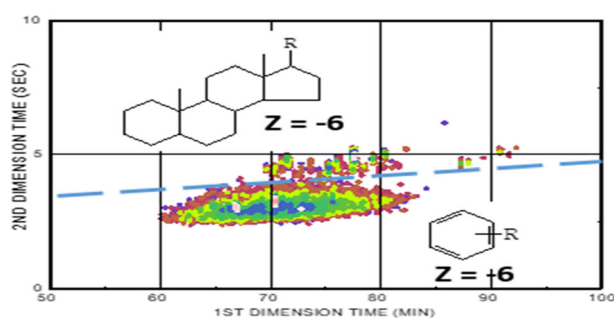


## RESEARCH ARTICLE

# Compositional Analysis of Heavy Petroleum Distillates by Comprehensive Two-dimensional Gas Chromatography, Field Ionization and High-resolution Mass Spectrometry

Kuangnan Qian, Frank C. Wang

ExxonMobil Research Engineering Company, 1545 Route 22 East, Annandale, NJ 08801, USA



**Abstract.** We report recent progresses of combining comprehensive two-dimensional gas chromatography (2DGC or GC  $\times$  GC) separation, field ionization (FI), and time-of-flight mass spectrometry (TOF MS) for the detailed analysis of vacuum gas oil distillation (VGO) cuts. 2DGC separates petroleum molecules by the combination of boiling point and polarity. FI generates molecule ions-only mass spectra. TOF MS allows accurate mass analysis of hydrocarbon molecules. A new

data analysis strategy is implemented for compositional analysis. First, all masses were separated into nominal mass classes. Since petroleum homologues have unique Kendrick mass defects (KMD), KMD plots were generated for easy recognition of homologues series within each nominal mass class. Finally, KMD windows were imposed for complete resolution of petroleum molecules. Using this approach, a total of 16 hydrocarbon types, 14 sulfur types, and their carbon number distributions were determined in the three VGO distillation cuts. Two series of geological biomarkers were also revealed by the analysis.

**Keywords:** Petroleum composition, Field ionization, Two-dimensional GC, Time-of-flight MS, Kendrick mass defect, Petroleum biomarkers, Accurate mass, Vacuum gas oil, Nominal mass class

Received: 24 May 2019/Revised: 12 September 2019/Accepted: 11 October 2019/Published Online: 18 November 2019

## Introduction

Petroleum compositional analysis has been the subject of industrial research for many decades. The advances in analytical techniques continue to push the envelope of the compositional/structural details of petroleum and enabled many R,D and business applications across upstream, downstream, and chemicals sectors of the industry. For example, the structure-oriented lumping (SOL) compositional modeling [1] was based on petroleum composition described as a mixture of hydrocarbon molecules. These molecules were arranged by molecular class/homologue series directly derived from

analytical measurements. Petroleum geological biomarkers, derived from their biological precursors during the geological processes, were extensively used for the assessment of reservoir conditions, oil migration, maturity and quality, etc. [2]

One of the active research areas is the comprehensive two-dimensional gas chromatography (GC  $\times$  GC or 2DGC) [3, 4]. Petroleum compositional space was greatly expanded due to simultaneous separation of petroleum molecules via boiling point and polarity. In upstream applications, 2DGC has been successfully used to detect and quantify geological biomarkers in for petroleum explorations [5–9]. In downstream, 2DGC has now been used to provide full molecular compositions for lower boiling fractions, such as gasoline, jet fuel, and diesels [10–12]. It is increasingly difficult to resolve individual components as boiling point increases. 2DGC was typically used to generate lumped chemical type information for higher boiling fractions, such as whole crude, lube base stocks, and vacuum gas oils (VGO) [13–17]. By combining with elemental selective detectors, such as sulfur and nitrogen chemiluminescence

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s13361-019-02349-6>) contains supplementary material, which is available to authorized users.

Correspondence to: Kuangnan Qian;  
e-mail: Kuangnan.qian@exxonmobil.com

detector (SCD/NCD), certain heteroatom compound types can also be resolved and quantified [11, 12, 18]. In general, 2DGC applications of petroleum analysis are limited to samples with nominal boiling point below 1000 °F (or  $\sim 537$  °C)

As boiling point increases, complete resolution of petroleum molecules becomes challenging. Coupling 2DGC with mass spectrometry (2DGC-MS) proved to be very powerful to enhance compositional details [4, 7, 19–29]. Most of the 2DGC-MS characterization was done by electron impact ionization (EI). EI mass spectrometry fragmentation is very useful to identify chemical structures of the separated species and to resolve isomers/isobars. However, the extensive fragmentation by EI also makes it very challenging to quantify hydrocarbon molecules for compositional analysis. The use of soft ionization, such as field ionization [25, 30–34] and single-photon ionization [35–37] to minimize fragmentation, has been explored and successfully used in petroleum analysis by GC-ToF MS. Because of the lacking of fragmentation, mass spectrometry further separate chemical components by molecular weight, which served as another dimension of separation. When combined with GC or 2DGC, two- and three-dimensional separations have been achieved and illustrated in the recent publications [32, 34, 37, 38]. Recently, a combination of FI and EI was used as a universal biomarker analysis in geochemistry applications [23, 24]. Most recently, 2DGC was also coupled with multiple ionization methods, such as electron ionization, photon ionization, and chemical ionization for petroleum base oil analysis [26]. Most of the research focuses on the analysis of selected chemical types of interests. Complete accounting of all molecules in heavy petroleum distillates is challenging and needs a new data analysis approach.

The work described here invoked the use of nominal mass classes, accurate mass analysis, and Kendrick mass defect (KMD) in conjunction with 2DGC separation to determine full chemical compositions of vacuum gas oil distillates at molecular level. In this approach, all masses were first separated into nominal mass classes. KMD plots (KMD versus molecular weight) within each nominal mass class were generated for easy recognition of homologues series. KMD windows of identified homologues were then imposed to the 2DGC data for complete resolution and full accounts of petroleum molecules (petroleum types and carbon numbers). Byer et al [39] had shown that difficult-to-resolve isobaric mass overlap (e.g.,  $C_3/S_4$ ,  $\Delta M \sim 3.4$  mDa) can be resolved by 2DGC with moderate mass resolution. Here we demonstrated resolution of hydrocarbons and sulfur-containing molecules across the whole carbon number range of VGO without using ultra-high-resolution mass spectrometry such as Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) [40–42]. The latter is more difficult to couple with GC or 2DGC for high-speed analysis. In addition, it has not been successful to combine Field ionization with FTICR-MS for saturate hydrocarbon analysis because of the short life span of parent ions generated under FI conditions [42].

## Experimental

### Instrumentation

The instrument used in this work has been discussed in previous publications [34, 38]. In brief, the GC  $\times$  GC TOF MS system consists of an Agilent 7890 gas chromatograph (Agilent Technology, Wilmington, DE) with a flame ionization detector (FID) and a split/splitless inlet system. The two-dimensional capillary column system is a combination of low polarity first column (BPX-5, 30 m, 0.25 mm i.d., 1.0  $\mu$ m film) and a mid-polarity second column (BPX-50, 3 m, 0.10 mm i.d., 0.10  $\mu$ m film) (SGE Analytical Science., Austin, TX). A ZX1 looped jet thermal modulation assembly (Zoex Corp., Houston, TX) is installed between the two columns. In this work, the effluent of GC  $\times$  GC is split into two streams, one connected to a FID, the other one connected to the ion source of a JEOL MS via transfer line.

A 0.2- $\mu$ L sample was injected with 50:1 split at 300 °C in constant flow mode of 2.0 mL per minute helium. The oven was programmed from 45 to 315 °C at 3 °C/min for a total run time of 90 min. The hot jet is kept at 120 °C above the oven temperature and then kept constant at 390 °C. The MS transfer line and ion source were set at 350 °C and 150 °C, respectively. The modulation period was 10 s.

The MS system is a JMS-T100GCV 4G (JEOL, Tokyo, Japan), time-of-flight mass spectrometer (TOFMS) system with a field ionization (FI) source. The JMS system has an average mass resolution ( $R$ ) of  $\sim 5000$ – $10,000$  and a mass accuracy of  $\sim 5$ – $10$  ppm. Here mass resolution is defined as  $R = M/\Delta M_{FWHM}$ . The high resolution enabled differentiation of isobaric compound types by accurate masses. TOF MS with moderate mass resolution ( $R \sim 5$ – $10$  K) can differentiate  $H_{12}$  from C ( $\Delta M \sim 94$  mDa) and S from  $C_2H_8$  ( $\Delta M \sim 90$  mDa). However, it cannot resolve  $C_3$  from  $SH_4$  ( $\Delta M \sim 3.4$  mDa). Since petroleum homologue series (member of compounds are separated by 14 or the mass of a  $CH_2$  unit) have identical KMDs (see Table 1), the latter can be used to separate and identify petroleum compound types.

A mixture of compounds containing heptacosane, pentafluorobenzene, hexafluorobenzene, pentafluoriodobenzene, pentafluorochlorobenzene, perfluorotrimethylcyclohexane, xylene, and acetone was used to calibrate a mass range from 50 to 800 Da. Once a homologue series was identified, an internal calibration can be performed to further improve mass accuracy.

### Data Analysis

In hydrocarbon analysis, it is convenient to convert mass scale from IUPAC ( $C_{12} = 12.0000$ ) to a Kendrick mass scale [43–45] ( $CH_2 = 14.0000$ ) because petroleum contains homologues with a repeating unit of  $CH_2$ . In Kendrick scale, members of a homologue series are forced to have an identical mass defect. Equations (1) and (2) are used to convert mass scale from

**Table 1.** Nominal Mass Class ( $Z^*$ ), Hydrogen Deficiency ( $Z$ ), Double Bond Equivalent (DBE), and Kendrick Mass Defects (KMD) of Representative Petroleum Compound Types and Core Structures

Hydrocarbons						S Compounds					
$Z^*$	$Z$	DBE	Representative Compound Types	Representative Core Structures	KMD	$Z^*$	$Z$	DBE	Representative Compound Types	Representative Structures	KMD
2	2	0	Paraffins		-13.4	-8	2	0	Aliphatic sulfides		50.2
0	0	1	Cyclic Paraffins		0.0	-10	0	1	Cyclic sulfides		63.6
-2	-2	2	Dicyclic Paraffins		13.4	2	-2	2	Dicyclic sulfides		77.0
-4	-4	3	Tricyclic Paraffins		26.8	0	-4	3	Tricyclic sulfides / Thiophenes (T)		90.4
-6	-6	4	Benzenes/Tetracyclic Paraffins		40.2	-2	-6	4	Naphtheno Thiophenes		103.8
-8	-8	5	Tetralines/Indanes		53.6	-4	-8	5	Dinaphtheno Thiophenes		117.2
-10	-10	6	Dinaphtheno Benzenes		67.0	-6	-10	6	Benzothiophenes (BT)		130.6
2	-12	7	Naphthalenes		80.4	-8	-12	7	Naphtheno BT		144.0
0	-14	8	Bi-phenyls		93.8	-10	-14	8	Dinaphtheno BT		157.4
-2	-16	9	Fluorenes		107.2	2	-16	9	Dibenzothiophenes (DBT)		170.8
-4	-18	10	Phenanthrenes		120.6	0	-18	10	Naphtheno DBT		184.2
-6	-20	11	Acephenanthrenes		134.0	-2	-20	11	Dinaphtheno DBT		197.6
-8	-22	12	Fluoranthenes		147.4	-4	-22	12	Benzo DBT		211.0
-10	-24	13	Chrysenes		160.8	<b>S<sub>2</sub> Compounds</b>					
2	-26	14	Benzofluoracenes		174.2	$Z^*$	$Z$	DBE	Representative Compound Types	Representative Structures	KMD
0	-28	15	Perylenes		187.6	-6	-14	8	Dithiopheno Benzens		207.6

IUPAC to Kendrick mass scale and to calculate the Kendrick mass defects (KMD).

$$\text{Kendrick Mass} = \text{IUPAC Mass} \times 14/14.01565 \quad (1)$$

$$\text{KMD} = (\text{Nominal Mass} - \text{Kendrick Mass}) \times 1000 \quad (2)$$

Here the nominal mass is calculated as the integer number of mono-isotopic mass.

A new data analysis strategy is implemented for compositional analysis. The 2DGC-FI-TOF MS data were imported and processed by an internally developed data analysis and visualization routine. The routine calculates the first and second dimension retention times, the nominal mass classes ( $Z^*$ ), Kendrick mass, and Kendrick mass defect (KMD) for all components detected. Resolution, identification, and quantification of petroleum molecules were performed via multiple steps. First, all masses were separated into nominal mass classes (see discussion later). This separation allows easy recognition/identification of homologue series via KMD

plots. KMD windows were used to isolate the homologue series. Finally, any mass-overlapping isobaric homologues (e.g.,  $\text{C}_3/\text{SH}_4$  overlap) were separated by the 2DGC retention times. The intensities of fully resolved components were summed to generate normalized concentrations (total ion intensities normalized to 100%) and carbon number distributions of identified homologues.

Two-dimensional GC-FID images were processed by PhotoShop (Adobe System Inc., San Jose, CA).

### Samples

Crude oil is commonly distilled into various boiling point fractions, such as naphtha, middle distillate, and vacuum gas oil for detailed chemical and compositional analysis. Distillation reduces the molecular weight range and the variety of chemical types in the sample. In this study, VGO was further distilled into narrow cuts with a nominal 50 °F window. Three VGO cuts with nominal boiling points of 650–702 °F, 702–752 °F, and 752–777 °F were selected and analyzed.

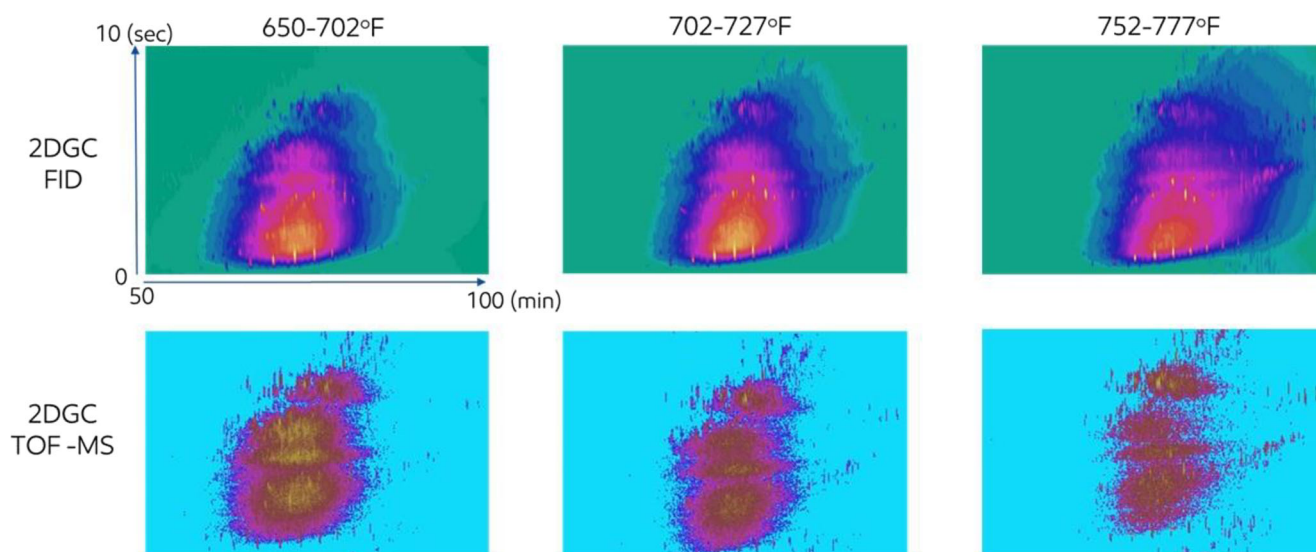


Figure 1. 2DGC images of three vacuum gas oil distillation cuts: 650–702 °F, 702–752 °F, and 752–777 °F. Top row: Data obtained by 2DGC-FID. Bottom row: Data obtained by 2DGC-FI-TOF MS

## Results and Discussions

### 2DGC-FID and 2DGC-FI-TOF MS Images

The 2DGC-FID and 2DGC-FI-TOF MS images of the three VGO distillation cuts are compared in Figure 1. The 1st dimension GC separates petroleum molecules roughly by boiling point or carbon number and the 2nd dimension separates them by polarity. In petroleum applications, the 2nd dimension separation is dominated by aromatic ring types (or unsaturations). The higher the aromatic ring classes, the longer the 2nd retention time. Figure 1 shows that as boiling points of the cuts increase, the 2DGC images moved toward the upper right and became broadened, indicating the increase in molecular size, polarity, and compositional complexity. VGO molecules eluted very closely in the 2DGC space (both 1st and 2nd dimension), making it difficult to resolve individual components by 2DGC-FID only. As will be shown in the later sections, soft ionization mass spectrometry can help to resolve the overlaps and greatly improve resolutions of petroleum compounds. Although the 2DGC-FID image pattern tracks that of TOF MS, we did notice that 2DGC-TOF MS image becomes narrower as boiling point increases. We attribute this to the limited dynamic range of MS relative to that of FID and lower FI sensitivity for higher boiling point fractions.

### Divide 2DGC-MS Data by Nominal Mass Classes

Before we analyze the 2DGC-MS data, it is useful to introduce the concept of nominal mass classes. Petroleum molecules are commonly expressed in general chemical formula  $C_nH_{2n+Z}X$ , where  $n$  is the carbon number and  $X$  is a heteroatom or a combination of heteroatoms.  $Z$  (number) is typically referred as hydrogen deficiency. Nominal mass is defined as the integer formula weight calculated using C (12), H (1), S (32), N (14), and O (16), respectively. All nominal masses of petroleum molecules can be grouped into 14 nominal mass classes (or

$Z^*$  class). Mathematically,  $Z^*$  can be calculated from the nominal mass as follows:

$$M = \text{modulus}(\text{nominal mass}/14) \quad (3)$$

$$Z^* = M - 14 \text{ if } M \geq 4 \text{ and} \quad (4)$$

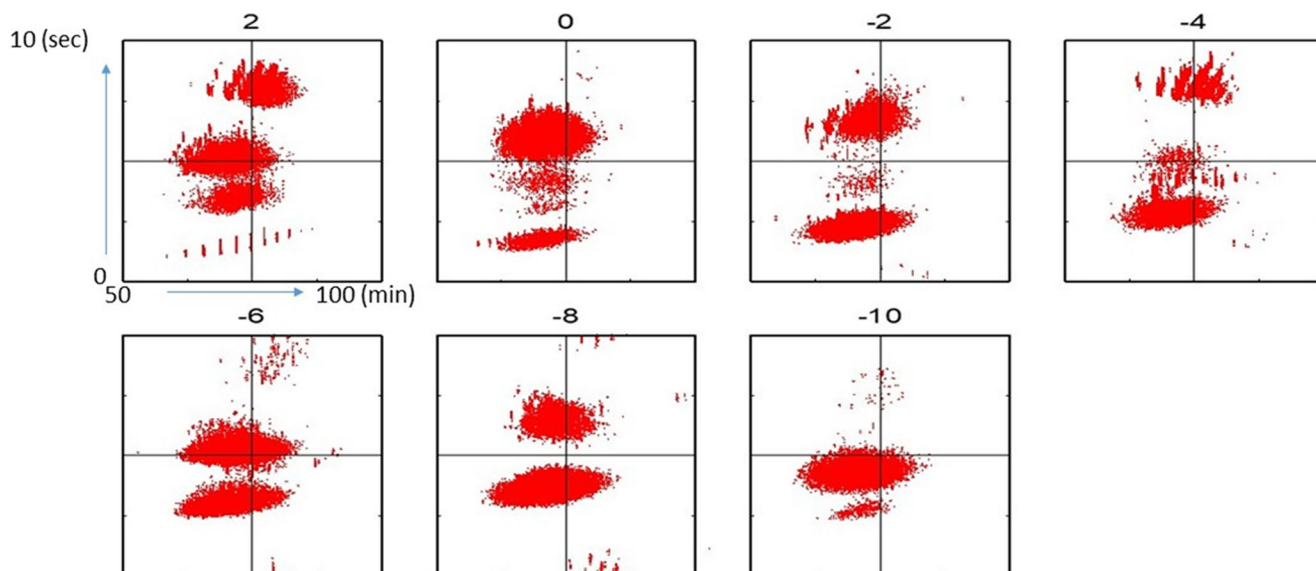
$$Z^* = M \text{ if } M < 4 \quad (5)$$

In this definition,  $Z^*$  equals to the hydrogen deficiency ( $Z$ -number) for the first 7 hydrocarbon classes ( $Z = 2, 0, -2, -4, -6, -8, -10$ ) as shown in Table 1.  $Z^*$  wrap-around occurs when  $Z$  is below  $-10$ .  $Z^*$  and  $Z$  are also different for heteroatom-containing-molecules. Consequently, each  $Z^*$  class can contain multiple homologue series of different  $Z$ -numbers. For example,  $Z^* = 2$  class contains paraffins ( $Z = 2$ ), alkylated naphthalenes ( $Z = -12$ ), dicyclic sulfides ( $Z = -2$  S), and dibenzothiophenes ( $Z = -16$  S) etc.

FI is a soft ionization method that generates mono-isotopic ( $^{12}\text{C}$ ) molecule ions and corresponding  $^{13}\text{C}$  isotope ions of the analytes. The molecule ions of hydrocarbons and sulfur-containing compounds present as even masses. Their mono  $^{13}\text{C}$  isotopes and nitrogen-containing compounds present as odd masses. Only even mass ions were analyzed in this work because the samples contain trace levels of nitrogen compounds. The even nominal masses (mono-isotopic molecular ions) are grouped into 7 nominal mass classes ( $Z^* = 2, 0, -2, -4, -6, -8$ , and  $-10$ , respectively).

The pre-separation of the masses by  $Z^*$  simplified the 2DGC-MS data analysis and visualization. Figure 2 shows the 2DGC-MS plots of the 650–702 °F cut for the individual nominal mass classes. To generate these plots, components with the same  $Z^*$  number were isolated and plotted by their 1st and 2nd dimension retention times. The chemical compound classes in the full 2DGC-MS image (Figure 1) cannot





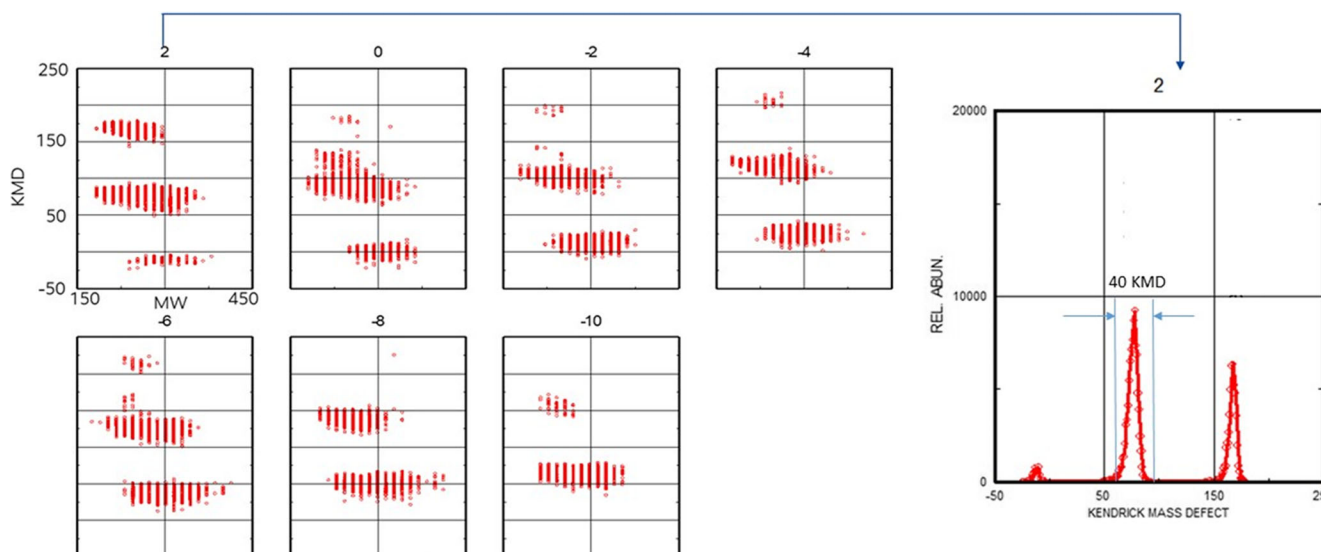
**Figure 2.** 2DGC-FI-TOF MS images of the 650–702 °F cut separated by nominal mass classes:  $Z^* = 2, -2, -4, -6, -8$ , and  $-10$ , respectively

be easily analyzed because of the presence of large number of components and their elution has significant overlaps. The display of the 2DGC-MS image by nominal mass classes revealed presence of multiple clusters of species. These are petroleum homologues of different core structures. For example, in the  $Z^* = 2$  group, normal paraffins can be easily recognized because of their distinguished pattern (single mono-isotopic molecule ions). In addition to n-paraffins, three additional compound groups with higher degrees of unsaturation (eluting at longer 2nd retention time scale) can be recognized. As discussed before, these are due to the presence of multiple homologues that share the same nominal mass series (or the same  $Z^*$ ). Similarly, one can plot the 2DGC-MS data for all the nominal mass classes, resulting separation of 20+ clusters of molecule ions. Although an experienced

analytical chemist may be able to guess the identities of these clusters based on the 2nd dimension elution and their nominal mass class. Positive identifications require accurate mass analysis.

### *Recognition of Petroleum Homologues by Kendrick Mass Defect Analysis*

The mass spectral profiles of the three VGO distillation cuts can be found in the support information Figure S1. As expected, FI generates molecular ion mass spectra (top row) with no detectable fragmentation. As boiling point increases, mass distribution shifted slightly to the right. The mass spectra are very complicated as illustrated in the KMD plots (bottom row) which plot KMD of all species against corresponding  $m/z$ .



**Figure 3.** Kendrick mass defect plots of the 650–702 °F cut separated by nominal mass classes:  $Z^* = 2, -2, -4, -6, -8$ , and  $-10$ , respectively. Inset shows KMD distributions of  $Z^* = 2$  box

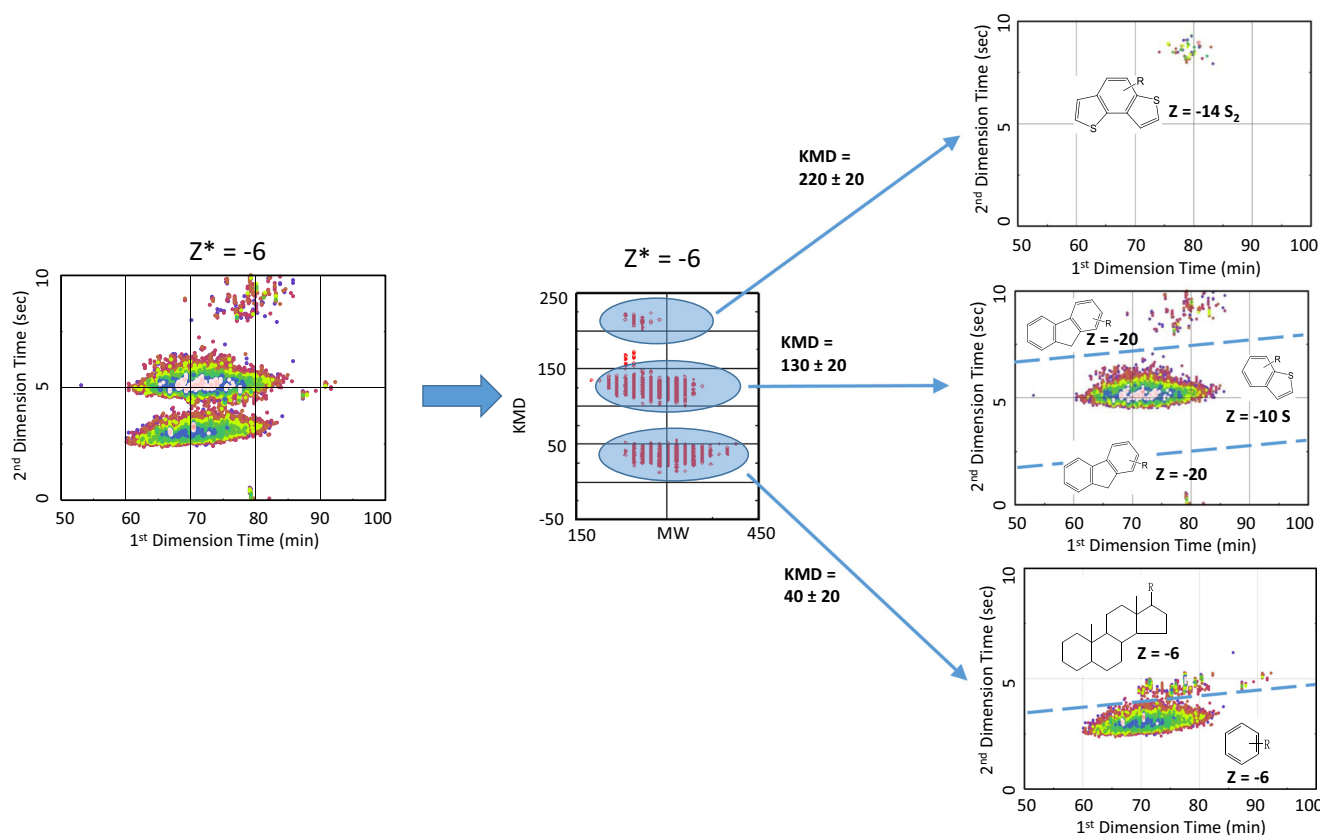


Figure 4.  $Z^* = -6$  nominal mass class further separated by KMDs of  $40 \pm 20$ ,  $130 \pm 20$ , and  $220 \pm 20$ , respectively

There are severe overlaps in KMDs which range from  $\sim -30$  to  $\sim 250$ , indicating the presence of a wide range of chemical compounds with different degrees of unsaturation and heteroatom contents.

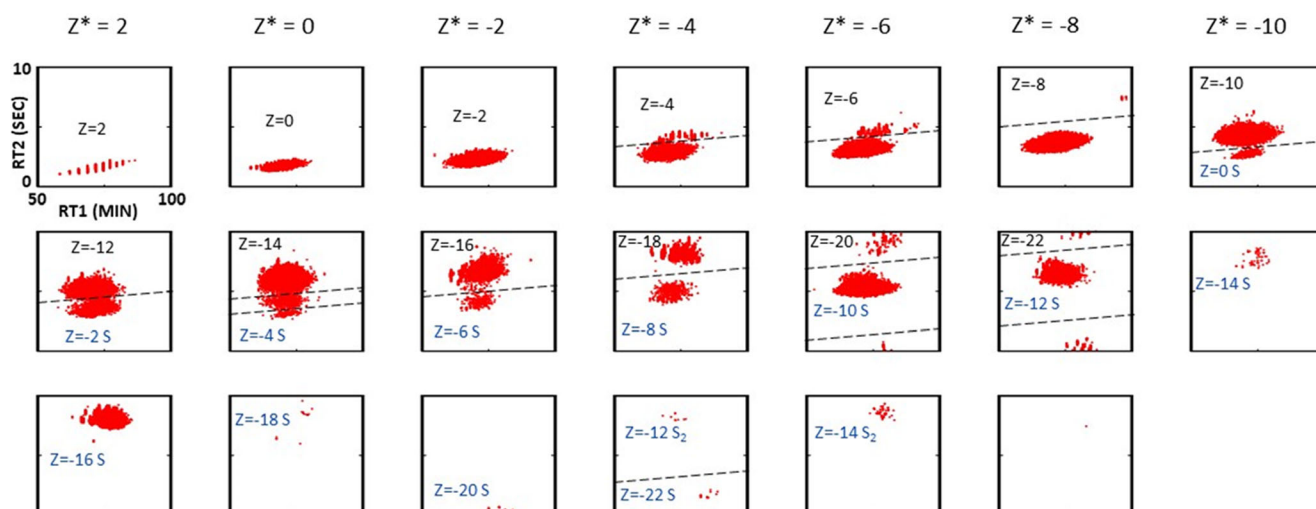
Pre-separation of mass peaks into nominal mass classes enabled easy recognition of homologues with unique KMD values. Figure 3 shows KMD plots of the 650–702 °F cut where KMDs are plotted against nominal masses (or molecular weights because all species are singly charged in FI) for the seven nominal mass classes. Each nominal mass class box contains multiple homologue series. For example, in the  $Z^* = 2$  box, there are three homologues with KMD values around  $-13$ ,  $80$ , and  $173$ , respectively. About 19 homologous in total can be recognized. The inset in the right bottom corner illustrates the KMD distributions of the three homologues in the  $Z^* = 2$  box. The peak widths indicate the spread of KMD values. The  $\Delta$ KMD at full width at half maximum is about  $\pm 10$  KMD unit. In this work, a KMD window of  $\pm 20$  KMD unit is used to select mass peaks for further 2DGC analysis.

General chemical formula ( $C_nH_{2n+Z}S_s$ ) may be assigned by matching the theoretical KMDs of petroleum compound types. Table 1 lists nominal mass class ( $Z^*$ ), hydrogen deficiency ( $Z$ ), double bond equivalent (DBE), and Kendrick mass defects (KMD) of representative petroleum compound types and core structures. For convenience, we normally shorthand the description of petroleum types by their  $Z$ -number and heteroatoms. For example, alkylated benzenes (homologue series) were shorthanded as “ $-6$  HC” and alkylated benzothiophenes

were shorthanded as “ $-10$  S.” Only a few of hydrocarbon (HC) compound types (with  $Z$  ranging from  $+2$  to  $-6$ ) can be assigned without ambiguity by TOF MS. Many compound types cannot be uniquely assigned because of  $C_3/SH_4$  overlap ( $\Delta M \sim 3.4$  mDa). For example, the mass peaks with KMD  $\sim 80$  in  $Z^* = 2$  group can be assigned either as naphthalenes ( $-12$  HC, KMD =  $80.4$ ) or dicyclic sulfides ( $0$  S, KMD =  $77$ ). TOF MS alone cannot resolve these two homologues.

#### Full Composition Analysis by Combining 2DGC and KMD Window Selection

More compositional information was revealed when one impose selected KMD windows in data analyses. Figure 4 shows the image plot of  $Z^* = -6$  class which was further resolved by selective display of KMD around  $40$ ,  $130$ , and  $220$ , respectively. At KMD of  $40 \pm 20$ , general chemical formula  $C_nH_{2n-6}$  (or “ $-6$  HC”) was assigned. Two clusters of molecules were revealed by the 2DGC separation. The bottom cluster is mainly attributed to alkyl benzenes. The top cluster is the sterane biomarkers which exhibited scattered distribution due to the unique structures retained from their biological precursors. Although the two groups of molecules have the same degree of unsaturation (identical chemical formula), biomarkers always eluted later than the non-biomarker molecules due to their bulky conformations. It should be noted that these compounds cannot be easily recognized without the KMD separation (as shown in Figure 2,  $Z^* = -6$  box).



**Figure 5.** 2DGC-FI-TOF MS images of the 650–702 °F cut of seven nominal mass classes separated by KMDs, using a process similar to that in Figure 4

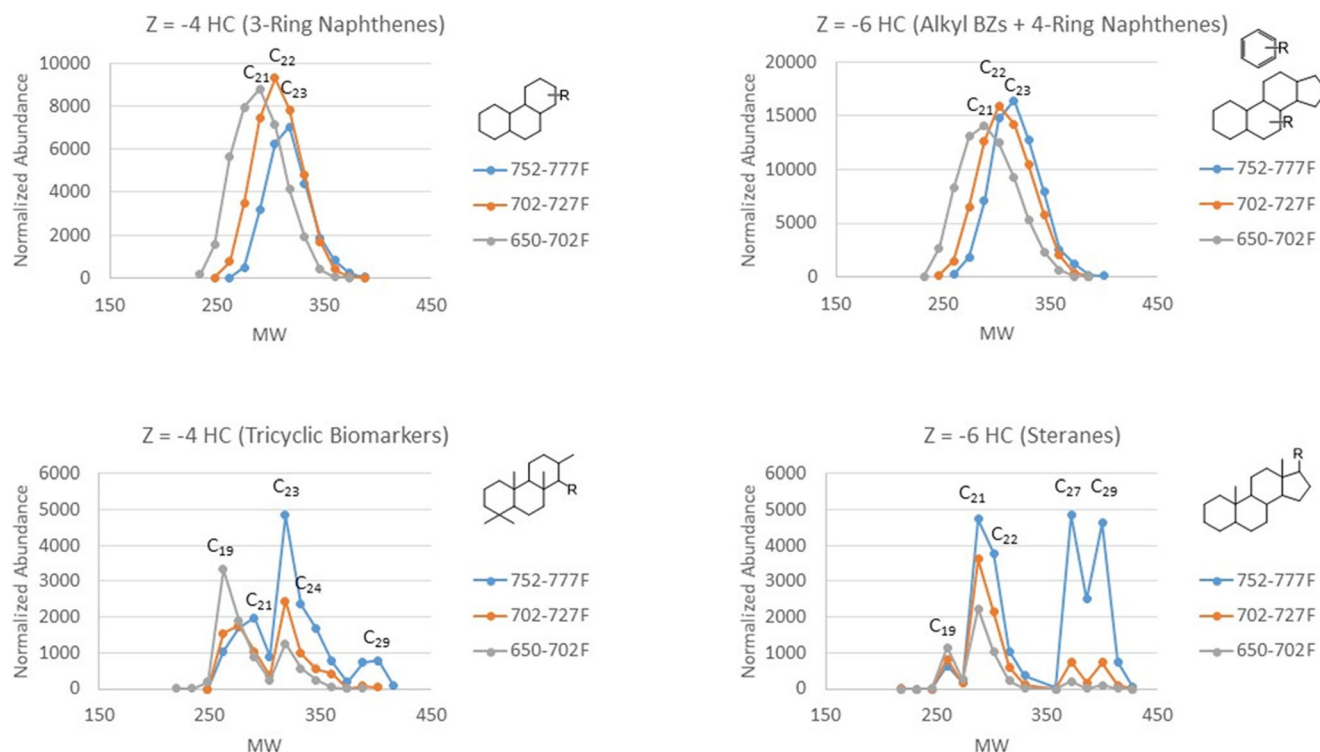
Two more groups of compounds were observed in the 2DGC separations of the KMD 130 window. These are identified as benzothiophenes, “– 10 S,” and acephenanthrenes, “– 20 HC.” The two groups were well resolved by the polarity separation as acephenanthrenes are 3-ring aromatics which elute much later than do benzothiophenes (2-ring aromatics) in the second dimensional separation. Note acephenanthrenes

have wrap arounds in the second dimension separation, appearing in both top and bottom sides of the image.

In the KMD of  $220 \pm 20$  window, we observed another group of compounds eluting in the 3-aromatic ring region. They were identified as dithiophenobenzenes (– 14 S<sub>2</sub>). It should be noted that these species would co-elute with the acephenanthrenes (– 20 HC) shown in the KMD 130 window.

**Table 2.** Petroleum Compound Types in the Three VGO Distillation Cuts (Abundances of All Compound Types Normalized 100%)

Compound types	Z	X	Cut 1 (650–702 °F)	Cut 2 (702–727 °F)	Cut 3 (752–777 °F)
Paraffins	2		0.5	0.6	1.1
Cyclic paraffins	0		1.9	1.9	1.2
Dicyclic paraffins			6.7	6.9	6.5
Tricyclic paraffins			3.8	3.6	2.4
Tricyclic biomarkers			0.8	0.9	1.7
Benzenes/tetracyclic paraffins			6.8	7	6.5
Tetracyclic biomarkers			0.5	0.9	2.4
Tetralines/indanes			12.1	13.2	14.4
Pentacyclic biomarkers			0	0.2	0.6
Dicyclic benzenes	– 10		8.7	7.9	7
Naphthalenes	– 12		7.1	4.5	2
Bi-phenyls	– 14		15.6	13.2	10.2
Fluorenes	– 16		6.5	5.9	4.2
Phenanthrenes	– 18		4.2	5.8	8.3
Acephenanthrenes	– 20		0.2	0.7	1.1
Fluoranthenes	– 22		0.2	0.4	0.8
Aliphatic sulfides/mercaptanes	2	S			
Cyclic sulfides	0	S	0.3	0.3	
Dicyclic sulfides	– 2	S	2.4	2.1	0.9
Tricyclic sulfides/thiophenes (T)		S	0.8	0.4	0.1
Naphtheno thiophenes		S	0.3	0.2	0
Dinaphtheno thiophenes		S	0.7	0.3	0.1
Benzothiophenes (BT)	– 10	S	10.3	10.1	11.2
Naphtheno BT	– 12	S	4.4	4.4	3.4
Dinaphtheno BT	– 14	S	0.1	0.1	0.1
Dibenzothiophenes (DBT)	– 16	S	5	8.2	13
Naphtheno DBT	– 18	S		0.1	0.1
Dinaphtheno DBT	– 20	S			0.2
Benzo DBT	– 22	S			0.4
Dithiopheno benzene	– 14	S <sub>2</sub>			0.1
Total			99.9	99.8	100



**Figure 6.** Molecular weight distributions of  $Z = -4$  and  $-6$  hydrocarbon classes in the 650–702 °F cut. Top: non-biomarker molecules. Bottom: biomarker molecules

The same approach can be repeated for all nominal mass classes to resolve petroleum types. The results are shown in Figure 5. About 30+ molecular types were recognized. They were divided into hydrocarbon types ( $Z = 2$  to  $-22$ ) and sulfur compound types ( $Z = 0$  to  $-16$ ). As  $Z$  decreases, molecules become more unsaturated and eluted progressively later in the second dimensional GC separation.

Starting from  $Z = -10$ , both sulfur and hydrocarbon molecules were observed. Sulfur molecules always elute earlier in the 2nd dimension because they are more saturated than their isobaric hydrocarbon counter parts. These molecules could not be resolved in one-dimensional GC-FI-TOF MS analysis [31] when mass is greater than 300. They were completely resolved by the 2DGC separation. It is interesting to note that there are three clusters of molecule ions in  $Z = -14/-4$  S box. The two “ $-4$  S” isomers were attributed thiophenes and tricyclic sulfides.

### Compositions of VGO Distillation Cuts

The normalized concentrations of the three VGO cuts (shown in Figure 1) from the 2DGC-FI-TOF MS measurements are summarized in Table 2. Since ionization efficiencies of petroleum compounds vary with core structures and molecular weights, the numbers do not represent absolute composition. Quantification may be achieved by applying response factor correction and by integration with 2DGC-FID results. A total of 16 hydrocarbon types and 14 sulfur types were listed and compared for the three boiling point cuts. Hydrocarbon types range from paraffins, 1–3 ring cyclic paraffins, and 1 to 4 ring

aromatics. Sulfur compounds include 1–3 ring cyclic sulfides and 1–3 ring aromatic thiophenes. The relative concentration of sulfur compounds and condensed aromatic species increase with boiling point of the cuts as expected.

Molecular weight or carbon number distributions of the petroleum compound types can be obtained by summation of the separated 2DGC-MS data shown in Figure 5. Figure 6 shows MW distributions of “ $-4$  HC” and “ $-6$  HC” petroleum compound types, respectively. The top distributions are the non-biomarker molecules and bottoms are biomarker molecules of the same general chemical formula. The mass distribution of non-biomarker molecules is relatively smooth (close to Gaussian) as we expected for petroleum distillation cuts. The distribution shifted toward the right side as boiling point of the cut increases as expected. The carbon numbers peaked at C21, C22, and C23 for the three boiling point cuts, respectively.

The mass distributions of the bio-marker molecules (bottom chart) are very spiky (with high intensities at certain carbon numbers), indicating favorable skeleton structures retained from their biological precursors. We observed core skeleton tricyclic and tetracyclic (steranes) structures. Both start with carbon number of 19. Tricyclic biomarker distributions peaked around C21, C23, C24, C28, and C29, respectively. Steranes peaked around C21, C22, C27, and C29, respectively. C27–C30 steranes were commonly found in crude oil and have been extensively applied in the biomarker analysis for geochemistry applications [24]. In this work, they were found to be concentrated in the higher boiling cuts. It is interesting to note that abundant C19, C21, and C22 steranes were also observed. They



are more concentrated in the lower boiling cuts. These short side chain biomarkers could be originated from cracking of the original biomarkers (e.g.,  $C_{27}$ – $C_{29}$  steranes) in the diagenesis process.

## Conclusions

We have shown that combined two-dimensional GC separation, soft ionization, and high-resolution mass spectrometry greatly enhanced compositional analysis of heavy petroleums. A new data analysis strategy is implemented for compositional analysis. By nominal mass class separation, selective KMD window, and 2DGC separation, 30+ petroleum homologues series were fully resolved. Difficult-to-resolve mass overlaps (e.g.,  $C_3/SH_4$ ,  $\Delta M \sim 3.4$  mDa) were clearly separated, revealing sulfide and aromatic sulfur compositions that are hard to separate from hydrocarbon signals in the one-dimensional GC-TOF MS analysis. Two series of biological markers were also revealed with unique carbon number distributions.

## Acknowledgements

Authors greatly appreciate the assistance of Ms. Tara Reddington for the 2DGC and 2D-GC-TOF MS experiments.

## References

- Quann, R.J., Jaffé, S.B.: Structure-oriented lumping: describing the chemistry of complex hydrocarbon mixtures. *Ind. Eng. Chem. Res.* **31**, 2483–2497 (1992)
- Peters, K.E., Walters, C.C., Moldowan, J.M.: *The Biomarker Guide*, 2nd edn. Cambridge University Press, Cambridge (2005)
- Pollo, B.J., Alexandrino, G.L., Augusto, F., Hantao, L.W.: The impact of comprehensive two-dimensional gas chromatography on oil, gas analysis: recent advances and applications in petroleum industry. *TrAC Trends Anal. Chem.* **105**, 202–217 (2018)
- Alam, M.S., Harrison, R.M.: Recent advances in the application of 2-dimensional gas chromatography with soft and hard ionisation time-of-flight mass spectrometry in environmental analysis. *Chem. Sci.* **7**, 3968–3977 (2016)
- Li, S., Cao, J., Hu, S., Zhang, D., Fan, R.: Analysis of terpanes in biodegraded oils from China using comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry. *Fuel* **133**, 153–162 (2014)
- Santos, L.C., da Cruz, G.F., Avila, B.M.F., Pereira, V.B., Azevedo, D.A.: Exploratory analysis of Campos Basin crude oils via geochemical parameters by comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry. *Energy Fuel* **32**, 10321–10332 (2018)
- Wang, Y., Ma, W., Zhou, N., Ren, J., Cao, J.: Analyzing crude oils from the Junggar Basin (NW China) using comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry (GC×GC-TOFMS). *Acta Geochim.* **36**, 66–73 (2017)
- Zhang, W., Zhu, S., He, S., Wang, Y.: Screening of oil sources by using comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry and multivariate statistical analysis. *J. Chromatogr. A* **1380**, 162–170 (2015)
- Zhang, W., Zhu, S., Pang, L., Gao, X., Zhu, G.-T., Li, D.: Determination of diamondoids in crude oils using gas purge microsyringe extraction with comprehensive two dimensional gas chromatography-time-of-flight mass spectrometry. *J. Chromatogr. A* **1478**, 75–83 (2016)
- Wang, F.C.-Y., DiSanzo, F.P., McElroy, F.C.: The applications of comprehensive two-dimensional gas chromatography in the petroleum industry. *Prepr. Am. Chem. Soc. Div. Petrol. Chem.* **49**, 4–8 (2004)
- Wang, F.C.-Y., Robbins, W.K., Di Sanzo, F.P., McElroy, F.C.: Speciation of sulfur-containing compounds in diesel by comprehensive two-dimensional gas chromatography. *J. Chromatogr. Sci.* **41**, 519–523 (2003)
- Wang, F.C.-Y., Robbins, W.K., Greaney, M.A.: Speciation of nitrogen-containing compounds in diesel fuel by comprehensive two-dimensional gas chromatography. *J. Sep. Sci.* **27**, 468–472 (2004)
- Dutriez, T., Courtiade, M., Thiebaut, D., Dulot, H., Hennion, M.-C.: Improved hydrocarbons analysis of heavy petroleum fractions by high temperature comprehensive two-dimensional gas chromatography. *Fuel* **89**, 2338–2345 (2010)
- Parker, A., McGregor, L., Barden, D.: GC × GC-FID for routine analysis of total petroleum hydrocarbons. *LCCG N. Am.* 14–15. UBM Life Sciences (2017)
- Striebig, R.C., Shafer, L.M., Adams, R.K., West, Z.J., DeWitt, M.J., Zabarnick, S.: Hydrocarbon group-type analysis of petroleum-derived and synthetic fuels using two-dimensional gas chromatography. *Energy Fuel* **28**, 5696–5706 (2014)
- Van, G.K.M., Pyl, S.P., Reyniers, M.-F., Vercammen, J., Beens, J., Marin, G.B.: On-line analysis of complex hydrocarbon mixtures using comprehensive two-dimensional gas chromatography. *J. Chromatogr. A* **1217**, 6623–6633 (2010)
- Wang, F.C.-Y., Zhang, L.: Chemical composition of group II lubricant oil studied by high-resolution gas chromatography and comprehensive two-dimensional gas chromatography. *Energy Fuel* **21**, 3477–3483 (2007)
- Machado, M.E.: Comprehensive two-dimensional gas chromatography for the analysis of nitrogen-containing compounds in fossil fuels: a review. *Talanta* **198**, 263–276 (2019)
- Li, S., Cao, J., Hu, S.: Analyzing hydrocarbon fractions in crude oils by two-dimensional gas chromatography/time-of-flight mass spectrometry under reversed-phase column system. *Fuel* **158**, 191–199 (2015)
- Lu, H., Shi, Q., Lu, J., Sheng, G., Peng, P.: a.; Hsu, C. S. Petroleum sulfur biomarkers analyzed by comprehensive two-dimensional gas chromatography sulfur-specific detection and mass spectrometry. *Energy Fuel* **27**, 7245–7251 (2013)
- Oliveira, C.R., Ferreira, A.A., Oliveira, C.J.F., Azevedo, D.A., Santos Neto, E.V., Aquino Neto, F.R.: Biomarkers in crude oil revealed by comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry: depositional paleoenvironment proxies. *Org. Geochem.* **46**, 154–164 (2012)
- Weng, N., Wang, H., Zhang, S., Zhu, G., Cai, D., Wan, S., Liu, J., Yang, Y.: Insight into unresolved complex mixtures of aromatic hydrocarbons in heavy oil via two-dimensional gas chromatography coupled with time-of-flight mass spectrometry analysis. *J. Chromatogr. A* **1398**, 94–107 (2015)
- Walters, C.C., Wang, F.C., Higgins, M.B., Madincea, M.E.: Universal biomarker analysis: aromatic hydrocarbons. *Org. Geochem.* **124**, 205–214 (2018)
- Walters, C.C., Wang, F.C., Higgins, M.B., Madincea, M.E.: Universal biomarker analysis using GC×GC with dual FID and ToF-MS (EI/FI) detection. *Org. Geochem.* **115**, 57–66 (2018)
- Genuit, W., Chaabani, H.: Comprehensive two-dimensional gas chromatography-field ionization time-of-flight mass spectrometry (GC×GC-FI-TOFMS) for detailed hydrocarbon middle distillate analysis. *Int. J. Mass Spectrom.* **413**, 27–32 (2017)
- Giri, A., Coutriade, M., Racaud, A., Stefanuto, P.-H., Okuda, K., Dane, J., Cody, R.B., Focant, J.-F.: Compositional elucidation of heavy petroleum base oil by GC × GC-EI/PI/CI/FI-TOFMS. *J. Mass Spectrom.* **54**, 148–157 (2019)
- Piotrowski, P.K., Tasker, T.L., Burgos, W.D., Dorman, F.L.: Applications of thermal desorption coupled to comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for hydrocarbon fingerprinting of hydraulically fractured shale rocks. *J. Chromatogr. A* **1579**, 99–105 (2018)
- Piotrowski, P.K., Weggler, B.A., Barth-Naftilan, E., Kelly, C.N., Zimmermann, R., Saiers, J.E., Dorman, F.L.: Non-Targeted chemical characterization of a Marcellus shale gas well through GC × GC with scripting algorithms and high-resolution time-of-flight mass spectrometry. *Fuel* **215**, 363–369 (2018)
- Piotrowski, P.K., Weggler, B.A., Yoxtheimer, D.A., Kelly, C.N., Barth-Naftilan, E., Saiers, J.E., Dorman, F.L.: Elucidating environmental fingerprinting mechanisms of unconventional gas development through hydrocarbon analysis. *Anal. Chem.* **90**, 5466–5473 (2018)

30. Androulakis, I.P., Weisel, M.D., Hsu, C.S., Qian, K., Green, L.A., Farrell, J.T., Nakakita, K.: An integrated approach for creating model diesel fuels. *Energy Fuel*. **19**, 111–119 (2005)
31. Qian, K., Dechert, G.J.: Recent advances in petroleum characterization by GC field ionization time-of-flight high-resolution mass spectrometry. *Anal. Chem.* **74**, 3977–3983 (2002)
32. Qian, K., Dechert, G.J., Edwards, K.E.: Deducing molecular compositions of petroleum products using GC-field ionization high resolution time of flight mass spectrometry. *Int. J. Mass Spectrom.* **265**, 230–236 (2007)
33. Ubukatal, M., Reichenbach, S.E., Tao, Q., Wu, Z., Dane, A.J., Cody, R.B.: Detailed type analysis of petroleum samples by using comprehensive two-dimensional gas chromatography/high-resolution mass spectrometry with field ionization. *Prepr. - Am. Chem. Soc., Div. Energy Fuels*. **60**, 642 (2015)
34. Wang, F.C.-Y., Qian, K., Green, L.A.: GC×MS of diesel: a two-dimensional separation approach. *Anal. Chem.* **77**, 2777–2785 (2005)
35. Eschner, M.S., Welthagen, W., Groeger, T.M., Gonin, M., Fuhrer, K., Zimmermann, R.: Comprehensive multidimensional separation methods by hyphenation of single-photon ionization time-of-flight mass spectrometry (SPI-TOF-MS) with GC and GC×GC. *Anal. Bioanal. Chem.* **398**, 1435–1445 (2010)
36. Welthagen, W., Mitschke, S., Muehlberger, F., Zimmermann, R.: One-dimensional and comprehensive two-dimensional gas chromatography coupled to soft photo ionization time-of-flight mass spectrometry: a two- and three-dimensional separation approach. *J. Chromatogr. A*. **1150**, 54–61 (2007)
37. Mitschke, S., Welthagen, W., Zimmermann, R.: Comprehensive gas chromatography-time-of-flight mass spectrometry using soft and selective photoionization techniques. *Anal. Chem.* **78**, 6364–6375 (2006)
38. Wang, F.C.-Y.: Comprehensive three-dimensional gas chromatography mass spectrometry separation of diesel. *J. Chromatogr. A*. **1489**, 126–133 (2017)
39. Byer, J.D., Siek, K., Jobst, K.: Distinguishing the C3 vs SH4 mass split by comprehensive two-dimensional gas chromatography-high resolution time-of-flight mass spectrometry. *Anal. Chem.* **88**, 6101–6104 (2016)
40. Qian, K., Robbins, W.K., Hughey, C.A., Cooper, H.J., Rodgers, R.P., Marshall, A.G.: Resolution and identification of elemental compositions for more than 3000 crude acids in heavy petroleum by negative-ion microelectrospray high-field Fourier transform ion cyclotron resonance mass spectrometry. *Energy Fuel*. **15**, 1505–1511 (2001)
41. Qian, K., Rodgers, R.P., Hendrickson, C.L., Emmett, M.R., Marshall, A.G.: Reading chemical fine print: resolution and identification of 3000 nitrogen-containing aromatic compounds from a single electrospray ionization Fourier transform ion cyclotron resonance mass spectrum of heavy petroleum. *Energy Fuel*. **15**, 492–498 (2001)
42. Schaub, T.M., Hendrickson, C.L., Qian, K., Quinn, J.P., Marshall, A.G.: High-resolution field desorption/ionization Fourier transform ion cyclotron resonance mass analysis of nonpolar molecules. *Anal. Chem.* **75**, 2172–2176 (2003)
43. Kendrick, E.: A mass scale based on  $\text{CH}_2 = 14.0000$  for high resolution mass spectrometry of organic compounds. *Anal. Chem.* **35**, 2146–2154 (1963)
44. Hsu, C.S., Qian, K., Chen, Y.C.: An innovative approach to data analysis in hydrocarbon characterization by online liquid chromatography-mass spectrometry. *Anal. Chim. Acta*. **264**, 79–89 (1992)
45. Hughey, C.A., Hendrickson, C.L., Rodgers, R.P., Marshall, A.G., Qian, K.: Kendrick mass defect spectrum: a compact visual analysis for ultrahigh-resolution broadband mass spectra. *Anal. Chem.* **73**, 4676–4681 (2001)