This paper investigates downlink transmission strategies in a NOMA-based ISAC system and develops beamforming algorithms specifically designed for the max-min SINR balancing and power minimization problems. Overall, the paper presents some new ideas, and the topic is interesting. However, there are several issues that need to be carefully addressed:

1. There is no convergence analysis provided throughout the paper. I suggest that the authors add convergence analysis for each proposed algorithm. In particular, the convergence of Algorithm 1 relies on the ability to optimally solve problem (10), i.e., problem (13) must always yield a rank-one optimal solution. However, the authors have not proved that the SDR is tight. Therefore, whether the convergence of Algorithm 1 can still be guaranteed remains unclear. Please provide a detailed explanation on this point.
2. The description of Algorithm 1 appears to be inaccurate. The feasibility of problem (13) alone does not imply that the current t is feasible for problem (8). Instead, the optimal value of problem (13) must be less than the total power constraint to ensure that the current t is feasible for problem (8). Please carefully review and correct this.
3. Constraint (8b) in problem (8) appears to be problematic. From the reviewer's perspective, the total transmit power constraint at each BS should not include the power directed toward the target, since the beamforming power toward the target originates from the BS’s transmit power. The authors are requested to clarify the physical meaning of constraint (8b) and provide appropriate references to support it.
4. The authors do not analyze the computational complexity of the proposed algorithms. This omission makes the structure of the paper incomplete. I recommend adding a formal computational complexity analysis for each algorithm.
5. For the max-min SINR balancing problem, the quartic transform described in [A] could also be applied. Therefore, what are the advantages of using the bisection search-based method? Does it lead to reduced computational complexity or improved performance? If the performance is improved, the authors are encouraged to compare it in the simulations.

[A] K. Shen and W. Yu, “Fractional Programming for Communication Systems—Part I: Power Control and Beamforming,” IEEE Transactions on Signal Processing, vol. 66, no. 10, pp. 2616–2630, May 2018.

1. The proof in Appendix B seems inaccurate. Specifically, what is the definition of rho? The left-hand side of (B.3) is a zero matrix; why is (B.3) premultiplied by W? Why does the second equality in (B.5) hold? Overall, I suggest the authors carefully review Appendix B and improve the rigor of the presentation.
2. There are several typos and imprecise sentences throughout the paper. For example, in the simulations, the authors claim that Algorithm 1 is optimal and attempt to verify this through numerical simulations. However, optimality cannot be demonstrated solely through simulations; a theoretical proof is also required.