

## Vehicle Cabin Climate MPC Parameter Tuning Using Constrained Contextual Bayesian Optimization (C-CMES)

Stenger, David ; Reuscher, Tim ; Vallery, H.

**DOI**

[10.1109/ITSC57777.2023.10422050](https://doi.org/10.1109/ITSC57777.2023.10422050)

**Publication date**

2023

**Document Version**

Final published version

**Published in**

Proceedings of the IEEE 26th International Conference on Intelligent Transportation Systems, ITSC 2023

**Citation (APA)**

Stenger, D., Reuscher, T., & Vallery, H. (2023). Vehicle Cabin Climate MPC Parameter Tuning Using Constrained Contextual Bayesian Optimization (C-CMES). In *Proceedings of the IEEE 26th International Conference on Intelligent Transportation Systems, ITSC 2023* (pp. 1598-1603). (IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC). IEEE.  
<https://doi.org/10.1109/ITSC57777.2023.10422050>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

***Green Open Access added to TU Delft Institutional Repository***

***'You share, we take care!' - Taverne project***

**<https://www.openaccess.nl/en/you-share-we-take-care>**

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

# Vehicle Cabin Climate MPC Parameter Tuning Using Constrained Contextual Bayesian Optimization (C-CMES)

David Stenger<sup>1,\*</sup>, Tim Reuscher<sup>1,\*</sup>, Heike Vallery<sup>1,2</sup>, and Dirk Abel<sup>1</sup>

**Abstract**—Climate-controlled cabins have for decades been standard in vehicles. Model Predictive Controllers (MPCs) have shown promising results in achieving temperature tracking in vehicle cabins and may improve upon model-free control performance. However, for the multi-zone climate control case, proper controller tuning is challenging, as externally, e.g., passenger-triggered changes in compressor setting and thus mass flow lead to degraded control performance. This paper presents a tuning method to automatically determine robust MPC parameters, as a function of the blower mass flow. Constrained contextual Bayesian optimization (BO) is used to derive policies minimizing a high-level cost function subject to constraints in a defined scenario. The proposed method leverages random disturbances and model-plant mismatch within the training episodes to generate controller parameters achieving robust disturbance rejection. The method contains a postprocessing step to achieve smooth policies that can be utilized in real-world applications. First, simulation results show that the mass flow-dependent policy outperforms a constant parametrization, while achieving the desired closed-loop behavior. Second, the robust tuning method greatly reduces worst-case overshoot and produces consistent closed-loop behavior under varying operating conditions.

## I. INTRODUCTION

Climate control for vehicles is important for reasons of comfort and safety. Thermally uncomfortable environments may lead to a decrease in driver's attention and a resulting increase in reaction times [1]. As a result, customers of modern vehicles, especially in higher-priced segments, expect a climate-control system capable of achieving high thermal comfort under varying operating conditions.

Traditionally, model-free approaches such as proportional-integral-derivative (PID) control are used for this purpose. However, they require large design and application engineering efforts for manual tuning, and closed-loop performance may be unsatisfactory. This is caused in part by the large number of degrees of freedom of the resulting control architecture, introduced by necessary decoupling controllers. It has been shown (e.g., [2], [3], [4]) that model predictive controllers (MPC) have the potential to improve control performance and reduce the number of degrees of freedom especially for multi-zone temperature control.

However, the correct tuning of the MPC hyper-parameters is critical for closed-loop control performance. In [4], the

authors presented an MPC structure for controlling the cabin temperature based on a user-selected mass flow. This control mode is challenging because the selected mass flow alters the system dynamics [5]. As a result, MPC parameters need to be adapted accordingly. Furthermore, unknown disturbances such as solar irradiation or ambient temperature need to be robustly rejected by the MPC. However, their influence is hard to model accurately [6]. Necessary model order reductions for control can increase the model error further [7] resulting in model-plant mismatch especially for lower-order control-oriented models. Due to the wide range in these disturbances, tuning a fixed parameter set by hand achieves satisfactory results only in selected operation scenarios.

Therefore, we propose to formulate the tuning problem as a constrained contextual optimization problem and approximately solve it with Bayesian optimization (BO). BO has been shown to be more sample-efficient than other black-box optimizers on various tuning problems in control [8].

In literature, BO was used to tune the internal model of an MPC (e.g., [9]) and other hyper-parameters such as objective function weights parametrization and prediction horizon (e.g., [10], [11]) and Multi-objective BO was used to address conflicting objectives ([12], [13]). Robustly constrained optimization w.r.t. different model-plant mismatches has been considered for example in [10].

Contextual BO refers to the tuning of parameters as a function of operating conditions, in this case, the user-defined mass flow. Contextual BO for MPCs has seen one experimental application by [14]. However, constraints were not considered. Literature on constrained contextual BO for other controller structures used either a combination with safe BO [15] or approaches based on trust regions [16].

We propose to combine BO with Constrained Max-value Entropy Search (CMES) [17] with contextual BO and therefore name this method C-CMES. A probabilistic constraint formulation ensures robustness against randomly generated disturbance trajectories and unknown model-plant mismatch. A post-processing step addresses smoothness requirements for the policy. The method, applied in simulation, highlights the practical applicability of automatic tuning based on BO for the domain of vehicle cabin climate control.

This paper is structured as follows: In Section II the cabin temperature control tuning problem is described from an application point of view. Cost functions are defined and properties given. The optimization method proposed (C-CMES) is described in Section III. Simulation results regarding performance, smoothness, and robustness are given in Section IV. A discussion of results is given in Section V.

<sup>1</sup> David Stenger, Tim Reuscher, Heike Vallery, and Dirk Abel are with the Institute of Automatic Control, RWTH Aachen University, {d.stenger, t.reuscher}@irt.rwth-aachen.de

<sup>2</sup> Heike Vallery is also with the Department of BioMechanical Engineering, Delft University of Technology, and with the Department for Rehabilitation Medicine, Erasmus MC, Rotterdam, The Netherlands.

\* These authors contributed equally and are listed in random order.

## II. PROBLEM STATEMENT

### A. MPC for Cabin Temperature Control

The vehicle cabin temperature control system considered in this work is shown in Fig. 1. The goal of the control loop is to track temperature references  $T_{\text{ref}}$  in three zones of the vehicle (zone 1: driver zone, zone 2: passenger zone, zone 3: front zone) and to effectively reject disturbances. This corresponds to a usual three-zone climatization setup [1]. For this purpose, a model predictive controller is used to modulate the temperature of inflowing air. The MPC's parameters  $\theta$  are chosen according to a policy (cf. (5)) based on the passenger-defined mass flow  $\dot{m}$ . The model underlying the control system was presented in [5] and is of the form:

$$\begin{pmatrix} \mathbf{T}_{\text{air}} \\ \mathbf{T}_{\text{solids}} \end{pmatrix}_{p+1} = \mathbf{g}(\mathbf{T}_{\text{air},p}, \mathbf{T}_{\text{solids},p}, \dot{m}_p, \mathbf{T}_{\text{mix},p}), \quad (1)$$

with air temperature  $\mathbf{T}_{\text{air},p}$  at time step  $p$ , mean solid component temperature  $\mathbf{T}_{\text{solids},p}$ , mass flow  $\dot{m}_p$  and mixing temperature entering the climatization system  $\mathbf{T}_{\text{mix},p}$ , each as a vector of three zonal values. The map  $\mathbf{g}: \mathbb{R}^{12} \mapsto \mathbb{R}^6$  is strongly nonlinear in mass flow and slightly nonlinear in air temperature [5]. For the manual control mode of the cabin temperature, which the passenger can choose, the mixing temperature is used, while the mass flow level remains controlled by the vehicle passengers. For this control setup, the mass flow is thus an exogenous input. This baseline control system was initially presented in [18] with extensions in [4]. As the open loop is multiple-input multiple-output (MIMO) with strongly coupled states, but mostly linear for constant mass flows, an MPC is used with a continuously (i.e. in each time step) linearized system model. The cost function of this MPC is:

$$J_{\text{MPC}} = \sum_{p=0}^{N_2-1} \|\mathbf{T}_p - \mathbf{T}_{\text{ref},p}\|_q + \|\Delta\mathbf{T}_{\text{mix},p}\|_{\lambda_p}, \quad (2)$$

with the MPC prediction horizon  $N_2$ , the cabin temperature reference  $\mathbf{T}_{\text{ref}}$ , the change in mixing temperature between steps  $\Delta\mathbf{T}_{\text{mix},p} = \mathbf{T}_{\text{mix},p} - \mathbf{T}_{\text{mix},p-1}$ , and weighting vector  $\lambda_p \in \mathbb{R}^3$ . To achieve offset-free tracking, an extended

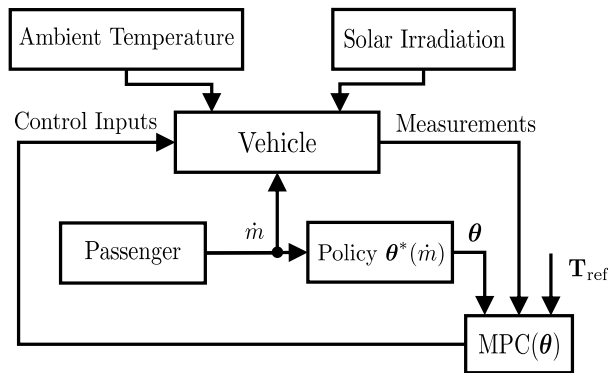


Fig. 1. Signal flow diagram of the control structure presented in [4] augmented with mass flow-dependent policy. Blower mass flow as passenger-defined exogenous input for this contribution.

Kalman filter is used to estimate input disturbances for all air temperature states. This filter is tuned by hand to find a compromise between convergence time and noise reduction. Details on observer tuning, the controller and more motivation can be found in [18], [4].

### B. MPC Tuning Problem

This section outlines the specific tuning problem addressed in this paper. However, the presented method treats the tuning problem as a black box. As a result, it can directly be applied to different degrees of freedom, training episodes, and high-level closed-loop control objectives.

a) *Degrees of Freedom:* The MPC's degrees of freedom are here reduced to the matrix  $\lambda$ , as horizons are predetermined by application needs and computation time limitations. Initial experiments have shown that  $\lambda_0$  should be chosen differently from the remaining entries to allow good tracking and disturbance rejection behavior (cf. (3)). For simplicity, we set the tuning for the third zone to be constant. As the priority for both front zones is assumed to be identical, no separate tuning parameters for these zones were chosen. Therefore, the matrix  $\lambda \in \mathbb{R}^{3 \times N_2}$  for tuning is as follows:

$$\lambda = (\lambda_{p=0} \quad \dots \quad \lambda_{p=N_2-1}) = \begin{pmatrix} \lambda_0 & \lambda & \dots & \lambda \\ \lambda_0 & \lambda & \dots & \lambda \\ 1 & 1 & \dots & 1 \end{pmatrix}. \quad (3)$$

We expect a similar influence of parameter changes in the log scale. This lines up with manual tuning experience and has been shown, e.g., by the authors in [10], Fig. 4. We thus choose the tuning parameter vector to be  $\theta = (\log(\lambda) \quad \log(\lambda_0))$ .

b) *Simulative Training Episode:* A simulation environment is used to evaluate the MPC performance. The optimizer repeatedly queries a simulative training episode with different values of the tuning parameters. The closed-loop behavior is then analyzed and fed back to the optimizer. This way, MPC parameters that achieve the desired closed-loop behavior consistently are searched for.

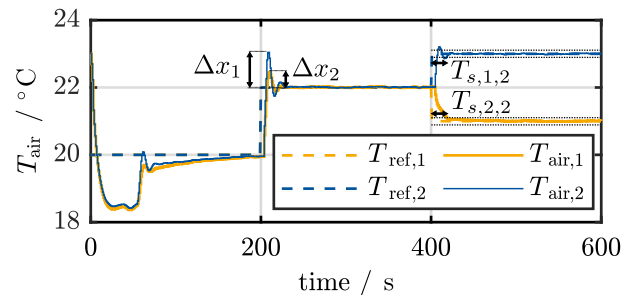


Fig. 2. Training episode including characteristics in time domain

Fig. 2 depicts the simulative training episode used in this work. After an initial convergence of the disturbance observer, the reference changes in both front zones. Afterwards, a reference change leads to deviating zone temperatures. One simulative episode takes about 30 s.

The simulation environment is designed such that it has similar properties as a real-world experiment. This way, parameters tuned in simulation may be applicable to the real system. Additionally, we can evaluate whether the tuning method can handle the challenges of real-world experiments such as noise, disturbances, and varying mass flows.

Therefore, the disturbances (environmental temperature and solar irradiation) are randomly sampled from a catalog of recorded trajectories for each objective function query. A random model plant mismatch is drawn for each function evaluation. All key parameters of the simulation model are uniformly sampled within  $\pm 30\%$  of their nominal values. A constant mass flow is randomly chosen for each episode. In a commercial setting, this random choice may be made by test drivers before release or by customers after release.

c) *Desired Closed-Loop Behaviour*: For vehicle cabin interior temperature controllers, first-order behaviour of the closed-loop control system is desirable (cf. Fig 4, right). Independent of mass flow, disturbances, model-plant mismatch, step height or deviation between individual zones' setpoints, the reference shall be reached

- 1) with short settling time  $T_s$ ,
- 2) and limited overshoot  $\Delta x$ .

Both characteristics are shown in Fig. 2. Settling time and overshoot are defined for both step responses and both zones. To reduce both characteristics to a scalar value, the following definition is chosen:

$$\Delta x = \max(\Delta x_1, \Delta x_2) \quad T_s = \sum_{i,j} T_{s,i,j} \quad (4)$$

with overshoot  $\Delta x_i$  for zone  $i$  and the settling time  $T_{s,i,j}$  for step  $j$  and zone  $i$  as a function of blower mass flow  $\dot{m}$  and MPC parameters  $\theta$ .

### C. Mathematical Optimization Problem Statement

To achieve the desired closed-loop behaviour, the tuning problem is formulated as a robustly constrained, contextual, and stochastic optimization problem:

$$\begin{aligned} \theta^*(s) &= \arg \min_{\theta \in \mathbb{R}^2} \mathbb{E}[J(\theta, s)] \\ \text{s.t.} \quad & P(g(\theta, s) \leq g_{\max}) \geq \delta \\ & \mathbf{s}_{\min} \leq \mathbf{s} \leq \mathbf{s}_{\max} \\ & \boldsymbol{\theta}_{\min} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{\max}, \end{aligned} \quad (5)$$

with maximum constraint  $g_{\max}$  and constraint satisfaction probability  $\delta$ , context  $s = \dot{m}$ , its lower bound  $\mathbf{s}_{\min}$ , its upper bound  $\mathbf{s}_{\max}$ , optimal parameter vector  $\theta^*(s)$ , parameter vector lower bound  $\boldsymbol{\theta}_{\min}$ , upper bound  $\boldsymbol{\theta}_{\max}$ . The constraints for  $\mathbf{s}$  can be chosen based on physical relations and for  $\boldsymbol{\theta}$  based on experience. The operators  $\mathbb{E}$  and  $P$  describe expected value and probability w.r.t. the randomly drawn disturbances and model-plant mismatch. Cost function  $J(\theta, s)$  and constraint  $g(\theta, s)$  are formulated as

$$J(\theta, s) = T_s \quad g(\theta, s) = \Delta x. \quad (6)$$

Note that  $J(\theta, s)$  relates to the hyperparameter optimization and is not the cost function of the MPC algorithm  $J_{\text{MPC}}$  in (2). The key properties and design choices of (5) are:

*Constrained vs. Multi-Objective Optimization*: Reducing settling time and overshoot (cf. (4)) are conflicting objectives. One way to address this conflict is to use single-objective unconstrained optimization with a weighted cost function. However, this generates an additional degree of freedom for the weighting between both terms. In practice, this results in tedious tuning by hand. As an alternative, Pareto optimization [12] can be used to produce a set of compromises between the two objectives. However, this is expected to require more objective function evaluations.

Instead, we propose to separate the terms and use only settling time for the cost function and the overshoot as a constraint value. This formulation yields the desired first-order behavior, as the settling time cost achieves fast approaches to the reference, while the overshoot constraint effectively cancels out any oscillations or higher-order behavior. Due to the probabilistic simulation, a robust constraint formulation

$$P(\Delta x(\theta, \dot{m}) \leq \Delta x_{\max}) \geq \delta \quad (7)$$

is chosen in (5). By choosing  $\delta \stackrel{!}{=} 1$ , the constraint can be guaranteed. This in practice leads to either infeasible behavior or very conservative tuning. Here, we chose  $\delta = 0.5$  and  $\delta = 0.93$  corresponding to  $0 \sigma$  and  $1.5 \sigma$  as non-robust and robust cases (cf. Fig. 3).

*Contextual Optimization*: The MPC's behavior heavily depends on the passenger-selected mass flow. Therefore, we search for optimal MPC parameters as a function of the mass flow, i.e., a policy  $\theta^*(\dot{m}) : \mathbb{R} \mapsto \mathbb{R}^2$  instead of one optimal parameterization for all mass flows. This can be termed gain scheduling in classical control engineering terms. Additionally, we require  $\theta^*(\dot{m})$  to be smooth. Sudden parameter changes with changing mass flow may lead to undesired closed-loop switching behavior. Therefore, the second derivative of the solution  $\frac{d^2 \theta^*}{d \dot{m}^2}$  should be small.

*Noisy Black-Box Optimization*: Information about the functions  $J(\theta, s)$  and  $g(\theta, s)$  can only be obtained by querying the simulation model with varying MPC parameters. Furthermore, evaluating the simulative episode twice with identical MPC parameters and context produces varying results, due to the randomly generated disturbances. The problem is thus a noisy black-box optimization problem. Critically, no analytical gradients can be given for the optimization. These properties restrict the class of usable optimization algorithms considerably.

*Expensive Function Evaluations*: Each objective function evaluation is expensive as it takes around thirty seconds. Therefore, the optimization algorithm needs to be sample-efficient, i.e., it needs to be able to find good solutions with as little objective function evaluations as possible.

*Simple Regret and Safety*: Here, we focus on the performance of the final policy  $\theta^*(\dot{m})$  after  $K$  iterations. Large cost function values during the optimization are not critical. Additionally, there are no immediate functional safety issues for exceeded constraint values during optimization. This is equivalent of minimizing simple regret instead of cumulative regret. Additionally, safe sampling is not required, although it can be achieved with BO [15].

### III. BO WITH C-CMES: CONTEXTUAL CONSTRAINED MAX-VALUE ENTROPY SEARCH

#### A. Optimization

BO has been shown to be more sample-efficient than other black-box optimizers [8]. Additionally, it naturally addresses noisy optimization problems. For a detailed introduction to BO, the reader is referred to [19] and [20]. Here, we extend constrained max-value entropy search (CMES)[17] to the contextual case to account for all challenges posed by (5). This section emphasises the main differences to unconstrained single-objective BO.

---

#### Algorithm 1 Constrained Contextual BO.

---

- 1: Generate an initial data set  $\mathcal{D}_1 = (\Theta_1, \mathcal{J}_1, \mathcal{G}_1, \mathcal{S}_1)$
  - 2: **for**  $k = 1; 2; \dots; K$  **do**
  - 3: Learn probabilistic surrogate models  $(\mathcal{GP}_k^J, \mathcal{GP}_k^G)$  using all past evaluations  $\mathcal{D}_k = (\Theta_k, \mathcal{J}_k, \mathcal{G}_k, \mathcal{S}_k)$
  - 4:  $\mathbf{s}_{k+1} \leftarrow \text{receiveContext}()$
  - 5: Select  $\boldsymbol{\theta}'_{k+1}$  by optimizing the acquisition function for context  $\mathbf{s}_{k+1}$  :
 
$$\boldsymbol{\theta}'_{k+1} = \arg \min_{\boldsymbol{\theta}_{\min} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{\max}} \text{CMES}(\boldsymbol{\theta}, \mathbf{s}_{k+1}, \Theta_k, \mathcal{GP}_k^G, \mathcal{GP}_k^J) \quad (8)$$
  - 6: Query expensive-to-evaluate simulation with  $\boldsymbol{\theta}'_{k+1}$  to obtain  $j'_{k+1}$  and  $g'_{k+1}$  for context  $\mathbf{s}_{k+1}$
  - 7: Augment data set with new evaluations to obtain  $\mathcal{D}_{k+1} = (\Theta_{k+1}, \mathcal{J}_{k+1}, \mathcal{G}_{k+1}, \mathcal{S}_{k+1})$
  - 8: **end for**
  - 9: Obtain final smooth policy  $\boldsymbol{\theta}^*(s)$  (Sec. III-B)
- 

First, in Step 1 of Algo. 1, an initial data set  $\mathcal{D}_1$  consisting of evaluated parameters  $\Theta_1$  for contexts  $\mathcal{S}_1$  and obtained objective function  $\mathcal{J}_1$  and constraint  $\mathcal{G}_1$  values is generated to obtain initial information about the latent objective function  $J(\boldsymbol{\theta}, \mathbf{s})$  and constraints  $g(\boldsymbol{\theta}, \mathbf{s})$ . Latent means that only noisy samples of both functions are available (cf. Sec. II-C).

The main optimization loop spans Steps 2 - 8. In each iteration Gaussian Process (GP) surrogate models of objective function and constraint function are generated (Step 3):

$$\begin{aligned} J(\boldsymbol{\theta}, \mathbf{s}) &\approx \tilde{J}_k(\boldsymbol{\theta}, \mathbf{s} \mid \mathcal{D}_k) \sim \mathcal{N}(\mu_{J,k}(\boldsymbol{\theta}, \mathbf{s}), \sigma_{J,k}^2(\boldsymbol{\theta}, \mathbf{s})), \\ g(\boldsymbol{\theta}, \mathbf{s}) &\approx \tilde{g}_k(\boldsymbol{\theta}, \mathbf{s} \mid \mathcal{D}_k) \sim \mathcal{N}(\mu_{g,k}(\boldsymbol{\theta}, \mathbf{s}), \sigma_{g,k}^2(\boldsymbol{\theta}, \mathbf{s})). \end{aligned} \quad (9)$$

We follow [21] in modelling the true unknown objective function  $J(\boldsymbol{\theta}, \mathbf{s})$  as a GP. The model of iteration  $k$ ,  $\mathcal{GP}_k^J$ , yields normally distributed predictions  $\tilde{J}_k(\boldsymbol{\theta}, z \mid \mathcal{D}_k)$  as a function of parameters and context. Predictive mean and predictive standard deviation are denoted as  $\mu$  and  $\sigma$ , respectively. Predictive uncertainty increases in areas of the parameter-context space, where the MPC performance has not been evaluated yet. For the GP models (cf. [22]), a constant priori mean, homoscedastic Gaussian likelihood, and an anisotropic squared exponential is used. Smooth box hyper-priors are placed on the kernel length scales in order to constrain them to sensible orders of magnitude. The

GP models hyperparameters are updated at each iteration using maximum a-posteriori estimation. For the constrained function  $g(\boldsymbol{\theta}, \mathbf{s})$  an independent GP model with identical settings is constructed.

The GP models are used to determine useful parameters  $\boldsymbol{\theta}'_{k+1}$  to be evaluated next. This decision also depends on the context of the next episode. As described in Sec. II, the context cannot be chosen by the optimizer and is instead received from the simulation environment (Step 4). When choosing the next sample point, a so-called acquisition function is used to balance between exploration and exploitation (Step 5). Here, we use CMES [17], an information theoretic acquisition function, based on max-value entropy search ([23]). In contrast to improved-based acquisition functions, CMES does not require a single *current best solution*. Additionally, in contrast to UCB that was used in [14] and [16], CMES naturally handles constraints and therefore does not require trust regions or SafeBO. In general, the acquisition function is multimodal. Therefore, it is optimized using a combination of random search and gradient-based optimization.

In Step 6, the expensive-to-evaluate simulative episode is queried with the suggested controller parameters in order to obtain the resulting system responses. After the data set is augmented (Step 7), the next iteration is started. The optimization is terminated after a given budget of  $K$  objective function evaluations is exhausted.

#### B. Post Processing for Smoothness

After the main optimization loop has terminated, final GP models  $\mathcal{GP}_K^J, \mathcal{GP}_K^G$  are available. Although they ideally have been refined in promising parameter regions, a final smooth policy  $\boldsymbol{\theta}^*(s)$  still needs to be obtained.

To achieve that, we discretize the context space  $\mathcal{S}_{\text{disc}} = \{s_1, \dots, s_N\}$ . The corresponding optimal parameter values are denoted as  $\boldsymbol{\theta}_1^*, \dots, \boldsymbol{\theta}_N^*$ . To obtain them the expected objective function value is minimized, while the probabilistic constraints are required to hold:

$$\begin{aligned} \min_{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N \in \mathbb{R}^d} & \sum_{n=1}^N \mu_{J,K}(\boldsymbol{\theta}_n, s_n) + \dots \\ & \dots \gamma \sum_{n=1}^{N-2} (\boldsymbol{\theta}_{n+2} - 2\boldsymbol{\theta}_{n+1} + \boldsymbol{\theta}_n) \\ \text{s.t.} & \boldsymbol{\theta}_{\min} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{\max} \\ & \psi \left( \frac{\mu_{g,K}(\boldsymbol{\theta}_n, s_n) - g_{\max}}{\sigma_{g,K}(\boldsymbol{\theta}_n, s_n)} \right) > \delta, n \in \{1, \dots, N\}, \end{aligned} \quad (10)$$

As an additional term, the second-order finite difference of the discretized optimal policy addresses the smoothness requirements. The weighting parameter  $\gamma$  needs to be adjusted manually, to get the desired smoothness. In practice, the discretized policy can be interpolated in order to extract parameters for each context. Eq. (10) is solved using gradient-based optimization. The optimal feasible solution for each of the discretized contexts is used as an initial guess.

## IV. SIMULATION RESULTS

### A. Policy Smoothing

The policy retrieved after 500 black-box evaluations is shown in Fig. 3. Both smooth and non-smooth policies, e.g., MPC parameters as a function of mass flow, are shown for the robust case in the top plot and the non-robust case (cf. (7)) in the bottom plot. For the non-robust case, the smoothing postprocessing leads to a mere smoothing of the solution. In the robust case, the smoothing fixes the solution to one of the optimal solutions between which the non-smooth solution appears to be switching. Up to about  $90 \text{ kg h}^{-1}$ , the behavior is similar to the non-robust case. Afterwards,  $\lambda$  drops, while  $\lambda_0$  increases sharply. The non-smooth policy switches multiple times between these two solutions. Without smoothing, this behavior would not be acceptable for a closed-loop operation of the controller.

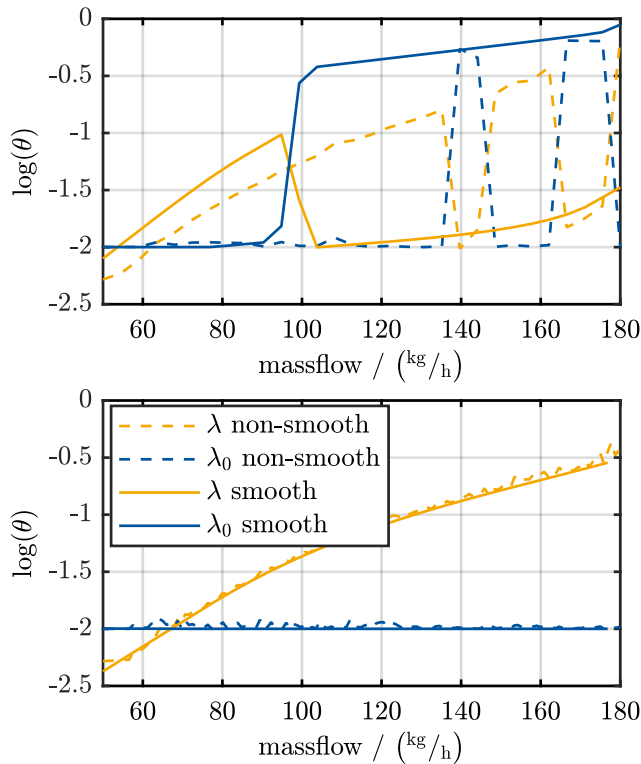


Fig. 3. Optimized context-/mass flow-dependent MPC parameterizations. **Top:** Robust policy ( $\delta = 0.93$ ). **Bottom:** Non-robust policy ( $\delta = 0.5$ ).

### B. Constant vs. Contextual Policy

The advantage of contextual over constant tuning is shown in Fig. 4. Using constant tuning, only at  $\dot{m} = 50 \text{ kg h}^{-1}$  the desired closed-loop behavior was achieved. The same tuning leads to  $0.3 \text{ K}$  overshoot at  $\dot{m} = 100 \text{ kg h}^{-1}$  and unwanted oscillations, as well as  $0.7 \text{ K}$  overshoot at  $\dot{m} = 150 \text{ kg h}^{-1}$ . In contrast, contextual tuning achieves satisfactory results for all mass flows. No overshoot and first-order behavior can be seen for all mass flows. Thus, contextual tuning is highly important to cabin temperature control.

### C. Robustness

To validate the robust behavior of the closed control loop, 50 episodes with model plant mismatch and disturbance sampling (cf. Sec. II-B.0.b) are simulated for the two smooth policies shown in Fig. 3. The individual model plant mismatch combinations were different to the ones drawn in policy generation. The results of these simulations are shown in Fig. 5. The distribution of cost function (i.e. settling time) values is shown in the top plot. The robust set shows overall a worse performance regarding cost function. Both mean and variance are higher. For the constraint shown in the bottom plot, mean and especially variance is strongly reduced. 94% of samples were below the threshold of  $0.05 \text{ K}$ , which matches the desired value of  $\delta = 93.7\%$ . The increased robustness regarding constraint compliance of the method is thereby validated. The case shown here puts a strong focus on robustness. The effect of decreasing delta and thus relaxing this focus is also shown in Fig. 5. The method allows an application-driven trade-off between the characteristics.

## V. CONCLUSION

This paper proposed a Bayesian optimization methodology that produces controller parameters, i.e., a policy, as a smooth function of the context. Furthermore, the policy is robust w.r.t. randomly drawn disturbance trajectories minimizing a high-level objective while fulfilling constraints.

The method is used to optimize the parameters of an MPC for vehicle climate control. Three main results are obtained from Monte Carlo validation simulations. First, a distinct post-processing step allows obtaining a smooth policy from the GP models fitted during optimization. Second, contextual optimization is elemental for operating an MPC under different conditions, in this case, different blower mass flows. Third, a robust formulation of the constraints enables BO to generate a policy that is robust to different ambient temperature and solar irradiation trajectories.

The results encourage an experimental application of the presented method on a physical vehicle in future work. The simulative training episode used here has the same structural properties as a real-world experiment, therefore easy transferability is expected. Additionally, other BO versions have been shown to be applicable in experiments in other applications, e.g. [14].

## REFERENCES

- [1] S. Daly, *Automotive air-conditioning and climate control systems*, 1st ed. Amsterdam: Butterworth-Heinemann, 2006.
- [2] H. Wang, I. Kolmanovsky, M. R. Amini, and J. Sun, "Model predictive climate control of connected and automated vehicles for improved energy efficiency," in *2018 Annual American Control Conference (ACC)*. Piscataway, NJ: IEEE, 2018.
- [3] I. Cvok, B. Škugor, and J. Deur, "Control trajectory optimisation and optimal control of an electric vehicle hvac system for favourable efficiency and thermal comfort," *Optimization and Engineering*, vol. 22, no. 1, pp. 83–102, 2021.
- [4] T. Reuscher, K. Poovendran, and D. Abel, "Model predictive zonal temperature control of a vehicle cabin," in *2021 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 2021, pp. 37–43.



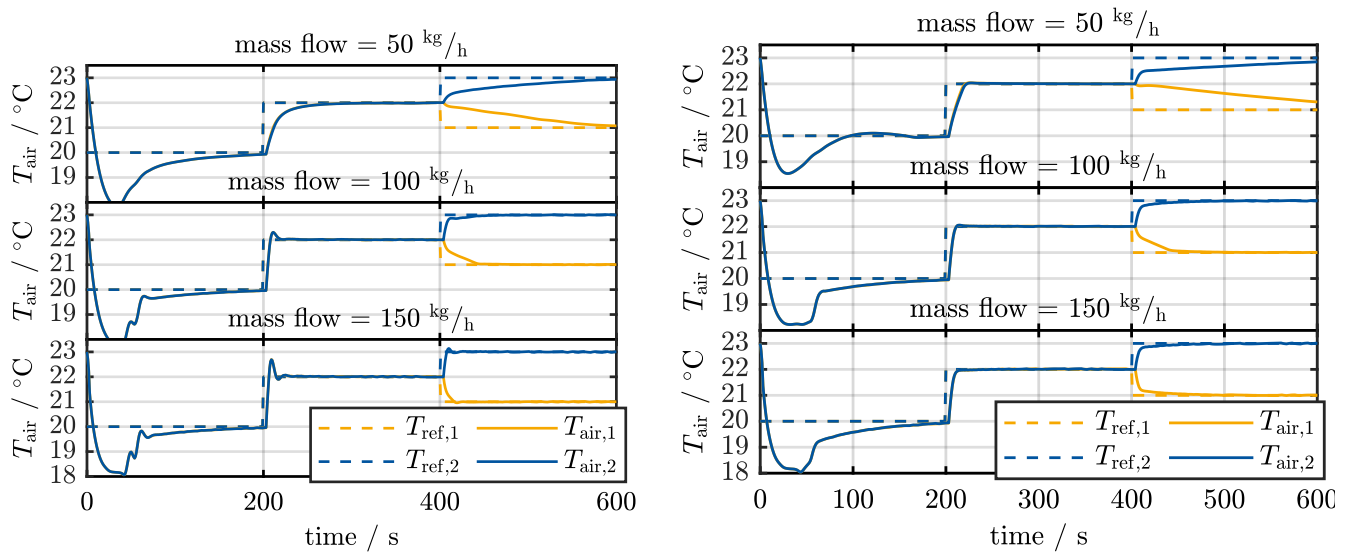


Fig. 4. Closed-loop time domain plots for the training episode with different mass flows and constant model-plant mismatch and disturbance. **Left:** Constant MPC parametrization. Overshoot is visible after reference steps. **Right:** Smooth and non-robust contextual MPC parametrization. Only minor overshoot after reference steps.

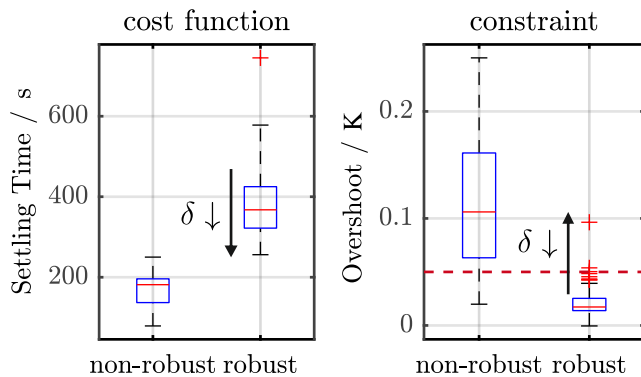


Fig. 5. Distribution of objective function and constraint values obtained from 50 validation samples. Dotted red line: allowed overshoot of 0.05 K

- [5] K. Poovendran, D. Abel, T. Reuscher, and V. Govender, "Vehicle cabin thermal multi-zone modelling for control," in *2020 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA)*. IEEE, 11/11/2020 - 11/13/2020, pp. 489–495.
- [6] Q. Zhang, S. E. Li, and K. Deng, *Automotive Air Conditioning*. Cham: Springer International Publishing, 2016.
- [7] D. Klemm, W. Roessner, N. Widdecke, and J. Wiedemann, "Reduced model of a vehicle cabin for transient thermal simulation," ser. SAE Technical Paper Series. SAE International, 2018.
- [8] D. Stenger and D. Abel, "Benchmark of bayesian optimization and metaheuristics for control engineering tuning problems with crash constraints," 2022. [Online]. Available: <https://arxiv.org/abs/2211.02571>
- [9] D. Piga, M. Forgione, S. Formentin, and A. Bemporad, "Performance-oriented model learning for data-driven mpc design," *IEEE Control Systems Letters*, vol. 3, no. 3, pp. 577–582, 2019.
- [10] D. Stenger, M. Ay, and D. Abel, "Robust parametrization of a model predictive controller for a cnc machining center using bayesian optimization," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 10388–10394, 2020.
- [11] O. Andersson, M. Wzorek, P. Rudol, and P. Doherty, "Model-predictive control with stochastic collision avoidance using bayesian policy optimization," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 4597–4604.
- [12] D. Stenger, R. Ritschel, F. Krabbes, R. Voßwinkel, and H. Richter, "What is the best way to optimally parameterize the mpc cost function for vehicle guidance?" *Mathematics*, vol. 11, no. 2, 2023. [Online]. Available: <https://www.mdpi.com/2227-7390/11/2/465>
- [13] G. Makrygiorgos, A. D. Bonzanini, V. Miller, and A. Mesbah, "Performance-oriented model learning for control via multi-objective bayesian optimization," *Computers & Chemical Engineering*, vol. 162, p. 107770, 2022.
- [14] L. P. Fröhlich, C. Küttel, E. Arcari, L. Hewing, M. N. Zeilinger, and A. Carron, "Model learning and contextual controller tuning for autonomous racing." [Online]. Available: <https://arxiv.org/pdf/2110.02710>
- [15] M. Fiducioso, S. Curi, B. Schumacher, M. Gwerder, and A. Krause, "Safe contextual bayesian optimization for sustainable room temperature pid control tuning." [Online]. Available: <http://arxiv.org/pdf/1906.12086v1>
- [16] Jinkyoo Park, "Contextual bayesian optimization with trust region (cbotr) and its application to cooperative wind farm control in region 2," *Sustainable Energy Technologies and Assessments*, vol. 38, p. 100679, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2213138819303108>
- [17] V. Perrone, I. Shcherbatyi, R. Jenatton, C. Archambeau, and M. Seeger, "Constrained bayesian optimization with max-value entropy search." [Online]. Available: <https://arxiv.org/pdf/1910.07003>
- [18] T. Reuscher, K. Poovendran, and D. Abel, "Control oriented modeling and nonlinear model predictive temperature control of a vehicle cabin," in *European Control Conference 2020*, A. Pogromsky and Y. Ebihara, Eds. Piscataway, NJ: IEEE, 2020, pp. 124–129.
- [19] R. Garnett, *Bayesian Optimization*. Cambridge University Press, 2022, in preparation.
- [20] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, "Taking the human out of the loop: A review of bayesian optimization," *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, 2016.
- [21] A. Krause and C. S. Ong, "Contextual gaussian process bandit optimization," in *Advances in Neural Information Processing Systems 24*, J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, Eds. Curran Associates, Inc, 2011, pp. 2447–2455. [Online]. Available: <http://papers.nips.cc/paper/4487-contextual-gaussian-process-bandit-optimization.pdf>
- [22] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, 2006. [Online]. Available: <http://www.gaussianprocess.org/gpml/>
- [23] Z. Wang and S. Jegelka, "Max-value entropy search for efficient bayesian optimization." [Online]. Available: <https://arxiv.org/pdf/1703.01968>