NNV$_{UB}$: NEURAL NETWORKS FOR INCLUSIVE V_{ub}

1604.07598

PAOLO GAMBINO
UNIVERSITÀ DI TORINO & INFN

Importance of $|V_{xb}|$

V_{cb} plays an important role in the determination of UT

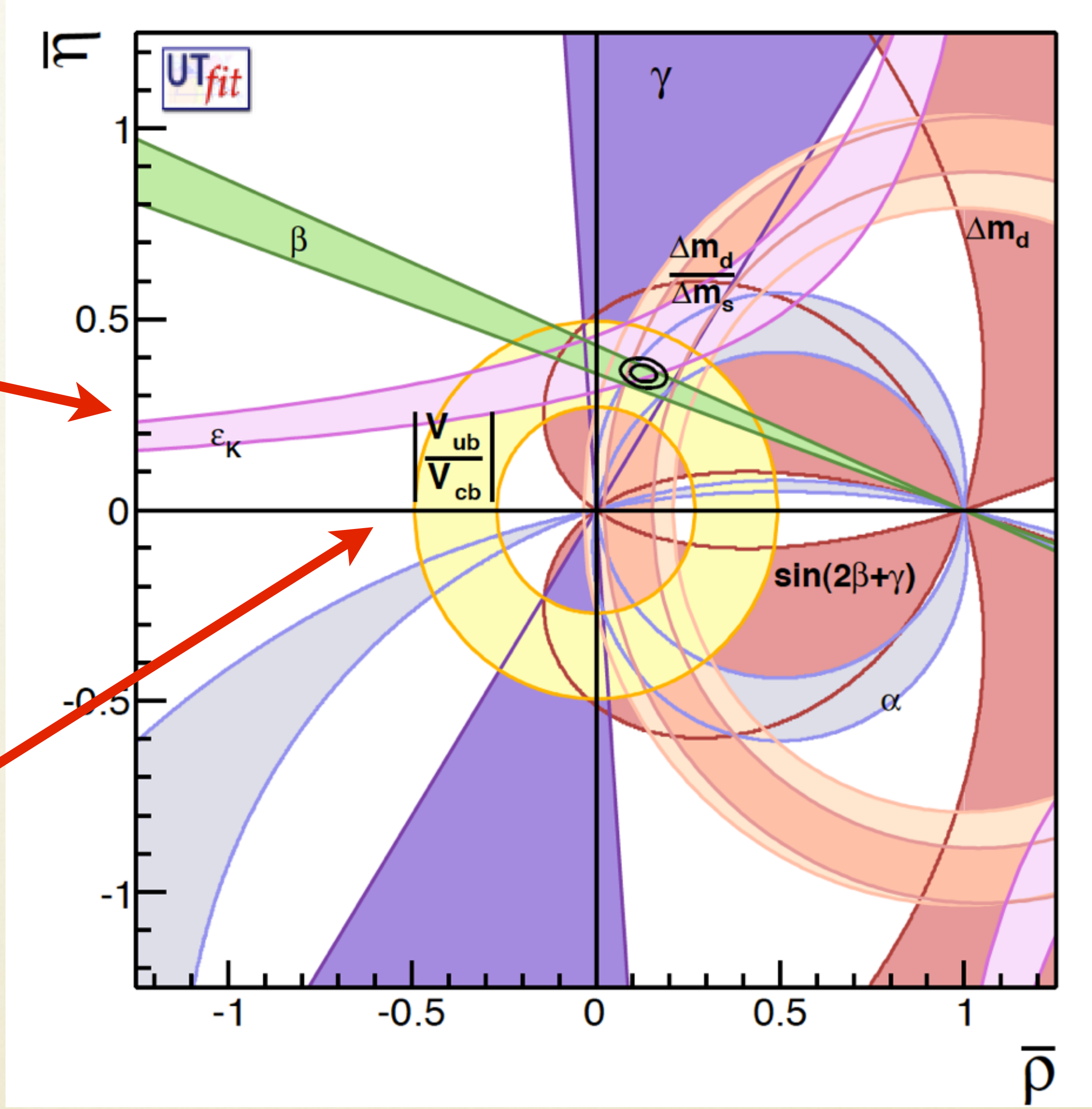
$$\varepsilon_K \approx x|V_{cb}|^4 + \ldots$$

and in the prediction of FCNC:

$$\propto |V_{tb}V_{ts}|^2 \simeq |V_{cb}|^2 \left[1 + O(\lambda^2)\right]$$

where it often dominates the theoretical uncertainty.

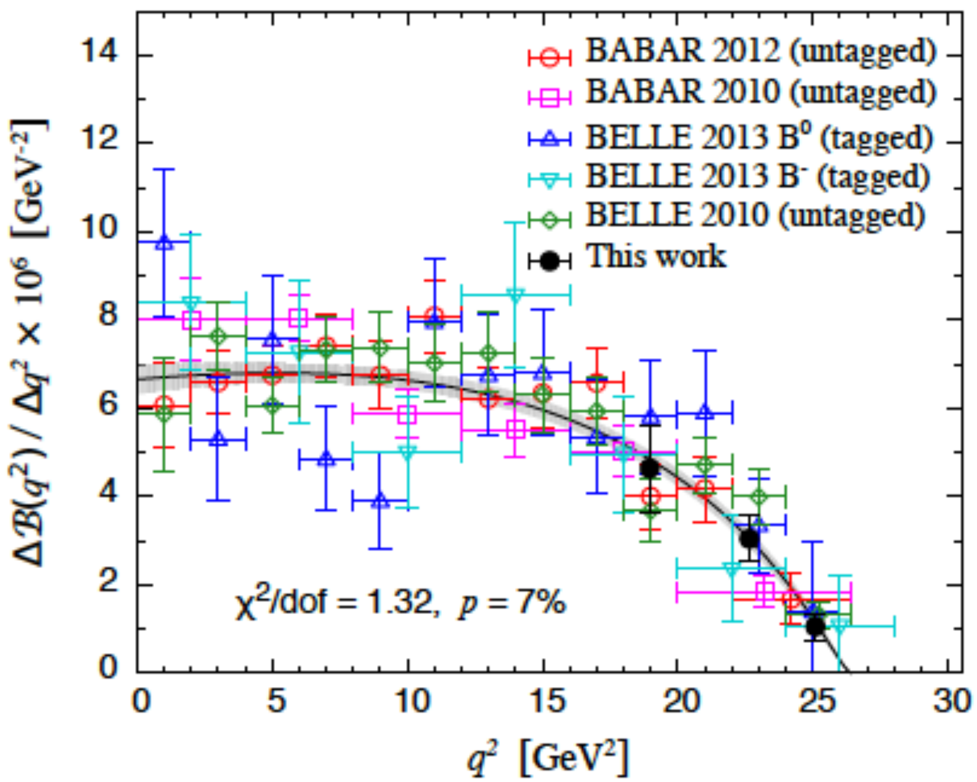
V_{ub}/V_{cb} constrains directly the UT



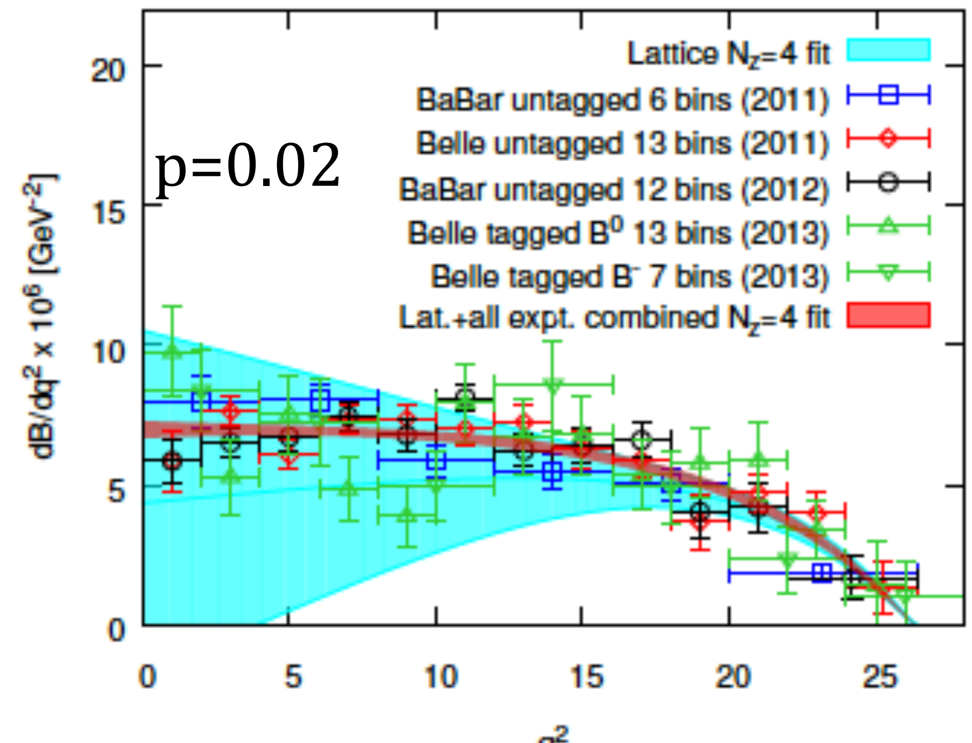
Since several years, exclusive decays prefer smaller $|V_{ub}|$ and $|V_{cb}|$ which cannot be SU(2)xU(1) invariant new physics Crivellin, Pokorski 1407.1320

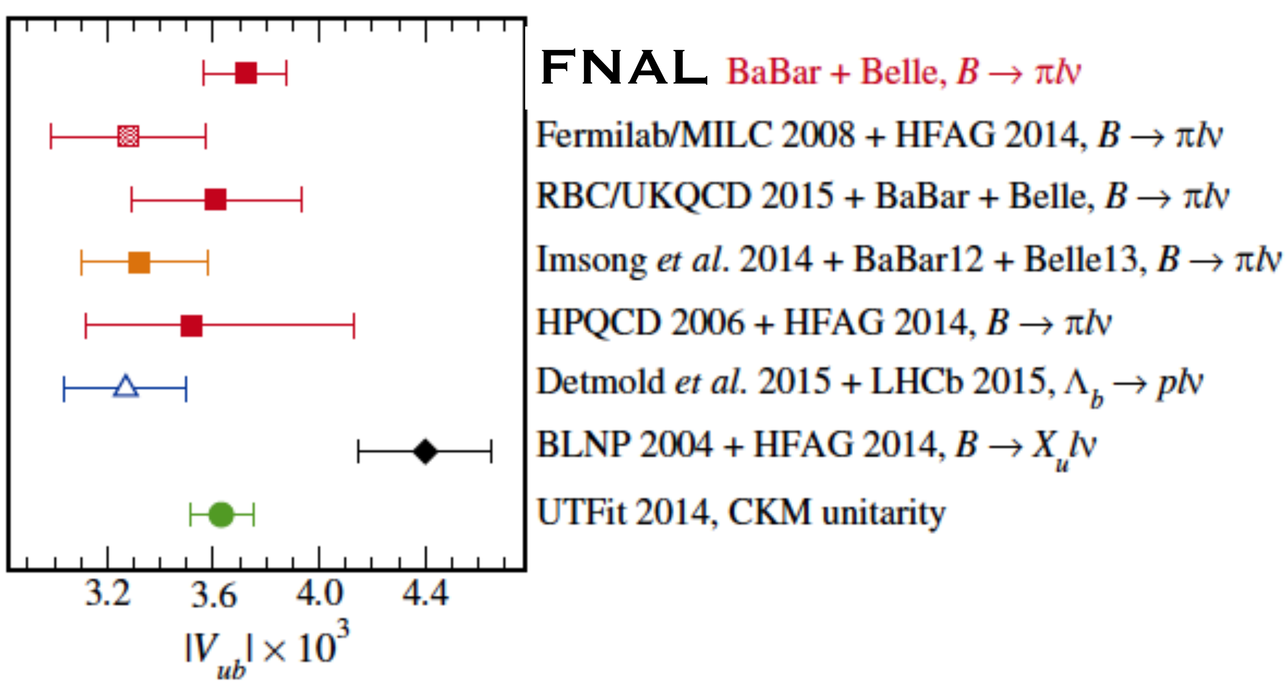
RECENT LATTICE $B \to \pi$ RESULTS

RBC/UKQCD 1501.05373



FNAL/MILC 1503.07839





FNAL 3.72(16) 10^{-3}
only 4.3% error
2.2σ from inclusive

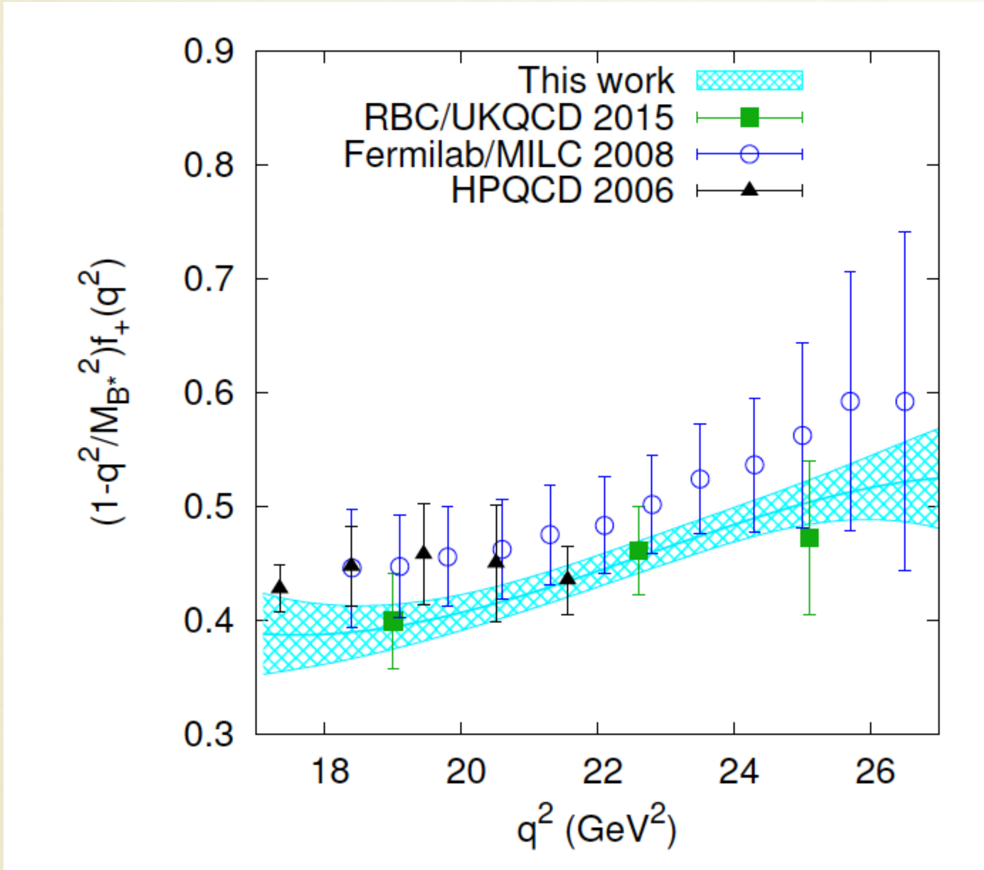
RBC/UKQCD 3.61(32) 10^{-3}
1.9σ from inclusive

LCSR 3.32(26) 10^{-3}
2.9σ from inclusive

LHCb depends on V_{cb} employed but low

Recent lattice results

1503.07839



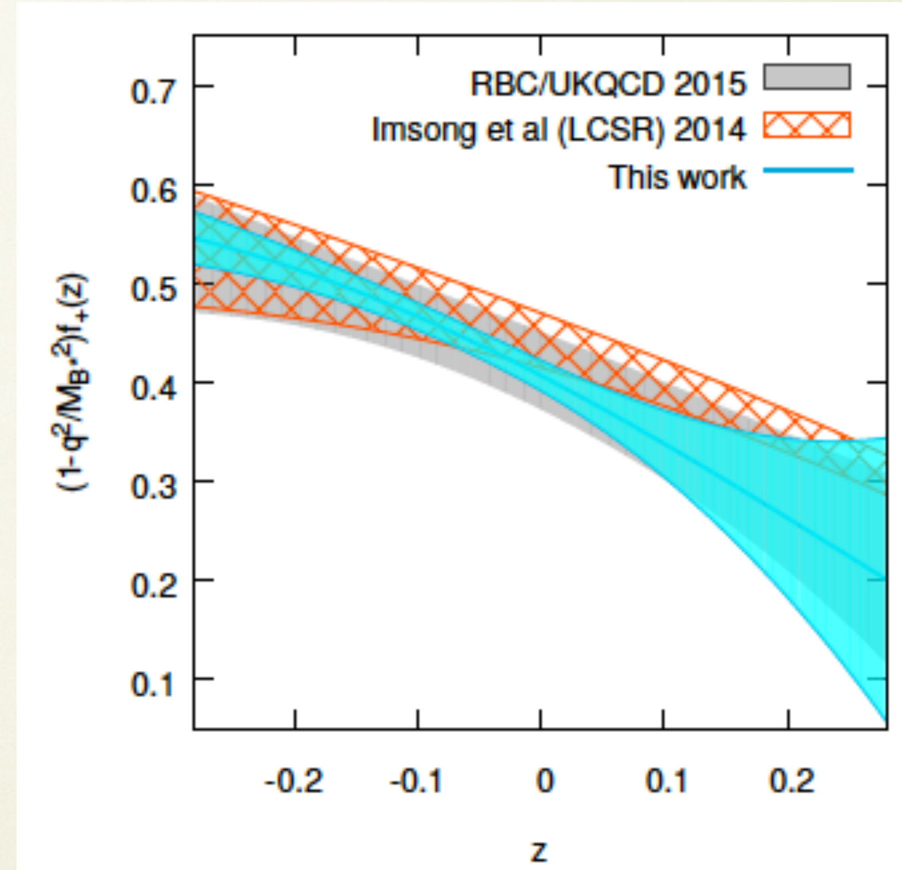


Table XVI. Results of the combined lattice+experiment fits with $N_z = 4$;.

Fit	χ^2/dof	dof	p value	b_0^+	b_1^+	b_2^+	b_3^+	$\lvert V_{ub}\rvert(\times 10^3)$
Lattice+exp.(all)	1.4	54	0.02	0.419(13)	−0.495(55)	−0.43(14)	0.22(31)	3.72(16)
Lattice+BaBar11 [7]	1.1	9	0.38	0.414(14)	−0.488(73)	−0.24(22)	1.33(44)	3.36(21)
Lattice+BaBar12 [8]	1.1	15	0.34	0.415(14)	−0.551(72)	−0.45(18)	0.27(41)	3.97(22)
Lattice+Belle11 [9]	0.9	16	0.55	0.412(13)	−0.574(65)	−0.40(16)	0.38(36)	4.03(21)
Lattice+Belle13 [10]	1.0	23	0.42	0.406(14)	−0.623(73)	−0.13(22)	0.92(45)	3.81(25)

Prospects: further improvements in LQCD, much more data @ BelleII, $B_s \to K l \nu$ and other channels @Belle-II and LHCb

Inclusive semileptonic B decays

OPE allows us to write inclusive observables as double series in Λ/m_b and α_s

$$M_i = M_i^{(0)} + \frac{\alpha_s}{\pi} M_i^{(1)} + \left(\frac{\alpha_s}{\pi}\right)^2 M_i^{(2)} + \left(M_i^{(\pi,0)} + \frac{\alpha_s}{\pi} M_i^{(\pi,1)}\right)\frac{\mu_\pi^2}{m_b^2}$$
$$+ \left(M_i^{(G,0)} + \frac{\alpha_s}{\pi} M_i^{(G,1)}\right)\frac{\mu_G^2}{m_b^2} + M_i^{(D,0)}\frac{\rho_D^3}{m_b^3} + M_i^{(LS,0)}\frac{\rho_{LS}^3}{m_b^3} + \dots$$

$$\mu_\pi^2(\mu) = \frac{1}{2M_B}\left\langle B\left|\bar{b}_v (i\vec{D})^2 b_v\right|B\right\rangle_\mu \qquad \mu_G^2(\mu) = \frac{1}{2M_B}\left\langle B\left|\bar{b}_v \frac{i}{2}\sigma_{\mu\nu}G^{\mu\nu} b_v\right|B\right\rangle_\mu$$

OPE valid for inclusive enough measurements, away from perturbative singularities ➠ semileptonic width, moments

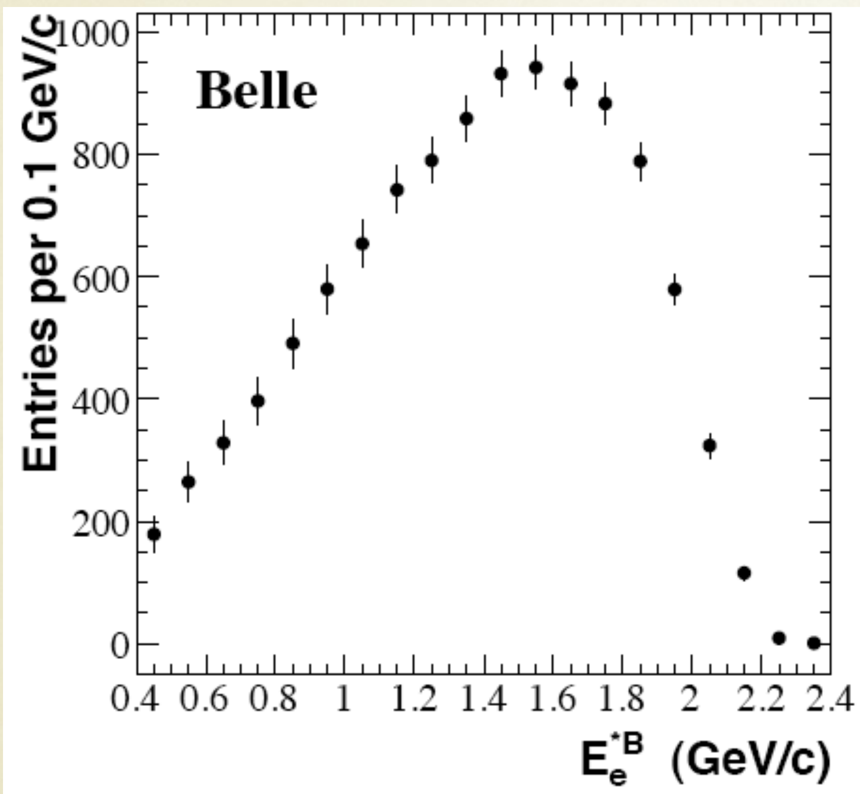
Current fits includes 6 non-pert parameters

$m_{b,c}$ $\quad \mu^2_{\pi,G}$ $\quad \rho^3_{D,LS}$

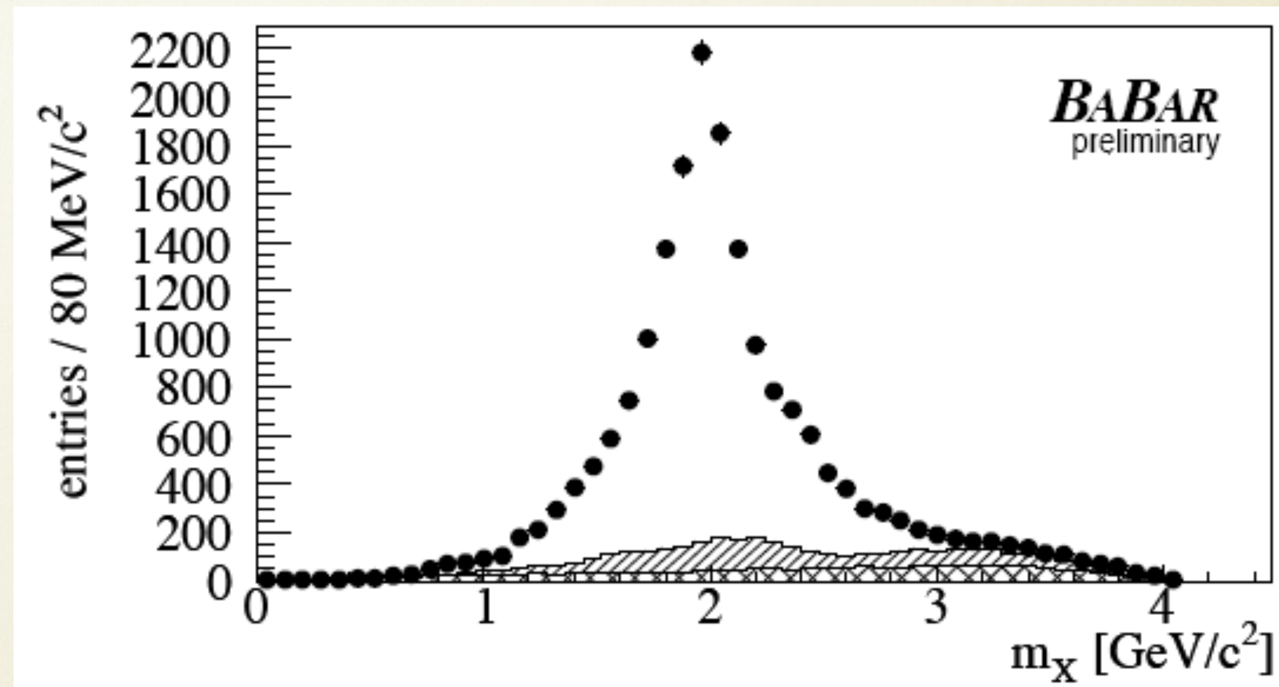
and all known corrections up to $O(\Lambda^3/m_b^3)$

Extraction of the OPE parameters

E_l spectrum



hadronic mass spectrum



*Global **shape** parameters (first moments of the distributions) tell us about m_b, m_c and the B structure, total **rate** about $|V_{cb}|$*

OPE parameters describe universal properties of the B meson and of the quarks → useful in many applications (rare decays, V_{ub},...)

FIT RESULTS

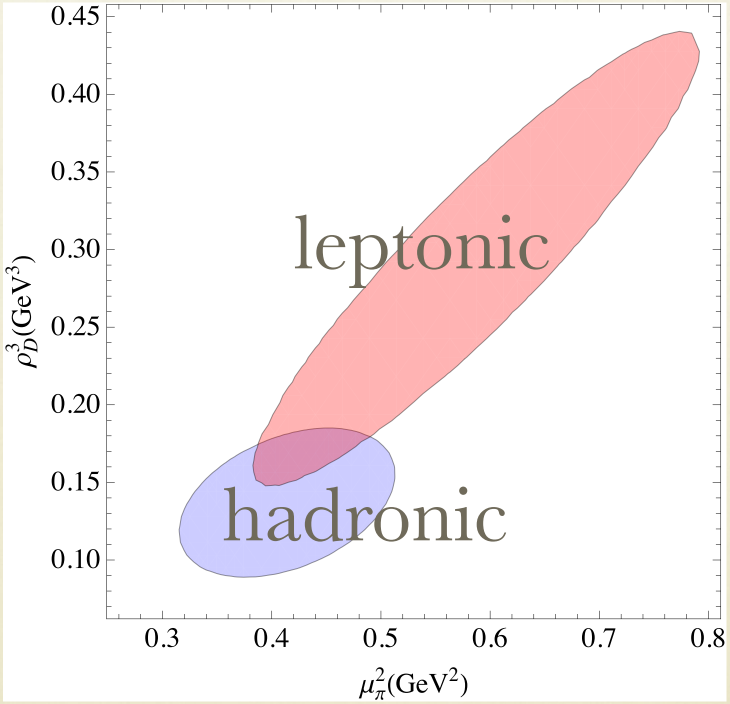
m_b^{kin}	$\overline{m}_c(3\,\text{GeV})$	μ_π^2	ρ_D^3	μ_G^2	ρ_{LS}^3	$\text{BR}_{c\ell\nu}$	$10^3\lvert V_{cb}\rvert$
4.553	0.987	0.465	0.170	0.332	-0.150	10.65	42.21
0.020	0.013	0.068	0.038	0.062	0.096	0.16	0.78

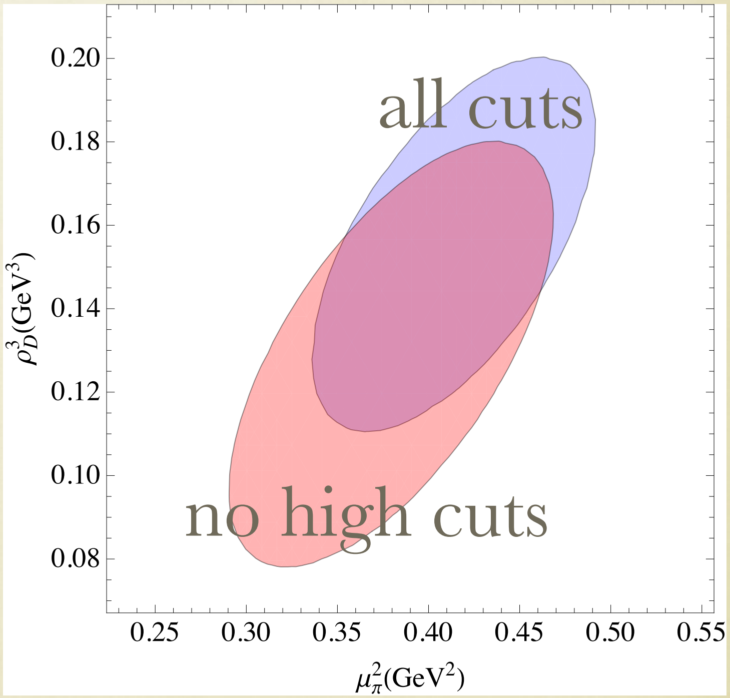
Alberti et al, 1411.6560

WITHOUT MASS CONSTRAINTS

$$m_b^{kin}(1\text{GeV}) - 0.85\,\overline{m}_c(3\text{GeV}) = 3.714 \pm 0.018\,\text{GeV}$$

- results depend little on assumption for correlations and choice of inputs, 1.8% determination of V_{cb}
- 20-30% determination of the OPE parameters





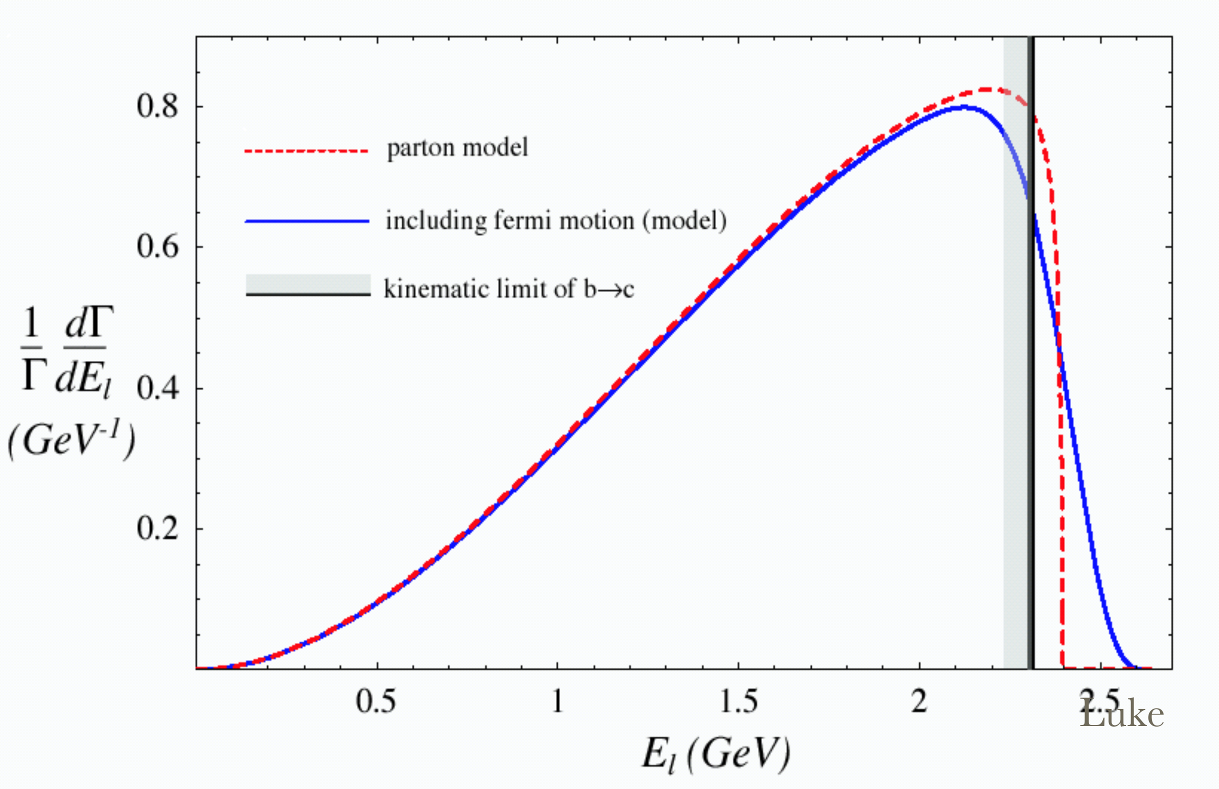
CUTS IN $B \to X_u l \nu$

Experiments often use kinematic cuts to avoid the $b \to c l \nu$ background:

$m_X < M_D$ $\quad E_l > (M_B^2 - M_D^2)/2M_B$ $\quad q^2 > (M_B - M_D)^2$ ...

The cuts destroy convergence of the OPE that works so well in b→c. OPE expected to work only away from pert singularities

Rate becomes sensitive to *local* b-quark wave function properties like Fermi motion. Dominant non-pert contributions can be resummed into a **SHAPE FUNCTION** f(k+). Equivalently the SF is seen to emerge from soft gluon resummation



How to access the SF?

$$\frac{d^3\Gamma}{dp_+ dp_- dE_\ell} = \frac{G_F^2 |V_{ub}|^2}{192\pi^3} \int dk C(E_\ell, p_+, p_-, k) F(k) + O\left(\frac{\Lambda}{m_b}\right)$$

Subleading SFs

OPE constraints
e.g. at $q^2=0$

$$\int_{-\infty}^{\bar{\Lambda}} k^2 F(k)\, dk = \frac{\mu_\pi^2}{3} + O(\frac{\Lambda^3}{m_b})$$ etc.

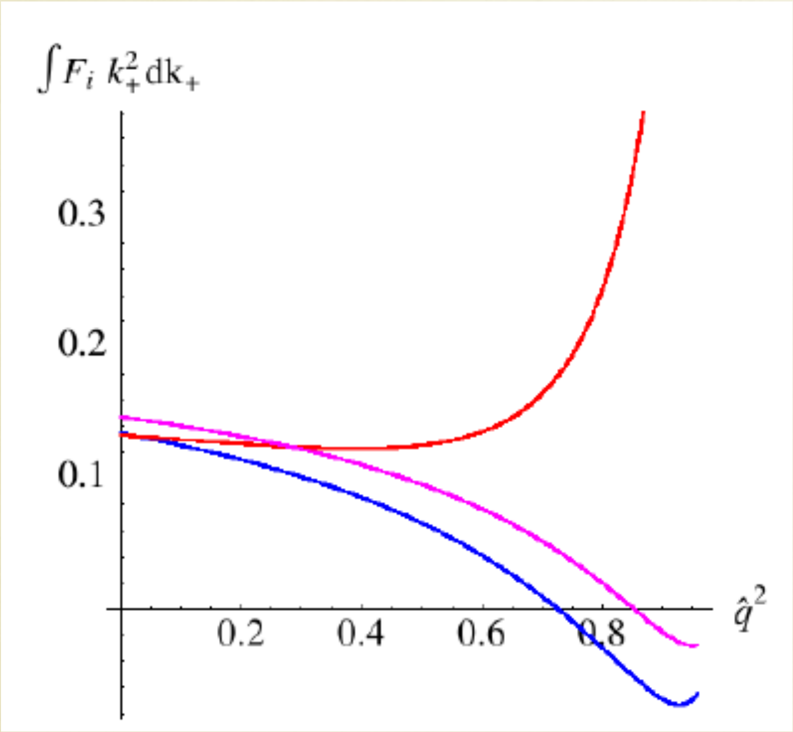
Predictions *based* on resummed pQCD DGE, ADFR	OPE constraints + parameterization without/with resummation GGOU, BLNP
OPE constraints + fit semileptonic (and radiative) data SIMBA, NNVub	

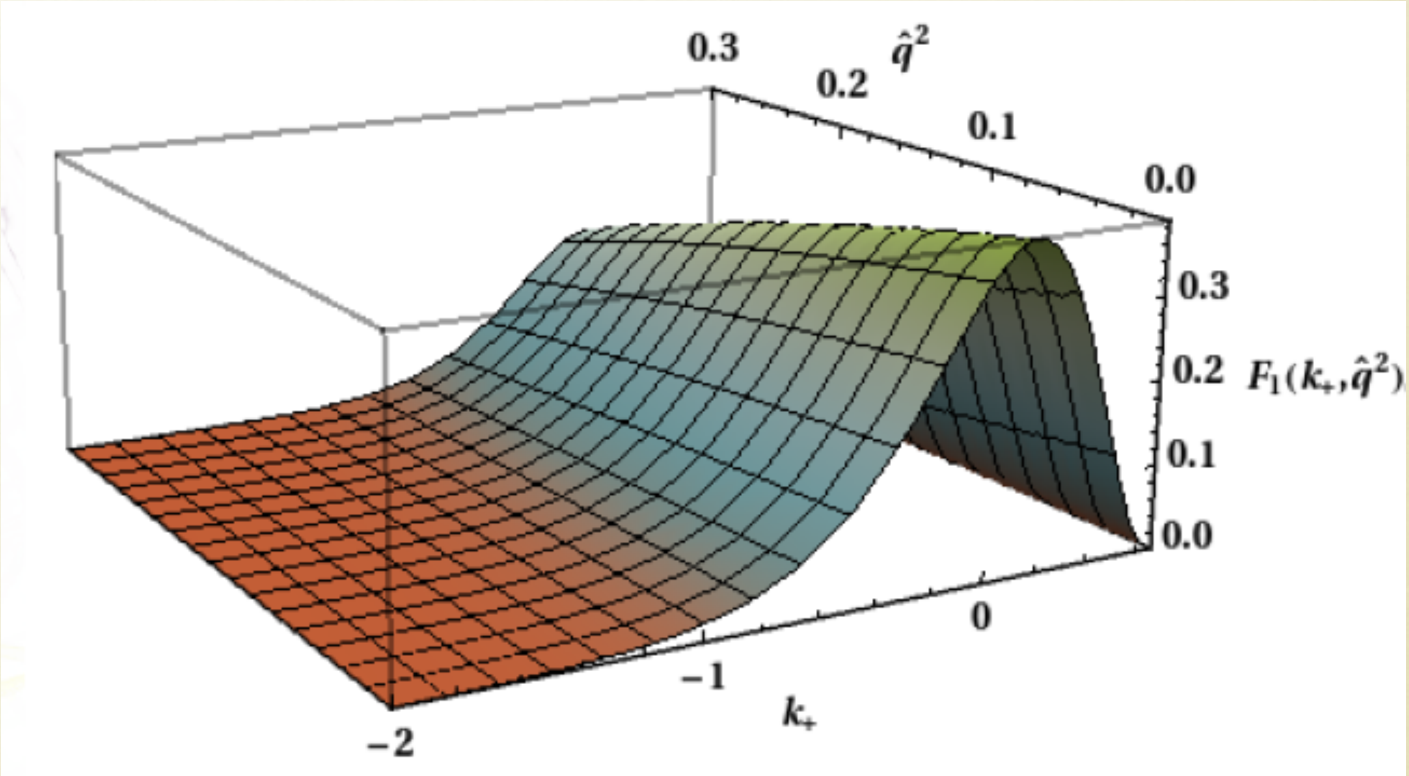
Shape functions in GGOU

$$W_i(q_0, q^2) \sim \int dk_+ \; F_i(k_+, q^2, \mu) W_i^{pert}\left[q_0 - \frac{k_+}{2}\left(1 - \frac{q^2}{m_b M_B}\right), q^2, \mu\right]$$

3 SFs, one for each form factor

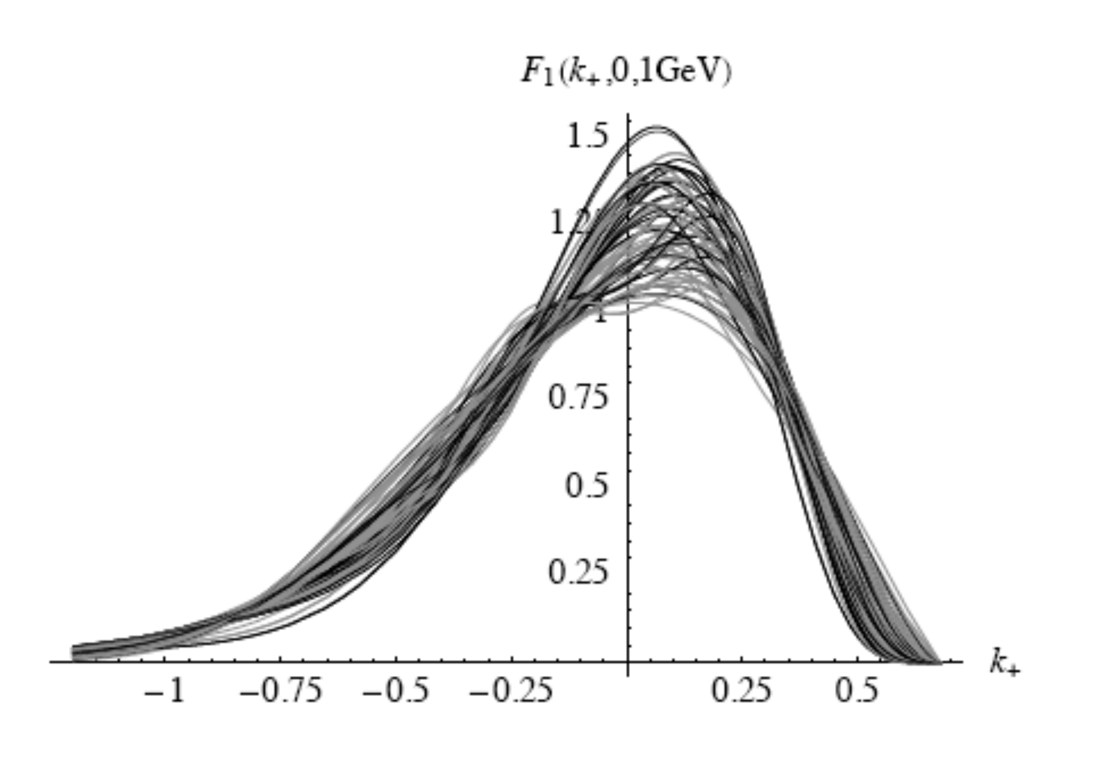
SF depend on q^2 through moments

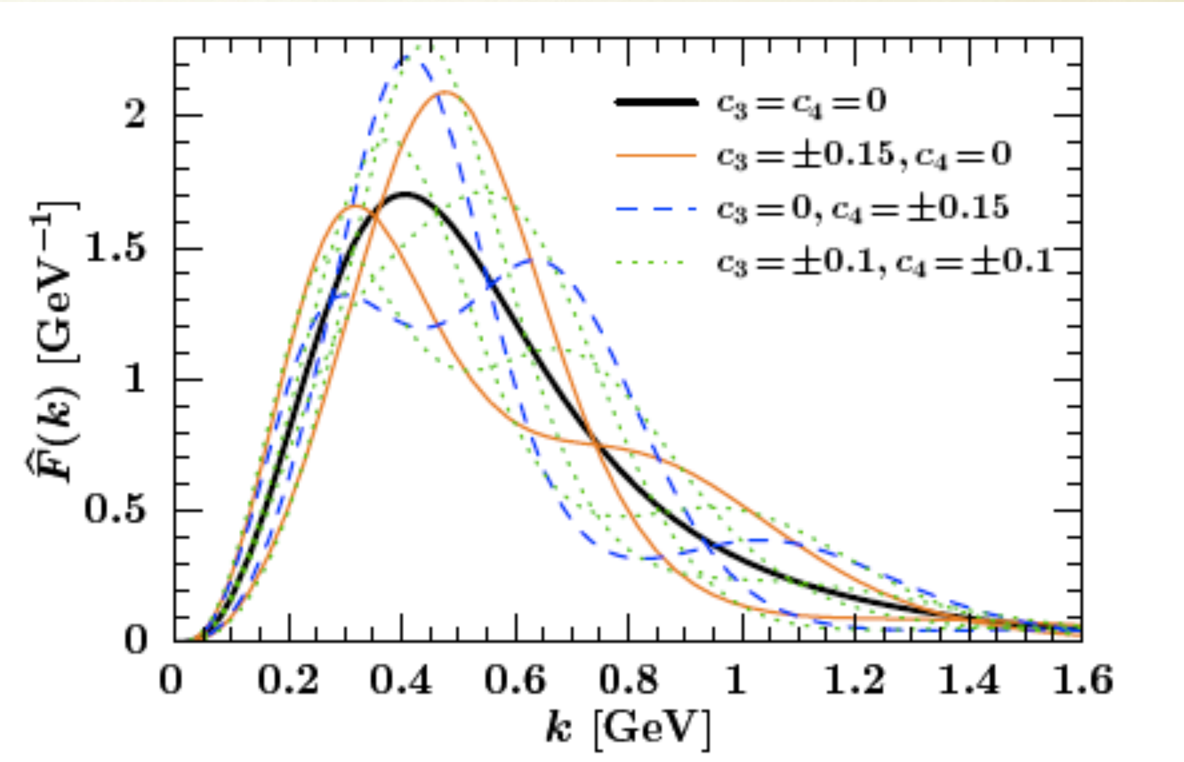




Each SF parameterized by simple 2-parameter functional forms

Functional forms





About 100 forms considered in GGOU, large variety, double max discarded. Small uncertainty (1-2%) on V_{ub}

A more systematic method by Ligeti et al. arXiv:0807.1926 Plot shows 9 SFs that satisfy all the first three moments

Only 2 parameters FF, is that good enough?

see SIMBA talk in this session

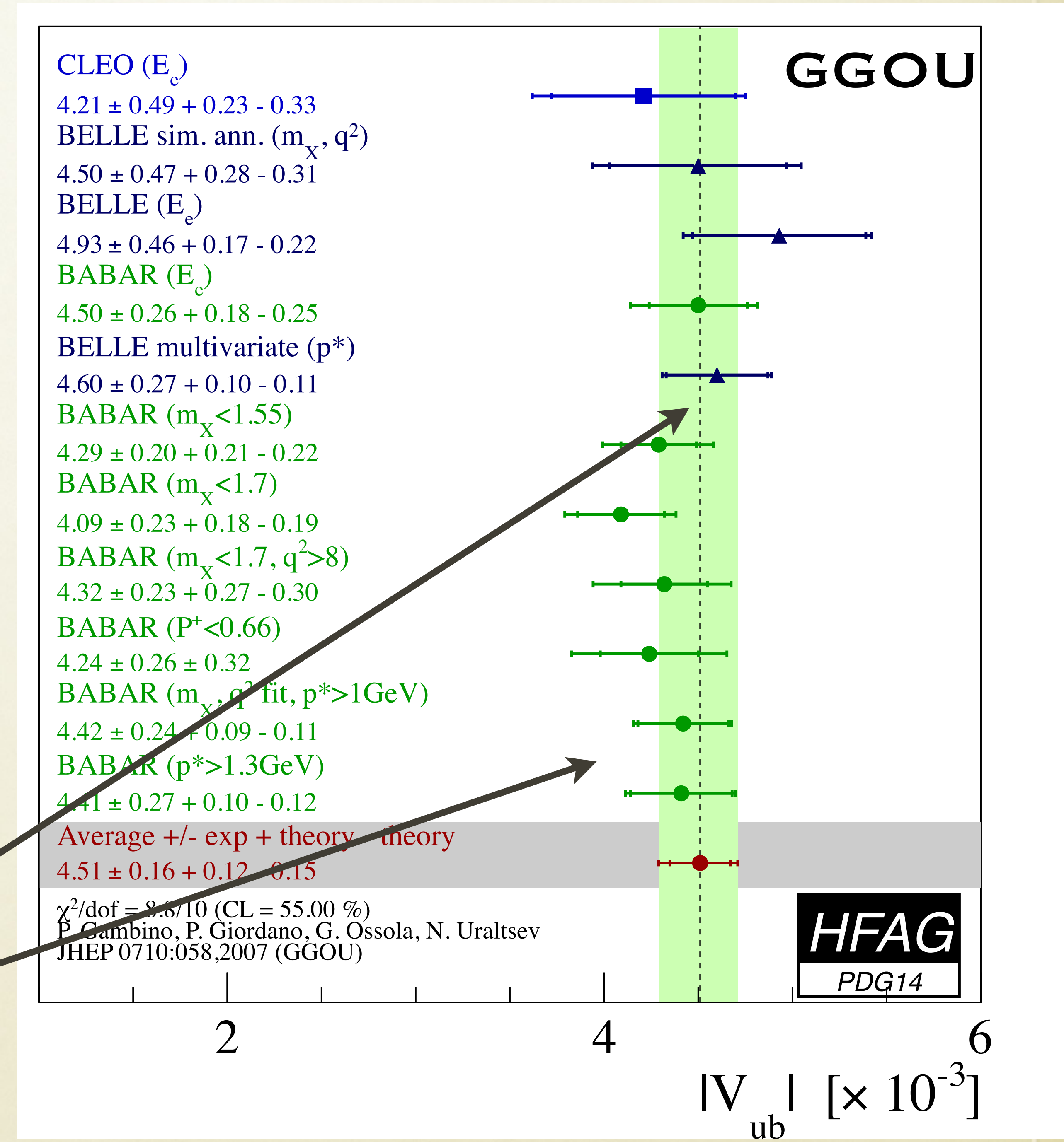
$|V_{ub}|$ DETERMINATIONS

Inclusive: 5% total error

HFAG 2014	Average lV
DGE	4.52(16)(16)
BLNP	4.45(16)(22)
GGOU	4.51(16)(15)

UT fit (without direct V_{ub}):
V_{ub}=3.66(12) 10^{-3}

Recent experimental results are theoretically cleanest (2%) but based on background and signal simulation…

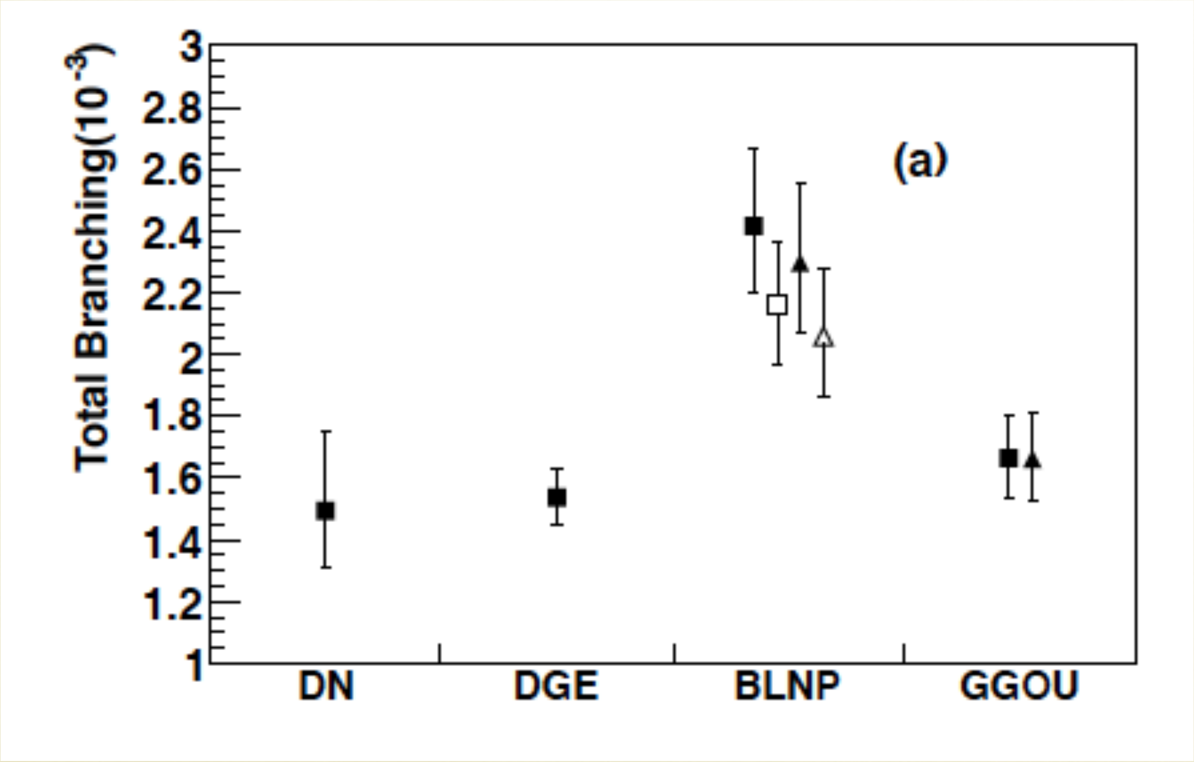


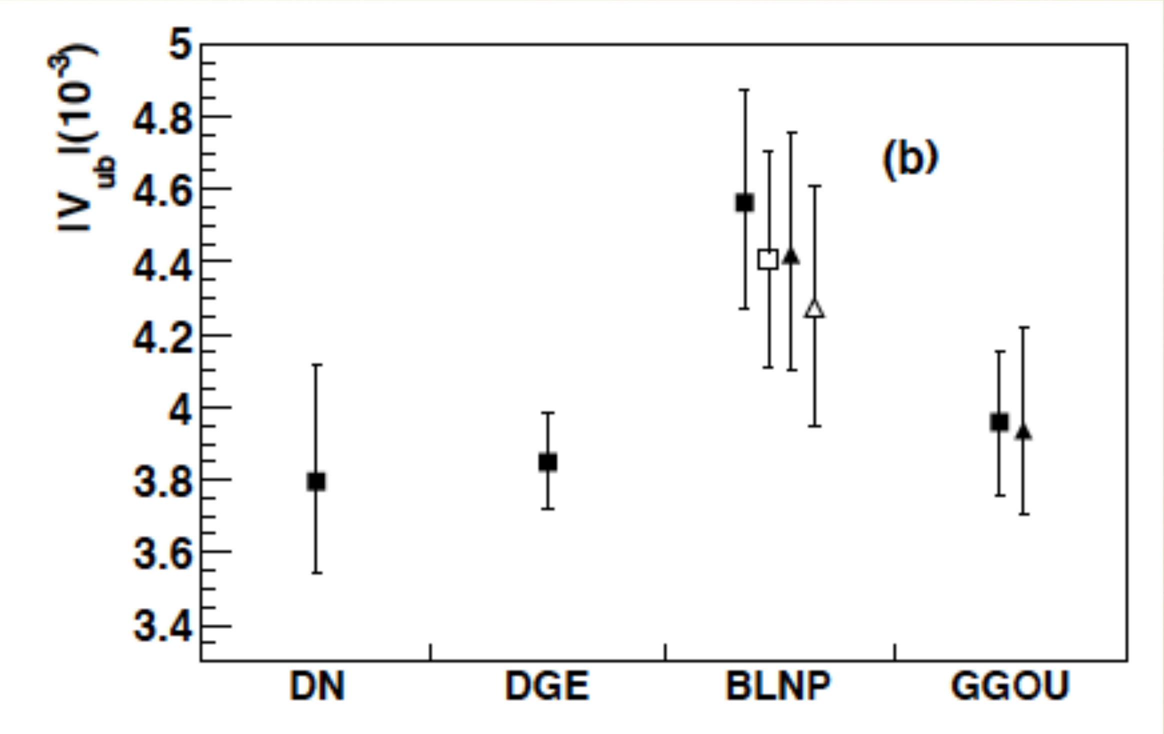
NEW Babar endpoint analysis

1611.05624

High sensitivity of the BR on the shape of the signal in the endpoint region. <u>Single most precise measurement to date</u>

$$\text{GGOU: } |V_{ub}| = (3.96 \pm 0.10_{exp} \pm 0.17_{th}) \times 10^{-3}$$





What happens if same is done in other BaBar analyses?

NB Belle multivariate analysis uses GGOU+DN for the inclusive part

Main issues in V_{ub} inclusive

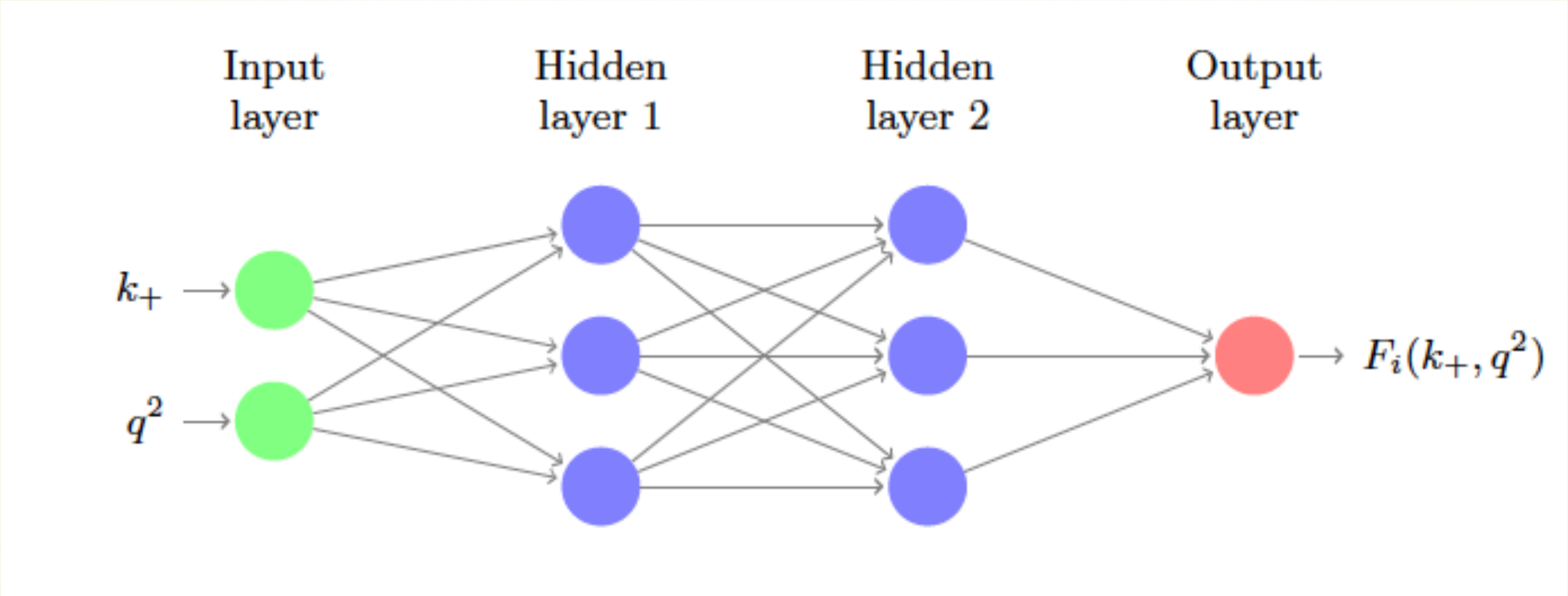
- Limited knowledge of **leading and subleading SFs**
- Nonperturbative effects in **high q^2 tail**, including Weak Annihilation (strongly constrained by charm semileptonic decays, < 1% in V_{ub})
- Potential role of **NNLO corrections**

Kamenik, PG
Ligeti,Luke, Manohar
Bigi et al

Brucherseifer, Caola, Melnikov
Greub, Neubert, Pecjak

The NNVub Project

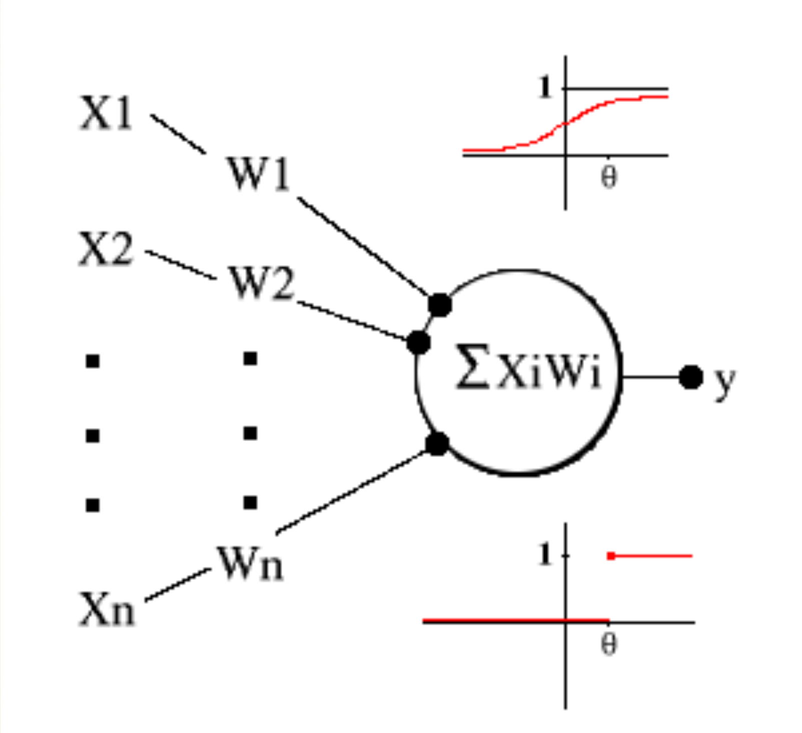
K.Healey, C. Mondino, PG, 1604.07598



- Use **Artificial Neural Networks** to parameterize shape functions without bias and extract V_{ub} from theoretical constraints and data, together with HQE parameters in a model independent way (without assumptions on functional form). Similar to NNPDF. Applies to $b \to u l \nu$, $b \to s \gamma$, $b \to s l^+ l^-$

- Belle-II will be able to measure some kinematic distributions, thus constraining directly the shape functions. NNVub will provide a flexible tool to analyse data.

Artificial Neural Networks



- NN provide **unbiased parameterization** of a continuous function: in the limit of infinite nodes they are universal approximators, highly non-linear functions
- Neuron **activates** if weighted input is positive, sigmoid gives smooth activation
- **Weights are trained** to reproduce desired response: random weights undergo random modifications, retaining only those that improve response (e.g. better χ^2): genetic algorithm → replicas
- Used in pattern recognition, **computationally intensive**, data-driven

Our Training Process

- Three Neural Networks : $F_{1,2,3}(k_+, q^2)$ (2 in/ 1 out)
- Goodness Of Fit (χ^2/dof) from moments:

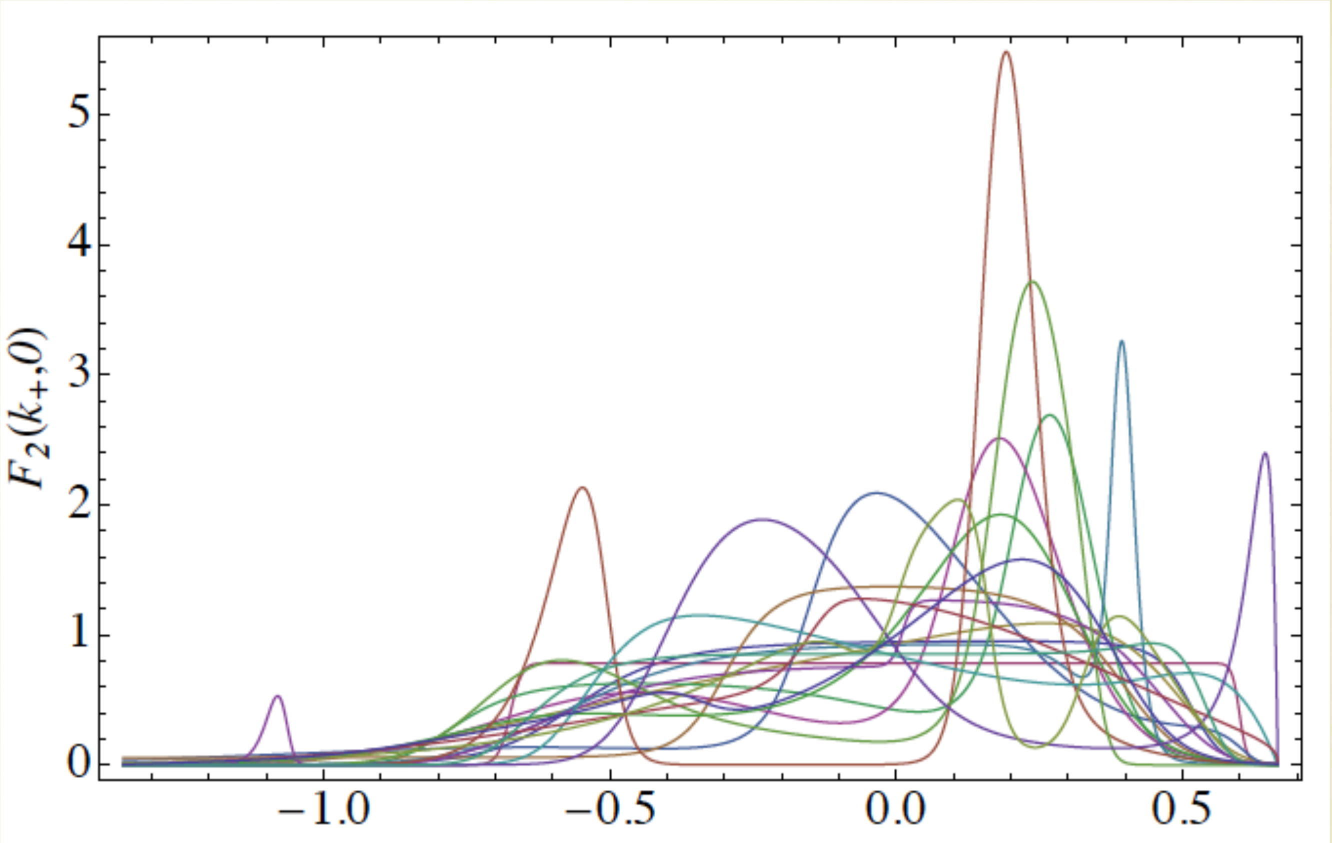
$$\int dk_+ \; k_+^n \; F_i(k_+, q^2) = \left(\frac{2\, m_b}{\Delta}\right)^n \left[\delta_{n0} + \frac{I_i^{(n),pow}}{I_i^{(0),\mathrm{tree}}}\right]$$

Generated over a sampling of $q^2 \in [0, 13\mathrm{GeV}^2]$

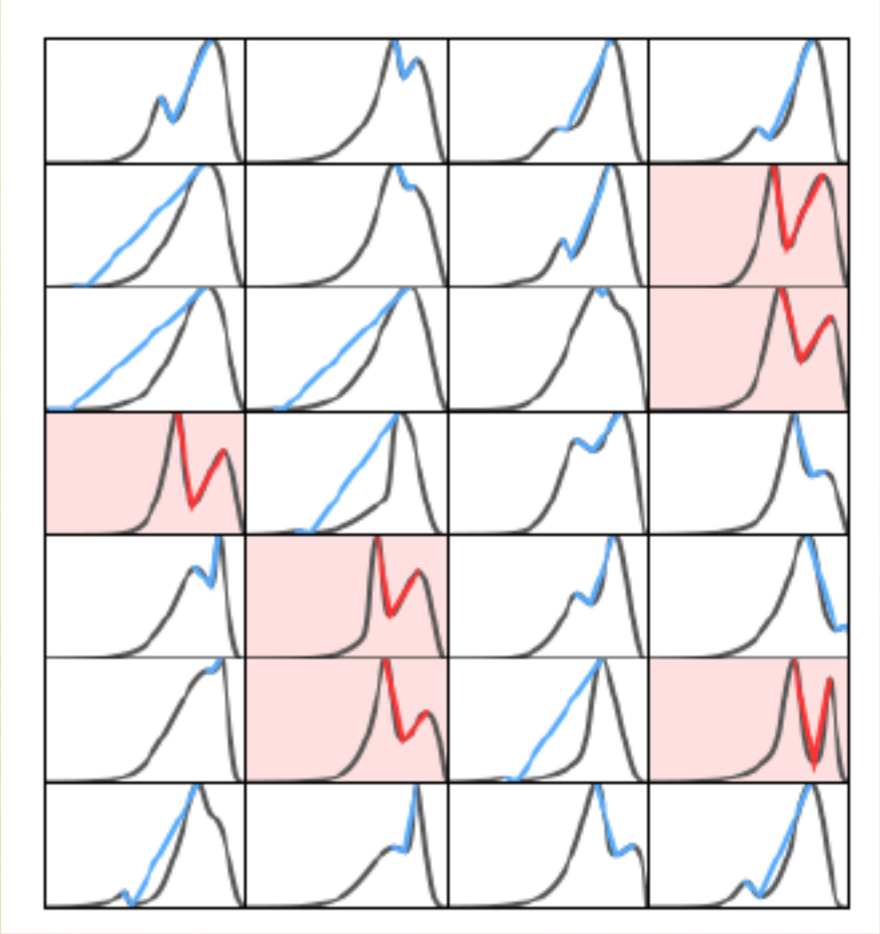
- $\chi^2 = \sum (desired - actual)_i [C^{-1}]_{ij} (desired - actual)_j$
- Stopping Criteria of $\chi^2/(dof) = 1$
- Genetic Algorithm :
  Randomly (MC) choose weights and vary them to create new children. Vary many times, choosing "best child", it becomes parent, repeat.

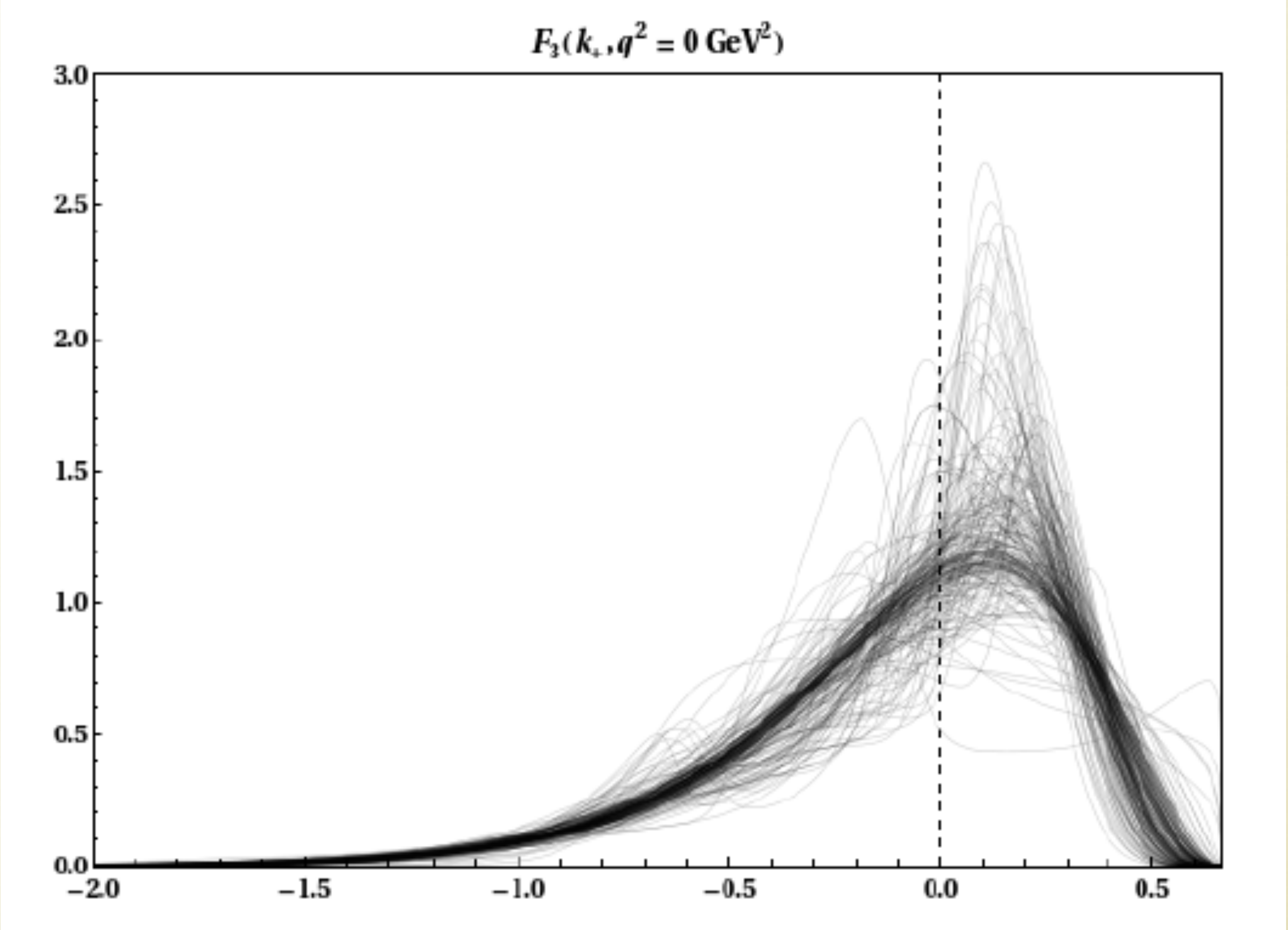
K.Healey

Selection of NN replicas trained on the first three moments only. They are not sufficient. But we know photon spectrum in bsgamma: single peak dominance, not too steep

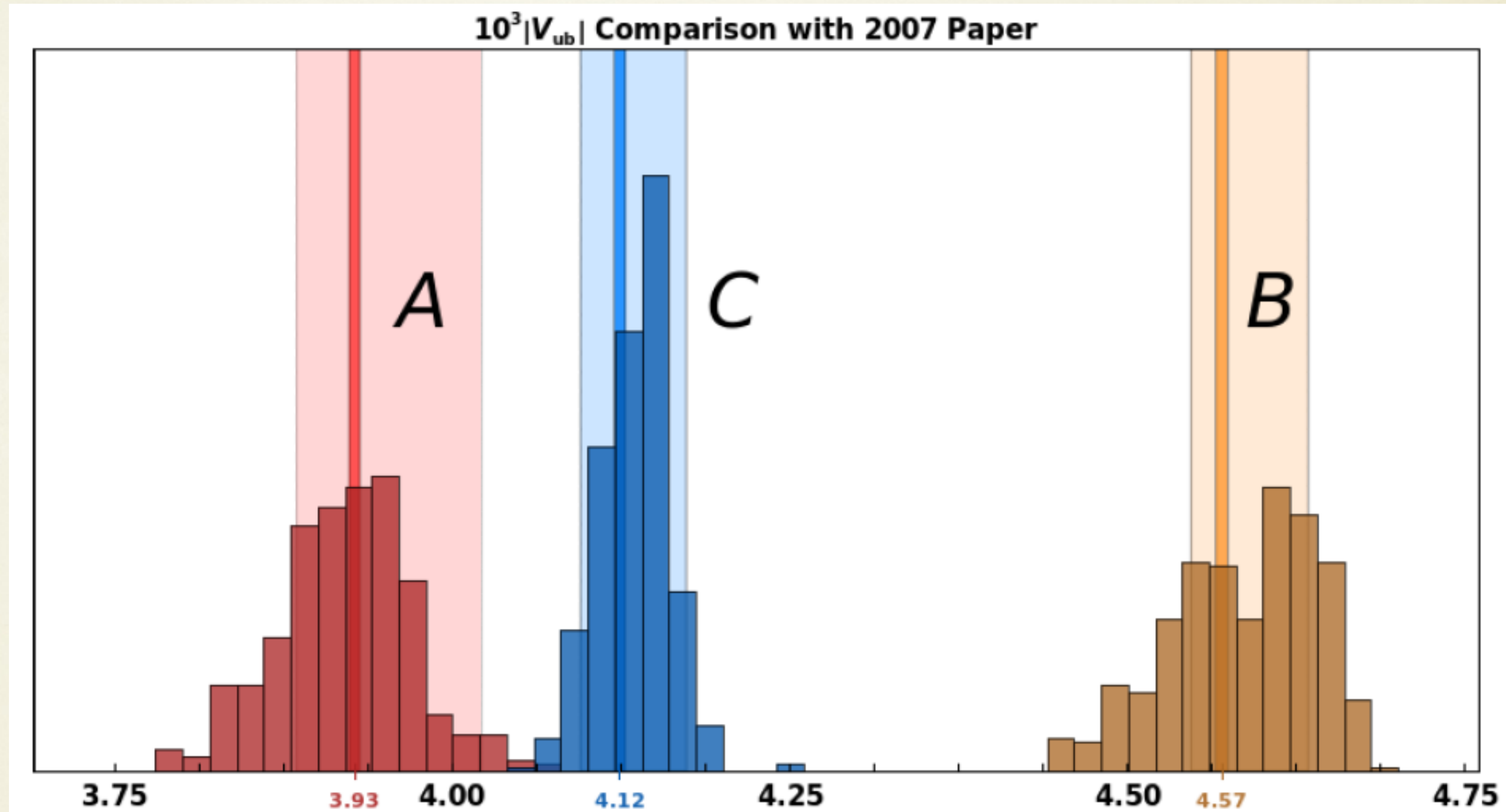


Beware: sampling can be biased by implementation, e.g. random initialization, or selection based on training speed





Comparison with 2007 paper, same inputs



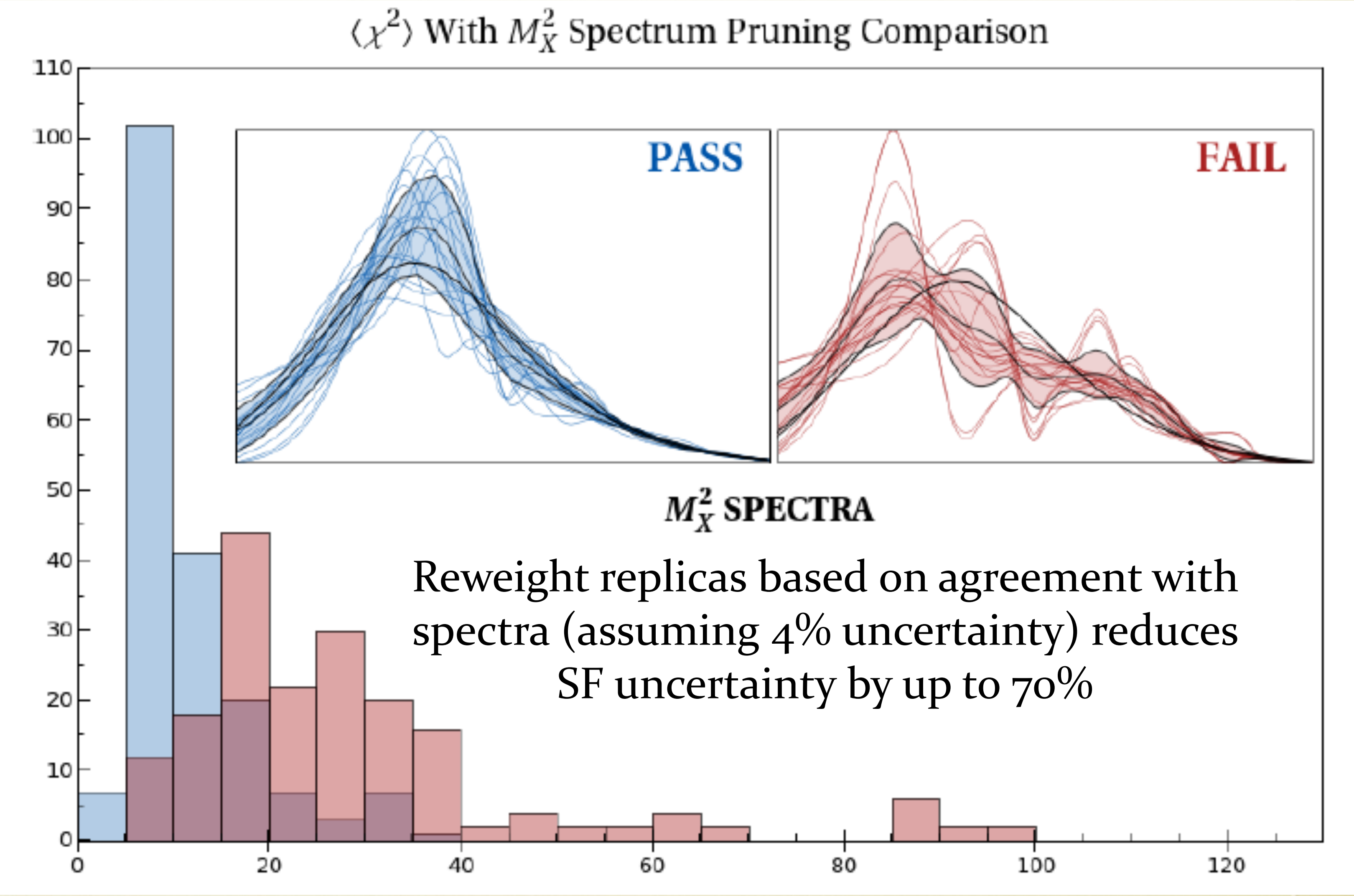
NNVub GGOU(HFAG 2014)

Experimental cuts (in GeV or GeV^2)	$\lvert V_{ub}\rvert \times 10^3$	$\lvert V_{ub}\rvert \times 10^3$[15]
$M_X < 1.55, E_\ell > 1.0$ Babar [44]	$4.30(20)(^{26}_{27})$	$4.29(20)(^{21}_{22})$
$M_X < 1.7, E_\ell > 1.0$ Babar [44]	$4.05(23)(^{19}_{20})$	$4.09(23)(^{18}_{19})$
$M_X \leq 1.7, q^2 > 8, E_\ell > 1.0$ Babar[44]	$4.23(23)(^{26}_{28})$	$4.32(23)(^{27}_{30})$
$E_\ell > 2.0$ Babar [41]	$4.47(26)(^{22}_{27})$	$4.50(26)(^{18}_{25})$
$E_\ell > 1.0$ Belle [45]	$4.58(27)(^{10}_{11})$	$4.60(27)(^{10}_{11})$

Inputs for constraints from sl fit by Alberti et al, 2014 with full uncertainties and correlations

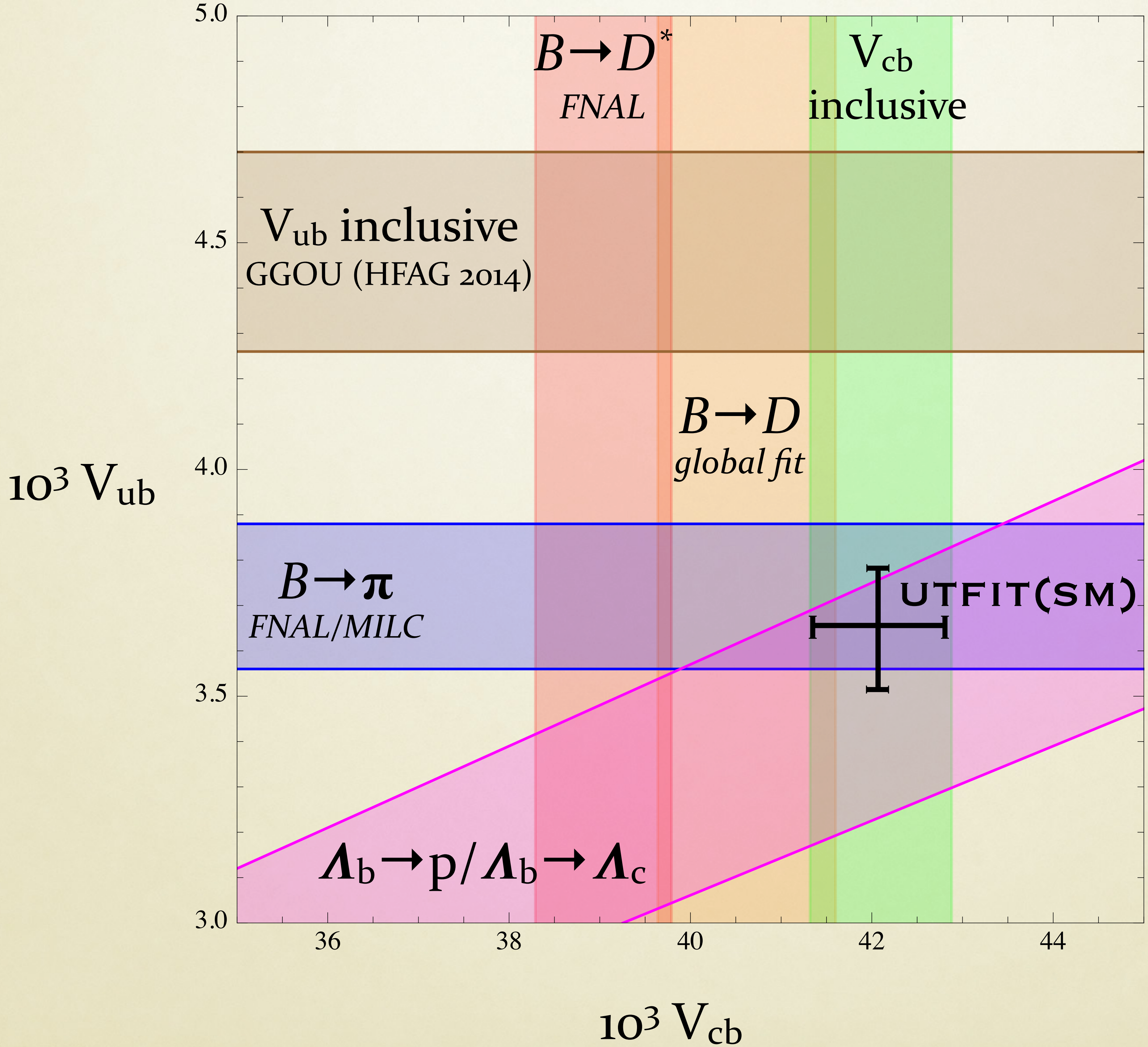
Prospects

- Learning @ Belle-II from kinematic distributions, e.g. M_X spectrum
- OPE parameters checked/ improved in $b \to u l \nu$ (moments): global NN+OPE fit
- alternative approach SIMBA
  Bernlochner, Ligeti, Stewart, F&K Tackmann
- include all relevant information with correlations
- check signal dependence at endpoint
- full phase space implementation of α_s^2 and α_s/m_b^2 corrections
- model/exclude high q^2 tail



At Belle-II we can expect to bring inclusive V_{ub} at almost the same level as V_{cb}

VISUAL SUMMARY



reasonable consistency among exclusive channels

not all results at the same level

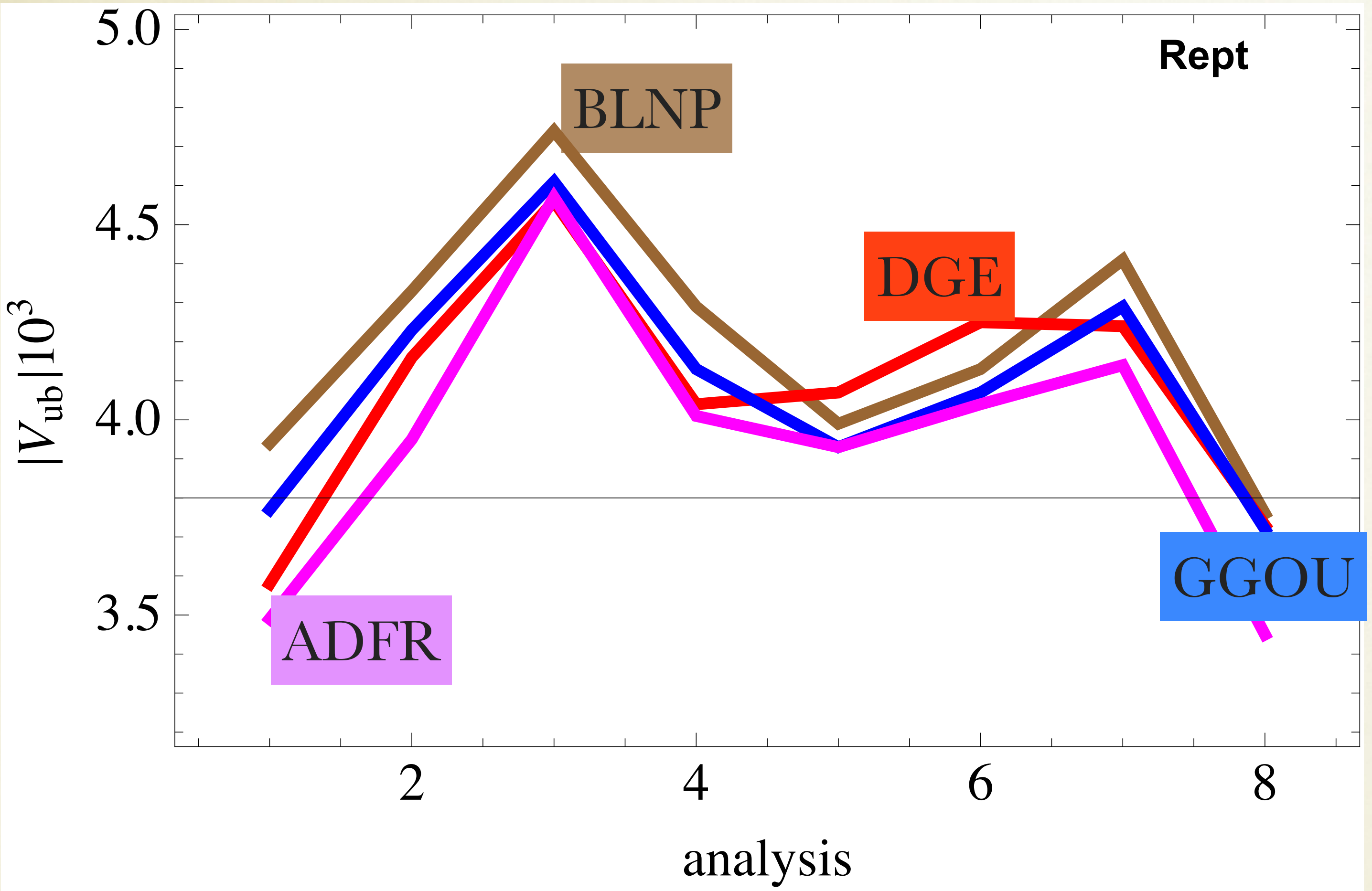
Summary

- **Exclusive/incl tension in V_{ub}** seems receding because of new lattice and endpoint Babar results. Significant progress will come with Belle-II and further LHCb data in various channels ($B\to\tau\nu$, baryons, ...).
- NNs have proven to be a **useful and flexible tool** to estimate SF uncertainties in $B\to X_u l\nu$. New constraints can be included by reweighting (instant) or retraining (slow).
- NNVub framework permits implementation of Belle-II experimental data and OPE constraints, reducing the SFs uncertainty. Comparison with data will **validate inclusive approach** to V_{ub} in a much more stringent way.

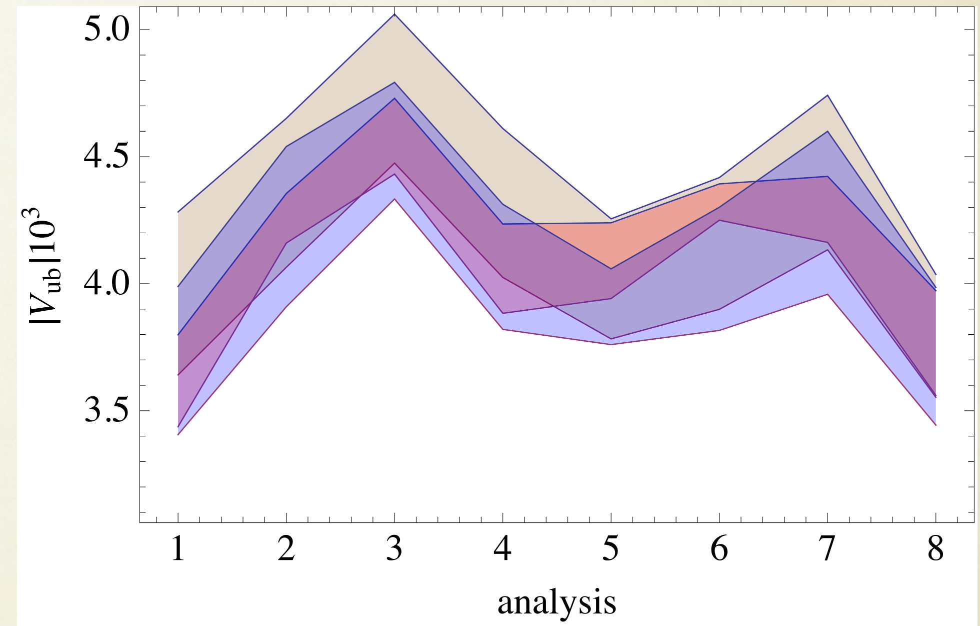
BACK-UP SLIDES

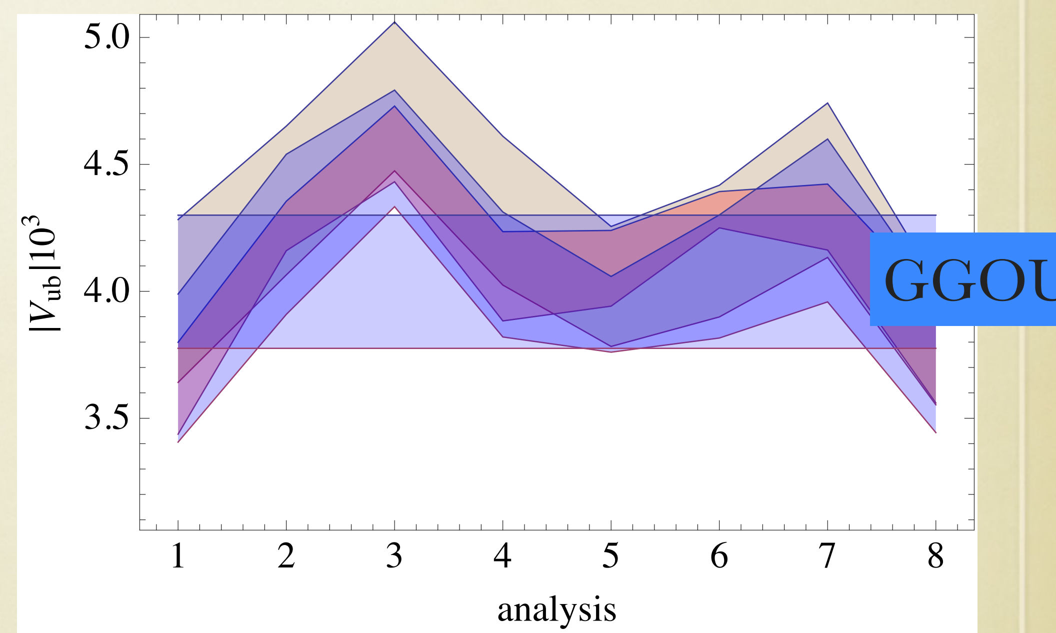
A global comparison

0907.5386, Phys Rept



only theory errors
(without common parametric)



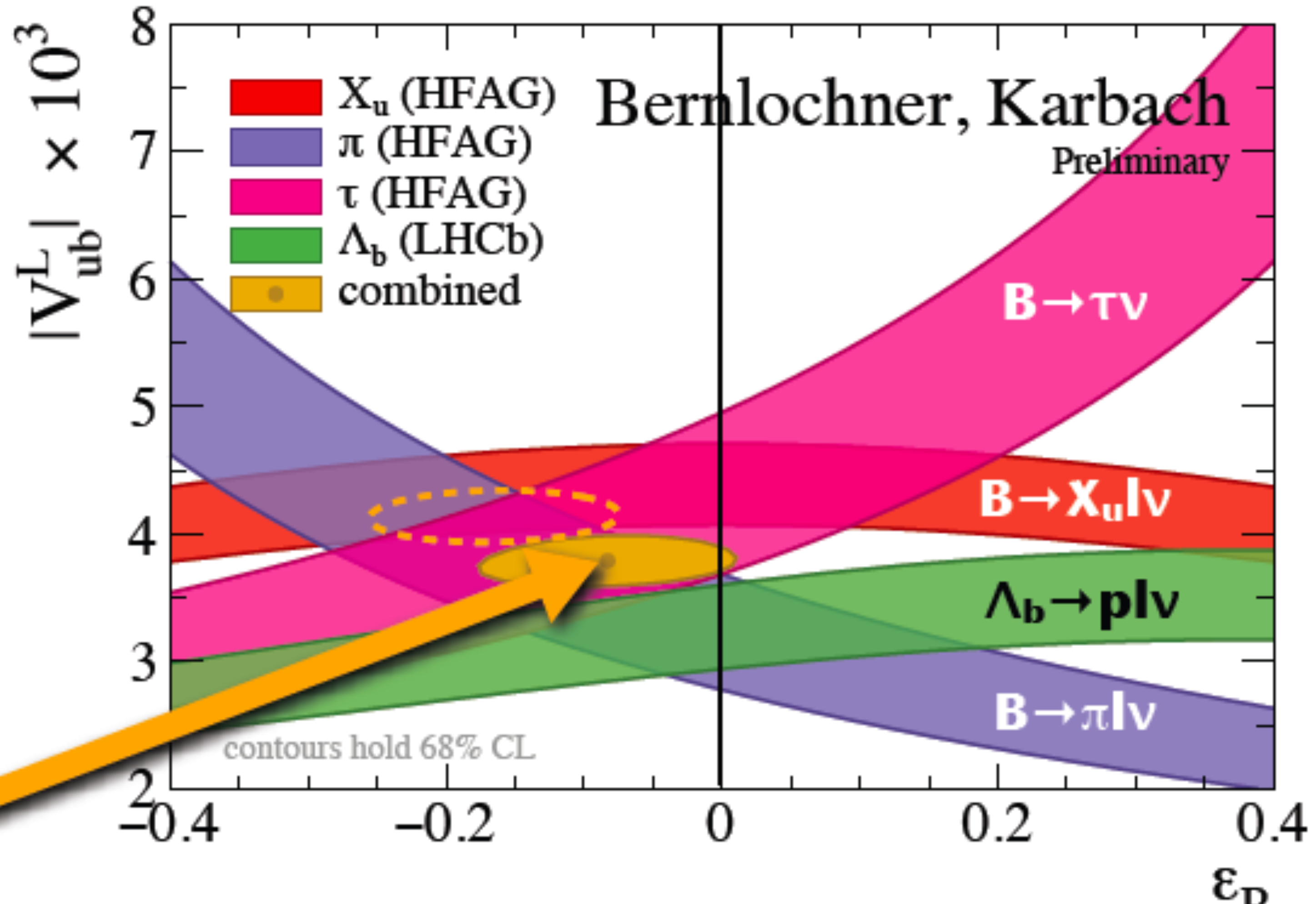


- common inputs (except ADFR)
- Overall good agreement **SPREAD WITHIN THEORY ERRORS**
- NNLO BLNP still missing: will push it up a bit
- Systematic offset of central values: normalization? to be investigated

RH currents don't help V_{ub} either

- Can ease $|V_{ub}|$ tension by allowing small right-handed contribution to Standard-Model weak current [Crivellin, PRD81 (2010) 031301]
- RH currents disfavored by Λ_b decays (taking $|V_{cb}|$ from $B \to D^* l\nu$ + HFAG to obtain $|V_{ub}|$)

p=0.03



[based on Bernlochner *et al.*, PRD 90, 094003 (2014)]

R. van de Water

Also here SU(2)xU(1) invariant NP cannot explain discrepancies 1407.1320