

On Tomography, Measurement Functionals, a Shallow-Water Model, and Representers; Or, What I Did Last Summer

Brian D. Dushaw
Applied Physics Laboratory,
University of Washington,
Seattle, Washington
dushaw@apl.washington.edu

I. INTRODUCTION

During the Summer 1999, I attended Andrew Bennett's summer school for "Inverse Methods and Data Assimilation." During this two-week school, I took it upon myself to calculate representer for tomography and the shallow-water model that was used to do the numerical exercises. These, then, are a summary of my notes in making this calculation. For this discussion, I have copied wholesale from Andrew's notes, without remorse. In addition, these are merely notes, so that they contain information conveyed by others and for which I cannot claim credit, e.g. Andrew's comments pertaining to Raleigh-Ritz variation method and Gary Egbert's comment on why a measurement of current does not necessarily give one a constraint on sea-surface elevation.

In his 1985 paper "Array Design by Inverse Methods", A. "Mr. Cryptic" Bennett calculated representer for reciprocal shooting tomography and a tidal model as part of a study on array design. Alas, this paper was not heuristically helpful, at least from the Dushaw point of view...these notes hopefully make the details of the issue more explicit (all in good humour here!!).

For much of the Summer School, a simple 1-D forward model was used, while the exercises used a linear shallow-water model. These notes begin with a review of the 1-D model in order to introduce language, notation and a description of measurement functionals, and then these notes will describe the adjoint equation for the linear shallow-water model. Representer are calculated using these adjoint equations. All of these sundry equations are reviewed in order for this document to be reasonably self-contained.

II. A SIMPLE INVERSE MODEL - NOTATION AND LANGUAGE

Let $u_F = u_F(x, t)$ be the solution of the forward problem:

$$\frac{\partial u_F}{\partial t} + c \frac{\partial u_F}{\partial x} = F + f \quad (1a)$$

for $0 \leq x \leq L$ and $0 \leq t \leq T$, with

$$u_F(x, 0) = I(x) + i(x) \quad (1b)$$

for $0 \leq x \leq L$, and

$$u_F(0, t) = B(t) + b(t) \quad (1c)$$

for $0 \leq t \leq T$, where f, i, b are errors in forcing, initial condition, and boundary conditions, respectively.

If a measurement m results in a datum d_m at a point of u , then the Euler-Lagrange equations for the local extrema \hat{u} of the penalty functional $J[u]$ (minimizing misfit with the data and errors in forcing, boundary and initial conditions with weights W) are

$$(backward) \left[\begin{array}{l} -\frac{\partial \lambda}{\partial t} - c \frac{\partial \lambda}{\partial x} = -w \sum_{m=1}^M (\hat{u}_m - d_m) \delta(x - x_m) \delta(t - t_m) \\ \lambda(x, T) = 0 \\ \lambda(L, t) = 0 \end{array} \right. \quad (2a)$$

$$(forward) \left[\begin{array}{l} \frac{\partial \hat{u}}{\partial t} + c \frac{\partial \hat{u}}{\partial x} = F + W_f^{-1} \lambda(x, t) \\ \hat{u}(x, 0) = I(x) + W_i^{-1} \lambda(x, 0) \\ \hat{u}(0, t) = B(t) + W_b^{-1} \lambda(0, t) \end{array} \right. \quad (2b)$$

Equation (2a) is the adjoint equation; the adjoint variable $\lambda \equiv W_f \left(\frac{\partial \hat{u}}{\partial t} + c \frac{\partial \hat{u}}{\partial x} - F \right)$

(defined in the course of minimizing $J[u]$.) The change in sign on the left hand side of (2a) from (1a) originates from an integration by parts in the process of minimizing $J[u]$. The best estimates for f, i , and b are

$$\hat{f}(x, t) \equiv W_f^{-1} \lambda(x, t), \quad \hat{i}(x) \equiv W_i^{-1} \lambda(x, 0), \quad \hat{b}(t) \equiv cW_b^{-1} \lambda(0, t). \quad (2c)$$

These equations may be solved for \hat{u} using representer functions. There are M of them, denoted $r_m(x, t)$, $1 \leq m \leq M$. Each representer has an "adjoint" $\alpha_m(x, t)$, satisfying

$$(backward\ m) \left[\begin{array}{l} -\frac{\partial \alpha_m}{\partial t} - c \frac{\partial \alpha_m}{\partial x} = \delta(x - x_m) \delta(t - t_m) \\ \alpha_m(x, T) = 0 \\ \alpha_m(L, t) = 0 \end{array} \right. \quad (3a)$$

This "adjoint" equation may be integrated backwards in time and space; it picks up a "bare" delta function "impulse" at the time and place of the m th measurement. The representers, $r_m(x, t)$ are found by integrating

$$(forward\ m) \left[\begin{array}{l} \frac{\partial r_m}{\partial t} + c \frac{\partial r_m}{\partial x} = W_f^{-1} \alpha_m(x, t) \\ r_m(x, 0) = W_i^{-1} \alpha_m(x, 0) \\ r_m(0, t) = cW_b^{-1} \alpha_m(0, t) \end{array} \right. \quad (3b)$$

forward in time and space, with forcing, initial and boundary conditions given by the α_m . Once the r_m are found, the solution \hat{u} is in the form

$$\hat{u}(x, t) = u_F(x, t) + \sum_{m=1}^M \beta_m r_m(x, t) \quad (4)$$

for constant coefficients β_m . Thus, the solution to the inverse problem the sum of the forward model run with the prescribed forcing and a linear superposition of the representers. Note that the representer is a function in space and time, and if the forward problem involved multiple fields such as sea surface elevation and currents, then THE representer would be the SET of functions corresponding to sea

surface elevation and currents [three days of labor before this dawned on me...]. The coefficients β_m are found by solution of the equation

$$(\mathbf{R} + w^{-1}\mathbf{I})\hat{\beta} = \mathbf{d} - \mathbf{u}_F \quad (5)$$

where the l^{th} column of the $M \times M$ representer matrix \mathbf{R} consists of the M measured values of the l^{th} representer function $r_l(x, t)$. With the solution of (5) for $\hat{\beta}$, the solution for \hat{u} may be found by straightforward integration of a few differential equations.

The representers have all the earmarks of Green's functions in that they are calculated using delta function impulses, and then a linear superposition of them gives the desired "best" solution. This linear superposition is the solution for the perturbation around the initial "best guess" field u_F ; u_F is also known to tomographers as the "reference" ocean. In the general case, the "reference" ocean is clearly time dependent, so generally and technically a new set of rays would have to be calculated at each time step.

This discussion has used only scalar, or diagonal, weights W , i.e. "white" covariances. The results may be (and should be) generalized to physically meaningful covariances so that the solution \hat{u} is smoothed in a physically meaningful way. The use of covariances is the answer to how to obtain better eigenfunction solutions in the Raleigh-Ritz variational problem for the eigenvalues of a quantum mechanical problem (e.g., Baym, 1969). With enforced covariances, the eigenfunction solutions may be a little better than the junky results that sometimes typify the "bare" Raleigh-Ritz method.

III. MEASUREMENT FUNCTIONALS REVIEWED

The above formalism pertains to point measurements, hence the delta functions in (2a) and (3a). However, as I have told the world until I was blue in the face, tomography is a line-integral measurement. This section generalizes the point measurements to an arbitrary measurements functional, and attempts to provide some meaning to this functional.

An arbitrary measurement functional $\Lambda_{m(y,s)}[u]$ of a field $u(y, s)$ is

$$\Lambda_{m(y,s)}[u] = \int dy \int ds K(y, s; x, t) u(y, s) \quad (6)$$

where $K(y, s; x, t)$ is the "kernel" of the measurement, and where the measurement

$m(y, s)$ pertains to a function of (y, s) .

For the measurement functional $\Lambda_{m(y,s)}[u]$, equation (3a) becomes

$$(backward\ m) \left[\begin{array}{l} -\frac{\partial \alpha_m}{\partial t} - c \frac{\partial \alpha_m}{\partial x} = \Lambda_{m(y,s)} \left[\delta(x-y)\delta(t-s) \right] \\ \alpha_m(x, T) = 0 \\ \alpha_m(L, t) = 0 \end{array} \right] \quad (7)$$

The equation for the representers r_m is as in (3b). But what does $\Lambda_{m(y,s)} \left[\delta(x-y)\delta(t-s) \right]$ mean?

For the case of point measurements of a field $u(x, t)$ the measurement functional is,

$$\Lambda_{m(x,t)}[u] = \int dx \int dt \delta(x-x_m)\delta(t-t_m)u(x, t) = u(x_m, t_m) \quad (8)$$

so that for a point measurement the "kernel," $K(x, t; x_m, t_m) = \delta(x-x_m)\delta(t-t_m)$. The right-hand side of (7) is then,

$$\begin{aligned} \Lambda_{m(y,s)} \left[\delta(x-y)\delta(t-s) \right] &= \int dy \int ds \delta(y-x_m)\delta(s-t_m)\delta(x-y)\delta(t-s) \quad (9) \\ &= \delta(x-x_m)\delta(t-t_m) \end{aligned}$$

With this latter expression, (7) reduces to (3a). The key to interpreting the measurement functional is the measurement "kernel," K . The right hand side of the first equation of (7) is a "delta function" associated with the particular data type. The example of tomography below will hopefully make this more clear.

IV. MEASUREMENT KERNELS FOR TOMOGRAPHY

Case A: A Line integral of sea-surface height. To begin with a simple case, I will discuss a line average of sea-surface height, i.e. a scalar field. This case is not entirely unphysical, because sound speed is sensitive to pressure, and, were it not

for the vast number of other competing phenomena, tomography could indeed measure the slight pressure variations caused by sea surface displacements. So, suppose we have a field of sea-surface height $q(x, y)$ and we make a measurement that averages along a line segment of this field. In the x direction, suppose the segment goes from x_1 to x_2 , and suppose the path of this segment, p , is $(x, y_0(x))$, where $y_0(x) = mx + b$ (Fig. 1). The measurement of q is

$$q \rightarrow \frac{1}{L} \int_p q(x, y) ds \quad (10)$$

Where L is the length of the line segment. The element of path length $ds = dx \sqrt{1 + \left(\frac{dy}{dx}\right)^2} = dx \sqrt{1 + m^2}$. Therefore, the measurement applied to the scalar field, $q(x, y)$, is

$$\Lambda[q] = \frac{1}{L} \int_0^T dt \int_{x_1}^{x_2} dx \int_0^Y \delta(y - y_0(x)) \delta(t - t_m) \sqrt{1 + m^2} q(x, y) \quad (11)$$

The measurement kernel is therefore,

$$K(x, y, t) = \delta(y - y_0(x)) \delta(t - t_m) (\theta(x - x_1) - \theta(x - x_2)) \sqrt{1 + m^2} \quad (12)$$

where $\theta(x)$ is the familiar theta function. Note that this is still a point measurement in time. This expression may be plugged directly into (7) to obtain the adjoint equation for a tomography measurement and the simple dynamics of that equation.

Case B: Don't forget the ray paths... In reality, tomography consists of ray paths that have 10–100 turning points along a slice of range and depth (Fig. 2). This ray path may be described by a path, Γ , which in the two dimensions of range and depth is $(x, z_0(x))$. Let us suppose that the measurement is by reciprocal tomography, and that the data are the difference between reciprocal travel times $T = \frac{1}{2} (T^+ - T^-)$. The measurement of current, u is therefore

$$T = \Lambda_{m(y,s)}[u] = \int_{\Gamma} \frac{u ds}{c_0^2(x, z)} \quad (13)$$

$$= \int_{x_1}^{x_2} dx \int_{-H}^0 dz \int dt \frac{\delta(z - z_0(x))\delta(t - t_m)u(x, z, t) a(x)}{c_0^2(x, z)}$$

where it is assumed that the current u is to good approximation the projection of the current $\vec{u}(x, y, z)$ along the ray path, e.g., vertical components are negligible, and $a(x) = \sqrt{1 + (dz/dx)}$ is the ray angle. The reader may perform a simple check by doing the integrals over z and t . Note that the arc length $ds = dx\sqrt{1 + (dz/dx)} = a(x) dx$. The measurement kernel is therefore,

$$K(x, z, t) = \frac{\delta(z - z_0(x))\delta(t - t_m)(\theta(x - x_1) - \theta(x - x_2)) a(x)}{c_0^2(x, z)} \quad (14)$$

Case C: A Line Integral of a Current Field. Neglecting the depth dimension again, let us consider a line-integral measurement of a current field, i.e. a vector field (Fig. 3). As in Case A above, the path p is $(x, y_0(x))$, and I introduce $\vec{\tau}$ which is a unit vector in the direction of the acoustic path, $\vec{\tau} = (\tau_x, \tau_y)$. The measurement of \vec{u} is

$$\vec{u} \rightarrow \frac{1}{L} \int_p \vec{u} \cdot \vec{\tau} ds \quad (15)$$

The measurement applied to the vector field is

$$\Lambda[\vec{u}] = \frac{1}{L} \int dt \int_{x_1}^{x_2} dx \int_0^Y \delta(y - y_0(x))\delta(t - t_m)\sqrt{1 + m^2} \vec{u} \cdot \vec{\tau} \quad (16)$$

Note that $\vec{u} \cdot \vec{\tau} = u_x \tau_x + u_y \tau_y$ and $\tau_x = 1/\sqrt{1 + m^2}$, $\tau_y = m/\sqrt{1 + m^2}$. The measurement contains an ambiguity between the x and y components of the flow field. The measurement kernel is a vector in this case; it is

$$\vec{K}(x, y, t) = \delta(y - y_0(x))\delta(t - t_m)(\theta(x - x_1) - \theta(x - x_2))\sqrt{1 + m^2}\vec{\tau} \quad (17)$$

Though I have written $\delta(t - t_m)$ throughout this section, this is not entirely true because of the finite time-of-flight of tomography signals. Most of the time this is unimportant, but one example of when it IS important is for ocean tides and O(1-hr-long) trans-Pacific acoustic transmissions.

V. THE ADJOINT AND REPRESENTER EQUATIONS FOR THE SHALLOW WATER MODEL

Assuming the reader is reasonably familiar with the shallow-water equations, I will skip a multitude of introductory formalities, and write the adjoint equation for a linear shallow water model and tomographic data directly:

$$-\frac{\partial \alpha_m^u}{\partial t} + f \alpha_m^v - H \frac{\partial \alpha_m^q}{\partial x} + r_u \alpha_m^u = \quad (18a)$$

$$\left[\begin{array}{l} \frac{1}{L} \tau_x \delta(t - t_m) \delta(y - y_0(x)) (\theta(x - x_1) - \theta(x - x_2)) \sqrt{1 + m^2} \quad (\text{for current}) \\ 0 \quad (\text{for sea - surface height}) \end{array} \right.$$

$$-\frac{\partial \alpha_m^v}{\partial t} - f \alpha_m^u - H \frac{\partial \alpha_m^q}{\partial y} + r_v \alpha_m^v = \quad (18b)$$

$$\left[\begin{array}{l} \frac{1}{L} \tau_y \delta(t - t_m) \delta(y - y_0(x)) (\theta(x - x_1) - \theta(x - x_2)) \sqrt{1 + m^2} \quad (\text{for current}) \\ 0 \quad (\text{for sea - surface height}) \end{array} \right.$$

$$-\frac{\partial \alpha_m^q}{\partial t} - g \left(\frac{\partial \alpha_m^u}{\partial x} + \frac{\partial \alpha_m^v}{\partial y} \right) + r_q \alpha_m^q = \quad (18c)$$

$$\left[\begin{array}{l} 0 \quad (\text{for current}) \\ \frac{1}{L} \delta(t - t_m) \delta(y - y_0(x)) (\theta(x - x_1) - \theta(x - x_2)) \sqrt{1 + m^2} \quad (\text{for sea - surface height}) \end{array} \right.$$

The shallow water equations here have drag coefficients $\{r_u, r_v, r_q\}$, and the

boundary conditions are periodic in the x direction and no-flow in the y direction, i.e. this is a channel. On the right-hand side, I have written the "delta function" appropriate for either a line-integral measurement of current, or a line-integral measurement of sea-surface height. These formidable looking expressions have a simple interpretation. Recall that in integrating the adjoint equation backwards in time and space for a point measurement the integration "picked up" a delta function impulse at the time and place of the measurement. In the case of tomography, the impulse is a knife-edge impulse - a segment delta function, if you will (Fig. 4). The impulse in the case of current is tempered by the projection τ_x or τ_y .

These equations may be used to solve for THE adjoint $\{\alpha_m^u(x, t), \alpha_m^v(x, t), \alpha_m^q(x, t)\}$, which may then be used to get THE representer $\{r_m^u(x, t), r_m^v(x, t), r_m^q(x, t)\}$ as described earlier.

The matrix R is used to calculate the coefficients $\hat{\beta}$. R is found by applying the measurement functional (a vector...) to the representer field \vec{r} , so that for a single measurement of current (i.e. a single transmission),

$$R = \frac{1}{L} \int dt \int_{x_1}^{x_2} dx \int_0^Y \delta(y - y_0(x)) \delta(t - t_m) \sqrt{1 + m^2} (r^u \tau_x + r^v \tau_y) \quad (19)$$

VI. THE REPRESENTER, $\{r_m^u(x, y, t), r_m^v(x, y, t), r_m^q(x, y, t)\}$, FOR TOMOGRAPHY AND A SHALLOW WATER MODEL

Before describing the representers for tomography, I will begin by first describing a suggestive run of the shallow-water model in the forward direction, and then the representer for a point measurement of sea-surface height. With this introductory discussion, a better understanding of the tomography representers will follow. As a reminder, the covariances employed here are "white" so no smoothing or additional structure is introduced to the representers through the covariances. The grid spacing for the model employed is 100-km, and the domain under consideration is 1000-km across channel and 2000-km along channel. Bear in mind that the grid spacing is crude and that the model implementation is perhaps not the best either, hence the results shown here are not perfect.

The run of the model in the forward direction (Fig. 5) starts out with no initial currents and a gaussian displacement of the sea surface. The figure shows a selected number of snapshots of the fields, equally spaced in time. The initial disturbance of the sea surface propagates nicely away from its origin, and later the effects of the periodic boundary conditions may be seen. The blue arrows at the top of each frame are, of course, the current vectors.

A representer for a point measurement of sea-surface height is shown in Figure 6. The time of the measurement is the third frame, which shows the delta-function impulse discussed earlier in the integration of the adjoint equation. This delta function "propagates" away from its origin, as for the forward model run, except that the representer influences the model solution at times before and after the measurement. The similarity with the forward model run, and recall that a steep gaussian may be used as an approximation for a delta function, is apparent.

Finally getting to the gratification of the tomographic representer, Figure 7 shows the representer for a tomographic measurement of sea-surface height. The tomographic path is as described in Fig. 4. Again, the third frame shows the "delta function" impulse, a segment impulse in this case. Most of the wigglyness of the representer is for wavenumbers perpendicular to the acoustic path. This property brings to mind the tomographic measurement of the internal tide, and one might consider at this point a reduced gravity model... The representer has low-wavenumber disturbances along the acoustic path.

Figure 8 shows the representer for a tomographic measurement of current at the time of the third frame. If Figure 7 had a delta function in elevation at the time of the measurement, then the third frame of Fig. 8 has the derivative of a delta function in elevation at the time of the measurement. A strong "current" along the path at the time of the measurement is also apparent in the representer, and as in all previous cases the disturbance propagates away from the time and place of the measurement.

In Figs. 7 and 8, note that a measurement of elevation (current) will influence the estimation of current (elevation), since the solution for the perturbations to the first guess field \vec{u}_F is a linear superposition of representers. However, this is true only if the dynamical errors f are weak; large errors in dynamics effectively decouple elevation and current (Dushaw, et al., 1997).

REFERENCES

- Baym, G., Lectures on Quantum Mechanics, Benjamin Cummings, Reading, Massachusetts, 1969.
- Bennett, A. F., Array Design by Inverse Methods, *Prog. Oceanogr.*, **15**, 129–156, 1985.
- Bennett, A. F., Inverse Methods and Data Assimilation, Lecture Notes for the 1999 Summer School, Oregon State University, College of Oceanic and Atmospheric Sciences, Corvallis, Oregon, July 1999, 235 pp.

6 October 1999

Bennett, A. F., *Inverse Methods in Physical Oceanography*, Cambridge University Press, New York, 1999.

Dushaw, B. D., G. D. Egbert, P. F. Worcester, B. D. Cornuelle, B. M. Howe, and K. Metzger, 1997. A TOPEX/POSEIDON global tidal model (TPXO.2) and barotropic tidal currents determined from long-range acoustic transmissions. *Prog. Oceanogr.*, **40**, 337–367.

ftp.oce.orst.edu, cd /dist/bennett/class - this might be the location for the code for the shallow-water model in a channel.

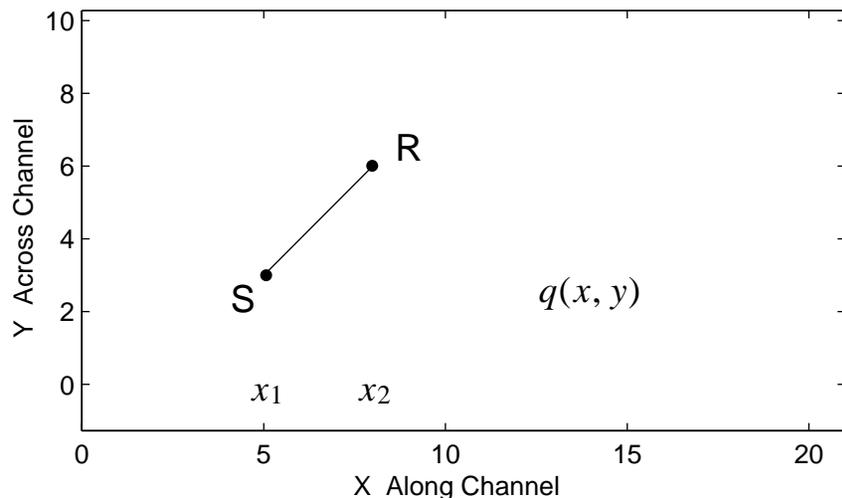


Figure 1. Line-integral measurement of the scalar field $q(x,y)$. S and R are acoustic source and receiver.

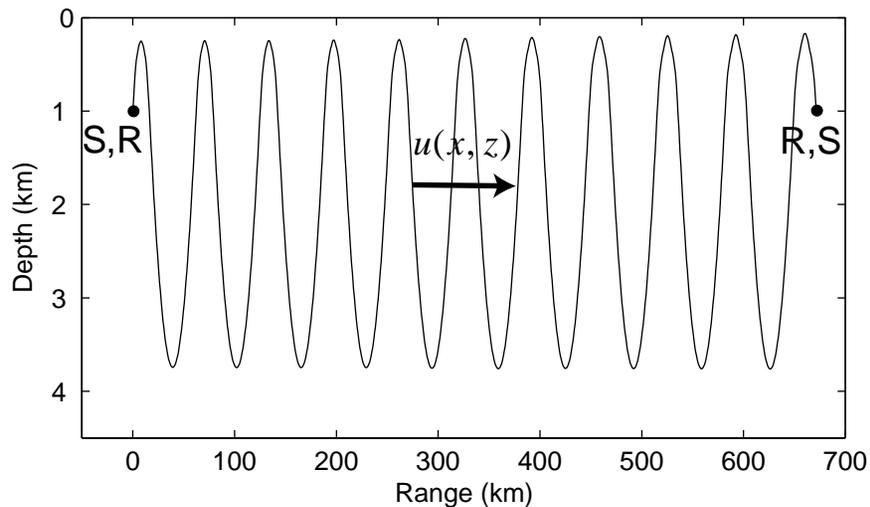


Figure 2. A ray path along a slice in range and depth for the region of the Sargasso Sea. Reciprocal tomography measures the current $u(x,z)$ averaged along the ray path.

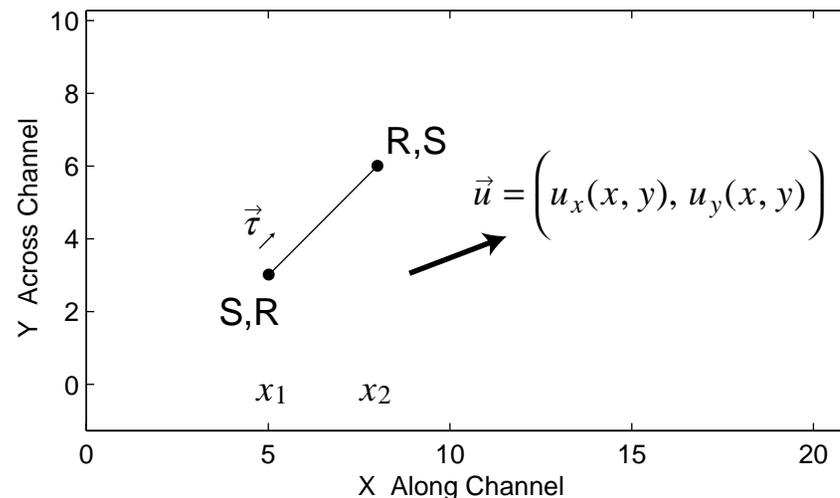


Figure 3. Line-integral measurement of the vector field $u(x,y)$ by reciprocal shooting. τ is a unit vector along the acoustic path.

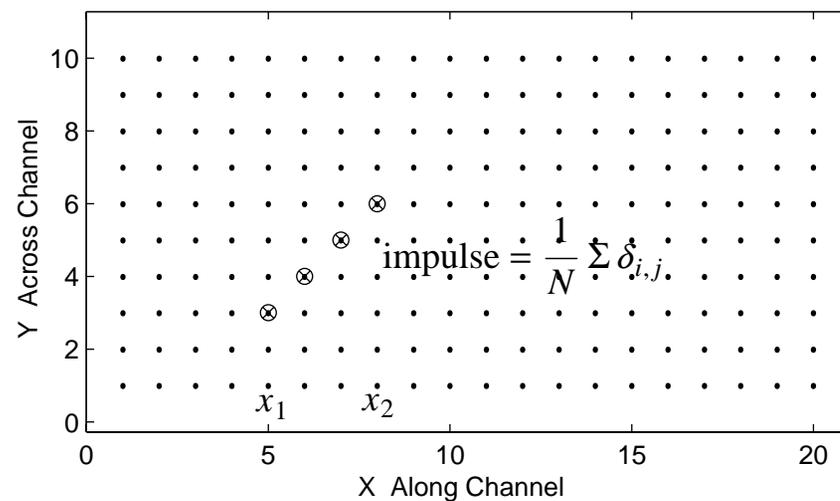


Figure 4. In the discretized implementation of the model, the "delta function" forcing on the right-hand side of (18) when calculating the adjoint is given as the average of the Kronecker delta functions at the time of the measurement.

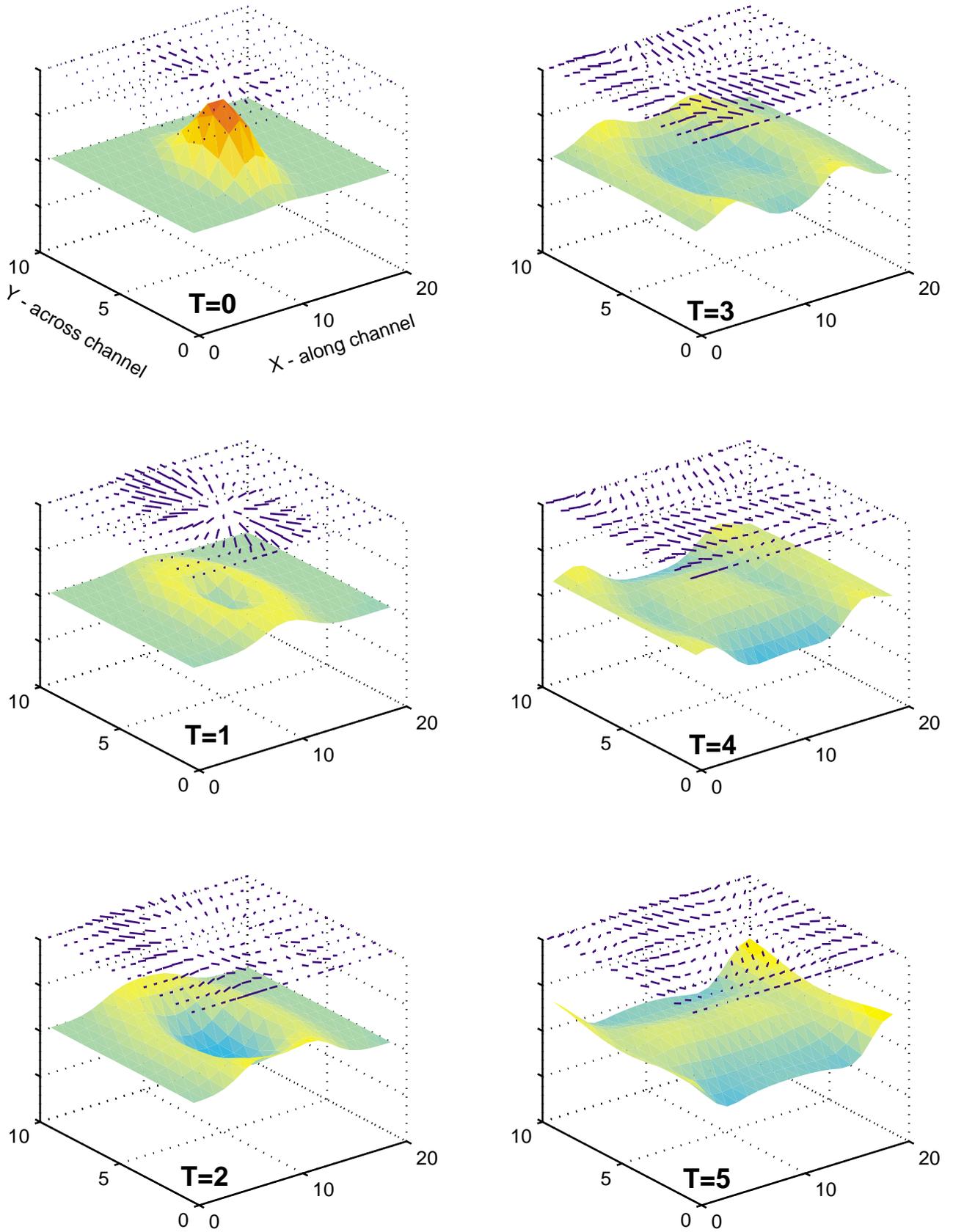


Figure 5. An example of a forward model run. The sea-surface begins with a gaussian displacement at $T=0$.

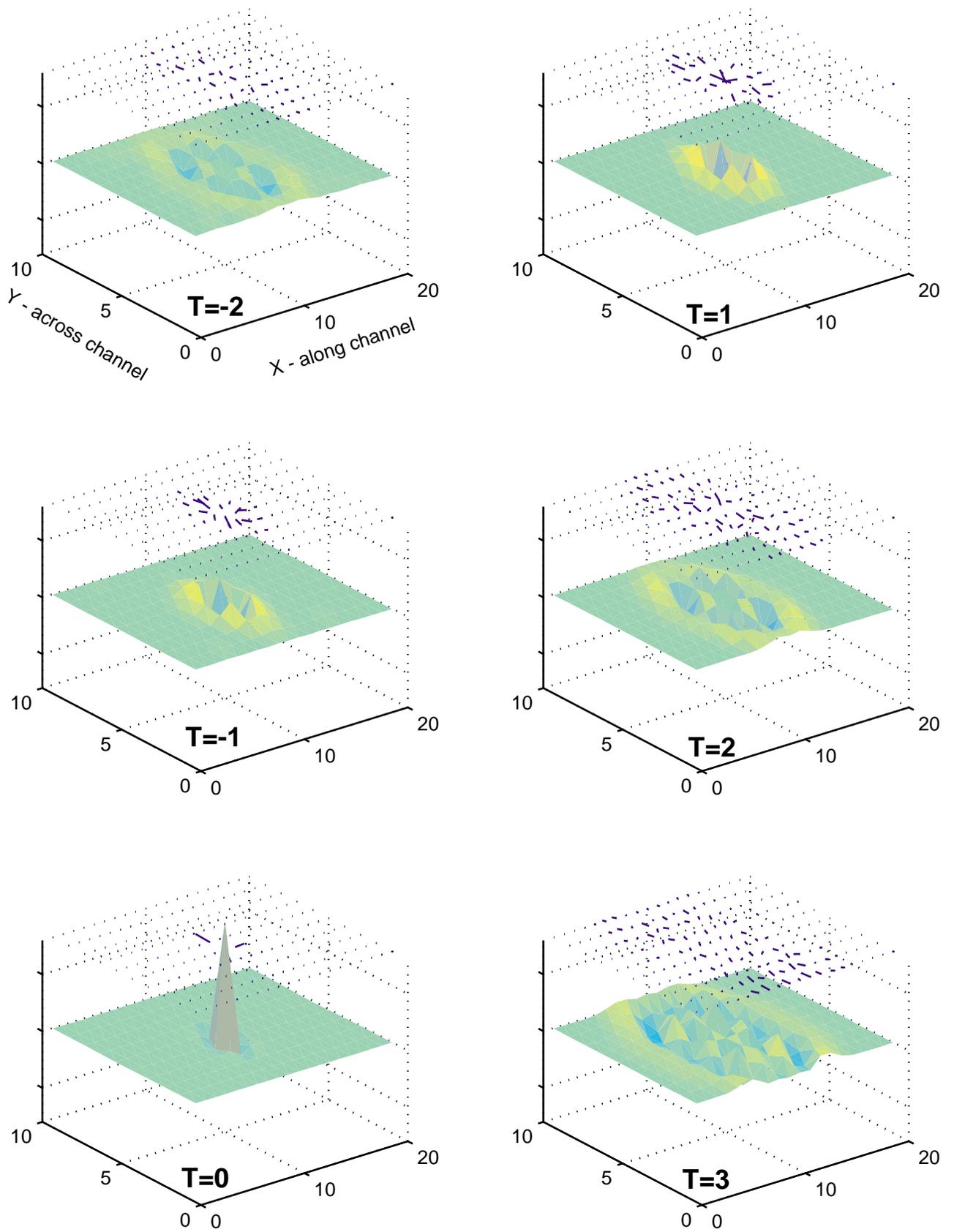


Figure 6. An example of a representer for a point measurement of sea-surface elevation at time $T=0$.

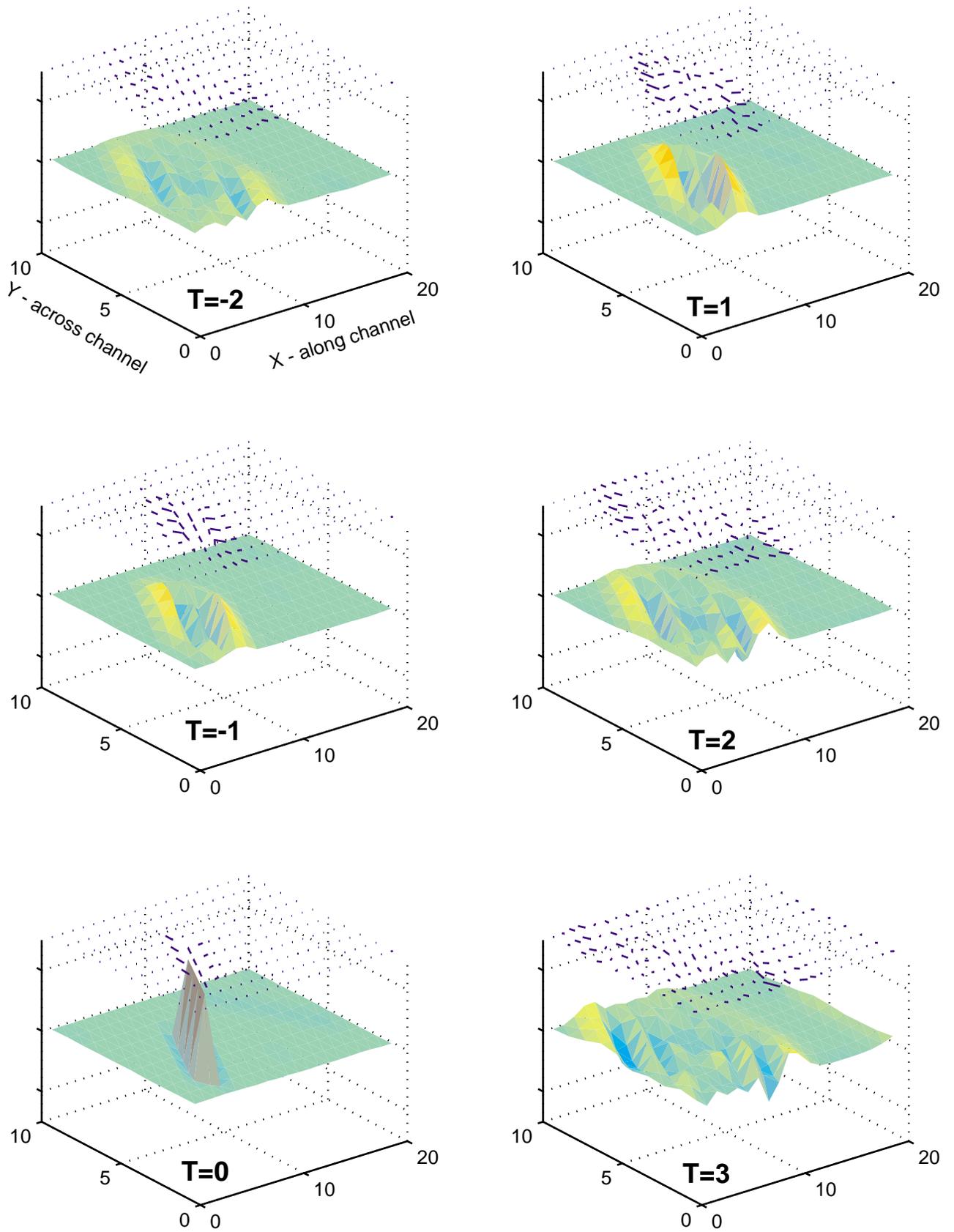


Figure 7. An example of a representer for a line-integral measurement of sea-surface elevation at time $T=0$.

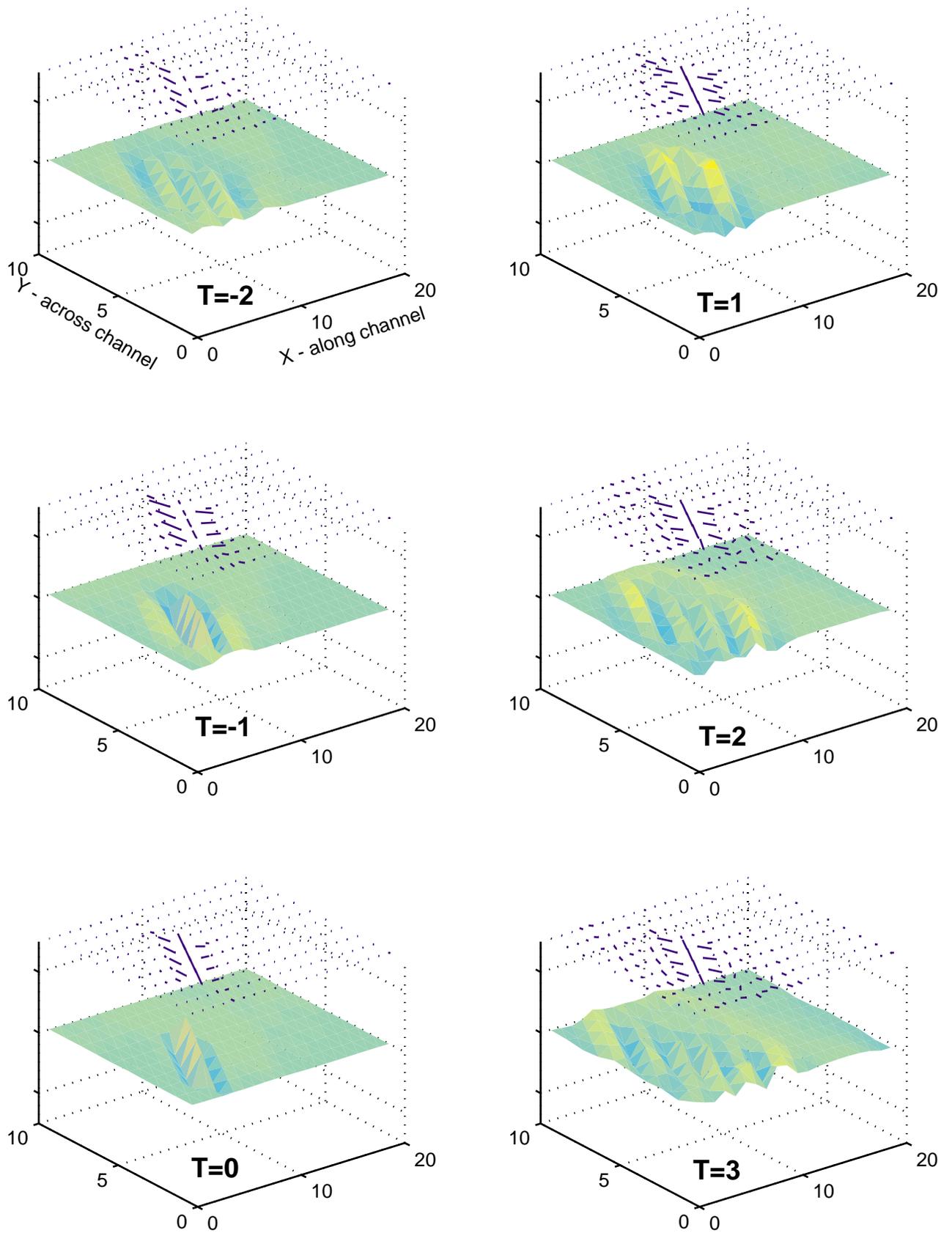


Figure 8. An example of a representer for a line-integral measurement of current at time $T=0$.