

A brief introduction to Bayesian Statistics through Astronomical Applications (Lecture 3)

Cecilia Mateu J.

Centro de Investigaciones de Astronomía (CIDA)

Universidad de Barcelona

22 de noviembre de 2017

More Examples

Coordinate Transformations

- ◆ Lets say we have a problem in which we know how to write our posterior $P(x)$ in a variable x , but in reality we are interested in knowing the posterior $P(y)$ for variable $y=f(x)$
- ◆ We know the integral of the probability must be conserved

$$\int P(x)dx = \int P(y)dy \qquad P(x)dx = P(y)dy$$

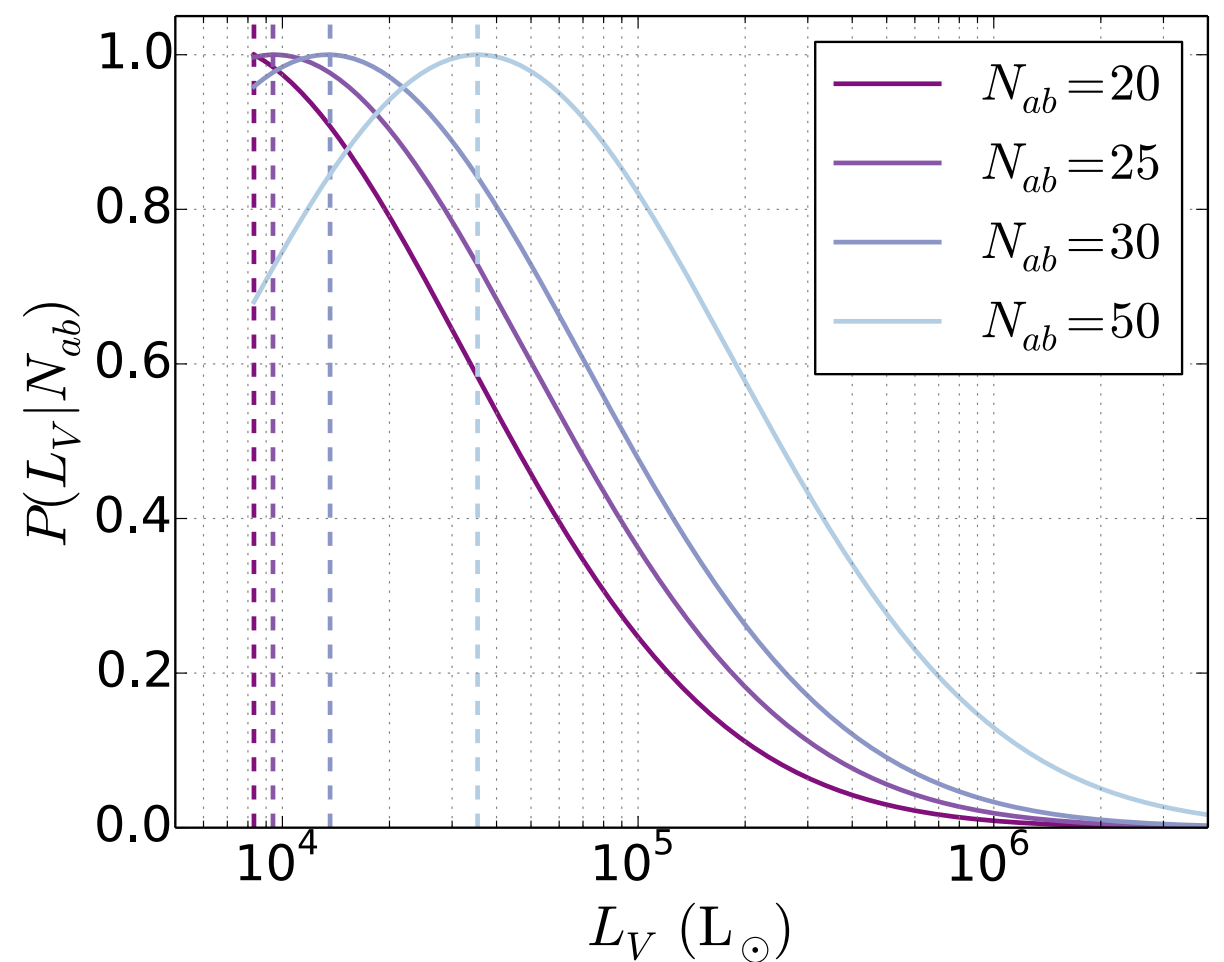
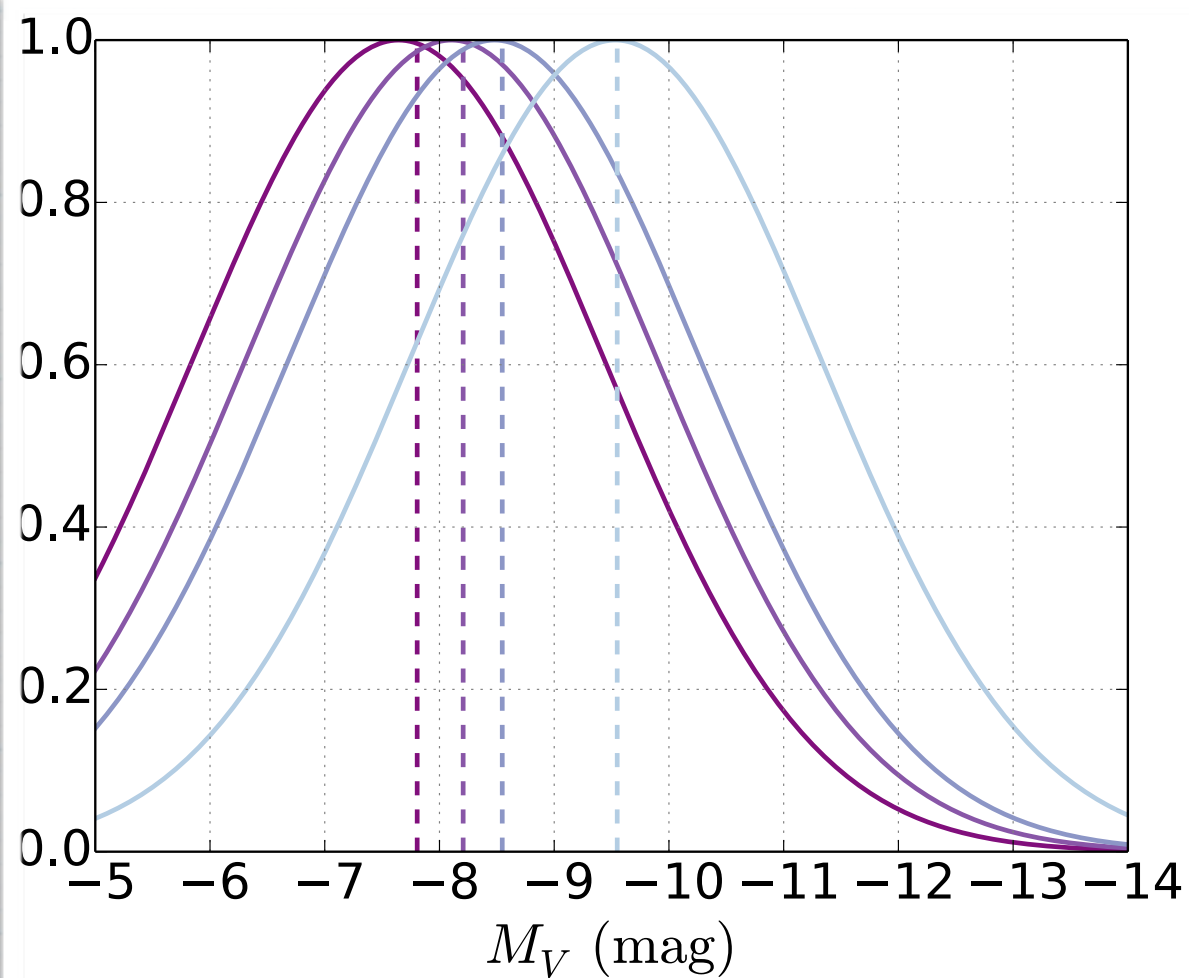
- ◆ This makes it easy, we know how to change variables inside an integral

$$P(y) = \left| \frac{dx}{dy} \right| P(x)$$

← Jacobian of the transformation $f(x)$

The RR Lyrae problem from L1

$$P(\dot{L}_V | \log N_{ab}) = P(M_V(L_V) | \log N_{ab}) / L_V$$



Error Propagation

Error Propagation

- ◆ Lets say we know the PDF for two variables X and Y

$$P(X) \quad p(Y) \quad |data, I$$

- ◆ Now we'd like to know what is the PDF for Z , where

$$Z = X + Y$$

- ◆ First, lets look at the case when X and Y are conditionally independent. If this is the case, the joint probability $P(X,Y)=P(X)p(Y)$, from the Marginalization rule this is:

$$P(Z) = \iint P(X)p(Y)dXdY \quad |data, I$$

for all X, Y
t.q. $Z=X+y$

Error Propagation

- ♦ First, let's look at the case when X and Y are conditionally independent. If this is the case, the joint probability $P(X, Y) = P(X)p(Y)$, from the Marginalization rule this is:


$$P(Z) = \iint P(X)p(Y)dXdY \quad |data, I$$

for all X, Y
t.q. $Z = X + Y$

$$P(Z) = \iint P(X)p(Y)\delta(Z - [X + Y])dXdY$$

$$P(Z) = \int P(X)p(Y = Z - X)dX$$

**note this is
the
convolution of
 $P(X)$ and $p(Y)$**



Error Propagation

- ♦ If $P(X)$ and $p(y)$ are gaussians, e.g. lets say like yesterday we have a model with gaussian uncertainties and a uniform prior,

$$P(X) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{(X-X_o)^2}{2\sigma_X^2}} \quad P(Y) = \frac{1}{\sqrt{2\pi}\sigma_Y} e^{-\frac{(Y-Y_o)^2}{2\sigma_Y^2}}$$

then

$$P(Z) = \int P(X)p(Y = Z - X)dX \quad |data, I$$

is the convolution of these two gaussians

$$P(Z) = \frac{1}{2\pi\sigma_X\sigma_Y} \int_{-\infty}^{+\infty} e^{-\frac{(X-X_o)^2}{2\sigma_X^2}} e^{-\frac{(Z-X-Y_o)^2}{2\sigma_Y^2}} dX$$

Error Propagation

- ◆ After completing squares and simplifying we get

$$P(Z) = \frac{1}{\sqrt{2\pi}\sigma_z} e^{-\frac{(Z-Z_o)^2}{2\sigma_z^2}}$$

where

$$Z_o = X_o + Y_o \quad \text{and} \quad \sigma_z^2 = \sigma_X^2 + \sigma_Y^2$$

so, the sum in quadrature rule is derived

- ◆ Lets go back to the general problem, where $Z=f(X,Y)$

Error Propagation

- ◆ In general, i.e. without assuming cond. independence of X, Y we have

$$P(Z) = \iint P(X, Y) dX dY$$

for all X, Y
t.q.
 $Z=f(X, Y)$

- ◆ Since $Z=f(X, Y)$ we have

$$P(Z) = \iint P(X, Y) \delta(Z - f(X, Y)) dX dY$$

- ◆ ... what follows will depend on the exact form of $P(X, Y)$...

More Examples

Fitting a density profile to data

- Another example. Lets say you have a catalogue of a particular stellar tracer, e.g. RR Lyrae stars, Red Clump stars, Cepheids or whatever.
- We'd like to use these tracers to study how stars are distributed in space in the different components in the Galaxy, i.e. their density profile ρ
- How do we compare and fit this function to a sample of stars?

Thick Disc and Halo RRL Density Profiles

- ♦ Thick Disc density profile

$$\rho_{tkd} = e^{-\frac{R-R_{\odot}}{h_R}} e^{-\frac{|z|}{h_z}}$$

- ♦ Halo density profile

$$\rho_H = \frac{1}{R_{\odot}^n} \left[R^2 + z^2 \right]^{n/2}$$

- ♦ Total density profile:

$$\rho = N_{mod} \left(\rho_{tkd} + C_{hd} \rho_H \right) = N_{mod} \hat{\rho}$$

$$N_{RR} = \iiint_{V_S} \rho(\vec{r}) dV$$

... but the density ρ is not a directly observable quantity! all we have are the total number N_{OBS} of RRLs in the survey volume V_S and their observed positions

Density Profiles: A Bayesian approach

- ◆ Our free parameters are:

$$\vec{\theta} = \{(h_z, h_R, n, \hat{C}_{hd}), N_{mod}\}$$

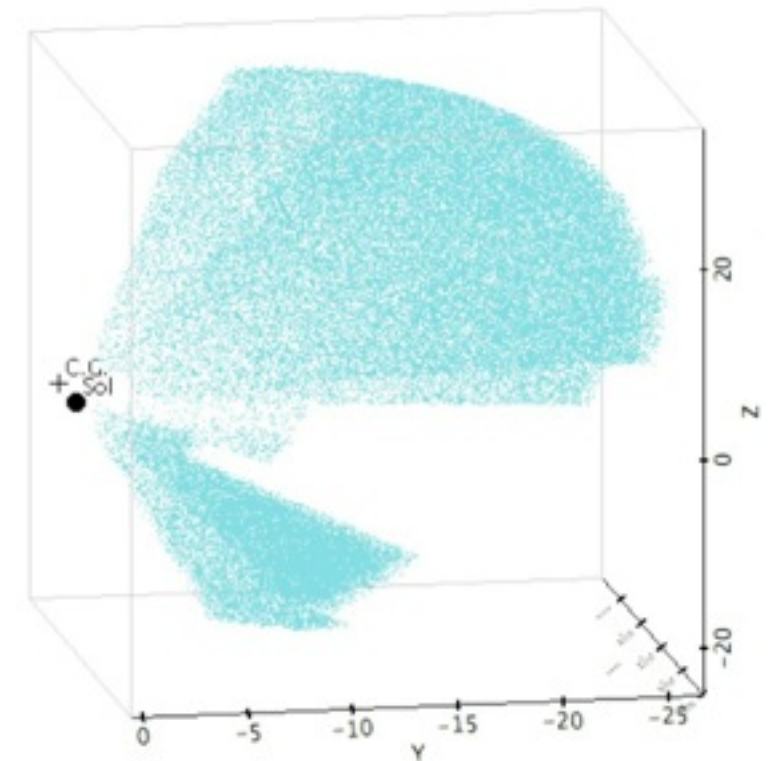
- ◆ Bayes's Theorem is

$$P(\vec{\theta} | N_{obs}, \{\vec{r}_i\}) = L(N_{obs}, \{\vec{r}_i\} | \vec{\theta}) P(\vec{\theta}) \quad | I$$

- ◆ Our data is the total number **N_{obs}** of RRLs in the survey volume, and the individual positions \vec{r}_i

- ◆ The Likelihood function can be expressed as

$$L(N_{obs}, \{\vec{r}_i\} | \vec{\theta}) = P(N_{obs} | \vec{\theta}) P(\{\vec{r}_i\} | \vec{\theta})$$



Density Profiles: A Bayesian approach

- Our free parameters are:

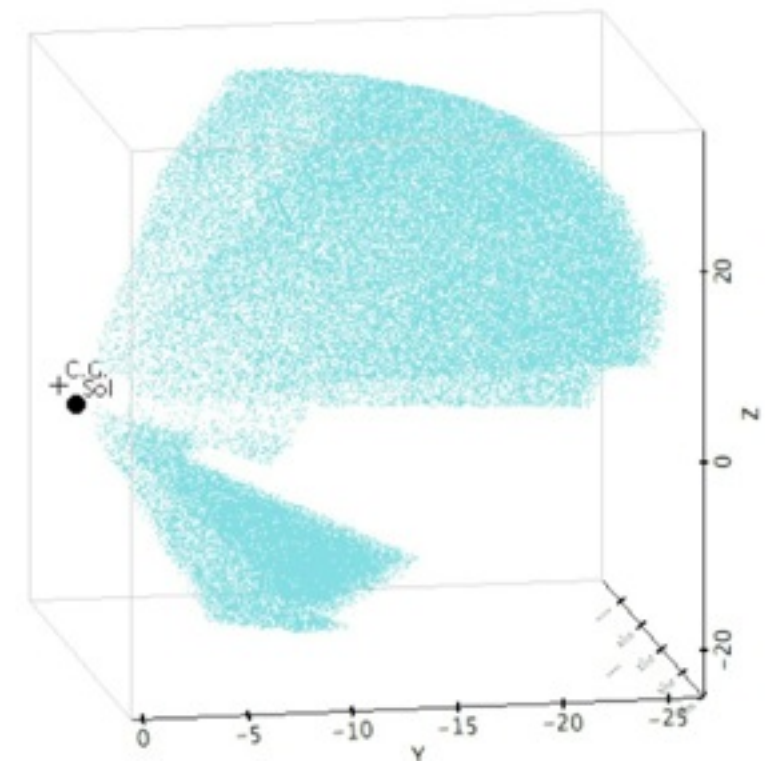
$$\vec{\theta} = \{(h_z, h_R, n, \hat{C}_{hd}), N_{mod}\}$$

- Our data is the total number N_{obs} of RRLs in the survey volume, and the individual positions \vec{r}_i
- The Likelihood function can be expressed as

$$L(N_{obs}, \{\vec{r}_i\} | \vec{\theta}) = \underbrace{P(N_{obs} | \vec{\theta})}_{e^{-N_{mod}} (N_{mod})^{N_{obs}}} \prod_{i=1}^{N_{obs}} \underbrace{P(\vec{r}_i | \vec{\theta})}_{\hat{\rho}(\vec{r}_i | \vec{\theta})}$$

Poisson probability of observing N_{obs} given predicted number N_{mod}

Density profile (normalized to integrate 1 over the survey volume)



Density Profiles: A Bayesian approach


- ◆ So we have

$$L(N_{obs}, \{\vec{r}_i\} | \vec{\theta}) = \left(e^{-N_{mod}} (N_{mod})^{N_{obs}} \right) \left(\prod_{i=1}^{N_{obs}} \hat{\rho}(\vec{r}_i | \vec{\theta}) \right)$$

- ◆ which we can write as a log-likelihood for a more convenient computation as:

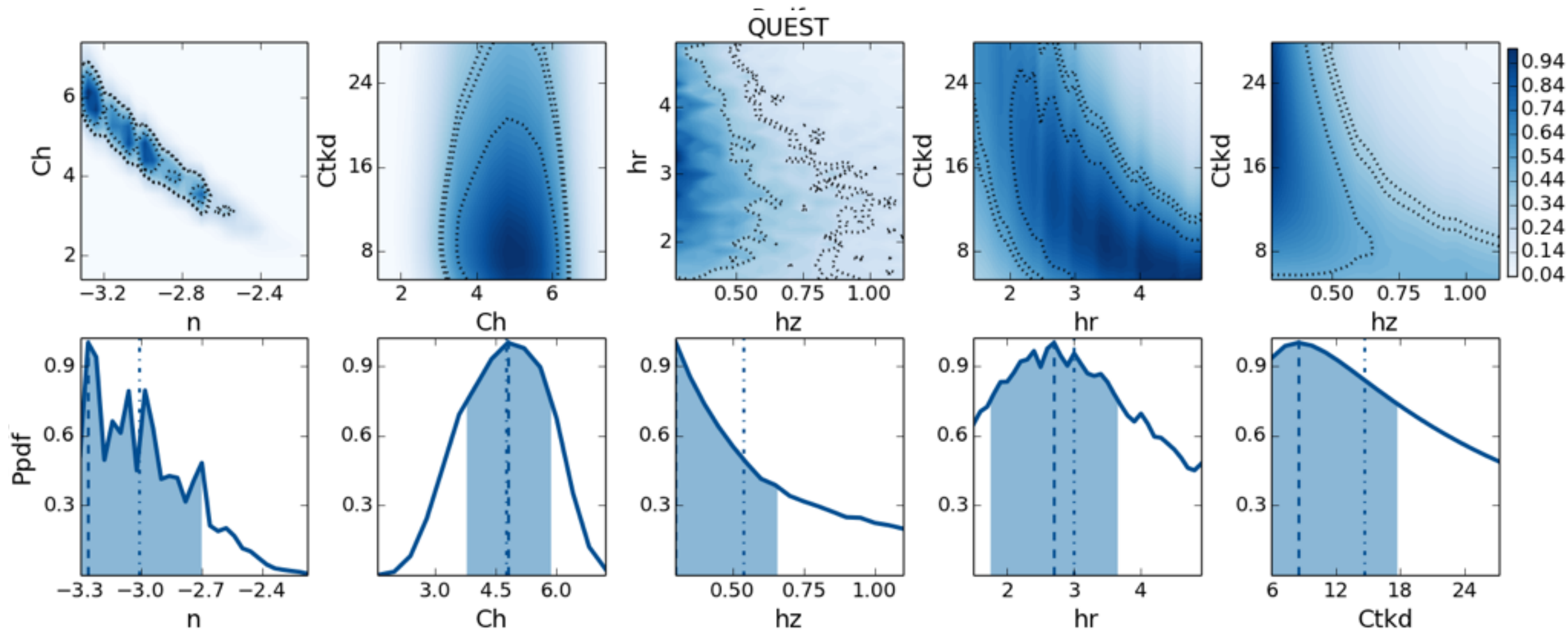
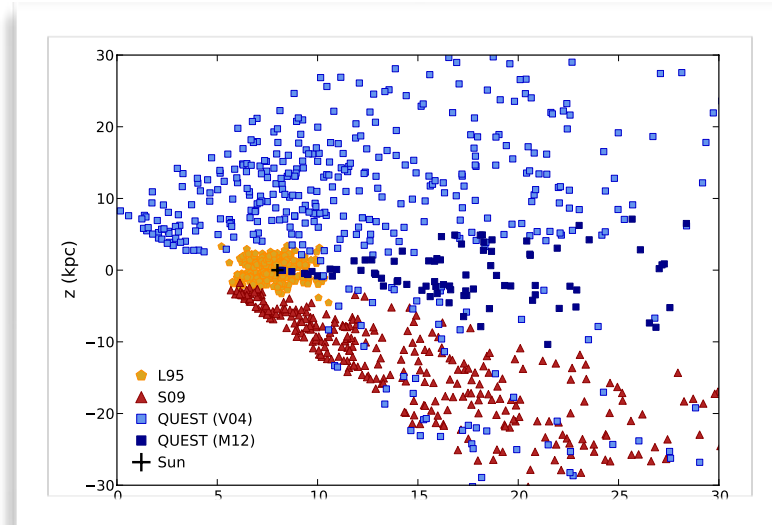
$$\ln L(N_{obs}, \{\vec{r}_i\} | \vec{\theta}) = \left(N_{obs} \ln N_{mod} - N_{mod} \right) + \sum_{i=1}^{N_{obs}} \ln \hat{\rho}(\vec{r}_i | \vec{\theta})$$

This is the computationally intensive part


$$N_{RR} = \iiint_{V_S} \rho(\vec{r}) dV$$

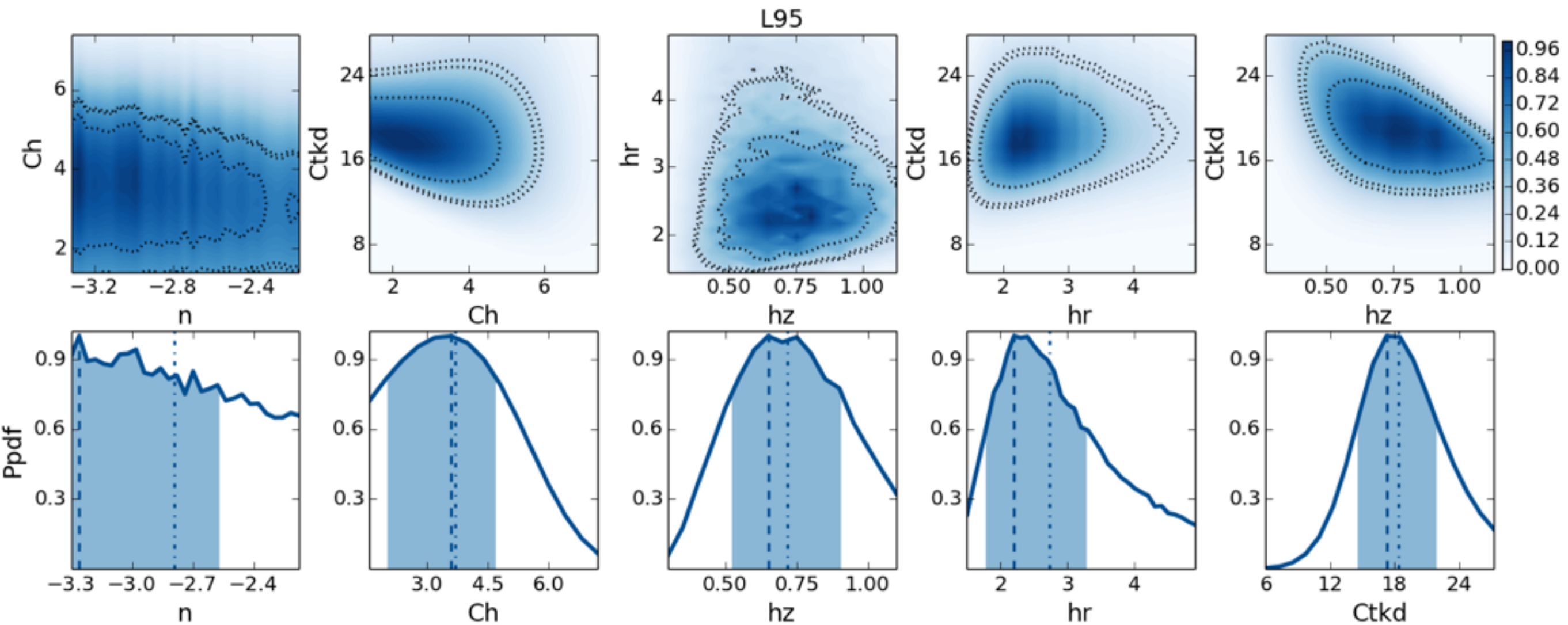
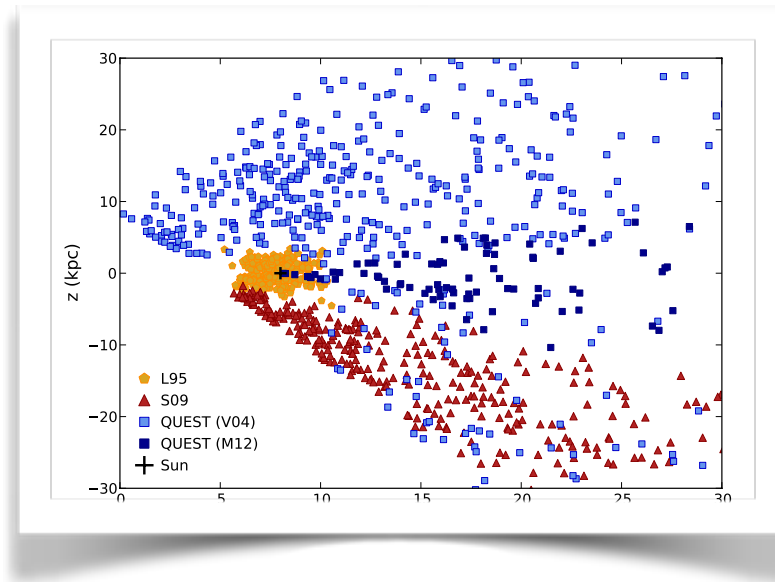
Density Profiles: One sample at a time

The marginal posteriors taking only the **QUEST sample** are as follows:



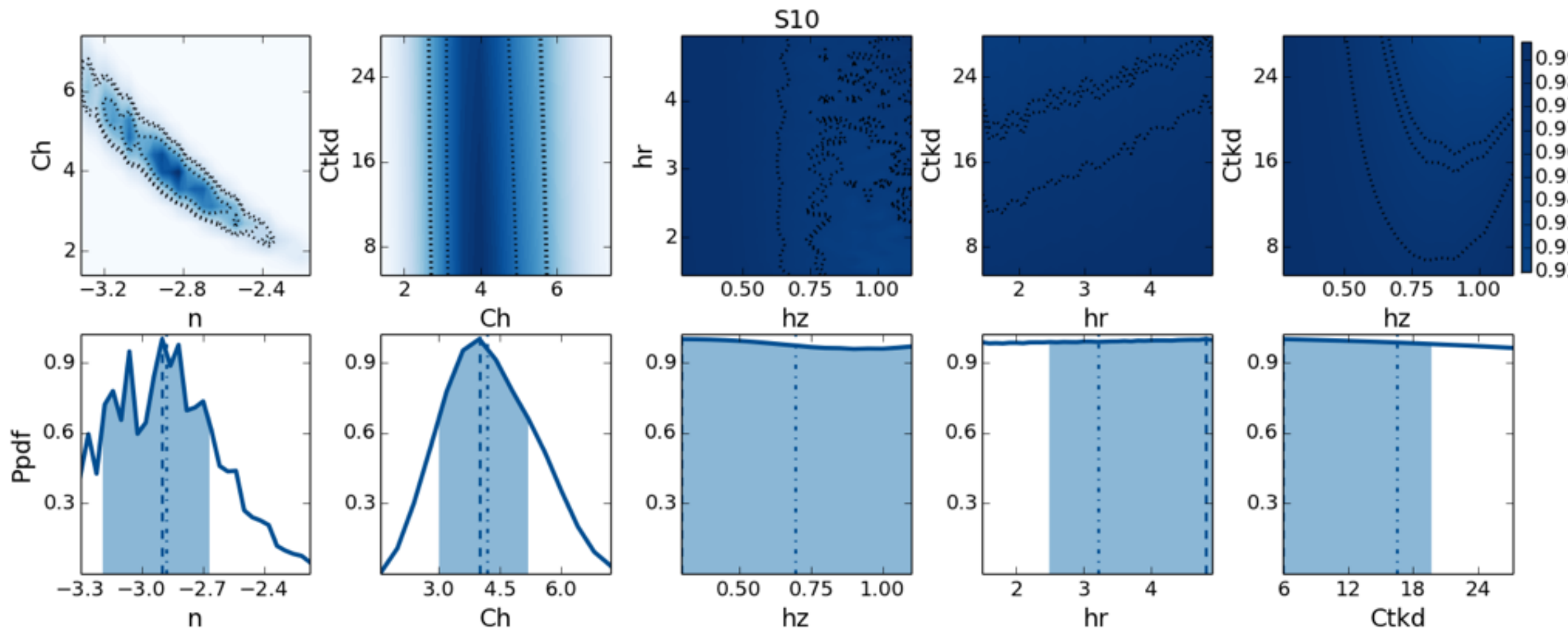
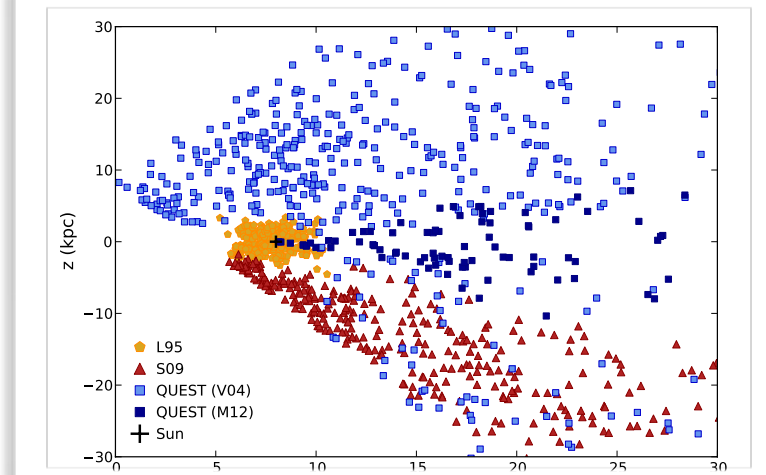
Density Profiles: One sample at a time

The marginal posteriors taking only the **Layden sample** are as follows:



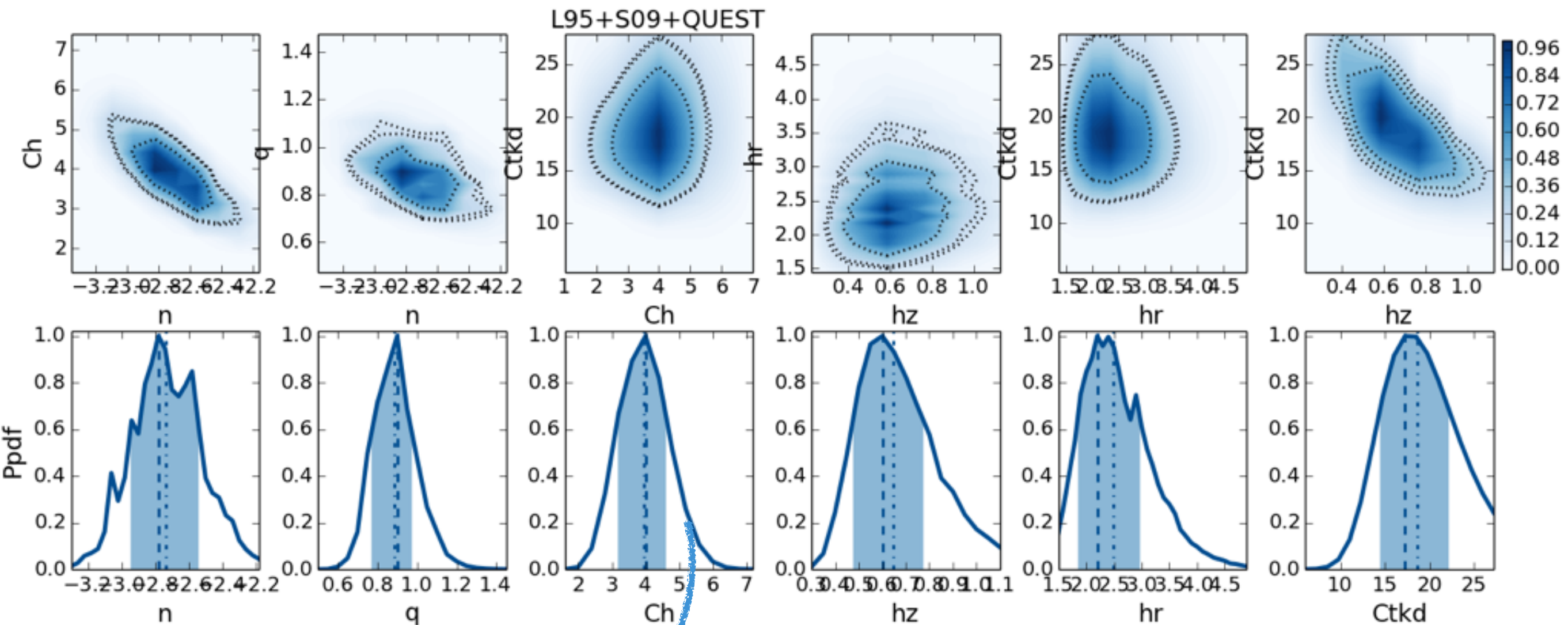
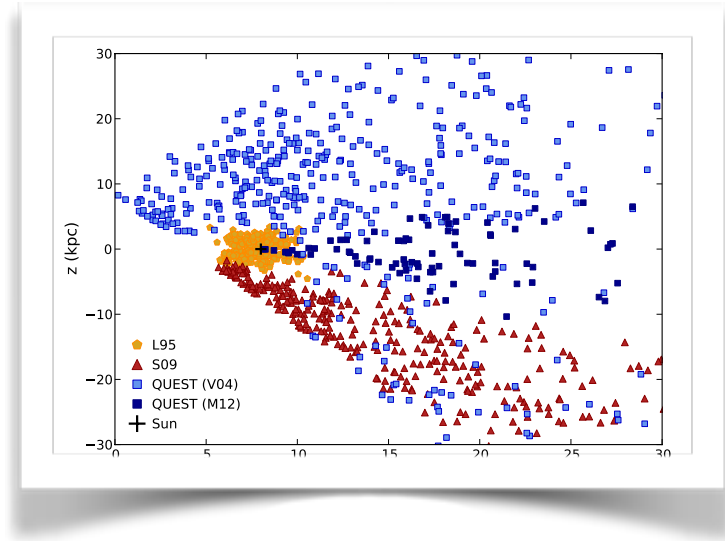
Density Profiles: One sample at a time

The marginal posteriors taking only the **Sesar sample** are as follows:



Density Profiles: Combined samples

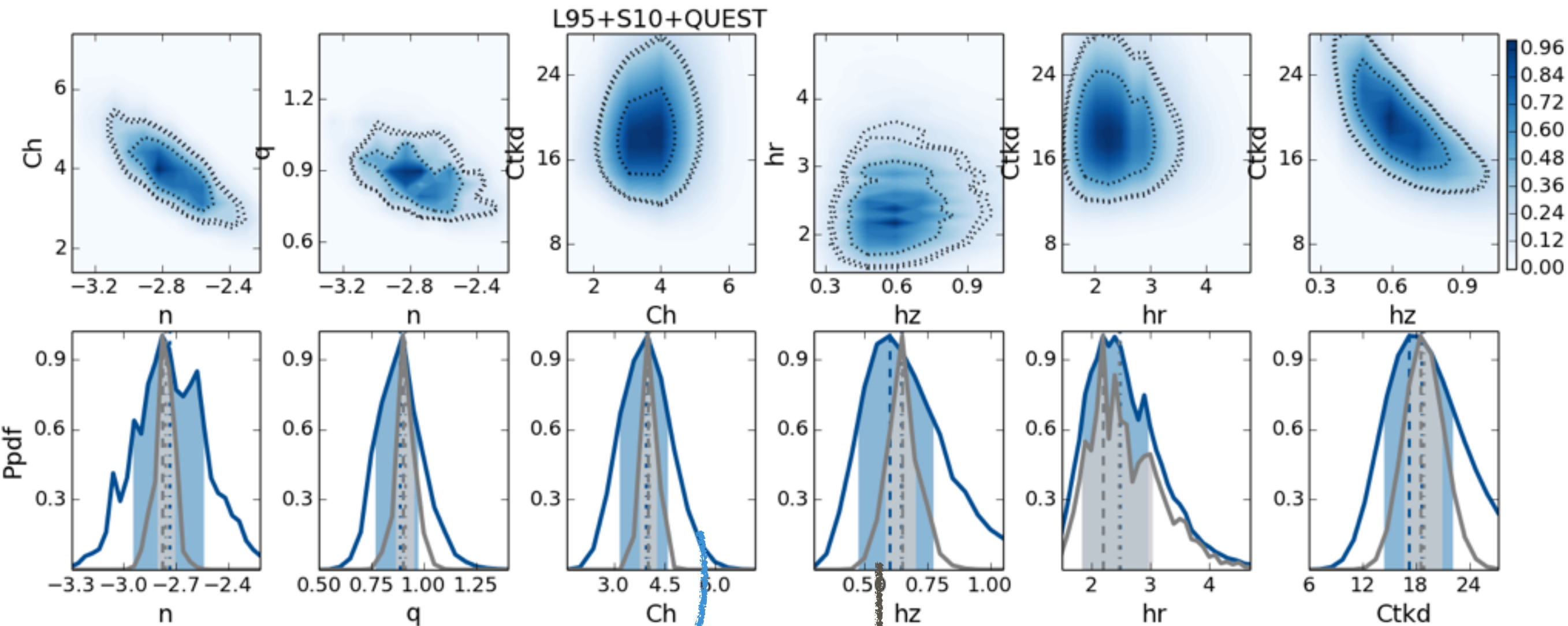
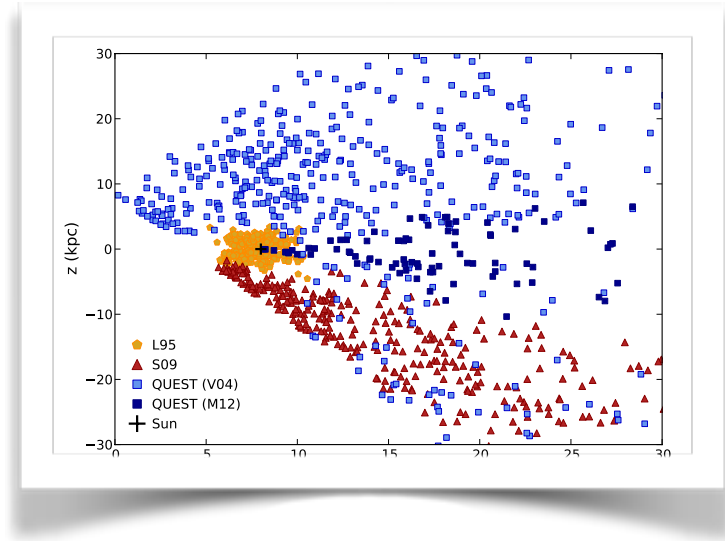
Combining three different samples we find these marginal posteriors (remember its the product of the individual pdfs in 5-D space, and then marginalizing):



marginal distribution

Density Profiles: Combined samples

Combining three different samples we find these marginal posteriors (remember its the product of the individual pdfs in 5-D space, and then marginalizing):



marginal distribution

joint distribution