Struct Multidisc Optim (2017) 55:925–943
DOI 10.1007/s00158-016-1546-7

RESEARCH PAPER

Efficient aerodynamic shape optimization of transonic wings using a parallel infilling strategy and surrogate models

J. Liu[1] **· W.-P. Song**[1] **· Z.-H. Han**[1] **· Y. Zhang**[1]

Received: 19 January 2015 / Revised: 14 June 2016 / Accepted: 20 July 2016 / Published online: 4 August 2016

Abstract Surrogate models are used to dramatically improve the design efficiency of numerical aerodynamic shape optimization, where high-fidelity, expensive computational fluid dynamics (CFD) is often employed. Traditionally, in adaptation, only one single sample point is chosen to update the surrogate model during each updating cycle, after the initial surrogate model is built. To enable the selection of multiple new samples at each updating cycle, a few parallel infilling strategies have been developed in recent years, in order to reduce the optimization wall clock time. In this article, an alternative parallel infilling strategy for surrogate-based constrained optimization is presented and demonstrated by the aerodynamic shape optimization of transonic wings. Different from existing methods in which multiple sample points are chosen by a single infill criterion, this article uses a combination of multiple infill criteria, with each criterion choosing a different sample point. Constrained drag minimizations of the ONERA-M6 and DLR-F4 wings are exercised to demonstrate the proposed method, including low-dimensional (6 design variables) and higher-dimensional problems (up to 48 design variables). The results show that, for surrogate-based optimization of transonic wings, the proposed method is more effective than the existing parallel infilling strategies, when the number of initial sample points are in the range from N_v to $8N_v$ (N_v here denotes the number of design variables). Each case is repeated 50 times to eliminate the effect of randomness in our results.

Keywords Aerodynamic shape optimization · Surrogate-based optimization · Surrogate models · Kriging · Wing design

✉ Z.-H. Han
hanzh@nwpu.edu.cn

J. Liu
liu.jun@csu.edu.cn

W.-P. Song
wpsong@nwpu.edu.cn

Y. Zhang
momozhang@126.com

[1] National Key Laboratory of Science and Technology on Aerodynamic Design and Research, School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China

1 Introduction

Nowadays, numerical optimization is playing an increasingly important role in aerodynamic design of aircraft, where high-fidelity computational fluid dynamics (CFD) is often employed. However, the main challenge is associated with the computational cost of the optimization process, which can be prohibitive when a large number of expensive CFD simulations are required. Therefore, the development of efficient aerodynamic shape optimization methods is of great interest.

CFD-driven numerical optimization methods can be classified into two categories: gradient-based methods with the gradients computed by the adjoint method (Jameson 1988 & 1997) and gradient-free methods. The gradient-based methods are usually very effective, but the solution optimality can be sensitive to the initial guesses and the method often becomes trapped into a local minimum (Chernukhin and Zingg 2013). Since aerodynamic functions are usually multi-modal (Keane 2006; Laurenceau 2008), gradient-free methods capable of finding global optimum are of great interest for an aerodynamic shape optimization. However, the computational cost associated with global optimization methods, such as genetic

algorithms (GAs) or particle swarm optimization (PSO), could easily become prohibitive with the increase of the number of design variables, if high-fidelity CFD is employed for functional evaluation.

The efficiency of a global optimization can be dramatically improved by using surrogate models (Jones 1998), such as polynomial regression (Vavalle 2007), radial basis functions (Sobester 2005), kriging (Forrester 2009), and support vector regression (Yun 2009), etc. A surrogate model is essentially an approximation model for the cost function or constraint function, which is built from the limited information obtained by probing (or sampling) the design space. Once the surrogate model is built, it can be used to replace the expensive CFD simulation for predicting the responses of the cost function or constraint function during the optimization process. Since the computational cost of evaluating a surrogate model is negligible, compared to a CFD simulation, the optimization efficiency could be greatly improved. But for this strategy of simply replacing CFD simulations with the surrogate models, the optimization results heavily rely on the global approximation accuracy. Especially, building a sufficiently accurate surrogate model for a high-dimensional aerodynamic function needs a huge number of CFD-evaluated sample points, which limits the use of this strategy to the problems with only a few design variables (less than around 10). A remedy is to use adaptive sampling techniques, in which initial surrogate models are built first and then new sampling data is repetitively added to update the models, either to improve the current optimal design (exploitation) or to improve the global accuracy of the model (exploration) (Forrester 2009; Koziel 2013). In general, this method only uses a small number of initial sample points, and adaptively chooses new sample point (s) in the vicinity of the optimum or the promising regions of the design space. In turn, the computation cost of finding the optimum can be dramatically reduced.

For an aerodynamic shape optimization, the initial surrogate models are usually not accurate enough, since a small number of initial sample points are not sufficient to represent the numerous peaks and valleys for typical aerodynamic functions. Consequently, the strategies about how to determine the sites of new sample points in the design space, i.e. infilling strategies (also called infill criteria), are crucial for the success of a surrogate-based optimization (Forrester 2009; Yao 2013). Currently, a few infill criteria (Forrester 2009; Jones 2001) are available, such as minimizing the predicted objective function (MP), maximizing the expected improvement function (EI) (Jones 2001), maximizing the probability of improvement function (PI) (Forrester 2009), minimizing the lower confidence bound (LCB) (Laurenceau 2010), maximizing the mean squared error (MSE) (Forrester 2009), etc. The EI method in conjunction with a kriging model, called efficient global optimization (EGO) method (Jones 1998), quickly gained popularity for aerodynamic and multidisciplinary design optimizations (Simpson 2001; Jeong 2005; Hoyle 2006; Goto 2008; Han 2010a, b and 2013; Song 2007; Kanazaki 2007). It is noted that, for the most of the researches, a single infill criterion is employed and only one new sample point is chosen and evaluated at each updating cycle. Obviously, this is not economic for the modern computers which always have multiple CPU cores. If multiple new sample points could be chosen and multiple CFD simulations are performed in parallel, the total wall-clock time for an optimization process can be dramatically reduced.

To this end, several parallel infill strategies have been made available in recent years. Ginsbourger et. al. (2007) developed a multi-point infilling criterion (called q-EI) using a generalized EI method defined by Schonlau (1997). In the method, q points are chosen in each updating cycle. However, a highly expensive computational method such as Monte-Carlo simulation is needed to derive the solutions of multiple EI criteria when $q>2$. Therefore, they proposed to use the concept of approximated q-EI method to reduce the computational burden. Viana et al. (2010) used the PI criterion to select n points from Monte Carlo sampling in the analytical test cases, and they found that the more sample points are infilled at each updating cycle, the less iterations are usually needed. Sobester (2004) used the method to pick out the best N_p local maxima of the EI function as new sample points to update the RBFs model and applied it to structural optimization. Laurenceau (2010) developed a method in which 3 sample points were chosen simultaneously at each updating cycle, by setting 3 different values for the user-defined parameter of a LCB criterion. This method was demonstrated by aerodynamic shape optimization of airfoils and wings. Parr (2012) proposed the method of choosing multiple points from the Pareto-front set of maximizing both the expected improvement (EI) and the probability of satisfying the constraints.

It is noted that these parallel infilling strategies use a single infill criterion to choose multiple sample points, and therefore the inherent drawback of this criterion could have a significant impact on the results. For example, the MP tends to exploit the region near the current best design and when it accurately converges to a local optimum, it is trapped there; the EI has a global performance but its local convergence is not satisfactory; for the LCB, it is hard to set the user-defined parameter for a new application; and for the MSE, it is a waste of sample points to purely increase the global accuracy of the surrogate model, even a local optimum can't be found if it is used in isolation. If multiple new sample points are chosen by a single infill criterion to refine the surrogate model at each updating cycle such as Laurenceau (2010), Schonlau (1997), Viana (2010), and Sobester (2004), the weakness of this infill criterion still exists. An alternative way is to choose multiple points by using multiple infill criteria. Yao (2013) developed a hybrid infill strategy by combining the LCB and DLI (Divergence from local Linear Interpolation) criteria.

Chaudhuri (2015) used the method of choosing multiple points by combining the PI and EI criterion, which was applied to the optimization of the experimental flapping wing and uncertainty quantifications.

Although the use of parallel infill strategies tends to a standard practice for a surrogate-based optimization, they haven't been well studied for higher-dimensional engineering problems, especially for strongly constrained aerodynamic design optimization with large number of design variables. This article is motivated by the inspiration of developing an alternative parallel infilling strategy for the surrogate-based global optimization and demonstrating it by low- and higher-dimensional constrained aerodynamic shape optimization of transonic wings. First, the existing infill sampling criteria such as EI, PI, LCB, and MP are extended to the constrained form, which are to be used simultaneously at each updating cycle of a surrogate-based optimization. Second, the parallel infilling strategy proposed in this article is verified by representative analytical test cases, and compared with existing parallel infilling strategies as well as the traditional single-point infilling criterion. Then, drag minimizations of the ONERA-M6 and DLR-F4 wings with multiple constraints are exercised to demonstrate the proposed method. The optimization of the M6 wing is performed in a comprehensive way: first, the planform shape of the wing is optimized with a fixed sectional shape (6 design variables); then, the sectional shapes are optimized with a fixed planform shape (24 design variables); third, the planform and sectional shapes are optimized simultaneously (30 design variables). The proposed method is further demonstrated by a higher-dimensional aerodynamic design problem (48 design variables), the design of a transport aircraft wing, DLR-F4, and also compared with three reference parallel infilling strategies. The results show that the proposed method is more effective than the exiting methods, with faster convergence and slightly lower optimal drag (2–3 drag counts).

2 Surrogate-based optimizer with parallel infilling strategy

An in-house surrogate-based generic optimizer called SurroOpt (Liu 2012a; Han 2013; Zhang 2016) is used for all the optimization study in this article. A number of surrogate models such as kriging, polynomial response surface model, radial basis functions, support vector regression, are available in the code and the kriging model is used in this study.

This section describes the main ingredients of the surrogate-based parallel optimizer in which the parallel infilling strategy is used, including the design of experiments, surrogate model, multiple-sample infill strategy, and the flowchart of the surrogate-based optimizer employed in this article.

2.1 Design of experiments

Design of experiments (DoE) (Giunta 2003) is used to generate sample points in the design space for constructing initial surrogate models. To obtain a surrogate model as accurate as possible with a limited number of pre-sampled points, the space-filling design is usually favored. In this article, Latin hypercube sampling (McKay 1979) is employed. Assuming that we have n_s sample points and N_v design parameters, the range of each parameter is divided into n_s bins with equal probability, resulting in $(N_v)^{n_s}$ bins in total. Then the n_s samples are randomly selected in the parameter space subject to the following requirements: (1) each sample is randomly placed inside a bin; (2) when projected to one dimension, there is no two points existing in the same bin. Assuming that the range of each parameter is [0, 1], the sample points can be obtained using the following formulation:

$$x_j^{(i)} = \frac{\pi_j^{(i)} + U_j^{(i)}}{n_s}, \quad 1 \le j \le N_v, \quad 1 \le i \le n_s, \tag{1}$$

where i denotes the i^{th} sample point, j denotes the j^{th} design variable, and U is a random number in [0,1]. For each j, π_j denotes a random permutation of $\{0, 1, \cdots, n_s - 1\}$.

2.2 Surrogate modelling

Kriging is a geostatistical interpolation method suggested by Krige (1951) and mathematically formulated by Matheron (1963). Kriging was widely used in the context of geostatistics. Another milestone in the development of kriging model is that in 1989, kriging was extended by Sacks et al. (1989) to the design and analysis of deterministic computer experiments. Then it was widely used as a surrogate modeling technique for predicting the output of a computer code in simulation-based analysis and optimization. There are several variants of the kriging model, in which ordinary kriging is the most commonly used one. The following is a brief description of ordinary kriging. The readers are referred to Ref. (Han et al. 2010a, b, 2012a, b, 2013; Han and Goertz 2012) for other kinds of kriging models.

2.2.1 Kriging predictor and mean squared error

Kriging treats the output of a deterministic computer experiment as the realization of a random process

$$Y(\mathbf{x}) = \beta_0 + Z(\mathbf{x}). \tag{2}$$

The stationary random process $Z(\mathbf{x})$ has mean zero and covariance of

$$Cov\left[Z(\mathbf{x}), Z\left(\mathbf{x}'\right)\right] = \sigma^2 R\left(\mathbf{x}, \mathbf{x}'\right), \tag{3}$$

where σ^2 is the process variance of $Z(\mathbf{x})$. It is assumed that $\sigma^2(\mathrm{x})\equiv\sigma^2$ for all $\mathbf{x}$ stationarity. And R is the spatial correlation function that only depends on the Euclidean distance between two sites $\mathbf{x}$ and $\mathbf{x}'$.

The observed functional responses are denoted $\mathbf{y}_s = \left(y^{(1)},\ldots,y^{(n_s)}\right)^{\mathrm{T}}$. Following the derivation of (Sacks 1989; Liu 2012a; Han 2012a, b and 2013), we can obtain the following kriging predictor

$$\hat{y}(\mathbf{x}) = \hat{\beta}_0 + \mathbf{r}^{\mathrm{T}}(\mathbf{x})\mathbf{R}^{-1}\left(\mathbf{y}_s - \mathbf{1}\hat{\beta}_0\right), \tag{4}$$

where $\mathbf{1}$ is a unit column vector filled with ones and

$$\hat{\beta}_0 = \left(\mathbf{1}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{1}\right)^{-1}\mathbf{1}\mathbf{R}^{-1}\mathbf{y}_s, \tag{5}$$

and

$$\begin{aligned}\mathbf{R}: &= \left[R\left(\mathbf{x}^{(i)},\mathbf{x}^{(j)}\right)\right]_{ij}\in\mathbb{R}^{n_s\times n_s},\\ \mathbf{r}(\mathbf{x}) &:= \left[R\left(\mathbf{x}^{(i)},\mathbf{x}\right)\right]_{i}\in\mathbb{R}^{n_s}.\end{aligned} \tag{6}$$

The mean squared error (MSE) of the kriging prediction at any untried $\mathbf{x}$ is

$$MSE\left[\hat{y}(\mathbf{x})\right] = \hat{\sigma}^2\left[1-\mathbf{r}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{r} + \left(1-\mathbf{1}\mathbf{R}^{-1}\mathbf{r}\right)^2\Big/\mathbf{1}^{\mathrm{T}}\mathbf{R}\mathbf{1}\right], \tag{7}$$

where

$$\hat{\sigma}^2 = \left(\mathbf{y}_s - \mathbf{1}\hat{\beta}_0\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{y}_s - \mathbf{1}\hat{\beta}_0\right)\Big/n_s. \tag{8}$$

2.2.2 Correlation models

The constructions of the correlation matrix $\mathbf{R}$ and the correlation vector $\boldsymbol{r}$ require the calculation of the correlation functions. The correlation function for random variables at two sites $\mathbf{x}^{(i)},\mathbf{x}^{(j)}$ is assumed to be only dependent on the spatial distance. Here we focus on a family of correlation models that is of the form

$$R\left(\mathbf{x}^{(i)},\mathbf{x}^{(j)}\right) = \prod_{k=1}^{n_y} R_k\left(\theta_k, \mathbf{x}_k^{(i)} - \mathbf{x}_k^{(j)}\right). \tag{9}$$

The correlation function used here is a cubic spline:

$$R_k = \begin{cases} 1-15\xi_k^2+30\xi_k^3 & \text{for } 0\le\xi_k\le 0.2\\ 1.25\left(1-\xi_k\right)^3 & \text{for } 0.2<\xi_k<1\\ 0 & \text{for } \xi_k\ge 1\end{cases} \tag{10}$$

where

$$\xi_k = \theta_k\left|\mathbf{x}_k^{(i)} - \mathbf{x}_k^{(j)}\right|. \tag{11}$$

2.2.3 Kriging fit

Hyper parameters of the kriging, $\theta = (\theta_1,\ldots,\theta_{N_v})$, whose dimensionality equals the dimensionality of the sampled data, can be tuned by solving a maximum likelihood estimation (MLE) problem:

$$\mathrm{MLE} = \arg\max_{\theta}\left(-\frac{1}{2}\left[n_s\ln\left(\hat{\sigma}^2\right) + \ln|\mathbf{R}|\right]\right). \tag{12}$$

In this paper, the quasi-Newton method (Byrd 1995) is used.

2.3 Multiple-point infill strategy and constraint handling

Five types of typical infill criteria are investigated in (Liu 2012b). Since it was shown that the MSE criterion performs much worse than the others, the other four infill criteria are used in this article. Accordingly, four new points can be obtained and evaluated in parallel at each updating cycle of a surrogate-based optimization. The new sampling data is augmented to the sampled database to update the surrogate models, which drives the optimization towards the global optimum.

2.3.1 Minimizing the prediction of surrogate models (MP)

This criterion directly replaces the cost function as well as the constraint functions with surrogate models and searches for the minimum objective based on the surrogate models. After building the surrogate models for the cost function and the constraint functions, the following optimization problem is solved:

$$\begin{aligned}&\textit{Minimize} \quad \hat{y}(\mathbf{x})\\ &s.t. \quad\ \hat{g}_i(\mathbf{x})\ge 0,\quad i=1,\cdots,N_G\end{aligned}, \tag{13}$$

where N_G is the number of the constraints. To solve this problem, both the genetic algorithm (GA) and gradient-based algorithms (including quasi-Newton BFGS and sequential quadratic programming (SQP)) are used. For GA, an efficient constraint handling method proposed by Deb (2000) is adopted. In this constraint handling method, the following fitness function is used:

$$F(\mathbf{x}) = \begin{cases}\hat{y}(\mathbf{x}) & \text{if } \hat{g}_i(\mathbf{x})\ge 0 \quad \forall i=1,2,\cdots,N_G\\ \hat{y}_{\max} + \sum_{i=1}^{N_G}\left\langle\hat{g}_i(\mathbf{x})\right\rangle & \text{otherwise}\end{cases} \tag{14}$$

where $\langle\rangle$ returns the absolute value of the operand if it is negative. The parameter $\hat{y}_{\max}$ is the objective function value of the worst feasible solution in the population, and $\hat{g}_i(\mathbf{x})$ is

normalized here. Note that there is no user-defined penalty parameter for this method. The detail of this method and its advantages over the traditional penalty function methods can be found in literature (Deb 2000). On the other hand, SQP can handle the constraints by itself.

The constrained form of the MP criterion is called CMP in this article. Practice suggests that it usually leads the optimizer to become trapped in a local optimum, or sometimes even before a local optimum is found. However, the optimization can be more accurately converged than that of other criteria.

2.3.2 Maximizing the expected improvement (EI)

The expected improvement is defined as the improvement we expect to achieve at an untried site $\mathbf{x}$. We assume a random variable $Y \sim N[\hat{y}(\mathbf{x}), s^2(\mathbf{x})]$, where $\hat{y}$ is the kriging predictor defined in Eq. (4), and s^2 is the MSE defined in Eq. (7). Let $y_{\min}$ be the current best real objective value, the improvement can be expressed as $I = y_{\min} - Y(\mathbf{x}) > 0$. Then the expected improvement is given by

$$E[I(\mathbf{x})] = \begin{cases} \left(y_{\min} - \hat{y}(\mathbf{x})\right)\Phi\left(\dfrac{y_{\min} - \hat{y}(\mathbf{x})}{s(\mathbf{x})}\right) \\ +s(\mathbf{x}) \times \phi\left(\dfrac{y_{\min} - \hat{y}(\mathbf{x})}{s(\mathbf{x})}\right) \quad \text{if} \quad s > 0 \\ 0 \qquad\qquad\qquad\qquad\qquad\quad \text{if} \quad s = 0 \end{cases} \tag{15}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the cumulative distribution and probability density function of a standard normal distribution, respectively.

For a constrained optimization, the EI method can be extended to the constrained form, which is depicted below. First, a kriging model for the constraint function $g(\mathbf{x})$ is constructed. Similar to the expected improvement, we can assume a random variable $G \sim N[\hat{g}(\mathbf{x}), s_g^2(\mathbf{x})]$ for the constraint function. The probability of satisfying the constraint can be written as

$$P[G > 0] = 1 - \Phi\left(\frac{-\hat{g}(\mathbf{x})}{s_g(\mathbf{x})}\right), \tag{16}$$

where s_g is the root mean squared error (RMSE) of the kriging model for the constraint function. Then the constrained expected improvement can be calculated by :

$$E_c[I(\mathbf{x})] = E[I(x) \cap G > 0] = E[I(\mathbf{x})] \times P[G > 0]. \tag{17}$$

For the problem with multiple constraints, the constrained expected improvement is obtained by multiplying each probability of satisfying the constraint, and the following problem is solved:

$$\textit{Maximize}: \quad E_c[I(\mathbf{x})] = E[I(\mathbf{x})] \times \prod_{i=1}^{N_G} P_i[G_i > 0]. \tag{18}$$

This constrained form of the EI criterion is denoted by CEI in this article. It soon got popularity for efficient global optimization, after it was proposed. Nevertheless, we observed that the EI function is highly multi-modal, and it is hard to search the maximum value in high-dimensional problems. Besides, even if the global optimum could be found, finding it will be expensive (Jones 2001).

2.3.3 Minimizing the lower confidence bounding (LCB)

The lower confidence bounding is defined as following:

$$LCB(\mathbf{x}) = \hat{y}(\mathbf{x}) - A \times s(\mathbf{x}), \tag{19}$$

where $s(\mathbf{x})$ denotes the standard deviation defined in Eq. (7), and A is an user-defined parameter, which can determine the relative weight of exploitation and exploration. A larger A makes the search tend to be more global; otherwise, a smaller A means the search is more local. In this article, A is set to 4, according to our best practice. In each iteration of updating the surrogate models, we solve the following constrained sub-optimization problem:

$$\begin{aligned} &\textit{Minimize}: \; LCB(\mathbf{x}) = \hat{y}(\mathbf{x}) - A \times s(\mathbf{x}) \\ &s.t. \quad \hat{g}_i(\mathbf{x}) \geq 0, \quad i = 1, \cdots, N_G \end{aligned}, \tag{20}$$

using the same optimization method and constraint handling method as used in CMP. Note that the constrained form of LCB criterion is denoted CLCB in this article.

2.3.4 Maximizing the probability of improvement (PI)

Similar to the EI criterion, we assume a random variable $Y \sim N[\hat{y}(\mathbf{x}), s^2(\mathbf{x})]$. Hence, the probability that the predicted objective value precedes the current best real objective value is as following:

$$P[Y < y_{\min}] = \Phi\left(\frac{y_{\min} - \hat{y}(\mathbf{x})}{s(\mathbf{x})}\right). \tag{21}$$

The constraints are handled in the same manner as the CEI criterion. The constrained form of the PI criterion is denoted by CPI in this article. For the CPI criterion, new sample point is obtained by solving the following sub-optimization problem:

$$\textit{Maxmize}: \quad \Phi\left(\frac{y_{\min} - \hat{y}(\mathbf{x})}{s(\mathbf{x})}\right) \times \prod_{i=1}^{N_G} P_i[G_i > 0]. \tag{22}$$

For this criterion, the strengths and weaknesses are similar to the CEI.

For the use of aforementioned multiple infill criteria, multiple new sample points can be obtained, with each criterion choosing a single point. Then the new sample points can be evaluated in parallel. As each infill criterion has pros and cons, employing the criteria simultaneously can complement each other. The EI, LCB, and PI can find different promising regions since they have distinct characteristics of balancing the exploitation and exploration. In addition, MP can exploit the region of the current optimum design and lead the optimizer to accurately converge to the optimum. It is noted that this method can be readily extended to add arbitrary number of sample points at each update cycle, by simply selecting multiple sample points for each of the four infill criteria using the methods from Laurenceau (2010), Schonlau (1997), Viana (2010), and Soberster (2004), which is beyond the scope of this article.

2.4 Flowchart of surrogate-based parallel optimization method

Similar to other "two-step" surrogate-based optimization methods, a certain number of initial sample points are chosen in the design space using a DoE, which is LHS in this article. Then the samples (each sample point corresponds to a wing shape) are evaluated by CFD in parallel. Kriging models are subsequently built for the cost function as well as the constraint function(s), based on the sampled data. Then 4 sub-optimization problems are executed individually in parallel: minimizing the kriging approximated cost function (CMP), maximizing the expected improvement (CEI), minimizing the lower confidence bounding (CLCB), maximizing the probability of improvement (CPI). After that, the 4 new points obtained by solving the sub-optimization problems are evaluated by CFD in parallel, and then the resulting data is augmented to the sample database to update the kriging models. The updating process is repeated until the global optimum is found or a limit to the total number of function evaluations is reached. Figure 1 shows the flowchart of this parallel optimizer.

3 Analytical test cases

To verify the effectiveness of the developed parallel infilling strategy, it is compared to three kinds of similar methods recently developed in the literature and the conventional single-point infilling EGO method as well.

3.1 Three existing parallel infilling strategies

3.1.1 Kriging Believer (KB)

For the Kriging Believer strategy (Ginsbourger 2010), the true responses at the site with maximum EI are replaced by the predicted values and the dummy sample data is augmented to the sample data set to rebuild the Kriging models. The iteration repeats until q sites have been chosen to run in parallel. In this research, the CEI criterion is used to choose the new sites.

3.1.2 Multiple point CEI (Multi-CEI)

First proposed by Sobester (2004), this strategy chooses the positions of multiple local maxima of the CEI function to update the surrogate models.

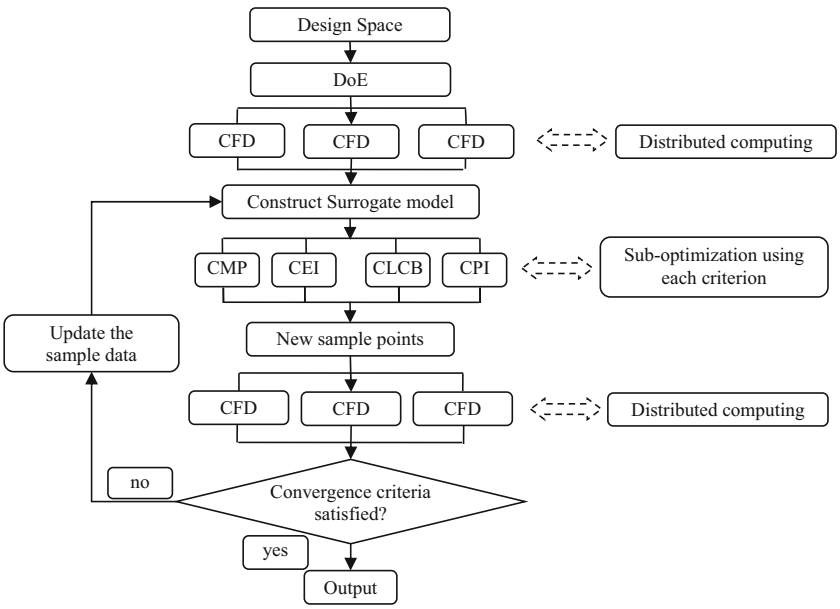
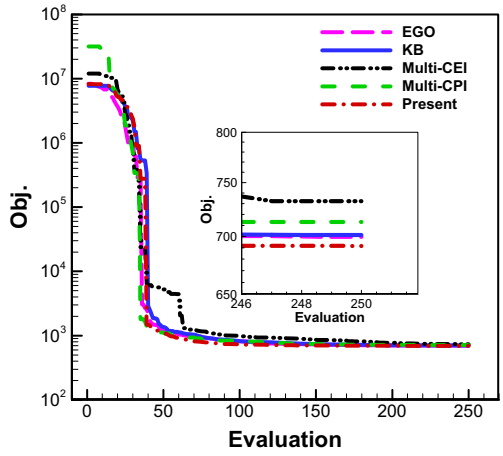
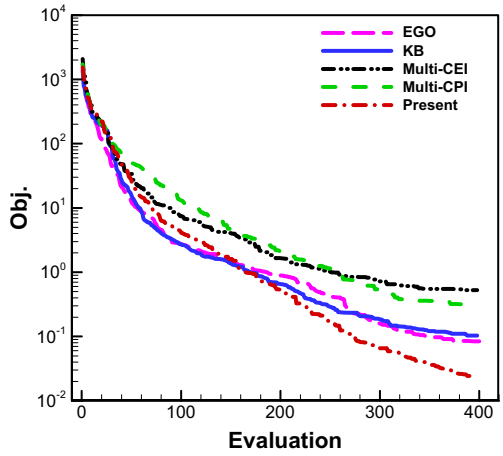


Fig. 1 Flowchart of the surrogate-based parallel optimizer, SurroOpt



(a) G9 problem



(b) Rosenbrock

Fig. 2 Averaged convergence of the five infilling sampling strategies

3.1.3 Multiple point CPI (Multi-CPI)

This strategy was first investigated by Viana and Haftka (2010), it chooses multiple points using the probability of

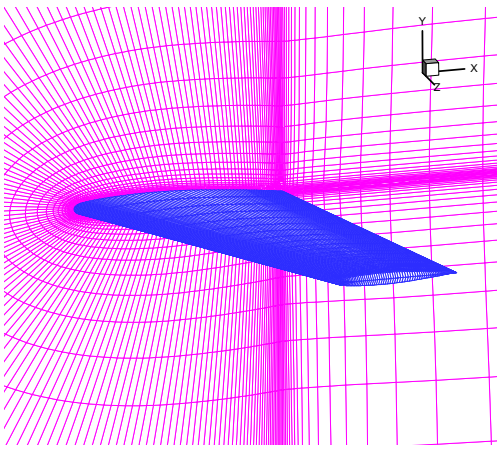
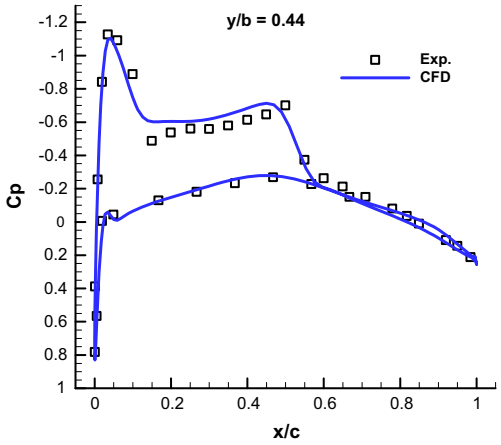
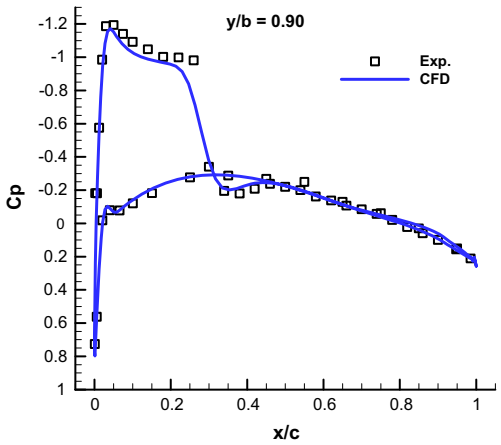


Fig. 3 Sketch of C-H grid for wing optimization



(a) 44% span



(b) 90% span

Fig. 4 Comparison of pressure distributions between CFD and experimental data for ONERA-M6 wing

the improvement criterion. Instead of picking the best set of q points from randomly generated datasets, q points are

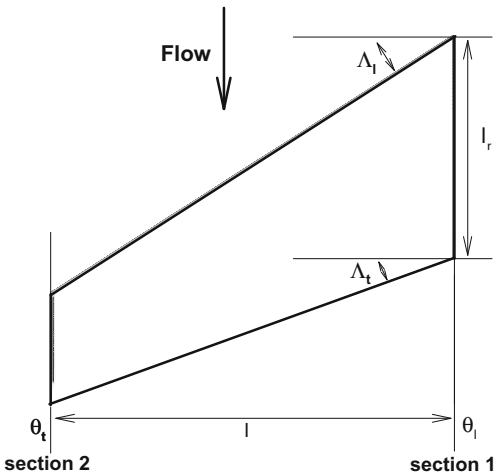


Fig. 5 Schematic diagram of wing design variables (for planform optimization)

Table 1 Lower and upper limits of the design variables for the planform shape optimization of M6 wing

Design variable	Baseline	Lower limit	Upper limit
l_r (chord length of wing root)	0.8059	0.70	0.90
ℓ (half span)	1.1963	1.10	1.35
Λ_l (sweep-back angle of leading edge)	30°	25°	35°
Λ_t (sweep-back angle of trailing edge)	15.8°	10°	20°
θ_r (twist angle of wing root)	0	0°	3°
θ_t (twist angle of wing tip)	0	−5°	0°

obtained by searching the q local maxima of the CPI function in this research.

For all of the three aforementioned strategies, q is set to four for convenience of comparison.

3.2 Problem definition

A strongly constrained global optimization problem G9 (Deb 2000) and a tough local optimization problem Rosenbrock are employed to carry out the comparison.

3.2.1 G9 problem

This problem has seven design variables and four nonlinear constraints. The mathematical model is of the form

$$\begin{aligned} Minimize: \quad & f(\boldsymbol{x}) = (x_1-10)^2 + 5(x_2-12)^2 + {x_3}^4 + 3(x_4-11)^2 + 10{x_5}^6 \\ & +7{x_6}^2 + {x_7}^4 - 4x_6x_7 - 10x_6 - 8x_7\big], \quad x_i = [-10, 10], i = 1, 7 \\ s.t.: \quad & g_1(\boldsymbol{x}) = 127 - 2{x_1}^2 - 3{x_2}^4 - x_3 - 4{x_4}^2 - 5x_5 \geq 0 \\ & g_2(\boldsymbol{x}) = 282 - 7x_1 - 3x_2 - 10{x_3}^2 - x_4 + x_5 \geq 0 \\ & g_3(\boldsymbol{x}) = 196 - 23x_1 - {x_2}^2 - 6{x_6}^2 + 8x_7 \geq 0 \\ & g_4(\boldsymbol{x}) = -4{x_1}^2 - {x_2}^2 + 3x_1x_2 - 2{x_3}^2 - 5x_6 + 11x_7 \geq 0 \end{aligned} \tag{23}$$

The optimal solution observed so far is $x^* = (2.3305, 1.9514, -0.4775, 4.3657, -0.6245, 1.0382, 1.5942)$, $f(x^*) = 680.6300573$.

Table 2 Comparison of the planform parameters between baseline and optimized wing (Case 1: planform shape optimization of M6)

Design variable	Baseline	Optimum
l_r	0.8059	0.7673
ℓ	1.1963	1.3500
Λ_l	30°	34.89°
Λ_t	15.8°	20.00°
θ_r	0	0.5556°
θ_t	0	−2.22°

3.2.2 Rosenbrock

The mathematical model of this problem is as following:

$$Minimize: \ f(\boldsymbol{x}) = \left(\sum_{i=1}^{N_v-1} \left[100 \times \left(x_{i+1} - {x_i}^2\right)^2 + (1-x_i)^2 \right] \right). \quad x_i \in [-2, 2], \quad i = 1, \cdots, N_v \tag{24}$$

The true optimal solution is $x^* = (1, \cdots, 1), \quad f(x^*) = 0$. Here we are concerned with the problem with 5 design variables ($N_v = 5$).

3.3 Results

For both problems, ten initial sample points are used and the limit of the total number of sample points is prescribed to 250 and 400 respectively. For each infilling strategy, the optimizations are repeated for 30 times and the averaged convergence histories are shown in Fig. 2. It is shown that for both problems the proposed parallel infilling strategy performs the best, followed by the EGO, Kriging believer, multi-CPI, and multi-CEI.

4 Wing design optimization

4.1 Flow analysis

The flow analyses are performed with an in-house code called PMNS3D. It solves the Reynolds-averaged Navier–Stokes (RANS) equations to simulate the flow around a 3D

Table 3 Comparison of the aerodynamic force coefficients and wing area for baseline and optimized wing (Case 1: planform optimization of M6)

	Baseline	Optimum
C_d	0.01751	0.01491 (−14.85 %)
C_l	0.23803	0.23805 (+0.01 %)
Area	0.75318	0.75325 (+0.01 %)

Table 4 Comparison of the predicted and the validated C_d and C_l at the obtained optimal point (Case 1)

	by CFD	by surrogate	error
C_d	0.01491	0.014907	0.03 %
C_l	0.23805	0.238033	0.006 %

configuration. The three-dimensional RANS equations used in PMNS3D are as follows:

$$\frac{\partial \mathbf{W}}{\partial t}+\frac{\partial \mathbf{E}}{\partial x}+\frac{\partial \mathbf{F}}{\partial y}+\frac{\partial \mathbf{G}}{\partial z}=\frac{\partial \mathbf{E}_v}{\partial x}+\frac{\partial \mathbf{F}_v}{\partial y}+\frac{\partial \mathbf{G}_v}{\partial z}, \tag{25}$$

where **W** are conservative variables, **E**, **F** and **G** are the inviscid flux terms, $\mathbf{E}_v$, $\mathbf{F}_v$ and $\mathbf{G}_v$ are the viscous flux terms respectively. With the structured grids of C-H topology, the equations are solved by using the cell-centered finite volume method. The second-order Jameson central scheme is used as the spatial discretization scheme and the Spalart-Allmaras one-equation turbulence model is used for turbulence closure. Implicit residual smoothing, local-time stepping and multigrid techniques are used to accelerate the solution to converge to the steady state. Figure 3 shows the C-H type grids used in this paper, which are generated automatically by our in-house code, with the grid distribution of 208 (chord-wise direction) × 48 (normal direction) × 48 (span direction). Figure 4 shows the comparison of pressure distributions of the ONERA-M6 wing between the CFD results obtained by PMNS3D and the wind-tunnel experimental results at the free-stream condition of $Ma = 0.8395$, $\mathrm{Re} = 11.72 \times 10^6$, $\alpha = 3.06°$. As one can see, the CFD results are in reasonably good agreement with the experimental data.

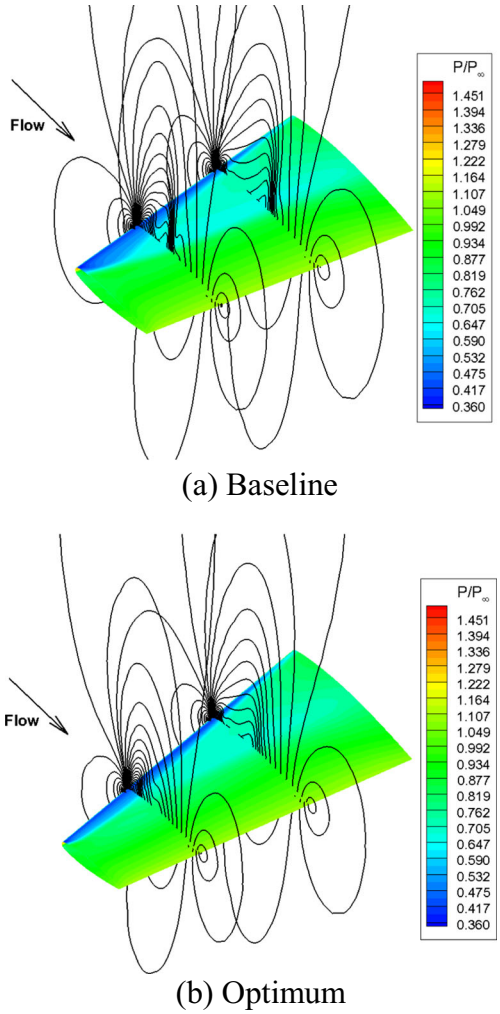


Fig. 6 Comparison of the pressure contours for baseline and the optimized shape (surface and 37.6, 72.7 % semi-span) (Case 1: planform shape optimization of M6 wing)

4.2 Geometric parameterization of wing

Generally, the wing shape contains two parts: planform shape and sectional shape. Once the planform parameters and the shapes of the specified control sections are determined, the shape at any span-wise location can be obtained by interpolation.

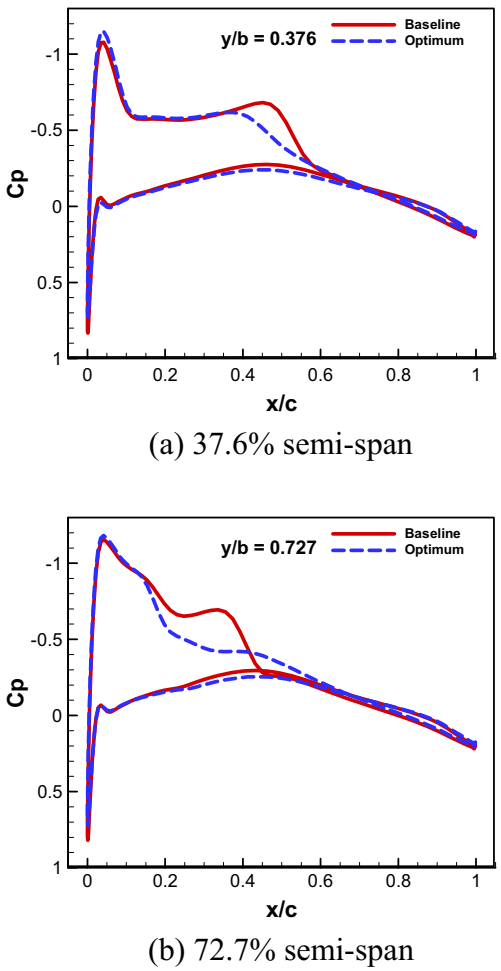


Fig. 7 Comparison of the pressure distributions for baseline and optimized shapes (Case 1: planform shape optimization of M6 wing)

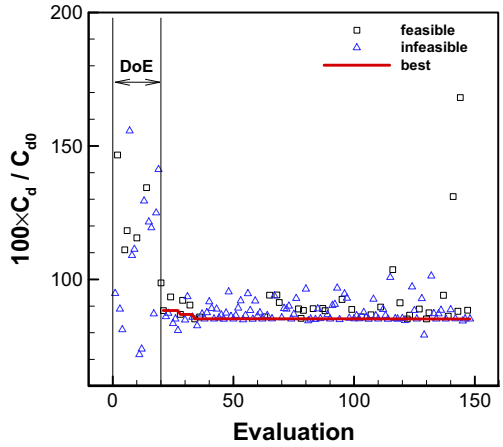


Fig. 8 Convergence history of objective function of the optimization process (Case 1: planform shape optimization of M6 wing)

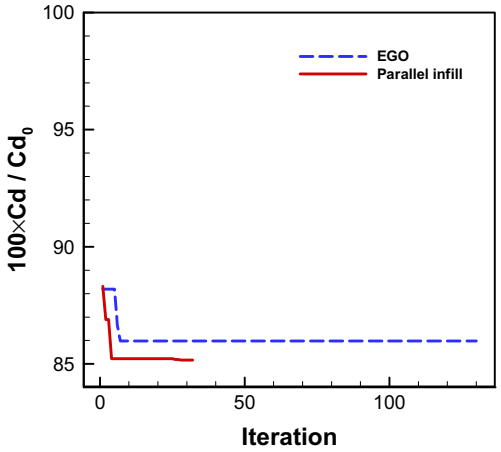


Fig. 9 Comparison of the convergence histories for EGO and parallel method (Case 1: planform shape optimization of M6 wing)

4.2.1 Planform parameterization

In general, there are 7 independent parameters to describe the shape of a trapezoidal wing: area of the wing *Area*, aspect ratio *AR*, root-to-tip ratio λ, leading-edge sweep angle Λ_l, dihedral angle θ_d, twist angle of the wing tip θ_t and the incident angle θ_m. For convenience, the first 4 parameters are replaced with the following equivalent 4 parameters: chord length of the root section l_r, half span ℓ, leading-edge sweep angle of Λ_l, and trailing edge sweep angle of Λ_t. Besides, the incident angle is replaced with the twist angle of wing root θ_r, since a single wing (without fuselage) is used in this article; and the dihedral angle is not considered. As a result, 6 independent parameters are considered and taken as the design variables for the planform shape optimization (see Fig. 5). The design variables and their ranges are shown in Table 1. The twist angle is prescribed to vary linearly from the root to tip for the planform shape optimization.

4.2.2 Wing section parameterization

Since the sectional shape of the wing can be considered as an airfoil, in this article, the CST (class function/shape function transformation) method proposed by Kulfun (2008) is used. The CST method can describe an airfoil with a small number of variables.

Table 5 Comparison of the drag reduction and CPU time for the EGO and parallel method (Case 1: planform shape optimization of M6 wing)

	Drag reduction	CPU time
EGO	14.02 %	197 h,17 min
Parallel infill	14.85 %	48 h,28 min

The CST describes an airfoil shape (y coordinates of upper or lower surfaces) with the product of a class function $C(x/c)$ and a shape function $S(x/c)$:

$$\frac{y}{c} = C\left(\frac{x}{c}\right)\cdot S\left(\frac{x}{c}\right) + \frac{x}{c}\cdot\frac{\Delta z_{te}}{c}, \tag{26}$$

where c is the chord length, and $C(x/c)$ has the form of

$$C\left(\frac{x}{c}\right) = \left(\frac{x}{c}\right)^{N_1}\cdot\left[1-\frac{x}{c}\right]^{N_2}, \quad 0\le\frac{x}{c}\le 1\,. \tag{27}$$

Usually, $N_1 = 0.5$ and $N_2 = 1.0$ are used for a round nose and sharp trailing edge airfoil.

The shape function is the weighted sum of *nth*-order Bernstein polynomial

$$S\left(\frac{x}{c}\right) = \sum_{i=0}^{n}\left[A_i\cdot K_{i,n}\cdot\left(\frac{x}{c}\right)^{i}\cdot\left(1-\frac{x}{c}\right)^{n-i}\right], \tag{28}$$

where $K_{i,n}$ are the coefficients of the Bernstein polynomial

$$K_{i,n}\equiv\binom{n}{i}\equiv\frac{n!}{i!(n-i)!}, \tag{29}$$

and A_i are the weights, which are the design variables for the sectional shape optimization.

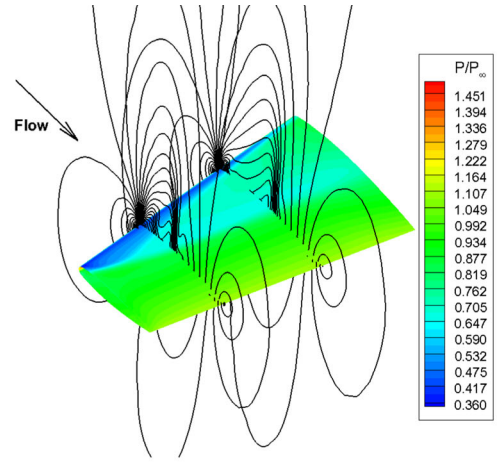
Table 6 Comparison of the aerodynamic force coefficients and geometric parameters for baseline and optimized shapes (Case 2: section shape optimization of M6 wing)

	Baseline	Optimum
C_d	0.01751	0.01562(−10.79 %)
C_l	0.23803	0.23804 (+0.01 %)
$Thick_r$	0.09786	0.09793 (+0.07 %)
$Thick_t$	0.09786	0.09799 (+0.13 %)

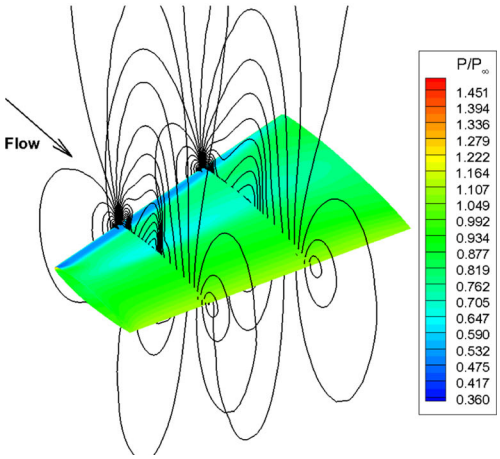
Table 7 Comparison of the predicted and the validated C_d and C_l at the obtained optimal point (Case 2)

	by CFD	by surrogate	error
C_d	0.01562	0.01567	0.28 %
C_l	0.23804	0.238033	0.003 %

It is evident from Eq. (28) that, for the CST with nth-order Bernstein polynomial, there are $n+1$ design variables for upper and lower surface respectively, so there are $2n+2$ design variables in total for an airfoil or a wing section. The accuracy of fitting an airfoil or wing section is higher if larger n is used. Kulfan (2008) suggested that 4 ~ 6 is enough to fit a conventional airfoil with sufficient accuracy. Hence, 5th-order Bernstein polynomials are used in this article.



(a) Baseline



(b) Optimum

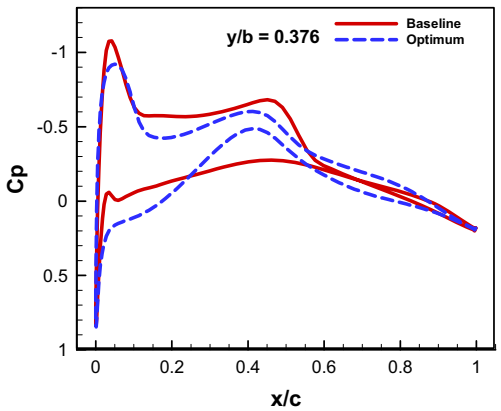
Fig. 10 Comparison of the pressure contours for baseline and the optimized shapes (surface and 37.6, 72.7 % semi-span) (Case 2: sectional shape optimization of M6 wing)

4.3 Design cases

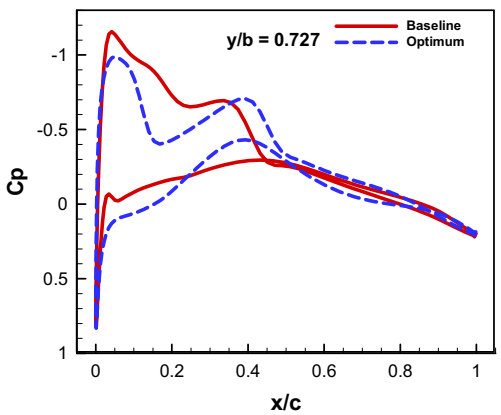
Four cases are exercised in this research, three of which are drag minimizations of an ONERA-M6 wing. First, the planform shape is optimized with fixed section shapes (case 1, 6 design variables); then, the section shapes are optimized with fixed planform shape (case 2, 24 design variables); third, the planform and section shapes are optimized simultaneously (case 3, 30 design variables). The fourth case is the sectional shape optimization of a DLR-F4 wing (48 design variables).

4.3.1 Planform shape optimization of M6 wing

In this case, we take the six parameters given in sub-section 4.2.1) as design variables. The objective is to minimize the drag at the free-stream condition of $Ma = 0.8395$, $\mathrm{Re} = 11.72 \times 10^6$, $\alpha = 3.06°$, with two constraints about the lift



(a) 37.6% semi-span



(b) 72.7% semi-span

Fig. 11 Comparison of the pressure distributions for baseline and optimized shapes (Case 2: sectional shape optimization of M6 wing)

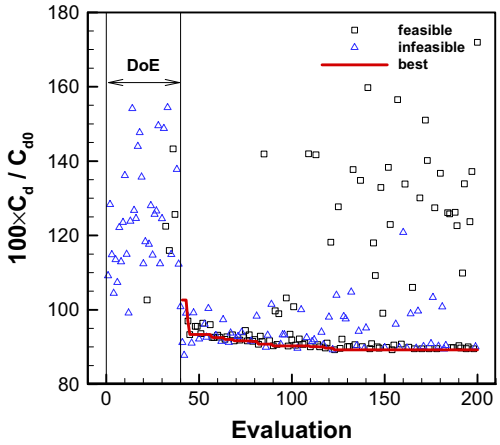


Fig. 12 Convergence history of objective function of the optimization process (Case 2: sectional shape optimization of M6 wing)

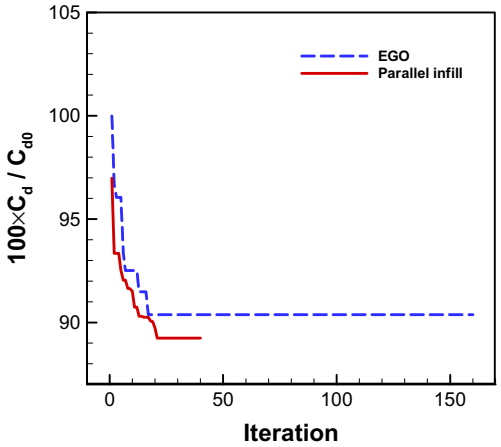


Fig. 13 Comparison of convergence histories of EGO and Parallel method (Case 2: sectional shape optimization of M6 wing)

coefficient and the area of wing. The mathematical model is as following:

$$\begin{aligned} &Minimize: \quad C_d \\ &s.t. \quad \begin{cases} (1) \quad C_l \geq C_{l0} \\ (2) \quad Area \geq Area_0 \end{cases} . \end{aligned} \tag{30}$$

20 initial sample points are selected using the LHS, and the maximum allowable number of CFD calls is set to 150. The optimization experienced 48 h 28 min using a PC with a 4-cores, Intel i7 processor (3.4 GHz). The comparison of the baseline and optimized planform parameters is shown in Table 2.

Table 3 gives the comparison of aerodynamic performance and wing area between the baseline and optimized wing. From the two tables, one can see that the root chord length is decreased while the half-span length is increased; and the sweep angles for leading edge and trailing edge are both increased. As a result, the induced drag and wave drag are reduced due to the increased aspect ratio and sweep angle are increased. Table 3 also shows that the total drag coefficient is reduced by 14.85 %, while both two constraints are strictly satisfied. Table 4 gives the predicted aerodynamic coefficients by kriging at the best site, which shows that the surrogate have satisfactory accuracy of approximation at this site.

Figure 6 compares the pressure contours between the baseline and optimized wing. It is shown that the aspect ratio and sweep-back angle are both increased, resulting in smaller induced drag and wave drag. Figure 7 compares the pressure distributions of two spanwise sections. It can also be seen that the shock at each section is weakened for the optimized wing. Figure 8 shows the convergence history of the objective function during the optimization process. It should be noted that, the black square and blue delta symbols denote the feasible and infeasible solutions, respectively, while the solid line denotes the best feasible solution obtained so far.

This problem is also solved by single-point infilling EGO method with the same number of initial samples and limit number of total CFD evaluations. Only one CPU core is employed since one point is infilled per updating cycle. As a result, the drag is reduced by 14.02 %, which is slightly less than that of the parallel infilling method proposed in this article. Besides, it costs 197 h 17 min, almost 4 times as long as that of the parallel infilling method (also shown in Table 5). The convergence of the objective function value with respect to the iteration number is shown in Fig. 9. It is observed that less iterations are needed for the parallel infilling method.

Table 8 Comparison of the drag reduction and CPU time for EGO and parallel method (Case 2: sectional shape optimization of M6 wing)

	Drag reduction	CPU time
EGO	9.62 %	190 h,7 min
Parallel infill	10.79 %	60 h,14 min

4.3.2 Sectional shape optimization of M6 wing

For this case, the planform parameters are taken from the baseline and only the root and tip sections are parameterized using the CST of 5th order, resulting in 12 variables for each section and 24 in total. The objective here is also to minimize the drag at the same free-stream condition as in case 1 and

Table 9 Comparison of the aerodynamic force coefficients and geometry parameters of the baseline and optimized shapes (Case 3: planform and sectional shape optimization of M6 wing)

	Baseline	Optimum
C_d	0.01751	0.01277 (−27.07 %)
C_l	0.23803	0.23814 (+0.05 %)
$Thick_r$	0.09786	0.09786 (+0.00 %)
$Thick_t$	0.09786	0.10021 (+2.40 %)
$Area$	0.75318	0.75529 (+0.28 %)

Table 10 Comparison of the predicted and the validated C_d and C_l at the obtained optimal point (Case 3)

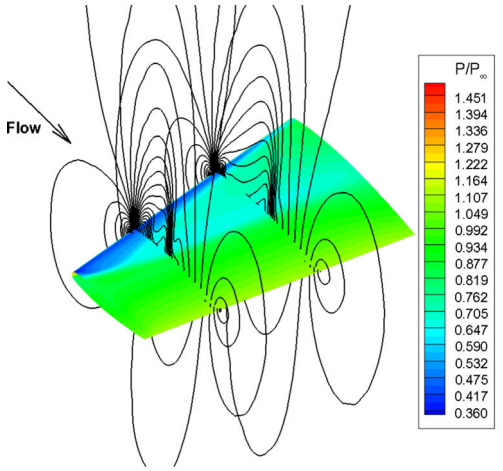
	by CFD	by surrogate	error
C_d	0.01 277	0.01273	0.32 %
C_l	0.23814	0.238302	0.07 %

with 3 constraints on the lift coefficient, the maximum thickness-to-chord ratio of the root and tip sections. The mathematical model is

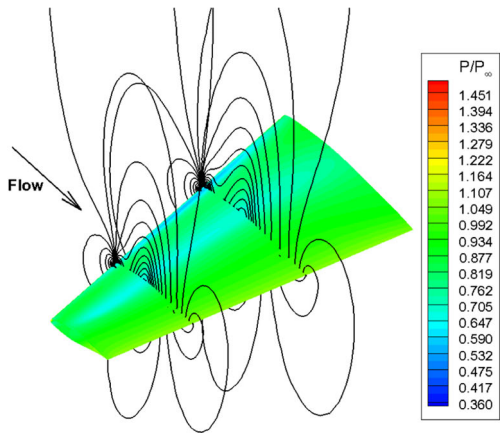
$$\begin{aligned} &\textit{Minimize} \quad C_d \\ &s.t. \quad \begin{aligned} &(1) \quad C_l \geq C_{l0} \\ &(2) \quad Thick_r \geq Thick_{r0} \\ &(3) \quad Thick_t \geq Thick_{t0} \end{aligned} \end{aligned} \tag{31}$$

where $Thick_r$ denotes the maximum thickness-to-chord ratio of the root section, while $Thick_t$ denotes the tip thickness.

40 initial points are selected by the LHS, and totally 200 CFD simulations have been run during the entire optimization process. This costs 60 h 14 min in total using 4 CPU

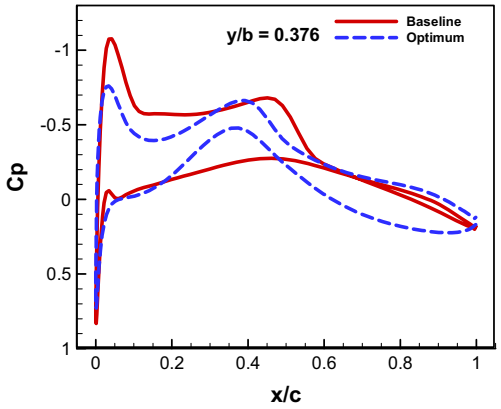


(a) Baseline

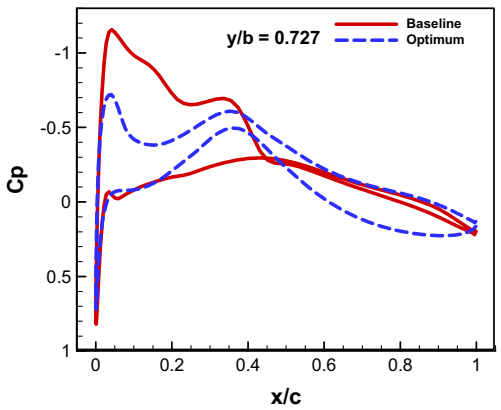


(b) Optimum

Fig. 14 Comparison of the pressure contours for baseline and the optimized shapes (surface and 37.6, 72.7 % semi-span) (Case 3: planform and sectional shape optimization of M6 wing)



(a) 37.6% semi-span



(b) 72.7% semi-span

Fig. 15 Comparison of the pressure distributions for baseline and optimized shapes (Case 3: planform and sectional shape optimization of M6)

cores. Table 6 compares the aerodynamic performance and thickness of control sections between the baseline and optimized wings. We see that the drag is reduced by 10.79 % with all the constraints strictly satisfied. Table 7 also gives the predicted aerodynamic coefficients by kriging at the

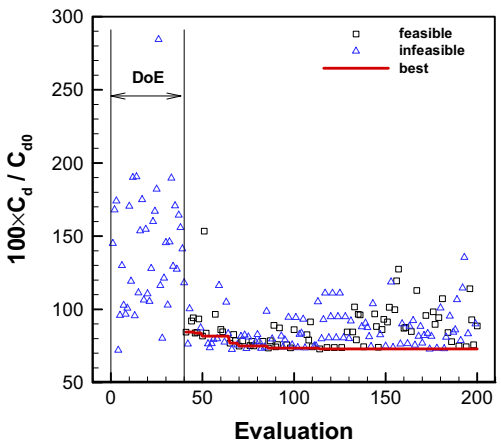


Fig. 16 Convergence history of objective function of the optimization process (Case 3: planform and sectional shape optimization of M6 wing)

Table 11 Comparison of the drag reduction and CPU time for EGO and parallel method (Case 3: planform and sectional shape optimization of M6 wing)

	Drag reduction	CPU time
EGO	24.44 %	272 h,20 min
Parallel infill	27.07 %	68 h,50 min

optimum site, which shows that the surrogate models are accurate enough at this site.

Figure 10 shows the comparison of baseline and optimized pressure contours. It is evident that the shock at each section is weaker for the optimized shape, which results in notable drag reduction. The weakened shock is further checked by comparison of the pressure distributions at two spanwise sections of the wing, and the results are shown in Fig. 11. The convergence history of the objective function is shown in Fig. 12. In this figure, the blue triangles and black squares represent the infeasible and feasible points, respectively, while the red line stands for history of observed best solution so far.

Similar to the previous case, the conventional single-point infilling EGO is also performed as the reference. The result shows that the drag is reduced by 9.62 %, which is less than that of the proposed parallel infill method in this article. Besides, it costs 190 h 7 min, about 3 times longer than that of the parallel infilling method. The comparison of drag reduction as well as the CPU time is shown in Table 8, and the comparison of convergence histories is shown in Fig. 13. It is shown that the proposed method gains more drag reduction and requires much less iterations, when compared with the single-point infilling EGO, which indicates the benefits of using the parallel infill strategy.

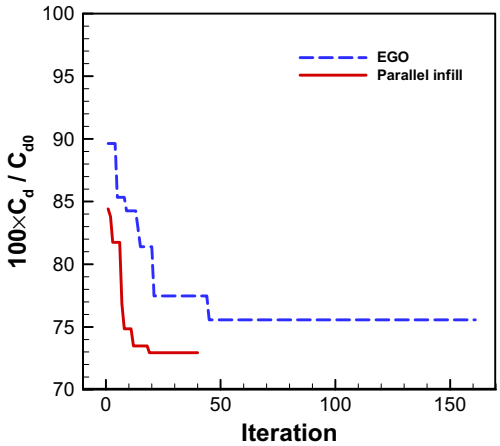


Fig. 17 Comparison of the convergence histories for EGO and Parallel infill method (Case 3: planform and sectional shape optimization of M6 wing)

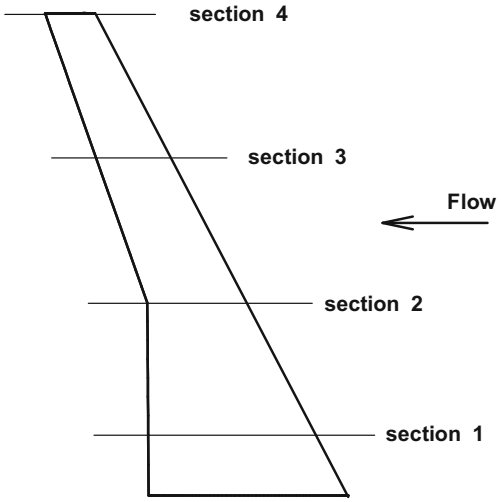


Fig. 18 Design sections of DLR-F4 wing

4.3.3 Planform and sectional shape optimization of M6 wing

This case combines the previous two, with the same objective and at the same free-stream condition. There are 30 design variables and 4 constraints in total. The constraints are associated with the lift coefficient, maximum thickness-to-chord ratio of the root and tip sections and area of the wing. The following formula gives the mathematical model of the optimization problem:

$$\begin{aligned} &Minimize: \quad C_d \\ &s.t. \quad : \begin{cases} (1) \quad C_l \geq C_{l0} \\ (2) \quad Thick_r \geq Thick_{r0} \\ (3) \quad Thick_t \geq Thick_{t0} \\ (4) \quad Area \geq Are\, a_0 \end{cases} . \end{aligned} \tag{32}$$

Initially, 40 sample points are selected by LHS and the maximum allowable number of CFD simulations is set to 200. The optimization process takes 68 h 50 min on a PC with a 4-cores Intel i7 (3.4GHz) processor. Table 9 shows the comparison of the aerodynamic coefficients and geometric parameters between the baseline and optimized wings. The drag is reduced by 27.07 %. The surrogate-model predicted and CFD-validated

Table 12 Comparison of the aerodynamic force coefficients and geometry parameters for baseline and optimized shapes (Case 4: sectional shape optimization of DLR-F4 wing)

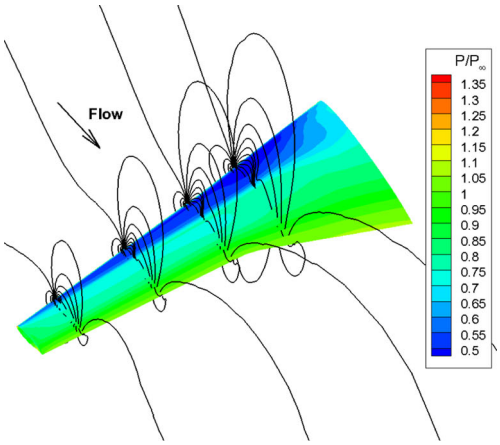
	Baseline	Optimum
C_d	0.02798	0.02636 (−5.79 %)
C_l	0.54626	0.54718 (+0.17 %)
C_m	−0.08462	−0.10870
$Thick_1$	0.15036	0.15124 (+0.59 %)
$Thick_2$	0.12200	0.12830 (+5.16 %)
$Thick_3$	0.12189	0.12412 (+1.83 %)
$Thick_4$	0.12182	0.12271 (+3.56 %)

Table 13 Comparison of the predicted and the validated C_d and C_l at the obtained optimal point (Case 4)

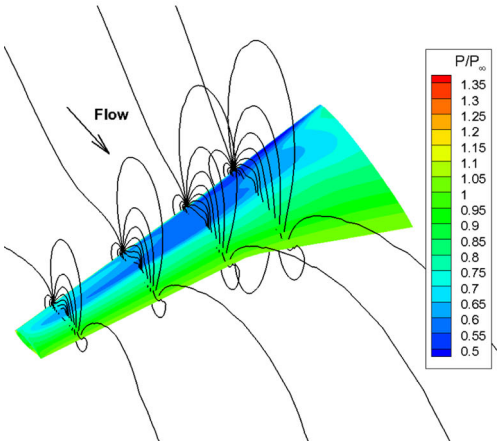
	by CFD	by surrogate	error
C_d	0.02636	0.02628	0.32 %
C_l	0.54718	0.54641	0.14 %

aerodynamic coefficients are compared in Table 10. It is shown that the surrogate models are sufficiently accurate at the site of the optimum solution.

The comparison of pressure contours is given in Fig. 14. One can see that both of the two shocks are almost removed, due to the simultaneous modification of the planform shape and sectional shape. Figure 15 compares the pressure distributions of two spanwise sections, it is evident that, there exists aft-loading for the optimized wing, which is similar to supercritical wings. Figure 16 shows the convergence history of the objective function. This figure shows that, although none of the initial sample points are feasible, the optimizer can find good results that strictly satisfy all the constraints by the use of constrained infill criteria.

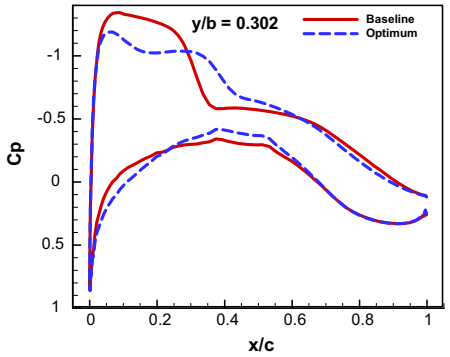


(a) Baseline

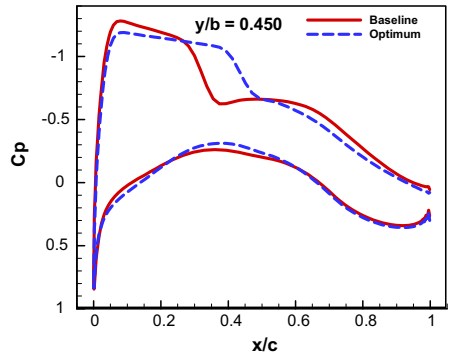


(b) Optimum

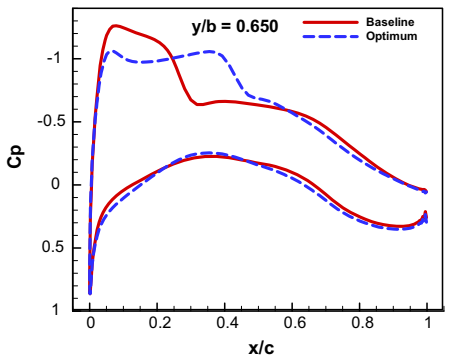
Fig. 19 Comparison of the pressure contours for baseline and the optimized wing shapes (surface and 30.2, 45.0, 65.0, 86.9 % semi-span) (Case 4: sectional shape optimization of DLR-F4 wing)



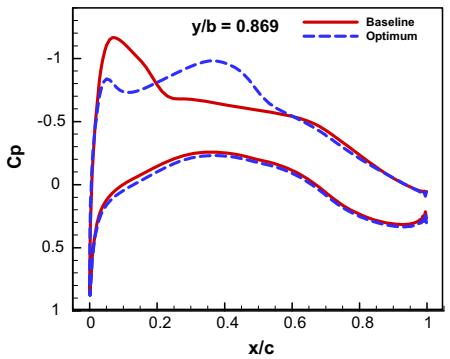
(a) 30.2% semi-span



(b) 45.0% semi-span



(c) 65.0% semi-span



(d) 86.9% semi-span

Fig. 20 Comparison of pressure distributions of the baseline and optimized wing shapes (Case 4: sectional shape optimization of DLR-F4)

In the case of conventional single-point infilling EGO, the drag reduction is 24.44 %, at a cost of 272 h 20 min using a single CPU core. The drag reduction is less than that of the parallel infilling method, and the CPU time is almost 3 times longer (also shown in Table 11). Comparison of the convergences of the objective value with respect to the iteration number is shown in Fig. 17. Similar to the previous cases, much less iterations are needed for the proposed method by the use of the parallel infilling strategy.

4.3.4 Sectional shape optimization of DLR-F4 wing

The DLR-F4 is a wing-body configuration from the 1st Drag Prediction Workshop (DPW-1). In this test case, only the wing is considered, without including the fuselage. Four sections at different spanwise locations are taken as the control sections for the shape optimization (Redeker 1994), which is shown in Fig. 18. Note that section 1 is taken as the wing root.

The 5th-order CST method is used to parameterize each of the sections, resulting in 48 design variables in total. The objective is to minimize the drag at the free-stream condition of $Ma = 0.75$, $\mathrm{Re} = 3.0 \times 10^6$, $\alpha = 0°$, with 6 constraints on the lift coefficient, the maximum thickness-to-chord ratio of the 4 sections, and the pitching moment coefficient. For the calculation of pitching moment coefficient, the mean aerodynamic chord (m.a.c.) is taken as the reference length and the quarter of the m.a.c. is taken as the reference point. The following formulation shows the mathematical model of the optimization:

$$\begin{aligned} &\textit{Minimize} : \quad C_d \\ &\textit{s.t.} \quad : \quad \begin{array}{ll} (1)\ \ C_l \geq C_{l0}, & (2)\ |C_m| \leq 0.12 \\ (3)\ \ Thick_1 \geq Thick_{1,0}, & (4)\ \ Thick_2 \geq Thick_{2,0} \\ (5)\ \ Thick_3 \geq Thick_{3,0}, & (6)\ \ Thick_4 \geq Thick_{4,0} \end{array} \end{aligned} \tag{33}$$

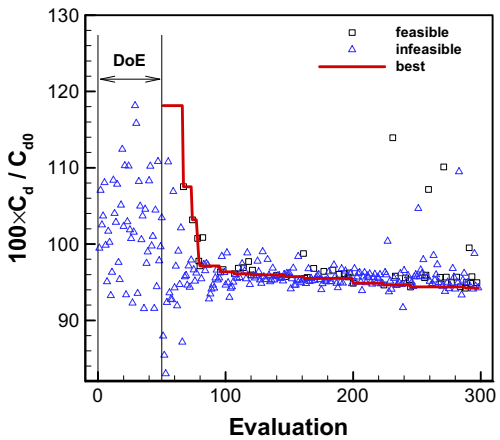
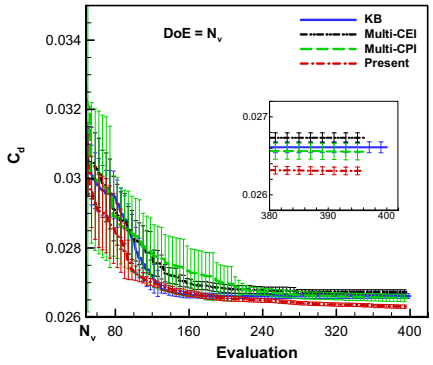
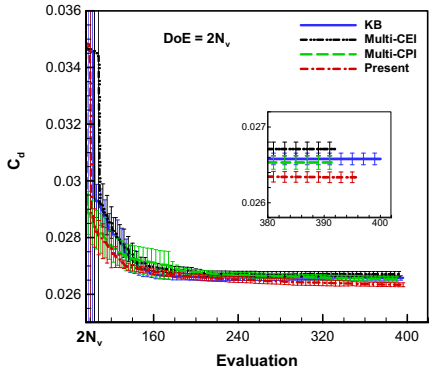


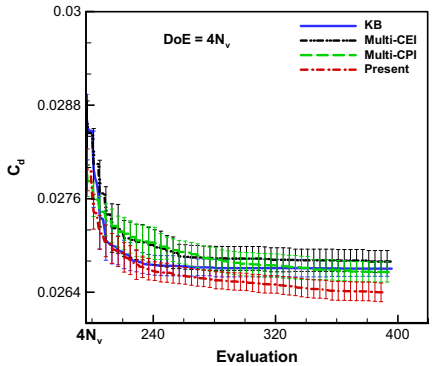
Fig. 21 Convergence history of objective function of the optimization process (Case 4: sectional shape optimization of DLR-F4 wing)



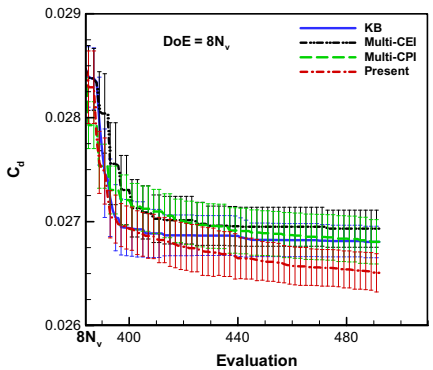
(a) Initial sample = 48



(b) Initial sample = 96



(c) Initial sample = 192



(d) Initial sample = 384

Fig. 22 Mean and standard deviation of the drag of the repeated 50 runs during the optimizations with different number of initial samples points (Case 4)

In the optimization process, 50 initial sample points are selected by the LHS, and the maximum number of CFD simulations is set to 300. The whole optimization process costs 125 h 36 min in total, when 4 CPU cores are employed. Table 12 compares the objective and constraint function values between the baseline and the optimized wings. One can see that the drag is reduced by 5.79 % and all the constraints are strictly satisfied. The approximation accuracy of the surrogate models is checked at the site of the optimum, as shown in Table 13. The absolute error for drag coefficient is as small as 1.2 counts and relative error is about 0.3 %, which confirms that through infilling sampling the accuracy of the surrogate models are dramatically improved in the vicinity of the optimum, even for the optimization problem with number of design variables as large as around 50 .

Figures 19 and 20 respectively depict the pressure contours and surface pressure distributions, for the baseline and optimized wings. It can be observed that for the optimized wing, the shock is weaker, thus the wave drag is reduced. Figure 21 shows the convergence history of the drag coefficient during the optimization process. Although none of the initial sample points is feasible, because of the strict constraints, the optimizer is still able to find good feasible solutions by repetitively infilling new sample points, which demonstrates the capacity of the proposed method for constraint handling.

This optimization problem was also carried out using the aforementioned three reference parallel infilling strategies, and the results are compared with that of using the method proposed in this article. For each parallel infilling strategy, the optimization is performed using different number of initial sample points, in the range from N_v to $8N_v$, to clarify the effect of the initial sampling. Furthermore, in order to eliminate the effect of randomness, each optimization case is repeated by 50 times and the statistical results, i.e. mean and standard deviations, are obtained. Note that since this comparison is very costly, we moved all the optimization stuff to the TianHe-1A supercomputer of the National Supercomputer Center in Tianjin, China.

Figure 22 depicts the convergence histories for each of the infilling strategies. The lines represent the averaged drag from 50 runs ($\overline{C}_d$), and the vertical bars are corresponding to the standard deviation of drag (σ_{Cd}). The length of a vertical bar equals σ_{Cd} with its center locates at $\overline{C}_d$. It is shown that the presented parallel infilling strategy outperforms the others, for

Table 14 Mean of the optimized drag of the repeated 50 runs, initial sample = N_v (Case 4: DLR-F4 wing)

method	Baseline (CFD)	Optimum (CFD)
KB	0.02798	0.02661
Multi-CEI	0.02798	0.02673
Multi-CPI	0.02798	0.02655
Present method	0.02798	0.02631

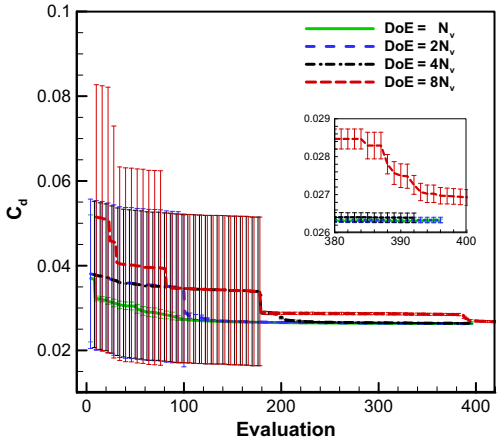


Fig. 23 Comparison of the convergence histories with different initial sample points using the presented parallel infilling strategy

the cases with different numbers of the initial sample points. Table 14 gives the mean drag coefficients of optimized wings from the repeated 50 runs with the number of initial sample points set as one time of the number of design variable. From this table, it can also be seen that, the present method proposed in this article achieves a lowest drag.

The effect of number of initial sample points is further studied in Fig. 23, which shows the convergence histories with different number of initial sample points using the presented parallel infilling strategy. As one can see, the convergence rate is quite different, but the final optima seem to be insensitive to the number of the initial sample points. It suggests that in the case that the computational budget is limited, setting number of initial sample points to N_v or $2N_v$ is likely to be a favorable choice.

5 Conclusions

In this article, an alternative parallel infilling strategy for a surrogate-based optimization is developed. The key feature of this strategy is to use multiple infill criteria simultaneously to choose multiple new sample points, which are to be evaluated in parallel at each updating cycle. In this infilling strategy, each criterion is extended to consider the constraints, which makes it well suited to problems with multiple constraints.

The proposed method is demonstrated by four cases of strongly-constrained aerodynamic shape optimization of transonic wings, including both planform and section shape design problems and with the number of design variables in the range from 6 to 48. The results verify that, the proposed parallel infilling strategy is more effective than three existing reference parallel infilling strategies, when the number of initial sample points are in the range from N_v to $8N_v$ (where N_v here denotes the number of design variables). Each case is repeated 50 times to eliminate the effect of randomness in our results. This proposed method can be readily extended to adding arbitrary number of new sample points at each

updating cycle, by simply selecting multiple sample points from each of the four infill criteria.

Acknowledgments This work was supported by the National Natural Science Foundation of China (Grant No. 11272265) and the Fundamental Research Funds for the Central Universities (Grant No. 3102015BJ(II) MYZ18). The Authors would like to thank the National Supercomputer Center in Tianjin for the use of computation resource of TianHe-1A, and Mr. Chen Fu for his effort helping us to adapt the optimization framework to TianHe-1A. The Authors also would like to thank the anonymous reviewers and editor for their thoughtful comments and suggestions for improving this article. The authors are grateful for Dr. Richard Dwight from Delft University of Technology who helped us to improve the language from the perspective of an English native speaker.

References

Byrd RH, Lu P, Nocedal J, Zhu C (1995) A limited memory algorithm for bound constrained optimization. SIAM J Sci Comput 16:1190–1208

Chaudhuri A, Haftka RT, Ifju P et al (2015) Experimental flapping wing optimization and uncertainty quantification using limited samples. Struct Multidiscip Optim 51(4):597–970

Chernukhin O, Zingg DW (2013) Multimodality and global optimization in aerodynamic design. AIAA J 51(6):1342–1354

Deb K (2000) An efficient constraint handling method for genetic algorithms. Comput Methods Appl Mech Eng 186:311–338

Forrester AIJ, Keane AJ (2009) Recent advances in surrogate-based optimization. Prog Aerosp Sci 45(1–3):50–79. doi:10.1016/j.paerosci.2008.11.001

Ginsbourger D, Le Riche R, and Carraro L (2007) A multi-points criterion for deterministic parallel global optimization based on kriging, The International Conference on Nonconvex Programming: Local and Global Approaches, Rouen, France

Ginsbourger D, Le Riche R, Carraro L (2010) Kriging is well-suited to parallelize optimization. In: Computational intelligence in expensive optimization problems. Springer, Berlin, pp 131–162

Giunta AA, Wojtkiewicz SFJ and Eldred MS (2003) Overview of modern design of experiments methods for computational simulations, AIAA 2003–649

Goto Y, Jeong S, Obayashi S, and Kohama Y (2008) Design space exploration of supersonic formation flying focusing on drag minimization. J Aircr 45(2): doi: 10.2514/1.28766

Han Z-H, Goertz S (2012) A hierarchical kriging model for variable-fidelity surrogate modeling. AIAA J 50(9):1885–1896. doi:10.2514/1.J051354

Han Z-H, Zhang K-S, Song W-P, Qiao Z-D (2010a) Optimization of active flow control over an airfoil using a surrogate-management framework. J Aircr 47(2):603–612. doi:10.2514/1.45899

Han Z-H, Zimmermann R, and Goertz S (2010b) A new cokriging method for variable-fidelity surrogate modeling of aerodynamic data, 48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, Orlando, FL, AIAA Paper 2010–1225

Han Z-H, Zimmermann R, Goertz S (2012a) An alternative cokriging model for variable-fidelity surrogate modeling. AIAA J 50(5): 1205–1210. doi:10.2514/1.J051243

Han Z-H, Goertz S, Zimmermann R (2012b) Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function. Aerosp Sci Technol 25(1):177–189. doi:10.1016/j.ast.2012.01.006

Han Z-H, Zhang K-S, Liu J, and Song W-P (2013) Surrogate-based aerodynamic shape optimization with application to wind turbine airfoils, 51th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, Grapevine, Texas, AIAA Paper 2013–1108

Hoyle N, Bressloff NW, Keane AJ (2006) Design optimization of a two-dimensional subsonic engine air intake. AIAA J 44(11): doi: 10.2514/1.16123

Jameson A (1988) Aerodynamic design via control theory. J Sci Comput 3:3

Jameson A, Pierce NA, and Martinelli L (1997) Optimum aerodynamic design using the Navier-Stokes equations, AIAA Paper 97–0101

Jeong S, Murayama M, Yamamoto K (2005) Efficient optimization design method using kriging model. J Aircr 42(2): doi: 10.2514/1.6386

Jones DR (2001) A taxonomy of global optimization methods based on response surfaces. J Glob Optim 21:345–383

Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. J Glob Optim 13:455–492

Kanazaki M, Tanaka K, Jeong S, Yamamoto K (2007) Multi-objective aerodynamic exploration of elements' setting for high-lift airfoil using kriging model. J Aircr 44(3): doi: 10.2514/1.25422

Keane AJ (2006) Statistical improvement criteria for use in multiobjective design optimization. AIAA J 44:4

Koziel S, Leifsson L (2013) Surrogate-based aerodynamic shape optimization by variable-resolution models. AIAA J 51(1): doi: 10.2514/1.j051583

Krige DG (1951) A statistical approach to some basic mine valuations problems on the Witwatersrand. J Chem Metall Min Eng Soc S Afr 52(6):119–139

Kulfan BM (2008) Universal parametric geometry representation method. J Aircr 45(1):142–158

Laurenceau J, Sagaut P (2008) Building Efficient response surfaces of aerodynamic functions with kriging and cokriging. AIAA J 46(2): 498–507

Laurenceau J, Meaux M, Montagnac M, Sagaut P (2010) Comparison of gradient-based and gradient-enhanced response-surface-based optimizers. AIAA J 48(5): doi: 10.2514/1.45331

Liu J, Han Z-H, and Song W-P (2012a) Efficient kriging-based optimization design of transonic airfoils: some key issues, 50th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, Nashville, Tennessee, AIAA Paper 2012–0967, Jan. 2012

Liu J, Han Z-H, and Song W-P (2012b) Comparison of infill sampling criteria in kriging-based aerodynamic optimization, 28th International Congress of the Aeronautical Sciences, Brisbane, Australia

Matheron GM (1963) Principles of geostatistics. Econ Geol 58(8):1246–1266

McKay MD, Beckman RJ, Conover WJ (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21(2): 239–245

Redeker G (1994), DLR-F4 wing body configuration, In AGARD, A Selection of Experimental Test Cases for the Validation of CFD Codes, Vol. 2

Sacks J, Welch WJ, Mitchell TJ, Wynn HP (1989) Design and analysis of computer experiments. Stat Sci 4:409–423

Schonlau M (1997) Computer experiments and global optimization, PhD thesis, University of Waterloo

Simpson TW, Mauery TM, Korte JJ, Mistree F (2001) Kriging models for global approximation in simulation-based multidisciplinary design optimization. AIAA J 39:12

Sobester A, Leary SJ, Keane AJ (2004) A parallel updating scheme for approximation and optimization high fidelity computer simulations. Struct Multidiscip Optim 27:371–383. doi:10.1007/s00158-004-0397-9

Sobester A, Leary SJ, Keane AJ (2005) On the design of optimization strategies based on global response surface approximation models. J Glob Optim 33:31–59. doi:10.1007/s10898-004-6733-1

Song W, Keane AJ (2007) Surrogate-based aerodynamic shape optimization of a civil aircraft engine nacelle. AIAA J 45(10): doi: 10.2514/1.30015

Vavalle A, Qin N (2007) Iterative response surface based optimization scheme for transonic airfoil design. J Aircr 44(2): doi: 10.2514/1.19688

Viana FAC, and Haftka RT (2010) Surrogate-based optimization with parallel simulations using the probability of improvement, 13th AIAA/ISSMO Multidisciplinary Analysis Optimization Conference, Fort Worth, Texas, AIAA-2010-9392

Yao W, Chen X, Huang Y, van Michel M (2013) A surrogate based optimization method with RBF neural network enhanced by linear interpolation and hybrid infill strategy. Optim Methods Softw. doi:10.1080/10556788.2013.777722

Yun Y, Yoon M, and Nakayama H (2009) Multi-objective optimization based on meta-modeling by using support vector regression. Optim Eng 10:doi: 10.1007/s11081-008-9063-1

Zhang Y, Han Z -H, Shi L -X, and Song W -P (2016) Multi-round surrogate-based optimization for benchmark aerodynamic design problems, 54th AIAA Aerospace Science Meeting, 4–8 January 2016, San Diego, California, USA, AIAA Paper 2016–1545