An Experimental Comparison of Bayesian Optimization for Bipedal Locomotion

Roberto Calandra\textsuperscript{1}, André Seyfarth\textsuperscript{2}, Jan Peters\textsuperscript{1,3} and Marc Peter Deisenroth\textsuperscript{4,1}

Abstract—The design of gaits and corresponding control policies for bipedal walkers is a key challenge in robot locomotion. Even when a viable controller parametrization already exists, finding near-optimal parameters can be daunting. The use of automatic gait optimization methods greatly reduces the need for human expertise and time-consuming design processes. Many different approaches to automatic gait optimization have been suggested to date. However, no extensive comparison among them has yet been performed. In this paper, we present some common methods for automatic gait optimization in bipedal locomotion, and analyze their strengths and weaknesses. We experimentally evaluated these gait optimization methods on a bipedal robot, in more than 1800 experimental evaluations. In particular, we analyzed Bayesian optimization in different configurations, including various acquisition functions.

I. INTRODUCTION

Bipedal locomotion is not robust, fast or energetically efficient [11] compared to wheeled or quadrupedal locomotion [25]. However, bipedal locomotion in robotics possesses various advantages: 1) high versatility with a wide range of gaits, from walking in rough terrains to fast running; 2) bipedal robots can naturally make use of environments designed for humans (e.g., stairs) without requiring architectonic modifications.

Key challenges in bipedal locomotion include balance control, foot placement, and gait optimization. In this paper, we focus on black-box gait optimization, i.e., finding good parameters for the gait of a biped, without the need of a dynamics model. Hence, we assume that a suitable controller to generate the desired gait has already been designed, but that appropriate gait parameters for the controller still need to be found. Due to the partially unpredictable effects and correlations among the gait parameters, gait optimization is often an empirical, time-consuming and strongly robot-specific process. In practice, gait optimization often translates into a trial-and-error process where choosing the parameters becomes either an educated guessing by a human expert or a systematic search, such as grid search. As a result, gait optimization may require considerable expert experience, engineering efforts and time-consuming experiments. Additionally, the effectiveness of the resulting gait parameters is limited by the circumstances and human insights speculated during the search process.

Therefore, a change in the environment (e.g., different floor surfaces), a variation in the hardware response (e.g., performance decline, substitution of a component or differences in the calibration) or the choice of a performance criterion (e.g., walking speed, energy efficiency, robustness), which differs from the one used during the controller design process, can require searching for new and more appropriate gait parameters.

By formulating the search for appropriate gait parameters as an optimization problem, it is possible to automate gait optimization. Automatic gait optimization is a valuable and principled approach to designing controllers and reduces the need for engineering expert knowledge. To date, various automatic gait optimization methods have been used in locomotion to design efficient gaits including: gradient descent methods [27], Bayesian optimization [18], [28], [5], genetic algorithms [6], particle swarm optimization [20] and others [12]. In the context of robotics, and specifically gait optimization, the number of experiments that can be performed on a real system is small. Each experiment can be costly, requires a long time, and inevitably leads to wear and tear of the robot’s hardware. Hence, it is crucial for the chosen optimization method to limit the number of experiments to perform while reliably finding near-optimal parameters. In the gait optimization literature, the results of the proposed methods are frequently presented without evaluating alternative methods. Additionally, each proposed method typically is evaluated on customized hardware. Therefore, any comparison inferred by a reader can give only a limited insight into the performance of different gait optimization methods. Furthermore, only rarely the characteristics, advantages and limitations of each method are clearly presented.

In this paper, we present an experimental evaluation and
comparison of common automatic optimization methods in the context of gait optimization. In Section [II] we formalize gait search as an optimization problem and introduce some popular optimization methods: grid search, pure random search, Bayesian optimization and gradient-descent-based methods. Furthermore, we discuss the properties and characteristics of the different methods. In Section [III] we present the experimental set-up and the results obtained using the bio-inspired dynamical bipedal walker Fox shown in Figure 1. We conclude in Section [IV] with some final remarks and possible future directions for gait optimization. Clearly, the experimental results obtained with the optimization methods considered in this paper are more generally applicable in robotics, for example, in grasping, imitation learning and robot-human interaction.

II. METHODS FOR GAIT OPTIMIZATION

The search for appropriate parameters of a controller can be formulated as an optimization problem, such as the minimization

\[
\text{minimize } f(\theta) \quad (1)
\]

of an objective function \( f(\cdot) \) with respect to the parameters \( \theta \). In the case of gait optimization, the objective function \( f \) encodes a performance criterion, such as walking speed, energy efficiency or robustness, while \( \theta \) are the parameters of the gait controller. Hence, evaluating the objective function \( f \) for a given set of parameters requires a physical interaction with the robot.

The considered gait optimization problem has the following properties:

1) **Zero-order objective function.** Each evaluation of the objective function \( f(\cdot) \) returns only the value of the function \( f(\theta) \), but no gradient information \( \nabla_{\theta} f = df(\theta)/d\theta \) with respect to the parameters. The use of first-order optimization methods, which make use of gradient information, is generally desirable in optimization as it leads to faster convergence than zero-order methods. Thus, it is common for zero-order objective functions to approximate the gradient using finite differences. However, finite differences require evaluating the objective function \( f \) multiple times. Since each evaluation requires interactions with the robot, the number of robot experiments quickly becomes excessive, rendering first-order methods (e.g., the whole family of gradient descent-based methods) unsuitable for our task.

2) **Stochastic objective function.** The evaluation of the objective function is inherently stochastic due to noisy measurements, variable initial conditions and system uncertainties (e.g., slack). Therefore, any suitable optimization method needs to take into consideration that two evaluations of the same parameters \( \theta \) can yield two different values \( f_1(\theta) \neq f_2(\theta) \).

3) **Global optimization.** Ideally, we strive to find the global minimum of the objective function in (1). However, no assumption can be made about the presence of multiple local minima or the convexity of the objective function.

These characteristics make this family of problems a challenging optimization task. Table I shows the methods commonly used to solve such a family of problems in the context of gait optimization. In the following, we introduce and analyze these optimization methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Order optimizer</th>
<th>Global optimizer</th>
<th>Re-usability evaluations</th>
<th>Stochasticity assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid Search</td>
<td>Zero-order</td>
<td>Global</td>
<td>Limited</td>
<td>No*</td>
</tr>
<tr>
<td>Pure Random Search</td>
<td>First-order</td>
<td>Local</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Gradient-descent Family</td>
<td>First-order</td>
<td>Global</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Bayesian Optimization</td>
<td>Zero-order</td>
<td>Global</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**TABLE I:** Comparison of various gait optimization methods. (*) extensions exist for stochastic optimization, but they are impractical in terms of number of evaluations required.
parallelization. Despite its simplicity pure random search is an attractive approach due to the statistical guarantees of global convergence. Given \( n \) evaluations and an optimum having a relative space \( a \) (i.e., size of the space of the optimum with respect to the size of the overall parameters space), the probability \( S \) of finding the global optimum is

\[
S = 1 - (1 - a)^n. \tag{3}
\]

Typically, such probabilistic convergence guarantees require an impractical (i.e., tending to infinite) number of evaluations. However, pure random search often performs surprisingly well in finding good approximate solutions, even compared to more complicated and modern optimization approaches, e.g., in high-dimensional problems where many irrelevant dimensions exist \([1]\).

Similarly to grid search, pure random search does not explicitly consider the case of stochastic objective functions, in which case it can fail to converge. Also for pure random search, there are extensions which can deal with stochastic optimization \([7]\), but they are impractical due to the even higher numbers of evaluations required. Adding new evaluations, if an optimization proved unsatisfactory, can be easily done by sampling new parameters and evaluating them as well.

C. Gradient-descent Family

Gradient-descent-based optimization is a wide family of optimization methods. They are based on gradient descent (GD), sometimes called “steepest descent”, which is a first-order iterative optimization method. At each iteration \( t \), a set of parameters \( \theta \) is selected based on the update rule

\[
\theta_{t+1} = \theta_t + \epsilon_t \nabla_{\theta_t} f, \tag{4}
\]

where the learning rate \( \epsilon_t \) (also called step size) is a free parameter and \( \nabla_{\theta_t} f \) is the gradient of the objective function with respect to the parameters \( \theta_t \). The selection of \( \epsilon_t \) is an important choice that can considerably affect the optimization process. While a constant \( \epsilon_t \) can be used, it would lead to an inefficient optimization, given that it would have to be tuned based on the particular specifics of the problem we want to optimize. A second possibility is the use of an adaptive step size (e.g., one that decreases through time). An appropriate adaptive learning rate toward the end of the optimization stabilizes the optimization process in a near-optimal minimum, rather than zig-zagging around. A common choice is to select an optimal \( \epsilon \) for each iteration using a linesearch. Other methods belonging to the gradient-descent family are conjugate gradients and L-BFGS \([4]\).

There are important limitations of gradient-descent methods for the formulated optimization problem. 1) Computing the gradient \( \nabla_{\theta_t} f \) is undesirable in our case as it increases the number of evaluations required. 2) Gradient-descent methods are local: since they follow the gradient, they only find the minimum in whose basin of attraction the optimization started. Hence, the initialization of the parameters \( \theta_0 \) becomes crucial and global optima are rarely found. Gradient-descent methods are occasionally used for global optimization by running multiple optimizations initialized by different parameters \( \theta_0 \). However, this technique clearly requires many more evaluations, and therefore is impractical. Since each new set of parameters \( \theta_{t+1} \) depends only on the previous set of parameters \( \theta_t \), gradient-descent methods do not allow for any re-use of past evaluations.

D. Bayesian Optimization

Bayesian optimization is an iterative model-based global optimization method \([16], [14], [21], [2]\). In iterative model-based global optimization methods, after each evaluation of the objective function \( f \), a model that maps parameters \( \theta \) to corresponding function evaluations \( f(\theta) \) is built. From the resulting model the response surface \( \hat{f}(\cdot) \) is predicted and used for a “virtual” optimization process

\[
\minimize_{\theta \in \mathbb{R}^D} \hat{f}(\theta). \tag{5}
\]

In this context, “virtual” indicates that optimizing the response surface \( \hat{f}(\cdot) \) with respect to the parameters \( \theta \) does not need interactions with the real system, but only evaluations of the learned model. Only when a new set of parameters \( \theta^* \) has been selected from the virtual optimization process of the response surface \( \hat{f} \), they are evaluated on the real objective function \( f \). The resulting value of the objective function, together with the corresponding parameters \( \{\theta^*, f(\theta^*)\} \), are used to update the model of the objective function.

A variety of different models, such as linear functions or splines \([14]\), have been used in the past to model \( f \). In Bayesian optimization, probabilistic models are used. The use of a probabilistic model allows to model noisy observations and to explicitly take the uncertainty about the model itself into account. Additionally, such a probabilistic framework allows to use priors that encode available expert knowledge or information from related systems, such as optimal parameter priors to changes in the system, e.g., after replacing a motor or changing the walking surface. In this paper, we use Gaussian processes (GPs) as probabilistic model for Bayesian optimization.

When using a probabilistic model, the response surface \( \hat{f}(\cdot) \) in Equation (5) is a probability distribution. Therefore, the response surface \( \hat{f} \) cannot be optimized directly. Instead, an acquisition function \( \alpha(\cdot) \) is necessary for the virtual optimization of the probabilistic GP. The purpose of

Algorithm 1: Bayesian optimization

1: \( T \leftarrow \{ \theta, f(\theta) \} \) if available
2: Prior \( \leftarrow \) if available: Prior of the GP model
3: while optimize do
4: \( \text{Train GP model from } T \)
5: \( \text{Compute response surface } \hat{f}(\cdot) \)
6: \( \text{Compute acquisition surface } \alpha(\cdot) \)
7: \( \text{Find } \theta^* \text{ that optimizes } \alpha(\cdot) \)
8: \( \text{Evaluate } f \text{ at } \theta^* \)
9: \( \text{Add } \{\theta^*, f(\theta^*)\} \text{ to } T \)
the acquisition function is two-fold: 1) It maps the GP onto a single surface called the acquisition surface \( \alpha(\theta) \), which is subsequently optimized. Thereby, the minimization of the objective function from Equation (1) can be rephrased as the minimization of the acquisition surface
\[
\min_{\theta} \alpha(\theta)
\]  

(6)

2) The GP expresses model uncertainty, which is used to trade off exploration and exploitation in the optimization. In Bayesian optimization, as summarized in Algorithm 1, a GP model \( \theta \rightarrow f(\theta) \) is learned from the data set \( T = \{\theta, f(\theta)\} \) composed by the parameters \( \theta \) and the corresponding measurements \( f(\theta) \) of the objective function (Line 3 of Algorithm 1). This model is used to predict the response surface \( \hat{f}(\cdot) \) (Line 4 of Algorithm 1) and the corresponding acquisition surface \( \alpha(\theta) \) (Line 5 of Algorithm 1), once the response surface \( f(\cdot) \) is mapped through the acquisition function \( \alpha(\cdot) \). Using a global optimizer, the minimum \( \theta^* \) of the acquisition surface \( \alpha(\theta) \) is computed (Line 6 of Algorithm 1) without any evaluation of the objective function, e.g., no robot interaction is required, see Equation (6). This minimum \( \theta^* \) is evaluated on the robot (Line 7 of Algorithm 1), as well as the prior of the GP model (Line 2 of Algorithm 1).

1) Gaussian Process Model for Objective Function: To create the model that maps \( \theta \rightarrow f(\theta) \), we make use of Bayesian non-parametric Gaussian Process regression [22]. Such a GP is a distribution over functions
\[
f \sim \text{GP}(m_f, k_f)
\]

(7)

fully defined by a prior mean \( m_f \) and a covariance function \( k_f \). As prior mean, we choose \( m_f \equiv 0 \), while the chosen covariance function \( k_f \) is the squared exponential with automatic relevance determination and Gaussian noise
\[
k_f(\theta_p, \theta_q) = \sigma_f^2 \exp\left(-\frac{1}{2}(\theta_p - \theta_q)^T \Lambda^{-1}(\theta_p - \theta_q)\right) + \sigma_w^2 \delta_{pq}
\]

(8)

with \( \Lambda = \text{diag}(l_1^2, \ldots, l_D^2) \). Here, \( l_i \) are the characteristic length-scales, \( \sigma_f^2 \) is the variance of the latent function \( f(\cdot) \) and \( \sigma_w^2 \) is the noise variance. The GP predictive distribution is
\[
p(f(\theta)|T, \theta) = \mathcal{N}(\mu(\theta), \sigma(\theta)^2)
\]

(9)

where the mean \( \mu(\theta) \) and the variance \( \sigma(\theta) \) are respectively computed as
\[
\mu(\theta) = k_x^T K^{-1} y, \quad \sigma(\theta) = k_{xx} - k_x^T K^{-1} k_x.
\]

(10)

Given \( n \) training inputs \( \mathbf{X} = [\theta_1, \ldots, \theta_n] \) and corresponding training targets \( \mathbf{y} = [f(\theta_1), \ldots, f(\theta_n)] \), we define the training data set \( \mathbf{T} = \{\mathbf{X}, \mathbf{y}\} \). Moreover, \( K \) is the matrix composed of \( K_{ij} = k(\theta_i, \theta_j) \), \( k_{xx} = k(\theta, \theta) \) and \( k_x = k(\mathbf{X}, \cdot) \).

A practical issue, for both GP modeling and Bayesian optimization, is the choice of the hyperparameters of the GP model. The hyperparameters of a GP model are the parameters of the covariance function, in our case, the characteristic length-scales \( l_i \), the variance of the latent function \( \sigma_f^2 \) and the noise variance \( \sigma_w^2 \). In gait optimization, these hyperparameters are often fixed a priori [18]. There are suggestions [17] that fixing the hyperparameters can considerably speed up the convergence of Bayesian optimization. However, manually choosing the value of the hyperparameters requires extensive expert knowledge about the system that we want to optimize, which is often an unrealistic assumption. Additionally, it is unclear if fixing the hyperparameters provide an advantage that balances the risk of incorrectly approximate the real hyperparameters. An alternative common approach, which we employ in our experiments, is to automatically select the hyper-parameters by optimizing with respect to the marginal likelihood [22].

2) Acquisition Function: A number of acquisition functions \( \alpha(\cdot) \) exists, such as probability of improvement [16], expected improvement [19], upper confidence bound [5] and entropy-based improvements [13]. Experimental results [13] suggest that expected improvement on specific families of artificial functions performs better on average than probability of improvement and upper confidence bound. However, these results do not necessarily hold true for real-world problems such as gait optimization, where the objective functions are more complex to model. Probability of improvement [18], expected improvement [28] and upper confidence bound [5] have all been previously employed in gait optimization. However, no experimental comparison has been carried out in gait optimization yet, and it is still unclear why one of them should be chosen over the others.

a) Probability of Improvement (PI): Introduced by Kushner [16], the acquisition function PI is defined as
\[
\alpha(\theta) = \Phi\left(\frac{T - \mu(\theta)}{\sigma(\theta)}\right)
\]

(11)

where \( \Phi(\cdot) \) is the normal cumulative distribution function and \( T \) the target value. The target value \( T \) is often the minimum of all explored data plus, optionally, a positive constant (for a study of its effects, see [17]). PI is a function bounded by the interval \([0, 1]\). Hence, since the normal cumulative distribution function is monotonically increasing, to minimize PI is sufficient to minimize
\[
\alpha(\theta) = \frac{T - \mu(\theta)}{\sigma(\theta)}
\]

(12)

Intuitively, PI computes the probability (cumulative distribution) of the response surface in \( \theta \) to be better than the target value \( T \).

b) Expected Improvement (EI): Mockus [19] introduced EI, which can be seen as an extension of probability...
EI is formulated as

$$\alpha(\theta) = \sigma[u\Phi(u) + \phi(u)]; \quad u = \frac{T - \mu(\theta)}{\sigma(\theta)},$$

(13)

where $\phi(.)$ is the normal probability density function.

c) Upper Confidence Bound (UCB): UCB [8] is defined as

$$\alpha(\theta) = \mu(\theta) - \kappa \sigma(\theta).$$

(14)

The choice of the free parameter $\kappa$ is crucial as it determines the trade-off between exploration and exploitation. GP-UCB [26] suggests to automatically select $\kappa$ according to

$$\kappa = \sqrt{2 \log \left( \frac{n^{D/2 + 2} \pi^2}{3 \delta} \right)},$$

(15)

where $\delta$ is the number of past evaluations of the objective function $f$ and $D$ the dimensionality of the parameters $\theta$. Additionally, automatically selecting $\kappa$ allows to estimate regret bounds [26].

3) Optimizing the Acquisition Surface: Once the acquisition surface $\alpha(\theta)$ is computed (Line 5 of Algorithm 1), it is still necessary to find the parameters $\theta^*$ of its minimum (Line 6 of Algorithm 1). To find this minimum, we use a standard global optimizer. Note that the global optimization problem in Equation (6) is different from the original global optimization problem defined in Equation (1). First, the measurements in Equation (6) are noise free because the objective function in Equation (14) is an analytical model. Second, there is no restriction in terms of how many evaluations we can perform. Evaluating the acquisition surface only requires interactions with the model, but not with a physical system, such as a robot. Third, we can compute the derivatives of any order, either with finite differences or analytically. Therefore, we are no longer restricted to zero-order optimization methods, and any global optimizer can be used. In our experiments we used DIRECT [15] to find the approximate global minimum, followed by L-BFGS [4] to refine it.

III. COMPARATIVE EVALUATION ON THE Fox ROBOT

In this section, we first introduce the bio-inspired dynamical bipedal walker Fox used as our robotic evaluation platform. Afterward, we describe the experimental set-up and give details for each of the optimization methods employed. Finally, we present the comparison of the experimental evaluations and discuss it.

A. Experimental Set-up

To validate our Bayesian gait optimization approach we used the 2-D dynamic walker Fox, shown in Figure 1. This robot consists of a rudimentary trunk, two legs, made of a rigid segment connected by a knee joint to a telescopic leg spring, and two spherical feet with touch sensors [23]. Fox is equipped with low-cost metal-gear DC motors at both hip and knee joints. Together they drive four actuated degrees of freedom. Moreover, there are six sensors on the robot: two on the hip joints, two on the knee joints, and one under each foot. The sensors on the hip and knee joints return voltage measurements corresponding to angular positions of the leg segments. The touch sensor under each foot returns binary ground contact signals. The walker is mounted on a boom that enforces planar, circular motion. An additional sensor in the boom measures the angular position of the walker, i.e., the position of the walker on the circle.

The controller of the walker is a finite state machine (FSM) with four states (see Figure 3): two for the swing phases of each leg [24]. These states control the actions performed by each of the four actuators, which were extension, flexion or holding of the joint. The transition between states is regulated by the contact sensors and thresholds based on the angles of the joints. For the optimization process, we identified four parameters of the controller crucial for the resulting gait: More parameters would make a comparison with grid search unfeasible, due to the excessive number of evaluations required. These four gait parameters consist of thresholds values of the FSM (two for each leg), and directly influence the position of the hips. The remaining parameters were chosen such that only a small set of parameters space would lead to a properly walking gait. To achieve a stable gait in this parameters space it was necessary to make use of the inertia of the legs to compensate the low torques applied.

B. Performance Metric

A common metric used in gait optimization is the average walking velocity $\bar{v} = \Delta x/\Delta t$, with $\Delta x$ being the walked distance during the time $\Delta t$. However, this metric can be misleading since fast gaits can also be sensitive to intrinsic perturbations that lead to falls. Therefore, the objective function $f$ to be minimized was defined as

$$f(\theta) = -\frac{1}{N} \sum_{i=1}^{N} \tilde{v}_i(\theta),$$

(16)
Contact with Left Foot Contact with Right Foot

Values of the objective function in Figure 4a. The optimization process of GP-UCB is limited between 6 and 9 hours of hands-on experimental time.

(i.e., uniformly randomly sampled sets of parameters), thus, we used the first three evaluations from pure random search for a total of 81 evaluations (from Equation (2)). For grid search optimization, we used 3 evaluations along each of the four dimensions for a total of 64 evaluations (from Equation (2)).

Experimental Results

The maximum walking speed of Fox evaluated during the gait optimization process for the different methods is shown in Figure 4a. The optimization process of GP-UCB is limited to 57 evaluations due to a mechanical failure that forcefully interrupted the experiment. Values of the objective function below 0.1 m/s indicate that the robot fell down after a single step. Values between 0.1 and 0.15 m/s indicate that the robot could walk for multiple steps but showed systematic falls. Between 0.15 m/s and 0.25 m/s only occasional falls occurred. Above 0.25 m/s the achieved gait was robust and did not present any fall. From the results, we see that both grid search and random search performed poorly, finding a maximum that can only barely walk. We notice that Bayesian optimization using any of the different acquisition functions performed considerably better. Bayesian optimization using PI and GP-UCB achieved robust gaits with similar walking speed, while GP-UCB being slightly faster in finding the maximum. On the other hand, Bayesian optimization using EI did not lead to robust gaits. The reason of this result were the inaccuracies of the model of the underlying objective function. In fact, the automatically selected hyperparameters had overly long length-scales (see Equation (8)), which resulted in an inappropriate model and, thus, evaluating parameters of little interest.

This result is unexpected as EI is considered a versatile acquisition function, and there are experimental results [13], which suggest that EI on specific families of artificial functions performs better than GP-UCB and PI. We speculate that these results do not necessarily apply to complex real-world objective functions, such as the one we optimized. For example, if the objective function does not possess a single natural length-scale, but multiple length-scales. In particular, during the empirical evaluation, we observed that EI behaved excessively greedily, exploring the parameters space insufficiently. In turn, this insufficient exploration resulted in overly long length-scales and an inappropriate GP model. Therefore, we hypothesize that in case of an objective function which is complicated to model, the EI acquisition function might not perform as well as for easier artificial functions.

As a second comparison, we studied the effects of manually fixing hyperparameters. Thereby, we manually fixed the hyperparameters of the GP models to reasonable values, based on our expert knowledge. Figure 4b shows the performance of the different acquisition functions when using the manually fixed hyperparameters. From these results, it can be observed that all the different acquisition functions, when using the fixed hyperparameters, found similar sub-optimal solutions. The reason is that for all three acquisition functions with fixed hyperparameters, one parameter reached only a sub-optimal value. This observation suggests that, at least for that one parameter, the wrong length-scales prevented the creation of an accurate model and, therefore, the optimization process was hindered. As a confirmation of this hypothesis, using all the evaluations performed, we trained a GP model and automatically selected the hyperparameters using the marginal likelihood. The resulting values of the hyperparameters were, in some cases, half of the manually selected values, therefore, suggesting that the chosen length-scales were a rough approximation of the real ones. Both GP-UCB and PI using fixed hyperparameters performed worse than the respective cases with automatic hyperparameters selection because the longer length-scales limited the exploration of these acquisition functions. In contrast, for EI
The use of fixed hyperparameters was beneficial. In fact, the fixed length-scales were smaller than the automatically selected ones and, therefore, more exploration was performed. The hyperparameters of the GP model directly influence the amount of exploration performed by the acquisition functions. Hence, fixing the hyperparameters using expert knowledge can be an attractive choice, since forcing the right amount of exploration can speed up the optimization process. However, the presented experimental results also show that a poor choice of hyperparameters can potentially harm the optimization process by limiting the exploration and leading to sub-optimal solutions.

In Table II, we present the comparison of the optimum found by all the evaluated methods. Variations of up to 0.04 m/s can depend on the presence of noise in the experiments. Additionally, it should be noticed that the real noise of the objective function is not Gaussian. In fact, configurations of the parameters that produce periodic falls after a single step or stable gaits, typically behave consistently across various experiments and, therefore, present smaller noise (0.01 m/s). For parameters that produce unstable gaits with occasional falls the noise can be larger, typically up to 0.04 m/s. A visual representation between two of the parameters of the parameter space as predicted using the data collected from all the over 1800 evaluations is shown in Figure 5. It can be noticed that this space is complex and non-convex, and, therefore, unsuitable for local optimization methods (e.g., gradient-based methods).

The GP modeling capabilities are often overlooked when evaluating Bayesian optimization’s performances, with all the emphasis on the use of different acquisition functions. Following the results of our experimental evaluation, we speculate that for complex objective functions there exists a strict and yet unexplored connection between the exploration properties of the acquisition function and the capabilities of GP modeling. The performance of an acquisition function depends on the capabilities of properly modeling the function, and vice-versa a proper modeling take place only when the acquisition function evaluate relevant parameters.

### Table II: Maximum average walking speeds [m/s] found by the different optimization methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Maximum during optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid search</td>
<td>0.148</td>
</tr>
<tr>
<td>Pure random search</td>
<td>0.142</td>
</tr>
<tr>
<td>BO: PI</td>
<td>0.328</td>
</tr>
<tr>
<td>BO: EI</td>
<td>0.232</td>
</tr>
<tr>
<td>BO: GP-UCB</td>
<td><strong>0.337</strong></td>
</tr>
<tr>
<td>BO: PI (fixed Hyp)</td>
<td>0.254</td>
</tr>
<tr>
<td>BO: EI (fixed Hyp)</td>
<td>0.266</td>
</tr>
<tr>
<td>BO: GP-UCB (fixed Hyp)</td>
<td>0.255</td>
</tr>
</tbody>
</table>

The maximum obtained using PI is qualitatively similar to 0.337 m/s. A visual representation between two of the parameters of the parameter space as predicted using the data collected from all the over 1800 evaluations is shown in Figure 5. It can be noticed that this space is complex and non-convex, and, therefore, unsuitable for local optimization methods (e.g., gradient-based methods).

The GP modeling capabilities are often overlooked when evaluating Bayesian optimization’s performances, with all the emphasis on the use of different acquisition functions. Following the results of our experimental evaluation, we speculate that for complex objective functions there exists a strict and yet unexplored connection between the exploration properties of the acquisition function and the capabilities of GP modeling. The performance of an acquisition function depends on the capabilities of properly modeling the function, and vice-versa a proper modeling take place only when the acquisition function evaluate relevant parameters.

### IV. Discussion & Conclusion

Gait optimization is a key research topic relevant for efficient bipedal locomotion. However, due to the presence of ad-hoc solutions and different hardware, the proposed optimization methods had not yet been thoroughly compared to other methods. In this paper, we presented a survey of some gait optimization methods and discussed their strengths and limitations. To compare the different gait optimization methods, we performed over 1800 evaluations on a real bipedal walker. Experimental results show that Bayesian optimization performed considerably better than grid or random search. In the context of Bayesian optimization, we firstly compared different acquisition functions. While GP-UCB had the best performances, EI greedily chose the parameters to evaluate and therefore performed poorly. Secondly, we compared the manually fixing hyperparameters against automatically selecting them. The results showed that manually fixing the hyperparameters can strongly influence the outcome of the optimization process. These experimental findings emphasize the importance of proper modeling and the exploration-exploitation trade-off in Bayesian optimization.
The presence of multiple local minima motivate the needs of using global optimization methods.

Bayesian optimization is a promising method for efficient optimization, especially in fields like locomotion where only few evaluations can be performed before wearing out the hardware. Future research in Bayesian optimization should focus on obtaining more accurate and appropriate response surface models. Since the experimental noise is often not constant, a logical extension would be the use of heteroscedastic GP models. Additionally, further studies are required to fully understand the joint influence of different acquisition functions and models, in the case of hard to model complex objective functions.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007–2013) under grant agreements #270327 (CompLACS) and #600716 (CoDyCo) and from the Department of Computing, Imperial College London.

REFERENCES