Optimization and Engineering (2023) 24:2763–2772
https://doi.org/10.1007/s11081-023-09792-1

SHORT COMMUNICATION

Bayesian optimization for the spanwise oscillation of a gliding flat plate

Chunyu Wang[1,2] · Zhaoyue Xu[1,2] · Xinlei Zhang[1,2] · Shizhao Wang[1,2]

Received: 25 May 2022 / Revised: 11 February 2023 / Accepted: 11 February 2023 /
Published online: 28 February 2023

Abstract

Bayesian optimization is a powerful tool in the kinematic design of bio-inspired flight, where expensive high-fidelity numerical simulations are required. However, the traditional single acquisition strategy cannot adapt to all design problems, which have various input–output characteristics. The hybrid acquisition strategy improves the robustness by utilizing three types of acquisition functions, and each acquisition function is assigned three different balance parameters. The kinematics of a gliding flat plate that oscillates along the spanwise direction has been optimized to enhance the power efficiency by using the Bayesian optimization method. The power factor under the optimal spanwise oscillation is 1.97 times as large as that without spanwise oscillation, which indicates the good capability of the optimization method for bio-inspired locomotion.

Keywords Bayesian optimization · Hybrid acquisition strategy · Gaussian processes · Flapping wing · Micro air vehicles

✉ Shizhao Wang
wangsz@lnm.imech.ac.cn

Chunyu Wang
wangchunyu@imech.ac.cn

Zhaoyue Xu
xuzhaoyue@imech.ac.cn

Xinlei Zhang
zhangxinlei@imech.ac.cn

[1] The State Key Laboratory of Nonlinear Mechanics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China

[2] School of Engineering Sciences, University of Chinese Academy of Sciences, Beijing 101408, China

1 Introduction

The optimization design problem can be formulated as $\max_{\boldsymbol{\xi}\in\chi} f(\boldsymbol{\xi})$, where $\boldsymbol{\xi}$ is the design vector in space χ, and $f : \chi \to \mathbb{R}$ is the objective function to be optimized. In particular, we focus on the complex black-box system (Audet and Kokkolaras 2016), which features the implicit relation and the expensive target evaluations from the input to the output. A global efficient optimization is required to solve this kind of problem, which occurs in the design of unmanned aerial vehicles for optimal or compromised aerodynamic performances (Hassanalian and Abdelkefi 2017) and other aerospace scenarios (Borggaard and Burns 1997; Duraisamy et al. 2019; Hebbal et al. 2021; Nadarajah and Tatossian 2010).

Different optimization methods have been developed to find the best solution. Gradient-based methods have fast convergence but conduct local searches. Although the use of multiple start points allows the exploitation of more local regions, it causes an increasing computational burden. Evolutionary algorithms globally search the design space but need a large enough population size and generation size to converge. The unaffordable computational cost of the two methods described above can be significantly reduced by replacing the original physical model with the surrogate model (Marsden et al. 2004; Singh et al. 2017). The probabilistic surrogate model, which has remarkable advantages to expensive black-box problems, is used in the framework of Bayesian optimization (Hebbal et al. 2021; Shahriari et al. 2015). On the basis of a few sampled points, Bayesian optimization visualizes input–output relationships, estimates the optimum as well as its location, and suggests points to use for improving the next estimation by an acquisition strategy. However, different acquisition strategies have different effects in a given application. Despite limited experience, the choice of the acquisition strategy is still difficult when a new optimization task appears.

A feasible solution that avoids this choice is the hybrid acquisition strategy (Liu et al. 2021). The idea of this strategy is to incorporate several well-established acquisition functions into a portfolio and select the one with the best adaptability. Hoffman et al. (2011) proposed a full-information hedging strategy and demonstrated its superiority over the individual acquisition strategy through tests for standard functions, sampled functions, and a real physical problem. Vasconcelos et al. (2019) modified the algorithm by incorporating the memory factor and reward normalization. This method is expected to show high efficiency and robustness for a new biolocomotive application, where high fidelity numerical simulations are used to evaluate the aerodynamic performance. The aim of this work is to utilize the optimization method with high efficiency and robustness to find the optimal spanwise oscillating parameters for the power efficiency of a gliding rectangular flat plate.

The remainder of this paper is organized as follows. The Bayesian optimization with hybrid acquisition strategy is presented in Sect. 2. The optimization framework is applied to optimize the kinematics of the spanwise oscillating wing in Sect. 3. Finally, conclusions are drawn in Sect. 4.

2 Portfolio allocation framework of the Bayesian optimization method

2.1 Gaussian process model

The Gaussian process $GP(\mu_0, k)$ is constructed from the prior mean function $\mu_0 : \chi \mapsto \mathbb{R}$ and the positive-semidefinite covariance function $k : \chi \times \chi \mapsto \mathbb{R}$ (Shahriari et al. 2015). In this work, we adopt the smooth squared exponential covariance function $k(\boldsymbol{\xi}, \boldsymbol{\xi}') = {\sigma_f}^2 \exp\left(-\frac{1}{2l^2}\left\|\boldsymbol{\xi} - \boldsymbol{\xi}'\right\|^2\right)$, where ${\sigma_f}^2$ is the maximum covariance and l is the length parameter that controls the effect of the distance between the two design vectors $\boldsymbol{\xi}$ and $\boldsymbol{\xi}'$. The basic assumptions for the GP model can be written as

$$\mathbf{f}|\boldsymbol{\xi}_{1:n}, {\boldsymbol{\xi}^*}_{1:m} \sim \mathcal{N}(\mathbf{m}, \mathbf{K}), \tag{1}$$

$$\mathbf{q}|\mathbf{f}, \boldsymbol{\varepsilon} \sim \mathcal{N}\left(\mathbf{f}, \sigma^2\mathbf{I}\right), \tag{2}$$

where $\boldsymbol{\xi}_i$ is the known design vector (training point) and ${\boldsymbol{\xi}^*}_i$ is the unknown design vector (prediction point). Here, $\mathbf{f}$ and $\mathbf{q}$ are vectors consisting of unknown object values and noisy observation values on n training points $\boldsymbol{\xi}_{1:n}$ and m prediction points ${\boldsymbol{\xi}^*}_{1:m}$, respectively; $\mathbf{m}$ is the prior mean vector with elements $m_i = \mu_0\left(\hat{\boldsymbol{\xi}}_i\right)$, where $\hat{\boldsymbol{\xi}}_i$ is a training point or a prediction point; $\mathbf{K}$ is the prior covariance matrix with elements $K_{i,j} = k\left(\hat{\boldsymbol{\xi}}_i, \hat{\boldsymbol{\xi}}_j\right)$; and $\boldsymbol{\varepsilon}$ is the noise vector with its elements ε_i, which follow the Gaussian distribution $\mathcal{N}\left(0, \sigma^2\right)$. For ease of expression, some items in the relations (1) and (2) are expanded into partitioned forms, as shown below:

$$\mathbf{f} = \begin{bmatrix} f_{1:n} \\ {f^*}_{1:m} \end{bmatrix} = \begin{bmatrix} f\left(\boldsymbol{\xi}_{1:n}\right) \\ f\left({\boldsymbol{\xi}^*}_{1:m}\right) \end{bmatrix} \in \mathbb{R}^{n+m}, \ \mathbf{q} = \begin{bmatrix} q_{1:n} \\ {q^*}_{1:m} \end{bmatrix} \in \mathbb{R}^{n+m},$$

$$\mathbf{m} = \begin{bmatrix} m_{1:n} \\ {m^*}_{1:m} \end{bmatrix} = \begin{bmatrix} \mu_0\left(\boldsymbol{\xi}_{1:n}\right) \\ \mu_0\left({\boldsymbol{\xi}^*}_{1:m}\right) \end{bmatrix} \in \mathbb{R}^{n+m}, \ \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{1:n} \\ {\varepsilon^*}_{1:m} \end{bmatrix} \in \mathbb{R}^{n+m}, \tag{3}$$

$$\mathbf{K} = \begin{bmatrix} K_{n\times n} & {K^*}_{n\times m} \\ {K^*}_{m\times n} & {K^{**}}_{m\times m} \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}, \tag{4}$$

where $\mathbf{f}$, $\mathbf{q}$, $\mathbf{m}$ and $\boldsymbol{\varepsilon}$ are decomposed into the training and prediction subvectors and $\mathbf{K}$ is divided into three components $K_{n\times n}$, ${K^*}_{n\times m}$ (${K^*}_{m\times n}$ is the transpose of ${K^*}_{n\times m}$), and ${K^{**}}_{m\times m}$ interpreting the correlation of all pairs of points in $\left\{\left(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j\right)\right\}_{i,j=1:n}$, $\left\{\left(\boldsymbol{\xi}_i, {\boldsymbol{\xi}^*}_j\right)\right\}_{i=1:n,\ j=1:m}$, and $\left\{\left({\boldsymbol{\xi}^*}_i, {\boldsymbol{\xi}^*}_j\right)\right\}_{i,j=1:m}$, respectively.

Combining the relations (1) and (2) together with substituting Eqs. (3) and (4), we obtain

$$\begin{aligned} \begin{bmatrix} q_{1:n} \\ {q^*}_{1:m} \end{bmatrix} |\boldsymbol{\xi}_{1:n}, {\boldsymbol{\xi}^*}_{1:m} &= \left(\begin{bmatrix} f_{1:n} \\ {f^*}_{1:m} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1:n} \\ {\varepsilon^*}_{1:m} \end{bmatrix}\right) |\boldsymbol{\xi}_{1:n}, {\boldsymbol{\xi}^*}_{1:m} \\ &\sim \mathcal{N}\left(\begin{bmatrix} m_{1:n} \\ {m^*}_{1:m} \end{bmatrix}, \begin{bmatrix} K_{n\times n} + \sigma^2 I_n & {K^*}_{n\times m} \\ {K^*}_{m\times n} & {K^{**}}_{m\times m} + \sigma^2 I_m \end{bmatrix}\right). \end{aligned} \tag{5}$$

Then, we apply the property of conditioning Gaussians, resulting in the following distribution:

$$q^*{}_{1:m}|q_{1:n}, \boldsymbol{\xi}_{1:n}, \boldsymbol{\xi}^*{}_{1:m} \sim \mathcal{N}\left(\mu^*{}_{1:m}, \Sigma^*{}_{m\times m}\right), \tag{6}$$

where $\mu^*{}_{1:m}$ is the posterior mean vector and $\Sigma^*{}_{m\times m}$ is the posterior covariance matrix. They are expressed as

$$\mu^*{}_{1:m} = m^*{}_{1:m} + K^*{}_{m\times n}\left(K_{n\times n} + \sigma^2 I_n\right)^{-1}\left(q_{1:n} - m_{1:n}\right), \tag{7}$$

$$\Sigma^*{}_{m\times m} = \left(K^{**}{}_{m\times m} + \sigma^2 I_m\right) - K^*{}_{m\times n}\left(K_{n\times n} + \sigma^2 I_n\right)^{-1} K^*{}_{n\times m}. \tag{8}$$

Once the training set $\mathcal{D}_{1:n} = \left\{\left(\boldsymbol{\xi}_i, q_i\right)\right\}_{i=1:n}$ is given, the best estimate for $q^*{}_{1:m}$ and its uncertainty are calculated through Eqs. (7) and (8), respectively, which is the modeling of the GP. In particular, when only one prediction point is considered ($m = 1$), we can derive the posterior mean function $\mu\left(\boldsymbol{\xi}\right)$ and variance function $\sigma^2\left(\boldsymbol{\xi}\right)$ (or the standard deviation function $\sigma\left(\boldsymbol{\xi}\right)$). It is noted that we have used type II maximum likelihood (empirical Bayes) to optimize the covariance function hyperparameters for improving the prediction of the surrogate model. The hyperparameter optimization is handled by the efficient L-BFGS-B algorithm with given bounds for each hyperparameter.

2.2 Hybrid acquisition strategy

After the probabilistic surrogate model is built, we use the statistical information provided by the model to create the acquisition function reflecting the acquisition strategy. Maximizing this acquisition function can help to search for the promising optimal point, which forms the suboptimization problem that is formulated as

$$\boldsymbol{\xi}_{n+1} = \arg\max_{\boldsymbol{\xi}\in\chi} \alpha\left(\boldsymbol{\xi}; \mathcal{D}_{1:n}\right). \tag{9}$$

Here, we employ three confidence information-assisted strategies, including probability of improvement (PI), expected improvement (EI), and upper confidence bound (UCB). For an arbitrary design vector $\boldsymbol{\xi}$, these acquisition functions are expressed as follows:

$$\alpha_{PI}\left(\boldsymbol{\xi};\ \mathcal{D}_{1:n}\right) = prob\left(f\left(\boldsymbol{\xi}\right) \geq \tau + \zeta_{PI}\right) = \Phi\left(\frac{\mu\left(\boldsymbol{\xi}\right) - \tau - \zeta_{PI}}{\sigma\left(\boldsymbol{\xi}\right)}\right), \tag{10a}$$

$$\alpha_{EI}\left(\boldsymbol{\xi};\ \mathcal{D}_{1:n}\right) = \begin{cases} \left(\mu\left(\boldsymbol{\xi}\right) - \tau - \zeta_{EI}\right)\cdot\Phi\left(\frac{\mu(\boldsymbol{\xi})-\tau-\zeta_{EI}}{\sigma(\boldsymbol{\xi})}\right) & \\ \quad +\sigma\left(\boldsymbol{\xi}\right)\cdot\phi\left(\frac{\mu(\boldsymbol{\xi})-\tau-\zeta_{EI}}{\sigma(\boldsymbol{\xi})}\right) & \sigma\left(\boldsymbol{\xi}\right) > 0 \\ 0 & \sigma\left(\boldsymbol{\xi}\right) = 0 \end{cases}, \tag{10b}$$

$$\alpha_{UCB}\left(\boldsymbol{\xi};\ \mathcal{D}_{1:n}\right) = \mu\left(\boldsymbol{\xi}\right) + \zeta_{UCB}\cdot\sigma\left(\boldsymbol{\xi}\right), \tag{10c}$$

where τ is the incumbent optimal target, $\Phi\left(\boldsymbol{.}\right)$ is the standard normal cumulative distribution function, and $\phi\left(\boldsymbol{.}\right)$ is the standard normal probability density function.

In addition, ζ_{PI}, ζ_{EI}, and ζ_{UCB} denote the balance parameters used in the tradeoff between global exploration and local exploitation.

Different acquisition criteria have different adaptability to a model with specific spatial characteristics. Moreover, the preferred strategy may change with the advancement of sequential optimization. Therefore, compared with using the single constant acquisition function, a better alternative is to dynamically pick a superior function from the prescribed portfolio. The robust strategy is leveraged in this study. We define an acquisition function portfolio that contains the PI, EI, and UCB. Each type of acquisition function is assigned three different balance parameters, i.e. {0.00, 0.01, 0.10} for PI, {0.00, 0.01, 0.10} for EI, and {1.00, 1.50, 2.00} for UCB.

2.3 Framework integration

The GP-Hedge algorithm (Hoffman et al. 2011) and its modification (Vasconcelos et al. 2019) adopt the portfolio allocation framework to improve the robustness of the Bayesian optimization. The frame structure is detailed as follows. The superscript n represents the iteration index ($n = 0$ means objects generated in the initialization) and the subscript j represents the acquisition function index ($j = 1, 2, \cdots, 9$). The coarse optimization landscape is needed to find a reasonable acquisition function in the beginning and ensure a sufficient optimization efficiency, so we incorporate the design of experiments (DOE) based on a space-filling sampling criterion into the initialization. The technique used here is the optimal Latin hypercube method where the space-filling sampling criterion makes sampling points evenly distributed in the design space. For the design vector with two variables, that is, $\boldsymbol{\xi}_i = \left[A_i \; k_i\right]^T$, 20 points are sampled by the DOE technique and evaluated by the CFD simulations, resulting in the initial training set $\mathcal{D}^0 = \left\{\left(\boldsymbol{\xi}_i, q_i\right)\right\}_{i=1:n_0}$ ($n_0 = 20$). Then, we create the initial surrogate model GP^0 (its posterior mean function and covariance function are $\mu^0(\boldsymbol{\xi})$ and $\sigma^0(\boldsymbol{\xi})$, respectively) and the corresponding acquisition functions $\alpha_j^0(\boldsymbol{\xi})$ by Gaussian process regression (GPR). As in the original algorithm, the initial gains G_j^0 are set to 0.

The aforementioned initialization is followed by a loop composed of 'suggest', 'evaluate', and 'update' steps. The first step is used to query the potential point $\boldsymbol{\xi}_{j_{best}}{}^n$ from the candidates $\boldsymbol{\xi}_j{}^n$ ($j = 1, 2, \cdots, 9$) according to the specified criterion, the second step is used to evaluate the selected point $\boldsymbol{\xi}_{j_{best}}{}^n(\boldsymbol{\xi}_{n_0+n})$ through the numerical simulation (CFD) for obtaining its target value q_{n_0+n}, and the third step is used to sequentially update the data-set $\mathcal{D}^n$, the surrogate model GP^n, the acquisition functions $\alpha_j{}^n$, and the gains $G_j{}^n$. It needs to be emphasized that differences exist between iterations $n = 1$ and $n > 1$. In the 'suggest' step, the criterion to select $\boldsymbol{\xi}_{j_{best}}{}^n$ when $n = 1$ is maximizing the initial posterior mean function $\mu^0(\boldsymbol{\xi})$, while the point $\boldsymbol{\xi}_{j_{best}}{}^n$ can be found with the probability $P_j{}^n = \exp\left(\eta \cdot g_j{}^{n-1}\right) / \sum_{j'=1}^{9} \exp\left(\eta \cdot g_j'{}^{n-1}\right)$, where hyperparameter $\eta = 4.0$ and normalized gains $g_j{}^{n-1} = [G_j{}^{n-1} - \max_j\left(G_j{}^{n-1}\right)] / [\max_j\left(G_j{}^{n-1}\right) - \min_j\left(G_j{}^{n-1}\right)]$, when $n > 1$. In the 'update' step, the gains $G_j{}^n$ are equal to the current rewards $r_j{}^n = \mu^n\left(\boldsymbol{\xi}_j{}^n\right)$ for $n = 1$, while $G_j{}^n = r_j{}^n + m \cdot G_j{}^{n-1}$ for $n > 1$. Here, the memory factor m of 0.70 is used to reduce the effect of the

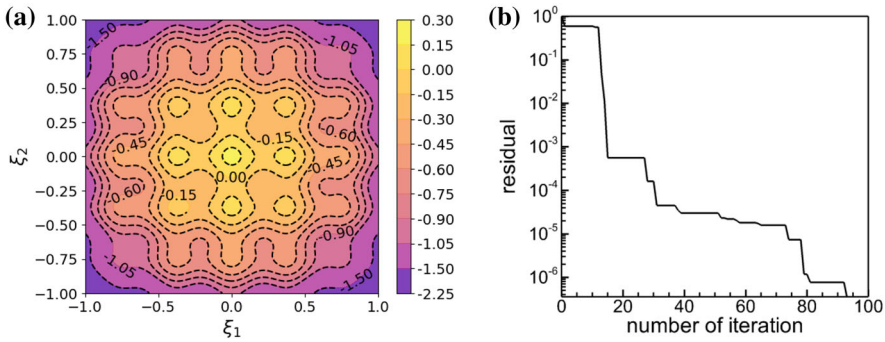

Fig. 1 Validation of the Bayesian optimization framework. **a** Contour of the cosine mixture function used to test the framework. **b** Convergence history of the test function optimization

past rewards. In this study, the loop stops when the given number of iterations is reached. Alternatively, the automated stopping criterion based on several properties may be better for the trade-off between optimization accuracy and computational cost. In addition, the manual stopping mode can also be optional to increase the flexibility of operation, especially when problem-specific expert knowledge is involved.

The setup of the framework is tested by the cosine mixture function $f(\boldsymbol{\xi}) = 0.1\left[\cos\left(5\pi\xi_1\right) + \cos\left(5\pi\xi_2\right)\right] - \left({\xi_1}^2 + {\xi_2}^2\right)$ (see Fig. 1a), a benchmark function with comparable optimization complexity for the numerical example in next section. The convergence history shown in Fig. 1b verifies the validity of the method. Furthermore, when the optimization loop iterates only 15 times, the incumbent maximum has a value of 0.199836. Therefore, the maximum number of iterations is set to 15, which is considered sufficient in the current study.

3 Spanwise oscillating optimization of a gliding flat-plate wing

We consider a simplified geometric model in typical flight conditions as follows. A rectangular flat plate with an aspect ratio ($AR = 2$) glides forward at a constant speed U and angle of attack ($AoA = 25°$), as shown in Fig. 2a. At the same time, different spanwise oscillations characterized by oscillation amplitude and frequency are imposed on the plate. Following the definition of spanwise motion in Wang et al. (2015), the center of the plate changes its position with time in harmonic form as

$$y(t) = A\sin(2kt), \tag{11}$$

where y is the coordinate in the spanwise direction, t is the time, A is the oscillation amplitude, and k is the reduced frequency (the four variables are non-dimensional). We define a design space χ with two design variables of A and k. Each design variable is set between 0 and 1 in the current work.

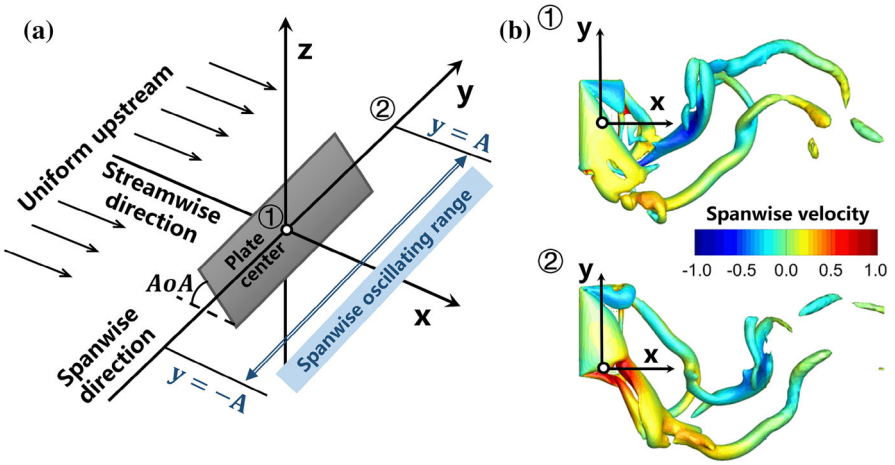

Fig. 2 Schematic of the gliding flat plate with spanwise oscillation (**a**) and the vortex structures around it (**b**). In subplot **b**, the flow structures at the equilibrium position ① and the positive maximum displacement position ② are identified by a Q-criterion of 0.75

The evaluation of the unsteady aerodynamic target function is based on the incompressible Navier–Stokes equations in the following dimensionless form:

$$\nabla \cdot \mathbf{u} = 0, \tag{12}$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \frac{1}{\mathrm{Re}} \nabla^2 \mathbf{u} + \mathbf{f}_b, \tag{13}$$

where $\mathbf{u}$ is the flow velocity vector and p is the static pressure. Re is the Reynolds number based on the chord length and the freestream velocity. We investigate the flow under $\mathrm{Re} = 300$ because the unsteady effect dominates the flow. $\mathbf{f}_b$ is the Eulerian force density used for the immersed boundary method (Wang and Zhang 2011). Numerically, the time advancement and the spacial discretization are implemented by the second-order Runge–Kutta scheme and the second-order finite volume formulation, respectively. The flow around the flat-plate wing can be fully resolved by the above strategies. The numerical treatment and mesh are reported at great length in previous work (Wang et al. 2015; Wang and Zhang 2011). Accordingly, we can obtain the time history of the drag coefficient ($C_D(t)$), the lift coefficient ($C_L(t)$), and the side-force coefficient ($C_S(t)$). The efficiency of endurance is considered here. The derived power factor is the objective function measuring the efficiency. Its expression is

$$PF = \frac{\overline{C_L(t)}^{1.5}}{\overline{C_D(t)} + \frac{1}{T}\int_0^T C_S(t)\, 2Ak\,(-\cos(2kt))\, dt}, \tag{14}$$

where the overbar denotes the period-averaged force coefficient and T is the dimensionless oscillating period. Briefly, the aerodynamic optimization problem can be formulated such that (1) the design variables are the oscillating amplitude A and the

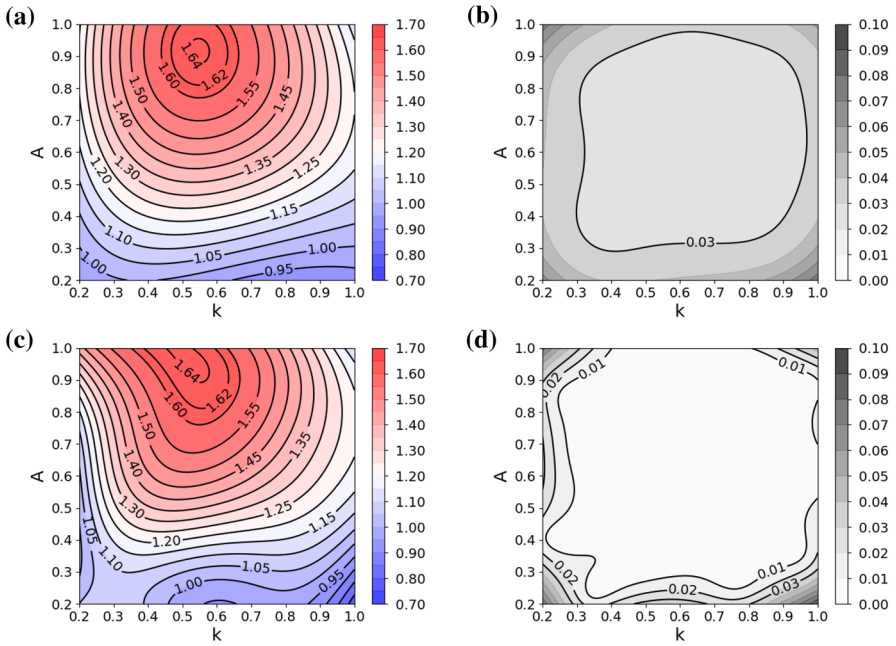

Fig. 3 Distribution of the power factor in the spanwise oscillating parameter space (k, A). **a** Mean (Initialization). **b** Uncertainty (Initialization). **c** Mean (The 15*th* iteration). **d** Uncertainty (The 15*th* iteration)

oscillating frequency k, and (2) the objective function to be maximized is the power factor PF.

The iteration process advances with two processes: searching for the optimum and refining the model. Figure 3 a and c show the estimation for the power factor at the initialization and the 15*th* iteration, respectively. The optimization locates the maximal power factor of 1.65 (which is 1.97 times as large as that under the pure glide condition) at the point where $A = 1.00$ and $k = 0.49$. The vortex structures under the optimal parameter configuration are identified with the Q-criterion, as is shown in Fig. 2b. The estimation uncertainty of the power factor is exhibited in Fig. 3b and d. The standard deviation interpreting the uncertainty decreases from the range (0.02, 0.03) to the range (0.00, 0.01) in the most part of investigated region when the GP model develops from the initialization to the 15*th* iteration. This indicates that the accuracy of the model improves with the involvement of additional points.

The Bayesian optimization method also has advantages explained by the following examples. Figure 4 shows the simulated distribution of the power factor. The power factor values on the scattered points come from the evaluation of the GP model, and these displayed contours are drawn with the hypothesis that the evaluation values are true values. The contour lines are far from smoothness in Fig. 4a, where 81 sampled points are evaluated. The increase of the sampled points to 289 improves the smoothness of the contour lines, as shown in Fig. 4b. It is seen that simple parameter studies require more sample points (only 35 sample points are used in Fig. 3c and d) to obtain

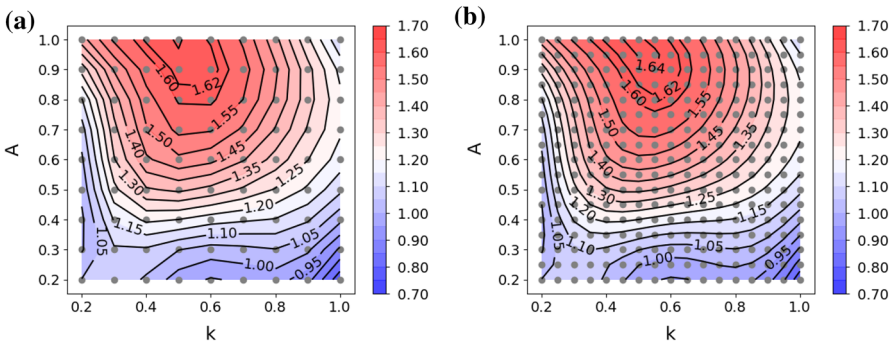

Fig. 4 Design points (gray circles) obtained by uniform sampling and corresponding power factor contours. **a** Sparse sampling (9 × 9 sampled points). **b** Dense sampling (17 × 17 sampled points)

comparative information about the power factor distribution and to find the approximate optimal value. This highlights the efficiency of exploring the design space based on the present optimization method.

4 Conclusion

The Bayesian optimization method with the hybrid acquisition strategy is pioneered to be applied in the bio-inspired locomotion. The hybrid acquisition strategy improves the robustness of the method with three types of acquisition functions and three different balance parameters for each type of acquisition function. The portfolio allocation framework is used to design the spanwise oscillation of a gliding flat plate with the objective of maximizing the power factor. A maximum power factor of 1.65 is obtained at the non-dimensional oscillating amplitude of 1.00 and the reduced frequency of 0.49. The results show that the Bayesian optimization method is applicable to optimize the kinematics of the bio-inspired flight. The more complex kinematics of bio-inspired micro air vehicles is expected to be designed by this Bayesian optimization method. Intelligent hybrid acquisition strategies with different memory modes for past rewards will be investigated in the next work, because certain suggested design points at the iterations far from the latest iteration may have an important effect on the optimization process. Another future research focus is developing the multi-objective optimization capabilities of the current framework.

Acknowledgements This work is supported by the NSFC Basic Science Center Program for 'Multiscale Problems in Nonlinear Mechanics' No. 11988102, the National Natural Science Foundation of China Nos. 11922214 and 92252203. The computations are conducted on Tianhe-1 at the National Supercomputer Center in Tianjin.

References

Audet C, Kokkolaras M (2016) Blackbox and derivative-free optimization: theory, algorithms and applications. Optim Eng 17(1):1–2. https://doi.org/10.1007/s11081-016-9307-4

Borggaard J, Burns J (1997) A PDE sensitivity equation method for optimal aerodynamic design. J Comput Phys 136(2):366–384. https://doi.org/10.1006/jcph.1997.5743
Duraisamy K, Iaccarino G, Xiao H (2019) Turbulence modeling in the age of data. Annu Rev Fluid Mech 51:357–377. https://doi.org/10.1146/annurev-fluid-010518-040547
Hassanalian M, Abdelkefi A (2017) Classifications, applications, and design challenges of drones: a review. Prog Aerosp Sci 91:99–131. https://doi.org/10.1016/j.paerosci.2017.04.003
Hebbal A, Brevault L, Balesdent M et al (2021) Bayesian optimization using deep Gaussian processes with applications to aerospace system design. Optim Eng 22(1):321–361. https://doi.org/10.1007/s11081-020-09517-8
Hoffman M, Brochu E, de Freitas N (2011) Portfolio allocation for Bayesian optimization. In: Proceedings of the Twenty-Seventh conference on uncertainty in artificial intelligence, pp 327–336
Liu Z, An W, Qu X et al (2021) Portfolio-based Bayesian optimization for airfoil design. AIAA J 59(6):1975–1989. https://doi.org/10.2514/1.J059812
Marsden AL, Wang M, Dennis JE et al (2004) Optimal aeroacoustic shape design using the surrogate management framework. Optim Eng 5(2):235–262. https://doi.org/10.1023/B:OPTE.0000033376.89159.65
Nadarajah SK, Tatossian C (2010) Multi-objective aerodynamic shape optimization for unsteady viscous flows. Optim Eng 11(1):67–106. https://doi.org/10.1007/s11081-008-9036-4
Shahriari B, Swersky K, Wang Z et al (2015) Taking the human out of the loop: a review of Bayesian optimization. Proc IEEE 104(1):148–175. https://doi.org/10.1109/JPROC.2015.2494218
Singh P, Couckuyt I, Elsayed K et al (2017) Multi-objective geometry optimization of a gas cyclone using triple-fidelity co-Kriging surrogate models. J Optim Theory Appl 175(1):172–193. https://doi.org/10.1007/s10957-017-1114-3
Vasconcelos TdeP, de Souza DA, Mattos CL, et al (2019) No-PASt-BO: Normalized portfolio allocation strategy for Bayesian optimization. In: 2019 IEEE 31st international conference on tools with artificial intelligence (ICTAI), IEEE, pp 561–568
Wang S, Zhang X (2011) An immersed boundary method based on discrete stream function formulation for two-and three-dimensional incompressible flows. J Comput Phys 230(9):3479–3499. https://doi.org/10.1016/j.jcp.2011.01.045
Wang S, He G, Zhang X (2015) Lift enhancement on spanwise oscillating flat-plates in low-Reynolds-number flows. Phys Fluids 27(6):061,901. https://doi.org/10.1063/1.4922236