We thank Rachel Silvern and Daniel Jacob for their interest in our paper, and for challenging our work with thought-provoking arguments (that we thoroughly enjoy responding to!). Before a point-by-point response to the issues raised, we would like to articulate the main points of our work:

- Including NVCs in the thermodynamic model largely resolves the molar ratio discrepancy, based on our data set, which is representative of the southeast. Only small amounts of NVC are often required, therefore the practice of omitting NVCs from fine mode calculations (which may seem unavoidable for many datasets) induces important biases in molar ratios, which have to be considered in any relevant interpretations (especially on the role of organics).
- The bias in R (ISORROPIA predicted R with Na^+ minus ISORROPIA predicted R without Na^+, where $R = NH_4^+/SO_4^{2-}$, mole/mole) is highly correlated with measured Na^+ ($r^2 = 0.93$), but not correlated with OA mass or OA mass fraction. Furthermore, the difference in observed R from a ratio of 2 (R observed minus 2) is correlated with NVCs and not correlated with OA mass or OA mass fraction. Both results provide strong evidence for NVCs, and not OA, as the underlying driver of the molar ratio discrepancy.

We now proceed with detailed responses (in normal font) to the comments (in italics).

Guo et al. (2017) criticize the results from the recent paper by Silvern et al. (2017) that built on the previous work by Kim et al. (2015). Silvern et al. (2017) found that the ammonium-sulfate molar ratio (R) in aerosol over the Southeast US in summer averages only 1.04 ± 0.21 mol mol-1, despite a large excess of ammonia, whereas thermodynamic models such as ISORROPIA predict that R should approach 2 under these conditions. Silvern et al. (2017) further showed that R steadily decreased during the 2003-2013 period even as SO2 emissions decreased, inconsistent with thermodynamic models.

Steadily decreasing R was reported in the southeast by Hidy et al. (2014) (see Fig 5) and in Weber et al. (2016) (see Fig 3). Weber et al. (2016) provided the first explanation; based on simple thermodynamic principles it could be explained by the contrast in volatility between ammonia/ammonium and sulfate. The Silvern et al. (2017) film is an alternative explanation that accounts for both the R trend and reconciles observed and thermodynamic model-predicted molar ratios. From analysis of our data sets, the latter may be explained by NVCs.

In the first study, Weber et al. (2016) over-predicted R_{SO4} due to limited consideration of NVCs. Despite that, the decreasing trend of R_{SO4} was captured by ISORROPIA (see Fig 2 in Weber et al. (2016)). In this study, we show a good agreement between measured and ISORROPIA-predicted R_{SO4} with inferred Na^+ from a charge balance, which is useful when NVC measurements are not available (Fig 5a). We do note that while not including NVC results in a decreasing R_{SO4} trend, the rate is lower than when NVCs are included (contrast is shown in this paper Fig 5a green vs blue lines). These analyses further suggest the thermodynamic model can explain the observed R_{SO4} trends, with no need for organic films.

1. Both Kim et al. (2015) and Silvern et al. (2017) included observed NVCs in their charge balance ratios and found them to have little effect because NVCs are present in very low concentrations relative to NH4+. Guo et al. (2017) do not acknowledge that both Kim et al. (2015) previously examined the effect of NVCs. The effect of NVCs was shown specifically in Figure 7 of Kim et al. (2015) and in p. 5110 of Silvern et al. (2017). So the inclusion of NVCs in the charge balance calculation is not new. It is also discussed by Pye et al. (ACPD 2017, https://doi.org/10.5194/acp-2017-623). Guo et al. (2017) may consider them 'underappreciated' (viz. the title) but they must invoke 'inferred' NVC concentrations over 4.5 times measured values in order to reconcile R at their SOAS site with thermodynamic calculations (bottom of page 4).

To clarify, the question goes beyond how NVCs affect the measured molar ratio or charge balance, but how they affect the thermodynamic model prediction of R. For example, Kim et al. (2015) Fig 7 only shows measured R that includes NVC, no thermodynamic modeling with NVCs is included (as far as we can tell). Silvern et al. (2017) does include a thermodynamic analysis.

The assertion that much higher NVC concentrations are needed than observed to reconcile R is incorrect. Here we show that when measured NVC > measurement LOD (i.e., reliable NVC concentrations, an admittedly a small fraction of the study period) there is good agreement between measured and ISORROPIA-predicted R (Fig 1b and 1c, and associated text; no difference at a significance level of $\alpha = 0.05$). Issues arise when the NVC concentrations are below or near LOD and they have to be estimated (this is not surprising). Most defensible conclusions can be made for periods with the most reliable data (NVCs above LODs). Assessments and critiques that focus solely on only periods of the least reliable data sets (NVCs below LODs) are weak. However, even periods with NVCs below LODs, often provide consistent results. For example:

- For SOAS, when measured Na^+ was below LOD (0.07 $\mu g\ m^{-3}$), a Na^+ of 0.1-0.3 $\mu g\ m^{-3}$ was needed to bring measured and predicted R into agreement, which is near the detection limit and so difficult to measure. Despite this, the trends are the same and the predicted R is generally smaller than the observed, and so including the estimated LVCs in this case over compensated.
- For WINTER data, inferred NVC results in agreement in R prediction with the observation, but inferred Na^+ is smaller than observed PM_1 Na^+, 0.15 vs. 0.23 $\mu g\ m^{-3}$, stated in the caption of Fig 4. Again, the results are uncertain since measurements of NVC were not measured online in that study.
- For the case (we guess) where the 4.5 factor is referred to, inferred $Na^+ = 0.28 \pm 0.18\ \mu g\ m^{-3}$ vs. measured $Na^+ = 0.06 \pm 0.18\ \mu g\ m^{-3}$. This does cause an under-prediction of R (predicted $R = 1.43 \pm 0.32$ vs. measured $R = 1.70 \pm 0.23$ (± SD of the data)). T-test shows a statistical difference, but because the results are not based on measured NVCs we are reluctant to make inferences on the cause of the discrepancy.

Our main finding is that the NVCs, even at low levels, have significant impact on predicted molar ratios (i.e., predicted ammonium concentration), in contrast to findings of Kim et al. (2015), Silvern et al. (2017), and Pye et al. (2017). If our findings are correct, 'underappreciated' is appropriate. We will cite the above papers with their findings in the revised draft.

Why the different conclusions? Maybe use of different (and possibly inconsistent) data sets? We note that NVC measurement limitations (measurement uncertainties) could be a cause. Also, we have used what we feel is one of the best data sets, which allows more precise, time-resolved analysis and hence more robust conclusions: Online PILS-IC for anions and cations, CIMS ammonia in our analysis, versus Kim et al. (2015), Silvern et al. (2017), or Pye et al. (2017) that use a combination of AMS data that does not include NVCs and CSN 24-hr filter $PM_{2.5}$ data which does include NVCs, but have known negative NH_4^+ biases (R will be biased low).

All of our results can be examined as we have provided the ISORROPIA-II input files in the supplement.

2. The 'inferred Na+' from Guo et al. (2017) seems like a self-fulfilling prophecy. Here Na+ (taken as a proxy for NVCs) is specified not from observations but from a charge balance 2[SO42-]+[NO3-]-[NH4+] excluding [H+]. Of course in that case the predicted R will match the observed values but that is nothing more than fitting NVC to match the observed R. The problem is that these fitted values of NVCs are much higher than observed. Guo et al. [2017] make the case that [H+] in the charge balance must be small but this is based on inferred Na+ so the argument seems circular.

We understand the point raised, but do not agree; the "self-fulfilling prophecy" would occur if we assumed NH_4 and NO_3 were not allowed to equilibrate with the (observed) gas phase and volatilize, or if the pH of the computed aerosol were very close to neutral. Neither happens, as ions volatilize and the aerosol becomes acidic (with a pH between 0 and 2 in this study). Therefore, we stand by our suggestion that the NVC derived from a charge balance to the data poses an upper limit for NVCs.

3. Guo et al. (2017) start from a value R = 1.70 mol mol-1 for a select time range at their SOAS site, and it does not take much correction to bring this value to approaching 2 through inclusion of NVCs. However, the CSN observations across the Southeast and the AMS observations from SOAS reported by Silvern et al. (2017) indicate much lower mean values of ***1.04 ± 0.21*** *mol mol-1 and 0.93 ± 0.29 mol mol-1 respectively. This is not acknowledged by Guo et al. (2017). Silvern et al. (2017) noted the discrepancy with the higher values of R observed at the SEARCH sites and analyzed the differences in the measurement protocols, concluding that the R values could be biased low in the CSN data and biased high in the SEARCH data. Weber et al. (2016) reported a value of R = 1.54 for one representative SEARCH site in summer 2013 and Guo et al. (2015) reported R = 1.4 for SOAS. Pye et al. (2017) also pointed out this large difference in R between data sets. Guo et al. (2017) should acknowledge this difference in observed values of R between datasets because it has major implications for their assertion that NVCs can solve the problem.*

The inclusion of NVCs moves the predicted R away from 2 into agreement with measured R. The selected 13-days period of the SOAS study is representative of the SOAS PILS-IC $PM_{2.5}$ data as it is 13 days of the 17 days (76%) of the $PM_{2.5}$ data. The first 4 days of the study in which $PM_{2.5}$ data is available, 06/07-06/10, were more affected by precipitation and so not included, but still gave a good prediction of NH_3, as shown by Fig 10 in Guo et al. (2015).

We did not selectively pick data that had R values close to 2. The measured R for the selected period (1.70 ± 0.23) is representative of the whole SOAS campaign (1.60 ± 0.23 for all of SOAS by MARGA measurement (Allen et al., 2015), the other $PM_{2.5}$ R data = 1.90 ± 1.01 (all points) and 1.68 ± 0.47 (excluding R over 3, 8% points) also for all of SOAS by SEARCH $PM_{2.5}$ online measurement) (± is the SD of the data). The PILS-IC measured R during the SOAS study falls right in the range of the long-term SEARCH observations at CTR; in 2013, summer mean R and SD was 1.63 ± 0.16. Hidy et al. (2014) plot R for all SEARCH sites over many years (Fig 5). For all SEARCH sites in the southeast, Silvern et al. (2017) give 1.62 ± 0.17. Our smaller SOAS data set is consistent with all these data. The CSN data, which Silvern et al. (2017) also used and reports at an R of 1.44 ± 0.34 (the 1.04 ± 0.21 in the comment is misquoted; that is for CSN R_{SO4} value), is lower due to the known low NH_4^+ bias due to CSN use of nylon filters. In our analysis, we have focused what we believed are the best data (PILS vs. Filter). Finally, we also have shown that a relatively small change in NVC can lead to substantial change in predicted R. For example, Fig 3a shows that the change in R is quite sensitive to concentrations of NVCs; the slope in ΔR vs Na^+ concentration (in $\mu g\ m^{-3}$) is near 2.

We agree that there are differences in R observations (ignoring CSN due to a known bias). A significant fraction of the differences in R between data sets can be attributed to the differences in measured particle size ranges (in contrast to what Pye et al. (2017) finds). A unique feature of our SOAS data is that we have both PM_1 and $PM_{2.5}$ results measured with the same instrumentation (only cyclone inlet changed, $PM_{2.5}$ PILS-IC data for the 1st half and PM_1 for the 2nd half). PM_1 vs $PM_{2.5}$ has an effect on measured R. For example, see the wide variation in SOAS measured R in the figure below. For the 1st half, the PILS-IC and MARGA $PM_{2.5}$ R tend to agree, AMS PM_1 R is much lower, for the 2nd half there is much less difference between AMS and PILS-IC PM_1 R compared to the larger difference between these PM_1 and MARGA $PM_{2.5}$ R. Overall, the figure shows an increase in R from PM_1 to $PM_{2.5}$ measurements with minor differences possibly due to instrumentation issues. (Aside: The comparison of AMS vs PILS-IC for the whole SOAS study can be found at Guo et al. (2015) in Fig 2. MARGA sulfate and ammonium are also highly correlated to and close to PILS-IC; ODR fits, MARGA SO_4^{2-} = (1.00 ± 0.01) PILS-IC SO_4^{2-} + (0.51 ± 0.02 $\mu g\ m^{-3}$), $r^2 = 0.96$; MARGA NH_4^+ = (1.04 ± 0.01) PILS-IC NH_4^+ + (0.21 ± 0.01 $\mu g\ m^{-3}$), $r^2 = 0.91$, see our comment to Pye et al. (2017) for a plot.)

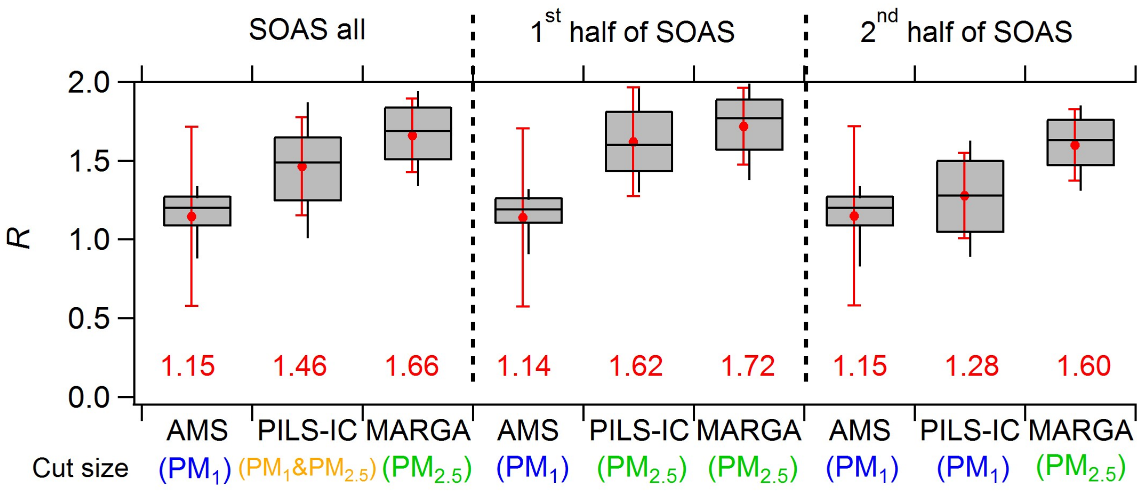

Figure I. Comparison of measured R by AMS, PILS-IC, and MARGA for (left) the entire SOAS campaign, (center) the 1st half of SOAS, and (right) the 2nd half of SOAS. PILS-IC measured $PM_{2.5}$ in the 1st half and switched to PM_1 in the 2nd half of the study, while AMS measured PM_1 and MARGA measured $PM_{2.5}$ for the entire study. For a consistent comparison, we excluded the data if any R from the three instruments is not available (therefore, the statistics may be slightly different from the above in the text). 10% and 90% percentiles are shown as vertical black sticks; 25%, 50% (median) and 75% percentiles are shown as horizontal black sticks and grey shades. Means are marked by solid red circles and numbers are listed above each instrument. Red error bars indicate propagated measurement uncertainties calculated with quadrature sum of squares based on measurement uncertainties of AMS 35%, PILS-IC 15%, and MARGA 10% (assumed based on Rumsey et al. (2014).

4. The most puzzling and interesting behavior of R noted by Silvern et al. (2017) is its decrease over the 2003-2013 period as SO2 emissions decrease while ammonia emissions stay flat. Guo et al. (2017) need to acknowledge this discussion of the trend by Silvern et al (2017). The top panel of Figure 5 in Guo et al. (2017) shows that ISORROPIA cannot properly predict the trend, supporting the argument made by Silvern et al. (2017) that observed trends show a departure from equilibrium behavior. In order to capture observed trends, Guo et al. (2017) require n an increasing concentration of NVCs with time (an increasing 'inferred Na+') that is inconsistent with observed concentrations at SEARCH sites.

We agree this is interesting; this was also noted by Weber et al. (2016) together with its explanation based on thermodynamic principles. We do not require increasing NVCs with time; in fact we show (Figure 5c) that Na^+ *decreases* over time (with a linear trend of -0.030 ± 0.009 μg m^{-3} yr^{-1}, $r^2 = 0.49$, consistent with the SEARCH observations). What is happening is that SO_4^{2-} is decreasing more rapidly than Na^+, so that in relative terms, the latter (and NVCs overall) become an increasingly important driver of the aerosol thermodynamic state (slope = 0.5% yr^{-1}, $r^2 = 0.32$).

5. They suggest that Silvern et al. (2017) did not think of NVCs. This is apparent in the title "The underappreciated role. . ." and the fact that nowhere in the paper is it acknowledged that Silvern et al. (2017) and previously Kim et al. (2015) quantified the effect of NVCs on R – see comment 1 above.

The importance of NVC in aerosol thermodynamics was known for decades (that is why ISORROPIA-II was developed to begin with!), and we will make this very clear in the revised version. However, effects of NVC in the fine mode (especially in the size ranges 1.0-2.5 μm) *is underappreciated*, because they are assumed unimportant or not constrained well enough by measurements. By having access to arguably the most precise, comprehensive and high resolution (in time) datasets allows us to test and resolve this issue (vs. using the $PM_{2.5}$ datasets), because pH (and R) can drastically change between PM_1 and $PM_{2.5}$, owing to the effects of NVC (Fang et al., 2017; Guo et al., 2017).

6. They suggest (specifically in the abstract) that Silvern et al. (2017) is solely about the organic film hypothesis with claims that retardation of equilibrium would somehow not apply to water vapor or nitric acid. In fact, Silvern et al. (2017) only presented the organic film as a tentative hypothesis to explain the decrease in R over the 2003-2013 period (see their Abstract). Also, they did not claim that it would not retard equilibrium for H2O and HNO3, and instead simply discussed the implications for H2O and HNO3 as an open question (see their last paragraph).

We did not mean to imply that Silvern et al. (2017) is *only* about the film hypothesis. However, the film hypothesis is one of the few mechanisms suggested that could be tested. We support that films are not likely because: *i*) the difference in measured *R* from 2 is anticorrelated with organics altogether (with implications for the role of organics that go beyond just acting as a kinetic barrier), but correlated with NVC (or Na^+) concentrations, and, *ii*) it is hard to justify organic films selectively inhibiting NH_3 from reaching equilibrium, but allow equally diffusive and sticky H_2O and HNO_3 molecules to "pass through" establish gas-particle equilibrium. Furthermore, Weber et al. (2016) provided the basis for explaining the *R* trend, based on the thermodynamics of the inorganics alone.

7. They imply that Silvern et al. (2017) do not understand the basics of H2SO4- NH3 thermodynamics and the implications for pH, but this seems based on inadequate reading of the paper. Contrary to what Guo et al. (2017) state (see for example page 5, lines 20-22), Silvern et al. (2017) did not claim that thermodynamics imply complete neutralization of sufate by ammonia resulting in elevated pH. In fact Silvern et al. (2017) took pains to show that the thermodynamic argument is not about pH but about R. See their Figure 1 which shows a pH value below 2 even when total ammonia is in large excess. Guo et al. (2017) claim that Silvern et al. (2017) are using R as "acidity proxy" to get at pH (page 2, lines 9-12) but Silvern et al. (2017) in fact emphasize that their paper is not about pH. The Guo et al. (2017) paper suggests that the controversy is about pH (page 8, lines 5-7) but it is not; it's about R.

"Incomplete neutralization" was the focus of Silvern et al. (2017), and central to that is the thermodynamic state of the system. pH is an important expression of the state, and governs the partitioning of ammonia and nitric acid onto the aerosol. Therefore, thermodynamics (and pH) controls the aerosol NH_4^+ (and if relevant, NO_3^-) concentration and *R* (= NH_4^+/SO_4^{2-}). To state that "it is about *R* and not about pH" is therefore incorrect or at least incomplete, as *R* is decoupled from aerosol acidity only when liquid water is absent. We will make sure that these points are clarified in the revised manuscript.

References:

Allen, H. M., Draper, D. C., Ayres, B. R., Ault, A., Bondy, A., Takahama, S., Modini, R. L., Baumann, K., Edgerton, E., Knote, C., Laskin, A., Wang, B., and Fry, J. L.: Influence of crustal dust and sea spray supermicron particle concentrations and acidity on inorganic NO_3^- aerosol during the 2013 Southern Oxidant and Aerosol Study, Atmospheric Chemistry and Physics, 15, 10669-10685, doi: 10.5194/acp-15-10669-2015, 2015.

Fang, T., Guo, H., Zeng, L., Verma, V., Nenes, A., and Weber, R. J.: Highly Acidic Ambient Particles, Soluble Metals, and Oxidative Potential: A Link between Sulfate and Aerosol Toxicity, Environmental science & technology, 51, 2611-2620, doi: 10.1021/acs.est.6b06151, 2017.

Guo, H., Xu, L., Bougiatioti, A., Cerully, K. M., Capps, S. L., Hite, J. R., Carlton, A. G., Lee, S. H., Bergin, M. H., Ng, N. L., Nenes, A., and Weber, R. J.: Fine-particle water and pH in the southeastern United States, Atmospheric Chemistry and Physics, 15, 5211-5228, doi: 10.5194/acp-15-5211-2015, 2015.

Guo, H., Liu, J., Froyd, K. D., Roberts, J. M., Veres, P. R., Hayes, P. L., Jimenez, J. L., Nenes, A., and Weber, R. J.: Fine particle pH and gas–particle phase partitioning of inorganic species in Pasadena, California, during the 2010 CalNex campaign, Atmospheric Chemistry and Physics, 17, 5703-5719, doi: 10.5194/acp-17-5703-2017, 2017.

Hidy, G. M., Blanchard, C. L., Baumann, K., Edgerton, E., Tanenbaum, S., Shaw, S., Knipping, E., Tombach, I., Jansen, J., and Walters, J.: Chemical climatology of the southeastern United States, 1999-2013, Atmospheric Chemistry and Physics, 14, 11893-11914, doi: 10.5194/acp-14-11893-2014, 2014.

Kim, P. S., Jacob, D. J., Fisher, J. A., Travis, K., Yu, K., Zhu, L., Yantosca, R. M., Sulprizio, M. P., Jimenez, J. L., Campuzano-Jost, P., Froyd, K. D., Liao, J., Hair, J. W., Fenn, M. A., Butler, C. F., Wagner, N. L., Gordon, T. D., Welti, A., Wennberg, P. O., Crounse, J. D., St Clair, J. M., Teng, A. P., Millet, D. B., Schwarz, J. P., Markovic, M. Z., and Perring, A. E.: Sources, seasonality, and trends of southeast US aerosol: an integrated analysis of surface, aircraft, and satellite observations with the GEOS-Chem chemical transport model, Atmospheric Chemistry and Physics, 15, 10411-10433, doi: 10.5194/acp-15-10411-2015, 2015.

Pye, H. O. T., Zuend, A., Fry, J. L., Isaacman-VanWertz, G., Capps, S. L., Appel, K. W., Foroutan, H., Xu, L., Ng, N. L., and Goldstein, A. H.: Coupling of organic and inorganic aerosol systems and the effect on gas-particle partitioning in the southeastern United States, Atmospheric Chemistry and Physics Discussions, 1-25, doi: 10.5194/acp-2017-623, 2017.

Rumsey, I. C., Cowen, K. A., Walker, J. T., Kelly, T. J., Hanft, E. A., Mishoe, K., Rogers, C., Proost, R., Beachley, G. M., Lear, G., Frelink, T., and Otjes, R. P.: An assessment of the performance of the Monitor for AeRosols and GAses in ambient air (MARGA): a semi-continuous method for soluble compounds, Atmospheric Chemistry and Physics, 14, 5639-5658, doi: 10.5194/acp-14-5639-2014, 2014.

Silvern, R. F., Jacob, D. J., Kim, P. S., Marais, E. A., Turner, J. R., Campuzano-Jost, P., and Jimenez, J. L.: Inconsistency of ammonium–sulfate aerosol ratios with thermodynamic models in the eastern US: a possible role of organic aerosol, Atmospheric Chemistry and Physics, 17, 5107-5118, doi: 10.5194/acp-17-5107-2017, 2017.

Weber, R. J., Guo, H., Russell, A. G., and Nenes, A.: High aerosol acidity despite declining atmospheric sulfate concentrations over the past 15 years, Nature Geoscience, 9, 282-285, doi: 10.1038/ngeo2665, 2016.