The SDSS-III Baryon Acoustic Oscillation Survey (BOSS)

What? BOSS has collected the spectra of 1.5 million galaxies and 300,000 quasars in order to map the structure of the Universe in 3D and to measure the clustering of galaxies in the Universe on very large scales.

Why? BOSS astronomers use these measurements of galaxy clustering to test predictions of the Big Bang Theory and to better understand the mysterious Dark Energy, which appears to be driving the accelerating expansion of the Universe.

How? How do the locations of the holes on the aluminum BOSS plate end up telling us about large-scale structure in the Universe? Let's find out!

The early universe was filled with a hot, dense plasma of electrons and baryons (protons and neutrons), which was almost perfectly smooth… *but not quite.*

Around each slightly over-dense point in the early Universe, the plasma oscillated in spherical waves, due to the counteracting forces of gravity and pressure. These oscillations are described as acoustic, since they are analogous to sound waves, which create pressure differences in air.

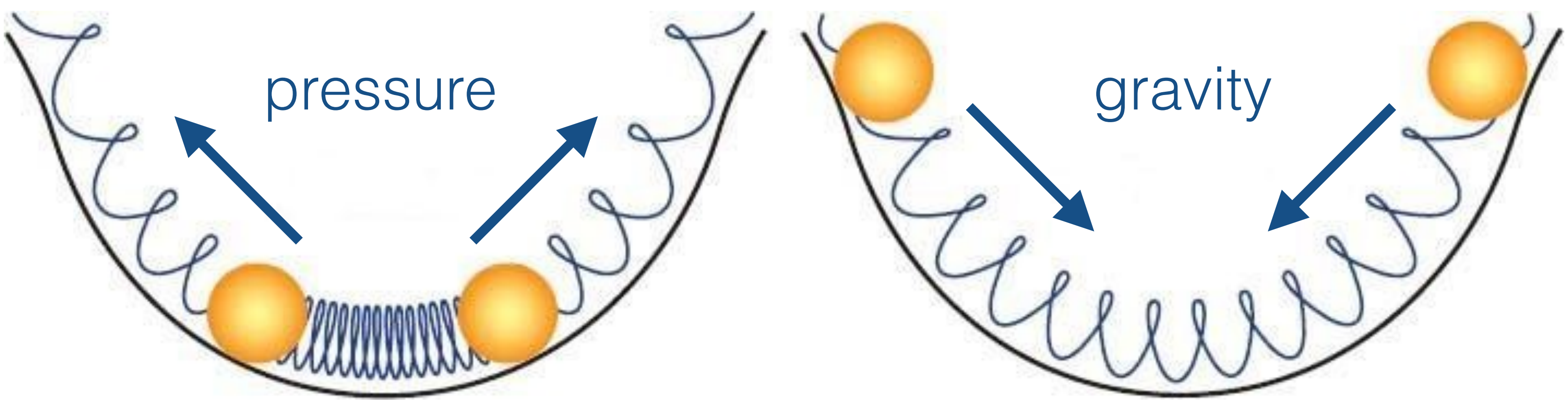
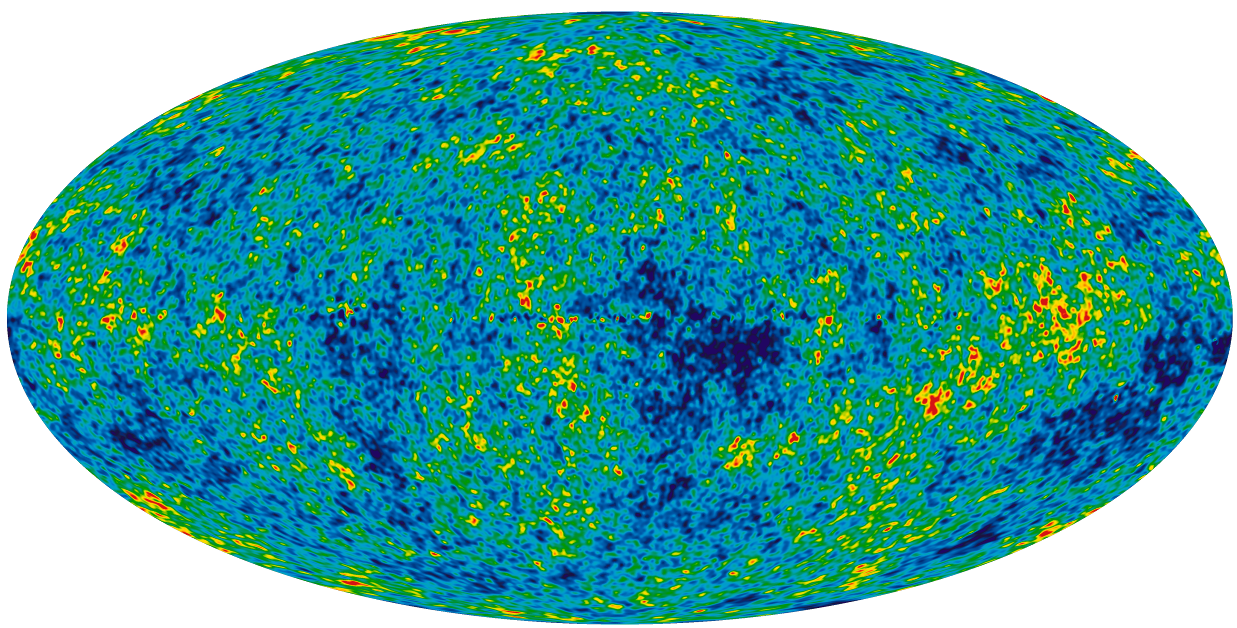


Image Credit: Wayne Hu, University of Chicago

At these early times photons were trapped, unable to travel very far without interacting with the plasma.

The universe expanded and cooled with time. About 379,000 years after the big bang, it had cooled enough to allow neutral Hydrogen to form out of the plasma. Photons that had previously been trapped by the plasma could then escape for the first time and free-stream through space. Those photons released after Hydrogen recombination produce our earliest image of the universe — a map of the cosmic microwave background (CMB).

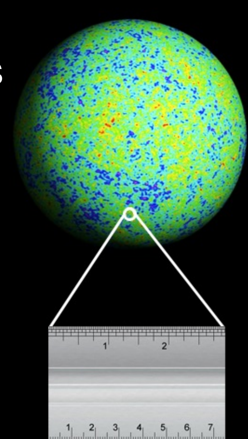


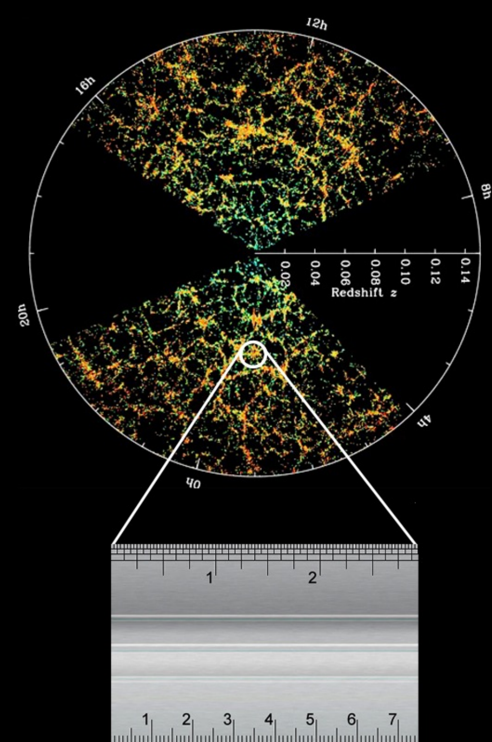
A map of the CMB, made with 9 years of data from the WMAP satellite (2012).

The initial density fluctuations, shown as color differences in the map of the CMB, acted like seeds from which the first galaxies formed and grew into the largest structures in the universe.

The scale of density fluctuations seen in the CMB provide a “standard ruler” for physical size, which we can compare to the scale of galaxy clustering in the modern universe to precisely measure how space has expanded over time.

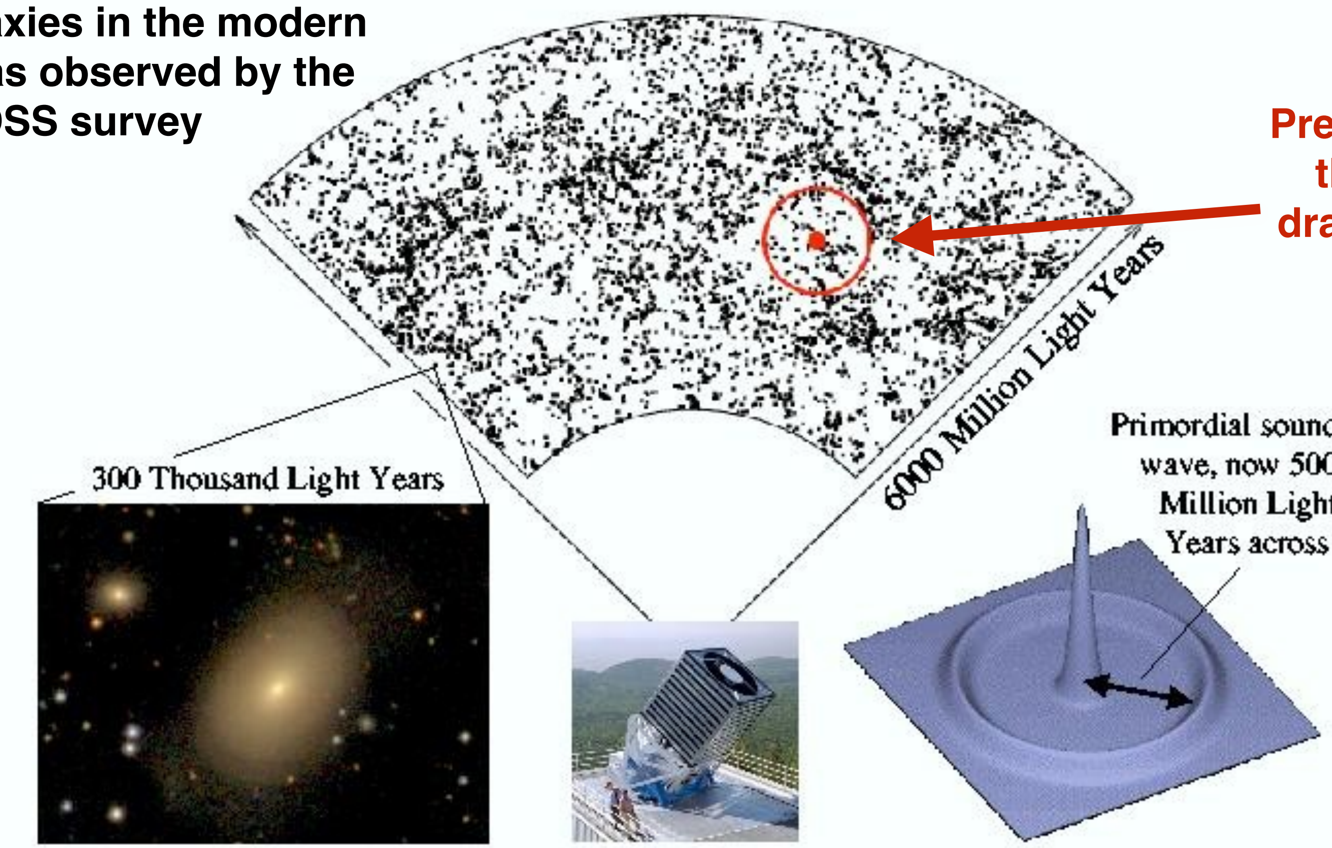
Map of density fluctuations in the early universe, as observed in the cosmic microwave background (CMB) radiation with the WMAP satellite.

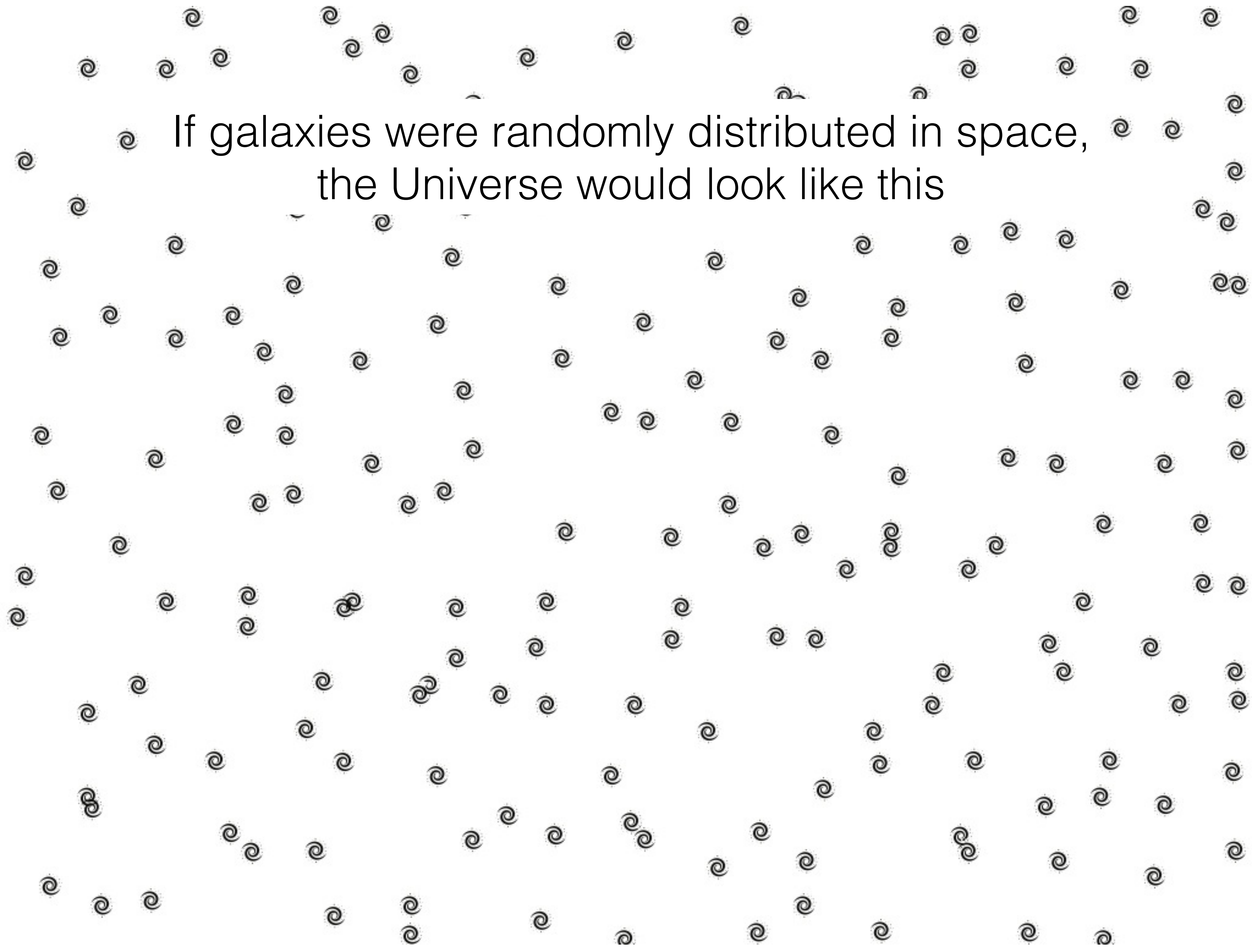




A partial map of galaxies in the modern universe, as observed by the BOSS survey

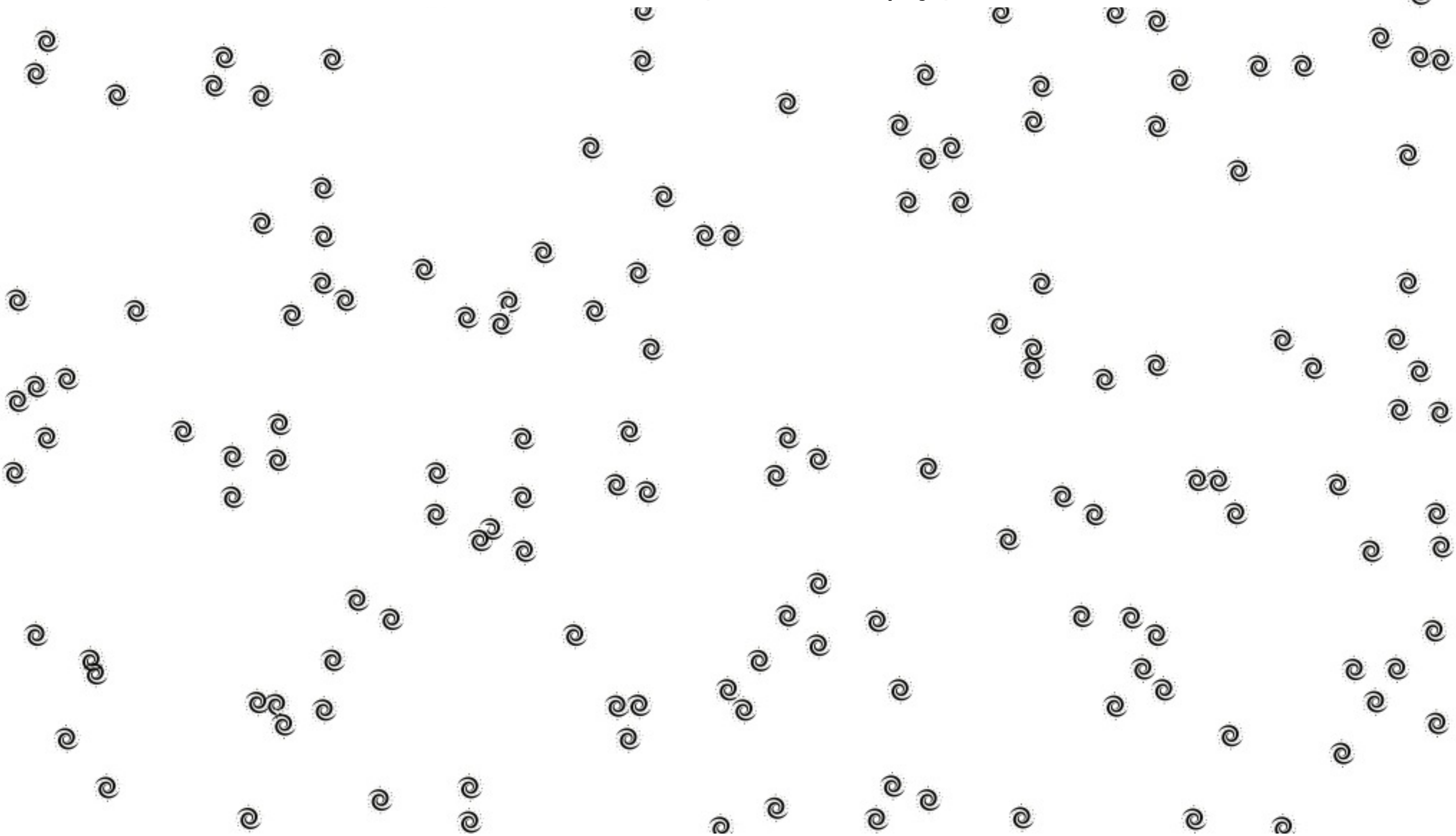
The BOSS Survey aimed to make a precise measurement of the clustering of galaxies over a wide range of scales in order to test whether galaxies are *excessively clustered* on a particular large scale called the *baryon acoustic peak*. This peak is predicted to exist at the physical distance reached by the primordial sound waves in the early universe when Hydrogen recombination occurred.



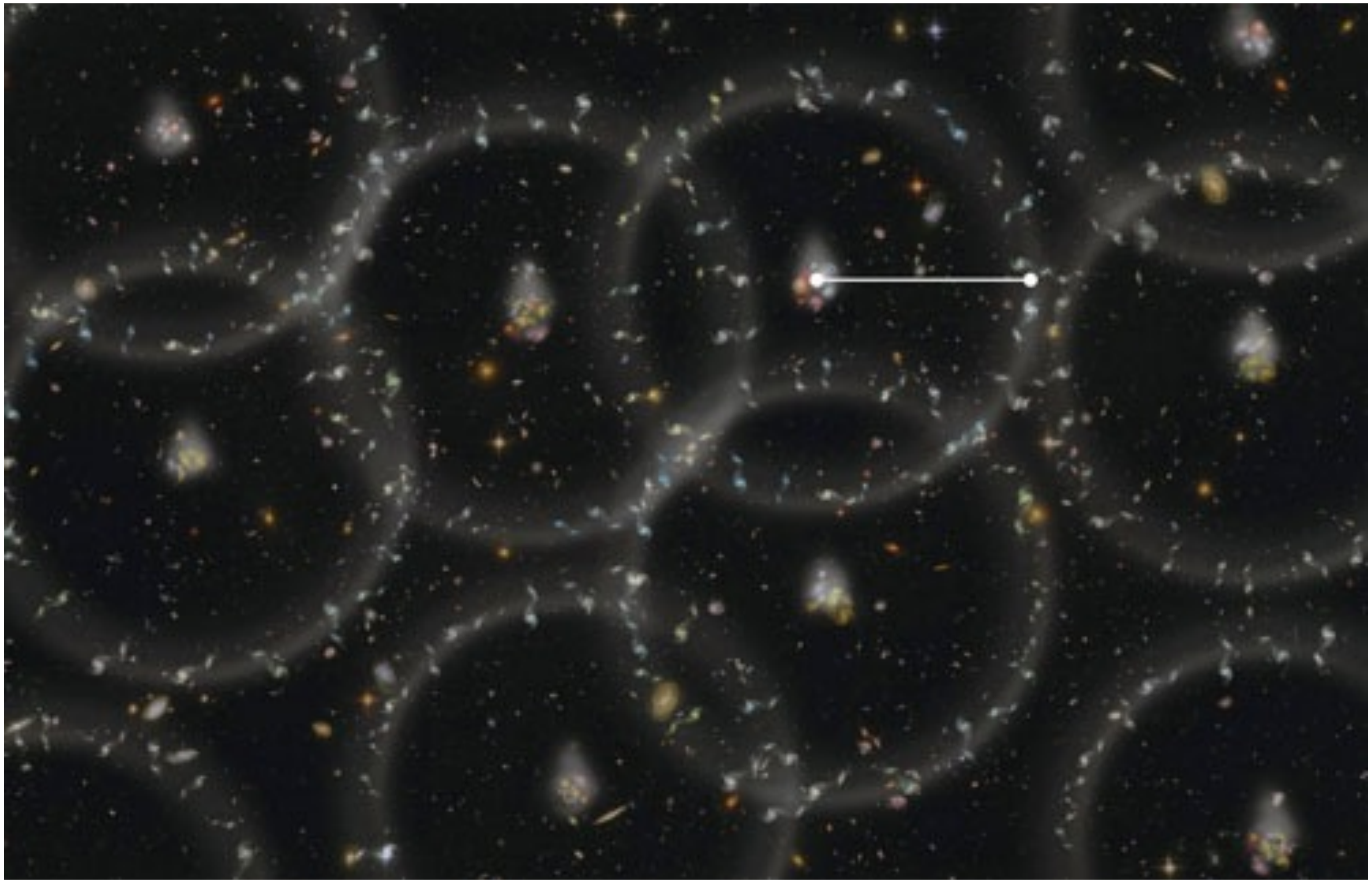


If galaxies were randomly distributed in space,
the Universe would look like this

But gravity causes real galaxies to clump together over time, so the real Universe looks more “clustered” (or clumpy) like this:



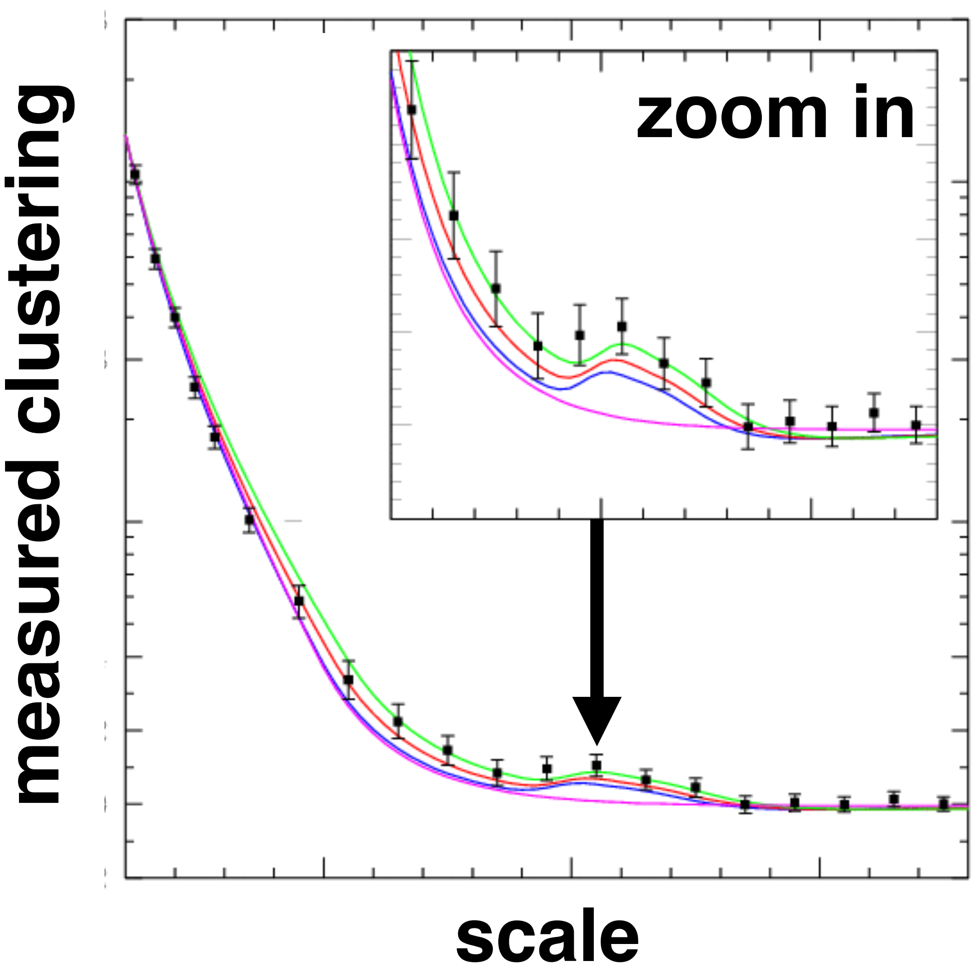
A universe in which galaxies are excessively clustered at a certain *scale* *(distance between galaxies)* might look something like this*:



*In this illustration, the excess clustering is greatly exaggerated to make it appear obvious around certain points in space.

At the particular scale of the baryon acoustic peak (BAO), the clustering of galaxies is indeed slightly excessive.

BUT, it takes hundreds of thousands of galaxies to see this feature and to measure it precisely.



Astronomers can use some surprisingly easy math to measure the clustering of galaxies.

One way to do this is to measure something called the "two-point correlation":

This measurement computes *on average* how many more galaxies you find at a certain distance around galaxies in the real Universe, compared to what you would expect if galaxies were distributed randomly in space.

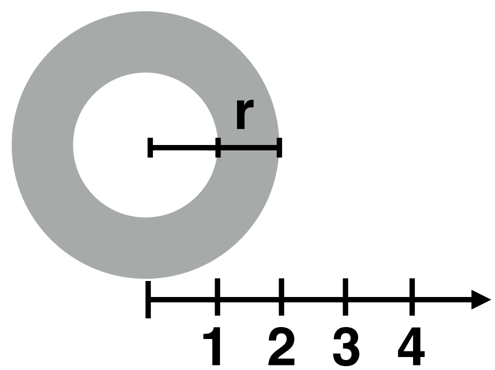
As a mathematical equation, the simplest "two-point correlation function" **w(r)** looks like this:

$$w(r) = \frac{N_{real}}{N_{random}} - 1$$

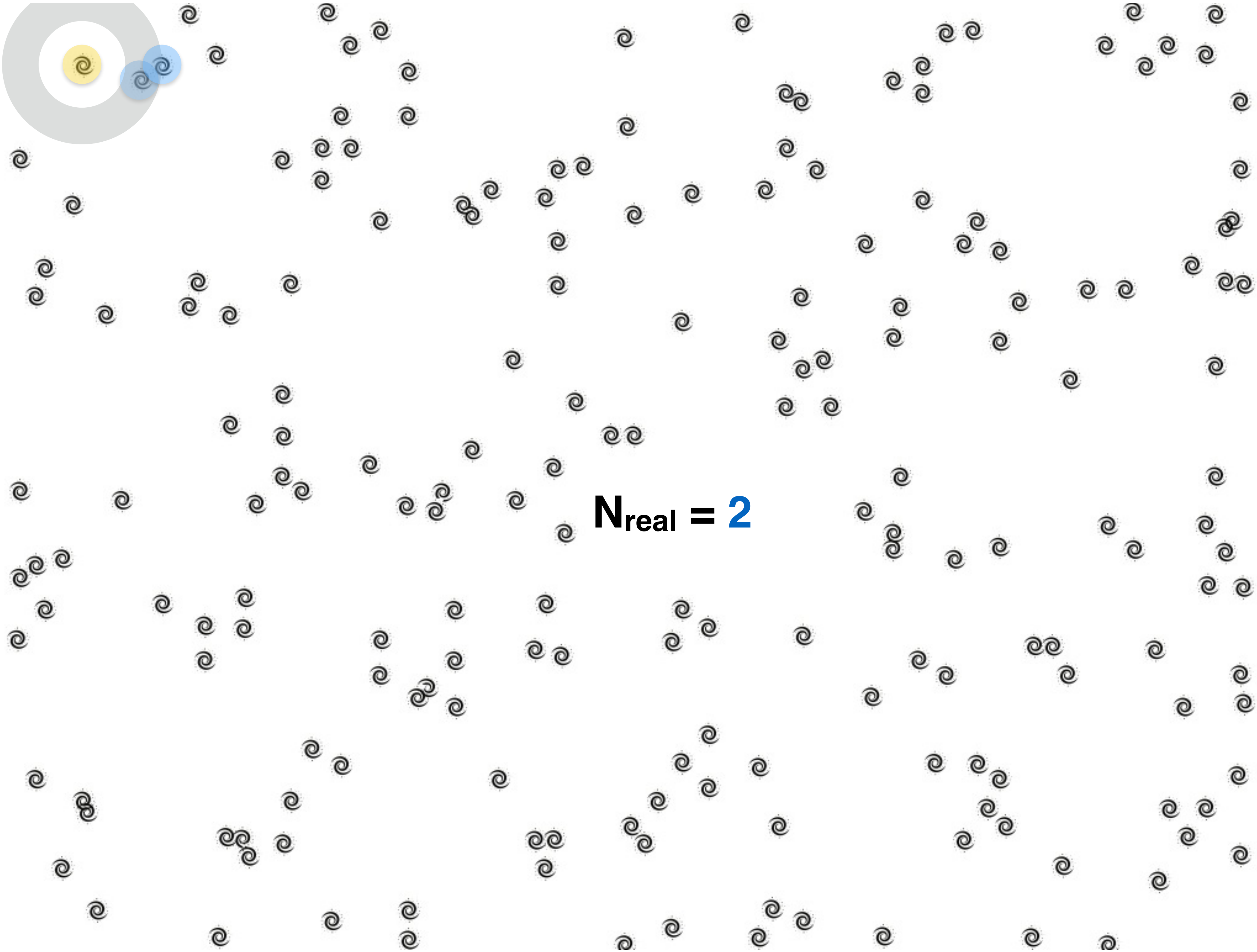
where $\boldsymbol{N_{real}}$ is a sum of the number of galaxies in the real Universe that are a certain distance **(r)** away from other galaxies, and $\boldsymbol{N_{random}}$ is the number of galaxies in a random distribution that are the same distance (r) away from other galaxies.

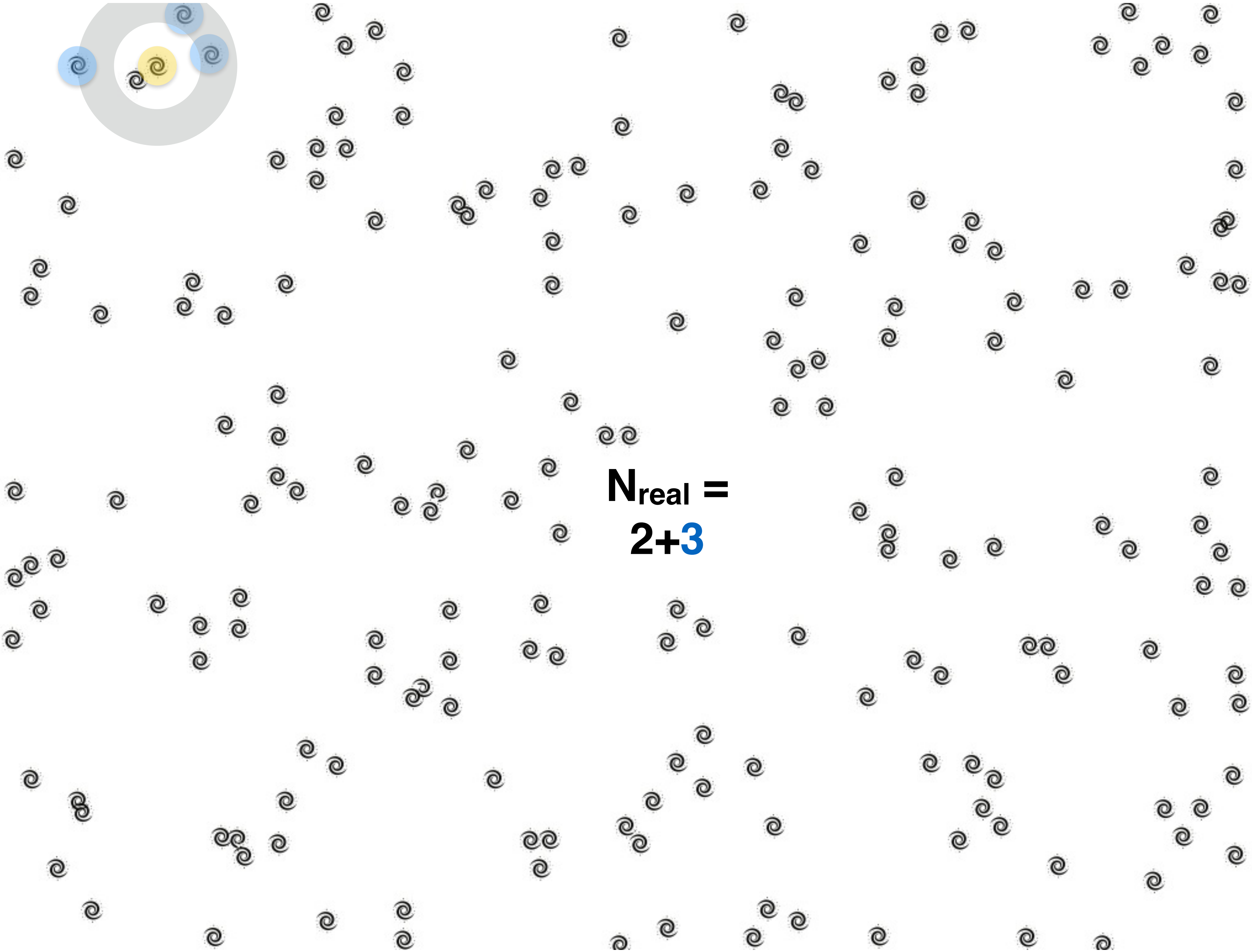
Let's try measuring it!

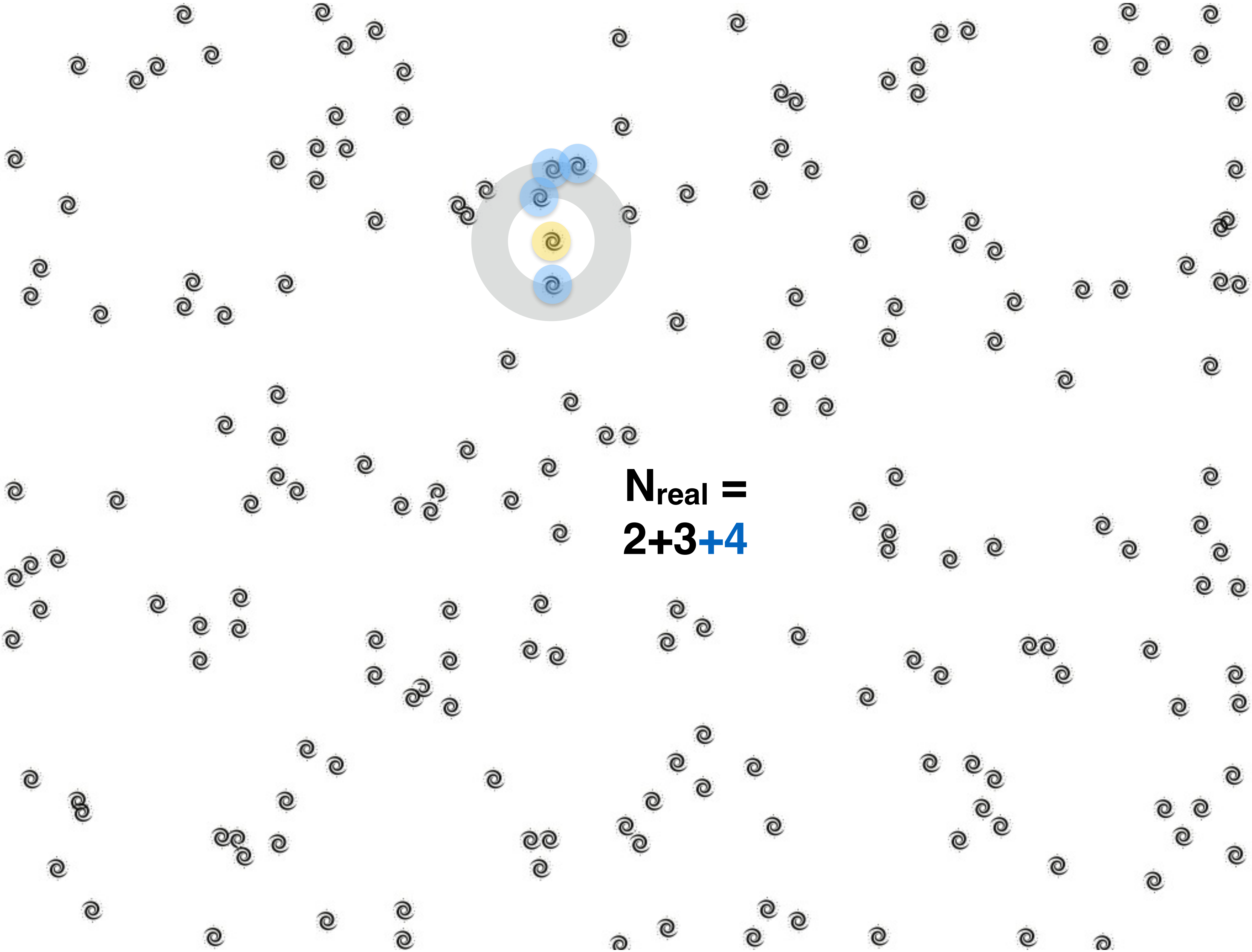
We’ll start by counting the galaxies in the “real Universe” that are a certain distance (let’s say, between 1 and 2 units in radius, r) away from other galaxies.



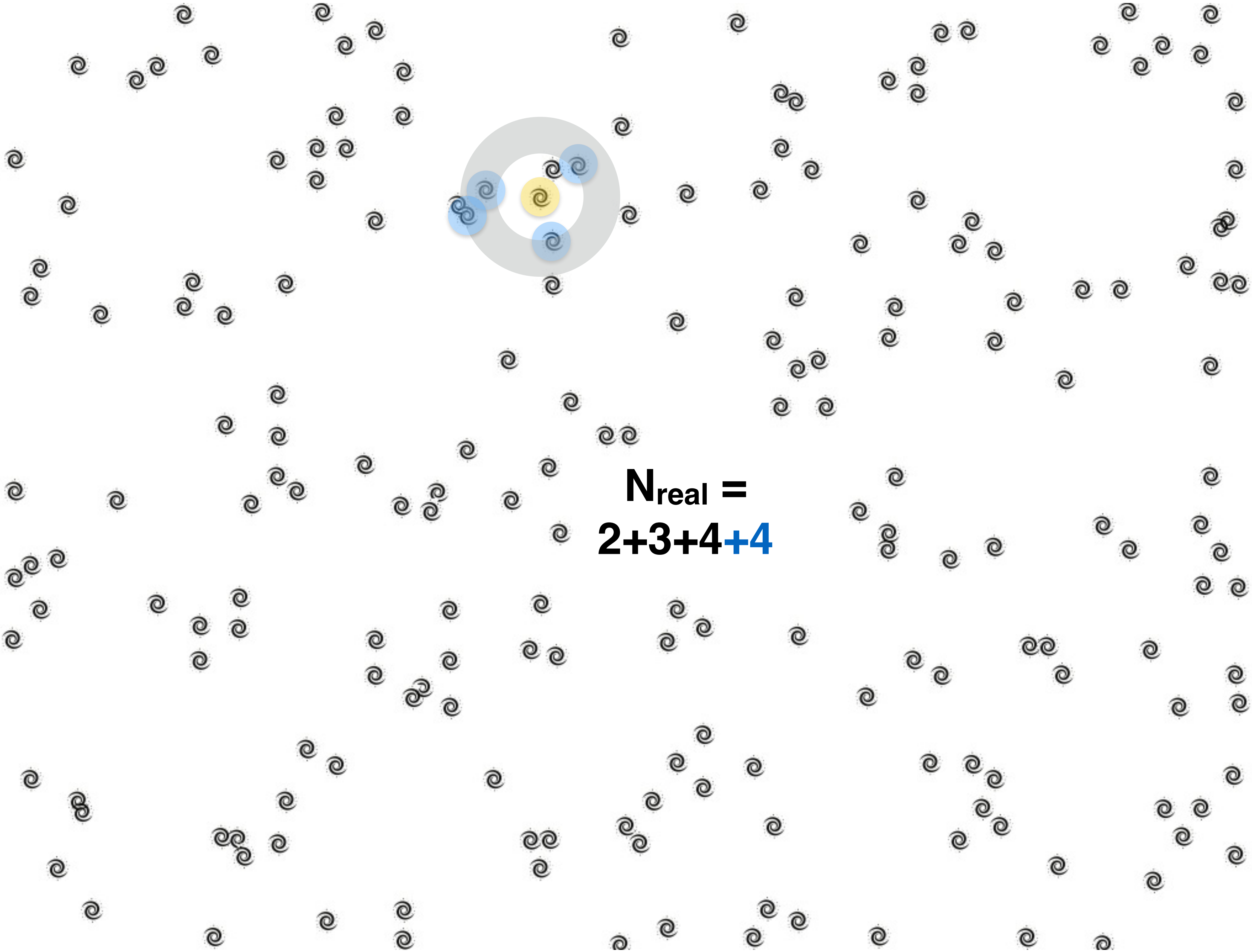
Place the center of this ring on as many galaxies as you can (trying not to repeat any), and count up all the other galaxies whose centers fall within the ring.


$N_{real} = 2$

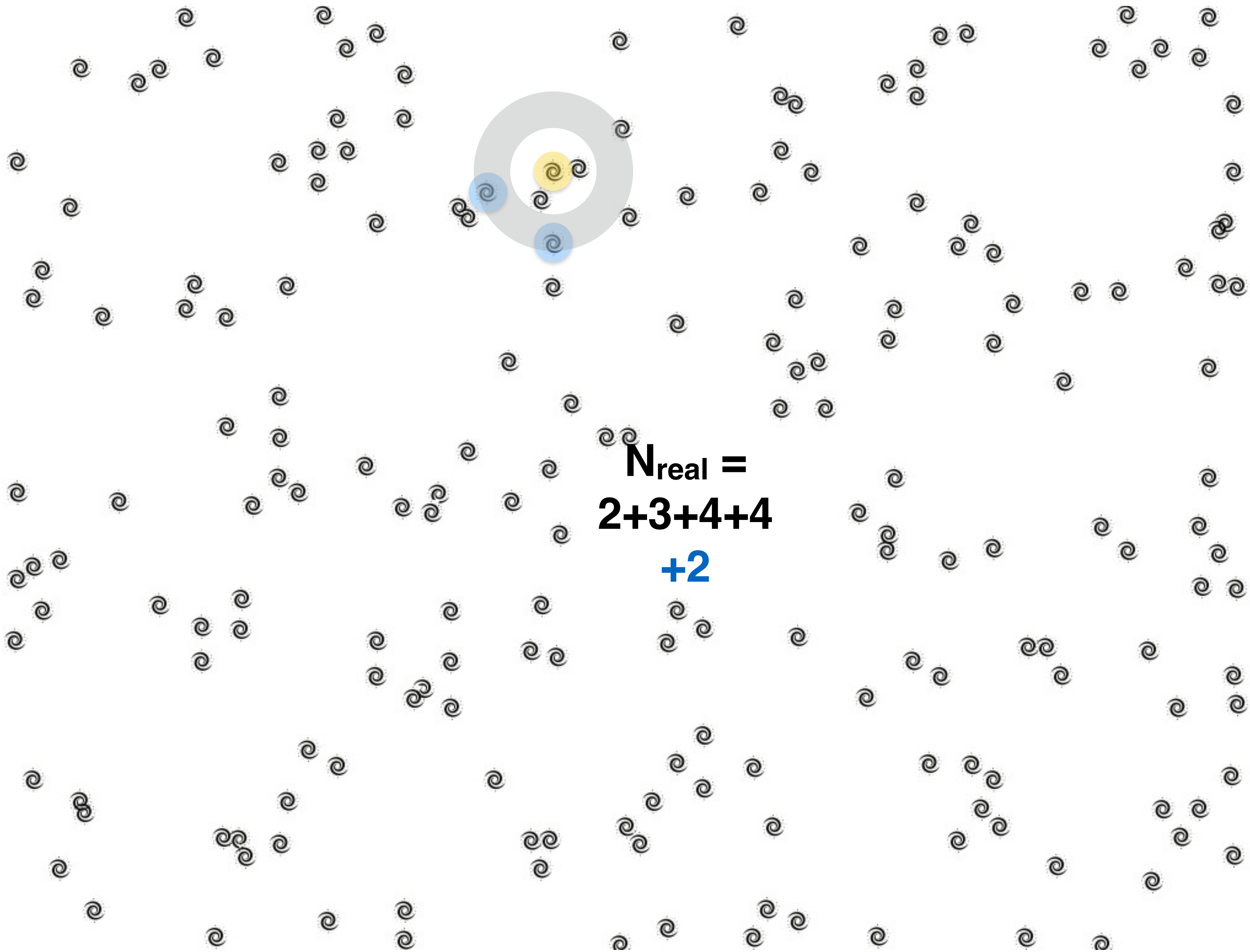

N_{real} =
2+3



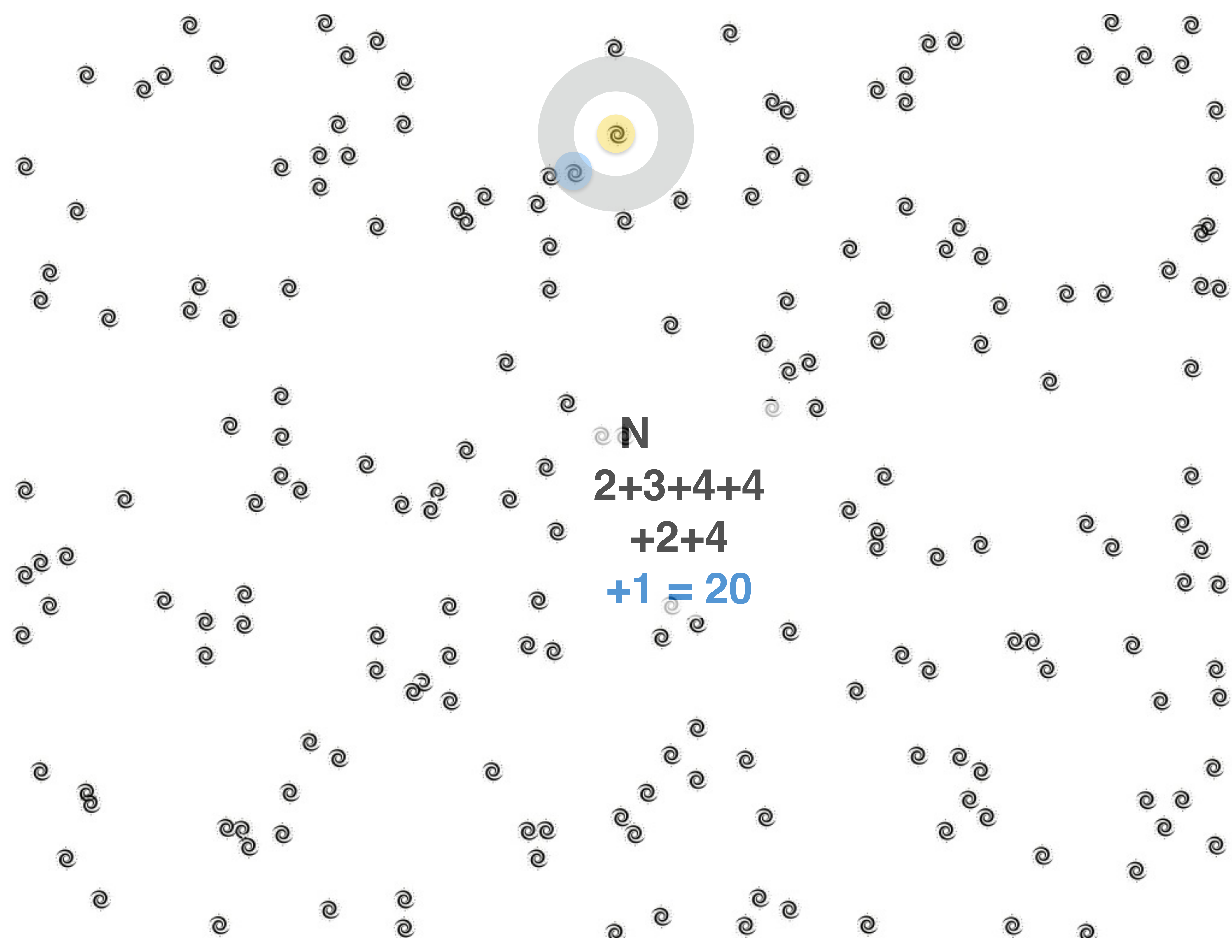
N_real =
2+3+4



N_{real} =
2+3+4+4


N_{real} =
2+3+4+4
+2

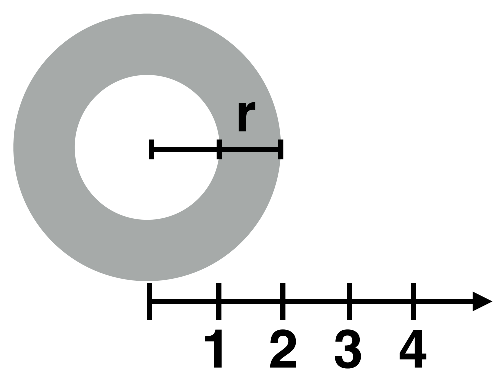

N_{real} =
2+3+4+4
+2+4

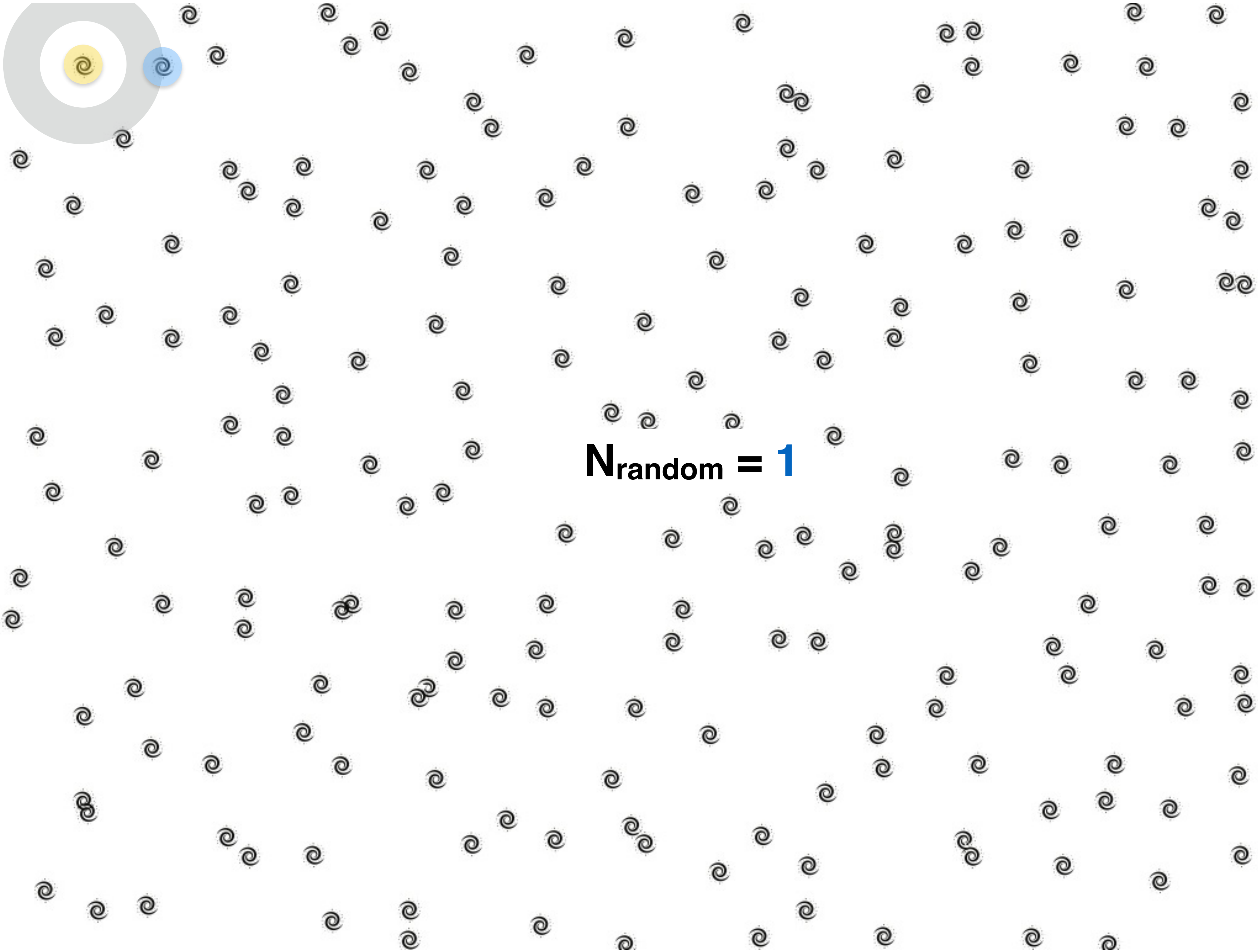

N
2+3+4+4
+2+4
+1 = 20

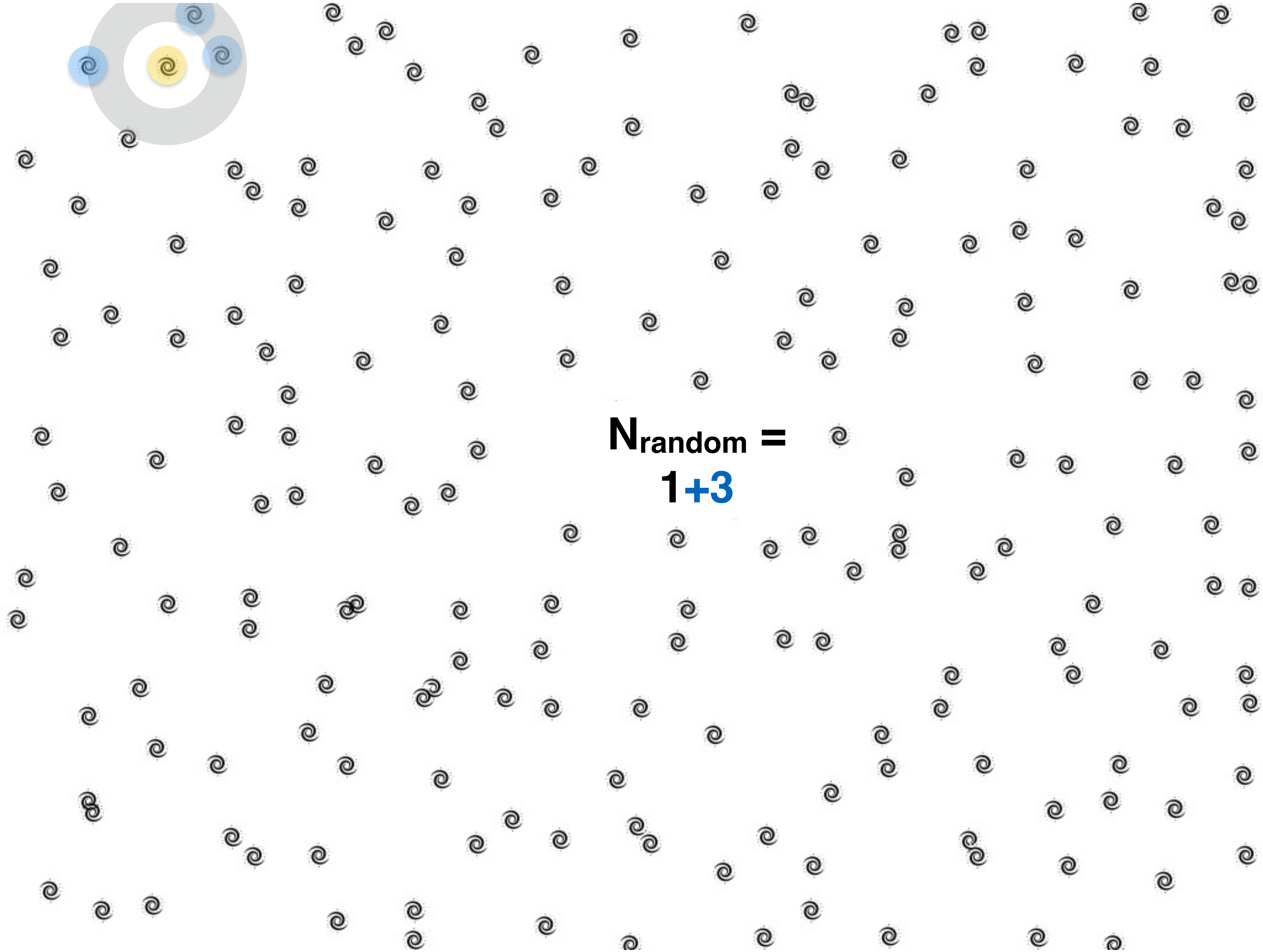
In our “real Universe” we made measurements around 8 galaxies within a distance of r=1, and counted a total of 20:

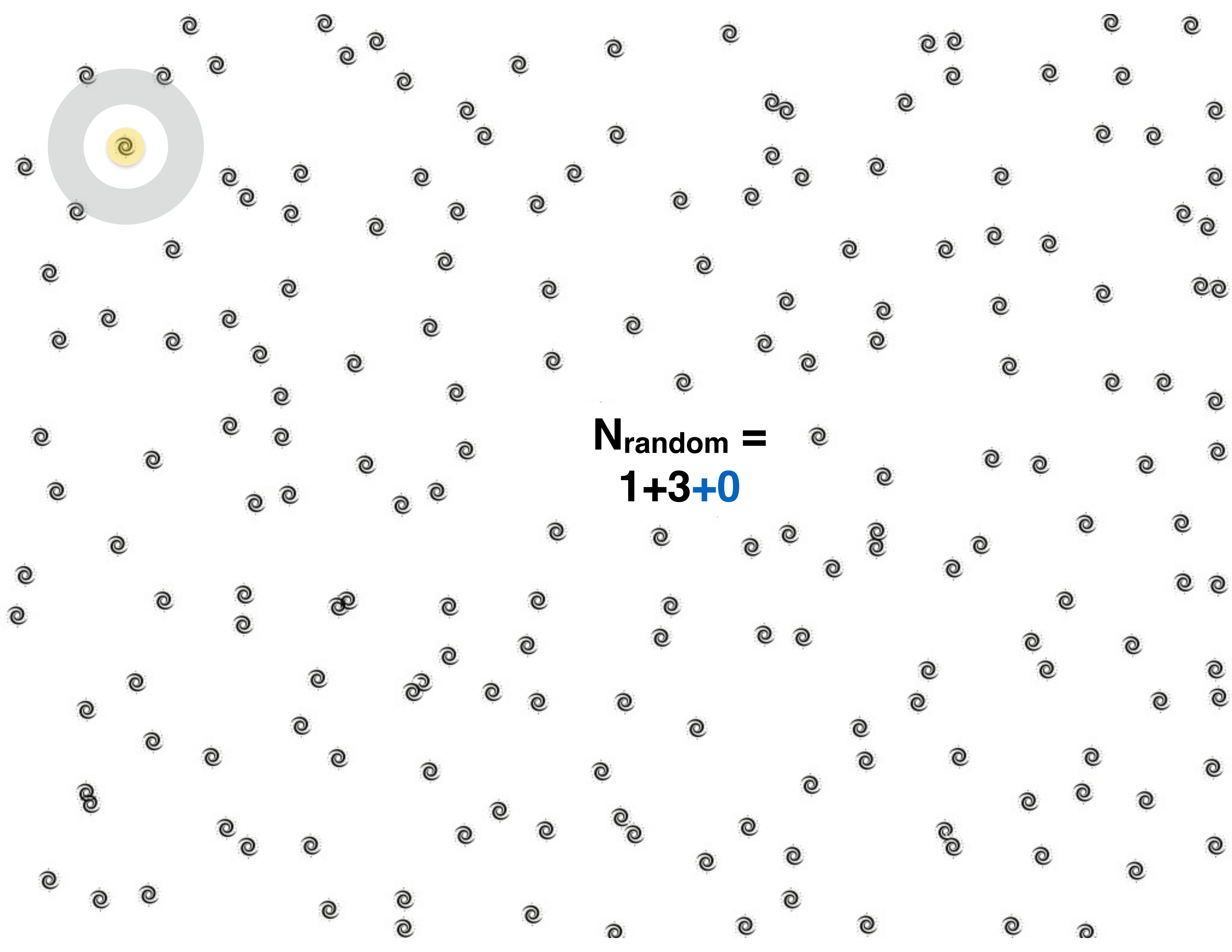
N_{real} = 20

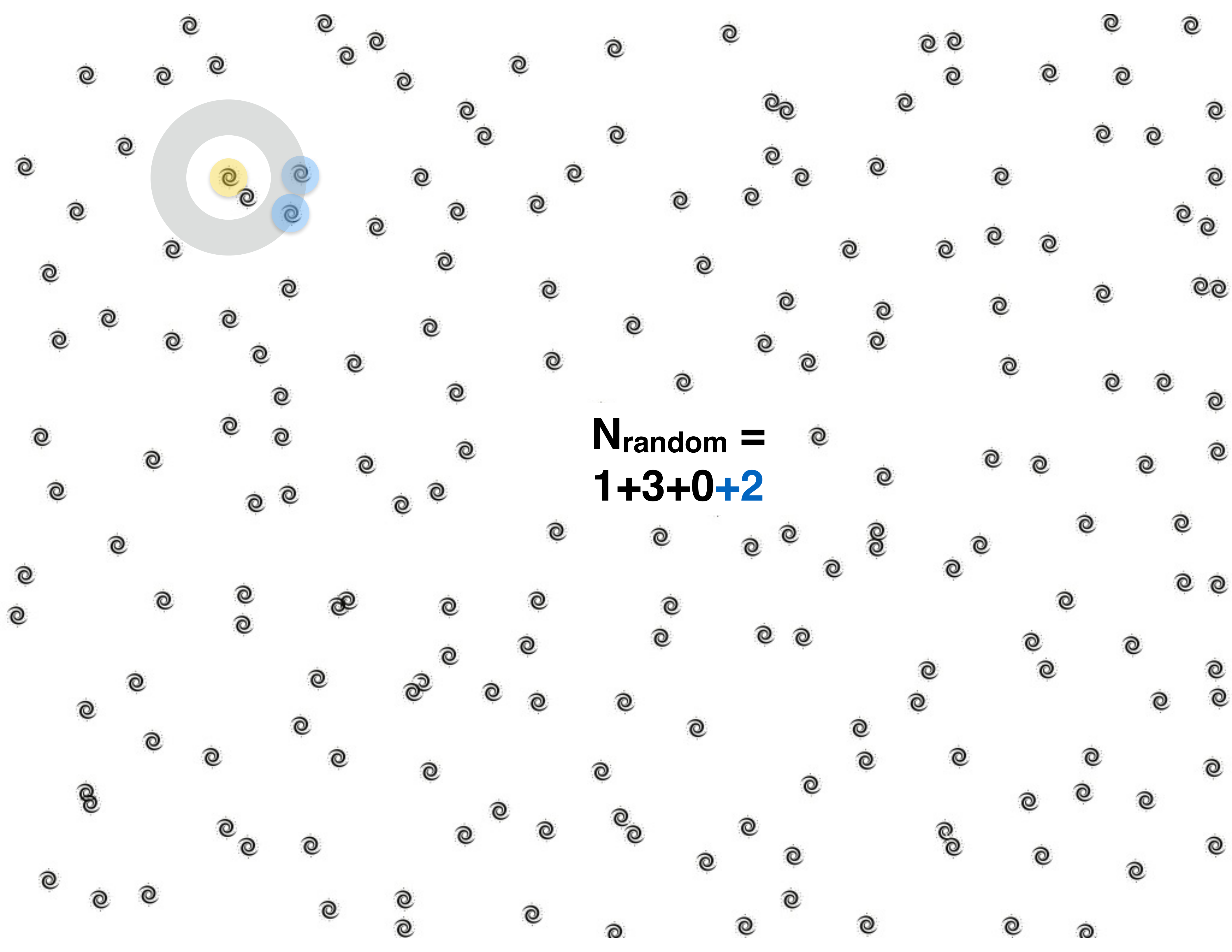
Let now do the same exercise for the “random distribution”. Be sure to do the same number of measurements (8), and to try not to center the measurement on the same galaxy more than once.

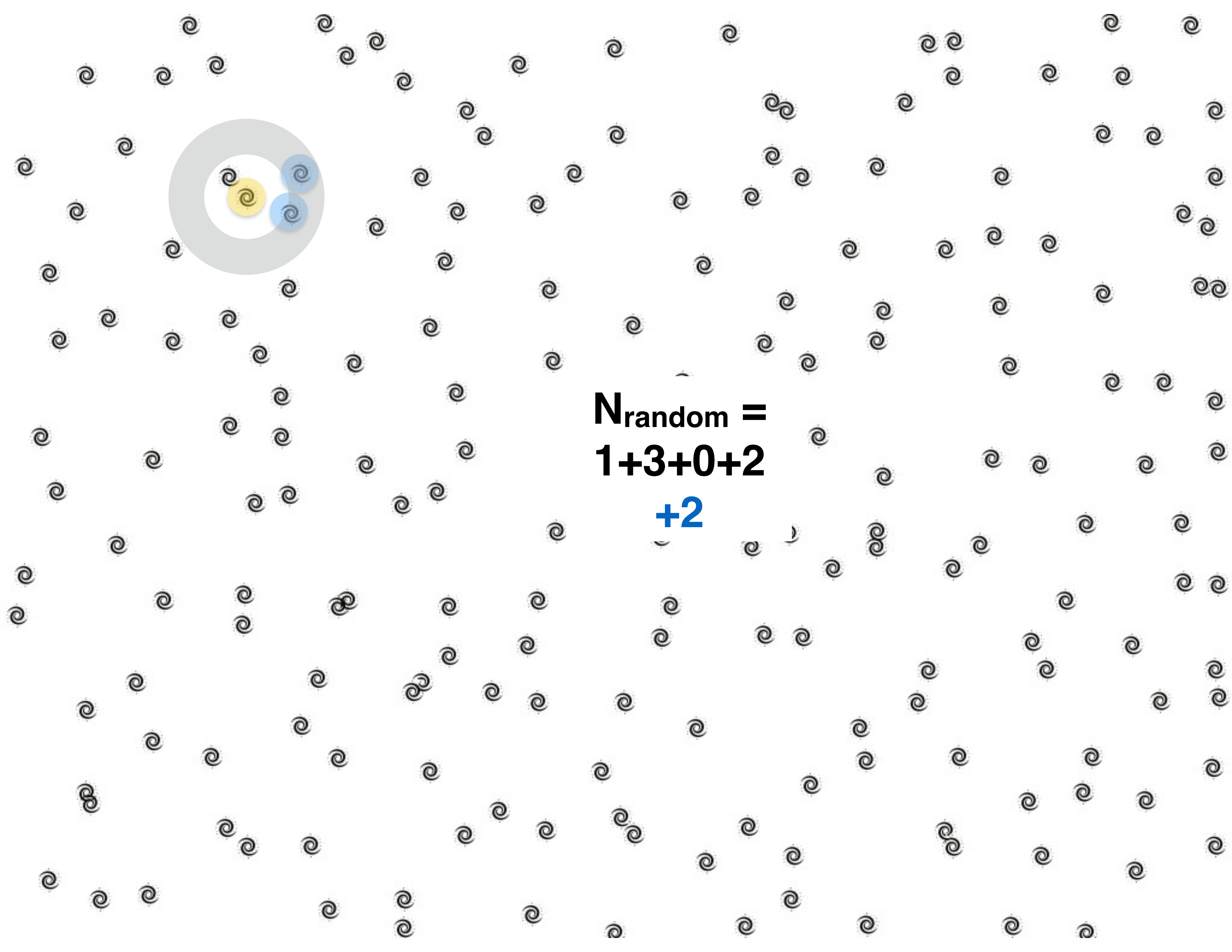


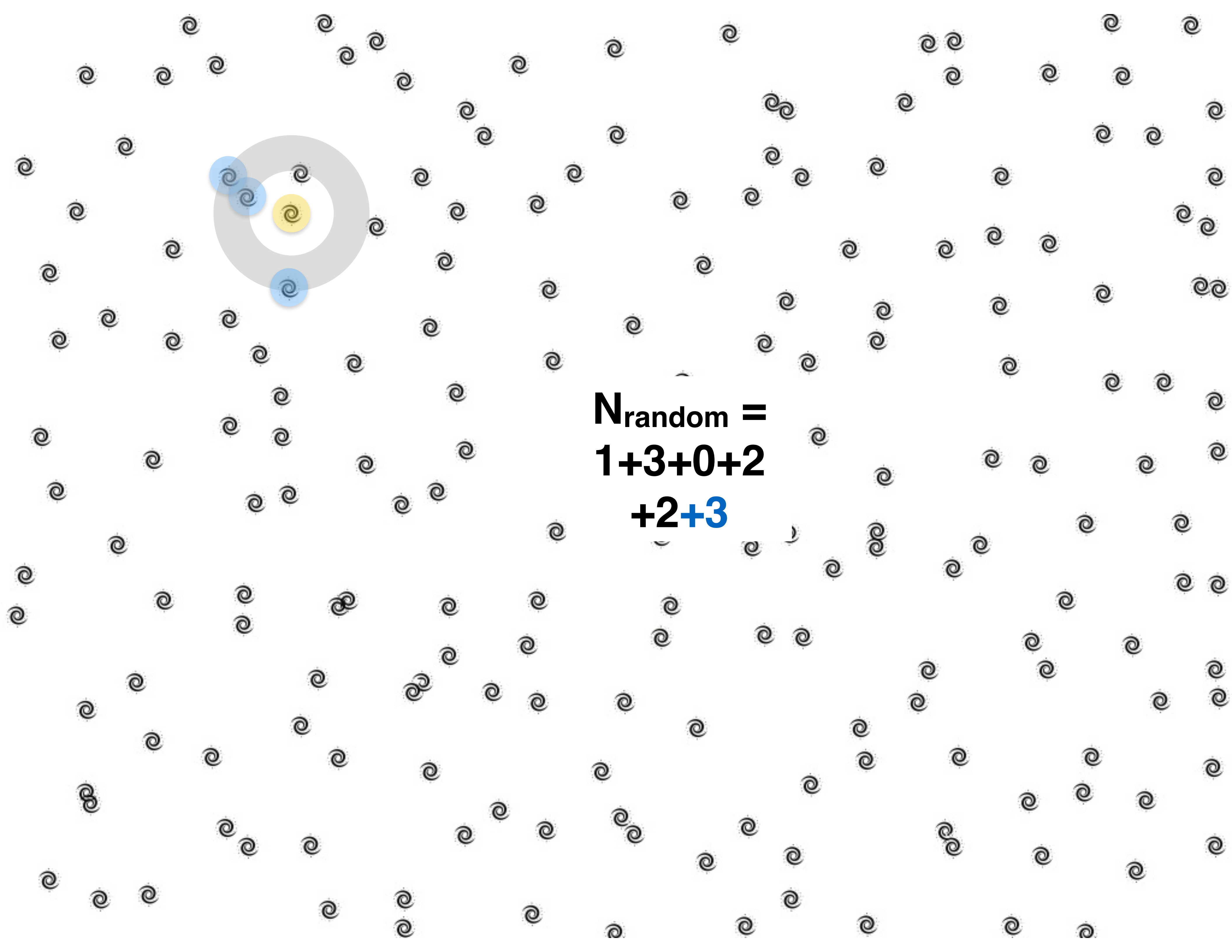

Nrandom = 1


N_{random} =
1+3


N_{random} =
1+3+0


N_{random} =
1+3+0+2


N_{random} =
1+3+0+2
+2



N_{random} =
1+3+0+2
+2+3


N_{random} =
1+3+0+2
+2+3+3


N_{random} =
1+3+0+2
+2+3+3+2
=16

In our “random Universe” we made measurements around 8 galaxies within a distance of r between 1 and 2 *(or, r=1.5)*, and counted a total of 16:

N_{random} = 16

We now have enough information to calculate the “two-point correlation function” **w(r) at r=1.5**:

$$w(r = 1.5) = \frac{N_{real}}{N_{random}} - 1$$

$$= \frac{20}{16} - 1$$

$$= 1.25 - 1.$$

$$= 0.25$$

$$w(r = 1.5) = \frac{N_{real}}{N_{random}} - 1 = 0.25$$

So what does this “clustering” measurement of 0.25 actually tell us?

Well, if galaxies in our “real” Universe were not clustered at all, then the measurement of $\boldsymbol{N_{real}}$ would be approximately the same as our measurement for $\boldsymbol{N_{random}}$... meaning:

$$\boldsymbol{N_{real} = N_{random}}$$

$$w(r = 1.5) = \frac{N_{real}}{N_{random}} - 1$$

$$= \frac{1}{1} - 1$$

$$= 1 - 1$$

$$= 0$$

$$w(r=1.5)=\frac{N_{real}}{N_{random}}-1=0.25$$

So what does this "clustering" measurement of 0.25 actually tell us?

The fact that 0.25 > 0.0 tells us that the "real" Universe we've measured is <u>more clustered than a random distribution</u>.

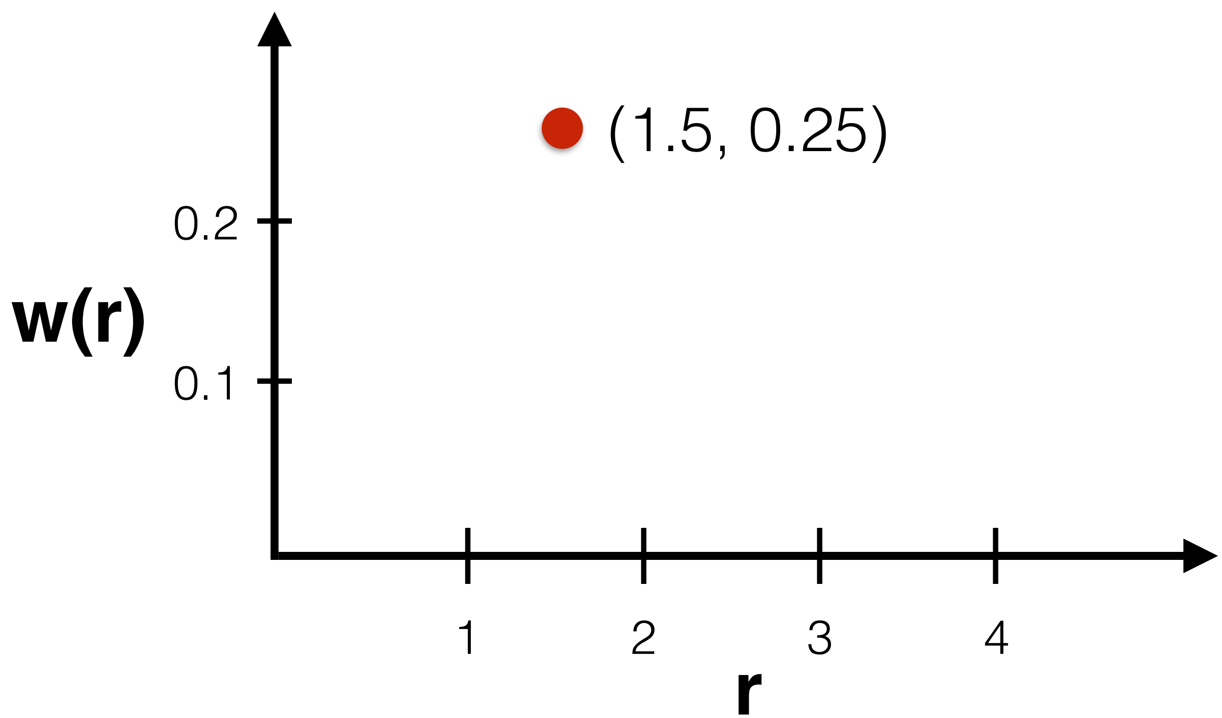
Remember now that
the “two-point correlation function” **w(r)** looks like this:

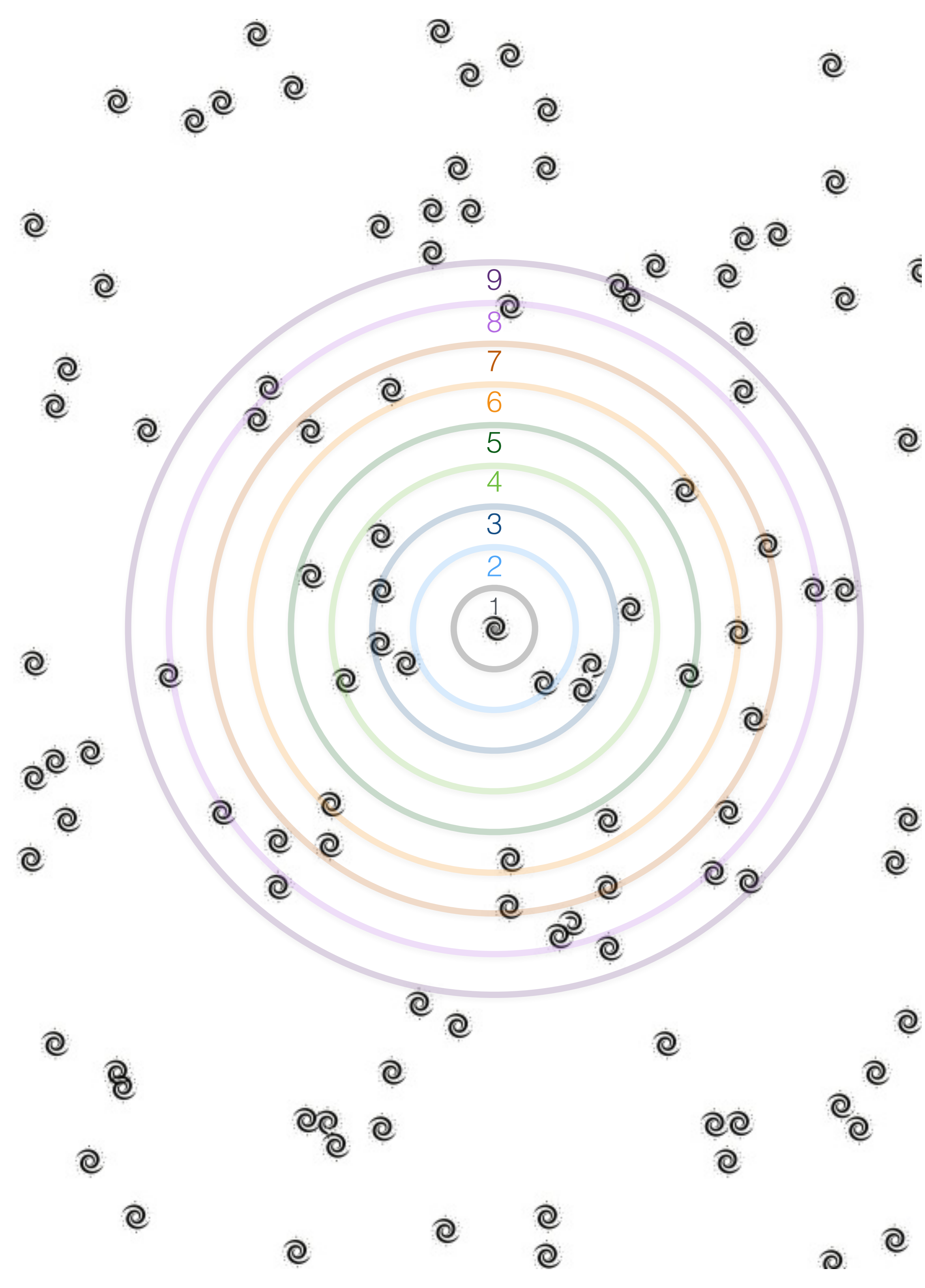
$$w(r) = \frac{N_{real}}{N_{random}} - 1$$

which means, we really want to do this same counting technique for a range of distances, to see how the amount of clustering depends on the physical scale (r).

$$w(r) = \frac{N_{real}}{N_{random}} - 1$$

We have already made a measurement at r=1.5, but we would need to make some more!





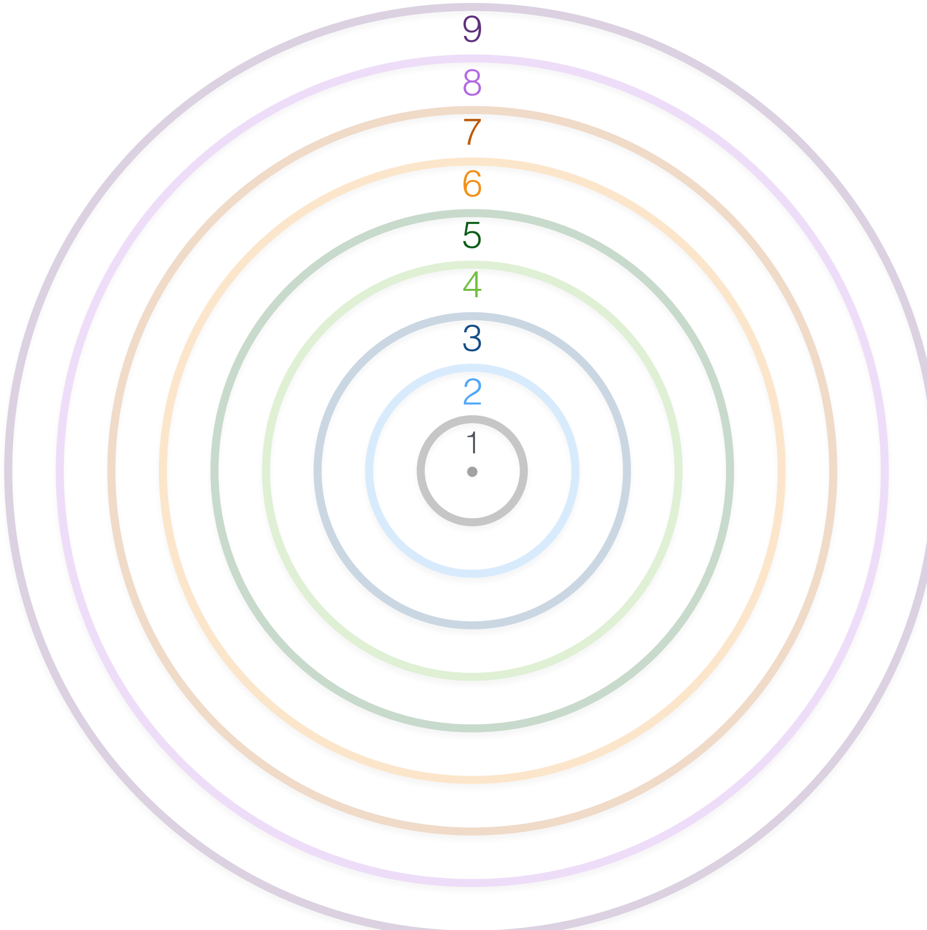
Directions:

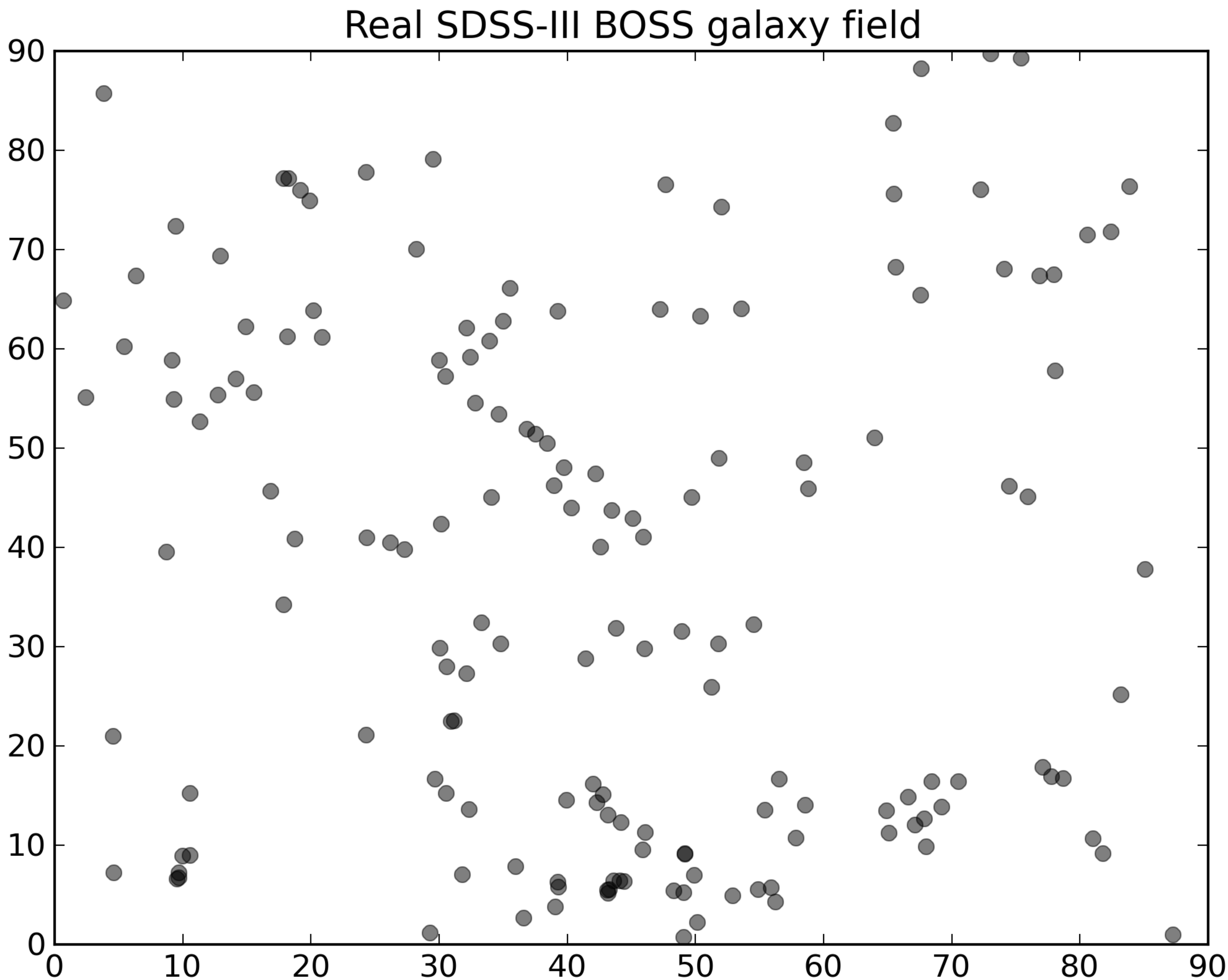
1. Print the rainbow-colored annulus tool on the next page onto a transparency.

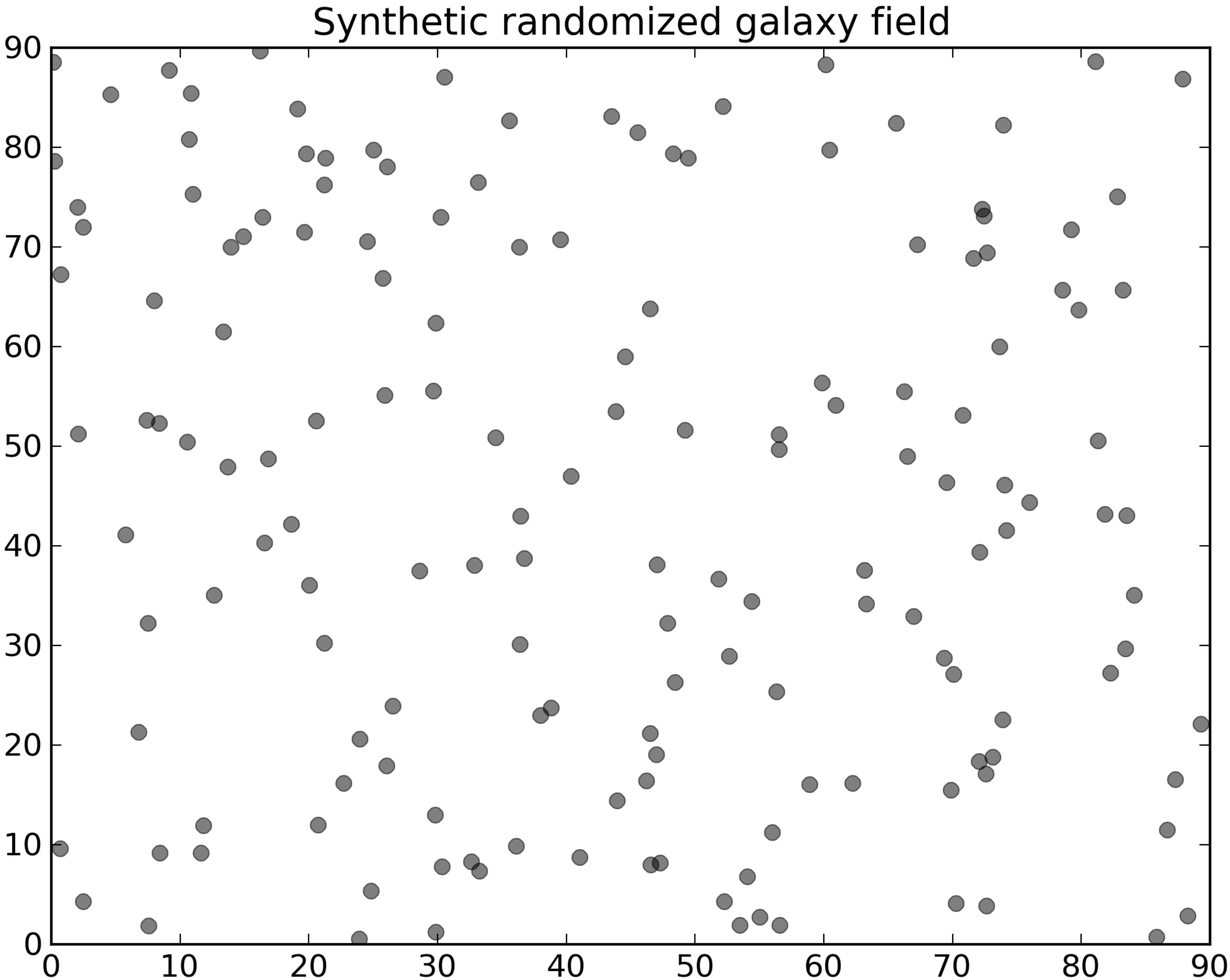
2. Then try measuring the two-point correlation function over a range of scales, using the real and random galaxy fields printed on the following pages.

3. Compare the shape you see to the SDSS measurements made with millions of galaxies. Are there similarities?

Print on transparency




Real SDSS-III BOSS galaxy field
0
10
20
30
40
50
60
70
80
90



Synthetic randomized galaxy field
0
10
20
30
40
50
60
70
80
90

Mid/Post-Activity Discussion:

1) *This is tedious!! How do astronomers do this sort of counting for observations of millions of galaxies?*

Astronomers in the SDSS like Prof. Alex Szalay (shown below) use powerful computers to measure the clustering in vast catalogs of galaxies in reasonable amounts of time. This way, the measuring happens fast, and no one gets bored!

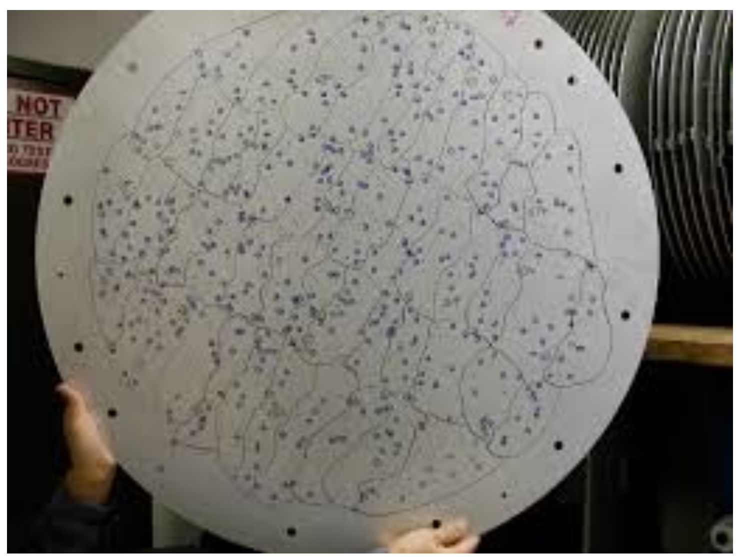
Prof. Alexander Szalay of Johns Hopkins University and the SDSS
(Photo: Joey Pulone / The Chronicle)

Mid/Post-Activity Discussion:

2) How do the holes on the aluminum plug plates relate to measurements of galaxy clustering?

Most of the holes on a BOSS plug plate line up with the positions of distant galaxies and quasars. Astronomers in the SDSS use the positions of these galaxies are used to measure the real clustering of large-scale structure in the Universe. With modern computers, these measurements are relatively quick and easy, but that only gives us the N_real measurement for the cross-correlation equation.

Determining what an unclustered Universe might look like, in order to measure N_random (the denominator for our clustering measurement), is often the hard part! To do this, SDSS astronomers use computers to make vast catalogs of random positions on the sky where galaxies could theoretically be observed. Often dust or bright stars obscure our view of distant galaxies, so making these maps realistic is tricky and also critical to generating accurate measurements of the BAO.



Mid/Post-Activity Discussion:

3) Who are the SDSS astronomers?

Astronomers in the SDSS are a friendly and diverse, international group of scientists, engineers, and technicians — who range in experience from undergraduate students to senior scientists.

2014 SDSS-III Collaboration Meeting
Park City, Utah

Activity Extension: Understanding Signal-to-Noise

As you've probably discovered, measuring the two-point correlation of galaxy clustering by hand is a very time-consuming exercise! And depending on how many measurements you've made, the results may look pretty noisy.

Often it takes a large number of measurements to see a clear signal in data. Take for example, a game in which many people guess the number of jelly beans in a jar. Each person's guess (measurement) will be imprecise on its own… but taken together, a normal distribution of measurements will be peaked toward precisely the right answer!

If you have many people working on the two-point correlation measurement activity, <u>try working individually or in small groups</u>. Each individual or group will naturally choose different galaxies on which to center their counting.

Make a plot of each individual/group's two-point correlation measurements, and then compare it to the **<u>average</u>** of everyone's combined. Do the measurements look less scattered? What if you quantify the noise? Try approximating the Poisson error bars for the individual/group and combined average measurements. How do the errors compare between the individual/group measurements and the average of them all?

Activity Extension: Understanding Cosmic Variance

Astronomers now know from observations that the universe is generally **homogeneous** (similar in its makeup) and **isotropic** (similar in the distribution of its contents in all directions). But one still needs large samples to measure these average properties. For instance, if you only measured the clustering of galaxies in a one or two very small parts of the sky (like Areas 1 & 2 in the picture below), you might find the universe to be much more or much less clustered than it truly is on average. This problem is referred to in astronomy as "**cosmic variance**".

You might have already noticed that your measurement of the two-point correlation in this activity will partly depend on which galaxies in the field you included in your sample. This is why one individual/group's measurements might not be the same as another's, and it contributes to the overall "noise" in the two-point correlation measurement.

The size of the differences between the results from each individual/group can be used to measure the effects of cosmic variance.

Combining measurements from each individual/group to determine an average two-point correlation:

1) increases the signal-to-noise, and
2) helps to remove the noisy effects of cosmic variance and get closer to the true physical measurement.