Performance of Oscillator Ising Machines on Realistic MU-MIMO Decoding Problems

Jaijeet Roychowdhury (jr@berkeley.edu)
University of California, Berkeley
Joachim Wabnig
Nokia Bell Labs
K. Pavan Srinath
Nokia Bell Labs

Article

Keywords: Oscillator Ising Machines, Multi-User MIMO signals, decoding

Posted Date: September 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-840171/v1

Performance of Oscillator Ising Machines on Realistic MU-MIMO Decoding Problems

Jaijeet Roychowdhury[*], Joachim Wabnig[†] and K. Pavan Srinath[‡]

Abstract

Ising machines have recently been attracting attention due to their apparent ability to solve difficult combinatorial problems using analog operational principles. Oscillator Ising Machines (OIM) are especially attractive because they can be implemented easily as integrated circuits (ICs) in standard CMOS electronics. We explore the performance of OIM for decoding noisy Multi-User MIMO signals, a problem of considerable interest in modern telecommunications. Our results indicate that OIM-based decoding achieves error rates almost as good as the optimal Maximum Likelihood method, over a wide range of practical signal-to-noise (SNR) values. At high SNR values, OIM achieves ~20x fewer errors than LMMSE, a decoding method used widely in industry today. We also investigate the influence of parameter precision on decoding performance, finding that using 6 or more bits of precision largely retains OIM's advantages across all SNR values. We estimate that straightforward CMOS OIM implementations can easily solve MU-MIMO decoding problems in under 10ns, more than 100x faster than current industrial requirements. We conclude that oscillator Ising machines can be effective for real-world applications, possibly serving as an important enabler for future telecommunication standards. Our results and data provide guidance for designing hardware OIM prototypes specialized for MU-MIMO decoding.

1 Introduction

Combinatorial optimization (CO) is an enabling technology in many fields that impact modern life, including communication networks, drug/vaccine design, healthcare, delivery logistics, smart grids, *etc.*. However, practical problem sizes have kept outpacing available computational power by large margins. As a result, there has long been interest in ways to speed up CO.

Many practically-important CO problems are computationally difficult (*e.g.*, NP-complete [1]). Such problems can be recast [2] in a standard mathematical form, the Ising model [3]. The model is simply a weighted graph, *i.e.*, a collection of nodes/vertices and branches/edges between pairs of nodes, with each branch having a real-number weight. Each node (termed a "spin" in this context) is allowed to take one of two values, either 1 or -1. Associated with this graph is a cost function, called the Ising Hamiltonian, obtained by multiplying the weight of each branch by the values of the two spins it connects to, and summing over all branches. Ising Hamiltonians are sometimes interpreted as an "energy" associated with a given configuration of the spins, although in many situations (such as in this paper) they are merely a mathematical cost function, with no connection with energy in physics. The "Ising problem" is to find spin configurations with the minimum possible Hamiltonian value. For many practical problems, finding a spin configuration with a Hamiltonian close to the minimum possible is also useful.

Simple though the Ising problem is to state, it has proven (since its inception nearly 100 years ago) to have remarkable power. For example, the Ising model was developed, and first used, to explain how ferromagnetic behaviour in magnetic materials emerges [4, 5]. Since then, it was found that many difficult computational problems in various disciplines can be translated to the Ising problem. Examples include protein folding in biology, finding the optimal artificial neural network for a given set of training data, optimal strategies for playing the game Go, and many NP-hard/complete graph-theoretic problems, including the discrete maximum-likelihood (M.L.) problem — indeed, all

[*]EECS Department, University of California, Berkeley, CA, USA.
[†]Nokia Bell Labs Research, Cambridge, UK.
[‡]Nokia Bell Labs, France.

21 of Karp's list of NP-complete problems [1] have Ising incarnations [2]. (Note that we use M.L. as an acronym for Maximum Likelihood throughout this paper, in an attempt to avoid confusion with ML, commonly used for Machine Learning.)

Unsurprisingly, solving the Ising problem computationally is itself NP-hard/complete [6, 7]. However, over the last decade, a class of hardware Ising solvers (known as **Ising machines**) has emerged as a promising means to accelerate solutions to these classically difficult computational problems. The premise of Ising machines is that specialized hardware implementing the Ising computational model can solve many classes of NP-complete problems faster than classical algorithms (such as semidefinite programming [8] and simulated annealing [9, 10]) run on digital computers. Ising machines first came into prominence with the D-Wave quantum annealer and the Coherent Ising Machine (CIM). A D-Wave quantum annealer [11] with 2000 spins has been available commercially for several years, with a 5000-spin version recently announced. CIM [12, 13] with 2000 spins has been successfully demonstrated at NTT Research Labs [14], with larger systems under active development. Though without question *tours-de-force* of technology and science that have established the field of Ising machines and inspired follow-on technologies, D-Wave quantum annealers and CIM are not ideally suited for all applications, being physically large, expensive, and difficult to miniaturize or scale to larger problems. For example, the CIM/DOPO scheme involves pulsed lasers and frequency doubling crystals, and is about the size of a rack for a size-2000 machine [14]; D-Wave machines require an operating temperature under 80mK, are the size of several large racks, and are said to cost in the range of $15M.

In 2016, Wang and the first author discovered that networks of oscillators can solve Ising problems [15]. In their scheme, each of the N binary variables (spins) of the Ising problem is implemented by an oscillator. The information needed to find a solution of the Ising problem is encoded in the phase (relative time delay) of each oscillator. The oscillators are coupled together in a network (Figure 1), with coupling strengths that correspond to the weights in the Ising Hamiltonian. They proved that such systems have an "energy function" (more precisely, Lyapunov function) that matches a given Ising problem's energy landscape; and that the Lyapunov function closely approximates the Ising Hamiltonian when the oscillators' phases are binarized, using a mechanism called sub-harmonic injection locking (SHIL). They also showed, crucially, that such Oscillator Ising Machines (OIMs) naturally find troughs of their Lyapunov landscapes, and that this innate property can be leveraged to find good (*i.e.*, near minimum-Hamiltonian) solutions of Ising problems. This ability stems from collective behaviour involving two types of *injection locking*, a generic synchronization-inducing property of oscillators.

In [16, 17], OIM was evaluated on the widely-used G-set [18, 19], comprising 51 benchmark problems for the NP-complete MAX-CUT problem, with sizes ranging from 800 to 3000 spins. OIM matched the best results from 4 other algorithms on 29 of the problems, finding better results on 17. Prototype electronic hardware implementations of OIM have been built that find good solutions of Ising problems in milliseconds [17, 20]. From a practical deployment perspective, OIM has compelling advantages over previous Ising machines. It can be implemented using plain electronics (in particular, standard CMOS in non-cutting-edge technologies) in very small form factors, especially compared to CIM and quantum annealers. OIMs are orders of magnitude cheaper than prior Ising machines – this, together with small size and easy mass production, greatly broadens the potential applicability of Ising machines. Because electronic oscillator frequencies can be calibrated easily, the impact of ever-present variability in electronic components can essentially be eliminated. Importantly, OIM is a purely classical — not quantum — scheme that is immediately practical; its power derives primarily from system-level interactions of oscillators, not from device technology *per se*. Hence it can use larger/older/cheaper technology nodes, while being able to fully reap the benefits of smaller ones and, indeed, make use of novel nanodevice technologies

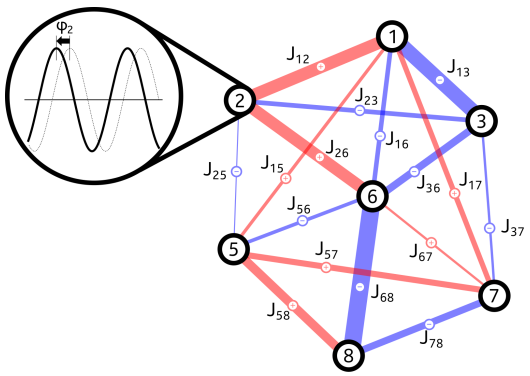

Fig. 1: An example network of 8 oscillators with various positive or negative coupling strengths J_{ij}. Each oscillator can be characterized by its phase relative to a reference oscillator (*e.g.*, oscillator 8).

that become practically viable. (Note that D-Wave's quantum annealing based Ising machine has not shown quantum advantage on practical problems yet; only on carefully designed abstract ones [21].)

In this paper, we report the performance of OIM on an important problem in wireless communications, the MU-MIMO (Multi-User Multiple-Input-Multiple-Output) decoding problem. As shown in Figure 2, modern wireless communication settings involve multiple users with single/multiple transmit antennas, using the same resources (time and frequency) to transmit to a receiver equipped with multiple receive antennas. As a result, each received signal consists of a noisy superposition of several users' transmitted symbols. Recovering the originally-sent symbols from received signals involves solving a hard CO problem (the MU-MIMO decoding problem [22, 23]) to infer the most likely set of transmitted symbols, given the set of signals received. Solving exactly for the most likely transmitted symbols, *i.e.*, the M.L. (Maximum Likelihood) solution, is too computationally expensive to be practical; hence, heuristic methods that use much less computation, such as LMMSE ("linear minimum mean-square error"), are universally used even though they do not recover transmitted symbols as accurately as M.L..

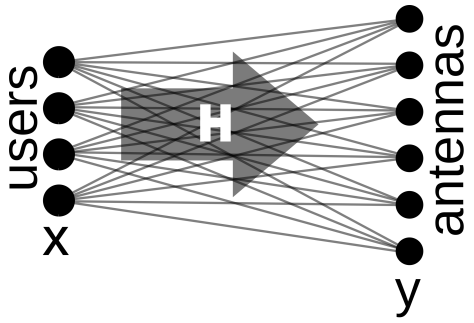

Fig. 2: An illustration of a Multiple User Multiple In Multiple Out (MU-MIMO) setup. Multiple users transmit their data, $\vec{x}$, to a receiver with multiple antennas, where the signal $\vec{y}$ is measured. Transmission occurs over several paths, characterized by the channel matrix H.

From a hardware perspective, it is highly desirable for MU-MIMO decoders to be physically small and inexpensive so that they are practical for size- and cost-sensitive products, such as cellular basestations. For example, in modern cellular network installations, space constraints dictate that units must fit in a fraction of a standard rackmount unit, even in configurations where they are implemented remotely. These considerations, which place strong constraints on MU-MIMO decoding technologies, favour OIM over other Ising machine schemes.

2 Results

We evaluate OIM (in simulation) on an extensive set of 550,000 MU-MIMO decoding problems (for a 16×64 QPSK setup) and compare our results against M.L. and LMMSE. Because of their suitability for CMOS IC fabrication, we use ring oscillators [24–26] for OIM's spins. We find that OIM achieves symbol error rates (SER) that are very close to optimal results from M.L.; both are more than an order of magnitude better than those from LMMSE at high SNR values (*i.e.*, for the more challenging problems). These results indicate the promise of OIM for solving real, practical problems and provide motivation for integrated circuit (IC) realizations of OIM specialized for MU-MIMO decoding, which may benefit 6G and future standards. Towards this end, we also evaluate ring-oscillator OIM with the coupling coefficients J_{ij} quantized to varying numbers of discrete levels, since quantized coupling is typically necessary for chip implementation. We find that using 256 levels (8 bits of quantization) results in essentially no performance degradation, while 64 levels (6 bits) leads to acceptable performance, especially when the SNR (Signal-to-Noise Ratio) of received MU-MIMO signals is low. Our results, and conclusions about the promise of OIM for MU-MIMO decoding, differ considerably from those reported in [27], wherein poor BER performance observed for OIM (as well as CIM) motivates the authors to devise a Regularized Ising (RI) formulation that is reported to improve BER performance, but still not to the extent of approaching our results.

2.1 Performance of OIM on practical MU-MIMO decoding problems

We evaluate OIM on 11 sets of test problems. Each set corresponds to a specific SNR (signal-to-noise ratio) at the receiving antennas; the 11 sets of problems have SNR (in dB, *i.e.*, $10\log_{10}(\text{actual SNR})$) varying from -1 to 9. Problem sets with lower SNRs generally have more symbol errors; those with higher SNRs have fewer. For each SNR value, the test problem set consists of 1000 different channel matrices H, for each of which 50 pairs of transmitted symbol vectors $\vec{x}$ and received signal vectors $\vec{y}$ are available. Thus, there are 50,000 decoding problems for each SNR value, or a total of 550,000 problems in all.

The channel matrices H for the problems, generated following [28, Section IV-A], capture correlations between users in a fading environment more realistically than the commonly-used independent,

identically distributed (i.i.d.) Rayleigh fading model in the literature — in essence, [28] takes into account the fact that users closer to one another tend to have more-correlated channels than further-away users. The data symbol vector $\vec{x}$ for each problem was generated randomly, each being a QPSK (Quadrature Phase Shift Keying, [29]) symbol chosen independently with a probability of 0.25. Synthetically-generated additive white Gaussian noise (AWGN) $\vec{w}$ was added to form the received signal, as $\vec{y} = H\vec{x} + \vec{w}$.

The MU-MIMO modulation scheme from which the problems are derived is QPSK with 16 users transmitting independently. Each QPSK symbol, which can take 4 values, is encoded as 2 binary symbols/spins; thus, $\vec{x}$ consists of 16 pairs of binary spins stacked one over the other. 64 receiving antennas, each capable of producing a complex number in the QPSK constellation, are used. The problem thus becomes identical to a BPSK one (converted to all-real matrices/vectors) with $N_t = 32$ and $N_r = 128$ (using terminology from Sec. 4.3, below). More precisely, the transmitted symbol vector $\vec{x}$ for each problem consists of $N_t = 32$ binary symbols, while each received signal vector $\vec{y}$ consists of $N_r = 128$ real numbers; *i.e.*, $\vec{x} \in \{\pm 1\}^{32}$, $\vec{y} \in \mathbf{R}^{128}$, $H \in \mathbf{R}^{128\times 32}$, $\hat{H} \in \mathbf{R}^{128\times 33}$ and $J \triangleq -\hat{H}^T\hat{H} \in \mathbf{R}^{33\times 33}$. QPSK encoding needs to be taken into account in the calculation of symbol error rates (SER), *i.e.*, any change to a symbol (a single or double bit error) should be counted as a single "symbol error".

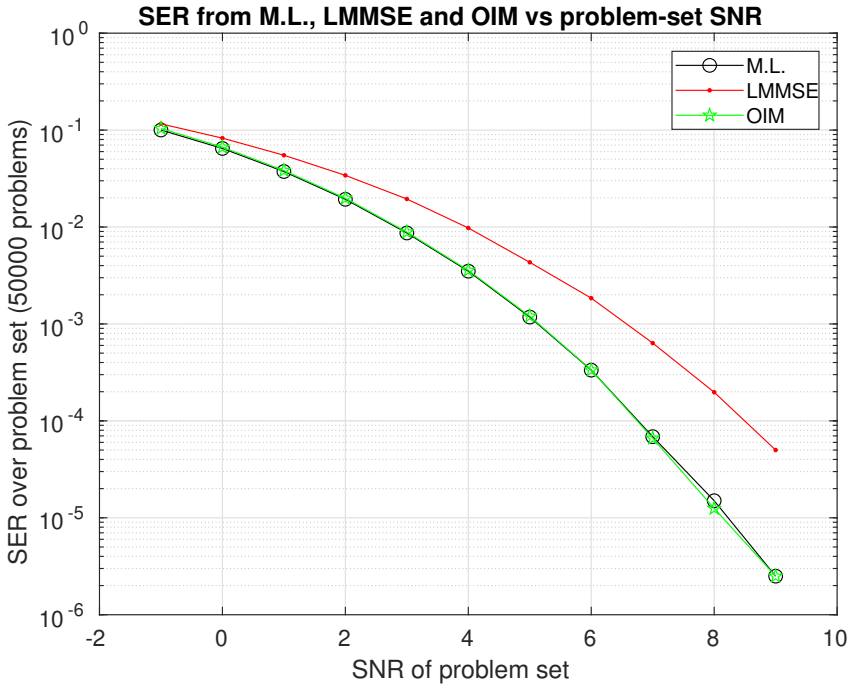

Fig. 3: Decoding performance of OIM *vs.* other methods. The symbol error rate (SER) is shown as a function of signal-to-noise ratio (SNR) for M.L., LMMSE and OIM decoders (data given in Table 1). OIM decoding closely matches the performance of the M.L. decoder over the whole range of signal to noise ratios and surpasses the LMMSE decoder at high SNR by more than an order of magnitude.

Figure 3 and Table 1 shows the average SER (over all problems in each SNR set) from ML, LMMSE, and OIM. For OIM, SER numbers were obtained by numerical simulation of the generalized Kuramoto equations [16, 17], using a C++ implementation of the code in [15].

Examining the data reveals several interesting features:

1. For high SNR values, the absolute number of errors over all 50000 test cases is very low. For example, at SNR=9, the optimal result from Maximum Likelihood features only two bit errors (out of $32 \times 50000 = 1,600,000$ possibilities). At lower SNR values, there are many more bit errors, *e.g.*, at SNR=-1, there are $160,240$ bit errors. It is helpful to keep these absolute numbers of bit errors in mind when assessing performance; high-SNR cases are much more challenging than low-SNR ones.

2. The performance of LMMSE varies from about 16% worse than M.L. at the lowest SNR of -1, to almost $20\times$ worse for the SNR=9 set.

3. In contrast, **SER numbers from OIM using ring oscillators are very close to M.L. for all SNR sets**, *i.e.*, not more than 4% over M.L., which is significantly better than LMMSE for every SNR set. At higher SNRs (the more challenging cases with few bit errors) in particular, OIM does particularly well, *e.g.*, matching M.L. exactly at SNR=9; indeed, each of 11 runs of the SNR=9 set of 50,000 problems yielded exactly the same SER of 2.5×10^{-6}. Interestingly, for SNR=8 and 7, OIM features *fewer* bit errors than M.L.; though surprising, this is possible since OIM features a judicious amount of noise/randomness in its operation [17].

2.2 Effect of coupling quantization on OIM performance

As noted in the Introduction, hardware implementations of OIM offer considerable promise on account of miniaturizability/small size, low cost, *etc.*, compared to other prevailing Ising machine schemes. For integrated circuit implementation, it is usually necessary to quantize continuous-valued couplings (J_{ij}); in hardware, these couplings are implemented using a set of B resistors, where B is the number of bits used to choose a resistance/coupling value from one of $L = 2^B$ quantized possibilities. While values as high as $B \sim 12$ can be realized in practice, lowering B makes IC design and fabrication significantly easier. Below, we examine the effect of changing B on the SER performance of OIM.

Figure 4 shows the absolute values of the 33×33 coupling matrices for one of the 50,000 problems from the SNR 9 set. As can be seen, the entries in the last row and column (which stem from the "external magnetic field" terms $H\vec{y}$), are about a factor of 4 larger than the other values in the matrix. Similar patterns are seen in the coupling matrices of all the problems. This suggests that from an accuracy standpoint, it is advantageous to use one set of quantized values for the last row and column, and another set for the remainder of the matrix — this is easy to implement in IC hardware. We adopt this quantization scheme, *i.e.*, with the same B but different sets of resistance values for the two sets.

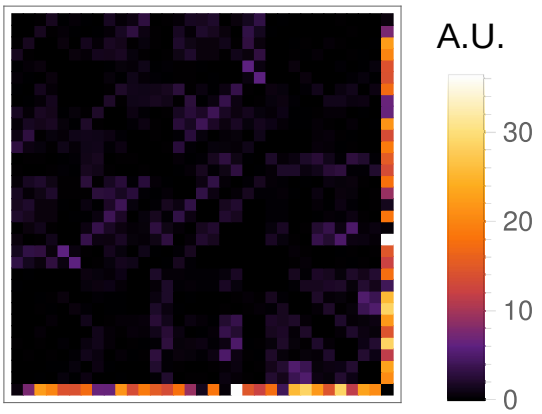

Fig. 4: Absolute value of an example coupling matrix used in the decoding (for SNR 9; all the other SNR sets are virtually identical in pattern). The terms in the last row and column are realized as coupling to the last oscillator and are typically much larger than the remaining ones. This justifies using a different quantization scale for the last row and column.

Figure 5 (data in Tables 2 and 3) shows ring-oscillator-based OIM's performance with quantized couplings; B, the number of bits used for quantization, is varied from 9 down to 4. It can be seen that SER performance degradation (over Maximum Likelihood with no quantization) is essentially negligible for 9 and 8 bits of quantization. Using 6 bits of quantization still yields significant improvements over LMMSE across all the problems, while $B = 5$ remains competitive against LMMSE. These results, indicating that implementing OIM in IC hardware is practical, can help guide design tradeoffs.

3 Discussion

The results presented above show the suitability of oscillator Ising machines for real-world decoding tasks in telecommunications. For the problem sizes investigated, we achieve decoding performance identical or close to the optimal Maximum Likelihood decoder. Our results differ considerably from those of Singh *et. al.* [27, Appendix D]. Their implementation of OIM is reportedly unable to achieve bit error rates (BER) less than about 2×10^{-2} at any value of SNR. Singh *et. al.* also report similar behaviour from the Coherent Ising Machine, motivating them to devise regularization schemes (aided by cheaply-computed approximate solutions) to improve the performance of CIM and OIM. Even

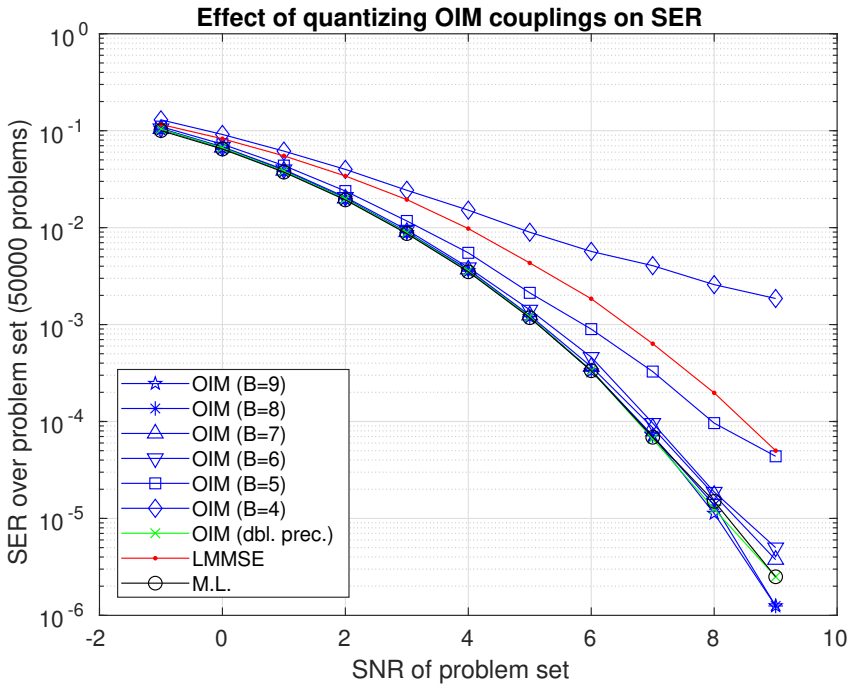

Fig. 5: Symbol error rate as a function of signal-to-noise ratio for different quantizations of the OIM coupling weights, compared to OIM with coupling weights in double precision, LMMSE and M.L. (data in Tables 2 and 3). OIM with 9-bit through 6-bit quantizations yields SERs from near-optimal to acceptable. Decoding performance deteriorates quickly below 6 bits of quantization.

with regularization, they report a BER at SNR 9 for OIM that is more than an order of magnitude larger than from the Sphere Decoder, an implementation of M.L. [27, Fig. 15, Appendix D]. In contrast, our results (Figure 3 and Table 1), which achieve SERs in the range $\sim [10^{-6}, 10^{-1}]$ on a large set of realistic benchmark problems spanning a range of SNR values actually encountered in practice, are within 4% of M.L.'s across all SNR values.

Modern communication systems operate at high data rates, requiring a decoding problem to be solved in, *e.g.*, 1μs (this is an aggressive decoding time target, applicable, *e.g.*, to 6G with enhanced data rate requirements). From Figure 8 below, it is apparent that OIM solves the decoding problem in well under 10 cycles of oscillation. CMOS ring oscillators with frequencies in excess of 1GHz are easily fabricated in well-established, widely used, industrial technologies today — *e.g.*, more than 15 years ago, oscillation frequencies of 3.5GHz were achieved in 65nm CMOS technology [30].

The above considerations suggest that decoding performance very similar to M.L. can easily be achieved by OIM in under 10ns, using today's hardware technologies and circuits — this is $100\times$ faster than 1μs, itself an aggressive target by current standards. Note that the complexity of M.L. varies for different data samples even within the same MU-MIMO configuration, making practical implementation difficult if fixed decoding delay, within reasonable limits of computation, is required. OIM's 10ns decoding times would be a significant improvement, and a powerful enabler for future standards such as 6G, which stipulates much greater data rates than current 5G specifications. For example, while current 5G standards support 12 transmission layers (total number of transmit antennas for all simultaneously transmitting users), 6G is expected to expand this by a factor of ~5, *e.g.*, to 64 transmission layers. Our results indicate that OIM will easily be able to handle such expansion. Our work thus provides concrete motivation for building and demonstrating CMOS IC implementations of OIMs specialized to solve the MU-MIMO decoding problem. If such hardware designs achieve results similar to this work, we believe it will be the first demonstration of an Ising machine solving an important real-world problem competitively.

4 Methods

4.1 The Ising Hamiltonian

The Ising Hamiltonian is obtained by multiplying the weight of each branch by the values of the two spins it connects to and summing over all branches, *i.e.*,

$$C(s_1,\cdots,s_n) \triangleq -\frac{1}{2}\sum_{i,j=1}^{n} J_{ij}\,s_i s_j, \tag{1}$$

where $s_i \in \{-1,\ +1\}$, $i = 1,\cdots,n$, are the n spins, with the weights J_{ij} obeying $J_{ij} = J_{ji}$ and $J_{ii} = 0$. Note that an alternative version of the Ising Hamiltonian uses so-called "external magnetic field" terms comprised of a linear combination of the spins, *i.e.*,

$$\tilde{C}(s_1,\cdots,s_n) \triangleq -\left[\frac{1}{2}\sum_{i,j=1}^{n} J_{ij}\,s_i s_j + \sum_{i=1}^{n} B_i s_i\right]. \tag{2}$$

By adding one more spin, $s_{n+1} \equiv 1$ and defining

$$J_{n+1,i} = J_{i,n+1} \triangleq B_i, \quad i = 1,\cdots,n, \quad \text{with } J_{n+1,n+1} \triangleq 0, \tag{3}$$

it is easily shown that (2) is equivalent to (1), *i.e.*,

$$\tilde{C}(s_1,\cdots,s_n) \equiv C(s_1,\cdots,s_n,s_{n+1}=1). \tag{4}$$

Thus the form (1), which we use here, is general enough to capture external magnetic field terms.

4.2 Oscillator Ising Machines

As mentioned in the Introduction and illustrated in Figure 1, an OIM (Oscillator Ising Machine) is a networked (*i.e.*, coupled) group of oscillators. If properly designed, such a system can serve as an effective Ising machine due to collective behaviour enabled by *injection locking*, a nonlinear synchronization phenomenon generically exhibited by oscillators. In fact, Oscillator Ising Machines embody the power of synchronization as an enabler for "complex, self-organizing systems, where vast numbers of components interact simultaneously", as prophesied by Steven Strogatz [31, 32] almost 20 years ago. Below, we outline the key ideas behind making networked oscillators solve Ising problems.

An oscillator (more precisely, a self-sustaining, asymptotically orbitally stable nonlinear oscillator [33]) is anything that generates periodic signals "on its own". Examples abound in engineering and nature, from grandfather clocks to flashing fireflies to LC and ring oscillators in electronics. For example, the waveform on the bottom left of Figure 6 depicts the output of an undisturbed sinusoidal oscillator, with natural angular frequency ω_0; though in many practical oscillators, the periodic waveform generated is not sinusoidal but is often, *e.g.*, square- or sawtooth-like in shape. Under the right circumstances, if an oscillator is disturbed by an external input with a frequency ω_1 close to ω_0, as illustrated by the waveform at the top left of Figure 6, it will spontaneously change its natural frequency to exactly match that of the external input. When this happens, moreover, the external input and the oscillator's output waveform become synchronized ("phase locked") to each other, as illustrated at the right of Figure 6. This phenomenon, which has a long and rich history dating back to at least 1672 [34], is known today as injection locking, or more precisely, as fundamental-harmonic injection locking (FHIL). It can be shown [35] that if FHIL occurs, the phase difference between the injection and the oscillation waveforms will be a single fixed number, *i.e.*, there cannot be two or more different phases at which the waveforms lock stably. Moreover, if the difference between the frequency of the external input and the oscillator's natural frequency is small, the level of external injection required to induce locking is typically also small, *e.g.*, often much smaller than

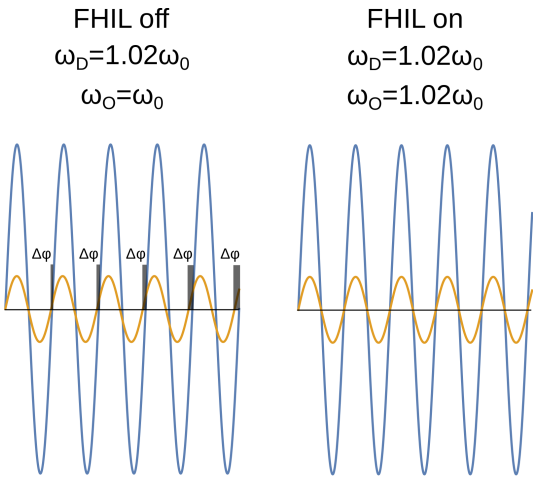

Fig. 6: Illustration of fundamental-harmonic injection locking (FHIL). The interaction of a self-sustaining nonlinear oscillator (with free-running frequency $\omega_O = \omega_0$) with a driving signal of slightly higher frequency ($\omega_D = 1.02\omega_0$), leads to a shift in frequency of the oscillator ($\omega_O = \omega_D$) and locking of the oscillators' phases.

the natural oscillation it influences [36].

FHIL is only one possible type of injection locking; interesting synchronization behaviours also manifest when the injected signal's frequency is near an integral multiple of the oscillator's natural frequency ω_0. For example, if the injection frequency is close to twice the natural frequency, *i.e.*, $\omega_1 \simeq 2\omega_0$, frequency- and phase-locking can also occur; this is called 2-SHIL (2^{nd} sub-harmonic injection locking). In 2-SHIL, the oscillator changes its natural frequency to precisely *half* of ω_1; the resulting waveform is also phase locked to the injection signal, as illustrated in Figure 7. A key difference between FHIL and 2-SHIL is that in the latter, there are *two* possible values of relative phase between the injection and oscillation waveforms at which (stable) lock can occur [35]; moreover, these two phase locks are always separated by 180°. In OIM, the two 180°-separated phase locks in 2-SHIL correspond to Ising spins $+1$ and -1. FHIL and 2-SHIL are both crucial for making networked oscillator systems function as Ising machines.

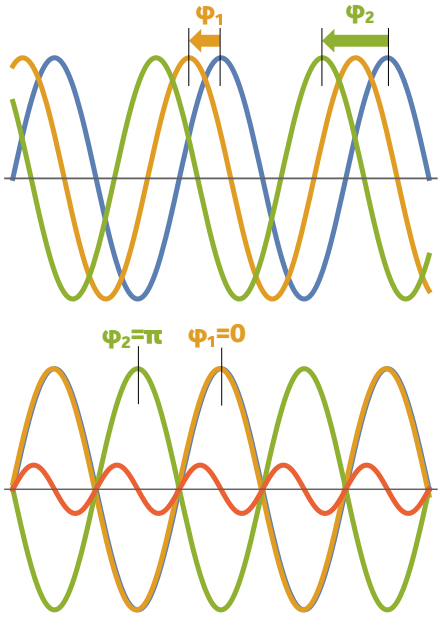

Fig. 7: Illustration of 2^{nd}-subharmonic injection locking (2-SHIL). Without the SUNC signal, interacting self-sustaining oscillators settle in a fixed phase relationship according to their coupling. When a 2-SHIL signal of sufficient amplitude is introduced, the phases lock at either π or 0.

For a system of coupled oscillators, such as the one shown in Figure 1, the external injection to each oscillator is a sum of the perturbations from each neighbour to which it is coupled. In an electronic context, with couplings implemented by resistors, the sum of currents entering the oscillator through the coupling resistors serves as its external injection. If the frequencies of the oscillators are close enough to each other, FHIL will make all lock to a common frequency [37]. For OIM, however, an additional common external signal, of fixed frequency set to about twice that of the average natural frequency of the oscillators, is also injected into each oscillator. This injection, termed SYNC, is used to induce 2-SHIL, *i.e.*, phase lock at one of two binary values separated by 180°. If the amplitude of SYNC is low (or 0), then FHIL between the oscillator dominates; if it is high, then 2-SHIL, ie, phase binarization, dominates. Making the two types of injection lock "compete", by increasing and decreasing the amplitude of SYNC periodically, is an important facet of OIM's operation.

A useful mathematical model for the coupled oscillator system with SYNC injection is the generalized Kuramoto equation [38],

$$\frac{1}{\omega_0}\frac{d\phi_i}{dt} = K_s\, z_s\left(2\phi_i(t)\right) + \sum_{j=1}^{N} J_{ij}\, z_c\left(\phi_i(t) - \phi_j(t)\right), \tag{5}$$

shown for the simplified case where all oscillators have the same natural (angular) frequency, ω_0. N is the number of oscillators in the system; $\phi_i(t)$ is the phase of the i^{th} oscillator; $z_c(\cdot)$ is a 2π-periodic function that captures the FHIL dynamics of the system, with its shape determined by the nature of the oscillators, the shape of oscillation waveforms, *etc.*; $z_s(\cdot)$, is, similarly, a function that captures 2-SHIL dynamics; with K_s represents the amplitude of the SYNC signal; and J_{ij} is the coupling between the i^{th} and j^{th} oscillator, the same as in (1). If K_s, the amplitude of SYNC, is kept constant with time, it can be shown that the phases in (5) always evolve to naturally minimize the Lyapunov function

$$L(\phi_1, \cdots, \phi_N) \triangleq \sum_{i=1}^{N}\sum_{k=1}^{N}\left[\frac{I_s(2\phi_i) + I_s(2\phi_k)}{2N} + J_{ik}\, I_c(\phi_i - \phi_k)\right], \tag{6}$$

where $I_s(\cdot)$ and $I_c(\cdot)$ are integrals of $f_s(\cdot)$ and $f_c(\cdot)$, respectively [38]. Such minima, reached naturally for any fixed value of K_s, are *local* minima. The importance of varying K_s periodically between low and high values lies in that it enables the system to progress to lower and lower local minima. Crucially, it can be shown that when K_s is high, the Lyapunov function approximates the Ising Hamiltonian. Thus, the coupled oscillator system, with periodic variation of K_s, evolves to find good solutions of the Ising problem.

4.3 Casting the MU-MIMO decoding problem in Ising form

A succinct development of the relation between the MU-MIMO and Ising problems follows (a more detailed exposition can be found in [23]). Given a BPSK MU-MIMO system with N_t transmitters (users) and N_r receivers, define a vector of transmitted symbols to be

$$\vec{x} = \left[x_1, \cdots, x_{N_t}\right]^T, \tag{7}$$

where $x_i \in \{\pm 1\}$ are N_t simultaneously transmitted symbols. Define $H \in \mathbf{R}^{N_r \times N_t}$ to be the channel transmission matrix, and $\vec{y} \in \mathbf{R}^{N_r}$ to be the vector of received signals. In an ideal situation, the received signal would be $\vec{y} = H\vec{x}$. However, in reality, the received signal deviates from this ideal due to corruption by noise, *i.e.*,

$$\vec{y} = H\vec{x} + \vec{w}, \tag{8}$$

where $\vec{w}$ represents additive white Gaussian noise (AWGN).

An optimal solution of the MU-MIMO decoding problem, *i.e.*, the Maximum Likelihood (M.L.) solution, is a transmitted symbol vector $\vec{x}^*$ that minimizes the error from the ideally-received signal, *i.e.*

$$\vec{x}^* = \underset{\vec{x} \in \{\pm 1\}^{N_t}}{\arg\min} \|\vec{y} - H\vec{x}\|^2. \tag{9}$$

To frame the MU-MIMO decoding problem in Ising form, we first augment the number of transmitted symbols by one to define the spin vector

$$\vec{s} \triangleq \left[\underbrace{x_1}_{s_1}, \cdots, \underbrace{x_{N_t}}_{s_{N_t}}, \underbrace{1}_{s_{N_t+1}}\right]^T = \begin{bmatrix}\vec{x} \\ 1\end{bmatrix}, \tag{10}$$

where we use the terminology $s_i \equiv x_i$, $i = 1, \cdots, N_t$ to emphasize that the transmitted symbols serve as spins for the Ising version of the problem. Note that the last spin of $\vec{s}$, s_{N_t+1}, is fixed at 1. Using this, define the Ising Hamiltonian to be

$$C_I(\vec{s}) \triangleq -\frac{1}{2} \sum_{k=1}^{1+N_t} \sum_{j=1}^{1+N_t} J_{kj}\, s_k s_j. \tag{11}$$

Next, define the matrices

$$\hat{H} = \left[H, \vec{y}\right] \in \mathbf{R}^{N_t \times (N_t+1)}, \quad J = -\hat{H}^T \hat{H} \in \mathbf{R}^{(N_t+1)\times(N_t+1)}. \tag{12}$$

With the above definitions, it is easily shown that the Ising Hamiltonian $C_I(\vec{s})$, as given by (11), equals the error being minimized by M.L. in (9), *i.e.*,

$$C_I(\vec{s}) = \|H\vec{x} - \vec{y}\|^2, \tag{13}$$

where J_{kj} in (11) is the $(k,j)^{\text{th}}$ element of J in (12). (13) implies that solving the Ising problem, *i.e.*, finding

$$\vec{s}^* = \underset{\vec{s} \in \{\pm 1\}^{N_t+1}}{\arg\min}\ C_I(\vec{s}), \quad \text{subject to } s_{N_t+1} = 1, \tag{14}$$

is equivalent to finding the Maximum Likehood solution (9) of the MU-MIMO decoding problem [22]. Because the Hamiltonian remains unchanged when all spins are flipped, any solution with $s_{N_t+1} = -1$ is easily converted to one with $s_{N_t+1} = 1$, simply by flipping all the spins.

4.4 Simulating OIM

The results in Sec. 2 were obtained by simulating (5). The functions $z_c(\cdot)$ and $z_s(\cdot)$ were based on circuit simulations of a ring oscillator circuit. The numerical simulation algorithm in [15], recoded in C++ for efficiency and usability, was used for the simulations. The SYNC signal's amplitude was varied from low to high once over the length of the simulations (about 5 oscillation cycles). Each SNR set (50000 problems) was run in parallel on a 40-processor Linux system with Intel Xeon E5-2670 CPUs running at 2.5GHz; each problem required about 3s of wall time (single threaded) to complete.

Figure 8 shows a sample phase evolution plot for one of the problems with SNR=-1, started with random initial phases. The time t is in units of $\frac{1}{\omega_0}$, *i.e.*, one cycle of oscillation corresponds to 2π such time units, with each simulation run corresponding to about 5 cycles of oscillation. Synchronization of groups of oscillators due to FHIL can be seen in the regions $t < 10$ and $t > 25$ or so, when the amplitude of SYNC is low. For t roughly in the range $[12, 22]$, phase binarization due to SHIL can be clearly seen, with phases clustering into two groups separated by π. Note that phases should be interpreted modulo 2π, *e.g.*, a phase of $-\pi$ is the same as π, and a phase of 0 is the same as 2π. The final Ising solution is obtained simply by thresholding the phases at the end of the simulation to the nearest 2π shift of 0 (spin = -1) or π (spin = $+1$), followed by flipping all the spins if the last spin is

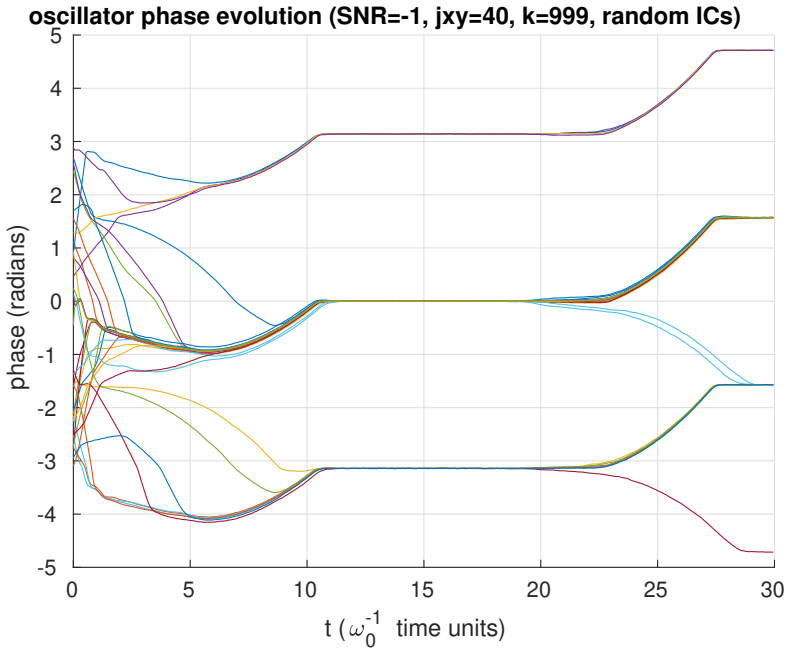

Fig. 8: Phase evolution of all 33 oscillators for an example problem from the SNR=−1 set. Time is measured in units of $1/\omega_0$, *i.e.*, one cycle of oscillation corresponds to $t = 2\pi$. Once SYNC is ramped up, the oscillators settle to "binarized" solutions over just a few oscillations.

−1.

References

[1] R. M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103. Springer, 1972.

[2] Andrew Lucas. Ising formulations of many NP problems. *Frontiers in Physics*, 2:5, 2014.

[3] E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, 1925.

[4] Stephen G. Brush. History of the Lenz-Ising Model. *Rev. Mod. Phys.*, 39(4):883–893, Oct 1967.

[5] S.M. Bhattacharjee and A. Khare. Fifty Years of the Exact solution of the Two-dimensional Ising Model by Onsager. *arXiv preprint cond-mat/9511003*, 1995.

[6] Johannes Schneider and Scott Kirkpatrick. *Stochastic Optimization*. Springer Science & Business Media, 2007.

[7] Francisco Barahona. On the computational complexity of Ising spin glass models. *Journal of Physics A: Mathematical and General*, 15(10):3241, 1982.

[8] Lieven Vandenberghe and Stephen Boyd. Semidefinite Programming. *SIAM Review*, 38(1):49–95, 1996.

[9] Bernd Gärtner and Jirí Matousek. *Approximation Algorithms and Semidefinite Programming*. Springer, 2014.

[10] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *SCIENCE*, 220(4598):671–680, 1983.

[11] Zhengbing Bian, Fabian Chudak, William G Macready, and Geordie Rose. The Ising model: teaching an old problem new tricks. *D-Wave Systems*, 2, 2010.

[12] Z. Wang, A. Marandi, K. Wen, R. L. Byer and Y. Yamamoto. Coherent Ising machine based on degenerate optical parametric oscillators. *Physical Review A*, 88(6):063853, 2013.

[13] Y. Haribara, S. Utsunomiya and Y. Yamamoto. Computational principle and performance evaluation of Coherent Ising Machine based on degenerate optical parametric oscillator network. *Entropy*, 18(4):151, 2016.

[14] T. Inagaki, Y. Haribara, K. Igarashi, T. Sonobe, S. Tamate, T. Honjo, A. Marandi, P. L. McMahon,

T. Umeki, K. Enbutsu and others. A Coherent Ising machine for 2000-node Optimization Problems. *Science*, 354(6312):603–606, 2016.

[15] Tianshi Wang and Jaijeet Roychowdhury. Oscillator-based Ising Machine. *arXiv:1709.08102*, 2017.

[16] Tianshi Wang and Jaijeet Roychowdhury. OIM: Oscillator-based Ising Machines for Solving Combinatorial Optimisation Problems. In *Proc. UCNC*, LNCS sublibrary: Theoretical computer science and general issues. Springer, June 2019. Preprint available at arXiv:1903.07163 [cs.ET].

[17] Tianshi Wang, Leon Wu, Parth Nobel, and Jaijeet Roychowdhury. Solving combinatorial optimisation problems using oscillator based Ising machines. *Natural Computing*, pages 1–20, April 2021.

[18] P. Festa, P. M. Pardalos, M. G. C. Resende, and C. C. Ribeiro. Randomized heuristics for the MAX-CUT problem. *Optimization methods and software*, 17(6):1033–1058, 2002.

[19] The G-set benchmarks for MAX-CUT. Website: http://grafo.etsii.urjc.es/optsicom/maxcut.

[20] T. Wang, L. Wu and J. Roychowdhury. New Computational Results and Hardware Prototypes for Oscillator-based Ising Machines. In *Proc. IEEE DAC*, pages 239:1–239:2, 2019.

[21] Tameem Albash and Daniel A. Lidar. Demonstration of a scaling advantage for a quantum annealer over simulated annealing. *Phys. Rev. X*, 8, Jul 2018.

[22] Mohamed Oussama Damen, Hesham El Gamal, and Giuseppe Caire. On maximum-likelihood detection and the search for the closest lattice point. *IEEE Transactions on Information Theory*, 49(10):2389–2402, 2003.

[23] M. Kim, D. Venturelli, and K. Jamieson. Leveraging quantum annealing for large MIMO processing in centralized radio access networks. In *Proceedings of the ACM Special Interest Group on Data Communication*, pages 241–255. ACM, 2019.

[24] T. Wang and J. Roychowdhury. Design Tools for Oscillator-based Computing Systems. In *Proc. IEEE DAC*, pages 188:1–188:6, 2015.

[25] J. Roychowdhury. Boolean Computation Using Self-Sustaining Nonlinear Oscillators. *Proceedings of the IEEE*, 103(11):1958–1969, Nov 2015.

[26] S. Srivastava and J. Roychowdhury. Analytical Equations for Nonlinear Phase Errors and Jitter in Ring Oscillators. *IEEE Transactions on Circuits and Systems*, 54(10):2321–2329, October 2007.

[27] Abhishek Kumar Singh, Kyle Jamieson, Davide Venturelli, and Peter McMahon. Ising Machines' Dynamics and Regularization for Near-Optimal Large and Massive MIMO Detection. *arXiv:2105.10535v1*, May 2021.

[28] Mathieu Goutay, Fayçal Ait Aoudia, and Jakob Hoydis. Deep Hypernetwork-based MIMO Detection. In *21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5. IEEE, 2020.

[29] Debabrata Saha and Theodore G Birdsall. Quadrature-quadrature phase-shift keying. *IEEE Transactions on Communications*, 37(5):437–448, 1989.

[30] M. Grozing, B. Phillip, and M. Berroth. CMOS ring oscillator with quadrature outputs and 100 MHz to 3.5 GHz tuning range. In *Proc. IEEE ESSCIRC*, pages 679–682, 2003.

[31] Steve Nadis. Complex systems: All together now. *Nature*, 421(6925):780–783, 2003.

[32] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.

[33] A. Demir, A. Mehrotra, and J. Roychowdhury. Phase Noise in Oscillators: a Unifying Theory and Numerical Methods for Characterization. *IEEE Trans. Ckts. Syst. – I: Fund. Th. Appl.*, 47(5):655–674, May 2000.

[34] Christian Huygens. *Horologium Oscillatorium*. Apud F. Muget, Paris, France, 1672. (Observations of injection locking between grandfather clocks).

[35] A. Neogy and J. Roychowdhury. Analysis and Design of Sub-harmonically Injection Locked Oscillators. In *Proc. IEEE DATE*, Mar 2012.

[36] Y. Wan, X. Lai, and J. Roychowdhury. Understanding injection locking in negative-resistance lc oscillators intuitively using nonlinear feedback analysis. In *Proc. IEEE CICC*, pages 729–732, 18-21 Sept. 2005.

[37] Jaijeet Roychowdhury and Palak Bhushan. Hierarchical abstraction of weakly coupled synchronized oscillator networks. *International Journal for Numerical Methods in Engineering*, 2015.

[38] T. Wang and J. Roychowdhury. OIM: Oscillator-based Ising Machines for Solving Combinatorial Optimisation Problems. In *arXiv:1903.07163*, 2019.

A Data (Tables)

SNR	SER: M.L.	LMMSE		OIM (ring osc)	
		SER	%>M.L.	SER	%>M.L.
-1	1.0015E-01	1.1605E-01	15.88%	1.0392E-01	3.76%
0	6.4646E-02	8.2568E-02	27.72%	6.6638E-02	3.08%
1	3.7364E-02	5.5036E-02	47.30%	3.8319E-02	2.56%
2	1.9318E-02	3.4112E-02	76.58%	1.9800E-02	2.50%
3	8.6775E-03	1.9469E-02	124.36%	8.8700E-03	2.22%
4	3.5025E-03	9.7750E-03	179.09%	3.5713E-03	1.96%
5	1.1775E-03	4.3238E-03	267.20%	1.2038E-03	2.23%
6	3.3375E-04	1.8463E-03	453.20%	3.3500E-04	0.37%
7	6.8750E-05	6.3500E-04	823.64%	6.6250E-05	-3.64%
8	1.5000E-05	1.9750E-04	1216.67%	1.2500E-05	-16.67%
9	2.5000E-06	5.0000E-05	1900.00%	2.5000E-06	0.00%

Table 1: Comparison of Symbol Error Rates between Maximum Likelihood (M.L.) decoding, Linear Minimum Mean Squared Error (LMMSE) decoding, and OIM using CMOS ring oscillators. Double precision accuracy is used to represent the OIM coupling weights. The "%>M.L." columns indicate how much greater SERs are over Maximum Likelihood decoding.

SNR	9 bits		8 bits		7 bits	
	SER	%>M.L.	SER	%>M.L.	SER	%>M.L.
-1	1.0413E-01	3.97%	1.0406E-01	3.91%	1.0432E-01	4.16%
0	6.6559E-02	2.96%	6.6634E-02	3.07%	6.7040E-02	3.70%
1	3.8499E-02	3.04%	3.8456E-02	2.92%	3.8799E-02	3.84%
2	1.9768E-02	2.33%	1.9964E-02	3.34%	2.0014E-02	3.60%
3	8.8513E-03	2.00%	8.8738E-03	2.26%	9.0188E-03	3.93%
4	3.5888E-03	2.46%	3.5525E-03	1.43%	3.7063E-03	5.82%
5	1.2000E-03	1.91%	1.2063E-03	2.44%	1.2438E-03	5.63%
6	3.3250E-04	-0.37%	3.3625E-04	0.75%	3.6750E-04	10.11%
7	7.1250E-05	3.64%	7.1250E-05	3.64%	8.6250E-05	25.45%
8	1.1250E-05	-25.00%	1.3750E-05	-8.33%	1.7500E-05	16.67%
9	1.2500E-06	-50.00%	1.2500E-06	-50.00%	3.7500E-06	50.00%

Table 2: SER results from ring oscillator OIM using 9, 8 and 7 bits to represent coupling weights. The "%>M.L." columns indicate how much greater quantized-coupling-OIM SERs are over Maximum Likelihood decoding.

SNR	6 bits		5 bits		4 bits	
	SER	%>M.L.	SER	%>M.L.	SER	%>M.L.
-1	1.0531E-01	5.16%	1.1017E-01	10.01%	1.2946E-01	29.26%
0	6.8118E-02	5.37%	7.2540E-02	12.21%	9.1473E-02	41.50%
1	3.9763E-02	6.42%	4.3665E-02	16.86%	6.1594E-02	106.56%
2	2.0663E-02	6.96%	2.3853E-02	23.47%	3.9904E-02	106.56%
3	9.4813E-03	9.26%	1.1715E-02	35.00%	2.4290E-02	179.92%
4	3.8800E-03	10.78%	5.4975E-03	56.96%	1.5149E-02	332.51%
5	1.4188E-03	20.49%	2.1250E-03	80.47%	8.9675E-03	661.57%
6	4.6125E-04	38.20%	8.9750E-04	168.91%	5.6838E-03	1603.00%
7	9.6250E-05	40.00%	3.2750E-04	376.36%	4.0525E-03	5794.55%
8	1.8750E-05	25.00%	9.6250E-05	541.67%	2.5825E-03	17116.67%
9	5.0000E-06	100.00%	4.3750E-05	1650.00%	1.8575E-03	74200.00%

Table 3: SER results from ring oscillator OIM using 6, 5 and 4 bits to represent coupling weights. The "%>M.L." columns indicate how much greater quantized-coupling-OIM SERs are over Maximum Likelihood decoding.