Reviewer #1:

General comments:This study utilizes ultra-high-performance liquid chromatography-high-resolution mass spectrometry (UHPLC-HRMS) to compare the molecular composition of $PM_{2.5}$ in a tunnel environment with that of exhaust from a single controlled gasoline vehicle on a chassis dynamometer (CD). The results indicate that while the controlled tailpipe emissions can represent the dominant CHO/CHON molecular baseline in tunnel aerosols, they fail to comprehensively cover the sulfur-rich molecular communities (CHOS/CHONS) retained under real-world road conditions, as well as specific components associated with tire wear particles (TWP) and low-molecular-weight CHN. Overall, the study provides rich data and employs methods such as hierarchical cluster analysis (HCA), volcano plots, and Kendrick mass defect (KMD) network analysis to deeply mine the HRMS data, presenting clear arguments. However, the manuscript requires further improvement regarding the representativeness of the experimental design, the rationality of the statistical analysis, and the limitations of the qualitative results.

Response: We appreciate the reviewer's positive assessment of the dataset and molecular composition analysis. In the revised manuscript, we have clarified the rationale for using a single gasoline vehicle as a controlled exhaust reference and the scope of its comparison with tunnel aerosol. We have also revised the statistical analyses and their methodological descriptions, and more clearly distinguished formula-level observations from compound identification and source interpretation.

Major comments

1. Lines 70–90 in the main text indicate that the exhaust samples were derived from three operating conditions of a single gasoline vehicle, whereas the tunnel samples represent the combined effects of a mixed fleet and the tunnel environment. Even though approximately 80% of the vehicles in the tunnel were gasoline-powered, it cannot be assumed that this single test vehicle represents the actual emission profile of the gasoline fleet. Therefore, the molecular formulas "detected in the tunnel but not in the test vehicle" could originate from non-tailpipe emissions or environmental transformations, but they could equally originate from the tailpipe

emissions of other vehicles. The authors are suggested to clearly define the scope of conclusions that can be supported by this single-vehicle exhaust control.

Response: We agree that measurements from a single gasoline vehicle under three operating conditions do not characterize between-vehicle variability in the tunnel fleet. A gasoline passenger car was selected because gasoline-powered cars accounted for approximately 80% of the recorded vehicles (Figure S1). Other recorded categories included gasoline motorcycles, electric vehicles and gasoline/CNG bi-fuel taxis; no diesel vehicles were recorded during the campaign. Even so, formulas retained in tunnel aerosol but not in the tested exhaust may also reflect tailpipe emissions from other vehicles, including other gasoline vehicles, or emissions under operating conditions not examined here, as well as non-tailpipe inputs and environmental transformations.

Accordingly, we have revised the interpretation of the tunnel–CD comparison to distinguish reference coverage from source attribution. Within each ionization mode, "tunnel-only" (T–O) now explicitly denotes formulas retained in at least one tunnel sample but in none of the three CD samples. The passage in Section 3.3 linking the T–O composition to additional traffic-related inputs and chemical processing within the tunnel has been revised to describe the observed compositional contrast without assigning its origin. TCR is likewise interpreted as formula-level signal coverage by the tested exhaust reference. Within this scope, the comparison identifies a CHO/CHON-rich formula subset shared by the measured profiles and differences in reference coverage among elemental classes and ionization modes. HCA further describes the temporal and compositional grouping of the selected tunnel formulas. These findings provide molecular targets for further source characterization, while their possible origins require evidence beyond nondetection in the tested vehicle. We have made this distinction explicit in the Abstract, Sections 2.1 and 3.3, and Section 4. The principal revised text now reads as follows:

"The molecular compositions of tunnel $PM_{2.5}$ and tailpipe $PM_{2.5}$ **from one gasoline passenger car under three steady conditions** were compared using high-resolution mass spectrometry." (Abstract, Lines 17–18)

"These findings **show** class- and mode-dependent coverage of tunnel aerosol by the **controlled tailpipe profile** and identify sulfur-rich and other compositionally distinct **groups. Molecular reference profiles for traffic $PM_{2.5}$ should therefore combine controlled tailpipe measurements, tunnel or on-road profiles, and reference materials from lubricants, tire wear and road dust, while oxidation, gas–particle partitioning and secondary formation need to be assessed separately.**" (Abstract, Lines 26–30)

"This study aimed to quantify the molecular coverage of tunnel aerosol by a controlled **single-vehicle** tailpipe profile and to identify chemically distinct molecular **groups** retained in the tunnel." (Introduction, Lines 64–66)

"The selection of a gasoline passenger car was informed by the predominance of gasoline-powered cars in the monitored tunnel fleet (approximately 80%; **Figure S1**). No diesel vehicles were recorded during the tunnel sampling campaign. A **SAIC Volkswagen Lavida 1.6 L** passenger car **(2017 model series) meeting the China V emission standard and** fueled with commercial RON 92 gasoline was operated on a two-roller chassis dynamometer at 0, 20, and 40 km h^{-1}." (Section 2.1, Lines 81–85)

"Within each ionization mode, T–O denotes **molecular formulas** retained in at least one tunnel sample but in none of the three CD samples. CD–O represented the corresponding set difference (Figure 4a, fractions are by formula number unless otherwise noted)." (Section 3.3, Lines 285–287)

The revised definition of TCR is quoted in our response to Reviewer #3, Major Comment 2.

"These contrasts identify molecular groups not covered by the measured reference, while CD nondetection alone cannot establish their sources or formation pathways. Formulas without retained CD signals may also be emitted by other vehicles, including other gasoline vehicles, or under operating conditions not examined here; non-tailpipe inputs and environmental transformations are also possible." (Section 4, Lines 427–431)

"For traffic $PM_{2.5}$ source characterization, controlled tailpipe profiles should therefore be regarded as one component of a broader molecular reference framework rather than a complete representation of traffic aerosol. **The present single-vehicle reference, obtained under three steady-state conditions, does not characterize between-vehicle emission variability or transient emissions during acceleration, deceleration and stop-and-go operation. A broader reference library should encompass** multiple vehicles and fuels, steady and transient operation, non-tailpipe materials, and paired background measurements. As exhaust emissions decrease under tighter controls and fleet electrification, these **non-exhaust and road-associated** processes will remain linked to vehicle activity, vehicle weight, tire composition, road surface conditions, and tunnel ventilation (Timmers and Achten, 2016; Fussell et al., 2022). **The sulfur-containing formulas concentrated in C2** and the formulas associated with **tire wear reference formulas** and low-molecular-mass CHN **formulas** in C3 **provide composition-based targets for developing** complementary traffic aerosol reference profiles, pending **compound-level validation**. Future work should test these candidates using authentic standards, MS/MS spectra, road dust, tire wear, and lubricating oil reference materials, and should integrate **LC–HRMS with GC×GC–MS and targeted LC–MS/MS** to quantify molecular burdens and evaluate marker stability across vehicle technologies, driving conditions, and road environments." (Section 4, Lines 438–450)

2. The tunnel samples were collected using a high-volume sampler at approximately 1.13 m^3/min for 3 hours per sample, whereas the exhaust samples were collected using three parallel MiniVol samplers at 5 L/min. Differences in the total air volume, particle loading, filter size, and matrix between the two sampling systems could reach orders of magnitude. Furthermore, Lines 91–94 state that "a quarter of each quartz fiber filter (4.3 cm^2)" was used for extraction. This area aligns with the 47 mm exhaust filter, but it is unclear if it corresponds to the high-volume tunnel filter. If the mass of particles entering the extraction and injection processes differs significantly, the greater number of T-O formulas and higher "formula richness" observed in the tunnel samples might partially reflect differences in detection limits rather than true chemical diversity.

Related sentences from the original submission

"The quartz filters were extracted and analyzed by UHPLC–HRMS for organic molecular composition in this study." (Section 2.1)

"A quarter of each quartz fiber filter (4.3 cm^2 in area) was extracted three times by sonication for 15 min each using 5 mL of HPLC-grade acetonitrile (Fisher Scientific) per extraction." (Section 2.3)

Response: We agree that the original description of the extracted filter portions was ambiguous and that the loading difference between the two sample sets must be considered. Our analysis shows that the larger T–O set does not result from a detection advantage of the more heavily loaded tunnel extracts. Instead, it reflects cumulative occurrence across 28 tunnel filters compared with three CD samples. Each of the three parallel MiniVol samplers collected one separate filter (0.60 m^3 per filter), so their volumes were not combined. For the tunnel samples, a 47 mm disc was punched from within the particle-deposited area of the large quartz filter, and one quarter of this disc was extracted. For the CD samples, one quarter of the whole 47 mm quartz filter was extracted. The revised method text now reads as follows:

"… to collect $PM_{2.5}$ on one 47 mm polytetrafluoroethylene membrane filter and two 47 mm quartz fiber filters (Whatman). **Each sampler collected one filter; one quartz filter was used for UHPLC–HRMS analysis, and the other for OC measurement.**" (Section 2.1, Lines 89–91)

"**For each tunnel sample, a 47 mm disc was punched from within the particle-deposited area of the quartz filter, and one quarter of this disc (approximately 4.3 cm^2) was extracted. For each CD sample, one quarter of the whole 47 mm quartz filter (approximately 2.84 cm^2 particle-deposited) was extracted. Each portion** was extracted three times by sonication for 15 min using 5 mL of HPLC-grade acetonitrile (Fisher Scientific) per extraction. … concentrated to 1 mL under a gentle stream of nitrogen**, and 5 µL was injected without further dilution or adjustment to equal $PM_{2.5}$ or OC loading.**" (Section 2.3, Lines 108–113)

Based on the recorded net masses on the HRMS filters, each tunnel injection represented approximately 5–18 times more $PM_{2.5}$ than each CD injection (Table R1). Tunnel estimates assume uniform deposition over a reference exposed area of 406.5 cm^2 (40 CFR Part 50, Appendix B, Section 7.1.2); the actual exposed area was not independently verified. Despite their lower estimated $PM_{2.5}$ loading, individual CD samples contained more formulas than the reported tunnel mean in both modes.

Table R1. Estimated filter-$PM_{2.5}$ equivalents per injection and formula counts per sample.

Sample set	$PM_{2.5}$ per 5 µL injection (µg)	Formulas per sample, ESI−	Formulas per sample, ESI+
Tunnel samples	0.98–2.4	363 ± 112	1,061 ± 152
CD (3 conditions)	0.136–0.201	582–731	2,051–2,379

With equal numbers of samples, the 3,276 possible three-filter tunnel subsets also contained fewer formulas (mean union counts of 612.2 in ESI− and 1,596.6 in ESI+) than the three CD samples (978 and 2,948). A higher loading need not yield more detected formulas, because more co-extracted material also intensifies charge competition in ESI and can raise detection thresholds in complex matrices (Nguyen et al., 2013).

Conversely, organosulfur formulas were detected in the CD extracts, but this alone does not establish comparable detectability for all sulfur-containing formulas. Differences in matrix composition may also affect ESI response (Patriarca et al., 2020; Liang et al., 2025). We therefore retain the observed elemental-class contrasts as characteristics of the detected profiles and distinguish cumulative counts from single-sample richness. Section 3.3 now reads as follows:

"T–O contained more formulas than CD–O in both ESI− (1,108 **versus** 666) and ESI+ (2,189 **versus** 1,859)**. These categories describe cumulative detections across 28 tunnel filters and three CD operating-condition samples (Figure 4a).**"

(Section 3.3, Lines 287–289)

The revised Abstract wording now reads as follows:

"In contrast, formulas **detected only in** tunnel **samples** had higher proportions of CHOS and CHONS than those **detected only in** the controlled tailpipe profile."

(Abstract, Lines 19–21)

The corresponding wording in Section 4 now reads as follows:

"**Formula occurrence and TCR characterize the molecular profiles detected under the present sampling and UHPLC–HRMS conditions and are influenced by sampling coverage, filter-associated $PM_{2.5}$ loading, and analytical detectability (Patriarca et al., 2020; Liang et al., 2025).**" (Section 4, Lines 431–434)

Additional manuscript references are listed in the following paragraphs.

Liang, S., Wang, Y., Chen, H., Chan, W., and Yu, J. Z.: Accurate quantification of multifunctional $C_{2–3}$ organosulfates in atmospheric aerosols using liquid chromatography–electrospray ionization mass spectrometry: Overcoming matrix effects and underestimation, Environ. Sci. Technol., 59, 12812–12821, https://doi.org/10.1021/acs.est.5c01846, 2025.

Patriarca, C., Balderrama, A., Može, M., Sjöberg, P. J. R., Bergquist, J., Tranvik, L. J., and Hawkes, J. A.: Investigating the ionization of dissolved organic matter by electrospray, Anal. Chem., 92, 14210–14218, https://doi.org/10.1021/acs.analchem.0c03438, 2020.

References cited only in this response are listed in the following paragraphs.

Nguyen, T. B., Nizkorodov, S. A., Laskin, A., and Laskin, J.: An approach toward quantification of organic compounds in complex environmental samples using high-resolution electrospray ionization mass spectrometry, Anal. Methods, 5, 72–80, https://doi.org/10.1039/C2AY25682G, 2013.
U.S. EPA: Reference method for the determination of suspended particulate matter in the atmosphere (high-volume method), 40 CFR Part 50, Appendix B, Section 7.1.2. https://www.ecfr.gov/current/title-40/chapter-I/subchapter-C/part-50/appendix-Appendix%20B%20to%20Part%2050.

3. The controlled exhaust molecular profile in this study was derived solely from chassis dynamometer tests of a single passenger car using RON 92 gasoline under three steady-state operating conditions (0, 20, 40 km/h). The fleet composition in the real tunnel is highly complex; Figure S1 in the supplement shows that the fleet includes gasoline cars, gasoline motorcycles, dual-fuel taxis, and electric vehicles. The steady-state emission profile of a single vehicle (lacking transient characteristics of acceleration/deceleration and mixed-vehicle types) is highly inadequate as a representative baseline for comparison. It is recommended that the authors discuss this limitation more deeply in the Discussion section and emphasize in the Conclusion that building a comprehensive traffic $PM_{2.5}$ molecular library in the future will require integrated data from more vehicle models and more complex operating conditions.

Response: We agree that the steady-state profile of one gasoline passenger car cannot represent the full emission variability of the mixed tunnel fleet. The tested profile provides a defined exhaust reference for examining molecular-formula overlap and elemental-class differences under the specified conditions. As explained in our response to Reviewer #1, Major Comment 1, the observed fleet composition motivated the selection of a gasoline passenger car without establishing its representativeness of the entire fleet. The revised Section 4 passage specifying these limitations and reference-library requirements is quoted in our response to Reviewer #1, Major Comment 1.

4. The manuscript alternates among the terms "formula," "formula assignment," and "formula-ionization-mode assignment," but it does not explain how cases where the same neutral formula corresponds to multiple retention time peaks are handled. It is also unclear whether TCC (Tunnel-CD Common) matching is based solely on the molecular formula or if it requires alignment of both m/z and

retention time. Structural isomers with the same molecular formula may have different sources; calculating the intersection based solely on the formula would overestimate the true chemical overlap between the exhaust and the tunnel samples. The authors need to clarify whether the comparative unit is the molecular formula or the chromatographic feature.

Response: We agree with the reviewer and now use "formula" consistently, including in the mode-specific tunnel–CD comparison and in HCA; the term "formula–ionization-mode assignment" has been removed. Within each sample and ionization mode, peak areas mapped to the same neutral formula were summed across peaks at different RTs and assigned adduct forms, and each formula was counted once. Although feature alignment used m/z and RT, the occurrence comparison in Figure 4a used neutral-formula identity within each mode without requiring RT correspondence. The original reference-membership rule used for Figure 4b is specified in our response to Reviewer #3, Major Comment 2. This criterion can combine different isomers and therefore favors overlap. With the retained signals, peak-area weights and denominator held fixed, adding RT correspondence could only maintain or reduce TCR. The reported non-overlap is therefore a lower bound with respect to matching stringency. The revised method and comparison text now reads as follows:
"…**Peak areas assigned to the same neutral molecular formula were summed across chromatographic features and adduct forms within each sample and ionization mode before signal screening and blank correction, and each formula was counted once per mode. "Formula" hereafter denotes a neutral molecular formula detected within one ionization mode and is the unit of all counts and comparisons; a neutral formula detected in both modes is counted as one formula in each mode. Formula-level comparisons do not distinguish structural isomers.**" (Section 2.4, Lines 132–136)
"**For the mode-specific occurrence comparison in Figure 4a, TCC was defined as the intersection of formulas retained in at least one tunnel sample and at least one CD sample within the same ionization mode, using neutral-formula identity without requiring retention-time correspondence.**" (Section 3.3, Lines 282–285)
The same term has been applied in Section 2.5, Section 3.4, the Figure 6 caption, Section 4, and the Supplement.

5. Tunnel samples encompass background air inputs, diesel/other vehicle emissions, lubricating oil, brake and tire wear, road dust resuspension, tunnel wall reservoirs, and in-tunnel chemical processes. The manuscript lacks background

samples from outside the tunnel and does not include reference materials for tires, lubricating oil, road dust, or exhaust from other vehicles. Therefore, being "not detected in the CD profile of this single vehicle" merely indicates that these compounds are not covered by this limited reference profile; it does not prove they originate from non-tailpipe emissions or near-road transformations. Sulfur-containing formulas could also be collectively influenced by fuel/lubricating oil, other vehicle types, analytical blanks, and matrix effects.

Response: We agree that nondetection in the CD profile of a single vehicle shows only that a formula is not covered by this reference profile; it does not show a non-tailpipe origin or a near-road transformation. Outside-tunnel background aerosol and reference materials for tires, lubricating oil, road dust, and other vehicles were not collected, so the contributions of the possible origins of the tunnel-only formulas cannot be separated with our data. The data support only a statement about coverage. Sulfur-containing classes had lower TCRs than CHO and CHON, and T–O formulas contained a higher fraction of sulfur-containing classes than CD–O formulas. We have therefore revised the statements in Section 3.3 that linked these formulas to additional traffic-related inputs, in-tunnel chemical processing, or near-road processing. Section 3.3 now reads as follows:

"The higher fraction of sulfur-containing formulas in T–O indicates a chemically distinct sulfur-containing molecular group in tunnel aerosol **that is not covered by the CD profile.**" (Section 3.3, Lines 291–292)

"These distributions indicate that CD–O is enriched in longer-chain species consistent with fresh tailpipe emissions, whereas T–O emphasizes mid-chain constituents **not represented in the CD profile**." (Section 3.3, Lines 325–327)

Regarding the factors that could influence the sulfur-containing formulas, fuel and lubricating oil can reach the particle phase through tailpipe exhaust (Worton et al., 2014) and were not separately characterized, so their presence does not by itself indicate a non-tailpipe origin. Other vehicle types are addressed in our response to Reviewer #1, Major Comment 1. For analytical blanks, a formula was retained only when its sample area was at least five times the blank estimate, calculated from the three tunnel field blanks for tunnel samples and the two CD system-background records for CD samples, and entries failing this criterion are treated as nondetections rather than evidence of absence (Section 2.3). Blank correction removes procedural background but not ambient aerosol entering the tunnel. For matrix effects, sulfur-containing formulas were detected in the CD extracts, but matrix-dependent ESI response may still influence detection, as discussed in our response to Reviewer #1,

Major Comment 2, so the coverage result is stated as conditional on the present sampling and UHPLC–HRMS conditions (Section 4). The collection of the CD blanks is now described in Section 2.1 (see our response to Reviewer #1, Minor Comment 3), whose text now reads as follows:

"**Two system-background blanks were also collected during the CD campaign by sampling the dilution system without tailpipe exhaust input.**" (Section 2.1, Lines 91–92)

6. The TCR (Tailpipe Coverage Ratio) is essentially "the proportion of the tunnel HRMS peak area corresponding to molecular formulas detected at least once in the single-vehicle exhaust." This may not be equivalent to the actual mass contribution or source contribution of the exhaust to the tunnel $PM_{2.5}$. The authors should clarify this distinction.

Related sentences from the original submission "The tailpipe coverage ratio (TCR) was calculated separately within each ionization mode and elemental class as TCC peak area divided by the sum of TCC and T–O peak areas (Figure 4b)." (Section 3.3)

"(b) Tailpipe coverage ratio (TCR), calculated from tunnel peak areas as TCC/(TCC + T–O)." (Figure 4 caption)

Response: We agree that TCR is not equivalent to the mass or source contribution of exhaust to tunnel $PM_{2.5}$. Peak areas were not converted to mass concentrations, and formula sharing alone does not establish a common source. The separately calculated ratios describe class- and mode-specific coverage of the detected tunnel signal by the tested reference. The corresponding Section 3.3 definition and Figure 4 caption are quoted in our response to Reviewer #3, Major Comment 2.

7. The premise that "a controlled single-vehicle exhaust profile cannot fully represent actual traffic aerosols" is inherently a reasonable expectation. The manuscript needs to further highlight findings that go beyond this basic expectation. Please clarify what new scientific insights the clustering analysis provides compared to a direct molecular profile comparison.

Response: We agree that incomplete coverage by a single-vehicle exhaust profile is expected. HCA adds information on how that reference gap is distributed among the selected tunnel formulas. Direct comparison identifies formula occurrence and class- and mode-specific coverage; clustering shows that the unmatched category occurs

within the principal CHO/CHON-rich group as well as in smaller groups with different compositions.

T–O formulas contributed 33.14% and 41.37% of the pooled normalized C1 area in ESI+ and ESI−, respectively (Table S7). Thus, similar predominant elemental classes did not imply complete formula coverage within the principal group. C2 had a higher sulfur-containing formula fraction (52.7%), while two low-molecular-mass CHN formulas contributed 62.98% of C3 ESI+ area. These smaller groups were compositionally informative despite their low signal shares. C1 accounted for 98.22% and 95.35% of the selected ESI+ and ESI− area (Table S6). Tables S7 and S8 are reproduced in our response to Reviewer #3, Minor Comment 4.

Reference development therefore requires both extending formula coverage within the principal group and representing the distinctive composition of smaller groups. Neither objective is captured adequately by overall profile similarity alone.

The groups describe the grouping of formulas within a preselected dataset, and their membership depends on preprocessing and day sampling (see our response to Reviewer #3, Major Comment 7). The revised wording now reads as follows:

"Tailpipe coverage **ratios** were high for CHO and CHON formulas in both ionization modes (0.77–0.82 and 0.84–0.92, respectively) but much lower for the combined CHOS and CHONS class (0.10–0.15). **Hierarchical clustering of selected formulas showed that tunnel-only formulas contributed 33%–41% of the signal of the dominant shared CHO/CHON baseline and formed most of two smaller groups concentrated in tunnel samples, characterized respectively by sulfur-containing formulas and by formulas associated with tire wear reference formulas together with a low-molecular-mass CHN response in positive-ion mode.**" (Abstract, Lines 21–26)

"It contained most formulas shared with the controlled tailpipe **profile. TCC formulas contributed 61.40%** and **58.63%** of the **pooled normalized** C1 peak **area in ESI+ and ESI−, respectively. The corresponding T–O fractions were 33.14%** and **41.37%** (Table **S7**). **Of** the C1 formulas**, 414 (70.77%)** were classified as **T–O (Table S8). Thus, agreement in broad elemental** composition **coexisted with incomplete formula-level reference coverage.**" (Section 3.4, Lines 362–365)

The discussion of group membership and signal magnitudes is quoted in our response to Reviewer #3, Major Comment 7.

"**Whether additional vehicles, fuels and operating conditions recover such formulas remains to be tested.**" (Section 4, Line 431)

The complementary reference targets are specified in the common Section 4 passage quoted in our response to Reviewer #1, Major Comment 1.

Minor comments

1. Line 40: "These processes collectively reduce the specificity of bulk chemical indicators and limit the effectiveness of traditional receptor models based on averaged aerosol composition metrics." The meaning of this sentence is somewhat unclear, and the logical flow of the context needs to be better organized.

Response: The original sentence was intended to explain that bulk aerosol measures can integrate contributions from several traffic-related inputs and can also change during post-emission processing; they are therefore not unique indicators of tailpipe exhaust. The wording did not clearly identify either these inputs and processes or the measures concerned. We have removed the statement on receptor-model effectiveness because this study did not evaluate receptor-model performance. We have replaced the sentence at original Line 40 with an explicit statement of the limitation of bulk measures and the purpose of molecular-formula comparison. The revised text now reads as follows:

"**Consequently, bulk measures such as organic carbon (OC) and OA mass combine organic material from tailpipe exhaust, lubricant- and fuel-related organics, tire and road wear and secondary formation into a single value, and these values and the OC/EC ratio continue to change after emission as partitioning and oxidation proceed. Bulk composition alone therefore cannot show which constituents of real-road traffic aerosol are represented by a tailpipe profile measured under controlled conditions. Molecular-level characterization resolves the detectable fraction of this organic material into individual molecular formulas with defined elemental composition, including nitrogen and sulfur content, and thus allows this comparison to be made at the molecular-formula level.**" (Introduction, Lines 43–49)

2. Could you please provide specific details about the test vehicle, such as the vehicle model, age, mileage, emission standard, engine load, and engine thermal state (e.g., hot/cold start)?

Response: We have added the requested information to Section 2.1 and to a new Table S1 in the Supplement. The test vehicle is a SAIC Volkswagen Lavida 1.6 L passenger car (2017 model series) with an EA211 CSR engine and port fuel injection, meeting the China V emission standard and fitted with a three-way catalytic converter

and no gasoline particulate filter. It was manufactured on 24 December 2018 and first registered in April 2019, and its odometer reading was 9,684 km on 24 June 2019. Engine load was not recorded in the available test records. Sampling at idle, 20 and 40 km h^{-1} followed a hot start, defined as the engine having passed the cold start period of European real driving emissions regulation (European Commission, 2017).

The relevant manuscript text now reads as follows:

"A **SAIC Volkswagen Lavida 1.6 L** passenger car **(2017 model series) meeting the China V emission standard and** fueled with commercial RON 92 gasoline was operated on a two-roller chassis dynamometer at 0, 20, and 40 km h^{-1}. **Sampling was conducted on 24–25 June 2019 under hot-start conditions, after the cold-start period (European Commission, 2017), nominally for 120 min per condition, with 0.60 m^3 collected per filter. Vehicle and sampling details are provided in Table S1.**" (Section 2.1, Lines 83–87)

The new Table S1 is reproduced below.

Table S1. Vehicle and CD sampling information.

Parameter	Information
Vehicle model	SAIC Volkswagen Lavida, previous-generation 1.6 L series
Model series	2017-model series
Manufacturing date	24 December 2018
First registration	April 2019
Engine family and code	EA211, CSR
Engine displacement	1,598 cm^3
Aspiration	Naturally aspirated
Fuel injection	Multi-point port fuel injection (MPI/PFI)
Emission standard	China V
Exhaust aftertreatment	Three-way catalytic converter (TWC)
Gasoline particulate filter	Not **fitted**
Fuel	Commercial RON 92 gasoline
Lubricant	**Not recorded**
Odometer	9,684 km, recorded on 24 June 2019
CD testing dates	24–25 June 2019 (20 km h^{-1} on 24 June; idle and 40 km h^{-1} on 25 June)
Operating conditions	Idle, 20 and 40 km h^{-1}
Engine load	**Not recorded**
Engine thermal state	Hot-start sampling
Sampling duration	Nominally 120 min per condition
Sampling volume	0.60 m^3 per analyzed CD filter
Exhaust dilution ratio	**Not recorded**
Net particulate mass on the corresponding whole quartz filters	161 μg at idle; 126 μg at 20 km h^{-1}; 109 μg at 40 km h^{-1}

Additional manuscript references are listed in the following paragraphs.
European Commission: Commission Regulation (EU) 2017/1154 of 7 June 2017 amending Regulation (EU) 2017/1151 as regards real driving emissions from light passenger and commercial vehicles (Euro 6), Off. J. Eur. Union, L 175, 708–732, 2017.

3. The Experimental Methods section mentions that 3 field blanks were used to correct the tunnel samples. Please clarify whether blank filters were simultaneously collected for the dilution channel or background air during the chassis dynamometer tests and used for background subtraction.

Response: Yes, two system-background blanks were collected during the CD campaign and were used for background subtraction of the CD samples. For each blank, the dilution channel and MiniVol samplers were operated in the same configuration as during the exhaust tests, but without tailpipe exhaust input, for 120 min at 5 L min^{-1} on quartz filters from the same batch. Because the dilution air was drawn from the test hall without filtration, each blank therefore includes contributions from both the dilution air and the sampling system. Section 2.1 now reads as follows:
"**Two system-background blanks were also collected during the CD campaign by sampling the dilution system without tailpipe exhaust input.**"
(Section 2.1, Lines 91–92)
The two CD blanks were extracted and analyzed by UHPLC–HRMS using the same procedure as the tunnel samples. Section 2.3 now reads as follows:
"Three field blanks and two CD system-background blanks were extracted and analyzed using the same procedure as the samples." (Section 2.3, Lines 120–121)

4. In the peak area bar charts in Figure 1, the data bars for the CD tests are multiplied by a scaling factor of 0.2 (x 0.2). It is recommended to briefly explain in the main text or the figure caption why this scaling factor was applied (e.g., because the CD sampling concentration was much higher than the tunnel environmental concentration) to prevent misunderstandings.

Response: We have added the reason for the scaling to the Figure 1 caption. The three CD bars in the peak area rows are multiplied by 0.2 for display because the CD summed peak areas were larger than those of the tunnel samples and would otherwise flatten the tunnel bars. The factor was not used in any calculation and does not correct for sampling volume, dilution or loading, so absolute bar heights are not comparable

between CD and tunnel samples. We also changed the weather wording in the same caption to refer to days, consistent with Section 3.2 (see our response to Reviewer #3, Minor Comment 3).

"The three rightmost bars in each panel denote the CD test samples. In the peak area **rows**, these three bars are **multiplied** by 0.2 **for display because the CD summed peak areas were larger than those of the tunnel samples. The factor was not used in any calculation and does not correct for sampling volume, dilution or loading, so absolute bar heights are not comparable between CD and tunnel samples, whereas class proportions within each bar are unaffected.**" (Figure 1 caption, Lines 180–184)