# The Reverse Ising Problem

Isaac Martin and Andrew Moore

Summer 2023

## Contents

**1 Foundations** **1**
1.1 Circuits . . . 1
1.2 The Reverse Ising Problem . . . 2
1.3 Constraint Sets for Designing Ising Systems . . . 3
1.4 The Augmented Approach . . . 7
Randomly Sampling Ising Hamiltonians: choosing auxiliary functions which are Ising functions 9
Finding neutralizable Ising functions . . . 10
Reducing the auxiliary function search space using the mixed-integer augmented constraints . 10
1.5 Symmetries of Circuits and Systems . . . 10

**2 Boolean Circuits and Functions** **11**
2.1 Concepts from Boolean Algebra . . . 12
2.2 Ising systems which solve boolean circuits . . . 12

**3 Hierarchical Clustering** **12**

**4 Pseudo-boolean optimization and polynomial fitting** **13**

# 1 Foundations

## 1.1 Circuits

Throughout this document we set $\Sigma = \{-1, 1\}$ to be the set of possible spins ($\Sigma$ for **s**pin) taken by a single vertex in an Ising graph. For any set $X$, we denote by $\Sigma^X$ the set of all functions $\sigma : X \longrightarrow \Sigma$ and note, by setting $n = |X|$, that $\Sigma^X$ can equivalently be viewed as the set of $n$-tuples valued in $\{-1, 1\}$, $\Sigma^X \cong \{(\sigma_1, ..., \sigma_n) \in \Sigma^n\}$ or as the set of binary strings of length $n$ under the identification $\Sigma \leftrightarrow \{0, 1\}$. We first review a few elementary facts.

**Proposition 1.1.** Let $X$ be any set.

(a) If $\alpha : N \hookrightarrow X$ then we have a map $\mathrm{res}_{XN}\, \Sigma^X \longrightarrow \Sigma^N$ given $\mathrm{res}_{XN}(\sigma) = \sigma \circ \alpha$. We denote $\mathrm{res}_{XN}(\sigma) = \sigma|_N$.

(b) If $N, M \subseteq X$ are disjoint, then $\Sigma^N \times \Sigma^M \cong \Sigma^{N \cup M}$. em

(c) If $N, M \subseteq X$ are not necessarily disjoint, then $\Sigma^N \times_{N \cap M} \Sigma^M \cong \Sigma^{N \cup M}$.

***Proof.***

(a) Trivial; statement of fact.

(b) This is a special case of (c).

(c) The object $\Sigma^{N\cup M}$ fits into the following diagram

$$\begin{array}{ccc} \Sigma^{N\cup M} & \xrightarrow{\text{res}} & \Sigma^{N} \\ \downarrow{\scriptstyle \text{res}} & & \downarrow{\scriptstyle \text{res}} \\ \Sigma^{M} & \xrightarrow{\text{res}} & \Sigma^{N\cap M} \end{array}$$

where all maps are the appropriate restriction maps. In fact, $\Sigma^{N\cup M}$ together with restriction maps is the universal object of this diagram; suppose we have another object $P$ with maps $q_1 : P \longrightarrow \Sigma^M$ and $q_2 : P \longrightarrow \Sigma^N$ making this commute. Then $\text{res}(q_1(p)) = \text{res}(q_2(p))$ so the map $u : P \longrightarrow \Sigma^{N\cup M}$ given by

$$u(p)(x) = \begin{cases} q_1(p)(x) & x \in M \\ q_2(p)(x) & x \in N \end{cases}$$

is well defined on $M \cap N$. Hence $\Sigma^{N\cup M}$ is the pullback of $\Sigma^M$ and $\Sigma^N$ with restriction maps to $\Sigma^{N\cap M}$.

□

**Definition 1.2.** A **circuit** is a tuple $(N, M, f)$ where

- $N, M \subseteq X$ are arbitrary subsets of some universal set $X$ and are almost always chosen so that $N\cap M = \varnothing$. Both $N$ and $M$ should be finite and we call their elements **input** and **output** spins respectively.
- $f : \Sigma^N \longrightarrow \Sigma^M$ is an arbitrary function called the *logic* function.

We make special note of the scenarios in which $N$ and $M$ are not disjoint. Additionally, we define the following terminology.

- $A = X \setminus (N \cup M)$ is the **set of auxiliary spins** of $X$.
- $\Sigma^X$ is the **spin space**, $\Sigma^N$ is the **input space**, $\Sigma^M$ the **output space** and $\Sigma^N$ the **auxiliary space** of $X$. Elements of $\Sigma^X$ are called **spin states**, elements of $\Sigma^N$ **input states**, etc.
- $\mathcal{R}(f) = \{\sigma \in \Sigma^X \mid f(\sigma|_N) = \sigma|_M\}$ is the set of **correct** or **right** spin states.
- $\mathcal{W}(f) = \Sigma^X \setminus \mathcal{R}(f)$ is the set of **wrong** spin states.
- $\mathcal{L}(\sigma) = \mathcal{L}(\sigma|_N) = \{\sigma' \in \Sigma^X \mid \sigma'|_N = \sigma|_N\}$ is the **input level** of $\sigma$. It is useful to talk about the input level of both an input state $\sigma|_N$ and a full spin state $\sigma$, so we write $\mathcal{L}(\sigma)$ and $\mathcal{L}(\sigma|_N)$ to mean the same thing.

At this point it is worth remarking that this object $(N, M, f)$ in no way carries the data of an Ising system. An Ising system is a collection of spins with connections between them, together with a fixed quadratic Hamiltonian which predicts the dynamics of the system. What we have defined has no Hamiltonian, no notion of dynamics and no real semblance of a graph structure; it has only a function which specifies a desired output to each input and a lot of extra space ($X \setminus (N \cup M)$ with which to begin adding additional structure. An abstract circuit, as defined, is nothing other than a skeleton for an Ising system we wish to design.

## 1.2 The Reverse Ising Problem

**Definition 1.3.** An **Ising system** is a pair $(X, H)$, often referred to as simply $X$, where

- $X \subseteq \mathbb{N}$ is a set whose elements are called **spins**,
- $H \in \mathbb{R}[X]$ is a quadratic polynomial called the *Hamiltonian* of $X$.

The **state space** of $X$ is $\Sigma^X$. An Ising system $X$ in state $\sigma \in \Sigma^X$ has energy $H(\sigma)$ given by evaluating the Hamiltonian at $\sigma$.

An Ising system is inherently probabilistic. The probability that an Ising system $X$ is in state $\sigma \in \Sigma^X$ is given by the **configuration probability**

$$P_\beta = \frac{e^{-\beta H(\sigma)}}{Z_\beta}$$

where $\beta = (k_B T)^{-1} \geq 0$ is inverse temperature, $k_B$ is the Boltzmann constant and the normalization constant $Z_\beta$ is the partition function $Z_\beta = \sum_{\sigma \in \Sigma^X} e^{-\beta H(\sigma)}$. Notice that probability is maximized whenever the Hamiltonian is minimized, hence low energy states are more probable than high energy states.

With this in mind, for an Ising system with $(X, H)$ with $X$ finite and a subset $U \subseteq X$, we define the **minimizer with respect to** $U$ to be the function $m_U : \Sigma^U \longrightarrow \Sigma^X$ defined

$$f_U(\tau) = \underset{\sigma \in \Sigma^X,\ \sigma|_U = \tau}{\arg\min} H(\sigma).$$

It answers the question: "if the states of $U$ are held fixed, then what is the most likely state of $X$?" We likewise define the **minimizer logic** of $U$ to be the function $f_U : \Sigma^U \longrightarrow \Sigma^{U^c}$ given $f_U(\tau) = m_U(\tau)|_{U^c}$.

We would like to design Ising systems with the following features:

(1) A subset $N \subseteq X$ of spins whose state can be fixed

(2) A subset $M \subseteq X$ whose states vary freely with dynamics

(3) For a choice $\sigma_N \in \Sigma^N$, the most likely spin state in $\sigma_M \in \Sigma^M$ is $f(\sigma_N)$, where $f : \Sigma^N \longrightarrow \Sigma^M$ is some function.

Stated another way, given an abstract circuit $(N, M, f)$, we want to design Ising systems such that for every choice of input state $\sigma|_N$, the state $\sigma'$ which minimizes energy among all states matching $\sigma$ in input is a correct spin state. That is,

$$\underset{\sigma' \in \mathcal{L}(\sigma)}{\arg\min} H(\sigma') \in \mathcal{R}(f) \ \text{ for all } \ \sigma \in \Sigma^X.$$

**Definition 1.4.** Fix an abstract circuit $(N, M, f)$. We say that an Ising system $(X, H)$ **solves** $(N, M, f)$ (in that it solves the reverse Ising problem on $(N, M, f)$) if for each $\sigma \in \Sigma^X$

$$\tau = \underset{\sigma' \in \mathcal{L}(\sigma)}{\arg\min}\ H(\sigma') \implies \tau \in \mathcal{R}(f).$$

In other words, a choice of Hamiltonian $H$ for $X$ solves the circuit if the minimizer of the Hamiltonian among all states with matching input states has as its output component the correct output as specified by the circuit logic $f$.

An **Ising circuit** is an abstract circuit $(N, M, f)$ with $N, M \subset X$ together with an Ising system $(X, H)$ such that $(X, H)$ solves $(N, M, f)$.

## 1.3 Constraint Sets for Designing Ising Systems

Let us first consider the simplest case of a circuit $(N, M, f)$ with $X = N \cup M$ and $N$ and $M$ disjoint.

**Lemma 1.5.** The circuit $X$ is solvable with an Ising system if and only if there exists a Hamiltonian $H$ such that for all $s \in \Sigma^N$ and $t \neq f(s)$ $H(s, t) > H(s, f(s))$.

***Proof.*** Obvious, is essentially the definition of the Ising system solvability. □

It is not hard to find examples of circuits which are not solvable, for instance, the naive XOR circuit with $|X| = 3$ is not solvable with an Ising system.

For the remainder of this section, suppose that $X = N \cup M \cup A$ is a finite set and $N$, $M$, and $A$ are all disjoint. We call elements of this extra set $A$ **auxiliary spins**.

**Proposition 1.6.** Let $X \subseteq \mathbb{N}$ be infinite and $N$, $M$ be finite disjoint subsets. For any choice of $f$, the circuit $(N, M, f)$ is solvable with an Ising system. Since $|X| > |N \cup M|$ in this case, we sometimes say that $X$ is **solvable with auxiliary spins**.

Notice that this lemma says nothing about the *number* of auxiliary spins needed to solve a circuit; in general, it can be quite large.

***Proof.*** Take the hamming objective function ham : $\Sigma^{N \cup M} \longrightarrow \mathbb{R}$ defined to be the hamming distance from $\sigma$ to the (unique) correct spin state whose $N$ coordinates match those of $\sigma$:

$$\text{ham}(\sigma) = d(\sigma_M, f(\sigma_N)).$$

This has minimum value 0 obtained precisely at spins with correct output coordinates. It is also a pseudo-boolean function and hence can be written uniquely as a multilinear polynomial. Add auxiliary variables until the degree of this polynomial is 2 using, for instance, Rosenberg reduction. The obtained quadratic will be an Ising Hamiltonian in $|N \cup M \cup A|$ variables where $A$ is the set of auxiliary variables added during the reduction step. □

The task of finding an Ising system which solves some abstract circuit $(N, M, f)$ can thus be thought of as finding a sufficiently large cardinality for $A$ and then solving a mixed non-linear optimization problem to find valid a valid quadratic Hamiltonian together with auxiliary states which occur at the desired minimizers of input levels. The following lemma demonstrates one way one might attempt to solve an abstract circuit with auxiliaries.

**Lemma 1.7.** The circuit $(N, M, f)$ is solvable with an Ising system if and only if there exists a function $g : \Sigma^N \longrightarrow \Sigma^A$ and a Hamiltonian $H$ such that for all $\sigma \in \Sigma^N$, $\eta \in \Sigma^A$ and $f(\sigma) \neq \omega \in \Sigma^M$, $H(\sigma, \omega, \eta) > H(\sigma, f(\sigma), g(\sigma))$.

The image of the function $g$ is called the **auxiliary array** of $X$. If we fix a specific spin $a \in A$, then the set $\{g(\sigma)_a\}_{\sigma \in \Sigma^N}$ is called the *auxiliary vector of* $a$.

***Proof.*** Again, obvious. □

A solution to a circuit employing Lemma 1.7 consists of two parts: obtaining a feasible auxiliary array $g$ and identifying choices of Hamiltonians such that the constraints in Lemma 1.7 are all satisfied. Although it is always possible, choosing a feasible $g$ is quite difficult, and doing so in such a way that minimizes the cardinality of $A$ is even harder. If a feasible $g$ is known, solving for $H$ is a linear programming problem; however, because the number of constraints in Lemma 1.7 grows exponentially in $N$, $M$ and $A$, for problems on the order of $|X| \approx 100$ there is no computer on earth that can actually solve the LP-problem as stated.

**Definition 1.8** (Constraint Sets for Reverse Ising)**.** Let $(N, M, f)$ be an abstract circuit with $X = N \cup M \cup A$ and $N, M, A$ all disjoint. We say that an Ising system $(X, H)$ together with a choice $g : \Sigma^N \longrightarrow \Sigma^A$ of auxiliary array satisfies

- **Weak Constraints** if for all $\sigma_N \in \Sigma^N$ and $\sigma_M \neq f(\sigma_N)$

  $$H(\sigma_N, \sigma_M, \sigma_A) > H(\sigma_N, f(\sigma_N), g(\sigma_N)).$$

  There are $2^{M+A} - 2^A$ constraints per input level, $2^N(2^{M+A} - 2^A)$ total.

- **Full Constraints** if for all $\sigma_N \in \Sigma^N$ and $\sigma_M \neq f(\sigma_N)$ *or* $\sigma_A \neq g(\sigma_N)$

  $$H(\sigma_N, \sigma_M, \sigma_A) > H(\sigma_N, f(\sigma_N), g(\sigma_N)).$$

  There are $2^{M+A} - 1$ constraints per input level, $2^N(2^{M+A} - 1)$ total.

- **F-Augmented Constraints** if for some function $F : \Sigma^N \times \Sigma^M \longrightarrow \Sigma^A$

$$H(\sigma_N, \sigma_M, F(\sigma_N, \sigma_M)) > H(\sigma_N, f(\sigma_N), F(\sigma_N, f(\sigma_N))).$$

  There are $2^M - 1$ constraints per input level or $2^N(2^M - 1)$ constraints total.

**Lemma 1.9.** Let $(N, M, f)$ be an abstract circuit with $X = N \cup M \cup A$ and all $N, M$ and $A$ disjoint. This circuit is solved by an Ising system $(X, H)$ if and only if there is an auxiliary array $g : \Sigma^N \longrightarrow \Sigma^A$ such that $H$ satisfies full constraints.

***Proof.*** The reverse implication is clear, as every constraint in Lemma 1.7 is a constraint in Lemma 1.9.

For the forward implication, suppose $(N, M, f)$ is solved by an Ising system $(X, H)$. Defining $g$ to be the auxiliary component of the minimizer with respect to $N$, $g(\sigma_N) = m_N(\sigma_N)|_A$, does the trick. □

The next lemma addresses the use of the $F$-augmented constraints, and is the most useful result in this section. It is the primary result underpinning the authors' approach to the reverse Ising problem.

**Lemma 1.10.** Let $(N, M, f)$ again be an abstract circuit. There exists an Ising system which solves this circuit if and only if there is some function $F : \Sigma^N \times \Sigma^M \longrightarrow \Sigma^A$ such that both

(a) the new circuit $(N \cup M, A, F)$ is solvable by an Ising system with Hamiltonian $R$ with the following additional property:

$$R(\sigma_N, \sigma_M, F(\sigma_N, \sigma_M)) \geq R(\sigma_N, f(\sigma_N), F(\sigma_N, f(\sigma_N))) \qquad (\dagger)$$

for all $\sigma_N$ and $\sigma_M$. We call this the **weak neutralizability condition.** (If the inequality is instead an equality, we call this the **strong neutralizability condition.** Likewise, if such an Ising system (X,R) exists, we correspondingly say that $F$ is weakly neutralizable or strongly neutralizable.) We call the system $(X, R)$ the **auxiliary system** and the circuit $(N \cup M, A, F)$ the **auxiliary circuit**.

(b) there is an Ising system $(X, H)$ which satisfies $F$-augmented constraints. We call $(X, H)$ the **base system** and the circuit $(N, M, f)$ the **base circuit**.

**Remark 1.11.** Before we prove this lemma, a remark is in order. It may seem at first glance that we have made the situation worse – Lemma 1.10 splits the task of finding a single Ising system which solves $(N, M, f)$ into the task of finding two Ising systems satisfying different constraint sets. However, the reader may have noticed that the $F$-augmented constraints are far weaker than the other two constraints sets given in Definition 1.8. This Lemma explains why they are at all useful; the auxiliary system $(X, R)$ allows the weak constraints to be replaced by the far simpler augmented constraints. Though the system $(X, R)$ needs to solve the circuit $(N \cup M, A, F)$ *and* be at least weakly neutralizable, the complexity of this circuit is directly controlled by $F$. These two observations together mean the reverse Ising problem reduces to building up a theory for choosing $F$ given a base circuit $(N, M, f)$.

***Proof of Lemma 1.10.*** Throughout this proof let $\sigma, \omega$ and $\eta$ denote elements in $\Sigma^N$, $\Sigma^M$ and $\Sigma^A$ respectively. Suppose first that the circuit $(N, M, f)$ is solvable by an Ising system $(X, H)$. Define $F : \Sigma^N \times \Sigma^M \longrightarrow \Sigma^A$ to be the auxiliary component of the minimizer with respect to $N \cup M$:

$$F(\sigma, \omega) = m_{N \cup M}(\sigma, \omega)|_A := \underset{\eta \in \Sigma^A}{\arg\min}\, H(\sigma, \omega, \eta).$$

By definition of $F$, the circuit $(N \cup M, A, F)$ is solvable by the Ising system $(X, H)$. Furthermore, since $H$ satisfies the weak constraints, for some value of $\eta' \in \Sigma^A$ we have that

$$H(\sigma, \omega, \eta) > H(\sigma, f(\sigma), \eta') \geq H(\sigma, f(\sigma), F(\sigma, f(\sigma)))$$

for all $\eta \neq \eta'$ where the second from the definition of $F$. Thus, in particular, it satisfies the weak neutralizability condition.

Since $(X, H)$ solves the circuit $(N, M, f)$, by Lemma 1.7 we have that

$$H(\sigma, \omega, \eta) > H(\sigma, f(\sigma), F(\sigma, f(\sigma)))$$

for all $\omega \neq f(\sigma)$, and all $\eta$ so in particular,

$$H(\sigma, \omega, F(\sigma, \omega)) > H(\sigma, f(\sigma), F(\sigma, f(\sigma)))$$

Thus $(X, H)$ satisfies the $F$-augmented constraints.

Now suppose that $F$ is an arbitrary function such that $(N \cup M, A, F)$ is an abstract circuit solvable by an Ising system $(X, R)$ whose Hamiltonian $R$ satisfies (†) and that $(X, S)$ is an Ising system with Hamiltonian $S$ which satisfies the $F$-augmented constraints. Consider the family of Ising Hamiltonians $H_\lambda = S + \lambda R$ parameterized by $\lambda$. We show that for sufficiently large $\lambda$, $H_\lambda$, $(X, H_\lambda)$ together with auxiliary array $g(\sigma) = F(\sigma, f(\sigma))$ satisfies the weak constraints and hence solves the circuit $(N, M, f)$.

Fix $\sigma$ and $\omega \neq f(\sigma)$, and consider first the case that $\eta = F(\sigma, \omega)$. Then

$$\begin{aligned}
&H_\lambda(\sigma, \omega, \eta) - H_\lambda(\sigma, f(\sigma), g(\sigma)) > 0 \\
&\quad \iff S(\sigma, \omega, \eta) - S(\sigma, f(\sigma), g(\sigma)) + \lambda(R(\sigma, \omega, \eta) - R(\sigma, f(\sigma), g(\sigma))) > 0 \\
&\quad \iff S(\sigma, \omega, F(\sigma, \omega)) - S(\sigma, f(\sigma), F(\sigma, f(\sigma))) + \lambda R(\sigma, \omega, F(\sigma, \omega)) - \lambda R(\sigma, f(\sigma), F(\sigma, f(\sigma))) > 0 \\
&\quad \iff S(\sigma, \omega, F(\sigma, \omega)) - S(\sigma, f(\sigma), F(\sigma, f(\sigma))) > 0.
\end{aligned}$$

The final condition in the above chain of bi-conditionals holds irrespective of the value of $\lambda$ since $S$ satisfies the $F$-augmented constraints. Now suppose that $\eta \neq F(\sigma, \omega)$. Set

$$\alpha = \min_{\substack{\omega \in \Sigma^M \\ \omega \neq f(\sigma)}} R(\sigma, \omega, \eta) - R(\sigma, f(\sigma), F(\sigma, f(\sigma))),$$

noting that by (†), the assumption that $(X, R)$ solves $(N \cup M, A, F)$ and because $\eta \neq F(\sigma, \omega)$ we have

$$R(\sigma, \omega, \eta) > R(\sigma, \omega, F(\sigma, \omega)) \geq R(\sigma, f(\sigma), g(\sigma))$$

which in turn implies that $\alpha > 0$. Additionally set

$$\beta = \max_{\sigma \in \Sigma^X} S(\sigma, f(\sigma), F(\sigma, f(\sigma))) - S(\sigma, \omega, \eta).$$

Then

$$\begin{aligned}
&H_\lambda(\sigma, \omega, \eta) - H_\lambda(\sigma, f(\sigma), F(\sigma, f(\sigma))) > 0 \\
&\qquad \iff \\
&S(\sigma, \omega, \eta) \; - \; S(\sigma, f(\sigma), F(\sigma, f(\sigma))) \\
&\qquad + \; \lambda R(\sigma, \omega, \eta) \; - \; \lambda R(\sigma, f(\sigma), F(\sigma, f(\sigma))) \; > \; 0 \\
&\qquad \iff \\
&\lambda \; > \; \frac{S(\sigma, f(\sigma), F(\sigma, f(\sigma))) - S(\sigma, \omega, \eta)}{R(\sigma, \omega, \eta) - R(\sigma, f(\sigma), F(\sigma, f(\sigma)))}.
\end{aligned}$$

Choosing $\lambda > \beta/\alpha$ ensures this is satisfied for all $\sigma \in \Sigma^X$. □

The following theorem summarizes our results up to this point.

**Theorem 1.12.** *Let $(N, M, f)$ be an abstract circuit. Then the following are equivalent.*

*(i) There exists an Ising system $(X, H)$ which solves $(N, M, f)$.*

*(ii) There exists an Ising system $(X, H)$ which satisfies the weak constraints of $f$.*

*(iii) There exists an Ising system $(X, H)$ which satisfies the full constraints of $f$.*

*(iv) There exists a function $F : \Sigma^N \times \Sigma^M \longrightarrow \Sigma^A$, an Ising system $(X, S)$ satisfying the $F$-augmented constraints and an Ising system $(X, R)$ which solves the circuit $(N \cup M, A, F)$ and has the weak neutralizability property.*

***Proof.*** Lemma 1.7 proves *(i)* $\Leftrightarrow$ *(ii)*, Lemma 1.9 proves *(i)* $\Leftrightarrow$ *(iii)* and Lemma 1.10 proves *(ii)* $\Leftrightarrow$ *(iv)*. □

**Remark 1.13.** Of these four equivalences, *(iv)* is perhaps the most opaque. One might wonder what utility it actually buys, especially since the assumption that $R$ has the weak neutralizability condition and solves the circuit $(N \cup M, A, F)$ is awfully close to simply demanding that $R$ satisfies the weak constraints of $f$. The Hamiltonian $S$ appears only to be used to "fix" the places where $R$ fails to satisfy the strong inequality version of the weak neutralizability condition. All of this begs the question: how is this ever useful?

1. **Augmented constraints do not scale exponentially in $A$.** Both the weak and full constraints grow exponentially in $A$, but the constraint matrices of the $F$-augmented constraints always have a fixed number of rows and grow quadratically columnwise due to the addition of quadratic coefficients. Hence this problem scales on the order of $\mathcal{O}(|N \cup M \cup A|^2)$.

2. **Good choices of $F$ are known.** If $F$ depends only on a few spins, then cooking up reasonable Hamiltonians $R$ becomes far easier. See examples 1.14 and 1.15.

3. **Improvements to $F$ can be made iteratively.** If a function $F_1 : \Sigma^N \times \Sigma^M \longrightarrow \Sigma^{A_1}$ is weakly neutralizable and $F_2 : \Sigma^N \times \Sigma^{M \cup A_1} \longrightarrow \Sigma^{A_2}$ is weakly neutralizable, then the composite function

$$F(\sigma, \omega) = (F_1(\sigma, \omega), F_2(\sigma, \omega, F_1(\sigma, \omega))) \in \Sigma^{A_1 \cup A_2}$$

is also strongly neutralizable. This means the threshold constraint approach to the reverse Ising problem lends itself well to iterative methods. In particular, due to the observation that the solvability of a boolean circuit by an Ising system is equivalent to linear separability, we can attach auxiliary spins one at a time (see Section 2).

**Example 1.14.** Suppose $F : \Sigma^N \times \Sigma^M \longrightarrow \Sigma^A$ is constant in the $\Sigma^M$ component. Then any Ising system $(X, R)$ which solves the circuit $(N \cup M, A, R)$ is weakly neutralizable.

**Example 1.15.** Let $(\{a, b\}, \{c\}, AND)$ be the 1-bit AND circuit. There exists an Ising system $(\{a, b, c\}, R)$ which solves the circuit and is strongly neutralizable.

## 1.4 The Augmented Approach

This section is dedicated to discussing general algorithms which utilize the $F$-augmented constraints and Lemma 1.10. As is quickly becoming standard, fix a decomposition $X = N \cup M \cup A$ with $N$, $M$ and $A$ pairwise disjoint and an abstract circuit $(N, M, f)$, the **base circuit** of our augmented system. Our task is to choose an auxiliary function $F$ such that

(a) there exists an Ising system which satisfies the $F$-augmented constraints of $(N, M, f)$

(b) $F$ is an Ising function, i.e. $(N \cup M, A, F)$ is a solvable circuit

(c) $F$ satisfies weak neutralizability.

All of these properties can be checked using linear programming for a fixed choice of $F$, but they depend on different portions of the Hamiltonian. To illustrate this, consider the following decomposition of an Ising Hamiltonian on $X$.

Denote by $x$ the indeterminants with indices in $N$, $y$ the indeterminants with indices in $M$ and $z$ the indeterminants with indices in $A$. An arbitrary Ising Hamiltonian on $X$ without a constant term can then be written

$$H(x,y,z) = h_x(x) + h_y(x) + h_z(z) + p_x(x) + p_y(y) + p_z(z) + q_{xy}(x,y) + q_{xz}(x,z) + q_{yz}(y,z) \quad (1)$$

where the $h$ terms denote linear polynomials, the $p$ terms denote self-interaction quadratic forms and the $q$ terms denote cross-type interaction quadratic terms (e.g. input/output interactions, etc.).

Imagine we've chosen an auxiliary function $F$. Condition (b) on $F$ demands that the weak constraints $(N \cup M, A, F)$ be solvable without auxiliary variables, which is the following familiar constraint set:

$$H(x,y,F(x,y)) < H(x,y,z) \text{ for all } F(x,y) \neq z \in \Sigma^A.$$

Subtracting $H(x,y,F(x,y))$ from both sides and writing our Hamiltonian in the form of Equation 1 gives us

$$q_{yz}(y, z - F(x,y)) + q_{xz}(x, z - F(x,y)) + p_z(z - F(x,y)) + h_z(z - F(x,y)) > 0$$

for all $F(x,y) \neq x \in \Sigma^A$, meaning property (b) is entirely determined by $q_x z$, $q_{yz}$, $p_z$ and $h_z$; the terms which depend on $z$. A similar argument demonstrates that neutralizability conditions, both strong and weak forms, depend on all terms involving $y$ and $z$.

We summarize this discussion with a definition and a remark.

**Definition 1.16.** Let $F : \Sigma^N \times \Sigma^M \longrightarrow \Sigma^A$ be an auxiliary circuit. We saw that $F$ has the...

(a) ...**augmented property** if there exists an Ising Hamiltonian $R$ on $X$ which solves the auxiliary circuit $(N \cup M, A, F)$. In this case we say that $F$ is an *augmented* auxiliary function.

(b) ...**circuit property** if there exists an Ising Hamiltonian $R$ on $X$ which solves the auxiliary circuit $(N \cup M, A, F)$. In this case we say that $F$ is an *Ising* auxiliary function.

(c) ...**weak** (resp. **strong**) **neutralizability** property if there exists a Hamiltonian $R$ which satisfies the weak (resp. strong) neutralizability conditions. In this case we say that $F$ is either a *weakly neutralizable* auxiliary function with respect to the circuit $(N, M, f)$ or a *strongly neutralizable* auxiliary function. Strong neutralizability does not depend on the circuit $(N, M, f)$ and hence need not make reference to it.

**Remark 1.17.** Continue letting $x$, $y$ and $z$ denote indeterminants indexed in $N$, $M$ and $A$ respectively. An auxiliary function $F : \Sigma^N \times \Sigma^M \longrightarrow \Sigma^A$ is...

(a) ...augmented if there exists a Hamiltonian $S$ such that

$$S(x, f(x), F(x, f(x))) < S(x,y,F(x,y)) \text{ for all } x \in \Sigma^N \text{ and } y \neq f(x)$$

(b) ...an Ising function if there exists a Hamiltonian $R$ such that

$$R(x,y,F(x,y)) < R(x,y,z) \text{ for all } (x,y) \in \Sigma^N \times \Sigma^M \text{ and } z \neq F(x,y),$$

which is the constrained linear programming problem from Lemma 1.5 depending only on $h_z$, $p_z$, $q_{xz}$ and $q_{yz}$.

(c) ...weakly neutralizable function with respect to $N, M, f)$ if there exists an Ising Hamiltonian $R$ on $X$ such that

$$R(x, f(x), F(x, f(x))) \leq R(x,y,F(x,y)) \text{ for all } x \in \Sigma^N \text{ and } f(x) \neq y \in \Sigma^M.$$

Both cases depend on $h_y, h_z, p_y, p_z, q_{xy}, q_{xz}$ and $q_{yz}$. The weak case is a constrained linear programming problem while the second is a system of linear equations.

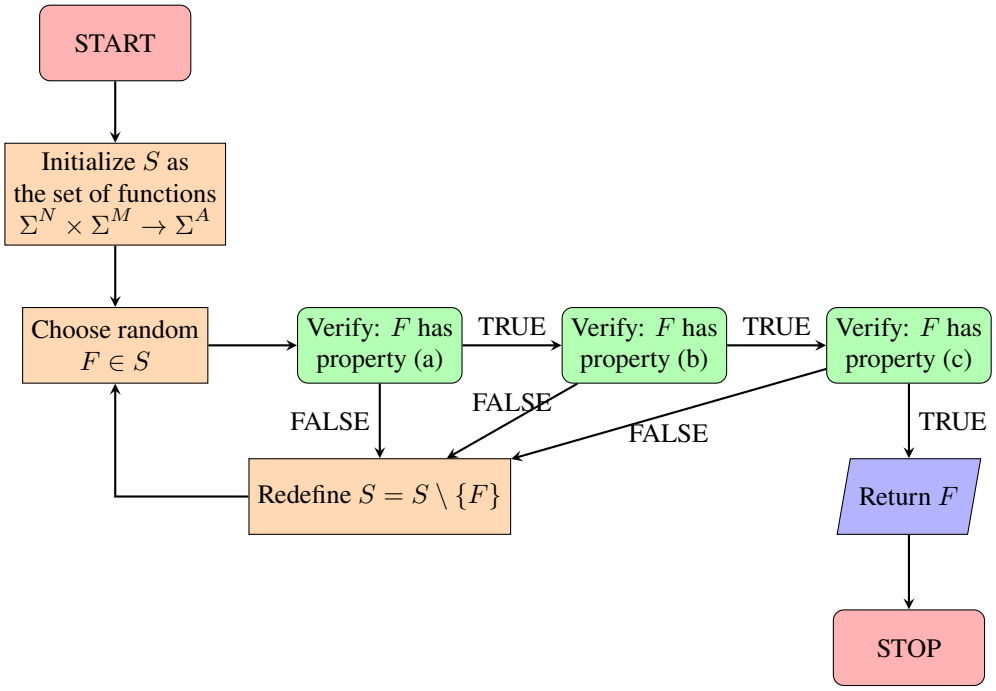


Figure 1: The most naïve version of AugSolve, an algorithm which solves a circuit $(N, M, f)$ using the augmented constraints and Lemma 1.10

Remark 1.17 should be interpreted to mean that, given a choice of $F$, we know how to check whether or not it satisfies the conditions requisite for solving the circuit $(N, M, f)$ using the augmented approach. If auxiliary functions satisfying all three properties were common, then we could simply search randomly through all the possibilities until we found a neutralizable augmented Ising function – reverse Ising solved! Such an algorithm would look like this:

However, auxiliary functions which are simultaneously augmented, Ising and neutralizable are exceedingly rare and the space of all auxiliary functions is absurdly large. We therefore need smarter ways of searching for auxiliary functions $F$ if this is to work. Here are some general ideas for improvement:

**Idea 1.18** (Improvements to AugSolve)**.**

(i) Instead of choosing $F$ randomly, choose auxiliary functions which are guaranteed to satisfy one of the properties (a), (b) and (c). If possible, this would both drastically shrink the search space $S$ and reduce computation time by removing one of the property verifications. Several versions of this improvement are discussed in the remainder of this section.

(ii) Some choices of $F$ are redundant; for example, $F$ and its "inverse" $-F$ either both satisfy properties (a) - (c) or both fail to satisfy the same property. Hence, when $F$ is removed from $S$ once might as well remove $-F$ too.

**Randomly Sampling Ising Hamiltonians: choosing auxiliary functions which are Ising functions**

Recall that the values of $F$ don't matter – we care about them only insofar as they allow us to solve the circuit in question. Given this, we can easily find auxiliary functions which are also Ising functions – simply choose a Hamiltonian $R$ on $X$ and define

$$F(x, y) = \arg\min_{z \in \Sigma^A} H(x, y, z).$$

In the case that $|A| = 1$, discussed further in Section 2, this is equivalent to choosing a random hyperplane in $\mathbb{R}^{N \cup M}$ and using it to partition the cube $\Sigma^{N \cup M}$ into two subsets. This produces the following variation of AugSolve:

$$\left\langle \text{Algo with [Random Choice] step swapped for } \left[\text{Sample Random Ising Hamiltonian from } \mathbb{R}^P\right] \right\rangle$$

This easily produces auxiliary functions which satisfy property (b), but there is still lots of redundancy. The Hamiltonian $R$ is entirely parameterized by its coefficients, and small perturbations in these coefficients will produce distinct Hamiltonians without changing the auxiliary function $F$. If it were possible to partition the parameter space $\mathbb{R}^P$ of $R$ into distinct regions depending on the auxiliary function $F$ induced, then this method on its own might prove quite powerful. As of yet, we do not understand this partition.

**Finding neutralizable Ising functions**

Given a Hamiltonian $R$ on $X$ it is theoretically (if not computationally) straightforward to check whether the auxiliary function $F$ it induces is strongly neutralizable; check whether the ground states associated to a single $x$ are all equal. This requires solving a boolean optimization problem, however; famously difficult.

Instead, consider if for small cardinalities $|A|$, each time we generate a Hamiltonian $R$ we cache the result of the neutralizability check. Over time, we build up a library of auxiliary functions which satisfy both the Ising and neutralizable properties. These can then be "glued"together to form larger auxiliary functions which are still both neutralizable and Ising.

This method has proven successful even for quite limited libraries of neutralizable Ising functions. An example of this is the "AND auxiliary" method. An additional drawback, however, is that the process of desigining a large auxiliary function stitched together from smaller auxiliary functions again involves iterating over a large search space. Designing auxiliary functions with fewer auxiliaries requires a larger library of these smaller functions, and consequently, creates a larger search space.

Another approach for designing neutralizable Ising functions might be to build a pseudo boolean function with the desired properties and reduce it until it is quadratic. This will likely lead to an explosion in the number of auxiliaries, however.

**Reducing the auxiliary function search space using the mixed-integer augmented constraints**

Searching directly for auxiliary functions $F$ such that the $F$-augmented constraints of $N, M, f)$ are satisfiable is a special type of auxiliary search. However, as the number of auxiliaries grows, no columns are added, so perhaps finding "augmented-feasible" auxiliaries is easier than finding typical feasible auxiliary arrays. Even if the full search is still prohibitively expensive, perhaps deductions can be made which reduce the space of possible auxiliary arrays. This could be done by performing Fourier-Motskin on a reduced augmented constraint set for example.

## 1.5 Symmetries of Circuits and Systems

<UNDER CONSTRUCTION>

Let $\mathrm{Circ}(N, M)$ denote the set of all circuits $(N, M, f)$ and $\mathrm{Ising}(X)$ the set of all Ising systems $(X, H)$. We then define the **Z-space** or **Zing-space** of $N, M \subset X$ by

$$\mathfrak{Z}_X(N, M) = \mathrm{Circ}(N, M) \times \mathrm{Ising}(X).$$

If $X$ is understood, we often suppress it in the notation and simply write $\mathfrak{Z}(N, M)$. The $\mathfrak{Z}$ stands for "Zing" as in the pronunciation of "Ising" as "I-zing". It also stands for "Zombiecupcake".

In this section we develop notions of maps and turn $\mathfrak{Z}(N, M)$ into a category. We are particularly interested in the *symmetries* or *automorphisms* in $\mathfrak{Z}(N, M)$.

**Definition 1.19.** An **Ising circuit morphism** or simply an **Ising map** is a pair of maps $(\alpha_1, \alpha_2)$ with $\alpha_1 : \mathrm{Circ}(N, M) \longrightarrow \mathrm{Circ}(N, M)$ and $\alpha_2 : \mathrm{Ising}(X) \longrightarrow \mathrm{Ising}(X)$ such that

$$(X, H) \text{ solves } (N, M, f) \Leftrightarrow \alpha_2(X, H) \text{ solves } \alpha_1(N, M, f).$$

A pair of maps like this is precisely the same as a map on the corresponding product, so we typically denote both maps by $\alpha$ and instead think of $\alpha$ as operating on $\mathfrak{Z}(N, M)$:

$$\alpha : \mathfrak{Z}(N, M) \longrightarrow \mathfrak{Z}(N, M).$$

Morphisms of Ising circuits is a function on $\mathfrak{Z}(N, M)$ which *preserves the relation of solvability* between circuits and Ising systems.

**Example 1.20.** Let $\sigma \in \Sigma^X$ be a spin state of $X$. By viewing $\Sigma^X = (\mathbb{Z}^\times)^{|X|}$, multiplication by $\sigma$ induces a group action of $\Sigma^X$ on itself. We then also have an action functions $f : \Sigma^N \longrightarrow \Sigma^M$ given by

$$(\sigma f)(s) = \sigma_M \cdot f(\sigma_N \cdot s).$$

This defines a map $\alpha_\sigma : \mathrm{Circ}(N, M) \longrightarrow \mathrm{Circ}(N, M)$. Correspondingly, if $H$ is an Ising Hamiltonian whose linear terms are denoted $h_i$ and quadratic terms $J_{ij}$ for $i, j \in X$, then by defining $\sigma \cdot H$ to be the Hamiltonian with linear and quadratic terms $h'_i$ and $J'_{ij}$ respectively given

$$J'_{ij} = \sigma_i \sigma_j J_{ij}, \qquad h'_i = \sigma_i h_i$$

we obtain a map $\alpha_\sigma : \mathrm{Ising}(X) \longrightarrow \mathrm{Ising}(X)$. For any spin $s \in \Sigma^X$ we have

$$H(\sigma \cdot s) = \sum_{i \in X} \sigma_i s_i h_i \; + \; \sum_{i<j} \sigma_i s_i \sigma_j s_j J_{ij} = \sum_{i \in X} (\sigma_i h_i) s_i \; + \; \sum_{i<j} (\sigma_i \sigma_j J_{ij}) s_i s_j = (\sigma \cdot H)(s),$$

hence $(s_N, s_M, s_A)$ minimizes $\sigma \cdot H$ on the input level $\mathcal{L}(s_N)$ if and only if $(\sigma_N s_N, \sigma_M s_M, \sigma_A s_A)$ minimizes $H$ on the input level $\mathcal{L}(\sigma_N \cdot s_N)$. This implies that

$$(X, H) \;\text{ solves }\; (N, M, f) \;\Leftrightarrow\; \alpha_\sigma(X, H) \;\text{ solves }\; \alpha_\sigma(N, M, f).$$

The actions of $\Sigma^X$ on $\mathrm{Circ}(N, M)$ and $\mathrm{Ising}(X)$ are both referred to as **spin actions**.

Notice that if $\alpha$ is an Ising symmetry and $\alpha'$ is another Ising symmetry, then $\alpha \circ \alpha'$ is again an Ising symmetry. Since the identity map $\mathrm{id}_{\mathfrak{Z}(N,M)}$ is an Ising morphism, we see that $\mathfrak{Z}(N, M)$ forms a category. Notice also that if $\alpha$ is a bijective Ising symmetry then its inverse $\alpha^{-1}$ is as well, so $\mathrm{Aut}(\mathfrak{Z}(N, M))$ is a group under function composition.

Consider the implications of this setup to the reverse Ising problem. If the group $\mathrm{Aut}(\mathfrak{Z}(N, M))$ were well understood, then every solution $(X, H)$ to $(N, M, f)$ would yield a family of Ising circuits in $\mathfrak{Z}(N, M)$ given by the orbit of $((X, H), (N, M, f))$ under $\mathrm{Aut}(\mathfrak{Z}(N, M))$. This is particularly useful when one isn't interested in solving a specific circuit but rather in understanding the space $\mathfrak{Z}(N, M)$ as a whole for fixed $N$ and $M$, for instance, when designing auxiliary circuits for use with $F$-augmented constraints. The rest of this section is therefore dedicated to developing an understanding of $\mathfrak{Z}(N, M)$.

Let us first imagine enumerating all possible elements of $\mathrm{Aut}(\mathfrak{Z}(N, M))$. One approach might be to write down all bijective maps on $\mathrm{Circ}(N, M)$ and then check which preserve the solvability relation. As a set, $\mathrm{Circ}(N, M)$ is the same as the collection of all functions $f : \Sigma^N \longrightarrow \Sigma^M$, hence it has size $2^{|N| \cdot 2^{|M|}}$. There are then $\left(2^{|N| \cdot 2^{|M|}}\right)!$ many bijective functions on $\mathrm{Circ}(N, M)$ – a bit too many to enumerate.

# 2 Boolean Circuits and Functions

<Under construction>

A **boolean function** is a map $f : \Sigma^N \longrightarrow \Sigma$. A **boolean circuit** is an abstract circuit $(N, M, f)$ where $|M| = 1$, and hence where $f$ is a boolean function.

**Remark 2.1** (Disambiguation of 0,1 and -1,1 $\Sigma$ representations)**.** Boolean algebra appears in many disparate fields of math – game theory, decision theory, coding theory, artificial intelligence and distributed algorithms to name a few – and hence conventions can vary quite a bit. In particular, the set $\Sigma$ assumes various roles and alternates between $\{-1, 1\}$ and $\{0, 1\}$ representations. Results in section 2 do not depend on whether $\Sigma = \{0, 1\}$ or $\Sigma = \{-1, 1\}$, and we will take care to note which definition is currently in use when ambiguity arises. Here is a context-dependent key for swapping between the $\{0, 1\}$ and $\{-1, 1\}$ representations.

- When $\Sigma$ is regarded as a 2-element set with no additional structure, any bijection between $\{0, 1\}$ and $\{-1, 1\}$ will suffice.
- When $\Sigma$ is regarded as a subset of $\mathbb{R}$, the change of variables $x \mapsto 2x - 1$ swaps the roles of $0$ and $-1$. This situation arises when considering pseudo-boolean functions $f : \Sigma^N \longrightarrow \mathbb{R}$, all of which have unique polynomial representations and hence depend on whether $-1$ or $0$ in use (consider evaluating an Ising Hamiltonian on strings of $0, 1$, for instance). It also comes up when one treats a boolean function $f : \Sigma^N \longrightarrow \Sigma$ as a labeling of the corners of the hypercube $\Sigma^N$, in which case the $\{-1, 1\}$ convention is more convenient as it makes the SVM formulas slightly more concise.
- When $\Sigma$ is regarded as the unique group $G$ of order 2, there is only identification between $\{0, 1\}$ and $\{-1, 1\}$ which makes sense: $0 \mapsto 1$ and $1 \mapsto -1$. This is because, $0$ is the identity element when $G$ is regarded as the additive group $\mathbb{Z}/2\mathbb{Z}$, but $1$ is the identity when $G$ is regarded as the multiplicative group $\mathbb{Z}^\times$. Annoyingly, this is distinct from the typical identification, and hence special care must be taken to avoid confusion in situations where the group structure of $\Sigma$ is relevant. The change of variables formula given above can be modified to $x \mapsto 1 - 2x$ if one wishes it to match the group isomorphism.

## 2.1 Concepts from Boolean Algebra

<Under construction>

**Definition 2.2.** Let $f : \Sigma^N \longrightarrow \Sigma$ be a boolean function. The **dual** of $f$ is a boolean function $f^d$ of the same dimension defined

$$f^d(x) = \overline{f(\overline{x})}.$$

The **self-dual** of $f$ is a boolean function $f^{sd}$ of dimension $N + 1$ defined

$$f^{sd}(x_0, x_1, ..., x_n) = \overline{x}_0 \cdot f^d(x_1, ..., x_n) \ + \ x_0 \cdot f(x_1, ..., x_n).$$

The additional bit $x_0$ "turns $f^d$ and $f$ on and off".

## 2.2 Ising systems which solve boolean circuits

**Proposition 2.3.** Let $(N, y, f)$ be an abstract circuit with one output – i.e. let $f$ be a boolean function. There exists an Ising system $(X, H)$ with $X = N \cup y$ if and only if the set of true vectors $T = f^{-1}(1)$ is linearly separable from the set of false vectors $F = X \setminus F = f^{-1}(-1)$.

***Proof.*** View the set $\Sigma^N$ as a set of tuples consisting of $+1/-1$ in each entry with every data point labeled $\square$

# 3 Hierarchical Clustering

A choice of "auxililary array" can be thought of as a function $a : \Sigma^N \longrightarrow \Sigma^A$ which assigns an auxiliary state to each input state. The collection of germs of this function partitions $\Sigma^N$ into subsets of inputs which share the same auxiliary state. To be clear, we're talking about the partition

$$\{a^{-1}(\alpha)\}_{\alpha \in \Sigma^A}$$

of input spin space. If the choice of auxiliary array $a$ makes the Ising circuit feasible, then we say $a$ solves the Ising circuit.

This simple observation motivates a simple question: *can a partition of $\Sigma^N$ be found such that it matches the partition produced by some feasible auxiliary array using only the logic of the Ising circuit?* Producing such a partition is a clustering problem on input spinspace.

# 4 Pseudo-boolean optimization and polynomial fitting

A pseudo boolean function (PBF) is any function $f : \Sigma^N \longrightarrow \mathbb{R}$, typically defined so that $\Sigma = \{0, 1\}$. It is a well known fact that any such PBF can be uniquely represented by a multilinear polynomial in $n$ variables [pseudo-boolean optimization Boros, Hammer]; that is, a polynomial

$$g(x_1, ..., x_n) = \sum_{S \subset [n]} a_S \prod_{j \in S} x_j$$

with $a_S \in \mathbb{R}$ which equals $f$ pointwise on $\{0, 1\}^n$. To be clear, here $S$ iterates over all subsets of $[n] = \{1, ..., n\}$.

It is another well-known fact that the optimization of any pseudo-boolean function can be reduced in polynomial time to an optimization problem on a quadratic polynomial. The original method for accomplishing this was first written by Rosenberg, and since then a reputable zoo of alternative algorithms have been introduced. Most methods share the same basic idea: reduce degree $\geq 3$ monomial terms appearing in the polynomial $g$ by introducing auxiliary variables subject to constraints.

<copy Rosenberg algorithm from Boros, Hammer pg 168>

**Theorem 4.1.** *Let $f$ be a multilinear polynomial in $n$ variables. There exists an algorithm* REDUCE *which produces a multilinear polynomial $g$ in $n + a$ variables such that*

$$\min_{(\mathbf{x},\mathbf{a}) \in \mathbb{B}^n \times \mathbb{B}^a} g(\mathbf{x}, \mathbf{a}) = \min_{\mathbf{x} \in \mathbb{B}^n} f(\mathbf{x})$$

*and if $(\mathbf{x}, \mathbf{a}) = \arg\min_{(\mathbf{x},\mathbf{a}) \in \mathbb{B}^n \times \mathbb{B}^a} g(\mathbf{x}, \mathbf{a})$ then $\mathbf{x} = \arg\min_{\mathbf{x} \in \mathbb{B}^n} f(\mathbf{x})$.*

***Boros Hammer Pseudo Boolean Optimization 2002.*** □

We need a slightly stronger statement however.

**Theorem 4.2.** *Let $f : \Sigma^N \longrightarrow \Sigma^M$ be a circuit. Then there exists an Ising circuit with auxiliary spins given by Hamiltonian $H$ which solves $f$.*

***Proof.*** Fix $G = N \cup M$ and consider the hamming objective function $\mathrm{ham} : \Sigma^N \times \Sigma^M \longrightarrow \mathbb{R}$ defined

$$\mathrm{ham}(s, t) = d(t, f(s))$$

where $d(t, f(s))$ is the Hamming distance between $t$ and the correct output $f(s)$. Then there exists some multilinear polynomial $g$ in $|G|$ variables which recovers $\mathrm{ham}$ pointwise. We now apply Rosenberg reduction to $g$ and set $H$ equal to the terminal quadratic polynomial we obtain. All that remains to show is that on any input level $s$ the output which minimizes $H$ is $f(s)$.

Fix an input $s$ and suppose that the minimizer of $g^k(s, \cdot)$ has output coordinates $f(s)$. To obtain $g^{k+1}$ we replace some pair $x_i x_j$ by $x_{k+1}$ and add the expression $M(x_i x_j - 2x_i x_{k+1} - 2x_j x_{k+1} + 3x_{k+1})$. Observe that this expression is zero if $x_i x_j = x_{k+1}$ and is strictly positive otherwise. It follows that $g^k(\mathbf{x}) = g^{k+1}(\mathbf{x}, x_{k+1})$ if $x_{k+1} = x_i x_j$ and $g^k(\mathbf{x}) < g^{k+1}(\mathbf{x}, x_{k+1})$ if $x_{k+1} \neq x_i x_j$. Hence the minimizer of $g^{k+1}$ on input level $s$ also has the correct output coordinates, and inductively, we conclude that $H$ is an Ising Hamiltonian reproducing the circuit $f$. □