

## Article

# An Order Reduction Design Framework for Higher-Order Binary Markov Random Fields

Zhuo Chen <sup>1</sup>, Hongyu Yang <sup>1,\*</sup> and Yanli Liu <sup>2</sup>

<sup>1</sup> Sichuan University National Key Laboratory of Fundamental Science on Synthetic Vision, Sichuan University, No. 29 Wangjiang Road, Chengdu 610065, China; zhuochen1989@outlook.com

<sup>2</sup> College of Computer Science, Sichuan University, No. 29 Wangjiang Road, Chengdu 610065, China

\* Correspondence: yanghongyu@scu.edu.cn

**Abstract:** The order reduction method is an important approach to optimize *higher-order binary Markov random fields* (HoMRFs), which are widely used in information theory, machine learning and image analysis. It transforms an HoMRF into an equivalent and easier *reduced first-order binary Markov random field* (RMRF) by elaborately setting the coefficients and auxiliary variables of RMRF. However, designing order reduction methods is difficult, and no previous study has investigated this design issue. In this paper, we propose an order reduction design framework to study this problem for the first time. Through study, we find that the design difficulty mainly lies in that the coefficients and variables of RMRF must be set simultaneously. Therefore, the proposed framework decomposes the design difficulty into two processes, and each process mainly considers the coefficients or auxiliary variables of RMRF. Some valuable properties are also proven. Based on our framework, a new family of 14 order reduction methods is provided. Experiments, such as synthetic data and image denoising, demonstrate the superiority of our method.

**Keywords:** Markov random field; discrete optimization; energy minimization; order reduction method; higher-order binary MRF; image desnoising



**Citation:** Chen, Z.; Yang, H.; Liu, Y. An Order Reduction Design Framework for Higher-Order Binary Markov Random Fields. *Entropy* **2023**, *25*, 535. <https://doi.org/10.3390/e25030535>

Academic Editor: Lei Wang

Received: 19 January 2023

Revised: 7 March 2023

Accepted: 16 March 2023

Published: 20 March 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The *higher-order binary Markov random field* (HoMRF) is a non-convex optimization model widely used in the fields of economy, information theory, quantum computing, machine learning and image analysis [1–14]. Recently, a new order reduction method has been developed to optimize HoMRF energies. The order reduction method transforms the HoMRF into a *reduced quadratic binary Markov random field* (RMRF) by elaborately setting the coefficients and auxiliary variables of RMRF to ensure the equivalence between HoMRF and RMRF. By equivalently transforming a complex HoMRF into an easier RMRF, the order reduction method achieves remarkable performance and thus has attracted increasing attention from researchers.

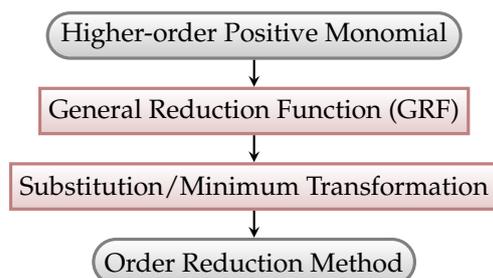
Since it is difficult to reduce an HoMRF into an RMRF in a single step, the mainstream approach is to transform a higher-order monomial into a sum of linear and quadratic monomials, and then iteratively perform this operation for all higher-order monomials of the HoMRF until a RMRF is obtained. According to the higher-order monomials to be reduced, previous order reduction methods can be divided into two categories: the methods of reducing higher-order negative monomials, and the methods of reducing higher-order positive monomials. In first category, Freedman et al. [15] developed an order reduction method for higher-order negative monomials for the first time. Then, Anthony et al. [16–18] proposed an order reduction method that is a more general form of the method [15]. Both two methods require only a binary auxiliary variable to produce submodular monomials, which can be easily minimized in polynomial time [19]. Anthony et al. [16–18] and Yip et al. [20] also developed two alternative methods to reduce higher-order negative monomials. Since

these two methods produce nonsubmodular monomials that make minimizing RMRF to be an NP-hard problem, they are rarely used in practice.

Compared with reducing higher-order negative monomials, designing the order reduction method for higher-order positive monomials is much more difficult [1,2,16–18,21,22]. Ishikawa [1,2] firstly developed a method of reducing higher-order positive monomials, by summarizing the rules in a large number of order reduction methods from the third-order to seventh-order positive monomials. Further, Anthony et al. [16–18] initiated a systematic study of lower and upper bounds on the number of binary auxiliary variables required by the order reduction methods. Nevertheless, Anthony et al. [16–18] only proved the existence of these upper and lower bounds, and failed to design any specific order reduction method that matches the corresponding upper and lower bounds. Until recently, Boros et al. [21,22] provided four order reduction methods, which utilize less auxiliary variables than Ishikawa [1,2]. Compared with the method of reducing higher-order negative monomial, the method of reducing higher-order positive monomials suffer from two critical problems: the nonsubmodular quadratic monomials obtained lead to the NP-hard problem of minimizing RMRF energies, and the introduced multiple binary auxiliary variables increase the computational complexity.

Therefore, designing more powerful order reduction methods for higher-order positive monomials is important and challenging. We observe that the main obstacle of designing order reduction methods for higher-order positive monomials is the interdependence between the coefficients and auxiliary variables of RMRF. Studies [21,22] also noticed the interdependence of auxiliary variables and coefficients. When designing order reduction methods that use logarithmic auxiliary variables, they proved that the coefficients increase exponentially. However, they did not further research the interdependence issue. Such interdependence forces us to consider the auxiliary variables and coefficients of RMRF simultaneously rather than separately, which greatly increases the design difficulty.

In this paper, we propose a unified design framework to reduce higher-order positive monomials. The proposed framework not only significantly decreases the difficulty of designing order reduction methods, but also provides 14 order reduction methods that allow applications in different fields to pick up their best method. Since the main difficulty of reducing HoMRF is due to the interdependence between the coefficients and auxiliary variables of RMRF, our core idea is to decompose this difficulty into two easier processes, where each process considers the coefficients or auxiliary variables separately. The flowchart of the proposed framework is shown in Figure 1. First, we generalize previous order reduction methods and propose a novel *general reduction function* (GRF) that requires only one auxiliary variable for any higher-order positive monomial. Different from the previous order reduction methods that utilize multiple binary auxiliary variables, our integral auxiliary variable is fixed, hence we can focus on setting the coefficients of RMRF. We rigorously prove the properties of the minimum value range of the integral auxiliary variable. Second, since the coefficients are already considered in the first process, we propose substitution and minimum transformations that convert the integral auxiliary variable of GRF into binary auxiliary variables. With the proposed transformations, we can concentrate on setting the binary auxiliary variables of RMRF. Based on the framework, a new family of 14 order reduction methods is proposed. Moreover, some state-of-the-art reduction methods can be easily derived from the proposed framework. We believe that our novel framework may take the design of order reduction methods to a new level. Comparison experiments show the superiority of our work.



**Figure 1.** Graphical flowchart of the order reduction design framework for the higher-order positive monomial.

The main contributions of our study are as follows:

1. We propose an order reduction design framework that divides the complex design of order reduction methods into two simpler processes. Compared with existing works, our framework significantly decreases the design difficulty for order reduction methods.
2. A novel GRF is developed to generalize previous order reduction methods. Unlike the previous methods with multiple binary auxiliary variables, GRF utilizes an integral auxiliary variable. Some valuable properties of GRF are also rigorously proved.
3. Two sets of substitution and minimum transformations are developed to produce more order reduction methods. A variety of 14 order reduction methods are produced to enable applications in different fields to choose their most suitable method. Moreover, four state-of-the-art order reduction methods can be easily derived from our work.

## 2. Notation and Related Work

In this section, we provide some basic terminologies, as well as the notations and backgrounds necessary for understanding the subsequent sections.

### 2.1. Higher-Order Binary Markov Random Field

In this paper, we focus on the binary Markov random field, since a Markov random field with any multi-labels can be converted to a binary Markov random field via standard techniques [2,23–26]. An HoMRF, also known as a pseudo-Boolean function [19,27], is a binary Markov random field whose order is greater than two. It is defined as  $E : \{0, 1\}^n \rightarrow \mathbb{R}$ . Here,  $n$  is the number of variables and  $\mathbb{R}$  is the set of real numbers. The energy function  $E$  can be uniquely expressed as multilinear polynomials [19]:

$$E(x_1, \dots, x_n) = \sum_{S \subseteq [n]} c_S \prod_{i \in S} x_i \quad (1)$$

where  $x_1, \dots, x_n$  are binary variables,  $[n] = \{1, 2, \dots, n\}$ ,  $c_S \in \mathbb{R}$  is the coefficient of  $c_S \prod_{i \in S} x_i$ . In particular, when  $S = \emptyset$ ,  $c_S$  denotes the constant of  $E(x_1, \dots, x_n)$ . The size of  $S$  is denoted by  $|S|$ . If  $|S| > 2$ , the monomial  $c_S \prod_{i \in S} x_i$  is a higher-order monomial. For brevity, a general higher-order negative monomial is denoted as  $-x_1 x_2 \cdots x_{|S|}$ , and a general higher-order positive monomial is denoted as  $x_1 x_2 \cdots x_{|S|}$ ,  $2 < |S| \leq n$ . Denote the complement of  $x_i$  as  $\bar{x}_i = 1 - x_i$ ,  $i \in [n]$ .

Although minimizing a general energy function is generally an NP-hard problem [19,28], the submodular energy function can be minimized globally in polynomial time. Proposition 24 in [19] is commonly used to judge whether  $E$  satisfies submodular.

As the higher orders make an HoMRF more difficult to optimize, the study [29] first proposes a substitution-based method to transform an HoMRF into a *reduced binary Markov random field* (RMRF). However, there are already experiments showing its poor performance [1,2]. Although some works can effectively reduce an HoMRF [27,30–35] or minimize an HoMRF [36–43], they can only be applied to some special HoMRFs, which

limits their applications. In this paper, we focus on reducing the general HoMRF. In reference [18], the RMRF of an HoMRF is formally defined as follows.

**Definition 1.** Given an HoMRF  $E(x_1, \dots, x_n)$ ,  $R$  is an RMRF of  $E(x_1, \dots, x_n)$  if  $R : \{0, 1\}^n \times \{0, 1\}^l \rightarrow \mathbb{R}$  is a quadratic polynomial depending on  $x_1, \dots, x_n$  and on  $l$  auxiliary variables  $y_1, \dots, y_l$  such that

$$E(x_1, \dots, x_n) = \min_{y_1, \dots, y_l \in \{0, 1\}^m} R(x_1, \dots, x_n, y_1, \dots, y_l), \forall x_1, \dots, x_n \in \{0, 1\}^n. \quad (2)$$

### 2.2. Related Works of Order Reduction Methods

As shown in Definition 1, the goal of reducing the HoMRF  $E$  is to design an equivalent RMRF  $R$  for  $E$ . However, directly reducing  $E$  into  $R$  is extremely difficult. A mainstream order reduction method is to transform the higher-order monomials of  $E$  one by one until  $R$  is obtained. Note that an important property of the higher-order monomial  $c_S x_1 x_2 \dots x_{|S|}$  is symmetry, which means that it is invariant under any permutation of the coordinates  $\{1, 2, \dots, m\}$  of the variable  $x$ . Most of reduction methods [16,18,21,22] utilize the symmetry to reduce higher-order monomials based on the following concept:

**Definition 2.** A function  $f(x, y) : \{0, 1\}^{|S|} \times \{0, 1\}^m \rightarrow \mathbb{R}$  is called the reduction function (RF) of  $-x_1 x_2 \dots x_{|S|}$  or  $x_1 x_2 \dots x_{|S|}$  if it is a symmetric quadratic polynomial about  $x_1, x_2, \dots, x_{|S|}$  and satisfies

$$-x_1 x_2 \dots x_{|S|} = \min_{y \in \{0, 1\}^m} f(x, y) \quad (3)$$

or

$$x_1 x_2 \dots x_{|S|} = \min_{y \in \{0, 1\}^m} f(x, y) \quad (4)$$

where  $x = (x_1, x_2, \dots, x_{|S|})^\top$  and  $y = (y_1, y_2, \dots, y_m)^\top$ .

We introduce the order reduction methods of higher-order negative terms and higher-order positive terms, respectively. First, for the higher-order negative monomial, Freedman et al. [15] are the first to provide:

$$-x_1 x_2 \dots x_{|S|} = \min_{y \in \{0, 1\}^m} (|S| - 1 - \sum_{i=1}^{|S|} x_i) y. \quad (5)$$

This method produces a submodular [19] RF and requires only one auxiliary variable. Thus, it achieves excellent experimental performance [1–4]. A more general form of Equation (5) is proposed by paper [17]:

$$-x_1 x_2 \dots x_{|S|} = \min_{y \in \{0, 1\}^m} (c|S| - 1 - c \sum_{i=1}^{|S|} x_i) y \quad (6)$$

where  $c \geq 1$  is the coefficient. To reduce the higher-order negative monomial, Anthony et al. [18] provide another method with one auxiliary variable, but their method do not satisfy submodular; alternatively, Yip et al. [20] propose a method that is an equivalent transformation of Equation (5).

Second, to reduce the higher-order positive monomial, order reduction method generates a nonsubmodular RF and requires multiple auxiliary variables, which is the main reason for the difficulty in minimizing the RMRF. A lot of efforts have been devoted to designing the RF for a higher-order positive monomial. Ishikawa [1,2] observes the

RF of  $x_1x_2x_3, x_1x_2x_3x_4, \dots, x_1x_2x_3x_4x_5x_6x_7$  and then summarizes the first order reduction method *high order clique reduction* (HOCR) for a general higher-order positive monomial:

$$x_1x_2 \cdots x_{|S|} = \min_{y_1, \dots, y_m \in \{0,1\}^m} \sum_{i=1}^m \left( c_{i,|S|} \left( 2i - \sum_{i=1}^{|S|} x_i \right) - 1 \right) y_i + \sum_{1 \leq i < j \leq |S|} x_i x_j \quad (7)$$

where

$$m = \lfloor \frac{|S| - 1}{2} \rfloor, c_{i,|S|} = \begin{cases} 1, & |S| \text{ is odd and } m = i, \\ 2, & \text{otherwise.} \end{cases}$$

The researches [21,22] systematically study the upper and lower bounds of the number of auxiliary variables for several classes of specially structured pseudo-Boolean functions, and provide two order reduction methods that require a logarithmic number of auxiliary variables. One is the *logarithmic reduction type-1* (LogR-1) method:

$$x_1x_2 \cdots x_{|S|} = \min_{y_0, \dots, y_{m-1} \in \{0,1\}} \left( c + \sum_{i=1}^{|S|} x_i - \sum_{i=0}^{m-1} 2^i y_i \right)^2 \quad (8)$$

where  $m = \lceil \log_2(|S|) \rceil$  and  $c = 2^m - |S|$ . Another is the *logarithmic reduction type-2* (LogR-2) method

$$x_1x_2 \cdots x_{|S|} = \min_{y_1, \dots, y_m \in \{0,1\}} \frac{1}{2} \left( c + \sum_{i=1}^{|S|} x_i - \sum_{i=1}^m 2^i y_i \right) \left( c + \sum_{i=1}^{|S|} x_i - \sum_{i=1}^m 2^i y_i - 1 \right) \quad (9)$$

where  $m = \lceil \log_2(|S|) \rceil - 1$  and  $c = 2^{m+1} - |S|$ . However, they notice that Equations (8) and (9) require exponential coefficients, which have a negative impact on computational performance. Thus, the following *linear reduction* (LinR) method is developed:

$$x_1x_2 \cdots x_{|S|} = \min_{y_1, \dots, y_m \in \{0,1\}} \frac{1}{2} \left( \sum_{i=1}^{|S|} x_i - cy_1 - 2 \sum_{i=2}^m y_i \right) \left( \sum_{i=1}^{|S|} x_i - cy_1 - 2 \sum_{i=2}^m y_i - 1 \right) \quad (10)$$

where  $\frac{|S|}{4} \leq m \leq \frac{|S|}{2}$  and  $c = |S| - 2m$ . Finally, they propose a *square root reduction* (SqrtR) method that nicely matches the lower bound in [17]:

$$x_1x_2 \cdots x_{|S|} = \min_{\substack{y_0, \dots, y_{c-1} \in \{0,1\} \\ y'_0, \dots, y'_{c-1} \in \{0,1\}}} \left( y_{i_0} y'_{i'_0} + \lambda \left( (1 - \sum_{i=0}^{c-1} y_i)^2 + (1 - \sum_{j=0}^{c-1} y'_j)^2 \right) + \lambda \left( \sum_{i=1}^{|S|} x_i - (c \sum_{i=0}^{c-1} i \cdot y_i + \sum_{j=0}^{c-1} j \cdot y'_j) \right)^2 \right) \quad (11)$$

where  $m = 2 \lceil \sqrt{|S| + 1} \rceil$  is the number of auxiliary variables,  $c = \frac{m}{2}$ , and  $\lambda$  is a large number.  $y_{i_0}$  and  $y'_{i'_0}$  must satisfy the condition  $c \cdot i_0 \cdot y_{i_0} + j_0 \cdot y'_{j_0} = |S|$ .

### 3. The First Process of the Proposed Framework: General Reduction Function (GRF)

In this section, we define a novel reduction function  $g(x, z)$  and demonstrate the convenience of designing it in Section 3.1. Then, we prove some critical properties of  $g(x, z)$  in Section 3.2.

#### 3.1. General Reduction Function (GRF)

Based on Definition 2, setting the coefficients and auxiliary variables of RMRF is greatly simplified to designing  $f(x, y)$ . However, designing  $f(x, y)$  is still very difficult, mostly owing to the interdependence between the coefficients and the auxiliary variables of  $f(x, y)$ . Such interdependence has rarely been studied, yet it must be considered when designing  $f(x, y)$ . Therefore, to avoid considering both coefficients and auxiliary variables,

we propose a novel function  $g(x, z)$  with a fixed auxiliary variable  $z$ , which allows us to focus on designing the coefficient.

**Definition 3.** A function  $g(x, z) : \{0, 1\}^{|S|} \times \mathbb{W} \rightarrow \mathbb{R}$  is called the general reduction function (GRF) of  $x_1 x_2 \cdots x_{|S|}$  if it is a symmetric polynomial about  $x_1, x_2, \dots, x_{|S|}$  and satisfies

$$x_1 x_2 \cdots x_{|S|} = \min_{z \in \mathbb{W}} g(x, z) \tag{12}$$

where  $x = (x_1, x_2, \dots, x_{|S|})^T$ ;  $\mathbb{W} = \{0, 1, \dots, |\mathbb{W}| - 1\}$  is the value space of  $z$  and  $|\mathbb{W}|$  is the size of  $\mathbb{W}$ .

Since the higher-order negative monomial has the excellent reduction method Equations (5) and (6), Definition 3 only considers the higher-order positive monomial. Furthermore,  $g(x, z)$  relaxes the restriction of  $f(x, y)$  in two ways:  $g(x, z)$  can be more than quadratic and  $z$  has not be binary. Since  $g(x, z)$  is symmetric about  $x_1, \dots, x_{|S|}$ , we denote an auxiliary function  $h : \{0, 1, \dots, |S|\} \times \mathbb{W} \rightarrow \mathbb{R}$  of  $g(x, z)$ :

$$\begin{aligned} g(x, z) &= g_1(z) + \sum_{i=1}^{|S|} g_2(z) x_i + \sum_{1 \leq i < j \leq |S|} c x_i x_j \\ &= g_1(z) + g_2(z) k + c \theta(k) = h(k, z) \end{aligned} \tag{13}$$

where  $g_1 : \mathbb{W} \rightarrow \mathbb{R}$  and  $g_2 : \mathbb{W} \rightarrow \mathbb{R}$  are polynomials;  $k = \sum_{i=1}^{|S|} x_i$  and  $\theta(k) = k(k - 1)/2$ ;  $c \in \mathbb{R}$  is the coefficient.

Since  $g(x, z)$  is a GRF of  $x_1 x_2 \cdots x_{|S|}$ , then  $x_1 x_2 \cdots x_{|S|} = \min_{z \in \mathbb{W}} h(k, z) = \mathbb{1}(k = |S|)$ , where the indicator function  $\mathbb{1}(k = |S|) = 1 \Leftrightarrow k = |S|$ . Obviously,  $h(k, z)$  is simple with only two variables. Thus, we can easily utilize it to formulate three specific forms of  $g(x, z)$ .

1. Let  $\mathbb{W}_1 = \{0, 1, \dots, |S|\}$ . Then, the general reduction type-1 (GR-1) method is

$$x_1 x_2 \cdots x_{|S|} = \min_{z \in \mathbb{W}_1} \left( \mathbb{1}(z = |S|) + \lambda \|z - x_1 - \cdots - x_{|S|}\|_2^2 \right) \tag{14}$$

where  $\|\cdot\|_2$  is the 2-norm,  $\lambda > 1$  is a large number and the indicator function  $\mathbb{1}(z = |S|) = 1 \Leftrightarrow z = |S|$ .

2. Let  $\mathbb{W}_2 = \{0, 1, \dots, |S| - 1\}$ . Then, the general reduction type-2 (GR-2) method is

$$x_1 x_2 \cdots x_{|S|} = \min_{z \in \mathbb{W}_2} \|z - x_1 - \cdots - x_{|S|}\|_2^2. \tag{15}$$

3. Let  $\mathbb{W}_3 = \{0, 1, \dots, \lfloor \frac{|S|-1}{2} \rfloor\}$ . Then, the general reduction type-3 (GR-3) method is

$$x_1 x_2 \cdots x_{|S|} = \min_{z \in \mathbb{W}_3} \frac{1}{2} \left( 2z - \sum_{i=1}^{|S|} x_i \right) \left( 2z + (-1)^{|S|} - \sum_{i=1}^{|S|} x_i \right). \tag{16}$$

### 3.2. Properties of GRF

The computational complexity of minimization algorithms, such as the sequential tree-reweighted algorithm (TRW-S) [44,45], depend heavily on the size of  $\mathbb{W}$ . Therefore, it is natural to hope that the size of  $\mathbb{W}$  is as small as possible. Before the discussion, we prove some important properties of  $g(x, z)$ .

**Theorem 1.** Let  $g(x, z)$  be the GRF of  $x_1 x_2 \cdots x_{|S|}$  and  $h(k, z)$  be the auxiliary function of  $g(x, z)$ .

1.  $g_1(z) \geq 0$  and  $g_1(z) \neq 0$ ;
2.  $c > 0$ ;
3.  $g_2(z) \leq 0$ .

**Proof.**

1. Prove that  $g_1(z) \geq 0$ . Based on the definition of  $h(k, z)$ ,  $\min_{z \in \mathbb{W}} h(k, z) = \mathbb{1}(k = |S|)$ . If  $k = 0$ , then  $\min_{z \in \mathbb{W}} h(0, z) = 0$ . According to Equation (13),  $h(0, z) = g_1(z)$ . Thus,  $\min_{z \in \mathbb{W}} h(0, z) = \min_{z \in \mathbb{W}} g_1(z) = 0 \Rightarrow g_1(z) \geq 0$ .
2. Prove that  $c > 0$ . Since  $\min_{z \in \mathbb{W}} h(k, z) = \mathbb{1}(k = |S|)$ , if  $k = 1$ , then  $\min_{z \in \mathbb{W}} h(1, z) = 0$ . According to Equation (13),  $h(1, z) = g_1(z) + g_2(z)$ . Suppose an integral  $z_0 \in \mathbb{W}$  such that

$$\min_{z \in \mathbb{W}} h(1, z) = \min_{z \in \mathbb{W}} (g_1(z) + g_2(z)) = g_2(z_0) + g_1(z_0) = 0 \Rightarrow g_2(z_0) = -g_1(z_0).$$

Similarly, if  $k = |S|$ , then  $\min_{z \in \mathbb{W}} h(|S|, z) = 1$ . When  $z = z_0$ , we have

$$h(|S|, z_0) \geq \min_{z \in \mathbb{W}} h(|S|, z) = 1.$$

According to Equation (13), the above inequality is

$$g_1(z_0) + |S|g_2(z_0) + \frac{|S|(|S| - 1)}{2}c \geq 1.$$

Since  $g_2(z_0) = -g_1(z_0)$ ,  $g_1(z_0) \geq 0$  and  $|S| > 2$ , the above inequality is transformed as

$$\begin{aligned} &g_1(z_0) - |S|g_1(z_0) + \frac{|S|(|S| - 1)}{2}c \geq 1 \\ \Rightarrow &\frac{|S|(|S| - 1)}{2}c - (|S| - 1)g_1(z_0) \geq 1 \\ \Rightarrow &c \geq \frac{2 + 2(|S| - 1)g_1(z_0)}{|S|(|S| - 1)} > 0. \end{aligned}$$

3. Prove that  $g_2(z) \leq 0$ . We prove this case by contradiction. Suppose  $z = z_0 \in \mathbb{W}$  such that  $g_2(z_0) > 0$  and  $\min_{z \in \mathbb{W}} h(k, z) = h(k, z_0) = \mathbb{1}(k = |S|)$ . We discuss it in two cases according to the value of  $k$ .

- (a) If  $1 \leq k < |S|$ , then  $h(k, z_0) = 0$ . According to Equation (13), we have

$$\begin{aligned} h(k, z_0) &= g_1(z_0) + kg_2(z_0) + \frac{k(k - 1)}{2}c = 0 \\ \Rightarrow g_2(z_0) &= -\frac{1}{k}g_1(z_0) - \frac{k - 1}{2}c. \end{aligned}$$

where  $1 \leq k < |S|$ . Since  $g_1(z_0) \geq 0$  and  $c > 0$ , then  $g_2(z_0) < 0$ , contradicting the assumption that  $g_2(z_0) > 0$ .

- (b) If  $k = |S|$ , then  $h(|S|, z_0) = 1$ . Suppose  $z_0 \neq z_1 \in \mathbb{W}$  such that

$$\min_{z \in \mathbb{W}} h(1, z) = h(1, z_1) = g_2(z_1) + g_1(z_1) = 0 \Rightarrow g_2(z_1) = -g_1(z_1).$$

Directly from the definition of  $h(k, z)$ , we can write

$$\begin{aligned} h(|S|, z_1) &\geq \min_{z \in \mathbb{W}} h(|S|, z) = h(|S|, z_0) \\ \Rightarrow g_1(z_1) + |S|g_2(z_1) + \frac{|S|(|S| - 1)}{2}c &\geq g_1(z_0) + |S|g_2(z_0) + \frac{|S|(|S| - 1)}{2}c \\ \Rightarrow g_1(z_1) + |S|g_2(z_1) &\geq g_1(z_0) + |S|g_2(z_0). \end{aligned}$$

Since  $|S| > 2$ ,  $g_2(z_1) = -g_1(z_1)$ ,  $g_1(z) \geq 0$  and  $g_2(z_0) > 0$ , we have

$$\begin{aligned} (1 - |S|)g_1(z_1) &\geq g_1(z_0) + |S|g_2(z_0) \\ \Rightarrow g_1(z_1) &\leq \frac{g_1(z_0) + |S|g_2(z_0)}{1 - |S|} < 0 \end{aligned}$$

which contradicts the fact that  $g_1(z) \geq 0$ . Therefore,  $g_2(z) \leq 0$ .

4. Prove that  $g_1(z) \neq 0$ . We prove this case by contradiction. Suppose that  $g_1(z) \equiv 0$ . Based on the definition of  $h(k, z)$  and Equation (13), there exists  $z_0 \in \mathbb{Z}$  such that

$$\min_{z \in \mathbb{Z}} h(1, z) = \min_{z \in \mathbb{Z}} (g_1(z) + g_2(z)) = \min_{z \in \mathbb{Z}} g_2(z) = g_2(z_0) = 0.$$

In other words,

$$g_2(z) \geq \min_{z \in \mathbb{Z}} g_2(z) = g_2(z_0) = 0.$$

If  $g_2(z) > 0$ , it contradicts the fact that  $g_2(z) \leq 0$ ; if  $g_2(z) = 0$ , then  $\min_{z \in \mathbb{W}} h(k, z) = \theta(k)c$  that cannot be equal to  $\mathbb{1}(k = |S|)$ , contradicting the definition of  $h(k, z)$ . Thus,  $g_1(z) \neq 0$ .

□

**Theorem 2.** Let  $g(x, z)$  be the GRF of  $x_1 x_2 \cdots x_{|S|}$  and  $h(k, z)$  be the auxiliary function of  $g(x, z)$ . There exists  $z_0 \in \mathbb{W}$  such that at most three integers  $0 \leq k_1 < k_2 < k_3 \leq |S|$  satisfy  $h(k_1, z_0) = 0$ ,  $h(k_2, z_0) = 0$ ,  $h(k_3, z_0) = \mathbb{1}(k_3 = |S|)$ .

**Proof.** According to the description in this theorem,  $\exists z_0 \in \mathbb{W}$  is such that  $h(k_1, z_0) = 0$ ,  $h(k_2, z_0) = 0$ ,  $h(k_3, z_0) = \mathbb{1}(k_3 = |S|)$ , where three integers  $0 \leq k_1 < k_2 < k_3 \leq |S|$ . Based on the Equation (13), we have

$$\begin{cases} g_1(z_0) + k_1 g_2(z_0) + \frac{k_1(k_1-1)}{2}c = 0, \\ g_1(z_0) + k_2 g_2(z_0) + \frac{k_2(k_2-1)}{2}c = 0, \\ g_1(z_0) + k_3 g_2(z_0) + \frac{k_3(k_3-1)}{2}c = \mathbb{1}(k_3 = |S|). \end{cases} \tag{17}$$

If  $k_3 \neq |S|$ , then  $c = 0$  which contradicts Theorem 1; if  $k_3 = |S|$ , there exists a solution of Equation (17) such that  $h(k_1, z_0) = 0$ ,  $h(k_2, z_0) = 0$ ,  $h(k_3, z_0) = \mathbb{1}(k_3 = |S|) = 1$ . Therefore, There exists  $z_0 \in \mathbb{W}$  such that three integers  $0 \leq k_1 < k_2 < k_3 \leq |S|$  satisfy  $h(k_1, z_0) = 0$ ,  $h(k_2, z_0) = 0$ ,  $h(k_3, z_0) = \mathbb{1}(k_3 = |S|)$ .

Now we prove that  $z_0 \in \mathbb{W}$  corresponds to at most three integers. We prove this by contradiction. Suppose that there exists  $z_0 \in \mathbb{W}$  such that four integers  $0 \leq k_1 < k_2 < k_3 < k_4 \leq |S|$  satisfy  $h(k_1, z_0) = 0$ ,  $h(k_2, z_0) = 0$ ,  $h(k_3, z_0) = 0$  and  $h(k_4, z_0) = \mathbb{1}(k_4 = |S|)$ . Then, according to the Equation (13), we have

$$\begin{cases} g_1(z_0) + k_1 g_2(z_0) + \frac{k_1(k_1-1)}{2}c = 0, \\ g_1(z_0) + k_2 g_2(z_0) + \frac{k_2(k_2-1)}{2}c = 0, \\ g_1(z_0) + k_3 g_2(z_0) + \frac{k_3(k_3-1)}{2}c = 0, \\ g_1(z_0) + k_4 g_2(z_0) + \frac{k_4(k_4-1)}{2}c = \mathbb{1}(k_4 = |S|). \end{cases} \tag{18}$$

To hold the equivalence of Equation (18),  $k_4$  must be less than  $|S|$ , i.e.,  $k_4 < |S|$ . At this point,  $c$  must be zero, which contradicts Theorem 1. For the case that there are more than four integers, the proof is similar. □

**Corollary 1.** Let  $g(\mathbf{x}, z)$  be the GRF of  $x_1 x_2 \cdots x_{|S|}$  and  $h(k, z)$  be the auxiliary function of  $g(\mathbf{x}, z)$ . If  $z_0 \in \mathbb{W}$  is such that  $h(k_1, z_0) = 0, h(k_2, z_0) = 0, h(k_3, z_0) = \mathbb{1}(k_3 = |S|)$ , where three integers  $0 \leq k_1 < k_2 < k_3 \leq |S|$ , then  $k_3 = |S|$  and

$$c = \frac{2}{(|S| - k_1)(|S| - k_2)}$$

**Proof.** Based on Theorem 2, to make Equation (17) hold and not contradict the fact that  $c > 0, k_3$  must be equal to  $|S|$  and the solution of Equation (17) is

$$g_1(z_0) = \frac{k_1 k_2}{(|S| - k_1)(|S| - k_2)}$$

$$g_2(z_0) = \frac{1 - k_1 - k_2}{(|S| - k_1)(|S| - k_2)}$$

$$c = \frac{2}{(|S| - k_1)(|S| - k_2)}.$$

□

**Theorem 3.** Let  $g(\mathbf{x}, z)$  be a GRF of  $x_1 x_2 \cdots x_{|S|}$  and  $\mathbb{W}$  be the value space of  $z$ . The minimum size of  $\mathbb{W}$  is  $\lceil \frac{|S|}{2} \rceil$ .

**Proof.** In order to obtain the minimum size of  $\mathbb{W}$ , we should let the value  $z_0 \in \mathbb{W}$  correspond to as many values of  $k$  as possible to satisfy  $\min_{z \in \mathbb{W}} h(k, z_0) = \mathbb{1}(k = |S|)$ . To do this, we choose one value  $z_0 \in \mathbb{W}$  such that  $\min_{z \in \mathbb{W}} h(k_1, z_0) = 0, \min_{z \in \mathbb{W}} h(k_2, z_0) = 0$  and  $\min_{z \in \mathbb{W}} h(|S|, z_0) = 1$ , where  $0 \leq k_1 < k_2 < |S|$ , based on Theorem 2. For  $\forall z_1 \in \mathbb{W}$  and  $z_1 \neq z_0$ , let  $z_1$  correspond to two integers  $0 \leq j_1 < j_2 < |S|, j_1 \neq k_1$  and  $j_2 \neq k_2$  such that  $\min_{z \in \mathbb{W}} h(j_1, z_1) = 0$  and  $\min_{z \in \mathbb{W}} h(j_2, z_1) = 0$  (or  $\min_{z \in \mathbb{W}} h(j_1, z_1) = 0, \min_{z \in \mathbb{W}} h(j_2, z_1) = 0$  and  $\min_{z \in \mathbb{W}} h(|S|, z_1) = 1$ ). Therefore, for  $k \in \{0, 1, \dots, |S|\}$ , we need at least  $\lceil \frac{|S|}{2} \rceil$  values in  $\mathbb{W}$  to ensure  $\min_{z \in \mathbb{W}} h(k, z) = \mathbb{1}(k = |S|)$ . The theorem is proved. □

Apparently, the size of  $\mathbb{W}_3$  in Equation (16) is  $\lceil \frac{|S|}{2} \rceil$  that matches the minimum size of  $\mathbb{W}$ . Although Equations (14) and (15) do not satisfy the minimum size of  $\mathbb{W}$ , they can not only derive some state-of-the-art order reduction methods shown in the next section, but also help to better analyze our framework shown in the experiments.

#### 4. The Second Process of the Proposed Framework: Transformation from GRF to RF

This section proposes two transformations that convert the GRF  $g(\mathbf{x}, z)$  to the RF  $f(\mathbf{x}, \mathbf{y})$  in Sections 4.1 and 4.2.

##### 4.1. Substitution Transformation

The first is the substitution transformation, which defines  $z(\mathbf{y})$  and then substitutes  $z$  in  $g(\mathbf{x}, z)$  with  $z(\mathbf{y})$ :

$$g(\mathbf{x}, z) \xrightarrow{z=z(\mathbf{y})} g(\mathbf{x}, z(\mathbf{y})) = f(\mathbf{x}, \mathbf{y}). \tag{19}$$

To ensure that  $f(\mathbf{x}, \mathbf{y})$  is quadratic,  $z(\mathbf{y})$  must be linear and cover all values of  $z$ . Thus, we provide three forms:

$$z(\mathbf{y}) = (|\mathbb{W}| - m)y_1 + \sum_{i=2}^m y_i, \quad \frac{|\mathbb{W}|}{2} \leq m \leq |\mathbb{W}| - 1 \tag{20}$$

$$z(\mathbf{y}) = (|\mathbb{W}| - 2m + 1)y_1 + \sum_{i=2}^m 2y_i, \quad \frac{|\mathbb{W}|}{4} \leq m \leq \frac{|\mathbb{W}|}{2} \tag{21}$$

$$z(\mathbf{y}) = \sum_{i=1}^m 2^{i-1}y_i, \quad m = \lceil \log_2(|\mathbb{W}|) \rceil \tag{22}$$

Then, substituting  $z$  in Equations (15) and (16) with Equation (20) or (21) or (22), we obtain six novel order reduction methods. By inserting Equation (20) into Equation (15), the *substitution reduction type-1* (SR-1) method is defined as:

$$\begin{aligned} x_1 x_2 \cdots x_{|S|} &= \min_{\mathbf{y} \in \{0,1\}^m} g(\mathbf{x}, z(\mathbf{y})) = \min_{\mathbf{y} \in \{0,1\}^m} \|z(\mathbf{y}) - x_1 - \cdots - x_{|S|}\|_2^2 \\ &= \min_{\mathbf{y} \in \{0,1\}^m} \|(|S| - m)y_1 + \sum_{i=2}^m y_i - x_1 - \cdots - x_{|S|}\|_2^2, \quad \frac{|S|}{2} \leq m \leq |S| - 1 \end{aligned} \tag{23}$$

Similarly, we substitute  $z$  in Equation (15) with Equation (21) and the *substitution reduction type-2* (SR-2) method is

$$x_1 x_2 \cdots x_{|S|} = \min_{\mathbf{y} \in \{0,1\}^m} \|(|S| - 2m + 1)y_1 + \sum_{i=2}^m 2y_i - x_1 - \cdots - x_{|S|}\|_2^2, \quad \frac{|S|}{4} \leq m \leq \frac{|S|}{2}. \tag{24}$$

Substituting  $z$  in Equation (15) with Equation (22), the *substitution reduction type-3* (SR-3) method is

$$x_1 x_2 \cdots x_{|S|} = \min_{\mathbf{y} \in \{0,1\}^m} \left\| \sum_{i=1}^m 2^{i-1}y_i - x_1 - \cdots - x_{|S|} \right\|_2^2, \quad m = \lceil \log_2(|S|) \rceil. \tag{25}$$

Substituting  $z$  in Equation (16) with Equation (20), the *substitution reduction type-4* (SR-4) method is

$$\begin{aligned} x_1 x_2 \cdots x_{|S|} &= \min_{\mathbf{y} \in \{0,1\}^m} \frac{1}{2} (2(\lceil \frac{|S|}{2} \rceil - m)y_1 + \sum_{i=2}^d 2y_i - \sum_{i=1}^m x_i) \cdot \\ &\quad (2(\lceil \frac{|S|}{2} \rceil - m)y_1 + \sum_{i=2}^d 2y_i + (-1)^{|S|} - \sum_{i=1}^{|S|} x_i) \end{aligned} \tag{26}$$

where  $\frac{\lceil |S| \rceil}{2} \leq m \leq \lceil \frac{|S|}{2} \rceil - 1$ . Substituting  $z$  in Equation (16) with Equation (21), the *substitution reduction type-5* (SR-5) method is

$$\begin{aligned} x_1 x_2 \cdots x_{|S|} &= \min_{\mathbf{y} \in \{0,1\}^m} \frac{1}{2} (2(\lceil \frac{|S|}{2} \rceil - 2m + 1)y_1 + \sum_{i=2}^m 4y_i - \sum_{i=1}^{|S|} x_i) \cdot \\ &\quad (2(\lceil \frac{|S|}{2} \rceil - 2m + 1)y_1 + \sum_{i=2}^m 4y_i + (-1)^{|S|} - \sum_{i=1}^{|S|} x_i) \end{aligned} \tag{27}$$

where  $\frac{\lceil |S| \rceil}{4} \leq m \leq \frac{\lceil |S| \rceil}{2}$ . Substituting  $z$  in Equation (16) with Equation (22), the *substitution reduction type-6* (SR-6) method is

$$x_1 x_2 \cdots x_{|S|} = \min_{\mathbf{y} \in \{0,1\}^m} \frac{1}{2} \left( \sum_{i=1}^m 2^i y_i - \sum_{i=1}^{|S|} x_i \right) \left( \sum_{i=1}^m 2^i y_i + (-1)^{|S|} - \sum_{i=1}^{|S|} x_i \right) \tag{28}$$

where  $m = \lceil \log_2(\lceil \frac{|S|}{2} \rceil) \rceil$ .

The RFs transformed from Equations (20)–(22) is simpler and clearer than some state-of-the-art reduction methods. For example, substituting  $z$  in Equation (15) with

$$z(\mathbf{y}) = \sum_{i=0}^{m-1} 2^i y_i - 2^m + |S| + \frac{1 - (-1)^{|S|}}{2}$$

where  $m = \lceil \log_1(|S|) \rceil$ , we obtain LogR-1 Equation (8) that has an extra  $-2^m + |S|$  and thus is more complex than SR-3 Equation (25). Substituting  $z$  in Equation (16) with

$$z(\mathbf{y}) = \sum_{i=1}^m 2^{i-1} y_i - 2^m + \frac{|S|}{2} + \frac{1 - (-1)^{|S|}}{4}$$

where  $m = \lceil \log_1(|S|) \rceil - 1$ , we obtain LogR-2 Equation (9) that has an extra  $-2^m + \frac{|S|}{2}$  and thus is more complex than SR-6 Equation (28). Note that LinR can be viewed as substituting  $z$  in Equation (16) with

$$z(\mathbf{y}) = \frac{1}{2}(|S| - 2m)y_1 + \sum_{i=2}^m y_i + \frac{1 - (-1)^{|S|}}{4}$$

where  $|S|/4 \leq m \leq |S|/2$ .

#### 4.2. Minimum Transformation

Unlike Equation (19), the minimum transformation utilizes the minimum operation to obtain  $f(\mathbf{x}, \mathbf{y})$ :

$$g(\mathbf{x}, z) \xrightarrow{z=z(\mathbf{y})} g(\mathbf{x}, z(\mathbf{y})) = \min_{\mathbf{y} \in \{0,1\}^d} f(\mathbf{x}, \mathbf{y}). \tag{29}$$

Compared with the proposed substitution transformation, the minimum transformation is capable of avoiding the exponential coefficients in Equation (22) and transforming Equation (14) into an RF. There are two specific forms.

First, suppose a vector of auxiliary variables:  $\mathbf{y} = (y_1, \dots, y_m) \in \{0,1\}^m$  where  $m = |\mathbb{W}| - 1$ . We define

$$z(\mathbf{y}) = y_1 + y_1 y_2 + \dots + y_1 y_2 \dots y_m. \tag{30}$$

By inserting Equation (30) into Equation (14), we have

$$\begin{aligned} g(\mathbf{x}, z(\mathbf{y})) &= \mathbb{1}(z(\mathbf{y}) = |S|) + z^2(\mathbf{y}) - 2z(\mathbf{y}) \sum_{i=1}^{|S|} x_i + \left(\sum_{i=1}^{|S|} x_i\right)^2 \\ &= \mathbb{1}\left(\sum_{i=1}^{|S|} \prod_{j=1}^i y_j = |S|\right) + \sum_{i=1}^{|S|} (2i - 1 - 2 \sum_{i=1}^{|S|} x_i) \prod_{j=1}^i y_j + \left(\sum_{i=1}^{|S|} x_i\right)^2 \\ &= \min_{\mathbf{y} \in \{0,1\}^{|S|}} y_{|S|} + \sum_{i=1}^{|S|} (2i - 1 - 2 \sum_{i=1}^{|S|} x_i) y_i + \left(\sum_{i=1}^{|S|} x_i\right)^2 \\ &= \min_{\mathbf{y} \in \{0,1\}^{|S|}} f(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Thus, the *minimum reduction type-1* (MR-1) method is

$$x_1 x_2 \dots x_{|S|} = \min_{z \in \mathbb{W}} g(\mathbf{x}, z) = \min_{\mathbf{y} \in \{0,1\}^{|S|}} y_{|S|} + \sum_{i=1}^{|S|} (2i - 1 - 2 \sum_{i=1}^{|S|} x_i) y_i + \left(\sum_{i=1}^{|S|} x_i\right)^2. \tag{31}$$

Similarly, for the Equation (15) case, we have the *minimum reduction type-2* (MR-2) method:

$$x_1 x_2 \cdots x_{|S|} = \min_{y \in \{0,1\}^{|S|-1}} \sum_{i=1}^{|S|-1} (2i - 1 - 2 \sum_{i=1}^{|S|} x_i) y_i + (\sum_{i=1}^{|S|} x_i)^2. \tag{32}$$

For the Equation (16) case, we have the *minimum reduction type-3* (MR-3) method:

$$x_1 x_2 \cdots x_{|S|} = \min_{y \in \{0,1\}^{\lfloor \frac{|S|-1}{2} \rfloor}} \sum_{i=1}^{\lfloor \frac{|S|-1}{2} \rfloor} (4i - 2 + (-1)^{|S|} - 2 \sum_{i=1}^{|S|} x_i) y_i + \sum_{1 \leq i < j \leq |S|} x_i x_j + \frac{1 - (-1)^{|S|}}{2} \sum_{i=1}^{|S|} x_i. \tag{33}$$

Second, suppose a vector of auxiliary variables:  $\mathbf{y} = (y_1, \dots, y_m, y'_1, \dots, y'_m) \in \{0, 1\}^{2m}$  where  $m = \lceil \sqrt{\lceil |S| \rceil} \rceil$ . We define

$$z(\mathbf{y}) = m \sum_{i=1}^m (i - 1) \cdot y_i + \sum_{j=1}^m (j - 1) \cdot y'_j. \tag{34}$$

To ensure the equivalence of Equation (29), we add a penalty term in  $f(\mathbf{x}, \mathbf{y})$ , which is:

$$x_1 x_2 \cdots x_{|S|} = \min_{y \in \{0,1\}^{2m}} f(\mathbf{x}, \mathbf{y}) = \min_{y \in \{0,1\}^{2m}} g(\mathbf{x}, z(\mathbf{y})) + \lambda \left( (1 - \sum_{i=1}^m y_i)^2 + (1 - \sum_{j=1}^m y'_j)^2 \right) \tag{35}$$

where  $\lambda$  is a large number. For Equation (14),  $\mathbb{1}(z(\mathbf{y}) = |S|) = y_{i_0} y'_{j_0}$  where  $y_{i_0}$  and  $y'_{j_0}$  satisfy  $m(i_0 - 1) \cdot y_{i_0} + (j_0 - 1) \cdot y'_{j_0} = |S|$ . By substituting  $z$  and  $\mathbb{1}(z = |S|)$  in Equation (14) with  $z(\mathbf{y})$  and  $\mathbb{1}(z(\mathbf{y}) = |S|)$ , the SqrtR Equation (11) is obtained. Similarly, substituting  $z$  in Equation (15) with Equation (34), we obtain the *minimum reduction type-4* (MR-4) method:

$$x_1 x_2 \cdots x_{|S|} = \min_{y \in \{0,1\}^{2m}} \left\| m \sum_{i=1}^m (i - 1) \cdot y_i + \sum_{j=1}^m (j - 1) \cdot y'_j - \sum_{k=1}^{|S|} x_k \right\|^2 + \lambda \left( (1 - \sum_{i=1}^m y_i)^2 + (1 - \sum_{j=1}^m y'_j)^2 \right) \tag{36}$$

where  $m = \lceil \sqrt{\lceil |S| \rceil} \rceil$ . Finally, substituting  $z$  in Equation (16) with with Equation (34), we obtain the *minimum reduction type-5* (MR-5) method:

$$x_1 x_2 \cdots x_{|S|} = \min_{y \in \{0,1\}^{2m}} \frac{1}{2} \left( 2m \sum_{i=1}^m (i - 1) \cdot y_i + \sum_{j=1}^m 2(j - 1) \cdot y'_j - \sum_{k=1}^{|S|} x_k \right) \cdot \left( 2m \sum_{i=1}^m (i - 1) \cdot y_i + \sum_{j=1}^m 2(j - 1) \cdot y'_j + (-1)^{|S|} - \sum_{k=1}^{|S|} x_k \right) + \lambda \left( (1 - \sum_{i=1}^m y_i)^2 + (1 - \sum_{j=1}^m y'_j)^2 \right) \tag{37}$$

where  $m = \lceil \sqrt{\lceil \frac{|S|}{2} \rceil} \rceil$ .

### 5. Experiments and Discussions

In this section, we compare the proposed method (GR-1, GR-2, GR-3, SR-1, SR-2, SR-3, SR-4, SR-5, SR-6, MR-1, MR-2, MR-3, MR-4 and MR-5) with state-of-the-art order reduction methods (HO CR, LogR-1, LogR-2, LinR and SqrtR) on synthetic data and image

denoising. Table 1 summarizes the order reduction methods developed in this paper. To ensure the fairness of the comparison, we utilize the classical algorithm TRW-S [44,45] to minimize the RMRFs reduced by these order reduction methods, and conduct the comparison experiments with the same hardware and software environment (AMD Ryzen 5800U, RAM 16GB, Nvidia RTX 3050 and MATLAB 2022a).

**Table 1.** Comparison of the proposed 14 order reduction methods for the higher-order positive monomial  $x_1x_2 \cdots x_{|S|}$ .

Methods	GRF	Transformation	Number of Auxiliary Variables	Type of Auxiliary Variables
GR-1 Equation (14)	(14)	None	$m = 1$	Integral
GR-2 Equation (15)	(15)	None	$m = 1$	Integral
GR-3 Equation (16)	(16)	None	$m = 1$	Integral
SR-1 Equation (23)	(15)	(20)	$\frac{ S }{2} \leq m \leq  S  - 1$	Binary
SR-2 Equation (24)	(15)	(21)	$\frac{ S }{4} \leq m \leq \frac{ S }{2}$	Binary
SR-3 Equation (25)	(15)	(22)	$m = \lceil \log_2( S ) \rceil$	Binary
SR-4 Equation (26)	(16)	(20)	$\frac{\lceil  S  \rceil}{2} \leq m \leq \lceil \frac{ S }{2} \rceil - 1$	Binary
SR-5 Equation (27)	(16)	(21)	$\frac{\lceil \frac{ S }{2} \rceil}{4} \leq m \leq \frac{\lceil  S  \rceil}{2}$	Binary
SR-6 Equation (28)	(16)	(22)	$m = \lceil \log_2(\lceil \frac{ S }{2} \rceil) \rceil$	Binary
MR-1 Equation (31)	(14)	(30)	$m =  S $	Binary
MR-2 Equation (32)	(15)	(30)	$m =  S  - 1$	Binary
MR-3 Equation (33)	(16)	(30)	$m = \lfloor \frac{ S -1}{2} \rfloor$	Binary
MR-4 Equation (36)	(15)	(34)	$m = \lceil \sqrt{ S } \rceil$	Binary
MR-5 Equation (37)	(16)	(34)	$m = \lceil \sqrt{\lceil \frac{ S }{2} \rceil} \rceil$	Binary

5.1. Synthetic Data Experiments

In the first experiment, we follow the approach in the paper [46] to generate a series of synthetic HoMRFs, which have the following form:

$$E(x_1, x_2, \dots, x_n) = \sum_{i=1}^n c_i x_i + \sum_{S \subseteq [n], |S| > 2} E_S(x_{S_1}, x_{S_2}, \dots, x_{S_{|S|}}) \tag{38}$$

where  $x_1, x_2, \dots, x_n \in \{0, 1\}$ ,  $c_i \in \mathbb{R}$  is the coefficient, and  $E_S : \{0, 1\}^{|S|} \rightarrow \mathbb{R}$ . We synthesize HoMRFs of order three to seven. For each order, we generate 500 HoMRFs with  $n = 50$ , 500 HoMRFs with  $n = 200$ , and 10 HoMRFs with  $n = 1000$ . The variables  $x_1, x_2, \dots, x_n$  are randomly sampled from a uniform distribution, and the coefficient  $c_i$  and the values of  $E_S$  are randomly sampled from a standard Gaussian distribution. The energy results of each order achieved by different order reduction methods are shown in Table 2, Table 3 and Table 4, respectively. Note that for SqrtR, MR-4 and MR-5,  $\lambda = 5$ .

**Table 2.** Comparison of energy ( $\times 10$ ) averaged over 500 HoMRFs with 50 variables, achieved by different order reduction methods. We highlight the best performance in bold.

Methods	3rd-Order	4th-Order	5th-Order	6th-Order	7th-Order
HOCR Equation (7)	<b>-9.98</b>	-10.03	<b>-13.04</b>	<b>-15.72</b>	<b>-19.22</b>
LogR-1 Equation (8)	-6.22	-3.72	-3.21	-6.53	-7.49
LogR-2 Equation (9)	-8.82	-10.03	-10.12	-8.98	-9.99
LinR Equation (10)	<b>-9.98</b>	-10.03	-10.10	-10.30	-11.89
SqrtR Equation (11)	-2.89	-3.46	-5.08	-6.82	-8.45
GR-1 Equation (14)	-0.59	-0.16	-0.06	-0.02	-0.01
GR-2 Equation (15)	-0.59	-0.16	-0.06	-0.02	-0.01

Table 2. Cont.

Methods	3rd-Order	4th-Order	5th-Order	6th-Order	7th-Order
GR-3 Equation (16)	−7.92	<b>−10.21</b>	−10.99	−15.41	−17.17
SR-1 Equation (23)	−6.19	−5.38	−6.16	−6.73	−7.56
SR-2 Equation (24)	−4.86	−4.85	−5.76	−7.15	−6.94
SR-3 Equation (25)	−4.19	−3.72	−7.91	−8.30	−8.68
SR-4 Equation (26)	−8.82	−10.03	−7.72	−10.3	−7.95
SR-5 Equation (27)	−8.82	−10.03	−10.02	−14.35	−11.79
SR-6 Equation (28)	−8.82	−10.03	−6.73	−10.22	−9.99
MR-1 Equation (31)	−6.01	−6.35	−8.89	−12.18	−14.56
MR-2 Equation (32)	−5.95	−6.53	−8.97	−11.75	−14.46
MR-3 Equation (33)	<b>−9.98</b>	−10.03	<b>−13.04</b>	<b>−15.72</b>	<b>−19.22</b>
MR-4 Equation (36)	−5.19	−4.22	−7.84	−9.34	−8.49
MR-5 Equation (37)	−6.68	−8.93	−9.08	−11.59	−12.01

The order reduction methods below the solid line are designed by our framework.

Table 3. Comparison of energy ( $\times 10^2$ ) averaged over 500 HoMRFs with 200 variables, achieved by different order reduction methods. We highlight the best performance in bold.

Methods	3rd-Order	4th-Order	5th-Order	6th-Order	7th-Order
HOCR Equation (7)	<b>−4.11</b>	−4.00	<b>−4.86</b>	<b>−5.74</b>	<b>−6.88</b>
LogR-1 Equation (8)	−2.61	−1.44	−1.28	−2.68	−3.28
LogR-2 Equation (9)	−3.52	−4.00	−4.18	−3.69	−4.10
LinR Equation (10)	<b>−4.11</b>	−4.00	−3.81	−3.75	−4.37
SqrtR Equation (11)	−1.12	−1.39	−2.20	−3.09	−3.96
GR-1 Equation (14)	−0.18	−0.05	−0.01	−0.01	0.00
GR-2 Equation (15)	−0.18	−0.05	−0.01	−0.01	0.00
GR-3 Equation (16)	−3.24	<b>−4.05</b>	−4.13	−5.52	−6.15
SR-1 Equation (23)	−2.57	−2.11	−2.22	−2.24	−2.36
SR-2 Equation (24)	−1.91	−1.86	−2.21	−2.67	−2.66
SR-3 Equation (25)	−1.67	−1.44	−3.26	−3.55	−3.78
SR-4 Equation (26)	−3.52	−4.00	−2.76	−3.75	−2.51
SR-5 Equation (27)	−3.52	−4.00	−4.07	−5.41	−4.82
SR-6 Equation (28)	−3.52	−4.00	−2.46	−3.81	−4.11
MR-1 Equation (31)	−2.36	−2.36	−3.24	−4.14	−5.21
MR-2 Equation (32)	−2.34	−2.41	−3.21	−4.00	−5.00
MR-3 Equation (33)	<b>−4.11</b>	−4.00	<b>−4.86</b>	<b>−5.74</b>	<b>−6.88</b>
MR-4 Equation (36)	−2.10	−1.62	−3.30	−3.86	−3.47
MR-5 Equation (37)	−2.75	−3.54	−3.69	−4.64	−5.05

The order reduction methods below the solid line are designed by our framework.

Table 4. Comparison of energy ( $\times 10^3$ ) averaged over 10 HoMRFs with 1000 variables, achieved by different order reduction methods. We highlight the best performance in bold.

Methods	3rd-Order	4th-Order	5th-Order	6th-Order	7th-Order
HOCR Equation (7)	<b>−1.97</b>	−1.92	<b>−2.21</b>	<b>−2.66</b>	<b>−3.15</b>
LogR-1 Equation (8)	−1.25	−0.66	−0.60	−1.30	−1.60
LogR-2 Equation (9)	−1.65	−1.92	−2.02	−1.72	−1.86
LinR Equation (10)	<b>−1.97</b>	−1.92	−1.80	−1.64	−1.95
SqrtR Equation (11)	−0.51	−0.64	−1.01	−1.40	−1.84
GR-1 Equation (14)	−0.07	−0.02	−0.01	0.00	0.00
GR-2 Equation (15)	−0.07	−0.02	−0.01	0.00	0.00
GR-3 Equation (16)	−1.57	<b>−1.94</b>	−2.00	−2.57	−2.91
SR-1 Equation (23)	−1.22	−0.97	−0.96	−0.99	−0.93
SR-2 Equation (24)	−0.86	−0.85	−1.06	−1.27	−1.21
SR-3 Equation (25)	−0.74	−0.66	−1.61	−1.72	−1.85
SR-4 Equation (26)	−1.65	−1.92	−1.29	−1.64	−1.01

Table 4. Cont.

Methods	3rd-Order	4th-Order	5th-Order	6th-Order	7th-Order
SR-5 Equation (27)	−1.65	−1.92	−1.99	−2.54	−2.30
SR-6 Equation (28)	−1.65	−1.92	−1.12	−1.75	−1.86
MR-1 Equation (31)	−1.09	−1.08	−1.45	−1.82	−2.20
MR-2 Equation (32)	−1.09	−1.06	−1.45	−1.79	−2.17
MR-3 Equation (33)	<b>−1.97</b>	−1.92	<b>−2.21</b>	<b>−2.66</b>	<b>−3.15</b>
MR-4 Equation (36)	−0.97	−0.73	−1.53	−1.80	−1.46
MR-5 Equation (37)	−1.29	−1.66	−1.78	−2.24	−2.44

The order reduction methods below the solid line are designed by our framework.

First, from Tables 2–4, it can be seen that for the third-order HoMRFs, the best energy performance is obtained by HO CR, LinR and MR-3; for the fourth-order HoMRFs, the best energy performance is obtained by GR-3; for the fifth-order to seventh-order HoMRFs, the best energy performance is obtained by HO CR and MR-3. As shown in Equations (7), (16) and (33), these methods with the best performance all have fewer number of auxiliary variables and smaller coefficients, which is strongly supported by the research [18,21,22,46]. Although the performance of HO CR is identical to that of MR-3, designing HO CR is more difficult than MR-3. Since HO CR is designed by summarizing the RF  $f(x, y)$  of the third-order to seventh-order positive monomials [1,2], it is heuristic. Therefore, it must be fortunate enough to heuristically find the required RFs from a large number of RFs. As for LinR, it is designed by a complex mathematical theory in the references [21,22]. In contrast, our design framework explicitly gives two processes, each of which is much more rigorous than HO CR [1,2] and simpler than LinR [21,22].

Moreover, as we discussed in Section 4.2, the state-of-the-art order reduction methods LogR-1 and LogR-2 [21,22] can be derived from our design framework. We also present two methods SR-3 and SR-6 that are more concise forms of LogR-1 and LogR-2. For the energy performance, SR-3 and SR-6 outperform LogR-1 and LogR-2, respectively, shown in Tables 2–4. It shows that the proposed framework produces superior order reduction methods than state-of-the-art ones.

Third, to validate the importance of the proposed theorems in Section 3.2, we compare the energy performance of GR-1 Equation (14), GR-2 Equation (15) and GR-3 Equation (16). As seen in Tables 2–4, while GR-3 achieves the best energy performance on the fourth-order HoMRFs and the second best energy performance on the fifth-order to seventh-order HoMRFs, GR-1 and GR-2 perform poorly. This result shows that GR-3 is more preferable than GR-1 and GR-2. Furthermore, the RFs transformed from GR-3 are also more preferable than those from GR-1 and GR-2: SR-4, SR-5 and SR-6 are superior to SR-1, SR-2 and SR-3 in energy performance, respectively; MR-3 outperforms MR-1 and MR-2, and MR-5 outperforms SqrtR and MR-4. All these results indicate the importance of GR-3 and demonstrate that the size of value space  $|\mathbb{W}|$  does significantly affect the efficiency of the reduction method.

Finally, we compare the energy performance of substitution transformations Equations (20)–(22) and minimum transformations Equations (30) and (34) in Section 3.2. As shown in Tables 2–4, SR-4, SR-5 and SR-6 have exactly the same energy performance on the third-order and fourth-order HoMRFs, since their formulas are the same in lower order HoMRFs. For the fifth-order, sixth-order and seventh-order HoMRFs, SR-4, SR-5 and SR-6 have different energy performance, and the performance of SR-5 is superior to that of SR-4 and SR-6. It demonstrates that for the GR-3, the substitution transformation Equation (21) is better than other transformations Equations (20) and (22). Similarly, it can be seen from SR-1, SR-2 and SR-3 that no substitution transformations Equations (20)–(22) has obvious advantages for GR-2. Furthermore, MR-1 and MR-3 show the better energy than SqrtR and MR-5 respectively in any case, and MR-2 shows the better energy than MR-4 in most cases. This result reveals that with the same GRF, the minimum transformation Equation (30) is more preferable than the minimum transformation Equation (34). We can conclude that

no matter what transformation  $g(x, z)$  chooses, the fewer auxiliary variables and smaller coefficients lead to superior results.

### 5.2. Image Denoising

In this section, we conduct the second experiment on the benchmark [1,2], *Fields of Expert* (FoE) [47,48], which is a widely used HoMRF model for image denoising. FoE represents the prior probability of an image distribution as:

$$E(x_1, \dots, x_n) = \sum_{i=1}^n \frac{(a_i - x_i)^2}{2\sigma^2} + \sum_{i=1}^n \alpha_i \log \left[ 1 + \frac{1}{2} (A_i * x_i) \right]$$

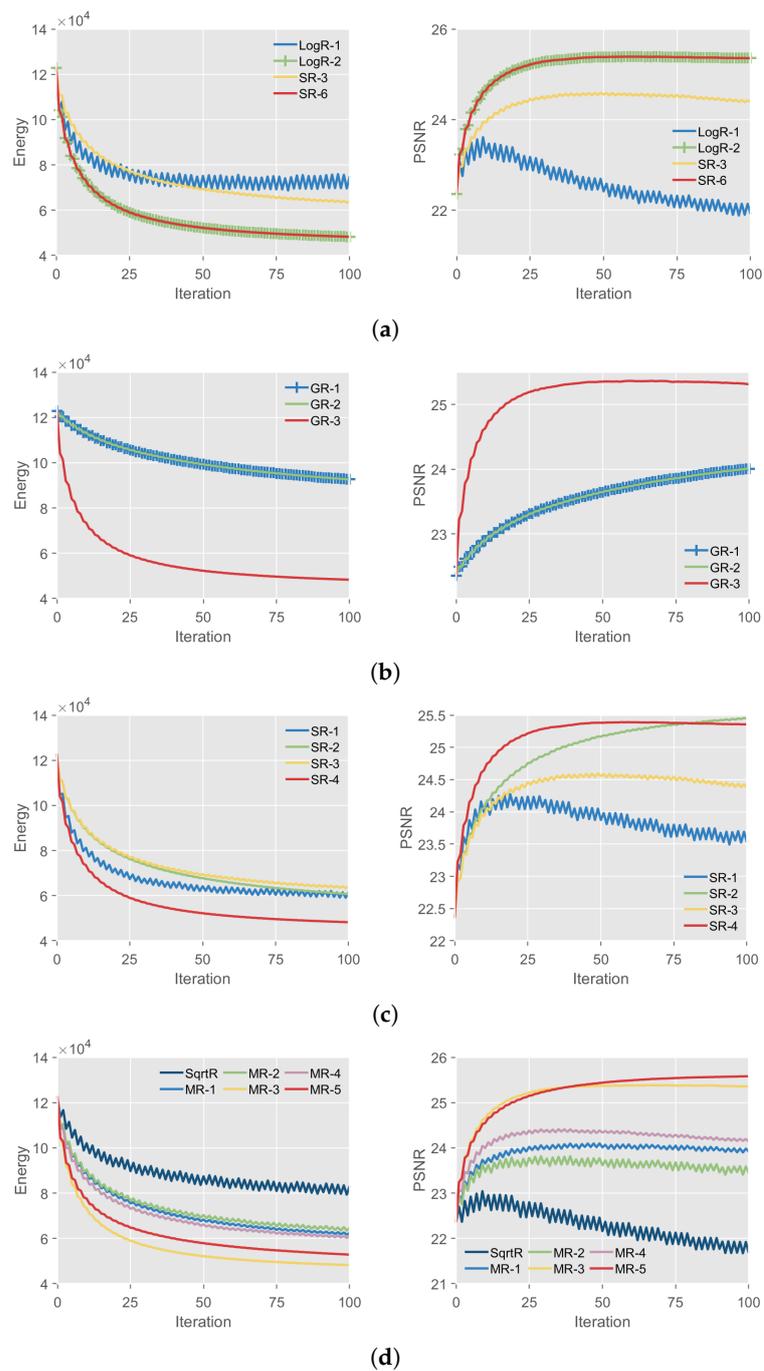
where  $n$  is the number of pixels of the final image,  $a_i$  is the value of the  $i$ -th pixel in the noise image, and  $A_i * x_i$  denotes the result of convolving the patch at pixel  $i$  with filter  $A_i$ . The parameters  $A_i$  and  $\alpha_i$  are learned from a database of natural images [47,48]. For more details, please refer to [1,2,47,48]. In the benchmark [1,2], 10 images from the widely known Berkeley Segmentation Database (BSDS500) [47–49] are selected. Gaussian noise with a standard deviation  $\sigma$  is added to each image. For efficiency, we shrink the image size by half. To make a fair comparison, we set the same starting point and maximum number of 100 iterations for order reduction methods.

We utilize the energy and the *peak signal to noise ratio* (PSNR) and the calculation time to compare the performance of different reduction methods. The results are shown in Table 5. We also plot the curves of the energy and PSNR with iterations shown in Figure 2. A visual example of image denoising is illustrated in Figure 3. Note that for the FoE model, HOCR and LinR are the same. Thus, they have the same performance in terms of energy and PSNR and differ only slightly in terms of calculation time. The same is true for SR-4, SR-5 and SR-6. For SqrtR, MR-4 and MR-5,  $\lambda = 5$ .

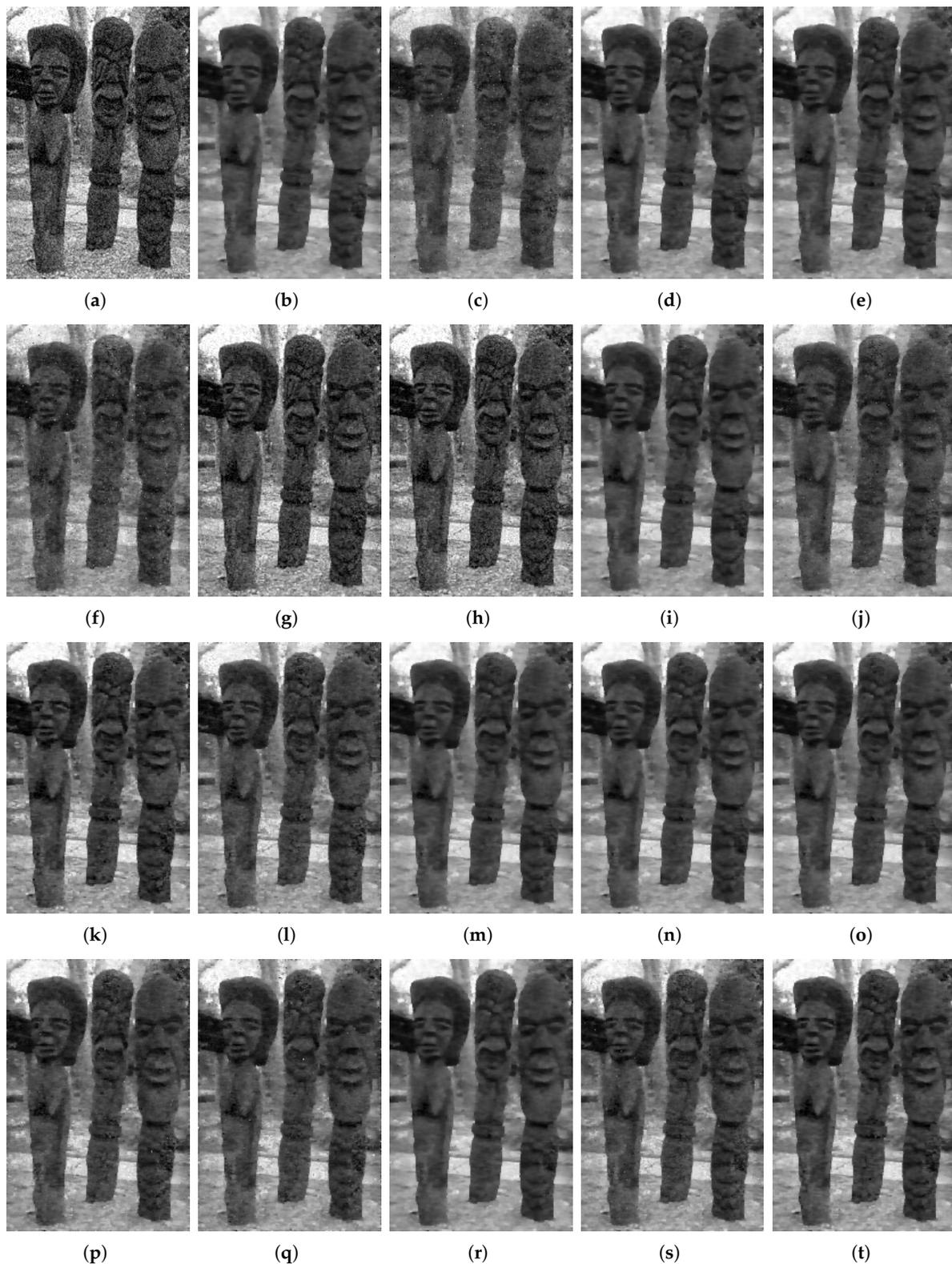
**Table 5.** Comparison of energy, PSNR and calculation time achieved by different order reduction methods, averaged over the benchmark [1,2]. We highlight the best performance in bold.

Methods	Energy ( $\times 10^4$ )			PSNR			Time ( $\times 10^2$ s)		
	$\sigma = 15$	$\sigma = 20$	$\sigma = 25$	$\sigma = 15$	$\sigma = 20$	$\sigma = 25$	$\sigma = 15$	$\sigma = 20$	$\sigma = 25$
HOCR Equation (7)	<b>5.24</b>	<b>4.74</b>	<b>4.45</b>	27.02	25.18	24.05	3.51	3.76	3.62
LogR-1 Equation (8)	7.36	7.77	7.02	25.52	21.91	22.17	5.47	5.69	5.41
LogR-2 Equation (9)	5.27	4.83	4.16	27.09	25.36	24.30	3.52	3.74	3.63
LinR Equation (10)	<b>5.24</b>	<b>4.74</b>	<b>4.45</b>	27.02	25.18	24.05	3.51	3.78	3.62
SqrtR Equation (11)	8.08	8.31	8.62	25.20	21.68	21.35	9.08	9.02	8.66
GR-1 Equation (14)	8.13	9.26	10.55	26.18	24.01	22.24	4.09	4.35	4.19
GR-2 Equation (15)	8.13	9.26	10.55	26.18	24.01	22.24	4.08	4.35	4.18
GR-3 Equation (16)	5.28	4.83	4.59	27.08	25.31	24.28	5.59	5.63	5.61
SR-1 Equation (23)	6.28	6.16	5.82	26.05	23.53	23.04	5.92	5.76	5.84
SR-2 Equation (24)	6.03	6.07	6.38	27.10	25.45	24.16	3.89	3.82	3.90
SR-3 Equation (25)	6.57	6.39	6.46	26.20	24.38	23.42	5.57	5.41	5.48
SR-4 Equation (26)	5.27	4.83	4.58	27.09	25.36	24.30	3.47	3.48	3.56
SR-5 Equation (27)	5.27	4.83	4.58	27.09	25.36	24.30	3.47	3.47	3.55
SR-6 Equation (28)	5.27	4.83	4.58	27.09	25.36	24.30	3.47	3.47	3.55
MR-1 Equation (31)	6.31	6.24	6.30	25.94	23.92	22.88	6.63	6.65	6.77
MR-2 Equation (32)	6.51	6.50	6.47	25.70	23.44	22.53	5.27	5.36	5.39
MR-3 Equation (33)	5.27	4.82	4.58	27.09	25.35	24.30	<b>3.46</b>	<b>3.45</b>	<b>3.53</b>
MR-4 Equation (36)	6.29	6.11	6.04	26.17	24.14	23.21	7.59	7.69	7.61
MR-5 Equation (37)	5.58	5.29	5.25	<b>27.17</b>	<b>25.58</b>	<b>24.48</b>	7.60	7.72	7.62

The order reduction methods below the solid line are designed by our framework.



**Figure 2.** Energy and PSNR with iterations on the benchmark [1,2], averaged over the denoising results. (a) Comparison of order reduction methods LogR-1, LogR-2, SR-3 and SR-6. (b) Comparison of order reduction methods GR-1, GR-2 and GR-3. (c) Comparison of order reduction methods SR-1, SR-2, SR-3, SR-4, SR-5 and SR-6. Note that SR-4, SR-5 and SR-6 are identical for the FoE model. (d) Comparison of order reduction methods SqrtR, MR-1, MR-2, MR-3, MR-4 and MR-5.



**Figure 3.** Visual example of image denoising. (b–f) Denoising results achieved by state-of-the-art order reduction methods. (g–t) Denoising results achieved by the proposed order reduction methods. (a) Noise Image. (b) HOOCR Equation (7). (c) LogN-1 Equation (8). (d) LogN-2 Equation (9). (e) LinR Equation (10). (f) SqrtR Equation (11). (g) GR-1 Equation (14). (h) GR-2 Equation (15). (i) GR-3 Equation (16). (j) SR-1 Equation (23). (k) SR-2 Equation (24). (l) SR-3 Equation (25). (m) SR-4 Equation (26). (n) SR-5 Equation (27). (o) SR-6 Equation (28). (p) MR-1 Equation (31). (q) MR-2 Equation (32). (r) MR-3 Equation (33). (s) MR-4 Equation (36). (t) MR-5 Equation (37).

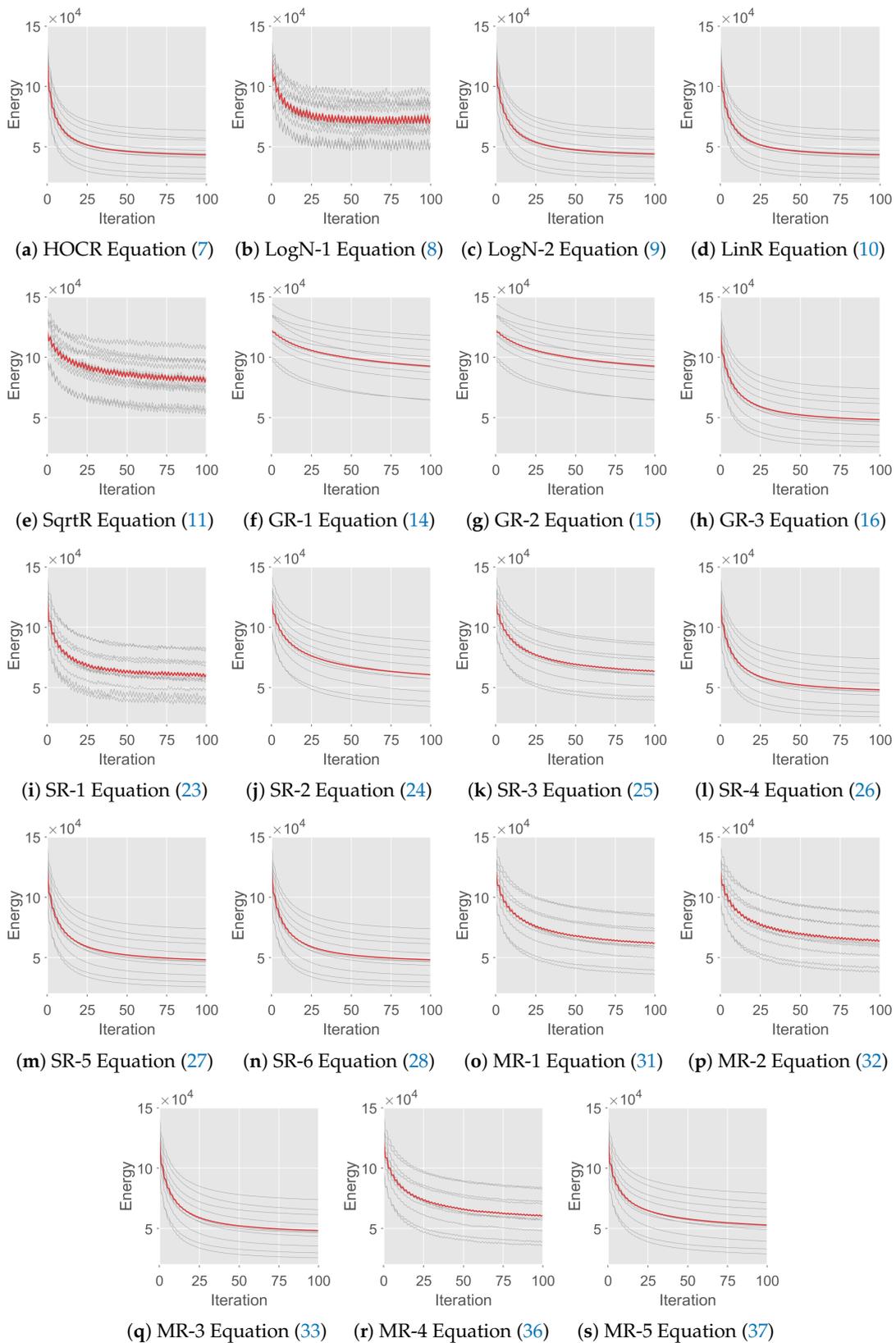
First, as shown in Table 5, the best performance of the energy, PSNR and calculation time is obtained by LinR, MR-5 and MR-3, respectively. For the denoised images with  $\sigma = 20$ , MR-3 has 1.68% higher energy than LinR and 0.89% lower PSNR than MR-5. However, the calculation time of MR-3 is 8.73% and 55.31% lower than that of LinR and MR-5, respectively. Thus, MR-3 balances energy, PSNR and calculation time and achieves the overall best performance. Similarly, for the denoised images with  $\sigma = \{15, 25\}$ , MR-3 also obtains the overall best performance. As discussed in the synthetic data experiment, MR-3 also has the advantage of lower design difficulty, which fully demonstrates the superiority of the proposed framework.

Moreover, as we discussed in Section 4.2, the proposed SR-3 and SR-6 are simpler and clearer than LogR-1 and LogR-2. It can be seen from the Figure 2a and Table 5 that SR-3 outperforms LogR-1 in energy, PSNR and calculation time, while SR-6 outperforms LogR-2 in calculation time. The results indicate that the proposed framework can improve the performance of some state-of-the-art order reduction methods.

Third, we compare different GR-1 Equation (14), GR-2 Equation (15) and GR-3 Equation (16) in Section 3.1. As shown in Figure 2b and Table 5, GR-3 has better performance than GR-1 and GR-2 in terms of energy, PSNR and calculation time. In Figure 3, GR-1 and GR-2 keep more noise, while GR3 preserves the important edge information while removing noise, which demonstrates the superiority of Equation (16). Similarly, MR-3 outperforms MR-1 and MR-2, and MR-5 outperforms SqrtR and MR-4, shown in Figures 2d and 3. This shows that the RF  $f(x, y)$  transformed from Equation (16) is superior to that from Equations (14) and (15). These results indicate that  $|\mathbb{W}|$  significantly affect the efficiency of the reduction method and that proving the theorems in Section 2.2 is important.

Forth, we compare different transformations Equations (20)–(22), (30) and (34) in Section 3.2. As shown in Figure 2c and Table 5, SR-2 performs best among SR-1 and SR-3 in energy, PSNR and calculation time. The visual results in Figure 3 also show that SR-2 outperforms SR-1 and SR-3, which demonstrates that Equation (21) is more preferable than Equations (20) and (22). Similarly, MR-1 outperforms SqrtR in energy, PSNR and calculation time; MR-2 and MR-3 outperform MR-4 and MR-5, respectively, in energy and calculation time. This result demonstrates the significant advantage of Equation (30) over Equation (34) in energy and calculation time, while for PSNR, Equation (34) is slightly better than Equation (30).

Moreover, we analyze the stability and convergence of different order reduction methods. To numerically investigate the stability of order reduction methods, the input images in the benchmark [1,2] are added with Gaussian noise of different standard deviations  $\sigma = \{15, 20, 25\}$ . The results are shown in Table 5. As it can be seen, different order reduction methods show almost the same hierarchy of denoising performances for  $\sigma = \{15, 20, 25\}$ , which demonstrates that the order reduction methods exhibit a very robust response against the standard deviation changes. In Figure 4, to show the convergence of different order reduction methods, we plot the energy changes obtained by different order reduction methods for all images in the benchmark [1,2]. The experimental results show that all order reduction methods converge for all images in the benchmark [1,2]. Among these methods, the curves of LogR-1, SqrtR and SR-1 are not smooth enough, indicating that they are more susceptible to the influence of fusion move [1,2], which is a standard technique that converts Markov random field with any multi-labels into a binary Markov random field. Moreover, most of the proposed order reduction methods (GR-1, GR-2, GR-3, SR-2, SR-3, SR-4, SR-5, SR-6, MR-1, MR-2, MR-3, MR-4 and MR-5) converge smoothly to their respective energies, which illustrates the superiority of the proposed design framework.



**Figure 4.** Energy with iterations achieved by different order reduction methods. In each subplot, every gray curve indicates the energy of one image in the benchmark [1,2], and the red curve indicates the average energy of all images in the benchmark [1,2].

Finally, we summarize two important conclusions that are helpful for designing effective reduction methods. First, the fewer the auxiliary variables  $\mathbf{y}$  in  $f(\mathbf{x}, \mathbf{y})$ , the better the experimental performance. For example, as shown in Figure 2 and Table 5, LinR, LogR-2 and MR-3 with the least  $\mathbf{y}$  in  $f(\mathbf{x}, \mathbf{y})$  have excellent performance. Second, the smaller coefficients in  $f(\mathbf{x}, \mathbf{y})$  are also beneficial to the experimental performance, especially the calculation time. For example, MR-3 has the fastest calculation time since it has the fewest coefficients  $\mathbf{y}$  in  $f(\mathbf{x}, \mathbf{y})$ .

## 6. Conclusions

In this paper, we propose a novel framework that significantly reduces the design difficulty for order reduction methods. The framework generalizes previous order reduction methods and proposes a novel GRF, which is then transformed into more RFs. The experimental results validate the superiority of the proposed design framework. There are several interesting directions for future work. First, although the proposed order reduction methods are able to reduce HoMRF with any order, minimizing the RMRF reduced from HoMRF would fail when the order of HoMRF is particularly high. Therefore, the efficiency of order reduction methods is the priority for future research. Second, there is a few of applications for order reduction methods. We plan to introduce the order reduction method into more fields.

**Author Contributions:** Conceptualization, Z.C.; Methodology, Z.C.; Software, Z.C.; Validation, H.Y. and Y.L.; Formal analysis, Z.C.; Investigation, Z.C.; Resources, H.Y. and Y.L.; Data curation, Z.C.; Writing—original draft, Z.C.; Writing—review & editing, Y.L.; Visualization, Z.C.; Supervision, H.Y.; Project administration, H.Y.; Funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Natural Science Foundation of China (grant number U20A20161).

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ishikawa, H. Higher-order clique reduction in binary graph cut. In Proceedings of the CVPR, Miami, FL, USA, 20–25 June 2009; pp. 2993–3000.
2. Ishikawa, H. Transformation of general binary MRF minimization to the first-order case. *IEEE TPAMI* **2011**, *33*, 1234–1249. [[CrossRef](#)]
3. Fix, A.; Gruber, A.; Boros, E.; Zabih, R. A graph cut algorithm for higher-order Markov Random Fields. In Proceedings of the ICCV, Barcelona, Spain, 6–13 November 2011; pp. 1020–1027.
4. Fix, A.; Gruber, A.; Boros, E.; Zabih, R. A Hypergraph-Based Reduction for Higher-Order Binary Markov Random Fields. *IEEE TPAMI* **2015**, *37*, 1387–1395. [[CrossRef](#)]
5. Shen, J.; Peng, J.; Dong, X.; Shao, L.; Porikli, F. Higher Order Energies for Image Segmentation. *IEEE TIP* **2017**, *26*, 4911–4922. [[CrossRef](#)] [[PubMed](#)]
6. Messaoud, S.; Kumar, M.; Schwing, A.G. Can We Learn Heuristics for Graphical Model Inference Using Reinforcement Learning? In Proceedings of the CVPR, Seattle, WA, USA, 14–19 June 2020; pp. 7586–7596.
7. Shen, R.; Tang, B.; Lodi, A.; Tramontani, A.; Ayed, I.B. An ILP Model for Multi-Label MRFs With Connectivity Constraints. *IEEE TIP* **2020**, *29*, 6909–6917. [[CrossRef](#)]
8. Cofré, R.; Maldonado, C. Information Entropy Production of Maximum Entropy Markov Chains from Spike Trains. *Entropy* **2018**, *20*, 34. [[CrossRef](#)]
9. Auzina, I.A.; Tomczak, J.M. Approximate Bayesian Computation for Discrete Spaces. *Entropy* **2021**, *23*, 312. [[CrossRef](#)]
10. Romanov, E.; Ordentlich, O. On Compressed Sensing of Binary Signals for the Unsourced Random Access Channel. *Entropy* **2021**, *23*, 605. [[CrossRef](#)]
11. Sioofy Khoojine, A.; Shadabfar, M.; Edrisi Tabriz, Y. A Mutual Information-Based Network Autoregressive Model for Crude Oil Price Forecasting Using Open-High-Low-Close Prices. *Mathematics* **2022**, *10*, 3172. [[CrossRef](#)]

12. Yarkoni, S.; Raponi, E.; Bäck, T.; Schmitt, S. Quantum Annealing for Industry Applications: Introduction and Review. *Rep. Prog. Phys.* **2022**, *85*, 104001. [[CrossRef](#)] [[PubMed](#)]
13. Deng, Z.; Wang, X.; Dong, B. Quantum computing for future real-time building HVAC controls. *Appl. Energy* **2023**, *334*, 120621. [[CrossRef](#)]
14. Calude, C.S.; Heidari, S.; Sifakis, J. What perceptron neural networks are (not) good for? *Inf. Sci.* **2023**, *621*, 844–857. [[CrossRef](#)]
15. Freedman, D.; Drineas, P. Energy minimization via graph cuts: Settling what is possible. In Proceedings of the CVPR, San Diego, CA, USA, 20–26 June 2005; Volume 2, pp. 939–946.
16. Gruber, A. Algorithmic and Complexity Results for Boolean and Pseudo-Boolean Functions. Ph.D. Thesis, New Brunswick Rutgers, The State University of New Jersey, New Brunswick, NJ, USA, 2015.
17. Anthony, M.; Boros, E.; Crama, Y.; Gruber, A. Quadratic reformulations of symmetric pseudo-Boolean functions. *Discret. Appl. Math.* **2016**, *203*, 1–12. [[CrossRef](#)]
18. Anthony, M.; Boros, E.; Crama, Y.; Gruber, A. Quadratic reformulations of nonlinear binary optimization problems. *Math. Program.* **2017**, *162*, 115–144. [[CrossRef](#)]
19. Boros, E.; Hammer, P.L. Pseudo-boolean optimization. *Discret. Appl. Math.* **2002**, *123*, 155–225. [[CrossRef](#)]
20. Yip, K.W.; Xu, H.; Koenig, S.; Kumar, T.K.S. Quadratic Reformulation of Nonlinear Pseudo-Boolean Functions via the Constraint Composite Graph. In *Integration of Constraint Programming, Artificial Intelligence, and Operations Research: 16th International Conference, CPAIOR 2019, Proceedings 16, Thessaloniki, Greece, 4–7 June 2019*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 643–660.
21. Boros, E.; Crama, Y.; Rodríguez-Heck, E. Quadratic reformulations of symmetric pseudo-Boolean functions: Sub-linear bounds on the number of auxiliary variables. In Proceedings of the International Symposium on Artificial Intelligence and Mathematics, Fort Lauderdale, FL, USA, 3–5 January 2018.
22. Boros, E.; Crama, Y.; Rodríguez-Heck, E. Compact quadratic reformulations for pseudo-Boolean functions. *J. Comb. Optim.* **2020**, *39*, 687–707. [[CrossRef](#)]
23. Boykov, Y.; Veksler, O.; Zabih, R. Fast approximate energy minimization via graph cuts. *IEEE TPAMI* **2001**, *23*, 1222–1239. [[CrossRef](#)]
24. Lempitsky, V.; Rother, C.; Blake, A. LogCut—Efficient Graph Cut Optimization for Markov Random Fields. In Proceedings of the ICCV, Rio De Janeiro, Brazil, 14–21 October 2007; pp. 1–8.
25. Lempitsky, V.S.; Rother, C.; Roth, S.; Blake, A. Fusion Moves for Markov Random Field Optimization. *IEEE TPAMI* **2010**, *32*, 1392–1405. [[CrossRef](#)]
26. Schlesinger, D.; FLACH, B. *Transforming an Arbitrary Minsum Problem into a Binary One*; Technical Report TUD-FI06-01; Dresden University of Technology: Dresden, Germany, 2006.
27. Verma, A.; Lewis, M. Optimal quadratic reformulations of fourth degree Pseudo-Boolean functions. *Optim. Lett.* **2020**, *14*, 1557–1569. [[CrossRef](#)]
28. Kolmogorov, V.; Zabih, R. What energy functions can be minimized via graph cuts. *IEEE TPAMI* **2004**, *26*, 147–159. [[CrossRef](#)]
29. Rosenberg, I.G. Reduction of bivalent maximization to the quadratic case. *Cah. Cent. Tudes Rech. Opér* **1975**, *17*, 71–74.
30. Bardet, M. On the Complexity of a Grobner Basis Algorithm. In Proceedings of the Algorithms Seminar, Le Chesnay-Rocquencourt, France, 25 November 2002; pp. 85–92.
31. Ishikawa, H. Higher-Order Clique Reduction without Auxiliary Variables. In Proceedings of the CVPR, Columbus, OH, USA, 23–28 June 2014; pp. 1362–1369.
32. Tanburn, R.; Lunt, O.; Dattani, N.S. Crushing runtimes in adiabatic quantum computation with Energy Landscape Manipulation (ELM): Application to Quantum Factoring. *arXiv* **2015**, arXiv:1510.07420.
33. Tanburn, R.; Okada, E.; Dattani, N. 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. *arXiv* **2015**, arXiv:1508.04816.
34. Okada, E.; Tanburn, R.; Dattani, N.S. 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. *arXiv* **2015**, arXiv:1508.07190.
35. Andres, B.; Di Gregorio, S.; Iрмаi, J.; Lange, J.H. A polyhedral study of lifted multicuts. *Discret. Optim.* **2023**, *47*, 100757. [[CrossRef](#)]
36. Arora, C.; Banerjee, S.; Kalra, P.; Maheshwari, S.N. Generic Cuts: An Efficient Algorithm for Optimal Inference in Higher Order MRF-MAP. In Proceedings of the ECCV, Florence, Italy, 7–13 October 2012; pp. 17–30.
37. Arora, C.; Banerjee, S.; Kalra, P.K.; Maheshwari, S. Generalized Flows for Optimal Inference in Higher Order MRF-MAP. *IEEE TPAMI* **2015**, *37*, 1323–1335. [[CrossRef](#)] [[PubMed](#)]
38. Shanu, I.; Arora, C.; Maheshwari, S. Inference in Higher Order MRF-MAP Problems with Small and Large Cliques. In Proceedings of the CVPR, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7883–7891.
39. Shanu, I.; Bharti, S.; Arora, C.; Maheshwari, S.N. An Inference Algorithm for Multi-label MRF-MAP Problems with Clique Size 100. In Proceedings of the ECCV, Glasgow, UK, 23–28 August 2020; pp. 257–274.
40. Kahl, F.; Strandmark, P. Generalized roof duality for pseudo-boolean optimization. In Proceedings of the ICCV, Barcelona, Spain, 6–13 November 2011; pp. 255–262.

41. Kannan, H.; Komodakis, N.; Paragios, N. Newton-Type Methods for Inference in Higher-Order Markov Random Fields. In Proceedings of the CVPR, Honolulu, HI, USA, 21–26 July 2017; pp. 7224–7233.
42. Ke, C.; Honorio, J. Exact Inference in High-order Structured Prediction. *arXiv* **2023**, arXiv:2302.03236.
43. Elloumi, S.; Lambert, A.; Lazare, A. Solving unconstrained 0-1 polynomial programs through quadratic convex reformulation. *J. Glob. Optim.* **2021**, *80*, 231–248. [[CrossRef](#)]
44. Kolmogorov, V. Convergent tree-reweighted message passing for energy minimization. *IEEE TPAMI* **2006**, *28*, 1568–1583. [[CrossRef](#)] [[PubMed](#)]
45. Kolmogorov, V. A New Look at Reweighted Message Passing. *IEEE TPAMI* **2015**, *37*, 919–930. [[CrossRef](#)]
46. Gallagher, A.C.; Batra, D.; Parikh, D. Inference for order reduction in Markov random fields. In Proceedings of the CVPR, Colorado Springs, CO, USA, 20–25 June 2011; pp. 1857–1864.
47. Roth, S.; Black, M.J. Fields of Experts: A framework for learning image priors. In Proceedings of the CVPR, San Diego, CA, USA, 20–26 June 2005; Volume 2, pp. 860–867.
48. Roth, S.; Black, M.J. Fields of Experts. *IJCV* **2009**, *82*, 205–229. [[CrossRef](#)]
49. Arbelaez, P.A.; Maire, M.; Fowlkes, C.C.; Malik, J. Contour Detection and Hierarchical Image Segmentation. *IEEE TPAMI* **2011**, *33*, 898–916. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.