Answer to Report 1

We thank the Referee for the positive comments on the originality and interest of our proposal. In the following list, we provide answer to the main points raised by the Referee, which led to useful modifications and additions in the manuscript.

> *A brief general introduction to tensor networks and MPS techniques should be included in the manuscript. A more detailed description of the definition of the bond dimension should also be given. In addition, it should be mentioned in the main text that a study on the dependence of the results on this dimension can be found in Appendix C and the main conclusions of this study should also be discussed in the main text (in particular, how to choose chi in practice).*

A brief introduction to tensor networks, mainly focusing on MPS and bond dimension definitions, is now provided in the Introduction, including a simple sketch of a generic MPS, explaining the difference between auxiliary and physical indices. Moreover, in the beginning of the Results section, we now clearly state our choice of bond dimension $\chi = 10$, which is the value used throughout the paper, unless otherwise stated. We also refer the reader to Appendix C for a more detailed discussion of the dependence of our results on the choice of χ, which is related to the entanglement production during the annealing process.

> *I didn't understand why trotterization was used in the ED procedures if it leads to worse results for large time steps. The authors should at least explain why this approximation is considered for comparison with the performances of ED.*

Essentially, trotterization is a necessary step to implement the quantum annealing framework on a gate-based quantum device, which features only a discrete set of native (universal) quantum gates. Usually, digitized Quantum Annealing is defined as the discrete quantum evolution obtained by time-discretization *and* trotterization: here, these approximations are applied both in our ED and MPS simulations. A second-order Trotter approximation was found to yield only negligible improvement on the accuracy of ED simulations (data not shown). We added a clarifying sentence in Sec. 2.2.

> *A comparison of the computation times and maximum system sizes achieved with the implementations of the different methods tested (ED with/without trotterization and MPS) should be included.*

A clear comparison of the maximum system size N achieved by MPS and ED is now included at the beginning of the Results section. In principle, for an ED approach the non-Trotterized dQA (considered in Sec. 3.3) is computationally more expensive than Trotterized dQA (because of the non-sparseness of the time evolution matrix). However, in practice, we were able to reach similar system sizes in both cases (up to $N \simeq 20$). Concerning the computation times, our MPS simulations ($\chi = 10$) are usually much faster than ED: e.g. for $N = 21$, $N_\xi = 17$, $P = 1000$, $T = 1000$ our MPS algorithm ($\chi = 10$) runs in ≈ 10 min, whereas the ED code needs a few hours ($\approx 2 - 3$ h). Since the computational complexity of our MPS simulation scales polynomially in the system size (in contrast with the usual exponential complexity of ED), the computational time gain clearly increases by scaling N up.

I'm a little confused by the Hopfield example. For the capacity given, the equilibrium phase should be the metastable recovery phase. This means that the equilibrium phase of this model should be the spin glass. Why does the QA strategy recover the patterns?

We thank the Referee for pointing out this issue. Indeed, in the thermodynamic limit $N, N_\xi \to \infty, \alpha = N_\xi/N$, the system at $\alpha = 0.13$ belongs to the metastable recovery phase. In the Hopfield example in Figure 20 (panel *a*), we verified that our MPS simulations match the performance of ED. In Figure 20 (panel *b*) we performed a single large-scale simulation for $N = 100$ and $N_\xi = 13$. Here, we had drawn a dashed line (representing the average energy of a state perfectly retrieving a single stored pattern) only as a *reference value*. Indeed, we expected the energy of the true global minimum to be only slightly lower than this reference value. In the new version of the paper, we replace the previous line with a new one, representing an estimate of the exact ground state energy obtained with state-of-the-art classical optimization methods. To get this, we used the online solver http://spinglass.uni-bonn.de/. As expected, this estimate of the ground state energy per site (-1.145) is slightly lower than our original reference value (-1.12).

Finally, in this metastable regime, we observed the expected behavior regarding the overlap with the patterns. Indeed, we checked that the wavefunction $|\psi(s = 1)\rangle$ has almost vanishing overlap with the $\{\boldsymbol{\xi}_\mu\}_1^{N_\xi}$, whereas it is peaked on other classical configurations that are the spurious attractors of the Hopfield landscape.

This is a form one: The authors consider 3 different optimization problems, but only two are discussed in the main text. Throughout the text, however, there's sometimes talk of 2 problems, and sometimes of 3. In the conclusion, only the perceptron and the Hopfield are mentioned. I feel that all 3 problems could well be discussed in the main text, but if the authors prefer to leave the Hopfield problem in the appendix, the text should be revised to be consistent.

We thank the Referee for pointing out this lack of clarity in the previous version of the manuscript. This inconsistency has been fixed. We decided to leave the discussion of the Hopfield model in the Appendix, since the main results are well exemplified by the perceptron. On the other hand, the p-spin model is left in the main text, since it constitutes a useful benchmark on the accuracy of our MPS methods up to large system-size. We revised the main text accordingly, in particular the Conclusion.