JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Resolving Coffee Roasting-Degree Phases Based on the Analysis of Volatile Compounds in the Roasting Off-Gas by Photoionization Time-of-Flight Mass Spectrometry (PI-TOFMS) and Statistical Data Analysis: Toward a PI-TOFMS Roasting Model

Hendryk Czech,[†] Claudia Schepler,[†] Sophie Klingbeil,[†,‡] Sven Ehlert,[†,§] Jessalin Howell,[∥] and Ralf Zimmermann*[,†,‡,⊥]

[†]Joint Mass Spectrometry Center, Chair of Analytical Chemistry, Institute of Chemistry, University of Rostock, Dr.-Lorenz-Weg 1, 18059 Rostock, Germany

[‡]Helmholtz Virtual Institute of Complex Molecular Systems in Environmental Health (HICE), www.hice-vi.eu, D-85764 Neuherberg, Germany

[§]Photonion GmbH, 19061 Schwerin, Germany

[∥]The J.M. Smucker Company, 1 Strawberry Lane, Orrville, Ohio 44667, United States

[⊥]Joint Mass Spectrometry Center, Cooperation Group "Comprehensive Molecular Analytics", Institute of Ecological Chemistry, Helmholtz Zentrum München—German Research Center for Environmental Health, 85764 Neuherberg, Germany

Supporting Information

ABSTRACT: Coffee beans of two cultivars, Arabica (Mexico) and Robusta (Vietnam), were roasted in a small-scale drum roaster at different temperature profiles. Evolving volatile compounds out of the roasting off-gas were analyzed by photoionization mass spectrometry at four different wavelengths, either with single-photon ionization (SPI) or resonance-enhanced multiphoton ionization (REMPI). The different analyte selectivities at the four wavelengths and their relevance for the examination of the roasting process were discussed. Furthermore, intensities of observed m/z were grouped by non-negative matrix factorization (NMF) to reveal the temporal evolutions of four roasting phases ("evaporation", "early roast", "late roast", and "overroast") from NMF scores and the corresponding molecular composition from the NMF factor loadings, giving chemically sound results concerning the roasting phases. Finally, linear classifiers were constructed from real mass spectra at maximum NMF scores by linear discriminant analysis to obtain quantities which are simple to measure for real-time analysis of the roasting process.

KEYWORDS: *beverage, single-photon ionization (SPI), resonance-enhanced multiphoton ionization (REMPI), process control, roasting phase*

INTRODUCTION

Coffee is known as a popular and worldwide consumed beverage with an extremely complex flavor. Numerous influences such as the cultivar (Arabica, Robusta), cultivation of the coffee plants, processing of the coffee beans, or brewing contribute to differences in the formation of flavor compounds inside the coffee beans and ultimately the taste of the resulting cup.[1] Green coffee beans contain about 300 volatile compounds and lack in color and characteristic flavor compared to roasted coffee. Both color and flavor are formed during the roasting process through predominantly Strecker and Maillard reactions, leading to more than 500 compounds.[2]

The roasting process can be roughly divided in three phases: (1) an endothermic drying phase characterized by the removal of moisture, (2) the actual roasting phase with a number of complex pyrolytic reactions, a dramatic change in the chemical composition of the beans, and the formation of a large number of substances associated with the flavor and taste of coffee, and finally (3) a rapid cooling phase to stop the exothermic part of the roasting using air or water as cooling agent.[3] The time-dependent release of volatiles in the roasting off-gas contains valuable information about the status of the roasting process and related flavor-forming reactions.[4−6]

Online and real-time measurement techniques with sufficient time resolution and limits of detection are demanded to monitor the roasting process in terms of roast degree. Direct inlet mass spectrometric techniques with time-of-flight mass spectrometer as mass analyzer and chemical ionization (CI), proton-transfer-reaction (PTR),[6−9] and photoionization (PI)[4,10−12] meet the requirements and were applied to monitor roasting off-gas components. Both PTR and PI are regarded as soft ionization techniques, leading to mainly molecular or quasimolecular ions which facilitate the interpretation of mass spectra of complex VOC and SVOC mixtures. Additionally, the ionization of gaseous bulk components such as nitrogen,

Received: April 14, 2016
Revised: May 23, 2016
Accepted: June 6, 2016

DOI: 10.1021/acs.jafc.6b01683
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

oxygen, carbon dioxide, or argon is suppressed. However, analyte selectivities are different due to different ionization mechanisms.[13−15]

PI techniques can be further divided into single-photon (SPI) and resonance-enhanced multiphoton ionization (REMPI). REMPI refers to a selective ionization technique for aromatic compounds, while SPI is regarded as a more universal technique which ionizes compounds with lower ionization energy than the photon energy.[13,15] The presented study involved a PI-TOF-MS with both SPI and REMPI at four different wavelengths to monitor the roasting process of green coffee beans in a laboratory scale roaster.

Generally, the guarantee of a constant quality and different properties of a commercial product are major interests of industry. In particular, coffee roasting companies aim to supply a constant quality of different discrete roast degrees (e.g., light, medium, or dark roast) for which a change in the chemical composition and consequently the taste of the brewed coffee can be observed. Thus, a method to predict the roast degree based on indicators analyzed in real-time is desirable.

This study ties in with previous studies of Wieland et al., Ruosi et al., and Liberto et al. about roast degree control.[9,16,17] A new concept for the identification of roasting phases for different roasting conditions together with an analysis technique for online determination of roasting phase transitions is presented. Photoionization mass spectra at four wavelengths to cover different analyte selectivity were processed in non-negative matrix factorization (NMF) to figure out the temporal evolution of four roasting phases in score data and their associated m/z located in the factor loadings. At the maximum of NMF score contribution, m/z with the most different temporal behavior between two subsequent roasting phases were revealed by calculating Fisher ratios. Finally, data points of pairs of m/z with the largest Euclidean distance between their centers were submitted to linear discriminant analysis (LDA) to derive simple classifiers for roasting phase transitions in real-time analysis. This concept aims at a PI-TOFMS data based roasting model, allowing the real-time determination of the roasting degree and quality at industrial coffee roasting processes in the future.

MATERIALS AND METHODS

Green Coffee Beans, Roaster, and Roasting Procedure. Green Arabica coffee beans (*Coffea arabica*) from Mexico and green Robusta coffee beans (*Coffea canephora*) from Vietnam were kindly supplied by the J.M. Smucker Company. The raw coffee beans were roasted by electrical heated single drum sample roaster PRE 1Z (Probat Burns, Vernon Hills, IL, USA) for the roasting of small batches (up to 3.5 oz per coffee batch). The roaster was equipped with temperature readout to monitor the temperature inside the roaster every minute.

Every experiment was started at an initial temperature of 200 °C by filling 100 g of green coffee beans into the drum. By choosing heating steps and times, three different roasting profiles were applied ("slow roast", step 3.5, about 20 min; "medium roast", step 4, about 11 min; "fast roast", step 6, about 6 min). In all experiments, the coffee beans were intentionally overroasted to identify a possible boundary between dark roasted and overroasted coffee beans.

Sampling with μ-Probe. A μ-probe sampling setup developed by Hertz et al.[18] was used as sampling interface between the roaster and a flexible heated transfer line of a PI-TOF-MS.[4,12] In brief, the μ-probe is composed of a conical shaped heated aluminum base body which is coupled to the heated transfer line by a heated adapter. The centerpiece of the μ-probe is made of a small stainless steel capillary (ID 0.2 mm/OD 0.4 mm), which is connected to the transfer capillary via a capillary union. The μ-probe capillary sticks out of the conically shaped base body by approximately 0.5 mm.

Photoionization Time-of-Flight Mass Spectrometer (PI-TOFMS). The self-built instrumental setup used in this work has already been described in detail elsewhere,[19,20] so only a short explanation is given here. The PI-TOFMS can be operated in two modes: SPI and REMPI. In the first instance, 355 nm photons were initially generated by frequency tripling of the fundamental radiation of 1064 nm of a Nd:YAG-laser (Continuum, Santa Clara, CA, USA, repetition rate 10 Hz, pulse width 3−5 ns, 22.5 mJ at 355 nm), Nd:YAG laser pulses. For REMPI, about 90% of the 355 nm output is guided to an optical parametric oscillator (OPO, VISIR 2 + SHG, GWU-Lasertechnik GmbH, Erftstadt, Germany) by a beam splitter with a thermally stabilized β-barium borate (β-BBO) crystal to create UV-photons for REMPI. Because of the OPO, a UV-photon range from 220 to 355 nm is accessible (Figure 1).

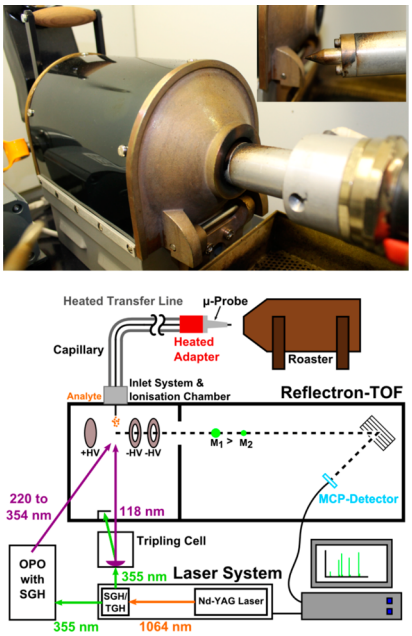

Figure 1. Experimental setup of drum roaster with μ-probe sampling (top) and instrumental setup of PI-TOFMS in SPI and REMPI mode (bottom).

In this study, roasting off-gas REMPI-TOFMS measurements were carried out at 266, 248, and 227 nm to cover a broad range of analyte selectivities. In particular, 266 and 248 nm were easily available by the fourth harmonic generation of a Nd:YAG laser and a KrF laser, respectively, so a tunable laser was not categorically necessary. Additionally, measurements at 227 nm were also carried out because of its higher cross sections for some potential analytes, especially nitrogen-containing compounds. To enable single photon ionization (SPI) at 118 nm, laser pulses were generated by pumping a third harmonic generation (THG) gas cell with the remaining 10% radiation of 355 nm.

The laser beam was focused underneath the inlet needle, which was connected to a heated sampling line (250 °C) to generate an effusive molecular beam in the ion source. Ions were guided into the flight tube of a reflectron TOF-MS (Kaesdorf Instrumente für Forschung und Industrie, Munich, Germany) and detected by a microchannel plate (MCP, Chevron Plate, Burle Electro-Optics Inc.). Two PC cards with a linear intensity range covering 5 orders of magnitude (Acquiris, Agilent Technologies, Basel, Switzerland, 250 MHz, 1 GS/s, 128 kb) enabled acquisition of 10 mass spectra per second.[21] The raw data was finally processed by in-house software based on Labview programming environment (National Instruments, Austin, TX, USA). For data evaluation, 10 consecutive mass spectra were averaged, leading to a final time resolution of 1 s.

DOI: 10.1021/acs.jafc.6b01683
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

Experiments, Data Treatment, and Statistical Analysis. In total, 87 roast experiments were carried out, allocated to roast conditions, and applied photoionization wavelengths as shown in Table 1. A scheme of the overall data treatment and statistical workflow can be found in the section Supporting Information (Figure S1). Before running statistical analysis, the parent ion of caffeine (m/z 194), parent ion with H-loss by ionization (m/z 193), and first two ^{13}C-peaks (m/z 195 and 196) were removed from the mass spectra for two reasons: (1) On the basis of our data, the evaporation of caffeine gives little information about the roast phase because of too high variances between the single roasting experiments. (2) SPI and REMPI are very sensitive ionization techniques for the very abundant caffeine, resulting in high abundances of its respective m/z which hamper statistical analyses. The mass spectra without caffeine-related m/z were subsequently normalized to their total intensity at each point of time to be independent from fluctuations in laser performance. Moreover, overestimation of "overroasting" with significantly higher amounts of roasting off-gas components is avoided.

Table 1. Number of Experiments Carried out for Roasting Conditions, Cultivars, and Photoionization Wavelengths

	Robusta (Vietnam)			Arabica (Mexico)		
	fast	medium	slow	fast	medium	slow
118 nm	4	6	6	5	5	2
227 nm	6			3		
248 nm	9			5		
266 nm	5	5	6	7	7	6

Non-Negative Matrix Factorization (NMF). In previous studies, different temporal evolutions were observed for evolving compounds during coffee roasting.[5,8,11,12,22] Non-negative matrix factorization (NMF) was applied to pool those compounds into classes. Generally, NMF partitions iteratively a non-negative m-by-t matrix $\mathbf{M}$ into a m-by-k matrix $\mathbf{W}$ (hereinafter referred to as factor loadings) and a k-by-t matrix $\mathbf{H}$ (hereinafter referred to as scores). The variable k refers to the rank of the NMF solution but can be practically regarded as the number of processes to identify, i.e., roasting phases, and has to be predefined. In this context, the matrix dimension m stands for m/z and t for the total roasting time. A rank of four was the highest-rank NMF result which does not only cover a mathematical but also a physically meaningful solution. In the following, the four roasting phases are called "evaporation", "early roast", "late roast", and "overroast".

$\mathbf{W}$ and $\mathbf{H}$ are computed by an alternating least-squares (ALS) algorithm to minimize the cost function $f(W,H)_k$

$$f(W, H)_k = \frac{1}{2}\|M - WH\|_F^2 \quad (1)$$

where $\|X\|_F$ computes the Frobenius-norm of a non-negative matrix $\mathbf{X}$:

$$\|X\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{t}|X_{ij}|^2} \quad (2)$$

Consequently, the product of $\mathbf{W}$ and $\mathbf{H}$ is an approximation of the original data matrix $\mathbf{M}$. Because the iteration starts with random initial values $\mathbf{W_0}$ and $\mathbf{H_0}$ for $\mathbf{W}$ and $\mathbf{H}$, the NMF may lead to different solutions when repeated if the algorithm converges in a local minima for $f(W,H)_k$. To improve reproducibility of NMF, initial values for $\mathbf{W_0}$ and $\mathbf{H_0}$ are optimized by a multiplicative update algorithm,[23] which is slower, but more sensitive for initial value optimization, before running the ALS algorithm.[24] The temporal evolution of the roasting stages are illustrated by calculating the relative proportions $\mathbf{h_t}$ of the absolute score values $\mathbf{H_t}$ for each of the k element at any point of time t_i.

$$h_{t_i} = \frac{H_t}{\sum_{i=1}^{k} H_i} \quad (3)$$

Subsequently, the duration of the roasting was converted in percent of total roasting time to ensure comparability between the different temperature profiles of roasting.

In short, relative NMF scores $\mathbf{h_t}$ refer to roasting phase contributions and NMF factor loadings $\mathbf{W_i}$ to representative photoionization mass spectra for the respective roasting phase. The intraclass consistency of the roasting phase determination at each wavelength was proved by principal component analysis (PCA) of temporarily normalized relative score $\mathbf{h_{t,i}}$ and factor loadings $\mathbf{W_i}$.

Fisher Ratio, Euclidean Distance Optimization, and LDA. Real mass spectra at the time of maximum NMF score contribution were extracted from the data matrix of each roasting experiment except for "evaporation". Because of low overall intensities, mean spectra over the whole phase "evaporation" instead of spectra from maximum relative NMF scores were chosen. Thereby, four m-by-n matrices $\mathbf{P_j}$ (m, m/z; n, number of roasting experiments at one roasting condition and photoionization wavelength; j, roasting phase) containing real mass spectra of maximum roasting phase contribution were obtained for each set of roasting condition and photoionization wavelength. Single and double outliers of m/z for every matrix $\mathbf{P_j}$ were removed by Grubbs' Test[25] at a significant level $\alpha = 0.05$. Subsequently, Fisher ratios $F_{m/z}$[26] for every m/z were calculated

$$F_{m/z} = \frac{(\bar{m}_i - \bar{m}_{j+1})^2}{\mathrm{var}_j + \mathrm{var}_{j+1}} \quad (4)$$

(where $\bar{m}_i$ and $\bar{m}_{i+1}$ correspond to the mean intensities of m/z, and var_i and var_{i+1} to the variances of consecutive roasting phases j and $j+1$ in $\mathbf{P_j}$, respectively) to figure out m/z with most different behavior in two consecutive roasting phases. According to eq 4, the Fisher ratio $F_{m/z}$ becomes large if the difference of the means between two phases is high and the intraphase variance is small. Only m/z with abundances above the detection limit in more than 50% of the experiments were considered. Plotting the intensities of two m/z of roasting phase j versus roasting phase $j+1$, two point clouds were obtained (not shown) which cover a Euclidean distance d between its centers. Distances d_i were calculated for every possible combination of the five pairs of m/z with highest $F_{m/z}$ for a roasting phase transition; pairs of equal m/z were not considered. For linear discriminant analysis (LDA), the combination of m/z with the maximum sum of d_i was chosen. The linear classifier function was further used to calculate dynamic upper or lower limits of one m/z based on the intensity of a second m/z. The exceedance or deceedance of that limit determines a transition and the beginning of the next roasting phase $j+1$. A roasting phase transition is defined if three classifier functions have been crossed.

NMF, LDA, and PCA were performed with Matlab 2014b Statistic Toolbox (The MathWorks, Natick, MA, USA).

RESULTS AND DISCUSSION

Temperature−Time Profiles. The temperature was recorded for each roast experiment. Figure 2 (bottom left) displays the averaged values together with the corresponding standard deviation for the three roasting profiles.

Starting from 200 °C, the temperature declined after filling the beans into the roaster because of a heat uptake by the beans. Depending on the chosen heating step, the temperature starts to reincrease for "fast roast" 2 min after filling (highest heating step), followed by "medium" and "slow roast" (lowest heating step).

These temperature trends agree well with previous studies of Wieland et al. and Gloess et al.[8,9]

Qualitative Detection of Volatile Species in Roasting Off-Gas at Different Wavelengths. The assignments of the m/z values are based on previous studies with photoionization[4,10−12] (and references therein) and only described briefly at this point. Although the ionization energies and

DOI: 10.1021/acs.jafc.6b01683
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

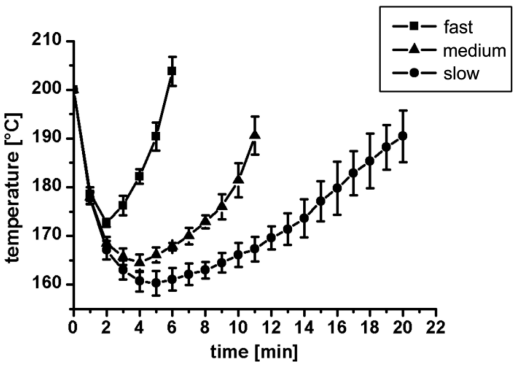

Figure 2. Corresponding temperatures and roasting time for "fast", "medium", and "slow roast" with dots representing mean values and error bars the standard deviation ($\pm\sigma$).

photoionization cross sections (PICS) of a certain compound shrink the number of possible analytes, it is not possible to distinctly identify compounds. Thus, for example, gas chromatography mass spectrometry (GC-MS) studies were used in the mentioned references to assign molecular structures. In cases where more than one substance could be assigned to one m/z, the tentative most likely structure was chosen (Table 2).

All spectra (Figure 3) cover intensities for caffeine (m/z 194), one of the most abundant nonprotein nitrogen-containing compounds in coffee beans. Apart from caffeine, the SPI spectrum contains many compound classes such as carbonyls (m/z 44 acetaldehyde, m/z 86 butanedione, m/z 96 furfural), aromatic, and aliphatic amines (m/z 79 pyridine, m/z 59 C_3-amine), alcohols (m/z 74 pyruvic alcohol, m/z 98 furfuryl alcohol), and thiols (m/z 48 methanethiol).

The REMPI spectra for 266 and 248 nm are dominated by phenolic compounds (m/z 94, m/z 110, m/z 124, m/z 150, and m/z 164) due to the optical selectivity of REMPI for this kind of species. Moreover, heterocyclic compounds like furfural (m/z 96) and indole (m/z 117) were detected as well. Both REMPI spectra at 266 and 248 nm revealed mainly the same substances but with different intensities due to different PIC.

With REMPI at 227 nm, clear spectra are generated with main signals at m/z 59 (C_3-alkylated amine), m/z 99 (succimide), m/z 136 (vinyl-1,2-benzenediole), m/z 150 (4-vinylguaiacol), and m/z 194 (caffeine). The signal for water (m/z 18) is caused by photon-induced electron ionization, which unintentionally occurs when the laser beam hits a metallic surface. Secondary electrons become accelerated in the electrostatic extraction field and lead to the ionization of water molecules that are usually not detectable due to its high ionization energy.

Rapid Discrimination between Arabica and Robusta. In the upper mass range, the pentacyclic diterpenes kahweol (m/z 314), cafestol (m/z 316), and 16-*O*-methylcafestol (m/z 330) can be found, which simplify the discrimination between the cultivars Arabica and Robusta (Supporting Information, Figure S2). Both cultivars contain cafestol, whereas 16-*O*-methylcafestol is only found in Robusta beans. By contrast, the amount of kahweol in Robusta is negligible compared to Arabica.[27] Although kahweol and cafestol are difficult to sample because of their low volatility and stability, they both eliminate water and form anhydrous kahweol (m/z 296) and anhydrous cafestol (m/z 298), which can be observed in SPI mass

Table 2. Assignments of Detected m/z Values for SPI and REMPI during the Whole Roasting Procedure Based on Previous Studies of Photoionization[4,10−12a]

SPI at 118 nm	REMPI at 266 nm	REMPI at 248 nm	REMPI at 227 nm
34 hydrogen sulfide	**66** fragment from phenol	**59** C3-alkylated amines	**59** C3-alkylated amines*
43 $C_2H_3O^+$	**77** fragment from phenolic derivatives	**94** phenol, methylpyrazine	**67** pyrrole*
44 acetaldehyde	**94** phenol, methylpyrazine	**96** furfural	**99** succinimide*
58 acetone, propanal, ethanedial	**109** fragment from guaiacol (M-CH_3)	**110** dihydroxybenzenes, 1,2-benzenediole, methylfurfural	**136** vinyl-1,2-benzenediol*
70 pentene, butenal	**110** dihydroxybenzenes, 1,2-benzenediole, methylfurfural	**117** Indole	**150** vinylguaiacol*
74 butanol, propionic acid, pyruvic alcohol	**117** indole	**120** 2-phenylacetaldehyde	**156** C_2-naphthalene*
79 pyridine	**120** 2-phenylacetaldehyde	**126** hydroxymethylfurfural, benzenetriole	**170** C_3-naphthalene*
86 2,3-butadione, pentanone, pentanal, butanedione, methyl-butenol, butyrolactone	**124** guaiacol, methylbenzenediole	**136** vinyl-1,2-benzenediol	**184** C_4-naphthalene*
95 formylpyrrole, C_2-alkylpyrrole	**126** hydroxymethylfurfural, benzenetriole	**150** vinylguaiacol	**194** caffeine*
98 furfuryl alcohol, octene	**136** vinyl-1,2-benzenediol	**164** dimethxyostyrene*	**198** C_5-naphthalene*
110 dihydroxybenzenes, 1.2-benzenediole, methylfurfural, acetylfuran	**150** vinylguaiacol	**176** 2,2′-methylenbis(5-methylfuran)*	**212** C_6-naphthalene*
126 hydroxymethylfurfural, benzenetriole, maltol	**152** vanillin	**180** caffeic acid*	
128 furaneol			
144 octanoic acid, dihydrohydroxymaltol, phenylfuran	**164** 3,4-dimethoxystyrene	**194** caffeine	
194 caffeine	**176** 2,2′-methylenbis(5-methylfuran)*		
256 hexadecanoic acid	**180** caffeic acid*		
280 linoleic acid	**194** caffeine		
284 octadecanoic acid			
312 eicosanoic acid			
340 docosanoic acid			

[a]Please note that for 227 nm, all molecules were (*) tentatively assigned due to the absence of comparable literature data.

DOI: 10.1021/acs.jafc.6b01683
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

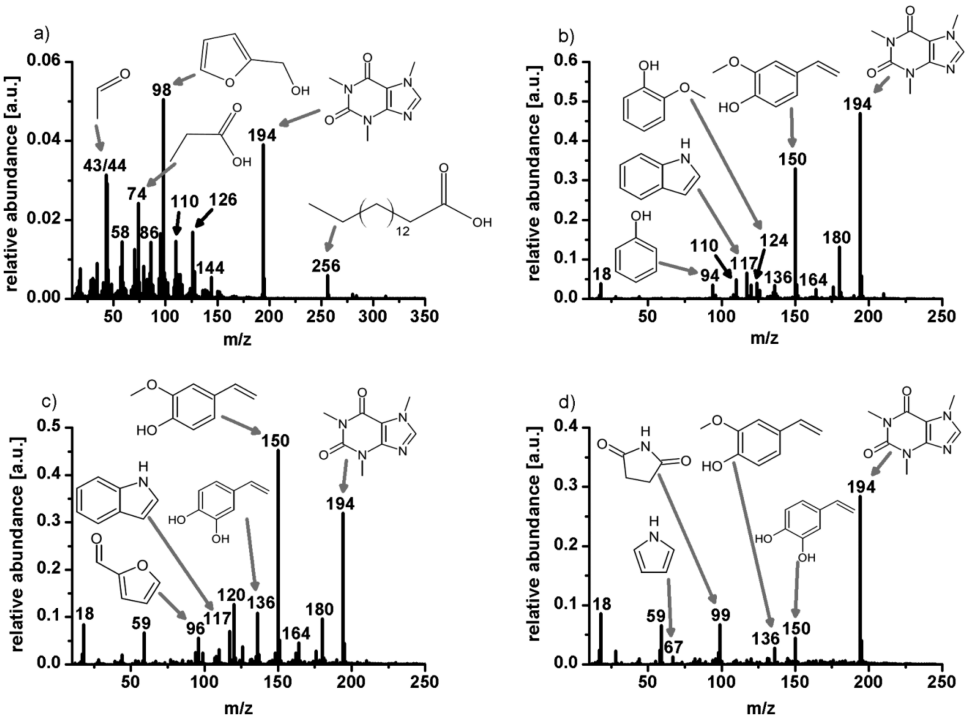

Figure 3. PI mass spectra with (a) SPI at 118 nm, (b) REMPI at 227 nm, (c) REMPI at 248 nm, and (d) REMPI and 266 nm of representative "fast roast" experiments with Arabica beans are depicted. For all mass spectra, the intensities were averaged over the whole roasting time (about 6 min). The most likely chemical structures are assigned to the peaks, illustrating the dependence between analyte selectivity and wavelength.

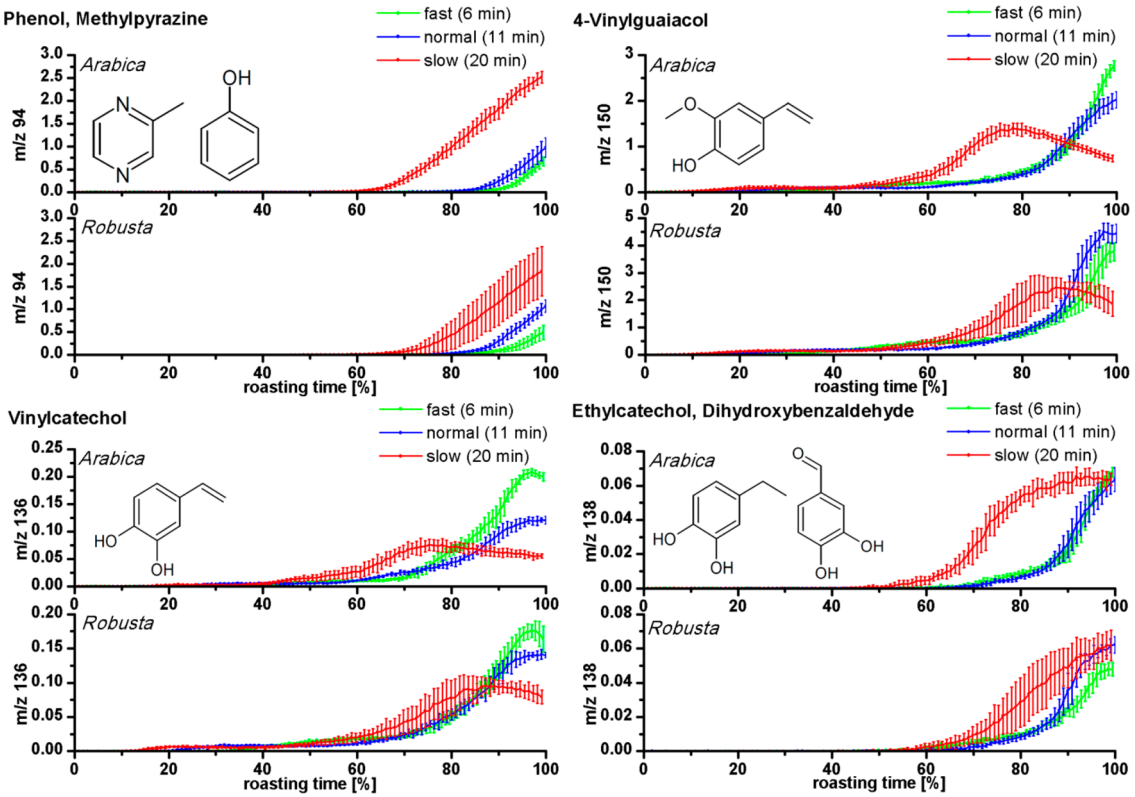

Figure 4. Temporal evolution of phenol/methylpyrazine (top left), 4-vinylguaiacol (top right), vinylcatechol (bottom left), and ethylcatechol/dihydroxybenzaldehyde (bottom right) depending on roasting conditions for both cultivars analyzed by REMPI at 266 nm.

spectra.[4] Therefore, only in the SPI spectrum of Arabica beans both m/z appeared whereas in the Robusta spectrum only m/z 298 was present, so anhydrous kahweol was considered to be a potential marker for Arabica derived from single bean roasting.[4] The REMPI spectra reveal no striking differences at the first sight, but by a closer look at higher mass values (>200 amu),

DOI: 10.1021/acs.jafc.6b01683
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

signals of higher m/z can be observed for Arabica beans at 266 and 248 nm, which was proposed to originate from flavonoids and polyphenols.[12] Small amounts of coffee beans can test for the presences of kahweol and cafestol in a coupled system of a thermal balance and SPI-TOFMS.[28]

Temporal Evolutions of Single Species during Different Roasting Conditions. The high time resolution of online direct mass spectrometric techniques enable the investigation of chemical reactions involved in the coffee roasting process.

Degradation Products of Chlorogenic Acids. Decarboxylation and degradation of chlorogenic acids are important reaction pathways during the roasting of coffee beans. In particular, the degradation of the chlorogenic acid 5-feruloylquinic acid (5-FQA) leads to the formation of phenolic compounds, which can be well monitored by REMPI-TOFMS. On the basis of REMPI at 266 nm, Dorfner et al.[5,10] developed two possible reaction pathways. The "low activation energy" pathway occurs at temperatures below 120 °C at the beginning of the roasting process. First, 5-FQA-lactone is hydrolyzed to ferulic acid, followed by the formation of 4-vinylguaiacol through decarboxylation. Finally, the polymerization at the vinyl groups takes place to form melanoidins. The "high activation energy" reaction pathway occurs later in the advanced roasting process when the water content of the beans has been reduced and the temperature has increased. Under these conditions, the oxidation of 4-vinylguaiacol to vanillin leads to the enhanced formation of guaiacol and phenol detected at the end of one roasting experiment. The different time intensity profiles for vinylguaiacol, phenol, and guaiacol are illustrated in Figure 4. In particular, the slow roast condition led to considerably different profiles than the fast and medium roasting conditions. On the one hand, the intensities of fast and medium roast conditions for both substances increase at the end of the roasting process, on the other hand, the slow conditions lead to profiles that are more in conjunction with the described reaction pathways: the intensity of vinylguaiacol increased due to the decarboxylation of ferulic acid, followed by a decay because of further thermal degradation to phenol and guaiacol.

The time intensity profiles of vinylcatechol and ethylcatechol (Figure 4, bottom) reveal similar trends when comparing the roasting conditions. Slow roast conditions led to noticeable different temporal profiles, whereas medium and fast roasting conditions result in almost similar curves. The profiles of some species for slow roast conditions are in good compliance with findings from Müller et al., who investigated the formation of vinylcatechol by the degradation of caffeoyl quinic acid. Moreover, 4-ethylcatechol and 4-methylcatechol are solely generated by the thermal breakdown of vinylcatechol whereas catechol is formed by the degradation of quinic acid and 4-vinyl-1,2-benzenediol as well.[29]

Either compounds in the coffee roasting off-gas are ingredients of the green coffee bean and evaporate during roasting or originate from chemical conversions. Although not every m/z can be assigned to at least one chemical structure, mass traces can be divided into groups depending on their temporal evolution to identify different stages of the coffee roast.

Systematic Grouping of m/z Traces by NMF. By performing NMF on the time-resolved mass spectrometric data for each wavelength as described in a previous section, the evolution of four roasting phases were identified by sorting of the relative NMF score according to the temporal appearance of the respective maxima (example SPI at 118 nm in Figure 5, bottom left). The roasting seems to be an interaction between different

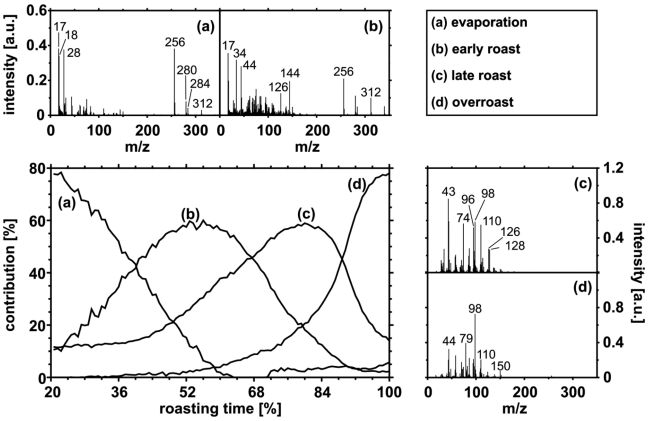

Figure 5. Temporal evolution of four roasting phases of one representative "medium roast" derived from relative NMF scores $\mathbf{h}_t$ (bottom left) with corresponding NMF factor loadings for each roasting phase "evaporation" (a), "early roast" (b), "late roast" (c), and "overroast" (d), analyzed at 118 nm. Time axis was normalized to the reduced total roasting time. The first 20% of the total roasting time were ignored due to absence of detected compounds.

subprocesses whereby one subprocess dominates. To verify the consistency of the obtained temporal evolutions for every roast condition and coffee type, PCA based on correlation was performed on all time-normalized relative NMF scores $\mathbf{h}_t$, i.e., temporal evolutions of the roasting phases, for each wavelength (Figure S3). In the four PCA score subplots, the four roasting phases are clearly separated, whereby different separation performances exist for each wavelength due to different analyte selectivities: for example, "early roast" appears earlier for REMPI at 227 nm than for REMPI at 248 nm. In the PCA score plot of SPI, data at 118 nm overlaps between the point clouds of the phases occur. SPI analyses included not only "fast roast" but also "medium" and "slow roast". In particular, "slow roast" led to another type of time profile for some compounds compared to almost similar time profiles for "fast" and "medium roast" (Figure 4), which lead to different trends of the relative scores and deteriorates the separation power. Interestingly, NMF solution of rank four was not appropriate for "fast" and "medium roast" with REMPI at 266 nm, which is known to be specifically selective for phenolic species originating from the degradation of chlorogenic acids.[10] Hence, solely the degradation of chlorogenic acid is deficient to define limits for the roasting phases because the relative NMF scores $\mathbf{h}_{t,2}$ and $\mathbf{h}_{t,3}$, which were intended to represent "early roast" and "late roast", showed no interpretable temporal evolution. However, a NMF solution of rank three generated reasonable results including the phases "evaporation", "medium roast", and "overroast" (not shown). Finally, we regard SPI data as valuable for further investigations, whereas REMPI can be carried out with less technical effort and fewer signals in the mass spectra for simple monitoring the roasting process.

The NMF factor loadings (SPI at 118 nm in Figure 5, top and right-hand side) can be regarded as representative mass spectra for each roasting phase and were examined by grouping in a PCA to recover known roasting phase indicators and identify possible new markers. In the Supporting Information (Figure S4), PCA biplots of NMF factor loadings from SPI and

DOI: 10.1021/acs.jafc.6b01683
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

REMPI analyses are depicted in which known marker substances can be found for every roasting phase. For example, in SPI spectra, hexadecanoic acid (m/z 256), as one of the most abundant fatty acids in green and roasted coffee,[30] appears together with water (m/z 18, ionized by photoinduced electrons) in "evaporation" due to drying and distillation of the beans while they take up thermal energy in the first 2 to 5 min. A degradation product of carbohydrates is hydroxymethylfurfural (m/z 126), which has its highest contribution between "early" and "late roast" and decays toward "overroast", in agreement with "light roast" in the study of ref 31. Furthermore, catechol (m/z 110) and phenol (m/z 94), which have been identified as degradation products of quinic and caffeic acid,[32,33] occur with their highest abundances during "late roast" and "overroast", similar to the darker roast experiments called "city roast" or "French roast" by Moon et al.[31] Pyridine (m/z 79) is a well-known marker for overroasted coffee beans from the ongoing decomposition of 1-methyl-3-pyridiniumcarboxylat (trigonelline).[34] Because REMPI ionizes only aromatic compounds and to a lower extent aliphatic amines, mainly decomposition products of chlorogenic acids can be observed. With REMPI at 266 nm, our PCA result for "slow roast" in terms of roasting phases (Supporting Information, Figure S4, bottom right) showed a similar picture to Dorfner et al., who derived possible reaction pathways of the most prominent compounds (vinyl-guaiacol (m/z 150), indole (m/z 117), caffeic acid (m/z 180), guaiacol (m/z 124), and phenol (m/z 94)) over the roast.[10] Taking all these indicators into account, we concluded that the NMF of PI-TOFMS data gives chemically sound results concerning the roasting phases. Similar to REMPI at 266 nm, REMPI at 248 nm favors the ionization of monocyclic aromatic compounds. In contrast to the distinct roasting phase separation by the NMF scores $\mathbf{h}_t$ (Supporting Information, Figure S3), no divergence between "early roast" and "late roast" in the space of the first three principal components was obtained by performing PCA on NMF factor loadings (not shown). Hence, slight changes in ratios between m/z determine the roasting phases rather than uniquely appearing m/z (Supporting Information, Figure S4).

When shortening the REMPI-wavelength one step further to 227 nm, the ionization selectivity is shifted toward aliphatic and aromatic amines as well as two-ring aromatic hydrocarbons and low-substituted furans. In particular, amines have been increasingly focused due to their association with enhanced amino acid concentrations in defected coffee beans.[35] The identification of the first two roasting phases was strongly driven by the homologue series of alkylated naphthalenes (m/z 156 C_2- to m/z 212 C_6-naphthalenes), which might originate from pyrolysis of coffee oils or related thermolabile substances on heating elements inside the roaster. The third roasting phase "late roast" was characterized by oxygenated species such as vinylguaiacol (m/z 150), methylfuran (m/z 82), and 4-vinyl-1,2-benzenediol (m/z 136), whereby "overroast" altered the molecular signature of the roasting off-gas to C_3-amines (m/z 59), pyrrole (m/z 67), and methylthiazole (m/z 99).

Construction of Linear Classifiers for Real-Time Roasting Phase Transitions: Toward a PI-TOFMS Roasting Degree Model. The breakdown of NMF results into five pairs of m/z for every phase transition reduces computing time and enables online and real-time roasting phase identification. The m/z of one pair exist in a relation whose limit is described by the linear classifier function. If three of five relations of m/z are exceeded or deceeded, a roasting phase transition is determined. The ascertained roasting phase limits calculated by NMF and LDA are depicted in Figure 6.

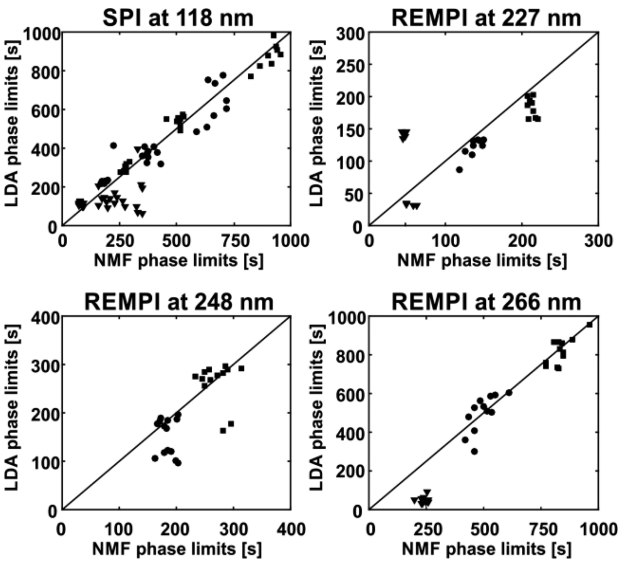

Figure 6. Comparison of the phase limits of "evaporation": "early roast" (triangle), "early roast–late roast" (circle), and "late roast–overroast" (square) determined by NMF and LDA.

At the beginning of the roasting process, the rate of evolving roasting products is small. Because of very low and fluctuating intensities at early roasting times, the residuals for the roasting phase transition between "evaporation" and "early roast" (in blue) becomes relatively high, which was compensated by consideration of an empirical offset of 30 s to prevent a transition determination far away from limits calculated by NMF. For all REMPI analyses, differentiation between the first two phases failed because no m/z, which have distinct higher abundances at the beginning of the roasting or solely appear in the first phase, could be observed. However, all REMPI analyses ended up with results comparable to SPI for the other two roasting phase transitions.

All three roasting conditions were treated the same and submitted to LDA, but especially for the transition between the first two phases large differences occur. In contrast to REMPI, SPI can detect compounds which appear with their highest abundance during the start of roasting, such as fatty acids. Taking the different overall intensities for the three roasting conditions into consideration, the large variances for the transition determination between "evaporation" and "early roast" become reasonable. However, when treating all three roasting conditions as single data sets, LDA results become closer to NMF results (Figure 7).

Although the variables submitted to LDA, i.e., the pairs of m/z, have been assigned to most likely compounds in a previous chapter of this study, they must be regarded as signals from photoionization mass spectrometry because possible isobaric compounds as well as the PICs and isotopes affect the observed intensities and consequently the result from the LDA. In addition to this technique, the concept presented in this study can be simply applied to other suitable analytical online techniques, such as repetitive ultrafast gas chromatography–mass spectrometry,[36] and further advanced by quasisimultaneous ionization with SPI and REMPI.[20]

DOI: 10.1021/acs.jafc.6b01683
J. Agric. Food Chem. XXXX, XXX, XXX–XXX

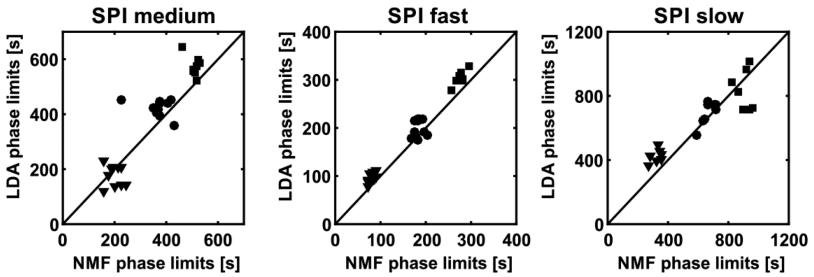

Figure 7. Comparison of the phase limits of "evaporation": "early roast" (triangle), "early roast−late roast" (circle), and "late roast−overroast" (square) determined by NMF and LDA only for SPI at 118 nm and single roasting conditions.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jafc.6b01683.

Statistical workflow and most suitable m/z for classification and corresponding molecular assignments (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +49 (0)381 498−6460. Fax: +49 (0) 381 498−118 6527. E-mail: ralf.zimmermann@uni-rostock.de, ralf.zimmermann@helmholtz-muenchen.de.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge Christian Scholz and Romy Hertz-Schünemann (University of Rostock) for their technical support in coffee roasting and mass spectrometric analyses.

■ REFERENCES

(1) Sunarharum, W. B.; Williams, D. J.; Smyth, H. E. Complexity of coffee flavor: A compositional and sensory perspective. *Food Res. Int.* **2014**, *62*, 315−325.

(2) Flament, I.; Bessière-Thomas, Y. *Coffee Flavor Chemistry*; John Wiley & Sons, Ltd.: Chichester, UK, 2002.

(3) Buffo, R. A.; Cardelli-Freire, C. Coffee flavour: an overview. *Flavour Fragrance J.* **2004**, *19*, 99−104.

(4) Hertz-Schünemann, R.; Streibel, T.; Ehlert, S.; Zimmermann, R. Looking into individual coffee beans during the roasting process: Direct micro-probe sampling on-line photo-ionisation mass spectrometric analysis of coffee roasting gases. *Anal. Bioanal. Chem.* **2013**, *405*, 7083−7096.

(5) Dorfner, R.; Ferge, T.; Kettrup, A.; Zimmermann, R.; Yeretzian, C. Real-time monitoring of 4-vinylguaiacol, guaiacol, and phenol during coffee roasting by resonant laser ionization time-of-flight mass spectrometry. *J. Agric. Food Chem.* **2003**, *51*, 5768−5773.

(6) Yeretzian, C.; Jordan, A.; Badoud, R.; Lindinger, W. From the green bean to the cup of coffee: Investigating coffee roasting by on-line monitoring of volatiles. *Eur. Food Res. Technol.* **2002**, *214*, 92−104.

(7) Lindinger, W.; Hansel, A.; Jordan, A. On-line monitoring of volatile organic compounds at pptv levels by means of proton-transfer-reaction mass spectrometry (PTR-MS) - Medical applications, food control and environmental research. *Int. J. Mass Spectrom. Ion Processes* **1998**, *173*, 191−241.

(8) Gloess, A. N.; Vietri, A.; Wieland, F.; Smrke, S.; Schönbächler, B.; López, J. A. S.; Petrozzi, S.; Bongers, S.; Koziorowski, T.; Yeretzian, C. Evidence of different flavour formation dynamics by roasting coffee from different origins: On-line analysis with PTR-ToF-MS. *Int. J. Mass Spectrom.* **2014**, *365−366*, 324−337.

(9) Wieland, F.; Gloess, A. N.; Keller, M.; Wetzel, A.; Schenker, S.; Yeretzian, C. Online monitoring of coffee roasting by proton transfer reaction time-of-flight mass spectrometry (PTR-ToF-MS): Towards a real-time process control for a consistent roast profile. *Anal. Bioanal. Chem.* **2012**, *402*, 2531−2543.

(10) Dorfner, R.; Ferge, T.; Yeretzian, C.; Kettrup, A.; Zimmermann, R. Laser Mass Spectrometry as On-Line Sensor for Industrial Process Analysis: Process Control of Coffee Roasting. *Anal. Chem.* **2004**, *76*, 1386−1402.

(11) Zimmermann, R.; Heger, H. J.; Yeretzian, C.; Nagel, H.; Boesl, U. Application of laser ionization mass spectrometry for on-line monitoring of volatiles in the headspace of food products: Roasting and brewing of coffee. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1975−1979.

(12) Hertz-Schünemann, R.; Dorfner, R.; Yeretzian, C.; Streibel, T.; Zimmermann, R. On-line process monitoring of coffee roasting by resonant laser ionisation time-of-flight mass spectrometry: Bridging the gap from industrial batch roasting to flavour formation inside an individual coffee bean. *J. Mass Spectrom.* **2013**, *48*, 1253−1265.

(13) Boesl, U. Laser mass spectrometry for environmental and industrial chemical trace analysis. *J. Mass Spectrom.* **2000**, *35*, 289−304.

(14) De Gouw, J.; Warneke, C. Measurements of volatile organic compounds in the earth's atmosphere using proton-transfer-reaction mass spectrometry. *Mass Spectrom. Rev.* **2007**, *26*, 223−257.

(15) Hanley, L.; Zimmermann, R. Light and molecular ions: The emergence of vacuum UV single-photon ionization in MS. *Anal. Chem.* **2009**, *81*, 4174−4182.

(16) Ruosi, M. R.; Cordero, C.; Cagliero, C.; Rubiolo, P.; Bicchi, C.; Sgorbini, B.; Liberto, E. A further tool to monitor the coffee roasting process: Aroma composition and chemical indices. *J. Agric. Food Chem.* **2012**, *60*, 11283−11291.

(17) Liberto, E.; Ruosi, M. R.; Cordero, C.; Rubiolo, P.; Bicchi, C.; Sgorbini, B. Non-separative Headspace Solid Phase Microextraction-Mass Spectrometry Profile as a Marker To Monitor Coffee Roasting Degree. *J. Agric. Food Chem.* **2013**, *61*, 1652−1660.

(18) Hertz, R.; Streibel, T.; Liu, C.; McAdam, K.; Zimmermann, R. Microprobe sampling-Photo ionization-time-of-flight mass spectrometry for in situ chemical analysis of pyrolysis and combustion gases: Examination of the thermo-chemical processes within a burning cigarette. *Anal. Chim. Acta* **2012**, *714*, 104−113.

(19) Mitschke, S.; Adam, T.; Streibel, T.; Baker, R. R.; Zimmermann, R. Application of time-of-flight mass spectrometry with laser-based photoionization methods for time-resolved on-line analysis of mainstream cigarette smoke. *Anal. Chem.* **2005**, *77*, 2288−2296.

(20) Mühlberger, F.; Hafner, K.; Kaesdorf, S.; Ferge, T.; Zimmermann, R. Comprehensive on-line characterization of complex gas mixtures by quasi-simultaneous resonance-enhanced multiphoton ionization, vacuum-UV single-photon ionization, and electron impact ionization in a time-of-flight mass spectrometer: Setup and instrument characterization. *Anal. Chem.* **2004**, *76*, 6753−6764.

(21) Adam, T.; McAughey, J.; McGrath, C.; Mocker, C.; Zimmermann, R. Simultaneous on-line size and chemical analysis of gas phase and particulate phase of cigarette mainstream smoke. *Anal. Bioanal. Chem.* **2009**, *394*, 1193−1203.

DOI: 10.1021/acs.jafc.6b01683
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

(22) Baggenstoss, J.; Poisson, L.; Kaegi, R.; Perren, R.; Escher, F. Coffee roasting and aroma formation: Application of different time-temperature conditions. *J. Agric. Food Chem.* **2008**, *56*, 5836−5846.
(23) Lee, D. D.; Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature* **1999**, *401*, 788−791.
(24) Berry, M. W.; Browne, M.; Langville, A. N.; Pauca, V. P.; Plemmons, R. J. Algorithms and applications for approximate nonnegative matrix factorization. *Computational Statistics and Data Analysis* **2007**, *52*, 155−173.
(25) Massart, D. L.; Vandeginste, B. G. M.; Buydens, L. M. C.; de Jong, S.; Lewi, P. J.; Smeyers-Verbeke, J. *Handbook of Chemometrics and Qualimetrics*, 1st ed.; Elsevier Science: Amsterdam, 1997; Part A, p 112f.
(26) Fisher, R. A. The use of multiple measurements in taxonomic problems. *Annals of Eugenics* **1936**, *7*, 179−188.
(27) Kurzrock, T.; Speer, K. Diterpenes and diterpene esters in coffee. *Food Rev. Int.* **2001**, *17*, 433−450.
(28) Fischer, M.; Wohlfahrt, S.; Varga, J.; Saraji-Bozorgzad, M.; Matuschek, G.; Denner, T.; Zimmermann, R. Evolved gas analysis by single photon ionization-mass spectrometry. *J. Therm. Anal. Calorim.* **2014**, *116*, 1461−1469.
(29) Müller, C.; Lang, R.; Hofmann, T. Quantitative precursor studies on Di- and trihydroxybenzene formation during coffee roasting using ″in bean″ model experiments and stable isotope dilution analysis. *J. Agric. Food Chem.* **2006**, *54*, 10086−10091.
(30) Budryn, G.; Nebesny, E.; Zyzelewicz, D.; Oracz, J.; Miskiewicz, K.; Rosicka-Kaczmarek, J. Influence of roasting conditions on fatty acids and oxidative changes of Robusta coffee oil. *Eur. J. Lipid Sci. Technol.* **2012**, *114*, 1052−1061.
(31) Moon, J. K.; Shibamoto, T. Role of Roasting Conditions in the Profile of Volatile Flavor Chemicals Formed from Coffee Beans. *J. Agric. Food Chem.* **2009**, *57*, 5823−5831.
(32) Kamiyama, M.; Moon, J. K.; Jang, H. W.; Shibamoto, T. Role of degradation products of chlorogenic acid in the antioxidant activity of roasted coffee. *J. Agric. Food Chem.* **2015**, *63*, 1996−2005.
(33) Moon, J. K.; Shibamoto, T. Formation of volatile chemicals from thermal degradation of less volatile coffee components: Quinic acid, caffeic acid, and chlorogenic acid. *J. Agric. Food Chem.* **2010**, *58*, 5465−5470.
(34) Stadler, R. H.; Varga, N.; Hau, J.; Arce Vera, F.; Welti, D. H. Alkylpyridiniums. 1. Formation in model systems via thermal degradation of trigonelline. *J. Agric. Food Chem.* **2002**, *50*, 1192−1199.
(35) Oliveira, S. D.; Franca, A. S.; Glória, M. B. A.; Borges, M. L. A. The effect of roasting on the presence of bioactive amines in coffees of different qualities. *Food Chem.* **2005**, *90*, 287−291.
(36) Fischer, M.; Wohlfahrt, S.; Varga, J.; Matuschek, G.; Saraji-Bozorgzad, M. R.; Denner, T.; Walte, A.; Zimmermann, R. Optically Heated Ultra-Fast-Cycling Gas Chromatography Module for Separation of Direct Sampling and Online Monitoring Applications. *Anal. Chem.* **2015**, *87*, 8634−8639.

DOI: 10.1021/acs.jafc.6b01683
J. Agric. Food Chem. XXXX, XXX, XXX−XXX