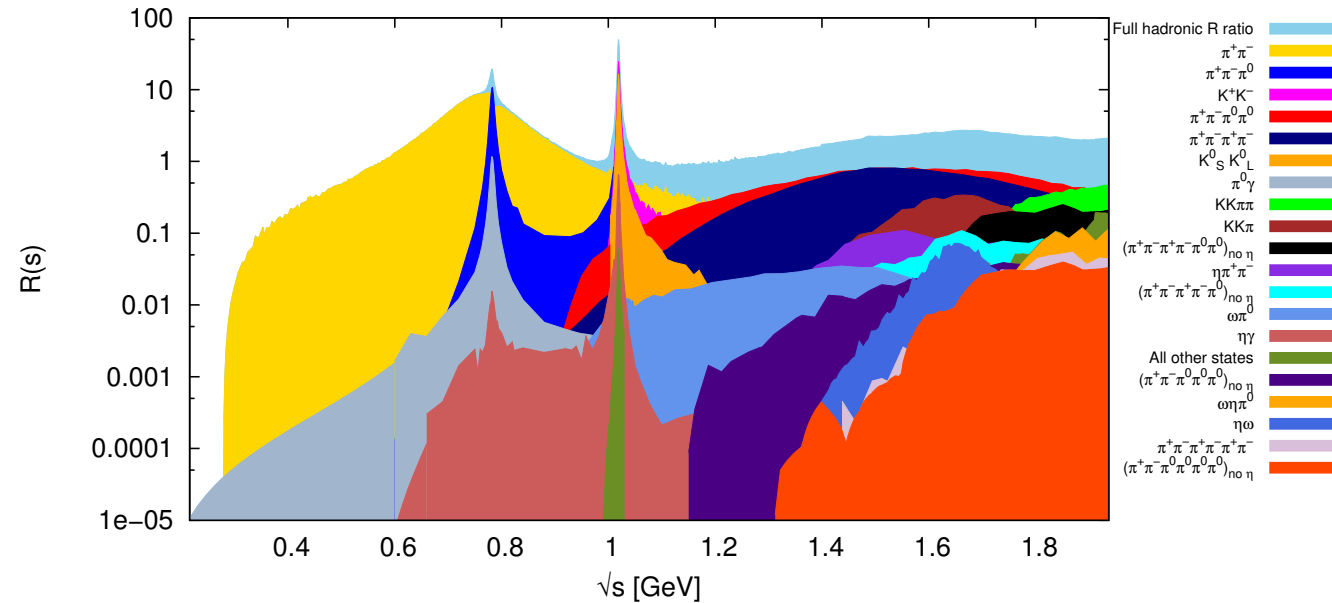


KNT Data-Driven HVP: Update

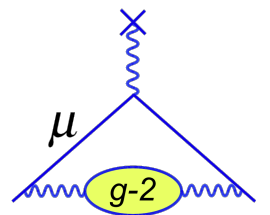


Alex Keshavarzi
University of Manchester

With Daisuke Nomura and Thomas Teubner

Muon g-2 Theory Initiative Workshop
University of Edinburgh

6th September 2022

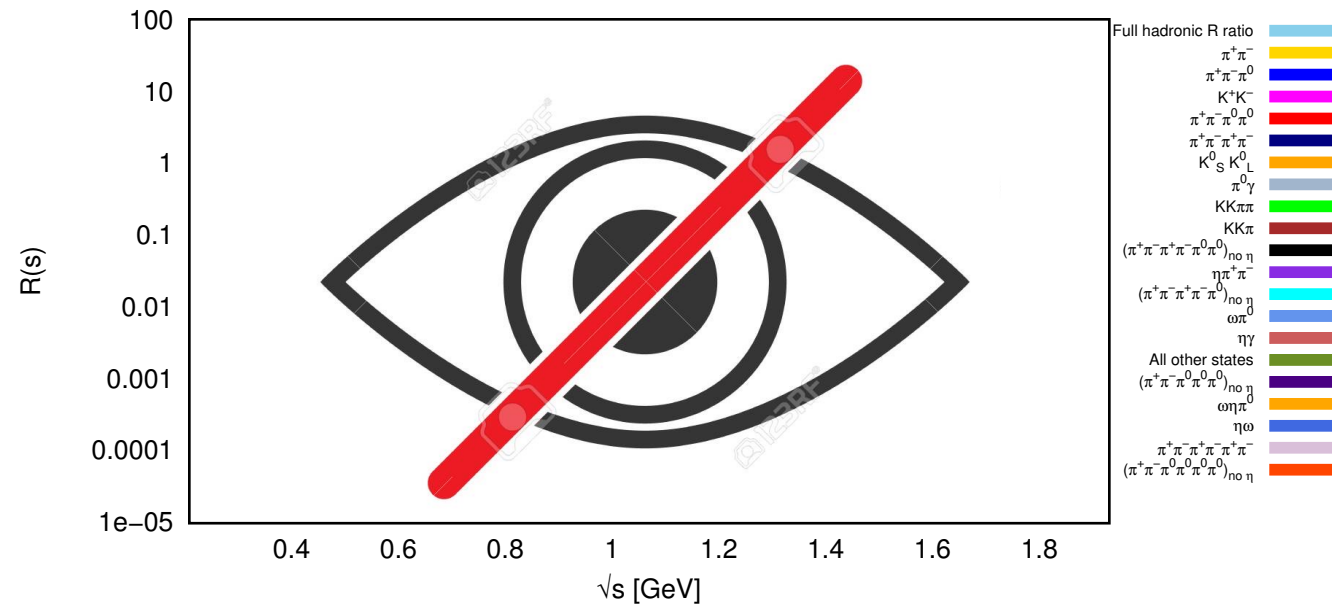


The update is...

The update is...

**... that we have exactly zero new results
to show you.**

KNT Data-Driven HVP: Blind Analysis

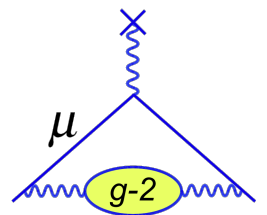


Alex Keshavarzi
University of Manchester

With Daisuke Nomura and Thomas Teubner

Muon g-2 Theory Initiative Workshop
University of Edinburgh

6th September 2022



Why do we want to blind our analysis?

- We want to minimise conscious and unconscious bias as the stakes are high:
 - Data-driven HVP is a sub-percent precision game.
 - New data could resolve current data tensions (e.g. BaBar vs KLOE in $\pi^+\pi^-$).
 - Everything now rests on lattice vs. data-driven HVP.
 - Current discrepancies (SM vs. FNAL, lattice vs. dispersive) = new physics potential.
 - FNAL Experiment and Lattice HVP are now blinded. Dispersive should be too.
- It will better the KNT/dispersive analysis:
 - Force us to scrutinise our analyses further and in different ways.
 - Additional reassurance to the community that our analysis is robust.
- It will stop a_μ being the primary observable / figure of merit:
 - Will place other observables (e.g. $\Delta\alpha_{\text{had}}$, Euclidean windows) on a more even footing.
- It will introduce external oversight:
 - Have an external / non-KNT person set the blinding offsets.
 - That external will be a party to the unblinding.
- Now is a good time:
 - We're about to start a new analysis.
 - No new data since KNT19 has been included in any final capacity/configuration. Can start from a clean slate.

Blinding aims 1

The blinding procedure/implementation should:

- Blind the results of our analysis in a meaningful and robust procedure that **removes our ability to know the results of data combinations that could lead us to bias**.
- **Not introduce new biases** to those results.
- **Preserve the combination methodology** as it currently exists, e.g. clustering & scans, fitting, the VP iteration.
- Allow us **to optimise our data combinations based on figures of merit** other than integrated mean values, e.g. error, χ^2 ...
- Have **more than one layer of blinding**, allowing for a relative unblinding and cross checks before a full unblinding.
 - **Distinctly blind each hadronic channel** for which we perform a combination.
 - This will provide an **extra layer of blinding** where the individual channels won't move relative to each other.
 - **Will allow for a relative unblinding** stage where, once analyses on individual channels are complete, individual channel offsets are removed and instead we have a single set of offsets for all channels.
- Apply **different blinding to integrated results and plots** containing the same data to minimise bias.
- **Blind both integrals and plots in an energy-dependent way**, with energy-dependent adjustments that aren't too extreme.

Blinding aims 2

The blinding procedure/implementation should:

- Be as water-tight as possible, with ideally **no way for us to disentangle what the introduced blinding offsets are.**
- Be **introduced and removed by an external person** outside of the KNT collaboration.
- **Retain our ability to know the results of individual data sets:**
 - Impossible for us to retrospectively blind the experimental data. Their analysis is complete, is not in our control, and **adjusting the data on input could lead to biases in the analysis.**
 - **New data will already be integrated over by the experiments** in their publications.
 - The **outputs of the KNT analysis are what we want to blind.**
 - It has proved extremely important in the past for the community to **be able to cross check new data** to ensure there are no issues.
- Be universal for any analyser (e.g. **not machine dependent**).
- **Be as simple as possible** for both internal and external contributors, whilst ensuring all the other aims that are listed above.
- Be **checked and ratified by someone / people with experience in blinding analyses.**

A good faith acknowledgment

The implementation of blinding in our analysis must come with an element of trust:

- We have **taken every precaution** to ensure that the blinding we have implemented is fool-proof.
- It has been **stress-tested and checked by external (non-KNT) parties**.
 - E.g. Mark Lancaster (blinding for the FNAL Muon g-2 Experiment).
- We are not a large collaboration:
 - **No analysis hierarchy in which the blinding can be embedded.**
 - **No additional blinding to be imposed externally** (e.g. FNAL Muon g-2 Clock Blinding).
- We **acknowledge** that, for example, there is nothing stopping us from simply not including the blinding software:
 - We hope you trust in us that we are implementing this because **we believe it to be the best scientific practice** and so will **not undermine its integrity**.



The blinding procedure

- **Blind any quantity that we “see”.** We want to remove our knowledge of:
 - Changes in direction of results: overall sign, \pm .
 - Changes in the size results: multiplicative scale factor.
 - Energy-dependent changes to results: s-dependent adjustment.
- The blinded quantities that we “see” are:
 - **Integrated values.** Offsets to blind integrated values in an energy-dependent way.
 - **Plots.** Different offsets to blind plots in an energy-dependent way.
- **No data that enter analyses are changed/adjusted/blinded**, but are not viewable.
 - Ensures that **we meet our aims** and analysis will not be impacted, e.g. VP iteration calculation.
- **Blinding offsets will be different for each channel**, until the relative unblinding imposes the same offsets to all.

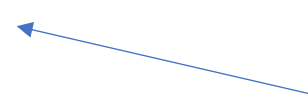
Blinding Adjustment	Integrals	Plots
Overall sign, \pm	✓	✗
Multiplicative scale factor	✓	✓
s-dependent adjustment	✓	✓

The blinding procedure

Integral: $I^{\text{blind}} = a b \int ds (s + s_0)^c f(s) \sigma_{\text{had}}(s) \quad ; \quad f(s) = \text{integration kernel, e.g. } K(s)/(4\pi^3)$

Plotted cross section: $\sigma_{\text{had}}^{\text{blind}}(s) = \tilde{b} (s + \tilde{s}_0)^{\tilde{c}} \sigma_{\text{had}}(s) \quad ; \quad \tilde{b} \neq b, \tilde{c} \neq c, \tilde{s}_0 \neq s_0$

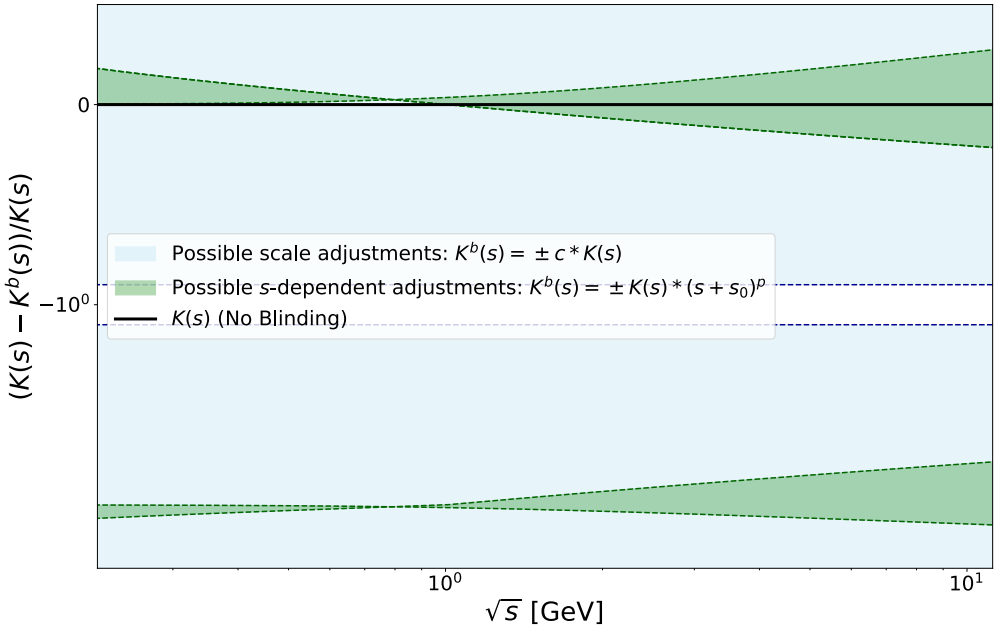
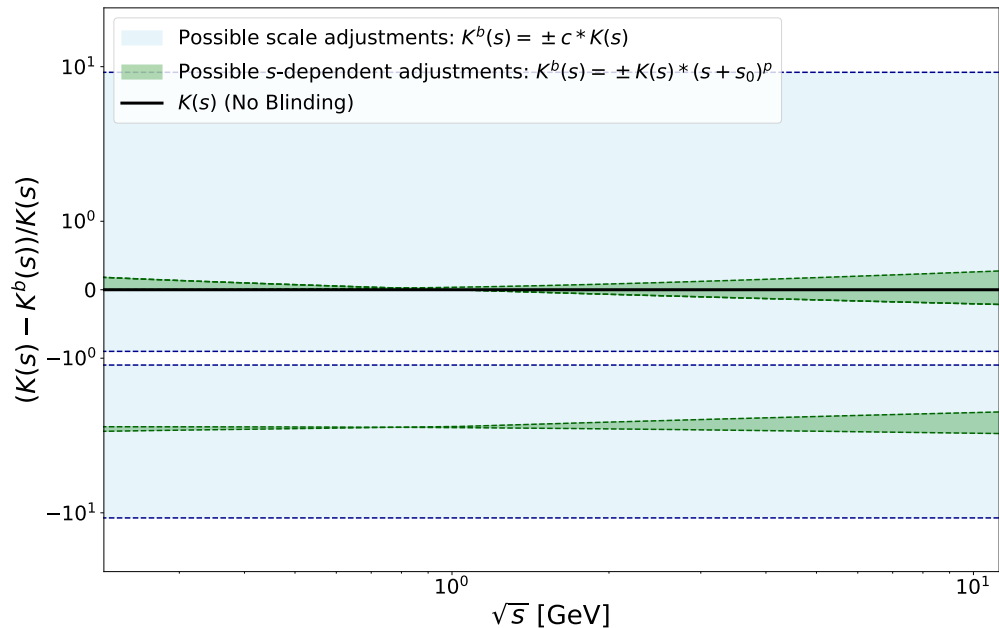
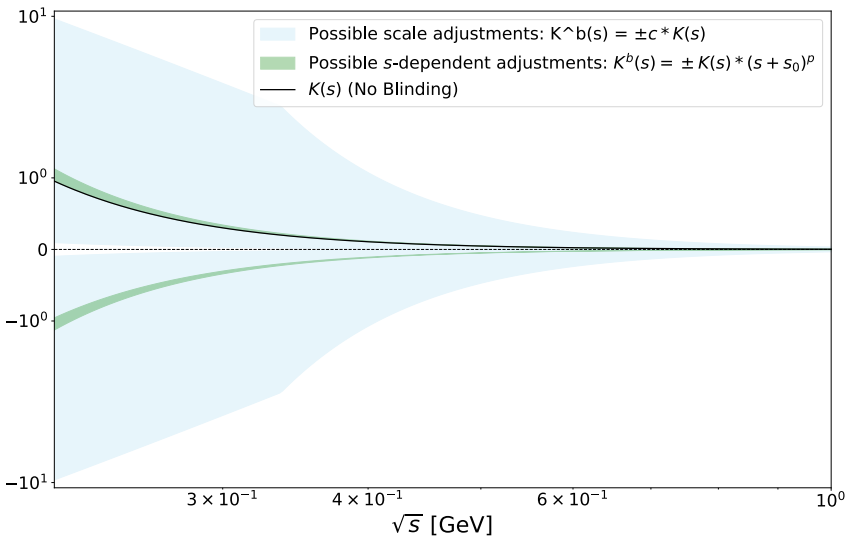
Blinding variable	Purpose	Values	Comment
a	Overall sign	± 1	
b, \tilde{b}	Overall scale factor	Between 0.1-0.9 or 1.1-10	Avoid scaling of 1
c, \tilde{c}	Power adjustment to energy dependence	Between -0.05 --> -0.01 or 0.01 --> 0.05	Avoid power of zero
s_0, \tilde{s}_0	Additive shift in energy dependence	Between -0.01 GeV ² --> 1 GeV ²	Avoid knowledge at $s = 1 \text{ GeV}^2 \rightarrow (1 \text{ GeV}^2)^c = 1 \text{ GeV}^2$



Each of these is set by a random number generated fixed by an externally set seed that results in a number in the given range

Possible ranges of blinding

Blinding variable	Purpose	Values	Comment
a	Overall sign	± 1	
b, \tilde{b}	Overall scale factor	Between 0.1-0.9 or 1.1-10	Avoid scaling of 1
c, \tilde{c}	Power adjustment to energy dependence	Between -0.05 --> -0.01 or 0.01 --> 0.05	Avoid power of zero
s_0, \tilde{s}_0	Additive shift in energy dependence	Between -0.01 GeV ² --> 1 GeV ²	Avoid knowledge at $s = 1 \text{ GeV}^2 \rightarrow (1 \text{ GeV}^2)^c = 1 \text{ GeV}^2$



The blinding seeds

Five blinding seeds are set externally to result in random numbers in those ranges:

- $s1 \rightarrow \text{sign} - a$
- $s2 \rightarrow \text{scale} - b, \tilde{b}$
- $s3 \rightarrow \text{s-offset} - s_0, \tilde{s}_0$
- $s4 \rightarrow \text{power} - c, \tilde{c}$
- $s5 \rightarrow \text{channel number map:}$

$$I^{\text{blind}} = a b \int ds (s + s_0)^c f(s) \sigma_{\text{had}}(s)$$

$$\sigma_{\text{had}}^{\text{blind}}(s) = \tilde{b} (s + \tilde{s}_0)^{\tilde{c}} \sigma_{\text{had}}(s)$$

Blinding variable	Values
a	± 1
b, \tilde{b}	Between 0.1-0.9 or 1.1-10
s_0, \tilde{s}_0	Between $-0.01 \text{ GeV}^2 \rightarrow 1 \text{ GeV}^2$
c, \tilde{c}	Between $-0.05 \rightarrow -0.01$ or $0.01 \rightarrow 0.05$

- Generates a random number between 1-100 that assigns each channel a unique number.
- Unique number is concatenated to the beginning (integrals) or end (plots) of each blinding string.

- Results in **two unique sets of blinding seeds for each hadronic channel:**

→ Integrals = $\{s5s1, s5s2, s5s3, s5s4\}$. Plots = $\{s1s5, s2s5, s3s5, s4s5\}$.

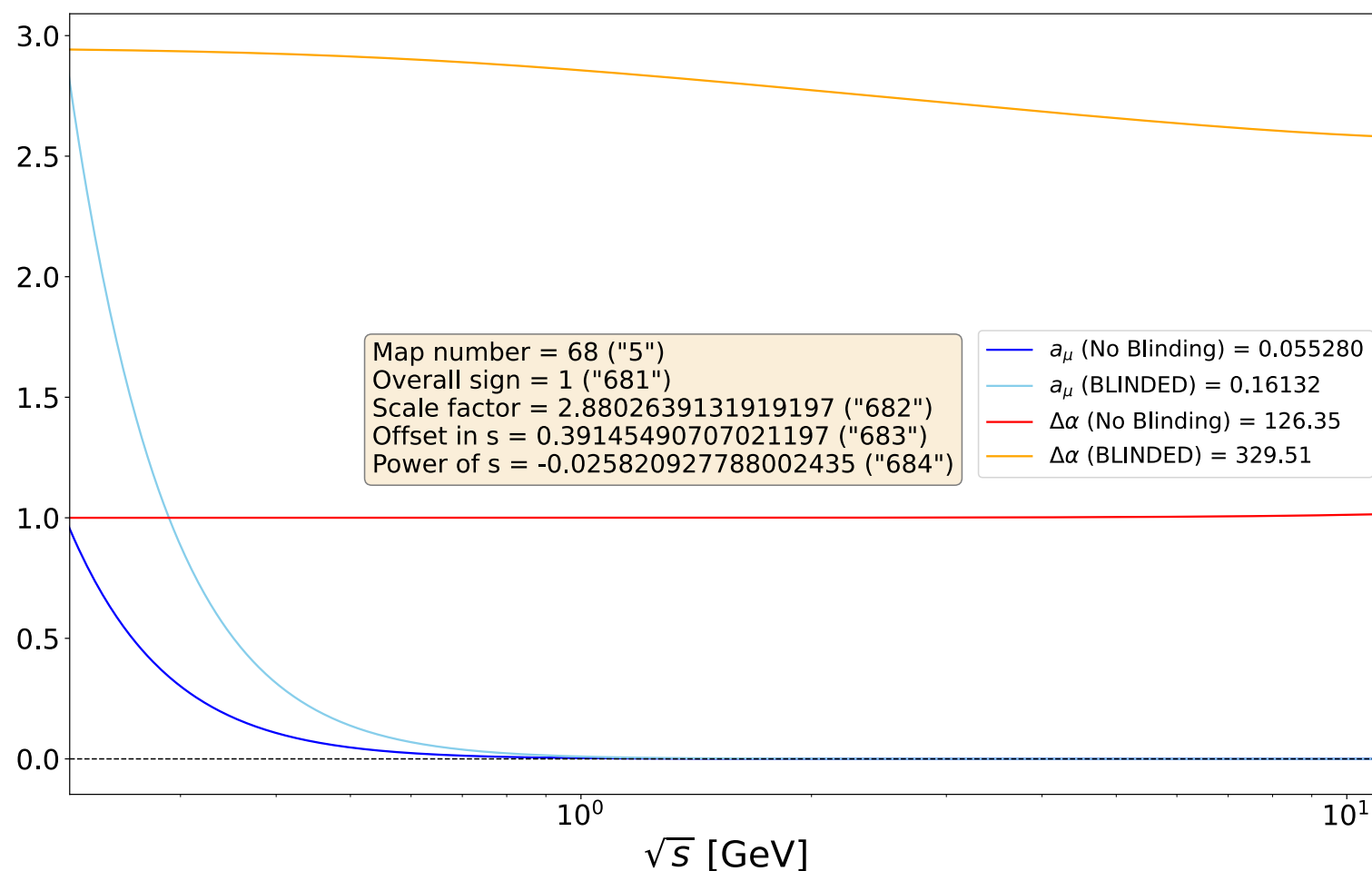
Example: $s1 = 1, s2 = 2, s3 = 3, s4 = 4, \underline{s5 = 5}$

KNT Channel	KNT Channel number	Channel number map ($s5 = 5 \rightarrow 68$)	Resulting integral seeds	Resulting plot seeds
Inclusive	0	$0+68 = 68$	$s1 = 681, s2 = 682, s3 = 683...$	$s2 = 268, s3 = 368...$
$\eta\pi^+\pi^-$	1	$1+68 = 69$	$s1 = 691, s2 = 692, s3 = 693...$	$s2 = 269, s3 = 369...$
K^+K^-	2	$2+68 = 70$	$s1 = 701, s2 = 702, s3 = 703...$	$s2 = 270, s3 = 370...$

Blinding seeds example: integral kernels

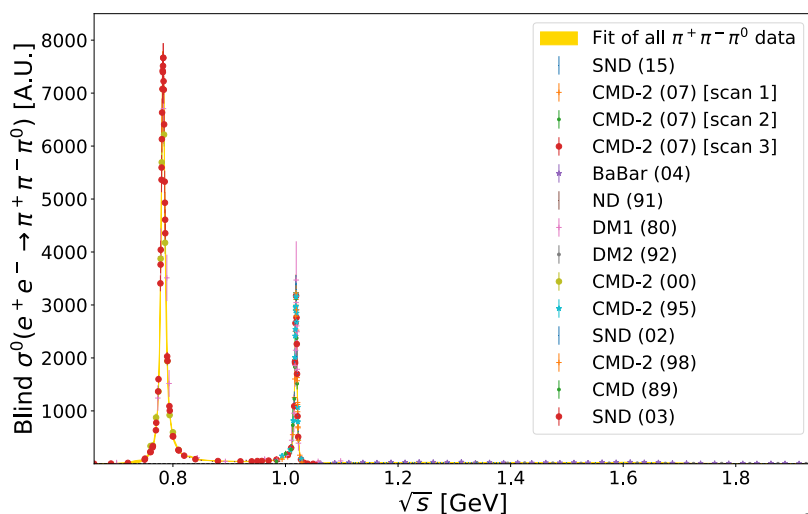
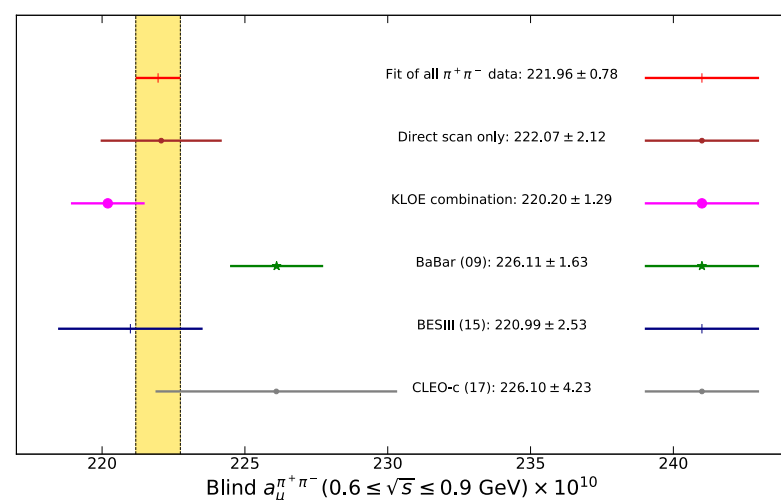
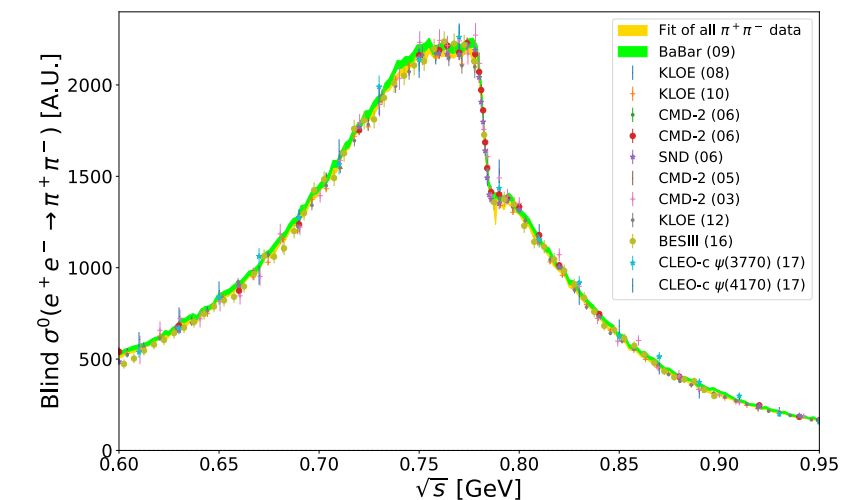
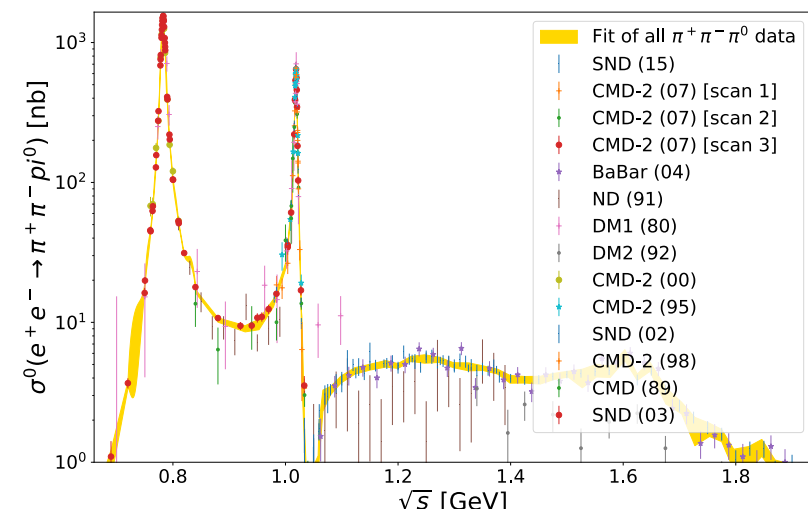
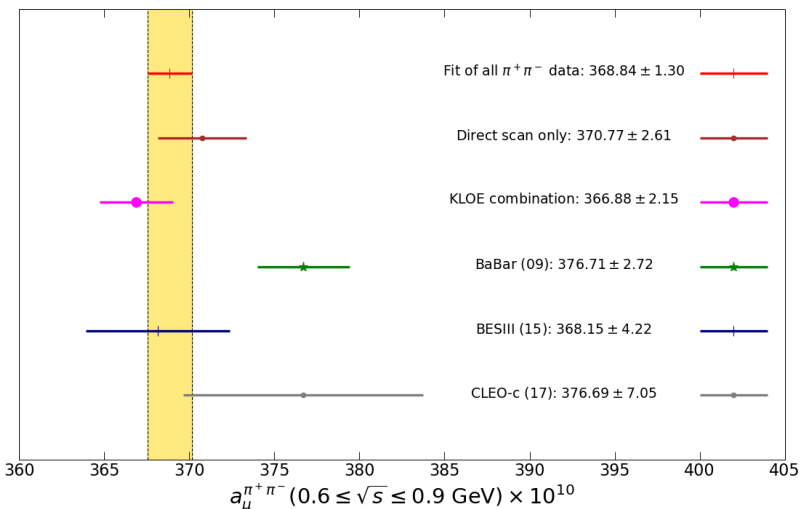
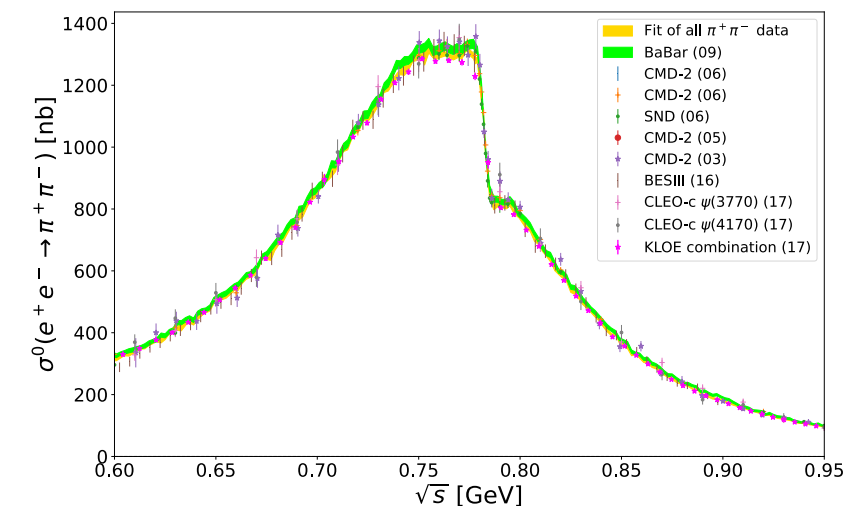
Example: $s_1 = 1$, $s_2 = 2$, $s_3 = 3$, $s_4 = 4$, $s_5 = 5$

Channel number	Channel number map ($s_5 = 5 \rightarrow 68$)	Resulting integral seeds	Resulting plot seeds
0	$0 + 68 = 68$	$s_1 = 681, s_2 = 682, s_3 = 683 \dots$	$s_2 = 268, s_3 = 368 \dots$



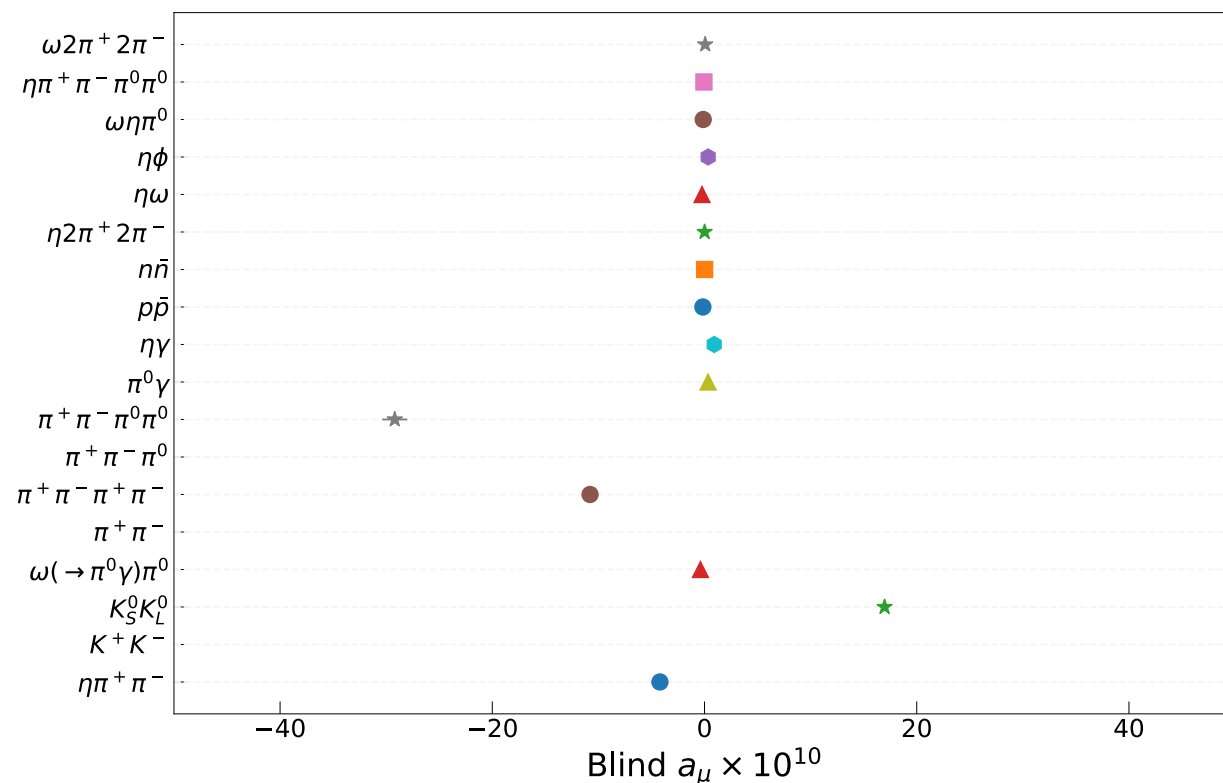
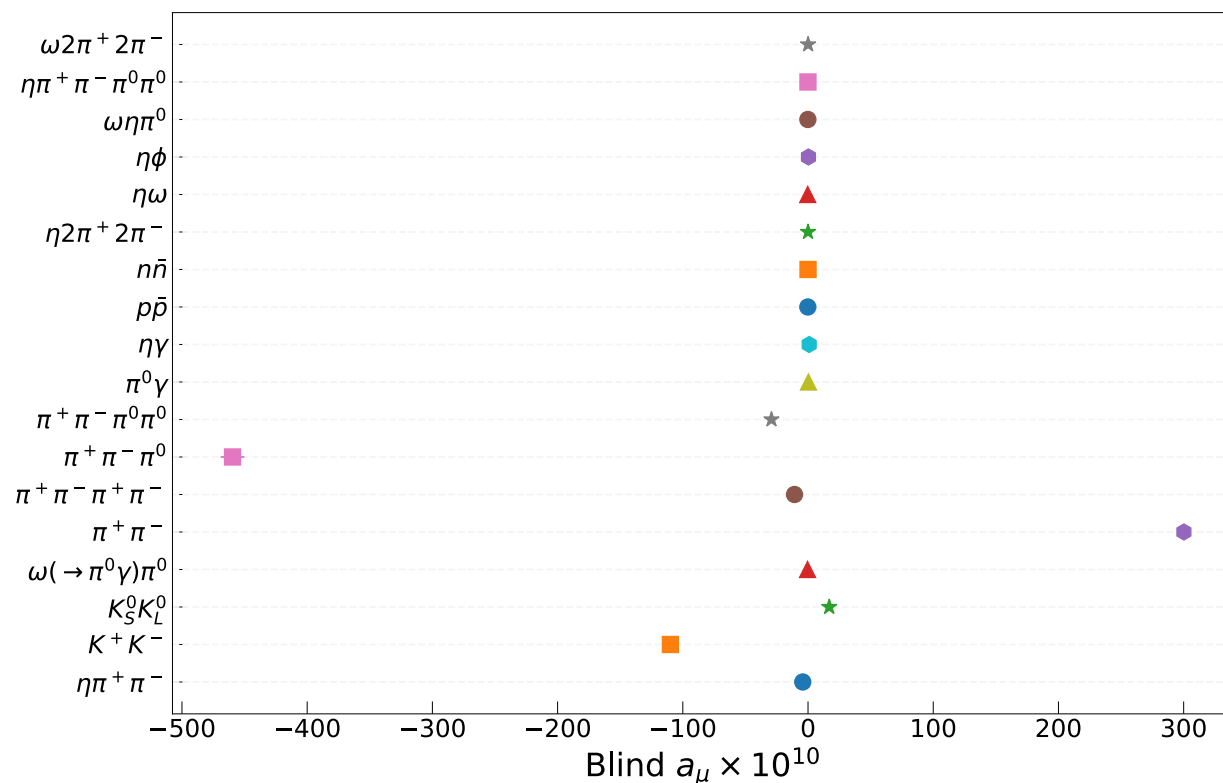
Blinding seeds example: KNT19

New example: s1=11111, s2=2222, s3=333, s4 = 44, s5 = 5



Blinding seeds example: KNT19

Example: $s_1=11111$, $s_2=2222$, $s_3=333$, $s_4 = 44$, $s_5 = 5$



(Un)Blinding implementation

The code works on multiple OS's and is machine-independent. It is a mix of languages:

- Python – random number generators.
- C++ - interface between the KNT HVP (FORTRAN) and the Python.
 - Blinding seeds set in a header file.
 - Map-seed concatenation for integrals/plots.
 - Check blinding seed/values are within allowed ranges.
- FORTRAN – initialises the code and implements the blinding:
 - Calls C++/FORTRAN and sets blinding values.
 - Allows for initial user input to:
 - **Relatively unblind** (remove channel map number).
 - (remove offsets entirely).
 - Returns a check **Fully unblind** sum on use.

```
*****
!!!!!!!!!!!!!! YOU HAVE CHOSEN TO BLIND YOUR ANALYSIS !!!!!!!!!!!!!!!

      Integrated values will be blinded according to:
      integrate [ ds * +/-a*(s+s0)^b * K(s) * sigma*(s) ]

*****

Choose from one of the following options (please enter 1,2 or 3):

  1) Fully blinded (individual blinding for each channel)
  2) Relative unblinding (common blinding between channels)
  3) Full unblinding

(Note that options 2) and 3) will need you to enter blinding seeds to continue)
```

(Note: unblinding requires knowing and entering the seeds and cannot happen without external input!)

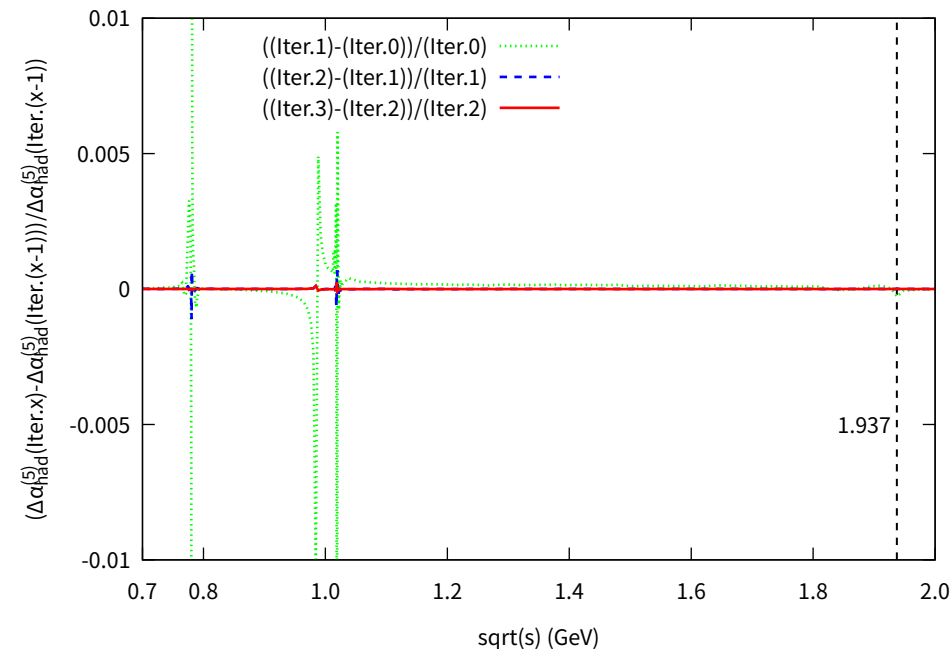
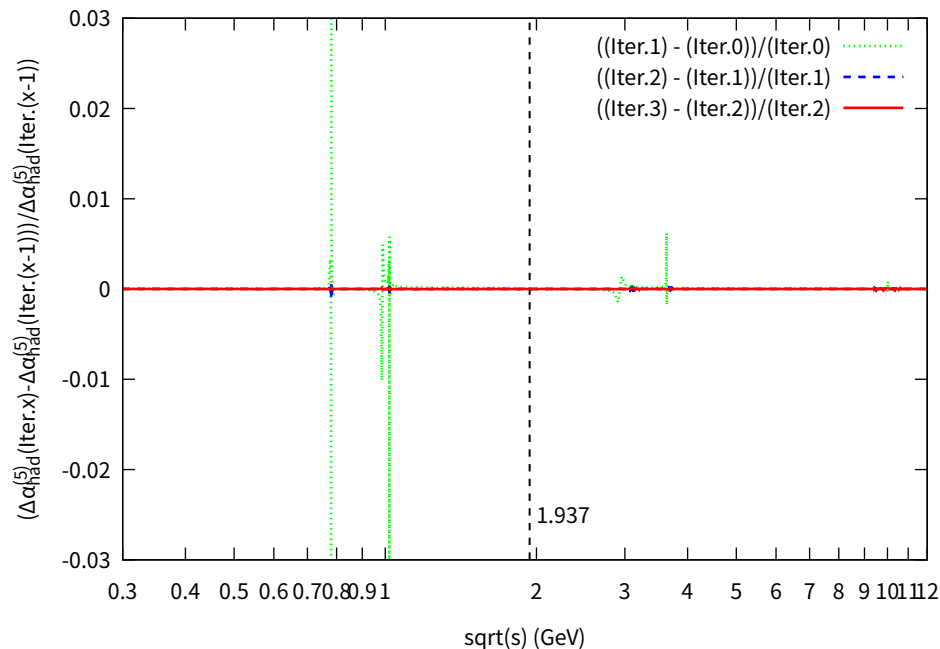
Once compiled, only the compiler object files are accessible to us!

- Cannot access the seeds, blinding values, or any other variables.
 - All private and not accessible in the compiled objects.
 - We cannot set variables to turn on/off the blinding once it's included in the software.

A test run: blinded VP iteration

For KNT19, we did not perform a VP iteration. We have **done this now in a blind way**.

- Towards the end of an analysis, the VP iteration will be performed following relative unblinding, once all channel combinations have been finalised.
 - The VP iteration will then **serve as a final cross check** before full unblinding.
 - Using a set of blinding seeds, we have performed a blinded VP iteration on the KNT19 data.
- We **successfully reached convergence** and have produced a new VP routine without compromising the blinding.



This means we now start our new blind KNT analysis with updated VP corrections produced under blind conditions (which in principle blinds our results relative to KNT19).

Blinding implementation: the external

The steps to set the blinding seeds and implement them in the analysis **should all be performed with the external present and recorded:**

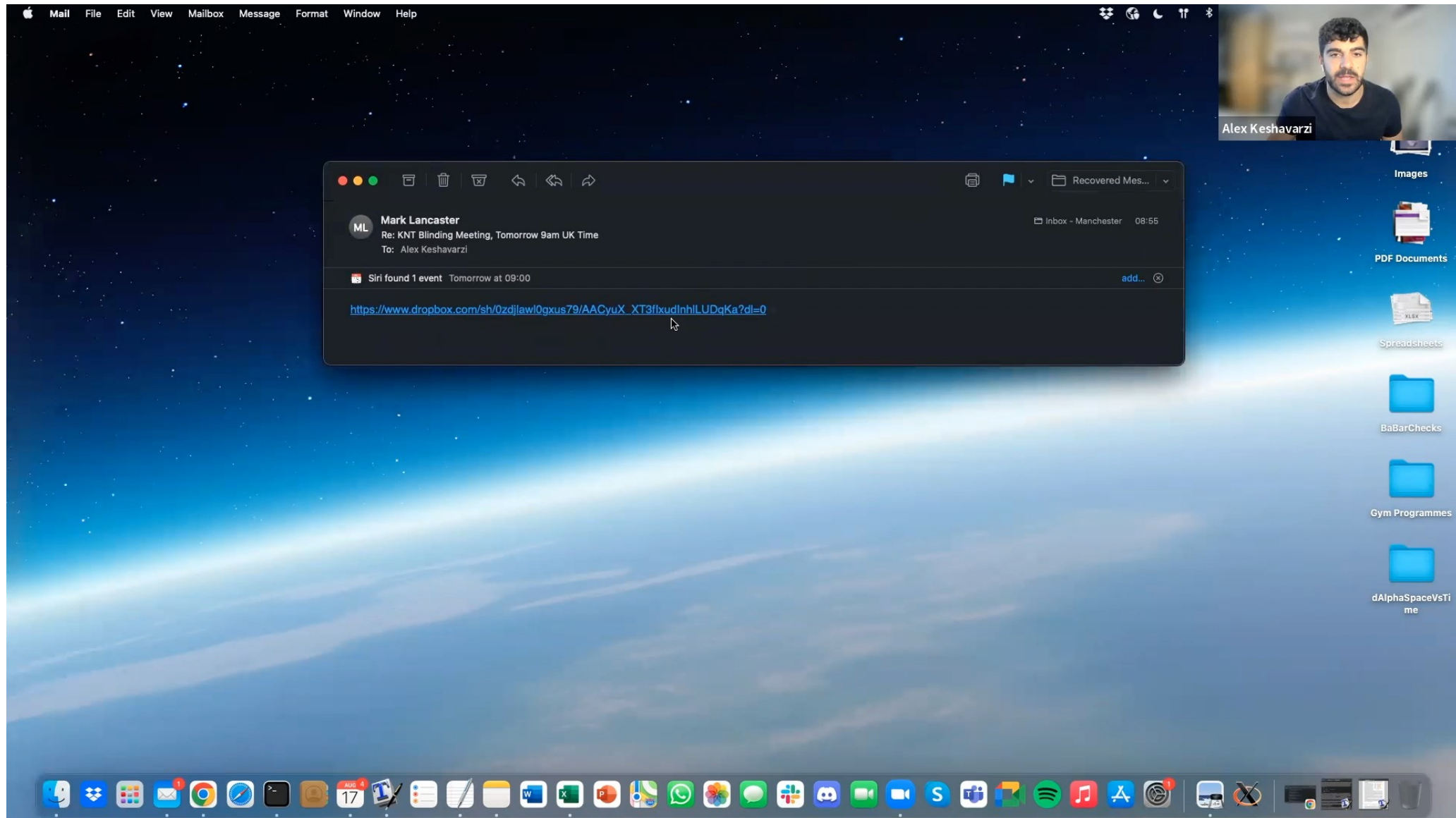
1. The external receives the blinding software.
 - i. **External sets blinding** seeds in the header file.
 - ii. External runs a script which **encrypts all sources files** with a provided password.
2. External sends analyser the encrypted source files and a script to perform an extraction.
 - i. The **analyser can download the file only once**.
3. Analyser runs the extraction script.
 - i. **External will provide the analyser with the password** to unencrypt.
 - ii. The script will then compile all files, provide a checksum and **delete all source files, leaving only the object files**.
4. Analyser can then compile KNT routines against the object files and **will run blinded**.

The blinding offsets for our new analysis were set by Mark Lancaster.

They were implemented on 17/08/22.

The KNT analysis is now fully blinded.

Implementing the blinding for real...



The new, blind KNT analysis

The following (and any new) data sets will be included in the now blind KNT analysis as an update to KNT19:

- $\pi^+\pi^-\pi^0$, BESIII (2019), arXiv:1912.11208
- $\pi^+\pi^-$ [covariance matrix erratum], BESIII (2020), Phys.Lett.B 812 (2021) 135982 (erratum)
- $K^+K^-\pi^0$, SND (2020), Eur.Phys.J.C 80 (2020) 12, 1139
- $e^+e^-\pi^0$ (res. only), SND (2020), Eur.Phys.J.C 80 (2020) 11, 1008
- $\pi^+\pi^-$, SND (2020), JHEP 01 (2021) 113
- $\eta\omega \rightarrow \pi^0\gamma$, SND (2020), Eur.Phys.J.C 80 (2020) 11, 1008
- $\pi^+\pi^-\pi^0$, SND (2020), Eur.Phys.J.C 80 (2020) 10, 993
- $\pi^+\pi^-\pi^0$, BaBar (2021), Phys.Rev.D 104 (2021) 11, 112003
- $\pi^+\pi^-2\pi^0\omega$, BaBar (2021), Phys. Rev. D 103, 092001
- $\eta\omega$, SND (2021), Eur.Phys.J.C 82 (2022) 2, 168
- $\eta\omega$, BaBar (2021), Phys.Rev.D 104 (2021) 11, 112004
- $\pi^+\pi^-\pi^0\eta$, BaBar (2021), Phys.Rev.D 104 (2021) 11, 112004
- $\omega\pi^0$, BaBar (2021), Phys. Rev. D 103, 092001
- $\pi^+\pi^-4\pi^0$, BaBar (2021), Phys.Rev.D 104 (2021) 11, 112004
- $\pi^+\pi^-\pi^0\eta$, BaBar (2021), Phys.Rev.D 103 (2021) 9, 092001
- $\pi^+\pi^-3\pi^0\eta$, BaBar (2021), Phys.Rev.D 104 (2021) 11, 112004
- $2\pi^+2\pi^-3\pi^0$, BaBar (2021), Phys. Rev. D 103, 092001
- $\omega3\pi^0$, BaBar (2021), Phys.Rev.D 104 (2021) 11, 112004
- $\pi^+\pi^-\pi^+\pi^-\eta$, BaBar (2021), Phys. Rev. D 103, 092001
- Inclusive $R(s)$, BESIII (2021), Phys.Rev.Lett. 128 (2022) 6, 062004
- $n\bar{n}$, SND (2022), arXiv:2206.13047
- $K^0_S K^0 \pi^0$, CMD-3 (2022), arXiv:2207.04615
- $KK^0\pi^0$, BaBar (2022). arXiv:2207.10340

Conclusions

- The next KNT analysis is a blind analysis.
 - So, **no new results until unblinding**.
 - Only expect blinded comparisons in given channels.
- **Each hadronic channel has a distinct blinding**.
 - Allows for two layers of blinding.
 - Procedure blinds integrals and plots for each separately **in an energy-dependent way**.
 - **No data are adjusted** → integrity of data combination is not at risk.
- Routine has been devised so there is **no way to disentangle / un-do the blinding** once it is included.
- Procedure has **been stress-tested for KNT19 and performs as expected**.
- A **blinded VP iteration of KNT19 data** has resulted in new VP routine ready for a new, blind analysis.
- **Blinding was implemented with Mark Lancaster as the external**.
 - He alone knows the blinding seeds.
 - He will oversee unblinding at the appropriate time.
- **New data since KNT19, and new data to be released, are to be included under this new blind paradigm.**

We are still happy to produce/share results with the community, but will of course be based on only KNT19 data.

Backups