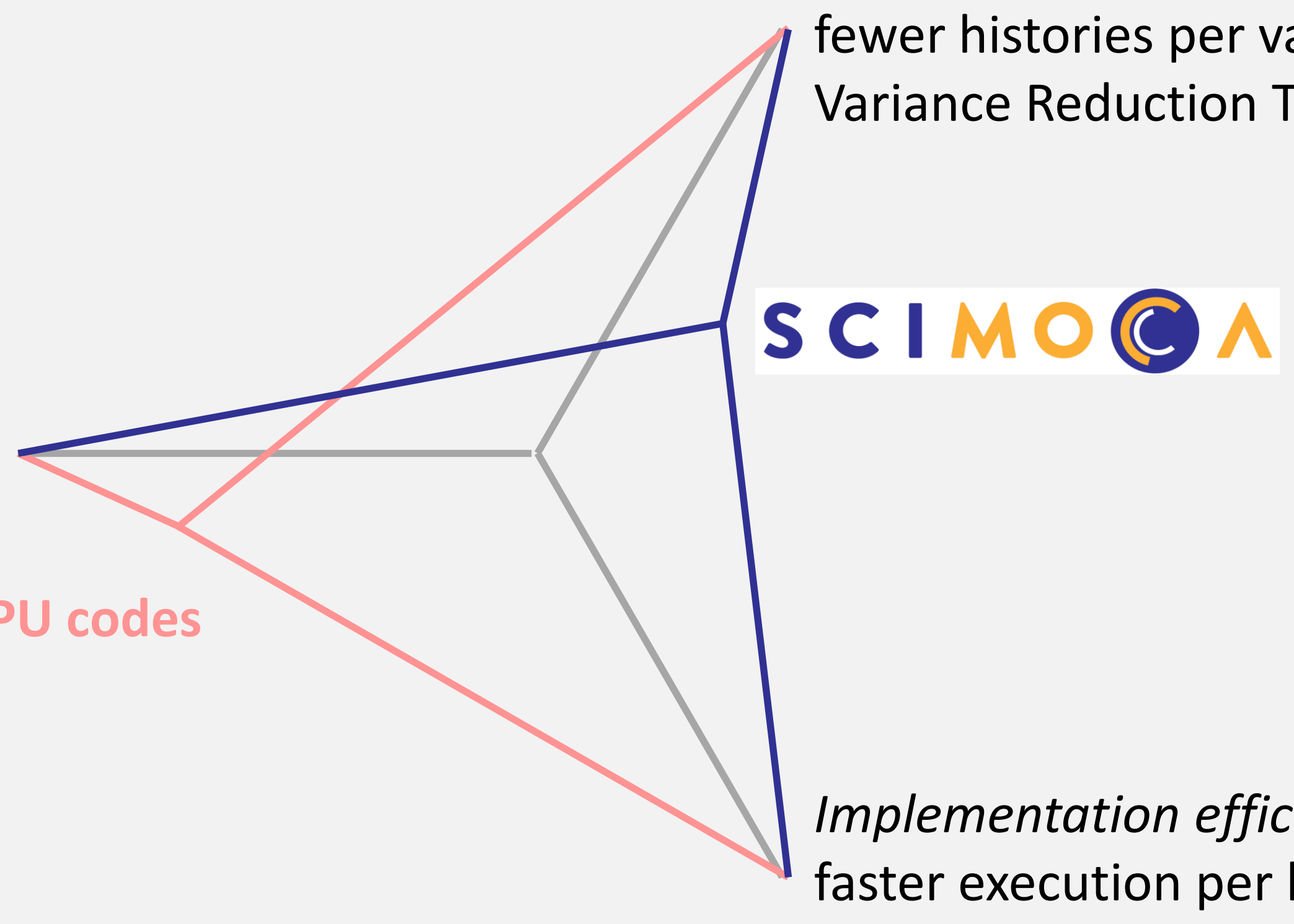
Towards real-time Monte Carlo dose computation: muscle or brain?

M. Alber[1,2] ,N. Saito[1,2], M. Söhn[2]

[1]Department of Radiation Oncology, Universitätsklinikum Heidelberg; [2]DScientific RT GmbH, Munich

How to speed up Monte Carlo?



Algorithm efficiency:
fewer histories per variance
Variance Reduction Techniques (VRT)

Hardware parallelization:
engage more processing units

GPU codes

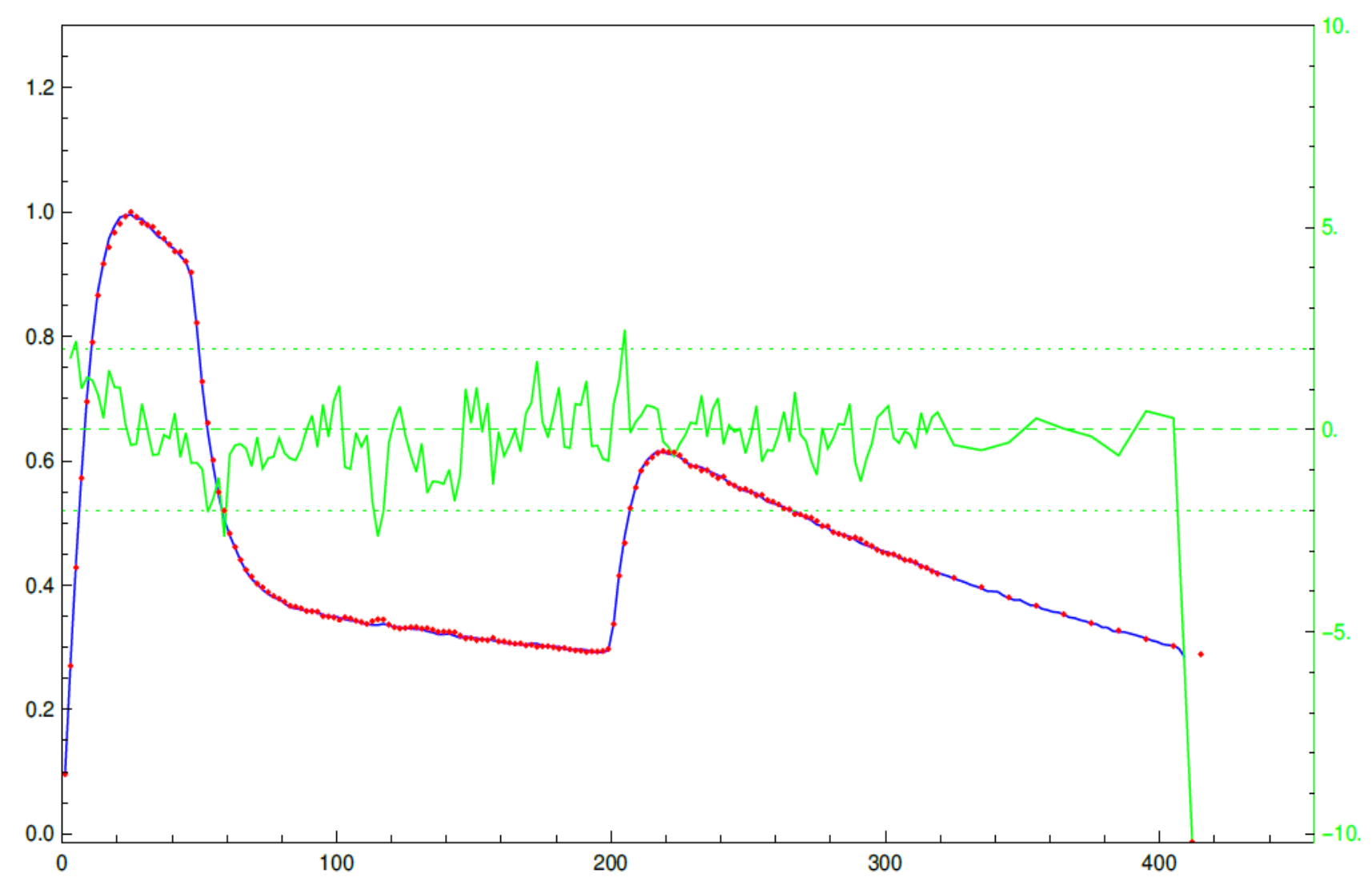
SCIMOCA

Implementation efficiency:
faster execution per history
hardware specific coding

Real-time Monte Carlo dose computation will soon be essential for planning and quality assurance of online-adapted treatment plans. By parallelization in GPUs (muscle) and CPUs (brain), this goal is in reach. However, muscle and brain need very specific code optimization for full performance.

Speed and uncompromised accuracy

10x10 mm² field, 6 MeV mono-energetic point source, 150 mm slab of ICRU-lung, 0.25 g/cm³ from 50 mm depth. Blue: SciMoCa, Red: EGSnrc.

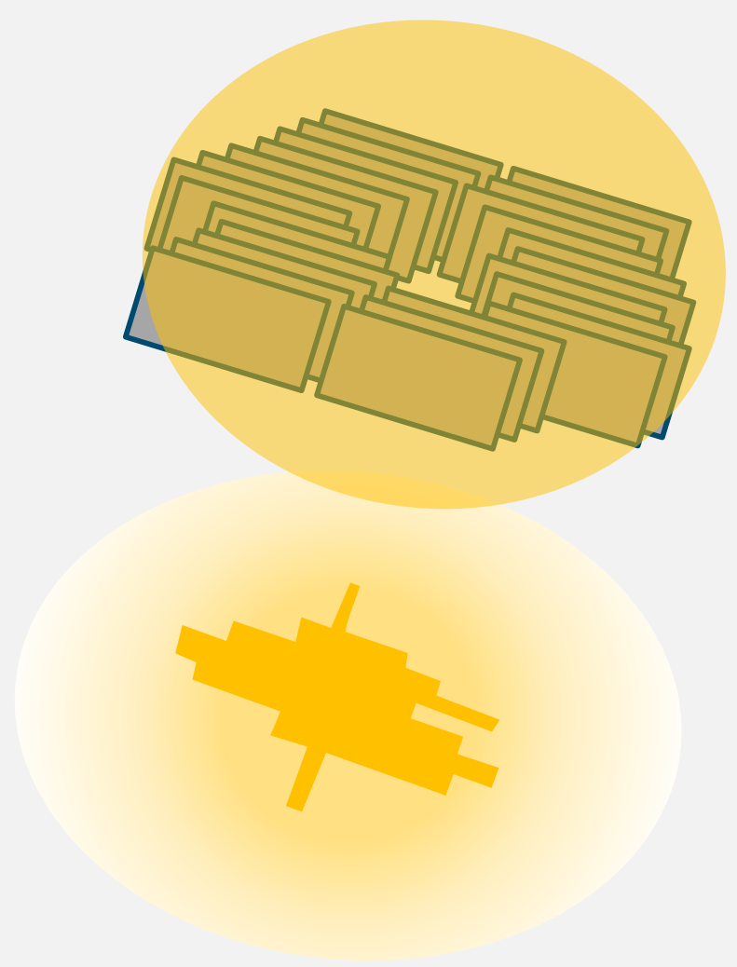


Clinical Monte Carlo:
The accelerator head is crucial for performance

Accelerator head simulations are inherently inefficient:

- complex geometries and diverse materials
- many absorbed particles and secondaries
- highly diverse linac designs challenge code optimization

Overall performance is driven by radiation source, collimator model and patient model.



Efficiency:
aperture / max field size

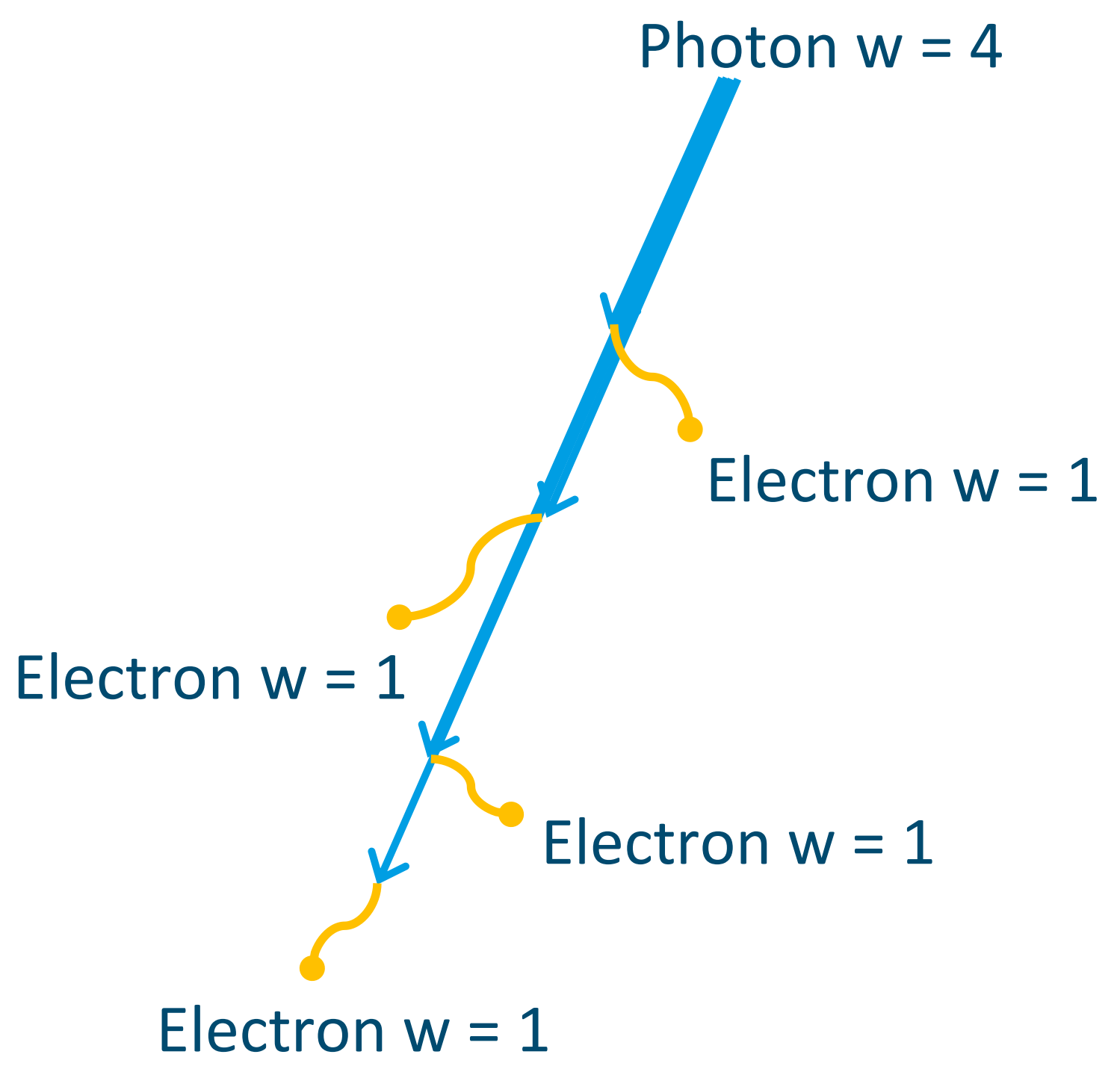
Complexity:
leaf shape, leakage, scatter

Variability:
hundreds of linac designs

SciMoCa supports all Varian, Elekta, and Siemens linacs, CyberKnife and Tomotherapy:

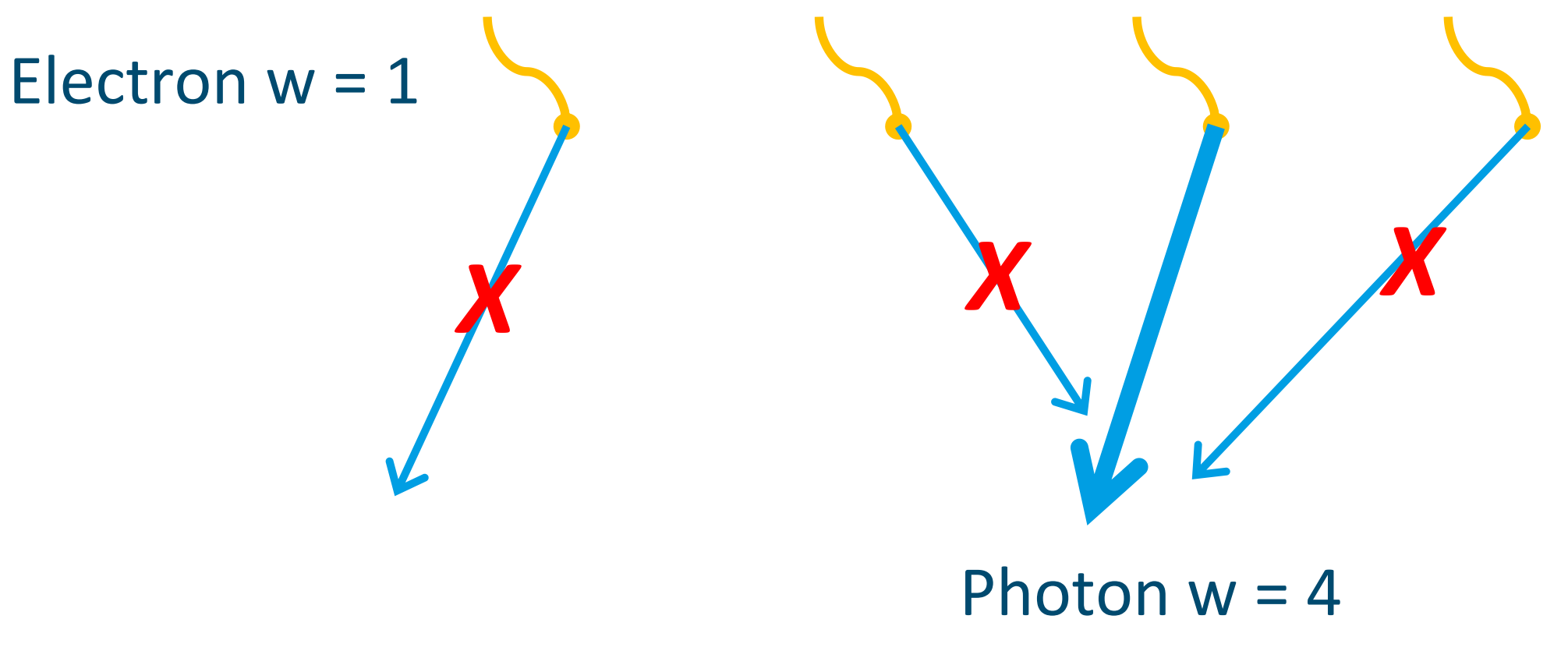
16 MLC Types
26 Beam qualities
41 Flattening Filter designs

Variance reduction techniques work brilliantly for accelerator heads

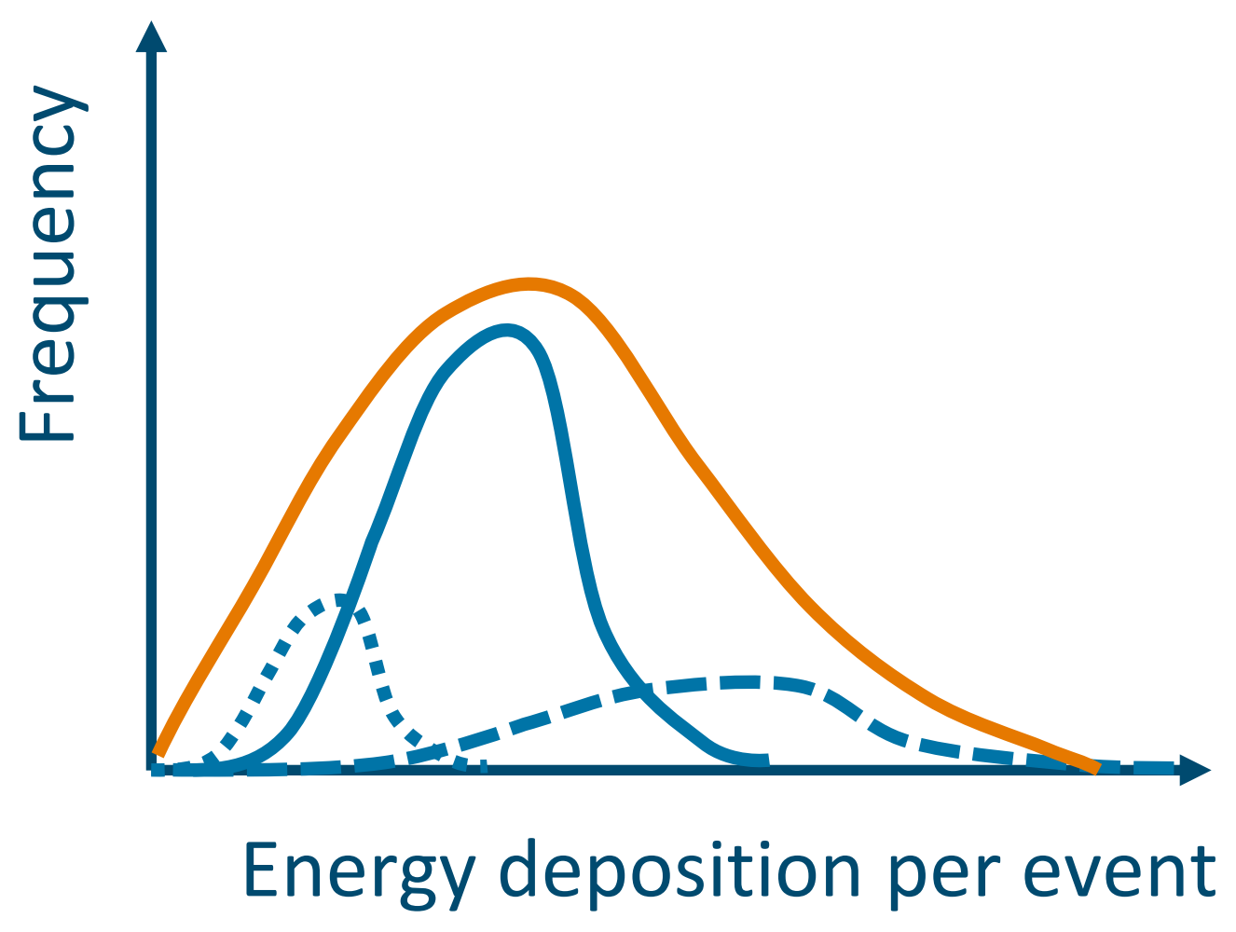


Variance reduction techniques utilize statistical particle weights to sample the interactions more efficiently:

- **particle splitting** and history repetition re-use sub-sets of a particle history to save repeat operations
  each split reduces particle weight
  Example: Photon traverses a leaf
- Russian Roulette discards some less important sub-sets of a particle history and gives higher weight to others
  each discard increases particle weight
  Example: Photon scatters in flattening filter



The cost of unbalanced particle weight manipulation: convergence efficiency drops



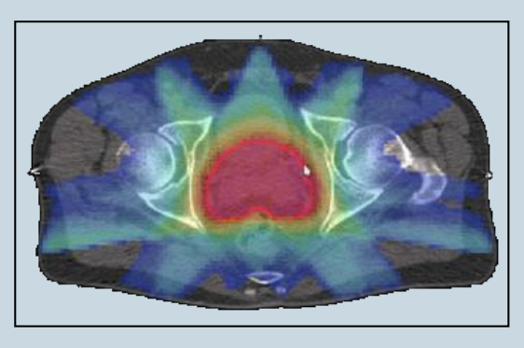
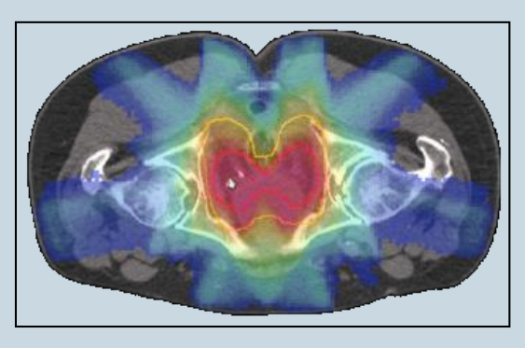
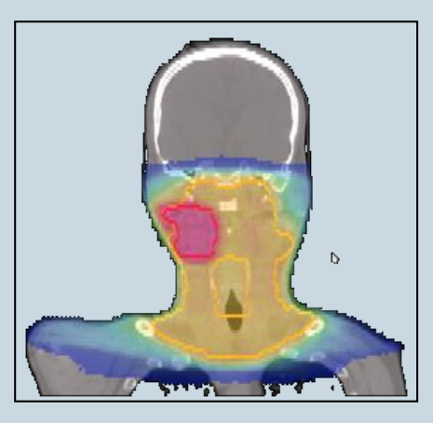
Energy deposition per event in a voxel (tally):
solid line: presumed distribution
dotted, dashed: for particle weights 0.5 and 2
orange: overall tally distribution following VRT

Voxel uncertainty = error of mean

Broadening the tally distribution requires more histories for the same uncertainty

VRT of source- and patient model need to be tuned optimally (case dependent)

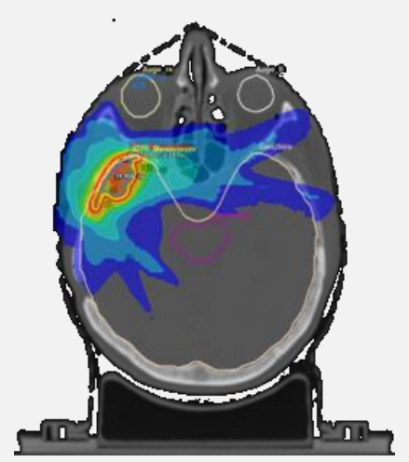
Balanced VRT: timings and scaling performance

	prostate, step & shoot (8 beams, 44 segments)	**prostate/LN, dMLC (7 beams, 140 control points)**	**head & neck, VMAT (2 arcs, 293 control points)**	
PTV volume	193.3 cc	SIB-case with 2 PTV volumes: 979.8 cc; 159.9 cc	SIB-case with 2 PTV volumes: 834.4 cc; 131.6 cc	
voxel size / uncertainty	3 mm / 1%	3 mm / 1%	3 mm / 1%	2 mm / 1%
calc time 16 cores	***15.8 sec***	***55.6 sec***	***40.9 sec***	***118.9 sec***
calc time 44 cores	***5.6 sec***	***18.2 sec***	***14.2 sec***	***39.3 sec***
				

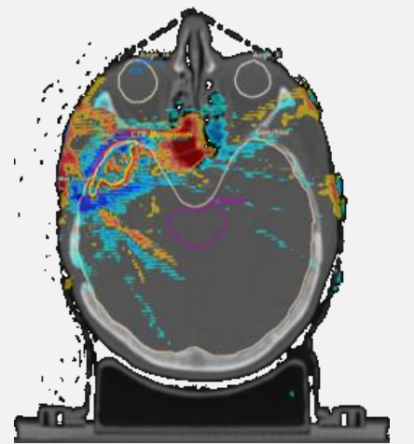
VRT employed in SciMoCa patient model:

Feature	Value/Reference	Similar to
electron cut-off energy for last Multiple Scatter step	< 240 keV	
fractional energy loss of electron Multiple Scatter step	0.12	
bremsstrahlung production cut-off energy	> 6 keV	
photon cut-off energy (local energy deposit)	< 60 keV	
minimum/maximum particle weight (Russian Roulette ratio)	0.5 < w < 2.0	
maximum photon energy	< 25 MeV	
KERMA-approximation threshold energy	< 1.0 MeV	
Material properties	ICRU 46	XVMC
Material property computation	Kawrakow 1996, Fippel 1999	VMC, XVMC, VMC++
Photon effects	Photoelectric absorption, Compton scatter, Pair production (Kawrakow 2000a)	XVMC, VMC++
Electron effects	Elastic scatter, Møller, Bremsstrahlung (Kawrakow 1996, 2000a)	XVMC, VMC++
Positron effects	Elastic scatter, Bhabha, Bremsstrahlung (Kawrakow 1996, 2000a)	XVMC, VMC++
Multiple Scatter theory	Kawrakow 2000b	EGSnrc, VMC++
Multiple Scatter boundary crossing	Kawrakow 1997, 2001	XVMC, VMC++
Variance reduction techniques	Woodcock tracking, adaptive history repetition, adaptive particle splitting, Russian Roulette, KERMA-approximation	XVMC, VMC++

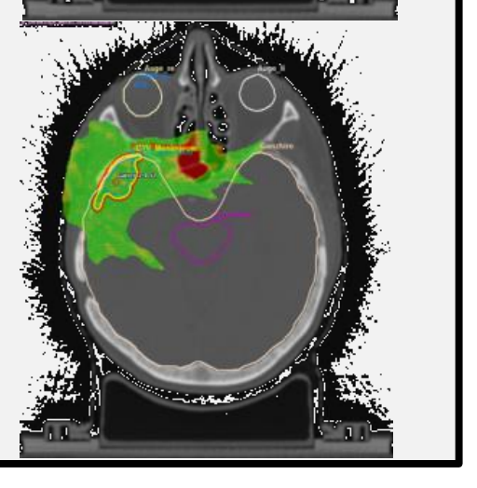
CyberKnife
Incise2
SciMoCa



Difference
Multiplan
SciMoCa



Gamma
1%/1mm



Fippel 1999: M. Fippel: Med. Phys. 26, 1466 (1999)
Kawrakow 1996: I. Kawrakow, M. Fippel, K. Friedrich: Med. Phys. 23, 445 (1996)
Kawrakow 1997: I. Kawrakow: Med. Phys. 24, 505 (1997)
Kawrakow 2000a: I. Kawrakow, M. Fippel: Phys. Med. Biol. 45, 2163 (2000)
Kawrakow 2000b: I. Kawrakow: Med. Phys. 27, 485 (2000)
Kawrakow 2001: I. Kawrakow in: A. Kling et al. (eds.), *Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications,* Springer-Verlag Berlin Heidelberg (2001)

Conclusions

Source and collimator simulation increases the complexity of MC:
advantage CPU
VRT tuning causes thread divergence:
advantage CPU
High computational load, low memory access:
high scalability on CPU
Real-time MC is rapidly becoming reality both with GPU and CPU.