Bayesian Inference in Analysing Results from Lattice QCD

ETMC Meeting Fall 2019

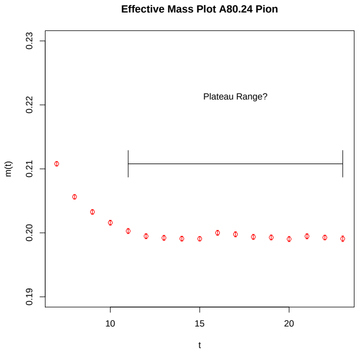
Matthias Fischer

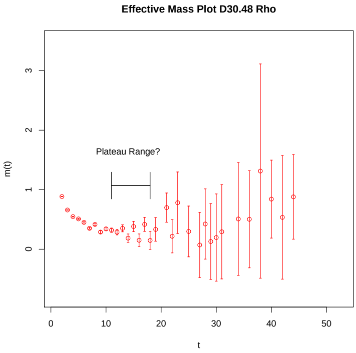
Bonn University

23. September 2019

Exemplary Effective Mass Data

Main aim in LQCD data analysis:
Find valid model range, estimate parameters and reliable error estimates.





$\Rightarrow$ **Solution**: joint posterior distribution, e.g.: $\pi(m, \vec{t}|\mathcal{D}, I)$

Bayes' Theorem and Bayesian Inference

Modified Bayes' Theorem

$$P(X|Y,I) = \frac{P(Y|X,I)P(X|I)}{P(Y|I)}, \quad I \text{ contains background information.}$$

Bayes' Theorem in Parameter Estimation

$$\pi(\theta_1,...,\theta_d|\mathcal{D},I) = \pi(\vec{\theta}|\mathcal{D},I) = \frac{\pi(\mathcal{D}|\vec{\theta},I)\pi(\vec{\theta}|I)}{\pi(\mathcal{D}|I)}$$

- $\pi(\vec{\theta}|\mathcal{D},I) =$ posterior
- $\pi(\mathcal{D}|\vec{\theta},I) =$ likelihood
- $\pi(\vec{\theta}|I) =$ prior
- $\pi(\mathcal{D}|I) =$ evidence, marginalized likelihood

Some Basic Probability Theory

Fully marginalized posterior distribution for θ_i:

$$\int \pi(\theta_1, ..., \theta_d|\mathcal{D}, I) \prod_{j \neq i} \mathrm{d}\theta_j = \pi(\theta_i|\mathcal{D}, I)$$

Aiming for fully marginalized effective mass distribution:

$$\pi(m|\mathcal{D}, I) \, .$$

$\Rightarrow$ Median provides effective mass parameter value and width error estimate.

Advantage: Error estimate incorporates statistical and systematical error.
Disadvantage: Fully marginalized distribution not analytically solvable.
↳ **Apply Monte Carlo Methods**.

Markov Chain Monte Carlo in Bayesian Inference

Draw samples from joint posterior distribution: $\pi(\vec{\theta}|\mathcal{D}, I)$.

- HMC Algorithm: Appropriate for high dimensional parameter space.
- Random Walk Metropolis Algorithm: Simplest sampling scheme.

Algorithm 1 Random Walk Metropolis-Hastings Algorithm

1: Input: $\vec{\theta}_n$
2: Draw: $\vec{\theta}^\star \sim \mathcal{N}(\vec{\theta}_n, \vec{\sigma})$
3:
$$\vec{\theta}_{n+1} = \begin{cases} \vec{\theta}^\star & \text{with probability:} \quad \min\left\{\frac{\pi(\mathcal{D}|\vec{\theta}^\star, I)\pi(\vec{\theta}^\star|I)}{\pi(\mathcal{D}|\vec{\theta}_n, I)\pi(\vec{\theta}_n|I)},\ 1\right\} \\ \vec{\theta}_n & \text{otherwise}\ . \end{cases}$$

$\Rightarrow$ Histogram of θ_i provides estimate of fully marginalized distribution.

Assumed Model

Signal function: $f(\vec{\theta}, t) = m_{eff}^{(1)}(t) = m \cdot (1 + A\mathrm{e}^{(-\delta E \cdot t)}),\ \vec{\theta} = (m, A, \delta E)$.

- Valid within possible plateau range $I_t = [t_0, t_1]$.
- Describes subset $\mathcal{D}_{[t_0,t_1]}$ of $\mathcal{D}$.
- Fit range parameter $\vec{t} = (t_0, t_1),\ t_0, t_1 \in \mathbb{N}$.

Need model $\mathcal{M}$ which describes $\mathcal{D}$ and not just subsets:
$\Rightarrow$ Introduce *artificial functions* $f_1(\vec{\alpha}, t)$ and $f_2(\vec{\alpha}, t)$ describing left out data.
$\Rightarrow$ Fit range parameter $\vec{t}$ becomes usual model function parameter.

$$\begin{aligned} M(\vec{\theta}, \vec{t}, \vec{\alpha}, t) = & f_1(\vec{\alpha}, t)\Theta((t_0 - 1) - t) \\ & + f(\vec{\theta}, t)\Theta(t - t_0)\Theta(t_1 - t) \\ & + f_2(\vec{\alpha}, t)\Theta(t - (t_1 + 1))\,. \end{aligned}$$

- $\Theta(\cdot)$ denotes *Heaviside-Theta*.

Bayes' Theorem for the Assumed Model

$$\pi(\vec{\theta}, \vec{t}, \vec{\alpha}|\mathcal{D}, \mathcal{M}, I') = \frac{\pi(\mathcal{D}|\vec{\theta}, \vec{t}, \vec{\alpha}, \mathcal{M}, I')\pi(\vec{\alpha}|\vec{\theta}, \vec{t}, \mathcal{M}, I')\pi(\vec{\theta}, \vec{t}|\mathcal{M}, I')}{\pi(\mathcal{D}|\mathcal{M}, I')}$$

Task: Find describing *likelihood* and *priors*.

- Latter equation contains *artificial parameter* $\vec{\alpha}$.
- Separated prior for $\vec{\alpha}$ because not accessible to us.
- Prior for $\vec{\theta}, \vec{t}$ can be assigned as usual.

Solution:

- Uniform priors for $\vec{\theta}$ and $\vec{t}$.
- Prior for $\vec{\alpha}$ will be implicitly incorporated into likelihood.

Likelihood for the Assumed Model

$$\pi(\mathcal{D}|\vec{\theta},\vec{t},\vec{\alpha},\mathcal{M},I') = C \cdot \underbrace{\exp(-\chi_1^2/2)}_{\text{from } f_1(\vec{\alpha},t)} \cdot \underbrace{\exp(-\chi^2/2)}_{\text{from } f(\vec{\theta},t)} \cdot \underbrace{\exp(-\chi_2^2/2)}_{\text{from } f_2(\vec{\alpha},t)}$$

- C: normalization constant.
- Different model ranges assumed to be uncorrelated.

$$\chi^2 = \sum_{i,j=t_0}^{t_1} (\mathcal{D}_i - f(\vec{\theta},t_i))\Sigma^{-1}_{ij,\mathcal{D}_{[t_0,t_1]}}(\mathcal{D}_j - f(\vec{\theta},t_j)),$$

$$\Sigma^{-1}_{ij,\mathcal{D}_{[t_0,t_1]}} \quad \text{inverse covariance matrix .}$$

- χ_1^2 and χ_2^2 drawn from corresponding χ^2-distribution.

Likelihood for the Artificial Functions

Definition χ^2-distribution, $x = \chi^2$ and ν being the degrees of freedom:

$$b(x,\nu) = \begin{cases} \frac{x^{\frac{\nu}{2}-1}\exp(-x/2)}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)}\,, & \text{for } x > 0\,, \\ 0\,, & \text{otherwise}\,. \end{cases}$$

Artificial functions are the true underlying model in the left out parts:
$\Rightarrow$ $\chi^2_{1,2}$ are χ^2 distributed.

$$\chi_1^2 \sim b(\chi_1^2, t_0 - 1) \quad \text{and} \quad \chi_2^2 \sim b(\chi_2^2, N_T - t_1 - 1)$$

What about the degree of freedom?:
$\Rightarrow$ Taken care of by prior $\pi(\vec{\alpha}|\vec{\theta}, \vec{t}, \mathcal{M}, I')$.

Newly Developed Algorithm

Algorithm 2 Fit Range and Parameter Sampling Algorithm

```
Start with an initial configuration $\{\vec{\theta}, \vec{t}\}$.
while Statistic not sufficient do
    for $\theta_i$ in $\vec{\theta}$ do
        Draw: $\chi_1^2 \sim b(\chi_1^2, t_0 - 1)$, $\chi_2^2 \sim b(\chi_2^2, N_T - t_1 - 1), \theta_i^\star \sim \mathcal{N}(\theta_i, \delta_i)$
        Compute $\pi(\vec{\theta}^\star, \vec{t}, \vec{\alpha}|\mathcal{D}, \mathcal{M}, I')$, do Metropolis update.
    end for
    Draw $\Delta t_i \sim \mathcal{N}(0, \delta_t)$, $\chi_1^2 \sim b(\chi_1^2, t_0^\star - 1)$, $\chi_2^2 \sim b(\chi_2^2, N_T - t_1^\star - 1)$
    Compute $\pi(\vec{\theta}, \vec{t}^\star, \vec{\alpha}|\mathcal{D}, \mathcal{M}, I')$, do Metropolis update.
    Store Metropolis updates.
end while
```

Fit range update: $\vec{t}^\star = \vec{t} + \vec{\Delta t}$.

Minimal plateau range: four data points.

Summary of the New Method

- Aiming for desired joint posterior distribution $\pi(m, \vec{t}|\mathcal{D}, \mathcal{M}, I')$.
- Adding artificial functions to signal function describes all data.
- Through artificial functions, $\vec{t}$ does not affect the data.
- $\chi^2_{1,2}$ serves as a penalty for shrinking the fit range.
- MC Simulation samples fit range and signal function parameters.
- Histogram of m yields desired fully marginalized distribution.

Pion Data in A80.24, Sample t_0 and t_1

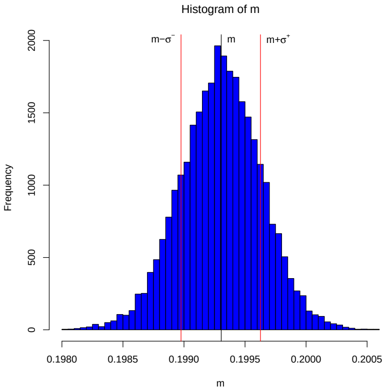


Figure: Effective mass distribution for sampled t_0 and t_1 for an A80.24 pion.

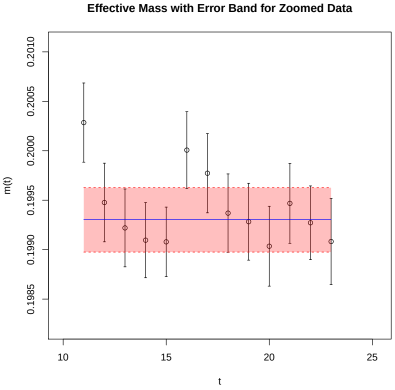


Figure: Error band received from the effective mass distribution.

Pion Data in A80.24, Sample t_0 and t_1

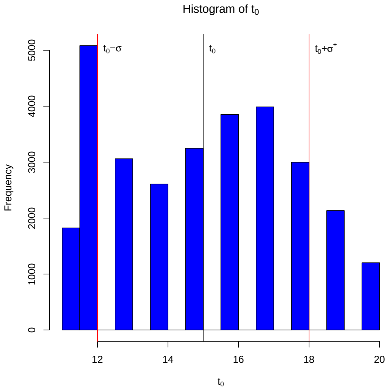


Figure: Distribution of t_0 for a pion in A80.24

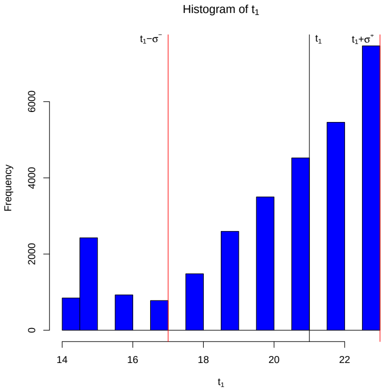


Figure: Distribution of t_1 for a pion in A80.24

Pion Data in A80.24, Sample t_0 and t_1

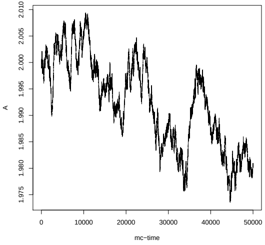


Figure: History of Markov chain of A for a pion in A80.24.

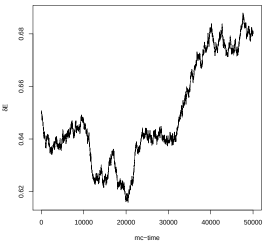


Figure: History of Markov chain of δE for a pion in A80.24.

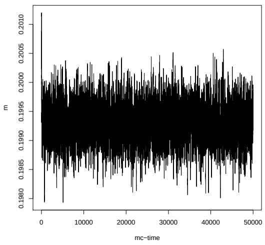


Figure: History of Markov chain of m for a pion in A80.24.

Pion Data in A80.24, Sample t_0 and t_1

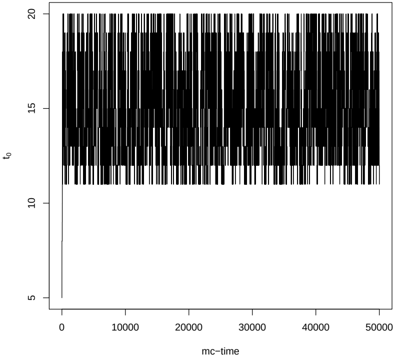


Figure: History of Markov chain of t_0 for a pion in A80.24.

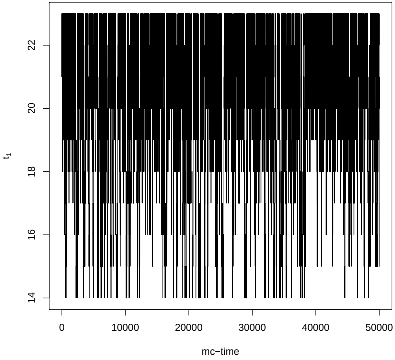


Figure: History of Markov chain of t_1 for a pion in A80.24.

Pion Data in A80.24, Different Methods

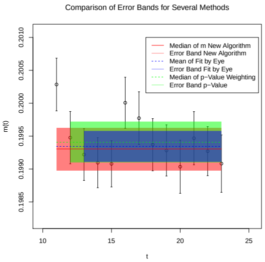


Figure: Effective mass estimates and error bands for t_0 and t_1 variation.

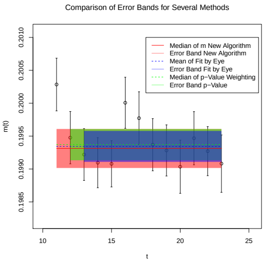


Figure: Effective mass estimates and error bands for t_0 variation and fixed t_1.

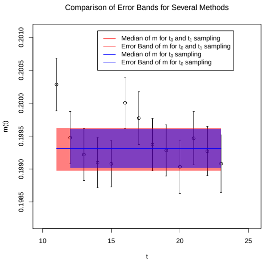


Figure: Comparison of t_0 and t_1 sampling to sample t_0 and fix t_1 to the end for the new method.

Rho Data in B55.32, Sample t_0 and t_1

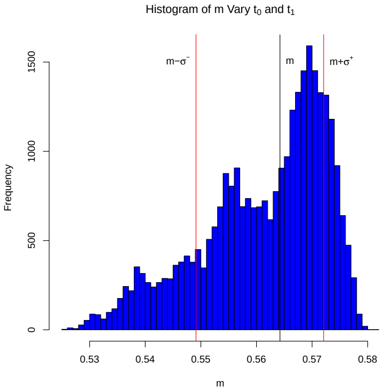


Figure: Distribution of m for one ρ-pc in B55.32.

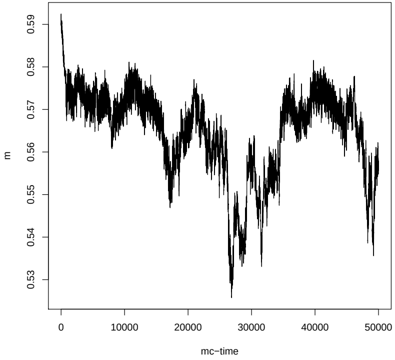


Figure: MCMC history of m for one ρ-pc in B55.32.

Rho Data in B55.32, Sample t_0 and t_1

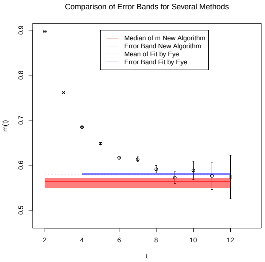


Figure: Error bands from the different analysis methods for one ρ-pc in B55.32.

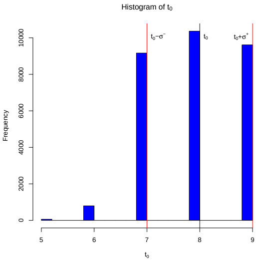


Figure: Distribution of t_0 for one ρ-pc in B55.32.

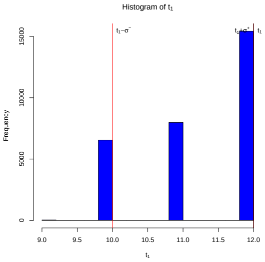


Figure: Distribution of t_1 for one ρ-pc in B55.32.

Rho Data Spectrum Comparison

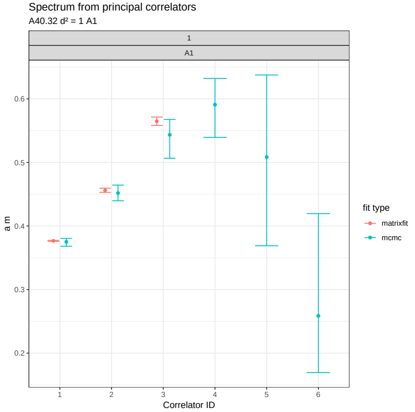


Figure: Spectrum for a ρ-GEVP in A40.32, $P^2 = 1$, projected to irrep $A1$ and a StN-ratio noise cut at 50 %.

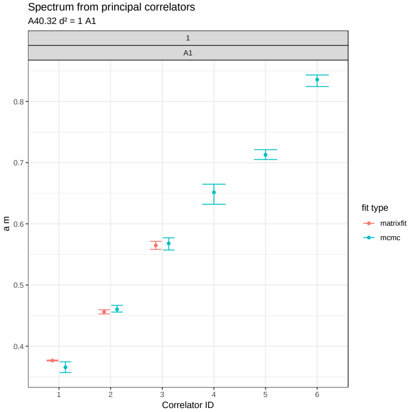


Figure: Spectrum for a ρ-GEVP in A40.32, $P^2 = 1$, projected to irrep $A1$ and a StN-ratio noise cut at 20 %.

Summary, Conclusion and Outlook

- Developed algorithm samples signal function and fit range parameters.
- Error estimates larger than for one fit range.
- Effective mass distribution not necessarily normally distributed.
- Results dependent on noise cut.
- Model selection to compare 0^{th} to 1^{st} order effective mass expansion.
- Extension to correlator fits.
- Improve sampling scheme.

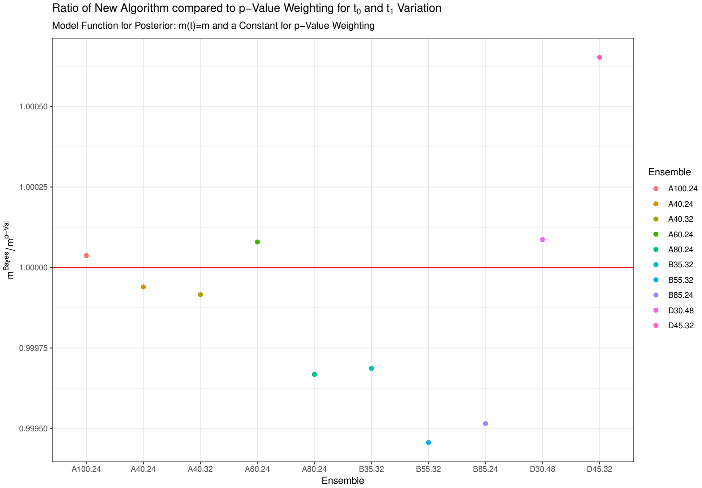
Thank you for your attention!
Do you have questions?

Literature

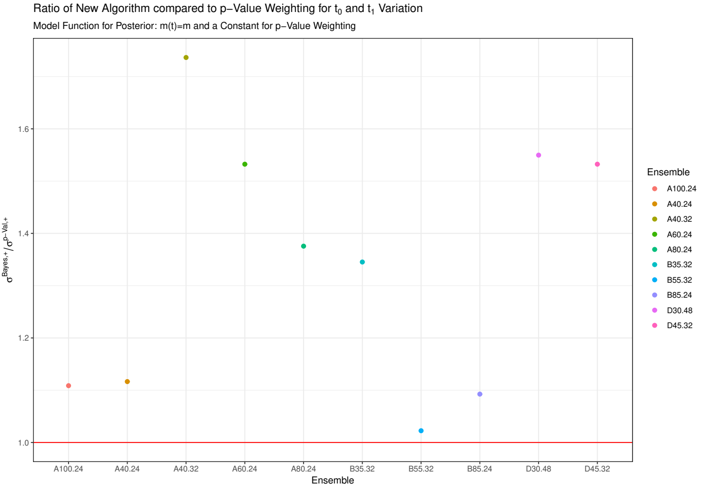
- Sivia, Data Analysis - A Bayesian Tutorial, 2. edition, Oxford University Press
- Carlin & Louis, Bayesian Methods for Data Analysis, 3. edition, CRC Press

Appendix

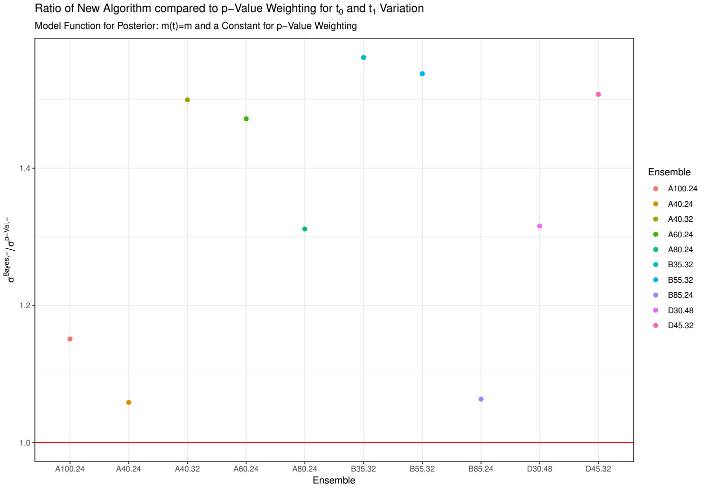
Ratio Plots Median Pion, Compare p-value



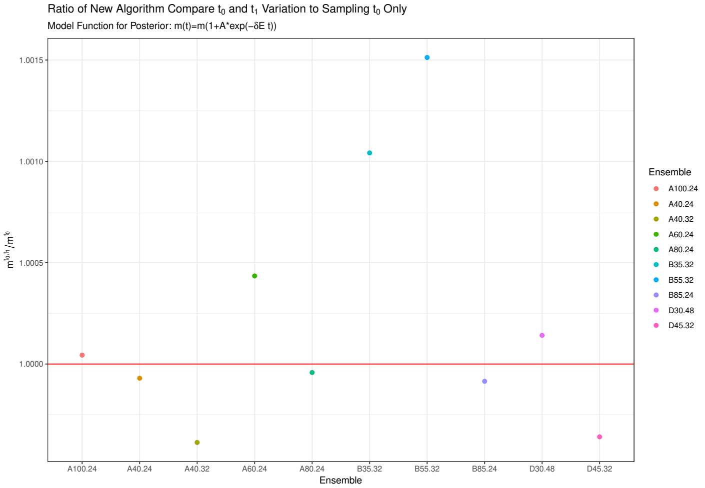
Ratio Plots Upper Error Pion, Compare *p*-value



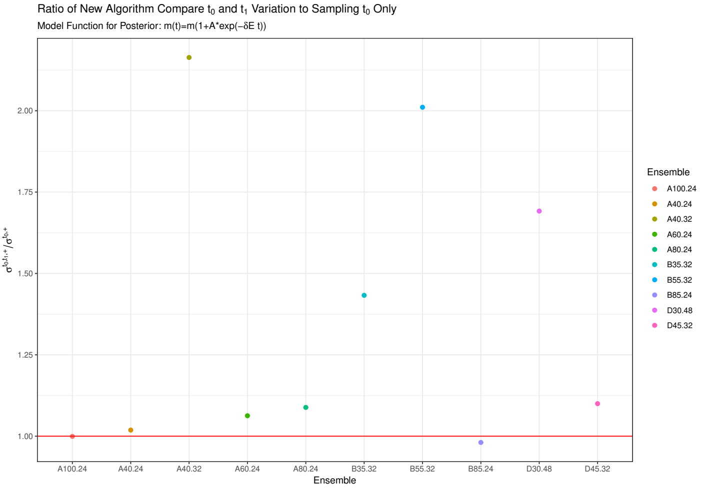
Ratio Plots Lower Error Pion, Compare p-value



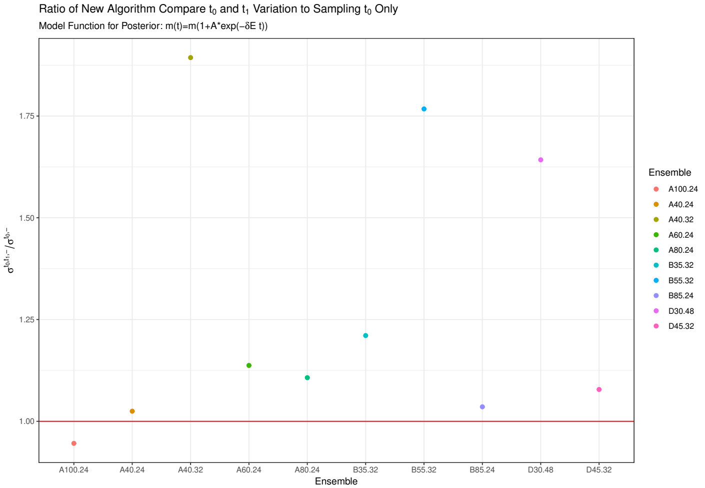
Ratio Plots Median Pion, $\vec{t}$ to t_0 Sampling



Ratio Plots Upper Error Pion, $\vec{t}$ to t_0 Sampling



Ratio Plots Lower Error Pion, $\vec{t}$ to t_0 Sampling



Pion Data in A80.24, Sample t_0

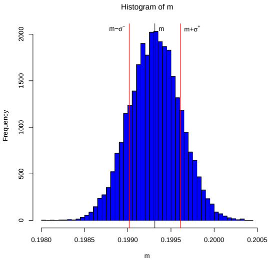


Figure: Effective mass distribution for sampling t_0 and fix t_1 for an A80.24 pion.

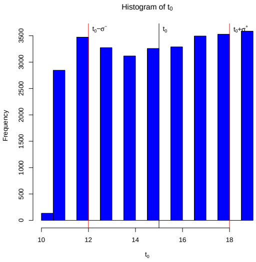


Figure: Distribution of t_0 for a pion in A80.24, sampling t_0, fix t_1

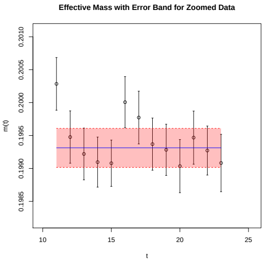


Figure: Error band received from the effective mass distribution.

Pion Data in A80.24, Sample t_0 and t_1

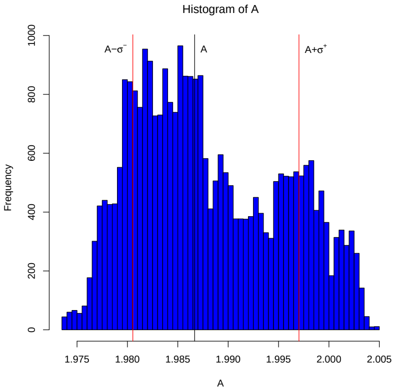


Figure: Distribution of A for a pion in A80.24

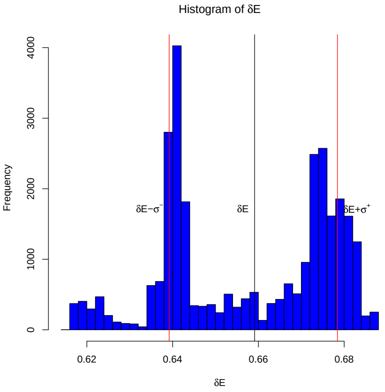


Figure: Distribution of δE for a pion in A80.24

Rho Data in A60.24, Sample t_0 and t_1

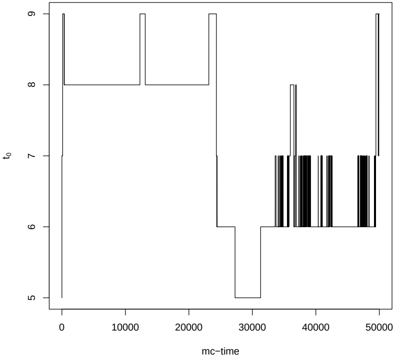


Figure: MCMC history of t_0 for one ρ-pc in A60.24.

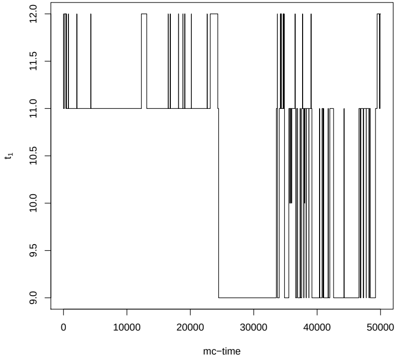


Figure: MCMC history of t_1 for one ρ-pc in A60.24.

Landscape Plots `A80.24`

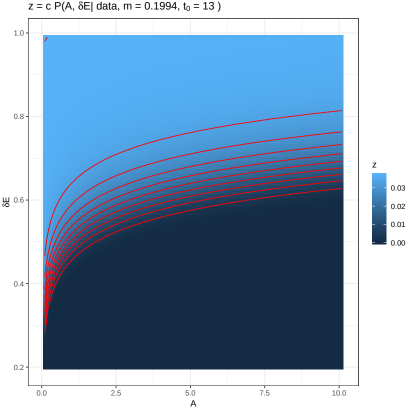


Figure: Posterior potential for A and δE with t_0 in the plateau.

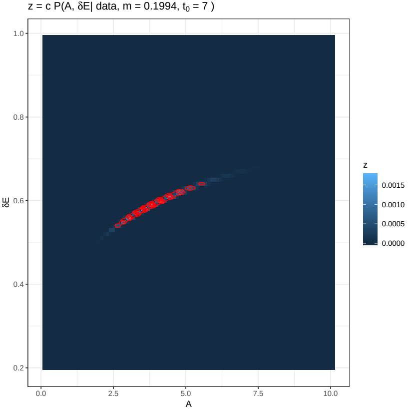


Figure: Posterior potential for A and δE with t_0 before the plateau.

Landscape Plots `A80.24`

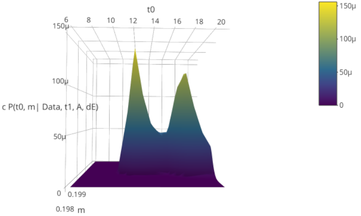


Figure: Posterior potential for t_0 and m. Projection of t_0.

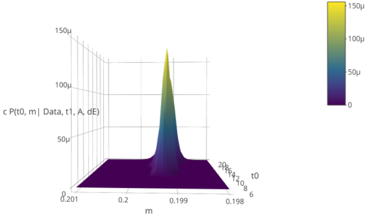


Figure: Posterior potential for t_0 and m. Projection of m.

Change-Point Analysis A80.24

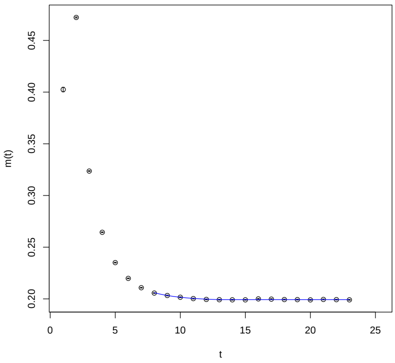


Figure: Fit result with medians of distributions for a pion

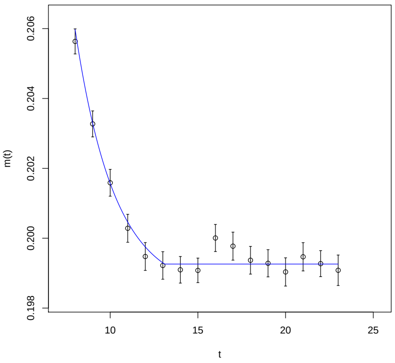


Figure: Zoomed fit result with medians of distributions for a pion

Change-Point Analysis A80.24

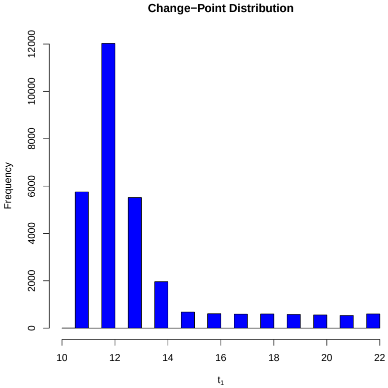


Figure: Change-point distribution with $t_0 = 8$ and $t_2 = 23$ for A80.24.

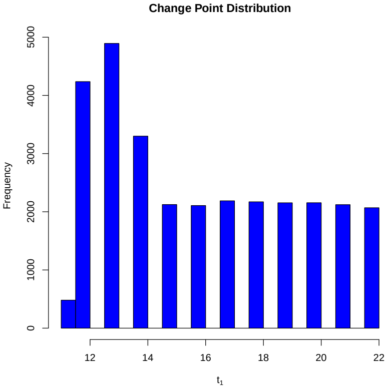


Figure: Change-point distribution with $t_0 = 5$ and $t_2 = 23$ for A80.24.