+1}(X_{n+1})|X_n]) \\ &\leq \sum_{i=0}^{n-1} \lambda^i r(X_i) + \lambda^n J_n(X_n) = Z_n \end{aligned}$$

Since by definition $J_n(x) \geq g(x)$, we have

$$\mathbb{E}\left[\sum_{i=0}^{\tau-1} \lambda^i r(X_i) + \lambda^\tau g(X_\tau)\right] \leq \mathbb{E}[Z_\tau] \leq \mathbb{E}[Z_0] = J_0(X_0)$$

for any stopping time $\tau \geq 0$.

We now prove that $J_0(X_0)$ is exactly equal to $\mathbb{E}[Z_{\tau^*}]$ for optimal stopping time τ^* .

Consider the following for $0 \leq n < N$

$$\begin{aligned}
& \mathbb{E}[Z_{\min(\tau^*, n+1)} | X_n, X_{n-1}, \dots, X_0] \\
&= 1_{\{\tau^* \leq n\}} Z_{\tau^*} + 1_{\{\tau^* > n\}} \mathbb{E}[Z_{n+1} | X_n, X_{n-1}, \dots, X_0] \\
&= 1_{\{\tau^* \leq n\}} Z_{\tau^*} + 1_{\{\tau^* > n\}} \left(\sum_{i=0}^n \lambda^i r(X_i) + \lambda^{n+1} \mathbb{E}[J_{n+1}(X_{n+1}) | X_n] \right) \\
&= 1_{\{\tau^* \leq n\}} Z_{\tau^*} + 1_{\{\tau^* > n\}} \left(\sum_{i=0}^{n-1} \lambda^i r(X_i) + \lambda^n J_n(X_n) \right) \quad (\tau^* > n \leftrightarrow J_n(X_n) = r(x) + \lambda \mathbb{E}[J_{n+1}(X_{n+1})]) \\
&= 1_{\{\tau^* \leq n\}} Z_{\tau^*} + 1_{\{\tau^* > n\}} Z_n \\
&= Z_{\min(\tau^*, n)}
\end{aligned}$$

It follows that

$$\begin{aligned}
J_0(X_0) &= \mathbb{E}[Z_0] = \mathbb{E}[Z_{\min(\tau^*, 0)}] = \dots = \mathbb{E}[Z_{\min(\tau^*, N)}] = \mathbb{E}[Z_{\tau^*}] \\
&= \mathbb{E}\left[\sum_{i=0}^{\tau^*-1} \lambda^i r(X_i) + \lambda^{\tau^*} J_{\tau^*}(X_{\tau^*})\right] = \mathbb{E}\left[\sum_{i=0}^{\tau^*-1} \lambda^i r(X_i) + \lambda^{\tau^*} g(X_{\tau^*})\right]
\end{aligned}$$

and $J_{\tau^*}(X_{\tau^*}) = g(X_{\tau^*})$ at τ^* .

Appendix D

Simulation Results

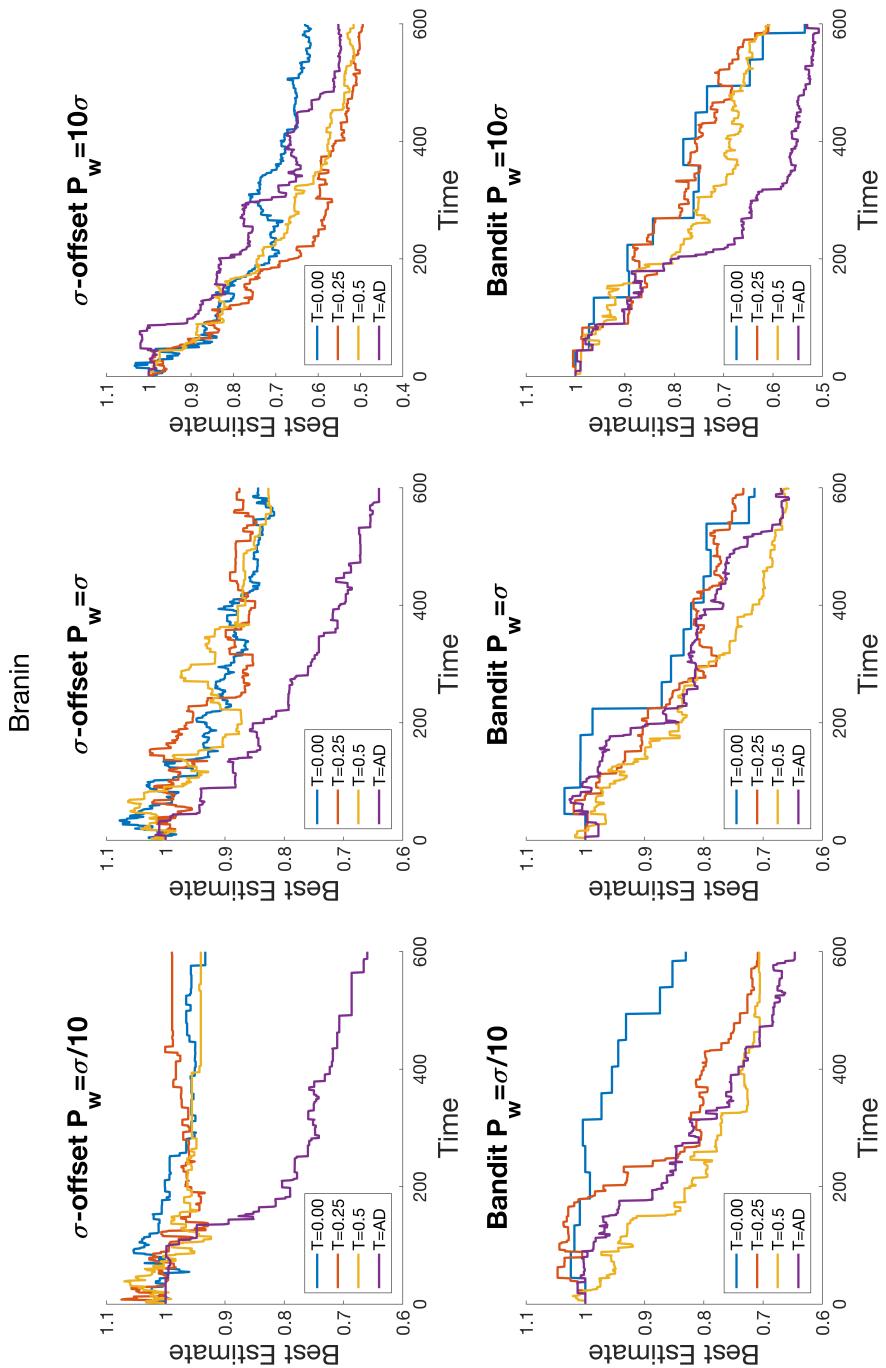


Figure D.1: Branin Function

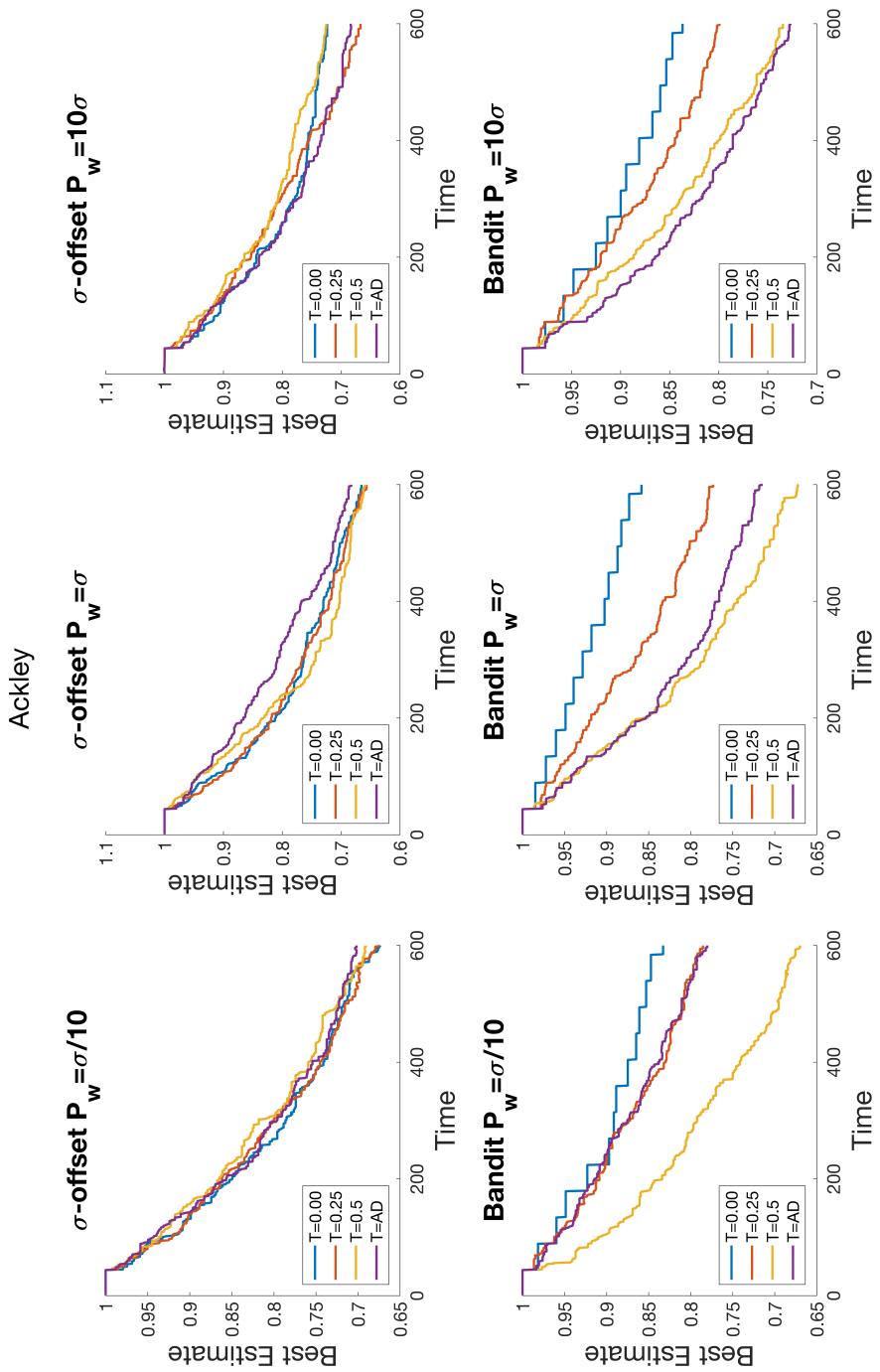


Figure D.2: Ackley Function

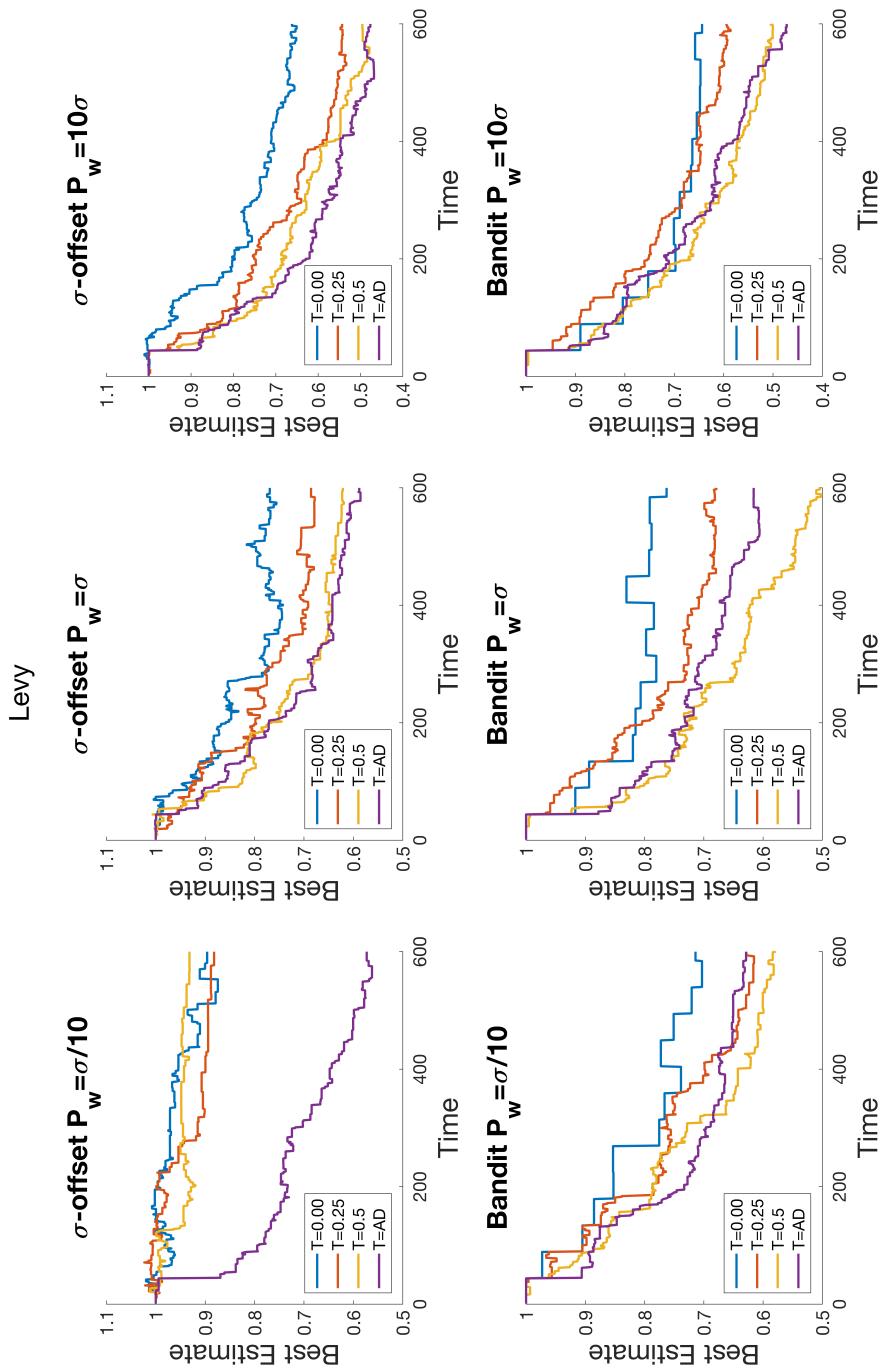


Figure D.3: Levy Function

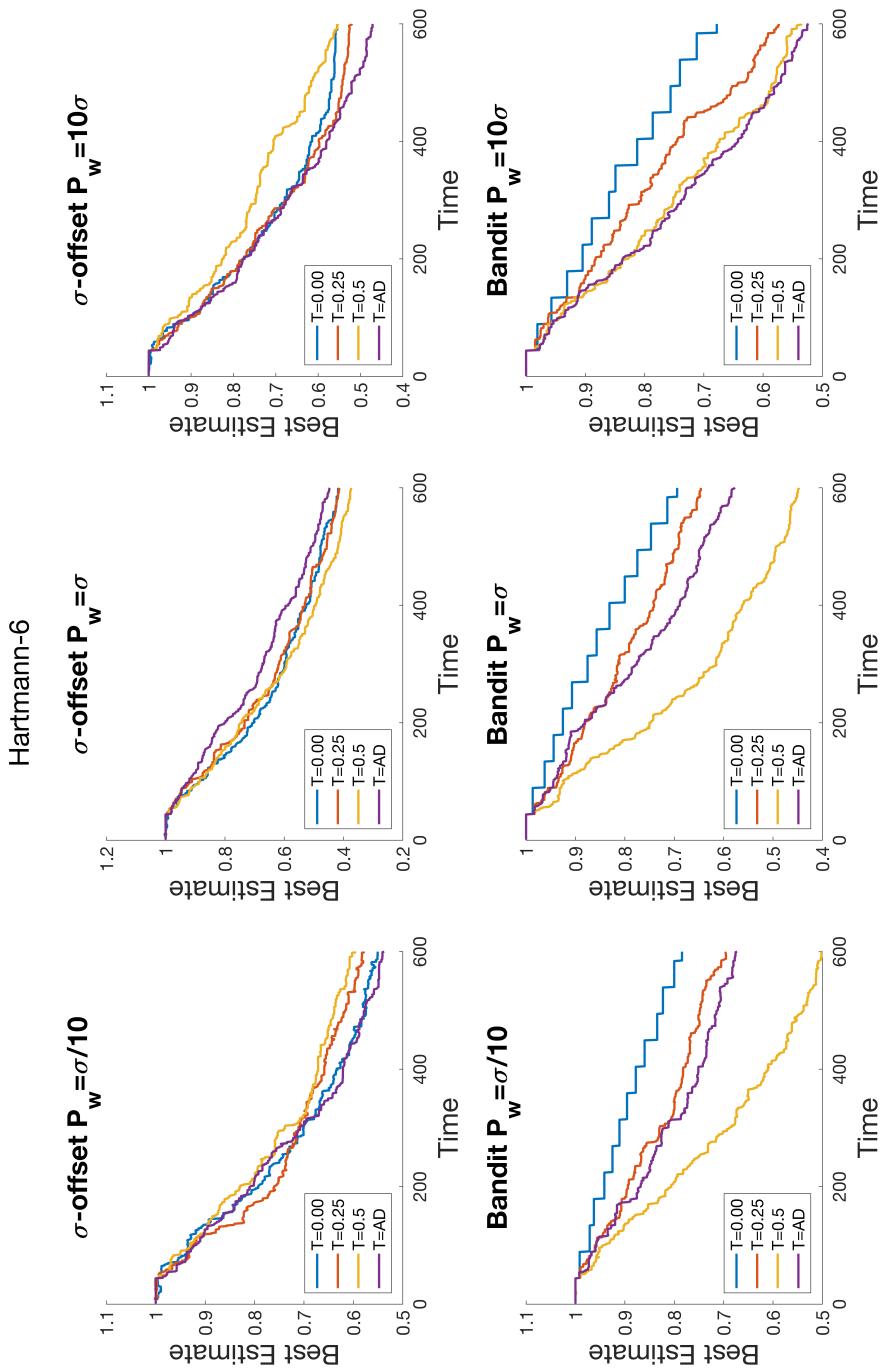


Figure D.4: Hartmann-6 Function