Proceedings Paper:

Binois, M. orcid.org/0000-0002-7225-1680, Branke, J., Fieldsend, J. orcid.org/0000-0002-0683-2583 et al. (1 more author) (2025) Decoupled design of experiments for expensive multi-objective problems. In: Festa, P., Ferone, D., Pastore, T. and Pisacane, O., (eds.) Learning and Intelligent Optimization: 18th International Conference, LION 18, Ischia Island, Italy, June 9–13, 2024, Revised Selected Papers. 18th International Conference, LION 18, 09-13 Jun 2024, Ischia Island, Italy. Lecture Notes in Computer Science, LNCS 14990. Springer Nature Switzerland, pp. 37-50. ISBN: 9783031756221. ISSN: 0302-9743. EISSN: 1611-3349.

https://doi.org/10.1007/978-3-031-75623-8_4

Decoupled Design of Experiments for Expensive Multi-objective Problems

Mickaël Binois[1][0000−0002−7225−1680],
Jürgen Branke[2][0000−0001−5880−1925],
Jonathan Fieldsend[3][0000−0002−0683−2583], and
Robin C. Purshouse[4][0000−0001−5880−1925]

[1] Inria Centre at Université Côte d'Azur `mickael.binois@inria.fr`
[2] University of Warwick `juergen.branke@wbs.ac.uk`
[3] University of Exeter `j.e.fieldsend@exeter.ac.uk`
[4] University of Sheffield `r.purshouse@sheffield.ac.uk`

Abstract. In this paper we look at the experimental design for multi-objective problems, where the objectives can be evaluated independently (decoupled) and thus it may make sense to evaluate different solutions for each objective if the objectives have different evaluation costs and/or different landscape characteristics. We propose to iteratively add design points in a way that minimises the total integrated mean squared prediction error assuming a Gaussian process response surface model, and show that allowing decoupled evaluations can lead to significantly better Pareto front estimations than a coupled design of experiments if the evaluation costs of the objectives are different. We also find that our approach of minimising mean squared prediction error yields significantly better results than standard Latin Hypercube designs even if the evaluation costs and landscape characteristics of the objectives are the same.

Keywords: Expensive optimisation · Varying costs · Multi-objective experimental design.

1 Introduction

Fundamental to the performance of surrogate-based optimisation frameworks is the need to construct an initial model based on a carefully selected set of initial designs and any prior system knowledge. This is both in the case of Bayesian optimisation (BO), which uses and iteratively updates model(s) mapping decision vectors to predicted performance criteria values, and for evolutionary computation approaches which involve surrogates. The selection and construction of initial designs, which are often treated separately to the decision vectors queried during the subsequent optimisation process, are usually referred to as the *design of experiments* (or DoE for short). This is because these decision vectors are selected to—in some fashion—be maximally informative on the global underlying process, rather than being biased towards particular regions.

Without any prior information regarding the properties of the objective function(s) such DoE for model fitting are commonly based around *space filling* sequences such as Latin hypercube sampling (LHS) [15] or Sobol sequences [16], as purely random sampling tends to naturally result in clusters, which do not serve model fitting well, particularly when the budget for sampling is tight.

Where there are multiple criteria being modelled, this leads to an interesting and under-explored question: *should one evaluate all initial designs fully, or instead selectively evaluate a subset of objectives per design, allowing a greater number of locations to be partially evaluated when building the model(s)?* A few works have looked at *decoupling* objective evaluations during the search process—particularly where there are different costs associated with each objective, but this can also be advantageous where there is a difference in the complexity of the functions being modelled (e.g. one being smooth and slowly changing, the other being rugged and fast changing). As such, this appears to be a promising direction for further investigation and research, as even small improvements in such areas can effectively lead to large savings for expensive optimisation problems. A possible drawback, on the other hand, is that the Pareto dominance cannot be determined for sure on decoupled designs (only with some confidence depending on the accuracy of the surrogate model prediction).

The remainder of the paper is set out as follows. In Section 2 we introduce existing work and methods relating to decoupled and cost-aware multi-objective optimisation, and highlight how our work relates to these. Section 3 presents results of the proposed approach with different problem configurations, and highlights the circumstances where there appears to be a significant benefit to decoupling the DoE locations. In Section 4 we discuss the results, and highlight future research directions.

2 Related Work

In single objective optimisation, there are various papers taking into account the cost of evaluating a solution where this cost depends on the solution evaluated. The de facto standard is to divide the acquisition function value by the corresponding cost value (e.g., [17]). [13] demonstrates that this is not always a good choice, and proposes an alternative mechanism. In particular, they propose an initial space-filling design that takes cost into account, by iteratively and greedily adding points that are inexpensive to evaluate but have a large distance from points already chosen. During optimisation, their algorithm reduces the emphasis on cost, starting with the standard division by cost, then slowly changing into a standard acquisition function optimisation without considering cost. In [12], the authors propose a non-myopic approach to BO with cost considerations.

A small number of existing works have considered decoupled and/or cost-aware multi-objective optimisation—some of which have considered these factors during the initial DoE phase. Below we discuss the most relevant approaches. A wider survey on the topic of objectives with different costs can be found in [2].

In [1], a user can define a cost ranking of the decision variables, e.g. in the case that decision variables represent the amount of an ingredient, and the ingredients have different costs. The acquisition function then favours solutions with small values in particular for the expensive variables, and this preference is reduced over the course of the run, eventually removing cost considerations.

Hernández-Lobato and colleagues proposed the *Predictive Entropy Search for Multi-Objective Bayesian Optimization* (PESMO) method [10]. PESMO uses predictive entropy search as the acquisition function. This function represents each objective using an additive component, which enables a decoupled evaluation approach to be adopted. The approach was subsequently extended to also consider constraints (again where decoupling is possible) [9].

Suzuki et al. developed the *Pareto-frontier entropy search* (PFES) approach [18]. PFES is also an entropy approach but considers the entropy in objective-space rather than decision-space, which is computationally simpler. This method also includes cost in evaluating the objectives by including cost in the denominator of the acquisition function. Like PESMO, the approach is easily extended to consider decoupled evaluations.

The Joint Entropy Search (JES) proposed in [19] is also able to take into account different costs and decoupled evaluations, although the authors did not actually experiment with it because they expected little benefit from decoupled evaluations.

Iqbal and colleagues proposed the *Flexible Multi-Objective Bayesian Optimization* (FlexiBo) algorithm [11]. The approach uses a decoupled evaluation in the Bayesian optimisation run but a coupled initial DoE procedure. It additionally learns the solution-dependent cost function for each objective. FlexiBo estimates for each individual optimistic and pessimistic objective values, which are identical if the objective has been evaluated. From that, it computes an optimistic and pessimistic Pareto front as the boundaries of the "Pareto region", and uses an acquisition function that estimates the expected reduction in the volume of this Pareto region, divided by the respective cost.

Buckingham et al. extended the multi-attribute Knowledge Gradient [3] to the case where objectives can be evaluated independently [6]. The authors demonstrate the benefit of independent evaluation not only when the computational costs for objectives differ, but also when the lengthscales of the modelled landscapes (which determine the smoothness of the landscape) differ. Independently, [8] propose to adapt a hypervolume-based Knowledge Gradient approach to allow for decoupled evaluation of the objectives.

A slightly different problem is considered in [14,5], where one objective is much cheaper (essentially free) to evaluate than the other. They directly incorporate evaluation of the cheap objectives into a pair of hypervolume-based acquisition functions for BO. Consequently, the cheap objectives are evaluated many times while the acquisition function is optimised.

A summary of the different approaches is shown in Table 1.

Table 1: Existing methods for decoupled cost-aware multi-objective optimisation

Approach	Design of experiments Decoupled?	Design of experiments Cost-aware?	Optimisation Decoupled?	Optimisation Cost-aware?	Acquisition function
PESMO [10]	✓	✗	✓	✗	predictive entropy search
PFES [18]	✗	✗	✓	✓	cost-weighted Pareto frontier entropy
FlexiBO [11]	✗	✗	✓	✓	cost-weighted objective space entropy
C-MOKG [6]	✗	✗	✓	✓	cost-weighted multi-objective knowledge gradient
CA-MOBO [1]	✗	✓	✗	✓	cost-weighted Tchebycheff scalarised UCB
HV-KG [8]	✗	✗	✓	✓	cost-weighted hypervolume knowlege gradient
JES [19]	✗	✗	✓	✓	joint entropy search
This paper	✓	✓	✗	✗	N/A

3 Empirical work

In this section we consider a range of different properties/configurations of a problem which may influence the effectiveness of a decoupled DoE, and investigate these empirically. LHS designs are generated using the `R` package `lhs` [7] with the maximin option.

3.1 Initial DoE when evaluations are decoupled

We begin with an illustration of a greatly simplified case, where the costs of querying each of two objectives are the same. The two objective functions are generated by Gaussian process models (GPs)—so we are assured that emulation by a trained GP will fit the modelling assumptions, and we also directly utilise the hyperparameters of the objective function GPs, removing the effect of having to infer these, so there is no model mismatch (i.e. our model is perfectly capable of modelling the generating process).

Our goal is to study the effect of coupled versus decoupled designs of experiments (DoE) on the uncertainty on the Pareto front in this very controlled problem configuration, before moving towards a more realistic scenario. We generate samples from a GP model for each objective and use it as the ground truth for fitting the GP approximation models. An example of the generating models and respective mapping to the objective space is given in Figure 1.

In Figure 2 an example is shown where the DoE for the first objective is the same while the second objective is either coupled (left panel) or decoupled (right

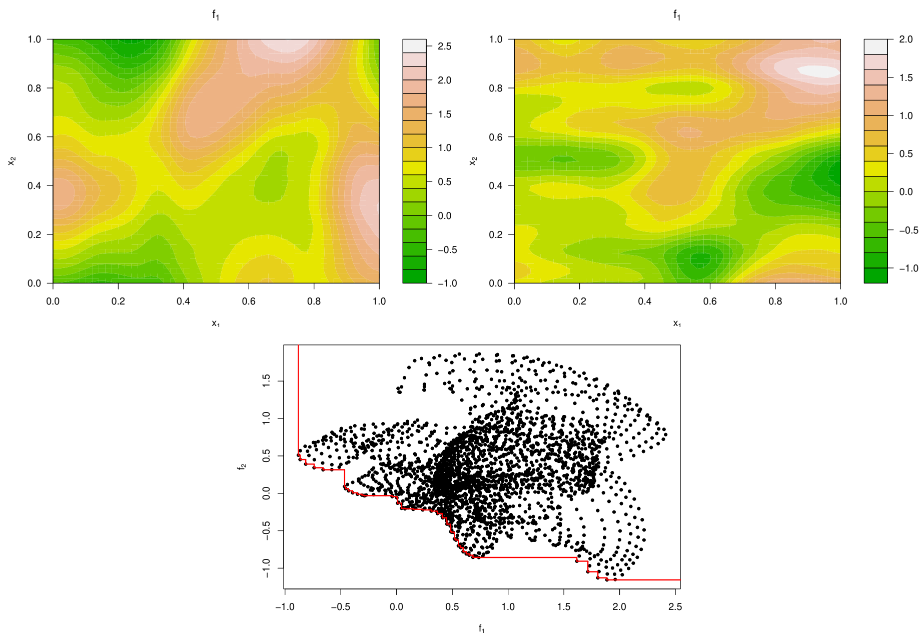

Fig. 1: Top: two realisations of Gaussian process priors, with Matérn 5/2 covariance kernel, with lengthscale hyperparameters $(0.3, 0.4)$ (resp. $(0.4, 0.2)$) for f_1 (resp. f_2), and unit variance. Bottom: corresponding image in the objective space (grid sampled), with the estimated Pareto front highlighted in red.

panel). The decoupled DoE of the second objective is obtained by augmenting the first objective DoE while maintaining the LHS structure. Attainment functions are obtained by taking a joint sample on a 51×51 grid from the GP posterior for each objective conditioned on the observations, then determining the non-dominated observations. The q-Attainment front is then representing the area that is dominated by a fraction q of all the estimated Pareto fronts generated. One visible effect is that when both objectives are jointly evaluated, the area that is dominated (attainment value $= 1$) is larger. This is probably because in the decoupled case, solutions are never surely dominated (even though the domination probability is extremely low), as no location has been queried under *both* objective functions (this can be further seen with the left panel having triangles denoting locations with a pair of known objective values, and the right panel having no triangles).

To help measure the uncertainty on the Pareto front associated with the fitted GPs, we use below the so called Vorob'ev deviation (VD), a set based variance metric that measures the variability of the q-Attainment fronts relative to the true frontier—see, e.g., [4] for further details on its properties. Algorithm 1 summarises the testing procedure.

Algorithm 1: Pseudo-code for the testing procedure

1: Generate the first design of experiments X_1 for objective 1.
2: **if** *Coupled case* **then**
| $X_2 = X_1$, the DoE of the second objective is the same.
end
else
| Generate X_2 the second DoE. (Decoupled case)
end
3: Build GP models.
4: Generate s conditional samples on some designs X_s from all GPs.
5: Compute the s sets of non-dominated points on couples of samples from the different GPs.
6: Compute the corresponding Vorob'ev deviation.

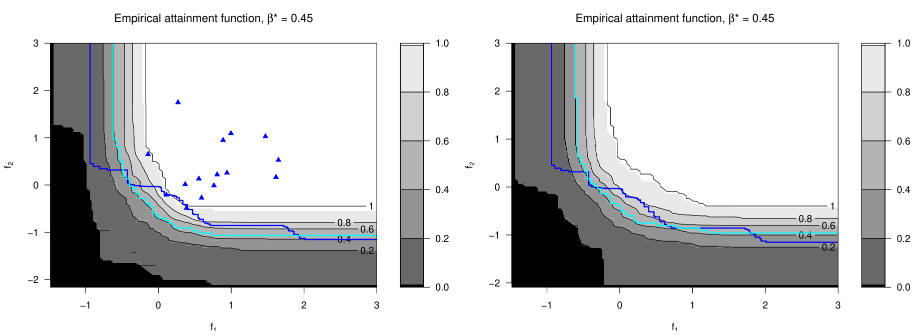

Fig. 2: Attainment function representation in the coupled (left) and decoupled (right) cases. The blue triangles mark observations in the coupled case, where both objectives are evaluated. The cyan line represents the estimated Pareto front of the GP while the reference Pareto front is in blue.

Figure 3 shows the Vorob'ev deviation of the coupled and decoupled designs for two cases. In the left panel, the design for each objective is uniformly random, while in the right panel, a LHS design is used for the first objective, and then an augmented LHS is used for the second objective. The panels show the results of 11 independent runs, with 10 replications for the design of the second objective in case of decoupled design (visualised as boxplots).

As expected, LHS designs (right panel) lead to slightly lower Vorob'ev deviations than random uniform designs (left panel), in particular for the coupled case. The larger benefit in case of the coupled design is probably due to the fact that a cluster or gap in the sample space of the first objective is unlikely to be duplicated for the second objective in the decoupled design. When LHS is used, the coupled design (red dots) seems to yield a lower Vorob'ev deviation than the decoupled designs, possibly due to the effect mentioned above on the size of the known dominated region. This difficulty in precisely estimating the Pareto front may also pose challenges for the optimisation procedure, as a reference Pareto front is generally required by acquisition functions. Note, however, that in these experiments we assume equal cost of sampling the two objectives, and equal lengthscales of the two objectives. As we see later, in other cases decoupling may be beneficial.

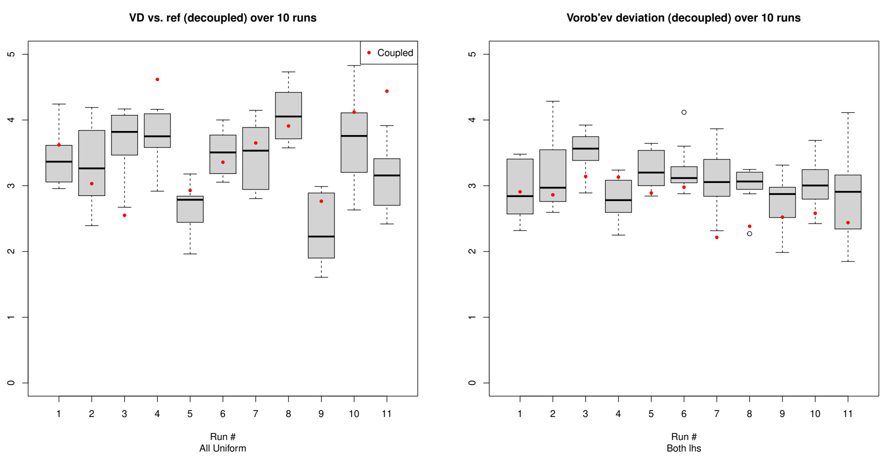

Fig. 3: Vorob'ev deviation against the true Pareto front. Boxplots are for decoupled designs, over 11 different runs and 10 replications per run for the second stage design in the decoupled case; red dots are the coupled designs. In the left figure, the designs are uniformly sampled, while in the right figure, an augmented LHS is used to complement the first LHS design. The value of the coupled design is in red.

3.2 Initial DoE when evaluations have different costs

Now let us assume the cost is different between different objectives f_1 and f_2. The first tasks are to define the total time budget for experiments and get relative costs of f_1 and f_2. We will then consider four alternative approaches to DoE, including a coupled LHS baseline.

1. (Fixed LHS) Rather than sampling incrementally, we use a fixed coupled LHS design of the required size.
2. (Coupled) Both objective functions are evaluated together.
3. (Decoupled naïve) Both objective functions are evaluated the same number of times, but at differing locations (generated by Augmented LHS using the `optAugmentLHS` function from the `R` package `lhs` [7]).
4. (Decoupled) The allocation of total budget to the two functions depends on lengthscales and relative costs, according to Eq. 1. Objectives with smaller lengthscales and smaller cost are sampled more often.

Considering how to split the computational budget, let us consider the simplest case of optimising a (weighted) sum of two objectives. In such a case, if we want to minimise integrated mean squared prediction error (IMSPE) assuming Gaussian process surrogate models with identical lengthscales, then it is not possible to improve beyond coupled sampling, as the variances of the two functions just add up, and the optimal design for each function would be the same. However, if the costs or lengthscales are different, then we could use IMSPE to determine an appropriate allocation of the budget to the two functions by choosing the number of samples n_1 allocated to objective 1 such that the following is minimised:

$$\min \frac{\text{IMSPE}_1(n_1)}{c_1 \times n_1} + \frac{\text{IMSPE}_2(N - n_1)}{c_2 \times (N - n_1)}, \tag{1}$$

where N is the total budget, IMSPE_1 (IMSPE_2) and $c_1(c_2)$ are the IMSPE and cost of evaluating objective $f_1(f_2)$, respectively.

In practice, if the lengthscales are not known, they may be estimated from initial data and Eq. 1 may be optimised sequentially rather than all at once. That is, in the coupled case we iteratively sample the solution that, if both its objectives are evaluated, reduces the IMSPE the most. For the decoupled case, we sample the solution and objective which maximally reduces the IMSPE as calculated in Eq. 1.

As in the previous section, we rely on GP samples in a two-dimensional decision variable space to define a ground truth. We start with the same four coupled initial designs for each objective in the various cases, based on LHS, then add additional samples in a way that minimises IMSPE. For the coupled option, a discrete search over a thousand uniformly sampled candidates is performed at each iteration. As for the decoupled version, a local optimisation is conducted from the best out of one hundred uniformly sampled candidates.

Figure 4 shows the results for the case that the lengthscales for both objectives are equal and known, here, (0.3, 0.4) for the Matérn 5/2 covariance kernel. In the left column, the evaluation cost for both objectives is the same, in the middle column the second objective is five times more expensive, and in the right column, the second objective is 10 times more expensive.

The top row depicts the IMSPE separately for each objective. In the case of equal cost to evaluate both objectives, also the decoupled designs reduce IMSPE equally for both objectives. However, if the evaluation cost for the objectives differ, the decoupled design samples the cheaper f_1 (black +) more often, reducing its IMSPE much more than the IMSPE of the expensive f_2 (red +).

The following rows 2-4 show aggregated performance metrics, namely the average IMSPE, the average root mean squared error (RMSE), and the Vorob'ev deviation. The results are consistent across all metrics: if the costs of the different objectives are equal, the IMSPE-minimising coupled and decoupled approaches perform very similarly (and best), not only with respect to IMSPE but also RMSE and VD. Next best is the fixed LHS and then, significantly worse, the naïve design. This is interesting as it suggests that iteratively choosing points to minimise IMSPE (decoupled as well as coupled) yields not only a lower IMSPE but also a lower Vorob'ev deviation than a standard LHS of the same size. Decoupling naïvely by two augmenting LHS is clearly worst.

If we look at cases of different costs (f_2 five times as expensive (middle column) or 10 times as expensive (right column)), the differences become more pronounced, and the decoupled design clearly beats the coupled design, as it can sample the cheaper objective more often and make better use of the available budget.

Similar results are obtained if the lengthscales are estimated and updated every iteration, see Figure 5. Note, however, that in these experiments we took 20 initial samples based on LHS, as more data is needed to estimate lengthscales reliably.

Additional experiments (see Appendix) with different lengthscales for the two objectives ($(0.3, 0.4)$ for the first, $(0.4, 0.2)$ for the second objective) do not seem to show a significant differences, but this may simply be because the chosen lengthscales were still quite similar.

Finally, we treat the case when the costs are varying depending on both x and the objectives, and we know the cost function. The cost functions correspond as well to samples from GPs, this time with lengthscales (0.5, 0.8) and (0.6, 0.7) for the respective objectives, while for the objective values we again use $(0.3, 0.4)$ for the first, $(0.4, 0.2)$ for the second objective. The results are given in Figure 6. The decoupled strategy is again more efficient to reduce the IMSPE the fastest, but there is no noticeable difference in terms of Vorob'ev deviation. The naïve decoupled design does not make use of the cost information and is thus clearly worse. Note that the fixed LHS strategy is not sensible here, as it is necessary to learn about the evaluation cost and use this information in an incremental design.

4 Discussion and future research ideas

In this paper, we have examined the possibility of improving the quality of the surrogate models obtained through a DoE in case of multi-objective optimisation where the evaluation of the different objectives can be decoupled. We found that for the case of equal costs and lengthscales for the two objectives, decoupling the evaluations (i.e. evaluating different solutions on different objectives) did tend to worsen the quality of the Pareto front estimate as measured by Vorob'ev deviation. However, when objectives had different costs, decoupling could improve results substantially in terms of total IMPSE, RMSE, and Vorob'ev deviation. Interestingly, we found that even in the case of equal costs and lengthscales, allocating samples iteratively by minimising IMSPE yielded better IMSPE, RMSE and Vorob'ev deviation than using an equally sized LHS design.

While in this paper we have only considered the case of two objectives, we see no reason why our conclusions should be any different also for more than two objectives. Indeed, one might expect that the greater flexibility in terms of which objectives to evaluate for a solution could lead to even larger benefits of decoupled experimental designs. However, we leave the experimental confirma-

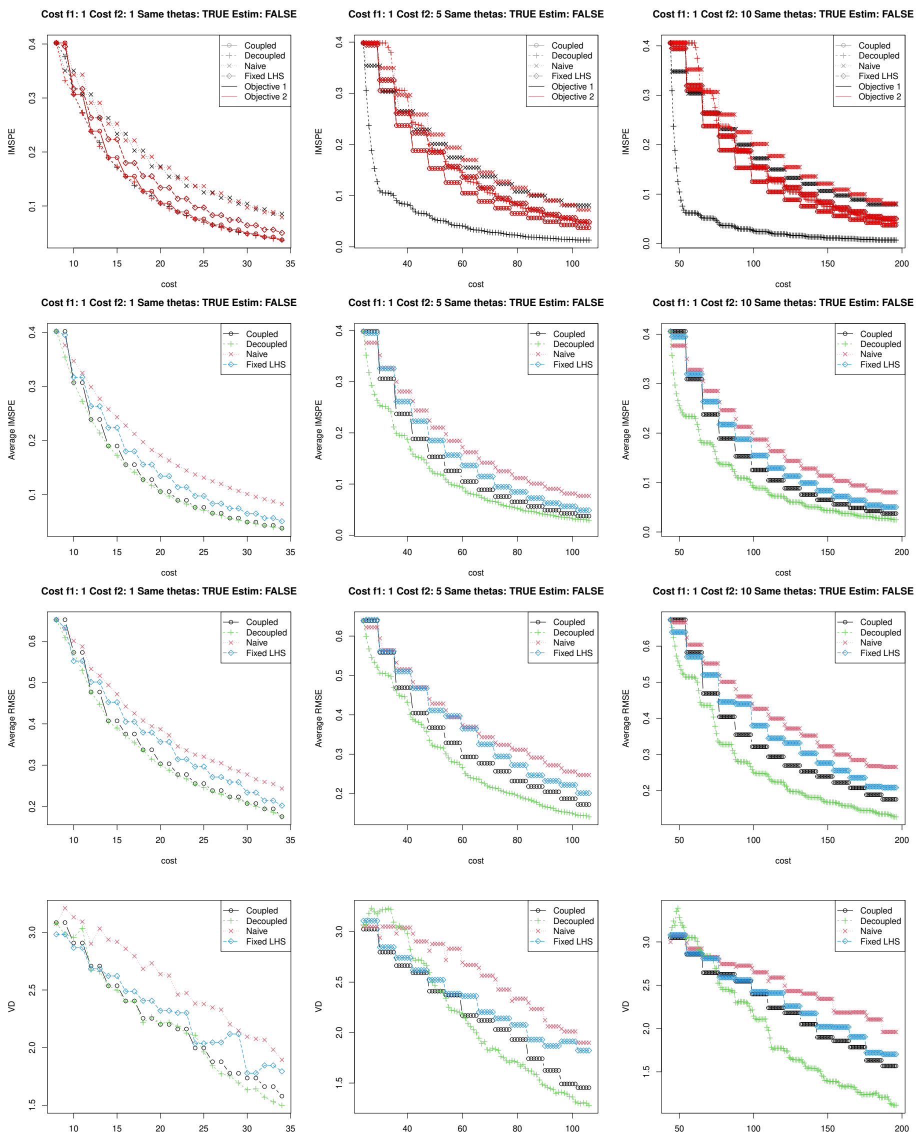

Fig. 4: Results of different metrics depending on the cost incurred. In the left column, both objectives have equal cost, in the middle column the cost for f_2 is 5 times as high, and in the right column, the cost for f_2 is 10 times as high. In this figure, lengthscales are equal and assumed known.

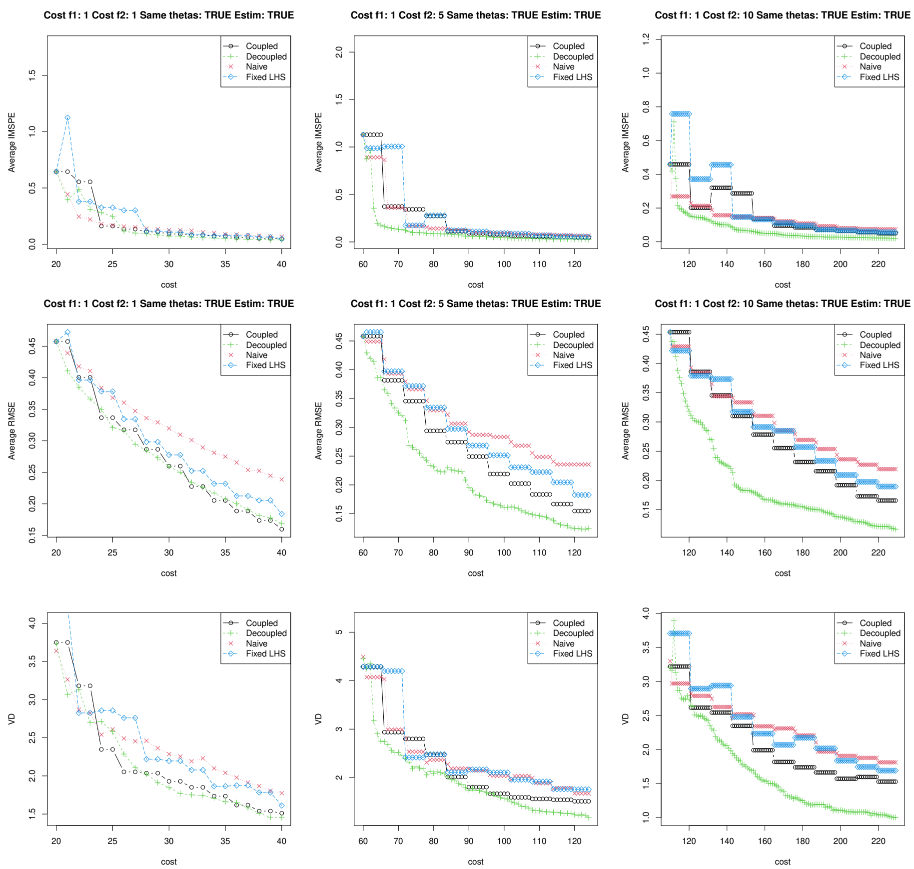

Fig. 5: Results of different metrics depending on the cost allocated. In the left column, both objectives have equal cost, in the middle column the cost for f_2 is 5 times as high, and in the right column, the cost for f_2 is 10 times as high. In this figure, lengthscales are equal but unknown (learned and in every iteration).

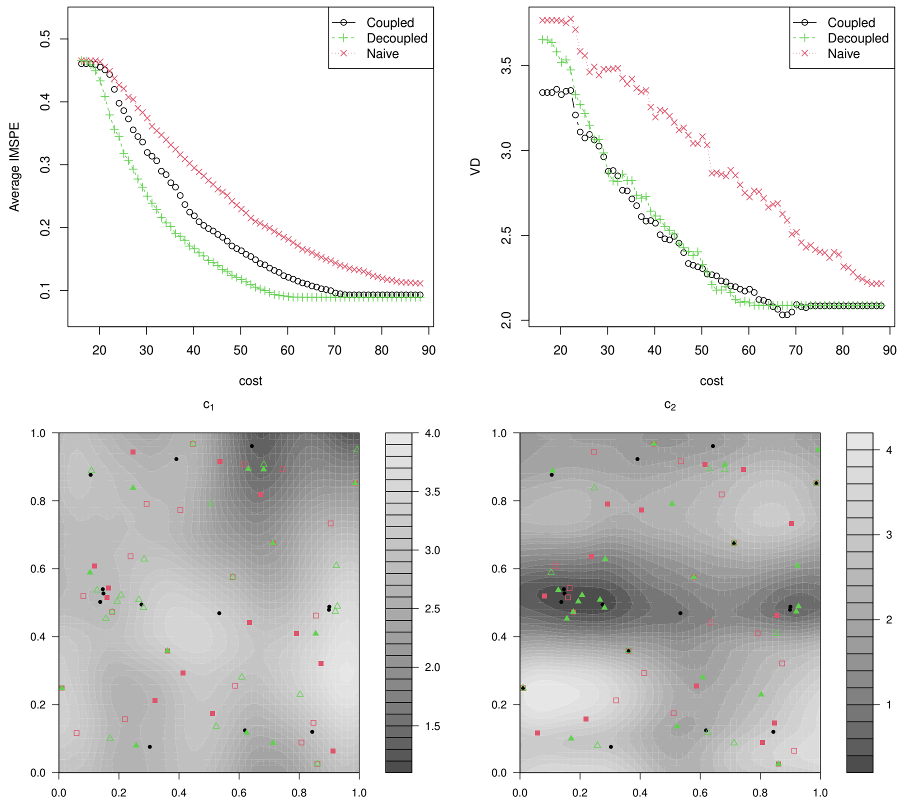

Fig. 6: Variable cost for each objective. Top left (resp. right): IMSPE (resp. VD) vs. cost for the various strategies. Bottom: cost surface and evaluated points for each strategy: black dots for coupled, red squares for naïve and green triangles for decoupled. Empty squares and triangles are evaluated on the other objective. naïve doesn't take into account the cost.

tion of this hypothesis to future work. Our results use GP generated functions to avoid the issue of model mismatch. However, it would be good to confirm results also on other types of functions. Finally, in the future we plan to investigate other sampling strategies such as taking into account the posterior of the first objective when deciding where to evaluate the second objective, or to learn the cost landscape (if the cost depends on the solution evaluated) on the fly.

The code for reproducing the results is available at `https://github.com/mbinois/DecoupledDoe`.

5 Acknowledgements

This report benefited from wider discussions within the “Surrogates” working group of the Dagsthul Seminar 23361 Multiobjective Optimization on a Budget. This group’s members included Thomas Bäck, Mickaël Binois, Jürgen Branke, Jonathan Fieldsend, Ekhine Irurozki, Pascal Kerschke, Boris Naujoks, Robin Purshouse, Tea Tusar, Vanessa Volz, Hao Wang and Kaifeng Yang. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission. RCP’s contribution was supported by SIPHER (MR/S037578/1), a UK Prevention Research Partnership funded by the UK Research and Innovation Councils, the Department of Health and Social Care (England) and the UK devolved administrations, and leading health research charities https://ukprp.org/.

References

1. Abdolshah, M., Shilton, A., Rana, S., Gupta, S., Venkatesh, S.: Cost-aware multi-objective Bayesian optimisation. arXiv preprint arXiv:1909.03600 (2019)
2. Allmendinger, R., Knowles, J.: Heterogeneous objectives: State-of-the-art and future research. arXiv arXiv:2103.15546 (2 2021)
3. Astudillo, R., Frazier, P.: Multi-attribute Bayesian optimization under utility uncertainty. In: Proceedings of the NIPS Workshop on Bayesian Optimization. vol. 172 (2017)
4. Binois, M., Ginsbourger, D., Roustant, O.: Quantifying uncertainty on Pareto fronts with Gaussian process conditional simulations. European Journal of Operational Research **243**(2), 386–394 (2015)
5. Binois, M., Picheny, V.: GPareto: An R package for Gaussian-process-based multi-objective optimization and analysis. Journal of Statistical Software **89**(1), 1–30 (2019)
6. Buckingham, J.M., Gonzalez, S.R., Branke, J.: Bayesian optimization of multiple objectives with different latencies. arXiv preprint arXiv:2302.01310 (2023)
7. Carnell, R.: lhs: Latin Hypercube Samples (2022), `https://CRAN.R-project.org/package=lhs`, r package version 1.1.6
8. Daulton, S., Balandat, M., Bakshy, E.: Hypervolume knowledge gradient: a lookahead approach for multi-objective Bayesian optimization with partial information. In: International Conference on Machine Learning. pp. 7167–7204. PMLR (2023)
9. Garrido-Merchán, E.C., Hernández-Lobato, D.: Predictive entropy search for multi-objective Bayesian optimization with constraints. Neurocomputing **361**, 50–68 (2019)
10. Hernández-Lobato, D., Hernandez-Lobato, J., Shah, A., Adams, R.: Predictive entropy search for multi-objective Bayesian optimization. In: International Conference on Machine Learning. pp. 1492–1501. PMLR (2016)
11. Iqbal, M.S., Su, J., Kotthoff, L., Jamshidi, P.: FlexiBO: A decoupled cost-aware multi-objective optimization approach for deep neural networks. J. Artif. Intell. Res. **77**, 645–682 (2023)
12. Lee, E.H., Eriksson, D., Perrone, V., Seeger, M.: A nonmyopic approach to cost-constrained Bayesian optimization. In: Uncertainty in Artificial Intelligence. pp. 568–577. PMLR (2021)

13. Lee, E.H., Perrone, V., Archambeau, C., Seeger, M.: Cost-aware Bayesian optimization. arXiv preprint arXiv:2003.10870 (2020)
14. Loka, N., Couckuyt, I., Garbuglia, F., Spina, D., Nieuwenhuyse, I.V., Dhaene, T.: Bi-objective Bayesian optimization of engineering problems with cheap and expensive cost functions. Engineering with Computers (1 2022)
15. M. D. Mckay, R.J.B., Conover, W.J.: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics **42**(1), 55–61 (2000)
16. Niederreiter, H.: Random number generation and quasi-Monte Carlo methods. SIAM (1992)
17. Snoek, J., Larochelle, H., Adams, R.P.: Practical Bayesian optimization of machine learning algorithms. Advances in neural information processing systems **25** (2012)
18. Suzuki, S., Takeno, S., Tamura, T., Shitara, K., Karasuyama, M.: Multi-objective Bayesian optimization using Pareto-frontier entropy. In: International Conference on Machine Learning. pp. 9279–9288. PMLR (2020)
19. Tu, B., Gandy, A., Kantas, N., Shafei, B.: Joint entropy search for multi-objective Bayesian optimization. Advances in Neural Information Processing Systems **35**, 9922–9938 (2022)

A Results if objectives have different lengthscales

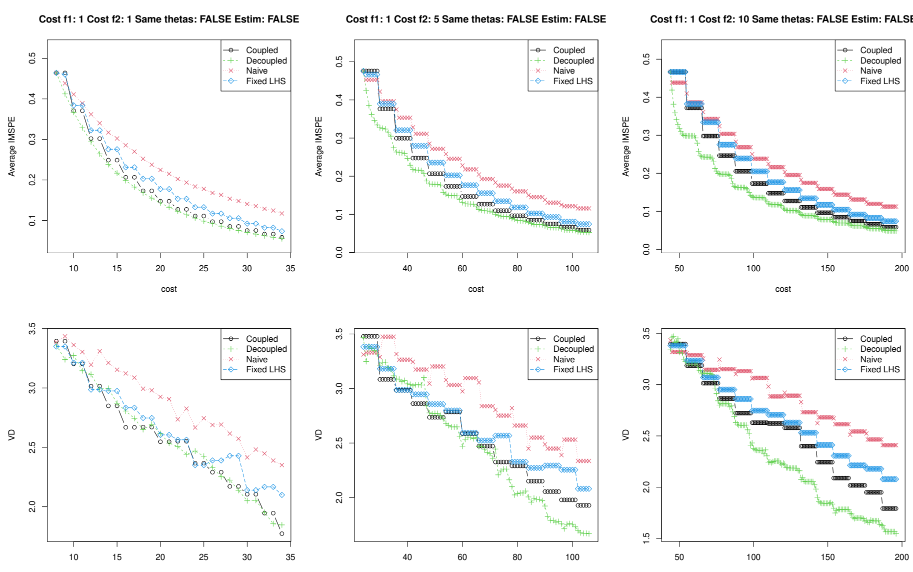

Fig. 7: Results of different metrics depending on the cost incurred. In the left column, both objectives have equal cost, in the middle column the cost for f_2 is 5 times as high, and in the right column, the cost for f_2 is 10 times as high. In this figure, lengthscales are equal and assumed known.

[6] observe that allowing decoupled evaluation of objectives is beneficial if the different objective functions have different lengthscales, i.e., if one objective is smooth and varying slowly, while the other is highly multimodal. Intuitively, one would like to allocate more samples to the more difficult objective. To test this, we have also run experiments where objectives have different lengthscales, in particular we used $(0.3, 0.4)$ for the first, $(0.4, 0.2)$ for the second objective. Results are summarised in Figure 7 for know lengthscales and Figure 8 for learned lengthscales. The results are very similar to the results with equal lengthscales reported above, which may be due to the fact that the lengthscales chosen were too similar to observe a significant difference.

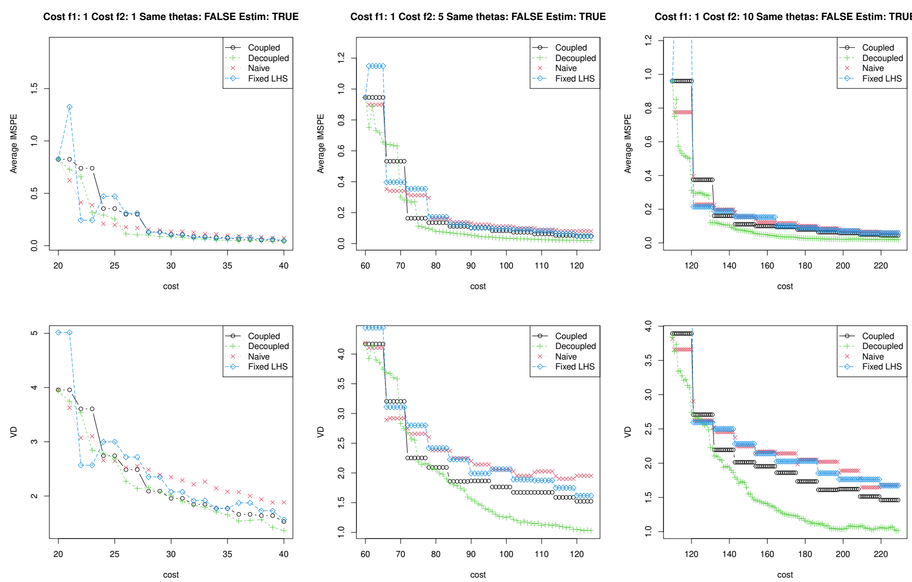

Fig. 8: Results of different metrics depending on the cost allocated. In the left column, both objectives have equal cost, in the middle column the cost for f_2 is 5 times as high, and in the right column, the cost for f_2 is 10 times as high. In this figure, lengthscales are equal but unknown (learned and in every iteration).