Operations Research Letters 49 (2021) 106–112

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

# Minimum-error classes for matching parts

Thomas A. Weber

*Chair of Operations, Economics and Strategy, École Polytechnique Fédérale de Lausanne, Station 5, CH-1015 Lausanne, Switzerland*

ARTICLE INFO

*Article history:*
Received 20 August 2020
Received in revised form 2 December 2020
Accepted 3 December 2020
Available online 8 December 2020

*Keywords:*
Error minimization
Matching partitions
Robust optimization
Selective assembly
Stochastic production

ABSTRACT

This paper examines the binning of two types of parts with random characteristics, so that a componentwise monotonic evaluation criterion exhibits a minimum deviation to a given target value over all possible realizations. The optimal matching classes are balanced in the sense that the maximum error needs to be the same over all matching classes. This condition allows for a complete solution of the minimum-error matching-class design problem in closed form.



## 1. Introduction

In assembly situations, components of different types have to fit each other so they can be put together (or "mated"). This is not so much an absolute requirement as it is relative. That is, the *absolute* diameter of a piston that needs to fit into a cylindrical shaft does not matter as much as its value *relative* to the diameter of the shaft. The former can never be larger than the latter, and it should in fact be slightly smaller—thus allowing for sufficient play and lubrication (as long as it is not too small). And since both components come in fluctuating dimensions that are impossible (or too expensive) to exactly control, it is necessary to group different parts of each component type into classes, in such a way that elements of the same "matching class" fit each other at the best possible tolerance. To optimize this tolerance relative to a given target value of a monotonic criterion function is the subject of this paper. Our approach is thereby "robust," in the sense that it aims at minimizing the worst-case matching error over all available matching classes, thus guaranteeing a minimum matching error relative to all possible statistical distributions of the underlying physical parts characteristics.

### 1.1. Literature

The classification of components into classes of similar characteristics (also known as "gauging"), to perform random matching of parts in corresponding classes, has been in practical use for at least a century [4, p. 18]. This "selective assembly" allows the production of systems at a lower tolerance than would be feasible, given the random variability in the dimensions of the constituent components [12]. The latter tolerances might be much larger than needed for the production of high-precision systems. Standard approaches to finding the best breakpoints for the bins have been based *either* on distributional quantiles so as to balance the same number of parts in each class [6] *or* on an equidistant partition of the physical characteristics [16]. On the other hand, more recent algorithms focus on finding the optimal breakpoints for the different bins by minimizing various expected-loss functions [13–15]. Meanwhile, it has also been suggested to avoid the use of bins altogether and proceed by direct matching [5]. By contrast, our approach is distribution-free focusing on the minimization of the maximum error across all matching classes, as evaluated by a criterion function. Instead of convexity of a criterion that usually takes the difference of the characteristics as an argument (corresponding to the piston-shaft example mentioned earlier), we allow for a (smooth) criterion function that is monotonic in the component characteristics. This allows for assemblies that penalize deviations from a system-performance criterion such as a prespecified resonance frequency for an electrical oscillator (see Section 4). The structural restriction to monotonic criterion functions may require a suitable variable transformation; its main benefit is to allow for a closed-form solution of our minimum-error matching-class design problem.

Matching the evaluation of characteristics by a real-valued criterion function $f$ to a given target value as best as possible is similar to approximating the graph of $f$ with a finite two-dimensional grid. In a one-dimensional setting, this is akin to approximating a smooth function by a piecewise constant function [2,8–10]; such piecewise constant approximation applies to a nontrivial class of problems in grouping, spacing, and stratification, connecting the fields of approximation theory and statistics [7]. The key difference for us is that the error is not necessarily computed by a set norm (e.g., $\|\cdot\|_1$, $\|\cdot\|_2$, or $\|\cdot\|_\infty$) but rather

*E-mail address:* thomas.weber@epfl.ch.

https://doi.org/10.1016/j.orl.2020.12.003

by the criterion function itself. That is, the same norm-distance of a vector of characteristics to the graph of the criterion function produces generically different errors, depending on the location in the space of characteristics. For example, when the quotient of two (scalar and positive) characteristics is used as a criterion, then the same absolute variations of the characteristics when they are both small matters more than when they are both large. [To illustrate this last point, assume that the target value of the quotient criterion is 1 and fix the deviations $\Delta_1, \Delta_2$ in the interval $(-1, 1)$; then $(1+\Delta_1)/(1-\Delta_2)$ deviates from the target value by $\varepsilon = |(\Delta_1 + \Delta_2)/(1 - \Delta_2)|$, whereas $(10+\Delta_1)/(10-\Delta_2)$ deviates only by $\hat{\varepsilon} = |(\Delta_1 + \Delta_2)/(10 - \Delta_2)| = \varepsilon(1 - \Delta_2)/(10 - \Delta_2) < \varepsilon$.]

Our developments present the first distribution-free binning results for selective assembly, thus guaranteeing system performance, no matter what the small-sample distributions of the constituent part types might be. It is also worth noting that the results are obtained in closed form and that the use of the optimal matching classes can yield substantial improvements over a naive equidistant binning approach.

### 1.2. Outline

The paper proceeds as follows. Section 2 introduces the criterion function, the matching partition, and their basic properties. In Section 3, we formulate the minimum-error matching-class design problem, construct optimality conditions, and derive a closed-form solution. Section 4 provides practical applications, and Section 5 concludes.

## 2. Preliminaries

Let $f(x, y)$ be a continuously differentiable *criterion function*, such that

$$f_x(x, y) < 0 < f_y(x, y), \tag{1}$$

for all $(x, y) \in \mathcal{X}\times\mathcal{Y}$, where $\mathcal{X} = [a, b]$ and $\mathcal{Y} = [c, d]$ are interval domains with real-valued boundaries $a, b, c, d$ such that $a < b$ and $c < d$, and where partial derivatives are denoted by indices of function symbols (e.g., $f_x = \partial f/\partial x$). The interval domains are partitioned into $N > 1$ subintervals each, by means of the vectors $\theta = (\theta_0, \ldots, \theta_N)$ and $\vartheta = (\vartheta_0, \ldots, \vartheta_N)$ (both elements of $\mathbb{R}^{N+1}$), such that the following *order condition* holds:

$$a = \theta_0 < \theta_1 < \cdots < \theta_N = b \quad \text{and} \quad c = \vartheta_0 < \vartheta_1 < \cdots < \vartheta_N = d. \tag{2}$$

More specifically, the $N$ classes of characteristics are represented by disjoint intervals,

$$\left\{ \begin{array}{lcl} \mathcal{X}_1 & = & [\theta_0, \theta_1] \\ \mathcal{Y}_1 & = & [\vartheta_0, \vartheta_1] \end{array} \right\} \quad \text{and} \quad \left\{ \begin{array}{l} \mathcal{X}_k = (\theta_{k-1}, \theta_k] \\ \mathcal{Y}_k = (\vartheta_{k-1}, \vartheta_k] \end{array} \right\}, \quad k \in \{2, \ldots, N\},$$

so $\mathcal{X}_k \cap \mathcal{X}_j = \mathcal{Y}_k \cap \mathcal{Y}_j = \emptyset$, for all $k, j \in \mathcal{N} = \{1, \ldots, n\}$ with $k \neq j$, and

$$\bigcup_{k=1}^{N} \mathcal{X}_k = \mathcal{X} \quad \text{and} \quad \bigcup_{k=1}^{N} \mathcal{Y}_k = \mathcal{Y}.$$

Fig. 1 illustrates the situation. The range of the criterion function on the space of characteristics $\mathcal{X}\times\mathcal{Y}$ is denoted by $\mathcal{Z} = [m, M]$, where the extrema $m$ and $M$ of $f$ are given by

$$m = f(b, c) = \min_{(x,y)\in\mathcal{X}\times\mathcal{Y}} f(x, y), \tag{3}$$

and

$$M = f(a, d) = \max_{(x,y)\in\mathcal{X}\times\mathcal{Y}} f(x, y), \tag{4}$$

respectively. On the other hand, in any given *joint class* $k \in \mathcal{N}$ *of matched characteristics* (i.e., $\mathcal{X}_k \times \mathcal{Y}_k$), the class-specific extrema of the criterion function are

$$m_k = f(\theta_k, \vartheta_{k-1}) = \min_{(x,y)\in\mathcal{X}_k\times\mathcal{Y}_k} f(x, y), \tag{5}$$

and

$$M_k = f(\theta_{k-1}, \vartheta_k) = \max_{(x,y)\in\mathcal{X}_k\times\mathcal{Y}_k} f(x, y), \tag{6}$$

which implies that $m \le m_k < M_k \le M$. Because $f$ is continuous and all relevant domains are compact, the extrema of $f$ are attained by the Weierstrass theorem; see, e.g., [3, p. 540].

By the monotonicity of the criterion function $f$ in (1), there exist the componentwise inverse functions $g(\cdot|y)$ and $h(\cdot|x)$, such that

$$x = g(f(x, y)|y) \quad \text{and} \quad y = h(f(x, y)|x), \tag{7}$$

for all $(x, y) \in \mathcal{X}\times\mathcal{Y}$. By the inverse function theorem see, e.g., [1], the functions $g, h$ are also continuously differentiable, and they inherit the monotonicity properties of $f$.

**Lemma 1.** *For any $(x, y) \in \mathcal{X} \times \mathcal{Y}$, the functions $g(\cdot|y)$ and $h(\cdot|x)$ are strictly monotonic on $\mathcal{Z}$. More specifically,*

$$g_z(z|y) = \frac{1}{f_x(g(z|y), y)} < 0 < \frac{1}{f_y(x, h(z|x))} = h_z(z|x), \tag{8}$$

*for all $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$.*

**Proof.** By definition of the inverse of $f(x, y)$ with respect to $x$ it is $f(g(z|y), y) = z$, for all $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Differentiating this identity with respect to $z$ yields

$$f_x(g(z|y), y)\, g_z(z|y) \equiv 1. \tag{9}$$

Similarly, $f(x, h(z|x)) \equiv z$ implies

$$f_y(x, h(z|x))\, h_z(z|x) \equiv 1. \tag{10}$$

Taking into account the monotonicities of the criterion function $f$ in Eq. (1), the result in Eq. (8) follows directly from combining Eqs. (9) and (10), which completes our proof. □

**Lemma 2.** *For any $z \in \mathcal{Z}$, the function $h(z|\cdot)$ and the function $g(z|\cdot)$ are strictly increasing on $\mathcal{X}$ and $\mathcal{Y}$, respectively. More specifically,*

$$h_x(z|x) = -\frac{f_x(x, h(z|x))}{f_y(x, h(z|x))} > 0, \quad \textit{and} \quad g_y(z|y) = -\frac{f_y(g(z|y), y)}{f_x(g(z|y), y)} > 0, \tag{11}$$

*for all $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$.*

**Proof.** By Eq. (7), it is $y = h(f(x, y)|x)$ for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$. Differentiating this identity with respect to $x$ yields

$$0 = h_z(z|x) f_x(x, h(z|x)) + h_x(z|x) = \frac{f_x(x, h(z|x))}{f_y(x, h(z|x))} + h_x(z|x),$$

for all $(x, z) \in \mathcal{X} \times \mathcal{Z}$, where we have set $z = f(x, y)$, $y = h(z|x)$, and taken into account Eq. (8) in Lemma 1. This implies the first relation in (11). The proof of the second relation is completely symmetric. □

**Remark 1.** In the case where $f$ violates the monotonicity condition (1) it may be possible to still use our framework after a suitable variable transformation. Indeed, as long as the original criterion function $f$ is monotonic in both arguments (i.e., both $\mathrm{sgn}(f_x)$ and $\mathrm{sgn}(f_y)$ are constant), one can simply "invert" the definition of one of the characteristics (e.g., by using an inverse

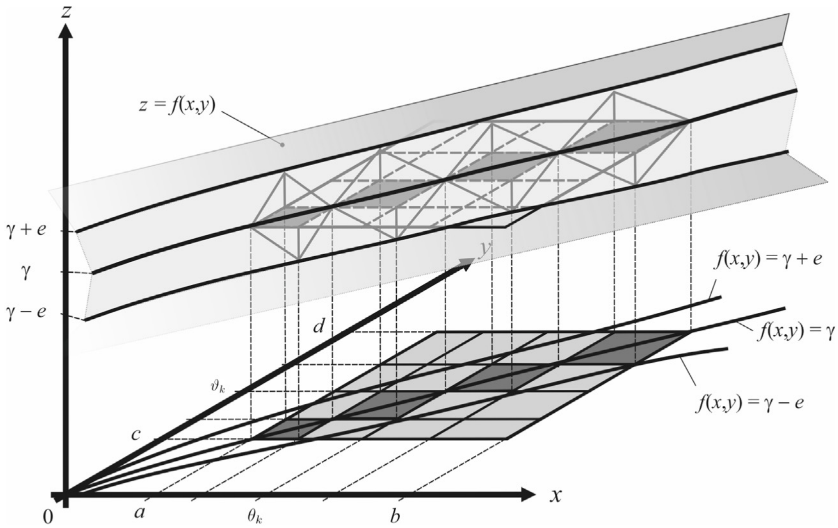



**Fig. 1.** Partition $(\theta, \vartheta)$ of the space of characteristics $\mathcal{X} \times \mathcal{Y} = [a, b] \times [c, d]$ relative to criterion function $f(x, y)$ with target value $\gamma$, resulting in the global matching error $e = e(\theta, \vartheta)$.

in addition or multiplication). For example, if $f(x, y)$ is such that

$$\max\left\{f_x(x, y), f_y(x, y)\right\} < 0,$$

for all $x, y > 0$, then one can introduce the inverse characteristic $\hat{y} = 1/y$ and consider instead the criterion function $\hat{f}(x, \hat{y}) = f(x, 1/\hat{y})$ which does satisfy the monotonicity condition (1); see Example 2 in Section 4, as well as Remark 5.

**Remark 2.** The modified criterion function $\hat{f}(x, y) = f(x, y) - \gamma$ satisfies the monotonicity condition (1) (with $f$ replaced by $\hat{f}$) whenever $f$ does. Hence, without any loss of generality, it is possible to normalize the target value $\gamma$ to zero, simply folding it into the criterion. While this reduces the number of problem parameters, we chose in our presentation to keep $\gamma$ in play, so as not to obscure any physical meaning the criterion function may carry.

## 3. Matching

The *matching error* $e(x, y)$ for a given tuple of characteristics $(x, y) \in \mathcal{X} \times \mathcal{Y}$ is the distance of the corresponding criterion $f(x, y)$ to its prespecified ideal value $\gamma \in (m, M)$:

$$\varphi(x, y) = |f(x, y) - \gamma|.$$

For any given joint class $k \in \mathcal{N}$, the *class-k matching error* is

$$e_k(\theta, \vartheta) = \max_{(x,y)\in\mathcal{X}_k\times\mathcal{Y}_k} \varphi(x, y), \qquad (12)$$

leading to a *global matching error*, defined as the maximal class-specific matching error:

$$e(\theta, \vartheta) = \max\{e_1(\theta, \vartheta), \ldots, e_N(\theta, \vartheta)\}. \qquad (13)$$

Both class-specific and global matching errors critically depend on the design of the set of matching classes $\{\mathcal{X}_k \times \mathcal{Y}_k\}_{k=1}^N$. The latter is defined by $(\theta, \vartheta)$ and is subject to optimization.

### 3.1. Problem

Given a tuple $(a, b, c, d)$ of *criterion-compatible boundary values*, i.e., satisfying

$$c = h(\gamma|a) < h(\gamma|b) = d, \qquad (14)$$

the *minimum-error matching-class design problem*,

$$e^* = \min_{(\theta,\vartheta)} e(\theta, \vartheta), \quad \text{s.t. (2)}, \qquad (*)$$

amounts to finding an optimal joint partition $(\theta^*, \vartheta^*)$, satisfying the order condition (2), so as to minimize the absolute deviation of the criterion function from its ideal value $\gamma$. The strict inequality in the criterion-compatibility condition (14) obtains for any boundary values $a, b$ with $a < b$, since $h(z|\cdot)$ is strictly increasing for all $z \in \mathcal{Z}$, by Lemma 2. Instead of $h(\gamma|b) = d$, one may only require $d \in [h(\gamma|b), h(\gamma + e^*|b)]$ to avoid overspecification of the minimum-error matching-class design problem.

### 3.2. Optimality conditions

The existence of a solution to (*) will be shown by explicit construction. For this, it is important to first note that any solution $(\theta^*, \vartheta^*)$ must necessarily produce equal class-specific matching errors $e_k^* = e_k(\theta^*, \vartheta^*)$ for all $k \in \mathcal{N}$.

**Theorem 1.** *For criterion-compatible boundary values, at any solution $(\theta^*, \vartheta^*)$ of the minimum-error matching-class design problem (*), any class-specific matching error $e_k^*$ equals the optimal global matching error $e^*$, i.e., $e^* = e_1^* = \cdots = e_N^*$.*

**Proof.** For any finite matching-class cardinality $N \geq 2$, it is not possible that the global matching error $e^*$ vanishes. Thus, assume that $e^* > 0$. By Eq. (13) the error in any matching class cannot exceed $e^* = e(\theta^*, \vartheta^*)$. Suppose that there is a matching class $k \in \mathcal{N}$ such that $e_k^* < e^*$. Thus, by Eqs. (5) and (6) it is

$$\gamma - m_k = \gamma - f(\theta_k, \vartheta_{k-1}) < e^* \quad \text{and} \quad M_k - \gamma = f(\theta_{k-1}, \vartheta_k) - \gamma < e^*,$$

or equivalently (using the monotonicity in Lemma 1),

$$\theta_k < g(\gamma - e^*|e_{k-1}^*) \quad \text{and} \quad \vartheta_k < h(\gamma + e^*|\theta_{k-1}).$$

The fact that $\gamma - f(\theta_k, \vartheta_{k-1}) < e^*$ implies that there is a number $s > 0$ such that $\gamma - f(\theta_k, \vartheta_{k-1}) + s = e^*$, so $\theta_k = g(e^* - \gamma - s|\vartheta_{k-1})$ by Eq. (7), whence Eq. (8) in Lemma 1 yields $\theta_k < g(\gamma - e^*|\vartheta_{k-1})$. The second of the two preceding inequalities is derived in an analogous manner. Overall, it is therefore possible to increase the class boundary $(\theta_k, \vartheta_k)$ to $(\hat{\theta}_k, \hat{\vartheta}_k)$ with $\theta_k < \hat{\theta}_k < \theta_{k+1}$ and $\vartheta_k < \hat{\vartheta}_k < \vartheta_{k+1}$, so as to decrease the class-specific matching error $e_k^*$ to $\hat{e}_k^* \in (e_k^*, e^*)$, without having a direct impact on the global matching error. Yet, there is an indirect impact, as the class-boundary shift decreases the error in class $k+1$:

$$\hat{M}_{k+1} = f(\hat{\theta}_k, \vartheta_{k+1}) < f(\theta_k, \vartheta_{k+1}) = M_{k+1},$$

and

$$\hat{m}_{k+1} = f(\theta_{k+1}, \hat{\vartheta}_k) > f(\theta_{k+1}, \vartheta_k) = m_{k+1}.$$

Thus, the reduced error $\hat{e}^*_{k+1} < e^*_{k+1}$ in matching class $k+1$ can be used to decrease the error in matching class $k+2$ and so forth, up to matching class $N$. Similarly, it is possible to propagate a lower error in matching class $k$ also to matching classes with indices smaller than $k$, by downwardly adjusting the class boundaries $\theta_{k-1}$ to $\hat{\theta}_{k-1} \in (\theta_{k-2}, \theta_{k-1})$ and $\vartheta_{k-1}$ to $\hat{\vartheta}_{k-1} \in (\vartheta_{k-2}, \vartheta_{k-1})$. Continuing this procedure all the way to matching class 1 lowers the class-specific matching error from $e^*_j$ to $\hat{e}^*_j < e^*_j$ for all $j < k$. Hence, the resulting optimal global matching error must be smaller than $e^*$, which is a contradiction. Therefore, all of the individual matching-class errors $e^*_k$ must be equal to the optimal global matching error $e^*$. □

### 3.3. Solution

An analytical solution to the minimum-error matching-class design problem is obtained by propagating a uniform class-specific matching error between the boundaries of the parts characteristics. For any given number of matching classes, this yields by Theorem 1 the optimal global matching error $e^*$, with which it is possible to determine the optimal breakpoints, that is, the tuple $(\theta^*, \vartheta^*)$ which defines the optimal matching classes.

**Theorem 2.** *A solution $(\theta^*, \vartheta^*)$ of the minimum-error matching-class design problem (*) is given by*

$$\theta^*_k = g(\gamma - \varepsilon|\vartheta^*_{k-1}), \quad \text{and} \quad \vartheta^*_k = \min\{h(\gamma, b), h(\gamma + \varepsilon|\theta^*_{k-1})\}, \quad (15)$$

*for all $k \in \mathcal{N}$ (given even cardinality; $N = 2n \geq 2$), with $(\theta^*_0, \vartheta^*_0) = (a, c) = (a, h(\gamma|a))$ and $(\theta^*_N, \vartheta^*_N) = (b, d)$. The optimal global matching error $e^*$ is uniquely determined by*

$$b = \bar{G}(a|e^*), \quad (16)$$

*where $\bar{G}(\cdot|\varepsilon) = G^n(\cdot|\varepsilon)$ denotes the n-fold application of $G(\cdot|\varepsilon) = g(\gamma - \varepsilon|h(\gamma + \varepsilon|\cdot))$.*

**Proof.** Assuming that $\varepsilon$ corresponds to the value of the global matching error, $\varepsilon = e(\theta, \vartheta)$, propagating this constant error through the matching classes yields

$$f(\theta_{k-1}, \vartheta_k) = \gamma + \varepsilon, \quad (17)$$

and

$$f(\theta_k, \vartheta_{k-1}) = \gamma - \varepsilon, \quad (18)$$

for all $k \in \mathcal{N}$. The equality of upper and lower error in Eqs. (17) and (18) is without loss of generality. If the errors are different (except at the boundary involving $d$), e.g., equal to $\varepsilon_1$ and $\varepsilon_2$ with $0 < \varepsilon_1 < \varepsilon_2$, then by Theorem 1 it is $\varepsilon = \varepsilon_2$ and the boundary corresponding to $\varepsilon_1$ can be moved to increase the size of class $k$ without changing the global matching error. This flexibility (or "slack") can only help to reduce the global approximation error, e.g., by also moving the boundary corresponding to $\varepsilon_2$ (and successively of the adjacent matching classes) to take advantage of the slack.

Eqs. (17) and (18), together with the compatibility condition (14) and the given values for $a$ and $b$, provide $2N + 3$ conditions to determine $(\theta, \vartheta) \in \mathbb{R}^{2N+2}$, together with the optimal global matching error $\varepsilon$. Note that Eq. (12) is automatically satisfied and does not provide any additional information.

Applying the componentwise inverse functions in Eq. (7) to the optimality conditions (17) and (18) yields

$$\theta_k = g(\gamma - \varepsilon|\vartheta_{k-1}), \quad \text{and} \quad \vartheta_k = h(\gamma + \varepsilon|\theta_{k-1}), \quad (19)$$

for all $k \in \mathcal{N}$. By combining these relations to isolate the propagation by component-type, one obtains

$$\theta_k = G(\theta_{k-2}|\varepsilon) = g(\gamma - \varepsilon|h(\gamma + \varepsilon|\theta_{k-2})), \quad (20)$$

and

$$\vartheta_k = H(\vartheta_{k-2}|\varepsilon) = \min\{h(\gamma|b), h(\gamma + \varepsilon|g(\gamma - \varepsilon|\vartheta_{k-2}))\}, \quad (21)$$

for all $k \in \{2, \dots, N\}$. Fig. 2 shows the different ways of computing the gridpoint-vectors $\theta$ and $\vartheta$, starting with the lower bounds $a$ and $c$ of $\mathcal{X}$ and $\mathcal{Y}$, respectively. Assuming an even number $N = 2n$ of matching classes for an integer $n \geq 1$, one obtains the global matching error $\varepsilon = e^*$ as solution of

$$b = \bar{G}(a|e^*),$$

where $\bar{G}(\cdot|\varepsilon) = G^n(\cdot|\varepsilon)$ denotes the $n$-fold application of $G(\cdot|\varepsilon)$. By the strict monotonicity of $G$ this solution is unique. The missing values of $(\theta^*, \vartheta^*)$ are determined by

$$\theta^*_k = g(\gamma - \varepsilon|\vartheta^*_{k-1}), \quad \text{and} \quad \vartheta^*_k = \min\{h(\gamma, b), h(\gamma + \varepsilon|\theta^*_{k-1})\},$$

for all $k \in \mathcal{N}$, with

$$(\theta^*_0, \vartheta^*_0) = (a, c) = (a, h(\gamma|a)),$$

and

$$(\theta^*_N, \vartheta^*_N) = (b, d),$$

which concludes our proof. □

**Remark 3.** In case $N = 2n + 1$ is odd (for some integer $n \geq 1$), the boundary value $b = g(\gamma - e^*|\bar{H}(c|e^*))$ yields the global matching error $e^*$, where $\bar{H}(\cdot|e^*)$ denotes the $n$-fold application of $H(\cdot|e^*)$, and the remaining calculations proceed similarly. The corresponding formulas are omitted, as in practical settings one can usually restrict attention to a class cardinality $N$ which is even.

**Remark 4.** The optimal global matching error $e^*$ in Theorem 2 is necessarily decreasing in the cardinality $N$ of the matching partition. This is evident, as the global matching error cannot increase by adding points to the grid $(\theta^*, \vartheta^*)$ but must strictly decrease because of the strict monotonicity of $f$ and the balancedness of the class-specific matching errors guaranteed by Theorem 1. Naturally, $e^* \downarrow 0$ for $N \to \infty$, since the class of piecewise constant functions is dense in the class of continuously differentiable functions, so that the criterion function can be approximated arbitrarily closely, provided that the cardinality of the matching partition is sufficiently large. A simple way to guarantee a rate of convergence is by majorizing $e^*$ with an upper bound for the class-$k$ matching error $e_k(\theta, \vartheta)$ in Eq. (12), obtained using a uniform partition of the form $(\theta, \vartheta)$ with $\theta_k = a + (k/N)(b - a)$ and $\vartheta_k = c + (k/N)(d - c)$, for $k \in \{0, \dots, N\}$. Clearly, taking into account Eqs. (5) and (6), for any given $N \geq 2$ it is:

$$\begin{aligned} e_k \leq M_k - m_k &= \big(f(\theta_{k-1}, \vartheta_k) - f(\theta_k, \vartheta_k)\big) + \big(f(\theta_k, \vartheta_k) - f(\theta_k, \vartheta_{k-1})\big), \\ &\leq \big|f(\theta_{k-1}, \vartheta_k) - f(\theta_k, \vartheta_k)\big| + \big|f(\theta_k, \vartheta_k) - f(\theta_k, \vartheta_{k-1})\big|, \\ &\leq K_x(\theta_k - \theta_{k-1}) + K_y(\vartheta_k - \vartheta_{k-1}), \\ &= \frac{K_x(b-a)}{N} + \frac{K_y(d-c)}{N} = \frac{K}{N}, \end{aligned}$$

for $k \in \mathcal{N}$, where we have set $K = K_x(b-a) + K_y(d-c) > 0$, with the Lipschitz constants

$$K_x = \max_{(x,y)\in\mathcal{X}\times\mathcal{Y}} |f_x(x, y)| \quad \text{and} \quad K_y = \max_{(x,y)\in\mathcal{X}\times\mathcal{Y}} |f_y(x, y)|.$$

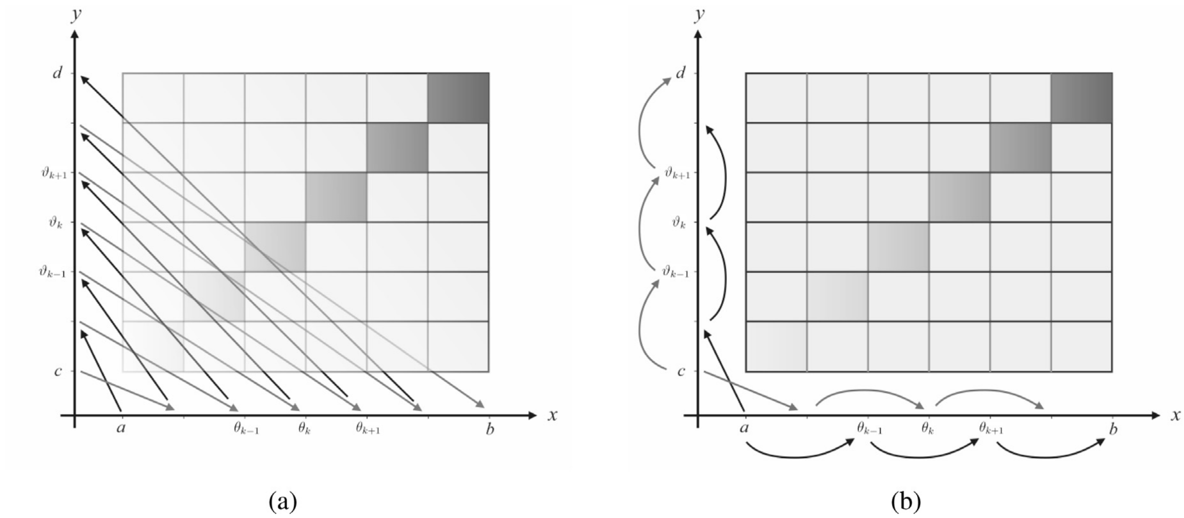



**Fig. 2.** Gridpoint computation (a) using Eq. (19); or, (b) using Eqs. (20) and (21).

Thus, the optimal global matching error $e^*$ is majorized by $K/N$, which therefore guarantees at least the linear convergence rate 1. This convergence rate is achieved exactly when the criterion function $f$ is affine.

## 4. Examples

The following two applications illustrate the solution to the minimum-error matching-class design problem (*), the first being a standard and the second a nonstandard application of selective assembly relative to the extant literature. For the latter application, we compare the optimal solution, in terms of its optimal global matching error, to the naive approach of an equidistant partition in the space of characteristics.

**Example 1** ( *Mechanical Fit*)**.** Let $x$ be the diameter of a piston and $y$ the diameter of a shaft designed to hold it. Then the criterion function $f(x, y) = y - x$ describes the available "play" between the two components, with inverses $g(z|y) = y - z$ and $h(z|x) = x + z$, for all $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Note that $\mathcal{Z} = [m, M]$ with $m = f(b, c) = c - b$ and $M = f(a, d) = d - a$. This implies

$$G(\theta|\varepsilon) = g(\gamma - \varepsilon|\gamma + \varepsilon + \theta) = \theta + 2\varepsilon.$$

Hence, by Eq. (16) it is $b = \bar{G}(a|e^*) = G^n(a|e^*) = a + 2ne^* = a + Ne^*$, which implies that

$$e^* = \frac{b - a}{N},$$

independent of the target value $\gamma > 0$. In addition, by the compatibility condition (14) it is $c = h(\gamma|a) = a + \gamma$. Thus, $H(\vartheta|\varepsilon) = h(\gamma + \varepsilon|\vartheta - \gamma + \varepsilon) = \vartheta + 2\varepsilon$, so $d = H^n(c|e^*) = c + 2ne^* = c + b - a = b + \gamma$. Theorem 2 yields

$$\theta_k^* = a + \frac{k}{N}(b - a) \quad \text{and} \quad \vartheta_k^* = a + \frac{k}{N}(b - a) + \gamma = \theta_k^* + \gamma,$$

for all $k \in \mathcal{N}$. Note that in this particular application the criterion $f(x, y)$ must be nonnegative for all admissible values of $x$ and $y$, which implies the requirement that $e^* < \gamma$, or equivalently that the number $N$ of matching classes exceeds $\lfloor (b - a)/\gamma \rfloor$ at least weakly.

**Example 2** (*Electrical Oscillator*)**.** An electromagnetic coil of inductance $x$ and a capacitor of elastance $y$ (measured as the inverse of its capacitance; see Remark 5) set in a parallel circuit produce a harmonic oscillator of resonance frequency $f(x, y) = (1/(2\pi))\sqrt{y/x}$. The inverses corresponding to this criterion function are $g(z|y) = (y/z^2)/(4\pi^2)$ and $h(z|x) = 4\pi^2 xz^2$, for all $(x, y, z)$ in $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, where $m = f(b, c) = (1/(2\pi))\sqrt{c/b}$ and $M = f(a, d) = (1/(2\pi))\sqrt{d/a}$ are the global extrema as in Eqs. (3) and (4). Therefore,

$$G(\theta|\varepsilon) = g(\gamma - \varepsilon|4\pi^2(\gamma + \varepsilon)^2\theta) = \left(\frac{\gamma + \varepsilon}{\gamma - \varepsilon}\right)^2 \theta.$$

As a result, by Eq. (16) of Theorem 2 it is

$$b = \bar{G}(a|e^*) = \left(\frac{\gamma + e^*}{\gamma - e^*}\right)^N a,$$

which yields the optimal global matching error,

$$e^* = \left(\frac{\left(\frac{b}{a}\right)^{1/N} - 1}{\left(\frac{b}{a}\right)^{1/N} + 1}\right)\gamma < \gamma.$$

In particular, we have that

$$\frac{\gamma + e^*}{\gamma - e^*} = \left(\frac{b}{a}\right)^{1/N},$$

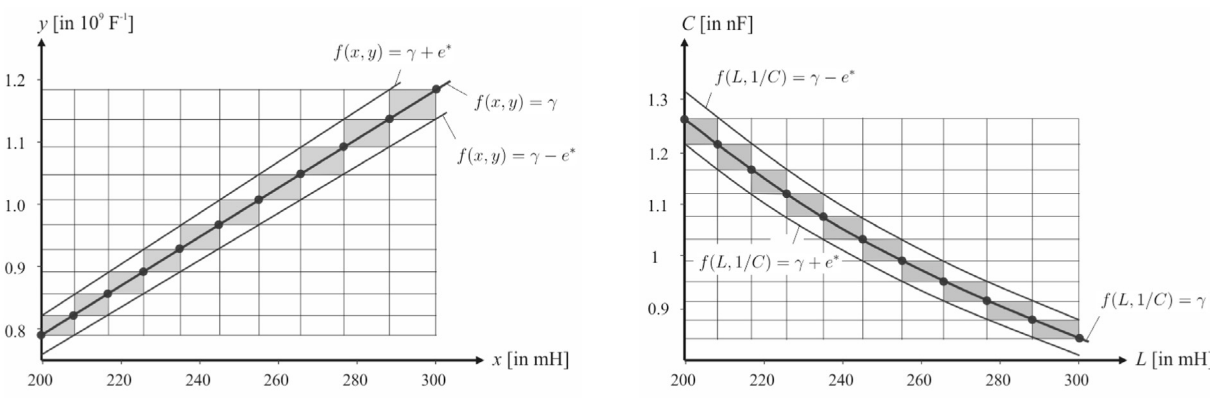
which, for all $k \in \{1, \ldots, n\}$, implies by Eq. (15) of Theorem 2 the optimal partition $(\theta^*, \vartheta^*)$ with the components

$$\vartheta_{2k-1}^* = 4\pi^2 \left(\frac{b}{a}\right)^{2(k-1)/N} (\gamma + e^*)^2 a \quad \text{and} \quad \vartheta_{2k}^* = 4\pi^2\gamma^2 \left(\frac{b}{a}\right)^{2k/N} a,$$

for $\vartheta^*$, as well as

$$\theta_{2k-1}^* = \left(\frac{b}{a}\right)^{2(k-1)/N} \frac{\gamma^2 a}{(\gamma - e^*)^2} \quad \text{and} \quad \theta_{2k}^* = \left(\frac{b}{a}\right)^{2k/N} a,$$

for $\theta^*$. Note that $c = 4\pi^2\gamma^2 a = h(\gamma|a)$ and $d = \vartheta_{2n}^* = 4\pi^2\gamma^2 b = h(\gamma|b)$, so that the compatibility condition (14) is satisfied. There are numerous practical applications requiring a precise tuning of the circuit to a given desired resonance frequency $\gamma$, despite tolerances in the characteristics $(x, y)$ of the components (e.g., in a tuning device for musical instruments: $\gamma = 440$ Hz; in a classical Pultec equalizer there are different switchable circuits, e.g., with $\gamma$ in $\{3, 4, 5, 8, 10, 12, 16\}$ kHz for the "high" frequency spectrum). Fig. 3(a) shows the optimal partition for an LC-oscillator with a target resonance frequency of $\gamma = 10$ kHz and $N = 10$ matching classes, resulting in an optimal global matching error of $e^* \approx 0.2$ kHz (i.e., within about 2% of the target). The underlying relative variability of the components' physical characteristics is




(a) (b)

**Fig. 3.** Optimal matching partition in Example 2 (for $a = 200$ mH, $b = 300$ mH, and $\gamma = 10$ kHz): (a) in the (transformed) variables $(x, y)$; and, (b) in the standard variables $(L, C) = (x, 1/y)$.

±20% which would imply a worst-case error of about 2.25 kHz (i.e., deviating from the target by more than 22%) in the absence of selective assembly with component matching. Let us now compare the optimal solution against the naive equidistant partition of the interval domains with

$$\hat{\theta} = \left(a + (b-a)\left(\frac{k}{N}\right)\right)_{k=0}^{N} \quad \text{and} \quad \hat{\vartheta} = \left(c + (d-c)\left(\frac{k}{N}\right)\right)_{k=0}^{N} = 4\pi^2\gamma^2\hat{\theta}.$$

The matching error corresponding to the equidistant partition is

$$e(\hat{\theta}, \hat{\vartheta}) = e_1(\hat{\theta}, \hat{\vartheta}) = \max\{\gamma - \hat{m}_1, \hat{M}_1 - \gamma\},$$

where

$$\hat{m}_1 = f(\hat{\theta}_1, c) = \gamma\sqrt{\frac{aN}{a(N-1)+b}} \quad \text{and} \quad \hat{M}_1 = f(a, \hat{\vartheta}_1) = \gamma\sqrt{\frac{a(N-1)+b}{aN}}.$$

Hence, since $a(N-1) + b > aN$,

$$e(\hat{\theta}, \hat{\vartheta}) = \max\{\gamma - \hat{m}_1, \hat{M}_1 - \gamma\} = \hat{M}_1 - \gamma = \gamma\left(\sqrt{1 + \frac{(b/a) - 1}{N}} - 1\right).$$

The resulting relative error improvement of the optimal partition over the naive equidistant partition,

$$\rho = 1 - \frac{e^*}{e(\hat{\theta}, \hat{\vartheta})},$$

is increasing in the endpoint ratio $(b/a)$ of the underlying dispersion interval $[a, b]$, and independent of the target value $\gamma$; see Fig. 4. Indeed, with sufficient attribute dispersion (i.e., when $b/a$ grows large), the relative error improvement approaches 100% (i.e., $\lim_{(b/a)\to\infty} \rho = 1$).

**Remark 5.** As alluded to in earlier sections, Example 2 illustrates the occasional need for a change of variables for the criterion function to satisfy the monotonicity condition (1). Indeed, in standard notation the resonance frequency $f$ of an electric "LC-oscillator" is given by $\omega = 2\pi f = 1/\sqrt{LC}$, where $L$ is the inductance of the electromagnetic coil (measured in *henry* (H), or equivalently in $kg\,(m/s)^2/A^2$), $C$ is the capacitance of the capacitor (measured in *farad* (F), or equivalently in $A^2s^4/(kg\,m^2)$), and $\omega$ is the angular frequency (measured in $1/s$). In this formulation, the criterion $f$ is decreasing in both $L$ and $C$. Yet, by choosing the variables $x = L$ and $y = 1/C$, the criterion function $f(x, y)$ satisfies the monotonicity condition (1) and all of the preceding

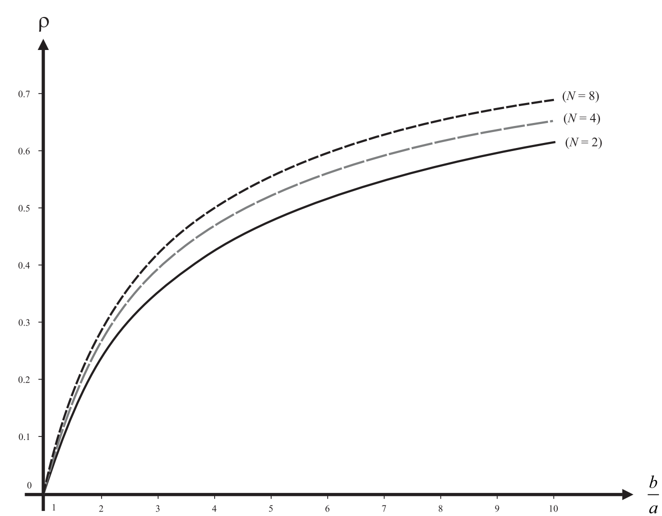



**Fig. 4.** Relative error improvement of optimal partition over equidistant partition in Example 2.

results apply; see Fig. 3(b). In general, such "monotonicity transformations" can be expected to work for physically motivated criterion functions, at least locally; for nonmonotonic criteria it may be useful to reorder the underlying physical characteristics so as to reestablish the monotonicity condition. Instead of the exact criterion function $f(x, y)$ one may also use an additive approximation which is exact at a designated point $(x_0, y_0)$ in the interior of $\mathcal{X} \times \mathcal{Y}$; for details, see [17].

## 5. Conclusion

The presented min–max approach to designing the matching classes for selective assembly is robust, for it does not depend on any specific distributional assumptions. Within corresponding matching classes, the parts are interchangeable and will always mate at a tolerance that guarantees the criterion is within $e^*$ of the target value $\gamma$, where $e^*$ is the optimal global matching error as determined in Theorem 2. Provided that $e^*$ is acceptable, the resulting procedure therefore guarantees zero defects in mating component fits. In a case where $e^*$ is not acceptable across the entire spectrum, the user could increase the cardinality $N$ of the matching partition, which would invariably decrease the optimal global matching error (to any desired specification tolerance)—at the expense of an increased number of sorting bins in the assembly process.

According to Theorem 1 the optimal matching classes, described by Eq. (15), are necessarily "balanced" (or "saturated") in the sense that the class-specific matching errors are all identical (and equal to $e^*$), a property that is common to the optimization of Leontief-type utilities [11], which simplifies the solution to the minimum-error matching-class design problem (*) significantly.

Finally, it is important to realize that the monotonicity requirement (1) is quite natural, as well-adapted parameters usually have monotone effects on the evaluation criterion; see Remarks 1 and 5 for simple variable transformations to obtain the desired monotonicity properties of the criterion function.

## References


[1] J.-P. Aubin, I. Ekeland, Applied Nonlinear Analysis, Wiley, New York, NY, 1984.

[2] R.E. Bellman, On the approximation of curves by line segments using dynamic programming, Commun. ACM 4 (6) (1961) 284.

[3] D.P. Bertsekas, Nonlinear Programming, Athena Scientific, Belmont, MA, 1995.

[4] E. Buckingham, Principles of Interchangeable Manufacturing, Industrial Press, New York, NY, 1921.

[5] C.R. Coullard, A.B. Gamble, P.C. Jones, Matching problems in selective assembly operations, Ann. Oper. Res. 76 (1998) 95–107.

[6] D.J. Desmond, C.A. Setty, Simplification of selective assembly, Int. J. Prod. Res. 1 (3) (1962) 3–18.

[7] R.L. Eubank, Optimal grouping, spacing, stratification, and piecewise constant approximation, SIAM Rev. 30 (3) (1988) 404–420.

[8] J. Fülöp, M. Prill, On the minimax approximation in the class of univariate piecewise constant functions, Oper. Res. Lett. 12 (5) (1992) 307–312.

[9] F. de Gournay, J. Kahn, L. Lebrat, Approximation of curves with piecewise constant or piecewise linear functions, Working Paper, University of Toulouse, Toulouse, France, 2019, arXiv:1909.04582.

[10] H. Konno, T. Kuno, Best piecewise constant approximation of a function of single variable, Oper. Res. Lett. 7 (4) (1988) 205–210.

[11] W.W. Leontief, The Structure of American Economy 1919–1929: An Empirical Analysis of Equilibrium Analysis, Harvard University Press, Cambridge, MA, 1941.

[12] E.M. Mansoor, Selective assembly – its analysis and applications, Int. J. Prod. Res. 1 (1) (1961) 13–24.

[13] S. Matsuura, Optimal partitioning of probability distributions under general convex loss functions in selective assembly, Commun. Stat. Theory Methods 40 (9) (2011) 1545–1560.

[14] S. Matsuura, N. Shinozaki, Optimal binning strategies under squared error loss in selective assembly with a tolerance constraint, Commun. Stat. Theory Methods 39 (4) (2010) 592–605.

[15] D. Mease, V.N. Nair, A. Sudijanto, Selective assembly in manufacturing: Statistical issues and optimal binning strategies, Technometrics 46 (2) (2004) 165–175.

[16] G.A. Pugh, Partitioning for selective assembly, Comput. Ind. Eng. 11 (11–4) (1986) 175–179.

[17] B.H. Strulovici, T.A. Weber, Additive envelopes of continuous functions, Oper. Res. Lett. 38 (3) (2010) 165–168.