Engineering Optimization

ISSN: 0305-215X (Print) 1029-0273 (Online) Journal homepage: http://www.tandfonline.com/loi/geno20

Horsetail matching: a flexible approach to optimization under uncertainty

L. W. Cook & J. P. Jarrett

To cite this article: L. W. Cook & J. P. Jarrett (2017): Horsetail matching: a flexible approach to optimization under uncertainty, Engineering Optimization, DOI: 10.1080/0305215X.2017.1327581

To link to this article: http://dx.doi.org/10.1080/0305215X.2017.1327581

Published online: 06 Jun 2017.

ENGINEERING OPTIMIZATION, 2017
https://doi.org/10.1080/0305215X.2017.1327581

Taylor & Francis
Taylor & Francis Group

Horsetail matching: a flexible approach to optimization under uncertainty*

L. W. Cook and J. P. Jarrett

Department of Engineering, University of Cambridge, Cambridge, UK

ABSTRACT

It is important to design engineering systems to be robust with respect to uncertainties in the design process. Often, this is done by considering statistical moments, but over-reliance on statistical moments when formulating a robust optimization can produce designs that are stochastically dominated by other feasible designs. This article instead proposes a formulation for optimization under uncertainty that minimizes the difference between a design's cumulative distribution function and a target. A standard target is proposed that produces stochastically non-dominated designs, but the formulation also offers enough flexibility to recover existing approaches for robust optimization. A numerical implementation is developed that employs kernels to give a differentiable objective function. The method is applied to algebraic test problems and a robust transonic airfoil design problem where it is compared to multi-objective, weighted-sum and density matching approaches to robust optimization; several advantages over these existing methods are demonstrated.

ARTICLE HISTORY
Received 25 November 2016
Accepted 28 April 2017

KEYWORDS
optimization under uncertainty; horsetail matching; robust optimization; density matching; probabilistic methods

1. Introduction

A traditional optimization considers a quantity of interest of a system q as a function of controllable design variables $\boldsymbol{x}$, to find a design that minimizes q. However, in practice, a computational simulation of q will be affected by uncontrollable uncertainties, $\boldsymbol{u}$, from a variety of sources (Beyer and Sendhoff 2007; Kennedy and O'Hagan 2001). A design optimized deterministically is often sensitive to variations in $\boldsymbol{u}$ and will see a degraded performance when realized (Huyse, Padula, Lewis, and Li 2002; Keane and Nair 2005). Therefore the importance of including uncertainties in the design process is becoming increasingly recognized: a designer instead defines a measure of the behaviour of the quantity of interest q under uncertainty as the objective function to optimize.

Formulating this problem effectively for engineering design is the field of robust optimization (RO); a good overview of available RO methods is provided in Beyer and Sendhoff (2007). Many methods rely on statistical moments or a single point on the uncertainty distribution, but this can lead to designs that a designer would not select if they had access to the distributions of all possible designs (discussed further in Section 2). Instead, other methods attempt to optimize the entire distribution of the quantity of interest. The most direct application of this philosophy is the recently developed density matching approach presented in Seshadri, Constantine, Iccarino, and Parks (2016), where a distance metric between a design's probability density function (PDF) and a designer-specified target PDF is minimized. Additionally, in Petrone, Iaccarino, and Quagliarella (2011), an approach is presented that minimizes the area between sections of a design's cumulative distribution function (CDF)

CONTACT L. W. Cook lwc24@cam.ac.uk

*Horsetail matching is available as a python package from http://www-edc.eng.cam.ac.uk/aerotools/horsetailmatching/

and an ideal target in a multi-objective formulation. This evolved into a formulation using the generalized inverse CDF (Quagliarella, Petrone, and Iaccarino 2014), where the values of the inverse at different CDF values are treated as objectives in a multi-objective formulation.

This article builds on these ideas and proposes a single objective, differentiable approach that minimizes the difference between a design's CDF and a target. It is shown to out-perform the density matching approach and overcome limitations of moment-based approaches while retaining significant flexibility.

2. Background

Many current applications of RO treat the first two statistical moments (mean μ and variance σ^2) as separate objectives and utilize techniques from the field of multi-objective (MO) optimization, since the moments are often competing objectives (Jin and Sendhoff 2003; Keane 2009). A pure MO approach can be used in order to find the robust Pareto front (which gives the trade-off between mean performance and robustness) (Dodson and Parks 2009; Ghisu, Jarrett, and Parks 2011; Keane 2009; Lee, Periaux, Onate, Gonzalez, and Qin 2011), but this is computationally expensive, so often they are combined into a single objective using a weighted sum (WS) (Lee and Kwon 2006; Padulo, Campobasso, and Guenov 2011; Zhang and Hosder 2013). Alternatively, the robust counterpart approach (sometimes known as the robust regularization approach) (Ryan, Lewis, and Yu 2015; Yao, Chen, Luo, Van Tooren, and Guo 2011) is essentially a 'minimax' optimization where the worst case of q over the uncertainty space, $\mathcal{U}$, is minimized; it is therefore thought of as a conservative approach.

However, μ and σ^2 are not independent: they are both properties of the underlying distribution of q. Therefore the authors argue there is only ever so much penalty to μ a designer is willing to accept in order to decrease σ^2, and the MO approach does not take this into account. In this article, the qualitative purpose of doing a basic robust optimization is taken to be finding a design that *maximizes the likelihood of achieving as good a performance as possible*, in which case it makes more sense to deal with probability distributions in their entirety instead of considering μ and σ^2 as competing.

To illustrate this, consider the PDFs in Figure 1(a), which represent designs on a hypothetical robust Pareto front. Even though C is significantly more robust than A, since there is no overlap between the PDFs, design A is guaranteed to give a better quantity of interest than C, and so a designer is unlikely to choose C as the final design. The PDF of design B overlaps that of A, but it appears that there is quite a large penalty to the mean.

Using the PDFs it is hard to quantify which design would be preferred, but now consider Figure 1(b): the CDFs of designs A and B do not cross at any point, meaning that at any given value

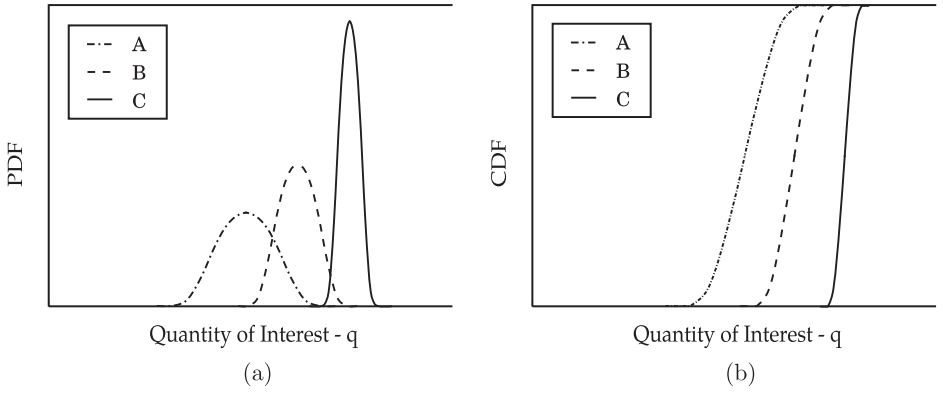

Figure 1. Hypothetical PDFs and CDFs for designs on a robust Pareto front resulting from a traditional MO robust optimization. (a) PDFs. (b) CDFs.

for the quantity of interest q, design A is always more likely to achieve this value or better than design B (from the definition of the CDF). The authors argue that design A is a superior design under the basic robust optimization philosophy. This is known as stochastic dominance, and its importance has been recognized in fields such as decision making under uncertainty and stochastic programming (Levy 2015; Shapiro, Dentcheva, and Ruszczynski 2009); but the authors argue that it is a concept that is often overlooked in the optimization under uncertainty literature for engineering design. This article uses the following definition of stochastic dominance.

Definition 2.1: Design x_A *stochastically dominates* design x_B (or design x_B is *stochastically dominated by* design x_A) if

$$\forall q \in \mathcal{Q}, \quad F_{x_A}(q) > F_{x_B}(q), \tag{1}$$

or equivalently

$$\forall h \in [0,1], \quad F^{-1}_{x_A}(h) < F^{-1}_{x_B}(h), \tag{2}$$

where $F_x(q) : \mathcal{Q} \to [0,1]$ and $F_x^{-1}(h) : [0,1] \to \mathcal{Q}$ are, respectively, the CDF and the inverse CDF of q for a design x (the inverse exists because the CDF is non-decreasing by definition), and $\mathcal{Q}$ is the set of feasible values of q.

From these arguments, finding the full robust Pareto front can be unnecessary and waste computational effort, since often there will be many designs that are inferior according to Definition 1. This philosophy to robust design is also evident in the development of the CDF-based methods of Petrone *et al.* (2011) and Quagliarella *et al.* (2014), where the value of q for different sections/values of the CDF are minimized or traded-off: these methods avoid stochastically dominated designs. Additionally, only considering mean and variance does not take into account higher-order moments or the tails of the distribution, and so loses information which might be important to a designer; this was one of the motivations behind the density matching approach in Seshadri *et al.* (2016).

Often it is too computationally expensive to perform a full MO approach to find the robust Pareto front, and so the weighted-sum approach is ubiquitous in the literature on robust optimization in practical engineering applications, since it has the appeal of being a single objective formulation and so is tractable with efficient gradient-based optimization algorithms. However, it has several drawbacks, even in pure multi-objective optimization (Marler and Arora 2010). Primarily, the weights must be set *a priori* by the designer, and it is difficult to know beforehand where on the Pareto front the optimum solution will be, since the shape of the front and relative magnitude of the two objectives on the front are not known before performing the optimization. This drawback is made worse in the case of robust optimization since, as discussed, the robust Pareto front is likely to contain stochastically dominated designs, to which a weighted-sum optimization could converge under certain combinations of weights. Alternatively, robustness can be controlled via a constraint on the variance, and just the mean can be optimized, but this approach can still give rise to stochastically dominated designs if the constraint on the variance is too strict.

Definition 2.1 suggests the use of the CDF in a robust optimization formulation in order to avoid stochastically dominated designs, and this is part of the motivation behind the development of the proposed horsetail matching technique.

3. Horsetail matching

As illustrated in Figure 2, the horsetail matching concept involves minimizing the difference between a design's CDF and a target.

The approach is named horsetail matching because the CDF is a special case, where all the uncertainties are probabilistic, of a more general quantification of uncertainty which consists of the bounds on the CDF and is sometimes referred to as a 'horsetail plot' (or as a 'p-box'). Planned future work

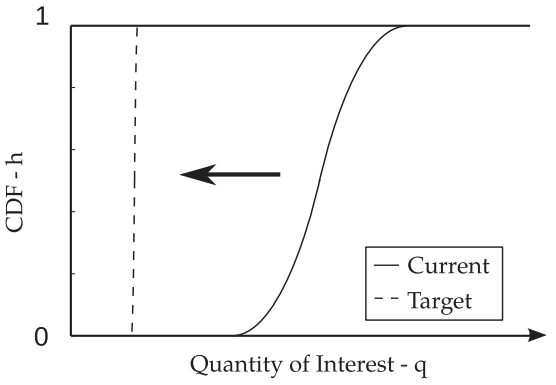

Figure 2. The horsetail matching concept (essentially CDF matching under probabilistic uncertainties).

aims to extend the proposed formulation to this more general case of optimization under uncertainty, and so the name 'horsetail matching' is introduced here (the term horsetail does not imply any assumption on the shape of the distributions).

3.1. Formulation

The quantity of interest, $q \in \mathcal{Q}$, is assumed to be a continuously differentiable (of class C_1), scalar-valued function of the n_x design variables $\boldsymbol{x}$ and n_u probabilistic uncertainties $\boldsymbol{u}$ which are assumed to be independent and each defined by a given probability distribution. The design variables must lie in the design space $\mathcal{X}$, which is defined by upper bounds $\boldsymbol{x}^u$, lower bounds $\boldsymbol{x}^l$ and n_g inequality constraints g_j: $\mathcal{X} : \{\boldsymbol{x} \in \mathbb{R}^{n_x} | x_k^l < x_k < x_k^u \forall k = 1, \ldots, n_x \text{ and } g_j(\boldsymbol{x}) \leq 0 \forall j = 1, \ldots, n_g\}$. At a given design $\boldsymbol{x}$, let $q_x(\boldsymbol{u}) = q(\boldsymbol{x}, \boldsymbol{u})$, let $F_x(q) : \mathcal{Q} \rightarrow [0, 1]$ and $F_x^{-1}(h) : [0, 1] \rightarrow \mathcal{Q}$ be, respectively, the CDF and inverse CDF of q_x, and let the target be given by $t(h) : [0, 1] \rightarrow \mathbb{R}$. The following L_2 norm is proposed as the measure of difference between a design's CDF and the target:

$$d_{hm}(\boldsymbol{x}, t) = \left(\int_0^1 \left(F_x^{-1}(h) - t(h) \right)^2 \mathrm{d}h \right)^{1/2}, \tag{3}$$

where $h \in [0, 1]$ represents the CDF value. The optimization problem becomes finding $\boldsymbol{x}^*$ such that

$$\boldsymbol{x}^* = \underset{\boldsymbol{x} \in \mathcal{X}}{\operatorname{argmin}}\, d_{hm}(\boldsymbol{x}; t), \tag{4}$$

where the value x^* corresponds to the optimal design under uncertainty: its behaviour under uncertainty is as close as possible to that specified in the target.

The integral in Equation (3) implies that the inverse CDF is well defined. In the implementation of horsetail matching proposed in Section 4, this condition is both always satisfied and unnecessary, however in general the metric in Equation (3) can be well defined for CDFs that are not strictly monotonically increasing by using the generalized inverse CDF: $F_x^{-1}(h) = \inf \{q_x \in \mathbb{R} : F_x(q) > h\}$.

3.2. Discussion

Utilizing the entire distribution of a quantity of interest avoids losing information by extracting just the first couple of moments. However, it was noted in the development of density matching (Seshadri *et al.* 2016) that requiring a target in an optimization formulation placed a lot of responsibility on the designer, since if the target is not feasible then density matching performs poorly (discussed further

in Section 5.2); so it might seem that requiring a target for horsetail matching restricts the approach. In contrast, the target provides horsetail matching with considerable flexibility.

To capture the basic robust optimization philosophy of 'maximizing the likelihood of achieving as good a performance as possible' a standard target is proposed where the target is set to a constant value of q_{ideal} that is beyond achievable. Here a standard target is defined to be $t(h) = q_{\text{ideal}}$ where $q_{\text{ideal}} \leq \inf(\mathcal{Q})$ (recall that $\mathcal{Q}$ is the set of all feasible values of q). In most engineering design problems, it is trivial to identify a value for q_{ideal} (*e.g.* zero cost, 100% efficiency, zero weight). Under this standard target, the horsetail matching metric in Equation (3) has the following important properties.

Property 3.1: When a standard target is used ($t(h) = q_{\text{ideal}} \leq \inf(\mathcal{Q})$), then the minimizer of d_{hm} will not be stochastically dominated by any other design $x \in \mathcal{X}$.

Proof summary: This follows directly from Definition 2.1. ■

Property 3.2: When a standard target is used ($t(h) = q_{\text{ideal}} \leq \inf(\mathcal{Q})$), then the minimizer of d_{hm} will lie on the Pareto front of μ and σ^2.

Proof summary: For a given design $\boldsymbol{x}$, let $q \in \mathcal{Q} \subseteq \mathbb{R}^+$, and define s as the realization of the random variable S such that $s = q + c$. Also define F_s, and F_s^{-1} where $F_s(q + c) = F_x(q)$ and $F_s^{-1}(h) = F_x^{-1}(h) + c$ for a constant c such that $t_s(h) = t(h) + c = 0$. Use a change of variables to show that minimizing d_{hm} is equivalent to minimizing the second statistical moment of s, $\mathbb{E}(S^2)$, noting that $s > 0$ and $d_{hm} > 0$:

$$\begin{aligned} d_{hm}^2 &= \int_{F(0)=0}^{F(\infty)=1} \left(F_x^{-1}(h) - t(h)\right)^2 \mathrm{d}h = \int_{F_s(0)=0}^{F_s(\infty)=1} \left(F_s^{-1}(h) - 0\right)^2 \mathrm{d}h \\ &= \int_0^\infty \left(F_s^{-1}(F_s(s))\right)^2 \frac{\mathrm{d}F_s(s)}{\mathrm{d}s}\,\mathrm{d}s = \int_0^\infty s^2 \frac{\mathrm{d}F_s(s)}{\mathrm{d}s}\,\mathrm{d}s = \mathbb{E}(S^2). \end{aligned} \tag{5}$$

Since s is just a translation of q by c: $\mathbb{V}(S) = \sigma^2 = d_{hm}^2 - \mathbb{E}(S)^2$ and $\mathbb{E}(S) = \mu + c$. Therefore reducing σ^2 with μ fixed must also reduce d_{hm}, and similarly reducing μ with σ^2 fixed must also reduce d_{hm}. Hence there cannot be a design that dominates $\boldsymbol{x}$ that does not also reduce d_{hm}, and so the minimizer of d_{hm} must lie on the Pareto front of μ and σ^2. ■

Properties 3.1 and 3.2 illustrate how using this metric along with a standard target is well suited to maximizing the likelihood of achieving as good a performance as possible, by obtaining designs on the robust Pareto front but avoiding stochastically dominated designs.

Further, by modifying the standard target, several different design scenarios can be considered, offering considerable flexibility to a designer; these are listed in the following and are illustrated in Figure 3.

- *Risk-averse*. By modifying the shape of the standard target, a designer can specify a preference for risk-averse designs so that robustness is preferred over possible high performance: skewing the shape of $t(h)$ for h close to 1 to lower values of q emphasizes minimizing the worst cases of q over the CDF.
- *Risk-seeking*. A designer can alternatively emphasize possible performance over robustness by modifying the standard target in the opposite sense to the risk-averse targets.
- *Feasible distribution*. In some applications, a designer might care more about higher moments of individual CDFs such as skewness. In this case, target distributions over feasible ranges of q with desirable higher-order moment properties can be provided (this was the main advantage of the density matching approach of Seshadri *et al.*, 2016).

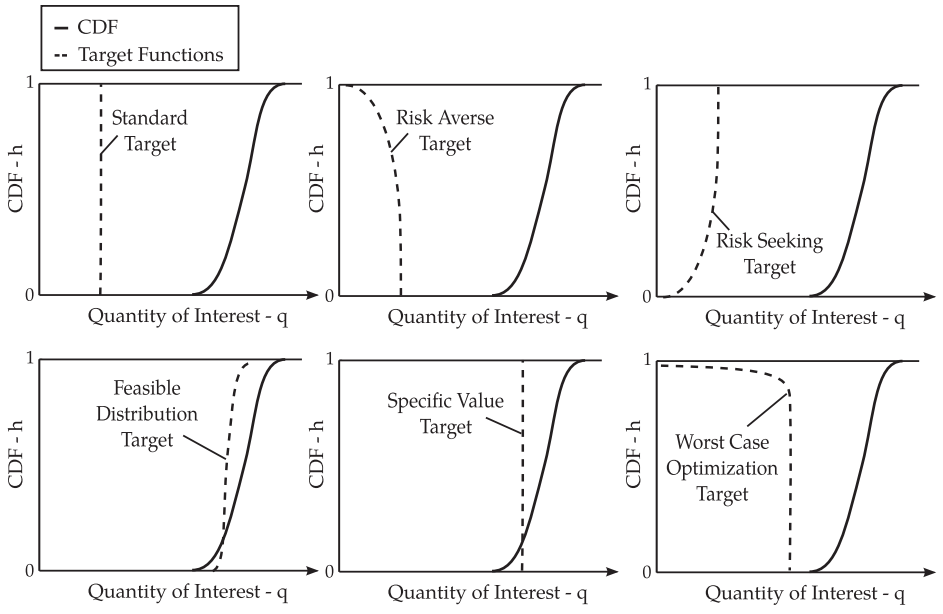

Figure 3. Examples of targets that specify different design scenarios using the horsetail matching formulation.

- *Specific value*. In other applications, for example where a component of a larger system is being designed, pure minimization may not be what is required and instead a target value of q_{targ} is desired: this is implemented in the horsetail matching formulation by setting $t(h) = q_{\text{targ}}$.
- *Worst-case optimization*. If the risk-averse target is taken to the extreme, such that $t(1) \to -\infty$, then only the worst-case value is optimized and the horsetail matching formulation reduces back to the robust regularization approach (Beyer and Sendhoff 2007; Ryan *et al.* 2015).

Although the point on the Pareto front of μ and σ^2 obtained using the standard target is arbitrary, by using different magnitudes of risk-averse or risk-seeking targets the minimizer of d_{hm} will move along the Pareto front (assuming there is a trade-off for the problem being considered): this is explored further in Section 5.1. This can be considered as an alternative to changing the weightings on μ and σ^2 in the weighted-sum formulation of robust optimization. However, by using these targets, there is no risk of obtaining stochastically dominated designs, and a preference for risk-seeking behaviour can be specified—something not possible when only considering the first two statistical moments.

It is possible to use a generalized L_p norm instead of the proposed L_2 norm as a difference metric between a CDF and the target, defined as

$$d_p = \left(\int_0^1 |F_x^{-1}(h) - t_0(h)|^p \, \mathrm{d}h \right)^{1/p}. \tag{6}$$

For example, using an L_1 norm minimizes the area between the current distribution and the target, and when the target $t(h)$ is a constant value this is essentially the 'robustness index' proposed in Petrone *et al.* (2011). Minimizing this area is equivalent to minimizing the mean, since the mean can be obtained from the CDF for strictly positive q (which can be enforced via a simple shift or transformation for practical problems) by

$$\mu = \int_0^\infty (1 - F_x(q)) \, \mathrm{d}q \equiv \int_0^1 F_x^{-1}(h) \, \mathrm{d}h, \tag{7}$$

and so $p = 1$ takes no account of the variance or higher-order moments of the distribution. As p is increased, sections of the CDF further from the target would be penalized more heavily, emphasizing robustness more than mean performance.

However, exactly the same effect can be achieved by varying the shape of the target, as shown with Property 3.3.

Property 3.3: Given a design, $\boldsymbol{x}^*$, that is an optimum of the L_p norm of the difference between an inverse CDF and a specified target $t_0(h)$, a target $t(h)$ exists such that $\boldsymbol{x}^*$ is also an optimum of the L_2 norm (the horsetail matching metric d_{hm}) under this target.

Proof summary: Let $f^* = F^{-1}_{\boldsymbol{x}^*}(h)$ be the inverse CDF at a design, $\boldsymbol{x}^*$, that is an optimum of d_p. It must also be an optimum of d_p^p since there is a one-to-one correspondence between d_p and d_p^p. The gradient of d_p^p is given by

$$\frac{\mathrm{d}(d_p^p)}{\mathrm{d}\boldsymbol{x}} = \int_0^1 p(f^* - t_0)^{(p-1)} \frac{\mathrm{d}f^*}{\mathrm{d}\boldsymbol{x}} \,\mathrm{d}h. \tag{8}$$

At this point the gradient of d_{hm}^2 is given by

$$\frac{\mathrm{d}(d_{hm}^2)}{\mathrm{d}\boldsymbol{x}} = \int_0^1 2(f^* - t) \frac{\mathrm{d}f^*}{\mathrm{d}\boldsymbol{x}} \,\mathrm{d}h. \tag{9}$$

If $t = f^* - (p/2)(f^* - t_0)^{(p-1)}$, then these two gradients are equal. Therefore if $\boldsymbol{x}^*$ is a local optimum of d_p^p, it must also be a local optimum of d_{hm}^2, since it must satisfy the same first-order optimality conditions in both cases if the gradients are equal. Hence it is an optimum of d_{hm} as there is a one-to-one correspondence between d_{hm}^2 and d_{hm}. ■

Even though the target is unknown (since f^* is unknown) *a priori*, the fact that it exists illustrates that varying the power p in the norm to control robustness is equally as arbitrary as varying the shape of the target. However, as illustrated with Figure 3, varying the shape of the target offers additional flexibility to the designer and so using d_{hm} for all design scenarios is proposed so that the numerical implementation remains consistent. This is advantageous because using d_{hm} enables a differentiable implementation of the horsetail matching formulation that leverages gradient information on q (explored further in Section 4).

4. Numerical implementation

To approximate d_{hm}, a numerical evaluation of an integral which can be expressed in the two forms given in Equation (10) is required:

$$D = \int_0^1 \left(F_x^{-1}(h) - t(h)\right)^2 \mathrm{d}h = \int_{-\infty}^{\infty} \left(q - t(F_x(q))\right)^2 F_x'(q) \,\mathrm{d}q. \tag{10}$$

The goal is to find a numerical approximation, $\hat{D}$, to this integral in order to give an approximation, $\hat{d}_{hm} = \hat{D}^{1/2}$, to d_{hm}. Note that in the following ∇_x refers to the gradient vector of a quantity, and $\partial/\partial x_k$ refers to the gradient with respect to a single design variable—the k^{th} component of ∇_x.

4.1. Approximating the CDF curves

In order to find $\hat{D}$, first an expression for the CDF at a design $\boldsymbol{x}$, $F_x(q)$, is required. In many practical applications, a (nonlinear) simulation is used to evaluate q and an exact form of the CDF of an output under probabilistic uncertainties is not available, so a method of estimating the CDF is necessary.

Kernel density estimation (Scott 1992) was successfully applied to PDF matching in Seshadri *et al.* (2016); this work is built upon here. For horsetail matching, the integral of kernel functions is used and so it is not using kernel density estimation directly, rather kernels are used as a method of finding a smoothed, differentiable version of the empirical CDF (which can be recovered by using step function kernels). By evaluating $q_x(\boldsymbol{u})$ at M samples of the uncertainty to obtain values of $q_j = q_x(\boldsymbol{u}_j)$, $j = 1, \ldots, M$, the estimation of the CDF at a point q is given by Equation (11):

$$F_x(q) \simeq \frac{1}{M} \sum_{j=1}^{M} \Phi(q - q_j), \tag{11}$$

where $\Phi(r) = \int_{-\infty}^{r} K(r')\, \mathrm{d}r'$, for kernel function K. In this article, a Gaussian kernel is used, and so $\Phi(r)$ is the error function; both K and Φ use a bandwidth parameter b:

$$K(r) = \frac{1}{\sqrt{2\pi b^2}} \exp\left(-\frac{r^2}{2b^2}\right), \quad \Phi(r) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{r}{\sqrt{2b^2}}\right)\right). \tag{12}$$

Similarly, each component of the gradient, $\nabla_x(F_x(q))$, of this CDF approximation with respect to the design variables is given by Equation (13):

$$\frac{\partial F_x(q)}{\partial x_k} \simeq \frac{1}{M} \sum_{j=1}^{M} K(q - q_j)(-1) \frac{\partial q_j}{\partial x_k}, \tag{13}$$

which requires the sensitivities of the quantity of interest to the design variables at given values of the uncertainties $\boldsymbol{u}_j$, $\partial q_j / \partial x_k$.

The use of Gaussian kernels does not reflect an assumption on the type of distribution the CDF will be; it is simply so that a smooth, differentiable estimate of the CDF is obtained to facilitate fast convergence of gradient-based optimizers. If Gaussian kernels are considered to assume too much about the CDF, the integral of any kernel function could be used for Φ since its derivative is the original kernel function, K, so is defined everywhere. Gaussian kernels are used because they have been shown to give good optimization performance both in preliminary experiments and in the density matching approach of Seshadri *et al.* (2016).

Prior to an optimization, the kernel bandwidth for use in Equation (11) is selected and fixed throughout the optimization (otherwise the gradient of d_{hm} would have an additional term due to the changing bandwidth parameter). One way of selecting this bandwidth is to apply Scott's rule (Scott 1992) to samples at the initial design, but it is worth noting that a poor choice of bandwidth can lead to a highly non-smooth gradient (if b is too small) or can give smooth but erroneous values of d_{hm} and its gradient (if b is too large).

4.2. Evaluating the integral

To find $\hat{D}$, numerical quadrature of the integral in Equation (10) is performed, which can be expressed as a summation of N quadrature points and corresponding weights w_i:

$$\hat{D} = \sum_{i=1}^{N} w_i \big(F^{-1}(h_i) - t(h_i)\big)^2 = \sum_{i=1}^{N} w_i \big(q_i - t(F_x(q_i))\big)^2. \tag{14}$$

The trapezium rule is used for the numerical integration: N fixed points q_i are chosen, the kernel-based approximation to $F_x(q)$ in Equation (11) is used to obtain $h_i \simeq F(q_i)$ giving N pairs of (q_i, h_i), then the target is used to evaluate $t_i = t(h_i) = t(F(q_i))$ to obtain N pairs of (t_i, h_i) which are used along with the pairs of (q_i, h_i) in a trapezium rule integration of the second form of the integral

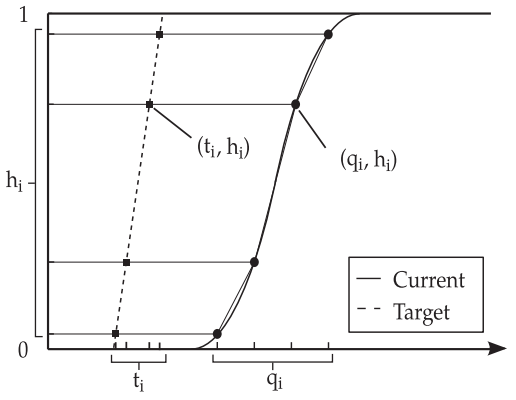

Figure 4. Illustration of the numerical integration procedure used to estimate the horsetail matching metric.

in Equation (10). The procedure is illustrated in Figure 4, and this numerical integration can be expressed as the vector–matrix–vector multiplication in Equation (15):

$$\hat{D} = (\boldsymbol{q} - \boldsymbol{t})^{\mathrm{T}} \boldsymbol{W} (\boldsymbol{q} - \boldsymbol{t}), \tag{15}$$

where $\boldsymbol{q}_i = q_i$ (fixed integration points), $\boldsymbol{t}_i = t(h_i)$, and the weighting matrix W is a diagonal matrix with entries given by

$$W_{i,i} = 0.5(h_{\min(i+1,N)} - h_{\max(1,i-1)}). \tag{16}$$

4.3. Evaluating the gradient

Expressing the numerical integration of the metric in the matrix form of Equation (15) facilitates the computation of the gradient of $\hat{D}$ efficiently:

$$\frac{\partial \hat{D}}{\partial x_k} = 2(\boldsymbol{q} - \boldsymbol{t})^{\mathrm{T}} \boldsymbol{W} \left(-\frac{\partial \boldsymbol{t}}{\partial x_k} \right) + (\boldsymbol{q} - \boldsymbol{t})^{\mathrm{T}} \frac{\partial \boldsymbol{W}}{\partial x_k} (\boldsymbol{q} - \boldsymbol{t}), \tag{17}$$

where

$$\left(\frac{\partial \boldsymbol{W}}{\partial x_k} \right)_{i,i} = \frac{\partial W_{i,i}}{\partial x_k} \left(e.g. \frac{\partial W_{1,1}}{\partial x_k} = 0.5 \left(\frac{\partial h_2}{\partial x_k} - \frac{\partial h_1}{\partial x_k} \right) \right), \tag{18}$$

$$\left(\frac{\partial \boldsymbol{t}}{\partial x_k} \right)_i = \frac{\partial t_i}{\partial x_k} = \frac{\mathrm{d}t(h)}{\mathrm{d}h} \frac{\partial h_i}{\partial x_k}. \tag{19}$$

In many practical applications, the simulation of $q(\boldsymbol{x}, \boldsymbol{u})$ will have existing capability to evaluate the sensitivities $\partial q_i / \partial x_k$ required by Equation (13) to find $\partial h_i / \partial x_k$, allowing horsetail matching to be implemented as a wrapper without any modification. In many cases this is especially important since sensitivities of q to the design variables are readily available at low computational cost—*e.g.* adjoints in CFD (Jameson, Martinelli, and Pierce 1998) can produce these sensitivities with one extra computational solve—and it is important to be able to use this information within an optimization to keep the computational cost feasible.

4.4. Response surfaces for computational efficiency

To get an accurate CDF to use in the formulation, a large number of samples of q_j should be obtained. However, when the simulation is computationally expensive, this is infeasible. Therefore a response surface to the quantity of interest over uncertainty space, $\mathcal{U}$, can be created using a small number of samples, and this response surface can then be sampled a large number of times in order to get a sufficiently resolved kernel density estimation. Similarly, a response surface can be fitted to each component of the sensitivity of q, $\partial q/\partial x_k$, in order to propagate the gradient. In this article, polynomial response surfaces are used, but any response surface can equally be used.

Prior to an optimization, M samples are drawn from the underlying probability distribution of $\boldsymbol{u}$, and these same samples are used throughout the optimization to evaluate the samples q_j. Since a response surface is being sampled, a large value of M can be selected to capture the CDF sufficiently such that the outcome of the optimization does not depend on the particular realization of these samples.

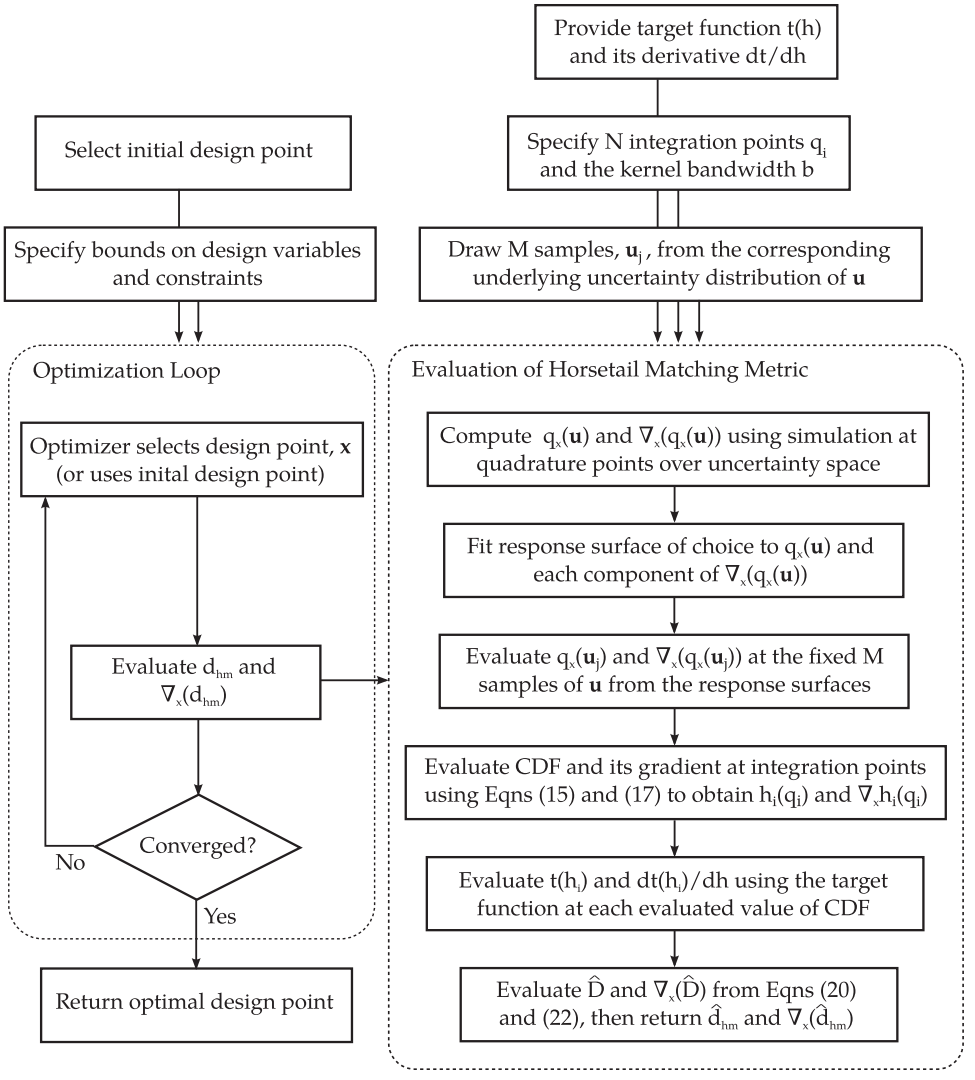

Figure 5. Flowchart outlining the implementation of horsetail matching optimization proposed in this article.

Many approaches for performing UQ within an optimization also rely on fitting a response surface to q over $\mathcal{U}$, such as non-intrusive polynomial chaos (Le Matre, Knio, Najm, and Ghanem 2001) and Kriging (Martin and Simpson 2005). Therefore, using a response surface in the numerical implementation proposed in this article is a limitation to the same extent as other methods that use response surfaces. However, it is still worth noting the computational cost of this proposed implementation may become infeasible when the dimensionality of $\mathcal{U}$ becomes too great for response surfaces to be effective; this is a limitation of the current implementation.

4.5. Overview

A flowchart of the numerical implementation is given in Figure 5.

5. Experiments and applications

5.1. Influence of the target

First, a simple algebraic problem is considered to illustrate how, when not being used as a feasible CDF, the target can be used to control preference for mean performance or robustness. The test problem uses one design variable, x (bounded by the interval $[0, 10]$), and has one uncertain parameter, u (uniformly distributed over the interval $[-1, 1]$), enabling the design space to be exhaustively searched:

$$q(x, u) = 1 + 8\arctan(x + 0.3) + \frac{1}{\arctan(x + 0.3)}(e^{1.5u} - 1). \tag{20}$$

For 30 values of x over the design space, a CDF is propagated as outlined in Section 4 using $M = 500$ and $N = 500$; from the response surface, the mean μ and standard deviation σ are also evaluated. In Figures 6(a) and 6(c), μ and σ for these designs are plotted in grey, giving the robust Pareto front, and in Figures 6(b) and 6(d) the corresponding CDFs are plotted in grey. The design that minimizes d_{hm} under the four targets given in Table 1 for two values for q_{ideal} are highlighted in Figure 6: for $q_{\text{ideal}} = -5$ in Figures 6(a) and 6(b), and for $q_{\text{ideal}} = -10$ in Figures 6(c) and 6(d).

It is clear from Figure 6 how the shape of the target can be used to control where on the Pareto front the horsetail matching optimum is located. It can also be seen that the majority of designs on the robust Pareto front give CDFs which are stochastically dominated according to Definition 2.1 by other designs: all the points to the right of the worst-case optimum on the Pareto front in Figure 6(a) are stochastically dominated by the worst-case optimum. This illustrates part of the motivation outlined in Section 2.

Additionally, moving the value of q_{ideal} further from the range of attainable values of q reduces the influence of the shape of the target on the optimum design (except the worst-case target). In this case the optimum designs move toward the minimum mean design, since moving the targets further from $\mathcal{Q}$ increases the contribution to d_{hm} of the sections of the CDF for h close to zero. Therefore, when $\mathcal{Q}$ is known, it is proposed that a designer chooses q_{ideal} to be as close as possible to feasible values of q but still beyond attainable in order to maximize the influence of the shape of the target, *i.e.* $q_{\text{ideal}} = \inf(\mathcal{Q})$. This choice corresponds to the examples of zero cost, zero weight, and 100% efficiency suggested for typical engineering design problems.

In the case where a designer does not have a good idea for an appropriate value to use as q_{ideal} in the standard target (a scenario that the authors argue is rare), an iterative approach could be adopted—a value of q_{ideal} is selected, and if it turns out that this guess was too conservative, a more ambitions value of q_{ideal} is selected and another optimization is performed. This can be repeated until the designer is content with the distribution of the optimum design. If, however, an appropriate target is selected initially, only a single optimization is required.

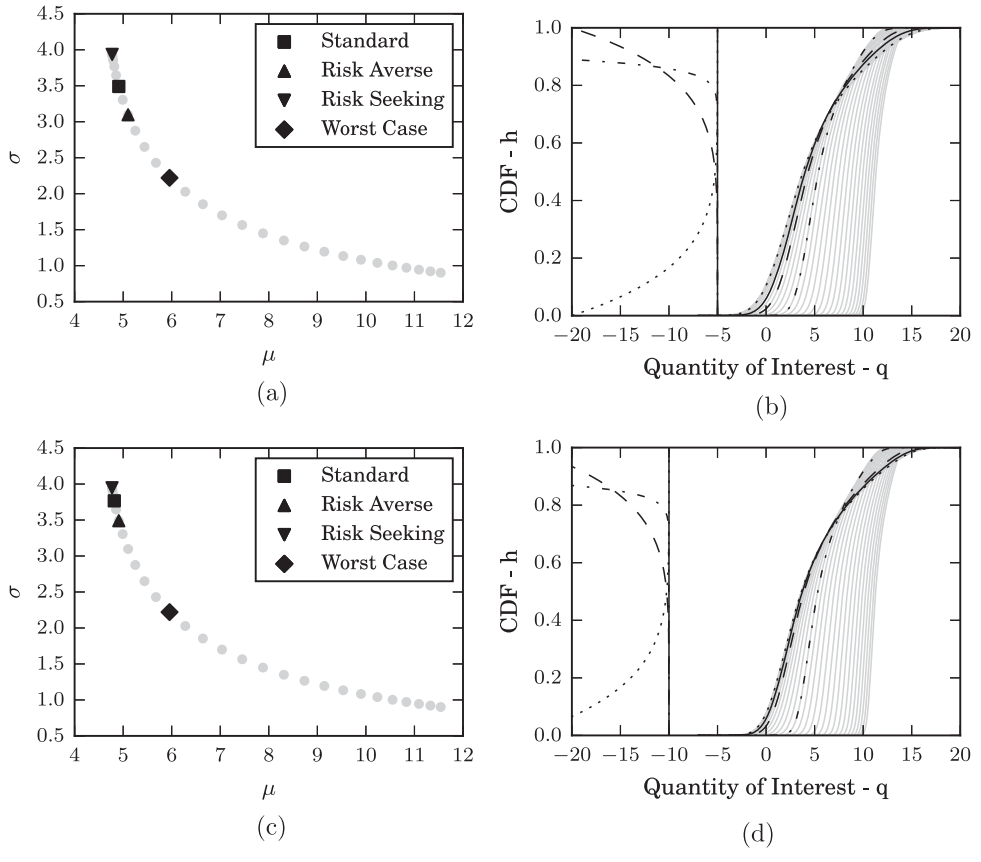

Figure 6. Robust Pareto front (left) with corresponding CDFs (right) for the test problem in Equation (20) in grey, along with the horsetail matching optimum designs under a standard target (square/solid), a risk-averse target (upward triangle/dashed), a risk-seeking target (downward triangle/dotted), and a worst-case optimization target (diamond/dotted and dashed). (a) Designs on the robust Pareto front for $q_{\text{ideal}} = -5$. (b) CDFs and targets for $q_{\text{ideal}} = -5$. (c) Designs on the robust Pareto front for $q_{\text{ideal}} = -10$. (d) CDFs and targets for $q_{\text{ideal}} = -10$.

Table 1. Equations for the different targets used for this test problem.

Target	Equation
Standard	$t(h) = q_{\text{ideal}}$
Risk averse	$t(h) = q_{\text{ideal}} - 15h^6$
Risk seeking	$t(h) = q_{\text{ideal}} - 15(1-h)^6$
Worst-case optimization	$t(h) = q_{\text{ideal}} - 500h^{30}$

5.2. Comparison with density matching

As mentioned in Section 1, density matching (Seshadri *et al.* 2016) is a recent method that the proposed approach builds upon, and so a detailed comparison to this method is warranted. Density matching sets a target PDF and minimizes a distance metric between a current design's PDF and the target using the (squared) L_2 norm:

$$d_{dm}(t, \boldsymbol{x}) = \int_{-\infty}^{+\infty} (t_{\text{pdf}}(q) - f_{\text{pdf}}(q))^2 \, \mathrm{d}q, \tag{21}$$

where f_{pdf} is the PDF of the quantity of interest q at the current design and t_{pdf} is the target PDF. This can also be implemented using kernel density estimation of the PDF, trapezium rule quadrature and matrix-based gradient—see Seshadri *et al.* (2016) for details.

Analytic illustration

One of the drawbacks of the density matching (DM) approach is the overlap problem, where if f_{pdf} and t_{pdf} do not overlap significantly, then there is very little information in the objective function about their similarity: d_{dm} is dominated by the shape of f_{pdf}. This can be illustrated by considering how the metrics d_{dm} and d_{hm} vary depending on the difference between the mean and variance of a design's distribution and that of the target. In Figure 7, contours of d_{hm} and d_{dm} are plotted over a design space consisting of a range of means and standard deviation values for a Gaussian distribution: $\mathcal{N}(\mu, \sigma^2)$ with the target also a Gaussian: $\mathcal{N}(\mu_{\text{target}} = 0, \sigma^2_{\text{target}} = 2^2)$.

The density matching landscape has large flat regions corresponding to where the distribution does not overlap the target significantly, and even when there is some overlap often the gradient points away from the minimum solution in design space (*e.g.* for $\mu \simeq 15$ and $\sigma \simeq 5$). In contrast, the horsetail matching (HM) landscape's gradient always points towards solutions closer in design space to the minimum (even when $\sigma < \sigma_{\text{target}}$). This makes the horsetail matching landscape easier to navigate for both gradient-based algorithms (*e.g.* SLSQP) and global algorithms (*e.g.* evolutionary algorithms), so it is expected that optimizers perform better (in terms of reaching the optimum solution using less computational effort) under a horsetail matching formulation.

In Seshadri *et al.* (2016), a two-step approach to optimizations is proposed to alleviate this overlap issue for density matching. For the first few steps of an optimization a large bandwidth is used for the kernel density estimate of the PDF, to give significant overlap, before switching to a more optimal bandwidth for the remainder of the optimization.

Numerical optimizations

To compare how the two approaches solve an optimization problem, a gradient-based optimizer with both HM and DM formulations is used on a nonlinear algebraic test function (to enable many optimizations to be run from random starting points). The test function has three design variables ($x_{1,2,3}$) contained in $\mathcal{X} = [-5, 5]^3$, and one uncertainty (u) uniformly distributed uncertainty over the interval $[-1, 1]$; q is to be minimized.

$$q(x_1, x_2, x_3, u) = 5 + x_1^2 + 2x_2u + x_3u^2 \tag{22}$$

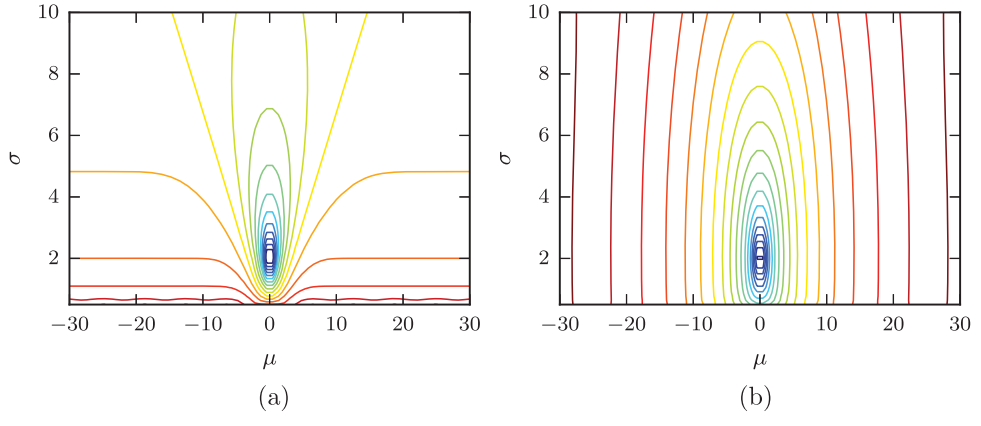

Figure 7. Density matching and horsetail matching metric contours (log scale) for $\mathcal{N}(\mu, \sigma^2)$ over a range of mean and variance values and a target distribution $\mathcal{N}(0, 2^2)$. (a) Density matching metric, d_{dm}. (b) Horsetail matching metric, d_{hm}.

An SLSQP optimization algorithm is run for 30 iterations, where one iteration corresponds to one evaluation of either the DM or HM metric and its gradient (implemented using the *SciPy.minimize* Python[tm] package[1] with the *SLSQP* option) from 50 randomly selected starting points.

A target that is slightly more ambitious than feasible distributions is used: a Gaussian with parameters $\mu = 3, \sigma = 1$. Figure 8 gives the convergence histories, of $\log_{10}(d_{hm})$ in both cases for comparison (the DM optimization is actually optimizing the d_{dm} metric), of each optimization run in grey and overlays the average of all 50 in black. It also gives the PDFs and CDFs of the initial design (plotted with the large initial bandwidth in the DM case), the target, and the solutions found by density matching (labelled 'DM soln') and horsetail matching (labelled 'HM soln') for an optimization run that found the global minimizer of the corresponding metrics.

In Figure 8, the HM optimizations converge more often, and faster on average, than the DM optimizations, illustrating improved computational efficiency. Furthermore, the global optimum under the DM metric is different from that under the HM metric. This is an important point: since the DM metric uses the L_2 norm integrated over q, it penalizes peaks in a design's PDF if they extend beyond the target, and rewards short and wide PDFs (non-robust designs) where the distribution does not overlap the target. In contrast, the HM metric intrinsically rewards peaks and penalizes tails in the PDF and hence prefers robust designs. This leads to the DM optimum being stochastically dominated by the HM optimum in Figure 8. Therefore, not only is the horsetail matching approach an easier, more computationally efficient problem for optimizers to solve than density matching, but it produces better designs when the target is more ambitious than what is actually achievable.

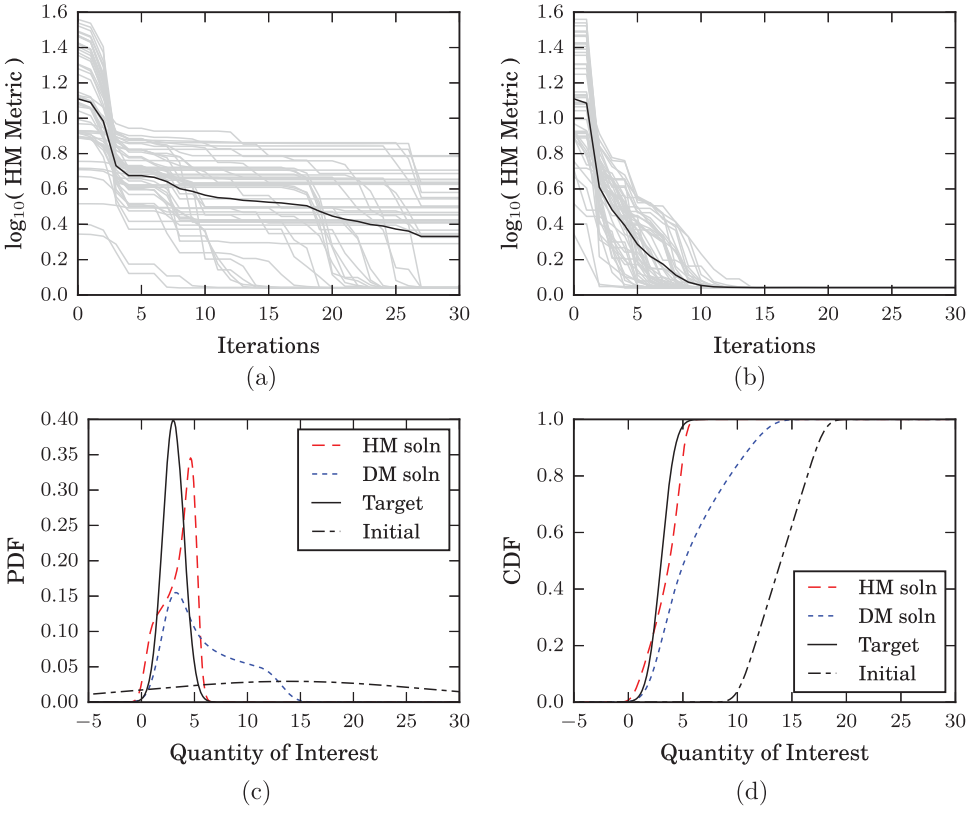

Figure 8. Results of gradient-based density matching and horsetail matching on the test problem with a Gaussian target distribution $\mathcal{N}(3,1)$. The metric values for the global optima are: d_{hm} for HM soln - 0.98, d_{hm} for DM soln - 7.06, d_{dm} for HM soln - 0.16, d_{dm} for DM soln - 0.088. (a) Density matching convergence. (b) Horsetail matching convergence. (c) PDFs of optimum designs. (d) CDFs of optimum designs.

5.3. Comparison with multi-objective RO and weighted-sum RO

Here, the nonlinear algebraic test problem in Equation (22) is used to compare horsetail matching to RO approaches that optimize mean and variance.

Firstly, a robust Pareto front (trading off mean and variance) is generated for the objective q using the NSGA-II algorithm (Deb, Pratap, Agarwal, and Meyarivan 2002) (implemented via the Python[tm] package *ECsPy*[2]) with a population size of 100 for 60 generations. Then a weighted-sum method is used to minimize $f_{ws} = \mu + \sigma^2$ (the WS metric) where the weighting on both μ and σ^2 is equal to one (also using the SLSQP method from the *SciPy.minimize* package, and using finite differencing to obtain the gradient). Finally, two horsetail matching optimizations are run, one with a standard target of $t(h) = -5$ (labelled 'HM' in Figure 9), and one with a uniform distribution target that has the same mean and variance as the weighted-sum optimum solution (labelled 'HM WS').

The results in μ, σ^2-space are plotted in Figure 9(a) and the CDFs are plotted in Figure 9(b), along with the CDF of the minimum μ design for comparison. Also, in Figure 10, optimization convergences for the weighted-sum approach and the HM approach under the standard target are compared from 50 random starting points.

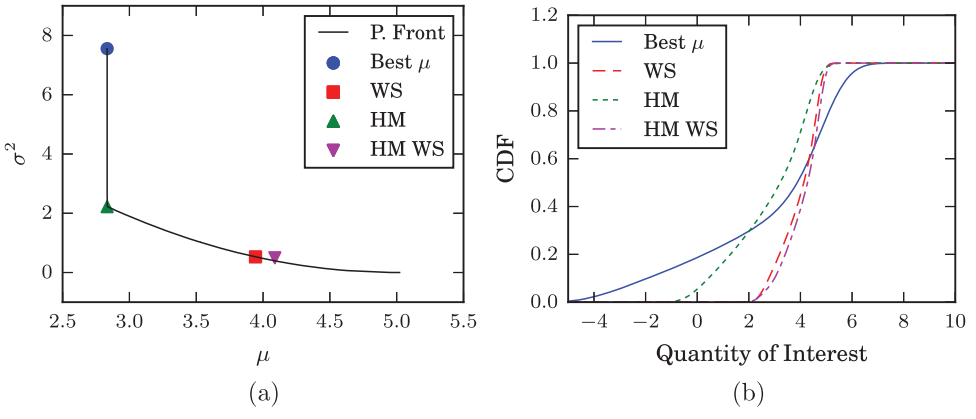

Figure 9. Robust Pareto front and optimum designs in μ, σ^2-space, and the CDFs of the optimum design. (a) Robust Pareto front. (b) CDFs of optimum designs.

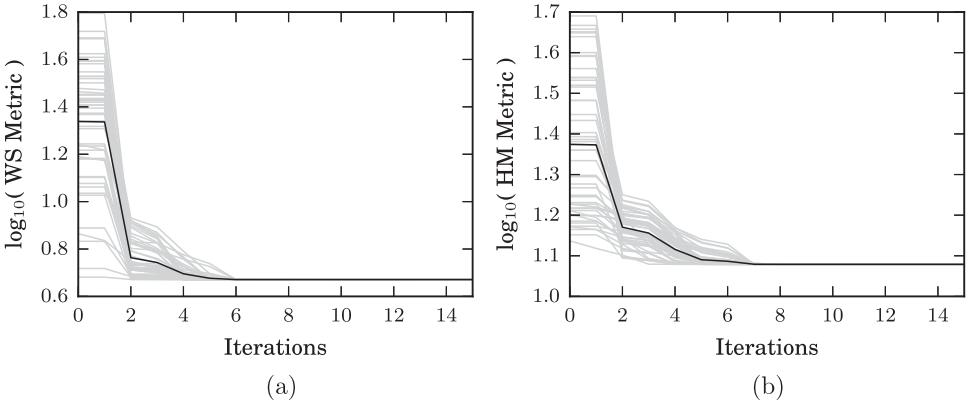

Figure 10. Convergence of weighted-sum approach (with $w_\mu = w_{\sigma^2} = 1$) and horsetail matching under the standard target ($t(h) = -5$). (a) Weighted-sum convergence. (b) Horsetail matching convergence.

From Figure 9 it can be seen that the WS approach produced a design on the Pareto front, and the HM approach with the uniform target found a similar design. It can also be seen that the design found by horsetail matching under a standard target produced a design that stochastically dominates this weighted-sum design. Although the weights for this optimization were chosen arbitrarily, it illustrates a limitation of both the MO and WS approaches to robust optimization discussed in Section 2: by treating mean and variance as separate objectives, the MO approach has wasted computational effort by finding designs on the Pareto front that are stochastically dominated, and similarly the WS method risks producing these designs under certain combinations of the weightings on μ and σ^2.

Additionally, comparing the minimum mean design and the HM optimum under the standard target: the CDFs are overlapping and neither is stochastically dominated by the other; however, the penalty to the mean is very small compared to the reduction in variance and so a designer is likely to choose the HM optimum over the minimum mean design since it is significantly more robust.

5.4. Lift-to-drag optimization of an Euler aerofoil

In order to validate these observations on a more practical test problem, here a robust aerofoil shape optimization is performed. The freely available SU2 CFD solver[3] is used, and a 2D transonic aerofoil under inviscid flow conditions is considered at an angle of attack of 5° and an uncertain Mach number uniformly distributed over the interval $[0.66, 0.69]$. The NACA0012 aerofoil and the unstructured mesh provided by SU2 for this aerofoil as a baseline are used, and a design space is parameterized using Hicks–Henne bump functions at the locations specified in Table 2—note that this is a similar design problem to that used to illustrate density matching in Seshadri *et al.* (2016).

The quantity of interest is lift-to-drag ratio, L/D, which is to be maximized, and so in this problem $q = 150 - L/D$ is minimized to keep q positive and the formulation as a minimization problem, therefore the CDFs of L/D plotted in Figure 11 are shifted complementary CDFs of q. This is mathematically equivalent to performing a horsetail matching optimization maximizing L/D, but a simple transformation is used to keep the implementation consistent.

First, a traditional MO robust design optimization is performed to obtain the Pareto front of $\mu_{L/D}$ and $\sigma_{L/D}$ to get an idea of the design space and for comparison purposes. For this, the NSGA-II algorithm is used with a population size of 50 for 60 generations. Third-order polynomial response surfaces are used at each design to propagate the statistical moments, so the NSGA-II optimization in total requires $50 \times 60 \times 4 = 12{,}000$ CFD runs. The resulting Pareto front is plotted in Figure 12 (labelled 'P. Front').

Next, horsetail matching and density matching are run using a feasible target with mean and standard deviation taken from a design on the Pareto front: a uniform target with $\mu_{L/D} = 60$ and $\sigma_{L/D} = 10$ (labelled 'HM' and 'DM'), and finally horsetail matching is run under a standard target where $t(h) = 10$ (equivalent to the target for L/D being 140, and labelled 'HM Stand.'). Fifth-order

Table 2. Surface (upper or lower), location, and limits (as a proportion of chord) for the Hicks–Henne bump functions that make up the design space.

Surface	Location	Limits
U	0.05	$\pm$ 0.001
U	0.15	$\pm$ 0.006
U	0.30	$\pm$ 0.009
U	0.45	$\pm$ 0.009
U	0.60	$\pm$ 0.006
U	0.80	$\pm$ 0.002
L	0.10	$\pm$ 0.001
L	0.30	$\pm$ 0.007
L	0.55	$\pm$ 0.007
L	0.80	$\pm$ 0.002

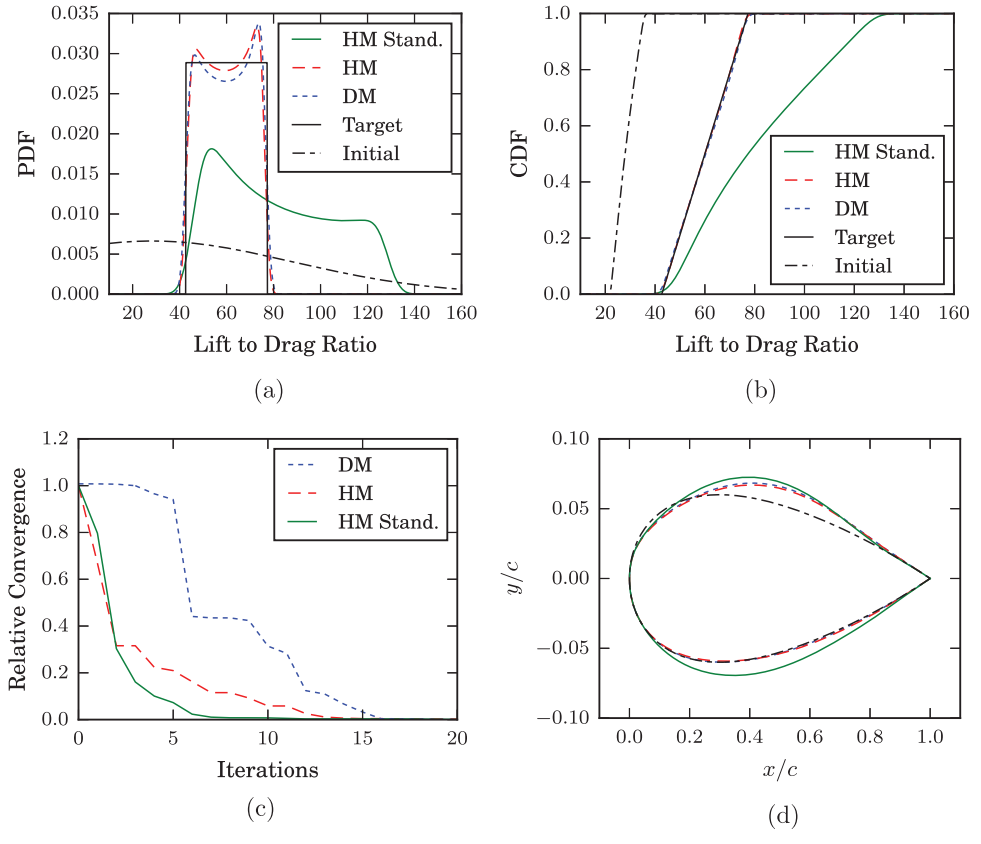

Figure 11. Distributions of *L*/*D*, optimization convergence and aerofoil shapes for the density matching and horsetail matching optimizations towards a uniform target and the horsetail matching optimization under a standard target. (a) PDFs. (b) CDFs. (c) Convergence of optimizations. (d) Optimum aerofoils.

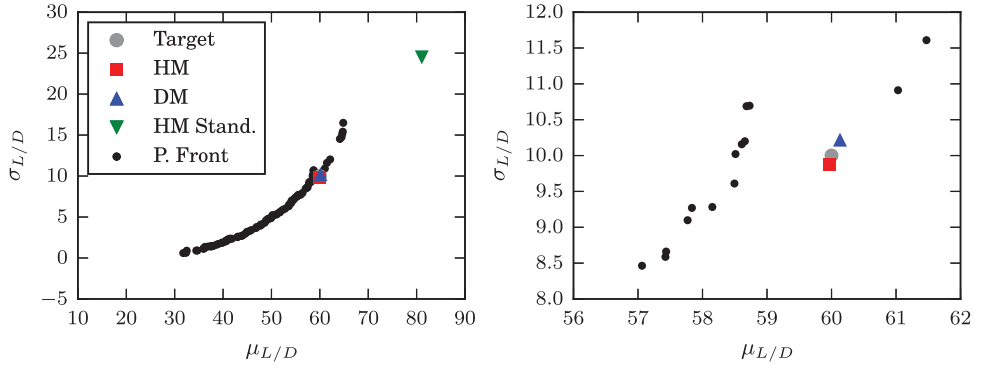

Figure 12. Robust Pareto front for the aerofoil problem (zoomed in on the right around the uniform target point).

polynomials are used to propagate d_{hm} and d_{dm}. The results are plotted in Figure 11, where the PDFs and CDFs of the initial, target and final designs are given, along with the corresponding aerofoil shapes and the optimization histories—where the relative metric $(d - d_{\text{final}})/(d_{\text{initial}} - d_{\text{final}})$ is given for ease of comparison. The values of μ and σ for the final designs are plotted in Figure 12.

When the target is on the Pareto front and so feasible, both HM and DM converge to an aerofoil design with a similar distribution to the target; however, it is observed that the computational efficiency of HM is superior, mainly due to the two-step heuristic used by DM to avoid non-overlapping PDFs.

Furthermore, when HM is used with a standard target, the optimum solution stochastically dominates the designs found using a target on the Pareto front, and it reaches this solution faster: by 7 objective function calls (corresponding to a total of 42 CFD and 42 adjoint runs, since a fifth-order polynomial is used) it has visibly converged. The MO approach to this robust optimization does not get close to this design as it is far beyond the end of the obtained Pareto front in Figure 12, and from Figure 11 the optimum aerofoil under this standard target is significantly different from the other solution, especially on the lower surface. This highlights the ability of the HM approach to find desirable designs from a robust design perspective.

6. Conclusions

Horsetail matching has been proposed as a formulation for optimization under uncertainty that minimizes the difference between a design's cumulative distribution function and a target. By applying the method to both an algebraic test function and a transonic aerofoil shape design problem, it is compared to density matching (the most comparable alternative method) as well as traditional methods of robust optimization. It is shown to give better designs at lower computational cost than density matching whilst also giving the designer the same flexibility. It is also shown to avoid stochastically dominated designs at a comparable computational cost to the weighted sum of statistical moments method.

The proposed implementation of the method relies on response surfaces for computational efficiency, therefore an alternative implementation for situations where the efficacy of response surfaces breaks down (*e.g.* a large number of uncertainties) would be desirable. Additionally, there exists a richer set of options for characterizing uncertainties than probability distributions, so an extension of the method able to optimize under different types of uncertainty would also be desirable. Such an extension is being considered for future work.

Notes

1. https://www.scipy.org/
2. https://pypi.python.org/pypi/ecspy
3. Stanford University Aerospace Design Lab SU2 unstructured solver: http://su2.stanford.edu/

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/L504920/1.

ORCID

L. W. Cook http://orcid.org/0000-0002-0033-1657

References

Beyer, H. G., and B. Sendhoff. 2007. "Robust Optimization—A Comprehensive Survey." *Computer Methods in Applied Mechanics and Engineering* 196 (33–34): 3190–3218. doi:10.1016/j.cma.2007.03.003

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan. 2002. "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II." *IEEE Transactions on Evolutionary Computation* 6 (2): 182–197. doi:10.1109/4235.996017.

Dodson, M., and G. T. Parks. 2009. "Robust Aerodynamic Design Optimization Using Polynomial Chaos." *Journal of Aircraft* 46 (2): 635–646. doi:10.2514/1.39419

Ghisu, T., J. P. Jarrett, and G. T. Parks. 2011. "Robust Design Optimization of Airfoils With Respect to Ice Accretion." *Journal of Aircraft* 48 (1): 287–304. doi:10.2514/1.C031100

Huyse, L., S. Padula, M. Lewis, and W. Li. 2002. "Probabilistic Approach to Free-Form Airfoil Shape Optimization Under Uncertainty." *AIAA Journal* 40 (9): 1764–1772.

Jameson, A., L. Martinelli, and N. A. Pierce. 1998. "Optimum Aerodynamic Design Using the Navier–Stokes Equations." *Theoretical and Computational Fluid Dynamics* 10 (1–4): 213–237. doi:10.1007/s001620050060

Jin, Y., and B. Sendhoff. 2003. "Trade-off between Performance and Robustness: An Evolutionary Multiobjective Approach." In *Proceedings of the Second International Conference on Evolutionary Multi-criterion Optimization (EMO 2003)*, Faro, Portugal, 8–11. April 2003 (pp. 237–251). Lecture Notes in Computer Science 2632. Berlin: Springer-Verlag. doi:10.1007/3-540-36970-8_17

Keane, A. J. 2009. "Comparison of Several Optimization Strategies for Robust Turbine Blade Design." *Journal of Propulsion and Power* 25 (5): 1092–1099. doi:10.2514/1.38673

Keane, A. J., and P. B. Nair. 2005. *Computational approaches for aerospace design: The pursuit of excellence*. New York: Wiley.

Kennedy, M. C., and A. O'Hagan. 2001. "Bayesian Calibration of Computer Models." *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 63 (3): 425–464. doi:10.1111/1467-9868.00294

Le Matre, O. P., O. M. Knio, H. N. Najm, and R. G. Ghanem. 2001. "A Stochastic Projection Method for Fluid Flow. I: Basic Formulation." *Journal of Computational Physics* 173 (2): 481–511. doi:10.1006/jcph.2001.6889

Lee, S. W., and O. J. Kwon. 2006. "Robust Airfoil Shape Optimization Using Design for Six Sigma." *Journal of Aircraft* 43 (3): 843–846. doi:10.2514/1.17359

Lee, D. S., J. Periaux, E. Onate, L. F. Gonzalez, and N. Qin. 2011. "Active Transonic Aerofoil Design Optimization Using Robust Multiobjective Evolutionary Algorithms." *Journal of Aircraft* 48 (3): 1084–1094. doi:10.2514/1.C031237

Levy, H. 2015. *Stochastic dominance—Investment decision making under uncertainty*. Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-21708-6

Marler, R. T., and J. S. Arora. 2010. "The Weighted Sum Method for Multi-Objective Optimization: New Insights." *Structural and Multidisciplinary Optimization* 41 (6): 853–862. doi:10.1007/s00158-009-0460-7

Martin, J. D., and T. W. Simpson. 2005. "Use of Kriging Models to Approximate Deterministic Computer Models." *AIAA Journal* 43 (4): 853–863. doi:10.2514/1.8650

Padulo, M., M. S. Campobasso, and M. D. Guenov. 2011. "Novel Uncertainty Propagation Method for Robust Aerodynamic Design." *AIAA Journal* 49 (3): 530–543. doi:10.2514/1.J050448

Petrone, G., G. Iaccarino., and D. Quagliarella. 2011. "Robustness Criteria in Optimization under Uncertainty." In *Proceedings of the Conference on Evolutionary and Deterministic Methods for Design, Optimization and Control with Applications to Industrial and Societal Problems (EUROGEN 2011)*, edited by C. Poloni, D. Quagliarella, J. Périaux, N. Gauger and K. Giannakoglou, 586–603. Capua: CIRA. doi:10.13140/2.1.1083.2643

Quagliarella, D., G. Petrone, and G. Iaccarino. 2014. "Optimization under Uncertainty Using the Generalized Inverse Distribution Function." In *Modeling, Simulation and Optimization for Science and Technology* (pp. 171–190). Vol. 34 of the Springer series on Computational Methods in Applied Sciences. Dordrecht: Springer Netherlands. doi:10.1007/978-94-017-9054-3 10

Ryan, K. M., M. J. Lewis, and K. H. Yu. 2015. "Comparison of Robust Optimization Methods Applied to Hypersonic Vehicle Design." *Journal of Aircraft* 52 (5): 1510–1523. doi:10.2514/1.C032986

Scott, D. W. 1992. *Multivariate density estimation: Theory, practice and visualization*. Wiley series on Probability and Statistics. New York: Wiley.

Seshadri, P., P. Constantine, G. Iccarino, and G. Parks. 2016. "A Density-Matching Approach for Optimization Under Uncertainty." *Computer Methods in Applied Mechanics and Engineering* 305: 562–578. doi:10.1016/j.cma.2016.03.006

Shapiro, A., D. Dentcheva, and A. Ruszczynski. 2009. *Lectures on stochastic programming: Modeling and theory*. Philadelphia, PA: SIAM.

Yao, W., X. Chen, W. Luo, M. VanTooren, and J. Guo. 2011. "Review of Uncertainty-Based Multidisciplinary Design Optimization Methods for Aerospace Vehicles." *Progress in Aerospace Sciences* 47 (6): 450–479. doi:10.1016/j.paerosci.2011.05.001

Zhang, Y., and S. Hosder. 2013. "Robust Design Optimization Under Mixed Uncertainties With Stochastic Expansions." *Journal of Mechanical Design* 135 (8): Paper No. 081005. doi:10.1115/1.4024230