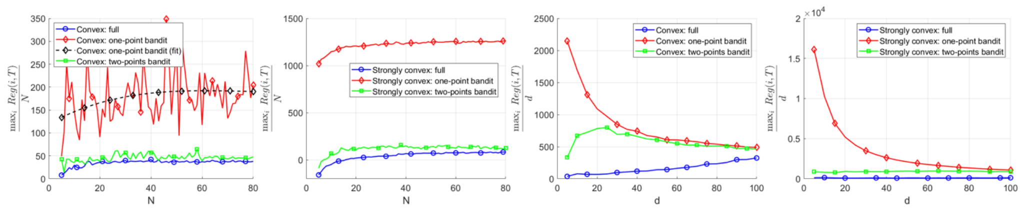
We thank all the reviewers for the thoughtful feedback. We are encouraged that three reviewers voted to accept, and acknowledge that the paper is well written and clearly presented ([R1],[R2],[R3],[R4]), “first to study distributed OCO in random graphs" ([R2]), “novel" and “interesting and potentially related to some real life distributed tasks" ([R3]), and has a “complete characterization" with a “nontrivial" and solid analysis ([R2],[R3],[R4]). We respond to main comments below and will address all feedback in the main paper.

Motivation and example ([R2], [R3], [R4]). We consider distributed OCO over Erdos-Renyi graphs motivated by the following. (1) Distributed optimization and learning over random graphs has been widely studied since it is a validated model in practical communication networks, e.g., for modeling package losses in wireless communications. (2) Erdos-Renyi graphs give us information about complex systems which exist in the real world. The Internet or social networks provide the example at the moment, and it is equally plausible to think about traffic flows, electrical systems or interacting biological processes ([1]). (3) It has been found that Erdos-Renyi random graphs can outperform fully connected graphs in some distributed training tasks ([2]). Moreover, in the dense network, we might want to avoid communicating along each edge per iteration to decrease communication, for which Erdos-Renyi graphs allow a more refined tradeoff between computation and communication costs. (4) Erdos-Renyi graphs allow us to intuitively have a deep theoretical characterization of how the graph topology and connectivity probability influence the OCO network regrets, and paves the way to study distributed OCO over random graphs with a power-law degree distribution by preferential attachment. A prominent motivating example is the distributed online learning through random social interactions for exploiting the streaming but private healthy data generated from wearable personal tracking device ([3]), which also motivates our simulation study with *bodyfat* dataset.

Discussions of theoretical results on factors, such as N ([R1], [R4]), graph topology ([R3]), and p ([R4]). (1) The average regret increases with N, possibly because the increasing node number would increase the nodes' information heterogeneity and make the network regret minimization harder as Reviewer 1 suggested. Compared with the previous work by Lobel and Ozdaglar for distributed optimization over random networks, which provided bounds that grew exponentially in N, we achieved polynomial scaling on network size. (2) The derived regrets showed the inverse dependence on the spectral gap of the expected network, which is quite natural since it is well-known to determine the mixing rates in random walks on graphs, and the information propagation over Erdos-Renyi graphs is closely tied to the random walk on the expected network. (3) It remains open what is the optimal order of dependence on factors N, d, and p in the regrets. (4) We give a simulation study to show how N and d impact the algorithm performance.



Technical challenges ([R1],[R2]) By making use of the 9th page allowed in the camera-ready version, we will elaborate on as many proof outlines as possible in the main body. The proof novelty induced from the Erdos-Renyi graphs lies in showing that the expected consensus error depends on the inverse spectral gap of the expected network.

Simulation Justification([R2],[R4]) We adopt *bodyfat* dataset in numerical studies since distributed OCO over Erdor-Renyi graphs is a practical and preferred learning framework for exploiting personal healthy data from wearable personal digital device or personal tracking device. Concerns on scalability and privacy make the distributed learning a preferable method than centralized ones, while the streaming of the data entails online learning.

Other discussions in simulations [R4] At the beginning steps of Figure 2, there is not enough accumulated data to adapt a good solution in the one-point bandit case, hence the performance can get deteriorated in the first few steps. This is accordant with theoretical results that the time averaged regret goes to zero as T goes to infinity. In Figure 3 and Figure 4, the regret is decreasing with the increment of p, while the decreasing magnitude depends on the graph topology. MReg is the maximal of expected regret of all nodes, hence is taken as a representative performance index in simulations.

[1]: Franceschetti & Meester. Random Networks for Communication: From Statistical Physics to Information Systems. Cambridge University Press. (2008) doi:10.1017/CBO9780511619632.

[2]: Adjodah, Dhaval, et al. "Communication topologies between learning agents in deep reinforcement learning." Arxiv preprint at arXiv:1902.06740.

[3]: Brisimi, Theodora, et al. "Federated learning of predictive models from federated electronic health records." International Journal of Medical Informatics 112 (2018): 59-67.