Modern Signal Processing
MSRI Publications
Volume **46**, 2003

Diffuse Tomography as a Source of Challenging Nonlinear Inverse Problems for a General Class of Networks

F. ALBERTO GRÜNBAUM

Abstract. Diffuse tomography refers to the use of probes in the infrared part of the energy spectrum to obtain images of highly scattering media. There are important potential medical applications and a host of difficult mathematical issues in connection with this highly nonlinear inverse problem. Taking into account scattering gives a problem with many more unknowns, as well as pieces of data, than in the simpler linearized situation. The aim of this paper is to show that in some very simplified discrete model, reckoning with scattering gives an inversion problem whose solution can be reduced to that of a *finite* number of linear inversion problems. We see here that at least for the model in question, the proportion of variables that can be solved for is higher in the nonlinear case than in the linear one. We also notice that this gives a highly nontrivial problem in what can be called *network tomography*.

1. Introduction

Optical, or *diffuse*, tomography, refers to the use of low energy probes to obtain images of highly scattering media.

The main motivation for this line of work is, at present, the use of an infrared laser to obtain images of diagnostic value. There is a proposal to use this in neonatal clinics to measure oxygen content in the brains of premature babies as well as in the case of repeated mammography. With the discovery of highly specific markers that respond well in the optical or infrared region there are many potential applications of this emerging area; see [A1; A2].

There are a number of physically reasonable models that have been used in the formulation of the associated direct and inverse problems. These models are based on some approximation to a wave propagation model, such as the so-called *diffusion approximation*, or a transport equation model resulting in some type of linear Boltzmann equation. See [A1; A2; D; NW] for recent surveys of work

The author was supported in part by NSF Grant # FD99711151.

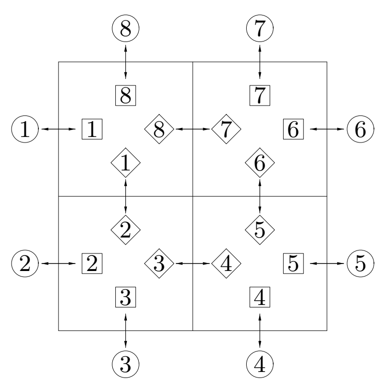
in this area. These papers give a detailed description of the physically relevant formulations that different authors have considered.

Our Markov chain formulation, going back to [G1; GP1; SGKZ], is different from those contained in these papers. We model the evolution of a photon as it moves through tissue by means of a Markov chain. At any (discrete) instant of time a photon occupies one of the states of the chain. These states are meant to represent a discretization of phase space, i.e. they encode position as well as velocity of a photon at a given time. The chain has three kinds of states: incoming states (which are meant to represent source positions surrounding the object of interest), hidden states (which are meant to represent the positions and velocities of photons inside the tissue) and finally, outgoing states(which represent detectors surrounding the object). We should also add an absorbing state at the center of each pixel to indicate that a photon "entering the pixel" can die in it. Instead of adding these extra states we simply do not assume that the sum of the one-step transition probabilities from a state should add to one. The difference between one and this sum is the probability of being absorbed into the pixel in question when coming into it from the corresponding state.

The direct problem would consist of determining different "input-output" quantities once the one-step transition probability matrix of our Markov chain has been given.

The resulting inverse problem amounts to reconstructing the one-step transition probability matrix for our Markov chain (with three kinds of states) from boundary measurements. This model is too simple and too general to faithfully reflect the physics of diffuse tomography but could be of interest in other set-ups. It gives a difficult class of *nonlinear* inverse problems for a certain *general class of networks* with a complex pattern of connections which are motivated by the diffuse tomography picture.

Since our model is the result of a discretization both in the positions occupied by a photon as well as the direction in which it is moving, the states will be indicated below by arrows placed at the boundaries of each pixel and pointing in one of four possible directions. One of the smallest cases of interest in dimension two is this:

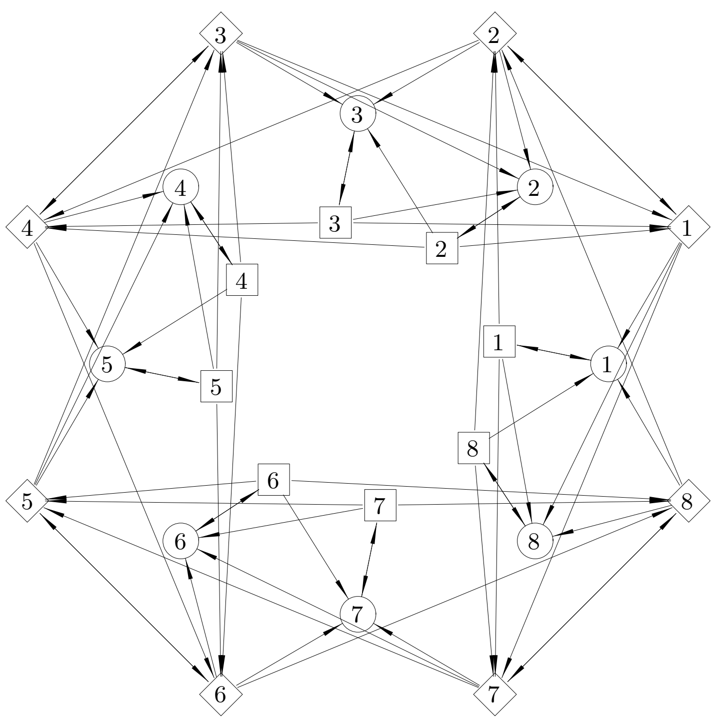


This simple model features four pixels, eight source positions, eight detector positions as well as eight hidden states. In this figure, incoming states are labeled by numbers enclosed in squares, outgoing states are labeled by numbers enclosed in circles, and hidden states are labeled by numbers enclosed in diamonds. The possible one step transitions are indicated in the next section, whereas the figure below displays (by means of arrows, as explained earlier) *only* the eight states of each kind.

In [G4] a discussion can be found of the corresponding smallest case in dimension three, where pixels are replaced by voxels and we have six different directions for our states.

The physics, or what is left of it, is best compressed into a *multiterminal network* where the nodes are the states of our Markov chain and the oriented edges indicate one-step transitions (with unknown probabilities) between the corresponding nodes. This is what a probabilist would call a state diagram.

As an example, here is the network corresponding to the physical model shown on the previous page (for clarity, when two nodes are joined by two opposite edges, we draw a single edge with arrows at both ends):



Notice that there is an underlying linear dynamics governed by the (unknown) one-step transition probability matrix of our Markov chain, but the inversion problem of interest is still nonlinear.

A remarkable feature of this simple model is that, at least for systems arising from very coarse tomographic discretizations, it gives an exactly solvable system of nonlinear equations, i.e., a certain number of unknowns are expressible in terms of the data and a number of free parameters. The advantages of this rather uncommon situation are clear: for instance it is possible to go beyond iterative methods of solution, which are very common for nonlinear problems.

In both the two-dimensional and three-dimensional situations we can consider as data the *photon count for a source-detector pair* which is defined as the probability that a photon that started at the source in question emerges at the detector in question regardless of the number of steps involved. If we assume that every one-step transition takes one unit of time we can consider the *time-of-flight* as a random variable associated to each incoming-outgoing pair. The photon count is the moment of order zero of this collection of random variables.

In Section 2 we see how far one can go using only the moment of order zero of time of flight. Section 3 considers the situation when we also use a small part of the information contained in the first moment of this collection of random variables. Section 4 deals with the issue of dealing with those variables that cannot be solved from the data. Finally Section 5 alludes to the fact that this same machinery can be applied in the non-physical situation when the dimension is neither two nor three but arbitrary.

It is also instructive in each case to consider the *standard tomographic* linear problem when scattering is completely ignored and a photon can only be absorbed in a pixel or continue in its straight-line trajectory. In this case each one of the four pixels, conveniently labeled $(1,1)$, $(1,2)$, $(2,1)$ and $(2,2)$ as the entries of a 2×2 matrix, is characterized by one parameter, its absorption probability.

The results regarding the ratio between the number of variables we can solve for and the total number of unknowns for each one of these scenarios are given below.

In the two-dimensional case, using four pixels (see figure on page 138) there are three situations:

(1) The linear one where scattering is ignored, gives a problem with 4 unknowns and 4 pieces of data, of which only three are independent and allows one to solve for 3 out of 4 unknowns.
(2) The general model discussed above (as in [GP1; GP2]) allows one to solve for 48 out of a total of 64 unknowns, leaving the ratio of $\frac{3}{4}$ unchanged.
(3) The use of time-of-flight information, which is discussed in Section 4, as well as in [G3], [GM1] gives a slightly better ratio, namely $\frac{56}{64} = \frac{7}{8}$.

When this comparison is done in dimension three, with a total of eight voxels, we get three situations:

(1) The linear version of the problem (scattering being ruled out) gives a system of 12 equations in 8 unknowns which can be solved for 7 of them in terms of one arbitrary parameter, giving a ratio of $\frac{7}{8}$.

(2) The general model (discussed in Sections 2 and 3) yields a system of 576 nonlinear equations in 288 variables that can be solved for 240 of them, with a ratio of $\frac{240}{288} = \frac{5}{6}$.
(3) The use of time-of-flight information (discussed in Section 4) raises the ratio to $\frac{264}{288} = \frac{11}{12}$. This shows that the consideration of a fully nonlinear problem can (in some sense) lead to a better determined problem than the corresponding linearized one.

We do not consider here the important issues of the difficulty in solving these systems or the sensitivity to errors of the corresponding problem.

For a very nice and up-to-date discussion of work in this area one can see [A1], [A2], [D], [NW]. These papers give a detailed description of the physically relevant formulations that different authors have considered. For an early reference in the area of *network tomography* see [V]. For similar problems in an area of great practical interest see the recent article [CHNY].

Remark This is an appropriate place to mention an oversight in [G4]. The labeling of the states given in the introduction to that paper does not correspond to the one used in [G4, Section 3]. The labeling used in the introduction to [G4] represents an improvement over the one used in [G4, Section 3]. The results in [G4] are correct, but some of the inversion formulas are unduly complicated since they are written down using a more complicated labeling scheme. When we use the labeling given in the introduction to [G4] we can reduce the entire problem to a set of equivalent linear ones, obviating the last nonlinear step in [G4]. This is reported in [GM2].

2. General Framework and Some Results

The one-step transition probability matrix P is naturally broken up into blocks that connect different types of states. We denote by P_{IO} the block dealing with a one-step transition from an arbitrary incoming state to an arbitrary outgoing state. P_{HH} denotes the corresponding block connecting hidden to hidden states, P_{IH} the one connecting incoming to hidden states and finally P_{HO} accounts for one-step transitions between hidden and outgoing states. For completeness we give these matrices below.

$$P_{HH} = \begin{pmatrix} 0 & \mathrm{N11S} & 0 & 0 & 0 & 0 & \mathrm{N11E} & 0 \\ \mathrm{S21N} & 0 & 0 & \mathrm{S21E} & 0 & 0 & 0 & 0 \\ \mathrm{W21N} & 0 & 0 & \mathrm{W21E} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathrm{E22W} & 0 & 0 & \mathrm{E22N} & 0 & 0 \\ 0 & 0 & \mathrm{S22W} & 0 & 0 & \mathrm{S22N} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathrm{N12S} & 0 & 0 & \mathrm{N12W} \\ 0 & 0 & 0 & 0 & \mathrm{E12S} & 0 & 0 & \mathrm{E12W} \\ 0 & \mathrm{W11S} & 0 & 0 & 0 & 0 & \mathrm{W11E} & 0 \end{pmatrix};$$

$$P_{HO} = \begin{pmatrix} \text{N11W} & 0 & 0 & 0 & 0 & 0 & 0 & \text{N11N} \\ 0 & \text{S21W} & \text{S21S} & 0 & 0 & 0 & 0 & 0 \\ 0 & \text{W21W} & \text{W21S} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \text{E22S} & \text{E22E} & 0 & 0 & 0 \\ 0 & 0 & 0 & \text{S22S} & \text{S22E} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \text{N12E} & \text{N12N} & 0 \\ 0 & 0 & 0 & 0 & 0 & \text{E12E} & \text{E12N} & 0 \\ \text{W11W} & 0 & 0 & 0 & 0 & 0 & 0 & \text{W11N} \end{pmatrix};$$

$$P_{IH} = \begin{pmatrix} 0 & \text{E11S} & 0 & 0 & 0 & 0 & \text{E11E} & 0 \\ \text{E21N} & 0 & 0 & \text{E21E} & 0 & 0 & 0 & 0 \\ \text{N21N} & 0 & 0 & \text{N21E} & 0 & 0 & 0 & 0 \\ 0 & 0 & \text{N22W} & 0 & 0 & \text{N22N} & 0 & 0 \\ 0 & 0 & \text{W22W} & 0 & 0 & \text{W22N} & 0 & 0 \\ 0 & 0 & 0 & 0 & \text{W12S} & 0 & 0 & \text{W12W} \\ 0 & 0 & 0 & 0 & \text{S12S} & 0 & 0 & \text{S12W} \\ 0 & \text{S11S} & 0 & 0 & 0 & 0 & \text{S11E} & 0 \end{pmatrix};$$

$$P_{IO} = \begin{pmatrix} \text{E11W} & 0 & 0 & 0 & 0 & 0 & 0 & \text{E11N} \\ 0 & \text{E21W} & \text{E21S} & 0 & 0 & 0 & 0 & 0 \\ 0 & \text{N21W} & \text{N21S} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \text{N22S} & \text{N22E} & 0 & 0 & 0 \\ 0 & 0 & 0 & \text{W22S} & \text{W22E} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \text{W12E} & \text{W12N} & 0 \\ 0 & 0 & 0 & 0 & 0 & \text{S12E} & \text{S12N} & 0 \\ \text{S11W} & 0 & 0 & 0 & 0 & 0 & 0 & \text{S11N} \end{pmatrix}.$$

The choice of names for the variables in P is meant to indicate the corresponding transitions, for instance N11S means that we enter pixel $(1, 1)$ going north and exit it going south. It is convenient to refer to the figure on page 138 at this point.

Just as in [GP1], [GP2] we find it convenient to introduce matrices A, X, Y, W by means of

$$A = P_{HO}^{-1},$$
$$P_{IO} = XA^{-1},\ P_{HH} = A^{-1}W,\ P_{IH} = XA^{-1}W - Y.$$

The transformation, for a given P_{HO}, from the matrices P_{HH}, P_{IO}, P_{IH} to the matrices W, X, Y was introduced by S. Patch in [P3]. Notice that from A, X, W and Y it is possible to recover (in that order) the matrices P_{HO}, P_{IO}, P_{HH} and, finally, P_{IH}.

One advantage of introducing these matrices is that the input-output relation

$$Q_{IO} = P_{IO} + P_{IH}(I - P_{HH})^{-1}P_{HO}$$

can be rewritten, by multiplying both sides first by A on the right and then by $(I - A^{-1}W)$ on the right again, in the form

$$Q_{IO}(A - W) = X - Y.$$

In [GP1], [GP2] we exploited the block structure of the matrices A, W, X, Y to show that once Q_{IO} is given then A is arbitrary. After choosing A, it is then possible to derive explicit formulas for X, Y and W.

In the three-dimensional case the situation is a bit better, although the equations that we have to handle are naturally harder to deal with. We find that the matrix A can no longer be picked arbitrarily but only 2/3 of it is arbitrary. This means that using *photon count* alone it is possible to express 24 of the 72 entries in the matrix A in terms of the data and 48 free parameters in A. By the *photon count* matrix we refer to the matrix whose entries are given by the probabilities that a photon that starts at a given source position would emerge from the tissue at a specified detector position. For details consult [G4] and [GM2].

3. Using the First Moment of Time-of-Flight

Now we go beyond the photon count and consider the first moment of the time-of- flight. As observed in the introduction the moment of order zero of this collection of random variables (one for each source-detector pair) gives the photon count matrix Q_{IO}.

If we denote the expression

$$P_{IH}(I - P_{HH})^{-2}P_{HO}$$

by R, we have:

LEMMA. The first moment of the "time-of-flight" can be expressed as

$$Q_{IO} + R.$$

PROOF. Start from the observation that the j-th moment of the time of flight is given by

$$Q_{IO}^{(j)} = P_{IO} + \sum_{k=0}^{\infty} P_{IH}P_{HH}^{k}P_{HO}(k+2)^{j}. \tag{3–1}$$

In particular, if $j = 0$ we recover (after an appropriate summation of the corresponding geometric series) the expression for $Q_{IO} \equiv Q_{IO}^{(0)}$ given in Section 2. We will return to this expression later in this section.

For $j = 1$ we get

$$\begin{aligned} Q_{IO}^{(1)} &= P_{IO} + 2P_{IH}(I - P_{HH})^{-1}P_{HO} + P_{IH}P_{HH}(I - P_{HH})^{-2}P_{HO} \\ &= Q_{IO}^{(0)} + P_{IH}(I - P_{HH})^{-2}[I - P_{HH} + P_{HH}]P_{HO} \\ &= Q_{IO}^{(0)} + R. \end{aligned}$$

□

Since Q_{IO} is taken as data we can consider R as the extra information provided by the expected value of time of flight.

Observe now that we have the relation

$$Q_{IO}A - X(A) = R(A - W(A)).$$

This follows, for instance, by noticing that each side of this identity is given by $P_{IH}(I - P_{HH})^{-1}$.

In the two-dimensional case ([GP2; GM1]) this concludes the job since we can use some of the entries of the matrix R to determine the ratios among eight pairs of the entries in A. Explicit formulas are given in [GM1].

The three-dimensional case has been given a first treatment in [G4]. By using the labeling mentioned in the introduction to that paper it is possible to obtain explicit formulas similar to those mentioned above. For details see [GM2].

It is very important to notice that in *either dimension* the entire problem of determining the blocks in P admits a natural "gauge transformation" given exactly by a diagonal matrix D. Consider the transformation that goes from a given set of blocks, to a new one given by the relations

$$\begin{aligned}\tilde{P}_{IO} &= P_{IO},\\ \tilde{P}_{IH} &= P_{IH}D^{-1},\\ \tilde{P}_{HH} &= DP_{HH}D^{-1},\\ \tilde{P}_{HO} &= DP_{HO}.\end{aligned}$$

Notice that this gauge transformation preserves the required block structure of all the matrices in question. Moreover the probability of going from an arbitrary incoming state to an arbitrary outgoing state in m steps, given by the matrix P_{IO} if $m = 1$ and by $P_{IH}P_{HH}^{m-2}P_{HO}$ if $m \geq 2$, is clearly invariant under the transformation mentioned above. It follows then by referring to (3–1) for the j-th moment of the time of flight distribution that this is not affected by this gauge.

In conclusion, we have shown that the zeroth and first moments of the time-of-flight distribution determine the matrix P up to the choice of the arbitrary diagonal matrix D introduced above.

4. Taking into Account a Physical Model

An important question remains: how should the values of the 24 free parameters be picked (or the 8 free parameters in dimension two)? A similar question was discussed in [GP2] where we considered the effect of imposing on our very general model the assumption of "microscopic reversibility", i.e., a one-step transition from a state (of our Markov chain) given by the vector $\mathbf{v}$ to a state given by the vector $\mathbf{w}$ has the same probability as a transition from the sates given by the vectors $-\mathbf{w}$ and $-\mathbf{v}$ respectively. On the other hand, in [G2], [GZ] we considered the case of isotropic scattering. Each one of these cases leads to a dramatic reduction in the number of free parameters.

It is tempting to make some of these simplifying assumptions at the very *beginning* of the process, thereby reducing the number of unknowns. Experience seems to indicate that the possibility of reducing the *already nonlinear* system of equations to a linear one is greatly enhanced by making use of these assumptions at the end of the process.

5. A Network Tomography Problem for the Hypercube

The two-dimensional and three-dimensional problems discussed above have a firm foundation in diffuse tomography. It is however possible to go to higher dimensions and consider the corresponding *d- dimensional hypercube* and the network that goes along with it. By using the techniques in [GM1] and [GM2] it is possible to see that by measuring the first two moments (zeroth and first) of time-of-flight we can determine everything explicitly up to a total of $d\,2^d$ free parameters. This happens to be the dimension of the gauge that appears at the end of Section 3, and thus this result is optimal. Details will appear in [GM3].

Acknowledgments. We thank the editors for useful suggestions on ways to improve the presentation.

References

[A1] S. Arridge, "Optical tomography in medical imaging", *Inverse Problems* **15** (1999), R41–R93.

[A2] S. Arridge and J. C. Hebden, "Optical imaging in medicine, II: Modelling and reconstruction", *Phys. Med. Biol.* **42** (1997), 841–853.

[D] O. Dorn, "A transport-backtransport method for optical tomography", *Inverse Problems* **14** (1998), 1107–1130.

[G1] F. A. Grünbaum, "Tomography with diffusion", pp. 16–21 in *Inverse Problems in Action*, edited by P. C. Sabatier, Springer, Berlin.

[G2] F. A. Grünbaum, "Diffuse tomography: the isotropic case", *Inverse Problems* **8** (1992), 409–419.

[G3] F. A. Grünbaum, "Diffuse tomography: using time-of-flight information in a two-dimensional model", *Int. J. Imaging Technology* **11** (2001), 283–286.

[G4] F. A. Grünbaum, "A nonlinear inverse problem inspired by three-dimensional diffuse tomography", *Inverse Problems* **17** (2001), 1907–1922.

[GM1] F. A. Grünbaum and L. Matusevich, "Explicit inversion formulas for a model in diffuse tomography", *Adv. Appl. Math.* **29** (2002), 172–183.

[GM2] F. A. Grünbaum and L. Matusevich, "A nonlinear inverse problem inspired by 3-dimensional diffuse tomography", *Int. J. Imaging Technology* **12** (2002), 198–203.

[GM3] F. A. Grünbaum and L. Matusevich, "A network tomography problem related to the hypercube", in preparation.

[GP1] F. A. Grünbaum and S. Patch, “The use of Grassmann identities for inversion of a general model in diffuse tomography”, in *Proceedings of the Lapland Conference on Inverse Problems*, Saariselka, Finland, June 1992.

[GP2] F. A. Grünbaum and S. Patch, “Simplification of a general model in diffuse tomography”, pp. 744–754 in *Inverse problems in scattering and imaging*, edited by M. A. Fiddy, Proc. SPIE **176**, 1992.

[GP3] F. A. Grünbaum and S. Patch, “How many parameters can one solve for in diffuse tomography?”, in *Proceedings of the IMA Workshop on Inverse Problems in Waves and Scattering*, March 1995.

[GZ] F. A. Grünbaum and J. Zubelli, “Diffuse tomography: computational aspects of the isotropic case”, *Inverse Problems* **8** (1992), 421–433.

[NW] F. Natterer and F. Wubbeling, *Mathematical methods in image reconstruction*, SIAM Monographs on Mathematical Modeling and Computation, SIAM, Philadelphia, 2001.

[P1] S. Patch, “Recursive recovery of a family of Markov transition probabilities from boundary value data”, *J. Math. Phys.* **36**:7 (July 1995), 3395–3412.

[P2] S. Patch, “A recursive algorithm for diffuse planar tomography”, Chapter 20 in *Discrete Tomography: Foundations, Algorithms, and Applications*, edited by G. Herman and A. Kuba, Birkhäuser, Boston, 1999.

[P3] S. Patch, “Recursive recovery of Markov transition probabilities from boundary value data”, Ph.D. thesis, UC Berkeley, 1994.

[SGKZ] J. Singer, F. A. Grünbaum, P. Kohn and J. Zubelli, “Image reconstruction of the interior of bodies that diffuse radiation”, *Science* **248** (1990), 990–993.

[V] J. Vardi, “Network tomography: estimating source-destination traffic intensities from link data”, *J. Amer. Stat. Assoc.*, **91** (1996), 365–377.

[CHNY] M. Coates, A. Hero, R. Nowak and B. Yu, “Internet tomography”, *Signal Processing Magazine*, **19**:3 (2002), 47–65.

F. Alberto Grünbaum
Department of Mathematics
University of California
Berkeley, CA 94720
United States
grunbaum@math.berkeley.edu