

## Gravothermal Collapse of Self-Interacting Dark Matter Halos and the Origin of Massive Black Holes

Shmuel Balberg\* and Stuart L. Shapiro†

*Physics Department, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801*  
(Received 29 October 2001; published 22 February 2002)

Black hole formation is an inevitable consequence of relativistic core collapse following the gravothermal catastrophe in self-interacting dark matter (SIDM) halos. Very massive SIDM halos form supermassive black holes (SMBHs)  $\geq 10^6 M_\odot$  directly. Smaller halos believed to form by redshift  $z = 5$  produce seed black holes of  $(10^2 - 10^3) M_\odot$  which can merge and/or accrete to reach the observational SMBH range. This scenario for SMBH formation requires no baryons, no prior star formation, and no other black hole seed mechanism.

DOI: 10.1103/PhysRevLett.88.101301

PACS numbers: 95.35.+d, 97.60.Lf, 98.62.Ai, 98.62.Gq

Black holes in the centers of galaxies are inferred to have masses of  $(10^6 - 10^9) M_\odot$ , and, to date, dozens of candidates have been discovered [1]. Black holes in this mass range are the likely power sources in quasars and active galactic nuclei [2–4]. Recent observational data have identified a strong correlation between the inferred black hole mass and the stellar velocity dispersion in the host galaxy bulge [5]; a correlation has also been found between the black hole mass and the mass of the bulge [6]. Both correlations suggest that formation and evolution of the central black hole and the bulge of the host galaxy may be closely related.

The origin of these supermassive black holes (SMBHs) is uncertain. Proposed scenarios include the collapse of a supermassive star [7], possibly built up by collisions and mergers of ordinary stars [8]. Another scenario involves the collapse of the core of a dense relativistic cluster (e.g., neutron stars or stellar mass black holes) following gravothermal evolution. The cluster core evaporates mass to an extended halo on a gravitational scattering *relaxation* time scale, while the core density and velocity dispersion grow (the “gravothermal catastrophe” [9,10]). Zel’dovich and Podurets originally conjectured [11] and Shapiro and Teukolsky [12] subsequently demonstrated that a nearly collisionless gas in virial equilibrium like a star cluster experiences a radial instability to collapse on a *dynamical* time scale when its core becomes sufficiently relativistic. As the instability sets in, the core and its immediate surroundings undergo catastrophic collapse to a black hole and the ambient halo settles into a new dynamical equilibrium state around the central hole.

Here we report that the formation of a black hole in the center of a galactic halo is a natural and inevitable consequence of the gravothermal evolution of self-interacting dark matter (SIDM) halos. The possibility that dark matter particles are “self-interacting” has been revived recently [13]. Studies of SIDM via  $N$ -body simulations [14,15] and via a gravothermal approach [16] have confirmed that SIDM halos are more consistent with observations: They exhibit a flat density core, rather than a cuspy one that

arises for cold dark matter [17,18]. An isolated SIDM halo evolves gravothermally, since the thermal relaxation time scale due to collisions is shorter than the typical ages of cosmological halos. Unlike a star cluster, a SIDM halo retains an appreciable core mass as it evolves towards the relativistic instability [16].

*I. Gravothermal evolution and collapse of a SIDM halo.*—The structure of a relaxed SIDM halo can be defined by the central mass density  $\rho_c$ , the (one-dimensional) velocity dispersion  $v_c$ , and the total halo mass  $M_{\text{tot}}$ . For a given cosmology, the background density,  $\bar{\rho}$ , is a function only of redshift. We assume that the central density contrast  $x \equiv \rho_c(t=0)/\bar{\rho}$  at halo virialization is independent of redshift and halo mass, which is consistent with the simulations of Davè *et al.* [15] and recent work on truncated isothermal spheres by Shapiro, Iliev, and Raga [19]. The independent parameters of a SIDM halo are then  $M_{\text{tot}}$ ,  $x$ , and  $z_0$ , the redshift of virialization. Define mass and radius scales,  $M_0$  and  $R_0$ , respectively, where  $M_0 = 4\pi R_0^3 \rho_c$  and  $v_c^2 = GM_0/R_0$ , and also the ratio  $g \equiv M_{\text{tot}}/M_0$ , which is a unique function of  $x$ . The central velocity satisfies

$$v_c^2(t=0) = (4\pi)^{1/3} G [M_{\text{tot}}/g(x)]^{2/3} [x\bar{\rho}(z_0)]^{1/3}. \quad (1)$$

*Evolutionary Stages:* The properties in the halo depend critically on the ratio of the collisional mean free path,  $\lambda$ , to the gravitational scale height,  $H$ :

$$\lambda = (\rho\sigma)^{-1}, \quad H = \left( \frac{v^2}{4\pi G\rho} \right)^{1/2}, \quad (2)$$

where  $\rho$  is the local mass density,  $v$  is the local one-dimensional velocity dispersion, and  $\sigma$  is the SIDM cross section per unit mass. In the long mean free path (lmfp) region  $\lambda \gg H$ , particles perform several orbits between collisions; in the short mean free path (smfp) region  $\lambda \ll H$ , particle motion is severely restrained, and heat transfer proceeds through diffusion, as in a fluid. Simulations show that a halo core must form in the lmfp regime [13]; otherwise the particles generate strong shocks, whereby entropy loss to the halo destroys the central core structure [20].

The requirement that at formation the core satisfies  $(\lambda/H)_c \geq 1$  yields an *upper* limit for the halo mass:

$$M_{\text{tot}} \leq 4\pi[\bar{\rho}(z_0)x]^{-2}g(x)\sigma^{-3}. \quad (3)$$

Black hole formation: Gravothermal evolution of a SIDM halo is shown in Fig. 1 [16]. While the extended halo ( $\rho \propto r^{-2.19}$ ) remains in the lmf limit, the core becomes increasingly dense. The core eventually bifurcates into a smfp inner part which is fluidlike ( $\lambda \ll H$ ), and a lmf outer part which is nearly static. The transition region corresponds to  $\lambda/H \sim 1$ . It is the inner core and its immediate surroundings which collapse to a black hole following the onset of the relativistic instability, leaving the outer core and extended halo in dynamical equilibrium about the central hole [12].

Evaporation of mass due to collisions from a lmf core into the halo is rapid with  $d \log M_c / d \log(v_c^2) \approx -4.27$  [16]. Once the core becomes very dense, with central  $\lambda/H \geq 100$ , evaporation is limited to a surface effect, which is much less efficient. We find [16] that, if the core of a SIDM halo forms with  $\lambda/H \geq 1$ , its mass decreases by about 1 order of magnitude prior to reaching the smfp limit (while the central velocity hardly changes), after which the core mass decreases according to  $d \log M_c / d \log(v_c^2) \approx -0.85$ . This estimate gives the

inner core mass  $M_{\text{coll}}$  at the onset of relativistic instability ( $v_c \approx c/3$ ):

$$\log_{10}\left(\frac{M_{\text{coll}}}{M_c(0)}\right) \approx -1 + 0.85 \log_{10}\left(\frac{v_c^2(0)}{10^{10} \text{ km}^2 \text{ s}^{-2}}\right). \quad (4)$$

Core lifetime: The lifetime of the core until collapse is [16]  $t_{\text{coll}} \approx 290t_r(0)$ , where  $t_r(0)$  is the collision relaxation time scale in the core at formation,

$$t_r = \frac{1}{a}[\rho_c(t=0)v_c(t=0)\sigma]^{-1}, \quad (5)$$

and  $a \approx 2.26$  for purely elastic, hard-sphere collisions and a Maxwell-Boltzmann velocity distribution. Requiring that the core of a halo formed at redshift  $z_0$  collapses by redshift  $z_1$  yields a *lower* limit on the mass of the SIDM halo:

$$M_{\text{tot}} \geq \left(\frac{290}{a}\right)^3 (4\pi G^3)^{-1/2} \frac{[\bar{\rho}(z)x]^{-7/2}g(x)\sigma^{-3}}{[t(z_1) - t(z_0)]^3}. \quad (6)$$

*II. A direct scenario for SMBH formation.*—Since the core virializes with nonrelativistic velocities,  $v_c \leq (10^2 - 10^3) \text{ km s}^{-1}$ , the core mass at collapse will be several orders of magnitude smaller than  $M_c(0) \approx 10^{-2}M_{\text{tot}}$  according to Eq. (4). Consequently, gravothermal evolution directly culminating in a SMBH with a mass above  $10^6 M_\odot$  requires that the total mass of the progenitor SIDM halo exceeds  $10^{12} M_\odot$ . Equations (3) and (6) determine the range of halo masses formed at redshift  $z_0$  that undergo core collapse by redshift  $z_1$ . These masses increase for smaller values of the ratio  $x$ , i.e., for smaller central density contrasts at halo formation.

We examine this scenario by setting  $x$  to the relatively small value of  $x = 1.8 \times 10^4$ , which leads to  $g \approx 206$ . This is the value derived for a truncated isothermal sphere of dark matter [19] for an Einstein–de Sitter universe, and also applies for all but very small redshifts for a flat universe with a finite cosmological constant [21]. In Fig. 2 we show the range of halo masses which experience core collapse (upper frame) as a function of formation redshift  $z_0$ . Since the lower limit depends also on the desired redshift at collapse,  $z_1$ , we show three variations:  $z_1 = 6, 3$ , and  $0$ . The lower frame shows the corresponding core ( $\approx$  black hole) masses at collapse, using Eqs. (1) and (4). The cosmological model is  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$ , and  $h = 0.65$ , and we set  $\sigma = 5 \text{ cm}^2 \text{ g}^{-1}$  [15].

We find that SMBHs of  $(10^6 - 10^7)M_\odot$  can form in massive halos which experience core collapse at intermediate and low redshifts. At redshift  $z_1 = 6$ , our model limits the initial black hole mass to  $M \approx 5 \times 10^5 M_\odot$ . Subsequent accretion onto the black hole must be invoked to reach masses of  $10^9 M_\odot$  [22], especially to reconcile with recent observations of high redshift quasars [23]. Very young halos are practically excluded as candidates for reaching core collapse by redshift  $z = 0$ : Halos formed later than  $z \approx 1.5$  will not have reached core collapse, regardless of mass, and only cores of extremely massive halos ( $M_{\text{tot}} \geq 10^{14} M_\odot$ ) could have collapsed if formed after  $z = 3$ .

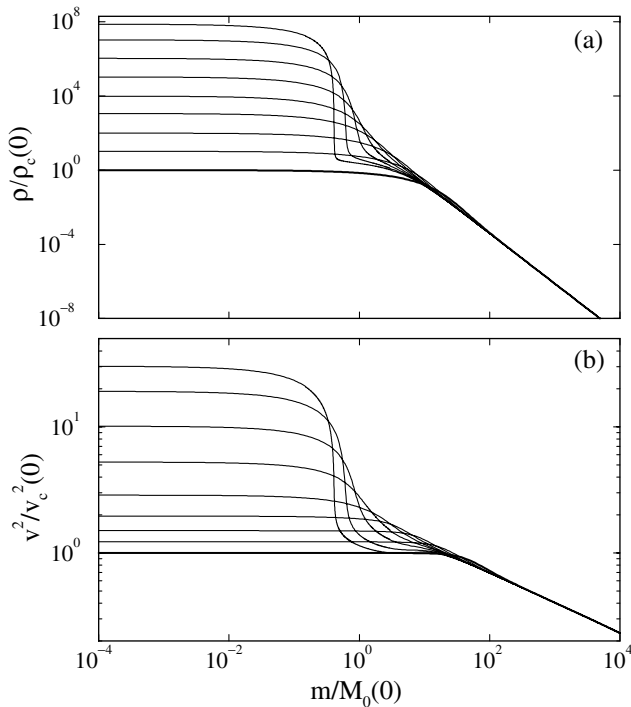


FIG. 1. Snapshots of the (a) density and (b) velocity dispersion profiles of a SIDM halo with  $\sigma = 0.674\pi R_0^2/M_0$  (for  $a = 2.26$ ) at selected times during gravothermal evolution [16]. The bottom line corresponds to the profile at  $t = 0$ . Subsequent profiles correspond to  $t/t_r(0) = 264.0, 287, 289.5, 289.90, 290.12, 290.30, 290.49$ , and  $290.63$ .

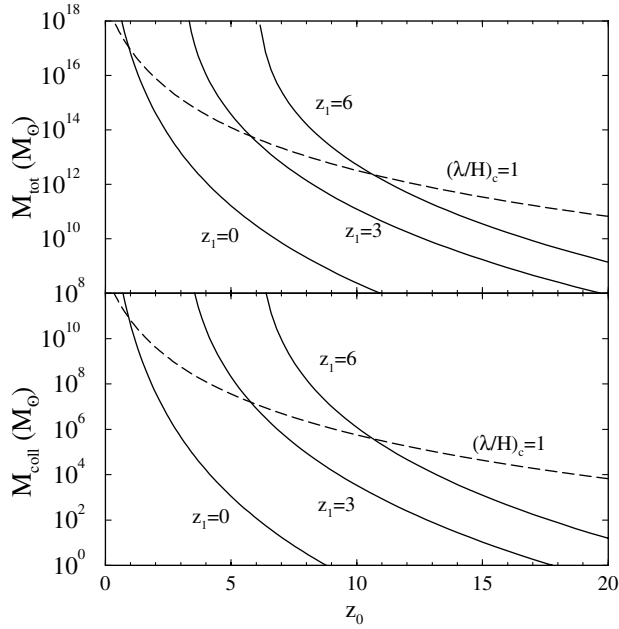


FIG. 2. Masses of SIDM halos that can undergo core collapse by redshifts  $z_1 = 6, 3,$  and  $0$  (upper frame), and the corresponding core mass at collapse (lower frame), as functions of formation redshift  $z_0$ . Dashed curve corresponds to the upper limit [Eq. (3)] and solid curves to the lower limits [Eq. (6)] as functions of  $z_0$  and  $z_1$ . In this model the density contrast  $x \equiv \rho_c(t=0)/\bar{\rho} = 1.8 \times 10^4$ .

*III. A bottom-up scenario for SMBH formation.*—Large values of  $M_{\text{coll}}$  require very massive halos in the direct scenario. Virialization of massive halos, especially at high redshift, is not favored in the currently accepted picture of structure formation [24,25]. The smaller mass halos formed preferentially in current theory can undergo core collapse in a Hubble time provided the fiducial value of  $x$  is larger [see Eqs. (3) and (6)]. In Fig. 3 we show similar results to those of Fig. 2, except that now  $x = 10^6$  (and  $g \approx 902$ ). This value appears to be in better agreement with the numerical simulations of Ref. [15], at least for low redshifts.

The relations  $M_{\text{min}} \propto x^{-7/2}$  and  $M_{\text{max}} \propto x^{-2}$  show why core collapse shifts to relatively low mass halos for  $x = 10^6$ . For this  $x$ , any halo which virializes at redshift  $z \geq 5$  will have undergone core collapse. However, young halos formed at low redshifts will still be safe from core collapse, as we found also for  $x = 1.8 \times 10^4$ . In this scenario, core collapse even by redshift  $z = 6$  is quite possible, for halos with masses in the range  $(10^7-10^{10})M_\odot$  virialized at redshifts  $z = 7-10$ . The masses of the collapsed cores are small, and, for halo masses considered realistic in the Press-Schechter formalism [24],  $M_{\text{coll}} \leq (10^2-10^3)M_\odot$ . In this case, gravothermal core collapse does not lead directly to the formation of SMBHs found in the centers of galaxies, but the resulting intermediate mass black holes can serve as seeds for SMBH buildup through mergers [3,22], or accretion.

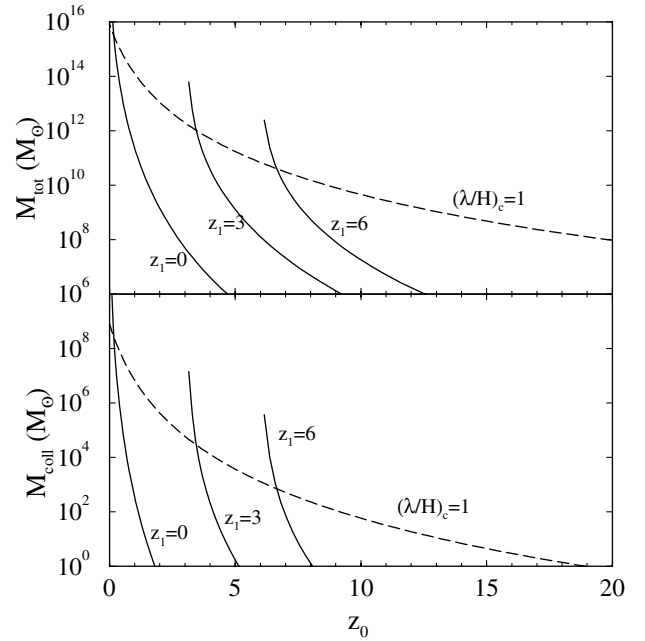


FIG. 3. Same as Fig. 2, except that  $x \equiv \rho_c(t=0)/\bar{\rho} = 10^6$ .

*IV. Discussion.*—Gravothermal core collapse in a SIDM halo triggers the formation of a SMBH at its center. The black hole forms when the inner core becomes relativistic and dynamically unstable. The mass of the collapsing core will be  $10^{-8}-10^{-6}$  times the total mass of the halo. Forming SMBHs by core collapse in SIDM halos requires no baryons, no prior epoch of star formation, and no other black hole seed mechanism.

For massive ( $>10^{13}M_\odot$ ) halos which virialize at modest overdensities, SMBHs with masses of  $(10^6-10^7)M_\odot$  form directly through gravothermal collapse. An alternative scenario where the smaller SIDM halos preferred in current cosmological models reach core collapse in a Hubble time is also possible and arises if the typical overdensity of the halos at virialization is large. In this case, any halo which forms prior to redshift  $z = 5$  produces a black hole by  $z = 0$ , and black hole formation at high redshift is also possible. However, these halos give rise only to low and intermediate mass black holes,  $\leq (10^2-10^3)M_\odot$ . These might be seeds of SMBHs through multiple halo mergers [22]. Even if only a few percent of the halos achieve core collapse by  $z = 5$ , this initial black hole population is sufficient to generate the observed SMBH spectrum eventually [26]. Coalescence of black holes in halo mergers is a likely source of gravitational waves, potentially detectable with the Laser Interferometer Gravitational Wave Observatory (LIGO) as well as the planned Laser Interferometer Space Antenna (LISA). If most of the SMBH population arises through multiple mergers of lower mass seed black holes, the rate may be as large as several per year [26].

The newly formed black hole may grow through accretion of SIDM from the halo. The main reservoir for accretion is the outer core, which is characterized by  $\lambda/H > 1$ ,

and includes a mass  $\geq 10^{-3}M_{\text{tot}}$ . If the ambient SIDM halo eventually dominates the spherical bulge of galaxies, then the accretion of the entire outer core would provide a natural means of producing the ratio  $M_{\text{BH}}/M_{\text{bulge}}$  observed in nearby galaxies. It would also explain the origin of the most massive ( $>10^9M_{\odot}$ ) central black holes. The possibility that SIDM accretion onto seed black holes is the origin of SMBHs was also proposed in [27], but there it is assumed that the seeds arise from supernova explosions in massive stars rather than the gravothermal scenario discussed here. Their quantitative results depend crucially on an assumed singular power-law density profile of the SIDM, which has yet to be verified self-consistently.

Young halos, especially low mass ones, could not have reached core collapse by  $z = 0$ . This allows SIDM halos to explain the flat density cores observed in some dwarf and low-surface-brightness galaxies, where the inferred values of central density and velocity dispersion are  $\rho_c \approx 0.02M_{\odot} \text{pc}^{-3}$  and  $v_c \lesssim 10^7 \text{cm s}^{-1}$ , yielding a gravothermal core lifetime that greatly exceeds the Hubble time.

To identify the main route for SMBH formation, it would be useful to determine the SMBH population as a function of redshift. More data about the presence or absence of central black holes in dwarf galaxies would also be useful.

We thank Y. Birnboim and P.R. Shapiro for useful discussions. This work was supported in part by NSF Grant No. PHY-0090310 and NASA Grants No. NAG5-8418 and No. NAG5-10781 at the University of Illinois at Urbana-Champaign.

---

\*Also at Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel.

†Also at Department of Astronomy and NCSA, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

- [1] J. Kormendy and K. Gebhardt, in Proceedings of The 20th Symposium on Relativistic Astrophysics, edited by H. Martel and J. C. Wheeler (astro-ph/0105230, to be published); D. Richstone *et al.*, *Nature (London)* **395**, A14 (1998); J. Magorrian *et al.*, *Astron. J.* **115**, 2285 (1998).
- [2] Ya. B. Zel'dovich, *Sov. Phys. Dokl.* **9**, 195 (1964); E. E. Salpeter, *Astrophys. J.* **140**, 796 (1964).
- [3] M. J. Rees, *Annu. Rev. Astron. Astrophys.* **22**, 471 (1984).
- [4] A. Laor, *Astrophys. J.* **505**, L83 (1998); A. Wandel, *Astrophys. J.* **519**, L39 (1999).
- [5] L. Ferrarese and D. Merrit, *Astrophys. J.* **539**, L9 (2000); K. Gebhardt *et al.*, *Astrophys. J.* **539**, L13 (2000).
- [6] A. Laor, *Astrophys. J.* **553**, 677 (2001).
- [7] See [3] and references therein; T. W. Baumgarte and S. L. Shapiro, *Astrophys. J.* **526**, 941 (1999); K. C. B. New and S. L. Shapiro, Proceedings of the Third LISA Symposium [Classical Quantum Gravity **18**, 3965 (2001)].
- [8] H. M. Lee, *Astrophys. J.* **319**, 801 (1987); G. D. Quinlan and S. L. Shapiro, *Astrophys. J.* **356**, 483 (1990).
- [9] D. Lynden-Bell and R. Wood, *Mon. Not. R. Astron. Soc.* **138**, 495 (1968).
- [10] A. P. Lightman and S. L. Shapiro, *Rev. Mod. Phys.* **50**, 437 (1978); L. Spitzer, *Dynamical Evolution of Globular Clusters* (Princeton University Press, Princeton, NJ, 1987).
- [11] Ya. B. Zel'dovich and M. A. Podurets, *Astron. Zh.* **42**, 963 (1965) [*Sov. Astron.* **9**, 742 (1966)].
- [12] S. L. Shapiro and S. A. Teukolsky, *Astrophys. J.* **292**, L41 (1985); *Astrophys. J.* **307**, 575 (1986); *Philos. Trans. R. Soc. London A* **340**, 365 (1992).
- [13] D. N. Spergel and P. J. Steinhardt, *Phys. Rev. Lett.* **84**, 3760 (2000). See also E. D. Carlson, M. E. Machacek, and L. J. Hall, *Astrophys. J.* **398**, 43 (1992); M. E. Machacek, *Astrophys. J.* **431**, 41 (1994).
- [14] N. Yoshida, V. Springel, S. D. M. White, and G. Tormen, *Astrophys. J.* **544**, L87 (2000).
- [15] R. Davè, P. J. Steinhardt, and B. J. Wandelt, *Astrophys. J.* **547**, 574 (2001).
- [16] S. Balberg, S. L. Shapiro, and S. Inagaki, astro-ph/0110561 [*Astrophys. J.* (to be published)].
- [17] J. F. Navarro, C. S. Frenk, and S. D. M. White, *Astrophys. J.* **490**, 493 (2001).
- [18] B. Moore *et al.*, *Astrophys. J.* **524**, L19 (1999).
- [19] P. R. Shapiro, I. T. Iliev, and A. C. Raga, *Mon. Not. R. Astron. Soc.* **307**, 203 (1999).
- [20] B. Moore *et al.*, *Astrophys. J.* **535**, L21 (2000).
- [21] I. T. Iliev and P. R. Shapiro, *Mon. Not. R. Astron. Soc.* **365**, 468 (2001).
- [22] Z. Haiman and A. Loeb, *Astrophys. J.* **552**, 459 (2001).
- [23] X. Fan *et al.*, *Astron. J.* **122**, 2833 (2001); R. H. Becker *et al.*, *ibid.* **122**, 2850 (2001).
- [24] W. H. Press and P. L. Schechter, *Astrophys. J.* **190**, 253 (1974).
- [25] C. Lacey and S. Cole, *Mon. Not. R. Astron. Soc.* **262**, 627 (1993); *Mon. Not. R. Astron. Soc.* **271**, 676 (1994).
- [26] K. Menou, Z. Haiman, and V. K. Narayanan, *Astrophys. J.* **558**, 535 (2001).
- [27] J. P. Ostriker, *Phys. Rev. Lett.* **84**, 5258 (2000); J. F. Hennawi and J. P. Ostriker, astro-ph/0108203, 2001 (to be published).