G. Androulakis and Tiju Cherian John: Quantum f-divergences via Nussbaum-Szkoła and Applications to f-divergence Inequalities

Referee report

The paper reviews the construction and properties of the Nussbaum-Szkola distributions and their relation to quantum f-divergences, extended to infinite dimensional Hilbert spaces. It is well written and could be very useful, especially the newly added Section 2 which surveys the most important applications of these divergences in quantum information theory. As suggested in Sec. 4, there are many classical inequalities for the f-divergences that can be easily extended to the quantum case using these constructions and that might be of importance.

As a review paper, this work certainly deserves publication. I have only a couple of minor comments listed below.

Some comments

1. p.4, "pairwise linearly independent states" looks strange (a pair of states is linearly dependent only if they coincide). It is more precise to say that any two states have linearly independent supports
2. p.8 and elsewhere: "Total variation". In the classical case, the given expression is equal to the L_1-distance and the total variation distance is 1/2 of this. In the quantum case, however, the given f-divergence is very different from the L_1-distance, which cannot be expressed as any f-divergence (as noticed in Ref. [21], Corollary 7.4). Are there any uses of the divergence $V(\rho\|\sigma)$ in the quantum case? I suggest that the authors include some comments on this.
3. As mentioned in the paper, the Nussbaum-Szkola distributions in infinite dimensions were used also in [12], what is the difference in the present treatment?
4. Theorem 3.8: What is the role of the assumption that the function f is convex or concave? One could use any function, in principle, would the statement of this theorem still hold? On the other hand, in the quantum context, it is usually assumed that the function f is operator convex (or concave), to obtain suitable monotonicity properties.
5. p.12, line 9: one of the $\mathrm{Ker}\,\sigma$'s should be $\mathrm{Ker}\,\rho$.
6. p.13, line 4 of Sec. 4: "same notations" -> the same notations. Also line 9: "the the"
7. p.13: the squared Hellinger distance is mentioned before its definition. Better include a reference to Appendix A.
8. p.17, last line: a - sign seems to be missing
9. p.18, Example 2, displayed equation: D_α should be D_{f_α}
10. p.19, line 2: "of the same" seems redundant.