We thank the reviewers for their insightful and positive feedback. We are encouraged that they found our idea to be important (R3 , R4 , R7), theoretically grounded (R3 , R4 , R7), clear (R3 , R4), and novel (R4 , R7); and our analysis to be rigorous and insightful (R4 , R7). We are glad that both R4 and R7 recognized the importance and novelty of this work and suggested an acceptance. Though R3 has recommended a rejection, the weaknesses identified by R3 are in fact clarifying questions. Very interestingly, all the pointed weaknesses/questions are in fact the biggest strengths of our framework. Hence, we are hopeful that R3 will champion our paper for acceptance.

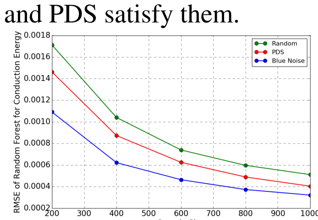
R3 Contributions and Significance: Currently, there do not exist any theoretical framework that allows the study of generalization of generic task-agnostic sample designs (e.g., space-filling). Our work proposes a fundamentally novel approach to this problem and significantly advances our understanding of sample design in the context of ML. This was achieved by connecting ideas from disparate fields – spectral analysis of sampling design (from DOE) and generalization error (from ML) using a statistical mechanics framework. In this process, we have made new contributions to both fields. In ML, our analysis enabled us go beyond the generalization analysis of random i.i.d. designs or other simple probabilistic variants. In DOE, it provided the much needed theoretical underpinning for the apparent superiority of space-filling designs over random designs. Further, this is the first work to construct optimal blue noise designs in dimensions $d > 2$. Finally, we derived new results for high-dim. PDS and blue noise (see supp mat). We hope this addresses R3 's concerns regarding the significance of our contributions.

R3 It is not clear how these results generalize beyond the specific samplers explored. Unlike existing approaches, our framework supports generic sample designs and this is **precisely our major contribution**. By combining Theorem 2 with Eq. 5, one can calculate the generalization gap of arbitrary sample designs (see line 156). We continued making our analysis more focused in rest of the paper by deriving results for SOTA samplers (as pointed our by R4 and R7). One can use our tools to study other complex designs (e.g., Latin Hypercube, Quasi Monte Carlo) that is currently not possible with traditional approaches.

R3 In practice, are the spectral properties of samplers often known? Spectral properties (i.e., power spectra) of any sample design can be computed (line 94 and Figure 1 in supp mat). While in some cases, their analytical forms (or upper bounds) are known – our theory is not limited to analytical PSDs, thus making it **broadly applicable**.

R3 Is it the spectral properties that are optimal, or they are picking up something else? We are not sure if we understand this question clearly. Every sample design can be represented in its spectral form and our analysis suggests that for an ideal design, the power spectrum must approach to zero power at low freq. (line 266). As these claims are proved rigorously, there are no confounding variables as it might have been the case with only empirical results.

R3 Blue noise optimize spectral properties-what are optimal spectral properties? seems potentially circular. This result is not circular in any way. Our guidelines for optimal power spectra form (line 265) are derived for any generic sample design (Proposition 3 and Theorem 2) and it so happens that blue noise and PDS satisfy them.



R3 Experiments are bit sparse. In addition to neural network regression on synthetic functions and real-world scientific simulator in the paper, we further experimented with (Randomforests, GBT), and results were consistent (will add to revision). Result that random design matches PDS (Fig S3b) was an artifact and is removed by increasing the number of independent realizations from 20 to 30. Due to theoretical nature of this work, we hope the lack of large scale experiments will not be considered as a weakness.

R3 Certain relationships not clear (why relate PSD and PCF). Theorem 1 along with realizability conditions establish a fundamental relationship between the PSD and PCF of designs. This relationship is required for constructing optimal forms of PDS and blue noise, and carrying out analysis only on realizable spectras (not every power spectra has corresponding sample design, though, other way around is always true) (see line 111).

R3 Some proof explanations were a bit sparse (e.g. props 4, 5). These results can be obtained simply by plugging in best- and worst-case PSD (as defined in the paper) in Eq 9. We will add this additional step in the revised version.

R4 , R7 Comparison with [17]. The only overlap between both papers is the formal construction of PDS, which is needed for Lemma 8 of our paper. While [17] and other space-filling papers design PDS heuristically/empirically, our analysis is entirely different and theoretically grounded.

R4 . Limitation of theory (or potential extensions) are (e.g., probabilistic bounds, manifold sampling, adaptive sampling). Theory allows one to optimize the zero region (see Lemma2). Loss specific analysis and distinguishing PDS and Blue Noise empirically are excellent suggestions and will be added as the future work in the revised version.

R7 . Technical reason behind best and worst case choices is given in [6], and will be clarified in the revised version. The definition of isotropic and homogeneous designs will also be clarified. Sorry for the typo – the realizability condition will be written in its generic form instead of the radial form.