

## Response to Reviewer #1

In this manuscript, the authors present a new framework with which to identify and characterize atmospherically relevant organic compounds based on a combination of ion mobility and molecular mass. Though a wide variety of such two-dimensional frameworks have been employed to parameterize and simplify descriptions of atmospheric mixtures, the technique proposed by the authors is unique and valuable in its ability to characterize compounds based on structural features and to separate isomeric species. The authors have done an excellent and detailed job of exploring intrinsic spatial relationships in this parameter space. Continued application of the tools and techniques described in this manuscript will likely provide more molecular and chemical information than has previously been available.

*We thank reviewer #1 for the constructive and insightful comments. Our point-by-point responses can be found below, with reviewer comments in **black**, our responses in **blue**, alongside the relevant revisions to the manuscript in **red**.*

### General comments:

While this manuscript builds a strong foundation for the application of these techniques to atmospheric samples, no attempt is made to apply these techniques to complex mixtures of unknowns. The title, abstract, and some portions of the introduction should be re-framed to highlight what is actually in this manuscript and focus less on what the authors hope to do with these tools in the future. It is implied or, in the case of the title stated explicitly, that this paper is about the “Characterization of Atmospheric Organic Aerosol.” Given the home institutions of the authors, I have no doubt this is the goal and am excited to see it applied to ambient mixtures. However, without more detailed attempts to apply this technique to atmospheric mixtures or at least detailed discussion, the language and title of this manuscript should be changed to focus more on “atmospherically relevant organic compounds,” or “highly oxidized small organic compounds,” or “characterizing functionality of organic compounds.” The detailed description of the framework also needs some added clarity – see detailed comments below.

**[Responses]** The reviewer has raised a good point. As we have not applied this framework to deconvolute the complexity of atmospheric aerosol mixture, the focus of ‘characterizing organic aerosols’ becomes irrelevant to the scope of the current study. On the other hand, this framework is not limited to the condensed phase speciation but can be applied to chemicals in the environment at all forms, with the use of appropriate ionization schemes. In fact, developing

alternative ionization schemes that are suitable for gas-phase measurement and also compatible with IMS drift tube is one of our current research focuses. As suggested, we have changed the language of the entire text to focus on the atmospherically relevant organic compounds, as all the functional groups characterized in the study have been identified as major components in the atmosphere.

We would also like to note that progresses have been made in applying this framework to ambient organic aerosol mixtures. In fact, the application of IMS-MS to a rather simple SOA system that is generated from the reactive uptake of IEPOX onto acidified ammonium sulfate seed particles has been demonstrated by our recent study (Krechmer et al. AMT, 2016). Figure 5c in Krechmer et al. (2016) shows the distribution of IEPOX derived organosulfate, along with its dimers and trimers, on the 2-D space. A trend line for the major IEPOX monomers, dimers, and trimers is clearly visible. The characteristic fragment upon CID, i.e., sulfate ( $m/z$  97), adds further confirmation on the chemical identities of the IEPOX derived products in the particle phase.

**[Changes]** Title: A Novel Framework for Molecular Characterization of Atmospherically Relevant Organic Compounds Based on Collision Cross Section and Mass-to-Charge Ratio

Abstract: ... “A new metric is introduced for representing the molecular signature of atmospherically relevant organic compounds.” ... “Reactions involving changes in functionalization and fragmentation can be represented by the directionalities along or across these trend lines, thus allowing for the interpretation of atmospheric transformation mechanisms of organic species.” ...

Introduction: Organic species in the atmosphere — their chemical transformation, mass transport, and phase transitions — are essential for the interaction and coevolution of life and climate (Pöschl and Shiraiwa, 2015). Organic species are released into the atmosphere through biogenic processes and anthropogenic activities. Once in the atmosphere, organic species actively evolve via multiphase chemistry and gas-particle phase conversion. The complexity and dynamic behaviors of organic species have prevented our capability to accurately predict their levels, temporal and spatial variability, and oxidation dynamics associated with the formation and evolution of organic aerosols in the atmosphere.

Minor comments:

Line 40: “scatter plots” is not a verb

**[Responses]** We refer to the definition of van Krevelen diagram in Heald et al., GRL, (2010): “*The Van Krevelen diagram was developed to illustrate how elemental composition changes during coal formation [Van Krevelen, 1950]. The diagram cross plots the hydrogen to carbon atomic ratio (H:C) and the oxygen to carbon atomic ratio (O:C).*”

Line 64: Re-word, perhaps use “and subsequent interactions” in place of “as well as”

**[Responses]** Revised as suggested.

Line 169: This phrase is awkwardly broken up and should be re-worded: “the instrument standard (the reduced mobility of such a standard is not affected by contaminants in the buffer gas) is needed”

**[Responses]** Agreed and revised as follows.

**[Changes]** Line 168-173: In view of these uncertainties, the instrument standard is needed to provide an accurate constraint on the instrumental parameters, such as voltage, drift length, pressure, and temperature.

$$K_0 \times t_d = \frac{L_d^2}{V_d} \frac{P}{1013.25} \frac{273.15}{T} = C_i \quad (5)$$

Tetraethyl ammonium chloride (TEA) is used here as the instrument standard, as its reduced mobility is not affected by contaminants in the buffer gas.

Lines 184-200 describe the apparent crux of this framework, but some points could be made clearer. In particular, explicitly relating the measured parameters to the calculated parameters would be very helpful. For instance, in Eq. 6, it would be useful to re-frame in terms of  $t_d$  since that is what is actually being measured, instead of  $K_0$  and  $v_d$ . What is the functional form of this relationship, considering all of the terms in the equation? Is it collision cross section generally linear with drift time?

**[Responses]** It is the arrival time ( $t_a$ ) that is ultimately measured under certain potential gradients applied to the drift tube ( $V_d$ ). The arrival time ( $t_a$ ) is the sum of the time that the ion spent in the drift tube ( $t_d$ ) and the time for the ion from the exit of the drift tube to the detector ( $t_0$ ). Knowing  $t_a$  and  $V_d$ , the mobility ( $K$ ) can be derived through linear regression of Equation (3). The reviewer is correct that the collision cross section is a linear function of the drift time. But since the drift time is not a directly measured quantity, one step of linear regression is

required to derive  $t_d$  or  $K$ . In view of the potential complexity in the calculations, we have attached Matlab codes in calculating the collision cross section in the Supplementary Material. In addition, the physical meaning of all parameters in the manuscript is attached in the Appendix.

**[Changes]** Line 198: Matlab codes for calculating  $\Omega_{N_2}$  are given in the Supplement.

Line 187: Is  $z=1$  assumed for all ions? In contrast to other ionization techniques, ESI can under some conditions yield a distribution of charges – is this an issue and to what extent would it change the results?

**[Responses]** Yes, we have carefully verified that for all the standards characterized in this study, no multiple charged ions are observed in the mass spectra. Electrospray ionization is conducive to the formation of singly charged small molecules (those we care most about) but is also well known for producing multiply charged species of larger molecules. Fortunately, software available with all electrospray mass spectrometers facilitates the molecular weight calculations necessary to determine the mass of the multiply charged species. In terms of distribution of singly charged vs. multiply charged ions on the space, previous studies (e.g. *Pringle et al., Int. J. Mass, 2007*) have demonstrated that they are well resolved because they lie along trend lines with different slopes.

Line 190: Is thermal velocity calculated as a function of molecular mass?

**[Responses]** Yes, it is a function of molecular mass and temperature. Matlab codes for calculating this quantity is attached in the Supplementary Material.

Line 193: How are the mass fractions calculated? How might this work for a mixture of unknowns with poorly defined sensitivities?

**[Responses]** We define  $\hat{m}$  and  $\hat{M}$  are molecular mass fractions of the ion and the buffer gas molecule ( $N_2$ ), respectively:

$$\hat{m} = \frac{m}{m + M} \quad \hat{M} = \frac{M}{m + M}$$

where  $m$  and  $M$  are the molecular masses for the ion and the buffer gas molecule ( $N_2$ ), respectively. Note that the word ‘mass’ here is not the actual concentration for unknowns in a mixture, but the molecular weight for unknowns, which are derived from the measured mass-to-charge ratio. We apologize for the confusion and have clarified this in the manuscript.

Line 214: Define or clarify “(12,4) potential”

**[Responses]** We define the (12,4) potential along with Equation (9).

**[Changes]** Line 214-215: The potential during interaction includes a long-range polarization term and a short-range repulsion term (Mason et al., 1972).”

Line 245-249: The core model, consisting of a (12-4) central potential displaced from the origin, is used to represent interactions of polyatomic ions with N<sub>2</sub> molecules (Mason et al., 1972). The (12-4) central potential includes a repulsive  $r^{-12}$  term, which describes the Pauli repulsion at short ranges due to overlapping electron orbitals, as well as an attractive  $r^{-4}$  term, which describes attractions at long ranges due to ion induced dipole:

Sections 3.1 and 3.2 could perhaps be switched, as they are discussed in Section 3 and in Section 4 in the opposite order.

**[Responses]** We would like to maintain the original order for Section 3.1 and 3.2, as the core model is basically a simplified version of the trajectory method. We feel that Equation (8) should be introduced first to give the readers a complete overview on the description of ion-neutral interaction potentials before moving on to Equation (9), which simplifies the ion structure as a sphere and ignores the  $r^{-6}$  term (attraction at long ranges due to van der Waals force) as it becomes inappreciable compared with attraction due to ion induced dipole.

Line 301: It is not clear to me that “functionalization and fragmentation can be represented by an intrinsic directionality” as claimed by the authors. As the authors note, the connected markers shown in Figure 2 represent addition of non-functionalized carbon atoms, but this is not the form that atmospheric functionalization takes. Instead, addition of carbon moves up and to the right, but addition of functional groups appears to move down. What would a vector of functionalization or fragmentation look like in this space? This question is particularly important if the authors intend to keep their focus on using this approach to characterize complex mixtures.

**[Responses]** As the reviewer suggested, we have added a vector of functionalization and fragmentation on the 2-D space in Figure 3. Here we define the functionalization as the addition of oxygen-containing functional groups and fragmentation as the cleavage of C-C bonds. As we show that molecules with lower collision cross section are generally much denser and more functionalized, the trajectory for functionalization on the space would be mostly downward, with a shift to the right hand side considering the increase of molecular mass. The trajectory for

fragmentation can be illustrated by the trendline for C<sub>8</sub>-C<sub>18</sub> mono-carboxylic acid shown in Figure 2. As fragmentation only involves changes in carbon number, the corresponding trajectory would more or less follow this trendline.

**[Changes]** Line 329-342: 4.3 Trajectories for Atmospheric Transformation Processes

Functionalization (the addition of oxygen-containing functional groups) and fragmentation (the oxidative cleavage of C–C bonds) are key processes during atmospheric transformation of organics. Reactions involving changes in functionalization and fragmentation can be represented by directionalities on the  $\Omega_{N_2} - m/z$  space, as illustrated by the distribution pattern of carboxylic acids in Figure 3. Addition of one carbon atom always leads to an increase in mass and collision cross section, with a generic slope of approximately 5 Å<sup>2</sup>/Th. Although the addition of one oxygen atom in the form of a carbonyl group results in a similar increase in the molecular mass, it leads to a shallower slope compared with that from expanding the carbon chain. Addition of carboxylic or hydroxyl groups does not necessarily lead to an increase in the collision cross section, as the formation of the intramolecular hydrogen bonding (O – H···O<sup>−</sup>) could result in a more compact conformation of the molecule. In general, fragmentation moves materials to the bottom left and functionalization to the right on the space.

Line 345: While the trend lines described in the Section 4.2 provide an important characterization of this framework, and a useful test of the core model, it's not completely clear they would be particularly help in identifying unknown species, as implied in this sentence. As demonstrated by Figure S5, there is substantial overlap between the regions of all of the trend lines – if one were handed an unknown, its location in this space alone would not provide much information on its family. These lines are perhaps useful for identifying family of species, and demonstrate the utility of the core model for helping to understand its location in the space, but as written this sentence is a bit of an overstatement without some explanation or support. Section 4.3, on the other hand, does indeed seem very promising for identifying unknowns...

**[Responses]** The following steps illustrate how we could potentially identify unknowns in a mixture:

- 1). We first try both positive (+) and negative (−) ESI modes and identify the ideal mode in which the sample would yield intensive signals. As we have demonstrated in Section 4.4, amines, aldehydes, and esters can only be detected in positive (+) mode, while carboxylic acids and sulfate yield high signal to noise ratios in negative (−) mode. This gives us primary information on the functionalities an unknown might contain, and more importantly, avoids

overlaps that would have occurred if one mixed data from two modes together, as shown in Figure S5.

2). We then measure the collision cross section ( $\Omega_{N_2}$ ) of the sample and map its location on the  $\Omega_{N_2} - m/z$  space. Based on the predicted trend lines that are constrained by chemical standards, we could possibly identify the chemical families to which the unknowns belong.

3). If there are overlaps of species even in one ESI mode, we need to use the collision induced dissociation (CID) function to identify the characteristic fragments for the overlapping species. For example, *mono*-carboxylic acids yield  $CO_2$  (44 amu), di-carboxylic acids yield both  $CO_2$  and  $H_2O$  (18 amu), sulfates yield  $HSO_3^-$  (97 amu).

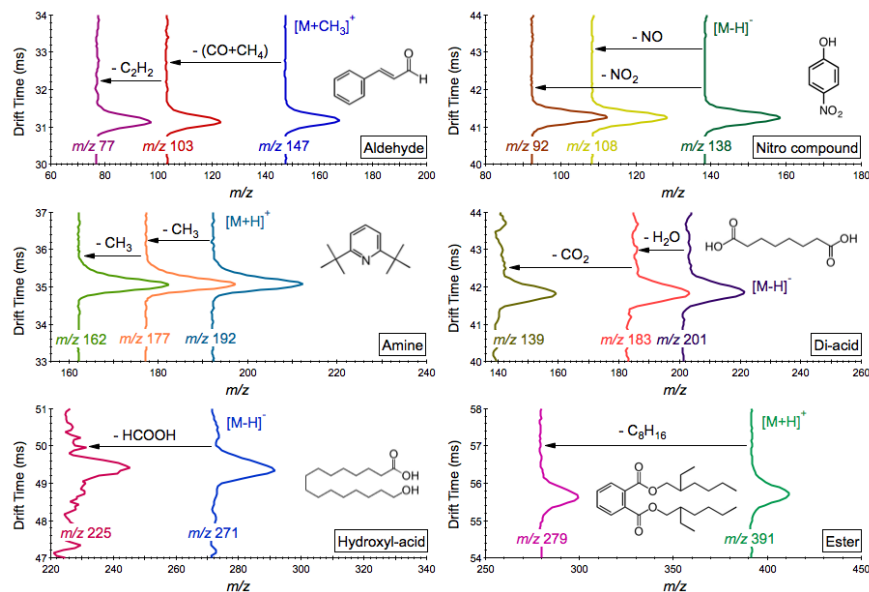
In summary, a combined knowledge on the mass to charge ratio of the unknowns, the location of the unknowns on the 2-D space, as well as the fragmentation patterns of unknowns, could provide information on the chemical classes to which the unknowns belong as well as the functional groups the unknowns contain.

**[Changes]** We have rewritten Section 4.1 and 4.2, by clarifying the application of 2-D space to the identification of chemical classes not the molecular structures of unknowns.

Line 318-328: The demonstrated  $\Omega_{N_2} - m/z$  trend lines provide a useful tool for categorization of structurally related compounds. Mapping out the locations and distribution patterns for various functionalities on the 2-D space would therefore facilitate classification of chemical classes for unknown compounds. It is likely that trend lines extracted from a complex organic mixture overlap and, as a result, the distribution pattern of unknowns on the space alone would not provide sufficient information on their molecular identities. In this case, the fragmentation pattern of unknowns upon collision induced dissociation (CID) needs to be explored for the functionality identification, as discussed in detail in Section 4.4. As it is highly unlikely that two distinct molecules will produce identical IMS, MS, as well as CID-based MS spectra, the 2-D framework therefore virtually ensures reliable identification of species of atmospheric interest.

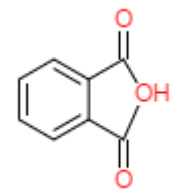
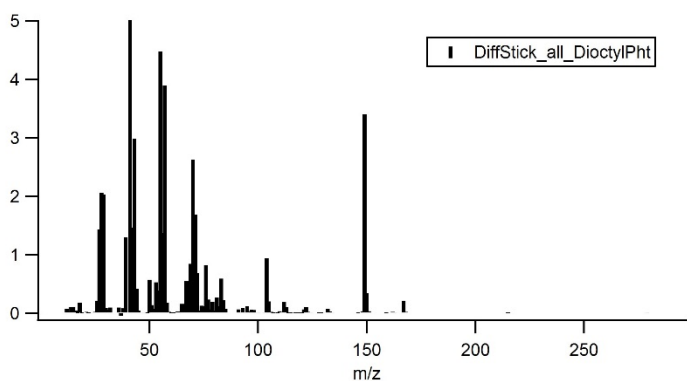
Line 386: A dominant peak at 108 is mentioned but not shown in Figure 4.

**[Responses]** We have added the measured drift time for ion at  $m/z$  108 in Figure 4, also given below:



Line 398: The there is no O-O bond in dioctyl phthalate. Do the authors mean the carbonyl-oxygen bond? Interestingly (and relatedly), this 149 peak is the dominant peak in EI spectra.

**[Responses]** Yes, it should be carbonyl-oxygen bond. We have corrected this in the manuscript. Yes, the  $m/z$  149 peak is also dominant in the EI spectra of dioctyl phthalate, as shown in the Aerodyne Mass Spectrometer (AMS) collected mass spectra below. We tentatively propose the following structure for the  $m/z$  149 ion, and perhaps it is also produced from the fragmentation during EI.



*Proposed structure for  $m/z$  149*

*Data source: Aiken et. al. Anal. Chem., 2007.*