This is a repository copy of Gearbox design for uncertain load requirements using active robust optimization.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/96602/

Version: Accepted Version

**Article:**

https://doi.org/10.1080/0305215X.2015.1031659

---

**Reuse**
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Gearbox Design for Uncertain Load Requirements Using Active Robust Optimization

Shaul Salomon*,†, Gideon Avigad‡, Robin C. Purshouse†, and Peter J. Fleming†

†Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, UK
‡Department of Mechanical Engineering, ORT Braude College of Engineering, Karmiel, IL

16 Mar 2015

Abstract
Design and optimization of gear transmissions have been intensively studied, but surprisingly the robustness of the resulting optimal design to uncertain loads has never been considered. Active Robust (AR) optimization is a methodology to design products that attain robustness to uncertain or changing environmental conditions through adaptation. In this study the AR methodology is utilized to optimize the number of transmissions, as well as their gearing ratios, for an uncertain load demand. The problem is formulated as a bi-objective optimization problem where the objectives are to satisfy the load demand in the most energy efficient manner and to minimize production cost. The results show that this approach can find a set of robust designs, revealing a trade-off between energy efficiency and production cost. This can serve as a useful decision-making tool for the gearbox design process, as well as for other applications.

keywords: Gearbox design; adaptive design; multi-objective optimization; robust optimization; active robustness.

1 Introduction
One of today’s engineers greatest challenge is the development of energy efficient products to cope with limited resources. In systems that include a gearbox, careful design of this component can enhance the efficiency of the system. A gearbox is an assembly of gears with different ratios that provides speed and torque conversions from a motor to another device. With the use of a gearbox, a single motor can meet a span of load demands, which are combinations of required speed and torque. There is a unique gearing ratio for every given motor that will result in the least energy consumption for a specific load demand. Usually a geared system operates under a range of possible loads. If optimality with respect to energy consumption is targeted, the gearbox should include an infinite number of gears in order to accommodate all loads within this range. Naturally it is not possible to produce such a gearbox, and anyway, a gearbox with too many gears has more drawbacks than advantages (e.g. dimensions, weight, costs). Therefore gearboxes used in real applications are made of a finite number of gears (typically up to six in the auto industry), where each gear covers a different range of the load demands (e.g. high reduction for high torque and low speed, and vice versa). The gearbox’s gearing ratios should allow for the satisfaction of each possible load by one of the gears in a reasonably efficient

*Corresponding author. Email: s.salomon@sheffield.ac.uk
manner. Therefore, the choice of the gears determine the overall performance of the gearbox. This choice can be supported by an optimization procedure for minimum energy consumption.

Some previous studies on gearbox optimization can be found in the literature. Guzzella and Amstutz (1999) presented a computer aided engineering tool for modelling and optimization of a hybrid vehicle. They showed an example of optimizing the transmission ratios for minimum fuel consumption. The model is deterministic, and the ratios are optimized for a single, arbitrarily chosen, load cycle. Roos, Johansson, and Wikander (2006) suggested an optimization procedure for selecting a motor and gearhead for mechatronic applications to maximize one of the following objectives: peak power, output torque or energy efficiency. This approach is suitable for a single gear system and not for a gearbox with several gears. The choice of the gearhead was conducted according to the worst case of the expected load scenarios. Swantner and Campbell (2012) developed a framework for gearbox optimization that searches among different types of gears (helical, conic, worm, etc.), topologies, materials and sizing parameters. The gearbox was optimized for minimum dimensions, considering a set of functional constraints. Other problem setting for single objective gearbox optimization include minimum variation from a given set of transmission ratios (Mogulapalli, Magrab, and Tsai, 1992), minimum volume or weight (Yokota, Taguchi, and Gen, 1998; Savsani, Rao, and Vakharia, 2010), minimum vibration (Inoue, Townsend, and Coy, 1992) or minimum center distance between input and output shafts (Li, Symmons, and Cockerham, 1996).

Some multi-objective gearbox optimization studies can also be found in the literature. Osyczka (1978) formulated a problem to minimize simultaneously four objective functions: volume of elements, peripheral velocity between gears, width of gearbox, and center distance. Wang (1994) considered center distance, weight, tooth deflection, and gear life as objective functions. Thompson, Gupta, and Shukla (2000) optimized for minimum volume and surface fatigue life. Kurapati and Azarm (2000) optimized a gearbox for minimum volume and minimum stress in the output shaft. Deb, Pratap, and Moitra (2000) designed a compound gear train to achieve a specific gear ratio. The objectives of the gear train design were minimum error between the obtained gear ratio and the required gear ratio and maximum size of any of the gears. Deb and Jain (2003) have optimized an 18-speed, 5-shafts gearbox for two, three and four objectives. Among the objectives were power, volume, center distance and variation from desired output speed. The same optimization problem was used by Deb (2003) to demonstrate how design principles can be extracted by investigating the relations between design variables of the Pareto optimal solutions in the design space. Li et al. (2008) optimized a two-stage gear reducer for minimum dimensions, minimum contact stress and minimum transmission precision errors.

The optimization involved within all studies above was conducted for given reduction ratios, or at least for a given speed-torque scenario or cycle. However, most applications that include a gearbox (such as vehicles) are subjected to a large span of uncertain load requirements, as a result of a variety of possible environmental conditions. The stochastic nature of the required torque and speed must be considered during the design phase. In order to optimize a gearbox for uncertain load requirements, a robust optimization (RO) procedure should be considered. A robust solution is a solution that can maintain good performance over the various scenarios associated with the involved uncertainties. Robustness is usually attained at the price of not achieving peak performance in any specific scenario, and the success of a solution to a robust optimization problem is measured according to a certain criterion such as its mean or worst performance (Paenke, Branke, and Jin, 2006). In this study, a gearbox is optimized for minimum energy consumption where the load demand is uncertain. A robust set of transmission ratios is searched for to maximize the system’s efficiency considering the uncertain load domain.

In many RO problems, in order to ensure robustness, a solution may include some properties that reduce the possible negative influences caused by uncontrolled
parameters’ variations (e.g. thick insulation may reduce fluctuations of an oven internal temperature, caused by changes in the ambient temperature). When this is the case, robustness is passively attained without any action required from the user. A gearbox, however, cannot be optimized for robustness with this approach, since its performance does not solely depend on its preliminary design. The performance is also influenced by the manner in which the gearbox is being operated. A gearbox with a good selection of gearing ratios for a span of load scenarios can be very inefficient if it is not being used properly. For best performance, the proper gear in the set has to be selected for each realization of the uncertain load demand. When cruising on the highway, the best efficiency is achieved with the highest gear (say sixth). A driver that uses the fifth gear for this scenario does not operate the gearbox in an optimal manner. Hence, robustness to the uncertain load demand is actively attained by selecting the proper gear for each load scenario. The selection of the optimal gear for each scenario can be made either manually by a skilled user, or with the use of a controller in the case of an automatic transmission.

The **active robustness** methodology (AR), recently introduced by Salomon et al. (2014), provides the required tools to conduct a robust optimization for a gearbox. AR aims at products that attain robustness to a changing or uncertain environment through adaptation. Such products are termed as adaptive products. The AR methodology assumes that an adaptive product possess some properties that can be modified by its user. These properties allow the product to adapt to environmental changes in order to enhance optimality. The adaptability of a geared system is provided by the user’s ability to change the gear ratios by altering the engaging wheels. This adaptability is taken into account at the evaluation of a candidate solution; it is evaluated according to its best possible performance for each scenario of the uncertain parameters. For the example above, it is assumed that the driver uses the sixth gear while cruising on the highway and second gear when carrying a heavy load up the hill. Since enhanced adaptability usually comes with a price (e.g., a gearbox with more gears would be more expensive), the objectives of an Active Robust Optimization Problem (AROP) are the solution’s best possible performance, evaluated at different scenarios of the uncertainties involved, and its cost.

The problem formulated in this paper is the optimization of a gearbox for a random variate of torque and speed requirements. Both the number of gears and their characteristics are optimized in order to minimize the overall energy consumption and gearbox cost. The solution to the problem is a set of gearboxes with a trade-off between energy efficiency and low cost. The AR optimization approach is demonstrated with a power system of a DC motor and a simple two stage reduction gearbox. The approach can be adopted to other geared systems such as vehicles, motorcycles, wind turbines, industrial and agricultural machinery.

The reminder of the paper is organised as follows: In Section 2 the required background on Robust Optimization and Active Robust Optimization is provided. In Section 3 an example system of a DC motor and a two-stage reduction gearbox is presented, and its model is described. The AROP for optimizing this gearbox is formulated in Section 4, and its solution is presented and analysed in Section 5. Finally, a discussion is given in Section 6 covering the advantages of the presented approach, and how the methods could be further extended to provide efficient support for adaptive complex engineering solutions.

### 2 Background

#### 2.1 Multi-Objective Optimization

Multi-objective optimization problems (MOPs) arise in many real-world applications, where multiple conflicting objectives should be simultaneously optimized. In the absence of prior subjective preference, the solution to such problems is a set of optimal “trade off” solutions rather than a single solution. This set is also called “Pareto
optimal set” or “non-dominated set”. A non-dominated solution is a solution where none of the other solutions is better than it with respect to all of the objective functions.

Mathematically, a MOP can be defined as:

$$\min_{x \in \mathcal{X}} \zeta(x, p) = [f_1(x, p), \ldots, f_m(x, p)],$$  

where \(x\) is an \(n_x\)-dimensional vector of decision variables in some feasible region \(\mathcal{X} \subset \mathbb{R}^{n_x}\), \(p\) is an \(n_p\)-dimensional vector of environmental parameters that are independent of the design variables \(x\) and \(\zeta\) is an \(m\)-dimensional performance vector.

The following define the Pareto optimal set, which is the solution to a MOP:

- A vector \(a = [a_1, \ldots, a_n]\) is said to dominate another vector \(b = [b_1, \ldots, b_n]\) (denoted as \(a \prec b\)) if and only if for all \(i \in 1, \ldots, n\): \(a_i \leq b_i\) and there exists \(i \in 1, \ldots, n\): \(a_i < b_i\).

- A solution \(x \in \mathcal{X}\) is said to be Pareto optimal in \(\mathcal{X}\) if and only if there does not exist \(\hat{x} \in \mathcal{X}: \zeta(\hat{x}, p) \prec \zeta(x, p)\).

The Pareto optimal set (PS) is the set of all Pareto optimal solutions, i.e.,

$$PS = \{x \in \mathcal{X} | \not\exists \hat{x} \in \mathcal{X}: \zeta(\hat{x}, p) \prec \zeta(x, p)\}.$$  

- The Pareto optimal front (PF) is the set of objective vectors corresponding to the solutions in the PS, i.e.,

$$PF = \{\zeta(x, p) | x \in PS\}.$$  

2.2 Robust Optimization

Robust performance design tries to ensure that performance requirements are met and constraints are not violated due to system uncertainties and variations. The uncertainties may be epistemic, resulting from missing information about the system, or aleatory, where the system’s variables inherently change within a range of possible values. Fundamentally, robust optimization is concerned with minimizing the effect of such variations without eliminating the source of the uncertainty or variation (Phadke, 1989).

The performance vector \(\zeta\) in Equation (1) might possess uncertain values due to several sources of uncertainties, which can be categorised according to Beyer and Sendhoff (2007) as follows:

1. **Changing environmental and operating conditions.** In this case, the values of some uncontrollable parameters \(p\) are uncertain. The reasons for uncertainty might be incomplete knowledge concerning these parameters, or expected changes in parameter values during system operation.

2. **Production tolerances and deterioration.** These uncertainties occur when the actual values of design variables differ from their nominal values. The deviation might occur during production (manufacturing tolerances) or during operation (deterioration). Here, the \(x\) variables in Equation (1) are the source of uncertainty.

3. **Uncertainties in the system output.** The actual value of the performance vector \(\zeta\) might differ from its measured or simulated value, due to measurement noise or model inaccuracies, respectively.

When uncertainties are involved within an optimization task, the objective and constraint functions, which define optimality and feasibility, become uncertain too. To assess the uncertain functions, robustness and reliability are considered (Schüeller and Jensen, 2008). Robustness can be seen as having good performance (i.e. objective function values) regardless of the realisation of the uncertain conditions. Reliability is concerned with remaining feasible despite the uncertainties involved.
This study aims at a robust design for changing operating conditions. The related robust optimization problem can be formulated as:

$$\min_{x \in \mathcal{X}} F(x, \mathbf{P}),$$

where \( x \) is an \( n_x \)-dimensional vector of decision variables in some feasible region \( \mathcal{X} \subset \mathbb{R}^{n_x} \), \( \mathbf{P} \) is an \( n_p \)-dimensional vector random variate, of uncertain environmental parameters that are independent of the design variables \( x \), and \( F(x, \mathbf{P}) \) is a distribution of objective function values that correspond to the variate of the uncertain parameters \( \mathbf{P} \).

In a robust optimization scheme, the random objective function is evaluated according to a robustness criterion, denoted by an indicator \( \phi [F] \). Three classes of criteria are presented in the following.

Worst-case optimization, also known as robust optimization in the operational research literature (Bertsimas, Brown, and Caramanis, 2011) or minmax optimization (Alicino and Vasile, 2014), considers the worst performance of a candidate solution over the entire range of uncertainties. The worst-case indicator for a minimization problem can be written as:

$$\phi_w [F(x, \mathbf{P})] := \max_{\mathbf{p} \in \mathbf{P}} F(x, \mathbf{p}).$$

The robust optimization problem in Equation (2) then reads:

$$\min_{x \in \mathcal{X}} \max_{\mathbf{p} \in \mathbf{P}} F(x, \mathbf{p}).$$

To address the tendency of this approach to produce over-conservative solutions, Jiang, Wang, and Guan (2012) suggested a method for controlling the conservatism of the search by reducing the size of the uncertainty interval with a tuneable parameter. Branke and Rosenbusch (2008) suggested an evolutionary algorithm for worst-case optimization that simultaneously searches for the robust solution and the worst-case scenario by co-evolving the population of scenarios alongside the candidate solutions.

Aggregation methods use an integral measure that amalgamates the possible values of the uncertain objective function. The most common aggregated indicators are the expected value of the objective function or its variance – see the review by Beyer and Sendhoff (2007). When the distribution of the uncertain parameters can be described by the probability density function \( \rho(p) \), the mean value criterion can be computed by:

$$\phi_m [F(x, \mathbf{P})] := \int_{\mathbf{p} \in \mathbf{P}} f(x, \mathbf{p}) \rho(\mathbf{p}) d\mathbf{p},$$

where \( f(x, \mathbf{p}) \) is a deterministic model for the objective function. Commonly in real world problems, Equation (5) cannot be analytically derived for the following reasons: i) the distribution of the uncertain parameters is not known and needs to be derived from empirical data, and/or ii) it is not feasible to analytically propagate the uncertainties to form the uncertain objective function. Monte-Carlo sampling can then be used for these cases to represent the random variate \( \mathbf{P} \) as a sampled set \( \mathbf{P} \) of size \( k \).

The mean value then becomes:

$$\phi_m [F(x, \mathbf{P})] := \frac{1}{k} \sum_{i=1}^{k} f(x, \mathbf{p}_i),$$

where \( \mathbf{p}_i \) is the \( i \)th sample in \( \mathbf{P} \). Kang, Lee, and Lee (2012) have considered the expected value with a partial mean of costs to solve a process design robust optimization problem. Kumar et al. (2008) have used Bayesian Monte-Carlo sampling to construct a sampled representation for the performance of candidate compressor blades. They considered both the mean value and the variance as a multi-objective optimization
problem, and used a multi-objective evolutionary algorithm to search for robust solutions. An alternative formulation is to aggregate the mean and variance into a single objective function (e.g. Lee and Park, 2001).

Beyer and Sendhoff (2007) suggested a criterion that uses the probability distribution of the objective function directly as a robustness measure. This is done by setting a performance goal, and maximising the probability for achieving this goal, i.e. for the function value to be better than a desired threshold. Considering a performance threshold \( q \), a threshold probability indicator can be defined as:

\[
\phi_{tp}[F(x, p)] := \Pr(F(x, p) < q).
\] (7)

Reliability-based design aims at minimizing the risk of failure during the product expected lifecycle (Schüller and Jensen, 2008). In the context of design optimization, it can be seen as minimizing the risk of violating the problem’s constraints. The criteria mentioned above for robustness can also be used to assess reliability by applying them to the constraint functions. A conservative worst-case approach was used by several authors (e.g. Avigad and Coello, 2010; Albert et al., 2011). The “six-sigma” methodology (see Brady and Allen, 2006) suggests a goal of 3.4 defects per million products, which sets a threshold probability for reliability.

2.3 Active Robustness Optimization Methodology

The AR methodology (Salomon et al., 2014), is a special case of robust optimization, where the product has some adjustable properties that can be modified by the user after the optimized design has been realized. These adjustable variables allow the product to adapt to variations in the uncontrolled parameters, so it can actively suppress their negative effect. The methodology makes a distinction between three types of variables: design variables, denoted as \( x \), adjustable variables, denoted as \( y \) and uncontrollable stochastic parameters \( P \). A single realized vector of uncertain parameters from the random variate \( P \) is denoted as \( p \).

In a conventional robust optimization problem, each realization \( p \) is associated with a corresponding objective function value \( f(x, p) \), and a solution \( x \) is associated with a distribution of objective function values that correspond to the variate of the uncertain parameters \( P \). This distribution is denoted as \( F(x, P) \). In active robust optimization, for every realization of the uncertain environment, the performance also depends on the value of the adjustable variables \( y \), i.e., \( f \equiv f(x, y, p) \). Since the adjustable variables’ values can be selected after \( p \) is realized, the solution can improve its performance by adapting its adjustable variables to the new conditions. In order to evaluate the solution’s performance according to the robust optimization methodology, it is conceivable that the \( y \) vector that yields the best performance for each realization of the uncertainties will be selected. This can be expressed as the optimal configuration \( y^* \):

\[
y^* = \arg\min_{y \in y(x)} f(x, y, p),
\] (8)

where \( y(x) \) is the solution’s domain of adjustable variables, also termed as the solution’s adaptability.

Considering the entire environmental uncertainty, a one-to-one mapping between the scenarios in \( P \) and the optimal configurations in \( y(x) \) can be defined as:

\[
Y^* = \arg\min_{y \in y(x)} F(x, y, P).
\] (9)

Assuming a solution will always adapt to its optimal configuration, its performance can be described by the following variate:

\[
F(x, P) \equiv F(x, Y^*, P).
\] (10)
An *Active Robust Optimization Problem* (AROP) optimizes a performance indicator $\phi$ for the variate $F(x, Y^*, P)$. It is denoted as $\phi(x, Y^*, P)$. Since enhanced performance usually increases the costs of the product, the aim of an AROP is to find solutions that are both robust and inexpensive. Therefore the AROP is a multi-objective problem that simultaneously optimizes the performance indicator $\phi$ and the solution’s cost.

The cost function for the gearbox that is used in this study only depends on the gearbox’s preliminary design, i.e., the number of gears and their specifications. Therefore it is not affected by the uncertain load demand and has a deterministic value. The general definition of an AROP considers a stochastic distribution of the cost function, but in this case it is denoted as $c(x)$.

Following the above, the *Active Robust Optimization Problem* is formulated:

$$\min_{x \in X} \zeta(x, P) = [\phi(x, Y^*, P), c(x)]$$  \hfill (11)

$$\text{where } Y^* = \arg\min_{y \in Y} F(x, y, P).$$  \hfill (12)

It is a multi-stage problem. In order to compute the objective function $\phi$ in Equation (11), the problem in Equation (12) has to be solved for every solution $x$ with the entire environment universe $P$. In a typical implementation the environmental uncertainty $P$ is sampled using Monte Carlo methods. This sample, $\bar{P}$, leads to sample-based representations of $Y^*$ and $F$ – denoted $\bar{Y}^*$ and $\bar{F}$ respectively. This leads to an estimated performance vector $\hat{\zeta}$.

### 3 Motor and Gear System

The problem at hand is the optimization of a gearbox for a span of torque-speed scenarios. A DC motor of type Maxon A-max 32 is to convey a torque $\tau_L$ at speed $\omega_L$. In order to do so, it is coupled with a gearbox as shown in Figure 1. The motor’s output shaft (white) rotates at speed $\omega_m$ and transmits a torque $\tau_m$. It is firmly connected to a cogwheel (black) that is constantly coupled to the layshaft. The layshaft consists of a shaft and $N$ gears (gray), rotating together as a single piece. $N$ gears (white) are also attached to the load shaft (black) with bearings, so they are free to rotate around it. The gears are constantly coupled to the layshaft and rotate at different speeds, depending on the gearing ratio. A collar (not shown in the figure) is connected, through splines, to the load shaft and spins with it. It can slide along the shaft to engage any of the gears, by fitting teeth called “dog teeth” into holes on the sides of the gears. In that manner the power is transferred to the load through a certain gear, with the desired reduction ratio.
The aim of this study is to optimize the gearbox to achieve good performance over a variety of possible load scenarios. Several objectives might be considered: monetary costs, energy efficiency for different loads and the transient behaviour of the gearbox (e.g. energy consumption during speed transitions and time required to change the system’s speed). A problem formulation that considers all of the aforementioned objectives is very complex and challenging. However, in order to demonstrate the features and concerns of the active robustness approach, at this stage it is sufficient to focus on a more restricted formulation of the gearbox optimization problem. Therefore, only the steady-state behaviour of the gearbox is addressed in this study.

The number of gears in the gearbox, $N$, and the number of teeth in each $i^{th}$ gear, $z_i$, are to be optimized. The objectives considered are minimum energy consumption and minimum manufacturing cost of the gearbox. The system is evaluated at steady-state, i.e., operating at the torque-speed scenarios. The power required for each scenario is considered, while the objective is to find the set of gears that will require the minimum average invested power over all scenarios. For every scenario, the gearbox is evaluated by the the smallest possible value of input power. This value is achieved by transmitting the power through the most suitable gear in the box.

### 3.1 Model Formulation

In this section, the model for the motor and gearbox system is presented according to Krishnan (2001), and the required performance measures are derived.

The motor armature current can be described by applying Kirchoff’s voltage law over the armature circuit:

$$ V = L\dot{I} + rI + k_v\omega_m, $$

where $V$ is the input voltage, $L$ is the coil inductance, $I$ is the armature current, $r$ is the armature resistance and $k_v$ is the velocity constant. The ordinary differential equation describing the motor’s angular velocity as related to the torques acting on the motor’s output shaft is:

$$ J_m\dot{\omega}_m = k_tI - b_m\omega_m - \tau, $$

where $J_m$ is the rotor’s inertia, $k_t$ is the torque constant and $b_m$ is the motor’s damping coefficient associated with the mechanical rotation. Since this study only deals with the gearbox’s performance at steady-state, the derivatives of $I$ and $\omega_m$ are considered as zero.

There are two speed reductions between the motor and the load. The first is from the motor shaft to the layshaft. This reduction ratio, denoted as $n_1$, is $z_l/z_m$, where $z_m$ is the number of teeth in the motor shaft cogwheel and $z_l$ is the number of teeth in the layshaft cogwheel. The second reduction, denoted as $n_2$, is from the layshaft to the load shaft. Each gear on the load shaft rotates at a different speed according to its gearing ratio $n_{2,i} = z_{g,i}/z_{l,i}$, where $z_{g,i}$ is the number of teeth of the $i^{th}$ gear’s load shaft cogwheel and $z_{l,i}$ is the number of teeth of its matching layshaft wheel. $n_2$ depends on the selected gear, and it can be one of the values \{${n_{2,1}, \ldots, n_{2,N}}$\}. The total reduction ratio from the motor to the load is $n = n_1 \times n_2$, and the load speed $\omega = \omega_m/n$. The motor and load shafts are coaxial, and the modules for all cogwheels are identical. Therefore, the total number of teeth $N_t$ for each gearing couple is identical:

$$ N_t = z_l + z_m = z_{g,i} + z_{l,i}, \quad \forall i \in 1, \ldots, N. $$

At steady-state, Equation (14) can be reflected to the load shaft as follows:

$$ 0 = nk_tI - (b_g + n^2b_m)\omega - \tau, $$

where $\tau$ is the load’s torque and $b_g$ is the gear’s damping coefficient with respect to the load’s speed.
If $\omega$ from Equation (16) is known, the armature current can be derived:

$$I = \frac{(b_g + n^2 b_m)\omega + \tau}{nk_t}. \quad (17)$$

Once the current is known, and after neglecting $\dot{I}$, the required voltage can be derived from Equation (13):

$$V = rI + nk_v \omega. \quad (18)$$

The invested electrical power is:

$$s = VI. \quad (19)$$

It is conceivable that manufacturing costs depend on the number of wheels in the gearbox, their size, and overheads. A function of this type is suggested for this generic problem to demonstrate how the various costs can be quantified:

$$c = \alpha N^\beta + \gamma \sum_{i=1}^{N} (z_{l,i}^2 + z_{g,i}^2) + \delta, \quad (20)$$

where $\alpha$, $\beta$, $\gamma$ and $\delta$ are constants. The first term considers the number of gears. It takes into account their influence on the costs of components such as the housing and shafts. The second term relates to the cogwheels material costs, which are proportional to the square of the number of teeth in each wheel. The third represents the overheads. In practice, other cost functions could be used.

4 Problem Definition

The gearbox optimization problem, formulated as an AROP, is the search for the number of gears $N$ and the number of teeth in each gear $z_{g,i}$ that minimize the production cost $c$ and the power input $s$. According to the AR methodology, introduced in Section 2, the variables are sorted into three vectors:

- $x$ is a vector with the variables that define the gearbox, namely the number of gears and their teeth number. These variables can be selected before the gearbox is produced, but cannot be altered by the user during its life cycle. The variables in $x$ are the problem’s design variables.

- $y$ is a vector with the adjustable variables. It includes the variables that can be adjusted by the gearbox’s user: the selected gear $i$ and the supplied voltage $V$. The decisions how to adjust these variables are made according to the load’s demand, and can be supported by an optimization procedure. For example, a high reduction ratio will be chosen for low speed, and a low ratio for high speeds, while the voltage is adjusted to maintain the desired velocity for the given torque.

- $p$ is a vector with all the environmental parameters that affect performance and are independent of the design variables. Some of the parameters in this problem are considered as deterministic, but some possess uncertain values. The uncertainty for $\omega$ and $\tau$ is aleatory, since they inherently vary within a range of possible load scenarios. The random variates of $\omega$ and $\tau$ are denoted as $\Omega$ and $\mathcal{T}$, respectively. Some values of the motor parameters are given tolerances by the supplier. The terminal resistance $r$ has a tolerance of 5% and the motor resistance $b_m$ has a tolerance of 10%. Additionally, the gearbox damping $b_g$ can be only estimated, and therefore it is treated as an epistemic uncertainty. The random variates of $r$, $b_m$ and $b_g$ are denoted as $R$, $B_m$ and $B_g$, respectively. The resulting variate of $p$ is denoted as $P$. 


A certain load scenario might have more than one feasible configuration. When the gearbox (represented by x) is evaluated for each scenario, the optimal configuration (the one that requires the least input power) is considered. This configuration is denoted as \( y^* \), and it consists of the optimal transmission \( i \) and input voltage \( V \) for the given scenario. The variate of optimal configurations that correspond to the variate \( P \) is termed as \( Y^* \). Since the input power varies according to the uncertain parameters (this can be denoted as \( S(x, Y^*, P) \)), a robust optimization criterion is used in order to assess its value. The mean value is a reasonable candidate for this purpose, as it captures the efficiency of the gearbox when it operates over the entire range of expected load scenarios. It is denoted as \( \pi(x, Y^*, P) \).

Following the above, the AROP is formulated:

\[
\min_{x \in \mathcal{R}} \zeta(x, P) = \{ \pi(x, Y^*, P), c(x) \},
\]

\[ Y^* = \arg\min_{y \in \mathcal{Y}(x)} S(y, P), \]

subject to:

\[ I \leq I_{nom}, \]

\[ z_{g,i} + z_{t,i} = N_t, \quad \forall i = 1, \ldots, N, \]

where:

\[ x = [N, z_{g,1}, \ldots, z_{g,i}, \ldots, z_{g,N}], \]

\[ y = [i, V], \]

\[ P = [\Omega, T, R, B_m, B_g, k_r, k_t, I_{nom}, v_1, N_t, \alpha, \beta, \gamma, \delta]. \]

The constraints are evaluated according to Equations (17) and (18), and the objectives according to Equations (19) and (20). \( I_{nom} \), the nominal current, is the highest continuous current that does not damage the motor. It is significantly smaller than the motor’s stall current.

By operating with maximum input power (i.e. with maximum voltage and current), for each velocity \( \omega \) there is a single transmission ratio \( n \) that would allow the maximum torque, denoted as \( \tau_{max}(\omega) \). This torque can be derived from Equations (16) and (18) by replacing \( I \) with \( I_{nom} \) and \( V \) with \( V_{max} \):

\[
\tau_{max}(\omega) = \max_{n \in \mathcal{Y}} nk_t I_{nom} - (b_g + n^2 b_m)\omega, \\
\text{subject to:} \quad rI_{nom} + nk_c \omega = V_{max},
\]

where \( \mathcal{Y} \subseteq \mathcal{R} \) is the range of possible reduction ratios for this problem. Since a gearbox in the above AROP consists of a finite number of gears, it cannot operate at \( \tau_{max} \) for most of the velocities. In order to obtain feasible solutions with five gears or less, the domain of possible scenarios in this example is assumed to be in the range of \( 0 \leq \tau(\omega) \leq 0.55\tau_{max}(\omega) \). The effects of this assumption on the obtained solutions’ robustness are further discussed in Section 5.2.

Some information on the probability of load scenarios is usually known in a typical gearbox design (e.g. drive cycle information in vehicle design). In this generic example this kind of information is not available, and therefore a uniform distribution is assumed. The other uncertainties are treated in a similar manner: A uniform distribution is assumed for \( R \) and \( B_m \), since the tolerance information provided by the manufacturer only specifies the boundaries for the actual property values, but does not specify their distribution. The epistemic uncertainty regarding \( b_g \) also results in a uniform distribution of \( B_g \) within an estimated interval.

Monte-Carlo sampling is used to represent the uncertain parameter domain \( P \). A set \( \hat{P} \) of size \( k \), is constructed by a random sampling of \( P \) with an even probability. In this example, \( \hat{P} \) consists of \( k = 1,000 \) scenarios. The choice of sample size is further investigated in Section 5.2. Figure 2 depicts the domain of load scenarios \( \Omega \) and \( T \), together with their samples in \( \hat{P} \) and the curve \( \tau_{max}(\omega) \).
The parameter values and the limits of search variables and uncertainties are presented in Table 1. The values and tolerances for the motor parameters were taken from the online catalog of Maxon (2014). Note that the upper limit of the selected gear $i$ is $N$, meaning that different gearboxes possess different domains of adjustable variables. This notion is manifested in the problem definition as $y \in Y(x)$.

5 Simulation Results

The discrete search space consists of 1,099,252 different combinations of gears (2–5 gears, 43 possibilities for the number of teeth in each gear: $C_{43}^2 + C_{43}^3 + C_{43}^4 + C_{43}^5$). The constraints and objective functions depend on the number of teeth $z$, so they only have to be evaluated 43 times for each of the 1000 sampled scenarios. As a result, it is feasible to find the true Pareto optimal solutions to the above problem by evaluating all of the solutions. The entire simulation took less than one minute, using standard desktop computing equipment.

A feasible solution is a gearbox that has at least one gear that does not violate the constraints for each of the scenarios (i.e., $I \leq I_{nom}$ and $V \leq V_{max}$). Figure 3 depicts the objective space of the AROP. There are 194,861 feasible solutions (marked with gray dots), and the 103 non-dominated solutions are marked with black dots. It is noticed that the solutions are grouped into three clusters with a different price range for each number of gears. The three clusters correspond to $N \in \{3, 4, 5\}$, where fewer gears are related with a lower cost. None of the solutions with $N = 2$ is feasible.

5.1 A Comparison Between an Optimal Solution and a Non-Optimal Solution

For a better understanding of the results obtained by the AR approach, two candidate solutions are examined: one that belongs to the Pareto optimal front, and another that does not. Consider a scenario where lowest energy consumption is desired for a given budget limitation. For the sake of this example, a budget limit of $243 per unit is arbitrarily chosen. The gearbox with the best performance for that cost is marked in Figure 3 as Solution A. This solution consists of five gears with $z_{2,A} = \{59, 49, 41, 34, 24\}$ and corresponding transmission ratios $n_A = \{9.02, 5.07, 3.38, 2.37, 1.38\}$. 

---

**Figure 2:** The possible domain of torque-speed scenarios, and a representative set randomly sampled with an even probability.
Another solution with the same cost is marked in Figure 3 as Solution B. The gears of this solution are $z_{2,B} = \{57, 40, 34, 33, 21\}$, and its corresponding transmission ratios are $n_B = \{7.96, 3.21, 2.37, 2.25, 1.14\}$.

Figure 4 depicts the set of optimal transmission ratio at every sampled scenario for both solutions. Each transmission is marked in the figure with a different marker. This set is in fact the set $Y^\star$ from Equation (21), that correspond to the sampled set of load scenarios $P$, in Figure 2. It is observed that the reduction ratios of Solution A almost form a geometrical series, where each consecutive ratio is divided by 1.6 approximately. The resulting $Y^\star(x_A)$ is such that all gears are optimal for a similar number of load scenarios. Solution B on the other hand has two gears with very similar ratios. It can be seen in Figure 4(b) that the third and the fourth gears are barely used. These gears do not contribute much to the gearbox’s efficiency, but significantly increase its cost. As can be seen in Figure 3, there are gearboxes with four gears that achieve the same or better efficiency as Solution B.

Figure 5 depicts the lowest power consumption for every sampled scenario, $s(x, Y^\star, P)$. This consumption is achieved by using the optimal gear for each load scenario (those in Figure 4). It can be seen that Solution A uses less energy at many load scenarios compared to Solution B. This is depicted by the darker shades of many of the scenarios in Figure 5(b). In order to assess the robustness, the mean input power $\pi(x, Y^\star, P)$ is used as the robustness criterion for this AROP. It is calculated by averaging the values of all points in Figure 5. The results are $\pi(x_A, Y^\star, P) = 5.23\, \text{W}$ and $\pi(x_B, Y^\star, P) = 5.47\, \text{W}$. Considering both solutions cost the same, this confirms Solution A’s superiority over Solution B. Given a budget limitation of $243$, Solution A should be preferred by the decision maker.

### 5.2 Robustness of the Obtained Solutions

In this section the sensitivity of the AROP’s solution to several factors of the problem formulation is examined. Two aspects are considered with respect to different robustness metrics and parameter settings: i) the optimality of a specific solution, and ii) the difference between two alternative solutions. For this purpose, three tests
Figure 3: The objectives values of all feasible solutions to the problem in Equation (21) and Pareto front.

Figure 4: Optimal transmission ratio for every sampled scenario.

Figure 5: Lowest power consumption for every sampled scenario.
are performed. The first relates to the robustness of the solutions to epistemic uncertainty, namely the unknown range of load scenarios. The second test relates to the robustness of the solutions to a different robustness metric. The third test examines the sensitivity to the sampling size.

Sensitivity to Epistemic Uncertainty

The domain of load scenarios is bounded between $0 \leq \tau \leq 0.55 \cdot \tau_{\text{max}}(\omega)$. The choice of $55\%$ is arbitrary, and it reflects an assumption made to quantify an epistemic uncertainty about the load. Similarly, the upper bound for $T$ could be a function $a \cdot \tau_{\text{max}}(\omega)$ with a different value of $a$. The Pareto frontiers for several values of $a$ can be seen in Figure 6. For $a = 40\%$, the Pareto set consists of solutions with two, three, four and five gears, whereas for $a = 70\%$ the only feasible solutions are those with five gears. For percentiles larger than 70\% there are no feasible solutions within the search domain.

To examine the effect of the choice of maximum torque percentile on the problem’s solution, the three solutions from Figure 3 are plotted for every percentile in Figure 6. Solutions $A$ and $C$, who belong to the Pareto set for $a = 55\%$, are also Pareto optimal for all other values of $a$ smaller than 65\%. Solution $B$ remains dominated by both Solutions $A$ and $C$. When very high performance is required (i.e. maximum torque percentiles of 65\% or higher), both Solution $A$ and Solution $C$ become infeasible.

It can be concluded that the mean value, as a robustness metric, is not sensitive to the maximum torque percentile. On the other hand, the reliability of the solutions, i.e. their probability to remain feasible, is sensitive to the presence of extreme loading scenarios.

Sensitivity to Preferences

The threshold probability metric is used to examine the sensitivity of the solutions to different performance goals. It is defined for the above AROP as the probability for a solution to consume less energy than a predefined threshold:

$$\phi_{tp} = \Pr(S < q),$$

where $q$ is the performance goal. The aim is to maximize $\phi_{tp}$.

Figure 7 depicts the results of the AROP described in Section 4, when $\phi_{tp}$ is considered as the robustness metric, and the goal performance is set to $q = 5W$. The same three solutions from Figure 3 are also shown here. Solution $A$, whose mean power consumption is the best for its price, is not optimal any more when
Figure 7: The objectives values of all feasible solutions and Pareto front, for maximizing the threshold probability $\phi_{tp} = \Pr(S < 11W)$.

Figure 8: Pareto frontiers for different thresholds $q$.

the probability of especially poor performance is considered. Solution $A$ manages to satisfy the goal for 98.6% of the sampled scenarios, while another solution with the same price satisfies 99% of the scenarios. It is up to the decision maker to determine whether the difference between 98.6% and 99% is significant or not.

Solutions $B$ and $C$ are consistent with the other robustness metric. Solution $B$ is far from optimal, and Solution $C$ is still Pareto optimal. This consistency is maintained for different values of the threshold $q$, as can be seen in Figure 8. Figure 8 also demonstrates that setting an over ambitious target results in a smaller probability of fulfilment by any solution.

Sensitivity to the Sampled Representation of Uncertainties

The random variates are represented in this study with a sampled set, using Monte-Carlo methods. The following experiment was conducted in order to verify that 1,000 samples are enough to provide a reliable evaluation of the solutions’ statistics: Solutions $A$ and $C$ were evaluated for their mean power consumption over 5,000 different sampled sets with sizes varying from $k = 100$ to $k = 100,000$. Figure 9(a) depicts the metric values of the solutions for every sample size. It is evident from the
results that a large number of samples is required for the sampling error to converge. For both solutions, the standard deviation is 15%, 6%, 2% and 0.5% of the mean value, for sample sizes of $k = 100$, $k = 1,000$, $k = 10,000$, and $k = 100,000$, respectively. If an accurate estimate is required for the actual power consumption, a large sample size must be used (i.e. larger than $k = 1,000$ that was used in this study).

On the other hand, a comparison between two candidate solutions can be based on a much smaller sampled set. Although the values of $\pi(x, Y^*, \overline{P})$ may change considerably between two consequent realisations of $\overline{P}$, a similar change will occur for all candidate solutions. This can be seen in Figure 9(a) where the “funnels” of the two solutions seem like exact replicas with a constant bias. The difference in performance between the two solutions $\Delta \pi(\overline{P})$ is defined:

$$\Delta \pi(\overline{P}) = \pi(x_C, Y^*, \overline{P}) - \pi(x_A, Y^*, \overline{P})$$

Figure 9(b) depicts the value of $\Delta \pi(\overline{P})$ for every evaluated sampled set. It can be seen that $\Delta \pi$ converges to 200mW. For a sampling size of $k = 100$, the standard deviation of $\Delta \pi$ is 25mW, which is only 12% of the actual difference. This means that it can be argued with confidence that Solution A has better performance than Solution C, based on a sample size of $k = 100$.

Based on the results from this experiment, it can be concluded that the solution to the AROP (i.e. the set of Pareto optimal solutions) is not sensitive to the sample size. The Pareto front shown in Figure 3 might be shifted along the $\pi$ axes for different sampled representations of the uncertainties, but the same (or very similar) solutions would always be identified.

6 Conclusions

This study is the first of its kind to extend gearbox design optimization to consider the realities of uncertain load demand. It demonstrates how the stochastic nature of the uncertain load demand can be fully catered for during the optimization process using an Active Robustness approach. A set of optimal solutions with a trade-off between cost and efficiency was identified, and the advantages of a gearbox from this set over a non-optimal one were shown. The robustness of the obtained Pareto optimal solutions to several aspects of the problem formulation was verified.

The approach takes account of – and exploits – user influence on system performance, but presently assumes that the user is able to operate the gearbox in an optimal manner to achieve best performance. Of course, this assumption can only be fully validated if a skilled user or a well tuned controller activates the gearbox. This raises an important issue of how to train this user or controller to achieve best performance, which is identified as a priority for further research.
Computational complexity is a concern for the AR approach demonstrated in this study. This case study used very simple analytic functions to evaluate each candidate solution. Therefore the real solution to the AROP could be found almost instantly. When applying this method to real world applications, every function evaluation might require extensive computational effort. In this case, efficient optimization algorithms would be required, and the uncertainties may need to be described by methods other than Monte-Carlo sampling. However, the large amount of function evaluations required to solve a typical AROP is a feasible prospect for real industrial problems. Since the problem is solved off-line, before the product goes to manufacturing, super-computing facilities are likely to be available, and a reasonable time-scale for solving the problem might be days or even a few weeks.

Adaptability is the solution's ability to react to changes in its environment by adjusting itself to a configuration that improves its performance. In this study the gearbox's adaptability was evaluated by only considering its performance at each of the sampled load scenarios, i.e., at steady-state. However, the Active Robustness methodology, presented by Salomon et al. (2014), considers adaptability in a wider sense. In addition to its performance at steady-state, the solution's transient behaviour during adaptation to environmental changes is also considered. For the problem presented in this paper, an environmental change is a change in demand from one load scenario to another. Although the optimal configurations can be found for both scenarios, the gearing ratios and input voltages applied while changing between these configurations may have a substantial impact on the solution's performance. This notion was deliberately not considered in the current study in order to focus on basic aspects of the approach. An important extension to this work would be to examine the transient behaviour when evaluating a candidate solution. Additional objectives such as acceleration and energy consumption during adaptation can be examined by doing so. The Optimal Adaptation method (Salomon et al., 2013) can be used to search for adaptation trajectories that optimize these objectives.

The transient extension to the problem formulation requires extra considerations with respect to computational complexity. The two main reasons for this are: (a) A change between any two scenarios can be made by infinite possible gear sequences and voltage trajectories. This requires a search for the optimal trajectory in order to be consistent with the AR approach. This kind of search is usually computationally expensive. (b) Each adaptation between two scenarios has to be examined. The number of possible adaptations between $k$ scenarios are $k(k-1)$. For the sampled set of 1,000 scenarios used in this study, there will be 999,000 adaptations to examine for each solution, implying a requirement to solve 999,000 optimization problems. As a part of future research, special attention should be given to model simplification and finding reliable ways to reduce the number of evaluated adaptations, e.g. by using efficient algorithms and sampling methods.

This initial study of gearbox optimization is based on a simple DC motor and gearbox. This is advantageous in focusing the presentation on the Active Robustness approach rather than, for example, constraint handling, and enables the objective functions to be calculated analytically. Additional applications for the AR methodology will be demonstrated in future publications, including more complex real-world geared systems.

Acknowledgement

This research was supported by a Marie Curie International Research Staff Exchange Scheme Fellowship within the Seventh European Community Framework Programme. The first author acknowledges support from Ort Braude College of Engineering, Israel and the support of the Anglo-Israel Association. The first and second authors acknowledge the hospitality and support of the Mechanical and Material Engineering Department at the University of Western Ontario, Canada.
References


