We thank the reviewers for their critical assessment of our work, and for the overall positive reception. Below we reply to each review to clarify points and provide additional insights into the problem and methods, which we hope will better highlight the technical novelty and contribution of our work.

Reviewer 1

- To the first point in weaknesses. We agree on this important point and we use cross-validation to verify this assumption. For instance, in our real-world applications (described below for R2), the model fit of LCE-A is much better than other models. Note also that the ABR simulator uses real throughput traces and video data, and the results there indicate a robustness of assuming a shared spatial kernel.
- To the second point. Thanks for raising this! The number of parameters is $(1 + m) \times C + d + m$ for embedding dimension m. Although LCE-A has more parameters, the model complexity is not necessarily higher. The kernel $k^z(E(c), E(c'))$ imposes regularization in the model and encourages correlations across component functions. Lines 124–128 in Sec 2.2 discuss the connection between the regularization of a GP prior and the function norm in the RKHS. With stronger correlation there is a larger norm and more regularization. We will clarify this in the paper.

Reviewer 2

- To the first point in comments. Thanks for asking! We have applied this method to two real problems: 1) tuning a live video playback controller under different connection qualities, to minimize stall time while maintaining high quality. There are 6 controller parameters and 5 connection qualities (from excellent to poor). 2) tuning data fetch parameters for an app surface, contextualized for different countries and connection types (4g, wifi, etc). The goal was to reduce data utilization without affecting app performance. For both problems, performance metrics had to be evaluated by A/B tests that took several days (though multiple design points could be run in parallel). Metrics could not be logged at the level of context, meaning typical contextual BO methods could not be used. For both problems, the LCE-A model had the best cross-validation performance and so was used for the optimization. LCE-A found the best policy compared with random search and standard BO, and significantly improved system performance.
- To the second point. The embedding sizes are set to be 1 in the simulation studies. If LCE-A fails to learn the embedding and the corresponding correlation structure, then it essentially falls back to SAC, which doesn't borrow strength across contexts; we will add discussion of this. As noted, when possible, pre-training can accelerate learning in the small-data regime, and is something we will investigate in future work.
- The two context buckets have very different network conditions and their bandwidths are unstable; thus the QoEs have larger variances. We will clarify this in the paper.
- Thanks for the suggestion in related work! We will add this into the discussion. Our work assumes an additive structure of component functions and fits nicely into the framework of grey-box BO $g(h(\mathbf{x}))$. Here, g is an additive function and h consists of unknown context functions which are learnt via the new LCE kernel.

Reviewer 3

- Thanks for the reference [TSC11]! We were unaware of this paper, but it is indeed related and our work nicely extends the methods developed there. Eqn. (1) of TSC11 is the problem that we solve as well. When using the LCE-A and SAC kernel, the acquisition function was the same as their expected policy score improvement. Our work extends TSC11 by handling the situation where the outcomes $f(x_e, x)$ (i.e. $f_c(x_c)$ in our paper) are not observable under each environment (i.e. context) separately. We show how to optimize Eqn. (1) of TSC11 when given only total score, which is critical for settings where the controller operates across heterogeneous environments and score cannot be easily attributed to each (the notorious credit assignment problem). We will add discussion of TSC11 and its area of work, and we believe that the connection to it broadens and strengthens the technical contribution of our paper. We hope that clarifying this point has made the contribution of our work more apparent.
- On policy performance in easy/difficult contexts. In our simulation studies, the easy contexts are the dense contexts with the majority of the weight and the difficult contexts are the sparse contexts with small weights. We show that LCE-A improves performance in difficult contexts by learning the context correlation structure, e.g. in Fig. 4(b).
- To the second question in comments. In our current studies, embeddings are learnt similar to inferring the matrix B inside an ICM kernel by maximizing likelihood. We will clarify this.

Reviewer 11

- Thanks for your suggestions on improving organization of the paper! For Q3, the matrix B is estimated by maximizing marginal likelihood, as is typical for multi-task GP fitting.
- For Q4, the struggle BO has with high dimensions is with the number of parameters in the feature space. That is remaining constant across these models. LCE-M does have additional hyperparameters, but these tend to increase regularization and do not cause problems of the sort faced by high-dimensional BO (see response to R1).