analytical chemistry

Article

pubs.acs.org/ac

Effectiveness of Global, Low-Degree Polynomial Transformations for GCxGC Data Alignment

Davis W. Rempe,[†] Stephen E. Reichenbach,[*,†,‡] Qingping Tao,[‡] Chiara Cordero,[§] Wayne E. Rathbun,[∥] and Cláudia Alcaraz Zini[#]

[†]University of Nebraska, Lincoln, Lincoln Nebraska 68588-0115, United States

[‡]GC Image, LLC, Lincoln Nebraska 68505-7403, United States

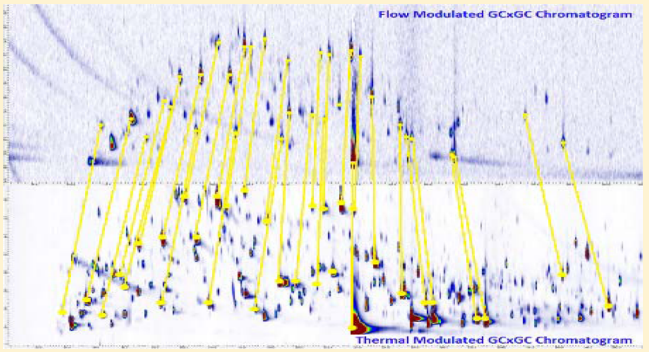
[§]Università degli Studi di Torino, I-10125 Torino, Italy

[∥]Honeywell UOP, 25 East Algonquin Road Des Plaines, Illinois 60017-5016, United States

[#]Universidade Federal do Rio Grande do Sul, 91501-970 Porto Alegre, Rio Grande do Sul, Brazil

Supporting Information

ABSTRACT: As columns age and differ between systems, retention times for comprehensive two-dimensional gas chromatography (GCxGC) may vary between runs. To properly analyze GCxGC chromatograms, it often is desirable to align the retention times of chromatographic features, such as analyte peaks, between chromatograms. Previous work by the authors has shown that global, low-degree polynomial transformation functions, namely affine, second-degree polynomial, and third-degree polynomial, are effective for aligning pairs of two-dimensional chromatograms acquired with dual second columns and detectors (GC×2GC). This work assesses the experimental performance of these global methods on more general GCxGC chromatogram pairs and compares their performance to that of a recent, robust, local alignment algorithm for GCxGC data [Gros et al. *Anal. Chem.* **2012**, *84*, 9033]. Measuring performance with the root-mean-square (RMS) residual differences in retention times for matched peaks suggests that global, low-degree polynomial transformations outperform the local algorithm given a sufficiently large set of alignment points, and are able to improve misalignment by over 95% based on a lower-bound benchmark of inherent variability. However, with small sets of alignment points, the local method demonstrated lower error rates (although with greater computational overhead). For GCxGC chromatogram pairs with only slight initial misalignment, none of the global or local methods performed well. In some cases with initial misalignment near the inherent variability of the system, these methods worsened alignment, suggesting that it may be better not to perform alignment in such cases.



This work assesses the performance of global, low-degree polynomial transformations, namely, affine, second-degree polynomial, and third-degree polynomial,[1] for retention-time alignment between chromatograms obtained by comprehensive two-dimensional gas chromatography (GCxGC). It also compares the performance of these global methods to that of a recent, robust, local alignment algorithm for GCxGC chromatograms proposed by Gros et al.[2]

Because of column aging and other run-to-run system variations, retention times may vary between GCxGC chromatograms, even when acquired on the same system. To mitigate this issue, it may be necessary to perform chromatographic alignment by mapping the retention times of one chromatogram to the times of another chromatogram. Alignment methods can be classified as "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".[1]

Previous work[1] by the authors investigated global, low-degree polynomial transformation functions for aligning chromatogram pairs acquired by comprehensive two-dimensional gas chromatography with one first-dimension (^{1}D) column and two parallel second-dimension (^{2}D) columns (GC×2GC).[3−6] The chromatogram pairs aligned in that work came from the same run with one ^{2}D column to a flame ionization detector (FID) and another ^{2}D column to a mass spectrometer (MS). These chromatogram pairs had significant variations in the ^{2}D retention times. For GC×2GC, low-degree polynomial mapping functions outperformed affine transformations. These polynomial functions were able to approach benchmarks

Received: June 10, 2016
Accepted: September 18, 2016
Published: September 18, 2016

DOI: 10.1021/acs.analchem.6b02254
Anal. Chem. 2016, 88, 10028−10035

for retention-time root-mean-square residual error (RMSE) between chromatograms based on consecutive replicate sample runs on the same system and detector.

The present work investigates the performance of these global, low-degree polynomial transformations more generally to align GCxGC chromatograms, that is: do the retention-times RMSE between two chromatograms after alignment approach the noise benchmark? To assess these global methods, chromatograms from three sets of data are used for alignment. Each data set varies an important chromatographic parameter: the date and time, the sample, and the instrument configuration. This allows the alignment methods to be tested across a wide range of situations. The first set of chromatograms was produced from the same diesel sample run over a period of about two and a half years. These chromatograms have moderate initial misalignment due to system and column variations. The second set of chromatograms was produced from samples of three different wine vintages that were run in a period of days. These chromatograms have minimal initial misalignment from run-to-run system variability. The last set of chromatograms was produced from a single cocoa sample, but on systems with two different modulation technologies: flow and thermal. These chromatograms are extremely misaligned due to different system configurations, namely, different modulators, column dimensions, and carrier gas flow.

Additionally, this research compares the performance of these global functions to that of a high-performing local alignment algorithm.[2] Comparing global and local methods may show whether retention-time differences between GCxGC chromatograms are systemic and therefore well-suited to simple, global functions, or if the differences are too complex and require more sophisticated local methods for alignment. For this, a local alignment method developed by Gros et al.[2] is evaluated. Although there are other available local alignment methods, their work indicates that compared to two other local alignment methods, their robust algorithm "performs the best overall in terms of decreased retention time deviations of matching analytes".[2] Gros et al. compared their method to one developed by Pierce et al.,[7] which was "the first published alignment algorithm for the correction of shifts resulting from uncontrollable variations for whole GCxGC chromatograms",[2] and to two-dimensional correlation optimized warping (2-D COW),[8] a multidimensional extension of the original COW algorithm.[9] As evidenced by these results, the method described by Gros is a high-performing local method.

The experimental methods for testing the various alignment functions follow previous work.[1] The effectiveness of the alignment methods is measured in terms of the RMSE of the postalignment retention times for pairs of matched peaks in two GCxGC chromatograms. The error that the alignment methods aim to reach is the benchmark RMSE, computed between pairs of chromatograms from consecutive replicate sample runs on the same system. This benchmark is based on the assumption that the retention-times differences between consecutive replicate sample runs on the same system are unpredictable random noise. Cross-validation experiments are used to evaluate all methods: affine, second-degree polynomial, third-degree polynomial, and the local algorithm from Gros et al. To get an unbiased indicator of performance, these tests use one set of matched peak-pairs to fit (or train) the alignment functions, and a different, disjoint set to measure (or test) the postalignment RMSE.

EXPERIMENTAL SECTION

Samples. Three different sample types are used to assess performance of the data alignment algorithms. The first is a single distillate diesel sample. The sample was run four different times on the same system over a period of about two and a half years to produce a set of GCxGC chromatograms. Each of these runs were far apart in time, so the chromatograms have moderate misalignments from column differences, such as aging and replacement. The lower-bound benchmark RMSE was determined from a set of four consecutive replicate runs with the same diesel sample on the same system.

The second set of chromatograms came from samples of three different wine vintages. All samples were run within a period of 3 days as part of a study at the Universidade Federal do Rio Grande do Sul related to the characterization of commercial Merlot wines from the Brazilian Campanha region. All samples were from the same Merlot brand, but from different years: 2011, 2012, and 2013. Each sample was run on the system twice consecutively, which provides the replicate runs for determining the alignment benchmark. Because all runs were within a short time period on the same system, the misalignments are relatively small.

The third set of chromatograms came from a single Trinitario cocoa nib sample from Ecuador. The sample was run as part of a study at the Università degli Studi di Torino in Turin, Italy, that focuses on the sensomic characterization of cocoa samples from different botanical and geographical origins. Two chromatograms were first acquired on the system using a reverse-inject differential flow modulator.[10] The same sample was again run about four months later to acquire three more chromatograms, but this time with a loop-type thermal modulator. The flow-modulated GCxGC runs were preliminary experiments under unoptimized conditions, making alignment even more difficult. The sample was run consecutively on each modulation platform, so there are replicate runs to determine the alignment benchmarks. Varying the modulation technologies between these sets of runs results in chromatograms with extreme misalignment, much larger than that seen in the diesel chromatograms, particularly in the ^{1}D.

Instrumentation. For analysis of the diesel sample, all run conditions were in accordance with UOP 990,[11] with a modulation period of 8 s and sampling with a flame ionization detector (FID) at 200 Hz, on a LECO GCxGC-FID system (LECO Corp., St. Joseph, MI) with Agilent 6890 GC (Agilent Technologies, Little Falls, DE).

For analysis of the volatile fraction of wine samples, headspace solid-phase microextraction (HS-SPME) was performed with one mL of wine, 0.3 g of sodium chloride at 55 °C (±0.9), and a DVB/CAR/PDMS fiber (Supelco, Bellefonte, PA) in 20 mL headspace screw-capped glass vials.[12] The system was a LECO GCxGC with an Agilent 6890N and time-of-flight mass spectrometric detector (TOFMS). The modulation cycle was 7 s with spectra from 45 to 450 m/z acquired at about 100 Hz.

For analysis of the cocoa nib, the GCxGC experimental conditions were different for each modulation technology. The GC×2GC-MS/FID runs with reverse-inject differential flow modulation used an Agilent 7890B GC unit coupled to an Agilent 5977A fast quadrupole MS detector operating in EI mode at 70 eV, and a fast FID. The modulation cycle was 3 s with spectra from 40 to 240 m/z acquired at about 35 Hz. The GCxGC-MS runs with thermal modulation used an Agilent

DOI: 10.1021/acs.analchem.6b02254
Anal. Chem. 2016, 88, 10028−10035

6890 unit with a Zoex loop-type modulator (Zoex Corp., Lincoln, NE) coupled to an Agilent 5975C MS detector operating in EI mode at 70 eV. The modulation cycle was 3 s with spectra from 40 to 240 m/z acquired at about 29 Hz.

Additional details of the instrumental conditions for all systems are included in the Supporting Information.

Data Preprocessing. Data preprocessing was performed using GC Image GCxGC Edition Software (R2.6 alpha build) from GC Image, LLC (Lincoln, NE).[13]

For the diesel chromatograms, phase-shifting, baseline correction, and peak detection were performed.[14] Automated bidirectional peak matching created initial lists of corresponding peaks between all pairs of chromatograms. The lists were edited manually to increase the number and temporal coverage and to ensure correct correspondences, resulting in a total of 112 peaks that were matched across all eight chromatograms (four runs well separated in time and four consecutive replicate runs). Because manual verification was being performed, loose matching criteria were used for creating the initial list to increase the number of prospective peak-pairs and minimize bias. After manual editing, the peaks are well-distributed across the retention times of the chromatograms (Figure S11).

For the wine chromatograms, baseline correction and peak detection were performed. Using the same process as for the diesel sample, a total of 78 peaks were selected and confirmed by MS to correspond across all six chromatograms (Figure S12).

Chromatograms acquired from the cocoa sample yielded fewer corresponding peak-pairs using these peak matching techniques. After baseline correction and peak detection, 33 peaks were confirmed by MS across all five chromatograms (Figures S13 and S14).

Evaluation Metric. The primary evaluation metric is the RMSE of the postalignment retention times across the peak sets for pairs of chromatograms. This metric is described in detail in previous work.[1] A blob's retention times indicate its data point with the maximal signal value, that is, its apex.

Transformation Models. The first evaluation is with no alignment function applied, that is, the initial misalignment. The transformation functions of the affine, second-degree polynomial, and third-degree polynomial global alignment methods are given in previous work.[1] Each global function requires a minimum number of alignment peak-pairs to determine the parameters: three peak-pairs for affine, six peak-pairs for second-degree polynomial, and ten peak-pairs for third-degree polynomial. For numbers of peak-pairs larger than the minimum number, the optimal parameters minimize the RMSE of fitted pairs.

The local alignment method of Gros et al.[2] also uses corresponding peak-pairs for alignment. These peak-pairs are referred to as alignment points. This algorithm guarantees that these points are perfectly aligned in the final chromatogram produced. Based on these alignment points, displacements for the rest of the data are estimated in both dimensions. In the ^{1}D, displacements are linearly interpolated between alignment points. In the ^{2}D, displacements are estimated using Sibson natural-neighbor interpolation,[15] based on Voronoi diagrams. For interpolation in the ^{2}D, the algorithm requires the typical peak width (tpw) for both dimensions. This is the number of data-points that make up approximately two standard deviations of a peak.[16] In the diesel experiments, tpws of 2 and 40 data-points (0.267 min and 0.2 s) were used for the ^{1}D and ^{2}D, respectively. For the wine samples, tpws of 2 and 17 (0.23 min and 0.17 s) data-points were used. For the cocoa samples, tpws of 5 (0.25 min) and 6 data-points (0.17 s for flow modulation, 0.21 s for thermal) were used. These tpws were roughly determined by visual examination of typical peaks near the center of the chromatogram. This process follows the documentation to users from Gros et al.[16] The final step of the algorithm reinterpolates the signal values for all pixels and applies a deformation correction. This part of the algorithm was not executed during the cross-validation testing in this paper, because the focus here is on comparing the retention-times alignments with those of the global methods and not on the separate step of intensity interpolation.

Evaluation Methodology. The evaluation methodology follows previous work.[1] Within the alignment points used, the transformations fit the noise as well as the alignment peaks, which is a problem of overfitting. To get an unbiased estimate of a method's performance, a cross-validation technique is employed. The set of corresponding peak-pairs is partitioned into two disjoint sets: a training and testing set. The training set is used as the alignment points for fitting the methods, and the testing set is used to measure their performance. Measuring the error across testing-set peak-pairs after alignment is a good unbiased indicator of the method's performance, as the transformation was not fit to these peak-pairs and their inherent noise.

The experiments are run for every training set size from 3 peak-pairs (the minimum size for affine transformations) to all of the matched peak-pairs, at which point the test set is null. For each training set size, 100 trials are run. The training and testing sets are randomly generated at each trial (and are disjoint complements of the peak-pairs set). Because of the random selection of peak-pairs, the training set may not be well-distributed across the entire chromatogram. The alignment is also done both forward and backward, i.e., peaks from chromatogram 1 are fit to those in chromatogram 2 and vice versa. The reported RMSE for each training set size is the average RMSE over all 200 trials (with 100 in each direction).

Performance Benchmarks. The global alignment methods are assessed in two ways. First, does the method approach the benchmark error set by the consecutive replicate runs? Second, does the method perform better than the local alignment algorithm? For the first question, the misalignment between consecutive replicate runs can be used as a benchmark indicating the lower bound of alignment performance due to systemic noise. Any misalignment between two replicate chromatograms acquired one after another with the same sample on the same system can be considered the level of random retention-times noise inherent to the system itself.

The degree to which an alignment method approaches the benchmark is measured by its percent improvement I_p. For a specific alignment method, let S be the set of postalignment average RMSEs for every testing set size and $\min\{s\}$, $s \in S$, be the best average RMSE achieved for any testing set size. Then that method's percent improvement is defined as

$$I_p = \frac{m_0 - \min_{s\in S}\{s\}}{m_0 - m_b} \times 100 \quad (1)$$

where m_0 is the average testing set RMSE over all trials with no alignment function applied (i.e., the initial misalignment) and m_b is the benchmark RMSE from consecutive replicate runs.

Comparing global performance to the local method is done in multiple ways. If the alignment methods have a RMSE less

DOI: 10.1021/acs.analchem.6b02254
Anal. Chem. 2016, 88, 10028−10035

than that of the local method, they can be said to perform better. The computational overhead (i.e., run-time) of an alignment algorithm is another useful comparison. It is also important to take into consideration how many peak-pairs are required in order to achieve (or nearly achieve) the method's maximal performance. It may be desired to have a method that can align two chromatograms relatively well using fewer alignment points, rather than one that can achieve a slightly smaller RMSE but which requires more alignment points.

Ideally, the methods should be compared on their performance for specific data sets of interest. For generality, the data sets used here offer a wide range of initial misalignment, from negligible to severely misaligned, so the alignment performance can be considered relative to the initial misalignment. Additionally, each data set varies a different GCxGC chromatogram acquisition parameter. The first varies the analysis over time, the second varies the sample, and the third varies the GCxGC instrument with different modulation platforms.

Execution Methodology. Experiments were run on the Crane cluster of the Holland Computing Center[17] located on the University of Nebraska-Lincoln campus. The cluster has a total of 452 nodes with 64 GB of RAM each. In each of the 16 cores within a single node, there are two Intel Xeon E5-2670 2.60 GHz processors.

All alignment methods were implemented in MATLAB. Part of the MATLAB implementation of the local algorithm from Gros et al. was parallelized in order to run much faster across 16 cores on the Crane cluster. Even with the speed boost, and without executing the resampling portion of the algorithm, the local method was more computationally expensive than the simpler global functions.

For the case of 105 peak-pairs for aligning two 1199 × 1600 diesel chromatograms, fitting the second-degree polynomial to the peak-pairs and computing the transformation for every data point required 0.1906 s. By comparison, the local algorithm required 8.5971 s to compute the displacements for every data point. Of course, the computation-time difference is smaller if fewer retention times must be transformed (e.g., as would be required to transform a template). However, as these timing results illustrate, the global function requires significantly less computation for larger alignment problems.

RESULTS AND DISCUSSION

Time-Varied Data Results. Chromatograms acquired from the diesel sample were used to test the alignment methods on time-varied data. Tests were performed on chromatograms from four consecutive replicate runs on the same diesel sample to establish a benchmark for the alignment methods. These four chromatograms are labeled runs 17, 18, 19, and 20. The initial misalignment was recorded for consecutive runs: 17 and 18, 18 and 19, and 19 and 20. The results from the cross-validation benchmark tests between runs 18 and 19 are shown in Figure 1. These graphs show the retention-time RMSE for the testing set of peak-pairs for each alignment method as a function of the training-set size, that is, the number of alignment points used. Each alignment method is represented by a different colored line. The figures for the training sets and additional replicate results can be found in the Supporting Information (Figure S1).

In both chromatographic dimensions, as the training set size increases, the RMSE of the global functions generally decreases for the testing sets. This makes sense because larger training sets yield better estimates of the global misalignment (because overfitting to noise is reduced), producing the decrease seen in testing-set error.

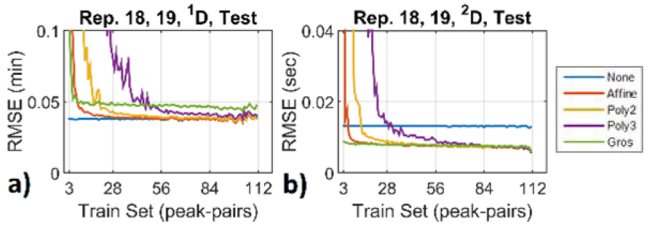


Figure 1. Cross-validation retention-time RMSE results as a function of training set size for consecutive replicate runs of a diesel sample. The RMSE is shown for (a) ^{1}D with the testing set and (b) ^{2}D with the testing set.

The RMSE of consecutive replicate runs, which provides our benchmark error, is the blue line in Figure 1. The ^{1}D graph (Figure 1a) shows that none of the alignment algorithms are able to improve upon this initial misalignment. In the ^{2}D (Figure 1b), there is only a small improvement of less than 0.01 s. This supports the claim that the initial misalignment of consecutive replicate runs indicates the inherent lower-bound limits on any alignment algorithm.

In the ^{1}D, with no alignment function applied, the RMSE averages 0.0375 min, which is the maximum initial misalignment seen in the ^{1}D across the three replicate tests. In the ^{2}D, the initial misalignment is about 0.0131 s. Across all three pairs of replicate runs, the average misalignment in the ^{1}D is 0.0243 min which is less than the modulator sampling noise level of 0.038 min. [The distillate analyses have a modulation cycle (P_M) of 8 s or P_M = 0.13 min. The standard deviation for random uniformly distributed residuals with respect to a single modulation interval is $12^{-1/2} \times P_M$, which is about 0.038 min for these data. This is the RMS retention-time noise level from the sampling effect of modulation and has implications for the benchmark RMSE in the ^{1}D.] The peaks in the diesel sample chromatograms are narrow in the ^{1}D, with a tpw of only about 2 modulations, which affects the choice of an alignment benchmark. An alignment method cannot be expected to achieve an RMSE better than the sampling noise, so 0.038 min is the benchmark value in the ^{1}D. Across all three pairs of replicate runs, the average misalignment in the ^{2}D is 0.0125 s. This value is the ^{2}D benchmark RMSE for the alignment methods being tested.

Next, the cross-validation tests were run on every pairwise combination of four chromatograms acquired over 2.5 years. Because of column aging, these chromatograms exhibit moderate misalignments. The results from one of these pairwise tests are shown in Figure 2. The names of the samples (January 20, 2011, and June 14, 2013) indicate the dates on which they were run; so, the chromatograms aligned in this figure were acquired about 2.5 years apart. Before any alignment is applied (the blue line in Figure 2), the RMSE is about 0.76 min in the ^{1}D and 0.24 s in the ^{2}D. The "None" function is excluded from plot (Figure 2a) to focus on performance of the alignment models.

The testing-set plots in Figure 2 show how the transformations affect peak-pairs that were not used for fitting, for an unbiased evaluation. In both dimensions, significant improvements are seen after applying both the global and local methods to the alignment of chromatograms 012011 and 061413. In the ^{1}D, the third-degree polynomial transformation achieves the smallest RMSE of 0.0641 min compared to the largest RMSE of

DOI: 10.1021/acs.analchem.6b02254
Anal. Chem. 2016, 88, 10028−10035

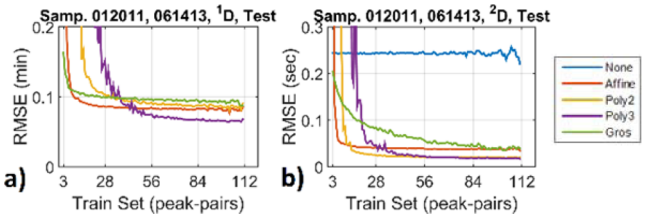


Figure 2. Cross-validation retention-time RMSE results as a function of training set size for chromatograms produced from the same diesel sample about 2.5 years apart. RMSE is shown for (a) ^{1}D with the testing set and (b) ^{2}D with the testing set. The names of the samples correspond to the acquisition date (January 20, 2011, and June 14, 2013).

0.0871 min for the local algorithm. The (best-performing) third-degree polynomial (0.0641 min) has a percent improvement of I_p = 96.4% using the benchmark of 0.038 min. Though it has the largest RMSE of the methods tested, resulting in a percent improvement of I_p = 93.2%, Gros' algorithm only requires about 10 peak-pairs to approach its peak performance. This is a smaller training-set size than required for the global methods to reach peak performance.

In the ^{2}D, the third-degree polynomial also achieves the best peak performance, with a minimum RMSE of 0.017 s (I_p = 98%), nearing the 0.0125 s benchmark, compared to 0.0346 s (I_p = 90.3%) for the local method. In the ^{2}D, Gros' algorithm takes much longer to reach its peak performance at around 85 peak-pairs, but it has a lower RMSE than the global functions when the training set size is small.

Training-set data and graphs similar to Figure 2 for all other cross-validation experiments can be found in the Supporting Information (Figure S4). The patterns discussed with Figure 2 are consistent across most of the experiments. Table 1 summarizes the results from all six cross-validation experiments run with the nonreplicate diesel chromatograms. Under "None" is the average initial misalignment (m_0 in eq 1). For each experiment, the minimum average testing set RMSE (min{s},s ∈ S, in eq 1) for each alignment method is shown. The bottom two rows present the averages for the minimum RMSE and percent improvement. Note that the top-performing method in terms of average minimum RMSE may not be the best in terms of average percent improvement (and vice versa), because the average percent improvement depends heavily on initial misalignment. Even if a method averages the smallest RMSE, it may not have the smallest RMSE in cases that the misalignment is very small, which negatively affects its average percent improvement.

On average, all three global alignment methods are able to reach a better peak performance and percent improvement than the local algorithm in both dimensions. The third-degree polynomial averages a 9.3% greater percent improvement than Gros' algorithm in the ^{1}D, and 7.6% greater in the ^{2}D. In the ^{1}D, the average percent improvement for all alignment methods is noticeably worse than the experiment discussed in Figure 2. For chromatogram pairs with a less significant initial misalignment in the ^{1}D (012011−090912, 012011−100412, and 090912−100412 in Table 1), the alignment methods tend to reach similar minimum RMSE values to the experiments with larger initial misalignments, causing the lower average percent improvements overall. For experiments with large misalignments (>0.7 min), like in Figure 2, the third-degree polynomial is consistently able to achieve a percent improvement over 95%. In the ^{2}D, both the second and third-degree polynomials average a percent improvement over 93% for all experiments.

There is a clear trade-off in terms of the number of alignment points used and the minimum RMSE reached for both the local and global methods. If using a very small number of alignment points (∼5), it may be preferable to use Gros' algorithm because it starts out at a much lower error than any of the global methods. Though the local method performs relatively well with a small number of alignment points, it is outperformed in both dimensions by the global methods when a larger numbers of alignment points are available. The number of peak-pairs at which the global methods overtake the local method varies between algorithms, and is larger in the ^{1}D than the ^{2}D. With just under 10 points or more, the affine transformation becomes a better choice, attaining a clear performance gain in both dimensions, on average. With around 30 pairs or more, the second-degree polynomial performance overtakes the local method. The third-degree polynomial improves upon the local method when about 50 alignment points or more are available. Though the third-degree polynomial is also able to outperform the second-degree (with ∼55 points), the performance gain is small. In terms of percent improvement, the second-degree actually averages better than the third-degree in the ^{1}D, and is within 1% in the ^{2}D. Therefore, for computational simplicity and because fewer alignment points are required, it may be preferable to use the second-degree function. This result is similar to that seen in previous work for GC×2GC.[1]

Table 1. Minimum Test-Set RMSE for Each Alignment Method in Both the First and Second Chromatographic Dimensions for All Six Experiments with the Nonreplicate Chromatograms from the Diesel Sample

	None (av.)		Affine		Poly2		Poly3		Gros et al.	
Chromatograms	^{1}D (min)	^{2}D (s)	^{1}D (min)	^{2}D (s)	^{1}D (min)	^{2}D (s)	^{1}D (min)	^{2}D (s)	^{1}D (min)	^{2}D (s)
012011−061413	0.7563	0.2414	0.0767	0.0344	0.0806	0.0184	0.0641	0.0170	0.0871	0.0346
012011−090912	0.1024	0.3982	0.0574	0.0147	0.0583	0.0130	0.0592	0.0131	0.0640	0.0435
012011−100412	0.0800	0.0569	0.0502	0.0257	0.0460	0.0225	0.0488	0.0223	0.0612	0.0283
061413−090912	0.8353	0.1819	0.0856	0.0367	0.0868	0.0223	0.0747	0.0221	0.0902	0.0209
061413−100412	0.7940	0.2905	0.0763	0.0558	0.0783	0.0331	0.0511	0.0282	0.0996	0.0558
090912−100412	0.0770	0.4386	0.0631	0.0292	0.0635	0.0247	0.0644	0.0241	0.0671	0.0578
Average	0.4408	0.2679	0.0682	0.0328	0.0689	0.0223	0.0604	0.0211	0.0782	0.0402
Average percent improvement (%)			76.7	87.7	77.8	93.1	77.3	93.6	68.0	86.0

[a]The "None" columns are the average initial misalignments, not the minimum. The third-degree polynomial function reaches the lowest error on average and Gros et al. has the highest error on average.

DOI: 10.1021/acs.analchem.6b02254
Anal. Chem. 2016, 88, 10028−10035

Sample-Varied Data Results. Chromatograms acquired from the three wine samples were used to test the alignment methods on sample-varied data. The benchmark RMSE for wine sample chromatographic alignment is established with pairs of consecutive replicate runs of the 2011, 2012, and 2013 vintages. The results for the 2011 sample replicate runs are shown in Figure 3. Training-set data and additional replicate

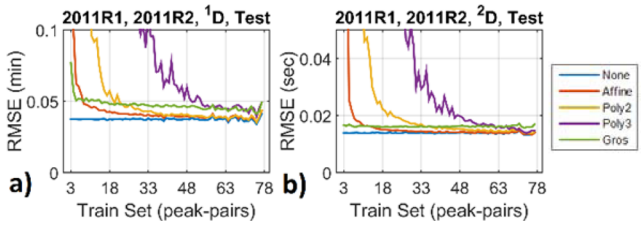


Figure 3. Cross-validation retention-time RMSE results as a function of training set size for consecutive replicate runs of the 2011 wine sample. The names correspond to the vintage year of the wine sample.

runs can be found in the Supporting Information (Figure S2). The titles of the graphs indicate which two chromatograms were aligned with the year of the sample followed by an "R" and the run number (1 or 2). Figure 3 shows aligned chromatograms for runs 1 and 2 from the 2011 sample. As seen in the testing-set plots, none of the alignment methods are able to improve on the initial misalignment. This indicates there is no systematic retention-time difference between the replicate runs, only retention-time noise. In the ^{1}D (Figure 3a), the RMSE with no alignment is about 0.0368 min, which is the maximum for any of the replicate sample runs. In the ^{2}D (Figure 3b), the initial misalignment is about 0.0137 s. Over the three sets of replicate runs from 2011, 2012, and 2013, the average misalignment in the ^{1}D is 0.03037 min, and in the ^{2}D is 0.01725 s. The average RMSE in the ^{1}D is less than the modulation sampling noise level of 0.034 min. [The wine analyses have a modulation cycle of 7 s or P_M = 0.117 min, so the RMS retention-time noise level from the sampling effect of modulation is 0.034 min.] The peaks detected in the wine chromatograms are very narrow, with a tpw of about 2 modulations, so the sampling noise must be considered. Therefore, 0.034 min is used as the ^{1}D benchmark RMSE for the alignment of the wine sample chromatograms. The average misalignment in the ^{2}D of 0.01725 s is the other benchmark for the alignment methods.

The chromatograms produced from the second run of each year's sample were tested in every pairwise combination with the other two years. All these samples were run within a span of 3 days, so the initial misalignment between them is small, due mainly to run-to-run random variations and sample differences for the different vintages. The results from aligning chromatograms from the 2011 and 2012 samples are shown in Figure 4. The RMSE between the chromatograms without any alignment functions applied is only about 0.0344 min in the ^{1}D (Figure 4a) and 0.02 s in the ^{2}D (Figure 4b). Both these values are just above the benchmark inherent noise threshold in each dimension, suggesting that the alignment methods should not be expected to improve much upon the initial misalignment. This is apparent in both the ^{1}D and ^{2}D testing-set plots which shows that none of the methods are able to improve the alignment more than a few thousandths of a minute and second, respectively. The minimum RMSE reached by Gros' algorithm in the ^{1}D is slightly worse than the initial misalignment.

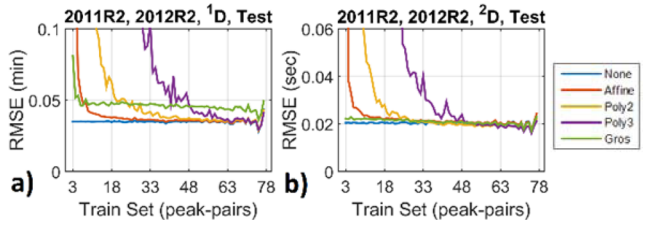


Figure 4. Cross-validation retention-time RMSE results as a function of training set size for alignment of two different wine sample chromatograms. The names correspond to the vintage year of the wine sample.

A table of results and graphs for all other cross-validation experiments can be found in the Supporting Information (Table S1 and Figures S5 and S6). On average, the initial misalignment in both chromatographic dimensions is close to the benchmark values and, as a result, none of the alignment methods achieve notable improvements. The third-degree polynomial even averages a slightly greater minimum value than the initial misalignment in the ^{1}D. These data then suggest that no method, global nor local, is able to perform well. If two chromatograms have only a small initial misalignment, it may be better not to perform any alignment operation at all.

Instrument-Varied Data Results. Chromatograms acquired from the single cocoa sample were used to test the alignment methods on data obtained with differing instruments. The benchmark RMSE values for the cocoa chromatogram alignments are established using two replicate sample runs with the flow modulator and three replicate runs with the thermal modulator. The results of the second replicate cross-validation experiment with the thermal-modulator chromatograms are shown in Figure 5. Training-set data and additional replicate

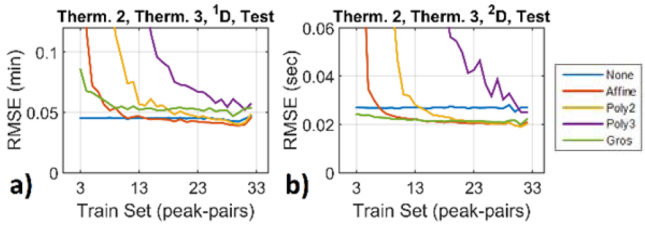


Figure 5. Cross-validation retention-time RMSE results as a function of training set size for consecutive replicate runs of a cocoa sample using a thermal modulator.

runs can be found in the Supporting Information (Figure S3). As expected, only negligible improvements in alignment are seen from any method in either chromatographic dimension. In the ^{1}D (Figure 5a), the average initial misalignment for this experiment is about 0.0438 min which is the maximum seen in any of the replicate experiments. In the ^{2}D (Figure 5b), the initial misalignment is about 0.026 s. Across all three replicate sample run experiments, the average misalignment in the ^{1}D is 0.0412 min, which is used as the benchmark. The modulation sampling noise level for these chromatograms does not greatly affect the benchmark because the peaks detected, with a tpw of about 5 modulations, are wider than those seen from the diesel and wine samples. [The cocoa analyses have a modulation cycle of 3 s or P_M = 0.05 min, so the RMS noise level from the

DOI: 10.1021/acs.analchem.6b02254
Anal. Chem. 2016, 88, 10028−10035

sampling effect of modulation is 0.0144 min.] The average ^{2}D misalignment is 0.0257 s, which is used as the benchmark.

Pairs of chromatograms, one from the two flow-modulator runs and one from the three thermal-modulator runs, were tested in every combination, totaling six experiments. The chromatograms in each experiment were acquired with two different modulators, so the initial misalignment is severe, especially in the ^{1}D because of the constraints posed by the differential flow modulation dynamics to carrier gas volumetric flow. The results from aligning the second flow-modulator chromatogram to the first thermal-modulator chromatogram are shown in Figure 6. The initial misalignment in the ^{1}D (the blue line) is excluded from plot (Figure 6a) because it is so large. In the ^{2}D, the initial misalignment hovers around 0.5 s, and is also excluded from plot (Figure 6b).

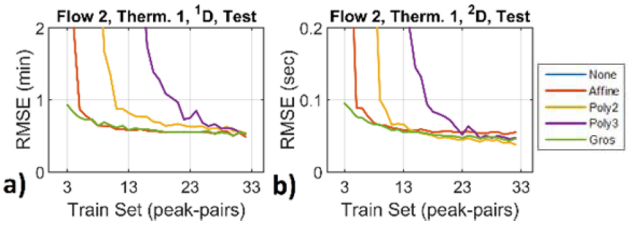


Figure 6. Cross-validation retention-time RMSE results as a function of training set size for chromatograms produced from the same cocoa sample but using two different modulation platforms.

In Figure 6, every method offers significant improvement in both dimensions. In the ^{1}D, the affine transformation function reaches the lowest error of 0.488 min (percent improvement I_p = 98.1%), just in front of Gros' algorithm at 0.503 min (I_p = 98.0%). The second and third-degree polynomial transformations are about the same at 0.537 and 0.527 min (I_p = 97.9%), respectively. The percent improvement from every method is good, even though the benchmark of 0.0412 min is not achieved (perhaps because the initial misalignment is so large).

That the affine transformation performs best suggests that the higher-degree polynomials were not fit with enough alignment points to reach peak performance. Similar to the diesel results, with fewer than 10 alignment points the local method outperforms the global methods in terms of RMSE. Around 10 peak-pairs, though, the affine transformation surpasses Gros' algorithm for a slight performance gain. Because the total number of corresponding peaks across the cocoa chromatograms is only 33, significantly fewer than for the diesel or wine chromatograms, the second and third-degree polynomials do not reach a lower RMSE than the affine transformation or local method. With training sets around 30 peak-pairs, though, they do approach these performances. Figure 6 is a good example of the potential advantages to using Gros' algorithm or the affine transformation when few alignment points are available.

In the ^{2}D, the second-degree polynomial reaches the lowest RMSE of 0.038 s (I_p = 97.4%), followed by Gros' algorithm at 0.043 s (I_p = 96.3%), the third-degree polynomial at 0.046 s (I_p = 95.7%), and the affine transformation at 0.052 s (I_p = 94.3%). Again, every alignment method attains a high percent improvement. The peak RMSE from the second-degree polynomial (0.038 s) also approaches the benchmark set at 0.0257 s. In the ^{2}D, the second-degree polynomial converges to its peak performance with fewer peak-pairs than in the ^{1}D, allowing it to surpass performance of the affine transformation and Gros' algorithm with about 15 alignment points. In terms of percent improvement, this performance gain is small. The third-degree polynomial does not have enough alignment points to be well fit, causing a slightly worse performance than both the second-degree polynomial and local method.

Table 2 shows a summary of the results from all six cross-validation experiments. Graphs from the other experiments are in the Supporting Information (Figure S7). The average case performance of the global and local methods closely mirrors the performances discussed with Figure 6. Although a global function was able to, on average, outperform the local method (affine in ^{1}D and second-degree polynomial in ^{2}D), the performance gain is minimal in terms of percent improvement. All methods perform well, averaging a percent improvement over 95%. In line with conclusions from the diesel alignment results, it may be preferable to use Gros' algorithm if very few alignment points are available, affine transformation when more than a few alignment points are available, and polynomial transformation when 30 or more alignment points are available.

CONCLUSIONS

This work indicates that low-degree polynomial transformation functions will, on average, outperform the local alignment method developed by Gros et al., if given a sufficient number of alignment points for a good fit. Looking at cross-validation tests run on diesel chromatograms, which were acquired at varying times, the global methods consistently achieve a lower peak RMSE than the local method. The cross-validation experiments

Table 2. Minimum Test-Set RMSE Reached by Each Alignment Method in Both the First and Second Chromatographic Dimensions for All Six Experiments Run with the Chromatograms from the Cocoa Sample

	None (av.)		Affine		Poly2		Poly3		Gros et al.	
Chromatograms	^{1}D (min)	^{2}D (s)	^{1}D (min)	^{2}D (s)	^{1}D (min)	^{2}D (s)	^{1}D (min)	^{2}D (s)	^{1}D (min)	^{2}D (s)
Flow 1-Thermal 1	23.2438	0.5204	0.4897	0.0491	0.5417	0.0332	0.5385	0.0358	0.5159	0.0408
Flow 1-Thermal 2	23.2396	0.5094	0.4822	0.0383	0.5265	0.0265	0.5159	0.0324	0.5268	0.0321
Flow 1-Thermal 3	23.2261	0.5271	0.4783	0.0455	0.5352	0.0273	0.5378	0.0362	0.5195	0.0378
Flow 2-Thermal 1	23.2566	0.4952	0.4879	0.0524	0.5367	0.0379	0.5274	0.0457	0.5025	0.0431
Flow 2-Thermal 2	23.2523	0.4842	0.4801	0.0420	0.5226	0.0316	0.5102	0.0427	0.5139	0.0355
Flow 2-Thermal 3	23.2389	0.5018	0.4757	0.0481	0.5304	0.0308	0.5276	0.0450	0.5057	0.0404
Average	23.2429	0.5064	0.4823	0.0459	0.5322	0.0312	0.5262	0.0396	0.5141	0.0383
Average percent improvement (%)			98.1	95.8	97.9	98.8	97.9	97.1	98.0	97.4

[a]The "None" columns are the average initial misalignments, not the minimum. All methods perform well as indicated by the high percent improvements.

DOI: 10.1021/acs.analchem.6b02254
Anal. Chem. 2016, 88, 10028−10035

run with the cocoa sample chromatograms, acquired with differing instrument configurations, support this conclusion, although the local algorithm still averaged a percent improvement of over 97% in both dimensions. In general, although the third-degree polynomial transformation consistently reaches the lowest minimum RMSE with sufficient fitting (requiring about 55 alignment points), the performance gain over the second-degree polynomial is not significant and may not be worth the extra computational cost.

The tests run on GCxGC chromatograms acquired from varying wine samples indicate that no alignment method, global or local, is able to significantly improve alignment when initial misalignment is close to the retention-times noise level. The third-degree polynomial and local method actually made the alignment slightly worse in several cases, suggesting that when misalignment is very small, it may be better not to apply any alignment operation.

This research suggests that for the purpose of chromatographic alignment between two GCxGC chromatograms, it may be preferable to use global, low-degree transformation functions such as second-degree polynomials rather than local methods when a sufficient number of alignment points are available. These global transformations show a better average performance and incur less computational overhead. However, if working with fewer than 10 alignment points, it may be better to use Gros' algorithm. In order to outperform Gros' algorithm, the affine transformation needed as many as 10 alignment points and the second-degree polynomial needed around 30 points.

The training set size at which the alignment methods reach their peak performance may be affected by how the alignment points are chosen. In the experiments presented here, these peak-pairs were chosen randomly from a large, well-distributed set, but choosing a subset that is better distributed across the range of retention times in a chromatogram may reduce the training set size required to approach peak performance. For the global methods, more distributed alignment points would allow the systemic misalignment trends to be modeled with fewer points. Although this would also help the local method, it is already approaching peak performance with very few points in most cases, suggesting that better distributed alignment points might reduce the set size at which the performance of the global methods overtakes the local method.

A final consideration is the problem of incorrect alignment points. With a local method, the associated error is localized but larger; whereas with a global method the associated error is smaller but global. If alignment point errors are possible, a global method with many alignment points to regularize the fit may be preferred.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.6b02254.

Results for additional examples and additional statistical results (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: reich@unl.edu.

Notes

The authors declare the following competing financial interest(s): Prof. Reichenbach has a financial interest in GC Image, LLC, which produces GC Image GCxGC Software.

■ ACKNOWLEDGMENTS

This material is based upon work supported in part by the National Science Foundation at GC Image, LLC, under Grant No. IP-1127264 and in part by the UCARE Program at the University of Nebraska, Lincoln.

■ REFERENCES

(1) Reichenbach, S. E.; Rempe, D. W.; Tao, Q.; Bressanello, D.; Liberto, E.; Bicchi, C.; Balducci, S.; Cordero, C. *Anal. Chem.* **2015**, *87*, 10056−10063.

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

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

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

(5) 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.

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

(7) Pierce, K. M.; Wood, L. F.; Wright, B. W.; Synovec, R. E. *Anal. Chem.* **2005**, *77*, 7735−7743.

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

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

(10) Cordero, C.; Rubiolo, P.; Cobelli, L.; Stani, G.; Miliazza, A.; Giardina, M.; Firor, R.; Bicchi, C. *J. Chromatogr., A* **2015**, *1417*, 79−95.

(11) UOP, LLC. *Organic Analysis of Distillate by Comprehensive Two-Dimensional Gas Chromatography with Flame Ionization Detection*, UOP 990-11; 2011.

(12) Welke, J. E.; Zanus, M.; Lazarotto, M.; Schmitt, K. G.; Zini, C. A. *J. Braz. Chem. Soc.* **2012**, *23*, 678−687.

(13) GC Image, LLC. *GC Image GCxGC Software*; GC Image, LLC: Lincoln NE, 2015.

(14) Reichenbach, S. Data acquisition, visualization, and analysis. In *Comprehensive Two Dimensional Gas Chromatography*; Ramos, L., Ed.; Elsevier, 2009; pp 77−106.

(15) Sibson, R. A Brief Description of Natural Neighbor Interpolation. In *Interpreting Multivariate Data*; Barnett, V., Ed.; John Wiley & Sons: New York, 1981; pp 21−36.

(16) Gros, J.; Arey, J. *Documentation for the Gros-Arey code to align GCxGC chromatograms as implemented in MATLAB*; EPFL: Switzerland, November, 2014.

(17) Holland Computing Center. HCC Documentation. https://hcc-docs.unl.edu/display/HCCDOC/HCC+Documentation (accessed April 12, 2016).

DOI: 10.1021/acs.analchem.6b02254
Anal. Chem. 2016, 88, 10028−10035