ICTP 2018:
Work with datasets

1. The table below gives a binned version of the most current supernova data (from Betoule et al. 2014).

z	DM	z	DM	z	DM
0.010	32.9538	0.051	36.6511	0.257	40.5649
0.012	33.8790	0.060	37.1580	0.302	40.9052
0.014	33.8421	0.070	37.4301	0.355	41.4214
0.016	34.1185	0.082	37.9566	0.418	41.7909
0.019	34.5934	0.097	38.2532	0.491	42.2314
0.023	34.9390	0.114	38.6128	0.578	42.6170
0.026	35.2520	0.134	39.0678	0.679	43.0527
0.031	35.7485	0.158	39.3414	0.799	43.5041
0.037	36.0697	0.186	39.7921	0.940	43.9725
0.043	36.4345	0.218	40.1565	1.105	44.5140
				1.300	44.8218

a) The quantity DM is proportional to the log of the apparent brightness. Make a plot of DM versus z; it should look like Figure E1 in Betoule et al. (2014).

b) The trend of DM with z should be matched by

$$M + 2.5\ \log_{10}\ d^2_{\rm Lum}(z|\Omega_m, \Omega_\Lambda)$$

where M, which represents the unknown luminosity of the standard candle, is a free parameter that is assumed to be independent of redshift. In the same plot as the data, plot curves which show $(\Omega_m, \Omega_\Lambda) = (1, 0)$, $(0.3, 0)$ and $(0.3, 0.7)$, and report your value of M.

c) In principle, $d_{\rm Lum}$ also depends on H_0. The expression above suggests that, if M were known a priori (i.e., if it were not a free parameter), then this data would constrain the value of H_0. However, even with perfect distance measurements there is a perfect degeneracy between the curvature Ω_k and the Hubble constant. Show this. (See Weinberg 1973.)

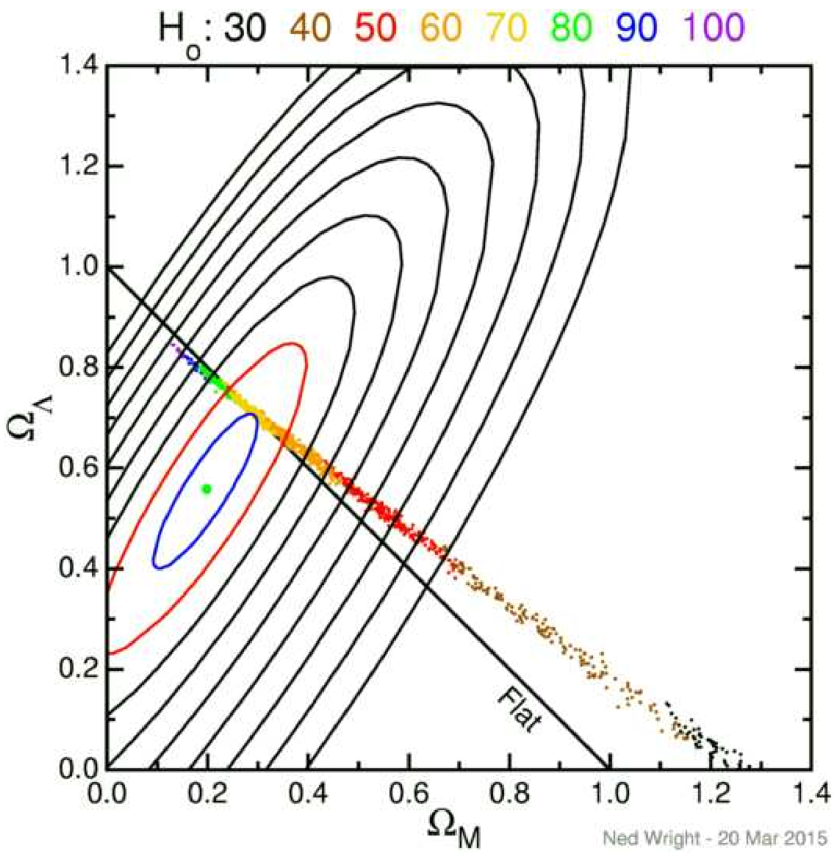

Figure 1: Cosmic complementarity: Combining datasets with orthogonal degeneracies is extremely useful.

2. The ellipses in the figure above show the constraints on the background geometry from SNae. The small colored dots show similar constraints from the CMB, for different values of H_0.

   (a) Each dataset has a rather large degeneracy in the $\Omega_{\rm M} - \Omega_\Lambda$ plane. Explain why.

   (b) For the CMB, small values of H_0 associated with large values of $\Omega_{\rm M}$. Why?

   (c) Do the SNae constraints depend on the value of H_0? Why or why not?

3. Page 24 of `growthCosmology.pdf` shows an SZ (thermal) and X-ray image of a cluster. For Bremmstrahlung radiation, the X-ray luminosity is $L_X \propto n_e^2 T_e^{1/2} r_{\rm vir}^3 \propto (f_g\, M_{\rm vir}/r_{\rm vir}^3)^2 T^{1/2}\, r_{\rm vir}^3 \propto f_g^2\, M_{\rm vir}\, T^{1/2} \propto f_g^2\, M_{\rm vir}^{4/3} \propto f_g^2\, T^2$. We won't derive this here, but comparison of the SZ and Xray observations allow a consistency check of the virial scalings, and/or a measure of the gas fraction in clusters, and/or another determination of the Hubble constant. See if you can convince yourself that $H_0 \propto \ell_x\, (T_e/T)^2/(\Delta T/T)^2_{\rm SZ}$ where ℓ_x is the observed X-ray flux.

4. I have provided the world's most recent measurement of $\xi(r)$ in the file `dr12_xi0.dat` (ignore the first and fourth columns). The file `camb_00394113_matterpower_z0.dat` is the CAMB power spectrum $P(k)$ for the current best ΛCDM cosmological parameters (from the CMB). Use it to compute $\xi(r)$, and compare with the measurement; when doing the comparison, allow the amplitude of $P(k)$ to be free. Report the value of the amplitude needed to make things match.

5. In class, we discussed why cluster abundances constrain cosmological parameters. This problem will give you a feel for what is involved. Figure 16 in the article `ChandraII.pdf` shows the current state of the art using X-ray clusters. (The build-up to Figure 16 – which you need not read carefully to do this problem – shows that there is a lot of groundwork required to even get to this point!) The $z = 0$ data from which the red points in this plot were constructed is `apj296878t2_ascii.txt`; the plot used M_Y, but M_G and M_T are two other viable estimates of the mass.

   (a) First, remake their measurement of the comoving density of halos using their equation (23):

   $$N(\geq M) = \sum_{M_i \geq M} 1/V(M_i).$$

   The factor of $V(M)$ in this expression is trying to account for the fact that it is difficult to see small mass halos which are far away, so we see low mass halos over a smaller volume than we see high-mass halos. If all halos were equally easy (or should I say difficult!) to observe, then $V(M)$ would be the same for all clusters. (Before moving on, you should convince yourself that in this case the constant value should be equal to the survey volume.) Unfortunately, their table does not give $V(M)$! The actual procedure for calculating $V(M)$ is rather involved, but the net result is the black curve in their Figure 14. This is reasonably well approximated by

$$(Vh^3/10^6\text{Mpc}^3) = (Mh/10^{14}M_\odot)^{2.5} \quad \text{if} \quad V \leq 9\times 10^8 h^{-3}\text{Mpc}^3$$

and by $V = 9\times 10^8 h^{-3}\text{Mpc}^3$ at larger V. (Feel free to build a better approximation for our 'precision' measurement!) Use this approximation with their list of halo masses to make your version of their Figure 16.

(b) The rest of their paper is about fitting a model to the measurements. We'll do a cruder test in a way which I hope highlights how one should exploit universality whenever one has reason to believe it is present.

Section 3 of `universalNm.pdf` gives you a fitting formula which the rest of the paper shows describes halo abundances in simulations rather well. It is in the 'universal' variable $\nu = \delta_c^2/\sigma^2(M)$ we discussed in class. Whereas `ChandraII.pdf` fit a model to $N(\geq M)$, universality says we could first transform to ν and then fit. To do this transformation, use equation (5) of `universalNm.pdf` for δ_c and the $P(k)$ you used for Problem Set 3 to get $\sigma^2(M)$. While it is tempting to simply set

$$N(\geq \nu) = \sum_{\nu_i \geq \nu} 1/V_i$$

the correct thing to do is to set

$$f(\geq \nu) = \sum_{\nu_i \geq \nu} (M_i/\bar\rho)/V_i$$

where $\bar\rho = \Omega_m \bar\rho_c$ is the comoving background density. Why?

(c) Make a plot of $\log_{10} f(\geq \nu)$ versus $\log_{10}\nu$. We can (and so should!) do better by accounting for the fact that the sample spans a range of z, and the halos may actually evolve over this range. Argue that setting

$$\nu_i = \frac{D_0^2}{D^2(z_i)}\frac{\delta_c^2(z_i)}{\sigma^2(M_i)}$$

accounts for this evolution. (E.g., the same M is mapped to a larger ν if it is at higher z.)

(d) Now compare to theory. This simply means you integrate equation (7) of `universalNm.pdf` to get $f(\geq \nu)$, which you can compare with the measurements. But the result depends on the parameters (A, a, p). Since the observations are of M_{500} and not $M_{\rm vir}$ or M_{200}, you should use the $500\rho_c$ values in Table 3 of `universalNm.pdf`. Compare this curve with the measurements.

(e) In principle, the mapping from measured masses to ν depends on the amplitude of the power spectrum, in a way which is loosely referred to as σ_8. Show what happens to the measurements if you change σ_8 from its default value by $\pm 10\%$.

Notice that the value of σ_8 is degenerate with the growth factor ratio. This is the basis of the cosmological test using clusters. If you had a sample of clusters at higher z, but you transformed to ν assuming $z = 0$, then the resulting $\log_{10} f(\geq \nu)$-$\log_{10} \nu$ plot will just be shifted, by $2 \log_{10}(D_0/D_z)$, with respect to the $z = 0$ symbols. In this way, universality would have allowed you to read-off the change in growth factor 'by eye'. In practice, small departures from universality add a slight complication. Talk to me if you are interested in this problem.

Although this is a good way to constrain σ_8 and the growth factor, it assumes that the background cosmology is known – after all, the step of convering from observed angles and redshifts to comoving volumes was done for us. (Indeed, strictly speaking, we should have first ensured that the CAMB cosmological model was the same as the one used in `ChandraII.pdf`.) Clearly, varying the cosmological model will change the measured M and $V(M)$ as well as the mapping to ν. Since the theory curve will change much less – it would not change at all if it were truly universal – this is how cluster counts constrain both the geometry of the universe and the shape and amplitude of $P(k)$.

Working with ν instead of mass M is a good way to see the Ω_m-σ_8 degeneracy shown in Figure 3 of `ChandraIII.pdf`. If $\nu\, f(\nu)$ truly were universal, then in plots of $\log_{10} f(\geq \nu)$-$\log_{10} \nu$, such as those you just made, changing Ω_m shifts the data vertically – because of the $M/(\Omega_m \bar{\rho}_c)$ weight applied to each object – but does not change the theory curve (because it is universal). On the other hand, changing σ_8 shifts all objects horizontally (since $\log\nu \propto \sigma_8^{-2}$) . Hence, if $\log_{10} f(\geq \nu)$-$\log_{10}\nu$ were a power-law, then there would be an exact degeneracy between Ω_m and σ_8: shifting the data upwards by reducing Ω_m can be exactly compensated by shifting the data leftwards by increasing σ_8). Since $\log_{10} f(\geq \nu)$-$\log_{10}\nu$ is a little curved, the degeneracy is not total; this is what gives the banana shape in Figure 3 of `ChandraIII.pdf`.

6. The paper `Fisher1995.pdf` contains a rather different derivation of the 'Kaiser factor' than the Zeldovich-based one we went through in class. The virtue of Fisher's calculation is that it shows how to proceed in the case of biased tracers. Fisher himself only considered the simplest case, in which the tracer field is linearly proportional to the matter field: $\delta_b = b\delta_m$. In this case, the Gaussian density and velocity fields are obviously still Gaussian. But it is not obvious that the streaming model will carry through, with a Gaussian for the distribution of pairwise velocities, if the biased density field is a *nonlinear* function of the matter field. To explore this, suppose that

$$1 + \delta_b \propto \exp(b\delta_m);$$

this is sometimes referred to as the Lognormal model.

(a) Set the constant of proportionality by requiring $\langle 1 + \delta_b \rangle = 1$.

(b) The correlation function for these biased tracers is given by $1 + \xi_b(r) \equiv \langle (1 + \delta_{b1})(1 + \delta_{b2})|r\rangle$, where the average is over $p(\delta_1, \delta_2|r)$. Express $\xi_b(r)$ in terms of $\Xi(r) \equiv \langle \delta_1\delta_2|r\rangle$, the correlation function of the underlying Gaussian field.

(c) Now repeat Fisher's calculation, but replace his pair-weight factor with $(1 + \delta_{b1})(1 + \delta_{b2})$. Your final expression should be in a form which shows if/how the pairwise velocity distribution is modified from that for the dark matter.

(d) Extra credit: How generic is this result? I.e., Can you generalize Fisher's calculation to the case in which $1 + \delta_b$ is an arbitrary monotonic function of δ_m?

With this in mind, remember that, in the spherical evolution model, the nonlinear $1 + \delta_{nl}$ is a monotonic function of the linear δ_m. This shows that one might think of the nonlinear field as a biased version of the linear one! Hence, any nonlinear bias transformation of the nonlinear field boils down to a different nonlinear transformation of the underlying Gaussian field. So the analysis above – weighted integrals over a multivariate Gaussian pdf – is surprisingly general.

7. The Cosmic Energy Equation: I had promised to show you what energy conservation looks like in an expanding universe. So, here we go ...

   (a) Define comoving coordinates

   $$\mathbf{x} \equiv \mathbf{r}/a(t) \qquad \text{and} \qquad \mathbf{v}_{\rm pec} \equiv a\dot{\mathbf{x}} = \dot{\mathbf{r}} - H\,\mathbf{r}.$$

   The Lagrangian for a single particle of mass m at position $\mathbf{r}$ is

   $$\mathcal{L} = \frac{ma^2\,\dot{\mathbf{x}}\cdot\dot{\mathbf{x}}}{2} - m\phi = \frac{mv_{\rm pec}^2}{2} - m\phi$$

   where

   $$\phi = -G\sum_j \frac{m_j}{|\,\mathbf{r}-\mathbf{r}_j|} + \left(\frac{\ddot{a}}{a} - \frac{\Lambda}{3}\right)\frac{r^2}{2}.$$

   The first term on the right hand side is familiar from Newtonian gravity. To see that the other terms are sensible, we would like to be sure that, in the limit in which the matter distribution is homogeneous, ϕ should not exert a force on a test particle. This means that we should check that $\nabla^2\phi = 0$ (right?!). Show that

   $$\nabla^2_{\mathbf{r}}\phi = 4\pi G\sum_j m_j\delta_{\rm D}(\mathbf{r}-\mathbf{r}_j) + 3\,\frac{\ddot{a}}{a} - \Lambda.$$

   The sum in first term on the right hand side is what we mean by the background density $\bar{\rho}(a) = \bar{\rho}_0/a^3$ (recall that the delta function has units of volume; $\bar{\rho}_0$ is the comoving background density). So $\nabla^2\phi = 0$ implies

   $$\frac{\ddot{a}}{a} = \frac{\Lambda}{3} - \frac{4\pi G}{3}\bar{\rho}(a)$$

   which is Friedmann's equation (Λ accelerates and matter decelerates the expansion). I.e., ϕ is sensible. Moreover, Friedmann's equation means that

   $$\phi = -G\sum_j \frac{m_j}{|\,\mathbf{r}-\mathbf{r}_j|} - \frac{G}{2r}\,\bar{\rho}(a)\frac{4\pi r^3}{3}.$$

This shows that we can think of the second term as representing $GM(< r)/2r$ if the mass were smoothly distributed. Note that $GM(< r)/r = GM(< x)/ax$. I.e., the comoving quantity evolves as a^{-1}; this will matter below.

Having convinced ourselves we have a sensible Lagrangian, we are ready to define the comoving momentum (i.e. the canonical momentum which is conjugate to the comoving coordinate $\mathbf{x}$):

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}} = ma^2\,\dot{\mathbf{x}} = a\,m\mathbf{v}_{\rm pec},$$

where the partial derivative means the positions are held fixed. Notice that this differs from the proper momentum measured by a comoving observer because of the extra factor of a.

The associated Hamiltonian is

$$\mathcal{H} = \mathbf{p}\cdot\dot{\mathbf{x}} - L = ma^2\,\dot{x}^2 - \frac{mv_{\rm pec}^2}{2} + m\phi = \frac{p^2}{2m\,a^2} + m\phi.$$

Notice the factor of a^{-2} on the first term, but not the second, which is arises from the fact that $\mathbf{p} = am\,\mathbf{v}_{\rm pec}$.

The Hamiltonian for a collection of particles is the sum of the individual contributions. However, whereas the term involving the momenta is just a straightforward sum, we must be careful to subtract the contribution from the background from the ϕ piece (just as we were careful when defining ϕ in the first place). Namely, if

$$H \equiv K + W,$$

with K depending on momenta and W on coordinates, then

$$K = \sum_i \frac{m_i v_i^2}{2} = \sum_i \frac{p_i^2}{2m_i\,a^2}$$

is the peculiar kinetic energy and

$$W = -\frac{G}{2}\int d\mathbf{r}_1 \int d\mathbf{r}_2\,\frac{[\rho(\mathbf{r}_1) - \bar{\rho}][\rho(\mathbf{r}_2) - \bar{\rho}]}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

is the potential energy. (If you write $\rho(\mathbf{r})$ as a sum over delta functions for the discrete particles, as in the expression for $\nabla^2\phi$, then the factor of two for the ϕ-term is to account for our double counting the same pair twice (recall ϕ_i is a sum over j, and it is understood that we should avoid $i = j$). You should convince yourself that none of these ρs are comoving, and this really is the right way to 'subtract-off' the contribution from the smooth background.

(b) Energy conservation means that $dH/dt = \partial H/\partial t$. Show that

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \qquad \text{implies} \qquad \frac{d(K+W)}{dt} = -\frac{\dot{a}}{a}\,(2K+W).$$

(Hint: The partial derivatives are to be taken at fixed momenta and at fixed positions. So the real problem is to show that $\partial K/\partial t = -2K\,\dot{a}/a$ and $\partial W/\partial t = -W\,\dot{a}/a$.)

(c) The expression above is the Layzer-Irvine cosmic energy equation for the evolution of the kinetic and gravitational potential energy of non-relativistic matter (such as CDM) which interacts only by gravity.

Non-interacting particles have $W = 0$. What does this equation say about this limit?

(d) Virialized structures have $W = -2K$, so the total energy $K + W$ is obviously conserved if everything is virialized. Is this a useful limit at late times in a Λ dominated universe? (Hint: What does 'structure has frozen out' imply?)

The Cosmic Energy equation is more than a nice generalization of the usual Newtonian energy conservation. In practice, it is used to ensure that energy is conserved in a numerical simulation (round-off error, force-smoothing, etc, are not causing trouble). It also provides a route to thinking about how to treat problems in which Λ also clusters.

Finally, although we won't use this here, notice that if we define $\langle W\rangle$ and $\langle K\rangle$ as the potential and kinetic energies per unit mass, then $\langle K\rangle$ is the variance of the distribution we discussed in class which replaces the Maxwell-Boltzmann, and

$$\langle W\rangle = -\frac{G\rho(a)}{2}\int dr\, 4\pi\, r^2\, \frac{\xi(r,a)}{r}.$$

This connects nicely to the two-point correlation function statistic we spent much of the last month discussing. In linear theory, $\langle W \rangle \propto \Omega H^2$ whereas $\langle K \rangle \propto f^2(\Omega) H^2 \approx \Omega^{8/7} H^2$, so the ratio of the two terms estimates Ω independent of the Hubble constant H. This may become an attractive estimate as more Kinetic SZ measurements become available. We also have a handle on how nonlinear clustering will affect W ad K; in the context of the halo model, both terms can be decomposed into 1- and 2-halo contributions. E.g., $\langle K \rangle$ is explicitly the sum of virial and linear theory motions. And, because halos are virialized, the 1-halo contribution will have $\langle W_{1h} \rangle = -2 \langle K_{1h} \rangle$, so the net result is actually rather straightforward.