priate because fluid flux across the vertical fracture–matrix boundary will be zero for an impermeable matrix block, but the vertical boundary will not necessarily be maintained at a constant concentration. These results lead to an estimate for the critical concentration difference for the onset of mode 2A convection

$$\Delta C_{\mathrm{crit}2A} = \frac{500.5 D_{\mathrm{av}} \nu H}{g\beta b^4}. \tag{5}$$

[17] The ratio of the two critical concentration differences is

$$\frac{\Delta C_{\mathrm{crit1}}}{\Delta C_{\mathrm{crit}2A}} = O\left(\lambda b/H^2\right). \tag{6}$$

[18] In the Landau order notation, we say $f(z) = O[g(z)]$ in some domain D if there exists a positive constant K such that $|f| \le K|g|$ for all z in D. Thus, roughly speaking, f has the same order of magnitude as g. The ratio in equation (6) is expected to be small compared to unity. We conclude that mode 1 is more unstable than mode 2A, and thus the latter is generally of lesser importance.

[19] However, the situation is dramatically different for the case of convection parallel to the fracture plane (mode 2B). In this case, and for the subcase of boundaries at constant concentration, Ra_{cr} is given by

$$Ra_{\mathrm{cr}} = \left(\frac{\pi^2}{4} + \alpha^2\right)^2, \tag{7}$$

where α is the wave number in the horizontal direction parallel to the fracture boundary planes. Equation (7) is a monotonically increasing function of α and so attains its minimum as α tends to zero, the minimum value being $\pi^4/16$.

[20] On the other hand, for the subcase of constant flux (and in particular zero fluid flux) boundaries, Ra_{cr} is still attained as α tends to zero, but the minimum value is zero. In fact, for small values of α,

$$Ra_{\mathrm{cr}} = O\left(3\alpha^2\right). \tag{8}$$

[21] Thus the overall critical value can be arbitrarily small, being limited only by the horizontal extent of the fractures:

$$\Delta C_{\mathrm{crit}2B} = \left(\frac{16 D_{\mathrm{av}} \nu H}{g\beta b^4}\right)\alpha^2. \tag{9}$$

[22] For example, if the fractures extend a distance L, then one has $\alpha = 2\pi/(2L/b) = \pi b/L$, and so

$$\Delta C_{\mathrm{crit}2B} = \frac{16\pi^2 D_{\mathrm{av}} \nu H}{g\beta b^2 L^2}. \tag{10}$$

[23] The ratio of the two critical concentration differences for mode 1 and mode 2B convection is

$$\frac{\Delta C_{\mathrm{crit1}}}{\Delta C_{\mathrm{crit}2B}} = O\left(L^2/b^2\right). \tag{11}$$

[24] One would expect that this expression would be normally considerably greater than unity. Then the overall conclusion is that mode 2B (with a typical streamline lying in a plane parallel to the fracture plane boundaries) is the favored mode of convection. Hence, the appropriate estimate for the critical concentration difference is that given by equation (10) for the mode of convection most favored amongst possible intrafracture and interfracture convection modes.

4. Discussion and Conclusion

[25] It is apparent that the most likely mode of convection in a low-permeability layer is mode 2B, with free convection occurring parallel to the plane of the fracture. The critical concentration difference required for the onset of convection in each mode determines the likelihood of each mode occurring in realistic hydrogeologic settings. If we assume a matrix porosity of ε = 0.05 and an aqueous diffusion coefficient of $D = 10^{-9}$ m^2 s^{-1}, then the effective matrix diffusion coefficient (a reasonable approximation for the average solute diffusivity) is $D_{\mathrm{av}} = 5 \times 10^{-11}$ m^2 s^{-1}. Other relevant parameters are $\beta = 0.7$ and $\nu = 10^{-6}$ m^2 s^{-1}. In the following paragraph, we define minimum and maximum bounds on fracture aperture, fracture spacing, fracture length, and shale layer thickness to estimate the range of critical concentration differences that would be required for convection to occur in each mode. The parameters chosen here are intended to be illustrative and representative of those which may be encountered more generally in fractured rock hydrogeology rather than those which may specifically exist in the Gulf of Mexico Basin. The present simple and idealized analysis is intended to be generalized to low-permeability media rather than limited only to the specific case of the Gulf of Mexico Basin. The intention here is to demonstrate using reasonable parameters in a simplified analysis that free convection may be plausible in an otherwise low-permeability layer when suitable combinations of salinity differences, shale layer thickness, and fracture properties exist. These findings are therefore applicable to, but not limited to, the Gulf of Mexico case study.

[26] The ranges for these demonstrative and representative parameters are chosen as follows:

Fracture aperture

$$b_{\min} = 10^{-6}\ \mathrm{m}; \quad b_{\max} = 10^{-3}\ \mathrm{m}$$

Fracture spacing

$$\lambda_{\min} = 0.1\ \mathrm{m}; \quad \lambda_{\max} = 100\ \mathrm{m}$$

Fracture length

$$L_{\min} = 1\ \mathrm{m}; \quad L_{\max} = 100\ \mathrm{m}$$

Shale layer thickness

$$H_{\min} = 1\ \mathrm{m}; \quad H_{\max} = 100\ \mathrm{m}$$

[27] We begin by considering the dominant convection mode 2B. The critical concentration difference required for convection can be calculated using equation (10) for a large range of possible choices for the fracture geometry factor for mode 2B convection as $f_{\mathrm{mode}2B} = H/(b^2 L^2)$. Using the range of parameters stated above for shale layer thickness, fracture aperture, and fracture length, we find that fracture geometry factors f range between 10^2 and 10^{14} m^{-3}. Substitution into equation (10) gives critical dimensionless concentration differences required for convection in the approximate range 10^{-13}–10^{-1}. At the lowest end of the range, this is an extremely small concentration difference and suggests that convection would be expected to occur parallel to the fracture for almost any nonzero concentration difference. At the highest end of the range, this is equivalent to a salinity of 100,000 ppm, which is a realistic concentration difference, especially in the Gulf of Mexico setting. On this basis, we can conclude that mode 2B convection is likely to occur in most hydrogeologic settings. We do not have to "beg the data" for this mode of free convection to occur.

[28] The calculation of critical concentration differences for mode 1 and mode 2A convections may seem somewhat superfluous given the ease by which mode 2B convection is expected to occur and the fact that it is expected to be the dominant mode. However, it is useful to examine under what conditions other convective modes can occur and their associated likelihood. We calculate the concentration difference required for mode 1 convection. We now define a new fracture geometry factor for mode 1 convection as $f_{\mathrm{mode}1} = \lambda/(b^3 H)$. Using the range of parameters stated above for fracture separation, fracture aperture, and shale layer thickness, we find that fracture geometry factors f range between 10^6 and 10^{20} m^{-3}. Equation (3) then gives critical dimensionless concentration differences required for convection in the approximate range 10^{-9}–10^5. At the lowest end of the range, this is an extremely small concentration difference and suggests that interfracture convection would be expected to occur on the layer scale for almost any nonzero concentration difference. At the highest end of this range, the required concentration differences are not possible. Concentration differences in the range 0.3–0.5 represent the upper limit of those that are physically allowed. This is because the precipitation of many salts (such as chlorides, carbonates, and sulphates) occurs at or below this stated concentration range of 0.3–0.5. As a result, a substantial range of the required concentration differences for free convection is not realistic. Compared with mode 2B convection, the range of physically permissible concentration differences is substantially smaller for mode 1 convection. This result suggests that convection in mode 1 is likely in many (but not all) situations. When it occurs, it is likely to occur in combination with mode 2B convection.

[29] Finally, we calculate the concentration difference required for mode 2A convection. We define a fracture geometry factor for mode 2A convection as $f_{\mathrm{mode}2A} = H/b^4$. Using the range of parameters stated above for shale layer thickness and fracture aperture, we find that fracture geometry factors f range between 10^{12} and 10^{26} m^{-3}. Equation (5) then gives critical dimensionless concentration differences required for convection in the approximate range 10^{-3}–10^{11}. At the lowest end of this range, a very modest concentration difference of 1000 ppm is required for the onset of convection. This is a very small concentration difference and would additionally require, for example, that the low-permeability layer contain fractures with apertures of about 1 mm in a layer of about 1 m thickness. Low-permeability layers can be substantially thicker, and fracture apertures are often thinner, and free convection in this mode would require concentration differences far greater than that which is physically possible in many hydrogeologic systems. Compared with mode 2B and mode 1 convection, the range of physically permissible concentration differences is substantially smaller for mode 2A convection and suggests that mode 2A convection is the least likely mode to occur.

[30] The major outcome of this paper is a summary of plausible arguments for the likelihood of free convection, potentially occurring in even low-permeability sediments such as those in the Gulf of Mexico Basin. Although completely based on theoretical evaluations and assessments regarding criteria for the onset of convection by using known analytical relationships of critical Rayleigh numbers, the paper offers a number of theoretical findings. The important generalized finding of this study is that all modes of free convection (parallel to the fracture plane, perpendicular to the fracture plane, and convection between fractures on the larger layer scale) are theoretically plausible for reasonable hydrogeologic parameters but that the dominant (and most likely) mode of convection is expected to be intrafracture convection parallel to the fracture plane (mode 2B). The least likely mode is intrafracture convection perpendicular to the fracture (mode 2A). The results suggest that episodes of density-driven flow may not be uncommon in the thick shale sequences, such as in the Gulf of Mexico Basin. While *Sharp et al.* [2001] showed that convection driven by salinity differences was possible when shale permeabilities are near the upper end of the expected range of values, these new results suggest that density-driven flow may be significantly more widespread because of the presence of fractures in the shale layers and the ease by which mode 2B convection may be expected to occur. Shales with matrix permeabilities at the lower end of *Neuzil*'s [1994] range originally thought not to permit convection may contain the fracture aperture and separation spatial scales required for convection to be a plausible fluid flow and solute transport mechanism. Intrafracture convection parallel to fracture planes is likely to permit free convection at low concentration differences. The propensity for density-driven convection is strongly influenced by permeability heterogeneities that may otherwise be insignificant in regional flow calculations or in models of basin evolution. This analysis provides conditions for the onset of convection only; further analyses are required to examine the temporal patterns of the growth and/or decay of instabilities associated with free convection once they are established. The density contrast is ideally assumed to be stable over very long geologic time. Eventually, the convection processes will be very slow, and other dynamic processes will dominate them in the Gulf of Mexico Basin. However, this transient effect is also expected to depend upon other largely unknown factors such as the frequency and magnitude of episodic expulsion of saline brines from depth. The important consequence of these findings is that free convection is expected to be pervasive (at least at some

point in time) when subvertical/vertical fracturing is present in a low-permeability layer. Such fracturing is expected to eventually dissipate the salinity inversions over timescales that are presently unknown. Similarly, if a salinity inversion is observed above or across a low-permeability layer, it is likely to be either a short-lived (transient) phenomenon on geologic timescales or a saline fluid that lies above a low-permeability layer that does not contain significant vertical fracturing. These findings are important in understanding both salinity inversion data and the possible solute transport processes that may occur in low-permeability shales such as in the Gulf of Mexico Basin. Clearly, more work is required to understand the transient persistence of such phenomena as well as to more intimately and explicitly connect these phenomena with the Gulf of Mexico case study.

[31] Some general remarks can also be made about the numerical simulation of free convection in fractured low-permeability media. Our results suggest that analyses not conducted in three dimensions underestimate the likelihood of the occurrence of convection. The simulation of variable density flow phenomena in 2-D fracture networks [e.g., *Shikaze et al.*, 1998; *Graf and Therrien*, 2007] does not capture the dominant mode of convection expected to occur in the fractured rock system (i.e., parallel to the fracture planes themselves) since only the interfracture convection mode is usually considered in that network-modeling framework. This is an important finding that does not appear to have been stated explicitly in previous published numerical modeling literature. It was, however, demonstrated numerically by *Shi* [2005], who observed that when free convection initiates, convection cells occur on the fracture plane with axes parallel (normal) to the fracture plane. *Shi* [2005] concluded that two-dimensional numerical models of flow and transport normal to fracture strike are unable to capture the most important pattern of convective motion. This observation is consistent with the results of our present analysis.

[32] It is recognized that the explicit simultaneous simulation of free convective transport through a network of fractures, and within the fractures themselves, is computationally demanding for a fractured rock flow and transport numerical model. Although the results of this study are suggestive of the need for a fully 3-D schematization to model recirculating convection modes, this is not always necessary and may represent an overkill for a number of cases. It is therefore important to clearly discuss the consequences of modeling free convective phenomena in these systems with reduced spatial dimensions. Cellular convection can be modeled in two dimensions if the vertical fracture is discretized by 2-D fracture elements and the diffusive flux from the matrix is handled by a source condition in the direction normal to the fracture plane. This has previously been preferred in modeling hot dry rock recirculation currents in single fractures [e.g., *Kolditz and Diersch*, 1993] and can be analogously adapted to salinity-driven free convection in fractures. Indeed, a 2-D modeling approach similar to that presented by *Kolditz and Diersch* [1993] would be extremely useful for analyzing intrafracture convection (mode 2B). Furthermore, a network of fractures discretized by 1-D fracture elements in a 2-D vertical cross section would also be able to model the interfracture convection (mode 1). It is clear that only the combination of all free convection modes would require a full 3-D modeling approach. As an intermediate step, 2-D models with 2-D and/or 1-D fracture elements appear to be preferable for analyzing the most favored convection modes before a fully 3-D model is constructed. We therefore emphasize that the role of 2-D models should not be underestimated. This is especially important since intra-fracture convection parallel to the fracture plane (mode 2B) has been shown in this analysis to be the most likely convection mode and is only a 2-D process.

[33] Previous analyses by *Simmons et al.* [2001], *Simmons* [2005], and *Nield and Simmons* [2007] have suggested that heterogeneity in geologic properties is critical in controlling the onset, growth, and/or decay of free convection in a hydrogeologic setting. A particular controlling feature of that geologic heterogeneity is its structure and interconnectedness. Future work is required to examine the spatiotemporal patterns of dense plume migration through significantly more complex and realistic 3-D fracture networks. It is entirely plausible that heterogeneity at another level (for example, larger-scale variations in shale permeability and variability in fracture spacing, aperture, orientation, and connectedness) will be vital in controlling whether or not instabilities will occur in field-based settings and how persistent they may be over larger spatial scales and longer temporal scales. These are largely unresolved matters in current literature. These previous theoretical findings suggest that there is a strong interplay between geological heterogeneity and free convective phenomena that means that free convective processes are not easily amenable to prediction. This also suggests that the analysis presented here may best be used as a semiquantitative guide to which modes of free convection may be expected, a priori, to be dominant. What is critical here is that we demonstrate using analytical relations and plausible parameters for a fractured shale system that free convection is a theoretically possible solute transport process in thick-shale sequences and that the dominant and most likely mode of free convection is parallel to vertical fracture planes.

[34] As discussed previously by *Simmons* [2005], a critical challenge that remains for this field of research is that there are limited observations of unstable dense plume phenomena and free convection in field-scale settings. *Simmons* [2005] notes that research papers often quote secondary evidence for their existence (e.g., the salt deficit in the salt lake may be accounted for by the slow downward convection of dense water; numerical experiments demonstrate the existence of a convection cell; abundant data indicate high fluid fluxes consistent with density-driven flow; and the system Rayleigh number is greater than the critical value required and, therefore, convection is assumed to exist) rather than direct or primary evidence of the free convective processes themselves. We are acutely aware that a major challenge remains to verify theoretical findings such as those presented here but also more broadly throughout free convection research literature, in field-based settings. This is vital to confirming or refuting the free convective hypothesis in hydrogeologic settings. Underpinned by several decades of theoretical and modeling analyses, the free convection hypothesis has a solid theoretical basis but to date has been neither explicitly nor completely observed through robust data collection in hydrogeologic settings

using primary hydrogeologic evidence of groundwater flow and salinity (i.e., convective circulations by measurement of flow directions and rates, temporal measurements of descending dense plumes, or mapping of spatial fingering patterns by detailed measurements of groundwater salinity). Nonetheless, simple and idealized analyses based upon theoretical Rayleigh stability criteria such as those presented in this paper may provide useful guidance for both future numerical modeling and field-based experimentation.

Notation

- b fracture aperture (L).
- C fluid concentration expressed as a mass fraction (M_S M^{-1}).
- C_U maximum value of concentration expressed as solute weight relative to weight of solution (M_S M^{-1}).
- C_L minimum value of concentration expressed as solute weight relative to weight of solution (M_S M^{-1}).
- D apparent molecular diffusivity of solutes in solution (L^2 T^{-1}).
- D_{av} average solute diffusivity (L^2 T^{-1}).
- f fracture geometry factor (L^{-3}).
- g acceleration due to gravity (L T^{-2}).
- H shale layer thickness (L).
- K_{av} average intrinsic permeability (L^2).
- L fracture length (L).
- Ra Rayleigh number of the system (dimensionless).
- β coefficient of density variability, $\rho_0^{-1}(\partial\rho/\partial C)$ (dimensionless).
- ΔC concentration difference between C_U and C_L (M_S M^{-1}).
- α wave number in the horizontal direction parallel to the fracture boundary plane.
- ε porosity of matrix block (dimensionless).
- λ wavelength (spacing) of the periodic fracture set (L).
- ρ fluid density (M L^{-3}).
- ν kinematic viscosity of the fluid (L^2 T^{-1}).

[35] **Acknowledgments.** Acknowledgment is made to the United States Department of Energy, Geosciences Research Program (grant DE-FG03-97ER14772), and to the Petroleum Research Fund of the American Chemical Society (grant 38949-AC9) for partial support of this research. We acknowledge the *Water Resources Research* Editor and Associate Editor who handled this manuscript as well as two anonymous reviewers for their constructive comments which assisted us in revising the paper.

References

Bjorlykke, K., A. Mo, and E. Palm (1988), Modeling of thermal convection in sedimentary basins and its relevance to diagenetic reactions, *Mar. Pet. Geol.*, *5*, 338–351.

Caltagirone, J. P. (1982), Convection in a porous medium, in *Convective Transport and Instability Phenomena*, edited by J. Zierep and H. Oertel, pp. 199–232, Braun, Karlsruhe, Germany.

Gershuni, G. Z., and E. M. Zhukhovitskii (1976), *Convective Stability of Incompressible Fluids*, 300 pp., Keter, Jerusalem.

Graf, T., and R. Therrien (2007), Variable-density groundwater flow and solute transport in irregular 2D fracture networks, *Adv. Water Resour.*, *30*, 455–468.

Kolditz, O., and H.-J. Diersch (1993), Quasi-steady-state strategy for numerical simulation of geothermal circulation in hot dry rock fractures, *Int. J. Non Linear Mech.*, *28*(4), 467–481.

Land, L. S. (1991), Evidence for vertical movement of fluids, Gulf Coast sedimentary basin, *Geophys. Res. Lett.*, *18*, 919–922.

Lapwood, E. R. (1948), Convection of fluids in a porous medium, *Proc. Cambridge Philos. Soc.*, *44*, 508–521.

Love, A. J., C. T. Simmons, and D. A. Nield (2007), Double-diffusive convection in groundwater wells, *Water Resour. Res.*, *43*, W08428, doi:10.1029/2007WR006001.

Malkovsky, V. I., and A. A. Pek (1997), Conditions for the onset of thermal convection of a homogeneous fluid in a vertical fault, *Petrology*, *5*(4), 381–387.

McKenna, T. E., and J. M. Sharp, Jr. (1997), Subsurface temperatures, fluid pressures, and salinities in the Rio Grande Embayment, Gulf of Mexico Basin, USA, *Proc. Int. Geol. Congr., 30th*, *8*, 263–274.

Morton, R. A., and L. S. Land (1987), Regional variations in formation water chemistry, Frio Formation (Oligocene), Texas Gulf Coast, *AAPG Bull.*, *71*, 191–206.

Murphy, H. D. (1979), Convective instabilities in vertical fractures and faults, *J. Geophys. Res.*, *84*, 6121–6130.

Neuzil, C. E. (1994), How permeable are clays and shales?, *Water Resour. Res.*, *30*(2), 145–150.

Nield, D. A., and C. T. Simmons (2007), A discussion on the effect of heterogeneity on the onset of convection in a porous medium, *Transp. Porous Media*, *68*, 413–421.

Prasad, A., and C. T. Simmons (2003), Unstable density-driven flow in heterogeneous porous media: A stochastic study of the Elder [1967b] "short heater" problem, *Water Resour. Res.*, *39*(1), 1007, doi:10.1029/2002WR001290.

Ronen, D., Y. Yechieli, and A. Kribus (1995), Buoyancy-induced flow of a tracer in vertical conduits, *Water Resour. Res.*, *31*(5), 1167–1174.

Schincariol, R. A., F. W. Schwartz, and C. A. Mendoza (1997), Instabilities in variable density flows: Stability analyses for homogeneous and heterogeneous media, *Water Resour. Res.*, *33*(1), 31–41.

Sharp, J. M., Jr., et al. (1988), Diagenetic processes in northwest Gulf of Mexico sediments, in *Diagenesis, II*, edited by G. V. Chilingarian and K. H. Wolf, chap. 1, pp. 43–113, Elsevier, New York.

Sharp, J. M., Jr., T. R. Fenstemaker, C. T. Simmons, T. E. McKenna, and J. K. Dickinson (2001), Potential thermohaline convection in a shale-rich sedimentary basin: Example from the Gulf of Mexico Basin in south Texas, *AAPG Bull.*, *85*, 2089–2110.

Shi, M. (2005), Characterizing heterogeneity in low-permeability strata and its control on fluid flow and solute transport by thermohaline free convection, Ph.D. thesis, 229 pp., Univ. of Tex. at Austin, Austin.

Shikaze, S. G., E. A. Sudicky, and F. W. Schwartz (1998), Density-dependent solute transport in discretely-fractured geologic media: Is prediction possible?, *J. Contam. Hydrol.*, *34*, 273–291.

Simmons, C. T. (2005), Variable density groundwater flow: From current challenges to future possibilities, *Hydrogeol. J.*, *13*(1), 116–119.

Simmons, C. T., T. R. Fenstemaker, and J. M. Sharp Jr. (2001), Variable-density groundwater flow and solute transport in heterogeneous porous media: Approaches, resolutions and future challenges, *J. Contam. Hydrol.*, *52*, 245–275.

Weatherill, D., C. T. Simmons, C. I. Voss, and N. I. Robinson (2004), Testing density-dependent groundwater models: Two-dimensional steady state unstable convection in infinite, finite and inclined porous layers, *Adv. Water Resour.*, *27*, 547–562.

Wood, J. R., and T. A. Hewett (1984), Reservoir diagenesis and convective fluid flow, in *Clastic Diagenesis*, edited by D. A. McDonald and R. C. Surdam, *AAPG Mem.*, *37*, 99–111.

D. A. Nield, Department of Engineering Science, University of Auckland, Auckland, 92019, New Zealand.

J. M. Sharp Jr., Department of Geological Sciences, University of Texas at Austin, Austin, TX 78712, USA.

C. T. Simmons, School of Chemistry, Physics and Earth Sciences, Flinders University, GPO Box 2100, Adelaide, SA 5001, Australia. (craig.simmons@flinders.edu.au)