ANALYSIS OF ERROR RESILIENCY OF BELIEF PROPAGATION IN COMPUTER VISION

Jungwook Choi[1]*, Ameya D. Patil*[2]*, Rob A. Rutenbar*[3]*, and Naresh R. Shanbhag*[4]

[1]IBM T. J. Watson Research Center, Yorktown Heights, NY 10598
[2,4]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign
[3]Department of Computer Science, University of Illinois at Urbana-Champaign

ABSTRACT

Probabilistic inference is a versatile tool to solve a large variety of pixel-labeling problems in computer vision such as stereo matching and image denoising. Belief Propagation (BP) is an effective method for such inference tasks, and has also shown attractive error-resilience properties—the ability to converge to usable solutions in the presence of low-level hardware errors. This is of increasing interest, as the looming end of Moore's Law scaling brings with it a vast increase in the statistical variability of nanoscale circuit fabrics. In this work we seek to understand why certain combinations of BP and error-resilience mechanisms work so well in practice. We focus on Algorithmic Noise Tolerance (ANT) techniques for the resilience mechanisms, and Max-Product BP for inference. We analyze the error characteristics of BP in this hardware context, derive novel asymptotic error bounds, and provide theoretical reasoning to explain why ANT works well in this BP context. Experimental results from detailed resilient-BP simulations for various stereo matching tasks offer empirical support for this analysis.

Index Terms— Probabilistic inference, algorithmic noise tolerance, computer vision, belief propagation

1. INTRODUCTION

We are living in the "Big Data" era where "enormous amounts of heterogeneous, semistructured and unstructured data are continually generated at unprecedented scale" [1]. To analyze and extract information from such data at scale, many tools from machine learning have been successfully employed. Among them, probabilistic graphical models (PGMs) are "an elegant framework which combines uncertainty (probabilities) and logical structure (independence constraints) to compactly represent complex, real-world phenomena" [2], and probabilistic inference derives useful insights from the PGM. More specifically, many "pixel-labeling" applications in computer vision such as stereo matching and image denoising can be mapped to maximum a posteriori (MAP) problems [3], which are successfully solved by one of the most popular probabilistic inference algorithms, Max-Product belief propagation (BP) [4].

The versatility of probabilistic inference has attracted active research on hardware implementation for the best performance per watt. However, technologies at the end of the Moore's law scaling road-map show increasing non-idealities due to statistical parameter variations. These cause erratic behavior of circuit components, hindering robust and energy efficient hardware implementation [5]. To effectively address this statistical upset, the statistical error compensation (SEC) techniques exploit the shape of the distribution of errors in the main computation and can achieve error resilience with minimal extra overhead [6]. Many efforts (e.g., [7]) have shown that SEC techniques such as Algorithmic Noise Tolerance (ANT) can not only significantly enhance the error resiliency of the overall hardware, but also achieve order of magnitude energy savings.

Probabilistic inference algorithms have inherent resiliency to small magnitude errors, and there have been many attempts to analyze this error resiliency and further enhance it. Ihler et al. [8] derived an error bound for BP, which implies inherent error resiliency, but it was only for the Sum-Product type. Varshney [9] also exploited the inherent robustness of BP and showed that, given enough iterations, BP-based LDPC decoding can be executed perfectly even in the presence of computation noise if the noise in the computation is bounded. Huang et al. [10] proposed two techniques (censoring and averaging) to improve error resiliency of BP and proved that these techniques allow BP to converge even in the presence of noise. However, their tolerable error is restricted by many constraints (e.g., censoring BP requires probability of incorrect output to be low, $\sim 5\%$). ANT has also been previously applied to significantly enhance the error resiliency of BP. Results obtained by Kim et al. [11] and Choi et al. [12] particularly demonstrate the effectiveness of ANT to handle high error rate ($20 \sim 30\%$) with little degradation in performance. But their claims were mainly based on empirical study; theoretical understanding about why ANT is able to achieve such favorable error enhancement with BP was missing.

In this paper, we provide new reasoning about the synergy between ANT and Max-Product BP, by theoretically deriving error bounds similar to what Ihler et al. [8] achieved in Sum-Product BP and providing empirical support for our analysis. The contributions of this paper are as follows: 1) We extend the error analysis framework of Ihler et al. [8] to drive asymptotic bounds on the dynamic range of message errors in Max-Product BP. 2) We use this analysis to provide an explanation for why ANT works well in the BP context. 3) We provide simulation results from various stereo matching tasks, not only to support our analysis, but also to claim that ANT enables the control of inference quality via an independent design variable separate from the choice of BP parameters and magnitude of errors.

The rest of the paper is organized as follows. Section 2 provides background information. Asymptotic bounds for Max-Product BP are derived in Section 3. Section 4 describes the supporting simulation results and Section 5 concludes the paper.

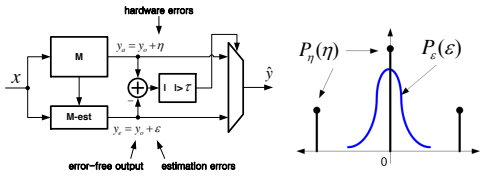
2. BACKGROUND

2.1. Algorithmic noise tolerance

Algorithmic noise tolerance (ANT) [11, 13] is a statistical error compensation technique that utilizes the statistics of errors to perform detection and estimation for error compensation. It also incorporates system level statistical metrics, such as signal-to-noise ratio (SNR) or bit error rate (BER).

As shown in Fig. 1(a), ANT incorporates a *main* block and an

This work was supported in part by Systems on Nanoscale Information fabriCs (SONIC), one of the six SRC STARnet Centers, sponsored by MARCO and DARPA



(a) (b)

Fig. 1. Block diagram: (a) algorithmic noise tolerance, and (b) error distributions.

estimator. The main block is permitted to make hardware/timing errors, but not the estimator. The estimator is a low-complexity block (typically 5%-to-20% of the main block complexity) generating a statistical estimate of the correct main block output, i.e.,

$$y_a = y_o + \eta \quad (1)$$

$$y_e = y_o + e \quad (2)$$

where y_a is the actual main block output, y_o is the error-free main block output, η is the hardware error, y_e is the estimator output, and e is the estimation error. Error distributions of the main block and the estimator are distinct, as shown in Fig. 1(b). Thus, the final output of an ANT system $\hat{y}$ can be obtained via the following decision rule:

$$\hat{y} = \begin{cases} y_a, & \text{if } |y_a - y_e| < \tau \\ y_e, & \text{otherwise} \end{cases} \quad (3)$$

where τ is an application-dependent parameter chosen to maximize the performance of ANT.

2.2. Probabilistic Inference for Computer Vision

Many pixel labeling problems in computer vision can be formulated as maximum a posteriori (MAP) problems. For example, stereo matching sweeps to find matching pixels from a pair of stereo images and statistically infer depth based on their pixel-wise horizontal displacement, which is inversely proportional to depth. Thus, stereo matching can naturally be formulated as a MAP problem [14], where we seek the most probable displacement (i.e., a set of disparity labels), $x = \{x_s, s \in \mathcal{V}\}$, given stereo images as observation, y, written as:

$$\arg\max_x P(x|y) = \arg\max_x P(y|x)P(x). \quad (4)$$

x_s is a label in a discrete domain χ associated with pixel s, and $\mathcal{V}$ corresponds to all the pixels. $P(x|y)$ is called the posterior, and it is rephrased as a product of the likelihood, $P(y|x)$, and the prior, $P(x)$, using Bayes rule.

For variety of MAP problems in computer vision applications, it is common that the likelihood and prior are factorized along an undirected graph model called a Markov Random Field (MRF), $G = (\mathcal{V}, \mathcal{E})$, where each node is a random variable associated with a pixel and each edge encodes the statistical relationship between the nodes, written as:

$$P(y|x)P(x) \propto \prod_{s\in\mathcal{V}} \phi_s(x_s, y_s) \prod_{(s,t)\in\mathcal{E}} \phi_{st}(x_s, x_t). \quad (5)$$

Factors ϕ_s and ϕ_{st} represent the likelihood of node s and the prior of edge (s,t), respectively. In order to encourage "smooth" results with label assignments, which is often preferred in pixel labeling problems, the prior term is represented as:

$$\phi_{st}(x_s, x_t) = \exp\{-\lambda \cdot \omega_{st} \cdot \min\{|x_s - x_t|^p, \theta_{max}\}\}, \quad (6)$$

Table 1. MRF parameter configurations for different Middlebury stereo matching tasks [3]

Tasks	Size	$\lvert\chi\rvert$	λ	ω_{st}	p	θ_{max}
Tsukuba	384x288	16	20	2	1	2
Venus	434x383	20	50	1	2	7
Teddy	450x375	60	10	3	1	1

where the parameters λ and ω_{st} represent coupling strength of entire edges and a specific edge st, respectively, θ_{max} is some bound in label disparity, and p is 1 or 2. (6) can represent linear/quadratic or truncated-linear/quadratic constraints, or the Potts model [3], which are the widely used priors in computer vision applications. For example, Table 1 shows various parameter configurations for different stereo matching tasks [3]. To gain intuition about error characteristics of probabilistic inference for computer vision, we analyze the impact of the coupling strength on the error resiliency of probabilistic inference.

To find the best label assignment in (4), we often minimize the objective called $Energy = -log(P(y|x)P(x))$, which is equivalent to maximizing the posterior. In general, this $energy$ minimization is intractable as the complexity of the problem grows exponentially with the size of the graph. However, one of the most popular probabilistic inference algorithm, Max-Product belief propagation (BP) [15], has been shown to effectively solve (4) via iteratively passing information about the best label (i.e., message) for all the nodes to their neighbors. In this work, we use a convergent variant of Max-Product BP called sequential tree-reweighted message passing (TRW-S) [16] for evaluation of our analysis.

3. ERROR RESILIENCY OF MAX-PRODUCT BP UNDER VARYING COUPLING STRENGTH

In this section, we analyze the error resiliency of Max-Product BP by deriving its asymptotic error bounds. We first define the message update rule for the Max-Product BP as

$$\begin{aligned} m_{st}(x_t) &\propto \max_{x_s}\{\phi(x_s, x_t) \cdot \phi_s(x_s) \cdot \prod_{u\in Nb(t)\backslash s} m_{us}(x_s)\} \\ &= \max_{x_s}\{\phi(x_s, x_t) \cdot M_{st}(x_s)\}, \end{aligned} \quad (7)$$

where $Nb(t)\backslash s$ is a set of nodes that are direct neighbors of a node t excluding s. Also, assume that all the entries of ϕ_s and ϕ_{st} have positive values and all the messages are initialized as a uniform distribution. We now consider the case when arithmetic errors occurred during the computation of (7), producing an erroneous message $\hat{m}_{st}(x_t)$. We define an error as a ratio of the error-free message and the erroneous message, $e(x) \triangleq \hat{m}(x)/m(x)$. Then, we can write an erroneous message update as

$$\begin{aligned} \hat{m}_{st}(x_t) &\propto \max_{x_s}\{\phi(x_s, x_t) \cdot \phi_s(x_s) \cdot \prod_{u\in Nb(t)\backslash s} \hat{m}_{us}(x_s)\} \\ &= \max_{x_s}\{\phi(x_s, x_t) \cdot M_{st}(x_s) \cdot E_{st}(x_s)\}. \end{aligned} \quad (8)$$

where $E_{st}(x_s) \triangleq \prod_{u\in Nb(t)\backslash s} e_{us}(x_s)$ and $e_{us}(x_s)$ is an error in message from node u to node s.

For Sum-Product BP (i.e., $\int$ instead of max in (7)), Ihler et al. have shown that the following bound holds [8]:

$$d(e_{st}) \leq \frac{d(\phi_{st})^2 d(E_{st}) + 1}{d(\phi_{st})^2 + d(E_{st})}, \quad (9)$$

where $d(e_{st})$ denotes the dynamic range of error e_{st}. The dynamic range of a vector $e(x)$ and a matrix $\phi(x,y)$ are defined as

$$d(e(x)) \triangleq \max_{a,b} \sqrt{\frac{e(a)}{e(b)}},\ d(\phi(x,y))^2 \triangleq \max_{a,b,c,d} \frac{\phi(a,b)}{\phi(c,d)}. \quad (10)$$

Therefore, $\log d(e)$ represents the largest difference among the entries of $e(x)$.

Equation (9) indicates that the dynamic range of errors on the updated messages is bounded by both the dynamic range of ϕ_{ts} and E_{ts}. This can be seen by describing limiting behavior; $d(e_{ts}) \leq d(\phi_{ts})^2$ as $d(E_{ts}) \to \infty$, and $d(e_{ts}) \leq d(E_{ts})$ as $d(\phi_{ts})^2 \to \infty$. In other words, the increasing dynamic range of error in the incoming messages has limited effect on the dynamic range of error in the output message, implying the intrinsic error resiliency of the Sum-Product BP.

The non-linearity of the $\max$ operation hinders direct extension of the error analysis approach in [8] to Max-Product BP, but previous work [12, 17] has shown that message passing inference based on Max-Product BP also exhibits inherent error resiliency. Hence, one can expect that a similar bound for the Max-Product inference would exist. In the following theorem, we establish asymptotic bounds for discrete label Max-Product BP for one particular form of the prior term, which is widely used in computer vision.

Theorem 1. *For discrete labels $x, y \in \chi$, if the prior term ϕ is defined as follows:*

$$\phi(x,y) \triangleq exp(-\lambda|x-y|^p),\ p = 1 \text{ or } 2.$$

Then, the following asymptotic bounds hold.

$$\begin{aligned} &\text{If } d(\phi)^2 \to \infty, \text{ then } d(e)^2 = d(E)^2, \\ &\text{If } d(E)^2 \to \infty, \text{ then } d(e)^2 \leq d(\phi)^4. \end{aligned} \quad (11)$$

Proof. We start from the definition of the dynamic range as follows:

$$\begin{aligned} d(e = \hat{m}/m)^2 = &\max_a \{\frac{\max_{x'}\{\phi(x',a)\cdot M(x')\cdot E(x')\}}{\max_x\{\phi(x,a)\cdot M(x)\}}\} \\ &\cdot \max_b \{\frac{\max_x\{\phi(x,b)\cdot M(x)\}}{\max_{x'}\{\phi(x',b)\cdot M(x')\cdot E(x')\}}\}. \end{aligned} \quad (12)$$

If $d(\phi)^2 \to \infty$ (i.e., $\lambda \to \infty$), ϕ dominates and the maximizing argument should be one that maximizes ϕ, which is when $x = y$ for $\phi(x,y)$. Therefore, $M(x)$ and $M(x')$ become either $M(a)$ or $M(b)$. By dividing the common factors $M(a)$ and $M(b)$, we get $d(e)^2 = \max_a\{E(a)\} \cdot \max_b\{1/E(b)\} = d(E)^2$.

Now consider the case when $\max_x E(x) \to \infty$ and thus $d(E)^2 \to \infty$. Assume $x^* = arg\max_x E(x)$. Since $\max_x E(x)$ now dominates, the maximizing argument should be one that maximizes E. Therefore, the following holds:

$$\begin{aligned} &\max_{x'}\{\phi(x',a)\cdot M(x')\cdot E(x')\} = \phi(x^*,a)\cdot M(x^*)\cdot E(x^*), \\ &\max_{x'}\{\phi(x',b)\cdot M(x')\cdot E(x')\} = \phi(x^*,b)\cdot M(x^*)\cdot E(x^*). \end{aligned} \quad (13)$$

By plugging (13) into $d(e)^2$,

$$\begin{aligned} d(e)^2 = &\max_a \{\frac{\phi(x^*,a)\cdot M(x^*)\cdot E(x^*)}{\max_x\{\phi(x,a)\cdot M(x)\}}\} \\ &\cdot \max_b \{\frac{\max_x\{\phi(x,b)\cdot M(x)\}}{\phi(x^*,b)\cdot M(x^*)\cdot E(x^*)}\}. \end{aligned} \quad (14)$$

Note that $M(x^*)$ and $E(x^*)$ are constants. By dividing the common constants and defining $\phi_{max} = 1$ and $\phi_{min} = exp(-\lambda \cdot ||\chi| - 1|^p)$, we can obtain the following upper bound,

$$\begin{aligned} &d(e)^2 \\ &\leq \max_a \{\frac{\phi_{max}}{\max_x\{\phi(x,a)\cdot M(x)\}}\} \cdot \max_b \{\frac{\max_x\{\phi(x,b)\cdot M(x)\}}{\phi_{min}}\} \\ &\leq \{\frac{\phi_{max}}{\phi_{min}\cdot \max_x\{M(x)\}}\} \cdot \{\frac{\phi_{max}\cdot \max_x\{M(x)\}}{\phi_{min}}\} = \frac{\phi_{max}^2}{\phi_{min}^2} \\ &= d(\phi)^4. \end{aligned}$$

□

In the above theorem, we can see that when the dynamic range of errors in the input messages is high, the dynamic range of errors in the output message is asymptotically bounded by $d(\phi)^4$ causing severe degradation in the inference quality for high coupling strength (λ). This qualitative observation will be made more precise in the following section where we experimentally show the impact of dynamic range of errors on the inference quality for TRW-S (one particular Max-Product type BP) and how ANT helps in enhancing its error resiliency.

4. EMPIRICAL SUPPORT FOR THEORETICAL ERROR ANALYSIS

In this section, we provide experimental results to support our theoretical analysis described in Section 3 and show how ANT can enhance error resiliency of BP regardless of choice of coupling strength. To evaluate the impact of error injection on the inference quality, we employ the same software simulator used in [12], but this time, we further change the coupling strengths λ of the prior term. We run TRW-S for the Tsukuba stereo matching task, but for simplicity in analysis, we use only a linear prior term (i.e., $p = 1$) with λ. Errors with varying magnitudes ($-2^{15} \sim 2^{15}$) are injected to the message computation with 10% injection rate. Fig. 2 shows the impact of the injected errors on the $energy$ minimization performance for the different coupling strength λ. We measure degradation of inference quality in terms of $energy$ increment (recall: for Max-Product BP, a larger $energy$ is a worse solution); to measure the $energy$ increment solely due to the error injection, we offset the erroneous $energy$ with the error free $energy$, i.e., $Energy_{erroneous}(\lambda) - Energy_{error-free}(\lambda)$.

Overall, the TRW-S inference is resilient to the small magnitude errors but vulnerable to the large magnitude ones. However, one thing to note is that, as the error magnitude becomes large, more degradation of inference quality can be observed as a larger λ is used. This trend is consistent with our analysis in (11) where as the dynamic range of input errors becomes dominant, its impact on the output error becomes more bounded by the dynamic range of the prior term, i.e., the coupling strength λ .

Next, we explore the impact of ANT on the quality of inference with the erroneous message computation. We use a simple LSB-truncated reduced precision replica (RPR) as our estimator of ANT; we experiment with different levels of estimator bit precisions (6 bit to 10 bit). Fig. 3(a) shows the degradation of inference quality vs. error magnitude for different λ. As can be seen, RPRs with larger than 8-bit precision effectively compensate for the errors and maintain low degradation of inference quality. Also, as λ increases, the larger degradation of inference quality is observed for the case without ANT. However, note that the error compensation of RPRs with 8-bit or larger precision is consistent regardless of λ. This unique

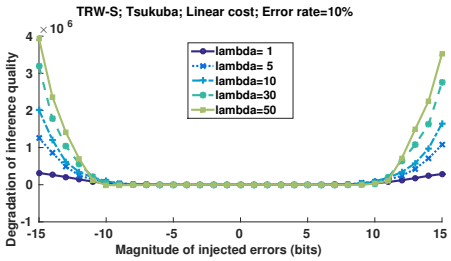
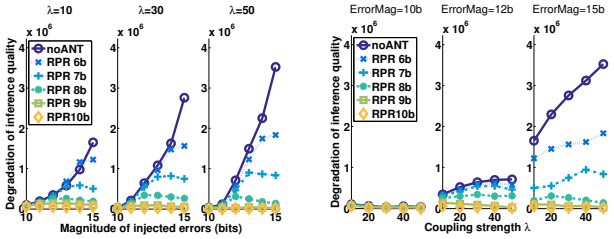


Fig. 2. Impact of error injection on inference quality for different coupling strengths.



(a) (b)

Fig. 3. (a) Inference quality vs. error injection for different coupling strength and ANT, (b) Inference quality vs. coupling strength for different error magnitude and ANT.

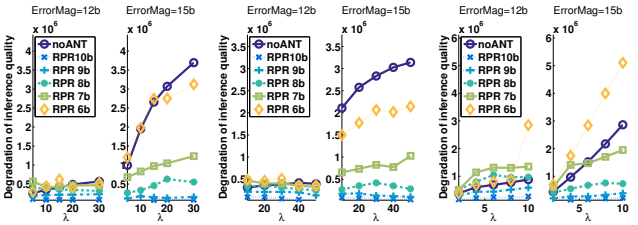
trend is more distinct if we plot the same graph as the degradation of inference quality vs. λ. As shown in Fig. 3(b), as the error magnitude increases, the degradation of inference quality increases more drastically as λ increases. However, once protected by RPRs, the degradation of inference quality becomes less affected by λ. In other words, ANT enhances the error resiliency of the message passing inference regardless of the coupling strength.

We can explain this trend using our error analysis. From (3), one can show that the difference between the output of ANT y_{ANT} and the error-free output y_o is bounded by the estimation error ϵ and the ANT threshold τ as follows:

$$|y_{ANT} - y_o| \leq \epsilon + \tau.$$

In other words, ANT can replace a large magnitude error with a small magnitude one, decreasing the dynamic range of error. It is to be noted that, in case of RPR, ϵ depends only upon estimator bit precision. Therefore, from (11), the errors on the output that were once bounded by the coupling strength due to large magnitude input errors now become bounded by the input errors with low dynamic ranges thanks to error compensation by ANT. Therefore, the error resilience of the message passing inference can be greatly enhanced by ANT regardless of the choice of coupling strength.

To verify our analysis, we perform stereo matching tasks with erroneous TRW-S under more realistic MRF parameter configurations; we run TRW-S for Tsukuba, Venus and Teddy using the different configurations specified in Table 1. In Fig. 4, which shows our experimental results, it can be observed that when ANT with estimator precision $\geq$ 8-bit is applied, inference quality is preserved for all values of coupling strength (λ) and dynamic range of injected errors. This observation supports our claim that ANT decouples the bound on dynamic range of errors in the output messages from dynamic range of errors in input messages and coupling strength, allowing the



(a) (b) (c)

Fig. 4. Inference quality vs. coupling strength for different error magnitude and ANT: (a) Tsukuba, (b) Venus, (c) Teddy.

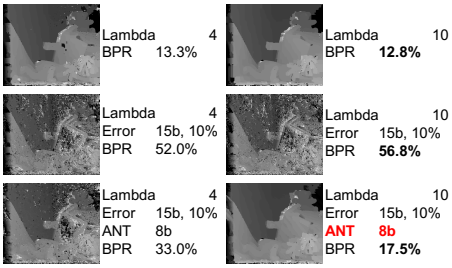


Fig. 5. Comparison of bad pixel ratio (BPR) for stereo matching of Teddy task using TRW-S with different λs.

inference quality to solely depend on estimator bit-precision. This also implies that with help of ANT, one can realize a variety of computer vision applications using various coupling strengths running on an erroneous computational fabric, since in the presence of ANT the inference quality can be controlled by estimator bit-precision for all possible choices of the coupling strength and dynamic range of errors.

We further consider the impact of errors during message passing inference on the perceptual quality of the results. Fig. 5 shows comparison of bad pixel ratio (BPR) for stereo matching of Teddy task using TRW-S with different λ (i.e., $\lambda = 4\ or\ 10$). When setting up an MRF for stereo matching, large λ is often preferred to encourage a smooth label assignment. The first row of Fig. 5 shows the disparity maps of Teddy task with different λ, and as expected, the larger λ achieves the lower BPR. However, as we discovered in the previous sections, the larger λ is, the more the quality of inference is affected by errors, resulting in higher BPR, as shown in the second row of Fig. 5. Then the last row of Fig. 5 shows that ANT with 8-bit precision RPRs can effectively compensate for the errors and reduce BPR. This demonstrates that ANT can be used to enable error resilient processing of TRW-S for a wider variety of MRF parameter configurations.

5. CONCLUSION

In this work, we analyze the error characteristics of Max-Product belief propagation (BP), and provide theoretical reasoning about why its error resiliency can be successfully enhanced by ANT. From our derivation of asymptotic error bounds of Max-Product BP, we claim that ANT can effectively complement BP to alleviate its error propagation. This claim is empirically supported by experimental results from the well known Middlebury stereo matching benchmarks.

6. REFERENCES

[1] Dunren Che, Mejdl Safran, and Zhiyong Peng, "From big data to big data mining: challenges, issues, and opportunities," in *Database Systems for Advanced Applications*. Springer, 2013, pp. 1–15.

[2] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*, The MIT Press, 2009.

[3] Richard Szeliski, Ramin Zabih, Daniel Scharstein, Olga Veksler, Vladimir Kolmogorov, Aseem Agarwala, Marshall Tappen, and Carsten Rother, "A comparative study of energy minimization methods for Markov random fields with smoothness-based priors," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 6, pp. 1068–1080, 2008.

[4] Pedro F Felzenszwalb and Daniel P Huttenlocher, "Efficient belief propagation for early vision," *International journal of computer vision*, vol. 70, no. 1, pp. 41–54, 2006.

[5] Subhasish Mitra, Kevin Brelsford, Young Moon Kim, Hsiao-Heng Kelin Lee, and Yanjing Li, "Robust system design to overcome cmos reliability challenges," *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*, vol. 1, no. 1, pp. 30–41, 2011.

[6] Naresh R Shanbhag, Rami A Abdallah, Rakesh Kumar, and Douglas L Jones, "Stochastic computation," in *Proceedings of the 47th Design Automation Conference*. ACM, 2010, pp. 859–864.

[7] Rami Abdallah, Naresh R Shanbhag, et al., "An energy-efficient ecg processor in 45-nm cmos using statistical error compensation," *Solid-State Circuits, IEEE Journal of*, vol. 48, no. 11, pp. 2882–2893, 2013.

[8] Alexander T Ihler, John W. Fisher III, and Alan S Willsky, "Loopy belief propagation: Convergence and effects of message errors," *Journal of Machine Learning Research*, pp. 905–936, 2005.

[9] Lav R Varshney, "Performance of ldpc codes under faulty iterative decoding," *Information Theory, IEEE Transactions on*, vol. 57, no. 7, pp. 4427–4444, 2011.

[10] Chu-Hsiang Huang, Yao Li, and Lara Dolecek, "Belief propagation algorithms on noisy hardware," *Communications, IEEE Transactions on*, 2015.

[11] Eric P Kim and Naresh R Shanbhag, "Energy-efficient LDPC decoders based on error-resiliency," in *IEEE Workshop on Signal Process. Syst. (SiPS)*, 2012, pp. 149–154.

[12] Jungwook Choi, Eric P Kim, Rob A. Rutenbar, and Naresh R Shanbhag, "Error resilient MRF message passing architecture for stereo matching," in *IEEE Workshop on Signal Process. Syst. (SiPS)*, 2013.

[13] Naresh R. Shanbhag, "Reliable and efficient system-on-a-chip design," vol. 37, no. 3, pp. 42–50, Mar. 2004.

[14] Jian Sun, Nan-Ning Zheng, and Heung-Yeung Shum, "Stereo matching using belief propagation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 7, pp. 787–800, 2003.

[15] M.F. Tappen and W.T. Freeman, "Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters," in *IEEE Int. Conf. on Comp. Vision*, 2003, pp. 900–906.

[16] Vladimir Kolmogorov, "Convergent tree-reweighted message passing for energy minimization," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 10, pp. 1568–1583, 2006.

[17] Eric P Kim, Jungwook Choi, Naresh R Shanbhag, and Rob Rutenbar, "A robust message passing based stereo matching kernel via system-level error resiliency," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 8331–8335.