Response to Referee #2:
We thank to reviewer for the positive feedback and valuable suggestions which greatly improved the quality of the manuscript.
We addressed the specific suggestions below (our replies in bold).

General comments:

Check nitrite/NO ¯ throughout manuscript 2

We checked for consistency and changed all nitrites to NO_2^-.

You note differences in process rates and between the communities exposed *in situ* to O2 gradients and in the O2 manipulation experiments (e.g. Line 463-4). I think at least some discussion is needed as to the potential effects of purging the samples with gas as described in refs below (e.g. Dalsgaard et al, deBrabandere et al, Holtappels et al, Stewart et al.)

We added a sentence to the method section, line 148-151: "He purging removed dissolved oxygen contamination which is likely introduced during sampling and the headspace prevents possible oxygen leakage from the rubber seals (DeBrabandere et al. 2012)" and in line 158/159: "Total incubation times were adjusted to prevent bottle effects, which become significant after 20 h based on respiration rate measurements (Tiano et al. 2014). "

Furthermore, we added a part in the discussion line 508-521. "Different responses of N_2O production rates to O_2 between *in situ* assemblages and incubation were not unexpected because different rates at different depths were likely not only due to O_2 differences but also other factors such as different organic matter fluxes and different amounts and types of N_2O producers at different depths. In addition, sampling with Niskin bottles and purging can induce stress responses (Stewart et al. 2012) and shift the richness and structure of the microbial community from the *in situ* community (Torres-Beltran et al. 2019), which can be one potential explanation for the different responses between manipulated oxygen and *in situ* oxygen experiments. The removal of other gases like H_2S during purging introduces another potential artefact. However, it is unlikely as measurable H_2S concentrations have mostly been found at very shallow coastal stations (< 100 m deep) (Callbeck et al. 2018), which was not the case in this study. On the contrary, high abundances (up to 12%) of sulfur oxidizing gamma proteobacteria, like SUP05 can be found in eddy-transported offshore waters where they actively contributed to autotrophic denitrification (Calbeck et al. 2018). In this study, it cannot differentiate between autotrophic or organotrophic denitrification, but a contribution of autotrophic denitrification in the eddy center is likely."

Can you be sure that there is no DNRA occurring in your experiments – in particular given the Lam et al. 2009 'Revising the N cycle...' paper also off the Peruvian coast. The presence of DNRA would complicate your isotope pairing experiments with 15NO3- and 15NO2- by transferring 15N into the NH4+ pool and you would get 'hybrid'N2O' of 15nh4+ and 15no2- forming 46N2O and be wrongly assigned. DNRA would also potentially dilute your 15NH4+ pool with 14N from background NO3- and alter the assumed 99% labelling in these experiments. I realise the contribution of AO to N2O production is small relative to denit, but the artefacts of DNRA on the rates/data should be discussed as it could lead to some N2O from AO being 'hidden'.

The reviewer raises a very important point and no, we cannot be sure that the occurrence of DNRA is impacting our results. We added this consideration to the manuscript in line 216 - 232. "Nevertheless, this assumption brings some initial considerations which need to be accounted for. There is a potential for overestimating hybrid N_2O production in $^{15}NO_2^-$ incubations by 5% in samples with high NO_3^- reduction rates. But in incubations from anoxic depths with high NO_3^- reduction rates, no hybrid N_2O production is found at all. For example, accounting for a decrease in f_N of the NO_3^- pool by active NO_2^- oxidation, the process with highest rates (Sun et al. 2017), had an effect of only ± 0.2 % on the final rate estimate. The presence of DNRA complicates ^{15}N-labelling incubations because it can change f in all three tracer experiments. In the vicinity of DNRA in $^{15}NO_3^-$ incubations, $^{15}NO_2^-$ and $^{15}NH_4^+$ can be produced from $^{15}NO_3^-$ which can contribute to $^{46}N_2O$ production by AO. Even when a maximum DNRA rate (20 nM d^{-1} in Lam et al. 2009) is assumed to produce 0.02 nM $^{15}NH_4^+$ in 24 h with all of it being oxidized to N_2O (max. N_2O production from AO 0.16 nM d^{-1}, this study), its contribution to $^{46}N_2O$ production is likely minor and within the standard error of N_2O production rates from NO_3^-. Hence an overestimation of the N_2O production rates is unlikely. The same applies in incubations with ^{15}N-NO_2^- when DNRA produces $^{15}NH_4^+$, additional $^{46}N_2O$ can be produced with a hybrid mechanism by AO not accounted for in the present rate calculations. In $^{15}NO_2^-$ incubations with high starting f (>0.7) the production of $^{14}NO_2^-$ by NO_3^- reduction (which decreases f) leads to an underestimation by max. 9%, whereas in incubations with a low f (<0.3) the effect is less with max. 3 % underestimation of N_2O production rates. In $^{15}NH_4^+$ incubations (f >0.9), max. DNRA rate would lead to an underestimation of 3.5 %. "

Specific comments:

Section 2.1: As with other papers with many sites, sampling points and manipulation experiments a written methods text quickly becomes very complicated with different additions, concentrations, replicates, time points etc. I think as a result of the text being quite confusing some information has been missed/is unclear. Adding a table of experiments, stations, variables, sampling routine (e.g. time points), number of replicates, other factors (e.g. whether O2 was measured in vials) would be informative/helpful to readers who are interested in comparing/replicating experiments.

It is correct, that such set ups can get confusing very quickly, but in table 1 stations, depths, measured variables and the kind of experiment performed are given. However, we added one column with the kind of tracer addition we did. The replicates and time points did not vary between experiments and hence is only stated in the test. We only measured oxygen in one bottle with each incubation per depth or treatment, which was also consistent and written in the text line 168/169.

Also Section 2.1: Missing info on NO3- and NO2- analyses (e.g. shown in Fig 2).

The measurement of nitrite and nitrate concentrations is given in line 130-132.

Line 145 (O2 manipulation experiments): Why was such a 'coarse' O2 range used compared to previous studies which use O2 manipulations generally below 1-2µM (e.g. Dalsgaard et al 2014, Bristow et al 2016)?

In Bristow et al. 2016 a and b the maximal oxygen concentration in their manipulation experiments was 10uM and 20uM dissolved oxygen, so we are not quite sure what the referee means. Dalsgaard et al 2014 performed a really nice microcosm experiment, where oxygen concentrations were monitored online in the flask they subsampled. In our case, each time point was a separate bottle making it impossible to use such an approach. For the experimental design in this study, it was important to choose oxygen levels where we can be sure that oxygen concentrations are different enough from each other that we can differentiate the two treatments (f.e. 100nM and 200nM would be tricky to tell apart with our standard deviations of 180nM and 240nM over 24h). We added a plot of oxygen over time into the supplements Figure S1.

Line 145 (O2 manipulation experiments): This is a bit confusing: '...headspace volume was adjusted depending on the amount of site water added...'. Do you mean that after the addition of different oxygenated water volumes you also wanted to end up with a 3mL headspace as in the 'natural gradient' O2 experiment? Please rephrase and explain more clearly.

Yes, that is exactly what we were trying to do. The sentence was rephrased (line 163 – 164) to "For the O_2 manipulation experiments, all serum bottles were He purged and after the addition of different amounts of air saturated site water a final headspace volume of 3 mL was achieved."

Line 153 (OM experiments): So only total N2O was measured in the OM experiments? Or were 15N substrates also added. Unclear as it is written now.

^{15}N substrates were also added in the organic matter addition experiments. We changed the text to : "For all experiments,.." in line 151 and adjusted line 174 as followed: "200µL of POC solution were added to each serum bottle before $^{15}N\text{-}NO_3^-$ or $^{14}N\text{-}NO_2^-$ tracer injection."

Line 166: Do you mean 'Ascarite' instead of Ascarid?

Yes, we mean "Ascarite" and it was changed.

Line 186-8: Rephrase to: 'If more single labelled N2O is produced than expected (...), a hyrid formation of one nitrogen atom from nh4+ and one from no2- (...) is assumed to be taking place se found in archaeal ammonia oxidizers'

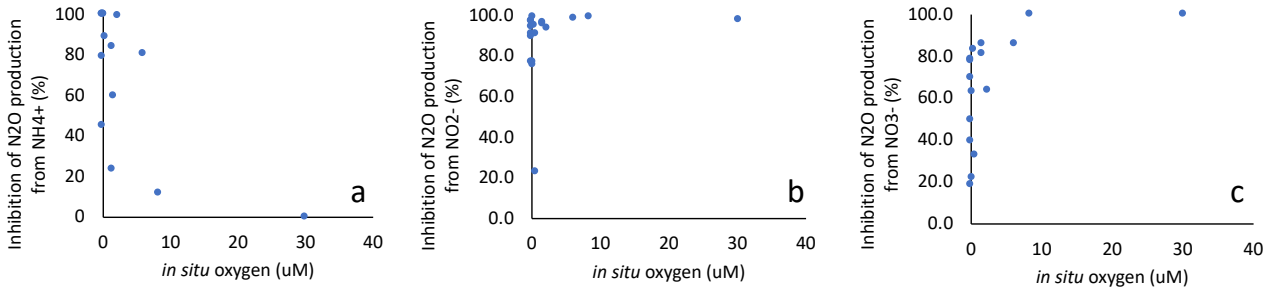
The sentence was rephrased as recommended.

Line 191: What about 45N2O formed from dilution from background 14NO3- and 14NO2- in samples? Then you will get 45N2O from 15NOx + 14NOx ... You note earlier (ca Line 140) that there is likely substantial 14NO3- (at least in some samples/depths) which will be reduced to 14NO2- and dilute your 15NO2- pool. Perhaps there is something I have missed in the text but this doesn't make sense to assume all 45N2O in incubations with 15NO2, especially in anoxic/low O2 manipulations where NOx can be respired.

Indeed, there is a potential for overestimating hybrid N2O production in 15NO2- incubations by 5% in samples with high NO3- reduction rates. But in incubations from anoxic depths with high NO3- reduction rates, no hybrid N2O production is found at

all. We added all potential problems which come with the assumption of a costant f into that section, starting line 217.

Section 3.3: Could the % inhibition of processes be plotted to help comparison to other relevant studies on O2 manipulation on AO/no2- ox/denit (e.g. Kalvelage et al 2011, Dalsgaard et al 2014, Bristow et al 2016). I think at least some short discussion is warranted in relation to O2 effects on processes in these previous papers.



We added the inhibition curves along the natural O2 gradient here, but we stay with the same figures in the main text as we want to show absolute rates. Section 3.3 is a results section, so we do not refer to papers there. However, in the discussion part 4.1 we cited Kalvelage et al. 2011 (line 525) and Dalsgaard et al. 2014 (line 522). Bristow et al. 2016 is added to the section 4.1. Out of these papers, only Dalsgaard et al. 2014 measured N2O production, hence the focus in the discussion is on their paper.

Section 3.4: This is a little confusing, additional 14NH4+ was also added along with the POC to experiments? Or was the POC filtered/rinsed after autoclaving?

The particles were in 0.2um filtered low nutrient seawater and during autoclaving some nitrogen from the particles was liberated into that solution. By adding 200uL of that concentrated POM solution with max. 1.56µM of NH4+ means we added minor amounts of NH4+ into our incubation bottles (0.3nM NH4+), which is neglectable form that perspective. It may be more important with respect to the change in N/C ratios. We adjusted the text, which reads now (line 386 – 389): "The autoclaving of the concentrated POM solution liberated NH_4^+ from the particles, reducing the N/C ratio of the particles compared to non-autoclaved particles (Table 2). The highest NH_4^+ accumulation is found in samples with the largest difference in N/C ratios between autoclaved and non-autoclaved particles (Table 2, 904-20m, 898-100 m)."

Line 466: In relation to the 'Unchanged N2O production with higher O2 levels in NO3- treatments...' sentence: Can anoxic niches be ruled out in these experiments? You do note the sampling being during low upwelling and chl period but the settling of small particulates during experiments may create anoxic/low O2 zones to sustain anaerobic processes.

Anoxic micro niches can never be fully ruled out, if not investigated. The Chlorophyll concentrations were in deed low for an upwelling area, max 5mg/m3, but on average 1mg/m3 and less. Figure 1, map of the study site was adjusted with Chlorophyll concentrations. The treatment was identical between depth profile samples and manipulation samples, so if the particles settle, they would settle in all of the bottles and create microniches in the samples from the depth profile as well. There is no

plausible explanation why more anoxic micro niches should be in the oxygen manipulations compared to the others.

Sentence line 470-472 Bristow et al 2016 should also be a ref here in relation to kinetics of multi-step processes

Bristow et al. 2016a and b were added.

Line 477: How can you be sure none of the N2O was consumed without further measurements (e.g. 15N-N2)? Production may just be much faster than consumption.

We are not able to say anything about N2 production, we can only assume. We added sufficient amounts of 44N2O carrier prior to the incubation to trap 15N-labelled N2O. If N2O reduction is taking place at high rates, we would see a decrease in the N2O pool over time. A plot with the mass 44, 45 and 46 over time was added to the supplements (Figure S2).

Line 512-515: Confusing sentences, consider rephrasing.

Sentence was rephrased (line 565 – 567): " While high N_2O yields are usually found in low O_2 waters (<6 µmol L^{-1}), in this study AO had also high yields at higher oxygen concentrations, 0.9 % at 30 µmol L^{-1} O_2 compared to previous studies (0.06% at > 50 µmol L^{-1} Ji et al. 2018a)."

Line 521: This is a bit of an oversimplification - because something is below detection doesn't necessarily mean nothing is happening, more likely a tight coupling between consumption and production (e.g. see Figure 4 in Klawonn et al 2019 and Figure 3 in Olofsson et al 2019 references). Could there be a dilution of your 15NH4+ pool to consider due to rapid cryptic cycling on shorter scales than your experiments? Ideally 15NH4+ and total NH4+ would be followed through the time series to check for dilution effects. Both show very rapid NH4+ turnover (within ~5h) in oligotrophic waters

The 15NH4+ substrate was not measured on the GC-IRMS because high ^{15}N label/ almost pure tracer is always problematic to analyze. We added the possibility of an overestimation of hybrid production to the method section line 217 and rephrased the wording here to (line 573-575) "Even though, *in situ* NH_4^+ is below detection in almost all water depths (f > 0.9), there remains the potential for $^{15}NH_4^+$ pool dilution by remineralization and DNRA during during the incubation. Despite below detection limit studies have shown fast turnover for NH_4^+ (Klawonn et al. 2019)."

Line 532: If measured, the accumulation and consumption of intermediates (e.g. NO2-) could also be used to imply biotic vs abiotic mechanisms (e.g. Betlach and Tiedje 1981 reference).

We measured NO2- concentrations and isotopic composition in the 15NH4+ treatments, but not other intermediates like NH2OH or NO. The change in concentration was below our detection limit 50nM. Abiotic N2O production was seen in the 15NO2- treatments in the anoxic depth. A supplementary figure was added Figure S9.

Line 560-3: Could a 15N recovery/inventory be calculated for the experiments (e.g. 15N recovery from initial substrate, measured intermediates and 15N-N2O?) This could help infer a % N2O production from denitrification which is important for putting the N2O production from denit in context – i.e. how do variations in O2 impact the proportion of N2O produced by denit relative to N2?

The reviewers make a good point, having both the N2 and N2O production from the same flask at low rates would be very nice. We do not think that there is a way we can come to a N2O /N2 yield without measuring N_2. The biological variations in the NO3- pool were so big that the little change in 15NO3- was too small to be detected. Therefore, the yield of N2O/NO2- was calculated.

Fig 4 b, c & Fig 6 b: consider zoomed-in insert of x-axis (e.g. similar to Fig S5)

Zoom ups are added into the figures.

Figure S5: Seems to be two different slopes here from manipulated vs natural O2 gradients – could also be discussed in relation to purging artefacts.

Both slopes are indicated in figure S7 now. Yes, this could be a purging artefact, but the scatter at the lower range is very high.