Full Paper

Least Absolute Shrinkage and Selection Operator as a Multivariate Calibration Tool for Simultaneous Determination of Diphenylamine and its Nitro Derivatives in Propellants

Ahmad Mani-Varnosfaderani,*[a] Masoud Soleimani,[a] and Naader Alizadeh[a]

Abstract: Determination of diphenylamine (DPA) and its nitro derivatives received great attention for storing, deposition and on-time usage of propellants. Herein, we present a novel and simple method for simultaneous determination of DPA and its nitro derivatives in solid propellants using UV-Vis spectroscopy and chemometrics techniques. The UV-Vis spectra of the analytes revealed strong overlap and it was difficult to determine them individually in their mixture without any separation and purification. To tackle the overlapping problem in collected spectra, analysis of first-order UV-Vis data was performed using multivariate calibration techniques. In this way, principle component regression (PCR), different modes of partial least square (PLS) and least absolute shrinkage and selection operator (LASSO) have been used for correlating the collected spectra to the concentration of DPAs in synthetic and real samples. The important variables were selected by confining the L_1-norm of the regression coefficients in multivariate model via the shrinkage and selection operator in LASSO approach. The results obtained by LASSO regression technique in this work were superior to those obtained by different modes of PLS algorithm. Moreover, it is shown that LASSO can be used as a reliable variable selection and modeling technique in multivariate calibration studies. Generally, the proposed strategy in this work is simple, non-destructive, low-cost and rapid and can be effectively applied for simultaneous determination of DPA and its nitro derivatives in solid propellants.

Keywords: Diphenylamine · multivariate calibration · least absolute shrinkage and selection operator · partial least square · UV-Vis

1 Introduction

Solid propellants are naturally degraded during their storage. This degradation process is related to sequential autocatalytic reactions which lead to formation of some active radicals. The activities of the produced radicals lead to denitration of the nitrocellulose which is typically the main component of the solid propellants. The recombination of these radicals produce NO_2 and HNO_3, being able to induce an autocatalytical phenomenon that finally lead to the spontaneous propellant deflagration. Generally, the storage of explosive materials such as propellants is of great importance and essential issue in inventory engineering. In order to have safe storage of the propellants containing DPA as a stabilizer, it is necessary to determine DPA content, precisely. Several factors such as oxidation by UV radiation, humidity, temperature changes and mechanical factors accelerate the radical formation and induce the degradation process.

In order to prevent this process, usually DPA and its urea-substituted derivatives such as akardite and centralite are predominantly used as stabilizers. Explosive materials are very sensitive to temperature and UV radiation. High temperature and high intensity of UV light if do not induce an explosion at least will autocatalytically release the nitrogen content of explosive materials. Without a stabilizer, the free radicals of the released nitrogen oxides will react with the nitrate esters not yet decomposed, leading to a catalytic effect on the decomposition rate. The decomposition induces a temperature increase, which may lead to auto-ignition, as well as a loss of energy, resulting in a decrease of muzzle velocity.

The stabilization of the propellant with DPA proceeds through a group of reactions that lead to the formation of N-nitroso-Diphenylamine (n-n-DPA), 2-nitro diphenylamine (2-nDPA) and 4-nitro diphenylamine (4-nDPA). Further degradation of DPA can lead to the formation of 2,4-dinitro Diphenylamine (2,4-dnDPA) [1]. This is schematically illustrated in Figure S1 in supporting material section. Further

[a] A. Mani-Varnosfaderani, M. Soleimani, N. Alizadeh
Department of Chemistry, Tarbiat Modares University, P.O. Box 14115–175, Tehran, Iran
Tel: +98-21-82884738
Fax: +98-21-82883755
*e-mail: a.mani@modares.ac.ir

degradation can even result to tri-, tetra- or hexa-nitro derivatives. However, with a normal degradation process, di- and tri-nitrate derivatives are rarely detected. Since, the concentrations of di- and tri-nitro derivatives of DPA are not too much in the early stages of the decomposition process, we focused on determination of the concentration of DPA and its mono-nitro derivatives. Because the concentration of 2,4-dnDPA is an indicator of the high degradation of the DPA in propellant, determination of this compound would be of great interest. Therefore, we finally limited our analysis to determination of DPA, N-n-DPA, 2-nDPA, 4-nDPA and 2,4-dnDPA.

In addition to military utilization, DPAs were and are still used in numerous industrial applications. Because of the wide variety of applications of DPA and its nitro derivatives, a large number of researches has been performed for determination and quantification of these compounds, especially DPA. Chromatographic methods such as liquid chromatography with UV-Vis absorption with atmospheric pressure chemical ionization (APCI) mass selective detection [2], fluorescence detection [3], electrochemical detection [4], gas chromatography [5] and supercritical liquid chromatography [6] have been reported for determination of DPA in different sample matrices. Simultaneous determination of DPA and N-n-DPA in solid propellants and in well water using fluorescence technique has been reported by Alizadeh *et al* [7]. The prediction of the age of the gunpowder has been performed using HPLC through direct determination of the DPA and its derivatives [8]. Moreover, determination of DPA and its derivatives have been reported in literature using tandem mass spectrometry [9], ion mobility spectroscopy (IMS) [10] and near-IR [11]. However, these methods require expensive instruments, advanced analytical skills and a considerable amount of time.

Instead of using methods that require tedious steps and considerable time, using rapid, simple and low cost methods such as UV-Vis spectroscopy can be very useful. When one uses first order data (e.g., UV-Vis spectra) the appearance of overlapping spectra and matrix effects may reduce the accuracy and precision of the analysis. Multivariate calibration methods have been proposed to overcome fundamental mathematical challenges occurred during analysis of these types of data [12]. Among different multivariate calibration techniques, projection-based methods such as principal component regression (PCR) and partial least squares (PLS) have attracted attention in recent years and these methods have been used for determination of different analytes in various types of matrices [13–22]. PLS has become the most popular multivariate calibration method because of the quality of its calibration models and the ease of its implementation [23]. Rather than projection based methods, more robust calibration techniques such as artificial neural networks (ANN) and support vector machines (SVM) have also been used by researchers for modeling the data with high degrees of overlap and interactions.

In the present contribution, different multivariate calibration methods including PCR, normal PLS, recursive weighted partial least square (rPLS) [24] and least absolute shrinkage and selection operator (LASSO) [25] have been used for simultaneous determination of DPA and its nitro-derivatives in synthetic and real samples. The results in this work revealed that with accurate determination of calibration wavelengths, it is possible to determine different derivatives of DPAs in synthetic and real samples, precisely. Moreover, LASSO algorithm has been suggested as a reliable and robust method for selection of important signal channels when dealing with highly overlapped UV-Vis spectra.

2 Material and Methods

2.1 Least Absolute Shrinkage and Selection Operator

Least absolute shrinkage and selection operator (LASSO) is a regularized regression technique and is based on L_1-minimizations of the regression coefficients [25]. This method controls the variance in the estimates of the regression coefficients through shrinkage of the coefficients. The LASSO method estimates the value of the regression coefficients (β_j) by minimizing the following objective function (i.e. LASSO function):

$$\text{LASSO} = \sum_{i=1}^{n} \left(y_i - \alpha - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \qquad (1)$$

where *y* and *x* are the dependent and independent variables in the regression problem, respectively. n and p represent number of observations and independent variables, respectively. Moreover, β_j is the regression coefficient, α is intercept and λ is a penalty term which controls the value of the shrinkage. Shrinkage penalty controls the magnitude of regression coefficients β_j according to equation (1). The λ parameter controls the impact of least squares and the penalty on the regression coefficient estimates. More generally, λ decides on the level of continuous shrinkage applied to the estimates. If λ is large, β_j coefficients are penalized highly towards zero (all absolute coefficients are penalized). Low value of λ imposes little penalty on β_j coefficients (i.e. least square criterion is applied). LASSO regression has been proven to have potential applications in the realm of chemistry. As an example, multicollinearity and feature selection issues have been investigated and LASSO was proposed as a reliable method to solve the multicollinearity and subset selection problems [26, 27]. In the present contribution, the LASSO regression technique has been applied for simultaneous determination of DPA in synthetic and real samples using the UV-Vis data. Implementation of penalty term on β_j coefficients in LASSO technique, led to automatic selection

**Propellants,
Explosives,
Pyrotechnics**

LASSO as a multivariate calibration tool for simultaneous determination of DPA and its nitro derivatives in propellants

of the predictors and therefore selective wavelengths have been chosen for final regression analysis.

For comparison purposes, the performances of LASSO regression technique have been compared with those of PCR, PLS and rPLS, in this work. More detailed theory behind LASSO and rPLS algorithms can be found in literature [24, 25].

2.2 Chemicals and Solvents

Pure solid standards of diphenylamine (DPA), 2-nitro-diphenylamine (2nDPA), 4-nitroDPA (4nDPA), N-Nitroso diphenylamine (N-n-DPA) and 2,4-dinitro diphenylamine (2,4-dnDPA) for the preparation of stock solutions were purchased from Sigma-Aldrich company. Acetonitrile and methanol were obtained from Merck. The physico-chemical properties of DPA and its nitro derivatives are given in Table 1.

Table 1. DPA and its nitro derivatives as target analytes and their physico-chemical properties.

Boiling point (°C)	Melting point (°C)	Molecular mass (g mol^{-1})	Purity (%)	Name
302	53	169.22	95 <	DPA
347	65	198.22	95 <	N-n-DPA
383	133	214.22	95 <	4-nDPA
346	74	214.22	95 <	2-nDPA
413	160	259.22	95 <	2,4-dnDPA

Stock solutions of each analyte with concentration of 5 μM in methanol were prepared and kept away from heat and light and then placed in a refrigerator at 4 °C. Forty-seven synthetic mixtures comprising of calibration (37 samples) and validation sets (10 samples) were prepared, in the concentration ranges of 1.0–40.0, 0.0–30, 0.0–25.0, 0.0–10.0, 0.0–10.0 for DPA, N-n-DPA, 2-nDPA, 4-nDPA and 2,4-dnDPA, respectively. To reduce the effects of uncontrolled factors, a central composite design (CCD) approach was used to build the calibration set. The CCD was used to obtain an orthogonal calibration design ($\alpha = 1.6$). The concentration of the validation set was adopted randomly. Standard solutions were obtained by diluting required volumes of stock solutions of analytes. The concentrations of the samples of the calibration and validation sets are given in Table 2. Two different propellants were analyzed as real samples after development of the multivariate models. These samples were named as P230 and P300. Both real samples were originally sampled 5 years ago from a military depository and wrapped in Aluminum foil and stored in dark in a refrigerator at 5 °C. In order to prepare each propellant sample, 30 mg of grinded propellants was placed in a filter paper, which was folded in the form of a bag, and then samples were extracted four times with 10 mL of methanol. Extraction time

was 10 min and the mixture was agitated in the ultrasonic bath. The prepared samples were kept in the dark before analysis. Standard addition method was used to eliminate the matrix effect for determination of the analytes in real samples. Certain amount of standard solution was added to each of the extracted solutions from propellants and diluted 10 times and then UV-Vis spectrum was recorded.

2.3 Apparatus

Absorbance measurements were performed on a diode array spectrophotometer (Scinco 2100) with the use of a 1.0 cm quartz cell. Spectra were acquired in the wavelength range from 210 to 605 nm with a resolution of 2 nm. All the data were recorded at room temperature. The Sartorius digital scale was used for weighing the analytes and samples with the precision of 0.01 mg. In total, spectral data has been recorded at 179 wavelengths. The Ultrasonic bath (LC30H-Elma) was used during the extraction procedure.

2.4 Data Analysis

The calculations in this work were implemented on a desktop computer with 'Windows 7 Pro' as the operating system, Intel(R) Core(TM) core i7 CPU and 16 GB of RAM. The standard functions of MATLAB were used for running PLS algorithm. Moreover, the *"LASSO"* function of MATLAB software was used for development of LASSO regression models and optimization of *"λ"* parameter. Design-Expert 7.1.6 Trial (Stat-Ease Inc., Minneapolis, USA) was used for designing the calibration and validation sets.

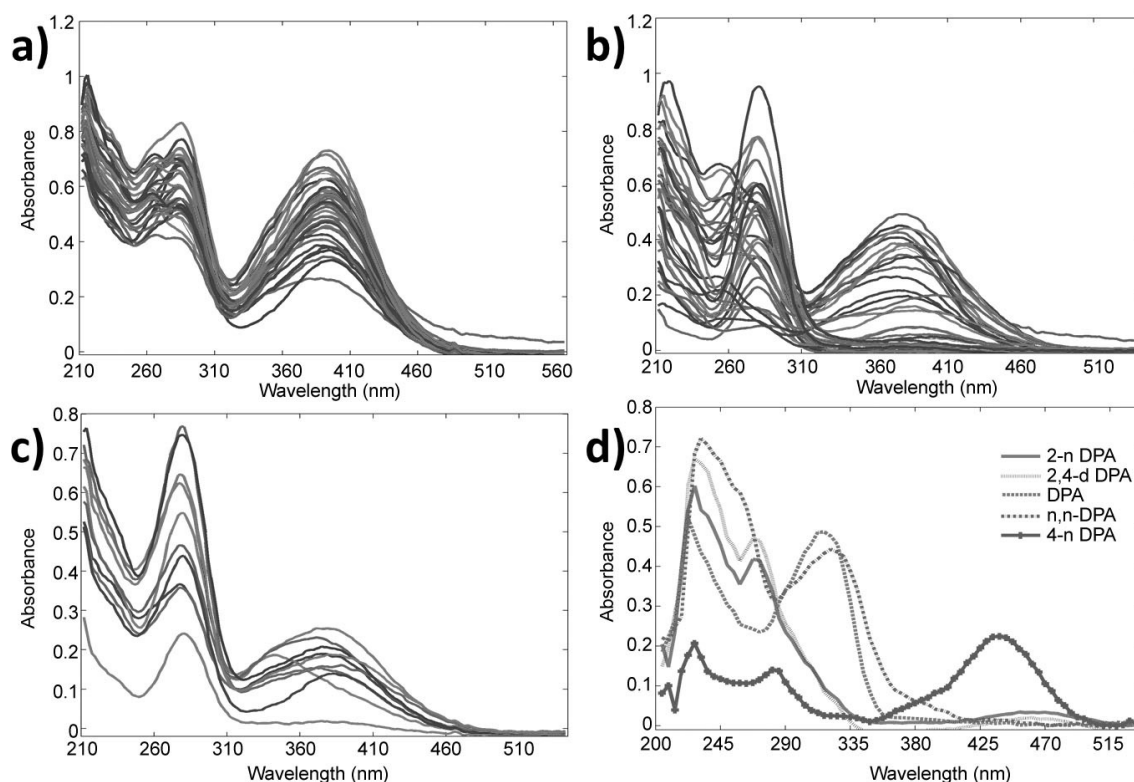## 3 Results and Discussion

3.1 Calibration Set

The raw UV-Vis spectra of different samples of the calibration set (with the concentration profiles built using CCD technique) are shown in Figure 1(a). Application of CCD design for making the calibration set yield good results in this work for predicting the concentration of analytes in validation set but when this calibration set was used for analysis of spiked real samples, the predicted concentration of some analytes were not reliable and accurate (average values of coefficient of determinations for prediction of spiked real samples were 0.74, 0.82, 0.65, 0.51 and 0.88 for PLS algorithm, respectively for DPA, N-n-DPA, 2-nDPA, 4-nDPA, and 2,4-DPA). This phenomenon was seen when different regression methods such as PLS, rPLS, and LASSO have been used. This problem was due to lack of enough spectral information in CCD-designed calibration set. CCD is a five level design and there can be lots of useful spectral information between the levels of the suggested discrete

Table 2. The design matrix of the calibration and validation sets constructed using CCD technique.

	Number	2-nDPA (μM)	2,4-dnDPA (μM)	DPA (μM)	N-n-DPA (μM)	4-nDPA (μM)
Calibration Set	1	7.8	7.8	20.5	6.4	9.5
	2	5.2	5.2	15.0	13.0	5.0
	3	2.6	2.6	20.5	6.4	9.5
	4	2.6	7.8	20.5	19.6	9.5
	5	7.8	2.6	9.5	6.4	9.5
	6	2.6	2.6	9.5	19.6	9.5
	7	7.8	7.8	9.5	19.6	9.5
	8	7.8	2.6	20.5	19.6	9.5
	9	2.6	7.8	9.5	6.4	9.5
	10	0.0	0.0	15.0	10.0	0.0
	11	0.5	0.0	10.0	5.0	0.0
	12	0.5	0.5	25.0	1.0	0.5
	13	0.0	0.0	10.0	1.0	5.0
	14	1.0	0.5	15.0	0.0	0.5
	15	0.5	0.0	25.0	5.0	0.0
	16	5.0	1.0	1.0	0.0	1.0
	17	5.0	1.0	5.0	0.5	3.0
	18	5.0	0.5	0.0	15.0	0.5
	19	0.0	0.0	0.5	15.0	1.0
	20	5.0	0.0	0.0	15.0	3.0
	21	10.0	5.0	0.0	0.0	5.0
	22	20.0	5.0	0.0	5.0	10.0
	23	25.0	10.0	5.0	0.0	5.0
	24	20.0	10.0	5.0	5.0	10.0
	25	5.0	1.0	20.0	15.0	0.5
	26	1.0	0.5	25.0	10.0	0.2
	27	0.2	0.2	30.0	5.0	0.2
	28	0.5	0.0	40.0	5.0	0.2
	29	0.1	0.1	10.0	30.0	0.1
	30	5.0	5.0	0.0	0.0	5.0
	31	10.0	5.0	5.0	10.0	1.0
	32	20.0	5.0	0.0	5.0	1.0
	33	15.0	10.0	0.0	5.0	5.0
	34	1.0	0.5	20.0	15.0	0.0
	35	0.5	0.2	25.0	10.0	0.5
	36	5.0	0.0	30.0	15.0	5.0
	37	0.5	0.5	40.0	20.0	0.5
Validation	1	10.0	8.0	20.0	5.0	5.0
	2	5.0	5.0	10.0	10.0	5.0
	3	1.0	10.0	20.0	8.0	0.0
	4	10.0	5.0	30.0	8.0	3.0
	5	3.0	8.0	15.0	8.0	5.0
	6	3.0	5.0	10.0	10.0	3.0
	7	5.0	5.0	25.0	15.0	5.0
	8	10.0	5.0	20.0	5.0	3.0
	9	5.0	0.0	15.0	10.0	5.0
	10	0.5	0.5	10.0	5.0	0.5

concentrations and this information is missed when CCD is applied. Therefore, we changed the design to a design made by random mixing of different concentrations of analytes in their linear dynamic ranges (37 solutions of the mixture of random concentration of analytes were prepared). The random generator operator of MATLAB software was used for generating the calibration samples. The UV-Vis spectra of the calibration samples obtained by mixing random concentrations of the analytes are illustrated in Figure 1(b). Comparison of the results in Figures 1(a) and 1(b) indicates that the ranges of the absorbances seen in UV-Vis spectra in the randomly-deigned calibration set are much wider than those obtained by the CCD algorithm. This can justify better predictions obtained when the randomly-designed calibration set has been used for prediction purposes. The results in this work suggests that when working with highly overlapped UV-Vis spectra, it is better to develop the calibration set using the randomly designed set of

LASSO as a multivariate calibration tool for simultaneous determination of DPA and its nitro derivatives in propellants

**Propellants,
Explosives,
Pyrotechnics**

Figure 1. The recorded UV-Vis spectra for (a) Calibration set designed using CCD approach, (b) Randomly designed Calibration set, (c) validation set (random design 10 samples) (d) pure standards of DPA and its nitro derivatives at different concentration levels.

concentrations rather than those proposed by CCD technique. The reason is the experimental gaps usually seen between the levels of concentrations in CCD levels. The UV-Vis spectra of the validation set (10 samples) are also shown in Figure 1(c).

3.2 The UV-Vis Spectra of DPA and its Nitro Derivatives

The UV-Vis spectra of pure DPA and its nitro derivatives are shown in Figure 1(d). As can be seen in this figure, there is a heavy overlap between the spectra of DPA and its derivatives. In addition, there is a narrow range of selective wavelengths for DPA (the most important analyte here) compared to other analytes. As it can be seen from the pure spectra of analytes and their mixtures, obtaining pure qualitative and quantitative information from target compounds using conventional univariate methods appears to be impossible. Multivariate calibration methods such as PCR and PLS may help for solving this complex problem. In addition, using of an appropriate variable selection method is useful when one works with a large number of overlapping variables (wavelengths). In this work, the recursive weighting and shrinkage strategies have been used for selection of informative and selective wavelengths for development of multivariate models.

3.3 Development of Multivariate Calibration Methods

In order to develop reliable and robust PCR and PLS models, it is necessary to find optimum values of the principal components (PCs) with maximum information for modeling signals rather than noises. In the present contribution, monte-carlo cross validation (MCCV) strategy has been used for finding optimum numbers of PCs. For each iteration of MCCV, one-third of the data set was used as a validation set. The procedure of random sampling was repeated 100 times. The variable selection and modelling procedures were performed using the data in the training set and the constructed model was used for the calculation of the concentration of samples in validation set. The average RMSE values (i. e. average of RMSE values for prediction of the concentrations of DPA, N-n-DPA, 2-nDPA, 4-nDPA and 2,4-dnDPA) for the training and validation sets for PCR and PLS models are illustrated in Figures 2a and 2b, respectively. As can be seen in these figures, selection of 6 PCs led to the best $RMSE_{MCCV}$ for PCR model and 8 latent variables (LVs) were selected as optimum numbers of variables for development of the PLS model. The results of PCR and PLS algorithms for prediction of the concentrations of DPA and its derivatives are given in Table 3. As can be seen in this Table, although PCR and PLS methods predicted the concentrations of DPA in different mixtures, but these two algo-
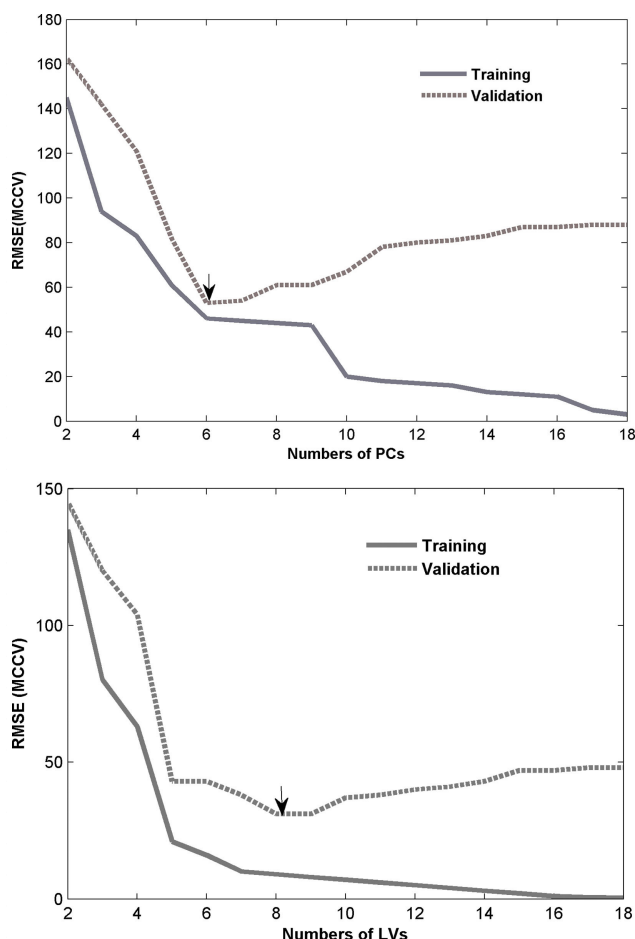
Figure 2. Root mean square error (RMSE) for monte-carlo cross validation (n = 100) against the numbers of a) principle components for PCR model and b) latent variables for PLS model. .

rithms fail to predict the concentrations of N-n-DPA, 4-nDPA and 2,4-dnDPA, precisely. This could be due to high degrees of overlapping in UV-Vis spectra of analytes, in this work.

In order to investigate about the effect of selective wavelengths on the performances of the models, recursive weighting and absolute shrinkage strategies have been used. The rPLS algorithm suggested by Rinnan et al. [24] has been used for implementation of recursive weighting in PLS algorithm. In rPLS algorithm, the regression coefficients (considered as variable weights) were recursively adjusted using the response vector (Y, i.e. concentration matrices in this work) to find the most important variables. In each iteration, the previous updated weighted X, (X_{R-1}), and Y (the response vector) are used to build a new PLS model; then the regression coefficients of the model stored in a vector (b), are used to reweight X_{R-1}:

$$X_R = X_{R-1} \times diag(b) \qquad (2)$$

where X_R is the new weighted data matrix. The important variables iteratively gain more weights, while the irrelevant variables to Y, are down-weighted. The algorithm converges to very limited number of wavelengths. In this method, the variables with the weights near zero are considered as unimportant variables, while large absolute weight values indicate the important ones. In the present work, a threshold equal to 10^{-4} was chosen to remove the variables with very small weights. The important selected wavelengths for determination of DPA and its nitro derivatives using rPLS algorithm are illustrated in Figure S2 in supporting material section. As can be seen in this figure, the rPLS algorithm could not find reasonable selective wavelengths for modeling the concentrations of 4-nDPA and 2,4-dnDPA. This is in agreement with poor prediction results for modeling the concentrations of these analytes using rPLS algorithm given in Table 3.

In order to find the selective and predictive wavelengths for modeling the concentration of DPAs in different mixtures, the least absolute shrinkage strategy has also been used, in this work. This is achieved by confining the L_1-norm of the coefficients of the regression model using the LASSO algorithm. The optimized values of λ parameter for model-

Table 3. Comparison of the performances of different multivariate calibration methods in this work for modeling the concentration of calibration and validation sets.

| | | Calibration | | | | Validation | | | |
		PCR	PLS	rPLS	LASSO	PCR	PLS	rPLS	LASSO
DPA	R^2	0.98	0.91	0.89	0.99	0.97	0.98	0.95	0.96
	RMSE*	0.91	0.80	0.62	0.38	0.78	0.66	0.63	0.57
N-n-DPA	R^2	0.92	0.90	0.88	0.97	0.90	0.90	0.70	0.95
	RMSE	0.81	0.89	0.11	0.11	0.93	0.86	1.72	0.78
2-nDPA	R^2	0.78	0.66	0.89	0.99	0.67	0.75	0.97	0.99
	RMSE	1.03	1.23	0.67	0.57	1.24	1.12	0.83	0.41
4-nDPA	R^2	0.95	0.95	0.88	0.96	0.91	0.93	0.76	0.93
	RMSE	0.52	0.62	1.09	0.62	0.58	0.55	1.17	0.60
2,4-dnDPA	R^2	0.96	0.92	0.88	0.99	0.98	0.98	0.87	0.98
	RMSE	0.69	0.79	0.33	0.70	0.73	0.75	0.94	0.61

* Root Mean Square Error

LASSO as a multivariate calibration tool for simultaneous determination of DPA and its nitro derivatives in propellants

Propellants, Explosives, Pyrotechnics

ing the concentration of DPA, N-n-DPA, 2-nDPA, 4-nDPA and 2,4-dnDPA using LASSO regression and monte-carlo cross-validation strategies were 0.143, 0.18, 0.15, 0.12 and 0.22, respectively. The shrinkage patterns of regression co-efficients for modeling the concentration of DPA and its ni-tro derivatives are shown in Figures S3–S7 in supporting material section. As can be seen in these figures, the LASSO algorithm selects informative and selective wavelengths for model developments and turns off the correlated and re-dundant variables, automatically. The selected wavelengths using the LASSO approach for modeling the concentrations of DPA and its nitro derivatives are illustrated in Figure 3. Comparison of results in this figure with those of Figure S2 reveals that the number of selected set of adjacent varia-bles using LASSO algorithm are too much less than those of rPLS algorithm. The adjacent variables are usually correlated and LASSO correctly selected less number of correlated wavelengths. In addition, the selected wavelengths using LASSO approach are more reasonable than those of rPLS al-gorithm. For example, LASSO selected some wavelengths in the range of 400–500 nm for modeling the concentration of 4-nDPA which justifies the selective wavelengths exists for 4-nDPA in this range (please see Figure 1d). Moreover, the LASSO approach selected different wavelengths in the range of 290–350 for modeling the concentration of DPA and N-n-DPA and this is in agreement with intrinsic se-lective wavelengths for these two analytes seen in Fig-ure 1d. The detailed specification of the PCR, PLS, rPLS and LASSO models are given in Table 2. As can be seen in this Table, the LASSO algorithm outperforms the rPLS, PLS and PCR models for simultaneous determination of DPAs. The superiority of LASSO over rPLS approach may be due to the ability of LASSO algorithm for thorough screening of corre-lated variables. The results in this section reveal that choos-ing a thorough strategy for selection of predictive wave-lengths is of major importance when dealing with very highly overlapping spectra for multivariate calibration. The shrinkage approach implemented in LASSO regression tech-nique, adequately selected informative and predictive wavelengths in this work for simultaneous determination of DPAs.

3.4 Analysis of Real Samples

After assessment of the prediction power of the LASSO ap-proach for describing the concentrations of DPAs in calibra-tion and validation sets in section 3.3, in this section LASSO algorithm has been used for determination of the concen-tration of DPAs in real samples. Two different propellants were analyzed as real samples (these samples were named as P230 and P300). Preparation of real samples before ac-quisition of their UV-Vis spectra is completely described in section 2.2. Adequate filtering and extraction steps were im-plemented before collecting UV-Vis spectra in order to mini-mize the effects of scattering and unknown interferences.
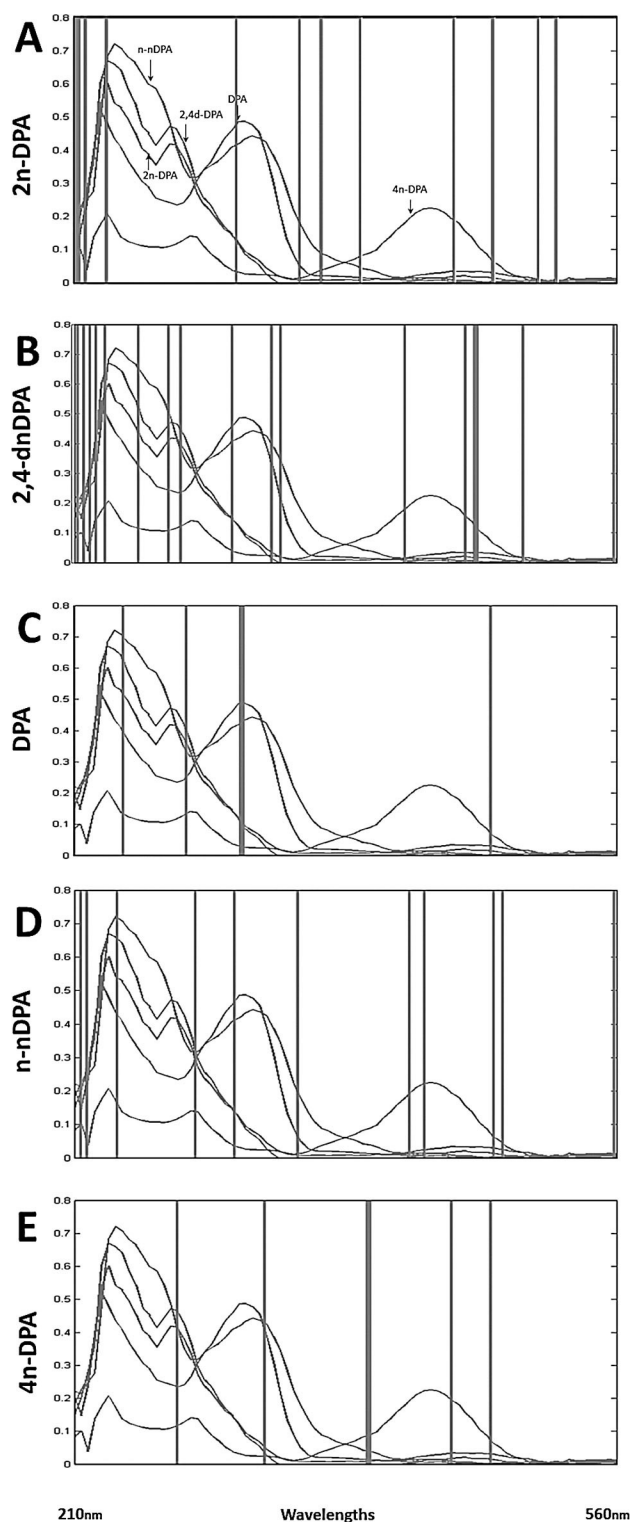
Figure 3. The selected wavelengths using LASSO algorithm for si-multaneous determination of DPA and its nitro derivatives.

The UV-Vis spectra of two real samples are shown in Fig-ure S8 in supplementary material section (five UV-Vis spec-tra were recorded for each sample).

Table 4 shows the results for determination of analytes in two types of propellants using collected UV-Vis spectra and LASSO algorithm. The RSD values (n = 5) were too large for all considered analytes except for DPA. The calculated concentrations for N-n-DPA, 2-nDPA, 4-nDPA and 2,4-dnDPA were low since there were no significant signals in recorded UV-Vis spectra for these analytes. It also justifies the large RSD values for these four analytes. Moreover, the calculated concentrations of these four analytes were less than the linear dynamic range we found using univariate calibration approach in λ_{max} for each analytes (data not shown). Inspection of Figure S8 reveals that there are no significant signals in wavelengths greater than 330 nm, while inspection of Figure 1(d) represents that N-n-DPA, 2-nDPA, 4-nDPA and 2,4-dnDPA have considerable absorbances in wavelengths longer than 330 nm. It implicitly reveals the very low concentration levels of N-n-DPA, 2-nDPA, 4-nDPA and 2,4-dnDPA in analyzed reals samples and this is in agreement with those obtained by the LASSO model (please see Table 4). In order to test the accuracy of the developed LASSO model for determination of DPAs in real samples, standard addition approach has been applied. The standard addition approach has been implemented for analysis of two real samples (P230 and P300). Figure 4 shows the plots of the predicted concentrations of analytes in real samples against the spiked concentrations. Good correlation coefficients, excellent linear dependencies and reasonable slopes seen in these plots imply that the developed method in this work is both sensitive and accurate. As can be seen in these figures, there are no significant intercepts in calibration curves except for DPA. It suggests that the concentration of N-n-DPA, 2-nDPA, 4-nDPA and 2,4-dnDPA are not considerable in analyzed real samples (P300). However, for DPA, a significant intercept is seen in calibration curve which implies that there exists notable amount of DPA in this real sample of propellant. Extrapolation of this calibration curve gave the value of 27.2 ppm as the value of DPA in this real sample which this is in agreement with those seen in Table 4.

Generally, the results in this work revealed that a thorough combination of a simple spectroscopic method like as UV-Vis and robust multivariate calibration methods can be used for the simultaneous analysis of complex mixtures of DPAs. It has been shown that selection of informative and predictive wavelengths significantly affects the predictive power of multivariate models. In this work, LASSO approach has been used for automatic selection of informative wavelengths to develop a robust model for simultaneous determination of DPAs, for the first time. It is notable that the performance of the proposed hybrid method in this work (i.e. UV-Vis/LASSO) is comparable with those of complex analytical instruments such as high-performance liquid chromatography-diode array detector (HPLC-DAD). Table 5 summarizes a comparison of this method with common analytical tools previously used by researchers for analysis of DPAs [1, 6, 7, 9, 26, and 27]. Inspection of this Table reveals that the developed method in this work (UV-Vis/LASSO) is comparable with liquid chromatographic methods in terms of LOD and RSD values. The main advantages of UV-Vis/LASSO approach over liquid chromatographic methods are its speed, simplicity, low cost and ease of implementation. However, higher sensitivities and better LODs were seen for fluorescence and tandem mass compared to UV-Vis/LASSO

Table 4. Results of determination of two real propellant samples (P230 and P300) using UV-Vis/LASSO approach.
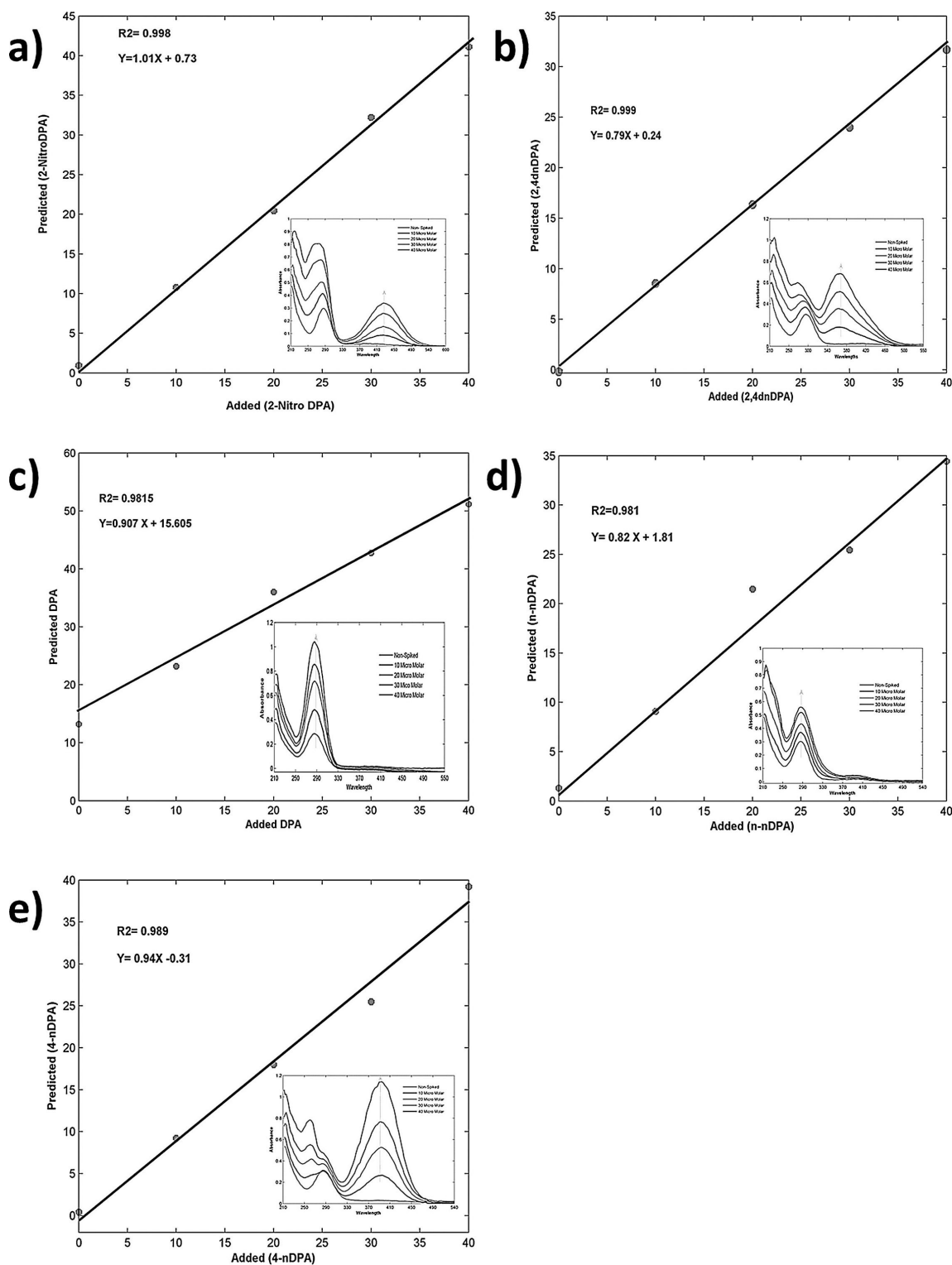
| | P300 | | | P230 | | | |
	C_{pred}[(a)]	%RSD	%M[(b)]	C_{pred}	%RSD	%M	LOD (ppm)
2-nDPA	1.60	32	0.08	1.23	38.36	0.04	0.49
2,4-dnDPA	0.73	58.05	0.05	0.74	58.40	0.03	0.82
DPA	24.65	8.20	1.04	26.29	6.92	0.74	0.79
N-n-DPA	2.26	65.21	0.11	3.93	55.48	0.13	1.16
4-nDPA	0.45	56.31	0.03	0.86	61.20	0.02	0.50

(a) C_{pred} is the mean concentration for five determinations. (b) %M is the mass percentage of analytes in propellant

Table 5. Comparison of UV-Vis/LASSO approach with techniques reported in literature for determination of DPAs.

Method	Sample	LDR (mg/kg)	LOD (mg/kg)	RSD (%)	R^2	Ref
SFC-SSME	Apple, Orange, pear	0.5–7.0	0.20	7.1	0.998	[6]
Spectrofluorimetry (SF-DS)	Water, gun powder	$(6.8–1500) \times 10^{-3}$	1.7×10^{-3}	3 <	–	[7]
CEC-MS (capillary electro chromatography)	Smokeless powder	–	0.6	< 14	–	[26]
HPLC/TS/MS	Soil sample	–	~50–100 ng	–	–	[27]
			0.024	–	–	[1]
		DPP (Differential pulse polarography)	–	–	Simple base solid propellant	
Tandem Mass	Smokeless gunpowder	$5.0–200.0 \times 10^{-3}$	1.0×10^{-3}	11.3 >	–	[9]
UV-Vis/LASSO	Solid Propellant	1–10	0.79	8.2	0.972	This work

LASSO as a multivariate calibration tool for simultaneous determination of DPA and its nitro derivatives in propellants

Figure 4. The predicted concentrations of analytes using LASSO algorithm against the spiked concentrations for a) 2-nDPA, b) 2,4-dnDPA, c) DPA, d) N-n-DPA, e) 4-nDPA.

approach, which this is due to intrinsic sensitivity of fluorescent and mass spectrometry. Since the amount of DPAs in propellant samples must be between 0.2–2 of mass percent therefore, intrinsic low of sensitivity of UV-Vis would

not affect the performance of the proposed approach in this work for simultaneous determination of DPAs in propellants as real samples.

4 Conclusions

DPA is a very important preservative usually used in propellants as effective anti-aging factor. Degradation of this molecule in Nitrogen-rich environments usually led to formation of mono- and di-nitro derivatives of DPA. Determination of DPA and its nitro derivatives is of major interest for tracking and monitoring the expiration time of propellants. The present study paves the way for development of a simple and effective method for simultaneous determination of DPA and its nitro-derivatives. In this line, the capabilities of UV-Vis spectroscopy combined with different multivariate calibration methods were explored for determination of DPA, in synthetic and real samples. The results revealed that thorough selection of informative wavelengths highly affects the performances of multivariate models for sample analysis. It is found that confining L_1-norm of the regression coefficients in multivariate models is a promising route for automatic selection of predictive and selective wavelengths, when dealing with very highly overlapping signals. Validation of multivariate models in this work revealed that the LASSO approach is superior over rPLS, PLS and PCR models for determination of DPA and its nitro derivatives. Generally, the proposed UV-Vis/LASSO approach in this work is simple, accurate and non-destructive and can be used as a reliable tool for tracking the expiration time of propellants in repositories.

References

[1] A. P. de Diego Martínez, M. Tascón, M. Vázquez, P. S. Batanero, Polarographic study on the evolution of the diphenylamine as stabiliser of the solid propellants, *Talanta* **2004**, *62*, 165–173.

[2] D. R. Rudell, J. P. Mattheis, J. K. Fellman, Evaluation of diphenylamine derivatives in apple peel using gradient reversed-phase liquid chromatography with ultraviolet–visible absorption and atmospheric pressure chemical ionization mass selective detection, *J. Chromatogr. A*, **2005**, *1081*, 202–209.

[3] B. Saad, N. H. Haniff, M. I. Saleh, N. H. Hashim, A. Abu, N. Ali, Determination of ortho-phenylphenol, diphenyl and diphenylamine in apples and oranges using HPLC with fluorescence detection, *Food Chem.* **2004**, *84*, 313–317.

[4] M. Olek, Determination of diphenylamine residues in apples, and 4-aminobiphenyl residues in diphenylamine, by high-performance liquid chromatography and electrochemical detection *J. Chromatogr. A* **1998**, *447*, 421–425.

[5] J. Garrido, M. De Alba, I. Jimenez, E. Cadado, M. L. Folgeiras, Gas-Chromatographic determination of diphenylamine in apples and pears- Method validation and results of Spanish-official-residue-monitoring-program 1995, *J. AOAC Int.* **1998**, *81*, 648–651.

[6] F. Rezaei, Y. Yamini, H. Asiabi, M. Moradi, Determination of diphenylamine residue in fruit samples by supercritical fluid extraction followed by vesicular based-supramolecular solvent microextraction, *J. Supercrit. Fluids* **2015**, *100*, 79–85.

[7] N. Alizadeh, A. Farokhcheh, Simultaneous determination of diphenylamine and nitrosodiphenylamine by photochemically induced fluorescence and synchronous fluorimetry using double scans method, *Talanta* **2014**, *121*, 239–246.

[8] M. López-López, J. C. Bravo, C. García-Ruiz, M. Torre, Diphenylamine and derivatives as predictors of gunpowder age by means of HPLC and statistical models, *Talanta* **2013**, *103*, 214–220.

[9] Y. Tong, Z. Wu, C. Yang, J. Yu, X. Zhang, S. Yang, X. Deng, Y. Xu, Y. Wen, Determination of diphenylamine stabilizer and its nitrated derivatives in smokeless gunpowder using a tandem MS method, *Analyst* **2001**, *126*, 480–484.

[10] C. West, G. Baron, J.-J. Minet, Detection of gunpowder stabilizers with ion mobility spectrometry, *Forensic Sci. Int.* **2007**, *166*, 91–101.

[11] Sh. Zhou, Zh. Wang, L.Lu, Q.Yin, L. Yu, G.Deng, Rapid quantification of stabilizing agents in single-base propellants using near infrared spectroscopy, *Infrared Phys. Technol.* **2016**, *77*, 1–7.

[12] H. Parastar, H. Shaye, Comparative study of partial least squares and multivariate curve resolution for simultaneous spectrophotometric determination of pharmaceuticals in environmental samples, *RSC Adv.* **2015**, *5*, 70017–70024.

[13] E. Dinç, D. Baleanu, G. Ioele, M. De Luca, G. Ragno, Multivariate analysis of paracetamol, propiphenazone, caffeine and thiamine in quaternary mixtures by PCR, PLS and ANN calibrations applied on wavelet transform data, *J. Pharm. Biomed. Anal.* **2008**, *48*, 1471–1475.

[14] Z. Rezaei, B. Hemmateenejad, S. Khabnadideh, M. Gorgin, Simultaneous spectrophotometric determination of carbamazepine and phenytoin in serum by PLS regression and comparison with HPLC, *Talanta* **2005**, *65*, 21–28.

[15] A. G. Rodriguez, A. G. De Torres, J. C. Pavon, C. B. Ojeda, Simultaneous determination of iron, cobalt, nickel and copper by UV-Visible spectrophotometry with multivariate calibration, *Talanta* **1998**, *47*, 463–470.

[16] B. Hemmateenejad, M. Akhond, F. Samari, A comparative study between PCR and PLS in simultaneous spectrophotometric determination of diphenylamine, aniline, and phenol: Effect of wavelength selection, *Spectrochim. Acta Part A* **2007**, *67*, 958–965.

[17] A. R. Khanchi, M. K. Mahani, M. Hajihosseini, M. G. Maragheh, M. Chaloosi, F. Bani, Simultaneous spectrophotometric determination of caffeine and theobromine in Iranian tea by artificial neural networks and its comparison with PLS, *Food Chem.* **2007**, *103*, 1062–1068.

[18] S. Şahin, C. Demir, Ş. Güçer, Simultaneous UV–vis spectrophotometric determination of disperse dyes in textile wastewater by partial least squares and principal component regression, *Dyes Pigm.* **2007**, *73*, 368–376.

LASSO as a multivariate calibration tool for simultaneous determination of DPA and its nitro derivatives in propellants

**Propellants,
Explosives,
Pyrotechnics**

[19] Y. Ni, G. Zhang, S. Kokot, Simultaneous spectrophotometric determination of maltol, ethyl maltol, vanillin and ethyl vanillin in foods by multivariate calibration and artificial neural networks, *Food Chem.* **2005**, *89*, 465–473.

[20] E. Dinc, D. Baleanu, Spectrophotometric quantitative determination of cilazapril and hydrochlorothiazide in tablets by chemometric methods, *J. Pharm. Biomed. Anal.* **2002**, *30*, 715–723.

[21] C.H .Zhang, Y. H. Yun, Zh.M. Zhang, Y. Z. Liang, Simultaneous determination of neutral and uronic sugars based on UV-vis spectrometry combined with PLS, *Int. J. Biol. Macromol.* **2016**, *87*, 290–294.

[22] A. Abbaspour, R. Mirzajani, Simultaneous determination of phenytoin, barbital and caffeine in pharmaceuticals by absorption (zero-order) UV spectra and first-order derivative spectra-multivariate calibration methods, *J. Pharm. Biomed. Anal.* **2005**, *38*, 420–427.

[23] S. Wold, M. Sjostrom, L. Eriksson, PLS-regression: a basic tool of Chemometrics, *Chemom. Intell. Lab. Syst.* **2001**, 58, 109–130.

[24] A. Rinnan, M. Andersson, C. Ridder, S. B. Engelsen, Recursive weighted partial least squares (rPLS): an efficient variable selection method using PLS, *J. Chemom.* **2014**, *28*, 439–447.

[25] R. Tibshirani, Regression shrinkage and selection via the Lasso, *J. R. Stat. Soc. B.* **1996**, *58*, 267–288.

[26] C. de Perre, I. Corbin, M. Blas, B. R. McCord, Seperation and identification of smokeless gunpowder additives by capillary electrochromatography, *J. Chromatogr. A.* **2012**, *1267*, 259–256.

[27] J. S. Ho, Th. A. Bellar, J. W. Eichelberger, W. L. Budde, Distinguishing diphenylamine and N-nitrosodiphenylamine in hazardous waste samples by using high-performance liquid chromatography/thermospray/mass spectrometry, *Environ. Sci. Technol.* **1990**, *24*, 1748–1751.