Hardware-In-the-Loop Robust Co-Design of an Underactuated System

Raphael Stöckner[1], and Shivesh Kumar[1]

***Abstract*— Control-morphology Co-Design seeks to concurrently optimize a robot's system dynamics and control, with the purpose of exploiting the interaction between both domains. Established works use a simulation-based approach that may or may not translate well to a physical system, strongly depending on the quality of the system model. We propose a Co-Design formulation that uses the region of attraction as the optimization objective, thus yielding a system with increased (guaranteed) safety margins. In a hardware-in-the-loop approach we explicitly include hardware experiments into the optimization, thus ensuring that these guarantees hold even after the sim2real step. Using Multi-Fidelity Gaussian Process Regression, we formulate a sample-efficient Co-Design optimization that uses both simulated and experiment data. This way, we avoid the necessity of a high-quality system identification while keeping the experiment time and costs low. We validate the approach on an inverted pendulum system, where it is shown to outperform an entirely simulation-based optimization on the real hardware.**

I. Introduction

Control-morphology co-design is an established technique to improve the behaviour of dynamical systems [1, 2, 3, 4, 5, 6, 7, 8]. It has shown the potential to increase the energy efficiency, robustness or speed [1, 2] in applications such as walking [9], manipulation [3] or general non-linear control [4]. However, almost all co-design approaches only compare the original and the optimized design in a zero-shot sim2real fashion [10, 2, 6]. This means they apply the results from simulation directly on the real system without considering the gap between simulation and reality (the "sim2real gap"). This can become problematic for multiple reasons. Wrongly identified parameters might skew the result away from the true optimum. Further, the dynamics might be well described on the original system, but different dynamics might dominate for the optimized parameters. In consequence, the validity of the result is in the best case weakened, or in the worst case the sim2real gap renders the results inapplicable in practice. While this problem is well-known e.g. in control or reinforcement learning, surprisingly, it has to the best of our knowledge not been approached yet for co-design.

Most established methods for co-design use either gradient-based methods [5, 6, 7] which need access to the analytical or estimated gradient. Alternatively, sampling-based methods such as genetic algorithms, CMA-ES [8, 10, 11] or simplex-based Nelder-Mead [12, 2] are used which need to optimize the target function. Neither is particularly suitable if experiments are involved in the process: experiments do not give natural gradients, and their evaluation is expensive and time-consuming. On the other hand, sample-efficient methods such as Gaussian Process Regression (GPR) have found usage for e.g. controller optimization on physical robots [9]. Some approaches [13, 14, 15] combine the experiments with simulations in order to reduce the number of experiments. The result is a Multi-Fidelity Gaussian Process Regression (MF-GPR), an optimization method that trades off between the cost of a real experiment (higher fidelity, higher cost) versus a simulation (lower fidelity, lower cost), and that decides automatically which experiment to perform at which fidelity. GPRs are not entirely new to co-design [3], but so far only used in simulation-based approaches.

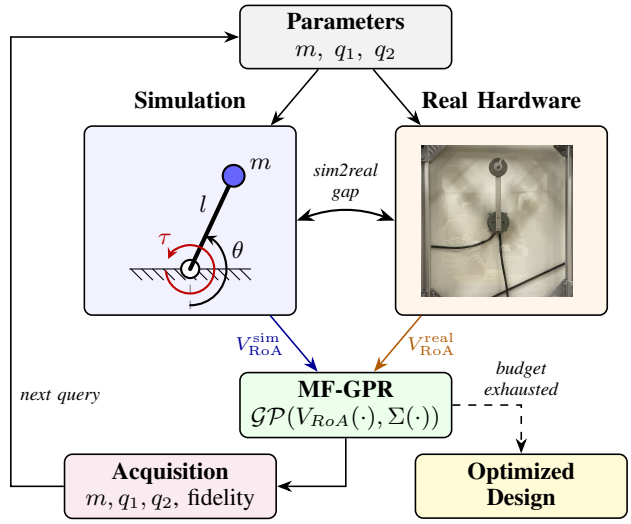

Fig. 1. Schematic of the co-design algorithm. Both simulation and experiment data are used to find a morphology and controller design that works robustly on the real hardware.

We present a novel approach for reliable co-design of a system's morphology and control that bridges the sim2real gap to a successful deployment on the real robot. The algorithm is designed to be robust enough that complete system identification is no longer required. This is realized by including experiment data directly into the co-design process, which to the best of our knowledge has not been proposed before. In this work, we use Multi-Fidelity Gaussian Process Regression which allows to explicitly trade off the cost of real-world experiments versus the lower information gain from simulations. This guides the optimization towards high-value experiment locations while relying on simulation elsewhere. The result is a sample-efficient co-design process that reduces the overall wall clock time. We show the validity

*This work was supported by the Swedish Foundation for Strategic Research (SSF) [Grant number: FFL24-0282].

[1] Division of Dynamics, Department of Mechanics and Maritime Sciences at Chalmers University of Technology 412 96 Gothenburg, Sweden {stockner},{shivesh.kumar}@chalmers.se

of our results on a real system and compare them to a completely simulation-based approach. Our algorithm significantly outperforms the simulation-based approach, which shows its robustness to noise and modelling errors.

II. Mathematical Preliminaries

In this proof-of-concept study, we consider a torque-controlled simple pendulum [16], shown in Fig. 1, with the state $x = [\theta, \dot{\theta}]^T \in \mathcal{X} \subseteq \mathbb{R}^2$ and input $u \in \mathcal{U} \subseteq \mathbb{R}$. The dynamics of the pendulum can be given as

$$\dot{x} = \begin{pmatrix} \dot{\theta} \\ I^{-1}(u - mlg\sin\theta - b_v\dot{\theta} - b_c \operatorname{sign}\dot{\theta}) \end{pmatrix}. \tag{1}$$

The system has an equilibrium at the upright position $x^* = \big(\pi, 0\big)^T$ which marks the desired setpoint. The pendulum is controlled around the setpoint by an LQR controller [17]. This well-established method from optimal control calculates the optimal controller based on the metric $J(u) = \int_0^\infty (x^T Q x + u^T R u) dt$ with the symmetric and positive definite cost matrices $Q = Q^T \succeq 0$ and $R = R^T \succeq 0$. The solution is an optimal linear feedback $u(x) = -Kx$ and a Lyapunov function $V(x) = x^T P x$ that guarantees stability for the linear, unconstrained system.

In our case of an input-constrained non-linear system, this guarantee holds in general only in a linearized region around x^*. This region can be formally described as the *region of attraction* (RoA), a set $\mathcal{R}_a$ of initial states around a fixed point x^* for which the system evolves towards said fixed point as $t \to \infty$. In general, the RoA cannot be computed directly, and is in practice often approximated as the sublevel set of a Lyapunov function $V(x)$ [18]

$$\mathcal{R}_a \approx \{x | V(x) < \rho\}, \tag{2}$$

where $\rho > 0$. This is a valid RoA estimate as long as $V(x) > 0$ and $\dot{V}(x) < 0$ for all $x \in \mathcal{R}_a \setminus \{0\}$, and the best approximation is found by calculating the maximum ρ satisfying these conditions. A way to find such an approximation is described in [19]. First, an RoA candidate of the form Eq. (2) is initialized where ρ_0 is large enough to overestimate the RoA. In the following, random points $x_s \in \mathcal{R}_a$ are sampled, and it is checked if the Lyapunov stability conditions are satisfied on the non-linear system. If the conditions are violated, the constraint ρ is tightened to $\rho = V(x_s)$. With a large enough number of samples, $\mathcal{R}_a$ converges to a valid RoA. An analytic estimation of RoA for torque-limited pendulum is described in [20]. However, we stick to a sampling-based method for the sake of generality.

III. Methodology

We formulate a gradient-free, sampling-based optimization that considers the evaluation at different fidelities, that is, in simulation and reality. The co-design problem formulation, comprising optimization variables, objectives and constraints, can be evaluated at any fidelity and is thus presented in Sec. III-A in a fidelity-agnostic way. The optimization algorithm, which explicitly exploits these multiple fidelities, is then described in Sec. III-B.

A. Co-Design Formulation

a) Parameters: For the optimization problem, we consider the mass m of the system and the LQR gains as decision variables. The mass is constrained by $m^{\min} \leq m \leq m^{\max}$. For the LQR gains $Q \in \mathbb{R}^{2\times 2}$ and $R \in \mathbb{R}$, we have principally five parameters to choose. However, LQR controllers are invariant to a scaling between Q and R and thus we choose R to a fixed positive value, which directly satisfies $R = R^T \succeq 0$. Further, we parametrize

$$Q = \begin{pmatrix} q_1 & 0 \\ 0 & q_2 \end{pmatrix} = \begin{pmatrix} q_1^{\max} e^{-\bar{q}_1} & 0 \\ 0 & q_2^{\max} e^{-\bar{q}_2} \end{pmatrix} \tag{3}$$

similar to [15]. We choose $0 < q_1^{\min} \leq q_i \leq q_1^{\max}, i = 1, 2$ which implies $\bar{q}_i \in \left[0, -\log\left(\frac{q_1^{\min}}{q_1^{\max}}\right)\right]$ and directly satisfies $Q = Q^T \succeq 0$. The parametrization takes into account that the effect of q_i is multiplicative and not additive: Scaling q_i by α has the inverse effect to scaling it by $\frac{1}{\alpha}$. For q_i, this displacement is very asymmetric, but for $\bar{q}_i$ it is equal and opposite, making $\bar{q}$ the natural formulation.

b) Objective: As co-design objective, we choose the volume V_{RoA} of the RoA $\mathcal{R}_a$ of the closed-loop system with parameters q_1, q_2 and m. A large RoA is generally considered a good proxy for robustness to disturbances and noise, making it a natural choice for a system that should perform well under real-world conditions [2, 4]. Practically, we use an approximation of the RoA as shown in Eq. (2) and choose the Lyapunov function $V(x) = (x - x^*)^T P (x - x^*)$. Due to the quadratic nature of $V(x)$, this results in an RoA $\mathcal{R}_a$ with ellipsoidal shape. The volume of $\mathcal{R}_a$ can thus be calculated as [2]

$$V_{RoA} = |det(\sqrt{\Lambda} W)|^{-1} \frac{\pi}{2} \tag{4}$$

where $\frac{1}{\rho} P = W^T \Lambda W$ is the eigendecomposition of $\frac{1}{\rho} P$.

The complete optimization problem can then be formulated as

$$\max_{m, \bar{q}_1, \bar{q}_2} \; V_{RoA} \tag{5a}$$

$$\text{subject to} \quad m \in [m_{min}, m_{max}] \tag{5b}$$

$$\bar{q}_i \in \left[0, -log\left(\frac{q_1^{\min}}{q_1^{\max}}\right)\right] \tag{5c}$$

where V_{RoA} is evaluated with respect to the system dynamics at the chosen mass m and an LQR controller with parameters $R = \text{const.}$, $Q = \text{diag}(q_1, q_1) = \text{diag}(q_1^{\max} e^{-\bar{q}_1}, q_2^{\max} e^{-\bar{q}_2})$.

B. Optimization Algorithm

For the optimization, we adopt a Multi-Fidelity Gaussian Process Regression (MF-GPR) surrogate model. GPR is a well-established method for making predictions of a distribution from a limited number of samples, with an intrinsic uncertainty measure. Through a so-called acquisition function, GPR enables goal-oriented exploration of the optimization variable space [21], e.g. based on the information gain from querying a given data point [22].

Multi-Fidelity GPR adds to this the possibility to combine information from different information sources. MF-GPR

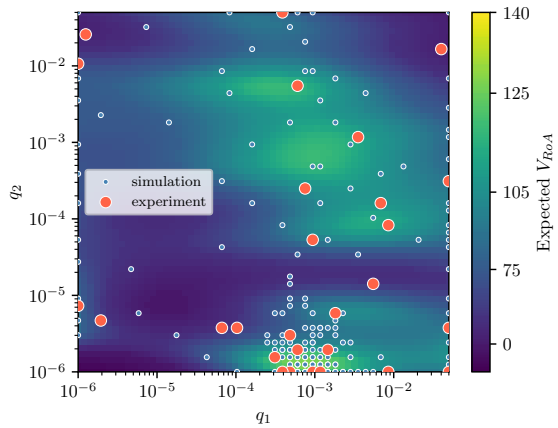

Fig. 2. Estimated Volume of the RoA at "experiment" fidelity for q_1, q_2 with m =70 g. Simulation samples in blue, experiment samples in red.

acquisition functions assign a cost to each fidelity and trade off information gain versus cost. The MF-BO algorithm adopted here, described in [23], supports parallel evaluation across multiple information sources and sampling from a finite discrete set of inputs rather than a continuous range.

C. Experiment Setup

The setup consists of two fidelities, a simulation of the pendulum system and the physical hardware on the open-access CloudPendulum[1] platform [24], as illustrated in Fig. 1. The simulation model uses a length of $0.05\,\mathrm{m}$, the mass as parametrized in the optimization and a moment of inertia of $I = ml^2$. No friction or noise effects are considered to show the algorithm's capability to handle rough models.

The real hardware is a torque-controlled inverted pendulum of length $0.05\,\mathrm{m}$ with a motor at the centre and a position and velocity sensor. The velocity is constrained to $\pm 50\,\mathrm{rad\,s^{-1}}$ and the torque to $\pm 0.04\,\mathrm{N\,m}$. The measurement noise in θ is low with $2.9 \times 10^{-7}\,\mathrm{rad}^2$, but high in $\dot{\theta}$ with $6.4 \times 10^{-2}\,\mathrm{rad}^2\,\mathrm{s}^{-2}$. The pendulum bob has an initial mass of $70\,\mathrm{g}$ which can be changed up to $150\,\mathrm{g}$ in 5 discrete steps by adding one to four $20\,\mathrm{g}$ plates.

To evaluate the objective function in simulation, 50000 sampling points are evaluated with the strategy described in Sec. II. For experimental evaluation, the same strategy using random samples is applied. Due to the noise in the velocity measurement, only initial positions with zero velocity are sampled. Using a built-in mechanism, the system is brought into the given initial position. From there, the controller is activated, the trajectory is recorded for $5\,\mathrm{s}$ and the RoA is updated in case of instability. Overall, the trajectories of 25 randomly sampled points are used.

The optimization is initialized with 25 simulations and three experiments at randomly chosen parameters, after which the MF-GPR algorithm automatically selects fidelity and parameters. It uses an evaluation cost of 1 for simulation and 5 for experiments, and collects 4 samples at a time. For experiments, the pendulum is modified manually with the

[1] https://cloudpendulum.m2.chalmers.se/

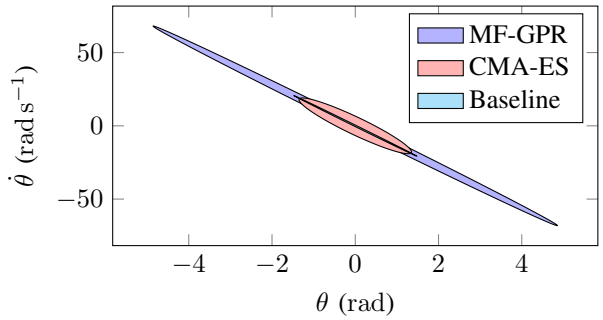

Fig. 3. Resulting $\mathcal{R}_a$ as measured on the hardware for MF-GPR ($q_1 = 7.3 \times 10^{-4}, q_2 = 1.3 \times 10^{-6}$), CMA-ES ($q_1 = 2.7 \times 10^{-3}, q_2 = 3.2 \times 10^{-4}$) and a baseline ($q_1 = 1 \times 10^{-4}, q_2 = 1 \times 10^{-6}$).

correct weight before the experiment starts. The algorithm terminates once an overall cost of 150 is reached.

The hyperparameters are chosen to $R = (1)$ and $q_i \in [0.0001, 5], i = 1, 2$. Since the discrete mass set necessitates discrete inputs, the control values are discretized to 50 log-spaced values in $\bar{q}$ space. With a continuously selectable mass during experiments, the whole MF-GPR formulation could as well be continuous instead. The MF-GPR algorithm uses a semiparametric latent factor model and a Gaussian kernel with automatic relevance determination as described and open-sourced in [23]. The kernel length scale is constrained to $[0.5, 5]$ in each dimension and optimized after four new data points. The noise variance of the measurements is set to $\sigma^2 = 0.01$ and the data is normalized.

IV. Results

The method finds the optimal parameters after 52 iterations with 199 simulations and 28 experiments (corresponding to ca. $52\,\mathrm{min}$ of experiment data) at $m = 70\,\mathrm{g}$, $q_1 = 7.3 \times 10^{-4}$ and $q_2 = 1.3 \times 10^{-6}$. This is expected, as a larger mass makes the torque constraint more restrictive. The low weight on the velocity $\dot{\theta}$ is equally expected due to the noise on the velocity which can negatively affect control performance. Notably, the noise is not visible in the simulation model. For the lowest mass, the surrogate is shown in Fig. 2 for learning from the whole data set.

In order to validate our claim that the hardware-based co-design is more robust to bad models than an entirely simulation-based method, we compare the performance on the physical hardware with an optimization using CMA-ES. This method, that has performed well in previous results [2], uses the same model as the MF-GPR model. It equally yields a mass of $70\,\mathrm{g}$, but depending on the initial parameters, Q lies between $\mathrm{diag}(1.0 \times 10^{-4}, 3.3 \times 10^{-4})$ and $\mathrm{diag}(2.7 \times 10^{-3}, 3.2 \times 10^{-4})$. This gives a V_{RoA} of 25.26, resp. 28.22 which is significantly smaller than the measured V_{RoA} of 47.43 for the MF-GPR. The resulting RoAs are shown in Fig. 3.

V. Conclusion and Outlook

We present results on including hardware-data into the co-design process. To the best of our knowledge, existing works

rely exclusively on simulation data, making them susceptible to poor sim2real transfer. As a solution, we propose to use an MF-GPR that includes both experiment and simulation data. Our initial results show that this approach improves robustness while keeping the number of experiments low. We plan on extending our results also on more challenging systems, such as the double pendulum in the lab. In addition, we want to explore different objective functions and approaches that use data more efficiently.

References

[1] Gabriel Bravo-Palacios, A. Del Prete, and Patrick M. Wensing. "One Robot for Many Tasks: Versatile Co-Design Through Stochastic Programming". In: *IEEE Robotics and Automation Letters* 5.2 (2020), pp. 1680–1687.

[2] Lasse Jenning Maywald et al. "Co-optimization of Acrobot Design and Controller for Increased Certifiable Stability". In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2022, pp. 2636–2641.

[3] Yifei Dong et al. "Co-designing tools and control policies for robust manipulation". In: *arXiv preprint arXiv:2409.11113* 1.2 (2024), p. 3.

[4] Federico Girlanda et al. "Robust Co-Design of Canonical Underactuated Systems for Increased Certifiable Stability". In: *2024 IEEE International Conference on Robotics and Automation (ICRA)*. 2024, pp. 13271–13277.

[5] Katja Mombaur. "Using optimization to create self-stable human-like running". In: *Robotica* 27.3 (2009), pp. 321–330.

[6] Andrew Spielberg et al. "Functional co-optimization of articulated robots". In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. 2017, pp. 5035–5042.

[7] Traiko Dinev et al. "Co-designing robots by differentiating motion solvers". In: *arXiv preprint arXiv:2103.04660* (2021), p. 106.

[8] Michael Chadwick et al. "Vitruvio: An Open-Source Leg Design Optimization Toolbox for Walking Robots". In: *IEEE Robotics and Automation Letters* 5.4 (2020), pp. 6318–6325.

[9] Roberto Calandra et al. "An experimental evaluation of Bayesian optimization on bipedal locomotion". In: *International Conference on Robotics and Automation*. 2014.

[10] Gabriele Fadini et al. "Computational design of energy-efficient legged robots: Optimizing for size and actuators". In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2021, pp. 9898–9904.

[11] Gabriele Fadini et al. "Co-designing versatile quadruped robots for dynamic and energy-efficient motions". In: *Robotica* 42.6 (2024), pp. 2004–2025.

[12] Durgesh Haribhau Salunkhe et al. "An efficient combined local and global search strategy for optimization of parallel kinematic mechanisms with joint limits and collision constraints". In: *Mechanism and Machine Theory* 173 (2022), p. 104796.

[13] Diego Rodriguez, André Brandenburger, and Sven Behnke. "Combining simulations and real-robot experiments for Bayesian optimization of bipedal gait stabilization". In: *Robot World Cup*. Springer, 2018, pp. 70–82.

[14] Akshara Rai et al. "Using simulation to improve sample-efficiency of Bayesian optimization for bipedal robots". In: *Journal of machine learning research* 20.49 (2019), pp. 1–24.

[15] Alonso Marco et al. "Virtual vs. real: Trading off simulations and physical experiments in reinforcement learning with Bayesian optimization". In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2017, pp. 1557–1563.

[16] Felix Wiebe et al. "Torque-limited simple pendulum: A toolkit for getting familiar with control algorithms in underactuated robotics". In: *Journal of Open Source Software* 7.74 (2022), p. 3884.

[17] Brian DO Anderson and John B Moore. *Optimal control: linear quadratic methods*. Courier Corporation, 2007.

[18] Hassan K Khalil and Jessy W Grizzle. *Nonlinear systems*. Vol. 3. Prentice hall Upper Saddle River, NJ, 2002.

[19] Esmaeil Najafi, Robert Babuška, and Gabriel AD Lopes. "A fast sampling method for estimating the domain of attraction". In: *Nonlinear dynamics* 86.2 (2016), pp. 823–834.

[20] Lukas Gross et al. "Analytic estimation of region of attraction of an LQR controller for torque limited simple pendulum". In: *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE. 2022, pp. 2695–2701.

[21] Eric Schulz, Maarten Speekenbrink, and Andreas Krause. "A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions". In: *Journal of Mathematical Psychology* 85 (2018), pp. 1–16.

[22] Philipp Hennig and Christian J Schuler. "Entropy search for information-efficient global optimization." In: *Journal of Machine Learning Research* 13.6 (2012).

[23] Shion Takeno et al. "Multi-fidelity Bayesian optimization with max-value entropy search and its parallelization". In: *International Conference on Machine Learning*. PMLR. 2020, pp. 9334–9345.

[24] Shivesh Kumar. "Swinging pendulums on the cloud: Digitalization of simulation & experimental infrastructure for feedback-based active learning". In: *Chalmers Konferens om undervisning och lärande (KUL2025)*. 2025.