SuperMOS: Combining Superconducting and CMOS Integrated Circuits

Erik P. DeBenedictis
Principal/CEO
Zettaflops LLC
Albuquerque, NM, USA
erikdebenedictis@gmail.com

Robert D. Wortman
Skywater Technologies Inc.
Bloomington, MN, USA
Robert.Wortman@skywatertechnology.com

Stephen R. Whiteley
CEO
Whiteley Research, Inc.
Sunnyvale, CA, USA
stevew@wrcad.com

Robert E. Freeman
Customer Success Group
Synopsys, Inc.
Sunnyvale, CA
robert.freeman@synopsys.com

***Abstract*—Moore's Law recently shifted from dimensional scaling to exploration of new materials, devices, and the third dimension. We continue this trend by combining superconducting and semiconductor technologies to exploit their synergies. While combining known subsystems from both domains may yield useful—but predictable—results, this paper focuses on novel use cases at the application level in artificial intelligence, quantum computing, supercomputing, and sensor fusion. It also outlines specific hybrid circuits uniquely enabled by SuperMOS. To realize SuperMOS, foundries must integrate a superconducting process at the back-end-of-line (BEOL) of a CMOS process—an approach already demonstrated at the physics and materials science levels.**

Keywords—SuperMOS, hybrid, CMOS, superconductor, Josephson junction, kinetic inductor, reversible logic

I. Introduction

CMOS integrated circuits transformed the world through exponentially increasing component density and rising speed. However, progress has slowed, and new methods are needed to scale beyond current limits. SuperMOS extends circuit scalability for applications amenable to cryogenic environments. By combining superconductors and semiconductors, we create a hybrid architecture termed "SuperMOS," which provides greater capabilities—though qualitatively different—than either technology alone. This paper explores use cases where one technology's strengths compensate for the weaknesses of the other.

Current integrated circuits contain either superconducting or CMOS components—but not both. Our approach enables monolithic fabrication of both types on a single die. We define SuperMOS (Fig. 1) as the seamless integration of superconducting wires, passive elements (e.g., inductors), and Josephson junctions with conventional CMOS circuits. A key innovation is cross-technology vias (Fig. 1b), which enable power delivery and up to THz signal transmission between the two layers. MIT-LL has performed experiments validating key steps for SuperMOS [1].

SuperMOS designers can leverage superconducting wires to minimize power and signal-distribution losses. They can also employ superconducting inductors to enable more efficient power handling, analog circuits, and reversible logic. Both Josephson junction-based single flux quantum (SFQ) logic and CMOS achieve top speeds just below a terahertz (Fig. 1b), opening a new trade space between speed and power for logic design optimization.

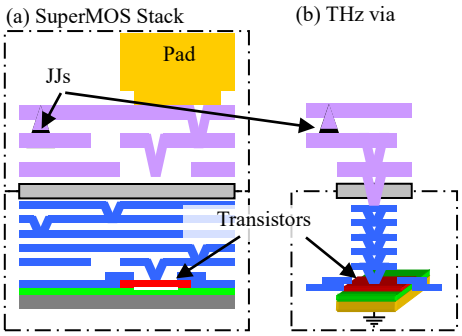


Fig. 1. (a) SuperMOS Stack comprised of a blue semiconductor integrated circuit with a Josephson junction (JJ) process on top (purple). (b) The interface (gray) includes a via that may carry signals up to about 1 THz.

Once SuperMOS becomes available, designers will be able to apply it beyond its initial motivation. In order of current national R&D priorities, this includes artificial intelligence (AI) processing, quantum computing, supercomputing, and optics/sensor fusion. Each section follows with a discussion of relevant circuit-level use cases.

A. Artificial Intelligence

The needs of artificial intelligence diverged from the CMOS roadmap some time ago when memristors entered the field to represent synaptic weights. Memristors introduced new circuit characteristics that directly mimic learning and inference behavior.

Recent interest in AI suggests continued exploration of novel devices. Josephson junctions are being considered because they can represent and compute with integers, based on the fact that superconducting loops hold only an integral number of flux quanta (note: this is not a qubit concept) [2].

However, a system of superconducting loops alone is insufficient. An AI processor requires the transfer of integer

synaptic weights to standard memory and external interfaces, where conventional memory must reside close to the loops to enable high-bandwidth, low-power data transfer.

Thus, the SuperMOS use case for artificial intelligence would include a conventional CMOS base seamlessly connected to Josephson junction circuitry that leverages the unique properties of these devices.

B. Quantum Computing

SuperMOS is a candidate technology for quantum-computer control. The top-level architecture for many quantum computers includes qubits at millikelvin temperatures controlled by tightly coupled control electronics at an intermediate temperature stage around one kelvin—with both supported by conventional electronics at room temperature.

There are two principal types of 4 K control electronics, each with strengths and weaknesses, where SuperMOS could combine the best aspects of both into a single, more scalable system.

Cryo-CMOS qubit controllers, such as Intel's Horse Ridge [3], use digital CMOS to control waveform synthesizers via analog transistor circuits. These systems are limited by heat dissipation from the CMOS controllers.

All-superconductor systems [4] use Josephson junction SFQ circuits for digital control and send pulses directly to the qubits. These systems face limitations due to large Josephson junction memory-storage cells, requiring substantially more memory for waveform storage.

The obvious and existing solution involves creating separate superconducting and semiconductor chips and connecting them using packaging technologies such as microbumps. This approach echoes the historical packaging of discrete transistors and passive devices—which led to Moore's Law in 1965. Thus, SuperMOS may represent the natural extension of integrated-circuit scaling to Josephson junctions, inheriting the advantages of integration.

C. Supercomputers

Josephson junction electronics were once envisioned for exascale-class supercomputers based on the low energy dissipation of Josephson junction logic. Decades ago, this vision was challenged by the reality that Josephson junctions are physically larger than semiconductors and would result in systems with orders of magnitude less memory compared to competing supercomputers with semiconductor memory. This limitation shifted the vision toward a superconducting processor with semiconductor memory—both housed within a cryostat.

The current limitation of this approach is inefficient data transfer between superconducting and semiconductor chips—requiring large, slow, and high-dissipation circuits—primarily in the SFQ-to-CMOS direction. For example, an SFQ pulse of about 1 mV must be amplified by a Josephson junction circuit before being transmitted over a line.

One SuperMOS solution would involve placing semiconductor memory circuitry as a base layer and Josephson junction logic on top.

An alternative solution would embed memory bus drivers (such as double data rate—DDR) in the semiconductor layer and integrate temperature-qualified but otherwise commodity-off-the-shelf (COTS) DRAM or stacked high-bandwidth memory (HBM) on nearby modules. The authors believe SuperMOS will offer greater cost-effectiveness than custom chip packaging—though commodity memory may remain cheaper.

Cryogenic reversible logic provides an additional computational option based on kinetic inductors in the superconductor layer. Since the 1990s, researchers have explored converting CMOS logic-gate circuits to reversible logic gates and fabricating them using the same semiconductor process. The concept relies on connecting the resulting circuit to an energy-recycling power supply that reduces power draw from the utility grid—either directly or by reducing heat dissipation in a cryostat and lowering cryocooler power consumption. This approach has not reached commercial rollout due to inadequate inductor technology for the energy-recycling power supply. However, kinetic inductors would enable practical cryogenic reversible logic.

All the use cases above would benefit from lower-power cryo-CMOS. Until recently, cryo-CMOS was simply room-temperature CMOS running within a cryostat. Recent work [5] has rebalanced various transistor parameters specifically for operation in cryogenic environments—resulting in more than an order-of-magnitude reduction in energy consumption and enabling effective operation below 1 K. The SuperMOS concept could extend this to rebalanced CMOS, enabling effective SuperMOS below 1 K—including reversible logic with energy-efficiency improvements beyond those achievable with rebalanced CMOS alone.

D. Optics and Sensor Fusion

With minor extensions, SuperMOS can include comprehensive optical capabilities. Superconducting processes support superconducting nanowire single-photon detectors (SNSPDs), which could become image sensors if arranged in arrays on the top surface—or serve as receivers for signals arriving via optical fiber. The semiconductor process can include lasers to generate optical signals.

These dual capabilities allow SuperMOS modules to communicate optically without requiring additional chips and associated packaging.

Furthermore, SuperMOS with a sensor array could detect and preprocess images, sending interpreted results electronically or via optical fiber.

The application-level use cases just explained will build on specific circuits. The sections below discuss circuits, or implementations of circuits, that are not feasible without SuperMOS.

E. Data-Movement Circuits

Let us compare the advantages of a SuperMOS module versus separate superconducting and semiconductor chips connected via wires or 3-D packaging technologies such as microbumps.

In existing designs, every signal path between a Josephson junction and a transistor introduces parasitic low-pass filtering or transmission-line effects. While engineers readily understand schematic diagrams with arbitrary connections of Josephson junctions, transistors, and passive elements, historically only all-superconducting or all-semiconductor chips have been fabricated—connected solely via bonding wires or transmission lines, as shown in Fig. 2a. At a signal frequency of 500 GHz—the maximum bandwidth of Josephson junctions and transistors—a 600 μm bonding wire would contain two bits at the same time, creating the parasitic impedance.

We estimate that a via connecting a Josephson junction to a transistor is approximately 10 μm—about one thirtieth of a 1 ps bit. At this scale, the via behaves effectively as a wire in a schematic diagram. Microbumps vary in size and typically exhibit poorer geometric control than direct vias, placing their performance somewhere in between.

Without parasitic components, significantly more efficient interconnects can be implemented.

Consider a use case motivated by spiking neural networks: routing neural spikes represented as SFQ pulses. For the static routing-information path on the bottom of Fig. 2b, CMOS would load routing information during power-up, then its clock would be stopped for hours or days while the routed path is used by many spikes.

The circuit receives a spike on the top left and relays it to the top right only if the CMOS routing information is at the correct level (the circuit acts as an AND gate where one input is a CMOS level). The incoming SFQ pulse is duplicated, and each copy is sent to a D flip-flop (DFF) via a path with delay. If the spike reaches the D input before it arrives at the clock input, the spike continues along the specified path; otherwise, it is suppressed.

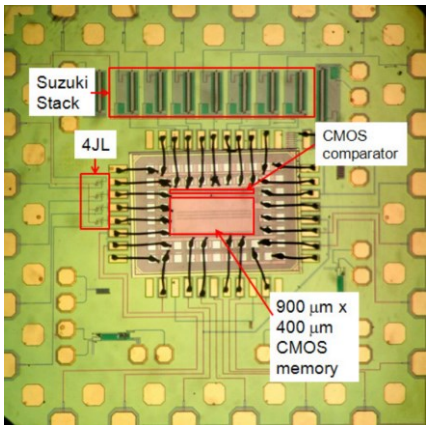
However, one copy of the spike makes a detour straight down through a via to a semiconductor varactor—a voltage-controlled variable capacitor. When the CMOS routing level places the varactor in a low-capacitance state, the SFQ pulse arrives at the D input sooner than when in a high-capacitance state—thereby controlling relative timing and determining whether the spike will be relayed.

CMOS dissipates significant dynamic power during system boot-up or routing path reconfiguration, but this accounts for a small fraction of total operating time. During the remainder of operation, CMOS operates at much lower static dissipation levels.

For each routed spike, three Josephson junction circuits generate and consume one SFQ pulse each, and the system should be considered to consume a share of the CMOS static dissipation.

The function in Fig. 2b could be implemented without SuperMOS using only Josephson junctions. The key difference is that large, high-performance Josephson junctions would occupy substantial chip area for a function expected to run once per hour or per day.

(a) Josephson-CMOS integration highlighting data transfer (precursor to SuperMOS). Bonding wires are 600 μm.



(b) Routing SFQ pulses as neural spikes.

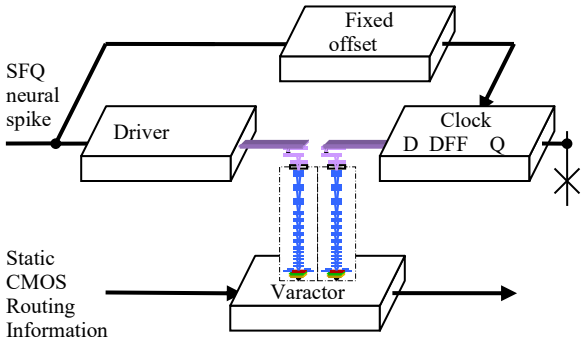


Fig. 2. Circuit for routing SFQ pulses representing neural spikes. The D Flip-Flop's clock and data both come from the same spike. The spike is relayed to the output if and only if the copy to the D input arrives before the one to the clock. However, the spike to D will have a variable delay based on loading by the varactor. Power consumption will be static CMOS plus several SFQ pulses. Without SuperMOS, the SFQ pulse would not have time to reach the CMOS chip and get back.

Historically, data transfer from CMOS to SFQ relied on DC-to-SFQ converter circuits. These circuits convert each CMOS pulse into an SFQ pulse—requiring the energy of one dynamic CMOS transition and one SFQ pulse per conversion. This is adequate for supercomputer memory where the data is about half ones and half zeros.

However, neural network routing presents different data statistics—in fact, statistics similar to FPGA reconfiguration. The data from the semiconductor side changes slowly, making it advantageous to exploit data patterns to reduce chip area at the expense of a negligible increase in energy consumption.

Data movement in the opposite direction—from superconducting to semiconductor chips—has been challenging due to mismatched voltage levels and impedances, exacerbated by high-speed operation. For instance, during SFQ-to-CMOS data transfer, a Suzuki stack (labeled in Fig. 2a) amplifies SFQ pulses by a significant factor—but still requires a transistor amplifier on the receiving side, biased to high currents, to boost the signal to CMOS-compatible levels. We have developed several circuits that exploit unique SuperMOS properties to implement this functionality, dramatically

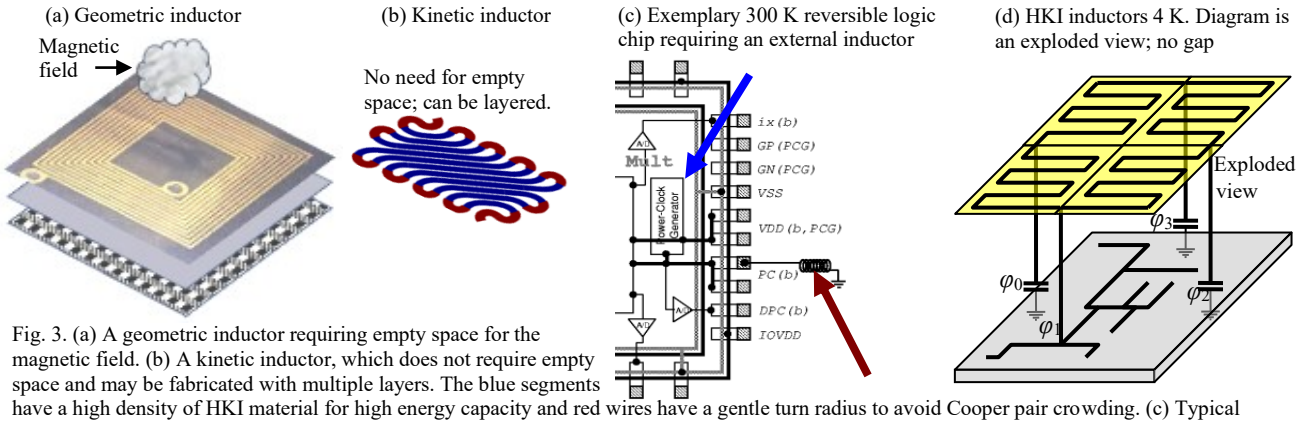


Fig. 3. (a) A geometric inductor requiring empty space for the magnetic field. (b) A kinetic inductor, which does not require empty space and may be fabricated with multiple layers. The blue segments have a high density of HKI material for high energy capacity and red wires have a gentle turn radius to avoid Cooper pair crowding. (c) Typical reversible logic test chip including a "power-clock generator" (blue arrow) but also requiring an external inductor (red arrow). (d) Reversible logic structure with kinetic inductors as the energy-recycling power supply layered on a transistor chip. Exploded view; actual fabrication is monolithic.

improving upon the baseline in Fig. 2a—but these are beyond the scope of this paper.

In summary, this section illustrates that the historical separation of superconducting and semiconductor fabrication—combined with bonding wires (Fig. 2a), whose device model is a low-pass filter—unavoidably inserts low-pass filtering on paths between Josephson junctions and transistors, restricting the range of realizable circuits. SuperMOS will drastically reduce the impact of these parasitic filters for relevant data. This section describes two previously unimplementable circuits.

A benefit of SuperMOS could be a period of innovation for circuits containing both Josephson junctions and semiconductors.

F. Power Circuits and Kinetic Inductors

Superconductors are ideal for low-loss power transmission and can serve as superior wires. This obvious use case will not be discussed further.

We propose using kinetic inductors and potentially reversible logic to reduce the 40-60% share of pins dedicated to power and ground (GND) in state-of-the-art CMOS processors and accelerators.

Superconducting processes enable fabrication of kinetic inductors with exceptional properties, suitable for integration alongside CMOS transistors within power-supply and power-amplifier circuits. Currently, such circuits rely on discrete off-chip inductors due to limitations in standard CMOS processes.

Traditionally, science classes teach inductors as wire coils that contain magnetic fields. When a current I flows through a wire, the stored energy is $E = \frac{1}{2}LI^2$, where the constant of proportionality L represents the coil's inductance. The magnetic field extends beyond the coil's surface by approximately its diameter (the cloud in Fig. 3a), so geometric inductors are usually placed on the surface of integrated circuits—except at very small inductance values.

Regardless of the storage mechanism, any circuit element whose energy equation is of the form $E = \frac{1}{2}LI^2$ qualifies as an inductor. In superconductors, current is carried by Cooper pairs. By controlling the number of Cooper pairs per unit length in a high-kinetic-inductance (HKI) wire, we regulate the pair velocity for a given current. This yields a kinetic-energy contribution from each Cooper pair described by $E_{\text{CooperPair}} = \frac{1}{2}mv^2$, where v is the pair velocity and proportional to current I. Due to the quadratic dependence in the energy expression, this kinetic energy directly contributes to the total inductance. The resulting inductance can be orders of magnitude greater than geometric inductance. Kinetic inductors (Fig. 3b) exhibit low loss (high Q) due to zero resistance and require no empty space.

Kinetic inductors have a maximum energy per unit area, given by $\frac{1}{2}L_{\square}I_c^2$ where $L_{\square}$ is the HKI layer inductance per square, and I_c is the maximum current per unit width of the wire. Asymptotically, the inductor in Fig. 3b has the same energy capacity as the blue rectangular region; this capacity remains unchanged when the blue wire's width varies. Although the resulting inductance changes, the energy capacity stays constant—a property that becomes important later.

Reversible logic represents a mid-level SuperMOS use case, enabling large-scale circuits that support the high-level use cases described above. Fig. 3c and 3d illustrate this concept. Often termed the "baby brother" of quantum computing, reversible logic has long been studied in the physics of computation field, defining a rigorous lower bound on power consumption. Commonly implemented with transistors from a CMOS process and driven by an AC clock that continuously recycles energy between cycles, early 1990s chips demonstrated 1–2 orders of magnitude higher energy efficiency compared with equivalent CMOS circuits, but they required an energy-recycling power supply.

Despite its theoretical promise, reversible logic has not achieved widespread adoption because scalable energy-recycling power supplies remain unavailable. Historical analysis by researchers working with spreadsheets or pencil-and-paper models showed that copper-trace conductivity was too low to yield meaningful wall-plug energy savings. External inductors were considered, illustrated in Fig. 3c by a red arrow pointing toward an external inductor [6], but as CMOS power densities rise, non-superconducting inductors are rapidly reaching the end of their window of opportunity.

Moreover, simple calculations reveal that such inductors would become prohibitively large, violating speed-of-light constraints required for high clock rates competitive with CMOS. Nonetheless, reversible logic attracted funding for roughly a dozen R&D projects over several decades; although these projects produced more reversible logic chips, none achieved a breakthrough in the energy-recycling power supply, so none progressed toward commercialization.

Under cryogenic operation, SuperMOS can host reversible logic circuits in the semiconductor layer and fill an equal-area region within the superconductor layer with kinetic inductors. This configuration satisfies reversible logic requirements, provided the kinetic-inductance layer has sufficient energy capacity to match the effective capacitance of the CMOS layer. The required energy capacity can be computed from the well-known formula $P = CV^2 f$, where P, V, and f are obtained from CMOS datasheets and C can be derived from the equation. Dividing $½CV^2$ by $½L_{\square}I_c^2$ yields the necessary number of kinetic-inductor layers.

For Intel's Horse Ridge processor, evaluation [7] indicates that two kinetic-inductor layers are necessary. These layers can be fabricated as part of a standard superconducting process. Thus, SuperMOS employing reversible logic with two kinetic-inductor layers could enable quantum computer controllers similar to those in the literature [3], while simultaneously reducing power consumption and allowing a larger number of qubits per system.

More than two kinetic layers may be achievable using a variant of the stairstep etching process [8] patterned into a multilayer configuration resembling Fig. 3b.

With N kinetic-inductor layers, SuperMOS could support cryogenic processors where power density grows with the capability to manufacture the kinetic inductors—essentially a scale up path.

This approach also extends to cryogenic sensor arrays that perform in-situ data preprocessing.

Furthermore, DARPA's Low-Temperature Logic Technology (LTLT) program [9] explored high-T_c cryo-CMOS for datacenter computing. The SuperMOS concept as presented here could create an LTLT-like processor operating at 4 K.

However, Fig. 3a depicts a high-T_c YBCO kinetic inductor [10]. The crystalline structure of YBCO and other high-T_c materials precludes co-fabrication with CMOS. Since these materials are used for power inductors, vias may be slow and large as long as they support high current. Future SuperMOS design could adopt a 3-D assembly approach that supports semiconductors plus high-T_c kinetic inductors—effectively a "CMOS+X" architecture, where X is a high-T_C kinetic inductor.

III. Conclusions

We have presented SuperMOS—a hybrid chip technology that combines superconducting and semiconductor components—and discussed both application-level and circuit-level use cases at 4 K. The application-level use cases include artificial intelligence, quantum computing, supercomputing, and integrated sensors.

If successful, the approach could be extended to higher temperatures (high T_c, liquid nitrogen).

An efficient N-layer kinetic inductor process based on stairstep etching would enable co-fabricated superconductors to supply power to cryo-CMOS at performance levels now found in data center processors; we believe this goal has not been explicitly articulated previously.

References

[1] MIT Lincoln Laboratory, "Superconducting integrated circuits," [Online]. Available: https://www.ll.mit.edu/research-and-development/advanced-technology/microsystems-prototyping-foundry/superconducting

[2] B. Primavera, "Superconducting optoelectronic circuits for neuromorphic computing," presented at the 16th Workshop on Low Temperature Electronics (WOLTE-16), Cagliari, Italy, June 2024. [Online]. Available: https://snf.ieeecsc.org/files/ieeecsc/slides/Primavera%20presentation.pdf

[3] S. Pellerano *et al.*, "Cryogenic CMOS for qubit control and readout," in *Proc. IEEE Custom Integrated Circuits Conf. (CICC)*, 2022, pp. 1-8. doi: 10.1109/CICC53496.2022.9772841. [Online]. Available: https://pure.tudelft.nl/ws/portalfiles/portal/122719163/Cryogenic_CMOS_for_Qubit_Control_and_ReadoutTaverne.pdf

[4] D. Yohannes *et al.*, "Materials and methods for fabricating superconducting quantum integrated circuits," U.S. Patent 11,991,935, May 21, 2024. [Online]. Available: https://patents.google.com/patent/US11991935B2/en

[5] N. Yurttagül *et al*., "Millikelvin Si-MOSFETs for quantum electronics," 2024, *arXiv:2410.01077*. [Online]. Available: https://arxiv.org/abs/2410.01077

[6] S. Kim *et al*., "A true single-phase 8-bit adiabatic multiplier," in *Proc. 38th Annu. Design Automation Conf. (DAC),* 2001, pp. 299-304. [Online]. Available: https://dl.acm.org/doi/10.1145/378239.378491

[7] E. DeBenedictis, "Scaling up reversible logic with HKI superconducting inductors," 2025, *arXiv:2504.09229*. [Online]. Available: https://arxiv.org/abs/2504.09229

[8] H. Xiang *et al*., "Hybrid stair-step Etch," U.S. Patent 9,741,563, Aug. 22, 2017 [Online]. Available: https://patents.google.com/patent/US9741563

[9] DARPA, "Low Temperature Logic Technology (LTLT)." [Online]. Available: https://www.darpa.mil/research/programs/low-temperature-logic-technology

[10] M. Brandl *et al*., "Cryogenic resonator design for trapped ion experiments in Paul traps," *Appl. Phys. B*, vol. 122, no. 5, p. 124, May 2016, doi: 10.1007/s00340-016-6430-z. [Online]. Available: https://arxiv.org/abs/1601.06699