

# Modèle d’Erreurs du Code LDPC de Rendement 3/4 dans le DVB-S2

Jean-Christophe SIBEL<sup>1</sup>, Matthieu CRUSSIÈRE<sup>2</sup>, Jean-François HÉLARD<sup>2</sup>

<sup>1</sup>CITI Lab, INSA de Lyon

<sup>2</sup>IETR, INSA de Rennes

jean-christophe.sibel@insa-lyon.fr,  
{matthieu.crussiere, jean-francois.helard}@insa-rennes.fr

**Résumé** – Dans cet article, nous proposons une analyse originale des échecs de décodage du BP pour le code LDPC de rendement 3/4 utilisé dans le standard DVB-S2. En montrant que les bits de faibles degrés sont impliqués dans plus de cycles courts que les autres bits, nous émettons l’hypothèse qu’ils sont responsables des échecs de décodage. Des tests expérimentaux dévoilent une dépendance en effet profonde entre le taux d’erreur binaire et le degré des bits du code. Ces résultats nous offrent alors un nouveau modèle d’erreurs de décodage du BP pour le code donné : à chaque RSB est associé une distribution des erreurs sur le degré des bits. Ce modèle pourrait ainsi être utilisé afin d’améliorer le design du code LDPC, permettant d’augmenter sa capacité de correction et donc de diminuer voire annuler le phénomène de plancher d’erreurs.

**Abstract** – In this paper, we propose an original analysis of failures of the BP algorithm when decoding the LDPC code of rate 3/4 from the DVB-S2 standard. Showing that bits of small degrees are mainly involved in short loops, we conjecture that they are responsible for BP failures. We then pursue a series of new experiments that exhibit a strong dependence of the bit-error rate behavior, especially the waterfall and the error floor, on the degree of bits. Thanks to these results, we propose a novel model of BP error events for the given LDPC code in such a way that to each SNR we associate a specific number of bit errors of each degree. In this way, our study would prove to be helpful to re-design codes in order to defeat the error floor and improve their correction capacity.

## 1 Introduction

La re-découverte des codes Low-Density Parity-Check (LDPC) par MacKay et Neal dans les années 90 [1] ainsi que la mise en application du décodeur Belief Propagation (BP) de faible complexité [2] ont ouvert la voie vers l’utilisation de communications sans fil robustes dont les transmissions par satellites. Nous nous intéressons dans cet article à la deuxième génération du standard Digital Video Broadcast by Satellite (DVB-S2). Les codes LDPC de ce standard étant irréguliers, la zone d’avalanche du Taux d’Erreur Binaire (TEB) peut être décalée vers de faibles Rapports Signal sur Bruit (RSB) [3], mais en contrepartie le plancher d’erreurs est très important, i.e. la pente du TEB est fortement réduite pour de forts RSB [4]. Différentes stratégies ont été proposées pour modéliser [5, 6], contrôler [7, 8] et annuler ou tout du moins diminuer ce phénomène. Dans le cadre de cette dernière stratégie, il a été montré que les petits trapping sets [9], structures topologiques cycliques, sont responsables de son émergence. Des algorithmes de décodage capable de gérer les trapping sets ont alors été mis en place tels que : les décodeurs itératifs à alphabet fini [10], très bons pour des codes réguliers de tailles modérées sur les canaux discret mais difficilement adaptables pour des codes très irréguliers avec des canaux continus ; l’algorithme du BP généralisé [11] capable de « casser » les trapping sets, indépendamment de la régularité ou irrégularité du code et de la nature du canal de transmission, et offrant une réduction significative du plancher

d’erreurs et du TEB en général mais présentant une complexité algorithmique trop grande. La stratégie adoptée dans le standard DVB-S2 consiste à pré-encoder l’information par un code de Bose-Chaudhuri-Hocquenghem (BCH) de telle façon que le décodeur BCH puisse « couper » le plancher d’erreurs, i.e. corriger toutes les erreurs pourvu qu’elles soient moins nombreuses que la capacité du code BCH. Cette méthode offre un bon compromis entre performance et complexité. Cependant, un décodeur capable de corriger le plancher d’erreurs sans l’ajout de bloc et sans un accroissement trop grand de la complexité serait plus approprié. Dans cet article, nous poursuivons l’analyse menée dans [12] afin de proposer une modélisation des événements d’erreur dans les zones d’avalanche et de plancher d’erreurs du TEB. Nous espérons par la suite proposer un design adapté des codes DVB-S2. Le code considéré dans cette étude est de rendement 3/4. L’article est organisé de la façon suivante. Le paragraphe 2 présente les notations, le paragraphe 3 montre les propriétés topologiques du code pour analyser les échecs de décodages. Le paragraphe 4 introduit le modèle.

## 2 Notations

Un code LDPC binaire  $\mathcal{C}$  de dimension  $K$  et de longueur  $N$  est un sous-espace du corps de Galois  $\mathbb{F}_2^N$  généré par  $2^K$  vecteurs appelé mots de code. Un mot de code  $\mathbf{x} \in \mathcal{C}$  satisfait à  $M$  équations de parité  $f_1, \dots, f_M$  représentant les lignes d’une

matrice de parité  $H$ , les colonnes étant les bits  $x_1, \dots, x_N$ . Nous représentons un code  $\mathcal{C}$  par son graphe de Tanner  $\mathcal{G} = (\mathbf{X} \cup \mathbf{F}, E)$  dans lequel le noeud de donnée  $X_i$  (resp. le noeud de parité  $F_a$ ) représente le bit  $x_i$  (resp. l'équation  $f_a$ ) et un bord  $e_{ia} \in E$  est tracé entre  $X_i$  et  $F_a$  si et seulement si  $x_i$  est un argument de  $f_a$ . Le degré d'un noeud est alors défini comme le nombre de bords qui lui sont adjacents :  $d_v(i)$  est le degré de  $X_i$  et  $d_c(a)$  est le degré de  $F_a$ . Nous appelons voisinage d'un noeud de donnée  $X_i$  l'ensemble  $\mathbf{F}_i$  des noeuds de parité qui lui sont connectés. De manière analogue est défini le voisinage  $\mathbf{X}_a$  du noeud de parité  $F_a$ . Le voisinage de donnée d'un noeud  $X_i$  est l'ensemble  $\mathbf{X}_i = \{X_j \in \mathbf{X} | \mathbf{F}_i \cap \mathbf{F}_j \neq \emptyset\}$ .

### 3 Codes DVB-S2

Nous étudions ici les propriétés du code, noté  $\mathcal{C}_{3/4}$ . Nous présentons les noeuds par lesquels transite la majorité des messages du BP, puis nous émettons une hypothèse sur le lien entre la topologie du graphe  $\mathcal{G}_{3/4}$  et les erreurs de décodage.

#### 3.1 Description Générale

Le code  $\mathcal{C}_{3/4}$  est de dimension  $K = 48600$  et de longueur  $N = 64800$ . L'encodage est rapide grâce à la construction de la matrice  $H_{3/4}$  en escalier [7], la redondance LDPC est ainsi représentée par des noeuds de donnée de degré 2.

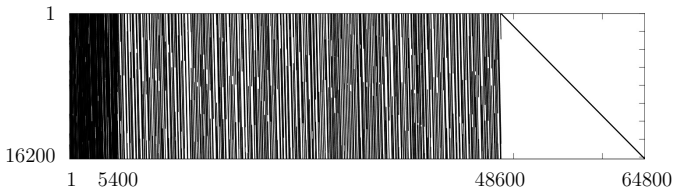


FIGURE 1 – Matrice  $H_{3/4}$ . L'axe horizontal (resp. vertical) contient les indexes des noeuds de donnée (resp. de parité).

$H_{3/4}$  est structurée en trois sous-matrices, cf. figure 1. Le tableau 1 renseigne les degrés et densités relatives de chacune de ces matrices, i.e. le nombre de leurs « 1 » divisé par  $M \times N$ .

TABLE 1 – Propriétés des sous-matrices de  $H_{3/4}$ .

	index du premier bit	degré $d_v$	densité relative $\rho$
$H_a$	0	12	$6.2 \cdot 10^{-5}$
$H_b$	5400	3	$1.2 \cdot 10^{-4}$
$H_c$	48600	2	$3.1 \cdot 10^{-5}$

Nous constatons que  $\rho_b \approx 2\rho_a$  et  $\rho_a \approx 2\rho_c$ , les messages du BP transigent principalement par les noeuds de degré 3. Ils interviennent donc probablement dans les événements qui conditionnent les performances du décodeur. Nous continuons notre étude en cherchant les liens entre les degrés des noeuds et les structures topologiques du graphe de Tanner  $\mathcal{G}_{3/4}$  de  $\mathcal{C}_{3/4}$ .

### 3.2 Topologie et Cycles

Le graphe  $\mathcal{G}_{3/4}$  présente de nombreux cycles qui rendent le BP sous-optimal [13]. Les dommages liés aux cycles courts sont particulièrement sévères, ces cycles étant souvent à la base des trapping sets qui renvoient des erreurs en quelques itérations [10]. Nous nous intéressons ici aux nombres de cycles pour chaque degré des noeuds. Nous généralisons tout d'abord la notion de voisinage en définissant  $\mathcal{G}_i^{(d)}$  le  $d$ -voisinage du noeud  $X_i$ . Tout élément de  $\mathcal{G}_i^{(d)}$  est un  $d$ -voisin de  $X_i$ , i.e. au plus  $d$  bords de  $\mathcal{G}$  sont nécessaires pour le lier à  $X_i$ . Par exemple :

$$\mathcal{G}_i^{(1)} = X_i \cup \mathbf{F}_i, \quad \mathcal{G}_i^{(2)} = X_i \cup \mathbf{X}_i \cup \mathbf{F}_i. \quad (1)$$

La première étape de notre algorithme de recherche de cycles est la construction de  $\mathcal{G}_i^{(d)}$  pour tout noeud de donnée  $X_i$ , pour une profondeur donnée  $d$ . La seconde étape consiste à calculer tous les chemins possible dans  $\mathcal{G}_i^{(d)}$  depuis  $X_i$  de taille au plus égal à un entier arbitraire  $l$ . Nous comptabilisons alors les chemins de taille  $l$  qui atteignent de nouveau  $X_i$  sans répétition d'un autre noeud. La précision de l'algorithme dépend de la profondeur  $d$  choisie. Dans notre étude, nous nous intéressons aux cycles courts uniquement, soit  $l \leq 8$ . La figure 2 indique le pourcentage moyen de cycles de taille 6, ou 6-cycles, en fonction de l'index des noeuds de donnée, avec  $d = 10$ . Le pourcentage de 8-cycles se déduit par complémentarité et il n'existe pas dans  $\mathcal{G}_{3/4}$  de 4-cycles d'après notre algorithme.

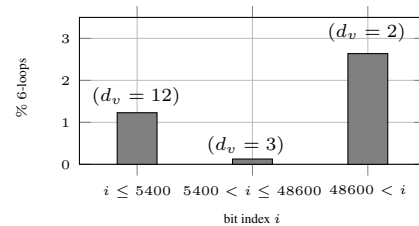


FIGURE 2 – Pourcentage de cycles de taille 6 dans  $\mathcal{C}_{3/4}$  avec une profondeur  $d = 10$ .

Les noeuds de degré 2 sont impliqués dans deux fois plus de cycles courts que les noeuds de degré 12, et vingt fois plus que les noeuds de degré 3. De manière générale, un noeud de donnée de degré 2 est particulièrement difficile à décoder lorsqu'il se retrouve piégé dans un cycle car, contrairement aux noeuds de degré supérieur, il n'a aucune connexion sortante, ses voisins faisant aussi partie du cycle. D'après les résultats précédents, nous pouvons donc supposer que les noeuds de degré 2 sont particulièrement difficiles à décoder dans  $\mathcal{C}_{3/4}$ . Nous poursuivons ainsi notre étude en regardant l'influence des degrés dans les événements d'erreurs de décodage.

### 4 Modélisation des Erreurs

Dans ce paragraphe, nous présentons une analyse expérimentale des degrés des noeuds liés aux erreurs de décodage, ce qui nous autorisera par la suite à créer un modèle d'erreurs statistique.

Toutes les expériences sont réalisées avec un canal de type additif à bruit blanc gaussien dont le RSB est  $E_s/N_0$  mesuré en dB, et une modulation 8-PSK.

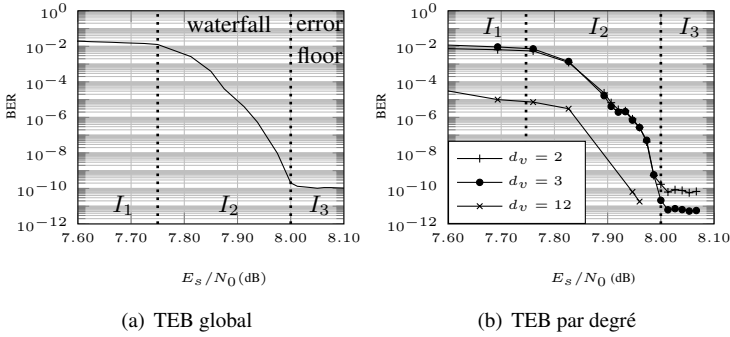


FIGURE 3 – TEB moyen du BP lors du décodage de  $C_{3/4}$ .

La figure 3(a) montre le TEB moyen du BP lors du décodage de  $C_{3/4}$ . Nous observons trois comportements différents, chacun associé à des valeurs de RSB particulières. Nous créons ainsi trois intervalles distincts :

- $I_1 = [7.60 \text{ dB} ; 7.75 \text{ dB}]$ ,
- $I_2 = [7.75 \text{ dB} ; 8.00 \text{ dB}]$  : avalanche,
- $I_3 = [8.00 \text{ dB} ; 8.10 \text{ dB}]$  : plancher d'erreurs.

Les valeurs du TEB pour des RSB en-dessous de 7.60 dB (resp. au-delà de 8.10 dB) sont similaires à celles observées dans  $I_1$  (resp. dans  $I_3$ ). Cette figure représente les deux propriétés liées à l'utilisation d'un code irrégulier : l'avalanche est très importante, le TEB descend de 8 décades en moins de 0.50 dB ; la pente du TEB après l'avalanche est cependant quasi-nulle, augmenter le RSB n'offre alors pas de meilleures performances.

La figure 3(b) montre le TEB moyen par degré. Les noeuds de degré 12 étant bien protégés grâce à leur forte connectivité aux noeuds de parité, ils sont significativement moins affectés par des erreurs de décodage que les autres noeuds de donnée. En revanche, les noeuds de degrés 2 et 3 sont responsables de la majorité des erreurs. Nous pouvons observer par la figure 4(a) la proportion d'erreurs  $\gamma_{d_v}$  pour chaque degré  $d_v \in \{2, 3, 12\}$ .

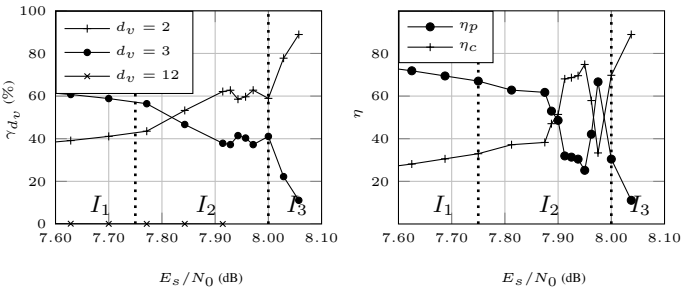


FIGURE 4 – Pourcentage moyen d'erreurs du BP lors du décodage de  $C_{3/4}$ .

Pour un RSB compris dans  $I_1$ , environ 60% (resp. 40 %) des

erreurs sont localisées sur des noeuds de degré 3 (resp. degré 2). La tendance s'inverse lorsque le RSB est dans  $I_2$  puis dans  $I_3$ , puisque plus de 85% des erreurs sont liées à des noeuds de degré 2. Ce résultat est cohérent avec l'analyse faite au paragraphe 3.2, le degré 2 est responsable du plancher d'erreurs.

Les cycles peuvent rendre faux un bit qui était correct à la réception. Nous proposons donc d'approfondir les résultats en regardant la proportion moyenne d'erreurs créées  $\eta_c$  comparée à la proportion moyenne d'erreurs persistantes  $\eta_p$ , i.e. les bits faux à la réception et restés faux en sortie du décodeur LDPC. La figure 4(b) indique que les erreurs persistantes dominent dans l'intervalle  $I_1$  et la première moitié de  $I_2$ . Nous supposons que le bruit du canal est alors trop important pour que les erreurs causées soit corrigées. Une fois que le RSB a dépassé la moitié de  $I_2$ , les erreurs sont alors majoritairement créées. Nous supposons que c'est l'influence des cycles qui est mis ici en relief. Par identification avec les précédentes observations, nous pouvons alors conjecturer que les noeuds de degré 3 sont porteurs des erreurs persistantes alors que ce sont les noeuds de degré 2 qui sont responsables de la création d'erreurs.

Il nous reste à présent à mesurer le nombre moyen de bits en erreur dans un événement d'erreurs. Considérant le mot de code nul  $\mathbf{x} = [0 \dots 0]$ , il s'agit simplement de relever le poids de Hamming  $w_H(\hat{\mathbf{x}})$  du mot de sortie  $\hat{\mathbf{x}}$ . Nous représentons ce poids moyen sur la figure 5(a).

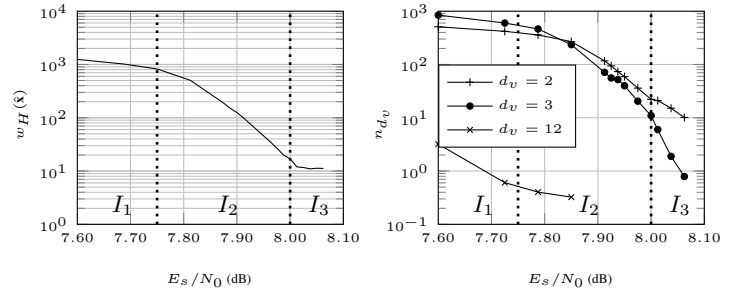


FIGURE 5 – Modèle de décodage de  $C_{3/4}$ .

Nous sommes à présent capable de dresser un modèle d'erreurs de décodage. Combinant les précédentes mesures, nous n'avons plus qu'à calculer le nombre moyen de bits en erreur  $n_{d_v}$  pour chaque degré  $d_v$  pour un RSB donné :

$$n_{d_v} = \gamma_{d_v} w_H(\hat{\mathbf{x}}) \quad (2)$$

L'utilisation de quantités moyennées peut renvoyer un nombre  $n_{d_v}$  non entier. Dans ce cas, nous seuillons arbitrairement à l'entier inférieur le plus proche. La figure 5(b) montre les quantités  $n_2$ ,  $n_3$  et  $n_{12}$  en fonction du RSB. Une propriété remarquable de ce modèle est la correspondance avec l'utilisation du BCH dans le standard DVB-S2. Le code BCH utilisé a une capacité de correction de 12 bits, i.e. il est capable de corriger au plus 12 erreurs sur le mot de sortie du décodeur LDPC. Nous rappelons que, la redondance LDPC étant tenue par les bits as-

sociés aux noeuds de degré 2, ces 12 erreurs ne concernent que les noeuds de degré 3 et 12. Comme les erreurs dans  $I_3$  ne sont portées que par les noeuds de degré 2 et 3 et que celles liées aux noeuds de degré 3 représentent moins de 10% du nombre total de bits en erreurs, le code BCH est pourrait avoir une capacité de correction réduite avec un résultat identique.

Dans le cas des autres intervalles  $I_1$  et  $I_2$ , la capacité de correction du BCH est dépassée. Par exemple, à 8.0 dB nous relevons  $n_3 = 11$  tandis qu'à 7.97 dB nous mesurons  $n_3 = 20$ . L'encodage BCH est donc inefficace, en moyenne, dès lors que le RSB descend en dessous de 8.0 dB. Le modèle d'erreurs que nous proposons ici pourrait s'avérer utile pour concevoir un code BCH, ou un ensemble de codes BCH, parfaitement adapté à l'évolution des erreurs de décodage LDPC. Il serait également le moyen de construire le code  $C_{3/4}$  de manière conjointe avec le résultat du décodeur itératif. Rappelons que les bits de redondance se calculent très vite, malgré la grande longueur des codes, grâce la forme en escalier de la matrice de parité, ce qui est un avantage considérable dans les applications pratiques. De plus, ils permettent d'équilibrer la complexité algorithmique de décodage due à la forte connectivité des premiers bits du code. Malgré ces bonnes propriétés, nous avons vu dans ce document que ces bits sont aussi responsables des échecs de décodage. Ainsi, un design différent des noeuds associés à la redondance LDPC à l'aide du modèle présenté serait une procédure prometteuse.

## 5 Conclusion

Dans cet article, nous avons montré une analyse des échecs du BP lors du décodage d'un code LDPC du standard DVB-S2. Nous avons vu par une étude de la topologie de ces codes que les noeuds de faibles degrés sont impliqués dans de nombreux cycles courts, influençant directement les performances de décodage du BP. En approfondissant l'analyse, nous avons présenté des résultats expérimentaux montrant la dépendance entre le RSB et le nombre d'erreurs par degré ainsi que par origine : les erreurs dans la zone d'avalanche semblent être liées au bruit de canal uniquement, le plancher d'erreurs proviendrait d'erreurs créées par les cycles courts. Nous avons terminé cet article par l'introduction d'un modèle statistique d'erreurs pour lequel à chaque RSB correspond un nombre moyen d'erreurs pour chaque degré du code. Un tel modèle indique les faiblesses du code et du décodeur, il pourrait s'avérer alors utile pour améliorer l'un et/ou l'autre procédé, donnant lieu à des transmissions DVB-S2 plus fiables et plus robustes. En outre, la méthode de modélisation des erreurs est suffisamment flexible pour être adapté à tout autre code LDPC dont on souhaiterait analyser voire améliorer les capacités de correction conjointement avec le décodeur BP.

Ce travail a été mené au sein du projet FUI OptisaT2.

## Références

- [1] D. J. C. MacKay, "Good Error-Correcting Codes Based on Very Sparse Matrices," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 399–431, 1995.
- [2] F.R. Kschischang, B. J. Frey and H.A. Loeliger, "Factor Graphs and the Sum-Product Algorithm," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [3] T. J. Richardson, M. A. Shokrollahi and R. L. Urbanke, "Design of Capacity-Approaching Irregular Low-Density Parity-Check Codes," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 619–637, 2001.
- [4] T.J. Richardson, "Error floors of LDPC codes," in *Proc. 41st Annual Allerton Conference on Communication, Control and Computing*, 2003, pp. 1426–1435.
- [5] L. Dolecek, P. Lee, Z. Zhang, V. Anantharam, B. Nikolic, M.J. Wainwright, "Predicting error floors of structured LDPC codes : deterministic bounds and estimates," *IEEE Journal of Selected Areas in Communications*, vol. 27, no. 6, pp. 908–917, 2009.
- [6] B.K. Butler and P.H. Siegel, "Error Floor Approximation for LDPC Codes in the AWGN Channel," in *Proc. 49th Annual Allerton Conference on Communication, Control and Computing*, 2011, pp. 204–211.
- [7] S.J. Johnson and S.R. Weller, "Constraining LDPC Degree Distributions for Improved Error Floor Performance," *IEEE Communications Letters*, vol. 10, no. 2, pp. 103–105, 2006.
- [8] D.V. Nguyen, S.H. Chilappagari, M.W. Marcellin and B. Vasic, "On the Construction of Structured LDPC Codes Free of Small Trapping Sets," *IEEE Transactions on Information Theory*, vol. 58, no. 4, pp. 2280–2302, 2012.
- [9] B. Vasic, S.K. Chilappagari, D.V. Nguyen and S.K. Planjery, "Trapping set ontology," in *Proc. Allerton Conference on Communication, Control and Computing*, 2009, pp. 1–7.
- [10] L. Danjean, D. Declercq, S. K. Planjery, B. Vasic, "On the selection of finite alphabet iterative decoders for LDPC codes on the BSC," in *Proc. IEEE Information Theory Workshop*, 2011, pp. 345–349.
- [11] J.-C. Sibel, S. Reynal and D. Declercq, "An application of Generalized Belief Propagation : splitting trapping sets in LDPC codes," in *Proc. IEEE Int. Symp. Inf. Theory*, 2014, pp. 706–710.
- [12] J.-C. Sibel, M. Crussière and J.-F. Héland, "Analysis of Decoding Failures of DVB-S2 LDPC Codes," in *Proc. IEEE Vehicular Technology Conference (VTC Fall)*, 2014, pp. 1–6.
- [13] K.P. Murphy, Y. Weiss and M.I. Jordan, "Loopy belief propagation for approximate inference : an empirical study," in *Proc. Uncertainty in Artificial Intelligence*, 1999, pp. 467–475.