In this paper we arbitrarily choose the approach suggested by Audze and Englelais\textsuperscript{14} to generate an experimental design. In \( k \) dimensions with \( N \) samples an \( N^k \) grid is generated. The points are then placed on this grid such that no two points lie along the same grid line and the quantity
\[
\sum_{i=1}^{N} \sum_{j=1}^{N} \frac{1}{d_{ij}^2}
\]
is minimized. Here \( d_{ij} \) refers to the distance between points \( i \) and \( j \).
For example, with \( k = 2 \) and \( N = 10 \) we obtain
\[
\begin{bmatrix}
8 & 10 & 4 & 6 & 2 & 3 & 9 & 5 & 7 & 1 \\
1 & 7 & 10 & 6 & 8 & 5 & 4 & 2 & 9 & 3
\end{bmatrix}
\]

Figure 1

**Approximation Methods**

The approximation method considered in this paper is a kriging model. The relationship between inputs \( x \) and observations (responses) \( y \) is expressed as
\[
y = \mu + \varepsilon(x)
\]
where \( \mu \), a constant, is the mean of the input responses and \( \varepsilon(x) \) is a Gaussian random function with zero mean and variance \( \sigma^2 \).

We do not assume that the \( \varepsilon \) are constant as in regression, but that these errors are correlated. The correlation between two points being related to some distance measure between corresponding points. The distance measure used here is
\[
d(x^{(i)}, x^{(j)}) = \sum_{k=1}^{k} \theta_k (x_k^{(i)} - x_k^{(j)})^2
\]
where \( \theta_k \) are hyperparameters yet to be determined.

The correlation between points \( x^{(i)} \) and \( x^{(j)} \) is defined by
\[
R(x^{(i)}, x^{(j)}) = \exp[-d(x^{(i)}, x^{(j)})].
\]
When we wish to sample at a new point \( x \), we form a vector of correlations between the new point and the previously sampled points
\[
r(x) = R(x, x^{(i)}).
\]
If \( x \) is close to \( x^{(i)} \), then these points are correlated and the predicted response will be strongly influenced by the response at \( x^{(i)} \).

Whereas if the points are far apart, the correlation is small and the predicted response will only be weakly influenced by the response at \( x^{(i)} \).

The prediction itself is given by
\[
y(x) = \mu + r^TR^{-1}(y - 1\mu)
\]
where the mean \( \mu \) is defined by
\[
\mu = \frac{1^TR^{-1}y}{1^TR^{-1}1}.
\]
The hyperparameters \( \theta_k \) are obtained by maximizing the likelihood of the sample. This is defined as
\[
\frac{1}{N \frac{N}{2} \left( \frac{2}{(2\pi)^2 \sigma^2 R^2} \right)} \exp \left[ -\frac{(y-1\mu)^T R^{-1}(y-1\mu)}{2\sigma^2} \right]
\]

\[
J^2(x) = \sigma^2 \left[ 1 + \left( 1 - I^T R^{-1} r \right)^2 \right]
\]

as this gives us a measure of the accuracy of the prediction at \( x \). The general strategy is shown in figure 2.

Figure 2

where the variance \( \sigma^2 \) is given by

\[
\sigma^2 = \frac{(y - 1\mu)^T R^{-1} (y - 1\mu)}{N}
\]

Another useful quantity is the mean squared error of prediction

Figure 3
Model Fusion

We now turn our attention to the use of approximation techniques in combining a large number of low fidelity analyses together with a small number of high fidelity analyses for constructing approximations that are inexpensive and accurate.

The general strategy is to consider the low fidelity model $f_a$ together with only selective calls to $f_e$. The inputs for calls to $f_e$ are again chosen using a design of experiments approach.

The low fidelity model provides some rough global information as to the response of the high fidelity model. The selective calls to $f_e$ are then made to build corrections to the low fidelity model. In this paper, the knowledge-based approach of Leary et al. is used for constructing approximations, the cheap model being included inside the approximation. In this way, accurate information about $f_e$ at a limited number of points is combined with the approximate global information provided by $f_a$, thus increasing the accuracy of the approximation. The approach is demonstrated on two simple finite element problems. The strategy in this case is shown in figure 3.

Example 1

As an initial example we consider the structure shown in figure 4, subjected to a uniformly distributed load $p_0$. The length $L$ is here taken as 1 metre. We wish to find the optimal values of $x_1$ and $x_2$ (width and height of cross-section) such that the volume $V$ is minimized subject to

$$\sigma_{\text{max}} < 100000 \text{ N/m}^2$$

and

$$0.05 \leq x_i \leq 0.1, \quad i=1,2.$$  

We consider a high fidelity model consisting of 100 finite elements and a low fidelity model consisting of just four elements. In this example the objective (volume) is cheap to compute whereas the stress (which forms one of the constraints) requires a finite element analysis. It is the maximum stress that we model here.

Clearly the "expensive model" here is not so complex that optimization is impractical. We therefore observe that the minimum occurs at

$$x_1 = 0.05m$$
$$x_2 = 0.083379m$$

the minimum volume here is

$$V = 0.010065\text{m}^3.$$  

This provides us with a benchmark for assessing the accuracy of the results obtained using approximation methods. Of course in a practical situation, this information will not be available.

The design space, objective and constraints for the low fidelity model are shown in figure 5, figure 6 shows the same information for the high fidelity model. Note here that the low fidelity model gives an inaccurate result due to the fact that it under-predicts stress.
poor. As the model is updated using successive approximations, predictions become more accurate and we begin to converge to the correct minimum. The algorithm terminates when either

1) the maximum allowable number of calls to $f_c$ is reached or

2) we have convergence.

Now consider the knowledge-based model, this approximates the high fidelity model by including the low fidelity model in the approximation. Results of optimization in this case are given in table 2.

<table>
<thead>
<tr>
<th>Point</th>
<th>$x_1 (\times 10^{-2})$</th>
<th>$x_2 (\times 10^{-2})$</th>
<th>$V$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DoE 1</td>
<td>5.0</td>
<td>5.0</td>
<td>0.6036</td>
</tr>
<tr>
<td>DoE 2</td>
<td>7.5</td>
<td>10.0</td>
<td>1.8107</td>
</tr>
<tr>
<td>DoE 3</td>
<td>10.0</td>
<td>7.5</td>
<td>1.8107</td>
</tr>
<tr>
<td>min 1</td>
<td>5.0</td>
<td>8.702</td>
<td>1.0504</td>
</tr>
<tr>
<td>min 2</td>
<td>5.0</td>
<td>8.395</td>
<td>1.0134</td>
</tr>
<tr>
<td>min 3</td>
<td>5.0</td>
<td>8.361</td>
<td>1.0093</td>
</tr>
<tr>
<td>min 4</td>
<td>5.0</td>
<td>8.352</td>
<td>1.0082</td>
</tr>
<tr>
<td>min 5</td>
<td>5.0</td>
<td>8.347</td>
<td>1.0076</td>
</tr>
<tr>
<td>min 6</td>
<td>5.0</td>
<td>8.345</td>
<td>1.0073</td>
</tr>
<tr>
<td>min 7</td>
<td>5.0</td>
<td>8.344</td>
<td>1.0072</td>
</tr>
<tr>
<td>min 8</td>
<td>5.0</td>
<td>8.343</td>
<td>1.0071</td>
</tr>
</tbody>
</table>

Table 2

Now observe the increased accuracy when also using the low fidelity model. The initial approximation is extremely accurate and we immediately focus near the optimum. Successive approximations again converge to the correct point. Convergence is quicker than in the previous case due to the extra information provided by the low fidelity model.

The longer-term goal is to produce inexpensive and accurate approximations to real world engineering systems in order to aid the use of optimization at the preliminary design stage. We will be interested in modeling spoked structures such as the one shown in figure 7. In this case it is the tail bearing housing of an aero-engine.

![Figure 7](image)
**Example 2**

This example considers the use of approximation methods applied to an idealization of a spoked structure with simple boundary conditions. The initial configuration together with the loading and constraints are shown in figure 8.

![Figure 8](image)

The structure is supported on the outer ring as shown, here the displacements are constrained to be zero. A pressure is exerted on one section of the inner ring as shown. These have been chosen for illustration of the method.

Such structures are typically designed subject to stress and stiffness constraints. We here consider stiffness. The goal is to reduce the weight of the structure by varying the thickness of the rings and the spokes such that the maximum displacement remains below a certain value. The design variables are the inner thickness

\[ 0.1 \leq x_1 \leq 1.2, \]

the spoke thickness

\[ 0.1 \leq x_2 \leq 3.0 \]

and the outer thickness

\[ 0.5 \leq x_3 \leq 8.0. \]

The radius of the outer ring is 200 units and the radius of the inner ring is 100 units. The depth of the rings is also taken as 100 units. We constrain the design to have a maximum displacement no greater than 0.1 units.

Admittedly, the present dimensions and boundary conditions are somewhat unrealistic. However, the goal here is simply to demonstrate the use of the approach on such structures.

Two models are considered, a low fidelity finite element model, shown in figure 9, consisting of 30 shell elements and 792 degrees of freedom and a high fidelity model, shown in figure 10, consisting of 1134 elements and 21816 degrees of freedom.

![Figure 9](image)

An approximation to both \( f_a \) and \( f_e \) is considered here. Computing \( f_a \) is cheap and the approximation is here based on over 1000 evaluations. The evaluation of \( f_e \) is more expensive and the
approximation here is based on only limited information.

The finite element analysis itself is performed using the ABAQUS package. Note it would be better to use $f_e$ directly (rather than an approximation) and this will be done once a successful interface is completed. Meanwhile a very accurate approximation to $f_e$ is considered.

Initially, 10 space filling points are constructed using the Audze and Engell algorithm. An approximation is first constructed using $f_e$ alone. The convergence is shown in table 3.

As before, we also look at the knowledge-based approach and convergence here is shown in table 4.

Again using the low fidelity model increases the rate of convergence allowing results to be obtained with fewer calls to the high fidelity model. The displacement of the low fidelity model at the best current point was 0.0975 units again indicating some misalignment between the two models. This is this misalignment that is successfully modeled.

Also shown in these tables are the actual displacements at the current best points, the constraint of 0.1 units is active and our approximate displacement gradually becomes more and more accurate (approaches 0.1).

<table>
<thead>
<tr>
<th>Point</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$V$</th>
<th>disp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>min 1</td>
<td>1.2</td>
<td>0.1</td>
<td>3.8097</td>
<td>564097</td>
<td>0.1256</td>
</tr>
<tr>
<td>min 2</td>
<td>1.1793</td>
<td>1.54613</td>
<td>3.6562</td>
<td>623000</td>
<td>0.0998</td>
</tr>
<tr>
<td>min 3</td>
<td>1.1796</td>
<td>1.2968</td>
<td>3.6446</td>
<td>613333</td>
<td>0.1013</td>
</tr>
<tr>
<td>min 4</td>
<td>1.2</td>
<td>2.3545</td>
<td>2.3872</td>
<td>331420</td>
<td>0.1141</td>
</tr>
<tr>
<td>min 5</td>
<td>1.1926</td>
<td>2.1380</td>
<td>2.7516</td>
<td>549151</td>
<td>0.1072</td>
</tr>
<tr>
<td>min 6</td>
<td>1.1997</td>
<td>2.3881</td>
<td>2.7968</td>
<td>568566</td>
<td>0.0999</td>
</tr>
<tr>
<td>min 7</td>
<td>1.1917</td>
<td>2.3867</td>
<td>2.7930</td>
<td>568679</td>
<td>0.1008</td>
</tr>
<tr>
<td>min 8</td>
<td>1.2</td>
<td>2.3817</td>
<td>2.7750</td>
<td>568770</td>
<td>0.1002</td>
</tr>
<tr>
<td>min 9</td>
<td>1.2</td>
<td>2.3833</td>
<td>2.7746</td>
<td>568220</td>
<td>0.1005</td>
</tr>
</tbody>
</table>

Table 3

<table>
<thead>
<tr>
<th>Point</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$V$</th>
<th>disp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>min 1</td>
<td>1.2</td>
<td>2.3205</td>
<td>2.7933</td>
<td>569920</td>
<td>0.10033</td>
</tr>
<tr>
<td>min 2</td>
<td>1.2</td>
<td>2.3500</td>
<td>2.7933</td>
<td>568116</td>
<td>0.0997</td>
</tr>
<tr>
<td>min 3</td>
<td>1.2</td>
<td>2.3493</td>
<td>2.7934</td>
<td>568021</td>
<td>0.09999</td>
</tr>
<tr>
<td>min 4</td>
<td>1.2</td>
<td>2.3482</td>
<td>2.7904</td>
<td>568024</td>
<td>0.1000</td>
</tr>
</tbody>
</table>

Table 4

To compare the results globally we also evaluate $f_e$ at many points ($m$) in the design space to assess the overall accuracy of our initial approximations. The correlation coefficient, defined by $r^2$ where

$$ r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} $$

$$ S_{xx} = \frac{\sum_{i=1}^{m} y_i^{(0)} y_i^{(0)} - \left(\sum_{i=1}^{m} y_i^{(0)}\right)^2}{m} $$

$$ S_{yy} = \frac{\sum_{i=1}^{m} y_i^{(0)} y_i^{(0)} - \left(\sum_{i=1}^{m} y_i^{(0)}\right)^2}{m} $$

$$ S_{xy} = \frac{\sum_{i=1}^{m} y_i^{(0)} y_i^{(0)} - \left(\sum_{i=1}^{m} y_i^{(0)} \sum_{i=1}^{m} y_i^{(0)}\right)}{m} $$

provides an indication of the overall accuracy. This coefficient varies between zero and one. The closer to one, the better the prediction.

For the model based on $f_e$ alone, we obtain

$$ r^2 = 3 \times 10^{-4} $$

indicating a very poor global prediction.

Using the knowledge-based approach gives a correlation coefficient of

$$ r^2 = 0.9998 $$

indicating a much better global accuracy. Displacement correlation is shown in figure 11.
Discussion and conclusions

The use of variable fidelity models in overcoming the computational burden associated with the optimization of complex high fidelity models has been described. This approach is seen to increase the accuracy of approximations. Once an approximation is constructed, an optimization is performed and we resample at the optimum of the approximation, then build a new approximation. In this way, as our optimization proceeds, our approximation becomes more accurate in points of interest.

The efficiency of the approach in the examples described was due to the fact that there was good correlation throughout the entire design space between the models of varying fidelity. Thus the knowledge-based approach quickly allowed us to home in to the correct point. This approach suffices if the two models have roughly similar sorts of behavior; it is reasonable to expect this as the two models are representing the same physical system.

If there is less global correlation between the two models, a better strategy may be to combine the knowledge-based methods with an expected improvement criterion, see Jones et al. for a description of expected improvement.

Additionally, it might be the case that derivatives of the response are available cheaply from a finite element analysis. If such information is available it can be incorporated into a kriging model. The application of these methods using derivatives within the context of multi-fidelity approximations is another area which we plan to investigate.

Finally, the application of these approaches to complex structures with realistic geometries and boundary conditions will be addressed.

Acknowledgements

We would like to thank Rolls-Royce Plc for the financial support the first author receives as part of the University Technology Partnership (UTP), the European Union for funding of the second author as part of the MMFSC project. We also thank Alison McMillan and Steve Wiseall of Rolls-Royce Plc for their industrial feedback, Cecil Armstrong and Willard McCune of Queens University Belfast for providing the models used in example 2.

References

13. M. D. Mackay, R. J. Beckman and W. J. Conover, A comparison of three methods for selecting values of input variables in the
15. ABAQUS version 6.1, Hibbit, Karlsson & Sorensen inc.