

1 We are grateful to the reviewers for their valuable feedback. They have clearly understood the paper well; we agree
2 with the assessment of both the strengths and weaknesses of the paper. In our revision, we will try our best to address
3 the concerns and execute the suggestions.

4 We first address the main concerns that shared by all reviewers.

5 • *When does double descent happen in regression?*

6 Double descent happens if the risk curve has a spike at $p = n$, typically caused by a near-zero minimum
7 eigenvalue of $\mathbf{X}_P^\top \mathbf{X}_P$. For example, when \mathbf{X} is an i.i.d. Gaussian matrix, and the noise is i.i.d. noise, then
8 there will be a spike at $p = n$ and double descent happens, regardless of whether \mathbf{X}_P contains the true features
9 or not.

10 • *When the second descent is better than the first one?*

11 In our current setting, our theorems show that the answer depends on the noise level and the decay rate of
12 the eigenvalues of the covariance matrix of \mathbf{X} . If the eigenvalues are all equal, the second descent achieves a
13 lower risk regardless of noise (we assume the variance of the noise does not scale with n). If the eigenvalues
14 decay too fast, i.e., $\kappa > 1$, the second descent achieves a lower risk only if there is no noise.

15 In general, we believe that if the noise level is high enough such that the noise becomes the dominant part of
16 the risk or the eigenvalues of covariance matrix decay too fast, then $p < n$ regime is better. Otherwise, the
17 $p > n$ regime may yield a lower risk.

18 **To Reviewer 1**

19 • We plan to discuss the results in the context of surrounding literature around interpolation and ridgeless
20 regression (including future directions).

21 • We plan to include more discussion of simulations results.

22 • We plan to add more appropriate figure titles.

23 **To Reviewer 2**

24 • We plan to include more discussion of simulation results.

25 • We plan to discuss to what extent we can answer the question in the title, specifically for our current setting,
26 and also what we hope to answer in future work.

27 **To Reviewer 3**

28 • We agree that our setting is non-standard for PCR, although it is plausibly realizable in a semi-supervised
29 learning setting. (We apologize for forgetting to emphasize this important point!) Unfortunately, we don't
30 have an answer for what happens in the standard PCR setting.

31 • We plan to discuss to what extent we can answer the question in the title, specifically for our current setting,
32 and also what we hope to answer in future work.