



Contents lists available at ScienceDirect

Coordination Chemistry Reviews

journal homepage: www.elsevier.com/locate/ccr

Review

Chemometric models for data processing in quantum dots-based photoluminescence methodologies

Rafael C. Castro, Ricardo N.M.J. Páscoa^{*}, M. Lúcia M.F.S. Saraiva, João L.M. Santos^{*}, David S. M. Ribeiro^{*}

LAQV, REQUIMTE, Department of Chemical Sciences, Laboratory of Applied Chemistry, Faculty of Pharmacy, University of Porto, Rua de Jorge Viterbo Ferreira n° 228, 4050-313 Porto, Portugal

ARTICLE INFO

Keywords:

Chemical analysis
Quantum dots
Chemometrics
First-order advantage
Second-order advantage

ABSTRACT

The appealing properties of quantum dots (QDs) have drawn the scientific community's attention, leading to extensive research on using these nanomaterials as sensing platforms for the detection and quantification of a variety of analytes in environmental, biological, pharmaceutical and food samples. Despite the multiple inventive strategies that can be used to develop efficient QDs-sensing schemes, the defiant reactivity of these nanomaterials, and their propensity to establish non-specific interactions, has significantly restrained their utilisation in situations demanding high selectivity, as is the case of the quantification of analytes in samples with interfering species or complex matrices, and in multiplexed detection. Several approaches have been proposed to overcome these selectivity issues, among which the chemometric analysis of photoluminescent (PL) data acquired from QDs-based analytical methodologies can be highlighted.

This review details the application of chemometric models in the characterization and optimization of QDs-based analytical procedures, as well as for the analysis of data obtained from QDs-based PL methodologies, discussing how they can be used to circumvent selectivity issues and pointing out the corresponding advantages

Abbreviations: (NH₄)₃L, triammonium-N-dithiocarboxyiminodiacetate; 4-NP, 4-nitrophenol; Act D, actinomycin D; ADF, analyte-dependent fluorophore; AFB1, aflatoxin B1; AFM, Atomic Force Microscopy; AgNCs, silver nanocluster; AgNPs, silver nanoparticles; AIS, AgInS₂; ALS, alternating least squares; ANN, artificial neural network; APIs, active pharmaceutical ingredients; Arg, arginine; AuNCs, gold nanoclusters; BDA, Bayes discriminant analysis; bPEI, polyethyleneimine; BSA, bovine serum albumin; CA, cysteamine; CB, calcein blue; CCD, central composite design; CDs, carbon dots; Cit, citric acid; COFs, covalent organic frameworks; CYS, L-cysteine; DAB, polypropyleneimine tetrahexacontamine dendrimer; DD-SIMCA, data-driven soft independent modelling of class analogy; DG, dodecyl gallate; DMAE, 2-dimethyl-aminoethanethiol; DNT, 2,4-dinitrotoluene; DT, Decision Table; DTAB, dodecyl trimethyl ammonium bromide; DTD, direct trilinear decomposition; EDA, Ethylenediamine; EDS, Energy-Dispersive x-ray Spectroscopy; EEM, excitation emission matrices; EET, exciton energy transfer; EFA, evolving factor analysis; FA, folic acid; FRET, Forster resonance energy transfer; Gln, glutamine; Gly, l-glycine; GSH, glutathione; GSSG, glutathione disulphide; HCA, hierarchical cluster analysis; HR, high resolution; HTD, hard trilinear decomposition; ILS, ionic liquids; LC-MS/MS, liquid chromatography with mass spectrometry detection; LDA, linear discriminate analysis; LEDs, light-emitting diodes; LS-SVM, least-squares support vector machine; LV, latent variables; Lys, l-lysine; MCR, multivariate curve resolution; MCR-ALS, multivariate curve resolution with alternating least squares; MDA, mahalanobis distance analysis; MES, 2-mercaptoethanesulfonate; MG, methyl gallate; MIPs, molecularly imprinted polymers; MLF-NNs, multilayer feed-forward neural networks; MLR, multiple linear regression; MNDC, mercapto-n-dodecanoate; MPA, mercaptopropionic acid; MWPLS-DA, moving window partial least squares discriminant analysis; NALC, N-Acetyl-L-cysteine; nano-ZnTPyP, Zn-nanoporphyryrin; NIR, Near-infrared; NPG, n-propyl gallate; N-PLS, N-way partial least-squares; OCPLS, one-class partial least squares; OTC, oxytetracycline; PAA, polyacrylate; PARAFAC, parallel factor analysis; PARAFAC2, Variant of the parallel factor analysis; PCA, principal component analysis; PCR, principal component regression; Pdots, semiconducting polymer dots; PEG, polyethylene glycol; PET, photoinduced electron transfer; PL, photoluminescent; PLS, partial least squares; PLS-DA, partial least squares discriminant analysis; PMAA, polymethacrylic acid; PSO-OSWLS SVM, optimized sample-weighted least-squares support vector machine based on particle swarm optimization; QDs, quantum dots; QY, quantum yield; R², coefficient of determination; RBFN, radial basis function neural network; RBF-NNs, radial basis function neural networks; RBL, residual bilinearization; RE, relative percentage error; RF, reference fluorophore; RhB, rhodamine B; RHD, rhodamine derivative; RMSEC, root mean square error of calibration; RMSECV, root mean square errors of cross-validation; RMSEP, root mean square error of prediction; ROS, reactive oxygen species; SEM, Scanning Electron Microscopy; Ser, l-serine; SIMCA, self-independent modelling of class analogy; SVD, singular value decomposition; SVM, support vector machine; TEM, Transmission Electron Microscopy; TG, alpha-thioglycerol; TGA, thioglycolic acid; TMA, thiomalic acid; TNP, 2,4,6-trinitrophenol; TNT, 2,4,6-trinitrotoluene; U-PLS, unfolded partial least-squares; U-PLS/RBL, Partial Least Squares followed by Residual Bilinearization; XRD, X-ray powder diffraction.

^{*} Corresponding authors.

E-mail addresses: rnascaoa@ff.up.pt (R.N.M.J. Páscoa), joalms@ff.up.pt (J.L.M. Santos), dsmribeiro@gmail.com (D.S.M. Ribeiro).

<https://doi.org/10.1016/j.ccr.2023.215605>

Received 24 October 2023; Accepted 1 December 2023

Available online 16 December 2023

0010-8545/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

and limitations. In this work, we provide insights not only about probe arrangement strategies that could be designed to obtain efficient QDs-based sensing platforms but also regarding the requirements that must be observed to select both the most suitable type of data and the most effective chemometric model to assure the objectives of the methodology. Related advantages, namely second and higher-order advantages, constraints, and application prospects are also discussed.

1. Introduction

The exceptionally attractive properties of quantum dots (QDs) make them one of the most auspicious nanomaterials in nanoscience and nanotechnology with a huge variety of applications [1,2], namely in bioimaging/biosensing [3–6], photocatalysis [7,8], chemical analysis [1,9–11], as light-emitting diodes (LEDs) or in light-harvesting systems [12,13]. This functional versatility is founded on their remarkable optoelectronic and chemical properties, which include: i) easily adjustable and controlled synthesis allowing to obtain nanomaterials of different composition, size and/or capping ligands; ii) wide excitation (due to the ability to absorb electromagnetic radiations in a wide range of wavelengths, usually with high molar attenuation coefficients) and narrow emission bands which allow the simultaneous excitation, at a single wavelength, of multiple QDs combined in a multi-emission probe without significant overlap between emission bands; iii) high quantum yields (QY) as well as impressive photochemical stability, when compared to other organic fluorophores, which provides enhanced sensitivity and permits kinetic-based procedures with long irradiation periods; iv) the surface chemistry of these nanoparticles can be effortlessly tailored enabling to adjust the selectivity and sensitivity of the sensing platform towards a specific target analyte through the QDs surface functionalization with distinct structures/biomolecules such as molecularly imprinted polymers (MIPs), aptamers and/or antibodies, etc [9,11,14,15]. Moreover, QDs could be combined with other organic or inorganic fluorescent materials, in a cross-reactive sensing array or in a multi-emitter nanoprobe, aiming at implementing multiplexed detection approaches [1,11]. The aforementioned features of QDs make them excellent fluorophores to be used in Forster Resonance Energy Transfer (FRET) schemes either as donors or acceptors [15].

In recent years, these nanomaterials have been successfully employed in some promising analytical methodologies, namely in ratiometric assays [16–18], and in visual [11,19,20] and multiplexed detections [4,10,21]. However, the reduced selectivity they usually exhibit towards the target analyte when their surfaces are not functionalized with a specific recognition element, is a major drawback that restrains a more widespread application in chemical analysis, particularly when it involves samples with complex matrices.

Photoluminescent (PL) data analysis through chemometric tools has been gaining relevance as a pertinent solution to understand more efficiently the interrelations between samples and analysed variables [10]. When used in the implementation of analytical methodologies based on QDs PL, the versatility of chemometric tools can allow not only the simultaneous detection of multiple analytes in the same sample, in a single assay, but also enable the quantification of the target analyte in complex sample matrices, prone to the occurrence of interfering species.

Alongside their use in QDs-based methodologies for analytical purposes, chemometrics can also be applied in the optimization of nanocrystals synthesis by assisting on the implementation of experimental designs aiming at defining the relative importance of each of the synthesis variables, as well as the relationship between them [14]. Moreover, the chemometric models have also allowed to predict some important characteristics of the QDs such as their morphology (shape) [22], composition [23] and size [24,25] through the analysis of data obtained by vibrational or absorption spectrophotometry, and fluorometry.

Chemometrics is a term introduced by Svante Wold in 1972 [26] to define the science of extracting valuable chemical information from

complex experimental systems, converting the obtained data into useful information [27]. In fact, univariate methods, in which one variable is analysed independently of the others, are not capable to recognize the inter-correlation between the multiple variables. Contrarily, multivariate strategies can consider different variables thus assuring a more detailed interpretation of the available data [27].

There are different chemometric models available in the literature, which can be used with distinct objectives, being their most common application for quantification and classification/discrimination purposes. The selection of a specific model depends not only on the aim of the work but also on the type of generated data (zero-, first-, second- or higher-order data) as well as whether it shows a linear or nonlinear behaviour. It should be highlighted that before the selection of a given chemometric model, it is crucial to perform exploratory data analysis for the detection of outliers. This is usually carried out by using principal component analysis (PCA). This model compresses the original data and retains the most important information, maximizing its variance. It is considered as an unsupervised analysis because the model has no previous knowledge of the original data [28]. Regarding quantification purposes, there are several chemometric models (ex: artificial neural networks- ANN; partial least squares- PLS; multivariate curve resolution - MCR; principal component regression- PCR) that can be used. All these models require information regarding the calibration samples (ex: concentration values) to quantify the analytes of interest. PLS and PCR are very similar in their working principles: both start building an inverse model although PCR estimates latent variables that are independent of the analyte, while with PLS the estimated latent variables are analyte dependent. This means that PCR does not use information regarding analyte concentrations or sample properties to estimate the latent variables [28]. The MCR model relies on a bilinear model decomposition to extract the relevant information and can be applied for quantification purposes through the application of the concentration correlation constraint [29]. In relation to ANN, this tool is often applied to nonlinear data, being capable of modelling complex functions and of “learning by example” [29]. For classification/discrimination purposes, linear discriminant analysis (LDA), partial least squares-discriminant analysis (PLS-DA) and soft independent modelling of class analogy (SIMCA) are the chemometric tools most widely used [28–30]. LDA maximizes the variance among the different classes and minimizes the variance within each class. However, in what represents a severe drawback, the number of variables needs to be smaller than the number of samples. Therefore, most of the time, PCA should be firstly used to reduce data dimension [31]. The working principle of PLS-DA is similar to PLS but the Y block contains a dummy matrix with the class membership of the samples, instead of sample properties values [32]. As this model performs data dimension reduction, and is capable of dealing with data where the variance within each class is higher than among different classes, it is considered to have a better performance than LDA [32]. SIMCA develops a separate model for each class and then provides a probability value for the classes that each sample can belong to. As SIMCA deals with each class independently, it is less affected by the number of existing classes, which is a limiting factor for LDA [33]. Moreover, SIMCA can deal with test samples that do not belong to any of the predefined classes, which does not happen with LDA and PLS-DA. These can therefore lead to misleading classifications as both models will always discriminate the test samples according to one of the predefined classes [34]. All these chemometric models are able to handle only linear data. When dealing with non-linear data for classification/

discrimination purposes, support vector machine (SVM) and ANN models are more appropriate.

Bearing this in mind, before the development of a new methodology exploring QDs sensing platforms and chemometrics for data analysis, the researchers need to select the type of data that will be generated (zero-, first-, second- or higher-order data). For example, if the samples that will be analysed have no interfering species, zero-order data (e.g.: a PL value at a specific wavelength) can be used. On the other hand, if the samples that will be analysed have known interfering species in their composition, first-order data (e.g.: PL values within a specific range) are the most suitable. Finally, the acquisition of second-order data (e.g.: PL values within a specific range within a specific period of time) is the best option when dealing with samples that have a complex matrix, with unknown interfering species. After knowing this, the researchers need to select the most suitable chemometric model, considering the purpose of the work (e.g.: quantification or classification/discrimination) and if the acquired data fit linear or non-linear models. However, in some situations, it is not possible to know in advance if the acquired data has a linear or non-linear behaviour. In those situations, the researchers should try linear models first because are simpler and faster than non-linear models and then analyse the results, to verify if it is necessary to implement non-linear models.

Despite the enormous potential of using chemometric models, not only for the optimization and characterization of QDs but also to circumvent selectivity issues in QDs-based chemical analyses, they are still slightly explored when compared to univariate methods. This review highlights the use of chemometric tools in both instances, discussing the advantages and limitations, and challenges and opportunities of combining QDs and chemometrics.

2. Optimization of QDs synthesis and/or characterization by applying chemometric tools

The optimization of nanocrystal synthesis is crucial to ascertain the experimental conditions that allow obtaining more efficient nanomaterials in terms of optical properties (fluorescent intensity, QY and PL decay lifetime), stability in solution (aqueous or organic solvent), morphology (size and shape), and reactivity (passivation of QDs surface with capping ligands containing selected terminal moieties to provide the required probing ability towards the target analyte).

Distinct synthetic procedures (organometallic, conventional aqueous synthesis, aqueous synthetic route assisted by microwave irradiation, among others) have been developed allowing the preparation of a wide assortment of size-controlled nanoparticles, which emit at a wide range of wavelengths (from visible to near-infrared – NIR - region of the electromagnetic spectrum). The rigorous control of all parameters involved in their synthesis is decisive for obtaining the desired optical and morphological characteristics. Indeed, only by controlling parameters such as the precursors' relative molar ratio, temperature, pH, reaction/irradiation/heating time, solvent, and nature of passivating agents, it is possible to obtain high quality nanomaterials adjusted to the required purpose.

The characterization of the nanocrystals in terms of morphology (size and shape) and structure (core composition, shell and capping layer) can be performed by advanced techniques such as Scanning Electron Microscopy (SEM), Transmission Electron Microscopy (TEM), Atomic Force Microscopy (AFM), X-Ray powder Diffraction (XRD), Energy-Dispersive x-ray Spectroscopy (EDS), among others. Nevertheless, simpler equipment, such as a spectrofluorometer, which are intrinsically easier to operate, faster, and more cost-effective, could be used to analyse the nanomaterials, generating data that upon processing by chemometric models allows to accurately assess the nanocrystal properties.

2.1. Optimization of the organometallic synthetic route and non-aqueous dispersed QDs characterization

One of the most common approaches to synthesize QDs is the organometallic synthetic route, which relies on the hot injection of semiconductor precursors in specific organic solvents followed by a controlled heating, in terms of temperature and time, of the obtained mixture. The obtained QDs show usually higher QY and a narrower particle size distribution (narrower emission bands) than those synthesized by aqueous synthetic routes [35,36]. Three different chemometric approaches [22–24] have been developed with the objective of obtaining a simpler and faster characterization of the nanocrystals (Table 1). The first example describes a methodology for the size determination of hydrophobic CdSe and hydrophilic CdSe/ZnS QDs, in a rapid and reliable way, based on the fluorescence maxima positions of the pure spectral profiles retrieved by multivariate curve resolution-alternating least squares (MCR-ALS) [24]. The authors obtained excitation-emission matrices (EEM), by using a measuring procedure that involved the excitation of the QDs at increasing wavelengths (within a pre-selected spectral range), which were further analysed by the respective chemometric model. The number of components was selected (the authors did not mention if this was performed manually or by using singular value decomposition, SVD) with the help of evolving factor analysis (EFA), seeking to appraise the number of fractions of differently-sized QDs in solution. EFA is a possible option for the estimation of the spectra components at the beginning of ALS optimization but it is more indicated for the analysis of evolving processes, which was not the case [37]. The subsequent application of MCR in the obtained data after ALS optimization, allowed not only a more exact assessment of the number of spectral components but also to obtain the pure spectral components. The number of pure spectral components is strictly related to the number of fractions of QDs of different diameter, present in the solutions. By using the maxima of the pure spectral profiles retrieved by MCR-ALS it was possible to estimate the average particle size. Despite the fact that TEM provides a more accurate determination of the referred size, its use for the rapid tracking of the QDs size distribution during the synthesis becomes unfeasible. In this sense, the described fluorometric methodology can be seen as an accessible and valuable alternative [24].

In the second example, variations in the absorbance spectra of CdSe QDs solutions, taking place during nanocrystals growth, were analysed by MCR-ALS. This allowed to evaluate the kinetics and mechanism of particle shape evolution [22]. First, the number of components, or in this case the number of reaction steps/number of particles, was assessed using SVD. Subsequently, the authors applied EFA to obtain an initial estimate of the concentration profiles of the possible particles present. Further analysis with MCR culminates in the attainment of the concentration and pure spectra profiles of the nanoparticles. Effectively, the analysis of concentration profiles retrieved by MCR-ALS revealed that it was possible to evaluate the shape of the nanoparticles during the synthetic process. Unfortunately, as there was an overlap between the concentration profiles of the formed species, the analysis of the pure spectral profiles retrieved by MCR-ALS was not feasible. Nonetheless, the proposed methodology can be envisioned as a valuable and simpler alternative to TEM and XRD measurements [22].

In the third work, the internal structure of 42 different QDs, composed of two chemical domains (CdS/CdSe) and arranged in 7 alloys and (core)shell structural classes, were accurately classified through Raman spectroscopy (Fig. 1a) followed by chemometric analysis employing multiple linear regression (MLR) after PCA [23]. The proposed strategy was able to predict the QY (within a range of up to 35 %) (Fig. 1b). The application of PCA before MLR was an expeditious way to circumvent the disadvantages of MLR, namely the need for a number of samples higher than the used wavelengths and the low degree of correlation or overlapping that should be displayed by the used data [28]. The obtained results revealed the potential of this approach as a straightforward, rapid, and non-destructive methodology for the

Table 1

Chemometric-assisted approaches for the characterization of QDs-based nanoparticles and for the optimization of its synthetic route.

Nanomaterial	Chemometric model	Experimental design	Synthetic route	Objective	Ref.
CdSe	MCR-ALS	n.a.	organometallic	particle size estimation	[24]
CdSe/ZnS	MCR-ALS	n.a.	organometallic	particle shape estimation	[22]
CdS:CdSe	PCA followed by MLR	n.a.	organometallic	internal structure	[23]
CdTe	n.a.	fractional factorial design and CCD	aqueous	synthesis optimization	[38]
AgInS ₂ /ZnS	n.a.	fractional factorial design and CCD	aqueous	synthesis optimization	[39]
CdTe	n.a.	two-level factorial design and Doehlert design	aqueous	synthesis optimization and size estimation	[25]
Mn-doped ZnS	n.a.	Decision table	aqueous	synthesis optimization	[40]

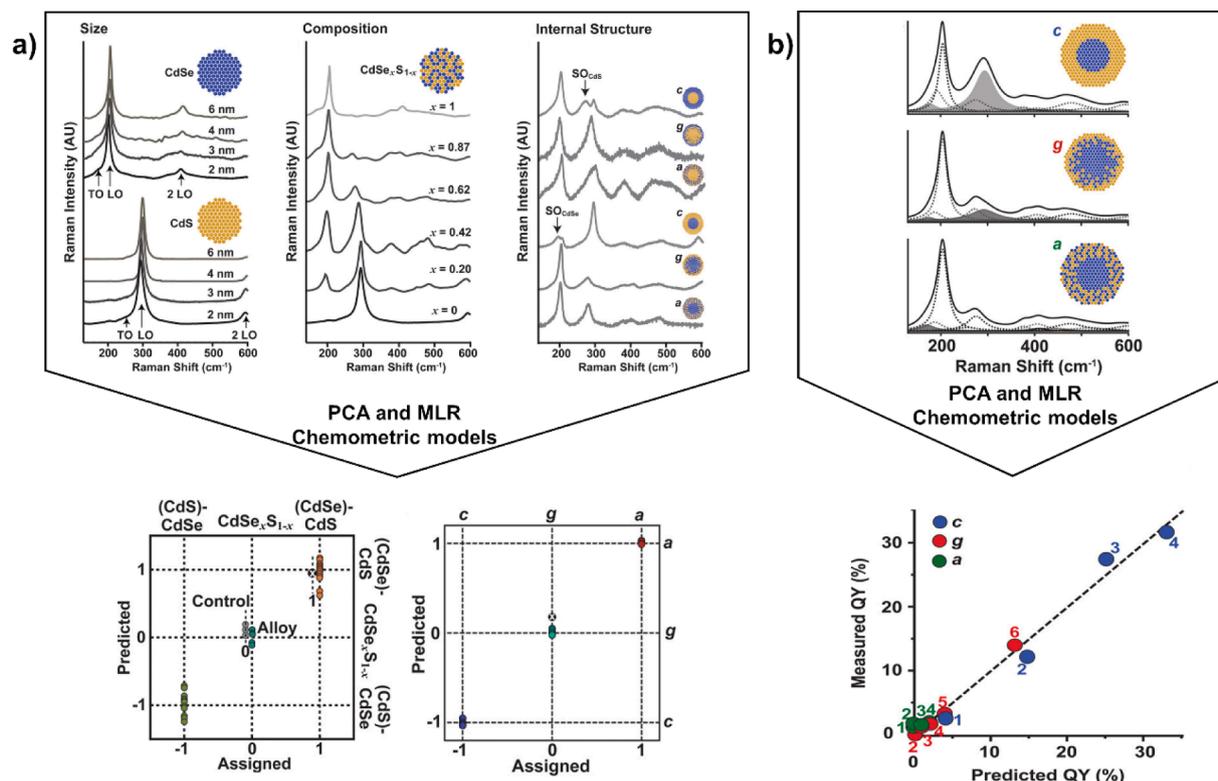


Fig. 1. Raman spectroscopic characterization of CdS:CdSe heterostructured nanocrystals and the corresponding chemometric analysis in order to predict the size, composition and internal structure a), as well as, the QY b). Adapted with permission from [23]. Copyright 2016 American Chemical Society.

monitoring of the structure and quality of the heterostructured nanocrystals during their synthesis. Moreover, it can enable the real-time adjustment of the nanomaterials structures and properties [23].

All these examples demonstrated that the use of chemometric tools for the analysis of experimental data obtained by vibrational spectroscopy, as well as by UV-Vis and fluorescence spectroscopy, allows to successfully characterize the nanomaterials in a simple and inexpensive way thus bypassing the need for advanced techniques that are usually more complex to handle, more expensive and not easily available to all research groups.

2.2. Optimization of the aqueous synthetic route and aqueous dispersed QDs characterization

Despite being the most used QDs synthetic route, the organometallic approach produces hydrophobic nanoparticles only soluble in organic media. Thus, for biological and clinical applications the QDs must be made water-soluble, which can be achieved by the modification of their surface. In alternative, QDs can be directly prepared via an aqueous synthetic route, which has evident advantages regarding the organometallic one since it is easier, faster, less expensive, more environmentally friendly and has a higher reproducibility [35,36]. In the aqueous

route, the QDs are passivated with thiol-based hydrophilic molecules which are not only more biocompatible but have also functional terminal groups that provide better chemical accessibility for further QDs surface functionalization. In addition, the obtained QDs can be precipitated, washed, and stored in the dry state under ambient conditions for longer periods of time while maintaining high stability.

The use of chemometric models in experimental design and multivariate optimization for the definition of a synthetic route is a valuable strategy to obtain highly luminescent, crystalline and monodisperse QDs, since the QDs synthesis is a multi-step process that requires strict control of all involved variables (Table 1). The key factor to obtain the ideal synthesis procedure is to uncover how the experimental variables affect the required response (i.e., QDs properties) and thus to outline the best conditions for the intended aim. This optimization is usually carried out by using the univariate approach which involved to study all experimental variables, individually, regardless of all remaining variables that are kept constant throughout the assays. However, this approach has major drawbacks, not only because it requires many runs to obtain an optimal response, but also because it does not consider the interaction between variables. In this sense, multivariate optimization methods have been regarded as advantageous alternatives since they study the influence of all experimental variables, either alone or upon

combined interaction effects, on the synthesis process, with a certain level of statistical reliability. Multivariate optimization is usually carried out in two steps: first, a screening experimental design is implemented aiming not only at evaluating and quantifying the effect of each individual experimental variable on the QDs optical properties but also to assess possible interactions between the involved variables; second, the most significant experimental variables are fine-tuned throughout the use of a suitable experimental design that fits a quadratic model to the data in order to identify the combination of optimum variables' levels. This strategy was applied by our research group for the development of one-stage synthetic route, assisted by microwave irradiation, for the aqueous synthesis of binary CdTe [38] and ternary AgInS₂/ZnS QDs [39]. The synthesis parameters, namely the precursors' relative molar ratio and the solutions' pH, were firstly screened by a two-level fractional factorial design to uncover those with the most significant influence on the QDs PL properties. Subsequently, a central composite design (CCD) was applied to obtain response surfaces for the quadratic model between the PL properties of the synthesized QDs and the parameters under study, allowing to uncover the optimum experimental conditions. For the optimization of binary CdTe QDs synthesis, QY was the outcome used to appraise the quality of the QDs and to design the mathematical model [38]. In the case of ternary AgInS₂/ZnS QDs, the maximum emission wavelength, QY, PL lifetime, and the elemental composition in terms of Ag:In and Zn:In ratios, were considered the best features to evaluate the QDs quality. Consequently, five models were predicted, one for each outcome of the synthesis process [39]. In another work, the aqueous synthesis of CdTe QDs using a conventional hydrothermal approach was also optimized through an experimental design [25]. Likewise the abovementioned works, a two-level factorial design was initially used to establish the significance of each of the synthesis parameters (i.e., temperature, pH, reaction time and precursor molar ratios). After that, a Doehlert design was applied to relate the most significant parameters (pH, temperature and reaction time) with the maximum absorbance wavelengths, allowing to predict the CdTe QDs diameters which were posteriorly successfully compared with high resolution (HR)-TEM and XRD measurements of the nanoparticles size [25]. Yang et al. proposed a Decision Table (DT) in Rough Set as a new chemometric approach for the optimization of the synthesis of Mn-doped ZnS QDs [40]. The DT was designed for the identification (and posterior reduction) of the synthesis variables that might be considered as non-critical based on the analysis of the observed PL intensity, which was used as a decision factor. After reducing the crucial experimental parameters from five to two, a second DT was applied to appraise the optimal synthesis strategy which was also able to pinpoint the core attribute as the most important parameter. The obtained results were then compared with those attained by using single factor analysis and orthogonal experiment, which confirmed these optimal experimental conditions. Nevertheless, the DT method allowed to achieve the optimal conditions with a lower number of experimental runs [40].

To conclude, characterizing nanocrystals usually involves complex techniques, but simpler and cheaper analytical tools like spectrofluorimeters, combined with chemometric models, offer a cost-effective alternative. By applying chemometric models to spectroscopy data, nanomaterial properties can be assessed without the need for advanced and expensive methods. In the synthesis of quantum dots (QDs), where precise control of variables is crucial, chemometric models play a key role. Unlike traditional univariate approaches that require many runs to achieve optimal responses, multivariate optimization considers variable interactions, providing a statistically reliable and efficient approach to achieve highly desirable QD properties. This strategic use of chemometric tools streamlines the characterization process and proves essential in optimizing QD synthesis for enhanced luminescence and uniformity.

Despite the clear advantages arising from using multivariate optimization models for the design of efficient and reproducible synthetic routes to obtain nanomaterials with excellent quality, in terms of high

luminescence, crystallinity and monodispersity, the use of the univariate approach is still predominant.

3. Application of chemometric models to analyse QDs-based PL analytical data

Although the use of chemometric tools regarding QDs synthesis is still scarce, their use in QDs-based PL analytical methodologies is increasing. This occurs because they allow to simultaneously relate unselective multiple instrumental signals with single or multiple analytes concentrations [41]. The acquired instrumental data for QDs-based PL methodologies can be classified according to their complexity. In 1994, Booksh and Kowalski categorized the analytical data, in terms of increasing complexity, as zeroth-order, first-order, second-order, and higher-order data, depending on the characteristics of the instrument or method that provided it [42]. It is expected that by increasing the complexity of the data collected per sample analysis, it would be possible to obtain more resolved quantitative or qualitative analytical estimations, with a higher selectivity [43].

Zeroth-order data are obtained when the instrument produces a single output per sample, such as a fluorescence intensity, at a single wavelength, per analyte concentration (Fig. 2a). This means that each sample is characterized by a single numeric value. For analytical purposes, it is important to understand that zeroth-order data, mostly used in univariate calibrations, require a full selectivity towards the analyte under determination and can only be used in the case of samples with known composition and without interfering species [44].

First-order data are obtained as a vector data for sample, which corresponds, for example, as to acquire a fluorescence emission spectrum at a fixed excitation wavelength (Fig. 2b). This spectrum corresponds to a set of intensity values at different wavelengths, forming a vector of data. The analysis of first-order data with suitable first-order multivariate calibration methods allows to circumvent the lack of selectivity upon the application of efficient mathematical algorithms. Effectively, this is acknowledged as the first-order advantage, which states that it is possible to quantify the analyte in samples with known interferences as long as those interferences are included in the standard solutions used in the calibration process. Hence, unexpected constituents can represent a complicated issue to handle with first-order data only [44,45].

In the case of second-order data, a data matrix is obtained per sample (Fig. 2c). Considering analytical methodologies using QDs as PL probe, this type of data can be obtained in two ways: i) by using a spectrofluorometer able to generate an EEM or by recording the evolution of the sample PL spectrum, at a fixed excitation wavelength, throughout time (Fig. 2c) (e.g. the kinetics of the QDs/sample interaction) [46]; ii) by resorting to instruments hyphenation using multi-block data analysis techniques. These allow to collect data of different nature from multiple instrumental modes, which contain extractable complementary information [41,47]. The exploration of second- and higher-order advantages allows to monitor the analyte even in the presence of unidentified interferences that were not included in the calibration set. This is known as the second-order advantage, which enables to circumvent the presence of unpredicted components in complex matrixes, such as biological, environmental or food samples. This represents a major benefit under an analytical perspective since the calibrated analytes can be selectively determined in the existence of uncalibrated constituents.

3.1. QDs -based analytical methodologies using first-order data

As abovementioned, the methodologies relying on first-order data allow the determination of the target analytes with increased selectivity when compared to those using zeroth-order data. Nevertheless, to be fully effective, these methods need to include the interfering species in the calibration step. This is only possible in samples with known and simple matrix composition.

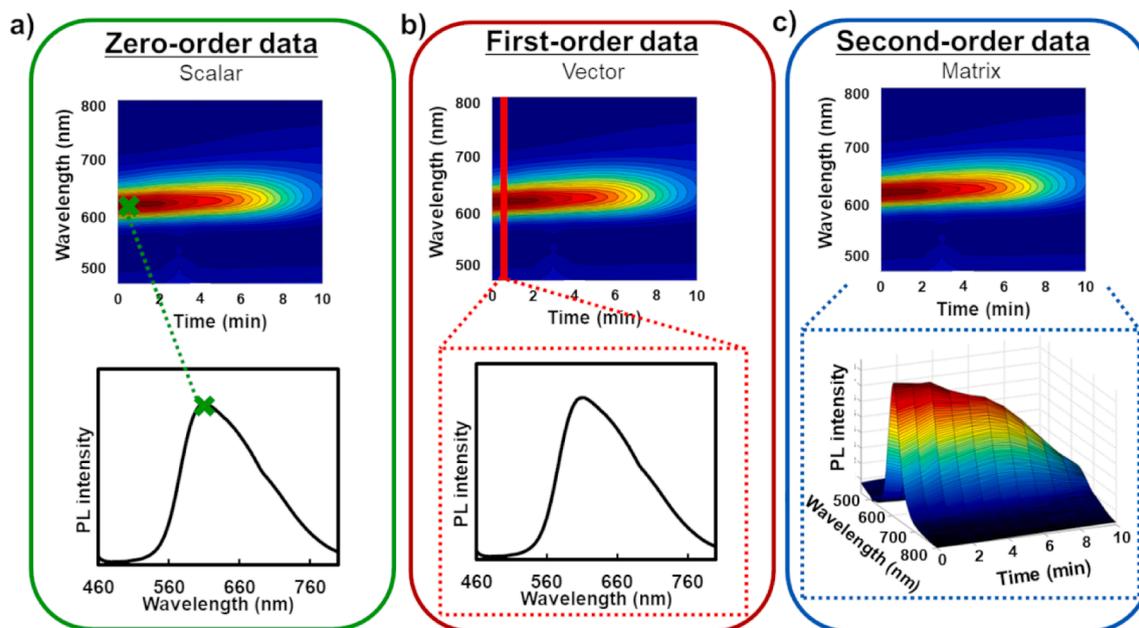


Fig. 2. Schematic representation of the different data structures that can be obtained for a sample in PL-based methodologies. Adapted with permission from [48]. Copyright 2023 MDPI.

The selection of both the most suitable chemometric model and the most appropriate modulation strategy to employ QDs as fluorescence probes is very important to guarantee the objectives of the analysis. In this sense, different strategies have already been explored to obtain an effective QDs-based sensing platform, namely:

- i) the use of single-emitter QDs which, upon interacting with the analyte, have their optical properties modulated (Fig. 3a). The interaction, which can generate different PL responses, is essentially determined by the QDs surface chemistry. Specifically, the reactivity of the nanocrystal towards the analyte depends on the different terminal functional groups of the capping ligand used to passivate the QD's surface. This way, the optical properties can be modulated either by changes in the PL intensity (enhancing or quenching) or by deviations in the maximum emission wavelength (hypsochromic or bathochromic effect).
- ii) the use of a multi-emitters probe resulting from the combination of different multicoloured fluorophores (QDs and other molecules) with distinct affinity towards the analyte. In this case, during the interaction process between probe and analyte, the involved PL emitters can respond individually, producing a specific analyte-response profile (Fig. 3b).
- iii) the use of a cross-reactive sensing array, composed of multiple individually arranged single emitters, to monitor multiple analytes. This sensor configuration returns a collective response of non-selective interactions between each single emitter probe and the target analytes (Fig. 3c). Likewise, the previous example, the emitters could be either QDs combined with other QDs exhibiting distinct reactivities towards the analytes, or QDs combined with fluorophores of different nature, such as organic dyes and fluorescent plasmonic nanoparticles.
- iv) the use of a cross-reactive sensing array composed of an arrangement of multiple multi-emitters probes. Each of these can be made of an assortment of QDs, with adjusted reactivity, or of a combination of QDs and other fluorophores (Fig. 3d).

The data collected from the abovementioned strategies can be analysed by proper chemometric models, which allows the extraction of useful analytical information seeking either authentication/

discrimination, identification, or quantification of analytes in real samples. The most used chemometric tools for authentication/discrimination purposes in QDs-based analytical methodologies are LDA, PLS-DA and SIMCA. Regarding identification, MCR is usually the best option as it enables the recovery of the spectral profiles of all species present in the sample. For quantification purposes, the most used chemometric tools are PCR, PLS and MCR-ALS. However, when the acquired data reveals a non-linear relationship, ANN and SMV are the best options both in terms of authentication/discrimination and quantification purposes.

3.1.1. First-order analysis for discrimination purposes

Most of the available analytical methodologies for discriminating samples from a characteristic group, seek the identification of different substances/species or the presence of an adulteration (Table 2). The selection of the sensing platform and of the most suitable chemometric model is crucial. As abovementioned, the most suitable chemometric models for authentication/discrimination purposes are PLS-DA and SIMCA, being MCR-ALS a very good option when the objective of the work is the identification and quantification of the species (through spectral profile recovery). An alternative chemometric model entitled data-driven soft independent modelling of class analogy (DD-SIMCA), which is one of the most suitable models for the discrimination between adulterated and authentic samples, should be also mentioned. This model works similarly to SIMCA but introduces a more robust estimation of two relevant statistics (namely the orthogonal and score distances), offering not only a friendly design for the acceptance area (thresholds) but also the possibility of selecting different thresholds in addition to the calculated one. Pertinent supplementary information regarding this chemometric model could be found elsewhere [53–56].

3.1.1.1. Single emitter probe as sensing strategy. Carbon quantum dots (CDs) were explored as a single-emitter probe, in combination with LDA, in two distinct analytical methodologies for the discrimination of five pesticides [57] and five food additives [58] in food samples (Table 2).

In the first case, the identification of the pesticides propanyl, parathion, dimethoate, chlorpyrifos and pirimicarb, was based on a FRET process using the CDs as donors and non-fluorescent silver nanoparticles (AgNPs) passivated with polyacrylate (PAA) and polyethyleneimine

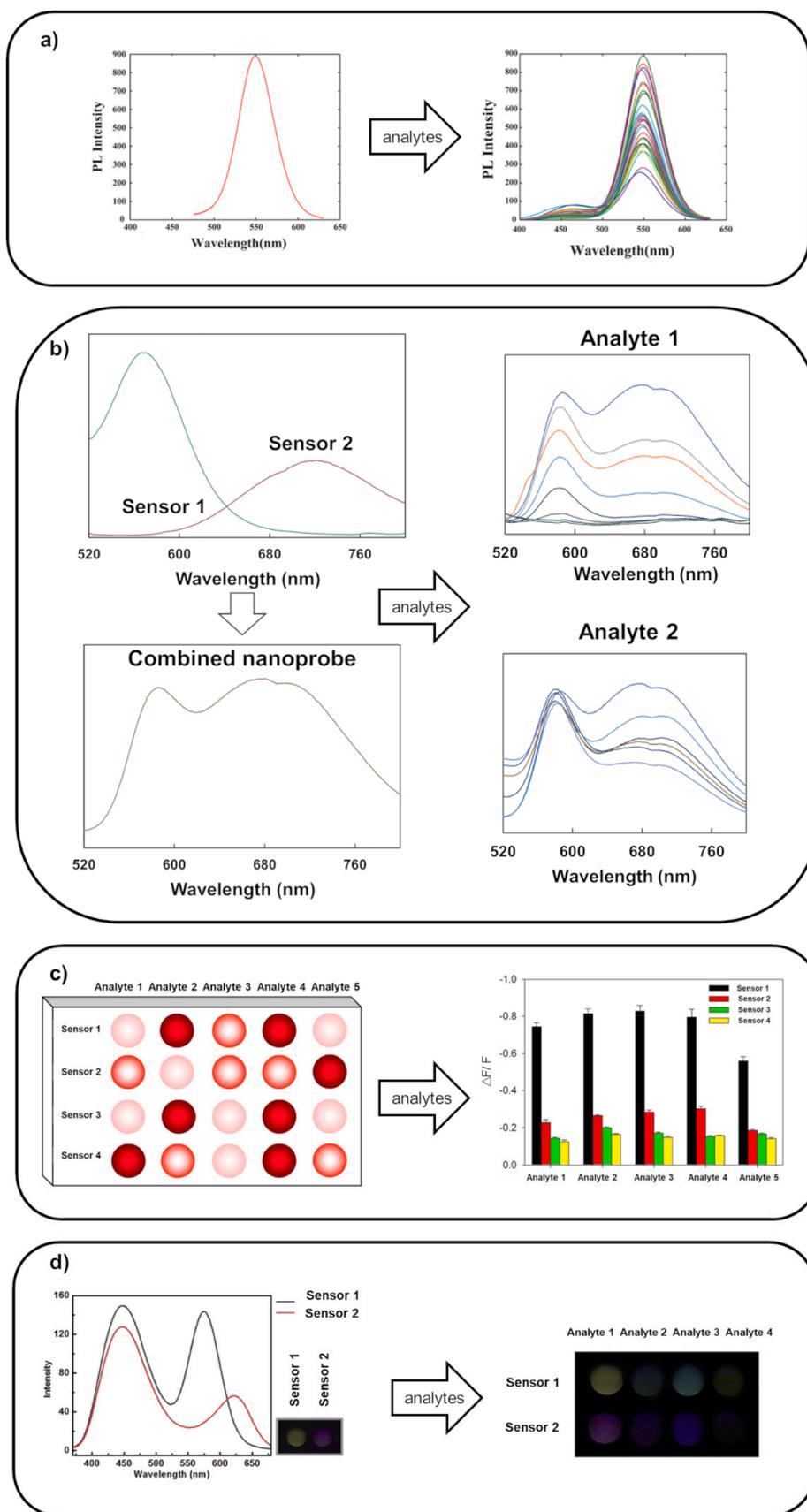


Fig. 3. Different strategies to obtain efficient QDs-based sensing platforms: a) Single emitter nanoprobe; b) multi-emission combined nanoprobe; c) cross-reactive sensing array composed of multiple individually arranged single emitters; and d) cross-reactive sensing array composed of an arrangement of multiple multi-emitters probes. Adapted with permission from [49,50,51] and [52]. Copyright 2014 American Chemical Society, 2017 Elsevier, 2019 Elsevier and 2020 Elsevier.

Table 2
Analytical approaches exploring the use of first-order data processed by chemometric tools.

Analyte	Sensing platform	Sensing strategy	Objective	Chemometric model	PL signal	LOD	Ref
propranolol, parathion, dimethoate, chlorpyrifos and pyrimicarb	CDs and AgNPs	Single emitter probe	discrimination	LDA	Quenching	250 ng mL ⁻¹	[57]
citric acid, lactic acid, ascorbic acid, sodium benzoate and potassium sorbate	CDs	Single emitter probe	discrimination	LDA	Quenching	252 ng mL ⁻¹	[58]
6 Jujube species	CDs	Single emitter probe	discrimination	PCA, BDA	Quenching	n.a.	[60]
Dianhong black tea	Co ²⁺ -modified CDs	Single emitter probe	discrimination	PCA, BDA	Quenching	n.a.	[62]
29 green teas	NALC-CdTe	Single emitter probe	discrimination	PLSDA	Quenching	n.a.	[49]
53 green teas	NALC-ZnCdSe@NALC-CdTe	Multi emitter probe	discrimination	PLSDA	Quenching	n.a.	[63]
Orange juice adulteration	NALC-ZnCdSe@NALC-CdTe	Multi-emitter nanoplatfom	discrimination	OCPLS	Quenching	5 % (w/w)	[64]
kiwifruit juice adulteration	NALC-ZnCdSe@NALC-CdSe	Multi-emitter nanoplatfom	discrimination	OCPLS	Quenching	2 % (w/w)	[65]
Herbal honey adulteration	NALC-ZnCdSe@NALC-CdSe	Multi-emitter nanoplatfom	discrimination	OCPLS	Quenching	1 % (w/w)	[66]
Dimethoate, dichlorvos and demeton	NALC-CdTe@ NALC-ZnCdSe	Multi-emitter nanoplatfom	discrimination	PLSDA	RGB colour change	n.a.	[67]
amikacin	Schiff base@ Rhod/SiO ₂ @Fluor/SiO ₂ @- MNDC-CdSe/CdS/ZnS	Multi-emitter nanoplatfom	discrimination	PCA, MDA	Both quenching and enhancing	n.a.	[68]
Adenine, Guanine, Cytosine, Uracil and Thymine	CA-CdTe NALC-CdTe DMAE-CdTe TGA-CdTe	Cross-reactive sensing array	discrimination	LDA and HCA	Quenching	0.34–0.67 mmol/L	[51]
Cu ²⁺ , Hg ²⁺ , Ag ⁺ and Cd ²⁺	NALC-Mn-ZnS Cit-Mn-ZnS MPA-Mn-ZnS (NH ₄) ₃ L-Mn-ZnS	Cross-reactive sensing array	discrimination	PCA and LDA	Quenching	0.3–2.7 pg mL ⁻¹	[69]
22 Baijiu liquors	MPA-CdTe, MPA-CdSe and GSH-Cu:CdS	Cross-reactive sensing array	discrimination	PCA, LDA, HCA and RBFN	Both quenching and enhancing	n.a.	[70]
Liquors of Baijiu	NALC-CdTe NALC-ZnSe NALC-ZnCdSe	Cross-reactive sensing array	discrimination	PCA, PLS-DA	RGB colour change	n.a.	[71]
Cu ²⁺ , Fe ³⁺ and Hg ²⁺	3 distinctly synthesized CDs with different N precursors (GlyCDs, LysCDs and SerCDs)	Cross-reactive sensing array	discrimination	LDA, HCA	Quenching	10 μmol L ⁻¹	[72]
Ag ⁺ , Cd ²⁺ , Cr ²⁺ , Fe ³⁺ , Hg ²⁺ and Pb ²⁺	7 distinctly synthesized CDs with different C and N precursors	Cross-reactive sensing array	discrimination	PCA, MDA	Both quenching and enhancing	>16 μmol L ⁻¹	[73]
Al ³⁺ , Cu ²⁺ , Co ²⁺ , Ni ²⁺ , Mg ²⁺ , Pb ²⁺ , Ba ²⁺ and Ca ²⁺	Pdots DG- Pdots NPG-Pdots MG- Pdots	Cross-reactive sensing array	discrimination	LDA	Both quenching and enhancing	1 mg mL ⁻¹	[75]
Ag ⁺ , Hg ²⁺ , Cu ²⁺ , Pb ²⁺ , Cr ³⁺ , Mn ²⁺ and Cd ²⁺	MPA-CdTe GSH-CdTe PMAA-AgNCs BSA-AuNCs RHD CB	Cross-reactive sensing array	discrimination	LDA	Both quenching and enhancing	n.a.	[76]
Fe ²⁺ , Fe ³⁺ , Cu ²⁺ , Co ²⁺ , Ni ²⁺ , Mn ²⁺ , Cd ²⁺ , Ca ²⁺ and Ag ⁺	MPA-Mn-ZnS TG-Mn-ZnS	Cross-reactive sensing array	discrimination	LDA, HCA	Both quenching and enhancing	<0.5 μmol L ⁻¹	[77]
GSSG, GSH, CYS and CA	Rhodamine B@NALC-CdTe CDs@NALC-CdTe	Cross-reactive sensing array of multi-emitter nanoplatfom	discrimination	HCA and LDA	Colour modulation	0.021, 0.028, 0.018 and 0.08 μmol L ⁻¹	[78]
Hg ²⁺ , Pb ²⁺ , Cd ²⁺ , Fe ³⁺ and Cu ²⁺	Rhodamine B@EDA-CDs Rhodamine B@urea-CDs Rhodamine B@Gly-CDs	Cross-reactive sensing array of multi-emitter nanoplatfom	discrimination	HCA and LDA	Colour modulation	n.a.	[79]
TNT, TNP, DNT	CDs@g-TMA-CdTe CDs@r-TMA-CdTe	Cross-reactive sensing array of multi-emitter nanoplatfom	discrimination	LDA	Quenching	5.0 μmol/L	[52]
Metolcarb, carbofuran and carbaryl	NALC-CdTe	Single emitter probe	discrimination and quantification	PLS-DA, PLS	RGB colour data	0.97 μg L ⁻¹ 0.89 μg L ⁻¹ 0.78 μg L ⁻¹	[80]

(continued on next page)

Table 2 (continued)

Analyte	Sensing platform	Sensing strategy	Objective	Chemometric model	PL signal	LOD	Ref
Dipterex, dursban, paraquat, methyl thiophanate and cartap	NALC-ZnCdSe@NALC-CdSe	Multi-emitter nanoplatform	discrimination and quantification	MWPLSDA and PSO-OWLS-SVM	Quenching	2×10^{-8} mol/L	[81]
Chinese baijiu 5 organic acids	TGA-CdTe GSH-CdTe NALC-CdTe CA-CdTe	Cross-reactive sensing array	discrimination and quantification	LDA PLS	Both quenching and enhancing	n.a.	[86]
Hg ²⁺	MPA-CdTe@MPA-CdTe	Multi-emitter nanoplatform	quantification	PLS	Quenching	n.a.	[87]
Cu ²⁺	MPA-CdTe@MPA-CdTe	Multi-emitter nanoplatform	quantification	PLS	Quenching	n.a.	[88]
Cu ²⁺ , Hg ²⁺ , Pb ²⁺	MPA-CdTe@MPA-CdTe	Multi-emitter nanoplatform	quantification	ANN and PLS	Quenching	1.9 $\mu\text{mol L}^{-1}$ and 4.1 $\mu\text{mol L}^{-1}$	[50]
FA, Fe ²⁺	CYS-CdTe@MPA-AgInS ₂	Multi-emitter nanoplatform	quantification	ANN and PLS	Quenching	2.01 and 1.88 $\mu\text{mol L}^{-1}$	[89]
TNT, 4-NP	CDs@CYS-CdTe	Multi-emitter nanoplatform	quantification	ANN and PLS	Quenching	2.01 and 1.88 $\mu\text{mol L}^{-1}$	[89]

(bPEI) as acceptors [57]. The sensing scheme relied on the displacement of the pesticides by the positive and negative surface-charged AgNPs, affecting the CDs photoluminescence. Combining this sensing strategy with the use of LDA for data analysis allowed to maximize pesticides discrimination because of the different fluorescence response patterns that each pesticide individually provided [57].

The second work was based on the interaction of a single emitter CDs sensing platform with different additives (ascorbic acid, lactic acid, citric acid, potassium sorbate and sodium benzoate), being the generated fluorescent data subsequently analysed through LDA [58]. Since the nanocrystals exhibited a distinct reactivity towards each target analyte, their respective PL response was unique (fingerprinting). This fact generated different response patterns for each analyte which were successively discriminated by applying LDA [58].

When using single-emitter QDs for discrimination purposes it is important to guarantee that the probe reactivity differs significantly for each of the analyte/sample under evaluation. The CDs reactivity depended on their surface chemistry, which can be fine-tuned by using distinct solvents or co-dopers during the synthetic route. In effect, CDs doping with heteroatoms is an efficient strategy to improve their optical and physicochemical properties, as well as to modify their electronic structures, tailor-making their reactivity to allow their application as suitable sensing platforms [59]. Additionally, the use of different solvents in the CD's synthesis affects the dehydration and carbonization process, which also influences the quantum size, oxidation degree and concentration of the nanomaterials, and, consequently, their reactivity [60].

By exploring the later example, Liu et al. used four different solvents (water, dimethylformamide, ethanol and formamide) for the CDs' synthesis. Distinctly coloured carbon-based nanomaterials were obtained which were later used for the discrimination of distinct jujube species [60]. The authors verified that in the presence of different metal ions, the PL intensity of all distinctly coloured CDs was progressively weakened, and that the highest quenching effect was observed in the presence of Fe³⁺. Moreover, they also estimated that the highest values for classification accuracy and cross-validation accuracy were obtained when using red-emitting CDs. Consequently, Bayesian discriminant analysis was employed for processing the analyte-induced CDs PL modulation data. Bayesian discriminant analysis (BDA) is similar to LDA: in LDA, the test samples are classified according to the lowest distance to each category, while in BDA the test samples are classified according to the lowest probability of misclassification. This means that BDA assumes a prior probability for each group, which is calculated using a probability density function [61]. The obtained results demonstrated that it was possible to distinguish and accurately predict jujube species according to their metal ions contents.

CDs were also used, along with BDA, in another analytical method

for the discrimination of 8 different Dianhong black tea grades [62]. The authors compared the discrimination efficiency resulting from doping N-CDs with Fe³⁺ and Co²⁺, and by using two different solvent extracts (water and ethanol). It was observed that the Co²⁺-modified CDs allowed to obtain higher values of classification accuracy when compared to Fe³⁺-modified CDs. In terms of the extraction solvent, ethanol yielded the best results. A total of 95 % of correct predictions were attained through cross-validation using Co²⁺-modified CDs and ethanol as solvent. The discrimination of the different teas was based on their polyphenols content. As polyphenols caused the PL quenching of the CDs, distinct responses were obtained for each Dianhong black tea grade depending on their composition [62].

As above discussed, one can expect that by increasing the amount of data collected per sample the qualitative and quantitative performance of the analysis would be also improved. The advantage of employing a multi-emitter probe, as an alternative of a single emitter, could be illustrated with two distinct works which aimed at the identification of green teas using binary and ternary QDs as PL probes and PLS-DA as chemometric model for data analysis [49,63]. When using N-Acetyl-L-cysteine (NALC)-CdTe as a single emitter, it was possible to discriminate 29 different green teas [49] while when using NALC-capped ZnCdSe and CdTe QDs, as a combined dual-emitter nanoprobe, the number of differentiated teas increased to 53 [63]. Due to the distinct composition and concentration of the diverse constituents of each green tea (namely amino acids, flavonoids and catechins), a different PL quenching magnitude was observed. In this sense, the addition of a second PL nanoprobe (NALC-ZnCdSe QDs) provided additional spectral information, due to their different reactivity, thus improving the classification accuracy of the used chemometric model.

3.1.1.2. Multi-emitter nanoplatform as sensing strategy. The amount of collected data can be also increased by designing multi-emitter sensing probes, combining multiple QDs, or QDs and other fluorophores, that show maximum PL emission at complementary wavelengths covering a broad spectrum.

Effectively, two, three, or, at most, four spectrally resolved QDs can be combined in the same sensing platform without spectral overlapping. This way, the selection of QDs with distinct sizes, compositions, arrangements, and/or surface chemistry may confer to each combined nanoprobe a specific pattern response towards a selected analyte. The same research team developed three similar works in which one-class partial least squares (OCPLS) was used as chemometric model and NALC-ZnCdSe and NALC-CdSe QDs were combined in a multi-emitter nanoprobe. The OCPLS is similar to SIMCA, as it is considered as a class modelling technique, although based on PLS, instead of PCA, for data reduction. Moreover, the components on OCPLS are not ascribed on explaining data variances, as in SIMCA. In terms of performance, it is

similar to SIMCA [34]. The dual-emission sensing platform was applied in the authentication of orange [64], kiwifruit juice [65] and honey [66] demonstrating to have huge potential for a sensitive and rapid screening of adulterants in complex samples (Table 2). In fact, for the authentication of orange juice regarding the presence of two adulterants (sucrose syrup 5 % w/w, and artificial fruit powder 2 %, w/w), a sensitivity and specificity of 97.8 % and 77 %, respectively, were obtained considering the best model. For the authentication of kiwi juice regarding the same adulterants (but both at 2 %, w/w), a sensitivity and specificity of 92.9 % and 83 %, respectively, were obtained. For the authentication of honey regarding the presence of four adulterants (glucose, sucrose and fructose syrups, and glucose-fructose syrup, all at 1 % w/w), a sensitivity and specificity of 94.9 % and 88.7 %, respectively, were obtained.

Wang et al. [67] explored a dual-emission sensing nanoplatform to discriminate the presence of three organophosphorus pesticides, demeton, dichlorvos and dimethoate. NALC-capped CdTe and ZnCdSe QDs were used as PL nanoprobes in combination with Zn-nanoporphyrin (nano-ZnTPyP), implementing a “turn-off-on” detection scheme. In the presence of nano-ZnTPyP, the PL intensity of the combined nanoprobe decreased via FRET process. Upon addition of the three pesticides, the PL emission was recovered, due to the affinity of the electron-rich group of these pesticides towards the positively charged dodecyl trimethyl ammonium bromide (DTAB), at the surface of nanoporphyrins, establishing electrostatic interactions that detached the nano-ZnTPyP from QDs surface, impairing the FRET process with the consequent PL recovery. The presence of the three pesticides generated different colour changes. The modulation of the RGB values was then analysed by PLS-DA, which allowed an accurate discrimination of the pesticides (100 % of correct predictions considering the test set) in complex sample matrices (apple and cabbage) [67]. The use of QDs-based visual sensing approaches has emerged, in the last years, as a promising and highly attractive alternative since it enables to monitor a sample, faster and *in-situ*, while avoiding expensive and complex instrumentation [11]. The use of a smartphone allows not only the acquisition of high-quality images but also the use of processing software (APPS) capable of converting the captured colour into RGB values. In addition, the use of the most suitable chemometric model enables the discrimination and quantification of chemical species with high accuracy and sensitivity, maintaining the portability of the naked-eye determination [10,11].

Divyanin et al. [68] developed a multi-emitter nanoprobe made of different combinations of up to five emitters (organic dyes and quantum dots), which were applied in the identification of mixtures of active pharmaceutical ingredients (amikacin, sulfamethoxazole, piracetam and chloramphenicol) through PCA. The most efficient nanoprobe was a four-emitter probe comprising a Schiff base, rhodamine (Rhod/SiO₂) and fluorescein (Fluor/SiO₂) attached to SiO₂ and mercapto-n-dodecanoate (MNDC)-capped CdSe/CdS/ZnS. With this four-emitter combined probe, the synthetic mixtures of 2 or 3 compounds clustered accordingly, demonstrating the possibility of applying this strategy in multiplexed detections in complex real samples [68].

3.1.1.3. Cross-reactive sensing array as sensing strategy. As mentioned before, when increasing the amount of PL data acquired, the analytes can be identified and discriminated with enhanced figures of merit. A strategy to enhance the available PL data information is through the use of multiple non-selective PL probes, assembled in a cross-reactive sensor array, and showing different responses towards the distinct analytical targets. In most cases, the sensing elements consisted of a combination of nanomaterials of the same nature, usually semiconductor QDs, but differing on composition, size, and surface chemistry. The latter is usually fine-tuned by using distinct thiol-based ligands containing carboxylic and amine functional moieties, which provide them with a distinguishable reactivity. To illustrate this strategy, two similar works can be highlighted [51,69] (Table 2). In the first work [51], four CdTe QDs passivated with different capping ligands, including NALC,

thioglycolic acid (TGA), cysteamine (CA) and 2-dimethyl-aminoethanethiol (DMAE) were used to distinguish among different nucleobases (cytosine, adenine, uracil, guanine and thymine) resorting to LDA and hierarchical cluster analysis (HCA) as chemometric methods for PL data analysis. In the presence of the nucleobases, an exciton energy transfer effect (EET) occurs leading to PL quenching via QDs aggregation. Moreover, the authors showed that the same sensing array could be used to discriminate other five rare bases, demonstrating the higher versatility of these sensing strategies [51]. In the second work, Mn-doped ZnS QDs passivated with NALC, triammonium-N-dithiocarboxyiminodiacetate ((NH₄)₃L), mercaptopropionic acid (MPA) and citric acid (Cit) were also used as a cross-reactive sensing array to differentiate 4 distinct metal ions in different mixtures [69]. The cross-reactive responses of the sensing array were analysed by LDA allowing to distinguish the metal ions at different concentration levels, in binary, ternary and quaternary mixtures [69].

Likewise different cappings were used to assure different reactivities, the combination of nanomaterials with distinct core compositions was also investigated. A cross-reactive sensing array composed of MPA-CdTe, MPA-CdSe and glutathione (GSH)-Cu:CdS QDs, was developed for the differentiation of 22 Baijiu liquors [70]. Upon interaction with the distinct liquors, the QDs exhibited dissimilar responses (either PL quenching or enhancing) due to the varied composition of the Baijiu samples, namely in terms of ethanol, ethyl acetate, organic acids, nitrogen and sulfur compounds content. The diverse PL responses were ascribed not only to the different terminal functional groups of the capping ligand at the QDs' surface (carboxylic acid and amines) but also to the distinct chemical composition of their core. The use of this cross-reactive sensing array allowed the correct discrimination of 22 Baijiu liquors through LDA, with 91 % of correct predictions considering the validation set [70].

In another sensing strategy, tricolour QDs with different core compositions, which include blue-emission ZnCdSe QDs, yellow-emission CdSe QDs and red-emission CdTe QDs, all passivated with the same capping ligand (NALC), were employed in the design of a paper sensor array [71]. By using this colorimetric array, the resultant RGB values were processed by PCA and PLS-DA models, enabling a successful visual discrimination of the evaluated samples. The authors observed that the higher the number of sensing elements of the cross-reactive sensing array, the higher the classification accuracy for the training and prediction samples. In fact, classification values (%) for the training/prediction sets of 96.85/85.71, 98.85/88.89 and 100/100 were obtained when using one, two and three sensing elements, respectively [71].

Cross-reactive sensing arrays that combined CDs with distinct surface chemistry were also investigated. CDs' surface chemistry can be tailored by resorting to different precursors during the synthesis. In two different works, CDs synthesized with distinct C and N precursors were used as PL sensing platforms for the differentiation of metal ions [72,73]. In the first work, nitrogen-doped carbon dots synthesized with different nitrogen precursors, namely L-lysine (Lys), L-glycine (Gly) and L-serine (Ser), and Cit as carbon precursor, were used in a sensing array that provided a fingerprinting pattern response for each metal ion [72], as a consequence of specific PL quenching magnitudes. The array response for each metal ion, at different concentrations, was analysed by HCA and by LDA, enabling the discrimination with 100 % of accuracy for three distinct metal ions (Cu²⁺, Fe³⁺ and Hg²⁺) in a ternary mixture considering the validation set [72]. In the other example [73], seven CDs synthesized with distinct C and N precursors were combined in a sensing array for the discrimination of six metal ions resorting to Mahalanobis distance analysis (MDA) and PCA. The authors observed that with the use of PCA, the initial array could be simplified to just a two-probe array without decreasing accuracy. This means that using only two sensing elements with dissimilar reactivities was enough to obtain fingerprints (response patterns) for the metal ions under evaluation, allowing an efficient discrimination. This CDs-based array was able to distinguish the metal ions even in the analysis of complex sample matrices such as

fetal calf serum or tap water [73].

Another organic water-dispersed nanoparticle that recently emerged due to its high photostability and brightness is the semiconducting polymer dots (Pdots), which can also be used as efficient analytical sensing elements [74]. Li et al. explored four Pdots sensing elements in a cross-reactive array for the discrimination of various metal ions (Al^{3+} , Ba^{2+} , Ca^{2+} , Co^{2+} , Cu^{2+} , Mg^{2+} , Ni^{2+} and Pb^{2+}) and brands of commercial water [75]. Pdots with and without the doping by three different gallates (methyl gallate – MG, dodecyl gallate – DG and n-propyl gallate – NPG) were assayed. The gallates generated different binding abilities towards the metal ions, yielding mixed fluorescence responses. These PL results allowed, by using LDA, to accurately discriminate eight metal ions, as well as a validation set of 24 samples of bottled water, with an accuracy of 100 % [75].

Other sensing schemes to enhance the cross-reactive array's sensitivity and discriminative capacity, could be explored. A possible strategy is to combine not only QDs with different reactivities towards the analytes but also to include fluorophores of a different nature, such as conventional organic dyes and their derivatives, or even plasmonic nanoparticles (i.e., fluorescent gold and silver nanoparticles). Kang et al. [76] developed a cross-reactive sensing platform wherein MPA and GSH capped CdTe QDs, bovine serum albumin (BSA) modified gold nanoclusters (AuNCs), calcein blue (CB), polymethacrylic acid (PMAA) modified silver nanoclusters (AgNCs) and a rhodamine derivative (RHD) were applied as PL sensing platforms [76]. The combination of different types of fluorophores, which were also functionalized with various reactive groups, allowed to obtain a wider assortment of PL responses. Effectively, the observed PL enhancing or quenching was analysed

through LDA, allowing the discrimination of seven heavy metal ions (Cd^{2+} , Mn^{2+} , Cr^{3+} , Ag^+ , Cu^{2+} , Pb^{2+} and Hg^{2+}) demonstrating the superior distinguishing capability of the proposed sensor array, except for Pb^{2+} and Cr^{3+} which were not differentiated [76].

An expeditious strategy to increase the available data information and, consequently, to enhance the cross-reactive array performance, is the use of selected molecules as modulators. Jing et al. [77] developed a sensing array using arginine (Arg) and glutamine (Gln) as modulators to increase the diversity of PL responses obtained for each analyte. MPA and alpha-thioglycerol (TG)-capped Mn-ZnS QDs were mixed with the amino acids to implement an array comprising six sensing elements (two QDs + [two QDs \times two amino acids]) (Fig. 4a). Due to the dissimilar affinity of the metal ions (Ag^+ , Ca^{2+} , Cd^{2+} , Cu^{2+} , Co^{2+} , Fe^{2+} , Fe^{3+} , Mn^{2+} and Ni^{2+}) to Gln, Arg and QDs surface, different chelates were formed, allowing to obtain a specific response for each metal (Fig. 4b). The PL data were analysed through LDA (Fig. 4c) and HCA (Fig. 4d) which allowed their discrimination. A classification rate of 100 % was obtained through LDA considering 15 unknown samples of spiked tap water. Together with the differentiation of nine metal ions at nanomolar concentrations, the oxidation states of ferrous (Fe^{2+}) and ferric (Fe^{3+}) species were also distinguished [77].

3.1.1.4. Cross-reactive sensing array based on multi-emitter nanoplatforms as sensing strategy. The amount of data collected during analysis can be significantly increased upon the combination of various multi-emitter probes assembled on a cross-reactive array, which would therefore allow the discrimination of a panoply of analytes. This strategy

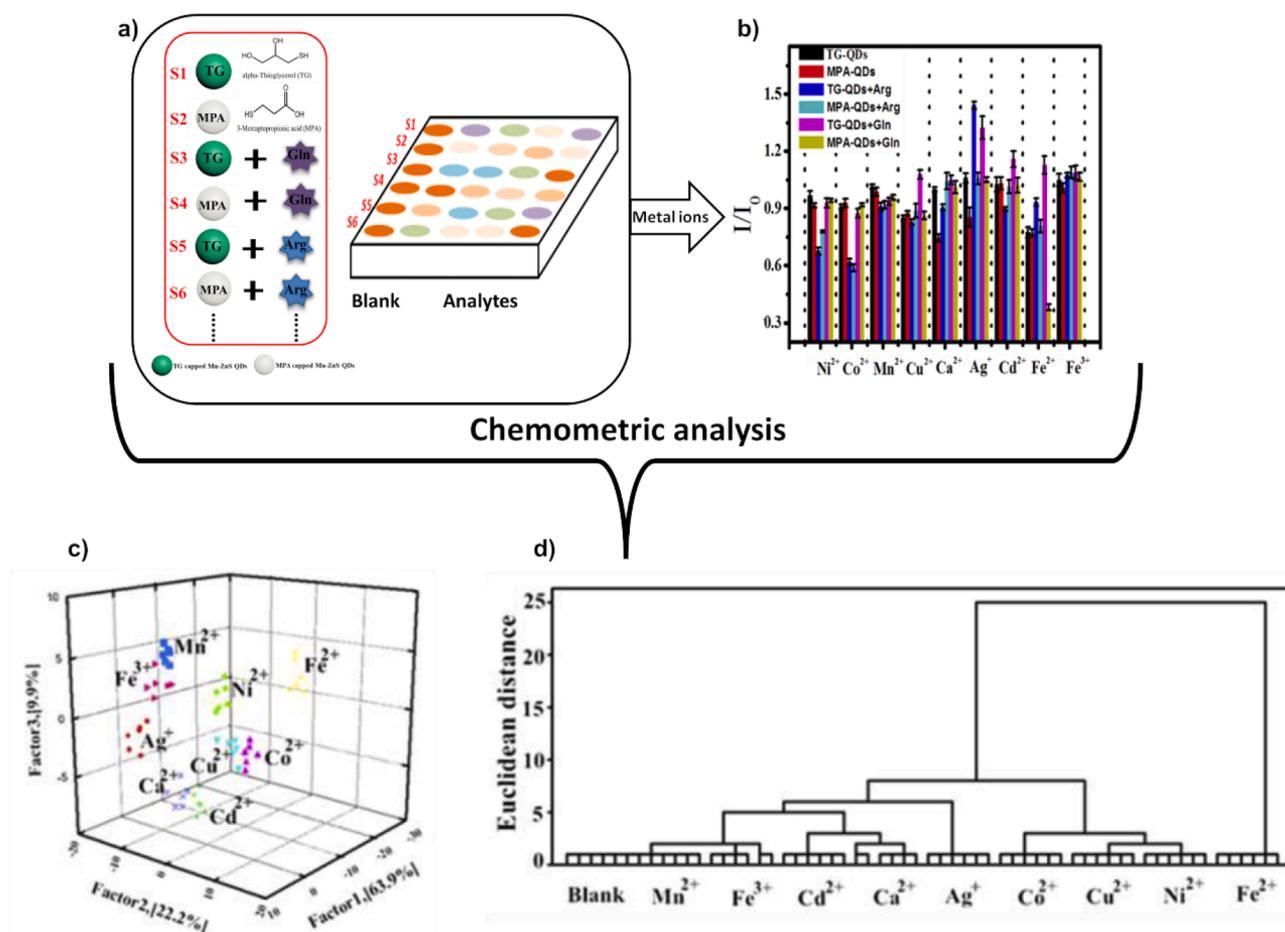


Fig. 4. Six cross-reactive sensing element encompassing the combination of QDs capped with different capping ligands and two amino acids a); distinct response patterns produced from the sensor array in the presence of different metal ions b); three-dimensional canonical score plot obtained from LDA c) and the dendrogram obtained from HCA for 9 metal ions d). Adapted with permission from [77]. Copyright 2017 Elsevier.

simultaneously increases the cross-reactive selectivity and simplifies the implementation of visual sensing approaches, since it enables visual detection relying on colour tonality modulation. In effect, changes in colour tonality are easier to identify than changes in colour brightness typical of QDs-based visual detections. The previously referred research group developed two visual and ratiometric-based sensors arrays for the differentiation of biothiols [78] and metal ions [79] (Table 2). In the first work [78], GSH, glutathione disulphide (GSSG), CA and cysteine (CYS) were identified and discriminated through HCA and LDA, respectively, using two ratiometric probes (blue-emitting CDs with red-emitting NALC-capped CdTe QDs and orange-emitting rhodamine B with green-emitting NALC-capped CdTe QDs). These two ratiometric probes were analysed in the absence and presence of 5 mol L⁻¹ NaOH, originating a four-element ratiometric sensor array. By adding the target analytes, and due to the high affinity of SH groups to the surface of CdTe QDs, the PL emission of these QDs was modulated, acting as analyte-dependent fluorophore (ADF), while the emission of the CDs and rhodamine persisted unchanged (reference fluorophore - RF). HCA and LDA were used for the identification and differentiation of these biothiols, and their mixtures, in human plasma [78] with a cross-validation accuracy of 89 %. The same research group, implemented a ratiometric sensor array for the identification and discrimination of five metal ions (Cd²⁺, Cu²⁺, Fe³⁺, Hg²⁺ and Pb²⁺) in spiked river water and digested salmon fish samples, through HCA and LDA, respectively [79]. Three distinct ratiometric probes composed of blue-emitting CDs stabilized with different capping agents (Gly, ethylenediamine (EDA) and urea) were used as ADF, while orange-emitting rhodamine B was used as RF. The metal ions were well-distinguished throughout RGB analysis by HCA and discriminated by LDA, which permitted the identification of a characteristic colour fingerprint-response pattern for each metal [79]. In fact, cross-validation accuracies of 91 % and 93 % were obtained for spiked fish and water samples, respectively. Results also showed a linear response regarding all metals' concentration, and a total Euclidean distance over a range of 10 to 100 µM. In a similar work, Ghasemi et al. [52] developed a two-emitter sensor array that was coupled to LDA data processing for the differentiation of widely-used nitroaromatics, namely 2,4,6-trinitrotoluene (TNT), 2,4-dinitrotoluene (DNT), and 2,4,6-trinitrophenol (TNP), in spiked well waters and soil samples. The sensor elements were composed of CDs combined with distinctly sized thiomalaic (TMA)-CdTe QDs. This cross-reactive sensing array comprising several multi-emitter nanoplatfoms allowed the discrimination of mixtures of structurally similar nitroaromatics in soil and groundwater samples, with a 100 % accuracy [52].

3.1.2. First-order data for quantification purposes

While some analyses only require a qualitative sample assessment, most of the time it is imperative to obtain quantitative information regarding sample contents. In these circumstances, chemometric tools other than the previously referred could provide a valuable service. An illustrative methodology exploring chemometric models for qualitative and quantitative analysis of PL data resorts to a CdTe/nano ZnTPyP sensing system for the discrimination and quantification, through PLS-DA and PLS, respectively, of carbamate pesticides (carbaryl, carbofuran and metolcarb) in spiked food matrices (tea, cabbage and apple) [80] (Table 2). The carbamate pesticides impaired the FRET process leading to a recovery of the CdTe QDs PL. An accuracy of 100 % of correct predictions was obtained with PLS-DA for all pesticides considering the testing set. Regarding PLS results, the authors claimed R² higher than 0.999 for all the pesticides over a linear range of 1 to 20 µg L⁻¹ and LOD of 0.91, 0.89 and 0.78 µg L⁻¹ for metolcarb, carbaryl and carbofuran, respectively, considering the testing set [80]. Another interesting work by Fan et al. [81] used NALC-capped CdTe and ZnCdSe QDs for the analysis of five pesticides (cartap, dipterex, dursban, methyl thiophar and paraquat) using moving window partial least squares discriminant analysis (MWPLS-DA) and optimized sample-weighted least-squares support vector machine based on particle swarm

optimization (PSO-OSWLS-SVM) model for qualitative and quantitative analysis, respectively [81]. MWPLS-DA is similar to the working principle of PLS-DA but instead of using the entire spectra, it selects just some useful spectral intervals through a moving window that is user-adjusted [82]. This way, not only the complexity and size of the calibration set is reduced, but also the influence of uninformative spectral variables in the final model, which contributes to a lack of accuracy, is minimized. The uninformative spectral variables result from the need of using a high number of latent variables (LV) to maintain model accuracy, when compared to informative spectral variables [82]. The PSO-OSWLS-SVM uses the PSO algorithm which calculates the samples weights of the training samples and hyper-parameters in SVM model, considering both the training and validation performances. This model is claimed to deal with nonlinearity data more effectively than PLS and to withstand overfitting more effectively than least-squares support vector machine (LS-SVM) [83]. In fact, when compared to PLS and LS-SVM, the obtained results are much lower in terms of root mean square error of calibration (RMSEC) or prediction (RMSEP) [83]. The SVM has been firstly proposed for classification problems, but its use has been extended to regression problems through the use of loss functions [84,85]. This type of chemometric model is able to deal with non-linear data through the use of a kernel function but the selection of the best kernel function requires the tuning of hyper-parameters, which is considered cumbersome and prone to overfitting, as aforementioned [83,84]. The LS-SVM reduces the computation time of SVM because it only requires to solve a few linear equations, in opposition to SVM that needs to solve quadratic problems, making the characterization of the optimal conditions much simpler [83–85]. Considering the obtained results, the use of MWPLS-DA model with a moving window of 60 data points yielded 100 % of correct classifications for all pesticides regarding the testing set. These results were better than those obtained with PLS-DA, considering the entire spectra (80 % of correct predictions). With respect to the use of PSO-OSWLS-SVM for the quantification of the pesticides, a R² higher than 0.999 was obtained for all pesticides, considering the testing set. Additionally, the authors performed a recovery assay with wastewater and tea samples and claimed recoveries percentages between 95 % and 104 % for all pesticides, which demonstrates the suitability and accuracy of the developed methodology.

In another work [86], a four-channel sensor array composed of distinctly passivated CdTe QDs was used to distinguish between commercial baijiu samples with different brands, aroma types, storage years and qualities, by using LDA. A classification rate of 100 % of correct predictions (cross-validation) was obtained for all parameters, except for the different types of brands, where a classification rate of 99.8 % was claimed by the authors (one sample was misclassified). In addition, five organic acids (acetic acid, butyric acid, caproic acid, heptanoic acid and lactic acid), generally present in baijiu, were accurately discriminated through LDA and quantified by resorting to PLS [86]. A classification rate of 100 % (cross-validation) was obtained. For the PLS results, the authors claimed accurate results for all the acids within the linear range of 10 to 80 µmol L⁻¹. The results regarding the coefficient of determination (R²) and root mean square error (RMSE) for calibration and cross-validation, as well as LOD and LOQ, were not mentioned by the authors.

Aiming at guaranteeing the quality of some samples (i.e., pharmaceutical and environmental), several analytical approaches for quantification purposes have also been carried out. Contrary to univariate calibration procedures wherein only the maximum emission wavelength is analysed (commonly through the Stern-Volmer equation that describes the PL quenching), the use of the full PL spectrum allows to obtain first-order data advantage, as mentioned previously. Again, the selection of the most appropriate chemometric model needs to consider not only the objective of the work, namely classification/discrimination or quantification, but also if the obtained data is linear or nonlinear.

Some multi-emitter nanoprobos have been developed for quantitative analysis, maximizing the extracted information from a single PL

spectrum. Our research group developed two works in which CdTe QDs emitting at complementary wavelengths and with dissimilar reactivities were combined for the determination of some metal ions [87,88]. In one of these works [87], distinct thiol-based ligands (MPA, GSH and 2-mercaptoethanesulfonate – MES) were employed to adjust reactivity thus obtaining a higher assortment of PL responses. As the different thiol-based capping ligands contain dissimilar terminal functional moieties, a characteristic response to each metal ion was observed. This way, the use of a dual-emission probe allowed to obtain a distinctive emission spectrum profile. By resorting to PLS, Cu^{2+} and Hg^{2+} metal ions were accurately quantified yielding R_{CV}^2 higher than 0.98 for both metal ions and RMSEC and root mean square errors of cross-validation (RMSECV) ranging from 0.0094 to 0.15 and 0.013–0.19 mg L^{-1} for Cu^{2+} and Hg^{2+} , respectively. Additionally, considering the sensitivity of the developed models, the authors concluded that the selection of dual-emitting nanoprobes composed of QDs with distinct capping ligands could be

useful to achieve a more sensitive quantification of the target analytes [87]. In the second work [88], a dual emission nanoprobe composed of two differently-sized CdTe QDs (both capped with the same capping ligand -MPA) was used to simultaneously quantify mutually interfering metal ions (Cu^{2+} , Hg^{2+} and Pb^{2+}). By applying PLS, it was possible to quantify the three metal ions in binary mixtures, with R_{CV}^2 ranging from 0.74 to 0.89 for all ions. However, when considering the three metal ions altogether in the same ternary mixture, the model was not able to quantify them accurately. Pb^{2+} and Hg^{2+} metal ions yielded R_{CV}^2 of 0.87 and 0.73, respectively, while for Cu^{2+} a R_{CV}^2 of 0.51 was obtained (revealing that it was not possible to perform an accurate quantification of this ion) [88].

Adjustment of QDs reactivity and selection of the most suitable chemometric model is crucial for the success of the analytical methodology. The combination of QDs with distinct reactivities allows obtaining a specific response profile for each analyte, making possible the

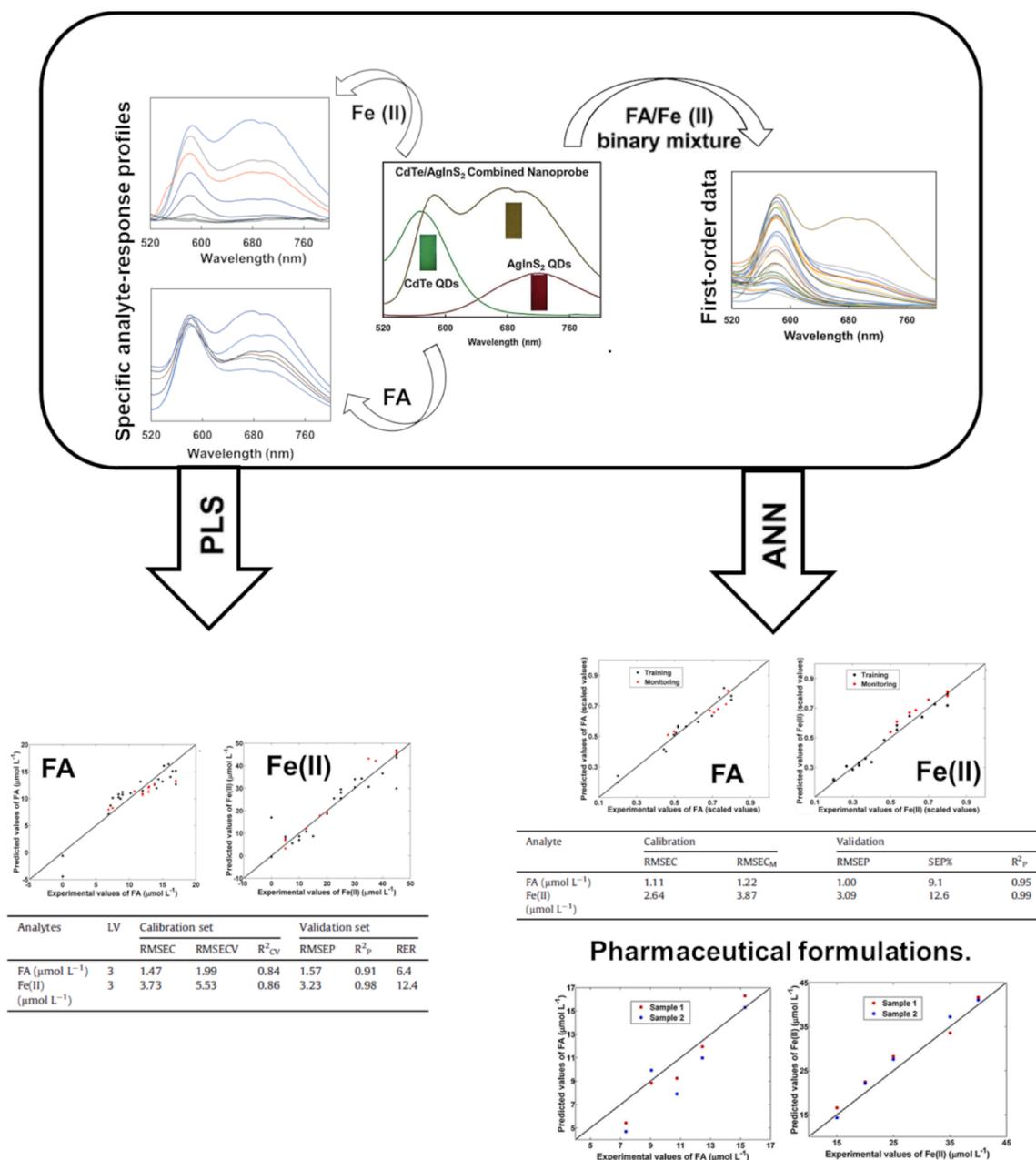


Fig. 5. Implementation of a dual-emission combined nanoprobe seeking to obtain specific analyte-response profiles which were analysed by PLS and ANN for the simultaneous determination of FA and Fe^{2+} in pharmaceutical formulations. Adapted with permission from [50]. Copyright 2020 Elsevier.

simultaneous determination of multiple analytes in one sample. In this sense, our research group employed a dual-emission nanoprobe encompassing CYS-CdTe and MPA-AgInS₂ QDs for the monitoring of folic acid (FA) and Fe²⁺ in a binary mixture (Fig. 5) [50] while Barati et al. combined CYS-CdTe QDs with CDs for the simultaneous determination of TNT and 4-nitrophenol (4-NP) [89]. In both works, one of the analytes (Fe²⁺ or 4-NP) led to a PL inhibition of both QDs PL while the remaining analytes did not cause a significant variation in the PL intensity of one of the QDs (FA to CYS-CdTe and TNT to CDs, respectively). Their characteristic emission spectra profiles were analysed through PLS and ANN models. In the first work [50], authors claimed R²_p of 0.91 and 0.98 when using PLS and R²_p of 0.95 and 0.99 when using ANN for the quantification of FA and Fe²⁺, respectively. Moreover, the authors tested the best developed model, in this case ANN, in two commercial samples obtaining standard errors of prediction below 9%. In the second work [89], the quantification of TNT and 4-NP in test samples yielded errors below 10% for both parameters and both chemometric models. However, the authors claimed lower results when using ANN. In conclusion, ANN exhibited a superior efficacy for the simultaneous determination of the analytes in binary mixtures, since a non-linear relationship between the obtained fluorometric data and the analytes concentration were observed [50,89].

The use of suitable chemometric models for processing PL first-order data enables the qualitative and quantitative analysis of real samples, even without the use of a specific surface functionalization procedure. However, some samples have a complex matrix with unknown interfering species, impairing an accurate and selective discrimination and/or quantification of the target analytes. In the next section, the acquisition of second and higher-order data that could allow circumventing the occurrence of uncalibrated interfering components, will be discussed.

3.2. QDs-based methodologies exploring second- and higher-order data analysis

Despite the advantages of using first-order data to circumvent selectivity issues, it should be emphasised that the target analyte determination in the occurrence of interfering substances can only be performed if these interferents are comprised in the calibration samples. In this case, the composition of the sample matrix must be previously known to guarantee an accurate quantification of the analyte. Otherwise, the first-order advantage cannot be availed. However, it is possible to determine analytes in samples with complex matrices and unknown composition by means of the use of second- or higher-order data, due to the second-order advantage in combination with suitable chemometric models [90,91].

In fact, when second-order data, such as three-dimensional data sets (e.g., EEM), is processed by suitable chemometric models they enable obtaining the second-order advantage which allows circumventing the occurrence of unpredicted sample components not included in the calibration step. This is a situation that occurs frequently when handling with samples of a more complex nature and with unknown constituents, such as food, environmental or biological samples, where it is very difficult to distinguish between the different components or anticipate the occurrence of interferences that can affect the quality of the analytical results. In these circumstances, second- or even higher-order data can be acquired to overcome the referred interferents' effects [46,47]. However, there are limitations that should be taken into consideration. For example, if the presence of uncalibrated species (interferents) suppresses the fluorescence signal or, in a different manner, enhances the fluorescence signal as to saturate the detector, it won't be possible to determine the analyte under study. In these extreme cases alternative strategies, such as a sample dilution, can be used to circumvent this problem.

3.2.1. Second-order data analysis exploring EEM

Second-order data in QDs-based methodologies can be obtained by recording the evolution of the PL spectrum throughout time (kinetic measurement of QDs-analyte interaction) (Fig. 6a), or by registering EEM of the QDs-analyte interaction, using a suitable fluorometer (Fig. 6b), which is by far the most used approach as it assures abundant spectral information. Indeed, EEMs can increase the specificity of fluorescence spectroscopy by accumulating PL emission spectra at several excitation wavelengths, thus providing a matrix of PL intensities [92].

Before discussing the most adequate chemometric models to deal with second-order QDs-based methodologies, it is important to clarify the concept of data trilinearity, which plays an important role in terms of data analysis. Trilinear second-order data implies that the data can be decomposed into three independent factors, being each factor a linear combination of two variables. Therefore, second-order data can only be considered trilinear if complies with the following three assumptions: i) the obtained signal for any sample is bilinear; ii) linearity correlated with the analyte concentration; and iii) the components profiles in all the samples are constant [95]. The first assumption requires that the obtained signal is proportional to the product of the two dimensions of the obtained matrix [90,91]. An example of bilinear data is the EEM signals in which the intensity of the obtained signal at a specific wavelength of excitation and emission is a function of the respective excitation and emission wavelengths [90,91]. Differently, the absorbance signals of the dissociation of weak acids at different pH values (as weak acids concentration is affected by pH), can be considered as non-bilinear. Therefore, if the dimensions of the obtained matrix are mutually dependent, the data are non-bilinear [90,91]. The second assumption signifies that the maximum signal measured at both data dimensions is directly proportional to the analyte concentration [95]. The last assumption states that the shape of the samples' profiles is the same in all the dimensions, only varying in the intensity due to the different concentrations of the analyte [95].

In this sense, EEM data can be generally classified as trilinear data. However, some circumstances, such as the presence of Rayleigh dispersion, Raman dispersion, second-order harmonic Rayleigh dispersion, the presence of inner filter effects [90,91] and deviations of the maximum emission band (hypsochromic or bathochromic effect), that may differ from sample to sample, contribute to the non-trilinearity of EEM data. Nonetheless, there are some strategies, for example, pre-processing of the EEM data to remove the non-trilinear signals, modelling with non-trilinear models, replacing the data values affected Rayleigh dispersion by zero or missing values, subtracting blank signal from sample signal or by defining the excitation and emission ranges so that the dispersion signals are reduced improving the signal-to-noise ratio, that can be used to circumvent the lack of trilinearity of EEM [90,91]. These strategies won't work in all situations, and therefore the non-trilinearity of EEM data might not be avoided. When just the excitation or the emission profiles of EEM data are affected, the data is known as non-trilinear of type 1, when both profiles are affected, the data is known as non-trilinear of type 2 and when the EEM data is not bilinear is known as non-trilinear of type 3 [90,91]. Strategies regarding the identification of trilinear or non-trilinear data are comprehensively discussed elsewhere [96]. The knowledge of the type of the generated EEM data, in terms of classification as trilinear, non-trilinear of type 1 or non-trilinear of type 2, and the objective of the work, are crucial to select the best chemometric model. In fact, some of the chemometric models are not able to deal with the lack of data trilinearity.

The most common chemometric models used for QDs-based methodologies exploring second-order advantage are PARAFAC, MCR-ALS and U-PLS/RBL. It is out of the scope of this manuscript to describe in detail how these chemometric models work and on what assumptions they are based on. For more details, please see the following references [90,91,95–98]. About these chemometric algorithms, it is pertinent to refer that PARAFAC is able to model trilinear data, MCR-ALS is able to model trilinear and non-trilinear data of type 1 and U-PLS/RBL is able to

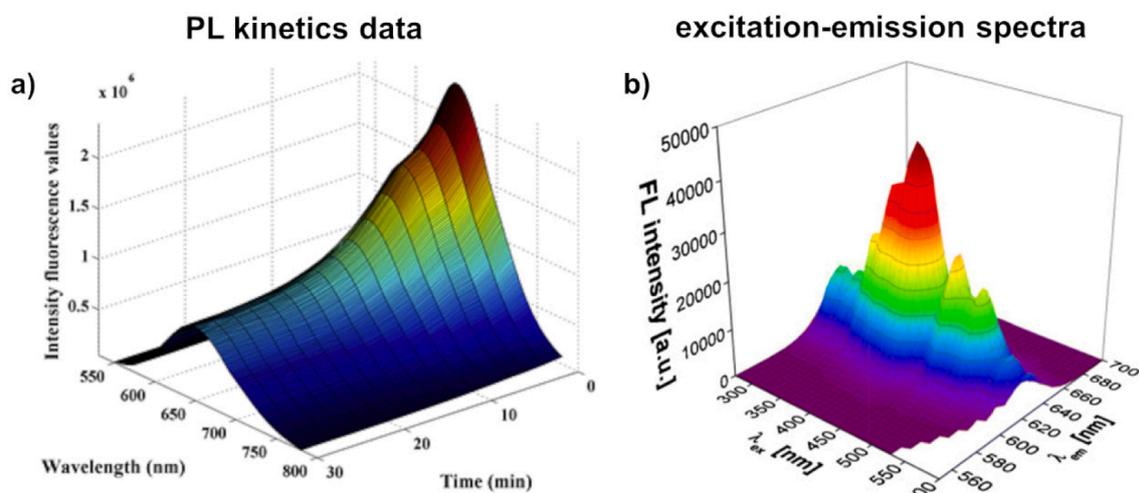


Fig. 6. Second-order data using QDs-based methodologies: a) PL kinetics data of TMA-AIS QDs upon the interaction with OTC throughout 30 min; and b) excitation-emission spectra (λ_{ex} ranged from 270 to 600 nm and λ_{em} ranged from 550 nm to 700 nm) of GSH-CdSeS/ZnS QDs. Adapted with permission from [93,94]. Copyright 2021 Springer and 2021 Elsevier.

model trilinear and non-trilinear data of type 1, 2 and 3 [90,91]. Concluding, MCR-ALS is more adequate for non-trilinear data of type 1 and U-PLS/RBL for non-trilinear data of type 2 and 3, although only one work [99] has described the suitability of U-PLS/RBL to model non-bilinear data. Albeit it might look appealing to apply U-PLS/RBL in all situations, due to its flexibility, its need for additional information regarding the process under analysis justifies the application of alternative chemometric models, like PARAFAC and MCR-ALS, that are able to provide more intuitive and meaningful chemical interpretations of the data modelled [90,91].

The number of works available in literature exploring second- or high-order data in QDs-based analytical methodologies is very reduced (Table 3), which clearly demands for deeper research in this field mostly because of the inherent benefits that it could provide, which are related, as previously referred, with achieving the second-order advantage. One of the first works evaluated the pH effect on differently-sized MPA-CdTe QDs [100] (Table 3). The obtained data were analysed with different chemometric models, namely PARAFAC, PARAFAC2 and MCR-ALS. The authors observed variations in the emission bands (maximum emission

wavelength red-shift), which meant that the obtained data was of non-trilinear type 1. This was confirmed throughout the comparison of singular magnitude values obtained using SVD. In this regard, MCR-ALS and PARAFAC2 demonstrated to be the most suitable chemometric models to deal with the obtained data. Additionally, the use of larger CdTe QDs revealed a higher pH sensitivity, probably because a nanocrystals populations with a bigger size, which consequently exhibited a greater size heterogeneity, was more affected in terms of PL emission as the pH increased [100]. The same group verified a quenching effect on GSH-CdTe QDs at increased concentrations of Pb^{2+} [101] and on DAB-CdS QDs at increased concentrations of nitromethane [102]. The chemometric model used for data analysis on both works was PARAFAC, as the presence of both the metal ion and the organic compound did not show any deviations in the obtained EEM data (which means the data can be considered trilinear). This chemometric model was able to estimate the excitation and emission spectra, the quenching profiles, and to obtain the second-order advantage. Another archetypal example of the use of EEM matrices to obtain second-order data sets [103] explored the versatility of MCR-ALS to deal with trilinear and non-trilinear data of

Table 3

QDs-based PL approaches using second or higher-order data analysed resorting to chemometric analysis.

Analytes	Sensing platform	Chemometric tool	Data set	LOD	Ref.
pH	MPA-CdTe	PARAFAC PARAFAC2 MCR-ALS	EEM	n.a.	[100]
Pb^{2+}	GSH-CdTe	PARAFAC	EEM	n.a.	[101]
Nitromethane	DAB-CdS	PARAFAC	EEM	n.a.	[102]
Hg^{2+}	PEG-CdS	MCR-ALS	EEM	n.a.	[103]
Ionic strength	DAB-CdS				
Act D	CdSe/ZnS	Restricted Tucker3 HTD	EEM	n.a.	[104]
6 neurotransmitters	GSH-CdSeS/ZnS	U-PLS U-PLS-DA	EEM	n.a.	[93]
Amino acids, oligopeptides, neurotransmitters	CdSe/ZnS	U-PLS-DA	EEM	n.a.	[111]
Cu^{2+}	CYS-CdS	MCR-ALS PLS	kinetic	13 nmol L ⁻¹	[113]
OTC	TMA-AgInS ₂	U-PLS	kinetic	0.144 $\mu\text{mol L}^{-1}$	[94]
AFB1	MPA-AgInS ₂	U-PLS	kinetic	1.2 $\mu\text{g L}^{-1}$	[114]
ASA	MES-CdTe@MPA-AgInS ₂	U-PLS N-PLS MLF-NNs RBF-NNs	kinetic	2.82 mg L ⁻¹ 3.10 mg L ⁻¹ 3.38 mg L ⁻¹ 3.26 mg L ⁻¹	[48]
Histamine	CDs@MPA-AgInS ₂	U-PLS N-PLS	kinetic	1.26 mg L ⁻¹	[115]
pesticide residues	Mn-ZnS and COFs	PCA	multi-block data from instrument hyphenation	0.15 $\mu\text{g mL}^{-1}$	[116]

type 1. In this work, CDs passivated with NALC and functionalized with Poly(ethylene glycol) (PEG), which were sensible to increasing concentrations of Hg^{2+} , and CdS QDs stabilized with TGA and attached to DAB, that were susceptible to the modification of ionic strength of the aqueous medium, were used as sensing platforms. The interaction between the metal ion and CDs only caused a PL decrease while, in the case of Cd-based nanocrystals, the increase of the ionic strength produced a simultaneous variation of the maximum excitation (347.9 nm to 367.8 nm) and emission (485.7 nm to 633.4 nm) wavelengths as well as a quenching effect. This bathochromic effect, observed in both excitation and emission spectra, and probably related to the aggregation of the nanoparticles due to changes in the QDs surface charges caused by the higher ionic strength values, obviously resulted in a significant deviation of the trilinearity of the data structures. Although this may suggest that the obtained data for the ionic strength was non-trilinear of type 2, the variation on the excitation profile was minimal. Otherwise, MCR-ALS would not be the right chemometric model. The results obtained with this sensing probe clearly demonstrated its adequacy for the detection of Hg^{2+} and ionic strength [103].

A multi-way analysis of a FRET process, involving CdSe/ZnS QDs, was developed to assess drug-DNA interactions with drug discovery purposes [104]. Commercial Cd-based QDs were conjugated with a synthetic oligonucleotide (corresponding to a specific region of c-Myc gene), in order to quantitatively describe the interaction of c-Myc duplex with actinomycin D. Effectively, actinomycin D was used as energy donor while QD-dsDNA were applied as energy acceptor. The obtained three-way fluorescence data were analysed resorting to restricted Tucker3 and hard trilinear decomposition (HTD). Unlike MCR-ALS, which decomposes second-order data through a bilinear model after data augmentation, the Tucker3 model is capable of directly decomposing second-order data. It has also the capacity of dealing with non-trilinear data, defining a different number of components for each mode in the second-order data (unlike PARAFAC models), and allowing the interaction between each mode [105]. In this sense, this model becomes very complex in practice, as its flexibility generates non-unique solutions and the presence of rational and intensity ambiguities [96,105,106]. Nonetheless, the application of several constraints can reduce these ambiguities but at the same time hinders its interpretation [106,107]. Tucker3 is not recommended for huge data sets because a higher number of components will be needed increasing the possible interactions between these components and making the optimization and interpretation very troublesome [106,107]. Indeed, as the profiles being differentiated are usually very similar to each other, the rotational

ambiguities increase [106,107]. More information about Tucker3 model can be found elsewhere [106,108]. The HTD is based on hard-modelling methods which take into account previous information regarding a possible chemical model. This is different from soft-modelling algorithms, such as MCR-ALS, PARAFAC, and direct trilinear decomposition (DTD), which does not use any a priori knowledge for model development. However, the use of restrictions in MCR-ALS, such as unimodality, non-negativity, selectivity, closure and kinetic model, demonstrated that this model allows to attain better results than hard-modelling models, being even capable of dealing with kinetic processes that have interfering species [109]. More information regarding HTD, can be found in the following reference [110]. In the abovementioned work [104], both chemometric models (HTD and restricted Tucker3) allowed the complete resolution of EEM data obtained from FRET process, as well as the estimation of the equilibrium constants of hybridization and complexation (with values similar to the ones previously present in the literature), despite the high spectral overlap between the absorption and emission bands of the acceptor and the donor, respectively [104]. It should be mentioned that the application of HTD model started with the information obtained by restricted Tucker3.

The work performed by Głowacz et al. [93] demonstrated that analytical estimations with superior selectivity can be attained when the complexity of the obtained data is increased (Fig. 7). The authors showed that depending on the type of acquired data (zeroth-, first- or second-order), the discrimination of 6 neurotransmitters (dopamine, norepinephrine, epinephrine, serotonin, GABA, and acetylcholine) upon interaction with GSH capped CdSe/ZnS and by using U-PLS-DA, can be significantly improved. Moreover, the quantification of dopamine, norepinephrine, and epinephrine using U-PLS was also successfully reported. Regarding the type of acquired data, authors concluded that when only a single fluorescence value of the quenching process was obtained at a maximum emission wavelength for a given concentration (zeroth-order data), the discrimination of the respective neurotransmitters based on non-specific interactions was not possible, as expected. However, when a spectrum (emission scan at fixed excitation) was used, a better discrimination between the compounds that contain a catechol group (dopamine, norepinephrine and epinephrine) was obtained using PLS-DA model, with an accuracy of 78 % considering the test set. Finally, when using full EEMs, the discrimination was improved to 89 % of accuracy, considering the test set. As abovementioned, the analysis of the EEM data with U-PLS enabled the determination of the catecholamine neurotransmitters at micromolar concentration level with a R^2 between 0.92 and 0.99 [93]. In a previous work developed by the same

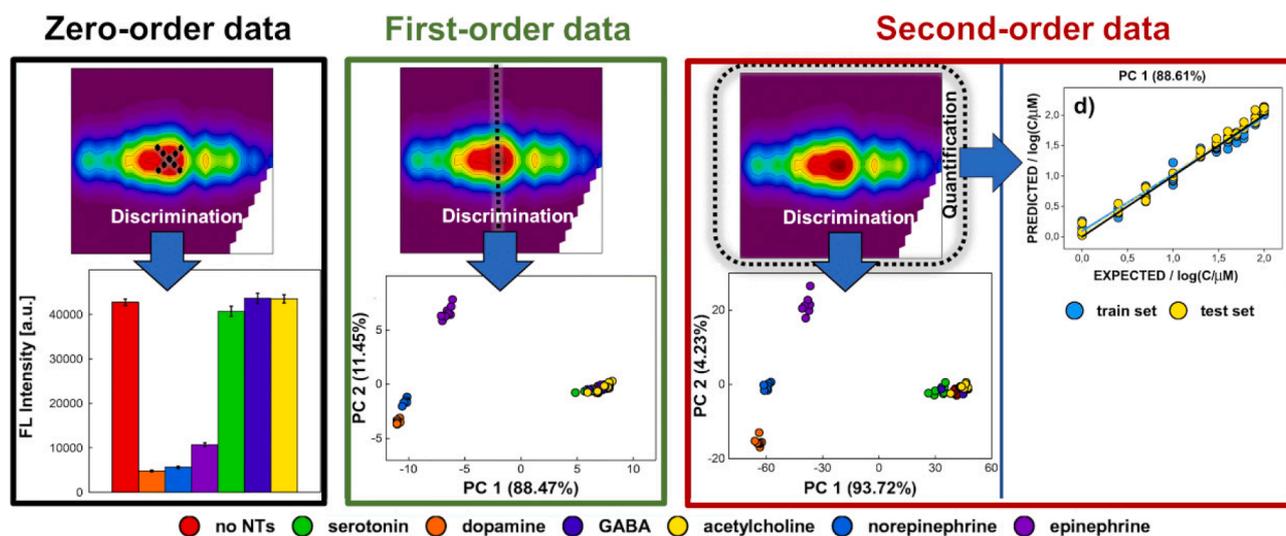


Fig. 7. Influence of the complexity of the PL data on the discrimination and/or quantification of five neurotransmitters. Adapted with permission of [93]. Copyright 2021 Springer.

research group, EEM data were used to obtain a fluorescence fingerprinting of different analytes, allowing the discrimination of several neurotransmitters, amino acids and oligopeptides at different pH, through U-PLS-DA [111]. An accuracy of around 80 % for neurotransmitters, amino acids and oligopeptides was obtained.

3.2.2. Second-order data analysis exploring kinetic approaches and multi-block data

Despite all advantages regarding second-order EEM data, its acquisition requires expensive and specific instrumentation [112]. In alternative, the time-based recording of QDs PL spectra modulation in the presence of increasing analyte concentrations, which can be easily surveyed by using a common fluorimeter, is a simple and expeditious way to also collect second-order data. Nevertheless, only a very restricted number of analytical methodologies has explored this possibility.

One of the first examples using three-dimensional kinetic data was developed by Abdollahi et al. [113] (Table 3). In this work, the modulation of the PL of CYS-capped CdS QDs in the presence of Cu^{2+} was studied throughout the time, and the obtained data were analysed

resorting to different chemometric models, namely MCR-ALS and PLS after row-wise augmentation of the second-order data. MCR-ALS was used to extract the spectral profiles of all chemical species present (the number of chemical species was estimated through EFA) while PLS was used for the metal ion quantification. The kinetic interaction of the nanoprobe with other ions (Ag^+ , Ni^{2+} and Hg^{2+}) was also studied. The authors observed that for each tested metal ion, a characteristic emission spectra profile was obtained. This fact demonstrates that different metal ions cause different kinetic behaviours in the CdS QDs PL, which can be very useful not only for the accurate quantification of Cu^{2+} in the presence of interfering species, but also for the simultaneous determination of co-existing metal ions that influence the QDs PL.

Our research group developed two chemometric-assisted kinetic determinations of oxytetracycline (OTC) [94] and aflatoxin B1 (AFB1) [114] using ternary AgInS_2 QDs as PL sensing elements and as a photocatalyst (ROS generators), respectively. In the first work, the presence of lactose monohydrate in commercial veterinary pharmaceutical formulations impaired the ratiometric determination of OTC. Aiming to circumvent the selectivity issues, the behaviour of the AIS QDs PL

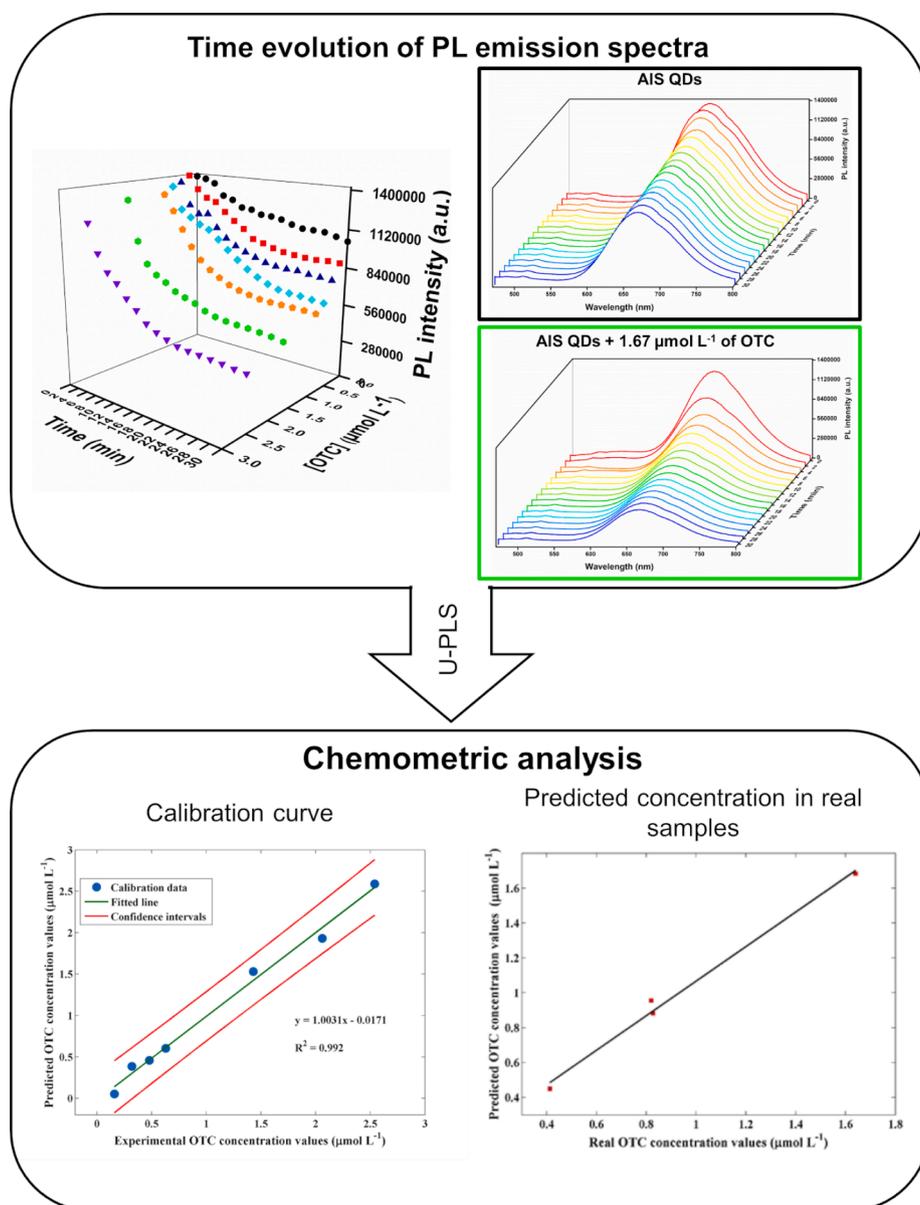


Fig. 8. Kinetic approach using the behaviour of AIS QDs in the presence of OTC through 30 min. and resorting to chemometric tools which allow the OTC quantification in pharmaceutical formulations. Adapted with permission from [94]. Copyright 2021 Elsevier.

quenching in the presence of OTC was carefully analysed throughout time (30 min). The implementation of a kinetic approach allowed the reduction of the LOD, enhancing sensitivity, as well as to acquire second-order data, which enabled the quantification of the drug in the presence of uncalibrated interferents (Fig. 8). The kinetic data was processed by U-PLS which assured the accurate determination of OTC on two different commercial formulations [94]. Indeed, a R_p^2 and a relative percentage error (RE) of 0.99 and 7.6 %, respectively, were obtained for the test set. In another work from our group [114], it was verified that, despite no perceptible variation on the QD's optical properties (meaning that the QDs did not directly behave as sensing elements), the fluorescence intensity corresponding to the AFB1 emission band gradually increased throughout the kinetic process. Effectively, reactive oxygen species (ROS) generated by the AIS QDs upon visible light irradiation led to the oxidation of AFB1 generating a photodegradation product with higher fluorescence intensity than the native fluorescence of AFB1. The acquired PL data were processed resorting to U-PLS and tested in two AFB1 commonly contaminated samples namely, maize and rice samples [114]. A R_p^2 of 0.96 and 0.99 and a RE of 8.0 and 5.9 were obtained for the spiked rice and maize samples, respectively, which attested the accuracy and suitability of the developed methodology. Moreover, the developed methodology was validated against a reference method, specifically liquid chromatography with mass spectrometry detection (LC-MS/MS), which yielded relative deviation values lower than 3 %.

Seeking the obtaining of a higher amount of data information to enable the development of more efficient and reliable chemometric models for quantitative analytical analysis, our research group proposed two different kinetic approaches for the determination of acetylsalicylic acid (ASA) [48] and histamine [115] exploring multi-emitter nanoparticles as sensing platforms. In the first case, a dual MES-CdTe/MPA-AgInS₂ QDs PL probe exhibiting overlapped emission bands was developed for ASA determination in which the obtained non-linear PL kinetic data were analysed and compared by N-way partial least-squares (N-PLS), U-PLS, radial basis function neural networks (RBF-NNs) and multilayer feed-forward neural networks (MLF-NNs) [48]. The emission intensity of both QDs of the combined probe was inhibited upon interacting with increasing concentrations of ASA. However, they exhibited different sensitivities towards the analyte, showing an affinity of MES-CdTe QDs higher than that of MPA-AIS ones. This fact enables to obtain a specific ASA-response profile, which, by applying a kinetic approach, facilitated its discrimination in samples where this drug is combined with other APIs. Our research team also demonstrated that the overlapping profile of the emission spectra can be mathematically decomposed through the use of chemometric models, allowing an accurate determination of ASA in pharmaceutical formulations even in the presence of other uncalibrated APIs. The comparison between the chemometric models showed no significant differences in terms of figures of merit, however, both PLS models required less computation time, were simpler to implement and the interpretation of the generated data was easier than those obtained by NNs [48]. In the second case, a ratiometric probe composed of CDs and MPA-AgInS₂ QDs was used for histamine determination in food samples, being the kinetic-spectral data analysed by N-PLS and U-PLS [115]. In the proposed ratiometric nanoprobe the AIS QDs acted as analyte-dependent fluorophore (PL quenching effect in the presence of the biogenic amine) while CDs were employed as reference fluorophore (no PL modulation upon the interaction with the analyte). The results obtained in this work demonstrated that both chemometric models enabled to accurately quantify histamine in complex samples such as tuna, hake fish and tomato. Nevertheless, the N-PLS model showed better results in terms of prediction than those obtained by the U-PLS model. A RE below 10 % with a R_p^2 higher than 0.9 as well as a LOD and LOQ of 1.26 and 3.80 mg L⁻¹ were obtained, respectively, which demonstrate that the proposed method can be successfully used to monitor histamine in foodstuffs even if the levels of this biogenic amine are below the maximum limit established in the EU and USA legislation.

The use of kinetic data only has analytical significance if variations in

the optical properties of the QDs are observed steadily throughout time. However, in some cases, upon mixing QDs and analyte, the optical properties of the former are almost instantly altered remaining subsequently unchanged for long periods of time, without analytical relevance. In these cases, and in alternative to the use of EEM or kinetic data, information obtained from multiple hyphenated sources can be explored. In effect, the use of multi-block data obtained from different detection instruments/sensing platforms enables acquiring complementary useful information, that guarantees reliable analyte determination. Obviously, for the merging of data acquired from multiple instruments, it is crucial to assure that the QDs optical properties remain unchanged throughout the analysis, being only affected by analyte interaction, or that the samples are measured at stable experimental conditions. Only one work has explored the use of multi-block QDs data for analytical purposes. Yuan et al. [116] developed a sensing approach in which phosphorescence, fluorescence and ultraviolet-visible spectroscopy were used to analyse the interaction between multiple pesticide residues and amino-silane and ionic liquids (ILs) co-modified Mn-ZnS QDs and covalent organic frameworks (COFs). The hyphenized three-dimensional optical signal was processed by resorting to PCA which enabled the differentiation of a panoply of pesticide residues in complex food samples [116]. The information regarding the PCA application and how the obtained data was assembled was not made available.

4. Conclusions and prospects

A careful and exhaustive revision concerning QDs-based methodologies using chemometric models was performed, including the optimization of the synthesis and characterization of QDs as well as its use in qualitative and quantitative analysis using first- and second-order data. This revision included a discussion of the most used chemometric models as well as their advantages in QDs-based methodologies. In fact, the analytical potential of QDs is clearly restrained by their high reactivity that result frequently in the establishment of non-specific interactions with both the analyte and the interfering species. This lack of selectivity could be amended by functionalising their surface with suitable molecular recognition mechanisms, which could be a cumbersome and expensive strategy. Alternatively, by guaranteeing the acquisition of complex analytical data, with plenty of chemical information, even if not visible or immediately perceptible, and by resorting to adequate chemometric models it is possible to extract valuable quantitative and qualitative information from all of the components in a sample, even in the presence of overlapping peaks or unpredicted interfering species, which would therefore assure the required selectivity.

Effectively, complex data like second or higher-order data can be very handy when analysing complex sample matrices, such as those found in the food and environmental industries, in which unknown constituents might occur, since the second-order advantage enables accurate analytes quantification, even in the presence of uncalibrated interferents. Unfortunately, only a few scientific works have been reported exploring the second- or higher-order data using QDs as a fluorescent analytical nanoplatform. It is expected that in a near future, the acquisition of three- or higher-array data using the response of combined nanoprobes (multiple QDs or fluorophores combined with QDs) and the subsequent PL data processing by using proper chemometric models might provide new possibilities and resourceful strategies for the simultaneous detection of multiple analytes in complex samples. Expected developments in multiplexed detection and array sensing could boost the application of pattern-based detection methods which, upon combination with multilinear pattern recognition methods would simplify the simultaneous analysis of multiple analytes in complex samples. Moreover, due to the versatility of QDs synthesis in terms of core composition, shells, cappings and coatings, functional groups, etc, it would be possible to develop small libraries of QDs probes that could be employed to build up sensing arrays and fingerprinting detection

schemes whose possible interactions, and respective outputs, are pre-determined.

On this subject, the synthesis of quantum dots, and the obtaining of high-quality nanomaterials, is determined by many variables whose weight and interconnection are not always discernible. In this field, as well as in the subsequent QDs characterization, the collecting of high value informative data and their processing by suitable chemometric methods could represent a major improvement in the implementation of expeditious and easily controlled synthetic approaches as these could provide chemically interpretable insights of evolving multi-component processes. Moreover, chemometric models could be a valuable tool for understanding and controlling the relationship between the coordination chemistry of QDs and their fluorescence properties. For instance, chemometrics can help analyze how variations in ligands affect the fluorescence data, or how their composition, size and shape affect their energy levels and bandgap, or their photostability, or even be used to predict the QDs emission wavelength and quantum yield, etc. Chemometric models can also be used to classify QDs seeking to identify and isolate specific populations, considered in terms of size and surface chemistry, for target applications.

Despite all advantages resulting from combining high-order data with chemometrics, there are also some possible drawbacks related with the complexity of this field, which can be challenging, particularly for researchers without a strong background in mathematics and statistics, who might require time and expertise to learn how to select the most adequate algorithm and how to use it effectively. In addition, chemometrics often requires extensive data preprocessing, including filtering, smoothing, and baseline correction, which can be time-consuming and introduce bias if not done carefully. For instance, when applying complex models to high-dimensional data, there's the risk of overfitting, which can lead to models that perform well on the training data but generalize poorly to new data leading to inaccurate results. In this regard, artificial Intelligence (AI) can provide valuable assistance when using chemometrics with second or high-order QDs-based fluorescence data either in automating and optimizing data pre-processing tasks, in extracting relevant information from the high-dimensional fluorescence data, in selecting the most appropriate model for a given dataset, in building predictive models that can anticipate trends or behaviors based on historical fluorescence data, which could be a significant advantage when dealing with evolving QDs chemical processes where the optical properties of the QDs or the fluorescence data evolve over time, etc.

In the realm of chemometric applications for quantum dot photoluminescence data analysis, the potential for new developments extends beyond the scope covered in this review. There is potential to incorporate QDs into innovative sensing platforms, exploring applications beyond traditional fields, such as environmental monitoring, medical diagnostics, or even wearable devices. Furthermore, the improvement of the specificity and sensitivity of sensors based on QDs through surface modifications or the incorporation of new materials has the potential to unveil new possibilities. Furthermore, the exploration of multifunctional QDs that not only exhibit photoluminescence but also possess other unique properties, like photocatalytic capabilities, also seems promising. This could lead to the development of advanced, versatile nanomaterials for a broader range of applications. Collaborations between researchers specializing in QDs and experts in diverse fields, such as biology, medicine, or environmental science, could foster interdisciplinary innovations. Moreover, efforts directed towards improving the scalability and cost-effectiveness of QD synthesis methods would be significant. This could facilitate the large-scale production of QDs for practical applications. Finally, exploring eco-friendly and sustainable approaches to QD synthesis aligns with the growing emphasis on green nanotechnology.

In conclusion, although this area does not seem very appealing, and it may look intimidating due to the complexity of the involved mathematics, the advantages that it provides and the increasing number of chemometric tools that are made available to the analytical chemist may

change its perspectives in the incoming years. It will be important to make efforts to create user-friendly software tools that facilitate the widespread adoption of chemometric methods among researchers in quantum dot studies.

CRediT authorship contribution statement

Rafael C. Castro: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Validation, Writing – original draft. **Ricardo N.M.J. Páscoa:** Conceptualization, Formal analysis, Validation, Methodology, Writing – original draft, Visualization, Investigation, Writing – review & editing, Supervision. **M. Lúcia M.F.S. Saraiva:** Funding acquisition, Validation, Supervision. **João L.M. Santos:** Funding acquisition, Validation, Writing – review & editing, Supervision. **David S.M. Ribeiro:** Conceptualization, Formal analysis, Validation, Methodology, Writing – original draft, Visualization, Investigation, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

Rafael C. Castro thanks FCT/MCTES (Fundação para a Ciência e Tecnologia and Ministério da Ciência, Tecnologia e Ensino Superior) and ESF (European Social Fund) through NORTE 2020 (Programa Operacional Região Norte) for his PhD grant ref. 2020.08465.BD. Ricardo N. M. J. Páscoa and David S. M. Ribeiro thank FCT (Fundação para a Ciência e Tecnologia) for funding through program DL 57/2016 – Norma transitória. This work received financial support from PT national funds (FCT/MCTES, Fundação para a Ciência e Tecnologia and Ministério da Ciência, Tecnologia e Ensino Superior) through the projects UIDB/50006/2020 and UIDP/50006/2020.

References

- [1] J.L.M. Santos, J.X. Soares, S.S.M. Rodrigues, D.S.M. Ribeiro, Semiconductor Quantum Dots in Chemical Analysis, in: Handbook of Smart Materials in Analytical Chemistry, 2019, pp. 309-343.
- [2] D.S.M. Ribeiro, R.C. Castro, J.X. Soares, J.L.M. Santos, *Microchem. J.* 155 (2020), 104728.
- [3] I.V. Martynenko, A.P. Litvin, F. Purcell-Milton, A.V. Baranov, A.V. Fedorov, Y. K. Gun'ko, *J. Mater. Chem. B* 5 (2017) 6701–6727.
- [4] S. Pandey, D. Bodas, *Adv. Colloid Interface Sci.* 278 (2020), 102137.
- [5] J. Fan, N. Li, F. Wang, Y. Lv, Q. Jin, M. Zhao, Y. Shi, R. Wu, H. Shen, L.S. Li, *Sens. Actuators B* 375 (2023), 132888.
- [6] Y. Lv, J. Fan, M. Zhao, R. Wu, L.S. Li, *Nanoscale* 15 (2023) 5560–5578.
- [7] Z.W. Heng, W.C. Chong, Y.L. Pang, C.H. Koo, *J. Environ. Chem. Eng.* 9 (2021), 105199.
- [8] B. Bajorowicz, M.P. Kobylański, A. Gołbiewska, J. Nadolna, A. Zaleska-Medynska, A. Malankowska, *Adv. Colloid Interface Sci.* 256 (2018) 352–372.
- [9] S.S.M. Rodrigues, D.S.M. Ribeiro, J.X. Soares, M.L.C. Passos, M.L.M.F.S. Saraiva, J.L.M. Santos, *Coord. Chem. Rev.* 330 (2017) 127–143.
- [10] R.C. Castro, M.L.M.F.S. Saraiva, J.L.M. Santos, D.S.M. Ribeiro, *Coord. Chem. Rev.* 448 (2021), 214181.
- [11] R.C. Castro, D.S.M. Ribeiro, J.L.M. Santos, *Coord. Chem. Rev.* 429 (2021), 213637.
- [12] E.H.M. Sakho, O.S. Oluwafemi, Chapter 11 - Quantum dots for solar cell applications, in: S. Thomas, E.H.M. Sakho, N. Kalarikkal, S.O. Oluwafemi, J. Wu (Eds.), *Nanomaterials for Solar Cell Applications*, Elsevier, 2019, pp. 377–415.
- [13] Z. Yang, M. Gao, W. Wu, X. Yang, X.W. Sun, J. Zhang, H.-C. Wang, R.-S. Liu, C.-Y. Han, H. Yang, W. Li, *Mater. Today* 24 (2019) 69–93.
- [14] Y.B. Monakhova, I.Y. Goryacheva, *TrAC Trends Anal. Chem.* 82 (2016) 164–174.
- [15] M. Cardoso Dos Santos, W.R. Algar, I.L. Medintz, N. Hildebrandt, *TrAC Trends Anal. Chem.* 125 (2020), 115819.
- [16] H. Jin, X. Jiang, Z. Sun, R. Gui, *Coord. Chem. Rev.* 431 (2021), 213694.

- [17] A. Bigdeli, F. Ghasemi, S. Abbasi-Moayed, M. Shahrajabian, N. Fahimi-Kashani, S. Jafarinejad, M.A. Farahmand Nejad, M.R. Hormozi-Nezhad, *Anal. Chim. Acta*, **1079** (2019) 30–58.
- [18] R. Gui, H. Jin, X. Bu, Y. Fu, Z. Wang, Q. Liu, *Coord. Chem. Rev.* **383** (2019) 82–103.
- [19] R.C. Castro, R.N.M.J. Páscoa, M.L.M.F.S. Saraiva, J.L.M. Santos, D.S.M. Ribeiro, *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **267** (2022), 120592.
- [20] R.C. Castro, J.X. Soares, D.S.M. Ribeiro, J.L.M. Santos, *Sens. Actuators B* **296** (2019), 126665.
- [21] G. Rousserie, A. Sukhanova, K. Even-Desrumeaux, F. Fleury, P. Chames, D. Baty, V. Oleinikov, M. Pluot, J.H.M. Cohen, I. Nabiev, *Crit. Rev. Oncol. Hematol.* **74** (2010) 1–15.
- [22] B. Hemmateenejad, M.R. Hormozi Nezhad, *J. Phys. Chem. C* **112** (2008) 18321–18324.
- [23] P. Mukherjee, S.J. Lim, T.P. Wrobel, R. Bhargava, A.M. Smith, *J. Am. Chem. Soc.* **138** (2016) 10887–10896.
- [24] D. Mutavdžić, J. Xu, G. Thakur, R. Triulzi, S. Kasas, M. Jeremić, R. Leblanc, K. Radotić, *Analyst* **136** (2011) 2391–2396.
- [25] E. Navarrete S.J. Román S.V. Rojas C.R. Henríquez N.R. Schrebler G.R. Córdova O.M. Bravo M.E. Muñoz C, *Arab. J. Chem.*, **12** (2019) 5103–5110.
- [26] S. Wold, *Technometrics* **16** (1974) 1–11.
- [27] P. Oliveri, M. Forina, Chapter 2 – Data analysis and chemometrics, in: Y. Picó (Ed.), *Chemical Analysis of Food: Techniques and Applications*, Academic Press, Boston, 2012, pp. 25–57.
- [28] A.C. Olivieri, *Introduction to Multivariate Calibration: A Practical Approach*, Springer, 2018.
- [29] S.D. Brown, R. Tauler, B. Walczak, *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*, Elsevier, 2020.
- [30] P. Geladi, *Spectrochim. Acta B Atmos. Spectrosc.* **58** (2003) 767–782.
- [31] L.A. Berrueta, R.M. Alonso-Salces, K. Héberger, *J. Chromatogr. A* **1158** (2007) 196–214.
- [32] M. Barker, W. Rayens, *J. Chemometr.* **17** (2003) 166–173.
- [33] S. Wold, M. Sjöström, *SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy*, in: *Chemometrics: Theory and Application*, American Chemical Society, 1977, pp. 243–282.
- [34] L. Xu, S.-M. Yan, C.-B. Cai, X.-P. Yu, *Chemom. Intel. Lab. Syst.* **126** (2013) 1–5.
- [35] S. Mussa Farkhani, A. Valizadeh, *IET Nanobiotechnol.* **8** (2014) 59–76.
- [36] Y. Pu, F. Cai, D. Wang, J.-X. Wang, J.-F. Chen, *Ind. Eng. Chem. Res.* **57** (2018) 1790–1802.
- [37] J. Jaumot, A. de Juan, R. Tauler, *Chemom. Intel. Lab. Syst.* **140** (2015) 1–12.
- [38] D.S.M. Ribeiro, G.C.S. de Souza, A. Melo, J.X. Soares, S.S.M. Rodrigues, A. N. Araújo, M.C.B.S.M. Montenegro, J.L.M. Santos, *J. Mater. Sci.* **52** (2017) 3208–3224.
- [39] J.X. Soares, K.D. Wegner, D.S.M. Ribeiro, A. Melo, I. Häusler, J.L.M. Santos, U. Resch-Genger, *Nano Res.* **13** (2020) 2438–2450.
- [40] S. Yang, W. Fan, T. Xiong, D. Wang, K. Qin, M. Fan, Z. Gong, *Chemom. Intel. Lab. Syst.* **182** (2018) 124–130.
- [41] A.C. Olivieri, *Anal. Chem.* **80** (2008) 5713–5720.
- [42] K.S. Booksh, B.R. Kowalski, *Anal. Chem.* **66** (1994) 782A–791A.
- [43] S. Mas, A. de Juan, R. Tauler, A.C. Olivieri, G.M. Escandar, *Talanta* **80** (2010) 1052–1067.
- [44] G.M. Escandar, A.C. Olivieri, N.M. Faber, H.C. Goicoechea, A. Muñoz de la Peña, R.J. Poppi, *TrAC Trends Anal. Chem.* **26** (2007) 752–765.
- [45] G. Ahmadi, R. Tauler, H. Abdollahi, *Chemom. Intel. Lab. Syst.* **142** (2015) 143–150.
- [46] H.-L. Wu, T. Wang, R.-Q. Yu, *TrAC Trends Anal. Chem.* **130** (2020), 115954.
- [47] P. Mishra, J.-M. Roger, D. Jouan-Rimbaud-Bouveresse, A. Biancolillo, F. Marini, A. Nordon, D.N. Rutledge, *TrAC Trends Anal. Chem.* **137** (2021), 116206.
- [48] R.C. Castro, R.N.M.J. Páscoa, M.L.M.F.S. Saraiva, J.L.M. Santos, D.S.M. Ribeiro, *Biosensors* **13** (2023) 437.
- [49] L. Liu, Y. Fan, H. Fu, F. Chen, C. Ni, J. Wang, Q. Yin, Q. Mu, T. Yang, Y. She, *Anal. Chim. Acta* **963** (2017) 119–128.
- [50] R.C. Castro, D.S.M. Ribeiro, R.N.M.J. Páscoa, J.X. Soares, S.J. Mazivila, J.L.M. Santos, *Anal. Chim. Acta* **1114** (2020) 29–41.
- [51] J. Liu, G. Li, X. Yang, K. Wang, L. Li, W. Liu, X. Shi, Y. Guo, *Anal. Chem.* **87** (2015) 876–883.
- [52] F. Ghasemi, M.R. Hormozi-Nezhad, *Talanta* **201** (2019) 230–236.
- [53] Y.V. Zontov, O.Y. Rodionova, S.V. Kucheryavskiy, A.L. Pomerantsev, *Chemom. Intel. Lab. Syst.* **167** (2017) 23–28.
- [54] A.L. Pomerantsev, O.Y. Rodionova, *J. Chemom.* **28** (2014) 429–438.
- [55] A.L. Pomerantsev, *J. Chemom.* **22** (2008) 601–609.
- [56] A.L. Pomerantsev, O.Y. Rodionova, *J. Chemom.* **28** (2014) 518–522.
- [57] S.V. Carneiro, V.H.R. de Queiroz, A.A.C. Cruz, L.M.U.D. Fechine, J.C. Denardin, R.M. Freire, R.F. do Nascimento, P.B.A. Fechine, *Sensors Actuators B: Chem.*, **301** (2019) 127149.
- [58] S.V. Carneiro, M.H.B. Holanda, H.O. Cunha, J.J.P. Oliveira, S.M.A. Pontes, A.A. C. Cruz, L.M.U.D. Fechine, T.A. Moura, A.R. Paschoal, R.A. Zambelli, R.M. Freire, P.B.A. Fechine, *J. Photochem. Photobiol. A Chem.* **411** (2021), 113198.
- [59] M.E. Lombardo, D. Benetti, V. La Carrubba, F. Rosei, *ECS Meeting Abstracts*, **MA2020-01** (2020) 1087.
- [60] G. Liu, D. Kong, J. Han, R. Zhou, Y. Gao, Z. Wu, L. Zhao, C. Wang, L. Wang, G. Lu, *Sens. Actuators B* **342** (2021), 129963.
- [61] Kurt Varmuza, P. Filzmoser, *Classification*, in: P.F. Kurt Varmuza (Ed.) *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press, Boca Raton, 2009.
- [62] J. Zhu, F. Zhu, L. Li, L. Cheng, L. Zhang, Y. Sun, X. Wan, Z. Zhang, *Food Chem.* **298** (2019), 125046.
- [63] O. Hu, L. Xu, H. Fu, T. Yang, Y. Fan, W. Lan, H. Tang, Y. Wu, L. Ma, D. Wu, Y. Wang, Z. Xiao, Y. She, *Anal. Chim. Acta* **1008** (2018) 103–110.
- [64] L. Xu, L. Wei, Q. Shi, C. Cai, H.Y. Fu, Y.B. She, *Food Anal. Meth.* **12** (2019) 2614–2622.
- [65] L. Xu, Q. Shi, D. Lu, L. Wei, H.-Y. Fu, Y. She, S. Xie, *Microchem. J.* **157** (2020), 105105.
- [66] L. Xu, D. Lu, Q. Shi, H. Chen, S. Xie, G. Li, H.-Y. Fu, Y.-B. She, *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **221** (2019), 117212.
- [67] Q. Wang, Q. Yin, Y. Fan, L. Zhang, Y. Xu, O. Hu, X. Guo, Q. Shi, H. Fu, Y. She, *Talanta* **199** (2019) 46–53.
- [68] N.N. Divyanin, A.V. Razina, E.A. Rukosueva, A.V. Garmash, M.K. Bekdemishev, *Microchem. J.* **135** (2017) 48–54.
- [69] Z. Jiao, P. Zhang, H. Chen, C. Li, L. Chen, H. Fan, F. Cheng, *Sens. Actuators B* **295** (2019) 110–116.
- [70] P. Yang, J. Li, P. Li, C. Hou, D. Huo, Y. Yang, S. Zhang, C. Shen, *Anal. Methods* **11** (2019) 4842–4850.
- [71] Y. Fan, L. Zhang, J. Jia, H. Chen, H. Fu, Y. She, *Sens. Actuators B* **319** (2020), 128260.
- [72] Z. Wang, C. Xu, Y. Lu, X. Chen, H. Yuan, G. Wei, G. Ye, J. Chen, *Sens. Actuators B* **241** (2017) 1324–1330.
- [73] Y. Wu, X. Liu, Q. Wu, J. Yi, G. Zhang, *Sens. Actuators B* **246** (2017) 680–685.
- [74] D.M. Mayder, C.M. Tonge, G.D. Nguyen, M.V. Tran, G. Tom, G.H. Darwish, R. Gupta, K. Lix, S. Kamal, W.R. Algar, S.A. Burke, Z.M. Hudson, *J. Am. Chem. Soc.* **143** (2021) 16976–16992.
- [75] J. Li, H. Huang, X. Sun, D. Song, J. Zhao, D. Hou, Y. Li, *Anal. Methods* **11** (2019) 3168–3174.
- [76] H. Kang, L. Lin, M. Rong, X. Chen, *Talanta* **129** (2014) 296–302.
- [77] W. Jing, Y. Lu, G. Yang, F. Wang, L. He, Y. Liu, *Anal. Chim. Acta* **985** (2017) 175–182.
- [78] S. Abbasi-Moayed, H. Golmohammadi, A. Bigdeli, M.R. Hormozi-Nezhad, *Analyst* **143** (2018) 3415–3424.
- [79] S. Abbasi-Moayed, H. Golmohammadi, M.R. Hormozi-Nezhad, *Nanoscale* **10** (2018) 2492–2502.
- [80] H. Chen, O. Hu, Y. Fan, L. Xu, L. Zhang, W. Lan, Y. Hu, X. Xie, L. Ma, Y. She, H. Fu, *Food Chem.* **327** (2020), 127075.
- [81] Y. Fan, L. Liu, D. Sun, H. Lan, H. Fu, T. Yang, Y. She, C. Ni, *Anal. Chim. Acta* **916** (2016) 84–91.
- [82] J.-H. Jiang, R.J. Berry, H.W. Siesler, Y. Ozaki, *Anal. Chem.* **74** (2002) 3555–3565.
- [83] H.-Y. Fu, H.-L. Wu, H.-Y. Zou, L.-J. Tang, L. Xu, C.-B. Cai, J.-F. Nie, R.-Q. Yu, *Anal. Methods* **2** (2010) 282–288.
- [84] J. Suykens, 3.26 – Kernel Methods, in: S. Brown, R. Tauler, B. Walczak (Eds.), *Comprehensive Chemometrics (second Edition)*, Elsevier, Oxford, 2009, pp. 555–566.
- [85] J.A.K. Suykens, *J. Vandewalle, Neural Process. Lett.* **9** (1999) 293–300.
- [86] H. Dai, J. Jia, Y. Fan, H. Chen, S. Wang, C. Shen, A. Li, L. Lu, C. Zhou, H. Fu, Y. She, *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **252** (2021), 119513.
- [87] D.S.M. Ribeiro, R.C. Castro, R.N.M.J. Páscoa, J.X. Soares, S.S.M. Rodrigues, J.L.M. Santos, *J. Lumin.* **207** (2019) 386–396.
- [88] D.B. Bittar, D.S.M. Ribeiro, R.N.M.J. Páscoa, J.X. Soares, S.S.M. Rodrigues, R. C. Castro, L. Pezza, H.R. Pezza, J.L.M. Santos, *Talanta* **174** (2017) 572–580.
- [89] A. Barati, M. Shamsipur, H. Abdollahi, *Anal. Methods* **6** (2014) 6577–6584.
- [90] A.C. Olivieri, G.M. Escandar, Chapter 1 – Calibration Scenarios, in: A.C. Olivieri, G.M. Escandar (Eds.), *Practical Three-Way Calibration*, Elsevier, Boston, 2014, pp. 1–9.
- [91] A.C. Olivieri, G.M. Escandar, Chapter 2 – Data Properties, in: A.C. Olivieri, G.M. Escandar (Eds.), *Practical Three-Way Calibration*, Elsevier, Boston, 2014, pp. 11–26.
- [92] J.W. Rutherford, N. Dawson-Elli, A.M. Manicone, G.V. Korshin, I.V. Novoselov, E. Seto, J.D. Posner, *Atmos. Environ.* **220** (2020), 117065.
- [93] K. Glowacz, M. Drozd, P. Ciosek-Skibińska, *Microchim. Acta* **188** (2021) 343.
- [94] R.C. Castro, R.N.M.J. Páscoa, M.L.M.F.S. Saraiva, J.L.M. Santos, D.S.M. Ribeiro, *Anal. Chim. Acta*, **1188** (2021) 339174.
- [95] A.C. Olivieri, G.M. Escandar, A.M.D.L. Peña, *TrAC Trends Anal. Chem.* **30** (2011) 607–617.
- [96] A. de Juan, R. Tauler, *J. Chemom.* **15** (2001) 749–771.
- [97] R. Bro, *Chemom. Intel. Lab. Syst.* **38** (1997) 149–171.
- [98] R. Tauler, *Chemom. Intel. Lab. Syst.* **30** (1995) 133–146.
- [99] A.V. Schenone, M.J. Culzoni, A.D. Campiglia, H.C. Goicoechea, *Anal. Bioanal. Chem.* **405** (2013) 8515–8523.
- [100] J.M.M. Leitão, H. Gonçalves, C. Mendonça, J.C.G. Esteves da Silva, *Anal. Chim. Acta* **628** (2008) 143–154.
- [101] H. Gonçalves, C. Mendonça, J.C.G. Esteves da Silva, *J. Fluoresc.* **19** (2008) 141.
- [102] B.B. Campos, M. Algarra, J.C.G. Esteves da Silva, *J. Fluoresc.* **20** (2010) 143–151.
- [103] J.M.M. Leitão, R. Tauler, J.C.G.E. da Silva, *J. Fluoresc.* **21** (2011) 1987–1996.
- [104] S. Gholami, M. Kompany-Zareh, *Anal. Bioanal. Chem.* **405** (2013) 6271–6280.
- [105] E. Peré-Trepát, A. Ginebreda, R. Tauler, *Chemom. Intel. Lab. Syst.* **88** (2007) 69–83.
- [106] L.R. Tucker, *Psychometrika* **31** (1966) 279–311.
- [107] R. Tauler, S. Lacorte, M. Guillardou, R. Cespedes, P. Viana, D. Barceló, *Environ. Toxicol. Chem.* **23** (2004) 565–575.
- [108] P. Geladi, *Chemom. Intel. Lab. Syst.* **7** (1989) 11–30.
- [109] A. de Juan, M. Maeder, M. Martínez, R. Tauler, *Chemom. Intel. Lab. Syst.* **54** (2000) 123–141.
- [110] Y.-M. Neuhold, M. Maeder, *J. Chemom.* **16** (2002) 218–227.

- [111] M. Zabadaj, P. Ciosek-Skibińska, *Sensors* 19 (2019) 3655.
- [112] M.A. Malik, E. Gatto, S. Macken, C. DiNatale, R. Paolesse, A. D'Amico, I. Lundström, D. Filippini, *Anal. Chim. Acta* 635 (2009) 196–201.
- [113] H. Abdollahi, M. Shamsipur, A. Barati, *Spectrochim. Acta Part A: Mol. Biomol. Spectros.* 127 (2014) 137–143.
- [114] R.C. Castro, R.N.M.J. Páscoa, M.L.M.F.S. Saraiva, R.A.S. Lapa, J.O. Fernandes, S. C. Cunha, J.L.M. Santos, D.S.M. Ribeiro, *Microchem. J.* 185 (2023), 108300.
- [115] R.C. Castro, R.N.M.J. Páscoa, M.L.M.F.S. Saraiva, S.C. Cunha, J.L.M. Santos, D.S. M. Ribeiro, *Food Control* 153 (2023), 109934.
- [116] X. Yuan, D. Zhang, X. Zhu, H. Liu, B. Sun, *Food Chem.* (2020), 128299.