1

Velocity Estimation in Mixtures using Tomography

E. Cahyono, M.E. Hochstenbach, J. Molenaar,
W.H.A. Schilders, G.M. Terra

Abstract

In oil production a lot of water is usually pumped up together with the oil. For many reasons the reduction of the water production is a very important issue. The method presented in this paper is meant to provide a necessary tool for this. Most drilling wells consist of a network of bore holes. Some of them may produce water, others oil or a mixture. At the moment the net flow of all bore holes together is brought to the surface. It is desirable to be able to detect how much water a specific bore hole contributes. If this amount surpasses a critical value one could then consider to close that bore hole. This leads to the question how the composition of the flow in a pipe can be determined *in situ*. In this paper we analyze how tomography techniques, well-known from medical applications, can be applied in the case of a bore hole. These techniques allow to measure instantaneously the mass distribution over a cross section of the pipe. For velocity estimation, the idea is to detect the mass distributions at two neighbouring cross sections at successive times. Correlating the obtained time series, one might be able to estimate the local velocity profile. The basic idea was already mentioned in literature before, but it was believed that the number of correlations to be evaluated is so huge, that the approach would fail in practice. In this paper we describe the mathematical details of the method and conclude that the number of time consuming calculations is not necessarily a limiting factor. In addition, suggestions are made to facilitate the use of tomography for velocity estimation.

Keywords

Velocity estimation in two-phase flow, oil well management, electrical impedance tomography.

1.1 Introduction

In oil production plants a lot of water is brought to the earth's surface, together with the oil. There are several reasons to avoid this as much as possible. Since one is only interested in oil, this has to be extracted from the mixture. The remaining water is still polluted, so it can not be dumped into the environment. Moreover, one is not allowed to extract too much water from the soil.

Therefore it is worthwhile to apply techniques to reduce the amount of water that is pumped up. One of these techniques is to use 'intelligent wells'. In fact, most wells consist of many subbranches. The key idea is to shut off a branch as soon as it starts to produce too much water. This strategy can be applied only if one is able to determine which branches produce mainly water. Hence one has to measure *in situ* the flow of water and oil through the pipelines separately. There are no satisfactory methods to do so yet. It is possible to measure the total flow in situ, and, of course, one can determine the fractions of oil and water at the surface, but one can not measure the flows of oil and water in situ separately.

This paper deals with a method which might overcome this problem by using electrical impedance tomography [2]. Measuring cross-sectional distributions of oil and water at two separate cross sections and looking at the correlations between these two distribution in time, one could estimate the velocity distribution in the pipe (see [3]). By coupling the information about the distributions of the fluid's components and the velocity distribution, the fluxes of water and oil can be determined separately.

The problem with this method, posed to us during the study group, lies in the number of correlations that have to be calculated in real time. This number was thought to scale with L^8 where L is the number[1] of electrodes used in the tomographic device. This would be too time consuming for the computer power that can be installed. One should bear in mind that the computer must work under very extreme conditions underground. Moreover, it is not possible to send the information to the surface to perform the computations there. Therefore we were asked to think about a way to reduce the number of calculations that have to be done.

In fact, our main contribution to the solution of this problem is that we dispute the scaling with L^8. We claim that the number of independent correlations scales with L^4. This is an enormous reduction in practice since the available computers may then deal with this problem in real time.

This paper is organized as follows. Section 1.2 deals with tomography. First, a mathematical model describing the situation, is introduced in Subsection 1.2.1. The ideas behind tomography are explained in Subsection 1.2.2, followed by a one-dimensional example in Subsection 1.2.3. Next, an efficient way to solve the tomographic problem is given in subsections 1.2.4 and 1.2.5. The subsequent section deals with the method used to derive estimates for the velocity distribution by correlating between two tomographic cross sections. In Section 1.4, our claim that the number of

[1] Typically $L \approx 32$.

calculations scales with L^4 is justified. Finally, conclusions are given in Section 1.5. An appendix is included to describe a reformulation of the correlations, which could be useful to reduce calculation times further.

1.2 Tomography

This section deals with electrical impedance tomography. A mathematical model describing the setup, is given in Subsection 1.2.1. Subsection 1.2.2 explains the idea behind the tomography after which a one-dimensional example is discussed in Subsection 1.2.3. Subsection 1.2.4 deals with a way to describe the cross-sectional conductivity distribution. This is used in Subsection 1.2.5 describing a particular method for solving the tomographic problem.

1.2.1 The mathematical model

In Figure 1.1, the cross section of a pipe with area Ω and boundary $\partial\Omega$ is sketched. The pipe contains a mixture of water and oil. We would like to use information obtained at the boundary $\partial\Omega$ to approximate the distribution of the fluids in Ω. To this end, L electrodes e_l are attached to the boundary $l = 1, \ldots, L$. Applying currents to the electrodes and measuring the resulting voltages at other electrodes, we obtain data. In the following we briefly discuss a mathematical model which allows us to estimate the oil and water distribution in Ω from these data.

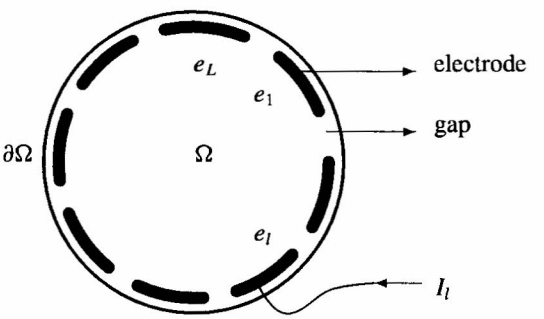

Figure 1.1: A cross section of the pipe

The electric potential or voltage $u(\mathbf{x})$ in Ω is governed by the equation

$$\nabla \cdot (\gamma(\mathbf{x}, \omega)\nabla u(\mathbf{x})) = 0. \tag{1.1}$$

Here $\mathbf{x}$ is a point in Ω, γ is the electric admittivity given by $\gamma(\mathbf{x}, \omega) = \sigma(\mathbf{x}, \omega) + i\omega\epsilon(\mathbf{x}, \omega)$, where σ is the electric conductivity, ϵ is the electric permittivity and ω is the angular frequency of the applied current. Note that the estimation procedure is independent of the value of ω. In fact, we consider the frequency as a fixed given quantity, determined by the apparatus. Equation (1.1) can be obtained from Maxwell's equations, see [2]. Note that the current density is given by $j = -\gamma\nabla u$ (local Ohm's law).

During the measurement one could apply currents on the electrodes and measure the resulting voltages, or, the other way around, apply voltages and measure the currents. In practice, the first approach is followed. The boundary conditions are determined by the currents which produce a current density on each electrode. The integral of this current density over the electrode is equal to the total current that flows through the electrode. Hence, we have

$$\int_{e_l} \gamma \frac{\partial u}{\partial \nu} \, \mathrm{d}s = I_l \qquad l = 1, \ldots, L, \tag{1.2}$$

where I_l is the (inward) current through the electrode e_l and $\frac{\partial}{\partial \nu}$ is the derivative taken along the outward normal. Through the gaps between the electrodes there is no current, so there we have

$$\gamma \frac{\partial u}{\partial \nu} = 0. \tag{1.3}$$

On the boundary, we have to take into account the electro-chemical effect that takes place at the contact between the electrodes and the fluids. There, the formation of a thin, highly resistive layer is observed. The *impedance* of this layer at electrode e_l is denoted by z_l. The boundary condition on electrode e_l reads as

$$u + z_l \gamma \frac{\partial u}{\partial \nu} = V_l \qquad l = 1, \ldots, L. \tag{1.4}$$

z_l is called the effective *contact impedance* or *surface impedance*.

Since the total amount of charge in the system remains conserved, it holds that

$$\sum_{l=1}^{L} I_l = 0. \tag{1.5}$$

The ground potential is chosen such that

$$\sum_{l=1}^{L} V_l = 0. \tag{1.6}$$

1.2.2 The idea behind tomography

This section deals with the principles of electrical impedance tomography. The setup described in the previous section is used to determine the composition of the material in the area Ω. Summarizing Section 1.2.1 we have the following equations for the voltage u in the cross section Ω:

$$\begin{aligned} \nabla \cdot \sigma \nabla u &= 0 \quad \text{in } \Omega, \\ u + z_l \sigma \frac{\partial u}{\partial \nu} &= V_l \quad \text{on electrode } e_l \text{ for } l = 1, \ldots, L, \\ \int_{e_l} \sigma \frac{\partial u}{\partial \nu} \, \mathrm{d}s &= I_l \quad \text{for } l = 1, \ldots, L, \\ \sigma \frac{\partial u}{\partial \nu} &= 0 \quad \text{through the gaps,} \end{aligned} \tag{1.7}$$

together with the calibration $\sum_{l=1}^{L} V_l = 0$ and mass conservation $\sum_{l=1}^{L} I_l = 0$. Note that the (complex-valued) admittivity γ has been replaced by the (real-valued) conductivity σ for simplicity. Like in [2], we consider only the problem of approximating the conductivity σ. Existence and uniqueness of the solution u, given the currents I_l and the conductivity distribution $\sigma(\mathbf{x})$, have been proved in [1]. This yields the potential distribution $u(\mathbf{x})$ and the voltages V_l on the electrodes. In case of tomography we would like to do it the other way around. We apply currents I_l, measure the voltages V_l, and from this we try to determine the distribution $\sigma(\mathbf{x})$. In fact, the potential distribution $u(\mathbf{x})$ itself is not of interest here. Only the voltages V_l at the electrodes are used eventually to find the conductivity distribution $\sigma(\mathbf{x})$.

For this purpose, $L-1$ independent current patterns are applied, specifying the currents I_l for $l = 1, 2, \ldots, L$ such that the condition $\sum_{l=1}^{L} I_l = 0$ holds, and the resulting voltages V_l, for $l = 1, 2, \ldots, L$, are measured. These patterns will be denoted by $I^k = (I_1^k, I_2^k, \ldots, I_L^k)$, with corresponding *measured* voltages $V^k = (V_1^k, V_2^k, \ldots, V_L^k)$, for $k = 1, 2, \ldots, L-1$. By way of illustration, suppose that current pattern I^k has positive (inward) current I_k^k through electrode k which flows outward through the opposite electrode. Other patterns are linearly dependent on these ones. Because the model (1.7) is linear in $u(\mathbf{x})$, V_l and I_l, those patterns do not provide additional information. Because of the choice of ground potential (the condition $\sum_{l=1}^{L} V_l^k = 0$), the number of independent measurements per current pattern is $L-1$. However, the total number of independent measurements is even less than $(L-1)^2$. Due to the symmetry under changing the direction of the currents, the number of independent measurements is $L(L-1)/2$, which is the number of degrees of freedom in a symmetric $(L-1) \times (L-1)$ matrix. Intuitively, this is because powering up electrode k and measuring at electrode l yields the same information as powering up electrode l and measuring at electrode k.

Given a conductivity distribution $\sigma(\mathbf{x})$, for each current pattern I^k, with $k = 1, 2, \ldots, L-1$, the model (1.7) can be solved in principal, usually numerically. In

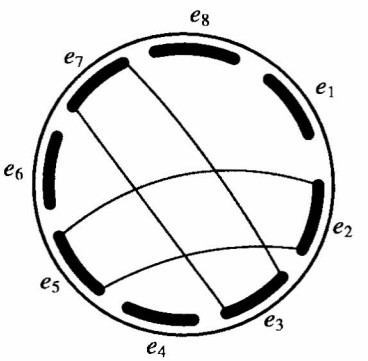

Figure 1.2: A sketch illustrating the sensing areas between e_2–e_5 and between e_3–e_7

particular, the corresponding voltages at the electrodes are calculated. We denote these *calculated* voltages by $U^k(\sigma)$. Just like the measured voltages V^k, each $U^k(\sigma)$ is itself a vector of voltages at all of the electrodes. Furthermore, it depends on the conductivity distribution $\sigma(\mathbf{x})$. The aim of tomography is to find the distribution $\sigma(\mathbf{x})$ for which the measured voltages V_l^k and the calculated voltages $U_l^k(\sigma)$ agree best. Note that the conductivity distribution is assumed to be constant on the timescale of the tomography. The tomographic intersection is considered to be a 'snapshot' of the fluid's composition at a certain time. This can be done because it is possible to switch between the different current patterns very fast electronically; adaptation to (1.7) is almost instantaneously.

Many algorithms have been proposed for reconstructing the distribution σ from the measured voltages V_l^k. An extensive list of methods and references can be found in [2, page 90]. For example, the linear backprojection method assumes that the measurement of the voltage at one electrode when powering up another depends only on the material inside a so-called "sensing area" from the former electrode to the latter, as illustrated in Figure 1.2. Such a sensing area exists for every combination of two electrodes, hence there are $L(L-1)/2$ of them. Note that this is the same as the previously mentioned number of independent measurements. The intersections of these areas lead to a mesh naturally. A linear formula is used to calculate the values of σ on this mesh from the measurements. This method is fast, but crude. In particular, because there are much more mesh cells (one for each intersection of two sensing areas) than independent measurements, these values of σ can not be independent; there must be relationships between the various values.

Another method is the Noser algorithm, which is discussed in Section 1.2.5.

1.2.3 A one-dimensional example

To get a better understanding of the mathematical model given in Section 1.2.1 and the ideas in the previous section, we shall work out the one-dimensional case in detail. We take for Ω the interval $[-1, 1]$, and for simplicity, we consider only the problem of reconstructing conductivity σ in case of vanishing surface impedance, so $z_l = 0$ for $l = 1, 2$ and vanishing permittivity ϵ. Hence, equations (1.1), (1.2), (1.4), (1.5) and (1.6) become

$$\partial_x (\sigma(x)\, \partial_x u(x)) = 0, \tag{1.8}$$

$$\sigma(-1) \left.\frac{\partial u}{\partial x}\right|_{x=-1} = -I_1, \quad \sigma(1) \left.\frac{\partial u}{\partial x}\right|_{x=1} = I_2, \tag{1.9}$$

$$u(-1) = V_1, \quad u(1) = V_2, \tag{1.10}$$

$$I_1 + I_2 = 0, \tag{1.11}$$

$$V_1 + V_2 = 0, \tag{1.12}$$

respectively. Note that the derivatives in (1.2) are taken along the outward normal. This leads to the minus sign in the first equation in (1.9).

We remark that equations (1.8)-(1.12) are linear in u, I_1, I_2, V_1 and V_2. Furthermore, from (1.11) and (1.12) it directly follows that I_2 and V_2 are not independent from I_1 and V_1. The one-dimensional system is in fact a highly reduced case since only one independent measurement is available. One can set the value of I_1 and measure the resulting V_1 value. This can be done at most once, since all other measurements do not yield extra information for the estimation of σ. The number of independent measurements to be used in the estimation is called the number of degrees of freedom of the system. This concept is more interesting in the two-dimensional case and is discussed in more detail in other sections. In this terminology, the one-dimensional case has one degree of freedom. This implies that if $\sigma(x)$ is expanded in a power series or a Fourier series on the interval $[-1, 1]$, only one coefficient of such a series can be estimated. Let us, for purpose of illustration, approximate $\sigma(x)$ by the constant σ_0, $\sigma_0 \neq 0$. Then from (1.8) we find that

$$u(x) = ax + b \tag{1.13}$$

for some constants a and b. From (1.9) we obtain that $a = -I_1/\sigma_0$ and by combining (1.10), (1.12) and (1.13) we conclude that $b = 0$. So,

$$u(x) = \frac{-I_1}{\sigma_0} x. \tag{1.14}$$

Evaluation at $x = 1$ yields for σ_0 the estimate

$$\sigma_0 = \frac{I_1}{V_1}, \tag{1.15}$$

which is just Ohm's law.

1.2.4 Representation of the conductivity

For given conductivity σ, the tomography model embodied in equations (1.1)–(1.6) is linear in the electric potential u, the measured values of the currents I_l, and potentials V_l of the boundary. It can be concluded (see [2] or Section 1.2.2) that for a system with L electrodes, at most $L(L-1)/2$ linearly independent measurements are available. These 'degrees of freedom' can be used to fit the distribution σ over the cross section by means of a fitting function that contains at most $L(L-1)/2$ parameters. A convenient approach is to expand $\sigma(t, \mathbf{x})$ in terms of a set of basis functions B_i, $i = 1, \ldots, N$, with $N \leq L(L-1)/2$, which are defined on the cross section. The estimate $\hat{\sigma}$ for σ is written as

$$\hat{\sigma}(t, \mathbf{x}) = \sum_{i=1}^{N} \sigma_i(t) B_i(\mathbf{x}) \; . \tag{1.16}$$

Note that the spatial dependence of σ is represented via the basis functions B_i, whereas the time dependence is included in the coefficients σ_i.

It is advantageous to take the B_i orthonormal, i.e.

$$(B_i, B_j) := \int_{\text{cross section}} B_i(\mathbf{x}) B_j(\mathbf{x}) \, \mathrm{d}\mathbf{x} = \delta_{ij} \; , \tag{1.17}$$

with $\delta_{i,j}$ the Kronecker delta, which is non-vanishing only if $i = j$. Then, it holds that the σ_i simply follow from the inner product

$$\sigma_i(t) = (\sigma, B_i) = \int_{\text{cross section}} \sigma(t, \mathbf{x}) B_i(\mathbf{x}) \, \mathrm{d}\mathbf{x} \; . \tag{1.18}$$

For the B_i several choices are possible. Natural candidates are obtained by dividing the cross section into N disjunct cells with areas A_i, $i = 1, \ldots, N$. The basis functions could then be defined as

$$B_i(\mathbf{x}) = \begin{cases} 1 & , \; x \in \text{cell } i \\ 0 & , \; x \notin \text{cell } i \end{cases} \tag{1.19}$$

Note that these basis functions are othogonal but not orthonormal. For this choice the coefficients σ_i have a clear geometrical interpretation:

$$\sigma_i(t) = \frac{1}{A_i} \int_{\text{cell } i} \sigma(t, \mathbf{x}) \, \mathrm{d}\mathbf{x} \; , \tag{1.20}$$

so $\sigma_i(t)$ is the average value of σ over cell i at time t. In this case the continuous function $\sigma(t, \mathbf{x})$ is approximated by N discrete levels.

If the cross section is the unit square, an alternative choice for the B_i is to make use of Fourier basis functions ($\mathbf{x} = (x, y)$, $n, m = 0, 1, 2, \ldots$)

$$\begin{aligned} &\cos(2\pi nx)\ \cos(2\pi my), \\ &\cos(2\pi nx)\ \sin(2\pi my), \\ &\sin(2\pi nx)\ \cos(2\pi my), \\ &\sin(2\pi nx)\ \sin(2\pi my). \end{aligned} \tag{1.21}$$

These functions form an orthogonal set with respect to the inner product introduced in (1.17) and can be easily normalized. For a cylindrical cross section it would be natural to use as basic functions products of Bessel functions and harmonics. In both cases the coefficients σ_i do not represent the amplitude of σ around a specific location, but measure the relative strength of the 'modes', i.e. the basis functions. Since the specific choice of the B_i is not a central issue of this project we do not work out this aspect in detail.

An important aspect to mention is the smoothness of the representation. If use is made of (1.19), the σ distribution is discretized using N cells. Plotting this discontinuous estimate of σ, we obtain a non-smooth plot. On the contrary, representation (1.21) is a continuous function over the cross section and the corresponding plot will look much more reliable. Smooth basis functions like the ones in (1.21) always lead to a smooth visualization. However, this does not necessarily imply that the accuracy is higher than that of a discontinuous representation. In the two examples given here the local basis functions (1.19) are discontinuous, whereas the global basis functions (1.21) are continuous. This need not be the case in general. It is also possible to choose localized smooth functions as well as discontinuous global basis functions.

1.2.5 The Noser Algorithm and extensions

In this section we discuss a specific tomography method, the Noser algorithm, and some generalizations and modifications. We have already seen in Section 1.2.2 that we have $L(L-1)/2$ degrees of freedom, the independent measurements. We can recover at most $L(L-1)/2$ degrees of freedom of σ.

In general it may be wise to recover less degrees of freedom, say $N \leq L(L-1)/2$, to reduce the influence of 'noise' (such as measurement and rounding errors) in the model. The idea is to approximate σ by a linear combination of N basisfunctions. A simple choice is to split up the domain, and take the indicator functions on mesh elements as basisfunctions. Then the approximation to σ will be a piecewise constant function, constant on each mesh element. These constants can be thought of as the average of σ on the mesh elements. But we may also consider more sophisticated systems of basisfunctions, see Section 1.2.4.

Once we have chosen our N basisfunctions, we would like to find an approximation $\sigma := (\sigma_1, \sigma_2, \ldots, \sigma_N)$, such that the corresponding voltage patterns $U^1(\sigma)$,

$U^2(\sigma), \ldots, U^{L-1}(\sigma)$ agree with the measured ones $V := (V^1, V^2, \ldots, V^{L-1})$ as much as possible[2]. For this goal, we want to minimize the least-squares functional

$$E(\sigma) = \sum_{k=1}^{L-1} \|U^k(\sigma) - V^k\|^2.$$

Differentiating with respect to the n-th component of σ, and setting the derivative equal to zero we obtain the condition

$$f_n(\sigma) := \frac{\partial E}{\partial \sigma_n}(\sigma) = 2 \sum_{k=1}^{L-1} (U^k(\sigma) - V^k) \cdot \frac{\partial U^k}{\partial \sigma_n}(\sigma) = 0.$$

These N nonlinear equations are summarized as

$$f(\sigma, V) = 0, \qquad f : \mathbb{R}^N \times \mathbb{R}^{L(L-1)/2} \to \mathbb{R}^N.$$

So, σ implicitly depends on the measurements V. Suppose that we start with $\sigma^{(0)}$ as an initial guess. The model then provides us with $U(\sigma^{(0)})$ and $D_\sigma U(\sigma^{(0)})$ (where the D denotes the total derivative). If our measurements were $V^{(0)} := U(\sigma^{(0)})$, then $\sigma^{(0)}$ would be a solution, but in general $V \neq V^{(0)}$.

The Implicit Function Theorem states that if $D_\sigma f(\sigma^{(0)}, V^{(0)})$ is invertible, then, in a neighborhood of $(\sigma^{(0)}, V^{(0)})$, there exists for all V a unique σ such that $f(\sigma, V) = 0$. This defines a function φ:

$$\sigma = \varphi(V), \qquad \varphi : \mathbb{R}^{L(L-1)/2} \to \mathbb{R}^N.$$

Moreover, the Implicit Function Theorem gives that

$$D\varphi(V) = -\left(D_\sigma f(\sigma^{(0)}, V^{(0)})\right)^{-1} \cdot D_V f(\sigma^{(0)}, V^{(0)}).$$

A Taylor expansion of φ around $V^{(0)}$ gives

$$\sigma = \sigma^{(0)} + D\varphi(V^{(0)})(V - V^{(0)}) + \text{ higher order terms.}$$

Neglecting the higher order terms we obtain a new approximation $\sigma^{(1)}$:

$$\sigma^{(1)} = \sigma^{(0)} - \left(D_\sigma f(\sigma^{(0)}, V^{(0)})\right)^{-1} \cdot D_V f(\sigma^{(0)}, V^{(0)}) \cdot (V - V^{(0)}). \tag{1.22}$$

Compare this with a Newton step applied to the function f, to find σ given the measurements V:

$$\sigma^{(1)} = \sigma^{(0)} - \left(D_\sigma f(\sigma^{(0)}, V)\right)^{-1} \cdot f(\sigma^{(0)}, V). \tag{1.23}$$

[2]Note that, with a slight abuse of notation σ is used to denote the vector of amplitudes $\sigma_i(t)$ in this section.

It can be shown that when we linearize (1.23) around the point $V^{(0)}$, we get exactly (1.22). For this reason, we call (1.22) a Newton step as well. When we take for the initial guess $\sigma^{(0)}$ a uniform conductivity in (1.22), we get the *Noser algorithm*. This method has the advantage that $\left(D_\sigma f(\sigma^{(0)}, V^{(0)})\right)^{-1} \cdot D_V f(\sigma^{(0)}, V^{(0)})$ can be computed (and stored) beforehand.

So, the Noser algorithm is in fact the first step of the following Newton algorithm:

Input: an initial guess $\sigma^{(0)}$ and a number of steps K
Output: an (hopefully better) approximation $\sigma^{(K)}$
For $s = 0, \ldots, K-1$ **do**
1. $\sigma^{(s+1)} = \sigma^{(s)} - \left(D_\sigma f(\sigma^{(s)}, V^{(s)})\right)^{-1} \cdot D_V f(\sigma^{(s)}, V^{(s)}) \cdot (V - V^{(s)})$
2. Compute $U(\sigma^{(s+1)})$ and $D_\sigma U(\sigma^{(s+1)})$ from the model

Newton's method to approximate σ.

In the first step of this algorithm,

$$M^{(s)} := \left(D_\sigma f(\sigma^{(s)}, V^{(s)})\right)^{-1} \cdot D_V f(\sigma^{(s)}, V^{(s)}) \tag{1.24}$$

looks troublesome. However, it may not be prohibitively expensive, considering the following remarks:

1. Suppose that we would like to compute $x := \left(D_\sigma f(\sigma^{(s)}, V^{(s)})\right)^{-1} y$ for a vector y. Instead of computing $\left(D_\sigma f(\sigma^{(s)}, V^{(s)})\right)^{-1}$, and multiplying y by this matrix, it is in general much cheaper to solve the linear system $D_\sigma f(\sigma^{(s)}, V^{(s)})x = y$. Numerical linear algebra provides us with a number of different approaches to solve this system. Which method is best depends on the properties of the system. It may also be solved approximately, the resulting method is called an *inexact Newton method*.

2. Instead of using $\left(D_\sigma f(\sigma^{(s)}, V^{(s)})\right)^{-1} \cdot D_V f(\sigma^{(s)}, V^{(s)})$ each step, we may also try cheaper (Newton) variants of the form $\sigma^{(s+1)} = \sigma^{(s)} - M \cdot (V - V^{(s)})$. For example, M may be taken $\left(D_\sigma f(\sigma^{(0)}, V^{(0)})\right)^{-1} \cdot D_V f(\sigma^{(0)}, V^{(0)})$ all the time, or updated cheaply (e.g. *Broyden's method*). Another idea is to approximate both Jacobians (e.g. by *finite-difference Newton*) or only one of them.

In conclusion, when we have reasonable approximations to the Jacobians (Step 1 of the algorithm), and when the model computations (Step 2 of the algorithm) are not too expensive, it may be worthwile to repeat Newton to get better tomographic approximations.

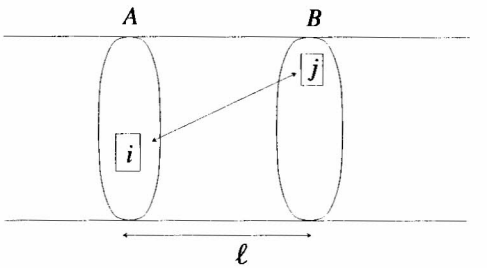

Figure 1.3: The experimental set-up consisting of a pipe and two tomographic sections

1.3 Velocity estimation

This section describes how results from two tomographic sections can be used to determine velocities of the fluids inside a pipe. In Subsection 1.3.1 the correlations are defined. Subsection 1.3.2 explains the estimation of the fluids' velocities using these correlations.

1.3.1 Correlations

The measurement device to be developed consists of two tomography units at a distance ℓ along the tube. Each unit yields an estimate for the conductivity profile over a cross section of the tube. The situation is sketched in Figure 1.3. The measurements and the analysis involved are performed at equidistant times $t_k = t_0 + k \cdot \Delta t$ for some time interval Δt. We assume that Δt is small with respect to the timescale of fluctuations in the conductivity σ. Although estimated conductivity profiles are available at discrete time points t_j, reliable estimates for the profiles at intermediate times can be obtained by interpolating if Δt is small enough. So, in the following we consider these profiles to be available for all times t. We will discern between the data obtained from the two tomography units by superscripts A and B.

As estimates for the expansion of the conductivity profiles $\sigma^A(t, \mathbf{x})$ and $\sigma^B(t, \mathbf{x})$ we use (1.16) with basis functions (1.19). This choice is not necessary, but it is made for clarity. For other choices of base functions essentially the same can be done. The interpretation however, is not immediately clear then. The measurements thus yield per grid cell i the averaged conductivities $\sigma_i^A(t)$ and $\sigma_i^B(t)$, $i = 1, \ldots, N$. We scale the conductivity such that $0 \leq \sigma_i \leq 1$, with $\sigma = 0$ corresponding to only water in cell i and $\sigma = 1$ to only oil. Intermediate values of σ_i then represent the ratio of the volume fractions of water and oil in cell i. The basic idea of the present approach is to estimate the velocities of water and oil from correlating the time series $\sigma_i^A(t)$ and $\sigma_j^B(t)$. If there is a strong correlation between $\sigma_i^A(t)$ and $\sigma_j^B(t + \tau_{ij})$ for some time interval τ_{ij}, it may be concluded that most of the flux through cell i of cross section A has arrived at cell j of cross section B after a time interval τ_{ij}. Since the distance

between these cells is known, this leads to an estimate for the velocity of this part of the volume flux. To calculate such a correlation reliably we have to average over a time window, say T. This interval should be such that the direction and intensity of the average velocity in cell i of cross section A does not vary much. On the other hand, T should be such long that the composition of this flux changes over this time interval, since otherwise no correlation at all can be detected. As for the notation we use the following conventions.

For the average:

$$\langle \sigma_i^A(t) \rangle = \frac{1}{T} \int_{t-T}^{t} \sigma_i^A(t') \, \mathrm{d}t' \tag{1.25}$$

and for the variance

$$\mathrm{var}(\sigma_i^A(t)) = \langle (\sigma_i^A(t))^2 \rangle - \langle \sigma_i^A(t) \rangle^2 \,, \tag{1.26}$$

with similar notation for σ^B. The correlation c_{ij} between $\sigma_i^A(t)$ and $\sigma_j^B(t+\tau)$ is defined as

$$c_{ij}(t,\tau) = \frac{\langle \, (\sigma_i^A(t) - \langle \sigma_i^A(t) \rangle) \, (\sigma_j^B(t+\tau) - \langle \sigma_j^B(t+\tau) \rangle) \, \rangle}{\sqrt{\mathrm{var}(\sigma_i^A(t)) \, \mathrm{var}(\sigma_j^B(t+\tau))}} \,. \tag{1.27}$$

Since both i and j run over N values, the number of correlations to be calculated is given by $N^2 = L^2(L-1)^2/4$. This number thus scales with L^4, not with L^8 as was thought for a long time; see also Section 1.4.

From the definition of correlation it follows that $-1 \leq c_{ij} \leq 1$. If the correlation between two signals equals one, this indicates that they vary in an identical way around their successive averages. The levels of the averages themselves are not relevant and there is also some freedom in the amplitudes of the fluctuations: two signals that fluctuate similarly apart from a multiplicative positive factor also yield a correlation equal to one. If the correlation equals zero, this indicates that the signals have nothing in common. They probably stem from completely independent sources. A correlation equal to minus one indicates that the fluctuations of the signals are similar apart from a negative multiplicative factor.

The correlations c_{ij} in (1.27) are expected to be non-negative in view of the physics of the system. If most of the flux through cell i of cross section A hits cell j of cross section B, c_{ij} will be close to one. If nothing of the flux through cell i hits cell j, c_{ij} will be close to zero. If a negative value for c_{ij} is calculated, the corresponding correlation must be spurious and then it is better to correct this by setting c_{ij} equal to zero.

1.3.2 Velocity estimation

In equation (1.27) we defined the delay τ to take into account that it takes some time for the flow to travel from cross section A to cross section B. This time will depend

on the local velocity at cell i of cross section A. The best estimate for τ, for mass transport from cell i to cell j, follows from

$$\tau_{ij}(t) = \text{value of } \tau \text{ for which } c_{ij}(t, \tau) \text{ is maximal.} \tag{1.28}$$

In practice it is neither possible nor necessary to scan all positive τ values, since the maximum speed of the flow is limited. If no fraction of the flux through cell i hits cell j the value of $c_{ij}(t, \tau)$ will in theory vanish for all τ, and τ_{ij} is undefined. In practice, the correlation will be very small in this case. Then, it suffices to set the corresponding velocity v_{ij}, defined underneath in (1.33), at zero, instead of applying (1.33) itself. Moreover note that, for determining $\tau_{ij}(t)$ from (1.28), it is possible to use

$$\tilde{c}_{ij}(t, \tau) = \frac{1}{T} \int_{t-T}^{t} \sigma_i^A(t') \, \sigma_j^B(t' + \tau) \, \mathrm{d}t' \tag{1.29}$$

instead of c_{ij}, assuming that the flow regime does not change considerably on the timescale related to τ, i.e. that $\langle \sigma_j^B(t + \tau) \rangle \approx \langle \sigma_j^B(t) \rangle$ and $\mathrm{var}(\sigma_j^B(t + \tau)) \approx \mathrm{var}(\sigma_j^B(t))$.

The flow is assumed to be such that all material passing cross section A passes cross section B after a short while. If we would take the correlation $c_{ij}(t, \tau_{ij})$ with τ_{ij} determined by (1.28) as a measure for the fraction of the flux through cell i hitting cell j, this would imply that

$$\sum_{j=1}^{N} c_{ij} = 1 \, . \tag{1.30}$$

In practice this condition will never be met, but one will find

$$\sum_{j=1}^{N} c_{ij} = d_i \neq 1 \, . \tag{1.31}$$

In order to correct for possible inaccuracies of the method, we normalize the c_{ij} by

$$\bar{c}_{ij} = \frac{c_{ij}}{d_i} \, , \tag{1.32}$$

so that (1.30) is satisfied for the $\bar{c}_{ij}$.

From τ_{ij} we can find an estimate for the velocity of the fraction of the flux through cell i hitting cell j. We are interested in the component v_{ij} parallel to the pipe axis. This component is estimated by

$$v_{ij}(t) = \frac{\ell}{\tau_{ij}(t)} \, . \tag{1.33}$$

The average axial velocity v_i of the flux through cell i is obtained by taking into account the velocities of all its fractions. This leads to

$$v_i(t) = \sum_{j=1}^{N} \overline{c}_{ij}(t)\, v_{ij}(t)\ . \tag{1.34}$$

The average velocity of the oil through cross section A in the time window $(t-T, t)$ is then given by

$$v_{\text{oil}}(t) = \sum_{i=1}^{N} \langle \sigma_i^A(t) \rangle\, v_i(t). \tag{1.35}$$

For the water component we find

$$v_{\text{water}}(t) = \sum_{i=1}^{N} (1 - \langle \sigma_i^A(t) \rangle)\, v_i(t)\ . \tag{1.36}$$

Note that the formulae provided here can not be found explicitly in the literature (e.g., cf. [3]).

1.4 Scaling of the number of correlations

As was already mentioned above, the main problem in applying tomography methods to velocity estimation was believed to be the number of correlations that have to be calculated. This number was thought to scale with L^8. The reasoning for this stems from the popularity of the linear backprojection method, explained in the last paragraph of Section 1.2.2. It calculates the values of the conductivity σ on the mesh arising from intersecting so-called "sensing areas". Because there are $N := L(L-1)/2$ sensing areas, the number of mesh cells is $N(N-1)/2$, which is of order L^4. Because correlations have to be calculated between all mesh cell in tomography unit A and all mesh cells in B, the number of correlations is the square of the number of mesh cells, hence of order L^8.

However, as was discussed in the previous sections, the number of degrees of freedom plays an important role. Because the number of independent measurements per tomography unit is $L(L-1)/2$, it is not possible to recover more than $L(L-1)/2$ independent values of σ. Apparently, many of the values of σ at the mesh cells calculated by the linear backprojection method are interdependent. For the Noser algorithm discussed in Section 1.2.5, it is very clear that no more than $L(L-1)/2$ independent components of σ are found. This means that the number of independent correlations that have to be calculated is no more $L^2(L-1)^2/4$, which is of order L^4. This is a dramatic reduction in comparison to the L^8-scaling. In fact, it renders the problem tractable in real time.

1.5 Conclusions and discussion

In this paper we considered estimating velocities in multiphase flow using tomography methods. In its most general setting, the velocity estimation problem is very hard and possibly fundamentally unsolvable. For example, if the flow reverses its direction in the vicinity of the measurement device, the estimation procedure outlined in this paper, fails. However, in practice the situaton will not be that bad; it is to be expected that under standard conditions the estimation yields reliable results.

The idea of applying tomography methods in this context is not new, but the approach was esteemed to be too time-consuming. We were asked to reduce the number of necessary calculations. This number is practically determined by the number of correlations to be evaluated between signals from different regions of two neighbouring cross sections. This number was thought to scale with L^8, where L is the number of electrodes in the tomographic device. It was known how to reduce this to a L^4-scaling in case of using linear reconstruction methods, such as the linear backprojection method. However, this method is too crude to acquire the needed accuracy. We were asked to make the reduction work also for other, nonlinear reconstruction methods. In fact, Appendix 1.6 reflects our work on this problem. However, we reached the far more interesting conclusion that the number of essentially different correlations always scales with L^4, and that the L^8-scaling is based on a wrong insight.

Effectively, this amounts to noting that there must be very many dependencies between the $O(L^8)$ correlations that were thought to be required. Using the linear backprojection method, one implicitly introduces many more (dependent) correlations. By using the Noser algorithm (Section 1.2.5) these dependent correlations are eliminated in a natural way, leading to $O(L^4)$ correlations. This solves the problem that was posed to us, because $O(L^4)$ is the best we can get without losing information.

Although the present findings make the the application of tomography methods a real option in oil well management, several problems still have to be overcome if the method is to be implemented in practice. For example, the calculations depend quite strongly on separation of timescales. The flowregime should not change much on the timescale T (see Section 1.3.1) over which the correlations are calculated. Nevertheless this timescale T must be large compared to the timelag τ to travel from cross section A to B. Several of such conflicting assumptions are to be made. Secondly, this method will not only measure the velocity of the fluid. Correlations caused by travelling and standing waves will also be found. The wave speed of travelling waves can not be distinguished from the velocity of the fluid itself by this method.

These and other considerations make clear that further research is still necessary in order to develop a robust tool for velocity estimation in two-phase flow based on tomography techniques.

Acknowledgement

The authors kindly acknowledge the fruitful and pleasant cooperation with Alex van der Spek from Shell during and after the study group.

1.6 Appendix: Efficient evaluation of correlations

This appendix deals with a method by which the number of correlations could be reduced if the reasoning leading to the scaling L^8 were true. The idea is to express the correlations between timeseries of σ in terms of timeseries of the measured voltages V_l^k. Because the number of independent measurements is $L(L-1)/2$ per tomography unit, whereas the number of timeseries of σ would be of order L^4, this would amount to the required reduction. However, according to the reasoning we adhere to, the number of independent measurements is the same as the number of independent timeseries of σ. Hence, no reduction is obtained, nor is it needed anymore. Nevertheless, this appendix is included, if only because we have devoted so much time to it. Here a method is proposed to calculate the correlations between the timeseries of σ at the mesh cells indirectly. The idea is that these correlations between timeseries of σ at sections A and B can be found from the correlations between the measured voltages at the electrodes. This could overcome the problem that the number of correlations that have to be calculated is too large (of order L^8) for the available computer power. Although in Section 1.4 the scaling L^8 is disputed, we still think this method is interesting by itself. It can be demonstrated most clearly for linear reconstruction methods. In this appendix the reasoning leading to the scaling L^8 is assumed to be correct. Note that this is important only when commenting on the number of necessary calculations.

Suppose that we use a linear reconstruction map, for example

$$\sigma^{(1)} = \sigma^{(0)} - M \cdot (V - V^{(0)}).$$

As in Section 1.2.5 the σ-values at the mesh cells are represented as components of an N-dimensional vector. The matrix M has dimensions $N \times L$. Moreover, $\sigma^{(0)}$ is some known distribution, e.g. the uniform distribution. Indeed one could define all parameters relative to this reference state, in which case the equation would read

$$\sigma = -M \cdot V. \tag{1.37}$$

We are interested in the correlations, see equation (1.29). Due to equation (1.37), i.e. $\sigma_i = -\sum_m M_{im} V_m$ (note that this is an abuse of notation: V_m denotes V_l^k and m indexes all $L(L-1)/2$ independent measurements), it can be written as

$$\begin{aligned} \tilde{c}_{ij}(t,\tau) &= \tfrac{1}{T}\int_{t-T}^{t} \sigma_i^A(t')\,\sigma_j^B(t'+\tau)\,\mathrm{d}t' \\ &= \textstyle\sum_{m,n} M_{im} M_{jn}\, \tfrac{1}{T}\int_{t-T}^{t} V_m^A(t')\,V_n^B(t'+\tau)\,\mathrm{d}t'. \end{aligned}$$

Note that at each cross section the same (precomputed) matrix M is used. This is justified by the fact that the shape of the cross sections is the same (usually circular). So it is possible to calculate the correlations for the measurements V only, and from this we can find the correlations between the values of σ in the cells by a simple linear expression. Because the number of cells is the square of the number of independent measurements according to "standard reasoning", this amounts to a reduction of computational work that has to be done.

Applying the Noser algorithm, discussed in Section 1.2.5, each step is linear in V. Although the resulting distribution $\sigma^{(K)}$ after some number K of steps, depends nonlinearly on the measured voltages V, this linearity of each single step can be used again to express the correlation for the cells in terms of those for the measurements. In fact, using the definition (1.24), the recursive formula in the algorithm becomes $\sigma_i^{(s+1)} = \sigma_i^{(s)} - \sum_m M_{im}^{(s)}(V_m - V_m^{(s)})$. Therefore the correlations can be updated each step using the following formula:

$$
\begin{aligned}
&\tilde{c}_{ij}^{(s+1)}(t,\tau) \\
&\quad = \frac{1}{T}\int_{t-T}^{t} \sigma_i^{(s+1),A}(t')\,\sigma_j^{(s+1),B}(t'+\tau)\,\mathrm{d}t' \\
&\quad = \frac{1}{T}\int_{t-T}^{t} \sigma_i^{(s),A}(t')\,\sigma_j^{(s),B}(t'+\tau)\,\mathrm{d}t' \\
&\qquad - \sum_m \frac{1}{T}\int_{t-T}^{t} M_{im}^{(s),A}(t')\,(V_m^A(t') - V_m^{(s),A}(t'))\,\sigma_j^{(s),B}(t'+\tau)\,\mathrm{d}t' \\
&\qquad - \sum_n \frac{1}{T}\int_{t-T}^{t} \sigma_i^{(s),A}(t')\,M_{jn}^{(s),B}(t'+\tau)\,(V_n^B(t'+\tau) - V_n^{(s),B}(t'+\tau))\,\mathrm{d}t' \\
&\qquad + \sum_{m,n} \frac{1}{T}\int_{t-T}^{t} M_{im}^{(s),A}(t')\,M_{jn}^{(s),B}(t'+\tau)\,(V_m^A(t') - V_m^{(s),A}(t')) \\
&\qquad\qquad (V_n^B(t'+\tau) - V_n^{(s),B}(t'+\tau))\,\mathrm{d}t'.
\end{aligned}
\tag{1.38}
$$

Note that this formula allows us to update the correlations one step in advance. One does not need to find $\sigma^{(s+1)}$ in order to calculate the correlations $\tilde{c}_{ij}^{(s+1)}$. However, this does not reduce the number of correlations because the matrices $M^{(s)}$ have to be included as they depend on time through $\sigma^{(s)}$ and $V^{(s)}$. Only if a constant matrix M would be used, as was proposed at the end of Section 1.2.5, this formula results in a linear expression for the correlations between cells in terms of the correlations between voltages. Moreover, in that case, all correlations $\tilde{c}_{ij}^{(s)}$ can be expressed in terms of correlations between the measured voltages, because the σ-distributions in the second and third terms in the righthand side of equation (1.38) can be eliminated,

using

$$\sigma_i^{(s)} = \sigma_i^{(0)} - \sum_{r=1}^{s} \sum_m M_{im}(V_m - V_m^{(r-1)})$$

and the fact that $\sigma_i^{(0)}$ can be omitted because it is constant. Of course, this also increases the number of correlations that have to be calculated, but only linearly in the number of steps s. Nevertheless, equation (1.38) reduces the number of correlations that have to be calculated as soon as the matrices $M^{(s)}$ can be kept out of the integrals because at least one of the σ-distributions has been replaced by voltages.

Bibliography

[1] E. Somersalo, M. Cheney and D. Isaacson. *Existence and uniqueness for electrode models for electric current computed tomography*. SIAM J. Appl. Math., 52 (1992), pp. 1023–1040.

[2] M. Cheney, D. Isaacson and J. C. Newell. *Electrical Impedance Tomography*. SIAM Review, Vol. 41 No. I, pp. 85–101.

[3] R. Thorn, S. M. Huang, C. G. Xie, J. A. Salkeld, A. Hunt and M. S. Beck. *Flow imaging for multi-component flow measurement*. Flow. Meas. Instrum., Vol. 1 October 1990, pp. 259–268.