Quadratization in Discrete Optimization and Quantum Mechanics

Richard Tanburn,[1] Nicholas Chancellor,[2] and Nike Dattani[1,3,4]

[1] *Oxford University, Department of Mathematics*
[2] *Durham University, Department of Physics*
[3] *McMaster University, Department of Chemistry and Chemical Biology*
[4] *Harvard-Smithsonian Center for Astrophysics*

A collaborative, evolving, open review paper on k-local to 2-local transformations (quadratizations) in classical computing, quantum annealing, and universal adiabatic quantum computing.

When optimizing discrete functions, it is often easier when the function is qudaratic than if it is of higher degree. But notice that the cubic and quadratic funcitons:

$$b_1b_2 + b_2b_3 + b_3b_4 - 4b_1b_2b_3 \quad \text{(cubic)}, \tag{1}$$
$$b_1b_2 + b_2b_3 + b_3b_4 + 4b_1 - 4b_1b_2 - 4b_1b_3 \quad \text{(quadratic)}, \tag{2}$$

where each b_i can either be 0 or 1, both never go below the value of -1, and all minima have the form $(b_1, b_2, b_3, b_4) = (1, 1, 1, b_4)$. Therefore if we are interested in the ground state of a discrete function of degree k, we may optimize either function and get exactly the same result. **Part I** gives more than 15 different ways to do this.

The binary variables b_i can be either of the eigenvalues of the matrix b below, which is related to the Pauli z matrix by $z = 2b - \mathbb{1}$. Other Pauli matrices are listed below:

$$b \equiv \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\; z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\; x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\; y \equiv \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix},\; \mathbb{1} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3}$$

Any Hermitian 2×2 matrix can be written as a linear combination of the Pauli matrices $z, x, y, \mathbb{1}$, so we can therefore describe the Hamiltonian of any number of spin-$^1/_2$ particles by a function of Pauli matrices acting on each particle, for instance:

$$z_1x_2 + x_2y_3 + z_3z_4 - 4y_1y_2z_3 \quad \text{(cubic)}, \tag{4}$$

where the coefficients tell us about the strengths of couplings between these particles. The Schrödinger equation tells us that the eigenvalues of the Hamiltonian are the allowed energy levels and their eigenvectors (wavefunctions) are the corresponding physical states. More generally these do not have to be spins but can be any type of qubits, and we can encode the solution to *any* problem in the ground state of a Hamiltonian, then solve the problem by finding the lowest energy state of the physical system (this is called adiabatic quantum computing).

Two-body physical interactions occur more naturally than many-body interactions so **Parts II-III** give more than 10 different ways to quadratize genreal Hamiltonians (some of these methods do not even require the s_i to be 2×2 matrices, meaning that we can have types of qudits that are not qubits).

The optimization problems of Eqs. (1)-(2) are specific cases of the type in Eq. (4) where only z matrices are present.

Part I

Diagonal Hamiltonians (pseudo-Boolean functions)

I. METHODS THAT INTRODUCE ZERO AUXILIARY VARIABLES

A. Deduction Reduction (Deduc-reduc)

Summary

We look for *deductions* (e.g. $b_1 b_2 = 0$) that must hold true at the global minimum. These can be found by *a priori* knowledge of the given problem, or by enumerating solutions of a small subset of the variables. We can then substitute high-order terms using the low-order terms of the deduction, and add on a penalty term to preserve the ground states [1].

Cost

- 0 auxiliary variables needed.
- For a particular value of m, we have $\binom{n}{m}$ different m-variable subsets of the n variable problem, and $\binom{n}{m}2^m$ evaluations of the objective function to find all possible m-variable deductions, whereas 2^n evaluations is enough to solve the entire problem. We therefore choose $m \lll n$.

Pros

- No auxiliary variables needed.

Cons

- When deductions cannot be determined naturally (as in the Ramsey number determination problem, see Example XII), deductions need to be found by 'brute force', which scales exponentially with respect to m. For highly connected systems (systems with a large number of non-zero α_{ij} coefficients), the value of m required to find even one deduction can be prohibitively large.

Example

Consider the Hamiltonian:

$$H_{4-\text{local}} = b_1 b_2 (4 + b_3 + b_3 b_4) + b_1 (b_3 - 3) + b_2 (1 - 2b_3 - b_4) + F(b_3, b_4, b_5, \ldots, b_N) \tag{5}$$

where F is any quadratic polynomial in b_i for $i \geq 3$. Since

$$H_{4-\text{local}}\left(1,1,b_3,b_4,\ldots\right) > H_{4-\text{local}}\left(0,0,b_3,b_4,\ldots\right), H_{4-\text{local}}\left(0,1,b_3,b_4,\ldots\right), H_{4-\text{local}}\left(1,0,b_3,b_4,\ldots\right), \tag{6}$$

it must be the case that $b_1 b_2 = 0$. Specifically, for the 4 assignments of (b_3, b_4), we see that $b_1 b_2 = 0$ at every minimum of $H_{4-\text{local}} - F$.

Using deduc-reduc we have:

$$H_{2-\text{local}} = 6 b_1 b_2 + b_1 (b_3 - 3) + b_2 (1 - 2b_3 - b_4) + F(b_3, b_4, b_5, \ldots, b_N), \tag{7}$$

which has the same global minima as $H_{4-\text{local}}$ but one fewer quartic and one fewer cubic term. The coefficient of $b_1 b_2$ was chosen as 6 because $6 \geq \max\left(4 + b_3 + b_3 b_4\right)$.

Bibliography

- Original paper, with more implementation details, and application to integer factorization: [1].

B. ELC Reduction (Ishikawa 2014)

Summary
An Excludable Local Configuration (ELC) is a partial assignment of variables that make it impossible to achieve the minimum. We can therefore add a term that corresponds to the energy of this ELC without changing the solution to the minimization problem. In practice we can eliminate all monomials with a variable in which a variable is set to 0, and reduce any variable set to 1. Given a general Hamiltonian we can try to find ELCs by enumerating solutions of a small subset of variables in the problem [2].

Cost
- 0 auxiliary variables needed.
- For a particular value of m, we have $\binom{n}{m}$ different m-variable subsets of the n variable problem, and $\binom{n}{m}2^m$ evaluations of the objective function to find all possible m-variable deductions, whereas 2^n evaluations is enough to solve the entire problem. We therefore choose $m \lll n$.
- Approximate methods exist which have been shown to be much faster and give good approximations to the global minimum [2].

Pros
- No auxiliary variables needed.

Cons
- No known way to find ELCs except by 'brute force', which scales exponentially with respect to m.
- ELCs do not always exist.

Example
Consider the Hamiltonian:

$$H_{3\text{-local}} = b_1 b_2 + b_2 b_3 + b_3 b_4 - 4 b_1 b_2 b_3. \tag{8}$$

If $b_1 b_2 b_3 = 0$, no assignment of our variables will we be able to reach a lower energy than if $b_1 b_2 b_3 = 1$. Hence this gives us *twelve* ELCs, and one example is $(b_1, b_2, b_3) = (1, 0, 0)$ which we can use to form the polynomial:

$$\begin{aligned} H_{2\text{-local}} &= H_{3\text{-local}} + 4b_1(1 - b_2)(1 - b_3) &(9)\\ &= b_1 b_2 + b_2 b_3 + b_3 b_4 + 4b_1 - 4b_1 b_2 - 4b_1 b_3. &(10)\end{aligned}$$

In both cases Eqs. (7) and (9), the only global minima occur when $b_1 b_2 b_3 = 1$.

Bibliography
- Original paper and application to computerized image denoising: [2].

C. Groebner Bases

Summary

Given a set of polynomials, a Groebner basis is another set of polynomials that have exactly the same zeros. The advantage of a Groebner basis is it has nicer algebraic properties than the original equations, in particular they tend to have smaller degree polynomials. The algorithms for calculating Groebner bases are generalizations of Euclid's algorithm for the polynomial greatest common divisor.

Work has been done in the field of 'Boolean Groebner bases', but while the variables are Boolean the coefficients of the functions are in $\mathbb{F}_2$ rather than $\mathbb{Q}$.

Cost

- 0 auxiliary variables needed.
- $\mathcal{O}\left(2^{2^n}\right)$ in general, $\mathcal{O}(d^{n^2})$ if the zeros of the equations form a set of discrete points, where d is the degree of the polynomial and n is the number of variables [3].

Pros

- No auxiliary variables needed.
- General method, which can be used for other rings, fields or types of variables.

Cons

- Best algorithms for finding Groebner bases scale double exponentially in n.
- Only works for Hamiltonians whose minimization corresponds to solving systems of discrete equations, as the method only preserves roots, not minima.

Example

Consider the following pair of equations:

$$b_1b_2b_3b_4 + b_1b_3 + b_2b_4 - b_3 = b_1 + b_1b_2 + b_3 - 2 = 0. \tag{11}$$

Feeding these to Mathematica's `GroebnerBasis` function, along with the binarizing $b_1(b_1 - 1) = \ldots = b_4(b_4 - 1) = 0$ constraints, gives a Groebner basis:

$$\{b_4b_3 - b_4, b_2 + b_3 - 1, b_1 - 1\}. \tag{12}$$

From this we can immediately read off the solutions $b_1 = 1$, $b_2 = 1 - b_3$ and reduce the problem to $b_3b_4 - b_4 = 0$. Solving this gives a final solution set of: $(b_1, b_2, b_3, b_4) = (1,0,1,0), (1,0,1,1), (1,1,0,0)$, which should be the same as the original 4-local problem.

Bibliography

- Reduction and embedding of factorizations of all bi-primes less than $200,000$: [4].

D. Application of ELM (Present Work)

Summary
We use the formula from [5] for representing any function of three binary variables:

$$
\begin{aligned}
f(b_1,b_2,b_3) = {} & (f(1,1,1)+f(1,0,0)-f(1,1,0)-f(1,0,1)-f(0,1,1)-f(0,0,0)+ && (13)\\
& f(0,0,1)+f(0,1,0))\,b_1b_2b_3+(f(0,1,1)+f(0,0,0)-f(0,0,1)-f(0,1,0))\,b_2b_3+ && (14)\\
& (f(1,0,1)+f(0,0,0)-f(0,0,1)-f(1,0,0))\,b_1b_3+ && (15)\\
& (f(1,1,0)+f(0,0,0)-f(1,0,0)-f(0,1,0))\,b_1b_2\ +(f(0,1,0)-f(0,0,0))\,b_2 && (16)\\
& +(f(1,0,0)-f(0,0,0))\,b_1+(f(0,0,1)-f(0,0,0))\,b_3+f(0,0,0). && (17)
\end{aligned}
$$

If the cubic term is zero, then the function becomes quadratic. We can use ELM (Energy Landscape Manipulation) to change the energy landscape *without* changing the ground state [6]. In this case, we apply ELM in order to make the cubic term zero.

Cost
- No auxiliary variables.
- May require many evaluations of the cubic function varying coefficients in order to find the right ELM coefficients.

Pros
- Can be generalized to arbitrary k-local functions, but the ELM constraints may become harder to achieve.
- Can quadratize an entire cubic function (or a cubic part of a more general function) with no auxiliary qubits.
- Can reproduce the full spectrum.

Cons
- May not always be possible.
- May require a local search to find appropriate deductions.

Example
In order to reduce the number of constraints required for the cubic term to be zero, we will assume that Deduc-Reduc told us that $(1-b_1)(1-b_2)+(1-b_2)(1-b_3)+(1-b_1)(1-b_3)=0$ when the overall function is minimized, which means the ground state only occurs when at least two variables are 1. This does not allow us to assign the linear terms, but assigns all quadratic terms to have $b_ib_j=1$. This also suggests that $b_1b_2b_3$ can also be reduced to a linear term, but we do not know whether it is b_1, b_2, or b_3. So we have the following constraint on the cubic term, after setting all $f(b_1,b_2,b_3)$ to zero if there is not at least two 1's:

$$f(1,1,1)-f(1,1,0)-f(1,0,1)-f(0,1,1)=0. \tag{18}$$

We now show an example of a function for which ELM has the power to enforce this constraint:

E. Split Reduction (Okada, Tanburn, Dattani, 2015)

Summary

It has been shown in [7] that, if multiple runs of a minimization algorithm is permitted, it is possible to reduce a lot of the problem by conditioning on the most connected variables. We call each of these operations a *split*.

Cost

Exponential in the number of splits, as the number of problems to solve doubles with every split.

Pros

- This method can be applied to any problem and can be very effective on problems with a few very connected variables.

Cons

- Exponential cost in the worst case.

Example

Consider the simple objective function

$$H = 1 + b_1b_2b_5 + b_1b_6b_7b_8 + b_3b_4b_8 - b_1b_3b_4. \tag{19}$$

In order to quadratize H, we first have to choose a variable to split over. In this case b_1 is the obvious choice since it is present in the most terms and contributes to the quartic term.

We then obtain two different problems:

$$H_0 = 1 + b_3b_4b_8 \tag{20}$$

$$H_1 = 1 + b_2b_5 + b_6b_7b_8 + b_3b_4b_8 - b_3b_4. \tag{21}$$

At this point, we could split H_0 again and solve it entirely, or use a variable we saved in the previous split to quadratize our only problem.

To solve H_1, we can split again on b_8, resulting in problems:

$$H_{1,0} = 1 + b_2b_5 + b_6b_7 \tag{22}$$

$$H_{1,1} = 1 + b_2b_5 + b_3b_4. \tag{23}$$

Now both of these problems are quadratic. Hence we have reduced our original, hard problem into 3 easy problems, requiring only 2 extra (much easier) runs of our minimization algorithm, and without needing any auxiliary variables.

Bibliography

- Original paper and application to Ramsey number determination: [7].

II. METHODS THAT INTRODUCE AUXILIARY VARIABLES TO QUADRATIZE A SINGLE NEGATIVE TERM (NEGATIVE TERM REDUCTIONS, NTR)

A. NTR-KZFD (Kolmogorov & Zabih, 2004; Freedman& Drineas, 2005)

Summary

For a negative term $-b_1 b_2 ... b_k$, introduce a single auxiliary variable b_a and make the substitution:

$$-b_1 b_2 \dots b_k \rightarrow (k-1)b_a - \sum_i b_i b_a. \tag{24}$$

Cost

- 1 auxiliary variable for each k-local term.

Pros

- All resulting quadratic terms are submodular (have negative coefficients).
- Can reduce arbitrary order terms with only 1 auxiliary.
- Reproduces the full spectrum.

Cons

- Only works for negative terms.

Example

$$H_{6\text{-local}} = -2b_1 b_2 b_3 b_4 b_5 b_6 + b_5 b_6, \tag{25}$$

has a unique minimum energy of -1 when all $b_i = 1$.

$$H_{2\text{-local}} = 2\left(5b_a - b_1 b_a - b_2 b_a - b_3 b_a - b_4 b_a - b_5 b_a - b_6 b_a\right) + b_5 b_6 \tag{26}$$

has the same unique minimum energy, and it occurs at the same place (all $b_i = 1$), with $b_a = 1$.

Alternate Forms

$$-b_1 b_2 \dots b_k = \min_{b_a}\left((k-1-\sum_i b_i)b_a\right) \tag{27}$$

$$\rightarrow \left((k-1-\sum_i b_i)b_a\right). \tag{28}$$

Alternate Names

- "Standard quadratization" of negative monomials [8].
- $s_k(b, b_a)$ [8]

Bibliography

- 2004: Kolmogorov and Zabih presented this for cubic terms [9].
- 2005: Generalized to arbitrary order by Freedman and Drineas [10].
- Discussion: [11], [12].

B. NTR-ABCG (Anthony, Boros, Crama, Gruber, 2014)

Summary

For a negative term $-b_1 b_2 ... b_k$, introduce a single auxiliary variable b_a and make the substitution:

$$-b_1 b_2 \dots b_k \to -\sum_i b_i - \sum_i b_i b_k - \sum_i b_i b_a + (k-1) b_k b_a. \tag{29}$$

Cost

- 1 auxiliary variable for each k-local term.
- 1 non-submodular term for each k-local termi (and it is quadratic).

Pros

- Can reduce arbitrary order terms with only 1 auxiliary.
- Reproduces the full spectrum.

Cons

- Only works for negative terms.
- Turns a symmetric term into a non-symmetric term (but only b_k is asymmetric).

Example

$$H_{6\text{-local}} = -2b_1 b_2 b_3 b_4 b_5 b_6 + b_5 b_6, \tag{30}$$

has a unique minimum energy of -1 when all $b_i = 1$.

$H_{2\text{-local}}$ has the same unique minimum energy, and it occurs at the same place (all $b_i = 1$), with $b_a = 1$.

Alternate Forms

$$-b_1 b_2 \dots b_k \to (k-1) b_k b_a - \sum_i b_i (b_a + b_k - 1) \tag{31}$$

$$= (k-2) b_a - \sum_i^{k-1} b_i \left(b_a + b_k - 1\right) \tag{32}$$

Alternate Names

- "Extended standard quadratization" of negative monomials [8].
- $s_k(b, b_a)^+$ [8]

Bibliography

- 2014, First presentation: [12, 13].
- Further discussion: [8].

C. NTR-ABCG-2 (Anthony, Boros, Crama, Gruber, 2016)

Summary

For a negative term $-b_1 b_2 ... b_k$, introduce a single auxiliary variable b_a and make the substitution:

$$-b_1 b_2 \ldots b_k \rightarrow (2k-1)\, b_a - 2\sum_i b_i b_a \tag{33}$$

Cost

- 1 auxiliary variable for each k-local term.
- 1 non-submodular term for each k-local term (and it is linear).

Pros

- Can reduce arbitrary order terms with only 1 auxiliary.
- Reproduces the full spectrum.
- The non-submodular term is linear as opposed to NTR-ABCG-1 whose non-submodular term is quadratic.
- Symmetric with respect to all non-auxiliary variables.

Cons

- Only works for negative terms.
- Turns a symmetric term into a non-symmetric term (but only b_k is asymmetric).
- Coefficients of quadratic terms are twice the size of their size in NTR-KZFD or NTR-ABCG-1, and roughly twice the size for the linear term.

Example

$$H_{6\text{-local}} = -2b_1 b_2 b_3 b_4 b_5 b_6 + b_5 b_6, \tag{34}$$

has a unique minimum energy of -1 when all $b_i = 1$.

$H_{2\text{-local}}$ has the same unique minimum energy, and it occurs at the same place (all $b_i = 1$), with $b_a = 1$.

Bibliography

- Discussion: [12].

D. Asymmetric Cubic Reduction

Summary

$$-b_1b_2b_3 \rightarrow b_a\left(-b_1+b_2+b_3\right)-b_1b_2-b_1b_3+b_1 \tag{35}$$

Cost
1 auxiliary variable per negative cubic term.

Pros
- Asymmetric which allows more flexibility in cancelling with other quadratics.

Cons
- Only works for negative cubic monomials.

Example

$$-b_1b_2b_3+b_1b_3-b_2=\min_{b_a}\left(b_a-b_1b_a-b_3b_a+b_2b_a+2b_1b_3\right)-b_2 \tag{36}$$

Alternate Forms

$$-b_1b_2b_3=\min_{b_a} b_a-b_1+b_2+b_3-b_1b_2-b_1b_3+b_1 \tag{37}$$

Bibliography
- Discussion: [11], [14].

III. METHODS THAT INTRODUCE AUXILIARY VARIABLES TO QUADRATIZE A SINGLE POSITIVE TERM (POSITIVE TERM REDUCTIONS, PTR)

A. Positive Term Reduction

Summary

By considering the negated literals $\bar{b}_i = 1 - b_i$, we recursively apply NTR-KZFD to $b_1 b_2 \ldots b_k = -\bar{b}_1 b_2 \ldots b_k + b_2 b_3 \ldots b_k$. The final identity is:

$$b_1 b_2 \ldots b_k \rightarrow \left(\sum_{i=1}^{k-2} b_{a_i} \left(k - i - 1 + b_i - \sum_{j=i+1}^{k} b_j \right) \right) + b_{k-1} b_k \tag{38}$$

Cost

- $k - 2$ auxiliary variables for each k-local term.

Pros

- Works for positive monomials.

Cons

- $k - 1$ non-submodular quadratic terms.

Example

$$b_1 b_2 b_3 b_4 \rightarrow b_{a_1} \left(2 + b_1 - b_2 - b_3 - b_4 \right) + b_{a_2} \left(1 + b_2 - b_3 - b_4 \right) + b_3 b_4 \tag{39}$$

Bibliography

- Summary: [15].

B. PTR-Ishikawa (Ishikawa, 2011)

Summary

This method re-writes a positive monomial using symmetric polynomials, so all possible quadratic terms are produced and they are all non-submodular:

$$b_1 ... b_k \rightarrow \left(\sum_{i=1}^{n_k} b_{a_i} \left(c_{i,d} \left(- \sum_{j=1}^{k} b_j + 2i \right) - 1 \right) + \sum_{i<j} b_i b_j \right) \tag{40}$$

where $n_k = \left\lfloor \frac{k-1}{2} \right\rfloor$ and $c_{i,k} = \begin{cases} 1, & i = n_d \text{ and } k \text{ is odd,} \\ 2, & \text{else.} \end{cases}$

Cost

- $\left\lfloor \frac{k-1}{2} \right\rfloor$ auxiliary variables for each k-order term
- $\mathcal{O}(kt)$ for a k-local Hamiltonian with t terms.

Pros

- Works for positive monomials.
- About half as many auxiliary variables for each k-order term as the previous method.
- Reproduces the full spectrum.

Cons

- $\mathcal{O}(k^2)$ quadratic terms are created, which may make chimerization more costly.
- $\frac{k(k-1)}{2}$ non-submodular terms.
- Worse than the previous method for quartics, with respect to submodularity.

Example

$$b_1 b_2 b_3 b_4 \rightarrow (3 - 2b_1 - 2b_2 - 2b_3 - 2b_4) b_a + b_1 b_2 + b_1 b_3 + b_1 b_4 + b_2 b_3 + b_2 b_4 + b_3 b_4 \tag{41}$$

Alternate Forms

For even k, and equivalent expression is given in [16]:

$$b_1 b_2 \dots b_k \rightarrow \sum_i b_i + \sum_{ij} b_i b_j + \sum_{2i} b_{a_{2i}} \left(4i - 2 - \sum_j b_j \right) \tag{42}$$

$$\rightarrow \sum_i b_i + 2 \sum_{2i} b_{a_{2i}} (2i - 1) + \sum_{ij} b_i b_j - \sum_{2i,j} b_j b_{a_{2i}} \tag{43}$$

Alternate Names

- "Ishikawa's Symmetric Reduction" [17].
- "Ishikawa Reduction"
- "Ishikawa"

Bibliography

- Original paper and application to image denoising: [11].
- Equivalent way of writing it for even k, shown in [16].

C. PTR-BCR-1 (Boros, Crama, and Rodríguez-Heck, 2018)

Summary

This is very similar to the alternative form of Ishikawa Reduction, but works for odd values of k, and is different from Ishikawa Reduction:

$$b_1b_2\ldots b_k \to \sum_i b_i + \sum_{2i-1}(4i-3)\,b_{a_{2i-1}} + \sum_{ij} b_ib_j - \sum_{2i-1,j} b_jb_{a_{2i-1}} \tag{44}$$

Cost

- Same number of auxiliaries as Ishikawa Reduction.

Pros

- Same as for Ishikawa Reduction.

Cons

- Same as for Ishikawa Reduction.
- Only works for odd k, but for even k we have an analogous method which is equivalent to Ishikawa Reduction.

Alternate Forms

$$b_1b_2\ldots b_k \to \sum_i b_i + \sum_{ij} b_ib_j + \sum_{2i-1} b_{a_{2i-1}}\left(4i-3-\sum_j b_j\right) \tag{45}$$

Bibliography

- Original paper: [16].

D. PTR-BCR-2 (Boros, Crama, and Rodríguez-Heck, 2018)

Summary
Let $m = \lceil \frac{k}{4} \rceil$,

$$b_1 b_2 \cdots b_k \to \alpha^b \sum_i b_i + \alpha^{b_{a,1}} \sum_i b_{a_i} + \alpha^{b_{a,2}} b_{a_m} + \alpha^{bb} \sum_{ij} b_i b_j + \alpha^{bb_{a,1}} \sum_i \sum_j^{m-1} b_i b_{a_j} + \alpha^{bb_{a,2}} \sum_i b_i b_{a_m} + \alpha^{b_{a,1} b_{a,1}} \sum_{ij}^{m-1} b_{a_i} b_{a_j} + \alpha^{b_{a,1} b_{a,2}} \sum_i^{m-1} b_{a_i} b_{a_m}, \tag{46}$$

where:

$$\begin{pmatrix} \alpha^b & \alpha^{bb_{a,1}} \\ \alpha^{b_{a,1}} & \alpha^{bb_{a,2}} \\ \alpha^{b_{a,2}} & \alpha^{b_{a,1} b_{a,1}} \\ \alpha^{bb} & \alpha^{b_{a,1} b_{a,2}} \end{pmatrix} = \begin{pmatrix} -1/2 & -1 \\ 1 & -2 \\ \frac{1}{2}(n - m + n^2 - 2mn + m^2) & -(n-m) \\ 1/2 & 4(n-m) \end{pmatrix}. \tag{47}$$

Cost
$\lceil \frac{k}{4} \rceil$ auxiliary qubits per positive monomial.

Pros
- Smallest number of auxiliary coefficients that scales linearly with k.
- Smaller coefficients than the logarithmic reduction.

Cons
- Introduces many non-submodular terms.

Example
We quadratize a quartic term with only 1 auxiliary (half as many as in PTR-Ishikawa):

$$b_1 b_2 b_3 b_4 \to \frac{1}{2} \left(b_1 + b_2 + b_3 + b_4 - 2b_{a_1} \right) \left(b_1 + b_2 + b_3 + b_4 - 2b_{a_1} - 1 \right) \tag{48}$$

Bibliography
- Original paper: [16].

E. PTR-BCR-3 (Boros, Crama, and Rodríguez-Heck, 2018)

Summary
Pick m such that $k < 2^{m+1}$,

$$b_1 b_2 \dots b_k \to \alpha + \alpha^b \sum_i b_i + \alpha^{b_{a_i}} \sum_i 2^{i-1} b_{a_i} + \alpha^{bb} \sum_{ij} b_i b_j + \alpha^{bb_a} \sum_{ij} b_i b_{a_j} + \alpha^{b_{a_i} b_{a_j}} b_{a_i} b_{a_j}, \tag{49}$$

where,

$$\begin{pmatrix} \alpha & \alpha^{bb} \\ \alpha^b & \alpha^{bb_a} \\ \alpha^{b_a} & \alpha^{b_a b_a} \end{pmatrix} = \begin{pmatrix} (2^m - k)^2 & 1 \\ 2(2^m - k) & 2^{j-1} \\ -2(2^m - k) & 2^{i+-2} \end{pmatrix}. \tag{50}$$

Cost
$\lceil \log k \rceil$ auxiliary qubits per positive monomial.

Pros
- Logarithmic number of auxiliary variables.

Cons
- Introduces all terms non-submodular except for the term linear in auxiliaries.

Example

$$b_1 b_2 b_3 b_4 \to \frac{1}{2}\left(4 + b_1 + b_2 + b_3 + b_4 - b_{a_1} - 2b_{a_2}\right)\left(3 + b_1 + b_2 + b_3 + b_4 - b_{a_1} - 2b_{a_2}\right) \tag{51}$$

Alternate Forms

$$b_1 b_2 \dots b_k \to \left(2^m - k + \sum_i b_i - \sum_i 2^{i-1} b_{a_i}\right)^2 \tag{52}$$

Bibliography
- Original paper (Theorem 4, special case of Theorem 1): [16].

F. PTR-BCR-4 (Boros, Crama, and Rodríguez-Heck, 2018)

Summary
Pick m such that $k \leq 2^{m+1}$,

$$b_1 \dots b_k \to \frac{1}{2}\left(2^{m+1} - k + \sum_i b_i - \sum_i^m 2^i b_{a_i}\right)\left(2^{m+1} - k + \sum_i b_i - \sum_i^m 2^i b_{a_i} - 1\right). \tag{53}$$

Cost
$\lceil \log k \rceil - 1$ auxiliary qubits per positive monomial.

Pros
- Logarithmic number of auxiliary variables.

Cons
- Introduces many non-submodular terms.

Example

$$b_1 b_2 b_3 b_4 \to \frac{1}{2}\left(b_1 + b_2 + b_3 + b_4 - 2b_a\right)\left(b_1 + b_2 + b_3 + b_4 - 2b_a - 1\right) \tag{54}$$

Alternate Forms
Let $X = \sum b_i$ and $N = 2^{m+1} - k$,

$$b_1 \dots b_k = \min_{b'_1, \dots b'_n} \frac{1}{2}\left(N + X - \sum_{i=1}^n 2^i b'_i\right)\left(N + X - \sum_{i=1}^n 2^i b'_i - 1\right). \tag{55}$$

Bibliography
- Original paper (Theorem 5): [16].

G. PTR-KZ (Kolmogorov & Zabih, 2004)

Summary

This method can be used to re-write positive or negative cubic terms in terms of 6 quadratic terms. The identity is given by:

$$b_1 b_2 b_3 \rightarrow 1 - (b_a + b_1 + b_2 + b_3) + b_a (b_1 + b_2 + b_3) + b_1 b_2 + b_1 b_3 + b_2 b_3 \tag{56}$$

Cost

- 1 auxiliary variable per positive or negative cubic term.

Pros

- Works on positive or negative monomials.
- Reproduces the full spectrum.

Cons

- Introduces all 6 possible non-submodular quadratic terms.

Alternate Names

- "Reduction by Minimum Selection" [11, 17].

Bibliography

- Original paper: [9].

H. Reduction by Minimum Selection (in terms of z)

Summary

The formula is almost the same as in the version PTR-KZ in terms of b, but with a factor of 2, a change of sign for the linear terms, and a slight change in the constant term:

$$\pm z_1 z_2 z_3 \rightarrow 3 \pm (z_1 + z_2 + z_3 + z_a) + 2z_a (z_1 + z_2 + z_3) + z_1 z_2 + z_1 z_3 + z_2 z_3. \tag{57}$$

Cost

- 1 auxiliary variable per positive or negative cubic term.

Pros

- Works on positive or negative monomials.
- Reproduces the full spectrum.

Cons

- Introduces all 6 possible non-submodular quadratic terms.

Bibliography

- 31 March 2016 published by Chancellor, Zohren, and Warburton without the constant term: [18].
- 8 April 2016 published independently by Leib, Zoller, and Lechner without the constant term [19, 20].

I. Asymmetric Reduction

Summary

Similar to other methods of reducing one term, this method can reduce a positive cubic monomial into quadratic terms using only one auxiliary variable, while introducing fewer non-submodular terms than the symmetric version.

The identity is given by:

$$\begin{aligned} b_1b_2b_3 &\rightarrow b_a - b_2b_a - b_3b_a + b_1b_a + b_2b_3 && (58) \\ &\rightarrow b_a - b_1b_a - b_3b_a + b_2b_a + b_1b_3 && (59) \\ &\rightarrow b_a - b_1b_a - b_2b_a + b_3b_a + b_1b_2. && (60) \\ & && (61) \end{aligned}$$

Cost

1 auxiliary variable per positive cubic term.

Pros

- Works on positive monomials.
- Fewer non-submodular terms than Ishikawa Reduction.

Cons

- Only been shown to work for cubics.

Example

$$b_1b_2b_3 + b_1b_3 - b_2 \rightarrow (b_a - b_1b_a - b_3b_a + b_2b_a + 2b_1b_3) - b_2 \tag{62}$$

Bibliography

- Original paper and application to computer vision: [17].

J. Symmetry Based Mappings

Summary

Auxilliary qubits can be made to "count" the number of logical qubits in the 1 configuration. By applying single qubit terms to the auxilliary qubits, the spectrum of *any* permutation symmetric Hamiltonian can be reproduced.

Cost

- For a k local coupler requires k auxilliary qubits.

Pros

- Natural flux qubit implementation [21].
- Single gadget can reproduce any permutation symmetric spectrum.
- High degree of symmetry means this method is natural for some kinds of quantum simulations [22].

Cons

- Requires coupling between all logical bits and from all logical bits to all auxilliary bits.
- Requires single body terms of increasing strength as k is increased.

Example

A 4 qubit gadget guarantees that the number of auxillary bits in the -1 state is equal to the number of logical bits in the 1 state

$$H_{4-\text{count}} = 4\sum_{i=2}^{4}\sum_{j=1}^{i-1} b_i b_j + 4\sum_{i=1}^{4}\sum_{j=1}^{4} b_i b_{a_j} - 15\sum_{i=1}^{4} b_i - 8\sum_{i=1}^{4} b_{a_i} + (5\,b_{a_1} + b_{a_2} - 3\,b_{a_3} - 7\,b_{a_4}) + 26 \quad (63)$$

This gadget can be expressed more naturally in terms of z:

$$H_{4-\text{count}} = \sum_{i=2}^{4}\sum_{j=1}^{i-1} z_i z_j - \frac{1}{2}\sum_{i=1}^{4} z_i + \sum_{i=1}^{4}\sum_{j=1}^{4} z_i z_{a_j} + \frac{1}{2}\left(5 z_{a_1} + z_{a_2} - 3\,z_{a_3} - 7 z_{a_4}\right). \quad (64)$$

To replicate the spectrum of $b_1 b_2 b_3 b_4$, we add

$$H_{2-\text{local}} = -b_{a_4} + \lambda H_{4-\text{count}}. \quad (65)$$

where λ is a large number.

For the spectrum of

$$\begin{aligned} z_1 z_2 z_3 z_4 = {} & 16\,b_1 b_2 b_3 b_4 - 8\,(b_1 b_2 b_3 + b_1 b_2 b_4 + b_1 b_3 b_4 + b_2 b_3 b_4) + \\ & 4\,(b_1 b_2 + b_1 b_3 + b_1 b_4 + b_2 b_3 + b_2 b_4 + b_3 b_4) - 2\,(b_1 + b_2 + b_3 + b_4) + 1, \end{aligned} \quad (66)$$

we implement,

$$H_{2-\text{local}} = 2\,b_{a_1} - 2\,b_{a_2} + 2\,b_{a_3} - 2\,b_{a_4} + \lambda H_{4-\text{count}}, \quad (67)$$

Bibliography

- Paper on flux qubit implementation: [21]
- Paper on max-k-sat mapping: [23]
- Talk including use in quantum simulation: [22]

K. Bit flipping (Ishikawa, 2011)

Summary

For any variable b, we can consider the negation $\bar{b} = 1 - b$. The process of exchanging b for $\bar{b}$ is called *flipping*. Using bit-flipping, an arbitrary function in n variables can be represented using at most $2^{(n-2)}(n-3) + 1$ variables, though this is a gross overestimate.

Can be used in many different ways:

1. Flipping positive terms and using II A, recursively;
2. For $\alpha < 0$, we can reduce $\alpha\bar{b}_1\bar{b}_2...\bar{b}_k$ very efficiently to submodular form using II A. A generalized version exists for arbitrary combinations of flips in the monomial which makes reduction entirely submodular [11];
3. When we have quadratized we can minimize the number of non-submodular terms by flipping.
4. We can make use of both b_i and $\bar{b}_i$ in the same Hamiltonian by adding on a sufficiently large penalty term: $\lambda(b_i + \bar{b}_i - 1)^2 = \lambda(1 + 2b_i\bar{b}_i - b_i - \bar{b}_i)$. This is similar to the ideas in reduction by substitution or deduc-reduc. In this way, given a quadratic in n variables we can make sure it only has at most n nonsubmodular terms if we are willing to use the extra n negation variables as well (so we have $2n$ variables in total).

Cost

- None, as replacing b_i with it's negation $\bar{b}_i$ costs nothing except a trivial symbolic expansion.

Pros

- Cheap and effective way of improving submodularity.
- Can be used to combine terms in clever ways, making other methods more efficient.

Cons

- Unless the form of the Hamiltonian is known, spotting these 'factorizations' using negations is difficult.
- We need an auxiliary variable for each b_i for which we also want to use $\bar{b}_i$ in the same Hamiltonian.

Example

By bit-flipping b_2 and b_4, i.e. substituting $b_2 = 1 - \bar{b}_2$ and $b_4 = 1 - \bar{b}_4$, we see that:

$$H = 3b_1b_2 + b_2b_3 + 2b_1b_4 - 4b_2b_4 \tag{68}$$

$$= -3b_1\bar{b}_2 - \bar{b}_2b_3 - 2b_1\bar{b}_4 - \bar{b}_2\bar{b}_4 + 5b_1 + b_3 + 4\bar{b}_2 + 4\bar{b}_4 - 4. \tag{69}$$

The first expression is highly non-submodular while the second is entirely submodular.

Bibliography

- Original paper: [11].

IV. METHODS THAT QUADRATIZE MULTIPLE TERMS WITH THE SAME AUXILIARIES (CASE 1: SYMMETRIC FUNCTION REDUCTIONS, SFR)

A symmetric function is one where if we switch any of the variable names (for example $b_1 \to b_5 \to b_8 \to b_1$), the function's output is unaffected.

A. SFR-ABCG-1 (Anthony, Boros, Crama, Gruber, 2014)

Summary
Any n-variable symmetric function $f(b_1, b_2, \ldots b_n) \equiv f(b)$ can be quadratized with $n-2$ auxiliaries:

$$f(b) \to -\alpha_0 - \alpha_0 \sum_i b_i + a_2 \sum_{ij} b_i b_j + 2 \sum_i (\alpha_i - c)\, b_{a_i} \left(2i - \frac{1}{2} - \sum_j b_j \right) \tag{70}$$

$$c = \begin{cases} \min(\alpha_{2j}) & , i \in \text{even} \\ \min(\alpha_{2j-1}) & , i \in \text{odd} \end{cases} \tag{71}$$

$$a_2 = \text{Determined from Page 12 of [24]} \tag{72}$$

Cost
- $n-2$ auxiliaries for any n-variable symmetric function.
- n^2 non-submodular quadratic terms (of the non-auxiliary variables).
- $n-2$ non-submodular linear terms (of the auxiliary variables).

Pros
- Quadratization is symmetric in all non-auxiliary variables (this is not always true, for example some of the methods in the NTR section).
- Reproduces the full spectrum.
- When there's a large number of terms, there's fewer auxiliary variables than quadratizing each positive monomial separately.

Cons
- Only works on a specific class of functions, although the quadratizations of arbitrary functions can be related to the quadratizations of symmetric functions on a larger number of variables.
- All quadratic terms of the non-auxiliary variables are non-sub-modular.
- All linear terms of the auxliaries are non-submodular.
- Not meant so much to be practical, but rather an easy proof of an upper bound on the number of needed auxiliaries.

Bibliography
- 2014, original paper (Theorem 4.1, with α_i from Corollary 2.3): [13].

B. SFR-BCR-1 (Boros, Crama, Rodríguez-Hector, 2018)

Summary

Any n-variable symmetric function $f(b_1, b_2, \dots b_n) \equiv f(b)$ that is non-zero only when $\sum b_i = c$ where $n/2 \le c \le n$, can be quadratized with $m = \lceil \log_2 c \rceil + 1$ auxiliary variables:

$$f(b) \rightarrow \alpha + \alpha^b \sum_i b_i + \alpha^{b_{a,1}} \sum_i^{m-1} b_{a_i} + \alpha^{b_{a,2}} b_{a_m} + \alpha^{bb} \sum_{ij} b_i b_j + \alpha^{bb_{a,1}} \sum_i \sum_j^{m-1} b_i b_{a_j} + \alpha^{bb_{a,2}} \sum_i b_i b_{a_m} + \alpha^{b_a b_a} \sum_{i,j}^{m-1} b_{a_i} b_{a_j}, \tag{73}$$

where:

$$\begin{pmatrix} \alpha & \alpha^{bb} \\ \alpha^b & \alpha^{bb_{a,1}} \\ \alpha^{b_{a,1}} & \alpha^{bb_{a,2}} \\ \alpha^{b_{a,2}} & \alpha^{b_a b_a} \end{pmatrix} = \begin{pmatrix} (c+1)^2 & 1 \\ -2(c+1) & -2^i \\ 2(c+1) & -2\left(2m - 2^{m-1} + 1\right) \\ 2(c+1)\left(2c - 2^{m-1} + 1\right) & 2^{i+j-2} \end{pmatrix}. \tag{74}$$

Cost

- $m = \lceil \log_2 c \rceil + 1$ auxiliary variables.
- $n^2 + m^2$ non-submodular quadratic terms (all possible quadratic terms involving only non-auxiliary or only auxiliary variables).
- m non-submodular linear terms (all possible linear terms involving auxiliaries).

Pros

- Small number of auxiliary terms

Cons

- Only works for a special class of functions
- Introduces many linear and even more quadratic non-submodular terms.

Example

By bit-flipping b_2 and b_4, i.e. substituting $b_2 = 1 - \bar{b}_2$ and $b_4 = 1 - \bar{b}_4$, we see that:

$$H = 3b_1b_2 + b_2b_3 + 2b_1b_4 - 4b_2b_4 \tag{75}$$

$$= -3b_1\bar{b}_2 - \bar{b}_2b_3 - 2b_1\bar{b}_4 - \bar{b}_2\bar{b}_4 + 5b_1 + b_3 + 4\bar{b}_2 + 4\bar{b}_4 - 4. \tag{76}$$

The first expression is highly non-submodular while the second is entirely submodular.

Alternate Forms

$$f(b_1, b_2, \dots, b_n) \rightarrow \left(-(c+1) + \sum_i b_i - \left(2c - 2^{m-1} + 1\right) b_{a_m} - \sum_i^{m-1} 2^{i-1} b_{a_i} \right)^2 \tag{77}$$

Bibliography

- 2018, original paper (Theorem 1): [16].

C. SFR-BCR-2 (Boros, Crama, Rodríguez-Hector, 2018)

Summary

Any n-variable symmetric function $f(b_1, b_2, \dots b_n) \equiv f(b)$ that is non-zero only when $\sum b_i = c$ where $0 \le c \le {}^{n}\!/_{2}$, can be quadratized with $m = \lceil \log_2(n-c) \rceil + 1$ auxiliary variables:

$$f(b) \to \left((m-1) - \sum_i b_i - \left(2(n-m) - 2^{\lceil \log(n-m) \rceil} + 1 \right) b_{a_{\lceil \log(n-m) \rceil + 1}} - \sum_i^{\lceil \log(n-m) \rceil} 2^{i-1} b_{a_i} \right)^2 \tag{78}$$

where:

$$\begin{pmatrix} \alpha & \alpha^{bb} \\ \alpha^{b} & \alpha^{bb_{a,1}} \\ \alpha^{b_{a,1}} & \alpha^{bb_{a,2}} \\ \alpha^{b_{a,2}} & \alpha^{b_a b_a} \end{pmatrix} = \begin{pmatrix} (c+1)^2 & 1 \\ -2(c+1) & -2^i \\ 2(c+1) & -2\left(2m - 2^{m-1} + 1\right) \\ 2(c+1)\left(2c - 2^{m-1} + 1\right) & 2^{i+j-2} \end{pmatrix}. \tag{79}$$

Cost

- $m = \lceil \log_2(n-c) \rceil + 1$ auxiliary variables.
- $n^2 + m^2$ non-submodular quadratic terms (all possible quadratic terms involving only non-auxiliary or only auxiliary variables).
- m non-submodular linear terms (all possible linear terms involving auxiliaries).

Pros

- Small number of auxiliary terms

Cons

- Only works for a special class of functions
- Introduces many linear and even more quadratic non-submodular terms.

Example

By bit-flipping b_2 and b_4, i.e. substituting $b_2 = 1 - \bar{b}_2$ and $b_4 = 1 - \bar{b}_4$, we see that:

$$H = 3b_1b_2 + b_2b_3 + 2b_1b_4 - 4b_2b_4 \tag{80}$$
$$= -3b_1\bar{b}_2 - \bar{b}_2b_3 - 2b_1\bar{b}_4 - \bar{b}_2\bar{b}_4 + 5b_1 + b_3 + 4\bar{b}_2 + 4\bar{b}_4 - 4. \tag{81}$$

The first expression is highly non-submodular while the second is entirely submodular.

Bibliography

- 2018, original paper (Theorem 1): [16].

D. SFR-BCR-3 (Boros, Crama, Rodríguez-Hector, 2018)

Summary

Any n-variable symmetric function $f(b_1, b_2, \dots b_n) \equiv f(b)$ that is non-zero only when $\sum b_i = c$ where $n/2 \leq c \leq n$, can be quadratized with $m = \lceil \log_2 c \rceil + 1$ auxiliary variables $f \rightarrow$:

$$\frac{1}{2}\left(\sum_i b_i - (m - 2^c)\, b_{a_{c+1}} - (m+1)((1 - b_{a_{c+1}}) - \sum_i^c 2^{i-1} b_{a_i}\right)\left(\sum_i b_i - (m - 2^c)\, b_{a_{c+1}} - (m+1)((1 - b_{a_{c+1}}) - \sum_i^c 2^{i-1} b_{a_i} - 1\right) \quad (82)$$

where:

$$\begin{pmatrix} \alpha & \alpha^{bb} \\ \alpha^{b} & \alpha^{bb_{a,1}} \\ \alpha^{b_{a,1}} & \alpha^{bb_{a,2}} \\ \alpha^{b_{a,2}} & \alpha^{b_a b_a} \end{pmatrix} = \begin{pmatrix} (c+1)^2 & 1 \\ -2(c+1) & -2^i \\ 2(c+1) & -2\left(2m - 2^{m-1} + 1\right) \\ 2(c+1)\left(2c - 2^{m-1} + 1\right) & 2^{i+j-2} \end{pmatrix}. \quad (83)$$

Cost

- $m = \lceil \log_2 c \rceil + 1$ auxiliary variables.
- $n^2 + m^2$ non-submodular quadratic terms (all possible quadratic terms involving only non-auxiliary or only auxiliary variables).
- m non-submodular linear terms (all possible linear terms involving auxiliaries).

Pros

- Small number of auxiliary terms

Cons

- Only works for a special class of functions
- Introduces many linear and even more quadratic non-submodular terms.

Example

By bit-flipping b_2 and b_4, i.e. substituting $b_2 = 1 - \bar{b}_2$ and $b_4 = 1 - \bar{b}_4$, we see that:

$$H = 3b_1b_2 + b_2b_3 + 2b_1b_4 - 4b_2b_4 \quad (84)$$
$$= -3b_1\bar{b}_2 - \bar{b}_2b_3 - 2b_1\bar{b}_4 - \bar{b}_2\bar{b}_4 + 5b_1 + b_3 + 4\bar{b}_2 + 4\bar{b}_4 - 4. \quad (85)$$

The first expression is highly non-submodular while the second is entirely submodular.

Alternate Forms

$$f(b_1, b_2, \dots, b_n) \rightarrow \begin{pmatrix} \frac{1}{2}\left(\sum_i b_i - (m - 2^c)\, b_{a_{c+1}} - (m+1)((1 - b_{a_{c+1}}) - \sum_i^c 2^{i-1} b_{a_i}\right) \\ 2 \end{pmatrix} \quad (86)$$

Bibliography

- 2018, original paper (Theorem 2): [16].

E. SFR-BCR-4 (Boros, Crama, Rodríguez-Hector, 2018)

Summary

Any n-variable symmetric function $f(b_1, b_2, \dots b_n) \equiv f(b)$ that is non-zero only when $\sum b_i = c$ where $0 \le c \le {}^n\!/_2$, can be quadratized with $m = \lceil \log_2(n-c) \rceil + 1$ auxiliary variables $f \to$:

$$\frac{1}{2}\left((c-1) - \sum_i b_i - \left(2(n-c) - 2^{m-1} + 1\right) b_{a_m} - \sum_i^{m-1} 2^{i-1} b_{a_i}\right)\left((c-1) - \sum_i b_i - \left(2(n-c) - 2^{m-1} + 1\right) b_{a_m} - \sum_i^{m-1} 2^{i-1} b_{a_i} - 1\right) \quad (87)$$

where:

$$\begin{pmatrix} \alpha & \alpha^{bb} \\ \alpha^{b} & \alpha^{bb_{a,1}} \\ \alpha^{b_{a,1}} & \alpha^{bb_{a,2}} \\ \alpha^{b_{a,2}} & \alpha^{b_a b_a} \end{pmatrix} = \begin{pmatrix} (c+1)^2 & 1 \\ -2(c+1) & -2^i \\ 2(c+1) & -2\left(2m - 2^{m-1} + 1\right) \\ 2(c+1)\left(2c - 2^{m-1} + 1\right) & 2^{i+j-2} \end{pmatrix}. \quad (88)$$

Cost

- $m = \lceil \log_2(n-c) \rceil + 1$ auxiliary variables.
- $n^2 + m^2$ non-submodular quadratic terms (all possible quadratic terms involving only non-auxiliary or only auxiliary variables).
- m non-submodular linear terms (all possible linear terms involving auxiliaries).

Pros

- Small number of auxiliary terms

Cons

- Only works for a special class of functions
- Introduces many linear and even more quadratic non-submodular terms.

Example

By bit-flipping b_2 and b_4, i.e. substituting $b_2 = 1 - \bar{b}_2$ and $b_4 = 1 - \bar{b}_4$, we see that:

$$H = 3b_1b_2 + b_2b_3 + 2b_1b_4 - 4b_2b_4 \quad (89)$$

$$= -3b_1\bar{b}_2 - \bar{b}_2b_3 - 2b_1\bar{b}_4 - \bar{b}_2\bar{b}_4 + 5b_1 + b_3 + 4\bar{b}_2 + 4\bar{b}_4 - 4. \quad (90)$$

The first expression is highly non-submodular while the second is entirely submodular.

Alternate Forms

$$f(b_1, b_2, \dots, b_n) \to \binom{\frac{1}{2}\left((c-1) - \sum_i b_i - \left(2(n-c) - 2^{m-1} + 1\right) b_{a_m} - \sum_i^{m-1} 2^{i-1} b_{a_i}\right)}{2} \quad (91)$$

Bibliography

- 2018, original paper (Theorem 2): [16].

F. SFR-BCR-5 (Boros, Crama, Rodríguez-Hector, 2018)

Summary

For an n-variable symmetric function that is a function of the sum of all variables $f(b_1, b_2, \ldots b_n) = f(\sum b_i)$, for some large value of $\lambda > \max(f)$, and $\lceil\sqrt{n+1}\rceil$:

$$
\begin{aligned}
f\left(\sum b_i\right) \rightarrow \sum_{ij}^{m} f\left((i-1)(m+1)+(j-1)\right) b_{a_i} b_{a_{c+j}} + \lambda\Bigg(\left(1-\sum_i^m b_{a_i}\right)^2 + \left(1-\sum_i^m b_{a_{c+i}}\right)^2 + \\
\left(\sum_i b_i - \left((m+1)\sum_i^m (i-1) y_{a_i} + \sum_i^m (i-1) b_{a_{c+i}}\right)\right)^2 + \left(\sum_i b_i - \left((m+1)\sum_i^m (i-1) y_{a_i} + \sum_i^m (i-1) b_{a_{c+i}}\right)\right)^2\Bigg)
\end{aligned} \tag{92}
$$

where:

$$
\begin{pmatrix} \alpha & \alpha^{bb} \\ \alpha^{b} & \alpha^{bb_{a,1}} \\ \alpha^{b_{a,1}} & \alpha^{bb_{a,2}} \\ \alpha^{b_{a,2}} & \alpha^{b_a b_a} \end{pmatrix} = \begin{pmatrix} (c+1)^2 & 1 \\ -2(c+1) & -2^i \\ 2(c+1) & -2\left(2m-2^{m-1}+1\right) \\ 2(c+1)\left(2c-2^{m-1}+1\right) & 2^{i+j-2} \end{pmatrix}. \tag{93}
$$

Cost

- $m = \lceil\sqrt{n+1}\rceil$ auxiliary variables.
- $n^2 + m^2$ non-submodular quadratic terms (all possible quadratic terms involving only non-auxiliary or only auxiliary variables).
- m non-submodular linear terms (all possible linear terms involving auxiliaries).

Pros

- Small number of auxiliary terms

Cons

- Only works for a special class of functions
- Introduces many linear and even more quadratic non-submodular terms.

Example

By bit-flipping b_2 and b_4, i.e. substituting $b_2 = 1 - \bar{b}_2$ and $b_4 = 1 - \bar{b}_4$, we see that:

$$
\begin{aligned}
H &= 3b_1b_2 + b_2b_3 + 2b_1b_4 - 4b_2b_4 &(94)\\
&= -3b_1\bar{b}_2 - \bar{b}_2b_3 - 2b_1\bar{b}_4 - \bar{b}_2\bar{b}_4 + 5b_1 + b_3 + 4\bar{b}_2 + 4\bar{b}_4 - 4. &(95)
\end{aligned}
$$

The first expression is highly non-submodular while the second is entirely submodular.

Alternate Forms

$$
f(b_1, b_2, \ldots, b_n) \rightarrow \left(\frac{\frac{1}{2}\left((c-1) - \sum_i b_i - \left(2(n-c) - 2^{m-1} + 1\right) b_{a_m} - \sum_i^{m-1} 2^{i-1} b_{a_i}\right)}{2} \right) \tag{96}
$$

Bibliography

- 2018, original paper (Theorem 9): [16].

G. SFR-BCR-6 (Boros, Crama, Rodríguez-Hector, 2018)

Summary

For an n-variable symmetric function that is a function of a *weighted* sum of all variables $f(b_1, b_2, \dots b_n) = f(\sum w_i b_i)$, for some large value of $\lambda > \max(f)$, and $\max(f(\sum w_i b_i)) < (m+1)^2$:

$$f\left(\sum w_i b_i\right) \rightarrow \sum_{ij}^{m} \alpha_{ij} b_{a_i} b_{a_{m+i}} + \lambda\Bigg(1 + \left(\sum_i w_i b_i - (m-1)\sum_i^m b_{a_i} + \sum_i^m b_{a_{c+i}}\right)^2 + \sum_i^{m-1}(1-b_{a_i})\, b_{a_{i+1}} + \sum_i^{m-1}(1-b_{a_{i+m}})\, b_{a_{i+m+1}}\Bigg) \tag{97}$$

where:

$$\sum_i^{\alpha}\sum_j^{\beta} \alpha_{ij} = f\left(\alpha(m+1) + \beta\right) \tag{98}$$

Cost

- $2m$ auxiliary variables, where $m > \sqrt{\max(f(\sum w_i b_i))} - 1$.
- $n^2 + m^2$ non-submodular quadratic terms (all possible quadratic terms involving only non-auxiliary or only auxiliary variables).
- m non-submodular linear terms (all possible linear terms involving auxiliaries).

Pros

- Small number of auxiliary terms

Cons

- Only works for a special class of functions
- Introduces many linear and even more quadratic non-submodular terms.

Example

By bit-flipping b_2 and b_4, i.e. substituting $b_2 = 1 - \bar{b}_2$ and $b_4 = 1 - \bar{b}_4$, we see that:

$$H = 3b_1b_2 + b_2b_3 + 2b_1b_4 - 4b_2b_4 \tag{99}$$

$$= -3b_1\bar{b}_2 - \bar{b}_2b_3 - 2b_1\bar{b}_4 - \bar{b}_2\bar{b}_4 + 5b_1 + b_3 + 4\bar{b}_2 + 4\bar{b}_4 - 4. \tag{100}$$

The first expression is highly non-submodular while the second is entirely submodular.

Alternate Forms

$$f(b_1, b_2, \dots, b_n) \rightarrow \begin{pmatrix} \frac{1}{2}\left((c-1) - \sum_i b_i - (2(n-c) - 2^{m-1} + 1)\, b_{a_m} - \sum_i^{m-1} 2^{i-1} b_{a_i}\right) \\ 2 \end{pmatrix} \tag{101}$$

Bibliography

- 2018, original paper (Theorem 10): [16].

H. SFR-ABCG-2 (Anthony, Boros, Crama, Gruber, 2014)

Summary

For any n-variable, k-local function that is non-zero only if $\sum b_i = 2m - 1$, we call it the "partity function" and it can be quadratized as follows:

$$f(b_1, b_2, \ldots, b_n) \to \sum_i b_i + 2\sum_{ij} b_i b_j + 4\sum_{2i-1}^{n-1} b_{a_i}\left(2i - 1 - \sum_j b_j\right). \tag{102}$$

Cost

- $m = 2\lfloor n/2 \rfloor$ auxiliary variables.
- $\lfloor 1.5n \rfloor$ non-submodular linear terms.
- n^2 non-submodular quadratic terms.

Pros

- Smaller number of auxiliary variables than the most naive methods.

Cons

- Only works for a special class of functions
- Introduces many linear and even more quadratic non-submodular terms (everything is non-submodular except for $0.5n^2$ quadratic terms involving the auxiliaries with the non-auxiliaries).
- non-submodular terms can be rather large compared to the submodular terms (about $4n$ times as big).

Example

By bit-flipping b_2 and b_4, i.e. substituting $b_2 = 1 - \bar{b}_2$ and $b_4 = 1 - \bar{b}_4$, we see that:

$$\begin{aligned} H &= 3b_1b_2 + b_2b_3 + 2b_1b_4 - 4b_2b_4 && (103) \\ &= -3b_1\bar{b}_2 - \bar{b}_2b_3 - 2b_1\bar{b}_4 - \bar{b}_2\bar{b}_4 + 5b_1 + b_3 + 4\bar{b}_2 + 4\bar{b}_4 - 4. && (104) \end{aligned}$$

The first expression is highly non-submodular while the second is entirely submodular.

Bibliography

- 2014, original paper (Theorem 4.6): [12, 13].

I. SFR-ABCG-3 (Anthony, Boros, Crama, Gruber, 2014)

Summary
The complement of the parity function can be quadratized as follow:

$$f(b_1, b_2, \ldots, b_n) \to 1 + 2\sum_{ij} b_i b_j - \sum_i b_i + 4\sum_{2i}^{n-1} b_{a_i}\left(i - \sum_j^n b_j\right) \quad (105)$$

Cost
- $m = 2\lfloor \frac{n-1}{2} \rfloor$ auxiliary variables.
- $\lfloor 0.5n \rfloor$ non-submodular linear terms.
- n^2 non-submodular quadratic terms.

Pros
- Smaller number of auxiliary variables than the most naive methods.
- Fewer non-submodular linear terms than in the analogous quadratization for its complement (the parity function).

Cons
- Only works for a special class of functions
- Introduces many linear and even more quadratic non-submodular terms (everything is non-submodular except for $0.5n^2$ quadratic terms involving the auxiliaries with the non-auxiliaries, and all n linear terms involving only the non-auxiliaries).
- non-submodular terms can be rather large compared to the submodular terms (about $4n$ times as big).

Example
By bit-flipping b_2 and b_4, i.e. substituting $b_2 = 1 - \bar{b}_2$ and $b_4 = 1 - \bar{b}_4$, we see that:

$$H = 3b_1b_2 + b_2b_3 + 2b_1b_4 - 4b_2b_4 \quad (106)$$
$$= -3b_1\bar{b}_2 - \bar{b}_2b_3 - 2b_1\bar{b}_4 - \bar{b}_2\bar{b}_4 + 5b_1 + b_3 + 4\bar{b}_2 + 4\bar{b}_4 - 4. \quad (107)$$

The first expression is highly non-submodular while the second is entirely submodular.

Bibliography
- 2014, original paper (Theorem 4.6): [12, 13].

J. Lower bounds for SFRs (Anthony, Boros, Crama, Gruber, 2014)

- There exist symmetric functions on n variables for which no quadratization can be done without at least $\Omega\left(\sqrt{n}\right)$ auxiliary variables (Theorem 5.3 from [12, 13].
- There exist symmetric functions on n variables for whcih no quadratization linear in the auxiliaries can be done without at least $\Omega\left(\frac{n}{\log_2(n)}\right)$ auxiliary variables (Theorem 5.5 from [12, 13]).
- The parity function on n variables cannot be quadratized without quadratic terms involving the auxiliary variables, unless there is at least $\sqrt{n/4 - 1}+1 = \Omega\left(\sqrt{n}\right)$ auxiliary variables (Theorem 5.6 from [12, 13]).

V. METHODS THAT QUADRATIZE MULTIPLE TERMS WITH THE SAME AUXILIARIES (CASE 2: ARBITRARY FUNCTIONS)

A. Reduction by Substitution (Rosenberg 1975)

Summary

Pick a variable pair (b_i, b_j) and substitute $b_i b_j$ with a new auxiliary variable $b_{a_{ij}}$. Enforce equality in the ground states by adding some scalar multiple of the penalty $P = b_i b_j - 2b_i b_{a_{ij}} - 2b_i b_{a_{ij}} + 3b_{a_{ij}}$ or similar. Since $P > 0$ if and only if $b_{a_{ij}} \neq b_i b_j$, the minimum of the new $(k-1)$-local function will satisfy $b_{a_{ij}} = b_i b_j$), which means that at the minimum, we have precisely the original function. Repeat $(k-2)$ times for each k-local term and the resulting function will be 2-local.

Cost

- 1 auxiliary variable per reduction.
- At most $k\,t$ auxiliary variables for a k-local Hamiltonian of t terms, but usually fewer.

Pros

- Variable can be used across the entire Hamiltonian, reducing many terms at once.
- Very easy to implement.
- Reproduces not only the ground state, but the full spectrum.

Cons

- Inefficient for single terms as it introduces many auxiliary variables compared to Ishikawa reduction, for example.
- Introduces quadratic terms with large positive coefficients, making them highly non-submodular.
- Determining optimal substitutions can be expensive.

Example

We pick the pair (b_1, b_2) and combine.

$$b_1 b_2 b_3 + b_1 b_2 b_4 \mapsto b_3 b_a + b_4 b_a + b_1 b_2 - 2b_1 b_a - 2b_1 b_a + 3b_a \tag{108}$$

Bibliography

- Original paper: [25]
- Re-discovered in the context of diagonal quantum Hamiltonians: [26].
- Used in: [27, 28].

B. FGBZ Reduction (Fix-Gruber-Boros-Zabih, 2011)

Summary
Here we consider a set C of variables which occur in multiple monomials throughout the Hamiltonian. Each application 'rips out' this common component from each term [29][15].

Let $\mathcal{H}$ be a set of monomials, where $C \subseteq H$ for each $H \in \mathcal{H}$ and each monomial H has a weight α_H. The algorithm comes in 2 parts: when all $\alpha_H > 0$ and when all $\alpha_H < 0$. Combining the 2 gives the final method:

1. $\alpha_H > 0$
$$\sum_{H\in\mathcal{H}} \alpha_H \prod_{j\in H} b_j = \min_{b_a} \left(\sum_{H\in\mathcal{H}} \alpha_H \right) b_a \prod_{j\in C} b_j + \sum_{H\in\mathcal{H}} \alpha_H (1 - b_a) \prod_{j\in H\setminus C} b_j \tag{109}$$

2. $\alpha_H < 0$
$$\sum_{H\in\mathcal{H}} \alpha_H \prod_{j\in H} b_j = \min_{b_a} \sum_{H\in\mathcal{H}} \alpha_H \left(1 - \prod_{j\in C} b_j - \prod_{j\in H\setminus C} b_j \right) b_a \tag{110}$$

Cost
- One auxiliary variable per application.
- In combination with II A, it can be used to make an algorithm which can reduce t positive monomials of degree d in n variables using $n + t(d-1)$ auxiliary variables in the worst case.

Pros
- Can reduce the connectivity of a Hamiltonian, as it breaks interactions between variables.

Cons
- $\alpha_H > 0$ method converts positive terms into negative ones of same order rather than reducing them, though these can then be reduced more easily.
- $\alpha_H < 0$ method only works for $|C| > 1$, and cannot quadratize cubic terms.

Example
First let $C = b_1$ and use the positive weight version:

$$\begin{aligned} b_1b_2b_3 + b_1b_2b_4 &\mapsto 2b_{a_1}b_1 + (1 - b_{a_1})b_2b_3 + (1 - b_{a_1})b_2b_4 && (111)\\ &= 2b_{a_1}b_1 + b_2b_3 + b_2b_4 - b_{a_1}b_2b_3 - b_{a_1}b_2b_4 && (112)\end{aligned}$$

now we can use II A:

$$\begin{aligned} -b_{a_1}b_2b_3 - b_{a_1}b_2b_4 &\mapsto 2b_{a_2} - b_{a_1}b_{a_2} - b_{a_2}b_2 - b_{a_2}b_3 + 2b_{a_2} - b_{a_1}b_{a_2} - b_{a_2}b_2 - b_{a_2}b_4 && (113)\\ &= 4b_{a_2} - 2b_{a_1}b_{a_2} - 2b_{a_2}b_2 - b_{a_2}b_3 - b_{a_2}b_4. && (114)\end{aligned}$$

Bibliography
- Original paper and application to image denoising: [29].

C. Flag Based SAT Mapping

Summary
This method uses gadgets to produce separate 3-SAT clauses which allow variables which 'flag' the state of pairs of other variables.

Cost
- Varies, but generally higher than other methods.

Pros
- Highly general and therefore conducive to proofs.

Cons
- Designed for generality rather than efficiency.

Example
To create a system which maps $b_1b_2b_3$, we use the following gadget (note that this is given in terms of z in the orginal work and translated to b here):

$$H_{3-\mathrm{SAT}}(b_1,b_2,b_3)=2\sum_{i=1}^{3}b_ib_{a_i}+2\sum_{i<j}^{3}b_{a_i}b_{a_j}-4\sum_{i=1}^{3}b_{a_i}-2\sum_{i=1}^{3}b_i+\frac{23}{2}. \tag{115}$$

Implementing $\lambda H_{3-\mathrm{SAT}}$, creates a situation where b_3 is a 'flag' for b_1 and b_2 in other words b_3 is constrained to be 1 in the low energy manifold if $b_1=0$ and $b_2=0$. It follows from the universality of $3-\mathrm{SAT}$ that these 'flag' clauses can be combined to map any spin Hamiltonian. To do this, we also need anti-ferromagnetic couplings to express the 'negated' variable, to do this, we define,

$$H_{\mathrm{AF}}(b_1,b_{\neg 1})=2\,b_1b_{\neg 1}-b_i-b_{\neg 1}+1. \tag{116}$$

As an explicit example, consider reproducing the spectrum of $z_1z_2z_3=(2\,b_1-1)(2\,b_2-1)(2\,b_3-1)$. In this case we need to assign a higher energy to the $(1,1,1)$, $(0,0,1)$, $(0,1,0)$, and $(1,0,0)$ states. A flag (b_f) which is forced into a higher energy state if these conditions are satisfied can be constructed from two instances of $H_{3-\mathrm{SAT}}$ and an auxilliary qubit, combining these leads to

$$\begin{gathered}H_{2-\mathrm{local}}(b_1,b_2,b_3)=\lambda[\sum_{i=1}^{3}H_{\mathrm{AF}}(b_i,b_{\neg i})+\sum_{i=1}^{8}H_{\mathrm{AF}}(b_{a_i},b_{\neg a_i})+H_{\mathrm{AF}}(b_f,b_{\neg f})\\ +H_{3-\mathrm{SAT}}(b_{\neg 1},b_{\neg 2},b_{\neg a_1})+H_{3-\mathrm{SAT}}(b_{a_1},b_{\neg 3},b_f)+H_{3-\mathrm{SAT}}(b_1,b_2,b_{\neg a_2})+H_{3-\mathrm{SAT}}(b_{a_2},b_3,b_{\neg f})\\ +H_{3-\mathrm{SAT}}(b_1,b_2,b_{\neg a_3})+H_{3-\mathrm{SAT}}(b_{a_3},b_{\neg 3},b_f)+H_{3-\mathrm{SAT}}(b_{\neg 1},b_{\neg 2},b_{\neg a_4})+H_{3-\mathrm{SAT}}(b_{a_4},b_3,b_{\neg f})\\ +H_{3-\mathrm{SAT}}(b_1,b_{\neg 2},b_{\neg a_5})+H_{3-\mathrm{SAT}}(b_{a_5},b_3,b_f)+H_{3-\mathrm{SAT}}(b_{\neg 1},b_2,b_{\neg a_6})+H_{3-\mathrm{SAT}}(b_{a_6},b_{\neg 3},b_{\neg f})\\ +H_{3-\mathrm{SAT}}(b_1,b_2,b_{\neg a_7})+H_{3-\mathrm{SAT}}(b_{a_7},b_{\neg 3},b_f)+H_{3-\mathrm{SAT}}(b_{\neg 1},b_{\neg 2},b_{\neg a_8})+H_{3-\mathrm{SAT}}(b_{a_8},b_3,b_{\neg f})]\\ -2\,b_f+1.\end{gathered}$$

The terms on the first line are all to enforce the relationship between a variable and its logical inverse. Each of the next four lines assigns a value to the flag variable b_f for a state and it's logical inverse, for instance the leftmost two terms of the second line enforce that $b_f = 0$ if $(b_1, b_2, b_3) = (0, 0, 0)$, while the right two terms enforce that $b_f = 1$ if $(b_1, b_2, b_3) = (1, 1, 1)$. Because there are $2^3 = 8$ possible bitstrings for (b_1, b_2, b_3), and each term to enforce a flag state requires two instances of $H_{3\text{-SAT}}$ (and two auxilliary bits), a total of 16 instances are required as well as 16 auxilliary bits. Finally the last line of this equation actually implements the penalty by adding an energy difference between the two states of b_f. Note that the total number of bits required to express this term is six for the logical bits and their inverses, 16 for the auxilliary qubits, and two for the flag and its inverse. The total number of bits to map this three local Hamiltonian is therefore 24.

Bibliography

- Paper showing the universality of the Ising spin models: [30].

D. Lower bounds for arbitrary functions (Anthony, Boros, Crama, Gruber, 2015)

- There exist functions on n variables for which no quadratization can be done without at least $\frac{2^{n/2}}{8} = \Omega\left(\sqrt{n}\right)$ auxiliary variables (Theorem 5.3 from [8].
- There exist symmetric functions on n variables for which no quadratization linear in the auxiliaries can be done without at least $\Omega\left(\frac{2^n}{n}\right)$ auxiliary variables (Theorem 5.5 from [8]).

VI. STRATEGIES FOR COMBINING METHODS

A. SCM-BCR (Boros, Crama, and Rodríguez-Heck, 2018)

Summary

Split a k-local monomial with odd k into a $(k-1)$-local term (with even degree) and a new odd k-local term which has negative coefficient:

$$b_1 b_2 \cdots b_k \to \prod_{i=1}^{k-1} b_i - \prod_{i=1}^{k-1} b_i (1 - b_k) \tag{117}$$

We can use any of the PTR methods for even k on the first term, and we can use any of the NTR methods on the second term. Can be generalized to split into different-degree factors when seeking an "optimum" quadratization. Can be generalized into more splits.

Cost

- Depends on the methods used for the PTR and NTR procedures.

Pros

- Very flexible.

Cons

- First turns one term into two terms, so might not be preferred when we wish to minimize the number of terms.

Bibliography

- Original paper: [16].

B. Decomposition into symmetric and anti-symmetric parts

Summary
Split a any function f into a symmetric part and anti-symmetric part:

$$f\left(b_1, b_2, \ldots, b_n\right) = f_{\text{symmetric}} + f_{\text{anti-symmetric}}, \tag{118}$$

$$f_{\text{symmetric}} \equiv \frac{1}{2}\left(f\left(b_1, b_2, \ldots, b_n\right) + f\left(1-b_1, 1-b_2, \ldots, 1-b_n\right)\right) \tag{119}$$

$$f_{\text{anti-symmetric}} \equiv \frac{1}{2}\left(f\left(b_1, b_2, \ldots, b_n\right) - f\left(1-b_1, 1-b_2, \ldots, 1-b_n\right)\right) \tag{120}$$

We can now use any of the methods described only for symmetric functions, on the symmetric part, and use the (perhaps less powerful) general methods on the anti-symmetric part.

Cost
- Depends on the methods used.

Pros
- Allows non-symmetric functions to benefit from techniques designed only for symmetric functions.

Cons
- May result in more terms than simply quadratizing the non-symmetric function directly.

Bibliography
- Discussed in: [14].

Part II

Hamiltonians quadratic in z and linear in x (Transverse Field Ising Hamiltonians)

The Ising Hamiltonian with a transverse field in the x direction is possible to implement in hardware:

$$H = \sum_i \left(\alpha_i^{(z)} z_i + \alpha_i^{(x)} x_i\right) + \sum_{ij} \left(\alpha_{ij}^{(zz)} z_i z_j\right). \tag{121}$$

C. ZZZ-TI-CBBK: Transvese Ising from ZZZ, by Cao, Babbush, Biamonte, and Kais (2015)

There is only one reduction in the literature for reducing a Hamiltonian term to the transverse Ising Hamiltonian, and it works on 3-local zzz terms, by introducing an auxiliary qubit with label a:

$$\alpha z_i z_j z_k \to \alpha^I + \alpha_i^z z_i + \alpha_j^z z_j + \alpha_k^z z_k + \alpha_a^z z_a + \alpha_a^x x_a + \alpha_{ia}^{zz} z_i z_a + \alpha_{ja}^{zz} z_j z_a + \alpha_{ka}^{zz} z_k z_a \tag{122}$$

$$\begin{aligned}
\alpha^I &= \tfrac{1}{2}\left(\Delta + \left(\tfrac{\alpha}{6}\right)^{2/5} \Delta^{3/5}\right)\\
\alpha_i^z &= -\tfrac{1}{2}\left(\left(\tfrac{7\alpha}{6} + \left(\tfrac{\alpha}{6}\right)^{3/5} \Delta^{2/5}\right) - \left(\tfrac{\alpha\Delta^4}{6}\right)^{1/5}\right)\\
\alpha_j^z &= \alpha_i^{(z)}\\
\alpha_k^z &= \alpha_i^{(z)}\\
\alpha_a^z &= \tfrac{1}{2}\left(\Delta - \left(\tfrac{\alpha}{6}\right)^{2/5} \Delta^{3/5}\right)\\
\alpha_a^x &= \left(\tfrac{\alpha\Delta^4}{6}\right)^{1/5}\\
\alpha_{ia}^{zz} &= -\tfrac{1}{2}\left(\left(\tfrac{7\alpha}{6} + \left(\tfrac{\alpha}{6}\right)^{3/5} \Delta^{2/5}\right) + \left(\tfrac{\alpha\Delta^4}{6}\right)^{1/5}\right)\\
\alpha_{ja}^{zz} &= \alpha_{ia}^{(zz)}\\
\alpha_{ka}^{zz} &= \alpha_{ja}^{(zz)}
\end{aligned}$$

Including all coefficients and factorizing, we get:

$$\alpha z_i z_j z_k \to \left(\Delta + \frac{\alpha\Delta^4}{6}^{1/5} (z_i + z_j + z_k)\right)\left(\frac{1 - z_a}{2}\right) + \frac{\alpha\Delta^4}{6}^{1/5} x_a \tag{123}$$

$$+ \left(\left(\frac{\alpha}{6}\right)^{2/5} \Delta^{3/5} - \left(\frac{7\alpha}{6} + \left(\frac{\alpha}{6}\right)^{3/5} \Delta^{2/5}\right)(z_i + z_j + z_k)\right)\left(\frac{1 + z_a}{2}\right) \tag{124}$$

The low-lying spectrum (eigenvalues ***and*** eigenvectors) of the right side of Eq. (121) will match those of the left side to within a spectral error of ϵ as long as $\Delta = \mathcal{O}(\epsilon^{-5})$.

Cost

- 1 auxiliary qubit
- 8 auxiliary terms not proportional to $\mathbb{1}$.

Part III

General Quantum Hamiltonians

VII. NON-PERTURBATIVE GADGETS

A. NP-OY (Ocko & Yoshida, 2011)

Summary
For the 8-body Hamiltonian:

$$H_{\text{8-body}} = -J\sum_{ij}\left(x_{ij3}x_{ij+1,2}x_{ij4}x_{ij+1,4}x_{i+1j1}x_{i+1j+1,1}x_{i+1j3}x_{ij+1,2}+\right.$$
$$\left.z_{ij1}z_{ij2}z_{ij3}z_{ij4}+z_{i-1j4}z_{ij1}+z_{ij2}z_{ij-1,3}+z_{ij4}z_{i+1j1}+z_{ij3}z_{ij+1,2}\right), \tag{125}$$

we define auxiliary qubits labeled by a_{ijk}, two auxiliaries for each pair ij: labeled a_{ij1} and a_{ij2}. Then the 8-body Hamiltonian has the same low-lying eigenspace as the 4-body Hamiltonian:

$$\begin{aligned}
H_{\text{4-body}} = -\sum_{ij}\alpha\left(z_{ij1}z_{ij2}z_{ij3}z_{ij4}+z_{i,j,-1,4}z_{ij1}+z_{ij2}z_{i,j-1,3}+z_{ij4}z_{i+1,j,1}+z_{ij3}z_{i,j+1,2}\right.\\
\left(1-z_{a_{ij1}}+z_{a_{ij2}}+z_{a_{ij1}}z_{a_{ij2}}\right)\left(z_{a_{i,j+1,1}}+z_{a_{i,j+1,2}}+z_{a_{i,j+1,1}}z_{a_{i,j+1,2}}-1\right)+\\
\left(1+z_{a_{ij1}}-z_{a_{ij2}}+z_{a_{ij1}}z_{a_{ij2}}\right)\left(1-z_{a_{i+1,j1}}-z_{a_{i+1,j2}}-z_{a_{i+1,j1}}z_{a_{i+1,j2}}\right)\Big)+\\
\frac{U}{2}\left(z_{a_{ij1}}+z_{a_{ij2}}+z_{a_{ij1}}z_{a_{ij2}}-1\right)+\\
\frac{t}{2}\Big(\left(x_{a_{ij2}}+z_{a_{ij1}}x_{a_{ij2}}\right)x_{ij3}x_{ij4}+\left(x_{a_{ij1}}x_{a_{ij2}}+y_{a_{ij1}}y_{a_{ij2}}\right)x_{i,j+1,2}x_{i,j+1,4}+\\
\left(x_{a_{ij2}}-z_{a_{ij1}}x_{a_{ij2}}\right)x_{i+1,j+1,1}x_{i+1,j+1,2}+\left(x_{a_{ij1}}x_{a_{ij2}}-y_{a_{ij1}}y_{a_{ij2}}\right)x_{i+1,j,1}x_{i+1,j,3}\Big)\Big).
\end{aligned} \tag{126}$$

Now by defining the following ququits (spin-3/2 particles, or 4-level systems):

$$s^{zz}_{ijki'j'k'} = z_{ijk}z_{i'j'k'} \tag{127}$$
$$s^{zz}_{a_{ij}1} = \left(1-z_{a1_{ij}}+z_{a2_{ij}}+z_{a1_{ij}}z_{a2_{ij}}\right) \tag{128}$$
$$s^{zz}_{a_{ij}2} = \left(z_{a1_{ij}}+z_{a2_{ij}}+z_{a1_{ij}}z_{a2_{ij}}-1\right) \tag{129}$$
$$s^{zz}_{a_{ij}3} = \left(1+z_{a1_{ij}}-z_{a2_{ij}}+z_{a1_{ij}}z_{a2_{ij}}\right) \tag{130}$$
$$s^{xz}_{a_{ij}1} = \left(x_{a2_{ij}}+z_{a1_{ij}}x_{a2_{ij}}\right) \tag{131}$$
$$s^{xz}_{a_{ij}2} = \left(x_{a2_{ij}}-z_{a1_{ij}}x_{a2_{ij}}\right) \tag{132}$$
$$s^{xx}_{ijki'j'k'} = x_{ijk}x_{i'j'k'} \tag{133}$$
$$s^{xy}_{a_{ij}1} = \left(x_{a1_{ij}}x_{a2_{ij}}+y_{a1_{ij}}y_{a2_{ij}}\right) \tag{134}$$
$$s^{xy}_{a_{ij}2} = \left(x_{a1_{ij}}x_{a2_{ij}}-y_{a1_{ij}}y_{a2_{ij}}\right) \tag{135}$$

NP-OY (Ocko & Yoshida, 2011) [Continued]

We can write the 4-body Hamiltonian on qubits as a 2-body Hamiltonian on ququits:

$$
\begin{aligned}
H_{\text{2-body}} = -\sum_{ij} \Big(\alpha \Big(& s^{zz}_{ij1ij2} s^{zz}_{ij3ij4} + s^{zz}_{ij-1,4ij1} + s^{zz}_{ij2ij-1,3} + s^{zz}_{ij4i+1j1} + s^{zz}_{ij3ij+1,2} + s^{zz}_{a_{ij}1} s^{zz}_{a_{ij+1}2} - s^{zz}_{a_{ij}3} s^{zz}_{a_{i+1j}3} \Big) \\
& + \frac{U}{2} s^{zz}_{a_{ij},1} + \frac{t}{2} \Big(s^{xz}_{a_{ij}1} s^{xx}_{ij3ij4} + s^{xy}_{a_{ij}1} s^{xx}_{ij+1,2ij+1,4} + s^{xz}_{a_{ij}2} s^{xx}_{i+1,j+1,1i+1,j+1,2} + s^{xy}_{a_{ij}2} s^{xx}_{i+1j1,i+1,j3} \Big) \Big).
\end{aligned}
\tag{136}
$$

The low-lying eigenspace of $H_{2-\text{body}}$ is *exactly* the same as for $H_{4-\text{local}}$.

Cost

- 2 auxiliary ququits for each pair ij.
- 6 more total terms (6 terms in the 8-body version becomes 12 terms: 11 of them 2-body and 1 of them 1-body).

Pros

- Non-perturbative. No prohibitive control precision requirement.
- Only two auxiliaries required for each pair ij.
- 8-body to 2-body transformation can be accomplished in 1 step, rather than a 1B1 gadget which would take 6 steps or an SD + $(3 \to 2)$ gadget combination which would take 4 steps.

Cons

- Increase in dimention from working with only 2-level systems (spin-1/2 particles or 2×2 matrices) to working with 4-level systems (spin-3/2 particles).
- Until now, only derived for a very specific Hamiltonian form.
- This appraoch may become more demanding for Hamiltonians that are more than 8-local.

Bibliography

- Original paper: [31].

B. NP-SJ (Subasi & Jarzynski, 2016)

Summary

Determine the k-local term, $H_{k\text{-local}}$, whose degree we wish to reduce, and factor it into two commuting factors: $H_{k'\text{-local}}H_{(k-k')\text{-local}}$, where k' can be as low as 0. Separate all terms that are at momst $(k-1)$-local into ones that commmute with one of these factors (it does not matter which one, but without loss of generality we assume it to be the $(k-k')$-local one) and ones that anti-commute with it:

$$H^{\text{commuting}}_{<k\text{-local}} + H^{\text{anti-commuting}}_{<k\text{-local}} + \alpha H_{k'\text{-local}}H_{(k-k')\text{-local}} \tag{137}$$

Introduce one auxiliary qubit labeled by a and the Hamiltonian:

$$\alpha x_a H_{k'\text{-local}} + H^{\text{commuting}}_{<k\text{-local}} + z_a H^{\text{anti-commuting}}_{<k\text{-local}} \tag{138}$$

no longer contains $H_{k\text{-local}}$ but $H^{\text{anti-commuting}}_{<k\text{-local}}$ is now one degree higher.

Cost

- 1 auxiliary qubit to reduce k-local term to $(k'+1)$-local where k' can even be 0-local, meaning the k-local term is reduced to a 1-local one.
- Raises the k-locality of $H^{\text{anti-commuting}}_{<k\text{-local}}$ by 1 during each application. It can become $(>k)$-local!

Pros

- Non-perturbative
- Can linearize a term of arbitrary degree in one step.
- Requires very few auxiliary qubits.

Cons

- Can introduce many new non-local terms as an expense for reducing only one k-local term.
- If the portion of the Hamiltonian that does not commute with the $(k-k')$-local term has termms of degree $k-1$ (which can happen if $k'=0$) they will all become k-local, so there is no guarantee that this method reduces k-locality.
- If any terms were more than 1-local, this method will not fully quadratize the Hamiltonian (it must be combined with other methods).
- It only works when the Hamiltonian's terms of degree at most $k-1$ all either commute or anti-commmute with the k-local term to be eliminated.

Example

$$4z_5 - 3x_1 + 2z_1y_2x_5 + 9x_1x_2x_3x_4 - x_1y_2z_3x_5 \rightarrow 9x_{a_1} + 4z_{a_4}z_5 - 3z_{a_3}x_1 - z_{a_3}x_{a_4}a3 + 2x_{a_3}x_5 \tag{139}$$

Bibliography

- Original paper, and description of the choices of terms and factors used for the given example [32].

C. NP-Nagaj-1 (Nagaj, 2010)

Summary The Feynman Hamiltonian can be written as [33]:

$$\frac{1}{4}\left(x_1x_2 - \mathrm{i}y_1x_2 + \mathrm{i}x_1y_2 + y_1y_2\right)U_{2-\text{local}} + \frac{1}{4}\left(x_1x_2 + \mathrm{i}y_1x_2 - \mathrm{i}x_1y_2 + y_1y_2\right)U_{2-\text{local}}^{\dagger}, \tag{140}$$

where $U_{2-\text{local}}$ is an arbitrary 2-local unitary matrix that acts on qubits different from the ones labeled by "1" and "2". This Hamiltonian that is 4-local on qubits can be transformed into one that is 2-local in qubits and qutrits. Here we show the 2-local Hamiltonian for the case where $U_{2-\text{local}} = \text{CNOT} \equiv \frac{1}{2}\left(\mathbb{1} + z_3 + x_4 - z_3x_4\right)$. We start with the specific 4-local Hamiltonian:

$$H_{4-\text{local}} = \frac{1}{4}\left(x_1x_2 + y_1y_2 + x_1x_2z_3 + x_1x_2x_4 + y_1y_2z_3 + y_1y_2x_4 - x_1x_2z_3x_4 - y_1y_2z_3x_4\right), \tag{141}$$

and after adding 4 auxiliary qubits labeled by a_1 to a_4 and 6 auxiliary qutrits labeled by a_5 to a_{10} and acted on by the Gell-Mann matrices λ_1 to λ_9, we get the following 2-local Hamiltonian:

$$H_{2-\text{local}} = {}^1\!/\!_2\,(2\lambda_{6,a_8} + x_1\lambda_{1,a_5} + y_1\lambda_{2,a_5} + \lambda_{6,a_5} - z_3\lambda_{6,a_5} + x_{a_1}\lambda_{4,a_5} + y_{a_1}\lambda_{5,a_5} + x_{a_1}\lambda_{1,a_6} + \tag{142}$$

$$y_{a_1}\lambda_{2,a_6} + 2x_4\lambda_{6,a_6} + x_{a_2}\lambda_{4,a_6} + y_{a_2}\lambda_{5,a_6} + x_{a_2}\lambda_{1,a_7} + y_{a_2}\lambda_{2,a_7} + \lambda_{6,a_7} - z_{a_1}\lambda_{6,a_7} + \tag{143}$$

$$x_2\lambda_{4,a_7} + y_2\lambda_{5,a_7} + x_1\lambda_{1,a_8} + y_1\lambda_{2,a_8} + z_3\lambda_{6,a_8} + x_{a_5}\lambda_{4,a_9} + y_{a_5}\lambda_{5,a_9} + \lambda_{6,a_9} + \tag{144}$$

$$x_{a_6}\lambda_{4,a_9} + y_{a_6}\lambda_{5,a_9} + x_{a_6}\lambda_{1,a_{10}} + y_{a_6}\lambda_{2,a_{10}} + \lambda_{6,a_{10}} + z_3\lambda_{6,a_{10}} + x_2\lambda_{4,a_{10}} + y_2\lambda_{5,a_{10}}), \tag{145}$$

whose low-lying spectrum is equivalent to the spectrum of $H_{2-\text{local}}$.

Cost

- 6 auxiliary qutrits and 4 auxiliary qubits
- 2 quartic, 4 cubic, and 2 quadratic terms becomes 27 quadratic terms and 5 linear terms in the Pauli-GellMann basis.

Pros

- Exact (non-perturbative). No special control precision demands.
- All coefficients are equal to each other, with a value of $^1\!/\!_2$, except one which is equal to 1.
- With more auxiliary qubits, can be further reduced to only containing qubits.

Cons

- Involves qutrits in all 32 terms.
- Only derived (so far) for the Feynman Hamiltonian.
- High overhead in terms of number of auxiliary qubits and number of terms.

Example

The transformation presented above was for the case of $U_{2-\text{local}} = \text{CNOT} \equiv \frac{1}{2}\left(\mathbb{1} + z_3 + x_4 - z_3x_4\right)$, but similar transformations can be derived for any arbitrary unitary matrix $U_{2-\text{local}}$.

Bibliography

- Original paper: [34].

D. NP-Nagaj-2 (Nagaj, 2012)

Summary

Similar to NP-Nagaj-1 but instead of using qutrits, we use two qubits for each qutrit, according to:

$$|0\rangle \to |00\rangle, \qquad |1\rangle \to \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right), \qquad |2\rangle \to \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right). \tag{146}$$

which leads to the following transformations:

$$|01\rangle\langle 10|_{ij} + h.c. \to \frac{1}{\sqrt{2}}\left(|01\rangle\langle 10|_{ij_1} + |01\rangle\langle 10|_{ij_2}\right) + h.c. \tag{147}$$

$$|02\rangle\langle 10|_{ij} + h.c. \to \frac{1}{\sqrt{2}}\left(|01\rangle\langle 10|_{ij_1} - |01\rangle\langle 10|_{ij_2}\right) + h.c. \tag{148}$$

$$|1\rangle\langle 2|_j + h.c. \to z_{j_1} - z_{j_2}, \tag{149}$$

and the following 2-local Hamiltonian involving only qubits:

$$\begin{aligned} H_{2-\mathrm{local}} = {}^1\!/_2 \big(& z_{a_5} - z_{a_6} + z_{a_9} - z_{a_{10}} - z_{a_3}z_{a_5} + z_{a_3}z_{a_6} - z_{a_3}z_{a_9} + z_{a_3}z_{a_{10}} + \\ & z_{a_{11}} - z_{a_{12}} + z_{a_{15}} - z_{a_{16}} + z_{a_3}z_{a_{11}} - z_{a_3}z_{a_{12}} + z_{a_3}z_{a_{15}} - z_{a_3}z_{a_{16}} \big) + x_4 z_{a_7} - x_4 z_{a_8} + z_{a_{13}} - z_{a_{14}} + \\ & {}^1\!/\!\sqrt{2} \big(x_1\lambda_{1,a_5} + y_1\lambda_{2,a_5} + x_1\lambda_{1,a_6} + y_1\lambda_{2,a_6} + x_{a_1}\lambda_{1,a_7} + y_{a_1}\lambda_{2,a_7} + x_{a_1}\lambda_{1,a_8} + y_1\lambda_{2,a_8} + \\ & x_{a_2}\lambda_{1,a_7} + y_{a_2}\lambda_{2,a_7} + x_{a_2}\lambda_{1,a_8} + y_{a_2}\lambda_{2,a_8} + x_{a_2}\lambda_{1,a_9} + y_{a_2}\lambda_{2,a_9} + x_{a_2}\lambda_{1,a_{10}} + \\ & y_{a_2}\lambda_{2,a_{10}} + x_1\lambda_{1,a_{11}} + y_1\lambda_{2,a_{11}} + x_{a_4}\lambda_{1,a_{15}} + y_{a_4}\lambda_{2,a_{15}} + x_{a_1}\lambda_{4,a_5} + y_{a_1}\lambda_{5,a_5} + \\ & x_{a_2}\lambda_{4,a_7} + y_{a_2}\lambda_{5,a_7} + x_2\lambda_{4,a_9} + y_2\lambda_{5,a_9} + x_{a_3}\lambda_{4,a_{11}} + y_{a_3}\lambda_{5,a_{11}} + x_{a_4}\lambda_{4,a_{13}} - \\ & y_{a_4}\lambda_{5,a_{13}} + x_{a_2}\lambda_{4,a_{15}} + y_{a_2}\lambda_{5,a_{15}} - x_{a_1}\lambda_{4,a_6} - y_{a_1}\lambda_{5,a_6} - x_{a_2}\lambda_{4,a_8} - y_{a_2}\lambda_{5,a_8} - x_2\lambda_{4,a_{10}} - \\ & -y_2\lambda_{5,a_{10}} - x_{a_3}\lambda_{4,a_{12}} - y_{a_3}\lambda_{5,a_{12}} - x_{a_4}\lambda_{4,a_{14}} - y_{a_4}\lambda_{5,a_{14}} - x_2\lambda_{4,a_{16}} - y_2\lambda_{5,a_{16}} \big), \end{aligned} \tag{150}$$

whose low-lying spectrum is equivalent to the spectrum of $H_{2-\mathrm{local}}$.

Cost

- 16 auxiliary qubits.

Pros

- Exact (non-perturbative). No special control precision demands.
- Only involves qubits (as opposed to NP-Nagaj-1 which contains qutrits and NP-OY which contains ququits.

Cons

- Only derived (so far) for the Feynman Hamiltonian.
- High overhead in terms of number of auxiliary qubits and number of terms.

Example

The transformation presented above was for the case of $U_{2-\mathrm{local}} = \mathrm{CNOT} \equiv \frac{1}{2}\left(\mathbb{1} + z_3 + x_4 - z_3x_4\right)$, but similar transformations can be derived for any arbitrary unitary matrix $U_{2-\mathrm{local}}$.

Bibliography

- Original paper: [35].

VIII. PERTURBATIVE $(3 \rightarrow 2)$ GADGETS

The first gadgets for arbitrary Hamiltonians acting on some number of qubits, were designed to reproduce the spectrum of a 3-local Hamiltonian in the low-lying spectrum of a 2-local Hamiltonian.

A. P$(3 \rightarrow 2)$-DC (Duan, Chen, 2011)

Summary

For any group of 3-local terms that can be factored into a product of three 1-local factors, we can define three auxiliary qubits (regardless of the number of qubits we have in total) labeled by a_i and make the transformation:

$$\prod_i^3 \sum_j \alpha_{ij} s_i \rightarrow \alpha + \alpha_i^{ss} \sum_i \left(\sum_j \alpha_{ij} s_{ij} \right)^2 + \alpha_i^{sx} \sum_i \sum_j \alpha_{ij} s_{ij} x_{a_i} + \alpha^{zz} \sum_{ij} z_{a_i} z_{a_j} \tag{151}$$

$$\alpha = \frac{1}{8\Delta} \tag{152}$$

$$\alpha^{ss} = \frac{1}{6\Delta^{1/3}} \tag{153}$$

$$\alpha^{sx} = -\frac{1}{6\Delta^{2/3}} \tag{154}$$

$$\alpha^{zz} = -\frac{1}{24\Delta} \tag{155}$$

The result will be a 2-local Hamiltonian whose low-lying spectrum is equivalent to the spectrum of $H_{3-\text{local}}$ to within ϵ as long as $\Delta = \Theta\left(\epsilon^{-3}\right)$.

Cost

- 1 auxiliary qubit for each group of 3-local terms that can be factored into three 1-local factors.
- $\Delta = \Theta\left(\epsilon^{-3}\right)$

Pros

- Very few auxiliary qubits needed

Cons

- Will not work for Hamiltonians that do not factorize appropriately.

Bibliography

- Original paper: [36]

B. $\mathbf{P}(3 \to 2)$-DC2 (Duan, Chen, 2011)

Summary

For any 3-local term (product of Pauli matrices s_i) in the Hamiltonian, we can define *one* auxiliary qubit labeled by a and make the transformation:

$$a \prod_{i}^{3} s_i \to \alpha + \alpha^s s_3 + \alpha^z z_a + \alpha^{ss} \left(s_1 + s_2\right)^2 + \alpha^{sz} s_3 z_a + +\alpha^{sx} \left(s_1 x_a + s_2 x_a\right) \tag{156}$$

$$\begin{pmatrix} \alpha & \alpha^s & \alpha^z \\ \alpha^{ss} & \alpha^{sz} & \alpha^{sx} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2\Delta} & a\left(\frac{1}{4\Delta^{2/3}} - 1\right) & a\left(\frac{1}{4\Delta^{2/3}} - 1\right) \\ \frac{1}{\Delta^{1/3}} & \frac{a}{4\Delta^{2/3}} & \frac{1}{\Delta^{2/3}} \end{pmatrix} \tag{157}$$

The result will be a 2-local Hamiltonian whose low-lying spectrum is equivalent to the spectrum of $H_{3-\text{local}}$ to within ϵ as long as $\Delta = \Theta\left(\epsilon^{-3}\right)$.

Cost

- 1 auxiliary qubit for each 3-local term.
- $\Delta = \Theta\left(\epsilon^{-3}\right)$

Example

$$x_1 z_2 y_3 - 3x_1 x_2 y_4 + z_1 x_2 \to \alpha + \alpha^z (z_{a_1} + z_{a_2}) + \alpha^y (y_3 + y_4) + \alpha^{zx}_{12} z_1 x_2 + \alpha^{zx} z_2 x_{a_1} + \alpha^{xx}_{11} x_1 x_2 \tag{158}$$

$$+ \alpha^{xx} \left(x_1 x_{a_1} + x_1 x_{a_2} + x_2 x_{a_2}\right) + \alpha^{yz} \left(y_3 z_{a_1} + y_4 z_{a_2}\right) \tag{159}$$

Bibliography

- Original paper: [36]

C. $\mathrm{P}(3 \to 2)$-KKR (Kempe, Kitaev, Regev, 2004)

Summary

For any 3-local term (product of commuting matrices s_i) in the Hamiltonian, we can define three auxiliary qubits labeled by a_i and make the transformation:

$$\prod_i^3 s_i \to \alpha + \alpha_i^{ss} \sum_i s_i^2 + \alpha_i^{sx} \sum_i s_i x_{a_i} + \alpha^{zz} \sum_{ij} z_{a_i} z_{a_j} \tag{160}$$

$$\alpha = -\frac{1}{8\Delta} \tag{161}$$

$$\alpha^{ss} = -\frac{1}{6\Delta^{1/3}} \tag{162}$$

$$\alpha^{sx} = \frac{1}{6\Delta^{2/3}} \tag{163}$$

$$\alpha^{zz} = \frac{1}{24\Delta} \tag{164}$$

The result will be a 2-local Hamiltonian whose low-lying spectrum is equivalent to the spectrum of $H_{3-\text{local}}$ to within ϵ as long as $\Delta = \Theta\left(\epsilon^{-3}\right)$.

Cost

- 3 auxiliary qubits for each 3-local term.
- $\Delta = \Omega\left(\epsilon^{-3}\right)$

Example

$$x_1 z_2 y_3 - 3x_1 x_2 y_4 + z_1 x_2 \to \alpha + \alpha^{zx}_{2a_{12}} z_2 x_{a_1 2} + \alpha^{xx}_{12} x_1 x_2 + \alpha^{xx}_{1a_{11}} x_1 x_{a_{11}} + \alpha^{xx}_{1a_{21}} x_1 x_{a_{21}} + \alpha^{xx}_{2a_{22}} x_2 x_{a_{22}} + \alpha^{yz}_{3a_{13}} y_3 x_{a_{13}} + \alpha^{yx}_{4a_{23}} y_4 x_{a_{23}} \tag{165}$$

D. P$(3\rightarrow 2)$-OT (Oliveira-Terhal, 2005)

Summary

For any 3-local term which is a product of 1-local matrices s_i, we can define one auxiliary qubit labeled by a and make the transformation:

$$a\prod_{i}^{3}s_i \rightarrow \alpha+\alpha_1^s s_1^2+\alpha_2^s s_2^2+\alpha_3^s s_3+\alpha_a^z z_a+\alpha_{12}^{ss}s_1s_2+\alpha_{13}^{ss}s_1^2s_3+\alpha_{23}^{ss}s_2^2s_3+\alpha_{3a}^{sz}s_3z_a+\alpha_{1a}^{sx}s_1x_a+\alpha_{2a}^{sx}s_2x_a \quad (166)$$

$$\begin{pmatrix}\alpha & \alpha_{12}^{ss}\\ \alpha_1^s & \alpha_{13}^{ss}\\ \alpha_2^s & \alpha_{23}^{ss}\\ \alpha_3^s & \alpha_{3a}^{sz}\\ \alpha_a^z & \alpha_{2a}^{sx}\\ \text{N/A} & \alpha_{2a}^{sx}\end{pmatrix}=\begin{pmatrix}\frac{\Delta}{2} & -\Delta^{1/3}\\ -\frac{\alpha^{2/3}\Delta^{1/3}}{2} & a\frac{1}{2}\\ \frac{\alpha^{2/3}\Delta^{1/3}}{2} & a\frac{1}{2}\\ -\frac{\alpha^{1/3}\Delta^{2/3}}{2} & \frac{\alpha^{1/3}\Delta^{2/3}}{2}\\ -\frac{\Delta}{2} & -\frac{\alpha^{1/3}\Delta^{2/3}}{\sqrt{2}}\\ \text{N/A} & \frac{\alpha^{1/3}\Delta^{2/3}}{\sqrt{2}}\end{pmatrix}. \quad (167)$$

A k-local Hamiltonian with a 3-local term replaced by this 2-local Hamiltonian will have an equivalent low-lying spectrum to within ϵ as long as $\Delta=\Omega\left(\epsilon^{-3}\right)$.

Cost

- 1 auxiliary qubit for each 3-local term.
- $\Delta=\Omega\left(\epsilon^{-3}\right)$

Example

Bibliography

- Original paper where $\alpha=1$: [37]. For arbitrary α see the 2005 v1 from arXiv, or [38]. Connection to improved version: [39].

IX. PERTURBATIVE 1-BY-1 GADGETS

A 1B1 gadget allows k-local terms to be quadratized one step at a time, where at each step the term's order is reduced by at most one. In each step, a k-local term is reduced to $(k-1)$-local, contrary to SD (sub-division) gadgets which can reduce k-local terms to $(1/2)$-local in one step.

A. P1B1-OT (Oliveira & Terhal, 2008)

Summary
We wish to reduce the k-local term:

$$H_{k\text{-local}} = \alpha \prod_{j}^{k} s_j. \tag{168}$$

Define one auxiliary qubit labeled by a and make the transformation:

$$H_{k\text{-local}} \to -\left(\frac{\alpha}{2}\right)^{1/3} \Delta^{2(1-r)} s_k \left(\frac{1-z_a}{2}\right) + \left(\frac{\alpha}{2}\right)^{1/3} \frac{\Delta^r}{\sqrt{2}} \left(s_{k-1} - s_{k-2}\right) x_a \tag{169}$$

$$+ \frac{1}{2}\left(\frac{\alpha}{2}\right)^{2/3} \left(\Delta^{r-1} s_{k-1} + \mathrm{sgn}(\alpha)\sqrt{2}\Delta^{-1/4} \prod_{j}^{k-2} s_j\right)^2 + \frac{\alpha}{4}\left(1 + 2\mathrm{sgn}^2\alpha\Delta^{3/2-2r}\right) s_k. \tag{170}$$

The result will be a $(k-1)$-local Hamiltonian with the same low-lying spectrum as $H_{k\text{-local}}$ to within ϵ as long as $\Delta = \Omega\left(\epsilon^{-3}\right)$.

Cost
- Only 1 auxiliary qubit.
- $\Delta = \Omega\left(\epsilon^{-3}\right)$

Example

Bibliography
- Described in: [39], based on: [37].

B. P1B1-CBBK (Cao, Babbush, Biamonte, Kais, 2015)

Summary
Define one auxiliary qubit labeled by a and make the transformation:

$$H_{k-\mathrm{local}} \to \left(\Delta + \left(\frac{\alpha}{2}\right)^{3/2} \Delta^{1/2} s_k\right)\left(\frac{1-z_a}{2}\right) \tag{171}$$

$$- \frac{\alpha^{2/3}}{2}\left(1 + \mathrm{sgn}^2\alpha\right)\left((2\alpha)^{2/3}\,\mathrm{sgn}^2\alpha + \alpha^{1/3} s_k - \sqrt[3]{2}\Delta^{1/2}\right)\left(\frac{1+z_a}{2}\right) \tag{172}$$

$$+ \left(\frac{\alpha}{2}\right)^{1/3} \Delta^{3/4}\left(\prod_{j}^{k-2} s_j - \mathrm{sgn}(\alpha) s_{k-1}\right) x_a + \mathrm{sgn}(\alpha)\sqrt[3]{2}\alpha^{2/3}\left(\Delta^{1/2} + \Delta^{3/2}\right)\prod_{j}^{k-1} s_j. \tag{173}$$

The result is $(k-1)$-local and its low-lying spectrum is the same as that of $H_{k-\mathrm{local}}$ when Δ is large enough.

Cost
- Only 1 auxiliary qubit.
- $\Delta = \Omega\left(\epsilon^{-3}\right)$

Example

Bibliography
- Described in: [39], based on: [37].

X. PERTURBATIVE SUBDIVISION GADGETS

Instead of recursively reducing k-local to $(k-1)$-local one reduction at a time, we can reduce k-local terms to $(k/2)$-local terms directly for even k, or to $(k+1)/2$-local terms directly for odd k. Since when k is odd we can add an identity operator to the k-local term to make it even, we will assume in the following that k is even, in order to avoid having to write floor and ceiling functions.

A. PSD-OT (Oliveira & Terhal, 2008)

Summary

We factor a k-local term into a product of three factors: operators H_1, H_2 acting on non-overlapping spaces, and scalar α. Then introduce an auxiliary qubit labelled by a and make the transformation:

$$H_{k\text{-local}} \to \Delta \frac{1-z_a}{2} + \frac{\alpha}{2}H_1^2 + \frac{\alpha}{2}H_2^2 + \sqrt{\frac{\alpha\Delta}{2}}\left(-H_1 + H_2\right)x_a. \tag{174}$$

The resulting Hamiltonian has a degree of 1 larger than the degree of whichever factor H_1 or H_2 has a larger degree, and the low-lying spectrum is equivalent to the original one to within $\mathcal{O}\left(\alpha\epsilon\right)$ for sufficiently large Δ.

Cost

- 1 auxiliary qubit for each k-local term that can be factored into two non-overlapping subspaces, is enough to reduce the degree down to $k/2 + 1$.

- $\Delta = \frac{\alpha\left(\|H_{(\text{else})} + \Omega(\sqrt{2})\max(\|H_1\|,\|H_2\|)\|\right)^6}{\epsilon^2} = \Omega\left(\alpha\epsilon^{-2}\right)$.

Pros

- Potentially very few auxiliary qubits needed.

Cons

- Requires the ability to factor k-local terms into non-overlapping subspaces that are at most $(k-2)$-local in order to reduce k-locality. This is not possible for $z_1x_2x_3 + z_2z_3x_4$, for example.

- Δ needs to be rather large.

- Cannot reduce 3-local to 2-local unless we generalize to a factor of 3 non-overlapping subspaces instead of 2. Needs to be combined with $3- \to 2$ gadgets, for example.

- A lot of work may be needed to find the optimal reduction, since each k-local term can be factored in many ways, and some of these ways may affect the ability to reduce other k-local terms.

Example

Bibliography

- Original paper where $\alpha = 1$: [37]. For arbitrary α see the 2005 v1 from arXiv, or [38]. Connection to improved version: [39].

B. PSD-CBBK (Cao, Babbush, Biamonte, Kais 2015)

Summary
For any k-local term, we can subdivide it into a product of two $(k/2)$ –local terms:

$$H_{k-\text{local}} = \alpha H_{1,(k/2)-\text{local}} H_{2,(k/2)-\text{local}} + H_{(k-1)-\text{local}}. \tag{175}$$

Define one qubit a and make the following Hamiltonian is $(k/2)$-local:

$$\Delta \frac{1 - z_a}{2} + |\alpha| \frac{1 + z_a}{2} + \sqrt{|\alpha|\Delta/2} \left(\text{sgn}(\alpha) H_{1,(k/2-\text{local})} - H_{2,(k/2-\text{local})} \right) x_a \tag{176}$$

The result is a $(k/2)$-local Hamiltonian with the same low-lying spectrum as $H_{k-\text{local}}$ for large enough Δ. The disadvantage is that Δ has to be larger.

Cost
- $\Delta \geq \left(\frac{2|\alpha|}{\epsilon} + 1 \right) \left(|\alpha| + \epsilon + 22||H_{(k-1)-\text{local}} \right)$

Pros
- only one qubit to reduce k to $\lceil k/2 \rceil + 1$

Cons
- Only beneficial for $k \geq 5$.

Example

Bibliography
- Original paper: [39].

XI. PERTURBATIVE DIRECT GADGETS

Here we do not reduce k by one order at a time (1B1 reduction) or by $k/2$ at a time (SD reduction), but we directly reduce k-local terms to 2-local terms.

A. PD-JF (Jordan & Farhi, 2008)

Summary
Express a sum of k-local terms as a sum of products of Pauli matrices s_{ij}, and define k auxiliary qubits laelled by a_{ij} for each term i, and make the transformmation:

$$\sum_i \alpha_i \prod_j^k s_{ij} \to \frac{-k(-\epsilon)^k}{(k-1)!} \sum_i \frac{1}{2} \left(k^2 - \sum_{jl}^k z_{a_{ij}} z_{a_{il}} \right) + \epsilon \left(\alpha_i s_{i1} x_{i1} + \sum_j^k s_{ij} x_{ij} \right) - f(\epsilon)\Pi, \tag{177}$$

for some polynomial $f(\lambda)$. The result is a 2-local Hamiltonian with the same low-lying spectrum to within ϵ^{k+1} for sufficiently smmall ϵ.

Cost
- Number of auxiliary qubits is tk for t terms.
- Unknown requirement for ϵ.

Pros
- All done in one step, so easier to implement than 1B1 and SD gadgets.

Cons
- Requires 2 more auxiliary qubits per term than 1B1-KKR.
- Unknown polynomial $f(\lambda)$

Example

Bibliography
- Original paper [40].

B. PD-BFBD (Brell, Flammia, Bartlett, Doherty, 2011)

Summary
The 4-body Hamiltonian:

$$H_{\text{4-local}} = -\sum_{ij}\left(z_{4i+1,j}z_{4i+2,j}z_{4i+3,j}z_{4i+4,j} + x_{4i+3,j}x_{4i+4,j}x_{4i+6,j}x_{4i+4,j+1}\right) \tag{178}$$

is transformed into the 2-body Hamiltonian:

$$H_{\text{2-local}} = -\sum_{ij}\left(x_{8i+4,j}x_{8i+6,j} + x_{8i+3,j+1}x_{8i+5,j+1} + z_{8i+4,j}z_{8i+3,j+1} + z_{8i+6,j}z_{8i+5,j+1}+\right. \tag{179}$$

$$= -\lambda\left(x_{8i+1,j}x_{8i+3,j} + x_{8i+2,j}x_{8i+4,j} + x_{8i+5,j}x_{8i+6} + x_{8i+7}x_{8i+8}\right. \tag{180}$$

$$= \left.\left. z_{8i+1,j}z_{8i+3,j} + z_{8i+2,j}z_{8i+4,j} + z_{8i+5,j}z_{8i+6} + z_{8i+7}z_{8i+8}\right)\right) \tag{181}$$

For $\lambda = \mathcal{O}\left(\epsilon^{-5}\right)$, the 2-local Hamiltonian has the same low-lying spectrum as the 4-local Hamiltonian, to within an error of ϵ.

Cost
- In total, uses four times the number of qubits of the original Hamiltonian.
- Unknown requirement for λ.

Pros
- All done in one step, so easier to implement than implementing two 1B1 gadgets.
- Very symmetric

Cons
- Ordinary 1B1 or SD followed by 3→2 gadgets would require half as many total qubits.
- Many 2-local terms.
- Perturbative, as opposed to NR-OY which is similar but does not involve any λ parameter.
- Required value of λ for it to work, is presently unknown.

Bibliography
- Original paper: [41].

Part IV

Appendix

XII. FURTHER EXAMPLES

Example Here we show how deductions can arise naturally from the Ramsey number problem. Consider $\mathcal{R}(4,3)$ with $N=4$ nodes. Consider a Hamiltonian:

$$H = (1-z_{12})(1-z_{13})(1-z_{23}) + \ldots + (1-z_{23})(1-z_{24})(1-z_{34}) + z_{12}z_{13}z_{14}z_{23}z_{24}z_{34}. \tag{182}$$

See [7] for full details of how we arrive at this Hamiltonian.

Since we are assuming we have no 3-independent sets, we know that $(1-z_{12})(1-z_{13})(1-z_{23}) = 0$, so $z_{12}z_{13}z_{23} = z_{12}z_{13} + z_{12}z_{23} + z_{13}z_{23} - z_{12} - z_{13} - z_{23} + 1$. This will be our deduction.

Using deduc-reduc we can substitute this into our 6-local term to get:

$$H = 2(1-z_{12})(1-z_{13})(1-z_{23}) + \ldots + (1-z_{23})(1-z_{24})(1-z_{34}) + \tag{183}$$

$$z_{14}z_{24}z_{34}(z_{12}z_{13} + z_{12}z_{23} + z_{13}z_{23} - z_{12} - z_{13} - z_{23} + 1). \tag{184}$$

We could repeat this process to remove all 5- and 4-local terms without adding any auxiliary qubits. Note in this case the error terms added by deduc-reduc already appear in our Hamiltonian.

XIII. $2 \to 2$ GADGETS

This review has only focused on $k-$local to $2-$local transformations where $k > 2$. There is also a large number of $2-$local to $2-$local transformations in the literature, which are used for various pruposes. Some of these are listed here:

- Gadgetization of any $2-$local Hamiltonian into $\{\mathbb{1}, z, x, zz, xx\}$ or $\{\mathbb{1}, z, x, zx\}$ [42]. Used for the proof that $xx + zz$ or xz is universal is enough for universal quantum computation. In other words, *any* computation can be transformed into a problem of finding the ground state of a $2-$local Hamiltonian containing terms from $\{\mathbb{1}, z, x, zz, xx\}$ or from $\{\mathbb{1}, z, x, zx\}$ with real coefficients, and the ground state can be found by adiabatic quantum computing with only polynomial time and space overhead over the best alternative algorithm for the problem.
- Transformation of any $2-$local Hamiltonian into $\{\mathbb{1}, z, x, zz, xx + yy\}$, without any perturbative gadgets, and only requiring the qubits to be connected in an almost 2D lattice [43].
- "Cross gadget", "fork gadget", and "triangle gadget" described in [37].
- Gadgetizeation of a $2-$local Hamiltonian with very strong couplings, into a $2-$local Hamiltonian with strengths in $\mathcal{O}\left(1/\text{poly}(\epsilon^{-1}, n)\right)$, and poly (ϵ^{-1}, n) auxiliary qubits and poly (ϵ^{-1}, n) new quadratic terms. [44].
- yy creation gadget: Simulation of yy terms using $\{\mathbb{1}, z, x, zz, xx\}$, with coupling strength restriction defined according to $\Delta = \Theta\left(\epsilon^{-4}\right)$ [39].

XIV. FURTHER REFERENCES

- Gadgets for pseudo-Boolean optimization problems, with reduced precision requirements: [45].
- Formalization of pseudo-Boolean gadgets in quantum language. [26].
- By adding more couplers and more auxiliary qubits, we can bring the error down arbitrarily low: [44].
- More toric code gadgets: [46].
- Parity adiabatic quantum computing (LHZ lattice): [47].
- Extensions of the LHZ scheme: [48].
- Minimizing k-local discrete functions with the help of continuous variable calculus [49].
- ORI graph which attempts to give optimal quadratizations [17].
- Survey on pseudo-Boolean optimization [50].
- Linearization of equations before they are squared [51], and its application to factoring numbers [52].
- Mentioned in [5] as an early application of quadratization: [53].

ACKNOWLEDGMENTS

- It is with immense pleasure that we thank Emile Okada of Cambridge University, who during his first year of udnergraduate study, worked with Nike Dattani and Richard Tanburn on quadratization of pseudo-boolean functions for quantum annealing, and in the first half of 2015 played a role in the development of the Deduc-Reduc and Split-Reduc and Groebner bases methods presented in this review.

- We thank Gernot Schaller of University of Berlin, who in December 2014 provided Nike Dattani with insights into his quadratization methods mentioned in this review paper, as well as for sharnig his Mathematica code which could be used to generate such quadratization formulas and others.

- We thank Mohammad Amin of D-Wave for pointing Nike Dattani to the paper [28] on determining Ramsey numbers on the D-Wave device, which contained what we call in this review "Reduction by substitution", later found through Ishikawa's paper to be from the much older 1970s paper by Rozenberg.

- We thank Catherine McGeoch of Amherst University and D-Wave, for helpful discussions with Nike Dattani in December 2014 about how to map quadratic pseudo-Boolean optimization problems onto the chimera graph of the D-Wave hardware and for pointing us to the important references of Vicky Choi. While chimerization is very different from quadratization, understanding that roughly n^2 variables would be needed to map a quadratic function of n variables, helped Nike Dattani and Richard Tanburn to appreciate how impotrant it is to be able to quadratize with as few variables as possible, and having this in mind throughout our studies helped inspire us in our goals towards "optimal quadratization".

- We thank Aritanan Gruber and Endre Boros of Rutgers University, who in August 2015 shared with Nike Dattani and Richard Tanburn some of their wisdom about sub-modularity, and Aritanan Gruber for pointing us to the then very recent paper of Hiroshi Ishikawa on what we call in this review "ELC reductions", which was also a valuable paper due to the references in it. We also thank him for helping us in our quest to determine whether or not "deduc-reduc" was a re-discovery by Richard, Emile, and Nike, or perhaps a novel quadratization scheme.

- We thank Toby Cathcart-Burn of Oxford University, who during the third year of undergraduate study, worked with Nike Dattani and Richard Tanburn and in Autumn 2015 and Winter 2016 helped us gain insights about the application of deduc-reduc and bit flipping to the problem of determining Ramsey numbers via discrete optimization, and for insights into the trade-offs between Ishikawa's symmetric reduction and reduction by substitution.

- We thank Hiroshi Ishikawa of Waseda University, who Nike Dattani had the memorable opportunity to visit in November 2015, and through discussions about the computer vision problem and neural network problem (two examples of real-world discrete optimization problems that benefit from quadratization), provided insights about the role of quadratization for calculations on classical computers. In particular, solving the computer vision problem in which he had experience solving on classical computers, was very different from the integer factoring problem and Ramsey number problem which we had been attempting to quadratize for D-Wave and NMR devices. He taught us that far more total (original plus auxiliary) variables can be tolerated on classical computers than on D-Wave machines or NMR systems, and approximate solutions to the optimization problems are acceptable (unlike for the factorization and Ramsey number problems in which we were interested). This gave us more insight into what trade-offs one might wish to prioritize when quadratizing optially. We also thank him for helping us in our quest to determine whether or not "deduc-reduc" was a re-discovery by Richard, Emile, and Nike, or perhaps a novel quadratization scheme.

- We thank Paul Ayers of McMaster University, who has been a major supporter of this review paper since the idea was presented to him in October 2016. At many points during the preparation of this review, we had prioritized other commitments and put preparation of this review aside. Paul's frequent encouragement was often what got us working on this review again. Without him, we are not certain this paper would have been completed by this time (or ever!). Paul also provided to us other interesting quantum mechianics problems in quantum chemistry and particle physics which served as a major motivation for continuing with this review when writing it was in a state of hiatous. These problems gave us more ideas about what trade-offs in quadratization might be important for some of its potential users in quantum chemistry and particle physics.

- Last but indubitably not least, we thank Jacob Biamonte of Skolkovo Institute of Technology, who Nike Dattani enjoyed visiting in Hangzhou in January 2017 and meeting at Harvard University in April 2018. Jacob provided us plenty of insights about perturbative gadgets, made valuable comments on early versions of our manuscript, and has been a major supporter of this review paper since the idea was presented to him in December 2016. At many points during the preparation of this review, we had prioritized other commitments and put preparation of this review aside. Jake's frequent encouragement was often what got us working on this review again. Without him, we are not certain this paper would have been completed by this time (or ever!).

[1] Richard Tanburn, Emile Okada, and Nike Dattani, "Reducing multi-qubit interactions in adiabatic quantum computation without adding auxiliary qubits. Part 1: The "deduc-reduc" method and its application to quantum factorization of numbers," (2015), arXiv:1508.04816.

[2] Hiroshi Ishikawa, "Higher-Order Clique Reduction without Auxiliary Variables," in *2014 IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2014) pp. 1362–1369.

[3] Magali Bardet, "On the Complexity of a Gröbner Basis Algorithm," Algorithms Seminar , 85–92 (2002).

[4] Raouf Dridi and Hedayat Alghassi, "Prime factorization using quantum annealing and computational algebraic geometry," (2016), arXiv:1604.05796.

[5] Asem M. Ali, Aly A. Farag, and Georgy L. Gimel'farb, "Optimizing Binary MRFs with Higher Order Cliques," (Springer, Berlin, Heidelberg, 2008) pp. 98–111.

[6] Richard Tanburn, Oliver Lunt, and Nikesh S. Dattani, "Crushing runtimes in adiabatic quantum computation with Energy Landscape Manipulation (ELM): Application to Quantum Factoring," Physical Review A (submitted). (2015), arXiv:1510.07420.

[7] Emile Okada, Richard Tanburn, and Nikesh S. Dattani, "Reducing multi-qubit interactions in adiabatic quantum computation without adding auxiliary qubits. Part 2: The "split-reduc" method and its application to quantum determination of Ramsey numbers," , 5 (2015), arXiv:1508.07190.

[8] Martin Anthony, Endre Boros, Yves Crama, and Aritanan Gruber, "Quadratic reformulations of nonlinear binary optimization problems," Mathematical Programming **162**, 115–144 (2017).

[9] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" IEEE Transactions on Pattern Analysis and Machine Intelligence **26**, 147–159 (2004).

[10] D. Freedman and P. Drineas, "Energy Minimization via Graph Cuts: Settling What is Possible," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 2 (IEEE, 2005) pp. 939–946.

[11] H Ishikawa, "Transformation of General Binary MRF Minimization to the First-Order Case," IEEE Transactions on Pattern Analysis and Machine Intelligence **33**, 1234–1249 (2011).

[12] Martin Anthony, Endre Boros, Yves Crama, and Aritanan Gruber, "Quadratization of symmetric pseudo-Boolean functions," Discrete Applied Mathematics **203**, 1–12 (2016).

[13] Martin Anthony, Endre Boros, Yves Crama, and Aritanan Gruber, "Quadratization of Symmetric Pseudo-Boolean Functions," (2014), arXiv:1404.6535.

[14] Kahl and Strandmark, "Generalized Roof Duality for Pseudo-Boolean Optimization," International Conference on Computer Vision (2011).

[15] Endre Boros and Aritanan Gruber, "On Quadratization of Pseudo-Boolean Functions," (2014), arXiv:1404.6538.

[16] Endre Boros, Yves Crama, and Elisabeth Rodríguez-Heck, "Quadratizations of symmetric pseudo-boolean functions: sub-linear bounds on the number of auxiliary variables," in *ISAIM* (2018).

[17] Andrew C. Gallagher, Dhruv Batra, and Devi Parikh, "Inference for order reduction in Markov random fields," in *CVPR 2011* (IEEE, 2011) pp. 1857–1864.

[18] Nicholas Chancellor, Stefan Zohren, and Paul A. Warburton, "Circuit design for multi-body interactions in superconducting quantum annealing system with applications to a scalable architecture," (2016), 10.1038/s41534-017-0022-6, arXiv:1603.09521.

[19] Martin Leib, Peter Zoller, and Wolfgang Lechner, "A Transmon Quantum Annealer: Decomposing Many-Body Ising Constraints Into Pair Interactions," (2016), arXiv:1604.02359.

[20] Martin Leib, Peter Zoller, and Wolfgang Lechner, "A transmon quantum annealer: decomposing many-body Ising constraints into pair interactions," Quantum Science and Technology **1**, 15008 (2016).

[21] N Chancellor, S Zohren, and P A Warburton, "Circuit design for multi-body interactions in superconducting quantum annealing systems with applications to a scalable architecture," npj Quantum Information **3**, 21 (2017).

[22] Nicholas Chancellor, "Max-k-SAT, Multi-Body Frustration, & Multi-Body Sampling on a Two Local Ising System," AQC 2016 https://www.youtube.com/watch?v=aC-6hg_h3EA (2016).

[23] N. Chancellor, S. Zohren, P.A. Warburton, S.C. Benjamin, and S. Roberts, "A Direct Mapping of Max k-SAT and High Order Parity Checks to a Chimera Graph," Scientific Reports **6** (2016), 10.1038/srep37107.

[24] Martin Anthony, Endre Boros, Yves Crama, and Aritanan Gruber, "Quadratic reformulations of nonlinear binary optimization problems," .

[25] I. G. Rosenberg, "Reduction of Bivalent Maximization to the Quadratic Case," Cahiers du Centre d'Etudes de Recherche Operationnelle **17**, 71–74 (1975).

[26] J D Biamonte, "Nonperturbative k-body to two-body commuting conversion Hamiltonians and embedding problem instances into Ising spins," Phys. Rev. A **77**, 52331 (2008).

[27] Alejandro Perdomo, Colin Truncik, Ivan Tubert-Brohman, Geordie Rose, and Alán Aspuru-Guzik, "Construction of model Hamiltonians for adiabatic quantum computation and its application to finding low-energy conformations of lattice protein models," Phys. Rev. A **78**, 12320 (2008).

[28] Zhengbing Bian, Fabian Chudak, William G. Macready, Lane Clark, and Frank Gaitan, "Experimental Determination of Ramsey Numbers," Physical Review Letters **111**, 130505 (2013).

[29] Alexander Fix, Aritanan Gruber, Endre Boros, and Ramin Zabih, "A graph cut algorithm for higher-order Markov Random Fields," in *2011 International Conference on Computer Vision* (IEEE, 2011) pp. 1020–1027.

[30] Gemma De las Cuevas and Toby S. Cubitt, "Simple universal models capture all classical spin physics," Science **351**, 1180—-1183 (2016).

[31] Samuel A. Ocko and Beni Yoshida, "Nonperturbative Gadget for Topological Quantum Codes," Physical Review Letters **107**, 250502 (2011).

[32] Yigit Subasi and Christopher Jarzynski, "Nonperturbative embedding for highly nonlocal Hamiltonians," Physical Review A **94**, 012342 (2016).

[33] Richard P. Feynman, "Quantum Mechanical Computers," Optics News **11**, 11 (1985).

[34] Daniel Nagaj, "Fast universal quantum computation with railroad-switch local Hamiltonians," Journal of Mathematical Physics **51**, 062201 (2010).

[35] Daniel Nagaj, "Universal two-body-Hamiltonian quantum computing," Physical Review A **85**, 032330 (2012).

[36] Qian-Heng Duan and Ping-Xing Chen, "Realization of universal adiabatic quantum computation with fewer physical resources," Physical Review A **84**, 042332 (2011).

[37] Roberto Oliveira and Barbara M. Terhal, "The complexity of quantum spin systems on a two-dimensional square lattice," Quantum Info. Comput. **8**, 900–924 (2008).

[38] Sergey Bravyi, David P. DiVincenzo, Daniel Loss, and Barbara M. Terhal, "Quantum Simulation of Many-Body Hamiltonians Using Perturbation Theory with Bounded-Strength Interactions," Physical Review Letters **101**, 070503 (2008).

[39] Yudong Cao, Ryan Babbush, Jacob Biamonte, and Sabre Kais, "Hamiltonian gadgets with reduced resource requirements," Physical Review A **91**, 012315 (2015).

[40] Stephen P. Jordan and Edward Farhi, "Perturbative gadgets at arbitrary orders," Physical Review A **77**, 062329 (2008).

[41] Courtney G Brell, Steven T Flammia, Stephen D Bartlett, and Andrew C Doherty, "Toric codes and quantum doubles from two-body Hamiltonians," New Journal of Physics **13**, 053039 (2011).

[42] Jacob D. Biamonte and Peter J. Love, "Realizable Hamiltonians for universal adiabatic quantum computers," Physical Review A **78**, 012352 (2008).

[43] Seth Lloyd and Barbara M Terhal, "Adiabatic and Hamiltonian computing on a 2D lattice with simple two-qubit interactions," New Journal of Physics **18**, 023042 (2016).

[44] Yudong Cao and Daniel Nagaj, *Quantum Information & Computation*, Vol. 15 (Rinton Press, 2015) pp. 1197–1222.

[45] Ryan Babbush, Bryan O'Gorman, and Alán Aspuru-Guzik, "Resource efficient gadgets for compiling adiabatic quantum optimization problems," Annalen der Physik **525**, 877–888 (2013).

[46] Courtney G Brell, Stephen D Bartlett, and Andrew C Doherty, "Perturbative 2-body parent Hamiltonians for projected entangled pair states," New Journal of Physics **16**, 123056 (2014).

[47] Wolfgang Lechner, Philipp Hauke, and Peter Zoller, "A quantum annealing architecture with all-to-all connectivity from local interactions," Science Advances **1** (2015), 10.1126/sciadv.1500838.

[48] Andrea Rocchetto, Simon C Benjamin, and Ying Li, "Stabilizers as a design tool for new forms of the Lechner-Hauke-Zoller annealer," Science Advances **2** (2016), 10.1126/sciadv.1601246.

[49] Jianbing Shen, Jianteng Peng, Xingping Dong, Ling Shao, and Fatih Porikli, "Higher Order Energies for Image Segmentation," IEEE Transactions on Image Processing **26**, 4911–4922 (2017).

[50] Endre Boros and Peter L. Hammer, "Pseudo-Boolean optimization," Discrete Applied Mathematics **123**, 155–225 (2002).

[51] Gernot Schaller and Ralf Schützhold, "The role of symmetries in adiabatic quantum algorithms," Quantum Information & Computation **10**, 109–140 (2010).

[52] Nanyang Xu, Jing Zhu, Dawei Lu, Xianyi Zhou, Xinhua Peng, and Jiangfeng Du, "Quantum Factorization of 143 on a Dipolar-Coupling Nuclear Magnetic Resonance System," Physical Review Letters **108**, 130501 (2012).

[53] William H. Cunningham, "Minimum cuts, modular functions, and matroid polyhedra," Networks **15**, 205–215 (1985).