FB23

Neural Networks Application in Hyperon Resonance Study

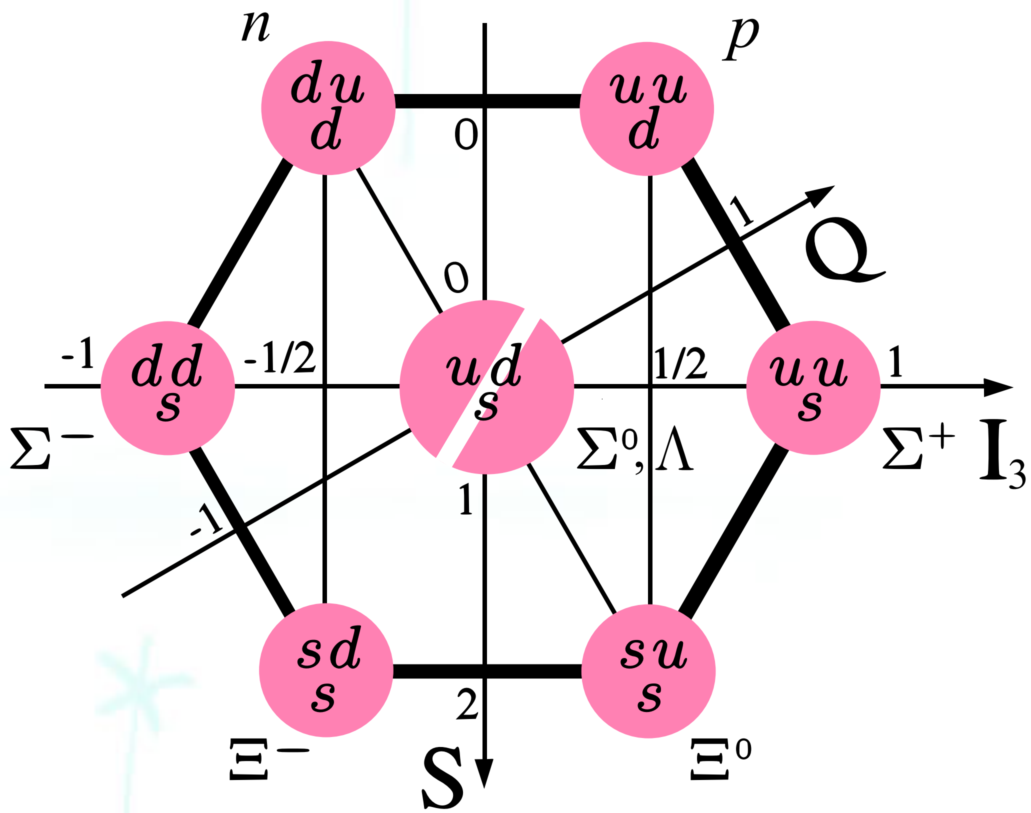
Jun Shi

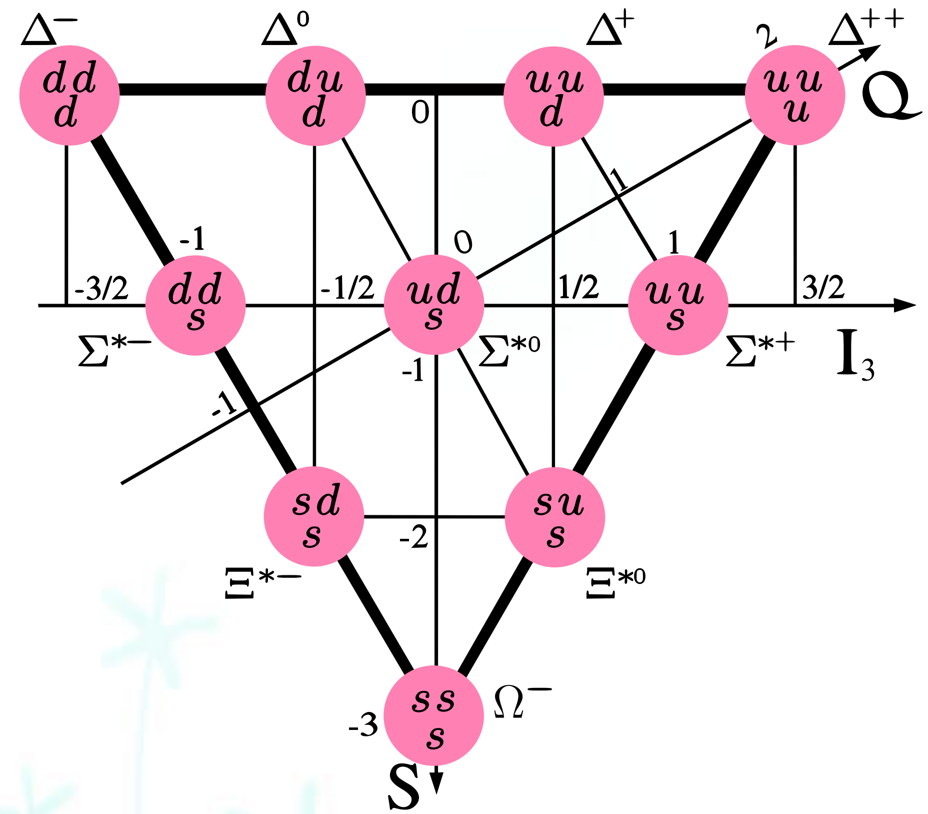
South China Normal University

(based on arXiv: 2305.01852)

September 25, 2024

Background



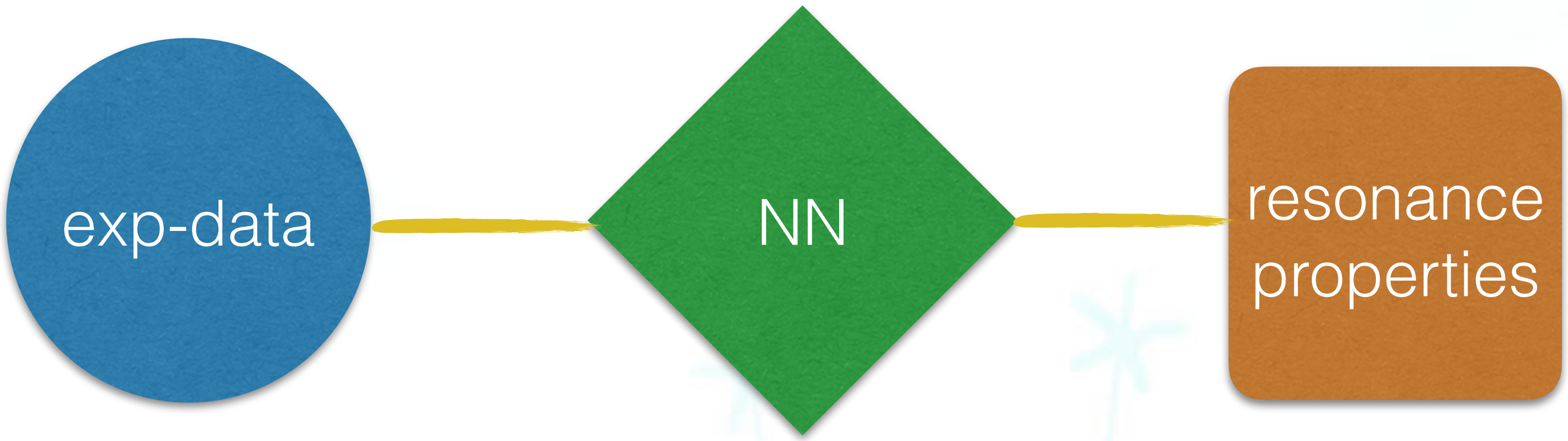


- Studying hadron spectra helps to understand their inner structure and the underlying dynamics
- Baryon resonances' properties (J^P, M, Γ, g) can be studied through scattering interactions
- Traditional analyzing method includes partial wave analysis by χ^2 fitting
- Neural Networks has been extensively applied in high energy physics

What can a NN do (in studying baryon resonances)?

In general, a NN can mimic an arbitrary function up to any required precision.

NN can be used to find out hidden relationships and correlations.



- ✓ **Classification → probability (J^P)**
- ✓ **Regression→parameters (M, Γ, g)**
- ✓ **No initial value dependence, more stable**
- ✓ **Statistical and systematic uncertainties**
- ✓ **Joint CA and RE**

NN: alternative option other than χ^2 fit!

NN application in a toy model

considering a toy model

$$A_{bg} = \frac{a + bq_{cm} + c\cos\theta + dq_{cm}\cos\theta + e\cos\theta^2 + fq_{cm}\cos\theta^2}{s + s_0}$$

$$A_{R,l} = \frac{g_R}{s - M_R^2 + iM_R\Gamma_R} P_l(\cos\theta)$$

$l = 0,1,2$

$$\frac{d\sigma}{d\cos\theta} = \frac{1}{s} |A_{bg} + \sum_i A_{Ri,l}|^2$$

$\sqrt{s} \in [1.3,\ 1.7]$ GeV
$M_R \in [1.4,\ 1.8]$ GeV
$\Gamma_R \in [0.01,\ 0.2]$ GeV

each data set contains 256 $\frac{d\sigma}{d\cos\theta}$ data

- ✓ **320000 training data sets**
- ✓ **40000 validating data sets (hyper parameter tuning)**
- ✓ **40000 testing data sets (final performance check)**

hyper parameter: learning rate, batch size, loss function weighted parameter…

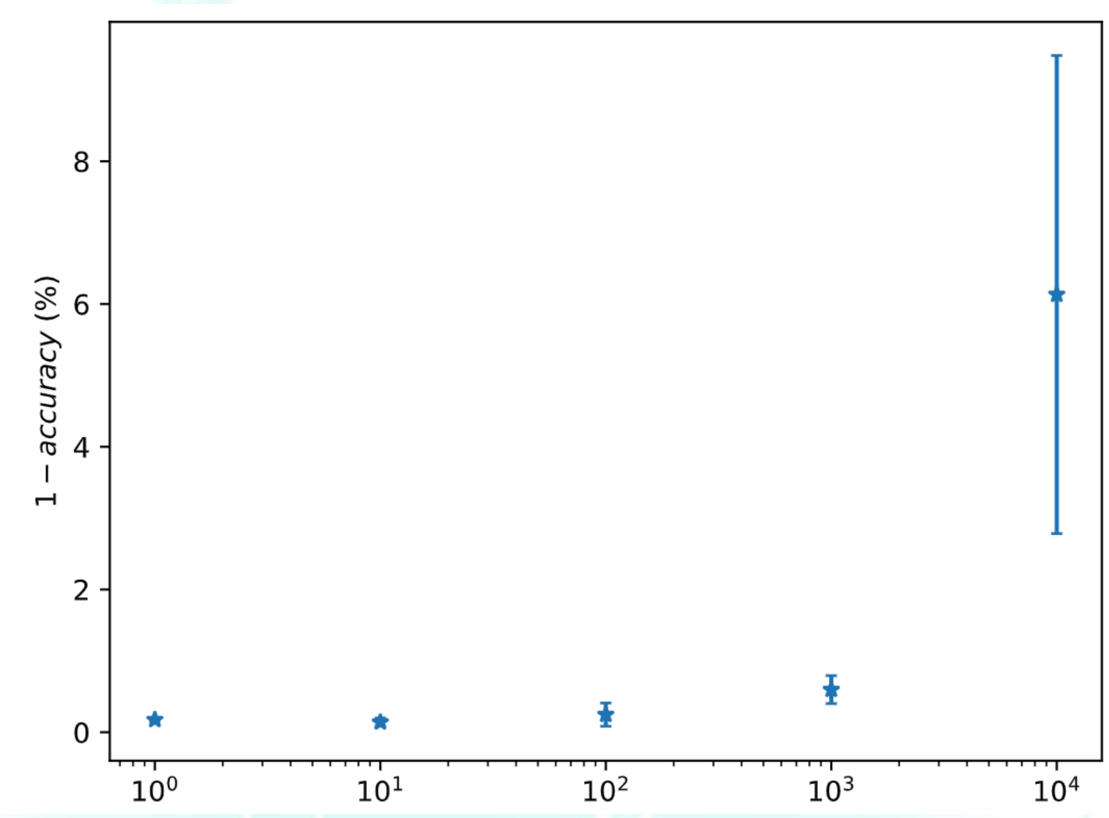
Loss function

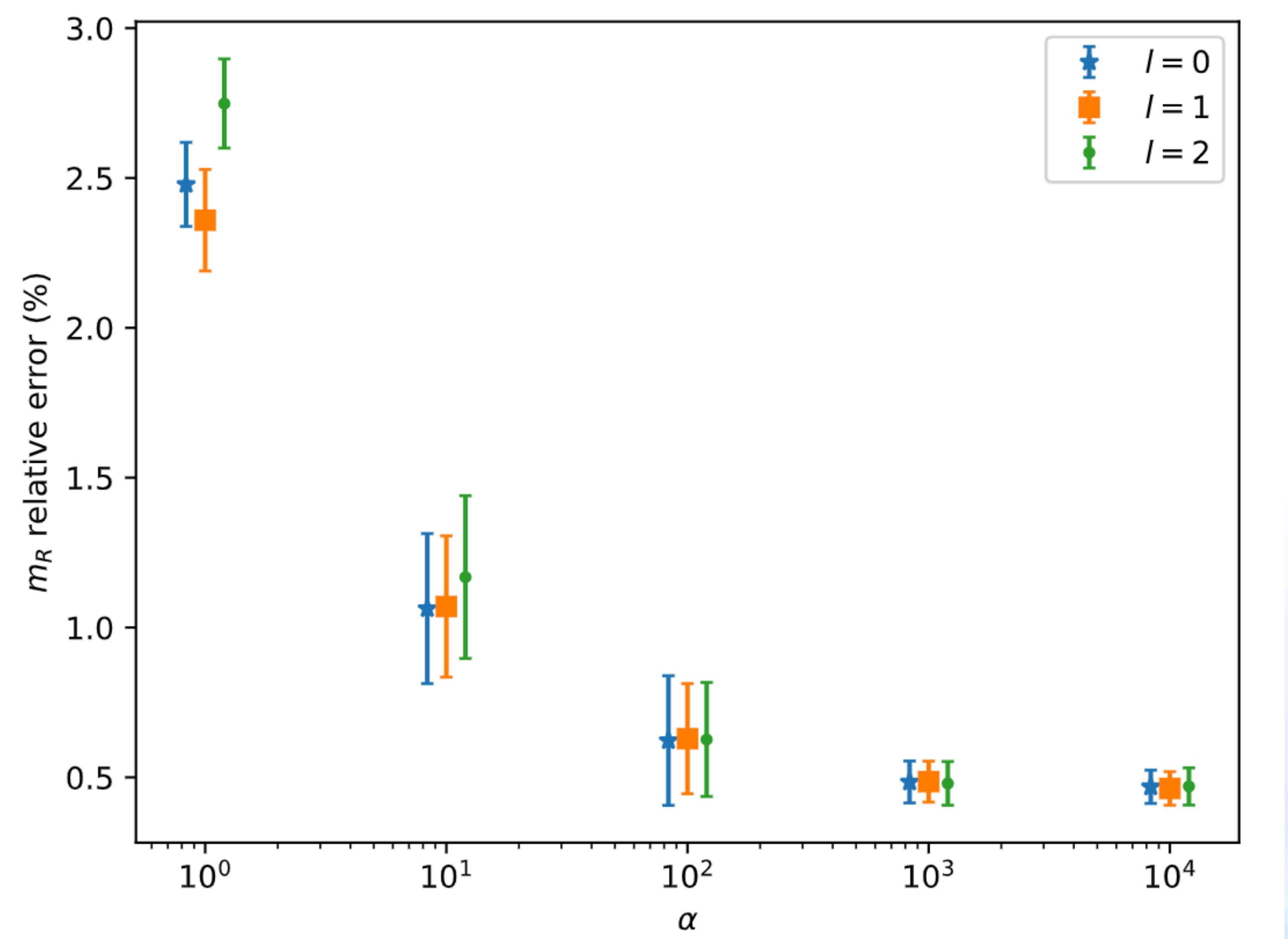
CA loss: cross entropy

RE loss: mean squared error

loss function for a joint CA and RE NN model: $\text{loss}_{CA} + \alpha \text{loss}_{RE}$

example when adding 1 additional resonance:





Procedure for more than one resonance

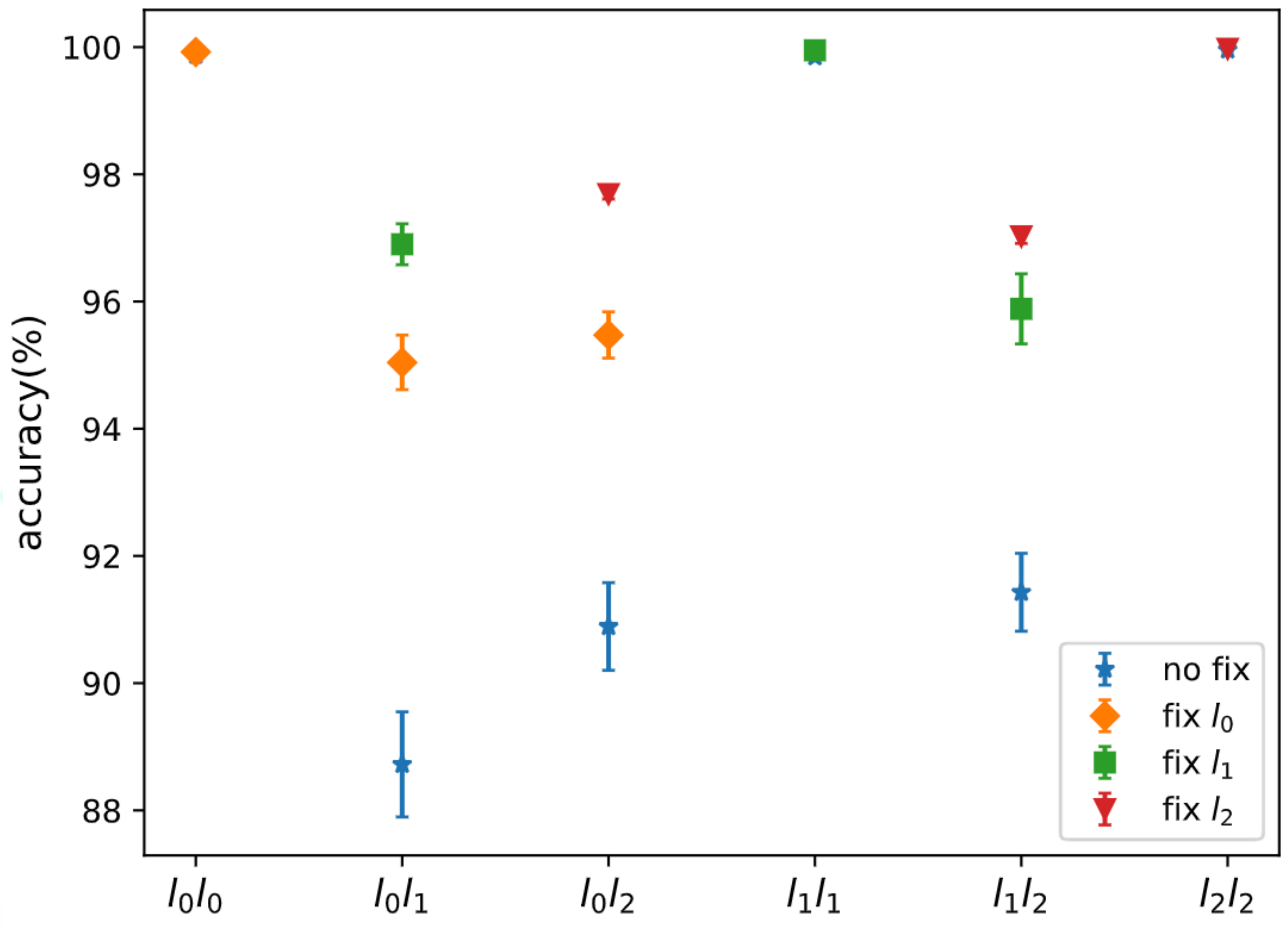
labels when adding two resonances:

all: $l_0l_0,\ l_0l_1,\ l_0l_2,\ l_1l_1,\ l_1l_1,\ l_2l_2$

$l_0X:\ \ l_0l_0,\ l_0l_1,\ l_0l_2$

$l_1X:\ \ l_0l_1,\ l_1l_1,\ l_1l_2$

$l_2X:\ \ l_0l_2,\ l_1l_2,\ l_2l_2$



- ✦ first add 1 resonance (1R)
- ✦ then 2R with one J^P fixed
- ✦ then 3R with two J^P fixed

Status of Σ resonance

$\Sigma(1385)$	$3/2^+$	****
$\Sigma(1580)$	$3/2^-$	*
$\Sigma(1620)$	$1/2^-$	*
$\Sigma(1660)$	$1/2^+$	***
$\Sigma(1670)$	$3/2^-$	****
$\Sigma(1750)$	$1/2^-$	***
$\Sigma(1775)$	$5/2^-$	****

$\Sigma(1780)$ was $\Sigma(1730)$	$3/2^+$	*
$\Sigma(1880)$	$1/2^+$	**
$\Sigma(1900)$	$1/2^-$	**
$\Sigma(1910)$ was $\Sigma(1940)$	$3/2^-$	***
$\Sigma(1915)$	$5/2^+$	****
$\Sigma(1940)$	$3/2^+$	*
$\Sigma(2010)$ was $\Sigma(2000)$	$3/2^-$	*

PDG

mass of lowest Σ $1/2^-$:

✓ classical quark model: ~1650 MeV

✓ pentaquark model: ~1400 MeV

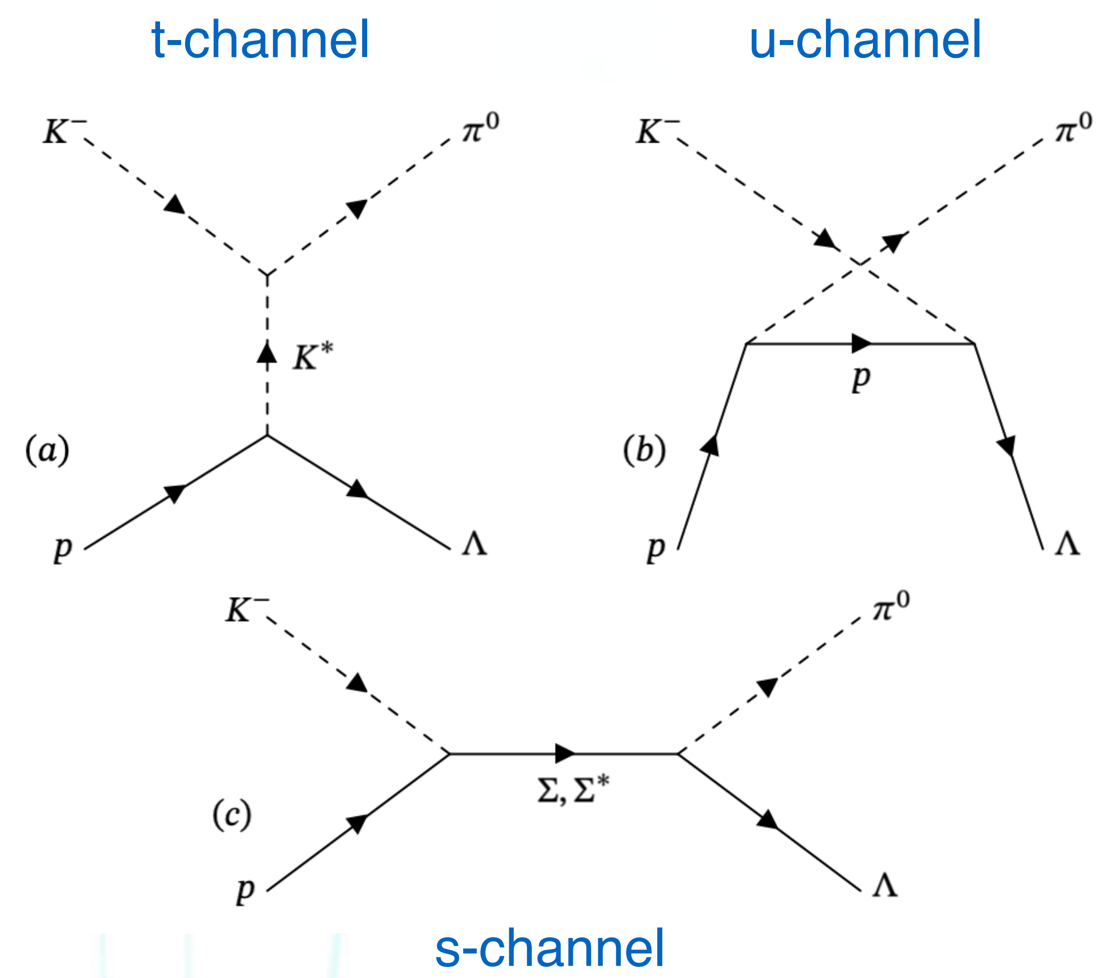
Zou, NPA 914(2013)

Σ resonances from $\bar{K}N$ interaction

$I(\bar{K}) = 1/2, I(N) = 1/2 \Rightarrow I_s = 0,\ 1$

$I(\pi) = 1, I(\pi) = 1, I(\Lambda) = 0 \Rightarrow I_s = 1$

$\Rightarrow$ s-channel only $\Sigma,\ \Sigma^*$



$\bar{K}N \rightarrow \pi\Lambda$ ideal reaction to study Σ resonances

Single channel analysis for comparison and feasibility check

χ^2 fitting-based PWA

Try ONE additional Σ resonance in the initial fits, pick the one with the smallest χ^2 as the most probable one. Get its parameters in the same fit.

Add another Σ resonance in the following fits while keeping the first picked one. Find out the second most probable resonance by the χ^2 criterion and determine the parameters.

Gao, Zou, Sibirtsev, Nucl.Phy.A 867(2011),
Gao, **Shi**, Zou, Phys. Rev. C 86(2012)

χ^2's of different candidates can be close.

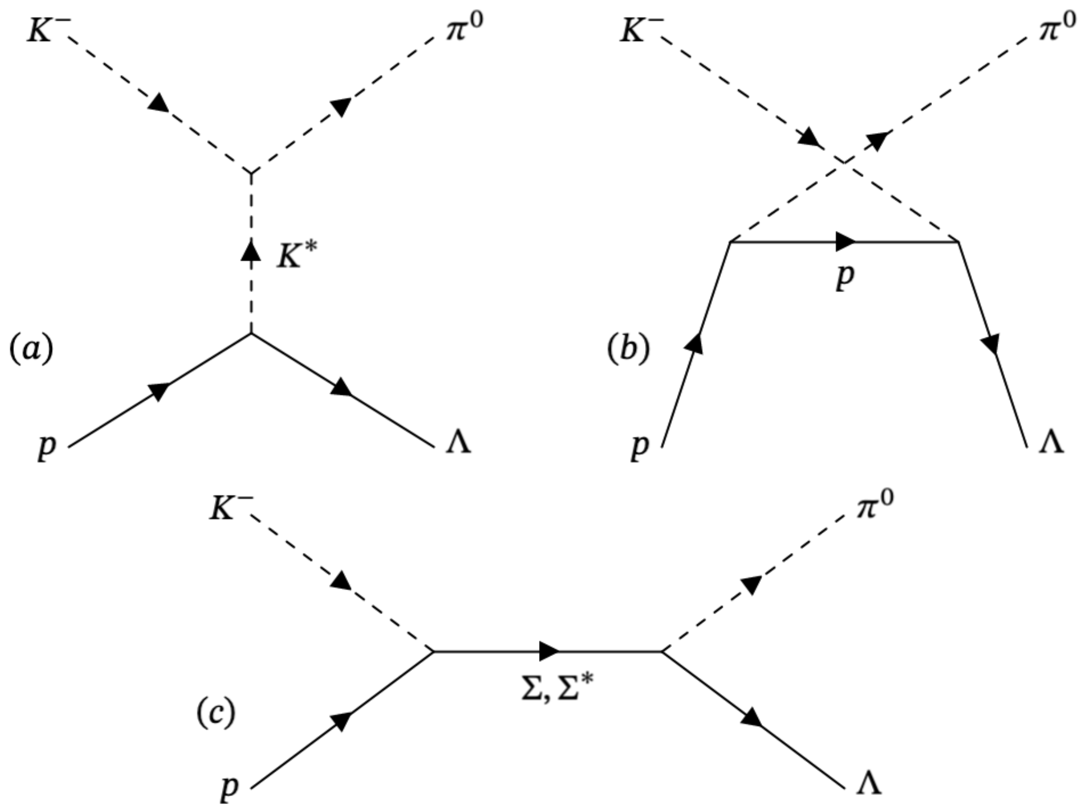
Probability?

From 2nd resonance, some of the errors are over 100% and the fit of some parameters reach boundaries.

Initial value dependence (systematic error) not taking into account

Theoretical formalism

t-channel u-channel



s-channel

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 s}\frac{\mathbf{q}}{\mathbf{k}}\overline{|\mathscr{M}|^2}$$

$$P_\Lambda = 2\mathrm{Im}\left(\mathscr{M}_{\frac{1}{2}\frac{1}{2}}\mathscr{M}^*_{\frac{1}{2}\frac{-1}{2}}\right)/\overline{|\mathscr{M}|^2}$$

Effective Lagrangian

$$\mathcal{L}_{KN\Sigma(\frac{1}{2}^+)} = \frac{g_{KN\Sigma}}{M_N + M_\Sigma}\partial_\mu\overline{K\Sigma}\cdot\tau\gamma^\mu\gamma_5 N + \mathrm{H.c.},$$

$$\mathcal{L}_{\Sigma(\frac{1}{2}^+)\Lambda\pi} = \frac{g_{\Sigma\Lambda\pi}}{M_\Lambda + M_\Sigma}\overline{\Lambda}\gamma^\mu\gamma_5\partial_\mu\pi\cdot\Sigma + \mathrm{H.c.},$$

$$\mathcal{L}_{KN\Sigma(\frac{3}{2}^+)} = \frac{f_{KN\Sigma}}{m_K}\partial_\mu\overline{K\Sigma}^\mu\cdot\tau N + \mathrm{H.c.},$$

$$\mathcal{L}_{\Sigma(\frac{3}{2}^+)\Lambda\pi} = \frac{f_{\Sigma\Lambda\pi}}{m_\pi}\partial_\mu\overline{\pi}\cdot\overline{\Sigma}^\mu\Lambda + \mathrm{H.c.},$$

Using Crystal Ball 2009 data

S. Prakhov et.al., Phys. Rev. C 80(2009)

Data generation

In total, $3 \times 4 \times 2.56\ \mathrm{M} = 30.72$ **million data sets.**

$g_{K^*N\Lambda}$	$g_{K^*N\Lambda}\kappa_{K^*N\Lambda}$	$g_{\pi NN}g_{KN\Lambda}$	$g_{KN\Sigma}g_{\Sigma\Lambda\pi}$
$(6.11, 4.26)$	$(16.3, 10.4)$	$(1/\sqrt{2}, \sqrt{2}) \times (-176)(\mathrm{SU3})$	$(1/\sqrt{2}, \sqrt{2}) \times 34.8(SU3)$
$f_{KN\Sigma^*}f_{\Sigma^*\Lambda\pi}$	$M_{\Sigma(1670)}$	$\Gamma_{\Sigma(1670)}$	$\Gamma_{\overline{K}N}\Gamma_{\pi\Lambda}/\Gamma_{\Sigma(1670)}$
-3.1(SU3)	$(1.665, 1.685)$	$(0.04,\ 0.1)$	$(0.018,\ 0.17)$

1R, 2R and 3R cases

$\Sigma\ 1/2^{\pm},\ 3/2^{\pm}$

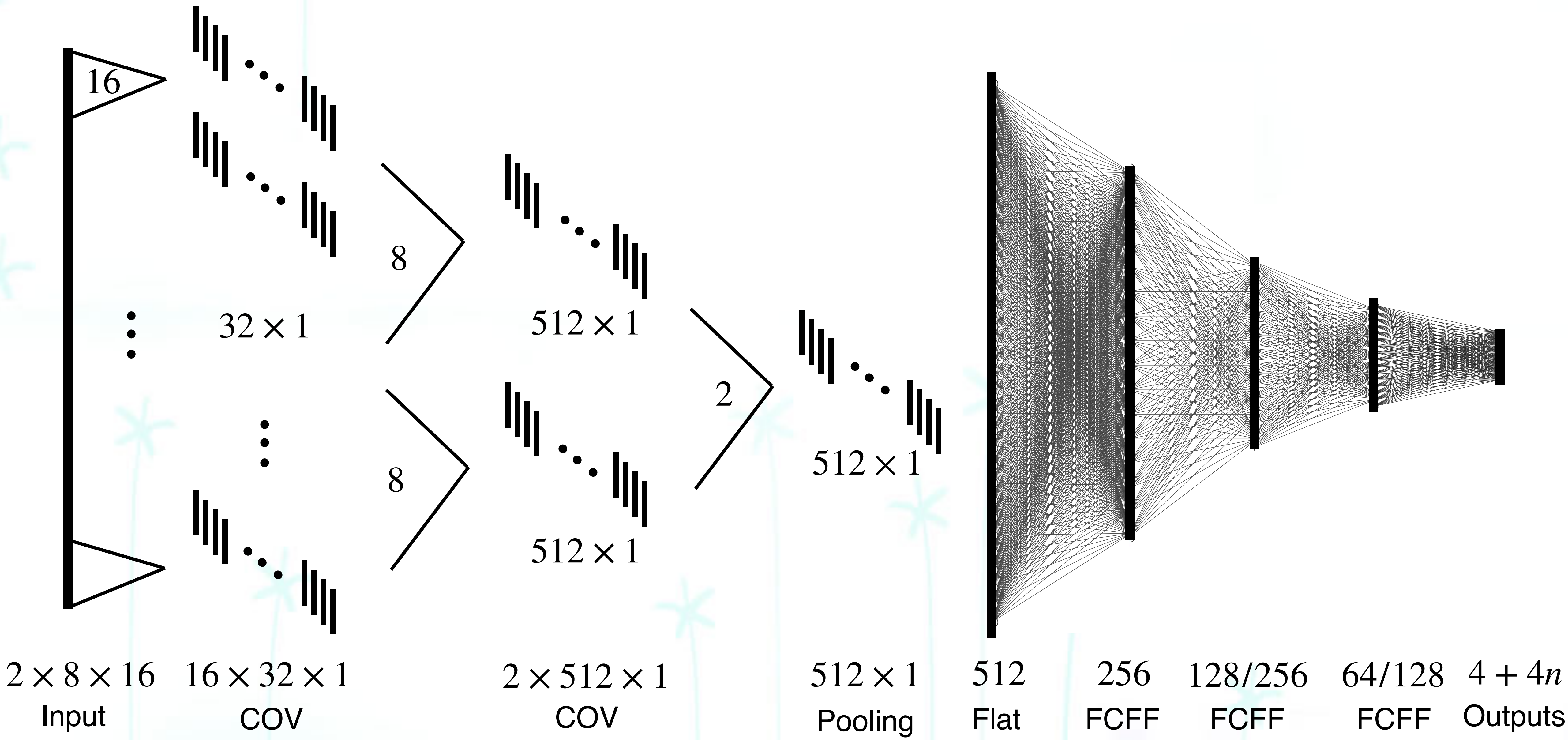
Random sampling in the allowed parameter space

Background: t-channel K^*, u-channel p, s-channel $\Sigma(1189)1/2^+$, $\Sigma(1385)3/2^+$, $\Sigma(1670)3/2^-$, $\Sigma(1775)5/2^-$

additional resonance(s):
$M \in [1.44,\ 1.9]$ GeV,
$\Gamma \in [0.01,\ 0.4]$ GeV
$g \in [-10,\ 10]$

- ✓ **10.00 M training data sets**
- ✓ **0.12 M validating data sets (hyper parameter tuning)**
- ✓ **0.12 M testing data sets (final performance check)**

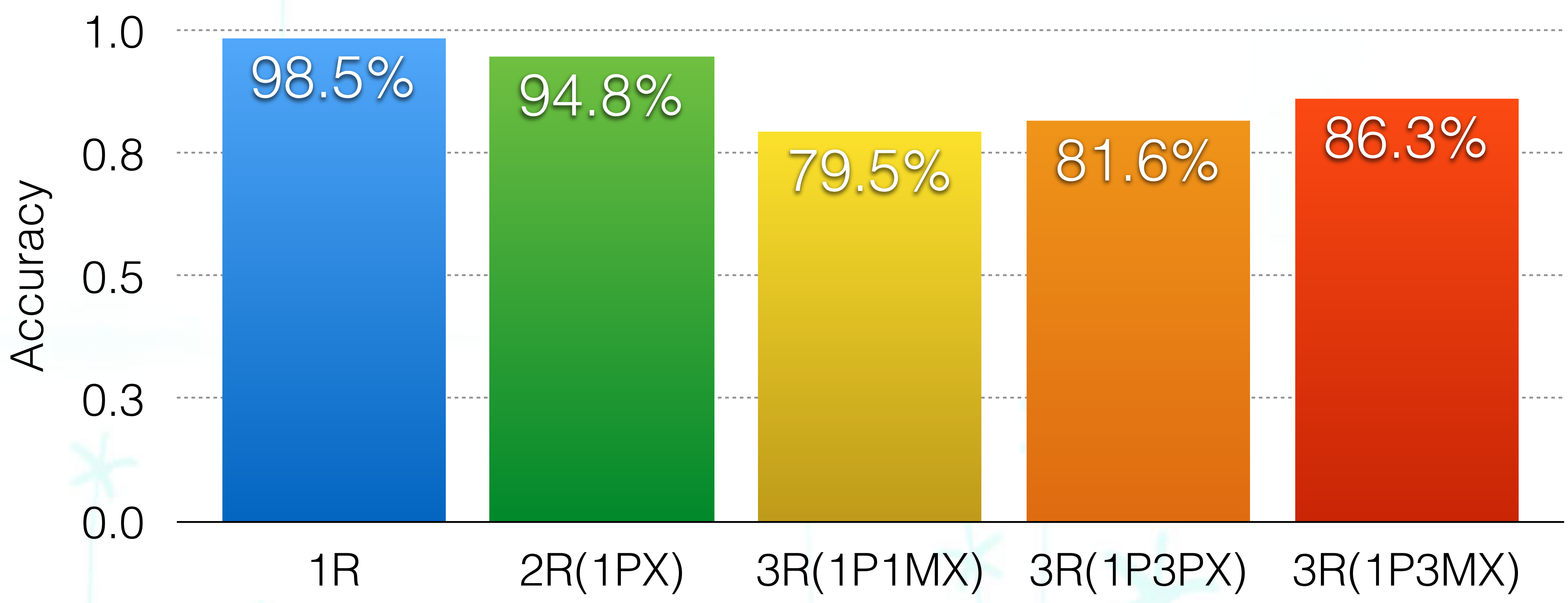
Data structure inspired joint NN model

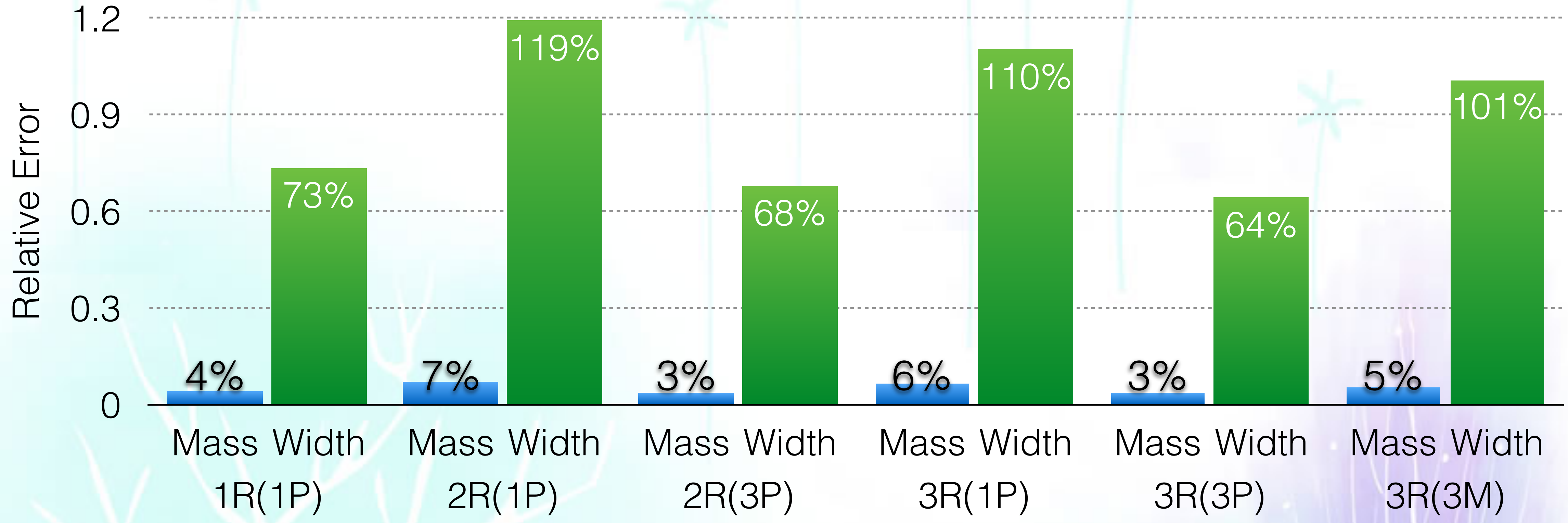


loss$_{CA}$ + α**loss**$_{RE}$ **with** $\alpha = 0.2$

Artificially expanded training data sets in Gaussian distribution (AED) for error tolerance

NN performance (on test sets)





Sophisticated error handling

20 models trained independently with different initial values in each case!

E1 = 1 - Accuracy on test sets (CA)

E1 = relative error on test sets (RE)

CA: probabilities of Σ^* with J^P

RE: properties of Σ^*

E2 = standard error of 20 trials

Systematic errors

E3 = standard deviation of 4000 sets of mock experimental data sampled assuming normal distribution

Statistical error

Quadratic sum for total error

Results for 1R case

CA

accuracy: 98.5

$1/2^+$	$1/2^-$	$3/2^+$	$3/2^-$
100.0(3.0)	0.0(0.6)	0.0(0.8)	0.0(1.6)

RE

	m/GeV	Γ/GeV
$\frac{1}{2}^+(1\mathrm{R},\ \frac{1}{2}^+)$	1.583(0.066)	0.155(0.115)

PDG $\Sigma(1660)1/2^+$: $M \in [1.64,\ 1.68]$ GeV, $\Gamma \in [0.1,\ 0.3]$ GeV

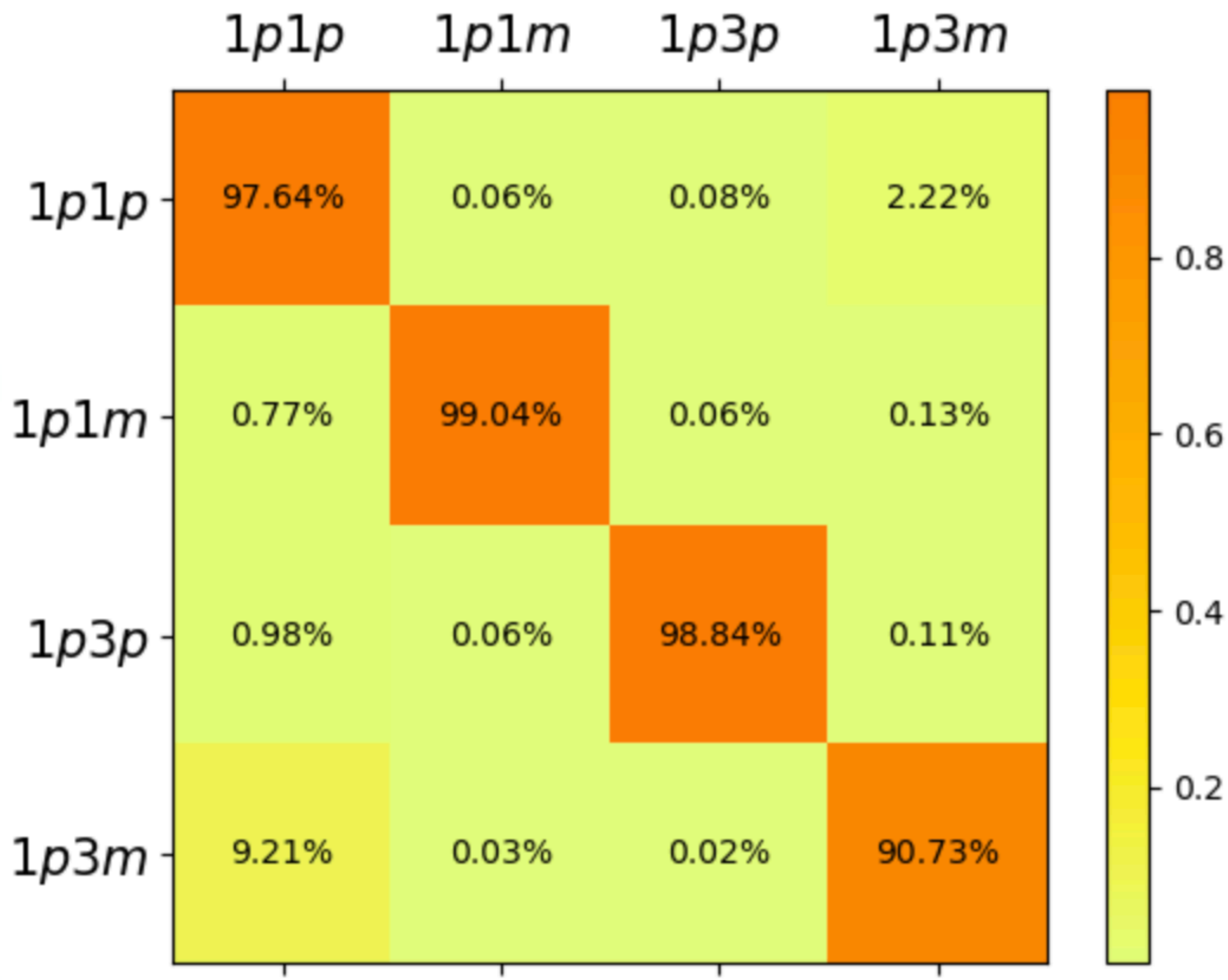
χ^2 fitting-based PWA $\Sigma\ 1/2^+$: $M = 1.633 \pm 3$ GeV, $\Gamma = 0.121^{+4}_{-7}$ GeV

Gao, Zou, Sibirtsev, Nucl.Phy.A 867(2011),
Gao, **Shi**, Zou, Phys. Rev. C 86(2012)

CA results for 2R

2R

accuracy: 94.8



$1/2^+1/2^+$	$1/2^+1/2^-$	$1/2^+3/2^+$	$1/2^+3/2^-$
0.0(3.4)	15.5(27.6)	72.2(23.6)	12.3(29.8)

CA results for 3R

1p1mX	$1/2^+1/2^-1/2^+$	$1/2^+1/2^-1/2^-$	$1/2^+1/2^-3/2^+$	$1/2^+1/2^-3/2^-$
acc: 79.5	0.0(15.8)	0.0(29.5)	0.0(5.7)	100.0(30.9)

1p3pX	$1/2^+3/2^+1/2^+$	$1/2^+3/2^+1/2^-$	$1/2^+3/2^+3/2^+$	$1/2^+3/2^+3/2^-$
acc: 81.6	1.6(17.2)	1.3(9.7)	0.3(28.3)	96.8(32.2)

1p3mX	$1/2^+3/2^-1/2^+$	$1/2^+3/2^-1/2^-$	$1/2^+3/2^-3/2^+$	$1/2^+3/2^-3/2^-$
acc: 86.3	0.3(27.9)	98.9(10.1)	0.6(5.7)	0.3(23.4)

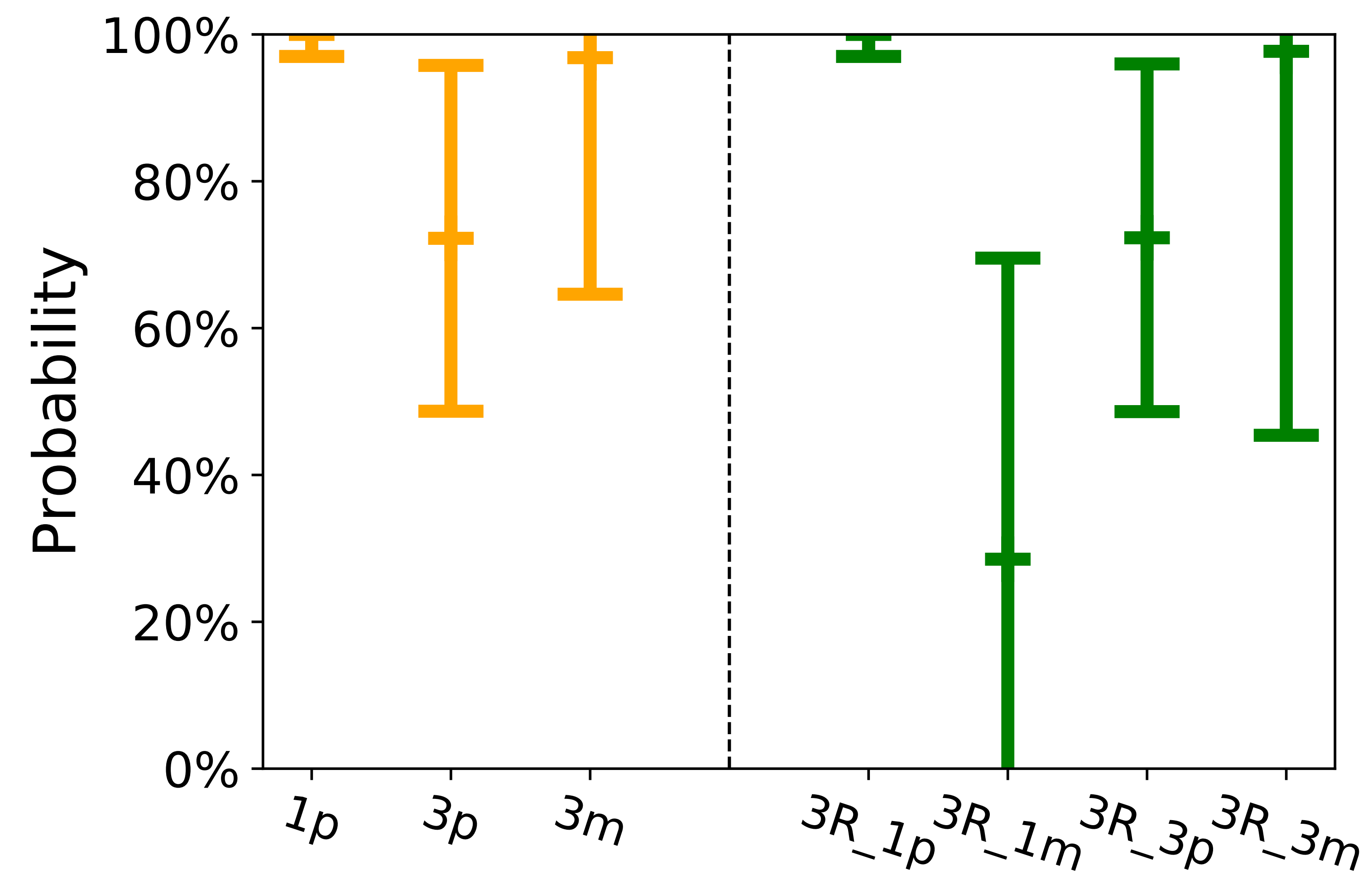
probability propagation: 1p1m3m 27.62(40.68)%, 1p3p3m 70.02(32.59)%

Final CA results

- **1R:** 1p $(1/2^+)$ **100(3)%**
- **1pX:** 3p $(3/2^+)$ **72.2(23.7)%**
- **1p3pX:** 3m $(3/2^-)$ **96.8(32.2)%**

Assuming 3 resonances are needed:

- 1p $(1/2^+)$ **: 100(3)%**
- 1m $(1/2^-)$ **: 28.5(41)%**
- 3p $(3/2^+)$ **: 72.3(23.7)%**
- 3m $(3/2^-)$**: 97.7(52.3)%**



Final predictions of Σ resonance parameters

Final prediction is from a weighted average of different combinations in 3R by probability

$\frac{1}{2}^+$	1.620(105)	0.173(192)
$\frac{1}{2}^-$	1.713(60)	0.183(123)
$\frac{3}{2}^+$	1.716(60)	0.245(159)
$\frac{3}{2}^-$	1.607(90)	0.201(205)

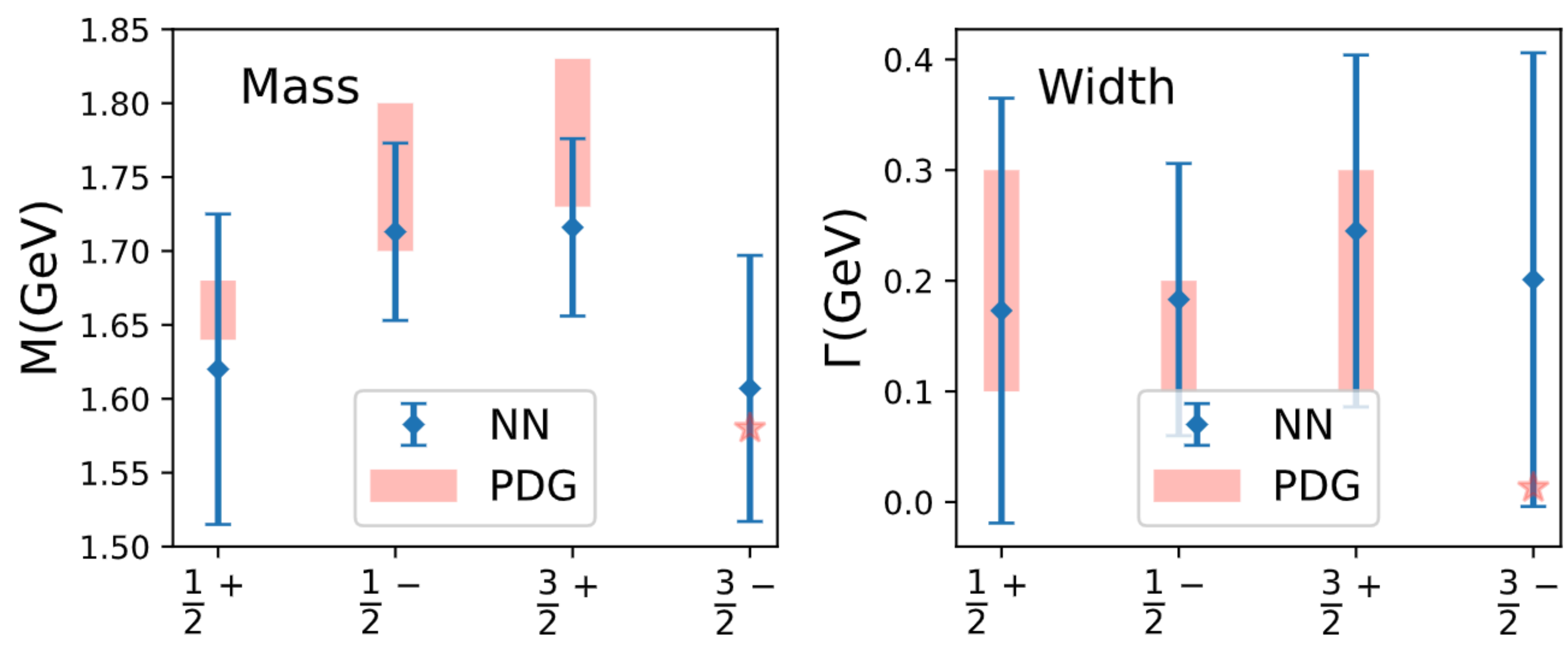
PDG:

*** $\Sigma(1660)1/2^+$

*** $\Sigma(1750)1/2^-$

* $\Sigma(1780)3/2^+$

* $\Sigma(1580)3/2^-$



Error source analysis

E1 = 1 - Accuracy on test sets (CA)

E1 = relative error on test sets (RE)

E2 = standard error of 20 trials

E3 = standard deviation of 4000 sets of mock experimental data sampled assuming normal distribution

CA of $1/2^+3/2^+$ in 2R:
72(2)(10)(21)

RE(mass) of $1/2^+3/2^+$ in 2R:
$1/2^+$:1.636(0.111)(0.004)(0.008)
$3/2^+$:1.705(0.059)(0.012)(0.022)

CA: E3 the biggest
1/2~1/3 uncertainties of exp to determine the most significant second resonance

RE: E1 the biggest
possibly larger training sets, more powerful NN or multi-channel PWA may improve

} better determine the lowest $1/2^-$

Summary and outlook

- ✓ We constructed a CA+RE joint NN to do PWA to study Σ resonances
- ✓ The NN works very well
- ✓ Comparing with χ^2 fitting-based PWA, NN can give probability, independent of initial values and more stable, signals for the second and third resonance are better
- ✓ Sophisticated probability propagation and error analysis
- ✓ NN is potentially another powerful tool in studying baryon resonances
- ✓ Further applications of NN-based PWA is expected with future experimental data and more sophisticated theoretical formulae

Thank you!