Unfolding: Point Estimation, Uncertainty Quantification and Future Directions

Mikael Kuusela

SAMSI / UNC Chapel Hill / Carnegie Mellon University

Statistical Methods for Physics Analysis in the 21st Century, QCHS18
Maynooth, Ireland

August 2, 2018

The unfolding problem

- Any differential cross section measurement is affected by the finite resolution of the particle detectors
  - This causes the observed spectrum of events to be "smeared" or "blurred" with respect to the true one
- The *unfolding problem* is to estimate the true spectrum using the smeared observations
- Ill-posed inverse problem with major methodological challenges

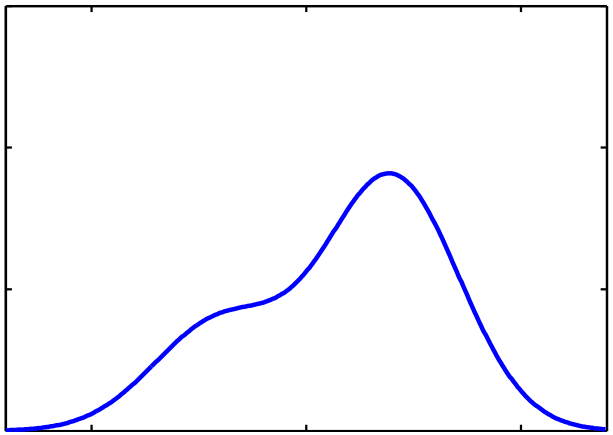
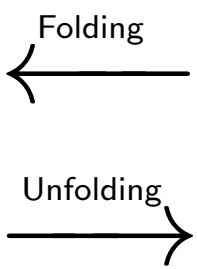


Figure: Smeared spectrum



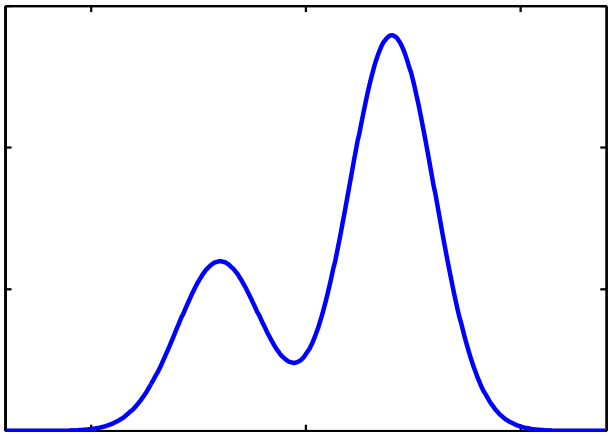


Figure: True spectrum

Problem formulation

- Let f be the true, particle-level spectrum and g the smeared, detector-level spectrum
  - Denote the true space by T and the smeared space by S (both taken to be intervals on the real line)
  - Mathematically f and g are the intensity functions of the underlying Poisson point process
- The two spectra are related by

$$g(s) = \int_T k(s,t) f(t)\, \mathrm{d}t,$$

where the smearing kernel k represents the response of the detector and is given by

$$k(s,t) = p(Y = s | X = t, X \text{ observed}) P(X \text{ observed} | X = t),$$

where X is a true event and Y the corresponding smeared event

Task: Infer the true spectrum f given smeared observations from g

Discretization

- Problem primarily discretized using histograms
- Let $\{T_i\}_{i=1}^{p}$ and $\{S_i\}_{i=1}^{n}$ be binnings of the true space T and the smeared space S
- Smeared histogram $\boldsymbol{y} = [y_1, \ldots, y_n]^{\mathrm{T}}$ with mean

$$\boldsymbol{\mu} = \left[\int_{S_1} g(s)\,\mathrm{d}s, \ldots, \int_{S_n} g(s)\,\mathrm{d}s\right]^{\mathrm{T}}$$

- Quantity of interest:

$$\boldsymbol{\lambda} = \left[\int_{T_1} f(t)\,\mathrm{d}t, \ldots, \int_{T_p} f(t)\,\mathrm{d}t\right]^{\mathrm{T}}$$

- The mean histograms are related by $\boldsymbol{\mu} = \boldsymbol{K}\boldsymbol{\lambda}$, where the elements of the *response matrix* $\boldsymbol{K}$ are given by

$$K_{i,j} = \frac{\int_{S_i}\int_{T_j} k(s,t) f(t)\,\mathrm{d}t\,\mathrm{d}s}{\int_{T_j} f(t)\,\mathrm{d}t} = P(\text{smeared event in bin } i \,|\, \text{true event in bin } j)$$

- The discretized statistical model becomes

$$\boldsymbol{y} \sim \mathrm{Poisson}(\boldsymbol{K}\boldsymbol{\lambda}),$$

where $\boldsymbol{K}$ is an ill-conditioned matrix

Discretization

- Problem primarily discretized using histograms
- Let $\{T_i\}_{i=1}^{p}$ and $\{S_i\}_{i=1}^{n}$ be binnings of the true space T and the smeared space S
- Smeared histogram $\boldsymbol{y} = [y_1, \ldots, y_n]^{\mathrm{T}}$ with mean

$$\boldsymbol{\mu} = \left[\int_{S_1} g(s)\,\mathrm{d}s, \ldots, \int_{S_n} g(s)\,\mathrm{d}s\right]^{\mathrm{T}}$$

- Quantity of interest:

$$\boldsymbol{\lambda} = \left[\int_{T_1} f(t)\,\mathrm{d}t, \ldots, \int_{T_p} f(t)\,\mathrm{d}t\right]^{\mathrm{T}}$$

- The mean histograms are related by $\boldsymbol{\mu} = \boldsymbol{K}\boldsymbol{\lambda}$, where the elements of the *response matrix* $\boldsymbol{K}$ are given by

$$K_{i,j} = \frac{\int_{S_i}\int_{T_j} k(s,t) f^{\mathrm{MC}}(t)\,\mathrm{d}t\,\mathrm{d}s}{\int_{T_j} f^{\mathrm{MC}}(t)\,\mathrm{d}t} = P(\text{smeared event in bin } i \,|\, \text{true event in bin } j)$$

- The discretized statistical model becomes

$$\boldsymbol{y} \sim \mathrm{Poisson}(\boldsymbol{K}\boldsymbol{\lambda}),$$

where $\boldsymbol{K}$ is an ill-conditioned matrix

Current unfolding methods

- Two main approaches (more information in the backup):
  1. Tikhonov regularization (i.e., SVD by Höcker and Kartvelishvili (1996) and TUnfold by Schmitt (2012)):

$$\min_{\boldsymbol{\lambda}\in\mathbb{R}^p} (\boldsymbol{y} - \boldsymbol{K}\boldsymbol{\lambda})^{\mathrm{T}} \hat{\boldsymbol{C}}^{-1} (\boldsymbol{y} - \boldsymbol{K}\boldsymbol{\lambda}) + \delta P(\boldsymbol{\lambda})$$

     with

$$P_{\mathrm{SVD}}(\boldsymbol{\lambda}) = \left\| \boldsymbol{L} \begin{bmatrix} \lambda_1/\lambda_1^{\mathrm{MC}} \\ \lambda_2/\lambda_2^{\mathrm{MC}} \\ \vdots \\ \lambda_p/\lambda_p^{\mathrm{MC}} \end{bmatrix} \right\|^2 \quad \text{or} \quad P_{\mathrm{TUnfold}}(\boldsymbol{\lambda}) = \|\boldsymbol{L}(\boldsymbol{\lambda} - \boldsymbol{\lambda}^{\mathrm{MC}})\|^2,$$

     where $\boldsymbol{L}$ is usually the discretized second derivative (also other choices possible)
  2. Expectation-maximization iteration with early stopping (D'Agostini, 1995):

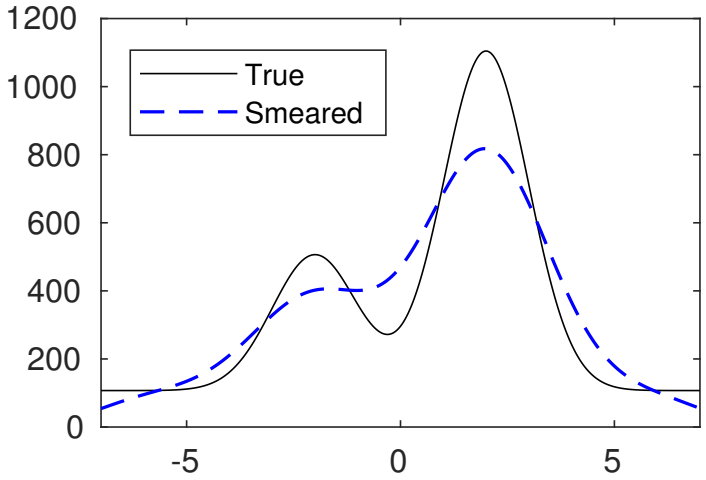
$$\lambda_j^{(t+1)} = \frac{\lambda_j^{(t)}}{\sum_{i=1}^n K_{i,j}} \sum_{i=1}^n \frac{K_{i,j} y_i}{\sum_{k=1}^p K_{i,k} \lambda_k^{(t)}}, \text{ with } \boldsymbol{\lambda}^{(0)} = \boldsymbol{\lambda}^{\mathrm{MC}}$$

- All these methods typically regularize by biasing towards a MC ansatz $\boldsymbol{\lambda}^{\mathrm{MC}}$
- Regularization strength controlled by the choice of δ in Tikhonov or by the number of iterations in D'Agostini
- Uncertainty quantification: $[\underline{\lambda}_i, \overline{\lambda}_i] = \left[\hat{\lambda}_i - z_{1-\alpha/2}\sqrt{\widehat{\mathrm{var}}\left(\hat{\lambda}_i\right)},\ \hat{\lambda}_i + z_{1-\alpha/2}\sqrt{\widehat{\mathrm{var}}\left(\hat{\lambda}_i\right)}\right]$, with $\widehat{\mathrm{var}}\left(\hat{\lambda}_i\right)$ estimated using error propagation or resampling

Choice of the regularization strength

- A key issue in unfolding is the choice of the regularization strength (δ in Tikhonov, # of iterations in D'Agostini)
  - The solution and especially the uncertainties depend heavily on this choice
- This choice should be done using an objective data-driven criterion
  - In particular, one must not rely on the software defaults for the regularization strength
- Many data-driven methods have been proposed:
  1. (Weighted/generalized) cross-validation (e.g., Green and Silverman, 1994)
  2. L-curve (Hansen, 1992)
  3. Empirical Bayes estimation (Kuusela and Panaretos, 2015)
  4. Goodness-of-fit test in the smeared space (Veklerov and Llacer, 1987)
  5. Akaike information criterion (Volobouev, 2015)
  6. Minimization of a global correlation coefficient (Schmitt, 2012)
  7. ...
- Limited experience about the relative merits of these in typical unfolding problems
- Important note: *All of these are designed for point estimation!*
  - Not necessarily optimal for uncertainty quantification

Simulation setup



$$f(t) = \lambda_{\text{tot}} \left\{ \pi_1 \mathcal{N}(t|-2,1) + \pi_2 \mathcal{N}(t|2,1) + \pi_3 \frac{1}{|E|} \right\}$$

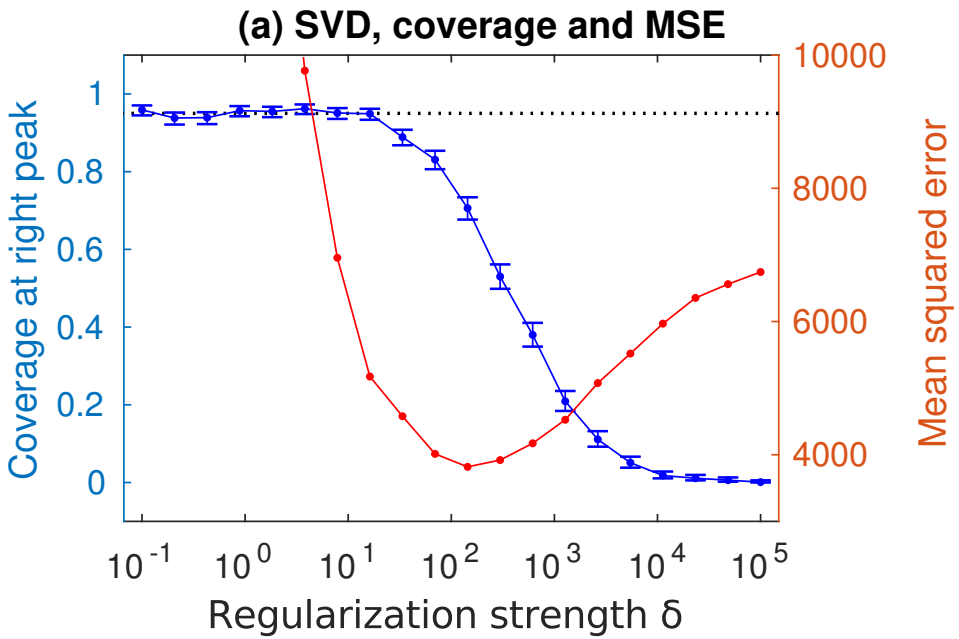
$$g(s) = \int_T \mathcal{N}(s-t|0,1) f(t)\, \mathrm{d}t$$

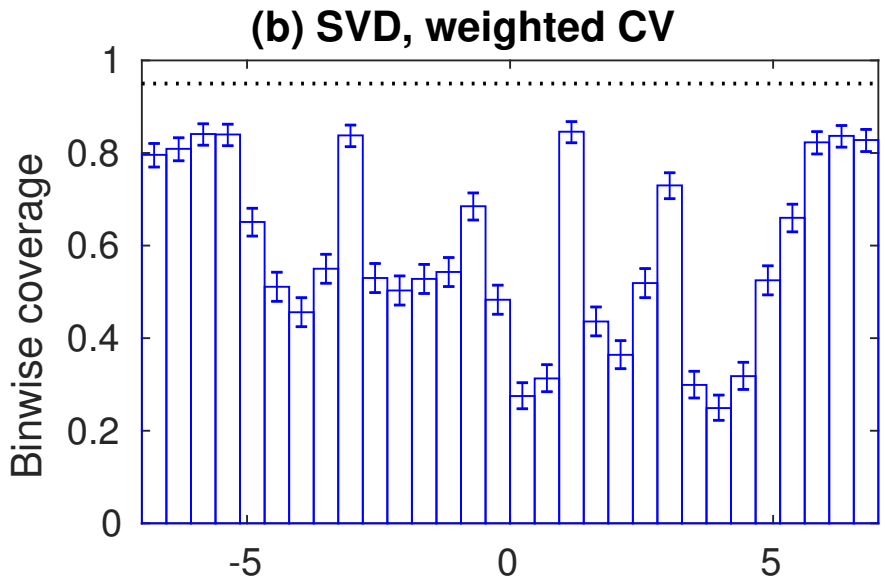
$$f^{\text{MC}}(t) = \lambda_{\text{tot}} \left\{ \pi_1 \mathcal{N}(t|-2,1.1^2) + \pi_2 \mathcal{N}(t|2,0.9^2) + \pi_3 \frac{1}{|E|} \right\}$$

Will focus on the following question: Do the unfolded confidence intervals have the advertised coverage probability

$$\mathrm{P}\left(\lambda_i \in \left[\underline{\lambda}_i(\boldsymbol{y}), \overline{\lambda}_i(\boldsymbol{y})\right]\right) \approx 1 - \alpha$$

Undercoverage of existing methods





- Optimal point estimation $\neq$ optimal uncertainty quantification
  - In terms of the uncertainties, standard methods for choosing δ tend to regularize too heavily
- Similar conclusions hold for other common methods (D'Agostini, TUnfold,...)

Undersmoothed unfolding

- A simple way to improve the coverage is to reduce the regularization strength δ from the value that is optimal for point estimation
  - In other words, adjust the bias-variance trade-off to the direction of less bias and more variance
- I introduced in Kuusela (2016) a data-driven technique for deciding *how much* one should undersmooth

Outline of undersmoothed UQ for unfolding

1. Choose a pilot estimate of δ using one of the standard data-driven methods (cross-validation, empirical Bayes, L-curve,...)
2. Reduce δ until intervals in all bins have estimated coverage greater than $1 - \alpha - \varepsilon$, for some small tolerance ε

- Coverage estimated using a hybrid plug-in approach; see Kuusela (2016)
- I have recently been working with Lyle Kim (a statistics student at UChicago) to implement undersmoothing as an extension of `TUnfold` V17.6
- The code is available at:

`https://github.com/lylejkim/UndersmoothedUnfolding`

Unfolded histograms, $\boldsymbol{\lambda}^{\mathrm{MC}} = 0$

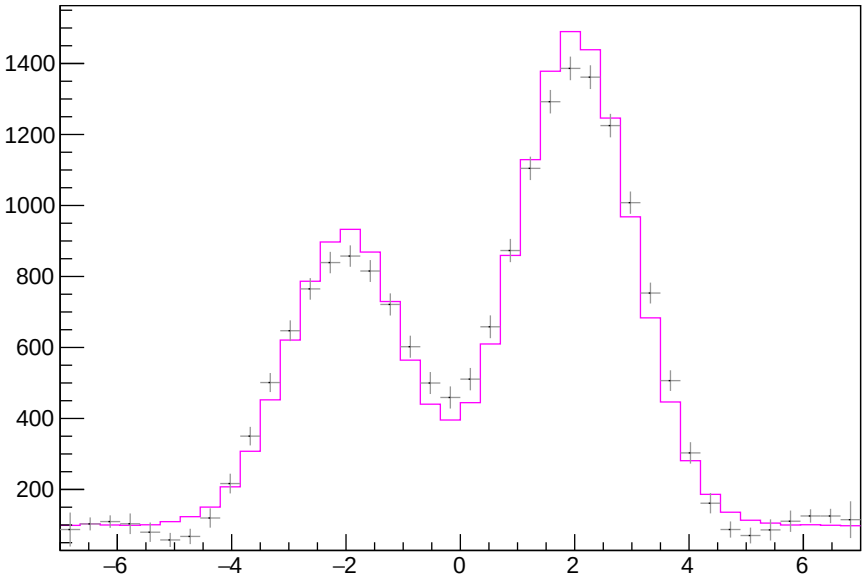


Figure: L-curve, $\tau = \sqrt{\delta}$=0.01186

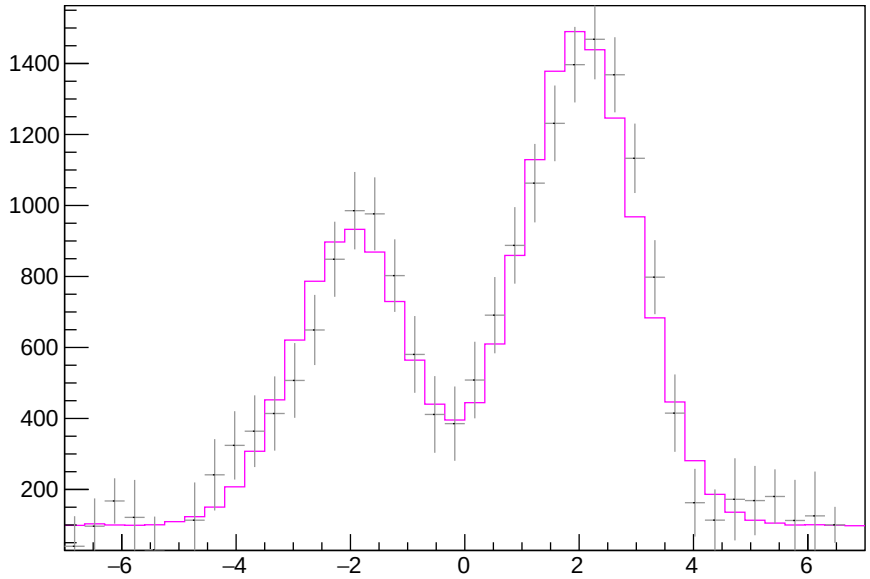


Figure: Undersmoothing, $\tau = \sqrt{\delta}$=0.00177

Binwise coverage, $\boldsymbol{\lambda}^{\mathrm{MC}} = 0$

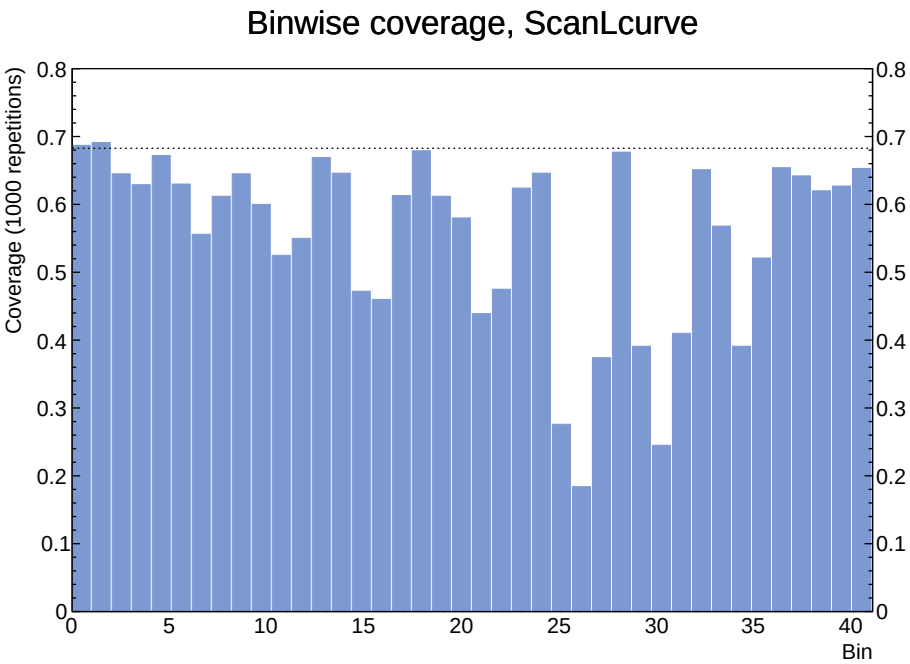


Figure: L-curve

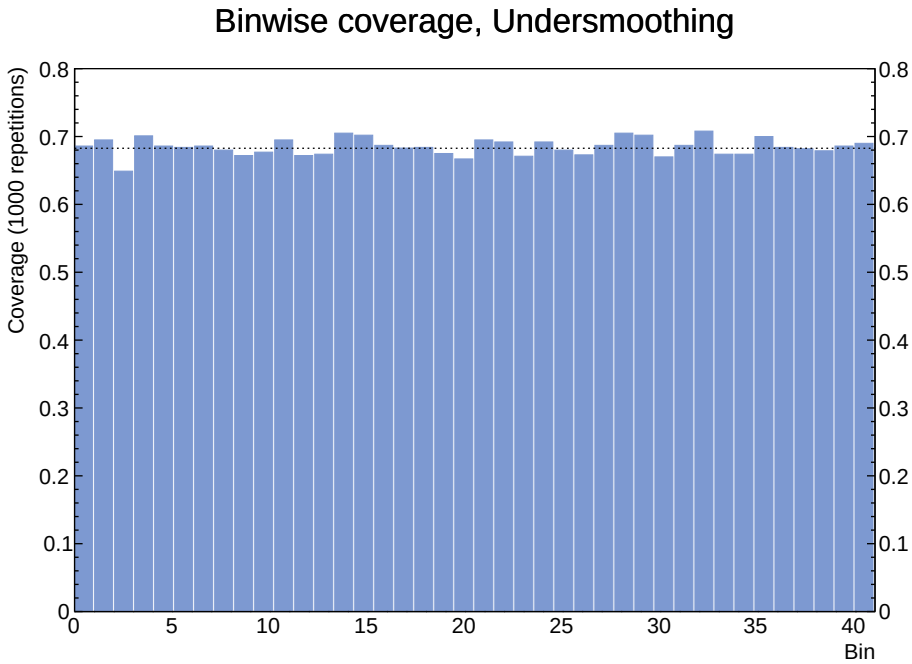


Figure: Undersmoothing

Future directions

- Undersmoothing can provide unfolded confidence intervals with reasonable length and coverage
  - Further improvement provided by iterative bias-corrections (Kuusela, 2016)
  - These debiasing approaches work in a wide variety of cases, but they do eventually break down in challenging enough scenarios (very low sample size, lots of smearing, f has difficult shape,...)
- An alternative approach is to regularize using shape constraints (Kuusela and Stark, 2017)
  - This approach yields guaranteed coverage if something is known about the shape of the true spectrum (monotonicity, convexity, unimodality,...)
- At the end of the day, *any regularization technique makes unverifiable assumptions about the true spectrum*
  - If these assumptions are not satisfied, the uncertainties will be wrong
- It seems to me that the fundamental problem is that we are asking too hard questions about the true spectrum
  - One simply cannot recover extremely detailed information about f without further outside knowledge
- So the question becomes: What features of f can be recovered based on the smeared data $\boldsymbol{y}$ and how to do this with *honest unregularized* uncertainties?

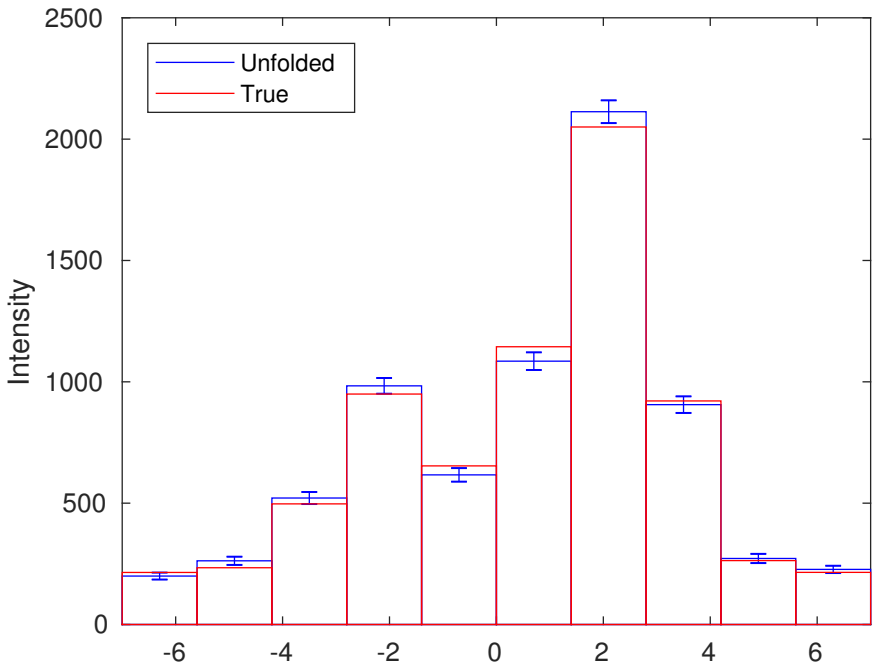
Wide bin unfolding?

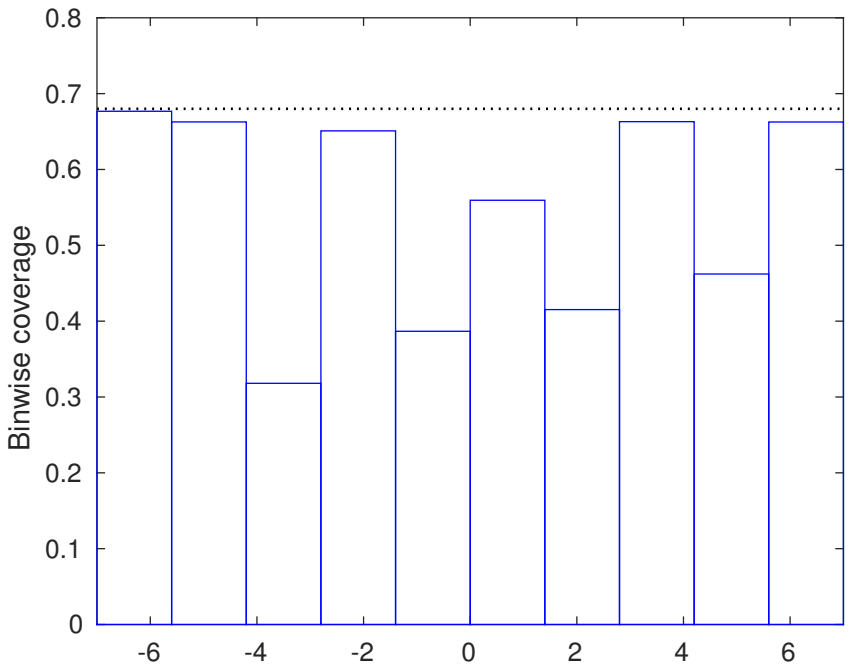
- One functional we should be able to recover without explicit regularization is the integral of f over a *wide* unfolded bin:

$$H[f] = \int_{T_j} f(t)\,\mathrm{d}t, \quad \text{width of } T_j \text{ large}$$

- But one cannot simply arbitrarily increase the particle-level bin size in the conventional approaches, since this increases the MC dependence of $\boldsymbol{K}$
- To circumvent this, *it is possible to first unfold with fine bins and then aggregate into wide bins*
- Let's see how this works!
  - Simulation setup: $\hat{\boldsymbol{\lambda}} = \boldsymbol{K}^{\dagger}\boldsymbol{y}$, convolution kernel $\mathcal{N}(0, 0.35^2)$, slightly different f^{MC}, otherwise as before

Wide bins, standard approach, perturbed MC

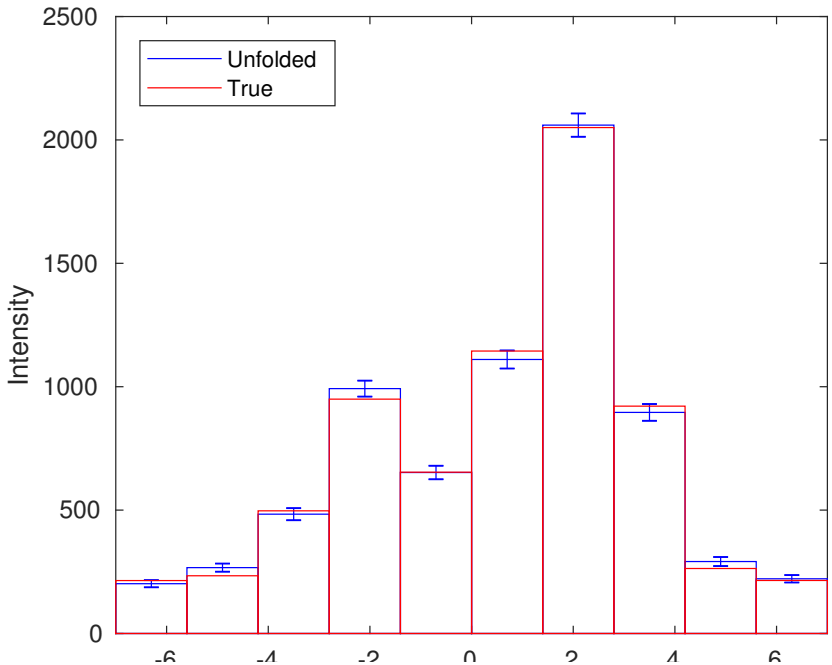


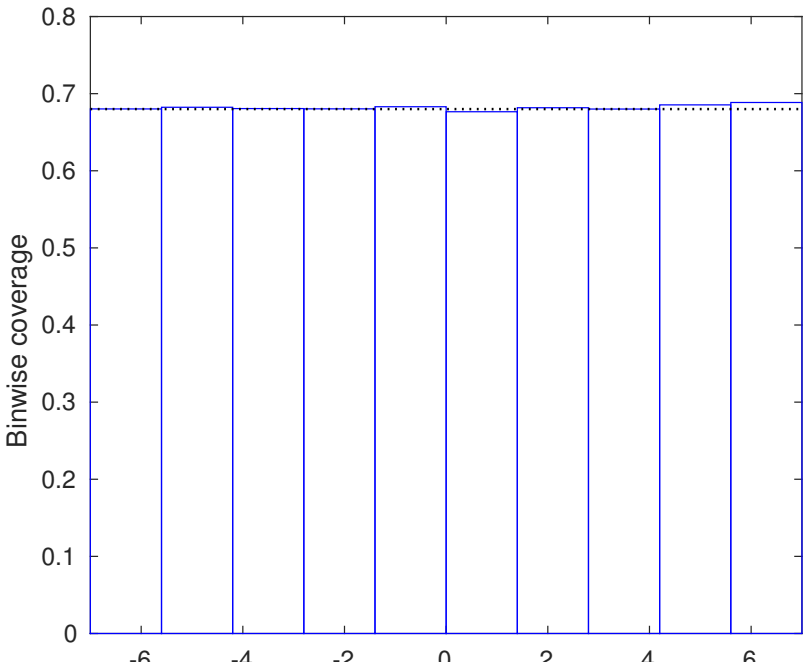


The response matrix $K_{i,j} = \frac{\int_{S_i}\int_{T_j} k(s,t) f^{\mathrm{MC}}(t)\,\mathrm{d}t\,\mathrm{d}s}{\int_{T_j} f^{\mathrm{MC}}(t)\,\mathrm{d}t}$ depends on f^{MC}

$\Rightarrow$ Undercoverage if $f^{\mathrm{MC}} \neq f$

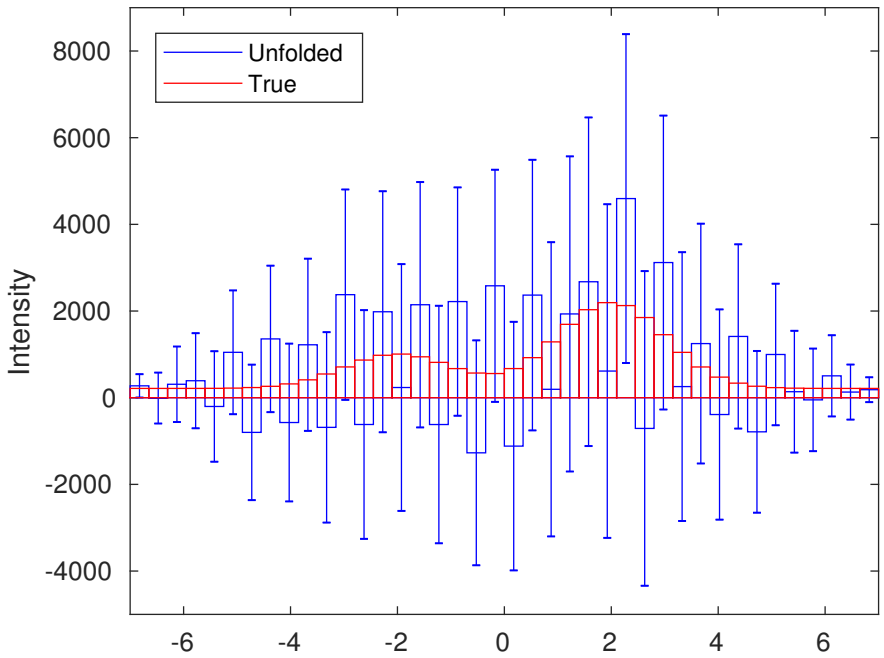
Wide bins, standard approach, correct MC

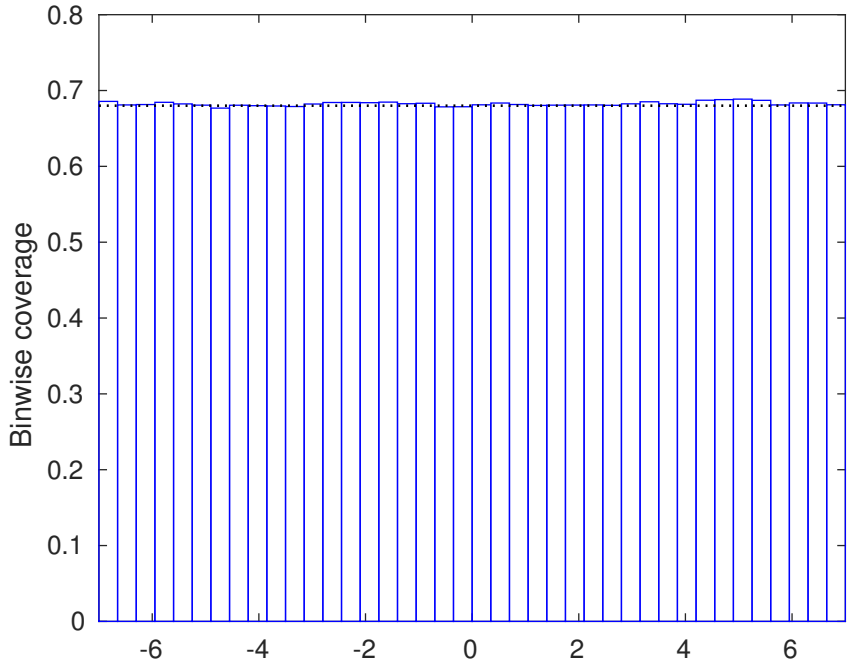




If $f^{\mathrm{MC}} = f$, coverage is correct

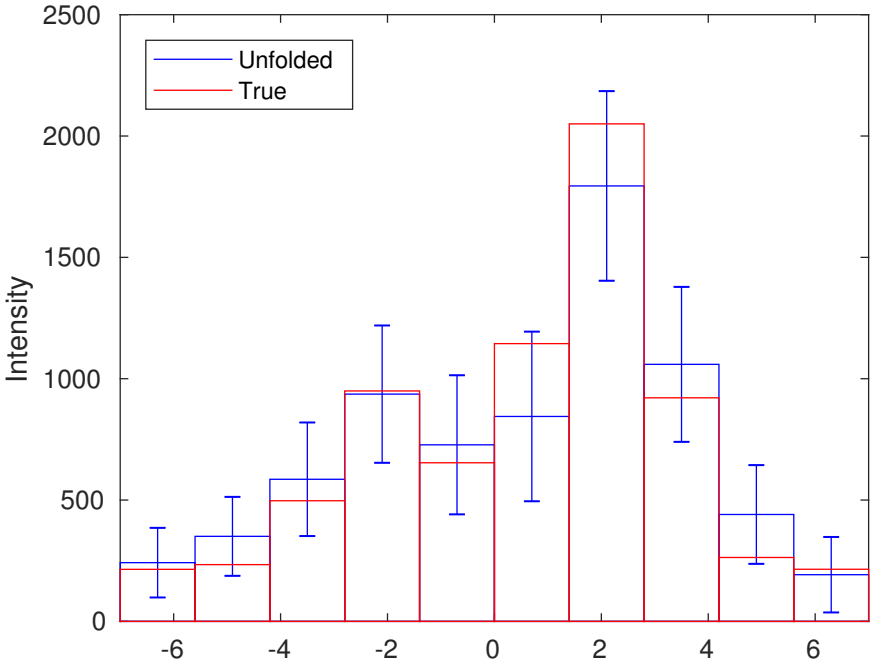
Fine bins, standard approach, perturbed MC

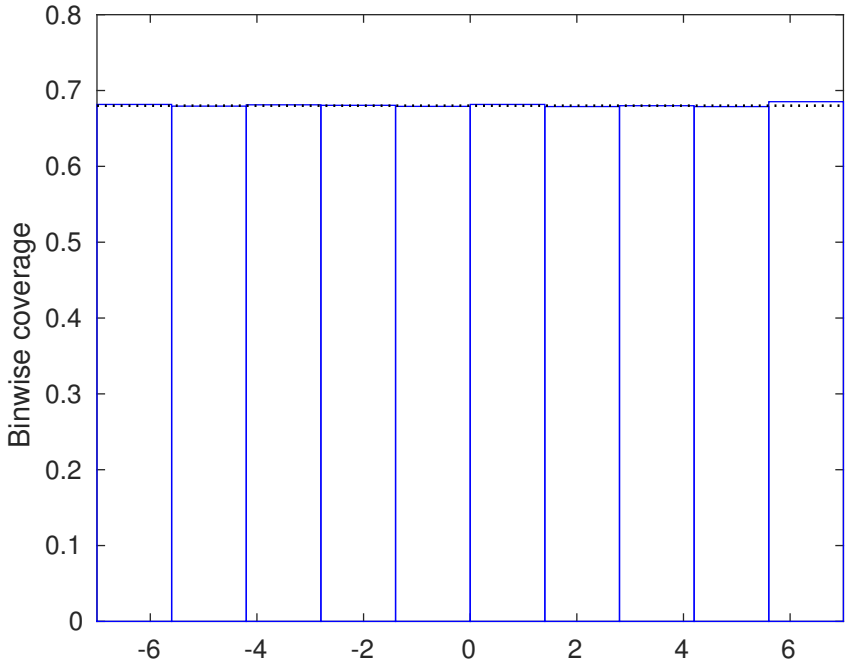




With narrow bins, less dependence on f^{MC} so coverage is correct, but the intervals are very wide

Wide bins via fine bins, perturbed MC





Wide bins via fine bins gives both correct coverage and intervals with reasonable length

Conclusions

- Coverage is a useful criterion for validating, optimizing and comparing unfolding methods
- Standard methods can have drastically lower coverage than expected, unless care is taken in the choice of the regularization strength
- Optimal point estimation $\neq$ optimal uncertainty quantification
- Undersmoothing (which is now available for ROOT) provides one way of obtaining intervals with both reasonable coverage and reasonable length
- Further progress is likely to involve a major rethinking of the role of regularization
  - Any regularization really means some amount of "cheating" on the uncertainties
  - We should probably think of ways to provide unregularized uncertainties on relevant, well-chosen functionals of f, such as integral over wide bins
  - A simple step in this direction is to first unfold with narrow bins and then compute the functionals

References I

T. Adye. Unfolding algorithms and tests using RooUnfold. In H. B. Prosper and L. Lyons, editors, *Proceedings of the PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding*, CERN-2011-006, pages 313–318, CERN, Geneva, Switzerland, 17–20 January 2011.

V. Blobel. Unfolding methods in high-energy physics experiments. In *CERN Yellow Report 85-09*, pages 88–127, 1985.

V. Blobel, The RUN manual: Regularized unfolding for high-energy physics experiments, OPAL Technical Note TN361, 1996.

G. Choudalakis. Fully Bayesian unfolding. arXiv:1201.4612v4 [physics.data-an], 2012.

G. D'Agostini, A multidimensional unfolding method based on Bayes' theorem, *Nuclear Instruments and Methods A*, 362:487–498, 1995.

A. P. Dempster, N. M. Laird, and D. B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.

P. J. Green and B. W. Silverman. *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman & Hall, 1994.

References II

P. Hall, Effect of bias estimation on coverage accuracy of bootstrap confidence intervals for a probability density, *The Annals of Statistics*, 20(2):675–694, 1992.

P. C. Hansen, Analysis of discrete ill-posed problems by means of the L-curve, *SIAM Review*, 34(4):561–580, 1992.

A. Höcker and V. Kartvelishvili, SVD approach to data unfolding, *Nuclear Instruments and Methods in Physics Research A*, 372:469–481, 1996.

M. Kuusela. *Uncertainty quantification in unfolding elementary particle spectra at the Large Hadron Collider*. PhD thesis, EPFL, 2016. Available online at: `https://infoscience.epfl.ch/record/220015`.

M. Kuusela and V. M. Panaretos, Statistical unfolding of elementary particle spectra: Empirical Bayes estimation and bias-corrected uncertainty quantification, *The Annals of Applied Statistics*, 9(3):1671–1705, 2015.

M. Kuusela and P. B. Stark, Shape-constrained uncertainty quantification in unfolding steeply falling elementary particle spectra, *The Annals of Applied Statistics*, 11(3): 1671–1710, 2017.

K. Lange and R. Carson, EM reconstruction algorithms for emission and transmission tomography, *Journal of Computer Assisted Tomography*, 8(2):306–316, 1984.

References III

L. B. Lucy, An iterative technique for the rectification of observed distributions, *Astronomical Journal*, 79(6):745–754, 1974.

B. Malaescu. An iterative, dynamically stabilized (IDS) method of data unfolding. In H. B. Prosper and L. Lyons, editors, *Proceedings of the PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding*, CERN-2011-006, pages 271–275, CERN, Geneva, Switzerland, 17–20 January 2011.

N. Milke, M. Doert, S. Klepser, D. Mazin, V. Blobel, and W. Rhode, Solving inverse problems with the unfolding program TRUEE: Examples in astroparticle physics, *Nuclear Instruments and Methods in Physics Research A*, 697:133–147, 2013.

W. H. Richardson, Bayesian-based iterative method of image restoration, *Journal of the Optical Society of America*, 62(1):55–59, 1972.

S. Schmitt, TUnfold, an algorithm for correcting migration effects in high energy physics, *Journal of Instrumentation*, 7:T10003, 2012.

L. A. Shepp and Y. Vardi, Maximum likelihood reconstruction for emission tomography, *IEEE Transactions on Medical Imaging*, 1(2):113–122, 1982.

Y. Vardi, L. A. Shepp, and L. Kaufman, A statistical model for positron emission tomography, *Journal of the American Statistical Association*, 80(389):8–20, 1985.

References IV

E. Veklerov and J. Llacer, Stopping rule for the MLE algorithm based on statistical hypothesis testing, *IEEE Transactions on Medical Imaging*, 6(4):313–319, 1987.

I. Volobouev. On the expectation-maximization unfolding with smoothing. arXiv:1408.6500v2 [physics.data-an], 2015.

Backup

Current methodology

Two main approaches to unfolding:

1. Tikhonov regularization (Höcker and Kartvelishvili, 1996; Schmitt, 2012)
2. Expectation-maximization iteration with early stopping (D'Agostini, 1995; Richardson, 1972; Lucy, 1974; Shepp and Vardi, 1982; Lange and Carson, 1984; Vardi et al., 1985)

Tikhonov regularization

- Tikhonov regularization estimates $\boldsymbol{\lambda}$ by solving:

$$\min_{\boldsymbol{\lambda}\in\mathbb{R}^p} (\boldsymbol{y} - \boldsymbol{K}\boldsymbol{\lambda})^{\mathrm{T}} \hat{\boldsymbol{C}}^{-1} (\boldsymbol{y} - \boldsymbol{K}\boldsymbol{\lambda}) + \delta P(\boldsymbol{\lambda})$$

- The first term as a Gaussian approximation to the Poisson log-likelihood
- The second term penalizes physically implausible solutions
- Common penalty terms:
  - Norm: $P(\boldsymbol{\lambda}) = \|\boldsymbol{\lambda}\|^2$
  - Curvature: $P(\boldsymbol{\lambda}) = \|\mathbf{L}\boldsymbol{\lambda}\|^2$, where $\mathbf{L}$ is a discretized 2nd derivative operator
  - SVD unfolding (Höcker and Kartvelishvili, 1996):

$$P(\boldsymbol{\lambda}) = \left\| \mathbf{L} \begin{bmatrix} \lambda_1/\lambda_1^{\mathrm{MC}} \\ \lambda_2/\lambda_2^{\mathrm{MC}} \\ \vdots \\ \lambda_p/\lambda_p^{\mathrm{MC}} \end{bmatrix} \right\|^2,$$

    where $\boldsymbol{\lambda}^{\mathrm{MC}}$ is a MC prediction for $\boldsymbol{\lambda}$
  - `TUnfold`[1] (Schmitt, 2012): $P(\boldsymbol{\lambda}) = \|\mathbf{L}(\boldsymbol{\lambda} - \boldsymbol{\lambda}^{\mathrm{MC}})\|^2$

[1] `TUnfold` implements also more general penalty terms

D'Agostini iteration

- Starting from some initial guess $\boldsymbol{\lambda}^{(0)} > \mathbf{0}$, iterate

$$\lambda_j^{(k+1)} = \frac{\lambda_j^{(k)}}{\sum_{i=1}^{n} K_{i,j}} \sum_{i=1}^{n} \frac{K_{i,j} y_i}{\sum_{l=1}^{p} K_{i,l} \lambda_l^{(k)}}$$

- Regularization by stopping the iteration before convergence:
  - $\hat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^{(K)}$ for some small number of iterations K
  - I.e., bias the solution towards $\boldsymbol{\lambda}^{(0)}$
  - Regularization strength controlled by the choice of K
- In `RooUnfold` (Adye, 2011), $\boldsymbol{\lambda}^{(0)} = \boldsymbol{\lambda}^{\mathrm{MC}}$

D'Agostini iteration

$$\lambda_j^{(k+1)} = \frac{\lambda_j^{(k)}}{\sum_{i=1}^{n} K_{i,j}} \sum_{i=1}^{n} \frac{K_{i,j} y_i}{\sum_{l=1}^{p} K_{i,l} \lambda_l^{(k)}}$$

- This iteration has been discovered in various fields, including optics (Richardson, 1972), astronomy (Lucy, 1974) and tomography (Shepp and Vardi, 1982; Lange and Carson, 1984; Vardi et al., 1985)
- In particle physics, it was popularized by D'Agostini (1995) who called it "Bayesian" unfolding
- But: This is in fact an expectation-maximization (EM) iteration (Dempster et al., 1977) for finding the *maximum likelihood estimator* of $\boldsymbol{\lambda}$ in the Poisson regression problem $\boldsymbol{y} \sim \mathrm{Poisson}(\boldsymbol{K}\boldsymbol{\lambda})$
- As $k \to \infty$, $\boldsymbol{\lambda}^{(k)} \to \hat{\boldsymbol{\lambda}}_{\mathrm{MLE}}$ (Vardi et al., 1985)
- *This is a fully frequentist technique for finding the (regularized) MLE*
  - The name "Bayesian" is an unfortunate misnomer

D'Agostini demo, $k = 0$

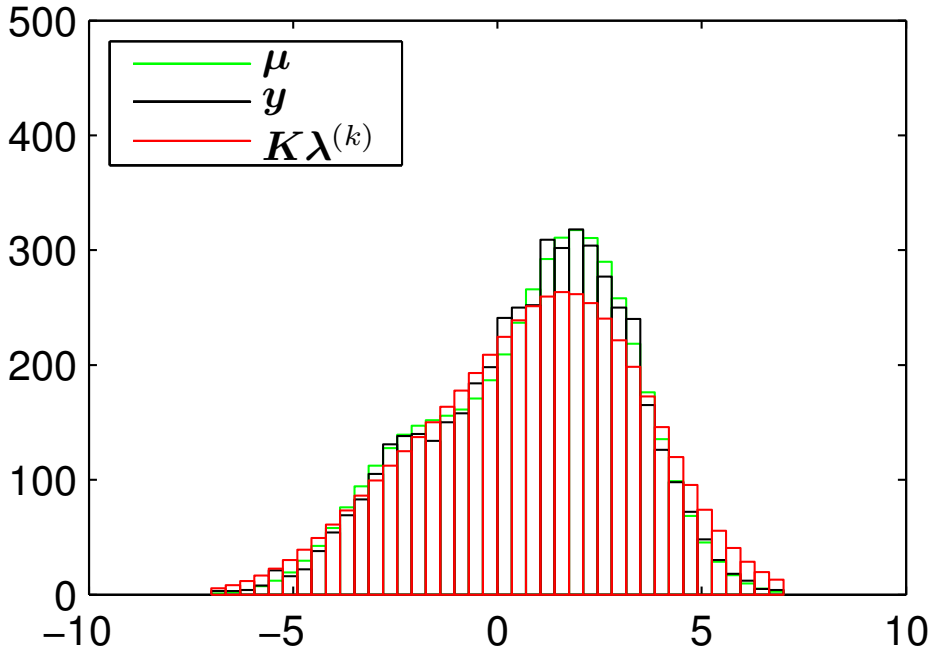


Figure: Smeared histogram

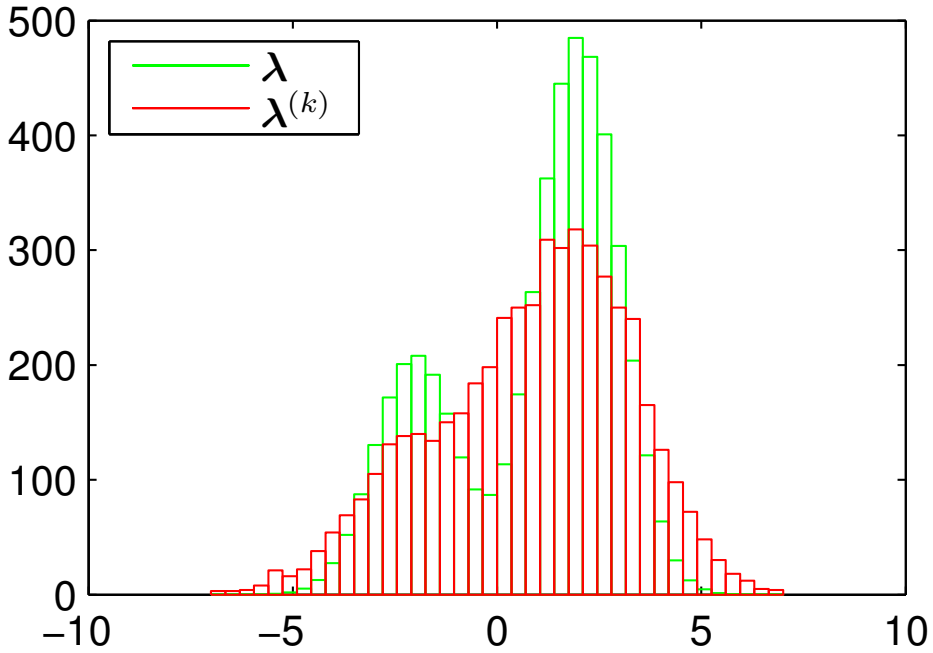


Figure: True histogram

D'Agostini demo, $k = 100$

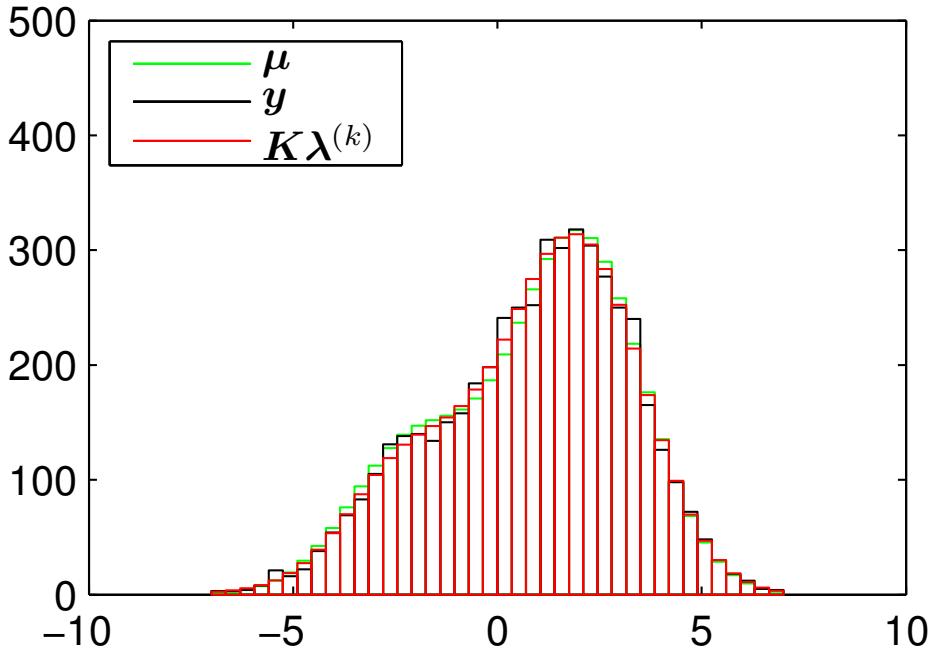


Figure: Smeared histogram

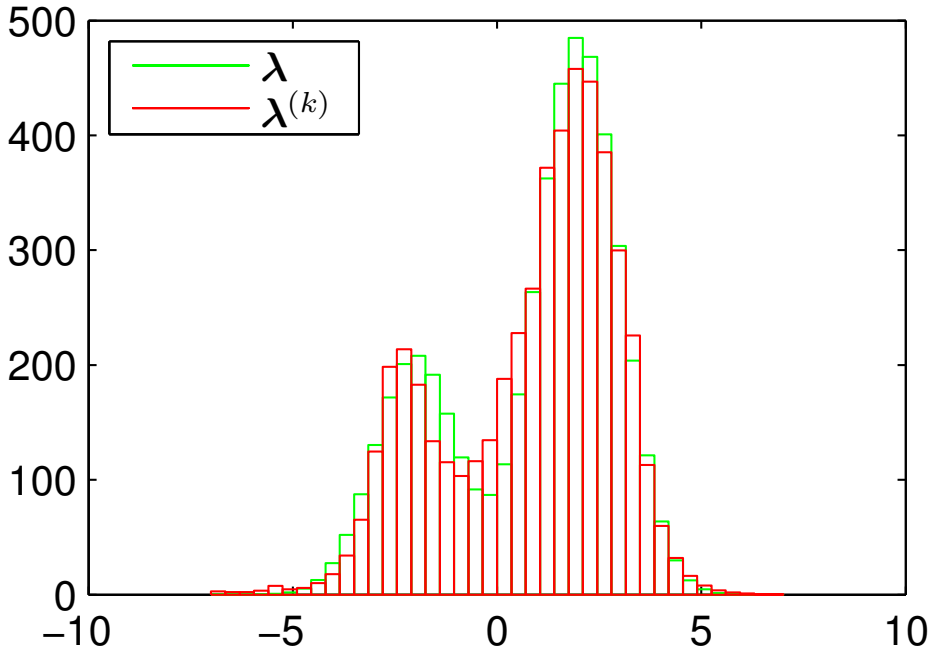


Figure: True histogram

D'Agostini demo, $k = 10000$

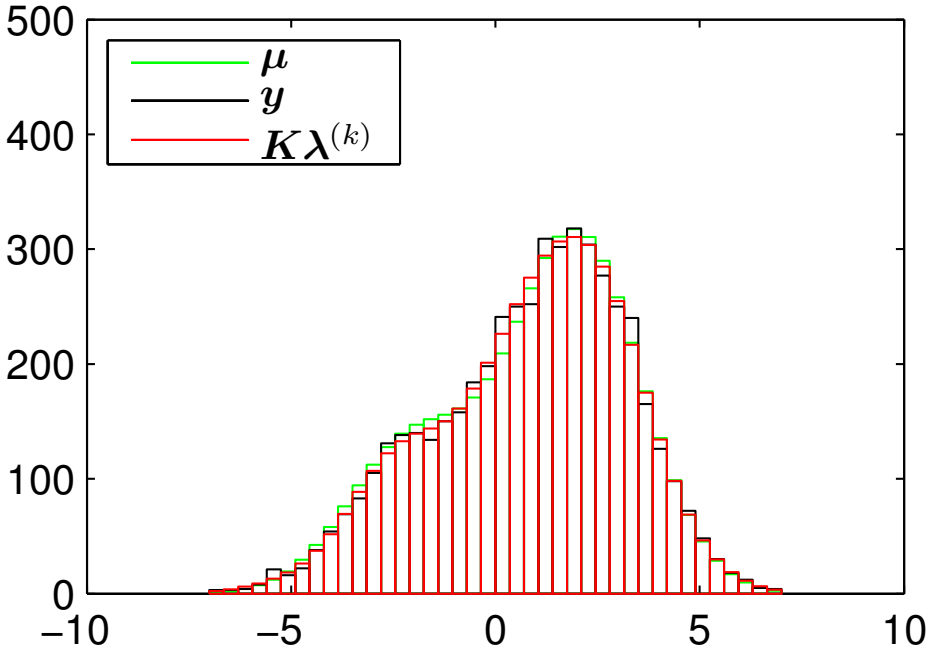


Figure: Smeared histogram

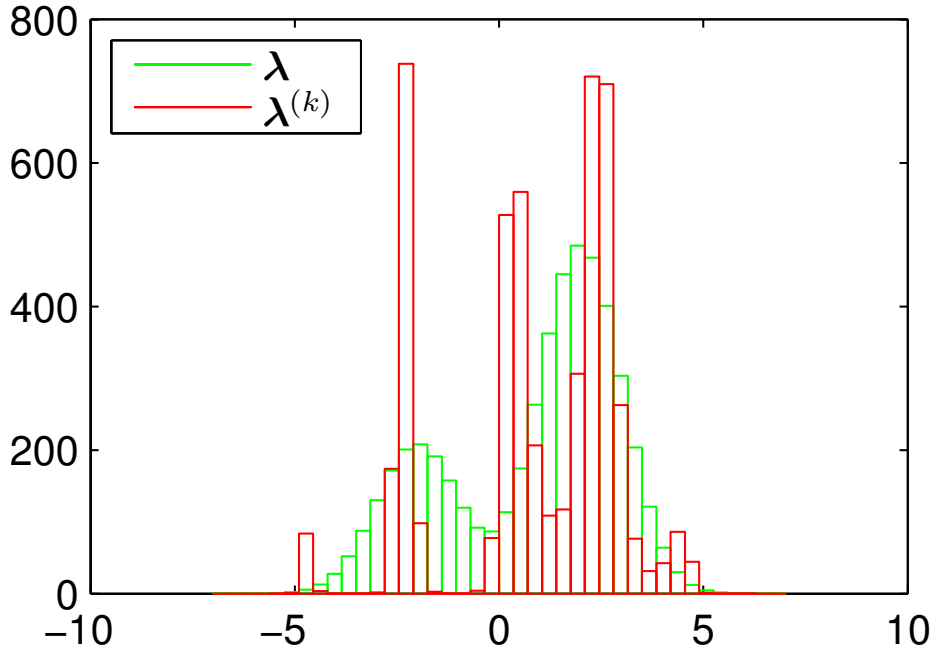


Figure: True histogram

D'Agostini demo, $k = 100000$

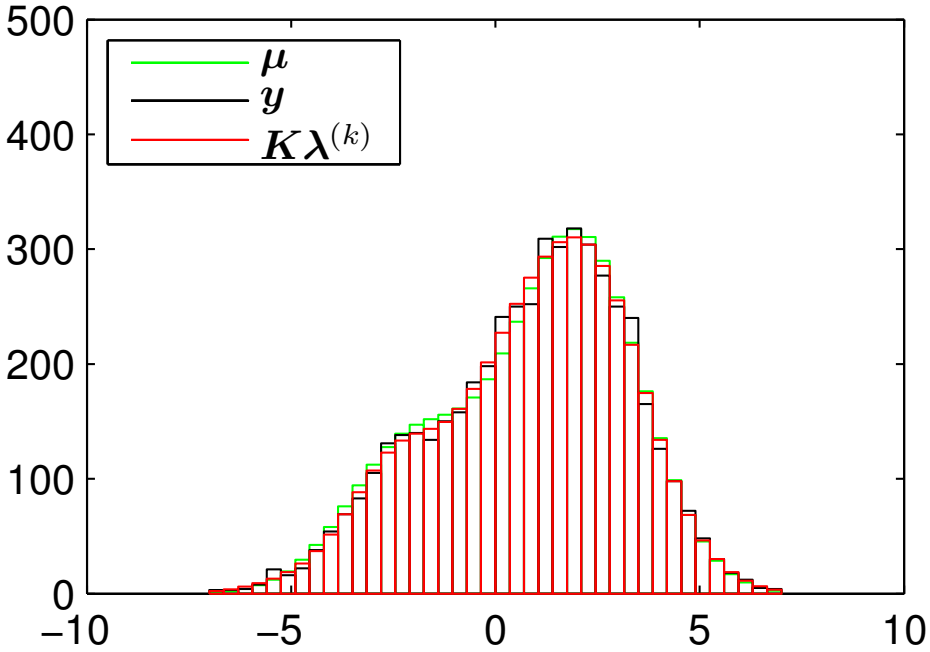


Figure: Smeared histogram

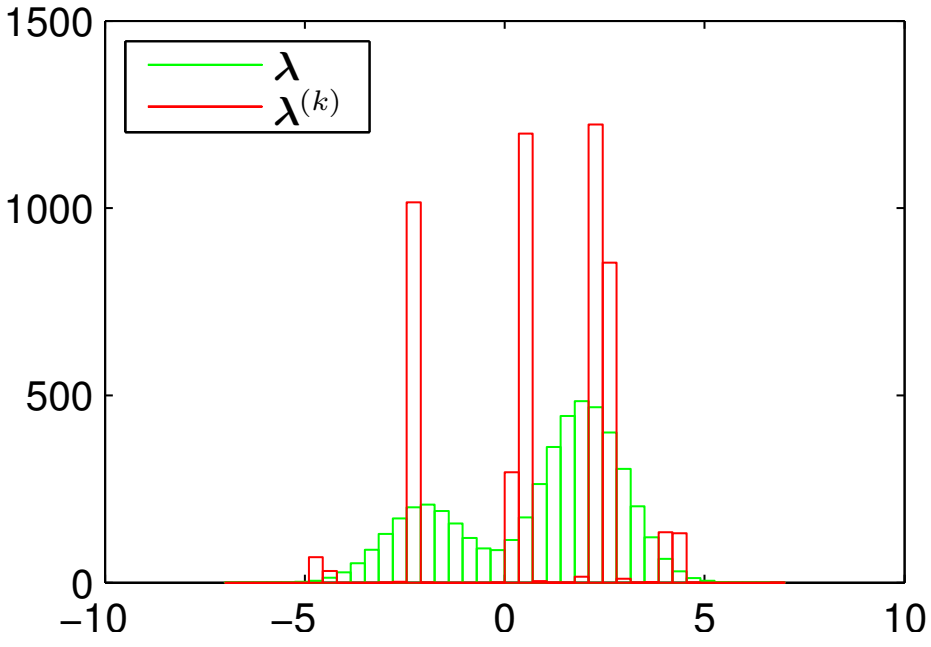
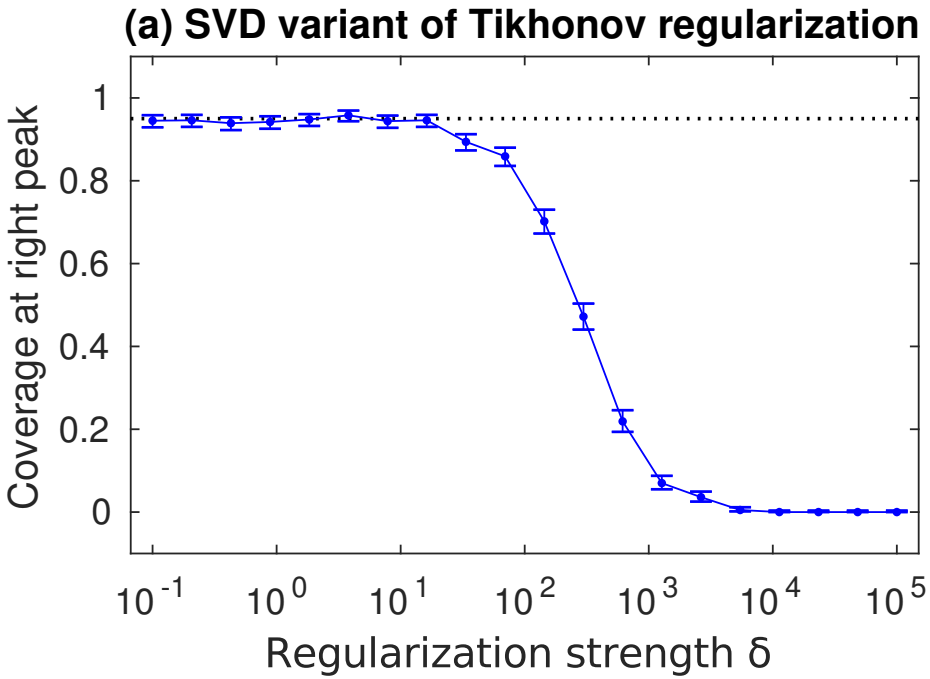


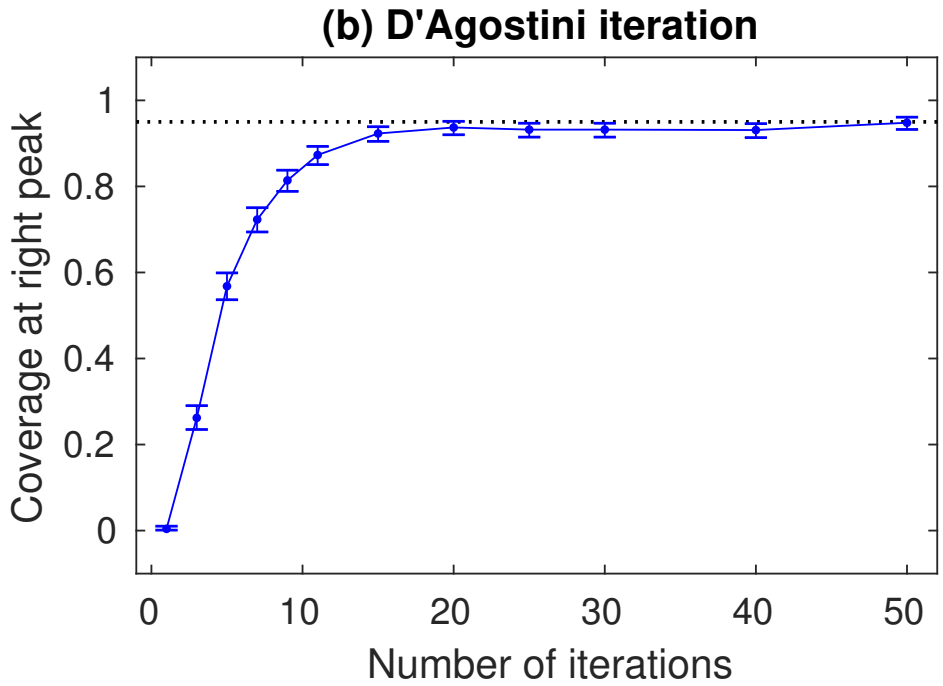
Figure: True histogram

Other methods

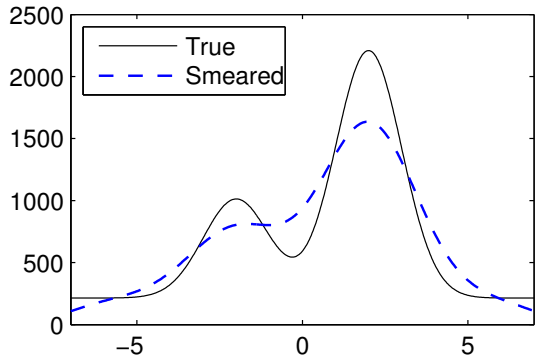
- Bin-by-bin correction factors
  - Attempts to unfold resolution effects by performing multiplicative efficiency corrections
  - This method is simply wrong and must not be used (it is no longer allowed in CMS)
- Fully Bayesian unfolding (Choudalakis, 2012)
  - Unfolding using Bayesian statistics where the prior regularizes the ill-posed problem
  - Certain priors lead to solutions similar to Tikhonov, but with Bayesian credible intervals as the uncertainties
  - Note: D'Agostini has nothing to do with proper Bayesian inference
- RUN/TRUEE (Blobel, 1985, 1996; Milke et al., 2013)
  - Penalized maximum likelihood with B-spline discretization
- Shape-constrained unfolding (Kuusela and Stark, 2017)
  - Correct-coverage uncertainties by imposing constraints on positivity, monotonicity and convexity
- Expectation-maximization with smoothing (Volobouev, 2015)
  - Adds a smoothing step to each iteration of D'Agostini
- Iterative dynamically stabilized unfolding (Malaescu, 2011)
  - Seems quite ad-hoc, with many free tuning parameters and unknown (at least to me) statistical properties
  - I have not seen this used in CMS, but it seems to be quite common in ATLAS
- ...

Coverage as a function of regularization strength (Kuusela, 2016)





Simulation setup:

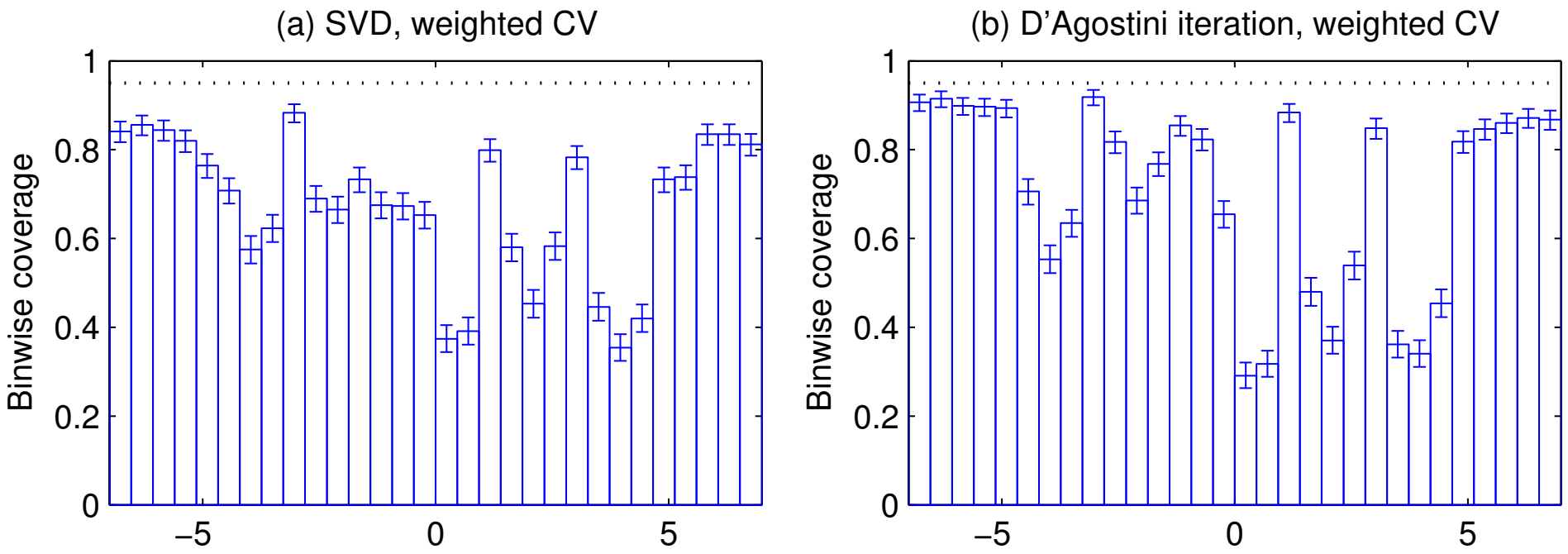


$$f(t) = \lambda_{\text{tot}} \left\{ \pi_1 \mathcal{N}(t|-2,1) + \pi_2 \mathcal{N}(t|2,1) + \pi_3 \frac{1}{|E|} \right\}$$

$$g(s) = \int_T \mathcal{N}(s-t|0,1) f(t)\, \mathrm{d}t$$

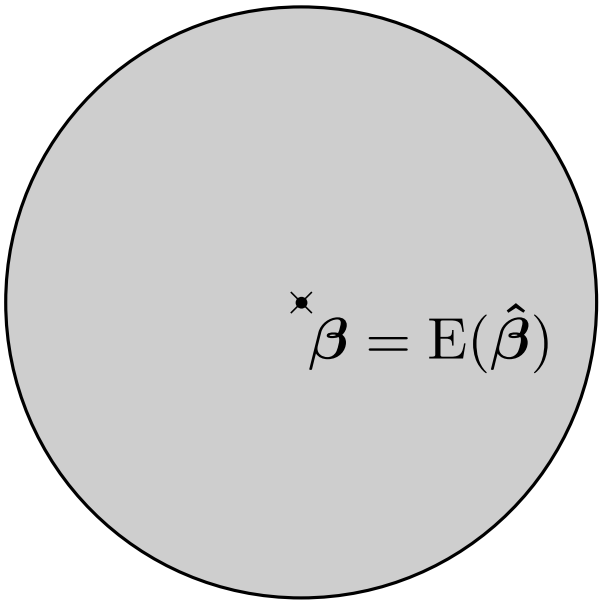
$$f^{\text{MC}}(t) = \lambda_{\text{tot}} \left\{ \pi_1 \mathcal{N}(t|-2,1.1^2) + \pi_2 \mathcal{N}(t|2,0.9^2) + \pi_3 \frac{1}{|E|} \right\}$$

Undercoverage of existing methods (Kuusela, 2016)



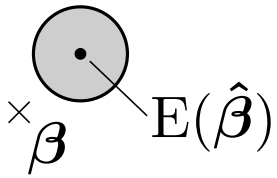
[The uncertainties tend to be especially badly estimated when $\boldsymbol{\lambda}^{MC}$ is close to the physical truth; some preliminary studies indicate that $\boldsymbol{\lambda}^{MC} = \text{const} \cdot \mathbf{1}$ performs better, but it is definitely not perfect either.]

Bias-variance trade-off and uncertainty quantification



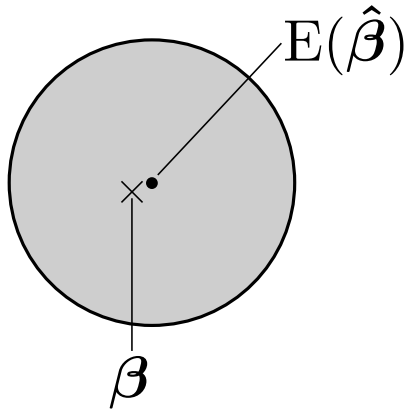
Unbiased,

coverage $= 1 - \alpha$



Optimal point estimation,

coverage $\ll 1 - \alpha$



Optimal UQ?,

coverage $= 1 - \alpha - \varepsilon$

Obtaining good coverage performance requires adjusting the bias-variance trade-off to the direction of less bias and more variance!

Coverage as a function of $\tau = \sqrt{\delta}$

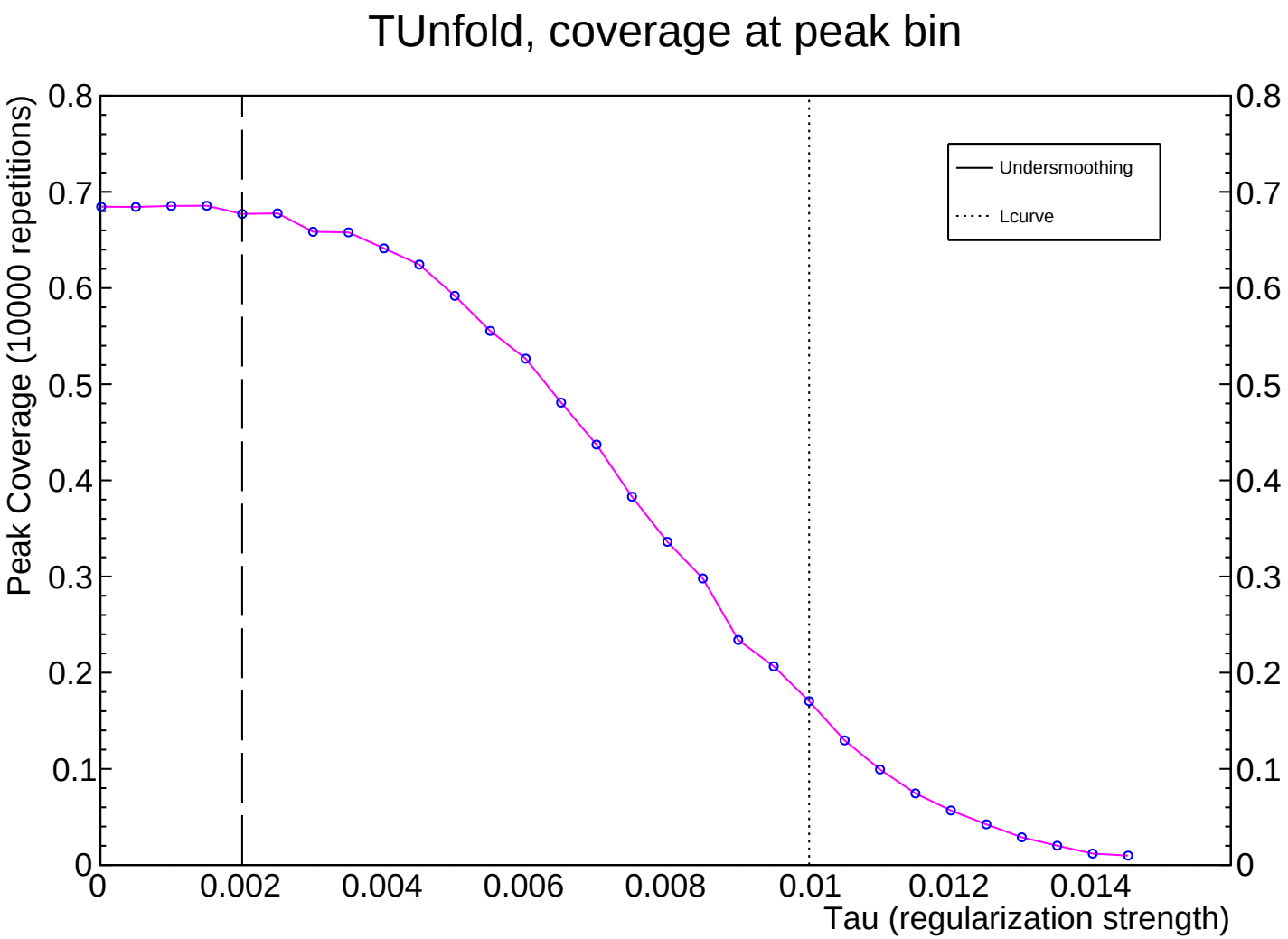
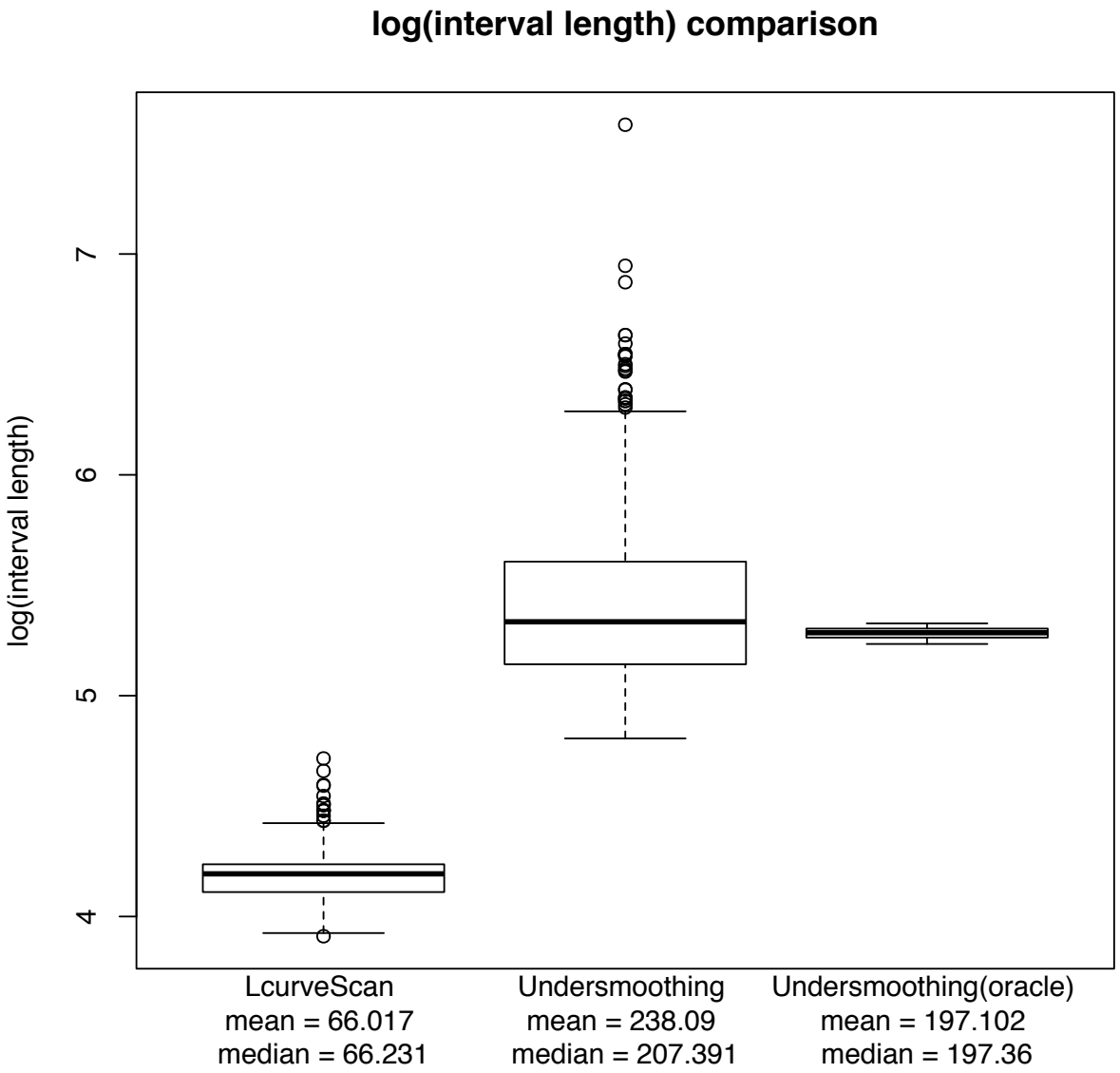
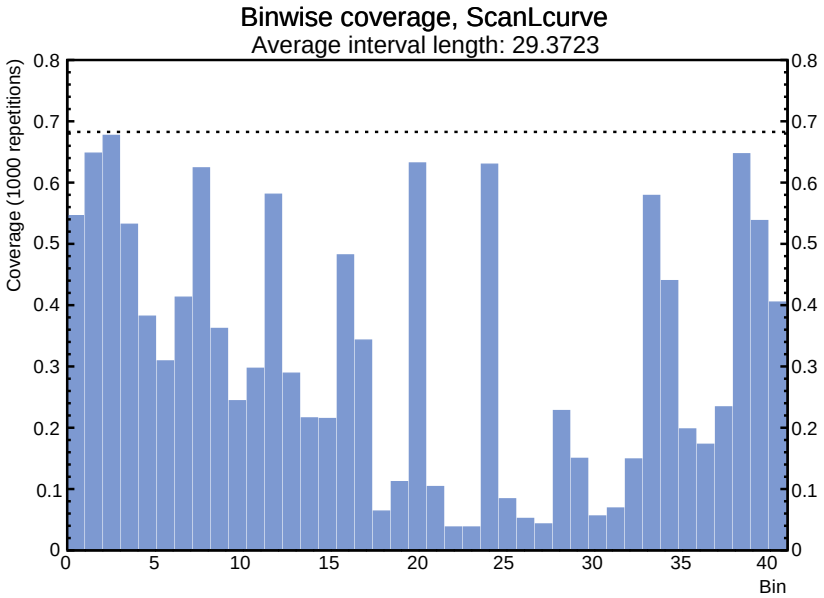


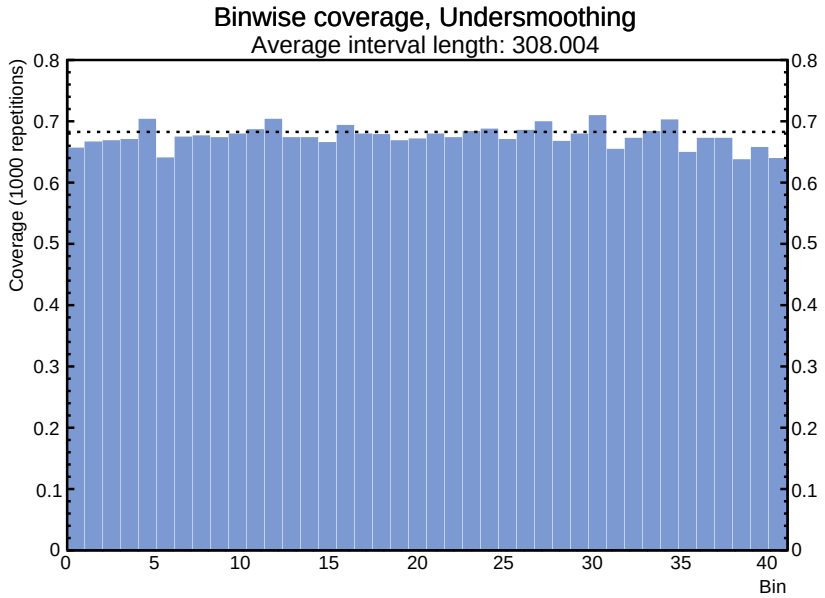
Figure: Coverage at the right peak of a bimodal density

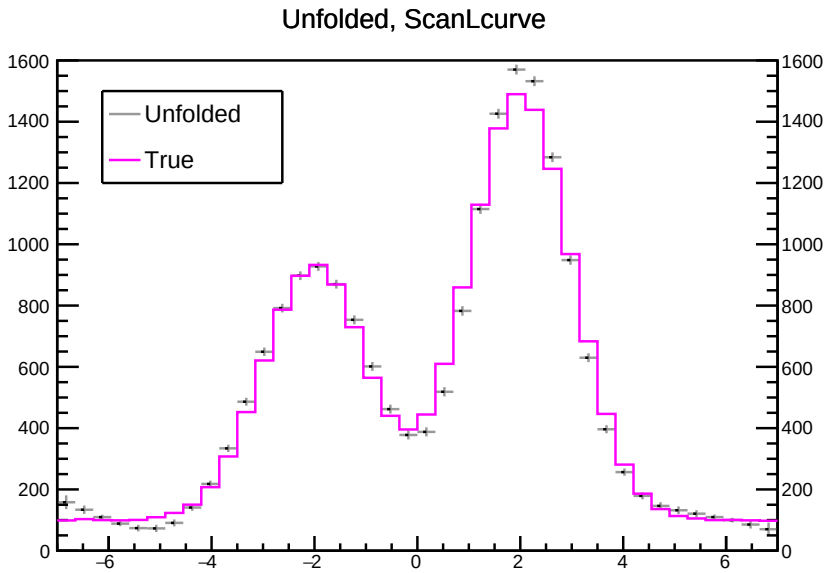
Interval lengths, $\boldsymbol{\lambda}^{\mathrm{MC}} = 0$

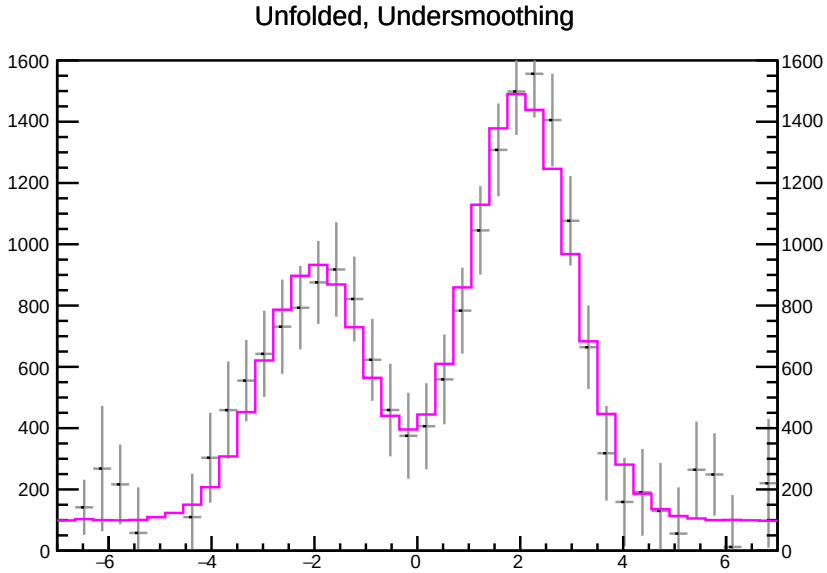


Histograms, coverage and interval lengths when $\boldsymbol{\lambda}^{\mathrm{MC}} \neq 0$







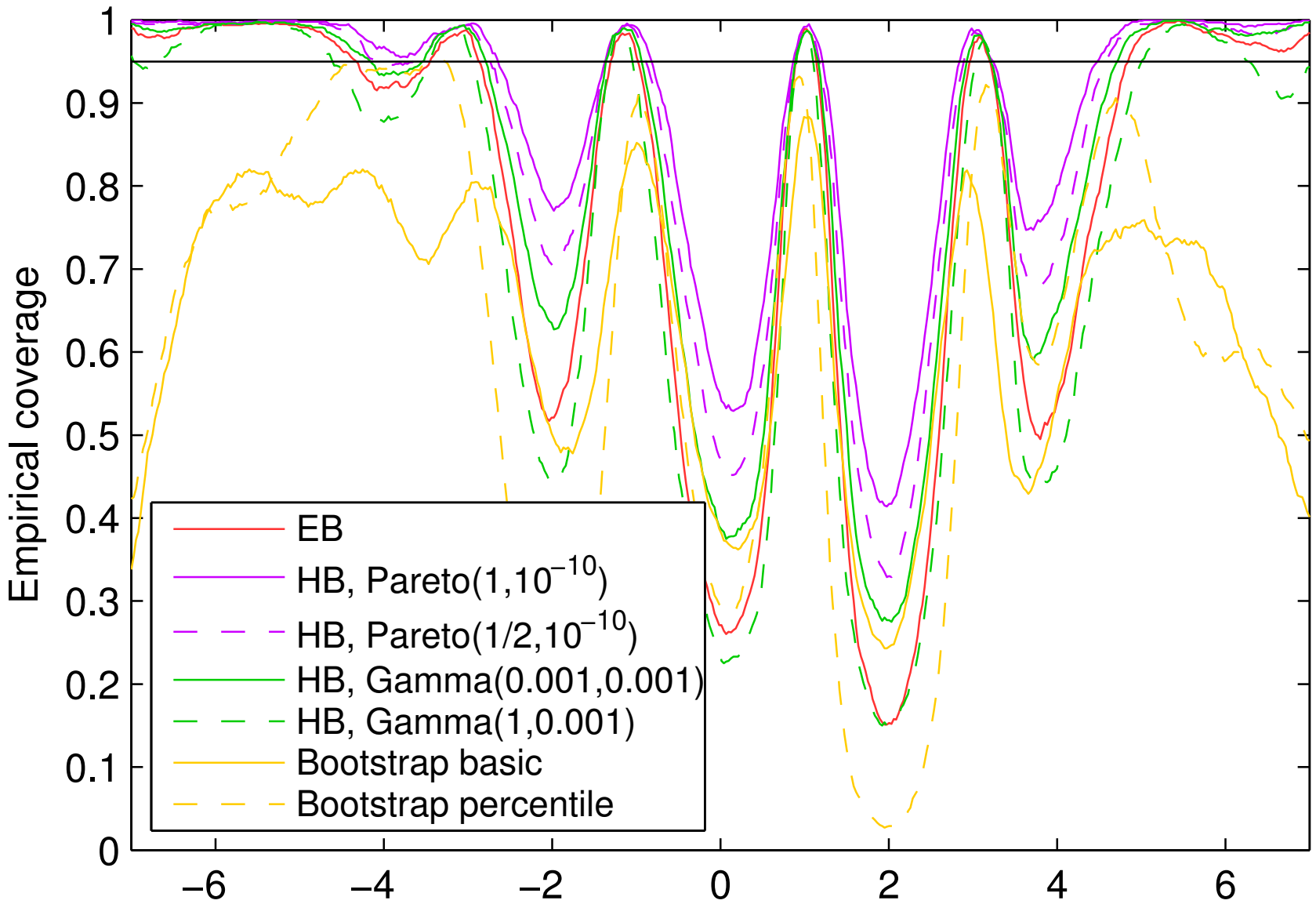


Coverage study from Kuusela (2016)

Method	Coverage at $t = 0$	Mean length
BC (data)	0.932 *(0.915, 0.947)*	0.079 *(0.077, 0.081)*
BC (oracle)	0.937 *(0.920, 0.951)*	0.064 *(0.064, 0.064)*
US (data)	0.933 *(0.916, 0.948)*	0.091 *(0.087, 0.095)*
US (oracle)	0.949 *(0.933, 0.962)*	0.070 *(0.070, 0.070)*
MMLE	0.478 *(0.447, 0.509)*	0.030 *(0.030, 0.030)*
MISE	0.359 *(0.329, 0.390)*	0.028
Unregularized	0.952 *(0.937, 0.964)*	40316

BC = iterative bias-correction
US = undersmoothing
MMLE = choose δ to maximize the marginal likelihood
MISE = choose δ to minimize the mean integrated squared error

UQ in inverse problems is challenging



[Kuusela and Panaretos (2015)]