Scale setting the möbius domain wall fermion on HISQ action with M_Ω and the gradient-flow scales

Phys. Rev. D 103, 054511 (2021)

Nolan Miller L. Carpenter E. Berkowitz C.C. Chang B. Hörz
D. Howarth H. Monge-Camacho E. Rinaldi D.A. Brantley C. Körber
C. Bouchard M.A. Clark A.S. Gambhir C.J. Monahan A. Nicholson
P. Vranas A. Walker-Loud

March 5, 2025

California Lattice Collaboration

Overview

Strategy

Results

Updates

Bayesian model averaging

Backup

- Mixed action: möbius domain wall fermions (valence) on highly-improved staggered quarks ($N_f = 2 + 1 + 1$)
- Work in isosymmetric limit
- m_Ω for physical scale
- SU(2) χPT & Taylor expansions (with model averaging)
- In addition to w_0, $\sqrt{t_0}$ calculate improved scales with leading-order discretization effects removed

$$\left.\frac{t^2\langle E(t)\rangle}{1+\sum_n C_{2n}\frac{a^{2n}}{t^n}}\right|_{t=t_{0,\mathrm{imp}}} = 0.3\,, \qquad t\frac{d}{dt}\left(\frac{t^2\langle E(t)\rangle}{1+\sum_n C_{2n}\frac{a^{2n}}{t^n}}\right)\Bigg|_{t=w^2_{0,\mathrm{imp}}} = 0.3$$

Why MDWF on HISQ?

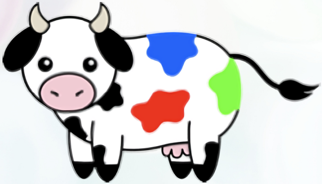
Strategy

- MDWF circumvents Nielsen-Ninomiya

Why MDWF on HISQ?

Strategy

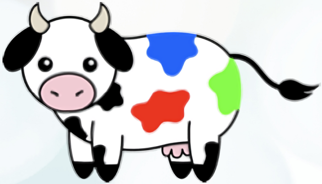
- MDWF circumvents Nielsen-Ninomiya
- Wanted to bootstrap our action with publicly availble data $\implies$ use MILC's HISQ action



Why MDWF on HISQ?

Strategy

- MDWF circumvents Nielsen-Ninomiya
- Wanted to bootstrap our action with publicly availble data $\implies$ use MILC's HISQ action
- QUDA provides excellent MDWF solvers: compute time optimized on GPUs such that 95+% of wall clock cycles are spent on MDWF solves, not contractions/smearing

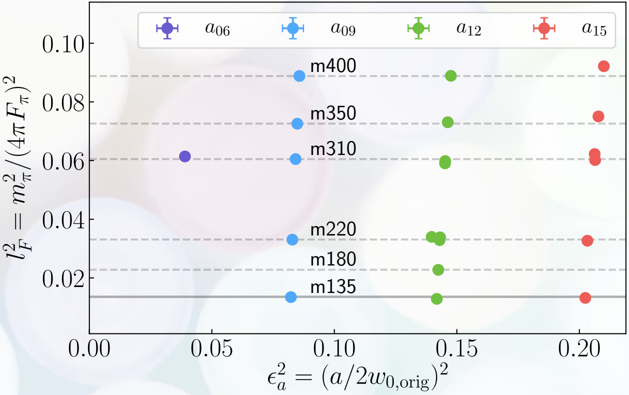


Why MDWF on HISQ?

Strategy

- MDWF circumvents Nielsen-Ninomiya
- Wanted to bootstrap our action with publicly availble data $\implies$ use MILC's HISQ action
- QUDA provides excellent MDWF solvers: compute time optimized on GPUs such that 95+% of wall clock cycles are spent on MDWF solves, not contractions/smearing
- Emprically, MDWF requires only around $\sim 1\%$ statistics compared to clover fermions for similar precision $\implies$ offsets cost from simulating 5th dimension, lack of multi-grid solver for MDWF

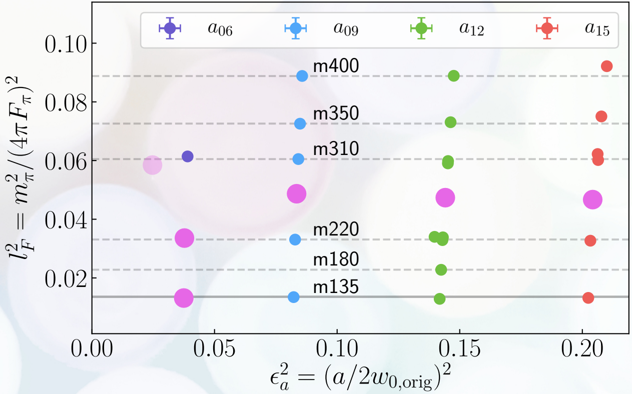
- 22 ensembles spanning $130 \text{ MeV} \lesssim m_\pi \lesssim 400 \text{ MeV}$, $0.06 \text{ fm} \lesssim a \lesssim 0.15 \text{ fm}$
- Strange dependence: `a12m220ms` tuned at $\sim 60\%$ physical strange quark mass
- Volume dependence: `a15m310/L` (2.4 fm vs 3.6 fm), `a12m310/XL` (2.9 fm vs 5.8 fm), `a12m310XL/a12m180L/a12m130` (same volume, different m_π)



Lever arms (2025)

Strategy

- 5^{+1}_{-0} new ensembles
- 1 at physical pion mass
- 2^{+1}_{-0} at/below the current finest lattice spacing



Why m_Ω?

Strategy

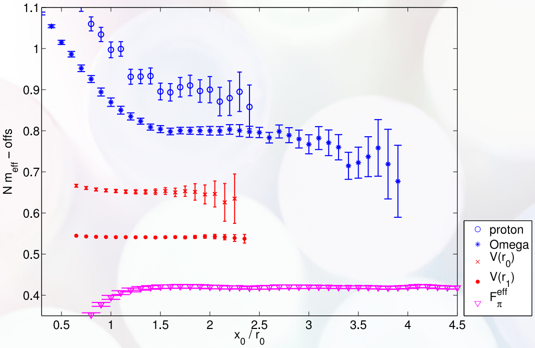
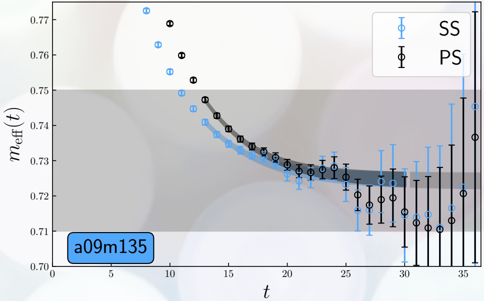


Figure: FLAG (2024)

- Stable in QCD+QED
- No valence light quarks $\implies$ milder chiral dependence for extrapolations
- Favorable Parisi-Lepage scaling compared to other baryons: SNR $\sim e^{-(m_\Omega - 3m_\pi/2)t}$
- (pre-Edinburgh) Experimental value for m_Ω more precise than decay constants
- (pre-Edinburgh) Experimental value for F_π containinated by QED, potential BSM effects

Fitting the Ω correlators

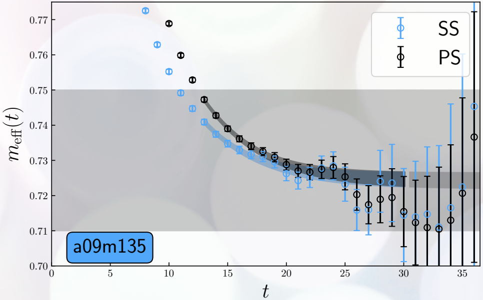
Strategy

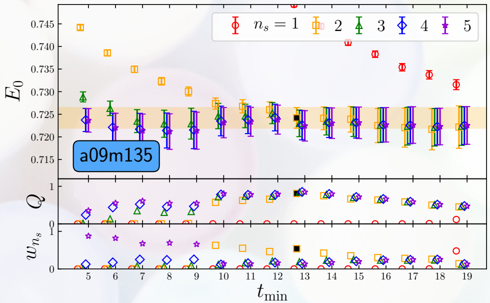


- Two different combinations of source/sinks, fit simultaneously

Fitting the Ω correlators

Strategy

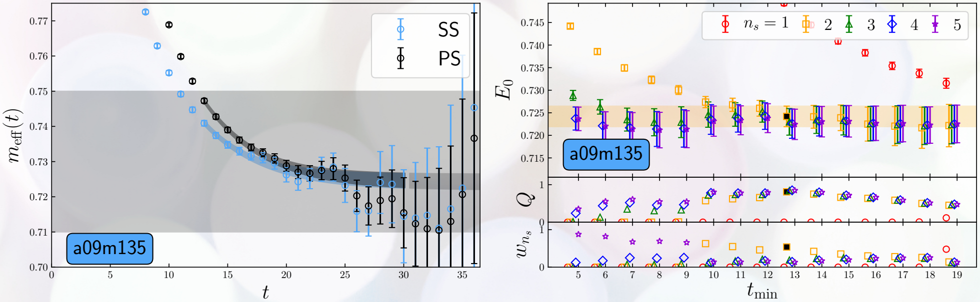




- Two different combinations of source/sinks, fit simultaneously
- Check for stability using multi-state fits up to $N = 5$

Fitting the Ω correlators

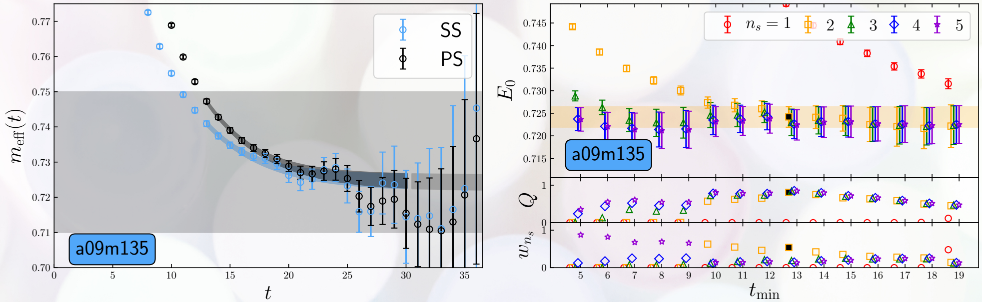
Strategy



- Two different combinations of source/sinks, fit simultaneously
- Check for stability using multi-state fits up to $N = 5$
- Stabilize fits with wide priors, e.g. $\delta E \sim \log \mathcal{N}(2m_\pi, m_\pi^2)$

Fitting the Ω correlators

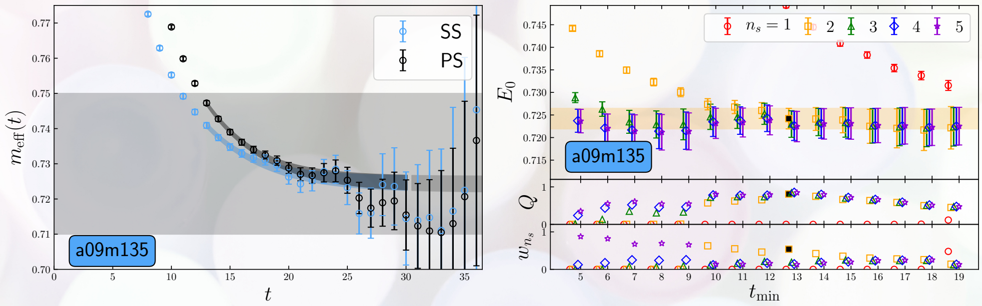
Strategy



- Two different combinations of source/sinks, fit simultaneously
- Check for stability using multi-state fits up to $N = 5$
- Stabilize fits with wide priors, e.g. $\delta E \sim \log \mathcal{N}(2m_\pi, m_\pi^2)$
- Use Bayes factors & relative stability for model selection

Fitting the Ω correlators

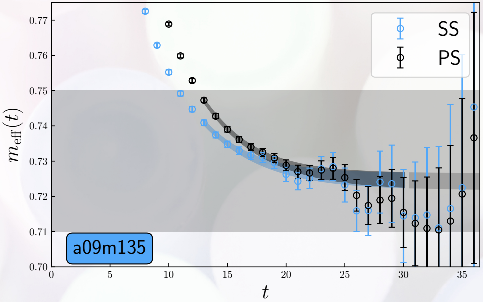
Strategy

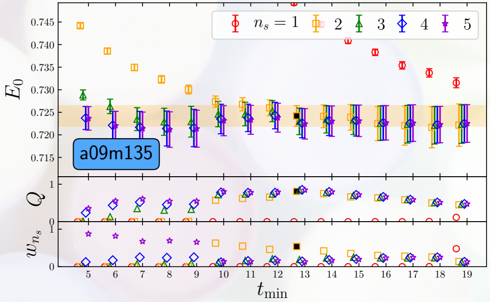


- Two different combinations of source/sinks, fit simultaneously
- Check for stability using multi-state fits up to $N = 5$
- Stabilize fits with wide priors, e.g. $\delta E \sim \log \mathcal{N}(2m_\pi, m_\pi^2)$
- Use Bayes factors & relative stability for model selection
- Try to optimize $t_{\rm min}^{\rm phys}$ over all ensembles simultaneously

Fitting the Ω correlators

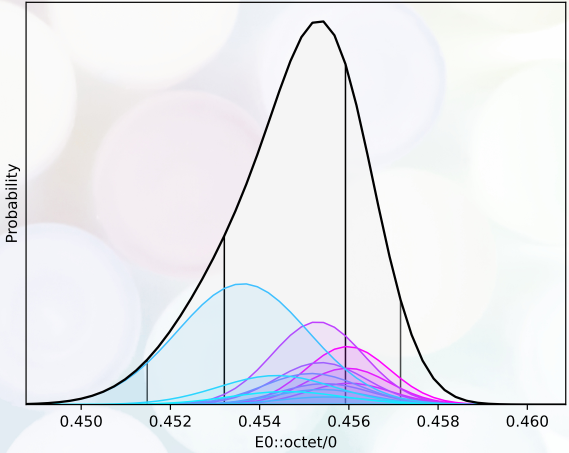
Strategy





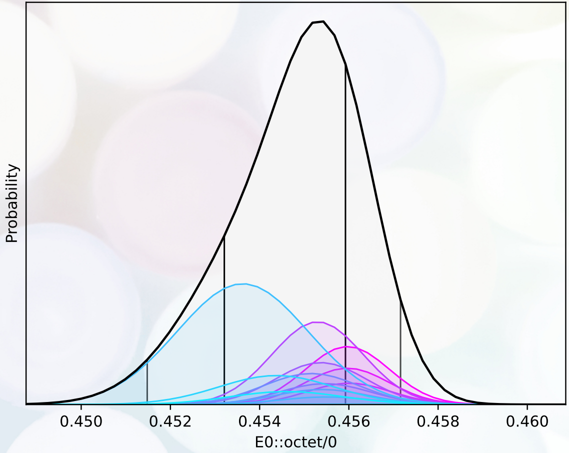
- Two different combinations of source/sinks, fit simultaneously
- Check for stability using multi-state fits up to $N = 5$
- Stabilize fits with wide priors, e.g. $\delta E \sim \log \mathcal{N}(2m_\pi, m_\pi^2)$
- Use Bayes factors & relative stability for model selection
- Try to optimize $t_{\rm min}^{\rm phys}$ over all ensembles simultaneously
- Cross-checks with 3+ people

- Model selection is not a replacement for model average!



Example from CLS ensemble B450

- Model selection is not a replacement for model average!
- Law of large numbers does not necessarily apply to systematic errors $\implies$ posterior will often be non-Gaussian

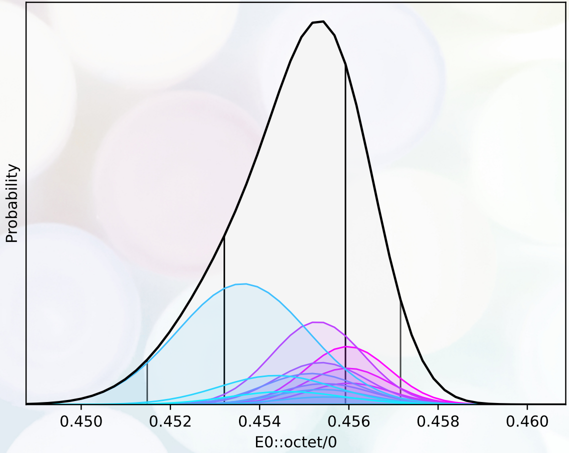


Example from CLS ensemble B450

Some systematic errors

Strategy

- Model selection is not a replacement for model average!
- Law of large numbers does not necessarily apply to systematic errors $\implies$ posterior will often be non-Gaussian
- Not sufficient to simply quote statistical/systematic errors – we must chain correlations forward

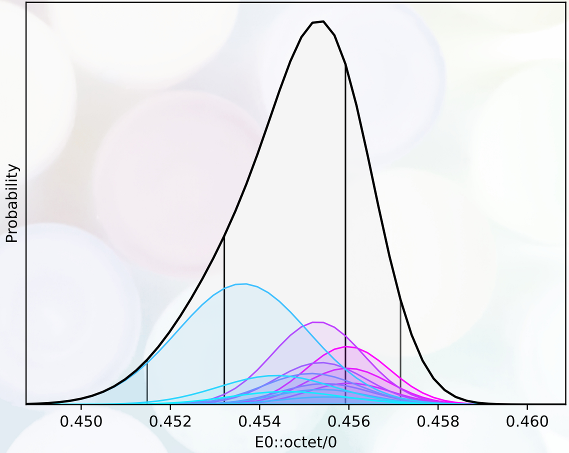


Example from CLS ensemble B450

Some systematic errors

Strategy

- Model selection is not a replacement for model average!
- Law of large numbers does not necessarily apply to systematic errors $\implies$ posterior will often be non-Gaussian
- Not sufficient to simply quote statistical/systematic errors – we must chain correlations forward
- Generalized linear least squares is unbiased, but *nonlinear* least squares is not



Example from CLS ensemble B450

Backbone for the analysis: choosing the appropriate effective field theory

Strategy

- Mixed-action EFT not advantageous per previous studies (e.g., F_K/F_π)
- SU(3) heavy baryon χPT converges too slowly $\implies$ use SU(2) heavy baryon χPT with only pions as degrees of freedom
- Fix ensembles near physical value of strange quark masses so that Taylor corrections are sufficient

Parameterizing the chiral dependence

Strategy

Generically write $O = O^{\chi} + O^{m_{qs} \neq m_{qs}^*} + O^{\text{disc}}$

- SU(3) heavy baryon χPT converges slowly $\implies$ use SU(2) heavy baryon χPT w/ Taylor corrections for strange quark mass mistuning
- Ω baryon has no terms non-analytic in m_π^2 until $\mathcal{O}(m_\pi^4)$, no odd powers of m_π^2 until $\mathcal{O}(m_\pi^7)$ $\implies$ no odd powers of m_π for $m_\Omega w_0$ to working order

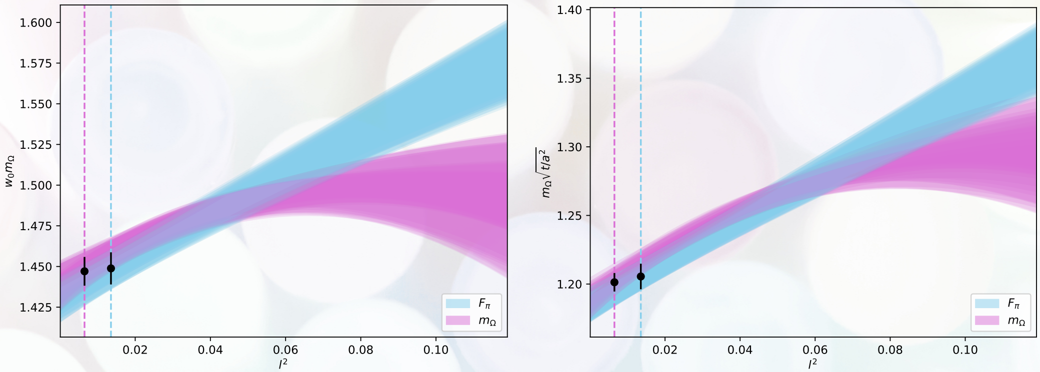
$$m_\Omega^\chi = m_0 \left\{ 1 + \alpha_2 \frac{m_\pi^2}{\Lambda_\chi^2} + \frac{m_\pi^4}{\Lambda_\chi^4}\left[\alpha_4 \lambda_\pi + \beta_4\right] + \frac{m_\pi^6}{\Lambda_\chi^6}\left[\alpha_6 \lambda_\pi^2 + \beta_6 \lambda_\pi + \gamma_6\right] + \dots \right\}$$

$$w_0^\chi = w_{0,\text{ch}} \left\{ 1 + k_1 \frac{m_\pi^2}{\Lambda_\chi^2} + \frac{m_\pi^4}{\Lambda_\chi^4}\left[k_2 \lambda_\pi + k_3\right] + \dots \right\}$$

where $\lambda_\pi = \log(m_\pi^2/\mu^2)$ (and we take $\mu = \Lambda = \{4\pi F_\pi, m_\Omega\}$)

Visualizing the chiral dependence (improved scales)

Strategy



- Despite having very different trajectories, both results agree at physical point

“Fixing” the renormalization scale

Strategy

- Chiral logs depend on the renormalization scale: $\lambda_l = \log(m_\pi^2/\mu^2)$
- One natural choice is to use the pion decay constant in the SU(2) limit, $\mu_0 = 4\pi F_0$
- Consider instead $\mu^{\text{ens}} = 4\pi F_\pi^{\text{ens}}$, with F_π taken on a given lattice ensemble.

$$F_\pi = 1 + \overbrace{\alpha\frac{m_\pi^2}{\Lambda^2} + \beta\frac{m_K^2}{\Lambda^2} - \frac{m_\pi^2}{\Lambda^2}\lambda_l - \frac{1}{2}\frac{m_K^2}{\Lambda^2}\lambda_K}^{\text{NLO}} + \dots$$

$$\implies \log\left(\frac{m_\pi^2}{\mu_0^2}\frac{\mu_0^2}{(\mu^{\text{ens}})^2}\right) = \log\left(\frac{m_\pi^2}{\mu_0^2}\right) - 2\delta_{F_\pi}^{\text{NLO}} + \mathcal{O}(m^4, m^4\lambda, m^4\lambda^2, a^2, a^2\alpha_S)$$

- To working order, we can aborb these terms into the N^3LO terms
- But at higher order, these corrections potentially become important

Generically write $O = O^{\chi} + O^{m_{qs} \neq m_{qs}^*} + O^{\text{disc}}$

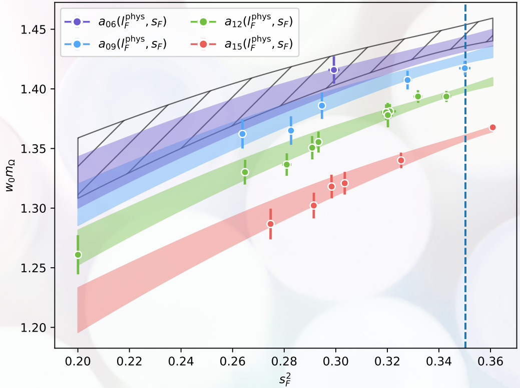
- All but what ensemble tuned near the strange quark mass $\implies$ a Taylor expansion for strange dependence should suffice

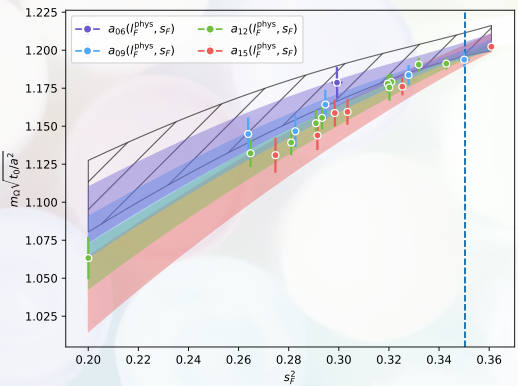
$$m_{\Omega}^{m_{qs} \neq m_{qs}^*} = c'_s s_{\Lambda}^2 + c'_{ss} s_{\Lambda}^4 + c'_{sss} s_{\Lambda}^6 + \dots$$
$$w_0^{m_{qs} \neq m_{qs}^*} = c''_s s_{\Lambda}^2 + c''_{ss} s_{\Lambda}^4 + c''_{sss} s_{\Lambda}^6 + \dots$$

where $s_{\Lambda}^2 = (2m_{\pi}^2 - m_K^2)/\Lambda_{\chi}^2$

Strange quark mass mistuning

Strategy





Representative model: $\mathrm{N}^3\mathrm{LO}$ w/ chiral logs, $\Lambda = 4\pi F_\pi$

Parameterizing the discretation effects (w/o FV corrections)

Strategy

Generically write $O = O^{\chi} + O^{m_{q_s} \neq m^*_{q_s}} + O^{\rm disc}$

- Similar Taylor expansion for a-dependence, but now must also include radiative corrections from the Symanzik effective theory

$$m_{\Omega}^{\rm disc} = \tilde{d}'_a \epsilon_a^2 \alpha_S + d'_a \epsilon_a^2 + d'_{aa} \epsilon_a^4 + d'_{aaa} \epsilon_a^6 \dots$$
$$w_0^{\rm disc} = \tilde{d}''_a \epsilon_a^2 \alpha_S + d''_a \epsilon_a^2 + d''_{aa} \epsilon_a^4 + d''_{aaa} \epsilon_a^6 \dots$$

Parameterizing the discretation effects (w/o FV corrections)

Strategy

Generically write $O = O^{\chi} + O^{m_{qs} \neq m^*_{qs}} + O^{\text{disc}}$

- Similar Taylor expansion for a-dependence, but now must also include radiative corrections from the Symanzik effective theory

$$
\begin{aligned}
m_\Omega^{\text{disc}} &= \tilde{d}'_a \epsilon_a^2 \alpha_S + d'_a \epsilon_a^2 + d'_{aa} \epsilon_a^4 + d'_{aaa} \epsilon_a^6 \ldots \\
w_0^{\text{disc}} &= \tilde{d}''_a \epsilon_a^2 \alpha_S + d''_a \epsilon_a^2 + d''_{aa} \epsilon_a^4 + d''_{aaa} \epsilon_a^6 \ldots
\end{aligned}
$$

More generally, $\delta_a^{\text{Symanzik}} \sim a^2 \alpha_S^{n+\gamma}$

- $n = 0$ for MDWF, $n = 1$ for HISQ
- γ unknown for our action

Proxies for the lattice spacing

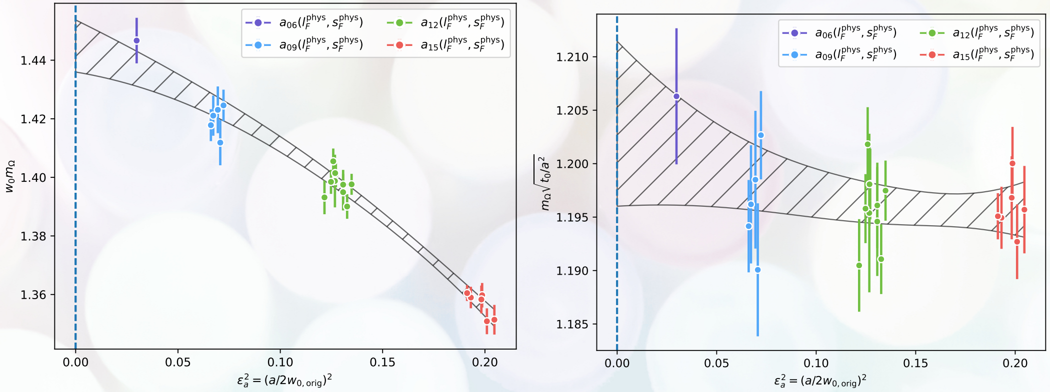
Strategy

- We define our small parameters such that $l_\Lambda^2 \sim s_\Lambda^2 \sim \epsilon_a^2$
- In previous work (for F_K/F_π) we parameterized the lattice dependence via $\epsilon^2_{a,\text{fixed}} = \frac{a^2}{(2w_{0,\text{orig}})^2}$
- But in this work, in order to give all the gradient flow scales equal footing, we also consider the parameterization

$$\epsilon^2_{a,\text{var}} = \begin{cases} \frac{a^2}{(2w_{0,\mathrm{orig}})^2}, & y = w_{0,\mathrm{orig}} m_\Omega \\ \frac{a^2}{(2w_{0,\mathrm{imp}})^2}, & y = w_{0,\mathrm{imp}} m_\Omega \\ \frac{a^2}{4t_{0,\mathrm{orig}}}, & y = \sqrt{t_{0,\mathrm{orig}}} m_\Omega \\ \frac{a^2}{4t_{0,\mathrm{imp}}}, & y = \sqrt{t_{0,\mathrm{imp}}} m_\Omega \end{cases}$$

Lattice spacing dependence

Strategy



Representative model: N^{3}LO w/ chiral logs, $\Lambda = 4\pi F_\pi$, $\epsilon_a = \epsilon_{a,\text{fixed}}$

Finite volume corrections

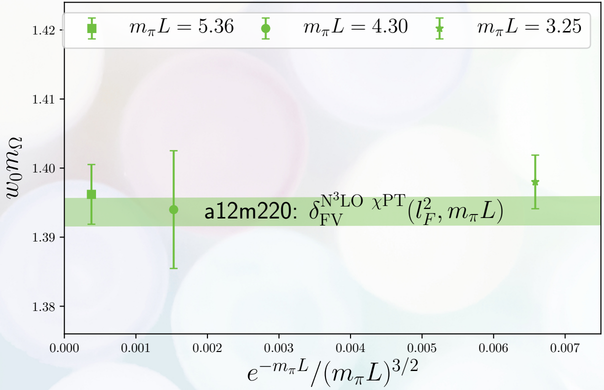
Strategy

Correct for finite volume by modifying the tadpole integral:

$$\log\left(\frac{m^2}{\mu^2}\right) \to \log\left(\frac{m^2}{\mu^2}\right) + 4k_1(mL)$$

where $k_1(mL) \sim e^{-ML}/(mL)^{3/2}$

- Determined corrections to N^3LO
- k_1 has no free parameters
- Conclusion: F.V. corrections are negligible



Generically write $O = O^{\chi} + O^{\mathrm{disc}} + O^{m_{qs} \neq m_{qs}^*}$

- Expand everything, remembering the cross-terms

$$w_0 m_\Omega = c_0 + \delta^{\mathrm{NLO}}_{ls,\Lambda} + \delta^{\mathrm{N^2LO}}_{ls,\Lambda} + \delta^{\mathrm{N^3LO}}_{ls,\Lambda} + \delta^{\mathrm{NLO}}_{a,\Lambda} + \delta^{\mathrm{N^2LO}}_{a,\Lambda} + \delta^{\mathrm{N^3LO}}_{a,\Lambda}$$

where

$$\delta^{\mathrm{NLO}}_{ls,\Lambda} = l^2_\Lambda c_l + s^2_\Lambda c_s\,,$$
$$\delta^{\mathrm{N^2LO}}_{ls,\Lambda} = l^4_\Lambda (c_{ll} + c^{ln}_{ll} \lambda_\pi) + l^2_\Lambda s^2_\Lambda c_{ls} + s^4_\Lambda c_{ss}\,,$$
$$\delta^{\mathrm{N^3LO}}_{ls,\Lambda} = l^6_\Lambda (c_{lll} + c^{ln}_{lll} \lambda_\pi + c^{ln^2}_{lll} \lambda^2_\pi) + l^4_\Lambda s^2_\Lambda \lambda_\pi c^{ln}_{lls} + l^4_\Lambda s^2_\Lambda c_{lls} + l^2_\Lambda s^4_\Lambda c_{lss} + s^6_\Lambda c_{sss}\,,$$

$$\delta^{\mathrm{NLO}}_{a,\Lambda} = d_a \epsilon^2_a + d'_a \alpha_S \epsilon^2_a\,,$$
$$\delta^{\mathrm{N^2LO}}_{a,\Lambda} = d_{aa} \epsilon^4_a + \epsilon^2_a \left(d_{al} l^2_\Lambda + d_{as} s^2_\Lambda\right)\,,$$
$$\delta^{\mathrm{N^3LO}}_{a,\Lambda} = d_{aaa} \epsilon^6_a + \epsilon^4_a (d_{aal} l^2_\Lambda + d_{aas} s^2_\Lambda) + \epsilon^2_a (d_{all} l^4_\Lambda + d_{als} l^2_\Lambda s^2_\Lambda + d_{ass} s^4_\Lambda)\,.$$

25 unknown LECs! Can we fit all of these terms?

- Instead of fitting the full expression, consider various truncations & model average

$$w_0 m_\Omega = \underbrace{c_0}_{\text{LO}} + \underbrace{\delta^{\text{NLO}}(l_F, s_F)}_{\text{chiral NLO}} + \underbrace{\delta^{\text{NLO}}_{a,F}}_{\text{disc NLO}}$$

$$w_0 m_\Omega = \underbrace{c_0}_{\text{LO}} + \underbrace{\delta^{\text{NLO}}(l_F, s_F)}_{\text{chiral NLO}} + \underbrace{\delta^{\text{N}^2\text{LO}}(l_F, s_F)}_{\text{chiral N}^2\text{LO}} + \underbrace{\delta^{\text{N}^2\text{LO}}_{\ln}}_{\text{chiral log N}^2\text{LO}} + \underbrace{\delta^{\text{NLO}}_{a,F}}_{\text{disc NLO}} + \underbrace{\delta^{\text{N}^2\text{LO}}_{a,F}}_{\text{disc N}^2\text{LO}} + \underbrace{\delta^{\text{N}^2\text{LO}}_{L,F}}_{\text{FV N}^2\text{LO}}$$

$$w_0 m_\Omega = \underbrace{c_0}_{\text{LO}} + \underbrace{\delta^{\text{NLO}}(l_F, s_F)}_{\text{chiral NLO}} + \underbrace{\delta^{\text{N}^2\text{LO}}(l_F, s_F)}_{\text{chiral N}^2\text{LO}} + \underbrace{\delta^{\text{N}^3\text{LO}}(l_F, s_F)}_{\text{chiral N}^3\text{LO}} + \underbrace{\delta^{\text{N}^2\text{LO}}_{\ln}}_{\text{chiral log N}^2\text{LO}} + \underbrace{\delta^{\text{N}^3\text{LO}}_{\ln}}_{\text{chiral logs N}^3\text{LO}} + \underbrace{\delta^{\text{NLO}}_{a,F}}_{\text{disc NLO}} + \underbrace{\delta^{\text{N}^2\text{LO}}_{a,F}}_{\text{disc N}^2\text{LO}} + \underbrace{\delta^{\text{N}^3\text{LO}}_{a,F}}_{\text{disc N}^3\text{LO}} + \underbrace{\delta^{\text{N}^2\text{LO}}_{L,F}}_{\text{FV N}^2\text{LO}} + \underbrace{\delta^{\text{N}^3\text{LO}}_{L,F}}_{\text{FV N}^3\text{LO}}$$

Setting priors with empirical Bayes

Strategy

- Stabilize fits with priors
- Because we tune to the physical strange quark mass, light quark mass dependence can be guesstimated by eye
- Assume higher-order LECs are of similar size
- But what about the rest?

Let $M = \{\Pi, f\}$ denote a model. Per Bayes's theorem:

$$p(\Pi|D, f) = \frac{p(D|\Pi, f)p(\Pi|f)}{p(D|f)}$$

Assuming a uniform distribution for the hyperpriors $p(\Pi|f)$:

$$\text{peak of } p(D|\Pi, f) \implies \text{peak of } p(\Pi|D, f)$$

where $p(D|\Pi, f) = \int d\theta\, P(D|\theta, \Pi, f)P(\theta|\Pi, f)$ is the marginal likelihood function

- Implemented in `lsqfit` via `lsqfit.empbayes_fit`

Setting priors with empirical Bayes

Strategy

Let $M = \{\Pi, f\}$ denote a model. Per Bayes's theorem:

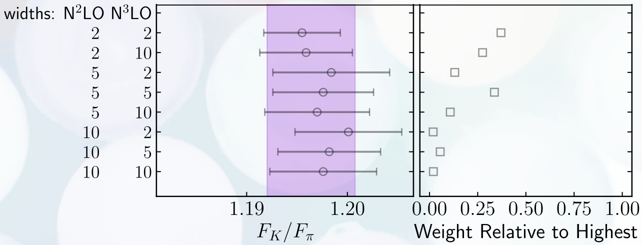
$$p(\Pi|D, f) = \frac{p(D|\Pi, f)p(\Pi|f)}{p(D|f)}$$

Assuming a uniform distribution for the hyperpriors $p(\Pi|f)$:

$$\text{peak of } p(D|\Pi, f) \implies \text{peak of } p(\Pi|D, f)$$

where $p(D|\Pi, f) = \int d\theta\, P(D|\theta, \Pi, f)P(\theta|\Pi, f)$ is the marginal likelihood function

- Shouldn't use for everything – only use for discretization terms
- Tune each model independently

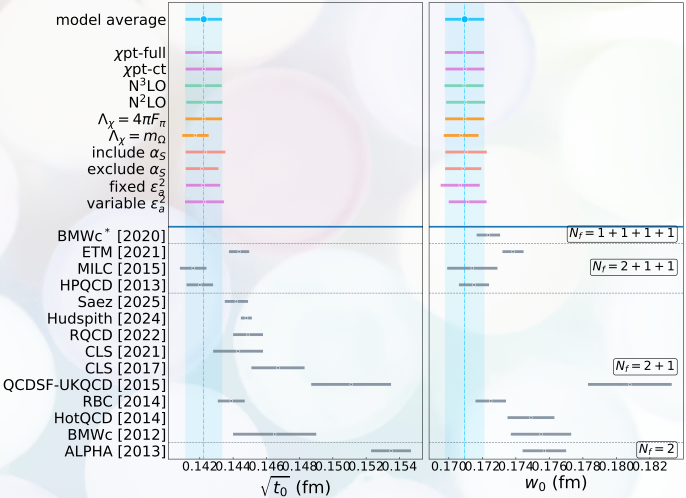


Results

Weight fits with Bayes factor:

- like a likelihood ratio, but Bayesian
- automatically penalizes more complicated models

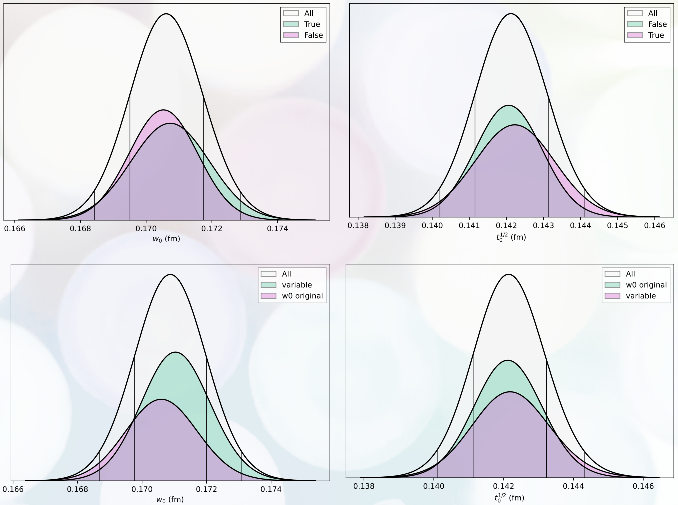
$\times 2$:	Expand to N^2LO or N^3LO
$\times 2$:	w/ or w/o chiral logs
$\times 2$:	incl./excl. F.V. corrections
$\times 2$:	incl./excl. $\alpha_S a^2$
$\times 2$:	$\Lambda = 4\pi F_\pi, M_\Omega$
$\times 2$:	$\epsilon_a = \epsilon_{a,\text{fixed}}, \epsilon_{a,\text{var}}$
64 :	total choices



Most important variations

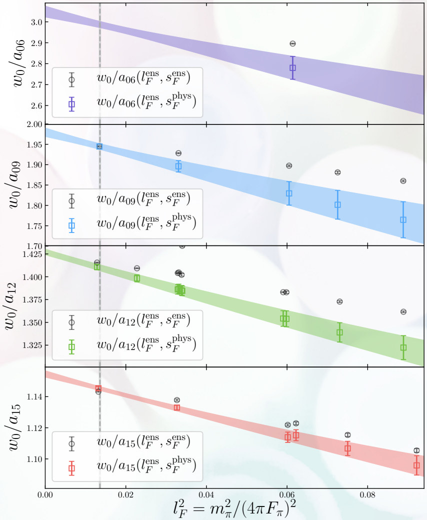
Results

- Models with $\Lambda_\chi = 4\pi F_\pi$ have much greater weight than $\Lambda_\chi = M_\Omega$
- Next largest variations come from including/excluding radiative $a^2\alpha_S$ discretization term (top), choice of $\epsilon_{a,\text{variable}}$ vs $\epsilon_{a,\text{fixed}}$ (bottom)



Measuring the lattice spacings

Results



We can determine the lattice spacing (in physical units) on each ensemble by fitting the N^{2}LO expression for each lattice spacing

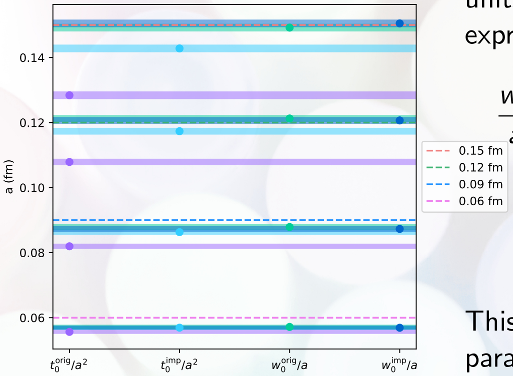
$$\frac{w_0}{a} = \frac{w_{0,\rm ch}}{a}\Big\{1 + k_l l_F^2 + k_s s_F^2 + k_a \epsilon_{a,\rm ch}^2 + k_{ll} l_F^4 + k_{lln} l_F^4 \ln(l_F^2) + k_{ls} l_F^2 s_F^2 + k_{ss} s_F^4 + k_{aa} \epsilon_{a,\rm ch}^4 + k_{al} l_F^2 \epsilon_{a,\rm ch}^2 + k_{as} s_F^2 \epsilon_{a,\rm ch}^2\Big\}$$

This requires us to use a different expansion parameter for ϵ_a

$$\epsilon_{a,\rm ch} = \frac{1}{(2 w_{0,\rm ch}/a)} .$$

Measuring the lattice spacings

Results



$$a = w_0/(w_{0,\mathrm{ch}}/a)$$

We can determine the lattice spacing (in physical units) on each ensemble by fitting the N^{2}LO expression for each lattice spacing

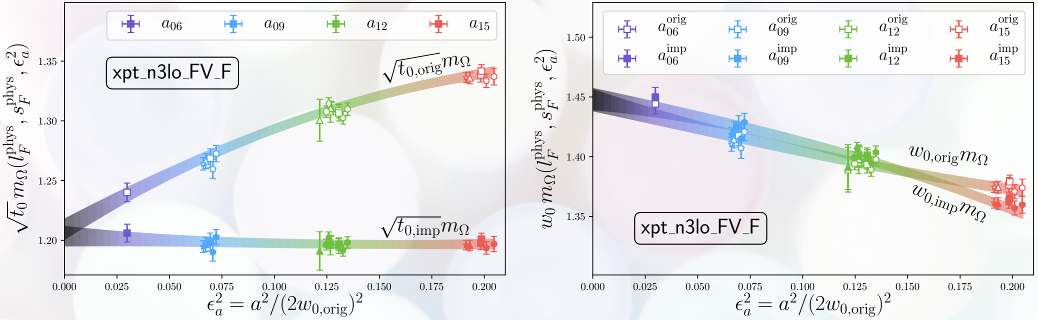
$$\frac{w_0}{a} = \frac{w_{0,\mathrm{ch}}}{a}\Big\{1 + k_l l_F^2 + k_s s_F^2 + k_a \epsilon_{a,\mathrm{ch}}^2 + k_{ll} l_F^4 + k_{lln} l_F^4 \ln(l_F^2) + k_{ls} l_F^2 s_F^2 + k_{ss} s_F^4 + k_{aa}\epsilon_{a,\mathrm{ch}}^4 + k_{al} l_F^2 \epsilon_{a,\mathrm{ch}}^2 + k_{as} s_F^2 \epsilon_{a,\mathrm{ch}}^2\Big\}$$

This requires us to use a different expansion parameter for ϵ_a

$$\epsilon_{a,\mathrm{ch}} = \frac{1}{(2w_{0,\mathrm{ch}}/a)} .$$

Continuum approach for different gradient flow scales

Results



- Although $\sqrt{t_{0,\mathrm{orig}}}$ and $\sqrt{t_{0,\mathrm{imp}}}$ agree in the continuum, their approaches are quite different!

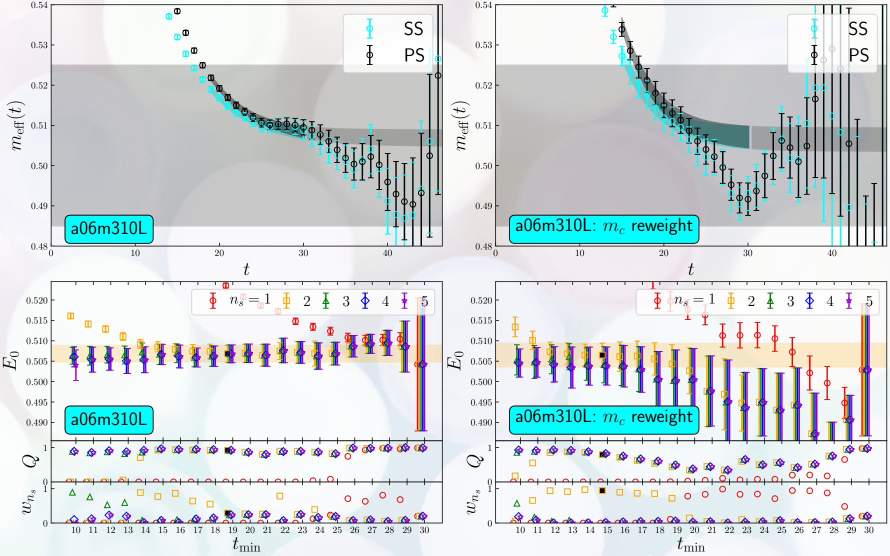
Charm reweighting (1/2)

Results

- On our finest ensemble, charm quark mass mistuned by about $\sim 10\%$ of the physical value
- Although this might sound rather mild, the size of the mistuning is $\sim$ the size of the strange quark mass
- Does the mistuning matter?

Charm reweighting (2/2)

Results



Updates

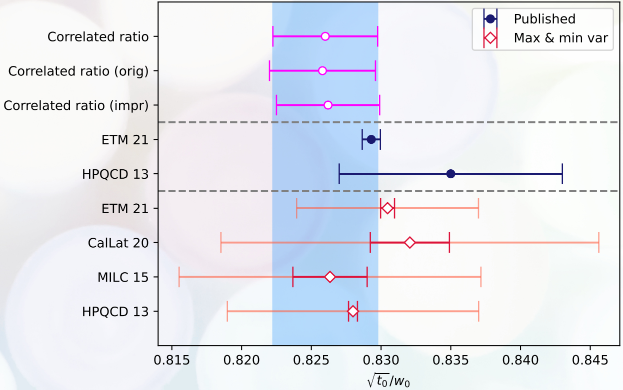
Ratio of flow scales $\sqrt{t_0}/w_0$ (preliminary)

Updates

- In CalLat20, fits for $m_\Omega w_0$ and $m_\Omega \sqrt{t_0}$ were treated separately $\rightarrow$ fit simultaneously to form correlated ratio $\sqrt{t_0}/w_0$ for each model
- This changes the model weights somewhat, leading to $< 1\sigma$ shifts in w_0, $\sqrt{t_0}$.

Ratio of flow scales $\sqrt{t_0}/w_0$ (preliminary)

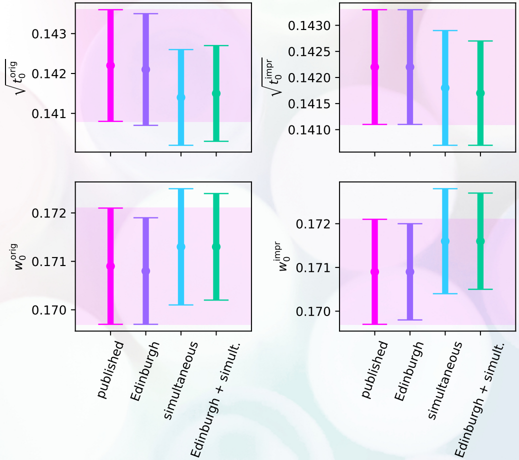
- In CalLat20, fits for $m_\Omega w_0$ and $m_\Omega \sqrt{t_0}$ were treated separately $\to$ fit simultaneously to form correlated ratio $\sqrt{t_0}/w_0$ for each model
- This changes the model weights somewhat, leading to $< 1\sigma$ shifts in w_0, $\sqrt{t_0}$.



$$\max/\min \text{ Var}\left(\frac{X}{Y}\right) = \left(\frac{X}{Y}\right)^2 \left[\left(\frac{\sigma_X}{X}\right)^2 + \left(\frac{\sigma_Y}{Y}\right)^2 \pm 2\frac{\sigma_X \sigma_Y}{XY}\right]$$

Redefining the physical point per the Edinburgh concensuSummarys (preliminary)

Updates



	Flag 2015	Edinburgh
m_π :	134.8(3) $\rightarrow$	135.0
m_K :	494.2(3) $\rightarrow$	494.6
f_π :	130.21(81) $\rightarrow$	130.5

- Using FLAG as input avoids QED, possible BSM corrections
- Edinburgh choice reduces physical point uncertainty in error budget, but numbers otherwise nearly identical

Bayesian model averaging

The issue of BMA can be roughly split into two parts:

- The estimation of model weights
- The procedure for computing the statistical, systematic errors

The issue of BMA can be roughly split into two parts:

- The estimation of model weights
- The procedure for computing the statistical, systematic errors

We concentrate on the latter point. *Desiderata* for our procedure:

- Able to separately estimate statistical, systematic errors
- Able to track correlations between different model-averaged observables
- Able to chain the results from model-averaged analysis $\mathcal{A}$ to a second model-averaged analysis $\mathcal{B}$

Quadrature

Bayesian model averaging

Popularized by Neil & Jay:

$$
\begin{aligned}
\mathrm{Var}[X] &= \mathrm{E}[X^2] - \mathrm{E}[X]^2 \\
&= \mathrm{E}[\mathrm{E}[X^2|M]] - \mathrm{E}[(\mathrm{E}[X|M])]^2 \\
&= \sum_M p(M)\left\langle X^2 \right\rangle_M - \left(\sum_M p(M)\left\langle X \right\rangle_M\right)^2 \\
&= \sum_M p(M)\mathrm{Var}(X)_M + \sum_M p(M)\left\langle X \right\rangle_M^2 - \left(\sum_M p(M)\left\langle X \right\rangle_M\right)^2
\end{aligned}
$$

Generalization to covariance follows directly from the law of total covariance:

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|M]]$$
$$\text{Cov}\,(X, Y) = \underbrace{\mathbb{E}[\text{Cov}\,(X, Y|\mathcal{M})]}_{\text{EVPC}} + \underbrace{\text{Cov}\,(\mathbb{E}[X|\mathcal{M}], \mathbb{E}[Y|\mathcal{M}])}_{\text{CHM}}$$

(EVPC: expectation value of process covariance; CHM: covariance of hypothetical means)

Is this model averaging?

Generalization to covariance follows directly from the law of total covariance:

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|M]]$$
$$\text{Cov}\,(X, Y) = \underbrace{\mathbb{E}[\text{Cov}\,(X, Y|\mathcal{M})]}_{\text{EVPC}} + \underbrace{\text{Cov}\,(\mathbb{E}[X|\mathcal{M}], \mathbb{E}[Y|\mathcal{M}])}_{\text{CHM}}$$

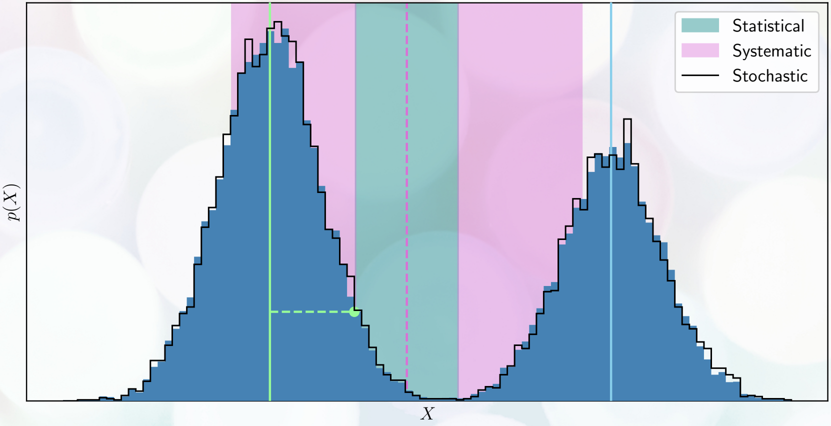
(EVPC: expectation value of process covariance; CHM: covariance of hypothetical means)
Is this model averaging? Sort of?

$$p(X, Y) = \sum_{\mathcal{M}} p(X, Y|\mathcal{M})p(\mathcal{M})$$

Total variance matches the marginal posterior, but none of the other moments are likely to match

Gaussian: Prediction

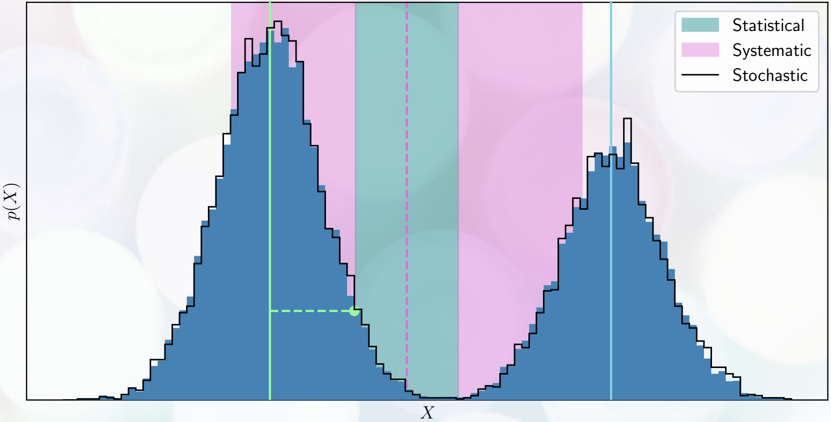
Bayesian model averaging



$$p(X, Y) \xrightarrow{\text{quadrature}} \mathcal{N}\left((\mathbb{E}[X], \mathbb{E}[Y]), \mathrm{EVPC}\,(X, Y) + \mathrm{CHM}\,(X, Y)\right)$$

Stochastic: Prediction

Bayesian model averaging



- Boyle et al (2024): for each resample, pick a model $M^r \sim p(M)$

Some other options

Bayesian model averaging

Desiderata for BMA:

- Able to separately estimate statistical, systematic errors
- Able to track correlations between different model-averaged observables
- Able to chain the results from model-averaged analysis $\mathcal{A}$ to a second model-averaged analysis $\mathcal{B}$

Some options:

- Fermilab/HPQCD/MILC [hep-lat/2411.09656]: requires quadrature, higher-order moments lost
- BMW [hep-lat/2407.10913]: parametric bootstrap; different predictions for statistical & systematic errors; can construct pathological examples
- Boyle et al [hep-lat/2406.19193]: stochastic, semi-parametric bootstrap; appears to underestimate/overestimate stat/sys errors compared to Fermilab
- Frison [hep-lat/2302.06550]: fully Bayesian; complicated

Summary & future work

- Finite volume corrections explicitly worked out to N^3LO, found to be negligible
- Even in cases of relatively severe charm quark mass mistuning, impact is relatively mild
- There is no standard procedure for BMA in our community
- For us, scale setting is our dominant uncertainty when extrapolating in physical units, e.g. in our current effort to measure the nucleon sigma term

Backup

Approximate marginal posterior as a Gaussian mixture:

$$X \sim \sum_M w_M N(\mu_M, \sigma_M) \to \sum_M w_M N(\mu_M, \sqrt{\lambda}\sigma_M)$$

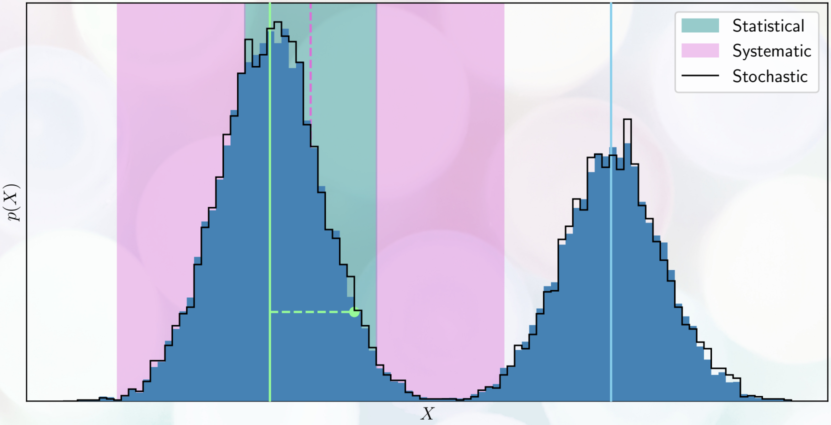
Let $I(\lambda) = Q(0.84; \lambda) - Q(0.16; \lambda)$ (with Q the quantile function) and

$$\sigma_{\mathrm{sys}}^2 + \lambda \sigma_{\mathrm{stat}}^2 \equiv I(\lambda)^2/4\,.$$

If the relationship between λ and $I(\lambda)^2$ is linear, statistical and systematic errors are uniquely defined by solving for two different values of λ.
Use median for best estimate.

BMW: Prediction

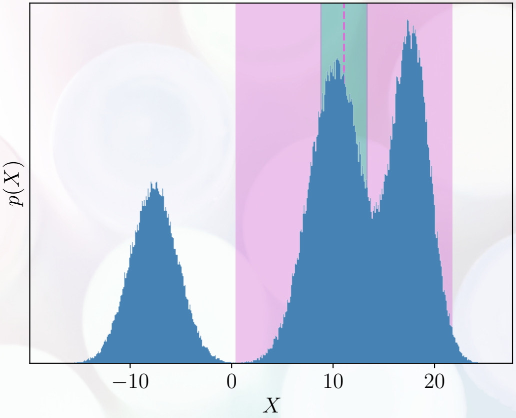
Backup

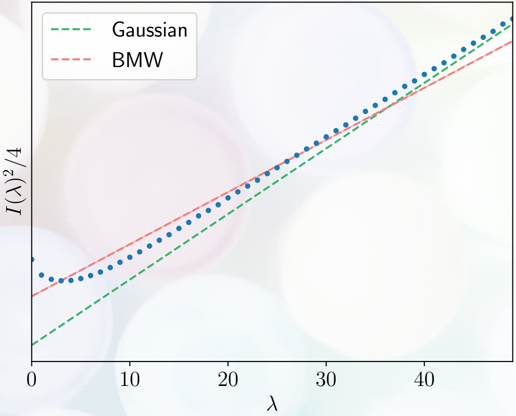


- Although error estimated with confidence intervals, this is not a confidence interval

BMW: A pathological example

Backup





- Must check that residuals are normal, but procedures still not guaranteed to match