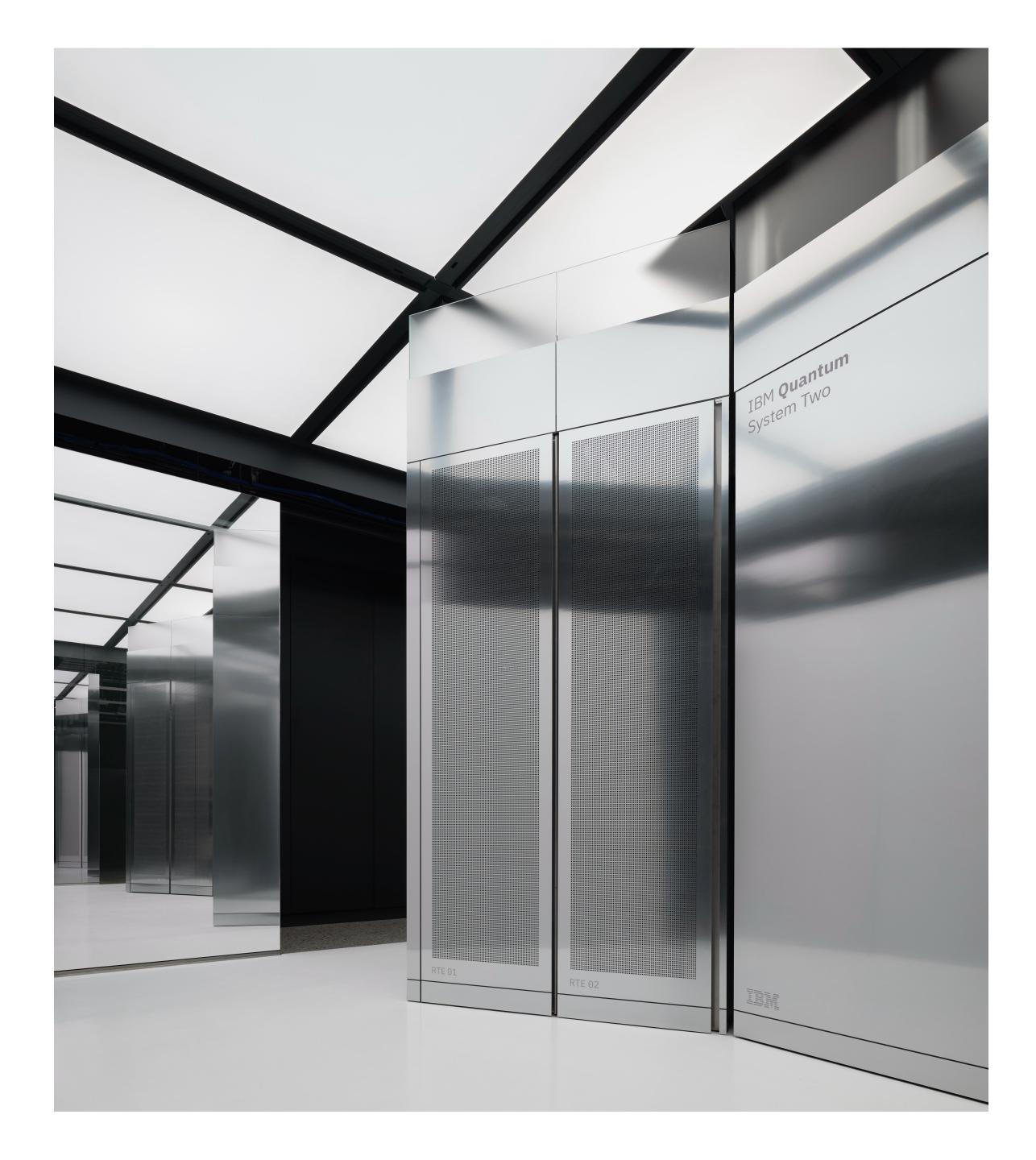
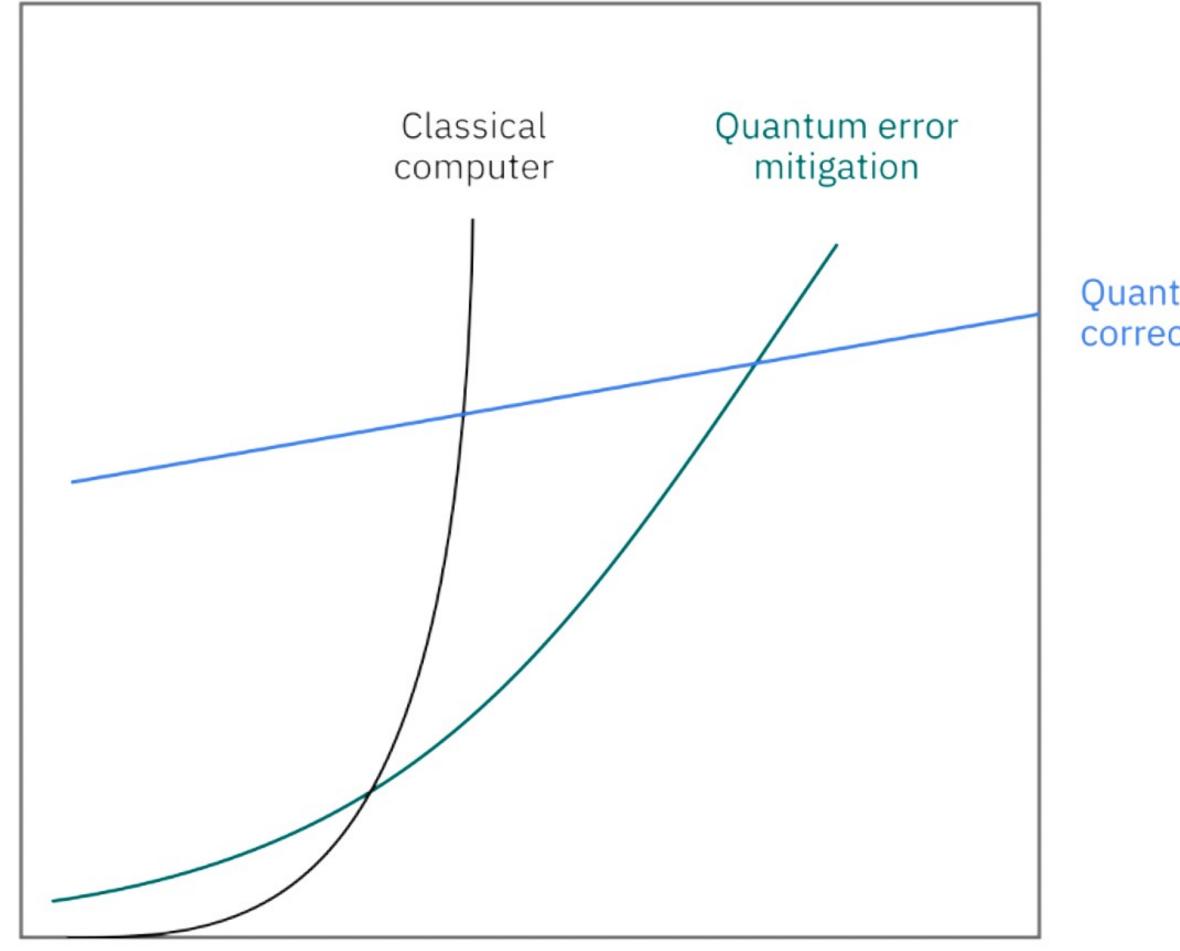
Quantum-centric Supercomputing for physics research

Vincent R. Pascuzzi Quantum Compilation Engineer IBM Quantum





Computational scaling of complex calculations

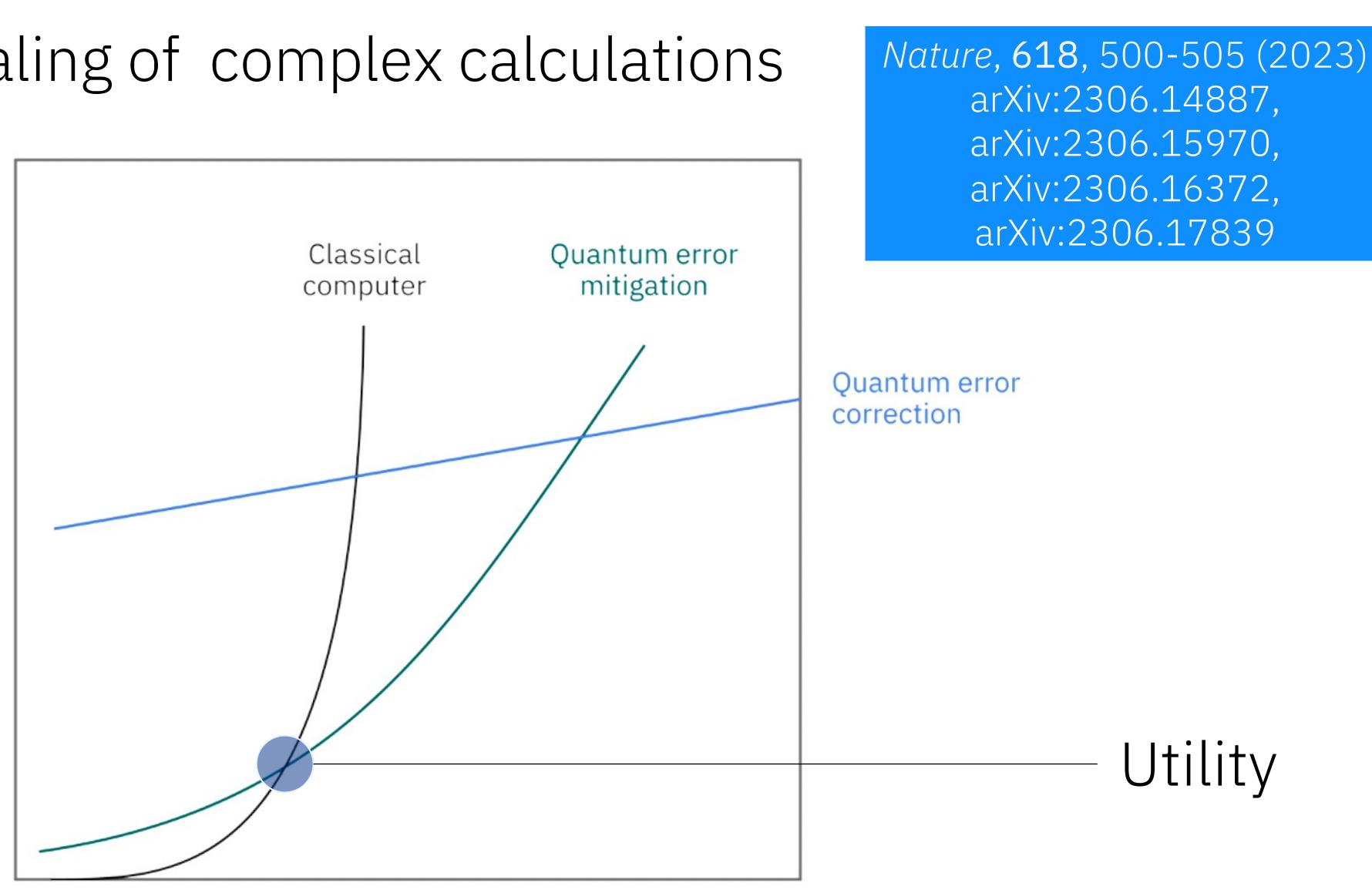


Quantum circuit complexity

Simulation cost

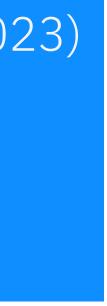
Quantum error correction

Computational scaling of complex calculations



Simulation cost

Quantum circuit complexity

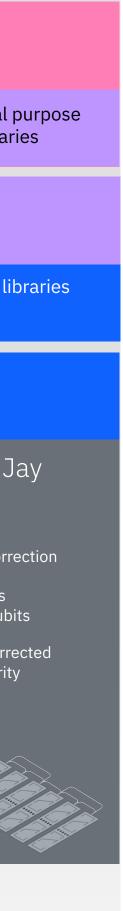


Development roadmap

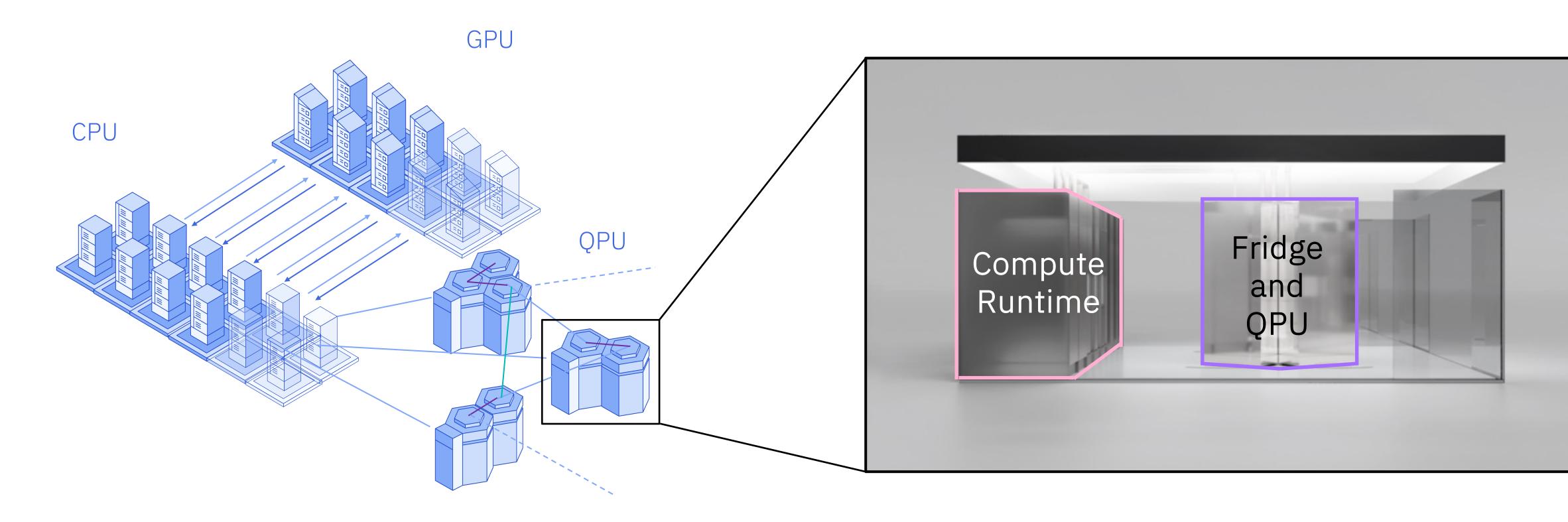
		2016-2019 🖌	2020 🥏	2021 🥪	2022 🗹	2023 🧹	2024	2025	2026	2027	2028	2029	2033
		Run quantum circuits on the IBM Quantum Platform	Release multi- dimensional roadmap publicly with initial aim focused on scaling	Enhancing quantum execution speed by 100x with Qiskit Runtime	Bring dynamic circuits to unlock more computations	Enhancing quantum execution speed by 5x with quantum serverless and Execution modes	Improving quantum circuit quality and speed to allow 5K gates with parametric circuits	Enhancing quantum execution speed and parallelization with partitioning and quantum modularity	Improving quantum circuit quality to allow 7.5K gates	Improving quantum circuit quality to allow 10K gates	Improving quantum circuit quality to allow 15K gates	Improving quantum circuit quality to allow 100M gates	Beyond 203 centric supe will include : logical qubit the full powe quantum co
	Data Scientists						Platform						
							Code 👌	Functions	Mapping Collections	Specific Libraries			General pu QC librarie
	Researchers					Middleware							
						Quantum Serverless	Transpiler 🕑 Service	Resource Management	Circuit Knitting x P	Intelligent Orchestration			Circuit libr
	Quantum Physicists	Qiskit Runtime			ime								
		IBM Quantum Experience	S	QASM3 📀	Dynamic <	Execution <	Heron う (5K)	Flamingo (5K)	Flamingo (7.5K)	Flamingo (10K)	Flamingo (15K)	Starling (100M)	Blue Ja (1B)
		Early 📀	Falcon	S	Eagle	S	Error Mitigation	Error Mitigation	Error Mitigation	Error Mitigation	Error Mitigation	Error Correction	Error Correc
								<u> </u>	<u> </u>				
		Canary 5 gubits	Benchmarking		Benchmarking		5k gates 133 qubits	5k gates 156 qubits	7.5k gates 156 qubits	10k gates 156 qubits	15k gates 156 qubits	100M gates 200 qubits	1B gates 2000 qubits
		5 qubits Albatross	Benchmarking 27 qubits		Benchmarking 127 qubits		5k gates	5k gates 156 qubits Quantum modular	7.5k gates 156 qubits Quantum modular	156 qubits Quantum modular	15k gates 156 qubits Quantum modular	100M gates	1B gates 2000 qubits Error correc modularity
		5 qubits Albatross 16 qubits					5k gates 133 qubits	5k gates 156 qubits	7.5k gates 156 qubits	156 qubits	15k gates 156 qubits	100M gates 200 qubits Error corrected	2000 qubits Error correc
	Executed by IBM On target	5 qubits Albatross					5k gates 133 qubits Classical modular Up to 133x3 =	5k gates 156 qubits Quantum modular Up to 156x7 =	7.5k gates 156 qubits Quantum modular Up to 156x7 =	156 qubits Quantum modular Up to 156x7 =	15k gates 156 qubits Quantum modular Up to 156x7 =	100M gates 200 qubits Error corrected	2000 qubits Error correc

3+

2033, quantumsupercomputers ude 1000's of ubits unlocking power of n computing



Quantum-centric Supercomputing (QCSC)



Delivering impactful quantum computing requires the interplay of quantum and classical resources at scale: HPC-assisted quantum computation to extract/boost useful signals in utility-scale experiments.



Quantum-centric Supercomputing (QCSC)

Quantum Computing for High-Energy Physics State of the Art and Challenges Summary of the QC4HEP Working Group

Alberto Di Meglio,^{1,*} Karl Jansen,^{2,3,†} Ivano Tavernelli,^{4,‡} Constantia Alexandrou,^{5,3} Srinivasan Arunachalam,⁶ Christian W. Bauer,⁷ Kerstin Borras,^{8,9} Stefano Carrazza,^{10,1} Arianna Crippa,^{2,11} Vincent Croft,¹² Roland de Putter,⁶ Andrea Delgado,¹³ Vedran Dunjko,¹² Daniel J. Egger,⁴ Elias Fernández-Combarro,¹⁴ Elina Fuchs,^{1,15,16} Lena Funcke,¹⁷ Daniel González-Cuadra,^{18,19} Michele Grossi,¹ Jad C. Halimeh,^{20, 21} Zoë Holmes,²² Stefan Kühn,² Denis Lacroix,²³ Randy Lewis,²⁴ Donatella Lucchesi,^{25, 26, 1} Miriam Lucio Martinez,^{27, 28} Federico Meloni,⁸ Antonio Mezzacapo,⁶ Simone Montangero,^{25, 26} Lento Nagano,²⁹ Vincent R. Pascuzzi,⁶ Voica Radescu,³⁰ Enrique Rico Ortega,^{31, 32, 33, 34} Alessandro Roggero,^{35,36} Julian Schuhmacher,⁴ Joao Seixas,^{37,38,39} Pietro Silvi,^{25,26} Panagiotis Spentzouris,⁴⁰ Francesco Tacchino,⁴ Kristan Temme,⁶ Koji Terashi,²⁹ Jordi Tura,^{12,41} Cenk Tüysüz,^{2,11} Sofia Vallecorsa,¹ Uwe-Jens Wiese,⁴² Shinjae Yoo,⁴³ and Jinglei Zhang^{44,45}

arXiv:2307.03236

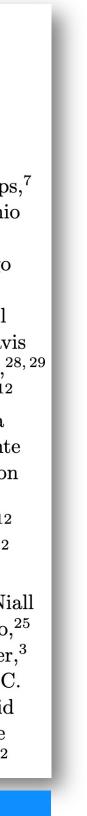
See also: arXiv:2312.09733 arXiv:2312.05344

Quantum-centric Supercomputing for Materials Science: A Perspective on Challenges and Future Directions

Yuri Alexeev,¹ Maximilian Amsler,² Marco Antonio Barroca,^{3,4} Sanzio Bassini,⁵ Torey Battelle,⁶ Daan Camps,⁷ David Casanova,^{8,9} Young jai Choi,¹⁰ Frederic T. Chong,¹¹ Charles Chung,¹² Christopher Codella,¹² Antonio D. Córcoles,¹² James Cruise,¹³ Alberto Di Meglio,¹⁴ Ivan Duran,¹² Thomas Eckl,² Sophia Economou,¹⁵ Stephan Eidenbenz,¹⁶ Bruce Elmegreen,¹² Clyde Fare,¹² Ismael Faro,¹² Cristina Sanz Fernández,¹² Rodrigo Neumann Barros Ferreira,³ Keisuke Fuji,¹⁷ Bryce Fuller,¹² Laura Gagliardi,^{18,1} Giulia Galli,^{11,1} Jennifer R. Glick,¹² Isacco Gobbi,¹⁹ Pranav Gokhale,²⁰ Salvador de la Puente Gonzalez,¹² Johannes Greiner,¹² Bill Gropp,²¹ Michele Grossi,¹⁴ Emmanuel Gull,²² Burns Healy,²³ Matthew R. Hermes,¹⁸ Benchen Huang,¹¹ Travis S. Humble,²⁴ Nobuyasu Ito,²⁵ Artur F. Izmaylov,^{26,27} Ali Javadi-Abhari,¹² Douglas Jennewein,⁶ Shantenu Jha,^{28,29} Liang Jiang,¹¹ Barbara Jones,³⁰ Wibe Albert de Jong,³¹ Petar Jurcevic,¹² William Kirby,³² Stefan Kister,¹² Masahiro Kitagawa,¹⁷ Joel Klassen,³³ Katherine Klymko,³⁴ Kwangwon Koh,³⁵ Masaaki Kondo,^{36,25} Doğa Murat Kürkçüoğlu,^{37, 38} Krzysztof Kurowski,³⁹ Teodoro Laino,⁴⁰ Ryan Landfield,²⁴ Matt Leininger,⁴¹ Vicente Leyton-Ortega,²⁴ Ang Li,^{42,43} Meifeng Lin,²⁸ Junyu Liu,^{11,44,45} Nicolas Lorente,^{46,8} Andre Luckow,⁴⁷ Simon Martiel,¹² Francisco Martin-Fernandez,¹² Margaret Martonosi,⁴⁸ Claire Marvinney,²⁴ Arcesio Castaneda Medina,¹⁹ Dirk Merten,¹⁹ Antonio Mezzacapo,^{12, *} Kristel Michielsen,⁴⁹ Abhishek Mitra,¹⁸ Tushar Mittal,¹² Kyungsun Moon,¹⁰ Joel Moore,^{50, 51} Sarah Mostame,¹² Mario Motta,¹² Young-Hye Na,¹² Yunseong Nam,¹² Prineha Narang,^{28,52} Yu-ya Ohnishi,⁵³ Daniele Ottaviani,⁵ Matthew Otten,⁵⁴ Scott Pakin,¹⁶ Vincent R. Pascuzzi,¹² Edwin Pednault,¹² Tomasz Piontek,³⁹ Jed Pitera,⁵⁵ Patrick Rall,³² Gokul Subramanian Ravi,²² Niall Robertson,⁵⁶ Matteo A. C. Rossi,⁵⁷ Piotr Rydlichowski,³⁹ Hoon Ryu,⁵⁸ Georgy Samsonidze,⁵⁹ Mitsuhisa Sato,²⁵ Nishant Saurabh,⁶⁰ Vidushi Sharma,¹² Kunal Sharma,¹² Soyoung Shin,⁶¹ George Slessman,⁶² Mathias Steiner,³ Iskandar Sitdikov,¹² In-Saeng Suh,²⁴ Eric D. Switzer,⁸ Wei Tang,⁴⁸ Joel Thompson,⁶³ Synge Todo,⁶⁴ Minh C. Tran,³² Dimitar Trenev,⁶⁵ Christian Trott,⁶⁶ Huan-Hsin Tseng,²⁸ Norm M. Tubman,⁶⁷ Esin Tureci,⁴⁸ David García Valiñas,¹² Sofia Vallecorsa,¹⁴ Christopher Wever,⁶⁸ Konrad Wojciechowski,³⁹ Xiaodi Wu,⁶⁹ Shinjae Yoo,²⁸ Nobuyuki Yoshioka,⁶⁴ Victor Wen-zhe Yu,¹ Seiji Yunoki,^{25,70} Sergiy Zhuk,⁵⁶ and Dmitry Zubarev¹²

arXiv:2312.09733







Quantum-centric Supercomputing (QCSC)

Quantum Computing for High-Energy Physics State of the Art and Challenges Summary of the QC4HEP Working Group

Alberto Di Meglio,^{1,*} Karl Jansen,^{2,3,†} Ivano Tavernelli,^{4,‡} Constantia Alexandrou,^{5,3} Srinivasan Arunachalam,⁶ Christian W. Bauer,⁷ Kerstin Borras,^{8,9} Stefano Carrazza,^{10,1} Arianna Crippa,^{2,11} Vincent Croft,¹² Roland de Putter,⁶ Andrea Delgado,¹³ Vedran Dunjko,¹² Daniel J. Egger,⁴ Elias Fernández-Combarro,¹⁴ Elina Fuchs,^{1,15,16} Lena Funcke,¹⁷ Daniel González-Cuadra,^{18,19} Michele Grossi,¹ Jad C. Halimeh,^{20, 21} Zoë Holmes,²² Stefan Kühn,² Denis Lacroix,²³ Randy Lewis,²⁴ Donatella Lucchesi,^{25, 26, 1} Miriam Lucio Martinez,^{27, 28} Federico Meloni,⁸ Antonio Mezzacapo,⁶ Simone Montangero,^{25, 26} Lento Nagano,²⁹ Vincent R. Pascuzzi,⁶ Voica Radescu,³⁰ Enrique Rico Ortega,^{31, 32, 33, 34} Alessandro Roggero,^{35,36} Julian Schuhmacher,⁴ Joao Seixas,^{37,38,39} Pietro Silvi,^{25,26} Panagiotis Spentzouris,⁴⁰ Francesco Tacchino,⁴ Kristan Temme,⁶ Koji Terashi,²⁹ Jordi Tura,^{12,41} Cenk Tüysüz,^{2,11} Sofia Vallecorsa,¹ Uwe-Jens Wiese,⁴² Shinjae Yoo,⁴³ and Jinglei Zhang^{44,45}

arXiv:2307.03236

The coming together of a wide range of expertise across disciplines to solve domain-specific problems.

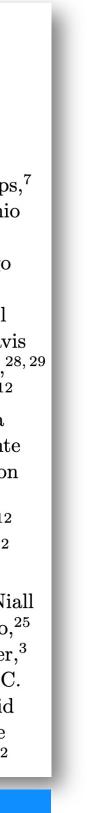
See also: arXiv:2312.09733 arXiv:2312.05344

Quantum-centric Supercomputing for Materials Science: A Perspective on Challenges and Future Directions

Yuri Alexeev,¹ Maximilian Amsler,² Marco Antonio Barroca,^{3,4} Sanzio Bassini,⁵ Torey Battelle,⁶ Daan Camps,⁷ David Casanova,^{8,9} Young jai Choi,¹⁰ Frederic T. Chong,¹¹ Charles Chung,¹² Christopher Codella,¹² Antonio D. Córcoles,¹² James Cruise,¹³ Alberto Di Meglio,¹⁴ Ivan Duran,¹² Thomas Eckl,² Sophia Economou,¹⁵ Stephan Eidenbenz,¹⁶ Bruce Elmegreen,¹² Clyde Fare,¹² Ismael Faro,¹² Cristina Sanz Fernández,¹² Rodrigo Neumann Barros Ferreira,³ Keisuke Fuji,¹⁷ Bryce Fuller,¹² Laura Gagliardi,^{18,1} Giulia Galli,^{11,1} Jennifer R. Glick,¹² Isacco Gobbi,¹⁹ Pranav Gokhale,²⁰ Salvador de la Puente Gonzalez,¹² Johannes Greiner,¹² Bill Gropp,²¹ Michele Grossi,¹⁴ Emmanuel Gull,²² Burns Healy,²³ Matthew R. Hermes,¹⁸ Benchen Huang,¹¹ Travis S. Humble,²⁴ Nobuyasu Ito,²⁵ Artur F. Izmaylov,^{26,27} Ali Javadi-Abhari,¹² Douglas Jennewein,⁶ Shantenu Jha,^{28,29} Liang Jiang,¹¹ Barbara Jones,³⁰ Wibe Albert de Jong,³¹ Petar Jurcevic,¹² William Kirby,³² Stefan Kister,¹² Masahiro Kitagawa,¹⁷ Joel Klassen,³³ Katherine Klymko,³⁴ Kwangwon Koh,³⁵ Masaaki Kondo,^{36,25} Doğa Murat Kürkçüoğlu,^{37, 38} Krzysztof Kurowski,³⁹ Teodoro Laino,⁴⁰ Ryan Landfield,²⁴ Matt Leininger,⁴¹ Vicente Leyton-Ortega,²⁴ Ang Li,^{42,43} Meifeng Lin,²⁸ Junyu Liu,^{11,44,45} Nicolas Lorente,^{46,8} Andre Luckow,⁴⁷ Simon Martiel,¹² Francisco Martin-Fernandez,¹² Margaret Martonosi,⁴⁸ Claire Marvinney,²⁴ Arcesio Castaneda Medina,¹⁹ Dirk Merten,¹⁹ Antonio Mezzacapo,^{12, *} Kristel Michielsen,⁴⁹ Abhishek Mitra,¹⁸ Tushar Mittal,¹² Kyungsun Moon,¹⁰ Joel Moore,^{50, 51} Sarah Mostame,¹² Mario Motta,¹² Young-Hye Na,¹² Yunseong Nam,¹² Prineha Narang,^{28,52} Yu-ya Ohnishi,⁵³ Daniele Ottaviani,⁵ Matthew Otten,⁵⁴ Scott Pakin,¹⁶ Vincent R. Pascuzzi,¹² Edwin Pednault,¹² Tomasz Piontek,³⁹ Jed Pitera,⁵⁵ Patrick Rall,³² Gokul Subramanian Ravi,²² Niall Robertson,⁵⁶ Matteo A. C. Rossi,⁵⁷ Piotr Rydlichowski,³⁹ Hoon Ryu,⁵⁸ Georgy Samsonidze,⁵⁹ Mitsuhisa Sato,²⁵ Nishant Saurabh,⁶⁰ Vidushi Sharma,¹² Kunal Sharma,¹² Soyoung Shin,⁶¹ George Slessman,⁶² Mathias Steiner,³ Iskandar Sitdikov,¹² In-Saeng Suh,²⁴ Eric D. Switzer,⁸ Wei Tang,⁴⁸ Joel Thompson,⁶³ Synge Todo,⁶⁴ Minh C. Tran,³² Dimitar Trenev,⁶⁵ Christian Trott,⁶⁶ Huan-Hsin Tseng,²⁸ Norm M. Tubman,⁶⁷ Esin Tureci,⁴⁸ David García Valiñas,¹² Sofia Vallecorsa,¹⁴ Christopher Wever,⁶⁸ Konrad Wojciechowski,³⁹ Xiaodi Wu,⁶⁹ Shinjae Yoo,²⁸ Nobuyuki Yoshioka,⁶⁴ Victor Wen-zhe Yu,¹ Seiji Yunoki,^{25,70} Sergiy Zhuk,⁵⁶ and Dmitry Zubarev¹²

arXiv:2312.09733





The no-nonsense path to quantum advantage

1. Run quantum circuits faster on quantum hardware

Chart a path to develop quantum technology (hardware + software) that runs noise-free estimators of quantum circuits faster than can be done using classical hardware alone

2. Map interesting problems to quantum circuits

We need applications that can only be solved with quantum circuits that are known to be difficult to simulate. This must be done collaboratively, with our clients and users.



The no-nonsense path to quantum advantage

1. Run quantum circuits faster on quantum hardware

Chart a path to develop quantum technology (hardware + software) that runs noise-free estimators of quantum circuits faster than can be done using classical hardware alone

2. Map interesting problems to quantum circuits

We need applications that can only be solved with quantum circuits that are known to be difficult to simulate. This must be done collaboratively, with our clients and users.

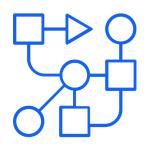


Workload management



Programmability

IBM Quantum



Hybrid classical-quantum workflows



Use case in physics



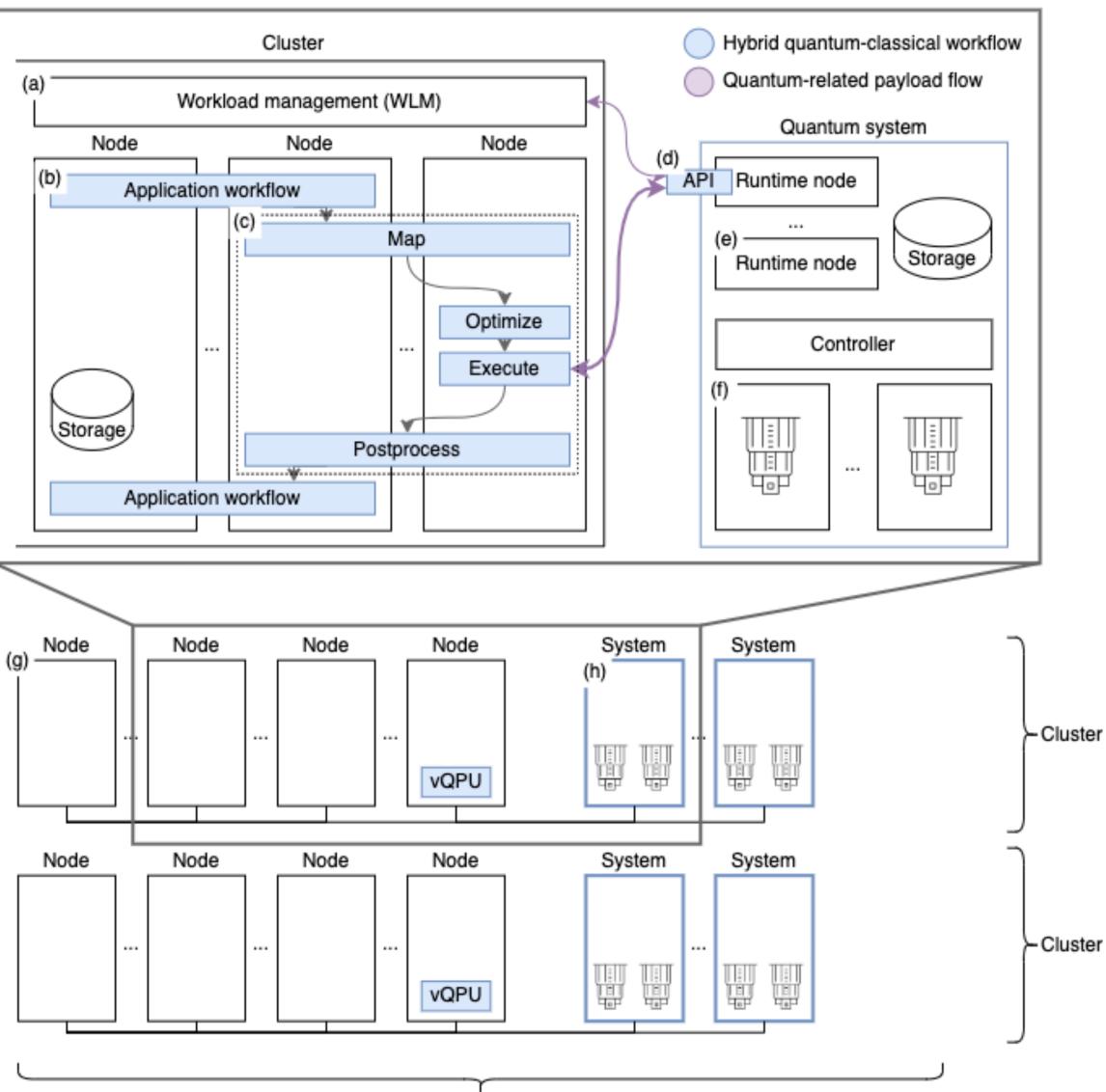
Workload management

QCSC comprises integrated classical and quantum resources

Consider WLM as conventional software system (e.g., SLURM) with extensions for interfacing to quantum devices

Worker nodes w/ 'virtual' QPU resource

Hybrid workfload spawned from login (classical) node; quantum circuit executed via low-latency, high-throughput interface ("Direct Access API")



Datacenter (co-located / multiregional)



Hybrid classical—quantum workflows



Near-term

"Weakly-coupled" regime

Latency between classical and quantum devices beyond qubit coherence time; e.g., distributed or colocated with 'slow' interconnect

Workflows comprising temporally decoupled tasks



Far-term

"Strongly-coupled" regime

High-speed bus (latency \ll coherence) connecting classical and quantum devices

Workflows, single large-scale hybrid applications with quantum accelerator complementing conventional HPC





Programmability



Near-term

"Loosely-coupled" regime

Desire is to offload quantum-accelerated kernels while Trivially separable classical and quantum executables, functions, etc.; latencies insignificant classical task(s) execute in parallel and/or asynchronously

Language-agnostic: "bag of tasks" to be executed on available resources; e.g., C/C++, Python (Qiskit) for quantum and AI/ML



Far-term

"Tightly-coupled" regime

Single-source application(s) part of a larger workflow that pervade all available resources; compiled language (C/C++), potentially with bindings





Programmability



Near-term

"Loosely-coupled" regime

Trivially separable classical and quantum executables,
functions, etc.; latencies insignificantDesire is to offload quantum-accelerated kernels while
classical task(s) execute in parallel and/or
asynchronously

Language-agnostic: "bag of tasks" to be executed on available resources; e.g., C/C++, Python (Qiskit) for quantum and AI/ML



Far-term

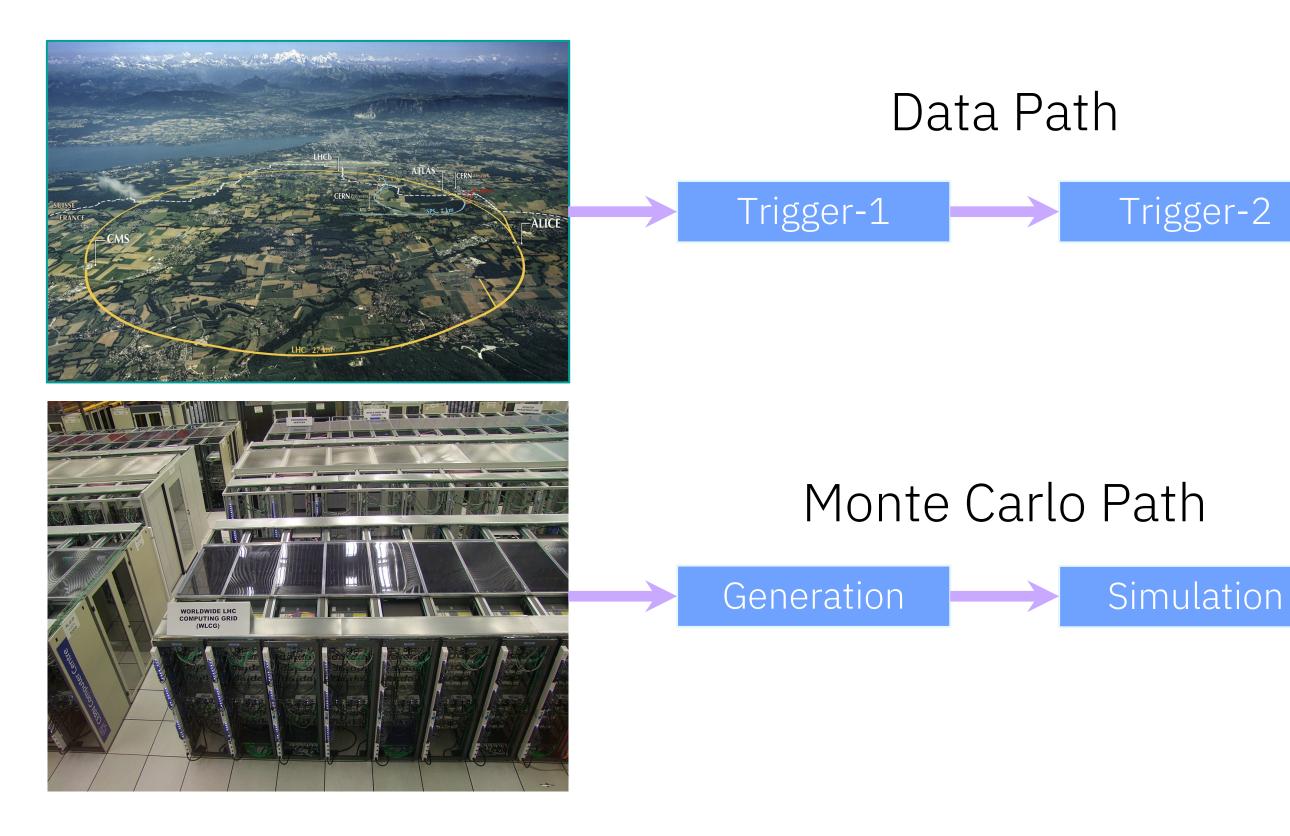
"Tightly-coupled" regime

Single-source application(s) part of a larger workflow that pervade all available resources; compiled language (C/C++), potentially with bindings

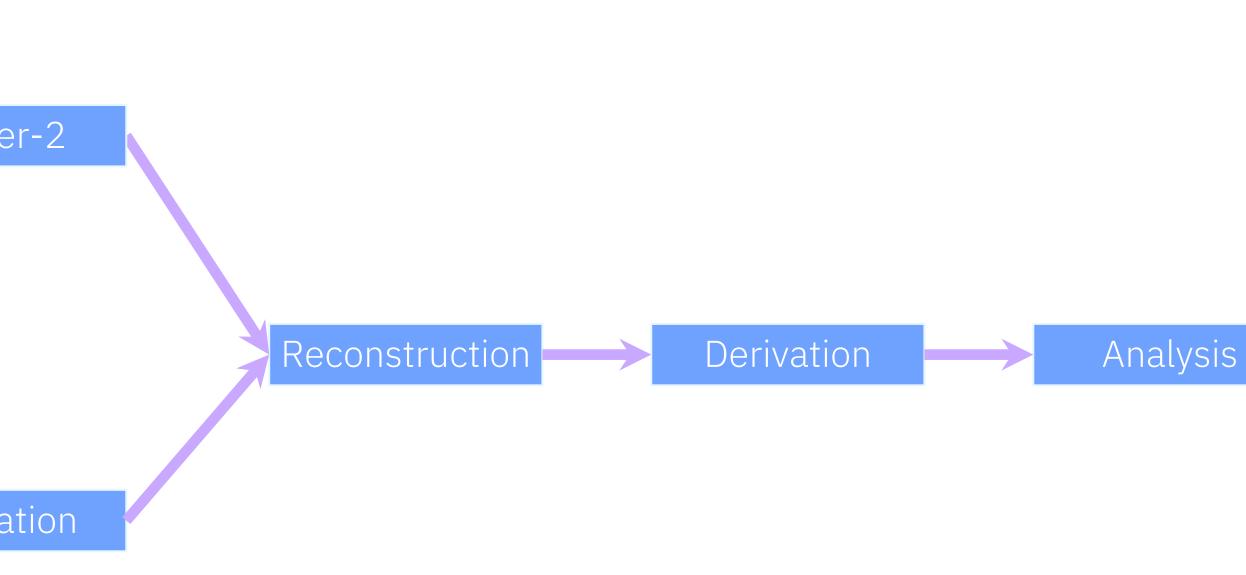
How will these systems be programmed?



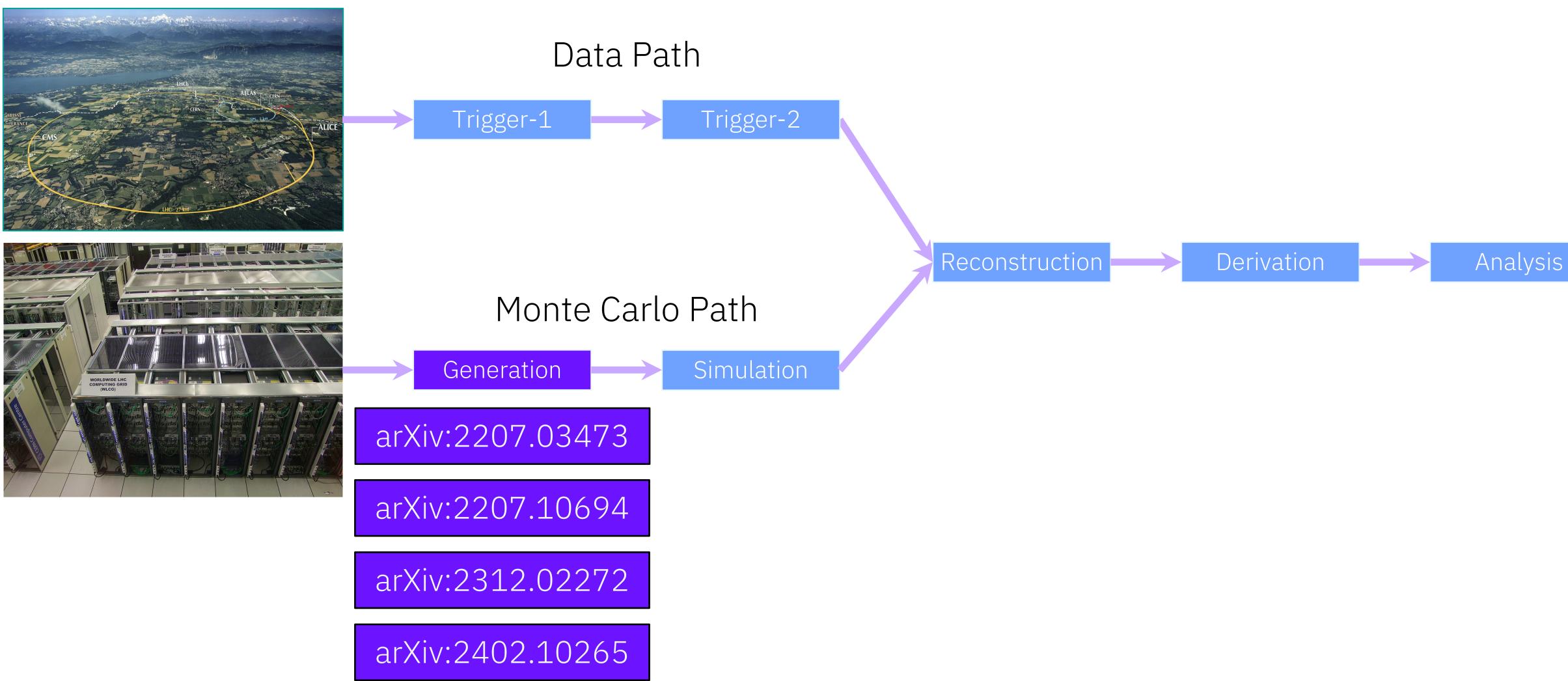




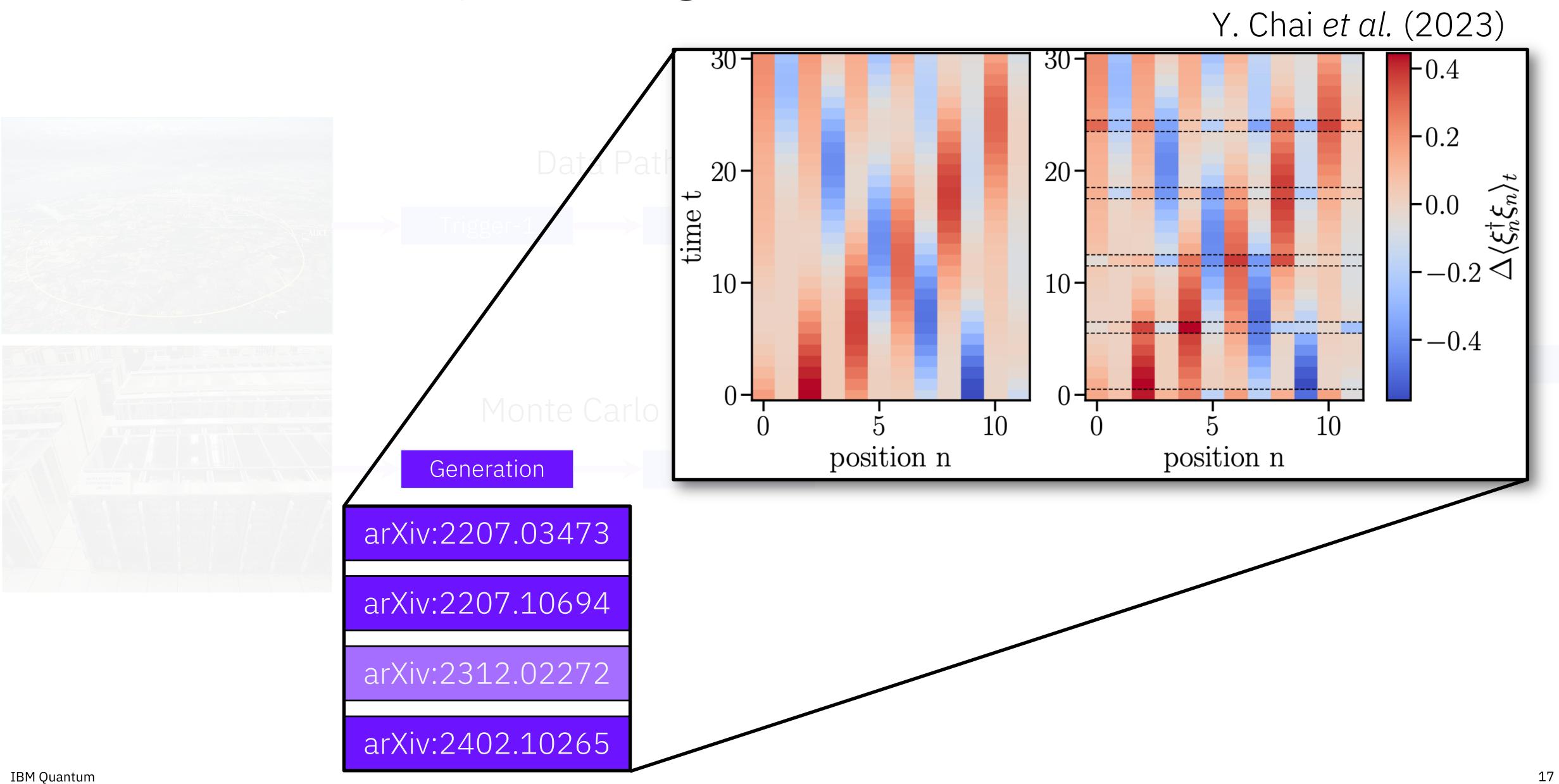
IBM Quantum

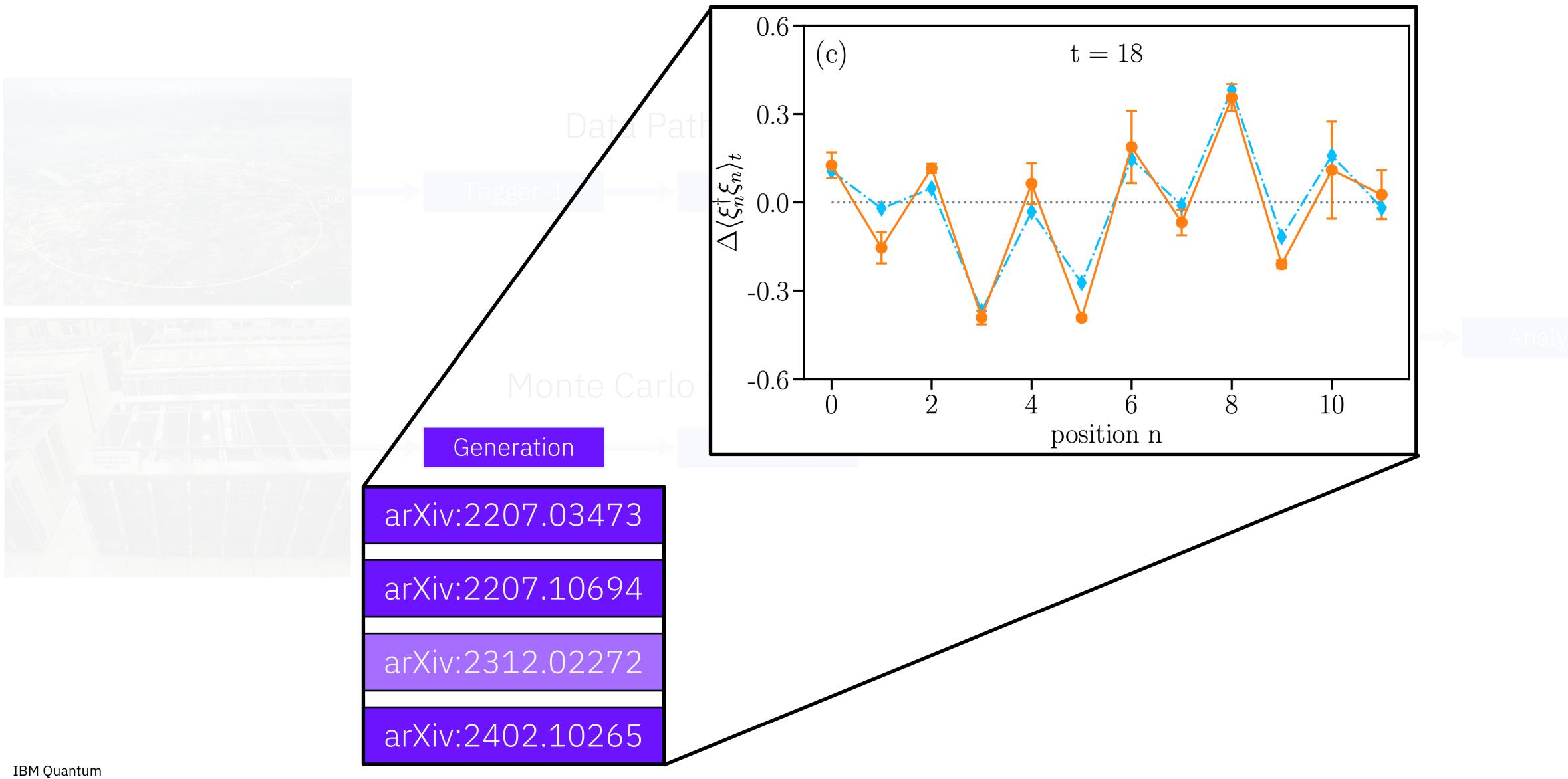








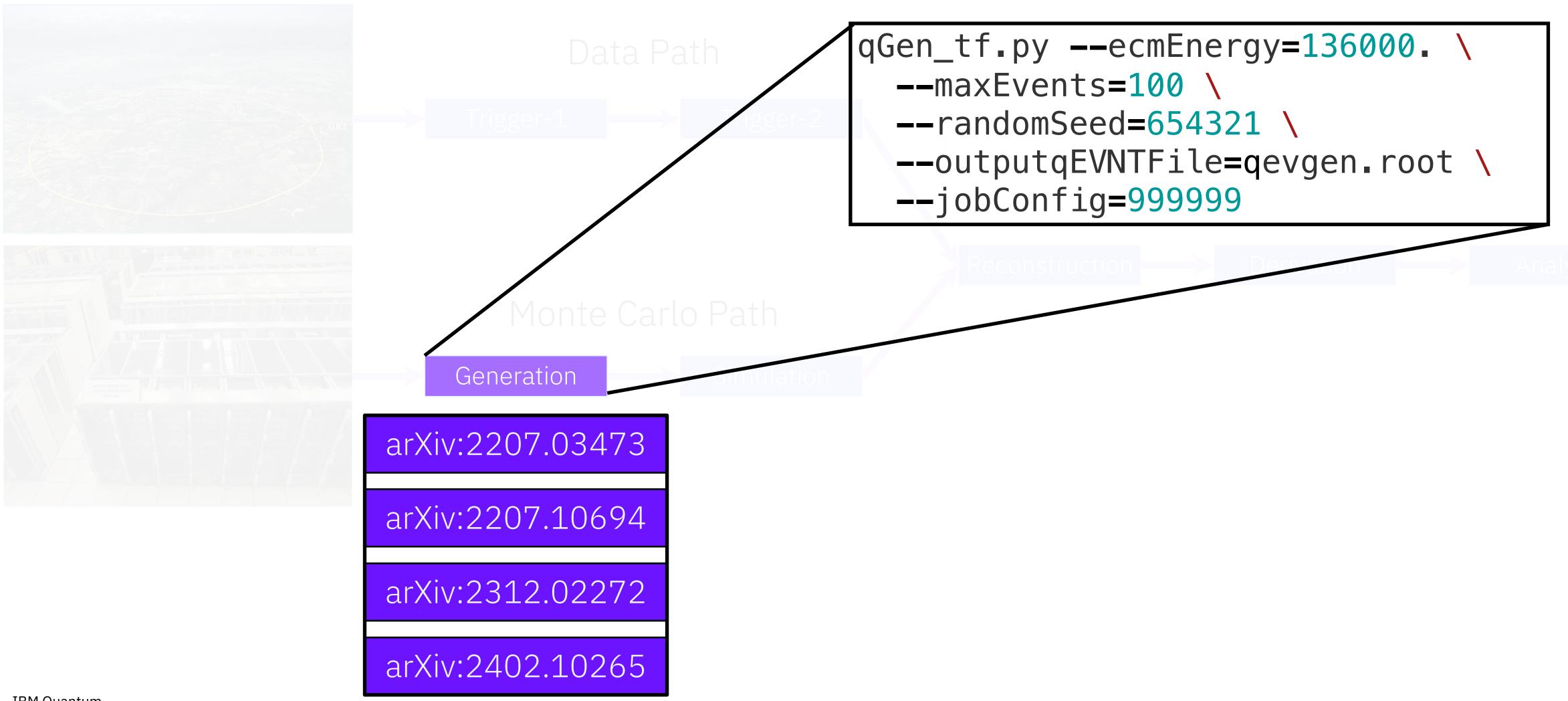




Y. Chai *et αl.* (2023)

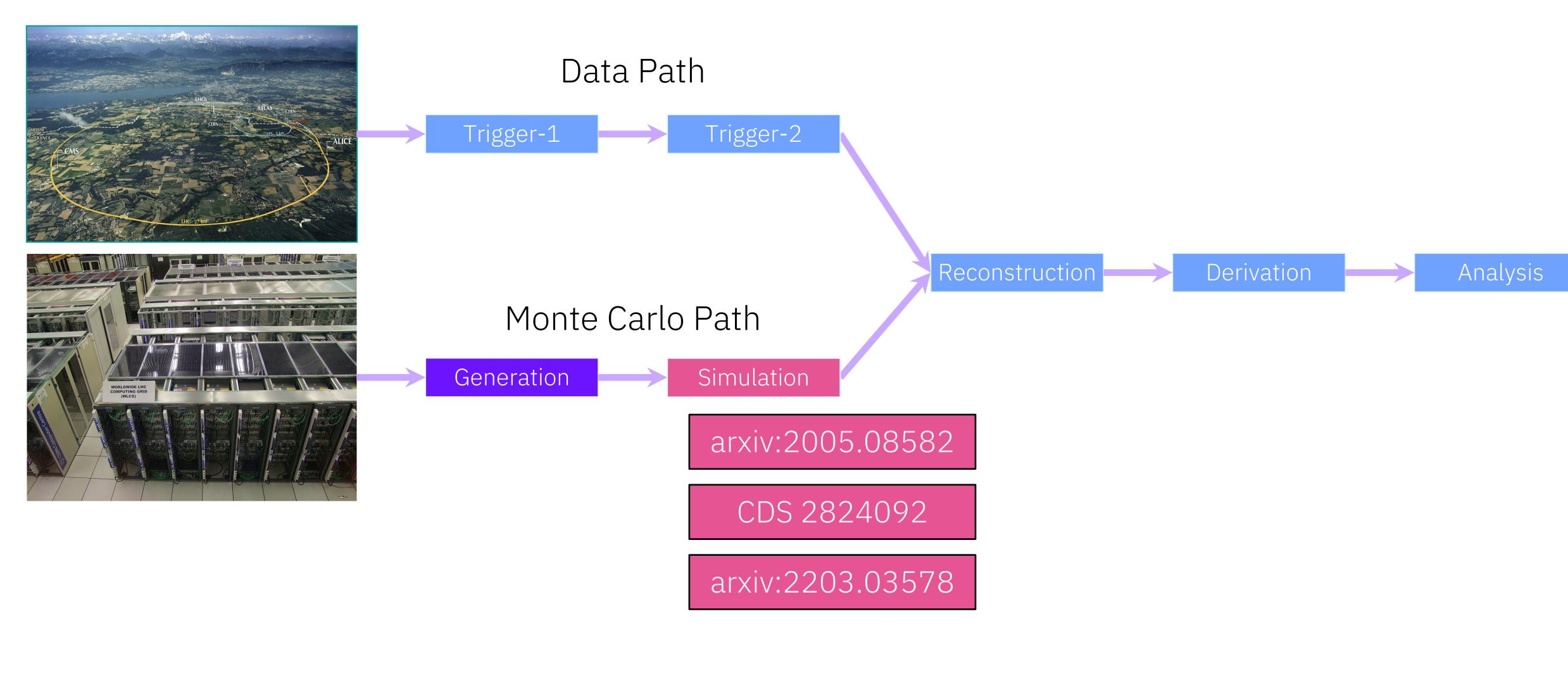




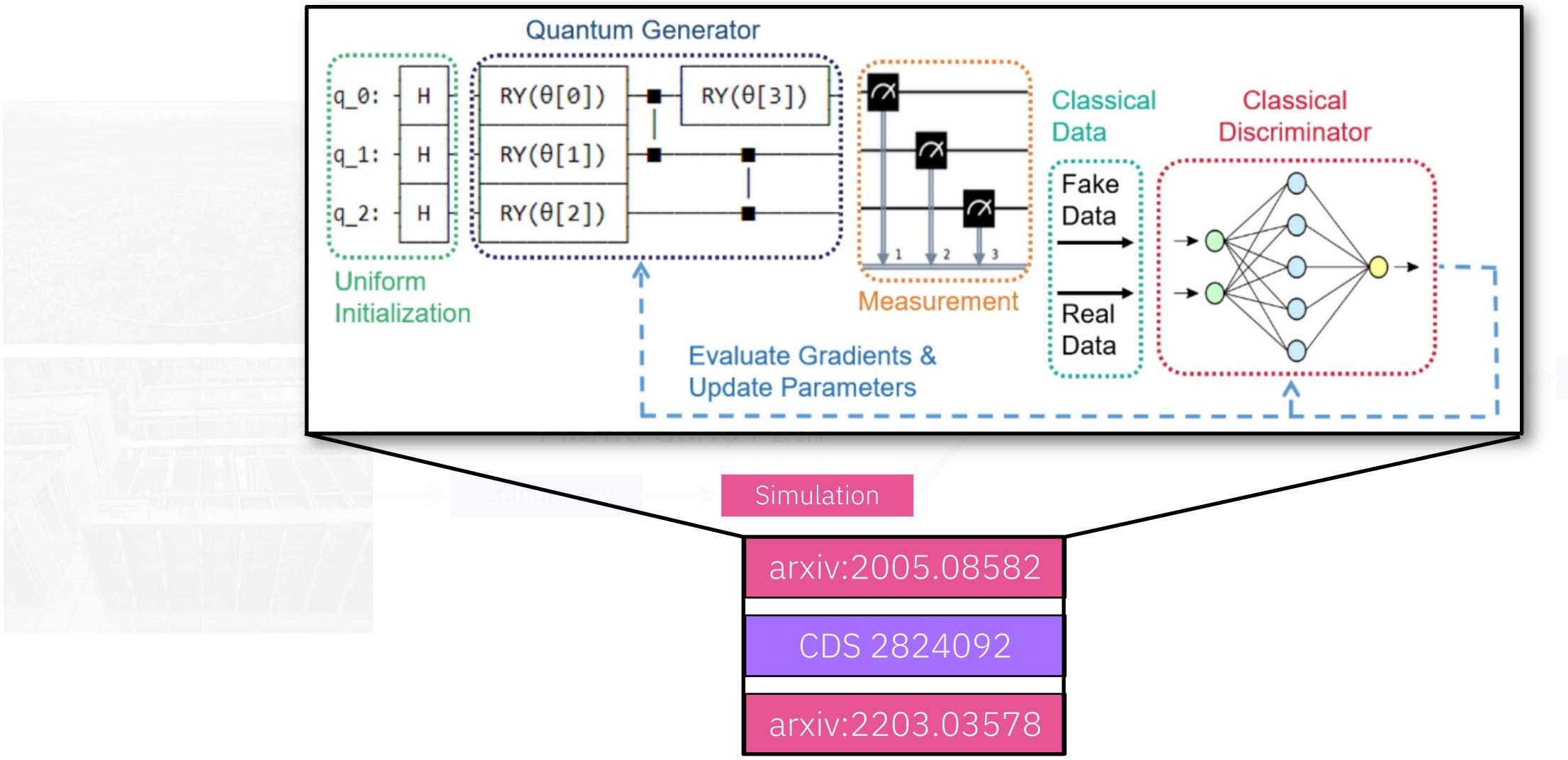


IBM Quantum



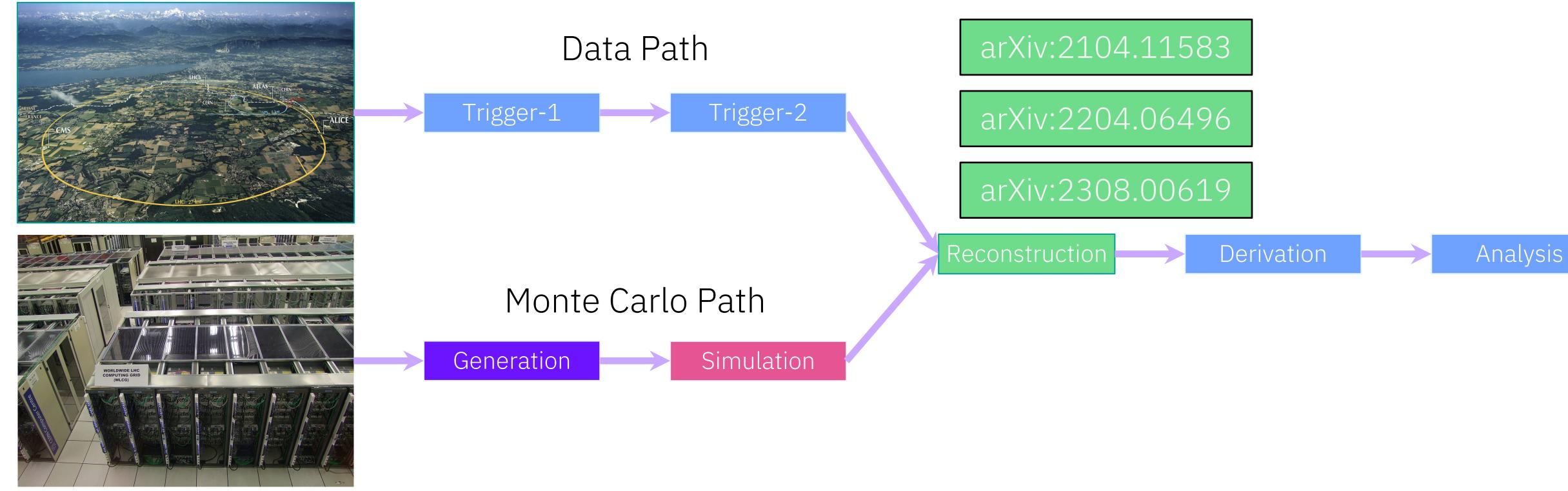




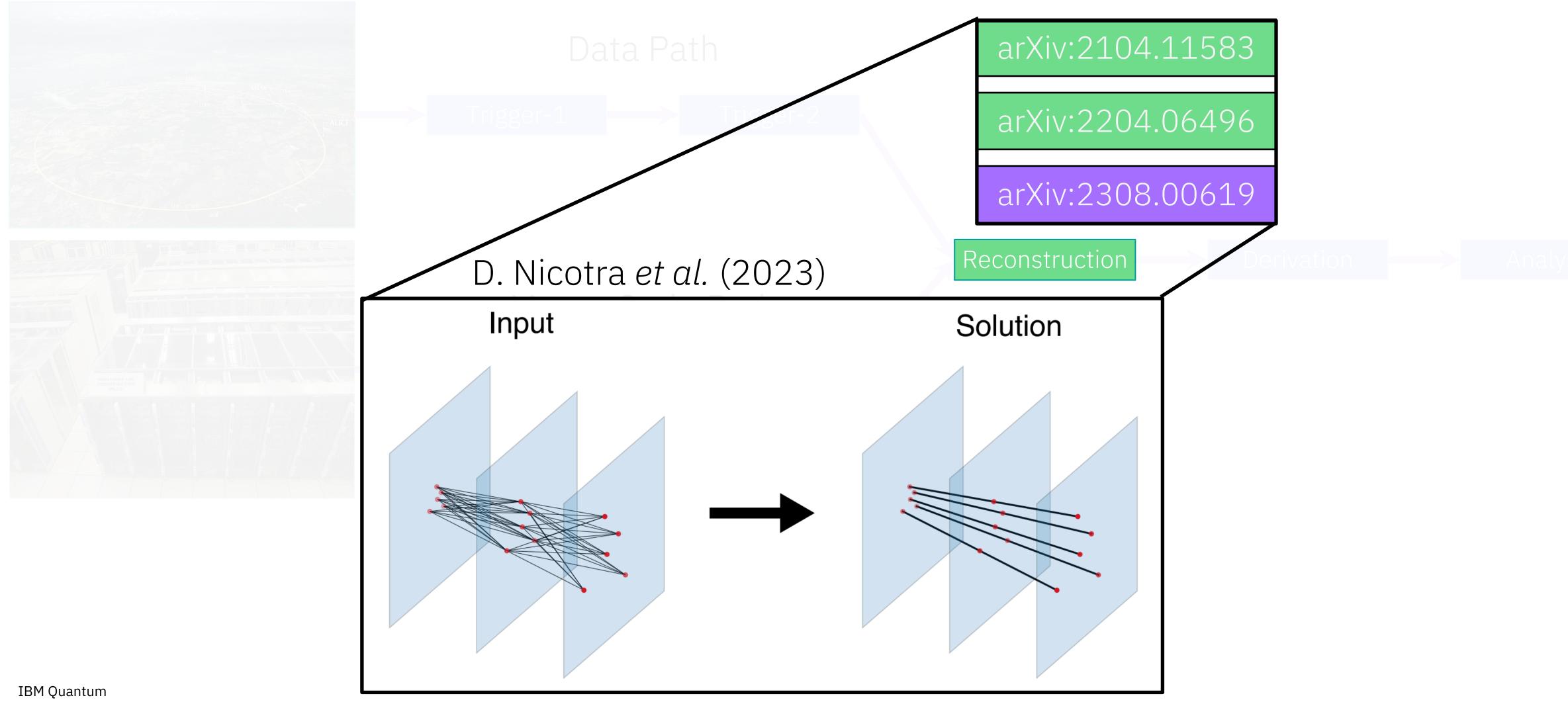


F. Rehm *et al.* (2021)

sis









D. Nicotra *et αl.* (2023)

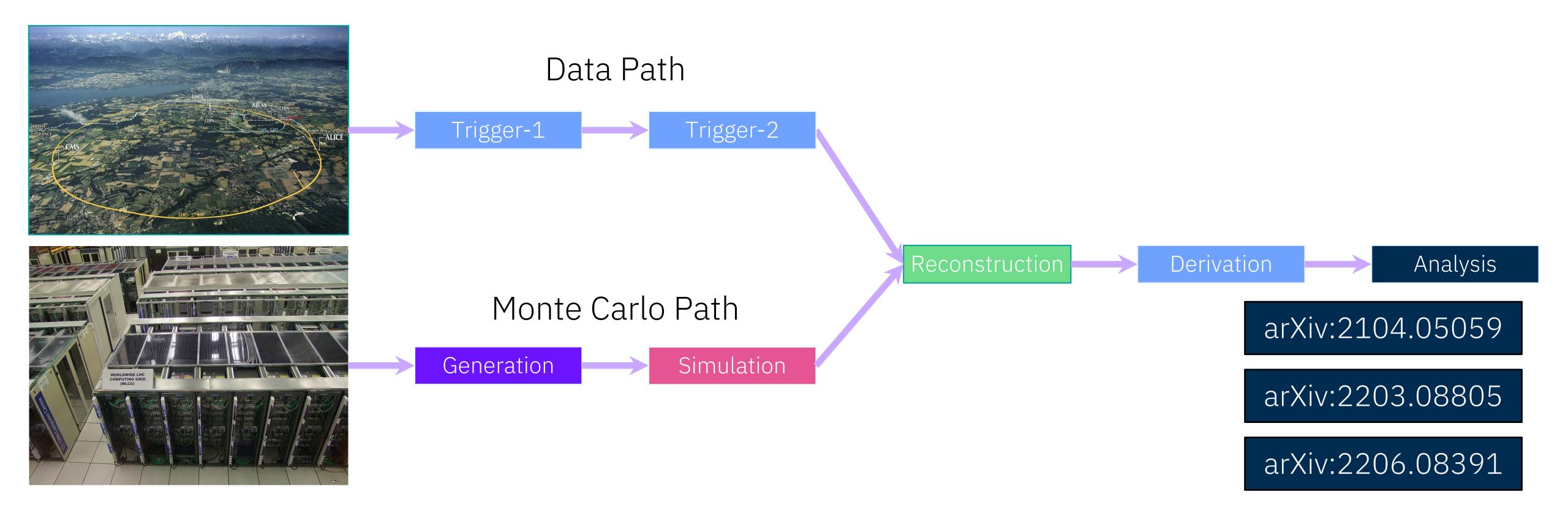
Layers	Particles	Doublets	Qubits	Depth	2-qubit gates
3	2	8	8	12 071	5 538
3	3	18	12	1 665 771	834 417
3	4	32	12	901 255	442 694
3	5	50	14	14 515 229	7 107 317
4	2	12	10	185 817	93 213
4	3	27	12	1 714 534	840 780
4	4	48	14	14 197 046	7 110 044

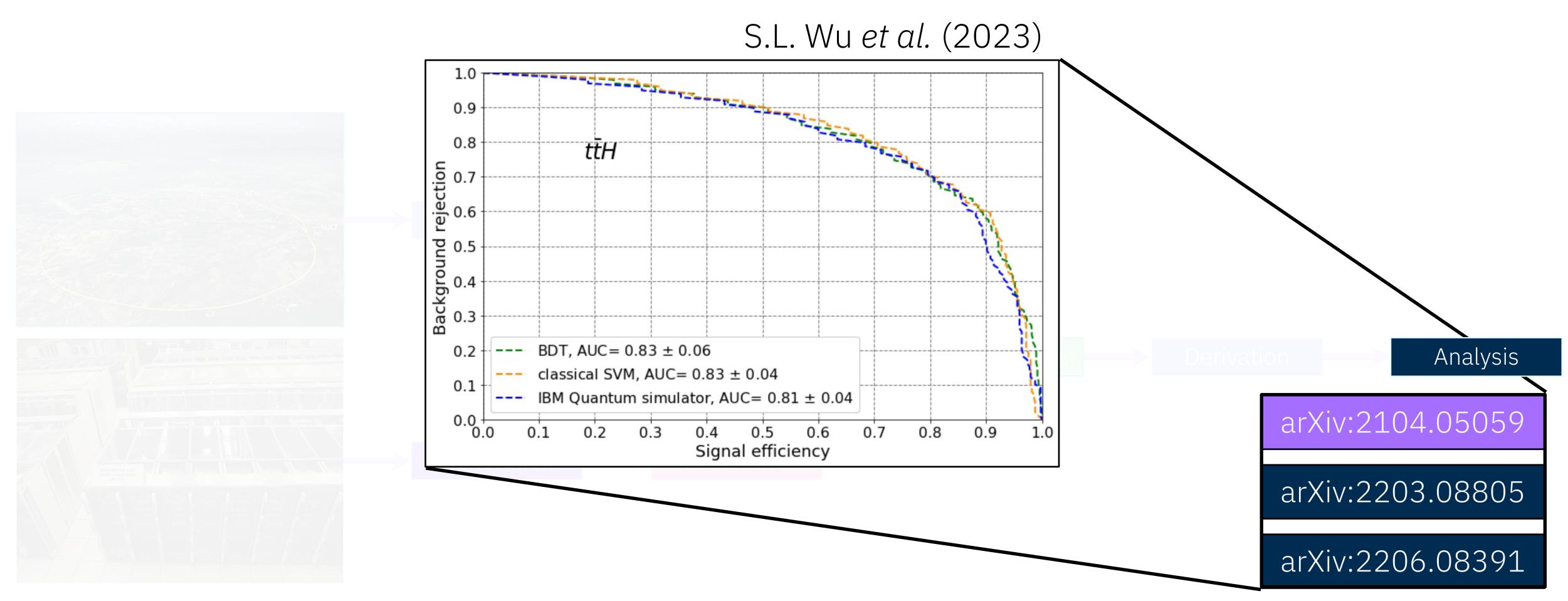


arXiv:2204.06496

arXiv:2308.00619

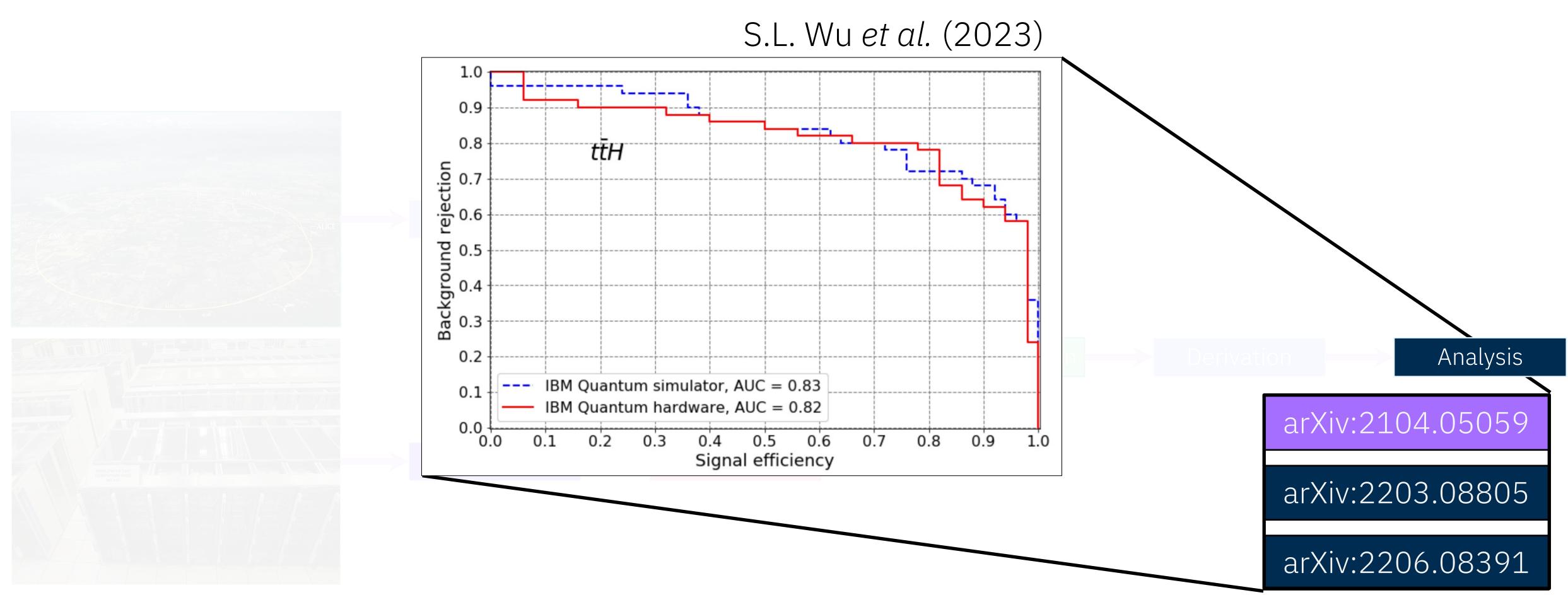
Reconstruction





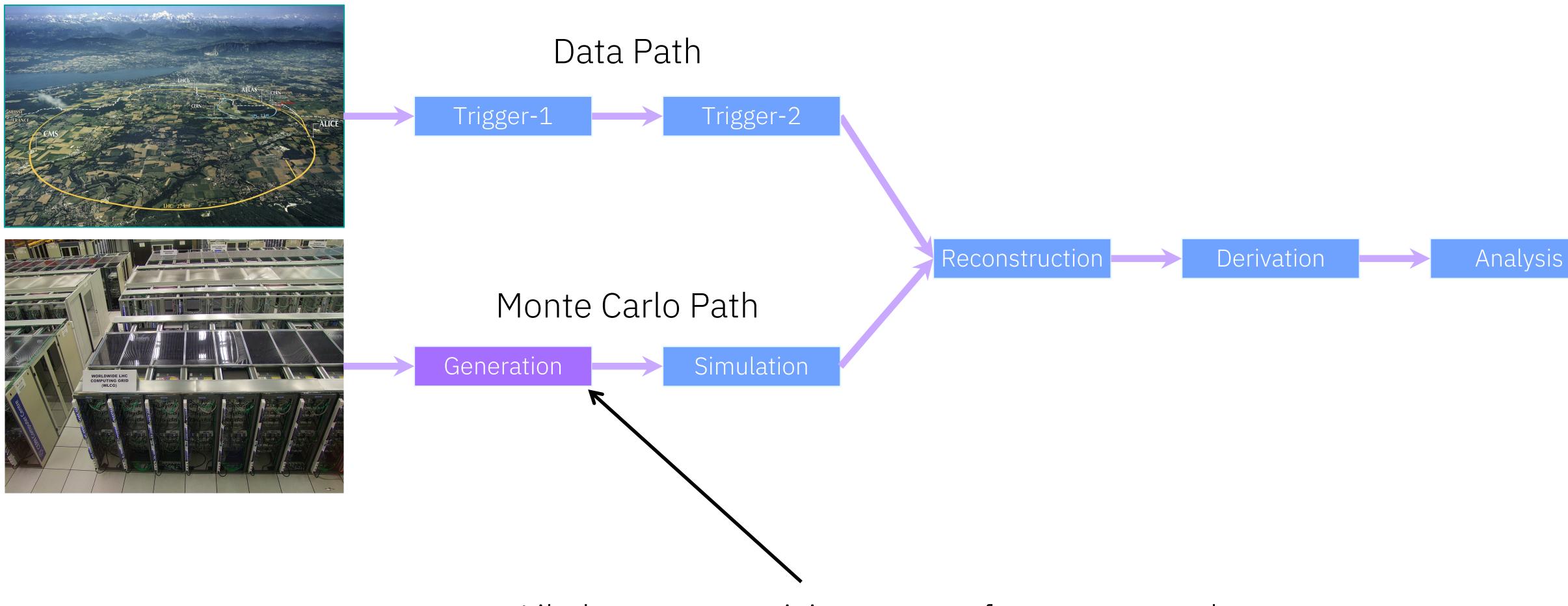












Likely most promising avenue for quantum advantage.



Summary

We are now in the era of utility, where quantum computing can provide reliable solutions at-par or beyond brute force classical computing methods.

QCSC to leverage quantum and classical computing devices to enable execution of hybrid workloads at utility-scale.

Domain scientists must find the hard problems to solve, cross-cutting expertise required to work on a solution.

Focus areas include workload management, HPC-quantum workflows, programmability, and use cases, all in the near- and far-term.

Many areas of physics already looking to quantum; theory or where quantum manifests in the problem are promising.

Keen to discuss with the community has in mind about other ideas and interests.



IBM Quantum

Backup

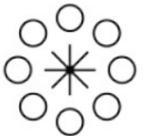


"Nature isn't classical, dammit, and if you want to make a simulation of Nature, you'd better make it quantum mechanical..."

Richard P. Feynman Simulating Physics with Computers, 1981

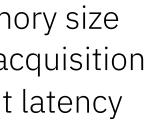
32

Error-resilience and time scales



		Middleware		Qiskit R	Runtime Quant	Quantum Hardware	
		$ \begin{array}{c} 0\\ 0\\ \\ \\ \\ \\ 0\\ \\ \\ \end{array} \end{array} $					
Timescale	Cloud loop time		Server loop time		Near-time	Real-time	
	hours	minutes	seconds	ms	µs-ns	NS	
Error resilience related function	Unitary parameter upd Custom QPU info share Logical circuit schedule optimization Final decode/reconstru	e e and compile	Compile probabilistic circuits (■) Noise model update (■■) Optimal qubit maps (■■) Multi-QPU job optimization	Physical parameter update Decoder update (●) Decode/Correct (●)	Probabilistic gates () Stretched gates () Dynamic logical ops () Fast decode, high BW () Decode data compression (Mid-circuit measurement feedforward (Non-local real operations (
Hardware challenges	Pipelining with QPUs Low latency HPC optimization Final level of decoding		Parallel QPU utilization Data BW Decode compute	FPGA Update AoT compiling	Fast reinitialization () Intra-qubit comm Local decode and compress	Local memory High-BW acqu Intra-qubit late	

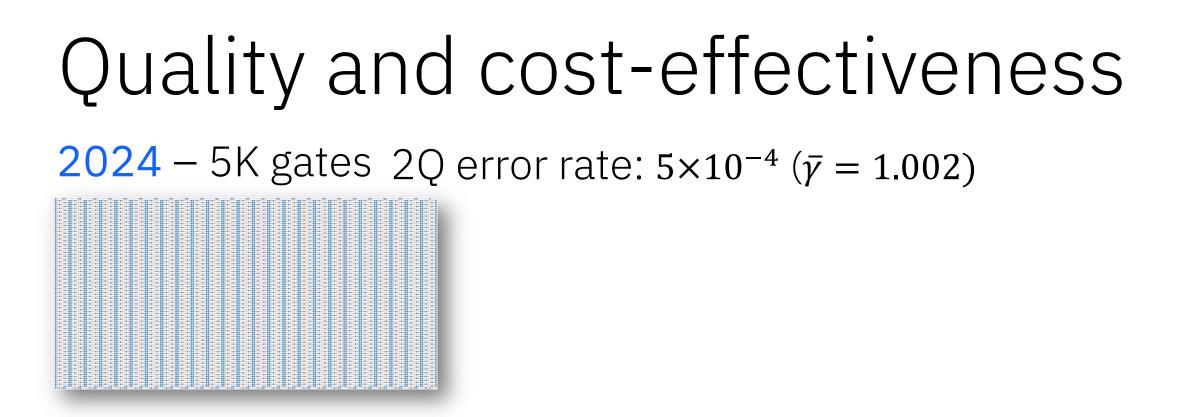




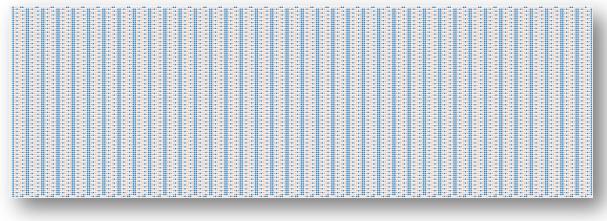




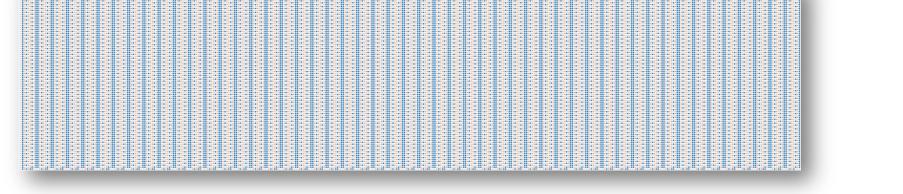




2026 – 7.5K gates 2Q error rate: 3.3×10^{-4} ($\bar{\gamma} = 1.001$)



2027 - 10K gates 2Q error rate: 2.5×10⁻⁴ ($\bar{\gamma} = 1.0009$)



2028 - 15K gates 2Q error rate: 1.6×10^{-4} ($\bar{\gamma} = 1.0006$)

IBM Quantum

Exponential growth of problem dimensionality places strict limits on classical compute scaling, and cost effectiveness of deploying ever growing resources; effective limit of ~50 qubits for direct simulation

There is no long-term value in CPU, GPU, or any other classical simulation of quantum circuits

Computations becomes more cost-effective to run on quantum hardware as error rates decrease

Crossover points for quantum cost effectiveness vs. GPUs

 $\bar{\gamma} = 1.01$ 2023: <u>460</u>

Internal Nvidia 4x DGX GH200 capable of 46Q direct simulation

2026: 32 $\bar{\gamma} = 1.001$

By 2026 it will be cheaper to run even trivial circuits on quantum hardware vs. GPUs





