Parameter identification for a model of chromatographic column.

Francois JAMES *
Mauricio SEPÚLVEDA †

June 14, 1993

Abstract

We are interested here in identifying the flux in a system of conservation laws. This problem is ill-posed, but we can deal with the indetermination by using an a priori problem as a constrained minimization of a cost function J. A Lagrangian formulation lead us to a formal computation of that gradient of J. We then perform an exact computation of the gradient of the discrete cost function. Finally, we give a few numerical results obtained throught a conjugate gradient method.

Key words : *Non linear hyperbolic systems - Inverse problems - Adjoint state - Constrained optimization - Gradient method.*

1 Introduction

Several separation problems, such as distillation and chromatography, involve matter exchange between two distinct phases in order to separate or analyse complex mixtures. Under several hypotheses, such processes can be modelled by a system of nonlinear hyperbolic conservation laws (see [15, 4]). The nonlinearity appearing in these equations comes precisely from the exchanges between phases, and we shall focus on this aspect in the following.

In particular, we shall assume the temperature to be constant, and also that the velocities of both phases are constant. We are left with the equations of mass conservation, and the nonlinear function involved here is called an *isotherm*: it represents a thermodynamical equilibrium state between the two phases. For our purpose, it will be a vector-valued function H, which gives the quantities in phase 2 as a function of the quantities in phase 1.

There are several models of isotherms: let us mention Sips and Freundlich ([14, 2], and also a very classical model given by Langmuir [8]. This last isotherm is interesting since it has good mathematical properties which allow a thorough study (see [10]). Unfortunately, it is no longer physically relevant as soon as the amount of matter increases: indeed this model does not take into account possible interactions between the components of the mixture.

It is therefore necessary to elaborate more complex models of isotherms. Let us mention several papers on the topic, which use statistical thermodynamics to model the phases,

*Département de Mathématiques, Université d'Orléans, B.P. 6759, F-45067 Orléans Cedex 2, FRANCE et Centre de Mathématiques Appliquées, Ecole Polytechnique.

†Département de Mathématiques, Université de Reims, Moulin de la Housse BP 347 - 51062 Reims Cedex, FRANCE et Centre de Mathématiques Appliquées, Ecole Polytechnique, F-91128 Palaiseau Cedex, FRANCE

and deduce from these models more relevant isotherms: [11, 12, 5]. The trouble with these models is that they depend on a large number of parameters, and are therefore very difficult to identify precisely. Note that this problem of determining the isotherms is very important in practice, and there are several experimental methods to deal with it (see [16], and the quoted references).

From the mathematical point of vue, we shall assume that our processes are modelled by a system of nonlinear conservation equations which we shall describe below. Within this framework, the isotherm appears as a part of the flux of the system, and we are naturally led to the problem of identifying this flux, by comparing the solution of our system to some experimental data. Section 2 gives a precise example of such a comparison, and states the model we are dealing with. In Section 3, we state the problem as a constrained minimization problem, and compute formally the gradient of the cost function. Applying the same ideas to discretized cost function, we obtain in Section 4 a discrete gradient, and apply it as conjugate gradient method to obtain some numerical results in Section 5.

2 The model and the identification problem

2.1 The model

We consider a 1-dimensional diphasic medium (the chromatographic or distillation column), in which phase 1 moves with a velocity $u > 0$, and phase 2 with a velocity $v \leq 0$. The case $v = 0$ would correspond to chromatography, and $v < 0$ to distillation. Let L be the length of the column, and consider a mixture of M components. We assume that

— u and v are constant;

— the temperature is constant;

— the process is quasi-static.

The first two assumptions eliminate the momentum and energy conservation equations. The third means that, at any time, the system constituted by the two phases is a thermodynamical equilibrium (see [15, 4]). In other words, at any time, the quantities in phase 2 are related by the isotherm to the quantities in phase 1. Therefore, the equations of matter conservation can be written

$$(\mathbf{c} + H(\mathbf{c}))_t + (u\mathbf{c} + vH(\mathbf{c}))_z = 0, \qquad (z,t) \in \Omega = (0,L) \times (0,T), \tag{1}$$

where $\mathbf{c} = (c_1, \dots, c_M) \geq 0$ is the vector of quantities of the components in phase 1.

We shall restrict ourselves to the case $v = 0$, that is to the case of chromatography. The case of distillation is indeed more difficult, since we have positive and negative eigenvalues for the Jacobian matrix of (1), and in this context boundary conditions cannot be well defined. If $v = 0$, we associate to (1) the following initial and boundary conditions

$$\mathbf{c}(z,0) = \mathbf{c}_{\text{init}}(z), \quad z \in [0,L], \qquad \mathbf{c}(0,t) = \mathbf{c}^{\text{inj}}(t), \quad t \geq 0, \tag{2}$$

where $\mathbf{c}_{\text{init}}(z)$ and $\mathbf{c}^{\text{inj}}(t)$ are given functions. Namely, $\mathbf{c}_{\text{init}}(z)$ is the *initial state* of the column –which is often a constant– and $\mathbf{c}^{\text{inj}}(t)$ is the *injection*, that is the quantity of matter injected in the column. We shall assume that the injection has a compact support, that is $\mathbf{c}^{\text{inj}}(t) = 0$ for $t \geq T_0 > 0$.

From its thermodynamical origin, the function H has the following property: the eigenvalues of its Jacobian matrix $H'(\mathbf{c})$ are real and positive (see [7, 4]). We can therefore perform the variable change $\mathbf{w} = \mathbf{c} + H(\mathbf{c})$, $\mathbf{w} = (w_1, \dots, w_M)$, where w_i is the total amount of component i in the system. We denote the reciprocal variable change by $\mathbf{c} = g(\mathbf{w})$, and set $F(\mathbf{w}) = ug(\mathbf{w}) + vH(g(\mathbf{w}))$. Thus, the system (1) can be rewritten in a more classical form:

$$\mathbf{w}_t + F(\mathbf{w})_z = 0, \qquad (z,t) \in \Omega = (0,L) \times (0,T), \tag{3}$$

where the injection and initial state become

$$\mathbf{w}(z,0) = \mathbf{w}^{\text{init}}(z), \quad z \in [0,L], \qquad \mathbf{w}(0,t) = \mathbf{w}_{\text{inj}}(t), \quad t \geq 0. \tag{4}$$

The vector-valued function H will be the unknown of our identification problem. As we shall see later, the problem of identifying H as a function, say of class $C^2(\mathbb{R}^n)$, is particulary ill-posed. But the construction of H from thermodynamics can give us several properties of H which help us to restrict the domain of admissible H. From a practical point of view, we can even compute explicit models for the function H, such that its shape is totally determined by a finite number of parameters $e_1, \ldots, e_N$, $N \geq M$. For instance, if we consider a gas-solid isotherm model for a single component mixture (see [5], and [6]), we have the following function:

$$H(c) = \frac{\mathcal{N}_2^*}{q} \frac{\sum\limits_{i=0}^{q} i\left(n_i e^{-\beta E_i}(Kc)^i\right)}{\sum\limits_{i=0}^{q} n_i e^{-\beta E_i}(Kc)^i}, \tag{5}$$

where q represent the degree of interaction between the molecules, $n_i = \dfrac{q!}{i!(q-i)!}$ is the degenerecence number, and $\beta = \frac{1}{RT}$. The finite number of parameters for the identification are the interaction energies E_i, $i = 1, \ldots, q$, and the Langnuir coefficient K. Now, if we consider a gas-solid isotherm model for two components mixture (see [5], and [6]), we have the following function:

$$H(c_1, c_2) = \frac{\mathcal{N}_2^*}{q} \left(\begin{array}{c} \dfrac{\sum\limits_{i+j\leq q} i\left(n_{ij} e^{-\beta E_{ij}}(K_1 c_1)^i (K_2 c_2)^j\right)}{\sum\limits_{i+j\leq q} n_{ij} e^{-\beta E_{ij}}(K_1 c_1)^i (K_2 c_2)^j} \\ \\ \dfrac{\sum\limits_{i+j\leq q} j\left(n_{ij} e^{-\beta E_{ij}}(K_1 c_1)^i (K_2 c_2)^j\right)}{\sum\limits_{i+j\leq q} n_{ij} e^{-\beta E_{ij}}(K_1 c_1)^i (K_2 c_2)^j} \end{array} \right), \tag{6}$$

where $n_{ij} = \dfrac{q!}{i!j!(q-i-j)!}$ is the degenerecence number. The finite number of parameters for the identification are the interaction energies E_{ij}, $0 \leq i+j \leq q$, and the Langnuir coefficient K_1, K_2.

We shall denote by $\mathbf{e}$ the vector with components $e_1, \ldots, e_N$, and emphasize if necessary the dependence of F with respect to $\mathbf{e}$ by $F(\mathbf{e})$. The parameters $e_1, \ldots, e_N$ have a physical meaning: they quantify the possible interactions between components in the stationary phase. Their identification is therefore an interesting problem from the point of view of the applications. Thus we are led to the problem of identifying a vector $\mathbf{e} \in \mathbb{R}^N$, in a way that we are going to precise now.

2.2 Identification

The principle of identification is generally to minimize a given "cost function" $\widetilde{J}(\mathbf{e})$, up to constraints on $\mathbf{e}$. The choice of $\widetilde{J}$ is of course fundamental in order to obtain a (possibly) well-posed problem. Even if the problem is well-posed, one can have trouble when approximating the solution: the convergence can be too slow to be acceptable. We do not intend here to give the best choice for $\widetilde{J}$. For our problem we shall limit ourselves to define $\widetilde{J}$ in a "natural" way.

When we have experimental data to compare with numerical simulation, we intend to consider an observation at the output $z = L$ of the column, that is a vector-valued function of the time $\mathbf{w}_{\text{obs}}(t) = (w_1^{\text{obs}}(t), \ldots, w_M^{\text{obs}}(t))$. A natural idea is now to find the vector of parameters $\mathbf{e} \in \mathbb{R}^N$ such that the solution of (3)-(4) in $z = L$, with the function $H(\mathbf{e})$, is the best approximation of the function $\mathbf{w}_{\text{obs}}$, in a sense to be specified. We are going to be led by physical arguments in our choice.

Notice first that the real observation would *not* be a vector-valued function, but a real-valued function, since the detectors used in chromatography give only the sum of the components of the solution. A realistic criterion would therefore act on $\sum_{i=1}^{M} w_i(L,t)$. For technical reasons that will appear later, we shall consider a simplified problem, assuming that we can observe every component of the solution.

To compare $\mathbf{w}(L,t)$ with $\mathbf{w}_{\text{obs}}(t)$, we consider a cost function J which measures, in a general way, a distance between both functions, and we define $\widetilde{J}$ by $\widetilde{J}(\mathbf{e}) \stackrel{\text{def}}{=} J(\mathbf{w}(\mathbf{e}))$, where $\mathbf{w}(\mathbf{e})$ denotes the solution of (3)-(4) with the function $H(\mathbf{e})$.

First we consider the following L^2-criterion $J = J_0$, which has no precise physical meaning, but allows us to perform some formal calculations on the function $\widetilde{J}$. So we set

$$J_0(\mathbf{w}) \stackrel{\text{def}}{=} \frac{1}{2}\|\mathbf{w}(L,\cdot) - \mathbf{w}_{\text{obs}}\|^2_{L^2(0,T)}, \quad \text{for all } \mathbf{w}, \tag{7}$$

On the other hand, for the chromatographic colum ($v = 0$) we have experimental data given by the flux $F(\mathbf{w}_{\text{obs}}) = u\mathbf{c}^{\text{obs}}$. In this case we are interested in a flux-L^2-criterion $J = \mathcal{J}_0$ which we define by the following form:

$$\mathcal{J}_0(\mathbf{w}) \stackrel{\text{def}}{=} \frac{1}{2}\|F(\mathbf{w}(L,\cdot)) - F(\mathbf{w}_{\text{obs}})\|^2_{L^2(0,T)}, \quad \text{for all } \mathbf{w}. \tag{8}$$

The identification problem becomes now

$$\min_{\mathbf{e}} J\left(\mathbf{w}(\mathbf{e})\right), \tag{9}$$

and we rewrite it as a constrained minimization problem on $\mathbf{w}$:

$$\min_{\mathbf{w}} J(\mathbf{w}), \qquad \mathbf{w} = \mathbf{w}(\mathbf{e}). \tag{10}$$

Since we are interested in the shape of the solutions, we ought to find more suitable criteria. Because we have a finite propagation speed, and the injection has a compact support, we have that $\mathbf{w}(L,\cdot)$ and $\mathbf{w}_{\text{obs}}$ have a compact support. Clearly, for a good identification we need that $\text{supp}(\mathbf{w}(L,\cdot)) \simeq \text{supp}(\mathbf{w}_{\text{obs}})$, but the criteria (7) and (8) do not ensure this support similarity. Furthermore, if we consider an initial guess such that the set $\text{supp}(\mathbf{w}(L,\cdot)) \cap \text{supp}(\mathbf{w}_{\text{obs}})$ is empty, then the constrained minimization problem (9), with the criterion (7) or (8) is numerically ill-posed (see [13]). If we have small measures $|\text{supp}(\mathbf{w}(L,\cdot))|$ and $|\text{supp}(\mathbf{w}_{\text{obs}})|$, then the support of each component $m = 1, \ldots, M$ is characterized by the first momentum for all component m. Consider the first momentum of a given function $\mathbf{X}(t) \in \mathbb{R}^M$:

$$\mu_1(\mathbf{X}) = \int_{-\infty}^{+\infty} t\mathbf{X}(t)dt \in \mathbb{R}^M. \tag{11}$$

If $\mathbf{x}$ is the concentration profile vector at the end of the column, this quantity is physically significant : $\mu_1(\mathbf{x})_m$ it is called the *retention time* of the component m, and can be measured. We introduce

$$J_1(\mathbf{w}) = \frac{1}{2}\|\mu_1(\mathbf{w}(L,.)) - \mu_1(\mathbf{w}_{\text{obs}})\|^2 \tag{12}$$

and the new cost function

$$J_\rho(\mathbf{w}) = J_0(\mathbf{w}) + \rho J_1(\mathbf{w}). \tag{13}$$

This implies of course a knowledge of each component of the mixture –which is not realistic, as we mentioned before.

One can imagine using the momentum of order $p \geq 1$

$$\mu_p(\mathbf{X}) = \int_{-\infty}^{+\infty} t^p\mathbf{X}(t)dt, \tag{14}$$

which gives rise to the cost function J by a linear combination between J_0, J_1, J_2, where

$$J_p(\mathbf{w}) = \frac{1}{2}\|\mu_p(\mathbf{w}(L,.)) - \mu_p(\mathbf{w}_{\text{obs}})\|^2, \quad p \geq 1. \tag{15}$$

However, only the momentum of order 2 seems to be reachable by physical measures in a few special cases. The criteria of high orders might be therefore theoretically interesting (if we know all the moments of a signal, we know the signal), but these orders are physically out of reach.

Similarly to the definition of the L^2-flux-criterion (8), we can introduce

$$\mathcal{J}_1(\mathbf{w}) = \frac{1}{2}\|\mu_1(F(\mathbf{w}(L,.))) - \mu_1(F(\mathbf{w}_{\text{obs}}))\|^2, \tag{16}$$

and the momentum-flux-criterion:

$$\mathcal{J}_\rho(\mathbf{w}) = \mathcal{J}_0(\mathbf{w}) + \rho\mathcal{J}_1(\mathbf{w}). \tag{17}$$

We are going now to compute, at least formally, the gradient of the function $\widetilde{J}$. A classical way to do this is to give a Lagrangian formulation of this problem (see for instance [1]). The following section is devoted to these calculations, and we shall see that serious problems of ill-posedness will occur.

3 Formal calculus of the gradient of the cost function

3.1 The Lagrangian formulation

Throughout this section, we assume $\mathbf{w}$ to be a smooth function. First, we rewrite our constraint on $\mathbf{w}$, that is "$\mathbf{w}$ is the solution of (3)-(4)", in a weak form. So, fix $T > 0$, such that $\mathbf{w}(z,T) = 0$ –this can be done since the injection is compactly supported–, let $\mathbf{p}$ be a smooth test function, and define $E(\mathbf{w}, \mathbf{p}; \mathbf{e})$ by

$$\begin{aligned} E(\mathbf{w}, \mathbf{p}; \mathbf{e}) \stackrel{\text{def}}{=} & -\int_0^T \int_0^L (\mathbf{w} \cdot \mathbf{p}_t + F(\mathbf{w}, \mathbf{e}) \cdot \mathbf{p}_z) \\ & + \int_{t=T} \mathbf{w} \cdot \mathbf{p} + \int_{z=L} F(\mathbf{w}) \cdot \mathbf{p} \\ & - \int_{t=0} \mathbf{w}_{\text{inj}}(t) \cdot \mathbf{p} - \int_{z=0} F(\mathbf{w}^{\text{init}}) \cdot \mathbf{p}, \end{aligned} \tag{18}$$

where the dot $\cdot$ denotes the scalar product in $\mathbb{R}^M$. The constraints are satisfied, that is $\mathbf{w} = \mathbf{w}(\mathbf{e})$, if and only if $E(\mathbf{w}, \mathbf{p}; \mathbf{e}) = 0$. We define now a Lagrangian for the problem (9) by

$$\mathcal{L}(\mathbf{w}, \mathbf{p}; \mathbf{e}) \stackrel{\text{def}}{=} J(\mathbf{w}) - E(\mathbf{w}, \mathbf{p}; \mathbf{e}). \tag{19}$$

The test function $\mathbf{p}$ will appear as a "generalized Lagrange multiplier" related to the constraint $\mathbf{w} = \mathbf{w}(\mathbf{e})$. Now, we notice that, since $E(\mathbf{w}(\mathbf{e}), \mathbf{p}; \mathbf{e}) = 0$,

$$\mathcal{L}(\mathbf{w}(\mathbf{e}), \mathbf{p}; \mathbf{e}) = J(\mathbf{w}(\mathbf{e})) = \widetilde{J}(\mathbf{e}), \tag{20}$$

so that, when derivating in the $\delta\mathbf{e}$ direction, we obtain

$$\widetilde{J}'(\mathbf{e}) = \left\langle \frac{\partial \mathcal{L}}{\partial \mathbf{w}}(\mathbf{w}(\mathbf{e}), \mathbf{p}; \mathbf{e}), \mathbf{w}'(\mathbf{e})\delta\mathbf{e} \right\rangle + \frac{\partial \mathcal{L}}{\partial \mathbf{e}}(\mathbf{w}(\mathbf{e}), \mathbf{p}; \mathbf{e})\delta\mathbf{e} \tag{21}$$

for every test function $\mathbf{p}$. Now the idea is that we cannot compute $\mathbf{w}'(\mathbf{e})\delta\mathbf{e}$, so we are going to choose a particular $\mathbf{p}$ such that

$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}}(\mathbf{w}(\mathbf{e}), \mathbf{p}; \mathbf{e}) \equiv 0. \tag{22}$$

This function $\mathbf{p}$ will be called the *adjoint state* associated to the Lagrangian (19).

This can also be viewed as a consequence of the inf-sup formulation for (9), which leads, at least formally, to cancel the gradient in $\mathbf{w}$ of $\mathcal{L}$. We proceed now to compute the adjoint state and the gradient of $\widetilde{J}$.

3.2 The adjoint state and the gradient of $\widetilde{J}$

Let $\delta\mathbf{w}$ be a smooth function, and perform a formal differentiation of E in the $\delta\mathbf{w}$ direction. For the time being, we give up the notation $F(\mathbf{w},\mathbf{e})$, and denote by $H'(\mathbf{w})$ the Jacobian matrix of $F(\mathbf{w})$. We obtain, since $\mathbf{w}_{\text{inj}}$ and $\mathbf{w}^{\text{init}}$ are given fixed functions,

$$\begin{aligned}\left\langle \frac{\partial E}{\partial \mathbf{w}}, \delta\mathbf{w}\right\rangle &= -\int_0^T\int_0^L \delta\mathbf{w}\cdot\left(\mathbf{p}_t + \partial_{\mathbf{w}}F(\mathbf{w},\mathbf{e})^\top\mathbf{p}_z\right)\\ &\quad+\int\limits_{t=T}\delta\mathbf{w}\cdot\mathbf{p} + \int\limits_{z=L}\delta\mathbf{w}\cdot\partial_{\mathbf{w}}F(\mathbf{w},\mathbf{e})^\top\mathbf{p}.\end{aligned} \tag{23}$$

For all smooth $\delta\mathbf{w}$ we have

$$\begin{aligned}\left\langle \frac{\partial \mathcal{L}}{\partial \mathbf{w}}, \delta\mathbf{w}\right\rangle &= \int_0^T\int_0^L \delta\mathbf{w}\cdot\left(\mathbf{p}_z + \partial_{\mathbf{w}}F(\mathbf{w},\mathbf{e})^\top\mathbf{p}_t\right)\\ &\quad+\int\limits_{z=L}\delta\mathbf{w}\cdot\left(\frac{\partial J}{\partial \mathbf{w}} - \partial_{\mathbf{w}}F(\mathbf{w},\mathbf{e})^\top\mathbf{p}\right)\\ &\quad-\int\limits_{t=T}\delta\mathbf{w}\cdot\mathbf{p}.\end{aligned} \tag{24}$$

The quantity $\frac{\partial\mathcal{L}}{\partial\mathbf{w}}$ is equal to zero when $\mathbf{p}$ is solution to the following backward adjoint problem

$$\mathbf{p}_t + \partial_{\mathbf{w}}F(\mathbf{w},\mathbf{e})^\top\mathbf{p}_z = 0, \tag{25}$$

with these final conditions of position and time,

$$\begin{cases}\partial_{\mathbf{w}}F(\mathbf{w}(L,t),\mathbf{e})^\top\mathbf{p}(L,t) = \left\langle \dfrac{\partial J}{\partial\mathbf{w}},\delta_t\right\rangle, & t\ge 0,\\ \mathbf{p}(z,T)=0, & z\in[0,L],\end{cases} \tag{26}$$

where δ_t is the Dirac measure in t. The relation $(26)_1$ is well defined since $\partial_{\mathbf{w}}F(\mathbf{w},\mathbf{e})^\top$ is invertible. For instance, if we consider the cost function J_ρ (13), then the final condition $(26)_1$ can be written

$$\begin{aligned}\partial_{\mathbf{w}}F(\mathbf{w}(L,t),\mathbf{e})^\top\mathbf{p}(L,t) &= \mathbf{w}(L,t)-\mathbf{w}_{\text{obs}}(t)\\ &\quad+\rho\left[\mu_1(\mathbf{w}(L,t))-\mu_1(\mathbf{w}_{\text{obs}}(t))\right]t, \quad t\ge 0.\end{aligned}$$

On the other hand, if we consider the flux-criterion $\mathcal{J}_\rho$ defined in (17), then we obtain

$$\begin{aligned}\mathbf{p}(L,t) &= F(\mathbf{w}(L,t))-F(\mathbf{w}_{\text{obs}}(t))\\ &\quad+\rho\left[\mu_1(F(\mathbf{w}(L,t)))-\mu_1(F(\mathbf{w}_{\text{obs}}(t)))\right]t, \quad t\ge 0.\end{aligned}$$

Always acting formally, we choose a solution $\mathbf{p}$ of (25)-(26), so that $\frac{\partial\mathcal{L}}{\partial\mathbf{w}}=0$, and we obtain the following expression for the gradient of $\widetilde{J}$:

$$\begin{aligned}\widetilde{J}'(\mathbf{e}) &= \frac{\partial\mathcal{L}}{\partial\mathbf{e}} = -\frac{\partial E}{\partial\mathbf{e}}\\ &\quad\int_0^T\int_0^L \partial_{\mathbf{e}}F\cdot\mathbf{p}_t + \int_0^L\partial_{\mathbf{e}}F(\mathbf{w}^{\text{init}}(z),\mathbf{e})\cdot\mathbf{p}(0,z).\end{aligned} \tag{27}$$

Formula (27) holds for any solution of (25)-(26), so the gradient is well-defined only if we have existence and uniqueness for the adjoint problem. This is clearly false if $\mathbf{w}$ has shocks as soon as the functions in (26) are not smooth, since we are led to a linear system with discontinuous coefficients. Then, the theory of characteristics clearly shows that $\mathbf{p}$ is not uniquely determined.

The computation of $\widetilde{J}'(\mathbf{e})$ needs problems (3)-(4) and (25)-(26) to be solved. Thus whatever conjugate gradient method we may try and apply in order to find the solution of the minimization problem, we shall face rather heavy computations.

4 Discretization of the identification problem

In this section we intend to give a discretization of the minimization problem. Rather than discretizing the "direct system" (3), the "adjoint system" (25) and the cost function, we prefer to discretize only the cost function and the direct system (3) (see [1]). Let us denote by J_Δ the discretized cost function. Then we follow exactly the same method as in the continuous case, that is that we compute a discrete Lagrangian, which will lead to a discrete adjoint state, and finally to a discrete gradient of J_Δ. This technique seems to be preferable for the stability of the numerical results.

Let Δz (resp. Δt) be a positive space (resp. time) step. These parameters will tend to 0, the ratio $\lambda = \Delta t/\Delta z$ keeping constant. For $n = 0, \ldots, N$, $j = 0, \ldots, J$, the sequence w_j^n is an approximation of the solution $\mathbf{w}$ at the point $(z_j = j\Delta z, t_n = n\Delta t)$. In the same way, we discretize $\mathbf{w}_{\text{inj}}(t)$, $\mathbf{w}^{\text{init}}(z)$, $\mathbf{w}_{\text{obs}}(t)$, by $\mathbf{w}_{\text{inj}}^n$, $\mathbf{w}_j^{\text{init}}$, $\mathbf{w}_{\text{obs}}^n$ respectively.

In the following, an index Δ will denote any discrete function depending on the discretization step. Then the minimization problem becomes

$$\min_{\mathbf{e}} J_\Delta\left(\mathbf{w}_j^n(\mathbf{e})\right), \tag{28}$$

where J_Δ is a discretization of J. For instance, we have the following discretization for the L^2-criterion

$$J_{0,\Delta}\left(\mathbf{w}_j^n\right) = \frac{1}{2}\sum_{n=0}^{N-1} \Delta t|\mathbf{w}_J^n - \mathbf{w}_{\text{obs}}^n|^2, \tag{29}$$

and the following discretization for the momentum of order p criterion

$$J_{p,\Delta}\left(\mathbf{w}_j^n\right) = \frac{1}{2}\|\mu_1(\mathbf{w}_J^n) - \mu_1(\mathbf{w}_{\text{obs}}^n)\|^2 \tag{30}$$

with $\left(w_j^n(\mathbf{e});\ j = 0, \ldots, J-1,\ n = 0, \ldots, N-1\right) \in \mathbb{R}^M$.

4.1 The identification problem for conservative schemes

We consider a conservative K-points scheme for the system (3)-(4):

$$\begin{cases} \mathbf{w}_j^{n+1} = \mathbf{w}_j^n - \lambda\left(g_{j+\frac{1}{2}}^n(\mathbf{e}) - g_{j-\frac{1}{2}}^n(\mathbf{e})\right), \\ \\ \mathbf{w}_0^n = \mathbf{w}_{\text{inj}}^n, \qquad \mathbf{w}_j^0 = \mathbf{w}_j^{\text{init}}, \end{cases} \tag{31}$$

where $g_{j+\frac{1}{2}}^n(\mathbf{e}) = g(\mathbf{w}_{j-K+1}^n, \ldots, \mathbf{w}_{j+K}^n, \mathbf{e})$, g being the numerical flux of the scheme.

The analogue of the adjoint state in the continuous case is now a discrete function $\mathbf{p}_\Delta = (\mathbf{p}_j^n;\ j = 0, \ldots, J-1,\ n = 0, \ldots, N-1)$, and appears also as "generalized Lagrange multiplier". Let us define a weak formulation for the scheme (31) by

$$
\begin{aligned}
E_\Delta\left(\mathbf{w}_j^n, \mathbf{p}_j^n, \mathbf{e}\right) &\stackrel{\text{def}}{=} \sum_{n,j} \left(\mathbf{w}_j^{n+1} - \mathbf{w}_j^n - \lambda(g_{j+\frac{1}{2}}^n(\mathbf{e}) - g_{j-\frac{1}{2}}^n(\mathbf{e}))\right) \cdot \mathbf{p}_j^{n+1} \\
&= -\sum_{n,j} \left(\mathbf{w}_j^n \cdot (\mathbf{p}_j^n - \mathbf{p}_j^{n+1}) + \lambda g_{j+\frac{1}{2}}^n(\mathbf{e}) \cdot (\mathbf{p}_j^{n+1} - \mathbf{p}_{j+1}^{n+1})\right) \\
&\quad + \sum_{j=0}^{J-1} \left(\mathbf{w}_j^N \cdot \mathbf{p}_j^N - \mathbf{w}_j^0 \cdot \mathbf{p}_j^0\right) \\
&\quad + \lambda \sum_{n=0}^{N-1} \left(g_{J-\frac{1}{2}}^n(\mathbf{e}) \cdot \mathbf{p}_J^n - g_{-\frac{1}{2}}^n(\mathbf{e}) \mathbf{p}_0^n\right),
\end{aligned} \tag{32}
$$

and we introduce the following discrete Lagrangian of the problem (28):

$$
L_\Delta\left(\mathbf{w}_j^n, \mathbf{p}_j^n, \mathbf{e}\right) \stackrel{\text{def}}{=} \frac{1}{\Delta z} J_\Delta\left(\mathbf{w}_j^n\right) - E_\Delta\left(\mathbf{w}_j^n, \mathbf{p}_j^n, \mathbf{e}\right). \tag{33}
$$

We assume that g is of class C^1, and differentiate E_Δ with respect to $\mathbf{w}_j^n$. We obtain

$$
\begin{aligned}
\frac{\partial E_\Delta}{\partial \mathbf{w}_j^n} &= \mathbf{p}_j^n - \mathbf{p}_j^{n+1} + \lambda \sum_{k=-K}^{K-1} \left(\frac{\partial}{\partial \mathbf{w}_j^n} g_{j+k+\frac{1}{2}}^n\right)^\top (\mathbf{p}_{j+k}^{n+1} - \mathbf{p}_{j+k+1}^{n+1}) \\
&\quad + \mathbf{p}_j^N \delta_{nN} + \lambda \left(\frac{\partial}{\partial \mathbf{w}_J^n} g_{J-\frac{1}{2}}^n\right)^\top \mathbf{p}_J^n \delta_{jJ}
\end{aligned} \tag{34}
$$

As previously, the $\mathbf{p}_j^n$ are chosen such that $\partial L_\Delta / \partial \mathbf{w}_j^n = 0$. This can be viewed as the inf-sup formulation of the problem (28). We have

$$
\begin{aligned}
-\frac{\partial L_\Delta}{\partial \mathbf{w}_j^n} &= \mathbf{p}_j^n - \mathbf{p}_j^{n+1} + \lambda \sum_{k=-K}^{K-1} \left(\frac{\partial}{\partial \mathbf{w}_j^n} g_{j+k+\frac{1}{2}}^n\right)^\top (\mathbf{p}_{j+k}^{n+1} - \mathbf{p}_{j+k+1}^{n+1}) \\
&\quad + \mathbf{p}_j^N \delta_{nN} + \lambda \left(\left(\frac{\partial}{\partial \mathbf{w}_J^n} g_{J-\frac{1}{2}}^n\right)^\top \mathbf{p}_J^n - \frac{\partial J_\Delta}{\partial \mathbf{w}_J^n}\right) \delta_{jJ}
\end{aligned} \tag{35}
$$

In order that $\frac{\partial}{\partial \mathbf{w}_j^n} L_\Delta = 0$, we choose $\mathbf{p}_j^n$ to be solution of the following 'adjoint scheme'

$$
\left\{
\begin{aligned}
&\mathbf{p}_j^n = \mathbf{p}_j^{n+1} - \lambda \sum_{k=-K}^{K-1} \frac{\partial}{\partial \mathbf{w}_j^n} g_{j+k+\frac{1}{2}}^n \cdot (\mathbf{p}_{j+k}^{n+1} - \mathbf{p}_{j+k+1}^{n+1}), \\
&\left(\frac{\partial}{\partial \mathbf{w}_J^n} g_{J-\frac{1}{2}}^n\right)^\top \mathbf{p}_J^n = \frac{\partial J_\Delta}{\partial \mathbf{w}_J^n}, \\
&\mathbf{p}_j^N = 0.
\end{aligned}
\right. \tag{36}
$$

With this choice for $\mathbf{p}_\Delta$, the discrete gradient of J_Δ is given by

$$
\begin{aligned}
\bar{J}'_\Delta(\mathbf{e}) &= -\frac{\partial E_\Delta}{\partial \mathbf{e}} \Delta z \\
&= -\Delta t \sum_{n,j} \frac{\partial}{\partial \mathbf{e}} g^n_{j+\frac{1}{2}} \cdot (\mathbf{p}^{n+1}_j - \mathbf{p}^{n+1}_{j+1}) + \Delta t \sum_{n=0}^{N-1} \frac{\partial}{\partial \mathbf{e}} g^n_{-\frac{1}{2}} \mathbf{p}^{n+1}_0 .
\end{aligned} \tag{37}
$$

As well as the discrete adjoint scheme (36), the discrete gradient is well defined, in contrast to the preceding formal calculus. We can therefore try a numerical resolution of our problem, by using for instance a conjugate gradient method. Now the question is: what happens when Δt and Δz go to 0? More precisely, if we have a convergence property for the adjoint scheme, what is the meaning of the limit? An attempt to answer these questions can be found in [13].

We give now two examples of computation of the discrete gradient: when a conservative scheme is given, it is sufficient to compute the derivative of the flux with respect to the coefficients $\mathbf{e}$ to obtain $\bar{J}'_\Delta$.

4.2 Examples

In the case of chromatography, keep in mind that the eigenvalues of the Jacobian of (3) are positive. Therefore, we are going to use upwind schemes. This simpler form allows explicit computations but, as we shall see, even in this case the adjoint scheme can look quite awful (for the Van Leer scheme).

4.2.1 Godunov scheme

The Godunov scheme for the system of chromatography is the simplest one can imagine (see [4]), since it can be boiled down to the standard upwind scheme :

$$
\begin{cases}
\mathbf{w}^{n+1}_j = \mathbf{w}^n_j - \lambda \left(F(\mathbf{w}^n_j, \mathbf{e}) - F(\mathbf{w}^n_{j-1}, \mathbf{e}) \right), \\[2ex]
\mathbf{w}^n_0 = \mathbf{w}^n_{\text{inj}}, \qquad \mathbf{w}^0_j = \mathbf{w}^{\text{init}}_j .
\end{cases} \tag{38}
$$

In this case the adjoint scheme is given by

$$
\begin{cases}
\mathbf{p}^n_j = \mathbf{p}^{n+1}_j - \lambda \left(\partial_{\mathbf{w}} F(\mathbf{w}^n_j, \mathbf{e}) \right)^\top \left(\mathbf{p}^{n+1}_j - \mathbf{p}^{n+1}_{j+1} \right), \\[2ex]
\mathbf{p}^n_J = \dfrac{\partial J_\Delta}{\partial \mathbf{w}^n_J}, \qquad \mathbf{p}^N_j = 0,
\end{cases} \tag{39}
$$

and the gradient is :

$$
\begin{aligned}
\bar{J}'_\Delta(\mathbf{e}) &= \Delta t \sum_{j=0}^{J-1} \sum_{n=0}^{N-1} \partial_{\mathbf{e}} F(\mathbf{w}^n_j, \mathbf{e}) \cdot (\mathbf{p}^{n+1}_{j+1} - \mathbf{p}^{n+1}_j) \\
&\quad + \Delta t \sum_{n=0}^{N-1} \partial_{\mathbf{e}} F(\mathbf{w}^{\text{init}}_j, \mathbf{e}) \cdot \mathbf{p}^{n+1}_0
\end{aligned} \tag{40}
$$

4.2.2 Van-Leer scheme

A possible formulation for a Van-Leer like scheme for the system of chromatography (see [4]) is the following:

$$\mathbf{w}_j^{n+1} = \mathbf{w}_j^n - \lambda \left(F(\mathbf{w}_j^{n+\frac{1}{2}}, \mathbf{e}) - F(\mathbf{w}_{j-1}^{n+\frac{1}{2}}, \mathbf{e}) \right), \tag{41}$$

$$\begin{cases} \mathbf{w}_j^{n+\frac{1}{2}} = \mathbf{w}_j^n - \frac{\lambda}{2} \left(F(\mathbf{w}_j^n, \mathbf{e}) - F(\mathbf{w}_{j-\frac{1}{2}}^n, \mathbf{e}) \right), \\ \mathbf{w}_{j-\frac{1}{2}}^n = \mathbf{w}_j^n - s_j^n \Delta z \end{cases} \tag{42}$$

where s_j^n is a slope corrector of min-mod type:

$$s_j^n = \begin{cases} \sigma \min \left(\dfrac{|\mathbf{w}_{j+1}^n - \mathbf{w}_j^n|}{2\Delta z}, \dfrac{|\mathbf{w}_j^n - \mathbf{w}_{j-1}^n|}{2\Delta z} \right) \\ \qquad \text{if } \operatorname{sgn}\left(\mathbf{w}_{j+1}^n - \mathbf{w}_j^n\right) = \operatorname{sgn}\left(\mathbf{w}_j^n - \mathbf{w}_{j-1}^n\right) = \sigma \\ 0 \qquad \text{otherwise} \end{cases} \tag{43}$$

Now, a tedious computation gives us the adjoint scheme:

$$\begin{aligned} \mathbf{p}_j^n \;=\; & \mathbf{p}_j^{n+1} - \lambda \left[1 - \frac{\lambda}{2} \left(\partial_{\mathbf{w}} F(\mathbf{w}_j^n) - (1 - \frac{\sigma_j^0}{2}) \partial_{\mathbf{w}} F(\mathbf{w}_{j-\frac{1}{2}}^n) \right) \right]^\top \\ & \partial_{\mathbf{w}} F(\mathbf{w}_j^{n+\frac{1}{2}})^\top \left(\mathbf{p}_j^{n+1} - \mathbf{p}_{j+1}^{n+1} \right) \\ & + \frac{\lambda^2}{4} \sigma_{j+1}^- \partial_{\mathbf{w}} F(\mathbf{w}_{j+\frac{1}{2}}^n)^\top \partial_{\mathbf{w}} F(\mathbf{w}_{j+1}^{n+\frac{1}{2}})^\top \left(\mathbf{p}_{j+1}^{n+1} - \mathbf{p}_{j+2}^{n+1} \right) \\ & + \frac{\lambda^2}{4} \sigma_{j-1}^+ \partial_{\mathbf{w}} F(\mathbf{w}_{j-\frac{3}{2}}^n)^\top \partial_{\mathbf{w}} F(\mathbf{w}_{j-1}^{n+\frac{1}{2}})^\top \left(\mathbf{p}_{j-1}^{n+1} - \mathbf{p}_j^{n+1} \right), \end{aligned} \tag{44}$$

with the following slope correctors

$$\sigma_j^0 = 2\Delta z \frac{\partial s_j^n}{\partial \mathbf{w}_0} = \begin{cases} -1 \quad \text{if} & \operatorname{sgn}\left(\mathbf{w}_{j+1}^n - \mathbf{w}_j^n\right) = \operatorname{sgn}\left(\mathbf{w}_j^n - \mathbf{w}_{j-1}^n\right), \\ & \text{and } |\mathbf{w}_{j+1}^n - \mathbf{w}_j^n| > |\mathbf{w}_j^n - \mathbf{w}_{j-1}^n| \\ 1 \quad \text{if} & \operatorname{sgn}\left(\mathbf{w}_{j+1}^n - \mathbf{w}_j^n\right) = \operatorname{sgn}\left(\mathbf{w}_j^n - \mathbf{w}_{j-1}^n\right), \\ & \text{and } |\mathbf{w}_{j+1}^n - \mathbf{w}_j^n| \le |\mathbf{w}_j^n - \mathbf{w}_{j-1}^n| \\ 0 \quad \text{otherwise} \end{cases} \tag{45}$$

$$\sigma_j^+ = 2\Delta z \frac{\partial s_j^n}{\partial \mathbf{w}_1} = \begin{cases} 1 \quad \text{if} & \operatorname{sgn}\left(\mathbf{w}_{j+1}^n - \mathbf{w}_j^n\right) = \operatorname{sgn}\left(\mathbf{w}_j^n - \mathbf{w}_{j-1}^n\right), \\ & \text{and } |\mathbf{w}_{j+1}^n - \mathbf{w}_j^n| \le |\mathbf{w}_j^n - \mathbf{w}_{j-1}^n| \\ 0 \quad \text{otherwise} \end{cases} \tag{46}$$

$$\sigma_j^- = 2\Delta z \frac{\partial s_j^n}{\partial \mathbf{w}_{-1}} = \begin{cases} -1 \quad \text{if} & \operatorname{sgn}\left(\mathbf{w}_{j+1}^n - \mathbf{w}_j^n\right) = \operatorname{sgn}\left(\mathbf{w}_j^n - \mathbf{w}_{j-1}^n\right), \\ & \text{and } |\mathbf{w}_{j+1}^n - \mathbf{w}_j^n| > |\mathbf{w}_j^n - \mathbf{w}_{j-1}^n| \\ 0 \quad \text{otherwise} \end{cases} \tag{47}$$

Therefore the gradient of $\bar{J}_\Delta$ is given by

$$\begin{cases} \bar{J}'_\Delta(\mathbf{e}) = \Delta t \sum_{n,j} \partial_{\mathbf{e}} F_j^n \cdot \left(\mathbf{p}_{j+1}^{n+1} - \mathbf{p}_j^{n+1}\right) + \Delta t \sum_{n=0}^{N-1} \partial_{\mathbf{e}} F_0^n \cdot \mathbf{p}_0^{n+1} \\ \partial_{\mathbf{e}} F_j^n = \frac{\partial F}{\partial \mathbf{e}}(\mathbf{w}_j^{n+\frac{1}{2}}) - \frac{\lambda}{2} \frac{\partial F}{\partial \mathbf{w}}(\mathbf{w}_j^{n+\frac{1}{2}}) \left(\frac{\partial F}{\partial \mathbf{e}}(\mathbf{w}_j^n) - \frac{\partial F}{\partial \mathbf{e}}(\mathbf{w}_{j-\frac{1}{2}}^n)\right) \end{cases} \tag{48}$$

Here $\mathbf{w}_{-1/2}^n$ is an extrapolation of the discrete solution at time $-1/2\Delta t$. Since we assume that all the data are with compact support, this quantity may be taken to be equal to 0.

One can see that, even though the direct scheme is quite simple, the adjoint scheme is definitely not a standard linear scheme.

4.2.3 Convergence results

In this section we give a result of convergence for numerical schemes, which is essentially formal, since we have no general results about the convergence for approximate solutions to systems of conservation laws. Let E be a compact set of $\mathbb{R}^N$. There exists at least one solution to that discrete identification problem

$$\min_{\mathbf{e} \in E} J_\Delta\left(\mathbf{w}_j^n(\mathbf{e})\right). \tag{49}$$

Now Let $\mathbf{e}_\Delta$ be a solution to (49) for each discretisation $\Delta z, \Delta t$. Then we obtain the following result :

Proposition 4.1 *Consider an entropic scheme in conservative form, consistent with* (3). *Assume that, for every* $\mathbf{e} \in E$, E *compact set of* $\mathbb{R}^N$, (3) *has an entropic solution. Assume that there exists a sequence* $\Delta_k t, \Delta_k z \to 0$ *such as*

$$\|\mathbf{w}_{\Delta_k}(\mathbf{e}_{\Delta_k})\|_{L^\infty(\Omega)} \leq C \tag{50}$$

$$\mathbf{w}_{\Delta_k}(\mathbf{e}_{\Delta_k}) \to \mathbf{w} \quad \text{in } L^1_{loc}(\Omega) \text{ and a.e.} \tag{51}$$

$$\mathbf{e}_{\Delta_k} \to \bar{\mathbf{e}} \quad \text{in } E. \tag{52}$$

Then $\mathbf{w}$ *is solution to* (3)-(4) *with* $\mathbf{e} = \bar{\mathbf{e}}$, *where* $\bar{e}$ *is a solution to the continuous identification problem*

$$\min_{\mathbf{e} \in E} J(\mathbf{w}(\mathbf{e})), \tag{53}$$

where J *is the cost function* J_ρ *or* $\mathcal{J}_\rho$, *with* $\rho \geq 0$. *Futhermore,*

$$J(\mathbf{w}_{\Delta_k}(\mathbf{e}_{\Delta_k})) \to J(\mathbf{w}(\bar{\mathbf{e}})), \quad \text{in } \mathbb{R}. \tag{54}$$

Remark : Hypotheses (50) and (51) are verified for the scalar case, if we suppose a numerical scheme with a Lipschitz continuous numerical flux $g^n_{j+\frac{1}{2}}$, with Lipschitz constant C_{Lip} independent from $\mathbf{e}$. In this case we have L^∞ and BV estimations for $\mathbf{w}_{\Delta_k}(\mathbf{e}_{\Delta_k})$ independent from $\mathbf{e}_\Delta$ (see [13]).

Proof Copy the proof of Lax-Wendroff's Proposition in [3] in order to prove that $\mathbf{w}$ is solution to (3)-(4).

We deduce the convergence of the p-momentum-criterion by the $L^1_{loc}(\Omega)$ and a.e. convergences (51), and the L^∞-stability (50). We have for $p=1$

$$\int_0^T t\mathbf{w}_{\Delta_k}(\mathbf{e}_{\Delta_k})(L,\cdot)dt \to \int_0^T t\mathbf{w}(\bar{\mathbf{e}})(L,\cdot)dt, \quad \text{in } \mathbb{R},$$

Thus, using the definition of the momentum of order 1, we deduce

$$\mu_1\left(\mathbf{w}_{\Delta_k}(L,\cdot)(\mathbf{e}_{\Delta_k})\right) \;\to\; \mu_1\left(\mathbf{w}(\bar{\mathbf{e}})(L,\cdot)\right), \quad \text{in } \mathbb{R},$$

and using the definition of J_1, we obtain

$$J_1\left(\mathbf{w}_{\Delta_k}(\mathbf{e}_{\Delta_k})\right) \;\to\; J_1\left(\mathbf{w}(\bar{\mathbf{e}})\right), \quad \text{in } \mathbb{R}.$$

We now turn to the convergence of the L^2-criterion J_0. We note

$$\begin{aligned}
\mathbf{p}^L_\Delta(\mathbf{e}_\Delta)(t) &= \mathbf{w}_\Delta(\mathbf{e}_\Delta)(L,t) - \mathbf{w}^{\text{obs}}(t),\\
\mathbf{p}^L(\bar{\mathbf{e}})(t) &= \mathbf{w}(\bar{\mathbf{e}})(L,t) - \mathbf{w}^{\text{obs}}(t),
\end{aligned}$$

and we write :

$$\begin{aligned}
J_0\left(\mathbf{w}_\Delta(\mathbf{e}_\Delta)\right) &= \int_0^T \mathbf{p}^L_{\Delta_k}(\mathbf{e}_{\Delta_k})\cdot\left(\mathbf{p}^L_{\Delta_k}(\mathbf{e}_{\Delta_k}) - \mathbf{p}^L(\bar{\mathbf{e}})\right)dt\\
&+ \int_0^T \mathbf{p}^L_{\Delta_k}(\mathbf{e}_{\Delta_k})\cdot\mathbf{p}^L(\bar{\mathbf{e}})dt
\end{aligned} \tag{55}$$

Using the $L^1_{loc}(\Omega)$ and a.e. convergences (51), and the L^∞-stability (50), we deduce

$$\begin{aligned}
&\left|\int_0^T \mathbf{p}^L_{\Delta_k}(\mathbf{e}_{\Delta_k})\cdot\left(\mathbf{p}^L_{\Delta_k}(\mathbf{e}_{\Delta_k}) - \mathbf{p}^L(\bar{\mathbf{e}})\right)dt\right| \le\\
&C\int_0^T |\mathbf{w}_{\Delta_k}(\mathbf{e}_{\Delta_k})(L,t) - \mathbf{w}(\bar{\mathbf{e}})(L,t)|dt \;\to\; 0,
\end{aligned} \tag{56}$$

when $\Delta_k z, \Delta_k t \to 0$, and we also deduce

$$\int_0^T \mathbf{p}^L_{\Delta_k}(\mathbf{e}_{\Delta_k})\cdot\mathbf{p}^L(\bar{\mathbf{e}})dt \to J_0\left(\mathbf{w}(\bar{\mathbf{e}})\right), \quad \text{when } \Delta_k z, \Delta_k t \to 0. \tag{57}$$

We replace the convergences (56), (57) in the relation (55), and we obtain

$$J_0\left(\mathbf{w}_{\Delta_k}(\mathbf{e}_{\Delta_k})\right) \;\to\; J_0\left(\mathbf{w}(\bar{\mathbf{e}})\right), \quad \text{when } \Delta_k x, \Delta_k t \to 0.$$

We have the convergence of (54) for the L^2-criterion and the momentum-criterion of order p. Finally, by the definition of the continuous identification problem (53) and the discretised identification problem (49), we deduce that $\mathbf{e}_{\Delta_k} \to \bar{\mathbf{e}}$ is a minimizing sequence, and then

$$J(\mathbf{w}(\bar{\mathbf{e}})) = \min_{\mathbf{e}\in\mathcal{K}} J(\mathbf{w}(\mathbf{e}))$$

Because the flux F is a continuous function, we can easily prove the same results for flux-criteria $\mathcal{J}_0$, and $\mathcal{J}_p$, for $p \geq 1$, thus the claim is verified. ∎

5 Numerical implementation

We propose in this section some remarks on the numerical implementation of the method, and two examples of identification. Unfortunately, real experimental data are quite difficult to obtain, so we deal only with two "school cases". By this we mean that the isotherm to identify is completely known, the observed peaks are results of a simulation. We are thus supposed to find this exact solution.

The first test case deals with a scalar conservation law - monocomponent chromatography - with several initial data. We observe a stability result when the mesh is refined. The second test concerns a system of two equations, and gives spectacular accuracy on the identified peaks.

5.1 Remarks about algorithms

Note first that it will never be possible to identify the flux as a function of $\mathbb{R}^n$ into $\mathbb{R}^n$, because of the severe ill-posedness mentioned before. Instead, we make use of the explicit models given in Section 2.1 (see [5], and [6]), thus being led to identify a finite number of parameters. This actually suppresses a part of the indetermination, by specifying a particular shape for the isotherm.

Then we limit ourselves during the computations to the Godunov scheme and the related adjoint scheme. We deal indeed with a non-linear optimization problem in $\mathbb{R}^N$, which we solve by a conjugate gradient method. The correct descent direction is chosen by a Polak-Ribière method [9]. Thus within each iteration of gradient, we have to minimize the cost J along several directions, that is to solve for each direction the direct problem (3). This leads to a rather heavy computational cost, so we use the simplest discretization scheme. The Van-Leer scheme is of order 2, but it increases drastically the computation time, and its adjoint scheme (44)-(47) is difficult to implement. It seems out of reach on a workstation.

Finally, the remarks about the cost function in section 2.2 lead us to use the function $\mathcal{J}_\rho$ defined by (17), with a careful choice of ρ. It turns out that the set of parameters $\mathbf{e}$ in the one component isotherm model (5), or the two components isotherm model (6), can be split into $\mathbf{e} = (\mathbf{K}, \mathbf{E})$, where the subset of Langmuir coefficients $\mathbf{K}$ acts essentially on the retention time, and the interaction energies $\mathbf{E}$ on the shape of the peaks. Namely, in the case of compactly supported data, the coefficients acting on the retention time are characterized by $\mathbf{K} = \left(\frac{\partial H}{\partial w_m}(0)\right)_{m=1,\dots M}$. Recall that

$$\mathcal{J}_\rho = \mathcal{J}_0 + \rho \mathcal{J}_1,$$

$\mathcal{J}_0$ being the L^2-criterion, which acts on the shape, $\mathcal{J}_1$ the momentum criterion, acting on the retention time on the support. It is therefore natural to split the minimization in the following way:

$$\min_{\mathbf{K}} \left[J_0 + \rho J_1\right] \left(\mathbf{w}(\mathbf{K}, \mathbf{E})\right), \tag{58}$$

denoting the solution $\bar{\mathbf{K}}$. Then, solve

$$\min_{\mathbf{E}} J_0 \left(\mathbf{w}(\bar{\mathbf{K}}, \mathbf{E})\right). \tag{59}$$

In this framework, we expect that $\rho J_1 \to 0$ when the number of iterations tends to $+\infty$. Thus we solve (9). This is the method we used with success in the computations.

As a last comment on the identification method, we should say that we applied a gradient method to minimize a functional which is non smooth and not globally convex. The gradient method may converge to a minimum provided the initial guess is "close enough" to the solution - this is a general feature in descent methods. Moreover we may converge to a local minimum of the functional, see an example in [13]. However, the discrete problem is well-posed and seems to converge: in particular, it gives a meaning to the adjoint state. One can interpret that in the scalar case in terms of nonconservative products of distributions (see [13] and a future publication).

5.2 Single component identification

Our first test case consists of identifying a single component isotherm. The observation is given here by a set of simulated peaks with several initial and injection data. Our purpose was also to test the stability of the method with respect to refinement of the discretization.

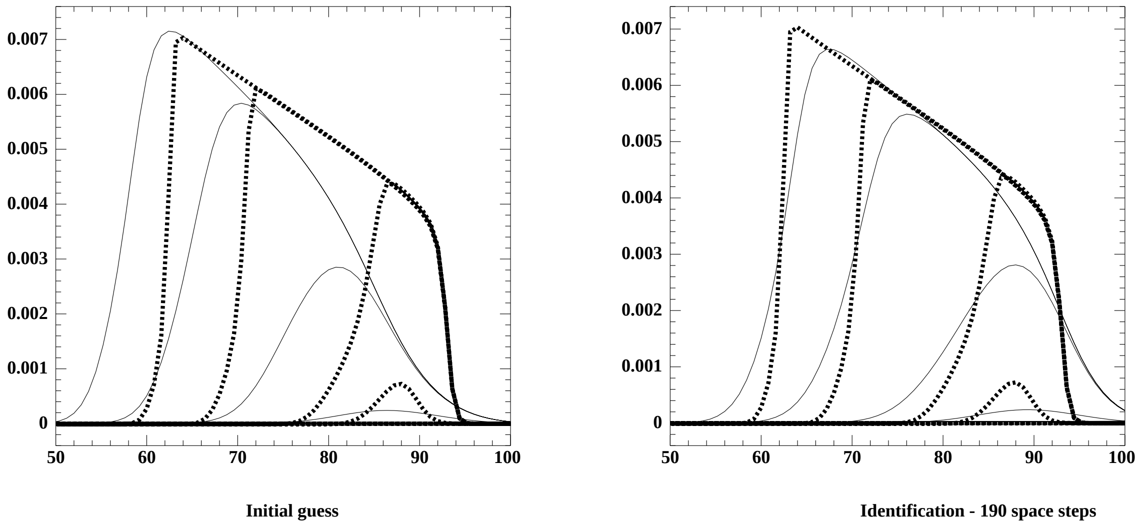

Figure 1: *Initial guess and identification ; 190 space steps*

In this respect, to avoid heavy computations, we used the following method. Apply first the method on a coarse grid. It converges quickly to a solution which is not very good (fig. 1) because the scheme has an important numerical viscosity. But we can use this solution as initial guess for another identification process, with a more refined mesh, and so on. The identification result is excellent on the peaks (fig. 2), and the isotherm itself (which is of course our main goal) is reasonably approximated.

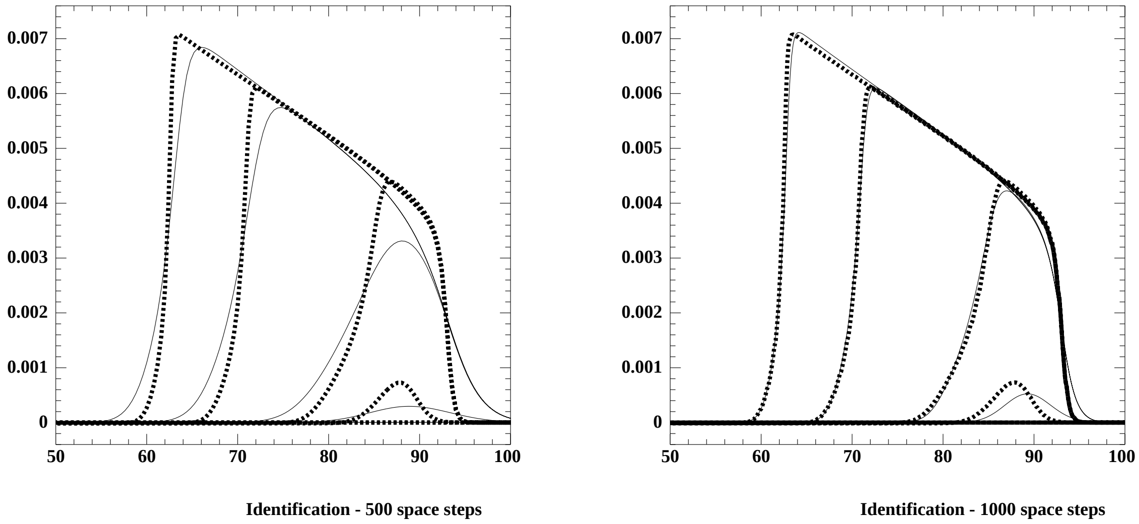

Figure 2: *Mesh refinement ; 500 and 1000 space steps*

A comparison between the identified isotherms is given in the Figure 3.

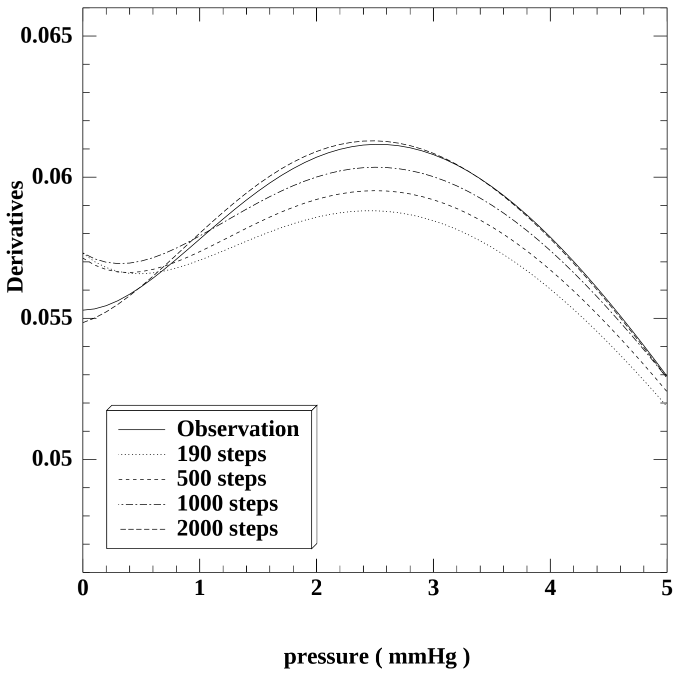

Figure 3: *Derivative of the isotherm with the identified parameters*

5.3 Another test case

Here we show an identification on a two component isotherm. Once again, the observation is the result of a set of simulations. The final result is very accurate on the peaks, but the computation time on a workstation becomes very long. This kind of problems deserve obviously large computers for reasonable exploitation.

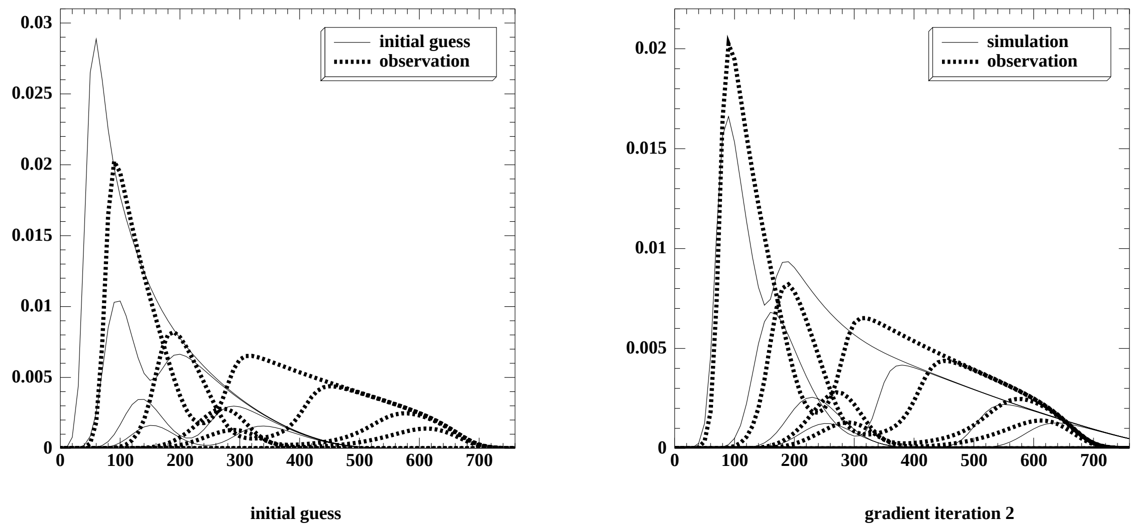

Figure 4: *Initial Guess, and result after the second iteration of the gradient.*

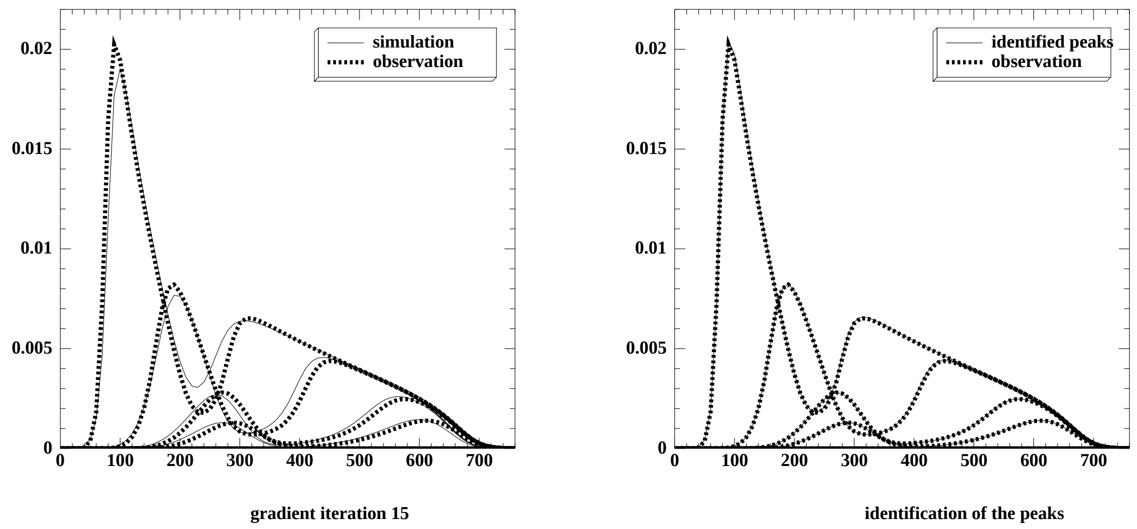

Figure 5: *Iteration # 15, and iteration # 18 (identification result).*

6 Conclusion

We proposed here an approach to a problem by nature ill-posed in uniqueness. However, the use of relevant models for the direct problem allows us to deal with the indetermination. The formal computations lead to numerical algorithms which converge to physically relevant solutions. The mathematical analysis is, on the other hand, out of reach, except in the scalar case.

Futher developments can be suggested, in particular one can try to discretize the continuous adjoint state, in contrast with our " exact gradient of discretized cost function" method. This may give the ability to optimize the computed adjoint state, and thus improve the convergence. We noticed the crucial choice of the initial guess for the gradient method.

A natural idea is to make use of global optimization methods to compute a predictor, and then perform the gradient. Simulated annealing and genetic algorithms are designed to globally minimize non smooth functionals. The latter can even deal with non unique extrema, and could even be used without the a priori modelling of the isotherm. We would like finally to remark that the identification problem is very close to the problems of control in coefficients. The techniques developed here might apply with few modifications.

References

[1] G. Chavent, *Identification of distributed parameter systems : about the output least square method, its implementation, and identifiability* Proceeding of the 5th. IFAC Symposium on Identification and System Parameter Estimations, Pergamon Press, Vol 1, pp. 85-97, 1979.

[2] Freundlich, *Colloid and Capillary Chemistry* Dutton, New-York, 1926.

[3] E. Godlewski, P.A. Raviart, *Hyperbolic systems of conservation laws.* Mathématiques et applications; n° 3/4. Paris, Ellipses, 1991.

[4] F. James, *Sur la modélisation mathématique des équilibres diphasiques et des colonnes de chromatographie.* Thèse de l'Ecole Polytechnique, 1990.

[5] F.James, P. Valentin, *An Analytic Model of Multicomponent Isothermal Adsorption Equilibrium*, Submitted to Continum Mechanics & Thermodynamics.

[6] F. James, M. Sepúlveda, P. Valentin, *Modèles de thermodynamique statistique pour un isotherme d'équilibre diphasique multicomposant.* Rapport interne n°223, CMAP Ecole Polytechnique, (1990).

[7] E. Kvaalen, L. Neel, D. Tondeur, *Directions of Quasi-static Mass and Energy Transfer Between Phases in Multicomponent Open Systems*, Chem. Eng. Sc. **40** n° 7 (1985) p. 1191-1204.

[8] I. Langmuir, *Jour. Am. Chem. Soc.*, **38** (1916), p.2221.

[9] E. Polak, G. Ribière, *Note sur la convergence de méthodes de directions conjuguées*, Rev. Fr. Inform Rech. Operation. (16-R1), 35-43 (1969).

[10] H.K. Rhee, R. Aris, N.R. Amundson, *On the Theory of Multicomponent Chromatography* Philos. Trans. Roy. Soc. London, **A 269** (1970), pp.419-455.

[11] R. Rota, G. Gamba, R. Paludetto, S. Carra, M. Morbidelli, *Generalized Statistical Model for the Multivomponent Adsorption Equilibria on Zeolithes.* Ind. Eng. Chem. Fundam., **27** (1988), pp.848-851

[12] D. Ruthven, F. Wong, *Generalized Statistical Model for the Prediction of Binary Adsorption Equilibria on Zeolithes.* Ind. Eng. Chem. Fundam., **24** (1985), pp.27-32

[13] M. Sepúlveda, *Identification de paramètres pour un système hyperbolique. Application à l'estimation des isothermes en chromatographie.* Thèse de l'Ecole Polytechnique, 1992.

[14] Sips, J. Chem. Phys., **18** (1950), p.1024.

[15] P. Valentin, G. Guiochon, *Propagation of Finite Concentration in Gas Chromatography.*Separation Science, **10** (1975), pp.245-305.

[16] P. Valentin, G. Guiochon, *Determination of Gas-liquid and Gas-solid Equilibrium Isotherms by Chromatography, I & II.* Jour. Chromatographic Sc.,**14** (1976), pp.56-63 & 132-139.