Response to referee 2

1. We agree that both references should be included. arXiv:2108.10817 was released on the same day as this manuscript, but arXiv:1403.4788 should have been included in the first place, which we apologize for. We have added both.

2. The point about running widths leading to potential gauge violations is of course completely correct. We would like to keep our remarks about it relatively brief, since this aspect is not by itself a main point of our paper. In the introduction on p4, we have therefore sought to change the relevant paragraph while trying not to lengthen it, as follows:

   > ... as well as options for allowing partial widths (and hence relative branching fractions) to vary with Q^2. The latter allows, for example, to account for kinematic thresholds and effects of running couplings across a reasonable range around the pole, but cannot be pushed too far, especially in the electroweak sector where masses and couplings are not independent of each other.

   One further point to clarify is that the masses of resonances produced by the EW shower are also distributed according to Breit-Wigner distributions with running widths. As is already explained in Appendix C, the gauge violations that lead to dangerous high-energy behaviour are taken care of in the same way as in the calculation of the EW antenna functions. Furthermore, Figure 2 (4 in the new version) shows that at high invariant mass, for the high-energy resonances that the EW shower tends to produce, the shower dominates over the Breit-Wigner distribution, meaning that the exact treatment of running width effects are of little consequence in that region. We have verified that this is true for all SM resonances. A note of this has been added to the resonance matching section.

3. In Vincia, the default cutoff scale for final-state evolution is 0.75 GeV, so the given definitions of Q_{res} are typically above the shower cutoff, for top, Z, and W resonances in the SM (which all have widths larger than 1 GeV). This means that, when one of the dynamical scale choices is selected, most of these resonances will still decay before the cutoff is reached. (Although the Q_{res} distributions are peaked at zero, the integrated probabilities below the shower cutoff are still less than 50%). We added a new figure 1 to highlight this, which shows the surviving fraction of t, Z, and W resonances as functions of the shower scale, for the three different dynamic-scale options. The region of typical shower cutoff values (0.5 - 1 GeV) is also illustrated.

   The reasoning for using a measure of offshellness is given in the introductory paragraph to section 2, specifically strong ordering of propagator virtualities, which dictates which diagrams are enhanced and which are suppressed. The default choice, eq.(4), explicitly represents the denominator structure of eq.(3); the alternatives are mainly provided to make it possible to study the sensitivity to variations on this choice. We have added the following two paragraphs to the beginning of section 2 to elaborate on this:

> The desire to connect with the strong-ordering criterion in the rest of the perturbative evolution, as the principle that should dictate the leading amplitude structures, leads us to prefer a dynamical scale choice for resonance decays, whereby resonances that are highly off shell will persist over shorter intervals in the evolution than ones that are almost on shell. We note that this has the consequence that the on-shell tail will be resolvable by soft photons or gluons, albeit suppressed by the survival fraction. To illustrate this, fig. 1 shows the survivial fractions (denoted Δ_R) as functions of evolution scale, for t, Z, and W resonances, for three different options for dynamical scale choices, all of which are roughly motivated by the propagator structure:
>
> $$i)\quad Q^2_{\mathrm{RES}} \quad = \quad (m - m_0)^2\ , \tag{1}$$
>
> $$ii)\quad Q^2_{\mathrm{RES}} \quad \overset{\text{default}}{\equiv} \quad \left(\frac{m^2 - m_0^2}{m_0}\right)^2 > 0\ , \tag{2}$$
>
> $$iii)\quad Q^2_{\mathrm{RES}} \quad \equiv \quad |m^2 - m_0^2|\ , \tag{3}$$
>
> where m_0 is the pole mass and m its BW-distributed counterpart. Near resonance, options *i)* and *ii)*, illustrated in the left and middle panes of fig. 1, are functionally almost equivalent, differing mainly just by an overall factor 2, while for option *iii)*, illustrated in the rightmost pane, $m = m_0 \pm \Gamma/2$ translates to $Q^2 \sim m_0\Gamma$, so that option is primarily intended to give an upper bound on the effect that interleaving could have.
>
> Alternatively, our model also allows for using a fixed scale, $Q_{\mathrm{RES}} \equiv \Gamma$, irrespective of offshellness. In that case, the resonance will not be resolved at all by any photons or gluons with scales $Q < \Gamma$. We regard this as a good starting point for the width dependence but have not selected it as our default since the fixed-scale choice by itself does not automatically extend strong ordering to the resonance propagators; this can only be achieved by allowing the choice to be dynamical. Our default choice, eq. (2), is constructed to have a *median* scale of $\langle Q_{\mathrm{RES}}\rangle = \Gamma$, while simultaneously respecting strong ordering event by event. This implies that soft quanta will be able to resolve the resonance with a suppressed magnitude $\propto \Delta_R$, which acts as a form factor.

We comment on the $Q_{\mathrm{res}} < Q_{\mathrm{cut}}$ issue under point 4.

4. We hope for the referee's understanding that it is not the intent of this paper, which already presents a significant body of work, to also develop possible non-perturbative aspects of the tail of long-lived (coloured) resonances. In this work, that tail is left to be treated just as it would have been in the conventional non-interleaved framework. For the record, in principle, our view is that yes, top hadrons could be formed from (part of) that tail, but we would probably not simply identify Q_{cut} with the formation of fully formed top hadrons. Rather, one still has to consider a range of scales, between Λ_{QCD} and Q_{cut}, with Λ_{QCD} a more appropriate formation time for actual hadronic states. In the interval one would be dealing with top quarks that have started to build up a confining field, but which decay before it is fully formed. This would be a fun project but not one that we believe we could complete without significantly delaying the publication of the work we have already done. Rather than speculate, we believe we do as much as we can, for now, by pointing out that this aspect in principle exists, but leave it for future studies to consider it more carefully.

5. IF branchings are indeed not included in the current implementation. The reason is that, contrary to QCD, no natural choice for recoiler selection as a result of colour ordering exists. One option to select recoilers is indeed presented in sec. 3.3, and a similar procedure was outlined in arXiv:1611.00788. For now, as we make no attempt to correctly describe coherent EW gauge boson emissions, and the number of included antennae is already very large, we made the choice to limit the shower to FF and II branchings, which is sufficient to include the relevant singular limits. We have added an explanation of this to the manuscript.

6. We agree that this concept is relevant for the construction of parton shower histories in the context of merging. In fact, as described in arXiv:2003.00702, Vincia's QCD shower is now sectorized, which means that only a single shower history path is associated with any particular phase space point. The overlap veto procedure has a similar function, in that it sectorizes the QCD and EW showers. The EW shower is currently not sectorized by itself, but if it were, the overlap veto would ensure only a single shower history, through either the QCD or the EW shower, would be associated with a phase space point. We have added a note of this to the manuscript.

7. The shaded bands are the statistical uncertainties. We added this to the caption of figure 3 (now figure 5), hoping that this provides enough clarity for the other figures.

8. Correct, we have fixed this and appreciate the level of detail the reviewer has inspected the manuscript with.

9. We thoroughly agree that an analysis of the logarithmic accuracy of the EW shower would be very interesting and highly desirable. The EW sector presents some unique features, such as the mentioned spin-dependent subleading logarithms, for which it would be very intersting to find out how the current spin treatment performs. However, the recent work of the PanScales collaboration, the Deductor shower and the work of for instance arXiv:2003.06400 have shown that the analysis of the logarithmic accuracy of a parton shower is not a straightforward task. For instance, a numerical analysis requires running the shower in extreme limits, for which Pythia and Vincia are not currently equipped. We consider such an analysis to be out of the scope of the current paper, and hope to be able to return to this topic in the form of a dedicated study in the future. We have added a brief paragraph at the end of the conclusion describing possible paths for future research.

10. The main difference is that the QED shower does not treat particle helicities, while this is required for the EW shower. While not a major obstacle, the inclusion of spin-dependent antenna functions in the coherent QED shower is not entirely straightworward, for instance due to the possibility of spin flips of massive particles due to photon emissions. We added a brief note of this to the manuscript.