Bayesian framework for model-data comparison with theoretical uncertainty quantification

Sunil Jaiswal

The Ohio State University

ISNET-11

ISNET-11: Information and Statistics in Nuclear Experiment and Theory

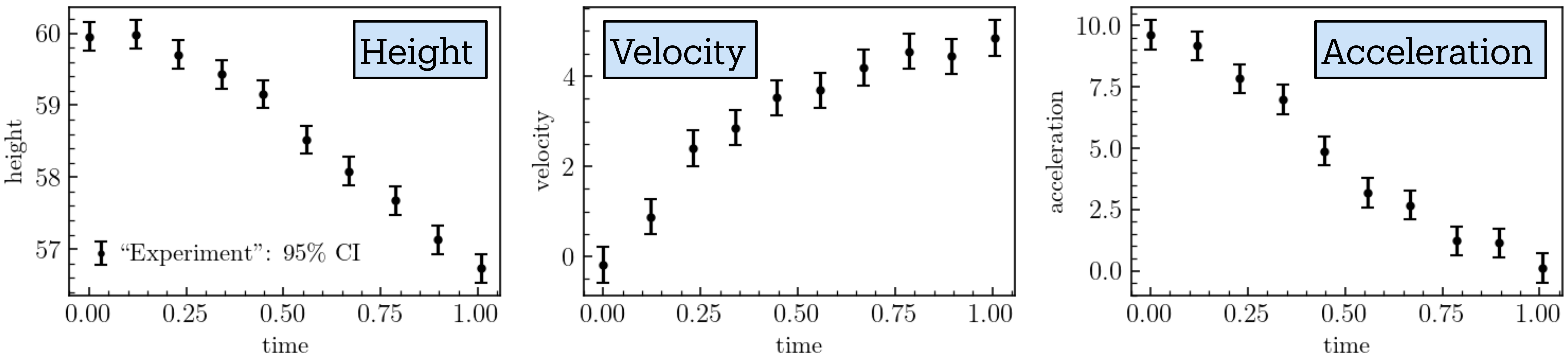
Based on arXiv: 2504.13144, 2509.19759

Collaborators: Richard Furnstahl, Ulrich Heinz, Matthew Pratola, Chun Shen

November 18, 2025

Motivation: A simple experiment

A ball is dropped from a tower. Height, velocity, acceleration are measured with time.



Drag force: $\mathbf{f}_D = -\,0.4v^2\,\hat{\mathbf{v}}$

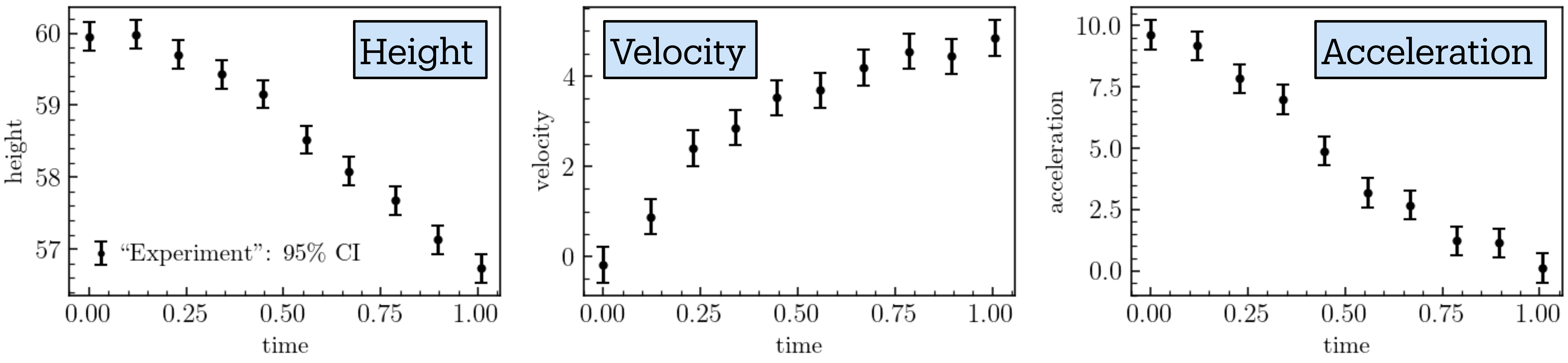
EoM:

$$a = \frac{d\mathbf{v}}{dt} = \mathbf{g} - 0.4v^2\hat{\mathbf{v}}\,, \quad \frac{dh}{dt} = -\,v$$

$h_0 = 0\,\mathrm{m}, v_0 = 0\,\mathrm{m/s}, \mathbf{g} = 9.8\,\mathrm{m/s}^2$

Motivation: A simple experiment

A ball is dropped from a tower. Height, velocity, acceleration are measured with time.



Drag force: $\mathbf{f}_D = -0.4v^2\,\hat{\mathbf{v}}$

EoM:

$$a = \frac{d\mathbf{v}}{dt} = \mathbf{g} - 0.4v^2\hat{\mathbf{v}}\,, \quad \frac{dh}{dt} = -v$$

$h_0 = 0\,\mathrm{m}, v_0 = 0\,\mathrm{m/s}, g = 9.8\,\mathrm{m/s^2}$

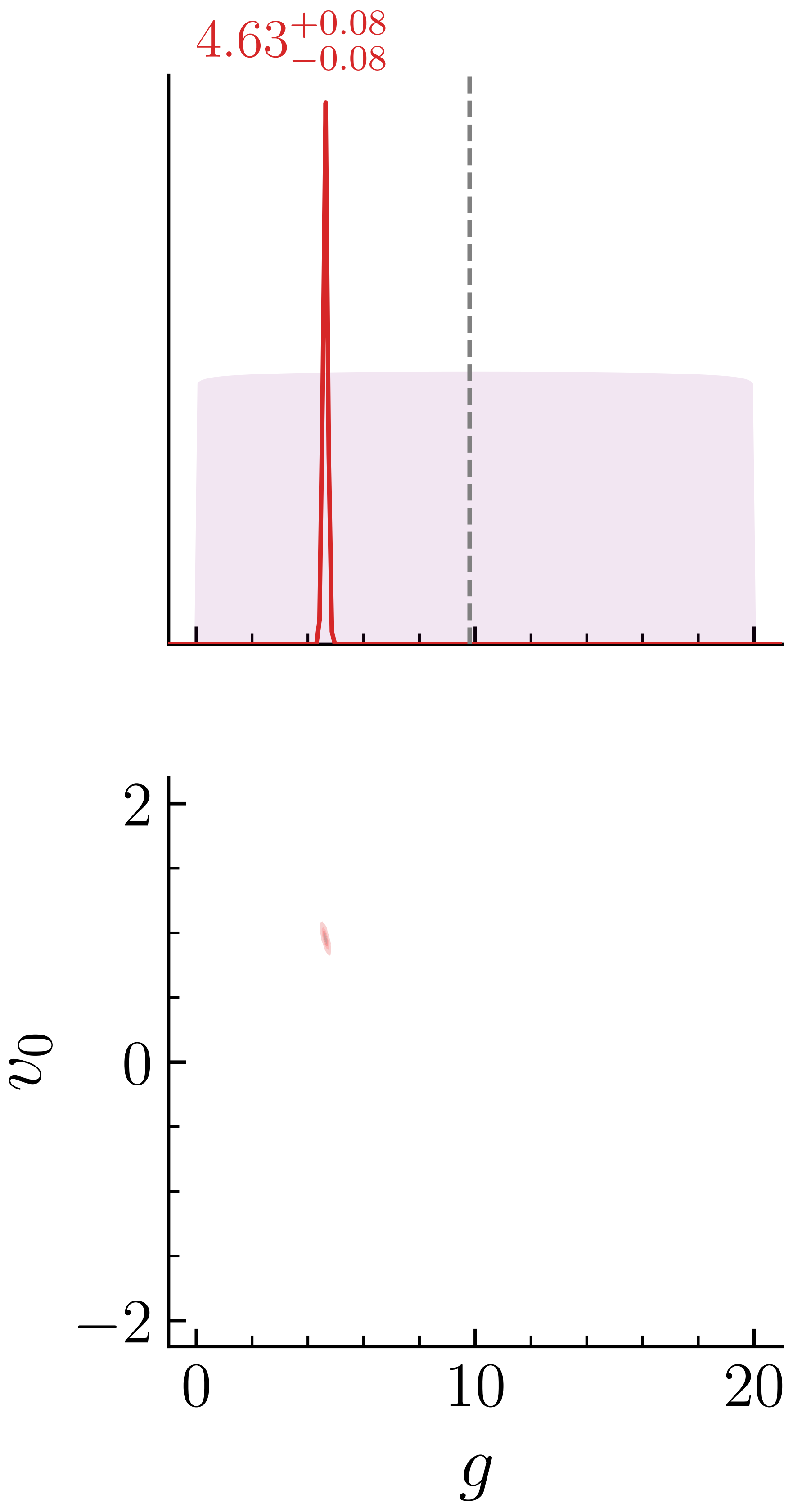
Theoretical model (theory):

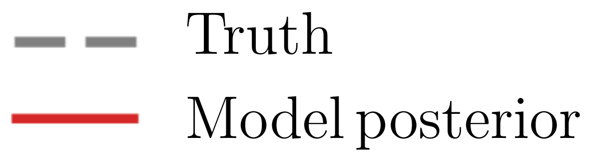
EoM: $a = g, \quad v = v_0 + gt, \quad h = h_0 - v_0 t - \frac{1}{2}gt^2$

2 model parameters : g, v_0

Goal: Infer the correct value $g = 9.8\,\mathrm{m/s^2}$ from data.

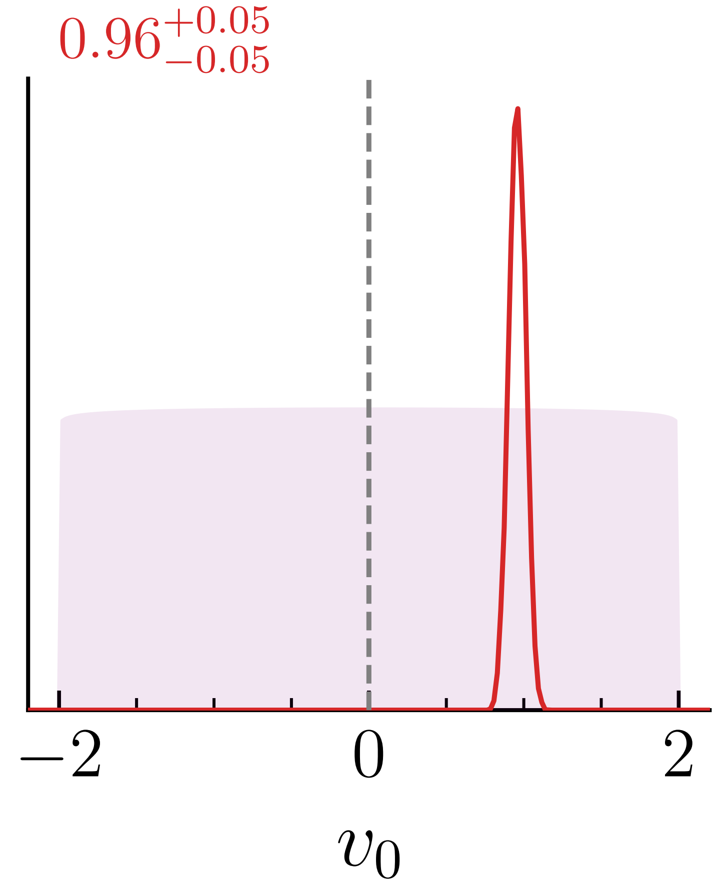
Model-data comparison: Bayesian inference

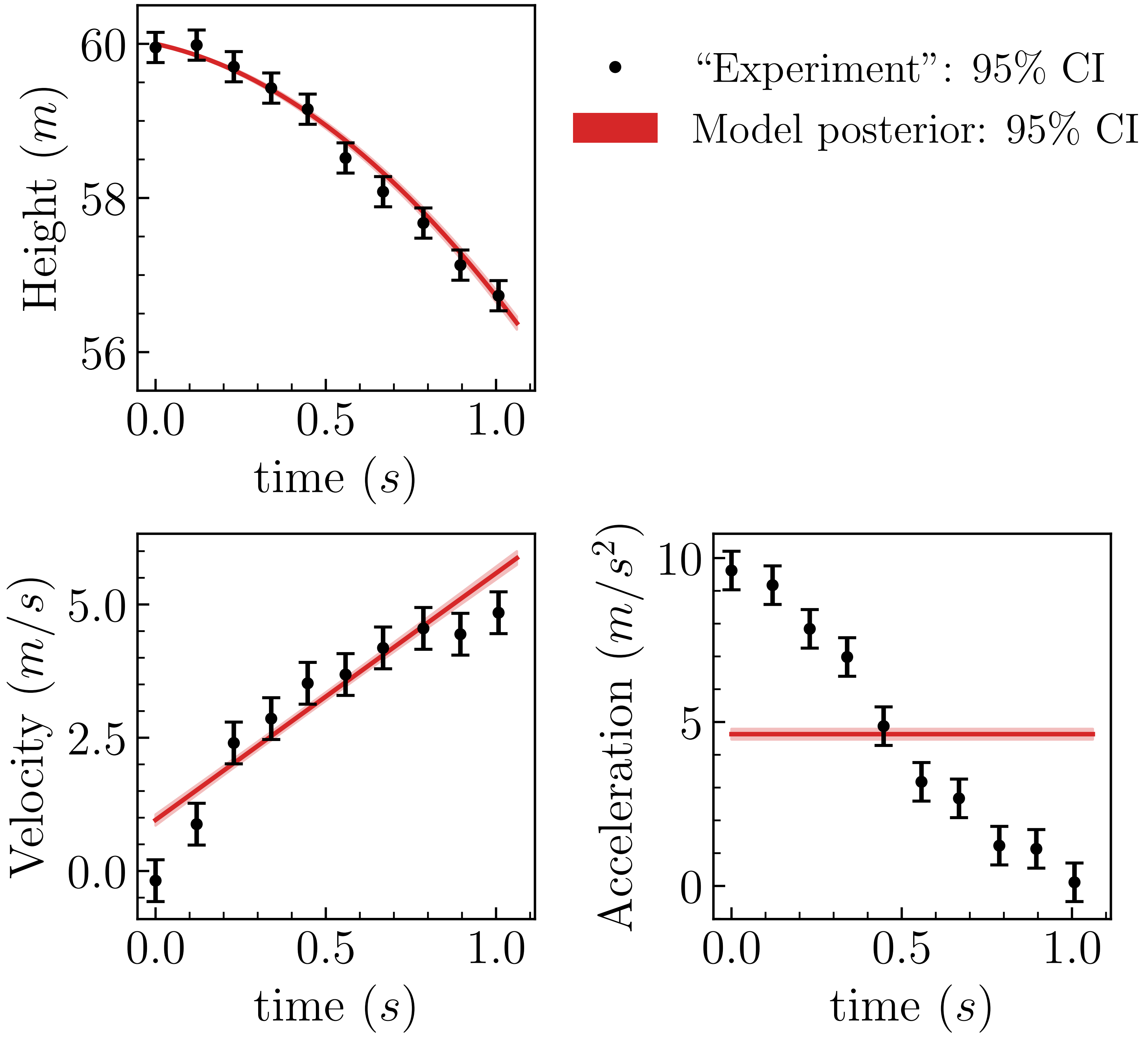




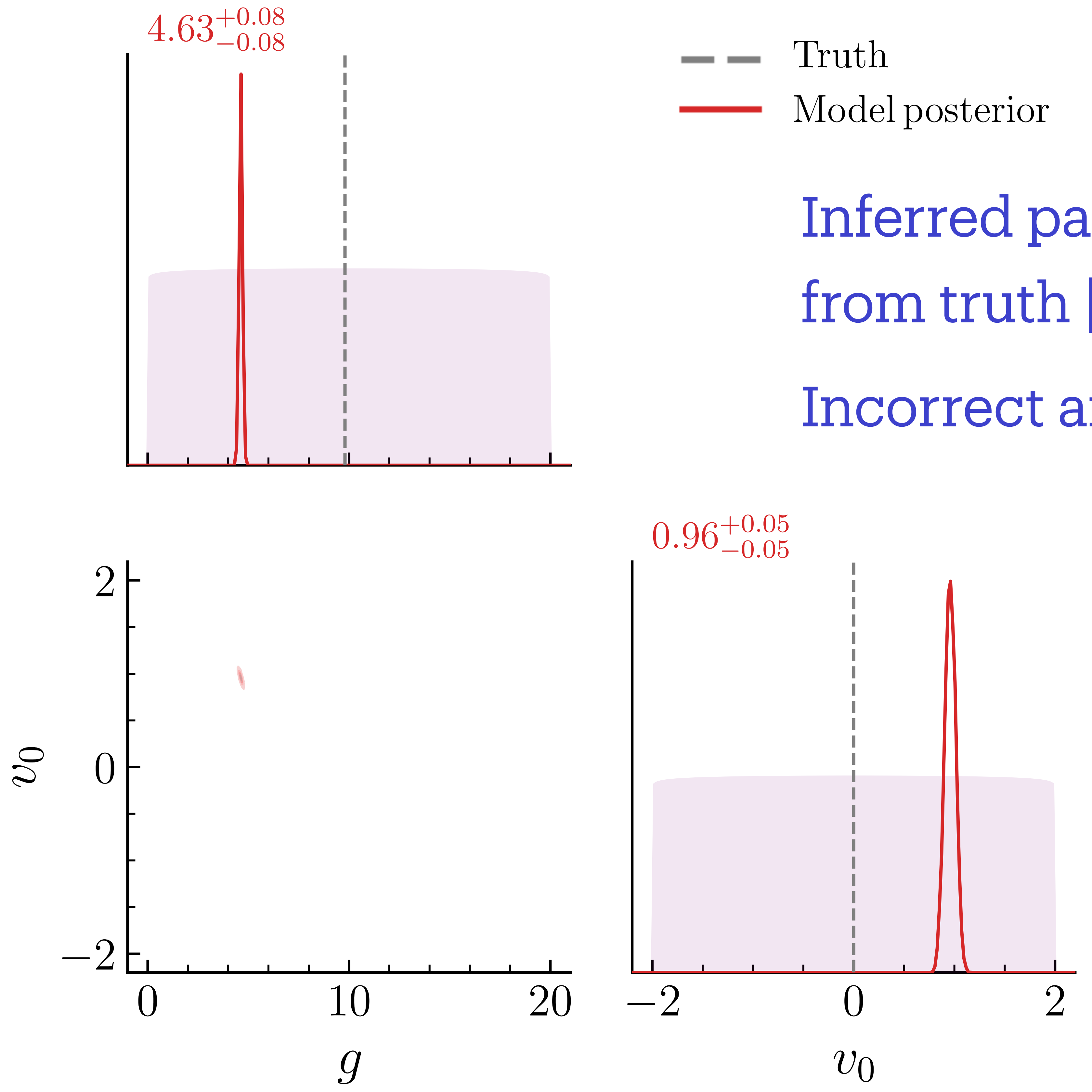
Inferred parameters far from truth [$g = 9.8,\ v_0 = 0$]

Incorrect and confident!



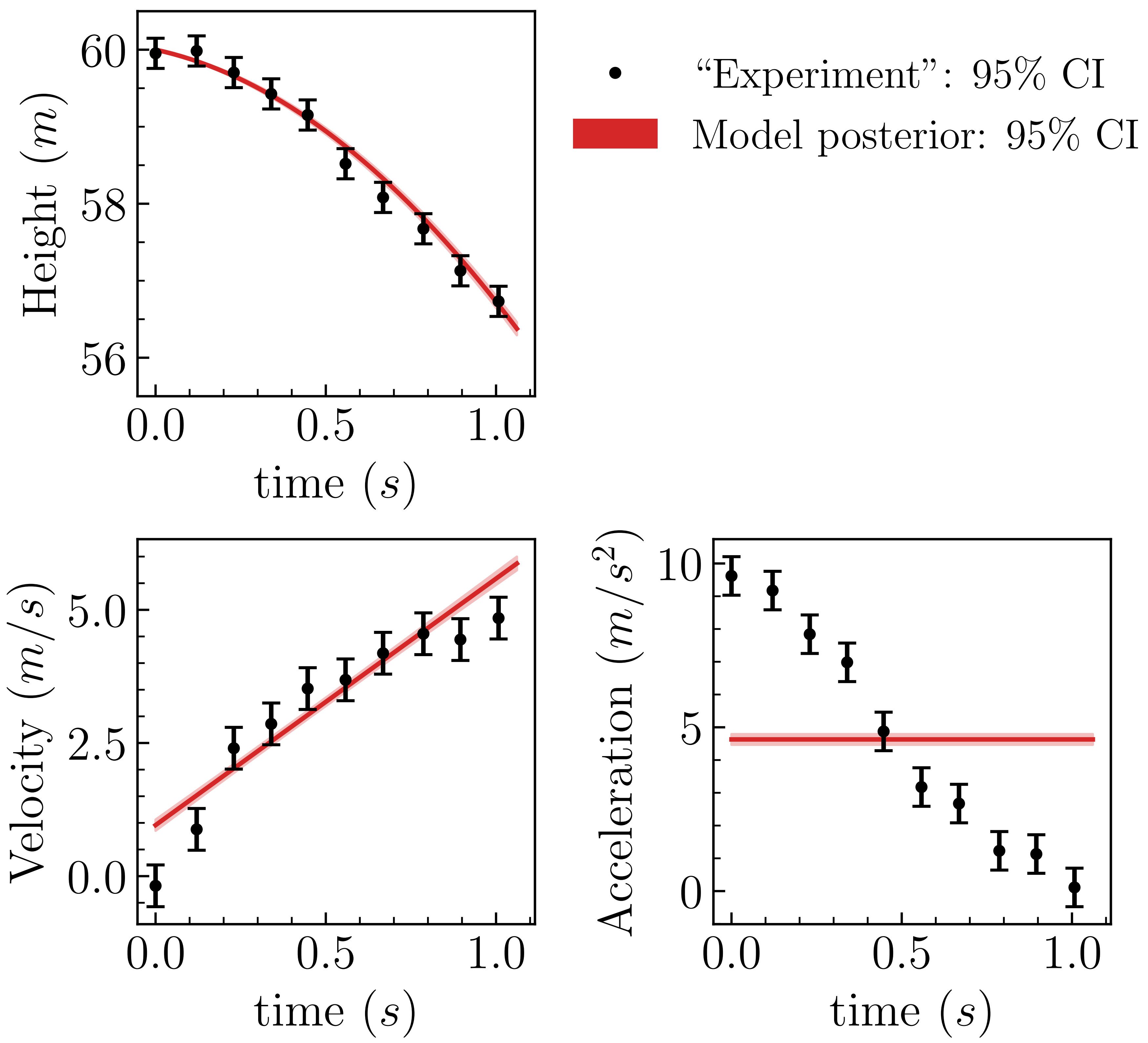


Model-data comparison: Bayesian inference



Inferred parameters far from truth [$g = 9.8,\ v_0 = 0$]

Incorrect and confident!



- Reason: Reality has air drag. Theory does not capture it and tries to fit data.
- Possible solution: Remove data at large times and fit. But where to draw the line?
- All theories are approximations of true underlying physics.
- **Need a method that correctly compares approximate theories with data, i.e., accounts for theoretical errors (uncertainties).**

Outline

- Model discrepancy framework
- Ball drop experiment with model discrepancy
- Real world application: Heavy-ion collisions
  - ➡ Applying model discrepancy framework to obtain constraints on transport properties of QGP
- Summary

Outline

- Model discrepancy framework
- Ball drop experiment with model discrepancy
- Real world application: Heavy-ion collisions
  - ➡ Applying model discrepancy framework to obtain constraints on transport properties of QGP
- Summary

Terminology: Model discrepancy = Theoretical error = Theoretical uncertainties
Theory = Physics model

Outline

- Model discrepancy framework
- Ball drop experiment with model discrepancy
- Real world application: Heavy-ion collisions
  - ➡ Applying model discrepancy framework to obtain constraints on transport properties of QGP
- Summary

Model discrepancy framework

- Every theoretical model is an **approximation** of the true underlying physics.

$$\underbrace{\zeta(x)}_{\text{Predictions from true theory}} = \underbrace{\eta(x, \theta)}_{\text{Predictions from approximate theory}} + \underbrace{\delta(x)}_{\text{Error of approximate theory predictions}}$$

Model discrepancy framework

- Every theoretical model is an **approximation** of the true underlying physics.

$$\underbrace{\zeta(x)}_{\text{Predictions from true theory}} = \underbrace{\eta(x, \theta)}_{\text{Predictions from approximate theory}} + \underbrace{\cancelto{0}{\delta(x)}}_{\text{Error of approximate theory predictions}}$$

- Usually, the error of theory is not known quantitatively.

- Therefore, conventional model-data comparison ignores theoretical uncertainty, implicitly assuming the model is perfect: $\delta(x) = 0$. Consequently, inferred **parameters can become mere fitting variables**, no longer corresponding to their true physical values.

Model discrepancy framework

- Every theoretical model is an **approximation** of the true underlying physics.

- Usually, the error of theory is not known quantitatively.

$$\underbrace{\zeta(x)}_{\text{Predictions from true theory}} = \underbrace{\eta(x,\theta)}_{\text{Predictions from approximate theory}} + \underbrace{\delta(x)}_{\text{Error of approximate theory predictions}}$$

- Therefore, conventional model-data comparison ignores theoretical uncertainty, implicitly assuming the model is perfect: $\delta(x) = 0$. Consequently, inferred **parameters can become mere fitting variables**, no longer corresponding to their true physical values.

- **Our approach :** SJ, C. Shen, R. J. Furnstahl, U. Heinz, and M. T. Pratola, PLB 870, 139946 (2025), arXiv: 2504.13144

Model discrepancy

M. Kennedy, A. O'Hagan doi:10.1111/1467-9868.00294

J. Brynjarsdóttir and A. O'Hagan, doi:10.1088/0266-5611/30/11/114007

D. Higdon, M. Kennedy, *et. al.*, doi:10.1137/S1064827503426693

Model discrepancy framework

- Every theoretical model is an **approximation** of the true underlying physics.

- Usually, the error of theory is not known quantitatively.

$$\underbrace{\zeta(x)}_{\text{Predictions from true theory}} = \underbrace{\eta(x,\theta)}_{\text{Predictions from approximate theory}} + \underbrace{\delta(x)}_{\text{Error of approximate theory predictions}}$$

- Therefore, conventional model-data comparison ignores theoretical uncertainty, implicitly assuming the model is perfect: $\delta(x) = 0$. Consequently, inferred **parameters can become mere fitting variables**, no longer corresponding to their true physical values.

- **Our approach :** SJ, C. Shen, R. J. Furnstahl, U. Heinz, and M. T. Pratola, PLB 870, 139946 (2025), arXiv: 2504.13144
  - ➡ We usually have some qualitative knowledge of the domain of validity of our theories. Example: "*We know the theory is more reliable in this regime than in that one*".
  - ➡ We use such important, qualitative knowledge of theory error to statistically model $\delta(x)$.
  - ➡ Model theory error with a Gaussian Process $\delta(\,\cdot\mid\phi) \sim \mathrm{GP}\left(\mathbf{0}, K(\,\cdot\,,\cdot\mid\phi)\right)$.

    By constructing covariance kernel K to reflect the prior qualitative knowledge of the error.

Model discrepancy

M. Kennedy, A. O'Hagan
doi:10.1111/1467-9868.00294

J. Brynjarsdóttir and A. O'Hagan,
doi:10.1088/0266-5611/30/11/114007

D. Higdon, M. Kennedy, *et. al.*,
doi:10.1137/S1064827503426693

Covariance kernel: central object of GP

A Gaussian process is a probability distribution over functions that fit a set of points. The functions have special properties determined by the covariance kernel.

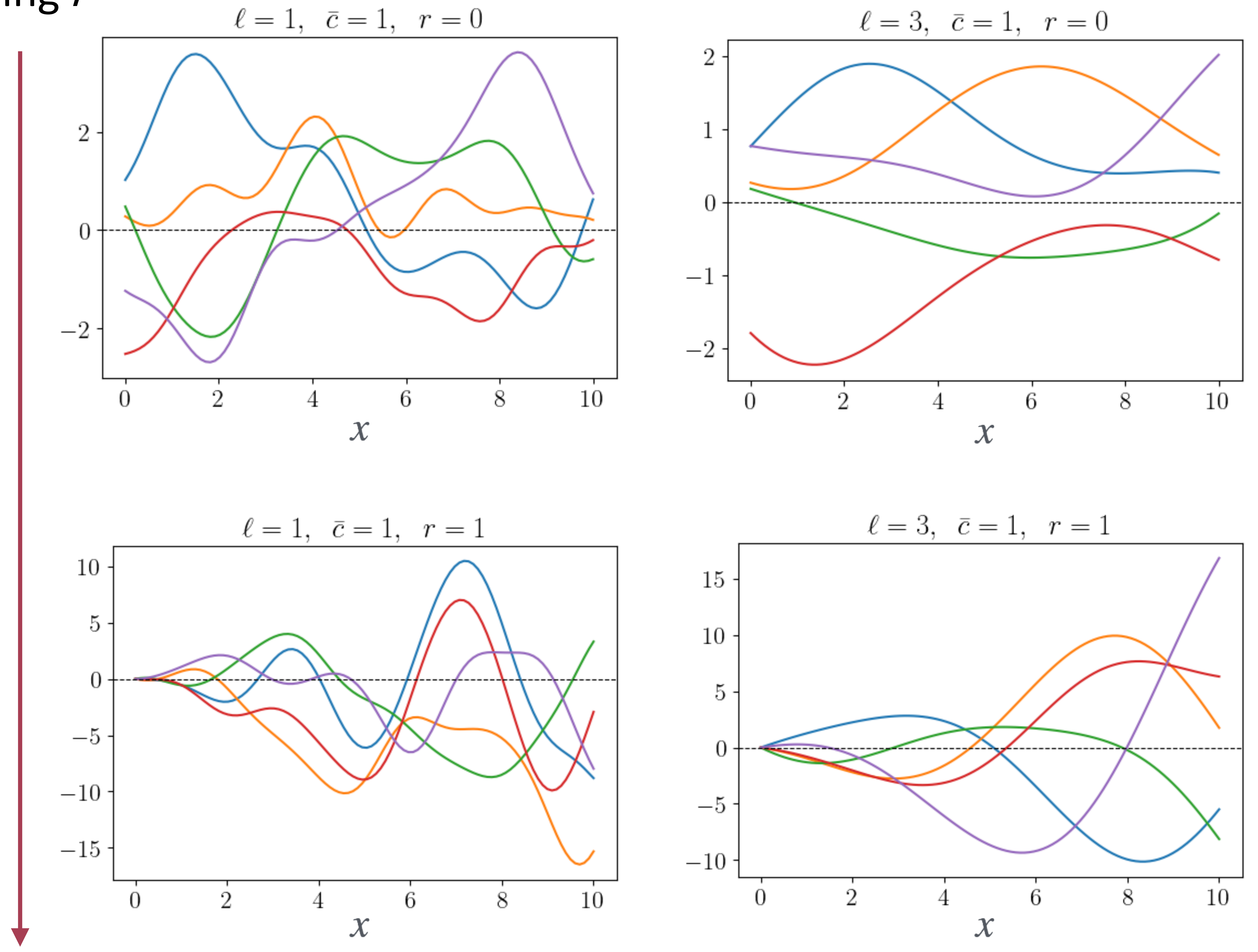
Consider the kernel:

$$K(x_i, x_j) = \bar{c}^2 (x_i x_j)^r \exp\left(-\frac{\|x_i - x_j\|^2}{2\ell^2}\right)$$

- The kernel, for fixed values of the parameters $(\bar{c}, r, \ell)$, generates covariance matrix for a multivariate gaussian distribution.
- ℓ: correlation length between two points.
- $\bar{c}$: controls the overall scale over which the functions vary. As $\bar{c} \to 0$, GP predictions are Dirac-delta distributions.
- r (≥ 0): **generates functions whose magnitude increase with** x.

Increasing ℓ

Increasing r



Ball drop experiment : with MD

- Theory EoM: $a = g, \quad v = v_0 + gt, \quad h = h_0 - v_0 t - \frac{1}{2}gt^2$

$\boxed{\text{2 model parameters : } g, v_0}$

- Statistical model:

$$\underbrace{y(t_i)}_{\text{Experimental observation}} = \underbrace{\eta(t_i, \theta)}_{\text{Theoretical prediction}} + \underbrace{\delta(t_i, \phi)}_{\text{Theoretical error}} + \underbrace{\epsilon(t_i)}_{\text{Observation error}}$$

- Model $\delta(\cdot \mid \phi) \sim \mathrm{GP}\left(\mathbf{0}, K(\cdot\,,\cdot \mid \phi)\right)$

Ball drop experiment : with MD

- Theory EoM: $a = g, \quad v = v_0 + gt, \quad h = h_0 - v_0 t - \frac{1}{2} g t^2$

2 model parameters : g, v_0

- Statistical model:

$$\underbrace{y(t_i)}_{\text{Experimental observation}} = \underbrace{\eta(t_i, \theta)}_{\text{Theoretical prediction}} + \underbrace{\delta(t_i, \phi)}_{\text{Theoretical error}} + \underbrace{\epsilon(t_i)}_{\text{Observation error}}$$

- Model $\delta(\cdot \mid \phi) \sim \mathrm{GP}\left(\mathbf{0}, K(\cdot\,,\cdot \mid \phi)\right)$

Prior knowledge about theory error:
Ignored drag => theory more correct for small velocities => **Theory is more correct at small t than at large t.**

Ball drop experiment : with MD

- Theory EoM: $a = g, \quad v = v_0 + gt, \quad h = h_0 - v_0 t - \dfrac{1}{2} g t^2$

2 model parameters : g, v_0

- Statistical model:

$$\underbrace{y(t_i)}_{\text{Experimental observation}} = \underbrace{\eta(t_i, \theta)}_{\text{Theoretical prediction}} + \underbrace{\delta(t_i, \phi)}_{\text{Theoretical error}} + \underbrace{\epsilon(t_i)}_{\text{Observation error}}$$

- Model $\delta(\cdot \mid \phi) \sim \mathrm{GP}\left(\mathbf{0}, K(\cdot, \cdot \mid \phi)\right)$

Prior knowledge about theory error:
Ignored drag => theory more correct for small velocities => **Theory is more correct at small t than at large t.**

- Theory error should grow as t increases. Consider the kernels (with $r \geq 0$):

SJ, C. Shen, R. J. Furnstahl, U. Heinz, and M. T. Pratola, PLB 870, 139946 (2025), arXiv: 2504.13144

➡ Kernel I: $K(t_i, t_j \mid \phi) \equiv s^2 + \bar{c}^2 (t_i t_j)^r \exp\left(-\dfrac{\|t_i - t_j\|^2}{2\ell^2}\right).$

Encodes the knowledge (conservative):
Theory is more reliable at small t than at large t.

➡ Kernel II: $K(t_i, t_j \mid \phi) \equiv \bar{c}^2 (t_i t_j)^r \exp\left(-\dfrac{\|t_i - t_j\|^2}{2\ell^2}\right).$

Encodes the knowledge (more informative):
Theory is correct at $t = 0$, but our confidence in its validity decreases as t increases

➡ Allows theory errors to be constant or increase with time (but not decrease).

➡ Non-vanishing theory error allowed even for Kernel I ($s \neq 0$).

➡ $s = \bar{c} = 0 \implies \delta = 0.$

Bayesian inference with MD

$$\underbrace{y(t_i)}_{\text{Experimental observation}} = \underbrace{\eta(t_i, \theta)}_{\text{Theoretical prediction}} + \underbrace{\delta(t_i, \phi)}_{\text{Theoretical error}} + \underbrace{\epsilon(t_i)}_{\text{Observation error}}$$

Bayesian updating of knowledge

$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)} \Longrightarrow \underbrace{\mathrm{pr}(\boldsymbol{\alpha}|\mathbf{y}_{\mathrm{exp}}, I)}_{\text{posterior}} \propto \underbrace{\mathrm{pr}(\mathbf{y}_{\mathrm{exp}}|\boldsymbol{\alpha}, I)}_{\text{likelihood}} \times \underbrace{\mathrm{pr}(\boldsymbol{\alpha}|I)}_{\text{prior}}$$

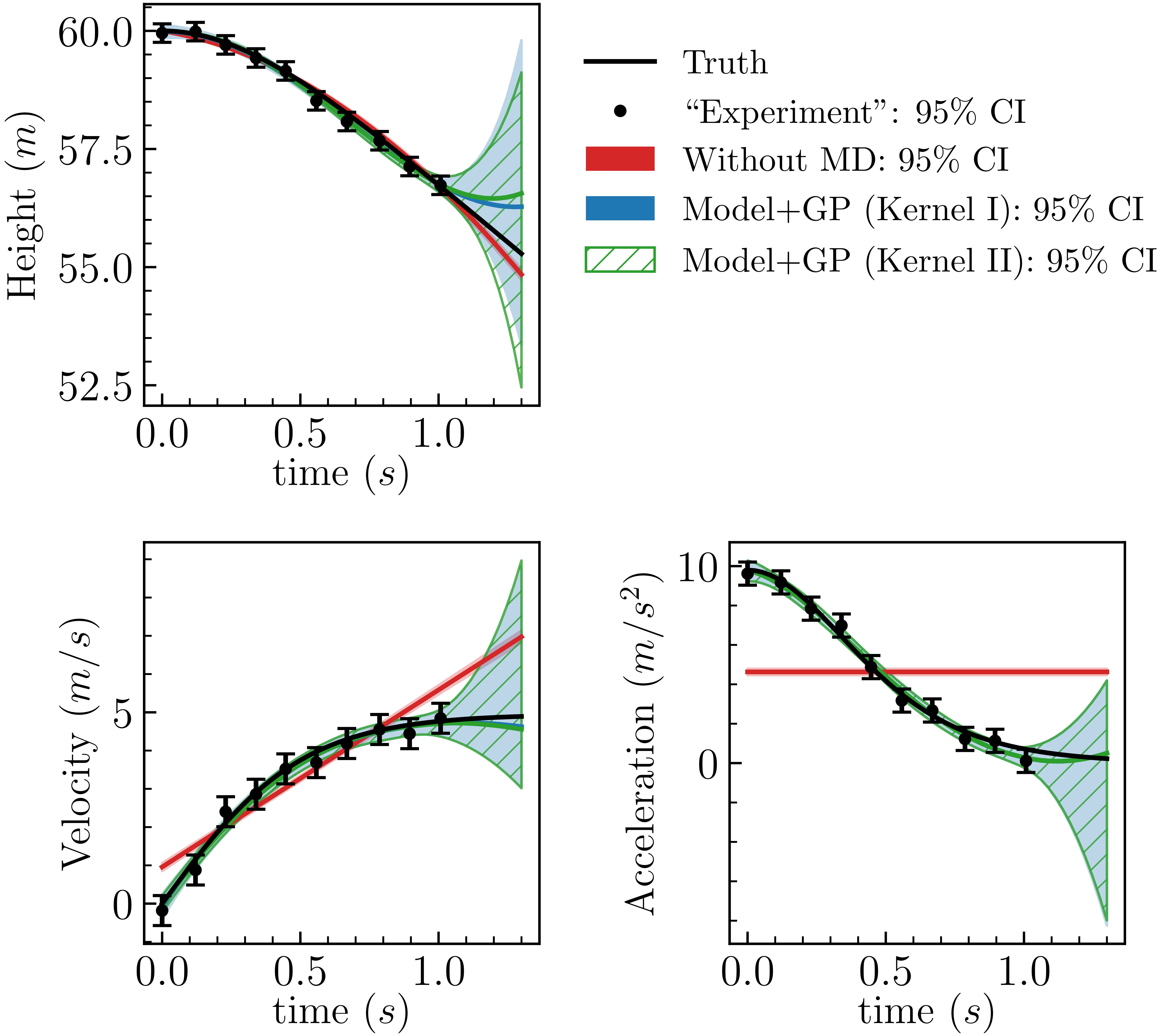
Fit $\zeta(t;\theta,\phi) \equiv \eta(t,\theta) + \delta(t,\phi)$ to data and estimate θ, ϕ simultaneously.

Note: individual discrepancy GP's (with hyperparameters) for each observable.

Beta priors mimicking uniform priors are considered for θ, ϕ.

Bayesian inference: Model + discrepancy predictions

SJ, C. Shen, R. J. Furnstahl, U. Heinz, and M. T. Pratola, PLB 870, 139946 (2025), arXiv: 2504.13144



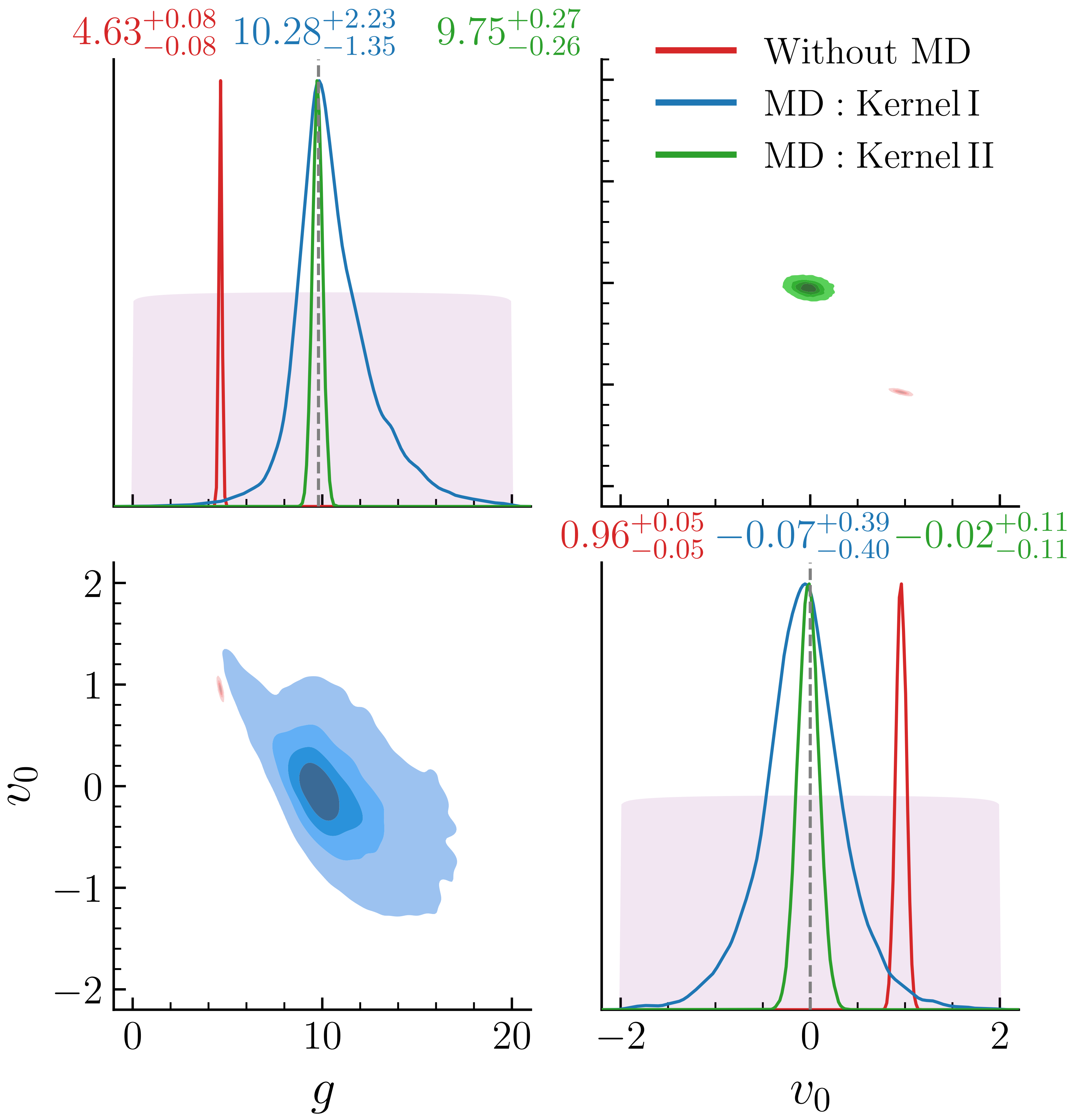
- Predictions from $\zeta(t;\theta,\phi) \equiv \eta(t,\theta) + \delta(t,\phi)$
- W/ MD predictions agrees with data.
- Extrapolation should be avoided.

➡ Kernel I: $K(t_i, t_j \mid \phi) \equiv s^2 + \bar{c}^2 (t_i t_j)^r \exp\left(-\frac{\|t_i - t_j\|^2}{2\ell^2}\right).$

➡ Kernel II: $K(t_i, t_j \mid \phi) \equiv \bar{c}^2 (t_i t_j)^r \exp\left(-\frac{\|t_i - t_j\|^2}{2\ell^2}\right).$

Bayesian inference: Model predictions

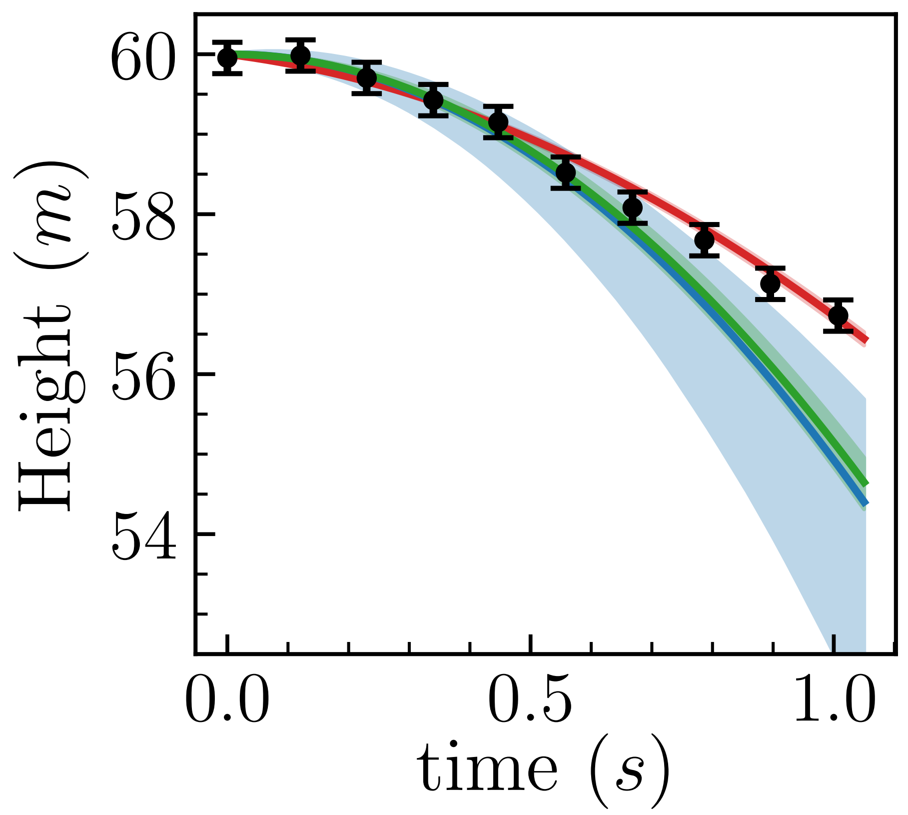
SJ, C. Shen, R. J. Furnstahl, U. Heinz, and M. T. Pratola, PLB 870, 139946 (2025), arXiv: 2504.13144

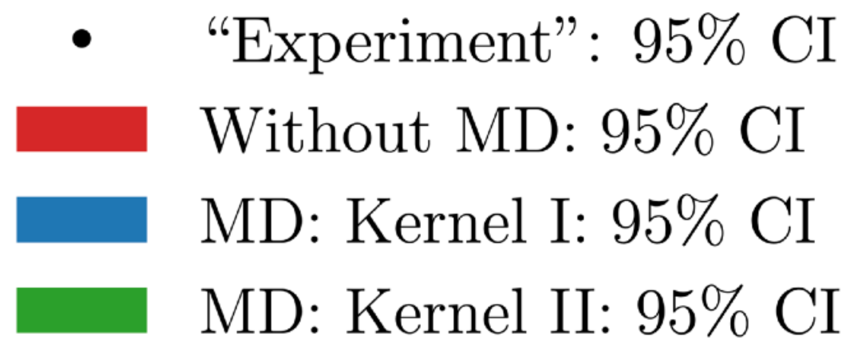


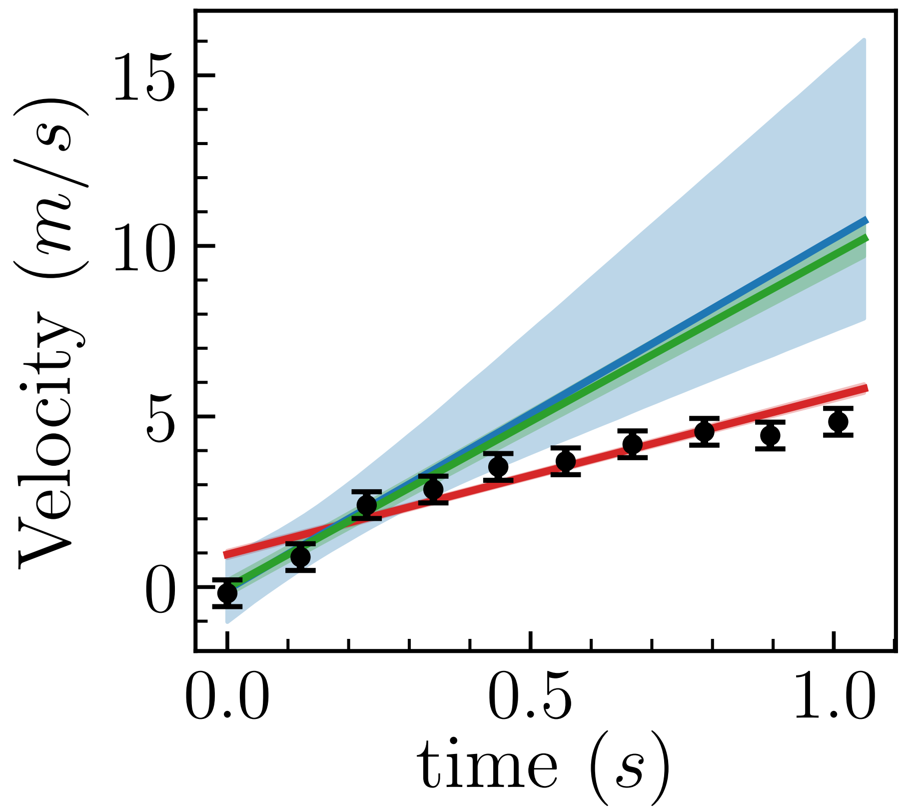
2 parameter inference

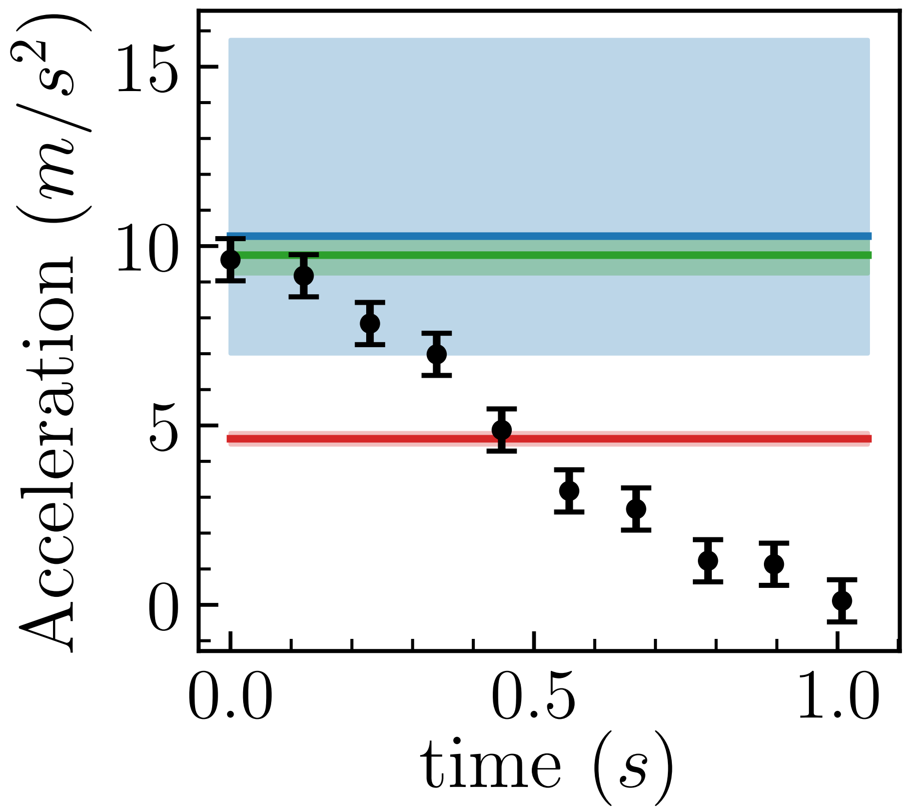
2+(3x4)=14 parameter inference

2+(3x3)=11 parameter inference



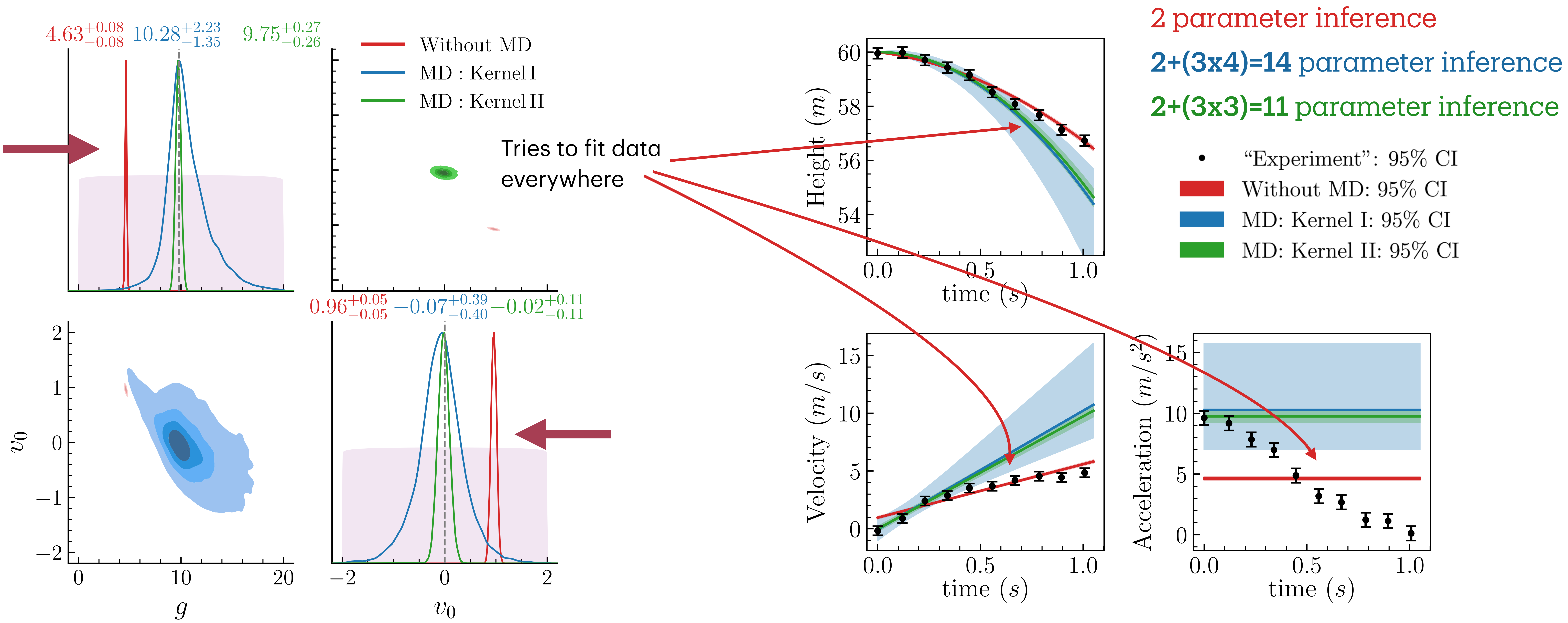






Bayesian inference: Model predictions

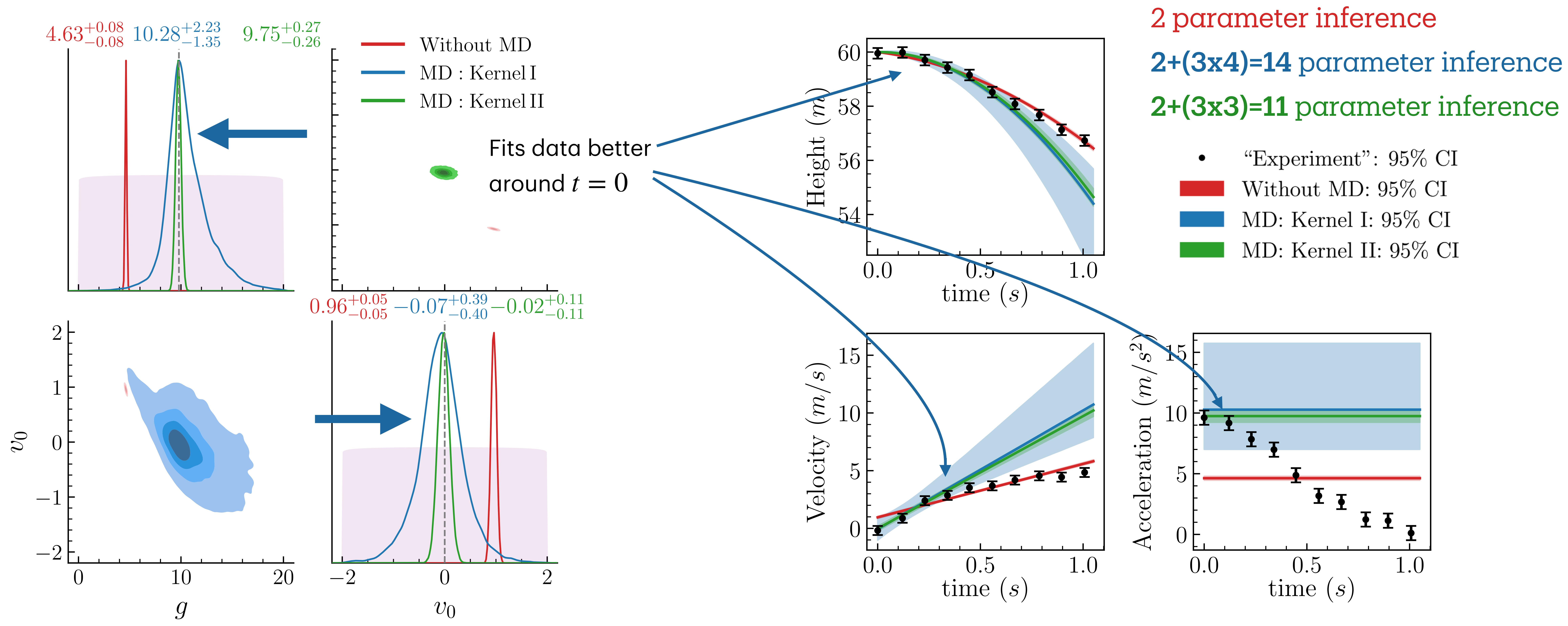
SJ, C. Shen, R. J. Furnstahl, U. Heinz, and M. T. Pratola, PLB 870, 139946 (2025), arXiv: 2504.13144



w/o MD (red): Inferred parameters far from truth [$g = 9.8,\ v_0 = 0$]; incorrect and confident.

Bayesian inference: Model predictions

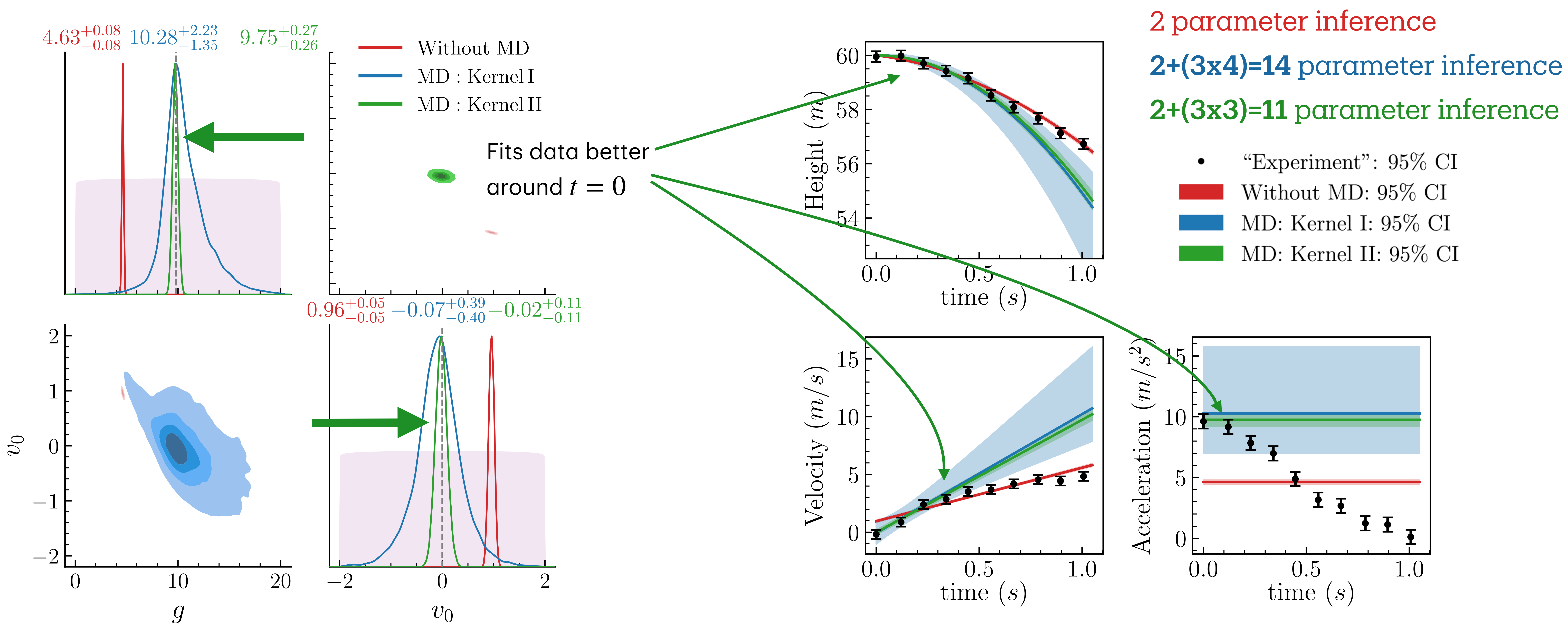
SJ, C. Shen, R. J. Furnstahl, U. Heinz, and M. T. Pratola, PLB 870, 139946 (2025), arXiv: 2504.13144



w/o MD (red): Inferred parameters far from truth $[g = 9.8,\ v_0 = 0]$; incorrect and confident.
w/ MD (blue — conservative kernel): $g,\ v_0$ true values inside posteriors.

Bayesian inference: Model predictions

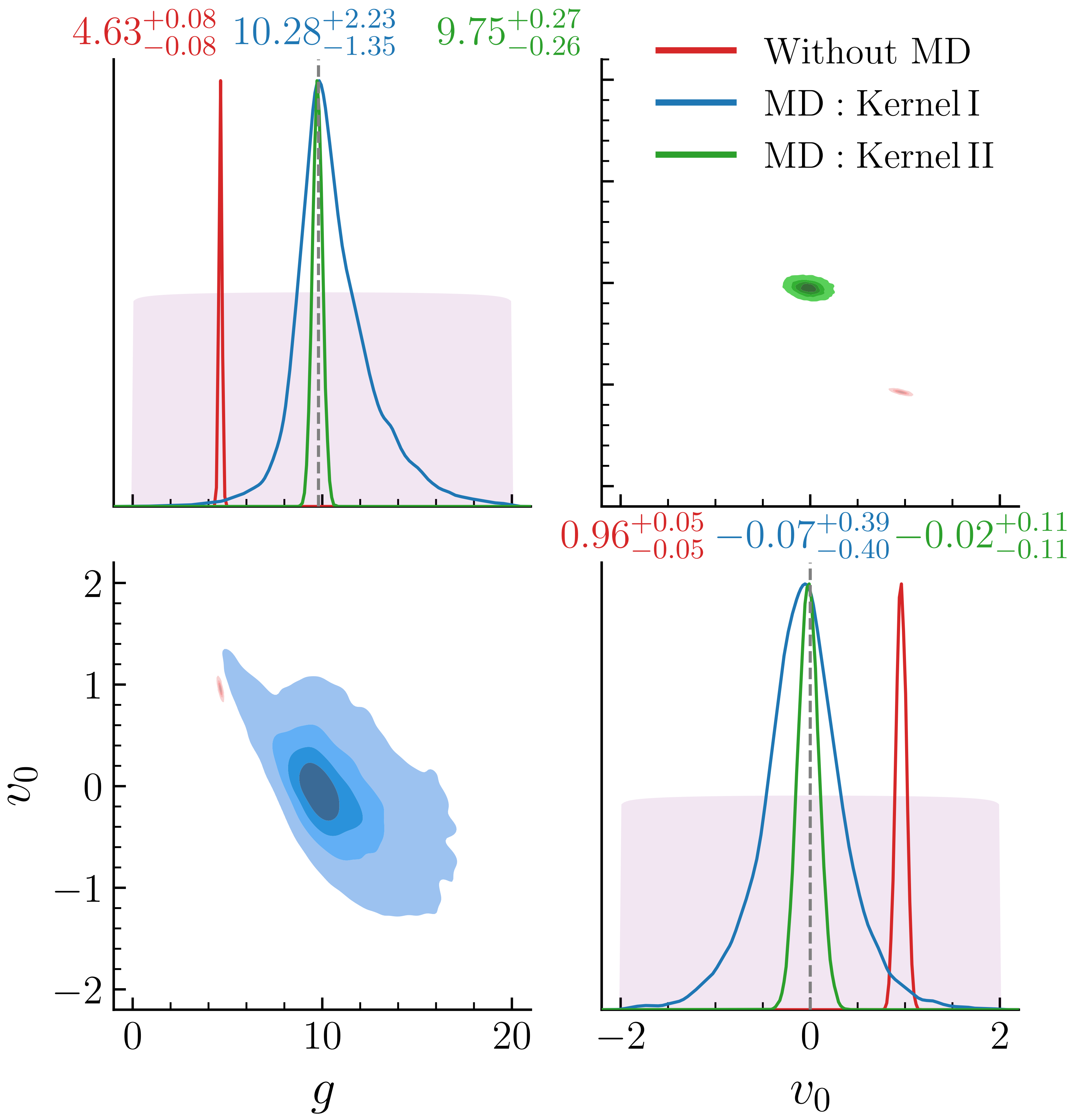
SJ, C. Shen, R. J. Furnstahl, U. Heinz, and M. T. Pratola, PLB 870, 139946 (2025), arXiv: 2504.13144



w/o MD (red): Inferred parameters far from truth [$g = 9.8,\ v_0 = 0$]; incorrect and confident.
w/ MD (blue — conservative kernel): $g,\ v_0$ true values inside posteriors.
w/ MD (Green — more informative of theory error): $g,\ v_0$ posteriors correct and confident!

Bayesian inference: Model predictions

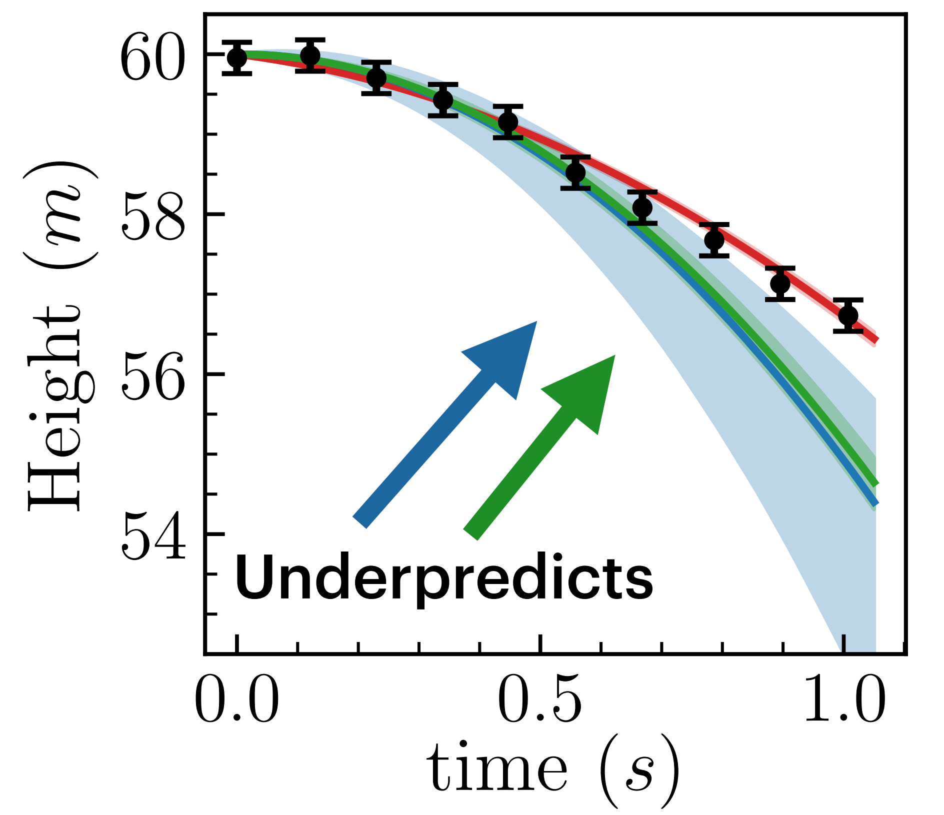
SJ, C. Shen, R. J. Furnstahl, U. Heinz, and M. T. Pratola, PLB 870, 139946 (2025), arXiv: 2504.13144

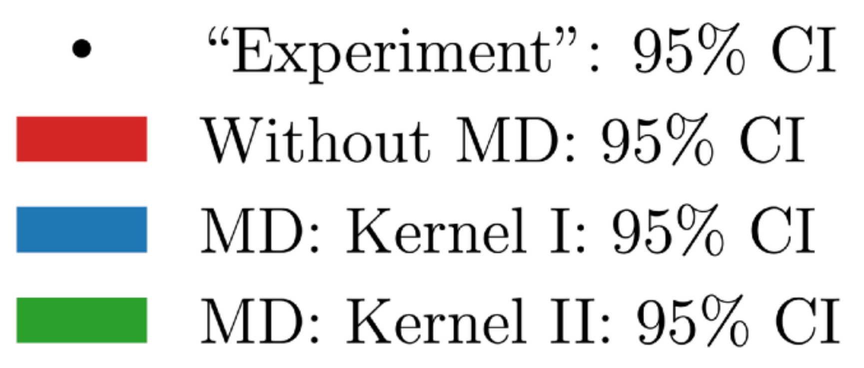


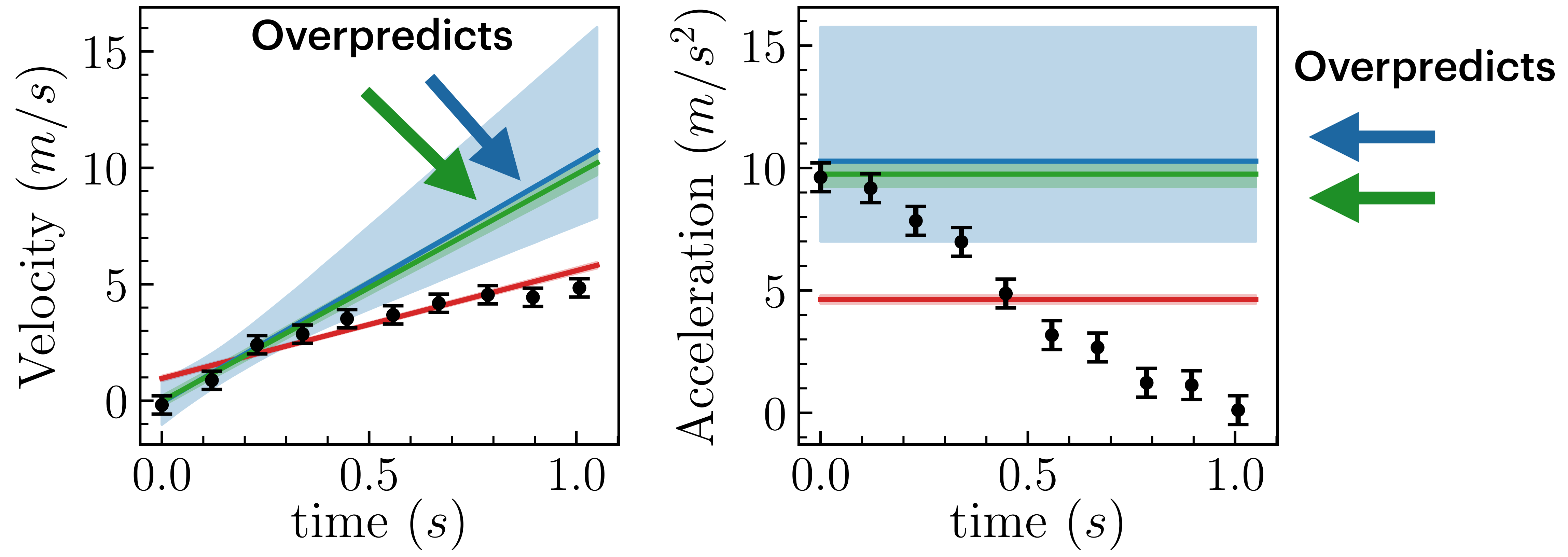
2 parameter inference

2+(3x4)=14 parameter inference

2+(3x3)=11 parameter inference







w/o MD (red): Inferred parameters far from truth [$g = 9.8,\ v_0 = 0$]; incorrect and confident.

w/ MD (blue — conservative kernel): $g,\ v_0$ true values inside posteriors.

w/ MD (Green — more informative of theory error): $g,\ v_0$ posteriors correct and confident!

Check: stability of posteriors

SJ, C. Shen, R. J. Furnstahl, U. Heinz, and M. T. Pratola, PLB 870, 139946 (2025), arXiv: 2504.13144

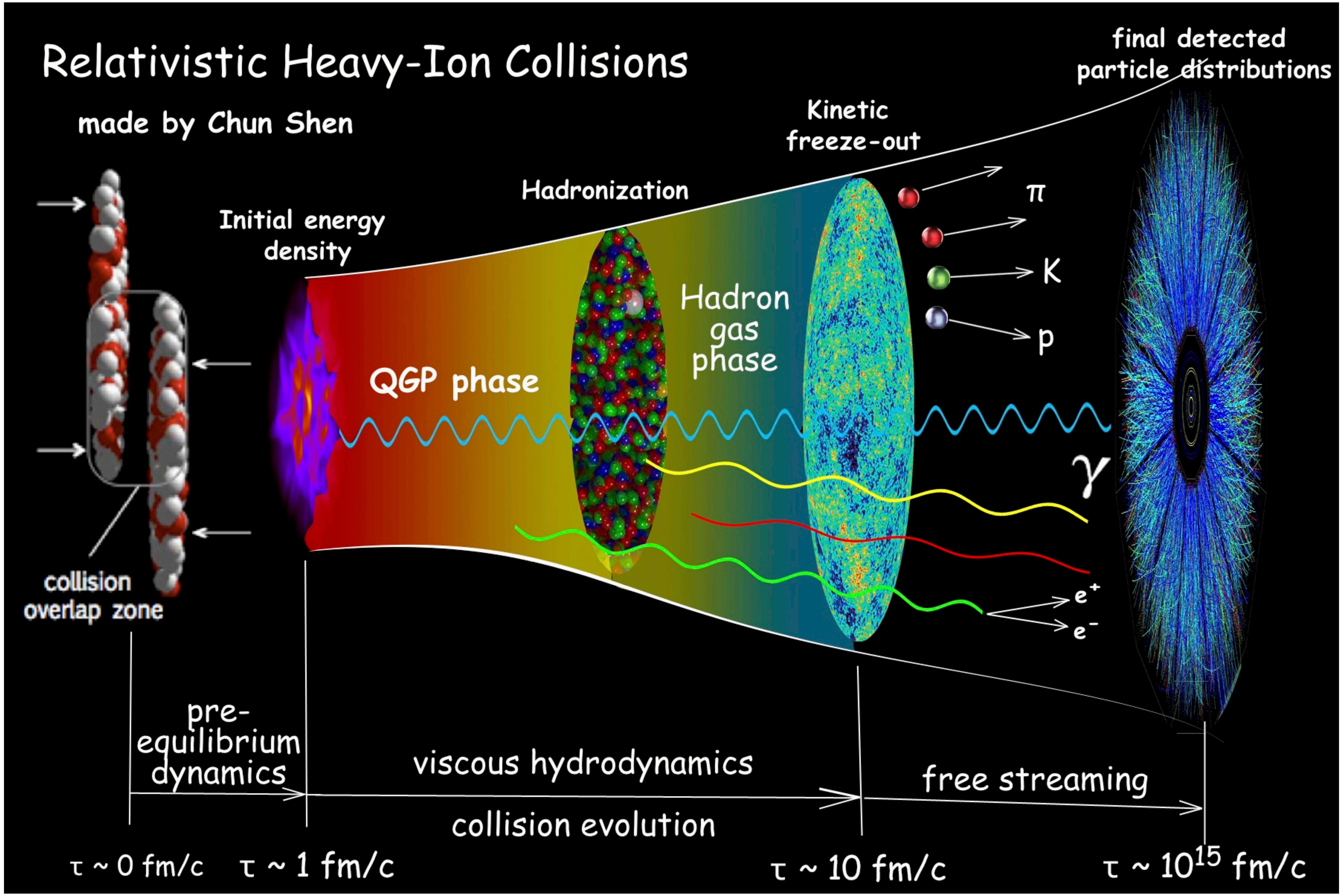
Calibrate with **height** → Calibrate with **height, velocity** → Calibrate with **all observables**

- w/o MD (red): Posteriors narrow around incorrect parameter value.
- w/ MD (blue, green): Posteriors narrow around correct parameter value.
  More informative Kernel II posteriors within Kernel I (conservative kernel) — Consistent.

Outline

- Model discrepancy framework
- Ball drop experiment with model discrepancy
- Real world application: Heavy-ion collisions
  - ➡ Applying model discrepancy framework to obtain constraints on transport properties of QGP
- Summary

Heavy-ion collisions: Many stages of evolution

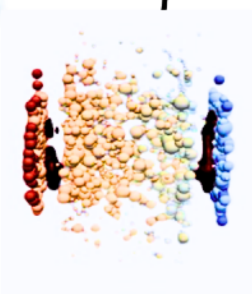


Multi-stage physics model

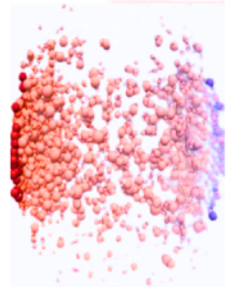
Initial conditions
- Trento
- IP-Glasma

Pre-Equilibrium
- Free Stream
- KoMPoST
- KTiSO

Viscous Hydrodynamics
- 2nd order viscous hydro
- Viscous anisotropic hydro



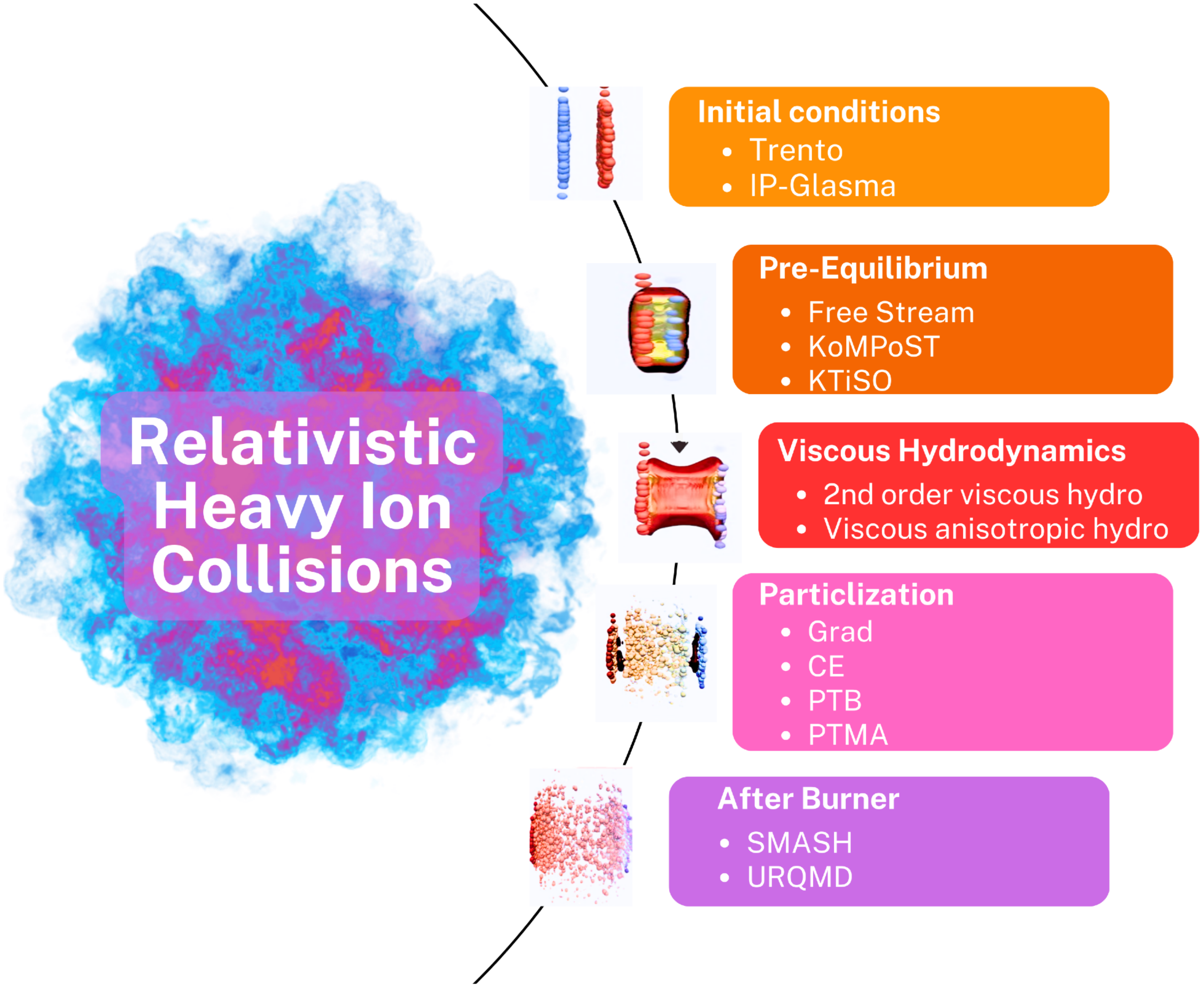
Particlization
- Grad
- CE
- PTB
- PTMA



After Burner
- SMASH
- URQMD

Slide adapted from D. Liyanage

Multi-stage physics model



All models are approximate with associated theoretical errors.

Errors from one stage can propagate to other stages.

Slide adapted from D. Liyanage

Multi-stage physics model

Initial conditions
- Trento
- IP-Glasma

Pre-Equilibrium
- Free Stream
- KoMPoST
- KTiSO

Viscous Hydrodynamics
- 2nd order viscous hydro
- Viscous anisotropic hydro

Particlization
- Grad
- CE
- PTB
- PTMA

After Burner
- SMASH
- URQMD

T$_R$ENTo
- N : normalization
- p : generalized mean power
- σ_k : multiplicity fluct. variance
- w : nucleon width
- d_{min} : min. dist. betw. nucleons

Free-stream
- τ_R : free-stream time scale
- α : energy density scaling of τ_R

MUSIC
- $(\eta/s)(T)$: shear viscosity
- $(\zeta/s)(T)$: bulk viscosity
- b_π: Shear relax time factor

GRAD **CE**

T_{sw}: switching temperature

SMASH

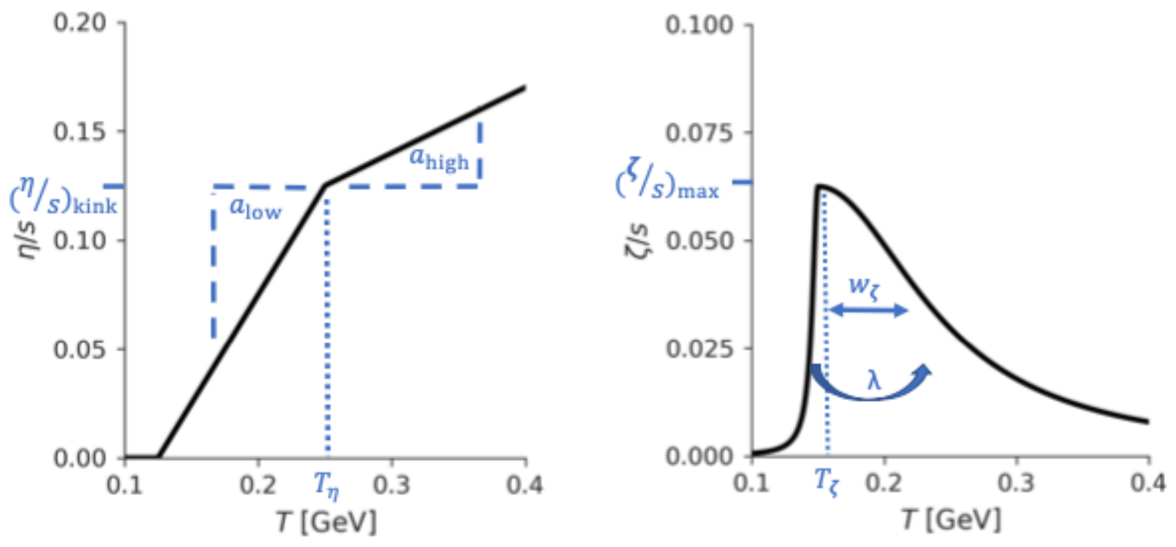
Pb - Pb at 2.76 TeV
Model predictions
for observables

JETSCAPE SIMS calibration

D. Everett *et al.* 2010.03928, 2011.01430

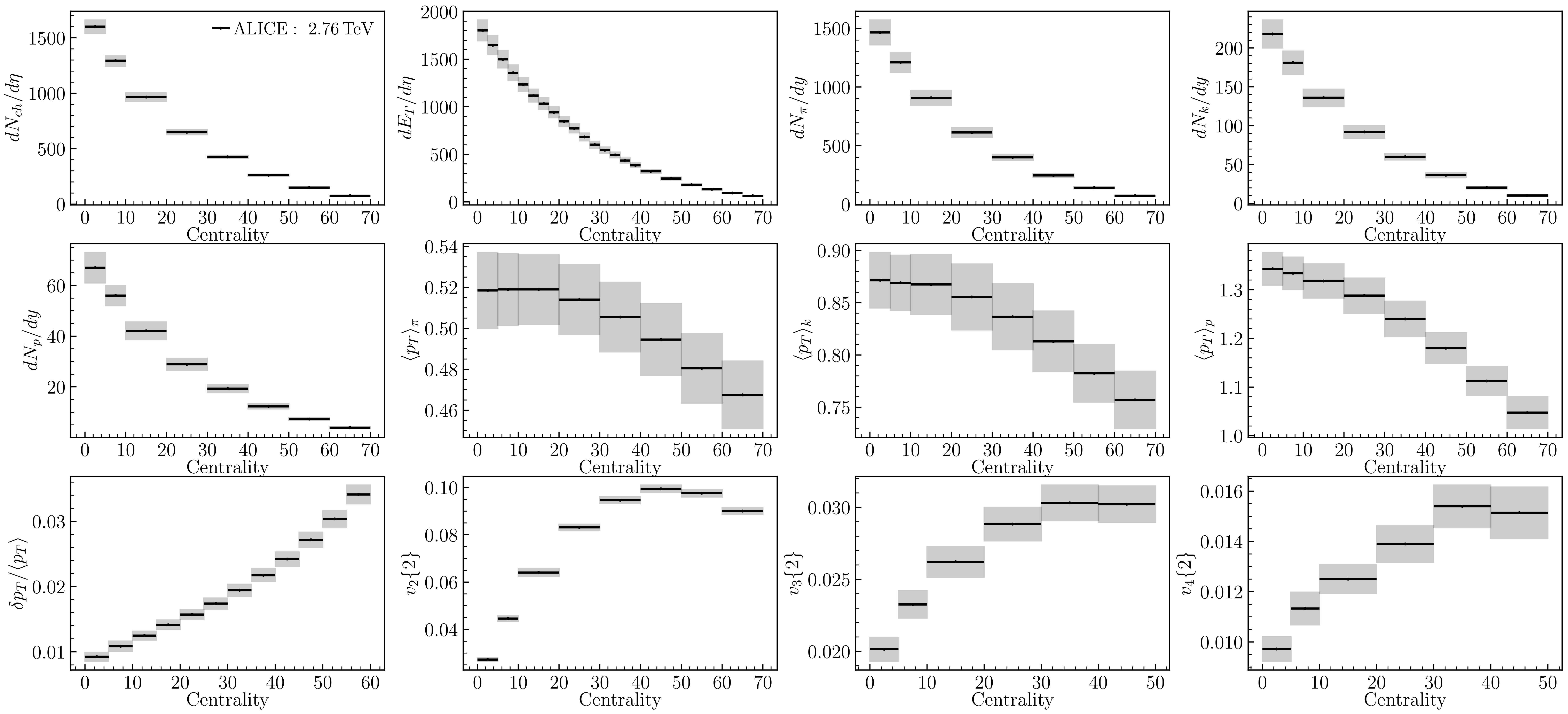
17 model parameters

Parameter name	Symbol	
Normalization	N	
Generalized mean	p	
Multiplicity fluctuation	σ_k	
Nucleon width	w	
Min. dist. btw. nucleons	$d^3_{\rm min}$	
Free-streaming time scale	τ_R	
Free-streaming energy dep.	α	
Temperature of (η/s) kink	$T_{\rm kink}$	$(\eta/s)(T)$
Low temp. slope of (η/s)	$a_{\rm low}$	$(\eta/s)(T)$
High temp. slope of (η/s)	$a_{\rm high}$	$(\eta/s)(T)$
(η/s) at kink	$(\eta/s)_{\rm kink}$	$(\eta/s)(T)$
Maximum of (ζ/s)	$(\zeta/s)_{\max}$	$(\zeta/s)(T)$
Temperature of (ζ/s) peak	T_ζ	$(\zeta/s)(T)$
Width of (ζ/s) peak	w_ζ	$(\zeta/s)(T)$
Asymmetry of (ζ/s) peak	λ_ζ	$(\zeta/s)(T)$
Shear relaxation time factor	b_π	
Particlization temperature	$T_{\rm sw}$	



Slide adapted from D. Liyanage

Experimental data: Pb+Pb at 2.76 TeV



SJ, arXiv: 2509.19759

(i) the number of charged hadrons per unit pseudorapidity $dN_{\rm ch}/d\eta$ [28]; (ii) the transverse energy per unit pseudorapidity $dE_T/d\eta$ [29]; (iii) the number of identified charged hadrons per unit rapidity dN_i/dy, $i \in \{\pi, k, p\}$ [30]; (iv) the mean transverse momenta of identified hadrons $\langle p_T \rangle_i$, $i \in \{\pi, k, p\}$ [30]; (v) the two-particle cumulant flow coefficients $v_n\{2\}$ for $n = 2, 3, 4$ [31]; (vi) the fluctuation in the mean transverse momentum $\delta p_T/\langle p_T \rangle$ [32].

[28] K. Aamodt, et al., Centrality dependence of the charged-particle multiplicity density at mid-rapidity in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, Phys. Rev. Lett. 106 (2011) 032301. arXiv:1012.1657, doi:10.1103/PhysRevLett.106.032301.

[29] J. Adam, et al., Measurement of transverse energy at midrapidity in Pb-Pb collisions at $\sqrt{s_{\rm NN}} = 2.76$ TeV, Phys. Rev. C94 (3) (2016) 034903. arXiv:1603.04775, doi:10.1103/PhysRevC.94.034903.

[30] B. Abelev, et al., Centrality dependence of π, K, p production in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, Phys. Rev. C88 (2013) 044910. arXiv:1303.0737, doi:10.1103/PhysRevC.88.044910.

[31] K. Aamodt, et al., Higher harmonic anisotropic flow measurements of charged particles in Pb-Pb collisions at $\sqrt{s_{NN}}$=2.76 TeV, Phys. Rev. Lett. 107 (2011) 032301. arXiv:1105.3865, doi:10.1103/PhysRevLett.107.032301.

[32] B. B. Abelev, et al., Event-by-event mean $p_{\rm T}$ fluctuations in pp and Pb-Pb collisions at the LHC, Eur. Phys. J. C74 (10) (2014) 3077. arXiv:1407.5530, doi:10.1140/epjc/s10052-014-3077-y.

12 observables. 110 observations

Modeling theoretical uncertainty (error) in heavy-ion

- Statistical model:

$$\underbrace{y(x_i)}_{\text{Experimental observation}} = \underbrace{\eta_{\text{mod}}(x_i, \theta)}_{\text{Theoretical prediction}} + \underbrace{\delta_{\text{MD}}(x_i, \phi)}_{\text{Theoretical error}} + \underbrace{\epsilon(x_i)}_{\text{Observation error}}$$

- Model $\delta(\,\cdot \mid \phi) \sim \text{GP}\left(\mathbf{0}, K(\,\cdot\,,\cdot \mid \phi)\right)$

Prior knowledge about theory error: **JETSCAPE model is more reliable at small centralities than at large centralities.**

Reliability of hydrodynamics degrades as one moves from the most central to more peripheral collisions (smaller overlap). Also, Grad and Chapman–Enskog particlization schemes perform better in more central collisions (near-equilibrium corrections).

- Consider the kernel: $K(x_i, x_j \mid \phi) \equiv s^2 + \bar{c}^2(x_i x_j)^r \exp\left(-\frac{\|x_i - x_j\|^2}{2\ell^2}\right), \quad (r \geq 0).$

➡ Allows theory errors to be constant or increase with centrality (but not decrease).

➡ Non-vanishing theory error allowed even for ultra-central collisions ($s \neq 0$).

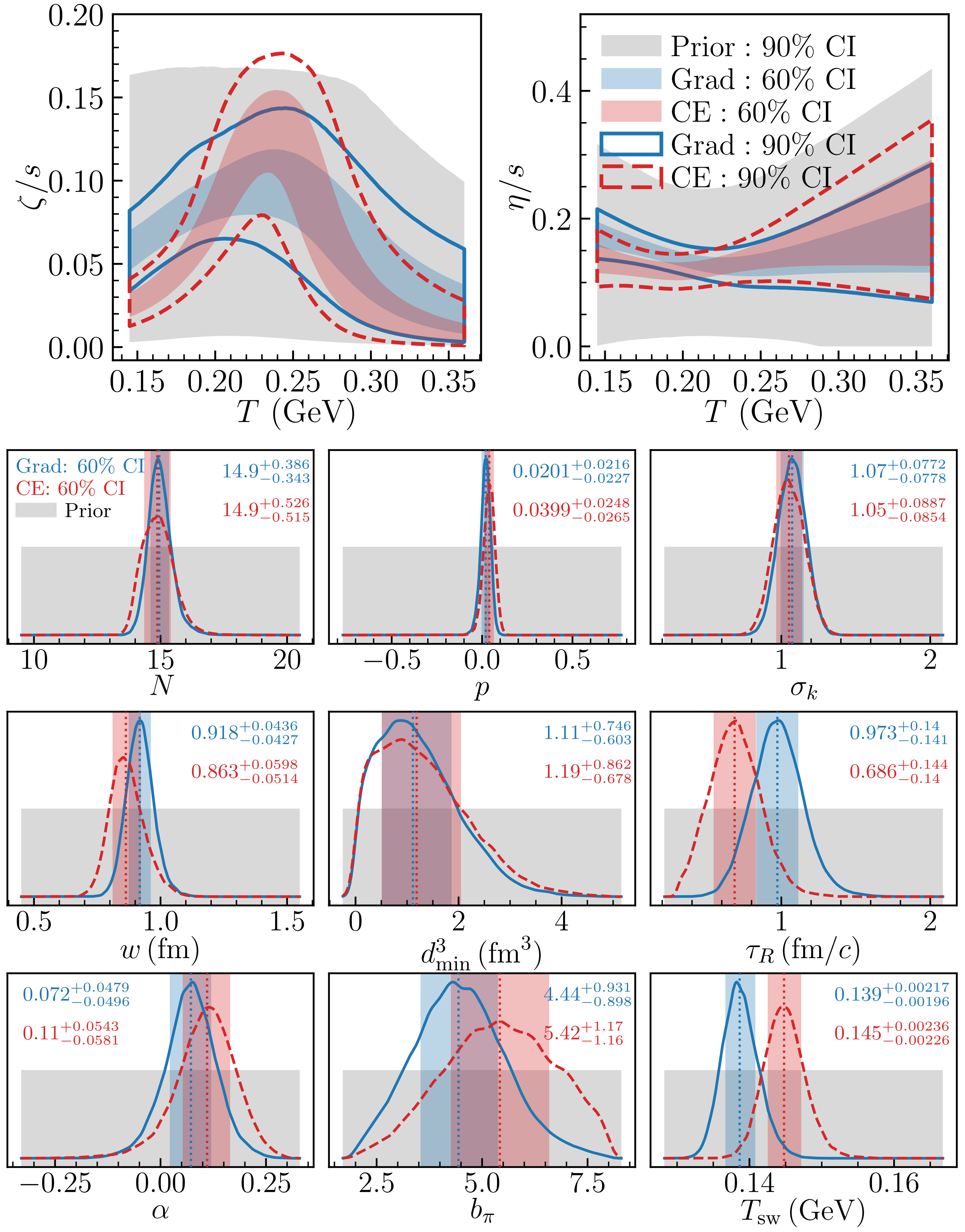
➡ $s = \bar{c} = 0 \implies \delta_{\text{MD}} = 0.$

Bayesian inference

SJ, arXiv: 2509.19759

Not accounting for theory uncertainties (w/o MD)

D. Everett *et al.* 2010.03928, 2011.01430



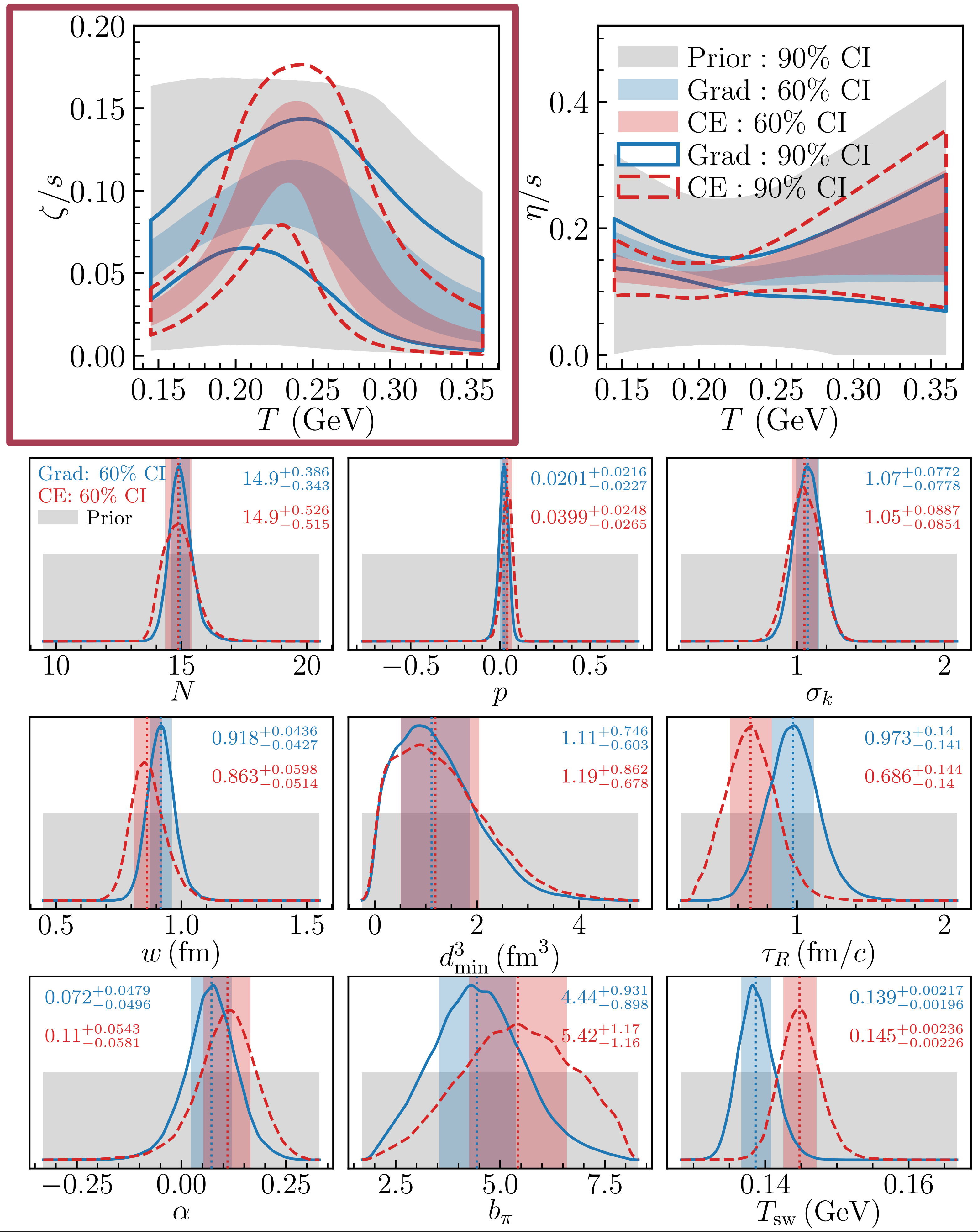
17 parameter Bayesian inference

Bayesian inference

SJ, arXiv: [2509.19759]

Not accounting for theory uncertainties (w/o MD)

D. Everett *et al.* 2010.03928, 2011.01430



17 parameter Bayesian inference

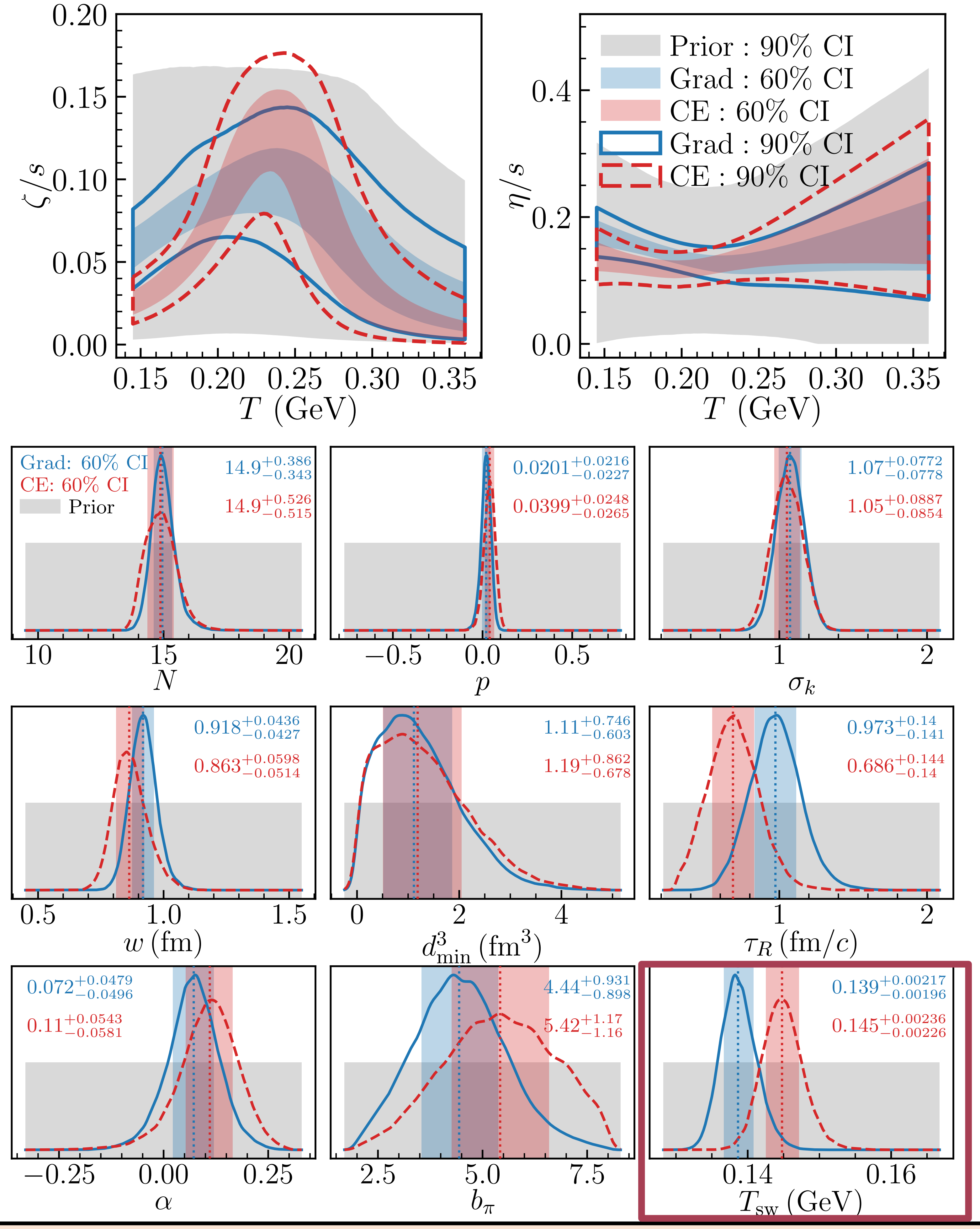
- Tension observed between the specific bulk viscosity posteriors from Grad (blue) and CE (red).

Bayesian inference

SJ, arXiv: 2509.19759

Not accounting for theory uncertainties (w/o MD)

D. Everett *et al.* 2010.03928, 2011.01430



17 parameter Bayesian inference

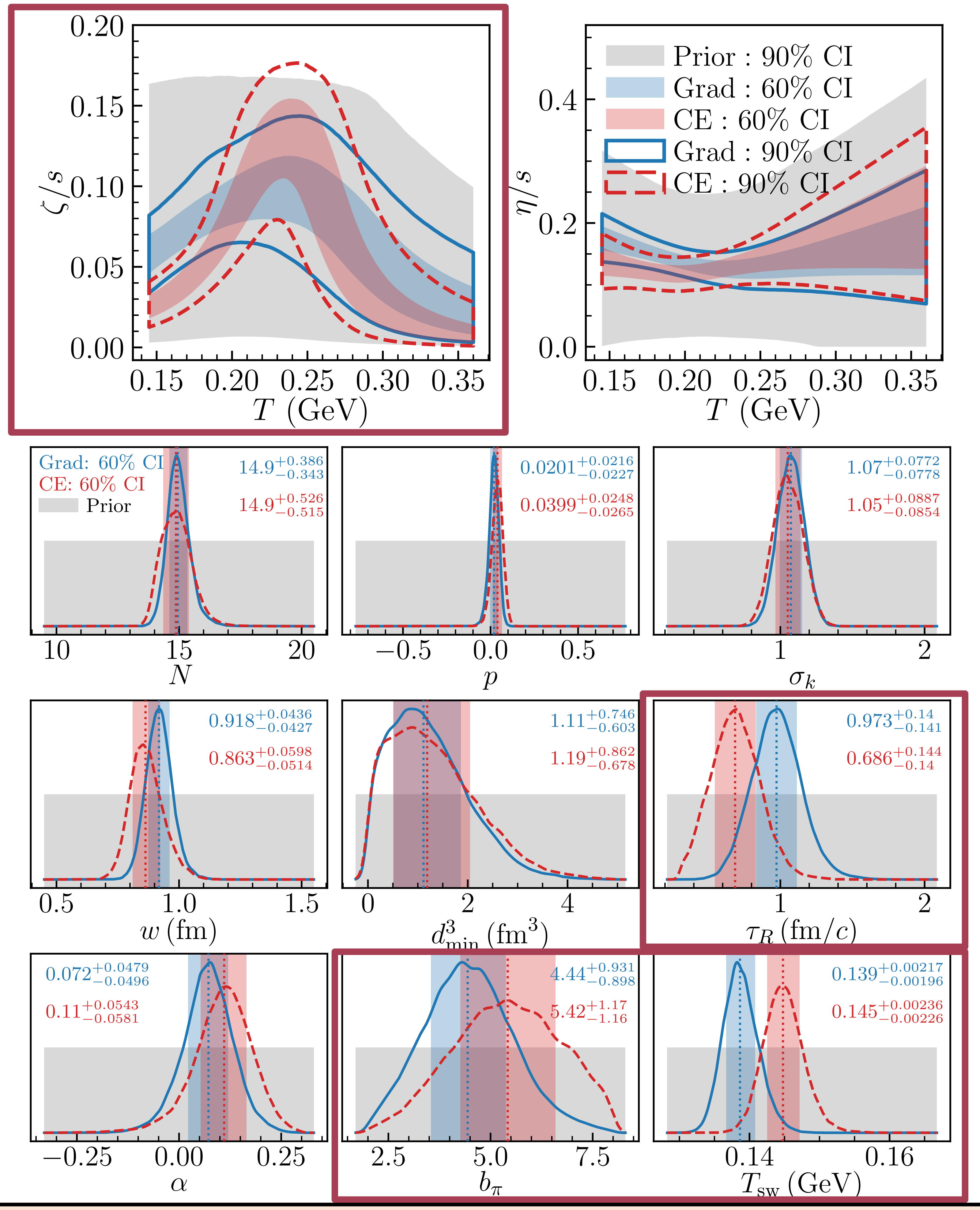
- Tension observed between the specific bulk viscosity posteriors from Grad (blue) and CE (red). Differences also visible for other parameters.
- Particlization temperature lower for Grad: $T_{\rm sw} = 138.65^{+2.17}_{-1.96}$ MeV, than for CE: $144.79^{+2.36}_{-2.26}$ MeV.

Bayesian inference

SJ, arXiv: 2509.19759

D. Everett *et al.* 2010.03928, 2011.01430

Not accounting for theory uncertainties (w/o MD)



17 parameter Bayesian inference

Error of theory propagates across stages. Results in different $T^{\mu\nu}$ evolution. Will effect studies where medium evolution needs to be considered (like jet studies).

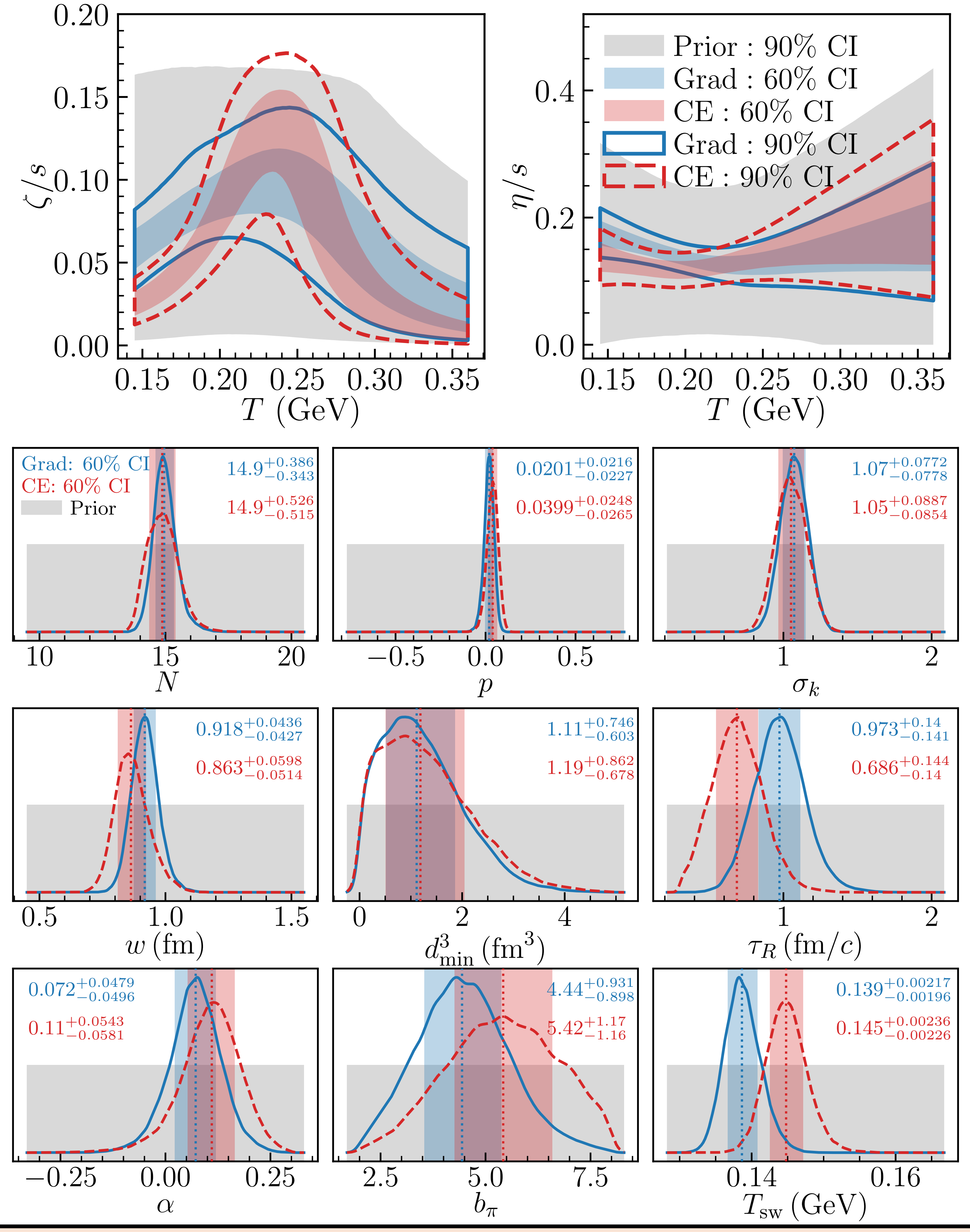
- Tension observed between the specific bulk viscosity posteriors from Grad (blue) and CE (red). Differences also visible for other parameters.
- Particlization temperature lower for Grad: $T_{\mathrm{sw}} = 138.65^{+2.17}_{-1.96}$ MeV, than for CE: $144.79^{+2.36}_{-2.26}$ MeV.

Bayesian inference

SJ, arXiv: 2509.19759

D. Everett *et al.* 2010.03928, 2011.01430

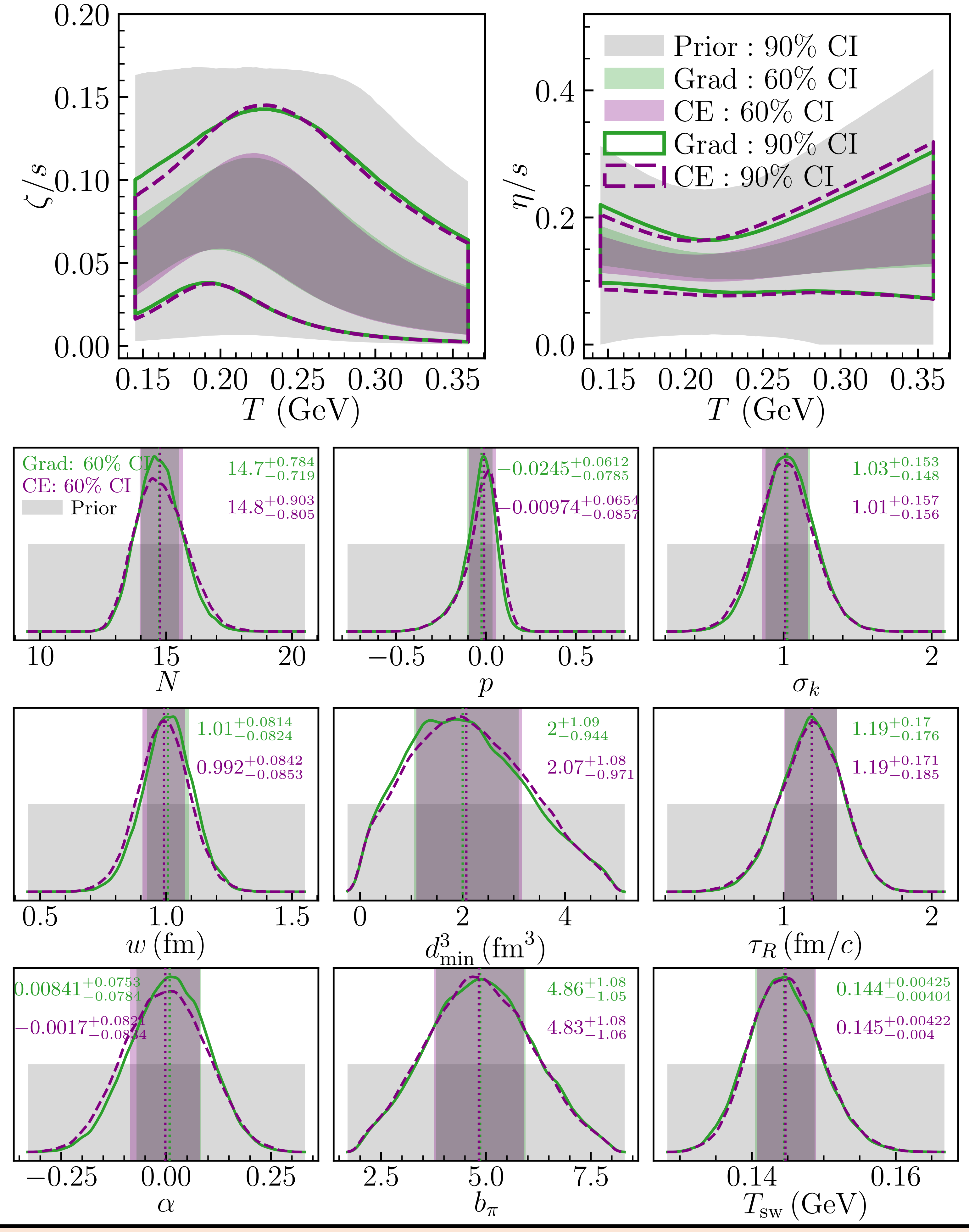
Not accounting for theory uncertainties (w/o MD)



17 parameter Bayesian inference

- Tension observed between the specific bulk viscosity posteriors from Grad (blue) and CE (red). Differences also visible for other parameters.
- Particlization temperature lower for Grad: $T_{\rm sw} = 138.65^{+2.17}_{-1.96}$ MeV, than for CE: $144.79^{+2.36}_{-2.26}$ MeV.

Accounting for theory uncertainties (w/ MD)



65 parameter Bayesian inference: 17 model parameters + 4 GP param per observable (12 observables ⇒ 48 GP param).

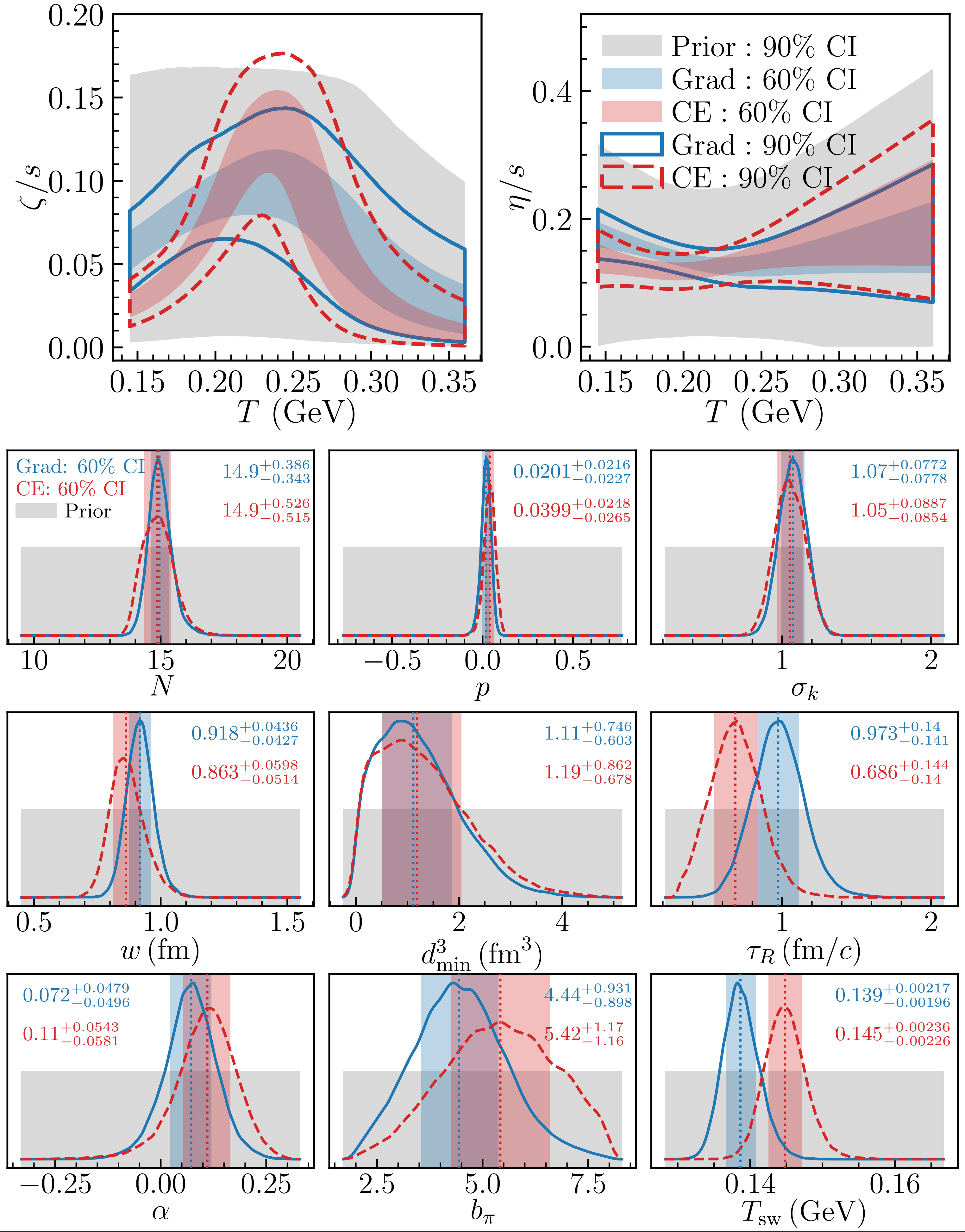
- Posteriors are almost indistinguishable for all parameters for Grad (green) and CE (purple).

Bayesian inference

SJ, arXiv: 2509.19759

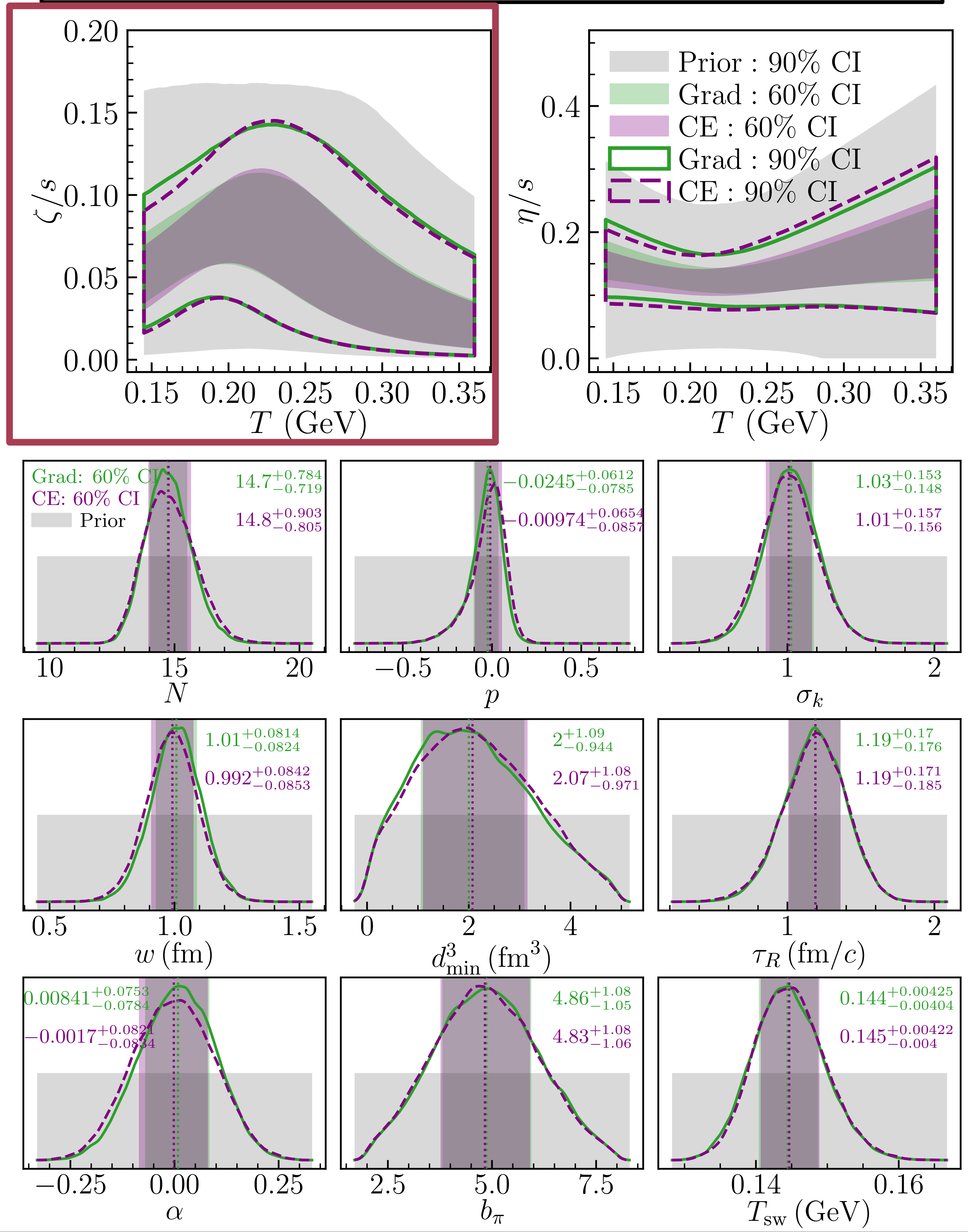
D. Everett *et al.* 2010.03928, 2011.01430

Not accounting for theory uncertainties (w/o MD)



17 parameter Bayesian inference

Accounting for theory uncertainties (w/ MD)



65 parameter Bayesian inference: 17 model parameters + 4 GP param per observable (12 observables ⇒ 48 GP param).

- Tension observed between the specific bulk viscosity posteriors from Grad (blue) and CE (red). Differences also visible for other parameters.
- Particlization temperature lower for Grad: $T_{\rm sw} = 138.65^{+2.17}_{-1.96}$ MeV, than for CE: $144.79^{+2.36}_{-2.26}$ MeV.

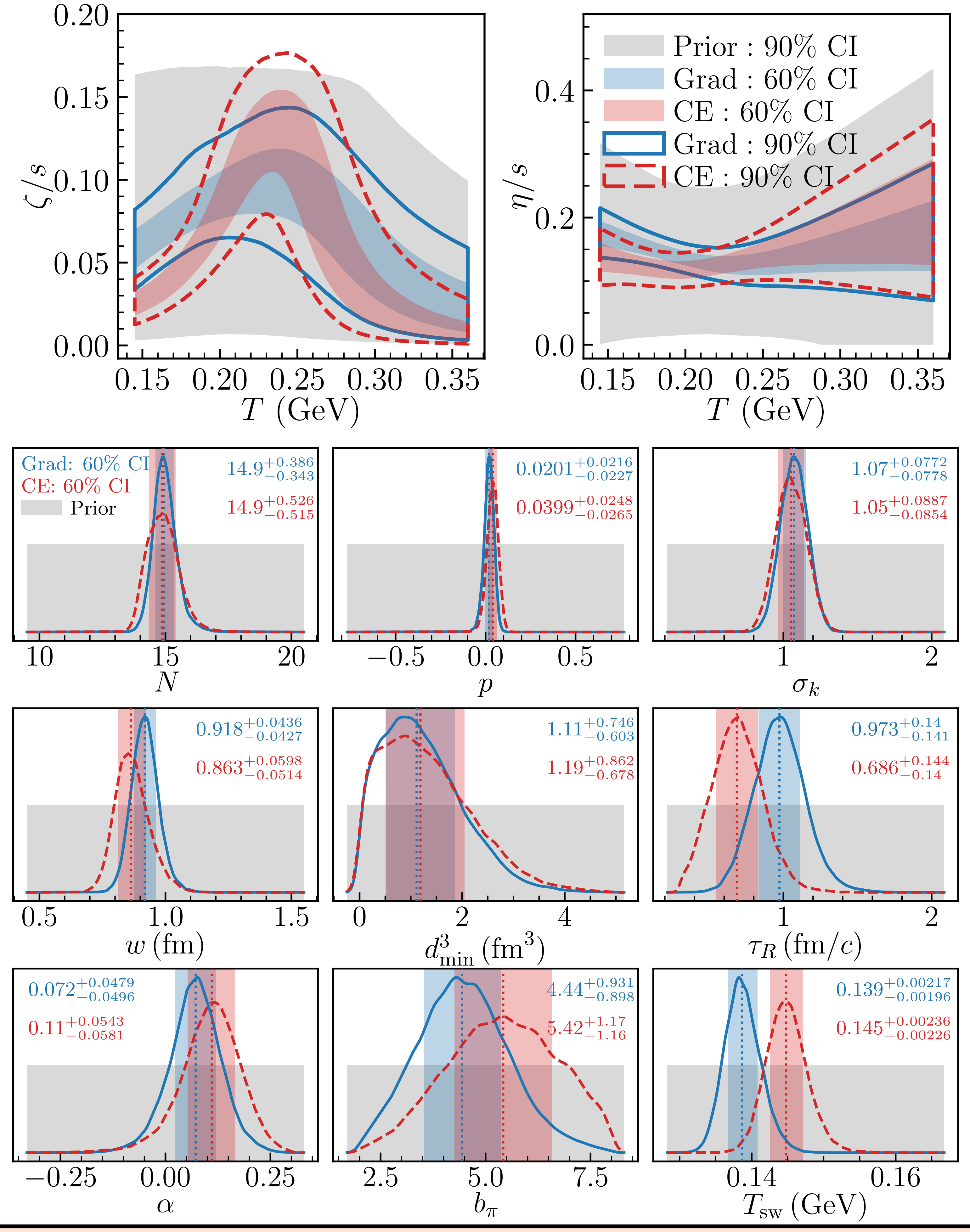
- Posteriors are almost indistinguishable for all parameters for Grad (green) and CE (purple).

Bayesian inference

SJ, arXiv: 2509.19759

D. Everett *et al.* 2010.03928, 2011.01430

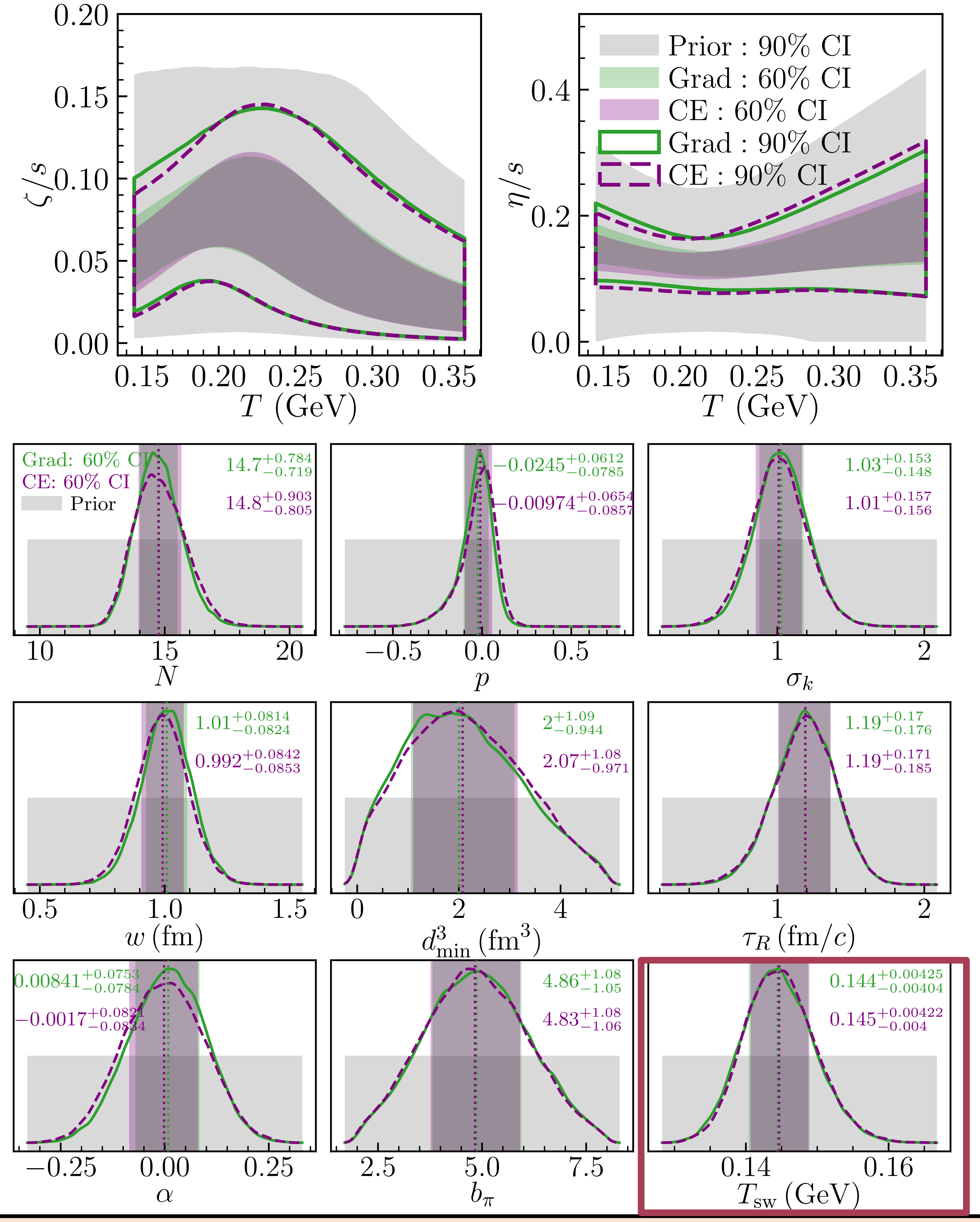
Not accounting for theory uncertainties (w/o MD)



17 parameter Bayesian inference

- Tension observed between the specific bulk viscosity posteriors from Grad (blue) and CE (red). Differences also visible for other parameters.
- Particlization temperature lower for Grad: $T_{\rm sw} = 138.65^{+2.17}_{-1.96}$ MeV, than for CE: $144.79^{+2.36}_{-2.26}$ MeV.

Accounting for theory uncertainties (w/ MD)



65 parameter Bayesian inference: 17 model parameters + 4 GP param per observable (12 observables ⇒ 48 GP param).

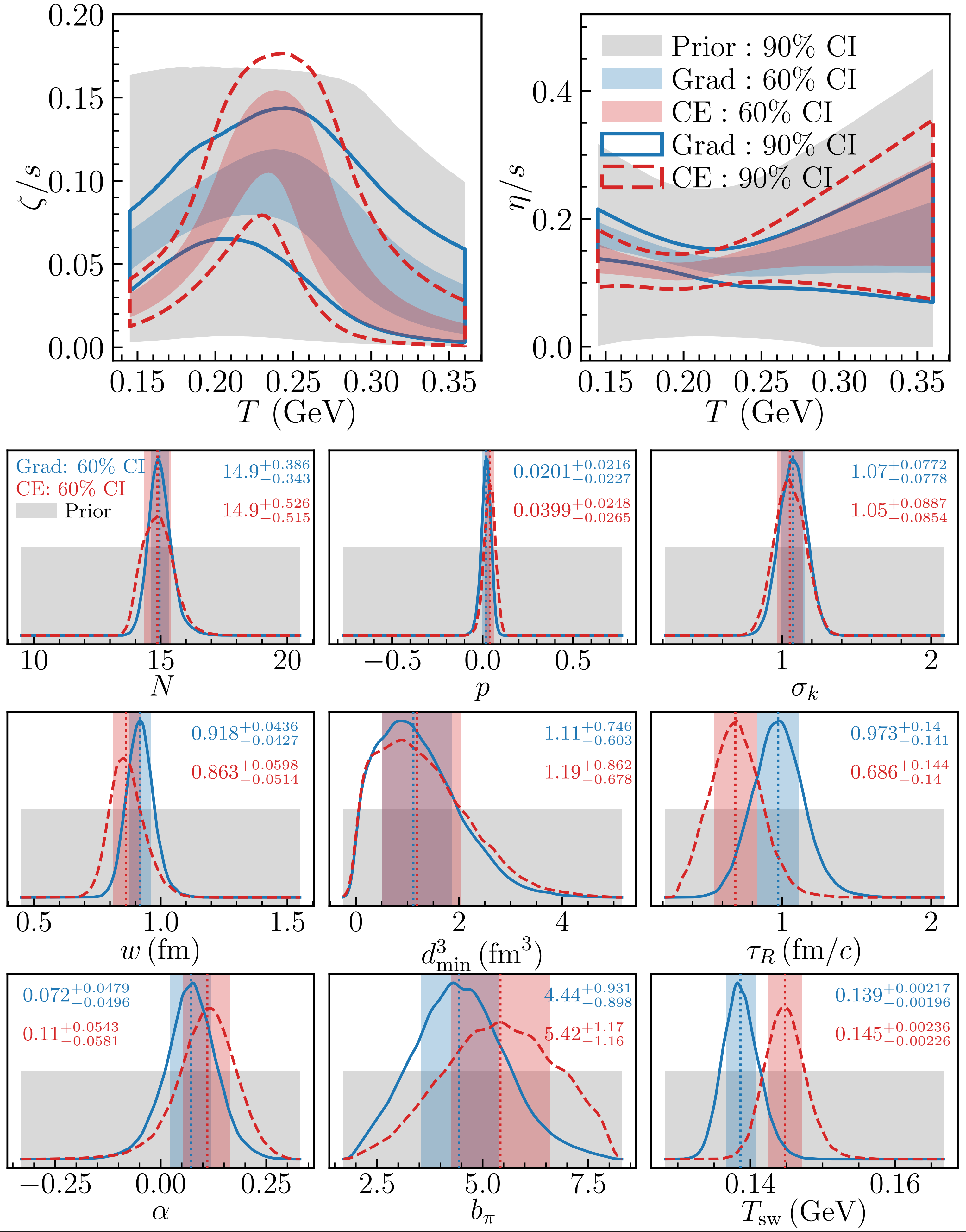
- Posteriors are almost indistinguishable for all parameters for Grad (green) and CE (purple).
- Particlization temperature consistent: $T_{\rm sw} = 144.46^{+4.25}_{-4.04}$ MeV (Grad) and $T_{\rm sw} = 144.66^{+4.22}_{-4.00}$ MeV (CE).

Bayesian inference

SJ, arXiv: 2509.19759

D. Everett *et al.* 2010.03928, 2011.01430

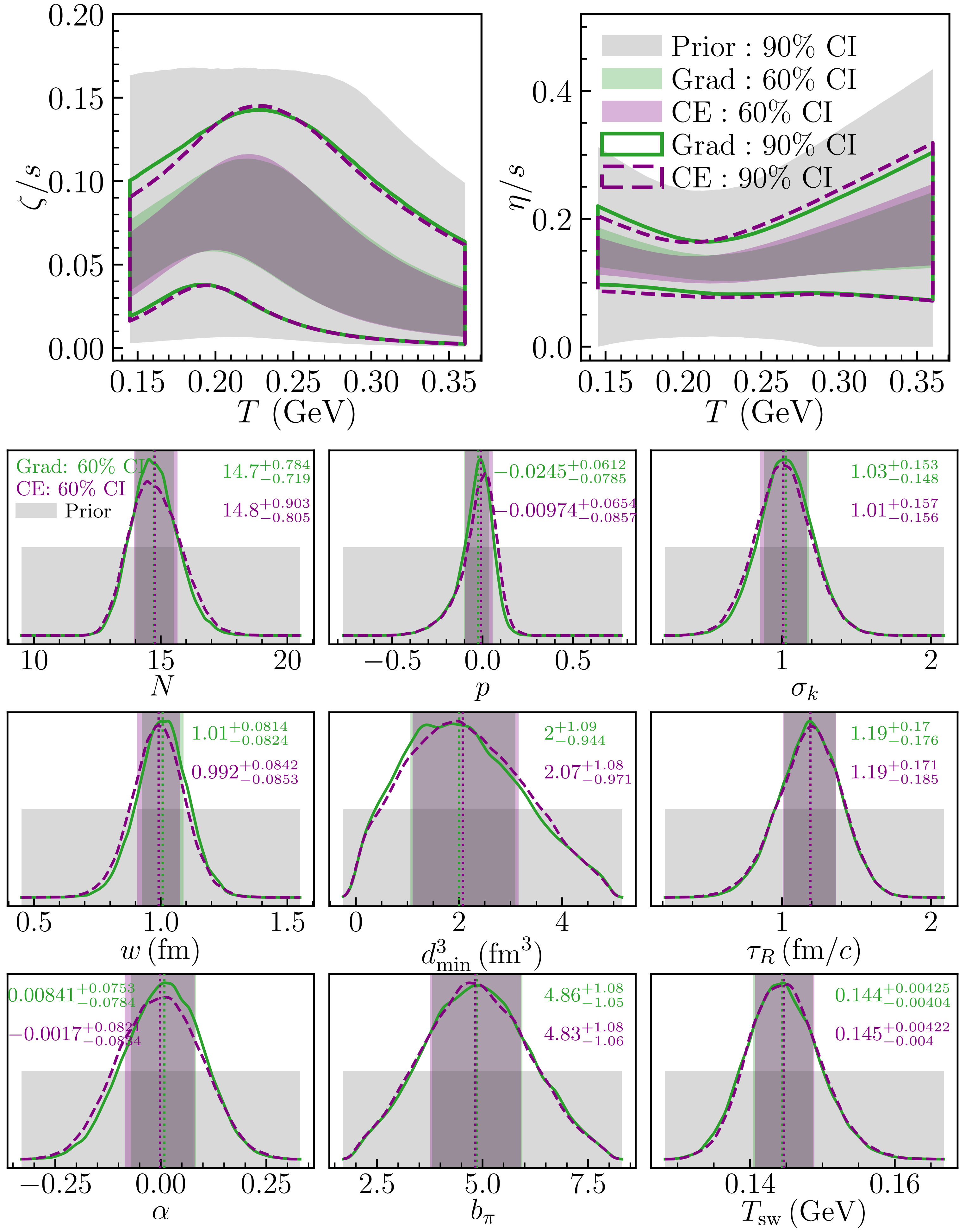
Not accounting for theory uncertainties (w/o MD)



17 parameter Bayesian inference

Error of theory is quantified and does not propagate across stages => same $T^{\mu\nu}$ evolution for Grad and CE.

Accounting for theory uncertainties (w/ MD)



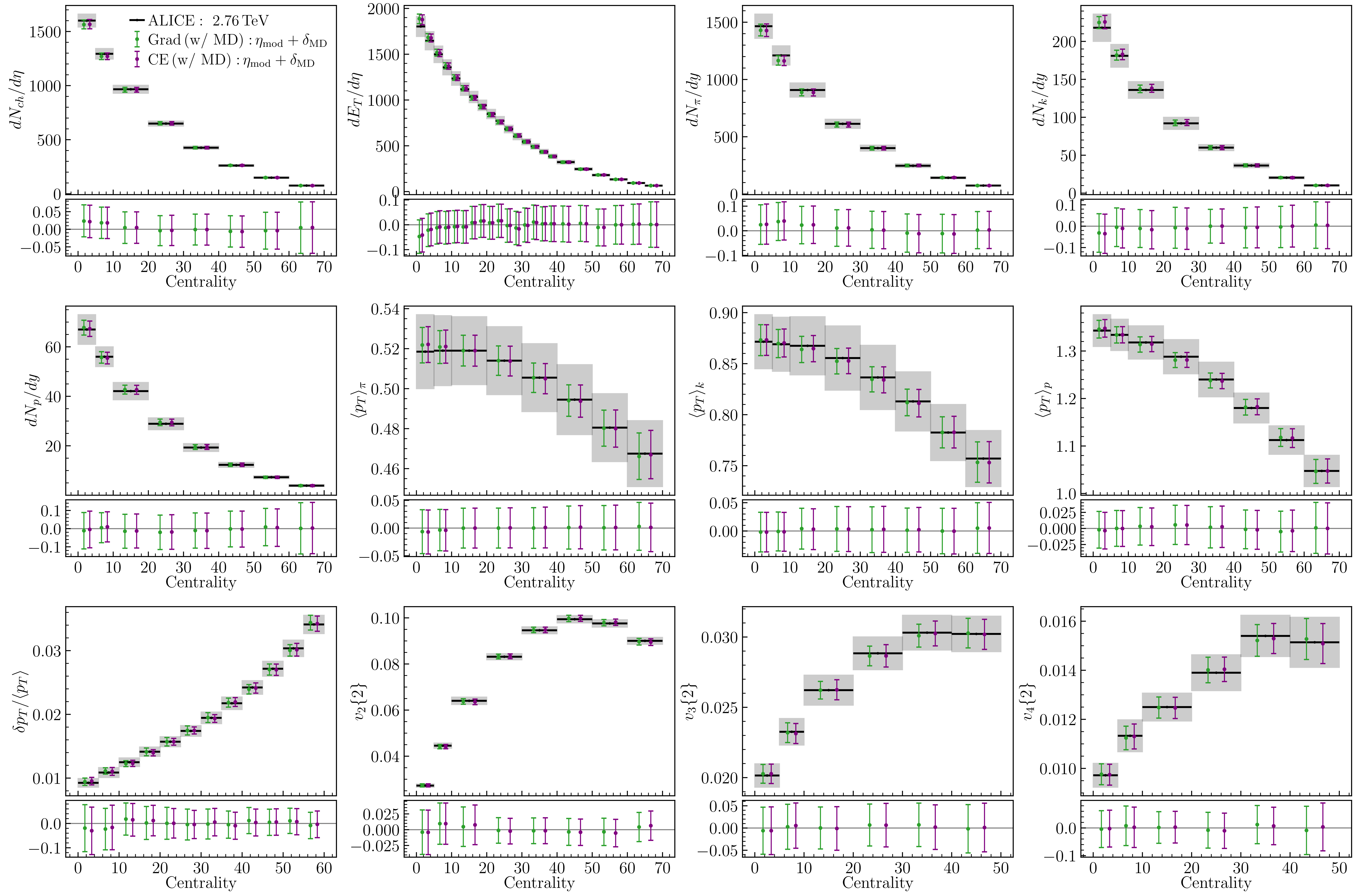
65 parameter Bayesian inference: 17 model parameters + 4 GP param per observable (12 observables ⇒ 48 GP param).

- Tension observed between the specific bulk viscosity posteriors from Grad (blue) and CE (red). Differences also visible for other parameters.
- Particlization temperature lower for Grad: $T_{\rm sw} = 138.65^{+2.17}_{-1.96}$ MeV, than for CE: $144.79^{+2.36}_{-2.26}$ MeV.

- Posteriors are almost indistinguishable for all parameters for Grad (green) and CE (purple).
- Particlization temperature consistent: $T_{\rm sw} = 144.46^{+4.25}_{-4.04}$ MeV (Grad) and $T_{\rm sw} = 144.66^{+4.22}_{-4.00}$ MeV (CE).

Model + discrepancy predictions

SJ, arXiv: 2509.19759



normalized error $(y_{\rm exp} - \eta_{\rm mod} - \delta_{\rm MD})/|\langle y_{\rm exp} \rangle|$

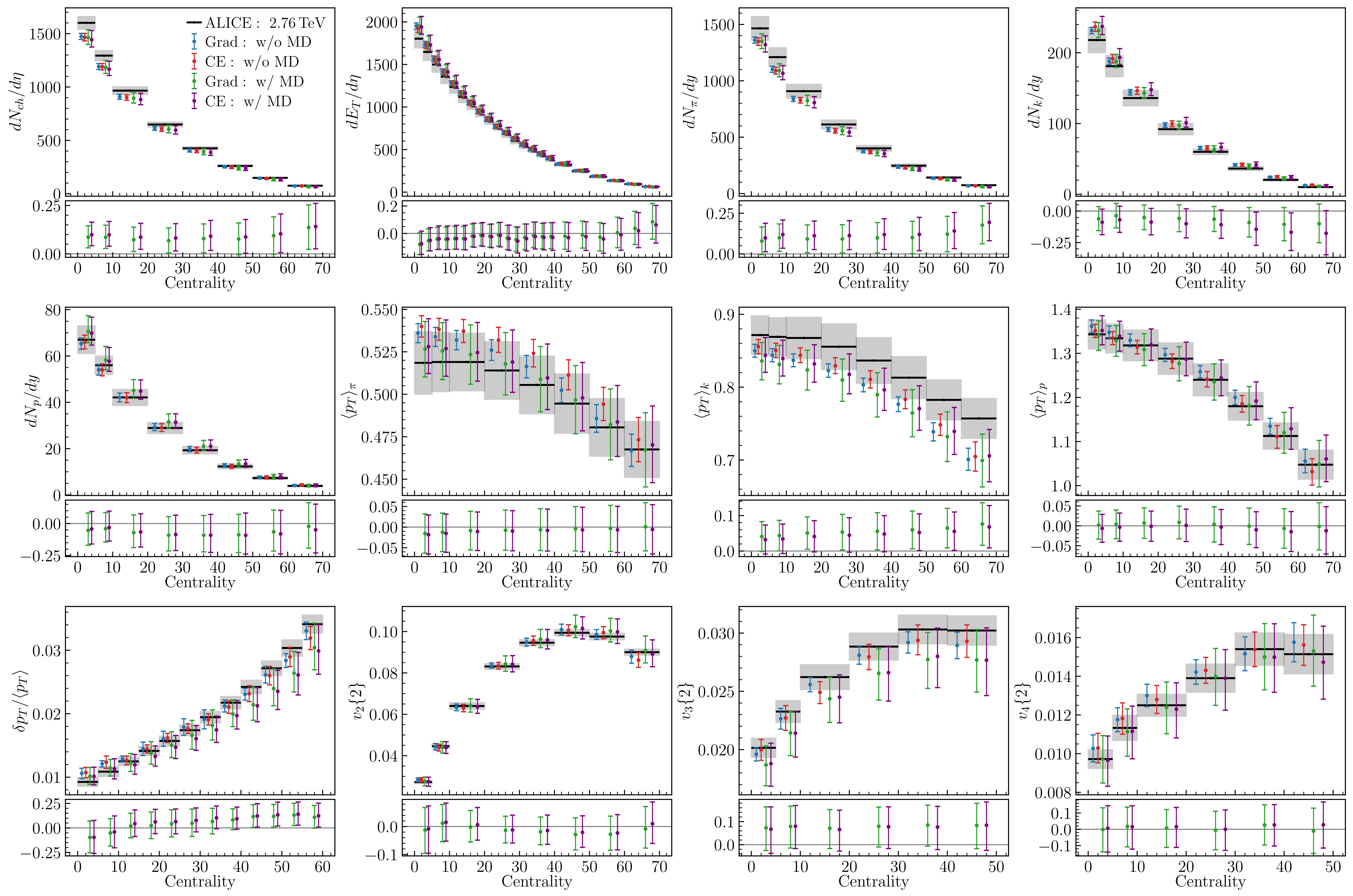
Predictions from

$\xi(x; \theta, \phi) \equiv \eta_{\rm mod}(x, \theta) + \delta_{\rm MD}(x, \phi)$

Excellent agreement with data.
Expected.

Model predictions

SJ, arXiv: 2509.19759

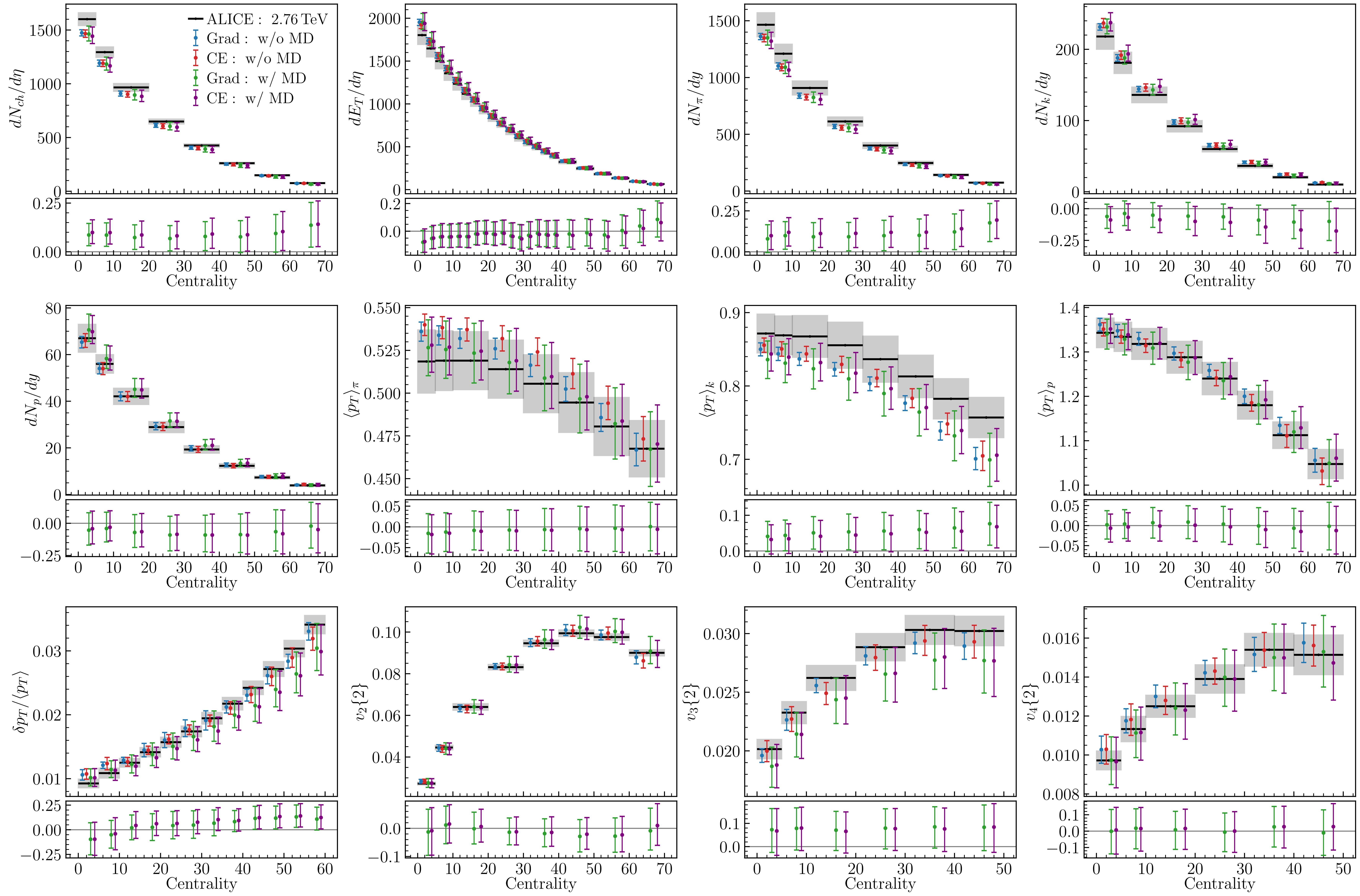


Predictions from $\eta_{\text{mod}}(x, \theta)$.

normalized discrepancies $\underbrace{(y_{\text{exp}} - \eta_{\text{mod}})/|\langle y_{\text{exp}}\rangle|}_{\approx\, \delta_{\text{MD}}(x,\phi)}$

Model predictions

SJ, arXiv: [2509.19759](https://arxiv.org/abs/2509.19759)



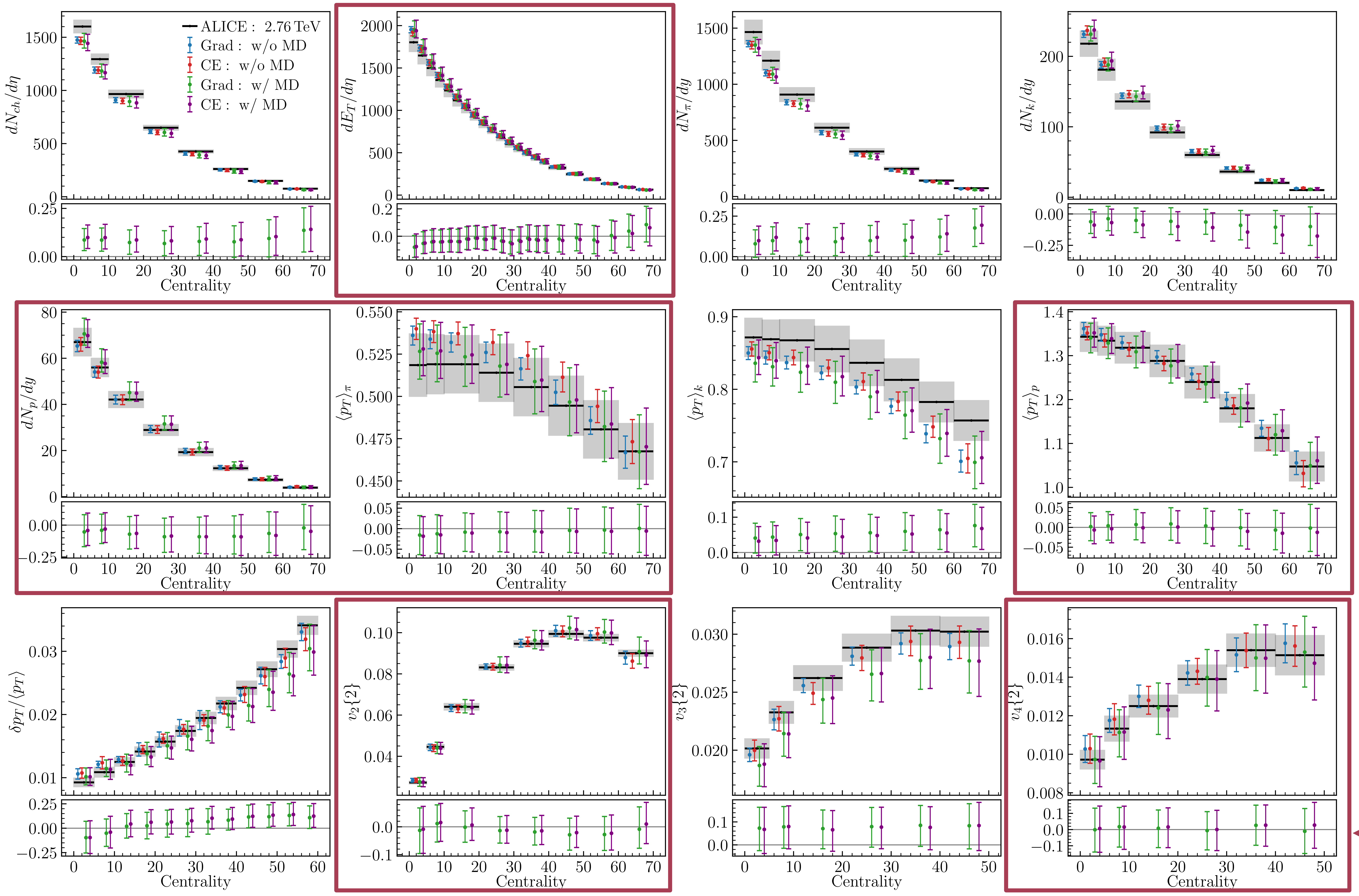
normalized discrepancies $\underbrace{(y_{\text{exp}} - \eta_{\text{mod}})/|\langle y_{\text{exp}} \rangle|}_{\approx\, \delta_{\text{MD}}(x,\phi)}$

Predictions from $\eta_{\text{mod}}(x, \theta)$.

- No clear preference for Grad or CE is observed.

Model predictions

SJ, arXiv: 2509.19759



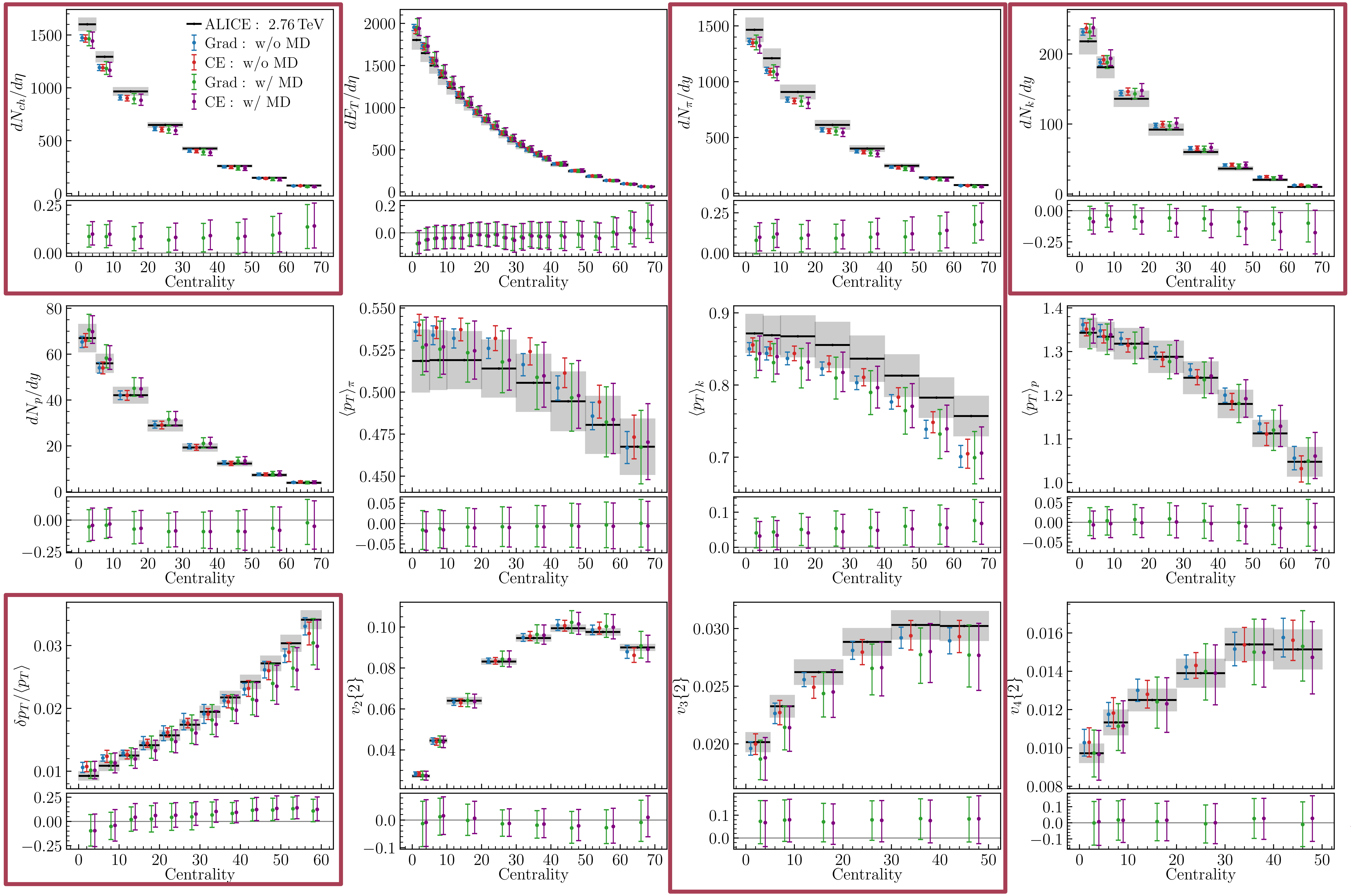
Predictions from $\eta_{\mathrm{mod}}(x, \theta)$.

- No clear preference for Grad or CE is observed.
- Theory predicts many observables correctly => learned error are small for these.

normalized discrepancies $\underbrace{(y_{\mathrm{exp}} - \eta_{\mathrm{mod}})/|\langle y_{\mathrm{exp}} \rangle|}_{\approx\, \delta_{\mathrm{MD}}(x, \phi)}$

Model predictions

SJ, arXiv: 2509.19759



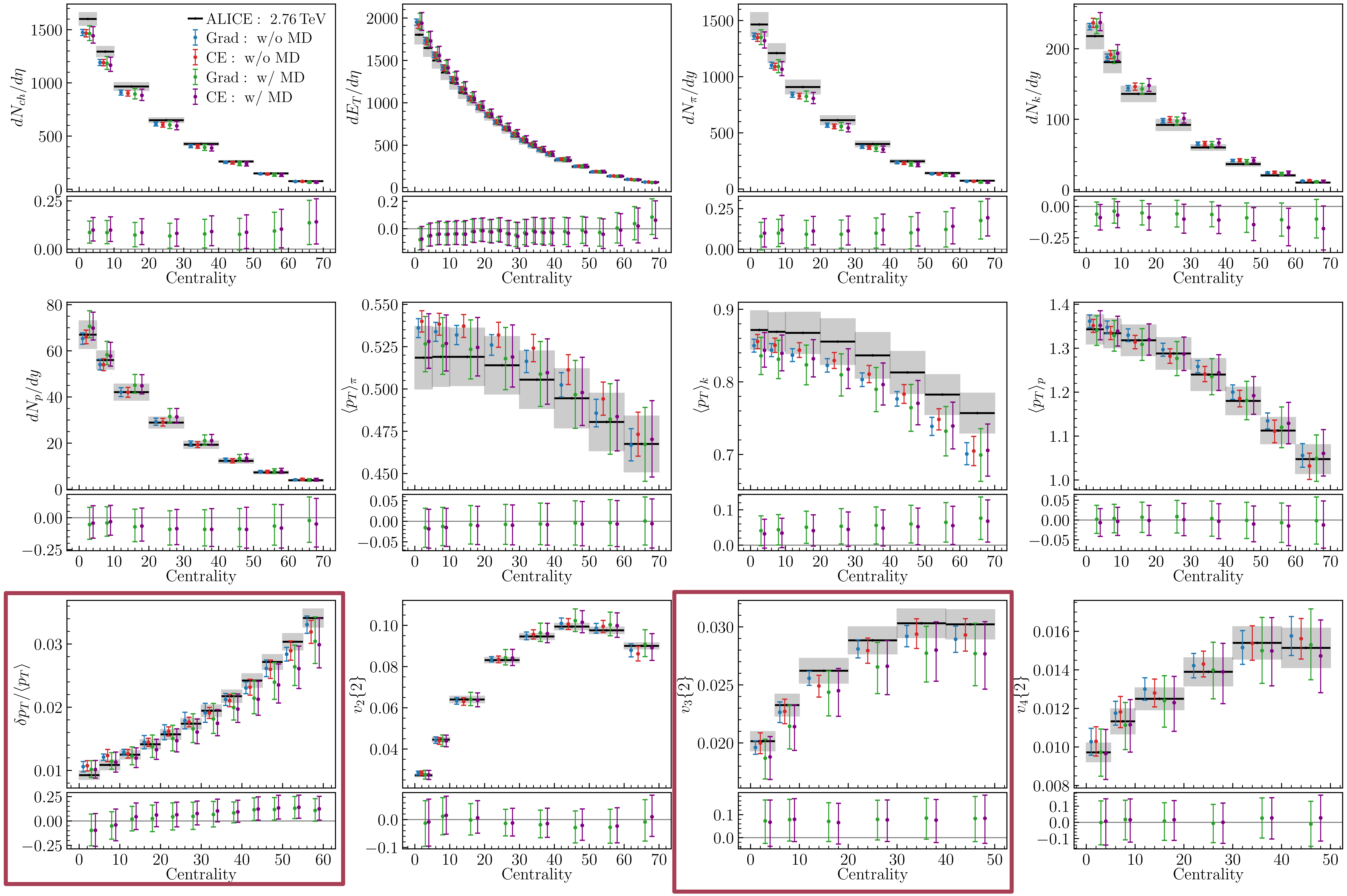
Predictions from $\eta_{\text{mod}}(x, \theta)$.

- No clear preference for Grad or CE is observed.
- Theory predicts many observables correctly => learned error are small for these.
- Deviations observed in w/ MD predictions => learned theory error are large.

normalized discrepancies $\underbrace{(y_{\text{exp}} - \eta_{\text{mod}})/|\langle y_{\text{exp}}\rangle|}_{\approx\, \delta_{\text{MD}}(x,\phi)}$

Model predictions

SJ, arXiv: 2509.19759



Predictions from $\eta_{\text{mod}}(x, \theta)$.

- No clear preference for Grad or CE is observed.
- Theory predicts many observables correctly => learned error are small for these.
- Deviations observed in w/ MD predictions => learned theory error are large.
- Particularly large deviations in $v_3\{2\}$ and $\delta p_T/\langle p_T \rangle$ suggest deficiencies in the initial-condition model.

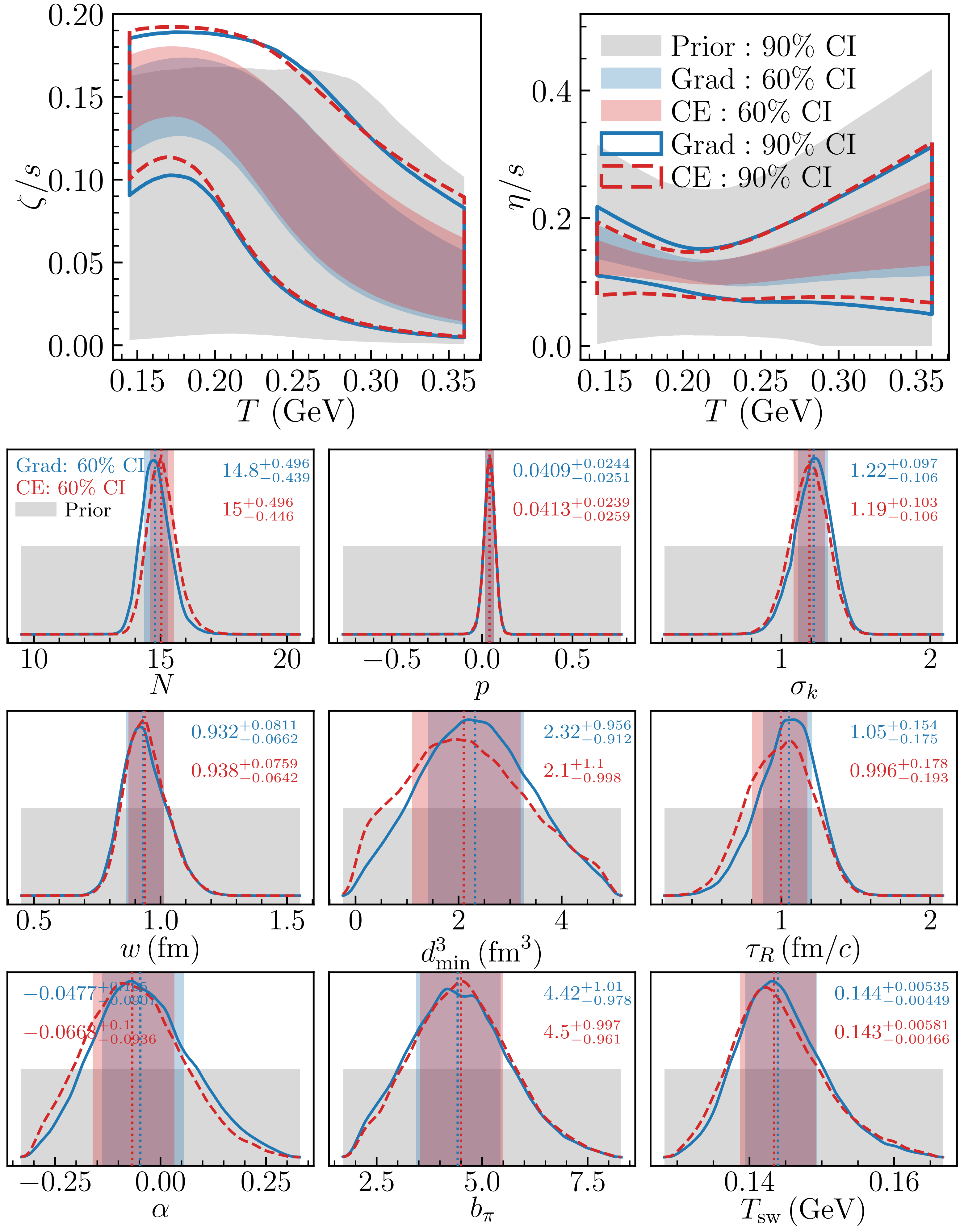
Guides understanding for targeted improvement in models.

normalized discrepancies $\underbrace{(y_{\text{exp}} - \eta_{\text{mod}})/|\langle y_{\text{exp}} \rangle|}_{\approx\, \delta_{\text{MD}}(x, \phi)}$

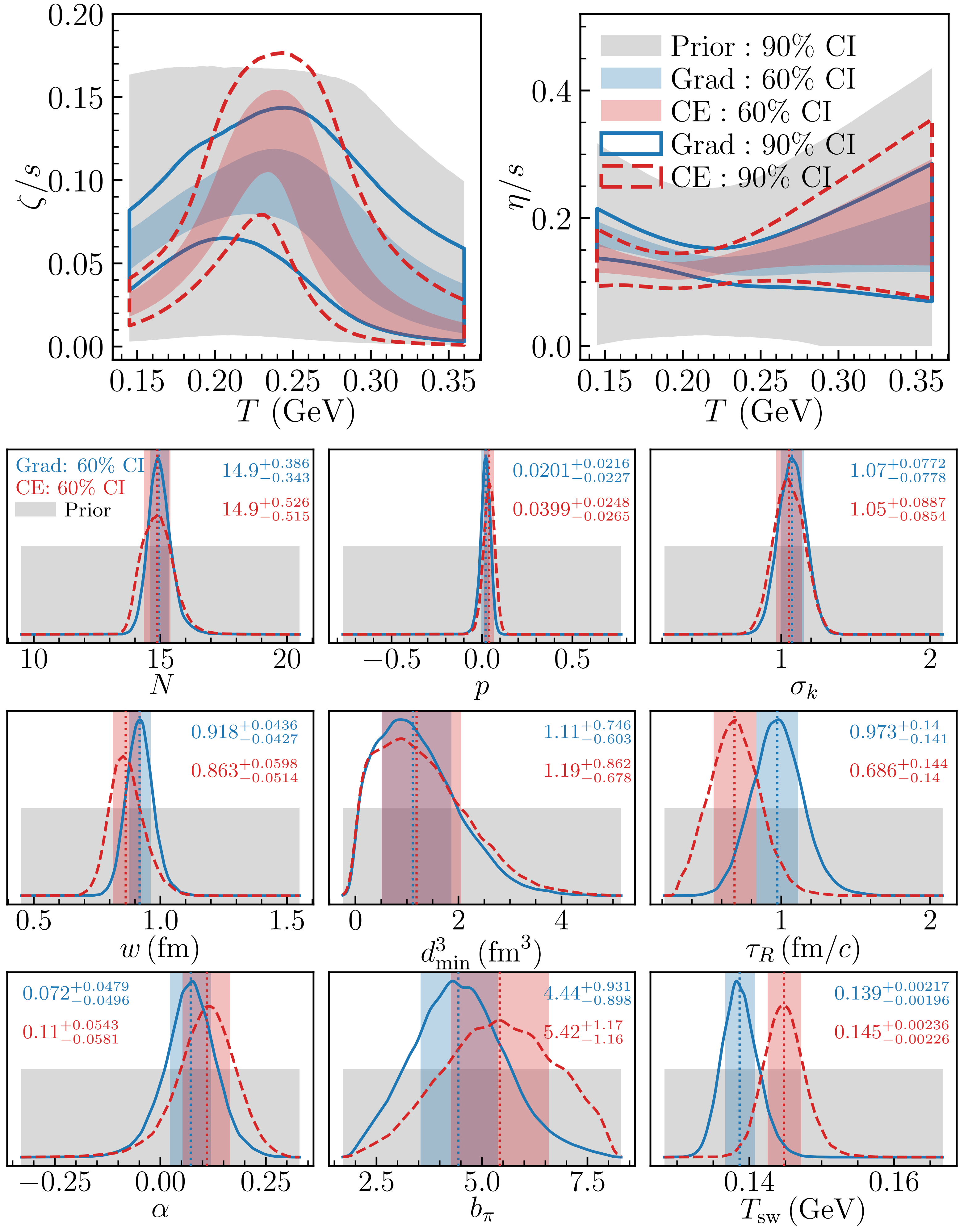
Check (w/o MD): Stability of posteriors

SJ, arXiv: 2509.19759

Calibration with subset of observables:
$dN_{\mathrm{ch}}/d\eta$, $dE_T/d\eta$, and $v_n\{2\}$ $(n = 2,3,4)$



Calibration with all observables



w/o MD

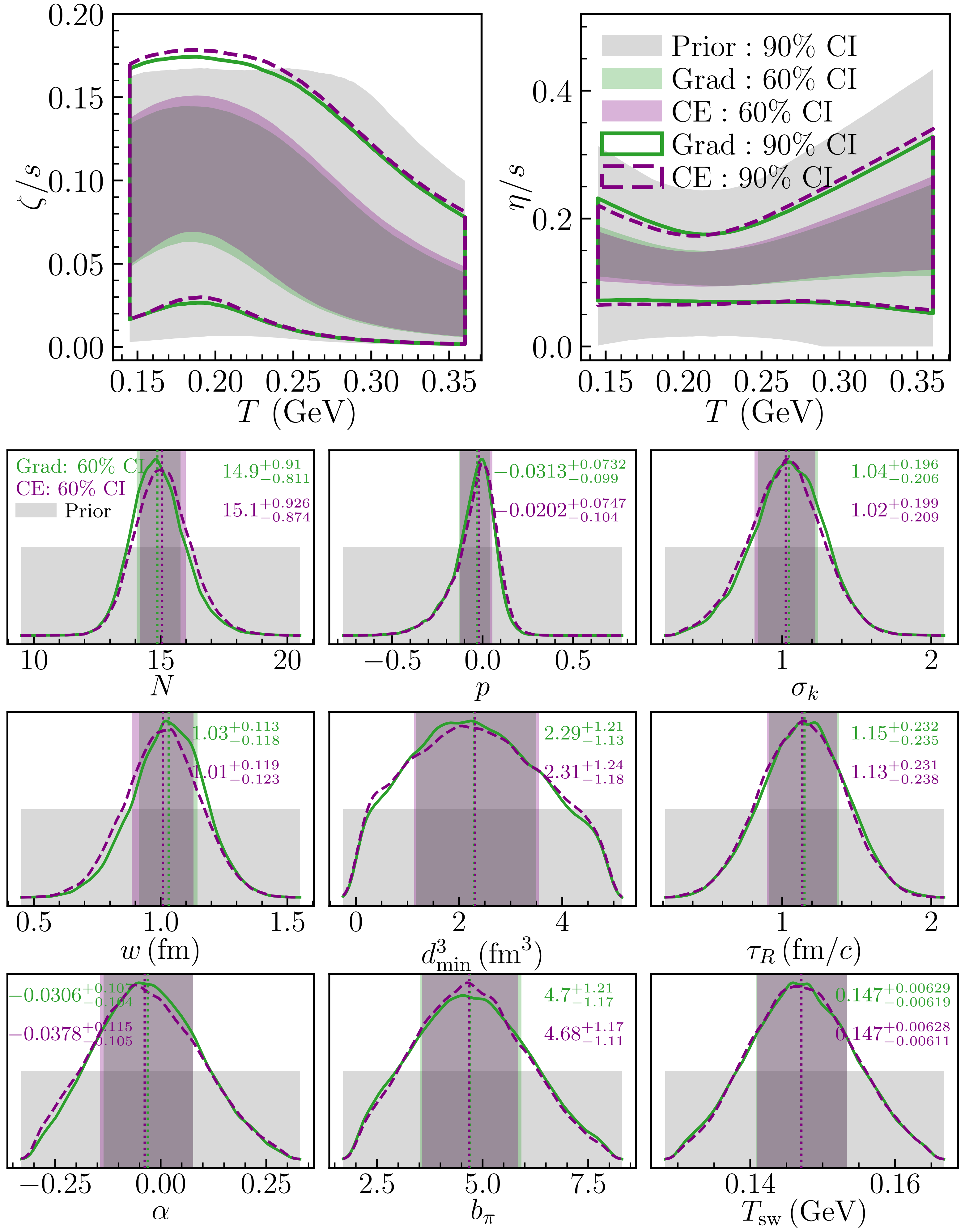
Posteriors for many parameters shift significantly when all measurements are included in the inference, indicating lack of robustness of the parameter estimate.

Check (w/ MD): stability of posteriors

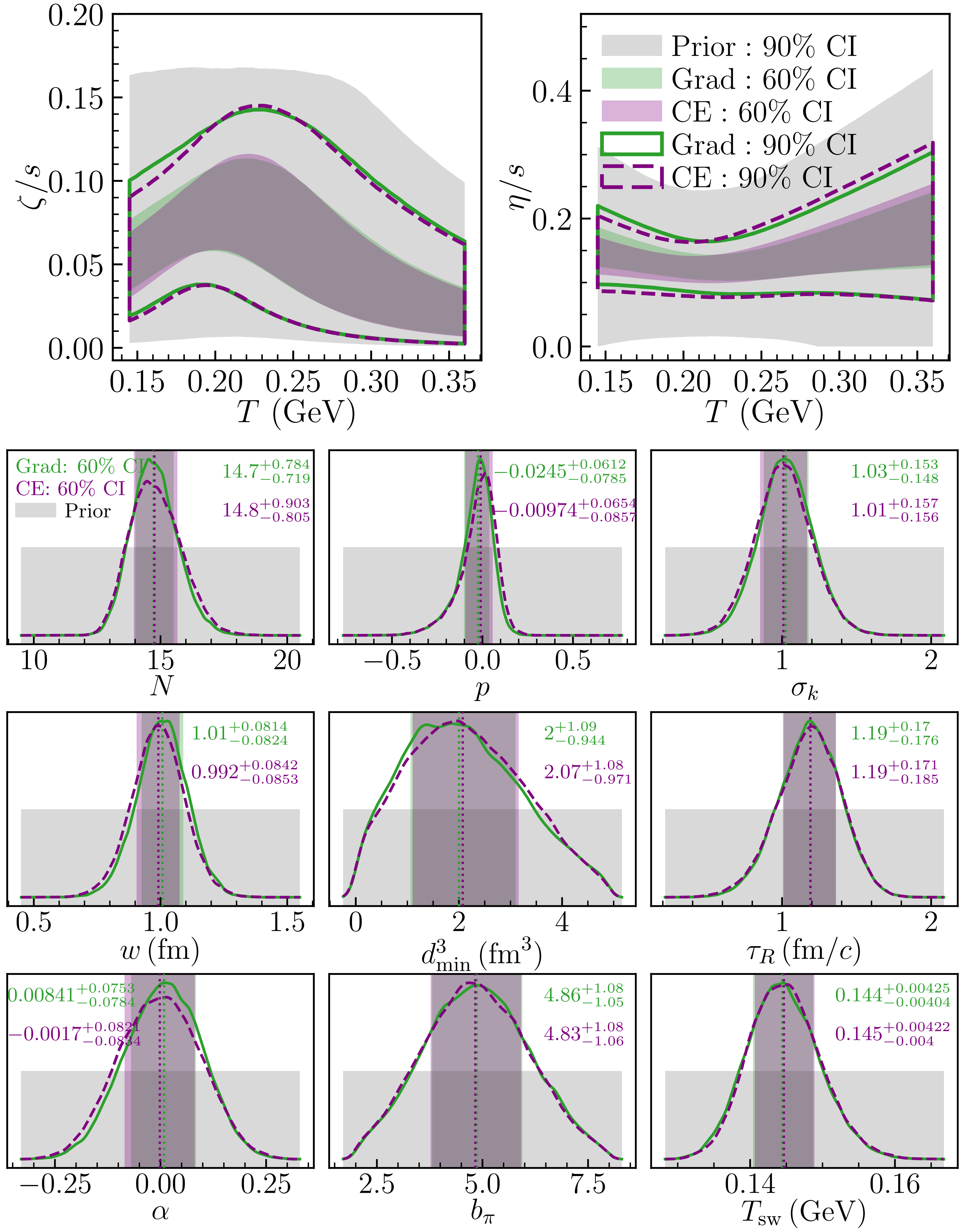
SJ, arXiv: 2509.19759

Calibration with subset of observables: $dN_{\mathrm{ch}}/d\eta$, $dE_T/d\eta$, and $v_n\{2\}$ $(n = 2,3,4)$

Calibration with all observables



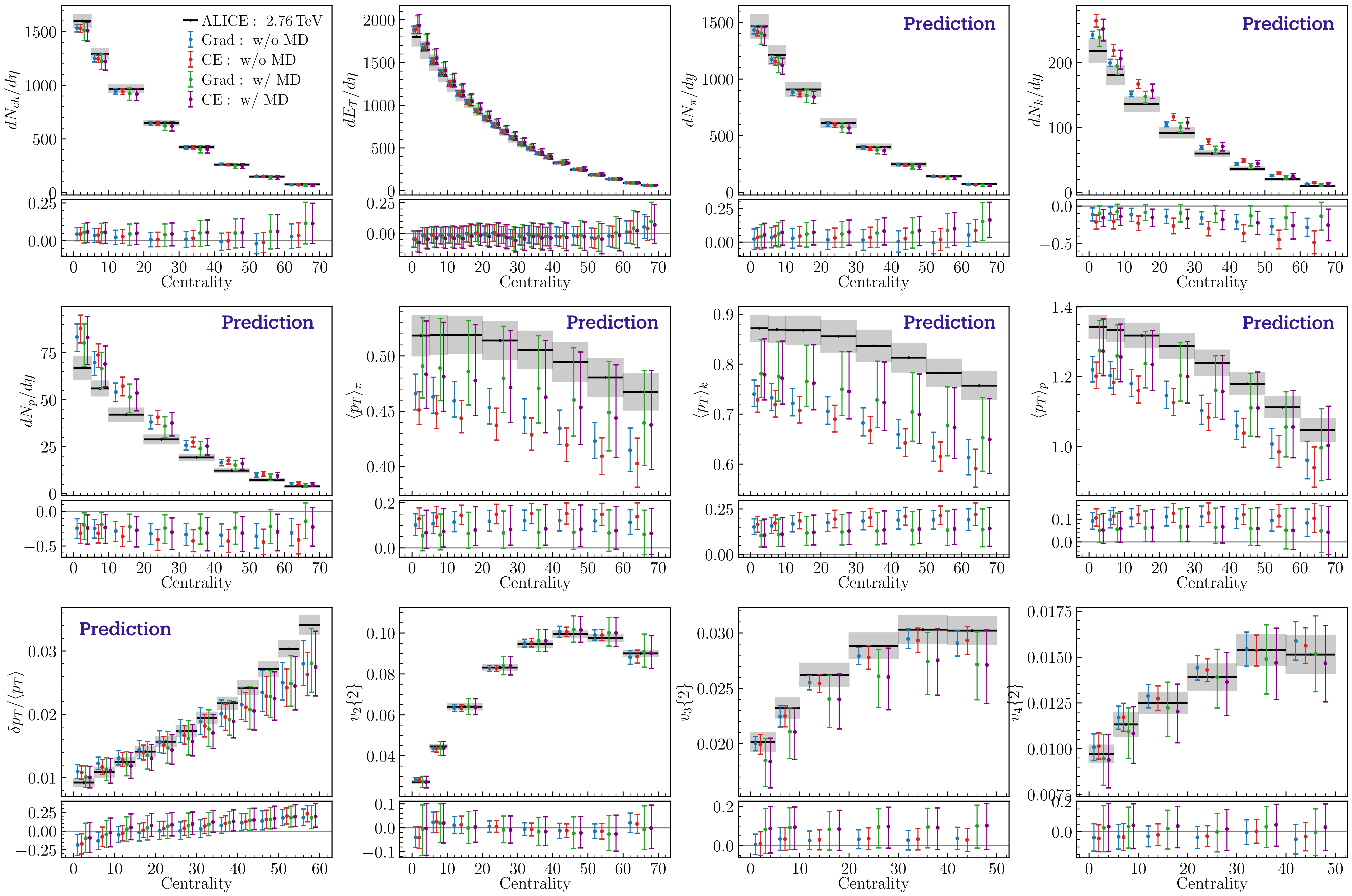
w/ MD



Posteriors tighten without shifting, indicating robust parameter inference.

Model predictions: calibrated on subset of observables

SJ, arXiv: 2509.19759



Predictions from $\eta_{\text{mod}}(x, \theta)$.

- Observables not used in calibration are labeled "Prediction".
- w/o MD predictions for $\langle p_T \rangle_i,\ i \in \{\pi, k, p\}$ are outside 68 % CI prediction interval, reflecting overfitting.
- w/ MD results show comparatively better agreement with data, albeit with large error bars.

normalized discrepancies $\underbrace{(y_{\text{exp}} - \eta_{\text{mod}})}_{\approx\, \delta_{\text{MD}}(x,\phi)}/|\langle y_{\text{exp}} \rangle|$

Outline

- Model discrepancy framework
- Ball drop experiment with model discrepancy
- Real world application: Heavy-ion collisions
  - ➡ Applying model discrepancy framework to obtain constraints on transport properties of QGP
- Summary

Summary

- Often the goal of model-data comparison is to extract physically meaningful model parameters, **not to fit data.**
- Quantifying theoretical uncertainties is **crucial** for correct inference of physical model parameters.
  - ➡ Implicit assumption of the model being universally valid is usually made in model-to-data comparison. This can lead to incorrect and misleading parameter estimates.
- Discussed a Bayesian framework that explicitly quantifies theoretical uncertainties by statistically modeling theory errors, guided by qualitative knowledge of the theory's domain of reliability.
- Derived robust, data-driven constraints on transport properties of QCD with quantified theory uncertainties.
- Framework is general and can be applied to various problems in science.
  Open source code is provided: https://github.com/sjaiswal-tifr/ModelDiscrepancy

Thank you!