Asymptotic distributions of the PLR in EFTs

SciPost Phys. Core 6, 013 (2023)

F. Bernlochner, D. Fry, S. Menary, E. Persson

stephen.menary@manchester.ac.uk

Introduction

1. EFTs

2. PLR and Wilks theorem

3. The problem: Wilks theorem sometimes does not apply to EFTs

4. Intuitions about how this happens, the consequences, and some solvable cases

5. Wrap-up

EFTs

- **EFTs** parameterise the **low-energy contributions** of unknown physics at mass scales higher than the experiment, without assuming *any specific* UV-complete model
  1. Reject SM: search for new physics and discover it's low-energy behaviour
  2. Constrain EFT param space: constrain specific theories post-hoc by their leading-order terms

$$\mathscr{L}_{\mathrm{EFT}} = \mathscr{L}_{\mathrm{SM}} + \sum_i^{(\mathrm{D5})} \frac{c_i}{\Lambda_i}\mathscr{O}_i + \sum_i^{(\mathrm{D6})} \frac{c_i}{\Lambda_i^2}\mathscr{O}_i + \dots$$

→ Feynman diagrams → Matrix elements $\mathscr{M} = \mathscr{M}_{\mathrm{SM}} + \sum_i c_i \mathscr{M}_i + \dots$

→ Cross-sections $\sigma \sim |\mathscr{M}|^2 = |\mathscr{M}_{\mathrm{SM}}|^2 + \sum_i c_i \mathscr{M}^*_{\mathrm{SM}}\mathscr{M}_i + \sum_i c_i^2 |\mathscr{M}_i|^2 + \sum_{ij} c_i c_j \mathscr{M}^*_i \mathscr{M}_j + \dots$

$$\sigma = \underbrace{\sigma^{\mathrm{SM}}}_{\text{SM}} + \underbrace{\sum_i c_i \sigma^{\mathrm{int}}}_{\text{Linear in } c_i} + \underbrace{\sum_i c_i^2 \sigma_i^{\mathrm{new}}}_{\text{Quadratic in } c_i} + \underbrace{\sum_{ij} c_i c_j \sigma_{ij}^{\mathrm{int}}}_{\text{Bilinear in } c_i,\, c_j} + \dots$$

Higher order perturbations

PLR

- N_c EFT params $c \in \mathbb{R}^{N_c}$
- N_x-bin diff cross-sections $x \in \mathbb{R}^{N_x}$

- **PLR test-statistic**

$$q_c = -2\log \frac{\mathscr{L}(x\,|\,c)}{\mathscr{L}(x\,|\,c_{\text{MLE}})} \qquad c_{\text{MLE}} = \arg\max_{c'} \mathscr{L}(c';x) \qquad CL(q_c^{\text{obs}}) = \int_{q_c^{\text{obs}}}^{\infty} p\left(q_c;c\right)\,\mathrm{d}q$$

  - To calculate frequentist CL, we must know $p(q_c)$ when c is true, for every c
  - **Wilks theorem**: p is a χ^2-distribution *under certain regularity conditions*

- Assume $\mathscr{L}(x\,|\,c)$ is **Gaussian** with **mean** $\mu(c)$ and **stat+syst covariance** Σ

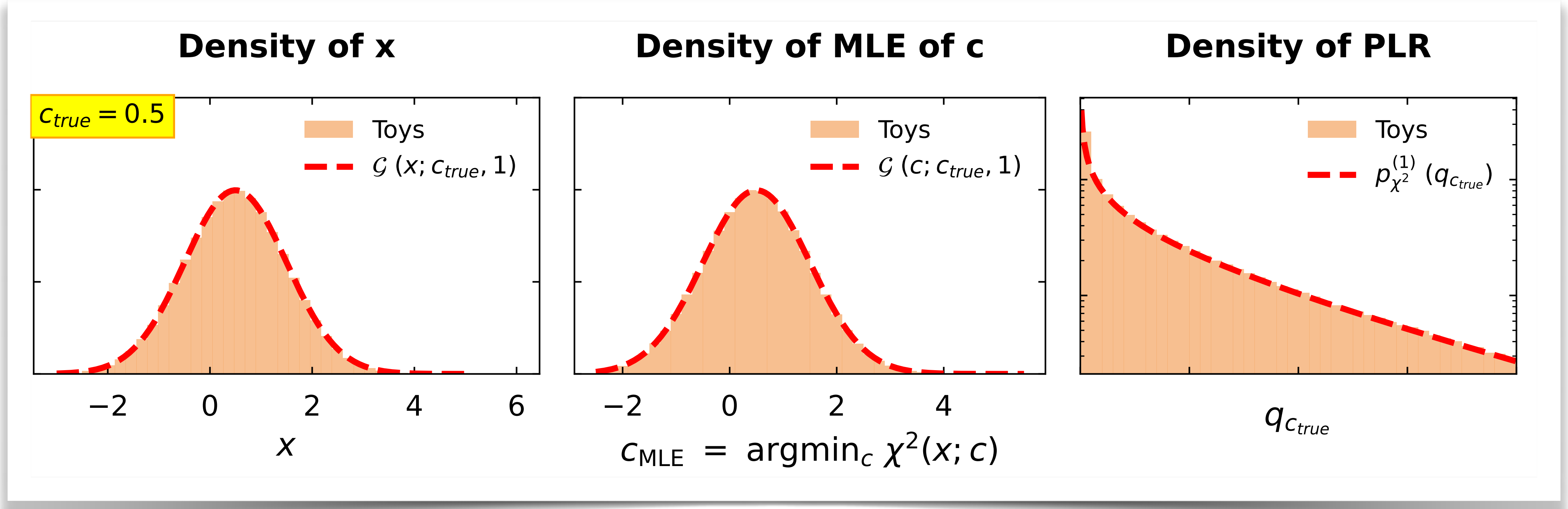
$$\mathscr{L}(x\,|\,c) = \frac{\exp\left[-\frac{1}{2}\chi^2(x;c)\right]}{\sqrt{2\pi}\,|\,\Sigma\,|} \qquad \chi^2(x;c) = \left(x-\mu(c)\right)^T\,\Sigma^{-1}\,\left(x-\mu(c)\right)$$

χ^2 is squared-length of vector $x-\mu(c)$ in flat space with metric $g=\Sigma^{-1}$

$$q_c = \chi^2(x;c) - \chi^2(x;c_{\text{MLE}})$$

The problem

- Wilks not valid if model $\mu(c')$ encounters **boundary** when optimising for MLE

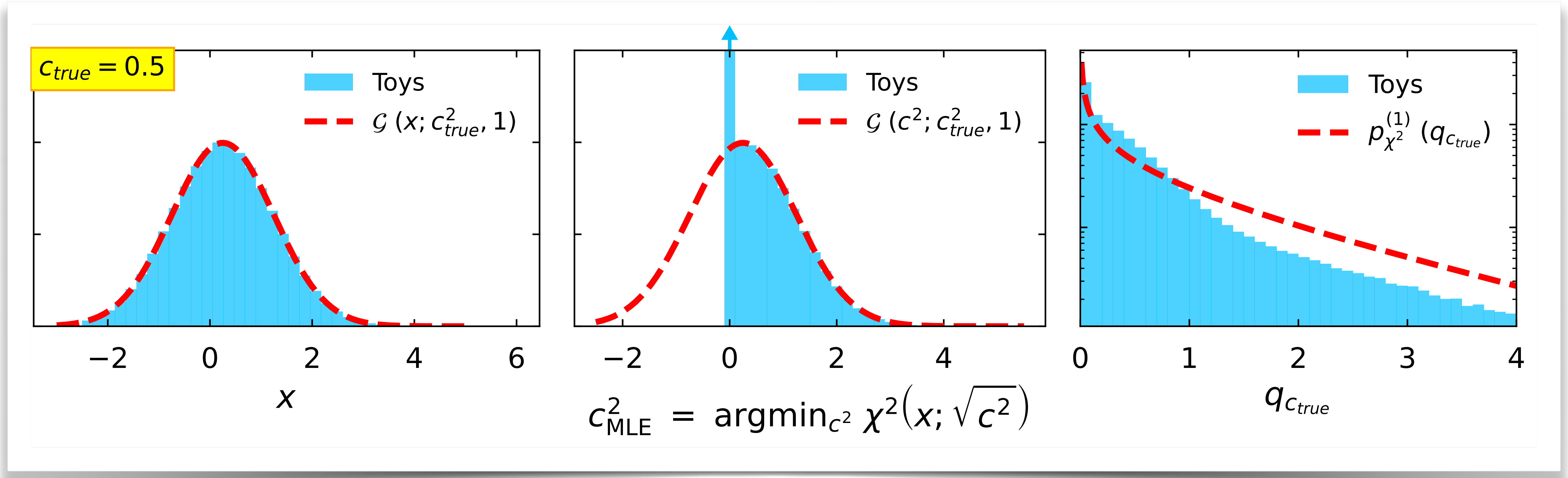


EFT dominated by linear part

$$\mu(c') = x^{\mathrm{SM}} + c' x^{\mathrm{int}}$$

c_{MLE} also Gaussian distributed because of linear relationship

$p(q_c)$ is a χ^2-distribution



EFT dominated by quadratic part

$$x = x^{\mathrm{SM}} + c'^2 x^{\mathrm{new}}$$

Model cannot capture fluctuations that go below SM, because $c'^2 x^{\mathrm{new}}$ is always +ve. Profiling encounters boundary at $c_{\mathrm{MLE}} = 0$

$p(q_c)$ is partly χ^2, partly something else

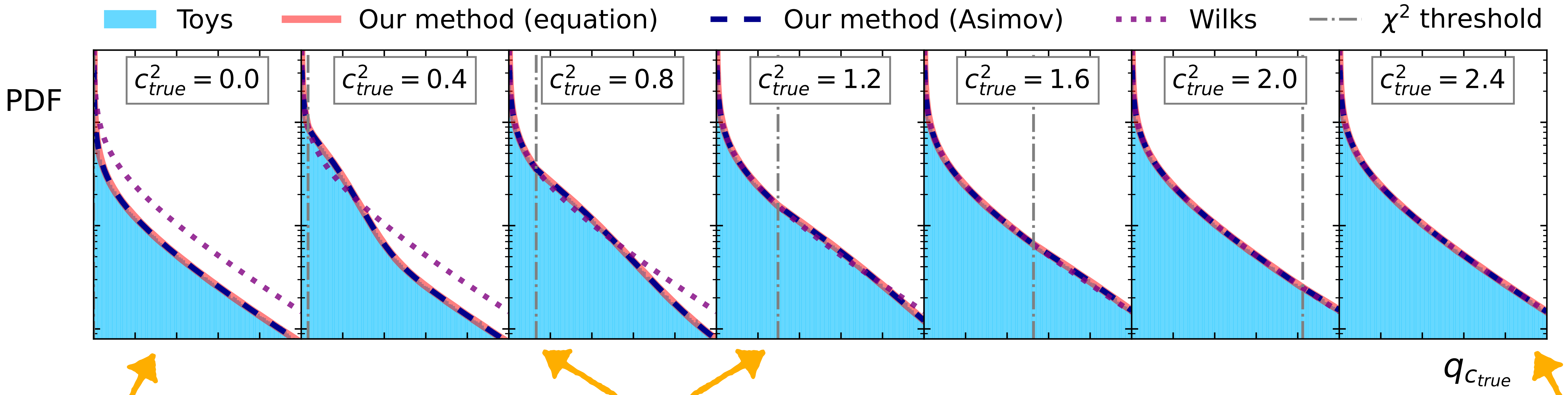
A single quadratic coefficient

Correct formula

$$p(q_c) = \begin{cases} p^{(1)}_{\chi^2}(q_c) & q_c < \frac{c^4}{\sigma^2_{c^2}} \\ \frac{1}{2} p^{(1)}_{\chi^2}(q_c) + \mathcal{G}\left(q_c; -\frac{c^4}{\sigma^2_{c^2}}, \frac{2c^2}{\sigma_{c^2}}\right) & q_c \geq \frac{c^4}{\sigma^2_{c^2}} \end{cases}$$

- Below some threshold, density is still χ^2
  - Intuition: small q_c = small data fluctuation = inside boundary
- Above threshold, density is mixture of χ^2 and Gaussian
  - Intuition: χ^2 if fluctuation ↑, Gaussian if ↓
- Parameter $\sigma^2_{c^2}$ can be **calculated from** $\{\mu, \Sigma\}$or **extracted from Asimov profile**

Toys

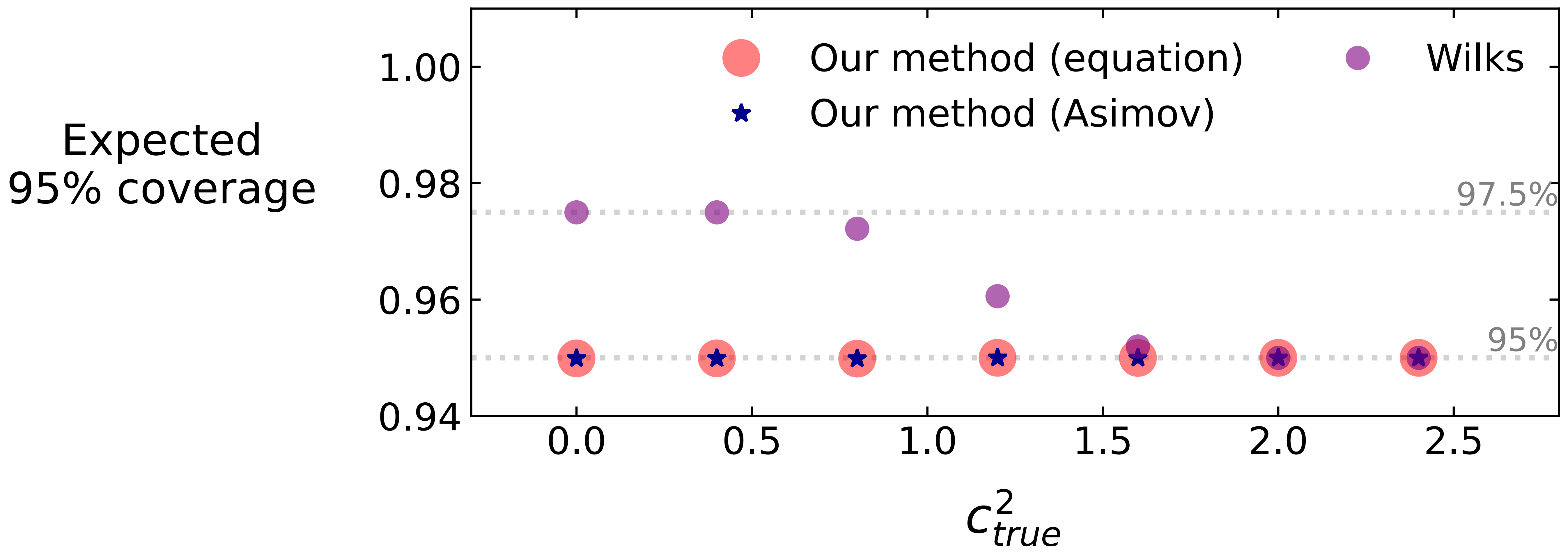


Threshold @ $q_c = 0$, density is $\frac{1}{2}\left(p^{(1)}_{\chi^2}(q_c) + \delta(q_c)\right)$

Intermediate threshold & mixed distributions

Threshold @ large q_c, density is $\approx p^{(1)}_{\chi^2}(q_c)$

Consequence

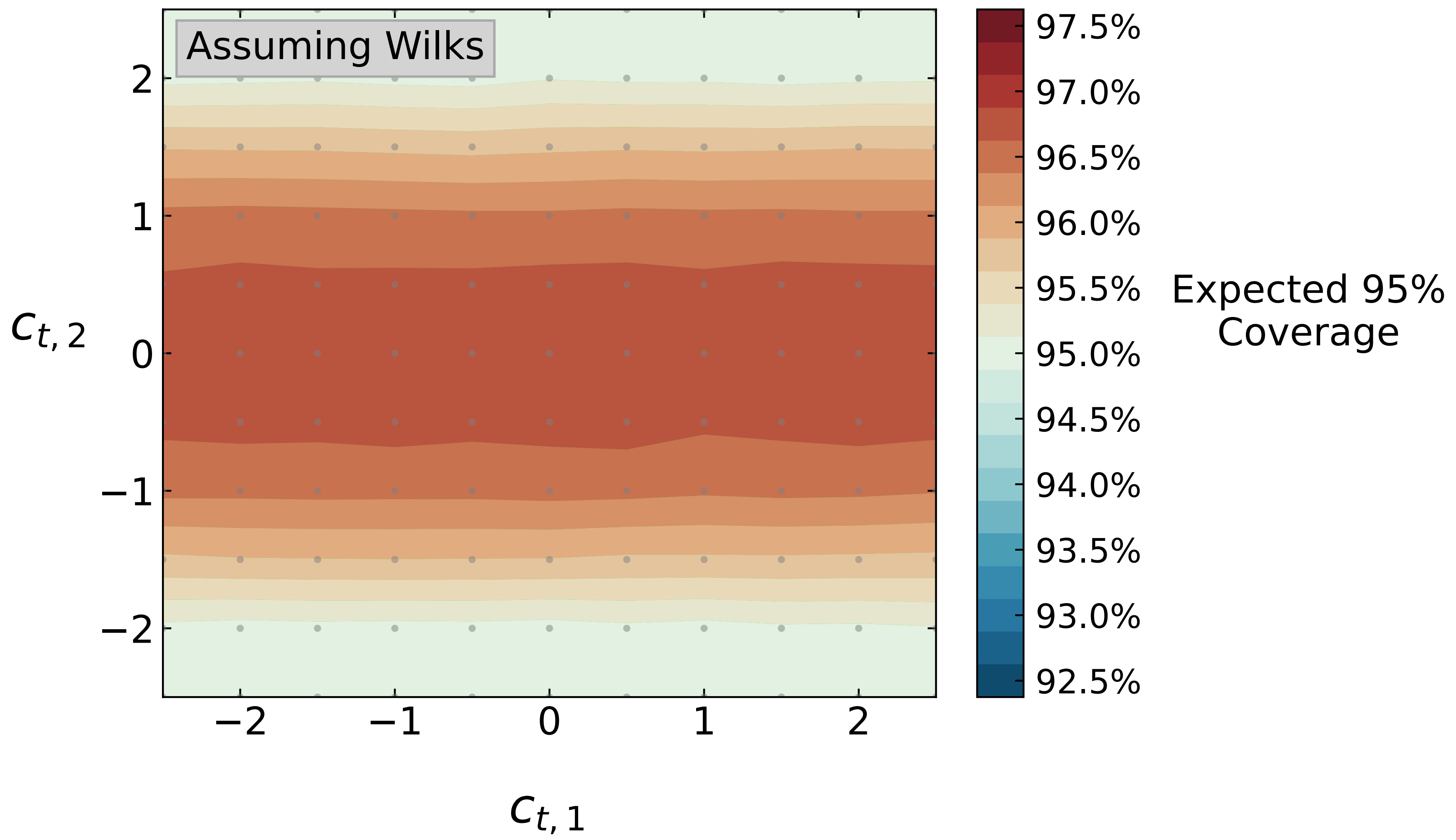


- **Overcoverage around the SM** if we assume Wilks
- **At SM, exclusion rate is half of target**, because half of datasets are “negative fluctuations” that encounter boundary
  - 95% CL actually has coverage of 97.5%
- 95% CL Wilks becomes correct for hypotheses that are distinguishable from SM at level of $\gtrsim 2\sigma$

Cases with 2 EFT coefficients

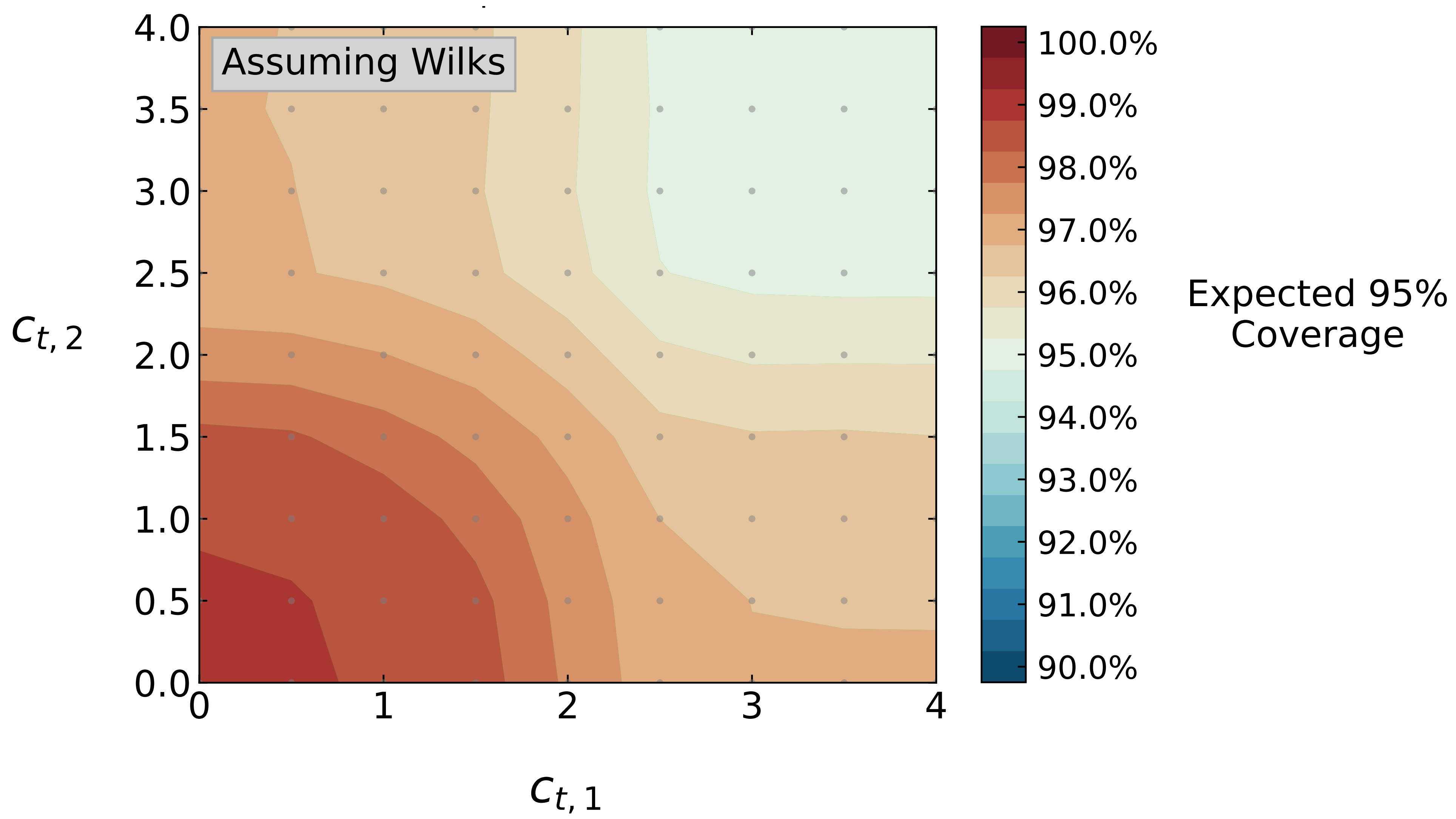
- Solution: partly closed-form, partly 1D numerical integration, very fast using Gaussian quadrature
- Again, params can be extracted from Asimov scan, and now include the *correlation coefficient* from the Hessian
- Intuition: for M quadratic params, divide param-space into 2^M regions, one which only 1 is free of boundary problem

Lin x quad



Target coverage: $95\,\%$. Actual: $95 - 97.5\,\%$ depending on correlation between x_1^{int} and x_2^{new}

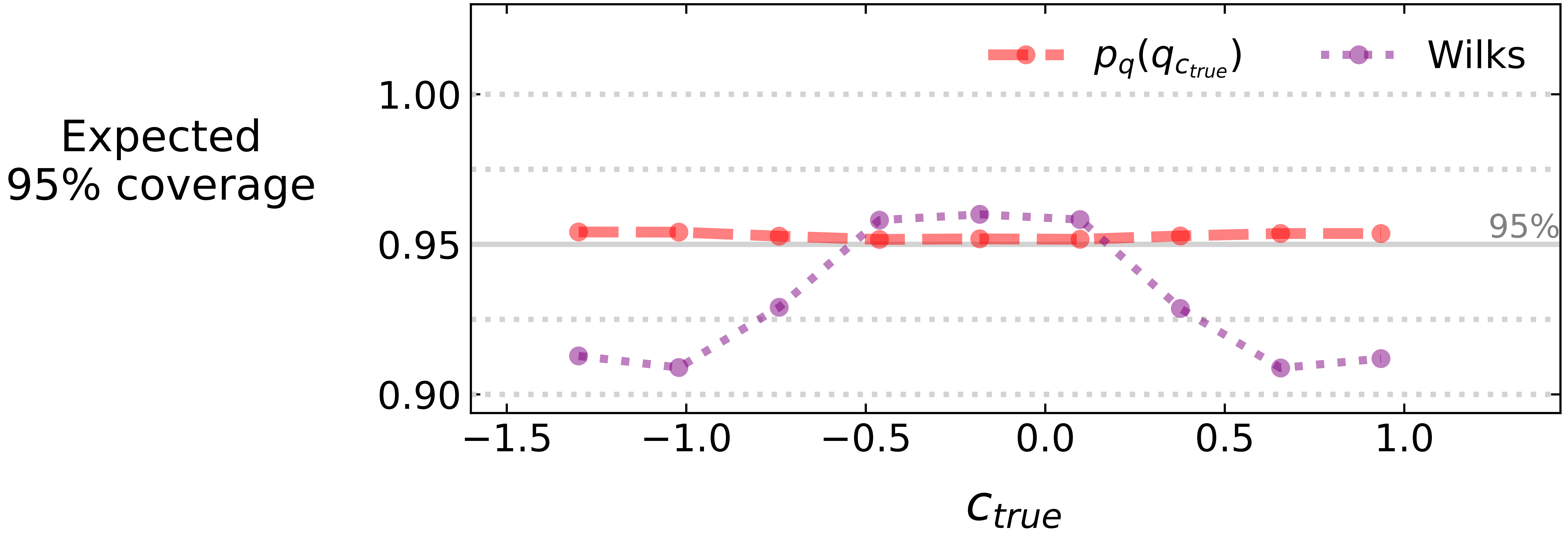
Quad x quad



Target coverage: $95\,\%$. Actual: up to $98.75\,\%$ depending on correlation between x_1^{new} and x_2^{new}

Large over-coverage at SM** if we have **many uncorrelated quad coefficients

A single lin+quad coefficient



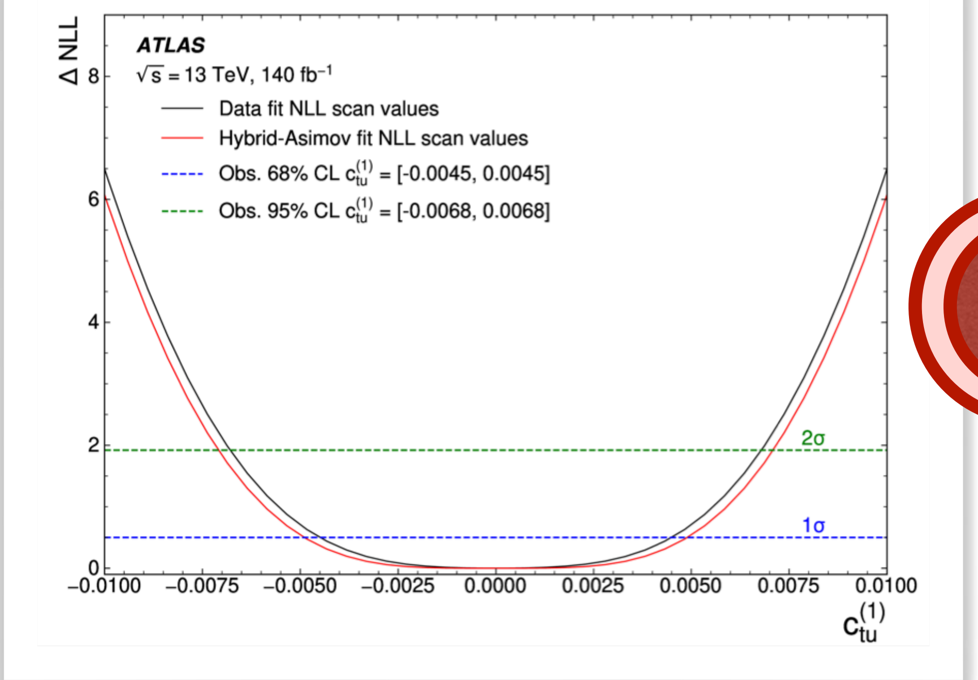
Purple: Assuming Wilks can cause **under-coverage** for lin+quad case

Red: We provide numerical method with ≈ right coverage, but computationally expensive

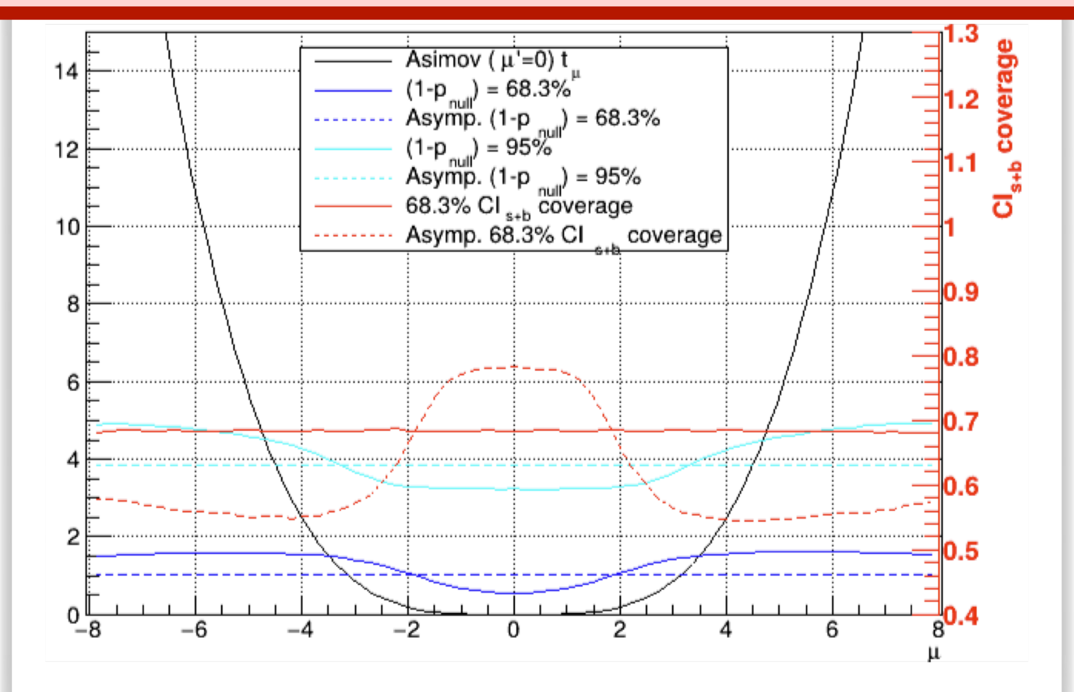
Experimental work

- **ATLAS and CMS investigating real-world impact and developing experimental checks/solutions**
  - Further details: https://indico.cern.ch/event/1452656/

1 Slight over-coverage in single-sign top [arXiv: 2409.14982]

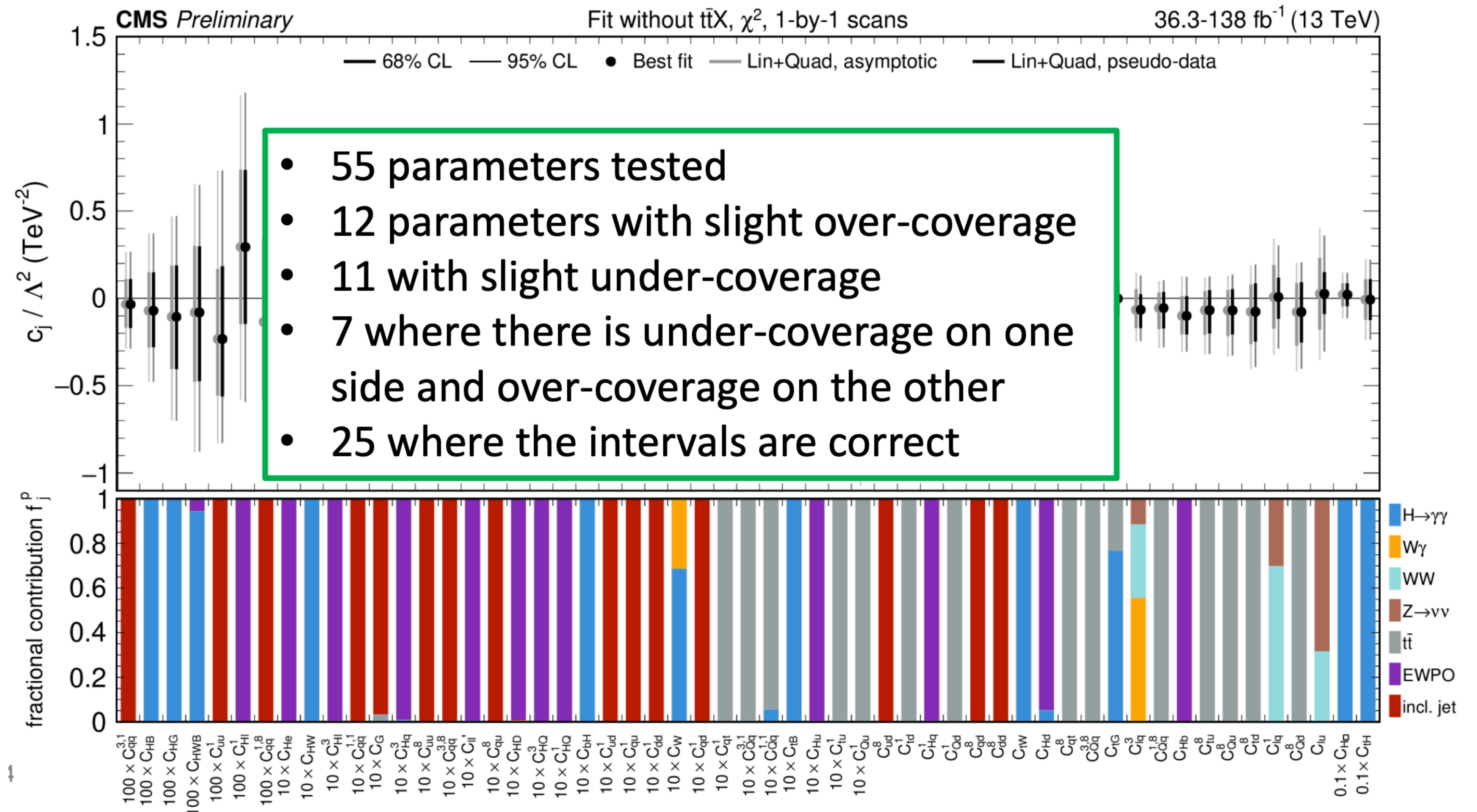


2 Ongoing 1D analysis: over-and under-coverage, with toy corrections



3 Toy methods difficult to scale to multiple coefficients

1 Global EFT fits with various over- and under-coverage [CMS-SMP-24-003]



- 55 parameters tested
- 12 parameters with slight over-coverage
- 11 with slight under-coverage
- 7 where there is under-coverage on one side and over-coverage on the other
- 25 where the intervals are correct

ATLAS *Credit: Gianna Loeschcke Centeno*

CMS *Credit: Brent R. Yates*

Summary

- **Wilks theorem doesn't apply to EFT fits that describe only positive deviations from SM**

- If all params dominated by linear **or** quadratic contributions:
    - *Over-coverage* around SM, conservative but reduces power to see new physics
    - Provided correct distributions for ≤ 2 EFT params

- If linear **and** quadratic contributions both relevant:
    - Can experience *under-coverage*, which is more problematic
    - Continuous transition between contributions makes life hard
    - Solved one-bin special case
    - Provided numerical method that may be computationally cheaper than toys

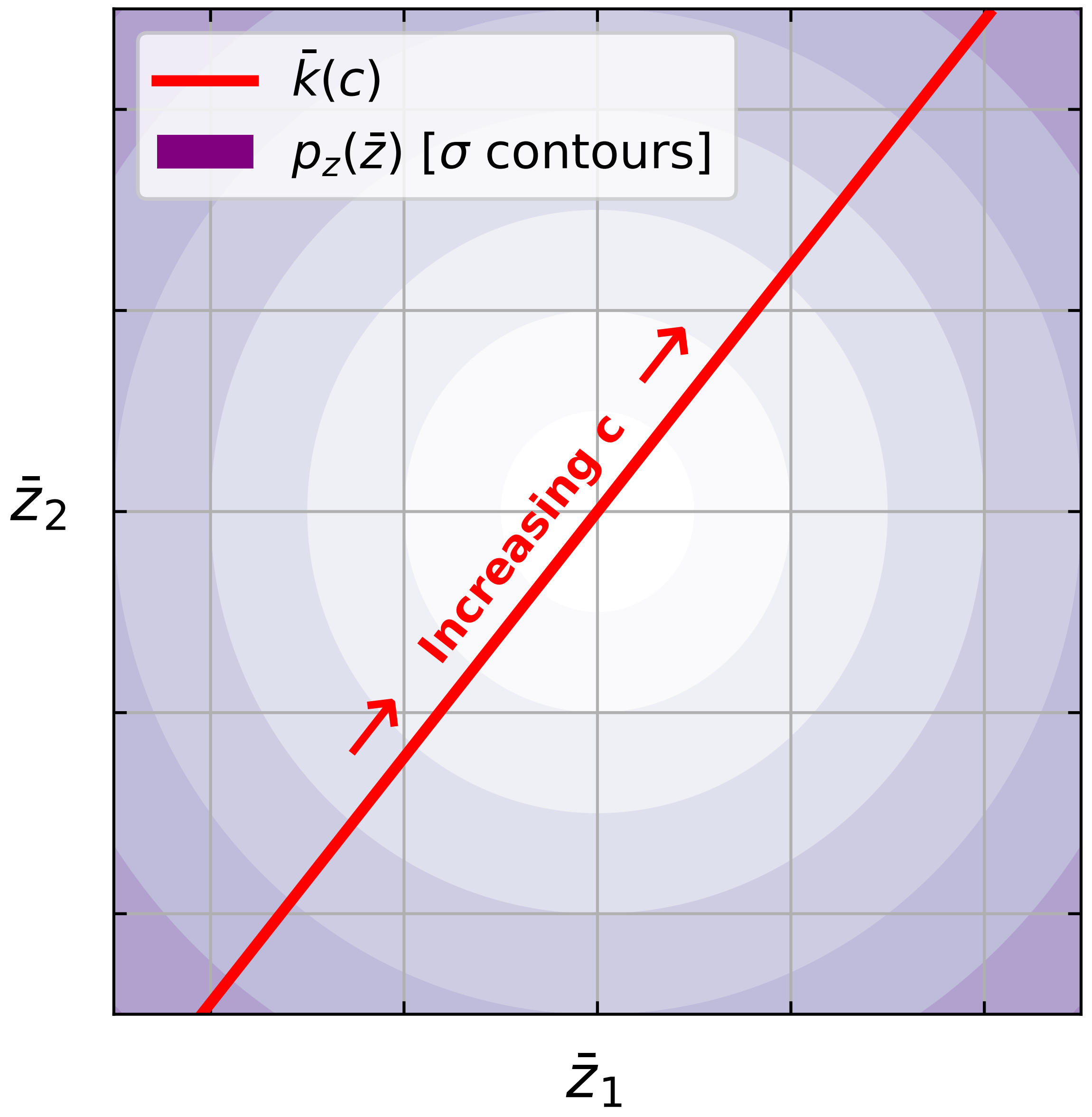
- Opportunities:
    1. Change-of-variables can be automated for arbitrary num lin *or* quad params: 2^M integration regions of which 1 is Wilks-like
    2. Analytic solution for lin+quad case, or extend numerical solution to multiple params
    3. Quadratic + quartic case, interference between EFT params, non-Gaussian NPs…

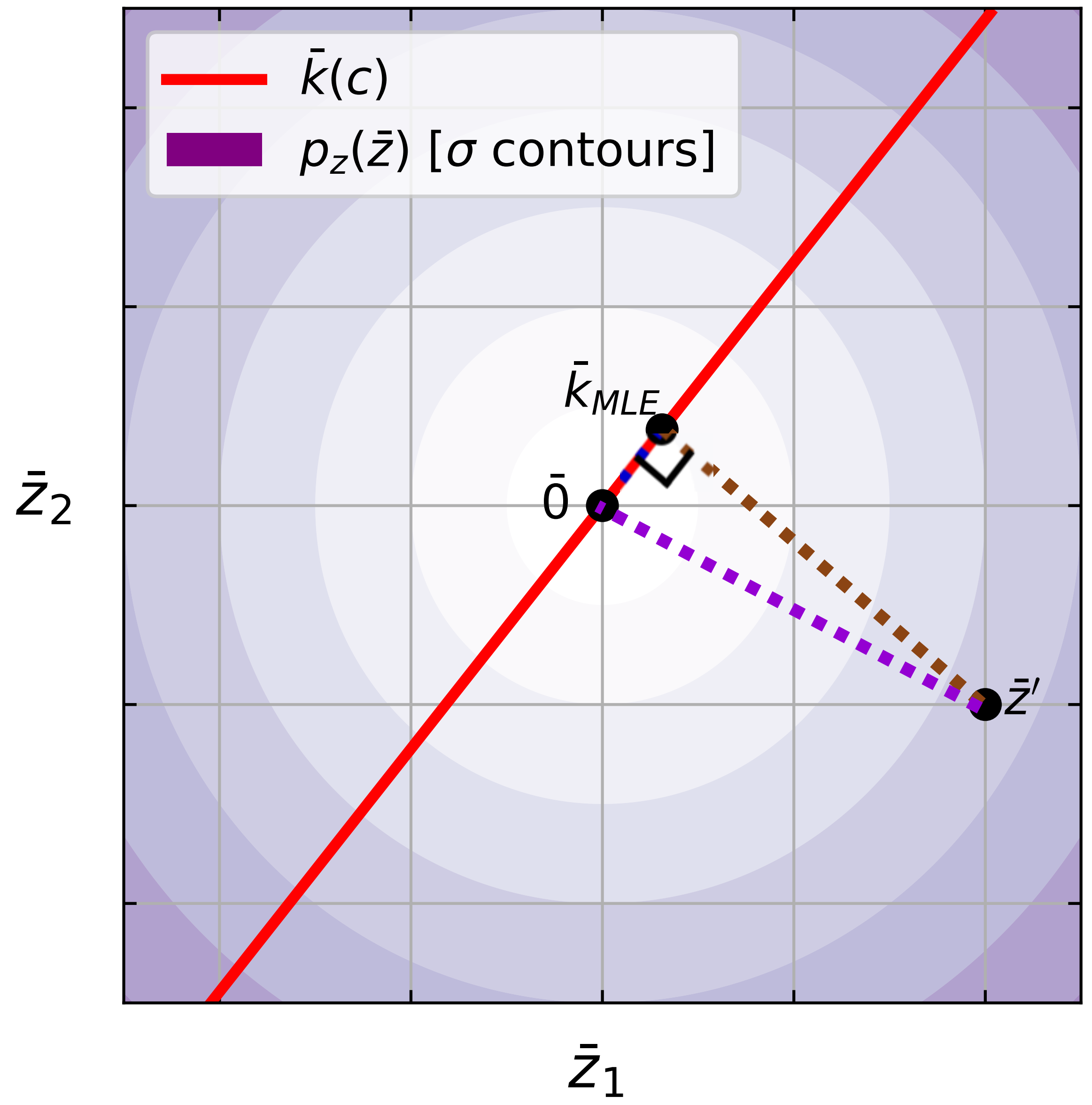
Backup

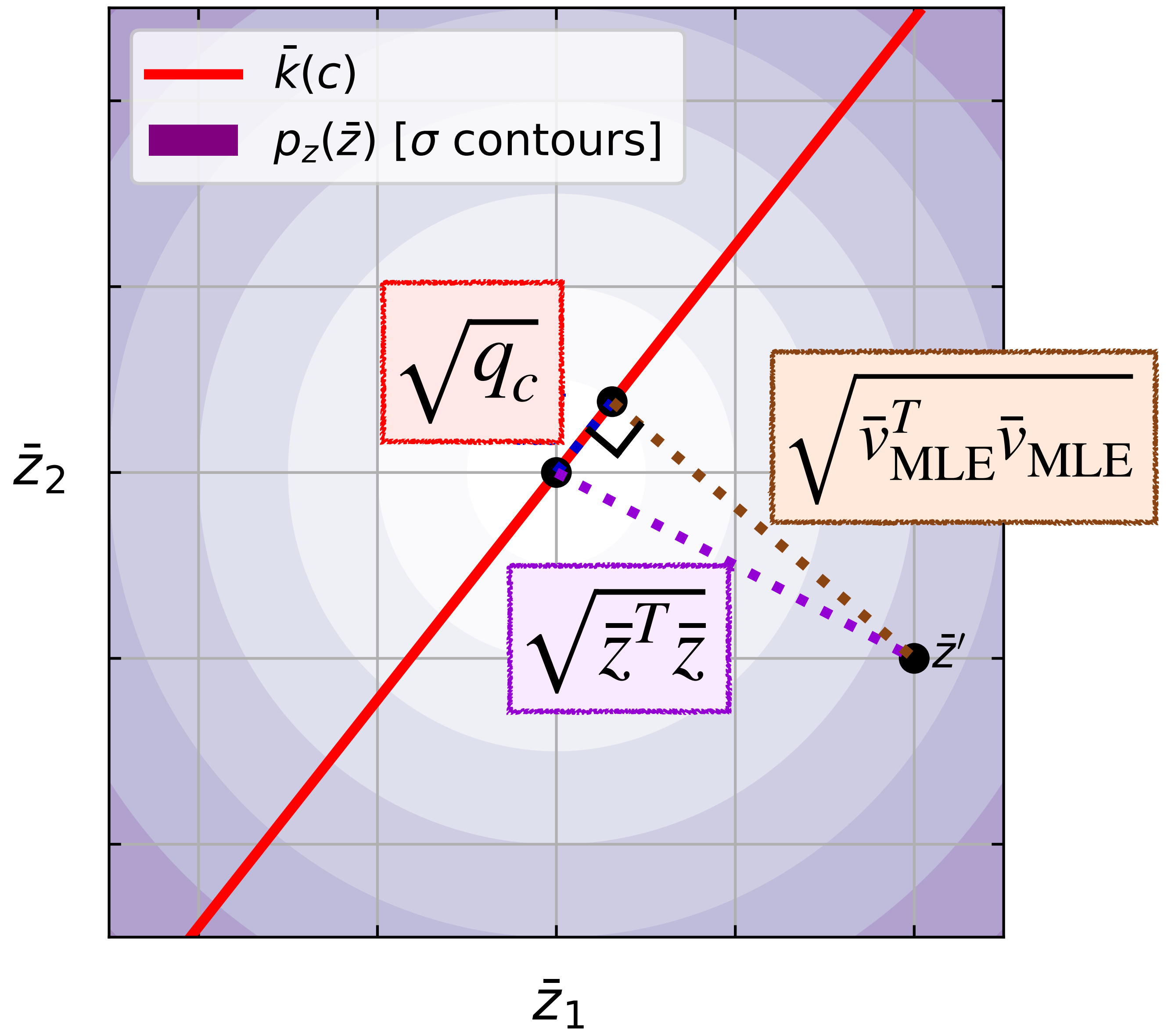
Geometric intuition

- Take observation vector $x \in \mathbb{R}^{N_x}$
- Make life easier by working in co-ordinates where $\mu(c)$ is the origin and Σ is identity
  - This turns x into **uncorrelated normally distributed random variables** $\bar{z} \in \mathbb{R}^{N_x}$
- Then $q_c \;=\; \bar{z}^T\bar{z} \;-\; \bar{v}_{\mathrm{MLE}}^T\bar{v}_{\mathrm{MLE}}$ where $\bar{v} = \bar{z} - \bar{k}(c')$ is residual between observation $\bar{z}$ and prediction $\bar{k}(c')$
  - i.e. **PLR is the difference in squared lengths** between observation vector and residual at MLE

Single linear case





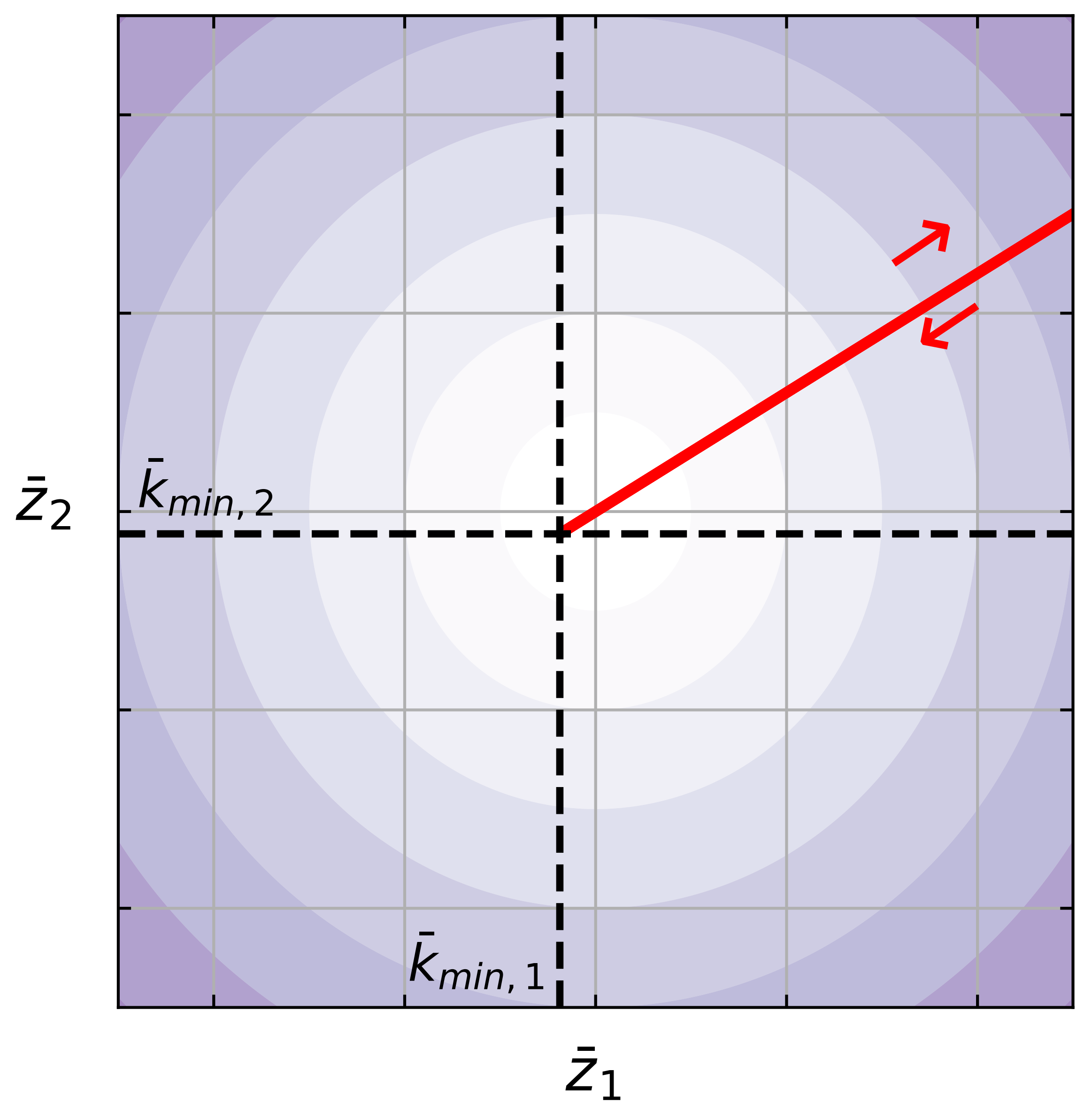


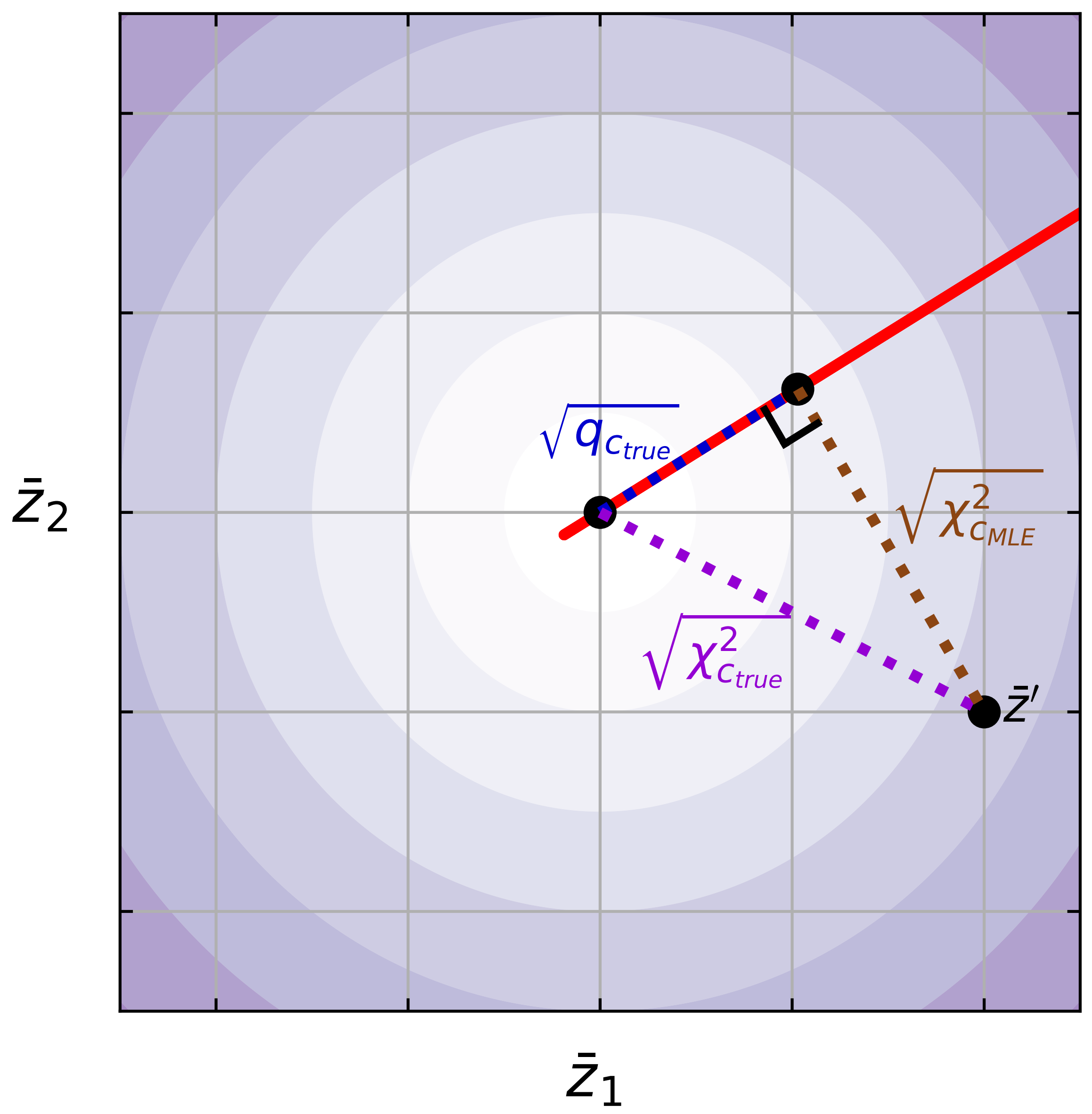
- Prediction curve $\bar{k}(c')$ traces out a straight line in $\bar{z}$-space

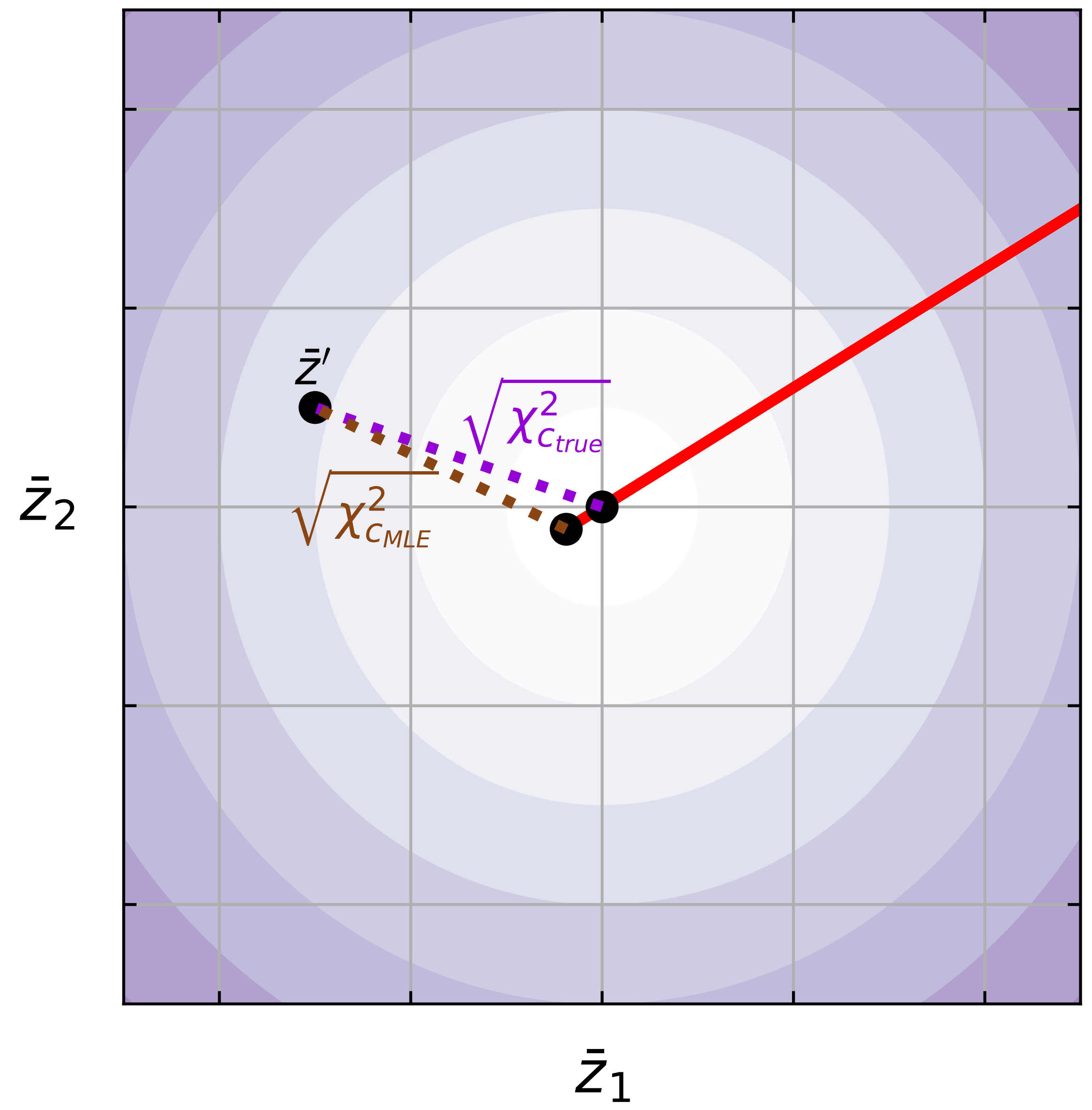
- Origin $\bar{0}$ is true c
- Observe some point $\bar{z}$
- MLE is the point of closest approach
- Three points trace a **right-angled triangle**

- $q_c = \bar{z}^T\bar{z} - \bar{v}_{\mathrm{MLE}}^T \bar{v}_{\mathrm{MLE}}$ is **Pythagoras theorem**
- This makes $\sqrt{q_c}$ a **projection of $\bar{z}$ onto 1D subspace**
- Projecting NₓD Normal dist gives 1D Normal, and squaring it makes q_c a χ^2-dist w/ 1 dof —> **Wilks**
- General: project to N_c-hyperplane —> χ^2 with N_c dof

Single quadratic case







- $\bar{k}(c')$ still linear in prediction space and passes through the origin
- Now turns around at SM

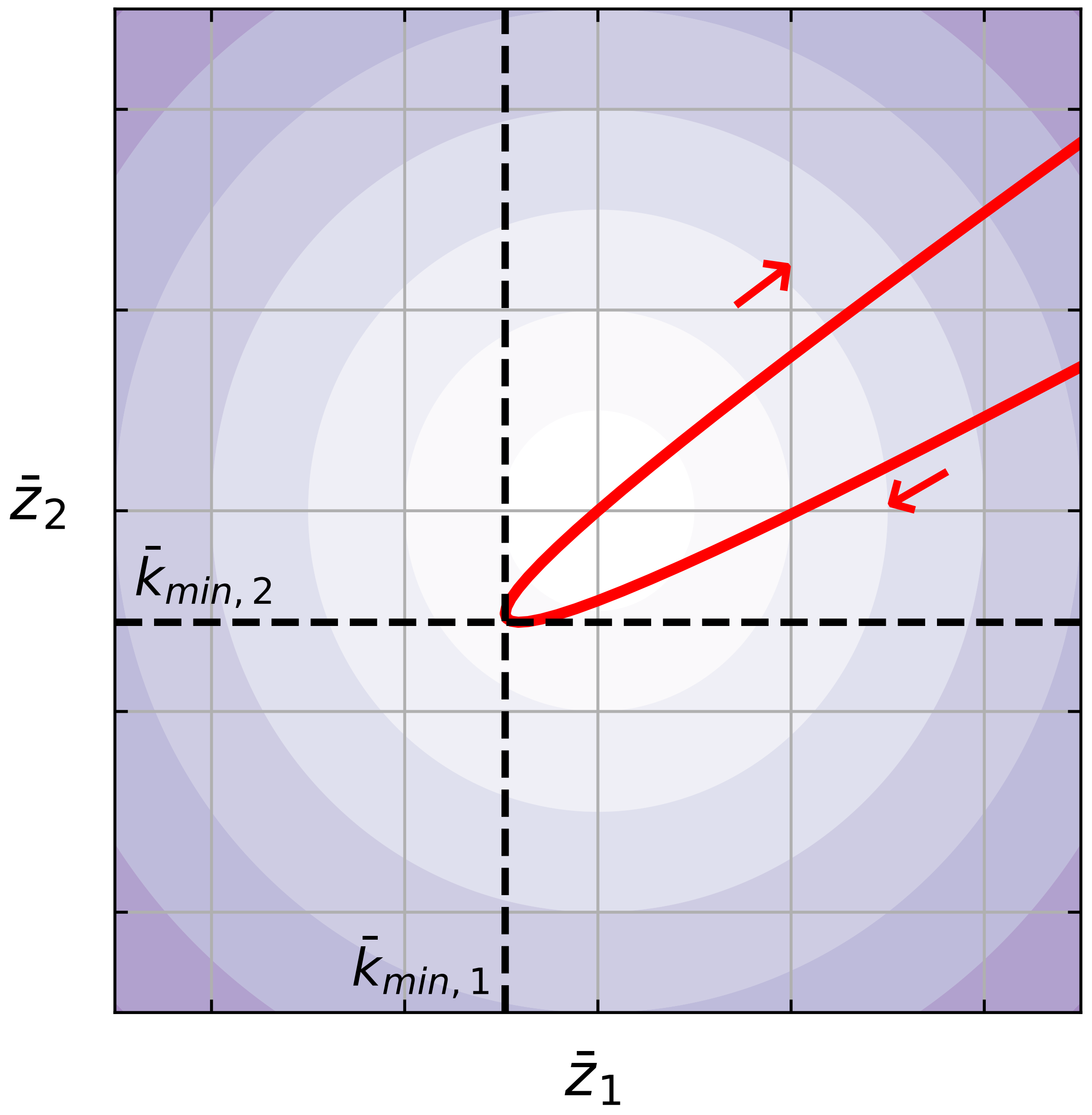
- When sample $\bar{z}$ falls in perpendicular subspace, same logic still holds, and we get χ^2 parts

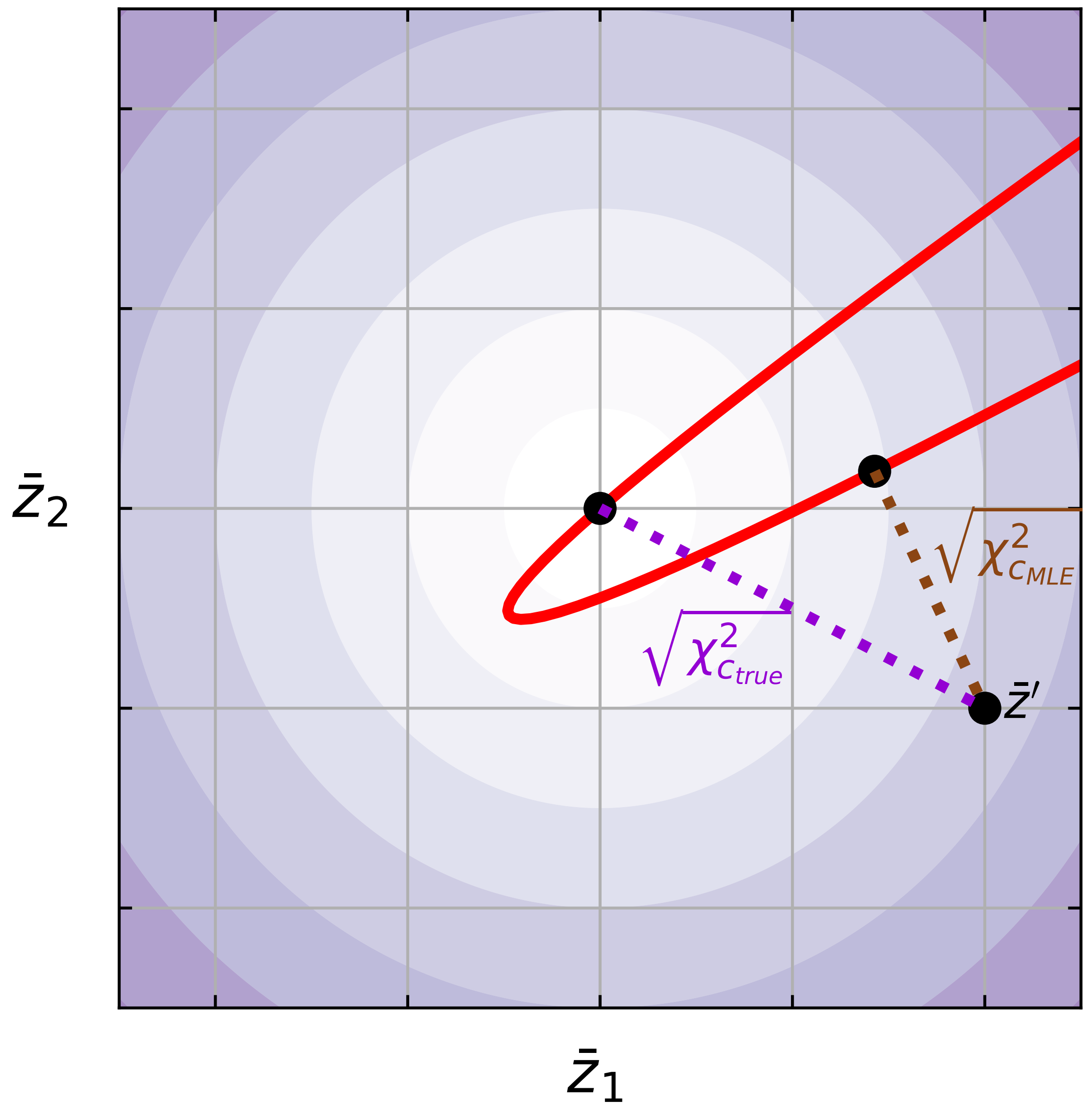
- When sample $\bar{z}$ falls elsewhere, the point-of-nearest-approach is the SM boundary point
- No-long a right-angled triangle —> no longer a 1D projection —> no longer χ^2
- q_c now solved using *scalene triangle* equation

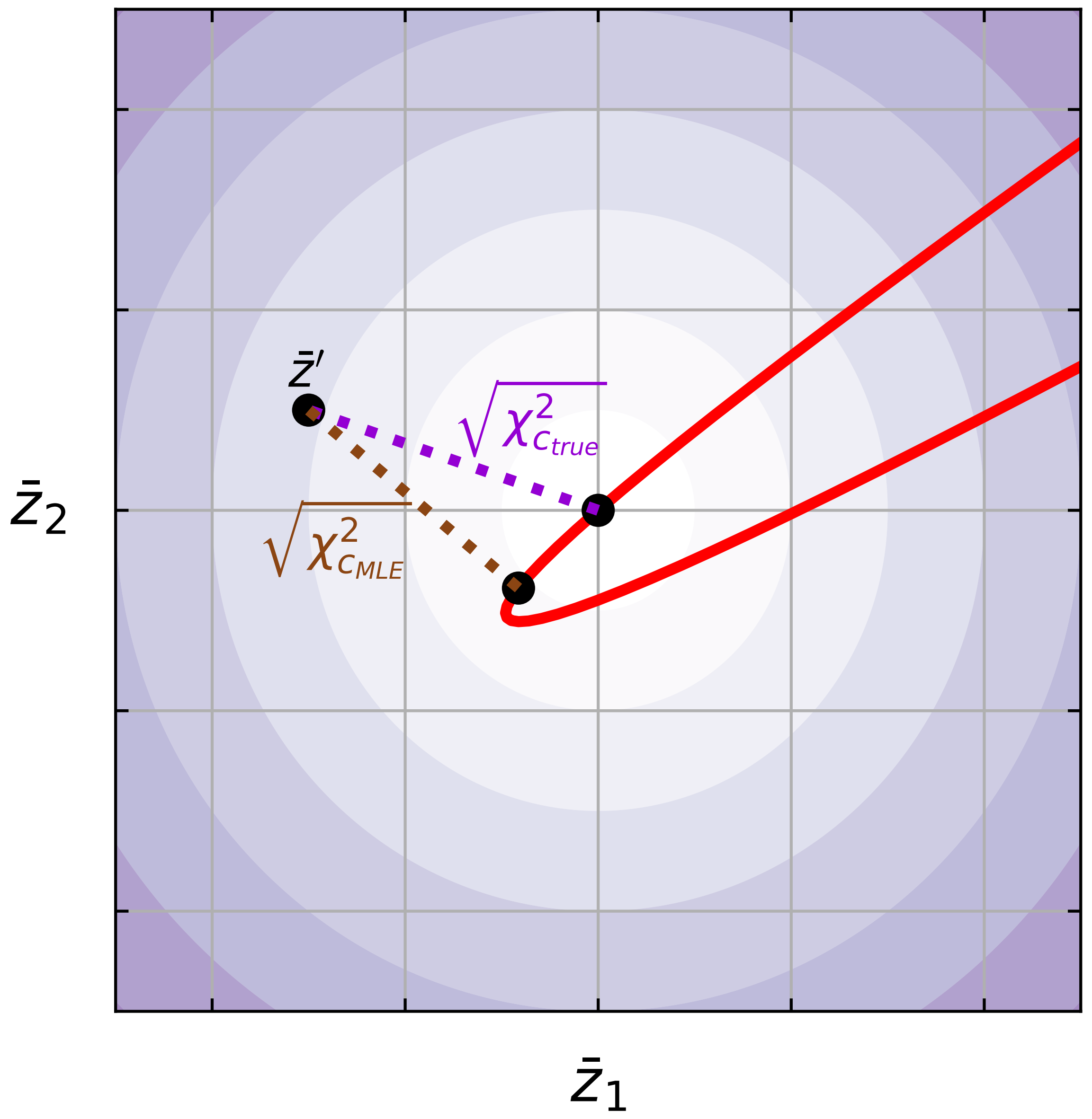
$$q_c \;=\; |\bar{z}'|^2 - \chi^2_{\text{MLE}} \;=\; -\,2c\,(\bar{z}'\cdot\bar{n}) - c^2\,\left|\bar{n}\right|^2$$

… where $\bar{z}'\cdot\bar{n}$ is Gaussian

Single lin+quad case







- As we profile c', the relative contributions of linear and quadratic components change in quadratic way, and $\bar{k}(c')$ traces a parabola
  - Parallel to $\bar{x}^{\text{int}}$ as $c' \to 0$
  - Parallel to $\bar{x}^{\text{new}}$ as $c'^2 \to \infty$

- **Lin+quad case is smooth continuum** —> no discrete modes (AFAIK)
  - Contrast to quad case, could be decomposed into two discrete modes, depending on whether MLE was a ⊥-distance or point projection
  - Forced to use change-of-variables formula

$$p(q_c) = \oint \left| \frac{\mathrm{d}\bar{z}}{\mathrm{d}q_c} \right| \mathcal{N}(\bar{z})\ \mathrm{d}\bar{z}(q_c)$$

Difficult to find $\bar{z}(q_c)$ analytically, instead we provide numerical method that scales slightly better than toys

Forward process $\bar{z} \rightarrow q_c$

Normally distributed random variables

$\bar{z} \sim \mathcal{N}(\bar{z})$

Scale+rotate+shift onto observation space

x

Analytically solve optimisation for MLE

c_{MLE}

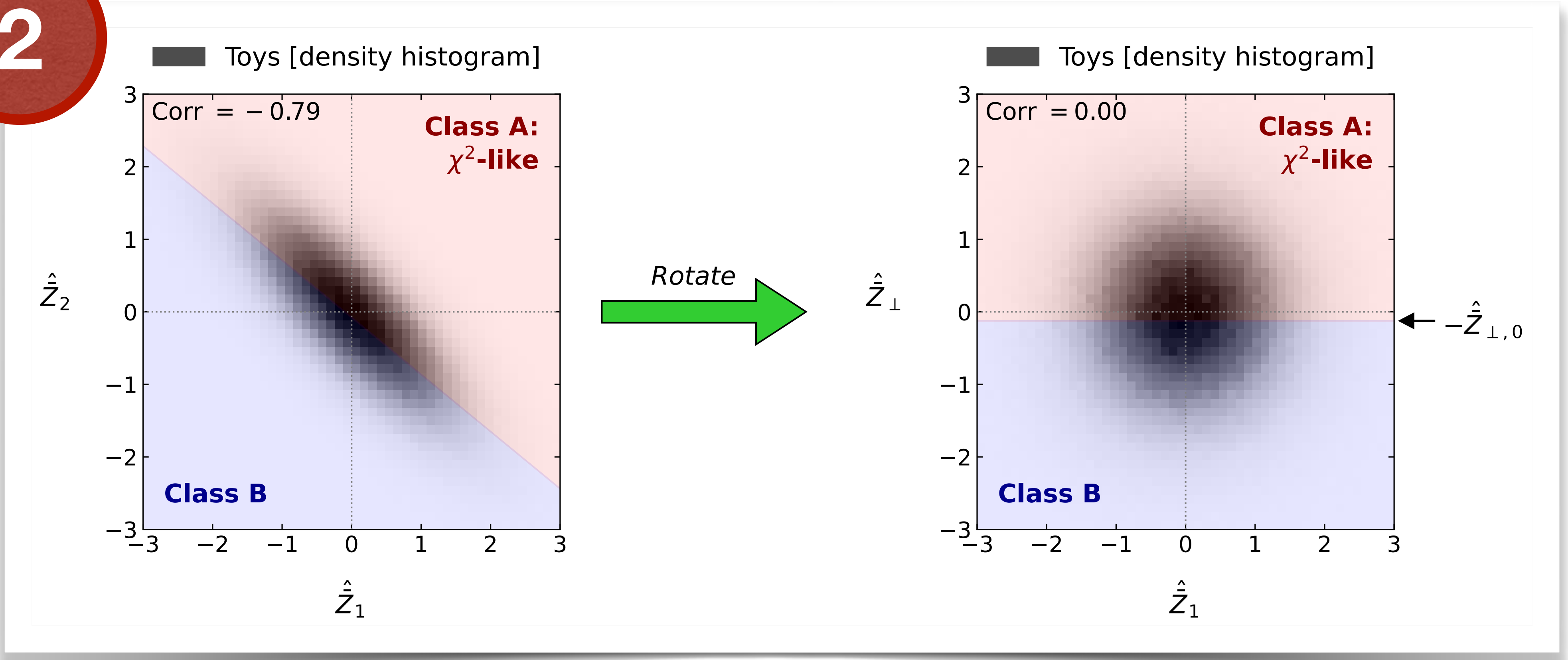
Calculate PLR

$$q_c = \chi^2(c; x) - \chi^2(c_{\text{MLE}}; x)$$

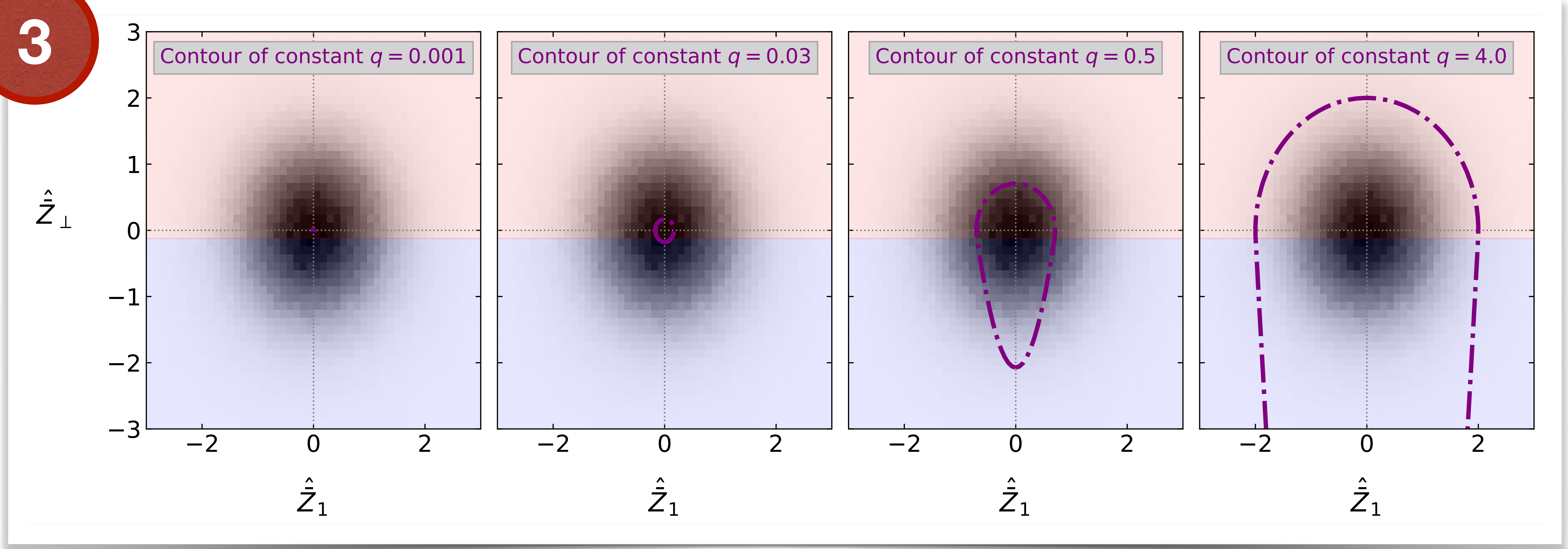
To calculate $p(q_c)$ using change-of-variables, we must solve inverse process $q_c \rightarrow \{x\}$ via latent step of c_{MLE}. For lin+quad case, this was too difficult to solve analytically, so we used numerical method. Jacobian can be calculated analytically or automatically with e.g. JAX.

Integration contours (lin x quad)

2



3



1

$$p_q\left(q_{c_{\text{true}}}\right) = \begin{cases} \frac{1}{2}p^{(1)}_{\chi^2}\left(q_{c_{\text{true}}}\right) + \frac{1}{2}p^{(2)}_{\chi^2}\left(q_{c_{\text{true}}}\right) & \text{if } c_{t,2} = 0\,, \\ p^{(2)}_{\chi^2}\left(q_{c_{\text{true}}}\right) & \text{if } q_{c_{\text{true}}} \leq \hat{\hat{Z}}^2_{\perp,0}\,, \\ \frac{1}{2}p^{(2)}_{\chi^2}\left(q_{c_{\text{true}}}\right) + \tilde{p}_1\left(q_{c_{\text{true}}}\right) + \tilde{p}_2\left(q_{c_{\text{true}}}\right) & \text{otherwise}\,, \end{cases} \tag{40}$$

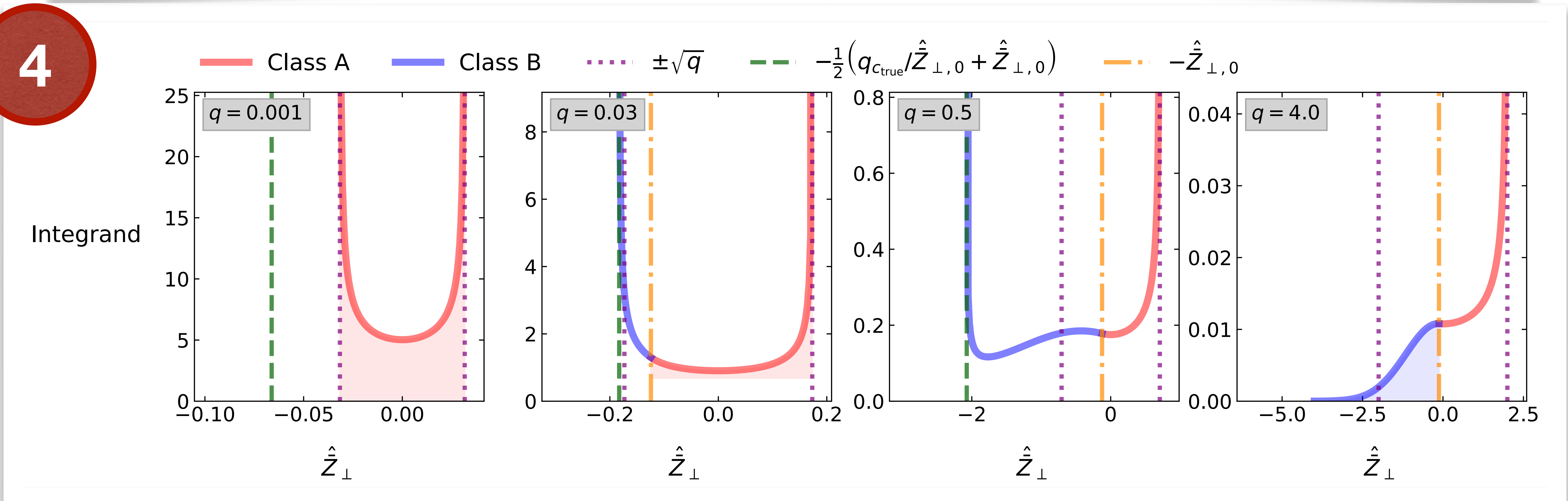
with

$$\hat{\hat{Z}}_{\perp,0} = \bar{N}_2\sqrt{1-\rho^2}\,c^2_{t,2}$$

$$\tilde{p}_1\left(q_{c_{\text{true}}}\right) = \int_0^{\hat{\hat{Z}}_{\perp,0}} p^{(1)}_{\chi^2}\left(q_{c_{\text{true}}} - \hat{\hat{Z}}^2_{\perp}\right)\mathcal{N}\left(\hat{\hat{Z}}_{\perp}\right)\mathrm{d}\hat{\hat{Z}}_{\perp}\,,$$

$$\tilde{p}_2\left(q_{c_{\text{true}}}\right) = \int_{\hat{\hat{Z}}_{\perp,0}}^{\frac{1}{2}\left(\frac{q_{c_{\text{true}}}}{\hat{\hat{Z}}_{\perp,0}} + \hat{\hat{Z}}_{\perp,0}\right)} p^{(1)}_{\chi^2}\left(q_{c_{\text{true}}} + \hat{\hat{Z}}_{\perp,0}{}^2 - 2\hat{\hat{Z}}_{\perp,0}\hat{\hat{Z}}_{\perp}\right)\mathcal{N}\left(\hat{\hat{Z}}_{\perp}\right)\mathrm{d}\hat{\hat{Z}}_{\perp}\,. \tag{41}$$

4



Integration contours (quad x quad)

1

$$p_q\left(q_{c_{\rm true}}\right) = \sum_{X\,\in\,\text{Classes}} \int_{\hat{\hat{Z}}_\perp} \left|\frac{\mathrm{d}\hat{\hat{Z}}_1^{(X)}}{\mathrm{d}q_{c_{\rm true}}}\right| \mathcal{N}\left(\hat{\hat{Z}}_1^{(X)}\right)\mathcal{N}\left(\hat{\hat{Z}}_\perp\right)\mathrm{d}\hat{\hat{Z}}_\perp$$

$$\text{Classes} = \{AA+,\ AA-,\ AB+,\ AB-,\ BA+,\ BA-,\ BB\}$$

AA $\quad \hat{\hat{Z}}_1 = \pm\sqrt{q_{c_{\rm true}} - \hat{\hat{Z}}_\perp^2}$

AB $\quad \hat{\hat{Z}}_1 = \pm\sqrt{q_{c_{\rm true}} + 2\sqrt{1-\rho^2}\,\kappa_2\,\hat{\hat{Z}}_\perp + (1-\rho^2)\kappa_2^2}$

BA

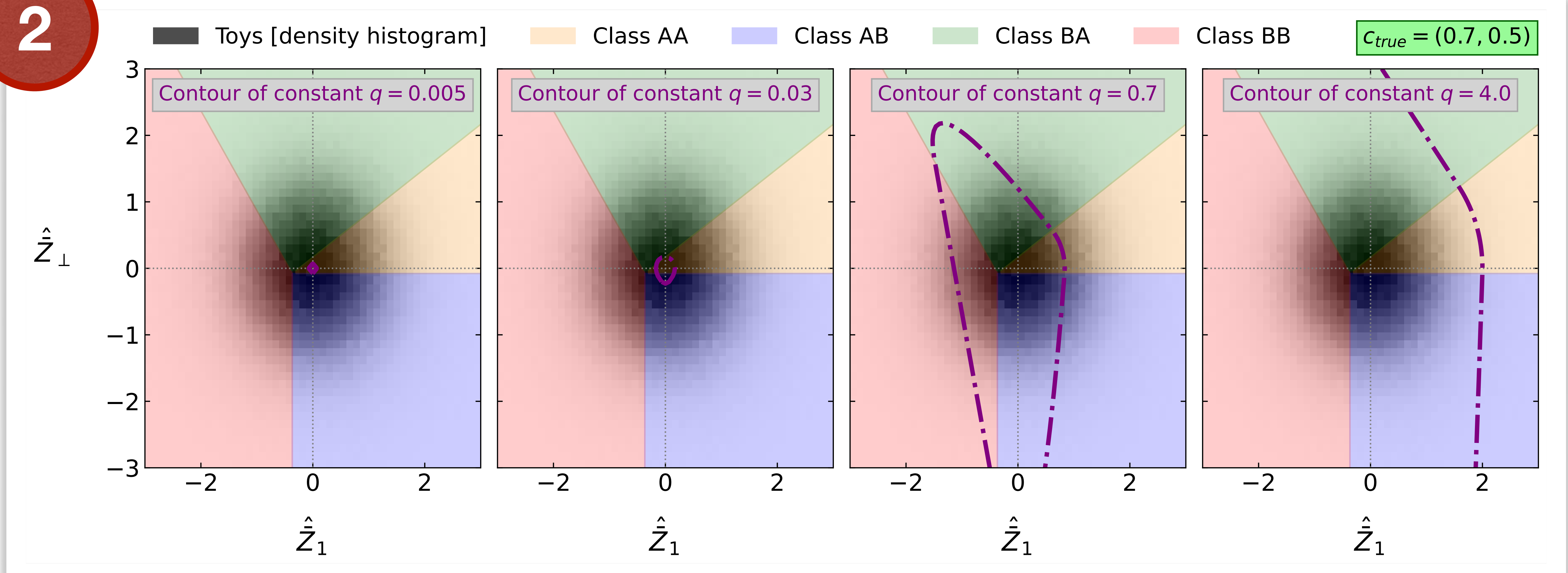
$$\alpha_1 = \rho^2, \qquad \beta_1 = 2\sqrt{1-\rho^2}\left(\rho\hat{\hat{Z}}_\perp - \sqrt{1-\rho^2}\kappa_1\right),$$

$$\gamma_1 = (1-\rho^2)\hat{\hat{Z}}_\perp^2 + 2\rho\sqrt{1-\rho^2}\,\kappa_1\,\hat{\hat{Z}}_\perp - (1-\rho^2)\kappa_1^2 - q_{c_{\rm true}}$$

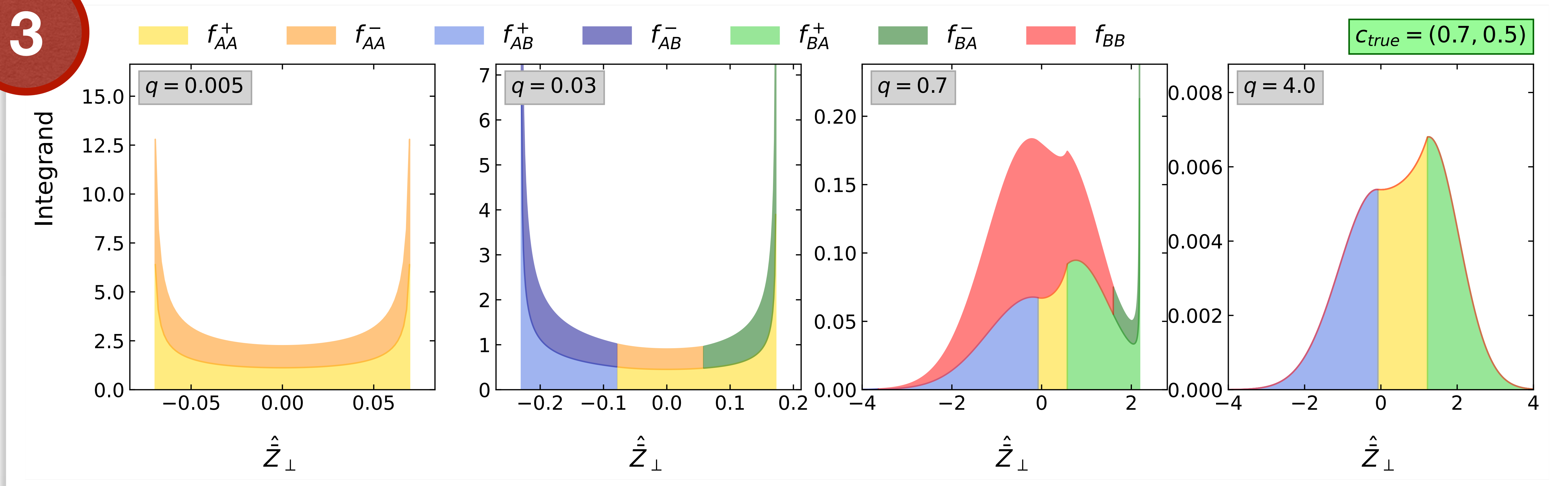
$$\hat{\hat{Z}}_1 = -\frac{1}{2\alpha_1}\beta_1 \pm \frac{1}{2\alpha_1}\sqrt{\beta_1^2 - 4\alpha_1\gamma_1},$$

BB $\quad \hat{\hat{Z}}_1 = \dfrac{1}{2(\kappa_1+\rho\kappa_2)}\left[-2\sqrt{1-\rho^2}\,\kappa_2\hat{\hat{Z}}_\perp - \left(\kappa_1^2+\kappa_2^2+2\rho\kappa_1\kappa_2+q_{c_{\rm true}}\right)\right]$

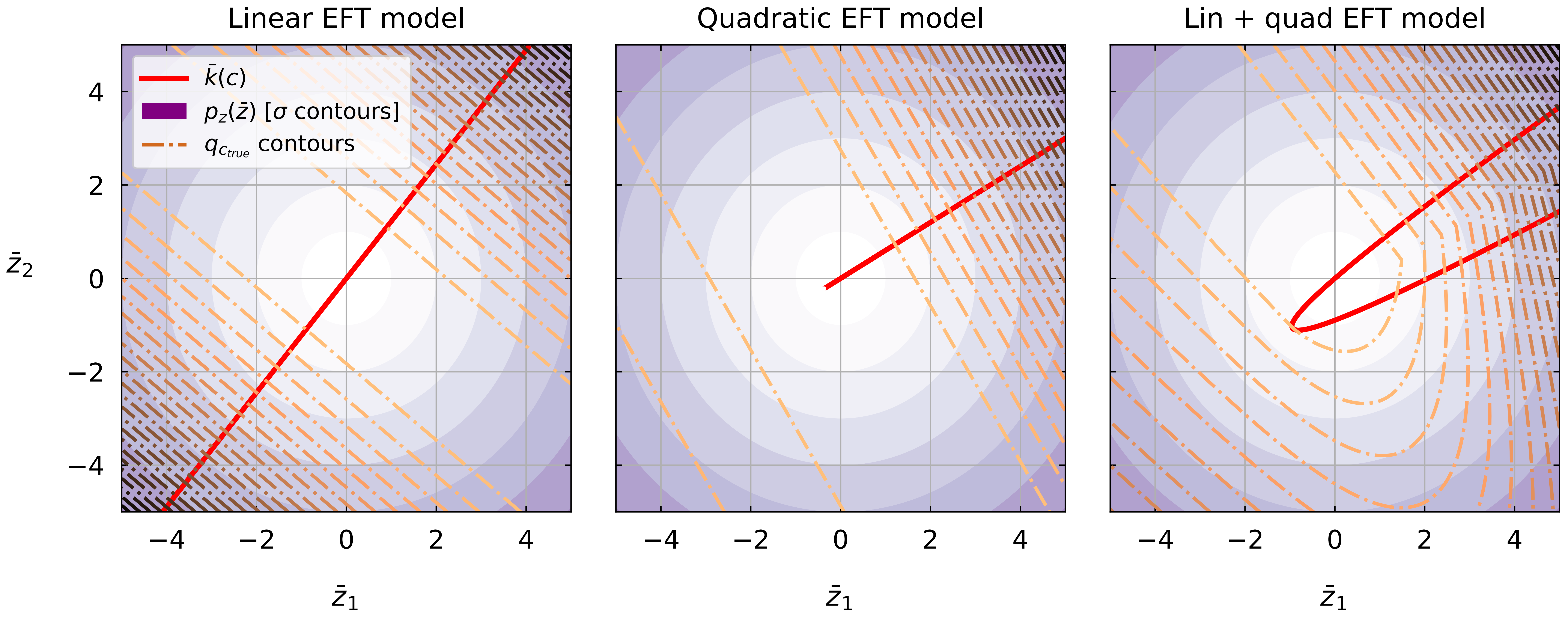
2



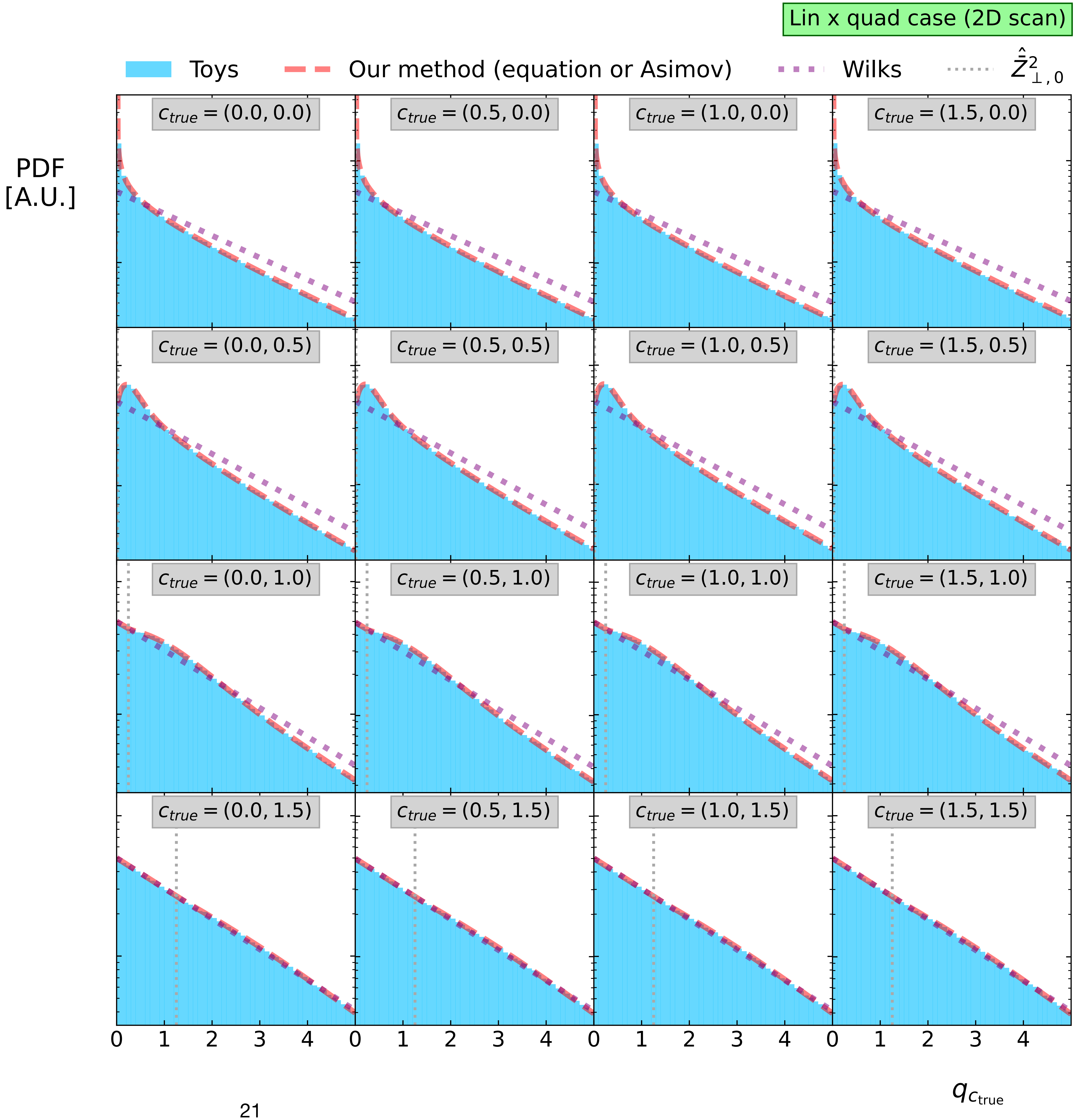
3



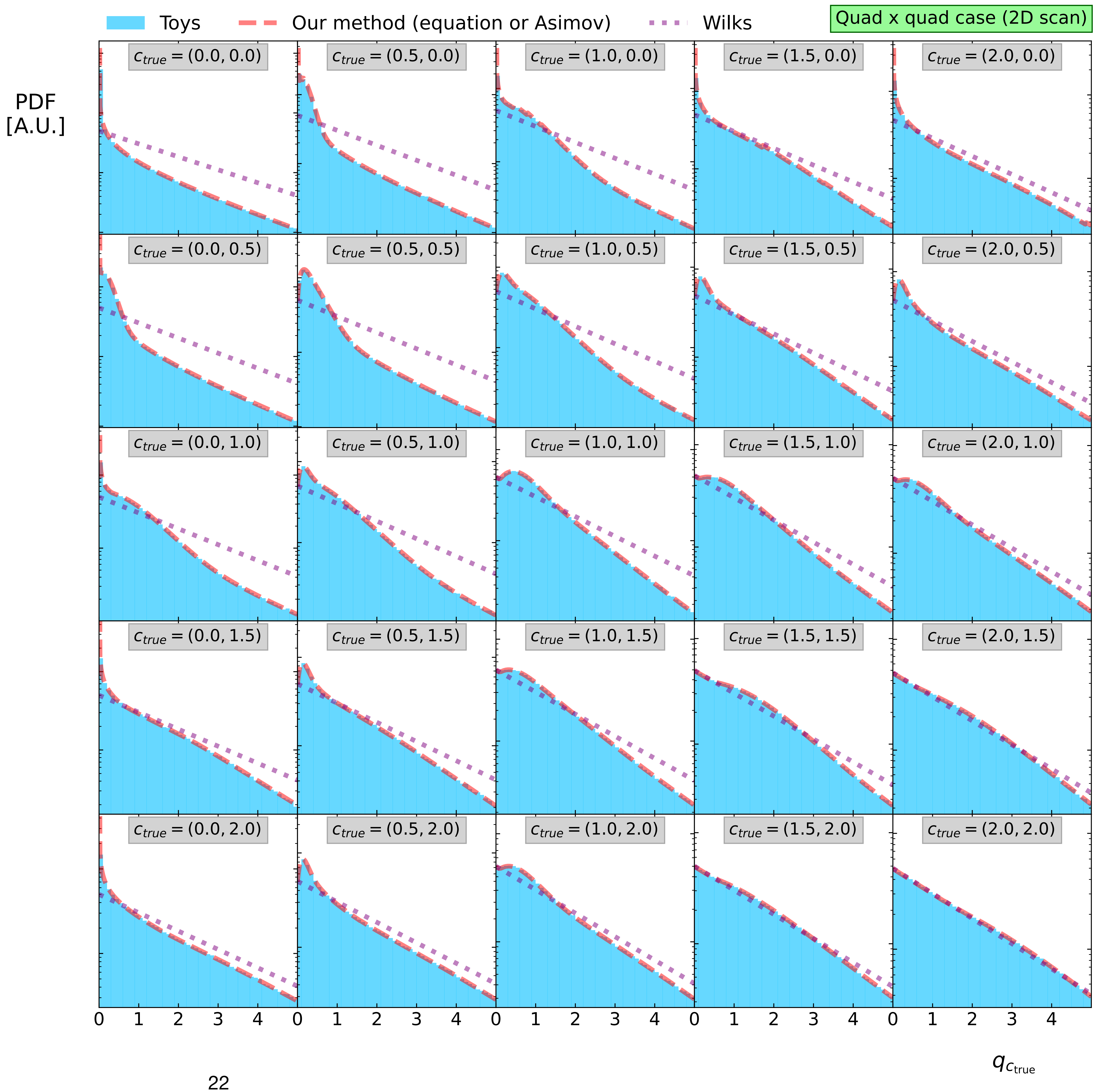
Integration contours (1D)



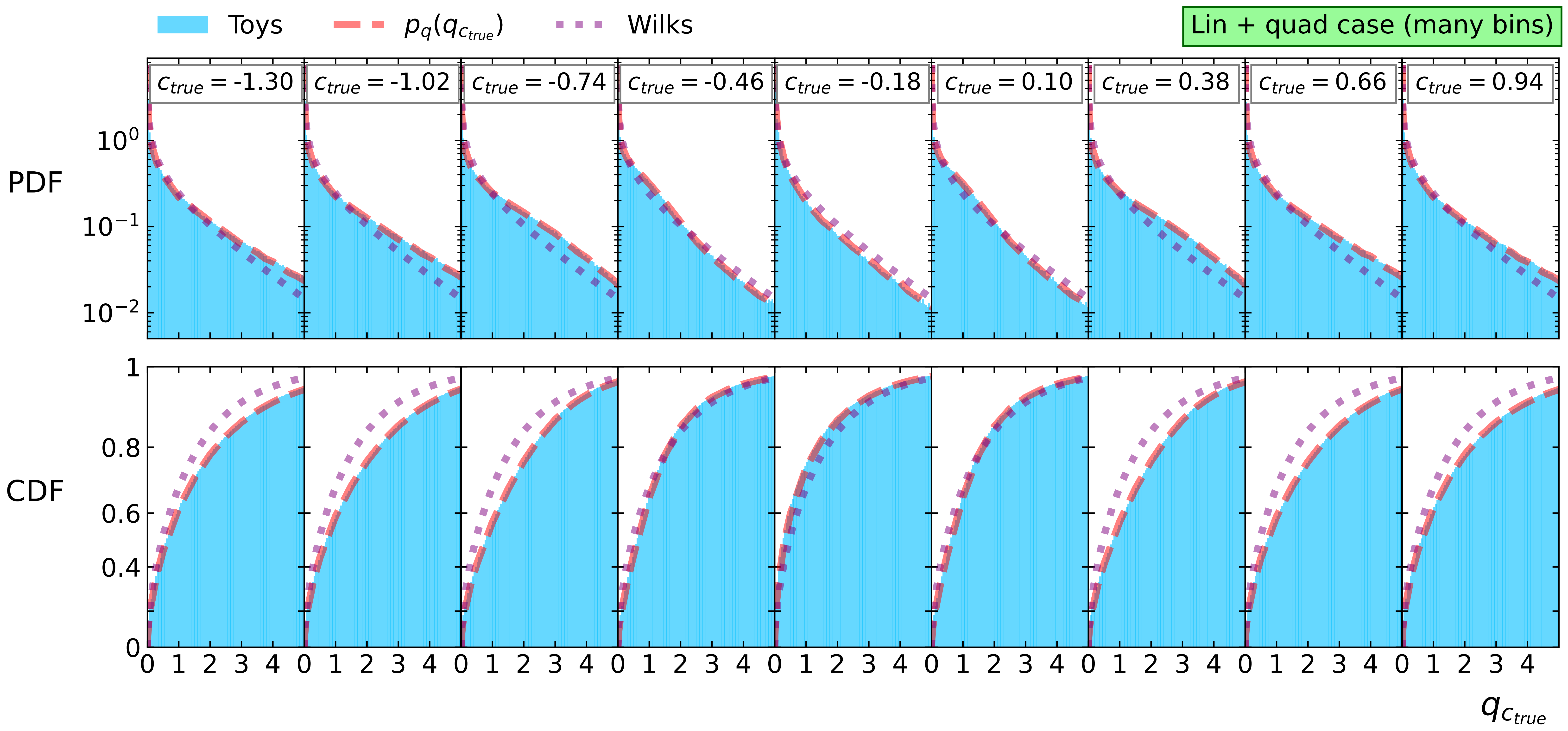
Lin x quad PDFs



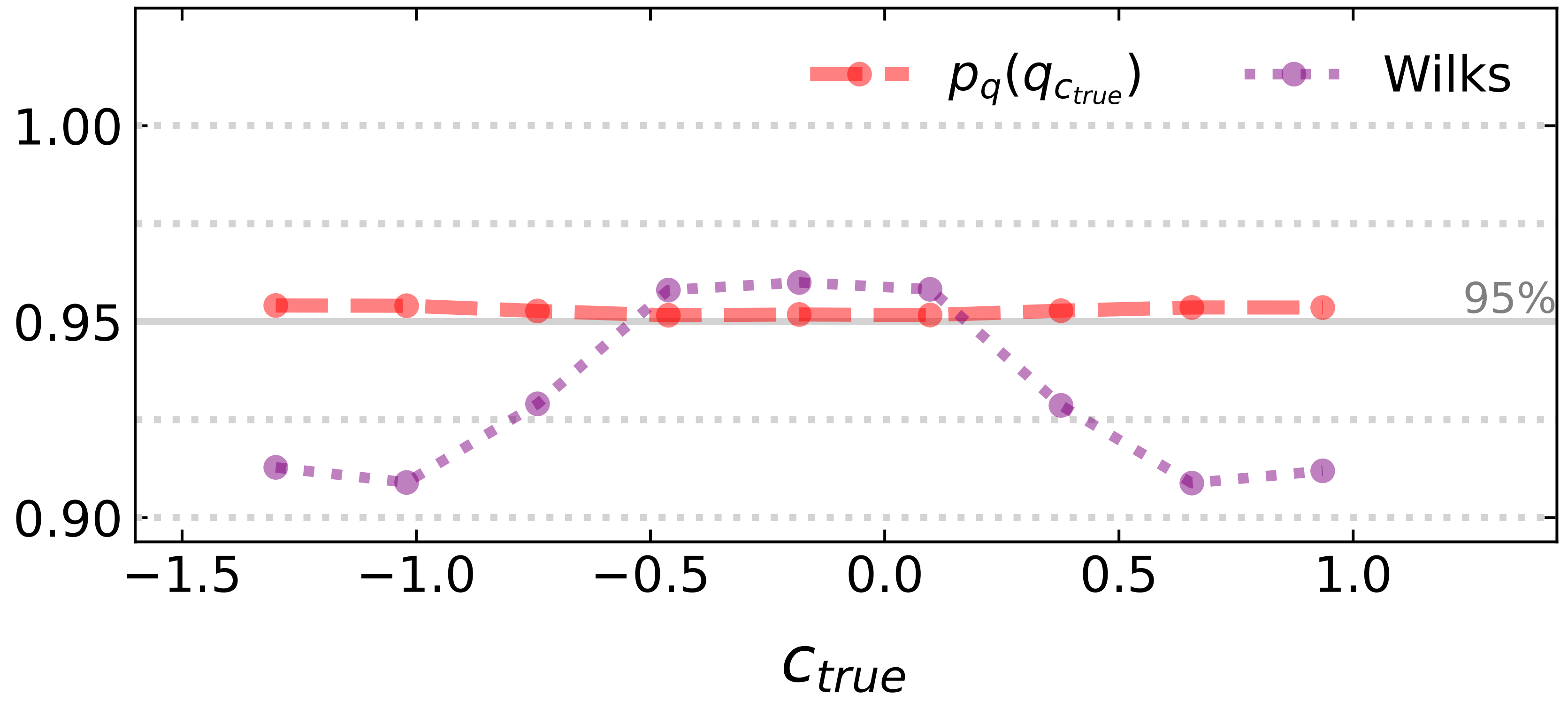
Quad x quad PDFs

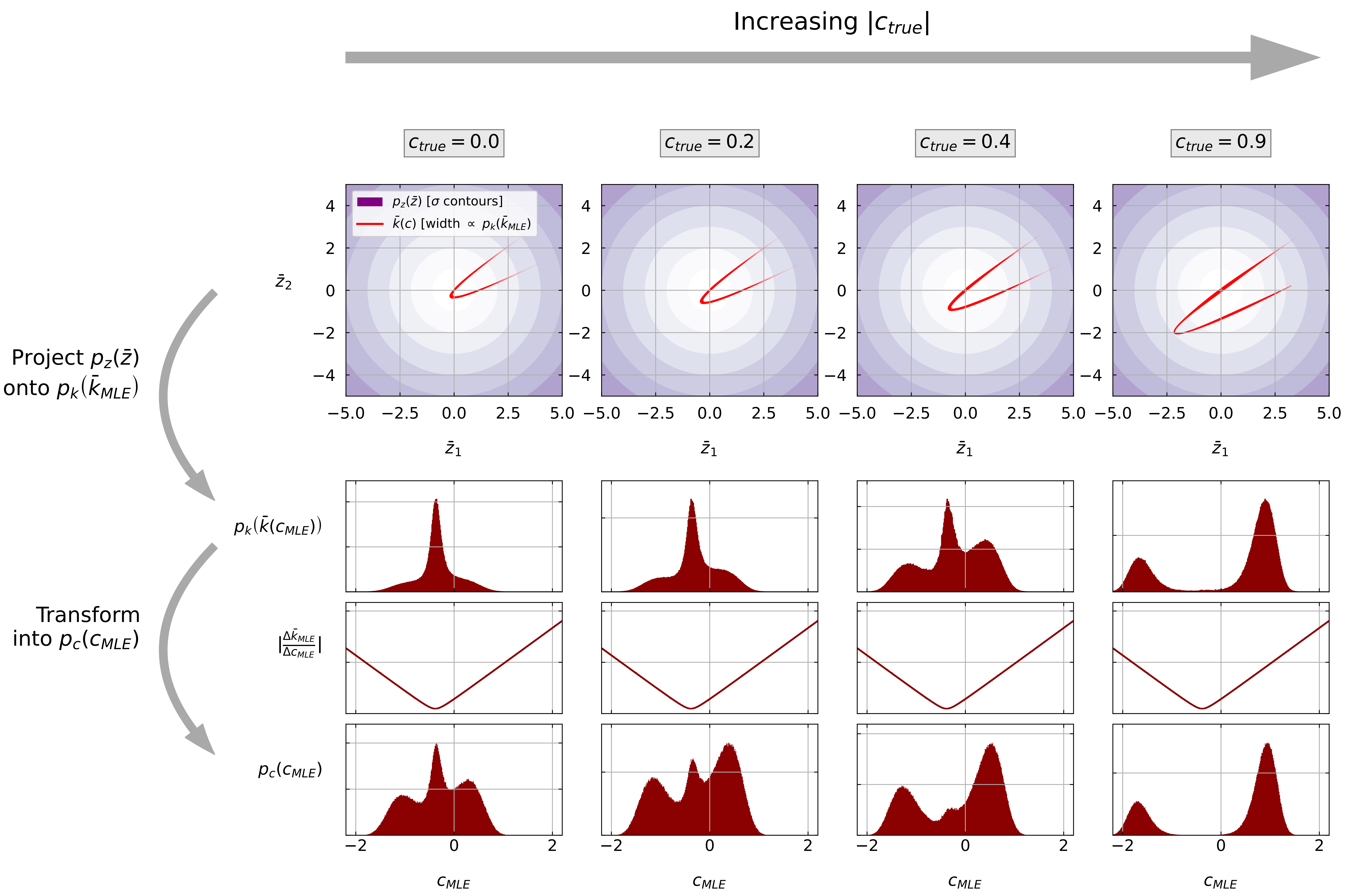


Lin + quad PDFs



Expected
95% coverage




Increasing $|c_{true}|$
$c_{true} = 0.0$
$c_{true} = 0.2$
$c_{true} = 0.4$
$c_{true} = 0.9$
$p_z(\bar{z})$ [σ contours]
$\bar{k}(c)$ [width $\propto$ $p_k(\bar{k}_{MLE})$]
$\bar{z}_2$
$\bar{z}_1$
Project $p_z(\bar{z})$ onto $p_k(\bar{k}_{MLE})$
Transform into $p_c(c_{MLE})$
$p_k(\bar{k}(c_{MLE}))$
$|\frac{\Delta \bar{k}_{MLE}}{\Delta c_{MLE}}|$
$p_c(c_{MLE})$
c_{MLE}