We appreciate the reviewers' feedback. It is encouraging that the reviewers highlighted how the work is effectively able to develop mathematical theory into practical verification algorithms (R1, R2) and the overall writing quality (R1, R2, R3). The main concerns were on presentation to a non-optimization audience (R1), clarifications on how it relates to specific works (R1), and potential variance in numerical results (R2). We address all of these concerns below and also make clarifications regarding comments on novelty (R2) and computational performance (R3).

[R1] **"The high-level ideas are clean and easy to understand for those with a background from [the discrete optimization] community... This paper is very dense and can be hard to read for some."** It is encouraging to hear this positive comment from R1 and we understand the concern regarding a different audience (although all other reviewers made positive comments on clarity). This was kept top of mind during writing but striking this balance is difficult given that we value mathematical rigor. To improve this, we will add a new figure as R1 suggested and revisit the writing, although we prefer to keep most of the math in the main text as we believe they are important contributions.

[R1] **"... paper reads like the algorithm-focused companion of [1] ", "I would like a little discussion of how this work builds on or differs from [1]".** We approached this in the footnote of p.4 and Appendix A (albeit briefly), but the final version will contain an expanded discussion that further clarifies how the manuscript differs in crucial ways from [1]. The central difference is that [1] works in the MIP space (i.e. with binary variables) whereas we work in the original space, with the connection that our formulation is a projection of the LP relaxation in [1]. Our main theoretical result is a minimal description of the new formulation, which can only be attained with a careful and data-dependent analysis of the structure of the ReLU (standard projection methods would lead to large, redundant formulations). This also requires the development of a more involved separation algorithm. Algorithmically, by working without binary variables, we are able to develop our fast propagation-based algorithm FastC2V, and no similar algorithm exists in [1].

[R1] **"how this work ... compares with [the work] by Lu and Pawan Kumar?"** Thanks, we were not aware of this work and we will cite it in the introduction. A detailed comparison is beyond the scope of our work given that we focus on relaxed verifiers (i.e. bounds) rather than branching-based exact verifiers. Nonetheless, our work could be used to directly improve theirs. First, exact verifiers typically use fast relaxed verifiers as a subroutine to produce bounds, and FastC2V can play that role. Second, our inequalities can improve the relaxation used in branching-based methods.

[R1] **"Another thing that would be nice is providing performance profiles ..."** We will include cactus plots (number of images verified per time) in the appendix of the final version (we slightly prefer these over performance profiles).

[R2] **"... there might potentially be a high variance in the numerical results."** We will add an additional section in the appendix investigating variance for networks trained with different initializations. We intend to keep the main experiments as is, so that we can keep the comparison with previous algorithms, and train networks ourselves for these supplementary experiments. As a tiny partial preview, we compared DeepPoly vs our FastC2V for 100 images in a 6x100 network in 10 adversarially trained networks with different random initializations (implementation and parameters are different from those in main text; we may adjust these further), and in all networks our FastC2V verifies more images than DeepPoly, averaging 49.0 vs 67.2 images verified, with standard deviations 3.41 and 4.09 respectively.

[R2] **"The work lacks novelty, as the certification problem has been extensively studied. In my opinion, it is not clear how much more can it be improved, and how relevant is in applications."** We strongly believe that the technical content of the paper is novel within the area of certification (at the very least). We provide for the first time the tightest possible convex relaxation for a ReLU neuron (without binary variables) and a new fast propagation-based algorithm that effectively leverages this new formulation, backed by computational evidence. We push the convex relaxation barrier from [2] and improve upon the kPoly method by [3] (both published in NeurIPS 2019). We believe that the steady abundance of work in the area of certification serves as evidence that it is active and relevant.

[R3] **"Similar performance compared to kPoly".** We would like to clarify that our computational contribution is much broader than this characterization. We highlight that FastC2V can verify more images than the strongest possible convex relaxation defined in [2] (i.e. solving the "triangle relaxation" LP) with ~5-50x faster solve time in our instances. As discussed in [2], this "triangle relaxation" is a barrier that restricts the great majority of verification algorithms. Even if one considers kPoly [3], a state-of-the-art algorithm that does bypass this barrier, we obtain better verification capability with our OptC2V. Finally, mathematical foundation aside, the algorithms themselves are simpler to implement (the pseudocode for FastC2V given in the appendix fits in a few pages) and depend less on hyperparameters than kPoly.

[R3] **Reproducibility.** R3 responded "no" to reproducibility whereas R1, R2 responded "yes". The algorithms are described in the main text and trained networks are publicly available, but please also refer to the appendix for pseudocode and implementation details. Furthermore, we plan to open source the code (as suggested by R1).

References. [1] Anderson et al., "Strong mixed-integer programming formulations for trained neural networks". 2020.
[2] Salman et al., "A convex relaxation barrier to tight robustness verification of neural networks". In NeurIPS 2019.
[3] Singh et al., "Beyond the single neuron convex barrier for neural network certification". In NeurIPS 2019.