right-hand side. The superscripts "−" ("+") denote one-sided limits from the left (right). Now $J = J(u, t_a, t_b)$. Varying the new cost function [i.e., (7) modified so that the integral from $t_a$ to $t_b$ is omitted] with respect to $u$, $t_a$, and $t_b$, and assuming $0 < t_1 < t_a$; $t_b < t_2 < T$, yields the following necessary conditions for $\hat{u}$ to be the minimizer of $J$:

(i) $0 \leq t \leq t_a$

$$\dot{\hat{u}} = -\hat{u} - \sin t + W^{-1}\lambda \tag{9}$$

subject to

$$\hat{u}(0) = u_0 + V^{-1}\lambda(0), \tag{10}$$

where the "adjoint" variable $\lambda$ satisfies the adjoint or Euler–Lagrange equation

$$-\dot{\lambda} = -\lambda + w\delta(t - t_1)[u_1 - \hat{u}(t_1)]. \tag{11}$$

(ii) $t_b \leq t \leq T$

$$\dot{\hat{u}} = -\hat{u} - \sin t + W^{-1}\lambda, \tag{12}$$

where $\lambda$ satisfies

$$-\dot{\lambda} = -\lambda + w\delta(t - t_2)[u_2 - \hat{u}(t_2)] \tag{13}$$

subject to the final condition

$$\lambda(T) = 0.$$

Following integration of (8) by parts, the boundary terms can be seen to contribute to the variation in $\delta t_a$ and $\delta t_b$ through the identity $u(t_a + \delta t_a) + \delta u(t_a + \delta t_a) = C$, which, to first order, leads to

$$\delta u(t_a^-) = -\dot{u}(t_a^-)\delta t_a = [C + \sin t_a - W^{-1}\lambda(t_a^-)]\delta t_a,$$

with a similar formula for $t_b$. The requirement that $J$ be stationary with respect to variations of $t_a$ and $t_b$ leads to

$$\lambda(t_a^-)[\lambda(t_a^-) - 2W(C + \sin t_a)] = 0, \tag{14}$$

$$\lambda(t_b^+)[\lambda(t_b^+) - 2W(C + \sin t_b)] = 0. \tag{15}$$

The missing conditions needed to connect and complete (9)–(15) are continuity conditions for $\hat{u}$ at $t_a$, $t_b$:

$$\hat{u}(t_a^-) = C, \tag{16}$$

$$\hat{u}(t_b^+) = C. \tag{17}$$

One solution strategy is as follows:

$$0 \leq t \leq t_a.$$

Equations (9) and (11) must be decoupled. To this end, introduce the "representer" function $r_1 = r_1(t)$ satisfying

$$\dot{r}_1 = -r_1 + W^{-1}\alpha_1 \tag{18}$$

subject to

$$r_1(0) = V^{-1}\alpha_1(0), \tag{19}$$

where the "representer adjoint" $\alpha_1$ satisfies

$$-\dot{\alpha}_1 = -\alpha_1 + \delta(t - t_1) \tag{20}$$

subject to

$$\alpha_1(t_a) = 0. \tag{21}$$

One obvious integration strategy is

(a) form estimates $\tilde{t}_a$ and $\tilde{\lambda}_a$ for $t_a$ and $\lambda(t_a)$, respectively;

(b) integrate (20) *backward* in time, from the condition (21) at $t = \tilde{t}_a$. This determines $\alpha_1$;

(c) integrate (18) *forward* in time from the condition (19) at $t = 0$. This determines $r_1$.

Next, let $u_c$ satisfy (3), subject to the initial condition $u_c(0) = u_0$ (a forward integration). Then elementary solutions of (9) and (11) show that in the interval $[0, \tilde{t}_a]$, $\hat{u}$ is given by

$$\hat{u}(t) = u_c(t) + W^{-1}\tilde{\lambda}_a \exp(-\tilde{t}_a)\sinh(t) + V^{-1}\tilde{\lambda}_a \exp(-t - \tilde{t}_a) + Ar_1(t), \tag{22}$$

where

$$A = (R_{11} + w^{-1})^{-1}[u_1 - u_c(t_1) - W^{-1}\tilde{\lambda}_a \exp(-\tilde{t}_a) \times \sinh(t_1) - V^{-1}\tilde{\lambda}_a \exp(-t_1 - \tilde{t}_a)]$$

and $R_{11} = r_1(t_1)$. The new estimate of $t_a$, which will almost certainly not be equal to $\tilde{t}_a$, is then determined by finding the first time at which $\hat{u} = C$. Once $t_a$ is known, the left-hand side of (14), which is a polynomial in $\lambda(t_a)$ may be solved for a new $\tilde{\lambda}_a$, and the process can be repeated. Obviously $\lambda(t_a) = 0$ is a solution to (14), but we shall shortly see an example in which the nonzero root of (14) yields an admissible $\hat{u}$.

Note that, while the model, the initial condition, the observation, the inequality constraint, and the adjoint equation are all linear, the conditions (16) and (17) become highly nonlinear equations for $t_a$ and $t_b$, respectively. Note also that each of the integrations is performed in the temporal direction, which leads to a stable problem; hence, the entire calculation is well posed. In more complex cases in which analytic solutions are not available, the above algorithm could be implemented with numerical integration schemes, and an iterative scheme could be used to refine estimates of $t_a$ and $t_b$ to satisfy (16) and (17).

From the outset, we can discern some properties of the solutions of (9) and (11) subject to (10) and (14). While there is no reason to expect $\dot{\hat{u}}$ to be continuous at $t_a$ (this is because we have assumed that the convective adjustment requires no time to take place), the estimated temperature $\hat{u}$ itself must be continuous, so we must have

$$\dot{\hat{u}}(t_a^-) = -(C + \sin t_a) + W^{-1}\lambda(t_a).$$

Clearly the system in the conductive state can never be cooler than the system in the convective state; there-

fore, the temperature must be nonincreasing at the onset of convection, and nondecreasing at the transition from the convective to the conductive state. This means that we must choose for $\lambda(t_a)$ the root of (14), which results in $\dot{\hat{u}} \leq 0$. It follows that the $\lambda(t_a) = 0$ root of (14) must correspond to $t_a$ with $(C + \sin t_a) > 0$, and the $\lambda(t_a) = 2W(C + \sin t_a)$ root must correspond to $t_a$ with $(C + \sin t_a) < 0$. This means that for candidate values of $t_a$ that are less than $\sin^{-1}(-C)$, we would expect to choose the nonzero root of (14), whereas we would choose the zero root for proposed values of $t_a$ between $\sin^{-1}(-C)$ and $\pi - \sin^{-1}(-C)$. Beyond this latter value, the zero solution does not lead to admissible solutions, and the nonzero solution has the wrong sign: if $C + \sin t_a < 0$, the resulting candidate solution will cool faster than it would in the case $\lambda(t_a) = 0$, and thus reach the constraint surface before, not after, the corresponding solution with $\lambda(t_a) = 0$.

A similar solution process could be devised for the interval $(t_b, T)$. In this case, the final condition $\lambda(T) = 0$ along with the evolution equation (13) for the adjoint variable and the null solution to (15) [i.e., $\lambda(t_b) = 0$] implies that $\lambda \equiv 0$ for $t_b \leq t \leq T$ and $\hat{u}(t_2) = u_2$. For some possible values of $u_2$ and most physically realistic choices of parameters this will lead to an admissible solution $\hat{u}$ in the interval $[t_b, T]$, which adds nothing to the total cost. This solution is constructed by integrating (3) backward from $u_2$ at time $t_2$ until the solution reaches the constraint. This is an unstable calculation, since solutions to (3) grow exponentially backward in time. Small changes in the observed value $u_2$ will result in large changes in $t_b$. On the other hand, if the observed temperature $u_2$ is sufficiently warm, this procedure may lead to a solution of (3) that never reaches the constraint. This is inconsistent with the overall strategy of constructing the optimal solution from three distinct pieces, since such a solution cannot be matched to the two pieces already constructed.

This solution is also questionable on physical grounds. If we write $\lambda \equiv 0$ for $t > t_b$, then (12) implies that the original estimate of the surface heat flux [i.e., the right-hand side of (1)] in that time interval is not modified by the assimilation process. The transition from the convective to the conductive state will therefore take place at a time different from the time at which the estimated surface heat flux changes sign. This solution is certainly optimal with respect to this cost function, but does not provide a mechanism for us to improve our estimate of the heat flux based on observations.

The nonzero solution of (15) may also lead to admissible solutions in some cases. This will happen when $C + \sin t_b > 0$. If this is the case, the influence of the adjoint solution will be to warm the water, rather than cool it, and thus the estimated solution will leave the constraint surface before the corresponding solution with $\lambda(t_b) = 0$. Such a solution may have $t_b < t_a$ for the case in which $\lambda(t_a) = 0$. On the other hand, it may be possible to match this solution with the nonzero solution of (14). Like the solution with $\lambda(t_b) = 0$, this solution will have a discontinuous derivative at $t = t_b$.

Solutions to the variational problem, along with a reference solution, are shown in Fig. 2. The reference solution was computed by integrating an equation similar to (3) subject to a given initial condition from $t = 0$ until the solution reached the constraint value $C = -0.5$. This determines the "true" time of onset of convection $t_a$. The equation satisfied by the reference solution differs from (3) in the initial interval by slight changes in the amplitude and phase of the forcing function. This reflects the fact that the forcing function in (3) represents our estimate of the surface heat flux, which is subject to error. The initial condition for the reference solution also differs from (5) as noted in the figure caption. At time $t = t_a$ the convective adjustment is invoked, and the reference solution remains at the constant value $C = -0.5$ until the modified right-hand side of (3), in this case $-[C + 1.1 \sin(t + 0.2)]$, becomes positive. This determines $t_b$, following the definition of the convective adjustment in Cox (1984, 1989). From $u(t_b) = C$ the equation is then integrated to the end of the simulation at $t = 4$.

The apparently poor agreement of the optimal solution with the measurement $u_1$ is due to the small value of $w$ relative to $V$ and $W$. These parameters describe a measurement that is not very trustworthy. By construction, the optimal trajectory passes through the point $(t_2, u_2)$. As noted in the description of the construction in this section, a greater (i.e., warmer) value of $u_2$ might have rendered this construction impossible, since the minimum of the solution of (3) subject to the condition that the solution pass through the observation point may be greater than $C$ and the system will not convect.

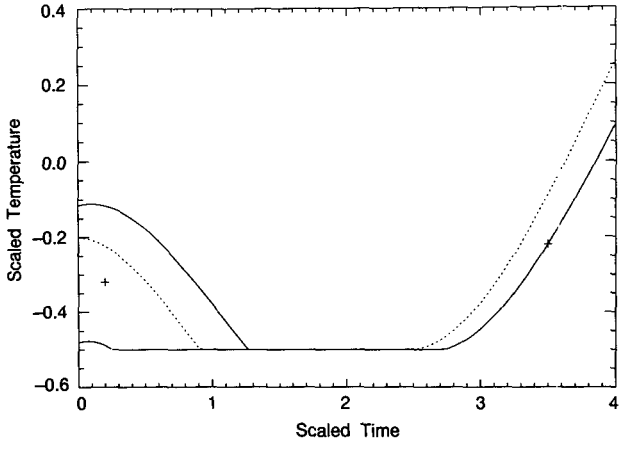


FIG. 2. Solutions to constrained optimization problem: $W = V = 100$; $w = 10$; $t_1 = 0.2$; $t_2 = 3.5$; $T = 4.0$; $C = -0.5$; $u_0 = -0.1$; $u_1 = -0.32$; and $u_2 = -0.35$. Dashed line shows "true" solution—that is, forward integration of Eq. (3)—with forcing amplitude of 1.1 and phase shifted by 0.2 rad. The reference initial condition is −0.2.

Figure 3 shows the cost function $J$, plotted as a function of $\lambda(t_a)$. The value of $t_a$ corresponding to a given $\lambda$ is calculated by setting the right-hand side of (22) equal to $C$ and solving for $t_a$ by a chord method. The nonzero root of (14) appears as an admissible cost maximum, while the desired solution is the one with $\lambda(t_a) = 0$. The kink in the curve near $\lambda(t_a) = -50$ is the point at which $t_a$ approaches $t_1$. In the small interval surrounding the cost maximum near $\lambda(t_a) = -50.65$, $t_a$ decreases to $t_1$. Beyond that point, $t_a$ increases smoothly and steadily. Each point on the curve in Fig. 3 represents a feasible candidate solution of the problem (3), (4). Most of these solutions are suboptimal, in that all except the cost minimizer and maximizer fail to satisfy (14).

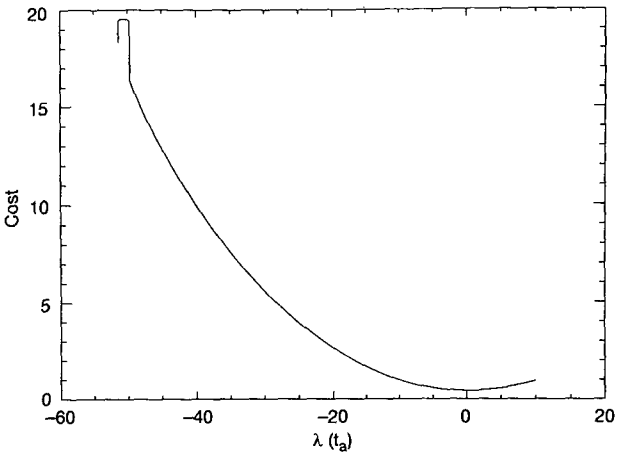


FIG. 3. Cost function $J$ vs adjoint variable $\lambda$ at time of entry to constraint boundary $t_a$ for the case shown in Fig. 2.

## 3. Alternative variational formulations

Whereas a given model, schematically represented by (3) and (4), is in general derived from physical principles, the choice of cost function depends on exactly what we want from our data assimilation system and on relative estimates of the accuracy of the data and the model. Our first choice of cost function—that is, (7) with the portion of the integral from $t_a$ to $t_b$ omitted—is a poor choice, leading as it does to nonunique solutions to the variational problem and to unstable calculations. If we have estimates $\hat{t}_a$ and $\hat{t}_b$ for the times of transition from conductive to convective states and back, we may write a new cost function:

$$J(u) = W\int_0^{t_a}(\dot{u} + u + \sin t)^2 dt + W\int_{t_b}^{T}(\dot{u} + u + \sin t)^2 dt + V[u(0) - u_0]^2 + w\sum_{j=1}^{2}[u(t_j) - u_j]^2 + \hat{V}_a(t_a - \hat{t}_a)^2 + \hat{V}_b(t_b - \hat{t}_b)^2. \quad (23)$$

Estimates of $\hat{V}_a$ and $\hat{V}_b$ could be derived from prior estimates of heat flux. The Euler–Lagrange equations for this new cost function are given by (9)–(13), with (14) and (15) replaced by

$$\lambda^2(t_a) - 2W\lambda(t_a)(C + \sin t_a) - 2W\hat{V}_a(t_a - \hat{t}_a) = 0$$

and

$$\lambda^2(t_b) - 2W\lambda(t_b)(C + \sin t_b) + 2W\hat{V}_b(t_b - \hat{t}_b) = 0.$$

We then have distinct roots for $\lambda$ at the transition points:

$$\lambda(t_a) = W(C + \sin t_a) \pm [W^2(C + \sin t_a)^2 + 2W\hat{V}_a(t_a - \hat{t}_a)]^{1/2} \quad (24)$$

$$\lambda(t_b) = W(C + \sin t_b) \pm [W^2(C + \sin t_b)^2 - 2W\hat{V}_b(t_b - \hat{t}_b)]^{1/2}. \quad (25)$$

This new choice of cost function relieves the difficulty of nonunique solutions to the variational problem in the leading interval. Substitution of the solutions of (24) into (9) shows that only the negative root of (24) leads to an admissible solution—that is, one with $\dot{\hat{u}}(t_a) \leq 0$. The negative root changes sign at $t_a = \hat{t}_a$ in the correct sense—that is, when $t_a < \hat{t}_a$, the solution must be forced down to the constraint sooner than the prior estimate—so $\lambda$ must be negative near $t = t_a$; when $t_a > \hat{t}_a$, $\lambda$ must be positive near $t_a$. Physically, this reflects the fact that if our prior estimate $\hat{t}_a$ of the time of onset of convection is later than the data would suggest, then the error in our prior estimate of the heat flux had the effect of making the water too warm, and correspondingly for the case in which $\hat{t}_a$ suggests that convection occurs sooner than it actually does.

A strategy similar to that adopted for the earlier minimization problem can be devised here. In leading interval $0 \leq t \leq t_a$ the same constructions shown in (18)–(21) may be used. In the trailing interval, even small values of $\hat{V}_b$ will result in stable calculation of $t_b$ by a process similar to that used to calculate $t_a$. Equations (12) and (13) may be decoupled by introducing the representer function $r_2 = r_2(t)$ satisfying

$$\dot{r}_2 = -r_2 + W^{-1}\alpha_2 \quad (26)$$

subject to

$$r_2(t_b) = 0. \quad (27)$$

In a fashion similar to the previous case, the representer adjoint satisfies

$$-\dot{\alpha}_2 = -\alpha_2 + \delta(t - t_2) \quad (28)$$

subject to

$$\alpha_2(T) = 0. \tag{29}$$

Clearly, the integration strategy is

(a) integrate (28) *backward* in time, from the condition (29) at $t = T$. This determines $\alpha_2$;

(b) integrate (26) *forward* in time from the condition (27) at $t = t_b$. This determines $r_2$, save for $t_b$.

Next, let $u_c$ satisfy (3), subject to condition (17) at $t = t_b$ (a forward integration). Then simple substitutions into (12) and (13) show that $\hat{u}$ is given by

$$\hat{u}(t) = u_c(t; t_b) + (R_{22} + w^{-1})^{-1}[u_2 - u_c(t_2; t_b)]r_2(t),$$

where $R_{22} = r_2(t_2)$. Finally, $t_b$ is determined by imposing (25). We must choose the negative root of (24) for $\lambda(t_a)$ and the positive root of (25) for $\lambda(t_b)$ in order to ensure feasible solutions, since we must have $\dot{u}(t_a) \leq 0$ and $\dot{u}(t_b) \geq 0$. This problem does not have the degeneracy of the one described in section 2.

With this choice of cost function, the transition time $t_b$ is not exactly the time at which the surface flux changes sign, and the solution to the cost minimization problem is again not smooth. This can be seen as a trade-off between the data misfit and error in the prior estimate of the heat flux.

Smoothness of the cost minimizer at $t_b$ could be enforced by replacing the last term in (23) with $\mu[-C - \sin t_b + W^{-1}\lambda(t_b)]$, where $\mu$ is a Lagrange multiplier. Recall that in this simple model, the prior estimate of the surface heat flux, appropriately scaled, is represented by $\sin t$. In the more general case, there will be no analytical representation of the rate of heating or cooling due to surface fluxes, and the expression $\sin t_b$ would be replaced by some general expression to be evaluated numerically, possibly coming from direct measurements.

The new Euler–Lagrange equations would consist of (9)–(12) and (24), with (25) replaced by

$$-\lambda^2 W^{-1} - \mu \cos t_b - \mu\lambda W^{-1} = 0$$

$$-C - \sin t_b + \lambda W^{-1} = 0,$$

where $\lambda$ is evaluated at $t = t_b$. The boundary term in the integration by parts of the second integral in (23) vanishes at the lower limit since $\hat{u}(t_b) = 0$ implies that $\delta u(t_b) = 0$.

This potentially attractive formulation of the variational problem appears tractable in this simple case, but generalization to more detailed and physically realistic models is complicated. Advective fluxes would have to be included in the constraint, and the temperatures of deep active layers would have to be included explicitly, as opposed to our simple formulation in which the deep water functions as a constant temperature reservoir that can supply or absorb heat indefinitely.

## 4. Optimization techniques

Whereas the general principle of constrained least-squares fitting can be adapted to the problem of data assimilation in models with convective adjustments, the method used in the above simple example would be cumbersome to apply to a large-scale ocean model in which static stability constraints are applied at each of a very large number of grid points at each time step.

One class of methods involves adding a term of the form $(1/\epsilon)(C - u)$ to the cost function whenever $u \leq C$, and seeking the limiting solution as $\epsilon$ goes to zero. These methods are referred to as penalty methods in the literature. They produce solutions that violate the constraint, if only by a small amount. This may not be acceptable in practical ocean models, which may become computationally unstable if the static stability constraint is violated. More generally, one could add a new Lagrange multiplier term to the cost function of the form $\mu(C - u)$, in which $\mu = 0$ if $u - C > 0$ and $\mu > 0$ if $u - C > 0$; see, for example, Stengel (1986, chapter 3). The constraint then appears in the adjoint equation in a natural way. The representer method can be adapted to this case, but neither approach provides an obvious way to avoid explicit calculation of the times $t_a$ and $t_b$ of entry to and exit from the convective state.

The examples considered here in which no cost is incurred by invoking the convective adjustment [i.e., the second integral in (8) is omitted from the cost function] are apparently more difficult than the case with all three terms in (8) included. Similar cases to that one were treated by Bryson and Ho (1975) and by Villalobos and Wahba (1987). Villalobos and Wahba, in particular, used a method that did not involve explicit calculation of the points of entry to and exit from the constraint surface. The analog of their method would be to impose the constraint on $\hat{u}$ at a finite number of specific times. The constraint is enforced by writing the solution as a sum of terms of the following form: the solution to the inhomogeneous problem; representers corresponding to the observations; exponential decay (i.e., basis functions for the null-space of the differential operator $L$, where $Lu = \dot{u} + u$), and representers corresponding to observations of the state at the points at which the constraint is to be imposed. This approach is similar in spirit to the introduction of bogus data (Thacker 1988). The coefficients of these latter representers can be chosen to force the solution to satisfy the constraint at the given points. The points at which the constraint was imposed form a grid in state space that can be refined. Villalobos and Wahba obtained good results for their problem with this method but did not address the issue of convergence as the grid on which the constraint was imposed was refined.

Finally, convex programming methods based on the Karush–Kuhn–Tucker theorem (see, e.g., Peressini et al. 1988) may be applicable in some cases. These meth-

ods involve searching the set of feasible points in state space—that is, those states that satisfy all constraints—for the state with minimal cost. Mackas et al. (1987) applied such a method to a linear problem with a small number of state variables. Straightforward application of these methods involves defining every point on the state function trajectory as a state variable.

This is clearly impractical for large-scale models of ocean circulation, but the rapid advances in computing may well render such calculations feasible in the near future. At present, these methods may find some application to steady-state problems such as those investigated by Tziperman and Thacker (1989).

*Acknowledgments.* This work was supported by ONR Contracts N00014-90-J-1125 (RNM), N00014-90-J-1142 (AFB) and National Science Foundation Grant OCE-8800004 (AFB). We thank Grace Wahba for bringing to our attention related work on smoothing splines. Ms. Barbara McVicar aided in the preparation of the manuscript.

## REFERENCES

Bryson, A. E., and Y.-C. Ho, 1975: *Applied Optimal Control.* Hemisphere Publishing, 481 pp.

Cox, M., 1984: A primitive equation, three-dimensional model of the ocean. GFDL Ocean Group Tech. Rep. No. 1, 143 pp.

——, 1989: An idealized model of the World Ocean. Part I: The global-scale water masses. *J. Phys. Oceanogr.*, **19,** 1730–1752.

Mackas, D. L., K. L. Denman, and A. F. Bennett, 1987: Least squares multiple tracer analysis of water mass composition. *J. Geophys. Res.*, **92,** 2907–2918.

Peressini, A. L., F. E. Sullivan, and J. J. Uhl Jr., 1988: *The Mathematics of Nonlinear Programming.* Springer-Verlag, 273 pp.

Stengel, R. F., 1986: *Stochastic Optimal Control.* John Wiley and Sons, 638 pp.

Thacker, W. C., 1988: Fitting models to inadequate data by enforcing spatial and temporal smoothness. *J. Geophys. Res.*, **93,** 10 655–10 665.

Tziperman, E., and W. C. Thacker, 1989: An optimal-control/adjoint-equations approach to studying the oceanic general circulation. *J. Phys. Oceanogr.*, **19,** 1471–1485.

Villalobos, M., and G. Wahba, 1987: Inequality-constrained multivariate smoothing with application to the estimation of posterior probabilities. *J. Amer. Stat. Assoc.*, **82,** 239–248.