*We again thank **Reviewer #8** for carefully reading our rebuttal and making further inquiries. Please find below our responses to the inquiries. Clarifications will be added to the manuscript if possible.*

Response to Reviewer #8:

COMMENT 1: *What's the definition of RP?*

REPLY 1: RP stands for the Resolution Probability, as mentioned at the end of Sec. 4.1. An RP score indicates how often an algorithm successfully resolves two closely spaced signals, given a number of trials [A]. Let θ_1 and θ_2 be the true DoAs, and $\widehat{\theta}_1$ and $\widehat{\theta}_2$ be the DoA estimates, respectively. In our work, the sources are considered resolved when the absolute DoA errors $|\widehat{\theta}_1 - \theta_1|$ and $|\widehat{\theta}_2 - \theta_2|$ are less than the resolution criterion, i.e., $\min\left(|\theta_1 - \theta_2|/2, 2^\circ\right)$, where $\min(a, b)$ returns the smaller value between a and b. We have indicated the criterion at the end of Sec. 4.2.2. To clarify, we will mention it earlier in the manuscript, e.g., at the end of Sec. 4.1.

[A] Q. T. Zhang, "Probability of resolution of the MUSIC algorithm," IEEE Trans. Signal Process., vol. 43, no. 4, pp. 978-987, 1995.

COMMENT 2: *In Fig. 2, the performance of the Capon-MUSIC-CNN in terms of RP is the worst one. While the ARMSE of that seems good based on Fig. 3. Please clarify.*

REPLY 2: Our experiments show that the performance of Capon-MUSIC-CNN is relatively poor in terms of RP. This can be attributed to the poor resolution performance of the classical Capon-MUSIC itself, especially under low SNR conditions (shown in Fig.1), which is used to generate the labels for the input covariance matrices for Capon-MUSIC-CNN. With our strict resolution criterion, as mentioned in the comment above, Capon-MUSIC-CNN performance is worst in terms of RP as the resolution criterion is not met in low SNR regions. Conversely, in the case of ARMSE, the error between the estimated and true DOAs is computed in every simulation (10^3 in total), regardless of how large or small the individual error may be, as no criteria like RP is used. Thus, in a nutshell, the performance of Capon-MUSIC-CNN degrades significantly for RP due to the use of strict RP criteria by us.

If the reply is hard to read, please find the PDF of this rebuttal with figures to justify the claim on the following link: https://github.com/MurtizaA

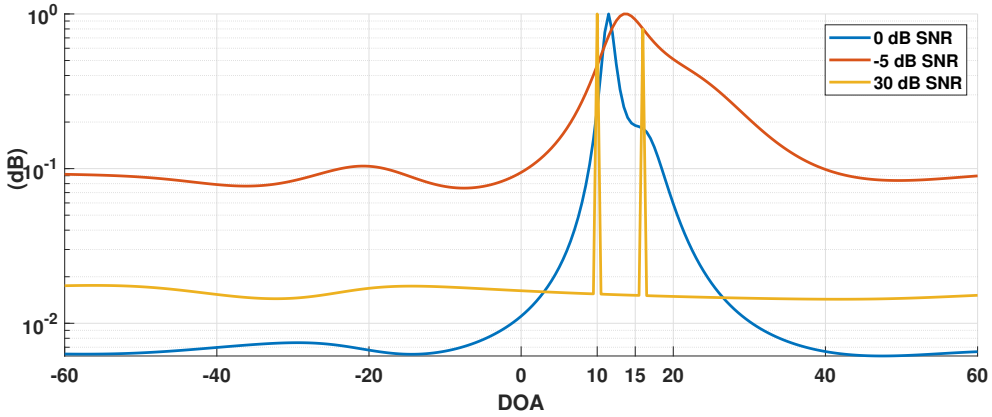

Figure 1: Pseudo-spectrum estimated by classical Capon-MUSIC at $\{-5, 0, 30\}$ dB SNR with sources are located at 10° and 16°, N=6, T=500