Stabilization of Climate Regimes by Noise in a Simple Model of the Thermohaline Circulation

ADAM HUGH MONAHAN

Institut für Mathematik, Humboldt-Universität zu Berlin, Berlin, Germany

(Manuscript received 20 March 2001, in final form 5 November 2001)

ABSTRACT

Salinity dynamics in a simple two-box model of the thermohaline circulation (THC) is considered. The model parameterizes fluctuating eddy transport and buoyancy forcing by two independent stochastic processes. The associated stationary probability density function is calculated analytically, and its structure is analyzed in the space of the three parameters of the model. It is found that over a broad range of model parameters in which the stationary density is technically bimodal, the population of one regime is very much larger than that of the other, so the system behaves effectively unimodally. This preferential population of one regime is denoted stabilization. This phenomenon is only relevant if the timescale of THC variability is less than the mean residence times of the destabilized regime, so that the system may be described by its stationary probability density. These average residence times are calculated, and it is found that stabilization occurs over a broad range of parameter values. The stabilization phenomenon has important consequences for the stability of the THC. It is shown that the inclusion of stochastic processes in the model results in random hysteresis responses to steady changes in freshwater forcing, such that the transitions between regimes generally occur some distance away from the bifurcation points at which transitions occur in the deterministic model.

1. Introduction

The thermohaline circulation (THC) of the World Ocean is believed to play a central role in climate variability on timescales from decades to millennia, through both its internal dynamics and its response to external forcing (see, e.g., Weaver and Hughes 1992). In the Atlantic Ocean, as it operates today, the THC has the net effect of transporting a tremendous quantity of heat northward, with significant consequences for the climate of northwest Europe (Rahmstorf 1999). However, it has become clear that the present configuration of the THC may not be the only one it can take. Investigations involving both simple heuristic models (e.g., Stommel 1961; Rahmstorf 1996; Scott et al. 1999; Stone and Krasovskiy 1999; Titz et al. 2001) and complex general circulation models (GCM; e.g., Manabe and Stouffer 1988; Hughes and Weaver 1994; Rahmstorf 1995, 1996; Tziperman 2000) indicate that the THC may display multiple regimes of circulation, transitions between which are very rapid relative to the length of time spent within a regime. Furthermore, evidence of rapid shifts in the climate state abounds in the geological record, with timescales from those of glaciation cycles to higher frequency, millennial scale fluctuations (e.g., Bond et al. 1997; Alley 2000; Stocker and Marchal 2000).

Simple box models have demonstrated a remarkably good ability to reproduce the multiple regimes of circulation demonstrated in full GCMs. They are an attractive conceptual tool because of both their extremely low computational cost and the small number of parameters that govern their dynamics. Considerable attention has been paid to both purely deterministic models (e.g., Stommel 1961; Maas 1994; Rahmstorf 1996; Scott et al. 1999) and to models with a stochastic component (e.g., Stommel and Young 1993; Cessi 1994; Griffies and Tziperman 1995; Lohmann and Schneider 1999; Timmermann and Lohmann 2000). For deterministic models, the natural framework of analysis is dynamical systems theory, and discussion centers on the nature and stability of attractors admitted by the model, and on the nature and location of bifurcations. Analysis of the stochastic class of models proceeds from the theory of stochastic differential equations (Gardiner 1997). Attention in these models has primarily been focused on the nature of the spectrum and on gross features of the stationary distribution.

A study combining perspectives from both stochastic analysis and dynamical systems theory was that of Timmermann and Lohmann (2000), in which changes in the qualitative behavior of the stationary distribution of the meridional salinity gradient with changes in bifurcation

Corresponding author address: Adam Hugh Monahan, School of Earth and Ocean Sciences, University of Victoria, P.O. Box 3055 STN CSC, Victoria, BC V8P 5C2, Canada.
E-mail: monahana@uvic.ca

parameters were mapped out. In this study, we extend the analysis of Timmermann and Lohmann using a somewhat more general model. We demonstrate that for broad, and physically reasonable, regions of parameter space, although the stationary distribution is technically bimodal, the population of one regime is so much greater than that of the other that the system behaves unimodally for all intents and purposes. This *stabilization* of one regime relative to another can have significant consequences for analyses of the stability of these regimes.

Section 2 below introduces the generalization of Stommel's two-box model, from which we derive a simpler approximation in section 3. The stationary distribution of this simpler model is obtained analytically, and its structure throughout phase space is analyzed. In section 4, the average residence times of the two regimes are calculated for those regions of phase space in which the stationary distribution is bimodal. An important piece of information provided by these residence times is the region of parameter space in which the stationary distribution is relevant for understanding the finite-time behavior of the model. Section 5 applies the above results to an analysis of the hysteresis behavior of the model, and discusses the consequences of these results for discussions of climate stability. A summary and conclusions are presented in section 6.

2. The model

We begin with a slightly modified Stommel (1961) two-box model for the thermohaline circulation of the North Atlantic:

$$\frac{d}{dt}\Delta T = -\frac{|q|}{V}\Delta T - \xi\Delta T + \Gamma(\Delta T_o - \Delta T) \quad (1)$$

$$\frac{d}{dt}\Delta S = -\frac{|q|}{V}\Delta S - \xi\Delta S + \frac{S_o}{h}(P - E). \quad (2)$$

In the above, ΔT and ΔS represent, respectively, the differences between the box-average temperature and salinity of the high-latitude and middle-latitude boxes. These boxes are taken to be of the same volume V and depth h. The strength of the volume-mean flow between the boxes, q, is given by the Stommel (1961) ansatz:

$$q = c(\alpha\Delta T - \beta\Delta S). \quad (3)$$

The salinity difference is driven by a prescribed freshwater flux $P - E$, which can in principle depend on the temperature difference ΔT, and the temperature difference is relaxed on the timescale Γ^{-1} to the ''climatological'' value ΔT_o. A reference salinity S_o is introduced to convert the freshwater flux into a salinity flux.

Traditional formulations of box models assume that the temperature and salinity are well mixed within each box, so that the only transport of these quantities is by the bulk flux q. Of course, the temperature and salinity fields of the ocean display considerable spatial structure, and eddy transports (up to gyre-scale circulations) could constitute a nontrivial term in the heat and salinity budgets of each box. We parameterize the eddy transport between the boxes in terms of the mean quantities ΔT and ΔS by introducing an effective eddy diffusivity ξ. We further assume that ξ fluctuates in time, and parameterize it as a stochastic process. This process is *not* constrained to be positive, so eddy transport between the boxes may be upgradient (e.g., Nakamura and Chao 2000). In what follows, the process ξ will be assumed to be an Ornstein–Uhlenbeck process:

$$\dot{\xi} = -\frac{1}{\tau^*}\xi + \frac{\Sigma_1}{\tau^*}\dot{W}_1, \quad (4)$$

where W_1 is a Wiener process (Gardiner 1997). Note that in the limit $\tau^* \to 0$, ξ becomes the white noise process $\Sigma_1 \dot{W}_1$.

In the interests of model parsimony, we have made three more assumptions that should be noted. First, we have assumed equal eddy diffusivities for temperature and salinity; in principle, the two diffusivities should be represented by separate (although presumably not independent) stochastic processes. As well, we have assumed that ξ is Gaussian and of mean zero. Relaxing the last assumption would involve the introduction of a parameter that could tune the relative frequency of upgradient and downgradient eddy fluxes.

Finally, we will generalize the model to introduce the effects of fluctuations in the freshwater flux:

$$\frac{S_o}{h}(P - E) \to \frac{S_o}{h}(P - E) + \Sigma_2\dot{W}_2, \quad (5)$$

where W_2 is a second Wiener process, independent of W_1. More generally, we could consider fluctuations in the freshwater flux with a finite autocorrelation time. However, these fluctuations are meant to represent atmospheric phenomena that presumably evolve on much shorter timescales than those of oceanic variability. In the interest of keeping the model as simple as possible, we will consider only white noise fluctuations in the freshwater flux.

Defining the nondimensional quantities

$$\hat{t} = \frac{c\alpha\Delta T_o}{V}t \quad (6)$$

$$x = \frac{1}{\Delta T_o}\Delta T \quad (7)$$

$$y = \frac{\beta}{\alpha\Delta T_o}\Delta S \quad (8)$$

$$\eta = \frac{V}{c\alpha\Delta T_o}\xi, \quad (9)$$

we nondimensionalize the equations (1), (2), and (4) to obtain:

$$\dot{x} = -|x - y|x - \eta x + \gamma(1 - x) \tag{10}$$

$$\dot{y} = -|x - y|y - \eta y + \mu + \sigma_2 \dot{W}_2 \tag{11}$$

$$\dot{\eta} = -\frac{1}{\tau}\eta + \frac{\sigma_1}{\tau}\dot{W}_1, \tag{12}$$

where from now on all time derivatives are with respect to $\hat{t}$. In the above,

$$\gamma = \frac{V}{c\alpha\Delta T_o}\Gamma \tag{13}$$

$$\mu = \frac{\beta V S_o(P - E)}{ch(\alpha\Delta T_o)^2} \tag{14}$$

$$\sigma_2 = \left[\frac{\beta^2 V}{c(\alpha\Delta T_o)^3}\right]^{1/2}\Sigma_2 \tag{15}$$

$$\tau = \frac{c\alpha\Delta T_o}{V}\tau_* \tag{16}$$

$$\sigma_1 = \left(\frac{V}{c\alpha\Delta T_o}\right)^{1/2}\Sigma_1. \tag{17}$$

To obtain Eqs. (15) and (17), we have used the fact that

$$\frac{dW}{dt} = \left(\frac{c\alpha\Delta T_o}{V}\right)^{1/2}\frac{dW}{d\hat{t}} \tag{18}$$

for a Wiener process W. Adopting the same parameter values as Timmermann and Lohmann (2000): $\alpha = 0.15$ K^{-1}, $c = 17 \times 10^6$ m^3 s^{-1}, $V = 2 \times 10^{15}$ m^3, and $\Delta T_o = 15$ K, we find that $t = (1.66\text{ yr})\,\hat{t}$.

Equations (10)–(12) constitute a stochastic differential equation (SDE) for the three-dimensional Markov process (x, y, η) whose associated probability density function satisfies a Fokker–Planck equation (FPE; Gardiner 1997). This system is still too complicated to admit analytic solutions; in particular, the associated FPE is a partial differential equation in one time and three state variables. The following section introduces a simplification of the model that admits analytic results.

3. Simplified model

The set of equations (10)–(12) describe an SDE in $\Re^3$ with five adjustable parameters. The following pair of approximations reduces the system to a one-dimensional SDE with three parameters. First, we will assume that the temperature relaxation timescale is very short, $\gamma \gg 1$, so that x does not deviate from unity. In this limit, the equation for y becomes

$$\dot{y} = -|1 - y|y - \eta y + \mu + \sigma_2 \dot{W}_2. \tag{19}$$

Second, we will assume that the autocorrelation timescale of the stochastic process η is also very short, $\tau \to 0$, so $\eta \to \sigma_1 \dot{W}_1$. Written in differential form, the resulting SDE for y is

$$dy = (-|1 - y|y + \mu)dt - \sigma_1 y \circ dW_1 + \sigma_2 dW_2, \tag{20}$$

where the stochastic differential $y \circ dW_2$ is interpreted in the Stratonovich sense [descriptions of the difference between Ito and Stratonovich SDEs are given in Penland (1996) and in Gardiner (1997)]. In the limit that $\sigma_1 \to 0$, this model is similar to that considered by Cessi (1994).

Equation (20), with $\sigma_2 = 0$, was obtained by Timmermann and Lohmann (2000) in a different manner. They did not consider oceanic eddy transports, but assumed that the temperature dynamics was completely described by an Ornstein–Uhlenbeck process, such that in the present notation $x = 1 + \eta$. However, to obtain equation (20), Timmermann and Lohmann were obliged to arbitrarily remove the process η from within the absolute value signs. This was necessary to produce a model to which the standard techniques of stochastic analysis could be applied, in the $\tau \to 0$ limit, but cannot be rigorously justified. In fact, the limiting distribution as $\tau \to 0$ for the model in which η remains within the absolute value operator can be shown to be a trivial delta function at the origin (Peter Imkeller 2000, personal communication). The white noise limit in the original formulation of Timmermann and Lohmann does not produce physically reasonable results.

The Fokker–Planck equation for the probability density function (PDF) of the process y, $p(y)$, is given by

$$\partial_t p = -\partial_y\left[\left(-|1 - y|y + \mu + \frac{\sigma_1^2}{2}y\right)p\right] + \frac{1}{2}\partial_{yy}^2[(\sigma_1^2 y^2 + \sigma_2^2)p]. \tag{21}$$

The stationary distribution p_s satisfies $\partial_t p_s = 0$ and is given by the ordinary differential equation

$$\frac{dp_s}{p_s} = 2\left(\frac{-|1 - y|y + \mu - \frac{1}{2}\sigma_1^2 y}{\sigma_1^2 y^2 + \sigma_2^2}\right)dy, \tag{22}$$

which when integrated gives for $y < 1$:

$$p_s(y) = N\exp\left\{\frac{-2}{\sigma_1^2}\left[-y + 1 + \left(\frac{\sigma_2}{\sigma_1} - \mu\frac{\sigma_1}{\sigma_2}\right) \times \left[\tan^{-1}\left(\frac{\sigma_1}{\sigma_2}y\right) - \tan^{-1}\left(\frac{\sigma_1}{\sigma_2}\right)\right] + \frac{\sigma_1^2 + 2}{4}\ln\left(\frac{\sigma_1^2 y^2 + \sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right)\right]\right\}, \tag{23}$$

and for $y > 1$:

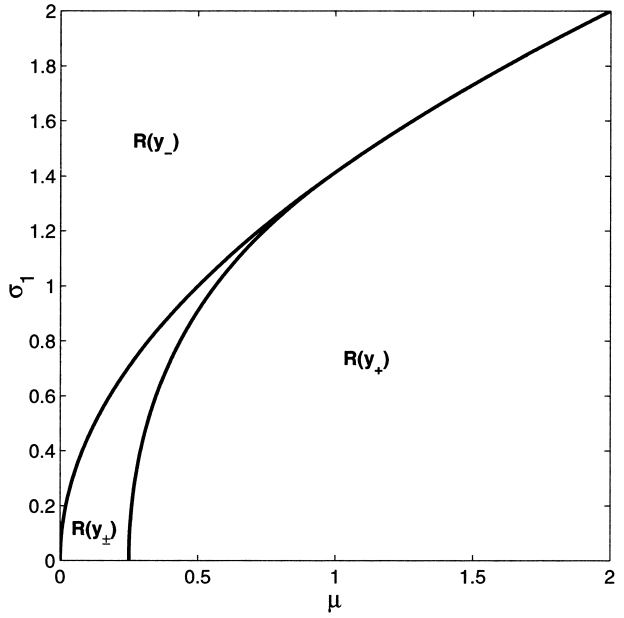

FIG. 1. Phase diagram of the number of peaks in the stationary distribution (23)–(24). In the regions $R(y_-)$ and $R(y_+)$ only the peaks at y_- and y_+ exist, respectively. In the region $R(y_\pm)$, the distribution is bimodal.

$$p_s(y) = N \exp\left\{\frac{-2}{\sigma_1^2}\left[y - 1 - \left(\frac{\sigma_2}{\sigma_1} + \mu\frac{\sigma_1}{\sigma_2}\right) \times \left[\tan^{-1}\left(\frac{\sigma_1}{\sigma_2}y\right) - \tan^{-1}\left(\frac{\sigma_1}{\sigma_2}\right)\right] + \frac{\sigma_1^2 - 2}{4}\ln\left(\frac{\sigma_1^2y^2 + \sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right)\right]\right\}. \quad (24)$$

The quantity N is a normalization constant. In the limit $\sigma_2 \to 0$, p_s reduces to the stationary distribution calculated by Timmermann and Lohmann (2000).

Having calculated the stationary distribution p_s, we now systematically investigate its behavior in the space of parameters μ, σ_1, and σ_2. We first note that, just as for a range of parameter values the deterministic Stommel model admits two stable fixed points, p_s is bimodal in a certain domain of parameter space. Extrema of p_s occur where its derivative with respect to y vanishes; from (22), this occurs when

$$0 = -|1 - y|y + \mu - \frac{1}{2}\sigma_1^2 y. \quad (25)$$

Note that σ_2 plays no role in determining the number or location of extrema in the stationary distribution p_s, as the noise process with which it is associated enters equation (20) additively. Equation (25) has three solutions for μ, σ_1 satisfying the inequalities:

$$-2 + 4\sqrt{\mu} < \sigma_1^2 < 2\mu < 2. \quad (26)$$

In this region, denoted $R(y_\pm)$, there are maxima at

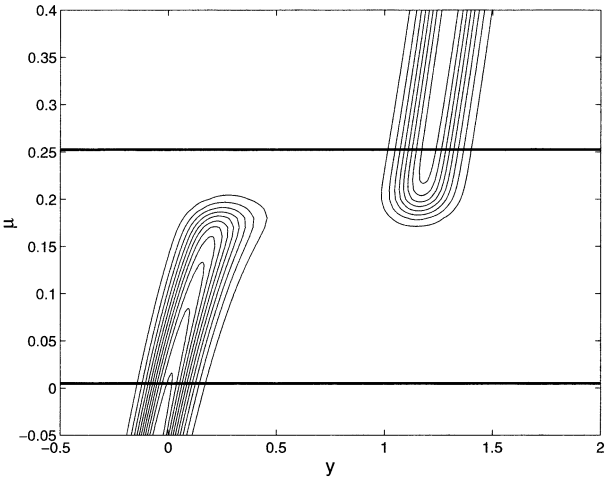

FIG. 2. Contour plot of stationary distribution $p_s(y)$ as a function of μ for $\sigma_1 = 0.1$, $\sigma_2 = 0.1$. Contours: 0.5, 1, 1.5, . . . , 6. The thick horizontal lines denote the boundaries between which p_s has two maxima.

$$y_- = \frac{1}{2} + \frac{\sigma_1^2}{4} - \sqrt{\left(\frac{1}{2} + \frac{\sigma_1^2}{4}\right)^2 - \mu} \quad (27)$$

$$y_+ = \frac{1}{2} - \frac{\sigma_1^2}{4} + \sqrt{\left(\frac{1}{2} + \frac{\sigma_1^2}{4}\right)^2 - \mu}, \quad (28)$$

and a minimum at

$$y_o = \frac{1}{2} + \frac{\sigma_1^2}{4} + \sqrt{\left(\frac{1}{2} + \frac{\sigma_1^2}{4}\right)^2 - \mu}. \quad (29)$$

Outside of $R(y_\pm)$, the stationary distribution is unimodal. In the regions $R(y_-)$ and $R(y_+)$ only the peaks at y_- and y_+ exist, respectively. These three regions are illustrated in Fig. 1. Note that along the line $2\mu = \sigma_1^2$ for $\sigma_1 > \sqrt{2}$, the points y_+ and y_- coincide with value 1.

Timmermann and Lohmann (2000) concentrated on the behavior of $p_s(y)$ for $\sigma_2 = 0$, $\mu = 0.25$, and $\sigma_1 > 0.3$. A calculation of the average rate at which the mean flow q changes sign (not shown) demonstrates that for σ_1 above 0.3, for most of the range of μ, sign changes occur on centennial or shorter timescales. Such rapid fluctuations in the sign of q are physically unreasonable. Consequently, in this paper, we will concentrate on the lower noise limit, $\sigma_1 < 0.2$. Similarly, for the fluctuations in freshwater flux, we will concentrate on the range $\sigma_2 < 0.2$.

Within the parameter region $R(y_\pm)$, the distribution $p_s(y)$ is bimodal. For the range of σ_1, σ_2, we are considering, however, throughout much of $R(y_\pm)$ the relative populations of each regime are very different. Figure 2 displays a plot of $p_s(y)$ as a function of μ for $\sigma_1 = 0.1$, $\sigma_2 = 0.1$. The thick horizontal lines in Fig. 2 correspond to the values of μ between which p_s has two maxima. It is clear that over most of the range for which p_s is bimodal, only one of the peaks is significantly

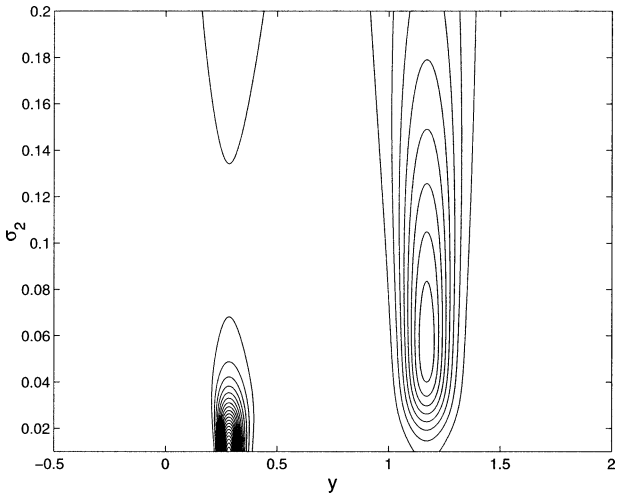

FIG. 3. Contour plot of stationary distribution $p_s(y)$ as a function of σ_2 for $\mu = 0.205$, $\sigma_1 = 0.1$. Contours: 0.5, 1, 1.5, . . . , 10.5.

populated. The stationary density is only effectively bimodal for a small range of μ near the value 0.19. At this pair of noise strengths (σ_1, σ_2), for most values of μ, the stationary behavior of the system will be more like that of a system with a unimodal distribution than with a bimodal. We will use the term *stabilization* to denote the preferred population, reflected in the stationary PDF, of one regime relative to the other. Of course, in the presence of noise, neither regime is strictly speaking stable; they are both metastable. However, a substantial asymmetry in the relative populations of the two regimes reflects a distinct preference of the system for one regime over the other; it is this phenomenon that we denote by the term stabilization.

Figure 3 contours $p_s(y)$ as a function of σ_2 for $\mu = 0.205$ and $\sigma_1 = 0.1$. Note that while σ_2 does not affect the positions of the extrema of p_s, as discussed above, it does affect the relative population of the two regimes, and in a complicated manner. For the smallest values of σ_2, the regime around y_- is more populated than that around y_+. As σ_2 increases, the mass shifts to the regime around y_+. However, as σ_2 increases further, the population of the regime around y_- begins to increase again.

The nontrivial dependence of the stationary PDF, p_s, on the parameters $(\mu, \sigma_1, \sigma_2)$ is succinctly illustrated through consideration of the first two moments of the distribution. Figures 4a and 4b contour, respectively, the mean and standard deviation of p_s as a function of σ_1 and σ_2 for $\mu = 0.19$. For these parameter values $y_- \simeq 0.25$ and $y_+ \simeq 1.16$. It is clear from Figs. 4a and 4b that for small values of σ_1, the regime around y_+ is favored, for most values of σ_2 considered. Conversely, for small values of σ_2, it is the regime around y_- that is more populated for most of the range of values of σ_1 under consideration. It is interesting to note the counterintuitive result that there are regions in parameter space in which increasing either σ_1, σ_2, or both, can actually lead to a *decrease* in the size of the fluctuations of y, as one regime becomes stabilized.

Figure 5 contours the mean and standard deviation of y for $\mu = 0.15$. For this value of μ, it is clear that for the range of σ_1, σ_2 considered, the PDF remains effectively unimodal around y_-. As well, the mean and variance of y are much more sensitive to σ_2 than they are to the value of σ_1. The dependence of the mean and standard deviation of y on σ_1 and σ_2 for $\mu = 0.23$ is displayed in Fig. 6. Here, for a broad range of values of σ_1 and σ_2, the distribution is effectively unimodal and concentrated around y_+. In contrast to the behavior at $\mu = 0.15$, the mean and standard deviation are more or less equally dependent on the values of σ_1 and σ_2.

The general picture emerges that for μ below (above) a certain window, the population of the regime around y_- (y_+) is stabilized. Within this window, the populations of the two regimes are of comparable size. We can characterize the location of this window of effective bimodality by defining the surface $\mu_{0.5}(\sigma_1, \sigma_2)$ as the value of μ, for a given (σ_1, σ_2), for which the stationary probability of y being in either regime is equal; that is,

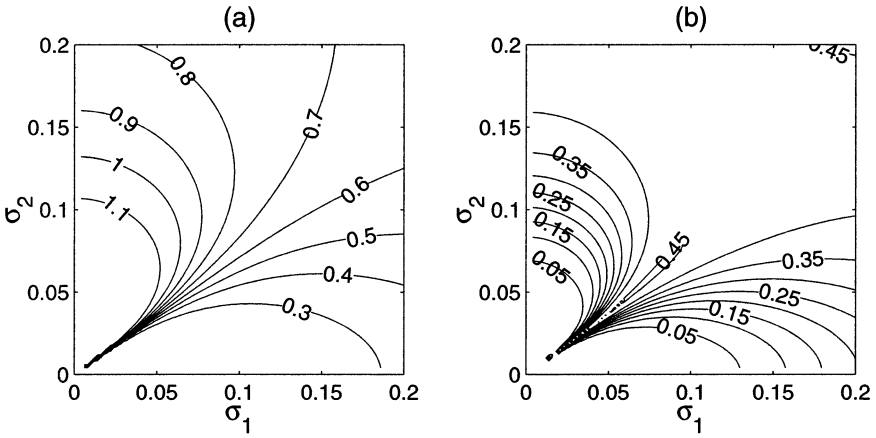

FIG. 4. Contour plot of (a) $E(y)$ and (b) $\{E[(y - E(y))]^2\}^{1/2}$, where $E(\cdot)$ denotes the expectation operator with respect to p_s, for $\mu = 0.19$. For these parameter values, $y_- \simeq 0.25$ and $y_+ \simeq 1.16$.

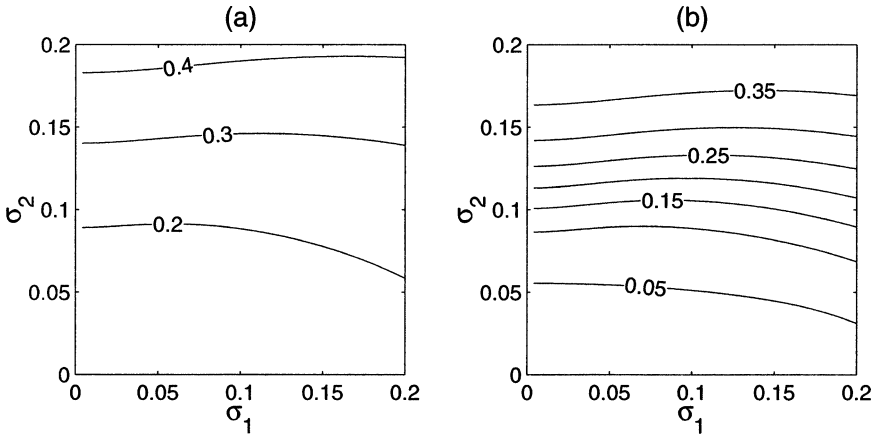

FIG. 5. As in Fig. 4 but for $\mu = 0.15$. For these parameter values, $y_- \simeq 0.18$ and $y_+ \simeq 1.13$.

$$\int_{-\infty}^{y_o} dy\, p_s(y;\, \sigma_1,\, \sigma_2,\, \mu_{0.5}) = \int_{y_o}^{\infty} dy\, p_s(y;\, \sigma_1,\, \sigma_2,\, \mu_{0.5}). \tag{30}$$

This surface is contoured as a function of σ_1 and σ_2 in Fig. 7a. For the range of σ_1, σ_2 considered, $\mu_{0.5}$ falls between 0.17 and 0.22. For $\mu < \mu_{0.5}$, the regime centered about y_- will dominate that about y_+, and conversely for $\mu > \mu_{0.5}$. Figure 7b characterizes the width of the window around $\mu_{0.5}$ within which the probabilities of being in either regime differ by less than a factor of 10. Note that this window is not symmetric about $\mu_{0.5}$. Clearly, the smaller the noise level, the sharper is the change in p_s as μ passes through $\mu_{0.5}$. Even for intermediate noise strength, the change may be rather abrupt, as is demonstrated in Fig. 2.

In summary, for the range of noise strengths σ_1, σ_2 under consideration, there exists a range of values of the parameter μ for which the regimes centered around y_- and y_+ are both significantly populated, as described by the stationary PDF p_s, and the distribution is effectively bimodal. As the noise strengths decrease, this window shrinks to a point. For values of μ below this window, almost all of the mass of p_s is concentrated around y_-, and the stationary PDF is effectively unimodal. For μ above this window, the stationary PDF is also essentially unimodal, with all of its mass concentrated around y_+. The phenomenon of stabilization provides an example of the manner in which fluctuations can be important in determining the mean state of a nonlinear stochastic system.

The phenomenon of stabilization is easily understood in the case of a 1D system moving in a two-well potential under the influence of additive noise:

$$\dot{x} = -\partial_x V + \sigma \dot{W}, \tag{31}$$

for which the stationary PDF is

$$p_s(y) = N \exp\left(-\frac{V}{2\sigma^2}\right). \tag{32}$$

If the potential has minima at x_- and x_+, then the PDF will have peaks at these points. The ratio of the peaks

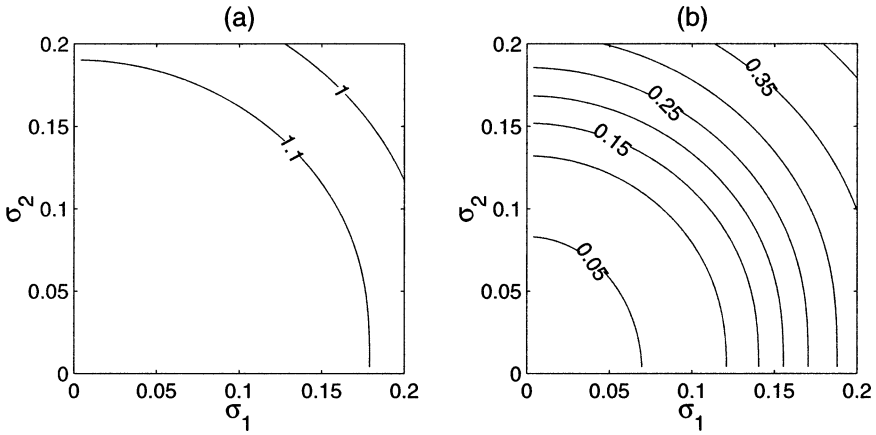

FIG. 6. As in Fig. 4 but for $\mu = 0.23$. For these parameter values, $y_- \simeq 0.36$ and $y_+ \simeq 1.19$.

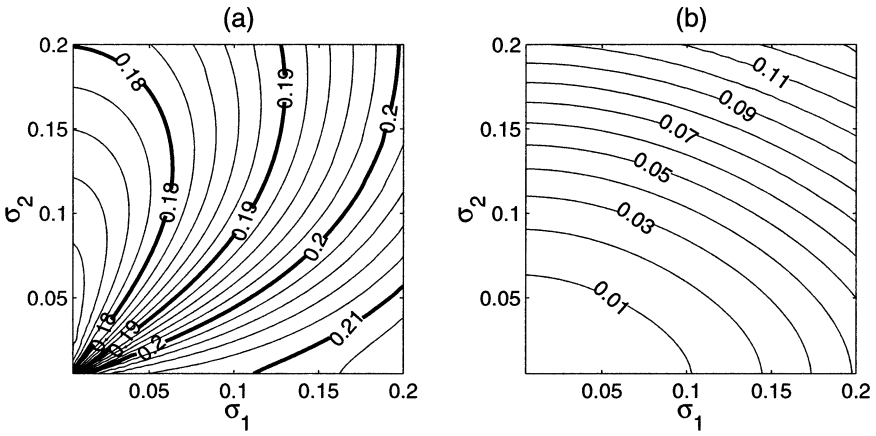

FIG. 7. (a) Contour plot of $\mu_{0.5}$, the value of μ such that the system is equally likely to be in either regime, as a function of σ_1 and σ_2. (b) Contour plot of the window in μ around $\mu_{0.5}$ for which the ratio of the stationary probabilities of being in one regime is less than a factor of 10 of being in the other.

of the PDF is determined by the difference in depth of the two potential wells:

$$\frac{p_s(x_+)}{p_s(x_-)} = \exp\left[\frac{V(x_-) - V(x_+)}{2\sigma^2}\right]. \tag{33}$$

For a fixed difference in well depth, the difference in populations between regimes is amplified as σ decreases, as was noted by Molteni and Tibaldi (1990) and Gardiner (1997).

Stabilization is clearly not limited to potential systems driven by additive noise; the model (20) is not driven just by additive white noise, so the above argument does not apply. Furthermore, numerical investigations (not shown) involving finite noise autocorrelation time τ, or temperature relaxation timescale γ^{-1}, for which the deterministic part of the dynamics does not correspond to the gradient of a potential, indicate that the phenomenon of stabilization is qualitatively unchanged.

So far we have considered only the stationary PDF p_s, which, starting from the initial condition $y = y(0)$ at $t = 0$, describes the distribution of y as $t \to \infty$. For finite times, the stationary distribution may not provide a complete description of the system, and must be supplemented by a consideration of the residence times within each of the regimes. This is discussed in the following section.

4. Residence times and the deterministic limit

In the limit that σ_1 and σ_2 become very small, the behavior of the stochastic model (20) should converge in some sense to that of the deterministic model

$$\dot{y} = -|1 - y|y + \mu. \tag{34}$$

This equation has two stable and one unstable fixed points for $0 < \mu < 0.25$, whose positions are given, respectively, by Eqs. (27)–(29) with $\sigma_1 = 0$. In a deterministic system with multiple stable fixed points, the asymptotic state of the system as $t \to \infty$ is determined entirely by the basin of attraction in which the initial conditions lie. For this simple system, the boundary between the basins of attraction of y_- and y_+ is the unstable fixed point y_0.

However, in the stochastic system, the asymptotic behavior is described by the stationary PDF, independent of the initial conditions. We have seen above that for small σ_1 and σ_2, within the range of values of μ for which the distribution is bimodal, there is only a narrow window in which the populations of the two regimes are of comparable size. In the limit that σ_1 and σ_2 go to zero, this window, centered on $\mu_{0.5}$, shrinks to a point: for $\mu < \mu_{0.5}$, the entire mass of p_s is concentrated around y_-, and for $\mu > \mu_{0.5}$ around y_+, independent of the initial conditions. Thus, the $t \to \infty$ description of the stochastic system for small σ_1 and σ_2 is completely different than that of the deterministic system. Compounding the problem is that the limiting value of $\mu_{0.5}$ for vanishing noise strengths is not unique, but depends on how the limit is taken. In particular,

$$\lim_{\sigma_1 \to 0} \lim_{\sigma_2 \to 0} \mu_{0.5} = 0.2086, \tag{35}$$

while

$$\lim_{\sigma_2 \to 0} \lim_{\sigma_1 \to 0} \mu_{0.5} = 0.1703. \tag{36}$$

The key to reconciling the stochastic and deterministic descriptions of the system in the limiting case lies in the consideration of the residence times. Denoting by $\langle T(y_-)\rangle$ and $\langle T(y_+)\rangle$ the respective average residence times of the system in the regimes around y_- and y_+, a standard result yields

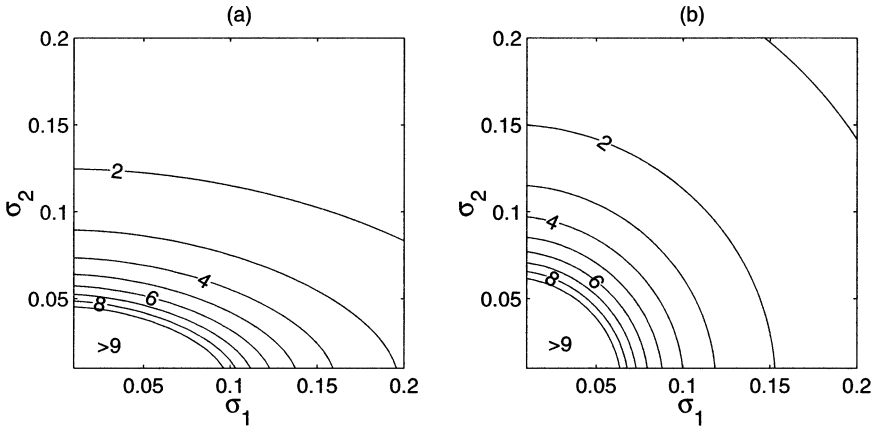

FIG. 8. Base 10 logarithm of mean residence times (a) $\langle T(y_-)\rangle$, (b) $\langle T(y_+)\rangle$, in nondimensional time units, as a function of σ_1 and σ_2 for $\mu = 0.19$.

$$\langle T(y_-)\rangle = \int_{y_-}^{y_o} \frac{dx}{p_s(x)B(x)} \int_{-\infty}^{x} dz\, p_s(z) \tag{37}$$

$$\langle T(y_+)\rangle = \int_{y_o}^{y_+} \frac{dx}{p_s(x)B(x)} \int_{x}^{-\infty} dz\, p_s(z), \tag{38}$$

where

$$B(y) = \sigma_1^2 y^2 + \sigma_2^2 \tag{39}$$

(Gardiner 1997). For the physical problem at hand, namely, the dynamics of the thermohaline circulation, there is a characteristic timescale $\tilde{T}$ such that if the average escape time for a regime is much greater than $\tilde{T}$, and if the initial conditions lie in this regime, then the system will for all practical considerations remain in this regime. If the average residence times of both regimes are much larger than $\tilde{T}$, then, practically speaking, the system will for the vast majority of realizations remain within the domain of attraction in which it started and the asymptotic behavior of the system will be described by only that part of the stationary distribution within this regime. The entire stationary distribution p_s is of relevance only if the system has time enough to sample the complete range of state space to which it theoretically has access. This will not be the case if the average escape times are much longer than any physically relevant timescale. As σ_1 and σ_2 go to zero, the average residence times of both regimes become infinite. It is in this sense that the deterministic limit arises as the noise strength vanishes. This can be expressed formally by the inequality:

$$\lim_{t\to\infty}\lim_{\text{noise}\to 0}(\text{dynamics}) \neq \lim_{\text{noise}\to 0}\lim_{t\to\infty}(\text{dynamics}). \tag{40}$$

It is the first pair of limits that yields the deterministic limit.

Figures 8–10 contour $\langle T(y_-)\rangle$ and $\langle T(y_+)\rangle$ (in nondimensional time units) as functions of σ_1 and σ_2 for $\mu = 0.19$, $\mu = 0.15$, and $\mu = 0.23$, respectively. Remembering that one nondimensional time unit corresponds to 1.66 yr, these figures may be essentially read as being in years. For millennial-scale variability of the THC, reasonable values of $\tilde{T}$ are on the order of 10^4–10^5 yr. For $\mu = 0.19$, there are broad regions of (σ_1, σ_2) space for which $\langle T(y_-)\rangle$ and $\langle T(y_+)\rangle$ are of about the same size, and both are less than $\tilde{T}$. Within such

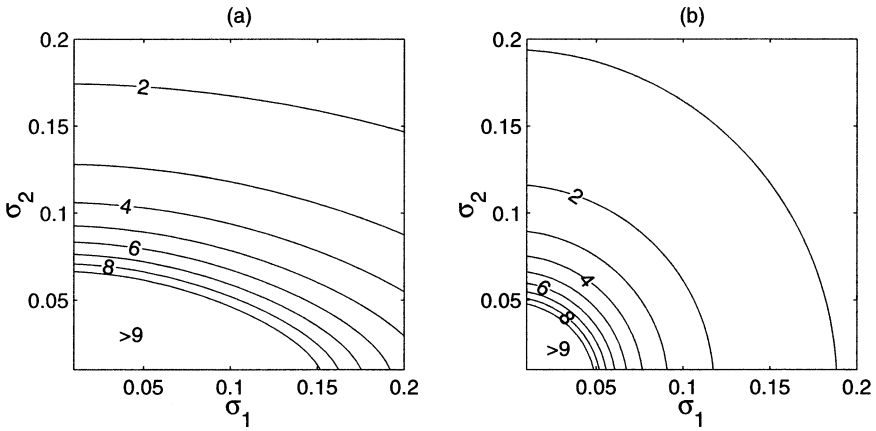

FIG. 9. As in Fig. 8 but for $\mu = 0.15$.

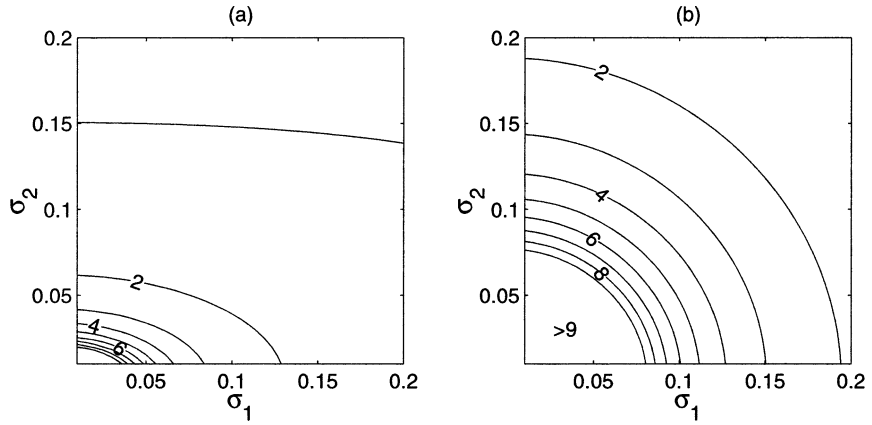

FIG. 10. As in Fig. 8 but for $\mu = 0.23$.

regions, the distribution of y will be realizably bimodal, independent of the initial conditions, and consistent with p_s. In other parts of (σ_1, σ_2) space, while both residence times are less than $\tilde{T}$, one is very much larger than the other. In these regions, the system will behave in an effectively unimodal fashion, as described by the stationary PDF, and independent of initial conditions. Finally, for smaller values of σ_1 and σ_2, the residence times of both regimes are much larger than $\tilde{T}$. For these parameter values, the stationary PDF (23)–(24) will be essentially irrelevant, and the system will fluctuate for all time (practically speaking) within the basin of attraction in which it started.

For $\mu = 0.15$ (Fig. 9), $\langle T(y_-)\rangle$ is everywhere larger than $\langle T(y_+)\rangle$, usually by at least an order of magnitude. For sufficiently large σ_1, σ_2, $\langle T(y_-)\rangle$ is less than $\tilde{T}$, and transitions to the regime centered around y_+ will occur. However, the time spent near y_+ will be typically much less than that spent near y_-, so at finite times the regime around y_- will be stabilized relative to that around y_+ at these noise levels. For $\mu = 0.23$, as is demonstrated in Fig. 10, $\langle T(y_-)\rangle$ is everywhere much smaller than $\langle T(y_+)\rangle$, and it is the regime around y_+ that is stabilized relative to that around y_- for sufficiently large σ_1, σ_2. For both $\mu = 0.15$ and $\mu = 0.23$, for small enough σ_1, σ_2, $\langle T(y_-)\rangle$ and $\langle T(y_+)\rangle$ are both much greater than $\tilde{T}$ and the description of the system by the stationary distribution will not be relevant for finite times.

Thus, for a broad range of values of σ_1, σ_2, and μ, the average residence times $\langle T(y_-)\rangle$ and $\langle T(y_+)\rangle$ are sufficiently different and sufficiently small that, on millennial timescales, one regime is stabilized relative to the other, irrespective of the initial conditions. The inclusion of noise, strong enough that the residence times are sufficiently small that the stationary distribution is relevant on finite timescales, produces a substantially different description of the system than the deterministic model, for which the system is bistable over a broad range of values of μ. The inclusion of fluctuations dramatically reduces this region of bistability. This difference in descriptions has a direct bearing on the issue of the stability of the THC, to which we now turn.

5. Hysteresis loops and climate stability

As mentioned above, the deterministic equation (34) has one unstable and two stable fixed points for $0 \leq \mu \leq 0.25$. At $\mu = 0$, the larger stable fixed point, y_+, merges with the unstable fixed point; for $\mu < 0$ only the smaller stable fixed point, y_-, remains. At $\mu = 0.25$, the smaller stable fixed point merges with the unstable fixed point in another bifurcation, and above $\mu = 0.25$ only the larger stable fixed point remains. The bifurcation diagram is displayed in Fig. 11a. In the deterministic limit, if μ is varied from below 0 to above 0.25 and back down to below zero, the hysteresis response illustrated in Fig. 11b is obtained. The qualitative character of this deterministic hysteresis loop is generic to all simple models of the thermohaline circulation that display multiple stable fixed points, and has been well-discussed in the literature (e.g., Rahmstorf 1996). In this deterministic loop, transitions between the two equilibria occur exactly at the bifurcation points where one equilibrium disappears.

As discussed above, a qualitative difference between the stochastic and deterministic models is the potential stabilization of one regime by the fluctuations. For the range of (σ_1, σ_2) considered, there is a range of μ, well below the bifurcation point at $y = 0.25$, within which the stationary distribution favors the regime around y_+. Thus, if the residence time of the regime around y_- is less than the timescale of the circuit around the loop, $\tilde{T}$, then the jump to the upper branch will occur at a random value of μ, on average well before the bifurcation point at $\mu = 0.25$. Similarly, moving down the hysteresis loop, as μ drops below $\mu_{0.5}$, the stationary distribution p_s will be dominated by the regime around y_-. If $\langle T(y_+)\rangle$ is smaller than $\tilde{T}$, then the jump down to the lower branch will most often occur at some random value of μ before the bifurcation point at $\mu = 0$. It

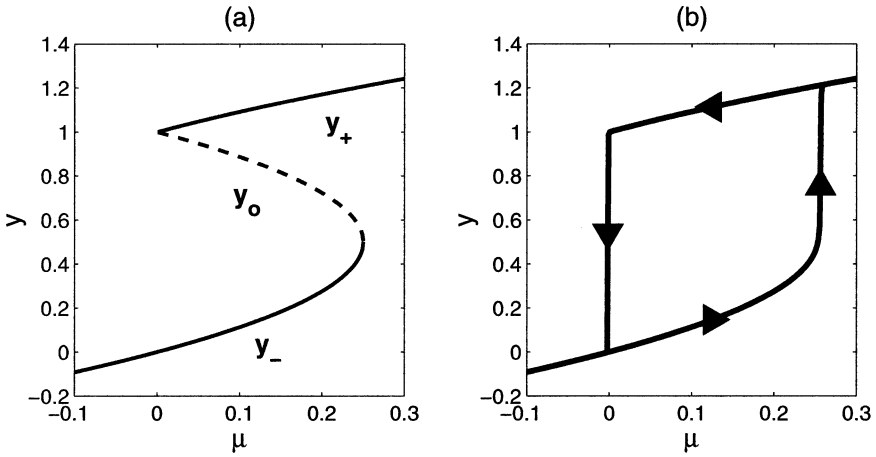

FIG. 11. (a) Plot of the unstable fixed point y_o and the two stable fixed points y_- and y_+ of the deterministic system (34) as functions of μ. There are bifurcations at $\mu = 0$ and $\mu = 0.25$. (b) Hysteresis curve for the deterministic model (34) obtained by increasing μ from -0.1 to 0.3 and then decreasing back to -0.1. The arrowheads illustrate the sense of movement around the curve.

follows that the hysteresis loop, which is now random, will be smaller in the stochastic case than in the deterministic limit. Such an effect was noted in the study of Wang et al. (1999), although it was not systematically investigated. The difference between the deterministic and stochastic hysteresis loops will be determined by the values of σ_1 and σ_2, which control the residence times of the two regimes, and by the rate at which μ is changed in the circuit around the loop.

Figure 12 displays sample hysteresis loops obtained by numerical solution of equation (20) with a time-varying value of μ, for four different pairs of noise strengths (σ_1, σ_2). The numerical simulations were carried out using a simple forward-Euler discretization

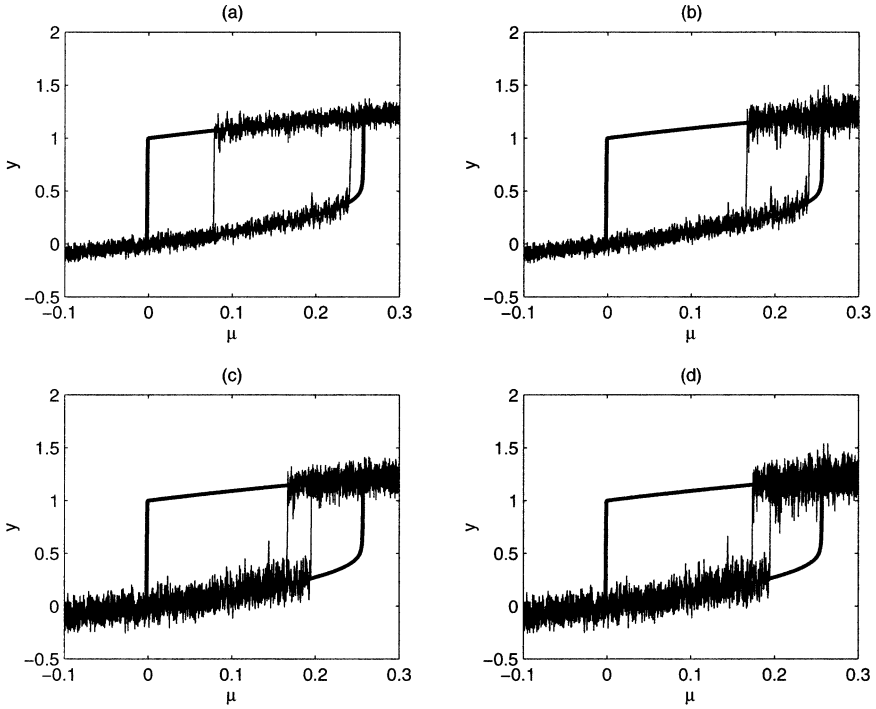

FIG. 12. Simulated hysteresis loops for the stochastic system (20) with noise levels (σ_1, σ_2): (a) (0.05, 0.05), (b) (0.1, 0.05), (c) (0.05, 0.1), (d) (0.1, 0.1). The thick curve is the deterministic hysteresis curve, plotted for comparison. The same realizations of the noise processes W_1 and W_2 were used in each simulation.

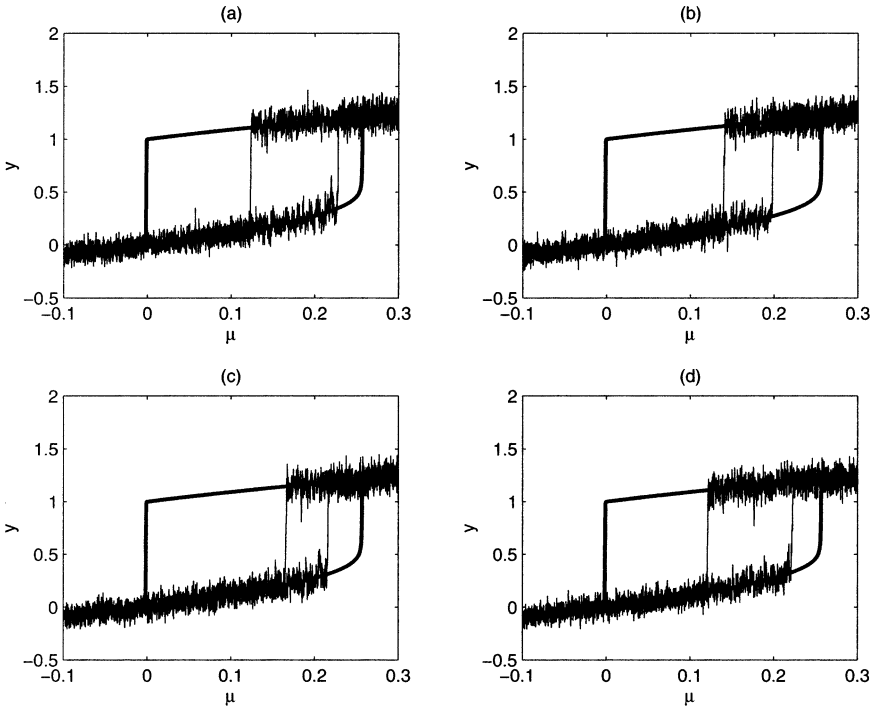

FIG. 13. As in Fig. 12 but with $(\sigma_1, \sigma_2) = (0.075, 0.075)$ for four different realizations of W_1 and W_2.

(Kloeden and Platen 1992). The parameter μ is moved from -0.1 to 0.3 over 2500 nondimensional time units, and back down to -0.1 at the same rate. To emphasise the differences induced by changing the noise levels alone, the same realizations of the Wiener processes W_1 and W_2 were used in all four simulations. Even for the fairly small noise levels $(\sigma_1, \sigma_2) = (0.05, 0.05)$, the hysteresis curve is substantially narrower than in the deterministic limit. As the strength of the noise increases, the hysteresis curves become narrower, on average. Figure 13 displays the random hysteresis curves obtained with four different realizations of the stochastic processes W_1 and W_2, at the fixed noise level $(\sigma_1, \sigma_2) = (0.075, 0.075)$. This figure demonstrates that, even for moderate noise levels, there is considerable variability in the values of μ at which the transitions occur. Figure 14 displays estimates of the PDFs of the values of μ for which the jumps between regimes occur, for $(\sigma_1, \sigma_2) = (0.075, 0.075)$, based on 1000 realizations of the hysteresis loop. For both transitions, the most likely value of μ at which the jump occurs is well separated from the deterministic bifurcation point. For smaller noise strengths, these distributions tighten and the most likely values of μ for the jumps approach the deterministic bifurcation points. For σ_1, σ_2 sufficiently small, the residence times of the regimes are for all μ larger than the timescale of the circuit around the loop, and the deterministic limit is recovered. On the other hand, for sufficiently large noise strengths, the residence times will be so much smaller than $\tilde{T}$ that more than two jumps between the regimes will be made, and the hysteresis loop structure is lost.

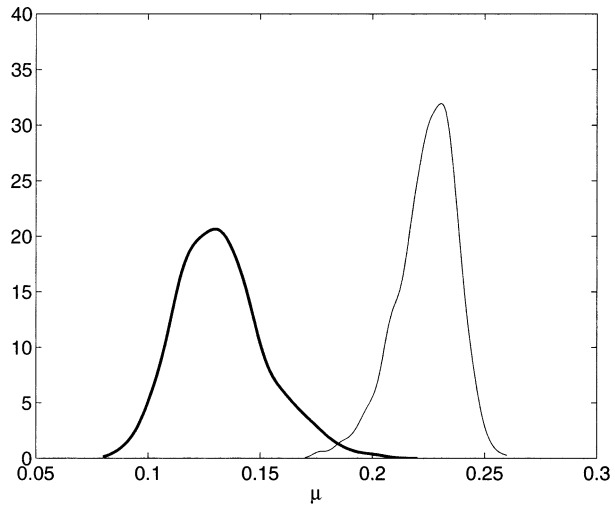

FIG. 14. Gaussian kernel estimates of the PDF of the value of μ at which the transition occurs from the regime around y_- to that around y_+ as μ increases (thin line), and the regime around y_+ to that around y_- as μ decreases (thick line), for $\sigma_1 = 0.075$, $\sigma_2 = 0.075$. The estimates are based on 1000 realizations of the hysteresis loop.

Multiple equilibria in the THC appear not only in simple box models, but also in sophisticated general circulation models (e.g., Manabe and Stouffer 1988; Hughes and Weaver 1994; Rahmstorf 1995, 1996, 1999; Tziperman 2000). Rahmstorf (1995, 1996) evaluated the

hysteresis loop of a full ocean GCM coupled to a simple atmosphere, and demonstrated that it could be fit rather well to the hysteresis loop produced by a simple box model. He was then able to estimate the position of the present climate on the box model hysteresis curve, and thereby estimate its proximity to the bifurcation point. However, we have demonstrated that in the presence of noise of even moderate strength, representing internal and external high-frequency variability, transitions between climate regimes may occur some distance away from the deterministic bifurcation points. Consequently, the distance between the present climate state and the deterministic bifurcation point may overestimate the stability of the present climate. Evidence for sensitivity of the hysteresis structure of a GCM to both noise realization and strength is given in Wang et al. (1999).

Furthermore, the results above suggest a potential problem with using even complex models to assess the stability of the THC, assuming that our stochastic parameterizations are appropriate representations of high-frequency variability. First, more than a single circuit of the hysteresis loop by the simulation would need to be carried out in order to estimate the distribution of the jump points. Admittedly, for models of intermediate complexity such as in Rahmstorf (1999), or for a full GCM, it is computationally very expensive to carry out multiple circuits around the hysteresis loop. However, this could potentially be important to determine the variability in the values of the control parameters at which the transitions occur. Second, even complex models tend to be biased toward low levels of internal variability. This bias can arise from either the finite resolution of the model or from the use of an uncoupled ocean model. No current ocean GCMs used for diagnosis of the stability of the thermohaline circulation resolve mesoscale eddies, and 2D ocean models do not resolve even gyre-scale circulations. Their internal noise levels are then presumably lower than those of the real ocean, perhaps substantially so. Furthermore, many studies of the stability of the thermohaline circulation use uncoupled ocean models, or ocean models coupled to toy atmospheres. It is a generic feature of uncoupled ocean or atmosphere models that they systematically underestimate the variability observed in nature, as discussed by Barsugli and Battisti (1998). Our simple stochastic box model suggests that an underestimate of the noise level in a more complex ocean model could lead to systematic errors in the range of control parameters at which the model is found to switch between climate equilibria. These systematic errors would presumably be removed through improved model resolution or coupling to an atmospheric model, or potentially through the introduction of stochastic subgrid-scale parameterizations (Palmer 2001).

6. Summary and conclusions

In this study, we have used a simple box model to consider the effects on the thermohaline circulation of both internal fluctuations, associated with unresolved internal eddy variability, and external fluctuations, associated with atmospheric processes. Because of the extreme simplicity of the model, a closed-form, analytic expression for the stationary probability density function p_s could be obtained from the associated Fokker–Planck equation. This stationary density was found to have a complicated dependence on the noise strengths σ_1 and σ_2, and on the parameter μ describing the differential freshwater forcing between the boxes. In particular, it was demonstrated that for much of the parameter space in which p_s is *technically* bimodal, in that it can be shown to have two extrema, it is *effectively* unimodal, in that one regime is very much more populated than the other. We denote this enhanced population of one regime as *stabilization.* For a broad range of noise levels within that part of parameter space in which the distribution is technically bimodal, the populations of the two regimes are comparable only for a small window of μ centered on the value $\mu_{0.5}$. Figure 7a illustrates that $\mu_{0.5}$ itself lies within a small range of values of μ. We see then that in a nonlinear stochastic system, the fluctuations may play a substantial role in determining the mean state of the system.

The stationary distribution p_s describes the behavior of the system in the infinite time limit. For finite times, it will only be relevant if the residence times of the two regimes are small compared to the the relevant timescale dictated by the physical process under consideration. Here, this process is the THC and the timescale is on the order of 10^4–10^5 yr. The simplicity of the model allows us to easily evaluate these residence times. For very small noise levels, the residence times are so large that the finite-time behavior of the system shares the feature with the deterministic system that the regime in which the system resides over the timescale of interest is determined by the initial conditions. However, there exists a broad range of parameter space for which the stationary distribution describes the behavior of the system and stabilization by fluctuations is relevant.

Stabilization of the system by fluctuations has a significant impact on the stability of the THC with regards to changes in freshwater forcing. In the deterministic limit, this may be described by a hysteresis loop, in which rapid transitions between regimes occur precisely at bifurcation points of the system. When fluctuations are added to the description, the hysteresis loop becomes random, and for intermediate noise strengths the transitions between regimes generally occur well away from the deterministic bifurcation points. This result suggests that fitting GCM output to simple deterministic models, as in Rahmstorf (1995, 1996), may overestimate the stability of the present climate. Furthermore, as GCMs generally are biased toward lower noise levels than those of the real world, this result also has bearing on GCM estimations of climate sensitivity.

These results provide another simple illustration of the idea put forward by Palmer (1999) that the effect

of small perturbations on the climate system, such as anthropogenic forcing, may be not to change the structure of the existing regimes of circulation, but their occupation statistics (assuming the system is sufficiently far from a bifurcation point). Each of the parameters μ, σ_1, σ_2 considered in this study could conceivably be altered by anthropogenic forcing, with consequent changes in the populations of the regimes, especially for μ in the neighborhood of $\mu_{0.5}$.

It is also worth noting that this study demonstrates the importance of accurate discretization schemes for the numerical solution of nonlinear SDEs. Inaccurate discretizations can lead to substantial systematic errors in the strength of the noise, as in Lagerloef (1995). The qualitative behavior of the model considered in this study is sensitive to the strengths of the stochastic fluctuations, and simulations carried out using an inaccurate discretization scheme could consequently be qualitatively inaccurate. The numerical solution of SDEs is described in detail in Kloeden and Platen (1992).

The analysis presented in this study also demonstrates the effects that stochastic fluctuations can have on the bifurcation structure of a dynamical system. The stationary PDF displays bifurcation points at which it changes from bimodal to unimodal; these bifurcations are the distributional equivalents of the deterministic bifurcations at which a stable fixed point disappears. For the range of noise strengths considered here, the bifurcations of the stationary PDF and the deterministic system occur at not very different values of μ. However, because of stabilization, the bifurcation behavior of individual realizations is generally quite different than in the deterministic case. An interesting analog occurs in the effect of stochastic perturbations on Hopf bifurcations. In a deterministic forward Hopf bifurcation from a fixed point attractor, the asymptotic behavior of the system changes from one in which there is no characteristic oscillatory timescale to one characterized by a dominant oscillatory timescale. In the stochastic setting, the Hopf bifurcation is associated with changes in the stationary PDF and Lyapunov exponents (Arnold et al. 1996), but individual realizations possess oscillatory timescales that change smoothly through the bifurcation point. In effect, the presence of fluctuations destroys the bifurcation in the power spectrum present in the deterministic limit, a result that is of significance to theories of ENSO dynamics (Penland et al. 2000) and of interdecadal THC variability (Rivin and Tziperman 1997). These results demonstrate that stochastic fluctuations can have significant effects on important aspects of the bifurcation structure of climatic dynamical systems.

The model we have investigated in this study is highly simplified, even by the standards of box modeling. Because of the extreme simplicity of the model, no attempt has been made to fit the parameters μ, σ_1, or σ_2 to data from observations or GCMs. Preliminary numerical investigations of the model (10)–(12) indicate that the stabilization phenomenon is qualitatively unchanged for nonzero γ^{-1} and τ. A physically meaningful quantitative examination of stabilization of THC regimes by noise would require more sophisticated models. A more dynamically significant, yet computationally efficient, class of models is the zonally averaged 2D ocean circulation models (e.g., Wright et al. 1998). This class of models could potentially be used for more quantitative, albeit numerical, studies of THC stabilization by noise.

Simple box models represent the most coarsely discretized extreme of models of ocean dynamics, and should be interpreted accordingly. While their *quantitative* features cannot be expected to be reliable, their *qualitative* behavior can be quite instructive. The remarkable correspondence between the general character of the behavior of box models and that of complicated GCMs demonstrated by Rahmstorf (1995, 1996) suggests that their apparent simplicity belies their ability to capture essential features of ocean dynamics. The stochastic model considered above introduces the phenomenon of stabilization that has important consequences for the dynamics of the model. The results of this simple model suggest an interesting experiment, in which a more complex model is run through as many hysteresis loops as computing power allows. If the stochastic parameterization of eddy variability is appropriate, the model should demonstrate random hysteresis loops, and this experiment would allow the assessment of their distribution. Such an experiment could be carried out using different model resolutions in order to vary the strength of the internal noise in the system. Evidence of a nondeterministic character of the hysteresis loops would have important consequences for any discussion of the stability of the THC.

Acknowledgments. The author would like to thank Peter Imkeller, Till Kuhlbrodt, Cécile Penland, Sven Titz, Stefan Rahmstorf, Lionel Pandolfo, Richard Greatbatch, Gerrit Lohmann, and Axel Timmermann, for their suggestions and helpful comments. The original manuscript was also greatly improved by the comments and suggestions of two anonymous referees. This work was supported by DFG Schwerpunktprogramm ''Inter-agierende stochastische Systeme von hoher Komplexität.''

REFERENCES

Alley, R. B., 2000: Ice-core evidence of abrupt climate changes. *Proc. Natl. Acad. Science,* **97,** 1331–1334.

Arnold, L., N. S. Namachchivaya, and K. R. Schenk-Hoppe, 1996: Toward an understanding of stochastic Hopf bifurcation: A case study. *Int. J. Bifur. Chaos,* **6,** 1947–1995.

Barsugli, J. J., and D. S. Battisti, 1998: The basic effects of atmosphere–ocean thermal coupling on midlatitude variability. *J. Atmos. Sci.,* **55,** 477–493.

Bond, G., and Coauthors, 1997: A pervasive millennial-scale cycle in North Atlantic Holocene and glacial climates. *Science,* **278,** 1257–1266.

Cessi, P., 1994: A simple box model of stochastically forced thermohaline flow. *J. Phys. Oceanogr.,* **24,** 1191–1920.

Gardiner, C. W., 1997: *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences.* Springer, 442 pp.

Griffies, S. M., and E. Tziperman, 1995: A linear thermohaline oscillator driven by stochastic atmospheric forcing. *J. Climate,* **8,** 2440–2453.

Hughes, T. M. C., and A. J. Weaver, 1994: Multiple equilibria of an asymmetric two-basin ocean model. *J. Phys. Oceanogr.,* **24,** 619–637.

Kloeden, P. E., and E. Platen, 1992: *Numerical Solution of Stochastic Differential Equations.* Springer, 632 pp.

Lagerloef, G. S. E., 1995: Interdecadal variations in the Alaska gyre. *J. Phys. Oceanogr.,* **25,** 2242–2258.

Lohmann, G., and J. Schneider, 1999: Dynamics and predictability of Stommel's box model: A phase space perspective with implications for decadal climate variability. *Tellus,* **51A,** 326–336.

Maas, L., 1994: A simple model for the three-dimensional, thermally and wind-driven ocean circulation. *Tellus,* **46A,** 671–680.

Manabe, S., and R. J. Stouffer, 1988: Two stable equilibria of a coupled ocean–atmosphere model. *J. Climate,* **1,** 841–866.

Molteni, F., and S. Tibaldi, 1990: Regimes in the wintertime circulation over the northern extratropics. II: Consequences for dynamical predictability. *Quart. J. Roy. Meteor. Soc.,* **116,** 1263–1288.

Nakamura, M., and Y. Chao, 2000: On the eddy isopycnal thickness diffusivity of the Gent–McWilliams subgrid mixing parameterisation. *J. Climate,* **13,** 502–510.

Palmer, T. N., 1999: A nonlinear dynamical perspective on climate prediction. *J. Climate,* **12,** 575–591.

——, 2001: A nonlinear dynamical perspective on model error: A proposal for nonlocal stochastic-dynamic parameterisation in weather and climate models. *Quart. J. Roy. Meteor. Soc.,* **127,** 279–304.

Penland, C., 1996: A stochastic model of IndoPacific sea surface temperature anomalies. *Physica D,* **98,** 534–558.

——, M. Flügel, and P. Chang, 2000: Identification of dynamical regimes in an intermediate coupled ocean–atmosphere model. *J. Climate,* **13,** 2105–2115.

Rahmstorf, S., 1995: Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle. *Nature,* **378,** 145–149.

——, 1996: On the freshwater forcing and transport of the Atlantic thermohaline circulation. *Climate Dyn.,* **12,** 799–811.

——, 1999: Rapid transitions of the thermohaline ocean circulation: A modelling perspective. *Reconstructing Ocean History: A Window Into the Future,* F. Abrantes and A. Mix, Eds., Kluwer Academic/Plenum Publishers, 139–149.

Rivin, I., and E. Tziperman, 1997: Linear versus self-sustained interdecadal thermohaline variability in a coupled box model. *J. Phys. Oceanogr.,* **27,** 1216–1232.

Scott, J., J. Marotzke, and P. Stone, 1999: Interhemispheric thermohaline circulation in a coupled box model. *J. Phys. Oceanogr.,* **29,** 351–365.

Stocker, T. F., and O. Marchal, 2000: Abrupt climate change in the computer: Is it real? *Proc. Natl. Acad. Sci.,* **97,** 1362–1365.

Stommel, H., 1961: Thermohaline convection with two stable regimes of flow. *Tellus,* **13,** 224–230.

——, and W. Young, 1993: The average *T–S* relation of a stochastically-forced box model. *J. Phys. Oceanogr.,* **23,** 151–158.

Stone, P. H., and Y. P. Krasovskiy, 1999: Stability of the interhemispheric thermohaline circulation in a coupled box model. *Dyn. Atmos. Oceans.,* **29,** 415–435.

Timmermann, A., and G. Lohmann, 2000: Noise-induced transitions in a simplified model of the thermohaline circulation. *J. Phys. Oceanogr.,* **30,** 1891–1900.

Titz, S., T. Kuhlbrodt, S. Rahmstorf, and U. Feudel, 2001: On freshwater-dependent bifurcations in box models of the interhemispheric thermohaline circulation. *Tellus,* in press.

Tziperman, E., 2000: Proximity of the present-day thermohaline circulation to an instability threshold. *J. Phys. Oceanogr.,* **30,** 90–104.

Wang, X., P. H. Stone, and J. Marotzke, 1999: Global thermohaline circulation. Part I: Sensitivity to atmospheric moisture transport. *J. Climate,* **12,** 71–82.

Weaver, A., and T. M. C. Hughes, 1992: Stability and variability of the thermohaline circulation and its link to climate. *Trends Phys. Oceanogr.,* **1,** 15–70.

Wright, D. G., T. F. Stocker, and D. Mercer, 1998: Closures used in zonally averaged ocean models. *J. Phys. Oceanogr.,* **28,** 791–804.