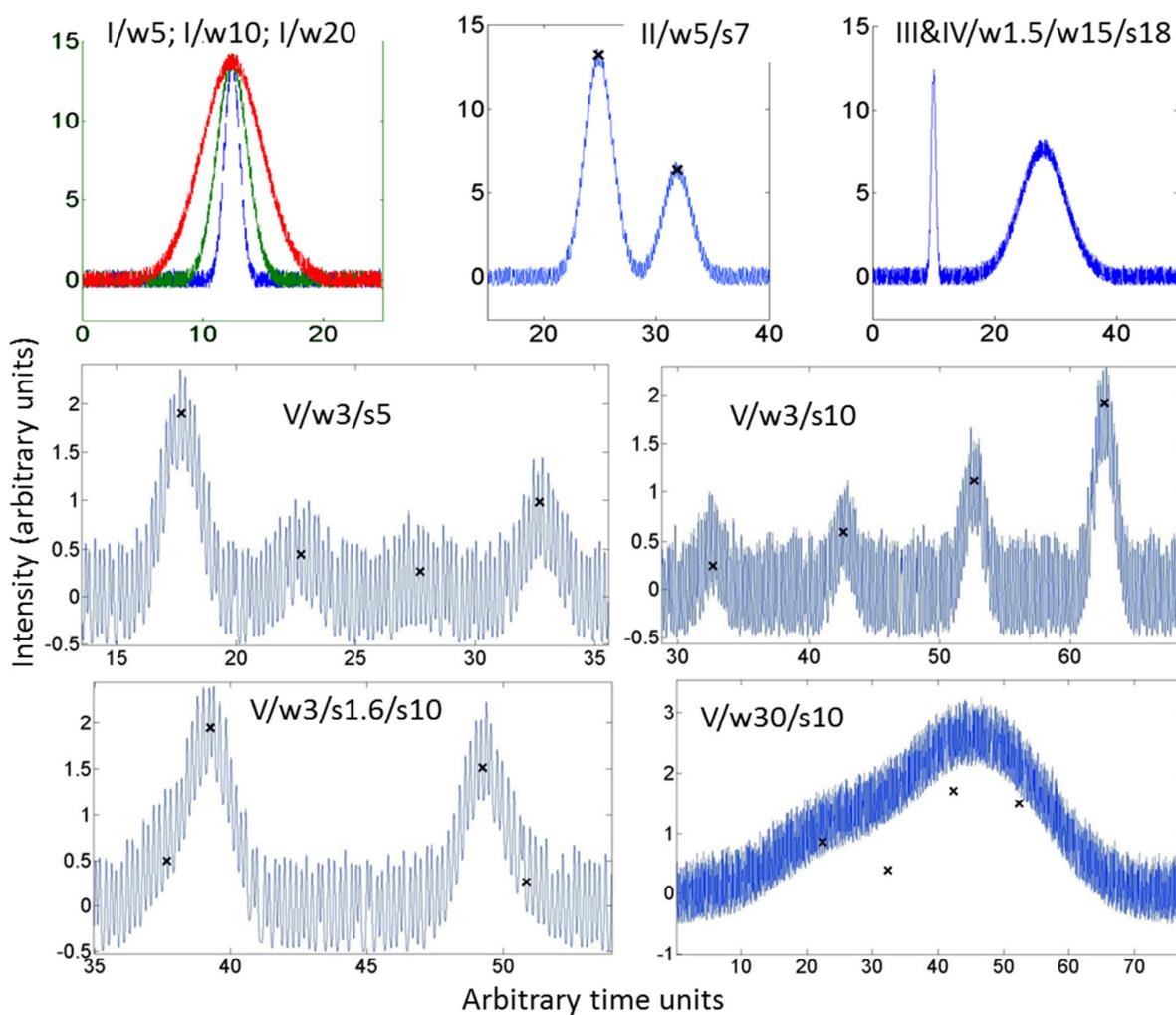548

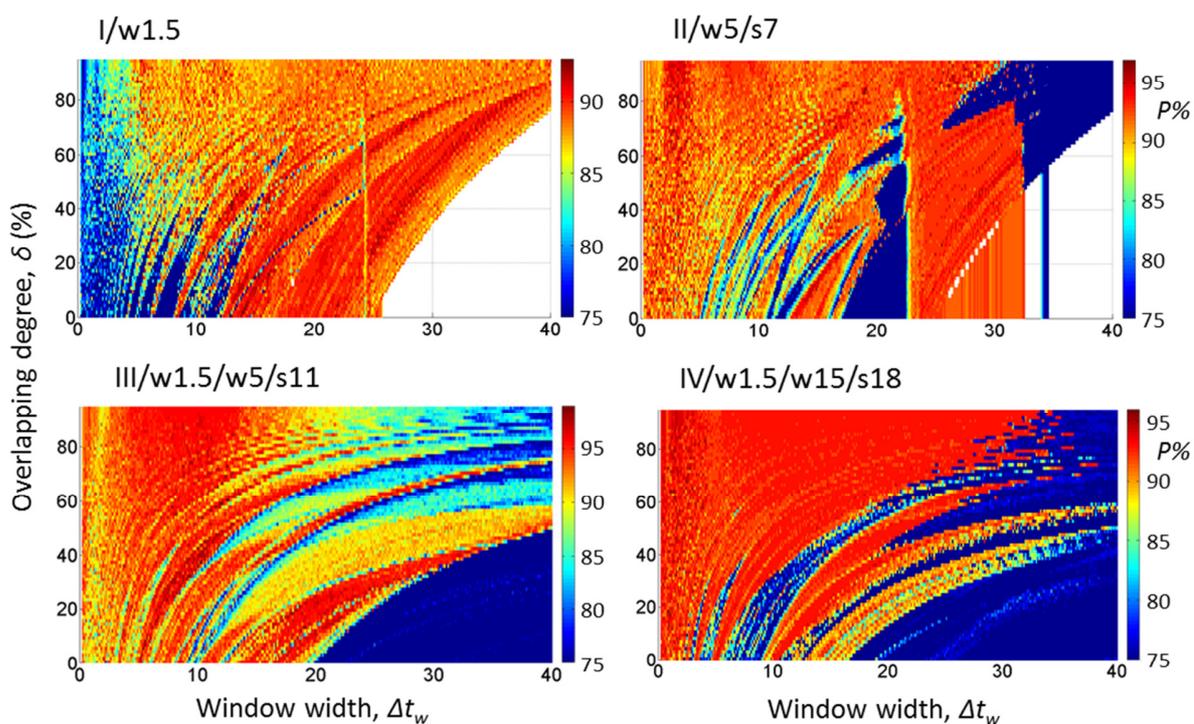**Electronic Supplementary Information (3 figures and 3 tables)**

549

550

551

552

553 **Fig. S1.** Simulated chromatograms according to Tables S1-S3; a chromatogram of one of

554 the classes is shown; the crosses indicate the location and height of the significant peaks for

555 the average chromatogram of the class.

556

557

558

559

Fig. S2. Overview of the response surfaces of four simulated experiments of the I-IV series
(see Tables S1-S3). Percentage of correctly classified chromatograms by leave-one-out
($P\%$) as a function of the window's width, Δt_w, and degree of overlapping of consecutive
windows, δ. At the right of the response surfaces, a scale indicating the color code for $P\%$
is given.

565

566
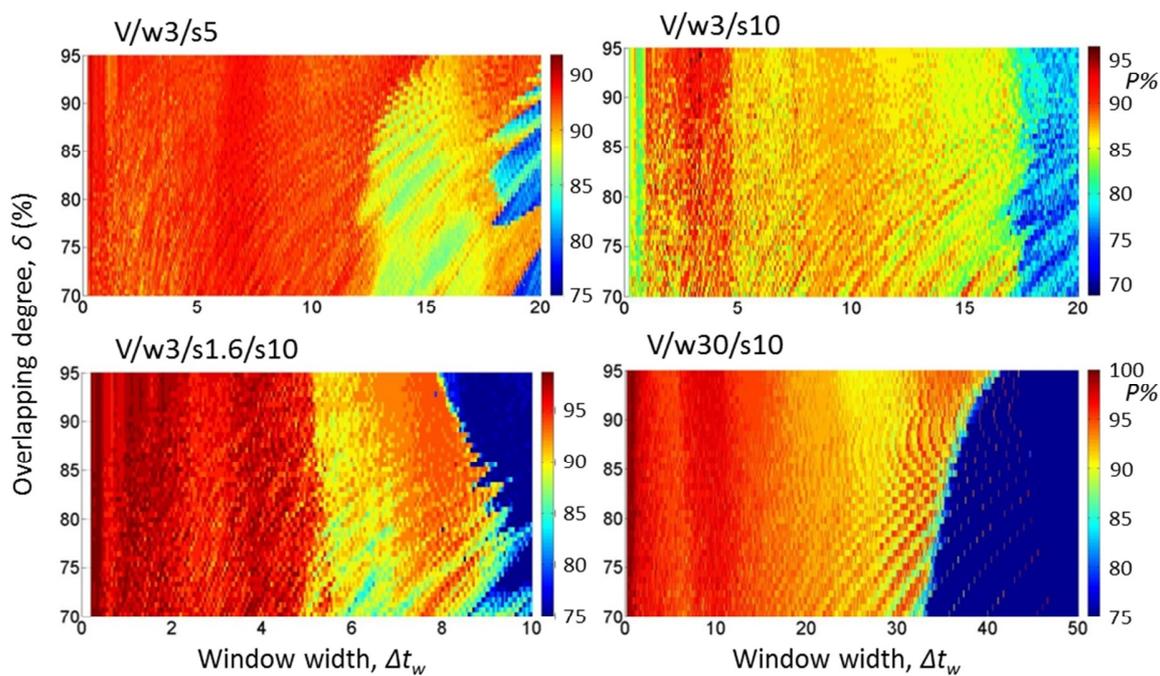
567

568

569 **Fig. S3.** Overview of the response surfaces of the four simulated experiments of the V

570 series (see Tables S1-S3). Other details as in Fig. S2.

571

572
573 **Table S1**
574 Parameters used in the generation of the baselines for series I to V [a]
575

Parameter	Symbol	I-IV	V
Time range	t_0; t_{max}	0; 50	0; 100
Number of baseline peaks (noise)	N_{np}	250	500
Number of points per peak within $\pm 3s$ (restricted to ≥ 20)	-	20	20
Average height of the baseline peaks	h_{np}	1	1
Width of baseline peaks (as 4 times the peak st. deviation)	$4s_{np}$	0.15	0.15
Standard deviation of peak height (to add random values to h_{np})	s_{nh}	0.2	0.2
Standard deviation of peak location (random lateral shifts)	s_{nt}	0.05	0.05

576
577 [a] As detailed in Tables S2 and S3, the series were distinguished by the number, width and relative locations of the significant peaks. A baseline
578 was a sum of N_{np} Gaussian functions, each function giving rise to a baseline peak within the t_0 to t_{max} range. Each baseline peak was defined
579 by 20 points which were evenly distributed between the $\pm 3s$ region around the peak maximum. The average height of the baseline peaks was
580 h_{np}, but the actual peak heights varied randomly with a standard deviation of s_{nh} around the average value. Evenly spaced initially, the baseline
581 peaks were laterally shifted at random with a standard deviation of s_{nt}. The origin of the vertical axis was located at the middle of the baseline
582 band by subtracting the average height of all the baseline points from all the points.
583

584

585 **Table S2**

586 Features of the chromatograms for series I to V and percentage of correctly classified chromatograms by the omniscient method, $P\%_{max}$

587

Series	Features of the chromatograms of the series [a]	Experiment [b]	Peak widths, $4s_{sp}$	Peak locations	$P\%_{max}$ [c]
I	A single peak; the peak width was varied	I/w1.5	1.5	20	89 (11)
		I/w5	5	20	89 (11)
		I/w20	20	20	89 (11)
II	Two peaks of the same width; the spacing between them was varied; the peak heights were different for the four classes (Table S3)	II/w5/s7	5	25; 32	94 (6)
		II/w5/s23	5	10; 33	94 (6)
III	Two peaks of different width; the spacing between them was varied; the peak heights were different for the four classes (Table S3)	III/w1.5/w5/s11	1.5; 5	10; 21	96 (4)
		III/w1.5/w15/s18	1.5; 15	10; 28	96 (4)
IV	Two peaks of different width; the height of the 1st peak was the same for classes B, C, D; the height of the 2nd peak was the same for classes A, B (Table S3)	IV/w1.5/w15/s18	1.5; 15	10; 28	93 (7)
V	Four peaks of the same width; both the spacing between them and the peak width were varied; the peak heights were different for all the classes (Table S3)	V/w3/s5	3	18; 23; 28; 33	91.2 (7)
		V/w3/s10	3	32; 42; 52; 62	96.2 (3)
		V/w3/s1.6/s10	3	38; 39.6; 48; 49.6	97.5 (2)
		V/w30/s10	30	22; 32; 42; 52	98.7 (1)

588

589 [a] The positive significant peaks added to the baselines were Gaussian functions limited to the $\pm 3s_{sp}$ range around their respective peak maxima. A Symlet10
590 wavelet was used to remove the tiny discontinuities which were produced at the ends of the $\pm 3s$ region around the maximum of both the baseline and
591 significant peaks. The average peak heights (h_{sp}) for the four A-D classes are indicated in Table S3.
592 [b] For each experiment, an acronym indicating the series and the peak width (w) and spacing (s) is used.
593 [c] Percentages over 4 classes x 25 chromatograms per class, except for series V where 4 classes x 20 chromatograms per class were simulated (the number of
594 critical cases is given between parentheses).

595

596

597 **Table S3**

598 Average height of the significant peaks, h_{sp}, for series I to V [a]

599

Class	I	II	III	IV	V/w3/s5	V/w3/s10	V/w3/s1.6/s10	V/w30/s10
A	7	7; 13	7; 13	13; 10	1.90; 0.45; 0.26;0.98	0.24; 0.59; 1.12; 1.92	0.50; 1.94; 1.52; 0.26	0.86; 0.40; 1.71; 1.50
B	9	9; 11	9 11	10; 10	1.70; 0.45; 0.26; 1.04	0.23; 0.77; 0.91; 1.87	0.36; 1.89; 1.76; 0.43	0.50; 0.56; 2.03; 1.15
C	11	11; 9	11; 9	10; 7	1.61; 0.59; 0.35; 1.16	0.00; 0.80; 1.37; 2.02	0.72; 1.95; 1.20; 0.63	0.79; 0.28; 1.52; 1.75
D	13	13; 7	13 7	10; 13	2.04; 0.24; 0.51; 1.16	0.05; 0.71; 0.84; 1.79	0.16; 2.25; 1.73; 0.32	0.76; 0.30; 1.94; 1.28

600

601 [a] Within each class, the peak heights of the chromatograms of series I to IV were made to vary at random with a standard deviation of $s_{ph} = 2$.
602 The peak heights of the chromatograms of series V were made to vary at random within a range equal to the 15% - 20% of the average peak
603 height.

604