We thank the reviewers for their work, especially these difficult times. Our paper uses mathematical tools that many researchers in ML/AI might not be very familiar with. We have added some background both in the main part and in the supplementary material (due to space restrictions, it is impossible to put all definitions in the main part).

[To Reviewer #1] We thank you for your careful reading, supportive comments, and clarifying the relation to prior works. Especially we appreciate your suggestion of underdamped Langevin for further consideration.

[To Reviewer #3] We are considering experiments based on stereographic projection of sphere to illustrate the equivalence of the implementations of our algorithm. Thanks for pointing this out.

[To Reviewer #4] Dimension of gradient and tangent vectors are equal to the dimension of the manifold. The noise vector is the Euler-Maruyama discretization of Riemannian Brownian motion.

[To Reviewer #5] The main points raised are:
1) Sampling on Riemannian manifold lacks of motivation and connection with ML;
2) The results of our paper carry over those of Vempala and Wibisono(2019) to manifold, and obtaining these results on manifold is straightforward.

Regarding the first point, we name a few papers to clarify the ecology of sampling on Riemannian manifold. Girolami and Calderhead[2011] "Riemannian manifold langevin and hamiltonian monte carlo methods", has 1200 citations, amongst, 30 from NIPS, 17 from ICML, 8 from JMLR, 9 from AISTATS, 5 from AAAI and 1 from COLT. Byrne and Girolami[2013]"Geodesic monte carlo on embedded manifold", 107 citations, 4 from NIPS,4 from AAAI, 3 from JMLR and 2 from ICML. Data from a rough counting through google scholar.
Brubaker et al [2012]"A family of MCMC methods on implicitly defined manifolds", AISTATS, 65 citations.
Patterson and Teh "Stochastic gradient Riemannian langevin dynamics on the prob. simplex", NIPS13, 198 citations.
Lan et al [2014]"Spherical Hamiltonian monte carlo for constrained target distributions", ICML, 40 citations.
Liu et al "Stochastic gradient geodesic MCMC methods", NIPS16, 19 citations.
Smith et al "Stochastic natural gradient descent draws posterior samples in function space", NIPS18.
Goyal, Shetty "Sampling and optimization ... in Riemannian manifolds of nonnegative curvature", COLT19.
Zhang et al "Wasserstein control of mirror langevin monte carlo", COLT20.
All the papers are about sampling on Riemannian manifold. It is a well established and active research area of ML.
"I suggest illustrations...occurs naturally," Sampling as a generic technique of generating random points has many applications. Sampling on manifold plays a role whenever constraints are introduced, according to the above papers, these constraints include: simplex with Fisher information metric, sphere, truncated set of $\mathbb{R}^d$ with Hessian metric, general embedded manifold in $\mathbb{R}^d$, etc. They have variety of application backgrounds: posterior sampling, molecule dynamics, truncated statistics, and an important example of embedded manifold is the matrix manifold with applications in robotics and computer vision. We motivate in our paper at line 21-27 with PCA, matrix completion, matrix factorization based on the work of Moitra and Risteski [2020], and continuous game theory based on the work of Domingo-Enrich et al[2020] where manifold langevin is used in zero-sum game, while our result is the potential game counterpart. "I did not understand...", they are connected since manifold Langevin descent-ascend is used in their paper to sample from stationary distribution, and our technique can give insight to the rate of convergence for their method.

About the second point, we argue that the analysis to generalize the result from Euclidean to Riemannian manifold is precisely the contribution, since to our best knowledge, this type of result does not exist in the aforementioned papers. To support the claim that this is not straightforward, we have to point out that the result and the approach of Vempala and Wibisono for $\mathbb{R}^d$ do not hold for general manifold. In fact, it is a very open ended problem for one to find reasonable conditions under which certain convergence results can be proved, and any success of doing this can be considered significant to some extent. For example, the aforementioned paper Goyal and Shetty, COLT 2019 and the following

Mangoubi and Smith [2018]"Rapid mixing of geodesic walk on manifold with positive curvature", Annals of Applied Probability, address the uniform distribution on manifold(and its subsets) with positive/non-negative curvature condition. Orthogonal to their approach, we put topological constraints(line 242: Assumption 1) instead of curvature constraints. Furthermore, the main technique in Vempala and Wibisono's is to control the error caused by discretization, while in our paper, the main technique is to control the error caused by linearization of a curved space, and it needs a careful treatment with very different toolkits(geodesic equations, sectional curvature, Riemann curvature, Jacobi field, parallel transport, etc) compared to Euclidean space. For example, the proof of Theorem 4.4 is based on Jacobi field method that has no trivial connection to Vempala and Wibisono's technique.

Again we emphasize that one contribution of our paper is to provide a general framework with necessary techniques that will be useful in other contexts as well, e.g. adjusted Langevin and Hamiltonian Monte Carlo with manifold constraints. This is also part of the role of a theory paper: to find techniques that might be leveraged by others in different contexts. The list of papers shows that ML community has a growing interest in manifold sampling but the convergence results are much fewer than that for Euclidean space. Filling up the technical gap is exactly the motivation and contribution.