We would like to thank all reviewers for their time and effort invested in reviewing our work and for the valuable feedback. We now turn to address each of the reviewers individual comments.

Reviewer #1: Thank you for your comments, for finding our results novel and for considering the question studied in this paper to be “fundamental".

We do not share your feelings regarding the claim that “the potential audience in the NeurIPS community is limited". We believe the ML community is eager for theory that pushes our understanding of such fundamental methods which are highly popular in our community. We give two past concrete examples: recent works on both FW with away-steps over polytopes (Lacoste Julien and Jaggi 2015, Garber and Hazan 2013, 2016) and FW over strongly-convex sets (Garber and Hazan 2015) were theoretical papers which have generated quite notable further research within the ML community. We believe this is due to the simple fact that those works, as we believe this current one also, presented simple yet powerful improvements to our current understanding of this very popular method. Note that there are already very recent works exploring the connections between strict complementarity and more-efficient optimization [2, 1, 3]

It is important to note that there is no practical need to verify the strict complementarity property since we do not present a new algorithm and the algorithm is independent of it.

Regarding your comment “presentation is unusually technical for machine learning venues", we would like to point out that all three reviewers have seem to perfectly understand the setting, the current state-of-affairs and contributions of this work. Nevertheless, we will make an effort to add some more explanations and discussions regarding applications of the results to standard Frank-Wolfe setups.

Additional feedback: 1. We believe the example in Table 2 demonstrates exactly this quite nicely. We can see a standard sparse recovery setup in which the strict complementarity parameter does not change substantially with the dimension, and so the benefit of the new bound over the previous which depends on the dimension is quite clear. We will add an appropriate discussion to clarify and emphasize this.

2. We will comment on the connection of our bound to previous FW lower-bounds.

3. Typos - thank you for catching these!

4. We will positively consider adding a conclusion section.

Reviewer #2: Thank you for you positive feedback and for for finding our results significant.

Regarding experiments, we have included a simple experiment to demonstrate the existence of substantial strict-complementarity in a classical sparse-recovery setting (Table 2 in the paper). This experiment also clearly shows the benefit of the new bound over the previous - the strict complementarity does not change substantially even though the dimension does. We will add an appropriate discussion to make this point clearer. Also, since the algorithm analyzed is not new and has been implemented in many recent papers on various applications, we do not see great importance for additional experiments, as our mission is mainly to better understand its fundamental convergence properties. Please also refer to our answer to Reviewer #1 (line 5).

Thanks you for catching these typos!

Reviewer #3: Thank you for you positive feedback, for your high appreciation of our work and for finding our results significant.

1. We are not sure there is a clear connection between these quantities. The pyramidal is a geometric property of the polytope, while strict complementarity obviously involves also the objective function.

2. Sparse recovery and applications: please see our response to Reviewer #2.

3. This work in only relevant for polytopes. We mentioned low-rank models to give further example of models in which a certain notion of sparsity is desired, beyond simply entry-wise sparsity.

References

[1] Lijun Ding, Jicong Fan, and Madeleine Udell. k fw: A frank-wolfe style algorithm with stronger subproblem oracles. *arXiv preprint arXiv:2006.16142*, 2020.

[2] Lijun Ding, Yingjie Fei, Qiantong Xu, and Chengrun Yang. Spectral frank-wolfe algorithm: Strict complementarity and linear convergence. *arXiv preprint arXiv:2006.01719*, 2020.

[3] Dan Garber. Linear convergence of frank-wolfe for rank-one matrix recovery without strong convexity. *arXiv preprint arXiv:1912.01467*, 2019.