Comment reviewer 1	Response	Manuscript change
At some points the wording and parameters seem a bit away from what the wind energy community uses. E.g. azimuthal rorque (yaw moment), wind strength (ambient wind speed), or full operation up to very high wind speeds (50 m/s) instead of the typical shut down at around 25 m/s.	We come from a slightly different field (therefore we have used a different notation), but agree that the wording should be adapted to match what is common in the wind energy community.	Changed the occurrences of “azimuthal torque” to “yaw torque”, “azimuthal misalignment” to “yaw misalitnment”, and “wind strength” to “wind speed”.
windturbines -> wind turbines	Yes, this should be two words.	Changed the occurrences of “windturbine” to “wind turbine”.
efficiency -> efficiently in line 20.	Agree.	Changed “efficiency” to “efficiently” in line 20.
Regarding line 21: Smaller blades have less inertia, but also less rotor area. The lock number, as an indicator should not change with a smaller blade of same design as a bigger one. Therefore, please check, whether your assumption of better temporal adaptations due to less intertia is true.	After investigating other literature such as https://ntrs.nasa.gov/api/citations/19740008659/downloads/19740008659.pdf and https://www.nrel.gov/docs/fy21osti/80125.pdf regarding the loads, and not finding any comments about how the radius affects the temporal adaptability to the wind we choose to remove the statement.	Changed line 21 to “[...]utilize the spatially varying wind field. Authors such as *citet*{Jamieson2011Innovation, Jamieson2014Structural, Sandhu2018Performance} have investigated and discussed other aspects of rotor scaling and multirotor setups.”
multirotor -> multi rotor	While it seems that both forms are used, a quick search reveals that multi-rotor is preferred in the literature.	Changed all occurrences of “multirotor” to “multi-rotor”.
How will this (furling) technically be achieved? Wouldn't a furling rotor crash into the space frame, carrying the rotor? A visualization/sketch of a furling multi rotor would be good to understand, how this works,	Agree, it should be explained better.	Updated lines 38-40 with: “Furling for a multi-rotor wind turbine can be achieved by rotating the whole support structure to produce a yaw offset as indicated by ψ in Fig. 1This technique will

considering each rotor to be attached to a space frame or other sort of support structure.		be used to illustrate an intriguing operation scheme for when the multi-rotor wind turbine needs to limit its power output.”
Please relate to a reference or two, where MPPT is seen as the typical strategy for wind turbines under partial load. Isn't the typical approach a torque control proportional to the rotor speed? Or are you referring to the yaw control here?	We are referring to the torque control that is proportional to the rotational velocity squared, and yes, it is a bit unclear in the text.	Updated line 47 to be clearer: This is typically achieved by using the well-known Maximum Power Point Tracking (MPPT) controller that dictates a generator torque proportional to the rotational velocity squared, as discussed in *Johnson et al. 2006*.
What is the reason to chose a small turbine, compared to today's turbine size? There are other virtual research turbines available in bigger size up to 20+MW.	We have chosen this turbine as it seems to be an established reference design, hoping that most readers have knowledge of it.	No change.
wind strength -> wind speed	Agree, already fixed.	Already fixed earlier.
In case you mean a force, I would recommend "F" instead of "f".	Ok.	Changed thrust force “f” to “F”.
Lacking specification of losses in line 79.	Yes, this could be explained better, even though it is not essential here.	Added “[...] losses, such as the drivetrain and generator losses, [...]”
If you create your own definition of TSR, then you would also define your optimal TSR values? In line 104.	Yes, good idea.	Added sentence on line 107: “At the optimum, where momentum theory predicts $v=\frac{W \cos{\psi}}{3}$, the modified TSR becomes $3/2$ times the traditional TSR.“
Please clarify, whether you apply your definition of	Agree, and use different symbol/notation for the modified TSR.	Specified that the TSR in line 147/148 is the traditional one, and

TSR or the typical definition here. Line 149.		renamed the modified to \hat{\lambda}.
closes -> closest, line 167	Agree.	Changed “closes” to “closest”.
To save space, figure 7 and 8 could be next to each other.	And some of the other figures could too, but from what I can see the journal uses a double-column format, which will fix the spacing issues.	No change, as this will be fixed with the double-column format.
Fig. 9: Units (m/s?) are missing. With flow you mean wind speed? It would be good to give this a variable name and highlight it in figure 4 for better understanding. Further it would be good to see the same figure for the perfectly aligned rotor.	Agree to lacking units. The net flow means the wind speed after the induced flow has been subtracted. Highlighting this is a good idea.	Added caption: “Mean axial wind speed (W cos ψ − v) in ms−1 with 45◦ yaw misalignment.”
Fig. 10. The difference in power for optimal and DMPPT are very small, whilst the differences in torque are very significant. Can you explain that?	This is an interesting question. The optimal solution creates a more uniform thrust distribution across the rotor array, because this reduces the losses that are aggregated from the upstream rotors. The wake of upstream rotors will affect all downstream rotors, thus reducing the thrust on the upwind-most rotors, one can extract more energy from all downwind rotors. Power is shifted around the array directly, while the thrusts are weighted by their horizontal displacement to form the torque. Thus the small redistribution of power results in small changes in thrust which in the torque are magnified by the horizontal displacement of the rotors.	Added after end of sentence on line 181: “The leveraging of the interactions also has the effect of reducing the yaw moment drastically for intermediate misalignments, as the optimal solution is to reduce the power on the upwind rotors so that the wind has more kinetic energy available for the downwind rotors. This results in a more even thrust distribution, which drastically affects the yaw moment because the thrusts are weighted by the horizontal distances to the center of the wind turbine. ”

Fig. 19. Wouldn't it lead to the same bending moment, but higher total power, when all rotors on top row are equally constrained?	In this case, the rotors must limit their power, and they do this by decreasing their TSRs. When the bending moment constraint is introduced, this it is desirable to operate with the highest possible power per thrust possible. The model predicts that this is achieved by reducing the power as little as possible in an L1 fashion. In other words, turning off one rotor yields a higher total power than if all rotors are slightly limited, because then all rotors operate at a lower power per thrust.	Added before last sentence in line 235: “The relation between power and thrust in the current case, where power is limited by reducing the TSR, enforces an L1 penalty on the system, favoring sparsity rather than reducing the power equally on all rotors.”
Fig. 21. Wouldn't it be more efficient to constrain all three rotors on the far left side?	That would depend on the objective function. As stated in the text, the uppermost left rotor gives the greatest reduction in yaw moment and bending moment, which in this case is preferred. Additionally, similarly to the bending moment constraint comment above, operating at max power per thrust is desired and when only controlling the TSR in a power-constrained case, the L1 sparse allocations are predicted to be better.	No change.
Why is the range of axial flow so high, when typical wind turbines shut down at ambient wind speed of ca 25 m/s?	The range is so large in order to make sure all operating conditions were covered by the model, but admittedly, 50 m/s might be too much. Nevertheless, we do not loose anything by including it.	No change.
Line 298. Please explain, why furling reduces the loads. It reads like a general statement, i.e. also considering dynamic	Furling reduces the thrust loading by reducing the axial component of the wind.	Added “[...] but also the thrust loading, as the axial component of the wind is reduced.”

loads. Dynamic loads might be severe under heavy yaw misalignement. And the typical "slowness" of furling might make it difficult to properly react on gusts.		
pending -> bending	Agree.	Changed “pending” to “bending”.

Comment reviewer 2	Response	Manuscript change
I think this paper reveals and very smartly addresses the fundamentally important aerodynamic effects of rotor interaction in a multi rotor system. This arises when adjacent rotors responding to local wind conditions operating at different loads (thrust coefficients) modify the flow field and create oblique average local flow angles. It also arises fundamentally as with a single rotor because the flow diverges across the system as a whole. If we consider a single actuator disc/rotor, the streamlines diverge increasingly from the center of the disc when the disc is loaded. The inflows are consequently yawed across the surface of the disc . If you imagine the same disc size then filled with multiple smaller discs/rotors, then, even if they are all loaded identically and in uniform far upstream flow, there will be angled inflow to all the rotors except in the center - with ever more complex interactions if	Yes, a more thorough discussion might help the reader. Then we can also highlight that the multi-rotor setup allows one to more efficiently adapt to the flow field.	Added in line 30: “[...] is believed to be significant based on their significance for multi-rotor helicopters Johnson (1994). Additionally, a multi-rotor setup with many smaller rotors will be able to better adapt to the local flow conditions than an equivalently big single rotor system. Thus, a multi-rotor can sample the wind field with greater fidelity than a single rotor system. This sampling gives rise to further interactions that are assumed to be of importance. The present work will include a simplified model [...]”.

the multi rotors are not uniformly loaded. It may help the reader if some discussion equivalent to what is just mentioned is included in the introduction. I think this would be better than the isolated statement (lines 29,30) asserting that the interactions are likely to be important in view of that being the case with multirotor helicopters which can be far different from wind turbines in their operational envelop.		
In Section 3, Modelling, it is understandable to use the NREL 5 MW as an extremely well documented design in the public domain and, as a simplification, to eliminate the overall effect of torque reaction on the structure by counter-rotation of adjacent turbines. At a later stage, the NREL 5 MW would be inadequate for comparisons of multi rotor systems with the largest single rotor systems. It is also unlikely that rotors would be designed for rotation in both directions. Blade production for example would then divide with one half of the population being of opposite hand to the other and with added manufacturing costs implicit.	For large-scale rotors we agree with the idea that rotors with opposing rotational directions will be plausible. However, if one uses high quantities of small rotors the extra cost of mirrored blade production can arguably be neglected in my opinion (without being an expert on the field).	No change.
It is asserted (line 65) that the net rated power of 5MW is equally divided among all the turbines. How is this done ? For example is it on the basis of equal total active swept area? If so, much prior analysis, modelling, wind tunnel test and very limited	The power and swept area is equally divided, and the blockage effect is not included as the inflow model does not include these effects in the present simple parametrization. The blockage effect is an	Changed line 65 to: “[...] the net rated power of 5 MW, and total swept area, is divided equally among all the turbines [...]”. Added after line 67: “The blockage

field testing indicates there is a blockage effect (applicable also to tidal turbine arrays) such that the power of the multi rotors will exceed the power of the equivalent (active area) single rotor. According to prior analyses, theoretical and numerical (CFD, vortex methods etc.) limited wind tunnel testing and minimal field experiments (Vestas) on wind turbines this is due to a blockage effect (recognized also as applying to tidal turbine arrays) which is predicted to be significant (power gains ~ 10% and thus possibly in a range more significant than the differences between independent MPPT power control of the turbines and optimized power control of the array) for a multi rotor system with many closely spaced rotors. I assume this is not accounted in your modelling as I cannot see how training data could be produced except say by extensive CFD analyses of the test case array? I think this needs some discussion or at least an acknowledgement whether the blockage affect is accounted.	interesting phenomenon and I agree that it should be commented on.	effect as evaluated in McTavish et al. (2015) is not included, as this would require a more complex model.”
Vertical variation in the angle of the net flow incident on each rotor is both inherent in the array being finite vertically and that variation is augmented in the case of differential rotor loading caused by wind shear which is later discussed. Regarding the statement (lines 94-95)	The skew angle is unfortunately only taken as an average over the whole array, so the vertical variation is not included. I agree, this should be made clear in the text.	Added after the text on line 95: “The simplified model uses a global skew angle, computed on the average wind speeds of all rotors. While this can be seen as a somewhat crude approximation, it is

and Figure 4 in section 3, is this vertical variation accounted? It would be useful to have text fully clarifying this.		believed to be suitable for the simplified analysis presented in this work. “
In Section 3.2, I would advise against the re-definition of tsr with the usual lambda symbol. Purely for clarity it is better to preserve the established definition and talk about local tsr with some prime/subscript or other modifier on the lambda symbol or some other symbol.	Agree.	Changed the modified TSR to use a hat.
In Section 4.3, line 175, the peak restoring moment is described as equivalent to a 3m upwind movement of the center of thrust. It is hard to interpret the significance of this. Maybe mention the height and width of the array or maybe the restoring moment could be compared to the maximum moment that can be generated with all the turbines on one side at rated thrust and all on the other switched off? The discussion in Section 4.3 about azimuthal stability is nevertheless very interesting and it is encouraging that yaw moments on the misaligned system are restoring.	The array size is seen in the pictures, but yes, the results can be elaborated on to make it clearer.	Modified and added on/after line 175: “[...] 3 m, slightly more than 2% of the multirotor width, upwind. A torque of similar magnitude can be obtained by turning off one of the upper and outermost rotors when the wind-turbine is aligned with the wind.”
In Section 6.1, line 290 to end, I think that "performs identically ……..turbine is more or less aligned" is too modest! Single large turbines would usually be controlled to keep alignment within ranges like ±10 or ±15 deg and maybe shut down as in a fault case if the error exceed 30 deg. Thus based on Fig 28, the SMPPT controller performs identically over the whole range of	Yes, the “more or less aligned” can be substituted for “the main operating conditions” or something along those lines to make the efficacy clearer. The blade design and loading considerations are very interesting, but not the main focus of this work, so we have decided	Changed 291 from “[...] multirotor problem, as long as the turbine is more or less aligned with the wind.” to “[...] multirotor problem for all main operating conditions.” Added at the end of 353: “Other consequences of furling such as

practical interest and this could be said. A reason for avoiding operating at large yaw error, in addition to incurring an obvious drop in power, is that cyclic loads in yaw and stall dynamic effects cause additional blade and system fatigue which may be unacceptable. That thought suggests that it would be interesting to see (as figures or tables) the distributions of resultant wind and skew angle χ over the multi rotor array in the aligned case. It would also be interesting to see what difference results when the aligned case is optimally controlled v MPPT on each turbine. An inference regarding the skewed flows may be that the rotors of a multi rotor system need to be designed for fatigue loading additional to what would apply if the same rotors were operated in isolation.	to not include this here as this would require higher fidelity models. This would be a nice idea for future work!	support structure and blade designs would also benefit from scientific attention.“