We thank the reviewers for their time and feedback, and will address all documented typos. In addition, we will clarify: (1) our usage of the condition number κ in Figure 2, (2) algebraic multiplicity of eigenvalues in the definition of pseudo-determinant (pdet), and (3) that w in Theorem 1 may be arbitrary.

To Reviewer #2 We agree that our results on MSE (mean-squared error) do not directly contradict the conclusions of Hastie et al. (2019) (specifically section 3.3 result 2) on the generalization risk, which are stated in terms of the mean-squared *prediction* error. We merely observe that our MSE expressions demonstrate that the minimum norm solution has a capacity for learning (in the sense of achieving MSE below that of the null estimator) even when the signal-to-noise ratio is 1. Thanks for pointing this out, we will clarify it further in the final version.

Regarding other applications of surrogate designs, our methodology should be useful whenever the inverse of a random matrix is being studied (e.g., for analyzing randomized Newton-type methods). It is also possible that other surrogate designs can be devised for analyzing other functions of random matrices.

To Reviewer #3 We agree that an answer to Conjecture 1 would be a strong contribution and are currently actively researching this. However, we believe Theorem 1 should be stated first because we believe it is the more important "exact expression" result which both requires less assumptions and is our starting point for investigating Conjecture 1.

With respect to Reviewer #3's comments on our broader impact statement, our results do not contradict Nakkiran et al. (2020) because they consider regularized i.i.d. designs whereas we consider unregularized i.i.d. / surrogate designs. To improve clarity, we will revise our broader impact statement to (1) make clear that the conclusion arises from the increasing MSE as n increases (in underdetermined $d/n > 1$ regime) and that (2) more data **may** lead to worse generalization, consistent with their earlier findings (Nakkiran et al., 2019) and result 1 of "Towards a more general characterization" subsection.

References

Hastie, T., Montanari, A., Rosset, S., and Tibshirani, R. J. (2019). Surprises in high-dimensional ridgeless least squares interpolation. *arXiv preprint arXiv:1903.08560*.

Nakkiran, P., Kaplun, G., Bansal, Y., Yang, T., Barak, B., and Sutskever, I. (2019). Deep double descent: Where bigger models and more data hurt. *arXiv preprint arXiv:1912.02292*.

Nakkiran, P., Venkat, P., Kakade, S., and Ma, T. (2020). Optimal regularization can mitigate double descent. *arXiv preprint arXiv:2003.01897*.