

Paths to Geant4 evolution: Refactoring, Reengineering and Physics

State-of-the-Art Simulation of Photon Interactions with Matter

Maria Grazia Pia

M. Batič, M. Han, S. Hauf, G. Hoff, C. H. Kim, M. Kuster,
P. Saracco, H. Seo, G. Weidenspointner, A. Zoglauer

INFN Genova, Italy

J. Stefan Institute, Ljubljana, Slovenia

Hanyang University, Seoul, Korea

PUCRS, Porto Alegre, Brazil

EU XFEL GmbH, Hamburg, Germany

MPI Halbleiterlabor, Munich, Germany

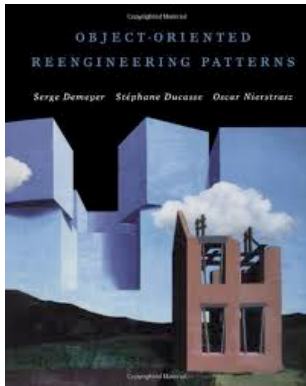
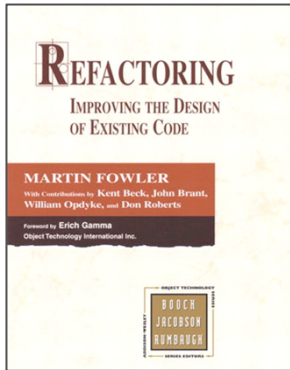
Space Sciences Laboratory, UC Berkeley, USA



NSS 2012

30 October 2012, Disneyland, USA





“The code slowly sinks from engineering to hacking”.

“Refactoring is the process of changing a software system in such a way that it does not alter the external behavior of the code yet improves its internal structure.”

“When you refactor you are improving the design of the code after it has been written.”

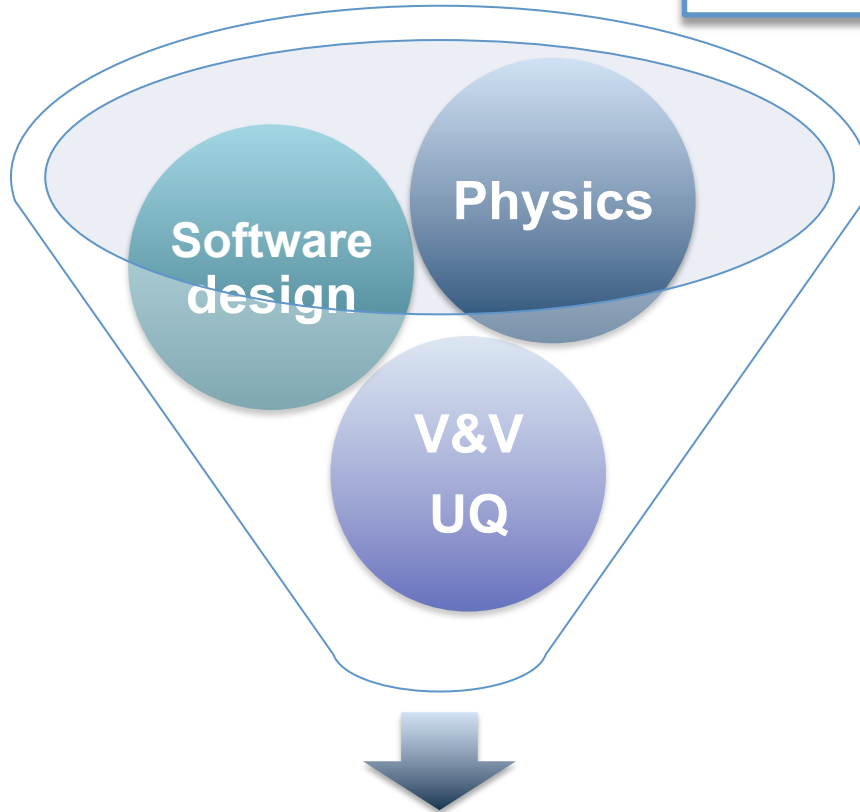
Reengineering “seeks to transform a legacy system into the system you would have built if you had the luxury of hindsight and could have known all the new requirements that you know today.”

“A legacy is something valuable that you have inherited.”

“With rapid development tools *and* rapid turnover in personnel, software systems can turn into legacies more quickly than you might imagine.”

R&D Project

Series of **pilot projects**
going on since 2008



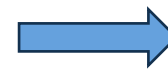
Focus on **physics** and
fundamental concepts
of particle transport

Highlights

No time to go into details

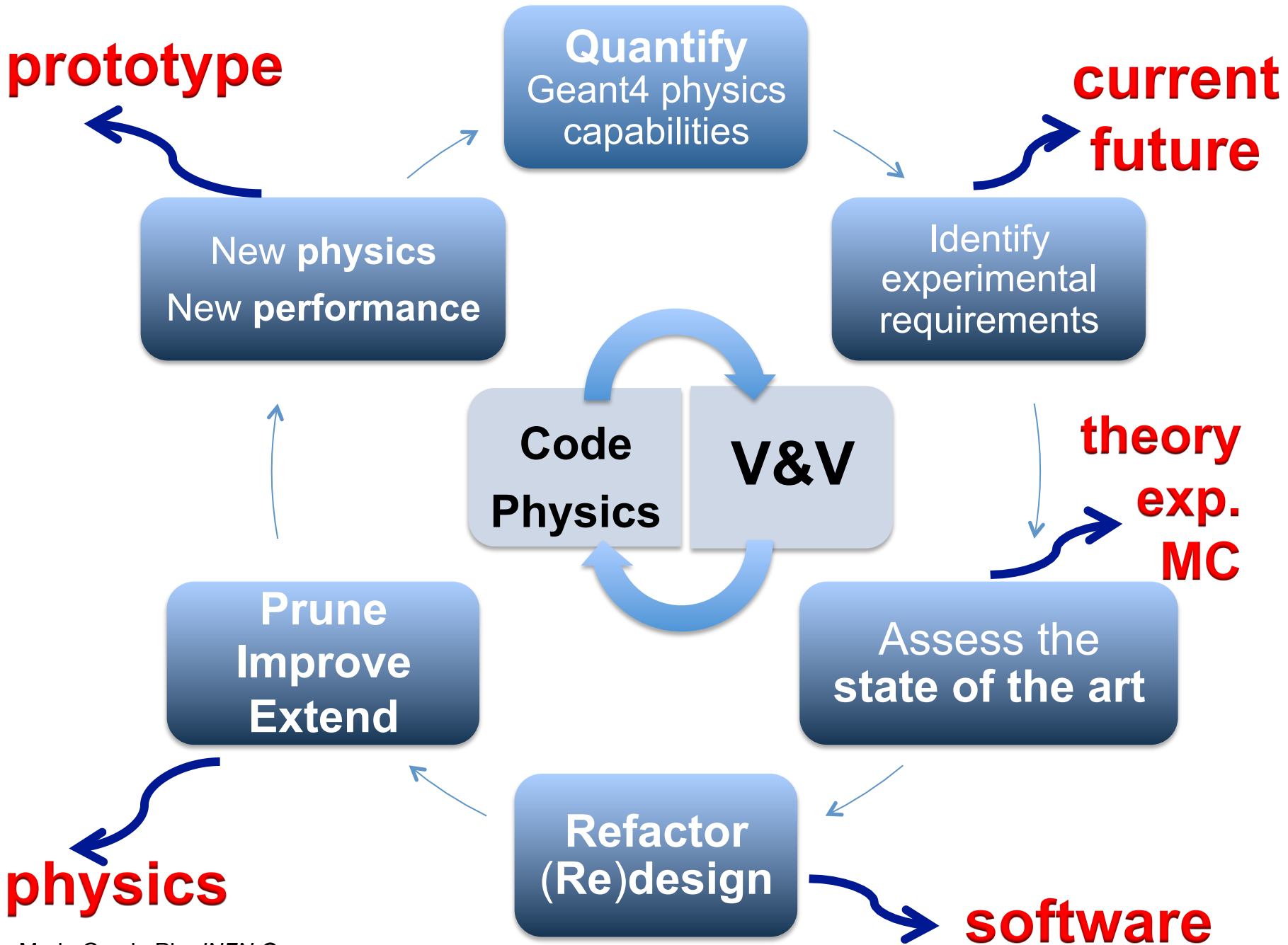
State-of-the-art, quantified

Geant 4



Archival
literature

Produce concrete deliverables





Smells

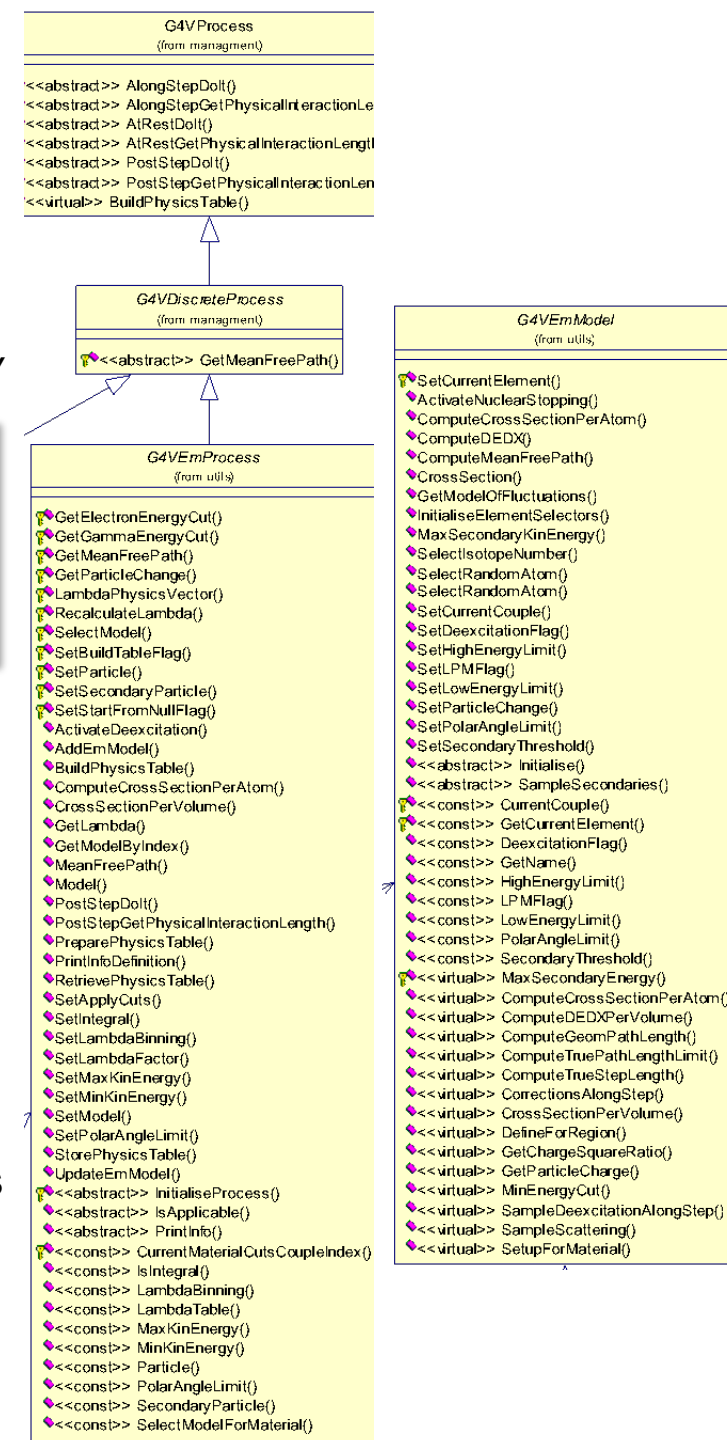
If it stinks, change it.

Grandma Beck, discussing child-rearing philosophy

M. Fowler, K. Beck et al.,

Refactoring: Improving the Design of Existing Code

- Duplicated Code
- Long Method
- Large Class
- Long Parameter List
- Divergent Change
- Shotgun Surgery
- Feature Envy
- Data Clumps
- Primitive Obsession
- Switch Statements
- Parallel Inheritance Hierarchies
- Lazy Class
- Speculative Generality
- Temporary Field
- Message Chains
- Middle Man
- Inappropriate Intimacy
- Alternative Classes with Different Interfaces
- Incomplete Library Class
- Data Class
- Refused Bequest



Electromagnetic smells

Coupling

σ_{tot} and final state modeling
have been decoupled in hadronic
physics design since RD44

Dependencies

on other parts of the software

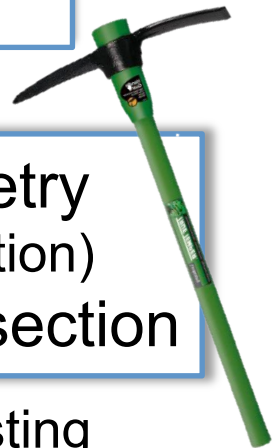
“model”

Total cross section

Whether a process occurs

Final state generation

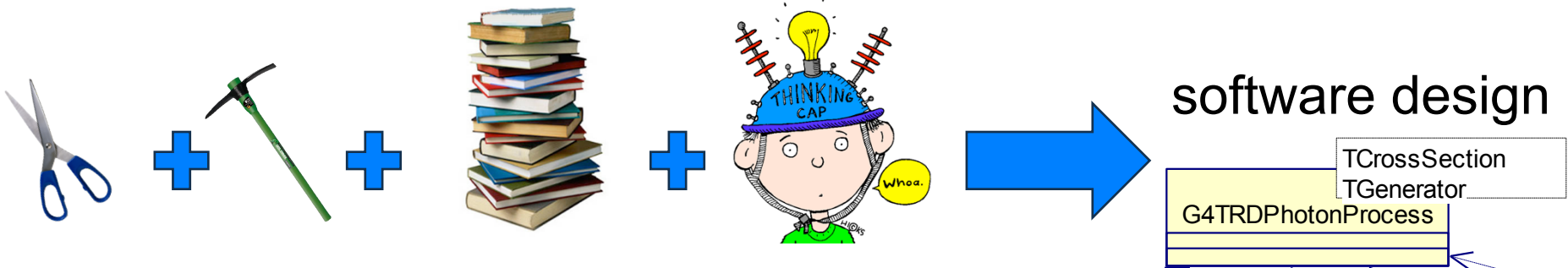
How a process occurs



One needs a geometry
(and a full scale application)
to test (verify) a cross section

Difficult to test \rightarrow no testing
often

Problem domain analysis
Improve domain decomposition



Benefits

Transparency

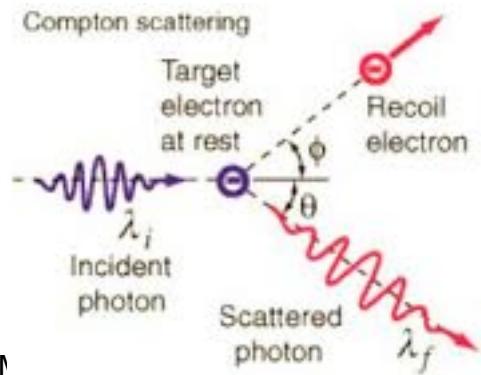
Ease of maintenance

Simplicity of **testing** for V&V

Basic physics V&V can be performed by means of **lightweight unit tests** → Easier exploration of new models
Easier quantification of accuracy

Bare-bone physical functionality

Decorations on top



Motion of atomic electrons
Doppler broadening
Binding effects

Vacancy creation
Atomic relaxation

scattering functions

Doppler profiles

Prune

Number one in the stink parade is duplicated code

M. Fowler,
Refactoring

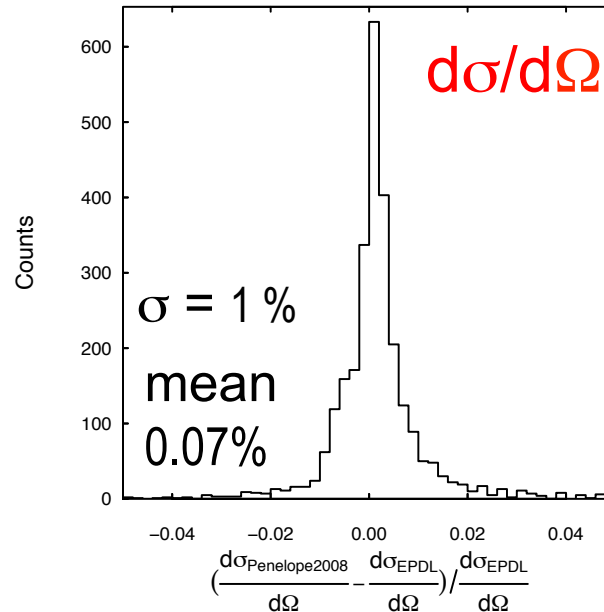
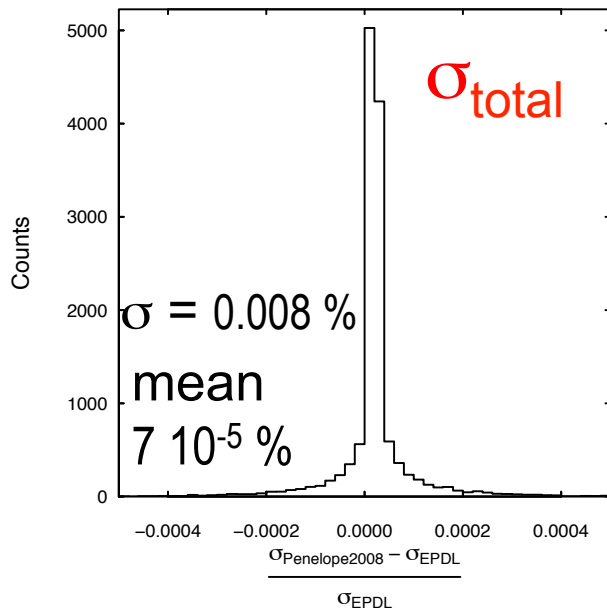
physics

Objective quantification of smell

Two Geant4 models, identical underlying physics content (*it used to be different*)

Efficiency w.r.t. experiment

“Livermore”	Penelope
EPDL97	EPDL97
0.38 ± 0.06	0.38 ± 0.06



Code bloat

Burden on

- Software design
- Maintenance
- User support

Unnecessary complexity

Trash and redo

Number one in the stink
parade is duplicated ~~code~~

numbers

1. Bearden & Burr (1967)
2. Carlson
3. EADL
4. Sevier
5. Tol 1978 (Shirley)
6. Tol 1996 (Larkins)
7. Williams

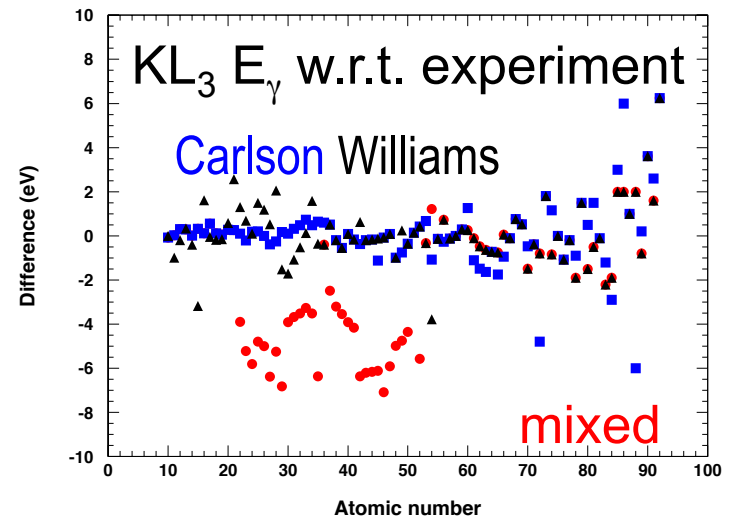
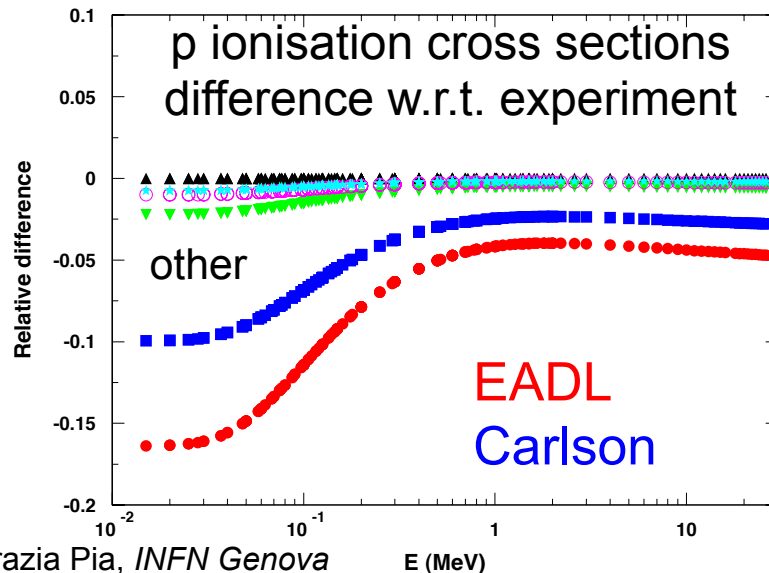
Atomic binding energies

Geant 4 { Carlson + Williams
EADL (Carlson
Shirley)

Source of epistemic uncertainties?

3246 IEEE TRANSACTIONS ON NUCLEAR SCIENCE, VOL. 58, NO. 6, DECEMBER 2011
Evaluation of Atomic Electron Binding Energies
for Monte Carlo Particle Transport
Maria Grazia Pia, Hee Seo, Matej Batic, Marcia Begalli, Chan Hyeong Kim, Lina Quintieri, and Paolo Saracco

23
pages



Algorithms

Popular belief

Physics model X is intrinsically slow

Baroque methods to combine it with “faster” lower precision models and limit its use to cases where one is willing to pay for higher precision

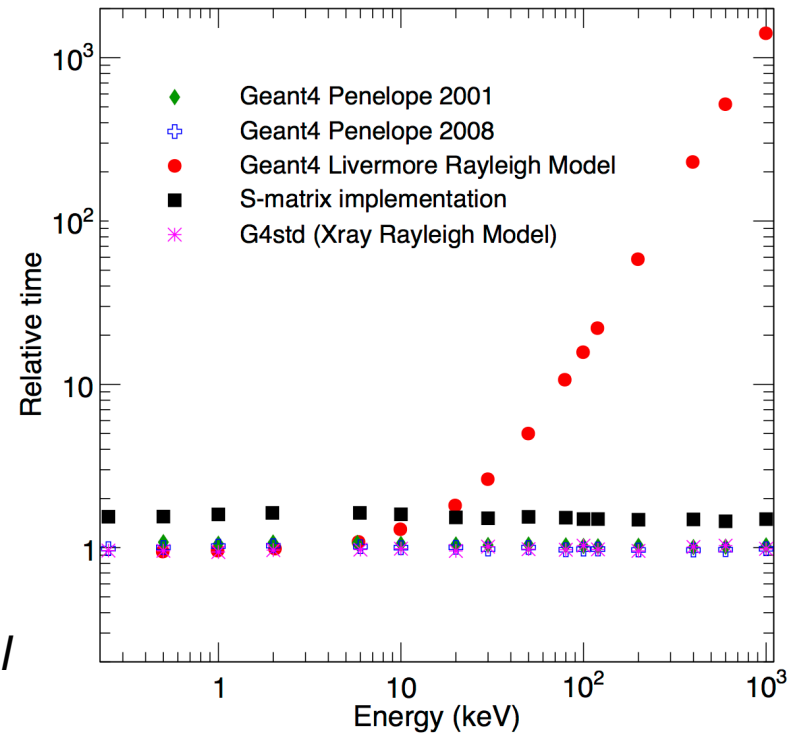
This design introduces an additional computational burden due to the effects of inheritance and the combination algorithms themselves

Truth

Physics model X is intrinsically fast

But its computationally fast physics functionality is spoiled by an inefficient sampling algorithm

Change the sampling algorithm!



- ▶ **No code smell**
- ▶ Spotted through in-depth **code review** in the course of **software validation**

The fastest algorithm

no algorithm at all

Shift modeling from algorithms to data

Merging models



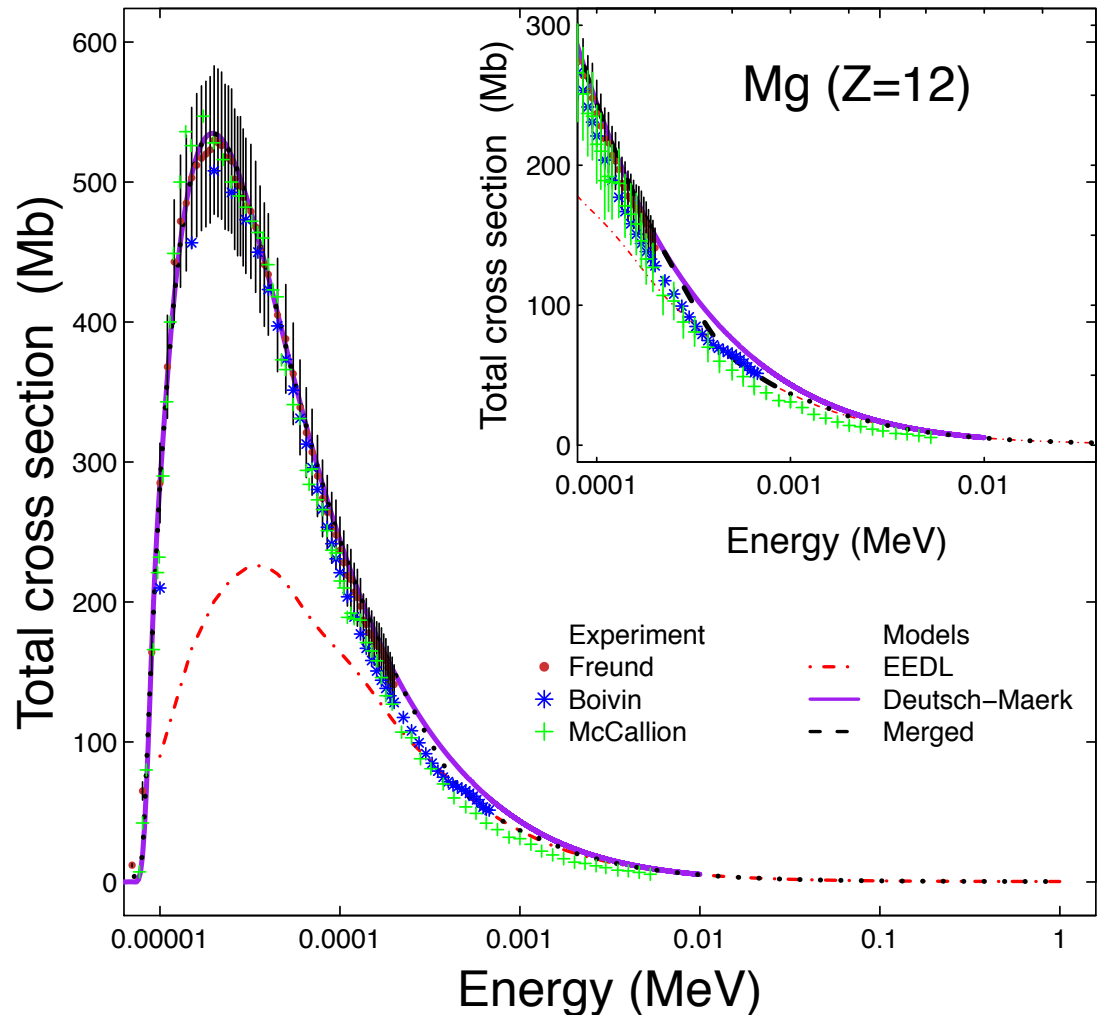
Smoothing data

Guidance from
experimental data
(when available)

Data libraries

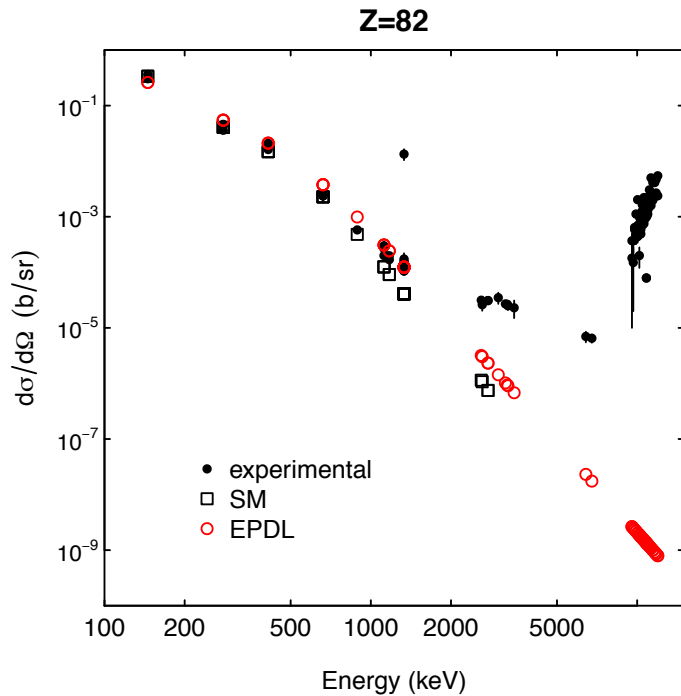
- PIXE ($\sigma_{\text{ionisation}}: p$)
- BEB-DM ($\sigma_{\text{ionisation}}: e^-$)
- Photon elastic
 - RTAB (L. Kissel)

Electron impact ionisation cross sections



Example: LOESS local polynomial regression fitting
Beware: not optimized! Mathematical support in progress

...no silver bullet



Photon elastic scattering
cross section at 90°

*Experimental data by Schumacher et al.,
Starr et al., Jackson et al., Moreh et al. etc.,
complete list in related publication*

Smoothing models (or data) that exhibit significantly different behaviour is not physically justified

← Here our understanding of the underlying physics phenomena fails

Experimental data themselves must be analyzed for inconsistencies and possible systematics

How to solve this kind of problems?

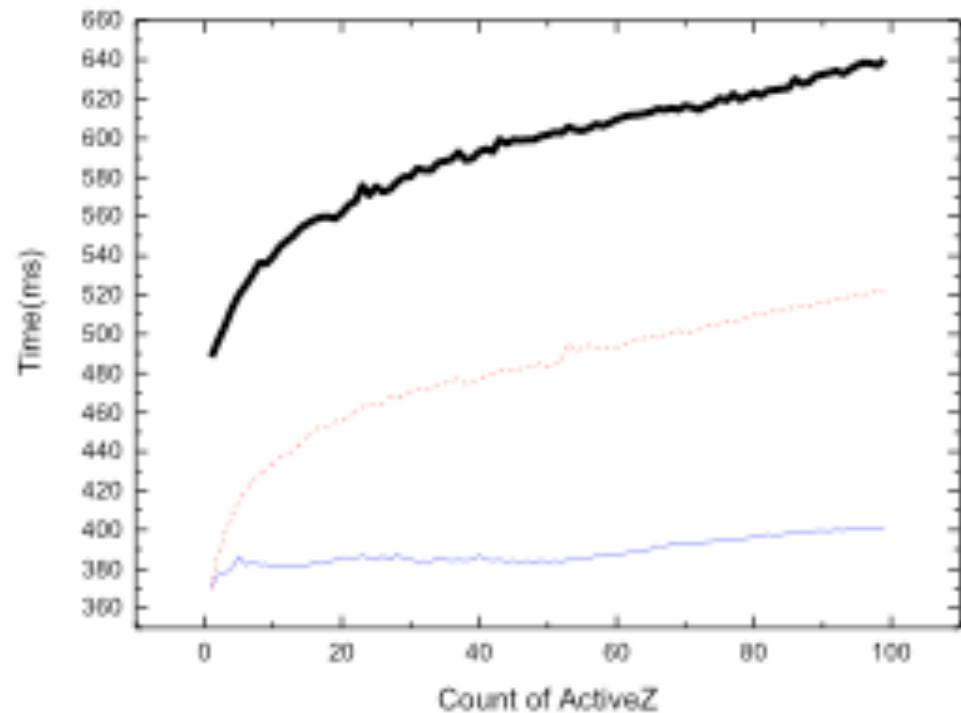
Better theoretical calculations
New, dedicated experimental data

No easy solution

Refactoring data management

- Today's technology
 - ...keeping an eye on the new C++ Standard

- Optimal container
- Pruning data
- Splitting files
- Software design

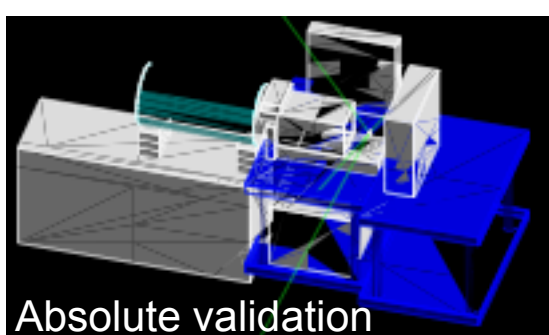


Big refactoring

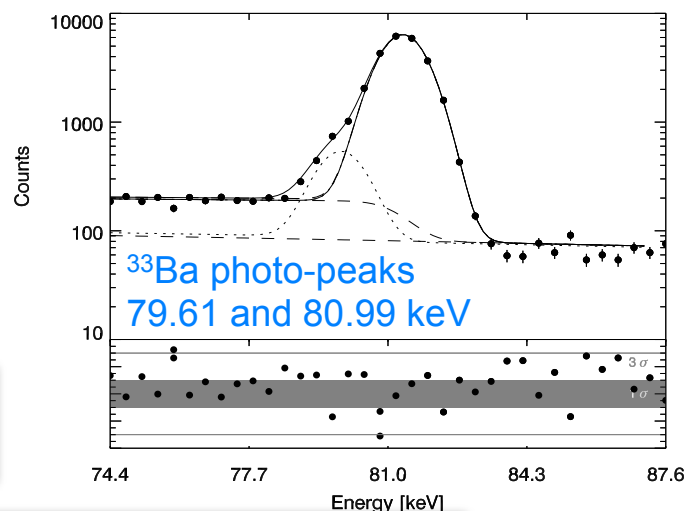
Geant4 Radioactive Decay

Well defined responsibilities and interactions





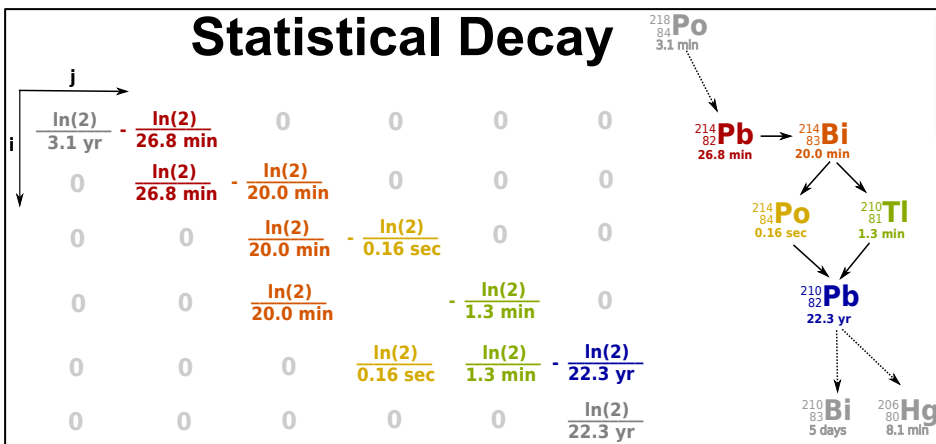
Absolute validation



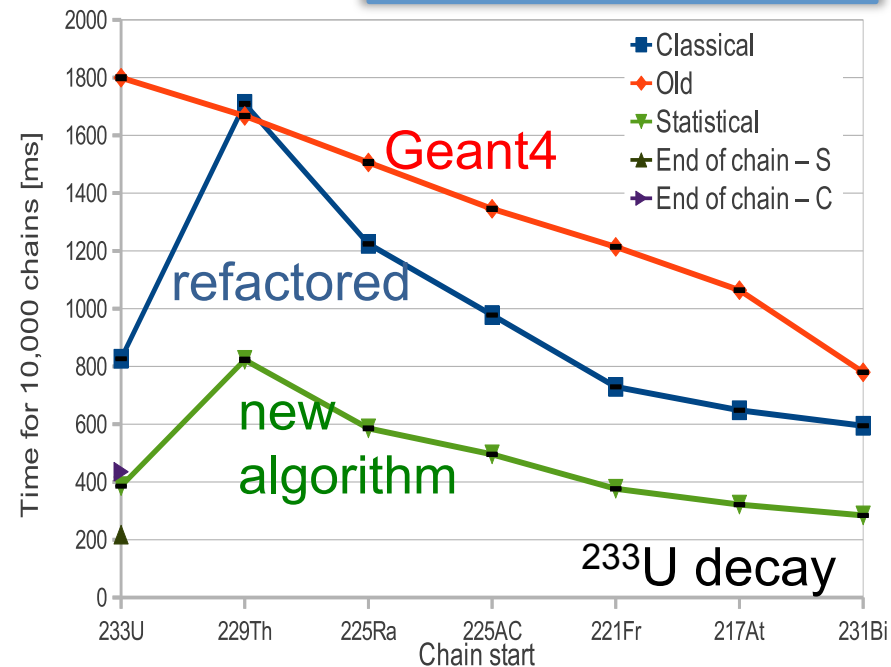
Experiment: Z. W. Bell (ORNL)

Validation

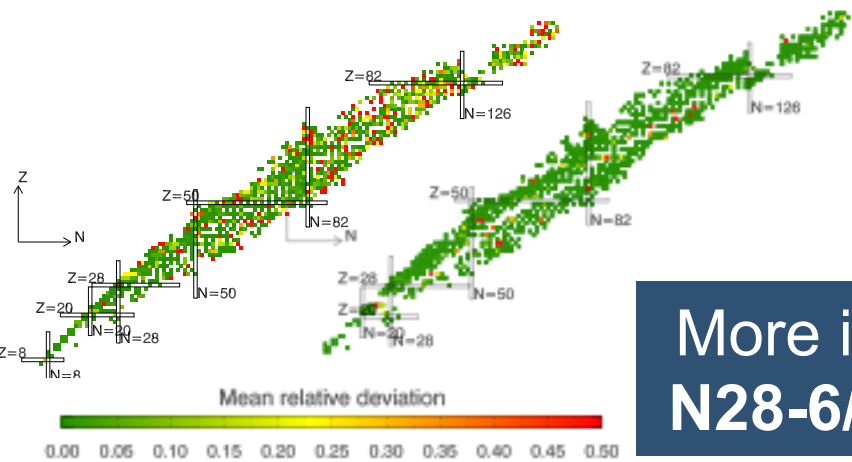
New algorithm



Performance



More in N28-6/7



Paths to Geant4 evolution: Refactoring, Reengineering and Physics

State-of-the-Art Simulation of Photon Interactions with Matter

Maria Grazia Pia

M. Batič, M. Han, S. Hauf, G. Hoff, C. H. Kim, M. Kuster,
P. Saracco, H. Seo, G. Weidenspointner, A. Zoglauer

INFN Genova, Italy

J. Stefan Institute, Ljubljana, Slovenia

Hanyang University, Seoul, Korea

PUCRS, Porto Alegre, Brazil

EU XFEL GmbH, Hamburg, Germany

MPI Halbleiterlabor, Munich, Germany

Space Sciences Laboratory, UC Berkeley, USA

NSS 2012

30 October 2012, Disneyland, USA



Refactoring to improve physics

Refactoring

M. Fowler, Refactoring

is the process of changing a software system in such a way that it **does not alter the external behavior** of the code yet **improves its internal structure**

Refactoring/reengineering legacy simulation code is an opportunity

- To complement software testing with **physics validation**
- To evaluate and compare the **accuracy** of various modeling options
- To estimate the trade-off between **accuracy** and **performance**
- To document the **state-of-the-art** for Monte Carlo transport

Electron-photon physics

Models and data libraries used by Geant4 and other Monte Carlo codes



Photon elastic scattering

Physics

- Rayleigh scattering
- Thomson nuclear scattering
- Delbrück scattering
- Nuclear resonance scattering

Form factor approximation

- **Non relativistic** form factors
(Hubbell 1977) **FLUKA, Geant4, MCNP, Penelope**
- **Relativistic** form factors
(Hubbell & Overbø) **EGS, ITS, XCOM**
- **Modified** form factors
(Schaupp et al.)
- + **anomalous scattering** factors

Geant 4

“Livermore”	<i>EPDL</i>
“Penelope 2008”	<i>EPDL</i>
“Penelope 2001”	<i>Baro et al.</i>
“Standard”	<i>Unknown</i>

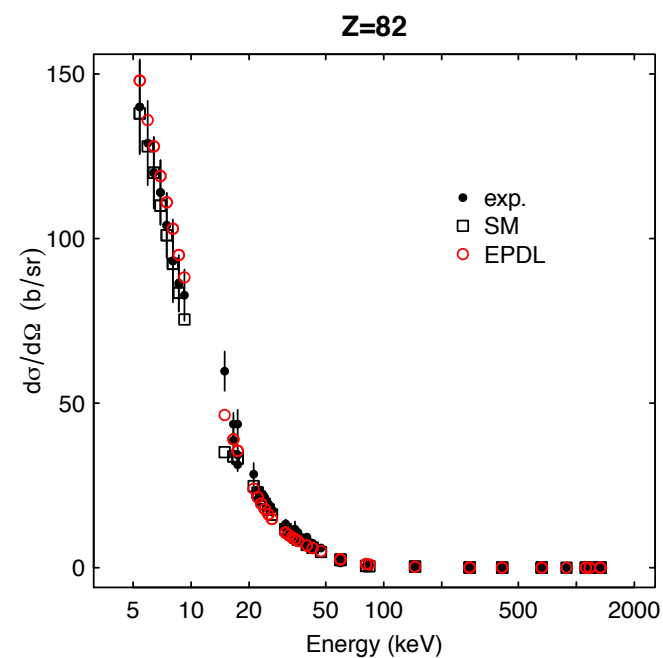
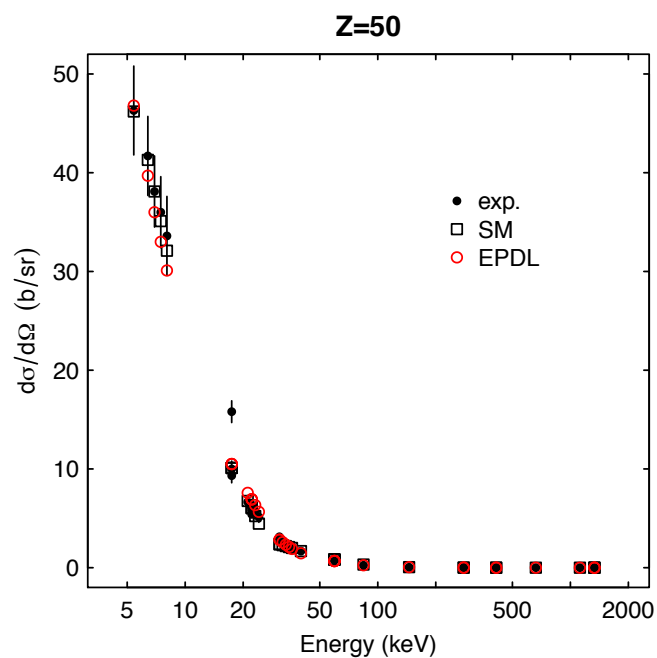
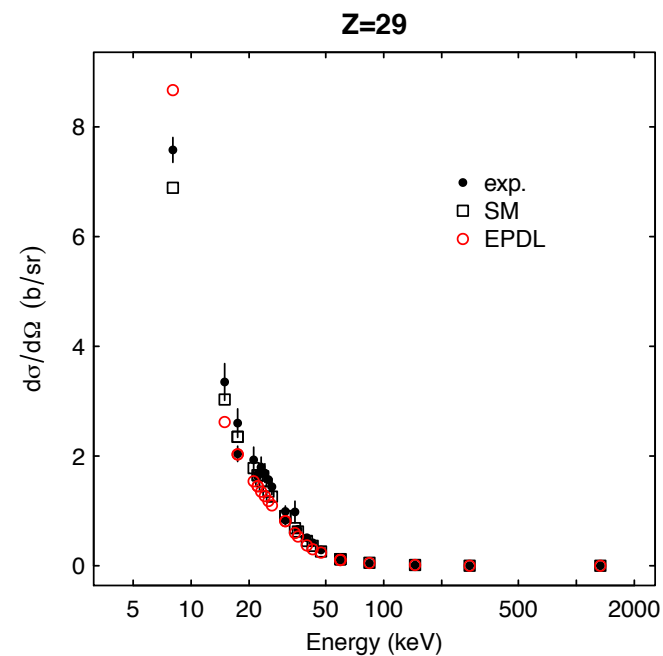
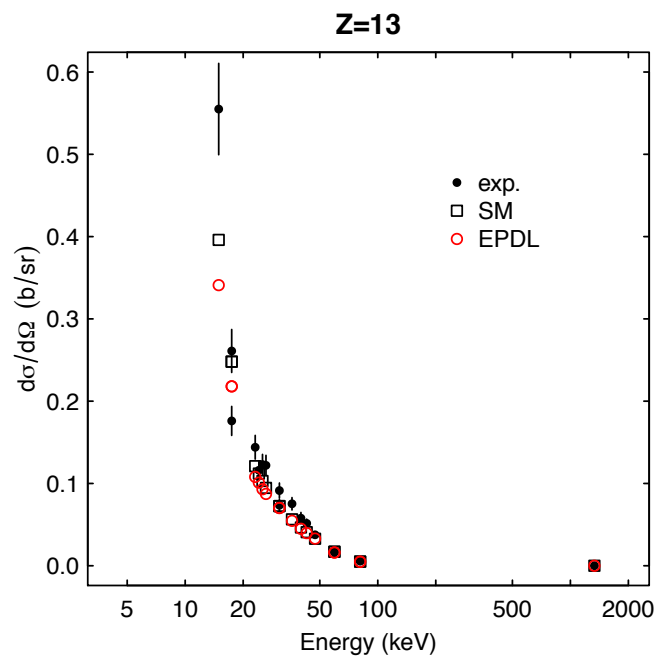
*EPDL: non relativistic form factors
Thomson scattering in total cross
section only*

2nd order S-matrix

- Calculations by L. Kissel
- Not yet exploited in Monte Carlo codes

Forward
peaked

Heavy
materials



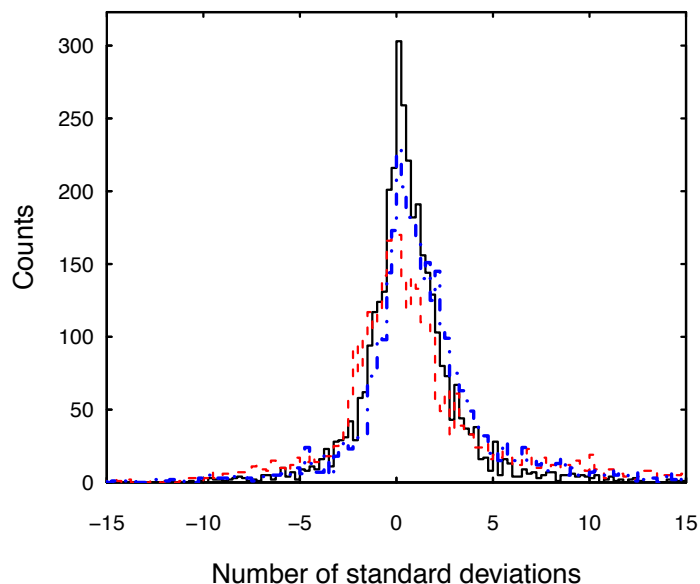
Comparison with experiment

> 4000 **differential cross section** measurements from the literature

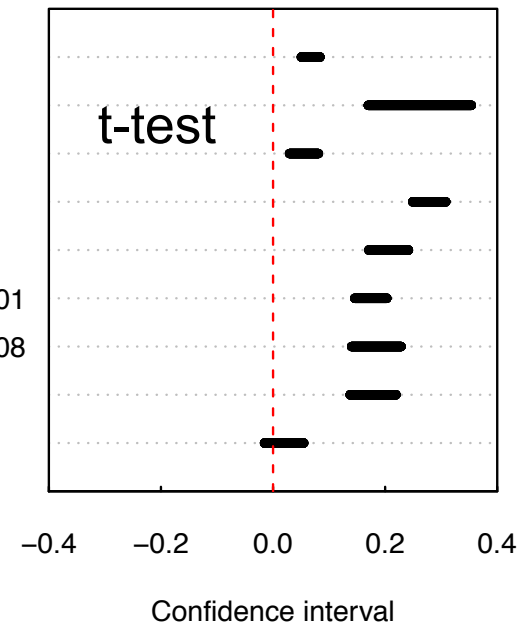
$$5.41 \text{ keV} \leq E \leq 39 \text{ MeV}$$

$$0.5^\circ \leq \theta \leq 165^\circ$$

69 target atoms, $3 \leq Z \leq 92$



MFASF
RFASF
MF
RF
NF
Penelope2001
Penelope2008
EPDL
SM

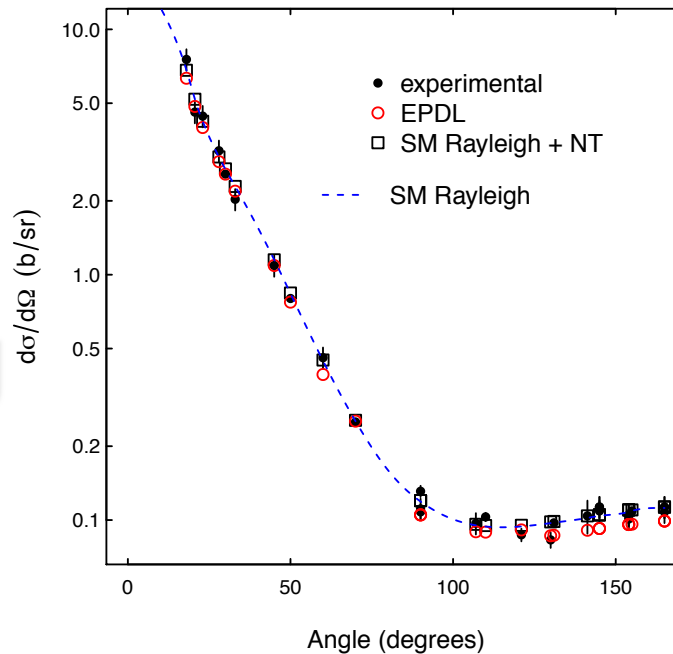


Statistical analysis of compatibility + categorical analysis

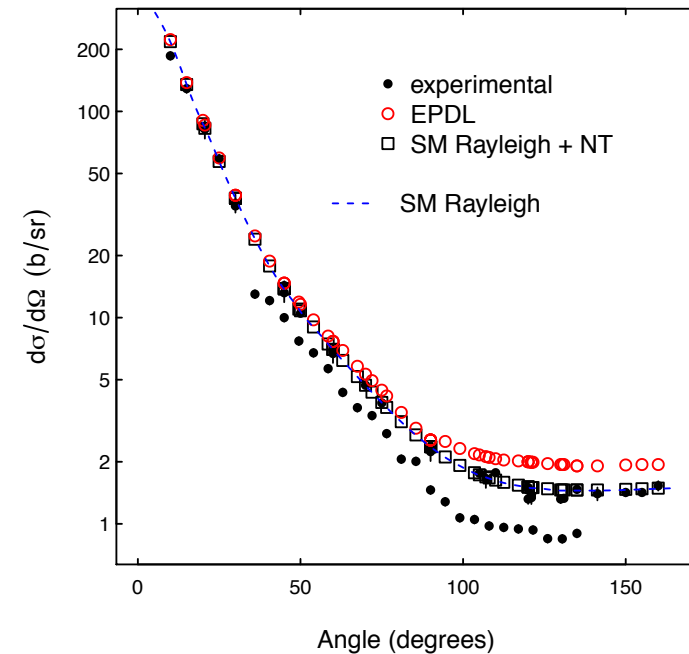
Nuclear Thomson scattering

Important to account for it especially at higher energy

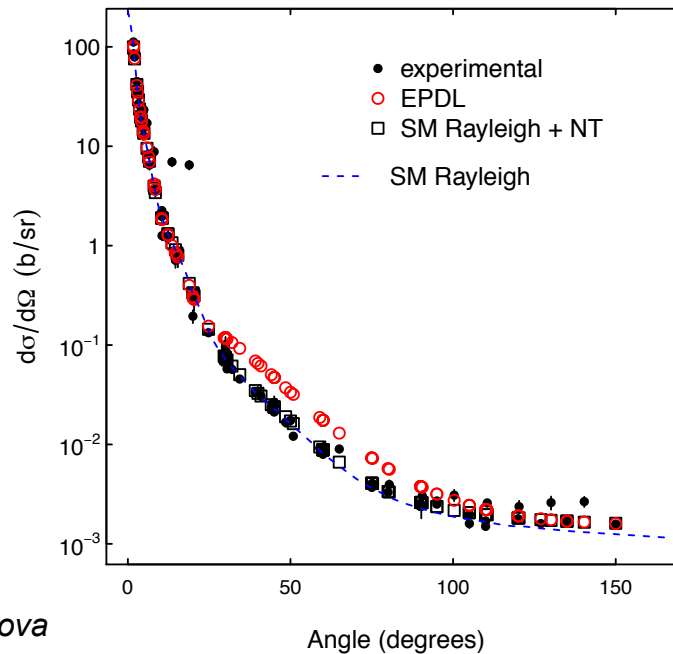
E=59.5 keV, Z=29



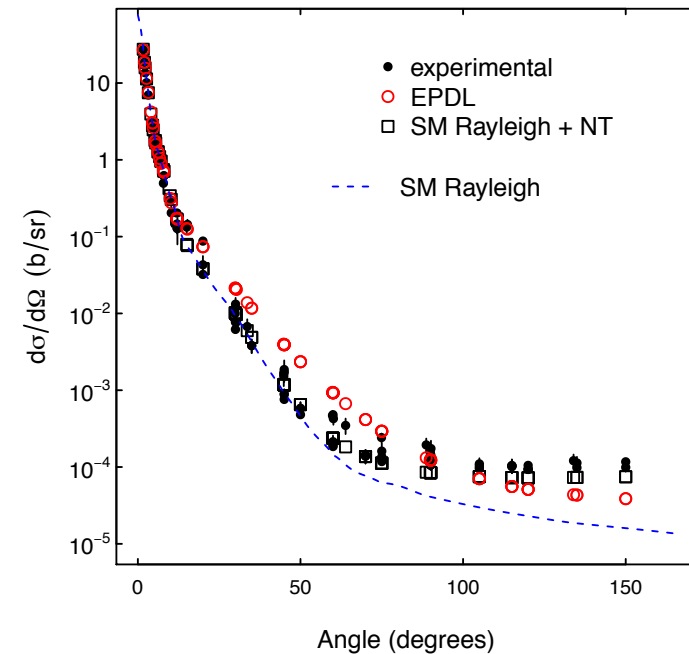
E=59.5 keV, Z=82



E=661.6 keV, Z=82



E=1.332 MeV, Z=82



Model validation

State of the art

Identified through quantitative validation w.r.t. experiment

Quantification

Compatibility with experiment: **GoF test**
(χ^2 test)

Comparison of models' accuracy: **categorical analysis**
(Fisher's exact test and χ^2 test)

Differential cross sections

	Penelope 2001	Penelope 2008	EPDL	Relativ. FF	Non-Rel. FF	Modified FF	MFF ASF	RFF ASF	SM NT
ϵ	0.27	0.38	0.38	0.25	0.35	0.49	0.52	0.48	0.77
error	± 0.05	± 0.06	± 0.06	± 0.05	± 0.06	± 0.06	± 0.06	± 0.06	± 0.05

ϵ = fraction of test cases compatible with experiment
0.01 significance in the χ^2 test

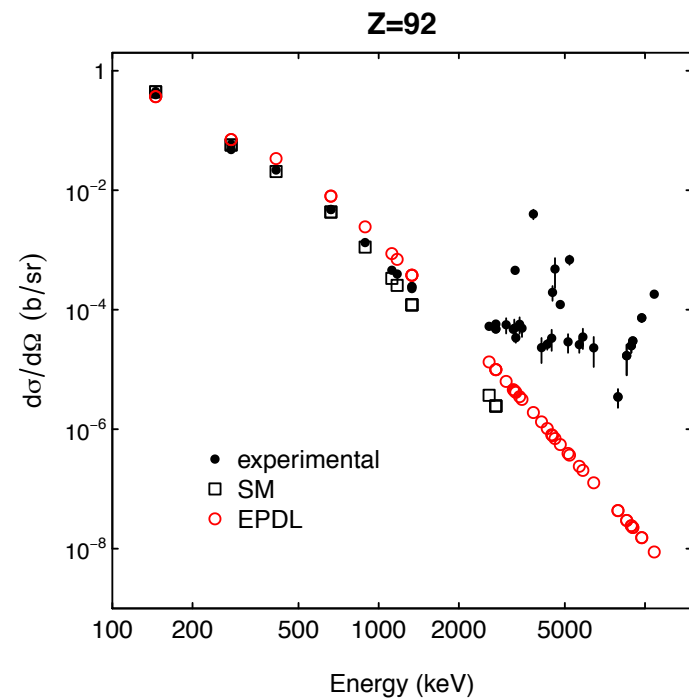
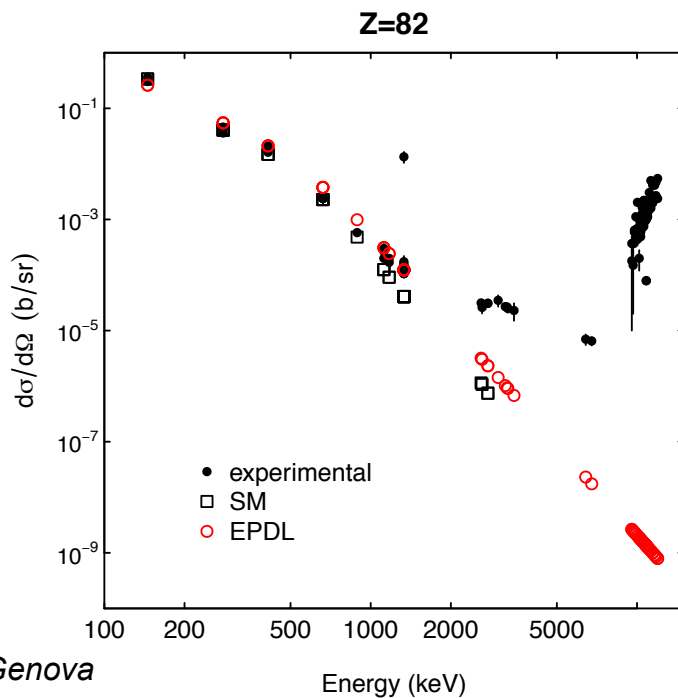
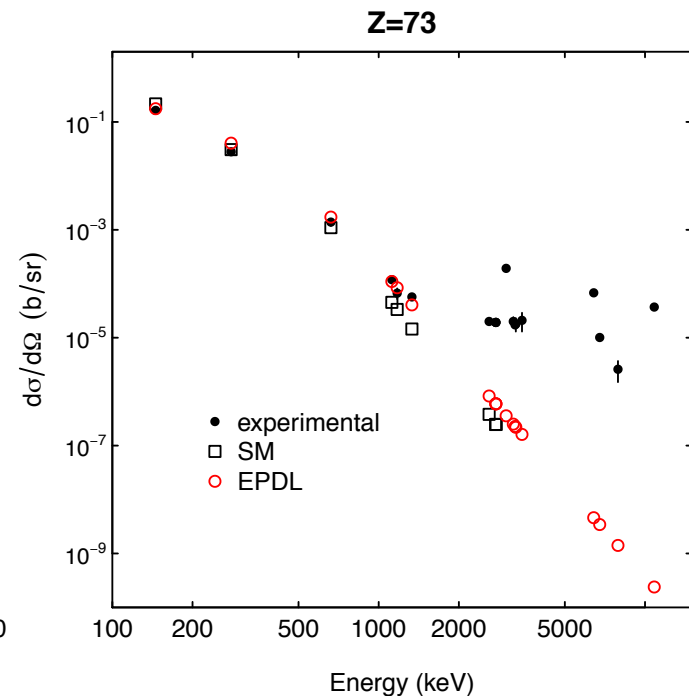
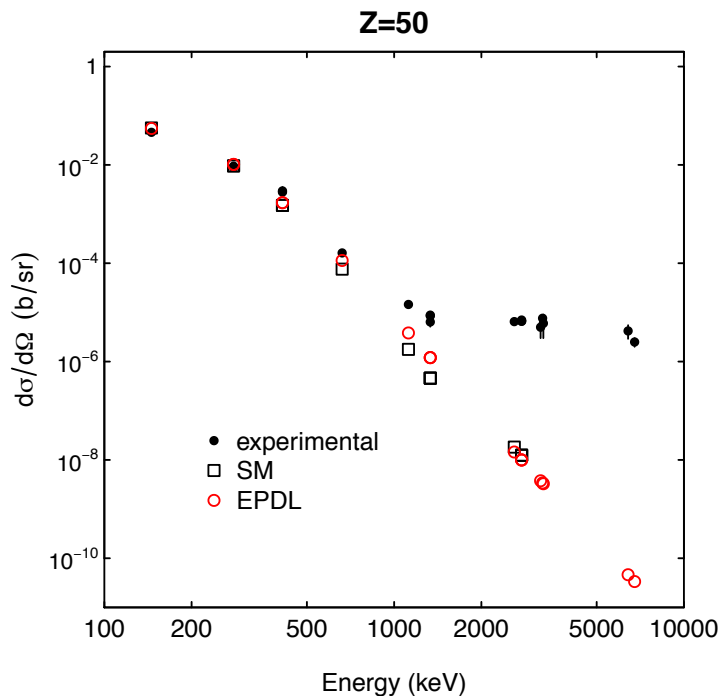


S-matrix

Modified form factors



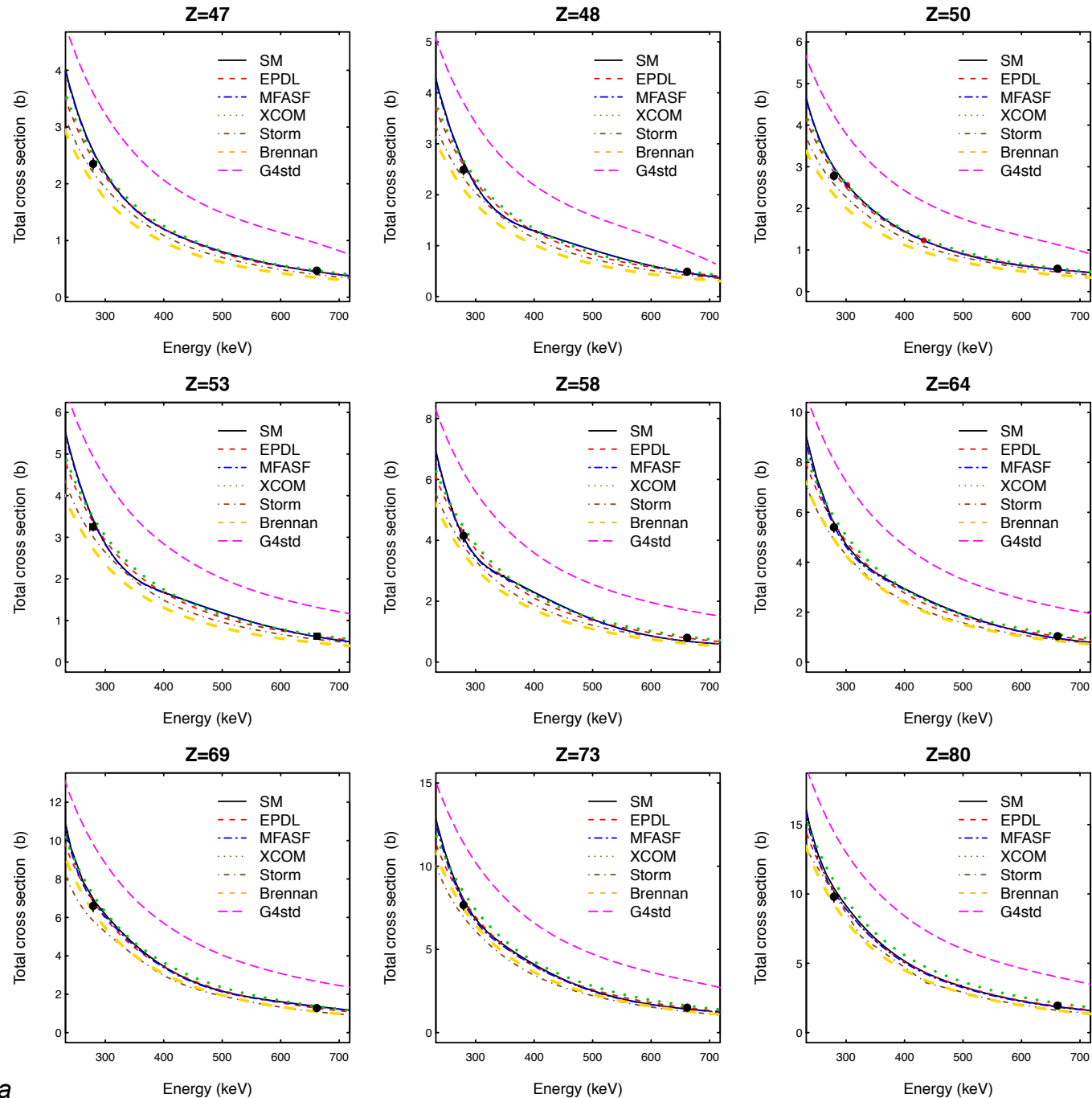
Other physics
to be
accounted for
at higher
energy



Total elastic cross section

- S-matrix
- EPDL
- MFASF
- XCOM
- Storm & Israel
- Brennan & Cowan
- Geant4 Standard

Scarce experimental data





Photoelectric effect

Total cross section

Shell cross sections

Photoelectron angular distribution

Total cross section

- EPDL
- Kissel (RTAB)
- Storm & Israel
- Biggs & Lighthill
- XCOM
- Brennan & Cowan
- McMaster

Geant 4

“Livermore”

EPDL

“Penelope 2008”

EPDL

“Penelope 2001”

EPDL

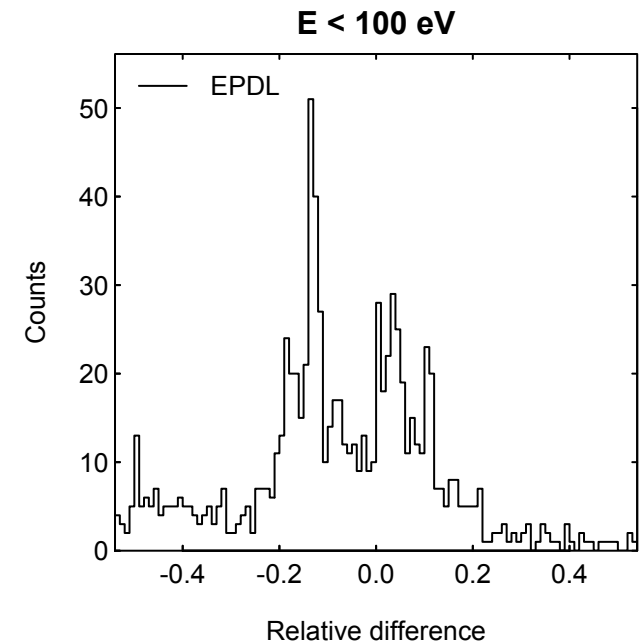
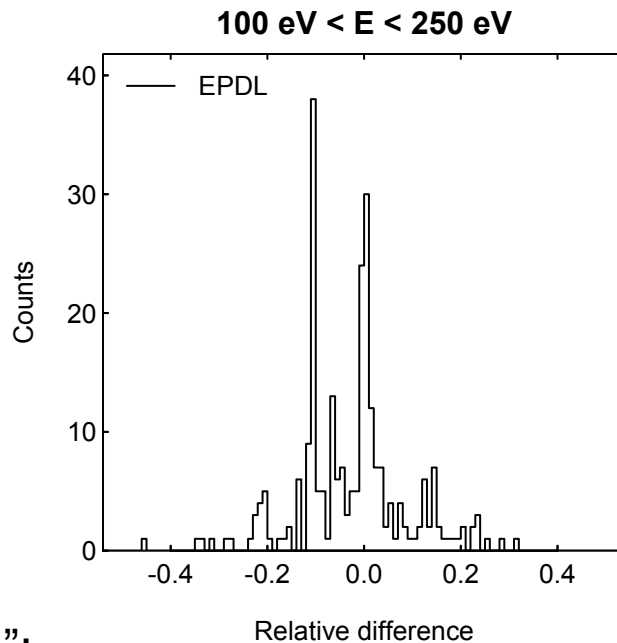
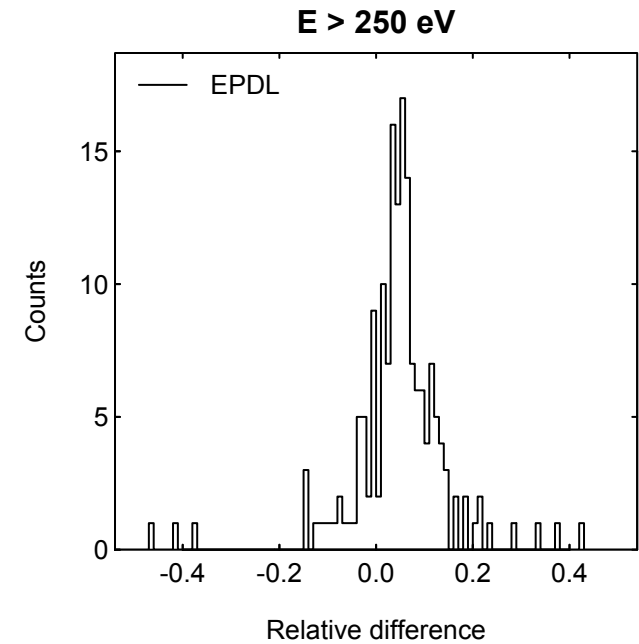
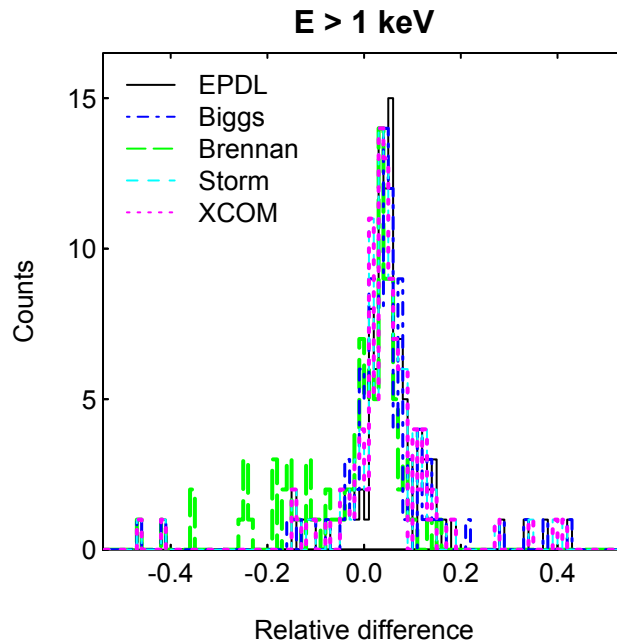
“Standard”

Biggs &
Lighthill

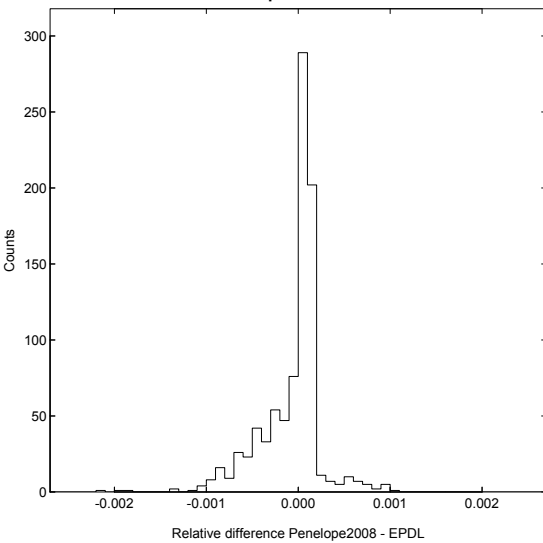
XCOM?

EPDL used down to **250 eV**
Never validated, nor optimized

Total cross section Difference w.r.t. experiment



Penelope2008 - EPDL



Geant4

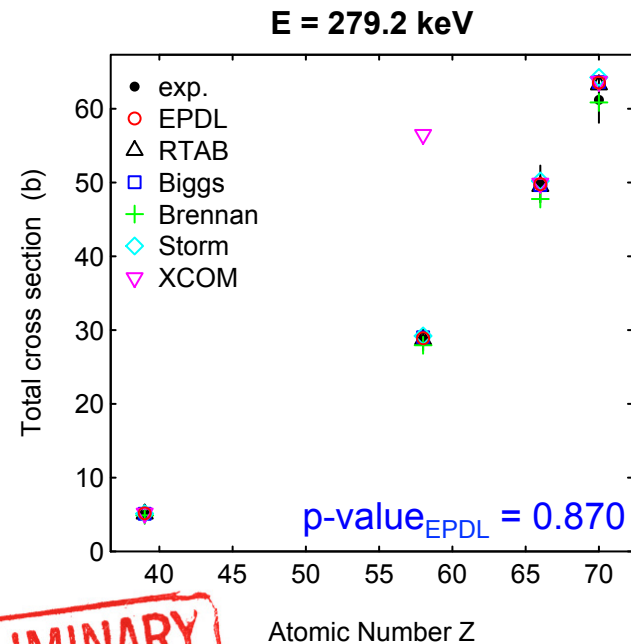
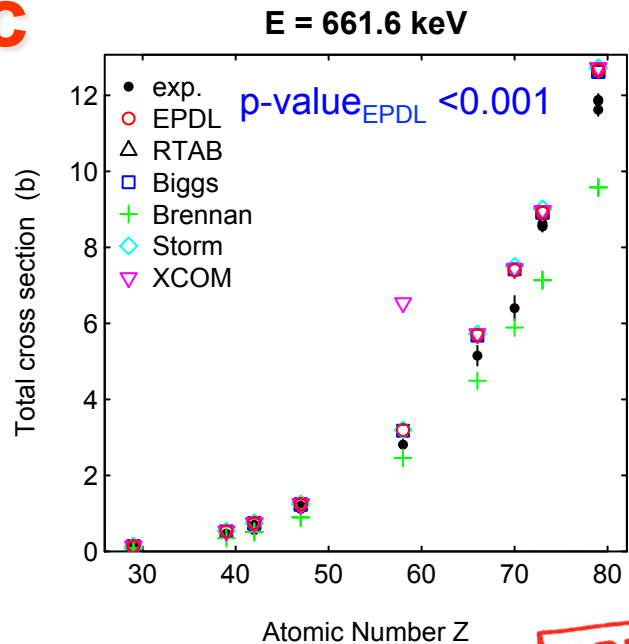
“Livermore” and “Penelope”:
both EPDL cross sections

Inadequate modeling or systematics in the measurements?

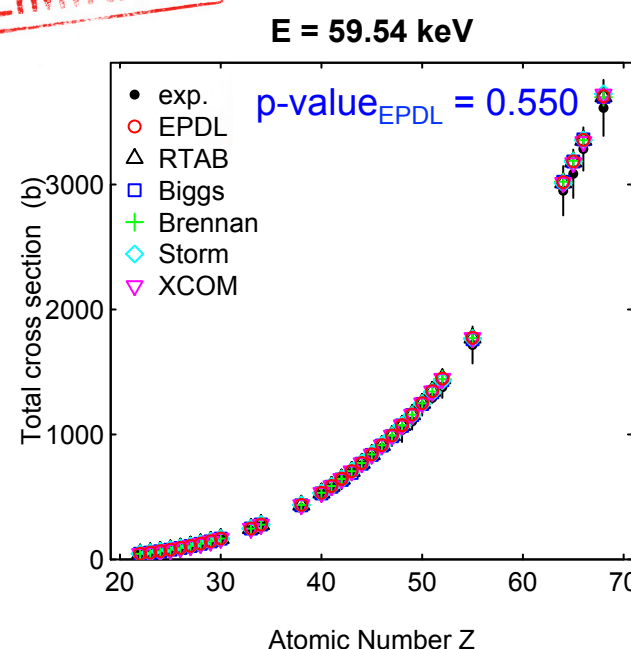
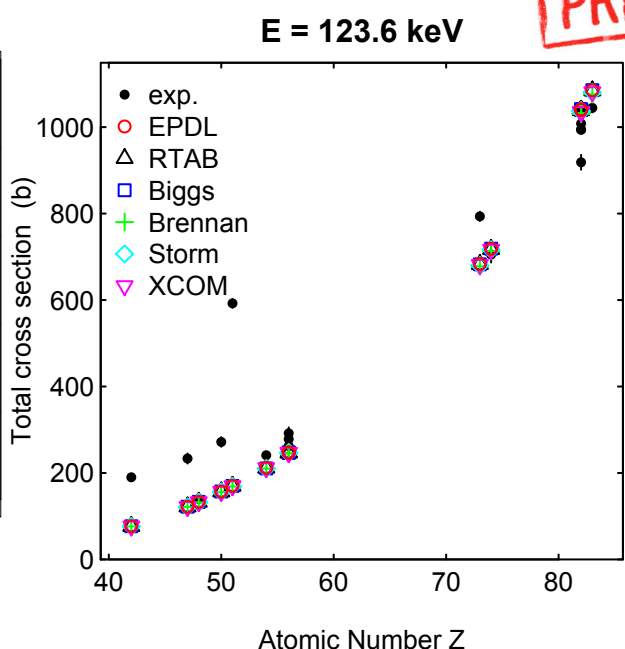
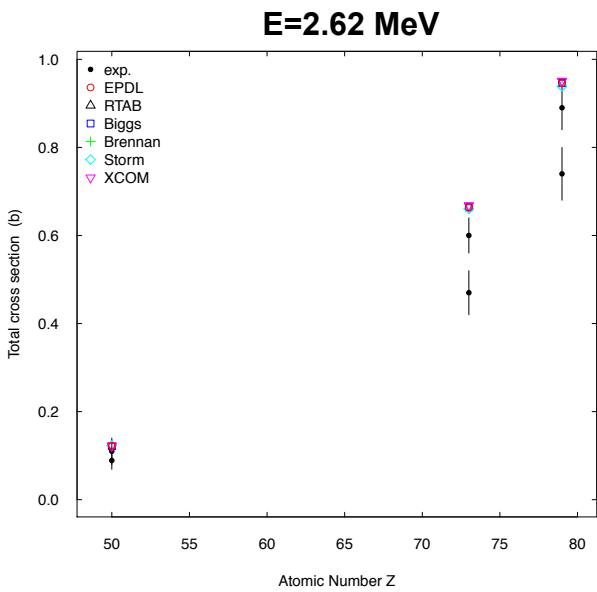
Photoelectric effect

Total cross section

$E > 50$ keV



PRELIMINARY



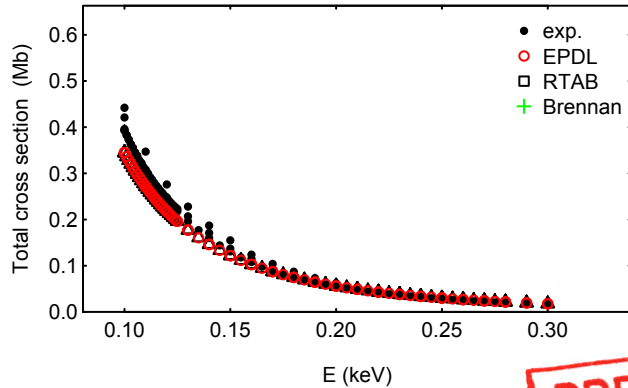
Photoelectric effect

Total cross section

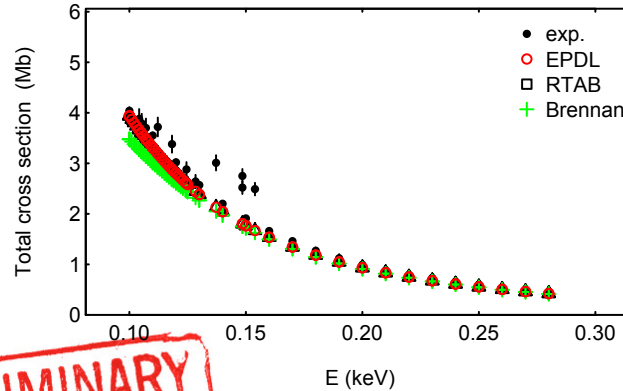
100 eV - 1 keV

Small experimental sample
Noble gases

Z = 2

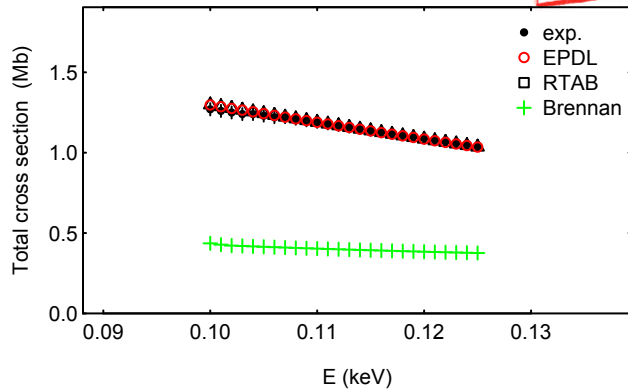


Z = 10

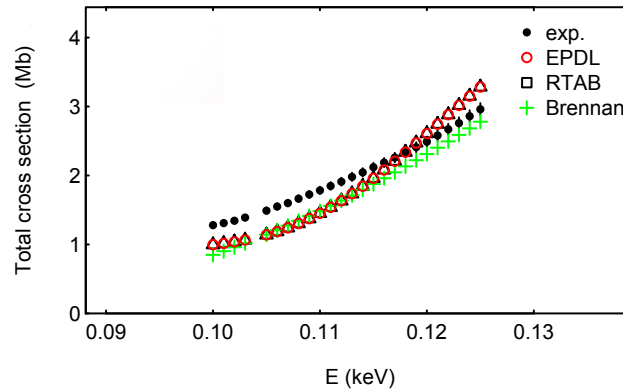


PRELIMINARY

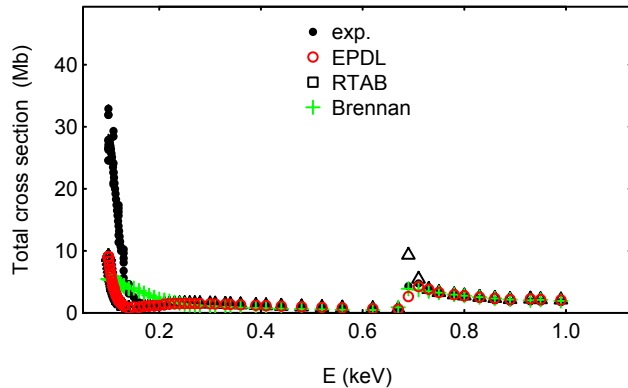
Z = 18



Z = 36



Z = 54



500 eV - 1 keV

p-value_{EPDL} = 0.593

250 - 500 eV

p-value_{EPDL} = 0.029

RTAB: similar to EPDL

Brennan-Cowan: less accurate

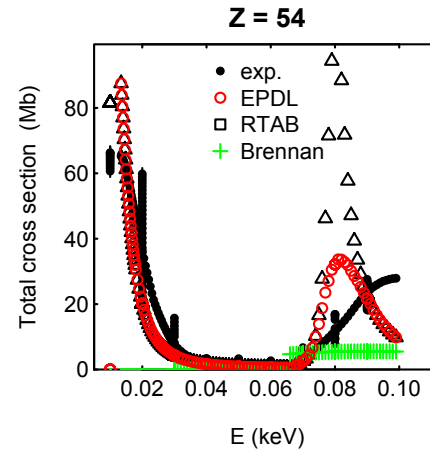
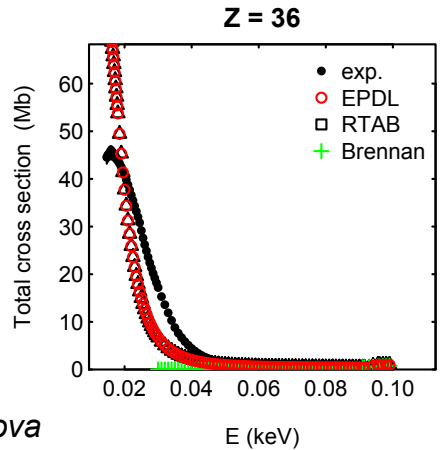
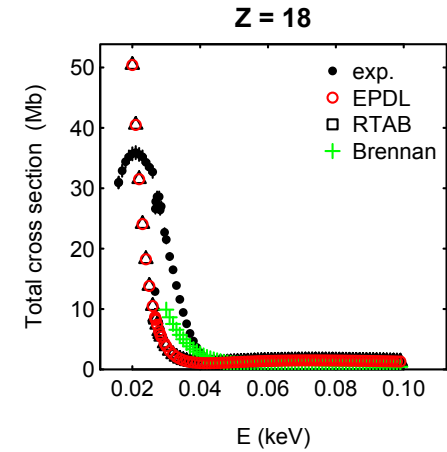
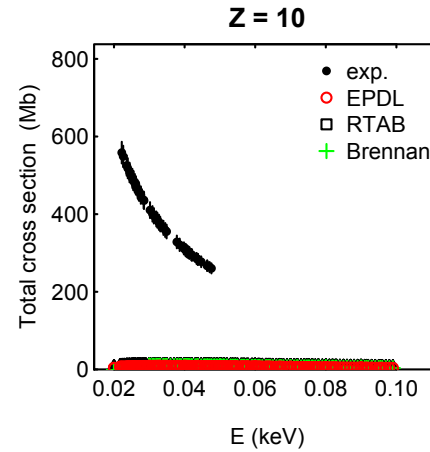
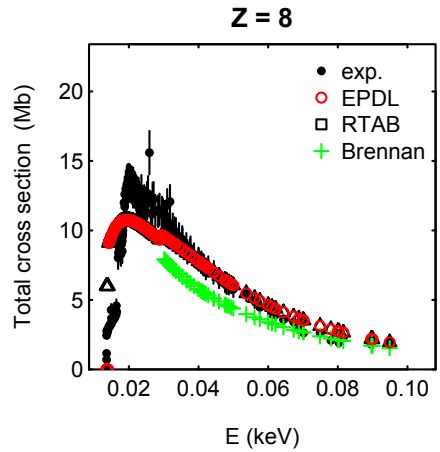
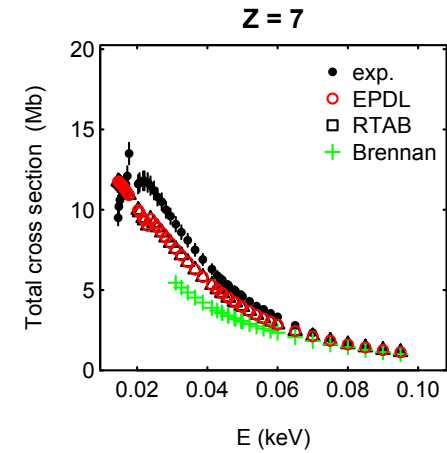
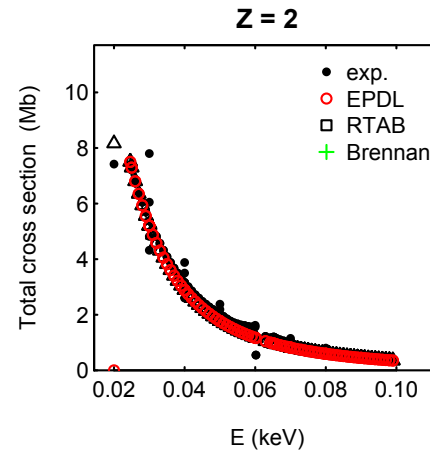
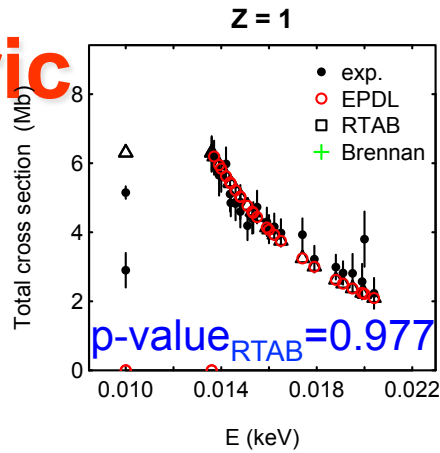
Photoelectric effect

PRELIMINARY

Total cross section

10-100 eV

Measured in gases only



Systematics in experiment?

Inadequate calculations?

Conclusions



A quantitatively validated, state-of-the-art
Monte Carlo code

a tremendous challenge
an inspiring motivation

Large effort

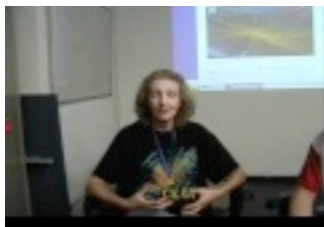
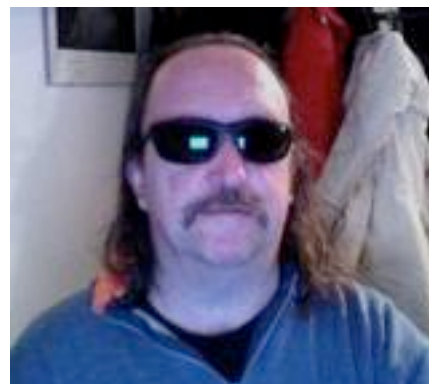
Supported by software design

Refactoring techniques and reengineering patterns contribute to **computational performance** and facilitate **validation**

Quantitative validation of physics models

more about it in the last talk of this session

Physics insight is the key



...a big **THANK YOU**
to the CERN Library!