

SCIENTIFIC REPORTS



OPEN

Statistical clumped isotope signatures

T. Röckmann¹, M. E. Popa¹, M. C. Krol^{1,2,3} & M. E. G. Hofmann¹

Received: 05 May 2016

Accepted: 22 July 2016

Published: 18 August 2016

High precision measurements of molecules containing more than one heavy isotope may provide novel constraints on element cycles in nature. These so-called clumped isotope signatures are reported relative to the random (stochastic) distribution of heavy isotopes over all available isotopocules of a molecule, which is the conventional reference. When multiple indistinguishable atoms of the same element are present in a molecule, this reference is calculated from the bulk (\approx average) isotopic composition of the involved atoms. We show here that this referencing convention leads to apparent negative clumped isotope anomalies (anti-clumping) when the indistinguishable atoms originate from isotopically different populations. Such statistical clumped isotope anomalies must occur in any system where two or more indistinguishable atoms of the same element, but with different isotopic composition, combine in a molecule. The size of the anti-clumping signal is closely related to the difference of the initial isotope ratios of the indistinguishable atoms that have combined. Therefore, a measured statistical clumped isotope anomaly, relative to an expected (e.g. thermodynamical) clumped isotope composition, may allow assessment of the heterogeneity of the isotopic pools of atoms that are the substrate for formation of molecules.

Analysis of the isotopic composition of molecules is one of the key tools for studying element cycles on earth. For the light elements H, C, N and O with relatively small heavy-to-light isotope ratios at natural abundance, the standard analytical instruments have largely limited isotope analysis to single-substituted isotopocules (isotopically substituted molecules). Studies of multiply substituted isotopocules, referred to as clumped isotopes, were only occasionally carried out, often using isotopically enriched substrates or labeling experiments^{1–5}. Recent analytical advancements using isotope ratio mass spectrometry^{6–8} or laser spectroscopy⁹ have enabled high precision measurements of clumped isotopes in several molecules such as CO₂, CH₄, O₂ and N₂O^{8,10–15} and the field is rapidly expanding.

Since multiply substituted isotopocules are thermodynamically more stable than single substituted ones, classical isotope theory predicts small but measurable positive clumped isotope anomalies for most molecules under natural conditions^{16–19}. These clumped isotope signatures depend on temperature, which is the basis of the new field of clumped isotope thermometry²⁰.

Yeung *et al.*¹⁴ and Wang *et al.*¹³ reported negative heavy isotope clumping in photosynthetic O₂ formation and in biogenic CH₄, respectively. Yeung *et al.*¹⁴ attributed the negative Δ values (see equation 3 for definition) in photosynthetic O₂ to different isotopic composition of the two O atoms originating from different sites in the oxygen evolving complex of photosystem II. Triggered by this observation we investigated this further and show here that negative clumping anomalies are necessarily expected whenever two or more indistinguishable atoms of the same element but with different isotopic composition combine in a molecule. The atoms do not need to share a common bond but can be at distant places in a molecule.

Yeung²¹ recently presented an analysis of such apparent statistical clumped isotope effects in combination with other isotope effects. In our paper, we restrict the analysis to statistical clumped isotope effects and phrase the calculations exclusively in terms of isotope ratios in order to elucidate the underlying general nature of these apparent isotope signatures. The fundamental origin of the apparent statistical clumped isotope effect is thoroughly presented and visualized geometrically. We then provide a general mathematical formalism for apparent statistical clumped isotope signatures in any multiple isotope system. Finally we demonstrate quantitatively how a certain measured statistical anti-clumping signal can be used to determine the isotopic heterogeneity of indistinguishable atoms in a molecule.

¹Institute for Marine and Atmospheric research Utrecht (IMAU), Utrecht University, Utrecht, The Netherlands.

²Wageningen University, Wageningen, Netherlands. ³SRON Netherlands Institute for Space Research, Utrecht, Netherlands. Correspondence and requests for materials should be addressed to T.R. (email: t.roeckmann@uu.nl)

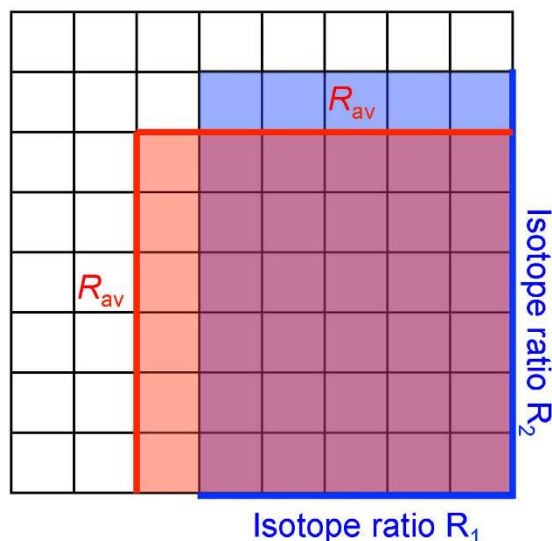


Figure 1. Geometric interpretation of the apparent statistical clumped isotope signatures. The quantity ${}^iR_{cl,random}$ that is used as a reference to calculate Δ_i (Equation 3) is the product of the individual isotope ratios R_1 and R_2 , indicated by the blue lines and blue area. However, for indistinguishable atoms in a molecule the individual ratios R_1 and R_2 cannot be determined independently, and they are both assigned the average ratio R_{av} (red lines). This leads to the red area for ${}^iR_{cl,random}$. The systematic error associated with using the red area instead of the blue area to calculate Δ_i causes the apparent negative clumped isotope signatures described in this paper.

Origin of the statistical negative clumped isotope signatures

We describe and calculate the statistical clumped isotope effects in terms of heavy-to-light isotope ratios iR of individual atoms and molecules, where the index i indicates the mass of the atom or molecule. The same letter R is used for both atomic isotope ratios and molecular isotopocule ratios. In particular, for molecules with multiple heavy isotopes (clumped isotopes), the heavy-to-light isotopocule ratio is defined as

$${}^iR_{cl} = \frac{\text{amount of multi-substituted isotopocule with mass } i}{\text{amount of light isotopocule}} \quad (1)$$

When two atoms with atomic heavy-to-light isotope ratios R_1 and R_2 (where R_i can be, e.g., ${}^2\text{H}/{}^1\text{H}$, ${}^{13}\text{C}/{}^{12}\text{C}$, ${}^{18}\text{O}/{}^{16}\text{O}$, etc.) combine in a molecule in a purely random manner, i.e., without any isotope effect, the ratio of molecules that include the heavy isotopes of both of these atoms relative to molecules including only light isotopes is simply the product of the atomic isotope ratios of the two atoms

$${}^iR_{cl,random} = R_1 R_2 \quad (2)$$

Clumped isotope signatures Δ_i are then by convention calculated as the relative difference between a certain (measured) clumped isotopocule ratio and the random clumped isotope ratio, usually reported in per mill (‰).

$$\Delta_i = \frac{{}^iR_{cl}}{{}^iR_{cl,random}} - 1 \quad (3)$$

The apparent negative statistical clumped isotope signatures that we describe in this paper are fundamentally related to this referencing convention, in particular the choice of the reference ratio ${}^iR_{cl,random}$ that is required to calculate Δ_i in Eq. 3. When the isotope ratios R_1 and R_2 of the involved atoms are individually known, ${}^iR_{cl,random} = R_1 \cdot R_2$ can be precisely calculated. This is always the case when heavy isotopes of different elements clump together (e.g. ${}^{13}\text{C}$ and ${}^{18}\text{O}$ in CO or in CO_2). It also holds when molecules of the same element, for which the individual isotope ratios are known, clump together (e.g. ${}^{15}\text{N}^\alpha$ and ${}^{15}\text{N}^\beta$ in N_2O , where α and β indicate the central and terminal position of the N atom in the linear NNO molecule, which can be determined independently^{22,23}). This is graphically illustrated in Fig. 1, where R_1 and R_2 are plotted on the x- and y-axis and their product, ${}^iR_{cl,random}$, is shown as the blue area.

When a molecule contains indistinguishable atoms of the same element, it is impossible to determine the individual atomic isotope ratios of these atoms. Nevertheless, the atoms may originate from isotopically distinct populations with isotope ratios R_1 and R_2 . To calculate the correct value of ${}^iR_{cl,random}$ we would therefore need to know the individual isotope ratios R_1 and R_2 . However, since these isotope ratios cannot be retrieved for indistinguishable atoms, it is common (and reasonable) to assign the bulk (\approx average, see below) isotopic composition of the atoms. Through this choice, the real stochastic clumped isotope ratio $R_1 \cdot R_2$ (blue area in Fig. 1) is substituted by the approximated value $R_{av} \cdot R_{av}$ (red area in Fig. 1). Both areas have the same perimeter $2 \cdot (R_1 + R_2) = 2 \cdot (R_{av} + R_{av})$, but the red square has a larger area than the blue rectangle, i.e., $R_{av} \cdot R_{av} > R_1 \cdot R_2$. Replacing the area

| Compound | Isotope Combinations | | |
|-------------------------------|---|---|-----------------|
| H ₂ | ² H ² H | | Anti-clumping |
| O ₂ | ¹⁸ O ¹⁸ O | ¹⁷ O ¹⁷ O | Anti-clumping |
| | ¹⁷ O ¹⁸ O | | Anti-clumping |
| N ₂ | ¹⁵ N ¹⁵ N | | Anti-clumping |
| H ₂ O | ¹ H ² H ¹⁸ O | ¹ H ² H ¹⁷ O | No effect |
| | ² H ² H ¹⁶ O | | Anti-clumping |
| CO | ¹³ C ¹⁸ O | ¹³ C ¹⁷ O | No effect |
| CO ₂ | ¹³ C ¹⁸ O ¹⁶ O | ¹³ C ¹⁷ O ¹⁶ O | No effect |
| | ¹² C ¹⁸ O ¹⁸ O | ¹² C ¹⁷ O ¹⁷ O | Anti-clumping |
| | ¹² C ¹⁷ O ¹⁸ O | | Anti-clumping |
| N ₂ O | ¹⁴ N ¹⁵ N ¹⁸ O | ¹⁴ N ¹⁵ N ¹⁷ O | No effect |
| | ¹⁵ N ¹⁵ N ¹⁶ O | | Anti-clumping* |
| CH ₄ | ¹³ CH ₃ ² H | | No effect |
| | ¹² CH ₂ ² H ₂ | | Anti-clumping |
| O ₃ | ¹⁸ O ¹⁶ O ¹⁸ O | ¹⁷ O ¹⁶ O ¹⁷ O | Anti-clumping |
| | ¹⁶ O ¹⁸ O ¹⁸ O | ¹⁶ O ¹⁷ O ¹⁷ O | Anti-clumping** |
| | ¹⁷ O ¹⁶ O ¹⁸ O | | Anti-clumping |
| | ¹⁶ O ¹⁷ O ¹⁸ O | ¹⁶ O ¹⁸ O ¹⁷ O | Anti-clumping** |
| NO ₂ | ¹⁸ O ¹⁵ N ¹⁶ O | ¹⁷ O ¹⁵ N ¹⁶ O | No effect |
| | ¹⁸ O ¹⁴ N ¹⁸ O | ¹⁷ O ¹⁴ N ¹⁷ O | Anti-clumping |
| | ¹⁷ O ¹⁴ N ¹⁸ O | | Anti-clumping |
| C ₂ H ₆ | ¹³ C ¹² C ² H ¹ H ₅ | | No effect |
| | ¹³ C ¹³ C ¹ H ₆ | | Anti-clumping |
| | ¹² C ¹² C ² H ₂ ¹ H ₄ | | Anti-clumping |

Table 1. Combinations of dual heavy isotope clumping in selected atmospheric compounds and information whether the corresponding clumped isotope signatures are affected by apparent statistical clumped isotope effects or not. *If the N atom positions cannot be individually determined. **If symmetric and asymmetric O atoms cannot be individually determined.

of the blue rectangle by the one of the red square in the denominator of Eq. 3 causes a systematic negative artifact. This produces the apparent negative statistical clumped isotope effect. The anti-clumping signal is the larger the more different the individual isotope ratios of the indistinguishable atoms are. In the following chapters we derive the general mathematical formalism to calculate these apparent statistical clumped isotope effects in any multi-isotope system. We also show that the measurement of statistical anti-clumping in principle allows quantifying the heterogeneity of the isotopic composition of indistinguishable atoms in a molecule, i.e. to reconstructing the blue area from the red area in Fig. 1.

We emphasize that the apparent anti-clumping signature is not related to a physical isotope effect, but is a mathematical artifact that originates from the referencing convention. It will never occur when the contributing atoms are distinguishable (thus never for atoms from different elements, e.g. for ¹³C-¹⁸O clumping in CO₂), but it will always occur when two indistinguishable atoms of the same element combine in a molecule (e.g. for ¹⁸O-¹⁸O clumping in CO₂). Table 1 shows a selection of common atmospheric molecules and specific clumping signatures for which statistical anti-clumping will occur or not occur, respectively.

Clumping of indistinguishable atoms in one molecule.

Molecules with two atoms of the same element (e.g. N₂, O₂, H₂). As mentioned above, two indistinguishable atoms of the same element in one molecule may generally originate from different reservoirs or involve different fractionation effects such that their isotope ratios R_1 and R_2 represented two distinct pools when the molecule formed. Since the two atoms are now indistinguishable, we cannot independently measure the isotope ratios R_1 and R_2 . In fact, in conventional isotope ratio measurements of single substituted isotopocules the arithmetic average ratio of the two ratios (e.g. $^{29}R = 2 \ ^{15}R_{av}$ for ¹⁵N measurements in N₂) is determined. For rare heavy isotopes ($R_1, R_2 \ll 1$), this average ratio $^{15}R_{av}$ is generally similar to the bulk isotope ratio R_{bulk} of the sample (see Supplementary Information). In this case, it is common and reasonable to assign $R_{bulk} \approx R_{av}$ to each of the indistinguishable atoms for further calculations. For the remainder of this paper, we use $R_{bulk} = R_{av}$, which considerably simplifies the formulas and removes the dependency of the apparent clumped isotope signal on the isotope ratio. The differences between using R_{bulk} and R_{av} are discussed in detail in the Supplementary Information.

The stochastically expected (random) ratio of isotopocules with two heavy atoms relative to the light isotopocules, from a population of atoms with average heavy isotope ratio of $R_{av} = \frac{R_1 + R_2}{2}$, is

$$R_{cl,random} = (R_{av})^2 = \left(\frac{R_1 + R_2}{2} \right)^2 \quad (4)$$

However, when the atoms represent different isotopic pools with possibly different isotope ratios, the real clumped isotope ratio, R_{cl} , of doubly substituted isotopocules relative to the light isotopocules is

$$R_{cl} = R_1 \cdot R_2 \quad (5)$$

The apparent statistical clumped isotope Δ is the relative difference between the real and the stochastically expected clumped isotope ratio

$$\Delta = \frac{R_{cl}}{R_{cl,random}} - 1 = \frac{R_1 \cdot R_2}{\left(\frac{R_1 + R_2}{2}\right)^2} - 1 = \frac{4R_1 \cdot R_2}{(R_1 + R_2)^2} - 1 = -\frac{(R_1 - R_2)^2}{(R_1 + R_2)^2} \leq 0 \quad (6)$$

The clumped isotope composition is always negative, except for the case $R_1 = R_2$, for which $\Delta = 0$. Thus, when two atoms of the same element with different isotopic composition combine in a molecule, the resulting molecule will always have an apparent negative clumping signature. The black curve in Fig. 2a shows the size of this quadratic statistical negative isotope clumping according to Eq. (6). The negative clumping signal Δ does not depend on the absolute value of the underlying isotope ratios, but only on the relative difference of the isotope ratios. In the following we refer to this effect as “statistical clumped isotope signature”. Eq. 6 was first derived for the case of formation of molecular O_2 in photosynthesis by Yeung *et al.*¹⁴, who indeed observed negative clumped isotope signals relative to the thermodynamically expected values for photosynthetic O_2 .

Generalization: Molecules with three or more atoms of the same element – complete substitution.

When a molecule contains three or more atoms of the same element, these atoms can generally represent populations with different isotope ratios R_1, R_2, \dots, R_n . We first consider the case of full heavy-isotope substitution, which is the generalization of the two-atom case presented above. As it is not possible to independently measure the individual isotope ratios R_i of the indistinguishable atoms, the arithmetic average isotope ratio $^{18}R_{av}$ ($\approx ^{18}R_{bulk}$, see Supplementary Information) is assigned to each of the indistinguishable atoms for further calculations

$$R_1 = \dots = R_n = R_{av} = \frac{\sum_{i=1}^n R_i}{n} \quad (7)$$

As the atoms are all assigned the same atomic isotope ratio R_{av} , the stochastically expected ratio of fully substituted isotopocules relative to non-substituted isotopocules from this population of atoms is the n -th power of R_{av} .

$$R_{full-cl,random} = \left(\frac{\sum_{i=1}^n R_i}{n}\right)^n \quad (8)$$

This is the n -dimensional equivalent of replacing blue rectangle in Fig. 1 by the red square. The real (=observed) ratio of fully-substituted isotopocules is the product of all isotope ratios involved, which is identical to the n -th power of the geometric mean of the isotope ratios

$$R_{full-cl} = \prod_{i=1}^n R_i = \left(\sqrt[n]{\prod_{i=1}^n R_i}\right)^n \quad (9)$$

Thus, the statistical clumped isotope signature for fully substituted isotopocules is

$$\Delta = \frac{R_{full-cl}}{R_{full-cl,random}} - 1 = \frac{\left(\sqrt[n]{\prod_{i=1}^n R_i}\right)^n}{\left(\frac{\sum_{i=1}^n R_i}{n}\right)^n} - 1 = \left(\frac{\text{geometric mean}}{\text{arithmetic mean}}\right)^n - 1 \leq 0 \quad (10)$$

This equation applies to any set of indistinguishable atoms in a molecule. Since the arithmetic mean is always larger or equal than the geometric mean, the statistical clumped isotope signal is always negative, except for the case where all ratios R_i are identical, in which case the arithmetic and geometric means are equal and thus $\Delta = 0$. Figure 2 shows the variation of Δ with the relative difference of the isotope ratios in the 2-, 3-, 4- and 10-atom systems. For the cases illustrated in Fig. 2a, only one isotope ratio is varied and the isotope ratios of all other atoms are kept constant and identical. For the same relative difference in isotope ratio of a single atom, the clumping signal decreases with increasing number of atoms. Although the 10-atom case may not be of much practical use, it is included to emphasize the point that the statistical negative heavy isotope clumping does not require the heavy isotopes to be linked directly by a common chemical bond. For example, statistical negative D-D clumping in ethane (C_2H_6 , Table 1) may involve pairs of hydrogen atoms at any of the 6 positions.

The isotope combinations presented in Fig. 2a all include the point where all isotope ratios are equal, which corresponds to $\Delta = 0\%$. However, in general, the isotope ratios of the atoms at different positions are not equal. Figures 2b,c show the clumping signal for the 3- and 4-atom cases when one ratio is varied again, and the other ratios are held constant, but at different values for the individual atoms. Now the situation that all isotope ratios are equal cannot occur and the parabola-shaped curves are shifted towards negative Δ values. The y-axis offset increases with increasing difference of the individual isotope ratios, thus with the heterogeneity of the isotopic composition of the individual indistinguishable atoms. The curves in Fig. 2b,c are selected 2-dimensional cross-sections of a multi-dimensional space, which illustrate the effect of varying one of the multiple isotope ratios relative to a fixed set of other ratios. In practice, a certain combination of isotope ratios among indistinguishable atoms will correspond to one single value of Δ and we will show below that the statistical clumped

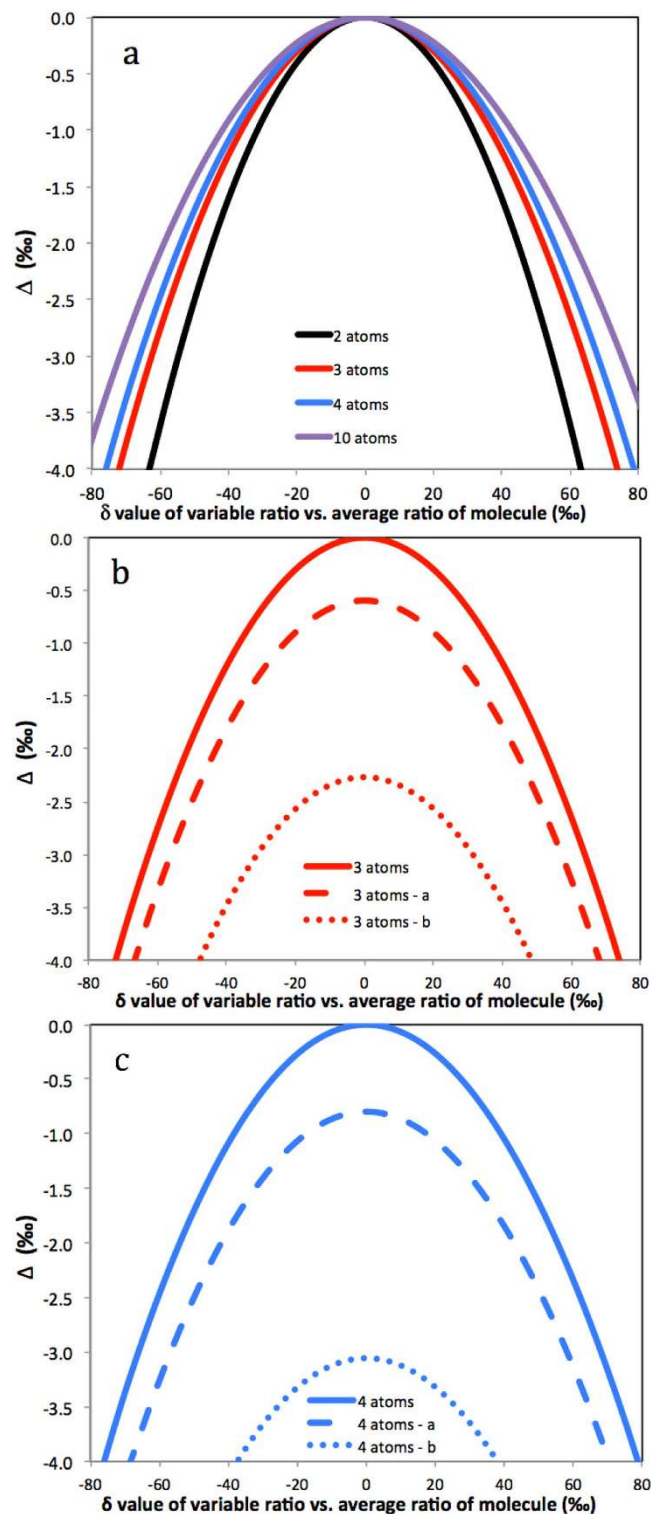


Figure 2. Dependence of the apparent statistical clumped isotope signal for combining two (black), three (red), four (blue) or 10 (purple) indistinguishable atoms in a molecule. In all cases one isotope ratio is varied to produce the relative difference of this ratio from the mean of all isotope ratios (x axis scale). In (a) and the solid lines in (b,c) the other isotope ratios are fixed and identical. This case includes the situation where all isotope ratios are equal and $\Delta = 0$ ‰. In (b) (for 3 indistinguishable atoms) and (c) (for 4 indistinguishable atoms) the solid lines are the same curves as in (a). The dashed lines show the case where one of the fixed isotope ratios is increased by 5% relative to the others, and the dotted line shows the case where one isotope ratio is increased by 10%.

isotope signal Δ is a measure for the heterogeneity of the isotopic composition of indistinguishable atoms in a molecule.

Molecules with three or more atoms of the same element – incomplete substitution. Clumping of two heavy isotopes in molecules with three indistinguishable atoms (e.g. ^{18}O - ^{18}O or ^{17}O - ^{17}O clumping in O_3 or NO_3). As a first example we consider molecules with three indistinguishable atoms and calculate the clumping signature of double substituted isotopocules. The three atoms generally represent three isotopically different populations with isotope ratios R_1 , R_2 and R_3 . However, as the atoms are indistinguishable, the individual ratios cannot be determined and for further calculations they are assigned the average isotope ratio, which can be determined from measurement of the single substituted isotopocules

$$R_{av} = \frac{R_1 + R_2 + R_3}{3} \quad (11)$$

The stochastically expected isotope ratio of isotopocules with exactly two out of three possible heavy isotopes relative to the light isotopocules from a population of indistinguishable atoms with assigned isotope ratio R_{av} is

$$R_{cl,2/3,random} = 3(R_{av})^2 \quad (12)$$

The factor 3 gives the number of possible permutations of two heavy isotopes over three atom positions. The real probability for finding a molecule with exactly two heavy atoms from the three atoms with isotope ratios R_1 , R_2 and R_3 is

$$R_{cl,2/3} = R_1R_2 + R_1R_3 + R_2R_3 \quad (13)$$

Thus, the statistical clumped isotope signal for clumping of two out of 3 possible heavy isotopes, $\Delta_{2/3}$, is

$$\begin{aligned} \Delta_{2/3} &= \frac{R_{cl,2/3}}{R_{cl,2/3,random}} - 1 = \frac{R_1R_2 + R_1R_3 + R_2R_3}{3(R_{av})^2} - 1 \\ &= \frac{3(R_1R_2 + R_1R_3 + R_2R_3)}{(R_1 + R_2 + R_3)^2} - 1 \\ &= \frac{3(R_1R_2 + R_1R_3 + R_2R_3) - (R_1^2 + R_2^2 + R_3^2 + 2R_1R_2 + 2R_1R_3 + 2R_2R_3)}{(R_1 + R_2 + R_3)^2} \\ &= -\frac{R_1^2 + R_2^2 + R_3^2 - (R_1R_2 + R_1R_3 + R_2R_3)}{(R_1 + R_2 + R_3)^2} \\ &= -\frac{(R_1 - R_2)^2 + (R_1 - R_3)^2 + (R_2 - R_3)^2}{2(R_1 + R_2 + R_3)^2} \leq 0 \end{aligned} \quad (14)$$

Since all terms in the numerator and denominator are squares and thus positive, Δ is always negative except for the case $R_1 = R_2 = R_3$ when $\Delta_{2/3} = 0$. Some examples for the clumping of two out of three heavy isotopes in a molecule are shown in Fig. 3a. The solid line again includes the case where all ratios are identical and $\Delta = 0$. The dashed and dotted lines show examples where one isotope ratio varies and the other ones are constant but different, so that the case that all are equal does not occur. Again, the dashed and dotted curves are shifted to more negative Δ values.

Clumping of two heavy isotopes in molecules with four indistinguishable atoms (e.g. D-D clumping in CH_4). We now consider molecules with four indistinguishable atoms and calculate the clumping signature of doubly-substituted isotopocules. The four atoms generally represent isotopically distinct pools with isotope ratios R_1 , R_2 , R_3 and R_4 . As the ratios cannot be determined individually, they are assigned the average atomic isotope ratio

$$R_{av} = \frac{R_1 + R_2 + R_3 + R_4}{4} \quad (15)$$

The stochastically expected probability for forming a molecule with exactly two out of four possible heavy isotopes from a population of indistinguishable atoms with assigned isotope ratio R_{av} is

$$R_{cl,2/4,random} = 6(R_{av})^2 \quad (16)$$

The factor 6 again gives the number of possible permutations of 2 heavy isotopes over 4 atom positions $\binom{4}{2} = \frac{4!}{2!(4-2)!} = 6$. However, the real probability to form an isotopocule with exactly two heavy atoms from the four atoms with different isotope ratios R_1 , R_2 , R_3 and R_4 is

$$R_{cl,2/4} = R_1R_2 + R_1R_3 + R_1R_4 + R_2R_3 + R_2R_4 + R_3R_4 \quad (17)$$

Thus, the clumped isotope signal Δ is

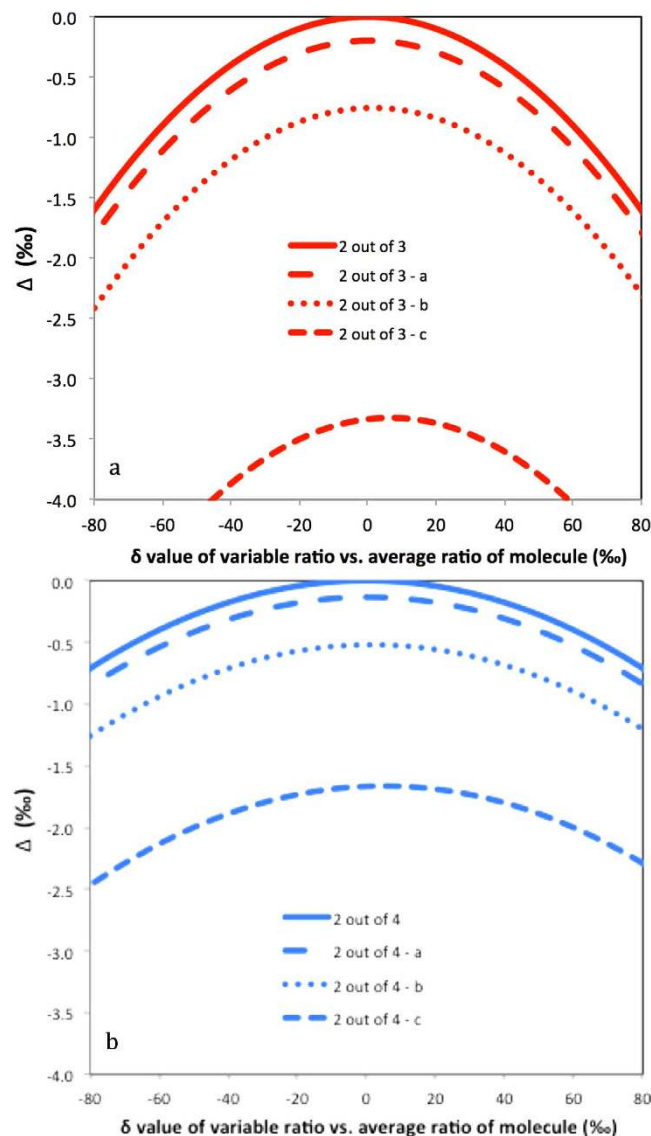


Figure 3. Statistical heavy isotope clumping of two out of three (a) and two out of four (b) heavy isotopes. In all cases shown, the isotope ratio of one atom is varied relative to the mean of the other ratios as indicated on the x axis and the other ratios are fixed. In the base case (solid line), the isotope ratios of the other atoms are identical. This includes the situation for which all isotope ratios are identical ($\Delta = 0\text{‰}$), in case a (dashed line) one of the fixed isotope ratios is increased by 5%, in case b (dotted line) one of the fixed isotope ratios is increased by 10% and in case c (long-dashed line) one of the fixed isotope ratio is increased by 10% and another ratio is decreased by 10%.

$$\begin{aligned}
 \Delta_{2/4} &= \frac{R_{cl,2/4}}{R_{cl,2/4,random}} - 1 = \frac{R_1R_2 + R_1R_3 + R_1R_4 + R_2R_3 + R_2R_4 + R_3R_4}{6(R_{av})^2} - 1 \\
 &= \frac{8(R_1R_2 + R_1R_3 + R_1R_4 + R_2R_3 + R_2R_4 + R_3R_4) - 3(R_1 + R_2 + R_3 + R_4)^2}{3(R_1 + R_2 + R_3 + R_4)^2} \\
 &= -\frac{3(R_1^2 + R_2^2 + R_3^2 + R_4^2) - 2(R_1R_2 + R_1R_3 + R_1R_4 + R_2R_3 + R_2R_4 + R_3R_4)}{3(R_1 + R_2 + R_3 + R_4)^2} \\
 &= -\frac{(R_1 - R_2)^2 + (R_1 - R_3)^2 + (R_1 - R_4)^2 + (R_2 - R_3)^2 + (R_2 - R_4)^2 + (R_3 - R_4)^2}{3(R_1 + R_2 + R_3 + R_4)^2} \leq 0
 \end{aligned} \tag{18}$$

Since Δ can again be expressed as a negative sum of squares, it is always negative, except for $R_1 = R_2 = R_3 = R_4$ where $\Delta_{2/4} = 0$. Some examples for the statistical clumped isotope effect of two out of four heavy isotopes in a molecule are shown in Fig. 3b. An important example for this case is the D-D clumping in methane. The reservoirs that supply the different hydrogen atoms in the formation of methane can vary considerably and significant apparent statistical anti-clumping is expected.

Generalization: Clumping of m heavy isotopes in molecules with n indistinguishable atoms.

For the general case of n indistinguishable atoms that represent isotopic pools with isotope ratios R_1, R_2, \dots, R_n , we assign again the arithmetic mean isotope ratio R_{av} to each of the atoms.

$$R_{av} = \frac{\sum_{i=1}^n R_i}{n} \quad (19)$$

The stochastically expected ratio of isotopocules with exactly m out of a possible n heavy atoms relative to the light isotopocules from a population of indistinguishable atoms with this average heavy isotope ratio R_{av} is

$$R_{cl,random} = \binom{n}{m} (R_{av})^m \quad (20)$$

The real ratio of isotopocules with m heavy isotopes relative to non-substituted isotopocules (R_{cl}) from n atoms with isotope ratios R_1, R_2, \dots, R_n is

$$R_{cl} = \sum_{\{j_1, \dots, j_m\} \subset \{1, \dots, n\}} \prod_{i=1}^m R_{j_i} \quad (21)$$

Thus, the clumped isotope signal Δ is

$$\Delta = \frac{R_{cl}}{R_{cl,random}} - 1 = \frac{\sum_{\{j_1, \dots, j_m\} \subset \{1, \dots, n\}} \prod_{i=1}^m R_{j_i}}{\binom{n}{m} (R_{av})^m} - 1 \quad (22)$$

Eq. 22 is the most general equation to calculate statistical negative isotope clumping in any multi-isotope system. In the case where all ratios R_i are identical, $R_{j_i} = R_{av}$, the numerator becomes identical to the denominator because the number of possible subsets $\{j_1, \dots, j_m\}$ out of $\{1, \dots, n\}$ is equal to the binomial coefficient $\binom{n}{m}$ and the products reduce to the factors $(R_{av})^m$. In this case, $\Delta = 0$.

More than 2 stable isotopes: ^{17}O - ^{18}O clumping for oxygen. So far we have formally only treated molecules with two stable isotopes. Oxygen has three stable isotopes ^{16}O , ^{17}O and ^{18}O . Clumping of heavy isotopes of the same sort (i.e., clumping of multiple ^{17}O atoms or multiple ^{18}O atoms in one molecule) follows the examples and general rules outlined above (see Table 1). However, it is also possible that the heavy ^{17}O and ^{18}O atoms clump together in one molecule.

O_2 and other molecules with two indistinguishable O atoms. In the case of molecular O_2 , the apparent statistical clumping signal for ^{17}O - ^{18}O clumping was derived in Yeung *et al.*¹⁴. Since the isotope ratios of the individual O atoms cannot be determined individually, they are assigned the average heavy isotope ratios

$$^{17}R_{av} = \frac{^{17}R_1 + ^{17}R_2}{2} \quad \text{and} \quad ^{18}R_{av} = \frac{^{18}R_1 + ^{18}R_2}{2} \quad (23)$$

The stochastically expected (random) ratio of molecules with one ^{17}O and one ^{18}O isotope relative to ^{16}O from this population of O atoms is

$$\begin{aligned} ^{35}R_{cl,random} &= \frac{^{35}(\text{O}_2)}{^{32}(\text{O}_2)} = 2^{17}R_{av} \, ^{18}R_{av} = 2 \frac{^{17}R_1 + ^{17}R_2}{2} \frac{^{18}R_1 + ^{18}R_2}{2} \\ &= \frac{1}{2} (^{17}R_1 + ^{17}R_2)(^{18}R_1 + ^{18}R_2) \end{aligned} \quad (24)$$

The real probability for forming ^{17}O ^{18}O molecules from two atoms with isotope ratio $^{17}R_1, ^{18}R_1$ and $^{17}R_2, ^{18}R_2$ is

$$^{35}R_{cl} = ^{17}R_1 ^{18}R_2 + ^{17}R_2 ^{18}R_1 \quad (25)$$

Therefore,

$$\begin{aligned} \Delta_{35} &= \frac{^{35}R_{cl}}{^{35}R_{cl,random}} - 1 = \frac{^{17}R_1 ^{18}R_2 + ^{17}R_2 ^{18}R_1}{\frac{1}{2} (^{17}R_1 + ^{17}R_2)(^{18}R_1 + ^{18}R_2)} - 1 \\ &= \frac{2^{17}R_1 ^{18}R_2 + ^{17}R_2 ^{18}R_1 - (^{17}R_1 ^{18}R_1 + ^{17}R_1 ^{18}R_2 + ^{17}R_2 ^{18}R_1 + ^{17}R_2 ^{18}R_2)}{(^{17}R_1 + ^{17}R_2)(^{18}R_1 + ^{18}R_2)} \\ &= - \frac{(^{17}R_1 - ^{17}R_2)(^{18}R_1 - ^{18}R_2)}{(^{17}R_1 + ^{17}R_2)(^{18}R_1 + ^{18}R_2)} \end{aligned} \quad (26)$$

For normal mass dependent fractionation, where $\frac{^{17}R_1}{^{17}R_2} = \left(\frac{^{18}R_1}{^{18}R_2}\right)^\beta$ ref. 24, with three-isotope exponent $\beta \approx 0.53$ ref. 25, $^{17}R_1$ and $^{18}R_1$ are either both smaller or both larger than $^{17}R_2$ and $^{18}R_2$, thus again the clumped isotope

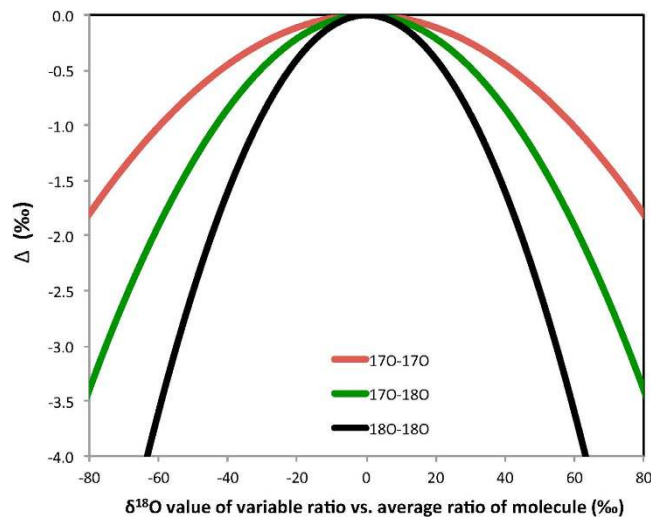


Figure 4. ^{17}O - ^{17}O , ^{17}O - ^{18}O and ^{18}O - ^{18}O clumping in O_2 as function of the $\delta^{18}\text{O}$ value of the variable isotope ratio versus the average isotope ratio in the molecule. Mass dependent fractionation according to

$$\frac{^{17}R_1}{^{17}R_2} = \left(\frac{^{18}R_1}{^{18}R_2} \right)^\beta \text{ with three-isotope exponent } \beta = 0.53 \text{ ref. 25 is assumed.}$$

signature $^{35}\Delta$ is always < 0 ($= 0$ if iR_1 and iR_2 are identical). This can also be shown by inserting the mass dependent fractionation relation as follows:

$$\begin{aligned} \Delta_{35} &= -\frac{(^{17}R_1 - ^{17}R_2)(^{18}R_1 - ^{18}R_2)}{(^{17}R_1 + ^{17}R_2)(^{18}R_1 + ^{18}R_2)} = -\frac{\left(^{17}R_2 \left(\frac{^{18}R_1}{^{18}R_2} \right)^\beta - ^{17}R_2 \right) (^{18}R_1 - ^{18}R_2)}{\left(^{17}R_2 \left(\frac{^{18}R_1}{^{18}R_2} \right)^\beta + ^{17}R_2 \right) (^{18}R_1 + ^{18}R_2)} \\ &= -\frac{\left(\left(\frac{^{18}R_1}{^{18}R_2} \right)^\beta - 1 \right) (^{18}R_1 - ^{18}R_2)}{\left(\left(\frac{^{18}R_1}{^{18}R_2} \right)^\beta + 1 \right) (^{18}R_1 + ^{18}R_2)} = -\frac{(^{18}R_1^\beta - ^{18}R_2^\beta) (^{18}R_1 - ^{18}R_2)}{(^{18}R_1^\beta + ^{18}R_2^\beta) (^{18}R_1 + ^{18}R_2)} \end{aligned} \quad (27)$$

As the atoms do not need to share the same bond to generate statistical isotope clumping, the equation also applies to ^{17}O - ^{18}O clumping of other molecules with 2 O atoms (importantly ^{17}O - ^{18}O clumping in CO_2). The resulting negative clumping signature is shown in Fig. 4.

Heavy isotope clumping for three indistinguishable oxygen atoms (e.g. O_3). When a molecule has three indistinguishable (or not distinguished) oxygen atoms, we assign the average ratios $^{17}R_{av} = \frac{^{17}R_1 + ^{17}R_2 + ^{17}R_3}{3}$ and $^{18}R_{av} = \frac{^{18}R_1 + ^{18}R_2 + ^{18}R_3}{3}$ to each of the atoms. The statistical clumping signatures for complete ^{17}O or ^{18}O substituted isotopocules and for ^{17}O - ^{17}O and ^{18}O - ^{18}O clumping can be derived according to the formalism derived above. Here we also consider the remaining combinations where ^{17}O and ^{18}O clump together in one molecule.

Since there are $3! = 6$ different possibilities to distribute the three distinguishable atoms (^{16}O , ^{17}O and ^{18}O) over the three positions, the stochastically expected clumping is

$$R_{17-18cl,random} = 6^{16}R_{av}^{17}R_{av}^{18}R_{av} = 6^{17}R_{av}^{18}R_{av} \quad (28)$$

The real clumping is calculated by considering explicitly all 6 configurations of ^{16}O , ^{17}O and ^{18}O

$$R_{17-18cl} = ^{17}R_1^{18}R_2 + ^{17}R_2^{18}R_1 + ^{17}R_1^{18}R_3 + ^{17}R_3^{18}R_1 + ^{17}R_2^{18}R_3 + ^{17}R_3^{18}R_2 \quad (29)$$

and thus

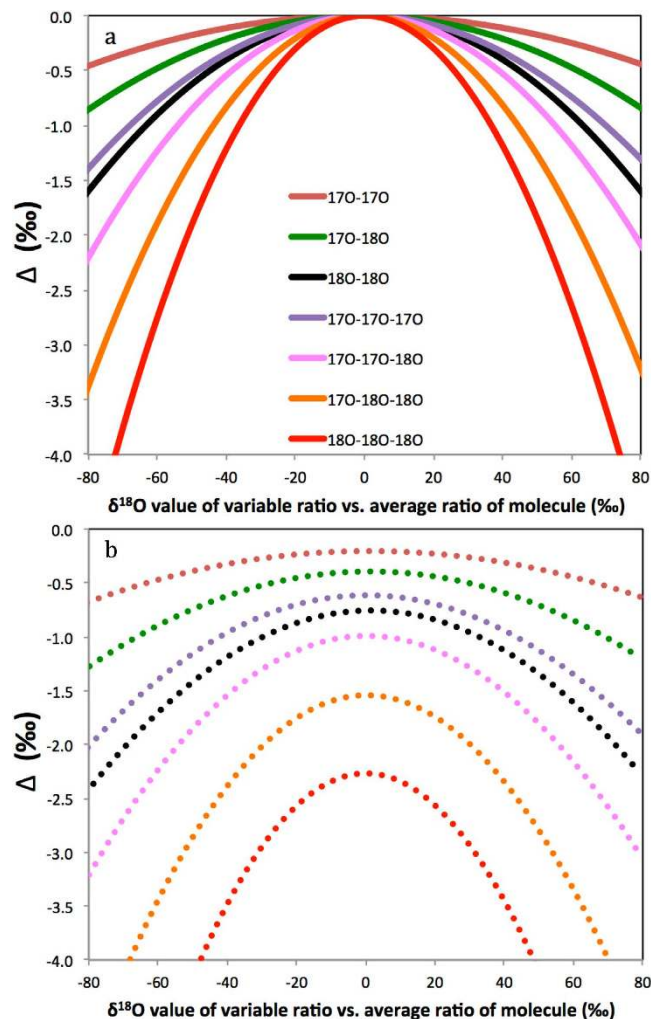


Figure 5. All combinations of ^{17}O - ^{18}O heavy isotope clumping for three indistinguishable O atoms in O_3 as function of the $\delta^{18}\text{O}$ value of the variable isotope ratio versus the average isotope ratio in the molecule. The curves for clumping of a single isotope are the same as in Figures 2 and 3 (red and black curves refer to red solid curves in Figures 2 and 3, respectively). In all scenarios two isotope ratios are kept fixed and the third isotope ratio is varied. In (a) the two fixed isotope ratios are identical, which includes the situation that all isotope ratios are identical and $\Delta = 0\text{‰}$, in (b) one of the fixed isotope ratios is 10% higher than the other. ^{17}O and ^{18}O ratios are linked via the mass dependent fractionation equation $\frac{^{17}R_1}{^{17}R_2} = \left(\frac{^{18}R_1}{^{18}R_2}\right)^\beta$.

$$\begin{aligned}
 \Delta_{17-18} &= \frac{R_{17-18cl}}{R_{17-18cl,random}} - 1 = \frac{^{17}R_1^{18}R_2 + ^{17}R_2^{18}R_1 + ^{17}R_1^{18}R_3 + ^{17}R_3^{18}R_1 + ^{17}R_2^{18}R_3 + ^{17}R_3^{18}R_2}{6^{17}R_{av}^{18}R_{av}} - 1 \\
 &= \frac{9(^{17}R_1^{18}R_2 + ^{17}R_2^{18}R_1 + ^{17}R_1^{18}R_3 + ^{17}R_3^{18}R_1 + ^{17}R_2^{18}R_3 + ^{17}R_3^{18}R_2)}{6(^{17}R_1 + ^{17}R_2 + ^{17}R_3)(^{18}R_1 + ^{18}R_2 + ^{18}R_3)} - 1 \\
 &= \frac{3(^{17}R_1^{18}R_2 + ^{17}R_2^{18}R_1 + ^{17}R_1^{18}R_3 + ^{17}R_3^{18}R_1 + ^{17}R_2^{18}R_3 + ^{17}R_3^{18}R_2) - 2(^{17}R_1 + ^{17}R_2 + ^{17}R_3)(^{18}R_1 + ^{18}R_2 + ^{18}R_3)}{2(^{17}R_1 + ^{17}R_2 + ^{17}R_3)(^{18}R_1 + ^{18}R_2 + ^{18}R_3)} \\
 &= \frac{(^{17}R_1^{18}R_2 + ^{17}R_2^{18}R_1 + ^{17}R_1^{18}R_3 + ^{17}R_3^{18}R_1 + ^{17}R_2^{18}R_3 + ^{17}R_3^{18}R_2) - 2(^{17}R_1^{18}R_1 + ^{17}R_2^{18}R_2 + ^{17}R_3^{18}R_3)}{2(^{17}R_1 + ^{17}R_2 + ^{17}R_3)(^{18}R_1 + ^{18}R_2 + ^{18}R_3)} \\
 &= -\frac{(^{17}R_1 - ^{17}R_2)(^{18}R_1 - ^{18}R_2) + (^{17}R_1 - ^{17}R_3)(^{18}R_1 - ^{18}R_3) + (^{17}R_2 - ^{17}R_3)(^{18}R_2 - ^{18}R_3)}{2(^{17}R_1 + ^{17}R_2 + ^{17}R_3)(^{18}R_1 + ^{18}R_2 + ^{18}R_3)} \quad (30)
 \end{aligned}$$

As argued above for O_2 , for normal mass dependent fractionation, $^{17}R_i$ and $^{18}R_i$ are either both smaller or both larger than $^{17}R_j$ and $^{18}R_j$, thus Δ is always < 0 ($= 0$ if all ratios are identical). Inserting the mass dependent fractionation relation $\frac{^{17}R_1}{^{17}R_2} = \left(\frac{^{18}R_1}{^{18}R_2}\right)^\beta$ again, Eq. 30 can be transformed to

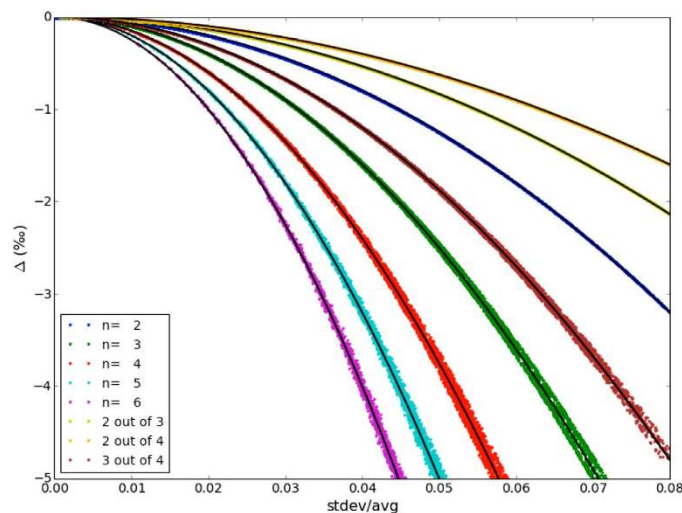


Figure 6. Relation between the apparent statistical heavy isotope clumping signal Δ and the relative standard deviation (stdev/average) of the involved indistinguishable atoms for randomly generated sets of isotope ratios. For each multi-isotope system, the points fall close to well-defined analytical curves (black lines) of the general form $\Delta_{m \text{ out of } n} = -\frac{m}{n} \frac{m-1}{2} \left(\frac{\text{stdev}}{\text{average}}\right)^2$. Note that the Δ values are given in ‰, i.e. the coefficients of the quadratic term in the fit have to be divided by 1000. As an example, for 2-atomic systems ($n = m = 2$) the fit is $\Delta_{m \text{ out of } n} = -500\text{‰} \left(\frac{\text{stdev}}{\text{average}}\right)^2$.

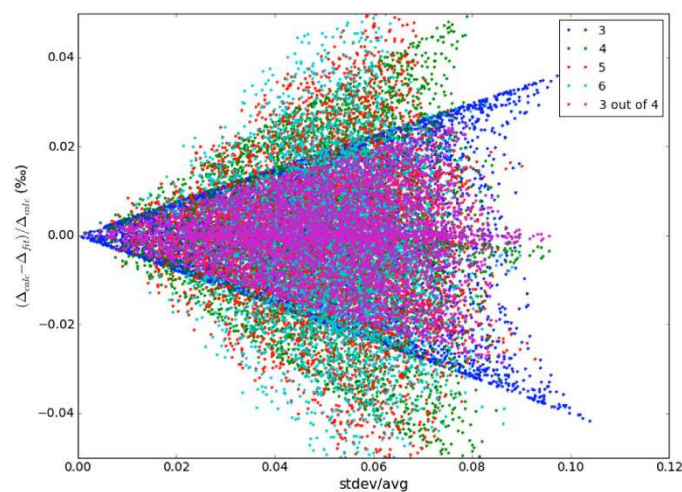


Figure 7. Difference between the analytical curves shown as black lines in Fig. 6 and the calculated value of Δ for the randomly generated sets of isotope ratios (colored points in Fig. 6). For each multi-isotope system, the outer envelope of the points gives the error within which the apparent statistical isotope clumping can be predicted for a certain set of isotope ratios when the relative standard deviation (stdev/average) of the isotope ratios of the indistinguishable atoms is known.

$$\Delta_{17-18} = -\frac{({}^{18}R_1^\beta - {}^{18}R_2^\beta)({}^{18}R_1 - {}^{18}R_2) + ({}^{18}R_1^\beta - {}^{18}R_3^\beta)({}^{18}R_1 - {}^{18}R_3) + ({}^{18}R_2^\beta - {}^{18}R_3^\beta)({}^{18}R_2 - {}^{18}R_3)}{2({}^{17}R_1 + {}^{17}R_2 + {}^{17}R_3)({}^{18}R_1 + {}^{18}R_2 + {}^{18}R_3)} \leq 0 \quad (31)$$

The corresponding values of Δ are shown in Fig. 5 together with all other possible isotope clumping combinations for ^{17}O and ^{18}O isotopes in O_3 . These are always calculated as the relative difference of the explicitly calculated clumped isotope ratio ${}^iR_{cl}$ from the stochastically expected heavy isotope ratio ${}^iR_{cl,random}$ (Equation 30). For example, for ^{17}O - ^{17}O - ^{18}O clumping there are three possible configurations to distribute the atoms (^{17}O , ^{17}O and ^{18}O) over the three positions, so the stochastically expected clumping is

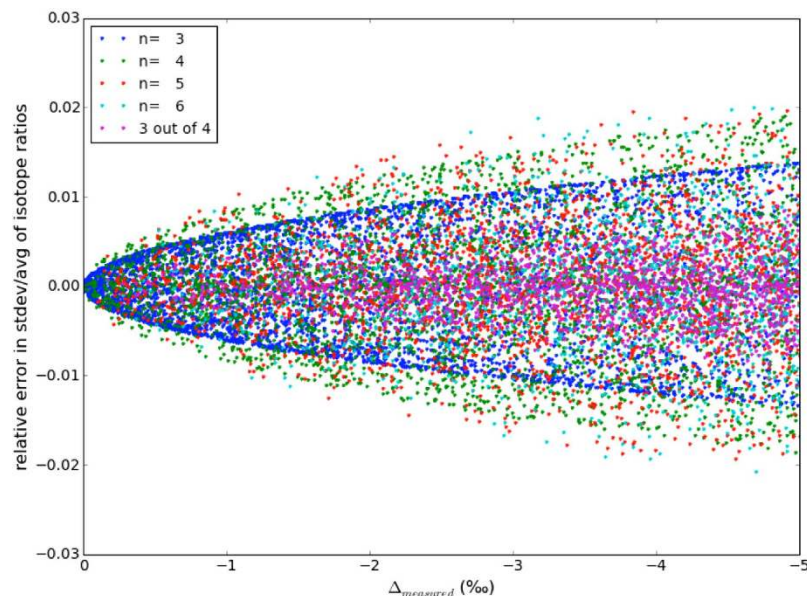


Figure 8. Relative difference of the heterogeneity (relative standard deviation) of the isotope ratios of indistinguishable atoms calculated using Eq. 38 from the true heterogeneity of the randomly chosen sets of isotope ratios. For each multi-isotope system, the outer envelope of the points gives the error within which the heterogeneity of indistinguishable atoms in a certain multi-isotope system can be calculated from the apparent statistical isotope clumping signal Δ .

$${}^{52}R_{cl,random} = 3({}^{17}R_{av})^2 {}^{18}R_{av} \quad (32)$$

The real clumping is calculated by considering explicitly all 3 configurations

$${}^{52}R_{cl} = {}^{17}R_1 {}^{17}R_2 {}^{18}R_3 + {}^{17}R_1 {}^{18}R_2 {}^{17}R_3 + {}^{18}R_1 {}^{17}R_2 {}^{17}R_3 \quad (33)$$

and thus

$$\Delta_{52} = \frac{{}^{17}R_1 {}^{17}R_2 {}^{18}R_3 + {}^{17}R_1 {}^{18}R_2 {}^{17}R_3 + {}^{18}R_1 {}^{17}R_2 {}^{17}R_3}{3({}^{17}R_{av})^2 {}^{18}R_{av}} - 1 \quad (34)$$

The corresponding equation for ${}^{17}\text{O}$ - ${}^{18}\text{O}$ - ${}^{18}\text{O}$ clumping is easily derived by exchanging ${}^{17}\text{O}$ and ${}^{18}\text{O}$.

Important examples of molecules with three oxygen atoms are O_3 and the nitrate ion NO_3^- . For O_3 , the central and terminal O atoms can be distinguished with suitable techniques^{2,26–29}. In these cases, statistical anti-clumping only occurs when the indistinguishable terminal O atoms are involved (Table 1). In many cases, however, the isotopic composition of O_3 is determined without position information^{30–33} and in this case the atoms can be treated as effectively indistinguishable and need to be treated according to the formalism developed here.

Statistical clumped isotope signals and the heterogeneity of the isotope ratios of indistinguishable atoms. The derivations above show that statistical combination of indistinguishable atoms in a molecule leads to apparent negative clumped isotope signals. The size of the apparent negative clumping in the molecule increases with increasing difference in isotopic composition between the individual atoms, so it may actually contain scientifically relevant information. In order to investigate this further, we created random (within a certain range) sets of isotope ratios for all multi-isotope systems with up to 5 indistinguishable atoms and calculated the apparent negative multi-isotope clumping signature Δ . Figure 6 shows that for each multi-isotope set all these random sets of isotope ratios yield Δ values that fall on distinct curves when Δ is plotted versus the relative standard deviation of the individual isotope ratios. For each of the multi-isotope clumping combinations with m out of n heavy isotopes, the curves can be parameterized as:

$$\Delta_{m \text{ out of } n} = -\frac{m}{n} \frac{m-1}{2} \left(\frac{\text{stdev}}{\text{average}} \right)^2 \quad (35)$$

Note that the Δ values in Fig. 6 are plotted as ‰, so the factor 1/2 in Eq. 35 corresponds to 500‰ in the fit curves.

In the case of $m=2$ (2 heavy isotopes in any multi-isotope system) the fit is perfect, but for more heavy atoms clumping in one molecule there is some scatter around these fit lines. This originates from the fact that there is no fixed analytical relation between the arithmetic and geometric means. Figure 7 shows the relative deviation of the explicitly calculated Δ values and the approximation using Eq. 35 for the randomly chosen sets of isotope ratios with known standard deviation and average values. The outer envelopes of the point clouds for each multi-isotope

system define the error with which the statistical clumped isotope signal Δ in a molecule can be predicted from Eq. 35 when the standard deviation and the mean of the individual isotope ratios are known.

Scientifically, the opposite relation

$$\left(\frac{\text{stdev}}{\text{average}}\right) = \sqrt{\frac{2n\Delta_{m \text{ out of } n}}{m(m-1)}} \quad (36)$$

is more attractive, since the isotopic variability among indistinguishable atoms in a molecule is usually not known, but Δ may be measurable¹⁴. This means that measurement of the statistical clumped isotope anomaly Δ could provide a novel tracer to determine the heterogeneity (quantified by the standard deviation) of the isotopic pools of indistinguishable atoms in a molecule. For example, in the 2-isotope system O₂, a Δ value of 1.5‰ below the thermodynamically expected value as measured by Yeung *et al.*¹⁴ would indicate a relative standard deviation in the isotope ratios of the two O atoms of about 5.5% (Fig. 6, blue curve).

Figure 8 shows the relative error that is made when Eq. 36 is used to calculate the relative standard deviation of the isotope ratios from the Δ value for randomly chosen sets of (known) isotope ratios. The outer envelope for each isotope system quantifies the error with which $\left(\frac{\text{stdev}}{\text{average}}\right)$ for a group of indistinguishable atoms can be derived from Δ . Above Δ values of -1.5‰ the relative error in $\left(\frac{\text{stdev}}{\text{average}}\right)$ is generally below 1%, and above Δ values of -5‰ the relative error is still only 2%. Thus, whereas we are not able to measure the isotopic composition of individual indistinguishable atoms, the apparent statistical isotope clumping provides a means to obtain information about the heterogeneity of the isotope ratios with quite good precision.

Conclusions

The statistical combination of indistinguishable atoms with different isotope ratios in a molecule always leads to apparent negative clumped isotope signals. We emphasize the term apparent, because this signal does not relate to a physical negative clumping process. The underlying reason is that in the calculation of the stochastic reference value for calculating Δ , the actual isotopic composition at each individual atom position is replaced by the average of the isotopic composition of all indistinguishable atoms (which is similar bulk isotopic composition of the molecule for small isotope ratios). Thus the apparent statistical Δ is by nature an artifact originating from our limitation to measure the isotope ratios of indistinguishable atoms. Using the formalism presented in this paper, this apparent statistical clumping signal can be calculated for any multi-isotope system.

We have calculated here the pure apparent statistical heavy isotope clumping values. In nature these apparent clumping signals will always occur in combination with thermodynamic heavy isotope clumping and possible other kinetic isotope effects that lead to clumped isotope anomalies (Yeung²¹). The statistical clumped isotope signatures will *always* occur whenever two or more indistinguishable atoms clump together in a molecule. For isotope heterogeneities of a few percent, the effect is of the same order of magnitude as the thermodynamic effects in many molecules. Thus, it is important to take these effects into consideration when interpreting isotopic clumping of indistinguishable atoms in nature. Furthermore, when the statistical clumping signature can be separated from other contributions, its magnitude provides quantitative information on the heterogeneity of the isotopic composition of the indistinguishable atoms in a molecule.

References

- Kaiser, J., Röckmann, T. & Brenninkmeijer, C. A. M. Assessment of ¹⁵N¹⁵N¹⁶O as a tracer of stratospheric processes. *Geophys. Res. Lett.* **30**, 10.1029/2002GL016073 (2003).
- Janssen, C., Guenther, J., Krankowsky, D. & Mauersberger, K. Relative formation rates of ⁵⁰O₃ and ⁵²O₃ in ¹⁶O-¹⁸O mixtures. *J. Chem. Phys.* **111**, 7179–7182 (1999).
- Mauersberger, K., Erbacher, B., Krankowsky, D., Günther, J. & Nickel, R. Ozone isotope enrichment: Isotopomer-specific rate coefficients. *Science* **283**, 370–372 (1999).
- Hauck, R. D. & Bouldin, D. R. Distribution of Isotopic Nitrogen in Nitrogen Gas During Denitrification. *Nature* **191**, 871–872 (1961).
- Mroz, E. J. *et al.* Detection of Multiply Deuterated Methane in the Atmosphere. *Geophys. Res. Lett.* **16**, 677–678, doi: 10.1029/Gl016i007p00677 (1989).
- Eiler, J. M. & Schauble, E. ¹⁸O¹³C¹⁶O in Earth's atmosphere. *Geochim. Cosmochim. Acta* **68**, 4767–4777, doi: 10.1016/j.gca.2004.05.035 (2004).
- Yeung, L. Y., Young, E. D. & Schauble, E. A. Measurements of ¹⁸O-¹⁸O and ¹⁷O-¹⁸O in the atmosphere and the role of isotope-exchange reactions. *J. Geophys. Res.* **117**, D18306, doi: 10.1029/12012JD017992 (2012).
- Eiler, J. M. *et al.* A high-resolution gas-source isotope ratio mass spectrometer. *Int J. Mass Spect.* **335**, 45–56 (2013).
- Ono, S. *et al.* Measurement of a Doubly Substituted Methane Isotopologue, ¹³CH₃D, by Tunable Infrared Laser Direct Absorption Spectroscopy. *Anal. Chem.* **86**, 6487–6494, doi: 10.1021/ac5010579 (2014).
- Stolper, D. A. *et al.* Formation temperatures of thermogenic and biogenic methane. *Science* **344**, 1500–1503, doi: 10.1126/science.1254509 (2014).
- Stolper, D. A. *et al.* Combined ¹³C-D and D-D clumping in methane: Methods and preliminary results. *Geochim. Cosmochim. Acta* **126**, 169–191, doi: 10.1016/j.gca.2013.10.045 (2014).
- Stolper, D. A. *et al.* Distinguishing and understanding thermogenic and biogenic sources of methane using multiply substituted isotopologues. *Geochim. Cosmochim. Acta* **161**, 219–247, doi: 10.1016/j.gca.2015.04.015 (2015).
- Wang, D. T. *et al.* Nonequilibrium clumped isotope signals in microbial methane. *Science* **348**, 428–431, doi: 10.1126/science.aaa4326 (2015).
- Yeung, L. Y., Ash, J. L. & Young, E. D. Biological signatures in clumped isotopes of O₂. *Science* **348**, 431–434, doi: 10.1126/science.aaa6284 (2015).
- Magyar, P. M., Orphan, V. J. & Eiler, J. M. Insights into Mechanisms of Nitrous Oxide Generation from Measurement of Nine N₂O Isotopologues *Goldschmidt Abstracts*, 2015, 1970 (2015).
- Urey, H. C. The thermodynamic properties of isotopic substances *J. Chem. Soc.* 562–581 (1947).

17. Bigeleisen, J. & Mayer, M. G. Calculation of equilibrium constants for isotopic exchange reactions. *J. Chem. Phys.* **15**, 261–267 (1947).
18. Richet, P., Bottinga, Y. & Javoy, M. A review of hydrogen, carbon, nitrogen, oxygen, sulphur, and chlorine stable isotope fractionation among gaseous molecules. *Ann. Rev. Earth Planet. Sci.* **5**, 65–110 (1977).
19. Wang, Z. G., Schauble, E. A. & Eiler, J. M. Equilibrium thermodynamics of multiply substituted isotopologues of molecular gases. *Geochim. Cosmochim. Acta* **68**, 4779–4797 (2004).
20. Eiler, J. M. “Clumped-isotope” geochemistry—The study of naturally-occurring, multiply-substituted isotopologues. *Earth Planet. Sci. Lett.* **262s**, 309–327 (2007).
21. Yeung, L. Y. Combinatorial effects on clumped isotopes and their significance in biogeochemistry. *Geochim. Cosmochim. Acta*, doi: 10.1016/j.gca.2015.1009.1020 (2016).
22. Brenninkmeijer, C. A. M. & Röckmann, T. Mass spectrometry of the intramolecular nitrogen isotope distribution of environmental nitrous oxide using fragment-ion analysis. *Rap. Commun. Mass Spectrom.* **13**, 2028–2033 (1999).
23. Toyoda, S. & Yoshida, N. Determination of nitrogen isotopomers of nitrous oxide on a modified isotope ratio mass spectrometer. *Anal. Chem.* **71**, 4711–4718 (1999).
24. Kaiser, J., Röckmann, T. & Brenninkmeijer, C. A. M. Contribution of mass-dependent fractionation to the oxygen isotope anomaly of atmospheric nitrous oxide. *J. Geophys. Res.* **109**, D03305, doi: dx.doi.org/10.1029/2003JD004088 (2004).
25. Young, E. D., Galy, A. & Nagahara, H. Kinetic and equilibrium mass-dependent isotope fractionation laws in nature and their geochemical and cosmochemical significance. *Geochim. Cosmochim. Acta* **66**, 1095–1104 (2002).
26. Vicars, W. C., Bhattacharya, S. K., Erbland, J. & Savarino, J. Measurement of the ^{17}O -excess ($\Delta^{17}\text{O}$) of tropospheric ozone using a nitrite-coated filter. *Rapid Commun. Mass Spectrom.* **26**, 1219–1231, doi: 10.1002/Rcm.6218 (2012).
27. Janssen, C. Intramolecular isotope distribution in heavy ozone (^{16}O - ^{18}O - ^{16}O and ^{16}O - ^{16}O - ^{18}O). *J. Geophys. Res.* **110**, D08308, doi: 10.1029/2004JD005479 (2005).
28. Larsen, R. W., Larsen, N. W., Nicolaisen, F. M., Sorensen, G. O. & Beukes, J. A. Measurements of ^{18}O -enriched ozone isotopomer abundances using high-resolution Fourier transform far-IR spectroscopy. *J. Mol. Spectrosc.* **200**, 235–247 (2000).
29. Johnson, D. G., Jucks, K. W., Traub, W. A. & Chance, K. V. Isotopic composition of stratospheric ozone. *J. Geophys. Res.* **105**, 9025–9031 (2000).
30. Shaheen, R., Janssen, C. & Röckmann, T. Investigations of the photochemical isotope equilibrium between O_2 , CO_2 and O_3 . *Atmos. Chem. Phys.* **7**, 495–509 (2007).
31. Johnston, J. C., Röckmann, T. & Brenninkmeijer, C. A. M. $\text{CO}_2 + \text{O}(^1\text{D})$ isotopic exchange: Laboratory and modeling studies. *J. Geophys. Res.* **105**, 15213–15229 (2000).
32. Früchtl, M., Janssen, C. & Röckmann, T. Experimental study on isotope fractionation effects in visible photolysis of O_3 and in the $\text{O} + \text{O}_3$ odd oxygen sink reaction. *J. Geophys. Res.* **120**, 4398–4416, doi: 10.1002/2014JD022944 (2015).
33. Chakraborty, S. & Bhattacharya, S. K. Oxygen isotopic fractionation during UV and visible light photodissociation of ozone. *J. Chem. Phys.* **118**, 2164–2172 (2003).

Acknowledgements

This work was carried out as part of the program of the Netherlands Earth System Science Centre (NESSC), financially supported by the Dutch Ministry of Education, Culture and Science (OCW).

Author Contributions

T.R., M.E.P. and M.E.G.H. developed the conceptual framework for the manuscript and discussed the content. M.C.K. contributed code to produce the figures. T.R. developed the mathematical formalism and wrote the manuscript with input from all authors.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Röckmann, T. *et al.* Statistical clumped isotope signatures. *Sci. Rep.* **6**, 31947; doi: 10.1038/srep31947 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016