

Comparative analysis of methods for reconstructing gas concentrations in multicomponent gas mixtures from data of laser photoacoustic gas analyzer

M.L. Belov, V.A. Gorodnichev, D.B. Dobritsa, and V.I. Kozintsev

*Scientific Research Institute of Radioelectronics and Laser Equipment, Moscow
Bauman State Technical University, Moscow*

Received July 26, 1999

A comparison of algorithms for reconstructing gas concentrations in multicomponent gas mixtures is presented. We describe the signal processing procedures in application to DIAL methods that are based on the regularization methods of solving the ill-posed inverse problems. Examples of processing the experimental data are given. It is shown that the use of regularization methods in data processing gives rather low errors in gas concentrations reconstructed. Different methods for selection of the regularization constant have been used. In most cases the smallest errors are achieved with the use of such methods as selection of the quasi-optimal regularization constant, modified method of discrepancy, and the method of discrepancy with additional independent information invoked.

Laser sensing methods are most promising for real-time remote and local gas analysis of multicomponent mixtures.¹⁻⁵ Gas analysis of a mixture can be divided into two steps: mixture identification (determination of whether one or several gases are present in the mixture) and measurement of its composition (obtaining quantitative information on concentrations of the mixture components). This paper is devoted to the second stage of gas analysis, namely, to the development of processing algorithms for obtaining the quantitative information on gas concentrations in a multicomponent mixture from the measurement data.

It is assumed that the composition of a gas mixture is known, absorption spectra of the gaseous constituents are known, and the optimal spectral measurement channels are determined (this is, certainly, a separate complicated problem; see, for example, Refs. 6 and 7). This paper continues our previous studies.⁸⁻¹⁰ It is devoted to comparative analysis of deterministic regularization methods (with different methods used for selection of the regularization constant) and statistical regularization methods as applied to the problem of reconstruction of gas concentrations from the data acquired with a laser photoacoustic gas analyzer.

The problem on reconstruction of the component concentrations from data measured with a laser photoacoustic gas analyzer is reduced to solution of a set of linear algebraic equations of the following form⁵:

$$\begin{cases} K_a(\lambda_1) + \sum_j^N C_j K_j(\lambda_1) = y(\lambda_1) \\ \dots \\ K_a(\lambda_M) + \sum_j^N C_j K_j(\lambda_M) = y(\lambda_M) \end{cases}, \quad (1)$$

where $y(\lambda_i)$ is the reduced signal measured at the wavelength λ_i ; $K_a(\lambda_i)$ is the coefficient of non-selective absorption at the wavelength λ_i ; $K_j(\lambda_i)$ is the absorption coefficient of the j th component at the wavelength λ_i ; C_j is the concentration of the j th component of the gas mixture; N is the total number of gas components in the mixture under analysis.

The parameters of the equations (1) sought are concentrations of the gas components and coefficients of non-selective absorption. The parameters $K_a(\lambda_i)$ only slightly depend on the wavelength. Therefore, it is usually believed that if the pair of channels chosen for the measurements is spectrally close enough, then the coefficients K_a for every pair can be considered constant. In this case, if we have \tilde{l} spectral channels for sounding of the gas mixture, the information from $\tilde{l}/2$ channels is needed for determination of the coefficients K_a .

The right-hand side of the system of equations (1) is always known accurate to a random error. As a result, the inverse operator is not stable. The situation can be resolved by using the processing methods based on the regularization of the solutions sought.^{11,12}

In the matrix form the system of equations (1) can be presented as^{11,12}

$$W\mathbf{x} = \mathbf{K}_a + K\mathbf{a} = \mathbf{y}, \quad (2)$$

where W is the matrix of the system (1); \mathbf{x} is the sought vector (N components of this vector correspond to gas concentrations); \mathbf{K}_a is the vector of coefficients of non-selective absorption; K is the matrix of absorption coefficients of the mixture components; \mathbf{a} is the vector of gas concentrations; \mathbf{y} is the right-hand side vector of the system of equations (1).

With the use of deterministic regularization methods for the system (1), the regularized solution is determined as^{11,12}:

$$\mathbf{x}_\alpha = (W^T W + \alpha E)^{-1} W^T \mathbf{y}, \quad (3)$$

where α is the regularization constant; \mathbf{x}_α is the regularized solution of the system (1); A is the unit matrix; $(W^T W + \alpha E)^{-1}$ is the matrix inverse to the matrix $W^T W + \alpha E$; T denotes transposition.

The main difficulty in constructing of regularized solutions is the selection of the regularization constant α . For this purpose several methods were used.

Let us have, instead of the exact right-hand side \mathbf{y} of Eq. (1), some its value $\tilde{\mathbf{y}}$ satisfying the condition

$$\| \mathbf{y} - \tilde{\mathbf{y}} \| \leq \delta,$$

where $\| \mathbf{y} \|$ is vector norm.

Introduce the function^{11,14}:

$$r(\alpha) = \rho(W\mathbf{x}_\alpha, \tilde{\mathbf{y}}). \tag{4}$$

The numerical solution of the equation

$$r(\alpha) = \delta^2 \tag{5}$$

gives the value of the regularization constant found on the basis of the discrepancy principle.

By the second method, the quasi-optimal regularization constant was found from the condition¹¹:

$$\inf_{\alpha} \| \eta_{\alpha} \|^2, \tag{6}$$

where $\eta_{\alpha} = \alpha \frac{d\mathbf{x}_{\alpha}}{d\alpha}$.

As the third method of constant selection, the modified method of discrepancy was used. Measurement noise in different spectral channels can differ widely. Therefore, for each gas its own regularization constant was used. This constant was sought in two steps. First, the regularization constant α_i for each gas was determined from Eq. (5) written for the spectral channel, in which this gas has the peak of absorption. Next, the final value of the regularization constant was found from the condition (6) in the vicinity of α_i determined at the first stage [in the interval between α_i and α found from Eq. (5) for all spectral channels].

The fourth method employs a more strict equation for the quasi-optimal regularization constant. In this equation α is determined from the condition¹¹:

$$\inf_{\alpha} \sup_y \| \eta_{\alpha} \|^2, \tag{7}$$

where \sup_y is sought over the set of realizations (calculations were made of 10 realizations) of the right-hand side of Eq. (1).

As the fifth method for selecting the parameter α , the method of ratios was used.¹¹ In this method, the regularization constant is taken equal to the least value from the set $\{\alpha_i\}$ of α values at which the functional

$$r_{\text{rel}}(\alpha) = \frac{r_1(\alpha)}{r(\alpha)}, \tag{8}$$

where $r_1(\alpha) = \| W\eta_{\alpha} - (W\mathbf{x}_{\alpha} - \tilde{\mathbf{y}}) \|^2$, is minimum.

In the sixth method the parameter α was selected by the minimum discrepancy with the additional independent information invoked. The parameter α is determined from the condition¹³

$$\min_{\alpha} \| \tilde{\mathbf{y}}_{\alpha} - \tilde{\mathbf{y}}_{\delta} \|, \tag{9}$$

where $\tilde{\mathbf{y}}_{\alpha} = W\mathbf{x}_{\alpha}$; $\tilde{\mathbf{y}}_{\delta}$ is the vector of the right-hand sides of Eq. (1) obtained from additional independent measurements.

If a series of measurements (set of realizations) is available, then \mathbf{x}_{α} and $\tilde{\mathbf{y}}_{\alpha}$ are determined from one realization, while other realizations are used as additional independent information.¹³ For a series of 10 experiments, we used the results of one experiment to determine \mathbf{x}_{α} and $\tilde{\mathbf{y}}$

α , and the results of other experiments were used to find the series mean value of $\tilde{\mathbf{y}}_{\delta}$.

The seventh and eighth methods use the statistical method of selection of the regularization constant based on the criterion of optimal regularized solution. The statistically regularized solution \mathbf{x}_{α} has the form^{11,14}

$$\mathbf{x}_{\alpha} = (\alpha G + W^T V_{\xi}^{-1} W)^{-1} W^T V_{\xi}^{-1} \tilde{\mathbf{y}}, \tag{10}$$

where G is the non-negative definite symmetric matrix.

Such α_w is taken as the value of the parameter α , which does not statistically contradict the optimal character of the regularized solution¹¹

$$V_e(\alpha_w) = V_{\xi} P^T(\alpha_w), \tag{11}$$

where V_e is the matrix of second moments of the discrepancy vector; $V_e = M[\mathbf{e} \times \mathbf{e}^T]$; \mathbf{e} is the discrepancy vector, $\mathbf{e} = \tilde{\mathbf{y}} - W\mathbf{x}_{\alpha}$; $\mathbf{x}_{\alpha} = T\tilde{\mathbf{y}}$; T is the regularizing operator; M denotes mathematical expectation; P is the discrepancy operator, $P = E - WT$; E is the unit matrix.

The hypothesis on the optimal character of the solution is rejected, if the identity (11) is not satisfied because of systematic errors associated with non optimal character of the selected regularization constants, rather than random errors due to the finite size of a sample. Consequently, if at $\alpha = \alpha_w$ the hypothesis (11) is accepted, then α_w can be considered as an estimate of the optimal value of the regularization constant.¹¹

To check the hypothesis (11), the statistics is introduced¹¹:

$$r_w(\alpha) = \mathbf{e}^T(\alpha) K^{-1}(\alpha) \mathbf{e}(\alpha), \tag{12}$$

where

$$K(\alpha) = V_{\xi} P^T(\alpha).$$

For the values of α which do not contradict Eq. (12), the properties of the statistics $r_w(\alpha)$ allow the distribution of $r_w(\alpha)$ to be approximated by the χ^2 -distribution with \tilde{l} degrees of freedom.¹¹ Then the check of the hypothesis (11) is reduced to addressing the question on whether or not $r_w(\alpha)$ satisfies the χ^2 -distribution with \tilde{l} degrees of freedom. This hypothesis is checked in the following way. Let us construct the interval

$$\theta_M(\beta) = \left[\vartheta_M\left(\frac{\beta}{2}\right), \vartheta_M\left(1 - \frac{\beta}{2}\right) \right], \tag{13}$$

where $\vartheta_M(\beta/2)$ is the quantile of the χ^2 -distribution of the level $\beta/2$; β is the probability of error of the first kind (the probability of rejecting the hypothesis by criterion in the case when the hypothesis is true).

If the value of $r_w(\alpha)$ falls within the interval $\theta_M(\beta)$, then the hypothesis (11) can be taken with the probability of first-kind error equal to β , and the value of α at which $r_w(\alpha) \in \theta_M(\beta)$ is taken as the value of the regularization constant.

If $\tilde{l} \leq 30$, then values of quantiles can be found in the table of probabilities of the χ^2 -distribution.¹⁵ At $\tilde{l} > 30$ the χ^2 -distribution is well approximated by the normal distribution with the mathematical expectation \tilde{l} and variance $2\tilde{l}$.

Thus, the solution (10) and the procedure used for determining the regularization constant give the sought vector \mathbf{x}_α . The difference between the vector \mathbf{x}_α and the true solution \mathbf{x} can be presented as follows^{11,14}

$$\mathbf{x}_\alpha - \mathbf{x} = \mathbf{b}_\alpha + \boldsymbol{\xi}_\alpha.$$

Here the random vector $\boldsymbol{\xi}_\alpha$ has zero mean and determines the effect of measurement noise on the random component of the error in the regularized solution. The vector \mathbf{b}_α characterizes the systematic error (bias) of the regularized solution and can be found from the following equation^{11,14}:

$$(W^T V_\xi^{-1} W + \alpha G) \mathbf{b}_\alpha = -\alpha G \mathbf{x}. \quad (14)$$

Since the true solution \mathbf{x} entering into the right-hand side of Eq. (14) is unknown, the vector \mathbf{b}_α can be only estimated as $\hat{\mathbf{b}}_\alpha$ with the regularized solution \mathbf{x}_α substituted, in Eq. (14), for \mathbf{x} . Thus determined vector $\hat{\mathbf{b}}_\alpha$ allows a new solution $\tilde{\mathbf{x}}_\alpha$ to be introduced, which can partially compensate for the bias of the solution \mathbf{x}_α (Refs. 11 and 14):

$$\tilde{\mathbf{x}}_\alpha = \mathbf{x}_\alpha - \hat{\mathbf{b}}_\alpha. \quad (15)$$

In calculations it was assumed that the measurement noise is uncorrelated and the noise correlation function has the form $V_\xi = \sigma_\xi^2 E$ (where σ_ξ^2 is the variance of noise, and E is the unit matrix). The matrix V_ξ^{-1} was used as the matrix G .

To check up the efficiency of the algorithms and to estimate the accuracy of reconstruction of gas concentrations in multicomponent mixtures, we used numerical simulation and processing of the photoacoustic measurement data. Operation of the algorithms for reconstructing gas concentrations was numerically simulated for multicomponent mixtures composed of three to ten components. The results of mathematical simulation showed that in the absence of noise the measurement data on concentrations of gas components can be accurately reconstructed by standard methods of solution of systems of linear algebraic equations and by the regularization procedures. As measurement noise increases, errors of reconstruction increase too.

For multicomponent mixtures (with more than five components), the processing procedures based on construction of the regularized solution give much lower errors. One of the essential advantages of these methods manifests itself in both of the cases when a signal value used in reconstruction comes either from a single measurement or it has been averaged (for example over 10 experiments).

The measurement data acquired with a photoacoustic gas analyzer for gas mixtures composed of three to six components have been processed. The processed results show that for three- and four-component mixtures the reconstruction errors are low and the regularization procedures do not increase the accuracy. For a five-component mixture the accuracy increases a little bit, and for a six-component mixture the accuracy of reconstruction increases significantly when using the regularization procedures.

Figures 1 and 2 show the reconstructed concentrations \tilde{N} of gases (ethylene in Fig. 1 and isopropanol in Fig. 2) contained in the six-component gas mixture of ethylene,

carbon dioxide, ammonia, methanol, ethanol, and isopropanol. Average (over 10 experiments) values of signals from the gas analyzer were used in processing. The measurements were conducted in the spectral channels at 10.532, 10.591, 9.329, 9.282, 9.317, 9.292, 9.658, 9.676, 9.753, 9.733, 10.303, and 10.182 μm wavelengths. The measurement technique and the laser photoacoustic gas analyzer were described in Ref. 8. The variance of the measurement noise in spectral channels was estimated from data of a series of test measurements. In processing the following relative rms values of the measurement noise were set: 0.01734 (10.532 μm), 0.03296 (9.292 μm), 0.02229 (9.676 μm), 0.00905 (9.733 μm), and 0.03537 (10.303 μm).

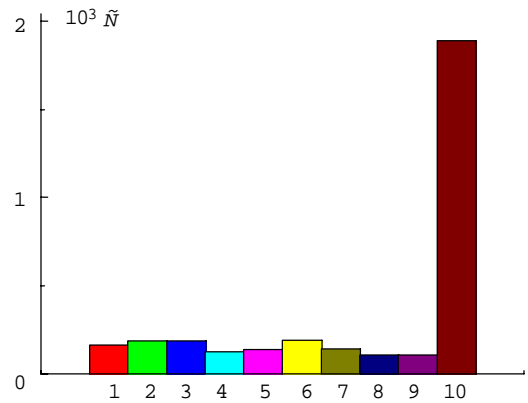


Fig. 1. The ethylene concentration reconstruction from the measurement data.

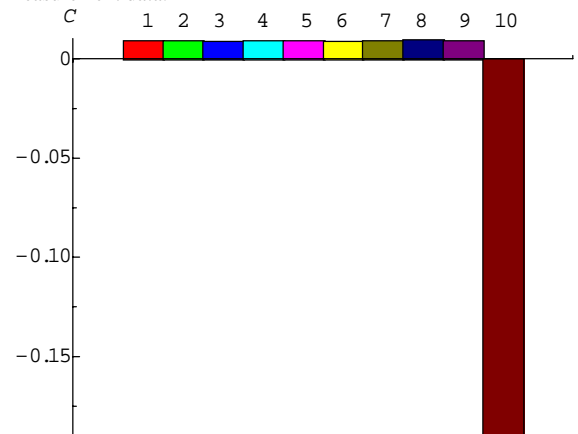


Fig. 2. The isopropanol concentration reconstructed from the measurement data.

Column 1 gives the actual value of the gas component concentration determined from the partial pressure at cell filling. The regularization constant was selected based on the principle of discrepancy (column 3), the modified principle of discrepancy (column 2), selection of the quasi-optimal regularization constant by Eqs. (6) and (7) (columns 4 and 5), the method of ratios (column 6), the principle of discrepancy with using an additional independent information (column 7), and the method of statistical regularization (columns 8 and 9; the regularized solution for column 8 was found from Eq. (10) and that for column 9 was found from Eq. (15)). Column 10

gives the gas concentration reconstructed by the standard method of solution of the system of linear algebraic equations (for solution we used the method of Gauss exclusion).

From Figs. 1 and 2 one can see that even in the case with pre-averaging over 10 experiments the reconstruction errors decrease significantly with the use of the processing procedures based on the regularization methods. The gas concentrations reconstructed by the standard method can not only essentially differ from the actual values, but they can even take negative values.

For a comparison among different algorithms based on deterministic and statistical regularization methods, Fig. 3 shows the relative errors of reconstruction δ (absolute value of the difference between the actual and reconstructed values of concentration divided by the actual concentration) of ethylene (I), ammonia (II), methanol (III), ethanol (IV), and isopropanol (V) (for the same gas mixture as in Figs. 1 and 2). The regularization constant was selected based on the principle of discrepancy (columns 2), modified principle of discrepancy (columns 1), selection of the quasi-optimal regularization constant by Eqs. (6) and (7) (columns 3 and 4), the method of ratios (columns 5), the principle of discrepancy with using an additional independent information (columns 6), the method of statistical regularization [columns 7 and 8; the regularized solution for column 7 was found from Eq. (10), and that for column 8 was found from Eq. (15)]. The errors of reconstruction by the standard method are not shown here, since they are too large.

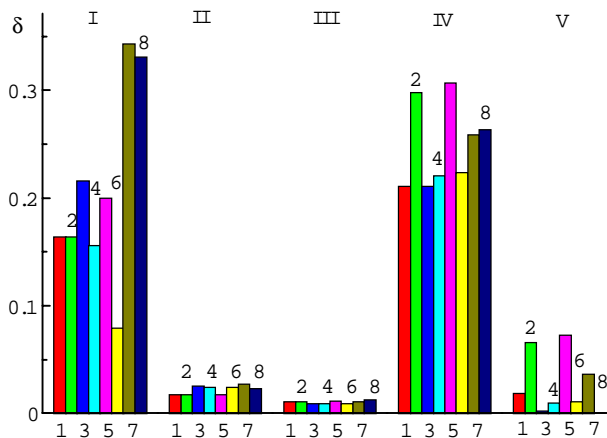


Fig. 3. Relative errors in reconstructed concentrations of the components of a six-component gas mixture.

The results of mathematical simulation and processing of the data show that the use of regularization procedures gives rather low level of errors for

multicomponent mixtures, as gas concentrations are reconstructed by both statistical and deterministic regularization methods. The significant advantage (over standard methods) of the methods based on construction of regularized solutions manifests itself in both of the cases when reconstruction uses the signal value obtained either in a single measurement or has been pre-averaged. The different methods of selection of the regularization constant can give both the close and widely different values of concentrations for different gases. The reconstruction accuracy also differs widely for different component gases. In most cases the least reconstruction errors are achieved in selecting the quasi-optimal regularization constant, as well as in applying the modified method of discrepancy, and the method of discrepancy with using an additional independent information.

References

1. Yu.S. Makushkin, A.A. Mitsel', and G.S. Khmel'nitskii, *Zh. Prikl. Spektrosk.* **35**, No. 5, 785–790 (1981).
2. R.M. Measures, *Laser Remote Sensing* (John Wiley and Sons, New York, 1987).
3. S.V. Ivanov, V.Ya. Panchenko, and T.B. Razumikhina, *Atmos. Oceanic Opt.* **6**, No. 8, 589–592 (1993).
4. Yu.N. Ponomarev, *Atmos. Oceanic Opt.* **8**, Nos. 1–2, 116–124 (1995).
5. M. Sigrist, M.Yu. Kataev, A.A. Mitsel', et al., *Atmos. Oceanic Opt.* **7**, Nos. 11–12, 795–799 (1994).
6. M.Yu. Kataev and A.A. Mitsel', *Avtometriya*, No. 4, 15–20 (1985).
7. M.Yu. Kataev, A.A. Mitsel', and S.R. Tarasova, *Atm. Opt.* **3**, No. 8, 763–771 (1990).
8. V.I. Kozintsev, *Atmos. Oceanic Opt.* **9**, No. 8, 689–691 (1996).
9. V.I. Kozintsev, *Atmos. Oceanic Opt.* **9**, No. 10, 867–870 (1996).
10. V.V. Avdienko, M.L. Belov, V.A. Gorodnichev, V.I. Kozintsev, and D.B. Dobritsa, *Zh. Prikl. Spektrosk.* **63**, No. 5, 755–759 (1996).
11. Yu.E. Voskoboinikov, N.G. Preobrazhenskii, and A.N. Sedel'nikov, *Mathematical Processing of Experiment in Molecular Gas Dynamics* (Nauka, Novosibirsk, 1984), 238 pp.
12. A.N. Tikhonov and V.Ya. Arsenin, *Methods of Solution of Ill-Posed Problems* (Nauka, Moscow, 1979), 288 pp.
13. I.I. Kochetov, *Zh. Vychisl. Mat. Mat. Fiz.* **16**, No. 2, 499–503 (1976).
14. Yu.E. Voskoboinikov and A.A. Mitsel', *Avtometriya*, No. 2, 67–72 (1982).
15. G.A. Korn and T.M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1961).