Non-Normal Effects on Salt Finger Growth

IAN EISENMAN

Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts

(Manuscript received 18 August 2004, in final form 8 October 2004)

ABSTRACT

Salt fingers, which occur because of the difference in diffusivities of salt and heat in water, may play an important role in ocean mixing and circulation. Previous studies have suggested the long-time dominance of initially fastest growing finger perturbations. Finger growth has been theoretically derived in terms of the normal modes of the idealized system, which include a growing mode and a pair of decaying internal wave modes. Because these normal modes are not orthogonal, however, transient effects can occur related to the interaction between the modes, as explained by the generalized stability theory of non-normal growth. Initial growth of a perturbation that is not along a normal mode can be faster than the leading normal mode. In this study, the effects of non-normal growth on salt finger formation are investigated. It is shown that some salt finger perturbations that are a superposition of the growing mode and the decaying modes initially grow faster than pure growing normal mode perturbations. These non-normal effects are found to be significant for up to 10 or more *e*-folding times of the growing normal mode. The generalization of the standard idealized salt finger growth dynamics to include non-normal effects is found to lead to fastest-growing fingers that agree less well with observed fully developed salt fingers than the fastest-growing normal mode previously investigated.

1. Introduction

When warm salty water overlies cool freshwater, the density is stably stratified in temperature but unstably stratified in salinity. If the temperature gradient dominates, the density stratification is stable, but the more rapid diffusion of heat than salt in water allows the potential energy stored in the salinity stratification to be released on small scales. If a small parcel of water is displaced downward, its heat will quickly diffuse to bring it into thermal equilibrium with the water around it, but it will remain salty and thus less buoyant than its surroundings. This leads to unstable perturbation growth. The destabilizing heat diffusion, which is most effective on smallest scales, can be balanced by stabilizing molecular viscous drag, which becomes significant on small spatial scales, leading to maximum growth of perturbations with a particular horizontal scale of variability (Stern 1960). Salty blobs of water with this characteristic width fall down like fingers with freshwater rising between them.

Experiments carried out in the laboratory using salt–sugar or heat–salt mixtures lead to readily observable narrow, vertically aligned pockets of salty or sugary water called salt fingers. Since solar heating of the upper ocean leads to salinity enhancement by evaporation, it is not surprising that much of the upper ocean is fingering favorable such that salt fingers a few centimeters wide and about a meter long are expected to exist and are evidenced in measurements (Schmitt 1994a). Although observational evidence of salt fingers in the ocean is often conflicting, many oceanographers today believe that salt fingering plays a major role in water-mass mixing and ocean circulation (Ruddick and Gargett 2003).

Salt fingers are believed to grow because of an instability related to the difference in diffusion rates of heat and salt in water, but scientific consensus on the mechanism for arresting their growth has not been reached (Kunze 2003). Interestingly, the basic salt finger equations (section 3) are linear, so the dynamics do not offer a simple nonlinear arresting mechanism. Proposed arresting mechanisms include turbulence and the nonlinearity in the seawater equation of state (Kunze 2003). In common among most suggested arresting mechanisms is a lack of discrimination among the salt fingers; that is, the fastest-growing fingers dominate the steady solution. This is encouraged by the close match between the theoretical fastest-growing normal mode and observed fully developed fingers in laboratory and numerical experiments (Fig. 6a, described below). The findings presented in this paper suggest that this match has appeared closer than it actually is because of a

Corresponding author address: Ian Eisenman, Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138.
E-mail: eisenman@fas.harvard.edu

mathematical simplification associated with examining only the growing normal mode.

Theoretical studies of the fastest-growing finger perturbations have all followed a common method in geophysical fluid dynamics stability problems in which the normal modes (eigenmodes of the dynamical operator) are examined and growth is identified with positive eigenvalues (Kunze 1987; Schmitt 1979b; Stern 1975). For systems with a non-normal propagator, however, it is known that transient effects can cause some perturbations to initially grow faster than the leading normal mode (Farrell 1988; Farrell and Ioannou 1996). These effects are typically called non-normal growth and have been examined mainly in the context of the dynamics of linearly stable systems including El Niño (Moore and Kleeman 1996; Penland and Sardeshmukh 1995), the wind-driven ocean circulation (Moore et al. 2002), and thermohaline circulation (Tziperman and Ioannou 2002). The propagator in the standard idealized dynamics of salt finger formation (a linearly unstable system) is non-normal, and although non-normal effects are typically fairly short lived, they may be significant on time scales relevant for salt finger formation.

The paper begins with an explanation of the tools of the generalized stability theory of non-normal growth relevant to this study (section 2). This is followed by a presentation of the standard equations of salt finger formation, expressed here in a three-dimensional (temperature, salinity, vertical velocity) perturbation vector space (section 3). A brief discussion of previous work in the context of the present formalism is given in section 4a. In section 4b, non-normal effects on the initial growth of wide salt fingers are analytically examined, and the full salt finger stability problem including non-normal effects is presented in section 4c. This is used to derive the spectrum of spatial variability for salt fingers initially excited by a white noise perturbation (section 4d), which is compared with the previous result that assumed only normal mode growth. In section 4e, the flux ratio (ratio of heat to salt contributions to perturbation buoyancy flux) for the fastest-growing finger is compared with experimental measurements and the normal mode growth result. The choice of norm used to quantify perturbation size and identify the fastest-growing salt finger is addressed in section 4f. Concluding remarks are given in section 5.

2. Generalized stability theory

The classical treatment of stability problems in geophysical fluid dynamics involves the identification of exponentially growing normal modes in a linearized system (e.g., Charney 1947; Eady 1949; Pedlosky 1987, chapter 7). Small perturbations about a mean state are examined according to

$$\frac{d\mathbf{v}}{dt} = \mathbf{K}\mathbf{v}, \tag{1}$$

where $\mathbf{v}(t)$ is the state vector and $\mathbf{K}$ is the linearized dynamical operator. The operator $\mathbf{K}$ generally includes partial derivatives in space. Sinusoidal spatial variation is typically assumed, and $\mathbf{K}$ becomes a matrix (with any remaining differential operators replaced by finite difference approximations); $\mathbf{K}$ is usually time independent, and the normal modes of the system, which are the eigenvectors of $\mathbf{K}$, evolve exponentially at a rate given by the associated eigenvalues. In the classical stability problem, the eigenvalue is considered as a function of the spatial wavenumber (dispersion relation), and a positive real part of the leading eigenvalue is associated with instability.

While this method is exactly accurate for assessing stability in the asymptotic limit $t \to \infty$, it neglects the possibility of transient non-normal growth, which can occur through interaction between the normal modes of the system (Farrell 1988). A broad generalization of classical stability theory to include transient non-normal effects is presented by Farrell and Ioannou (1996). What follows is a brief discussion of the aspects of this theory that will be used in section 4.

In the dynamical system (1), if $\mathbf{K}$ commutes with its Hermitian transpose, $\mathbf{K}\mathbf{K}^\dagger = \mathbf{K}^\dagger\mathbf{K}$, then $\mathbf{K}$ is normal and has a complete set of orthogonal eigenvectors. In this situation, no non-normal effects arise. If $\mathbf{K}\mathbf{K}^\dagger \neq \mathbf{K}^\dagger\mathbf{K}$, however, the normal modes of the dynamical system (eigenvectors of $\mathbf{K}$) are not orthogonal, and the decay of a vector's projection on a mode (or the effects of rotation for complex eigenmodes) may lead to overall growth of the vector.

Non-normal effects can lead to transient growth in a system with only decaying normal modes. Consider, for example, the two-dimensional system whose eigenvectors are given by $\mathbf{e_1}$ and $\mathbf{e_2}$ in Fig. 1a. An initial perturbation of unit length is labeled **ic**. Its projection on the eigenvectors is given by the dashed lines. If both eigenvectors decay but $\mathbf{e_2}$ decays significantly faster than $\mathbf{e_1}$, the perturbation **ic** will grow initially toward the point **p** before decaying approximately along $\mathbf{e_1}$. The initial condition has significant projections on both normal modes, which substantially cancel each other out; as one of the projections rapidly decays, the cancellation disappears and the length of the perturbation increases.

Similarly, non-normal effects can lead to transient growth enhancement in a system with normal mode instability. The effect is most dramatic when the decay is fast and the growth is slow. Considering this scenario in Fig. 1a, for example, as the projection of **ic** on $\mathbf{e_1}$ slowly grows, its projection on $\mathbf{e_2}$ quickly decays, causing the perturbation to shoot out toward the point **p** before slowly growing along $\mathbf{e_1}$.

The dynamical operator for the system in Fig. 1a can be written

$$\mathbf{K} = \begin{bmatrix} \lambda_1 & (\lambda_2 - \lambda_1)\cot(\theta) \\ 0 & \lambda_2 \end{bmatrix}, \tag{2}$$

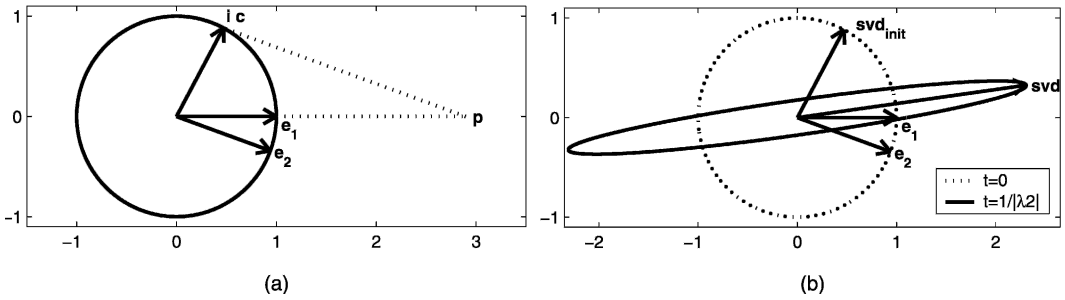

FIG. 1. (a) An initial perturbation, labeled **ic**, is considered in a dynamical system whose eigenvectors are $\mathbf{e_1}$ and $\mathbf{e_2}$. The dotted lines indicate the projection of the initial perturbation on the eigenvectors. If both eigenvectors decay in time but $\mathbf{e_2}$ decays significantly faster than $\mathbf{e_1}$, the perturbation will initially grow toward the point **p** before decaying. If $\mathbf{e_2}$ decays but $\mathbf{e_1}$ grows, the perturbation will initially grow rapidly toward **p**, faster than a perturbation aligned with the growing normal mode $\mathbf{e_1}$, before approaching the growth rate of $\mathbf{e_1}$ (cf. Farrell and Ioannou 1996, their Fig. 1). (b) An ensemble of initial perturbations of unit length, represented by the dotted circle, is evolved by the system in (a) with $\lambda_1 > 0$, $\lambda_2 < 0$, and $|\lambda_2|/\lambda_1 = 10$. The evolved perturbations after time $1/|\lambda_2|$ are indicated by the solid ellipse. The initial perturbation that experiences the most growth after time $1/|\lambda_2|$ is labeled svd_{init}, and its location after time $1/|\lambda_2|$ is labeled svd; these are the leading initial and final singular vectors.

where λ_1 and λ_2 are the eigenvalues and the angle between the eigenvectors is θ, which is 20° in Fig. 1a. Figure 1b illustrates the evolution of this system after time $1/|\lambda_2|$ when $\lambda_1 > 0$, $\lambda_2 < 0$, and $|\lambda_2|/\lambda_1 = 10$. An ensemble of initial perturbations of unit length, shown as a dotted circle, evolves into an ellipse, plotted as a solid curve. The initial perturbation that grows fastest during the duration $1/|\lambda_2|$ is labeled svd_{init}; its final location is labeled svd. Note that the perturbation along $\mathbf{e_1}$, which is the unstable normal mode perturbation, grew by 10% (barely discernible in the figure), while the perturbation svd_{init} grew by 133%.

More generally, if the dynamical system (1) is evolved by an autonomous ordinary differential equation (i.e., no time dependence or derivatives in $\mathbf{K}$), we can formally write the evolution of the state vector as

$$\mathbf{v}(t) = \Phi\mathbf{v}(0) \quad \text{and} \tag{3}$$

$$\Phi = \Phi(t) = \exp(\mathbf{K}t), \tag{4}$$

where exp() is the matrix exponential and Φ is the propagator of the system. We wish to find the initial condition $\mathbf{v}(0)$ that maximizes the norm of the state vector after time t,

$$|\mathbf{v}(t)|^2 = \mathbf{v}(t)^\dagger\mathbf{v}(t) = \mathbf{v}(0)^\dagger\Phi^\dagger\Phi\mathbf{v}(0), \tag{5}$$

subject to the constraint $|\mathbf{v}(0)| = 1$. This constrained variational problem can be solved by optimizing the functional (Farrell 1988):

$$\mathbf{v}(0)^\dagger\Phi^\dagger\Phi\mathbf{v}(0) + \lambda[1 - \mathbf{v}(0)^\dagger\mathbf{v}(0)]. \tag{6}$$

Setting the derivative with respect to $\mathbf{v}(0)$ to zero, we find that the vector $\mathbf{v}(0)$ that renders (6) stationary solves

$$\Phi^\dagger\Phi\mathbf{v}(0) = \lambda\mathbf{v}(0). \tag{7}$$

The initial condition leading to maximal growth after time t, then, is the leading eigenvector of $\Phi^\dagger\Phi$. Multiplying (7) from the left by $\mathbf{v}(0)^\dagger$ and comparing it with (5), we see that the Lagrange multiplier λ, which is the eigenvalue in (7), gives the square of the growth of $\mathbf{v}(t)$. The structure of the perturbation after time t can be computed by operating the propagator on the initial vector.

The problem posed here addresses the initial perturbation that maximizes the Euclidean norm, $|\mathbf{v}|^2 = \mathbf{v}^\dagger\mathbf{v}$. When maximizing a more general norm defined through the inner product norm kernel $\mathbf{X}$ as $|\mathbf{v}|^2 = \mathbf{v}^\dagger\mathbf{X}\mathbf{v}$, where the initial condition has unit norm under a possibly different norm kernel $\mathbf{Y}$, the eigenproblem (7) takes the form

$$\Phi^\dagger\mathbf{X}\Phi\mathbf{v}(0) = \lambda\mathbf{Y}\mathbf{v}(0). \tag{8}$$

The choice of a norm in the salt finger problem is addressed in section 4f. With the exception of that section, only the Euclidean norm is used in the following calculations.

The eigenproblem (7) can be equivalently expressed in terms of a singular value decomposition of the propagator,

$$\Phi = \mathbf{U}\Sigma\mathbf{V}^\dagger, \tag{9}$$

where Σ is a diagonal matrix of singular values that give the growth (square root of the eigenvalues of $\Phi^\dagger\Phi$) of initial singular vectors, which are columns of $\mathbf{V}$ (equal to the eigenvectors of $\Phi^\dagger\Phi$). The final singular vectors, columns of $\mathbf{U}$, give the evolved perturbations after time t when they are multiplied by the associated singular values.

To summarize, the growth of the leading normal mode after time t is

$$\text{eig}(\Phi) = \text{eig}(e^{\mathbf{K}t}) = e^{\text{eig}(\mathbf{K})t}, \tag{10}$$

where eig() represents the operation that takes the leading eigenvalue, and the matrix exponential is im-

plied. In normal mode stability theory, eig($\mathbf{\Phi}$) is identified as the greatest growth after time t. In the generalized stability approach, which includes transient effects, the greatest growth after time t is given by

$$\text{svd}(\mathbf{\Phi}) = [\text{eig}(\mathbf{\Phi}^\dagger\mathbf{\Phi})]^{1/2} = \text{svd}(e^{\mathbf{K}t}). \tag{11}$$

Here the operation svd( ) takes the leading singular value.

The generalized stability theory described here is typically applied in the context of transient growth in systems whose normal modes are all stable, as well as the growth of errors in numerical prediction models. It can be similarly applied, however, to study transient growth enhancement of unstable salt fingers. This leads to predicted salt finger growth after a fairly short time duration that significantly disagrees with the traditional predictions of normal mode stability theory.

3. Salt finger equations of motion

We consider here the growth of tall narrow salt fingers in a vertically unbounded fluid with uniform vertical temperature and salinity gradients, following Stern (1975), Schmitt (1979b), and Kunze (1987). The perturbations are independent of height and there is no horizontal velocity (assuming incompressibility and no mean flow, either of these assumptions will lead to the other by mass continuity), so we anticipate that the nonlinear advection terms in the momentum and diffusion equations will vanish and the system will be linear. The temperature and salinity diffusion equations and the Boussinesq approximated vertical momentum equation for perturbations ($\hat{T}$, $\hat{S}$, $\hat{w}$) about a mean state ($\overline{T}$, $\overline{S}$, 0) are

$$\frac{\partial\alpha\hat{T}}{\partial t_*} = \kappa_T\nabla^2\alpha\hat{T} - \hat{w}\,\alpha\overline{T}_z, \tag{12}$$

$$\frac{\partial\beta\hat{S}}{\partial t_*} = \kappa_S\nabla^2\beta\hat{S} - \hat{w}\beta\overline{S}_z, \quad \text{and} \tag{13}$$

$$\frac{\partial\hat{w}}{\partial t_*} = \nu\nabla^2\hat{w} - g(\beta\hat{S} - \alpha\hat{T}). \tag{14}$$

Here $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the two-dimensional Laplacian, $\alpha = -1/\rho\ \partial\rho/\partial T$ and $\beta = +1/\rho\ \partial\rho/\partial S$ are assumed constant, κ_T and κ_S are the molecular heat and salt diffusivities, ν is the kinematic viscosity, and the mean state gradients $\overline{T}_z$ and $\overline{S}_z$ are assumed to be constant. Anticipating nondimensionalization, the subscript in t_* indicates that the variable is dimensional.

We begin by assuming sinusoidal variation in space,

$$\begin{bmatrix} \alpha\hat{T}(\mathbf{x}, t) \\ \beta\hat{S}(\mathbf{x}, t) \\ \hat{w}(\mathbf{x}, t) \end{bmatrix} = \begin{bmatrix} \alpha T(t) \\ \beta S(t) \\ w(t) \end{bmatrix} e^{i\,\mathbf{m}\cdot\mathbf{x}}, \tag{15}$$

to reduce the partial differential equation to an ordinary differential equation. Here $\mathbf{m}$ is the wavenumber and $\mathbf{x}$ is the horizontal space vector.

The planform $e^{i\mathbf{m}\cdot\mathbf{x}}$ might represent a sheet as sin(kx), a square of the form sin(kx) sin(ly), or a linear combination of the two that can take on the form of triangles or hexagons or asymmetric modes (Schmitt 1994b), since the dynamic equations include only the norm of the horizontal wavenumber $m = |\mathbf{m}| = |k\hat{\mathbf{x}} + l\hat{\mathbf{y}}|$. A more general spatial structure can be obtained by considering a superposition of spatial Fourier modes in this linear problem.

We write the system (12)–(14) with (15) in the formalism of (1) as

$$\frac{d\mathbf{v}_*}{dt} = \mathbf{K}_*\mathbf{v}_*, \tag{16}$$

$$\mathbf{v}_*(t) = \begin{pmatrix} \alpha T \\ \beta S \\ w \end{pmatrix}, \quad \text{and} \tag{17}$$

$$\mathbf{K}_* = \begin{pmatrix} -\kappa_T m^2 & 0 & -\alpha\overline{T}_z \\ 0 & -\kappa_S m^2 & -\beta\overline{S}_z \\ g & -g & -\nu m^2 \end{pmatrix}. \tag{18}$$

It can be readily verified that $\mathbf{K}_*$ is non-normal by checking that $\mathbf{K}^\dagger\mathbf{K} \neq \mathbf{K}\mathbf{K}^\dagger$, so the normal modes are not orthogonal. Previous work (Kunze 1987; Schmitt 1979b; Stern 1975) examined the leading normal mode of the system to find the growth of perturbations—that is, $\mathbf{v}_*(t) = \mathbf{v}_{\mathbf{0}*}e^{\lambda t}$, $\lambda = \text{eig}(\mathbf{K}_*)$. Here we will generalize this to include trajectories that rotate in (αT, βS, w) space such that transient non-normal effects can be examined.

The system can be nondimensionalized by introducing several nondimensional parameters: the Prandtl number $\sigma = \nu/\kappa_T$, inverse Lewis number $\tau = \kappa_T/\kappa_S$, and density gradient ratio $R_\rho = (\alpha\overline{T}_z)/(\beta\overline{S}_z)$. The constant $\mu = g(\alpha\overline{T}_z)$ is used to scale time, $t = \sqrt{\mu}t_*$, as well as vertical velocity, $\omega = (\sqrt{\mu}/g)w$. The wavenumber is scaled following Schmitt (1979b) as $M^2 = (\sqrt{\nu\kappa_T}/\mu)m^2$, and the temperature and salinity perturbations are written $a = \alpha T$ and $b = \beta S$. This allows the system (16)–(18) to be rewritten

$$\frac{d\mathbf{v}}{dt} = \mathbf{K}\mathbf{v}, \tag{19}$$

$$\mathbf{v}(t) = \begin{pmatrix} a \\ b \\ \omega \end{pmatrix}, \quad \text{and} \tag{20}$$

$$\mathbf{K} = \begin{pmatrix} -\sigma^{-1/2}M^2 & 0 & -1 \\ 0 & -\tau^{-1}\sigma^{-1/2}M^2 & -R_\rho^{-1} \\ 1 & -1 & -\sigma^{1/2}M^2 \end{pmatrix}. \tag{21}$$

The parameters τ and σ are specified in any physical system. This leaves two parameters in addition to time: the horizontal wavenumber or finger width M and the mean state density gradient ratio R_ρ. When these parameters are specified, the evolution of the salt finger system in $(a, b, \omega) \sim (T, S, w)$ space is determined by (19).

4. Growth of salt fingers

The central questions to be examined in a theory of salt finger formation relate to the rate of growth and the characteristic structure of the finger that grows fastest, which is expressed in this formalism in terms of the value of M that leads to fastest growth (i.e., the spatial Fourier mode that grows fastest) and the direction of fastest growth in state vector space, $\mathbf{v}(t)/|\mathbf{v}(t)|$. Laboratory experiments often measure the salt finger flux ratio between heat and salt perturbation contributions to buoyancy,

$$R_F = \frac{\langle \hat{w}\alpha\hat{T}\rangle}{\langle \hat{w}\beta\hat{S}\rangle} = \frac{\alpha T}{\beta S} = \frac{a}{b}, \tag{22}$$

which is the slope of the projection of the state vector onto the temperature–salinity plane. Here the brackets indicate horizontal spatial averages, which are evaluated using (15).

a. Normal mode growth

The growth rates of the normal modes of the system (19)–(21) are given by the eigenvalues of the matrix $\mathbf{K}$. For supercritical salt fingers (i.e., fingers that exponentially grow),

$$1 < R_\rho < \tau. \tag{23}$$

This constraint applies because if $R_\rho < 1$, the density gradient ratio is unstable and dynamic convection occurs, while the salinity difference between fingers and their surroundings is too small to overcome the stable density stratification if $R_\rho > \tau$. When (23) is satisfied, the three eigenvalues correspond to an exponentially growing finger and a pair of oscillatory decaying fingers (internal waves) in the region of parameter space that includes the range of parameter values typically examined for salt fingers (a bifurcation occurs at the edge of this region, beyond which two real decaying modes replace the oscillatory modes). Two normal modes are required to describe the physical oscillation, since the oscillation corresponds to a complex eigenvalue of a real matrix and these come in conjugate pairs.

The eigenvalue equation for $\mathbf{K}$ is a cubic polynomial, the solution of which is algebraically messy, and studies of salt finger growth have typically avoided solving it. Stern (1975) took the limit $\sigma \gg 1$ and $\tau \gg 1$, which causes the time derivative of the vertical velocity to drop out of (14) and reduces the problem to two dimensions, allowing the cubic polynomial to be approximated as quadratic. Kunze (1987) used the same approximation. Schmitt (1979b) was able to restate the cubic eigenvalue equation for growth rate in terms of wavenumber into a quadratic equation for growth rate in terms of flux ratio. Backed by a physical argument, he found the flux ratio that maximized the growth rate and thus described the fastest-growing normal mode. All three normal modes, however, are important for an examination of non-normal salt finger growth, since the non-normal effects arise from interaction among the three modes.

b. Initial growth of a wide finger

Before considering the full problem for the evolution of a general perturbation to (19)–(21) after a finite amount of time, we consider here the *initial* growth of a finger perturbation of infinite width, $M \to 0$. The spatial dependence drops from the problem, and the ratio of mean state gradients, R_ρ, becomes the only parameter. This problem, which can be solved analytically, sheds light on the considerable difference between normal mode growth and non-normal growth which will be discussed in section 4c.

We begin by examining the normal modes of the system, eigenvectors of $\mathbf{K}$ (21) with $M = 0$. They are

$$\mathbf{e}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix},$$
$$\mathbf{e}_2 = d\begin{bmatrix}c^{-1}\\(R_\rho c)^{-1}\\1\end{bmatrix}, \quad \text{and} \tag{24}$$
$$\mathbf{e}_3 = d\begin{bmatrix}-c^{-1}\\-(R_\rho c)^{-1}\\1\end{bmatrix},$$

where

$$c = \sqrt{\frac{1}{R_\rho}} - 1 \quad \text{and} \quad d = \sqrt{\frac{R_\rho - R_\rho^2}{1 + R_\rho}}, \tag{25}$$

with associated eigenvalues

$$\lambda_1 = 0, \quad \lambda_2 = -c, \quad \text{and} \quad \lambda_3 = c. \tag{26}$$

For systems that can support salt fingers (23), c is purely imaginary. The first eigenvector corresponds to the normal mode typically associated with salt finger growth, which is stationary when $M = 0$, and the second two eigenvectors are a pair of internal wave normal modes with constant amplitude. No normal mode experiences growth.

There are perturbations to this system, however, that do initially grow. When all terms of order dt^2 or higher are dropped from (4), the propagator becomes

$$\Phi = \exp(\mathbf{K}\,dt) = 1 + \mathbf{K}\,dt. \tag{27}$$

The maximum possible growth after time dt is given by the leading singular value of Φ, svd(Φ), or equivalently the square root of the leading eigenvalue of

$$\Phi^{\dagger}\Phi = 1 + (\mathbf{K} + \mathbf{K}^{\dagger})dt. \tag{28}$$

Using (21) with $M = 0$, we find the square root of the leading eigenvalue of (28) to be

$$\text{svd}(\Phi) = 1 + \frac{1}{2}\left(1 + \frac{1}{R_{\rho}}\right)dt, \tag{29}$$

which corresponds to growth. The associated initial singular vector is $1/\sqrt{2}(0, -1, 1)$. It is not hard to see, on examination of the original dynamical equations in (12)–(14), that a perturbation of this type will lead to initial growth when the Laplacian vanishes, despite the fact that there is no growing normal mode. The full time evolution of the norm of an initial perturbation along this singular vector periodically grows and decays with a minimum value equal to the initial perturbation.

c. Growth of a general perturbation after finite time

For heat and salt in water, $\tau = 100$ and $\sigma = 7$. With these specified, we can ask how perturbations grow after a certain amount of time as a function of finger width M and mean state density gradient ratio R_{ρ}. It is helpful to define a meaningful time scale. When $R_{\rho} = 2$, it can readily be found numerically that the maximum normal mode growth rate—that is, eig($\mathbf{K}$) evaluated at the value of M that maximizes it (cf. Fig. 2a)—is 0.104 (recall that time was nondimensionalized in section 3). The time scale we will use here is the reciprocal of this value,

$$\varepsilon = 9.6, \tag{30}$$

which is the e-folding time of the fastest-growing normal mode when $R_{\rho} = 2$.

The growth rate of a perturbation can be computed as

$$G = \frac{1}{t}\log\left[\frac{|\mathbf{v}(t)|}{|\mathbf{v}(0)|}\right], \tag{31}$$

such that for the leading normal mode

$$G = \frac{1}{t}\log[\text{eig}(\Phi)] = \frac{1}{t}\log[e^{\text{eig}(\mathbf{K})t}] = \text{eig}(\mathbf{K}), \tag{32}$$

while for the perturbation that experiences the most growth

$$G = \frac{1}{t}\log[\text{svd}(e^{\mathbf{K}t})]. \tag{33}$$

Note that G is the average growth rate since $t = 0$ rather than the instantaneous rate.

Initially the fastest-growing perturbation (leading singular vector) differs significantly from the salt finger normal mode (leading eigenvector), but as $t \to \infty$ the former converges toward the latter (which is stationary in time), as discussed in section 2. In Fig. 2a, the growth rate G is plotted for the salt finger normal mode (dotted line) and the fastest-growing perturbation (solid line) at $t = 2\varepsilon$ for $R_{\rho} = 2$. Figure 2c is an equivalent plot with $R_{\rho} = 6$. The flux ratio (22) associated with each perturbation is plotted in Figs. 2b and 2d for $R_{\rho} = 2$ and $R_{\rho} = 6$, respectively. In Fig. 2c, the solid curve rises at small M and $M = 0$ is almost the maximum of G. This is not surprising in the light of section 4b, where it was found that when $M = 0$ non-normal growth occurs even though the normal modes are stable. In fact, $M = 0$ dominates growth in much of the large R_{ρ} parameter space when t is small. It is evident in Figs. 2a and 2c that non-normal effects cause the fastest-growing perturbation to grow at a faster rate and be shifted slightly in M in comparison with the normal mode. Figures 2b and 2d show that the flux ratio of the fastest-growing perturbation is not noticeably different from the growing normal mode except at very small M, but examination of the region near the value of M that maximizes G shows that the flux ratio that maximizes growth differs significantly from the normal mode flux ratio; that is, the

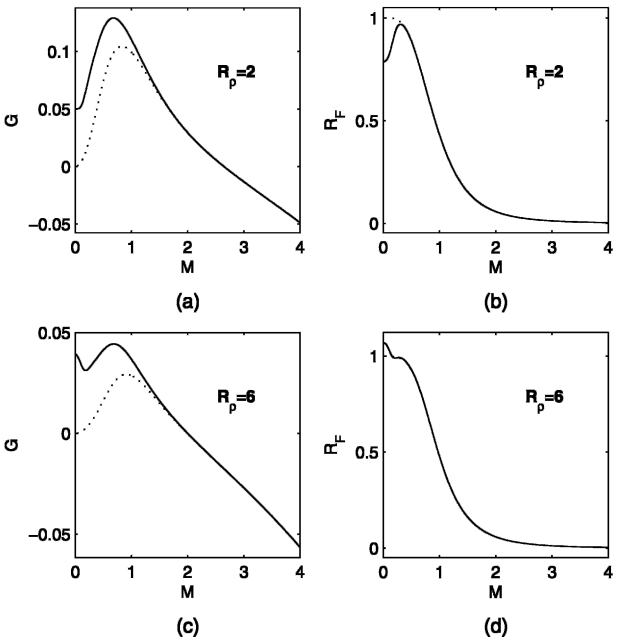

FIG. 2. (a) Growth rate G [(31) in text] at $t = 2\varepsilon$ vs horizontal wavenumber for the leading normal mode (dotted line) and the perturbation that experiences greatest growth after $t = 2\varepsilon$ (solid line) computed using singular value decomposition, with $R_{\rho} = 2$. Here ε [(30) in text] is the e-folding time of the fastest-growing finger when $R_{\rho} = 2$. (b) The flux ratios of the perturbations whose growth rates are plotted in (a). The normal mode (dotted line) and perturbation that grows most (solid line) are indistinguishable except at very small M. (c) Same as (a) but with $R_{\rho} = 6$. (d) Same as (b) but with $R_{\rho} = 6$. Note that the shift between the solid and dotted curves in the value of M that maximizes G, from (a), translates into a significant shift in R_F in (b) [equivalently in (c) and (d)].

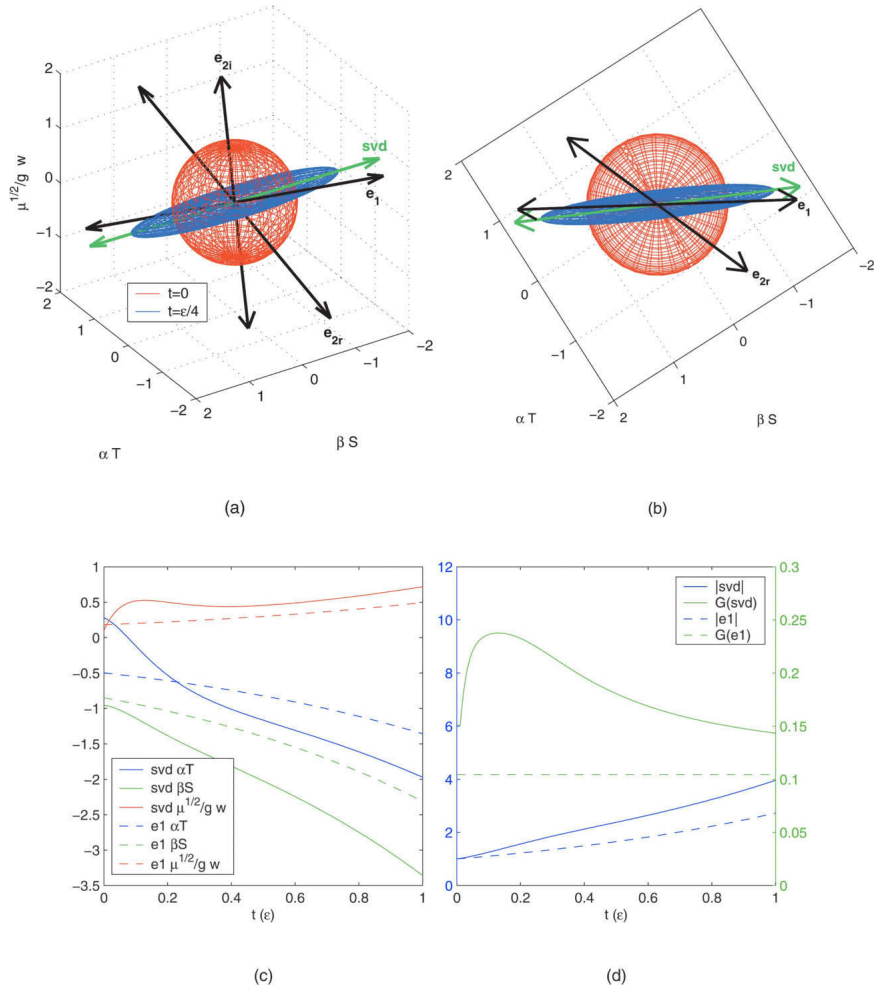

FIG. 3. Evolution of initial conditions. (a) An ensemble of initial conditions of unit norm in (T, S, w) space, represented by the red sphere, is evolved forward to $t = \varepsilon/4$ [(30) in text] with density gradient ratio $R_\rho = 2$ and wavenumber $M = 0.83$ (the value of M that maximizes normal mode growth for this value of R_ρ). The time-evolved perturbations are plotted in blue. The leading eigenvector is the black arrow $\mathbf{e_1}$. The other two eigenvectors are a complex conjugate pair with the real and imaginary parts plotted as $\mathbf{e_{2r}}$ and $\mathbf{e_{2i}}$. The perturbation that has grown the most, which is the leading final singular vector, is indicated by the green arrow labeled svd. The angle subtended by $\mathbf{e_1}$ and $\mathbf{e_{2r}}$ is 60°. Here $\mathbf{e_{2i}}$ is perpendicular to the αT axis and is at angles of 85° and 60° to $\mathbf{e_1}$ and $\mathbf{e_{2r}}$, respectively. Vector svd is 6.8° from $\mathbf{e_1}$, and the evolved perturbation along svd has grown by 73% as compared with a growth of 29% for the perturbation along the leading normal mode $\mathbf{e_1}$. (b) View of (a) from positive w axis. (c) Time series of the nondimensionalized temperature (blue), salinity (green), and vertical velocity (red) components of the perturbation that grows the most after $t = \varepsilon/4$ (solid lines) and the growing normal mode (dashed lines). (d) Time series of the norm (blue) of the perturbation in (T, S, w) space as well as the average growth rate G (green) [(31) in text] for the perturbation that grows the most after $t = \varepsilon/4$ (solid lines) and the growing normal mode (dashed lines). Note that the time range goes to $t = \varepsilon$.

difference in the value of M that maximizes G in Figs. 2a and 2c translates into a noticeable difference in R_F in Figs. 2b and 2d.

In Figs. 3a and 3b, an ensemble of initial conditions $\mathbf{v}(0)$ of unit norm in (T, S, w) space are represented by a red sphere. With $R_\rho = 2$ and $M = 0.83$ (the value that maximizes G for this R_ρ), they are evolved forward to $t = \varepsilon/4$ and plotted in blue. The normal modes are indicated by black vectors. The growing normal mode is $\mathbf{e_1}$, and the internal wave normal modes are $\mathbf{e_2}$ and $\mathbf{e_3}$, which are complex conjugates of each other; the real and imaginary parts are indicated in the figure by $\mathbf{e_{2r}}$ and $\mathbf{e_{2i}}$. The leading final singular vector, which is the perturbation that has grown the most, is marked by the

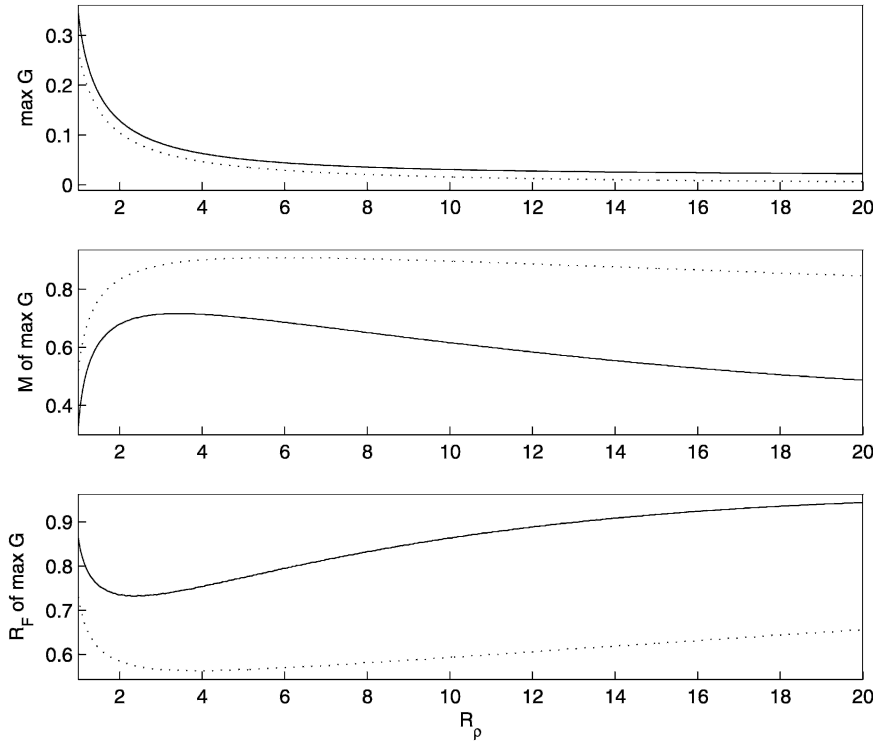

FIG. 4. Maximum growth rate G [(31) in text] and associated wavenumber and flux ratio as a function of mean state density gradient ratio. For each density gradient ratio, growth G is evaluated at a range of values of the wavenumber M and the wavenumber that leads to maximum growth is used. The solid line represents the perturbation that experiences maximum growth at $t = 2\varepsilon$ [(30) in text], and the dotted line represents the leading normal mode. Although the growth is only slightly affected by non-normal effects, the wavenumber that leads to this slightly higher growth rate differs significantly from the wavenumber that maximizes the normal mode growth. This leads to a flux ratio that also differs significantly from the normal mode, as expected by examination of Figs. 2b,d.

green arrow labeled svd. Note that this perturbation has grown significantly more than the perturbation along the growing normal mode $\mathbf{e_1}$. Figures 3c and 3d describe the time evolution between $t = 0$ and $t = \varepsilon$ for the growing normal mode and the perturbation that has grown most at $t = \varepsilon/4$. The evolution of the three components of each vector in (T, S, w) space is plotted, as well as the evolution of the norm and the growth G as defined by (31).

The peaks of the curves in Figs. 2a and 2c can readily be computed numerically, leading to a plot of the maximum growth G, as well as the wavenumber M and flux ratio R_F at which growth is maximized, as a function of mean state density gradient ratio R_ρ. These plots are presented in Fig. 4. Note that although the growth rate is only slightly affected by non-normal effects, the wavenumber that leads to the greatest growth rate differs significantly from the wavenumber that maximizes normal mode growth, which leads to a flux ratio that also differs significantly from the normal mode. The normal mode and non-normal flux ratio predictions are compared with observations in section 4e.

d. Spectrum of excited wavenumbers

We consider here the predicted spatial spectrum of salt fingers after an initial excitation of amplitude S_0 that is white in space (i.e., projects equally on all spatial modes $\mathbf{M}$), with density ratio $R_\rho = 2$, following the normal mode analysis of Schmitt (1979b).

Schmitt (1979b) examined this problem considering the growth of the unstable normal mode. In this case, the spectrum of the perturbation field at time t after the initial excitation is

$$F(M, t) = \frac{1}{2} S_0^2 |\mathrm{eig}(\Phi)|^2 = \frac{1}{2} S_0^2 e^{\{2\,\mathrm{eig}[\mathbf{K}(M)]t\}}. \quad (34)$$

To take non-normal growth effects into consideration, we replace (34) with

$$F(M, t) = \frac{1}{2} S_0^2 \{\mathrm{svd}[e^{\mathbf{K}(M)t}]\}^2. \quad (35)$$

An underlying assumption in Schmitt (1979b) is that for a given M, the initial perturbation is exactly aligned in (T, S, w) state space with the leading eigenvector. If

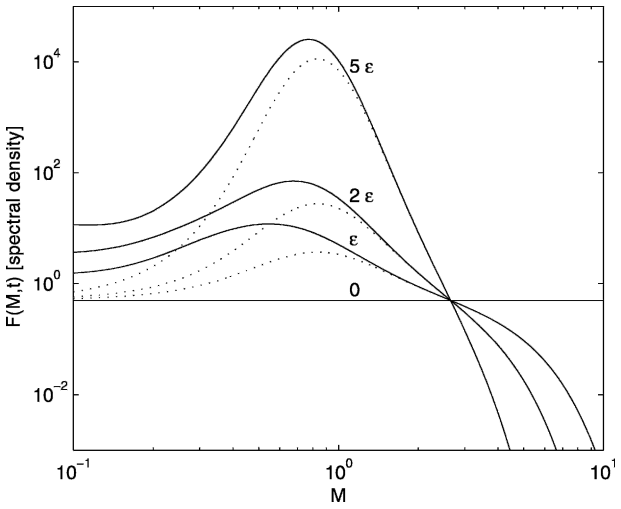

FIG. 5. Evolution of the salt finger spectrum. The spatial spectrum of salt finger perturbations excited by noise that is white in space is plotted after time 0, ε, 2ε, and 5ε, with $R_\rho = 2$, where ε is the *e*-folding time of the growing normal mode [(30) in text]. The dashed lines assume only normal mode growth, and the solid lines include non-normal growth effects (cf. Schmitt 1979b, his Fig. 7).

the perturbations randomly sample the state space in a uniform manner and the leading eigenvector strongly dominates growth, this is a reasonable assumption. Including the effect of perturbations that do not project significantly on the leading eigenvector should introduce only a minor correction. We carry this assumption here in the full non-normal analysis, assuming that the initial perturbation is aligned with the vector that will experience the greatest growth after time *t*.

Plots of the spectrum assuming only normal mode growth (dashed lines) and including non-normal effects (solid lines) are presented in Fig. 5 for several time durations after the initial perturbation. Including the effects of non-normal growth shifts the peaks slightly, enhances the peaks, and most significantly enhances the growth of wide salt fingers (small values of M), which is not surprising in light of section 4b.

e. Flux ratios: Comparison with observations

The salt finger flux ratio R_F in (22) is typically measured in experiments. Kunze (2003) compares the theoretical fastest-growing salt fingers with measurements from laboratory and numerical experiments in a plot of R_F versus the mean state density gradient ratio R_ρ (cf. Fig. 6a). It is interesting to note that the theoretical curve in Kunze (2003), from Kunze (1987), is an approximate solution as described in section 4a, and the exact solution (Schmitt 1979b) is not quite as good a fit to the observations (Fig. 6a).

When non-normal effects are included, the flux ratio of the fastest-growing finger after several *e*-folding times is a considerably worse fit to the measurements than the fastest-growing normal mode. In Fig. 6b, the flux ratio of the leading singular vector for times 2ε, 5ε, and 20ε, with ε given by (30), are plotted next to the growing normal mode. In contrast to the normal mode,

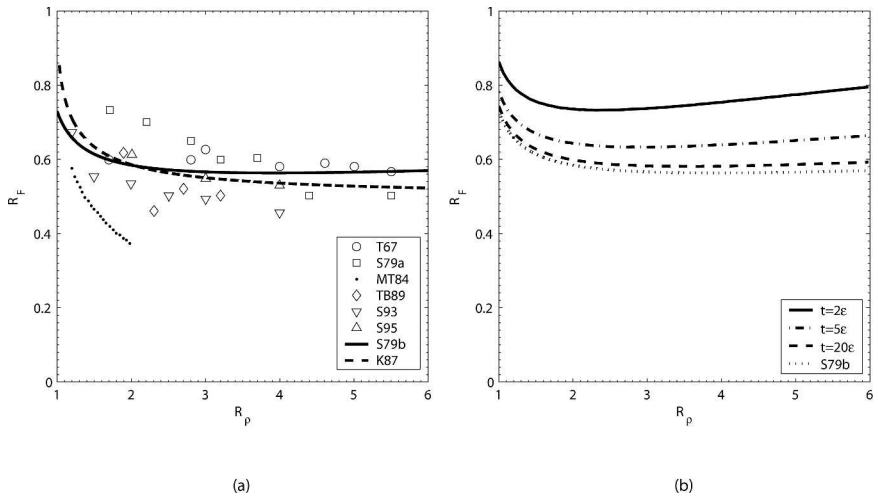

FIG. 6. Comparison of theoretical fastest-growing fingers with observed fingers. (a) Theoretical fastest normal mode growth curves of Kunze (1987) (approximate; dashed line) and Schmitt (1979b) (exact; solid line) for the salt finger flux ratio R_F vs the background density gradient ratio R_ρ are plotted next to measurements of fully developed salt fingers from laboratory experiments (McDougall and Taylor 1984; Schmitt 1979a; Taylor and Bucens 1989; Turner 1967) and numerical simulations (Shen 1993, 1995). Markers are identified in the legend by author initial and year (cf. Kunze 2003, his Fig. 3). (b) The final flux ratio for the fastest-growing perturbations including non-normal effects after times 2ε, 5ε, and 20ε plotted next to the growing normal mode (S79b). Note that when the normal mode approximation is not made, the fastest-growing curves agree less well with experimentally measured salt fingers.

the flux ratio is not constant in time for these perturbations, and only the final flux ratio is plotted for each time. The curves in Fig. 6b were computed by numerically maximizing the singular value in wavenumber M for each R_ρ and using the flux ratio of the leading final singular vector.

f. Choice of norm

In the preceding discussion, the size of a salt finger, which was used to identify the perturbation that grew most, was defined by the norm

$$|\mathbf{v}|^2 = \mathbf{v}^\dagger\mathbf{v} = (\alpha T)^2 + (\beta S)^2 + \frac{\alpha\overline{T}_z}{g}w^2 \tag{36}$$

(based on the nondimensionalization of section 3). The temperature and salinity perturbations are weighted by their contributions to the density, which is a reasonable weighting since density is the field most affected by salt fingers (although note that the temperature and salinity contribution to the norm is not the density perturbation but instead the sum of the squares of the temperature and salinity contributions to the density perturbation). Finding a reasonable weighting of the vertical velocity relative to the temperature and salinity is a more difficult issue. The weighting used in the results presented above is admittedly somewhat arbitrary, depending mostly on a convenient nondimensionalization.

The most compelling treatment of the norm might involve an initial condition with unit energy, and a separate norm that computes the density or salinity perturbation to evaluate the final magnitude of the salt finger. Both of these norms, however, introduce significant complications. The energy of a salt finger perturbation is

$$E = \frac{1}{2}\rho_T w^2 + \rho_T g h = \frac{1}{2}(\overline{\rho} + \rho)w^2 + (\overline{\rho} + \rho)g\frac{\rho}{\overline{\rho}_z}$$
$$= \overline{\rho}\left\{\frac{1}{2}w^2 + g\frac{\overline{\rho}}{\overline{\rho}_z}[(-\alpha T + \beta S) + (-\alpha T + \beta S)^2]\right\}. \tag{37}$$

Here $\rho_T = \overline{\rho} + \rho = \overline{\rho}(1 - \alpha T + \beta S)$ is the total density, $\overline{\rho}$ is the mean state density, and the perturbation scale height is $h = \rho/\overline{\rho}_z$. Terms higher than second order in the perturbation fields have been dropped in the last step. The energy given in (37) clearly cannot be expressed in terms of a vector norm using a matrix norm kernel as in (8), $|\mathbf{v}|^2 = \mathbf{v}^\dagger\mathbf{X}\mathbf{v}$, because the latter allows only quadratic multiples of the vector components while (37) includes both linear and quadratic terms. A norm that computes the density perturbation $\overline{\rho}(-\alpha T + \beta S)$, or the salinity perturbation S, is achievable using a matrix norm kernel, but the kernel is singular. While it is possible to regularize a singular norm kernel and compute singular vectors (Tziperman and Ioannou 2002), the density and salinity norms would not work for the initial condition in this problem because they have no contribution from vertical velocity w, allowing an arbitrarily large amount of energy to be stored in w in the initial condition.

The main weakness of the simple norm used in the preceding discussion is that it is unclear how the w component should be weighed against the T and S components. Conveniently, examination of Fig. 3a shows that, at least for some values of R_ρ and M, the w component of most initial perturbations quickly decays, so we may hope that the preceding results are not very sensitive to the weighting of w in the norm. Figure 7 shows that this is in fact the case. The results of Figs. 5 and 6 were reevaluated using the norm

$$|\mathbf{v}|^2 = (\alpha T)^2 + (\beta S)^2 + \frac{\alpha\overline{T}_z}{g}(\chi w)^2. \tag{38}$$

The weight of vertical velocity in the norm is reduced and then increased by a factor of 100 using $\chi = 0.1$ and $\chi = 10$, and the results are compared with the original $\chi = 1$ calculation. This was carried out using

$$\mathbf{X} = \mathbf{Y} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \chi \end{pmatrix} \tag{39}$$

in (8). Since we are using the same norm kernel for the initial condition and the final magnitude we are maximizing, and since $\mathbf{X}^{1/2}$ is Hermitian and invertible, the problem reduces to a formalism identical to (7) and (9) with the propagator Φ replaced by $\tilde{\Phi} = \mathbf{X}^{1/2}\Phi\mathbf{X}^{-1/2}$ and the singular vector $\mathbf{v}$ replaced by $\tilde{\mathbf{v}} = \mathbf{X}^{-1/2}\mathbf{v}$. The initial perturbation whose norm $\mathbf{v}^\dagger\mathbf{X}\mathbf{v}$ grows the most during time t, then, is $\mathbf{v}(t) = \mathbf{X}^{1/2}\tilde{\mathbf{v}}(t)$, where $\tilde{\mathbf{v}}(t)$ is the leading singular vector of $\tilde{\mathbf{\Phi}}$ (Farrell 1988). Figure 7 was created using these small adjustments to the calculations described in sections 4d and 4e.

5. Conclusions

Salt finger growth has traditionally been studied by examining only the leading normal mode, which is unstable. In this study, the two stable normal modes are also considered. It is found that because the dynamical operator is non-normal, the normal modes are not orthogonal, and a perturbation that is a superposition of the growing and decaying normal modes can grow faster than a perturbation that is exactly aligned with the leading normal mode. This non-normal growth is examined here using the methods of generalized stability theory (Farrell 1988; Farrell and Ioannou 1996), which have been previously applied mainly to linearly stable systems; their application to linearly unstable salt finger formation does not appear to have been examined before.

It is found here that the transient effects of non-normal growth are significant for typical salt fingers on time scales of up to more than 10 e-foldings after an initial perturbation. It is likely that by this time any

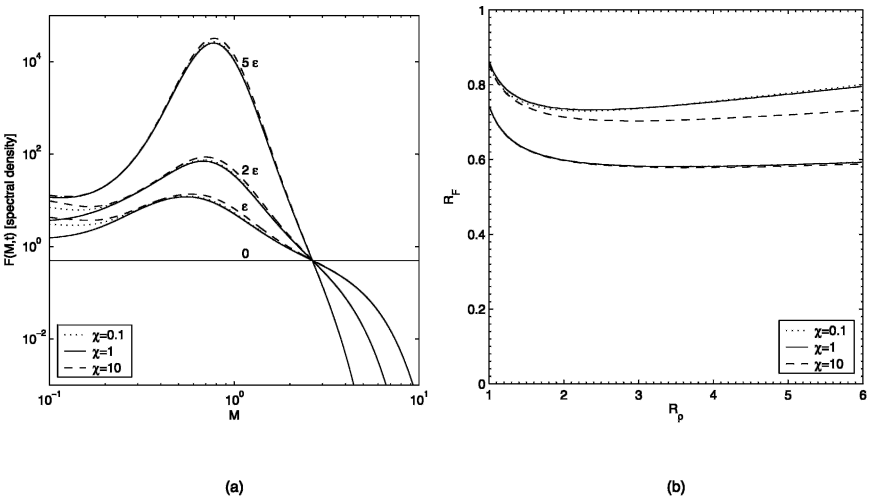

FIG. 7. Sensitivity of results to choice of norm. (a) The evolution of the salt finger spectrum including non-normal effects (Fig. 5) is reevaluated. The importance of vertical velocity in determining the norm is scaled by 1/100 (dotted line) and 100 (dashed line). Note that except for very wide fingers (small M), the curves are fairly indistinguishable from those evaluated using the original norm (solid line). (b) Final flux ratios at $t = 2\varepsilon$ and $t = 20\varepsilon$ (Fig. 6b) are plotted (solid line) next to curves evaluated using a norm where the contribution of vertical velocity is scaled by 1/100 (dotted line) and 100 (dashed line). With the exception of the dashed line for $t = 2\varepsilon$, the three curves for each time are nearly indistinguishable.

arresting mechanism that leads to steady finite amplitude salt fingers will have a significant effect. The agreement between observed fully developed salt fingers and the fastest-growing normal mode has suggested the long-time dominance of initially fastest growing fingers. The results of this study, however, suggest that the salt finger perturbations that dominate after many *e*-folding times do not closely resemble observed fully developed salt fingers. Thus the possibility of fully developed salt finger selection by the arresting mechanism rather than the initial growth dynamics is supported by this study.

The physical mechanism that allows some salt fingers to grow faster than the growing normal mode is explained by the interaction between the growing normal mode and the decaying internal wave normal modes. Based on this study, evidence of non-normal effects should be expected in numerical simulations of salt fingers. These results suggest that a simulation beginning with a system at equilibrium experiencing a small perturbation should be expected to demonstrate evidence of the flux ratio of the most developed salt fingers evolving similarly to Fig. 6b.

In this study, sinusoidal spatial variation was assumed. The singular value decomposition methods used here could be similarly applied, however, if the system were instead discretized in space. This would allow the identification of the salt finger spatial structure that grows fastest.

Acknowledgments. I thank Raymond Schmitt for introducing me to the dynamics of salt finger growth. Conversations with Raymond Schmitt and Timour Radko about the results of this research are also gratefully acknowledged. I am grateful to Eli Tziperman and Brian Farrell for reading over the manuscript and offering helpful comments and suggestions. This work was supported by the U.S. National Science Foundation, Climate Dynamics program, Grant ATM-0351123 to Eli Tziperman.

REFERENCES

Charney, J. G., 1947: The dynamics of long waves in a baroclinic westerly current. *J. Meteor.,* **4,** 135–163.

Eady, E. T., 1949: Long waves and cyclone waves. *Tellus,* **1,** 33–52.

Farrell, B., 1988: Optimal excitation of neutral Rossby waves. *J. Atmos. Sci.,* **45,** 163–172.

——, and P. J. Ioannou, 1996: Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci.,* **53,** 2025–2040.

Kunze, E., 1987: Limits on growing, finite-length salt fingers: A Richardson-number constraint. *J. Mar. Res.,* **45,** 533–556.

——, 2003: A review of oceanic salt-fingering theory. *Progress in Oceanography,* Vol. 56, Pergamon, 399–417.

McDougall, T., and J. Taylor, 1984: Flux measurements across a finger interface at low values of the stability ratio. *J. Mar. Res.,* **2,** 1–14.

Moore, A. M., and R. Kleeman, 1996: The dynamics of error growth and predictability in a coupled model of ENSO. *Quart. J. Roy. Meteor. Soc.,* **122,** 1405–1446.

——, C. L. Perez, and J. Zavala-Garay, 2002: A non-normal view of the wind-driven ocean circulation. *J. Phys. Oceanogr.,* **32,** 2681–2705.

Pedlosky, J., 1987: *Geophysical Fluid Dynamics.* 2d ed. Springer-Verlag, 710 pp.

Penland, C., and P. D. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate,* **8,** 1999–2024.

Ruddick, B., and A. E. Gargett, 2003: Oceanic double-infusion:

Introduction. *Progress in Oceanography,* Vol. 56, Pergamon, 381–393.
Schmitt, R. W., 1979a: Flux measurements on salt fingers at an interface. *J. Mar. Res.,* **37,** 419–436.
——, 1979b: Growth-rate of super-critical salt fingers. *Deep-Sea Res.,* **26,** 23–40.
——, 1994a: Double diffusion in oceanography. *Annu. Rev. Fluid Mech.,* **26,** 255–285.
——, 1994b: Triangular and asymmetric salt fingers. *J. Phys. Oceanogr.,* **24,** 855–860.
Shen, C. Y., 1993: Heat-salt finger fluxes across a density interface. *Phys. Fluids,* **5A,** 2633–2643.
——, 1995: Equilibrium salt-fingering convection. *Phys. Fluids,* **7,** 706–717.
Stern, M. E., 1960: The salt-fountain and thermohaline convection. *Tellus,* **12,** 172–175.
——, 1975: *Ocean Circulation Physics.* Academic Press, 246 pp.
Taylor, J., and P. Bucens, 1989: Laboratory experiments on the structure of salt fingers. *Deep-Sea Res.,* **36A,** 1675–1704.
Turner, J. S., 1967: Salt fingers across a density interface. *Deep-Sea Res.,* **14,** 599–611.
Tziperman, E., and P. J. Ioannou, 2002: Transient growth and optimal excitation of thermohaline variability. *J. Phys. Oceanogr.,* **32,** 3427–3435.