Reviewer comments and questions are shown in bold; responses are in plain text.

The paper describes experiments to validate the hypothesis that NOx and HOx are generated spatially separated in lightning. This is an important aspect when trying to consider the impact of lightning induced HOx generation on the global oxidizing capacity. The experiments support the hypothesis, but I find that many details are missing or poorly described. Maybe it is necessary to read the earlier papers to understand, but I feel more information in the manuscript would make the paper much better.

Adding a schematic view of the experiment would be helpful (even if there is a figure in Jenkins 2021 it would be good to have one in this manuscript as well).

A diagram of the experimental setup has been added as Figure S1 to the supplement. The subsequent supplemental figures have been renumbered in the text and supplement, and the following sentence has been added to text at Line 91: "A diagram of the laboratory setup is shown in Figure S1."

Figure S1. (A) Top-down diagram of the laboratory experimental setup showing the key components. (B) Side-view showing a close-up of the GTHOS inlet and Teflon line leading to the NO_x and O_3 analyzers, which sample from the same volume as GTHOS in a 1.3-cm diameter tube place over the GTHOS inlet (shown as two horizontal lines), along with the relative

positions of the flow tube, copper rod, and tungsten electrode. Neither (A) nor (B) is shown to scale.

And even after reading to the end of the paper, I am still wondering what you consider is the size of the core, ie. in your experiments what volume do you consider to contain NOx, and from which volume samples NOx and GTHOS?

We do not have a way to directly measure the core size in our laboratory sparks, but the diameter is no bigger than ~1 mm, based on the diameter of the visible light generated. The sparks are only ~ 7 mm long, based on the gap between the electrodes. If we treat the spark as roughly cylindrical, the volume of the core is ~5.5 mm^3.

The GTHOS and the NO_x and O_3 analyzers sample from the same volume, which is much larger than the spark size. Specifically, GTHOS has a sampling flow rate of ~4 standard liters per minute (slpm) or more depending on pressure, and the Teflon line leading to the NO_x and O_3 analyzers, located ~2 mm downstream of the GTHOS opening, pulls ~13.5 slpm over to the analyzers, from which NO_x and O_3 each sample 1 slpm. So all the analyzers sample from air that contains both the core and surrounding volume mixed together. There is also a short piece of Teflon tubing (1.3 cm diameter) affixed to the front of the GTHOS opening, which the NO_x-O_3 Teflon line points into, to ensure HO_x, NO_x, and O_3 are all sampled from the same volume of air.

When the spark occurs, the NO_x is contained in the very small volume in the center and begins diffusing out. HO_x is generated in the volume surrounding the center. This air containing NO_x in the center and HO_x in the surrounding volume travels together over to the analyzers. As the air travels forward, NO_x diffuses out and reacts with the HO_x generated in the surrounding air. For sparks generated in the same position, the NO_x will diffuse out about the same amount and has the same amount of time to react with HO_x before reaching the analyzers. As we move the spark back, NO_x has more time to diffuse out and more time to react with HO_x, leading to the observed HO_x decay. If the NO_x and HO_x were generated in the same place in the spark then we wouldn't detect the HO_x, because the amount of NO_x generated in the spark would titrate all the HO_x away before we could measure it. Thus, even though we are not sampling from only the core or only the surrounding air, we can still draw conclusions about where the NO_x and HO_x are being made in the spark.

The following has been added at Line 89: "A short piece of Teflon tubing (1.3 cm diameter x 2.5 cm long) was placed on the GTHOS inlet, and the opening of the Teflon tube leading to the NO_x and O_3 analyzers was positioned ~2 mm downstream of the GTHOS opening and facing into the short piece of Teflon tubing. This arrangement

ensured that GTHOS and the NO$_x$ and O$_3$ analyzers all sampled from the same volume."

For what reason did you do pressure dependent experiments? Initially I was thinking that you would look for increased diffusion of NOx out of the core volume with decreasing pressure, but you never draw any conclusions from the results obtained in pressure dependent experiments. Is the only difference the change in rate constants for pressure dependant reaction in MCM with decreasing pressure?

We conducted experiments at different pressures because lightning can occur at any pressure in the troposphere, and we wanted to look for any pressure dependent effects in the chemistry. NO$_x$ will diffuse more as the pressure drops, and some of the reaction rates are pressure dependent, but ultimately, our conclusions were the same for all the pressures we tested in these experiments.

The sentences starting at Line 98 have been modified as follows (modifications are underlined): "Because lightning can occur at any pressure in the troposphere, data were collected at pressures of 970 hPa, 770 hPa, 570 hPa, and 360 hPa (all within ± 2%) to cover most of the tropospheric pressures. Data were also collected at water vapor mixing ratios between 2000-2400 ppmv and temperatures between 289-294K."

Line 287 has also been modified as follows (modification is underlined): "Both the laboratory and model results across all the tested pressures confirm…"

Many details are missing: what is the time-gap between the different sparks, ie. what is the time window for the 10 sparks compared to the total measurement time such as shown in Figure 1?

There is ~75 ms between each spark in the packet, so the entire 10 spark packet is completed within ~675 ms.

The following has been added at Line 81: "…with ~75 ms between each spark in the packet,…"

Figure 1 (and the other figures) is not showing total measurement time. It is showing how much time elapses between generating each spark and the measurement of HO$_x$ from each spark. For example, at 970 hPa, it takes ~64, 120, 210, and 290 ms for the HO$_x$ generated by the sparks to reach the GTHOS inlet from each of the positions, based on the distance from the discharge to the inlet and the 50 standard liter per minute air flow rate in the tube. So if a spark packet is generated at position 3 in the flow tube, the air exposed to the sparks will take 210 ms to reach GTHOS. It also means

the HO_x generated by the spark has 210 ms of reaction time in the flow tube before we measure it.

And at what repetition rate you generate spark packets?

For each of the five positions in the flow tube, we generate four spark packets, with 5 second spacing between each packet. We then pause the sparking while the electrodes are pulled back in the flow tube to a new position. Once the electrodes have been repositioned, we again generate four spark packets with 5 second spacing. This process repeats for the remaining positions.

We have added the following to the manuscript at Line 94: "In each position, four spark packets were generated, with 5 second spacing between each packet."

It would be interesting to see the raw data of the time evolution of OH and HO2 measurements for a spark packet.

We have added Figure S2 to the supplement, which shows the OH and NO signals from the spark packets across each of the five positions.

We have also added to the main text at Line 108: "Figure S2 shows the OH and NO signals from the spark packets over time for one experiment."

Note that although we generate four spark packets in each position, there are signals from only 3 spark packets in the HO_x data. For one of the spark packets, we switch the GTHOS laser to a different wavelength from the one OH absorbs at, to confirm that the OH and HO_2 signals do not contain any electrical interference from the spark.

The following has also been added at Line 94: "For one of the four spark packets, the laser on GTHOS was switched to a wavelength slightly off the OH absorption wavelength to confirm the absence of electrical interference in the OH and HO_2 signals."
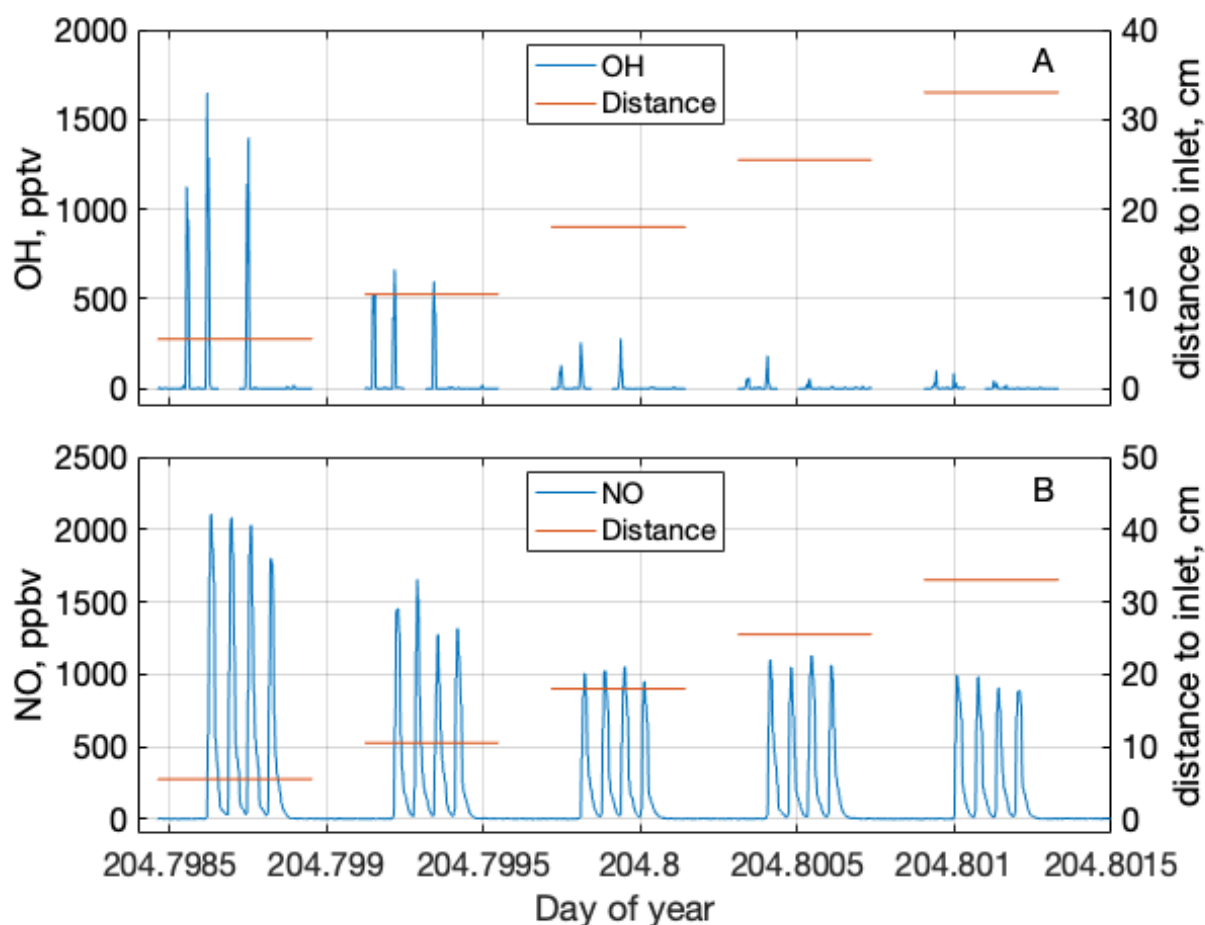
Figure S2. Change in OH (A) and NO (B) mixing ratios due to the spark discharges at each of the five discharge positions at 770 hPa and 0 ppbv of added NO. Each peak is from one spark packet containing 10 sparks. OH and NO mixing ratios are indicated by the blue lines and use the y-axes on the left side of their respective subplots, while the distances from the discharge to the GTHOS inlet and Teflon line leading to the NO_x analyzer are indicated by the orange lines and use the y-axes on the right side.

How the GTHOS measurements have been synchronized with the discharge?

There is no synchronization required. During the experiments GTHOS continuously makes measurements at a rate of 5 Hz, whether the discharge occurs or not. The HO_x signals from the discharges just create spikes in an ongoing measurement.

Over what time window do you integrate the peaks?

For the OH and HO_2 measurements, we integrate the signal from each 10-spark packet over 2.2 seconds. The actual OH and HO_2 peaks are only ~1.2 second wide, but we use

a broader window in the data processing to ensure we capture the whole peak. The HO_x measurements are made rapidly, so the HO_x signals from the spark packet do not spread out much.

For NO and total NO_x, we integrate over 4.8 seconds, which is about the whole width of these peaks. The air travels through a Teflon line and a filter before reaching the NO_x analyzer, which spreads out the NO and NO_x samples along the way. As a result, these signals come out broader than the HO_x signals, and we integrate over a bigger window.

The following has been added starting at Line 109: "For the OH and HO_2 measurements, the peaks were about ~1 second wide and were integrated over 2.2 seconds, while the NO and NO_x peaks were ~4.8 seconds wide and were also integrated over 4.8 seconds."

Over what distance do you move the discharge?

We move the discharge over a total distance of 27.5 cm, starting at 5.5 cm from the GTHOS inlet.

The following has been added to the manuscript at Line 94: "...over a total distance of 27.5 cm."

What is the time resolution of the NOx analyser and how well resolved are the NOx measurements for one spark packet?

The NO_x analyzer collects data at a rate of 2 Hz.

The following sentence has been added at Line 82: "The NO_x analyzer collected data at a rate of 2 Hz and the O_3 analyzer collected data at a rate of 1 Hz."

We cannot distinguish the signals from each individual spark in the 10-spark packet. If we could reliably detect the signal from a single spark on GTHOS and the NO_x and O_3 analyzers, we would just use single sparks instead of the spark packets. The reason for using 10 sparks has to do with generating a HO_x signal that is broad and intense enough for GTHOS to pick up consistently.

Essentially, the actual signal from a 10-spark packet for both HO_x and NO_x is a peak from the spark, following by a dip from the 75 ms gap between sparks, followed by another peak from the next spark, and so on. The air between sparks travels at least ~6.5 cm, which is enough distance so the that chemicals from one spark either do not overlap at all or overlap negligibly with the next. But because the instruments cannot

resolve the signals to this extent, we instead detect HO_x and NO_x signals that are a single peak, the accumulation of the entire 10 spark packet.

You say: " so any change in the NOx mixing ratio across the different positions was assumed to come from diffusion and not chemical loss. ": can you give an order of magnitude of the NOx loss over the different positions, maybe show a figure? At different pressures?

The change in NO_x mixing ratio decreases from the first position to the last position a factor of around 1.8 to 3.4, depending on pressure. This decrease is consistent with diffusion of the core over time.

A new Figure S3 has been added to the supplement showing the average change in NO_x over the different discharge positions, for all four pressures tested in the laboratory experiments.

The following has also been added to the text starting at Line 114: "The average change in NO_x over the different discharge positions in the flow tube is shown in Figure S3 for all four pressures tested."
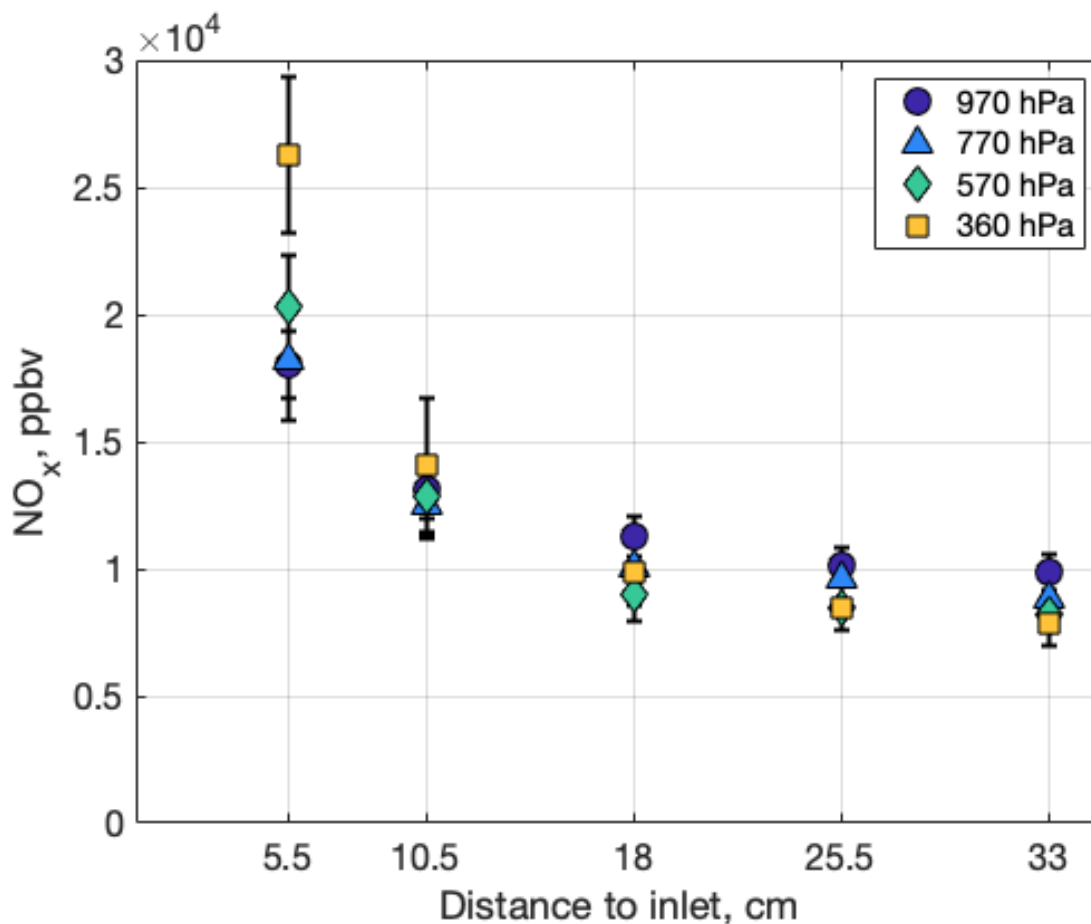
Figure S3. The average NO_x measured for the full 10-spark packet at each position in the flow tube for the experiments with no added NO. Error bars are the standard deviation from averaging the multiple measurements. Distance to the inlet is the distance from the opening of the Teflon tube that leads to the NO_x analyzer to the discharge in the flow tube.

How do you calibrate your HOx measurements?

We pass air at a known flow rate and with a known concentration of water vapor through 185 nm radiation from a mercury lamp. The 185 nm radiation photodissociates the water vapor and generates equal amounts of OH and HO_2 in the air, which is sent directly into the GTHOS inlet. Along with the water vapor concentration, we know the flux from the lamp, the exposure time of the air to the lamp (because we know the flow rate), the quantum yield, and the photolysis cross section of water vapor, so we know exactly how much OH and HO_2 is formed in the calibration system and thus measured by GTHOS. We vary the water vapor concentration to vary the OH and HO_2 signals during the calibration. We also use inlets with different sized openings to create different internal pressures in the GTHOS system and mimic the different pressures GTHOS encounters in the laboratory

experiments. For further details on GTHOS and its calibration procedures, please see Faloona et al., 2004, which is referenced in the manuscript. Because GTHOS has been in regular use for over 20 years now, and the details of its functioning and calibration are well-documented in Faloona et al. and many other publications, we will leave out discussion of the GTHOS calibration procedure in this manuscript. This calibration method is not unique to GTHOS – essentially every other group that measures OH uses some version of it.

What is the HO2 conversion factor within GTHOS?

Because these experiments were conducted without hydrocarbons and therefore there is no possibility of an RO_2 interference in the HO_2 measurement, we set the NO flow to maximize the conversion of HO_2 to OH, which based on the HO_x signals during calibration and is greater than ~95%.

Line 142: what do you means with "The model experiments ran for 0.5 seconds of experiment time"? How long was the experiment time?

Because all our laboratory measurements were made with a total reaction time of less than 0.5 seconds, we had the model experiments in F0AM simulate 0.5 seconds of chemistry. This time is not the model computation time.

This part of Line 142 in the text has been reworded to: "The model experiments were set to simulate 0.5 seconds of reaction time, enough to cover the longest reaction timescale of the laboratory experiments,…"

I understand that the conclusion is, that NOx is generated spatially different from HOx. My question is how do you take this into account when you measure HOx and NOx? Because it should be very sensitive to the position of the sampling points for GTHOS and the NOx analyser? Maybe add a Figure 4 type where you indicate the (estimated) size of the core as well as the precise position of the NOx and HOx sampling. Would it be possible to move the sampling points? Maybe at least for the NOx analyser?

In this experiment, it is not possible to sample the core and the surrounding air individually because, as detailed in a previous answer, the core is much smaller than the sampling volume of the instruments, and really any such air sampling instrument used by atmospheric chemists.

To add to our previous discussion of the sampling, we do align the spark in the center of the flow tube horizontally and vertically, which is also in line with the center of the GTHOS inlet and opening of the Teflon line leading to the NO_x and O_3 analyzers to within ~1 mm. Please see Figure S1b for a diagram showing the relative positions of the inlet, Teflon line, and spark. But the large sampling volume relative to the spark volume means even if it the spark gets a little off center, our results are unaffected. Thus, we are not very sensitive to the sampling positions for the analyzers. Even if we did move the sampling points (which is not possible for GTHOS), it would not affect any of our conclusions.

You correct both species for 15% due to not perfect sampling and diffusion, but I do not clearly understand how you can deduce from the NOx measurement that the correction factor should be the same for the HOx measurements.

We do not deduce from the NO_x measurements that the correction factor should be the same for the HO_x measurements. We deduce the correction factors are about the same between NO_x and HO_x because "OH and HO_2 have similar diffusion coefficients to NO_x" (Line 111). Because HO_x and NO_x have similar diffusion coefficients, we expect them to diffuse at a similar rate in the flow tube, and therefore we can apply the same diffusion corrections to both the HO_x and NO_x data. We cannot measure the HO_x sampling loss directly like we can for NO_x, so we make a reasonable assumption about what is happening with HO_x.

A reference has been added to the sentence starting at Line 111: "OH and HO_2 have similar diffusion coefficients to NO_x (Tang et al., 2014), so OH and HO_2 were also corrected up 15% to account for sampling." This reference has also been added to the reference list: Tang, M. J., Cox, R. A., and Kalberer, M.: Compliation and evaluation of gas phase diffusion coefficients of reactive trace gases in the atmosphere: volume 1. Inorganic compounds, Atmos. Chem. Phys., 14, 9233-9247, doi:10.5194/acp-14-9233-2014, 2014.

When you say " The laboratory air was found to contain ~20 ppbv of CO " I guess you mean the purified air that entered the reactor?

Correct.

This Line has been modified to: "The purified air used in the laboratory experiments was found to contain ~20 ppbv of CO (Thermo Scientific, 48i-TLE)…"

How did you measure and can you safely assume the absence of other trace gases in the purified air? As any remaining VOCs could generate many different

species in the discharge, that could react with OH, it seems to me important to verify the quality of the purified air.

We can safely assume the presence of any trace gases is minimal and negligible to our experiments for the following reasons:

1. In our first paper using this laboratory setup, we measured the OH reactivity of the purified air, and found it be ~1 s^{-1}. This reactivity is equivalent to the reactivity in the relatively clean free troposphere, where most lightning occurs (Jenkins et al., 2021). Since this initial study, there have been some upgrades to our air compressor and dryer that have further decreased this reactivity.

2. This decrease in reactivity was confirmed in a recently published paper where we used the same laboratory setup and purified air described in this paper to conduct a kinetic experiment looking at the OH + HO$_2$ reaction rate. In this kinetics paper, we found the OH reactivity due to impurities in our purified air to be ~0.35 s^{-1} (Brune & Jenkins, 2024). So we know the air we use the experiments is quite clean.

3. Lastly, even if we had a rather large impurity, its impact on the OH decays on the <0.5 second time scales we observe them over in the laboratory would be minimal. For example, if we run a version of the 3% NOx + 3% HONO model run that also includes 20 ppbv of formaldehyde — a large amount of formaldehyde that would only be present in a very polluted city and is certainly not present in our laboratory — the OH decays would not change much, as shown below.
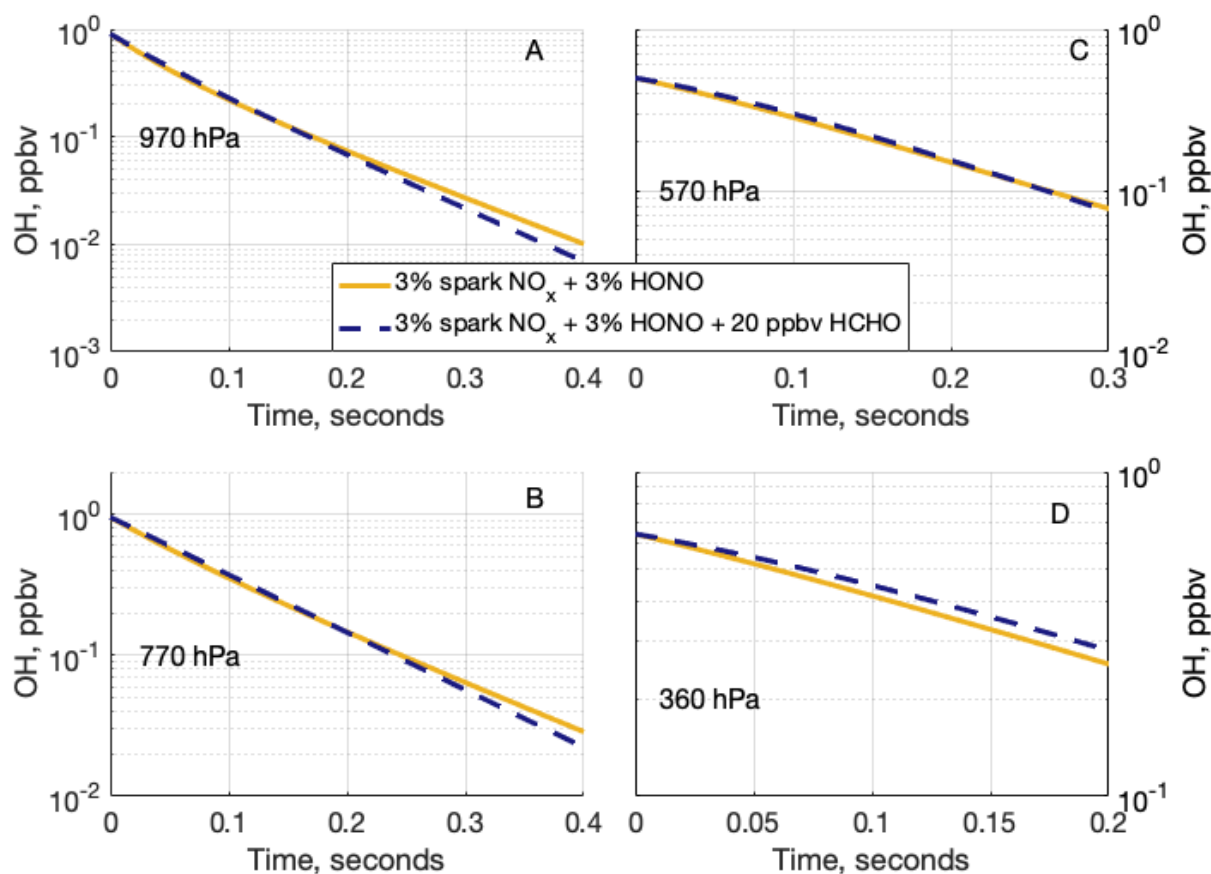
For these reasons, we are confident that the purified air is very clean and there are no impurities interfering without results.

The following has been added at Line 75: "Purified and dried air, with an OH reactivity of ~0.35 s^{-1} (Brune & Jenkins, 2024), was flowed through a bubbler..."

The Brune & Jenkins 2024 reference has also been added to the reference list: Brune, W. H., and Jenkins, J. M.: Is the reaction rate coefficient for OH + HO$_2$ → H$_2$O + O$_2$ dependent on water vapor?, JACS Au, 4, 4921–4926, doi: 10.1021/jacsau.4c00905, 2024.

Figure 1: why are there no measurements with higher NOx at the higher pressure? Any good reason? Because you say "OH and HO2 decays became progressively steeper, as shown Figure 1 (970 hPa and 360 hPa)" and to be really convincing it would have been interesting to see the trace at 970 hPa for 100 ppb NO at least. Did you not do the measurements at higher NO or were there not

enough data points available at higher NO (even though the 50ppb data look good enough to still expect good data also at 100ppb)?

The reason is as we explain at Line 118: "In some experiments, the HO_x decay was fast enough that usable HO_x data was not available at all 5 flow tube positions. If at least 3 positions had clear OH and HO_2 signals, the decay was included in the results; if only 2 positions or less were available, the data were not used in the results, as there was not enough confidence in the extrapolated fit." We did conduct experiments with added NO amounts of 100 ppbv and 250 ppbv at 970 hPa, but we only had good data at two positions and one position, respectively, because the decays were so rapid. So yes, we made the measurements, and we did not include them for the reason we stated at Line 118.

If you set aside the extrapolated data point at time zero, there were good data at four positions in the 0 ppbv case and good data at three positions in the 50 ppbv data at 970 hPa. So it is not surprising that we only had good data at two (one) positions at 100 ppbv (250 ppv) of added NO, which was not enough to fit a curve.

We have modified Line 118 as follows (modifications are underlined): "In some experiments, the HO_x decay was fast enough that <u>the HO_x data became too small and imprecise to use at farther discharge positions in the flow tube.</u>"

Even if it seems clear, maybe explain in the legend that HOx is the sum of OH and HOx, ie. the total signal together with (in the main text) some information on the calibration procedure of the HOx signal.

We define HO_x as the sum of OH and HO_2 in the Introduction, at Lines 28-29 "OH and HO_2 (hydrogen oxides or HO_x)"; we now modify this sentence to the following: "OH and HO_2 (together called the hydrogen oxides or HO_x)" for more clarity. Because we do define the HO_x abbreviation in the Introduction, we will not define it again in the legend. Additionally, for the reasons stated previously, we will not discuss the GTHOS calibration in this manuscript.

Line 161 and 168: I guess you talk about the Figure 2 and not Figure 3?

Yes, these Lines have been corrected to refer to Figure 2 instead of Figure 3.

It would be good to give the absolute concentration of spark NO and not just giving the % so the reader can have an idea of how much this is compared to the added NO.

Lines 188-191 have been modified as follows to include NO_x mixing ratios (modifications are underlined): "Here, adding 3% of the spark NO_x to the model (amounting to 61.6, 62.8, 69.9, and 90.7 ppbv of NO_x at 970, 770, 570, and 360 hPa, respectively) brings agreement within uncertainty to the laboratory HO_2 data, but the OH data are still overestimated by the model (Figure S10). Adding 5% (104, 105, and 117 ppbv at 970, 770, and 570 hPa, respectively) or 10% (303 ppbv at 360 hPa) of LNO_x instead brings measured-modelled agreement for OH, but the HO_2 data are then consistently underpredicted by the model (Figure S11)."

Line 248: The sentence "All the core LOH is also titrated to <1 pptv (our limit of detection in these experiments) over the same time frame the HONO is generated, so it would not be detected by GTHOS in the laboratory experiments, consistent with our observations." is not clear to me: I understand that you cannot make a difference between core and outer LOH in the measurement, so what is consistent with your observations?

It is true that if both core and outer LOH are present in the sample, it would all be mixed together and we would be unable to distinguish how much comes from inside vs. outside the core. But there should not be core LOH present in the sample at all; it should all decay away before we can measure it, leaving only the outer LOH.

In the laboratory experiments, we measure relatively equal amounts of LOH and LHO_2, and we make our first measurement 28-64 ms after the discharge, depending on pressure.

Modelling studies of the lightning hot channel or core have shown that:

1. Extremely large amounts of LOH is made in the lightning core, but the LHO_2 made is orders of magnitude less (Bhetanabhotla et al., 1985; Ripoll et al., 2014).
2. This core LHO_x should decay away within 10 ms (Bhetanabhotla et al., 1985).

Based on these modelling studies, we would not expect any core LHO_x to remain by the time we are able to make our first measurements. Therefore, when we do our model run in F0AM to check the feasibility of forming enough HONO, any core LOH we add to this model run must all decay away in less than 28-64 ms, or before we are able to make an LHO_x measurement in the laboratory. Should the core LOH remain after this timeframe, we would be able to detect it because LOH would be significantly greater than LHO_2, inconsistent with our observations and the lightning model predictions from Ripoll et al. and Bhetanabhotla et al. So the fact that all the core LOH we use in the model run is gone by the time we would start making measurements in the laboratory is consistent with our observations.

This Line has been modified and new sentences have been added as follows: "All the core LOH is also titrated to <1 pptv (our limit of detection in these experiments) over the same time frame the HONO is generated, so it would not be detected by GTHOS in the laboratory experiments. This model result is consistent with our laboratory observations because if substantial core LOH remains beyond the time the first measurement is made in the laboratory, then we would expect to detect significantly more LOH than LHO_2 during the experiments instead of the relatively equal amounts of LOH and LHO_2 that are actually detected. This result is also in line with the Bhetanabhotla et al. (1985) model prediction that all the core LOH should decay away very rapidly."

The following has also been added at Line 243: "Little to no HO_2 is expected to be generated in the hot channel (Bhetanabhotla et al., 1985; Ripoll et al., 2014)."

Line 81 : Should probably be 5 kHz.

5 Hz is correct; GTHOS collects data at a rate of 5 times per second.

The scaling in Figure 3c is missing.

Figure 3 has been revised to include the scaling on subplot C.

Figure S4 has also been revised to correct the marker shapes on subplot A.