Answer to Referee 4

[...] I believe the manuscript is an important contribution in the field of nonlinear response of materials. The studied effects are relatively unknown, and the level of the developed theory is clearly cutting edge, as it includes many-body excitonic effects that play an important role in 2D materials like the ones studied in the manuscript. Therefore, I recommend the publication in SciPost Physics.

We thank the referee for their evaluation which recognize the significance of our work.

1- For clarity, it would be nice if the authors could choose a different letter from "P", or at least its calligraphy, to denote polarization and momentum, otherwise the derivations are hard to follow at times. This is specially true in sections like 3.1 (page 6); right after Eq. 9, the authors refer to "complex Fourier components $\vec{p}_n$", I assume that they want to refer to the quantity defined in Eq. 9 $\vec{p}^{(n)}$, but the symbol they employ actually corresponds to single-particle momentum introduced above Eq. 6.

To improve clarity, in the present version we have changed the calligraphy for the polarization and changed the Fourier coefficients $p^{(n)}$ to $C^{(n)}$ in the same manner as in Eq. (12).

2- The description on how the FI-SHG is computed at small finite frequency could be improved. According to Eq. 8, it appears that the response at +-2w is proportional to the third-order susceptibility at frequency set(\nu,\omega\omega). Then, I do not entirely understand the following sentence "summing the $\chi(3)$ extracted by P $(2\omega +)$ and P $(2\omega -)$", one obtains the corresponding FI-SHG for low-frequency time-dependent pump fields." Do the authors mean that when \nu is positive (negative) it corresponds to 2w+ (2w-)? I believe this should be clarified.

A further minor point: the authors say "Among higher-order responses that can be extracted from Eq. (4), we look at the FI-SHG", but Eq. 4 explicitly discards higher-than-quadratic contributions. Could they clarify?

We now explicitly show higher order in Eq. 4 (now Eq. 6 in the new manuscript). In our procedure higher terms up to the machine precisions are extracted from the dynamics, though not explicitly used or plotted.

3- Why do the authors employ the approximate equality symbol in Eq. 5? To my understanding, this is the exact expression for the second-order susceptibility in perturbation theory.

The referee is right, we replaced the approximate equality with equality symbol.

4- The authors could provide more details on the symmetries of the two studied systems, e.g. their point group, and possibly provide an appropriate reference containing more details on the systems to help the interested reader.

We added a discussion and references on the symmetries of the two systems and their nonlinear response.

5- The authors rescale the response by a effective thickness d_eff. Is this procedure standard for quadratic responses? Do they expect any non-trivial dependence on this parameter?

For two-dimensional systems, the response is ill-defined. Then, to have a well-defined response function it is customary in experiments to assume an effective thickness, which is usually close to the layer-layer distance in the corresponding bulk system. In calculations, for two-dimensional systems we use a supercell approach and the response output by the calculation then depends on the supercell non-periodic dimension. In complete analogy with what is done in experiment (so making possible to compare with experimental results), instead of the arbitrary supercell non-periodic dimension, we use the effective thickness. The dependence on the thickness parameter is trivial, in the sense that is just a rescaling.