# Probabilistic Validation of a Stochastic Kinematic Model for an Eight-legged Robot

Konstantinos Karydis, Ioannis Poulakakis and Herbert G. Tanner

*Abstract*— **The paper suggests a new method for statistically validating, and selecting the parameters of a model for a miniature eight-legged robot. It is based on a novel adaptation of concepts and techniques originally developed in the context of robust control design using randomized algorithms. The proposed approach is data driven and offers probabilistic guarantees of model fidelity and descriptive capacity, checking models against experimental data. In principle, this method applies to a large class of physical processes, the available models of which may be in a variety of forms including sets of differential equations.**

## I. INTRODUCTION

The OctoRoACH depicted in Fig. 1 represents an instance of a paradigm for the design of bio-inspired, expendable, small-scale sensor platforms [1]. It is a highly-articulated mechanism produced through a relatively low-precision manufacturing process, and made of material with uncertain mechanical properties. As such, it resists accurate characterizations in the form of a deterministic dynamical model. On the other hand, state-of-the-art control methodologies for navigation, motion planning and cooperative behavior assume the availability of some low-dimensional model, typically in the form of a differential equation. However, the lack of knowledge of the actual dynamics of this robot, combined with the apparent stochastic behavior suggested by available experimental data, hinder the development of such task-specific controllers.

The complexity of this miniature robot, together with the observed variability between different available prototypes, render a direct approach to modeling based on first principles impractical, and raise questions about the accuracy and generality of the model. Even when the components of the design are represented in a simplified way (e.g., the torso as a single rigid body) attempts to construct accurate, three-dimensional, numerical simulations have had limited success [2]; instead, only two-dimensional sagittal plane MATLAB[TM] models have been developed. On the other hand, motion planning methods for multi-legged vehicles such as the RHex [3], assume[1] a unicycle model [5], and, similarly to the large body of work on wheeled platforms, apply it to describe the horizontal-plane motion.

A variety of bio-inspired models has been proposed in the relevant literature to capture the salient features of

[1]See [4] for a more formal reduction of the three-dimensional RHex dynamics into a two-dimensional unicycle.
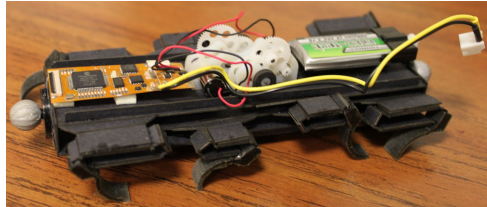


Fig. 1.   The OctoRoACH robot, designed at the University of California, Berkeley as part of the MAST CTA.

legged systems as they move through their environment [6]. Among them, spring-mass systems, including the Spring Loaded Inverted Pendulum (SLIP) [7], [8] and the Lateral Leg Spring (LLS) [9], [10] models have demonstrated the ability to predict the center-of-mass dynamics in the sagittal and horizontal planes of fast-moving robots (and animals) of various structures and morphologies. These models are dynamic, involving forces and accelerations, and do not fit well within the framework of many motion planning methods, which typically assume kinematic models. In order to make these planning tools relevant to legged platforms of small size, such as the OctoRoACH, we propose low-dimensional kinematic abstractions that are based on available experimental data.

From a general perspective, extracting meaningful system representations of physical processes based on input-output measurements lies within the field of system identification, and several methods are available to this end. For example, one method is to use available data to obtain a Linear Time Invariant (LTI) system model [11]. If nonlinear models are desired, then randomized methods can be employed to select nonlinear terms in ODE models as in [12]. More traditional methods, but perhaps less amenable to analysis, include Volterra series [13]–[15], or neural networks [16]–[18].

Although such approaches can in principle be applied in the context of the problem considered here, their potential in deriving suitable model representations is limited for a number of reasons. In more detail, linear system models, such as those in [11], represent local approximations of the dynamics around specific operating points. On the other hand, Volterra series and neural network representations lack the clarity and compactness of ODE models. Most important, none of these approaches can offer models that admit physically meaningful interpretations as in models derived based on first principles.

In our previous work [19], we introduced a reduced-order nonlinear abstraction for approximating the kinematics of

the OctoRoACH on the horizontal plane. This model reproduces straight-line and circular paths that resemble those observed in experiments with the OctoRoACH. Furthermore, it permits physical interpretations of its parameters, and couples them in an intuitive fashion to the motion of the robot. In that work, however, no formal justification—only mechanical intuition—has been provided in support of the proposed model. Furthermore, the approach in [19] cannot accommodate random effects that inevitably influence the robot's motion, and are further amplified due to its small size and low manufacturing cost.

The present paper attempts to fill the perceived gap between the physical process realized by the actual mechanism of Fig. 1 when in straight-line motion, and its suggested kinematic abstraction studied in [19]. Our approach is data driven in the sense that it does not depend on the specific type of system representation; rather, it depends on the experimental data of the physical process being modeled. We borrow concepts from the fields of robustness analysis [20], [21] and control synthesis [22]–[24] using randomized methods, and we adapt them to data-driven *model validation* and *parameter estimation*. These are instances of a more general problem, where the descriptive capacity and the fidelity of a model need to be quantified based on data.

Generally, randomized algorithms have long been employed to tackle several control problems where analytic solutions are impractical [25], [26], and, given sufficient structure [27], have found many applications beyond control systems [22], [28]. The underlying mathematical foundation for these methods is provided by statistical learning theory [27], [29]–[31]. Early mathematical results [32] are brought to bear [22], [23], [33] to determine a sufficient number of samples in a Monte Carlo approach to estimate empirical means, which in most cases—although with some important caveats [21], [33]—yields fairly reasonable sample sizes. This same Monte Carlo approach is used here to estimate the probability that a particular model instantiation satisfies the performance specifications.

Specifically, our approach hinges on checking stochastically parameterized models against a large body of experimental data. A Monte Carlo simulation is used to estimate the probability that a random instantiation of the model—that is, a solution produced by the model by randomly selecting the values of its parameters from a given distribution—is statistically consistent with our specifications, given available experimental data. As it turns out, through this process, we can tune the model's parameters to some nominal values so that the behavior of the model captures the one observed in experiments as closely as possible, while the probability of violating the specifications remains below a given threshold.

The rest of the paper is organized as follows. Section II first gives an overview of the model developed in [19], and then suggests nominal values for its parameters, derived based on experimental data collected for the case of straight-line motion. Section III extends the deterministic model to a stochastic one by introducing uncertainty into the robot's touchdown and liftoff angles. Section IV

presents an approach to probabilistically validate models against experimental data of the process they are supposed to describe, and Section V presents the results of the application of the proposed method on the stochastic model of the OctoRoACH. Section VI concludes the paper.

## II. A MODEL FOR THE OCTORoACH

### A. A Deterministic Kinematic Model

The OctoRoACH consists of a rectangular body with eight legs, designed in a way that ipsilateral[2] legs are driven by a single motor. The robot implements a gait that is composed of two alternating tetrapods. In particular, the legs $\{1, 2, 3, 4\}$ act in unison and form one tetrapod, and similarly the legs $\{5, 6, 7, 8\}$ form the other tetrapod, as shown in Fig. 2.

This mechanically induced synergy among the robot's legs motivated the kinematic approximation depicted in Fig. 3(a) which can be further reduced to the switching four-bar mechanism shown in Fig. 3(b). Note that, despite its simple kinematic nature, this model is capable of reproducing the planar straight-line motion of the robot in its nominal alternating tetrapod gait and it captures the meandering motion of the geometric center $G$, see Fig. 3, as it travels along a straight line parallel to the $y$ axis. More details about this reduced-order model of the OctoRoACH can be found in our previous work [19].

Average straight-line motion along the $y$ direction is produced assuming a 50% duty cycle for the two alternating

[2]Terminology: Ipsilateral means on the same side and contralateral means on the other side.
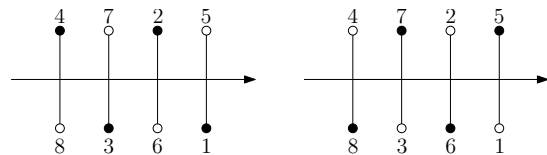


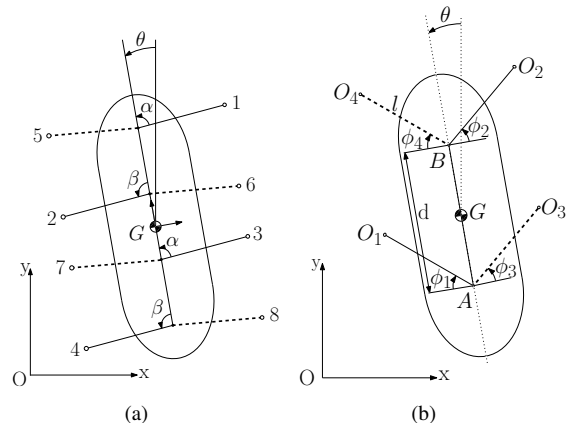Fig. 2.    Foot fall pattern of the OctoRoACH robot



Fig. 3.    (a) An eight-legged model for the OctoRoACH. By design, the legs are coupled in pairs, as indicated by the angles $\alpha$ and $\beta$. Thus, they form two tetrapods, the first of which consists of the legs $\{1, 2, 3, 4\}$ while the second contains $\{5, 6, 7, 8\}$. (b) The proposed abstraction that serves as our deterministic model. Due to the coupling, each tetrapod is reduced to a pair of legs – the tetrapod $\{1, 2, 3, 4\}$ corresponds to the legs $\{O_1, O_2\}$ (right pair) and the tetrapod $\{5, 6, 7, 8\}$ to the legs $\{O_3, O_4\}$ (left pair). When in contact with the ground, each pair forms a four-bar mechanism.

pairs of virtual legs, named "right" and "left," marked with solid and dashed lines in Fig. 3(b), respectively. Each abstract pair corresponds to one tetrapod, and because of the symmetry resulting from the equal duty cycle assumption, it suffices to analyze one of the two. The same assumption suggests that only a single pair is in contact with the ground at any given time instant.

Figure 4 presents the kinematic analysis of the right pair. With the notation of Fig. 4, the vector-loop equation [34],

$$\boldsymbol{R}_{AO_1} + \boldsymbol{R}_{O_1O_2} - \boldsymbol{R}_{AB} - \boldsymbol{R}_{BO_2} = 0 \tag{1}$$

can be expressed in complex number notation as

$$le^{j(\pi-\phi_1)} + a_R e^{j(q_1-\phi_1)} - de^{j(\pi/2)} - le^{j(\phi_2)} = 0 \tag{2}$$

which is equivalent to

$$\begin{aligned} -l\cos(\phi_1) + a_R\cos(q_1-\phi_1) - l\cos(\phi_2) = 0 \\ l\sin(\phi_1) + a_R\sin(q_1-\phi_1) - d - l\sin(\phi_2) = 0 \ . \end{aligned} \tag{3}$$

Differentiation of (2) results in

$$\begin{aligned} l\cos(\phi_1)\dot{\phi}_1 + a_R\cos(q_1-\phi_1)(\dot{q}_1-\dot{\phi}_1) - l\cos(\phi_2)\dot{\phi}_2 = 0 \\ l\sin(\phi_1)\dot{\phi}_1 - a_R\sin(q_1-\phi_1)(\dot{q}_1-\dot{\phi}_1) + l\sin(\phi_2)\dot{\phi}_2 = 0 \ . \end{aligned} \tag{4}$$

Note that the system has one degree of freedom, and its configuration is fully specified by numerically solving (3) for $q_1$ and $\phi_2$ given a value of $\phi_1$. The velocity of the platform can be obtained by solving (4) in a similar fashion; no numerical integration is needed.

The left pair is analyzed in a completely analogous way. In the place of $\phi_1, \phi_2, q_1, \dot{\phi}_1, \dot{\phi}_2, \dot{q}_1$ and $a_R$, we now have $\phi_3, \phi_4, q_3, \dot{\phi}_3, \dot{\phi}_4, \dot{q}_3$, and $a_L$, respectively. The resulting equations have exactly the same form.

Both legs in each pair are thought to touch and lift off the ground at specific angles, which we call *touchdown* and *liftoff* angles and are denoted by $\phi_i^{\text{td}}$ and $\phi_i^{\text{lo}}$, $i = 1, \ldots, 4$, respectively. A *sweep* angle $\psi_i$ is defined as the difference between the maximum and the minimum value for each $\phi_i$, that is,

$$\psi_i = |\phi_i^{\text{td}}| + |\phi_i^{\text{lo}}|, \tag{5}$$

and thus quantifies the range of values for each $\phi_i$. Although a variety of different combinations of touchdown, liftoff, and
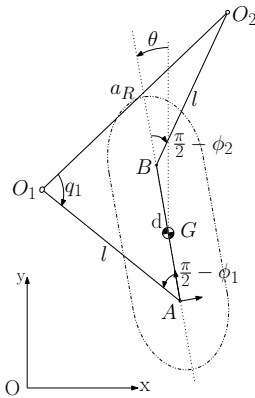


Fig. 4.  Analysis of the mechanism. $d$ is the distance between the hip points and $l$ the leg length. The angle $\theta$ is used to denote the orientation of the model with respect to some global frame of reference.

sweep angles can be used to produce very similar paths, it is not clear at this stage which particular combination enables the model to match better the available collection of experimental data. As ongoing work indicates, these parameter values are crucial in determining the nature and characteristics (path curvature) of motion paths produced by the model.

*B. Parameter Identification from Experimental Data*

We generate a collection of planar position measurement data, through experiments in which the robot is configured to move (ideally) along a straight line. The data is obtained with the use of a VICON$^{\text{TM}}$ motion capture system at a rate of 50Hz. The inputs to the robot are the gains of the two motors that actuate its legs, one for each side, and are both set equal to 40. Each run had a duration of 10 cycles, and we collected data from a total of 500 paths. The measured states are the planar position of the geometric center of the robot and its orientation. Tabulating measured states is impractical due to the size of the data, and for illustration purposes Fig. 5(a) depicts a random sample of 30 paths, while Fig. 5(b) shows average statistics for the whole collection.

All experiments were conducted on a rubber floor mat surface; in future work we plan to explore other terrain types as well. The initial position and orientation of the robot were manually set inside a designated area. The initial pose errors that are produced through this inexact procedure can be bounded using data statistics; we calculate the sample mean and standard deviation for the initial position along the x and y directions, as well as the same statistics for the initial orientation. These results are summarized in Table I. Furthermore, notice that these error bounds accommodate for measurement noise as well.

TABLE I
INITIAL POSE ERROR STATISTICS

| Measurement | Mean | Variance |
|---|---|---|
| Position along x-axis [cm]: | 0.053 | 0.083 |
| Position along y-axis [cm]: | 0.286 | 0.153 |
| Orientation [deg]: | 79.66 | 20.30 |

The model parameters to be estimated are included in

$$\zeta = [l, \phi_1^{\text{td,n}}, \phi_2^{\text{td,n}}, \phi_3^{\text{td,n}}, \phi_4^{\text{td,n}}], \tag{6}$$

where the superscript $^{\text{n}}$ stands for "nominal," while $^{\text{td}}$ is used to indicate that these angles are the tetrapod's touchdown angles. The domain of the parameter vector $\zeta$ is a subset of $\mathbb{R}_+ \times S^4$. One may choose to identify a larger set of model parameters—including, for instance, the distance $d$ between the hip points; this particular set is however sufficient to give the model granularity needed. The value of the parameter $d$ is set to be equal to 13 cm, which is the length of the torso of the actual platform. The liftoff angles were assumed to be

$$\phi_R^{\text{lo,n}} = -\min\{\phi_1^{\text{td,n}}, \phi_2^{\text{td,n}}\} \tag{7}$$

and

$$\phi_L^{\text{lo,n}} = -\min\{\phi_3^{\text{td,n}}, \phi_4^{\text{td,n}}\}, \tag{8}$$
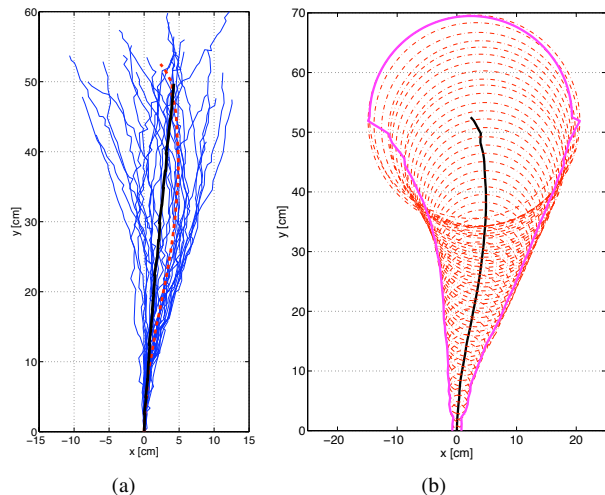
(a)                                (b)

Fig. 5. Experimental results for a duration of 10 cycles in mid-low speed. (a) 30 randomly selected paths of the platform. The red dashed curve is the average out of the whole data set of 500 observed paths (compare with the solid black curve on 5(b)). The black thick solid line corresponds to the output of the model using values for its parameters from Table II. (b) The statistics of the complete data set with 500 paths: the average is shown in the black solid curve; the dashed red circles mark a two-standard deviation region around each average position, and the solid magenta closed curve shows a confidence region at a 95% level, assuming a Student distribution. We need to design the stochastic model in such a way that its realizations cover the region subscribed.

because the duty cycle is assumed to be 50%; intuitively, all the legs that belong in the same tetrapod lift off the ground simultaneously.

Given a set of paths realized by the robot, the parameter vector $\zeta$ can then be selected through the solution of a least-squares optimization problem,

$$\min_{\zeta \in Z} \|p(\zeta) - w^{\text{ave}}\|^2 \qquad (9)$$

where $p(\zeta)$ denotes the path generated by the model corresponding to a particular value of the parameter vector $\zeta$, and $w^{\text{ave}}$ is the average path derived with data from all position trajectories observed experimentally. The solution, $\zeta^{\text{n}}$ produces the model output shown in Fig. 5(a). The numerical values of the components of $\zeta^{\text{n}}$ are given in Table II. These values allow the deterministic model to capture *on average* the behavior of the system shown in Fig. 5(a). Clearly, though, the deterministic model cannot express the variability between different experimentally observed paths.

TABLE II
NOMINAL MODEL PARAMETERS

| $d$ | $l$ | $\phi_1^{\text{td,n}}$ | $\phi_2^{\text{td,n}}$ | $\phi_3^{\text{td,n}}$ | $\phi_4^{\text{td,n}}$ | $\phi_R^{\text{lo,n}}$ | $\phi_L^{\text{lo,n}}$ |
|---|---|---|---|---|---|---|---|
| (cm) | (cm) | (deg) | (deg) | (deg) | (deg) | (deg) | (deg) |
| 13 | 2.1 | 45.3 | 39.6 | 39.9 | 47.0 | -39.6 | -39.9 |

III. INFUSING STOCHASTICITY

The position measurements indicate that individual paths can be considered as realizations of a stochastic process that describes the motion of the system. Let $\mathcal{W}$ denote the

collection of all these realizations, and $w$ an element in that collection; intuitively $w$ represents a single path of the robot. Our objective is to capture the stochastic nature of the process by modifying the model constructed in the previous section.

To achieve this, we infuse stochasticity in the touchdown and liftoff angles of each tetrapod, i.e., to the parameter array

$$\xi = [\phi_1^{\text{td}}, \phi_2^{\text{td}}, \phi_3^{\text{td}}, \phi_4^{\text{td}}, \phi_R^{\text{lo}}, \phi_L^{\text{lo}}], \qquad (10)$$

with $\xi \in \Xi$. The justification for this choice of parameters comes from the fact that the interaction of the robot's legs with the ground is inherently uncertain.

Insight from [19] suggests that certain combinations of the touchdown and liftoff angles in $\xi$ are more likely to produce straight-line paths. In our stochastic modeling, these combinations of touchdown and liftoff angles define the average value—i.e., the peak—of a multivariate normal[3] distribution. Hence, we assume that $\xi$ is a realization of

$$\tilde{\xi} = \eta^n + \begin{bmatrix} \sigma_1 \mathrm{I}_{4\times4} & \mathbf{0} \\ \mathbf{0} & \sigma_2 \mathrm{I}_{2\times2} \end{bmatrix} \cdot Y, \qquad (11)$$

where $Y$ is a six-dimensional random vector with statistically independent components, each drawn from a normal distribution of zero mean and unit variance, and

$$\eta^n = [\phi_1^{\text{td,n}}, \phi_2^{\text{td,n}}, \phi_3^{\text{td,n}}, \phi_4^{\text{td,n}}, \phi_R^{\text{lo,n}}, \phi_L^{\text{lo,n}}] \qquad (12)$$

is a constant average that contains the values of the touchdown and liftoff angles that produce straight-line paths with higher probability; these values are taken from Table II. In (11), $\sigma_1^2, \sigma_2^2$ correspond to the variances of the touchdown and liftoff angles, respectively. It is emphasized here that the variances $\sigma_1^2, \sigma_2^2$ are the only unknown parameters in our stochastic model. The following section shows how these values can be computed through a process that validates the combined model (3)–(11) against the experimental data.

IV. PROBABILISTIC MODEL VALIDATION

In the context of this paper, a model is valid when it is expressive enough to produce the whole range of behaviors observed in experiments, and its outputs follow the statistics of these behaviors at a given confidence level. In other words, we are interested in verifying whether the paths produced by (3)–(11) can cover the region outlined in Fig. 5(b), with the probability of a path produced by the model falling outside a desired confidence region being below a certain specification threshold. This threshold should not be too small since this would suggest that the model is too conservative.

In our analysis, whether a model instantiation[4] is representative of the behavior of the robot is captured by a binary decision function $g : \Xi \times \mathcal{W} \to \{0, 1\}$. Practically, this function accepts an instantiation of the model parameterized by some

---

[3]To provide more intuition, note if we had assumed a uniform distribution, that would have meant that *any* combination of parameters $\xi$ is equally likely to produce a straight-line motion, which is clearly not true.

[4]Here, the term "model instantiation" refers to the computation of a solution of our parameterized model at a randomly selected value of the parameter vector according to the normal distribution (11).

realization $\xi$ of (11), and a collection of elements from $\mathcal{W}$ which is called *multisample*, $\boldsymbol{w}_k = \{w^{(1)}, w^{(2)}, \ldots, w^{(k)}\}$, $k \in \mathbb{N}_+$, and returns 0 only if the output of the model is within a prespecified confidence interval at level $\gamma \in (0, 1)$, evaluated based on the experimental data collected. For $k < 30$, we assume a Student distribution for the distribution of paths within a multisample, and take $\gamma = 0.95$. The outline of the sketched region in Fig. 5(b) represents the boundaries within which a path produced by a given model instantiation would result in $g(\xi, \boldsymbol{w}_k) = 0$.

To maximally cover the support of $1 - g$, we attempt to *maximize* the random component in (11) to the extent that the probability that the decision function $g$ reports 1 for a random instantiation of model parameters is less or equal to a positive constant $\rho$. Minimizing $\sigma_1$, $\sigma_2$ instead would certainly meet this later requirement, but would produce an overly conservative model with paths around the known average, and without the ability to fully express the variability we observe in the data. In that case, the outcome would be the already known deterministic model (3)–(4).

The probability that $g = 1$ for a random model instantiation is described in the context of randomized robust control synthesis [21], [24] as the *probability of violation* of a specification constraint. Here, we adapt the definition of this probability to the context of the problem as follows. Given $\sigma_1, \sigma_2$ in (11), then for a parameter vector $\xi \in \Xi$ generated accordingly and multisamples $\boldsymbol{w}_k = \{w^{(1)}, w^{(2)}, \ldots, w^{(k)}\}$ of size $k \in \mathbb{N}_+$ drawn from a collection of experimental data $\mathcal{W}$, the probability of violation for the decision function $g : \Xi \times \mathcal{W} \to \{0, 1\}$ is defined as

$$E_k(\sigma_1, \sigma_2) \triangleq \mathrm{Pr}_{\mathcal{W}}\{w \in \mathcal{W} : g(\xi, \boldsymbol{w}_k) = 1, \ \xi \in \Xi\} \ , \quad (13)$$

where $\mathrm{Pr}_{\mathcal{W}}$ is a probability measure over $\mathcal{W}$.

In general, the probability of violation is difficult to be evaluated analytically, and one typically resorts to approximating it using an empirical average [21]; i.e.,

$$\hat{E}_k(\sigma_1, \sigma_2) = \frac{1}{n} \sum_{i=1}^{n} g(\xi_i, \boldsymbol{w}_k^{(i)}) \ , \ \ n \in \mathbb{N}_+ \ , \quad (14)$$

where $\xi_i$ are values generated by (11), and $\boldsymbol{w}_k^{(i)}$ are i.i.d. multisamples drawn from $\mathcal{W}$. Then the performance (fidelity) specification for the model is expressed using the empirical average of the probability of violation as $\hat{E}_k(\sigma_1, \sigma_2) \leq \rho$.

A constrained optimization problem can thus be formulated, the solution of which corresponds to the least conservative stochastic approximation of the physical process, for which the probability of violation of the performance specification for the model remains below $\rho$:

$$\max_{\sigma_1, \sigma_2} \hat{E}_k(\sigma_1, \sigma_2), \quad \text{subject to } \hat{E}_k(\sigma_1, \sigma_2) \leq \rho \ .$$

However, there is a natural monotonic relationship between $\sigma_1, \sigma_2$ and $E_k(\sigma_1, \sigma_2)$, since the larger the variance in (11), the higher the probability of violation. Therefore, the above problem can be reduced to

$$\max(\sigma_1 + \sigma_2), \quad \text{subject to } \hat{E}_k(\sigma_1, \sigma_2) \leq \rho \ . \quad (15)$$

Based on the same monotonic relationship between $\sigma_i$ and $E_k(\sigma_1, \sigma_2)$, the solution is expected at the (upper) boundary of the feasible set for $\sigma_i$.

We are thus seeking for the values of $\sigma_i$ for which $\hat{E}_k(\sigma_1, \sigma_2) \to \rho^-$. These can be found using a simple Monte Carlo simulation. Selecting an accuracy parameter $\varepsilon > 0$ we progressively increase $\sigma_i$ so that the empirical mean satisfies $\hat{E}_k(\sigma_1, \sigma_2) < \rho - \varepsilon$ with confidence $1 - \delta$. Based on [35], it suffices to select $n$ as follows

$$n \geq \frac{1}{2\epsilon^2} \log\left(\frac{2}{\delta}\right) \quad (16)$$

to approximate $E_k(\sigma_1, \sigma_2)$ with $\hat{E}_k(\sigma_1, \sigma_2)$ at this accuracy and confidence. Note that since each $\xi_i$ needs to be paired with a statistically *independent* multisample $\boldsymbol{w}_k$, one eventually needs to collect data from $n \cdot k$ different experiments. However, the same body of experimental data can be used for each different choice of $\sigma_i$.

## V. RESULTS

To illustrate the methodology we solve (15) for the case of the OctoRoACH with a collection $\mathcal{W}$ of experimental measurements from 500 paths. The upper bound for the probability of violation is selected at $\rho = 0.35$, and the parameters in (16) are $\epsilon = 0.15$ and $\delta = 0.27$. Then (16) suggests a sample of size $n = 45$. We thus create 45 mutually disjoint multisamples of size $k = 11$, by randomly grouping different paths in $\mathcal{W}$. Since $k < 30$, the statistics of each multisample is assumed to follow the Student distribution, and we select confidence intervals at a level $\gamma = 0.95$.

The solution of this problem suggests values of standard deviations, $\sigma_1$ and $\sigma_2$, (shown in Table III) which result in a parameterization of the model (3)–(11) that maximally covers the range of observed experimental values. These are denoted $\sigma_1^{\max}$, $\sigma_2^{\max}$, with the superscript reflecting the fact that the maximum of $\hat{E}_k(\sigma_1, \sigma_2)$ occurs for some maximum values of $\sigma_1$ and $\sigma_2$ that still keep the empirical average below the threshold $\rho$. However, the particular *pair* of values $(\sigma_1, \sigma_2)$ that maximize $\hat{E}_k$ is not obvious and this is the question that the algorithm answers.

TABLE III
PROBABLY NEAR MAXIMUM MODEL UNCERTAINTY

| $\sigma_1^{max}$ | $\sigma_2^{max}$ |
|---|---|
| 0.12 | 0.15 |

Figure 6 suggests that the model parameterized based on the values from Tables II-III produces paths that are within the desired confidence region. This test includes 45 different (random) instantiations of the model (3)–(11) with the chosen values for the parameters.

## VI. CONCLUSIONS

This work contributes to the development of a method that can be used to assess the fidelity of a model, and its capacity to capture a targeted physical process through available experimental data. We show that probabilistic guarantees
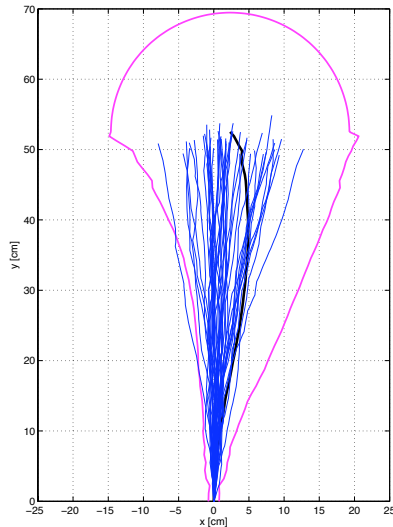
Fig. 6. Output of the stochastic model, tuned using Monte Carlo simulation. The 45 model instantiations, shown in blue, correspond to the values $\sigma_i^{\max}$ of Table III and are plotted over the average (in thick black) of the whole data collection. The 95% confidence level still allows a few paths to go outside the marked (in magenta) confidence region.

of model fidelity can be derived using techniques adapted from the field of robust control design using randomized algorithms. The proposed procedure is independent of the type and structure of the model considered, and it offers a systematic way to adjust the model parameters in order to maximize its expressiveness. This new method is applied to identify a stochastic kinematic model of the miniature robot OctoRoACH.

## REFERENCES

[1] A. Pullin, N. Kohut, D. Zarrouk, and R. Fearing, "Dynamic turning of 13 cm robot comparing tail and differential drive," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Saint Paul, MN, May 2012, pp. 5086–5093.

[2] Z. Lian, "Study of reinforcement learning methods to enable automatic tuning of state of the art legged robots," Electrical Engineering and Computer Sciences, University of California at Berkeley, Tech. Rep. UCB/EECS-2012-127, May 2012.

[3] U. Saranli, M. Bühler, and D. E. Koditschek, "RHex: A Simple and Highly Mobile Hexapod Robot," *The International Journal of Robotics Research*, vol. 20, no. 7, pp. 616–31, Jul. 2001.

[4] D. Panagou and H. Tanner, "Modeling of a Hexapod Robot; Kinematic Equivalence to a Unicycle," University of Delaware, Tech. Rep. UDMETR-2009-001, 2009.

[5] G. A. Lopes and D. E. Koditschek, "Visual Servoing for Nonholonomically Constrained Three Degree of Freedom Kinematic Systems," *The International Journal of Robotics Research*, vol. 26, no. 7, pp. 715–736, 2007.

[6] P. Holmes, R. J. Full, D. E. Koditschek, and J. Guckenheimer, "The Dynamics of Legged Locomotion: Models, Analyses, and Challenges," *SIAM Review*, vol. 48, no. 2, pp. 207–304, 2006.

[7] W. J. Schwind, "Spring Loaded Inverted Pendulum Running: A Plant Model," Ph.D. dissertation, University of Michigan, 1998.

[8] I. Poulakakis and J. W. Grizzle, "Modeling and Control of the Monopedal Running Robot Thumper," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Kobe, Japan, 2009, pp. 3327–3334.

[9] J. Schmitt and P. Holmes, "Mechanical models for insect locomotion: dynamics and stability in the horizontal plane I. Theory," *Biological Cybernetics*, vol. 83, no. 6, pp. 501–515, Nov. 2000.

[10] ——, "Mechanical models for insect locomotion: dynamics and stability in the horizontal plane - II. Application," *Biological Cybernetics*, vol. 83, no. 6, pp. 517–527, Nov. 2000.

[11] L. Ljung, "From data to model: a guided tour," in *Proceedings of the IEEE International Conference on Control*, vol. 1, Mar. 1994, pp. 422–430.

[12] M. Schmidt and H. Lipson, "Distilling free-form natural laws from experimental data," *Science*, vol. 324, pp. 81–85, Apr. 2009.

[13] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*. Malabar, FL: Krieger Publishing, 2006.

[14] W. Rugh, *Nonlinear System Theory: The Volterra/Wiener Approach*, ser. Johns Hopkins Series in Information Sciences and Systems. Baltimore, MD: Johns Hopkins University Press, 1981.

[15] T. Ogunfunmi, *Adaptive Nonlinear System Identification: The Volterra and Wiener Model Approaches*, ser. Signal and Communication Technology. New York, NY: Springer-Verlag, 2007.

[16] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Upper Saddle River, NJ: Prentice-Hall, 1999.

[17] M. Ibnkahla, "Statistical analysis of neural network modeling and identification of nonlinear systems with memory," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1508–1517, Jun. 2002.

[18] S. Billings and H.-L. Wei, "A new class of wavelet networks for nonlinear system identification," *IEEE Transactions on Neural Networks*, vol. 16, no. 4, pp. 862–873, Jul. 2005.

[19] K. Karydis, I. Poulakakis, and H. G. Tanner, "A switching kinematic model for an octapedal robot," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vilamoura, Algarve, Portugal, Oct. 2012, pp. 507–512.

[20] L. Ray and R. Stengel, "A monte Carlo approach to the analysis of control system robustness," *Automatica*, vol. 29, no. 1, pp. 229–236, Jan. 1993.

[21] T. Alamo, R. Tempo, and E. Camacho, "Randomized strategies for probabilistic solutions of uncertain feasibility and optimization problems," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2545–2559, Nov. 2009.

[22] M. Vidyasagar, *A Theory of Learning and Generalization: With Applications to Neural Networks and Control Systems*. London, UK: Springer Verlag, 1997.

[23] ——, "Statistical learning theory and randomized algorithms for control," *Control Systems Magazine*, vol. 18, no. 6, pp. 69–85, Dec. 1998.

[24] G. Calafiore, F. Dabbene, and R. Tempo, "Research on probabilistic methods for control system design," *Automatica*, vol. 47, no. 7, pp. 1279–1293, Jul. 2011.

[25] V. Blondel and J. Tsitsiklis, "NP-hardness of some linear control design problems," *SIAM Journal of Control and Optimization*, vol. 35, no. 6, pp. 2118–2127, Nov. 1997.

[26] ——, "A survey of computational complexity results in system and control," *Automatica*, vol. 36, no. 9, pp. 1249–1274, Sep. 2000.

[27] V. N. Vapnik and A. Chervonenkis, "On the uniform convergence of relative frequencies of events to their probabilities," *Theory of Probability and its Applications*, vol. 16, no. 2, pp. 264–280, 1971.

[28] E. D. Sontag, "VC dimension of neural networks," in *Neural Networks and Machine Learning*, C. Bishop, Ed. Springer, Berlin, 1998, pp. 69–95.

[29] R. Dudley, "Central limit theorems for empirical measures," *The Annals of Mathematical Statistics*, vol. 6, pp. 899–929, 1978.

[30] V. N. Vapnik, *Statistical Learning Theory*. New York, NY: Wiley, 1998.

[31] S. Mendelson, "A few notes on statistical learning theory," in *Advanced Lectures in Machine Learning*, S. Mendelson and E. J. Smola, Eds. Springer Verlag, 2003, vol. 2600, pp. 1–40.

[32] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *The Annals of Mathematical Statistics*, vol. 23, no. 4, pp. 493–507, Dec. 1952.

[33] V. Koltchinskii, C. Abdallah, M. Ariola, P. Dorato, and D. Panchenko, "Statistical learning control of uncertain systems: it is better than it seems," University of New Mexico, Tech. Rep. EECE-TR-00-001, 2000.

[34] R. Norton, *Design of Machinery*. New York, NY: McGraw Hill, 2008.

[35] M. Vidyasagar, "Randomized algorithms for robust controller synthesis using statistical learning theory," *Automatica*, vol. 37, no. 10, pp. 1515–1528, Oct. 2001.