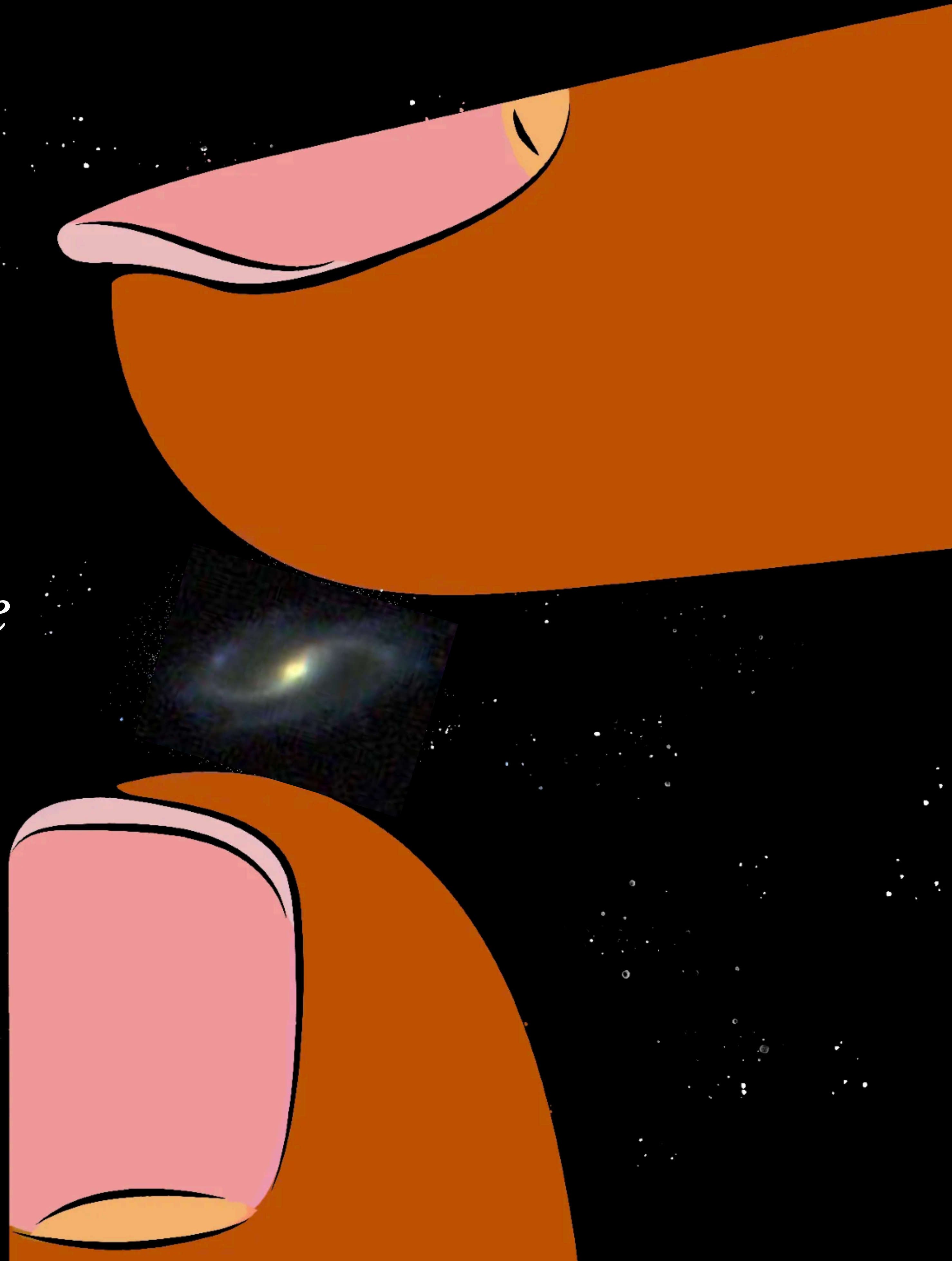
Cosmology with *Galaxy Photometry Alone*



Changhoon Hahn
changhoon.hahn@princeton.edu
Princeton University

ML IAP/CCA — November 30, 2023

The Astrophysical Journal, 929:132 (22pp), 2022 April 20 https://doi.org/10.3847/1538-4357/ac5d3f

Cosmology with One Galaxy?

Francisco Villaescusa-Navarro[1,2], Jupiter Ding[2], Shy Genel[1,3], Stephanie Tonnesen[1], Valentina La Torre[4], David N. Spergel[1,2], Romain Teyssier[2], Yin Li[1,5], Caroline Heneka[6], Pablo Lemos[7,8], Daniel Anglés-Alcázar[1,9], Daisuke Nagai[10], and Mark Vogelsberger[11]

[1] Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA; fvillaescusa@flatironinstitute.org
[2] Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA
[3] Columbia Astrophysics Laboratory, Columbia University, New York, NY 10027, USA
[4] Department of Physics and Astronomy, Tufts University, Medford, MA 02155, USA
[5] Center for Computational Mathematics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA
[6] University of Hamburg, Hamburg Observatory, Gojenbergsweg 112, D-21029 Hamburg, Germany
[7] Department of Physics and Astronomy, University of Sussex, Brighton, BN1 9QH, UK
[8] University College London, Gower Street, London, UK
[9] Department of Physics, University of Connecticut, 196 Auditorium Road, Storrs, CT 06269, USA
[10] Department of Physics, Yale University, New Haven, CT 06520, USA
[11] Kavli Institute for Astrophysics and Space Research, Department of Physics, MIT, Cambridge, MA 02139, USA
Received 2022 January 14; revised 2022 March 3; accepted 2022 March 11; published 2022 April 20

Abstract

Galaxies can be characterized by many internal properties such as stellar mass, gas metallicity, and star formation rate. We quantify the amount of cosmological and astrophysical information that the internal properties of individual galaxies and their host dark matter halos contain. We train neural networks using hundreds of thousands of galaxies from 2000 state-of-the-art hydrodynamic simulations with different cosmologies and astrophysical models of the CAMELS project to perform likelihood-free inference on the value of the cosmological and astrophysical parameters. We find that knowing the internal properties of a single galaxy allows our models to infer the value of $\Omega_{\rm m}$, at fixed $\Omega_{\rm b}$, with a $\sim$10% precision, while no constraint can be placed on σ_8. Our results hold for any type of galaxy, central or satellite, massive or dwarf, at all considered redshifts, $z \leqslant 3$, and they incorporate uncertainties in astrophysics as modeled in CAMELS. However, our models are not robust to changes in subgrid physics due to the large intrinsic differences the two considered models imprint on galaxy properties. We find that the stellar mass, stellar metallicity, and maximum circular velocity are among the most important galaxy properties to determine the value of $\Omega_{\rm m}$. We believe that our results can be explained by considering that changes in the value of $\Omega_{\rm m}$, or potentially $\Omega_{\rm b}/\Omega_{\rm m}$, affect the dark matter content of galaxies, which leaves a signature in galaxy properties distinct from the one induced by galactic processes. Our results suggest that the low-dimensional manifold hosting galaxy properties provides a tight direct link between cosmology and astrophysics.

Unified Astronomy Thesaurus concepts: Galaxy formation (595); Cosmological models (337); Astrostatistics (1882); Hydrodynamical simulations (767)

cosmology with *one galaxy*? trained using **CAMELS**

$$p(\,\Omega\,|\,\theta_g\,)$$

Villaescusa-Navarro et al. (2022)

cosmology with *one galaxy*? trained using **CAMELS**

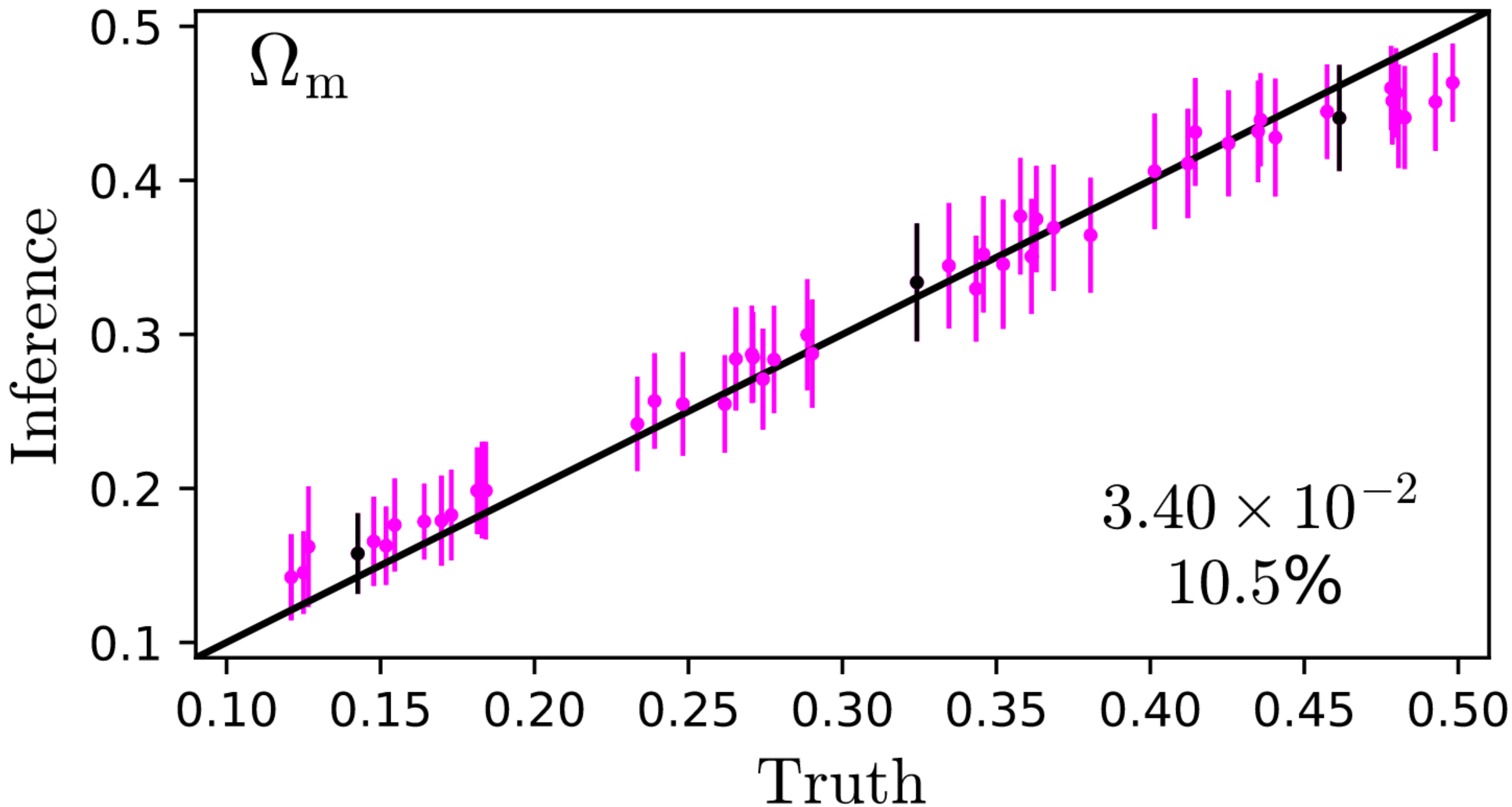
$$p(\Omega \mid \theta_g)$$

cosmological parameters *galaxy properties*

Ω_m, σ_8 $V_{\rm max}, M_*, M_{\rm gas}, Z_*, R_* \ldots$

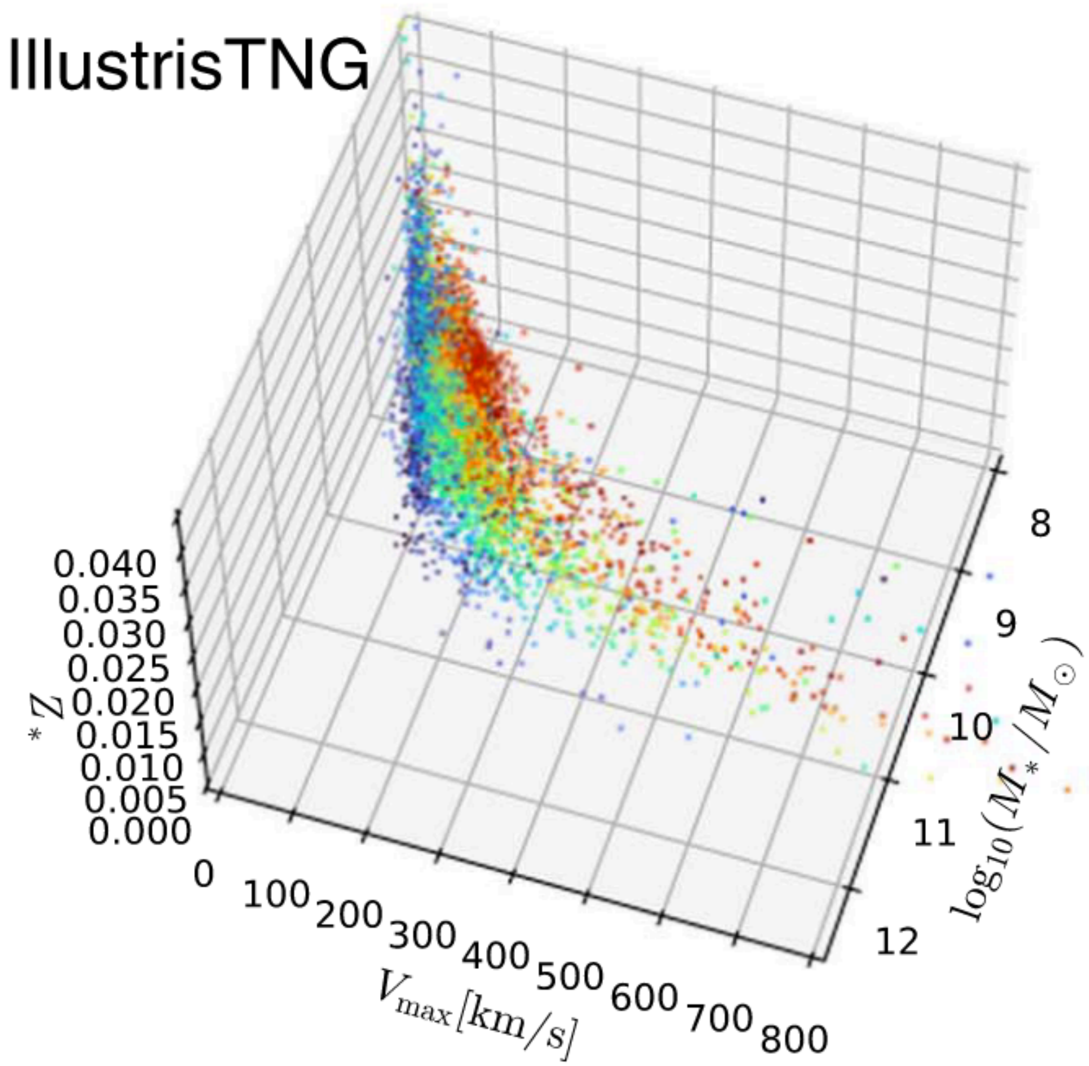
Villaescusa-Navarro et al. (2022)

cosmology with *one galaxy*?



Villaescusa-Navarro et al. (2022)

cosmology with *one galaxy*?



some low-dimensional manifold hosting θ_g?

Villaescusa-Navarro et al. (2022)

cosmology with *one galaxy*?

$$\frac{\Omega_m}{\Omega_b} \approx \frac{M_{\rm tot}}{M_b}$$

Villaescusa-Navarro et al. (2022)

cosmology with *one galaxy*?

$$\frac{\Omega_m}{\Omega_b} \approx \frac{M_{\rm tot}}{M_b} \quad = \quad \frac{M_{\rm tot}(V_{\rm max}\cdots)}{M_*/\epsilon_*(M_*, M_{\rm gas}, Z_*\ldots)}$$

star formation efficiency

Villaescusa-Navarro et al. (2022)

cosmology with *one galaxy*? — similar approach as *White et al. (**1993**)*

ARTICLES

The baryon content of galaxy clusters: a challenge to cosmological orthodoxy

Simon D. M. White*, Julio F. Navarro†, August E. Evrard‡ & Carlos S. Frenk†

* Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK
† Department of Physics, University of Durham, Durham DH1 3LE, UK
‡ Department of Physics, University of Michigan, Ann Arbor, Michigan 48109, USA

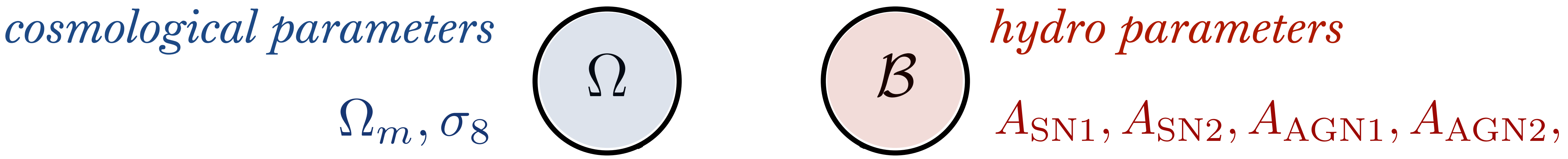
Baryonic matter constitutes a larger fraction of the total mass of rich galaxy clusters than is predicted by a combination of cosmic nucleosynthesis considerations (light-element formation during the Big Bang) and standard inflationary cosmology. This cannot be accounted for by gravitational and dissipative effects during cluster formation. Either the density of the Universe is less than that required for closure, or there is an error in the standard interpretation of element abundances.

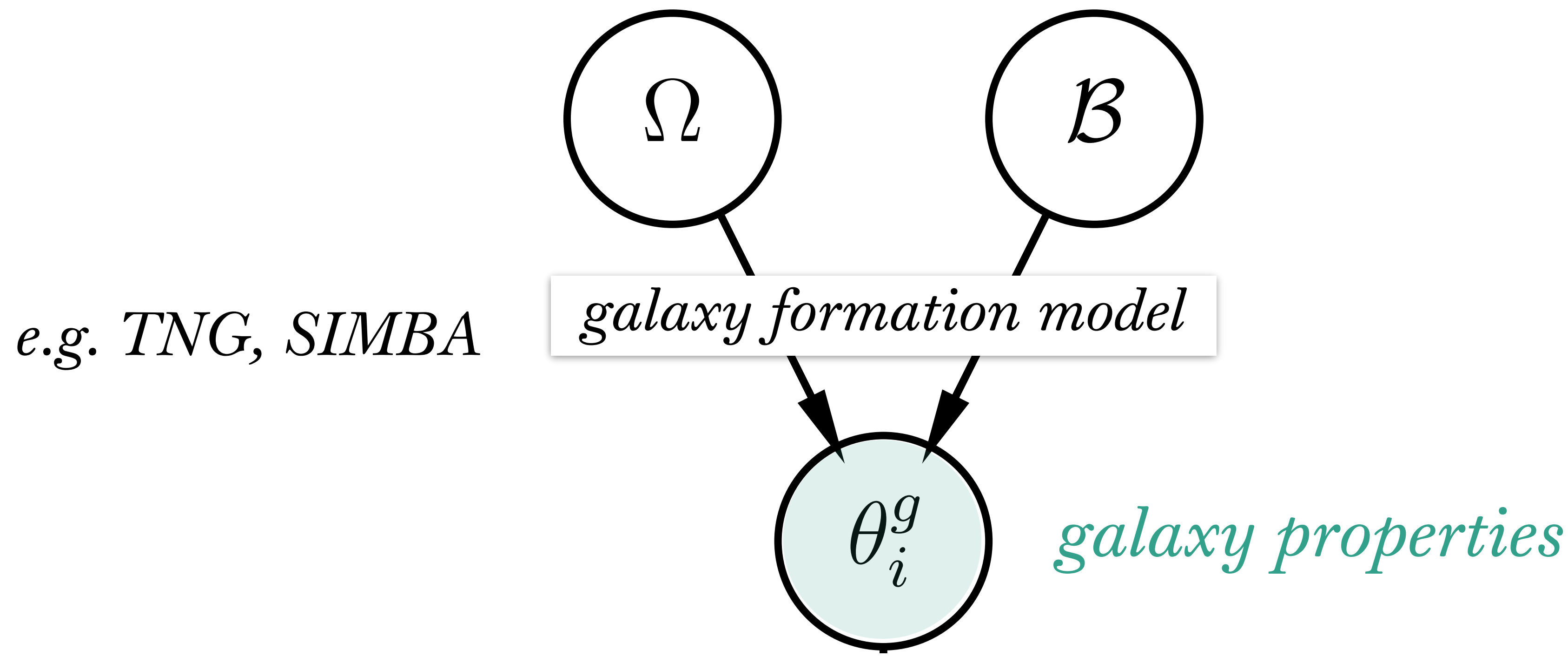
$$p(\,\Omega \mid \theta_g\,)$$

$\theta_g = \{V_{\rm max}, M_*, M_{\rm gas}, Z_*, R_* \ldots\}$ are ***not* observables**

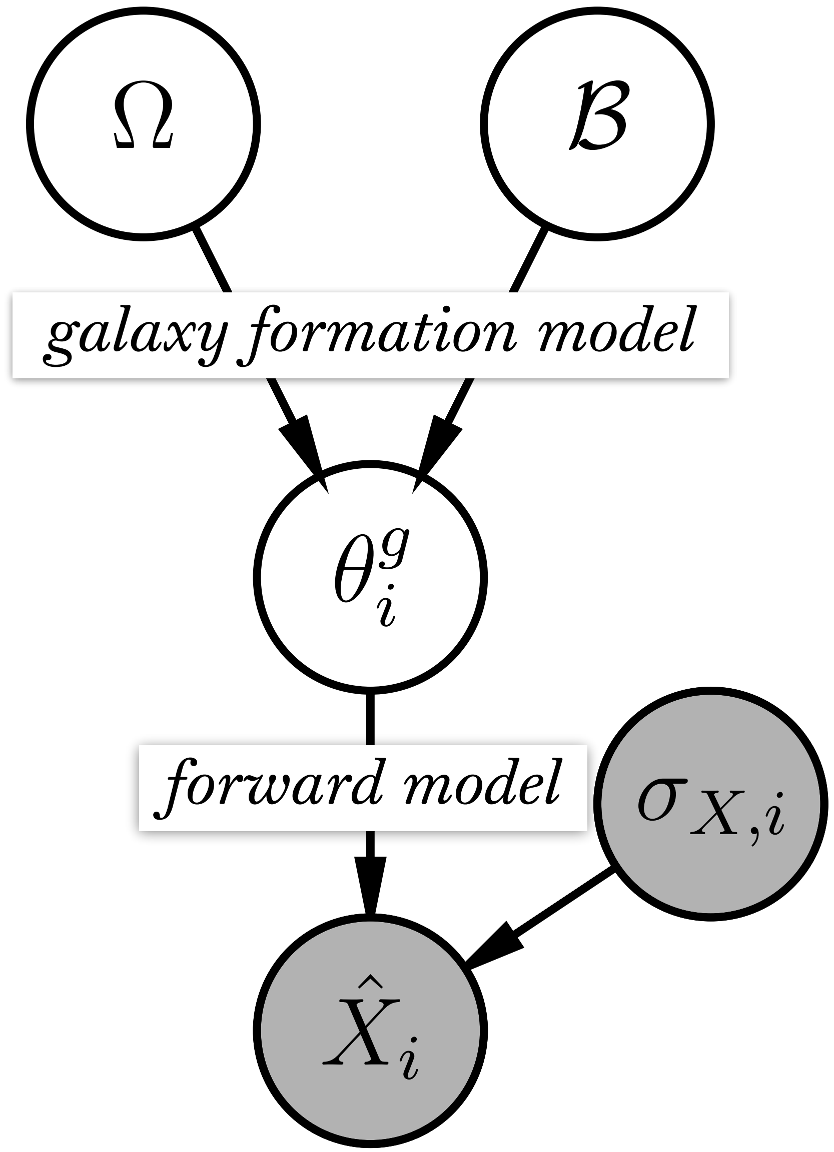
$$p(\,\Omega \mid X_i\,)\ \textbf{?}$$

cosmology with the X_i = **observables** of *one galaxy*?


cosmological parameters
Ω_m, σ_8
Ω
$\mathcal{B}$
hydro parameters
$A_{\rm SN1}, A_{\rm SN2}, A_{\rm AGN1}, A_{\rm AGN2},$



Ω
$\mathcal{B}$
e.g. TNG, SIMBA
galaxy formation model
θ_i^g
galaxy properties

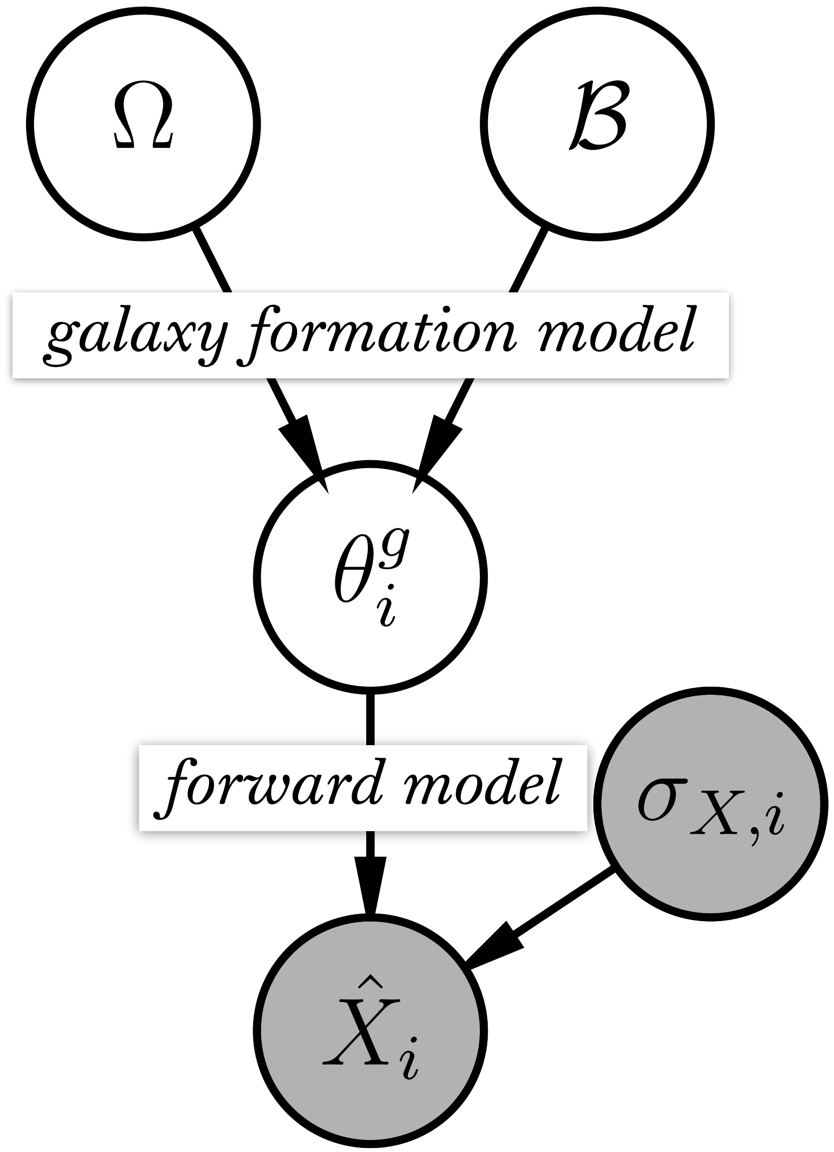


stellar population synthesis
dust model
noise model

Nelson et al. (2018a), Donnari et al. (2019)

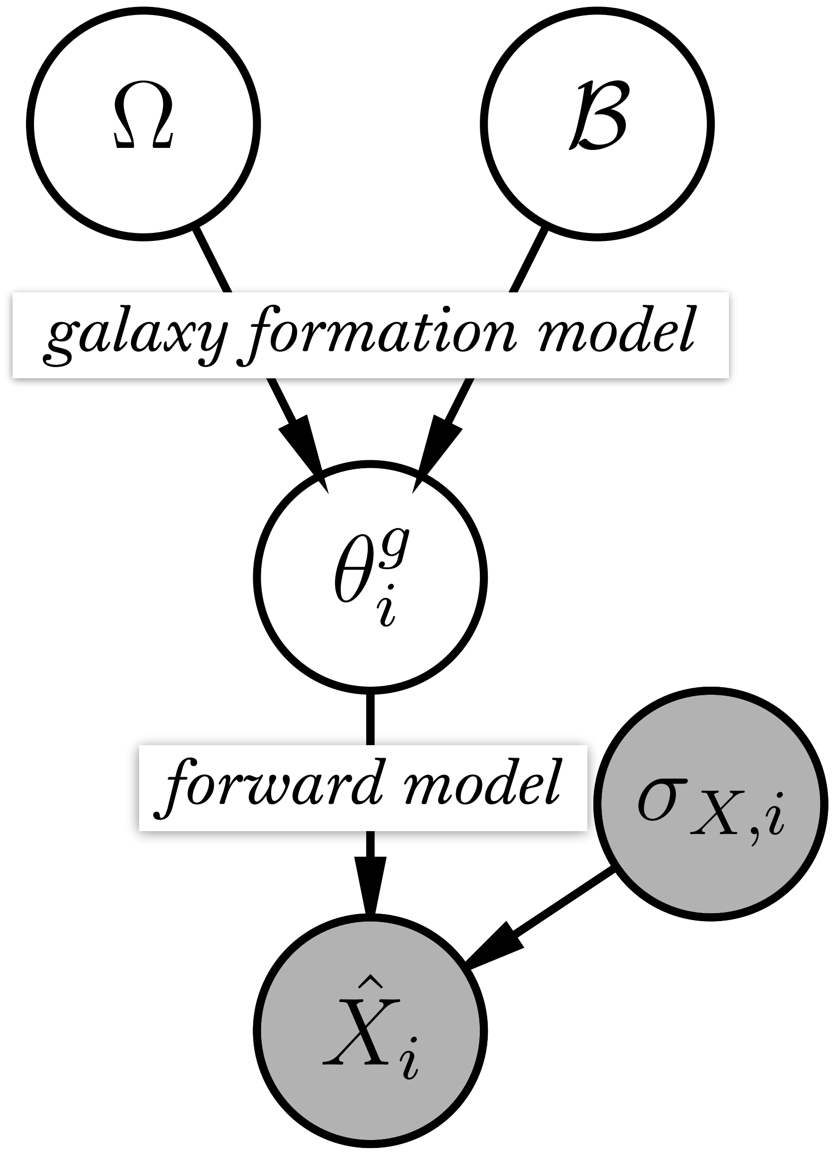
$$p(\theta_i^g \mid \Omega, \mathcal{B})$$

$$p(X_i \mid \theta_i^g)$$



$$p(\theta_i^g \mid \Omega, \mathcal{B})$$

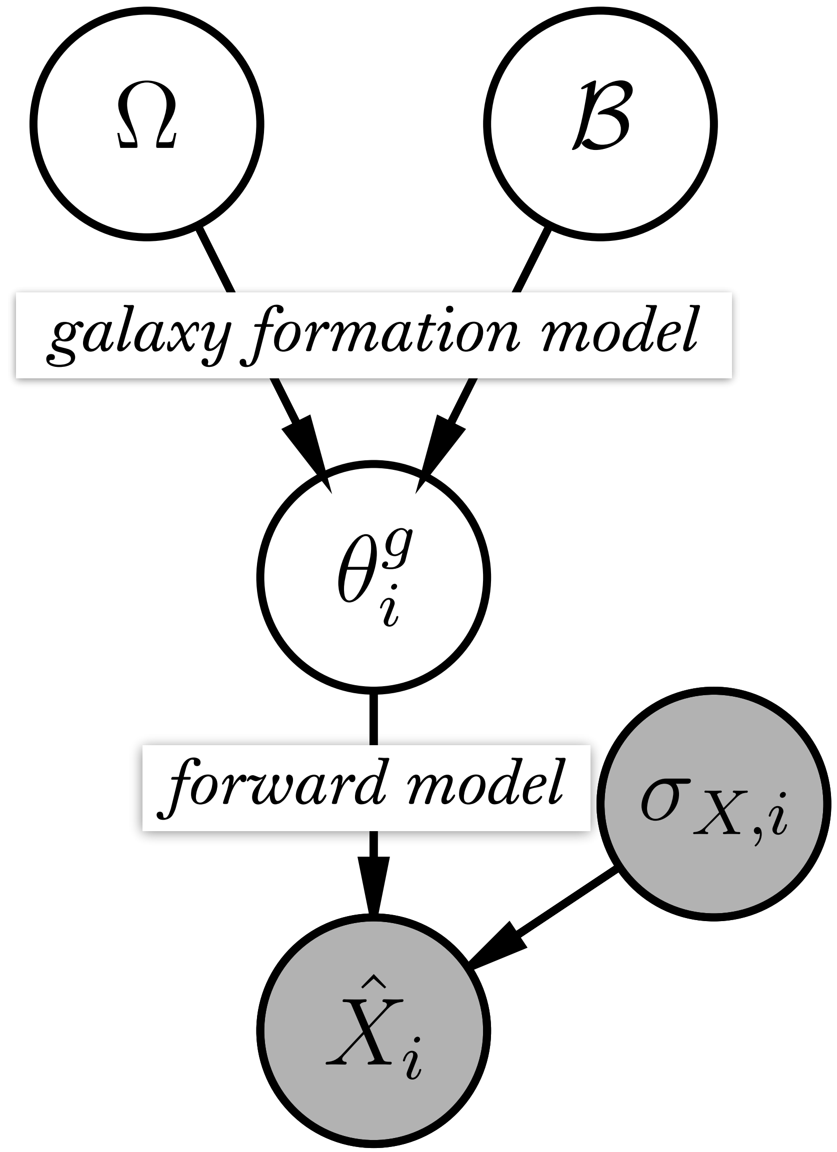
$$p(X_i \mid \theta_i^g)$$



$$p(\Omega, \mathcal{B} \mid X_i)$$

$$p(\theta_i^g \,|\, \Omega, \mathcal{B})$$

$$p(X_i \,|\, \theta_i^g)$$



$$p(\Omega, \mathcal{B} \,|\, X_i) = \int p(\Omega, \mathcal{B} \,|\, \theta_i^g)\; p(\theta_i^g \,|\, X_i)\; \mathrm{d}\theta_i^g$$

$$p(\Omega, \mathcal{B} \mid X_i) = \int p(\Omega, \mathcal{B} \mid \theta_i^g)\, p(\theta_i^g \mid X_i)\, \mathrm{d}\theta_i^g$$

cosmology with one galaxy *SED modeling*

$$p(\Omega, \mathcal{B} \mid X_i) = \int p(\Omega, \mathcal{B} \mid \theta_i^g)\ p(\theta_i^g \mid X_i)\ \mathrm{d}\theta_i^g$$

with **CAMELS** and ***neural density estimation***, we can directly estimate flexible model q with hyperparameters ϕ

$$\approx q_\phi(\Omega, \mathcal{B} \mid X_i)$$

$$p(\Omega, \mathcal{B} \mid X_i) = \int p(\Omega, \mathcal{B} \mid \theta_i^g)\ p(\theta_i^g \mid X_i)\ \mathrm{d}\theta_i^g$$

with **CAMELS** and ***neural density estimation***, we can directly estimate flexible model q with hyperparameters ϕ

$$\approx q_\phi(\Omega, \mathcal{B} \mid X_i)$$

$$\min_\phi D_{\mathrm{KL}}(\, p \ \| \ q_\phi) = \min_\phi \int p \ \log \left(\frac{p}{q_\phi} \right)$$

$$p(\Omega, \mathcal{B} \mid X_i) = \int p(\Omega, \mathcal{B} \mid \theta_i^g) \; p(\theta_i^g \mid X_i) \; \mathrm{d}\theta_i^g$$

with **CAMELS** and ***neural density estimation***, we can directly estimate flexible model q with hyperparameters ϕ

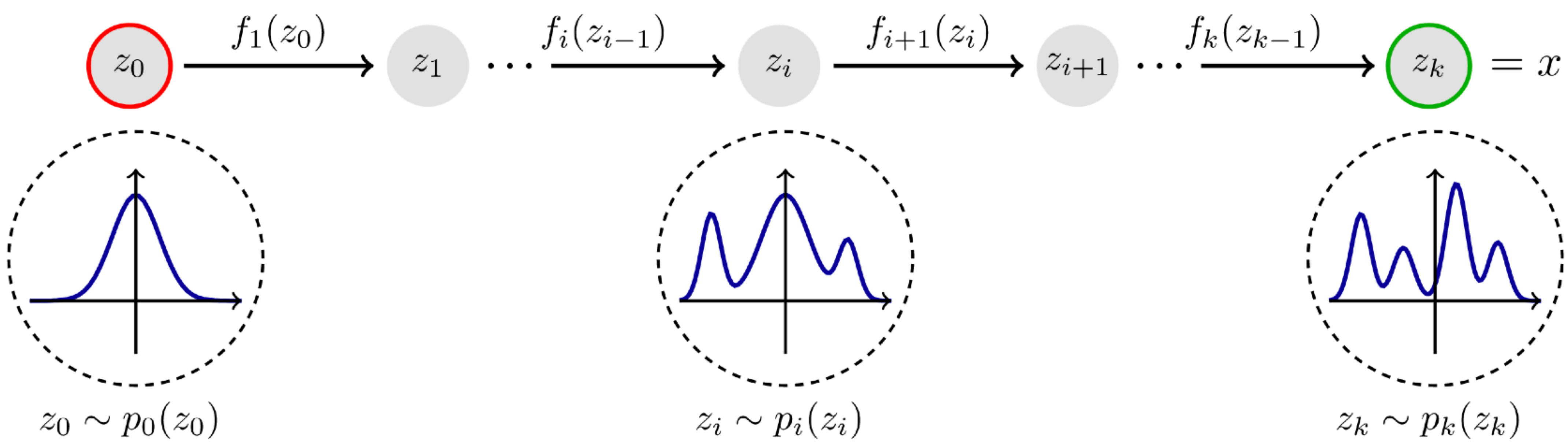
$$\approx q_\phi(\Omega, \mathcal{B} \mid X_i)$$

$$\min_\phi D_{\mathrm{KL}}(\, p \;\|\; q_\phi\,) = \min_\phi \int p \, \log\left(\frac{p}{q_\phi}\right) \approx \max_\phi \sum_i \log q_\phi(\Omega_i, \mathcal{B}_i \mid X_i)$$

galaxies in CAMELS forward model

$\{(\Omega', \mathcal{B}', X')\} \sim p(\Omega, \mathcal{B}, X)$

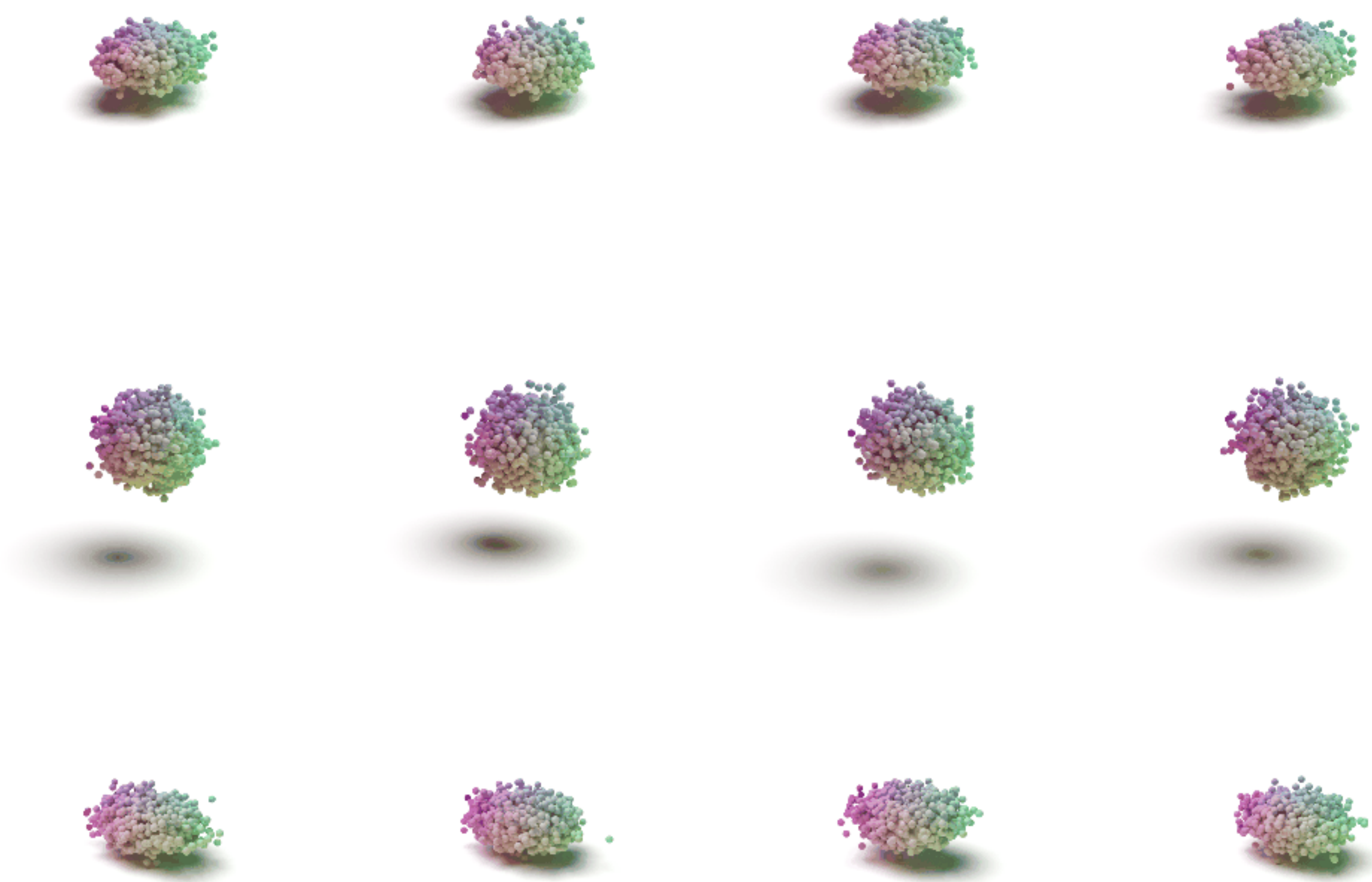
q_ϕ — **normalizing flows** are easy to evaluate and flexibly expressive



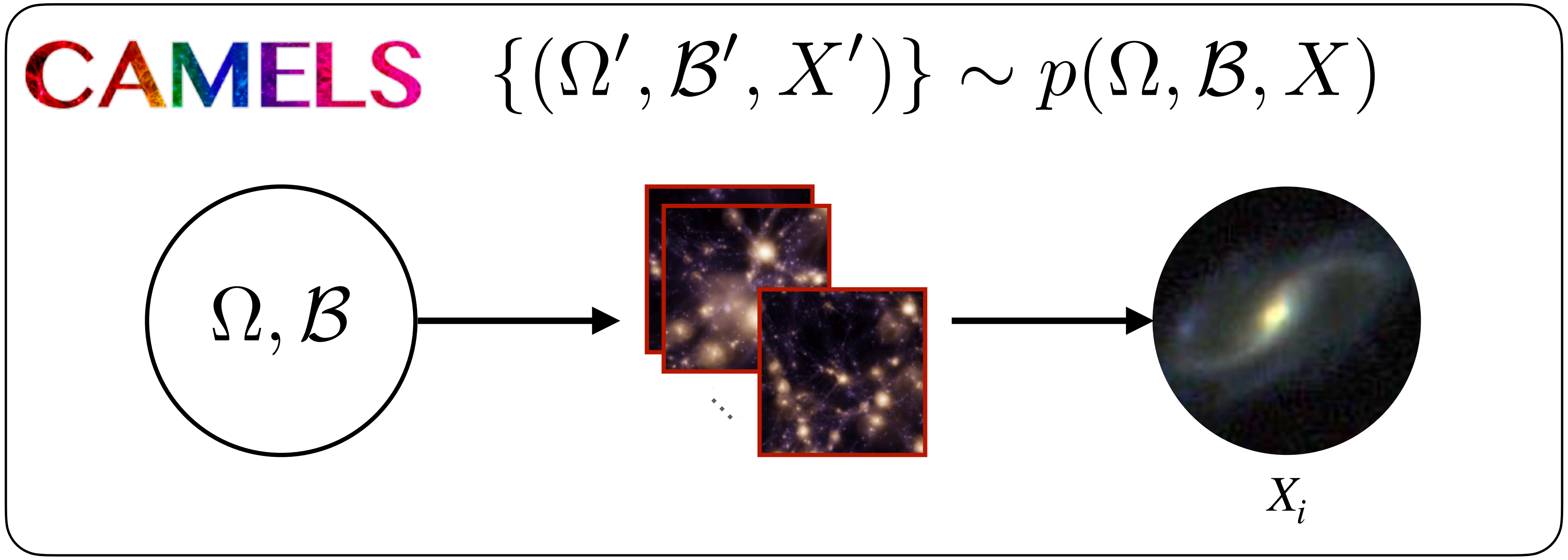
$z_i = f_i(z_{i-1})$ are *invertible* and *differentiable* transformations

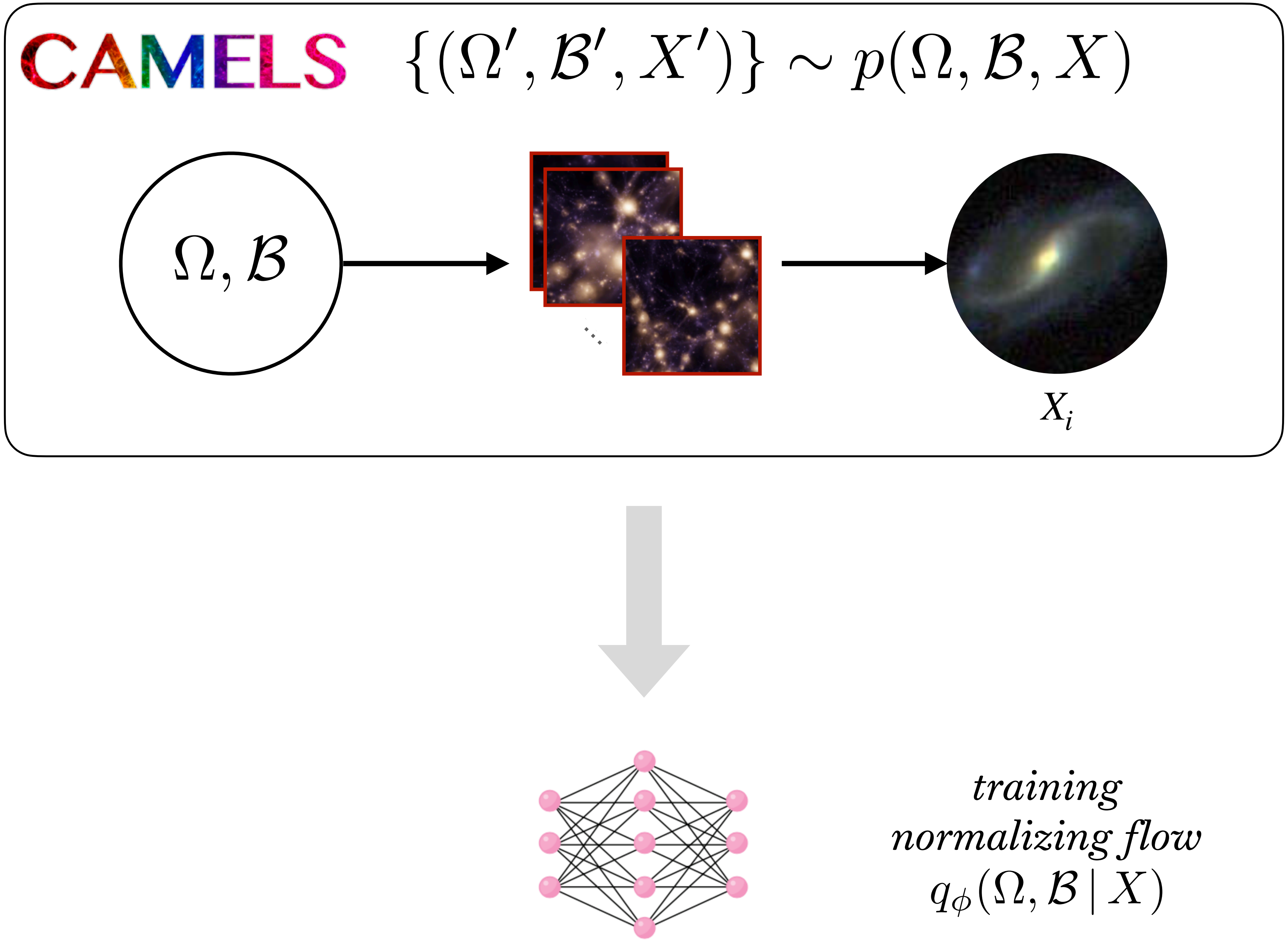
$$p(z_i) = p(z_{i-1}) \left| \det \left(\frac{\partial f_i^{-1}}{\partial z_i} \right) \right|$$

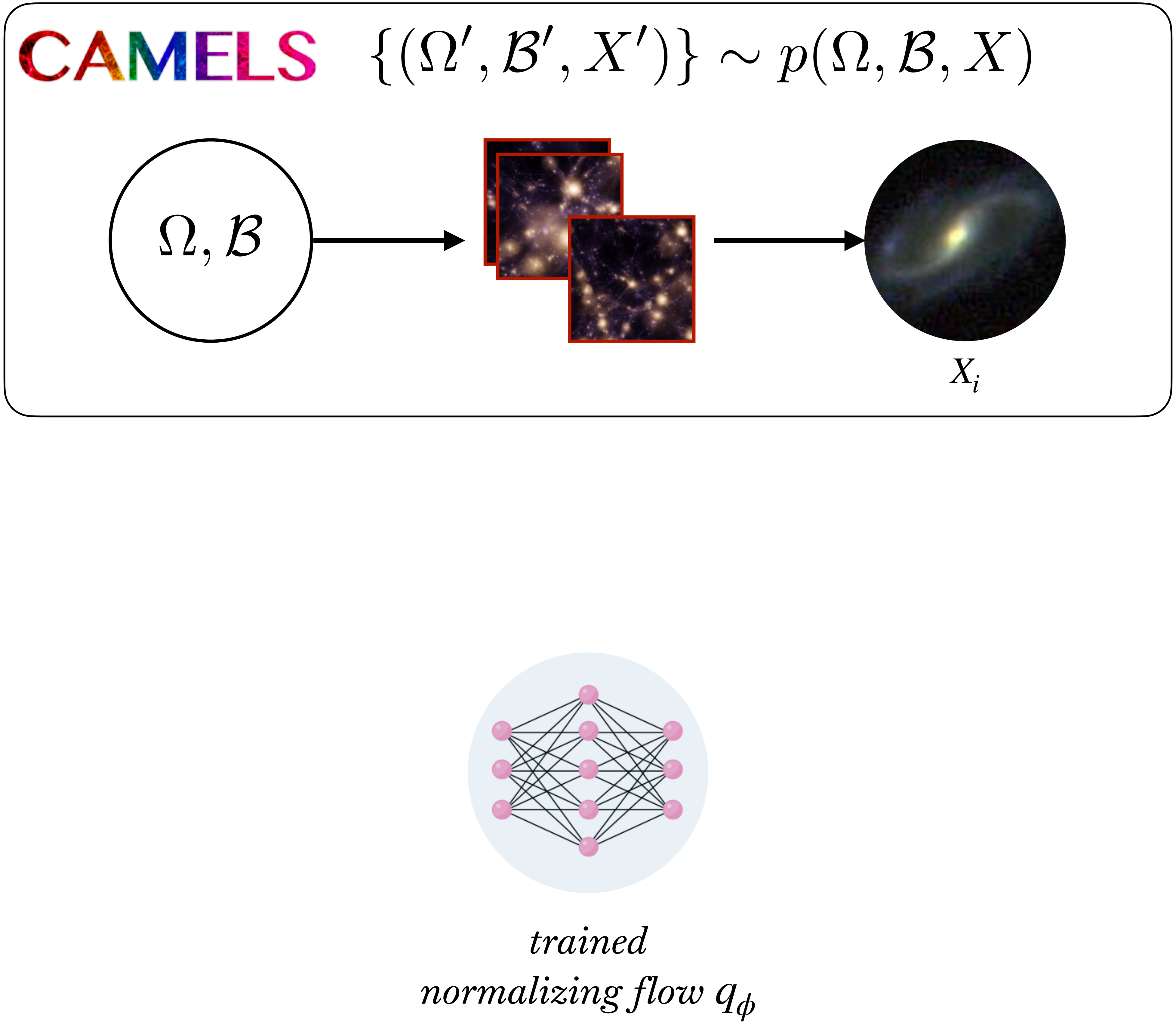
q_ϕ — **normalizing flows** are easy to evaluate and flexibly expressive



e.g. PointFlow (Yang et al. 2019)

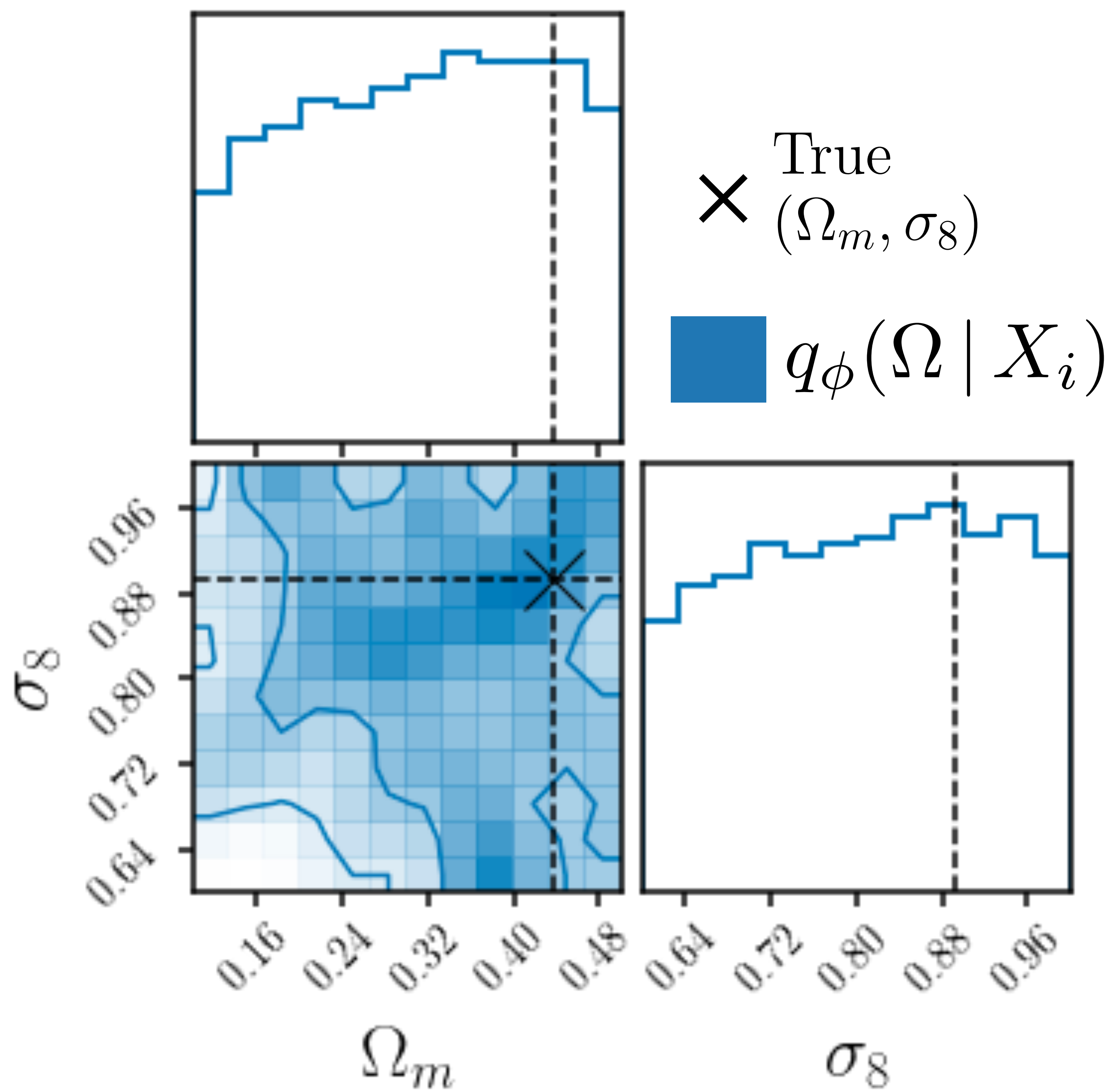

CAMELS
$\{(\Omega', \mathcal{B}', X')\} \sim p(\Omega, \mathcal{B}, X)$
$\Omega, \mathcal{B}$
X_i


CAMELS
$\{(\Omega', \mathcal{B}', X')\} \sim p(\Omega, \mathcal{B}, X)$
$\Omega, \mathcal{B}$
X_i
training
normalizing flow
$q_\phi(\Omega, \mathcal{B} \mid X)$

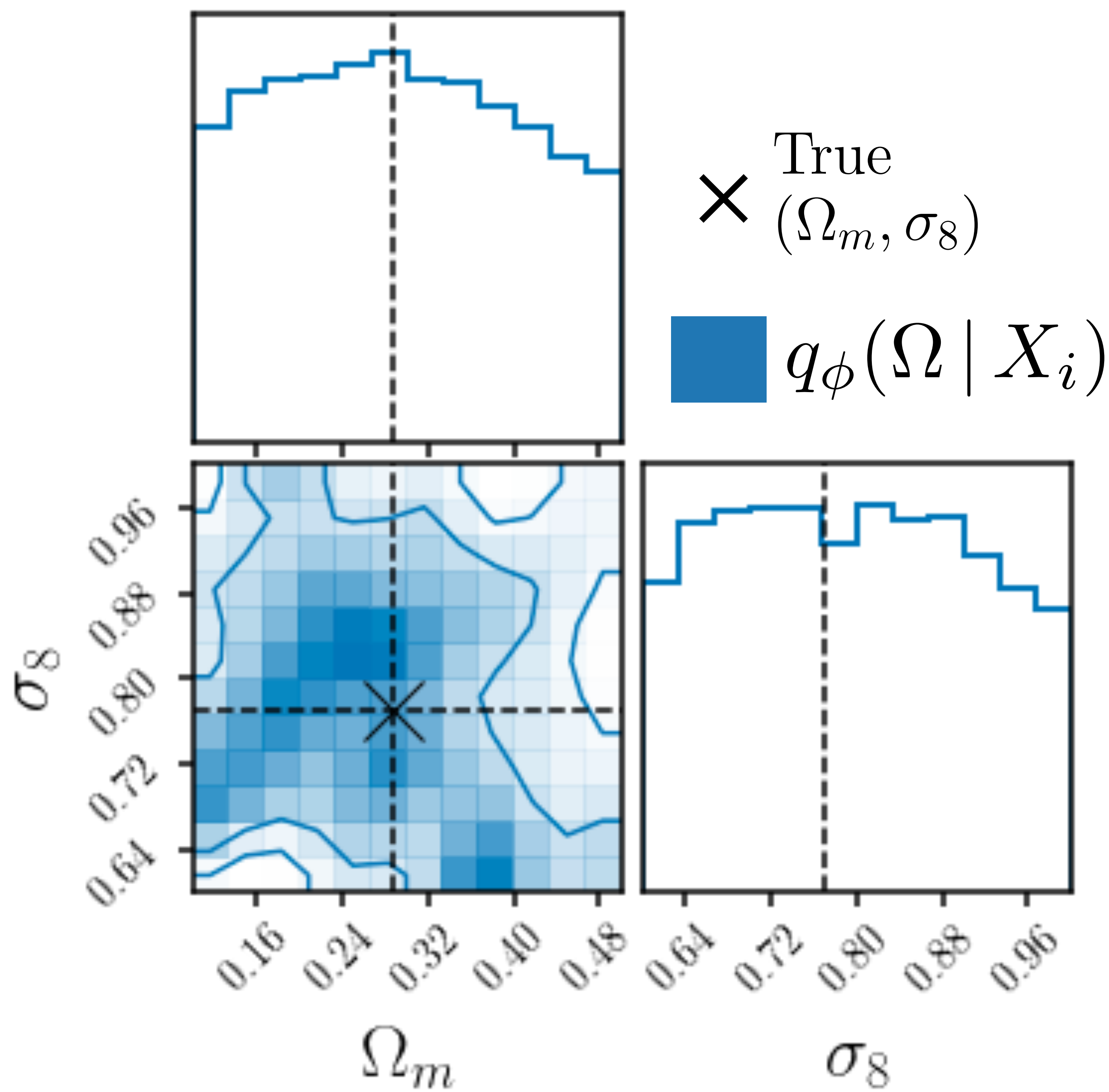


CAMELS $\{(\Omega', \mathcal{B}', X')\} \sim p(\Omega, \mathcal{B}, X)$
$\Omega, \mathcal{B}$
X_i
trained
normalizing flow q_ϕ

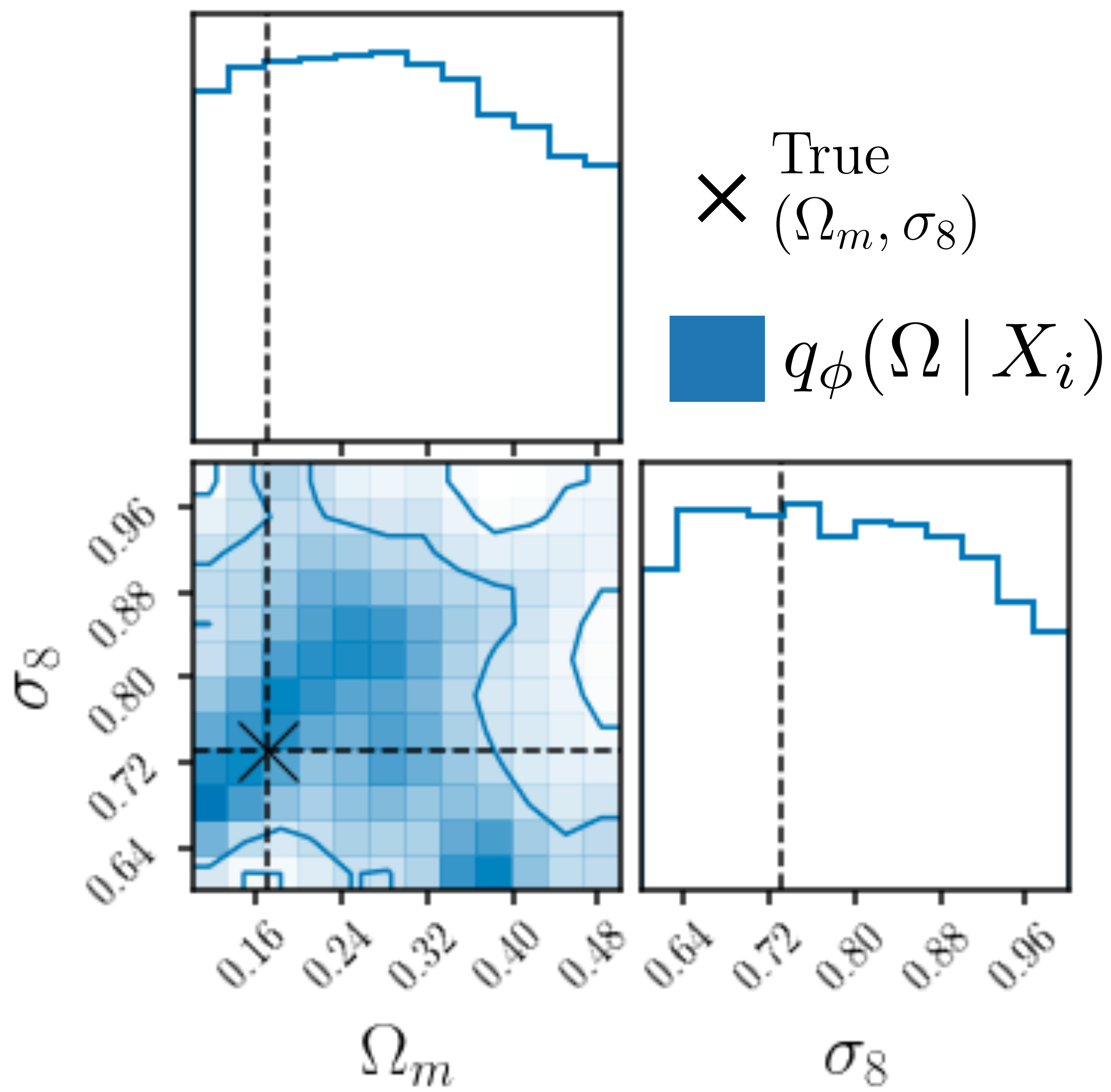
$p(\Omega \mid X_i) \approx q_\phi(\Omega \mid X_i)$ validation



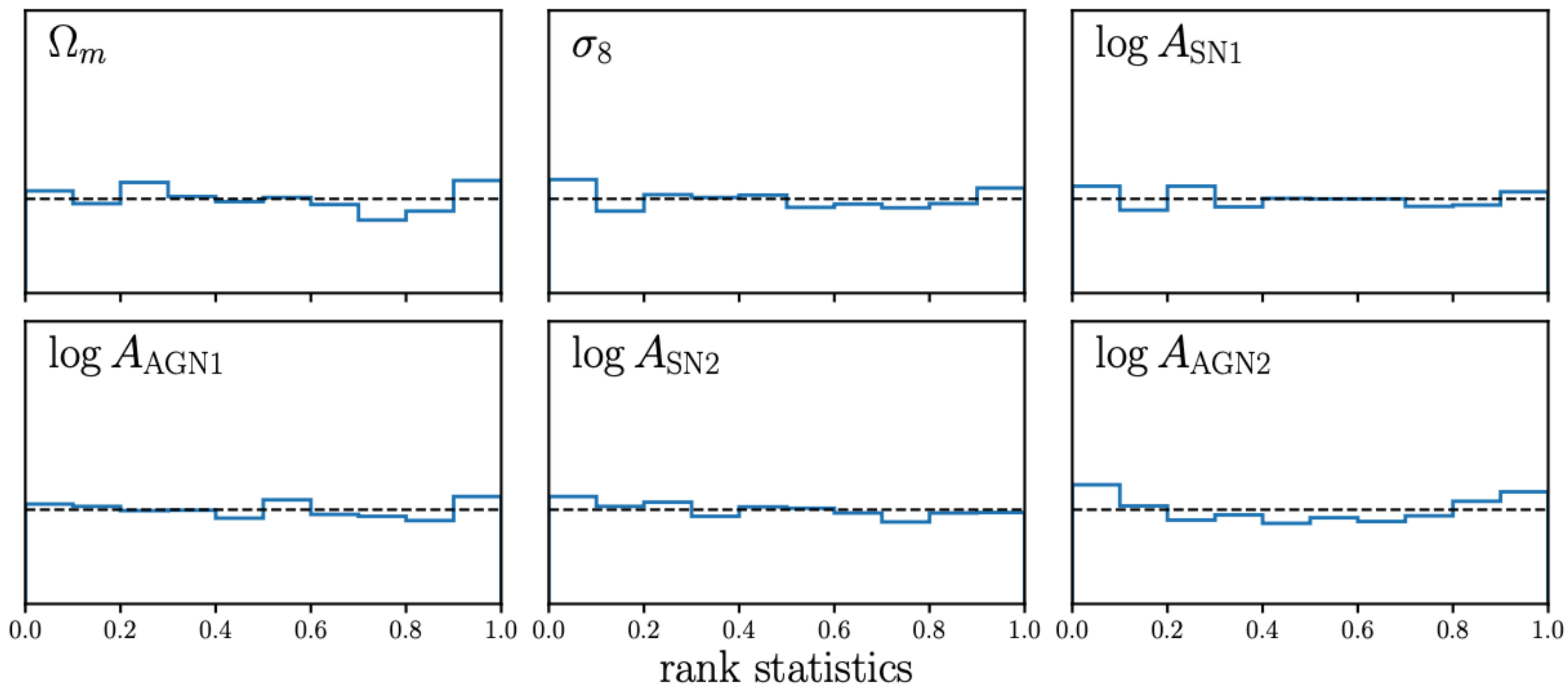
$p(\Omega \mid X_i) \approx q_\phi(\Omega \mid X_i)_{\text{validation}}$



$p(\Omega \mid X_i) \approx q_\phi(\Omega \mid X_i)$ validation

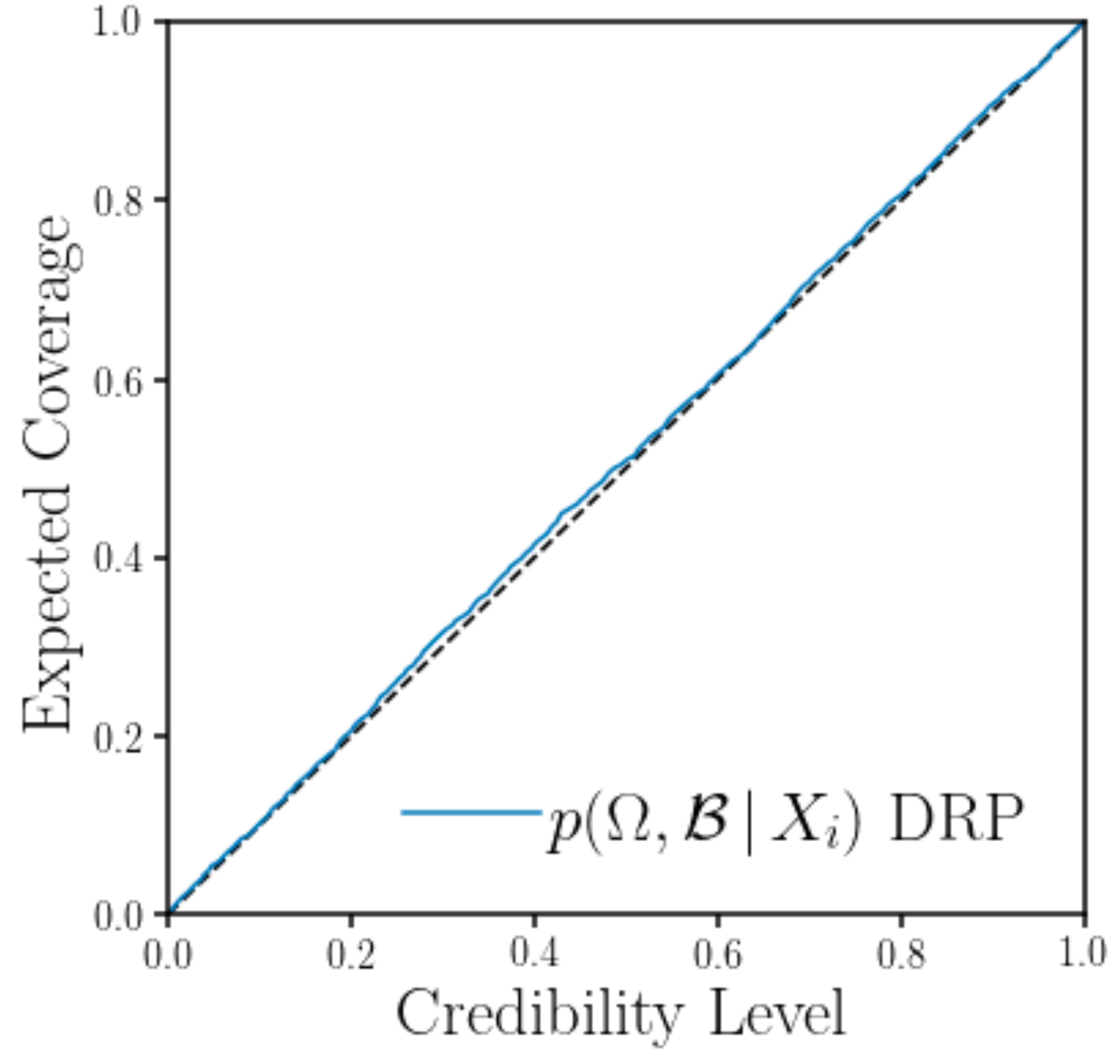


$p(\Omega \mid X_i) \approx q_\phi(\Omega \mid X_i)$ validation: **coverage tests**

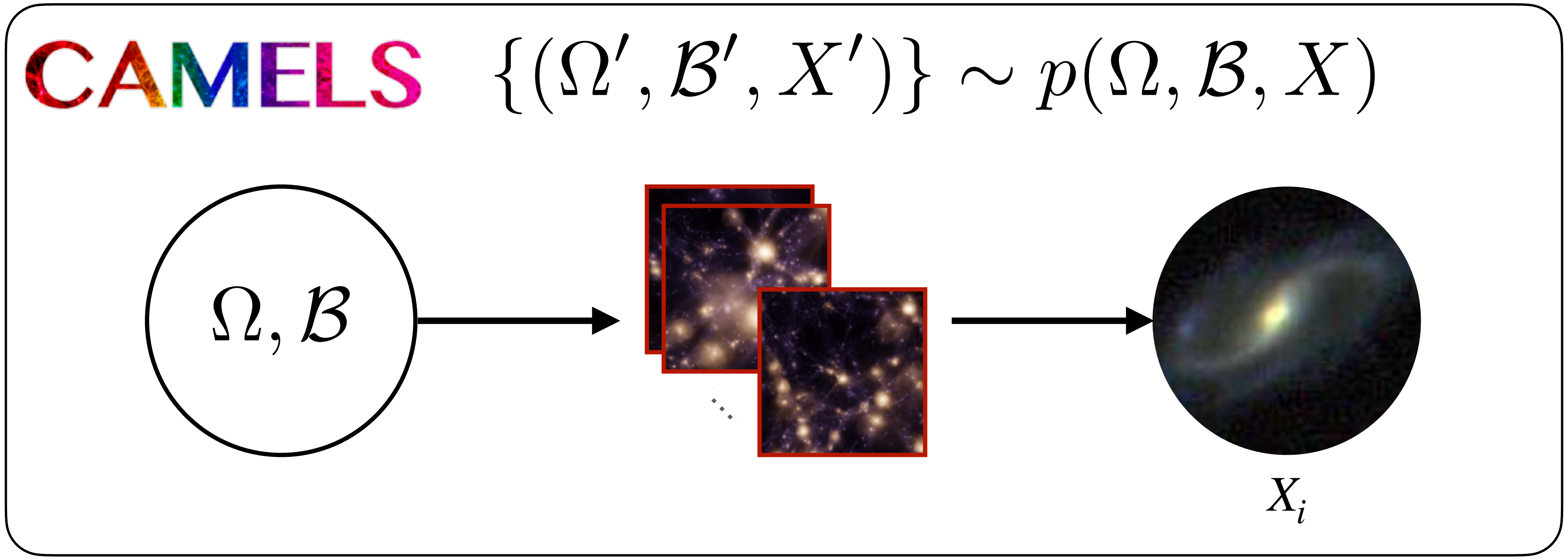


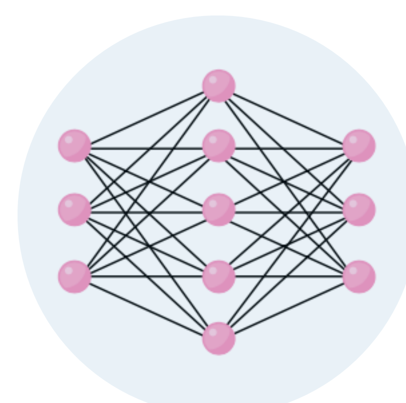
Talts et al. (2020)

$p(\Omega \mid X_i) \approx q_\phi(\Omega \mid X_i)$ validation: **coverage tests**

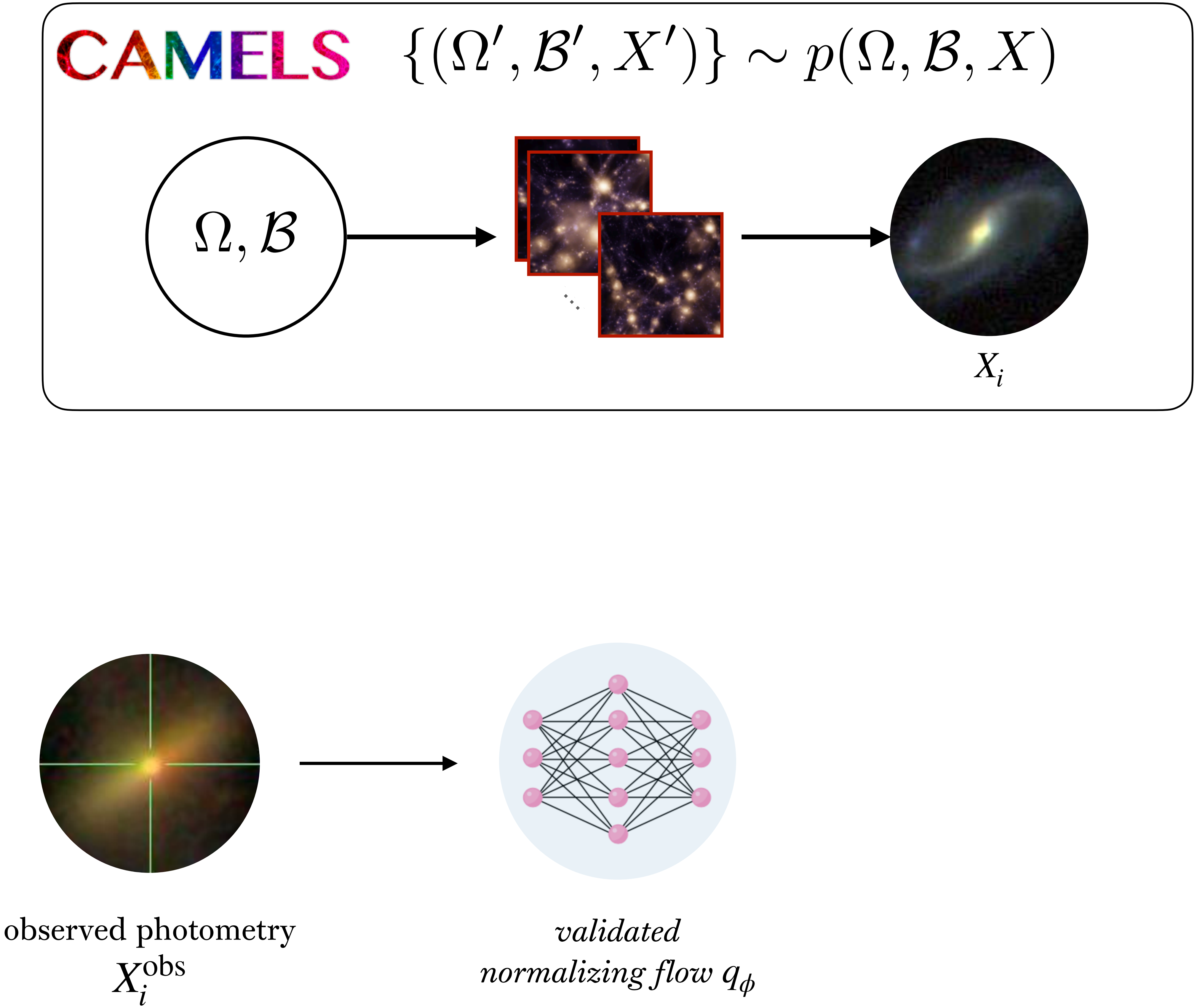


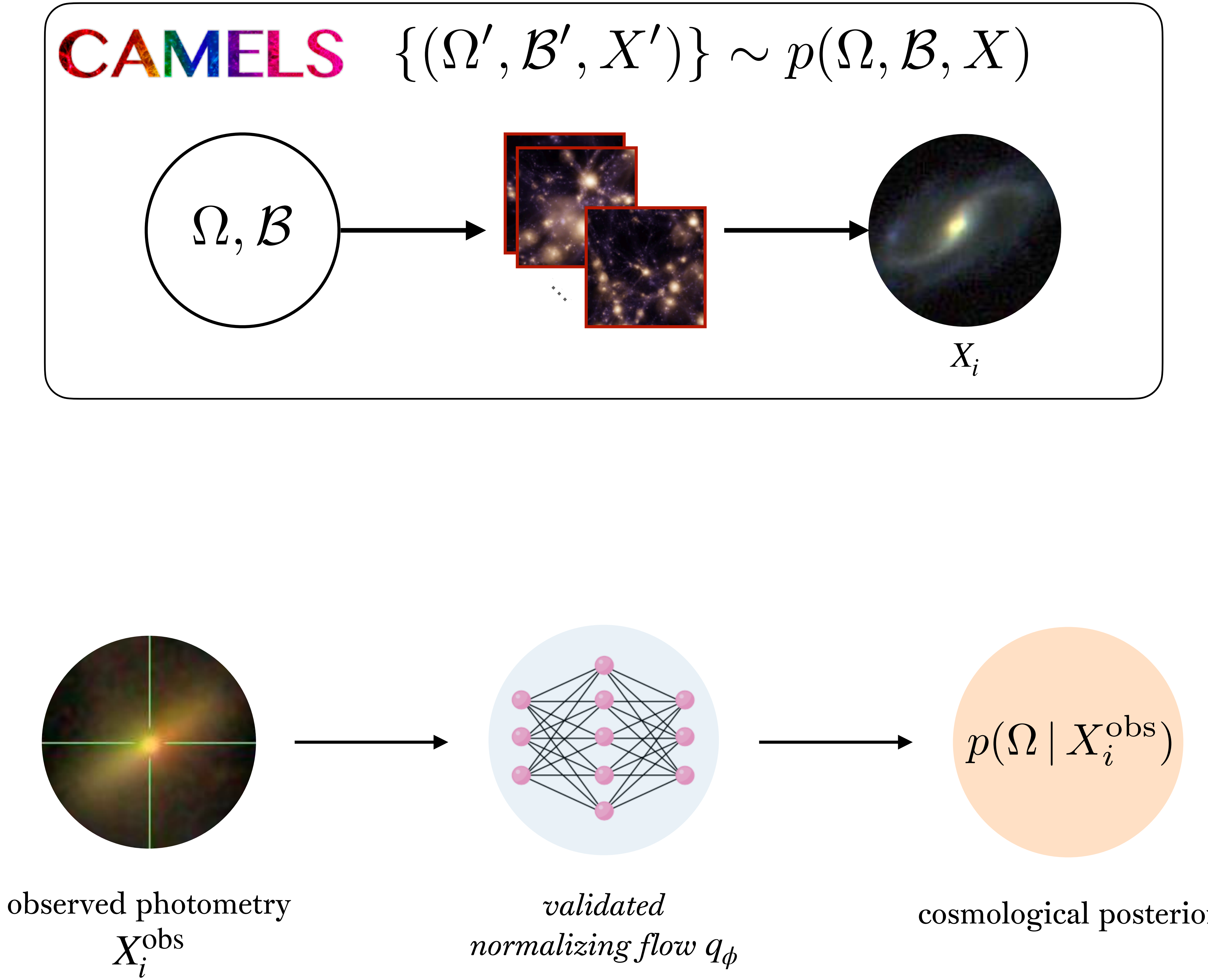
Lemos et al. (2023)



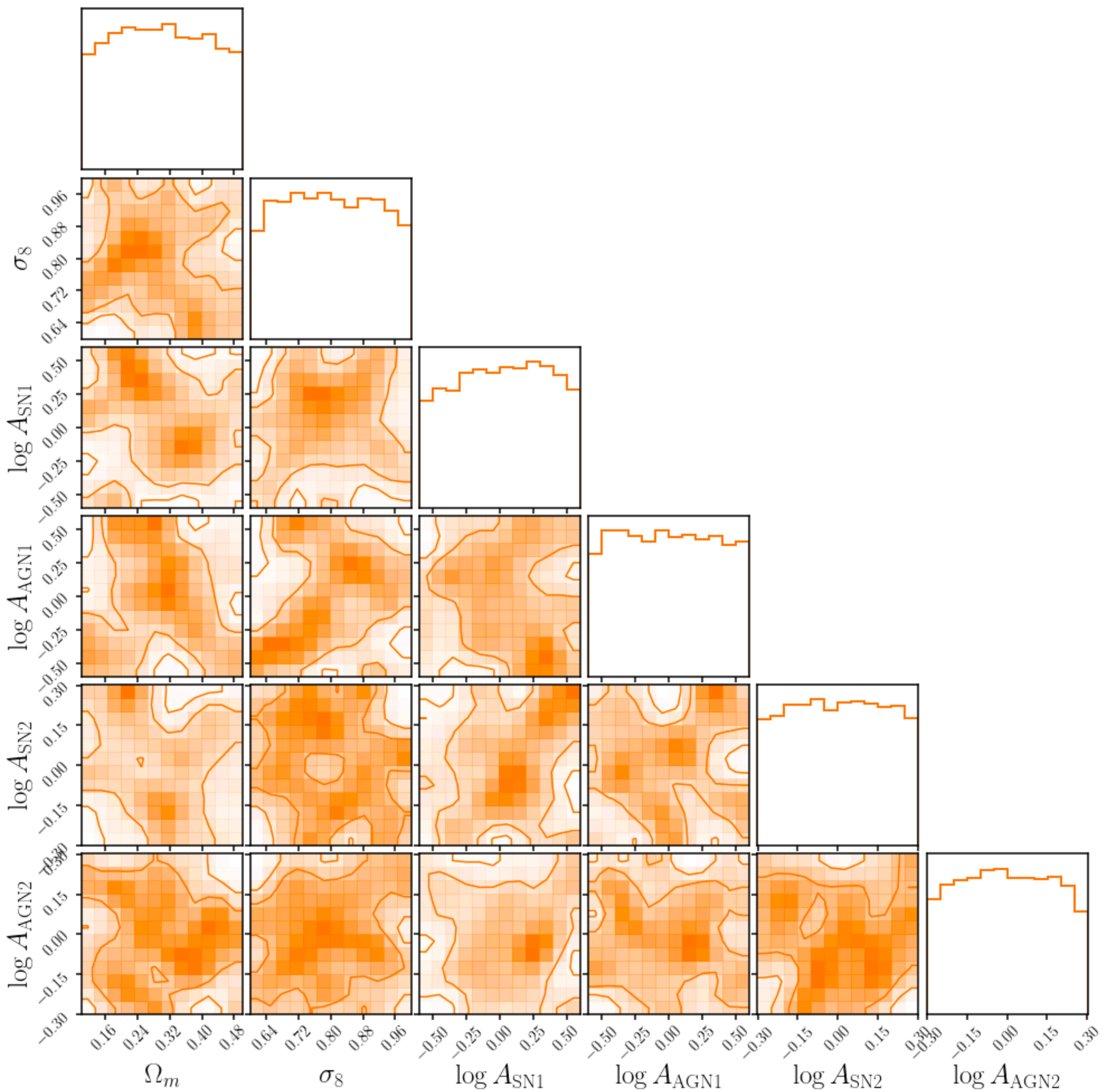


validated
normalizing flow q_ϕ

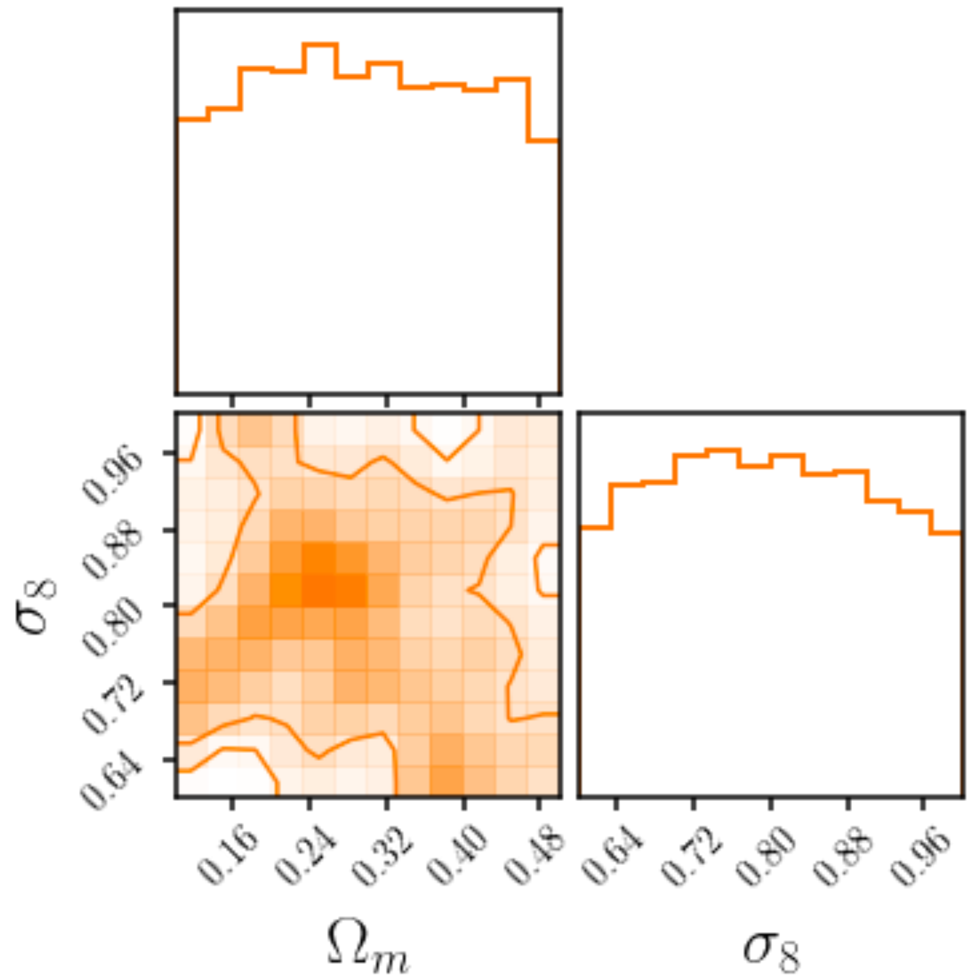

CAMELS
$\{(\Omega', \mathcal{B}', X')\} \sim p(\Omega, \mathcal{B}, X)$
$\Omega, \mathcal{B}$
X_i
observed photometry
X_i^{obs}
validated
normalizing flow q_ϕ


CAMELS
$\{(\Omega', \mathcal{B}', X')\} \sim p(\Omega, \mathcal{B}, X)$
$\Omega, \mathcal{B}$
X_i
observed photometry
X_i^{obs}
validated
normalizing flow q_ϕ
$p(\Omega \mid X_i^{\text{obs}})$
cosmological posterior

$p(\Omega, \mathcal{B} \mid X_i) \approx q_\phi(\Omega, \mathcal{B} \mid X_i)$ from *griz* optical photometry of a **single SDSS galaxy**

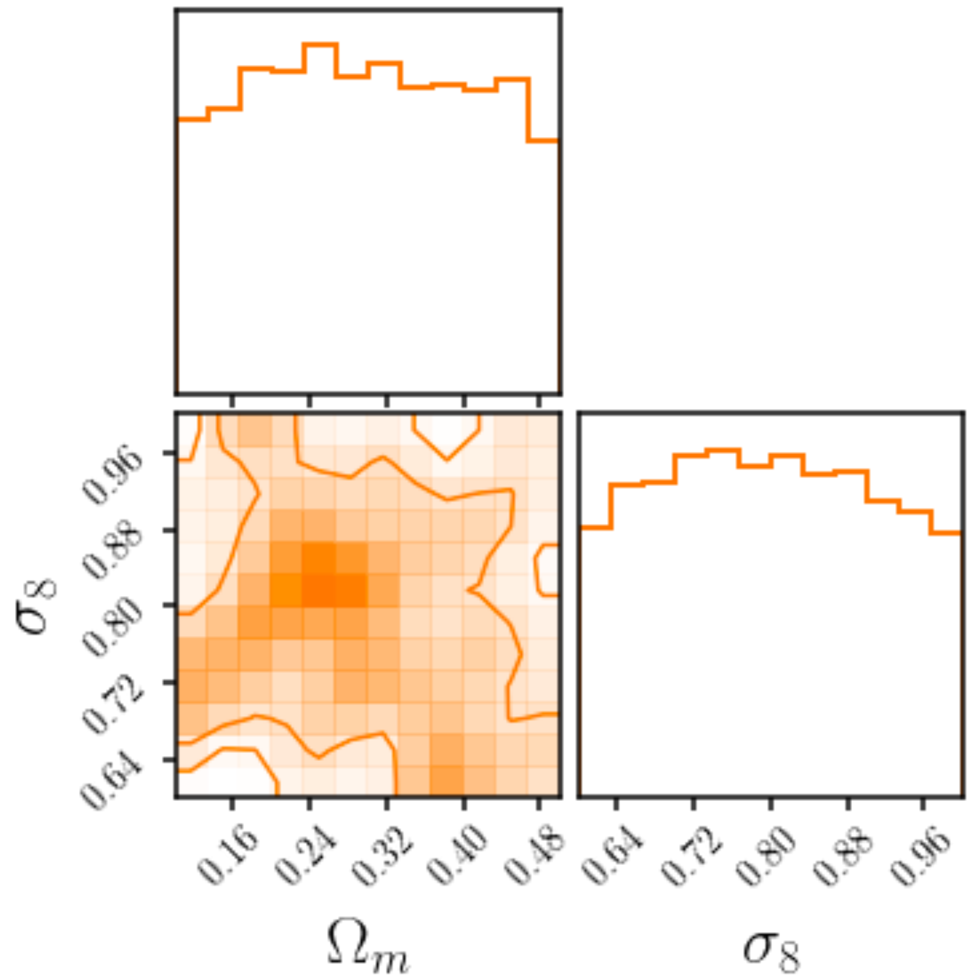


$p(\Omega, \mathcal{B} \mid X_i) \approx q_\phi(\Omega, \mathcal{B} \mid X_i)$ from *griz* optical photometry of a **single SDSS galaxy**



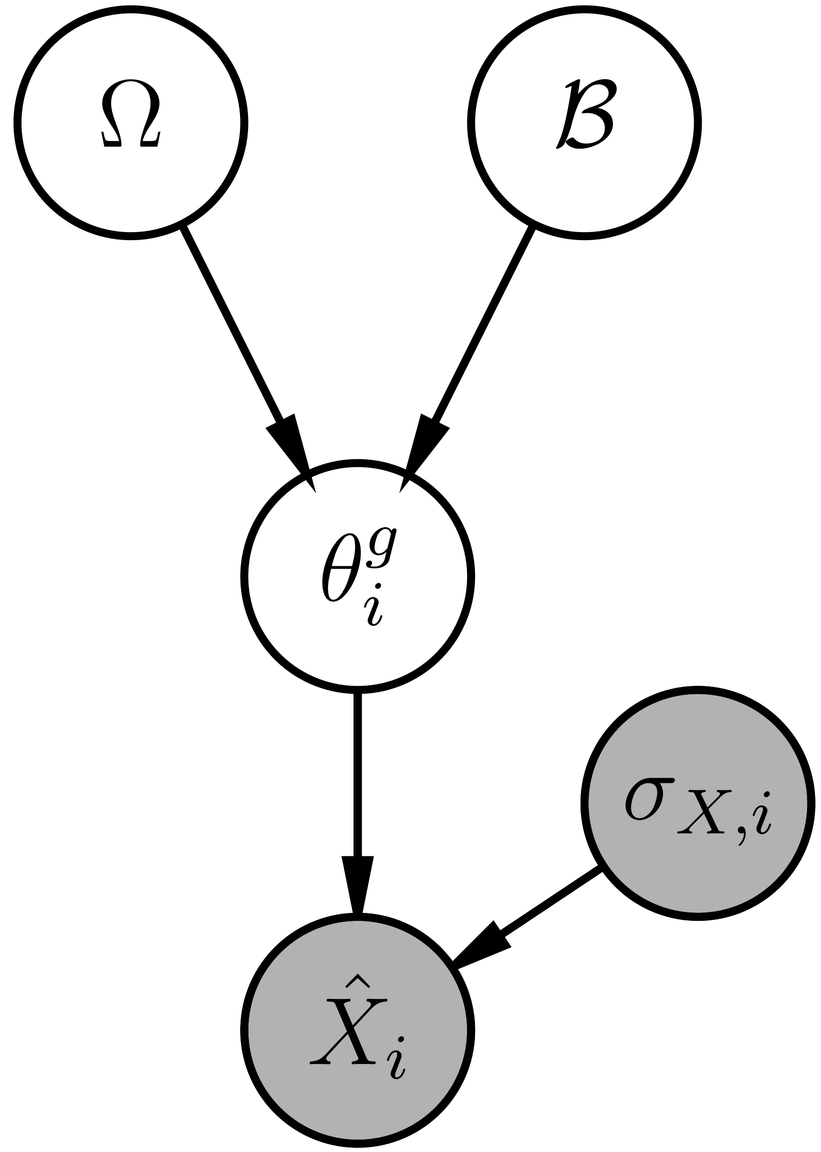
photometry of a single galaxy contains *limited* cosmological information

$p(\Omega, \mathcal{B} \mid X_i) \approx q_\phi(\Omega, \mathcal{B} \mid X_i)$ from *griz* optical photometry of a **single SDSS galaxy**

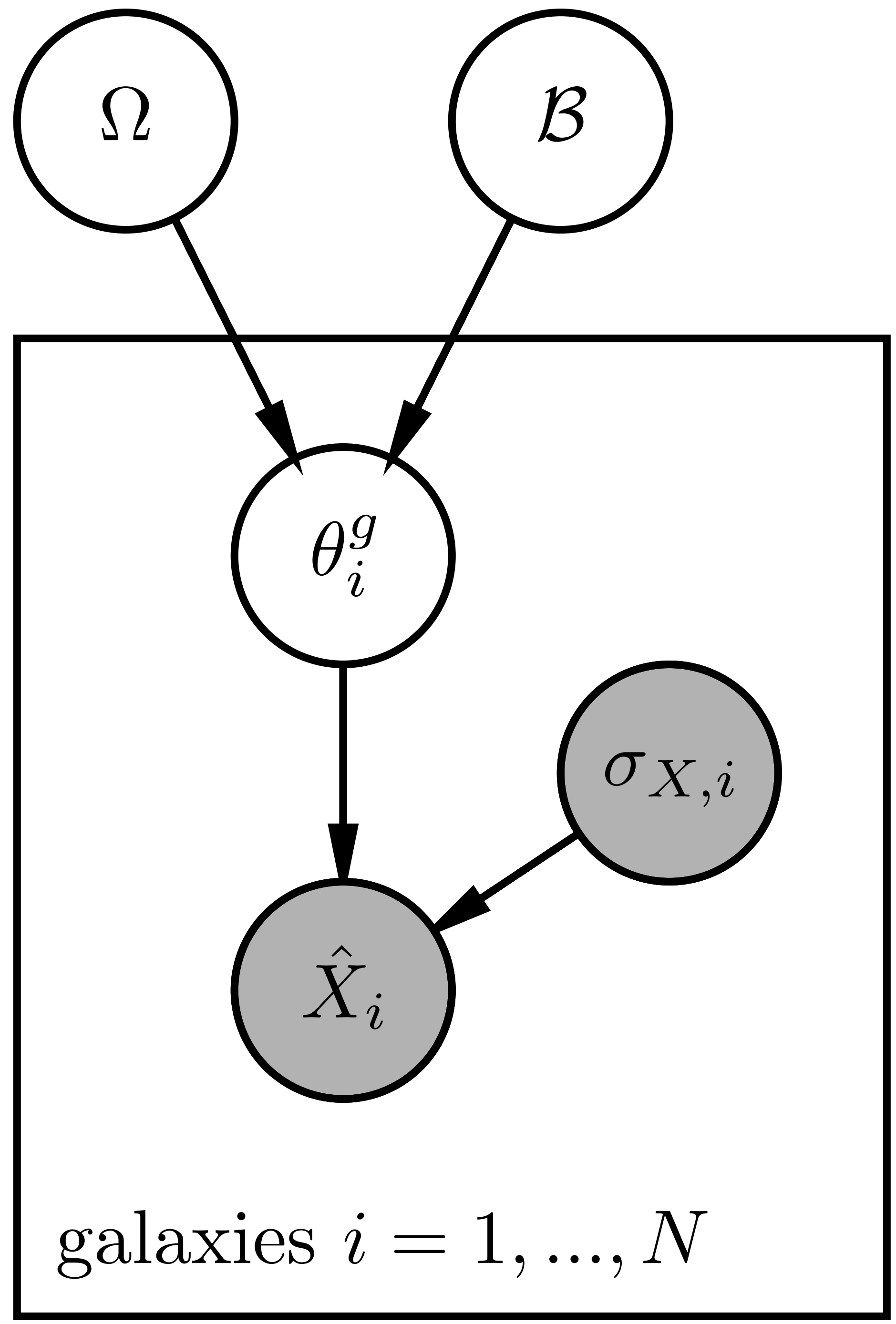


photometry of a single galaxy contains *limited* cosmological information

but not zero…

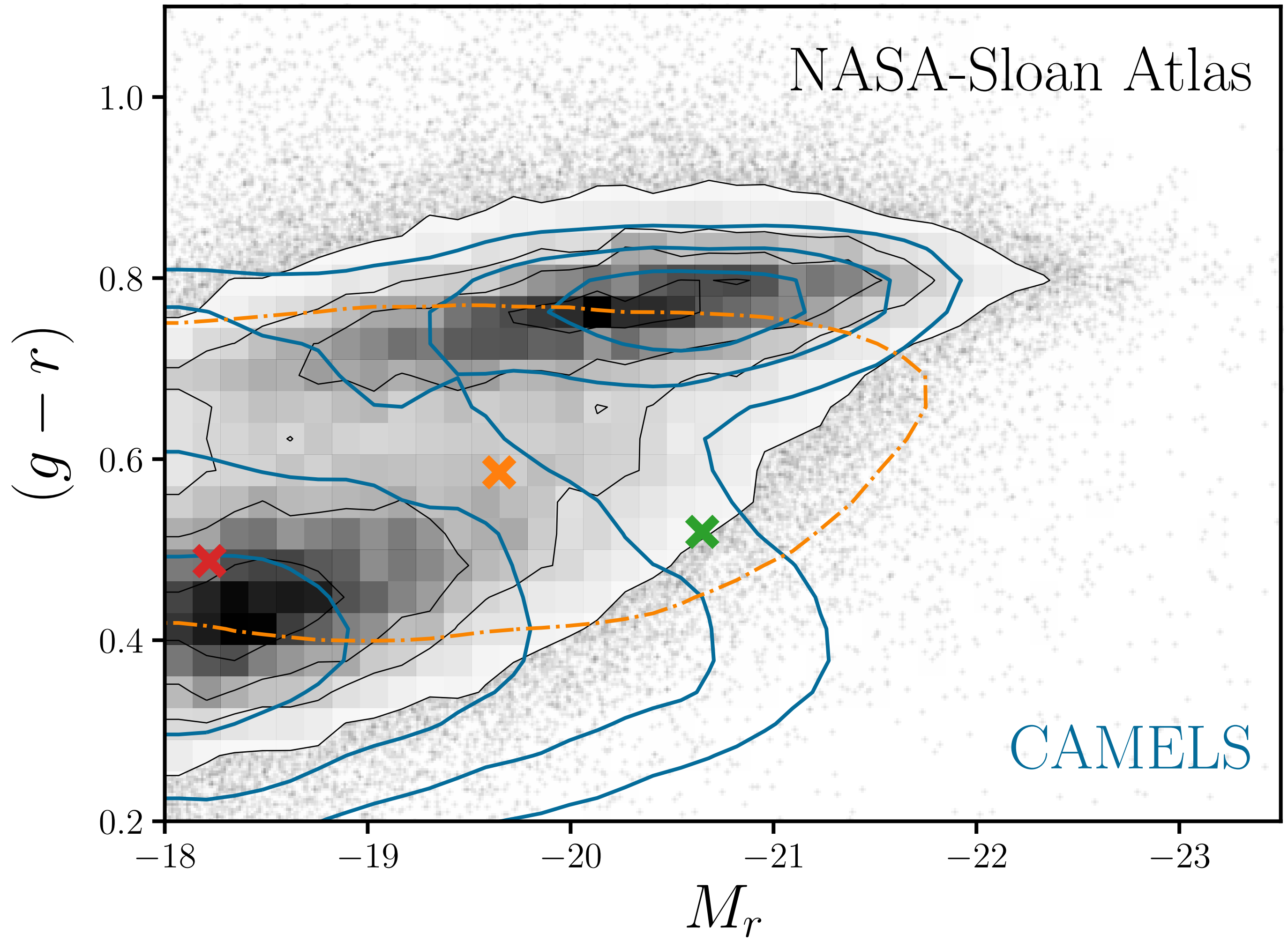


but we can combine the constraining power of *many* galaxies using **hierarchical population inference**



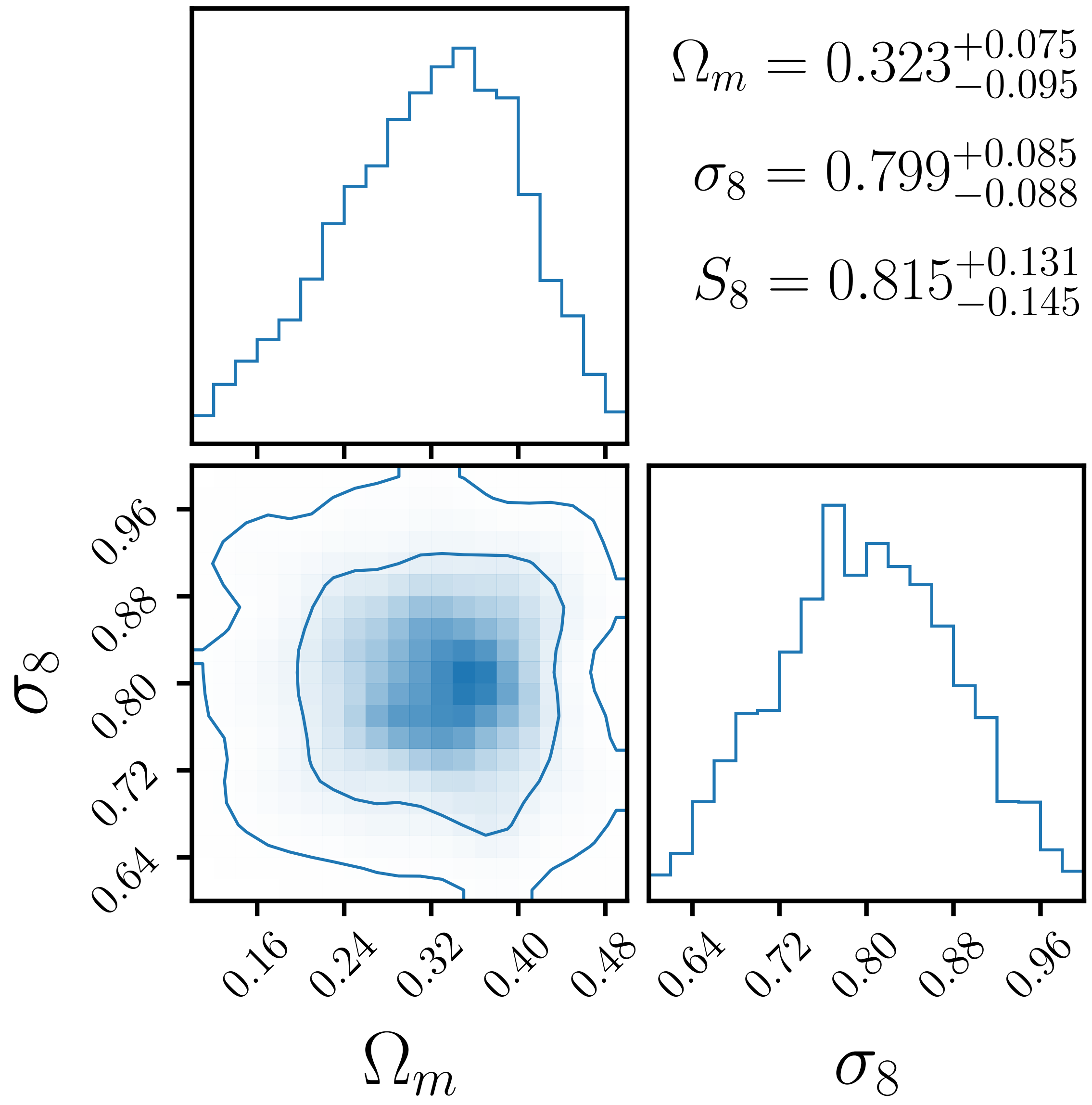
$$p(\Omega, \mathcal{B} \mid \{X_i\}) = p(\Omega, \mathcal{B})^{-(N-1)} \prod_{i=1}^{N} p(\Omega, \mathcal{B} \mid X_i)$$

N = *22,338 galaxies* from the NASA-Sloan Atlas



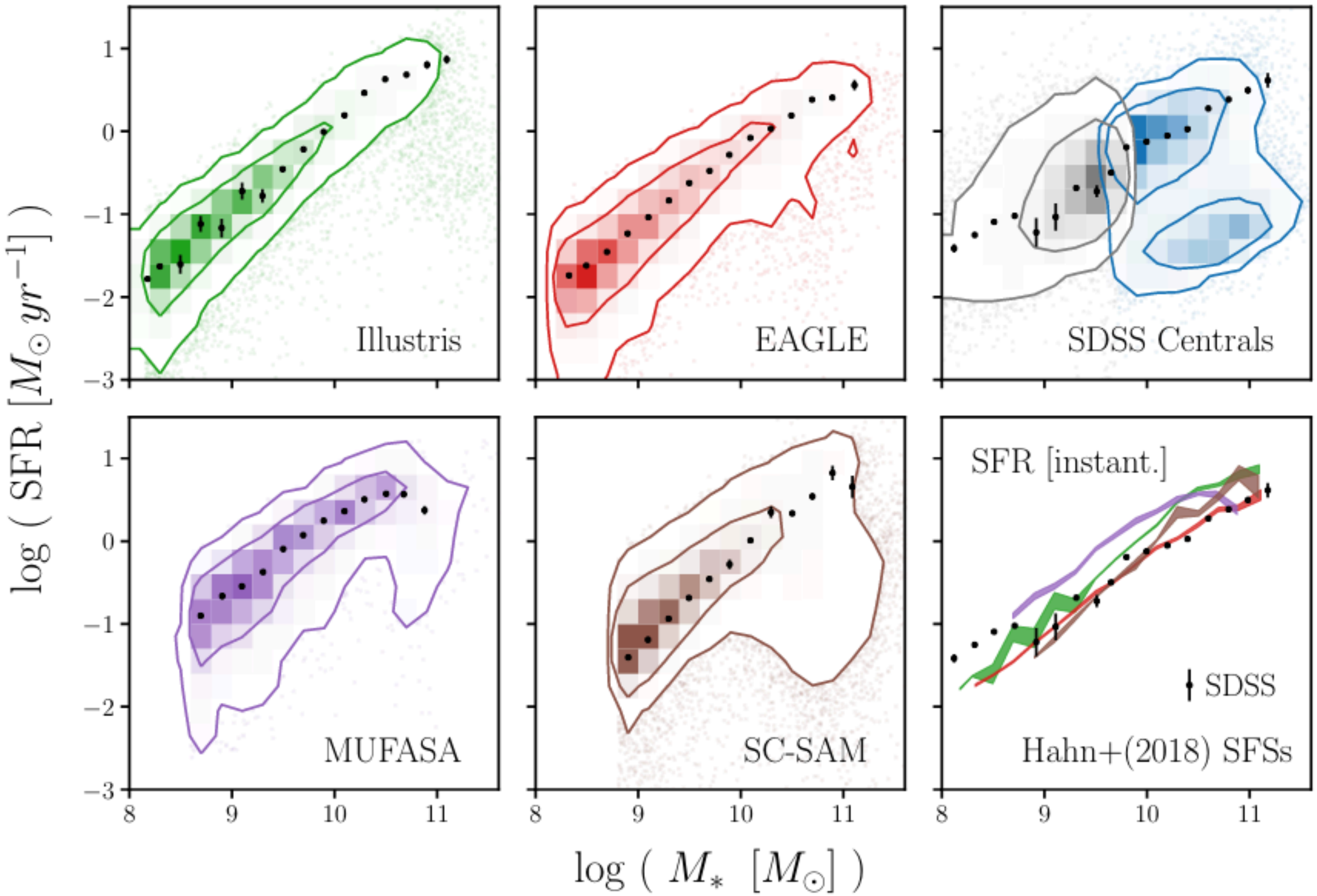
Hahn *et al. (2023h)*

cosmological constraints from *only* the observed photometry of *22,338* NASA-Sloan Atlas galaxies



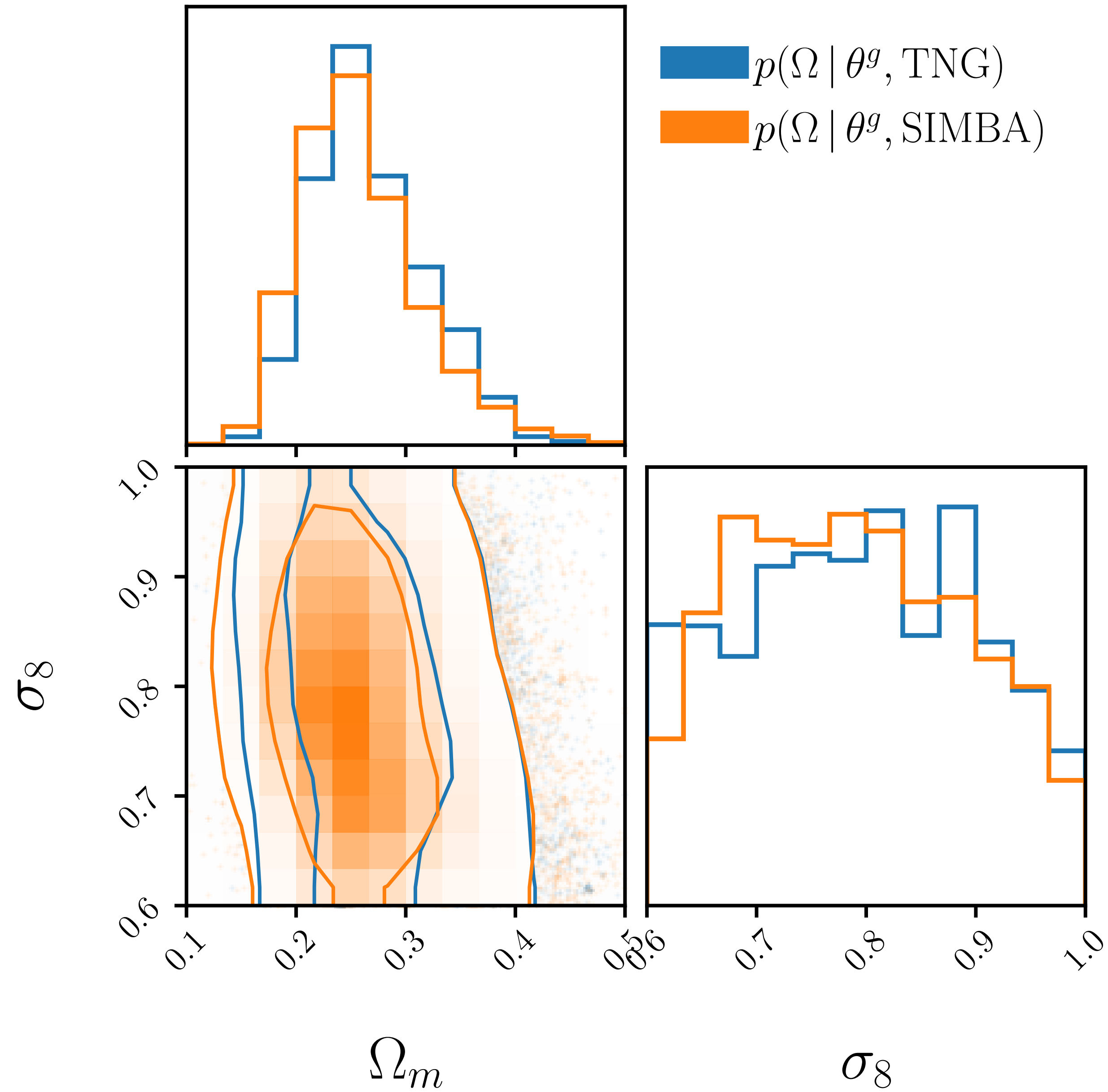
***Hahn** et al. (2023h)*

caveat: we assume a **galaxy formation model** (TNG) and an SED model



***Hahn** et al. (2019a)*

control regions: we can choose galaxy populations most *robust* to galaxy formation models — *e.g.* star-forming galaxy with $M_* \sim 10^{9.5} M_\odot$



***Hahn** et al. (2023h)*

caveat: we assume a galaxy formation model (TNG) and an **SED model**

$$f_\lambda = \int_{t'=0}^{t'=t} \mathrm{SFR}(t') \; f_{\mathrm{SSP}}(t',\, Z(t')\,)\; e^{-\tau_{\mathrm{dust}}(t')}\; \mathrm{d}t'$$

control regions: we can choose galaxy populations most *robust* to SED models

$$f_\lambda = \int_{t'=0}^{t'=t} \mathrm{SFR}(t')\ \underline{f_{\mathrm{SSP}}}(t',\ Z(t'))\ e^{-\tau_{\mathrm{dust}}(t')}\ \mathrm{d}t'$$

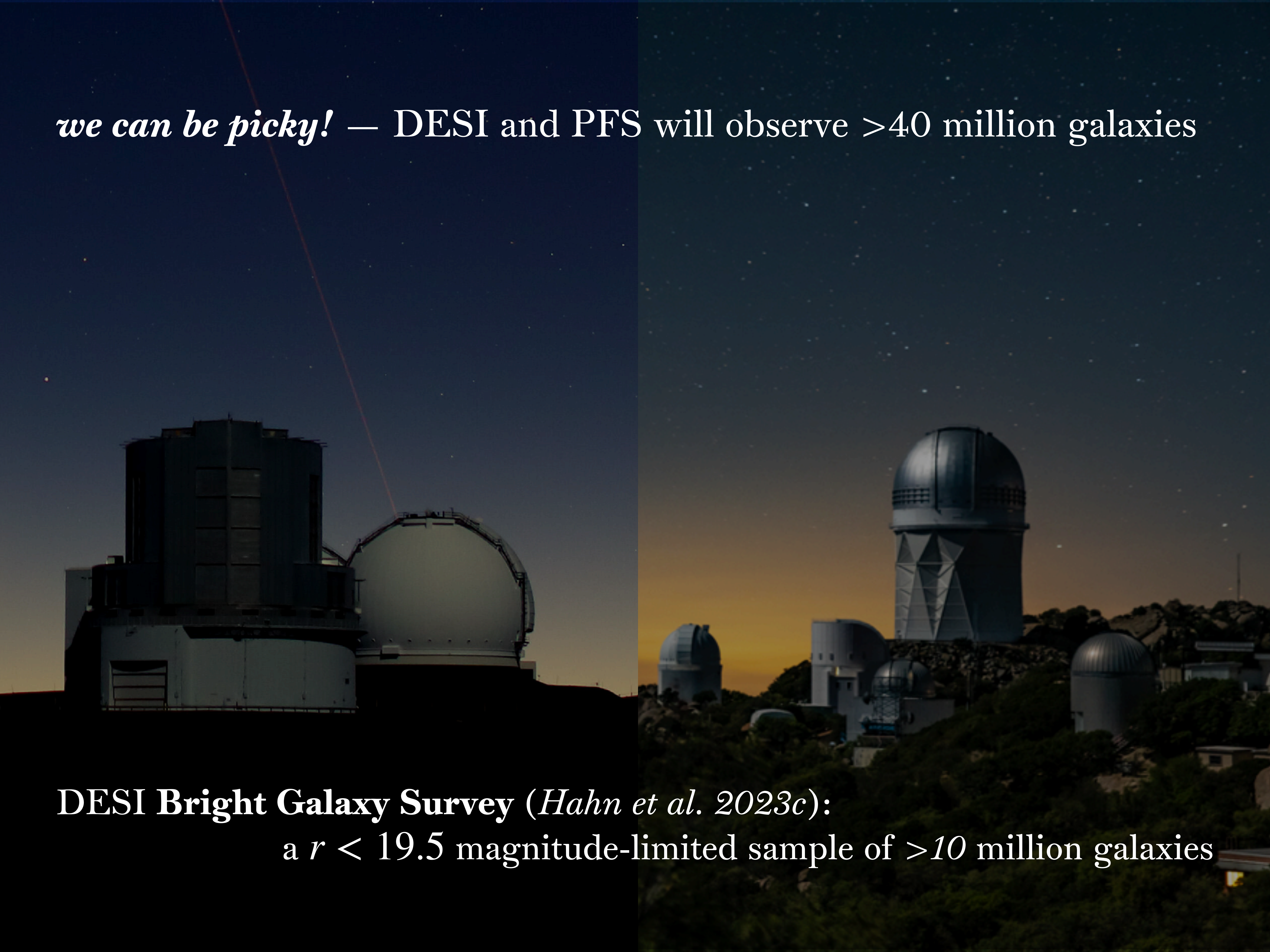
stellar evolution theory +
stellar spectral libraries +
initial mass function

focus on galaxies with *well-established IMF* control regions: we can choose galaxy populations most *robust* to SED models

$$f_\lambda = \int_{t'=0}^{t'=t} \mathrm{SFR}(t') \; f_{\mathrm{SSP}}(t', \; Z(t')\,) \; \underline{e^{-\tau_{\mathrm{dust}}(t')}} \; \mathrm{d}t'$$

dust attenuation curve,
star-to-dust geometry

focus on galaxies with *well-established IMF, with low WISE IR emission*



we can be picky! — DESI and PFS will observe >40 million galaxies

DESI **Bright Galaxy Survey** (*Hahn et al. 2023c*):
a $r < 19.5$ magnitude-limited sample of >*10* million galaxies

summary

the photometry of a single galaxy contains *some* cosmological information

with *neural density estimation, hierarchical population inference, and* CAMELS we can exploit this information from *thousands of galaxies*

control regions: we can target galaxy populations most ***robust*** to galaxy and SED modeling

DESI **Bright Galaxy Survey** will soon observe a diverse magnitude-limited sample of *>10 million galaxies* to choose from

ChangHoon Hahn
changhoon.hahn@princeton.edu