



Opinions Libres

le blog d'Olivier Ezratty

Decode Quantum with John Levy and Oleg Mukhanov from SEEQC

Welcome to the 77th episode of “Decode Quantum”. We continue our “international episodes”, this times with the cofounders of the US startup SEEQC, **John Levy** (CEO) and **Oleg Mukhanov** (CTO), who are creating superconducting control electronic circuits for solid state qubits. They are the only company in that field. We had a chance to meet our two guests in Elmsford in February this year and visit their lab and cleanroom.



Picture taken in February 2024 at SEEQC's HQ in Elmsford, New York State.

John Levy is the CEO and cofounder of SEEQC in 2019. Beforehand, he was the chairman of Hypres, the company that spun-out SEEQC for reasons he'll explain later. He was involved in various startups, like goTenna, and BioLite. He has been an investor and executive in the tech industry, having co-founded and led the tech practice at the venture capital firm L Capital Partners where he served on boards of over twelve tech companies. John took 3 companies public and oversaw numerous M&A transactions. During the 1990's he co-founded the computer vision company ePlanet which spun out of Interval Research, sponsored by Paul Allen, and ran the company for Intel, its main investor. John is a graduate of the Harvard Business School.

Oleg Mukhanov, is a cofounder and the CTO of SEEQC. He was before the CTO of Hypres since 1991, where he led developments in superconducting electronics. Oleg is a co-inventor of the Rapid Single Flux Quantum (RSFQ) technology, a superconducting circuit used for high-speed, energy-efficient computation useful for both classical and quantum processors. He did his PhD in Russia with Konstantin Likharev in the late 1980s, at Lomonosov Moscow State University (MSU), before moving to the US at the end of the USSR.

Questions

Fanny: I start with our first question, how did you land in quantum, when started your passion for this topic.

John: thank you for inviting both of us. It is a fun to look back at how we started things at SEEQC. I worked at a lab in Palo Alto called Interval Research that Paul Allen started in the mid-90s, and it was run by David Liddell. David Liddell, with Bob Metcalfe, developed the Xerox STAR which came out of Xerox PARC. It was the first computer that had a graphical user interface (GUI). It had a mouse. It had a laser printer. Ultimately, it was connected to Ethernet, which Bob Metcalfe invented. What I was able to do was to work with people at Interval who were the inventors of this new platform of personal computing. We also had the inventor of the laptop computer, and I saw what it took to build an entirely new computing platform. It was just an extraordinary group of people.

I have always been interested in next generation computing platforms. When I became the chairman of Hypres and began to see the potential for using superconducting electronics in computing, and then ultimately, applying that to quantum computing, I could see this had the potential to creating an entirely new computing platform. And I just can't resist. I mean, I had to step in. And with a partner like Oleg on the technical side, between the vision that we both had and then meeting our third cofounder, Matt Hutchings, in the UK, we could see the path to moving forward.

Oleg: back from 90s where I was with Hypres, we were always working with quantum phenomena called superconductivity, and we worked on the logic called single-flux quantum (SFQ) logic, so the quantum is in the work of the logic, because it used superconducting circuits, superconducting draws an effect. However, the logic itself was classical, so we were able to develop this circuit from the very nascent state to something useful, which ended up in a product, however, the quantum era of quantum computing started in the beginning of this millennium. It fascinated us to see how we can apply our technology to quantum computing. We saw that material side, and even the very element that we use for our digital circuits are the same. These are Josephson junction, but they are used as a non-dissipative non-linear element. So a little bit in different mode, than we use it for our circuits. So the question was how we can combine these two technologies to make something useful, to make something different and more scalable compared to what everybody was doing. And later, we came to a realization that there is a way to combine these two technologies, to enable scalability of quantum computing. So which is, is that overcoming the scalability problems in the conventional architecture of quantum computers so we decided to start up a company with John and Matt when it became obvious that, here's the moment when we can really make an impact in quantum computing area.

Fanny: thank you Oleg. Don't worry, Olivier is here, as usual, for technical and scientific questions. You can start to ask your questions, it's your time.

Olivier: can you describe your thesis that you did with Konstantin Likharev in Russia which was about SFQ if I understand well. What were the conditions of that work given it happened a while ago, during the very old time of the U.S.S.R.?

Oleg: that was indeed the U.S.S.R. I did my master's in a different institution, in Moscow Engineering and Physics Institute, where I worked on microelectronics based on semiconductor, pretty much standard technology, etc., and when I learned about the group in Moscow State, a group of Konstantin Likharev, who worked on a very different technology, which is much lower power and higher speed, I was fascinated that I wouldn't be a part of their millions and millions of developers working in semiconductor industry, but rather go to a narrow field where I can really make an impact. So that was my motivation to join the group and to work on the physics foundation of devices, rather than work on sort of already developed sort of field and making just different circuits. So, when I joined the group, I was given the task to see how to make data processing in

particular analytical digital conversion. Using single flux quantum (SFQ) digital circuits was kind of almost accidentally sort of invention. That was a very interesting experience for me since it was not the direct topic of my thesis work. It initially started as just to see how to simulate how single flux quantum pulses can navigate through different circuits and what you can do with them and then this sort of essentially byproduct topic eventually superseded the original idea and the original task and became the the main topic of my work so, so that in fact is very often in the history of various technology development when, the new directions are founded almost by accident and not not originally originally anticipated.

Olivier: Konstantin Likharev was specializing in condensed matter, but he was working both on what we would call today superconducting qubit or equivalent, but he was also working on silicon and spin electronics. So was it the case really? Because he's behind some of the work behind that as well.

Oleg: his group was always focused on just on junction circuits, working primarily on the physics foundation of devices, not necessarily, not really on digital circuits initially, on just junction itself, weak links, etc. He published seminal works on these topics and then almost simultaneously with, circuits the was discovered and the focused on so-called single electron type devices which eventually developed into the qubits but at a much later time. So, when we all moved to the USA, he was working on SFQ circuits and single electron devices which actually ended up not so practical, and then later in his career he switched into more AI type devices not necessarily superconducting but semiconducting and, so with room temperature type devices. Quantum computing was never, actually his favorite direction. He was somewhat critical of that particular field and mostly focusing on classical data processing.

Olivier: John, can you tell us about the backstory of the creation of SEEQC with Hypres?

John: first off, Hypres' work was concentrated in two areas before I became the chairman of the company. It was operating in instrumentation and measurement like voltage control, microscopes and then RF, radio frequency technology and related applications. While I was the chair, RF became the biggest part of Hypres.

Around 2014, we started working on the IARPA C3 program with the goal to build energy-efficient superconducting supercomputers. The Department of Energy had run a series of studies looking at what it was going to take to scale exascale computers from an energy perspective. We're seeing this play out now, when people are building these AI-based data centers with nuclear power plant requirements. The DOE and our team at Hypres saw this back in 2012, 2013, 2014. We were part of the IBM team, Oleg led that, working on logic, and on memory, with Raytheon BBN, and we competed against a team from Northrop Grumman. Over four years, we worked on that program, and it was during that time we developed more confidence around the concept of using SFQ in a chip-based approach focusing on high performance computing based on energy efficient technology to develop scalable systems.

At the same time, I had started a strategic planning process at Hypres to understand what it was going to take to build out a high-speed and quantum computing practice. At Hypres, we not only designed these energy efficient superconducting electronic chips, we also manufactured them and integrated them into full systems. We decided that for us to develop the focus that we needed and the investment that we required to build out our quantum computing company, we had to split the company into two: keep the RF and instrumentation work at Hypres, which was nearly all US government based, and spin-out the high speed and quantum computing design, test and integration work as well as the chip foundry, which required a much more commercial focus. So we decided to do that in 2017.

It took us a couple of years to implement this. There was a lot of sensitive technology that we needed to make sure stayed in the right place with the right controls so that we could safely build out a commercial company.

We worked on that from a legal and organizational perspective, and then ultimately, we raised seed level

financing to support the spin-out. We raised \$6.8M led by a venture capital company called BlueYard and in April 2019, we spun the company out, becoming SEEQC which stands for scalable, energy efficient quantum computing.

Olivier: when Hypres was working on RF, you could sell only to the US basically it was hard to export stuff which you try to do on SFQ?

John: yes, so a lot of the RF work that we were doing was for DoD and the intelligence community, and it was clear to us that this was a very limited market as opposed to working in commercial markets for computing, with a much larger application space, much bigger platforms, and where we could raise outside capital so we had more control over our ability to develop and grow.

Olivier: we're going to switch to SEEQC and what you do. I know you do that, both of you, very well. So educate our audience on why we need superconducting control electronics to drive a solid-state qubit. It's not accepted, I would say, broadly. So I'd love to get your pitch on that.

John: in the creation of new technology platforms, inventors typically develop different kinds of prototypes though it's not necessarily the case that those original prototypes are the technology that ultimately scale. There are often discontinuities in terms of development, which was the case in digital or classical computing. The very first classical computers were made with vacuum tubes and telephone switchboard connectivity, and each one of those vacuum tubes had multiple pins and every pin had a cable that had to be connected to every other cable. It may not have been clear at the time, but it became clear after the fact, that we were never going to scale classical computing by trying to build more efficient vacuum tubes with more cables. We needed something else. Just a year or two later at Bell Labs, not far from where the ENIAC computer at the University of Pennsylvania was developed, saw the invention of the first transistor, which led to the development of integrated circuits at Fairchild, to microprocessors at Intel, etc. Along the way, a slew of technical problems were discovered, resolved, and new categories of technologies were invented. The point is the discontinuity between the earliest classical computers and what we have today based on solid state devices, massive integration and the ability to manufacture at scale.

We think the same process is occurring in the quantum domain. We have multiple ways of building quantum computers. They're all in this very early stage, if you look at, say, superconducting quantum computers, you'll see racks of room temperature electronics, multiple cables connecting to each individual qubit and amazing complexity. Signals travel from cold temperature at milliKelvin up to room temperature and back, analog to digital and back. We have to amplify signals and filter them. There's latency. There's noise, they are expensive and error prone and slow et cetera, et cetera. And if you look at what a computer looks like today, at a small number of qubits, I don't know, between 75 and, say, 100, it's almost an impossibly complex system.

So we said, let's not do it that way. Let's not try to scale up existing systems. Let's take the core capabilities of our SFQ, our single-flux quantum technology, and put that on chip. Because we're operating at such low temperature and low power (which translates to low heat generation and the need to dissipate that heat dissipation), we can put our circuits in direct proximity to qubits. And by doing that, we can directly resolve issues around I/O, latency, speed, energy and heat dissipation. We can make our circuits run digitally which resolve a whole other set of issues like crosstalk and GPU chip integration, and we can scale whole systems to work in data centers.

When most people talk about scaling quantum computers they're often thinking about scaling the number of qubits, but we think about scaling as resolving all the parameters I just mentioned plus more: I/O, latency, speed, fidelity, coherence, cost, complexity, noise, manufacturability and supply chain. We think that in order to scale a quantum computer you have to resolve all of those, and that's why we've taken the approach we've

taken.

Oleg: the way quantum computers are designed now and being built and people trying to, it's somewhat natural again because the first quantum computer were built by physicists and if you look at their quantum computing setups now they exactly like that so it's a physics experiment where you have cryogenic objects if you talk about superconducting circuits, and then it's surrounded by expensive measurement equipment and control equipment, to induce sort of control pulses and read them out so, computers do not look like that right so computers are done very differently and fortunately we at hypers we matured we developed a mature technology which can be placed alongside with the qubits without disturbing their quantum state and this is single part quantum technology, uh due to their very low, power dissipation so none of the existing quantum computing player have this technology, as their core competency so that makes our company unique in this perspective so for us it was very natural to start integrating the control layer classical control layer and readout layer inside of the cryostat next to the qubit and by doing so eliminating all these cables etc if you look at the present system of the famous photos of google system which is the technology marvels because it works with such complexity and you see only wires there because the dilution refrigerator which is the cryogenic platform in order to cool qubits with chips to milli-kelvin temperature primarily pulls cables. That's what it does. That's why you need a larger and larger cooling platform, larger and larger dilution refrigerator, sometimes reaching some enormous size and power, just in order to cool cables, not cool qubits, because they really don't dissipate that much. We saw our main goal as to eliminating this, the roadblocks, the cables, which are standing on a way of scaling onto the computer. And that can be achieved by placing control elements inside of the cryostat and communicating with qubits via the result cables, so just using capacitive and magnetic wings. And that's what our technology actually does. was able to do. We didn't jump into this field right away, but first it was idea that we had to replicate all this room temperature electronics using superconducting circuit. So imagine that we would need to put millions and millions of joules and junctions in order to be able to generate this RF signals for qubit control etc. That would be a daunting task because that would require a lot of much much higher technology development level on our side until it was clear based on some academic work in the US etc. that you really don't have to do that. So you can use pure digital methods to do that. interact with qubits and also read them out on chip, basically directly converting quantum information to digital without necessity to bring everything to room temperature. So once we realized that, that triggered us to start up the company and implement this, quantum computer architecture, which looks like a computer, with the help of our superconducting.

John: an important consideration of our chips is that they are, as Oleg mentioned, classical digital chips. We use what we know about digital technology as a huge advantage to the performance of our computers and choose not to use conventional RF microwave control or RF based readout or RF multiplexing. What we're seeing is a massive reduction in crosstalk, which as you know is a significant source of noise for quantum systems. One can go through all kinds of contortions to address it with digital pre-distortion and other methods, but, what we see is with our SFQ digital chips we can greatly simplify the design of a quantum computer, get better performance and do so at a fraction of the cost to build a scaled system.

Having a digital chip-based system enables us to do that. And going back to the development of the microprocessor, it was the resolution of the cabling issue, the I/O issue, by using digital multiplexing that enabled the development of the classical microprocessor and the massive integration of the microprocessor by resolving interconnections and control lines. There's a wonderful paper around that, referred to as Rent's Rule. And then many years later, Jim Clark and his team at Intel, wrote a paper called "Rent's Rule and extensibility in quantum computing" that addressed this exact phenomenon. If you study history you'll see these patterns over and over again, and we're seeing it play out exactly in this way in the quantum domain.

Olivier: two weeks ago, I had a chance to meet with a guy from Google and he was doing the presentation on

QEC (error correction) and he mentioned publicly one amazing fact that they are planning to put 10,000 physical qubits on a single chip which is quite a challenge, and they plan to control that in a single fridge with control electronics and as many other vendors like IBM and some companies, they all plan to use cryo-CMOS control circuits which is the by default choice of many of these vendors. Could you explain why cryo-CMOS is a bad choice and why SFQ is a better choice and probably on top of that, explain us maybe also what are the challenges to build SFQ because all the companies have been discussing with said SFQ is interesting but it doesn't work yet so. I wish it worked since if your technology worked, it'd a boon with solving a lot of problems, simplifying the whole stuff.

Oleg: cryo-CMOS is a natural choice for many companies because it is very mature well. CMOS is very mature technology, so the necessary step is to migrate CMOS to cryogenic temperatures which again can be done. It is an engineering problem. They essentially recreate what was done at room temperature. All these mixed signal microwave circuits in cryo-CMOS is placed at 4K, but qubits are at 10 mK so you still have distance to traverse with your generated RF signal, so what what you've done in using cryo-CMOS you eliminated cables between fridge and room temperature electronics and you simplify room temperature electronics but within the cryostat nothing changed. So if you open it up you still see myriads of coaxial cables necessary to send signal down. And when people recreate room temperature electronics with cryo-CMOS, they have to sacrifice something. In order to save power dissipation, it helps to some way with some sort of the expenses on the performance, but does not solve entirely the problem. So our solution is to solve the problem in its entirety by placing SFQ chips next to the qubits, and that becomes possible with low power dissipation technology and also which is more important, that simplicity of the circuit allows to implement controller, even at the present state of SFQ technology development, which of course is less developed than CMOS.

Olivier: there is an exception on what you say, because the very low temperature, let's say 10 or 15 mK is the one for superconducting qubits but those companies will try to do more silicon quantum dot qubits plan to operate at a higher temperature like around 100 mK to maybe 500 mK and they think that they could run some electronics at this stage and not rely only on 4K. You have Pando Tree from Intel which was released recently as a prototype which demux RF signals at the qubit level. So, does it mean that there's a clear difference between silicon qubits and superconducting qubits or what you say is valid for silicon qubits as well extent?

Oleg: either you bring both technologies to some intermediate temperature, or you take one technology and try to scale it up or scale it down. Fundamentally, the silicon qubit can work at high temperature, and physics does not prohibit that question of performance. So far, I've heard proposals to reach maybe 1 Kelvin and place both technologies at that temperature. However, it's always a question of the cost to performance. Power dissipation of cryo-CMOS remains an issue; it's much higher than in SFQ technology, which still requires careful control to avoid disturbing the qubits' quantum state. We have expertise in superconducting electronics, which operate at lower power and can function below 5K. So, while it could be a competitive area, we wish them success.

John: controlling a qubit for us involves about three nW per qubit, which is significantly lower. Increasing CMOS speed means more power and heat generation, requiring heat dissipation, a tough trade-off. Operating at nWs, our chips can be clocked to 20 GHz, providing room for further development. Also, readout is often an overlooked functionality.

Oleg: I haven't seen advances in CMOS readout. CMOS may do readout better, but it's still dispersive, with signal amplification and cryogenic transfer to an analog-to-digital converter. Current cryo-CMOS circuits focus primarily on control, while readout remains as it is. In our case, we provide a full solution, addressing both control and readout, which is essential for utility-scale quantum computers. Readout fidelity directly impacts the overhead for a logical qubit; an error in readout means multiple measurements, increasing complexity in error decoding and quantum error correction.

John: our protocol is faster and we operate with lower latency. On the development side, we're preparing three papers showcasing our architecture, flux bias control, and readout as described by Oleg. These results should be surprising to the industry. Today, we operate SEEQC Orange, a full-stack, chip-based quantum computer, running algorithms on it, using much of our core SFQ technology.

Olivier: interestingly, D-Wave uses similar technology with SFQ to drive DC-driven qubit, with around 5,600 qubits, avoiding the cable clutter of companies like Google or IBM. That's a promising example.

John: our focus is on a full-stack, chip-based architecture with all core quantum computer functions. We prefer a holistic approach over specific features.

Olivier: earlier this year, you shared charts on energy footprints, particularly around 100 logical qubits on superconducting systems. Scaling to around 1,000 qubits, as planned by companies like Google and IBM, raises questions about energy differences from today's room-temperature electronics.

John: we've looked at this in various ways, discussing this topic with IBM and Google. For a 100,000-qubit quantum data center, just the chiller budget could require around 10 to 12 megawatts, excluding cooling and operational costs. Data centers could face tens of megawatts in energy needs. In contrast, our chip-based approach significantly reduces energy consumption, potentially down to tens of kilowatts instead of megawatts. SEEQC, which stands for scalable energy efficient quantum computing, emphasizes energy efficiency both at the chip and data center scalability levels.

Olivier: you mention a three-order magnitude difference in power consumption which is a massive advantage.

John: absolutely. A dilution refrigerator requires significant space for cables, each generating heat. We've solved this by minimizing cables in our setup, especially with SEEQC Orange and our upcoming SEEQC Yellow platform, which will be even more efficient. Our goal is a fully integrated quantum computer on a chip within a year.

Fanny: what developments have occurred since we last met in February?

John: we've made advancements in readout, programmable multiplexing, and launched SEEQC Orange. We've built on these achievements, and Oleg may add more.

Oleg: digital technology for qubit control introduced functionalities unavailable with analog control. We initially planned a demultiplexer, allowing one control driver per qubit. Then, we realized we'd built a programmable router, enabling simultaneous access to multiple qubits. Close proximity digital circuits eliminated issues like crosstalk, often challenging with superconducting circuits.

Olivier: scaling beyond 100 qubits without crosstalk is challenging.

Oleg: exactly. Experiments show we lack measurable crosstalk, an unexpected but welcome outcome. Additionally, we've solved problems that many struggle with, such as RF component elimination, improving system response and reducing costs.

John: we've leveraged these improvements, eliminating many expensive components and simplifying system operations. Moving to a chip-based approach has reduced costs while increasing responsiveness and uptime.

Olivier: how do you connect chips to room temperature control? Is it fiber, high-speed serial, or another form of connectivity?

Oleg: these are cables, but instead of sensitive RF signals, we use digital pulses, less demanding on cable

fidelity. We're developing superconducting multi-line flex cables without connectors for cryogenic packaging, a big advantage.

John: additionally, we're working with NVIDIA on direct digital links between our SFQ pulses and their CMOS chips, aiming for high-speed, low-latency connections. Our goal is to integrate quantum chips with NVIDIA's GPUs and CPUs for shared memory use.

Olivier: will error correction be handled by NVIDIA systems at room temperature, or will SFQ chips take this on?

Oleg: we foresee a multi-stage error correction approach. Initial checks and decoding will occur at cryogenic levels, reducing data rates. Complex cases would be sent to room temperature for easier handling.

John: our aim is to support diverse error correction schemes, providing a hardware system to make these approaches faster.

Olivier: wrapping up, could you discuss partnerships in Italy, the UK, and your broader market structure? If your technology scales, it could significantly restructure the industry.

John: initially, we aimed to build a top-tier quantum computer. However, investors suggested focusing on solving common industry challenges rather than competing directly with giants like IBM and Google. So, we pivoted to support superconducting, spin silicon, photonic, and even ion-based systems. While we haven't publicly announced partnerships yet, we work with some of the largest companies in quantum technology, aiming to be an integral player across the ecosystem.

Olivier: publicly, you work with OQC in the UK and QuantWare in the Netherlands. New vendors are entering the market, focusing on specific aspects of quantum chips. It's a novel approach.

John: yes, we've designed our own qubits and those for others, though we now concentrate on readout control, multiplexing, error correction, and GPU integration. Our goal is to be an integrated solution provider for the entire quantum ecosystem.

Fanny: thank you, Olivier, for preparing thoughtful questions, and thanks to John, Olivier, and Oleg for participating in Decode Quantum. To our listeners, see you soon.

Cet article a été publié le 26 novembre 2024 et édité en PDF le 29 novembre 2024.
(cc) Olivier Ezratty – “Opinions Libres” – <https://www.oezratty.net>