Alignment for Comprehensive Two-Dimensional Gas Chromatography with Dual Secondary Columns and Detectors

Stephen E. Reichenbach,[*,†] Davis W. Rempe,[†] Qingping Tao,[‡] Davide Bressanello,[§] Erica Liberto,[§] Carlo Bicchi,[§] Stefano Balducci,[∥] and Chiara Cordero[§]

[†]University of Nebraska−Lincoln, Lincoln, Nebraska 68588-0115, United States
[‡]GC Image, LLC, Lincoln, Nebraska 68505-7403, United States
[§]Università degli Studi di Torino, I-10125 Torino, Italy
[∥]La Sapienza University, 00185 Rome, Italy

S *Supporting Information*

ABSTRACT: In each sample run, comprehensive two-dimensional gas chromatography with dual secondary columns and detectors (GC × 2GC) provides complementary information in two chromatograms generated by its two detectors. For example, a flame ionization detector (FID) produces data that is especially effective for quantification and a mass spectrometer (MS) produces data that is especially useful for chemical-structure elucidation and compound identification. The greater information capacity of two detectors is most useful for difficult analyses, such as metabolomics, but using the joint information offered by the two complex two-dimensional chromatograms requires data fusion. In the case that the second columns are equivalent but flow conditions vary (e.g., related to the operative pressure of their different detectors), data fusion can be accomplished by

aligning the chromatographic data and/or chromatographic features such as peaks and retention-time windows. Chromatographic alignment requires a mapping from the retention times of one chromatogram to the retention times of the other chromatogram. This paper considers general issues and experimental performance for global two-dimensional mapping functions to align pairs of GC × 2GC chromatograms. Experimental results for GC × 2GC with FID and MS for metabolomic analyses of human urine samples suggest that low-degree polynomial mapping functions out-perform affine transformation (as measured by root-mean-square residuals for matched peaks) and achieve performance near a lower-bound benchmark of inherent variability. Third-degree polynomials slightly out-performed second-degree polynomials in these results, but second-degree polynomials performed nearly as well and may be preferred for parametric and computational simplicity as well as robustness.
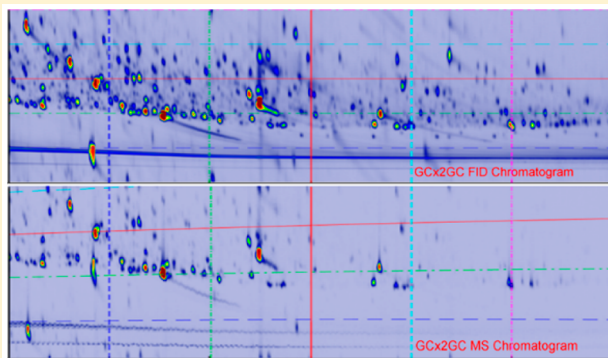
This work demonstrates global transformation functions to align the retention times of chromatograms acquired by comprehensive two-dimensional gas chromatography with one primary column and two parallel secondary columns (GC × 2GC), with a flame ionization detector (FID) for one secondary column and a mass spectrometer (MS) for the other, i.e., GC × 2GC-FID/MS. GC × 2GC increases separation efficiency and information capacity,[1−4] but the two resulting chromatograms may exhibit different retention-time patterns. Then, a retention-times transformation function is required to map between the data-points and/or features such as peaks in the two chromatograms.

Seeley et al. proposed a GC × 2GC system with two secondary columns with different selectivity characteristics, each connected with a separate FID, to "increase the resolution and qualitative information supplied by comprehensive two-dimensional gas chromatographic analysis."[1] They demonstrated results for mixtures of volatile organic compounds (VOCs)

and complex gaseous samples, including exhaled breath.[2] Peroni et al.[5,6] used multiple identical second columns to achieve optimum carrier gas velocity but employed only a single FID. Nicolotti, Bressanello, and Cordero similarly used identical second columns for improved gas linear velocities in both chromatographic dimensions but connected one column to a FID and the other column to a MS with auxiliary flow/pressure controller.[3] In a study of metabolic profiling and fingerprinting with mice urine samples, they concluded that "working in close-to-optimal 2D linear velocities and a doubled secondary column loading capacity, showed positive effects on overall system orthogonality, resolution, and fingerprinting accuracy."[4] The benefits of using both FID and MS in a single run include complementary data, with improved quantification

compared to MS alone and improved compound identification compared to FID alone, and the ability to cross-validate results.

GC × 2GC-FID/MS systems produce two complementary chromatograms for each sample run. The greater information capacity of two detectors is especially useful for difficult analyses, such as metabolomics, but using the joint information offered by the two chromatograms requires data fusion. In the case that the second columns are equivalent but flow conditions vary (e.g., related to the detectors' different operative pressures), data fusion can be accomplished by aligning the chromatographic data and/or chromatographic features such as peaks and retention-time windows. In a GC × 2GC-FID/MS system without auxiliary flow/pressure control at the outlet of the secondary column for the MS, there are differences between the two-dimensional (2D) retention-times patterns of the two detectors.[3] Outlet pressure control adds carrier volumetric flow to reduce these differences, but alignment still is required. The coherence of 2D patterns produced by FID/MS parallel detection was handled by affine transformation implemented in template matching[4] for analysis of the mice urine metabolome but with manual intervention to cross-align FID and MS template patterns. More effective alignment could reduce or eliminate the need for manual intervention.

A fundamental distinction among alignment methods is whether the underlying model is global or local, i.e., whether the geometric differences between chromatograms are characterized by a single function for the entire chromatogram or by a combination of many functions for different regions of the chromatogram. A brief review of local methods used to align GC × GC data is included in the Supporting Information.[7−15] Global functions may be able to capture systemic properties and structure that underlie retention-time differences. On the other hand, local functions may be able to capture retention-time variations that are not related to systemic properties and structure. Local functions typically offer greater representational power than simple global functions, which allows them to capture small-scale variations but also are more susceptible to overfitting of confounding input differences (e.g., compositional differences, artifacts, and noise) and so may be less robust than global functions. Global functions with many parameters have unlimited representational power but also have greater susceptibility to overfitting. Finally, global functions with few parameters are computationally simple. For these reasons, the ideal solution is an alignment method based on a simple global model with few parameters that is able to effectively capture chromatographic differences. Of course, if the chromatographic differences are too complex, a simple global function with few parameters may prove ineffective for modeling those differences. In such circumstances, local functions may be required.

In an early study of global functions for modeling chromatographic differences in GC × GC data, Ni et al.[16] found that four-parameter separable shift-and-scale and six-parameter bivariate affine (shift, scale, and shear) functions were effective for modeling retention-time differences related to oven temperature ramp rate and inlet gas pressure. They noted that these functions, optimized for least-squares fit, would be effective for pattern matching and normalizing retention times to retention indices. Van Mispelaar[17] parameterized a global bivariate second-degree polynomial function (with 12 parameters) using the locations of six peaks (supplying the minimum number of constraints for the second-degree polynomial), then used bilinear interpolation to resample the data values.

The goal of this research is to evaluate the performance of simple global alignment models for GC × 2GC in order to determine their representational effectiveness. As described in the Experimental Section, performance is quantified by the root-mean-square error (RMSE) for retention times of matched peaks in paired FID and MS chromatograms from a single GC × 2GC run. The performance benchmark is the RMSE between matched peaks for the same detector (i.e., FID with FID and MS with MS) in consecutive replicate runs, with the expectation that consecutive replicate runs exhibit negligible sample and chromatographic differences. Then, small misalignments for consecutive replicate runs can be deemed noise, thereby providing a performance benchmark. If the performance of a simple global method to align FID and MS chromatograms from GC × 2GC can approach the benchmark, then the residual differences can be attributed to noise rather than chromatographic misalignment. The experiments employed cross-validation to evaluate three simple global 2D alignment methods: affine, second-degree polynomial, and third-degree polynomial transformations.

■ EXPERIMENTAL SECTION

Samples. The urine samples were gathered for the Italian Diabetes Exercise Study 2 (IDES_2), which is assessing the effect of a behavioral intervention strategy on the promotion and maintenance of physical activity in adults with type 2 diabetes. IDES_2 is a randomized clinical trial that monitors objective measurable changes in sedentary time and physical activity over a 3-year period after behavioral intervention as compared with usual care. The study also monitors physical fitness, modifiable cardiovascular risk factors (HbA1c, lipids, blood pressure, C-reactive protein), and health related quality of life. The samples analyzed by GC × 2GC are for the first and fourth quartile of physical activity objectively measured at the baseline. Sample preparation followed a standard derivatization protocol of oximation/silylation.[18] Derivatization reagents, *O*-methylhydroxylamine hydrochloride (MOX) and *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA), and HPLC-grade solvents, methanol, pyridine, *n*-hexane, and dichloromethane, were supplied by Sigma-Aldrich (Milan, Italy).

Calibration standards of pyruvic acid, lactic acid, malonic acid, succinic acid, malic acid, 2-ketoglutaric acid, hippuric acid, L-alanine, L-valine, glycine, L-threonine, L-tyrosine, creatinine, phenylalanine, xylitol, ribitol, glycerol, fructose, galactose, glucose, mannitol, and myo-inositol; and the internal standard (ISTD) 4-fluorophenylalanine were from Sigma-Aldrich (Milan, Italy). Calibration solutions for quantitative determination of relevant analytes were prepared as in a previous protocol[4] at 2 mg/L, 10 mg/L, 50 mg/L, and 100 mg/L. The ISTD for data normalization and quality control, 4-fluorophenylalanine, was at 10 mg/L.

Instrumentation. Separations were performed with an Agilent 6890-5975C GC/MS fast quadrupole. The first-dimension (^{1}D) column was SE52 (30 m × 0.25 mm internal diameter (^1d_c) × 0.25 μm film thickness (^1d_f)) and the two parallel second-dimension (^{2}D) columns were OV1701 (1.4 m × 0.10 mm ^2d_c × 0.10 μm ^2d_f). Columns were from Mega (Legnano, Milan, Italy). The carrier gas was helium with an initial head pressure of 296 kPa (constant flow). The modulation period was 5 s with pulse time 0.350 ms. The FID acquisition frequency was 100 Hz and the MS performed full scans (50−350 m/z) at 24 Hz. The oven programming was 50 °C (1 min) to 300 °C (10 min) at 4.0 °C/min. A

microfluidic three-port splitter (SilFlow, SGE Analytical Science, Ringwood, Victoria, Australia) split the flow from the ^{1}D column into the two parallel ^{2}D columns where the first 0.60 m were wrapped together in the loop-type thermal modulator (Zoex Corporation, Houston, TX). A deactivated, restrictor capillary (0.17 m × 0.1 mm d_c) was connected to the outlet of the ^{2}D column for the MS.

Data Preprocessing. Preprocessing was performed with GC Image GCxGC Edition Software (R2.6 prerelease, GC Image, LLC, Lincoln, NE). The preprocessing operations were start-time adjustment, modulation-phase rotation, baseline correction, and peak detection.[19] For any two chromatograms, corresponding peaks were paired by preliminary bidirectional template matching and then interactive selection using a new graphical user interface (GUI) for interactive template matching and transformation.[20] For the calibration samples, 25 peaks were paired. For the urine samples, 156 peaks were paired. For each set of N_p paired peaks, the retention times of the peaks in the target chromatogram are denoted (x_i, y_i) and the peaks in the reference chromatogram are denoted (x'_i, y'_i) where i is the peak index from 1 to N_p.

Evaluation Metric. The difference of retention times for paired peaks in two chromatograms is two-dimensional, with a difference in each chromatographic dimension. The RMSE is computed as the square roots of the means of the squared relative retention-times differences:

$$\text{RMSE} = \left(\sqrt{\frac{1}{N_p} \sum_1^{N_p} (x_i - x_i')^2}, \sqrt{\frac{1}{N_p} \sum_1^{N_p} (y_i - y_i')^2} \right) \quad (1)$$

The retention-time differences in each dimension also are expressed relative to the average detected peak-width in that dimension (w_x, w_y), based on the logic that the peak widths are fundamental to chromatographic resolution. The relative RMSE is

$$\text{RRSME} = \left(\frac{\text{RMSE}_x}{w_x}, \frac{\text{RMSE}_y}{w_y} \right) \quad (2)$$

Transformation Models. Experiments evaluated initial misalignment (i.e., no transformation) and the performance of three transformation models, each generalizing the previous model with additional parameters. For each model, if the number of peak-pairs overconstrains the model, the model parameters are determined to minimize the RMSE for the peak-pairs set.[21]

For notational completeness, the identity transformation, which gives the initial misalignment, is defined as

$$f_0(x, y) = (x, y) \quad (3)$$

The affine transformation is linear scaling and shearing plus translation:

$$f_1(x, y) = (t_x + s_x x + h_x y, \ t_y + h_y x + s_y y) \quad (4)$$

where (s_x, s_y) are the scale parameters, (h_x, h_y) are the shear parameters, and (t_x, t_y) are the translation parameters. The affine transform preserves parallel lines and distance ratios. A minimum of three noncollinear two-dimensional peak-pairs (one peak of each pair from each chromatogram) are required to parameterize the affine transformation.

The second-degree polynomial adds three additional terms in each dimension:

$$f_2(x, y) = (t_x + s_x x + h_x y + a_x xy + b_x x^2 + c_x y^2, \ t_y + h_y x \\ + s_y y + a_y xy + b_y x^2 + c_y y^2) \quad (5)$$

A minimum of six peak-pairs are required to parameterize the second-degree polynomial transformation.

The third-degree polynomial adds four additional terms in each dimension:

$$f_3(x, y) = (t_x + s_x x + h_x y + a_x xy + b_x x^2 + c_x y^2 + \alpha_x x^2 y \\ + \beta_x xy^2 + \gamma_x x^3 + \delta_x y^3, \ t_y + h_y x + s_y y + a_y xy + b_y x^2 \\ + c_y y^2 + \alpha_y x^2 y + \beta_y xy^2 + \gamma_y x^3 + \delta_y y^3) \quad (6)$$

A minimum of ten peak-pairs are required to parameterize the third-degree polynomial transformation.

For GC × 2GC data, these models could be simplified by presuming that the first-column retention-times are identical and so transforming only the second-column retention times. However, for completeness, the experimental results shown here use the general two-dimensional models.

Because each successive transformation model subsumes the representational power of the lower-order models, performance is monotonically nondecreasing with order; i.e., higher-order polynomials always do at least as well as lower-order polynomials for a given set of peak-pairs. With no limit on the polynomial degree, arbitrarily large peak-pairs sets can be transformed with no residual error. However, if a lower-order polynomial successfully models the actual chromatographic transformation, then additional parameters merely increase overfitting and incur additional computation. So, for robustness and computational efficiency, lower-degree polynomials are preferred if they are representationally effective.

Evaluation Methodology. Given two chromatograms, a set of corresponding peak-pairs is used to determine the optimal transformation, so the RMSE of those points after the transformation is a biased indicator of performance for other points. In particular, the transformation is the best fit for the specific peak-pairs, including noise inherent in their peak locations. That the transformation fits the noise as well as the relevant chromatographic differences is called overfitting. Constraining the transform with many points tends to attenuate overfitting, but it is important to assess the level of overfitting, particularly for small peak-pairs sets.

Cross-validation can be used to provide an unbiased estimate of the transformation performance. In cross-validation, a peak-pairs set is partitioned into a training set, which is used to determine the transform, and a testing set, which is used to independently evaluate performance. To account for variability, results over multiple rounds of cross-validation, each performed using randomly generated partitions of the peak-pairs set, are averaged. Gros et al.[13] used leave-one-out cross-validation (i.e., training sets with all but one peak-pair), but the size of the training set can be varied to assess performance relative to the number of peak-pairs used to determine the transformation. Performance measures for both the training and testing sets are reported.

Each cross-validation result is computed across random partitions, for each transformation method (including no transformation), for both the training set and the testing set, at each training set size from 3 peak-pairs (the minimum size
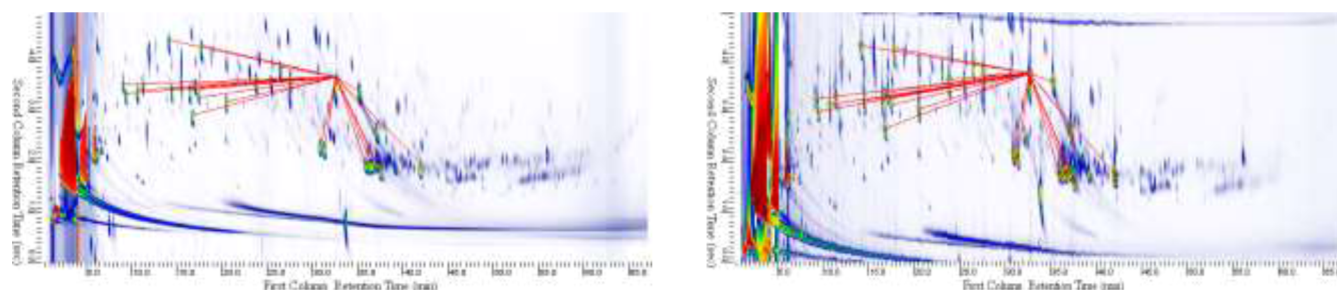
Figure 1. Two chromatograms from a GC × 2GC analysis of a calibration sample: (A) FID on the left and (B) MS on the right. Peaks in the 25 peak-pairs set are indicated with open circles. Lines connect each target peak to the ISTD, phenylalanine. Clockwise around the ISTD peak (from 12 o'clock in the MS chromatogram), the targets are L-tyrosine (1), Myo-inositol, L-tyrosine (2), galactose, D-mannitol, D-glucose, D-fructose, ribitol, xylitol, glycerol, L-threonine, L-glycine, lactic acid, L-leucine, L-valine, L-alanine, L-proline, creatinine, pyruvic acid, malic acid, L-phenylalanine, malonic acid, succinic acid, and 2-ketoglutaric acid.
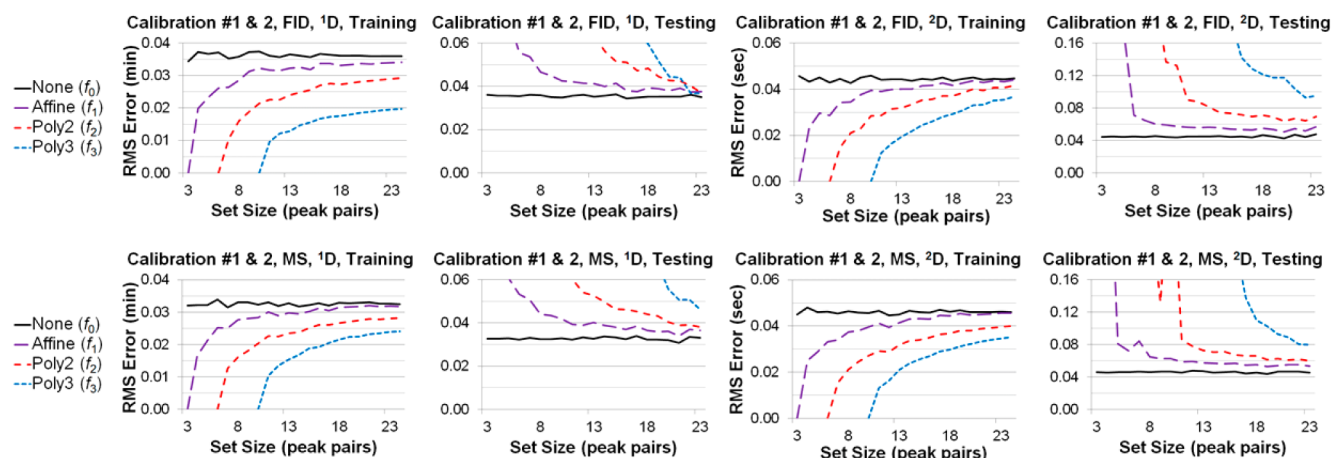
Figure 2. Cross-validation RMSE results as a function of the training set size for alignment of consecutive, replicate calibration chromatograms from the same detector. Columns from left to right are for ^{1}D with the training set, for ^{1}D with the testing set, for ^{2}D with the training set, and for ^{2}D with the testing set. The top row is for the FID chromatograms of calibration runs no. 1 and no. 2, and the bottom row is for the MS chromatograms of calibration runs no. 1 and no. 2.

for the affine transformation) to the total number of peak-pairs (which does not allow for a testing set), and for both directions of transformation (i.e., switching the target and reference chromatograms). The cross-validation RMSE is computed as

$$\overline{RMSE} = \left(\sqrt{\frac{1}{KN_P} \sum_1^K \sum_1^{N_P} (x_i - x_i')^2}, \sqrt{\frac{1}{KN_P} \sum_1^K \sum_1^{N_P} (y_i - y_i')^2} \right) \quad (7)$$

where K is the number of iterations of random partitioning. For the graphs shown in the Results and Discussion, $K = 100$ and the $\overline{RMSE}$ values for the transformations in both directions of the same cross-validation partition are averaged. Additional statistics related to worst-case performance (maximum RMSE) are included in the Supporting Information.

Performance Benchmarks. The goal is a transformation between FID and MS chromatograms from the same GC × 2GC run. The performance of each transformation model for this task is compared to a benchmark computed as the RMSE between matched peaks for the same detector (i.e., FID with FID and MS with MS) in consecutive replicate runs. Consecutive replicate runs have negligible sample and chromatographic differences (regarded as noise), so any transformation model for them should provide only negligible

reductions in RMSE. If the performance of a transformation model to align FID and MS chromatogram pairs can approach the benchmark of consecutive replicate runs, then the residual differences can be attributed to noise rather than chromatographic misalignment.

Aligning Chromatographic Features and Data. Transformation models are sufficient to map retention-times of peaks, windows, or other features from one chromatogram to align with another chromatogram. However, mapping data values in one chromatogram (or an entire chromatogram) to align with another chromatogram requires additional steps of resampling (and interpolation or approximation). The additional steps to align data values are not discussed here, but the alignment RMSE of data points left out during cross-validation is an indicator of retention-time mapping quality for data points which are not in the training set.

The computation required for low-degree polynomial mapping is small. For example, on a computer with an Intel i7-4770 CPU with 4 cores at 3.4 GHz and 16 GB RAM and Microsoft Windows 10 64-bit OS, Java bytecode execution required about 0.018 ms for optimal fitting of a second-degree polynomial to 156 matched peak-pairs (which is a fairly large number of alignment peaks) and about 0.067 ms to apply the second-degree polynomial to remap the retention times of 156 peaks (which is a rate of more than 2 M points per second). Presumably, compiled code could be faster.
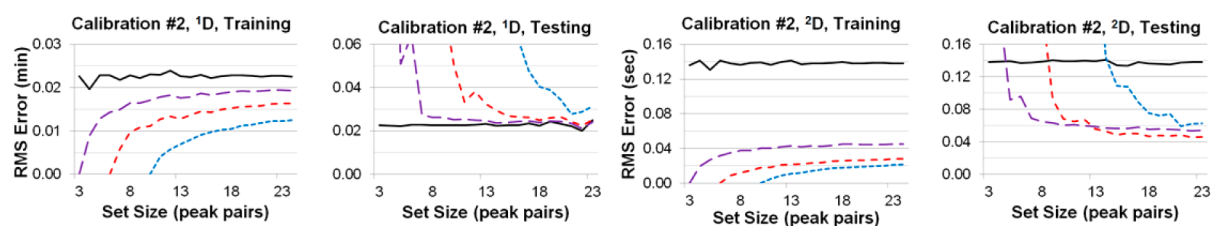
Figure 3. Cross-validation RMSE results as a function of the training set size for alignment of GC × 2GC calibration chromatograms (from different detectors). Columns from left to right are for ^{1}D with the training set, for ^{1}D with the testing set, for ^{2}D with the training set, and for ^{2}D with the testing set. The rightmost graph demonstrates the effectiveness of the transformations.
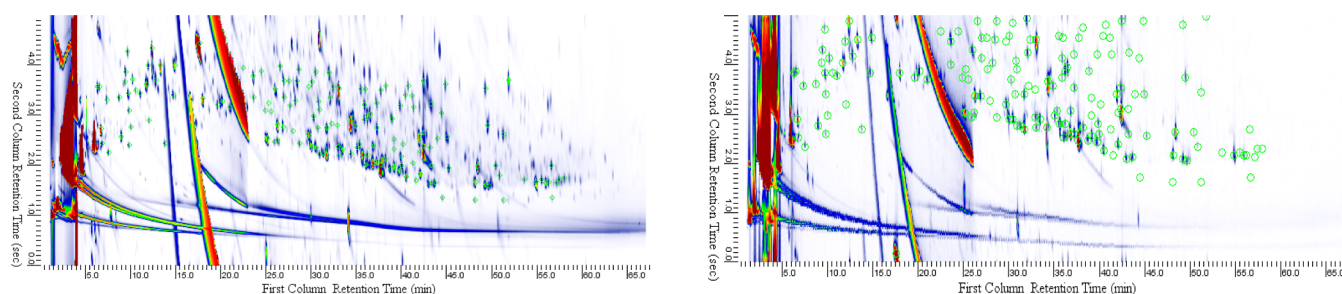
Figure 4. Two chromatograms from GC × 2GC analysis of a urine sample: (A) FID on the left and (B) MS on the right. Peaks in the 156 peak-pairs set are indicated with open circles.

RESULTS AND DISCUSSION

Results for Calibration Chromatograms. The first experiments presented here evaluated calibration runs with a set of informative target metabolites to be quantified in the human urine samples. Figure 1 shows images of the FID and MS chromatograms from one of the calibration runs (with concentration 100 mg/L). Open green circles indicate the analyte peaks used for alignment experiments. Red lines connect each analyte peak to the ISTD peak, which is indicated with a black circle. Peak identities and retention times are provided as Supporting Information.

The average peak widths, computed as the second-central moments (i.e., standard deviations), were determined from these peaks in six FID and MS chromatograms of three consecutive replicate calibration runs. The average peak standard-deviation widths were about $^1\sigma = 0.060$ min and $^2\sigma = 0.085$ s. In ^{1}D, the standard-deviation peak width was slightly less than the modulation cycle, $P_M = 0.083$ min, which has implications for the alignment benchmark. The standard deviation for randomly distributed positions over a single modulation interval is $12^{-1/2} \times P_M$, which is about 0.024 min (or $0.4 \times {}^1\sigma$) for these data. This is the RMS noise level from the sampling effect of modulation. In ^{2}D, the average peak width is about 8.5 times the sampling rate for the FID data and twice the sampling rate for the MS data. The RMS noise level from detector sampling is about 0.003 s (or $0.03 \times {}^2\sigma$) for the FID data and 0.012 s (or $0.14 \times {}^2\sigma$) for the MS data.

The benchmarks for the calibration runs were evaluated for consecutive replicate runs with respect to the same detector. As can be seen in Figure 2, for consecutive replicate runs, there was a small misalignment of peak-pairs (shown by the solid black lines labeled "None (f_0)"): for ^{1}D, about 0.035 min or $0.6 \times {}^1\sigma$; and, for ^{2}D, about 0.045 s or $0.5 \times {}^2\sigma$. In ^{1}D, the misalignment is only a little more than the stochastic modulation-sampling noise level (0.024 min). With the results in the Supporting Information, the ^{2}D misalignment averaged slightly more for MS than FID (0.045 s versus 0.039 s), which

is in line with the difference in detector sampling noise (a difference of 0.009 s). When aligning FID and MS chromatograms from the same run, we would like to approach these benchmarks: 0.035 min in ^{1}D and 0.045 s in ^{2}D.

The expectation that the misalignment in chromatograms from the same detector for consecutive replicate runs is due to noise is supported by the observation that the global transformations produced no improvements in alignment of the testing-set peaks. Figure 2 illustrates the performance of the alignment methods for both the training and testing sets (shown as colored, dashed lines). When the models are not overconstrained (peak-pairs sets of size 3, 6, and 10, respectively, for f_1, f_2, and f_3), they are able to perfectly align the peak-pairs in the training set. Those models are computed to fit the retention-times noise in the training-set peaks as well as any underlying transformation, so this perfect performance is misleading, as evidenced by the performance for the testing-set peaks. As more peak-pairs are included in the training set (which reduces overfitting to noise), the performance for the training set worsens, but the more meaningful performance for the testing-set peaks improves. However, the performance gains from using additional peak-pairs to optimize the model diminish as the model is more and more constrained; and, even with all but one of the peaks in the training set, none of the models improves alignment for the testing set for consecutive replicate runs.

Figure 3 illustrates the performance for aligning the two chromatograms from a GC × 2GC-FID/MS calibration run. As can be seen in the second of the four graphs, for ^{1}D, none of the models yields any improvement in the testing set, but this is expected because both the FID and MS chromatograms were generated for the same ^{1}D separation. Note also that the ^{1}D misalignment is approximately equal to the stochastic modulation sampling noise level (0.024 min). For ^{2}D, as seen in the right-most graph, all of the transformation models significantly improved alignment from about 0.14 s (or $1.6 \times {}^2\sigma$) before transformation to about 0.06 s (or $0.7 \times {}^2\sigma$) or less, with the second-degree polynomial approaching the benchmark
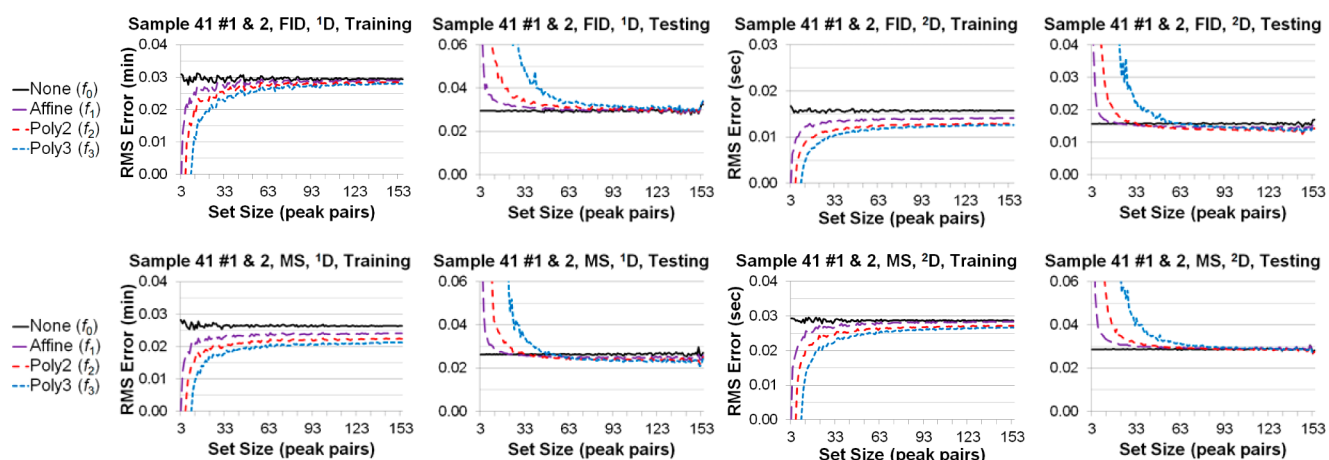
Figure 5. Cross-validation RMSE results as a function of the training set size for alignment of consecutive, replicate urine sample chromatograms from the same detector. Columns from left to right are for ^{1}D with the training set, for ^{1}D with the testing set, for ^{2}D with the training set, and for ^{2}D with the testing set. The top row is for the FID chromatograms of sample 41, runs no. 1 and no. 2, and the bottom row is for the MS chromatograms of sample 41, runs no. 1 and no. 2.
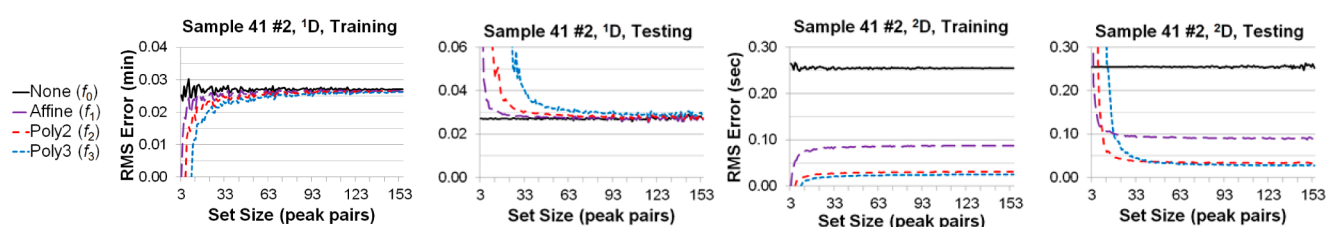
Figure 6. Cross-validation RMSE results as a function of the training set size for alignment of GC × 2GC urine sample chromatograms (from different detectors). Columns from left to right are for ^{1}D with the training set, for ^{1}D with the testing set, for ^{2}D with the training set, and for ^{2}D with the testing set. The rightmost graph demonstrates the effectiveness of the transformations.
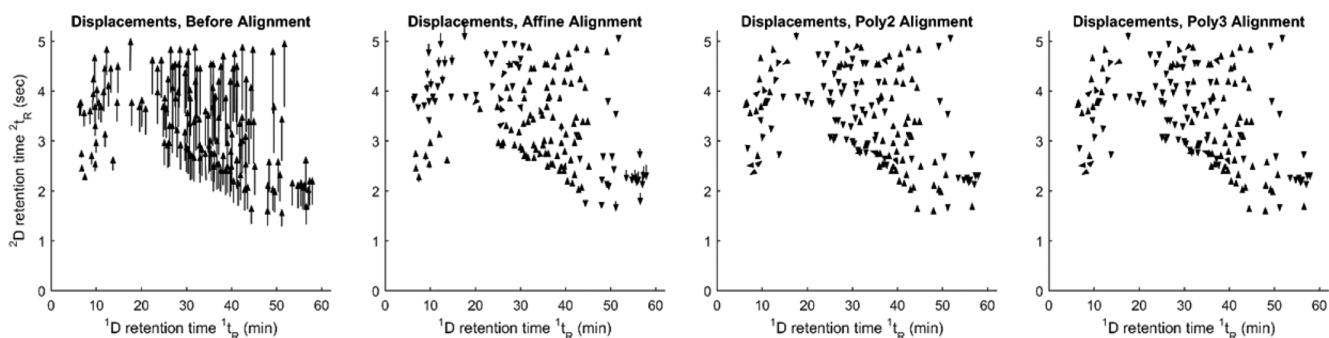
Figure 7. Misalignment vectors (from FID to MS) for 156 peak pairs in chromatograms from GC × 2GC analysis of a urine sample. Columns from left to right are for the four alignment transformations: none (f0), affine (f1), second-degree polynomial (f2), and third-degree polynomial (f3).

of 0.045 s. The performance of the third-degree polynomial is slightly worse than the benchmark; however, the peak-pairs set for the calibration data may not be large enough to provide robust constraints and is somewhat limited in covering the retention-times space, so more in-depth analyses are presented next in the results for urine samples, for which a larger peak-pairs set with wider retention-times ranges is used.

Results for Urine Sample Chromatograms. Figure 4 shows images of the FID and MS chromatograms from a GC × 2GC analysis of a urine sample. Open green circles indicate the analyte peaks used for the alignment experiments. This peak-pairs set is more than 6 times larger than the calibration peak-pairs set, with a wider retention-times distribution.

Figure 5 shows the alignment results for consecutive replicate runs of one of the urine samples, which are used to establish

benchmarks. For ^{1}D, the benchmark is about 0.03 min (or about $0.5 \times \, ^1\sigma$), and for ^{2}D, the benchmark is about 0.03 s (or about $0.35 \times \, ^2\sigma$) for the MS (which has a lower sampling frequency than the FID). These values are slightly less than the benchmarks for the calibration peaks (but still above the level of stochastic noise from sampling), a difference that seems related to the larger, more intense peaks in the calibration sample. As seen in the results for the calibration samples, none of the transformation models produced any improvement in alignment of consecutive replicate runs.

Figure 6 shows the alignment performance for the two chromatograms from a GC × 2GC-FID/MS run of a urine sample. For ^{1}D, the misalignment of the FID and MS chromatograms is at the benchmark (about 0.03 min or $0.5 \times \, ^1\sigma$) without transformation and, as expected, none of the
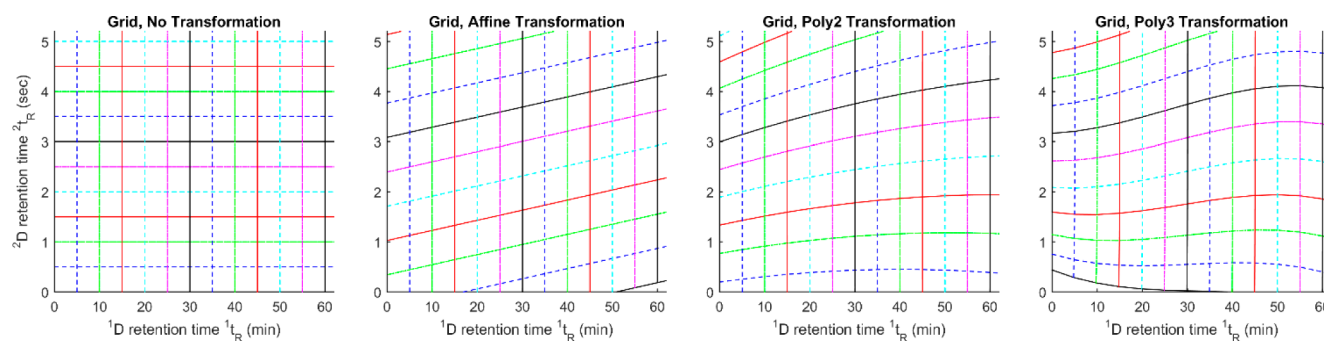
Figure 8. Transformation of grid lines using the models derived for Figure 7 (FID to MS mapping for GC × 2GC analysis of a urine sample). Columns from left to right are for the four alignment transformations: none (f_0), affine (f_1), second-degree polynomial (f_2), and third-degree polynomial (f_3).

models reduces misalignment. For ^{2}D, the misalignment before transformation is large, about 0.25 s or 2.9 × $^2\sigma$, about 8 times the benchmark. All of the transformation models significantly improve alignment.

As seen in the right-most graph in Figure 6, affine transformation reduces misalignment by about two-thirds, to 0.08 s (or 0.9 × $^2\sigma$) but does not achieve the benchmark. Both polynomial transformations reduce misalignment to about 0.03 s (or 0.35 × $^2\sigma$), which is the benchmark for consecutive replicate sample runs. The second-degree polynomial does not approach its peak performance until there are nearly 30 points in the training set, although it outperforms affine transformation with as few as 9 or 10 points. The third-degree polynomial requires nearly 20 points in the training set to outperform affine transformation and requires nearly 50 points to outperform the second-degree polynomial. Even when the third-degree polynomial overtakes the second-degree polynomial, the performance gain is small. Therefore, the second-degree polynomial may be preferred for robustness with a smaller training set and simpler computation, with little loss in performance.

Figure 7 shows the misalignment vectors between peak-pairs in the GC × 2GC chromatograms before and after alignment. The displacements before alignment increase both from bottom to top and from left to right. The affine transformation improves alignment for many peaks but has errors visible in the upper left and lower right. Both the polynomial transformations improve alignment across the retention-times space.

Figure 8 shows the transformation functions applied to a uniform rectangular grid. With the affine transformation, the parallel lines for constant ^{2}D retention times remain parallel after the transformation, so the nonlinear differences cannot be corrected. The polynomial transformations can track the nonlinear systemic differences and so outperform affine transformation.

◼ CONCLUSIONS AND FUTURE WORK

This research indicates that global, low-order polynomial transformations are effective for aligning chromatograms from GC × 2GC with duplicate secondary columns and that they improve on the performance of affine transformation. Second-degree polynomials performed nearly as well as third-degree polynomials and may be preferred for their simplicity, requirement of fewer landmark peak-pairs, and greater robustness for small sets of landmark peak-pairs. In the experiments here, although as few as six landmark peak-pairs are mathematically sufficient for determining the optimal

second-degree polynomial, more than 4 times as many may be required for near peak-performance for points not in the training set. We expect that these conclusions also would hold for GC × GC with dual parallel detectors (e.g., GC × GC-FID/MS or other detector combinations) for which the 2D chromatograms may have differing retention-times patterns (e.g., due to differences in splitter performance, connecting capillaries, operative pressures, and/or detectors). However, for GC × 2GC with different stationary phases in the secondary columns, simple retention-time transformation models could be ill-suited (be they local or global). This research also makes contributions in methodologies for evaluating alignment performance: demonstrating the use of consecutive replicate runs to establish alignment benchmarks related to inherent noise and using cross-validation with variable-sized training sets to assess performance as a function of the number of points used for alignment.

Although it is not the subject of this work, alignment of GC × GC chromatograms from different runs also is an important problem. For example, GC × GC retention-times patterns change as columns age and differ from one system to another. Global low-order polynomial transformations can be applied to the problem of aligning different GC × GC chromatograms,[17] but their performance for different situations has not been investigated deeply. We have begun experiments that use the assessment methods described here to evaluate the performance of global, low-order polynomials for aligning GC × GC chromatograms in varied situations, with comparisons to local alignment methods.

The next version of GC Image Software will have facilities (a) to compute the optimal parameters for a second-degree polynomial model from matched points determined by template matching, (b) to update a template based on a second-degree polynomial function computed after template matching, and (c) to transform a chromatogram according to a set of matched points (e.g., those determined by template matching or any other method) using a second-degree polynomial function and a selected resampling method.

◼ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.anal-chem.5b02718.

Results for additional examples and additional statistical results (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: reich@unl.edu.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Seeley, J. V.; Kramp, F. J.; Sharpe, K. S. *J. Sep. Sci.* **2001**, *24*, 444−450.

(2) Seeley, J. V.; Kramp, F. J.; Sharpe, K. S.; Seeley, S. K. *J. Sep. Sci.* **2002**, *25*, 53−59.

(3) Nicolotti, L.; Cordero, C.; Bressanello, D.; Cagliero, C.; Liberto, E.; Magagna, F.; Rubiolo, P.; Sgorbini, B.; Bicchi, C. *J. Chromatogr. A* **2014**, *1360*, 264−275.

(4) Bressanello, D.; Liberto, E.; Collino, M.; Reichenbach, S. E.; Benetti, E.; Chiazza, F.; Bicchi, C.; Cordero, C. *J. Chromatogr. A* **2014**, *1361*, 265−276.

(5) Peroni, D.; Sampat, A. A. S.; van Egmond, W.; de Koning, S.; Cochran, J.; Lautamo, R.; Janssen, H.-G. *J. Chromatogr. A* **2013**, *1317*, 3−11.

(6) Peroni, D.; Janssen, H.-G. *J. Chromatogr. A* **2014**, *1332*, 57−63.

(7) Fraga, C. G.; Prazen, B. J.; Synovec, R. E. *Anal. Chem.* **2001**, *73*, 5833−5840.

(8) Sakoe, H.; Chiba, S. *IEEE Trans. Acoust., Speech, Signal Process.* **1978**, *26*, 43−49.

(9) Nielsen, N.-P.; Carstensen, J. M.; Smedsgaard, J. *J. Chromatogr. A* **1998**, *805*, 17−35.

(10) Vial, J.; Noçairi, H.; Sassiat, P.; Mallipatu, S.; Cognon, G.; Thiébaut, D.; Teillet, B.; Rutledge, D. N. *J. Chromatogr. A* **2009**, *1216*, 2866−2872.

(11) Zhang, D.; Huang, X.; Regnier, F. E.; Zhang, M. *Anal. Chem.* **2008**, *80*, 2664−2671.

(12) Wang, B.; Fang, A.; Helm, J.; Bogdanov, B.; Pugh, S.; Libardoni, M.; Zhang, X. *Anal. Chem.* **2010**, *82*, 5069−5081.

(13) Gros, J.; Nabi, D.; Dimitriou-Christidis, P.; Rutler, R.; Arey, J. S. *Anal. Chem.* **2012**, *84*, 9033−9040.

(14) Sibson, R. In *Interpreting Multivariate Data*; Barnett, V., Ed.; John Wiley & Sons: New York, 1981; pp 21−36.

(15) Furbo, S.; Hansen, A. B.; Skov, T.; Christensen, J. H. *Anal. Chem.* **2014**, *86*, 7160−7170.

(16) Ni, M.; Reichenbach, S. E.; Visvanathan, A.; TerMaat, J.; Ledford, E. B. *J. Chromatogr. A* **2005**, *1086*, 165−170.

(17) van Mispelaar, V. *Chromametrics*; Universal Press: Veenendaal, The Netherlands, 2005.

(18) Zhang, Q.; Wang, G.; Du, Y.; Zhu, L.; Jiye, A. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2007**, *854*, 20−25.

(19) Reichenbach, S. E. In *Comprehensive Two Dimensional Gas Chromatography*; Ramos, L., Ed.; Elsevier: Amsterdam, The Netherlands, 2009; pp 77−106.

(20) GC Image, LLC. *GC Image GCxGC Software Users' Guide*; GC Image, LLC: Lincoln, NE, 2015.

(21) Wolfram Research, Inc. Least Squares Fitting-Polynomial. http://mathworld.wolfram.com/LeastSquaresFittingPolynomial.html (accessed August 25, 2015).