BAYESIAN OPTIMIZATION WITH GAUSSIAN PROCESS REGRESSION USING OBSERVATION-DEPENDENT UNCERTAINTIES FOR REDUCING FLOW INDUCED TRANSVERSE LOADS ON MULTI-STEP CYLINDERS

C. Tian
Institute of Fluid Mechanics (LSTM)
Friedrich-Alexander-Universität (FAU)
Erlangen-Nürnberg, 91058 Erlangen, Germany
cai.tian@fau.de
School of Naval Architecture, Ocean and Civil Engineering
Shanghai Jiao Tong University
Shanghai 200240, China
cai.tian@sjtu.edu.cn

S. Rezaeiravesh
Department of Fluids and Environment
The University of Manchester
Manchester M13 9PL, UK
saleh.rezaeiravesh@manchester.ac.uk

P. Schlatter
Institute of Fluid Mechanics (LSTM)
Friedrich-Alexander-Universität (FAU)
Erlangen-Nürnberg, 91058 Erlangen, Germany
philipp.schlatter@fau.de

1 ABSTRACT

Multi-step cylinders are widely used in marine riser systems, but the large-diameter buoyancy elements may amplify vortex-induced vibration by increasing the transverse fluid loading. Previous shape optimization using Bayesian optimization with Gaussian process regression (BO-GPR) showed that lift fluctuations of a multi-step cylinder can be strongly suppressed, but at a noticeable computational cost, even at a relatively low Reynolds number of 300. To improve the practicality of this approach, the present study introduces a heteroscedastic BO-GPR framework that explicitly accounts for observation-dependent uncertainties from finite-length unsteady CFD simulations. The statistical uncertainty of each sample is estimated from the lift-force time series using an autocorrelation-based integral-time-scale method and is incorporated into the Gaussian process surrogate as case-dependent observation noise. In addition, each optimization cycle adaptively decides whether to evaluate a new design point or refine an existing sample to reduce its uncertainty. When applied to optimizing the multi-step cylinder geometry, the proposed method achieves results comparable to conventional BO-GPR, while delivering a 50% improvement in efficiency.

2 Introduction

The cylindrical structure with several small- and large-diameter coaxial cylinders is called the multi-step cylinder, as shown in figure 1(a). Structures with a similar shape to a multi-step cylinder have been broadly used in marine riser systems. An example is the steel lazy wave riser (SLWRs) (Felisita *et al.*, 2017), shown in figure 1(b), where the large diameter buoyancy elements are attached to the small diameter riser. The additional force provided by the buoyancy elements could decompose the motion between the marine platform and the riser attached to the seabed. This leads to reduced dynamic responses and better fatigue performance, making SLWRs an attractive choice for the deep-water riser system (Ruan *et al.*, 2024).

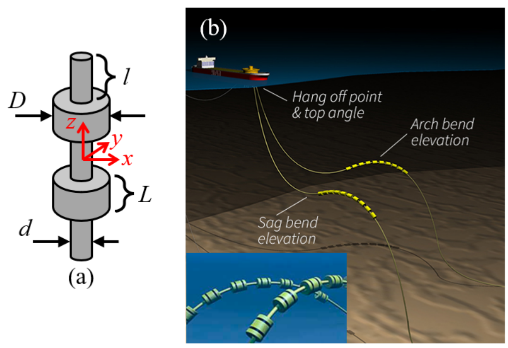

Figure 1. (a) A sketch of the multi-step cylinder geometry. (b) A view of steel lazy wave risers (retrieved from https://www.kongsberg.com).

However, the large-diameter cylindrical buoyancy has the risk of amplifying local vortex-induced vibration (VIV) responses because of the larger hydrodynamic fluctuating force induced by its larger diameter than that of the small-diameter cylinder. Compared to the bare uniform cylinder, the alternately arranged coaxial small and large cylinders significantly changed the vortex formation (Dunn & Tavoularis, 2006; Tian *et al.*, 2017), the structural loads (McClure *et al.*, 2015; Massaro & Schlatter, 2024; Tian *et al.*, 2020*b*), the turbulent characteristics in the wake (Tian *et al.*, 2020*a*; Massaro *et al.*, 2023*a*,*b*), and the fluid structure interactions (Ji *et al.*, 2019; Zhao *et al.*, 2024). These bring challenges to understanding and predicting the vortex-induced vibration (VIV) of the multi-step cylinders. VIV of risers, induced by ocean waves and currents, can cause extensive fatigue damage if not adequately mitigated. Therefore, previous research (Li *et al.*, 2017; Jhingran *et al.*, 2017) has mainly focused on the dynamic response of different multi-step cylinders, where the diameter ratio $2.5 < D/d < 5$ (where D and d represent the diameter of the large and small cylinders, respectively; see Figure 1(a)) is usually picked up. However, the exploration of structural load optimization of the multi-step cylinder is limited.

Considering that the transverse force (i.e., the lift force) is generally more important than the in-line force (i.e., the drag force) for VIV, Tian *et al.* (2026) applied the Bayesian optimization based on Gaussian process regression (BO-GPR) described by Morita *et al.* (2022) to explore the possibility of suppressing the transverse force (i.e., lift force) of the multi-step cylinder by properly designing its shape. After 26 simulations, a reduction of approximately 99% in the lift force was achieved. While this result demonstrates the approach's potential, it required more than 100,000 CPU hours, even at a relatively low Reynolds number of 300. For realistic flow conditions at much higher Reynolds numbers, the computational cost is expected to increase substantially due to the need for finer spatial and temporal resolutions. Therefore, improving the efficiency of the BOGP framework is a critical prerequisite for its practical application.

In the present work, we address this challenge by introducing an observation uncertainty model based on time-step iterations of CFD simulations. By explicitly accounting for the uncertainty arising from partially converged simulations, the proposed approach enables adaptive use of computational effort, thereby significantly improving the overall efficiency of the BOGP optimization process.

Methodology

The present work proposes a heteroscedastic Bayesian optimization framework for unsteady CFD design optimization. Direct numerical simulations (DNS) are carried out using the second-order finite-volume solver MGLET (Manhart, 2004) in the present work. An open-source version of MGLET is available at *https://github.com/kmturbulenz/mglet-base*. For the optimization framework, in contrast to a conventional Bayesian optimization with Gaussian process regression (Morita *et al.*, 2022), each evaluated sample is not represented only by a design-response pair, but by a triplet $\mathscr{D}_n = \{(\mathbf{q}_i, r_i, \sigma_i)\}_{i=1}^n$, where $\mathbf{q}_i$ is the design vector, r_i is the observed objective value, and σ_i is the case-dependent observation uncertainty. Another key difference from standard BO is that each optimization cycle in the present study does not necessarily launch a new design point. Instead, there are two possible actions: (1) NEW - start a new CFD simulation at a new design point; (2) REFINE - continue or refine an existing case in order to reduce its statistical uncertainty. The action decision is made automatically by comparing a new-point score and a refinement score, which will be explained in detail in the following.

Design parameters and objective function:

Design parameters and objective function: The design vector is written as $\mathbf{q} = [q_1, q_2]^\top \in \mathscr{Q}$, where q_1 and q_2 represent the lengths of the small and large cylinders, respectively. The admissible design space is defined as $\mathscr{Q} = \{(q_1, q_2) \mid q_1 \in [0.1, 5],\ q_2 \in [0.1, 3]\}$. The objective function is set up as $r(\mathbf{q}) = \log\left(1 + \frac{C_{L_{\mathrm{rms}}}(\mathbf{q})D}{\varepsilon q_2}\right)$ to make it vary within a small range even when the lift force coefficient $C_{L_{\mathrm{rms}}}$ changes in several orders. Here, $\varepsilon = 10^{-5}$ is used, the RMS lift force coefficient is defined as $C_{L_{\mathrm{rms}}} = \sqrt{\mathbb{E}\left[C_L^2\right]}$, where $C_L(t) = \frac{F_y(t)}{\rho U(q_1 d + q_2 D)}$ and $F_y(t)$ is the time series of the lift force.

Uncertainty estimation:

Since the objective value $r(\mathbf{q})$ is calculated from a finite-length unsteady CFD signal, it is affected by sampling uncertainty. Therefore, the observation at design point $\mathbf{q}_i$ is modeled as $r_i = f(\mathbf{q}_i) + \varepsilon_i \mid\ \varepsilon_i \sim \mathscr{N}(0, \sigma_i^2)$, where $f(\mathbf{q})$ denotes the latent noise-free objective, and σ_i is the case-dependent observation standard deviation. The uncertainty is estimated from the time series of the lift force $F_y(t)$ using an autocorrelation function (ACF) based integral-time-scale method (Oliver *et al.*, 2014; Mockett *et al.*, 2010).

The analysis is performed on the squared lift signal, $y(t) = C_L^2(t)$, such that the RMS lift coefficient is expressed as $C_{L_{\mathrm{rms}}} = \sqrt{\mathbb{E}[y]}$. For the discrete samples y_k in the selected statistical window, with average sampling interval Δt, the normalized autocorrelation function is defined as $\rho(\ell) = \frac{\sum_{k=1}^{N-\ell}(y_k - \bar{y})(y_{k+\ell} - \bar{y})}{\sum_{k=1}^{N}(y_k - \bar{y})^2}$, where $\bar{y}$ is the sample mean. The corresponding integral time scale is estimated as $\tau_{\mathrm{int}} = \Delta t\left(\frac{1}{2} + \sum_{\ell=1}^{\ell_c}\rho(\ell)\right)$, with ℓ_c taken as the first lag for which the autocorrelation becomes negative. The effective number of independent samples is then approximated by $N_{\mathrm{eff}} = \frac{T}{2\tau_{\mathrm{int}}}$, with $T = N\Delta t$, which accounts for the temporal correlation in the signal. Using this effective sample size, the standard error of the mean of y is estimated as $\mathrm{SE}(\bar{y}) = \sqrt{\frac{\mathrm{Var}(y)}{N_{\mathrm{eff}}}}$, and, by error propagation, the uncertainty of the RMS lift coefficient is obtained as $\mathrm{SE}(C_{L_{\mathrm{rms}}}) = \frac{\mathrm{SE}(\bar{y})}{2C_{L_{\mathrm{rms}}}}$.

Considering that BO uses $r(\mathbf{q})$ rather than $C_{L_{\mathrm{rms}}}$ directly, the uncertainty must be mapped from the RMS space to the objective space. This is done by interval propagation: $C_{L_{\mathrm{rms}}}^- = \max\left(C_{L_{\mathrm{rms}}} - \sigma_{C_L}, 0\right)$ and $C_{L_{\mathrm{rms}}}^+ = C_{L_{\mathrm{rms}}} + \sigma_{C_L}$, where $\sigma_{C_L} = \mathrm{SE}(C_{L_{\mathrm{rms}}})$. The corresponding lower and upper objective values are $r^- = \log\left(1 + \frac{C_{L_{\mathrm{rms}}}^-}{\varepsilon q_2}\right)$ and $r^+ = \log\left(1 + \frac{C_{L_{\mathrm{rms}}}^+}{\varepsilon q_2}\right)$ and the objective uncertainty is defined as $\sigma = \frac{r^+ - r^-}{2}$.

Heteroscedastic Gaussian process surrogate:

To improve robustness, the physical design space is first mapped to the unit square: $\mathbf{x} = \frac{\mathbf{q} - \mathbf{q}_{\min}}{\mathbf{q}_{\max} - \mathbf{q}_{\min}}$, where $\mathbf{x} \in [0,1]^2$. The surrogate model is then trained with the heteroscedastic observation model $r_i = f(\mathbf{x}_i) + \varepsilon_i$, where $\varepsilon_i \sim \mathscr{N}(0, \sigma_i^2)$ using a Gaussian process regressor with sample-wise noise variance σ_i^2. The covariance kernel is chosen as the Matérn-5/2. The kernel hyperparameters are optimized using the L-BFGS-B algorithm Zhu *et al.* (1997). For any candidate point $\mathbf{x}$, the GP provides the predictive mean $\mu(\mathbf{x})$ and standard deviation $s(\mathbf{x})$. Here, σ_i and $s(\mathbf{x})$ represent two different uncertainty notions: σ_i is the observation uncertainty of a computed CFD case, while $s(\mathbf{x})$ is the posterior uncertainty of the surrogate. This distinction is essential in the present method because the former enters the GP training noise term, whereas the latter is used to determine whether a NEW or a REFINE case will be conducted.

Action between NEW and REFINE:

A key difference from conventional Bayesian optimization is that the present framework does not necessarily launch a new design point at every iteration. Instead, two possible actions are considered simultaneously: *NEW*, which evaluates a previously unexplored design, and *REFINE*, which continues the simulation of an existing sample in order to reduce its observation uncertainty. The optimizer then selects the action with the larger expected benefit.

For a candidate new point $\mathbf{x}$, the corresponding score is defined as $\text{Score}_{\text{new}}(\mathbf{x}) = \frac{\text{EI}(\mathbf{x})}{C_{\text{new}}} \cdot \frac{1}{1+\sigma_{\text{new}}(\mathbf{x})/\sigma_{\text{ref}}}$, where $\text{EI}(\mathbf{x})$ is the Expected Improvement, C_{new} denotes the computational cost of launching a new short simulation (i.e., the total time-step a new simulation needs), and $\sigma_{\text{new}}(\mathbf{x})$ is the estimated uncertainty associated with that short run. In this way, a new point is preferred when it is both promising in terms of objective improvement and sufficiently reliable in terms of statistical accuracy. Since $\sigma_{\text{new}}(\mathbf{x})$ is unknown before running the simulation, it is approximated from nearby existing samples. In the present implementation, a distance-weighted k-nearest-neighbor estimate is used, $\sigma_{\text{new}}(\mathbf{x}) = \sum_{j\in\mathcal{N}_k(\mathbf{x})} w_j \sigma_j$, $w_j \propto \frac{1}{d(\mathbf{x},\mathbf{x}_j)}, k=3$.

For an already evaluated sample i, the objective of refinement is not to explore a new region of the design space, but to reduce the uncertainty of a case that may already be important for the optimization. To assess the value of a refinement action, a temporary heteroscedastic GP model is reconstructed by reducing the uncertainty of only one existing sample, while keeping all others unchanged. The refinement score is then defined as $\text{Score}_{\text{refine}}(i) = \frac{1}{C_{\text{ref}}}\left[\frac{1}{|\Omega_p|}\sum_{\mathbf{x}\in\Omega_p}\max\left(s_{\text{now}}(\mathbf{x}) - s_i^{\text{ref}}(\mathbf{x}), 0\right)\right] g_i$, where C_{ref} is the refinement cost (i.e., the time-step a refinement simulation needs), $s_{\text{now}}(\mathbf{x})$ is the current GP posterior standard deviation, and $s_i^{\text{ref}}(\mathbf{x})$ is the posterior standard deviation after assuming that the refinement will reduce 30% uncertainty (σ) of the sample i. The factor g_i assigns more importance to samples that are closer to the current best solution.

To avoid spending computational effort on clearly irrelevant regions, the refinement benefit is evaluated only over a promising region identified from the GP predictive mean. For the minimization problem considered here, this region is defined as $\Omega_p = \{\mathbf{x} \in \mathcal{Q} \mid \mu(\mathbf{x}) \leq Q_{0.30}(\mu)\}$. Only the candidate points belonging to the lowest 30% of the predictive mean field are retained. If no existed candidate is involved, all candidates will be selected.

The action decision is made by comparing the best NEW score and the best REFINE score, $\text{Score}^*_{\text{new}} = \max_{\mathbf{x}} \text{Score}_{\text{new}}(\mathbf{x})$, $\text{Score}^*_{\text{refine}} = \max_i \text{Score}_{\text{refine}}(i)$. Refinement is selected only when $\text{Score}^*_{\text{refine}} > \text{Score}^*_{\text{new}}$. Otherwise, a NEW action starts. In addition, refinement is blocked when the selected sample already has sufficiently small relative uncertainty, i.e., $\frac{\sigma_i}{|r_i|} < 3\%$. Then, the optimizer is forced to find the new point by maximizing $\text{Score}_{\text{new}}(\mathbf{x})$.

Overall workflow:

In general, the optimization loop follows the sequence:

1. read the current BO sample list;
2. build the heteroscedastic GP surrogate;
3. compute $\text{Score}_{\text{new}}$ and $\text{Score}_{\text{refine}}$;
4. decide whether to execute **NEW** or **REFINE**;
5. post-process the force signal to obtain $r(\mathbf{q})$ and $\sigma_r(\mathbf{q})$;
6. update the BO database and convergence state.

Results and discussions

Figure 2 shows the convergence of $r(\mathbf{q})$. The black line is from the present study, while the red line is taken from , where a conventional BO-GPR framework without observation-dependent uncertainties is applied. Except for the uncertainty treatment, both approaches share the same numerical setups (e.g., mesh strategy and CFD solver) and optimization settings (e.g., design variables and initial samples). It shows that $r(\mathbf{q})$ decreases from 12.2 to approximately 0.8 in both cases, reaching a comparable minimum level. According to $r(\mathbf{q}) = \log\left(1+\frac{C_{L_{\text{rms}}}(\mathbf{q})D}{\varepsilon q_2}\right)$, the corresponding $C_{L_{\text{rms}}}$ is reduced by approximately four orders of magnitude, from 10^{-1} to 10^{-5}, effectively approaching zero.

To reach a comparable optimum, the present BO-GPR requires four more iterations than the conventional BO-GPR, as shown in figure 2. This is mainly because the present framework permits objective evaluations based on shorter CFD simulations, for which the associated uncertainties remain relatively high. As a result, the objective estimates in the early stage of the optimization are less accurate, and additional iterations are therefore needed to identify promising candidates and further reduce their uncertainties through the REFINE process. By contrast, the conventional BO-GPR relies on fully converged CFD simulations with uniformly low uncertainty, so that each evaluation is more accurate, although substantially more expensive.

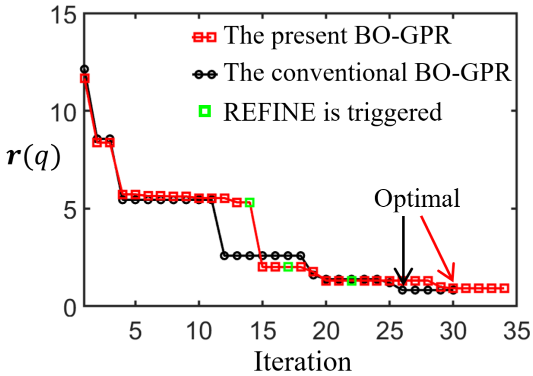

Figure 2. Convergence history of $r(\mathbf{q})$ for the present BO-GPR and the conventional BO-GPR (obtained from Tian *et al.* (2026)), with markers indicating the iterations where the REFINE process is triggered.

Despite the increase in the number of iterations, the present BO-GPR reduces the total computational cost by approximately 50% in terms of total time step (also reflecting the CPU hours) cost by all CFD simulations compared with the conventional BO-GPR, as shown in figure 3. In the conventional BO-GPR, each numerical simulation must be advanced for sufficiently long time steps to ensure that all cases reach a uniformly low uncertainty level. In contrast, the present BO-GPR explicitly accounts for the uncertainties induced by the relatively small time steps in each CFD simulation. This allows early termination of individual runs, resulting in substantial improvements in computational efficiency.

The green markers in figure 2 indicate the iterations where the REFINE process is triggered. REFINE is activated only 3 times out of 34 total iterations. This suggests that adopting fewer time steps per CFD simulation in the first place may enable more frequent REFINE steps, potentially leading to further reductions in computational cost. However, it should also be noted that fewer time steps may increase the number of required BO-GPR iterations due to larger uncertainty, and, in turn, the total optimization time. Therefore, how to better balance the cost of each individual simulation against the number of refinement and optimization iterations is a key issue for further improving the overall efficiency of the framework. This aspect will be explored in future work.

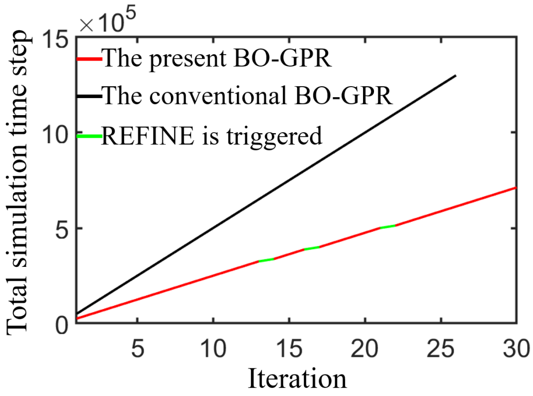

Figure 3. Accumulated total simulation time steps as a function of iteration for the present BO-GPR and the conventional BO-GPR, with REFINE-triggered iterations highlighted.

Figure 4(a) shows the predicted mean of the objective function $r(\boldsymbol{q})$ from the Gaussian Process surrogate model, while figure 4(b) shows the corresponding predicted standard deviation (uncertainty) of $r(q)$ with 95% confidence. In both plots, the black dashed lines connect the sequence of evaluated design points, illustrating the search trajectory of the BO-GPR algorithm. The black circles indicate the locations of individual sample points in the design space. At the optimal design, defined by $q_1 = l/D = 1.68$ and $q_2 = L/D = 0.4$, the corresponding predicted uncertainty is in the dark blue region (i.e. almost zero) in figure 4(b), indicating the convergence and reliability of BO-GPR at this location.

The objective landscape shown in figure 4(a) is clearly non-convex, with multiple low-response regions and strongly anisotropic variations. This is because the objective is derived from the RMS lift coefficient, which depends on the vortex dynamics and interactions in the near wake of the multi-step cylinder. Small changes in the lengths of the small (q_1) and large (q_2) cylinders could lead to substantial changes in the wake state and, consequently, in the lift fluctuations. This interpretation is further supported by figure 5, which shows vortex isosurfaces for different parameter combinations and corresponding levels of lift fluctuation. Figures 5 (c, d) show the final optimal geometries obtained by the conventional and the present BO-GPR, respectively. As the geometry changes from figure 5(a) to (d), both the final optimal geometries suppress the classical von Kármán vortex street and lead to negligible fluctuating lift force $C_{L,rms}$, although they are not identical. This indicates a near-optimal low-response region, where different BO trajectories may converge on distinct representative geometries while sharing the same flow states and physical mechanisms.

The multiple low-response regions seen in figure 4(a) represent different geometric configurations that achieve a similar low-lift flow regime, rather than fundamentally distinct optimal points. The non-convex shape does not just arise from the surrogate model but reflects the nonlinear and regime-dependent characteristics of the underlying flow physics. This inherent multi-modality in the wake flow highlights the value of Bayesian optimization for tackling global optimization problems that are expensive, non-convex, and noisy.

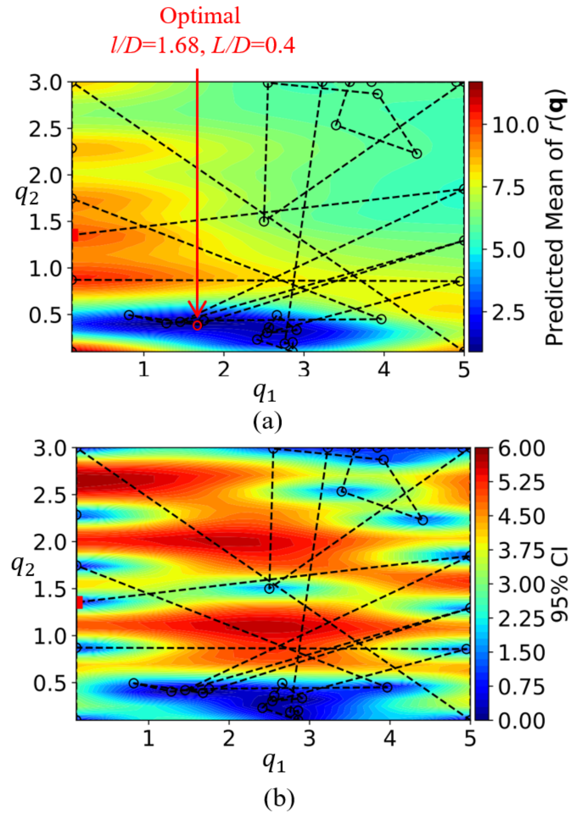

Figure 4. (a) The predicted mean of the objective function $r(q)$ from the Gaussian Process surrogate model; (b) The predicted standard deviation (uncertainty) of $r(q)$ with 95% confidence.

3 Conclusion

A heteroscedastic Bayesian optimization framework with Gaussian process regression was proposed to reduce flow-induced transverse loads on a multi-step cylinder while improving the efficiency of unsteady CFD-based optimization. In the proposed method, each evaluated sample is represented not only by its design parameters and objective value, but also by a case-dependent observation uncertainty estimated from the lift-force time series. This uncertainty is obtained using an autocorrelation-based integral-time-scale analysis and is incorporated directly into the Gaussian process surrogate. In contrast to a conventional BO-GPR procedure, the present framework can either evaluate a new design point or refine an existing sample, thereby improving computational efficiency.

The results show that the present BO-GPR reaches a comparable optimum to the conventional approach, with the objective decreasing from 12.2 to approximately 0.8, corresponding to an approximately four-order reduction in $C_{L_{\mathrm{rms}}}$ from 10^{-1} to 10^{-5}. The optimal design is found near $q_1 = l/D = 1.68$ and $q_2 = L/D = 0.4$, where the predicted uncertainty is very small, indicating reliable convergence. Although the proposed method entails four additional optimization iterations compared to conventional BO-GPR, the total computational cost is reduced by approximately 50% owing to the utilization of shorter CFD simulations in the early stages, with refinement applied only when necessary. Furthermore, the vortex structures indicate that the suppression of lift fluctuations is attributed to the disappearance of the classical von Kármán vortex street, which is consistent with previous publications.

One future work will focus on balancing shorter initial

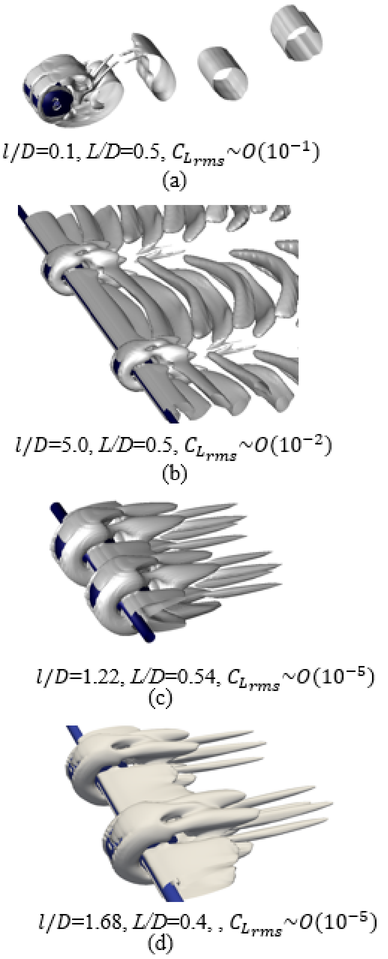

Figure 5. Instantaneous vortex structures behind the multi-step cylinder visualized by isosurface of $Q = 0.2$ (Hunt et al. 1988).

simulations, greater statistical uncertainty, and the resulting number of refinement and optimization iterations to further enhance the overall efficiency of the framework. Additionally, in the present study, the mesh is fixed for all CFD evaluations. Therefore, the uncertainty modeled here is the statistical uncertainty associated with finite-length unsteady force signals, rather than the spatial error of the CFD solver. This setup isolates the effect of observation-dependent statistical uncertainty on the BO process. A natural extension is to account for resolution-induced error as an additional uncertainty contribution, e.g. $\sigma^2_{\text{tot}} = \sigma^2_{\text{stat}} + \sigma^2_{\text{disc}}$, where σ_{disc} can be estimated from occasional refined-grid or refined-time-step calculations. An alternative is to treat numerical resolution itself as a fidelity variable within a multi-fidelity BO framework. Instead of merely penalizing low-resolution data with inflated noise, this approach models the correlation between different resolutions to actively correct the bias of coarse grids using sparse high-fidelity data. These will also be considered in future work.

3.1 Acknowledgment

The first author would like to thank the Alexander von Humboldt Foundation for financial support. The authors also gratefully acknowledge the scientific support and HPC resources provided by the Erlangen National High Performance Computing Center (NHR@FAU) of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU). The hardware is funded by the German Research Foundation (DFG).

REFERENCES

Dunn, W. & Tavoularis, S. 2006 Experimental studies of vortices shed from cylinders with a step-change in diameter. *Journal of Fluid Mechanics* **555**, 409–437.

Felisita, A., Gudmestad, O. T., Karunakaran, D. & Martinsen, L. O. 2017 Review of steel lazy wave riser concepts for the north sea. *Journal of Offshore Mechanics and Arctic Engineering* **139** (1), 011702.1–011702.15.

Jhingran, V., Zhang, H., Lie, H., Braaten, H. & Vandiver, J. K. 2017 Buoyancy spacing implications for fatigue damage due to vortex-induced vibrations on a steel lazy wave riser (slwr). In *Proceedings of OTC Offshore Technology Conference, Houston, Texas, USA, April 2012.*, pp. 369–384.

Ji, C., Cui, Y., Xu, D., Yang, X. & Srinil, N. 2019 Vortex-induced vibrations of dual-step cylinders with different diameter ratios in laminar flows. *Physics of Fluids* **31** (7).

Li, L., Fu, S., Yang, J., Ren, T. & Wang, X. 2017 Experimental investigation on vortex-induced vibration of risers with staggered buoyancy. In *Proceedings of the ASME 2011 30th International Conference on Ocean, Offshore and Arctic Engineering OMAE2011, Rotterdam, The Netherlands, June 19-24, 2011*, pp. 369–384.

Manhart, M. 2004 A zonal grid algorithm for DNS of turbulent boundary layers. *Computers & Fluids* **33**, 435–461.

Massaro, D., Peplinski, A. & Schlatter, P. 2023*a* Coherent structures in the turbulent stepped cylinder flow at Re= 5000. *International Journal of Heat and Fluid Flow* **102**, 109144.

Massaro, D., Peplinski, A. & Schlatter, P. 2023*b* The flow around a stepped cylinder with turbulent wake and stable shear layer. *Journal of Fluid Mechanics* **977**, A3.

Massaro, D. & Schlatter, P. 2024 Global stability of the flow past a stepped cylinder. *Journal of Fluid Mechanics* **988**, A1.

McClure, J., Morton, C. & Yarusevych, S. 2015 Flow development and structural loading on dual step cylinders in laminar shedding regime. *Physics of Fluids* **27**, 063602.

Mockett, C., Knacke, T. & Thiele, F. 2010 Detection of initial transient and estimation of statistical error in time-resolved turbulent flow data. In *In Proceedings of the 8th ERCOFTAC Symposium on Engineering Turbulence Modelling and Measurements (ETMM8)*, pp. 9–11. ERCOFTAC.

Morita, Y., Rezaeiravesh, S., Tabatabaei, N., Vinuesa, R., Fukagata, K. & Schlatter, P. 2022 Applying bayesian optimization with gaussian process regression to computational fluid dynamics problems. *Journal of Computational Physics* **449**, 110788.

Oliver, T. A., Malaya, N., Ulerich, R. & Moser, R. D. 2014 Estimating uncertainties in statistics computed from direct numerical simulation. *Physics of Fluids* **26** (3).

Ruan, W., Zhou, C., Yang, H., Wang, Z., Sun, B. & Bai, Y. 2024 A review on the recent process of lazy wave risers. *Journal of Marine Science and Engineering* **12** (11).

Tian, C., Jiang, F., Pettersen, B. & Andersson, H.I. 2017 Antisymmetric vortex interactions in the wake behind a step cylinder. *Physics of Fluids* **29** (10), 101704.

Tian, C., Jiang, F., Pettersen, B. & Andersson, H.I. 2020*a* Vortex dislocation mechanisms in the near wake of a step cylinder. *Journal of Fluid Mechanics* **891**, 393–441.

Tian, C., Massaro, D. & Schlatter, P. 2026 Bayesian optimization of the structural load of multi-step cylinders. *International Journal of Heat and Fluid Flow* **Under Review**.

Tian, C., Zhu, J., Jiang, F., Pettersen, B., Andersson, H.I. & Holmedal, L. E. 2020*b* Laminar flow over i-shaped dual-step cylinders. *Journal of Fluid Mechanics* **891**, 393–441.

Zhao, M., Zhang, Q. & Lei, C. 2024 A three-dimensional numerical investigation of vortex induced vibration of a step cylinder. *Physics of Fluids* **36** (6).

Zhu, C., Byrd, R. H., Lu, P. & Nocedal, J. 1997 Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software* **23** (4), 550–560.