(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 December 2024 (12.12.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/254413 A2

(51) International Patent Classification:
Not classified

(21) International Application Number:
PCT/US2024/032953

(22) International Filing Date:
07 June 2024 (07.06.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/506,895 08 June 2023 (08.06.2023) US

(71) Applicant: **BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM** [US/US]; 210 West 7th Street, Austin, Texas 78701 (US).

(72) Inventors: **ARMSTRONG, Daniel W.**; 611 Loch Chalet Court, Arlington, Texas 76012 (US). **WAHAB, Muhammad Farooq**; 801 Bering Drive, Apt 1117, Arlington, Texas 76013 (US). **HANDLOVIC, Troy T.**; 377 W Fork, Apt 3632, Irving, Texas 75039 (US). **ROY, Souvik**; 2608 Cinnamon Park Circle, Apt. 810, Arlington, Texas 76016 (US).

(74) Agent: **TIEFF, Michael** et al.; Meunier Carlin & Curfman LLC, 999 Peachtree Street NE, Suite 1300, Atlanta, Georgia 30309 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*
— *of inventorship (Rule 4.17(iv))*

Published:
— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

WO 2024/254413 A2

(54) Title: SYSTEMS AND METHODS FOR REMOVAL OF EXTRA-COLUMN BAND BROADENING IN CHROMATOGRAPHIC PEAKS

(57) Abstract: In some aspects, the techniques described herein relate to a chromatography device (500) for removing column band broadening including: a column (505); a detector (509); and at least one computing device (520), wherein the at least one computing device is adapted to: receive non-column data for each flow rate of a plurality of flow rates (600); cause an injection of a mixture of substances into the column at a selected flow rate from the plurality of flow rates (710); generate a first chromatogram based on the injection of the mixture of substances at the selected flow rate (720); and remove extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate a second chromatogram based on the first chromatogram with the removed extra column effects (750).

SYSTEMS AND METHODS FOR REMOVAL OF EXTRA-COLUMN BAND BROADENING IN CHROMATOGRAPHIC PEAKS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to and benefit of U.S. Provisional Application Serial No. 63/506,895, filed on June 8, 2023, and titled "SYSTEMS AND METHODS FOR REMOVAL OF EXTRA-COLUMN BAND BROADENING IN CHROMATOGRAPHIC PEAKS." The disclosure of which is hereby incorporated in its entirety.

BACKGROUND

[0002] Chemical separations by instrumental analysis have two main goals to (i) identify and (ii) quantitate components in a mixture. For any separation involving a chromatography column (packed or open tubular), the data collected relevant to these goals is always impaired by extra-column effects. Extra-column effects, or extra-column band broadening, arise from fluid flow phenomenon and detection electronics, which occur throughout the time the analyte spends outside the column during the chromatographic separation process. Although instrument manufacturers can reduce the instrument's physical volume and optimize the system's materials and geometries, there is not a single instrument in the world that does not have extra-column effects.

SUMMARY

[0003] Systems and methods to remove extra column effects by post-data acquisition mathematical calculations are provided. The proposed systems and methods are valid for most chromatographic separations, such as all gas chromatography (GC) techniques, and all liquid chromatography (LC) techniques. In addition, the technology works for all associated methods of detection such as ultraviolet (UV), circular dichroism (CD), all mass spectrometry (MS) detection techniques(QqQ, TOF, etc.), fluorescence, evaporative light scattering (ELSD), flame ionization detector, thermal conductivity detector and so forth. As separation columns become more efficient with the evolution of column technology and separations become faster, removing extra-column effects is increasingly relevant. This technology is expected to be of commercial interest to instrument

manufacturers globally. This technology will advance the use of narrow bore columns, microbore columns, miniaturized columns, and microfluidic instrument development.

[0004] In some aspects, the techniques described herein relate to a method for removing column band broadening in chromatography including: removing a column from a chromatography device; while the column is removed: causing a plurality of injections of a test substance into the chromatography device, wherein each injection is associated with a different flow rate of a plurality of flow rates; collecting non-column data for each flow rate of the plurality of flow rates; placing the column in the chromatography device; while the column is in the chromatography device: causing an injection of a mixture of substances into the chromatography device at a selected flow rate from the plurality of flow rates; generating a first chromatogram based on the injection of the mixture of substances at the selected flow rate; and removing extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate a second chromatogram based on the first chromatogram with the removed extra column effects.

[0005] In some aspects, the techniques described herein relate to a method, wherein removing the extra column effects from the first chromograph using the non-column data collected for the selected flow rate includes: using regularized deconvolution of the column data collected for the selected flow rate, followed by artifact and further noise removal by a mathematical filter.

[0006] In some aspects, the techniques described herein relate to a method further including displaying the second chromatogram.

[0007] In some aspects, the techniques described herein relate to a method, further including: displaying the first chromatogram; receiving an indication to remove the extra column effects from the first chromatogram; and removing the extra column effects from the first chromatogram using the non-column data in response to the received indication.

[0008] In some aspects, the techniques described herein relate to a method, wherein collecting non-column data for each flow rate of the plurality of flow rates includes: modeling peaks of the collected non-column data for each flow rate of the plurality of flow rates using a non-linear least squares regression to generate an asymmetric generalized normal; and modeling parameters of the asymmetric generalized normal vs. flow rate reciprocal using a linear least squares regression and interpolating the collected non-column

data with the modeled parameters to generate non-column data for any flow rate within a range of flow rates represented by the plurality of flow rates.

[0009] In some aspects, the techniques described herein relate to a method, wherein removing extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate the second chromatogram, includes: generating a circulant matrix from the non-column data for the selected flow rate; formulating a deconvolution problem using the first chromatogram and the circulant matrix, and solving the deconvolution problem using Tikhonov regularization to generate the second chromatogram.

[0010] In some aspects, the techniques described herein relate to a method, further including: denoising the second chromatogram using anisotropic diffusion.

[0011] In some aspects, the techniques described herein relate to a chromatography device for removing column band broadening including: a column; a detector; and at least one computing device, wherein the at least one computing device is adapted to: receive non-column data for each flow rate of a plurality of flow rates; cause an injection of a mixture of substances into the column at a selected flow rate from the plurality of flow rates; generate a first chromatogram based on the injection of the mixture of substances at the selected flow rate; and remove extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate a second chromatogram based on the first chromatogram with the removed extra column effects.

[0012] In some aspects, the techniques described herein relate to a chromatography device, further displays the second chromatogram.

[0013] In some aspects, the techniques described herein relate to a chromatography device, further including: displaying the first chromatogram at the selected flow rate; receiving an indication to remove the extra column effects from the first chromatogram; and removing the extra column effects from the first chromatogram using the non-column data in response to the received indication.

[0014] In some aspects, the techniques described herein relate to a chromatography device, wherein removing extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate the second chromatogram, includes: generating a circulant matrix from the non-column data for the selected flow rate; formulating a deconvolution problem using the circulant matrix and the first chromatogram;

and solving the deconvolution problem using Tikhonov regularization to generate the second chromatogram, followed by artifact and noise removal using anisotropic diffusion.

[0015] In some aspects, the techniques described herein relate to a chromatography device, further including: displaying the first chromatogram; receiving an indication to remove the extra column effects from the first chromatogram; and removing the extra column effects from the first chromatogram using the non-column data in response to the received indication.

[0016] In some aspects, the techniques described herein relate to a chromatography device, further including: removing the column; while the column is removed: causing a plurality of injections of a test substance into the chromatography device, wherein each injection is associated with a different flow rate of the plurality of flow rates; and collecting the non-column data for each flow rate of the plurality of flow rates.

[0017] In some aspects, the techniques described herein relate to a non-transitory computer-readable medium with computer executable instructions stored thereon than when executed by a chromatography device cause the chromatography device to: receive non-column data for each flow rate of a plurality of flow rates; cause an injection of a mixture of substances into a column of the chromatography device at a selected flow rate from the plurality of flow rates; generate a first chromatogram based on the injection of the mixture of substances at the selected flow rate; and remove extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate a second chromatogram based on the first chromatogram with the removed extra column effects.

[0018] In some aspects, the techniques described herein relate to a non-transitory computer-readable medium, further displays the second chromatogram.

[0019] In some aspects, the techniques described herein relate to a non-transitory computer-readable medium, further including: displaying the first chromatogram at the selected flow rate; receiving an indication to remove the extra column effects from the first chromatogram; and removing the extra column effects from the first chromatogram using the non-column data in response to the received indication.

[0020] In some aspects, the techniques described herein relate to a non-transitory computer-readable medium, wherein removing extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate

the second chromatogram, includes: generating a circulant matrix from the non-column data for the selected flow rate; formulating a deconvolution problem using the circulant matrix and the first chromatogram; and solving the deconvolution problem using Tikhonov regularization to generate the second chromatogram, followed by artifact and noise removal using anisotropic diffusion.

[0021] In some aspects, the techniques described herein relate to a chromatography device, further including: displaying the first chromatogram; receiving an indication to remove the extra column effects from the first chromatogram; and removing the extra column effects from the first chromatogram using the non-column data in response to the received indication.

[0022] In some aspects, the techniques described herein relate to a non-transitory computer-readable medium, further including: removing the column; while the column is removed: causing a plurality of injections of a test substance into the chromatography device, wherein each injection is associated with a different flow rate of the plurality of flow rates; and collecting the non-column data for each flow rate of the plurality of flow rates.

[0023] In some aspects, the techniques described herein relate to a non-transitory computer-readable medium, wherein the chromatography device is a gas or a liquid chromatography device.

[0024] Additional advantages of the invention will be set forth in part in the description which follows, and in part will be obvious from the description, or may be learned by practice of the invention. The advantages of the invention will be realized and attained by means of the elements and combinations particularly pointed out in the appended claims. It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention, as claimed.

BRIEF DESCRIPTION OF THE DRAWINGS

[0025] The accompanying figures, which are incorporated herein and form part of the specification, illustrate a system and method for removing column band broadening in chromatography. Together with the description, the figures further serve to explain the principles of the system and method described herein and thereby enable a person skilled in

the pertinent art to make and use the system and method for removing column band broadening in chromatography.

[0026] FIG. 1 is an illustration of graphs showing shifted hyperbolic trends for fit parameters of the GAN function versus flow rate, with the asymmetry term (a_3) locked at the average value;

[0027] FIG. 2 is an illustration of graphs showing experimentally collected naphthalene peaks (top) without a column and corresponding predicted peaks (bottom);

[0028] FIG. 3 is an illustration of graphs showing a comparison of different frequency and time (orange) domain deconvolution techniques for an ultrafast separation;

[0029] FIG. 4 is an illustration of graphs showing the application of Tikhonov regularized deconvolution with Perona-Malik filtering to the enantioseparation of R/S-5-methyl-5-phenylhydantoin;

[0030] FIG. 5 is an illustration of an example chromatography device;

[0031] FIG. 6 is an illustration of an example method for collecting non-column data for removing column band broadening in chromatography;

[0032] FIG. 7 is an illustration of an example method for removing column band broadening in chromatography using collected non-column data; and

[0033] FIG. 8 shows an exemplary computing environment in which example embodiments and aspects may be implemented.

DETAILED DESCRIPTION

[0034] Despite significant instrumentation advances in separation sciences, it is currently impossible to find a commercial liquid chromatography device that does not have extra-column band-broadening effects. An example device is the chromatography device 500 illustrated with respect to FIG. 5. As shown, the device 500 includes an injector 501, tubing 508, a frit 503, a column 505, an exit frit 504, connections 507, and a detector 509. The detector 509 is connected to a computing device 520 such as the computing device 800 illustrated with respect to FIG. 8.

[0035] When a sample mixture is injected into the device 500 by the injector 503, starting from a sample introduction port of the injector 501, the sample is broadened by the injector 501 and the tubing 508 connecting to the column 505. When the injected sample

reaches the head of the column 505, the sample is further broadened by the inlet frit 503. The "ideal" column 505 then acts as a Gaussian operator or a so-called action function on each element of the sample mixture. In practice, the column 505 adds peak delay, broadening, and higher moment shape changes. After exiting the packed bed of the column 505, further band broadening occurs at the exit frit 504 before the analyte bands undergo additional distortion in the post-column tubing of the connections 507, leading to the detector 509. The flow cell geometry and digital signal processing of the detector 509 impart a final observed peak shape of a resulting chromatogram. The broadening or peak shape distortion that occurs outside the column 505 in chromatography has been widely recognized in the literature for more than six decades, starting with the classical work of Sternberg in gas chromatography. The observed chromatogram can be modeled by the mathematical convolution of an extra-column function and the actual band shape.

[0036] With the availability of bonded sub-2 μm fully porous or superficially porous particles in short columns, plate heights, H, on the order of $2\times$ particle diameter, have been routinely observed in both chiral and achiral chromatography. Particle technology continues to challenge the engineering aspects of designing chromatographic instruments with low extra-column volumes and advanced digital filters. Reducing the physical dimensions of chromatographic systems by minimizing injection volumes, narrowing connection tubing, and optimizing flow cell geometry comes at a cost. Specifically, increased back pressure is a concern due to the quadratic relationship between pressure and tubing diameter, narrow tubings can be more prone to clogging, and a smaller flow path in spectroscopic measurements may require a compromise in detector sensitivity. In contrast, some detection schemes cannot employ low extra-column volume cells. For instance, the evaporative light scattering (ELSD) and circular dichroism (CD) detectors have significant detector (unspecified) volumes. The former has a large volume nebulization tube after the column, allowing only fine spray to enter the heating tube and light scattering cell. On the other hand, CD is a differential absorption technique; hence, inherently insensitive, and a longer path length is required.

[0037] Researchers have made several efforts to mathematically remove the instrumental convolution effects from chromatograms with variable success. Burke et al. applied deconvolution in the Fourier domain and demonstrated its advantages in liquid

chromatography using a no-column peak as an instrument response function. Fourier deconvolution for removing extra-column broadening was recently explored for UHPLC, for microbore columns, and in ultrafast separations. Fourier methods are attractive because they achieve deconvolution in a single step by dividing the Fourier transform of the signal with the Fourier transform of the broadening function, followed by the inverse transform. Unfortunately, this division process is often the source of many problems, in which noise is amplified, and occasionally, Not-a-Number (NaN) errors are displayed. The noise can be suppressed by applying a high-frequency filter in the Fourier domain, but artifacts can persist.

[0038] Although seminal work has been done on deconvolution mathematics in astronomical/medical imaging and spectroscopic techniques, wherein their applications have become routine, there is a lack of widespread acceptance of deconvolution science within the chromatography community. In this application is described a robust framework to automate the deconvolution of chromatograms while reducing artifacts in liquid chromatography to the point where peak efficiencies and symmetries acquired from routine HPLC instrumentation approach that of those acquired from low volume UHPLC systems. Several new approaches are proposed to address the ill-posed nature of this chromatographic deconvolution problem, including Tikhonov regularization. To model a noise-free instrumental response as a function of pump flow rates, a numerically stabilized generalized asymmetric normal (GAN) distribution is proposed. A training dataset of instrument response (acquired from no-column injections) is fit using the GAN model to extract parameters describing the instrumental distortion using the interior point optimization algorithm. With these parameters, one can generate instrument response functions without noise at a given flow rate.

Theory

[0039] Using deconvolution, we seek to remove every distortion except the portion of the peak that forms in the chromatographic separation in the column. Herein, the term vector denotes a sequence of discrete data points, e.g., the chromatographic signal. This problem can be mathematically represented as a convolution problem:

$$\mathbf{Cs} = \mathbf{o} \quad (1)$$

where $\mathbf{C}$ is a rectangular Toeplitz matrix corresponding to the convolution operator, $\mathbf{s}$ is the undistorted chromatogram, and $\mathbf{o}$ is the experimentally observed chromatogram represented as vectors. The inverse problem is to recover the unknown $\mathbf{s}$, from the given $\mathbf{o}$ and an estimated matrix $\mathbf{C}$.

$$\mathbf{s} = \mathbf{C^{-1}o} \quad (2)$$

where $\mathbf{C}^{-1}$ is the Moore-Penrose pseudoinverse. This inversion is equivalent to previously reported Fourier deconvolution methods without a filter function. In this work, the operations are conducted directly in the time domain. However, the problem is ill-posed because minor variations in the elements of the $\mathbf{C}$ matrix may cause huge noise variations in the recovered signal $\mathbf{s}$ during matrix inversion. Thus, numerical matrix inversion methods, like Gaussian elimination, lead to large noise amplitudes. In this context, the Tikhonov regularization process[17,18] can stabilize deconvolution and allow an estimate of the true chromatogram by formulating a minimization problem to minimize $\mathbf{\Phi}[\mathbf{s}]$, which is given as follows:

$$\mathbf{\Phi}[\mathbf{s}] = \|\mathbf{o} - \mathbf{Cs}\|^2 + \lambda\|\mathbf{s}\|^2 \quad (3)$$

where λ is a positive constant. Equation (3) proposes the problem as a regularized least-squares problem, where the first term in $\mathbf{\Phi}[\mathbf{s}]$ is the least-squares data fitting term, and the second term is a regularization term that controls wild oscillations in the estimated signal. The symbol $\|\cdot\|$ indicates the 2-norm of the vectors. Taking the first derivative of Equation (3) using vector calculus and setting the derivative equal to zero gives the normal equations,

$$(\mathbf{C^TC} + \boldsymbol{\lambda}\mathbf{I})\mathbf{s} = \mathbf{C^To} \quad (4)$$

and by taking the matrix inverse on both sides, we obtain,

$$\mathbf{Is} = \mathbf{s} = (\mathbf{C^TC} + \boldsymbol{\lambda}\mathbf{I})^{-1}\mathbf{C^To} \quad (5)$$

with $\mathbf{I}$ as the identity matrix. Equation (5) is a stabilized inversion of $\mathbf{C}$ to obtain $\mathbf{s}$. The matrix $\mathbf{C}$ is chosen specifically as a circulant matrix in this work. The instrument response row vector, $\mathbf{c}$, at a given flow rate, is generated by removing the column 505 and

connecting the tubings 508 with zero dead volume unions. Then the convolution matrix is constructed from the elements of **c** by a cyclic shifting of the elements of the vector **c**.

$$\mathbf{C} = \begin{bmatrix} c_0 & c_{n-1} & \cdots & c_2 & c_1 \\ c_1 & c_0 & c_{n-1} & & c_2 \\ \vdots & c_1 & c_0 & \ddots & \vdots \\ c_{n-2} & & \ddots & \ddots & c_{n-1} \\ c_{n-1} & c_{n-2} & \cdots & c_1 & c_0 \end{bmatrix} \quad (6)$$

[0040] Since the no-column peak is rapidly eluted and **c** has the same length as **o**, most of the elements of the vector **c** are numerically close to zero. Using the circulant matrix allows the data output length to be identical to the original chromatogram after deconvolution.

[0041] Different artifacts in a deconvolved chromatogram, such as ringing and negative dips, are common. These issues can be suppressed by the Perona-Malik (PM) anisotropic diffusion filtering method. PM filtering, also known as anisotropic filtering, is an advanced form of standard weighted filtering (e.g., Gaussian). While the Gaussian filtering smoothens out all regions in the signal uniformly, thus, killing signal peak resolution alongside oscillations, PM filtering smooths in a non-uniform way to preserve peak resolution according to equation 7:

$$s_t = \nabla \cdot (c(x,t)\nabla s); \;\; s(0) = s_0 \quad (7)$$

[0042] Here ∇ is the derivative of the signal, t is the parameter for the number of refinements of the unfiltered signal s_0. Starting from s_0, we generate new sets of chromatograms for s through the evolution equation (7) till a certain stage T. At that stage T, we obtain a signal that has sharp edges and no or minimal oscillations resulting in a high resolution reconstruction. We note that in the regions of the domain of the signal s where the magnitude of $|\nabla s| >> 1$, it usually resembles a peak, and, in that case, (7) is essentially $s_t = 0$ in those regions. This pure advection equation preserves the same solution on these regions from the previous refinement iteration. On the other hand, the regions where $|\nabla s| << 1$ are expected to correlate with noise oscillations. In those regions, (7) represents non-uniform filtering and, thus, smoothens out oscillations while preserving peak resolution. Such a filtering method has recently been employed in tomographic medical imaging to reconstruct 2D images with high contrast and resolution of the finest detail.

Experimental Results and Discussion

[0043] We employ the standard procedure for collecting the HPLC's response function but use the data only for modeling to generate noise-free instrument responses. Using raw experimental data as the deconvolving function has several practical downsides. The broadening function may be collected at every flow rate at which future data will be collected. It also may be collected at an identical sampling rate as future data. Further, any deviations in the baseline from a perfect zero will alter the area of the deconvolved chromatogram, and any imperfections in the experimental data will be transferred to the deconvolved chromatogram. Therefore, when developing an automated algorithm, all the above issues are circumvented by using a noise-free, perfectly zeroed baseline-modeled deconvolution response function (Equation 8).

[0044] To develop a modeling and prediction framework, experimental no-column data at various flow rates were collected on three different but optimized instruments (UHPLC-UV, HPLC-CD, HPLC-UV, see SI-1). This data used a pure methanol mobile phase with the same naphthalene sample. A concentration series on the UHPLC found that a 20 ppm solution has a suitable response without overloading the detector. For all instruments, 268 nm was used since it is one of the secondary absorbance maxima of naphthalene. Resultant chromatograms from the training set were fit in PeakLab using available peak functions, looking at the fit quality and patterns for the fit parameters related to flow rate. No common chromatography peak function (*e.g.,* Gaussian, exponentially modified Gaussian, Gaussian modified Gaussian, Giddings, etc.) satisfactorily fit all the no-column data. More complex peak functions, such as variations of the Haarhoff-Van der Linde (HVL) based function or the Non-Linear Chromatography (NLC) based function, had extremely confident fits but had parameters that did not trend clearly as a function of flow rate. The only peak function in PeakLab that was found to have satisfactory fits for all instruments and whose parameters had a clear trend as a function of flow rate was a generalized asymmetric normal (GAN).

[0045] GAN is a four-parameter log-normal distribution first discussed in 1917 by van Uven and reviewed in detail by Johnson and coworkers. The GAN can fit various distributions with possible variability in the higher moments using only four parameters

(area, location, width, and symmetry). However, the natural domain of the log function limits its applicability as a proper curve-fitting function for real signal values with noise. To stay in the valid interval of real numbers, the point at which the log term becomes negative is placed into a Heaviside theta function $\theta(x)$. This allows $\theta(x)$ to "step" to zero at the t value, where the log term becomes invalid with real negative arguments. This modification multiplies the undefined term by zero, outputting zero. The modified GAN is,

$$c[t] = \frac{a_0 \exp - \frac{\log\left[\exp\left(\frac{a_3^2}{2}\right) + \frac{(t-a_1)a_3}{a_2}\right]^2}{2a_3^2}}{\sqrt{2\pi}\left(\exp\left(\frac{a_3^2}{2}\right)a_2 + ta_3 - a_1a_3\right)} \theta\left(\left[\frac{\text{sign}[a_3]\left(\exp\left(\frac{a_3^2}{2}\right)a_2 + ta_3 - a_1a_3\right)}{a_3}\right]\right) \quad (8)$$

the function is parameterized so that a_0 is peak area, $a_1$is the mean of the asymmetric distribution,a_2 is the standard deviation of the pre-distorted Gaussian, a_3 is statistical symmetry (-1 to 1), and t is the time.

Instrumental Peak Shape Prediction and Formation of the Broadening Function

[0046] When fitting the GAN to the training dataset, the asymmetry term (a_3) had inconsistent behavior as a function of flow rate for most trials. To simplify the predictions, the asymmetry term was averaged after fitting all data for each instrument, and then the data were sequentially refitted with the a_3 term locked. Trends for the GAN parameters as a function of flow rate are displayed in the graphs 100 of FIG 1.

[0047] The flow rates were recorded in triplicate, and all data points were plotted. The trendlines are shifted rectangular hyperbolic functions following the Equation of:

$$a_i = \frac{m}{F} + b \qquad i = 0,1,2 \qquad (9)$$

where a_i is the fit parameter value, m is a constant, F is the flow rate, and b is the magnitude of the shift. The m and b values can be computed by linear regression of a_i versus inverse flow rate ($1/F$), in practice, b is near zero. By constructing these plots, no-column data can be rapidly collected at a time just long enough to secure the signal (< 30 s), and then the broadening function can be predicted at any flow rate on any time vector containing any

desired sampling rate. A convolution matrix can then be formed through cyclic shifting of the output (Equation (6)).

[0048] To make this process automated in the algorithm, a least-squares function was formulated to find the best parameter (a_0 — a_2) values for each peak. This least squares minimization problem is automatically solved using MATLAB's "fmincon" (Optimization Toolbox) function with the interior point optimization algorithm. The interior point algorithm minimizes an associated unconstrained optimization problem using an approximate logarithmic barrier function. The algorithm sets a threshold for the predicted change in the function to be small enough for an accurate minimization process. The lower and upper bounds of the parameters are set based on predicted parameter behaviors, which helps in an accurate solvability of the minimization problem. Convergence was reached for each chromatogram produced by every instrument tested without changing the bounds. Trendlines are computed using the "polyfit" (1st order, a_i vs. $1/F$) function, and m and b are extracted using the "polyval" function.

[0049] The data from this training set, displayed in the graphs 200 of FIG. 2, exhibits the vast differences in system peak size and shape across the same conditions. The UHPLC contains the smallest system volume (detector flow cell of 2.5 µL) and has the narrowest and symmetric no- column peaks. The HPLC-UV contains a moderate volume detector but still produces symmetric peaks. The circular dichroism detector allows for simultaneous CD and UV detection. Since the flow path and detector electronics are identical, a UV peak can be used to determine the broadening function for a CD chromatogram using a non-interacting achiral probe like naphthalene. The HPLC-CD (detector flow cell of 44 µL) produces the broadest peaks with a high level of asymmetry. At some flow rates in the HPLC-CD data, a shoulder can be seen on the trailing edge of the peak. Gritti and coworkers have investigated this behavior in capillaries (250 µm x 65 cm) using a butyrophenone probe and attribute these shoulders to an effect of when molecular diffusion does not have enough time or speed to average the concentration gradient found across the capillary. This results in a fast-moving zone in the center of the capillary, and a slow-moving zone on the wall, creating a lagging shoulder. However, by our modeling, this shoulder is removed.

[0050] The graphs 200 of FIG. 2 display the calculated peaks at each referenced flow rate using the automated algorithm and Equations (8&9). The predicted peaks

match well with the experimental peaks and contain perfect continuity in the function without noise and zero baselines. The improvement in peak shape can be seen most clearly by comparing the experimental peak for the HPLC-CD 0.4 mL min^{-1} to the calculated peak. It should be noted that these plots could not be calculated for the HPLC-ELSD as the manual injection system has too much variability. In these situations, or when an instrument is only operated at one flow rate, single point calibration can be used, inputting only one flow rate data in replicate to this same algorithm.

Comparison Of Time And Frequency Domain Deconvolution Methods

[0051] As mentioned in the theory section, deconvolution in the presence of noise, is an ill-posed problem that has been described as “violently unstable” by early pioneer N. E. Scofield in 1962. A comparison of different methods to deconvolve extra-column effects is displayed in the graphs 301, 303, 305, 307, 309, and 311 of FIG. 3 for an ultrafast (7 peaks 1 minute), achiral separation on a short narrow bore C-18 column (50x2.1 mm). This chromatogram is a difficult situation for deconvolution as the broadening function width approached the width of the first peak of the chromatogram. Deconvolution in the Fourier domain produces a wild noise amplification from dividing the FFT of the signal with small numbers by the FFT of the deconvolving function. Without apodization in FFT deconvolution, the output contains an immense magnitude of noise (on the order of 5×10^{13} μAU), burying all analytical information shown in the graph 301.

[0052] The simplest window shape to control this noise is a rectangle that effectively acts as a "frequency cut-off filter." This method has been used in extra-column corrections previously. A sharp frequency cut-off filter (rectangular window) is applied in the graph 303. Here a high amount of resolution is produced, and the noise is effectively suppressed, but the sharp truncation of the data in the frequency domain produces "sinc" artifacts. These dips produce a negative area for the peaks and make their positive regions artificially tall to account for the negative area. Alternative FFT windows, such as Gaussian and super-Gaussian windows, have also been used for extra-column removal. The Equation for Gaussian-based windows is:

$$\mathrm{G} = \exp\left[-\left(\frac{|k - Z|}{W}\right)^{p}\right], \qquad \mathrm{G}[k] = \text{circular shift}\left[\mathrm{G}, \frac{N}{2}\right] \tag{10}$$

where Z is the mean value of k, W is related to the window's width, and p is the power of the super-Gaussian (2 for Gaussian). A Gaussian window produces a smoother decay in the frequency domain, suppressing the dips seen with the rectangular window. The result of the application of Equation (10) is shown in the graph 305. Here, we see how the smoother decay produces a less distinct cut-off, and some undesirable frequencies bleed through (i.e., noise), producing the common "ringing" or Gibb's phenomena artifacts. These artifacts can be reduced or eliminated using a narrower Gaussian window (smaller W), but this operation comes at the cost of broader and shorter deconvolved peaks. The ringing artifacts would be even more predominant in frequency domain deconvolution outputs if the data were collected at a higher sampling rate.

[0053] Time domain deconvolution techniques like convolution matrix pseudo-inversion without regularization produce mathematically identical output to the graph 301. In practicality, this approach is often numerically more stable than frequency domain techniques producing a slightly better output, but it is slower and much more computationally demanding to perform on a computing device. To improve the speed of matrix operations, all numbers in the matrices formed can be treated with a conditional so that numbers smaller than double precision machine epsilon ($\varepsilon = 2.2204 \times 10^{-16}$) are replaced with zero prior to operations. This greatly enhanced the speed of all-time deconvolution methods tested; however, we recommend using these methods for 20,000 points or fewer in the chromatograms. The algorithm is best suited for fast separations since these are the most affected by extra-column effects.

[0054] The graph 307 presents the first chromatographic application of Tikhonov regularized deconvolution ($\lambda = 0.0001$) with the 2-norm penalty (Equation 5). Here, the ringing is suppressed, and the negative dips are smaller than in graph 303 but still noticeable. Increasing the weight of the regularization term will decrease these dips but broaden the peaks. Another constrained time domain technique takes a similar approach to Tikhonov regularization; however, the 2-norm penalty is replaced with the norm of the second difference of the chromatogram. Minimizing this function, just as was done for Equation 5, and solving for $\mathbf{x}$ provides the following Equation:

$$\mathbf{s} = (\mathbf{C}^{\mathrm{T}}\mathbf{C} + \boldsymbol{\lambda}\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1}\mathbf{C}^{\mathrm{T}}\mathbf{o} \quad (11)$$

where **D** is a matrix (second difference matrix can be used). Equation (11) is applied to produce the results shown in the graph 309. The graph 309 produces a similar result to the 2-norm penalty but requires greater weight to be placed on the penalty (λ = 0.005). Eilers and Ruckebusch have recently applied a combination of 2-norm penalties of signal vector and second difference in image processing, but our chromatographic data did not benefit from this approach. Alternatively, Eilers and Goeman have combined first and second difference penalties to reduce ringing in histograms. However, all methods involving the multiplication of difference matrices are computationally intensive for chromatograms containing tens of thousands of data points.

[0055] The disclosed method combines the linear deconvolution process with non-linear Perona- Malik filtering to fine-tune the actual output (the graph 307). This is shown in the graph 311. In cases where Tikhonov's raw output produces negative dips, the Perona-Malik algorithm is an excellent option for post-processing and removing those artifacts. The graph 311 has no noticeable ringing or dips after deconvolution and PM filtering. Removing these negative dips corrects the peak height and gives a more accurate indication of column-only efficiency. Popular smoothing algorithms in chemistry, such as the "perfect smoother" or Savitzky-Golay, cannot effectively remove such dips.

Practical Applications Of Tikhonov Regularization With Perona-Malik Filtering On A Variety Of Liquid Chromatographs

[0056] There is an industry-wide need for fast, high-efficiency, and high-throughput chiral separations to determine enantiomeric excess in different fields. Fast separations, with limited retention and therefore limited time in the column for broadening, are susceptible to extra-column effects. In the graphs 401, 403, 405, and 407 of FIG. 4, the same sample, *R/S*-5-methyl-5-phenyl-hydantoin, mobile phase, and column, is analyzed on four systems UHPLC-UV, HPLC-UV, HPLC-CD, HPLC- ELSD. One signal indicates the raw data, with the other indicating the "column-only" deconvolved result. Table 1, contains moment analysis (M_0 through M_4) using GAN fits for these peaks. Plates were calculated using moments to avoid the assumption of Gaussian shape, using the relationship of $N = M^2/M$.

On the UHPLC system (the graph 401), the first enantiomer has 17,200 plates raw (h = 2.15) and 17,800 plates deconvolved (h = 2.08) with the second enantiomer showing 11,200 plates raw and 11,300 plates deconvolved. The original resolution is 4.38 and after removing extra-column effects it becomes 4.46. Interestingly, the higher moments (3rd and 4th) remain nearly unchanged, indicating the instrument primarily added symmetric broadening.

Table 1: Moment analysis for *R*/*S*-5-methyl-5-phenylhydantoin on multiple instruments

	M_0, Area		M_1, Mean		$10^2\sqrt{M_2}$, Std. Dev.		M_3, Skewness		M_4, Kurtosis	
Peak	Raw	Decon.	Raw	Decon.	Raw	Decon.	Raw	Decon.	Raw	Decon.
A1	266	263	0.989	0.981	0.753	0.735 (-2.35%)	0.256	0.257	3.12	3.12
A2	302	299	1.18	1.17	1.12	1.10 (-1.10%)	0.431	0.436	3.33	3.34
B1	3.79	3.71	1.03	1.00	0.870	0.738 (-15.1%)	0.296	0.302	3.16	3.16
B2	3.56	3.50	1.23	1.20	1.26	1.16 (-7.48%)	0.502	0.555	3.45	3.55
C1	-335	-334	1.04	1.00	1.55	0.889 (-42.8%)	0.929	0.209	4.58	3.08
C2	333	320	1.25	1.21	2.02	1.29 (-35.9%)	0.940	0.274	4.61	3.13
D1	18300	17300	1.12	1.01	1.59	0.996 (-37.3%)	1.061	0.714	5.07	3.92
D2	19900	19200	1.32	1.22	1.93	1.55 (-19.9%)	0.992	1.103	4.80	5.24

Peak ID correlates to FIG. 4, where Peak A1 is the first peak in the graph 401, Peak A2 is the second peak in the graph 401, and so on. M_0 has units of µAU min for peaks A1 to B2 and µV min for peaks C1 to D2. M_1 and $\sqrt{M_2}$ have units of minutes. M_3 and M_4 are unitless. For a Gaussian distribution M_3 is zero and M_4 is three.

[0057] The instrument with the next smallest extra column volume is the HPLC-UV (the graph 403), where the first enantiomer has 14,000 plates (h = 2.64) and 18,400 plates deconvolved (h = 2.01) with the second enantiomer showing 9,600 plates raw and then 10,700 deconvolved. Here, the higher moments also remained relatively unfazed as well. The discrepancy between the column-only plate count between these two instruments can be attributed to the applied broadening functions being estimates, possibly since the flow path is continuous in the no- column data and disrupted with a column, and the effects that the instrument's digital filtering impart on the final peak shape. In this instance, the gain in resolution is from 4.07 to 4.56. It should be noted that with the proper removal of extra-column effects outlined herein, peak efficiencies and higher moments

obtained from ordinary HPLCs (using 2.7 µm SPP containing columns) can be equivalent to a highly optimized UHPLC. This result confirms that this technology can successfully deconvolve the widely varying no-column data shapes and produce true column-specific peaks, independent of vastly different instruments.

[0058] The graphs 405 and 407 display the separation on two large volume detectors (CD and ELSD). The CD chromatogram contains an efficiency of only 4,500 plates (h = 8.23) for the first peak and 3,850 plates for the second peak. The separation lost over 12,700 plates in comparison to the UHPLC separation. After deconvolution, over 8,000 plates are gained back on the first peak producing a deconvolved plate count of 12,600 (h = 2.93), and the second peak now has 8,730 plates. This result again confirms that the broadening function is an estimation, and the impact of filtering cannot be reversed. In this case, the higher moments are changed, with the skewness being reduced from 0.929 to 0.209 for the first peak and 0.940 to 0.274 for the second peak. In addition, the excess kurtosis is reduced from 1.58 to 0.08 for the first peak and 1.61 to 0.13 for the second peak. The deconvolved higher moments now fall within a range seen on the UHPLC. Therefore, the CD detector broadened the peak and added a high level of asymmetry.

[0059] The ELSD portion of the graph 407 shows an example of how discontinuities in the data will be enhanced and effectively amplified during deconvolution. Looking at the pointed regions of the graph 407, the shoulder and splits on the original chromatogram are more apparent and better resolved in the deconvolved chromatogram. After deconvolution, the chromatogram has 10,300 (h = 3.6) plates for the first peak and 6,190 plates for the second peak. In this case, these measures do not hold the highest accuracy as artifacts in the peaks alter the fitting capability. Therefore, there will be errors in the moment analysis, especially in the kurtosis and skewness. This can be seen in Table 1, where peak "D2" has significantly more skew and kurtosis after deconvolution as a direct result of the artifact on the trailing edge of the peak. Therefore, moment analysis should only be taken seriously in the presence of a strong fit assessed by residuals.

[0060] An indication of system volume and column-only data accuracy is the alignment of peaks across instrumentation after deconvolution. The shift in retention time between the raw chromatogram and the deconvolved chromatogram accounts for the time that the analyte spends outside the column. Therefore, when the chromatograms of the graphs 401, 403, 405, and 407 of FIG. 4 are deconvolved,

all peaks should be aligned. After true signal recovery, all instruments with an autosampler (excluding HPLC-ELSD) contained first moments for the first peak within a 1.07 second range, revealing that reproducibility between instruments is improved through the proposed algorithm. The HPLC-ELSD peaks are still not aligned due to manual injection and timing. Another point to note is that even after deconvolution, there is still a small amount of skew and excess kurtosis. This indicates that this chiral column is not an ideal Gaussian operator (under these conditions) and suffers from some of the characteristics of chiral chromatography that produce asymmetric peaks.

Example Methods

[0061] FIG. 6 is an illustration of an example method for collecting non-column data for removing column band broadening in chromatography. The method 600 may be implemented by the chromatography device 500 illustrated with respect to FIG. 5.

[0062] At 610, a column is removed from a chromatography device. The chromatography device 500 may be a commercial liquid or gas chromatography device.

[0063] At 620, a plurality of injections of a test substance are caused to be injected into the chromatography device 500 while the column is removed. The test substances may be injected by the injector 501 and may include substances such as naphthalene, thiourea, uracil, or any unretained pure compound which does not chemically interact with the parts of the chromatograph or tubings. Each injection of the substance may be at a different flow rate with a range of flow rates between the flow rates of interest to the user. Typically six different flow rates should be sufficient. The same injection can be replicated at a given flow rate.

[0064] At 630, non-column data is collected for each injection and flow rate. The non-column data may be collected from the detector 509 by the computing device 520. In some embodiments, as part of the collecting of the non-column data, the computing device 520 may model peaks of the collected non-column data for each flow rate using a non-linear least squares regression to generate an asymmetric generalized normal. The computing device 520 may further model parameters of the asymmetric generalized normal vs. flow rate reciprocal using a linear least squares regression and interpolate the collected non-

column data with the modeled parameters to generate non-column data for any flow rate within a the range of flow rates.

[0065] At 640, the non-column data collected for each flow rate may be stored by the computing device 520 and associated with the chromatography device 500. As will be described further with respect to the method 700, the non-column data collected for the chromatography device 500 may be later used to remove column band broadening from a subsequent chromatogram generated by the chromatography device 500 after the column 505 has been added back to the chromatography device 500.

[0066] FIG. 7 is an illustration of an example method for removing column band broadening in chromatography using collected non-column data. The method 700 may be implemented by the chromatography device 500 illustrated with respect to FIG. 5.

[0067] At 710, a mixture of substances in caused to be injected in a chromatography device at a selected flow rate. A user may desire to generate a chromatogram that shows the concentrations of some or all of the substances in the mixture. Each substance may be represented by a peak in the generated chromatogram. The chromatography device 500 may include a column 505. As described above, characteristics of the column 505 may cause broadening of the peaks in the resulting chromatogram.

[0068] At 720, a first chromatogram is generated based on the injection. The first chromatogram may be generated by the computing device 520 of the chromatography device 500 based on data received from the detector 509. Because the column 505 was included in the chromatography device 500, the first chromatogram may include column band broadening associated with the column 505. The data collected by the chromatography device 500 from the injection that is used to generate the first chromatogram is referred to as the column data.

[0069] At 730, the first chromatogram is displayed. The first chromatogram may be displayed to a user on a display associated with the chromatography device 500 and/or the computing device 520.

[0070] At 740, an indication to remove the extra column effects from the first chromatogram is received. The indication may be received from the user through the computing device 520 of the chromatography device 500. For example, the user may select a user interface element displayed by the computing device 520 that asks the user whether

they would like to see a version of the first chromatogram with the extra column effects removed.

[0071] At 750, in response to the indication, the extra column effects of the column are removed from the first chromograph using the non-column data. The non-column data may be the non-column data generated for the column at the selected flow rate used for the injection at 710. In some embodiments, the computing device 520 may remove the extra column effects from the first chromograph using regularized deconvolution of the column data using the non-column data associated with the same flow rate. In addition, the computing device 520 may perform noise removal from the resulting data using a mathematical filter.

[0072] In some embodiments, the computing device 520 may remove the extra column effects by first generating a circulant matrix from the non-column data collected for the selected flow rate. The computing device 520 may then formulate a deconvolution problem using the column data and the circulant matrix. Finaly, the computing device 520 may solve the deconvolution problem using Tikhonov regularization. The resulting data can then be used to generate the second chromatogram (i.e., the first chromatogram with the extra column effects removed).

[0073] At 760, the second chromatogram is displayed. The second chromatogram is displayed in the same on the display associated with the chromatography device 500 and/or the computing device 520. Depending on the embodiment, the user may switch between viewing the first and second chromatograms that are displayed or may display both chromatograms at the same time to allow for comparison.

[0074] FIG. 8 shows an exemplary computing environment in which example embodiments and aspects may be implemented. The computing device environment is only one example of a suitable computing environment and is not intended to suggest any limitation as to the scope of use or functionality.

[0075] Numerous other general purpose or special purpose computing devices environments or configurations may be used. Examples of well-known computing devices, environments, and/or configurations that may be suitable for use include, but are not limited to, personal computers, server computers, handheld or laptop devices, multiprocessor systems, microprocessor-based systems, network personal computers (PCs),

minicomputers, mainframe computers, embedded systems, distributed computing environments that include any of the above systems or devices, and the like.

[0076] Computer-executable instructions, such as program modules, being executed by a computer may be used. Generally, program modules include routines, programs, objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types. Distributed computing environments may be used where tasks are performed by remote processing devices that are linked through a communications network or other data transmission medium. In a distributed computing environment, program modules and other data may be located in both local and remote computer storage media including memory storage devices.

[0077] With reference to FIG. 8, an exemplary system for implementing aspects described herein includes a computing device, such as computing device 800. In its most basic configuration, computing device 800 typically includes at least one processing unit 802 and memory 804. Depending on the exact configuration and type of computing device, memory 804 may be volatile (such as random access memory (RAM)), non-volatile (such as read-only memory (ROM), flash memory, etc.), or some combination of the two. This most basic configuration is illustrated in FIG. 8 by dashed line 806.

[0078] Computing device 800 may have additional features/functionality. For example, computing device 800 may include additional storage (removable and/or non-removable) including, but not limited to, magnetic or optical disks or tape. Such additional storage is illustrated in FIG. 8 by removable storage 808 and non-removable storage 810.

[0079] Computing device 800 typically includes a variety of computer readable media. Computer readable media can be any available media that can be accessed by the device 800 and includes both volatile and non-volatile media, removable and non-removable media.

[0080] Computer storage media include volatile and non-volatile, and removable and non-removable media implemented in any method or technology for storage of information such as computer readable instructions, data structures, program modules or other data. Memory 804, removable storage 808, and non-removable storage 810 are all examples of computer storage media. Computer storage media include, but are not limited to, RAM, ROM, electrically erasable program read-only memory (EEPROM), flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage,

magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store the desired information and which can be accessed by computing device 800. Any such computer storage media may be part of computing device 800.

[0081] Computing device 800 may contain communication connection(s) 812 that allow the device to communicate with other devices. Computing device 800 may also have input device(s) 814 such as a keyboard, mouse, pen, voice input device, touch input device, etc. Output device(s) 816 such as a display, speakers, printer, etc. may also be included. All these devices are well known in the art and need not be discussed at length here.

[0082] It should be understood that the various techniques described herein may be implemented in connection with hardware components or software components or, where appropriate, with a combination of both. Illustrative types of hardware components that can be used include Field-programmable Gate Arrays (FPGAs), Application-specific Integrated Circuits (ASICs), Application-specific Standard Products (ASSPs), System-on-a-chip systems (SOCs), Complex Programmable Logic Devices (CPLDs), etc. The methods and apparatus of the presently disclosed subject matter, or certain aspects or portions thereof, may take the form of program code (i.e., instructions) embodied in tangible media, such as floppy diskettes, CD-ROMs, hard drives, or any other machine-readable storage medium where, when the program code is loaded into and executed by a machine, such as a computer, the machine becomes an apparatus for practicing the presently disclosed subject matter.

[0083] Although exemplary implementations may refer to utilizing aspects of the presently disclosed subject matter in the context of one or more stand-alone computer systems, the subject matter is not so limited, but rather may be implemented in connection with any computing environment, such as a network or distributed computing environment. Still further, aspects of the presently disclosed subject matter may be implemented in or across a plurality of processing chips or devices, and storage may similarly be effected across a plurality of devices. Such devices might include personal computers, network servers, and handheld devices, for example.

[0084] Although the subject matter has been described in language specific to structural features and/or methodological acts, it is to be understood that the subject matter defined in the appended claims is not necessarily limited to the specific features or

acts described above. Rather, the specific features and acts described above are disclosed as example forms of implementing the claims.

WHAT IS CLAIMED IS:

1. A method for removing column band broadening in chromatography comprising:

removing a column from a chromatography device;

while the column is removed:

causing a plurality of injections of a test substance into the chromatography device, wherein each injection is associated with a different flow rate of a plurality of flow rates;

collecting non-column data for each flow rate of the plurality of flow rates;

placing the column in the chromatography device;

while the column is in the chromatography device:

causing an injection of a mixture of substances into the chromatography device at a selected flow rate from the plurality of flow rates;

generating a first chromatogram based on the injection of the mixture of substances at the selected flow rate; and

removing extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate a second chromatogram based on the first chromatogram with the removed extra column effects.

2. The method of claim 1, wherein removing the extra column effects from the first chromatogram using the non-column data collected for the selected flow rate comprises:

using regularized deconvolution of the column data collected for the selected flow rate, followed by artifact and further noise removal by a mathematical filter.

3. The method of claim 1 further comprising displaying the second chromatogram.

4. The method of claim 1, further comprising:

displaying the first chromatogram;

receiving an indication to remove the extra column effects from the first chromatogram; and

removing the extra column effects from the first chromatogram using the non-column data in response to the received indication.

5. The method of claim 1, wherein collecting non-column data for each flow rate of the plurality of flow rates comprises:

modeling peaks of the collected non-column data for each flow rate of the plurality of flow rates using a non-linear least squares regression to generate an asymmetric generalized normal; and

modeling parameters of the asymmetric generalized normal vs. flow rate reciprocal using a linear least squares regression and interpolating the collected non-column data with the modeled parameters to generate non-column data for any flow rate within a range of flow rates represented by the plurality of flow rates.

6. The method of claim 1, wherein removing extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate the second chromatogram, comprises:

generating a circulant matrix from the non-column data for the selected flow rate; formulating a deconvolution problem using the first chromatogram and the circulant matrix, and

solving the deconvolution problem using Tikhonov regularization to generate the second chromatogram.

7. The method of claim 6, further comprising:

denoising the second chromatogram using anisotropic diffusion.

8. A chromatography device for removing column band broadening comprising:

a column;

a detector; and at least one computing device, wherein the at least one computing device is adapted to:

receive non-column data for each flow rate of a plurality of flow rates;

cause an injection of a mixture of substances into the column at a selected flow rate from the plurality of flow rates;

generate a first chromatogram based on the injection of the mixture of substances at the selected flow rate; and

remove extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate a second chromatogram based on the first chromatogram with the removed extra column effects.

9. The chromatography device of claim 8, further displays the second chromatogram.

10. The chromatography device of claim 8, further comprising:

displaying the first chromatogram at the selected flow rate;

receiving an indication to remove the extra column effects from the first chromatogram; and

removing the extra column effects from the first chromatogram using the non-column data in response to the received indication.

11. The chromatography device of claim 8, wherein removing extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate the second chromatogram, comprises:

generating a circulant matrix from the non-column data for the selected flow rate;

formulating a deconvolution problem using the circulant matrix and the first chromatogram; and

solving the deconvolution problem using Tikhonov regularization to generate the second chromatogram, followed by artifact and noise removal using anisotropic diffusion.

12. The chromatography device of claim 8, further comprising:

displaying the first chromatogram;

receiving an indication to remove the extra column effects from the first chromatogram; and

removing the extra column effects from the first chromatogram using the non-column data in response to the received indication.

13. The chromatography device of claim 8, further comprising:

removing the column;

while the column is removed:

causing a plurality of injections of a test substance into the chromatography device, wherein each injection is associated with a different flow rate of the plurality of flow rates; and

collecting the non-column data for each flow rate of the plurality of flow rates.

14. A non-transitory computer-readable medium with computer executable instructions stored thereon than when executed by a chromatography device cause the chromatography device to:

receive non-column data for each flow rate of a plurality of flow rates;

cause an injection of a mixture of substances into a column of the chromatography device at a selected flow rate from the plurality of flow rates;

generate a first chromatogram based on the injection of the mixture of substances at the selected flow rate; and

remove extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate a second chromatogram based on the first chromatogram with the removed extra column effects.

15. The non-transitory computer-readable medium of claim 14, further displays the second chromatogram.

16. The non-transitory computer-readable medium of claim 14, further comprising:

displaying the first chromatogram at the selected flow rate;

receiving an indication to remove the extra column effects from the first chromatogram; and

removing the extra column effects from the first chromatogram using the non-column data in response to the received indication.

17. The non-transitory computer-readable medium of claim 14, wherein removing extra column effects from the first chromatogram using the non-column data collected for the selected flow rate to generate the second chromatogram, comprises:

generating a circulant matrix from the non-column data for the selected flow rate;

formulating a deconvolution problem using the circulant matrix and the first chromatogram; and

solving the deconvolution problem using Tikhonov regularization to generate the second chromatogram, followed by artifact and noise removal using anisotropic diffusion.

18. The chromatography device of claim 14, further comprising:

displaying the first chromatogram;

receiving an indication to remove the extra column effects from the first chromatogram; and

removing the extra column effects from the first chromatogram using the non-column data in response to the received indication.

19. The non-transitory computer-readable medium of claim 14, further comprising:

removing the column;

while the column is removed:

causing a plurality of injections of a test substance into the chromatography device, wherein each injection is associated with a different flow rate of the plurality of flow rates; and

collecting the non-column data for each flow rate of the plurality of flow rates.

20. The non-transitory computer-readable medium of claim 14, wherein the chromatography device is a gas or a liquid chromatography device.

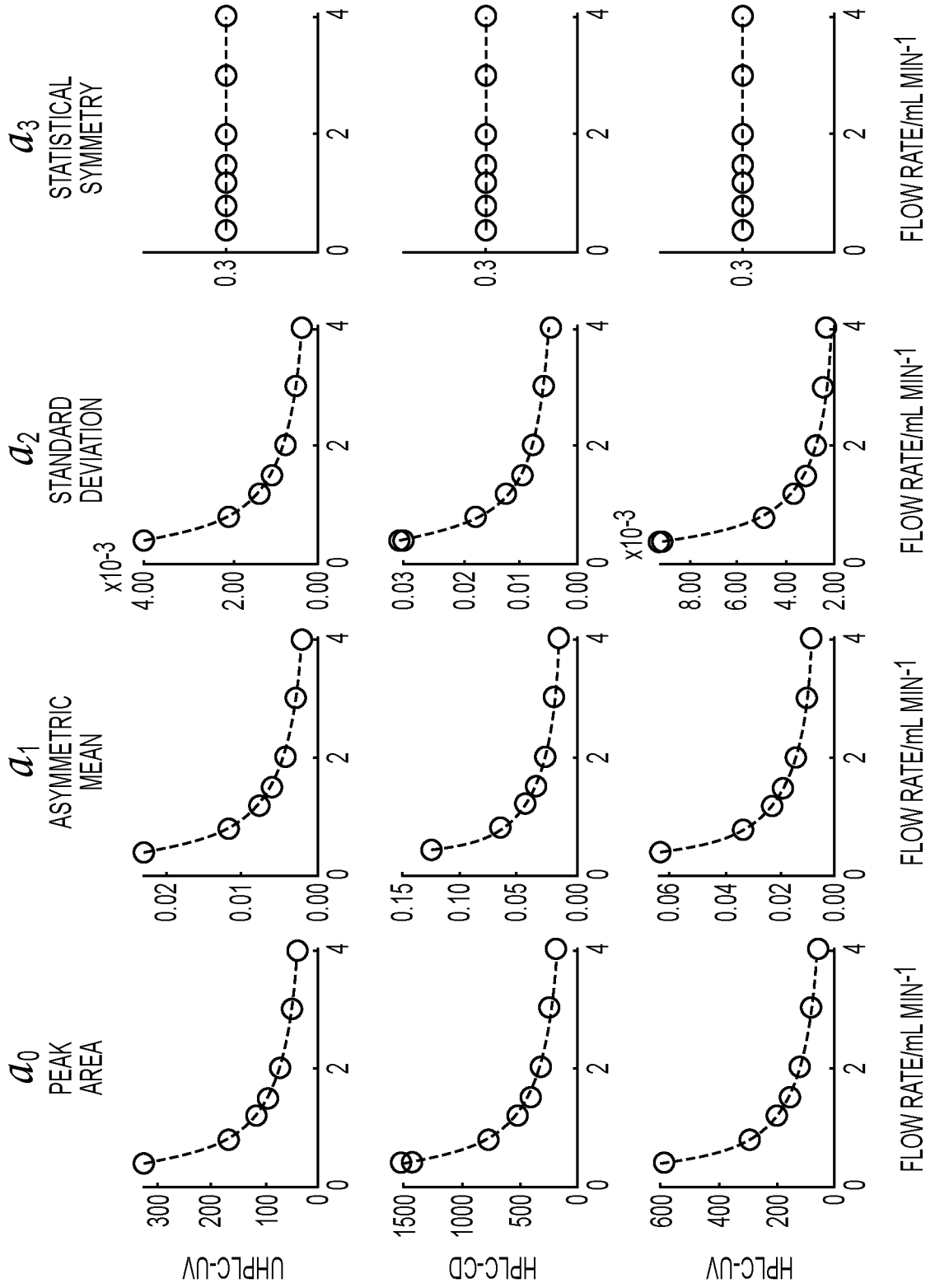
a_0
PEAK AREA
a_1
ASYMMETRIC MEAN
a_2
STANDARD DEVIATION
a_3
STATISTICAL SYMMETRY
UHPLC-UV
HPLC-CD
HPLC-UV
FLOW RATE/mL MIN-1
x10-3
100

FIG. 1

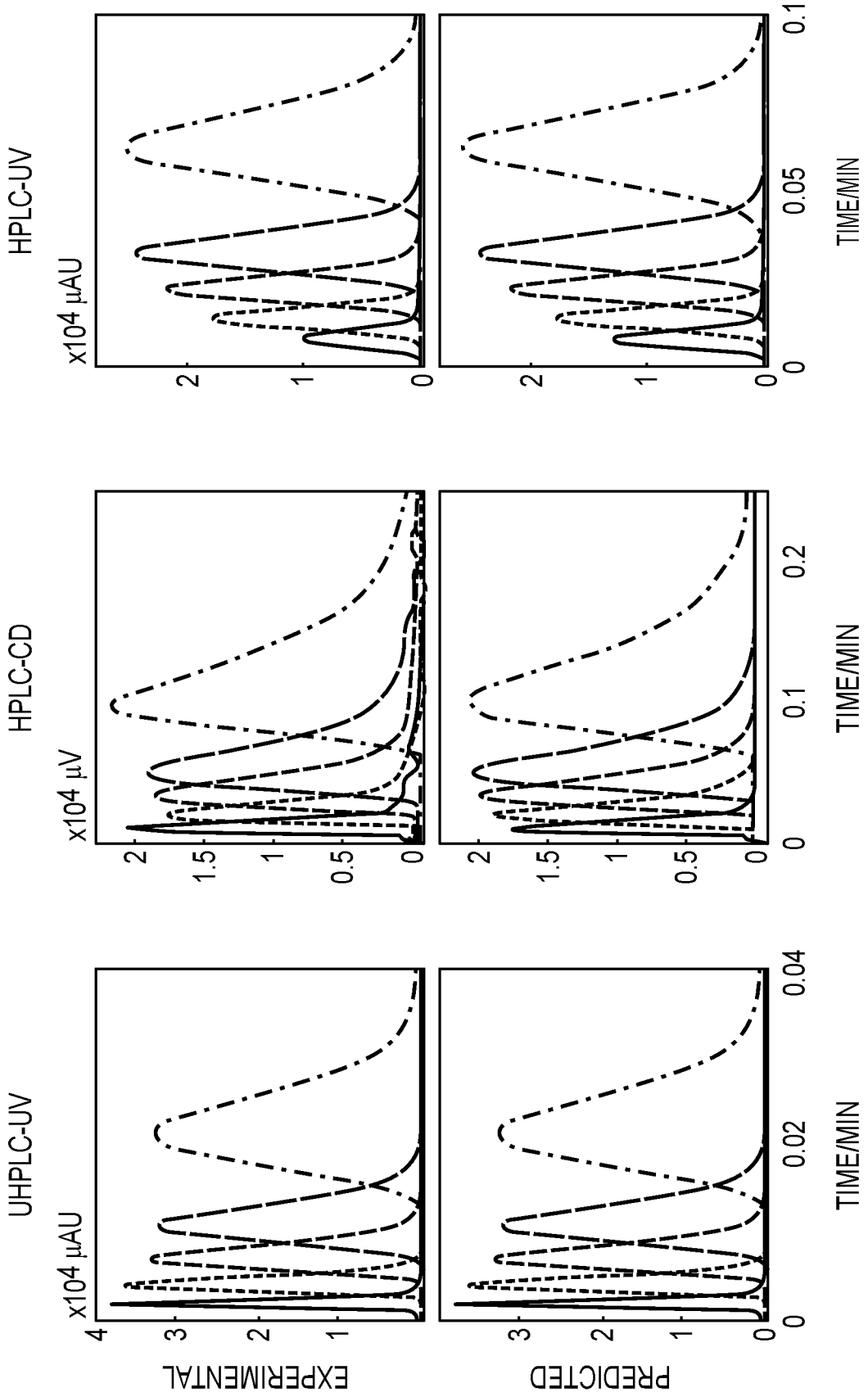
UHPLC-UV
HPLC-CD
HPLC-UV
x10^4 µAU
x10^4 µV
x10^4 µAU
EXPERIMENTAL
PREDICTED
TIME/MIN
TIME/MIN
TIME/MIN
200

FIG. 2

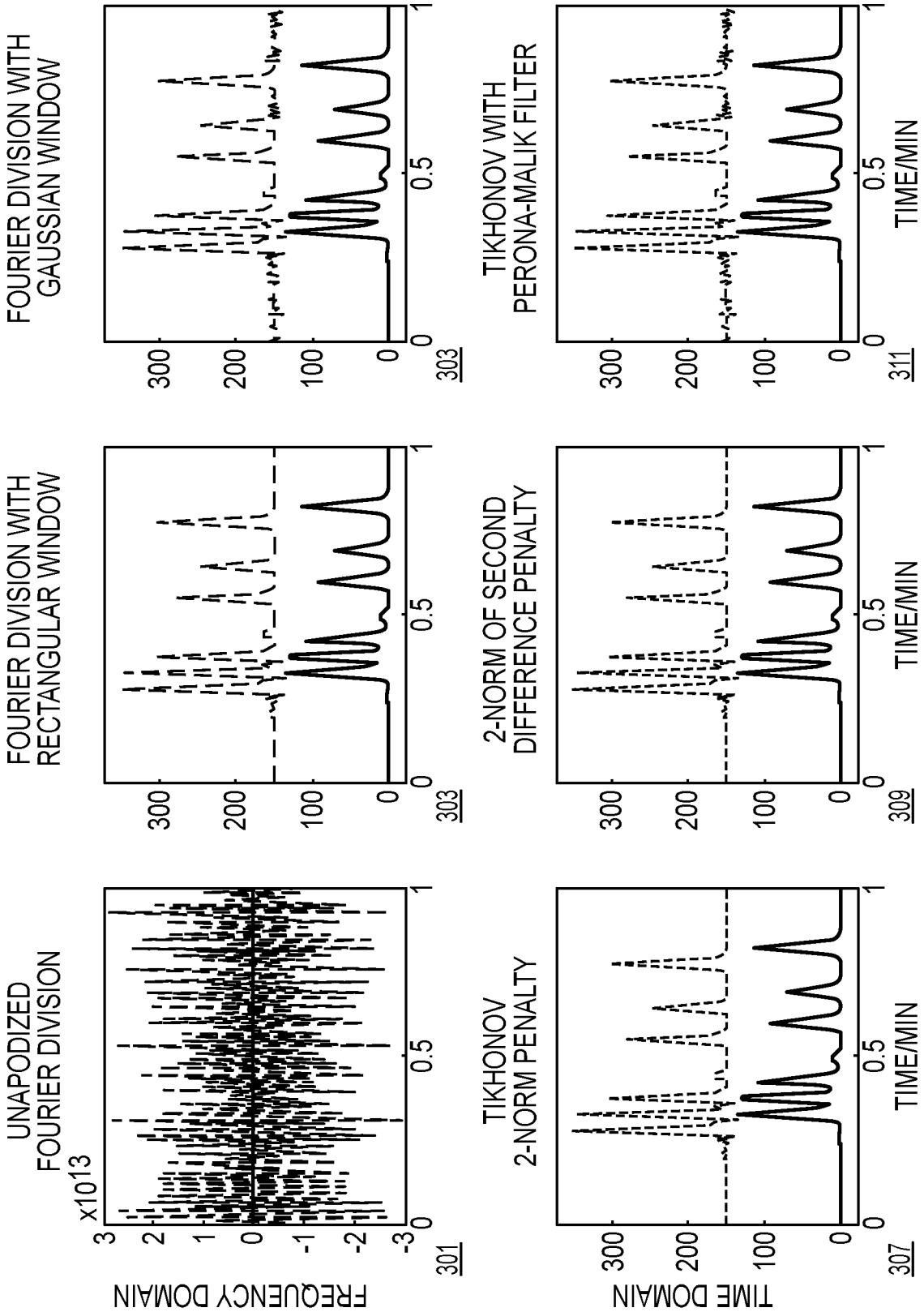
UNAPODIZED FOURIER DIVISION
FOURIER DIVISION WITH RECTANGULAR WINDOW
FOURIER DIVISION WITH GAUSSIAN WINDOW
TIKHONOV 2-NORM PENALTY
2-NORM OF SECOND DIFFERENCE PENALTY
TIKHONOV WITH PERONA-MALIK FILTER
FREQUENCY DOMAIN
TIME DOMAIN
TIME/MIN
x10^13
301
303
307
309
311

FIG. 3

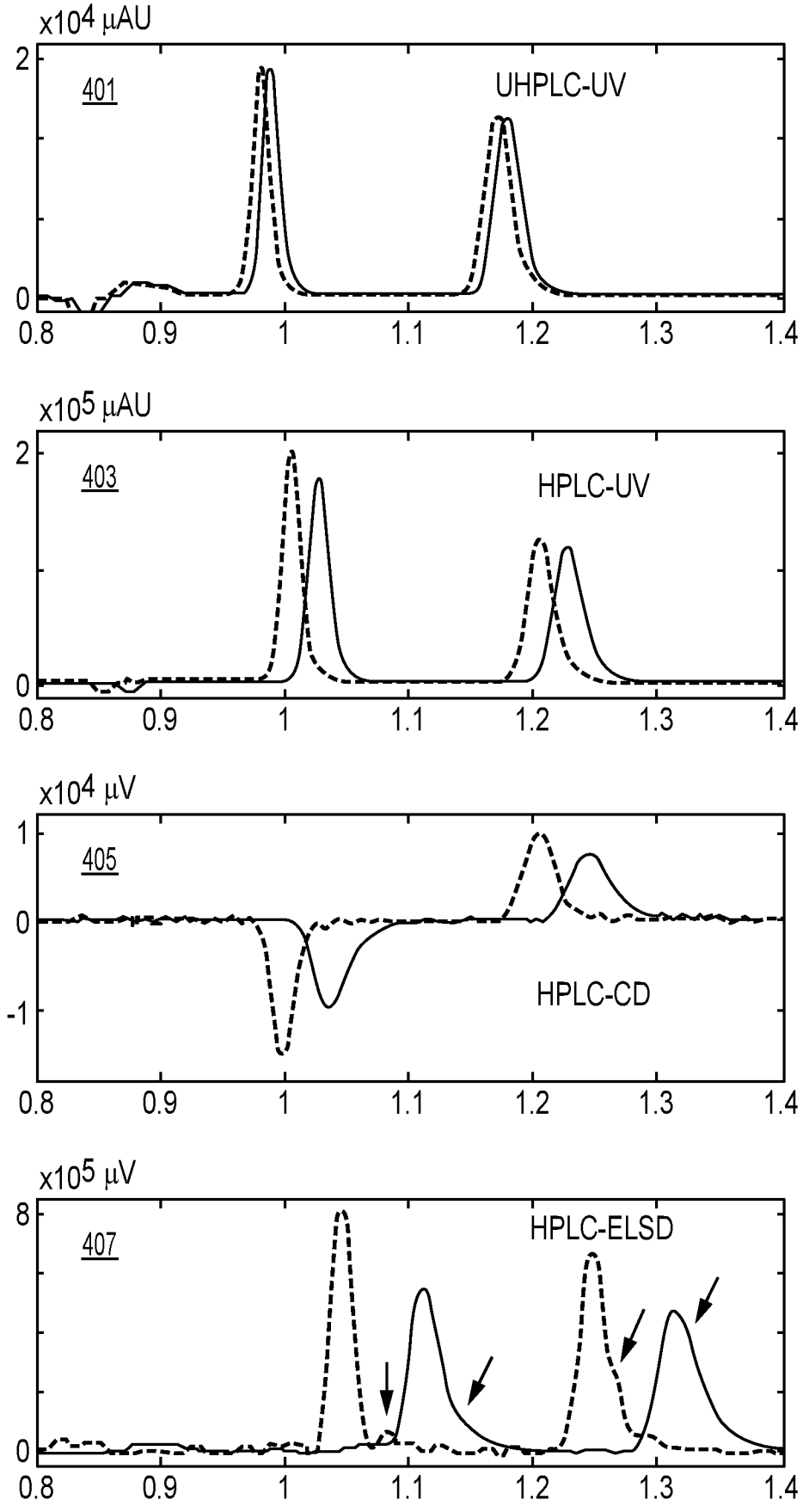
x10^4 µAU
2
0
401
UHPLC-UV
0.8
0.9
1
1.1
1.2
1.3
1.4
x10^5 µAU
2
0
403
HPLC-UV
0.8
0.9
1
1.1
1.2
1.3
1.4
x10^4 µV
1
0
-1
405
HPLC-CD
0.8
0.9
1
1.1
1.2
1.3
1.4
x10^5 µV
8
0
407
HPLC-ELSD
0.8
0.9
1
1.1
1.2
1.3
1.4
TIME/MIN

FIG. 4

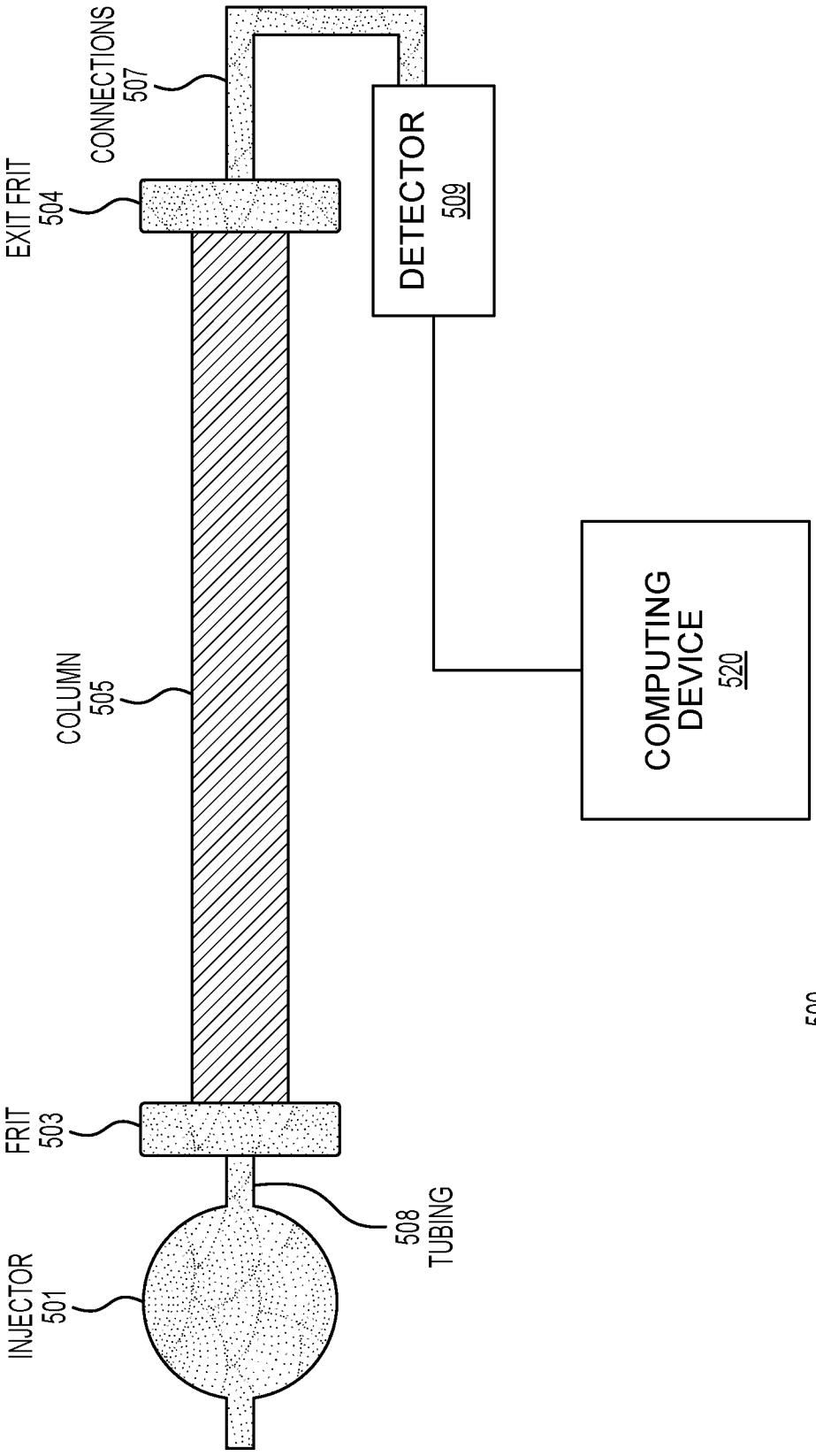
CONNECTIONS
507
EXIT FRIT
504
DETECTOR
509
COLUMN
505
COMPUTING
DEVICE
520
FRIT
503
508
TUBING
INJECTOR
501
500

FIG. 5

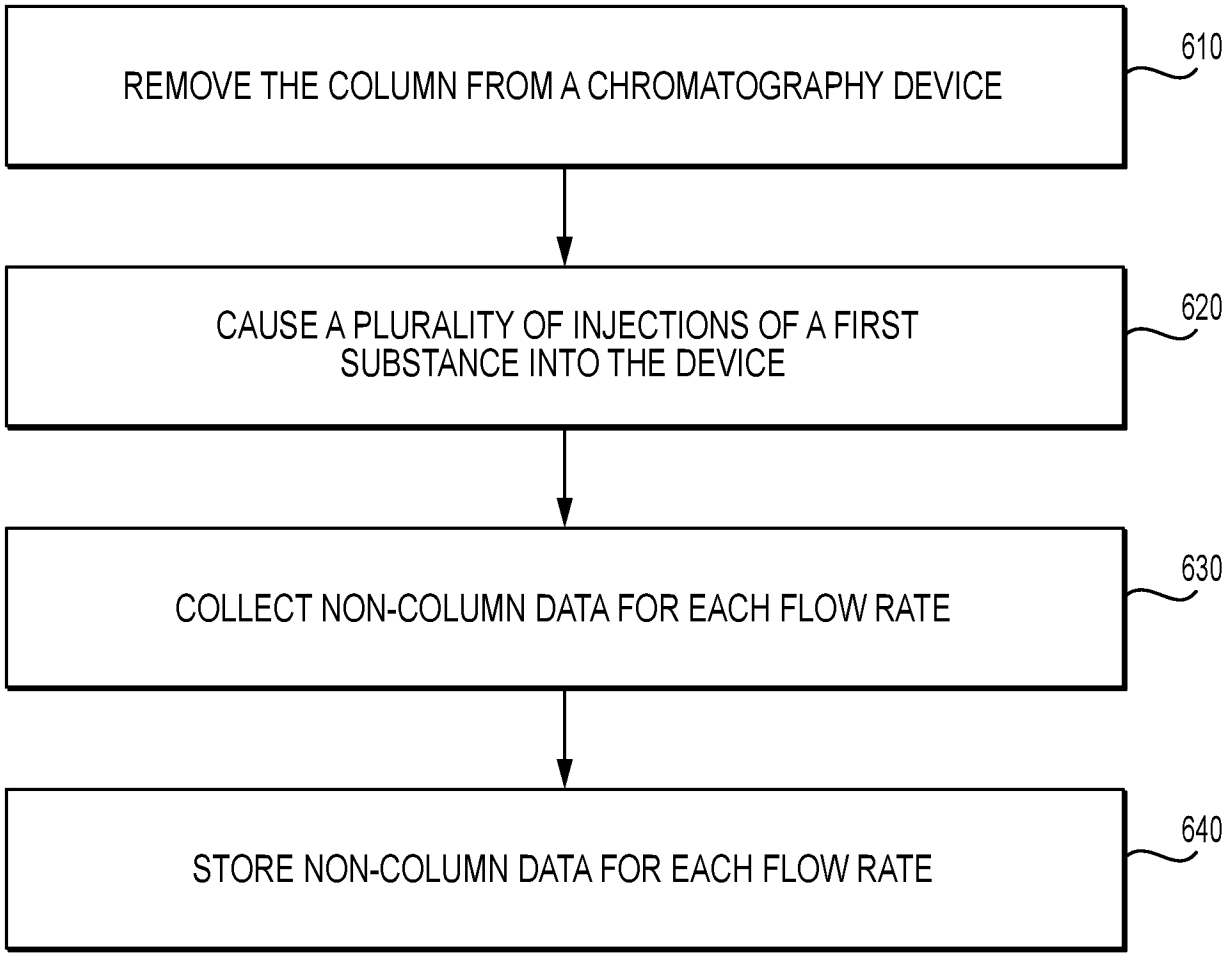
REMOVE THE COLUMN FROM A CHROMATOGRAPHY DEVICE
610
CAUSE A PLURALITY OF INJECTIONS OF A FIRST SUBSTANCE INTO THE DEVICE
620
COLLECT NON-COLUMN DATA FOR EACH FLOW RATE
630
STORE NON-COLUMN DATA FOR EACH FLOW RATE
640
600

FIG. 6

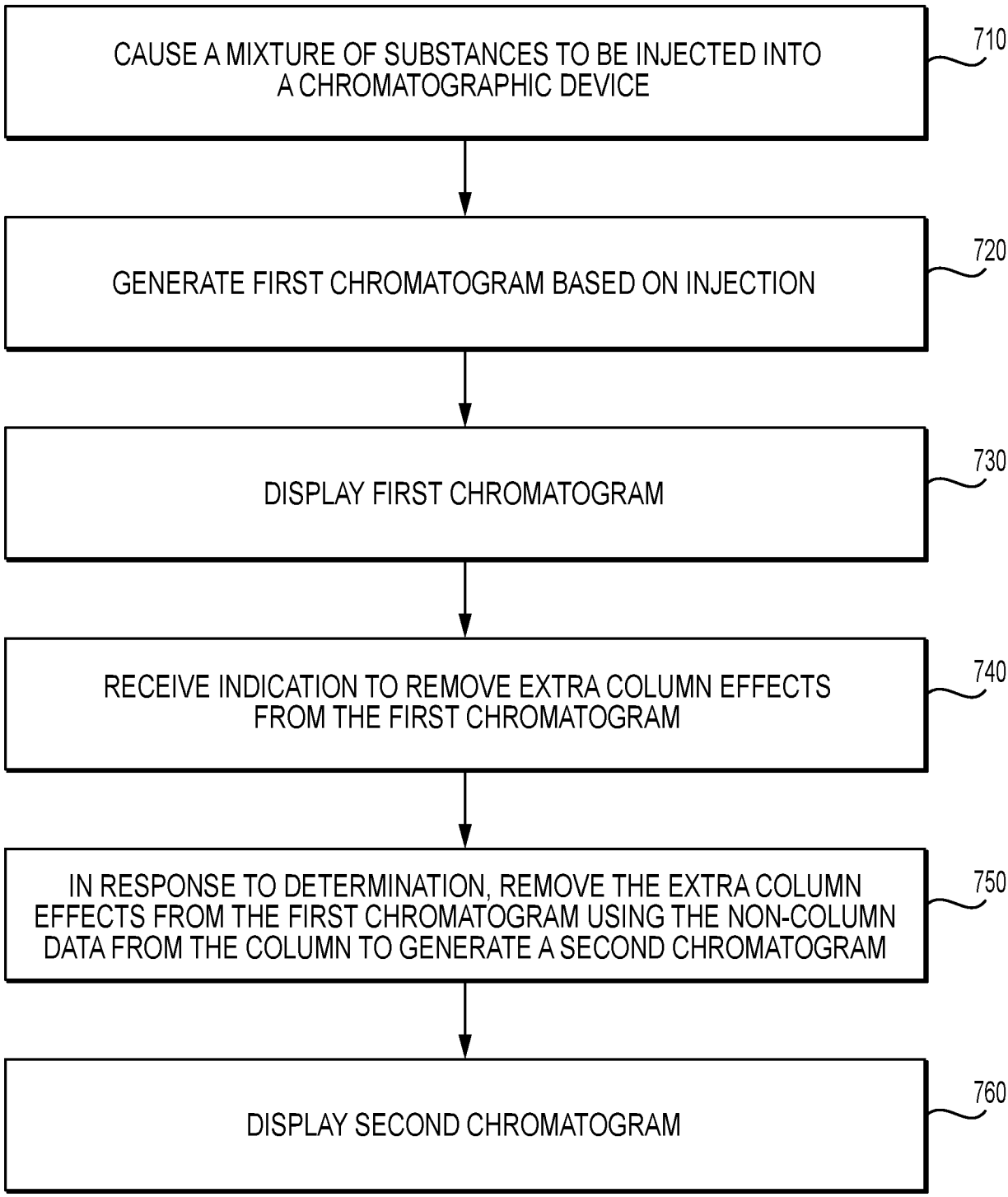
CAUSE A MIXTURE OF SUBSTANCES TO BE INJECTED INTO A CHROMATOGRAPHIC DEVICE
710
GENERATE FIRST CHROMATOGRAM BASED ON INJECTION
720
DISPLAY FIRST CHROMATOGRAM
730
RECEIVE INDICATION TO REMOVE EXTRA COLUMN EFFECTS FROM THE FIRST CHROMATOGRAM
740
IN RESPONSE TO DETERMINATION, REMOVE THE EXTRA COLUMN EFFECTS FROM THE FIRST CHROMATOGRAM USING THE NON-COLUMN DATA FROM THE COLUMN TO GENERATE A SECOND CHROMATOGRAM
750
DISPLAY SECOND CHROMATOGRAM
760
700

FIG. 7

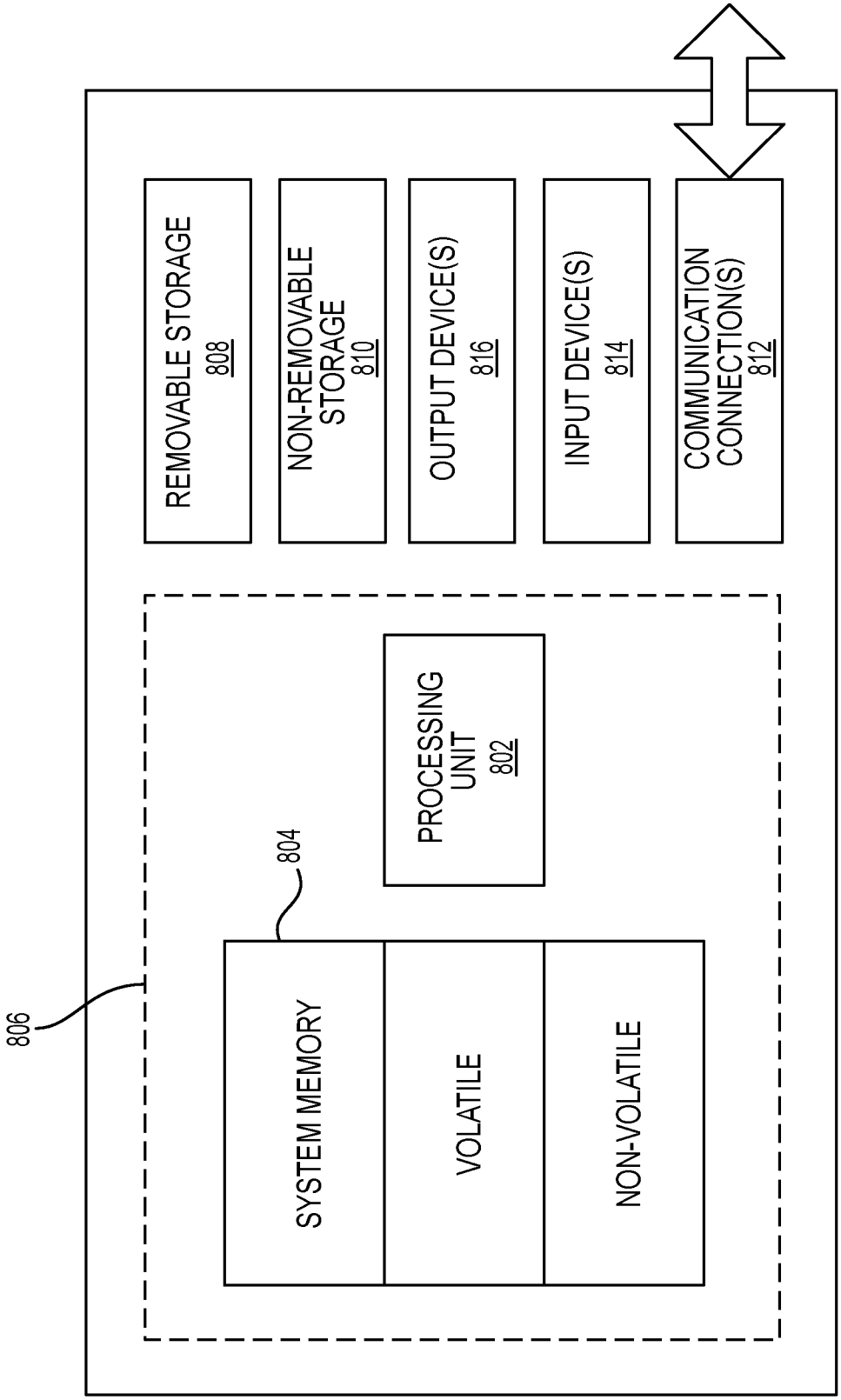
REMOVABLE STORAGE
808
NON-REMOVABLE STORAGE
810
OUTPUT DEVICE(S)
816
INPUT DEVICE(S)
814
COMMUNICATION CONNECTION(S)
812
PROCESSING UNIT
802
804
SYSTEM MEMORY
VOLATILE
NON-VOLATILE
806
800

FIG. 8