Generalized Adjoint for Physical Processes with Parameterized Discontinuities. Part IV: Problems in Time Discretization

QIN XU

Naval Research Laboratory, Monterey, California

(Manuscript received 29 October 1996, in final form 24 April 1997)

ABSTRACT

It is shown analytically and graphically that when parameterized on/off switches are triggered at discrete time levels by a threshold condition in a numerical model, the model solution is not continuously dependent on the initial state. Consequently, the response function and costfunction contain small zigzag discontinuities; their gradients contain delta functions and thus are not good approximations of the original continuous gradients. The problem is caused by the traditional time discretization and cannot be solved by the conventional treatment of on/off switches. To solve the problem, the traditional time discretization is modified with the switch time determined by interpolation as a continuous function of the initial state. With this modification, the response function and costfunction become continuous in the space of the initial state and their gradients can be accurately computed by the generalized adjoint.

1. Introduction

The classic adjoint formulations were recently generalized by Xu (1996a,b; 1997) for physical processes with parameterized discontinuities. In these previous studies, time continuous models were used to illustrate how the variations of the switchs cause delta functions in the tangent linear and adjoint equations and how these delta functions impact the tangent linear and adjoint solutions during the integrations of their respective equations under various types of threshold conditions. However, it is not clear how the generalized adjoint formulations obtained for time continuous models should be applied to time discrete numerical models. In a numerical model, on/off switches are triggered traditionally at discrete time levels by a threshold condition. In this case, as will be seen in this paper, the switch times and thus the discrete solution are not continuously dependent on the initial state. Consequently, the response function and costfunction contain small zigzag discontinuities and their gradients contain delta functions. These deteriorated properties cause difficulties in the tangent linearization and adjoint minimization. The related problems will be examined in this paper.

The paper is organized as follows. The time continuous model of Xu (1996a, henceforth referred to as X96a) is reviewed in the following section, which provides benchmarks for the discrete results in the later sections. Problems caused by the traditional time discretization are illustrated in section 3. Section 4 derives discrete tangent linear and adjoint equations for the traditionally discretized model equation and the resulting equations are used to analyze the problem in the conventional treatment of on/off switches. Section 5 shows how the problem can be solved by a modified time discretization scheme in which the switch time is determined, by interpolation, as a continuous function of the initial state. The principle results are summarized with conclusions in section 6.

Corresponding author address: Dr. Qin Xu, Naval Research Laboratory, Monterey, CA 93943-5502.
E-mail: xuq@nrlmry.navy.mil

2. Review of time continuous model

a. Model equation and generalized adjoint

The time evolution of the concerned variable x is described by the following equation:

$$d_t x = F + GH(x - x_c),$$
$$x = x_0 \quad \text{at } t = 0, \tag{2.1}$$

where G is the source term due to the parameterized process, F is the source term due to other processes, $H(\ )$ is the Heaviside unit-step function (Courant and Hilbert 1962, p. 622), and x_c is the threshold value for the parameterized process. As shown in X96a, $H(x - x_c)$ can be replaced by $H(t - \tau)$ for an on switch, or by $H(\tau - t)$ for an off switch, where τ denotes the switch time. In general, F and G are functions of (x, t). For the simple examples considered in this paper, G and F are assumed to be constant and satisfying $F > -G > 0$. In this case, the solution of (2.1) for $t > 0$ is given by

$$x = x_0 + Ft + [(t - \tau)H(t - \tau) + \tau H(-\tau)]G, \tag{2.2}$$

where $\tau = (x_c - x_0)/F$. Here, the initial state is not restricted by $x_0 < x_c$ and thus τ can be negative, so (2.2) is more general than (3.4) of X96a. When the initial state is within (or outside) the range of $x_c \geq x_0 \geq x_c - FT$, the solution in (2.2) has a single on switch (or has no switch) in [0, T].

As in (3.7)–(3.8) of X96a, corresponding to (2.1) the generalized tangent linear and adjoint operators have the following forms:

$$L = d_t - H'(t - \tau)G/F, \tag{2.3a}$$

$$L^* = -d_t - H'(t - \tau)G/F, \tag{2.3b}$$

where $H'(\ )$ is the unit delta function—the derivative of $H(\ )$. In association with the two operators in (2.3a,b), the left and right continuities are specified for the tangent linear variable δx and adjoint variable δx^*, respectively. When $x_0 < x_c$ and thus $\tau > 0$, the tangent linear equation $L\delta x = 0$ can be integrated as in X96a, and the solution is $\delta x = \delta x_0 + \delta x_0 H(t - \tau)G/F$. When $x_0 > x_c$ ($\tau < 0$), the tangent linear operator reduces to $L = d_t$ and the tangent linear solution is a trivial one: $\delta x = \delta x_0$. Combining these two situations, the tangent linear solution can be written into the following unified form

$$\delta x = \delta x_0 + \delta x_0[H(t - \tau) - H(-\tau)]G/F. \tag{2.4}$$

This tangent linear solution can be verified by the linear perturbation of the solution in (2.2).

b. Response function and costfunction

For the concerned sensitivity problem, we may consider the following simple form of response function:

$$R = \int_0^T r\, dt = x(t_i), \tag{2.5}$$

where r is a function of (x, t) and chosen to be $r = xH'(t - t_i)$ with $0 < t_i \leq T$. Substituting (2.2) into (2.5) gives $R = x(t_i) = x_0 + FT + [(x_0 + Ft_i - x_c)H(x_0 + Ft_i - x_c) - (x_0 - x_c)H(x_0 - x_c)]G/F$. The sensitivity of R with respect to the initial state is measured by the gradient of R; that is,

$$\partial R/\partial x_0 = 1 + [H(x_0 + Ft_i - x_c) - H(x_0 - x_c)]G/F. \tag{2.6}$$

As in (5.6)–(5.8) of X96a, the associated adjoint problem can be formulated by $L^*\delta x^* = \partial r/\partial x = H'(t - t_i)$ and $\partial x^*(T) = 0$ or, equivalently,

$$L^*\delta x^* = 0$$
$$\delta x^*(t_i) = 1, \tag{2.7}$$

where L^* is given by (2.3b). When $x_0 < x_c$ ($\tau > 0$), the adjoint solution of (2.7) is $\delta x^*(0) = 1 + H(x_0 + Ft_i - x_c)G/F$, which is the same as that of (5.8) in X96a. When $x_0 > x_c$ ($\tau < 0$), the adjoint operator reduces to $L^* = -d_t$ and the adjoint solution is a trivial one: $\delta x^*(0) = 1$. Combining these two situations, we have $\delta x^*(0) = 1 + [H(x_0 + Ft_i - x_c) - H(x_0 - x_c)]G/F$, which is exactly the same as the sensitivity measure in (2.6). This verifies that the adjoint solution of (2.7) gives the sensitivity measure in (2.6). The response function R and gradient $\partial R/\partial x_0$ are sketched by the thick curves (for $t_i = T$) in Fig. 1.

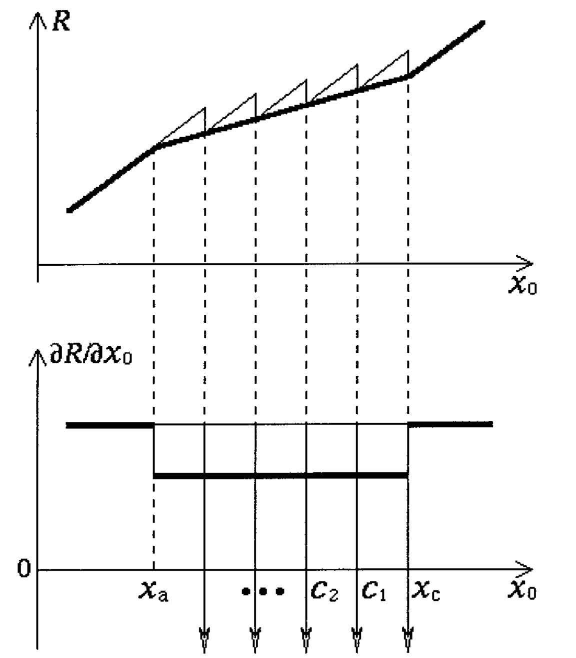

FIG. 1. Response function R (upper panel) and its gradient $\partial R/\partial x_0$ (lower panel). The continuous and discretized results are shown by the thick and thin lines, respectively. The spikes (delta functions) in the gradient are shown by thin arrows at the equally spaced discrete points $c_k \equiv x_c - k\Delta tF$ ($k = 0, 1, \cdots, N - 1$) on the x_0 axis (where $x_c = c_0$ and $x_a = c_N$).

For the concerned variational data assimilation problem, the costfunction is defined by

$$J = \int_0^T D^2\, dt, \tag{2.8}$$

where $D = x - x_{ob}$, x is the solution of (2.1), and x_{ob}, the observed value of x for the data assimilation period [0, T]. Substituting (2.4) into $\delta J = 2\int_0^T (D\delta x)\, dt$ yields the following gradient formulation:

$$\partial J/\partial x_0 = 2\int_0^T D\, dt + 2(G/F)[1 - H(-\tau)]\int_\tau^T D\, dt. \tag{2.9}$$

This formulation reduces to (3.6) of X96a when $x_0 \leq x_c$ ($\tau \geq 0$). One can verify that the gradient in (2.9) can be obtained by backward integrating the following adjoint system

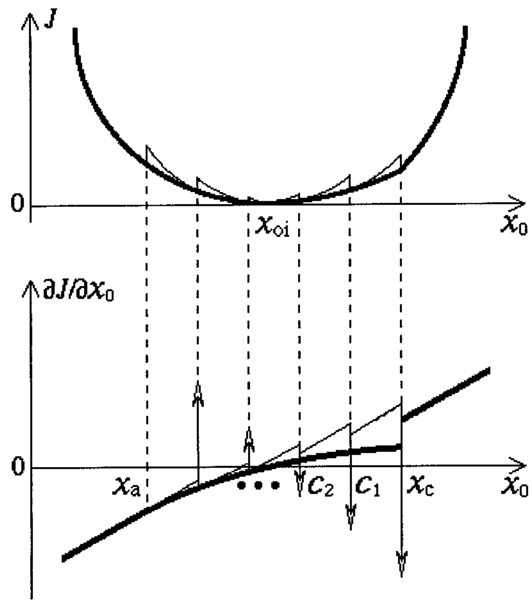

FIG. 2. As in Fig. 1 but for costfunction J and its gradient $\partial J/\partial x_0$. The observed initial state is denoted by x_{oi} on the x_0 axis.

$$L^*\delta x^* = 2D,$$
$$\delta x^*(T) = 0, \quad (2.10)$$

where L^* is given by (2.3b). By substituting (2.2) into (2.8) and (2.9), the costfunction and gradient can be expressed in terms of x_0 [see (6.7) of X96a]. The results are sketched by the thick curves in Fig. 2. The major features of these thick curves have been previously described and interpreted in X96a. These results will be called ''original'' or ''time continuous'' as they are used as benchmarks for the discrete results in the later sections.

3. Time discretization and related problems

a. Discrete model and solution

When time is discretized, an on switch is traditionally assigned to the nearest time level after the threshold condition is exceeded. With this tradition, (2.1) yields the following finite-difference equations

$$x_n = x_{n-1} + F\Delta t + G\Delta t H(x_{n-1} - x_c)$$
$$\text{for } n = 1, 2, \cdots, N, \quad (3.1a)$$

where the subscript n indicates the value at the nth time level, N is total number of time steps with $\Delta t \equiv T/N$, and the Heaviside unit-step function $H(\ )$ should be left (or right) continuous at the jump point if the threshold condition for the on switch is $x > x_c$ (or $x \geq x_c$). Explicitly, (3.1a) means the following steps of operation:

$$x_n = x_{n-1} + F\Delta t$$
$$\text{for } n = 1, 2, \cdots, m,$$
$$x_n = x_{n-1} + (F + G)\Delta t$$
$$\text{for } n = m + 1, m + 2, \cdots, N, \quad (3.1b)$$

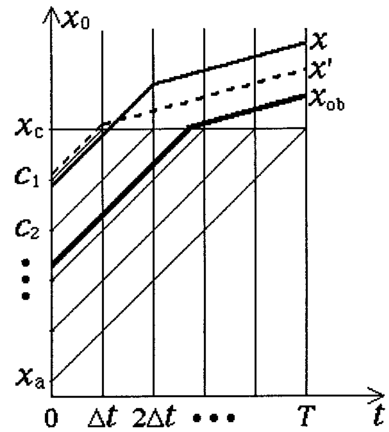

FIG. 3. Reference solution x (solid) with an initial state immediately below point c_1 on the x_0 axis, perturbed solution x' (dashed) with an initial state immediately above point c_1 on the x_0 axis, and observed true state x_{ob} (thick solid). As shown, when x_0 is perturbed to pass point c_1, the switch time jumps backward from time level 2 to 1 and the solution jumps by $G\Delta t$ at and after the second time level.

where m denotes the time level of switch and is determined by $x_m \geq x_c > x_{m-1}$, corresponding to the threshold condition $x \geq x_c$ for the on switch. [If $x > x_c$ is assumed for the on switch threshold condition, then m should be determined by $x_m > x_c \geq x_{m-1}$.]

Forward integrating (3.1a) yields the following discrete solution:

$$x_n = x_0 + nF\Delta t + G\Delta t \sum_{k=0}^{n-1} H(x_k - x_c)$$
$$\text{for } n = 1, 2, \cdots, N. \quad (3.2a)$$

Let $c_k \equiv x_c - k\Delta t F$ ($k = 0, 1, \cdots, N - 1$) denote the N equally spaced discrete points on the x_0 axis. When $x_0 = c_k$, (3.2a) gives $x_k = x_c$, which means that $\tau = k\Delta t$ and the solution reaches the threshold value exactly at the kth time level (see Fig. 3). Substituting $x_k - x_c = x_0 + k\Delta t F - x_c = x_0 - c_k$ into (3.2a) gives

$$x_n = x_0 + nF\Delta t + G\Delta t \sum_{k=0}^{n-1} H(x_0 - c_k)$$
$$\text{for } n = 1, 2, \cdots, N. \quad (3.2b)$$

When $x_0 < c_{n-1}$, (3.2b) reduces to $x_n = x_0 + nF\Delta t$. When x_0 increases and passes point c_{n-1} and thus x_{n-1} passes the threshold value x_c, the switch time jumps backward from time level n to $n - 1$ and x_n jumps by $G\Delta t$ according to (3.1). When x_0 further increases and passes point c_{n-2} and thus x_{n-2} passes x_c, the switch time jumps backward from time level $n - 1$ to $n - 2$, and x_{n-1} and x_n jump by $G\Delta t$ (see Fig. 3). As x_0 passes all the points c_k (from $k = n - 1$ to $k = 0$), x_n undergoes a sequence of jumps as shown by the summation terms in (3.2b).

b. Response function and costfunction

By choosing $t_i = i\Delta t$ in (2.5), we have $R = x(i\Delta t) = x_i$, so the response function is given by the discrete solution in (3.2b) at the time level of $n = i$. The gradient of the response function can be derived directly from $\partial(3.2\text{b})/\partial x_0$ and the result is

$$\partial R/\partial x_0 = \partial x_i/\partial x_0 = 1 + G\Delta t \sum_{k=0}^{i-1} H'(x_0 - c_k). \tag{3.3}$$

The response function and gradient are sketched by the thin curves in Fig. 1.

The costfunction in (2.8) can be discretized into the following form:

$$J = \sum_{n=0}^{N} D_n^2 \Delta t, \tag{3.4}$$

where $D_n = x_n - x_{\text{ob}}(n\Delta t)$. Substituting (3.2b) into $\partial(3.4)/\partial x_0$ gives

$$\partial J/\partial x_0 = 2D_0\Delta t + 2\Delta t \times \sum_{n=1}^{N} \left\{ D_n \left[1 + G\Delta t \sum_{k=0}^{n-1} H'(x_0 - c_k) \right] \right\}. \tag{3.5}$$

The results in (3.4) and (3.5) are sketched by the thin curves in Fig. 2.

The above results indicate that when the solution is computed by using the discrete model (3.1), x_n is not continuously dependent on the initial state x_0, so the computed response function and costfunction contain discontinuities and their gradients contain delta functions. As shown in Fig. 3, when x_0 moves between the discrete points $c_k = x_c - k\Delta tF$ $(k = 0, 1, \cdots, N - 1)$, the switch time does not change. Thus, as shown by (3.2)–(3.3) and thin curves in Fig. 1, the response function is a linear function of x_0 and its gradient is constant between the discrete points c_k. As shown by (3.4)–(3.5) and thin curves in Fig. 2, the costfunction is a parabola and its gradient is a linear function of x_0 between the discrete points c_k. However, when the initial value passes through one of the discrete points c_k, the switch time jumps by Δt, the response function and costfunction become discontinuous, and their gradients have delta functions (or, say, spikes as shown by the vertical arrows in Figs. 1 and 2). When $\Delta t \to 0$, the discrete response function and costfunction converge to their respective time-continuous counterparts in (2.5) and (2.8), but the discrete gradients in (3.3) and (3.5) are only weakly convergent (Courant and Hilbert 1962, pp. 777–778).

4. Discrete adjoint and conventional treatment

a. Discrete adjoint

The tangent linear equation for (3.1) has the following discrete form:

$$\delta x_n = \delta x_{n-1} + G\Delta t H'(x_{n-1} - x_c)\delta x_{n-1} \quad \text{for } n = 1, 2, \cdots, N. \tag{4.1}$$

Forward integrating (4.1) yields the following discrete tangent linear solution:

$$\delta x_n = \delta x_0 \left[1 + G\Delta t \sum_{k=0}^{n-1} H'(x_0 - c_k) \right], \tag{4.2}$$

where

$$x_k - x_c = x_0 - c_k \quad \text{and}$$

$$\prod_{k=0}^{n-1} [1 + G\Delta t H'(x_0 - c_k)] = \left[1 + G\Delta t \sum_{k=0}^{n-1} H'(x_0 - c_k) \right]$$

are used. This tangent linear solution can be verified by directly perturbing the solution in (3.2b).

Similar to its time-continuous counterpart in (2.7), the gradient of the response function $R = x_i$ can be given by the solution of the following discrete adjoint problem:

$$\begin{aligned} \delta x_i^* &= 1 \\ \delta x_{n-1}^* &= \delta x_n^* + G\Delta t H'(x_{n-1} - x_c)\delta x_n^* \\ &\quad \text{for } n = i, \cdots, 2, 1. \end{aligned} \tag{4.3}$$

Using the fact that $x_{n-1} - x_c = x_0 - c_{n-1}$ [see (3.2a)–(3.2b)], one can verify that the adjoint solution δx_0^* obtained by backward integrating (4.3) gives the gradient in (3.3).

Similar to its time-continuous counterpart in (2.10), the gradient of the costfunction in (3.4) can be given by the solution of the following discrete adjoint problem:

$$\begin{aligned} \delta x_{N+1}^* &= 0 \\ \delta x_{n-1}^* &= \delta x_n^* + G\Delta t H'(x_{n-1} - x_c)\delta x_n^* + 2\Delta t D_{n-1} \\ &\quad \text{for } n = N + 1, N, \cdots, 2, 1. \end{aligned} \tag{4.4}$$

One can verify that the adjoint solution δx_0^* obtained from (4.4) gives the gradient in (3.5).

b. Conventional treatment

The discrete tangent linear and adjoint operators in (4.1) and (4.3)–(4.4) contain delta functions. When these delta functions are ignored by the conventional treatment of on/off switches (see case 1 in section 3a of X96a), the tangent linear equation (4.1) reduces to $\delta x_n = \delta x_{n-1}$ and the adjoint equations (4.3) and (4.4) reduce to $\delta x_{n-1}^* = \delta x_n^*$ and $\delta x_{n-1}^* = \delta x_n^* + 2\Delta t D_n$, respectively. In this case, the response function gradient reduces to $\partial R/\partial x_0 = 1$, and the costfunction gradient reduces to

$$2 \sum_{n=0}^{N} D_n \Delta t.$$

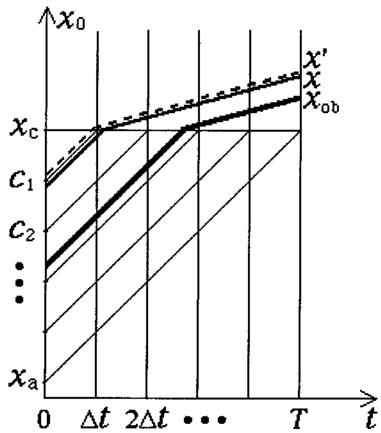

FIG. 4. As in Fig. 3 but the switch time is calculated by linear interpolation; thus, the switch time and the discrete solution vary continuously as x_0 is perturbed to pass point c_1.

The reduced gradients are clearly different from the original gradients in (3.3) and (3.5). When $\Delta t \to 0$, the reduced gradients do not converge to their respective time continuous counterparts in (2.6) and (2.9). In particular, as we can see from Figs. 1 and 2, the thin gradient curves (with the spikes eliminated) do not converge to their respective thick gradient curves. Thus, with the conventional treatment, the response function and costfunction are modified into continuous functions (not shown), but these modified functions are not valid approximations of the original response function and costfunction.

5. Modified discretization and generalized adjoint

a. Modified discretization

When time is discretized and the switch time is determined on discrete time levels, the response function and costfunction do not continuously depend on the initial state and thus cause problems in their gradient computations. The discontinuities in the response function and costfunction can be eliminated if the switch time is calculated by linear interpolation (see Fig. 4). In this case, (2.1) is discretized into the following forward model:

$$\begin{aligned} x_n &= x_{n-1} + F\Delta t \\ &\quad \text{for } n = 1, 2, \cdots, m-1 \\ x_n &= x_{n-1} + F\Delta t + G(\Delta t - \Delta\tau) \\ &\quad \text{for } n = m \\ x_n &= x_{n-1} + (F + G)\Delta t \\ &\quad \text{for } n = m+1, m+2, \cdots, N, \end{aligned} \tag{5.1}$$

where m is determined by $x_m \geq x_c > x_{m-1}$ as in (3.1), and $\Delta\tau$ is the intermediate time step determined by $x_c = x_{m-1} + F\Delta\tau$ through linear interpolation. Forward integrating (5.1) yields the following discrete solution:

$$x_n = x_0 + Fn\Delta t + G[(n\Delta t - \tau)H(n\Delta t - \tau) - \tau H(-\tau)] \quad \text{for } n = 1, 2, \cdots, N, \tag{5.2}$$

where $\tau = (x_c - x_0)/F = (m-1)\Delta t + \Delta\tau$ and $x_{m-1} = x_0 + F(m-1)\Delta t$ are used. This solution has the same value as the analytical solution (2.2) at each discrete time level. Here, since τ is a linear function of x_0, the discrete solution in (5.2) is a continuous function of the initial state.

b. Response function and costfunction

The discrete response function is given by the discrete solution in (5.2) at the time level of $n = i$; that is, $R = x(i\Delta t) = x_i$. The gradient of the response function can be derived from $\partial(5.2)/\partial x_0$; that is,

$$\begin{aligned} \partial R/\partial x_0 &= \partial x_i/\partial x_0 \\ &= 1 + [H(x_0 + Fi\Delta t - x_c) - H(x_0 - x_c)]G/F. \end{aligned} \tag{5.3}$$

This gradient is the same as the analytical result in (2.6) for $R = x(i\Delta t) = x_i$.

Since the discrete solution in (5.2) is a continuous function of x_0, the associated costfunction in (3.4) is also a continuous function of x_0. Substituting (5.2) into $\partial(3.4)/\partial x_0$ gives the gradient formulation

$$\partial J/\partial x_0 = 2\Delta t \sum_{n=0}^{N} D_n + 2\Delta t(G/F) \times [1 - H(x_0 - x_c)] \sum_{n=m}^{N} D_n, \tag{5.4}$$

where m is determined as in (5.1). As shown in (5.3) and (5.4), with the modified discretization, the response function and costfunction become continuous, and their gradients contains no delta function of x_0. When $\Delta t \to 0$, the gradients in (5.3) and (5.4) converge to their time-continuous counterparts in (2.6) and (2.9).

c. Tangent linear and adjoint operators

Corresponding to (5.1), the discrete tangent linear equation is

$$\begin{aligned} \delta x_n &= \delta x_{n-1} \\ &\quad \text{for } n = 1, 2, \cdots, N, \text{ but } n \neq m \\ \delta x_n &= \delta x_{n-1} + (G/F)\delta x_{n-1} \\ &\quad \text{for } n = m, \end{aligned} \tag{5.5}$$

where $\delta\tau = -\delta x_{m-1}/F$ is used. When $x_0 \geq x_c$ and thus $\tau \leq 0$, $m \leq 0$ and (5.5) contains only the first expression. Integrating (5.5) forward yields the following discrete tangent linear solution:

$$\delta x_n = \delta x_0 + [H(x_0 + Fn\Delta t - x_c) - H(x_0 - x_c)]\delta x_0 G/F. \tag{5.6}$$

This tangent linear solution can be verified by directly perturbing the solution in (5.2).

The gradient of the response function $R = x_i$ can be given by the solution of the following discrete adjoint problem:

$$\begin{aligned}
\delta x_i^* &= 1 \\
\delta x_{n-1}^* &= \delta x_n^* \\
&\quad \text{for } n = i, \cdots, 2, 1, \text{ but } n \neq m \\
\delta x_{n-1}^* &= \delta x_n^* + (G/F)\delta x_n^* \\
&\quad \text{for } n = m. \qquad (5.7)
\end{aligned}$$

The adjoint solution δx_0^* obtained from (5.7) gives the same gradient as in (5.3).

The gradient of the costfunction (3.4) can be given by the solution of the following discrete adjoint problem:

$$\begin{aligned}
\delta x_{N+1}^* &= 0 \\
\delta x_{n-1}^* &= \delta x_n^* + 2\Delta t D_{n-1} \\
&\quad \text{for } n = N+1, N, \cdots, 2, 1, \text{ but } n \neq m \\
\delta x_{n-1}^* &= \delta x_n^* + (G/F)\delta x_n^* + 2\Delta t D_{n-1} \\
&\quad \text{for } n = m. \qquad (5.8)
\end{aligned}$$

One can verify that the adjoint solution δx_0^* obtained from (5.8) gives the same gradient as in (5.4).

With the modified discretization, the solution in (5.2) is continuously dependent on the initial state x_0. When $\Delta t \to 0$, this discrete solution converges to the original solution in (2.2), the discretized tangent linear operator in (5.5) and associated adjoint operator in (5.7) or (5.8) converge to their respective time continuous counterparts in (2.3a) and (2.3b). The adjoint solution of (5.7) [or (5.8)] gives the gradient in (5.3) [or (5.4)], and this gradient converges to the original gradient in (2.6) [or (2.9)] as $\Delta t \to 0$.

6. Conclusions

When a numerical model contains discontinuities caused by parameterized on/off switches and is integrated numerically on discretized time levels, an on switch (or off switch) is traditionally assigned to the nearest time level after the threshold condition is (or is not) exceeded. This traditional time discretization makes the switch time and model solution not continuously dependent on the initial state. Consequently, the response function and costfunction contain small zigzag discontinuities and their gradients contain delta functions. The discontinuities in the response function imply spikes (delta functions) in the sensitivity measure, while the discontinuities in the costfunction manifest small zigzag barriers to the iterative procedures of global minimization. The related problems are examined analytically with graphical illustrations. The principle results can be summarized as follows:

- Generalized tangent linear and adjoint models can be derived for a traditionally discretized model [see (4.1) and (4.3)–(4.4)]. The derived adjoint solutions contain delta functions, just as the discrete gradients do, so they are not good approximations of the original analytical gradients.
- The problem caused by the traditional discretization cannot be solved by the conventional treatment of on/off switches. By ignoring the switch time perturbation, the conventional treatment virtually ignores the delta functions in the gradients. Thus, although the conventionally computed gradients converge as $\Delta t \to 0$, they do not converge to the original analytical gradients.
- The problem can be solved if the switch time is determined, by interpolation, as a continuous function of the initial state. With this modified discretization, the generalized tangent linear and adjoint models converge to their original time continuous counterparts as $\Delta t \to 0$.

When a parameterization scheme is designed for finite time steps, there is no continuous counterpart for the parameterized process, but the problem caused by the traditional time discretization remains essentially the same as illustrated in this paper. Since the problem is intrinsic to the traditional discretization (independent of the adjoint or any other method used for the gradient computation), we have either to modify the traditional discretization or to tolerate zigzag discontinuities in the response function or costfunction. In principle, a modification with time interpolation can be done even if the parameterization scheme is designed for finite time steps (because the interpolated time steps are still finite). The modified discretization not only solves the problem illustrated in this paper but also eliminates the on–off oscillations caused by the traditional discretization around a marginal state (see Fig. 4 of X96a).

Tangent linearization and adjoint are accurate for infinitesimal perturbations, but their practical applications consider finite perturbations with approximations. If the perturbation is finite and sufficiently larger than the change of the solution caused by one step jump of the switch time (that is, $|x_0' - x_0| \gg F\Delta t$ in Fig. 3), then a ''nonlocal'' (area-averaged) gradient may be considered for the coarse-grain geometry of the response function (or costfunction). Since zigzag discontinuities are smoothed in coarse-grain geometry, the problem caused by the traditional discretization may be solved without modifying the discretization scheme (but with reduced accuracy). Thus, the proposed modification in this paper may not be the only approach in solving the problem caused by the traditional discretization. Alternate approaches will be examined in follow-up papers.

Acknowledgments. I am thankful to Drs. Edward Barker, Roger Daley, Ronald Errico, Jidong Gao, and Hans Verlinde, and anonymous reviewers for comments that improved the presentation of the results. This work was supported by NSF Grant ATM-9417304, NOAA Contract NA67J0150, and U.S. Air Force Grant F49620-95-1-0320 to CIMMS at the University of Oklahoma, and by the Office of Naval Research (ONR) Program Element 0602435N to the Marine Meteorology Division of the Naval Research Laboratory at Monterey.

REFERENCES

Courant, R., and D. Hilbert, 1962: *Methods of Mathematical Physics.* Vol. II. Interscience, 830 pp.

Xu, Q., 1996a: Generalized adjoint for physical processes with parameterized discontinuities. Part I: Basic issues and heuristic examples. *J. Atmos. Sci.*, **53,** 1123–1142.

——, 1996b: Generalized adjoint for physical processes with parameterized discontinuities. Part II: Vector formulations and matching conditions. *J. Atmos. Sci.*, **53,** 1143–1155.

——, 1997: Generalized adjoint for physical processes with parameterized discontinuities. Part III: Mutiple threshold conditions. *J. Atmos. Sci.*, **54**, 2713–2721.