We thank the expert reviewers for their high quality reviews and their appreciation of the results. We have answered a few (excellent) technical questions below, and hope that in light of this, the reviewers would consider raising their scores.

0A: Novelty & Innovation: Cohen et al. propose solving the problem on the state-action covariance as an online **linear** optimization routine, a class for which the best known upper and lower bounds scale as $\sqrt{T}$. We use alternative convex parameterization for the controller, suggested in Agarwal et al. However, this formulation employs a **highly over-parameterized** representation for the controller (as evident via parameter counting) rendering the program potentially not strongly convex in the parameters, even if the costs are strongly convex in (x, u). This is the **main technical challenge**.

In this context, our results for OGD quantify the conditions under which the over-parameterization does not hinder strong convexity. Those for ONG exploit the specifics here to quantify more general conditions under which curvature is preserved in directions that matter for regret minimization. We'll add a detailed discussion to the paper upon revision.

0B: More general conditions enabling logarithmic regret for OGD: In actuality, the conditions that permit $\log T$ regret for OGD can be extended using the same techniques as presented in the paper with a more intricate analysis; given the reviewers' interest we will spell these out precisely in the final version. The result relaxes the assumption of stability of A to a more general notion of controllability:

Assumption 1. [Diagonal Strong Stability] A controller K is diagonally strongly stable if there exists a *complex* diagonal matrix L and a complex matrix Q such that $A - BK = QLQ^{-1}$ with $\|Q\|, \|Q^{-1}\| \leq \kappa$ and $\|L\| \leq 1 - \gamma$.

While the above is possibly more stringent than strong stability (Cohen et al), the fact that the above assumption is over complex diagonalization makes it quite general. The above extension comes about via a careful fine grained analysis of the Hessian aimed at bounding the effect (on strong convexity) of the time delayed influence of noise on controls introduced by including Kx in the control.

Reviewer #1 Please see **0A**.

1A: On strong stability: In stochastic systems, stable policies, of which strong stability is a quantification, ensure that the size of state ie. $\|x\|_2$ is bounded (independent of T). In this way, they permit operating with bounded instantaneous cost, which is a standard assumption in online learning.

Reviewer #3 Please see **0A**.

3A: Presentation & Writing: We absolutely agree with the importance of readability for theoretical results. To ensure the paper can be read as a stand-alone, we thought to include a running example of 1-d system to illustrate the proof ideas. Perhaps the reviewer can comment on this presentation idea.

3B: Computations concerning the ideal cost: As noted for Gaussian noise with quadratic costs, the gradient and function value are analytically computable. Beyond this, we can get a close estimate by averaging samples. OCO algorithms are robust to $poly(T)^{-1}$ errors in gradient and function value.

3C: Scaling with Noise: Indeed, the regret scales as square of the noise (as it should since the cost is strongly convex), as long as the condition number of noise covariance is fixed.

Reviewer #4 Please see **0A** and **0B**.

4A: Linear policies beyond quadratic costs: Beyond quadratic costs, as noted, the optimal policy might not be linear. But even the offline computation of such a policy poses computational difficulties. Therefore, we compare with and execute linear policies. It is a great future direction to find a good policy comparator class more generally.

4B: On Assumption 2.2 for OGD: Very insightful question. While it is true that the superimposition of a stable controller K renders the effective dynamics $A' = A - BK$ stable, the cost associated with executing a disturbance-action policy is not the same on the original system A (with superimposed K) and the system A'. This is because in the former, the cost (say $\|u_t\|^2$) associated with the action u_t pays for the actions of K in addition to the actions undertaken by the disturbance-action policy. This additional part to the cost makes the analysis particularly hard as it might reduce or eliminate the strong convexity in the disturbance-action parametrization.

However, as stated in **0B**, our technique is robust enough to tackle this issue, thereby extending generality of the results via a more fine-grained eigen characterization of the hessian.

4C: Typos: Indeed, Corollary 4.3 employs assumption 2.2. We agree that K should be an input to the algorithm. Note that when working with Assumption 2.2 (e.g. OGD), $K = 0$ suffices, and when working with Definition 2.4 (i.e. ONG), any strongly stable K suffices; the search for the later can posed as SDP feasibility as in Cohen et al.