Q. J. R. Meteorol. Soc. (2002), **128**, pp. 1–999 doi: 10.1256/qj.yy.n

Eulerian backtracking of atmospheric tracers: II Numerical aspects

By F. HOURDIN[1]*, O. TALAGRAND[1] and A. IDELKADI[1]
[1] *Laboratoire de Météorologie Dynamique, CNRS/IPSL, France*

(Received 1 January 2000; revised 31 January 2001)

KEYWORDS: Atmospheric transport, inverse methods, adjoint, backtracking

SUMMARY

In Part I of this paper, a mathematical equivalence was established between Eulerian backtracking or retro-transport, on the one hand, and adjoint transport with respect to an air-mass-weighted scalar product, on the other. The time symmetry which lies at the basis of this mathematical equivalence can however be lost through discretisation. That question is studied, and conditions are explicitly identified under which discretisation schemes possess the property of time symmetry. Particular consideration is given to the case of the LMDZ model. The linear schemes used for turbulent diffusion and sub-grid scale convection are symmetric. For the Van Leer advection scheme used in LMDZ, which is non linear, the question of time symmetry does not even make sense. Those facts are illustrated by numerical simulations performed in the conditions of the European Transport EXperiment (ETEX). For a model that is not time-symmetric, the question arises as to whether it is preferable, in practical applications, to use the exact numerical adjoint, or the retro-transport model. Numerical results obtained in the context of one-dimensional advection show that the presence of slope limiters in the Van Leer advection scheme can produce in some circumstances irrealistic (in particular, negative) adjoint sensitivities. The retro-transport equation, on the other hand, generally produces robust and realistic results, and always preserves the positivity of sensitivities. Retro-transport may therefore be preferable in sensitivity computations, even in the context of variational assimilation.

1. INTRODUCTION

In Part I of this paper (Hourdin et Talagrand, 2005), *retro-transport* of a perfect retro-tracer was introduced, which describes the reversed temporal evolution of a scalar field conserved along air trajectories. The retro-transport equation is derived from the direct equation through simple transformations. The direction of explicit advection is reversed. This backward advection can be formulated in a Lagrangian as well as in a Eulerian framework, a remark which is at the basis of the concept of Eulerian backtracking. Terms describing linear sources or sinks of tracers are kept unchanged in the retro-transport equation. Terms representing diffusion by unresolved time-symmetric motions of the transporting air are also unchanged. For parametrisation of non-symmetric subgrid-scale motions, the direction of the motions has to be changed in the reverse computation. This applies to the so-called *mass-flux parametrisations* of cumulus convection.

For a linear tracer, the retro-transport equation turns out to be the adjoint of the direct transport equation, with respect to the scalar product defined by integration with respect to air mass, $\langle \phi, \psi \rangle = \int \rho \phi \psi \ \mathbf{dx}dt$, where ϕ and ψ are tracer concentrations per unit mass of transporting air (ρ being the density of transporting air, and $\mathbf{dx}$ and dt volume and time elements respectively). The retro-transport allows to rewrite any measurement of the direct tracer (with mass concentration c) of the form

$$\mathcal{M} = \int_{\Omega \times \tau} \rho \mu c \ \mathbf{dx}dt \tag{1}$$

* Corresponding author, Laboratoire de Météorologie Dynamique, UPMC, Tour 45-55, 3eme étage, BP-99, Jussieu, 4 place Jussieu, 75 005 Paris, France, hourdin@lmd.jussieu.fr

as a function of sources and boundary conditions. By taking for c^* the solution of the adjoint transport equation defined by the condition that $c^* = 0$ at a time t_f posterior to the end of the measurement process ($\mu(\mathbf{x}, t) = 0$ for $t > t_f$) and along the outflow boundary $\partial\Omega_o$ of the physical domain Ω under consideration (defined by $\mathbf{v}.\mathbf{n} > 0$ where $\mathbf{v}$ is the wind vector and $\mathbf{n}$ is a unit outward vector normal to the boundary $\partial\Omega$) and with a zero surface flux, one obtains a receptor-oriented expression for the measurement:

$$\begin{aligned}\mathcal{M} &= \int_\Omega \rho c^* c_{|_{t_i}} \mathbf{dx} - \int_{\partial\Omega_i \times \tau} \rho c^* c \ \mathbf{v}.\mathbf{n}_{|\partial\Omega_i} \ dsdt \\ &+ \int_{\Omega\times\tau} \rho c^* \sigma \ \mathbf{dx}dt + \int_{\mathcal{S}\times\tau} \Sigma c^* dxdydt \end{aligned} \tag{2}$$

where $\tau = [t_i, t_f]$ is the time domain, $\partial\Omega_i$ is the inflow boundary ($\mathbf{v}.\mathbf{n} \leqslant 0$) and σ and Σ are 4D and surface sources of tracer respectively.

For a conservative tracer, c^* is, up to a scaling factor, the distribution at a given time in the past of the air which will be sampled later on for the measurement. When there is no inflow into nor outflow from the domain, this distribution can be combined with the direct tracer concentration at any time between the end of the emission (σ or Σ) and the beginning of the measurement process, to estimate the measurement as

$$\mathcal{M} = \int_\Omega \rho c c^* \mathbf{dx} \tag{3}$$

The purpose of this second Part is to analyse the equivalence of adjoint and retro- transport in the numerical world. Because of the spatial and temporal discretisations, the time symmetry which is at the basis of that mathematical equivalence may be lost. A discretized scheme will be time symmetric if the direct code used in retro-transport mode is identical with the exact numerical adjoint of the direct code. Only linear schemes can be time symmetric. This question of time symmetry of numerical algorithms is discussed here for the particular case of the tracer version of the LMDZ global climate model (described in section 2). A numerical illustration of compared direct and retro-transport computations is presented in the case of the ETEX-1 experiment (section 3). Then the time symmetry of numerical algorithms is checked (section 4). The linear schemes used in LMDZ are shown to be time symmetric. However, the symmetry is lost when the Van Leer (1977) finite volume advection scheme is used. This results from the non-linearities introduced in the scheme in order to ensure important properties such as positivity or monotony (a monotonic distribution remains monotonic after advection). Implications for inversion of atmospheric transport and variational assimilation are then discussed (section 5) before some conclusions are drawn (section 6).

2. Transport component of the LMDZ model

(*a*) *The LMDZ general circulation model*

LMDZ is a general circulation model developed at Laboratoire de Météorologie Dynamique (LMD) for climate studies. It is used in particular to study the coupling between chemistry, micro-physics and atmospheric dynamics on Earth, Mars and Titan (Hauglustaine et al., 2004; Lefèvre et al., 2004; Rannou et al., 2002).

It is based on a finite-difference formulation of the primitive equations of meteorology. Equations are discretised horizontally on a global Arakawa C-grid (see *e. g.*, Kasahara, 1977) which can be refined on a particular region (the "Z" of LMDZ standing for that "Zoom" capability). The vertical discretisation is based on the σ-pressure coordinate (see *e. g.*, Simmons and Burridge, 1981). With scalars defined at the centre of control volumes and winds at the interfaces, the grid is well suited for the implementation of conservative finite-volume schemes for advection. Here we use the LMDZ3 version of the Earth climate model which includes classical parameterisations for radiative transfer (Morcrette, 1984; Fouquart and Bonnel, 1980), turbulent eddy-diffusion in the boundary layer (Laval et al., 1981), convection (Tiedtke, 1989) and clouds (Le Treut and Li, 1991).

(*b*) *The tracer component*

The transport of atmospheric tracers was implemented in LMDZ by Hourdin and Armengaud (1999). The transport equation reads

$$\frac{\partial c}{\partial t} + \mathbf{v}.\mathbf{grad}\ c + \lambda c + \frac{1}{\rho}\frac{\partial\left(F_d + F_c\right)}{\partial z} = \sigma \tag{4}$$

with the continuity equation for the air

$$\frac{\partial \rho}{\partial t} + \mathrm{div}\left(\rho \mathbf{v}\right) = 0 \tag{5}$$

F_d and F_c correspond to diffusive and convective parametrisations of the vertical subgrid-scale flux of tracer c. The coefficient λ, which accounts for a linear sink (or source), can be either a constant (radioactive decay) or a function of space and time (scavenging by rain, chemical reaction with a prescribed oxidant, ...).

Large scale advection of tracers is computed on the basis of the Van Leer I scheme (Van Leer, 1977), a second-order finite-volume scheme with slope limiters which guarantee positivity and monotony. The diffusive flux reads

$$F_d = -K_z \rho \frac{\partial c}{\partial z} \quad \text{with } F_d|_{\mathrm{surf}} = \Sigma \quad \text{and } F_d|_{\mathrm{top}} = 0 \tag{6}$$

K_z being the eddy-diffusivity and Σ the surface emission of tracer. For convection, the Tiedtke (1989) mass-flux parametrisation separates the atmospheric column between an updraught ($\hat{}$) and a downdraught ($\check{}$). The convective flux reads

$$F_c = \hat{f}\hat{c} - \check{f}\check{c} - (\hat{f} - \check{f})c \tag{7}$$

where the tracer concentrations per unit mass of transporting air in the updraught ($\hat{c}$) and downdraught ($\check{c}$) are given by

$$\frac{\partial \hat{f}\hat{c}}{\partial z} = \hat{e}c - \hat{d}\hat{c} \tag{8}$$

$$-\frac{\partial \check{f}\check{c}}{\partial z} = \check{e}c - \check{d}\check{c} \tag{9}$$

with the associated continuity equations for convective mass fluxes

$$\frac{\partial \hat{f}}{\partial z} = \hat{e} - \hat{d} \tag{10}$$

$$-\frac{\partial \check{f}}{\partial z} = \check{e} - \check{d} \tag{11}$$

where the convective mass flux $\hat{f}$ (kg m^{-2} s^{-1}), entrainment $\hat{e}$ into and detrainment $\hat{d}$ from the updraught (resp. $\check{f}$, $\check{e}$ and $\check{d}$ for the downdraught), are positive quantities equal to zero at the upper and lower boundaries.

Given the space and time evolution of ρ, $\mathbf{v}$, λ, K_z, $\hat{e}$, $\hat{d}$, $\check{e}$ and $\check{d}$ ($\hat{f}$ and $\check{f}$ being computed according to equations (10) and (11)), the tracer concentration $c(\mathbf{x}, t)$ depends on the volume (σ) and surface (Σ) sources, as well as on the initial concentration $c(\mathbf{x}, t_i)$ at initial time t_i and tracer inflow across $\partial\Omega_i$.

(*c*) *Off-line and retro-transport mode*

In LMDZ, tracer transport can be computed either on-line, by "passing" directly the required meteorological variables ρ, $\mathbf{v}$, K_z, $\hat{e}$, $\hat{d}$, $\check{e}$, $\check{d}$ to the transport algorithms, or off-line, by reading the same variables from an archive, built-up in the course of a prior integration of the numerical model. When λ is also a 4D variable linked to meteorology (such as a rate of scavenging by rain), it is archived as well. The off-line mode then requires the storage of 9 3D variables (the vertical wind component being recomputed from the mass continuity equation) on the model grid. This storage is done with a time-step of 3 hours in the simulations presented below, the meteorological variables being averaged in an appropriate way between two successive archiving times.

The off-line version of LMDZ can be used either in direct or retro-transport mode. As explained in Part I, the retro-transport equation is obtained from the direct equations (4 to 11) through the following formal transformation: $(t, \rho, \mathbf{v}, \lambda, K_z, \hat{e}, \hat{d}, \check{e}, \check{d}) \rightarrow (-t, \rho, -\mathbf{v}, \lambda, K_z, \check{d}, \check{e}, \hat{d}, \hat{e})$. In practice, the retro-transport mode consists in running exactly the same off-line code as in the direct mode, but reading the meteorological archives backward in time while changing $\mathbf{v}$, $\hat{e}$, $\hat{d}$, $\check{e}$, $\check{d}$ into $-\mathbf{v}$, $\check{d}$, $\check{e}$, $\hat{d}$, $\hat{e}$.

3. Numerical illustration

We show a numerical illustration of Eulerian backtracking in the context of the European Transport EXperiment (ETEX).† On 23 October 1994, 340 kg of the insoluble gas Perfluoro-Methyl-Cyclo-Hexane (PMCH) were emitted continuously over a 12-hour period starting at 1600 UTC (t_0) at the top of an 8m-high tower located in Monterfil (Western France) in a strong and maintained westerly to south-westerly flow. A 4-day nudged ‡ integration of the LMDZ model is performed first, using a stretched grid with a mean horizontal resolution of about 120×120 km^2 over Europe. An archive is built with a 3-hour time-step as described above. The left column of Fig. 1 shows the time evolution of the surface concentration as obtained with the direct off-line transport model for the emission of 340 kg of PMCH at Monterfil, between t_0 and $t_0 + 3$h. Note that the combination of horizontal transport by sheared wind plus vertical mixing by turbulence explains the rapid horizontal dispersion of the tracer plume. The

† ETEX was organised by the World Meteorological Organisation, the International Atomic Energy Agency and the European Commission, to test the ability of transport models to predict the dispersion of a cloud of pollutant on continental scales.

‡ In order to remain as close as possible to the observed synoptic situation, the simulated wind and temperature are relaxed toward the meteorological analyses of the European Centre for Medium-range Weather Forecasts with a time constant of 2.5 hours. This method, often called nudging, is an economical alternative to full assimilation of meteorological observations (Jeuken et al., 1996).

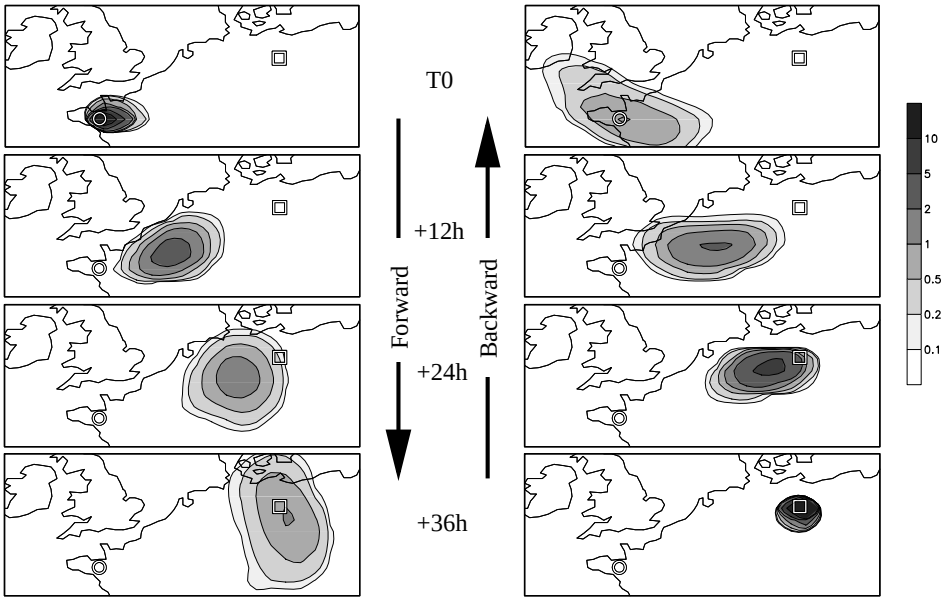

Figure 1. ETEX-1 experiment: snapshots of the simulated surface concentration of PMCH (in ng/m^3) over the 36-hour period following release time t_0. The time-step for injection and sampling is 3 hours (for instance the plume at $t_0 + 12$h corresponds in fact to an average from $t_0 + 12$ to $t_0 + 15$h). The left panels correspond to a direct simulation with injection at Monterfil (circle) between t_0 and $t_0 + 3$h of 340 kg of PMCH. The right panels correspond to a backward integration with retro-injection of the same quantity at station D05 (square) between $t_0 + 39$h and $t_0 + 36$h.

location of the source is the white circle. 24 hours after emission, the PMCH plume reaches station D05 (square in the figure). The peak concentration is obtained for the same station 36 hours after emission. On the right hand side, we show the results of a reverse transport computation for which the 340 kg of PMCH are injected at station D05 uniformly between $t_0 + 39$h and $t_0 + 36$h. Reciprocity is illustrated in Fig. 1 by the fact that the same concentration is observed at D05 36 hours after emission in the direct simulation (square, lowest panel on the left) and at Monterfil at emission time t_0 for the reverse simulation (circle, uppermost panel on the right).

For a more accurate check of the time symmetry, one can retro-inject one tracer for each measurement (every 3 hours in our case). In Fig. 2(a), we compare the time evolution of PMCH concentration at station D05 as actually observed on the occasion of ETEX, and as obtained from direct transport computation or through backward reconstruction. Here, the direct tracer is injected uniformly between t_0 and $t_0 + 12$h. Consistently, the retro-plumes are averaged over the same 12-hour period. The computation is done both with the standard version of the model based on the Van Leer I second-order finite volume scheme and with the first order upstream scheme originally proposed by Godunov (1959). The backward estimates are reconstructed with one retro-simulation for each 3-hour measurement. For the Van Leer scheme, the backward reconstruction (plus signs) does not exactly match the direct estimate (black squares) but the discrepancy is much less than between simulations and observations (full curve). For the

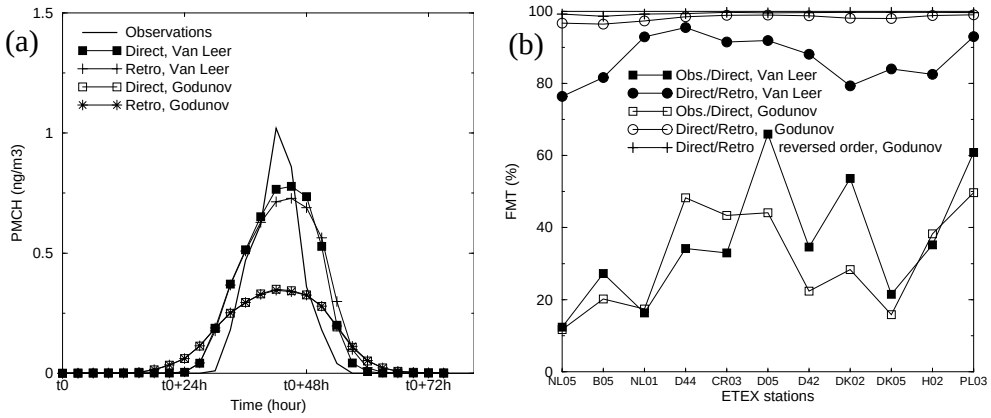

Figure 2. Numerical test of the time symmetry of atmospheric transport. Panel **a:** Time evolution of the PMCH concentration at station D05 as observed and simulated with the Van Leer I and Godunov schemes and reconstructed from a series of backward transport computations with both schemes. Panel **b:** FMTs for the 11 stations retained for model intercomparison by the ETEX team. The FMTs measuring the difference between direct simulation and observations for the Van Leer I and Godunov schemes are shown as well as that measuring the difference between the direct and reverse estimates.

Godunov scheme, the direct and backward estimates (white squares and stars) are indistinguishable at the figure accuracy.

The discrepancies are quantified by using the Figure of Merit in Time (FMT) defined, when two time series are plotted on the same graph, as the ratio of the overlap area between the two curves to the total area defined by the envelope. The FMT, which varies from 0 to 100%, depends both on the time shift between the two signals and on their relative magnitudes. For the 11 stations retained by the ETEX team for specific analysis on the time series, the mean FMT measuring the difference between the direct simulations and observations (black and white squares in Fig. 2(b)) is about 40%§. By comparison the FMT measuring the difference between the direct and retro-transport computation is systematically larger than 75% (black circles). The difference is still much smaller with the Godunov scheme (white circles, mean FMT of 98.2%). The symmetry can further be improved by inverting the sequence of individual operators (+ signs, mean FMT of 99.5%) as explained in Section 4e.

4. Time symmetry of the LMDZ transport algorithms

In LMDZ, large scale advection and parametrisations are applied sequentially. A model integration reduces to a sequence of transition steps from concentration field c^n to field c^{n+1}. A complete time-step is the computation of a number of successive transition steps. We are going to investigate the symmetry of each individual step by comparing its retro-transport and adjoint formulations.

(*a*) *Adjoint model*

Let us denote by $\langle \phi, \psi \rangle^n = \phi^T M^n \psi$ a discretised scalar product at stage n, where M^n is a symmetric positive definite matrix, and by L the linear tangent

§ This score is typical of good dispersion models according to the intercomparison study conducted in the frame of ETEX (Graziani et al., 1996; Klug et al., 1992). A systematic evaluation of the direct model is out of purpose here and will be presented elsewhere.

model which describes the forward evolution of a perturbation dc

$$dc^{n+1} = Ldc^n \tag{12}$$

The adjoint L^* of L with respect to scalar products M^n and M^{n+1}, is defined by $\langle L^*\phi, \psi\rangle^n = \langle \phi, L\psi\rangle^{n+1}$ for any ϕ and ψ, i. e.

$$L^* = (M^n)^{-1} L^T M^{n+1} \tag{13}$$

The corresponding adjoint model reads

$$c^*_n = L^* c^*_{n+1} \tag{14}$$

With $c^* = \nabla_c J$, that model describes the backward temporal evolution of the gradient of any function J with respect to the model state variables c (see e. g. Talagrand and Courtier, 1987). This is easily shown by identifying the expressions for the gradient at stages n and $n+1$, $dJ = \langle \nabla_{c^n} J, dc^n\rangle^n = \left\langle \nabla_{c^{n+1}} J, dc^{n+1}\right\rangle^{n+1}$.

For a linear model, a linear measurement of concentration $\mathcal{M} = \left\langle \mu, c^N\right\rangle^N$ can be evaluated at any time-step n between source and detection time (N) as $\mathcal{M} = \langle \nabla_{c^n}\mathcal{M}, c^n\rangle^n$. This, with $c^{*n} = \nabla_{c^n}\mathcal{M}$, is a numerical equivalent to Eq. (3).

(*b*) *Checking time symmetry of numerical algorithms*

Hence, the time symmetry of a linear numerical model is equivalent to the identity between the retro-transport mode of the numerical model R (obtained from the direct model L by simple physically-based transformations, here $(\mathbf{v}, \hat{e}, \hat{d}, \check{e}, \check{d}) \rightarrow (-\mathbf{v}, \check{d}, \check{e}, \hat{d}, \hat{e})$) and the exact adjoint code of the direct numerical model, $R = L^*$ where L^* is given by Eq. (13). For the air-mass-weighted scalar product, $M = \mathrm{diag}(m_i)$ where m_i is the air mass in grid cell i. If we take for matrix indices the classical convention that $\psi_i = \sum_j L_{i,j}\phi_j$, we find for the time symmetry the following condition between the entries of the direct and retro-transport matrices

$$m^n_i R_{i,j} = m^{n+1}_j L_{j,i} \tag{15}$$

For parametrisations which do not affect the large scale air density (turbulent mixing, convection, linear sink or advection by a non divergent wind field), so that $M^n = M^{n+1} = M$, time symmetry can be checked somewhat more easily by considering a flux-like form of the model

$$M\left(c^{n+1} - c^n\right) = Ac^n \tag{16}$$

which corresponds to $L = I + M^{-1}A$. From Eq. (13), $L^* = I + M^{-1}A^T$ so that the adjoint model (14) reads

$$M\left(c^{*n} - c^{*n+1}\right) = A^T c^{*n+1} \tag{17}$$

For a model of form (16), the time symmetry will thus be ensured in the numerical world if the matrix B associated with the retro-transport mode of the direct model is equal to the transpose matrix of the direct model, $B = A^T$.

The same condition, $B = A^T$, also insures time symmetry of a scheme of form

$$M\left(c^{n+1} - c^n\right) = A\left[\gamma c^n + (1-\gamma)\, c^{n+1}\right] \tag{18}$$

which covers the cases of explicit ($\gamma = 1$), implicit ($\gamma = 0$) and Crank-Nicholson ($\gamma = 1/2$) time integration for a same operator A. After transforming the equation above into

$$c^{n+1} = \left[I - (1-\gamma)\, M^{-1}A\right]^{-1} \left[I + \gamma M^{-1}A\right] c^n \tag{19}$$

(the inverse matrix on the right hand side exists for a small enough time-step) application of Eq. (13) yields

$$c^{*n} = M^{-1} \left[I + \gamma A^T M^{-1}\right] \left[I - (1-\gamma)\, A^T M^{-1}\right]^{-1} M\ c^{*n+1} \tag{20}$$

$$= \left[I + \gamma M^{-1}A^T\right] \left[I - (1-\gamma)\, M^{-1}A^T\right]^{-1} c^{*n+1} \tag{21}$$

The matrices $[I + aQ]$ and $[I + bQ]^{-1}$ commute for any matrix Q and scalars a and b so that Eq. (21) is of the form (19), A being replaced by A^T. This proves the announced result that the condition $B = A^T$ ensures time symmetry of scheme (18).

(*c*) *The Godunov advection scheme*

The time symmetry of the Godunov scheme (suggested by Fig. 2) is derived here in the simple case of 1-D advection with a non divergent wind $u > 0$, a constant air density ρ, a regular grid with grid spacing δx and time-step δt (the more general case of advection by a divergent wind is considered in Appendix A). Let us note by $\alpha = u\delta t/\delta x$ the Courant number. The Godunov scheme reads for grid cell i and times n and $n+1$

$$c_i^{n+1} - c_i^n = \alpha \left(c_{i-1}^n - c_i^n\right) \tag{22}$$

which is of the form (16) with $M = I$. The corresponding retro-transport scheme is obtained by inverting the roles of n and $n+1$, and changing u into $-u$. Because of the change of sign of u, the upwind scheme now reads

$$c^{*n}_i - c^{*n+1}_i = \alpha \left(c^{*n+1}_{i+1} - c^{*n+1}_i\right) \tag{23}$$

It is easily seen that the matrix defining the operator on the right-hand side of this equation is the transpose of the matrix defining the operator on the right-hand side of Eq. (22), which shows the time symmetry of scheme (22).

The extension of the Godunov scheme to three dimensions is done in LMDZ through time splitting. At each time-step, the advection is computed successively in the directions x, y, z, y, x (with half a time-step for each computation in the x and y directions). In each direction, both the tracer transport (Eq. (A.1) of appendix A) and air-mass continuity equation (Eq. (A.2)) are integrated. This ensures time symmetry of the global 3D scheme.

In addition to conserving the total tracer content and the monotony of the tracer field, the Godunov scheme thus ensures one additional physical property of transport, namely time symmetry. These desirable properties are obtained however at the price of a strong numerical diffusion. Scheme (22) can be rewritten

$$c_i^{n+1} - c_i^n = \frac{\alpha}{2} \left(c_{i-1}^n - c_{i+1}^n\right) + \frac{\alpha}{2} \left(c_{i-1}^n - 2c_i^n + c_{i+1}^n\right) \tag{24}$$

as the sum of a second-order centred scheme and a diffusion with diffusivity $\alpha(\delta x)^2/(2\delta t) = u\delta x/2$. The retro-transport scheme (23) similarly reads

$$c^{*n}_i - c^{*n+1}_i = -\frac{\alpha}{2} \left(c^{*n+1}_{i-1} - c^{*n+1}_{i+1}\right) + \frac{\alpha}{2} \left(c^{*n+1}_{i-1} - 2c^{*n+1}_i + c^{*n+1}_{i+1}\right) \tag{25}$$

Numerical diffusion acts as a diffusion in both the direct and retro-transport computations, exactly as turbulent diffusion does. Note also that the second-order finite-difference scheme obtained by ignoring the last term on the right hand side of Eq. (24) (or Eq. (25) for retro-transport) is also time-symmetric but does not preserve monotony nor positivity of the tracer field.

Finite volume schemes of higher order, such as the Van Leer I scheme used in LMDZ, also introduce some numerical diffusion to avoid spurious oscillations. The diffusivity then generally depends on the tracer field itself (stronger where tracer gradients are strong) thus breaking the linearity with respect to the tracer concentration, and hence the very possibility of time symmetry of atmospheric transport. Implications for inversion algorithms are discussed in section 5.

(*d*) *Physical parametrisations*

For turbulent diffusion, the LMDZ model uses an implicit time integration and centred finite differences on the vertical

$$m_i \left(c_i^{n+1} - c_i^n\right) = \tilde{K}_{i+1/2} \left(c_{i+1}^{n+1} - c_i^{n+1}\right) - \tilde{K}_{i-1/2} \left(c_i^{n+1} - c_{i-1}^{n+1}\right) \qquad (26)$$

(m is unchanged by mixing) where i is now the index of vertical discretisation and $\tilde{K}_{i+1/2}$ is an estimate of $K_z \rho \delta x \delta y \delta t / \delta_z$ at the interface between layers i and $i+1$. Equation (26) is of the form (18) with $\gamma = 0$, the matrix A being given by $A_{i,i-1} = \tilde{K}_{i-1/2}$, $A_{i,i} = -\tilde{K}_{i+1/2} - \tilde{K}_{i-1/2}$, $A_{i,i+1} = \tilde{K}_{i+1/2}$, and $A_{i,j} = 0$ for $|i-j| > 1$. We have shown in Part I that the same diffusion must be used for direct and retro transport. It is easily verified that $A^T = A$, which shows that the retro-transport form and the adjoint of scheme (26) are identical.

Linear sinks also conserve the mass of air and can be discretised through a diagonal matrix $A = \mathrm{diag}(m_i \lambda_i)$ which ensures time symmetry.

The convective scheme is also time-symmetric as shown in Appendix B.

(*e*) *Full model*

With the exception of the Van Leer advection scheme, still to be discussed, all schemes used in LMDZ for transport and diffusion are thus symmetric with respect to time. Note that the tracer concentrations c_i^n and air masses m_i^n are instantaneous state variables whereas U, λ, K_z and convective mass fluxes are transfer variables between stages n and $n+1$ (or $n+1$ and n in backward mode). The verification of time symmetry for the whole model, once the symmetry is ensured for each individual step, requires that, for each time step, the sequence of the individual processes is reversed. This is not done in the standard retro-transport version of LMDZ which keeps the same order of operations (large-scale advection, linear sinks, boundary layer mixing, cumulus convection in that order) both in the direct and retro-transport mode (there is no reason to make special efforts to ensure time symmetry which is broken when Van Leer I scheme is used). This essentially explains the remaining difference between the direct and retro-transport integrations for the Godunov scheme (white circles, Fig. 2(b)). As already mentioned, when the order of the sequence is reversed in the retro-transport computation (+ signs in the same figure), the averaged FMT for the 11 stations increases from 98.2 to 99.5%.

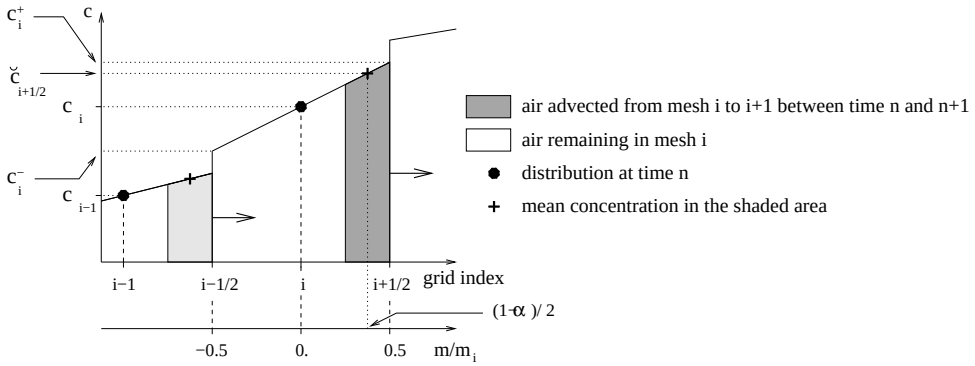

Figure 3. General principle of Van Leer I finite volume scheme and notations for the present paper. The vertical coordinate is the tracer mass concentration. The upper horizontal coordinate is the grid index, while the lower one is the mass counted from the centre of cell i and normalised by the mass m_i of that cell. The shaded areas represent the quantity of tracer which is advected through one cell interface during one time-step.

5. The case of algorithms which are not time-symmetric

When discretisation does not preserve the time symmetry of atmospheric transport, one can use either the adjoint or the retro-transport algorithm for practical applications. The retro-transport algorithm guarantees some important properties of the retro-transport equation (the same that hold for direct transport). This led us to adopt the retro-transport approach to evaluate the efficiency of the CTBT network (Hourdin and Issartel, 2000). One could hypothesise however that the adjoint algorithm, which provides the exact gradient of an objective function, should be more efficient in the case of variational assimilation. We investigate this question in a 1D configuration taking as an example the Van Leer (1977) advection scheme. We first compare various ways of computing sensitivities and then consider the case of variational assimilation. Note that Vukićević et al. (2001) also investigated the possibility of using either the adjoint of the direct code or the non-linear code used in backward mode in a 2D configuration.

(*a*) *Computation of sensitivities*

Van Leer (1977) proposed a whole hierarchy of advection schemes defined on the same general principle. First a polynomial subgrid scale tracer distribution is prescribed in each grid cell. Then the transfer of tracer between cells i and $i+1$ is computed as the product of the air mass transfer by an estimate of the average tracer concentration $\breve{c}_{i+1/2}$ in that air which is advected from cell i to $i+1$ during the time-step. That estimate is determined from the polynomial interpolation. For 1D advection on a regular grid with a non divergent wind field, the time evolution of the tracer concentration in cell i is then given by

$$c_i^{n+1} - c_i^n = \alpha \left(\breve{c}_{i-1/2}^n - \breve{c}_{i+1/2}^n \right) \qquad \text{with } \alpha = u\delta t/\delta x \tag{27}$$

For the case of an interpolated linear sub-grid scale distribution which will be considered here, with slope $(\delta c)_i$ (from which the concentration at cell boundaries

reads $c_i^{\pm} = c_i \pm (\delta c)_i/2$ as illustrated in Fig. 3), $\check{c}_{i+1/2}$ is given by

$$\check{c}_{i+1/2} = c_i + \frac{1}{2}(1-\alpha)(\delta c)_i \text{ if } \alpha > 0 \quad (28)$$

$$= c_{i+1} - \frac{1}{2}(1+\alpha)(\delta c)_{i+1} \text{ otherwise.} \quad (29)$$

In the Van Leer I scheme, used in the standard version of LMDZ, the slope is first estimated by centred finite differences. In order to ensure monotony, it is then imposed that the absolute value of the slope must be less than twice the absolute difference between c_i and either one of its neighbours. The full scheme thus reads

$$\text{if } \{ (c_{i+1} - c_i) \times (c_i - c_{i-1}) > 0 \} \text{ then} \quad (30)$$

$$(\delta c)_i = (c_{i+1} - c_{i-1})/2 \quad (31)$$

$$\text{if } \{ |(\delta c)_i| > 2|c_i - c_{i-1}| \} \text{ then } (\delta c)_i = 2(c_i - c_{i-1}) \quad (32)$$

$$\text{if } \{ |(\delta c)_i| > 2|c_{i+1} - c_i| \} \text{ then } (\delta c)_i = 2(c_{i+1} - c_i) \quad (33)$$

$$\text{else } (\delta c)_i = 0 \quad (34)$$

We show below the results produced by three different methods for computing the sensitivity of a measurement to the initial tracer distribution. The first two methods consist of backward integration of the retro-transport and adjoint models respectively. The third one consists of explicit perturbation of the initial tracer concentration, followed by integration of the full nonlinear model. That method requires one model integration per grid-point. It was verified also that direct integrations from perturbed initial states with the tangent linear model produce the same sensitivities as adjoint integration (results not shown).

For a time-symmetric linear system, the three estimates of the sensitivity should be equal. For a non time-symmetric linear system, the direct perturbation and adjoint estimates should be equal, the retro-transport estimate being different. Because of the presence of the if statements (30, 32 and 33), the Van Leer I scheme is nonlinear. As a consequence, the three estimates can be expected to be different. The retro-transport estimate depends only on the measurement function. Integration of the adjoint model requires prior storage of the results of the direct integration, which are then used for implementing the if statements in the course of the sensitivity computations. The adjoint estimate therefore depends not only on the measurement function but also on the particular direct solution in the vicinity of which the sensitivity is determined. As for the direct perturbation estimate, it depends on the initial state of the direct solution as well as on the amplitude of the initial perturbations.

Results are presented in Fig. 4. The 1D spatial domain is periodic with 60 grid-points, and the Courant number is $\alpha = 0.2$. Three different initial distributions of tracer have been used, *viz.*, total absence of tracer (flat state), a sine wave and an isolated square wave, both with amplitude one. Fig. 4(a) shows the result of direct integrations over $N=150$ time-steps, started from the sine and square waves respectively. The physically correct solution would be obtained by translating the initial field by $\alpha N = 30$ grid-points. The diffusive nature of the Van Leer I scheme is clearly visible, although the smoother sine wave is less affected.

Fig. 4(b) shows the computed sensitivity functions. The measurement is a uniform sampling over a 6-grid-point interval. The corresponding measurement function (i. e. the function μ of Eq. (1)) is shown by the thick black curve in the graphics. The sensitivity is computed with respect to tracer concentration

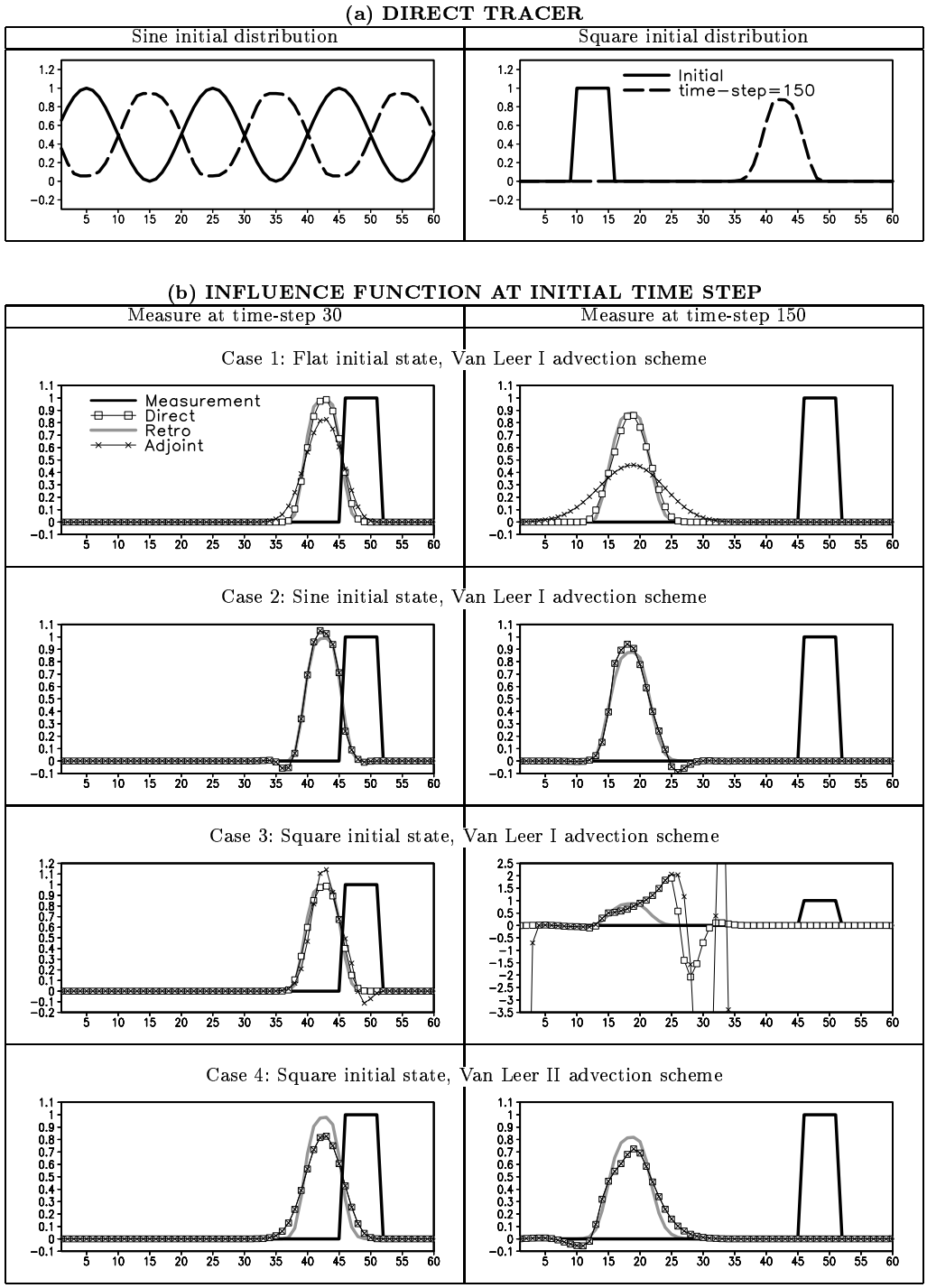

Figure 4. Computation of sensitivity functions.
Panel **a**: Direct transport. Left: sine wave. Right: isolated square wave. Full curve: initial distribution of tracer. Dashed curve: numerically computed distribution at time-step 150.
Panels **b**: Sensitivity functions at initial time for a measurement computed at time-step 30 (left column) or 150 (right column). The thick black curve, identical for all graphics, is the measurement function. The other curves show the sensitivity functions computed by the retro-transport model (thick grey curve), the adjoint model (crosses) and perturbation of the full direct model (squares). Different cases, corresponding to different initial tracer concentrations and different advection schemes, are shown from top to bottom (note the different vertical scales for case 3).

at initial time-step for a measurement computed at time-step 30 (left column) or 150 (right column). The physically correct sensitivities would be obtained by translating the measurement function by -6 and -30 grid-points respectively. For the direct estimate, a perturbation with amplitude $g = 10^{-8}$ is added in turn at each grid-point.

For case 1 (upper panels of Fig. 4(b)), corresponding to a flat initial state, condition (30) is never fulfilled. The slope is zero in the adjoint computation, which then reduces to the first-order Godunov scheme (Eqs 27 and 28 with $(\delta c) = 0$). The corresponding sensitivity function is subject to much more dissipation than the functions computed through retro-transport and explicit perturbation. The latter two, in addition to being much closer to the physically correct solution, are also mutually close. This classical situation, in which retro-plumes are computed a priori for a zero initial concentration, corresponds here to a singular limit in which the value of the sensitivity depends on arbitrary choices. Changing the $>$ sign to $\geqslant$ in Eq. (30) would lead to different values of the sensitivity.

Case 2, which corresponds to a sine wave tracer distribution, is less pathological. Condition (30) is verified at most grid-points, while slope limitation is active, at a given time-step, at only a few grid-points. The three estimates of the sensitivity function are significantly affected by numerical dissipation, but mutually close (the retro-transport sensitivity, which depends only on the measurement function, is the same as in case 1). It is to be noted however that the direct and adjoint integrations produce small, but unphysical, negative sensitivities (particularly visible in the sensitivities at retro time-step 150).

Case 3 corresponds to the isolated square wave. The retro-transport sensitivity is again the same as before. At retro time-step 30, the adjoint sensitivity starts showing spurious numerical oscillations, while the other two estimates are close to each other. At retro time-step 150, both the adjoint and direct computations exhibit totally unphysical (but different) oscillations, with large negative sensitivities. The instability of the adjoint computation can be explained as follows. After a number of time-steps in the direct simulation, the square distribution is no more a square. On both sides of the peak, the concentration is not exactly zero but decreases by typically one order of magnitude from one grid-point to the next. As a consequence, downstream of the peak, in the region where the peak of the sensitivity function is advected, condition (33) is reached everywhere. The adjoint scheme then corresponds to a direct transport scheme with $(\delta c)_i = 2(c_{i+1} - c_i)$. This scheme is distinctly pathological for two reasons: the finite difference is not centred and it corresponds to twice the slope that would be normally computed by finite differences.

To illustrate case 3 further, we show, in Fig. 5(a), direct sensitivities computed at retro time-step 150 (similar to the right panel of Fig. 4(b)) by using different values of the perturbation amplitude g ($g = 10^{-8}$ was used for all the results shown in Fig. 4). Small values of g produce results which, as can be expected, are close to the adjoint sensitivities. As g increases, the direct perturbation sensitivities tend to smooth out and to become progressively closer to the retro-transport sensitivities (and so to the physically correct sensitivities). We propose the following tentative explanation for this interesting phenomenon. Because of the presence in the code of 'if' statements associated with the slope limiters (32) and (33), each integration of the model goes through a sequence of bifurcations.

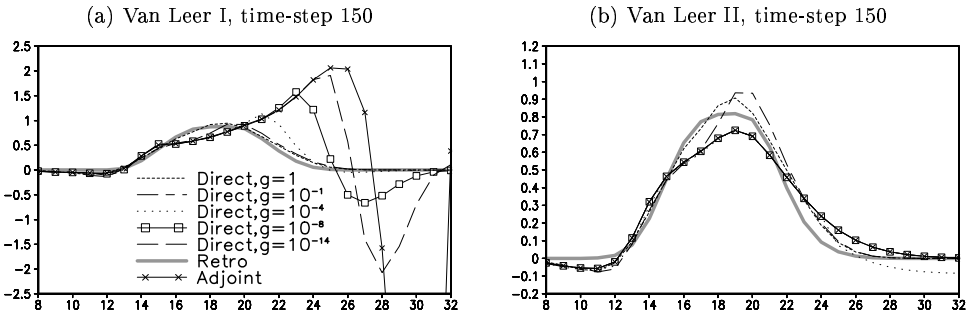

Figure 5. Computation of direct perturbation sensitivity functions with the Van Leer I (**a**) and II (**b**) advection schemes (same conditions as for the lower two panels of the right column of Fig. 4(b)). Sensitivity functions are shown for different values of the amplitude g of the initial perturbation. The curves $g = 10^{-8}$ as well as the adjoint and retro-transport sensitivity functions are also shown in Fig. 4.

For a small enough initial perturbation, that sequence of bifurcations is not modified (a small perturbation is not 'seen' by the slope limiters), and the adjoint sensitivity curve shows that the corresponding local sensitivity is very large. As the amplitude of the initial perturbation increases, the sequence of bifurcations is more and more modified, with large fluctuations in the local sensitivities. These large fluctuations tend to mutually cancel out, which is not surprising since the slope limiters are intended at maintaining a physically desirable (and, in a sense, stabilising) property of advection, *viz.*, monotony. Large initial perturbations tend to smooth out the local sensitivities, and result in smaller values of sensitivity. Note that retro-transport (which of course includes the effect of slope limiters) achieves the same goal at a much lower cost than direct perturbation.

Case 4 of Fig. 4 also corresponds to the isolated square wave, but the integration is now performed with the second of the Van Leer schemes, an elegant alternative to the first scheme in which the slope (31) is replaced by $(\delta c)_i = 2\,(c_{i+1} - c_i)\,(c_i - c_{i-1})\,/\,(c_{i+1} - c_{i-1})$. With that formulation, conditions (32) and (33) are never reached. This second scheme is slightly more diffusive than Van Leer I but its non linearities are weaker. The same effect is observed as for case 3, although much less pronounced. For $g \leqslant 0.0001$ and at the accuracy of the figure (Fig. 5(b)), direct perturbation sensitivities are indistinguishable from the adjoint-computed sensitivities. For $g = 1$ on the other hand, the direct sensitivities are closer to the retro-transport sensitivities (as with Van Leer I).

Other results (not shown) confirm that the retro-transport integration generally produces sensitivities which are close to the direct sensitivities obtained with large initial perturbations.

It is seen that in the cases when it is unambiguously defined, the adjoint gives, as it must, very accurate estimates of the numerical sensitivity of the direct transport algorithm. In strongly nonlinear systems however, those numerically exact estimates can be unphysical, with unrealistic oscillations and negative values. By comparison, the retro-transport algorithm, which is robust and preserves positivity (as well as monotony), produces sensitivity estimates which, even with the strongly nonlinear Van Leer scheme I, are sometimes more realistic and, in any case, never grossly erroneous.

(*b*) *Minimisation experiments*

The retro-transport and adjoint schemes are now compared in the possibly more testing conditions of variational assimilation, intended at reconstructing the tracer distribution from observations distributed in time. The experimental procedure used here, of the classical twin experiments type, is standard in evaluation of assimilation algorithms. A reference simulation, denoted y_i^n, is first performed over N time-steps, and taken as the truth. Synthetic observations of the form $\sum_i \mu_i y_i^n$ are then extracted from that truth by applying at each time-step n a measurement function $\mu = (\mu_i)$. For any solution c_i^n of the transport equation, the scalar objective function

$$J\,(y,\,c) = \sum_n \left[\sum_i \mu_i y_i^n - \sum_i \mu_i c_i^n \right]^2 \tag{35}$$

measures the misfit between that solution and the observations. The objective function is then minimised with respect to the tracer concentration c_i^0 at the initial time of the reference simulation.

The minimisation is performed using the code M1QN3 developed by Gilbert and Lemaréchal (1989). It is an iterative code, each step of which requires the explicit knowledge of at least an approximation of the local gradient of the objective function. It is of the quasi-Newton type, meaning that the sequence of computed local gradients is used to progressively build up an approximation of the inverse Hessian (matrix of second derivatives) of the objective function. Appropriate use of that approximate inverse Hessian makes the minimisation particularly efficient, at least if the gradient varies smoothly. Two series of experiments have been performed, in which the gradient was estimated either exactly by the adjoint model, or approximately by the retro-transport model.

Experiments are performed with the same 1D model used in the previous subsection. The length of the reference simulation, $N = 300$ time-steps, corresponds to one period of advection over the complete domain. The measurement function is the uniform local sampling represented in Fig. 4. The reference field is therefore completely sampled over the reference period, with the consequence that the minimum of the objective function (35), in addition to being equal to zero, is unique.

Results are presented in Fig. 6. The initial state of the reference is an isolated square wave, of the same width as the measurement function, and located half a period away from it. The Van Leer I scheme is used for advection. The minimisation is started from a state that is free of tracer. The minimisation is continued until the increment between two iterations becomes smaller than a prespecified threshold. Fig. 6(a) shows that, as can be expected, the minimisation leads to a smaller value of the objective function when performed with the adjoint model. However the first steps of the minimisation are faster with the retro-transport algorithm. The reconstruction of the initial concentration (shown in Fig. 6(c)) is also slightly better with the retro-transport algorithm (compare also the estimation errors in Fig. 6(b)). Both reconstructions are rather crude because measurements are available only at some distance from the source (which would be the case in many real applications). In fact, direct advection computations starting either from the exact initial state or from the initial states obtained after minimisation (shown in Fig. 6(c)) lead to almost the same tracer distribution at time-step 50 of the assimilation period (Fig. 6(d)). It is here the numerical

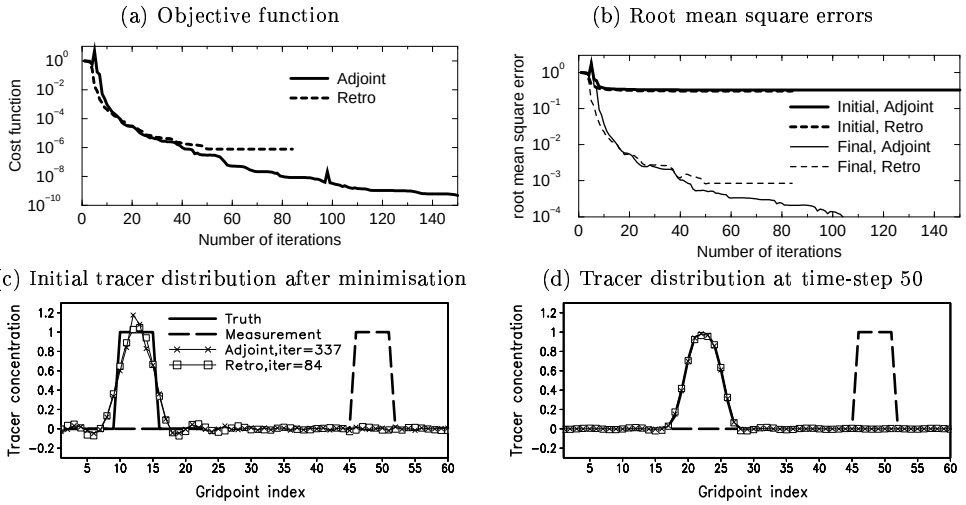

Figure 6. Results of experiments of variational assimilation of synthetic tracer observations. The initial truth to be reconstructed is represented by the full thick curve in Panel **c**. The objective function (Eq. (35)) to be minimised is computed with the measurement function represented by the thick dashed curve in panels **c** and **d**. The minimisation is started from a state free of tracer. Panel **a**: Variations of the value of the objective function with the number of iterations of the minimisation algorithm. Panel **b**: Variations of the estimation error at the initial and final times of the assimilation interval. Panel **c**: True and reconstructed (at the end of the minimisation) tracer concentrations at initial time. Panel **d**: Same thing as panel c, at time-step $n = 50$ into the assimilation interval.

diffusion that makes it difficult to reconstruct the initial tracer distribution. But a physically realistic turbulent diffusion would produce the same effect.

Other results (not shown) confirm the conclusions obtained from Fig. 6. With the retro-transport algorithm, the decrease of the objective function rapidly saturates to a relatively large value, while it most often goes on to much smaller values when using the adjoint. No significant difference is apparent however in terms of the quality of the reconstruction of the fields. The reconstruction of the truth at initial time is actually often slightly more accurate with the retro-transport algorithm. It must also be noted that there were some cases where the minimisation process did not converge when using the adjoint code, while it always converged with the retro-transport code. These results are consistent with the interpretation that has been given of the results of Fig. 5. The approximate smooth gradient allows rapid, but not perfectly accurate localisation of the minimum of the objective function. More accurate localisation requires the exact numerical gradient, which describes the small-scale variations of the objective function (Fig. 6(a)). Now, more accurate localisation does not seem to be practically useful (Fig. 6(b) to (d)). Worse than that, the exact gradient oscillates so much in some cases that the minimisation algorithm fails. The general conclusion is that, at least in the present example, an approximate but smooth gradient is preferable to an exact but rapidly oscillating one.

6. Conclusions

In Part I, a mathematical equivalence was established between Eulerian backtracking, or retro-transport, on the one hand, and adjoint transport with respect

to the air-mass-weighted scalar product on the other. Eulerian backtracking basically requires only changing the signs of a number of terms in the direct transport equation. This is of practical advantage in terms of development and maintenance of codes. The mathematical equivalence between the retro-transport and adjoint equations, or time symmetry, can however be lost through discretisation, and this second Part has been devoted to the numerical aspects of Eulerian backtracking.

It has been shown that time symmetry is ensured for the schemes used in the LMDZ model for eddy-diffusion and cumulus convection. As a general rule, ensuring time symmetry of schemes for sub-grid scale transport should not be difficult. For instance, exchange matrices used in some sophisticated mass flux schemes of cumulus convection (Emanuel, 1991), or transilient matrices of the type proposed by Stull (1984) for vertical transport in the planetary boundary layer would have to be simply replaced by their transposes.

The situation is not as simple for advection schemes. Some simple schemes such as the Godunov first order scheme or the second-order finite-difference scheme (obtained by ignoring the second term on the right-hand side of Eq. (24)) are symmetric. But both suffer from intrinsic deficiencies such as irrealistic diffusivity for the former and non positivity for the latter. Vukićević et al. (2001), who investigated the numerical properties of computations of 2D advection in the context of variational data assimilation, also found that one particular linear transport algorithm, the QUICK algorithm, is time symmetric. More sophisticated schemes, intended at guaranteeing desirable properties such as conservation of the total quantity of tracer, positivity, monotony, stability, or weak diffusivity, generally introduce non linearities, thus breaking the equivalence between backward and adjoint transport. In that case, the adjoint algorithm provides a sensitivity which depends on the direct tracer concentration and is numerically exact, but may be physically unrealistic. Retro-transport provides on the other hand a unique but approximate sensitivity, which has the distinct advantage of always being positive if the direct scheme preserves positivity. Positivity can be absolutely fundamental, for instance when combining retro-plumes from different stations in order to localise a source of pollution.

Numerical experiments performed in the simplest possible situation (1D advection) with two versions of the (nonlinear) Van Leer advection scheme strongly suggest that retro-transport, because of its robustness and its preservation of positivity, may be preferable to exact adjoint integration, even for the purpose of numerical minimisation. Note that Vukićević et al. (2001) reached similar conclusions with a different algorithm. That result is important. Concerning minimisation, experience has shown that even minor errors in the determination of the gradient can totally inhibit the minimisation. That need not be true if the approximate gradient is in some sense physically sufficiently 'sound'.

Acknowledgements

The ETEX simulations have been performed on the NEC-SX5 of the IDRIS/CNRS computer centre. The minimisation code M1QN3 was provided by J.-C. Gilbert. The graphics have been made with the user-friendly and public domain graphical package GrADS originally developed by Brian Dotty (COLA, support@grads.iges.org). The authors thank the anonymous referees for their constructive comments. A question by one of them stimulated the work described in section 5.

References

Emanuel, K. A. 1991 A scheme for representing cumulus convection in large-scale models, *J. Atmos. Sci.*, **48**, 2313–2335.

Fouquart, Y., and Bonnel, B. 1980 Computations of solar heating of the Earth's atmosphere: A new parametrization, *Contrib. Atmos. Phys.*, **53**, 35–62.

Gilbert, J.-C., and Lemaréchal, C. 1989 Some numerical experiments with variable-storage quasi-Newton algorithms, *Math. Program.*, **45**, 407–435.

Godunov, S. K. 1959 Finite-difference methods for the numerical computations of equations of gas dynamics, *Math. Sb.*, **7**, 271–290.

Graziani, C., Klug, W., and Mosca, S. 1996 Real-time long-range dispersion model evaluation of the ETEX release, Technical Report, Joint Research Center, Office for Official Publications of the European Communities.

Hauglustaine, D. A., Hourdin, F., Jourdain, L., Filiberti, M.-A., Walters, S., Lamarque, J.-F., and Holland, E. A. 2004 Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: Description and background tropospheric chemistry evaluation, *J. Geophys. Res.*, **109**, 4314-4357

Hourdin, F., and Armengaud, A. 1999 Test of a hierarchy of finite-volume schemes for transport of trace species in an atmospheric general circulation model, *Mon. Wea. Rev.*, **127**, 822–837.

Hourdin, F., and Issartel, J.-P. 2000 Sub-surface nuclear tests monitoring through the CTBT xenon network, *Geophys. Res. Lett.*, **27**, 2245–2248.

Hourdin, F., and Talagrand, O. 2005 Eulerian backtracking of atmospheric tracers: I Adjoint derivation and parametrization of subgrid-scale transport, *Q. J. R. Meteorol. Soc.*, this issue.

Jeuken, A. B. M., Siegmund, P. C., C., H. L., Feichter, J., and Bengtsson, L. 1996 On the potential of assimilating meteorological analyses in a global climate model for the purpose of model validation, *J. Geophys. Res.*, **101**, 939. 18,535–18,553.

Kasahara, A. 1977 Computational aspects of numerical models for weather prediction and climate simulation, in *Methods in computational physics*, edited by J. Chang, vol. 17, 1–66, Academic press, Inc.

Klug, W., Graziani, G., Grippa, G., Pierce, D., and Tassone, C. 1992 Evaluation of long range atmospheric transport models using environmental radioactivity data from the Chernobyl accident, *Elsevier Sciences Publishers*, 18–28.

Laval, K., Sadourny, R., and Serafini, Y. 1981 Land surface processes in a simplified general circulation model, *Geophys. Astrophys. Fluid Dyn.*, **17**, 129–150.

Le Treut, H., and Li, Z. X. 1991 Sensitivity of an atmospheric general circulation model to prescribed SST changes: Feedback effects associated with the simulation of cloud optical properties., *Climate Dynamics*, **5**, 175–187.

Lefèvre, F., Lebonnois, S., Montmessin, F., and Forget, F. 2004 Three-dimensional modeling of ozone on Mars, *J. Geophys. Res. (planets)*, **109**, 7004–7023.

Morcrette, J. J. 1984 Sur la paramétrisation du rayonnement dans les modèles de la circulation générale atmosphérique, Thèse de Doctorat d'Etat, Univ. des Sci. et Tech. de Lille, France

Rannou, P., Hourdin, F., and McKay, C. P. 2002 A wind origin for Titan's haze structure, *Nature*, **418**, 853–856.

Simmons, A. J., and Burridge, D. M. 1981 An energy and angular momentum conserving verical finite-difference scheme and hybrid vertical coordinates, *Mon. Wea. Rev.*, **109**, 758–766.

Stull, R. B. 1984 Transilient turbulence theory. Part I: The concept of eddy-mixing across finite distances, *J. Atmos. Sci.*, **41**, 3351–3367.

Talagrand, O., and Courtier, P. 1987 Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory, *Q. J. R. Meteorol. Soc.*, **113**, 1331–1328.

Tiedtke, M. 1989 A comprehensive mass flux scheme for cumulus parameterization in large-scale models, *Mon. Wea. Rev.*, **117**, 1179–1800.

Van Leer, B. 1977 Towards the ultimate conservative difference scheme : IV. A new approach to numerical convection, *J. Computational Phys.*, **23**, 276–299.

Vukićević , T., Steyskal, M., and Hecht, M. 2001 Properties of Advection Algorithms in the Context of Variational Data Assimilation, *Mon. Wea. Rev.*, **129**, 1221–1231.

A. SYMMETRY OF THE GODUNOV SCHEME FOR A DIVERGENT WIND

Let us introduce the mass transfer $U_{i+1/2}$ between grid cells i and $i+1$ and times n and $n+1$ (time integral of the total mass flux across the interface). Assuming, to fix ideas, that mass fluxes are positive, the upstream scheme reads

$$m_i^{n+1} c_i^{n+1} - m_i^n c_i^n = U_{i-1/2} c_{i-1}^n - U_{i+1/2} c_i^n \tag{A.1}$$

with

$$m_i^{n+1} - m_i^n = U_{i-1/2} - U_{i+1/2} \tag{A.2}$$

Equation (A.1) is of form (12), L being given by

$$L_{i,i} = \frac{m_i^n - U_{i+1/2}}{m_i^{n+1}} \text{ and } L_{i,i-1} = \frac{U_{i-1/2}}{m_i^{n+1}} \tag{A.3}$$

Backward upwind advection (by wind $-U$) corresponds to a matrix R with

$$R_{i,i} = \frac{m_i^{n+1} - U_{i-1/2}}{m_i^n} = \frac{m_i^n - U_{i+1/2}}{m_i^n} = \frac{m_i^{n+1}}{m_i^n} L_{i,i} \tag{A.4}$$

and

$$R_{i,i+1} = \frac{U_{i+1/2}}{m_i^n} = \frac{{m_{i+1}}^{n+1}}{m_i^n} L_{i+1,i} \tag{A.5}$$

This shows that R and L verify condition (15) and that the scheme A.1 is time symmetric.

B. SYMMETRY OF THE CONVECTION SCHEME

Here we derive the symmetry of the mass flux scheme, considering for simplicity the case of a unique updraught compensated by a slow subsidence in the environment. In the implementation retained in LMDZ, the transport terms are treated with an upstream scheme (numerical diffusion is not an issue here). The continuity equations for air and tracer in the updraught (Eqs. 10 and 8 respectively) are discretised along the vertical as

$$E_i + F_{i-1/2} = D_i + F_{i+1/2} \tag{B.6}$$

and

$$E_i c_i^n + F_{i-1/2} \hat{c}_{i-1} = \hat{c}_i \left(D_i + F_{i+1/2} \right) \tag{B.7}$$

where $E_i \simeq \hat{e} \delta z \delta t$ and $D_i \simeq \hat{d} \delta z \delta t$ are the lateral entrainment and detrainment into and from the updraught for layer i during time-step δt, and $F_{i+1/2} \simeq \hat{f} \delta t$ is the updraught mass transfer between layers i and $i+1$, with the conditions that F equals zero at the lower and upper boundaries. The time evolution of the large scale tracer concentration restricted to convective processes, obtained by combining Eq. (4) and Eq. (7) with $\mathbf{v} = 0$, $\lambda = 0$, $F_d = 0$ and $\check{f} = 0$, which reads

$$\rho \frac{\partial c}{\partial t} = - \frac{\partial \hat{f} \left(\hat{c} - c \right)}{\partial z} \tag{B.8}$$

is discretised as

$$\begin{aligned} m_i c_i^{n+1} - m_i c_i^n &= F_{i-1/2} \hat{c}_{i-1} - F_{i+1/2} \hat{c}_i + F_{i+1/2} c_{i+1}^n - F_{i-1/2} c_i^n & \text{(B.9)} \\ &= D_i \hat{c}_i - E_i c_i^n + F_{i+1/2} c_{i+1}^n - F_{i-1/2} c_i^n & \text{(B.10)} \end{aligned}$$

This is of form (16). Since the model is linear, the entries $A_{i,j}$ of the corresponding matrix A can be obtained by computing $m_i c_i^{n+1} - m_i c_i^n$ in layer i for a tracer injection restricted to layer j ($c_k^n = \delta_{k,j}$). Let us first compute the concentration in the updraught for this particular injection. Below the injection ($k < j$), $\hat{c}_k = 0$. At the layer of injection the concentration in the updraught is given by $\hat{c}_j \left(D_j + F_{j+1/2} \right) = E_j c_j^n$. Above the injection ($k > j$), $\hat{c}_k \left(D_k + F_{k+1/2} \right) = \hat{c}_{k-1} F_{k-1/2}$ and, through iteration,

$$\hat{c}_i = \prod_{k=j+1}^{i} \frac{F_{k-1/2}}{D_k + F_{k+1/2}} \hat{c}_j = \frac{\prod_{k=j+1}^{i} F_{k-1/2}}{\prod_{k=j}^{i} \left(D_k + F_{k+1/2} \right)} E_j c_j^n \ . \tag{B.11}$$

The tracer in the updraught is then detrained in the environment in layer i, yielding, for $i > j$,

$$A_{i,j} = D_i \hat{c}_i = \frac{\prod_{k=j+1}^{i} F_{k-1/2}}{\prod_{k=j}^{i} \left(D_k + F_{k+1/2}\right)} E_j \, D_i \tag{B.12}$$

In the injection layer $i = j$, a similar computation shows that

$$A_{j,j} = -F_{j-1/2} - E_j + \frac{D_j E_j}{\left(D_j + F_{j+1/2}\right)} \tag{B.13}$$

The large scale concentration is also affected by subsidence just below the layer of injection ($m_{j-1} c_{j-1}^{n+1} = F_{j-1/2}$) so that

$$A_{j-1,j} = F_{j-1/2} \quad \text{and} \quad A_{i,j} = 0 \quad \text{for} \quad i < j-1. \tag{B.14}$$

The scheme for retro-transport is obtained by inverting the roles of D and E and changing the direction of the vertical propagation. It is also of form (16), the matrix A being replaced by the matrix B obtained by performing the corresponding modifications on Eqs (B.12–B.14). This yields, for retro-injection in layer j,

$$B_{i,j} = \frac{\prod_{k=i}^{j-1} F_{k+1/2}}{\prod_{k=i}^{j} \left(E_k + F_{k-1/2}\right)} D_j E_i \quad \text{for} \quad i < j \ , \tag{B.15}$$

$$B_{j,j} = -F_{j+1/2} - D_j + \frac{E_j D_j}{\left(E_j + F_{j-1/2}\right)} \ , \tag{B.16}$$

$$B_{j+1,j} = F_{j+1/2} \quad \text{and} \quad B_{i,j} = 0 \quad \text{for} \quad i > j+1. \tag{B.17}$$

Using the continuity equation (B.6), one verifies that $B = A^T$, which shows the equivalence of the retro and adjoint transport algorithms.