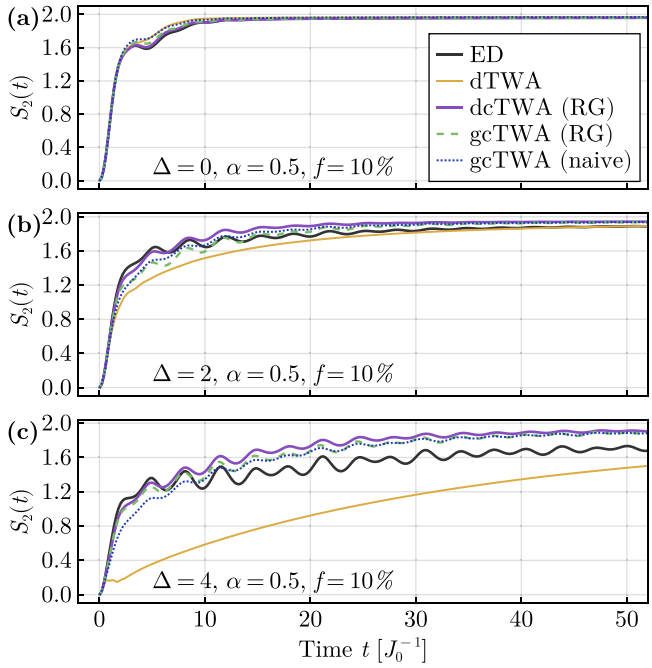

FIG. 10. The average Rényi entropy $\langle S_2(t)\rangle$ is calculated over all possible pairs of two sites. This analysis is performed with the same parameter settings as in Fig. 8.

excellent agreement with the exact results, while the naive clustering gcTWA and dTWA fail to capture the dynamics. This stark contrast to the long-range interacting systems likely originates in the much broader distribution of couplings caused by the much shorter interaction range. Since the RG clustering scheme incorporates the strongest of the relevant couplings, the system can show deviations only at very late timescales.

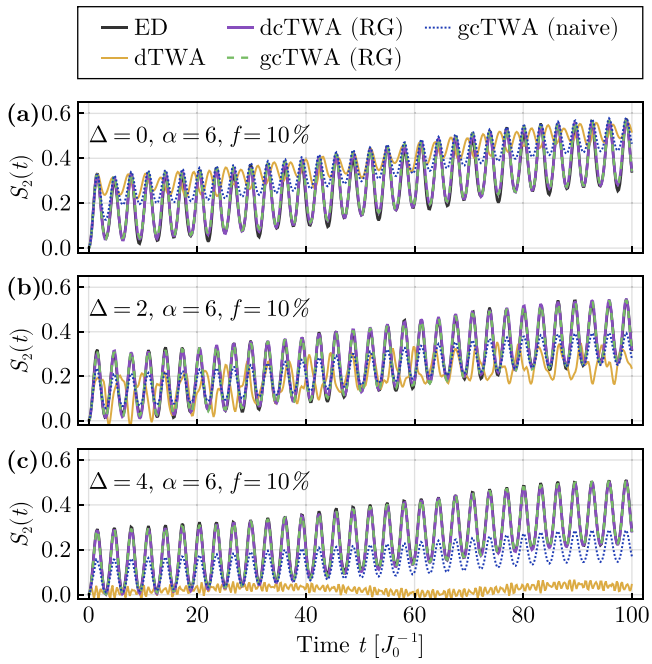

FIG. 11. Same as Fig. 10 but for $\alpha = 6.0$.

In order to understand why dTWA struggles to accurately capture the dynamics of even a single pair of spin interacting via an XXZ Hamiltonian $H = J(\sigma_x^1\sigma_x^2 + \sigma_y^1\sigma_y^2) + \Delta\sigma_z^1\sigma_z^2$ we need to consider its spectrum. Eigenstates of H are the maximally entangled Bell states $|\pm\rangle = (|\uparrow\downarrow\rangle \pm |\downarrow\uparrow\rangle)/\sqrt{2}$ at energies $E_\pm = \pm 2J - \Delta$ and the polarized states $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$ with energy $E_p = \Delta$. Taking the Néel state $|\uparrow\downarrow\rangle = (|+\rangle + |-\rangle)/\sqrt{2}$ as initial state, the exact quantum dynamics only populates the two maximally entangled eigenstates. Since their energetic splitting depends on J only, the exact dynamics is independent of Δ and just encompasses the coherent flipping of both spins $|\uparrow\downarrow\rangle \leftrightarrow |\downarrow\uparrow\rangle$. dTWA essentially has access only to single-body terms and thus needs to approximate this process by two steps which will couple to the polarized states. This gives an intuitive understanding of the dependence on Δ for dTWA. The precise nature of this relation is quite intricate and not akin to, e.g., a two-photon transition. For further analysis of the two-spin case with dTWA, we refer to Appendix B. At this point, we want to remark that even cTWA of course captures the dynamics exactly if the two spins are part of the same cluster, but still the state is not represented exactly at all times. As we have shown in Appendix A, dcTWA can only represent states where the Wigner function is non-negative but one can ensure that all observables within a cluster have correct means and only higher moments deviate.

In summary, we find that even in very long-range systems and for strong Ising interactions the RG-inspired clustering yields quite accurate results at early and intermediate times. At late times, we see some deviations that increase with the strength of the Ising couplings which likely signals the breakdown of the pair approximation in this regime. For short-range interactions, the cTWA methods with RG clustering yield basically exact results in all cases studied here. Conversely, dTWA and gcTWA with the naive clustering strategy struggle due to the competition of the Ising and hopping interactions. We did not see a significant difference between discrete and Gaussian sampling in these settings.

C. Statistical error analysis

To highlight the merits of the discrete sampling scheme, we study the convergence of the staggered magnetization Monte Carlo samples by extracting standard deviation of the staggered magnetization across 10 000 trajectories of a single disorder shot. While previous analyses did not show large differences in result between the sampling schemes, Fig. 12(a) reveals the higher accuracy of the discrete sampling schemes which leads to a reduced number of samples required to achieve a given level of precision. Averaged over the timescale shown, we report approximately 8% smaller standard deviation for dcTWA with cluster size 2 and 15% reduction for cluster size 4. This translates to approximately 16%, and 28% fewer trajectories needed to achieve similar levels of accuracy.

We repeat this analysis for the Renyi entropy, where we estimate the standard deviation from 100 sets of 100 trajectories each [cf. Fig. 12(b)]. Again by averaging, we find a similar reduction of 14% and 29% reduction in standard deviation for cluster sizes 2 and 4, respectively.